Ф.Р.Гантмахер

ТЕОРИЯ МАТРИЦ

Содержание

Предисловие автора к первому изданию 7
Предисловие редактора ко второму издания 10

ЧАСТЬ I

ОСНОВЫ ТЕОРИИ

Глава I. Матрицы и действия над ними 13
§ 1. Матрицы. Основные обозначения 13
§ 2. Сложение и умножение прямоугольных матриц 15
§ 3. Квадратные матрицы 24
§ 4. Ассоциированные матрицы. Миноры обратной матрицы 30
§ 5. Обращение прямоугольных матриц. Псевдообратная матрица 32
Глава II. Алгоритм Гаусса и некоторые его применения 41
§ 1. Метод исключения Гаусса 41
§ 2. Механическая интерпретация алгоритма Гаусса 45
§ 3. Детерминантное тождество Сильвестра 47
§ 4. Разложение квадратной матрицы на треугольные множители 49
§ 5. Разбиение матрицы на блоки. Техника оперирования с блочными матрицами. Обобщенный алгоритм Гаусса 55
Глава III. Линейные операторы в *n*-мерном векторном пространстве 65
§ 1. Векторное пространство 65
§ 2. Линейный оператор, отображающий *n*-мерное пространство в *m*-мерное 70
§ 3. Сложение и умножение линейных операторов 71
§ 4. Преобразование координат 73
§ 5. Эквивалентные матрицы. Ранг оператора. Неравенства Сильвестра 74
§ 6. Линейные операторы, отображающие *n*-мерное пространство само в себя 79
§ 7. Характеристические числа и собственные векторы линейного оператора 82
§ 8. Линейные операторы простой структуры 84
Глава IV. Характеристический и минимальный многочлены матрицы 87
§ 1. Сложение и умножение матричных многочленов 87
§ 2. Правое и левое деление матричных многочленов. Обобщенная теорема Безу 89
§ 3. Характеристический многочлен матрицы. Присоединенная матрица 92
§ 4. Метод Д. К. Фаддеева одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы 96
§ 5. Минимальный многочлен матрицы 98
Глава V. Функции от матрицы 103
§ 1. Определение функции от матрицы 103
§ 2. Интерполяционный многочлен Лагранжа — Сильвестра 108

§ 3. Другие формы определения $f(A)$. Компоненты матрицы A 111
§ 4. Представление функций от матриц рядами 115
§ 5. Некоторые свойства функций от матриц 119
§ 6. Применение функций от матрицы к интегрированию системы линейных дифференциальных уравнений с постоянными коэффициентами 124
§ 7. Устойчивость движения в случае линейной системы 130
Глава VI. Эквивалентные преобразования многочленных матриц. Аналитическая теория элементарных делителей 135
§ 1. Элементарные преобразования многочленной матрицы 135
§ 2. Канонический вид λ-матрицы 139
§ 3. Инвариантные многочлены и элементарные делители многочленной матрицы 143
§ 4. Эквивалентность линейных двучленов 148
§ 5. Критерий подобия матриц 149
§ 6. Нормальные формы матрицы 151
§ 7. Элементарные делители матрицы $f(A)$ 155
§ 8. Общий метод построения преобразующей матрицы 159
§ 9. Второй метод построения преобразующей матрицы 162
Глава VII. Структура линейного оператора в n-мерном пространстве (геометрическая теория элементарных делителей) 171
§ 1. Минимальный многочлен вектора пространства (относительно заданного линейного оператора) 171
§ 2. Расщепление на инвариантные подпространства с взаимно простыми минимальными многочленами 173
§ 3. Сравнение. Надпространство 175
§ 4. Расщепление пространства на циклические инвариантные подпространства 177
§ 5. Нормальная форма матрицы 182
§ 6. Инвариантные многочлены. Элементарные делители 184
§ 7. Нормальная жорданова форма матрицы 188
§ 8. Метод акад. А. Н. Крылова преобразования векового уравнения 190
Глава VIII. Матричные уравнения 199
§ 1. Уравнение $AX = XB$ 199
§ 2. Частный случай: $A = B$. Перестановочные матрицы 203
§ 3. Уравнение $AX — XB = C$ 207
§ 4. Скалярное уравнение $f(X)=0$ 207
§ 5. Матричное многочленное уравнение 209
§ 6. Извлечение корня m-й степени из неособенной матрицы 212
§ 7. Извлечение корня m-й степени из особенной матрицы 215
§ 8. Логарифм матрицы 219
Глава IX. Линейные операторы в унитарном пространстве 222
§ 1. Общие соображения 222

§ 2. Метризация пространства 222
§ 3. Критерий Грама линейной зависимости векторов 225
§ 4. Ортогональное проектирование 227
§ 5. Геометрический смысл определителя Грама и некоторые неравенства 229
§ 6. Ортогонализация ряда векторов 233
§ 7. Ортонормированный базис 237
§ 8. Сопряженный оператор 239
§ 9. Нормальные операторы в унитарном пространстве 243
§ 10. Спектр нормальных, эрмитовых, унитарных операторов 245
§ 11. Неотрицательные и положительно определенные эрмитовы операторы 248
§ 12. Полярное разложение линейного оператора в унитарном пространстве. Формулы Кэли 249
§ 13. Линейные операторы в евклидовом пространстве 254
§ 14. Полярное разложение оператора и формулы Кэли в евклидовом пространстве 260
§ 15. Коммутирующие нормальные операторы 263
§ 16. Псевдообратный оператор 265
Глава X. Квадратичные и эрмитовы формы 267
§ 1. Преобразование переменных в квадратичной форме 267
§ 2. Приведение квадратичной формы к сумме квадратов. Закон инерции 269
§ 3. Метод Лагранжа приведения квадратичной формы к сумме квадратов. Формула Якоби 271
§ 4. Положительные квадратичные формы 276
§ 5. Приведение квадратичной формы к главным осям 279
§ 6. Пучок квадратичных форм 281
§ 7. Экстремальные свойства характеристических чисел регулярного пучка форм 286
§ 8. Малые колебания системы с n степенями свободы 293
§ 9. Эрмитовы формы 297
§ 10. Ганкелевы формы 301

ЧАСТЬ II

СПЕЦИАЛЬНЫЕ ВОПРОСЫ И ПРИЛОЖЕНИЯ

Глава XI. Комплексные симметрические, кососимметрические и ортогональные матрицы 313
§ 1. Некоторые формулы для комплексных ортогональных и унитарных матриц 313
§ 2. Полярное разложение комплексной матрицы 317
§ 3. Нормальная форма комплексной симметрической матрицы 319
§ 4. Нормальная форма комплексной кососимметрической матрицы 322
§ 5. Нормальная форма комплексной ортогональной матрицы 327
Глава XII. Сингулярные пучки матриц 331
§ 1. Введение 331

§ 2. Регулярный пучок матриц 332
§ 3. Сингулярные пучки. Теорема о приведении 335
§ 4. Каноническая форма сингулярного пучка матриц 340
§ 5. Минимальные индексы пучка. Критерий строгой эквивалентности пучков 342
§ 6. Сингулярные пучки квадратичных форм 345
§ 7. Приложения к дифференциальным уравнениям 348
Глава XIII. Матрицы с неотрицательными элементами 352
§ 1. Общие свойства 352
§ 2. Спектральные свойства неразложимых неотрицательных матриц 354
§ 3. Разложимые матрицы 365
§ 4. Нормальная форма разложимой матрицы 372
§ 5. Примитивные и импримитивные матрицы 377
§ 6. Стохастические матрицы 381
§ 7. Предельные вероятности для однородной цепи Маркова с конечным числом состояний 385
§ 8. Вполне неотрицательные матрицы 394
§ 9. Осцилляционные матрицы 398
Глава XIV. Различные критерии регулярности и локализации собственных значений 406
§ 1. Критерий регулярности Адамара и его обобщения 406
§ 2. Норма матрицы 409
§ 3. Распространение критерия Адамара на блочные матрицы 412
§ 4. Критерий регулярности Фидлера 414
§ 5. Круги Гершгорина и другие области локализации 415
Глава XV. Приложения теории матриц к исследованию систем линейных дифференциальных уравнений 419
§ 1. Системы линейных дифференциальных уравнений с переменными коэффициентами. Общие понятия 419
§ 2. Преобразование Ляпунова 422
§ 3. Приводимые системы 423
§ 4. Каноническая форма приводимой системы. Теорема Еругина 426
§ 5. Матрицант 429
§ 6. Мультипликативный интеграл. Инфинитезимальное исчисление Вольтерра 433
§ 7. Дифференциальные системы в комплексной области. Общие свойства 437
§ 8. Мультипликативный интеграл в комплексной области 439
§ 9. Изолированная особая точка 443
§ 10. Регулярная особая точка 448
§ 11. Приводимые аналитические системы 461
§ 12. Аналитические функции от многих матриц и их применение к исследованию дифференциальных систем. Работы И. А. Лаппо-Данилевского 465

Глава XVI. Проблема Рауса — Гурвица и смежные вопросы **468**
§ 1. Введение 468
§ 2. Индексы Коши 469
§ 3. Алгоритмы Рауса 472
§ 4. Особые случаи. Примеры 476
§ 5. Теорема Ляпунова 479
§ 6. Теорема Рауса — Гурвица 483
§ 7. Формула Орландо 488
§ 8. Особые случаи в теореме Рауса — Гурвица 490
§ 9. Метод квадратичных форм. Определение числа различных вещественных корней многочлена 493
§ 10. Бесконечные ганкелевы матрицы конечного ранга 495
§ 11. Определение индекса произвольной рациональной дроби через коэффициенты числителя и знаменателя 498
§ 12. Второе доказательство теоремы Рауса — Гурвица 504
§ 13. Некоторые дополнения к теореме Рауса — Гурвица. Критерий устойчивости Льенара и Шипара 508
§ 14. Некоторые свойства многочлена Гурвица. Теорема Стильтьеса. Представление многочленов Гурвица при помощи непрерывных дробей 512
§ 15. Область устойчивости. Параметры Маркова 518
§ 16. Связь с проблемой моментов 521
§ 17. Связь между определителями Гурвица и определителями Маркова 525
§ 18. Теоремы Маркова и Чебышева 526
§ 19. Обобщенная задача Рауса — Гурвица 533
Добавление. Неравенства для собственных и сингулярных чисел (В.Б.Лидский) 535
Литература 560
Предметный указатель **572**

Предметный указатель

Алгоритм Гаусса 41
- - обобщенный 55
- - механическая интерпретация 45
- Рауса 472
Базис векторного пространства 66
- ортонормированный 224
Вектор 65
- главный пучка квадратичных форм 281
- - эрмитовых форм 301
- собственный линейного оператора 82
- - матрицы 82
- аннулирующий многочлен 171
- координаты 67
- минимальный многочлен 172
- норма 409
- относительный аннулирующий многочлен 177
- - минимальный многочлен 177
Векторы линейно зависимые и линейно независимые см. Зависимость линейных векторов
Вероятность абсолютная 390
- - предельная 391

- - средняя 393
- переходная 381
- - предельная 386
- - средняя 393
Вычитание матриц 16
Делитель элементарный квадратной матрицы 147
- - λ-матрицы 145
Дефект линейного оператора 77
- матрицы 77
Дискриминант квадратичной формы 267
- эрмитовой формы 298
Дополнение ортогональное 241
Жорданова форма см. Форма нормальная квадратной матрицы жорданова (верхняя)
Зависимость линейная векторов 66
- - критерий Грама 225
Задача Рауса — Гурвица обобщенная 533
Закон инерции квадратичных форм 269
- - эрмитовых форм 299
Индекс импримитивности 377
- Коши 469
Интеграл мультипликативный 434
- - в комплексной области 439
- от матрицы 127
Компонента линейного оператора кососимметрическая 255
- - симметрическая 255
Конгруэнтность пучков симметрических матриц 345
- симметрических матриц 269
Координаты вектора 67
Корень из матрицы 212
Критерий Адамара 406
- - для блочных матриц 413
- Грама линейной зависимости векторов 225
- Льенара и Шипара 508
- Ляпунова 425
- подобия матриц 150
- Рауса 475
- Рауса — Гурвица 486
- Фидлера 416
- эквивалентности λ-матриц 144
- - строгой регулярных пучков матриц 332
- - сингулярных пучков матриц 344
Линейный оператор см. Оператор линейный
Логарифм матрицы 219
λ-матрица 135
- инвариантные многочлены 143
- ранг 143
- элементарные делители 145
- - операции (левые и правые) 135, 136
Матрица 13
- ассоциированная *p*-я 30
- бесконечная вполне положительная 524
- - ганкелева 495
- блочная 55
- Вандермонда обобщенная 395
- вполне неотрицательная 395
- - положительная 395
- ганкелева 301
- главная пучка квадратичных форм 282
- Гурвица 483
- двояко стохастическая 536
- диагональная 14
- единичная 24
- жорданова (верхняя и нижняя) 153
- пдемпотептная 208
- импримитивная 377
- цнволютивная 81
- интегральная 418
- квадратная 13
- - аннулирующий многочлен 98
- - дефект 77
- - естественная нормальная форма (первая и вторая) 152

- - жорданова нормальная форма (верхняя и нижняя) 153, 154, 189, 190
- - инвариантные многочлены 147
- - компоненты 111
- - минимальный многочлен 98
- - перестановка рядов 352
- - порядок 13
- - разложение на треугольные множители 49
- - след 96
- - функция от нее 103
- - характеристический многочлен 92
- - элементарные делители 145
- квазидиагональная 56
- квазитроугольная (верхняя и нижняя) 56
- комплексная кососимметрическая 313
- - нормальная форма 322
- - неособенная, полярное разложение 317
- - ортогональная 313
- - нормальная форма 327
- - симметрическая 313
- - нормальная форма 319
- кососимметрическая 29
- Ляпунова 422
- многочленная см. λ-матрица
- - квадратная 87
- неособенная 26
- неотрицательная 352
- неразложимая 352
- - индекс импримитивности 377
- нильпотентная 208
- нормальная 244
- обратная 26
- ограниченная 132
- ортогональная 238
- особенная 26
- осцилляционная 398
Матрица положительная 352
- преобразующая 150
- примитивная 377
- присоединенная 92
- - приведенная 99
- проекционная 81
- простой структуры 85
- прямоугольная 13
- псевдообратная 32
- разложимая 352
- - нормальная форма 372
- Рауса 484
- симметрическая 29
- сопряженная 29
- столбцовая 14
- стохастическая 381
- строчная 14
- транспонированная 29
- треугольная (верхняя и нижняя) 28
- унитарная 238, 244
- фундаментальная 86
- характеристическая 92
- эрмитова 244
- якобиева 395
- интеграл 127
- корень из нее 212
- логарифм 219
- минор 14
- многочлен от нее 24
- производная 124
- ранг 14
- степень 24
- функция от нее см. Функция от матрицы
Матрицант 430
Матрицы перестановочные (коммутирующие) 18
- подобные 80
- вычитание 16
- сложение 15
- умножение 17
- - на число 16
Метод Гревилля нахождения псевдообратной матрицы 39
- Крылова 190

- Лагранжа приведения к сумме квадратов квадратичной формы 271
- - эрмитовой формы 299
- построения преобразующей матрицы 159
- Фаддеева 96
- Якоби приведения к сумме квадратов
квадратичной формы 272
- - эрмитовой формы 299
Метрика 223
- евклидова 225
- эрмитова 223
- - неотрицательная 223
- - положительно определенная 223
Минор 14
- главный 14
- почти главный 398
Многочлен аннулирующий вектора 171
- - матрицы 98
- - пространства 172
Многочлен инвариантный квадратной матрицы 145
- - λ-матрицы 143
- - интерполяционный Лагранжа—Сильвестра 104
- матричный 87
- - регулярный 87
- - левое и правое деление 87
- минимальный вектора 172
- - относительный 177
- - матрицы 98
- - пространства 172
- от матрицы 24
- характеристический 82, 92
Модуль линейного оператора левый 249
- - правый 249
Неравенства Вейля 251, 542
- Неймана — Хорна 539
- Сильвестра 78
Неравенство Адамара 230
- - обобщенное 231
- Бесселя 235
- Буняковского 232
- детерминантное для вполне неотрицательных матриц 396
Норма вектора 409
- линейного оператора 410
- матрицы 410
Оператор линейный 70
- - в R 79
- - вещественный 256
- - инволютивный 81
- - кососимметрический 255
- - неособенный 81
- - нормальный в евклидовом пространстве 254
- - в унитарном пространстве 242
- - ортогональный 255
- - второго рода 255
- - первого рода 255
- - проекционный 81
- - простой структуры 84
- - псевдообратный 265
- - симметрический 254
- - неотрицательный 254
- - положительно определенный 254
- - сопряженный 240
- - транспонированный 254
- - унитарный 243
- - эрмитов 243
- - неотрицательный 248
- - положительно определенный 248
- - дефект 77
- - инвариантное подпространство 173
- - кососимметрическая и симметрическая компоненты 255
- - левый и правый модули 249
- - полярное разложение в евклидовом пространстве 260

- - полярное разложение в унитарном пространстве 249
Оператор линейный, ранг 77
- - собственные векторы 82
- - характеристические (собственные)числа 82
Операторы линейные, сложение 71
- - умножение 71
- - умножение на число 73
Операция элементарная 135
- - Левая 136
- - правая 136
Определитель Грама 226
- - геометрический смысл 229
- Гурвипа 485
- Маркова 521, 525
Ортогонализация ряда векторов 233
Параметры Маркова 518
Перестановка рядов в квадратной матрице 352
Подпространство 68
- инвариантное 173
- координатное 353
- циклическое 178
Порядок матрицы 13
Предел последовательности матриц 49
Преобразование координат 70
- - ортогональное 238
- - унитарное 238
- линейное 15
- Ляпунова 428
Приведение квадратичной формы к главным осям 279
- - к сумме квадратов 271
- эрмитовой формы к главным осям 300
Проблема Рауса — Гурвица 468
Проектирование ортогональное 227
Проекционная матрица 81
Проекционный оператор 81
Производная мультипликативная 434
- от матрицы 124
Пространство 65
- бесконечномерное 66
- векторное см. Пространство евклидово 225
- конечномерное 66
- унитарное 223
- - ортоиормпрованный базис 224
- аннулирующий многочлен 171
- базис 66
- метрика см. Метрика
- минимальный многочлен 172
Псевдообратная матрица 32
Пучок квадратичных форм 281
- - регулярный 281
- - главная матрица 282
- - главный вектор 281
- - характеристические числа 281
- - характеристическое уравнение 281
- матриц 331
- - регулярный 332
- - каноническая форма 334
- - сингулярный 332
- - каноническая форма 340
Пучок матриц сингулярный, ранг 335
- эрмитовых форм 301
- - регулярный 301
- - главный вектор 301
- - характеристические числа 301
- - характеристическое уравнение 301
Равенство Парсеваля 237
Разложение квадратной матрицы на треугольные множители 49
- полярное комплексной матрицы 317
- - линейного оператора в евклидовом пространстве 260
- - в унитарном пространстве 249
Ранг квадратичной формы 268
- линейного оператора 77
- матрицы 14
- - многочленной 143
- сингулярного пучка матриц 335
- эрмитовой формы 298
Ряд векторов ортогональный 233

- - полный 236
- - ортогонализация 233
Сигнатура квадратичной формы 270
- эрмитовой формы 299
Система линейных дифференциальных уравнений с переменными коэффициентами 419
- - приводимая 423
- - аналитическая 461
- - интегральная матрица 419
- - преобразование Ляпунова 422
- - с постоянными коэффициентами 124
Системы векторов биортонормированные 241
Скелетное разложение матрицы 32
След матрицы 96
Сложение линейных операторов 71
- матриц 15
Собственное число см. Число собственное Степень матрицы 24
Схема Рауса 474
Сходимость в среднем 236
Теорема Безу обобщенная 92
- Гамильтона — Кэли 93
- Гершгорина 415
- Еругина 426
- Кронекера об ассоциированных матрицах 86
- Ляпунова 479
- Маркова 528
- Неймана — Хорна 541
- Ольги Тауски 409
- о расщеплении вторая 179
- - первая 173
Теорема о расщеплении третья 181
- Перрона 354
- Рауса 475
- Рауса — Гурвица 483
- Стильтьеса 512
- Фробениуса о ганкелевых формах 305
- - о неотрицательных матрицах 355
- Чебышева — Маркова 531
- Штурма 471
- Шура 242
Тождество Сильвестра детерминантное 47
Умножение линейных операторов 71
- матриц 17
- на число линейного оператора 73
- - матрицы 16
Уравнение матричное многочленное 209
- характеристическое (вековое) матрицы 82
- - пучка квадратичных форм 281
- - эрмитовых форм 301
Форма билинейная 267
- квадратичная 225, 267
- - вещественная 267
- - ганкелева 301
- - неотрицательная 276
- - неположительная 276
- - отрицательно определенная 276
- - положительно определенная 225, 276
- - сингулярная 267
- - дискриминант 267
- - закон инерции 269
- - приведение к главным осям 279
- - приведение к сумме квадратов 271
- - ранг 268
- - сигнатура 270
- - формула Якоби 272
- нормальная квадратной матрицы естественная первая 152
- - вторая 152
- - жорданова верхняя 153, 189
- - нижняя 154, 190
- - комплексной кососимметрической матрицы 322
- - ортогональной матрицы 327

- - симметрической матрицы 319
- - разложимой матрицы 372
- эрмитова 224, 297
- - билинейная 297
- - неотрицательная 224, 300
- - неположительная 300
- - отрицательно определенная 300
- - положительно определенная 224, 300
- - сингулярная 298
- - дискриминант 298
- - закон инерции 299
- - приведение к главным осям 300
- - формула Якоби 299
Формула Бине — Коши 20
- Орландо 488
Формула основная для функции от матрицы 111
- Чебыщева — Маркова 532
- Якоби для квадратичной формы 272
- - для эрмитовой формы 299
Формулы Кэли 252
- - в евклидовом пространстве 261
Функция от матрицы 103
- - основная формула 111
- от многих матриц аналитическая 465
Характеристическое число см. Число характеристическое
Цепочка векторов жорданова 188
- - нижняя 189
Цепь Маркова однородная 381
- - ациклическая 386
- - неразложимая 386
- - правильная 386
- - разложимая 386
- - регулярная 386
Цепь Маркова однородная циклическая 386
Число характеристическое (собственное) линейного оператора 82
- - пучка квадратичных форм 281
- - экстремальные свойства 286
- - эрмитовых форм 301
- - экстремальные свойства 301
Эквивалентность линейных двучленов 148
- λ-матриц 137
- - левая 137
- - правая 137
- - критерий 144
- рядов векторов 233
- строгая пучков матриц 331
- - регулярных, критерий 332
- - сингулярных, критерий 344
Элементарный делитель см. Делитель элементарный

ПРЕДИСЛОВИЕ АВТОРА К ПЕРВОМУ ИЗДАНИЮ

В настоящее время матричное исчисление широко применяется в различных областях математики, механики, теоретической физики, теоретической электротехники и т. д. В то же время ни в советской, ни в иностранной литературе нет книги, которая достаточно полно освещала бы как вопросы теории матриц, так и разнообразные ее приложения. Данная книга представляет собой попытку восполнить этот пробел в математической литературе. В основе книги лежат курсы лекций по теории матриц и ее приложениям, читанные автором в разное время на протяжении последних 17 лет в Московском Государственном университете им. М. В. Ломоносова, в Тбилисском Государственном университете и в Московском физико-техническом институте.

Книга рассчитана не только на математиков (студентов, аспирантов, научных работников), но и на специалистов в смежных областях (физиков, инженеров-исследователей), интересующихся математикой и ее приложениями. Поэтому автор стремился сделать изложение материала возможно более доступным, предполагая у читателя только знакомство с теорией определителей и курсом высшей математики в объеме программы втуза. Лишь отдельные параграфы в последних главах книги требуют дополнительных математических знаний у читателя. Кроме того, автор старался сделать изложение отдельных глав возможно более независимым друг от друга. Так, например, глава V «Функции от матрицы» не опирается на материал, помещенный в главах II и III. В тех же местах главы V, где впервые используются основные понятия, введенные в главе IV, имеются соответствующие ссылки. Таким образом, читатель, уже знакомый с элементами теории матриц, имеет возможность непосредственно приступить к чтению интересующих его глав книги.

Книга состоит из двух частей, содержащих 15 глав.

В главах I и III приводятся первоначальные сведения о матрицах и линейных операторах и устанавливается связь между операторами и матрицами.

В главе II излагаются теоретические основы метода исключения Гаусса и связанных с ним эффективных методов решения системы n линейных уравнений при большом n. В этой же главе читатель знакомится с техникой оперирования с матрицами, разбитыми на прямоугольные «клетки» или «блоки».

В главе IV вводятся имеющие фундаментальное значение «характеристический» и «минимальный» многочлены квадратной матрицы, «присоединенная» и «приведенная присоединенная» матрицы.

В главе V, посвященной функциям от матрицы, даются самое общее определение и конкретные способы вычисления $f(A)$, где $f(\lambda)$ — функция скалярного аргумента λ, а A — квадратная матрица. Понятие функции от матрицы используется в §§ 5, 6 этой главы для нахождения и полного

исследования решения системы линейных дифференциальных уравнений первого порядка с постоянными коэффициентами. Как понятие о функции от матрицы, так и связанное с ним исследование системы линейных дифференциальных уравнений с постоянными коэффициентами первого порядка опираются только на понятие о минимальном многочлене матрицы и не используют (в отличие от обычного изложения) так называемой «теории элементарных делителей», которая излагается в последующих главах VI и VII.

Первые пять глав охватывают некоторый цикл сведений о матрицах и их применениях. Более глубокие вопросы теории матриц связаны с приведением матрицы к нормальной форме. Это приведение проводится на основе теории элементарных делителей Вейерштрасса. Ввиду важности этой теории в книге даны два ее изложения: аналитическое — в главе VI и геометрическое — в главе VII. Обращаем внимание читателя на §§ 7 и 8 главы VI, в которых рассматриваются эффективные методы нахождения матрицы, преобразующей данную матрицу к нормальной форме. В § 8 главы VII подробно исследуется метод акад. А. Н. Крылова для практического вычисления коэффициентов характеристического многочлена.

В главе VIII решаются матричные уравнения некоторых типов. Здесь же рассматривается задача об определении всех матриц, перестановочных с данной, и детально изучаются многозначные функции от матрицы $\sqrt[m]{A}$, $\ln A$.

Главы IX и X посвящены теории линейных операторов в унитарном пространстве и теории квадратичных и эрмитовых форм. Эти главы не опираются на теорию элементарных делителей Вейерштрасса и используют из предыдущего материала лишь основные сведения о матрицах и линейных операторах, изложенные в первых трех главах книги. В § 9 главы X дается приложение теории форм к исследованию главных колебаний системы с n степенями свободы. В § 10 этой же главы приведены тонкие исследования Фробениуса по теории ганкелевых форм. Эти исследования применяются в дальнейшем в главе XV при рассмотрении особых случаев в проблеме Рауса — Гурвица.

Последние пять глав составляют вторую часть книги. В главе XI определяются нормальные формы для комплексных симметрических, кососимметрических и ортогональных матриц и устанавливаются интересные связи этих матриц с вещественными матрицами тех же классов и с унитарными матрицами.

В главе XII излагается общая теория пучков матриц вида $A + \lambda B$, где A и B — произвольные прямоугольные матрицы одних и тех же размеров. Подобно тому как исследование регулярных пучков матриц $A + \lambda B$ проводится на основе теории элементарных делителей Вейерштрасса, изучение сингулярных пучков опирается на теорию минимальных индексов Кронекера, которая является как бы дальнейшим развитием теории элементарных делителей Вейерштрасса. С помощью теории Кронекера (автору кажется, что ему удалось упростить изложение этой теории) в главе XII устанавливается каноническая форма пучка матриц $A + \lambda B$ в самом общем случае. Полученные результаты применяются к исследованию системы линейных дифференциальных уравнений с постоянными коэффициентами.

В главе XIII излагаются замечательные спектральные свойства матриц с неотрицательными элементами и рассматриваются две важные

области применений матриц этого класса: 1) однородные цепи Маркова в теории вероятностей и 2) осцилляционные свойства упругих колебаний в механике. Матричный метод исследования однородных цепей Маркова получил свое развитие в работах В. И. Романовского [29] и опирается на тот факт, что матрица переходных вероятностей в однородной цепи Маркова с конечным числом состояний является матрицей с неотрицательными элементами специального типа («стохастическая матрица»).

Осцилляционные свойства упругих колебаний связаны с другим важным классом неотрицательных матриц — с «осцилляционными матрицами». Эти матрицы и их приложения были исследованы М. Г. Крейном совместно с автором настоящей книги. В главе XIII изложены только некоторые основные результаты из этой области. Подробное же изложение всего этого материала читатель найдет в монографии [7].

В главе XV собраны приложения теории матриц к системам дифференциальных уравнений с переменными коэффициентами. В этой главе центральное место (§§ 5—9) занимают теория мультипликативного интеграла и связанное с ним инфинитезимальное исчисление Вольтерра. Эти вопросы почти совсем не освещены в советской математической литературе. В первых параграфах и в § 11 изучаются приводимые (по Ляпунову) системы в связи с задачей об устойчивости движения и приводятся некоторые результаты Н. П. Еругина. §§ 9—11 относятся к аналитической теории систем дифференциальных уравнений. Здесь выясняется ошибочность основной теоремы Биркгоффа, которую обычно используют для исследования решения системы дифференциальных уравнений в окрестности особой точки, и устанавливается канонический вид решения в случае регулярной особой точки.

В § 12 главы XV в обзорном порядке излагаются некоторые результаты фундаментальных исследований И. А. Лаппо-Данилевского по аналитическим функциям от многих матриц и их применениям к дифференциальным системам.

Последняя глава (XVI) посвящена применениям теории квадратичных форм (и, в частности, ганкелевых форм) к проблеме Рауса — Гурвица об определении числа корней многочлена, лежащих в правой полуплоскости ($\operatorname{Re} z > 0$). В первых параграфах этой главы приводится классическая трактовка вопроса. В § 5 дана теорема А. М. Ляпунова, в которой устанавливается критерий устойчивости, эквивалентный критерию Рауса — Гурвица. Наряду с критерием устойчивости Рауса — Гурвица в § 11 этой главы выводится сравнительно мало известный критерий Льенара и Шипара, в котором число детерминантных неравенств примерно вдвое меньше, нежели в критерии Рауса — Гурвица.

В конце главы XVI показана тесная связь с задачами устойчивости двух замечательных теорем А. А. Маркова и П. Л. Чебышева, которые были получены знаменитыми авторами на основе теории разложения в ряд по убывающим степеням аргумента некоторых непрерывных дробей специального типа. Здесь же дается матричное доказательство этих теорем.

Таков краткий перечень содержания настоящей книги.

В заключение автор приносит свою искреннюю благодарность Д. К. Фаддееву, В. П. Потапову и Д. М. Котелянскому, прочитавшим рукопись книги и сделавшим много существенных замечаний, которые были учтены автором при подготовке книги к печати. Автор выражает также свою благодарность М. Г. Крейну и А. И. Узкову за ценные советы, использованные автором при написании книги.

ПРЕДИСЛОВИЕ РЕДАКТОРА КО ВТОРОМУ ИЗДАНИЮ

Среди существующей литературы по теории матриц монография Ф. Р. Гантмахера занимает общепризнанно одно из лучших мест. Это объясняется систематичностью, широтой рассмотренных вопросов и четкостью изложения. Первое издание этой книги, вышедшее в 1953—1954 гг., было затем переведено на немецкий и английский языки.

В последние годы своей жизни Ф. Р. Гантмахер очень много времени уделил пересмотру и расширению этой книги. Изменения, сделанные им, частично касаются стиля (приведение некоторых терминов в соответствие с новыми традициями, улучшение отдельных доказательств и т. д.). Однако помимо этого было добавлено много нового материала, главным образом во второй, специальной, части книги. Отдельная новая глава XIV («Различные критерии регулярности и локализация собственных значений») посвящена различным методам приближенного отыскания собственных значений. Добавлены также § 5 гл. V («Некоторые свойства функций от матриц»), § 17 гл. XVI («Связь между определителями Гурвица и определителями Маркова») и два параграфа (§ 5 гл. I, § 16 гл. IX) о псевдообратных операторах и матрицах.

Известно, что автор был намерен включить в свою книгу ряд недавно разработанных вопросов, связанных с комбинаторикой собственных значений в алгебре матриц. К этим вопросам относится, в частности, задача о распределении собственных значений суммы и произведения двух матриц, а также известные неравенства Вейля и их обобщения. В настоящем издании соответствующее добавление было написано В. Б. Лидским, которому принадлежит одна из первых работ в этом направлении. В. Б. Лидский также принимал участие в подготовке и редактировании второго издания этой книги.

Можно надеяться, что некоторое увеличение объема книги не затруднит ее чтения, но, напротив, доставит читателям много интересной и ценной информации.

Д. П. Желобенко

ЧАСТЬ ПЕРВАЯ

ОСНОВЫ ТЕОРИИ

ГЛАВА I

МАТРИЦЫ И ДЕЙСТВИЯ НАД НИМИ

§ 1. Матрицы. Основные обозначения

1. Пусть дано некоторое числовое поле K[1]).

Определение 1. Прямоугольную таблицу чисел из поля K

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right\| \tag{1}$$

будем называть *матрицей*. Если $m=n$, то матрица называется *квадратной*, а число m, равное n, — ее *порядком*. В общем же случае матрица называется *прямоугольной* (с размерами $m \times n$) или $m \times n$-матрицей. Числа, составляющие матрицу, называются ее *элементами*.

Обозначения. При двухиндексном обозначении элементов первый индекс всегда указывает номер строки, а второй индекс — номер столбца, на пересечении которых стоит данный элемент.

Наряду с обозначениями матрицы (1) будем употреблять и сокращенное обозначение:

$$\| a_{ik} \| \quad (i=1, 2, \dots, m;\ k=1, 2, \dots, n).$$

Часто матрицу (1) будем обозначать также одной буквой, например матрица A. Если A — квадратная матрица порядка n, то будем писать: $A=\| a_{ik} \|_1^n$. Определитель квадратной матрицы $A=\| a_{ik} \|_1^n$ будем обозначать так: $| a_{ik} |_1^n$ или $| A |$.

Введем сокращенные обозначения для определителей, составленных из элементов данной матрицы:

$$A \begin{pmatrix} i_1 i_2 & \dots & i_p \\ k_1 k_2 & \dots & k_p \end{pmatrix} = \left| \begin{array}{cccc} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \dots & a_{i_2 k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_p k_1} & a_{i_p k_2} & \dots & a_{i_p k_p} \end{array} \right|. \tag{2}$$

[1]) Под *числовым полем* понимают любую совокупность чисел, в пределах которой всегда выполнимы и однозначны четыре операции: сложение, вычитание, умножение и деление на число, отличное от нуля.

Примерами числовых полей могут служить совокупность всех рациональных чисел, совокупность всех действительных чисел или совокупность всех комплексных чисел.

Предполагается, что все встречающиеся в дальнейшем числа принадлежат данному исходному числовому полю.

Определитель (3) называется *минором* p-го порядка матрицы A, если $1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant m$ и $1 \leqslant k_1 < k_2 < \ldots < k_p \leqslant n$.

$m \times n$-матрица $A = \| a_{ik} \|$ имеет $C_m^p C_n^p$ миноров p-го порядка

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \quad \left(\begin{matrix} 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant m \\ 1 \leqslant k_1 < k_2 < \ldots < k_p \leqslant n \end{matrix};\ p \leqslant m, n\right). \tag{2'}$$

Миноры (2'), у которых $i_1 = k_1$, $i_2 = k_2$, ..., $i_p = k_p$, называются *главными*.

В обозначениях (2) определитель квадратной матрицы $A = \| a_{ik} \|_1^n$ запишется так:

$$|A| = A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

Наибольший из порядков отличных от нуля миноров, порождаемых матрицей, называется *рангом* матрицы. Если r — ранг прямоугольной матрицы A с размерами $m \times n$, то, очевидно, $r \leqslant m$, n.

Прямоугольную матрицу, состоящую из одного столбца

$$\left\| \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} \right\|,$$

мы будем называть *столбцевой* и обозначать так: $(x_1, x_2, \ldots, x_n)$.

Прямоугольную матрицу, состоящую из одной строки

$$\| z_1, z_2, \ldots, z_n \|,$$

мы будем называть *строчной* и обозначать так: $[z_1, z_2, \ldots, z_n]$.

Квадратную матрицу, у которой все элементы, расположенные вне главной диагонали, равны нулю,

$$\left\| \begin{matrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & d_n \end{matrix} \right\|$$

мы будем называть *диагональной* и обозначать так: $\| d_i \delta_{ik} \|_1^n$[1]) или

$$\{d_1, d_2, \ldots, d_n\}.$$

Введем еще специальные обозначения для строк и столбцов $m \times n$-матрицы $A = \| a_{ik} \|$. Будем обозначать i-ю строку матрицы A через $a_{i.}$, а j-й столбец — через $a_{.j}$:

$$\begin{aligned} a_{i.} &= [a_{i1}, a_{i2}, \ldots, a_{in}], \\ a_{.j} &= [a_{1j}, a_{2j}, \ldots, a_{mj}] \\ &(i = 1, \ldots, m;\ j = 1, \ldots, n). \end{aligned} \tag{3}$$

---

[1]) Здесь δ_{ik} — символ Кронекера: $\delta_{ik} = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k). \end{cases}$

Пусть m величин $y_1, y_2, \ldots, y_m$ выражаются линейно и однородно через n других величин $x_1, x_2, \ldots, x_n$:

$$\left.\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{aligned}\right\} \tag{4}$$

или, в сокращенной записи,

$$y_i = \sum_{k=1}^{n} a_{ik}x_k \quad (i = 1, 2, \ldots, m). \tag{4'}$$

Преобразование величин $x_1, x_2, \ldots, x_n$ в величины $y_1, y_2, \ldots, y_m$ при помощи формул (4) называется *линейным преобразованием*. Коэффициенты этого преобразования образуют $m \times n$-матрицу (1). Задание линейного преобразования (4) однозначно определяет матрицу (1) и наоборот.

В следующем параграфе, исходя из свойств линейных преобразований (4), мы определим основные операции над прямоугольными матрицами.

§ 2. Сложение и умножение прямоугольных матриц

Определим основные операции над матрицами: сложение матриц, умножение матрицы на число и умножение матриц.

1. Пусть величины $y_1, y_2, \ldots, y_m$ выражаются через величины $x_1, x_2 \ldots, x_n$ при помощи линейного преобразования

$$y_i = \sum_{k=1}^{n} a_{ik}x_k \qquad (i = 1, 2, \ldots, m), \tag{5}$$

а величины $z_1, z_2, \ldots, z_m$ — через те же величины $x_1, x_2, \ldots, x_n$ при помощи преобразования

$$z_i = \sum_{k=1}^{n} b_{ik}x_k \qquad (i = 1, 2, \ldots, m). \tag{6}$$

Тогда

$$y_i + z_i = \sum_{k=1}^{n} (a_{ik} + b_{ik})\, x_k \qquad (i = 1, 2, \ldots, m). \tag{7}$$

В соответствии с этим мы устанавливаем

Определение 2. *Суммой двух прямоугольных матриц* $A = \| a_{ik} \|$ и $B = \| b_{ik} \|$ одинаковых размеров $m \times n$ называется матрица $C = \| c_{ik} \|$ тех же размеров, элементы которой равны суммам соответствующих элементов данной матрицы:

$$C = A + B,$$

если

$$c_{ik} = a_{ik} + b_{ik} \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n).$$

Операция нахождения суммы данных матриц называется *сложением матриц*.

Пример.

$$\begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} + \begin{Vmatrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{Vmatrix} = \begin{Vmatrix} a_1+c_1 & a_2+c_2 & a_3+c_3 \\ b_1+d_1 & b_2+d_2 & b_3+d_3 \end{Vmatrix}.$$

Согласно определению 2, складывать можно только прямоугольные матрицы одинаковых размеров.

В силу этого же определения матрица коэффициентов в преобразовании (7) есть сумма матриц коэффициентов в преобразованиях (5) и (6).

Из определения сложения матриц непосредственно следует, что эта операция обладает переместительным и сочетательным свойствами:

1° $A+B=B+A,$

2° $(A+B)+C=A+(B+C).$

Здесь A, B, C — произвольные прямоугольные матрицы одинаковых размеров.

Операция сложения матриц естественным образом распространяется на случай любого числа слагаемых.

2. Умножим в преобразовании (5) величины $y_1, y_2, \ldots, y_m$ на некоторое число α из K. Тогда

$$\alpha y_i = \sum_{k=1}^{n} (\alpha a_{ik}) x_k \qquad (i=1, 2, \ldots, m).$$

В соответствии с этим имеет место

Определение 3. *Произведением матрицы* $A=\|a_{ik}\|$ $(i=1, 2, \ldots, m;$ $k=1, 2, \ldots, n)$ *на число* α из K называется матрица $C=\|c_{ik}\|$ $(i=1, 2, \ldots, m;$ $k=1, 2, \ldots, n)$, элементы которой получаются из соответствующих элементов матрицы A умножением на число α:

$$C=\alpha A,$$

если

$$c_{ik}=\alpha a_{ik} \qquad (i=1, 2, \ldots, m;\ k=1, 2, \ldots, n).$$

Операция нахождения произведения матрицы на число называется *умножением матрицы на число.*

Пример.

$$\alpha \begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} = \begin{Vmatrix} \alpha a_1 & \alpha a_2 & \alpha a_3 \\ \alpha b_1 & \alpha b_2 & \alpha b_3 \end{Vmatrix}.$$

Легко видеть, что

1° $\alpha(A+B)=\alpha A+\alpha B,$

2° $(\alpha+\beta)A=\alpha A+\beta A,$

3° $(\alpha\beta)A=\alpha(\beta A).$

Здесь A, B — прямоугольные матрицы одинаковых размеров, α, β — числа из поля K.

Разность $A-B$ двух прямоугольных матриц одинаковых размеров определяется равенством

$$A-B=A+(-1)B.$$

Если A — квадратная матрица порядка n, а α — число из K, то[1])

$$|\alpha A|=\alpha^n |A|.$$

[1]) Здесь символы $|A|$ и $|\alpha A|$ обозначают определители матриц A и αA (см. стр. 13).

3. Пусть величины $z_1, z_2, \ldots, z_m$ выражаются через величины $y_1, y_2, \ldots, y_n$ при помощи преобразования

$$z_i = \sum_{k=1}^{n} a_{ik} y_k \qquad (i = 1, 2, \ldots, m), \tag{8}$$

а величины $y_1, \ldots, y_n$ выражаются через величины $x_1, x_2, \ldots, x_q$ при помощи формул

$$y_k = \sum_{j=1}^{q} b_{kj} x_j \qquad (k = 1, 2, \ldots, n). \tag{9}$$

Тогда, подставляя эти выражения для y_k $(k = 1, 2, \ldots, n)$ в формулы (8), мы выразим $z_1, z_2, \ldots, z_m$ через $x_1, x_2, \ldots, x_q$ при помощи «составного» преобразования:

$$z_i = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{q} b_{kj} x_j = \sum_{j=1}^{q} \Big(\sum_{k=1}^{n} a_{ik} b_{kj} \Big) x_j \qquad (i = 1, 2, \ldots, m). \tag{10}$$

В соответствии с этим имеет место

Определение 4. *Произведением двух прямоугольных матриц*

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix}, \qquad B = \begin{Vmatrix} b_{11} & b_{12} & \ldots & b_{1q} \\ b_{21} & b_{22} & \ldots & b_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{nq} \end{Vmatrix}$$

называется матрица

$$C = \begin{Vmatrix} c_{11} & c_{12} & \ldots & c_{1q} \\ c_{21} & c_{22} & \ldots & c_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1} & c_{m2} & \ldots & c_{mq} \end{Vmatrix},$$

у которой элемент c_{ij}, стоящий на пересечении i-й строки и j-го столбца, равен «произведению» i-й строки первой матрицы A на j-й столбец второй матрицы B[1]):

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, q). \tag{11}$$

Операция нахождения произведения данных матриц называется *умножением матриц*.

Пример.

$$\begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} \begin{Vmatrix} c_1 & d_1 & e_1 & f_1 \\ c_2 & d_2 & e_2 & f_2 \\ c_3 & d_3 & e_3 & f_3 \end{Vmatrix} =$$

$$= \begin{Vmatrix} a_1c_1 + a_2c_2 + a_3c_3, & a_1d_1 + a_2d_2 + a_3d_3, & a_1e_1 + a_2e_2 + a_3e_3, & a_1f_1 + a_2f_2 + a_3f_3 \\ b_1c_1 + b_2c_2 + b_3c_3, & b_1d_1 + b_2d_2 + b_3d_3, & b_1e_1 + b_2e_2 + b_3e_3, & b_1f_1 + b_2f_2 + b_3f_3 \end{Vmatrix}.$$

[1]) Под произведением двух рядов чисел $a_1, a_2, \ldots, a_n$ и $b_1, b_2, \ldots, b_n$ мы понимаем сумму произведений соответствующих чисел этих рядов: $\sum_{i=1}^{n} a_i b_i$.

По определению 4 матрица коэффициентов в преобразовании (10) равна произведению матрицы коэффициентов в (8) на матрицу коэффициентов (9).

Заметим, что *операция умножения двух прямоугольных матриц выполнима лишь в том случае, когда число столбцов в первом сомножителе равно числу строк во втором.* В частности, умножение всегда выполнимо, если оба сомножителя — квадратные матрицы одного и того же порядка. Обратим внимание читателя и на то, что даже в этом частном случае умножение матриц не обладает переместительным свойством. Так, например,

$$\begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix} \begin{Vmatrix} 2 & 0 \\ 3 & -1 \end{Vmatrix} = \begin{Vmatrix} 8 & -2 \\ 18 & -4 \end{Vmatrix}, \quad \begin{Vmatrix} 2 & 0 \\ 3 & -1 \end{Vmatrix} \begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix} = \begin{Vmatrix} 2 & 4 \\ 0 & 2 \end{Vmatrix}.$$

Если $AB = BA$, то матрицы A и B называются *перестановочными* или *коммутирующими* между собой.

Пример. Матрицы

$$A = \begin{Vmatrix} 1 & 2 \\ -2 & 0 \end{Vmatrix} \text{ и } B = \begin{Vmatrix} -3 & 2 \\ -2 & -4 \end{Vmatrix}$$

перестановочны между собой, так как

$$AB = \begin{Vmatrix} -7 & -6 \\ 6 & -4 \end{Vmatrix}, \quad BA = \begin{Vmatrix} -7 & -6 \\ 6 & -4 \end{Vmatrix}.$$

Легко проверяется сочетательное свойство умножения матриц, а также распределительное свойство умножения относительно сложения:

$$\begin{aligned} &1^\circ && (AB)C = A(BC), \\ &2^\circ && (A+B)C = AC + BC, \\ &3^\circ && A(B+C) = AB + AC. \end{aligned} \tag{12}$$

Операция умножения матриц естественным образом распространяется на случай нескольких сомножителей.

4. Если воспользоваться произведением прямоугольных матриц, то линейное преобразование

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{aligned} \tag{13}$$

можно записать одним матричным равенством

$$\begin{Vmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{Vmatrix} = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Vmatrix},$$

или в сокращенной записи:

$$y = Ax. \tag{13'}$$

Здесь $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ — столбцевые матрицы, $A = \| a_{ik} \|$ — прямоугольная матрица размеров $m \times n$.

Равенства (13) выражают собой тот факт, что столбец y является линейной комбинацией столбцов матрицы A с коэффициентами $x_1, x_2, \ldots, x_n$:

$$y = x_1 a_{\cdot 1} + x_2 a_{\cdot 2} + \ldots + x_n a_{\cdot n} = \sum_{k=1}^{n} x_k a_{\cdot k}. \qquad (13'')$$

Вернемся теперь к равенствам (11), которые эквивалентны одному матричному равенству

$$C = AB. \qquad (14)$$

Эти равенства могут быть записаны в виде

$$c_{\cdot j} = \sum_{k=1}^{n} b_{kj} a_{\cdot k} \qquad (j = 1, 2, \ldots, q) \qquad (14')$$

или в виде

$$c_{i\cdot} = \sum_{k=1}^{n} a_{ik} b_{k\cdot} \qquad (i = 1, 2, \ldots, m). \qquad (14'')$$

Таким образом, любой j-й столбец матрицы-произведения $C = AB$ является линейной комбинацией столбцов первого сомножителя, т. е. матрицы A, причем коэффициенты этой линейной зависимости образуют j-й столбец во втором сомножителе B. Аналогично, любая i-я строка в матрице C является линейной комбинацией строк матрицы B, а коэффициентами этой линейной зависимости являются элементы i-й строки матрицы A[1]).

Остановимся еще на том частном случае, когда в произведении $C = AB$ второй сомножитель является квадратной и притом диагональной матрицей $B = \{d_1, d_2, \ldots, d_n\}$. Тогда из формул (11) следует:

$$c_{ij} = a_{ij} d_j \qquad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n),$$

т. е.

$$\begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} \begin{Vmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_n \end{Vmatrix} = \begin{Vmatrix} a_{11}d_1 & a_{12}d_2 & \ldots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \ldots & a_{2n}d_n \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1}d_1 & a_{m2}d_2 & \ldots & a_{mn}d_n \end{Vmatrix}.$$

Аналогично

$$\begin{Vmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_m \end{Vmatrix} \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} = \begin{Vmatrix} d_1a_{11} & d_1a_{12} & \ldots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \ldots & d_2a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ d_ma_{m1} & d_ma_{m2} & \ldots & d_ma_{mn} \end{Vmatrix}.$$

Таким образом, *при умножении прямоугольной матрицы A справа (слева) на диагональную матрицу $\{d_1, d_2, \ldots\}$ все столбцы (соответственно строки) матрицы A помножаются на числа $d_1, d_2, \ldots$*

[1]) Следовательно, матричные уравнения $AX = C$, где A и C — заданные соответственно $m \times n$- и $m \times q$-матрицы, а X — искомая $n \times q$-матрица, имеет решение в том и только в том случае, когда столбцы матрицы C являются линейными комбинациями столбцов матрицы A. Для уравнения $XB = C$ необходимое и достаточное условие существования решения X состоит в том, что строки матрицы C должны быть линейными комбинациями строк матрицы B.

5. Пусть квадратная матрица $C=\|c_{ij}\|_1^m$ является произведением двух прямоугольных матриц $A=\|a_{ik}\|$ и $B=\|b_{kj}\|$ соответственно размеров $m\times n$ и $n\times m$:

$$\begin{Vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{Vmatrix} = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix} \cdot \begin{Vmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nm} \end{Vmatrix}, \tag{15}$$

т. е.

$$c_{ij}=\sum_{\alpha=1}^{n} a_{i\alpha}b_{\alpha j} \qquad (i, j=1, 2, \dots, m). \tag{15'}$$

Установим важную формулу Бине—Коши, выражающую определитель $|C|$ через миноры матриц A и B:

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \sum_{1\leqslant k_1<k_2<\dots<k_m\leqslant n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \cdot \begin{vmatrix} b_{k_11} & \dots & b_{k_1m} \\ \dots & \dots & \dots \\ b_{k_m1} & \dots & b_{k_mm} \end{vmatrix} \tag{16}$$

или в специальных обозначениях (см. стр. 13):

$$C\begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1\leqslant k_1<k_2<\dots<k_n\leqslant n} A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \tag{16'}$$

Согласно этой формуле *определитель матрицы C равен сумме произведений всевозможных миноров максимального (m-го) порядка*[1]) *матрицы A на соответствующие миноры того же порядка матрицы B.*

Вывод формулы Бине—Коши. На основании формулы (15') определитель матрицы C можно представить в виде

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \begin{vmatrix} \sum_{\alpha_1=1}^{n} a_{1\alpha_1}b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^{n} a_{1\alpha_m}b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum_{\alpha_1=1}^{n} a_{m\alpha_1}b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^{n} a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1, \dots, \alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1}b_{\alpha_1 1} & \dots & a_{1\alpha_m}b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1}b_{\alpha_1 1} & \dots & a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1, \dots, \alpha_m=1}^{n} A\begin{pmatrix} 1 & 2 & \dots & m \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \end{pmatrix} b_{\alpha_1 1}b_{\alpha_2 2}\dots b_{\alpha_m m}. \tag{16''}$$

Если $m>n$, то среди чисел $\alpha_1, \alpha_2, \dots, \alpha_m$ всегда найдутся равные между собой числа и, следовательно, каждое слагаемое в правой части равенства (16'') будет равно нулю. Значит, в этом случае $|C|=0$.

Пусть теперь $m\leqslant n$. Тогда в сумме, стоящей в правой части равенства (16''), будут равны нулю те слагаемые, у которых хотя бы два из индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ равны между собой. Все же остальные слагаемые этой суммы можно разбить на группы по $m!$ слагаемых

[1]) Если $m>n$, то матрицы A и B не имеют миноров m-го порядка. В этом случае правые части формул (16) и (16') следует заменить нулями.

в каждой, объединяя в одну группу те слагаемые, которые отличаются друг от друга только порядком индексов $\alpha_1, \alpha_2, \ldots, \alpha_m$ (индексы $\alpha_1, \alpha_2, \ldots, \alpha_m$ в пределах каждой группы слагаемых имеют одну и ту же совокупность значений). Тогда в пределах одной такой группы сумма соответствующих слагаемых будет равна [1])

$$\sum \varepsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \ldots b_{\alpha_m m} =$$

$$= A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} \sum \varepsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) b_{\alpha_1 1} b_{\alpha_2 2} \ldots b_{\alpha_m m} =$$

$$= A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \ldots & k_m \\ 1 & 2 & \ldots & m \end{pmatrix}.$$

Поэтому из (16″) получаем (16′).

Пример 1.

$$\begin{Vmatrix} a_1c_1+a_2c_2+\ldots+a_nc_n & a_1d_1+a_2d_2+\ldots+a_nd_n \\ b_1c_1+b_2c_2+\ldots+b_nc_n & b_1d_1+b_2d_2+\ldots+b_nd_n \end{Vmatrix} = \begin{Vmatrix} a_1 & a_2 & \ldots & a_n \\ b_1 & b_2 & \ldots & b_n \end{Vmatrix} \begin{Vmatrix} c_1 & d_1 \\ c_2 & d_2 \\ \vdots & \vdots \\ c_n & d_n \end{Vmatrix}.$$

Поэтому формула (16) дает так называемое тождество Коши

$$\begin{vmatrix} a_1c_1+a_2c_2+\ldots+a_nc_n & a_1d_1+a_2d_2+\ldots+a_nd_n \\ b_1c_1+b_2c_2+\ldots+b_nc_n & b_1d_1+b_2d_2+\ldots+b_nd_n \end{vmatrix} = \sum_{1 \leqslant i < k \leqslant n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix} \begin{vmatrix} c_i & d_i \\ c_k & d_k \end{vmatrix}.$$

Полагая в этом тождестве $a_i = c_i$, $b_i = d_i$ $(i = 1, 2, \ldots, n)$, получим:

$$\begin{vmatrix} a_1^2+a_2^2+\ldots+a_n^2 & a_1b_1+a_2b_2+\ldots+a_nb_n \\ a_1b_1+a_2b_2+\ldots+a_nb_n & b_1^2+b_2^2+\ldots+b_n^2 \end{vmatrix} = \sum_{1 \leqslant i < k \leqslant n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix}^2.$$

В случае, когда a_i и b_i $(i = 1, 2, \ldots, n)$ — вещественные числа, отсюда следует известное неравенство

$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2 \leqslant (a_1^2+a_2^2+\ldots+a_n^2)(b_1^2+b_2^2+\ldots+b_n^2).$$

При этом знак равенства имеет место тогда и только тогда, когда все числа a_i пропорциональны соответствующим числам b_i $(i = 1, 2, \ldots, n)$.

Пример 2.

$$\begin{Vmatrix} a_1c_1+b_1d_1 & \ldots & a_1c_n+b_1d_n \\ \cdots & \cdots & \cdots \\ a_nc_1+b_nd_1 & \ldots & a_nc_n+b_nd_n \end{Vmatrix} = \begin{Vmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{Vmatrix} \begin{Vmatrix} c_1 & \ldots & c_n \\ d_1 & \ldots & d_n \end{Vmatrix}.$$

Поэтому [2]) при $n > 2$

$$\begin{vmatrix} a_1c_1+b_1d_1 & \ldots & a_1c_n+b_1d_n \\ \cdots & \cdots & \cdots \\ a_nc_1+b_nd_1 & \ldots & a_nc_n+b_nd_n \end{vmatrix} = 0.$$

---

[1]) Здесь $k_1 < k_2 < \ldots < k_m$ — нормальное расположение индексов $\alpha_1, \alpha_2, \ldots, \alpha_m$, а $\varepsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) = (-1)^N$, где N — число транспозиций индексов, необходимых для преобразования перестановки $\alpha_1, \alpha_2, \ldots, \alpha_m$ к нормальному расположению $k_1 < k_2 < \ldots < k_m$.

[2]) См. сноску на стр. 20.

Рассмотрим частный случай, когда A и B — квадратные матрицы одного и того же порядка n, и положим в (16′) $m=n$. Тогда приходим к известной теореме об умножении определителей:

$$C\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix}=A\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix}B\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix},$$

или в других обозначениях:

$$|C|=|AB|=|A|\cdot|B|. \tag{17}$$

Таким образом, *определитель произведения двух квадратных матриц равен произведению определителей перемножаемых матриц.*

6. Формула Бине — Коши дает возможность в самом общем случае выразить миноры произведения двух прямоугольных матриц через миноры сомножителей. Пусть

$$A=\|a_{ik}\|,\quad B=\|b_{kj}\|,\quad C=\|c_{ij}\|$$
$$(i=1,\ 2,\ \dots,\ m;\ k=1,\ 2,\ \dots,\ n;\ j=1,\ 2,\ \dots,\ q)$$

и

$$C=AB.$$

Рассмотрим произвольный минор матрицы C:

$$C\begin{pmatrix}i_1 & i_2 & \dots & i_p\\ j_1 & j_2 & \dots & j_p\end{pmatrix}\quad\left(1\leqslant\begin{matrix}i_1<i_2<\dots<i_p\leqslant m\\ j_1<j_2<\dots<j_p\leqslant q\end{matrix};\ p\leqslant m,\ q\right).$$

Матрица, составленная из элементов этого минора, представляет собой произведение двух прямоугольных матриц

$$\begin{Vmatrix}a_{i_1 1} & a_{i_1 2} & \dots & a_{i_1 n}\\ \dots & \dots & \dots & \dots\\ a_{i_p 1} & a_{i_p 2} & \dots & a_{i_p n}\end{Vmatrix},\qquad\begin{Vmatrix}b_{1 j_1} & \dots & b_{1 j_p}\\ b_{2 j_1} & \dots & b_{2 j_p}\\ \dots & \dots & \dots\\ b_{n j_1} & & b_{n j_p}\end{Vmatrix}.$$

Поэтому, применяя формулу Бине—Коши, получаем:

$$C\begin{pmatrix}i_1 & i_2 & \dots & i_p\\ j_1 & j_2 & \dots & j_p\end{pmatrix}=\sum_{1\leqslant k_1<k_2<\dots<k_p\leqslant n}A\begin{pmatrix}i_1 & i_2 & \dots & i_p\\ k_1 & k_2 & \dots & k_p\end{pmatrix}B\begin{pmatrix}k_1 & k_2 & \dots & k_p\\ j_1 & j_2 & \dots & j_p\end{pmatrix}^{1)} \tag{18}$$

При $p=1$ формула (18) переходит в формулу (11). При $p>1$ формула (18) является естественным обобщением формулы (11).

Отметим еще одно следствие из формулы (18).

Ранг произведения двух прямоугольных матриц не превосходит ранга любого из сомножителей.

Если $C=AB$ и r_A, r_B, r_C — ранги матриц A, B, C, то

$$r_C\leqslant r_A,\ r_B.$$

7. Если X — решение матричного уравнения $AX=C$ (размеры матриц A, X и C соответственно $m\times n$, $n\times q$ и $m\times q$), то $r_X\geqslant r_C$. Покажем, что *среди решений матричного уравнения $AX=C$ существует решение X_0 минимального ранга, для которого $r_{X_0}=r_C$.*

---

[1]) Из той же формулы Бине — Коши следует, что миноры p-го порядка матрицы C при $p>n$ (если миноры таких порядков имеются) все равны нулю. В этом случае правую часть формулы (18) следует заменить нулем. См. сноску на стр. 20.

Действительно, пусть $r=r_C$. Тогда среди столбцов матрицы C имеется r линейно независимых [1]). Пусть для конкретности первые r столбцов $C_{\cdot 1}, \ldots, C_{\cdot r}$ линейно независимы, а остальные столбцы $C_{\cdot r+1}, \ldots, C_{\cdot q}$ являются линейными комбинациями первых r:

$$C_{\cdot j}=\sum_{k=1}^{r} \alpha_{jk} C_{\cdot k} \qquad (j=r+1, \ldots, q). \tag{19}$$

Пусть X — произвольное решение уравнения $AX=C$. Тогда (см. стр. 19)

$$AX_{\cdot k}=C_{\cdot k} \qquad (k=1, \ldots, r). \tag{20}$$

Определим столбцы $\widetilde{X}_{\cdot r+1}, \ldots, \widetilde{X}_{\cdot q}$ равенствами

$$X_{\cdot j}=\sum_{k=1}^{r} \alpha_{jk} X_{\cdot k} \qquad (j=r+1, \ldots, q).$$

Умножая эти равенства слева почленно на A, в силу равенств (19) и (20) находим:

$$A\widetilde{X}_{\cdot j}=C_{\cdot j} \qquad (j=r+1, \ldots, q). \tag{20'}$$

Система из q равенств (20) и (20′) эквивалентна одному матричному равенству

$$AX_0=C,$$

где $X_0=(X_{\cdot 1}, \ldots, X_{\cdot r}, \widetilde{X}_{\cdot r+1}, \ldots, \widetilde{X}_{\cdot q})$ — матрица ранга r [2]).

Решение X_0 минимального ранга r_C матричного уравнения $AX=C$ всегда представимо в виде

$$X_0=VC,$$

где V — некоторая $n \times m$-матрица.

Действительно, из равенства $AX_0=C$ следует, что строки матрицы C являются линейными комбинациями строк матрицы X_0. Поскольку как среди строк матрицы C, так и среди строк матрицы X_0 имеется одно и то же число r_C линейно независимых [3]), то и, обратно, строки матрицы X_0 являются линейными комбинациями строк матрицы C, а отсюда уже следует равенство $X_0=VC$.

Докажем теперь следующее предложение [4]).

Матричное уравнение

$$AXB=C, \tag{21}$$

где A, B — заданные, а X — искомая прямоугольная матрица [5]), *имеет решение в том и только том случае, когда одновременно имеют решения матричные уравнения*

$$AY=C, \qquad ZB=C, \tag{22}$$

т. е. когда столбцы матрицы C являются линейными комбинациями столбцов матрицы A, а строки матрицы C являются линейными комбинациями строк матрицы B.

В самом деле, если матрица X — решение уравнения (21), то матрицы $Y=XB$ и $Z=AX$ являются решениями уравнений (22).

Обратно, пусть существуют решения Y, Z уравнений (22). Тогда первое из этих уравнений имеет решение Y_0 минимального ранга r_C, которое по доказанному представимо в виде

$$Y_0=VC.$$

Поэтому

$$C=AY_0=AVC=AVZB.$$

Тогда матрица $X=VZ$ будет решением уравнения (21).

---

[1]) Мы ссылаемся на хорошо известное положение: ранг матрицы равен числу линейно независимых столбцов (строк) матрицы. Доказательство этого положения приведено в гл. III, стр. 68.

[2]) В матрице X_0 последние $n-r$ столбцов являются линейными комбинациями первых r; первые же r столбцов $X_{\cdot 1}, \ldots, X_{\cdot r}$ линейно независимы, так как линейная зависимость между этими столбцами в силу равенств (20) повлекла бы линейную зависимость между столбцами $C_{\cdot 1}, \ldots, C_{\cdot r}$.

[3]) См. сноску на стр. 19.

[4]) См. [232], [199].

[5]) Предполагается, что размеры матриц A, X, B, C таковы, что произведение AXB имеет смысл и имеет размеры матрицы C.

§ 3. Квадратные матрицы

1. Квадратную матрицу n-го порядка, у которой на главной диагонали стоят единицы, а все остальные элементы равны нулю, будем называть *единичной матрицей* и обозначать через $E^{(n)}$ или просто E. Название «единичная матрица» связано со следующим свойством матрицы E: для любой прямоугольной матрицы

$$A = \| a_{ik} \| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

имеют место равенства

$$E^{(m)} A = A E^{(n)} = A.$$

Очевидно,

$$E^{(n)} = \| \delta_{ik} \|_1^n.$$

Пусть $A = \| a_{ik} \|_1^n$ — квадратная матрица. Тогда *степень матрицы* определяется обычным образом:

$$A^p = \underbrace{AA \ldots A}_{p \text{ раз}} \quad (p = 1, 2, \ldots); \qquad A^0 = E.$$

Из сочетательного свойства умножения матриц следует:

$$A^q = A^{p+q}.$$

Здесь p, q — произвольные целые неотрицательные числа.

Рассмотрим многочлен (целую рациональную функцию) с коэффициентами из поля K:

$$f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + \ldots + \alpha_m.$$

Тогда под $f(A)$ будем понимать матрицу

$$f(A) = \alpha_0 A^m + \alpha_1 A^{m-1} + \ldots + \alpha_m E.$$

Так определяется *многочлен от матрицы*.

Пусть многочлен $f(t)$ равен произведению многочленов $h(t)$ и $g(t)$:

$$f(t) = h(t)\, g(t).$$

Многочлен $f(t)$ получается из $h(t)$ и $g(t)$ путем почленного перемножения и приведения подобных членов. При этом используется правило перемножения степеней: $t^p t^q = t^{p+q}$. Так как все эти действия правомерны и при замене скалярной величины t на матрицу A, то

$$f(A) = (A)\, g(A).$$

Отсюда, в частности,

$$h(A)\, g(A) = g(A)\, h(A)^1),$$

т. е. *два многочлена от одной и той же матрицы всегда перестановочны между собой.*

---

1) Так как каждое из этих произведений равно одному и тому же $f(A)$, поскольку и $g(t)\, h(t) = f(t)$. Следует отметить, что подстановка матриц в алгебраическое тождество с несколькими переменными неправомерна. Впрочем, перестановочные между собой матрицы можно подставлять и в этом случае.

Примеры.

Условимся *p-й наддиагональю* (*поддиагональю*) в прямоугольной матрице $A=\|a_{ik}\|$ называть ряд элементов a_{ik}, у которых $k-i=p$ (соответственно $i-k=p$). Обозначим через H квадратную матрицу n-го порядка, у которой элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Тогда

$$H=\begin{Vmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & \dots & 0 \end{Vmatrix}, \qquad H^2=\begin{Vmatrix} 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & \dots & 0 \end{Vmatrix} \text{ и т. д.;}$$

$$H^p=0 \quad (p\geqslant n).$$

В силу этих равенств, если

$$f(t)=a_0+a_1t+a_2t^2+\dots+a_{n-1}t^{n-1}+\dots$$

—многочлен относительно t, то

$$f(H)=a_0E+a_1H+a_2H^2+\dots=\begin{Vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & a_0 & a_1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_2 \\ \cdot & \cdot & \cdot & \cdot & a_1 \\ 0 & 0 & 0 & \dots & a_0 \end{Vmatrix}.$$

Аналогично, если F—квадратная матрица n-го порядка, у которой все элементы первой поддиагонали равны единице, а все остальные элементы равны нулю, то

$$f(F)=a_0E+a_1F+a_2F^2+\dots=\begin{Vmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 \\ a_{n-1} & \dots & a_1 & a_0 \end{Vmatrix}.$$

Предлагаем читателю проверить следующие свойства матриц H и F:

1° *В результате умножения произвольной* $m\times n$*-матрицы* A *слева на матрицу* H (*матрицу* F) m*-го порядка все строки матрицы* A *поднимаются* (*опускаются*) *на одно место вверх* (*вниз*), *первая* (*последняя*) *строка матрицы* A *исчезает, а последняя* (*первая*) *строка произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{Vmatrix}\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix}=\begin{Vmatrix} b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 0 \end{Vmatrix},$$

$$\begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix}=\begin{Vmatrix} 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{Vmatrix}.$$

2° *В результате умножения произвольной* $m\times n$*-матрицы* A *справа на матрицу* H (F) n*-го порядка все столбцы матрицы* A *сдвигаются вправо* (*влево*) *на одно место, при этом последний* (*первый*) *столбец матрицы* A *исчезает, а первый*

(*последний*) *столбец произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & b_1 & b_2 & b_3 \\ 0 & c_1 & c_2 & c_3 \end{Vmatrix},$$

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Vmatrix} = \begin{Vmatrix} a_2 & a_3 & a_4 & 0 \\ b_2 & b_3 & b_4 & 0 \\ c_2 & c_3 & c_4 & 0 \end{Vmatrix}.$$

2. Квадратную матрицу будем называть *особенной*, если $|A|=0$. В противном случае квадратная матрица A называется *неособенной*.

Пусть $A=\|a_{ik}\|_1^n$ — неособенная матрица ($|A|\neq 0$). Рассмотрим линейное преобразование с матрицей коэффициентов A

$$y_i=\sum_{k=1}^{n} a_{ik}x_k \quad (i=1,\ 2,\ \ldots,n). \tag{23}$$

Рассматривая равенства (23) как уравнения относительно $x_1, x_2, \ldots, x_n$ и замечая, что определитель системы уравнений (23) по условию отличен от нуля, мы можем однозначно по известным формулам выразить $x_1, x_2, \ldots, x_n$ через $y_1, y_2, \ldots, y_n$:

$$x_i=\frac{1}{|A|}\begin{Vmatrix} a_{11} \ldots a_{1,\,i-1}\ y_1\ a_{1,i+1} \ldots a_{1n} \\ a_{21} \ldots a_{2,\,i-1}\ y_2\ a_{2,\,i+1} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{n,\,i-1}\ y_n\ a_{n,\,i+1} \ldots a_{nn} \end{Vmatrix} \equiv \sum_{k=1}^{n} a_{ik}^{(-1)}y_k \ (i=1,\ 2,\ \ldots,n). \tag{24}$$

Мы получили «обратное» преобразование для (23). Матрицу коэффициентов этого преобразования

$$A^{-1}=\|a_{ik}^{(-1)}\|_1^n$$

мы назовем *обратной матрицей* для матрицы A. Из (24) легко усмотреть, что

$$a_{ik}^{(-1)}=\frac{A_{ki}}{|A|} \quad (i,\ k=1,\ 2,\ \ldots,n), \tag{25}$$

где A_{ki} — алгебраическое дополнение (адъюнкта) элемента a_{ki} в определителе $|A|$ $(i,\ k=1,\ 2,\ \ldots,n)$.

Так, например, если

$$A=\begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{Vmatrix} \text{ и } |A|\neq 0,$$

то

$$A^{-1}=\frac{1}{|A|}\begin{Vmatrix} b_2c_3-b_3c_2 & a_3c_2-a_2c_3 & a_2b_3-a_3b_2 \\ b_3c_1-b_1c_3 & a_1c_3-a_3c_1 & a_3b_1-a_1b_3 \\ b_1c_2-b_2c_1 & a_2c_1-a_1c_2 & a_1b_2-a_2b_1 \end{Vmatrix}.$$

Образуя составное преобразование из данного преобразования (23) и обратного (24) в одном и в другом порядке, мы в обоих случаях

получаем тождественное преобразование (с единичной матрицей коэффициентов); поэтому

$$AA^{-1}=A^{-1}A=E. \tag{26}$$

В справедливости равенств (26) можно убедиться и непосредственным перемножением матриц A и A^{-1}. Действительно, в силу (25) [1])

$$[AA^{-1}]_{ij}=\sum_{k=1}^{n}a_{ik}a_{kj}^{(-1)}=\frac{1}{|A|}\sum_{k=1}^{n}a_{ik}A_{jk}=\delta_{ij}\quad(i,\ j=1,\ 2,\ \ldots,n).$$

Аналогично

$$[A^{-1}A]_{ij}=\sum_{k=1}^{n}a_{ij}^{(-1)}a_{kj}=\frac{1}{|A|}\sum_{k=1}^{n}A_{ki}a_{kj}=\delta_{ij}\quad(i,\ j=1,\ 2,\ \ldots,n).$$

Нетрудно видеть, что матричные уравнения

$$AX=E\quad\text{и}\quad XA=E\qquad(|A|\neq 0) \tag{27}$$

никаких других решений, кроме решения $X=A^{-1}$ не имеют. Действительно, умножая обе части первого уравнения слева, а второго — справа на A^{-1} и используя сочетательное свойство произведения матриц, а также равенства (26), мы в обоих случаях получим [2]):

$$X=A^{-1}.$$

Этим же способом доказывается, что каждое из матричных уравнений

$$AX=B,\quad XA=B\qquad(|A|\neq 0), \tag{28}$$

где X и B — прямоугольные матрицы равных размеров, A — квадратная матрица соответствующего размера, имеет одно и только одно решение:

$$X=A^{-1}B\text{ и соответственно }X=BA^{-1}. \tag{29}$$

Матрицы (29) являются как бы «левым» и «правым» частными от «деления» матрицы B на матрицу A. Из (28) и (29) следует соответственно (см. стр. 22) $r_B\leqslant r_X$ и $r_X\leqslant r_B$, т. е. $r_X=r_B$. Сопоставляя с (28), имеем:

При умножении прямоугольной матрицы слева или справа на неособенную матрицу ранг исходной матрицы не изменяется.

Заметим еще, что из (26) вытекает $|A||A^{-1}|=1$, т. е.

$$|A^{-1}|=\frac{1}{|A|}.$$

Для произведения двух неособенных матриц имеем:

$$(AB)^{-1}=B^{-1}A^{-1}. \tag{30}$$

---

[1]) Здесь мы используем известное свойство определителя, согласно которому сумма произведений элементов какого-либо столбца на их адъюнкты равна величине определителя, а сумма произведений элементов столбца на адъюнкты соответствующих элементов другого столбца равна нулю.

[2]) Если A — особенная матрица, то уравнения (27) не имеют решений. Действительно, если бы какое-либо из этих уравнений имело решение $X=\|x_{ik}\|_1^n$, то тогда было бы в силу теоремы об умножении определителей [см. формулу (17)] $|A||X|=|E|=1$, что невозможно при $|A|=0$.

3. Все матрицы n-го порядка образуют кольцо [1]) с единичным элементом E. Поскольку в этом кольце определена операция умножения на число из поля K и существует базис из n^2 линейно независимых матриц, через которые линейно выражаются все матрицы n-го порядка [2]), то кольцо матриц n-го порядка является *алгеброй* [3]).

Все квадратные матрицы n-го порядка образуют коммутативную группу относительно операции сложения [4]). Все неособенные матрицы n-го порядка образуют (некоммутативную) группу относительно операции умножения.

Квадратная матрица $A=\|a_{ik}\|_1^n$ называется *верхней треугольной* (*нижней треугольной*), если равны нулю все элементы матрицы, расположенные под главной диагональю (над главной диагональю):

$$A=\begin{Vmatrix} a_{11} & a_{12} \dots a_{1n} \\ 0 & a_{22} \dots a_{2n} \\ \cdot & \cdot \quad \cdot \\ \cdot & \cdot \quad \cdot \\ \cdot & \cdot \quad \cdot \\ 0 & 0 \dots a_{nn} \end{Vmatrix}, \qquad A=\begin{Vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}.$$

(1) (2)

Диагональная матрица является частным случаем как верхней, так и нижней треугольной матрицы.

Так как определитель треугольной матрицы равен произведению ее диагональных элементов, то треугольная (и, в частности, диагональная) матрица является неособенной только тогда, когда все ее диагональные элементы отличны от нуля.

Легко проверить, что сумма и произведение двух диагональных (верхних треугольных, нижних треугольных) матриц есть диагональная (соответственно верхняя треугольная, нижняя треугольная) матрица и что обратная матрица для неособенной диагональной (верхней треугольной, нижней треугольной) матрицы является матрицей того же типа. Поэтому

1° Все диагональные, все верхние треугольные, все нижние треугольные матрицы n-го порядка образуют три коммутативные группы относительно операции сложения.

2° Все неособенные диагональные матрицы образуют коммутативную группу относительно умножения.

3° Все неособенные верхние (нижние) треугольные матрицы составляют группу (некоммутативную) относительно умножения.

---

[1]) *Кольцом* называется совокупность элементов, в которой определены и всегда однозначно выполнимы две операции: «сложение» двух элементов (с переместительным и сочетательным свойствами) и «умножение» двух элементов (с сочетательным и распределительным относительно сложения свойствами), причем сложение обратимо. См., например, [20], стр. 15, 25 и 100 или [39], стр. 333.

[2]) Действительно, произвольная матрица $A=\|a_{ik}\|_1^n$ с элементами из K представима в виде $A=\sum_{i,k=1}^{n} a_{ik}E_{ik}$, где E_{ik} — матрица n-го порядка, у которой на пересечении i-й строки и k-го столбца стоит 1, а все остальные элементы равны 0.

[3]) См., например, [20], стр. 101.

[4]) *Группой* называется всякая совокупность объектов, в которой установлена операция, относящая любым двум элементам a и b совокупности определенный третий элемент $a \times b$ той же совокупности, если 1) операция обладает сочетательным свойством $[(a \times b) \times c = a \times (b \times c)]$, 2) существует в совокупности единичный элемент e $(a \times e = e \times a = a)$ и 3) для любого элемента a совокупности существует обратный элемент a^{-1} $(a \times a^{-1} = a^{-1} \times a = e)$. Группа называется *коммутативной* или *абелевой*, если групповая операция обладает переместительным свойством. Относительно понятия группы см., например, [20], стр. 310 и далее.

4. В заключение этого параграфа укажем на две важные операции над матрицами — *транспонирование* матрицы и переход к *сопряженной* матрице.

Если $A = \| a_{ik} \|$ $(i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$, то транспонированная матрица A' определяется равенством $A' = \| a'_{ik} \|$, где $a'_{ki} = a_{ik}$ $(i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$. Сопряженная же матрица $A^* = \| a^*_{ik} \|$, где $a^*_{ki} = \overline{a'_{ki}} = \overline{a}_{ik}$ $(i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$[1]. Если матрица A имеет размеры $m \times n$, то матрицы A' и A^* имеют размеры $n \times m$.

Легко проверяются свойства[2]):

1° $(A + B)' = A' + B'$, $(A + B)^* = A^* + B^*$.

2° $(\alpha A)' = \alpha A'$, $(\alpha A)^* = \overline{\alpha} A^*$.

3° $(AB)' = B'A'$, $(AB)^* = B^*A^*$.

4° $(A^{-1})' = (A')^{-1}$, $(A^{-1})^* = (A^*)^{-1}$.

5° $(A')' = A$, $(A^*)^* = A$.

Если квадратная матрица $S = \| s_{ik} \|_1^n$ совпадает со своей транспонированной $(S' = S)$, то такая матрица называется *симметрической*. Если же квадратная матрица $H = \| h_{ik} \|$ совпадает со своей сопряженной $(H^* = H)$, то она называется *эрмитовой*. В симметрической матрице элементы, симметрично расположенные относительно главной диагонали, равны, а в эрмитовой они комплексно сопряжены между собой. Диагональные элементы эрмитовой матрицы всегда вещественны. Заметим, что произведение двух симметрических (эрмитовых) матриц, вообще говоря, не является симметрической (эрмитовой) матрицей. В силу 3° это имеет место только в том случае, когда данные две симметрические или эрмитовы матрицы перестановочны между собой.

Если A — *вещественная* матрица, т. е. матрица с вещественными элементами, то $A^* = A'$. Эрмитова вещественная матрица всегда является симметрической.

С каждой прямоугольной матрицей $A = \| a_{ik} \|$ размеров $m \times n$ связаны две эрмитовы матрицы AA^* и A^*A размеров $m \times m$ и $n \times n$. Любое из равенств $AA^* = 0$ или $A^*A = 0$ влечет за собой[3]) равенство $A = 0$.

Если квадратная матрица $K = \| k_{ij} \|_1^n$ отличается множителем -1 от своей транспонированной $(K' = -K)$, то такая матрица называется *кососимметрической*. В кососимметрической матрице любые два элемента, расположенные симметрично относительно главной диагонали, отличаются друг от друга множителем -1, а диагональные элементы равны нулю. Из 3° следует, что произведение двух перестановочных между собой кососимметрических матриц является симметрической матрицей[4]).

---

[1]) Чертой обозначается переход к комплексно сопряженной величине.

[2]) В формулах 1°, 2°, 3° и 5° A, B — произвольные прямоугольные матрицы, для которых соответствующие операции выполнимы, а α — произвольное комплексное число. В формуле 4° A — произвольная квадратная неособенная матрица.

[3]) Это следует из того, что сумма диагональных элементов в каждой из матриц AA^* и A^*A равна $\sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}|^2$.

[4]) Относительно представления квадратной матрицы A в виде произведения двух симметрических $(A = S_1 S_2)$ либо двух кососимметрических матриц $(A = K_1 K_2)$ см. [247].

§ 4. Ассоциированные матрицы. Миноры обратной матрицы

Пусть дана матрица $A = \| a_{ik} \|_1^n$. Рассмотрим всевозможные миноры p-го $(1 \leqslant p \leqslant n)$ порядка матрицы A:

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \left(1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n \right). \tag{31}$$

Число этих миноров равно N^2, где $N = C_n^p$ — число сочетаний из n по p. Для того чтобы расположить миноры (31) в квадратную таблицу, занумеруем в определенном (например, лексикографическом) порядке все сочетания по p из n индексов $1, 2, \dots, n$.

Если сочетания индексов $i_1 < i_2 < \dots < i_p$ и $k_1 < k_2 < \dots < k_p$ при этой нумерации будут иметь номера α и β, то минор (31) будем обозначать и так:

$$\mathfrak{a}_{\alpha\beta} = A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}.$$

Давая α и β независимо друг от друга все значения от 1 до N, мы охватим все миноры p-го порядка матрицы $A = \| a_{ik} \|_1^n$.

Квадратная матрица N-го порядка

$$\mathfrak{A}_p = \| \mathfrak{a}_{\alpha\beta} \|_1^N$$

называется p-й *ассоциированной матрицей* для матрицы $A = \| a_{ik} \|_1^n$; p может принимать значения $1, 2, \dots, n$. При этом $\mathfrak{A}_1 = A$, а матрица $\mathfrak{A}_n$ состоит из одного элемента, равного $|A|$.

Замечание. Порядок нумерации сочетаний индексов фиксируется раз навсегда и не связан с выбором матрицы A.

Пример. Пусть

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

Перенумеруем все сочетания из четырех индексов 1, 2, 3, 4 по два, расположив их в следующем порядке:

$$(1\ 2)\quad (1\ 3)\quad (1\ 4)\quad (2\ 3)\quad (2\ 4)\quad (3\ 4).$$

Тогда

$$\mathfrak{A}_2 = \begin{Vmatrix} A\binom{1\ 2}{1\ 2} & A\binom{1\ 2}{1\ 3} & A\binom{1\ 2}{1\ 4} & A\binom{1\ 2}{2\ 3} & A\binom{1\ 2}{2\ 4} & A\binom{1\ 2}{3\ 4} \\ A\binom{1\ 3}{1\ 2} & A\binom{1\ 3}{1\ 3} & A\binom{1\ 3}{1\ 4} & A\binom{1\ 3}{2\ 3} & A\binom{1\ 3}{2\ 4} & A\binom{1\ 3}{3\ 4} \\ A\binom{1\ 4}{1\ 2} & A\binom{1\ 4}{1\ 3} & A\binom{1\ 4}{1\ 4} & A\binom{1\ 4}{2\ 3} & A\binom{1\ 4}{2\ 4} & A\binom{1\ 4}{3\ 4} \\ A\binom{2\ 3}{1\ 2} & A\binom{2\ 3}{1\ 3} & A\binom{2\ 3}{1\ 4} & A\binom{2\ 3}{2\ 3} & A\binom{2\ 3}{2\ 4} & A\binom{2\ 3}{3\ 4} \\ A\binom{2\ 4}{1\ 2} & A\binom{2\ 4}{1\ 3} & A\binom{2\ 4}{1\ 4} & A\binom{2\ 4}{2\ 3} & A\binom{2\ 4}{2\ 4} & A\binom{2\ 4}{3\ 4} \\ A\binom{3\ 4}{1\ 2} & A\binom{3\ 4}{1\ 3} & A\binom{3\ 4}{1\ 4} & A\binom{3\ 4}{2\ 3} & A\binom{3\ 4}{2\ 4} & A\binom{3\ 4}{3\ 4} \end{Vmatrix}.$$

Отметим некоторые свойства ассоциированных матриц:

1° *Из* $C = AB$ *следует* $\mathfrak{C}_p = \mathfrak{A}_p \mathfrak{B}_p$ $(p = 1, 2, \dots, n)$.

Действительно, выражая миноры p-го порядка $(1 \leqslant p \leqslant n)$ матрицы-произведения C через миноры того же порядка матриц-сомножителей по формуле (18), будем иметь:

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_p & k_2 & \dots & k_p \end{pmatrix} = \sum_{1 \leqslant l_1 < l_2 < \dots < l_p \leqslant n} A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ l_1 & l_2 & \dots & l_p \end{pmatrix} B\begin{pmatrix} l_1 & l_2 & \dots & l_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}$$

$$\left(1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n\right). \tag{32}$$

Очевидно, в обозначениях этого параграфа равенство (32) может быть записано так:

$$\mathfrak{c}_{\alpha\beta} = \sum_{\lambda=1}^{N} \mathfrak{a}_{\alpha\lambda}\mathfrak{b}_{\lambda\beta} \qquad (\alpha, \beta = 1, 2, \dots, N)$$

(здесь α, β, λ — номера сочетаний индексов $i_1 < i_2 < \dots < i_p$; $k_1 < k_2 < \dots < k_p$; $l_1 < l_2 < \dots < l_p$). Отсюда

$$\mathfrak{E}_p = \mathfrak{A}_p \mathfrak{B}_p \quad (p = 1, 2, \dots, n).$$

2° *Из* $B = A^{-1}$ *следует*: $\mathfrak{B}_p = \mathfrak{A}_p^{-1}$ $(p = 1, 2, \dots, n)$.

Это предложение непосредственно вытекает из предыдущего, если там положить $C = E$ и обратить внимание на то, что $\mathfrak{E}_p$ — единичная матрица порядка $N = C_n^p$.

Из предложения 2° вытекает важная формула, выражающая миноры обратной матрицы через миноры данной матрицы:

Если $B = A^{-1}$, то при любых $(1 \leqslant) \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} (\leqslant n)$

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \frac{(-1)^{\sum_{\nu=1}^{p} i_\nu + \sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k_1' & k_2' & \dots & k_{n-p}' \\ i_1' & i_2' & \dots & i_{n-p}' \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}}, \tag{33}$$

где $i_1 < i_2 < \dots < i_p$ вместе с $i_1' < i_2' < \dots < i_{n-p}'$, а $k_1 < k_2 < \dots < k_p$ вместе с $k_1' < k_2' < \dots < k_{n-p}'$ составляют полную систему индексов $1, 2, \dots, n$.

Действительно, из $AB = E$ следует:

$$\mathfrak{A}_p \mathfrak{B}_p = \mathfrak{E}_p$$

или в более подробной записи:

$$\sum_{\alpha=1}^{N} \mathfrak{a}_{\gamma\alpha}\mathfrak{b}_{\alpha\beta} = \delta_{\gamma\beta} = \begin{cases} 1 & (\gamma = \beta), \\ 0 & (\gamma \neq \beta). \end{cases} \tag{34}$$

Равенства (34) могут быть записаны еще так:

$$\sum_{1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n} A\begin{pmatrix} j_1 & j_2 & \dots & j_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{cases} 1, \text{ если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 = 0, \\ 0, \text{ если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 > 0 \end{cases} \tag{34'}$$

$$\left(1 \leqslant \begin{matrix} j_1 < j_2 < \dots < j_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n\right).$$

С другой стороны, применяя к определителю $|A|$ известные разложения Лапласа, получаем:

$$\sum_{1\leqslant i_1<i_2<\ldots<i_p\leqslant n} A\begin{pmatrix} j_1 j_2 \ldots j_p \\ i_1 i_2 \ldots i_p \end{pmatrix}\cdot(-1)^{\sum_{\nu=1}^{p} i_\nu+\sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k_1' k_2' \ldots k_{n-p}' \\ i_1' i_2' \ldots i_{n-p}' \end{pmatrix} =$$

$$= \begin{cases} |A|, & \text{если } \sum\limits_{\nu=1}^{p}(j_\nu-k_\nu)^2=0, \\ 0, & \text{если } \sum\limits_{\nu=1}^{p}(j_\nu-k_\nu)^2>0, \end{cases} \tag{35}$$

где $i_1'<i_2'<\ldots<i_{n-p}'$ вместе с $i_1<i_2<\ldots<i_p$, а $k_1'<k_2'<\ldots<k_{n-p}'$ вместе с $k_1<k_2<\ldots<k_p$ образуют полную систему индексов $1, 2, \ldots, n$. Сопоставление (35) с (34′) и (34) показывает, что равенства (34) удовлетворятся, если вместо $\mathfrak{b}_{\alpha\beta}$ взять не $B\begin{pmatrix} i_1 i_2 \ldots i_p \\ k_1 k_2 \ldots k_p \end{pmatrix}$, а

$$\frac{(-1)^{\sum_{\nu=1}^{p} i_\nu+\sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k_1' k_2' \ldots k_{n-p}' \\ i_1' i_2' \ldots i_{n-p}' \end{pmatrix}}{A\begin{pmatrix} 1\,2\ldots n \\ 1\,2\ldots n \end{pmatrix}}.$$

Так как из системы (34) элементы $\mathfrak{b}_{\alpha\beta}$ обратной матрицы для $\mathfrak{A}_p$ определяются однозначно, то имеет место равенство (33).

§ 5. Обращение прямоугольных матриц. Псевдообратная матрица

Если A — квадратная и неособенная матрица, то для нее существует обратная матрица A^{-1}. Если же A — не квадратная, а прямоугольная $m\times n$-матрица $(m\neq n)$ или квадратная, но особенная, то матрица A не имеет обратной и символ A^{-1} не имеет смысла. Однако, как будет показано далее, для произвольной прямоугольной матрицы A существует «псевдообратная» матрица A^{+}, которая обладает некоторыми свойствами обратной матрицы и имеет важные применения при решении системы линейных уравнений. В случае, когда A — квадратная неособенная матрица, псевдообратная матрица A^{+} совпадает с обратной A^{-1}[1]).

1. Скелетное разложение матрицы. В дальнейшем мы будем пользоваться представлением произвольной прямоугольной $m\times n$-матрицы $A=\|a_{ik}\|$ ранга r в виде произведения двух матриц B и C, имеющих соответственно размеры $m\times r$ и $r\times n$:

$$A=BC=\begin{Vmatrix} b_{11} \ldots b_{1r} \\ b_{21} \ldots b_{2r} \\ \ldots\ldots \\ b_{m1} \ldots b_{mr} \end{Vmatrix}\begin{Vmatrix} c_{11}\ c_{12} \ldots c_{1n} \\ \ldots\ldots\ldots \\ c_{r1}\ c_{r2} \ldots c_{rn} \end{Vmatrix} \quad (r=r_A). \tag{36}$$

[1]) Приведенное далее в этом параграфе определение псевдообратной матрицы было дано в 1920 г. Муром, указавшим на важные применения этого понятия [214]. Позже независимо от Мура в несколько иной форме псевдообратная матрица определялась и исследовалась в работах Бьерхаммара [159], Пенрозе [227] и других авторов.

Здесь ранги сомножителей B и C обязательно равны рангу произведения A, $r_B = r_C = r$. Действительно (см. стр. 22), $r \leqslant r_B$, r_C. Но ранги r_B и r_C не могут превосходить r, так как r — один из размеров матриц B и C. Поэтому $r_B = r_C = r$.

Для того чтобы получить разложение (36), достаточно в качестве столбцов матрицы B взять любые r линейно независимых столбцов матрицы A, либо любые r линейно независимых столбцов, через которые линейно выражаются столбцы матрицы A [1]). Тогда произвольный j-й столбец матрицы A будет линейной комбинацией столбцов матрицы B с коэффициентами $c_{1j}, c_{2j}, \ldots, c_{rj}$; эти коэффициенты и образуют j-й столбец матрицы C ($j = 1, \ldots, n$, см. стр. 19) [2]).

Поскольку матрицы B и C имеют максимально возможный ранг r, то квадратные матрицы B^*B и CC^* являются неособенными:

$$|B^*B| \neq 0, \quad |CC^*| \neq 0. \tag{37}$$

Действительно, пусть столбец x — произвольное решение уравнения

$$B^*Bx = 0. \tag{38}$$

Помножим это уравнение слева на строку x^*. Тогда $x^*B^*Bx = (Bx)^* Bx = 0$. Отсюда [3]) следует $Bx = 0$ и (поскольку Bx — линейная комбинация линейно независимых столбцов матрицы B; ср. с формулой (13′)) $x = 0$. Из того, что уравнение (38) имеет только нулевое решение $x = 0$, вытекает, что $|B^*B| \neq 0$. Аналогично устанавливается второе неравенство (37) [4]).

Разложение (36) будем называть *скелетным* разложением матрицы A.

2. Существование и единственность псевдообратной матрицы. Рассмотрим матричное уравнение

$$AXA = A. \tag{39}$$

Если A — квадратная неособенная матрица, то это уравнение имеет единственное решение $X = A^{-1}$. Если же A — произвольная прямоугольная $m \times n$-матрица, то искомое решение X имеет размеры $n \times m$, но не определяется однозначно. В общем случае уравнение (39) имеет бесчисленное множество решений. Ниже будет показано, что среди этих решений имеется только одно, обладающее тем свойством, что его строки и столбцы являются линейными комбинациями соответственно строк и столбцов сопряженной матрицы A^*. Именно это решение мы будем называть псевдообратной матрицей для A и обозначать через A^+.

---

[1]) Мы исходим из известного положения: в матрице A ранга r имеется r линейно независимых столбцов, через которые линейно (т. е. в виде линейных комбинаций с числовыми коэффициентами из данного поля) выражаются все остальные столбцы. Аналогичное утверждение имеет место и для строк. Подробнее об этом см. гл. III, § 1.

[2]) Совершенно так же строками матрицы C могут быть любые r строк, через которые выражаются в виде линейных комбинаций все строки матрицы A. Тогда коэффициенты этих линейных комбинаций образуют строки матрицы B.

[3]) См. конец § 3.

[4]) Неравенства (37) также непосредственно следуют из формулы Бине — Коши. Согласно этой формуле определитель $|B^*B|$ ($|CC^*|$) равен сумме квадратов модулей всех миноров r-го порядка матрицы B (соответственно C).

Определение 5. Матрица A^+ размеров $n \times m$ называется *псевдообратной* для $m \times n$-матрицы A, если выполняются равенства [1])

$$AA^+A = A, \tag{40}$$

$$A^+ = UA^* = A^*V, \tag{41}$$

где U и V — некоторые матрицы

Докажем сначала, что для данной матрицы A не может существовать двух различных псевдообратных матриц A_1^+ и A_2^+. Действительно, из равенств

$$AA_1^+A = AA_2^+A = A, \quad A_1^+ = U_1A^* = A^*V_1, \quad A_2^+ = U_2A^* = A^*V_2,$$

полагая $D = A_2^+ - A_1^+$, $U = U_2 - U_1$, $V = V_2 - V_1$, найдем:

$$ADA = 0, \quad D = UA^* = A^*V.$$

Отсюда

$$(DA)^* DA = A^*D^*DA = A^*V^*ADA = 0$$

и, следовательно (см. конец § 3),

$$DA = 0.$$

Но тогда $DD^* = DAU = 0$, т. е. $D = A_2^+ - A_1^+ = 0$.

Для того чтобы установить существование матрицы A^+, мы воспользуемся скелетным разложением (36) и будем искать сначала псевдообратные матрицы B^+ и C^+ [2]). Так как по определению должны иметь место равенства

$$BB^+B = B, \quad B^+ = \widehat{U}B^*, \tag{42}$$

где $\widehat{U}$ — некоторая матрица, то

$$B\widehat{U}B^*B = B.$$

Умножая слева на B^* и замечая, что B^*B — неособенная квадратная матрица, найдем:

$$\widehat{U} = (B^*B)^{-1}.$$

Но тогда второе из равенств (42) дает искомое выражение для B^+:

$$B^+ = (B^*B)^{-1}B^*. \tag{43}$$

Совершенно аналогично найдем:

$$C^+ = C^*(CC^*)^{-1}. \tag{44}$$

Покажем теперь, что матрица

$$A^+ = C^+B^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^* \tag{45}$$

удовлетворяет условиям (40), (41) и, следовательно, является псевдообратной матрицей для A.

---

[1]) Условия (41) означают, что строки (столбцы) матрицы A^+ являются линейными комбинациями строк (столбцов) матрицы A^* (см. сноску к стр. 19). Условия (41) могут быть заменены одним условием $A^+ = A^*WA^*$, где W — некоторая матрица (см. конец § 2).

[2]) Из определения 5 сразу следует, что если $A = 0$, то и $A^+ = 0$. Поэтому в дальнейшем предполагается, что $A \neq 0$, и потому $r = r_A > 0$.

В самом деле,

$$AA^+A = BCC^*(CC^*)^{-1}(B^*B)^{-1}B^*BC = BC = A.$$

С другой стороны, из равенств (43), (44) и (45) с учетом равенства $A^* = C^*B^*$, полагая $K = (CC^*)^{-1}(B^*B)^{-1}$, находим

$$A^+ = C^*KB^* = C^*K(CC^*)^{-1}CC^*B^* = UC^*B^* = UA^*,$$

$$A^+ = C^*KB^* = C^*B^*B(B^*B)^{-1}KB^* = C^*B^*V = A^*V,$$

где

$$U = C^*K(CC^*)^{-1}C, \qquad V = B(B^*B)^{-1}KB^*.$$

Таким образом доказано, что для произвольной прямоугольной матрицы A существует одна и только одна псевдообратная матрица A^+, которая определяется формулой (45), где B и C — сомножители в скелетном разложении $A = BC$ матрицы A [1]). Из самого определения псевдообратной матрицы непосредственно следует, что в случае квадратной неособенной матрицы A псевдообратная матрица A^+ совпадает с обратной A^{-1}.

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Здесь $r = 2$. Примем в качестве столбцов матрицы B первые два столбца матрицы A. Тогда

$$A = BC = \begin{Vmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{Vmatrix} \begin{Vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}$$

и

$$B^*B = \begin{Vmatrix} 2 & -3 \\ -3 & 6 \end{Vmatrix}, \qquad (B^*B)^{-1} = \begin{Vmatrix} 2 & 1 \\ 1 & 2/3 \end{Vmatrix},$$

$$CC^* = \begin{Vmatrix} 3 & 0 \\ 0 & 3 \end{Vmatrix}, \qquad (CC^*)^{-1} = \begin{Vmatrix} 1/3 & 0 \\ 0 & 1/3 \end{Vmatrix} = \frac{1}{3}E.$$

Поэтому согласно формуле (45)

$$A^+ = \frac{1}{3} \begin{Vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \end{Vmatrix} \begin{Vmatrix} 2 & 1 \\ 1 & 2/3 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix}.$$

3. Свойства псевдообратной матрицы. Отметим следующие свойства псевдообратной матрицы:

1° $(A^*)^+ = (A^+)^*$;

2° $(A^+)^+ = A$;

3° $(AA^+)^* = AA^+$, $(AA^+)^2 = AA^+$;

4° $(A^+A)^* = A^+A$, $(A^+A)^2 = A^+A$.

---

[1]) Разложение (36) не определяет однозначно сомножителей BC. Однако поскольку, как было доказано, существует только одна псевдообратная матрица A^+, формула (45) при всех скелетных разложениях матрицы A дает одно и то же значение для A^+. В гл. II, § 5 будет изложен другой метод вычисления псевдообратной матрицы, использующий разбиение исходной матрицы на блоки (см. формулу (101) на стр. 64).

Первое свойство означает, что операции перехода к сопряженной и к псевдообратной матрице перестановочны между собой. Равенство 2° выражает собой взаимность понятия псевдообратной матрицы, так как согласно 2° псевдообратной матрицей для A^+ является исходная матрица A. Согласно равенствам 3° и 4° матрицы AA^+ и A^+A являются эрмитовыми и *инволютивными* (квадрат каждой из этих матриц равен самой матрице).

Для вывода равенства 1° воспользуемся скелетным разложением (36): $A=BC$. Тогда равенство $A^*=C^*B^*$ дает скелетное разложение матрицы A^*. Поэтому, заменяя в формуле (45) матрицу B на C^*, а матрицу C на B^*, получим:

$$(A^*)^+=B(B^*B)^{-1}(CC^*)^{-1}C=[C^*(CC^*)^{-1}(B^*B)^{-1}B^*]^*=(A^+)^*.$$

Равенства $A^+=C^+B^+$, $B^+=(B^*B)^{-1}B^*$, $C^+=C^*(CC^*)^{-1}$ являются скелетными разложениями. Следовательно,

$$(A^+)^+=(B^+)^+(C^+)^+=(B^*)^+B^*BCC^*(C^*)^+.$$

Используя свойство 1°, а также выражения для B^+ и C^+, найдем:

$$(A^+)^+=B(B^*B)^{-1}B^*BCC^*(CC^*)^{-1}C=BC=A.$$

Справедливость равенств 3° и 4° проверяется непосредственно путем подстановки в эти равенства вместо A^+ соответствующего выражения из формулы (45).

Заметим, что в общем случае, когда разложение $A=BC$ не является скелетным, не всегда имеет место равенство $A^+=C^+B^+$. Так, например,

$$A=\|\,1\,\|=\|\,0\quad 1\,\|\left\|\begin{matrix}1\\1\end{matrix}\right\|=BC.$$

Здесь

$$A^+=A^{-1}=\|\,1\,\|,\qquad B^+=\|\,0\quad 1\,\|^+=(\|\,1\,\|\cdot\|\,0\quad 1\,\|)^+=\left\|\begin{matrix}1\\0\end{matrix}\right\|\|\,1\,\|=\left\|\begin{matrix}1\\0\end{matrix}\right\|,$$

$$C^+=\left\|\begin{matrix}1\\1\end{matrix}\right\|^+=\left(\left\|\begin{matrix}1\\1\end{matrix}\right\|\cdot\|\,1\,\|\right)^+=\|\,1\,\|\cdot\|\,2\,\|^{-1}\|\,1\quad 1\,\|=\|\,1/2\quad 1/2\,\|.$$

Поэтому

$$C^+B^+=\|\,1/2\quad 1/2\,\|\begin{matrix}0\\1\end{matrix}\|\,\|=\|\,1/2\,\|\neq A^+.$$

4. **Наилучшее приближенное решение** (по методу наименьших квадратов). Рассмотрим произвольную систему линейных уравнений

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=y_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=y_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=y_m,\end{aligned}\tag{46}$$

или в матричной записи

$$Ax=y.\tag{46'}$$

Здесь $y_1, y_2, \ldots, y_m$ — заданные числа, а $x_1, x_2, \ldots, x_n$ — искомые.

В общем случае система (46) может быть и несовместной.

Столбец

$$x^0=(x^0, x_2^0, \ldots, x_n^0)\tag{47}$$

называется *наилучшим приближенным решением* системы (46), если при значениях $x_1 = x_1^0$, $x_2 = x_2^0$, ..., $x_n = x_n^0$ «квадратичное отклонение»

$$|y - Ax|^2 = \sum_{i=1}^{m} \left| y_i - \sum_{k=1}^{n} a_{ik} x_k \right|^2 \tag{48}$$

достигает своего наименьшего значения и среди всех столбцов x, для которых это отклонение имеет минимальное значение, столбец x^0 имеет наименьшую «длину», т. е. для этого столбца величина

$$|x|^2 = x^* x = \sum_{i=1}^{n} |x_i|^2 \tag{49}$$

имеет наименьшее значение.

Покажем, что система (46) всегда имеет одно и только одно наилучшее приближенное решение и это приближенное решение определяется по формуле

$$x^0 = A^+ y, \tag{50}$$

где A^+ — псевдообратная матрица для матрицы A.

Для этого рассмотрим произвольный столбец x и положим

$$y - Ax = u + v,$$

где

$$u = y - Ax^0 = y - AA^+ y, \qquad v = A(x^0 - x). \tag{51}$$

Тогда

$$|y - Ax|^2 = (y - Ax)^* (y - Ax) = \\ = (u + v)^* (u + v) = u^* u + v^* u + u^* v + v^* v. \tag{52}$$

Но

$$v^* u = (x^0 - x)^* A^* (y - AA^+ y) = (x^0 - x)^* (A^* - A^* AA^+) y. \tag{53}$$

Исходя из разложения (36) и формулы (45), найдем:

$$A^* AA^+ = C^* B^* BCC^* (CC^*)^{-1} (B^* B)^{-1} B^* = C^* B^* = A^*.$$

Поэтому из равенства (53) следует

$$v^* u = 0, \tag{54}$$

но тогда и

$$u^* v = (v^* u)^* = 0. \tag{54'}$$

Поэтому из равенства (52) находим:

$$|y - Ax|^2 = |u|^2 + |v|^2 = |y - Ax^0|^2 + |A(x^0 - x)|^2, \tag{55}$$

и, следовательно, для любого столбца x

$$|y - Ax| \geqslant |y - Ax^0|. \tag{56}$$

Пусть теперь

$$|y - Ax| = |y - Ax^0|;$$

тогда согласно равенству (55)

$$Az = 0, \tag{57}$$

где

$$z = x - x^0.$$

С другой стороны,

$$|x|^2=(x^0+z)^*(x^0+z)=|x^0|^2+|z|^2+(x^0)^*z+z^*x^0. \tag{58}$$

Вспоминая, что $A^+=A^*V$ (см. определение 5), получим в силу (57):

$$(x^0)^*z=(A^+y)^*z=(A^*Vy)^*z=y^*V^*Az=0. \tag{59}$$

Но тогда и

$$z^*x^0=((x^0)^*z)^*=0.$$

Поэтому из равенства (58) находим:

$$|x|^2=|x^0|^2+|z|^2,$$

и, следовательно,

$$|x|^2\geqslant|x^0|^2, \tag{60}$$

причем знак $=$ имеет место только при $z=0$, т. е. при $x=x^0$, где $x^0=A^+y$.

Пример. Найти наилучшее приближенное решение (по методу наименьших квадратов) системы линейных уравнений:

$$\begin{aligned} x_1-x_2+2x_3 &=3,\\ -x_1+2x_2-3x_3+x_4&=6,\\ x_2-x_3+x_4&=0. \end{aligned}$$

Здесь

$$A=\begin{Vmatrix} 1 & -1 & 2 & 0\\ -1 & 2 & -3 & 1\\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Но тогда (см. пример на стр. 35)

$$A^+=\begin{Vmatrix} 1/3 & 0 & 1/3\\ 1/9 & 1/9 & 2/9\\ 2/9 & -1/9 & 1/9\\ 4/9 & 1/9 & 5/9 \end{Vmatrix},$$

и потому

$$x^0=\begin{Vmatrix} x_1^0\\ x_2^0\\ x_3^0\\ x_4^0 \end{Vmatrix}=\begin{Vmatrix} 1/3 & 0 & 1/3\\ 1/9 & 1/9 & 2/9\\ 2/9 & -1/9 & 1/9\\ 4/9 & 1/9 & 5/9 \end{Vmatrix}\begin{Vmatrix} 3\\ 6\\ 0 \end{Vmatrix}.$$

Следовательно,

$$x_1^0=1,\quad x_2^0=1,\quad x_3^0=0,\quad x_4^0=2.$$

Определим норму $\|A\|$ $m\times n$-матрицы $A=\|a_{ik}\|$ как неотрицательное число, задаваемое формулой

$$\|A\|^2=\sum_{i,\,k}|a_{ik}|^2. \tag{61}$$

При этом очевидно, что

$$\|A\|^2=\sum_{k=1}^{n}|A_{\cdot k}|^2=\sum_{i=1}^{n}|A_{i\cdot}|^2. \tag{61'}$$

Рассмотрим матричное уравнение

$$AX=Y, \tag{62}$$

где A и Y — заданные $m\times n$- и $m\times p$-матрицы, а X — искомая $n\times p$-матрица.

Определим наилучшее приближенное решение X^0 уравнения (62) из условия

$$\|Y-AX^0\|=\min\|Y-AX\|,$$

причем в случае, когда

$$\|Y-AX\|=\|Y-AX^0\|,$$

требуется, чтобы

$$\|X^0\|\leqslant\|X\|.$$

Из соотношений

$$\|Y-AX\|^2=\sum_{k=1}^{p}|Y_{\cdot k}-AX_{\cdot k}|^2, \tag{63}$$

$$\|X\|^2=\sum_{k=1}^{p}|X_{\cdot k}|^2 \tag{64}$$

следует, что k-й столбец искомой матрицы $X^0_{\cdot k}$ должен быть наилучшим приближенным решением системы линейных уравнений

$$AX_{\cdot k}=Y_{\cdot k}.$$

Поэтому

$$X^0_{\cdot k}=A^+Y_{\cdot k}.$$

Поскольку это равенство справедливо при любом $k=1, \ldots, p$, то

$$X^0=A^+Y. \tag{65}$$

Таким образом, уравнение (62) всегда имеет одно и только одно наилучшее приближенное решение, определяемое формулой (65).

В частном случае, когда $Y=E$ — единичная матрица m-го порядка, имеем $X^0=A^+$. Следовательно, *псевдообратная матрица A^+ является наилучшим приближенным решением (по методу наименьших квадратов) матричного уравнения*

$$AX=E.$$

Это свойство псевдообратной матрицы A^+ может быть принято в качестве ее определения.

5. Метод Гревилля последовательного нахождения псевдообратной матрицы состоит в следующем. Пусть a_k — k-й столбец в $m\times n$-матрице A, $A_k=(a_1, \ldots, a_k)$ — матрица, образованная первыми k столбцами матрицы A. b_k — последняя строка в матрице A_k^+ ($k=1, \ldots, n$, $A_1=a_1$, $A_n=A$). Тогда [1])

$$A_1^+=a_1^+=\frac{a_1^*}{a_1^*a_1} \tag{66}$$

и для $k>1$ имеют место рекуррентные формулы

$$A_k^+=\begin{pmatrix}B_k\\ b_k\end{pmatrix},\quad B_k=A_{k-1}^+-d_kb_k,\quad d_k=A_{k-1}^+a_k; \tag{67}$$

при этом, если $c_k=a_k-A_{k-1}d_k\neq 0$, то

$$b_k=c_k^+=(a_k-A_{k-1}d_k)^+; \tag{68}$$

[1]) Если $A_1=a_1=0$, то и $A_1^+=0$.

если же $c_k = 0$, т. е. $a_k = A_{k-1}d_k$, то

$$b_k = (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^+. \tag{69}$$

Предлагаем читателю проверить, что матрица $\begin{pmatrix} B_k \\ b_k \end{pmatrix}$ является псевдообратной для матрицы A_k^+, если матрица B_k и строка b_k определяются формулами (61)—(64). Этот метод не требует вычисления детерминантов и может быть использован для вычисления обратной матрицы.

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \\ 0 & 1 & 1 \end{Vmatrix}.$$

Заметим, что для каждой вещественной матрицы M мы можем писать M' вместо M^*. Тогда

$$A_1^+ = (A_1'A_1)^{-1}A_1' = \frac{1}{6}A_1' = \| {}^1/_6,\ -{}^1/_6,\ {}^1/_3,\ 0 \|,$$

$$d_2 = A_1^+ a_2 = -{}^3/_2, \qquad c_2 = a_2 - A_1 d_2 = \begin{Vmatrix} {}^1/_2 \\ {}^1/_2 \\ 0 \\ 1 \end{Vmatrix},$$

$$b_2 = c_2^+ = (c_2'c_2)^{-1}c_2' = \frac{2}{3}c_2' = \| {}^1/_3,\ {}^1/_3,\ 0,\ {}^2/_3 \|,$$

$$B_2 = A_1^+ - d_2 b_2 = \| {}^2/_3,\ {}^1/_3,\ {}^1/_3,\ 1 \|.$$

Таким образом,

$$A_2^+ = \begin{Vmatrix} {}^2/_3 & {}^1/_3 & {}^1/_3 & 1 \\ {}^1/_3 & {}^1/_3 & 0 & {}^2/_3 \end{Vmatrix}.$$

Далее

$$d_3 = A_2^+ a_3 = \begin{Vmatrix} 1 \\ 1 \end{Vmatrix} \quad \text{и} \quad c_3 = a_3 - A_2 d_3 = 0.$$

Поэтому

$$b_3 = (1 + d_3'd_3)^{-1}d_3'A_2^+ = \| {}^1/_3,\ {}^1/_3 \| A_2^+ = \| {}^1/_3\ {}^2/_9\ {}^1/_9\ {}^5/_9 \|$$

и

$$B_3 = A_2^+ - d_3 b_3 = \begin{Vmatrix} {}^2/_3 & {}^1/_3 & {}^1/_3 & 1 \\ {}^1/_3 & {}^1/_3 & 0 & {}^2/_3 \end{Vmatrix} - \begin{Vmatrix} {}^1/_3 & {}^2/_9 & {}^1/_9 & {}^5/_9 \\ {}^1/_3 & {}^2/_9 & {}^1/_9 & {}^5/_9 \end{Vmatrix} = \begin{Vmatrix} {}^1/_3 & {}^1/_9 & {}^2/_9 & {}^4/_9 \\ 0 & {}^1/_9 & -{}^1/_9 & {}^1/_9 \end{Vmatrix}$$

$$A^+ = A_3^+ = \begin{Vmatrix} {}^1/_3 & {}^1/_9 & {}^2/_9 & {}^4/_9 \\ 0 & {}^1/_9 & -{}^1/_9 & {}^1/_9 \\ {}^1/_3 & {}^2/_9 & {}^1/_9 & {}^5/_9 \end{Vmatrix}.$$

ГЛАВА II

АЛГОРИТМ ГАУССА И НЕКОТОРЫЕ ЕГО ПРИМЕНЕНИЯ

§ 1. Метод исключения Гаусса

1. Пусть дана система n линейных уравнений с n неизвестными $x_1, x_2, \ldots, x_n$ и правыми частями $y_1, y_2, \ldots, y_n$:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= y_n. \end{aligned}\right\} \tag{1}$$

В матричной форме эта система может быть записана так:

$$Ax = y. \tag{1'}$$

Здесь $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ — столбцы и $A = \| a_{ik} \|_1^n$ — квадратная матрица коэффициентов.

Если A — неособенная матрица, то можно написать:

$$x = A^{-1}y \tag{2}$$

или в развернутом виде:

$$x_i = \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \quad (i = 1, 2, \ldots, n). \tag{2'}$$

Таким образом, задача вычисления элементов обратной матрицы $A^{-1} = \| a_{ik}^{(-1)} \|_1^n$ эквивалентна задаче решения системы уравнений (1) при любых правых частях $y_1, y_2, \ldots, y_n$. Элементы обратной матрицы определяются формулами (25) главы I. Однако фактическое вычисление элементов матрицы A^{-1} по этим формулам при большом n весьма затруднительно. Поэтому большое практическое значение имеют эффективные методы вычисления элементов обратной матрицы и, следовательно, решения системы линейных уравнений [1]).

В настоящей главе мы изложим теоретические основы некоторых из этих методов, представляющих собой разновидности метода исключения Гаусса, знакомство с которым у читателя началось еще в курсе алгебры средней школы.

---

[1]) Для подробного ознакомления с этими методами мы рекомендуем обратиться к книге Фаддеевых [33], а также к циклу статей, напечатанных в «Успехах математических наук», т. 5, вып. 3 (1950).

2. Пусть в системе уравнений (1) $a_{11} \neq 0$. Мы исключим x_1 из всех уравнений, начиная со 2-го, для чего ко второму уравнению почленно прибавим первое, помноженное на $-\frac{a_{21}}{a_{11}}$, к третьему почленно прибавим первое, помноженное на $-\frac{a_{31}}{a_{11}}$, и т. д. После этого система уравнений (1) заменится эквивалентной системой

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n2}^{(1)}x_2 + \ldots + a_{nn}^{(1)}x_n &= y_n^{(1)}. \end{aligned}\right\} \tag{3}$$

Коэффициенты при неизвестных и свободные члены в последних $(n-1)$ уравнениях определяются формулами

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}}a_{1j}, \qquad y_i^{(1)} = y_i - \frac{a_{i1}}{a_{11}}y_1 \qquad (i,\ j = 2,\ \ldots,\ n). \tag{3'}$$

Пусть $a_{22}^{(1)} \neq 0$. Тогда таким же образом мы исключим x_2 из последних $n-2$ уравнений системы (3) и получим систему уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots & \\ a_{n3}^{(1)}x_3 + \ldots + a_{nn}^{(2)}x_n &= y_n^{(2)}. \end{aligned}\right\} \tag{4}$$

При этом новые коэффициенты и правые части связаны с предыдущими формулами:

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}a_{2j}^{(1)}, \qquad y_i^{(2)} = y_i^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}y_2^{(1)} \qquad (i,\ j = 3,\ \ldots,\ n). \tag{5}$$

Продолжая этот алгоритм далее, мы на $(n-1)$-м этапе приведем исходную систему (1) к треугольной рекуррентной системе

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots & \\ a_{nn}^{(n-1)}x_n &= y_n^{(n-1)}. \end{aligned}\right\} \tag{6}$$

Это приведение выполнимо в том и только в том случае, если в процессе приведения все числа a_{11}, $a_{22}^{(1)}$, $a_{33}^{(2)}$, $\ldots$, $a_{n-1,\,n-1}^{(n-2)}$ оказываются отличными от нуля.

Изложенный нами алгоритм Гаусса состоит из однотипных операций, которые легко выполняются на современных счетных машинах.

3. Выразим коэффициенты и правые части приведенной системы через коэффициенты и правые части исходной системы (1). При этом мы не будем предполагать, что в процессе приведения все числа a_{11}, $a_{22}^{(1)}$, $\ldots$, $a_{n-1,\,n-1}^{(n-2)}$ оказываются отличными от нуля, а рассмотрим общий случай, когда первые p из этих чисел отличны от нуля:

$$a_{11} \neq 0,\ a_{22}^{(1)} \neq 0,\ \ldots,\ a_{pp}^{(p-1)} \neq 0 \qquad (p \leqslant n-1), \tag{7}$$

что дает возможность (на p-м этапе приведения) привести исходную систему уравнений к виду

$$\left.\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots + a_{1n}x_n &= y_1,\\
a_{22}^{(1)}x_2 + \ldots\ldots\ldots + a_{2n}^{(1)}x_n &= y_2^{(1)},\\
\ldots\ldots\ldots

Если условия (7) выполнены для данного значения p, то такие же условия выполнены для любого меньшего значения p. Поэтому формулы (13) имеют место не только для данного значения p, но и для всех меньших значений p. То же можно сказать и о формуле (11). Поэтому вместо этой формулы можно написать равенства

$$A\begin{pmatrix}1\\1\end{pmatrix}=a_{11},\quad A\begin{pmatrix}1&2\\1&2\end{pmatrix}=a_{11}a_{22}^{(1)},\quad A\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=a_{11}a_{22}^{(1)}a_{33}^{(2)},\ \ldots \tag{14}$$

Таким образом, условия (7), т. е. необходимые и достаточные условия выполнимости первых p этапов алгоритма Гаусса, могут быть записаны в виде следующих неравенств:

$$A\begin{pmatrix}1\\1\end{pmatrix}\neq 0,\quad A\begin{pmatrix}1&2\\1&2\end{pmatrix}\neq 0,\ \ldots,\quad A\begin{pmatrix}1&2&\ldots&p\\1&2&\ldots&p\end{pmatrix}\neq 0. \tag{15}$$

Тогда из (14) находим:

$$a_{11}=A\begin{pmatrix}1\\1\end{pmatrix},\quad a_{22}^{(1)}=\frac{A\begin{pmatrix}1&2\\1&2\end{pmatrix}}{A\begin{pmatrix}1\\1\end{pmatrix}},$$

$$a_{33}^{(2)}=\frac{A\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}}{A\begin{pmatrix}1&2\\1&2\end{pmatrix}},\ \ldots,\ a_{pp}^{(p-1)}=\frac{A\begin{pmatrix}1&2&\ldots&p\\1&2&\ldots&p\end{pmatrix}}{A\begin{pmatrix}1&2&\ldots&p-1\\1&2&\ldots&p-1\end{pmatrix}}. \tag{16}$$

Для того чтобы в алгоритме исключения Гаусса можно было последовательно исключить $x_1, x_2, \ldots, x_p$, нужно, чтобы все величины (16) были отличны от нуля, т. е. чтобы выполнялись неравенства (15). В то же время формулы для $a_{ik}^{(p)}$ имеют смысл, если выполняется только последнее из условий (15).

4. Пусть матрица коэффициентов в системе уравнений (1) имеет ранг r. Тогда надлежащей перестановкой уравнений и изменением нумерации неизвестных можно добиться выполнения неравенств

$$A\begin{pmatrix}1&2&\ldots&j\\1&2&\ldots&j\end{pmatrix}\neq 0\qquad (j=1, 2, \ldots, r). \tag{17}$$

Это позволяет последовательно исключить $x_1, x_2, \ldots, x_r$ и получить систему уравнений

$$\left.\begin{aligned}
a_{11}x_1+a_{12}x_2+\ldots\ldots\ldots\ldots+a_{1n}x_n &= y_1,\\
a_{22}^{(1)}x_2+\ldots\ldots\ldots\ldots+a_{2n}^{(1)}x_n &= y_2^{(1)},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
a_{rr}^{(r-1)}x_r+\ldots\ldots\ldots+a_{rn}^{(r-1)}x_n &= y_r^{(r-1)},\\
a_{r+1,\,r+1}^{(r)}x_{r+1}+\ldots+a_{r+1,\,n}^{(r)}x_n &= y_{r+1}^{(r)},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
a_{n,\,r+1}^{(r)}x_{r+1}+\ldots+a_{nn}^{(r)}x_n &= y_n^{(r)}.
\end{aligned}\right\} \tag{18}$$

Здесь коэффициенты определяются по формулам (13). Из этих формул, поскольку ранг матрицы $A=\|a_{ik}\|_1^n$ равен r, следует, что

$$a_{ik}^{(r)}=0\qquad (i, k=r+1, \ldots, n)$$

и матрица G_r, получающаяся из матрицы $A=\|a_{ik}\|_1^n$, после применения r-этапного алгоритма исключения Гаусса, имеет вид

$$G_r=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2r}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{Vmatrix}. \tag{19}$$

Последние $n-r$ уравнений (18) сводятся к условиям совместности

$$y_i^{(r)}=0 \qquad (i=r+1, \dots, n). \tag{20}$$

Заметим, что столбец свободных членов при алгоритме исключения подвергается таким же преобразованиям, как и любой столбец коэффициентов. Поэтому, дополняя матрицу $A=\|a_{ik}\|_1^n$ $(n+1)$-м столбцом из свободных членов, мы получим:

$$y_i^{(p)}=\frac{A\begin{pmatrix}1 \dots p & i\\ 1 \dots p & n+1\end{pmatrix}}{A\begin{pmatrix}1 \dots p\\ 1 \dots p\end{pmatrix}} \qquad (i=1, 2, \dots, n;\ p=1, 2, \dots, r). \tag{21}$$

В частности, условия совместности (20) сводятся к известным условиям

$$A\begin{pmatrix}1 \dots r & r+j\\ 1 \dots r & n+1\end{pmatrix}=0 \qquad (j=1, 2, \dots, n-r). \tag{22}$$

Если $r=n$, т. е. матрица $A=\|a_{ik}\|_1^n$ неособенная, и

$$A\begin{pmatrix}1 & 2 \dots j\\ 1 & 2 \dots j\end{pmatrix}\neq 0 \qquad (j=1, 2, \dots, n),$$

то при помощи алгоритма Гаусса можно последовательно исключить $x_1, x_2, \dots, x_{n-1}$ и привести систему уравнений к виду (6).

§ 2. Механическая интерпретация алгоритма Гаусса

Рассмотрим произвольную упругую статическую систему S, закрепленную на краях (например, струну, стержень, многопролетный стержень, мембрану, пластину или дискретную систему), и возьмем на ней n точек (1), (2), ..., (n).

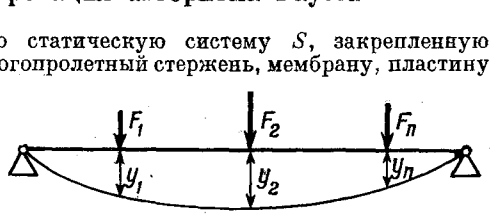


Рис. 1.

Мы будем рассматривать перемещения (прогибы) $y_1, y_2, \dots, y_n$ точек (1), (2), ..., (n) системы S под действием сил $F_1, F_2, \dots, F_n$, приложенных в этих же точках. Мы будем предполагать, что силы и перемещения параллельны одному и тому же направлению и потому определяются своими алгебраическими величинами (рис. 1).

Кроме того, мы примем, что имеет место принцип линейного наложения сил:

1° *При суммарном наложении двух систем сил соответствующие прогибы складываются.*

2° *При умножении величин всех сил на одно и то же вещественное число все прогибы умножаются на это число.*

Обозначим через a_{ik} коэффициент влияния точки (k) на точку (i), т. е. прогиб в точке (i) под действием единичной силы, приложенной в точке (k) $(i, k=1, 2, \ldots, n)$ (рис. 2). Тогда при совместном действии сил $F_1, F_2, \ldots, F_n$ прогибы $y_1, y_2, \ldots, y_n$ определятся по формулам

$$\sum_{k=1}^{n} a_{ik}F_k = y_i \qquad (i=1, 2, \ldots, n). \tag{23}$$

Сопоставляя (23) с исходной системой (1), мы задачу отыскания решения системы уравнений (1) можем интерпретировать так:

Даны прогибы $y_1, y_2, \ldots, y_n$. *Ищутся соответствующие силы* $F_1, F_2, \ldots, F_n$.

Обозначим через S_p статическую систему, получающуюся из S введением p неподвижных шарнирных опор в точках (1), (2) ..., (p) $(p \leqslant n)$. Коэффициенты влияния для оставшихся подвижных точек $(p+1), \ldots, (n)$ системы S_p обозначим через

$$a_{ik}^{(p)} \qquad (i, k=p+1, \ldots, n)$$

(см. рис. 3 для $p=1$).

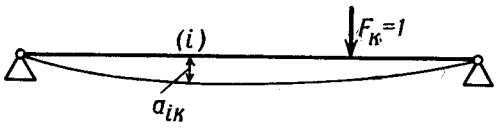


Рис. 2.

Коэффициент $a_{ik}^{(p)}$ можно рассматривать как прогиб в точке (i) системы S при действии единичной силы в точке (k) и сил реакций $R_1, R_2, \ldots, R_p$ в закрепленных точках (1), (2), ..., (p). Поэтому

$$a_{ik}^{(p)} = R_1a_{i1} + \ldots + R_pa_{ip} + a_{ik}. \tag{24}$$

С другой стороны, при этих же силах прогибы системы S в точках (1), (2), ..., (p) равны нулю:

$$\left.\begin{aligned} R_1a_{11} + \ldots + R_pa_{1p} + a_{1k} &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ R_1a_{p1} + \ldots + R_pa_{pp} + a_{pk} &= 0. \end{aligned}\right\} \tag{25}$$

Если

$$A\begin{pmatrix} 1 & 2 \ldots p \\ 1 & 2 \ldots p \end{pmatrix} \neq 0,$$

то мы можем из (25) определить $R_1, R_2, \ldots, R_p$ и полученные выражения подставить в (24). Это исключение $R_1, R_2, \ldots, R_p$ можно сделать и так. К системе равенств (25) прибавим равенство (24), записанное в виде

$$R_1a_{i1} + \ldots + R_pa_{ip} + a_{ik} - a_{ik}^{(p)} = 0. \tag{24'}$$

Рассматривая (25) и (24') как систему $p+1$ однородных уравнений, имеющую ненулевое решение $R_1, R_2, \ldots, R_p, R_{p+1}=1$, получаем, что определитель этой системы равен нулю:

$$\begin{vmatrix} a_{11} \ldots a_{1p} & a_{1k} \\ \cdots\cdots\cdots & \cdots \\ a_{p1} \ldots a_{pp} & a_{pk} \\ a_{i1} \ldots a_{ip} & a_{ik} - a_{ik}^{(p)} \end{vmatrix} = 0.$$

Отсюда

$$a_{ik}^{(p)} = \frac{A\begin{pmatrix} 1 & 2 \ldots p & i \\ 1 & 2 \ldots p & k \end{pmatrix}}{A\begin{pmatrix} 1 & 2 \ldots p \\ 1 & 2 \ldots p \end{pmatrix}} \qquad (i, k=p+1, \ldots, n). \tag{26}$$

По этим формулам коэффициенты влияния «опорной» системы S_p выражаются через коэффициенты влияния исходной системы S.

Но формулы (26) совпадают с формулами (13) предыдущего параграфа. Поэтому *для любого* p $(\leqslant n-1)$ *коэффициенты* $a_{ik}^{(p)}$ $(i, k=p+1, \ldots, n)$ *в алгоритме Гаусса являются коэффициентами влияния опорной системы* S_p.

В справедливости этого основного положения можно убедиться из чисто механических соображений, не опираясь на алгебраический вывод формул (13). Для этого рассмотрим сначала частный случай одной опоры: $p=1$ (рис. 3). В этом случае коэффициенты влияния системы S_1 определятся по формулам [полагаем $p=1$ в (26)]:

$$a_{ik}^{(1)}=\frac{A\begin{pmatrix}1 & i\\ 1 & k\end{pmatrix}}{A\begin{pmatrix}1\\ 1\end{pmatrix}}=a_{ik}-\frac{a_{i1}}{a_{11}}a_{1k}\qquad (i,\ k=1,\ 2,\ \ldots,\ n).$$

Эти формулы совпадают с формулами (3′).

Таким образом, если коэффициенты a_{ik} $(i,\ k=1,\ 2,\ \ldots,\ n)$ в системе уравнений (1) являются коэффициентами влияния статической системы S, то коэффициенты $a_{ik}^{(1)}$ $(i,\ k=2,\ \ldots,\ n)$ в алгоритме Гаусса являются коэффициентами влияния системы S_1. Применяя эти же соображения к системе S_1 и вводя в ней вторую опору в точке (2), получим, что коэффициенты $a_{ik}^{(2)}$ $(i,\ k=3,\ \ldots,\ n)$ в системе уравнений (4) являются коэффициентами влияния опорной системы S_2, и вообще для любого $p\ (\leqslant n-1)$ коэффициенты $a_{ik}^{(p)}$ $(i,\ k=p+1,\ \ldots,\ n)$ в алгоритме Гаусса являются коэффициентами влияния опорной системы S_p.

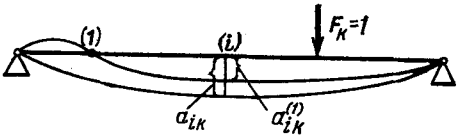


Рис. 3.

Из механических соображений очевидно, что *последовательное введение p опор равносильно одновременному введению этих опор.*

Замечание. Обращаем внимание на то, что при механической интерпретации алгоритма исключения не было необходимости предполагать, что точки, в которых рассматриваются прогибы, совпадают с точками приложения сил $F_1, F_2, \ldots, F_n$. Можно считать, что $y_1, y_2, \ldots, y_n$ — прогибы точек (1), (2), ..., (n), а силы $F_1, F_2, \ldots, F_n$ приложены в точках (1′), (2′), ..., (n′). Тогда a_{ik} — коэффициент влияния точки (k') на точку (i). В этом случае вместо опоры в точке (j) следует рассматривать обобщенную опору в точках (j), (j'), при которой прогиб в точке (j) поддерживается все время равным нулю за счет надлежащим образом выбранной вспомогательной силы R_j в точке (j'). Условие возможности введения p обобщенных опор в точках (1), (1′); (2), (2′); ...; (p), (p'), т. е. возможность удовлетворить условиям $y_1=0$, $y_2=0$, ..., $y_p=0$ при любых $F_{p+1}, \ldots, F_n$ за счет надлежащих $R_1=F_1, \ldots, R_p=F_p$, выражается неравенством

$$A\begin{pmatrix}1 & 2 & \ldots & p\\ 1 & 2 & \ldots & p\end{pmatrix}\neq 0.$$

§ 3. Детерминантное тождество Сильвестра

В § 1 путем сопоставления матриц A и G_p мы пришли к равенствам (10) и (11).

Эти равенства позволяют сразу получить важное детерминантное тождество Сильвестра. Действительно, из (10) и (11) находим:

$$|A|=A\begin{pmatrix}1 & 2 & \ldots & n\\ 1 & 2 & \ldots & n\end{pmatrix}=A\begin{pmatrix}1 & 2 & \ldots & p\\ 1 & 2 & \ldots & p\end{pmatrix}\begin{vmatrix}a_{p+1,\,p+1}^{(p)} & \ldots & a_{p+1,\,n}^{(p)}\\ \ldots & \ldots & \ldots\\ a_{n,\,p+1}^{(p)} & \ldots & a_{nn}^{(p)}\end{vmatrix}. \qquad (27)$$

Введем в рассмотрение окаймляющие минор $A\begin{pmatrix}1 & 2 & \ldots & p\\ 1 & 2 & \ldots & p\end{pmatrix}$ определители

$$b_{ik}=A\begin{pmatrix}1 & 2 & \ldots & p & i\\ 1 & 2 & \ldots & p & k\end{pmatrix}\qquad (i,\ k=p+1,\ \ldots,\ n).$$

Матрицу, составленную из этих определителей, обозначим через

$$B=\|\,b_{ik}\,\|_{p+1}^{n}.$$

Тогда согласно формулам (13)

$$\begin{vmatrix} a^{(p)}_{p+1,\,p+1} & \dots & a^{(p)}_{p+1,\,n} \\ \dots\dots\dots\dots \\ a^{(p)}_{n,\,p+1} & \dots & a^{(p)}_{nn} \end{vmatrix} = \frac{\begin{vmatrix} b_{p+1,\,p+1} & \dots & b_{p+1,\,n} \\ \dots\dots\dots\dots \\ b_{n,\,p+1} & \dots & b_{nn} \end{vmatrix}}{\left[A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}\right]^{n-p}} = \frac{|B|}{\left[A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}\right]^{n-p}}.$$

Поэтому равенство (27) может быть записано так:

$$|B| = \left[A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}\right]^{n-p-1} |A|. \tag{28}$$

Это и есть детерминантное тождество Сильвестра. Оно выражает определитель $|B|$, составленный из окаймляющих определителей, через исходный определитель и окаймляемый минор.

Равенство (28) было нами установлено для матриц $A = \|a_{ik}\|_1^n$, элементы которых удовлетворяют неравенствам

$$A\begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \qquad (j = 1, 2, \dots, p). \tag{29}$$

Однако из «соображений непрерывности» следует, что эти ограничения можно отбросить и что тождество Сильвестра справедливо для любой матрицы $A = \|a_{ik}\|_1^n$. В самом деле, пусть неравенства (29) не выполняются. Введем матрицу

$$A_\varepsilon = A + \varepsilon E.$$

Очевидно, $\lim\limits_{\varepsilon \to 0} A_\varepsilon = A$. С другой стороны, миноры

$$A_\varepsilon\begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} = \varepsilon^j + \dots \qquad (j = 1, 2, \dots, p)$$

представляют собой p не равных тождественно нулю многочленов относительно ε. Поэтому можно выбрать такую последовательность $\varepsilon_m \to 0$, что

$$A_{\varepsilon_m}\begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \qquad (j = 1, 2, \dots, p;\ m = 1, 2, \dots).$$

Для матрицы A_{ε_m} мы можем написать тождество (28). Переходя в обеих частях этого тождества к пределу при $m \to \infty$, мы получим тождество Сильвестра для предельной матрицы $A = \lim\limits_{m \to \infty} A_{\varepsilon_m}$[1]).

Если мы тождество (28) применим к определителю

$$A\begin{pmatrix} 1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q \end{pmatrix} \quad \left(p < \begin{matrix} i_1 < i_2 < \dots < i_q \\ k_1 < k_2 < \dots < k_q \end{matrix} \leqslant n\right),$$

то мы получим удобный для применений вид тождества Сильвестра:

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ k_1 & k_2 & \dots & k_q \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}\right]^{q-1} A\begin{pmatrix} 1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q \end{pmatrix}. \tag{30}$$

[1]) Под *пределом* (при $p \to \infty$) *последовательности матриц* $B_p = \|b_{ik}^{(p)}\|$ понимают матрицу $B = \|b_{ik}\|_1^n$, где $b_{ik} = \lim\limits_{p \to \infty} b_{ik}^{(p)}$ $(i, k = 1, 2, \dots, n)$.

§ 4. Разложение квадратной матрицы на треугольные множители

1. Пусть дана матрица $A = \|a_{ik}\|_1^n$ ранга r. Введем следующие обозначения для последовательных главных миноров этой матрицы:

$$D_k = A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \qquad (k = 1, 2, \dots, n).$$

Допустим, что имеют место условия выполнимости алгоритма Гаусса:

$$D_k \neq 0 \qquad (k = 1, 2, \dots, r).$$

Обозначим через G матрицу коэффициентов системы уравнений (18), к которой приводится система уравнений

$$\sum_{k=1}^{n} a_{ik}x_k = y_i \qquad (i = 1, 2, \dots, n)$$

методом исключения Гаусса. Матрица G имеет верхнюю треугольную форму, причем элементы ее первых r строк определяются формулами (13), а элементы последних $n-r$ строк все равны нулю[1]):

$$G = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{Vmatrix}.$$

Переход от матрицы A к матрице G совершался при помощи некоторого числа N операций следующего типа: к i-й строке матрицы прибавлялась j-я $(j < i)$ строка, предварительно помноженная на некоторое число α. Такая операция равносильна умножению преобразуемой матрицы слева на матрицу

$$\begin{Vmatrix} & & & (j) & & (i) & \\ 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & & \dots & & \dots & & \dots \\ & & & 1 & & & \\ \dots & & \dots & & \dots & & \dots \\ 0 & \dots & \alpha & \dots & 1 & \dots & 0 \\ \dots & & \dots & & \dots & & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{Vmatrix}. \qquad (31)$$

В этой матрице на главной диагонали стоят единицы, а все остальные элементы, за исключением элемента α, равны нулю.

[1]) Матрица G совпадает с матрицей G_r (стр. 45).

Таким образом,

$$G = W_N \dots W_2 W_1 A,$$

где каждая из матриц $W_1, W_2, \dots, W_N$ имеет вид (31) и, следовательно, является нижней треугольной матрицей с диагональными элементами, равными 1.

Пусть

$$W = W_N \dots W_2 W_1. \tag{32}$$

Тогда

$$G = WA. \tag{33}$$

Матрицу W будем называть *преобразующей* матрицей для матрицы A в методе исключения Гаусса. Обе матрицы, G и W, однозначно определяются заданием матрицы A. Из (32) следует, что W — нижняя треугольная матрица с диагональными элементами, равными 1 (см. стр. 28).

Поскольку W — неособенная матрица, то из (33) находим:

$$A = W^{-1}G. \tag{33'}$$

Мы представили матрицу A в виде произведения нижней треугольной матрицы W^{-1} на верхнюю треугольную матрицу G. Вопрос о разложении матрицы A на множители такого типа полностью выясняется следующей теоремой:

Теорема 1. *Всякую матрицу* $A = \|a_{ik}\|_1^n$ *ранга* r, *у которой первые* r *последовательных главных миноров отличны от нуля,*

$$D_k = A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \qquad (k = 1, 2, \dots, r), \tag{34}$$

можно представить в виде произведения нижней треугольной матрицы B *на верхнюю треугольную матрицу* C

$$A = BC = \begin{Vmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{Vmatrix} \begin{Vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{Vmatrix}. \tag{35}$$

При этом

$$b_{11}c_{11} = D_1, \quad b_{22}c_{22} = \frac{D_2}{D_1}, \quad \dots, \quad b_{rr}c_{rr} = \frac{D_r}{D_{r-1}}. \tag{36}$$

Первым r *диагональным элементам матриц* B *и* C *можно дать произвольные значения, удовлетворяющие условиям* (36).

Задание первых r *диагональных элементов матриц* B *и* C *определяет однозначно элементы первых* r *столбцов матрицы* B *и первых* r *строк матрицы* C. *Для этих элементов имеют место формулы*

$$b_{gk} = b_{kk}\frac{A\begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}}, \qquad c_{kg} = c_{kk}\frac{A\begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & g \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \tag{37}$$

$$(g = k, k+1, \dots, n; \ k = 1, 2, \dots, r).$$

В случае $r < n$ ($|A| = 0$) *в последних* $n - r$ *столбцах матрицы* B *можно все элементы положить равными нулю, а в последних* $n - r$ *строках матрицы* C *всем элементам дать произвольные значения, либо*

наоборот, последние $n-r$ строк матрицы C заполнить нулями, а последние $n-r$ столбцов матрицы B взять произвольными.

Доказательство. Возможность представления матрицы, удовлетворяющей условию (34), в виде произведения (35) была доказана выше [см. (33′)].

Пусть теперь B и C — произвольные нижняя и верхняя треугольные матрицы, произведение которых равно A. Пользуясь формулой для миноров произведения двух матриц, найдем:

$$A\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ 1 & 2 & \dots & k-1 & k\end{pmatrix}=\sum_{\alpha_1<\alpha_2<\dots<\alpha_k} B\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ \alpha_1 & \alpha_2 & \dots & \alpha_{k-1} & \alpha_k\end{pmatrix} C\begin{pmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_k\\ 1 & 2 & \dots & k\end{pmatrix} \tag{38}$$

$$(g=k, k+1, \dots, n, k=1, 2, \dots, r).$$

Поскольку C — верхняя треугольная матрица, то первые k столбцов матрицы C содержат только один отличный от нуля минор k-го порядка $C\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}$. Поэтому равенство (38) может быть записано так:

$$A\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ 1 & 2 & \dots & k-1 & k\end{pmatrix}=B\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ 1 & 2 & \dots & k-1 & k\end{pmatrix}C\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}=$$

$$=b_{11}b_{22}\dots b_{k-1,\,k-1}b_{gk}c_{11}c_{22}\dots c_{kk} \tag{39}$$

$$(g=k, k+1, \dots, n;\ k=1, 2, \dots, r).$$

Положим сначала здесь $g=k$. Тогда получим:

$$b_{11}b_{22}\dots b_{kk}c_{11}c_{22}\dots c_{kk}=D_k \qquad (k=1, 2, \dots, r), \tag{40}$$

откуда уже вытекают соотношения (36).

Не нарушая равенства (35), мы можем в нем умножить матрицу B справа на произвольную неособенную диагональную матрицу $M=\|\mu_i\delta_{ik}\|_1^n$, одновременно умножая матрицу C слева на $M^{-1}=\|\mu_i^{-1}\delta_{ik}\|_1^n$. Это равносильно умножению столбцов матрицы B соответственно на $\mu_1, \mu_2, \dots, \mu_n$ и строк матрицы C на $\mu_1^{-1}, \mu_2^{-1}, \dots, \mu_n^{-1}$. Поэтому диагональным элементам $b_{11}, \dots, b_{rr}, c_{11}, \dots, c_{rr}$ можно придать любые значения, удовлетворяющие условиям (36).

Далее, из (39) и (40) находим:

$$b_{gk}=b_{kk}\frac{A\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ 1 & 2 & \dots & k-1 & k\end{pmatrix}}{A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}} \qquad (g=k, k+1, \dots, n;\ k=1, 2, \dots, r),$$

т. е. первые формулы (37). Совершенно аналогично устанавливаются вторые формулы (37) для элементов матрицы C.

Обратим внимание на то, что при перемножении матриц B и C элементы b_{kg} последних $n-r$ столбцов матрицы B и элементы c_{gk} последних $n-r$ строк матрицы C перемножаются только между собой. Мы видели, что все элементы последних $n-r$ строк матрицы C можно выбрать равными нулю[1]). Тогда элементы последних $n-r$ столбцов матрицы B можно

[1]) Это следует из представления (33′). При этом диагональным элементам $b_{11}, \dots, b_{rr}, c_{11}, \dots, c_{rr}$, как было уже показано, можно придать любые значения, удовлетворяющие условиям (36) за счет надлежащих множителей $\mu_1, \mu_2, \dots, \mu_r$.

выбрать произвольными. Ясно, что произведение матриц B и C не изменится, если мы последние $n-r$ столбцов матрицы B возьмем нулевыми, а элементы последних $n-r$ строк матрицы C произвольными.

Теорема доказана.

Из доказанной теоремы вытекает ряд интересных следствий.

Следствие 1. *Элементы первых r столбцов матрицы B и первых r строк матрицы C связаны с элементами матрицы A рекуррентными соотношениями:*

$$\left.\begin{aligned} b_{ik} &= \frac{a_{ik} - \sum_{j=1}^{k-1} b_{ij}c_{jk}}{c_{kk}} \qquad (i \geqslant k;\ i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, r), \\ c_{ik} &= \frac{a_{ik} - \sum_{j=1}^{i-1} b_{ij}c_{jk}}{b_{ii}} \qquad (i \leqslant k;\ i = 1, 2, \ldots, r;\ k = 1, 2, \ldots, n). \end{aligned}\right\} \tag{41}$$

Соотношения (41) непосредственно следуют из матричного равенства (35) ими удобно пользоваться для фактического вычисления элементов матриц B и C.

Следствие 2. *Если $A = \| a_{ik} \|_1^n$ — неособенная матрица ($r = n$), удовлетворяющая условию* (34), *то в представлении* (35) *матрицы B и C определяются однозначно, как только диагональные элементы этих матриц выбраны в соответствии с условиями* (36).

Следствие 3. *Если $S = \| s_{ik} \|_1^n$ — симметрическая матрица ранга r и*

$$D_k = S\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix} \neq 0 \qquad (k = 1, 2, \ldots, r),$$

то

$$S = BB',$$

где $B = \| b_{ik} \|_1^n$ — нижняя треугольная матрица, в которой

$$b_{gk} = \begin{cases} \dfrac{1}{\sqrt{D_k D_{k-1}}} A\begin{pmatrix} 1 & 2 & \ldots & k-1 & g \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} & (g = k, k+1, \ldots, n;\ k = 1, 2, \ldots, r), \\ 0 & (g = k, k+1, \ldots, n;\ k = r+1, \ldots, n). \end{cases} \tag{42}$$

2. Пусть в представлении (35) у матрицы B элементы последних $n-r$ столбцов равны нулю. Тогда можно положить:

$$B = F \cdot \begin{Vmatrix} b_{11} & & & & 0 \\ & \ddots & & & \\ & & b_{rr} & & \\ & & & 0 & \\ 0 & & & & 0 \end{Vmatrix}, \qquad C = \begin{Vmatrix} c_{11} & & & & 0 \\ & \ddots & & & \\ & & c_{rr} & & \\ & & & 0 & \\ 0 & & & & 0 \end{Vmatrix} \cdot L, \tag{43}$$

где F — нижняя, а L — верхняя треугольная матрица; при этом первые r диагональных элементов у матриц F и L равны 1, а элементы последних $n-r$ столбцов матрицы F и последних $n-r$ строк матрицы L выбраны совершенно произвольно. Подставляя в (35) выражения (43) для B и C и используя равенства (36), придем к следующей теореме:

Теорема 2. *Всякая матрица* $A=\|a_{ik}\|_1^n$ *ранга* r, *у которой*

$$D_k=A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}\neq 0 \qquad (k=1,\ 2,\ \dots,\ r),$$

представима в виде произведения нижней треугольной матрицы F, *диагональной* D *и верхней треугольной* L:

$$A=FDL=\begin{Vmatrix}1 & 0 & \dots & 0\\ f_{21} & 1 & \dots & 0\\ \dots & \dots & \dots & \dots\\ f_{n1} & f_{n2} & \dots & 1\end{Vmatrix}\begin{Vmatrix}D_1 & & & & \\ & \frac{D_2}{D_1} & & & \\ & & \ddots & & \\ & & & \frac{D_r}{D_{r-1}} & \\ & & & & 0\ddots 0\end{Vmatrix}\begin{Vmatrix}1 & l_{12} & \dots & l_{1n}\\ 0 & 1 & \dots & l_{2n}\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & 1\end{Vmatrix}, \tag{44}$$

где

$$f_{gk}=\frac{A\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ 1 & 2 & \dots & k-1 & k\end{pmatrix}}{A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}},\quad l_{kg}=\frac{A\begin{pmatrix}1 & 2 & \dots & k-1 & k\\ 1 & 2 & \dots & k-1 & g\end{pmatrix}}{A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}} \tag{45}$$

$$(g=k+1,\ \dots,\ n;\ k=1,\ 2,\ \dots,\ r),$$

а f_{gk}, l_{kg} *произвольны при* $g=k+1,\ \dots,\ n;\ k=r+1,\ \dots,\ n$.

3. Метод исключения Гаусса, будучи применен к матрице $A=\|a_{ik}\|_1^n$ ранга r, для которой $D_k\neq 0$ $(k=1,\ 2,\ \dots,\ r)$, дает нам две матрицы: нижнюю треугольную матрицу W с диагональными элементами 1 и верхнюю треугольную матрицу G, у которой первые r диагональных элементов равны $D_1,\ \frac{D_2}{D_1},\ \dots,\ \frac{D_r}{D_{r-1}}$, а последние $n-r$ строк заполнены нулями. G — гауссова форма матрицы A, W — преобразующая матрица.

Для конкретного вычисления элементов матрицы W можно рекомендовать следующий прием.

Мы получим матрицу W, если к единичной матрице E применим все те преобразования (задаваемые матрицами $W_1,\ \dots,\ W_N$), которые мы в алгоритме Гаусса делали над матрицей A (в этом случае вместо произведения WA, равного G, мы будем иметь произведение WE, равное W). Поэтому к матрице A приписываем справа единичную матрицу E:

$$\begin{Vmatrix}a_{11} & \dots & a_{1n} & 1 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots\\ a_{n1} & \dots & a_{nn} & 0 & \dots & 1\end{Vmatrix}. \tag{46}$$

Применяя к этой прямоугольной матрице все преобразования алгоритма Гаусса, получим прямоугольную матрицу, состоящую из двух квадратных матриц G и W:

$$(G,\ W).$$

Таким образом, *применение алгоритма Гаусса к матрице* (46) *дает одновременно и матрицу* G *и матрицу* W.

Если A — неособенная матрица, т. е. $|A| \neq 0$, то и $|G| \neq 0$. В этом случае из (33) следует $A^{-1} = G^{-1}W$. Поскольку матрицы G и W определены при помощи алгоритма Гаусса, то нахождение обратной матрицы A^{-1} сводится к определению G^{-1} и умножению G^{-1} на W.

Хотя нахождение обратной матрицы G^{-1}, после того как определена матрица G, не представляет затруднений, поскольку G — треугольная матрица, тем не менее можно избежать этой операции. Для этого наряду с матрицами G и W введем аналогичные матрицы G_1 и W_1 для транспонированной матрицы A'. Тогда $A' = W_1^{-1}G_1$, т. е.

$$A = G_1'W_1'^{-1}. \tag{47}$$

Сопоставим между собой равенства (33′) и (44):

$$A = W^{-1}G, \quad A = FDL.$$

Эти равенства могут быть рассматриваемы как два различных разложения вида (35); при этом мы произведение DL рассматриваем как второй множитель C. Поскольку первые r диагональных элементов у первых множителей одинаковы (они равны 1), то первые r столбцов у них совпадают. Тогда, поскольку последние $n-r$ столбцов матрицы F могут быть выбраны произвольными, то выберем их так, чтобы

$$F = W^{-1}. \tag{48}$$

С другой стороны, сопоставление равенств (47) и (44):

$$A = G_1'W_1'^{-1}, \quad A = FDL$$

показывает, что можно так подобрать произвольные элементы в L, чтобы

$$L = W_1'^{-1}. \tag{49}$$

Подставляя в (44) вместо F и L их выражения из (48) и (49), получим:

$$A = W^{-1}DW_1'^{-1}. \tag{50}$$

Сопоставляя это равенство с равенствами (33′) и (47), мы найдем:

$$G = DW_1'^{-1}, \quad G_1' = W^{-1}D. \tag{51}$$

Введем в рассмотрение диагональную матрицу

$$\widehat{D} = \left\{\frac{1}{D_1}, \frac{D_1}{D_2}, \ldots, \frac{D_{r-1}}{D_r}, 0, \ldots, 0\right\}. \tag{52}$$

Тогда, поскольку

$$D = D\widehat{D}D,$$

то из (50) и (51) следует:

$$A = G_1'\widehat{D}G. \tag{53}$$

Формула (53) показывает, что разложение матрицы A на треугольные множители может быть получено применением алгоритма Гаусса к матрицам A и A'.

Пусть теперь A — неособенная матрица ($r = n$). Тогда $|D| \neq 0$, $\widehat{D} = D^{-1}$. Поэтому из (50) следует:

$$A^{-1} = W_1'\widehat{D}W. \tag{54}$$

Эта формула дает возможность эффективного вычисления обратной матрицы A^{-1} путем применения алгоритма Гаусса к прямоугольным матрицам

$$(A,\ E)\ (A',\ E).$$

В частном случае, когда вместо матрицы A возьмем симметрическую матрицу S, матрица G_1 совпадает с G, а матрица W_1 — с матрицей W, и потому формулы (53) и (54) принимают вид

$$S = G'\widehat{D}G, \tag{55}$$

$$S^{-1} = W'\widehat{D}W. \tag{56}$$

§ 5. Разбиение матрицы на блоки. Техника оперирования с блочными матрицами. Обобщенный алгоритм Гаусса

Часто приходится пользоваться матрицами, разбитыми на прямоугольные части — «клетки» или «блоки». Рассмотрению таких «блочных» матриц мы посвящаем настоящий параграф.

1. Пусть дана прямоугольная матрица

$$A = \| a_{ik} \| \qquad (i = 1,\ 2,\ \ldots,\ m;\ k = 1,\ 2,\ \ldots,\ n). \tag{57}$$

При помощи горизонтальных и вертикальных линий рассечем матрицу A на прямоугольные блоки:

$$A = \begin{pmatrix} \overbrace{A_{11}}^{n_1} & \overbrace{A_{12}}^{n_2} & \ldots & \overbrace{A_{1t}}^{n_t} \\ A_{21} & A_{22} & \ldots & A_{2t} \\ \ldots & \ldots & \ldots & \ldots \\ A_{s1} & A_{s2} & \ldots & A_{st} \end{pmatrix} \begin{matrix} \} m_1 \\ \} m_2 \\ \vdots \\ \} m_s \end{matrix} \tag{58}$$

Про матрицу (58) будем говорить, что она разбита на st блоков $A_{\alpha\beta}$ размером $m_\alpha \times n_\beta$ ($\alpha = 1,\ 2,\ \ldots,\ s$; $\beta = 1,\ 2,\ \ldots,\ t$) или что она представлена в виде блочной матрицы. Вместо (58) будем сокращенно писать:

$$A = (A_{\alpha\beta}) \qquad (\alpha = 1,\ 2,\ \ldots,\ s;\ \beta = 1,\ 2,\ \ldots,\ t). \tag{59}$$

В случае $s = t$ будем пользоваться и такой записью:

$$A = (A_{\alpha\beta})_1^s. \tag{60}$$

Действия над блочными матрицами производятся по тем же формальным правилам, как и в случае, когда вместо блоков имеем числовые элементы. Пусть, например, даны две прямоугольные матрицы одинаковых размеров и с одинаковым разбиением на блоки:

$$A = (A_{\alpha\beta}),\ B = (B_{\alpha\beta}) \qquad (\alpha = 1,\ 2,\ \ldots,\ s;\ \beta = 1,\ 2,\ \ldots,\ t). \tag{61}$$

Легко усмотреть, что

$$A + B = (A_{\alpha\beta} + B_{\alpha\beta}) \qquad (\alpha = 1,\ 2,\ \ldots,\ s;\ \beta = 1,\ 2,\ \ldots,\ t). \tag{62}$$

Подробнее остановимся на умножении блочных матриц. Известно (см. гл. I, стр. 17), что при умножении двух прямоугольных матриц A и B длина строк в первом сомножителе A должна совпадать с высотой столбцов во втором сомножителе B. Для возможности «блочного»

умножения этих матриц мы дополнительно потребуем, чтобы при разбиении на блоки все горизонтальные размеры в первом сомножителе совпадали с соответствующими вертикальными размерами во втором:

$$A=\begin{pmatrix} \overbrace{A_{11}}^{n_1} & \overbrace{A_{12}}^{n_2} & \dots & \overbrace{A_{1t}}^{n_t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{st} \end{pmatrix}\begin{matrix} \} m_1 \\ \} m_2 \\ \vdots \\ \} m_s \end{matrix}, \quad B=\begin{pmatrix} \overbrace{B_{11}}^{p_1} & \overbrace{B_{12}}^{p_2} & \dots & \overbrace{B_{1u}}^{p_u} \\ B_{21} & B_{22} & \dots & B_{2u} \\ \dots & \dots & \dots & \dots \\ B_{t1} & B_{t2} & \dots & B_{tu} \end{pmatrix}\begin{matrix} \} n_1 \\ \} n_2 \\ \vdots \\ \} n_t \end{matrix}. \tag{63}$$

Тогда легко проверить, что

$$AB=C=(C_{\alpha\beta}), \quad \text{где} \quad C_{\alpha\beta}=\sum_{\delta=1}^{t} A_{\alpha\delta}B_{\delta\beta} \qquad \begin{pmatrix}\alpha=1,\ 2,\ \dots,\ s\\ \beta=1,\ 2,\ \dots,\ u\end{pmatrix}. \tag{64}$$

Отдельно отметим тот частный случай, когда одним из сомножителей является *квазидиагональная* матрица. Пусть A — квазидиагональная матрица, т. е. $s=t$ и $A_{\alpha\beta}=0$ при $\alpha\neq\beta$. В этом случае формула (64) нам дает:

$$C_{\alpha\beta}=A_{\alpha\alpha}B_{\alpha\beta} \qquad (\alpha=1,\ 2,\ \dots,\ s;\ \beta=1,\ 2,\ \dots,\ u). \tag{65}$$

При умножении блочной матрицы слева на квазидиагональную матрицу строки блочной матрицы умножаются слева на соответствующие диагональные клетки квазидиагональной матрицы.

Пусть теперь B — квазидиагональная матрица, т. е. $t=u$ и $B_{\alpha\beta}=0$ при $\alpha\neq\beta$. Тогда из (64) получаем:

$$C_{\alpha\beta}=A_{\alpha\beta}B_{\beta\beta} \qquad (\alpha=1,\ 2,\ \dots,\ s;\ \beta=1,\ 2,\ \dots,\ u). \tag{66}$$

При умножении блочной матрицы справа на квазидиагональную все столбцы блочной матрицы умножаются справа на соответствующие диагональные клетки квазидиагональной матрицы.

Заметим, что умножение квадратных блочных матриц одного и того же порядка всегда выполнимо, когда сомножители разбиты на одинаковые квадратные схемы блоков и в каждом из сомножителей на диагональных местах стоят квадратные матрицы.

Блочная матрица (58) называется *верхней (нижней) квазитреугольной*, если $s=t$ и все $A_{\alpha\beta}=0$ при $\alpha>\beta$ (соответственно все $A_{\alpha\beta}=0$ при $\alpha<\beta$). Частным случаем квазитреугольной матрицы является квазидиагональная матрица.

Из формулы (64) легко усмотреть, что

Произведение двух верхних (нижних) квазитреугольных матриц является снова верхней (нижней) квазитреугольной матрицей[1]; *при этом диагональные блоки произведения получаются путем перемножения соответствующих диагональных блоков сомножителей.*

Действительно, полагая в (64) $s=t$ и

$$A_{\alpha\beta}=0, \quad B_{\alpha\beta}=0 \text{ при } \alpha<\beta,$$

найдем:

$$\left.\begin{aligned} &C_{\alpha\beta}=0 \text{ при } \alpha<\beta \\ \text{и} \quad &C_{\alpha\alpha}=A_{\alpha\alpha}B_{\alpha\alpha} \end{aligned}\right\} \qquad (\alpha,\ \beta=1,\ 2,\ \dots,\ s).$$

[1] При этом предполагается, что блочное умножение выполнимо.

Аналогично разбирается случай нижних квазитреугольных матриц.

Отметим правило вычисления определителя квазитреугольной матрицы. Это правило можно получить, исходя из разложения Лапласа.

Если A — квазитреугольная (в частности, квазидиагональная) матрица с квадратными диагональными блоками, то определитель этой матрицы равен произведению определителей диагональных блоков:

$$|A| = |A_{11}|\,|A_{22}| \dots |A_{ss}|. \tag{67}$$

2. Пусть дана блочная матрица

$$A = \begin{matrix} \overbrace{\phantom{A_{11}}}^{n_1} & \overbrace{\phantom{A_{12}}}^{n_2} & \dots & \overbrace{\phantom{A_{1t}}}^{n_t} \\ \end{matrix} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{st} \end{pmatrix} \begin{matrix} \} m_1 \\ \} m_2 \\ \vdots \\ \} m_s \end{matrix}. \tag{68}$$

Прибавим к α-й блочной строке β-ю строку, предварительно помноженную слева на прямоугольную матрицу X размером $m_\alpha \times m_\beta$. Получим блочную матрицу

$$B = \begin{pmatrix} A_{11} & \dots & A_{1t} \\ \dots & \dots & \dots \\ A_{\alpha 1} + XA_{\beta 1} & \dots & A_{\alpha t} + XA_{\beta t} \\ \dots & \dots & \dots \\ A_{\beta 1} & \dots & A_{\beta t} \\ \dots & \dots & \dots \\ A_{s1} & \dots & A_{st} \end{pmatrix}. \tag{69}$$

Введем вспомогательную квадратную матрицу V, представленную в виде следующей квадратной схемы блоков:

$$V = \begin{matrix} m_1 & \dots & m_\alpha & \dots & m_\beta & \dots & m_s \\ \end{matrix} \begin{pmatrix} E & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & E & \dots & X & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & E & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & E \end{pmatrix} \begin{matrix} \} m_1 \\ \vdots \\ \} m_\alpha \\ \vdots \\ \} m_\beta \\ \vdots \\ \} m_s \end{matrix} \tag{70}$$

В диагональных клетках матрицы V стоят единичные матрицы, порядки которых равны соответственно $m_1, m_2, \dots, m_s$; все недиагональные блоки матрицы V равны нулю, за исключением блока X, стоящего на пересечении α-й блочной строки с β-м блочным столбцом.

Нетрудно видеть, что

$$VA = B. \tag{71}$$

Отсюда, поскольку V — неособенная матрица, для рангов матриц A и B имеем [1]):

$$r_A = r_B. \tag{72}$$

[1]) См. стр. 22.

В частном случае, когда A — квадратная матрица, из (71) имеем:

$$|V|\,|A|=|B|. \tag{73}$$

Но определитель квазитреугольной матрицы V равен 1:

$$|V|=1. \tag{74}$$

Следовательно,

$$|A|=|B|. \tag{75}$$

К таким же выводам можно прийти, если к какому-либо столбцу матрицы (67) прибавлять другой столбец, предварительно помноженный справа на прямоугольную матрицу X надлежащих размеров.

Полученные результаты могут быть сформулированы в виде следующей теоремы.

Теорема 3. *Если в блочной матрице A к α-й блочной строке (столбцу) прибавить β-ю блочную строку (столбец), предварительно помноженную слева (справа) на прямоугольную матрицу X соответствующих размеров, то при этом преобразовании не изменятся ранг матрицы A, а также в случае, когда A — квадратная матрица, и определитель матрицы A.*

3. Рассмотрим теперь тот частный случай, когда в матрице A диагональный блок A_{11} — квадратная и притом неособенная матрица ($|A_{11}|\neq 0$).

К α-й строке матрицы A прибавим первую строку, помноженную слева на $-A_{\alpha 1}A_{11}^{-1}$ ($\alpha=2, \ldots, s$). Тогда получим матрицу

$$B_1=\begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1t} \\ 0 & A_{22}^{(1)} & \ldots & A_{2t}^{(1)} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & A_{s2}^{(1)} & \ldots & A_{st}^{(1)} \end{pmatrix}, \tag{76}$$

где

$$A_{\alpha\beta}^{(1)}=-A_{\alpha 1}A_{11}^{-1}A_{1\beta}+A_{\alpha\beta} \qquad (\alpha=2, \ldots, s;\ \beta=2, \ldots, t). \tag{77}$$

Если $A_{22}^{(1)}$ — квадратная неособенная матрица, то этот процесс можно продолжить. Мы приходим, таким образом, к обобщенному алгоритму Гаусса.

Пусть A — квадратная матрица. Тогда

$$|A|=|B_1|=|A_{11}|\begin{vmatrix} A_{22}^{(1)} & \ldots & A_{2t}^{(1)} \\ \ldots & \ldots & \ldots \\ A_{s2}^{(1)} & \ldots & A_{st}^{(1)} \end{vmatrix}. \tag{78}$$

Формула (78) сводит вычисление определителя $|A|$, состоящего из st блоков, к вычислению определителя меньшего порядка, состоящего из $(s-1)(t-1)$ блоков[1]).

Рассмотрим определитель Δ, разбитый на четыре блока:

$$\Delta=\begin{vmatrix} A & B \\ C & D \end{vmatrix}, \tag{79}$$

где A и D — квадратные матрицы.

[1]) Если $A_{22}^{(1)}$ — квадратная матрица и $|A_{22}^{(1)}|\neq 0$, то к полученному определителю из $(s-1)(t-1)$ блоков мы можем снова применить такое же преобразование и т. д.

Пусть $|A| \neq 0$. Тогда вычтем из второй строки первую, предварительно помноженную слева на $-CA^{-1}$. Получим:

$$\Delta = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} = |A|\,|D - CA^{-1}B|. \tag{I}$$

Точно так же, если $|D| \neq 0$, то мы вычтем в Δ из первой строки вторую, предварительно помноженную слева на $-BD^{-1}$. Получим:

$$\Delta = \begin{vmatrix} A - BD^{-1}C & 0 \\ C & D \end{vmatrix} = |A - BD^{-1}C|\,|D|. \tag{II}$$

В частном случае, когда все четыре матрицы A, B, C, D — квадратные (одного и того же порядка n), из (I) и (II) следуют формулы Шура, сводящие вычисление определителя $2n$-го порядка к вычислению определителя n-го порядка:

$$\Delta = |AD - ACA^{-1}B| \qquad (|A| \neq 0), \tag{Ia}$$

$$\Delta = |AD - BD^{-1}CD| \qquad (|D| \neq 0). \tag{IIa}$$

Если матрицы A и C перестановочны между собой, то из (Ia) следует:

$$\Delta = |AD - CB| \qquad \text{(при условии } AC = CA). \tag{Iб}$$

Точно так же, если C и D перестановочны между собой, то

$$\Delta = |AD - BC| \qquad \text{(при условии } CD = DC). \tag{IIб}$$

Формула (Iб) была получена в предположении $|A| \neq 0$, а формула (IIб) при условии $|D| \neq 0$. Однако, исходя из соображений непрерывности, эти ограничения можно отбросить.

Из формул (I)—(IIб) можно получить еще шесть формул, поменяв местами в правых частях A и D и одновременно B и C.

Пример.

$$\Delta = \begin{vmatrix} 1 & 0 & b_1 & b_2 \\ 0 & 1 & b_3 & b_4 \\ c_1 & c_2 & d_1 & d_2 \\ c_3 & c_4 & d_3 & d_4 \end{vmatrix}.$$

По формуле (Iб)

$$\Delta = \begin{vmatrix} d_1 - c_1b_1 - c_2b_3 & d_2 - c_1b_2 - c_2b_4 \\ d_3 - c_3b_1 - c_4b_3 & d_4 - c_3b_2 - c_4b_4 \end{vmatrix}.$$

4. Установим формулу Фробениуса[1]) для обращения блочной матрицы. Пусть неособенная квадратная матрица M ($|M| \neq 0$) разбита на блоки

$$M = \begin{matrix} \\ n \\ q \end{matrix} \begin{matrix} \overbrace{}^{n} & \overbrace{}^{q} \\ \end{matrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{80}$$

и пусть A — также неособенная квадратная матрица ($|A| \neq 0$). Требуется определить M^{-1}

[1]) См., например, [44].

Применим к матрице M обобщенный алгоритм Гаусса. Из второй блочной строки вычтем первую, предварительно помноженную слева на $-CA^{-1}$. Эта операция равносильна умножению матрицы M слева на матрицу[1] $\begin{pmatrix} E & 0 \\ X & E \end{pmatrix}$, где $X = -CA^{-1}$. Поэтому

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} M = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \tag{81}$$

Введем обозначение

$$H = D - CA^{-1}B$$

и заметим, что из равенства (81) следует:

$$|M| = |A|\,|H|. \tag{82}$$

Поэтому, поскольку $|M| \neq 0$, то и $|H| \neq 0$[2]. Переходя к обратным матрицам в равенстве (81), получим

$$M^{-1}\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1}. \tag{83}$$

Обратную матрицу для матрицы $\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}$ будем искать в виде $\begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix}$. Тогда из равенства

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}\begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$$

находим, что $U = -A^{-1}BH^{-1}$. Таким образом,

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix}. \tag{84}$$

Но тогда из равенства (83) находим

$$M^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1}\begin{pmatrix} E & 0 \\ -CA^{-1} & B \end{pmatrix} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix}\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}. \tag{85}$$

Выполняя умножение блочных матриц в правой части равенства (85), приходим к формуле Фробениуса

$$M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ -H^{-1}CA^{-1} & H^{-1} \end{pmatrix}, \tag{86}$$

где

$$H = D - CA^{-1}B. \tag{87}$$

Формула Фробениуса (86) сводит обращение матрицы порядка $n+q$ к обращению двух матриц порядка n и q и к операциям сложения и умножения матриц с размерами $n \times n$, $q \times q$, $n \times q$ и $q \times n$.

---

[1]) В первой блочной строке буква E обозначает единичную матрицу n-го порядка, а во второй—единичную матрицу порядка q.

[2]) Нам нет необходимости заранее устанавливать неособенность матрицы M, так как это свойство матрицы M следует из того, что $|A| \neq 0$ и $|H| \neq 0$. Если бы оказалось, что $|H| = 0$, то тогда и $|M| = 0$, и в этом случае не существует обратной матрицы M^{-1}.

Если предположить, что $|D| \neq 0$ (вместо $|A| \neq 0$) и поменять ролями матрицы A и D, то можно получить другой вид формулы Фробениуса:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} K^{-1} & -K^{-1}BD^{-1} \\ -D^{-1}CK^{-1} & D^{-1}+D^{-1}CK^{-1}BD^{-1} \end{pmatrix}, \tag{88}$$

где

$$K = A - BD^{-1}C. \tag{89}$$

Пример. Требуется найти элементы обратной матрицы для матрицы

$$M = \begin{Vmatrix} 1 & -1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 2 \end{Vmatrix}.$$

Полагаем

$$A = \begin{Vmatrix} 1 & -1 \\ -1 & 2 \end{Vmatrix}, \quad B = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}, \quad C = \begin{Vmatrix} 0 & 2 \\ 2 & 0 \end{Vmatrix}, \quad D = \begin{Vmatrix} 0 & 1 \\ 1 & 2 \end{Vmatrix}.$$

Находим последовательно

$$A^{-1} = \begin{Vmatrix} 2 & 1 \\ 1 & 1 \end{Vmatrix}, \quad CA^{-1} = \begin{Vmatrix} 0 & 2 \\ 2 & 0 \end{Vmatrix} \begin{Vmatrix} 2 & 1 \\ 1 & 1 \end{Vmatrix} = \begin{Vmatrix} 2 & 2 \\ 4 & 2 \end{Vmatrix},$$

$$H = D - CA^{-1}B = D - \begin{Vmatrix} 2 & 2 \\ 4 & 2 \end{Vmatrix} \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 1 \\ 1 & 2 \end{Vmatrix} - \begin{Vmatrix} -2 & 2 \\ -2 & 4 \end{Vmatrix} = \begin{Vmatrix} 2 & -1 \\ 3 & -2 \end{Vmatrix},$$

$$H^{-1} = \begin{Vmatrix} 2 & -1 \\ 3 & -2 \end{Vmatrix}, \quad A^{-1}B = \begin{Vmatrix} 2 & 1 \\ 1 & 1 \end{Vmatrix} \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} = \begin{Vmatrix} -1 & 2 \\ -1 & 1 \end{Vmatrix},$$

$$A^{-1}BH^{-1} = \begin{Vmatrix} -1 & 2 \\ -1 & 1 \end{Vmatrix} \begin{Vmatrix} 2 & -1 \\ 3 & -2 \end{Vmatrix} = \begin{Vmatrix} 4 & -3 \\ 1 & -1 \end{Vmatrix},$$

$$A^{-1}BH^{-1}CA^{-1} = \begin{Vmatrix} 4 & -3 \\ 1 & -1 \end{Vmatrix} \begin{Vmatrix} 2 & 2 \\ 4 & 2 \end{Vmatrix} = \begin{Vmatrix} -4 & 2 \\ -2 & 0 \end{Vmatrix},$$

$$A^{-1} + A^{-1}BH^{-1}CA^{-1} = \begin{Vmatrix} -2 & 3 \\ -1 & 1 \end{Vmatrix},$$

$$H^{-1}CA^{-1} = \begin{Vmatrix} 2 & -1 \\ 3 & -2 \end{Vmatrix} \begin{Vmatrix} 2 & 2 \\ 4 & 2 \end{Vmatrix} = \begin{Vmatrix} 0 & 2 \\ -2 & 2 \end{Vmatrix}.$$

Поэтому по формуле (86) находим:

$$M^{-1} = \begin{Vmatrix} -2 & 3 & -4 & 3 \\ -1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ 2 & -2 & 3 & -2 \end{Vmatrix}.$$

5. Из теоремы 3 вытекает также

Теорема 4. *Если прямоугольная матрица R представлена в блочном виде*

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{90}$$

где A — квадратная неособенная матрица порядка n ($|A| \neq 0$), то ранг

матрицы R равен n в том и только в том случае, когда

$$D = CA^{-1}B. \tag{91}$$

Доказательство. Вычтем из второй блочной строки матрицы R первую, предварительно помноженную слева на CA^{-1}. Тогда получим матрицу

$$T = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \tag{92}$$

Матрицы R и T согласно теореме 3 имеют один и тот же ранг. Ранг же матрицы T совпадает с рангом матрицы A (т. е. с n) тогда и только тогда, когда $D - CA^{-1}B = 0$, т. е. когда имеет место (91). Теорема доказана.

Из теоремы 4 вытекает алгоритм построения обратной матрицы A^{-1} и вообще произведения $CA^{-1}B$, где B, C — прямоугольные матрицы размером $n \times p$, $q \times n$ [1]).

Приведем при помощи алгоритма Гаусса [2]) матрицу

$$\begin{pmatrix} A & B \\ -C & 0 \end{pmatrix} \quad (|A| \neq 0) \tag{93}$$

к виду

$$\begin{pmatrix} G, & B_1 \\ 0, & X \end{pmatrix}. \tag{94}$$

Докажем, что

$$X = CA^{-1}B. \tag{95}$$

В самом деле, то же преобразование, которое было применено к матрице (93), приведет матрицу

$$\begin{pmatrix} A & B \\ -C & -CA^{-1}B \end{pmatrix} \tag{96}$$

к виду

$$\begin{pmatrix} G & B_1 \\ 0 & X - CA^{-1}B \end{pmatrix}. \tag{97}$$

Согласно теореме 4 матрица (96) имеет ранг n (n — порядок матрицы A). Но тогда и матрица (97) должна иметь ранг n. Отсюда $X - CA^{-1}B = 0$, т. е. имеет место (95).

В частности, если $B = y$, где y — столбцевая матрица, и $C = E$, то

$$X = A^{-1}y.$$

Следовательно, применяя алгоритм Гаусса к матрице

$$\begin{pmatrix} A & y \\ -E & 0 \end{pmatrix},$$

---

[1]) См. [88].

[2]) Здесь мы применяем к матрице (93) не весь алгоритм Гаусса, а только его первые n этапов, где n — порядок матрицы A. Это можно сделать, если выполняются условия (15) при $p = n$. Если же эти условия не выполнены, то, поскольку $|A| \neq 0$, мы можем так перенумеровать первые n строк (или первые n столбцов) матрицы (93), чтобы n этапов алгоритма Гаусса оказались выполнимыми. Такой видоизмененный алгоритм Гаусса иногда применяют и при выполнении условий (15) для $p = n$.

мы получаем решение системы уравнений

$$Ax = y.$$

Далее, если в (93) положить $B = C = E$, то после применения алгоритма Гаусса к матрице

$$\begin{pmatrix} A, & E \\ -E, & 0 \end{pmatrix}$$

получим:

$$\begin{pmatrix} G, & W \\ 0, & X \end{pmatrix},$$

где

$$X = A^{-1}.$$

Проиллюстрируем этот способ нахождения A^{-1} на следующем примере.

Пример. Пусть

$$A = \begin{Vmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 2 \end{Vmatrix}.$$

Требуется вычислить A^{-1}.
Применяем несколько видоизмененный метод исключения [1]) к матрице

$$\begin{Vmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{Vmatrix}.$$

Ко всем строкам прибавляем вторую строку с некоторым множителем и добиваемся того, чтобы все элементы первого столбца, кроме второго элемента, равнялись нулю. После этого ко всем строкам, кроме второй, прибавляем третью строку с некоторым множителем и достигаем того, чтобы во втором столбце все элементы, кроме второго и третьего, были равны нулю. После этого к последним трем строкам прибавляем первую строку с некоторым множителем и получаем матрицу вида

$$\begin{Vmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & 0 & 0 & -2 & -1 & 2 \\ 0 & 0 & 0 & 4 & 1 & -3 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{Vmatrix}.$$

Поэтому

$$A^{-1} = \begin{Vmatrix} -2 & -1 & 2 \\ 4 & 1 & -3 \\ 1 & 1 & -1 \end{Vmatrix}.$$

6. Разбиение матрицы на блоки может быть использовано также для нахождения псевдообратной матрицы (см. гл. I, § 5). Пусть снова

[1]) См. сноску [2]) на стр. 62.

прямоугольная матрица R представлена в виде

$$R=\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{98}$$

где A — неособенная квадратная матрица ($|A|\neq 0$) и $r_A=r_R$. Тогда согласно теореме 4 справедливо равенство $D=CA^{-1}B$, и потому

$$R=\begin{pmatrix} A \\ C \end{pmatrix} A^{-1}(AB). \tag{99}$$

Так как эта формула является результатом двух последовательных скелетных разложений (см. гл. I, § 5):

$$R=\begin{pmatrix} A \\ C \end{pmatrix}(E\ \ A^{-1}B),\quad (E\ \ A^{-1}B)=A^{-1}(AB),$$

то

$$R^+=(AB)^+(A^{-1})^+\begin{pmatrix} A \\ C \end{pmatrix}^+=(AB)^+A\begin{pmatrix} A \\ C \end{pmatrix}^+. \tag{100}$$

Применяя формулы (43) и (44) на стр. 34, окончательно найдем:

$$R^+=\begin{pmatrix} A^* \\ B^* \end{pmatrix}(AA^*+BB^*)^{-1}A\,(A^*A+C^*C)^{-1}(A^*C^*). \tag{101}$$

Формула (101) дает явное выражение для псевдообратной матрицы R^+ через блоки A, B, C.

Пример.

$$R=\left\|\begin{array}{cc|cc} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ \hline 0 & 1 & -1 & 1 \end{array}\right\|=\begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Здесь $r_R=2$ и

$$A=\begin{Vmatrix} 1 & -1 \\ -1 & 2 \end{Vmatrix},\quad B=\begin{Vmatrix} 2 & 0 \\ -3 & 1 \end{Vmatrix},\quad C=\|0\ \ 1\|,\quad D=\|-1\ \ 1\|,$$

$$(AA^*+BB^*)^{-1}=\left(3\begin{Vmatrix} 2 & -3 \\ -3 & 5 \end{Vmatrix}\right)^{-1}=\frac{1}{3}\begin{Vmatrix} 5 & 3 \\ 3 & 2 \end{Vmatrix},$$

$$(A^*A+C^*C)^{-1}=\begin{Vmatrix} 2 & -3 \\ -3 & 6 \end{Vmatrix}^{-1}=\frac{1}{3}\begin{Vmatrix} 6 & 3 \\ 3 & 2 \end{Vmatrix},$$

$$(AA^*+BB^*)^{-1}A\,(A^*A+C^*C)^{-1}=\frac{1}{9}\begin{Vmatrix} 15 & 8 \\ 9 & 5 \end{Vmatrix}.$$

Тогда по формуле (101)

$$R^+=\frac{1}{9}\begin{Vmatrix} 1 & -1 \\ -1 & 2 \\ 2 & -3 \\ 0 & 1 \end{Vmatrix}\begin{Vmatrix} 15 & 8 \\ 9 & 5 \end{Vmatrix}\begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{Vmatrix}=\begin{Vmatrix} \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{2}{9} & -\frac{1}{9} & \frac{1}{9} \\ \frac{4}{9} & \frac{1}{9} & \frac{5}{9} \end{Vmatrix}.$$

ГЛАВА III

ЛИНЕЙНЫЕ ОПЕРАТОРЫ В n-МЕРНОМ ВЕКТОРНОМ ПРОСТРАНСТВЕ

Матрицы составляют основной аналитический аппарат для изучения линейных операций в n-мерном пространстве. В свою очередь изучение этих операций дает возможность разбить все матрицы на классы и выявить важные свойства, присущие всем матрицам одного и того же класса.

В настоящей главе излагаются наиболее простые свойства линейных операторов в n-мерном пространстве. Дальнейшее исследование линейных операторов в n-мерном пространстве будет продолжено в главах VII и IX.

§ 1. Векторное пространство

1. Пусть дана некоторая совокупность $\boldsymbol{R}$ произвольных элементов x, y, z, ..., в которой определены две операции: операция «сложения» и операция «умножения на число из поля K»[1]. Допустим, что эти операции всегда выполнимы и однозначны в $\boldsymbol{R}$ и для любых элементов x, y, z из $\boldsymbol{R}$ и чисел α, β из K:

1° $$x+y=y+x.$$

2° $$(x+y)+z=x+(y+z).$$

3° Существует такой элемент $\mathbf{0}$ в $\boldsymbol{R}$, что произведение числа 0 на любой элемент x из $\boldsymbol{R}$ равно элементу $\mathbf{0}$:

$$0\cdot x=\mathbf{0}.$$

4° $$1\cdot x=x.$$

5° $$\alpha(\beta x)=(\alpha\beta)x.$$

6° $$(\alpha+\beta)x=\alpha x+\beta x.$$

7° $$\alpha(x+y)=\alpha x+\alpha y.$$

Определение 1. Совокупность элементов $\boldsymbol{R}$, в которой всегда выполнимы и однозначны две операции: «сложение» элементов и «умножение элемента из $\boldsymbol{R}$ на число из K», причем эти операторы удовлетворяют постулатам 1°—7°, мы будем называть *векторным пространством* (над полем K), а сами элементы — *векторами*[2].

---

[1]) Эти операции будем отмечать обычными знаками «$+$» и «$\cdot$», причем последний знак иногда не ставится, а только подразумевается.

[2]) Нетрудно видеть, что из свойств 1°—7° следуют все обычные свойства операций сложения и умножения на число. Так, например, при любом x из $\boldsymbol{R}$: $x+0=1\cdot x+0\cdot x=(1+0)x=1\cdot x=x$, $x+(-x)=0$, где $-x=(-1)\cdot x$ и т. п.

Определение 2. Векторы x, y, ..., u из R называются *линейно зависимыми*, если существуют такие числа α, β, ..., δ из K, не равные одновременно нулю, что

$$\alpha x+\beta y+\ldots+\delta u=0. \tag{1}$$

В случае, если не существует подобной линейной зависимости, векторы x, y, ..., u называются *линейно независимыми*.

Если векторы x, y, ..., u линейно зависимы, то один из векторов может быть представлен в виде линейной комбинации остальных с коэффициентами из поля K. Так, например, если в (1) $\alpha\neq 0$, то

$$x=-\frac{\beta}{\alpha}y-\ldots-\frac{\delta}{\alpha}u.$$

Определение 3. Пространство R называется *конечномерным*, а число n — *числом измерений* этого пространства, если в R существует n линейно независимых векторов, в то время как любые $n+1$ векторов из R линейно зависимы. Если же в пространстве можно найти линейно независимую систему из любого числа векторов, то пространство называется *бесконечномерным*.

В настоящей книге в основном изучаются конечномерные пространства.

Определение 4. Система из n линейно независимых заданных в определенном порядке векторов e_1, e_2, ..., e_n в n-мерном пространстве называется *базисом* этого пространства.

2. Пример 1. Совокупность обычных векторов (направленных геометрических отрезков) является трехмерным векторным пространством. Часть этого пространства, состоящая из векторов, параллельных некоторой плоскости, является двумерным пространством, а все векторы, параллельные некоторой прямой, образуют одномерное векторное пространство.

Пример 2. Столбец из n чисел поля K: $x=(x_1, x_2, \ldots, x_n)$ назовем вектором (n — фиксированное число). Основные операции определим как операции над столбцевыми матрицами:

$$(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n)=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n),$$
$$\alpha(x_1, x_2, \ldots, x_n)=(\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

Элементом нуль будет столбец (0, 0, ..., 0). Легко проверить, что все постулаты 1°—7° выполняются. Эти векторы образуют n-мерное пространство. В качестве базиса этого пространства можно, например, взять столбцы единичной матрицы n-го порядка:

$$(1, 0, \ldots, 0),\quad (0, 1, \ldots, 0),\quad \ldots,\quad (0, 0, \ldots, 1).$$

Пространство, рассмотренное в этом примере, часто называют *численным n-мерным пространством*.

Пример 3. Совокупность бесконечных последовательностей $(x_1, x_2, \ldots, x_n, \ldots)$, в которой операции определены естественным образом, т. е.

$$(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n, \ldots)=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n, \ldots),$$
$$\alpha(x_1, x_2, \ldots, x_n, \ldots)=(\alpha x_1, \alpha x_2, \ldots, \alpha x_n, \ldots),$$

представляет собой бесконечномерное пространство.

Пример 4. Совокупность многочленов $\alpha_0+\alpha_1 t+\ldots+\alpha_{n-1}t^{n-1}$ степени $<n$ с коэффициентами из K представляет собой n-мерное векторное пространство[1]. Базисом такого пространства является, например, система степеней t^0, t, t^2, ..., t^{n-1}.

Все же такие многочлены (без ограничения степени) образуют бесконечномерное пространство.

Пример 5. Все функции, определенные в замкнутом интервале $[a, b]$, образуют бесконечномерное пространство.

---

[1]) В качестве основных операций берутся обычное сложение многочленов и умножение многочлена на число.

3. Пусть векторы $e_1, e_2, \ldots, e_n$ образуют базис n-мерного векторного пространства R, а x — произвольный вектор этого пространства. Тогда векторы $x, e_1, e_2, \ldots, e_n$ линейно зависимы (ибо число их равно $n+1$):

$$\alpha_0 x + \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0,$$

где по крайней мере одно из чисел $\alpha_0, \alpha_1, \ldots, \alpha_n$ отлично от нуля. Однако в данном случае $\alpha_0 \neq 0$, так как векторы $e_1, e_2, \ldots, e_n$ не могут быть связаны линейной зависимостью. Поэтому

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n, \tag{2}$$

где $x_i = -\frac{\alpha_i}{\alpha_0}$ $(i = 1, 2, \ldots, n)$.

Заметим, что числа $x_1, x_2, \ldots, x_n$ однозначно определяются заданием вектора x и базиса $e_1, e_2, \ldots, e_n$. В самом деле, если наряду с (2) имеется другое разложение для вектора x:

$$x = x_1' e_1 + x_2' e_2 + \ldots + x_n' e_n, \tag{2'}$$

то, вычитая почленно (2) из (2'), получим:

$$(x_1' - x_1) e_1 + (x_2' - x_2) e_2 + \ldots + (x_n' - x_n) e_n = 0,$$

откуда в силу линейной независимости векторов базиса следует

$$x_1' - x_1 = x_2' - x_2 = \ldots = x_n' - x_n = 0,$$

т. е.

$$x_1' = x_1, \quad x_2' = x_2, \quad \ldots, \quad x_n' = x_n. \tag{3}$$

Числа $x_1, x_2, \ldots, x_n$ называются *координатами* вектора x в базисе $e_1, e_2, \ldots, e_n$.

Если

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{i=1}^{n} y_i e_i,$$

то

$$x + y = \sum_{i=1}^{n} (x_i + y_i) e_i, \qquad \alpha x = \sum_{i=1}^{n} \alpha x_i e_i,$$

т. е. *координаты суммы векторов получаются почленным сложением соответствующих координат слагаемых векторов и при умножении вектора на число α все координаты вектора умножаются на это число.*

4. Пусть векторы

$$x_k = \sum_{i=1}^{n} x_{ik} e_i \qquad (k = 1, 2, \ldots, m)$$

линейно зависимы, т. е.

$$\sum_{k=1}^{m} c_k x_k = 0, \tag{4}$$

где по крайней мере одно из чисел $c_1, c_2, \ldots, c_m$ не равно нулю.

Если вектор равен нулю, то все его координаты равны нулю. Поэтому векторное равенство (4) эквивалентно следующей системе скалярных равенств:

$$\left.\begin{aligned} c_1 x_{11} + c_2 x_{12} + \ldots + c_m x_{1m} &= 0, \\ c_1 x_{21} + c_2 x_{22} + \ldots + c_m x_{2m} &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ c_1 x_{n1} + c_2 x_{n2} + \ldots + c_m x_{nm} &= 0. \end{aligned}\right\} \tag{4'}$$

Эта система однородных линейных уравнений относительно $c_1, c_2, \ldots, c_m$, как известно, имеет ненулевое решение тогда и только тогда, когда ранг матрицы коэффициентов меньше числа неизвестных, т. е. меньше m. Поэтому равенство этого ранга числу m является необходимым и достаточным условием для линейной независимости векторов $x_1, x_2, \ldots, x_m$.

Таким образом, имеет место следующая

Теорема 1. *Для того чтобы векторы $x_1, x_2, \ldots, x_m$ были линейно независимы, необходимо и достаточно, чтобы ранг r матрицы, составленной из координат этих векторов в произвольном базисе,*

$$X = \begin{Vmatrix} x_{11} & x_{12} & \ldots & x_{1m} \\ x_{21} & x_{22} & \ldots & x_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1} & x_{n2} & \ldots & x_{nm} \end{Vmatrix}$$

был равен m, т. е. числу векторов.

Замечание. Линейная независимость векторов $x_1, x_2, \ldots, x_m$ означает линейную независимость столбцов матрицы X, поскольку в k-м столбце стоят координаты вектора x_k ($k = 1, 2, \ldots, m$). Поэтому, согласно теореме, если в прямоугольной матрице столбцы линейно независимы, то ранг матрицы равен числу столбцов. Отсюда следует, что в произвольной прямоугольной матрице максимальное число линейно независимых столбцов равно рангу матрицы. Кроме того, если мы транспонируем матрицу, т. е. строки делаем столбцами (и наоборот), то ранг матрицы при этом, очевидно, не меняется. Поэтому *в прямоугольной матрице всегда число линейно независимых столбцов равно числу линейно независимых строк и равно рангу матрицы.*

5. Если в n-мерном пространстве выбран базис $e_1, e_2, \ldots, e_n$, то каждому вектору x однозначно отвечает столбец $x = (x_1, x_2, \ldots, x_n)$, где $x_1, x_2, \ldots, x_n$ —координаты вектора x в данном базисе. Таким образом, задание базиса устанавливает взаимно однозначное соответствие между векторами произвольного n-мерного векторного пространства $\boldsymbol{R}$ и векторами n-мерного численного пространства $\boldsymbol{R}'$, рассмотренного в примере 2. При этом сумме векторов из $\boldsymbol{R}$ отвечает сумма соответствующих векторов из $\boldsymbol{R}'$. Аналогичное имеет место и для произведения вектора на число α из K. Другими словами, произвольное n-мерное векторное пространство *изоморфно* численному n-мерному пространству и, следовательно, *все векторные пространства одного и того же числа измерений n над одним и тем же числовым полем K изоморфны между собой.* Это означает, что с точностью до изоморфизма существует только одно n-мерное векторное пространство при заданном числовом поле.

Может возникнуть вопрос, зачем мы ввели «абстрактное» n-мерное пространство, если с точностью до изоморфизма оно совпадает с n-мерным численным пространством. Действительно, можно было бы определить вектор как систему n чисел, заданных в определенном порядке, и установить операцию над этими векторами, как это было сделано в примере 2. Но при этом смешались бы воедино свойства вектора, не зависящие от выбора базиса, со свойствами специального базиса. Например, равенство нулю всех координат вектора есть свойство самого вектора; оно не зависит от выбора базиса. Равенство между собой всех координат вектора не есть свойство самого вектора, ибо при изменении базиса оно исчезает. *Аксиоматическое определение векторного пространства непосредственно выделяет свойства векторов, не зависящие от выбора базиса.*

6. Если некоторая совокупность векторов $\boldsymbol{R}'$, составляющая часть $\boldsymbol{R}$, обладает тем свойством, что сумма любых двух векторов из $\boldsymbol{R}'$ и произведение любого вектора из $\boldsymbol{R}'$ на число $\alpha \in K$ всегда принадлежат $\boldsymbol{R}'$, то такое многообразие $\boldsymbol{R}'$ само является векторным пространством, *подпространством* в $\boldsymbol{R}$.

Если даны два подпространства R' и R'' в R и известно, что
1° R' и R'' не имеют общих векторов, кроме нуля, и
2° любой вектор x из R представляется в виде суммы

$$x = x' + x'' \qquad (x' \in R',\ x'' \in R''), \tag{5}$$

то мы будем говорить, что пространство R *расщепляется* на два подпространства R' и R'' и будем писать:

$$R = R' + R''. \tag{6}$$

Заметим, что условие 1° означает единственность представления (5). В самом деле, если бы для некоторого вектора x мы имели два разных представления в виде суммы слагаемых из R' и R'', представление (5) и представление

$$x = \tilde{x}' + \tilde{x}'' \qquad (\tilde{x}' \in R',\ \tilde{x}'' \in R''), \tag{7}$$

то, вычитая почленно (7) из (5), мы получили бы:

$$x' - \tilde{x}' = \tilde{x}'' - x'',$$

т. е. равенство между отличными от нуля векторами $x' - \tilde{x}' \in R'$ и $\tilde{x}'' - x'' \in R''$, что невозможно в силу 1°.

Таким образом, условие 1° можно заменить требованием единственности представления (5). В таком виде определение расщепления непосредственно распространяется на любое число слагаемых подпространств.

Пусть

$$R = R' + R''$$

и $e'_1, e'_2, \ldots, e'_{n'}$ и $e''_1, e''_2, \ldots, e''_{n''}$ — базисы соответственно в R' и R''. Тогда читатель без труда докажет, что все эти $n' + n''$ векторов линейно независимы и образуют базис в R, т. е. что из базисов слагаемых подпространств составляется базис всего пространства. В частности, отсюда будет следовать, что $n = n' + n''$.

Пример 1. Пусть в пространстве трех измерений даны три непараллельных одной и той же плоскости направления. Так как любой вектор в пространстве можно разложить на составляющие по этим трем направлениям, притом единственным образом, то

$$R = R' + R'' + R''',$$

где R — совокупность всех векторов нашего пространства, R' — совокупность всех векторов, параллельных первому направлению, R'' — второму, R''' — третьему. В данном случае $n = 3$, $n' = n'' = n''' = 1$.

Пример 2. Пусть в пространстве трех измерений даны плоскость и пересекающая ее прямая. Тогда

$$R = R' + R'',$$

где R — совокупность всех векторов нашего пространства, R' — совокупность всех векторов, параллельных заданной плоскости, и R'' — совокупность всех векторов, параллельных заданной прямой. В этом примере $n = 3$, $n' = 2$, $n'' = 1$.

Задание базиса $e_1, e_2, \ldots, e_n$ в пространстве R по существу означает некоторое расщепление всего пространства R на n одномерных подпространств.

§ 2. Линейный оператор, отображающий n-мерное пространство в m-мерное

1. Рассмотрим линейное преобразование

$$\left.\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{aligned}\right\} \tag{8}$$

коэффициенты которого принадлежат числовому полю K, и два векторных пространства над этим полем: n-мерное $\boldsymbol{R}$ и m-мерное $\boldsymbol{S}$. Выберем в $\boldsymbol{R}$ некоторый базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ и в $\boldsymbol{S}$ некоторый базис $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$. Тогда преобразование (8) относит каждому вектору $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$ из $\boldsymbol{R}$ некоторый вектор $\boldsymbol{y} = \sum_{k=1}^{m} y_k \boldsymbol{g}_k$ из $\boldsymbol{S}$, т. е. преобразование (8) определяет некоторый оператор $\boldsymbol{A}$, относящий вектору $\boldsymbol{x}$ вектор $\boldsymbol{y}$: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. Нетрудно видеть, что этот оператор $\boldsymbol{A}$ обладает свойством линейности, которое мы сформулируем так:

Определение 5. Оператор $\boldsymbol{A}$, отображающий $\boldsymbol{R}$ в $\boldsymbol{S}$, т. е. относящий каждому вектору $\boldsymbol{x}$ из $\boldsymbol{R}$ некоторый вектор $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ из $\boldsymbol{S}$, называется *линейным*, если для любых $\boldsymbol{x}$, $\boldsymbol{x}_1$ из $\boldsymbol{R}$ и α из K

$$\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{x}_1) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{x}_1, \quad \boldsymbol{A}(\alpha\boldsymbol{x}) = \alpha\boldsymbol{A}\boldsymbol{x}. \tag{9}$$

Таким образом, преобразование (8) при заданных базисах в $\boldsymbol{R}$ и $\boldsymbol{S}$ определяет некоторый линейный оператор, отображающий $\boldsymbol{R}$ в $\boldsymbol{S}$.

Покажем теперь обратное, т. е. что для произвольного линейного оператора $\boldsymbol{A}$, отображающего $\boldsymbol{R}$ в $\boldsymbol{S}$, и произвольных базисов $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ в $\boldsymbol{R}$ и $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$ в $\boldsymbol{S}$ существует такая прямоугольная матрица с элементами из поля K

$$\left\|\begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{matrix}\right\|, \tag{10}$$

что составленное при помощи этой матрицы линейное преобразование (1) выражает координаты преобразованного вектора $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ через координаты исходного вектора $\boldsymbol{x}$.

Действительно, применим оператор $\boldsymbol{A}$ к базисному вектору $\boldsymbol{e}_k$ и координаты полученного вектора $\boldsymbol{A}\boldsymbol{e}_k$ в базисе $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$ обозначим через $a_{1k}, a_{2k}, \ldots, a_{mk}$ $(k = 1, 2, \ldots, n)$:

$$\boldsymbol{A}\boldsymbol{e}_k = \sum_{i=1}^{m} a_{ik}\boldsymbol{g}_i \quad (k = 1, 2, \ldots, n). \tag{11}$$

Помножая обе части равенства (11) на x_k и суммируя в пределах от 1 до n, получим:

$$\sum_{k=1}^{n} x_k \boldsymbol{A}\boldsymbol{e}_k = \sum_{i=1}^{m} \Big(\sum_{k=1}^{n} a_{ik}x_k\Big)\boldsymbol{g}_i,$$

откуда

$$y = Ax = A\left(\sum_{k=1}^{n} x_k e_k\right) = \sum_{k=1}^{n} x_k Ae_k = \sum_{i=1}^{m} y_i g_i,$$

где

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, \ldots, m),$$

что и требовалось установить.

Таким образом, при заданных базисах в R и S *каждому линейному оператору* A, *отображающему* R *в* S, *отвечает некоторая прямоугольная матрица* (10) *с размерами* $m \times n$ *и, наоборот, каждой такой матрице отвечает некоторый линейный оператор, отображающий* R *в* S.

При этом в матрице A, отвечающей оператору A, *k-й столбец состоит из последовательных координат вектора* Ae_k $(k = 1, 2, \ldots, n)$.

Обозначим через $x = (x_1, x_2, \ldots, x_n)$ и $y = (y_1, y_2, \ldots, y_n)$ столбцы координат векторов x и y. Тогда векторному равенству

$$y = Ax$$

соответствует матричное равенство

$$y = Ax,$$

которое является матричной записью преобразования (8).

Пример.

Рассмотрим совокупность всех многочленов от t степени $\leqslant n-1$ с коэффициентами из числового поля K. Эта совокупность представляет собой некоторое n-мерное векторное пространство R_n (см. пример 4 на стр. 66). Точно так же многочлены от t степени $\leqslant n-2$ с коэффициентами из K образуют пространство R_{n-1}. Оператор дифференцирования $\frac{d}{dt}$ относит каждому многочлену из R_n некоторый многочлен из R_{n-1}. Таким образом, этот оператор отображает R_n в R_{n-1}. Оператор дифференцирования является линейным оператором, так как

$$\frac{d}{dt}[\varphi(t) + \psi(t)] = \frac{d\varphi(t)}{dt} + \frac{d\psi(t)}{dt}, \quad \frac{d}{dt}[\alpha\varphi(t)] = \alpha\frac{d\varphi(t)}{dt}.$$

В пространствах R_n и R_{n-1} выберем базисы из степеней t:

$$t^0 = 1, \ t, \ \ldots, \ t^{n-1} \text{ и } t^0 = 1, \ t, \ \ldots, \ t^{n-2}.$$

Пользуясь формулой (11), построим прямоугольную матрицу размером $(n-1) \times n$, соответствующую оператору дифференцирования $\frac{d}{dt}$ в этих базисах:

$$\begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & n-1 \end{Vmatrix}.$$

§ 3. Сложение и умножение линейных операторов

1. Пусть даны два линейных оператора A и B, отображающие R в S, и соответствующие им матрицы

$$A = \|a_{ik}\|, \quad B = \|b_{ik}\| \quad (i = 1, 2, \ldots, m; \ k = 1, 2, \ldots, n).$$

Определение 6. *Суммой операторов* $\boldsymbol{A}$ и $\boldsymbol{B}$ называется оператор $\boldsymbol{C}$, определяемый равенством

$$\boldsymbol{Cx} = \boldsymbol{Ax} + \boldsymbol{Bx} \quad (\boldsymbol{x} \in \boldsymbol{R})^{1}). \tag{12}$$

На основе этого определения легко проверяется, что сумма $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ линейных операторов $\boldsymbol{A}$ и $\boldsymbol{B}$ есть также линейный оператор.

Далее,

$$\boldsymbol{Ce}_k = \boldsymbol{Ae}_k + \boldsymbol{Be}_k = \sum_{k=1}^{n} (a_{ik} + b_{ik})\, \boldsymbol{e}_k.$$

Отсюда следует, что оператору $\boldsymbol{C}$ отвечает матрица $C = \| c_{ik} \|$, где $c_{ik} = a_{ik} + b_{ik}$ $(i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$, т. е. оператору $\boldsymbol{C}$ отвечает матрица

$$C = A + B. \tag{13}$$

К этому же выводу можно прийти из рассмотрения матричного равенства

$$Cx = Ax + Bx \tag{14}$$

(x — столбец координат вектора $\boldsymbol{x}$), соответствующего векторному равенству (12). Поскольку x — произвольный столбец, то из (14) следует (13).

2. Пусть даны три векторных пространства $\boldsymbol{R}$, $\boldsymbol{S}$ и $\boldsymbol{T}$ соответственно q, n и m измерений и два линейных оператора $\boldsymbol{A}$ и $\boldsymbol{B}$, из которых $\boldsymbol{B}$ отображает $\boldsymbol{R}$ в $\boldsymbol{S}$, а $\boldsymbol{A}$ отображает $\boldsymbol{S}$ в $\boldsymbol{T}$; в символической записи:

$$\boldsymbol{T} \xleftarrow{A} \boldsymbol{S} \xleftarrow{B} \boldsymbol{R}.$$

Определение 7. *Произведением операторов* $\boldsymbol{A}$ и $\boldsymbol{B}$ называется оператор $\boldsymbol{C}$, для которого при любом $\boldsymbol{x}$ из $\boldsymbol{R}$

$$\boldsymbol{Cx} = \boldsymbol{A}(\boldsymbol{Bx}) \quad (\boldsymbol{x} \in \boldsymbol{R}). \tag{15}$$

Оператор $\boldsymbol{C}$ отображает $\boldsymbol{R}$ в $\boldsymbol{T}$:

$$\boldsymbol{T} \xleftarrow{C=AB} \boldsymbol{R}.$$

Из линейности операторов $\boldsymbol{A}$ и $\boldsymbol{B}$ вытекает линейность оператора $\boldsymbol{C}$. Выберем в пространствах $\boldsymbol{R}$, $\boldsymbol{S}$, $\boldsymbol{T}$ произвольные базисы и обозначим через A, B, C матрицы, соответствующие операторам $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ при этом выборе базисов. Тогда векторным равенствам

$$\boldsymbol{z} = \boldsymbol{Ay},\ \boldsymbol{y} = \boldsymbol{Bx},\ \boldsymbol{z} = \boldsymbol{Cx} \tag{16}$$

будут соответствовать матричные равенства

$$z = Ay,\ y = Bx,\ z = Cx,$$

где x, y, z — столбцы координат векторов $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$. Отсюда находим:

$$Cx = A(Bx) = (AB)\, x,$$

и в силу произвольности столбца x

$$C = AB. \tag{17}$$

Таким образом, произведению $\boldsymbol{C} = \boldsymbol{AB}$ операторов $\boldsymbol{A}$ и $\boldsymbol{B}$ отвечает матрица $C = \| c_{ij} \|$ $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, q)$, равная произведению матриц A и B.

---

[1]) $\boldsymbol{x} \in \boldsymbol{R}$ означает, что элемент $\boldsymbol{x}$ принадлежит совокупности $\boldsymbol{R}$. Предполагается, что равенство (12) имеет место при любом $\boldsymbol{x}$ из $\boldsymbol{R}$.

Предоставляем читателю самому доказать, что оператору

$$\boldsymbol{C}=\alpha\boldsymbol{A}\ (\alpha\in K)^{1})$$

отвечает матрица

$$C=\alpha A.$$

Таким образом, мы видим, что в главе I действия над матрицами были определены так, что сумме линейных операторов $\boldsymbol{A}+\boldsymbol{B}$, произведениям $\boldsymbol{AB}$ и $\alpha\boldsymbol{A}$ отвечают соответственно матрицы $A+B$, AB и αA, где A и B — матрицы, соответствующие операторам $\boldsymbol{A}$ и $\boldsymbol{B}$, а α — число из K.

§ 4. Преобразование координат

Рассмотрим в n-мерном векторном пространстве два базиса: $e_1, e_2, \ldots, e_n$ («старый» базис) и $\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n$ («новый» базис).

Взаимное расположение векторов базиса определится, если задать координаты векторов одного базиса относительно другого.

Мы положим:

$$\left.\begin{aligned}
\tilde{e}_1 &= t_{11}e_1 + t_{21}e_2 + \ldots + t_{n1}e_n,\\
\tilde{e}_2 &= t_{12}e_1 + t_{22}e_2 + \ldots + t_{n2}e_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\tilde{e}_n &= t_{1n}e_1 + t_{2n}e_2 + \ldots + t_{nn}e_n,
\end{aligned}\right\} \tag{18}$$

или в сокращенной записи:

$$\tilde{e}_k = \sum_{i=1}^{n} t_{ik}e_i \qquad (k=1, 2, \ldots, n). \tag{18'}$$

Установим связь между координатами одного и того же вектора в различных базисах.

Пусть $x_1, x_2, \ldots, x_n$ и $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ — координаты вектора x соответственно в «старом» и «новом» базисах:

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{k=1}^{n} \tilde{x}_k \tilde{e}_k. \tag{19}$$

Подставим в (19) вместо векторов $\tilde{e}_k$ их выражения из (18). Получим:

$$x = \sum_{k=1}^{n} \tilde{x}_k \sum_{i=1}^{n} t_{ik}e_i = \sum_{i=1}^{n} \Big(\sum_{k=1}^{n} t_{ik}\tilde{x}_k\Big) e_i.$$

Сопоставляя это равенство с (19) и учитывая, что координаты вектора однозначно определяются заданием вектора и базиса, находим:

$$x_i = \sum_{k=1}^{n} t_{ik}\tilde{x}_k \qquad (i=1, 2, \ldots, n), \tag{20}$$

---

[1]) То есть оператору, для которого $\boldsymbol{Cx}=\alpha\boldsymbol{Ax}$ $(\boldsymbol{x}\in\boldsymbol{R})$.

или в подробной записи:

$$\left.\begin{aligned} x_1 &= t_{11}\widetilde{x}_1 + t_{12}\widetilde{x}_2 + \ldots + t_{1n}\widetilde{x}_n, \\ x_2 &= t_{21}\widetilde{x}_1 + t_{22}\widetilde{x}_2 + \ldots + t_{2n}\widetilde{x}_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= t_{n1}\widetilde{x}_1 + t_{n2}\widetilde{x}_2 + \ldots + t_{nn}\widetilde{x}_n. \end{aligned}\right\} \tag{21}$$

Формулы (21) определяют преобразование координат вектора при переходе от одного базиса к другому. Они выражают «старые» координаты через «новые». Матрица

$$T = \| t_{ik} \|_1^n \tag{22}$$

называется *матрицей преобразования координат* или *преобразующей матрицей*. В ней k-й столбец состоит из «старых» координат k-го «нового» базисного вектора. В этом можно убедиться из формулы (18) или непосредственно из формул (21), положив в последних $\widetilde{x}_k = 1$, $\widetilde{x}_i = 0$ при $i \neq k$.

Заметим, что матрица T неособенная, т. е.

$$|T| \neq 0. \tag{23}$$

Действительно, положив в (21) $x_1 = x_2 = \ldots = x_n = 0$, получим систему n линейных однородных уравнений с n неизвестными $\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n$ и с определителем $|T|$. Эта система может иметь только нулевое решение $\widetilde{x}_1 = 0$, $\widetilde{x}_2 = 0$, $\ldots$, $\widetilde{x}_n = 0$, так как в противном случае из (19) следовала бы линейная зависимость между векторами $\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n$. Поэтому $|T| \neq 0$ [1]).

Введем в рассмотрение столбцевые матрицы $x = (x_1, x_2, \ldots, x_n)$ и $\widetilde{x} = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n)$.

Тогда формулы преобразования координат (21) могут быть записаны в виде следующего матричного равенства:

$$x = T\widetilde{x}. \tag{24}$$

Помножая снова обе части этого равенства на T^{-1}, найдем выражение для обратного преобразования

$$\widetilde{x} = T^{-1}x. \tag{25}$$

§ 5. Эквивалентные матрицы. Ранг оператора. Неравенства Сильвестра

1. Пусть даны два векторных пространства $\boldsymbol{R}$ и $\boldsymbol{S}$, соответственно n и m измерений над числовым полем K, и линейный оператор $\boldsymbol{A}$, отображающий $\boldsymbol{R}$ в $\boldsymbol{S}$. В настоящем параграфе мы выясним, как меняется матрица A, соответствующая данному линейному оператору $\boldsymbol{A}$, при изменении базисов в $\boldsymbol{R}$ и $\boldsymbol{S}$.

Выберем в $\boldsymbol{R}$ и $\boldsymbol{S}$ произвольные базисы $e_1, e_2, \ldots, e_n$ и $g_1, g_2, \ldots, g_m$. В этих базисах оператору $\boldsymbol{A}$ будет соответствовать матрица $A = \| a_{ik} \|$ ($i = 1, 2, \ldots, m$; $k = 1, 2, \ldots, n$). Векторному равенству

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{26}$$

[1]) Неравенство (23) вытекает и из теоремы 1 (стр. 68), поскольку элементами матрицы T являются «старые» координаты линейно независимых векторов $\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n$.

соответствует матричное равенство

$$y = Ax, \tag{27}$$

где x и y — координатные столбцы для векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в базисах $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ и $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$.

Выберем теперь в $\boldsymbol{R}$ и $\boldsymbol{S}$ другие базисы $\widetilde{\boldsymbol{e}}_1, \widetilde{\boldsymbol{e}}_2, \ldots, \widetilde{\boldsymbol{e}}_n$ и $\widetilde{\boldsymbol{g}}_1, \widetilde{\boldsymbol{g}}_2, \ldots, \widetilde{\boldsymbol{g}}_m$. В новых базисах вместо x, y, A будем иметь: $\widetilde{x}$, $\widetilde{y}$, $\widetilde{A}$. При этом

$$\widetilde{y} = \widetilde{A}\,\widetilde{x}. \tag{28}$$

Обозначим через Q и N неособенные квадратные матрицы соответственно порядков n и m, осуществляющие преобразование координат в пространствах $\boldsymbol{R}$ и $\boldsymbol{S}$ при переходе от старых базисов к новым (см. § 4):

$$x = Q\widetilde{x}, \quad y = N\widetilde{y}. \tag{29}$$

Тогда из (27) и (29) получаем:

$$\widetilde{y} = N^{-1}y = N^{-1}Ax = N^{-1}AQ\widetilde{x}. \tag{30}$$

Полагая $P = N^{-1}$, мы из (28) и (30) находим:

$$\widetilde{A} = PAQ. \tag{31}$$

Определение 8. Две прямоугольные матрицы A и B одинаковых размеров называются *эквивалентными*, если существуют две неособенные квадратные матрицы P и Q такие [1]), что

$$B = PAQ. \tag{32}$$

Из (31) следует, что две матрицы, соответствующие одному и тому же линейному оператору $\boldsymbol{A}$ при различном выборе базисов в $\boldsymbol{R}$ и $\boldsymbol{S}$, всегда эквивалентны между собой. Нетрудно видеть, что и обратно, если матрица A отвечает оператору $\boldsymbol{A}$ при некоторых базисах в $\boldsymbol{R}$ и $\boldsymbol{S}$, а матрица B эквивалентна матрице A, то она отвечает тому же линейному оператору при некоторых других базисах в $\boldsymbol{R}$ и $\boldsymbol{S}$.

Таким образом, каждому линейному оператору, отображающему $\boldsymbol{R}$ в $\boldsymbol{S}$, соответствует класс эквивалентных между собой матриц с элементами из поля K.

2. Следующая теорема устанавливает критерий эквивалентности двух матриц:

Теорема 2. *Для того чтобы две прямоугольные матрицы одинаковых размеров были эквивалентны, необходимо и достаточно, чтобы эти матрицы имели один и тот же ранг.*

Доказательство. Условие необходимо. При умножении прямоугольной матрицы на какую-либо неособенную квадратную матрицу (слева или справа) ранг исходной прямоугольной матрицы не может измениться (см. гл. I, стр. 27). Поэтому из (32) следует

$$r_A = r_B.$$

Условие достаточно. Пусть A — прямоугольная матрица размера $m \times n$. Она определяет линейный оператор $\boldsymbol{A}$, отображающий пространство $\boldsymbol{R}$ с базисом $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ в пространство $\boldsymbol{S}$ с базисом

[1]) Если матрицы A и B имеют размеры $m \times n$, то в (32) квадратная матрица P имеет порядок m, а квадратная матрица Q — порядок n. Если элементы эквивалентных матриц A и B принадлежат некоторому числовому полю, то матрицы P и Q могут быть выбраны так, чтобы их элементы принадлежали тому же числовому полю.

$\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$. Обозначим через r число линейно независимых векторов среди векторов $\boldsymbol{Ae}_1, \boldsymbol{Ae}_2, \ldots, \boldsymbol{Ae}_n$. Не нарушая общности, можем считать, что линейно независимыми являются векторы $\boldsymbol{Ae}_1, \boldsymbol{Ae}_2, \ldots, \boldsymbol{Ae}_r$ [1]), а остальные, $\boldsymbol{Ae}_{r+1}, \ldots, \boldsymbol{Ae}_n$, выражаются линейно через них:

$$\boldsymbol{Ae}_k = \sum_{j=1}^{r} c_{kj} \boldsymbol{Ae}_j \qquad (k = r+1, \ldots, n). \tag{33}$$

Определим новый базис в $\boldsymbol{R}$ следующим образом:

$$\widetilde{\boldsymbol{e}}_i = \begin{cases} \boldsymbol{e}_i & (i = 1, 2, \ldots, r), \\ \boldsymbol{e}_i - \sum\limits_{j=1}^{r} c_{ij} \boldsymbol{e}_j & (i = r+1, \ldots, n). \end{cases} \tag{34}$$

Тогда в силу (33)

$$\boldsymbol{A}\widetilde{\boldsymbol{e}}_k = 0 \qquad (k = r+1, \ldots, n). \tag{35}$$

Далее положим:

$$\boldsymbol{A}\widetilde{\boldsymbol{e}}_j = \widetilde{\boldsymbol{g}}_j \qquad (j = 1, 2, \ldots, r). \tag{36}$$

Векторы $\widetilde{\boldsymbol{g}}_1, \widetilde{\boldsymbol{g}}_2, \ldots, \widetilde{\boldsymbol{g}}_r$ линейно независимы. Дополним их некоторыми векторами $\widetilde{\boldsymbol{g}}_{r+1}, \ldots, \widetilde{\boldsymbol{g}}_m$ до базиса $\widetilde{\boldsymbol{g}}_1, \widetilde{\boldsymbol{g}}_2, \ldots, \widetilde{\boldsymbol{g}}_m$ в $\boldsymbol{S}$.

Тогда матрица, отвечающая тому же оператору $\boldsymbol{A}$ в новых базисах $\widetilde{\boldsymbol{e}}_1, \ldots, \widetilde{\boldsymbol{e}}_n$; $\widetilde{\boldsymbol{g}}_1, \ldots, \widetilde{\boldsymbol{g}}_m$, согласно (35) и (36) будет иметь вид

$$I_r = \left\| \begin{array}{cccc} \overbrace{1 & 0 \ldots 0}^{r} & 0 \ldots & 0 \\ 0 & 1 \ldots 0 & 0 \ldots & 0 \\ 0 & 0 \ldots 1 & 0 \ldots & 0 \\ 0 & 0 \ldots 0 & 0 \ldots & 0 \\ \multicolumn{4}{c}{\ldots\ldots\ldots\ldots} \\ 0 & 0 \ldots 0 & 0 \ldots & 0 \end{array} \right\|. \tag{37}$$

В матрице I_r вдоль главной диагонали сверху вниз идут r единиц; все остальные элементы матрицы I_r равны нулю. Так как матрицы A и I_r соответствуют одному и тому же оператору $\boldsymbol{A}$, то они эквивалентны между собой. По доказанному эквивалентные матрицы имеют один и тот же ранг. Поэтому ранг исходной матрицы A равен r.

Мы показали, что произвольная прямоугольная матрица ранга r эквивалентна «канонической» матрице I_r. Но матрица I_r полностью определяется заданием размеров $m \times n$ и числа r. Поэтому все прямоугольные матрицы данных размеров $m \times n$ и данного ранга r эквивалентны одной и той же матрице I_r и, следовательно, эквивалентны между собой. Теорема доказана.

3. Пусть дан линейный оператор $\boldsymbol{A}$, отображающий n-мерное пространство $\boldsymbol{R}$ в m-мерное $\boldsymbol{S}$. Совокупность векторов вида $\boldsymbol{Ax}$, где $\boldsymbol{x} \in \boldsymbol{R}$, образует векторное пространство [2]). Это пространство мы будем обозначать

---

[1]) Этого можно достигнуть надлежащей нумерацией базисных векторов $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$.

[2]) Совокупность векторов вида $\boldsymbol{Ax}$ ($\boldsymbol{x} \in \boldsymbol{R}$) удовлетворяет постулатам 1°—7° § 1, поскольку сумма двух векторов вида $\boldsymbol{Ax}$ ($\boldsymbol{x} \in \boldsymbol{R}$) и произведение такого вектора на число снова дают вектор такого вида.

через $\boldsymbol{AR}$; оно составляет часть пространства $\boldsymbol{S}$ или, как говорят, является *подпространством* в пространстве $\boldsymbol{S}$.

Наряду с подпространством $\boldsymbol{AR}$ в $\boldsymbol{S}$ рассмотрим совокупность всех векторов $x \in \boldsymbol{R}$, удовлетворяющих уравнению

$$\boldsymbol{A}x = 0. \tag{38}$$

Эти векторы также образуют подпространство в $\boldsymbol{R}$; это подпространство мы обозначим через $\boldsymbol{N_A}$.

Определение 9. Если линейный оператор $\boldsymbol{A}$ отображает $\boldsymbol{R}$ в $\boldsymbol{S}$, то число измерений r пространства $\boldsymbol{AR}$ называется *рангом* оператора $\boldsymbol{A}$ [1]), а число измерений d пространства N_A, состоящего из всех векторов $x \in \boldsymbol{R}$, удовлетворяющих условию (38), — *дефектом* оператора $\boldsymbol{A}$.

Среди всех эквивалентных прямоугольных матриц, задающих данный оператор $\boldsymbol{A}$ в различных базисах, имеется каноническая матрица I_r [см. (37)]. Обозначим через $\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n$ и $\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_m$ соответствующие ей базисы в $\boldsymbol{R}$ и $\boldsymbol{S}$. Тогда

$$\boldsymbol{A}\tilde{e}_1 = \tilde{g}_1, \ \ldots, \ \boldsymbol{A}\tilde{e}_r = \tilde{g}_r, \quad \boldsymbol{A}\tilde{e}_{r+1} = \ldots = \boldsymbol{A}\tilde{e}_n = 0.$$

Из определения $\boldsymbol{AR}$ и $\boldsymbol{N_A}$ следует, что векторы $\tilde{g}_1, \ldots, \tilde{g}_r$ образуют базис в $\boldsymbol{AR}$, а векторы $\tilde{e}_{r+1}, \ldots, \tilde{e}_n$ составляют базис в N_A. Отсюда вытекает, что r — ранг оператора $\boldsymbol{A}$ и

$$d = n - r. \tag{39}$$

Если A — произвольная матрица, соответствующая оператору $\boldsymbol{A}$, то она эквивалентна I_r и, следовательно, имеет тот же ранг r. Таким образом, *ранг оператора $\boldsymbol{A}$ совпадает с рангом прямоугольной матрицы A*

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix},$$

определяющей оператор $\boldsymbol{A}$ в некоторых базисах $e_1, e_2, \ldots, e_n \in \boldsymbol{R}$ и $g_1, g_2, \ldots, g_m \in \boldsymbol{S}$.

В столбцах матрицы A стоят координаты векторов $\boldsymbol{A}e_1, \boldsymbol{A}e_2, \ldots, \boldsymbol{A}e_n$. Так как из $x = \sum_{i=1}^{n} x_i e_i$ следует $\boldsymbol{A}x = \sum_{i=1}^{n} x_i \boldsymbol{A}e_i$, то ранг оператора $\boldsymbol{A}$, т. е. число измерений $\boldsymbol{RA}$, равняется максимальному числу линейно независимых векторов среди $\boldsymbol{A}e_1, \boldsymbol{A}e_2, \ldots, \boldsymbol{A}e_n$. Таким образом, *ранг матрицы совпадает с числом линейно независимых столбцов матрицы.* Поскольку при транспонировании строки матрицы делаются столбцами, а ранг не меняется, то *число линейно независимых строк матрицы также равно рангу матрицы* [2]).

---

[1]) Число измерений пространства $\boldsymbol{AR}$ всегда $\leqslant$ числа измерений пространства $\boldsymbol{R}$, т. е. $r \leqslant n$. Это следует из того, что равенство $x = \sum_{i=1}^{n} x_i e_i$ ($e_1, e_2, \ldots, e_n$ — базис в $\boldsymbol{R}$) влечет равенство $\boldsymbol{A}x = \sum_{i=1}^{n} x_i \boldsymbol{A}e_i$.

[2]) К этим же выводам мы пришли в § 1 из других соображений (см. стр. 68).

4. Пусть даны два линейных оператора $\boldsymbol{A}$, $\boldsymbol{B}$ и их произведение $\boldsymbol{C}=\boldsymbol{AB}$.

Пусть оператор $\boldsymbol{B}$ отображает $\boldsymbol{R}$ в $\boldsymbol{S}$, а оператор $\boldsymbol{A}$ отображает $\boldsymbol{S}$ в $\boldsymbol{T}$. Тогда оператор $\boldsymbol{C}$ отображает $\boldsymbol{R}$ в $\boldsymbol{T}$:

$$\boldsymbol{T} \xleftarrow{A} \boldsymbol{S} \xleftarrow{B} \boldsymbol{R}, \quad \boldsymbol{T} \xleftarrow{C} \boldsymbol{R}.$$

Введем матрицы A, B, C, соответствующие операторам $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ при некотором выборе базисов в $\boldsymbol{R}$, $\boldsymbol{S}$ и $\boldsymbol{T}$. Тогда операторному равенству $\boldsymbol{C}=\boldsymbol{AB}$ будет соответствовать матричное равенство $C=AB$.

Обозначим через r_A, r_B, r_C ранги операторов $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, или, что то же, ранги матриц A, B, C. Эти числа определяют число измерений подпространств $\boldsymbol{AS}$, $\boldsymbol{BR}$, $\boldsymbol{A}(\boldsymbol{BR})$. Поскольку $\boldsymbol{BR}\subset\boldsymbol{S}$, то $\boldsymbol{A}(\boldsymbol{BR})\subset\boldsymbol{AS}$ [1]). Кроме того, число измерений $\boldsymbol{A}(\boldsymbol{BR})$ не может превосходить числа измерений $\boldsymbol{BR}$ [2]). Поэтому

$$r_C \leqslant r_A, \quad r_C \leqslant r_B.$$

Эти неравенства были нами получены в главе I, § 2 из формулы для миноров произведения двух матриц.

Рассмотрим оператор $\boldsymbol{A}$ как оператор, отображающий $\boldsymbol{BR}$ в $\boldsymbol{T}$. Тогда ранг этого оператора будет равен числу измерений пространства $\boldsymbol{A}(\boldsymbol{BR})$, т. е. r_C. Поэтому, применяя формулу (39), получим:

$$r_C = r_B - d_1, \tag{40}$$

где d_1 — максимальное число линейно независимых векторов из $\boldsymbol{BR}$, удовлетворяющих уравнению

$$\boldsymbol{Ax}=0. \tag{41}$$

Но все решения этого уравнения, принадлежащие $\boldsymbol{S}$, образуют подпространство d измерений, где

$$d = n - r_A \tag{42}$$

— дефект оператора $\boldsymbol{A}$, отображающего $\boldsymbol{S}$ в $\boldsymbol{T}$. Поскольку $\boldsymbol{BR}\subset\boldsymbol{S}$, то

$$d_1 \leqslant d. \tag{43}$$

Из (40), (42) и (43) находим:

$$r_A + r_B - n \leqslant r_C.$$

Таким образом, мы получили следующие неравенства Сильвестра для ранга произведения двух прямоугольных матриц A и B с размерами $m\times n$ и $n\times q$:

$$r_A + r_B - n \leqslant r_{AB} \leqslant r_A,\ r_B. \tag{44}$$

Если матричное уравнение $AXB=C$, где размеры прямоугольных матриц A, X, B — $m\times n$, $n\times p$, $p\times q$, имеет решение X (см. стр. 23), то из неравенств Сильвестра легко следует:

$$r_c \leqslant r_x \leqslant n + p - r_A - r_B.$$

Можно доказать, что если уравнение $AXB=C$ имеет какое-либо решение, то оно имеет решение любого ранга r, заключенного между числами r_C и $n+p-r_A-r_B$.

---

[1]) $\boldsymbol{R}\subset\boldsymbol{S}$ означает, что совокупность $\boldsymbol{R}$ составляет часть совокупности $\boldsymbol{S}$.

[2]) См. сноску [2]) на стр. 22.

§ 6. Линейные операторы, отображающие n-мерное пространство само в себя

1. Линейный оператор, отображающий n-мерное векторное пространство $\boldsymbol{R}$ само в себя (в данном случае $\boldsymbol{R} \equiv \boldsymbol{S}$, $n = m$), мы будем просто называть *линейным оператором в* $\boldsymbol{R}$.

Сумма двух линейных операторов в $\boldsymbol{R}$, а также произведение такого оператора на число — снова линейные операторы в $\boldsymbol{R}$. Умножение двух таких линейных операторов всегда выполнимо, и произведение их есть снова линейный оператор в $\boldsymbol{R}$. Таким образом, линейные операторы в $\boldsymbol{R}$ образуют кольцо [1]). В этом кольце имеется единичный оператор, т. е. оператор $\boldsymbol{E}$, для которого

$$\boldsymbol{E}x = x \quad (x \in \boldsymbol{R}).$$

При этом для произвольного оператора $\boldsymbol{A}$ в $\boldsymbol{R}$

$$\boldsymbol{E}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{E} = \boldsymbol{A}.$$

Если $\boldsymbol{A}$ — линейный оператор в $\boldsymbol{R}$, то имеет смысл $\boldsymbol{A}^2 = \boldsymbol{A}\boldsymbol{A}$, $\boldsymbol{A}^3 = \boldsymbol{A}\boldsymbol{A}\boldsymbol{A}$, ... и вообще $\boldsymbol{A}^m = \underbrace{\boldsymbol{A}\boldsymbol{A} \ldots \boldsymbol{A}}_{m \text{ раз}}$. Кроме того, полагаем $\boldsymbol{A}^0 = \boldsymbol{E}$. Тогда, как легко видеть, при любых целых неотрицательных p и q

$$\boldsymbol{A}^p \boldsymbol{A}^q = \boldsymbol{A}^{p+q}.$$

Пусть $f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + \ldots + \alpha_{m-1} t + \alpha_m$ — многочлен относительно скалярного аргумента t с коэффициентами из поля K. Тогда полагаем:

$$f(\boldsymbol{A}) = \alpha_0 \boldsymbol{A}^m + \alpha_1 \boldsymbol{A}^{m-1} + \ldots + \alpha_{m-1}\boldsymbol{A} + \alpha_m \boldsymbol{E}.$$

При этом $f(\boldsymbol{A})\, g(\boldsymbol{A}) = g(\boldsymbol{A})\, f(\boldsymbol{A})$ для любых двух многочленов $f(t)$ и $g(t)$.

Пусть

$$\boldsymbol{y} = \boldsymbol{A}x \quad (x,\ \boldsymbol{y} \in \boldsymbol{R}). \tag{45}$$

Обозначим через $x_1, x_2, \ldots, x_n$ координаты вектора x в произвольном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, а через $y_1, y_2, \ldots, y_n$ — координаты вектора $\boldsymbol{y}$ в том же базисе. Тогда

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, \ldots, n). \tag{46}$$

В базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ линейному оператору $\boldsymbol{A}$ отвечает квадратная матрица $A = \| a_{ik} \|_1^n$, [2]). Напомним читателю (см. стр. 71), что в k-м столбце этой матрицы стоят координаты вектора $\boldsymbol{A}\boldsymbol{e}_k$ $(k = 1, 2, \ldots, n)$, т. е.

$$\boldsymbol{A}\boldsymbol{e}_k = \sum_{i=1}^{n} a_{ik} \boldsymbol{e}_i \quad (k = 1, \ldots, n). \tag{47}$$

Имея координатные столбцы $x = (x_1, x_2, \ldots, x_n)$ и $y = (y_1, y_2, \ldots, y_n)$, мы преобразование (46) можем записать в матричной форме

$$y = Ax. \tag{48}$$

---

[1]) Это кольцо является алгеброй. Ср. гл. I, стр. 28.

[2]) См. стр. 70—71. В данном случае пространства $\boldsymbol{R}$ и $\boldsymbol{S}$ совпадают; точно также отождествлены базисы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ и $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m$ в этих пространствах.

Сумме и произведению двух операторов $\boldsymbol{A}$ и $\boldsymbol{B}$ отвечают сумма и произведение соответствующих квадратных матриц $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$. Произведению $\alpha \boldsymbol{A}$ соответствует матрица αA. Единичному оператору $\boldsymbol{E}$ отвечает квадратная единичная матрица $E = \| \delta_{ik} \|_1^n$. Таким образом, выбор базиса устанавливает *изоморфное соответствие между кольцом линейных операторов в* $\boldsymbol{R}$ *и кольцом квадратных матриц* n*-го порядка с элементами из* K. При этом соответствии многочлену $f(\boldsymbol{A})$ соответствует матрица $f(A)$.

2. Рассмотрим наряду с базисом $e_1, e_2, \ldots, e_n$ другой базис $\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n$ в $\boldsymbol{R}$. Тогда аналогично (48)

$$\widetilde{y} = \widetilde{A}\,\widetilde{x}, \tag{49}$$

где $\widetilde{x}$, $\widetilde{y}$ — столбцевые матрицы, составленные из координат векторов x, y в базисе $\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n$, а $\widetilde{A} = \| \widetilde{a}_{ik} \|_1^n$ — квадратная матрица, соответствующая оператору $\boldsymbol{A}$ в этом базисе. Запишем в матричной форме формулы преобразования координат

$$x = T\widetilde{x}, \quad y = T\widetilde{y}. \tag{50}$$

Тогда из (48) и (50) находим:

$$\widetilde{y} = T^{-1}AT\widetilde{x},$$

что в сопоставлении с (49) дает:

$$\widetilde{A} = T^{-1}AT. \tag{51}$$

Формула (51) представляет собой специальный частный случай формулы (31) на стр. 75 (в данном случае $P = T^{-1}$, $Q = T$).

Определение 10. Две матрицы A и B, связанные соотношением

$$B = T^{-1}AT, \tag{51'}$$

где T — некоторая неособенная матрица, называются *подобными*[1]).

Таким образом, мы показали, что *две матрицы, соответствующие одному и тому же линейному оператору в* $\boldsymbol{R}$ *при различных базисах, подобны между собой*, причем матрица T, связывающая эти матрицы, совпадает с матрицей преобразования координат при переходе от первого базиса ко второму (см. (50)).

Другими словами, линейному оператору в $\boldsymbol{R}$ отвечает целый класс подобных между собой матриц; эти матрицы представляют данный оператор в различных базисах.

Изучая свойства линейного оператора в $\boldsymbol{R}$, мы тем самым изучаем свойства матриц, присущие одновременно всему классу подобных матриц, т. е. изучаем свойства матриц, остающиеся неизменными инвариантными при переходе от данной матрицы к матрице, ей подобной.

Заметим еще, что две подобные матрицы имеют всегда равные определители. Действительно, из (51') следует, что

$$|B| = |T|^{-1}|A||T| = |A|. \tag{52}$$

Равенство $|B| = |A|$ является необходимым, но не достаточным условием для подобия матриц A и B

---

[1]) Матрицу T всегда можно выбрать так, чтобы ее элементы принадлежали основному числовому полю K, которому принадлежат элементы матриц A и B. Легко проверяются три свойства подобия матриц: рефлексивность (матрица A всегда подобна самой себе), симметричность (если A подобна B, то и B подобна A) и транзитивность (если A подобна B, B подобна C, то A подобна C).

В главе VI будет установлен критерий подобия двух матриц, т. е. будут даны необходимые и достаточные условия для того, чтобы две квадратные матрицы n-го порядка были подобны между собой.

Согласно равенству (52) мы можем под определителем линейного оператора A в R ($|A|$) понимать определитель любой матрицы, соответствующей данному оператору.

Если $|A|=0$ ($\neq 0$), то оператор A называется *особенным* (соответственно *неособенным*). Согласно этому определению в любом базисе особенному (неособенному) оператору отвечает особенная (соответственно неособенная) матрица. Для особенного оператора:

1) всегда существует вектор $x\neq 0$ такой, что $Ax=0$,

2) AR составляет правильную часть R. Для неособенного оператора:

1) из $Ax=0$ следует $x=0$;

2) $AR\equiv R$, т. е. векторы вида Ax ($x\in R$) заполняют все пространство R.

Другими словами, линейный оператор в R является особенным или неособенным в зависимости от того, больше или равен нулю его дефект.

3. Если A — неособенный оператор, то в равенстве $y=Ax$ задание вектора $y\in R$ однозначно определяет вектор $x\in R$. Действительно, существование вектора x следует из того, что векторы вида Ax ($x\in R$) заполняют все пространство R. С другой стороны, из равенств $y=Ax'$ и $y=Ax''$ ($x', x''\in R$) следует: $A(x'-x'')=Ax'-Ax''=0$ и отсюда: $x'-x''=0$, т. е. $x'=x''$. Поэтому, исходя из равенства $y=Ax$, можно определить обратный оператор A^{-1} равенством $x=A^{-1}y$. Легко видеть, что обратный оператор A^{-1} для линейного оператора A в R также является линейным оператором в R; при этом

$$AA^{-1}=A^{-1}A=E,$$

где E — единичный оператор. Если в некотором базисе неособенному оператору A отвечает неособенная матрица A, то в этом базисе обратному оператору A^{-1}, соответствует матрица A^{-1}.

Рассмотрим некоторые частные типы линейных операторов в R.

1°. Оператор J в R называется *инволютивным*, если $J^2=E$. Инволютивный оператор неособенный и для него $J^{-1}=J$. Инволютивному оператору в любом базисе соответствует *инволютивная матрица* J, т. е. матрица J, для которой $J^2=E$.

2°. Оператор P в R называется *проекционным*, если $P^2=P$. Пусть дано произвольное расщепление пространства R на два подпространства S и T: $R=S+T$. Тогда для любого вектора $x\in R$ имеет место разложение $x=x_S+x_T$, где $x_S\in S$, $x_T\in T$. Вектор x_S называется *проекцией вектора x на подпространство S параллельно подпространству T* [1]). Рассмотрим оператор P, осуществляющий проектирование пространства R на подпространство S параллельно подпространству T, т. е. оператор в R, определяемый равенством $Px=x_S$ для любого вектора $x\in R$. Очевидно, этот оператор является линейным, но он является и проективным, так как $Px=x_S$, $P^2x=Px_S=x_S$ и, следовательно, $(P^2-P)x=x_S-x_S=0$, т. е. $P^2=P$.

Легко проверяется и обратное утверждение. Произвольный проекционный оператор P в R осуществляет проектирование всего пространства R на подпространство $S=PR$ параллельно подпространству $T=(E-P)R$.

Любая натуральная степень проекционного оператора является проекционным оператором. Если P — проекционный оператор, то и $E-P$ — проекционный оператор, так как $(E-P)^2=E-2P+P^2=E-P$.

Квадратная матрица P называется *проекционной*, если $P^2=P$. Очевидно, в произвольном базисе проекционному оператору соответствует проекционная матрица.

[1]) Аналогично, вектор x_T — проекция вектора x на подпространство T параллельно подпространству S.

§ 7. Характеристические числа и собственные векторы линейного оператора

При исследовании структуры линейного оператора $\boldsymbol{A}$ в $\boldsymbol{R}$ большую роль играют векторы x, для которых

$$\boldsymbol{A}x = \lambda x \qquad (\lambda \in K, \quad x \neq 0). \tag{53}$$

Такие векторы называются *собственными векторами*, а соответствующие им числа λ — *характеристическими* или *собственными числами* оператора $\boldsymbol{A}$ (матрицы A).

Для нахождения характеристических чисел и собственных векторов оператора $\boldsymbol{A}$ выберем произвольно базис $e_1, e_2, \ldots, e_n$ в $\boldsymbol{R}$. Пусть $x = \sum_{i=k}^{n} x_i e_i$ и $A = \| a_{ik} \|_1^n$ — матрица, отвечающая оператору $\boldsymbol{A}$ в базисе $e_1, e_2, \ldots, e_n$. Тогда, приравнивая между собой соответственные координаты векторов, стоящих в левой и правой частях равенства (53), получим систему скалярных уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= \lambda x_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= \lambda x_n, \end{aligned}\right\} \tag{54}$$

которую можно записать и так:

$$\left.\begin{aligned} (a_{11} - \lambda)\, x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22} - \lambda)\, x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)\, x_n &= 0. \end{aligned}\right\} \tag{55}$$

Так как искомый вектор не должен быть равен нулю, то среди его координат $x_1, x_2, \ldots, x_n$ по крайней мере одна координата должна быть отлична от нуля.

Для того чтобы система линейных однородных уравнений (55) имела ненулевое решение, необходимо и достаточно, чтобы определитель этой системы был равен нулю:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{56}$$

Уравнение (56) представляет собой алгебраическое уравнение n-й степени относительно λ. Коэффициенты этого уравнения принадлежат тому же числовому полю, что и элементы матрицы $A = \| a_{ik} \|_1^n$, т. е. полю K.

Уравнение (56) часто встречается в различных проблемах геометрии, механики, астрономии, физики и носит название *характеристического уравнения* или *векового*[1]) *уравнения* матрицы $A = \| a_{ik} \|_1^n$ (левую часть этого уравнения называют *характеристическим многочленом*).

[1]) Такое название связано с тем, что это уравнение встречается при исследовании вековых возмущений планет.

Таким образом, каждое характеристическое число λ линейного оператора $\boldsymbol{A}$ является корнем характеристического уравнения (56). И наоборот, если некоторое число λ является корнем уравнения (56), то при этом значении λ система (55) и, следовательно, (54) имеет ненулевое решение $x_1, x_2, \ldots, x_n$, т. е. этому числу λ отвечает собственный вектор $\boldsymbol{x} = \sum x_i \boldsymbol{e}_i$ оператора $\boldsymbol{A}$.

Из сказанного следует, что любой линейный оператор $\boldsymbol{A}$ в $\boldsymbol{R}$ имеет не более чем n различных характеристических чисел.

Если K есть поле всех комплексных чисел, то любой линейный оператор в $\boldsymbol{R}$ всегда имеет по крайней мере один собственный вектор в $\boldsymbol{R}$ и соответствующее этому собственному вектору характеристическое число λ [1]). Это следует из основной теоремы алгебры, согласно которой алгебраическое уравнение (56) в поле комплексных чисел всегда имеет по крайней мере один корень.

Напишем уравнение (56) в развернутом виде

$$|A-\lambda E| \equiv (-\lambda)^n + S_1(-\lambda)^{n-1} + S_2(-\lambda)^{n-2} + \ldots + S_{n-1}(-\lambda) + S_n = 0; \qquad (57)$$

здесь, как нетрудно видеть,

$$S_1 = \sum_{i=1}^{n} a_{ii}, \quad S_2 = \sum_{1 \leqslant i < k \leqslant n} A \begin{pmatrix} i & k \\ i & k \end{pmatrix}, \ldots \qquad (58)$$

и вообще S_p равно сумме главных миноров p-го порядка матрицы $A = \| a_{ik} \|_1^n$ $(p = 1, 2, \ldots, n)$ [2]). В частности, $S_n = |A|$.

Обозначим через $\widetilde{A}$ матрицу, соответствующую тому же оператору $\boldsymbol{A}$ в другом базисе. Матрица $\widetilde{A}$ подобна матрице A:

$$\widetilde{A} = T^{-1}AT.$$

Отсюда

$$\widetilde{A} - \lambda E = T^{-1}(A - \lambda E)\, T$$

---

[1]) Это положение справедливо и в более общем случае, когда K — произвольное алгебраически замкнутое поле, т. е. такое поле, которому принадлежат корни всех алгебраических уравнений с коэффициентами из данного поля.

[2]) Степень $(-\lambda)^{n-p}$ имеется только в тех членах характеристического определителя (56), которые содержат какие-либо $n-p$ диагональных элементов

$$a_{j_1 j_1} - \lambda, \ a_{j_2 j_2} - \lambda, \ \ldots, \ a_{j_{n-p} j_{n-p}} - \lambda.$$

Произведение этих диагональных элементов входит в состав определителя (56) со множителем, равным главному минору

$$A \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix},$$

где индексы $i_1, i_2, \ldots, i_p$ вместе с индексами $j_1, j_2, \ldots, j_{n-p}$ образуют полную систему индексов $1, 2, \ldots, n$:

$$|A-\lambda E| = (a_{j_1 j_1} - \lambda)(a_{j_2 j_2} - \lambda) \ldots (a_{j_{n-p} j_{n-p}} - \lambda)\, A \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} + (*).$$

Здесь $(-\lambda)^{n-p}$ умножается на $A \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix}$. Перебирая всевозможные сочетания $j_1, j_2, \ldots, j_{n-p}$ по np и—з индексов $1, 2, \ldots, n$, мы получим в качестве коэффициента S_p при $(-\lambda)^{n-p}$ сумму всех главных миноров p-го порядка матрицы $\boldsymbol{A}$.

и, следовательно,

$$|\widetilde{A}-\lambda E|=|A-\lambda E|. \tag{59}$$

Таким образом, подобные матрицы A и $\widetilde{A}$ имеют один и тот же характеристический многочлен. Этот многочлен иногда называют характеристическим многочленом оператора $\boldsymbol{A}$ и обозначают через $|\boldsymbol{A}-\lambda\boldsymbol{E}|$.

Если x, y, z, ... — собственные векторы оператора $\boldsymbol{A}$, соответствующие одному и тому же характеристическому числу λ, а α, β, γ, ... — произвольные числа из K, то вектор $\alpha x+\beta y+\gamma z+\ldots$ либо равен нулю, либо также является собственным вектором оператора $\boldsymbol{A}$ при том же числе λ. Действительно, из

$$\boldsymbol{A}x=\lambda x,\ \boldsymbol{A}y=\lambda y,\ \boldsymbol{A}z=\lambda z,\ \ldots$$

следует:

$$\boldsymbol{A}(\alpha x+\beta y+\gamma z+\ldots)=\lambda(\alpha x+\beta y+\gamma z+\ldots).$$

Поэтому линейно независимые собственные векторы, отвечающие одному и тому же характеристическому числу λ, образуют базис некоторого «собственного» подпространства, каждый вектор которого есть собственный вектор при том же λ. В частности, каждый собственный вектор порождает одномерное собственное подпространство, «собственное направление».

Однако, если собственные векторы оператора $\boldsymbol{A}$ соответствуют различным характеристическим числам, то линейная комбинация этих собственных векторов, вообще говоря, не будет собственным вектором оператора $\boldsymbol{A}$.

Значение собственных векторов и характеристических чисел при исследовании линейных операторов будет проиллюстрировано в следующем параграфе на примере операторов простой структуры.

§ 8. Линейные операторы простой структуры

Начнем со следующей леммы.

Лемма. *Собственные векторы, соответствующие попарно различным характеристическим числам, всегда линейно независимы.*

Доказательство. Пусть

$$\boldsymbol{A}x_i=\lambda_i x_i \qquad (x_i\neq 0;\ \lambda_i\neq\lambda_k \text{ при } i\neq k;\ i,\ k=1,\ 2,\ \ldots,\ m) \tag{60}$$

и пусть

$$\sum_{i=1}^{m} c_i x_i=0. \tag{61}$$

Применяя к обеим частям этого равенства оператор $\boldsymbol{A}$, получим:

$$\sum_{i=1}^{m} c_i\lambda_i x_i=0. \tag{62}$$

Умножим обе части равенства (61) на λ_1 и вычтем почленно (61) из (62). Тогда получим:

$$\sum_{i=2}^{m} c_i(\lambda_i-\lambda_1)x_i=0. \tag{63}$$

Можно сказать, что равенство (63) было получено из (61) путем почленного применения оператора $\boldsymbol{A}-\lambda_1\boldsymbol{E}$. Применяя к (63) почленно операторы $\boldsymbol{A}-\lambda_2\boldsymbol{E}$, ..., $\boldsymbol{A}-\lambda_{m-1}\boldsymbol{E}$, мы придем к следующему равенству:

$$c_m(\lambda_m-\lambda_{m-1})(\lambda_m-\lambda_{m-2})\ldots(\lambda_m-\lambda_1)x_m=0,$$

откуда $c_m = 0$. Так как в (61) любое слагаемое может быть поставлено на последнее место, то в (61)

$$c_1 = c_2 = \ldots = c_m = 0,$$

т. е. между векторами x_1, x_2, ..., x_m нет линейной зависимости. Лемма доказана.

Если характеристическое уравнение оператора имеет n различных корней и эти корни принадлежат полю K, то на основании леммы собственные векторы, соответствующие этим корням, линейно независимы.

Определение 11. Линейный оператор $\boldsymbol{A}$ в $\boldsymbol{R}$ называется *оператором простой структуры*, если $\boldsymbol{A}$ имеет в $\boldsymbol{R}$ n линейно независимых собственных векторов, где n — число измерений.

Таким образом, линейный оператор в $\boldsymbol{R}$ имеет простую структуру, если все корни характеристического уравнения различны между собой и принадлежат полю K. Однако это условие не является необходимым. Существуют линейные операторы простой структуры, у которых характеристический многочлен имеет кратные корни.

Рассмотрим произвольный линейный оператор $\boldsymbol{A}$ простой структуры. Обозначим через $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$ базис в $\boldsymbol{R}$, состоящий из собственных векторов оператора, т. е.

$$\boldsymbol{A}\boldsymbol{g}_k = \lambda_k \boldsymbol{g}_k \qquad (k = 1, 2, \ldots, n).$$

Если

$$\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{g}_k, \text{ то } \boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{A}\boldsymbol{g}_k = \sum_{k=1}^{n} \lambda_k x_k \boldsymbol{g}_k.$$

Другими словами, воздействие оператора $\boldsymbol{A}$ простой структуры на вектор $\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{g}_k$ может быть описано следующим образом:

В n-мерном пространстве $\boldsymbol{R}$ существует n линейно независимых «направлений», вдоль которых оператор простой структуры $\boldsymbol{A}$ осуществляет «растяжение» с коэффициентами $\lambda_1, \lambda_2, \ldots, \lambda_n$. Произвольный вектор $\boldsymbol{x}$ может быть разложен на компоненты, идущие вдоль этих собственных направлений. Эти компоненты подвергаются соответствующим «растяжениям», после чего они в сумме дают вектор $\boldsymbol{A}\boldsymbol{x}$.

Нетрудно видеть, что оператору $\boldsymbol{A}$ в «собственном» базисе $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$ соответствует диагональная матрица

$$\widetilde{A} = \| \lambda_i \delta_{ik} \|_1^n.$$

Если мы через A обозначим матрицу, отвечающую оператору $\boldsymbol{A}$ в произвольном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, то

$$A = T \| \lambda_i \delta_{ik} \|_1^n T^{-1}. \tag{64}$$

Матрицу, подобную диагональной, будем называть *матрицей простой структуры*. Таким образом, оператору простой структуры в любом базисе отвечает матрица простой структуры и наоборот.

Матрица T в равенстве (64) осуществляет переход от базиса $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ к базису $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$. В k-м столбце матрицы T стоят координаты (в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$) собственного вектора $\boldsymbol{g}_k$, соответствующего характеристическому числу λ_k матрицы A ($k = 1, 2, \ldots, n$). Матрица T называется *фундаментальной* матрицей для матрицы A.

Равенство (64) перепишем так:

$$A = TLT^{-1} \qquad (L = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}). \tag{64'}$$

Переходя к p-м ассоциированным матрицам ($1 \leqslant p \leqslant n$), получим (см. гл. I, § 4):

$$\mathfrak{A}_p = \mathfrak{T}_p \Omega_p \mathfrak{T}_p^{-1}; \tag{65}$$

Ω_p — диагональная матрица N-го порядка ($N = C_n^p$), у которой на главной диагонали стоят всевозможные произведения по p из $\lambda_1, \lambda_2, \ldots, \lambda_n$. Сопоставление (65) с (64′) дает нам теорему:

Теорема 3. *Если матрица* $A = \| a_{ik} \|_1^n$ *имеет простую структуру, то при любом* $p \leqslant n$ *ассоциированная матрица* $\mathfrak{A}_p$ *также имеет простую структуру; при этом характеристическими числами матрицы* $\mathfrak{A}_p$ *являются всевозможные произведения по* p $\lambda_{i_1}\lambda_{i_2} \ldots \lambda_{i_p}$ $(1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n)$ *из характеристических чисел* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *матрицы* A, *а фундаментальной матрицей матрицы* $\mathfrak{A}_p$ *является ассоциированная* $\mathfrak{T}_p$ *для фундаментальной матрицы* T *матрицы* A.

Следствие. *Если характеристическому числу* λ_k *матрицы простой структуры* $A = \| a_{ik} \|_1^n$ *отвечает собственный вектор с координатами* $t_{1k}, t_{2k}, \ldots, t_{nk}$ $(k = 1, 2, \ldots, n)$ *и* $T = \| t_{ik} \|_1^n$, *то характеристическому числу* $\lambda_{k_1}\lambda_{k_2} \ldots \lambda_{k_p}$ $(1 \leqslant k_1 < k_2 < \ldots < k_p \leqslant n)$ *матрицы* $\mathfrak{A}_p$ *отвечает собственный вектор с координатами*

$$T \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \qquad (1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n). \tag{66}$$

Произвольную матрицу $A = \| a_{ik} \|_1^n$ можно представить в виде предела последовательности матриц A_m $(m \to \infty)$, каждая из которых не имеет кратных характеристических чисел и потому имеет простую структуру. Характеристические числа $\lambda_1^{(m)}, \lambda_2^{(m)}, \ldots, \lambda_n^{(m)}$ матрицы A_m в пределе при $m \to \infty$ переходят в характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ матрицы A, т. е.

$$\lim_{m \to \infty} \lambda_k^{(m)} = \lambda_k \qquad (k = 1, 2, \ldots, n).$$

Отсюда

$$\lim_{m \to \infty} \lambda_{k_1}^{(m)} \lambda_{k_2}^{(m)} \ldots \lambda_{k_p}^{(m)} = \lambda_{k_1}\lambda_{k_2} \ldots \lambda_{k_p} \qquad (1 \leqslant k_1 < k_2 < \ldots < k_p \leqslant n).$$

Так как, кроме того, $\lim\limits_{m \to \infty} \mathfrak{A}_{(m)\,p} = \mathfrak{A}_p$, то из теоремы 3 вытекает

Теорема 4 (Кронекера). *Если* $\lambda_1, \lambda_2, \ldots, \lambda_n$ — *полная система характеристических чисел произвольной матрицы* A, *то полная система характеристических чисел ассоциированной матрицы* $\mathfrak{A}_p$ *состоит из всевозможных произведений по* p *из чисел* $\lambda_1, \lambda_2, \ldots, \lambda_n$ $(p = 1, 2, \ldots, n)$.

В этом параграфе мы исследовали операторы и матрицы простой структуры. Изучение структуры операторов и матриц общего типа будет проведено в главах VI и VII.

ГЛАВА IV

ХАРАКТЕРИСТИЧЕСКИЙ И МИНИМАЛЬНЫЙ МНОГОЧЛЕНЫ МАТРИЦЫ

С каждой квадратной матрицей связаны два многочлена: характеристический и минимальный. Эти многочлены играют большую роль в различных вопросах теории матриц. Так, например, понятие о функции от матрицы, которое мы введем в следующей главе, будет целиком основываться на понятии о минимальном многочлене матрицы. В этой главе рассматриваются свойства характеристического и минимального многочлена. Этому исследованию предпосылаются основные сведения о многочленах с матричными коэффициентами и о действиях над ними.

§ 1. Сложение и умножение матричных многочленов

Рассмотрим квадратную *многочленную* матрицу $A(\lambda)$, т. е. квадратную матрицу, элементами которой являются многочлены относительно λ (с коэффициентами из данного числового поля K):

$$A(\lambda) = \| a_{ik}(\lambda) \|_1^n = \| a_{ik}^{(0)}\lambda^m + a_{ik}^{(1)}\lambda^{m-1} + \ldots + a_{ik}^{(m)} \|_1^n. \tag{1}$$

Матрицу $A(\lambda)$ можно представить в виде многочлена с матричными коэффициентами, расположенного по степеням λ:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m, \tag{2}$$

где

$$A_j = \| a_{ik}^{(j)} \|_1^n \qquad (j = 0, 1, \ldots, m). \tag{3}$$

Число m называется *степенью* многочлена, если $A_0 \neq 0$. Число n называется *порядком* многочлена. Многочлен (1) будем называть *регулярным*, если $|A_0| \neq 0$.

Многочлен с матричными коэффициентами мы будем иногда называть *матричным многочленом*. В отличие от матричного многочлена обычный многочлен со скалярными коэффициентами будем называть *скалярным многочленом*.

Рассмотрим основные действия над матричными многочленами. Пусть даны два матричных многочлена одного и того же порядка $A(\lambda)$ и $B(\lambda)$. Обозначим через m наибольшую из степеней этих многочленов. Эти многочлены можно записать в виде

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m,$$
$$B(\lambda) = B_0\lambda^m + B_1\lambda^{m-1} + \ldots + B_m.$$

Тогда

$$A(\lambda) \pm B(\lambda) = (A_0 \pm B_0)\lambda^m + (A_1 \pm B_1)\lambda^{m-1} + \ldots + A_m \pm B_m,$$

т. е. *сумма (разность) двух матричных многочленов одного и того же порядка может быть представлена в виде многочлена, степень которого не превосходит наибольшей из степеней данных многочленов.*

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ степеней m и p одного и того же порядка n:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \qquad (A_0 \neq 0),$$
$$B(\lambda) = B_0\lambda^p + B_1\lambda^{p-1} + \ldots + B_p \qquad (B_0 \neq 0).$$

Тогда

$$A(\lambda)B(\lambda) = A_0B_0\lambda^{m+p} + (A_0B_1 + A_1B_0)\lambda^{m+p-1} + \ldots + A_mB_p. \tag{4}$$

Если бы мы перемножили $B(\lambda)$ на $A(\lambda)$ (т. е. изменили бы порядок сомножителей), то мы получили бы, вообще говоря, другой многочлен.

Умножение матричных многочленов обладает еще одним специфичным свойством. В отличие от произведения скалярных многочленов произведение матричных многочленов (4) может иметь степень, меньшую $m+p$, т. е. меньшую суммы степеней сомножителей. Действительно, в (4) произведение матриц A_0B_0 может равняться нулю при $A_0 \neq 0$ и $B_0 \neq 0$. Однако, если хотя бы одна из матриц A_0 и B_0 неособенная, то из $A_0 \neq 0$ и $B_0 \neq 0$ следует: $A_0B_0 \neq 0$. Таким образом, *произведение двух матричных многочленов равно многочлену, степень которого меньше или равна сумме степеней сомножителей. Если хотя бы один из двух сомножителей регулярный многочлен, то в этом случае степень произведения всегда равна сумме степеней сомножителей.*

Матричный многочлен n-го порядка $A(\lambda)$ можно записать двояко:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \tag{5}$$

и

$$A(\lambda) = \lambda^m A_0 + \lambda^{m-1}A_1 + \ldots + A_m. \tag{5'}$$

Обе записи при скалярном λ дают один и тот же результат. Однако если мы пожелаем вместо скалярного аргумента λ подставить квадратную матрицу n-го порядка Λ, то результаты подстановок в (5) и (5') будут, вообще говоря, различны, так как степени матрицы Λ могут не быть перестановочными с матричными коэффициентами $A_0, A_1, \ldots, A_m$.

Положим

$$A(\Lambda) = A_0\Lambda^m + A_1\Lambda^{m-1} + \ldots + A_m \tag{6}$$

и

$$\widehat{A}(\Lambda) = \Lambda^m A_0 + \Lambda^{m-1}A_1 + \ldots + A_m \tag{6'}$$

и будем называть $A(\Lambda)$ правым, а $\widehat{A}(\Lambda)$ левым значением матричного многочлена $A(\lambda)$ при подстановке вместо λ матрицы Λ [1]).

Рассмотрим снова два матричных многочлена

$$A(\lambda) = \sum_{i=0}^{m} A_{m-i}\lambda^i, \quad B(\lambda) = \sum_{k=0}^{p} B_{p-k}\lambda^k$$

[1]) В «правом» значении $A(\lambda)$ степени матрицы Λ стоят справа от коэффициентов, а в «левом» — слева.

и их произведение

$$P(\lambda)=\sum_{i=0}^{m}\sum_{k=0}^{p}A_{m-i}\lambda^{i}B_{p-k}\lambda^{k}=\sum_{i=0}^{m}\sum_{k=0}^{p}A_{m-i}B_{p-k}\lambda^{i+k}=\sum_{j=0}^{m+p}\Big(\sum_{i+k=j}A_{m-i}B_{p-k}\Big)\lambda^{j}$$

и (7′)

$$P(\lambda)=\sum_{i=0}^{m}\sum_{k=0}^{p}\lambda^{i}A_{m-i}\lambda^{k}B_{p-k}=\sum_{i=0}^{m}\sum_{k=0}^{p}\lambda^{i+k}A_{m-i}B_{p-k}=\sum_{j=0}^{m+p}\lambda^{j}\sum_{i+k=j}A_{m-i}B_{p-k}. \quad (7'')$$

Преобразования в тождестве (7′) сохраняют свою силу при замене λ матрицей n-го порядка Λ, если только матрица Λ перестановочна со всеми матричными коэффициентами B_{p-k}[1]). Аналогично в тождестве (7″) можно заменить скаляр λ матрицей Λ, если матрица Λ перестановочна со всеми коэффициентами A_{m-i}. В первом случае получаем:

$$P(\Lambda)=A(\Lambda)B(\Lambda), \quad (8')$$

во втором

$$\widehat{P}(\Lambda)=\widehat{A}(\Lambda)\widehat{B}(\Lambda). \quad (8'')$$

Таким образом, *правое* (*левое*) *значение произведения двух матричных многочленов равно произведению правых* (*левых*) *значений сомножителей, если матрица-аргумент* Λ *перестановочна со всеми коэффициентами правого* (*левого*) *сомножителя.*

Если $S(\lambda)$ — сумма двух матричных многочленов n-го порядка $A(\lambda)$ и $B(\lambda)$, то при замене скаляра λ любой матрицей n-го порядка Λ всегда справедливы тождества

$$S(\Lambda)=A(\Lambda)+B(\Lambda), \qquad \widehat{S}(\Lambda)=\widehat{A}(\Lambda)+\widehat{B}(\Lambda). \quad (9)$$

§ 2. Правое и левое деление матричных многочленов. Обобщенная теорема Безу

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ одного и того же порядка n, причем $B(\lambda)$ — регулярный многочлен:

$$A(\lambda)=A_0\lambda^{m}+A_1\lambda^{m-1}+\ldots+A_m \qquad (A_0\neq 0),$$
$$B(\lambda)=B_0\lambda^{p}+B_1\lambda^{p-1}+\ldots+B_p \qquad (|B_0|\neq 0).$$

Мы будем говорить, что матричные многочлены $Q(\lambda)$ и $R(\lambda)$ являются соответственно *правым частным* и *правым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda)=Q(\lambda)B(\lambda)+R(\lambda) \quad (10)$$

и степень $R(\lambda)$ меньше степени $B(\lambda)$.

Совершенно аналогично будем называть многочлены $\widehat{Q}(\lambda)$ и $\widehat{R}(\lambda)$ соответственно *левым частным* и *левым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda)=B(\lambda)\widehat{Q}(\lambda)+\widehat{R}(\lambda) \quad (11)$$

и степень $\widehat{R}(\lambda)$ меньше степени $\widehat{B}(\lambda)$.

[1]) В этом случае и любая степень матрицы Λ перестановочна со всеми коэффициентами B_{p-k}.

Обратим внимание читателя, что при «правом» делении (т. е. при нахождении правого частного и правого остатка) в (5) на «делитель» $B(\lambda)$ частное $Q(\lambda)$ умножается справа, а при «левом» делении в (6) на делитель $B(\lambda)$ частное $\widehat{Q}(\lambda)$ умножается слева. В общем случае многочлены $Q(\lambda)$ и $R(\lambda)$ не совпадают с $\widehat{Q}(\lambda)$ и $\widehat{R}(\lambda)$.

Покажем, что *как правое, так и левое деление матричных многочленов одного и того же порядка всегда выполнимо и однозначно, если делитель — регулярный многочлен.*

Рассмотрим правое деление $A(\lambda)$ на $B(\lambda)$. Если $m < p$, можно положить $Q(\lambda) = 0$, $R(\lambda) = A(\lambda)$. В случае $m \geqslant p$ для нахождения частного $Q(\lambda)$ и остатка $R(\lambda)$ применим обычную схему деления многочлена на многочлен. «Разделим» старший член делимого $A_0\lambda^m$ на старший член делителя $B_0\lambda^p$. Получим старший член искомого частного $A_0B_0^{-1}\lambda^{m-p}$. Умножим этот член справа на делитель $B(\lambda)$ и полученное произведение вычтем из $A(\lambda)$. Найдем «первый остаток» $A^{(1)}(\lambda)$:

$$A(\lambda) = A_0B_0^{-1}\lambda^{m-p}B(\lambda) + A^{(1)}(\lambda). \tag{12}$$

Степень $m^{(1)}$ многочлена $A^{(1)}(\lambda)$ меньше m:

$$A^{(1)}(\lambda) = A_0^{(1)}\lambda^{m^{(1)}} + \ldots \quad (A_0^{(1)} \neq 0,\ m^{(1)} < m). \tag{13}$$

Если $m^{(1)} \geqslant p$, то, повторяя этот процесс, получаем:

$$\left.\begin{aligned} A^{(1)}(\lambda) &= A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p}B(\lambda) + A^{(2)}(\lambda), \\ A^{(2)}(\lambda) &= A_0^{(2)}\lambda^{m^{(2)}} + \ldots \qquad (m^{(2)} < m^{(1)}) \end{aligned}\right\} \tag{14}$$

и т. д.

Так как степени многочленов $A(\lambda)$, $A^{(1)}(\lambda)$, $A^{(2)}(\lambda)$, ... убывают, то на некотором этапе мы придем к остатку $R(\lambda)$, степень которого меньше p. Тогда из (12) — (14) будет следовать:

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda),$$

где

$$Q(\lambda) = A_0B_0^{-1}\lambda^{m-p} + A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p} + \ldots$$

Докажем теперь однозначность правого деления. Пусть одновременно

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \tag{15}$$

и

$$A(\lambda) = Q^*(\lambda)B(\lambda) + R^*(\lambda), \tag{15'}$$

где степени многочленов $R(\lambda)$ и $R^*(\lambda)$ меньше степени $B(\lambda)$, т. е. меньше p. Вычитая почленно (11) из (12), получим:

$$[Q(\lambda) - Q^*(\lambda)]B(\lambda) = R^*(\lambda) - R(\lambda). \tag{16}$$

Если бы $Q(\lambda) - Q^*(\lambda) \not\equiv 0$, то, поскольку $|B_0| \neq 0$, степень левой части равенства (16) равнялась бы сумме степеней $B(\lambda)$ и $Q(\lambda) - Q^*(\lambda)$ и потому была бы $\geqslant p$. Это невозможно, так как степень многочлена, стоящего в правой части равенства (16), меньше p. Таким образом, $Q(\lambda) - Q^*(\lambda) \equiv 0$, а тогда из (16) $R^*(\lambda) - R(\lambda) \equiv 0$, т. е.

$$Q(\lambda) = Q^*(\lambda), \quad R(\lambda) = R^*(\lambda).$$

Совершенно аналогично устанавливаются существование и единственность левого частного и левого остатка [1]).

Пример.

$$A(\lambda)=\begin{Vmatrix} \lambda^3+\lambda & 2\lambda^3+\lambda^2 \\ -\lambda^3-2\lambda^2+1 & 3\lambda^3+\lambda \end{Vmatrix}=$$

$$=\overbrace{\begin{Vmatrix} 1 & 2 \\ -1 & 3 \end{Vmatrix}}^{A_0}\lambda^3+\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\lambda^2+\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}\lambda+\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix},$$

$$B(\lambda)=\begin{Vmatrix} 2\lambda^2+3 & -\lambda^2+1 \\ -\lambda^2-1 & \lambda^2+2 \end{Vmatrix}=\overbrace{\begin{Vmatrix} 2 & -1 \\ -1 & 1 \end{Vmatrix}}^{B_0}\lambda^2+\begin{Vmatrix} 3 & 1 \\ -1 & 2 \end{Vmatrix},$$

$$|B_0|=1,\quad B_0^{-1}=\begin{Vmatrix} 1 & 1 \\ 1 & 2 \end{Vmatrix},\quad A_0B_0^{-1}=\begin{Vmatrix} 3 & 5 \\ 2 & 5 \end{Vmatrix},\quad A_0B_0^{-1}B(\lambda)=\begin{Vmatrix} \lambda^2+4 & 2\lambda^2+13 \\ -\lambda^2+1 & 3\lambda^2+12 \end{Vmatrix},$$

$$A^{(1)}(\lambda)=\begin{Vmatrix} \lambda^3+\lambda & 2\lambda^3+\lambda^2 \\ -\lambda^3-2\lambda^2+1 & 3\lambda^3+\lambda \end{Vmatrix}-\begin{Vmatrix} \lambda^3+4\lambda & 2\lambda^3+13\lambda \\ -\lambda^3+\lambda & 3\lambda^3+12\lambda \end{Vmatrix}=$$

$$=\begin{Vmatrix} -3\lambda & \lambda^2-13\lambda \\ -2\lambda^2-\lambda+1 & -11\lambda \end{Vmatrix},$$

$$A^{(1)}(\lambda)=\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\lambda^2+\begin{Vmatrix} -3 & -13 \\ -1 & -11 \end{Vmatrix}\lambda+\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix},$$

$$A_0^{(1)}B_0^{-1}=\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\cdot\begin{Vmatrix} 1 & 1 \\ 1 & 2 \end{Vmatrix}=\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix},$$

$$A_0^{(1)}B_0^{-1}B(\lambda)=\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix}\cdot\begin{Vmatrix} 2\lambda^2+3 & -\lambda^2+1 \\ -\lambda^2-1 & \lambda^2+2 \end{Vmatrix}=\begin{Vmatrix} 1 & \lambda^2+5 \\ -2\lambda^2-4 & -6 \end{Vmatrix},$$

$$R(\lambda)=A^{(1)}(\lambda)-A_0^{(1)}B_0^{-1}B(\lambda)=$$

$$=\begin{Vmatrix} -3\lambda & \lambda^2-13\lambda \\ -2\lambda^2-\lambda+1 & -11\lambda \end{Vmatrix}-\begin{Vmatrix} 1 & \lambda^2+5 \\ -2\lambda^2-4 & -6 \end{Vmatrix}=\begin{Vmatrix} -3\lambda-1 & -13\lambda-5 \\ -\lambda+5 & -11\lambda+6 \end{Vmatrix},$$

$$Q(\lambda)=A_0B_0^{-1}\lambda+A_0^{(1)}B_0^{-1}=\begin{Vmatrix} 3 & 5 \\ 2 & 5 \end{Vmatrix}\lambda+\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix}=\begin{Vmatrix} 3\lambda+1 & 5\lambda+2 \\ 2\lambda-2 & 5\lambda-2 \end{Vmatrix}.$$

Предлагаем читателю в качестве упражнения непосредственно проверить, что

$$A(\lambda)=Q(\lambda)B(\lambda)+R(\lambda).$$

---

[1]) Заметим, что возможность и однозначность левого деления $A(\lambda)$ на $B(\lambda)$ следует из возможности и однозначности правого деления транспонированных матриц $A'(\lambda)$ и $B'(\lambda)$. (Из регулярности $B(\lambda)$ вытекает регулярность $B'(\lambda)$.) Действительно, из

$$A'(\lambda)=Q_1(\lambda)B'(\lambda)+R_1(\lambda)$$

следует (см. гл. I, стр. 29):

$$A(\lambda)=B(\lambda)Q_1'(\lambda)+R_1'(\lambda). \qquad (11')$$

В силу этих же соображений левое деление $A(\lambda)$ на $B(\lambda)$ однозначно, так как из неоднозначности левого деления $A(\lambda)$ на $B(\lambda)$ следовала бы неоднозначность правого деления $A'(\lambda)$ на $B'(\lambda)$.

Сопоставление (11) и (11′) дает:

$$\widehat{Q}(\lambda)=Q_1'(\lambda),\quad \widehat{R}(\lambda)=R_1'(\lambda).$$

Рассмотрим произвольный матричный многочлен n-го порядка

$$F(\lambda)=F_0\lambda^m+F_1\lambda^{m-1}+\ldots+F_m \qquad (F_0\neq 0). \tag{17}$$

Разделим его на бином $\lambda E-A$ справа и слева:

$$F(\lambda)=Q(\lambda)(\lambda E-A)+R, \qquad F(\lambda)=(\lambda E-A)\widehat{Q}(\lambda)+\widehat{R}. \tag{18}$$

В данном случае правый остаток R и левый остаток $\widehat{R}$ не будут зависеть от λ. Для определения правого значения $F(A)$ и левого $\widehat{F}(A)$ можно соответственно в тождествах (18) заменить скаляр λ на матрицу A, поскольку матрица A перестановочна с матричными коэффициентами бинома $\lambda E-A$ (см. § 1):

$$F(A)=Q(A)(A-A)+R=R, \qquad \widehat{F}(A)=(A-A)\widehat{Q}(A)+\widehat{R}=\widehat{R}. \tag{19}$$

Нами доказана

Теорема 1 (обобщенная теорема Безу). *При правом (левом) делении матричного многочлена $F(\lambda)$ на бином $\lambda E-A$ остаток от деления равен $F(A)$ (соответственно $\widehat{F}(A)$).*

Из доказанной теоремы следует, что *многочлен $F(\lambda)$ делится без остатка справа (слева) на бином $\lambda E-A$ тогда и только тогда, когда $F(A)=0$ (соответственно $\widehat{F}(A)=0$).*

Пример. Пусть $A=\|a_{ik}\|_1^n$ и $f(\lambda)$ — многочлен относительно λ. Тогда

$$F(\lambda)=f(\lambda)E-f(A)$$

делится (слева и справа) без остатка на $\lambda E-A$. Это следует непосредственно из обобщенной теоремы Безу, поскольку в данном случае $F(A)=\widehat{F}(A)=0$.

§ 3. Характеристический многочлен матрицы. Присоединенная матрица

1. Рассмотрим матрицу $A=\|a_{ik}\|_1^n$. *Характеристической матрицей* для матрицы A называется матрица $\lambda E-A$. Определитель характеристической матрицы

$$\Delta(\lambda)=|\lambda E-A|=|\lambda\delta_{ik}-a_{ik}|_1^n$$

представляет собой скалярный многочлен относительно λ и называется *характеристическим многочленом* матрицы A (см. гл. III, § 7)[1]).

Матрицу $B(\lambda)=\|b_{ik}(\lambda)\|_1^n$, где $b_{ik}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ik}-a_{ki}$ в определителе $\Delta(\lambda)$, мы будем называть *присоединенной матрицей* для матрицы A.

Так, например, для матрицы

$$A=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

[1]) Этот многочлен отличается множителем $(-1)^n$ от многочлена $\Delta(\lambda)$, введенного в гл. III, § 7.

будем иметь:

$$\lambda E - A = \begin{Vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{Vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 + \ldots,$$

$$B(\lambda) = \begin{Vmatrix} \lambda^2 - (a_{22} + a_{33})\lambda + a_{22}a_{33} - a_{23}a_{32} & * & * \\ a_{21}\lambda + a_{23}a_{31} - a_{21}a_{33} & * & * \\ a_{31}\lambda + a_{21}a_{32} - a_{22}a_{31} & * & * \end{Vmatrix}.$$

Из приведенных определений следуют тождества относительно λ:

$$(\lambda E - A) B(\lambda) = \Delta(\lambda) E, \tag{20}$$

$$B(\lambda)(\lambda E - A) = \Delta(\lambda) E. \tag{20'}$$

Правые части этих равенств мы можем рассматривать как многочлены с матричными коэффициентами (каждый из этих коэффициентов равен произведению скаляра на единичную матрицу E). Многочленную матрицу $B(\lambda)$ можно также представить в виде многочлена, расположенного по степеням λ. Равенства (20) и (20') показывают, что $\Delta(\lambda)E$ делится слева и справа на $\lambda E - A$ без остатка. Согласно обобщенной теореме Безу это возможно лишь тогда, когда остаток $\Delta(A)E = \Delta(A)$ равен нулю. Нами доказана

Теорема 2 (Гамильтона—Кэли). *Всякая квадратная матрица A удовлетворяет своему характеристическому уравнению*, т. е.

$$\Delta(A) = 0. \tag{21}$$

Пример.

$$A = \begin{Vmatrix} 2 & 1 \\ -1 & 3 \end{Vmatrix},$$

$$\Delta(\lambda) = \begin{Vmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 3 \end{Vmatrix} = \lambda^2 - 5\lambda + 7,$$

$$\Delta(A) = A^2 - 5A + 7E = \begin{Vmatrix} 3 & 5 \\ -5 & 8 \end{Vmatrix} - 5\begin{Vmatrix} 2 & 1 \\ -1 & 3 \end{Vmatrix} + 7\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix} = 0.$$

2. Обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$ все характеристические числа матрицы A, т. е. все корни характеристического многочлена $\Delta(\lambda)$ (каждое из чисел λ_i повторяется в этом ряду столько раз, какова его кратность как корня многочлена $\Delta(\lambda)$). Тогда

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n). \tag{22}$$

Пусть дан произвольный скалярный многочлен $g(\mu)$. Найдем характеристические числа матрицы $g(A)$. Для этого разложим $g(\mu)$ на линейные множители

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2) \ldots (\mu - \mu_l). \tag{23}$$

Подставим в обе части этого тождества вместо μ матрицу A:

$$g(A) = a_0(A - \mu_1 E)(A - \mu_2 E) \ldots (A - \mu_l E). \tag{24}$$

Переходя к определителям в обеих частях равенства (24) и используя равенства (22) и (23), получим:

$$|g(A)| = a_0^n |A-\mu_1 E||A-\mu_2 E|\ldots|A-\mu_l E| =$$
$$= (-1)^{nl} a_0^n \Delta(\mu_1)\Delta(\mu_2)\ldots\Delta(\mu_l) =$$
$$= (-1)^{nl} a_0^n \prod_{i=1}^{l}\prod_{k=1}^{n}(\mu_i-\lambda_k) = g(\lambda_1)\,g(\lambda_2)\ldots g(\lambda_n).$$

Заменив в равенстве

$$|g(A)| = g(\lambda_1)\,g(\lambda_2)\ldots g(\lambda_n) \tag{25}$$

многочлен $g(\mu)$ на $\lambda - g(\mu)$, где λ — некоторый параметр, найдем:

$$|\lambda E - g(A)| = [\lambda - g(\lambda_1)][\lambda - g(\lambda_2)]\ldots[\lambda - g(\lambda_n)]. \tag{26}$$

Из этого равенства вытекает следующая

Теорема 3. *Если* $\lambda_1, \lambda_2, \ldots, \lambda_n$ — *все характеристические числа (с учетом кратностей) матрицы* A, *а* $g(\mu)$ — *некоторый скалярный многочлен, то* $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$ — *все характеристические числа матрицы* $g(A)$.

В частности, если матрица A имеет характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, то матрица A^k имеет характеристические числа $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ $(k = 0, 1, 2, 3, \ldots)$.

3. Укажем эффективную формулу, выражающую присоединенную матрицу $B(\lambda)$ через характеристический многочлен $\Delta(\lambda)$.

Пусть

$$\Delta(\lambda) = \lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \ldots - p_n. \tag{27}$$

Разность $\Delta(\lambda) - \Delta(\mu)$ делится без остатка на $\lambda - \mu$. Поэтому

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda)-\Delta(\mu)}{\lambda-\mu} = \lambda^{n-1} + (\mu - p_1)\lambda^{n-2} + (\mu^2 - p_1\mu - p_2)\lambda^{n-3} + \ldots \tag{28}$$

есть многочлен относительно λ и μ.

Тождество

$$\Delta(\lambda) - \Delta(\mu) = \delta(\lambda, \mu)(\lambda - \mu) \tag{29}$$

не нарушится, если в него вместо λ и μ подставить перестановочные между собой матрицы λE и A. Тогда, поскольку согласно теореме Гамильтона — Кэли $\Delta(A) = 0$,

$$\Delta(\lambda) E = \delta(\lambda E, A)(\lambda E - A). \tag{30}$$

Сопоставляя между собой равенства (20′) и (30), из однозначности частного получаем искомую формулу

$$B(\lambda) = \delta(\lambda E, A). \tag{31}$$

Отсюда в силу (28)

$$B(\lambda) = \lambda^{n-1} + B_1\lambda^{n-2} + B_2\lambda^{n-3} + \ldots + B_{n-1}, \tag{32}$$

где

$$B_1 = A - p_1 E, \quad B_2 = A^2 - p_1 A - p_2 E, \ldots$$

и вообще

$$B_k = A^k - p_1 A^{k-1} - p_2 A^{k-2} - \ldots - p_k E \qquad (k = 1, 2, \ldots, n-1). \tag{33}$$

Матрицы $B_1, B_2, \ldots, B_{n-1}$ можно вычислять последовательно, исходя из рекуррентных соотношений

$$B_k = AB_{k-1} - p_k E \qquad (k = 1, 2, \ldots, n-1;\ B_0 = E). \tag{34}$$

При этом

$$AB_{n-1} - p_n E = 0\ ^{1}). \tag{35}$$

Соотношения (34) и (35) непосредственно получаются из тождества (20), если в обеих частях этого тождества приравнять между собой коэффициенты при одинаковых степенях λ.

Если A — неособенная матрица, то

$$p_n = (-1)^{n-1} |A| \neq 0$$

и из (35) следует:

$$A^{-1} = \frac{1}{p_n} B_{n-1}. \tag{36}$$

Пусть λ_0 — характеристическое число матрицы A, т. е. $\Delta(\lambda_0) = 0$. Подставив в (20) вместо λ значение λ_0, найдем:

$$(\lambda_0 E - A) B(\lambda_0) = 0. \tag{37}$$

Допустим, что матрица $B(\lambda_0) \neq 0$, и обозначим через b любой ненулевой столбец этой матрицы. Тогда из (37) $(\lambda_0 E - A) b = 0$, или

$$Ab = \lambda_0 b. \tag{38}$$

Следовательно, любой ненулевой столбец матрицы $B(\lambda_0)$ определяет собственный вектор, соответствующий характеристическому числу λ_0 [2]).

Таким образом, *если коэффициенты характеристического многочлена известны, то присоединенная матрица может быть найдена по формуле* (31). *Если данная матрица A неособенная, то по формуле* (36) *находится обратная матрица A^{-1}. Если λ_0 — характеристическое число матрицы A, то ненулевые столбцы матрицы $B(\lambda_0)$ являются собственными векторами матрицы A для $\lambda = \lambda_0$.*

Пример.

$$A = \begin{Vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{Vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda-2 & 1 & -1 \\ 0 & \lambda-1 & -1 \\ 1 & -1 & \lambda-1 \end{vmatrix} = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^2 + \lambda(\mu - 4) + \mu^2 - 4\mu + 5,$$

$$B(\lambda) = \delta(\lambda E, A) = \lambda^2 E + \lambda \underbrace{(A - 4E)}_{B_1} + \underbrace{A^2 - 4A + 5E}_{B_2}.$$

---

[1]) Из (34) следуют равенства (33). Если в (35) подставить выражение для B_{n-1} из (33), то получим: $\Delta(A) = 0$. Этот вывод теоремы Гамильтона — Кэли не опирается явным образом на обобщенную теорему Безу, но неявно содержит в себе эту теорему.

[2]) См. гл. III, § 7. Если характеристическому числу λ_0 соответствует d_0 линейно независимых собственных векторов ($n - d_0$ — ранг матрицы $\lambda_0 E - A$), то ранг матрицы $B(\lambda_0)$ не превосходит d_0. В частности, если числу λ_0 соответствует только одно собственное направление, то в матрице $B(\lambda_0)$ элементы любых двух столбцов пропорциональны.

Но

$$B_1 = A - 4E = \begin{Vmatrix} -2 & -1 & 1 \\ 0 & -3 & 1 \\ -1 & 1 & -3 \end{Vmatrix},$$

$$B_2 = AB_1 + 5E = \begin{Vmatrix} 0 & 2 & -2 \\ -1 & +3 & -2 \\ 1 & -1 & 2 \end{Vmatrix},$$

$$B(\lambda) = \begin{Vmatrix} \lambda^2 - 2\lambda & -\lambda + 2 & \lambda - 2 \\ -1 & \lambda^2 - 3\lambda + 3 & \lambda - 2 \\ -\lambda + 1 & \lambda - 1 & \lambda^2 - 3\lambda + 2 \end{Vmatrix},$$

$$|A| = 2, \quad A^{-1} = \frac{1}{2} B_2 = \begin{Vmatrix} 0 & 1 & -1 \\ -\frac{1}{2} & \frac{3}{2} & -1 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix}.$$

Далее

$$\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2)$$

Первый столбец матрицы $B(+1)$ дает собственный вектор $(+1, +1, 0)$ для характеристического числа $\lambda = 1$.

Первый столбец матрицы $B(+2)$ дает собственный вектор $(0, 1, 1)$, соответствующий характеристическому числу $\lambda = 2$.

§ 4. Метод Д. К. Фаддеева одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы

Д. К. Фаддеев[1]) предложил метод одновременного определения скалярных коэффициентов $p_1, p_2, \ldots, p_n$ характеристического многочлена

$$\Delta(\lambda) = \lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \ldots - p_n \tag{39}$$

и матричных коэффициентов $B_1, B_2, \ldots, B_{n-1}$ присоединенной матрицы $B(\lambda)$.

Для изложения метода Д. К. Фаддеева[2]) введем понятие о следе матрицы.

Под *следом* матрицы $A = \| a_{ik} \|_1^n$ (обозначение: $\operatorname{Sp} A$) понимают сумму диагональных элементов этой матрицы:

$$\operatorname{Sp} A = \sum_{i=1}^{n} a_{ii}. \tag{40}$$

Нетрудно видеть, что

$$\operatorname{Sp} A = p_1 = \sum_{i=1}^{n} \lambda_i, \tag{41}$$

---

[1]) См. [32], стр. 160.

[2]) С другим эффективным методом вычисления коэффициентов характеристического многочлена, с методом А. Н. Крылова, мы познакомим читателя в гл. VII, § 8.

если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы A, т. е.

$$\Delta(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n). \tag{42}$$

Так как, согласно теореме 3, степень матрицы A^k имеет своими характеристическими числами степени $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ $(k = 0, 1, 2, \ldots)$, то

$$\operatorname{Sp} A^k = s_k = \sum_{i=1}^{n} \lambda_i^k \qquad (k = 0, 1, 2, \ldots). \tag{43}$$

Суммы s_k $(k = 1, 2, \ldots, n)$ степеней корней многочлена (39) связаны с коэффициентами этого уравнения формулами Ньютона [1])

$$kp_k = s_k - p_1 s_{k-1} - \ldots - p_{k-1} s_1 \qquad (k = 1, 2, \ldots, n). \tag{44}$$

Если вычислить следы $s_1, s_2, \ldots, s_n$ матриц $A, A^2, \ldots, A^n$, то затем можно из уравнений (44) последовательно определить коэффициенты $p_1, p_2, \ldots, p_n$. В этом состоит метод Леверрье определения коэффициентов характеристического многочлена по следам степеней матрицы.

Д. К. Фаддеев предложил вместо следов степеней $A, A^2, \ldots, A^n$ вычислить последовательно следы некоторых других матриц $A_1, A_2, \ldots, A_n$ и с их помощью определить $p_1, p_2, \ldots, p_n$ и $B_1, B_2, \ldots, B_n$ следующими формулами:

$$\left.\begin{array}{lll} A_1 = A & p_1 = \operatorname{Sp} A_1, & B_1 = A_1 - p_1 E, \\ A_2 = AB_1, & p_2 = \frac{1}{2} \operatorname{Sp} A_2, & B_2 = A_2 - p_2 E, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ A_{n-1} = AB_{n-2}, & p_{n-1} = \frac{1}{n-1} \operatorname{Sp} A_{n-1}, & B_{n-1} = A_{n-1} - p_{n-1} E, \\ A_n = AB_{n-1}, & p_n = \frac{1}{n} \operatorname{Sp} A_n, & B_n = A_n - p_n E = 0. \end{array}\right\} \tag{45}$$

Последнее равенство $B_n = A_n - p_n E = 0$ может быть использовано для контроля вычислений.

Для того чтобы убедиться, что числа $p_1, p_2, \ldots, p_n$ и матрицы $B_1, B_2, \ldots, B_{n-1}$, последовательно определяемые по формулам (45), являются коэффициентами $\Delta(\lambda)$ и $B(\lambda)$, заметим, что из (45) вытекают следующие формулы для A_k и B_k $(k = 1, 2, \ldots, n)$:

$$A_k = A^k - p_1 A^{k-1} - \ldots - p_{k-1} A, \quad B_k = A^k - p_1 A^{k-1} - \ldots - p_{k-1} A - p_k E. \tag{46}$$

Приравняем между собой следы левой и правой частей первой из этих формул; получим:

$$kp_k = s_k - p_1 s_{k-1} - \ldots - p_{k-1} s_1.$$

Но эти формулы совпадают с формулами Ньютона (44), по которым последовательно определяются коэффициенты характеристического многочлена $\Delta(\lambda)$. Следовательно, числа $p_1, p_2, \ldots, p_n$, определяемые по формулам (45), и являются коэффициентами $\Delta(\lambda)$. Но тогда вторые формулы (46) совпадают с формулами (33), по которым определяются матричные коэффициенты $B_1, B_2, \ldots, B_{n-1}$ присоединенной матрицы $B(\lambda)$. Следовательно, формулы (45) определяют и коэффициенты $B_1, B_2, \ldots, B_{n-1}$ матричного многочлена $B(\lambda)$.

[1]) См., например, [20], стр. 224.

Пример[1]).

$$A=\begin{Vmatrix} 2 & -1 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ \hline 2 & 2 & 4 & 3 \end{Vmatrix},\quad p_1=\operatorname{Sp} A=4,\quad B_1=A-4E=\begin{Vmatrix} -2 & -1 & 1 & 2 \\ 0 & -3 & 1 & 0 \\ -1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -4 \end{Vmatrix},$$

$$A_2=AB_1=\begin{Vmatrix} -3 & 4 & 0 & -3 \\ -1 & -2 & -2 & 1 \\ 2 & 0 & -2 & -5 \\ -3 & -3 & -1 & 3 \\ \hline -5 & -1 & -5 & -4 \end{Vmatrix},\quad p_2=\frac{1}{2}\operatorname{Sp} A_2=-2,\quad B_2=A_2+2E=\begin{Vmatrix} -1 & 4 & 0 & -3 \\ -1 & 0 & -2 & 1 \\ 2 & 0 & 0 & -5 \\ -3 & -3 & -1 & 3 \end{Vmatrix};$$

$$A_3=AB_2=\begin{Vmatrix} -5 & 2 & 0 & -2 \\ 1 & 0 & -2 & -4 \\ -1 & -7 & -3 & 4 \\ 0 & 4 & -2 & -7 \\ \hline -5 & -1 & -7 & -9 \end{Vmatrix},\quad p_3=\frac{1}{3}\operatorname{Sp} A_3=-5,\quad B_3=A_3+5E=\begin{Vmatrix} 0 & 2 & 0 & -2 \\ 1 & 5 & -2 & 4 \\ -1 & -7 & 2 & 4 \\ 0 & 4 & -2 & -2 \end{Vmatrix};$$

$$A_4=AB_3=\begin{Vmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{Vmatrix},\quad p_4=-2,$$

$$\Delta(\lambda)=\lambda^4-4\lambda^3+2\lambda^2+5\lambda+2,$$

$$|A|=2,\quad A^{-1}=\frac{1}{p_4}B_3=\begin{Vmatrix} 0 & -1 & 0 & 1 \\ -\frac{1}{2} & -\frac{5}{2} & 1 & -2 \\ \frac{1}{2} & \frac{7}{2} & -1 & -2 \\ 0 & -2 & 1 & 1 \end{Vmatrix}.$$

Замечание. Если мы хотим определить p_1, p_2, p_3, p_4 и только первые столбцы в B_1, B_2, B_3, то достаточно вычислить в A_2 элементы 1-го столбца и только диагональные элементы остальных столбцов, в A_3 — только элементы 1-го столбца, в A_4 — только два первых элемента 1-го столбца.

§ 5. Минимальный многочлен матрицы

Определение 1. Скалярный многочлен $f(\lambda)$ называется *аннулирующим многочленом* квадратной матрицы A, если

$$f(A)=0.$$

Аннулирующий многочлен $\psi(\lambda)$ наименьшей степени со старшим коэффициентом, равным единице, называется *минимальным многочленом* матрицы A.

Согласно теореме Гамильтона — Кэли характеристический многочлен $\Delta(\lambda)$ матрицы A является аннулирующим для этой матрицы. Однако, как будет показано ниже, в общем случае он не является минимальным.

[1]) Для контроля за вычислениями мы под каждой из матриц A_1, A_2, A_3 подписываем ее суммарную строку. Произведения суммарной строки первого сомножителя на столбцы второго должны дать элементы суммарной строки произведения.

Разделим произвольный аннулирующий многочлен $f(\lambda)$ на минимальный:

$$f(\lambda) = \psi(\lambda)\, q(\lambda) + r(\lambda),$$

где степень $r(\lambda)$ меньше степени $\psi(\lambda)$. Отсюда имеем:

$$f(A) = \psi(A)\, q(A) + r(A).$$

Поскольку $f(A) = 0$ и $\psi(A) = 0$, то, значит, и $r(A) = 0$. Но степень $r(\lambda)$ меньше степени минимального многочлена $\psi(\lambda)$. Поэтому $r(\lambda) \equiv 0$ [1]). Таким образом, *произвольный аннулирующий многочлен матрицы всегда делится без остатка на ее минимальный многочлен.*

Пусть два многочлена $\psi(\lambda)$ и $\psi^1(\lambda)$ являются минимальными для одной и той же матрицы. Тогда каждый из них делится на другой многочлен без остатка, т. е. эти многочлены отличаются постоянным множителем. Этот постоянный множитель равен единице, поскольку равны единице старшие коэффициенты в $\psi(\lambda)$ и $\psi^1(\lambda)$. Мы доказали *единственность минимального многочлена* для данной матрицы A.

Выведем формулу, связывающую минимальный многочлен с характеристическим.

Обозначим через $D_{n-1}(\lambda)$ наибольший общий делитель всех миноров $(n-1)$-го порядка характеристической матрицы $\lambda E - A$, т. е. наибольший общий делитель всех элементов присоединенной матрицы $B(\lambda) = \| b_{ik}(\lambda) \|_1^n$ (см. предыдущий параграф); при этом старший коэффициент в $D_{n-1}(\lambda)$ берем равным единице. Тогда

$$B(\lambda) = D_{n-1}(\lambda)\, C(\lambda), \tag{47}$$

где $C(\lambda)$ — некоторая многочленная матрица, *«приведенная» присоединенная матрица* для $\lambda E - A$. Из (20) и (47) находим:

$$\Delta(\lambda)\, E = (\lambda E - A)\, C(\lambda)\, D_{n-1}(\lambda). \tag{48}$$

Отсюда следует, что $\Delta(\lambda)$ делится без остатка на $D_{n-1}(\lambda)$ [2]):

$$\frac{\Delta(\lambda)}{D_{n-1}(\lambda)} = \psi(\lambda), \tag{49}$$

где $\psi(\lambda)$ — некоторый многочлен. Обе части тождества (48) можно сократить на $D_{n-1}(\lambda)$ [3]):

$$\psi(\lambda)\, E = (\lambda E - A)\, C(\lambda). \tag{50}$$

Поскольку $\psi(\lambda)\, E$ делится без остатка слева на $\lambda E - A$, то в силу обобщенной теоремы Безу

$$\psi(A) = 0.$$

Таким образом, многочлен $\psi(\lambda)$, определенный формулой (49), является аннулирующим многочленом для матрицы A. Докажем, что он является минимальным многочленом.

---

[1]) В противном случае существовал бы аннулирующий многочлен, степень которого была бы меньше степени минимального многочлена.

[2]) В этом можно и непосредственно убедиться, разлагая характеристический определитель $\Delta(\lambda)$ по элементам какой-либо строки.

[3]) В данном случае наряду с (50) имеет место и тождество [см. (20′)]

$$\psi(\lambda)\, E = C(\lambda)\, (\lambda E - A),$$

т. е. $C(\lambda)$ является одновременно и левым и правым частным от деления $\psi(\lambda)\, E$ на $\lambda E - A$.

Обозначим минимальный многочлен через $\psi^*(\lambda)$. Тогда $\psi(\lambda)$ делится без остатка на $\psi^*(\lambda)$:

$$\psi(\lambda) = \psi^*(\lambda)\chi(\lambda). \tag{51}$$

Поскольку $\psi^*(A) = 0$, то в силу обобщенной теоремы Безу матричный многочлен $\psi^*(\lambda)E$ делится слева без остатка на $\lambda E - A$:

$$\psi^*(\lambda)E = (\lambda E - A)C^*(\lambda). \tag{52}$$

Из (51) и (52) следует:

$$\psi(\lambda)E = (\lambda E - A)C^*(\lambda)\chi(\lambda). \tag{53}$$

Тождества (50) и (53) показывают, что как $C(\lambda)$, так и $C^*(\lambda)\chi(\lambda)$ являются левыми частными при делении $\psi(\lambda)E$ на $\lambda E - A$. В силу однозначности деления

$$C(\lambda) = C^*(\lambda)\chi(\lambda).$$

Отсюда следует, что $\chi(\lambda)$ является общим делителем всех элементов многочленной матрицы $C(\lambda)$. Но, с другой стороны, наибольший общий делитель всех элементов приведенной присоединенной матрицы $C(\lambda)$ равен единице, поскольку эта матрица была получена из $B(\lambda)$ путем деления на $D_{n-1}(\lambda)$. Поэтому $\chi(\lambda) = \text{const}$. Так как старшие коэффициенты в $\psi(\lambda)$ и $\psi^*(\lambda)$ равны единице, то в (51) $\chi(\lambda) = 1$, т. е. $\psi(\lambda) = \psi^*(\lambda)$, что и требовалось доказать.

Мы установили следующую формулу для минимального многочлена:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}. \tag{54}$$

Для приведенной присоединенной матрицы $C(\lambda)$ имеем формулу, аналогичную формуле (31) (на стр. 94):

$$C(\lambda) = \Psi(\lambda E, A), \tag{55}$$

где многочлен $\Psi(\lambda, \mu)$ определяется равенством

$$\Psi(\lambda, \mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu}\ ^{1}). \tag{56}$$

Кроме того,

$$(\lambda E - A)C(\lambda) = \psi(\lambda)E. \tag{57}$$

Переходя к определителям в обеих частях равенства (57), получаем:

$$\Delta(\lambda)\,|C(\lambda)| = [\psi(\lambda)]^n. \tag{58}$$

Таким образом, $\Delta(\lambda)$ делится без остатка на $\psi(\lambda)$, а некоторая степень $\psi(\lambda)$ делится без остатка на $\Delta(\lambda)$, т. е. совокупность всех различных между собой корней у многочлена $\Delta(\lambda)$ и $\psi(\lambda)$ одна и та же. Другими словами, *корнями $\psi(\lambda)$ служат все различные между собой характеристические числа матрицы A.*

Если

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \ldots (\lambda - \lambda_s)^{n_s} \tag{59}$$

$$(\lambda_i \neq \lambda_j \quad \text{при} \quad i \neq j; \quad n_i > 0,\ i, j = 1, 2, \ldots, s),$$

---

[1]) Формула (55) выводится совершенно так же, как и формула (31). В обе части тождества $\psi(\lambda) - \psi(\mu) = (\lambda - \mu)\Psi(\lambda, \mu)$ вместо λ и μ подставляются матрицы λE и A и полученное матричное равенство сопоставляется с (50).

то

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}, \tag{60}$$

где

$$0 < m_k \leqslant n_k \qquad (k = 1, 2, \dots, s). \tag{61}$$

Отметим еще одно свойство матрицы $C(\lambda)$. Пусть λ_0 — какое-либо характеристическое число матрицы $A = \| a_{ik} \|_1^n$. Тогда $\psi(\lambda_0) = 0$, и потому согласно (57)

$$(\lambda_0 E - A) C(\lambda_0) = 0. \tag{62}$$

Заметим, что всегда $C(\lambda_0) \neq 0$. Действительно, в противном случае все элементы приведенной присоединенной матрицы $C(\lambda)$ делились бы без остатка на $\lambda - \lambda_0$, что невозможно.

Обозначим через c любой ненулевой столбец матрицы $C(\lambda_0)$. Тогда из (62)

$$(\lambda_0 E - A) c = 0,$$

т. е.

$$Ac = \lambda_0 c. \tag{63}$$

Другими словами, любой ненулевой столбец матрицы $C(\lambda_0)$ (а такой столбец всегда имеется) определяет собственный вектор для $\lambda = \lambda_0$.

Пример.

$$A = \begin{Vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{Vmatrix},$$

$$\Delta(\lambda) = \begin{vmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{vmatrix} = \lambda^3 - 8\lambda^2 + 20\lambda - 16 = (\lambda - 2)^2 (\lambda - 4),$$

$$\delta(\lambda, \mu) = \frac{\Delta(\mu) - \Delta(\lambda)}{\mu - \lambda} = \mu^2 + \mu(\lambda - 8) + \lambda^2 - 8\lambda + 20,$$

$$B(\lambda) = A^2 + (\lambda - 8) A + (\lambda^2 - 8\lambda + 20) E =$$

$$= \begin{Vmatrix} 10 & -18 & 12 \\ -6 & 22 & -12 \\ -6 & 18 & -8 \end{Vmatrix} + (\lambda - 8) \begin{Vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{Vmatrix} + (\lambda^2 - 8\lambda + 20) \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{Vmatrix} =$$

$$= \begin{Vmatrix} \lambda^2 - 5\lambda + 6 & -3\lambda + 6 & 2\lambda - 4 \\ -\lambda + 2 & \lambda^2 - 3\lambda + 2 & -2\lambda + 4 \\ -\lambda + 2 & 3\lambda - 6 & \lambda^2 - 8\lambda + 2 \end{Vmatrix}.$$

Все элементы матрицы $B(\lambda)$ делятся на $D_2(\lambda) = \lambda - 2$. Сокращая этот множитель, получим:

$$C(\lambda) = \begin{Vmatrix} \lambda - 3 & -3 & 2 \\ -1 & \lambda - 1 & -2 \\ -1 & 3 & \lambda - 6 \end{Vmatrix}$$

и

$$\psi(\lambda) = \frac{\Delta(\lambda)}{\lambda - 2} = (\lambda - 2)(\lambda - 4).$$

Подставим в $C(\lambda)$ вместо λ значение $\lambda_0=2$:

$$C(2)=\begin{Vmatrix} -1 & -3 & 2 \\ -1 & 1 & -2 \\ -1 & 3 & -4 \end{Vmatrix}$$

Первый столбец дает нам собственный вектор (1, 1, 1) для $\lambda_0=2$. Второй столбец дает нам собственный вектор (—3, 1, 3) для того же характеристического числа $\lambda_0=2$. Третий столбец есть линейная комбинация первых двух.

Точно так же, полагая $\lambda=4$, из первого столбца матрицы $C(4)$ найдем собственный вектор (1, —1, —1), отвечающий характеристическому числу $\lambda_0=4$.

Обратим внимание читателя на то, что $\psi(\lambda)$ и $C(\lambda)$ можно было бы определить иначе.

Находим сначала $D_2(\lambda)$. $D_2(\lambda)$ может иметь своими корнями только числа 2 и 4. При $\lambda=4$ в $\Delta(\lambda)$ минор второго порядка $\begin{vmatrix} 1 & \lambda-5 \\ 1 & -3 \end{vmatrix}=-\lambda+2$ не обращается в нуль. Поэтому $D_2(4)\neq 0$. При $\lambda=2$ столбцы матрицы A становятся пропорциональными. Поэтому все миноры второго порядка в $\Delta(\lambda)$ при $\lambda=2$ равны нулю: $D_2(2)=0$. Так как вычисленный минор имеет первую степень, то $D_2(\lambda)$ не может делиться на $(\lambda-2)^2$. Следовательно,

$$D_2(\lambda)=\lambda-2.$$

Отсюда

$$\psi(\lambda)=\frac{\Delta(\lambda)}{\lambda-2}=(\lambda-2)(\lambda-4)=\lambda^2-6\lambda+8,$$

$$\psi(\lambda,\mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}=\mu+\lambda-6,$$

$$C(\lambda)=\psi(\lambda E, A)=A+(\lambda-6)E=\begin{Vmatrix} \lambda-3 & -3 & 2 \\ -1 & \lambda-1 & -2 \\ -1 & 3 & \lambda-6 \end{Vmatrix}.$$

ГЛАВА V

ФУНКЦИИ ОТ МАТРИЦЫ

§ 1. Определение функции от матрицы

Пусть даны квадратная матрица $A = \| a_{ik} \|_1^n$ и функция $f(\lambda)$ скалярного аргумента λ. Требуется определить, что следует понимать под $f(A)$, т. е. требуется распространить функцию $f(\lambda)$ и на матричные значения аргумента.

Решение этой задачи нам известно для простейшего случая, когда $f(\lambda) = \gamma_0\lambda^l + \gamma_1\lambda^{l-1} + \ldots + \gamma_l$ — многочлен относительно λ. В этом случае $f(A) = \gamma_0 A^l + \gamma_1 A^{l-1} + \ldots + \gamma_l E$. Исходя из этого частного случая, постараемся определить $f(A)$ в общем случае.

Обозначим через

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s} \tag{1}$$

минимальный многочлен матрицы A [1]) (здесь $\lambda_1, \lambda_2, \ldots, \lambda_s$ — все различные характеристические числа матрицы A). Степень этого многочлена $m = \sum_{k=1}^{s} m_k$.

Пусть два многочлена $g(\lambda)$ и $h(\lambda)$ таковы, что

$$g(A) = h(A). \tag{2}$$

Тогда разность $d(\lambda) = g(\lambda) - h(\lambda)$, будучи аннулирующим многочленом для матрицы A, делится на $\psi(\lambda)$ без остатка, что запишем так:

$$g(\lambda) \equiv h(\lambda) \pmod{\psi(\lambda)}. \tag{3}$$

Отсюда в силу (1)

$$d(\lambda_k) = 0, \quad d'(\lambda_k) = 0, \ \ldots, \ d^{(m_k-1)}(\lambda_k) = 0 \ (k = 1, 2, \ldots, s),$$

т. е.

$$g(\lambda_k) = h(\lambda_k), \ g'(\lambda_k) = h'(\lambda_k), \ \ldots, \ g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k) \ (k = 1, 2, \ldots, s). \tag{4}$$

m чисел

$$f(\lambda_k), \ f'(\lambda_k), \ \ldots, \ f^{(m_k-1)}(\lambda_k) \qquad (k = 1, 2, \ldots, s) \tag{5}$$

---

[1]) См. гл. IV, § 5.

будем условно называть *значениями функции* $f(\lambda)$ *на спектре матрицы* A и совокупность этих значений символически обозначать через $f(\Lambda_A)$. Если для функции $f(\lambda)$ существуют (т. е. имеют смысл) значения (5), то мы будем говорить, что *функция* $f(\lambda)$ *определена на спектре матрицы* A.

Равенства (4) показывают, что многочлены $g(\lambda)$ и $h(\lambda)$ имеют одни и те же значения на спектре матрицы A. В символической записи

$$g(\Lambda_A) = h(\Lambda_A).$$

Рассуждения наши обратимы: из (4) вытекают (3) и, следовательно, (2).

Таким образом, если задана матрица A, то значения многочлена $g(\lambda)$ на спектре матрицы A вполне определяют матрицу $g(A)$, т. е. все многочлены $g(\lambda)$, принимающие одни и те же значения на спектре матрицы A, имеют одно и то же матричное значение $g(A)$.

Мы потребуем, чтобы определение $f(A)$ в общем случае подчинялось такому же принципу: *значения функции* $f(\lambda)$ *на спектре матрицы* A *должны полностью определять* $f(A)$, *т. е. все функции* $f(\lambda)$, *имеющие одни и те же значения на спектре матрицы* A, *должны иметь одно и то же матричное значение* $f(A)$ [1]).

Но тогда, очевидно, для определения $f(A)$ в общем случае достаточно подыскать такой многочлен $g(\lambda)$, который принимал бы те же значения на спектре матрицы A, что и $f(\lambda)$ [2]), и положить:

$$f(A) = g(A).$$

Таким образом, приходим к следующему определению:

Определение 1. Если функция $f(\lambda)$ определена на спектре матрицы A, то

$$f(A) = g(A),$$

где $g(\lambda)$ — любой многочлен, принимающий на спектре матрицы A те же значения, что и $f(\lambda)$:

$$f(\Lambda_A) = g(\Lambda_A).$$

Среди всех многочленов с комплексными коэффициентами, принимающих те же значения на спектре, что и $f(\lambda)$, имеется один и только один многочлен $r(\lambda)$ степени $< m$ [3]). Этот многочлен $r(\lambda)$ однозначно определяется интерполяционными условиями:

$$r(\lambda_k) = f(\lambda_k),\ r'(\lambda_k) = f'(\lambda_k),\ \ldots,\ r^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k)\ \ (k = 1, 2, \ldots, s). \tag{6}$$

Многочлен $r(\lambda)$ называют *интерполяционным многочленом Лагранжа-Сильвестра* для функции $f(\lambda)$ на спектре матрицы A. Определение 1 можно еще сформулировать так:

---

[1]) Кроме того, предполагается что общее определение для $f(A)$ в частном случае, когда $f(\lambda)$ — многочлен, должно давать тот же результат, что и непосредственная подстановка в многочлен вместо λ матрицы A.

[2]) В § 2 будет показано, что такой интерполяционный многочлен всегда существует, и будет дан алгоритм для вычисления коэффициентов интерполяционного многочлена наименьшей степени.

[3]) Этот многочлен получается из любого другого многочлена, имеющего те же спектральные значения, что и остаток от деления на $\psi(\lambda)$.

Определение 1′. Пусть $f(\lambda)$ — функция, определенная на спектре матрицы A, а $r(\lambda)$ — соответствующий интерполяционный многочлен Лагранжа—Сильвестра. Тогда

$$f(A)=r(A).$$

Замечание. Если минимальный многочлен $\psi(\lambda)$ матрицы A не имеет кратных корней[1]) [в равенстве (1) $m_1=m_2=\ldots=m_s=1$; $s=m$], то для того, чтобы $f(A)$ имело смысл, достаточно, чтобы функция $f(\lambda)$ была определена в характеристических точках $\lambda_1, \lambda_2, \ldots, \lambda_m$. Если же $\psi(\lambda)$ имеет кратные корни, то в некоторых характеристических точках должны быть определены и производные от $f(\lambda)$ до известного порядка [см. (6)].

Пример 1. Рассмотрим матрицу[2])

$$H=\begin{Vmatrix} \overbrace{0 & 1 & 0 \ldots 0}^{n} \\ 0 & 0 & 1 \ldots 0 \\ \cdot & \cdot & \cdot \ldots \cdot \\ 0 & 0 & 0 \ldots 1 \\ 0 & 0 & 0 \ldots 0 \end{Vmatrix}.$$

Для нее минимальным многочленом будет λ^n. Поэтому значениями $f(\lambda)$ на спектре H будут числа $f(0)$, $f'(0)$, ..., $f^{(n-1)}(0)$, и многочлен $r(\lambda)$ будет иметь вид

$$r(\lambda)=f(0)+\frac{f'(0)}{1!}\lambda+\ldots+\frac{f^{(n-1)}(0)}{(n-1)!}\lambda^{n-1}.$$

Таким образом,

$$f(H)=f(0)E+\frac{f'(0)}{1!}H+\ldots+\frac{f^{(n-1)}(0)}{(n-1)!}H^{n-1}=\begin{Vmatrix} f(0) & \frac{f'(0)}{1!} & \cdots & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \frac{f'(0)}{1!} \\ \cdot & & & \cdot \\ 0 & 0 & \cdots & f(0) \end{Vmatrix}.$$

Пример 2. Рассмотрим матрицу

$$J=\begin{Vmatrix} \overbrace{\lambda_0 & 1 & 0 \ldots 0}^{n} \\ 0 & \lambda_0 & 1 \ldots 0 \\ \cdot & \cdot & \cdot \ldots \cdot \\ \cdot & & \cdot \ldots \cdot \\ \cdot & & \cdot \ldots \cdot \\ 0 & 0 & 0 \ldots 1 \\ 0 & 0 & 0 \ldots \lambda_0 \end{Vmatrix}.$$

Заметим, что $J=\lambda_0 E+H$ и, следовательно, $J-\lambda_0 E=H$. Минимальным многочленом для J, очевидно, будет $(\lambda-\lambda_0)^n$. Интерполяционный многочлен $r(\lambda)$ для функции $f(\lambda)$ определится равенством

$$r(\lambda)=f(\lambda_0)+\frac{f'(\lambda_0)}{1!}(\lambda-\lambda_0)+\ldots+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(\lambda-\lambda_0)^{n-1}.$$

---

[1]) В гл. VI будет выяснено, что в этом и только в этом случае матрица A является матрицей простой структуры (см. гл. III, § 8).

[2]) Свойства матрицы H разобраны в примере на стр. 24—25.

Поэтому

$$f(J)=r(J)=f(\lambda_0)E+\frac{f'(\lambda_0)}{1!}H+\ldots+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}H^{n-1}=$$

$$=\begin{Vmatrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdots & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & f(\lambda_0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{f'(\lambda_0)}{1!} \\ 0 & 0 & \cdots & f(\lambda_0) \end{Vmatrix}.$$

Отметим три свойства функции от матрицы.

1° Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы n-го порядка A, то $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ — полная система характеристических чисел матрицы $f(A)$.

В частном случае, когда $f(\lambda)$ — многочлен, это предложение было доказано на стр. 94. Доказательство для общего случая сводится к этому частному случаю, поскольку (в силу определения 1') $f(A)=r(A)$ и $f(\lambda_i)=r(\lambda_i)$ $(i=1, \ldots, n)$, где $r(\lambda)$ — интерполяционный многочлен Лагранжа — Сильвестра для функции $f(\lambda)$.

2° *Если две матрицы A и B подобны и матрица T преобразует A в B:*

$$B=T^{-1}AT,$$

то матрицы $f(A)$ и $f(B)$ подобны и та же матрица T преобразует $f(A)$ в $f(B)$

$$f(B)=T^{-1}f(A)T.$$

Действительно, две подобные матрицы имеют одинаковые минимальные многочлены [1]) и, следовательно, функция $f(\lambda)$ принимает одни и те же значения как на спектре матрицы A, так и на спектре матрицы B. Поэтому существует интерполяционный многочлен $r(\lambda)$ такой, что $f(A)=r(A)$, $f(B)=r(B)$. Но тогда из равенства $r(B)=T^{-1}r(A)T$ (см. сноску) следует:

$$f(B)=T^{-1}f(A)T.$$

3° *Если A — квазидиагональная матрица*

$$A=\{A_1, A_2, \ldots, A_u\},$$

то

$$f(A)=\{f(A_1), f(A_2), \ldots, f(A_u)\}.$$

Обозначим через $r(\lambda)$ интерполяционный многочлен Лагранжа — Сильвестра для функции $f(\lambda)$ на спектре матрицы A. Тогда, как легко видеть,

$$f(A)=r(A)=\{r(A_1), r(A_2), \ldots, r(A_u)\}. \tag{7}$$

С другой стороны, минимальный многочлен $\psi(\lambda)$ для A является аннулирующим многочленом для каждой из матриц $A_1, A_2, \ldots, A_u$. Поэтому из равенства

$$f(\Lambda_A)=r(\Lambda_A)$$

следует:

$$f(\Lambda_{A_1})=r(\Lambda_{A_1}), \ldots, f(\Lambda_{A_u})=r(\Lambda_{A_u}).$$

[1]) Из $B=T^{-1}AT$ следует $B^k=T^{-1}A^kT$ $(k=0, 1, 2, \ldots)$. Отсюда для любого многочлена $g(\lambda)$ имеем $g(B)=T^{-1}g(A)T$. Поэтому из $g(A)=0$ вытекает $g(B)=0$ и наоборот.

Поэтому

$$f(A_1)=r(A_1),\ \ldots,\ f(A_u)=r(A_u),$$

и равенство (7) можно записать так:

$$f(A)=\{f(A_1),\ f(A_2),\ \ldots,\ f(A_u)\}. \tag{8}$$

Пример 1. Если матрица простой структуры

$$A=T\{\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n\}T^{-1},$$

то

$$f(A)=T\{f(\lambda_1),\ f(\lambda_2),\ \ldots,\ f(\lambda_n)\}T^{-1}.$$

$f(A)$ имеет смысл, если функция $f(\lambda)$ определена в точках $\lambda_1\ \lambda_2,\ \ldots,\ \lambda_n$.

Пример 2. Матрица J имеет следующий квазидиагональный вид:

$$J=\left\|\begin{array}{cccc}
\overbrace{\begin{array}{ccccc}\lambda_1 & 1 & 0 & \ldots & 0\\ 0 & \lambda_1 & 1 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & 1\\ 0 & 0 & 0 & \ldots & \lambda_1\end{array}}^{\nu_1} & & & \\
 & \ddots & & \\
 & & \overbrace{\begin{array}{ccccc}\lambda_u & 1 & 0 & \ldots & 0\\ 0 & \lambda_u & 1 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & 1\\ 0 & 0 & 0 & \ldots & \lambda_u\end{array}}^{\nu_u} &
\end{array}\right\|.$$

В недиагональных блоках все элементы равны нулю. Согласно формуле (8) (см. также пример на стр. 105)

$$f(J)=\left\|\begin{array}{cccc}
\begin{array}{cccc}f(\lambda_1) & \frac{f'(\lambda_1)}{1!} & \ldots & \frac{f^{(\nu_1-1)}(\lambda_1)}{(\nu_1-1)!}\\ 0 & f(\lambda_1) & \ldots & \vdots\\ \vdots & \ddots & \ddots & \frac{f'(\lambda_1)}{1!}\\ 0 & 0 & \ldots & f(\lambda_1)\end{array} & & & \\
 & \ddots & & \\
 & & \begin{array}{cccc}f(\lambda_u) & \frac{f'(\lambda_u)}{1!} & \ldots & \frac{f^{(\nu_u-1)}(\lambda_u)}{(\nu_u-1)!}\\ 0 & f(\lambda_u) & \ldots & \vdots\\ \vdots & \ddots & \ddots & \frac{f'(\lambda_u)}{1!}\\ 0 & 0 & \ldots & f(\lambda_u)\end{array} &
\end{array}\right\|.$$

Здесь, как и в матрице J, все элементы в недиагональных блоках равны нулю [1].

[1] В дальнейшем (гл. VI, § 6 или гл. VII, § 7) будет установлено, что произвольная матрица $A=\|a_{ik}\|_1^n$ всегда подобна некоторой матрице вида J: $A=TJT^{-1}$. Поэтому (см. 2° на стр. 106) всегда $f(A)=Tf(J)T^{-1}$.

§ 2. Интерполяционный многочлен Лагранжа — Сильвестра

1. Рассмотрим сначала тот случай, когда характеристическое уравнение $|\lambda E - A| = 0$ не имеет кратных корней. Корни этого уравнения — характеристические числа матрицы A — обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$. Тогда

$$\psi(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n),$$

и условия (6) записываются так:

$$r(\lambda_k) = f(\lambda_k) \qquad (k = 1, 2, \ldots, n).$$

В этом случае $r(\lambda)$ является обычным интерполяционным многочленом Лагранжа для функции $f(\lambda)$ в точках $\lambda_1, \lambda_2, \ldots, \lambda_n$:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda-\lambda_1)\ldots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\ldots(\lambda-\lambda_n)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

Согласно определению 1′

$$f(A) = r(A) = \sum_{k=1}^{n} \frac{(A-\lambda_1 E)\ldots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\ldots(A-\lambda_n E)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

2. Допустим теперь, что характеристический многочлен имеет кратные корни, но минимальный многочлен, являющийся делителем характеристического, имеет только простые корни [1]):

$$\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_m).$$

В этом случае (как и в предыдущем) все показатели m_k в (1) равны единице, и равенства (6) принимают вид

$$r(\lambda_k) = f(\lambda_k) \qquad (k = 1, 2, \ldots, m).$$

$r(\lambda)$ снова является обычным интерполяционным многочленом Лагранжа и

$$f(A) = \sum_{k=1}^{m} \frac{(A-\lambda_1 E)\ldots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\ldots(A-\lambda_m E)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_m)} f(\lambda_k).$$

3. Рассмотрим общий случай:

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s} \qquad (m_1+m_2+\ldots+m_s = m).$$

Представим правильно-дробную функцию $\frac{r(\lambda)}{\psi(\lambda)}$ в виде суммы простых дробей:

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \left[\frac{\alpha_{k1}}{(\lambda-\lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda-\lambda_k)^{m_k-1}} + \ldots + \frac{\alpha_{k,m_k}}{\lambda-\lambda_k} \right], \tag{9}$$

где α_{kj} $(j = 1, 2, \ldots, m_k;\ k = 1, 2, \ldots, s)$ — некоторые числа.

Для определения числителей простых дробей α_{kj} умножим обе части этого равенства на $(\lambda-\lambda_k)^{m_k}$ и обозначим через $\psi^k(\lambda)$ многочлен $\frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}$.

[1]) См. сноску [1]) на стр. 105.

Получим:

$$\frac{r(\lambda)}{\psi^k(\lambda)} = \alpha_{k1} + \alpha_{k2}(\lambda-\lambda_k) + \ldots + \alpha_{km_k}(\lambda-\lambda_k)^{m_k-1} + \\ + (\lambda-\lambda_k)^{m_k}\varrho(\lambda) \quad (k=1, 2, \ldots, s), \tag{10}$$

где $\varrho(\lambda)$ — рациональная функция, регулярная при $\lambda=\lambda_k$ [1]).

Отсюда

$$\left.\begin{aligned} \alpha_{k1} &= \left[\frac{r(\lambda)}{\psi^k(\lambda)}\right]_{\lambda=\lambda_k}, \\ \alpha_{k2} &= \left[\frac{r(\lambda)}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} = r(\lambda_k)\left[\frac{1}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} + r'(\lambda_k)\frac{1}{\psi^k(\lambda_k)}, \ldots \quad (k=1,2,\ldots,s). \end{aligned}\right\} \tag{11}$$

Формулы (11) показывают, что числители α_{kj} в правой части равенства (9) выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, а эти значения нам известны: они равны соответствующим значениям функции $f(\lambda)$ и ее производных. Поэтому

$$\alpha_{k1} = \frac{f(\lambda_k)}{\psi^k(\lambda_k)}, \quad \alpha_{k2} = f(\lambda_k)\left[\frac{1}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} + f'(\lambda_k)\frac{1}{\psi^k(\lambda_k)}, \ldots \\ (k=1, 2, \ldots, s). \tag{12}$$

Формулы (12) можно еще сокращенно записать так:

$$\alpha_{kj} = \frac{1}{(j-1)!}\left[\frac{f(\lambda)}{\psi^k(\lambda)}\right]^{(j-1)}_{\lambda=\lambda_k} \quad (j=1, 2, \ldots, m_k; \quad k=1, 2, \ldots, s). \tag{13}$$

После того как все α_{kj} найдены, мы определяем $r(\lambda)$ из следующей формулы, которая получается умножением обеих частей равенства (9) на $\psi(\lambda)$:

$$r(\lambda) = \sum_{k=1}^{s}\left[\alpha_{k1} + \alpha_{k2}(\lambda-\lambda_k) + \ldots + \alpha_{k,m_k}(\lambda-\lambda_k)^{m_k-1}\right]\psi^k(\lambda). \tag{14}$$

В этой формуле выражение в квадратных скобках, стоящее в качестве множителя перед $\psi^k(\lambda)$, в силу (13) равно сумме первых m_k членов разложения Тейлора по степеням $(\lambda-\lambda_k)$ для функции $f(\lambda)/\psi^k(\lambda)$.

Пример.

$$\psi(\lambda) = (\lambda-\lambda_1)^2(\lambda-\lambda_2)^3 \quad (m=5).$$

Тогда

$$\frac{r(\lambda)}{\psi(\lambda)} = \frac{\alpha}{(\lambda-\lambda_1)^2} + \frac{\beta}{\lambda-\lambda_1} + \frac{\gamma}{(\lambda-\lambda_2)^3} + \frac{\delta}{(\lambda-\lambda_2)^2} + \frac{\varepsilon}{\lambda-\lambda_2}.$$

Отсюда

$$r(\lambda) = [\alpha+\beta(\lambda-\lambda_1)](\lambda-\lambda_2)^3 + [\gamma+\delta(\lambda-\lambda_2)+\varepsilon(\lambda-\lambda_2)^2](\lambda-\lambda_1)^2$$

и, следовательно,

$$r(A) = [\alpha E+\beta(A-\lambda_1 E)](A-\lambda_2 E)^3 + [\gamma E+\delta(A-\lambda_2 E)+\varepsilon(A-\lambda_2 E)^2](A-\lambda_1 E)^2.$$

---

[1]) То есть не обращающаяся в ∞ при $\lambda=\lambda_k$.

$\alpha, \beta, \gamma, \delta, \varepsilon$ найдем из следующих формул:

$$\alpha=\frac{f(\lambda_1)}{(\lambda_1-\lambda_2)^3}, \quad \beta=-\frac{3}{(\lambda_1-\lambda_2)^4}f(\lambda_1)+\frac{1}{(\lambda_1-\lambda_2)^3}f'(\lambda_1),$$

$$\gamma=\frac{f(\lambda_2)}{(\lambda_2-\lambda_1)^2}, \quad \delta=-\frac{2}{(\lambda_2-\lambda_1)^3}f(\lambda_2)+\frac{1}{(\lambda_2-\lambda_1)^2}f'(\lambda_2),$$

$$\varepsilon=\frac{3}{(\lambda_2-\lambda_1)^4}f(\lambda_2)-\frac{2}{(\lambda_2-\lambda_1)^3}f'(\lambda_2)+\frac{1}{2}\frac{1}{(\lambda_2-\lambda_1)^2}f''(\lambda_2).$$

Примечание 1. Интерполяционный многочлен Лагранжа — Сильвестра может быть получен предельным переходом из интерполяционного многочлена Лагранжа.

Пусть

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s}\left(m=\sum_{k=1}^{s}m_k\right).$$

Обозначим через $L(\lambda)$ интерполяционный многочлен Лагранжа, построенный для m точек

$$\lambda_1^{(1)}, \lambda_1^{(2)}, \ldots, \lambda_1^{(m_1)}; \lambda_2^{(1)}, \lambda_2^{(2)}, \ldots, \lambda_2^{(m_2)}; \ldots; \lambda_s^{(1)}, \lambda_s^{(2)}, \ldots, \lambda_s^{(m_s)}.$$

Тогда нетрудно показать, что искомый многочлен Лагранжа — Сильвестра определяется формулой

$$r(\lambda)=\lim_{\substack{\lambda_1^{(1)},\ldots,\lambda_1^{(m_1)}\to\lambda_1 \\ \cdots\cdots\cdots \\ \lambda_s^{(1)},\ldots,\lambda_s^{(m_s)}\to\lambda_s}} L(\lambda).$$

Примечание 2. Пусть $A=\|a_{ik}\|$ — *вещественная матрица*, т. е. матрица с вещественными элементами. Тогда минимальный многочлен $\psi(\lambda)$ имеет вещественные коэффициенты [1]) и его корни, т. е. характеристические числа λ_i, либо вещественны, либо попарно комплексно сопряжены, причем, если $\lambda_g=\overline{\lambda_h}$, то соответствующие кратности равны: $m_g=m_h$. Условимся говорить, что *функция* $f(\lambda)$ *вещественна на спектре матрицы* A, если для вещественного λ_i все ее значения на спектре $f(\lambda_i)$, $f'(\lambda_i)$, ... вещественны, а для двух комплексно сопряженных характеристических чисел λ_h и $\lambda_g=\overline{\lambda_h}$ соответствующие значения на спектре комплексно сопряжены: $f(\lambda_g)=\overline{f(\lambda_h)}$, $f'(\lambda_g)=\overline{f'(\lambda_h)}$, ... [2]). В этом случае $f(A)$ — *вещественная* матрица. Действительно, в данном случае согласно формулам (12) $\alpha_{i1}, \alpha_{i2}, \ldots$ — вещественные числа и $\alpha_{g1}=\overline{\alpha_{h1}}$, $\alpha_{g2}=\overline{\alpha_{h2}}, \ldots$; при этом для вещественного λ_i многочлен $\psi^i(\lambda)=\dfrac{\psi(\lambda)}{(\lambda-\lambda_i)^{m_i}}$ имеет вещественные коэффициенты, а коэффициенты многочленов $\psi^h(\lambda)$ и $\psi^g(\lambda)$ (при $\lambda_g=\overline{\lambda_h}$) — комплексно сопряжены. Поэтому в силу формулы (14) интерполяционный многочлен $r(\lambda)$ имеет вещественные коэффициенты. Но тогда $r(A)$, а значит и $f(A)=r(A)$, — вещественная матрица.

---

[1]) Это следует непосредственно из определения минимального многочлена либо из формулы (54) на стр. 100.

[2]) Функция, представляемая суммой степенного ряда с вещественными коэффициентами, вещественна на спектре любой матрицы, характеристические числа которой находятся внутри круга сходимости данного ряда.

§ 3. Другие формы определения $f(A)$. Компоненты матрицы A

Вернемся к формуле (14) для $r(\lambda)$. Подставляя в нее выражения (12) для коэффициентов α и объединяя члены, содержащие одно и то же значение функции $f(\lambda)$ или какой-либо ее производной, мы представим $r(\lambda)$ в виде

$$r(\lambda)=\sum_{k=1}^{s}[f(\lambda_k)\varphi_{k1}(\lambda)+f'(\lambda_k)\varphi_{k2}(\lambda)+\ldots+f^{(m_k-1)}(\lambda_k)\varphi_{k,m_k}(\lambda)]. \tag{15}$$

Здесь $\varphi_{kj}(\lambda)$ $(j=1,\ 2,\ \ldots,m_k;\ k=1,\ 2,\ \ldots,s)$ — легко вычисляемые многочлены от λ степени $<m$. Эти многочлены вполне определяются заданием $\psi(\lambda)$ и не зависят от выбора функции $f(\lambda)$. Число этих многочленов равно числу значений функции $f(\lambda)$ на спектре матрицы A, т. е. равно m [m — степень минимального многочлена $\psi(\lambda)$]. Функция $\varphi_{kj}(\lambda)$ представляет собой интерполяционный многочлен Лагранжа — Сильвестра для функции, у которой все значения на спектре матрицы A равны нулю, за исключением одного $f^{(j-1)}(\lambda_k)$, равного 1.

Из формулы (15) следует основная формула для $f(A)$:

$$f(A)=\sum_{k=1}^{s}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ldots+f^{(m_k-1)}(\lambda_k)Z_{km_k}], \tag{16}$$

где

$$Z_{kj}=\varphi_{kj}(A)\qquad(j=1,\ 2,\ \ldots,m_k;\ k=1,\ 2,\ \ldots,s). \tag{17}$$

Матрицы Z_{kj} вполне определяются заданием матрицы A и не зависят от выбора функции $f(\lambda)$. В правой части формулы (16) функция $f(\lambda)$ представлена только своими значениями на спектре матрицы A.

Матрицы Z_{kj} $(j=1,\ 2,\ \ldots,m_k;\ k=1,\ 2,\ \ldots,s)$ будем называть *составляющими матрицами* или *компонентами* данной матрицы A.

Компоненты матрицы A всегда линейно независимы.

Действительно, пусть

$$\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}Z_{kj}=\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}\varphi_{kj}(A)=0. \tag{18}$$

Определим интерполяционный многочлен $r(\lambda)$ из m условий

$$r^{(j-1)}(\lambda_k)=c_{kj}\qquad(j=1,\ 2,\ \ldots,\ m_k;\ k=1,\ 2,\ \ldots,\ s). \tag{19}$$

Тогда, согласно формуле (15),

$$r(\lambda)=\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}\varphi_{kj}(\lambda). \tag{20}$$

Из сопоставления формул (18) и (19) находим

$$r(A)=0. \tag{21}$$

Но степень интерполяционного многочлена $r(\lambda)$, задаваемого формулой (20), меньше m, т. е. меньше степени минимального многочлена $\psi(\lambda)$. Поэтому из равенства (21) следует тождество

$$r(\lambda)=0.$$

Но тогда согласно (19)

$$c_{kj}=0 \qquad (j=1, 2, \ldots, m_k;\ k=1, 2, \ldots, s),$$

что и требовалось доказать[1]).

Из линейной независимости составляющих матриц Z_{kj} следует, между прочим, что ни одна из этих матриц не равна нулю. Заметим еще, что любые две из компонент Z_{kj} перестановочны между собой и с матрицей A, поскольку все они суть скалярные многочлены от A.

Формулой (16) для $f(A)$ особенно удобно пользоваться тогда, когда приходится иметь дело с несколькими функциями от одной и той же матрицы A, либо когда функция $f(\lambda)$ зависит не только от λ, но и от некоторого параметра t. В последнем случае в правой части формулы (16) компоненты Z_{kj} не зависят от t и параметр t входит только в скалярные коэффициенты при этих матрицах.

В примере на стр. 109, где $\psi(\lambda)=(\lambda-\lambda_1)^2(\lambda-\lambda_2)^3$, мы можем $r(\lambda)$ представить в виде

$$r(\lambda)=f(\lambda_1)\varphi_{11}(\lambda)+f'(\lambda_1)\varphi_{12}(\lambda)+f(\lambda_2)\varphi_{21}(\lambda)+f'(\lambda_2)\varphi_{22}(\lambda)+f''(\lambda_2)\varphi_{23}(\lambda),$$

где

$$\varphi_{11}(\lambda)=\left(\frac{\lambda-\lambda_2}{\lambda_1-\lambda_2}\right)^3\left[1-\frac{3(\lambda-\lambda_1)}{\lambda_1-\lambda_2}\right], \quad \varphi_{12}(\lambda)=\frac{(\lambda-\lambda_1)(\lambda-\lambda_2)^3}{(\lambda_1-\lambda_2)^3},$$

$$\varphi_{21}(\lambda)=\left(\frac{\lambda-\lambda_1}{\lambda_2-\lambda_1}\right)^2\left[1-\frac{2(\lambda-\lambda_2)}{\lambda_2-\lambda_1}+\frac{3(\lambda-\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\right],$$

$$\varphi_{22}(\lambda)=\frac{(\lambda-\lambda_1)^2(\lambda-\lambda_2)}{(\lambda_2-\lambda_1)^2}\left[1-\frac{2(\lambda-\lambda_2)}{\lambda_2-\lambda_1}\right],$$

$$\varphi_{23}(\lambda)=\frac{(\lambda-\lambda_1)^2(\lambda-\lambda_2)^2}{2(\lambda_2-\lambda_1)^2}.$$

Поэтому

$$f(A)=f(\lambda_1)Z_{11}+f'(\lambda_1)Z_{12}+f(\lambda_2)Z_{21}+f'(\lambda_2)Z_{22}+f''(\lambda)Z_{23},$$

где

$$Z_{11}=\varphi_{11}(A)=\frac{1}{(\lambda_1-\lambda_2)^3}(A-\lambda_2E)^3\left[E-\frac{3}{\lambda_1-\lambda_2}(A-\lambda_1E)\right].$$

$$Z_{12}=\varphi_{12}(A)=\frac{1}{(\lambda_1-\lambda_2)^3}(A-\lambda_1E)(A-\lambda_2E)^3 \text{ и т. д.}$$

Если дана матрица A, то для конкретного нахождения компонент этой матрицы можно в основной формуле (16) положить $f(\mu)=\frac{1}{\lambda-\mu}$, где λ — некоторый параметр. Тогда получим:

$$(\lambda E-A)^{-1}=\frac{C(\lambda)}{\psi(\lambda)}=\sum_{k=1}^{s}\left[\frac{Z_{k1}}{\lambda-\lambda_k}+\frac{1!\,Z_{k2}}{(\lambda-\lambda_k)^2}+\ldots+\frac{(m_k-1)!\,Z_{km_k}}{(\lambda-\lambda_k)^{m_k}}\right], \qquad (22)$$

где $C(\lambda)$ — приведенная присоединенная матрица для $\lambda E-A$ (гл. IV, § 6)[2]).

---

[1]) Из доказанной линейной независимости матриц $Z_{kj}=\varphi_{kj}(A)$ следует также линейная независимость многочленов $\varphi_{kj}(\lambda)$; $j=1, 2, \ldots, m_k$; $k=1, 2, \ldots, s$.

[2]) При $f(\mu)=\frac{1}{\lambda-\mu}$ имеем $f(A)=(\lambda E-A)^{-1}$. Действительно, $f(A)=r(A)$, где $r(\mu)$ — интерполяционный многочлен Лагранжа—Сильвестра. Из того, что $f(\mu)$ и $r(\mu)$ совпадают на спектре матрицы A, следует, что на этом спектре совпадают $(\lambda-\mu)r(\mu)$ и $(\lambda-\mu)f(\mu)=1$. Отсюда $(\lambda E-A)r(A)=(\lambda E-A)f(A)=E$.

Матрицы $(j-1)!\,Z_{kj}$ являются числителями простейших дробей в разложении (22), и потому по аналогии с разложением (9) эти числители могут быть выражены через значения $C(\lambda)$ на спектре матрицы A по формулам, подобным (11):

$$(m_k-1)!\,Z_{km_k}=\frac{C(\lambda_k)}{\psi^k(\lambda_k)},\qquad (m_k-2)!\,Z_{k,m_k-1}=\left[\frac{C(\lambda)}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k},\ \ldots$$

Отсюда

$$Z_{kj}=\frac{1}{(j-1)!\,(m_k-j)!}\left[\frac{C(\lambda)}{\psi^k(\lambda)}\right]^{(m_k-j)}_{\lambda=\lambda_k}\quad (j=1,2,\ldots,m_k;\ k=1,2,\ldots,s). \tag{23}$$

Подставляя в (16) вместо составляющих матриц их выражения (22), мы можем основную формулу (16) представить в виде

$$f(A)=\sum_{k=1}^{s}\frac{1}{(m_k-1)!}\left[\frac{C(\lambda)}{\psi^k(\lambda)}f(\lambda)\right]^{(m_k-1)}_{\lambda=\lambda_k}. \tag{24}$$

Пример 1.

$$A=\left\|\begin{matrix}2&-1&1\\0&1&1\\-1&1&1\end{matrix}\right\|\begin{matrix}2^{1})\\2,\\1\end{matrix}\quad \lambda E-A=\left\|\begin{matrix}\lambda-2&1&-1\\0&\lambda-1&-1\\1&-1&\lambda-1\end{matrix}\right\|.$$

В данном случае $\Delta(\lambda)=|\lambda E-A|=(\lambda-1)^2(\lambda-2)$. Поскольку минор элемента [1, 2] в $\lambda E-A$ равен 1, то $D_2(\lambda)=1$, и потому

$$\psi(\lambda)=\Delta(\lambda)=(\lambda-1)^2(\lambda-2)=\lambda^3-4\lambda^2+5\lambda-2,$$

$$\Psi(\lambda,\mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}=\mu^2+(\lambda-4)\mu+\lambda^2-4\lambda+5$$

и

$$C(\lambda)=\Psi(\lambda E,A)=A^2+(\lambda-4)A+(\lambda^2-4\lambda+5)E=$$

$$=\left\|\begin{matrix}3&-2&2\\-1&2&2\\-3&3&1\end{matrix}\right\|\begin{matrix}3\\3+(\lambda-4)\\1\end{matrix}\left\|\begin{matrix}2&-1&1\\0&1&1\\-1&1&1\end{matrix}\right\|+(\lambda^2-4\lambda+5)\left\|\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}\right\|.$$

Основная формула в данном случае имеет вид

$$f[A]=f(1)Z_1+f'(1)Z_2+f(2)Z_3. \tag{25}$$

Полагая здесь $f(\mu)=\frac{1}{\lambda-\mu}$, находим:

$$(\lambda E-A)^{-1}=\frac{C(\lambda)}{\psi(\lambda)}=\frac{Z_{11}}{\lambda-1}+\frac{Z_{12}}{(\lambda-1)^2}+\frac{Z_{22}}{\lambda-2},$$

откуда

$$Z_1=-C(1)-C'(1),\quad Z_2=-C(1),\quad Z_3=C(2).$$

Пользуясь приведенным выше выражением для $C(\lambda)$, вычисляем Z_1, Z_2, Z_3 и подставляем полученные результаты в (25):

$$f(A)=f(1)\left\|\begin{matrix}1&0&0\\1&0&0\\1&-1&1\end{matrix}\right\|+f'(1)\left\|\begin{matrix}1&-1&1\\1&-1&1\\0&0&0\end{matrix}\right\|+f(2)\left\|\begin{matrix}0&0&0\\-1&1&0\\-1&1&0\end{matrix}\right\|=$$

$$=\left\|\begin{matrix}f(1)+f'(1)&-f'(1)&f'(1)\\f(1)+f'(1)-f(2)&-f'(1)+f(2)&f'(1)\\f(1)-f(2)&-f(1)+f(2)&f(1)\end{matrix}\right\|.$$

[1]) Курсивом набраны элементы контрольного суммарного столбца. Умножая строки матрицы A на суммарный столбец матрицы B, получаем элементы суммарного столбца произведения AB.

Пример 2. Покажем, как можно определить $f(A)$, исходя только из основной формулы. Пусть снова

$$A=\begin{Vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{Vmatrix}, \quad \psi(\lambda)=(\lambda-1)^2(\lambda-2).$$

Тогда

$$f(A)=f(1)\,Z_1+f'(1)\,Z_2+f(2)\,Z_3. \tag{25'}$$

Подставим в формулу (25') вместо $f(\lambda)$ последовательно 1, $\lambda-1$, $(\lambda-1)^2$:

$$Z_1+Z_3=E=\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$Z_2+Z_3=A-E=\begin{Vmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{Vmatrix}\begin{matrix} 1 \\ 1, \\ 0 \end{matrix}$$

$$Z_3=(A-E)^2=\begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix}\begin{matrix} 0 \\ 0. \\ 0 \end{matrix}$$

Вычитая из первых двух равенств почленно третье, определим все Z_k. Подставляя в (25'), получим выражение для $f(A)$.

Разобранные примеры иллюстрировали три способа практического нахождения $f(A)$. В первом способе мы находили интерполяционный многочлен $r(\lambda)$ и полагали $f(A)=r(A)$. Во втором способе мы, пользуясь разложением (22), выражали компоненты Z_{kj} в формуле (16) через значения приведенной присоединенной матрицы $C(\lambda)$ на спектре матрицы A. В третьем способе мы исходили из основной формулы (16) и подставляли в нее вместо $f(\lambda)$ последовательно некоторые простейшие многочлены; из полученных линейных уравнений определяли составляющие матрицы Z_{kj}.

Третий способ является, пожалуй, практически наиболее удобным. В общем виде его можно сформулировать так:

В формулу (16) вместо $f(\lambda)$ подставляем последовательно некоторые многочлены $g_1(\lambda)$, $g_2(\lambda)$, ..., $g_m(\lambda)$:

$$g_i(A)=\sum_{k=1}^{s}\left[g_i(\lambda_k)\,Z_{k1}+g_i'(\lambda_k)\,Z_{k2}+\ldots+g_i^{(m_k-1)}(\lambda_k)\,Z_{k,m_k}\right] \tag{26}$$
$$(i=1, 2, \ldots, m).$$

Из m уравнений (26) определяем m матриц Z_{kj} и подставляем полученные выражения в (16).

Результат исключения Z_{kj} из $(m+1)$-го равенства (26) и (16) может быть записан в виде

$$\begin{vmatrix} f(A) & f(\lambda_1) & \ldots & f^{(m_1-1)}(\lambda_1) & \ldots & f(\lambda_s) & \ldots & f^{(m_s-1)}(\lambda_s) \\ g_1(A) & g_1(\lambda_1) & \ldots & g_1^{(m_1-1)}(\lambda_1) & \ldots & g_1(\lambda_s) & \ldots & g_1^{(m_s-1)}(\lambda_s) \\ \cdot & \cdot & & & & & & \\ \cdot & \cdot & & & & & & \\ \cdot & \cdot & & & & & & \\ g_m(A) & g_m(\lambda_1) & \ldots & g_m^{(m_1-1)}(\lambda_1) & \ldots & g_m(\lambda_s) & \ldots & g_m^{(m_s-1)}(\lambda_s) \end{vmatrix}=0.$$

Разлагая этот определитель по элементам первого столбца, мы получим искомое выражение для $f(A)$. Здесь при $f(A)$ в качестве множителя будет стоять определитель $\Delta = |g_i^{(j)}(\lambda_k)|$ (в i-й строке определителя Δ стоят значения многочлена $g_i(\lambda)$ на спектре матрицы A; $i = 1, 2, \ldots, m$). Для того чтобы можно было определить $f(A)$, нужно, чтобы $\Delta \neq 0$. Это будет иметь место, если никакая линейная комбинация [1]) многочленов $g_1(\lambda)$, $g_2(\lambda)$, $\ldots$, $g_m(\lambda)$ не обращается сплошь в нуль на спектре матрицы A, т. е. не делится на $\psi(\lambda)$.

Условие $\Delta \neq 0$ всегда выполнено, если степени многочленов $g_1(\lambda)$, $g_2(\lambda)$, $\ldots$, $g_m(\lambda)$ соответственно равны $0, 1, 2, \ldots, m-1$ [2]).

В заключение отметим, что большие степени матрицы A^n удобно вычислять по основной формуле (16), заменяя в ней $f(\lambda)$ на λ^n [3]).

Пример. Дана матрица $A = \begin{Vmatrix} 5 & -4 \\ 4 & -3 \end{Vmatrix}$. Требуется вычислить элементы степени A^{100}. Минимальный многочлен $\psi(\lambda) = (\lambda-1)^2$.

Основная формула

$$f(A) = f(1)\, Z_1 + f'(1)\, Z_2.$$

Заменяя здесь $f(\lambda)$ на 1, а затем на $\lambda - 1$, получим:

$$Z_1 = E, \quad Z_2 = A - E.$$

Поэтому

$$f(A) = f(1)\, E + f'(1)\,(A - E).$$

Полагая $f(\lambda) = \lambda^{100}$, найдем:

$$A^{100} = E + 100\,(A-E) = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} + 100 \begin{Vmatrix} 4 & -4 \\ 4 & -4 \end{Vmatrix} = \begin{Vmatrix} 401 & -400 \\ 400 & -399 \end{Vmatrix}.$$

§ 4. Представление функций от матриц рядами

Пусть дана матрица $A = \| a_{ik} \|_1^n$ с минимальным многочленом $\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s} \left(m = \sum_{k=1}^{s} m_k\right)$. Пусть, далее, функция $f(\lambda)$ и последовательность функций $f_1(\lambda)$, $f_2(\lambda)$, $\ldots$, $f_p(\lambda)$, $\ldots$ определены на спектре матрицы A.

Мы будем говорить, что *последовательность функций* $f_p(\lambda)$ *при* $p \to \infty$ *стремится к некоторому пределу на спектре матрицы* A, если существуют пределы

$$\lim_{p\to\infty} f_p(\lambda_k), \quad \lim_{p\to\infty} f_p'(\lambda_k), \quad \ldots, \quad \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k) \qquad (k = 1, 2, \ldots, s).$$

Мы будем говорить, что *последовательность функций* $f_p(\lambda)$ *стремится при* $p \to \infty$ *к функции* $f(\lambda)$ *на спектре матрицы* A, и будем писать:

$$\lim_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A),$$

если

$$\lim_{p\to\infty} f_p(\lambda_k) = f(\lambda_k), \quad \lim_{p\to\infty} f_p'(\lambda_k) = f'(\lambda_k), \quad \ldots, \quad \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k)$$

$$(k = 1, 2, \ldots, s).$$

---

[1]) С не равными одновременно нулю коэффициентами.

[2]) В последнем примере $m = 3$, $g_1(\lambda) = 1$, $g_2(\lambda) = \lambda - 1$, $g_3(\lambda) = (\lambda-1)^2$.

[3]) Формулу (16) можно использовать и для вычисления обратной матрицы A^{-1}, полагая $f(\lambda) = \frac{1}{\lambda}$ или, что то же, полагая $\lambda = 0$ в формуле (22).

Основная формула

$$f(A)=\sum_{k=1}^{s}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ldots+f^{(m_k-1)}(\lambda_k)Z_{km_k}]$$

выражает $f(A)$ через значения $f(\lambda)$ на спектре матрицы A. Если рассматривать матрицу как вектор в пространстве n^2 измерений R_{n^2}, то из основной формулы в силу линейной независимости матриц Z_{kj} следует, что все $f(A)$ (при заданном A) образуют m-мерное подпространство в R_{n^2} с базисом Z_{kj} ($j=1, 2, \ldots, m_k$; $k=1, 2, \ldots, s$). В этом базисе «вектор» $f(A)$ имеет своими координатами m значений функции $f(\lambda)$ на спектре матрицы A.

Эти соображения делают совершенно прозрачной следующую теорему:

Теорема 1. *Для того чтобы последовательность матриц $f_p(A)$ при $p\to\infty$ стремилась к некоторому пределу, необходимо и достаточно, чтобы последовательность $f_p(\lambda)$ при $p\to\infty$ на спектре матрицы A стремилась к пределу, т. е. пределы*

$$\lim_{p\to\infty}f_p(A),\quad \lim_{p\to\infty}f_p(\Lambda_A)$$

всегда существуют одновременно. При этом равенство

$$\lim_{p\to\infty}f_p(\Lambda_A)=f(\Lambda_A) \tag{27}$$

влечет за собой равенство

$$\lim_{p\to\infty}f_p(A)=f(A) \tag{28}$$

и наоборот.

Доказательство 1. Если значения $f_p(\lambda)$ на спектре матрицы A при $p\to\infty$ стремится к предельным значениям, то из формулы

$$f_p(A)=\sum_{k=1}^{s}[f_p(\lambda_k)Z_{k1}+f_p'(\lambda_k)Z_{k2}+\ldots+f_p^{(m_k-1)}(\lambda_k)Z_{km_k}] \tag{29}$$

следует существование предела $\lim\limits_{p\to\infty}f_p(A)$. На основании этой же формулы и формулы (16) из (27) вытекает (28).

2. Обратно, пусть существует $\lim\limits_{p\to\infty}f_p(A)$. Так как m составляющих матриц Z линейно независимы, то мы можем из (29) выразить (в виде линейных форм) m значений $f_p(\lambda)$ на спектре матрицы A через m элементов матрицы $f_p(A)$. Отсюда следует существование предела $\lim\limits_{p\to\infty}f_p(\Lambda_A)$, и равенство (27) имеет место при наличии равенства (28).

Согласно установленной теореме, если последовательность многочленов $g_p(\lambda)$ ($p=1, 2, \ldots$) стремится к функции $f(\lambda)$ на спектре матрицы A, то

$$\lim_{p\to\infty}g_p(A)=f(A).$$

Эта формула подчеркивает естественность и общность данного нами определения $f(A)$. $f(A)$ всегда получается предельным переходом из $g_p(A)$ при $p\to\infty$, если только последовательность многочленов $g_p(\lambda)$ сходится к $f(\lambda)$ на спектре матрицы A. Последнее условие является необходимым для существования предела $\lim\limits_{p\to\infty}g_p(A)$ при $p\to\infty$.

Условимся говорить, что *ряд* $\sum_{p=0}^{\infty} u_p(\lambda)$ *сходится на спектре матрицы* A *к функции* $f(\lambda)$, и будем писать:

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A), \tag{30}$$

если все фигурирующие здесь функции определены на спектре матрицы A и имеют место равенства

$$f(\lambda_k) = \sum_{p=0}^{\infty} u_p(\lambda_k), \quad f'(\lambda_k) = \sum_{p=0}^{\infty} u_p'(\lambda_k), \ \ldots, \ f^{(m_k-1)}(\lambda_k) = \sum_{p=0}^{\infty} u_p^{(m_k-1)}(\lambda_k)$$
$$(k = 1, 2, \ldots, s),$$

причем в правых частях этих равенств стоят сходящиеся ряды. Другими словами, если положить

$$s_p(\lambda) = \sum_{q=0}^{p} u_q(\lambda) \qquad (p = 0, 1, 2, \ldots),$$

то равенство (30) равносильно равенству

$$f(\Lambda_A) = \lim_{p\to\infty} s_p(\Lambda_A). \tag{31}$$

Очевидно, что доказанной теореме можно дать следующую эквивалентную формулировку:

Теорема 1′. *Для того чтобы ряд* $\sum_{p=0}^{\infty} u_p(A)$ *сходился к некоторой матрице, необходимо и достаточно, чтобы ряд* $\sum_{p=0}^{\infty} u_p(\lambda)$ *сходился на спектре матрицы* A. *При этом из равенства*

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A)$$

следует равенство

$$f(A) = \sum_{p=0}^{\infty} u_p(A)$$

и наоборот.

Пусть дан степенной ряд с кругом сходимости $|\lambda - \lambda_0| < R$ и суммой $f(\lambda)$:

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p \qquad (|\lambda - \lambda_0| < R). \tag{32}$$

Так как степенной ряд можно почленно дифференцировать любое число раз внутри круга сходимости, то ряд (32) сходится на спектре любой матрицы, характеристические числа которой попадают внутрь круга сходимости.

Таким образом, имеет место

Теорема 2. *Если функция* $f(\lambda)$ *разлагается в степенной ряд в круге* $|\lambda - \lambda_0| < r$,

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p, \tag{33}$$

то это разложение сохраняет силу, если скалярный аргумент λ заменить любой матрицей A, характеристические числа которой лежат внутри круга сходимости.

Примечание. В этой теореме можно допустить, чтобы характеристическое число λ_k матрицы A попало на периферию круга сходимости, но при этом нужно дополнительно потребовать, чтобы m_k-1 раз почленно продифференцированный ряд (33) в точке $\lambda=\lambda_k$ сходился. Отсюда, как известно, уже следует сходимость j раз продифференцированного ряда (33) в точке λ_k к $f^{(j)}(\lambda_k)$ для $j=0, 1, 2, \ldots, m_k-1$.

Из доказанной теоремы вытекают, например, следующие разложения [1]):

$$e^A=\sum_{p=0}^{\infty}\frac{A^p}{p!}\,,\quad \cos A=\sum_{p=0}^{\infty}\frac{(-1)^p}{(2p)!}A^{2p},\quad \sin A=\sum_{p=0}^{\infty}(-1)^p\frac{A^{2p+1}}{(2p+1)!}\,,$$

$$\operatorname{ch} A=\sum_{p=0}^{\infty}\frac{A^{2p}}{(2p)!}\,,\qquad \operatorname{sh} A=\sum_{p=0}^{\infty}\frac{A^{2p+1}}{(2p+1)!}\,,$$

$$(E-A)^{-1}=\sum_{p=0}^{\infty}A^p\qquad (|\lambda_k|<1;\ k=1, 2, \ldots, s),$$

$$\ln A=\sum_{p=1}^{\infty}\frac{(-1)^{p-1}}{p}(A-E)^p\qquad (|\lambda_k-1|<1;\ k=1, 2, \ldots, s)$$

(здесь под $\ln\lambda$ мы понимаем так называемое главное значение многозначной функции $\operatorname{Ln}\lambda$, т. е. ветвь, для которой $\operatorname{Ln}1=0$).

Формула (22) на стр. 112 позволяет легко распространить интегральную формулу Коши для аналитических функций на функции от матриц.

Рассмотрим в плоскости комплексного переменного λ правильную область, ограниченную замкнутым контуром Γ и содержащую внутри себя характеристические числа $\lambda_1, \ldots, \lambda_n$ матрицы A. Возьмем произвольную аналитическую функцию $f(\lambda)$, регулярную в этой области (включая границу Γ). Тогда по известным формулам Коши [2])

$$f(\lambda_k)=\frac{1}{2\pi i}\int_\Gamma\frac{f(\lambda)}{\lambda-\lambda_k}\,d\lambda,\quad f'(\lambda_k)=\frac{1}{2\pi i}\int_\Gamma\frac{f(\lambda)}{(\lambda-\lambda_k)^2}\,d\lambda,\quad \ldots,\quad f^{(m_k-1)}(\lambda_k)=$$

$$=\frac{(m_k-1)!}{2\pi i}\int_\Gamma\frac{f(\lambda)}{(\lambda-\lambda_k)^{m_k}}\,d\lambda.$$

$$(k=1, \ldots, s).$$

Умножая обе части матричного равенства (22) на $\frac{f(\lambda)}{2\pi i}$ и интегрируя вдоль Γ [3]), получим:

$$\frac{1}{2\pi i}\int_\Gamma(\lambda E-A)^{-1}f(\lambda)\,d\lambda=\sum_{k=1}^{s}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ \ldots\ +f^{(m_k-1)}(\lambda_k)Z_{km_k}],$$

[1]) Разложения в первых двух строках имеют место при произвольной матрице A.

[2]) См., например, И. И. Привалов, Введение в теорию функций комплексного переменного, Физматгиз, 1960, стр. 166.

[3]) Интеграл от матрицы определяем как результат «поэлементного» интегрирования. Поэтому

$$\int_\Gamma(\lambda E-A)^{-1}f(\lambda)\,d\lambda=\left\|\int_\Gamma(\lambda E-A)^{-1}_{ik}f(\lambda)\,d\lambda\right\|_{i,\,k=1}^{n},$$

где

$$(\lambda E-A)^{-1}_{ik}=\frac{b_{ik}}{\Delta(\lambda)}\quad (i, k=1, \ldots, n)\text{—элементы матрицы } (\lambda E-A)^{-1}$$

(см. гл. IV, § 3).

что, согласно основной формуле (16), и дает:

$$f(A)=\frac{1}{2\pi i}\int_{\Gamma}(\lambda E-A)^{-1}f(\lambda)\,d\lambda. \tag{34}$$

Эти же рассуждения показывают, что интеграл, стоящий в правой части равенства (34), равен нулю (при $f(A)\neq 0$!), если все характеристические числа матрицы A расположены вне контура Γ, и равен

$$\sum_{k=1}^{q}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ldots+f^{(m_k-1)}(\lambda_k)Z_{km_k}],\qquad (q<s),$$

если характеристические числа $\lambda_1, \ldots, \lambda_q$ расположены внутри, а $\lambda_{q+1}, \ldots, \lambda_n$ — вне Γ.

Формулу (34) можно принять за определение аналитической функции от матрицы.

§ 5. Некоторые свойства функций от матриц

В этом параграфе мы докажем несколько предложений, позволяющих распространить тождества, связывающие функции от скалярного переменного на матричные значения аргумента.

1° *Пусть* $G(u_1, u_2, \ldots, u_l)$ — *многочлен относительно* $u_1, u_2, \ldots, u_l$; $f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)$ — *функции от* λ, *определенные на спектре матрицы* A, *и*

$$g(\lambda)\equiv G[f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)]. \tag{35}$$

Тогда из

$$g(\Lambda_A)=0,$$

следует

$$G[f_1(A), f_2(A), \ldots, f_l(A)]=0.$$

Действительно, обозначим через $r_1(\lambda), r_2(\lambda), \ldots, r_l(\lambda)$ интерполяционные многочлены Лагранжа — Сильвестра для $f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)$ и положим:

$$h(\lambda)=G[r_1(\lambda), r_2(\lambda), \ldots, r_l(\lambda)].$$

Тогда из (35) вытекает:

$$h(\Lambda_A)=0.$$

Отсюда следует, что

$$G[f_1(A), f_2(A), \ldots, f_l(A)]=G[r_1(A), r_2(A), \ldots, r_l(A)]=h(A)=0,$$

что и требовалось доказать.

Согласно предложению 1° из тождества

$$\cos^2\lambda+\sin^2\lambda=1$$

следует для любой матрицы A

$$\cos^2 A+\sin^2 A=E$$

(в данном случае: $G(u_1, u_2)=u_1^2+u_2^2-1$, $f_1(\lambda)=\cos\lambda$, $f_2(\lambda)=\sin\lambda$).

Точно так же для любой матрицы A

$$e^A e^{-A}=E,$$

т. е.

$$e^{-A}=(e^A)^{-1}.$$

Далее, для любой матрицы A

$$e^{iA} = \cos A + i \sin A.$$

Пусть дана неособенная матрица A ($|A| \neq 0$). Обозначим через $\sqrt{\lambda}$ однозначную ветвь многозначной функции $\sqrt{\lambda}$, определенную в области, не содержащей нуля и содержащей все характеристические числа матрицы A. Тогда имеет смысл $\sqrt{A}$. При этом из $(\sqrt{\lambda})^2 - \lambda = 0$ будет следовать [1]):

$$(\sqrt{A})^2 = A.$$

Пусть $f(\lambda) = \frac{1}{\lambda}$, и $A = \| a_{ik} \|_1^n$ — неособенная матрица. Тогда функция $f(\lambda)$ определена на спектре матрицы A, и потому в равенстве

$$\lambda f(\lambda) = 1$$

можно заменить λ на A:

$$A \cdot f(A) = E,$$

т. е.

$$f(A) = A^{-1}\ ^2).$$

Обозначая через $r(\lambda)$ интерполяционный многочлен для функции $\frac{1}{\lambda}$, мы сможем обратную матрицу A^{-1} представить в виде многочлена от данной:

$$A^{-1} = r(A).$$

Рассмотрим рациональную функцию $\varrho(\lambda) = \frac{g(\lambda)}{h(\lambda)}$, где $g(\lambda)$ и $h(\lambda)$ — взаимно простые многочлены относительно λ. Эта функция определена на спектре матрицы A в том и только в том случае, если характеристические числа матрицы A не являются корнями многочлена $h(\lambda)$, т. е. [3]) если $|h(A)| \neq 0$. При выполнении этого условия мы можем в тождестве

$$\varrho(\lambda) h(\lambda) = g(\lambda)$$

заменить λ на A:

$$\varrho(A) h(A) = g(A).$$

Отсюда

$$\varrho(A) = g(A) [h(A)]^{-1} = [h(A)]^{-1} g(A).$$

2° *Если составная функция*

$$g(\lambda) \equiv h[f(\lambda)]$$

определена на спектре матрицы A, *то*

$$g(A) = h[f(A)],$$

т. е. $g(A) = h(B)$, *где* $B = f(A)$. При доказательстве этого предложения, как и ранее, будем предполагать, что

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s}$$

---

[1]) В гл. VIII §§ 6 и 7 будет дано более общее определение $\sqrt{A}$ как произвольного решения матричного уравнения $X^2 = A$.

[2]) Этим положением мы уже пользовались на стр. 112. См. сноску [2]) к стр. 112.

[3]) См. (25) на стр. 94.

— минимальный многочлен матрицы A. Тогда значения функции $g(\lambda)$ на спектре матрицы A определяются по формулам[1])

$$\begin{aligned} &g(\lambda_k) = h(\mu_k), \quad g'(\lambda_k) = h'(\mu_k) f'(\lambda_k), \ \ldots, \\ &g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\mu_k) [f'(\lambda_k)]^{(m_k-1)} + \ \ldots \ + h'(\mu_k) f^{(m_k-1)}(\lambda_k), \end{aligned} \tag{36}$$

где $\mu_k = f(\lambda_k)$ $(k = 1, \ldots, s)$. Многочлен

$$\chi(\mu) = (\mu - \mu_1)^{m_1} (\mu - \mu_2)^{m_2} \ldots (\mu - \mu_s)^{m_s}$$

будет аннулирующим многочленом для матрицы B. Действительно, каждое число λ_k является корнем по крайней мере кратности m_k функции

$$q(\lambda) \equiv \chi[f(\lambda)] = \prod_{k=1}^{s} [f(\lambda) - f(\lambda_k)]^{m_k}.$$

Поэтому

$$q(\Lambda_A) = 0$$

и согласно 1°

$$q(A) = \chi[f(A)] = \chi(B) = 0.$$

Поэтому среди значений

$$h(\mu_1), \ h'(\mu_1), \ \ldots \ h^{(m_k-1)}(\mu_k) \qquad (k = 1, \ldots, s), \tag{37}$$

содержатся все значения функции $h(\mu)$ на спектре матрицы B. Исходя из значений (37), построим интерполяционный многочлен $r(\lambda)$ для функции $h(\lambda)$. Тогда, с одной стороны,

$$h(B) = r(B).$$

С другой стороны, как показывают формулы (36), функции $g(\lambda)$ и $g_1(\lambda) = r[f(\lambda)]$ будут равны на спектре матрицы A. Поэтому, применяя к разности $g(\lambda) - r[f(\lambda)]$ предложение 1°, получим:

$$g(A) - r[f(A)] = 0;$$

но тогда

$$g(A) = r[f(A)] = r(B) = h(B) = h[f(A)],$$

что и требовалось доказать.

Комбинируя предложения 1° и 2°, приходим к следующему обобщению предложения 1°.

3°. *Пусть*

$$g(\lambda) \equiv G[f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)],$$

где функции $f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)$ *определены на спектре матрицы* A, *а функция* $G(u_1, u_2, \ldots, u_l)$ *есть результат последовательного применения к величинам* $u_1, u_2, \ldots, u_l$ *операций сложения, умножения, умножения на число и замены величины произвольной функцией от нее. Тогда из*

$$g(\Lambda_A) = 0$$

---

[1]) Предполагается, что все входящие в эти формулы величины

$$f(\lambda_k), \ \ldots, f^{(m_k-1)}(\lambda_k), \ h(\mu_k), \ \ldots, \ h^{(m_k-1)}(\mu_k) \qquad (k = 1, \ldots, s)$$

имеют смысл.

следует

$$G[f_1(A),\ f_2,(A),\ \ldots,\ f_l(A)]=0.$$

Так, например, пусть A — неособенная матрица ($|A|\neq 0$). Обозначим через $\ln\lambda$ однозначную ветвь многозначной функции $\operatorname{Ln}\lambda$, определенную в некоторой области, не содержащей числа 0 и содержащей все характеристические числа матрицы A. Тогда в скалярном тождестве

$$e^{\ln\lambda}-\lambda=0$$

можно заменить скалярный аргумент λ на матрицу A:

$$e^{\ln A}-A=0,$$

т. е. $e^{\ln A}=A$. Другими словами, матрица $X=\ln A$ удовлетворяет матричному уравнению $e^X=A$, т. е. является «натуральным логарифмом» матрицы A.

Беря в качестве $\ln\lambda$ другие однозначные ветви многозначной функции $\operatorname{Ln}\lambda$, мы получим другие логарифмы матрицы A [1]). Пусть $A=\|a_{ik}\|_1^n$ — вещественная неособенная матрица. В главе VIII, § 8 будут установлены необходимые и достаточные условия для того, чтобы вещественная матрица имела вещественный натуральный логарифм. Здесь же мы рассмотрим два частных случая.

1) Матрица A не имеет вещественных отрицательных характеристических чисел. Обозначим через $\ln_0\lambda$ однозначную ветвь функции $\ln\lambda$ в комплексной λ-плоскости с разрезом вдоль отрицательной действительной оси, определяемую равенством

$$\ln_0\lambda=\ln r+i\varphi,\quad -\pi<\varphi<\pi\qquad(\lambda=re^{i\varphi}).$$

Функция $\ln_0\lambda$ принимает вещественные значения при положительных вещественных λ и комплексно сопряженные значения при комплексно сопряженных значениях λ. Поэтому функция $\ln_0\lambda$ вещественна на спектре матрицы A (см. стр. 110) и $\ln_0 A$ — вещественная матрица.

2) $A=B^2$, где B — вещественная матрица [2]). Наряду с функцией $\ln_0\lambda$ введем в рассмотрение две однозначные ветви функции $\ln\lambda$ в комплексной λ-плоскости с разрезом вдоль положительной действительной оси:

$$\ln_1\lambda=\ln r+i\varphi\qquad 0\leqslant\varphi<2\pi\qquad(\lambda=re^{i\varphi}),$$

$$\ln_2\lambda=\ln r+i\varphi\qquad -2\pi<\varphi\leqslant 0\qquad(\lambda=re^{i\varphi}).$$

Пусть матрица B имеет различные характеристические числа λ_k ($k=1,\ \ldots,\ s$). Выберем круговые окрестности G_k точек λ_k ($k=1,\ \ldots,\ s$) так, чтобы они не пересекались и не содержали начало $\lambda=0$. В области, составленной из этих окрестностей, определим функцию $f(\lambda)$ равенством

$$f(\lambda)=\ln_0\lambda^2,\ \text{если}\ \lambda\in G_k\ \text{и}\ \operatorname{Re}\lambda_k\neq 0,$$

$$f(\lambda)=\ln_1\lambda^2,\ \text{если}\ \lambda\in G_k\ \text{и}\ \operatorname{Re}\lambda_k=0,\ \operatorname{Im}\lambda_k>0,$$

$$f(\lambda)=\ln_2\lambda^2,\ \text{если}\ \lambda\in G_k\ \text{и}\ \operatorname{Re}\lambda_k=0,\ \operatorname{Im}\lambda_k<0.$$

---

[1]) Однако на этом пути мы не получим все логарифмы матрицы A. Общая формула, охватывающая все логарифмы матрицы A, будет дана в гл. VIII, § 8.

[2]) В этом случае матрица A имеет отрицательные характеристические числа, если матрица B имеет чисто мнимые характеристические числа.

Тогда функция $f(\lambda)$ представляет собой однозначную ветвь функции $\ln\lambda^2$, определенную и вещественную на спектре матрицы B. Поэтому $f(B)$ — вещественная матрица и

$$e^{f(B)} = B^2 = A,$$

т. е. матрица $f(B)$ является вещественным натуральным логарифмом матрицы A.

Примечание 1. Если A — линейный оператор в n-мерном пространстве R, то $f(A)$ определяется совершенно так же, как и $f(A)$:

$$f(A) = r(A),$$

где $r(\lambda)$ — интерполяционный многочлен Лагранжа — Сильвестра для $f(\lambda)$ на спектре оператора A (спектр оператора A определяется минимальным аннулирующим многочленом $\psi(\lambda)$ оператора A).

Согласно этому определению, если оператору A отвечает матрица $A = \|a_{ik}\|_1^n$ в некотором базисе пространства, то оператору $f(A)$ в том же базисе отвечает матрица $f(A)$. Все утверждения и формулировки этой главы, в которых фигурирует матрица A, остаются в силе и после замены матрицы A оператором A.

Примечание 2. Можно определить [1]) функцию от матрицы $f(A)$, исходя из характеристического многочлена

$$\Delta(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{n_k},$$

заменяя им минимальный многочлен $\psi(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{m_k}$. При этом полагают $f(A) = g(A)$, где $g(\lambda)$ — интерполяционный многочлен степени $< n$ по $\mathrm{mod}\,\Delta(\lambda)$ для функции $f(\lambda)$ [2]). Формулы (16), (22) и (24) заменяются формулами

$$f(A) = \sum_{k=1}^{s} f[(\lambda_k)\widehat{Z}_{k1} + f'(\lambda_k)\widehat{Z}_{k2} + \ldots + f^{(n_k-1)}(\lambda_k)\widehat{Z}_{kn_k}], \tag{16'}$$

$$(\lambda E - A)^{-1} = \frac{B(\lambda)}{\Delta(\lambda)} = \sum_{k=1}^{s}\left[\frac{\widehat{Z}_{k1}}{\lambda-\lambda_k} + \frac{1!\,\widehat{Z}_{k2}}{(\lambda-\lambda_k)^2} + \ldots + \frac{(n_k-1)!\,\widehat{Z}_{kn_k}}{(\lambda-\lambda_k)^{n_k-1}}\right], \tag{22'}$$

$$f(A) = \sum_{k=1}^{s} \frac{1}{(n_k-1)!}\left[\frac{B(\lambda)}{\Delta_k(\lambda)} f(\lambda)\right]^{(n_k-1)}_{\lambda=\lambda_k} \text{ [3])}, \tag{24'}$$

где

$$\Delta_k(\lambda) = \frac{\Delta(\lambda)}{(\lambda-\lambda_k)^{n_k}} \qquad (k = 1, 2, \ldots, s).$$

Однако в формуле (16′) значения $f^{(m_k)}(\lambda_k)$, $f^{(m_k+1)}(\lambda_k)$, …, $f^{(n_k-1)}(\lambda_k)$ входят лишь фиктивно, поскольку из сопоставления (22) с (22′) следует:

$$\widehat{Z}_{k1} = Z_{k1}, \ \ldots, \ \widehat{Z}_{km_k} = Z_{km_k}, \qquad \widehat{Z}_{km_k+1} = \ldots = \widehat{Z}_{kn_k} = 0.$$

---

[1]) См., например, В. Д. Мак-Миллан, Динамика твердого тела, ИЛ, М., 1951, стр. 403 и далее.

[2]) Многочлен $g(\lambda)$ не определяется однозначно равенством $f(A) = g(A)$ и условием «степень $< n$».

[3]) Частный случай формулы (24′), когда $f(\lambda) \equiv \lambda^h$, иногда называют формулой Перрона [см. [29], стр. 25—27].

§ 6. Применение функций от матрицы к интегрированию системы линейных дифференциальных уравнений с постоянными коэффициентами

1. Рассмотрим сначала систему однородных линейных дифференциальных уравнений с постоянными коэффициентами первого порядка:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n; \end{aligned}\right\} \tag{38}$$

здесь t — независимое переменное, $x_1, x_2, \ldots, x_n$ — неизвестные функции переменной t, a_{ik} $(i, k = 1, 2, \ldots, n)$ — комплексные числа.

Введем в рассмотрение квадратную матрицу $A = \|a_{ik}\|_1^n$, составленную из коэффициентов, и столбцевую матрицу $x = (x_1, x_2, \ldots, x_n)$. Тогда система уравнений (38) может быть записана в виде одного матричного дифференциального уравнения

$$\frac{dx}{dt} = Ax. \tag{39}$$

Здесь и в дальнейшем под *производной от матрицы* мы понимаем матрицу, получающуюся из данной путем замены всех элементов их производными. Поэтому $\frac{dx}{dt}$ — столбцевая матрица с элементами $\frac{dx_1}{dt}$, $\frac{dx_2}{dt}, \ldots, \frac{dx_n}{dt}$.

Будем искать решение системы дифференциальных уравнений, удовлетворяющее заданным начальным условиям:

$$x_1|_{t=0} = x_{10}, \quad x_2|_{t=0} = x_{20}, \ \ldots, \ x_n|_{t=0} = x_{n0},$$

или в сокращенной записи:

$$x|_{t=0} = x_0. \tag{40}$$

Разложим искомый столбец x в ряд Маклорена по степеням t:

$$x = x_0 + \dot{x}_0 t + \ddot{x}_0 \frac{t^2}{2!} + \ldots \left(\dot{x}_0 = \frac{dx}{dt}\Big|_{t=0}, \ \ddot{x}_0 = \frac{d^2x}{dt^2}\Big|_{t=0} \text{ и т. д.}\right). \tag{41}$$

Но из (39) почленным дифференцированием находим:

$$\frac{d^2x}{dt^2} = A\frac{dx}{dt} = A^2x, \quad \frac{d^3x}{dt^3} = A\frac{d^2x}{dt^2} = A^3x, \ \ldots \tag{42}$$

Подставляя в (39) и (42) значение $t = 0$, получим:

$$\dot{x}_0 = Ax_0, \quad \ddot{x}_0 = A^2x_0, \ \ldots$$

Теперь ряд (41) запишется так:

$$x = x_0 + tAx_0 + \frac{t^2}{2!}A^2x_0 + \ldots = e^{At}x_0. \tag{43}$$

Непосредственной подстановкой в (39) убеждаемся[1]) в том, что (43) есть решение дифференциального уравнения (39). Полагая в (43) $t=0$, найдем: $x|_{t=0}=x_0$.

Таким образом, формула (43) дает нам решение данной системы дифференциальных уравнений, удовлетворяющее начальным условиям (40).

Положим в (16) $f(\lambda)=e^{\lambda t}$. Тогда

$$e^{At}=\| q_{ik}(t) \|_1^n=\sum_{k=1}^{s}(Z_{k1}+Z_{k2}t+\ldots+Z_{km_k}t^{m_k-1})\,e^{\lambda_k t}. \tag{44}$$

Теперь решение (43) может быть записано в следующей форме:

$$\left.\begin{aligned} x_1&=q_{11}(t)\,x_{10}+q_{12}(t)\,x_{20}+\ldots+q_{1n}(t)\,x_{n0},\\ x_2&=q_{21}(t)\,x_{10}+q_{22}(t)\,x_{20}+\ldots+q_{2n}(t)\,x_{n0},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ x_n&=q_{n1}(t)\,x_{10}+q_{n2}(t)\,x_{20}+\ldots+q_{nn}(t)\,x_{n0}; \end{aligned}\right\} \tag{45}$$

здесь $x_{10}, x_{20}, \ldots, x_{n0}$ — произвольные постоянные, равные начальным значениям неизвестных функций $x_1, x_2, \ldots, x_n$.

Таким образом, интегрирование данной системы дифференциальных уравнений сведено к вычислению элементов матрицы e^{At}.

Если в качестве начального значения аргумента взять значение $t=t_0$, то формула (43) заменится формулой

$$x=e^{A(t-t_0)}x_0. \tag{46}$$

Пример.

$$\frac{dx_1}{dt}=3x_1-x_2+x_3,$$

$$\frac{dx_2}{dt}=2x_1+x_3,$$

$$\frac{dx_3}{dt}=x_1-x_2+2x_3.$$

Матрица коэффициентов:

$$A=\begin{Vmatrix} 3 & -1 & 1\\ 2 & 0 & 1\\ 1 & -1 & 2 \end{Vmatrix}.$$

Составим характеристический определитель

$$\Delta(\lambda)=-\begin{vmatrix} 3-\lambda & -1 & 1\\ 2 & -\lambda & 1\\ 1 & -1 & 2-\lambda \end{vmatrix}=(\lambda-1)(\lambda-2)^2.$$

Наибольший общий делитель миноров второго порядка этого определителя $D_2(\lambda)=1$. Поэтому

$$\psi(\lambda)=\Delta(\lambda)=(\lambda-1)(\lambda-2)^2.$$

Основная формула:

$$f(A)=f(1)\,Z_1+f(2)\,Z_2+f'(2)\,Z_3.$$

---

[1]) $\frac{d}{dt}(e^{At})=\frac{d}{dt}\left(E+At+\frac{A^2t^2}{2!}+\ldots\right)=A+A^2t+\frac{A^3t^2}{2!}+\ldots=Ae^{At}$.

Возьмем вместо $f(\lambda)$ последовательно 1, $\lambda-2$, $(\lambda-2)^2$. Получим:

$$Z_1+Z_2=E=\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix},$$

$$-Z_1+Z_3=A-2E=\left\|\begin{matrix}1&-1&1\\2&-2&1\\1&-1&0\end{matrix}\right\|\begin{matrix}1\\1.\\0\end{matrix}$$

$$Z_1=(A-2E)^2=\left\|\begin{matrix}0&0&0\\-1&1&0\\-1&1&0\end{matrix}\right\|\begin{matrix}0\\0.\\0\end{matrix}$$

Определенные отсюда Z_1, Z_2 и Z_3 подставляем в основную формулу:

$$f(A)=f(1)\left\|\begin{matrix}0&0&0\\-1&1&0\\-1&1&0\end{matrix}\right\|+f(2)\left\|\begin{matrix}1&0&0\\1&0&0\\1&-1&1\end{matrix}\right\|+f'(2)\left\|\begin{matrix}1&-1&1\\1&-1&1\\0&0&0\end{matrix}\right\|.$$

Заменяя здесь $f(\lambda)$ на $e^{\lambda t}$, будем иметь:

$$e^{At}=e^t\left\|\begin{matrix}0&0&0\\-1&1&0\\-1&1&0\end{matrix}\right\|+e^{2t}\left\|\begin{matrix}1&0&0\\1&0&0\\1&-1&1\end{matrix}\right\|+te^{2t}\left\|\begin{matrix}1&-1&1\\1&-1&1\\0&0&0\end{matrix}\right\|=\left\|\begin{matrix}(1+t)e^{2t}&-te^{2t}&te^{2t}\\-e^t+(1+t)e^{2t}&e^t-te^{2t}&te^{2t}\\-e^t+e^{2t}&e^t-e^{2t}&e^{2t}\end{matrix}\right\|.$$

Таким образом,

$$\begin{aligned}x_1&=C_1(1+t)e^{2t}-C_2te^{2t}+C_3te^{2t},\\x_2&=C_1[-e^t+(1+t)e^{2t}]+C_2(e^t-te^{2t})+C_3te^{2t},\\x_3&=C_1(-e^t+e^{2t})+C_2(e^t-e^{2t})+C_3e^{2t},\end{aligned}$$

где

$$C_1=x_{10},\quad C_2=x_{20},\quad C_3=x_{30}.$$

2. Рассмотрим теперь систему неоднородных линейных дифференциальных уравнений с постоянными коэффициентами:

$$\left.\begin{aligned}\frac{dx_1}{dt}&=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n+f_1(t),\\\frac{dx_2}{dt}&=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n+f_2(t),\\&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\\frac{dx_n}{dt}&=a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n+f_n(t),\end{aligned}\right\}\qquad(47)$$

где $f_i(t)$ $(i=1, 2, \ldots, n)$ — непрерывные функции в интервале $t_0\leqslant t\leqslant t_1$. Обозначая через $f(t)$ столбцевую матрицу с элементами $f_1(t)$, $f_2(t)$, $\ldots$, $f_n(t)$ и снова полагая $A=\|a_{ik}\|_1^n$, мы систему (47) запишем так:

$$\frac{dx}{dt}=Ax+f(t).\qquad(48)$$

Введем вместо x новый столбец неизвестных функций z, связанный с x соотношением

$$x=e^{At}z.\qquad(49)$$

Дифференцируя почленно (49) и подставляя полученное выражение для $\frac{dx}{dt}$ в (48), найдем[1]):

$$e^{At}\frac{dz}{dt}=f(t).\qquad(50)$$

[1]) См. сноску[1]) к стр. 125.

Отсюда

$$z(t) = c + \int_{t_0}^{t} e^{-A\tau} f(\tau)\, d\tau \text{ [1]} \tag{51}$$

и потому согласно (49)

$$x = e^{At}\left[c + \int_{t_0}^{t} e^{-A\tau} f(\tau)\, d\tau\right] = e^{At}c + \int_{t_0}^{t} e^{A(t-\tau)} f(\tau)\, d\tau; \tag{52}$$

здесь c — столбец с произвольными постоянными элементами.

Давая в (52) аргументу t значение t_0, найдем:

$$c = e^{-At_0}x_0,$$

и, следовательно, решение (52) может быть записано так:

$$x = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)} f(\tau)\, d\tau. \tag{53}$$

Полагая $e^{At} = \| q_{ij}(t) \|_1^n$, мы решение (53) можем записать в развернутом виде так:

$$\left.\begin{aligned} x_1 &= q_{11}(t-t_0)\,x_{10} + \ldots + q_{1n}(t-t_0)\,x_{n0} + \\ &\quad + \int_{t_0}^{t} [q_{11}(t-\tau) f_1(\tau) + \ldots + q_{1n}(t-\tau) f_n(\tau)]\, d\tau, \\ &\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= q_{n1}(t-t_0)\,x_{10} + \ldots + q_{nn}(t-t_0)\,x_{n0} + \\ &\quad + \int_{t_0}^{t} [q_{n1}(t-\tau) f_1(\tau) + \ldots + q_{nn}(t-\tau) f_n(\tau)]\, d\tau. \end{aligned}\right\} \tag{54}$$

3. Рассмотрим в качестве примера движение весомой материальной точки в пустоте вблизи поверхности Земли с учетом движения Земли. В этом случае, как известно [2]), ускорение точки относительно Земли определяется постоянной силой веса $m\boldsymbol{g}$ и инерционной кориолисовой силой $-2m\boldsymbol{\omega}\times\boldsymbol{v}$ ($\boldsymbol{v}$ — скорость точки относительно земли, $\boldsymbol{\omega}$ — постоянная угловая скорость земли) [3]). Поэтому дифференциальное уравнение движения точки имеет вид

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{g} - 2\boldsymbol{\omega}\times\boldsymbol{v}. \tag{55}$$

---

[1]) Как уже указывалось для частного случая (см. сноску [1]) к стр. 118), если дана матричная функция от скалярного аргумента $B(\tau) = \| b_{ik}(\tau) \|$ ($i = 1, 2, \ldots, m$; $k = 1, 2, \ldots, n$; $t_1 \leqslant \tau \leqslant t_2$), то *интеграл* $\int_{t_1}^{t_2} B(\tau)\, d\tau$ определяется естественным образом:

$$\int_{t_1}^{t_2} B(\tau)\, d\tau = \left\| \int_{t_1}^{t_2} b_{ik}(\tau)\, d\tau \right\| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n).$$

[2]) См., например, Г. К. Суслов, Теоретическая механика, Гостехиздат, 1944, стр. 141.

[3]) Здесь жирные буквы означают векторы, а символ «$\times$» означает векторное умножение.

Определим линейный оператор A в трехмерном евклидовом пространстве равенством

$$Ax=-2\omega\times x \tag{56}$$

и напишем вместо (55):

$$\frac{dv}{dt}=Av+g. \tag{57}$$

Сопоставляя (57) с (48), легко найдем по формуле (53):

$$v=e^{At}v_0+\int_0^t e^{At}\,dt\cdot g \qquad (v_0=v_{t=0}).$$

Интегрируя почленно, определим радиус-вектор движущейся точки:

$$r=r_0+\int_0^t e^{At}\,dt\,v_0+\int_0^t dt\int_0^t e^{At}\,dt\,g, \tag{58}$$

где

$$r_0=r_{t=0} \text{ и } v_0=v_{t=0}.$$

Подставляя вместо e^{At} сумму ряда

$$E+A\frac{t}{1!}+A^2\frac{t^2}{2!}+\dots$$

и заменяя оператор A его выражением из (56), будем иметь:

$$r=r_0+v_0t+\frac{1}{2}gt^2-\omega\times\left(v_0t^2+\frac{1}{3}gt^3\right)+\omega\times\left[\omega\times\left(\frac{2}{3}v_0t^3+\frac{1}{6}gt^4\right)\right]+\dots$$

Считая угловую скорость ω малой величиной (для Земли $\omega\approx 7{,}3\cdot 10^{-5}$ 1/сек.) и отбрасывая члены, содержащие вторую и высшие степени ω, мы для дополнительного отклонения точки, вызванного вращением Земли, получим приближенную формулу

$$d\approx-\omega\times\left(v_0t^2+\frac{1}{3}gt^3\right).$$

Возвращаясь к точному решению (58), вычислим e^{At}. Предварительно установим, что минимальный многочлен оператора A имеет вид

$$\psi(\lambda)=\lambda(\lambda^2+4\omega^2).$$

Действительно, из (56) находим:

$$A^2x=4\omega\times(\omega\times x)=4(\omega x)\,\omega-4\omega^2x,$$

$$A^3x=-2\omega\times A^2x=8\omega^2(\omega\times x).$$

Отсюда и из (56) следует, что операторы E, A, A^2 линейно независимы, а

$$A^3+4\omega^2A=0.$$

Минимальный многочлен $\psi(\lambda)$ имеет простые корни 0, $2\omega i$, $-2\omega i$. Интерполяционный многочлен Лагранжа для e^{At} имеет вид

$$1+\frac{\sin 2\omega t}{2\omega}\lambda+\frac{1-\cos 2\omega t}{4\omega^2}\lambda^2.$$

Тогда

$$e^{At}=E+\frac{\sin 2\omega t}{2\omega}A+\frac{1-\cos 2\omega t}{4\omega^2}A^2.$$

Подставляя это выражение для e^{At} в (58) и заменяя оператор A его выражением из (56), найдем:

$$\boldsymbol{r}=\boldsymbol{r}_0+\boldsymbol{v}_0t+\boldsymbol{g}\frac{t^2}{2}-\boldsymbol{\omega}\times\left(\frac{1-\cos 2\omega t}{2\omega^2}\boldsymbol{v}_0+\frac{2\omega t-\sin 2\omega t}{4\omega^3}\boldsymbol{g}\right)+$$

$$+\boldsymbol{\omega}\times\left[\boldsymbol{\omega}\times\left(\frac{2\omega t-\sin 2\omega t}{2\omega^3}\boldsymbol{v}_0+\frac{-1+2\omega^2t^2+\cos 2\omega t}{4\omega^4}\boldsymbol{g}\right)\right]. \qquad (59)$$

Рассмотрим частный случай $\boldsymbol{v}_0=0$. Тогда, раскрывая тройное векторное произведение[1],) получим:

$$\boldsymbol{r}=\boldsymbol{r}_0+\boldsymbol{g}\frac{t^2}{2}+\frac{2\omega t-\sin 2\omega t}{4\omega^3}(\boldsymbol{g}\times\boldsymbol{\omega})-\frac{\cos 2\omega t-1+2\omega^2t^2}{4\omega^3}(g\sin\varphi\boldsymbol{\omega}+\omega\boldsymbol{g}),$$

где φ — географическая широта в данном месте земли. Член

$$\frac{2\omega t-\sin 2\omega t}{4\omega^3}(\boldsymbol{g}\times\boldsymbol{\omega})$$

представляет собой отклонение, направленное перпендикулярно к плоскости меридиана на восток, а последнее слагаемое в правой части последней формулы дает отклонение, лежащее в плоскости меридиана и направленное от земной оси (перпендикулярно к ней).

4. Пусть теперь дана следующая система линейных дифференциальных уравнений второго порядка:

$$\left.\begin{aligned}\frac{d^2x_1}{dt^2}+a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ \frac{d^2x_2}{dt^2}+a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \frac{d^2x_n}{dt^2}+a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0,\end{aligned}\right\} \qquad (60)$$

где $a_{ik}\,(i,\,k=1,\,2,\,\ldots,\,n)$ — постоянные коэффициенты. Вводя снова столбец $x=(x_1,\,x_2,\,\ldots,\,x_n)$ и квадратную матрицу $A=\|a_{ik}\|_1^n$, мы перепишем систему (60) в матричном виде

$$\frac{d^2x}{dt^2}+Ax=0.$$

Рассмотрим сначала случай, когда $|A|\neq 0$. Если $n=1$, т. е. x и A — скаляры и $A\neq 0$, общее решение уравнения (60) может быть записано в виде

$$x=\cos\left(\sqrt{A}\,t\right)x_0+\left(\sqrt{A}\right)^{-1}\sin\left(\sqrt{A}\,t\right)\dot{x}_0, \qquad (61)$$

где $x_0=x_{t=0}$ и $\dot{x}_0=\left(\frac{dx}{dt}\right)_{t=0}$.

Непосредственной проверкой убеждаемся, что (61) представляет собой решение уравнения (60) при любом n, когда x — столбец, а A — неособенная квадратная матрица[2]). При этом мы опираемся на формулы

$$\left.\begin{aligned}\cos\left(\sqrt{A}t\right)=E-\frac{1}{2!}At^2+\frac{1}{4!}A^2t^4-\ldots,\\ \left(\sqrt{A}\right)^{-1}\sin\left(\sqrt{A}t\right)=t-\frac{1}{3!}At^3+\frac{1}{5!}A^2t^5-\ldots\end{aligned}\right\} \qquad (62)$$

---

[1]) По формуле $\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})=\boldsymbol{b}(\boldsymbol{ac})-\boldsymbol{c}(\boldsymbol{ab})$; при этом следует учесть, что вектор $\boldsymbol{g}$ направлен к центру Земли, откуда $\boldsymbol{\omega g}=-\omega g\sin\varphi$. — *Прим. ред.*

[2]) Под $\sqrt{A}$ мы здесь понимаем произвольную матрицу, квадрат которой равен A. $\sqrt{A}$ заведомо существует при $|A|\neq 0$ (см. стр. 120).

Формула (61) охватывает все решения системы (60) или (60′), поскольку начальные значения x_0 и $\dot{x}_0$ могут быть выбраны произвольно.

В формулах (62) правые части имеют смысл и при $|A|=0$. Поэтому (61) представляет собой общее решение данной системы дифференциальных уравнений и в случае, когда $|A|=0$, если только под функциями $\cos(\sqrt{A}\,t)$ и $(\sqrt{A})^{-1}\sin(\sqrt{A}\,t)$, входящими в состав этого выражения, понимать правые части формул (62).

Предоставляем читателю проверить, что общее решение неоднородной системы

$$\frac{d^2x}{dt^2}+Ax=f(t), \tag{63}$$

удовлетворяющее начальным условиям $x_{t=0}=x_0$ и $\left(\frac{dx}{dt}\right)_{t=0}=\dot{x}_0$, может быть записано в виде

$$x=\cos(\sqrt{A}\,t)\,x_0+(\sqrt{A})^{-1}\sin(\sqrt{A}\,t)\,\dot{x}_0+ \\ +(\sqrt{A})^{-1}\int_0^t \sin\left[\sqrt{A}\,(t-\tau)\right]f(\tau)\,d\tau. \tag{64}$$

Если в качестве начального момента времени берется $t=t_0$, то в формулах (61) и (64) следует заменить $\cos(\sqrt{A}\,t)$ и $\sin(\sqrt{A}\,t)$ на $\cos\sqrt{A}\,(t-t_0)$ и $\sin\sqrt{A}\,(t-t_0)$, а $\int_0^t$ на $\int_{t_0}^t$.

§ 7. Устойчивость движения в случае линейной системы

Пусть $x_1, x_2, \ldots, x_n$ — параметры, характеризующие отклонение «возмущенного» движения данной механической системы от исследуемого движения [1]), и пусть эти параметры удовлетворяют системе дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt}=f_i(x_1, x_2, \ldots, x_n, t) \quad (i=1, 2, \ldots, n); \tag{65}$$

здесь независимое переменное t — время, правые части $f_i(x_1, x_2, \ldots, x_n, t)$ — непрерывные функции в некоторой области значений $x_1, x_2, \ldots, x_n$ (содержащей точку $x_1=0$, $x_2=0, \ldots, x_n=0$) при всех $t>t_0$ (t_0 — начальный момент времени).

Введем определение устойчивости движения по Ляпунову [2]).

Исследуемое движение называется *устойчивым*, если для любого числа $\varepsilon>0$ можно указать число $\delta>0$ такое, что при любых начальных (при $t=t_0$) значениях параметров $x_{10}, x_{20}, \ldots, x_{n0}$, меньших по модулю числа δ, параметры $x_1, x_2, \ldots, x_n$ во все время движения ($t\geqslant t_0$) по модулю меньше ε, т. е. для любого $\varepsilon>0$ можно указать такое $\delta>0$, что из

$$|x_{i0}|<\delta \qquad (i=1, 2, \ldots, n) \tag{66}$$

---

[1]) В этих параметрах исследуемое движение характеризуется постоянными нулевыми значениями $x_1=0, x_2=0, \ldots, x_n=0$. Поэтому при математической трактовке вопроса говорят об устойчивости нулевого решения системы дифференциальных уравнений (65).

[2]) См. [22] стр. 13; [38], стр. 10—11 или [23], стр. 11—12.

следует:

$$|x_i(t)|<\varepsilon \qquad (t\geqslant t_0). \tag{67}$$

Если при этом дополнительно при некотором $\delta_0>0$ всегда $\lim\limits_{t\to+\infty} x_i(t)=0$ $(i=1, 2, \ldots, n)$, коль скоро $|x_{i0}|<\delta_0$ $(i=1, 2, \ldots, n)$, то исследуемое движение называется *асимптотически устойчивым.*

Рассмотрим теперь линейную систему, т. е. тот частный случай, когда система (65) является системой линейных однородных дифференциальных уравнений

$$\frac{dx_i}{dt}=\sum_{k=1}^{n} p_{ik}(t)\,x_k, \tag{68}$$

где $p_{ik}(t)$ — непрерывные функции при $t\geqslant t_0$ $(i, k=1, 2, \ldots, n)$.

В матричной записи система (68) запишется так:

$$\frac{dx}{dt}=P(t)\,x; \tag{68'}$$

здесь x — столбцевая матрица с элементами $x_1, x_2, \ldots, x_n$, а $P(t)=\|p_{ik}(t)\|_1^n$ — матрица коэффициентов.

Обозначим через

$$q_{1j}(t),\ q_{2j}(t),\ \ldots,\ q_{nj}(t) \qquad (j=1, 2, \ldots, n) \tag{69}$$

n линейно независимых решений системы (68)[1]. Матрицу $Q(t)=\|q_{ij}\|_1^n$, столбцами которой являются эти решения, называют *интегральной матрицей* системы (68).

Произвольное решение системы линейных однородных дифференциальных уравнений получается как линейная комбинация с постоянными коэффициентами из n линейно независимых решений:

$$x_i=\sum_{j=1}^{n} c_j q_{ij}(t) \qquad (t=1, 2, \ldots, n),$$

или в матричной записи:

$$x=Q(t)\,c, \tag{70}$$

где c — столбцевая матрица, элементами которой являются произвольные постоянные $c_1, c_2, \ldots, c_n$.

Выберем теперь специальную интегральную матрицу, для которой

$$Q(t_0)=E; \tag{71}$$

другими словами, при выборе n линейно независимых решений (69) будем исходить из следующих специальных начальных условий[2]):

$$q_{ij}(t_0)=\delta_{ij}=\begin{cases}0 & (i\neq j),\\ 1 & (i=j)\end{cases} \qquad (i, j=1, 2, \ldots, n).$$

Тогда, полагая в формуле (70) $t=t_0$, из (71) найдем:

$$x_0=c,$$

[1]) Здесь второй индекс j обозначает номер решения.

[2]) Любые начальные условия определяют, и притом однозначно, некоторое решение данной системы.

и потому формула (70) примет вид

$$x = Q(t)\, x_0, \tag{72}$$

или в развернутом виде

$$x_i = \sum_{j=1}^{n} q_{ij}(t)\, x_{j0} \qquad (i = 1,\ 2,\ \ldots,\ n). \tag{72'}$$

Рассмотрим три случая:

1. $Q(t)$ — *ограниченная матрица* в интервале $(t_0,\ +\infty)$, т. е. существует такое число M, что

$$|q_{ij}(t)| \leqslant M \qquad (t \geqslant t_0;\ i, j = 1,\ 2,\ \ldots,\ n).$$

В этом случае из (72') следует:

$$|x_i(t)| \leqslant nM \max |x_{j0}|.$$

Условие устойчивости выполняется. (Достаточно в (66), (67) взять $\delta < \frac{\varepsilon}{nM}$.) *Движение, характеризуемое нулевым решением* $x_1 = 0$, $x_2 = 0, \ldots, x_n = 0$, *устойчиво.*

2. $\lim\limits_{t \to +\infty} Q(t) = 0$. В этом случае матрица $Q(t)$ ограничена в интервале $(t_0,\ +\infty)$, и потому, как уже было выяснено, движение устойчиво. Кроме того, из (72) следует, что

$$\lim_{t \to +\infty} x(t) = 0$$

при любом x_0. *Движение асимптотически устойчиво.*

3. $Q(t)$ — *неограниченная матрица* в интервале $(t_0,\ +\infty)$. Это означает, что по крайней мере одна из функций $q_{ij}(t)$, например $q_{hk}(t)$, не ограничена в интервале $(t_0,\ \infty)$. Возьмем начальные условия $x_{10} = 0, \ldots, x_{k0} \neq 0, \ldots, x_{n0} = 0$. Тогда

$$x_h(t) = q_{hk}(t)\, x_{k0}.$$

Каким бы малым по модулю ни было x_{k0}, функция $x_h(t)$ будет не ограничена. Условие (67) не будет выполняться ни при одном δ. *Движение неустойчиво.*

Рассмотрим теперь частный случай, когда коэффициенты в системе (68) — постоянные числа:

$$P(t) = P = \text{const}. \tag{73}$$

В этом случае (см. § 5)

$$x = e^{P(t-t_0)}\, x_0. \tag{74}$$

Сопоставляя (74) с (72), находим, что в данном случае

$$Q(t) = e^{P(t-t_0)}. \tag{75}$$

Обозначим через

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s}$$

минимальный многочлен матрицы коэффициентов P.

Для исследования интегральной матрицы (75) воспользуемся формулой (16) на стр. 111. В данном случае $f(\lambda) = e^{\lambda(t-t_0)}$ (t рассматривается

как параметр), $f^{(j)}(\lambda_k) = (t-t_0)^j e^{\lambda_k (t-t_0)}$. Формула (16) дает:

$$e^{P(t-t_0)} = \sum_{k=1}^{s} [Z_{k1} + Z_{k2}(t-t_0) + \ldots + Z_{km_k}(t-t_0)^{m_k-1}] e^{\lambda_k (t-t_0)}. \quad (76)$$

Рассмотрим три случая:

1. $\operatorname{Re}\lambda_k \leqslant 0$ $(k=1, 2, \ldots, s)$, причем для тех λ_k, для которых $\operatorname{Re}\lambda_k = 0$, соответствующее $m_k = 1$ (т. е. чисто мнимые характеристические числа являются простыми корнями минимального многочлена).

2. $\operatorname{Re}\lambda_k < 0$ $(k=1, 2, \ldots, s)$.

3. При некотором k имеем $\operatorname{Re}\lambda_k > 0$ либо $\operatorname{Re}\lambda_k = 0$, но $m_k > 1$.

Из формулы (76) следует, что в случае 1 матрица $Q(t) = e^{P(t-t_0)}$ ограничена в интервале $(t_0, +\infty)$, в случае 2 $e^{P(t-t_0)} \to 0$ при $t \to +\infty$ и в случае 3 матрица $e^{P(t-t_0)}$ не ограничена в интервале $(t_0, +\infty)$.

Здесь особого рассмотрения требует лишь тот случай, когда в выражении (76) для $e^{P(t-t_0)}$ имеется несколько слагаемых максимального роста (при $t \to +\infty$), т. е. с максимальным $\operatorname{Re}\lambda_k = \alpha_0 \geqslant 0$ и (при данном $\operatorname{Re}\lambda_k = \alpha_0$) с максимальным значением $m_k = m_0$. Тогда выражение (76) можно представить в виде

$$e^{P(t-t_0)} = e^{\alpha_0 (t-t_0)} (t-t_0)^{m_0-1} \Big[\sum_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j (t-t_0)} + (*) \Big], \quad (77)$$

где $\beta_1, \beta_2, \ldots, \beta_r$ — различные вещественные числа, а $(*)$ обозначает матрицу, стремящуюся к нулю при $t \to +\infty$. Из этого представления вытекает, что матрица $e^{P(t-t_0)}$ не ограничена при $\alpha_0 + m_0 - 1 > 0$[1]), поскольку матрица $\sum_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j (t-t_0)}$ не может стремиться к нулю при $t \to +\infty$. В последнем мы убедимся, если докажем, что функция

$$f(t) = \sum_{j=1}^{r} c_j e^{i\beta_j t},$$

где c_j — комплексные числа, а β_j — вещественные и различные между собой числа, может стремиться к нулю при $t \to +\infty$ только в случае $f(t) \equiv 0$. Но, действительно,

$$\overline{f(t)} = \sum_{j=1}^{r} \bar{c}_j e^{-i\beta_j t}. \quad (78')$$

Перемножая почленно равенства (78) и (78′) и интегрируя по t в пределах от 0 до T, получим:

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T |f(t)|^2 \, dt = \sum_{j=1}^{r} |c_j|^2. \quad (79)$$

Но из $\lim\limits_{t\to+\infty} f(t) = 0$ вытекает, что и

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T |f(t)|^2 \, dt = 0.$$

Поэтому из равенства (79) находим, что $c_1 = c_2 = \ldots = c_r = 0$, т. е. $f(t) \equiv 0$.

Поэтому в случае 1 движение $(x_1 = 0, x_2 = 0, \ldots, x_n = 0)$ устойчиво, в случае 2 — асимптотически устойчиво и в случае 3 — неустойчиво.

---

[1]) Иначе говоря, при $\alpha_0 > 0$ либо при $\alpha_0 = 0$, но $m_0 > 1$. — *Прим. ред.*

Результаты исследования могут быть сформулированы в виде следующей теоремы[1]):

Теорема 3. *Нулевое решение линейной системы* (68) *при* $P=\text{const}$ *является устойчивым по Ляпунову, если* 1) *все характеристические числа матрицы* P *имеют отрицательные или нулевые вещественные части,* 2) *все характеристические числа с нулевыми вещественными частями, т. е. чисто мнимые характеристические числа (если таковые имеются), являются простыми корнями минимального многочлена матрицы* P, *и неустойчивым, если хотя бы одно из условий* 1), 2) *не выполняется.*

Нулевое решение линейной системы (68) *является асимптотически устойчивым в том и только в том случае, когда все характеристические числа матрицы* P *имеют отрицательные вещественные части.*

Приведенные выше соображения позволяют высказать суждение о характере интегральной матрицы $e^{P(t-t_0)}$ в общем случае при произвольных характеристических числах постоянной матрицы P.

Теорема 4. *Интегральная матрица* $e^{P(t-t_0)}$ *линейной системы* (58) *при* $P=\text{const}$ *всегда представима в виде*

$$e^{P(t-t_0)}=Z_-(t)+Z_0(t)+Z_+(t), \tag{80}$$

где 1) $\lim\limits_{t\to+\infty} Z_-(t)=0$, 2) $Z_0(t)$ *либо* $=\text{const}$, *либо является ограниченной матрицей в интервале* $(t_0, +\infty)$, *не имеющей предела при* $t\to+\infty$, 3) $Z_+(t)$ *либо* $\equiv 0$, *либо является неограниченной матрицей в интервале* $(t_0, +\infty)$.

Доказательство. Разобьем в правой части равенства (76) все m слагаемых на три группы. Обозначим через $Z_-(t)$ сумму всех слагаемых, содержащих множитель $e^{\lambda_k(t-t_k)}$ с $\operatorname{Re}\lambda_k<0$. Через $Z_t(t)$ обозначим сумму слагаемых, у которых либо $\operatorname{Re}\lambda_k>0$, либо $\operatorname{Re}\lambda_k=0$ при наличии множителя $(t-t_0)^\nu$ при $\nu>0$. Через $Z_0(t)$ обозначим сумму всех остальных слагаемых. Приведенные ранее соображения показывают, что $\lim\limits_{t\to+\infty} Z_-(t)=0$, а функция $Z_+(t)$ не ограничена, если только она не равна тождественно нулю. Функция же $Z_0(t)$ ограничена. Покажем, что из существования предела $\lim\limits_{t\to+\infty} Z_0(t)=B$ следует, что $Z_0(t)=\text{const}$. Действительно, разность $Z_0(t)-B$ может быть представлена в виде суммы $\sum\limits_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j(t-t_0)}$ из равенства (77). Относительно же суммы такого вида выше было показано, что она может иметь предел 0 при $t\to+\infty$ лишь тогда, когда она тождественно равна нулю.

Теорема 4 доказана.

---

[1]) О том, как уточняются критерии устойчивости и неустойчивости для линеаризованных систем (т. е. нелинейных систем, становящихся линейными после пренебрежения нелинейными членами), см. далее в гл. XIV, § 3.

ГЛАВА VI

ЭКВИВАЛЕНТНЫЕ ПРЕОБРАЗОВАНИЯ МНОГОЧЛЕННЫХ МАТРИЦ. АНАЛИТИЧЕСКАЯ ТЕОРИЯ ЭЛЕМЕНТАРНЫХ ДЕЛИТЕЛЕЙ

Первые три параграфа настоящей главы посвящены учению об эквивалентности многочленных матриц. На основе этого в последующих трех параграфах строится аналитическая теория элементарных делителей, т. е. теория приведения постоянной (немногочленной) квадратной матрицы A к нормальной форме $\widetilde{A}$ ($A = T\widetilde{A}T^{-1}$). В последних двух параграфах главы даны два метода построения преобразующей матрицы T.

§ 1. Элементарные преобразования многочленной матрицы

Определение 1. *Многочленной матрицей* или *λ-матрицей* называется прямоугольная матрица $A(\lambda)$, элементы которой суть многочлены от λ:

$$A(\lambda) = \| a_{ik}(\lambda) \| = \| a_{ik}^{(0)}\lambda^l + a_{ik}^{(1)}\lambda^{l-1} + \ldots + a_{ik}^{(l)} \| \quad \begin{pmatrix} i = 1, 2, \ldots, m; \\ k = 1, 2, \ldots, n \end{pmatrix};$$

здесь l — наибольшая из степеней многочленов $a_{ik}(\lambda)$.

Полагая

$$A_j = \| a_{ik}^{(j)} \| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n;\ j = 0, 1, \ldots, l),$$

мы можем представить многочленную матрицу $A(\lambda)$ в виде матричного многочлена относительно λ, т. е. в виде многочлена с матричными коэффициентами:

$$A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \ldots + A_{l-1}\lambda + A_l.$$

Введем в рассмотрение следующие *элементарные операции* над многочленной матрицей $A(\lambda)$:

1° Умножение какой-либо, например i-й, строки на число $c \neq 0$.

2° Прибавление к какой-либо, например i-й, строке другой, например j-й, строки, предварительно умноженной на произвольный многочлен $b(\lambda)$.

3° Перестановка местами любых двух строк, например i-й и j-й строк.

Предлагаем читателю проверить, что операции 1°, 2°, 3° равносильны умножению многочленной матрицы $A(\lambda)$ *слева* соответственно на

следующие квадратные матрицы порядка m[1]):

$$S' = \begin{Vmatrix} & & (i) & & \\ 1 & & \vdots & & 0 \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ 0 & & & & 1 \end{Vmatrix}, \quad S'' = \begin{Vmatrix} & & (j) & & \\ 1 & & \vdots & & 0 \\ & \ddots & & & \\ & & 1 \ldots b(\lambda) & & \\ & & & \ddots & \\ 0 & & & & 1 \end{Vmatrix} \begin{matrix} \\ \\ \\

$$T''' = \begin{Vmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & \dots & 1 & \dots & \\ & & & \vdots & & \vdots & & \\ & & & 1 & \dots & 0 & \dots & \\ & & & & & & \ddots & \\ 0 & & & & & & & 1 \end{Vmatrix} \begin{matrix} \\ \\ \\ \dots\ (i) \\ \\ \dots\ (j) \\ \\ \\ \end{matrix}$$

В результате применения правой элементарной операции матрица $A(\lambda)$ умножается *справа* на соответствующую матрицу T.

Матрицы типа S', S'', S''' (или, что то же, типа T', T'', T''') мы будем называть *элементарными матрицами*.

Определитель любой элементарной матрицы не зависит от λ и отличен от нуля. Поэтому для каждой левой (правой) элементарной операции существует обратная операция, которая также является левой (соответственно правой) элементарной операцией [1]).

Определение 2. Две многочленные матрицы $A(\lambda)$ и $B(\lambda)$ называются 1) *левоэквивалентными*, 2) *правоэквивалентными*, 3) *эквивалентными*, если одна из них получается из другой путем применения соответственно 1) левых элементарных операций, 2) правых элементарных операций, 3) левых и правых элементарных операций [2]).

Пусть матрица $B(\lambda)$ получается из $A(\lambda)$ при помощи левых элементарных операций, соответствующих матрицам $S_1, S_2, \dots, S_p$. Тогда

$$B(\lambda) = S_p S_{p-1} \dots S_1 A(\lambda). \tag{2}$$

Обозначая через $P(\lambda)$ произведение $S_p S_{p-1} \dots S_1$, мы равенство (2) запишем в виде

$$B(\lambda) = P(\lambda) A(\lambda), \tag{3}$$

где $P(\lambda)$, как и каждая из матриц $S_1, S_2, \dots, S_p$, имеет отличный от нуля постоянный [3]) определитель.

В следующем параграфе будет доказано, что каждая квадратная λ-матрица $P(\lambda)$ с постоянным отличным от нуля определителем может быть представлена в виде произведения элементарных матриц. Поэтому равенство (3) эквивалентно равенству (2) и потому означает левую эквивалентность матриц $A(\lambda)$ и $B(\lambda)$.

В случае правой эквивалентности многочленных матриц $A(\lambda)$ и $B(\lambda)$ вместо равенства (3) будем иметь равенство

$$B(\lambda) = A(\lambda) Q(\lambda), \tag{3'}$$

---

[1]) Отсюда следует, что если матрица $B(\lambda)$ получается из $A(\lambda)$ путем применения левых (правых; левых и правых) элементарных операций, то и обратно матрица $A(\lambda)$ может быть получена из $B(\lambda)$ путем применения элементарных операций такого же типа. Как левые элементарные операции, так и правые образуют группу.

[2]) Из определения следует, что левоэквивалентными, правоэквивалентными или просто эквивалентными могут быть только прямоугольные матрицы одинаковых размеров.

[3]) То есть не зависящий от λ.

а в случае (двусторонней) эквивалентности — равенство

$$B(\lambda) = P(\lambda)\, A(\lambda)\, Q(\lambda). \qquad (3'')$$

здесь опять $P(\lambda)$ и $Q(\lambda)$ — матрицы с отличными от нуля и не зависящими от λ определителями.

Таким образом, определение 2 можно заменить равносильным определением.

Определение 2′. Две прямоугольные λ-матрицы $A(\lambda)$ и $B(\lambda)$ называются 1) *левоэквивалентными*, 2) *правоэквивалентными*, 3) *эквивалентными*, если соответственно

$$1)\ B(\lambda) = P(\lambda)\, A(\lambda), \quad 2)\ B(\lambda) = A(\lambda)\, Q(\lambda), \quad 3)\ B(\lambda) = P(\lambda)\, A(\lambda)\, Q(\lambda),$$

где $P(\lambda)$ и $Q(\lambda)$ — многочленные квадратные матрицы с постоянными и отличными от нуля определителями.

Все введенные выше понятия проиллюстрируем на следующем важном примере.

Рассмотрим систему m линейных однородных дифференциальных уравнений l-го порядка с n неизвестными функциями $x_1, x_2, \ldots, x_n$ аргумента t с постоянными коэффициентами:

$$\left.\begin{aligned} a_{11}(D)\, x_1 + a_{12}(D)\, x_2 + \ldots + a_{1n}(D)\, x_n &= 0, \\ a_{21}(D)\, x_1 + a_{22}(D)\, x_2 + \ldots + a_{2n}(D)\, x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m1}(D)\, x_1 + a_{m2}(D)\, x_2 + \ldots + a_{mn}(D)\, x_n &= 0; \end{aligned}\right\} \qquad (4)$$

здесь

$$a_{ik}(D) = a_{ik}^{(0)} D^l + a_{ik}^{(1)} D^{l-1} + \ldots + a_{ik}^{(l)} \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

— многочлен относительно D с постоянными коэффициентами, $D = \frac{d}{dt}$ — оператор дифференцирования.

Матрица операторных коэффициентов

$$A(D) = \| a_{ik}(D) \| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)]$$

является многочленной матрицей или D-матрицей.

Очевидно, что левая элементарная операция 1° над матрицей $A(D)$ означает почленное умножение i-го дифференциального уравнения системы на число $c \neq 0$. Левая элементарная операция 2° означает почленное прибавление к i-му уравнению j-го уравнения, предварительно подвергнутого дифференциальной операции $b(D)$. Левая элементарная операция 3° означает перестановку местами i-го и j-го уравнений.

Таким образом, если в системе уравнений (4) матрицу операторных коэффициентов $A(D)$ заменить левоэквивалентной ей матрицей $B(D)$, то мы получим новую систему уравнений. Поскольку, обратно, исходная система может быть получена из этой системы при помощи аналогичных операций, то обе системы уравнений равносильные [1]).

Нетрудно проинтерпретировать на данном примере и правые элементарные операции. Первая из этих операций означает введение вместо

[1]) При этом принимается, что искомые функции $x_1, x_2, \ldots, x_n$ таковы, что производные от этих функций всех порядков, которые встречаются при преобразованиях, существуют. Тогда две системы уравнений с левоэквивалентными матрицами $A(D)$ и $B(D)$ имеют одни и те же решения.

одной из неизвестных функций x_i новой неизвестной функции $x_i' = \frac{1}{c} x_i$; вторая элементарная операция означает введение новой неизвестной функции $x_j' = x_j - b(D)\, x_i$ (вместо x_j); третья операция означает перемену местами в уравнениях членов, содержащих x_i и x_j (т. е. $x_i = x_j'$, $x_j = x_i'$).

§ 2. Канонический вид λ-матрицы

1. Выясним сначала, к какому сравнительно простому виду можно привести прямоугольную многочленную матрицу $A(\lambda)$ путем применения одних только левых элементарных операций.

Допустим, что в первом столбце матрицы $A(\lambda)$ имеются элементы, не равные тождественно нулю. Возьмем среди них многочлен наименьшей степени и путем перестановки строк сделаем его элементом $a_{11}(\lambda)$. После этого разделим многочлен a_{i1} на $a_{11}(\lambda)$; частное и остаток обозначим через $q_{i1}(\lambda)$ и $r_{i1}(\lambda)$

$$a_{i1}(\lambda) = a_{11}(\lambda)\, q_{i1}(\lambda) + r_{i1}(\lambda) \quad (i = 2, \ldots, m).$$

Вычтем теперь из i-й строки первую строку, предварительно умноженную на $q_{i1}(\lambda)$ $(i = 2, \ldots, m)$. Если при этом не все остатки $r_{i1}(\lambda)$ равны тождественно нулю, то тот из них, который не равен нулю и имеет наименьшую степень, может быть перестановкой строк поставлен на место $a_{11}(\lambda)$. В результате всех этих операций степень многочлена $a_{11}(\lambda)$ понизится.

Теперь мы снова повторим этот процесс и т. д. Так как степень многочлена $a_{11}(\lambda)$ конечна, то на некотором этапе этот процесс уже нельзя будет продолжить, т. е. на этом этапе все элементы $a_{21}(\lambda)$, $a_{31}(\lambda)$, $\ldots$ $\ldots$, $a_{m1}(\lambda)$ окажутся равными тождественно нулю.

После этого возьмем элемент $a_{22}(\lambda)$ и применим ту же процедуру к строкам с номерами $2, 3, \ldots, m$. Тогда добьемся того, что и $a_{32}(\lambda) = \ldots$ $\ldots = a_{m2}(\lambda) = 0$. Продолжая так далее, мы в конце концов приведем матрицу $A(\lambda)$ к следующему виду:

$$\left\| \begin{matrix} b_{11}(\lambda) & b_{12}(\lambda) & \ldots & b_{1m}(\lambda) & \ldots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \ldots & b_{2m}(\lambda) & \ldots & b_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_{mm}(\lambda) & \ldots & b_{mn}(\lambda) \end{matrix} \right\| \quad (m \leqslant n),$$

$$\left\| \begin{matrix} b_{11}(\lambda) & b_{12}(\lambda) & \ldots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \ldots & b_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_{nn}(\lambda) \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{matrix} \right\| \quad (m \geqslant n). \tag{5}$$

Если многочлен $b_{22}(\lambda)$ не равен тождественно нулю, то, применяя левую элементарную операцию второго типа, мы сделаем степень элемента $b_{12}(\lambda)$ меньшей, нежели степень $b_{22}(\lambda)$ (если $b_{22}(\lambda)$ имеет нулевую степень, то $b_{12}(\lambda)$ станет тождественно равен нулю). Точно так же, если $b_{33}(\lambda) \not\equiv 0$, то при помощи левых элементарных операций второго типа мы сделаем степени элементов $b_{13}(\lambda)$, $b_{23}(\lambda)$ меньшими, нежели степень $b_{33}(\lambda)$, не изменив при этом элемента $b_{12}(\lambda)$, и т. д.

Мы установили следующую теорему:

Т е о р е м а 1. *Произвольная прямоугольная многочленная матрица с размерами $m \times n$ при помощи левых элементарных операций всегда может быть приведена к виду* (5), *где многочлены* $b_{1k}(\lambda)$, $b_{2k}(\lambda)$, ..., $b_{k-1,k}(\lambda)$ *имеют меньшую степень, нежели* $b_{kk}(\lambda)$, *если только* $b_{kk}(\lambda) \not\equiv 0$, *и все равны тождественно нулю, если* $b_{kk}(\lambda) = \text{const} \neq 0$ ($k = 2, 3, \ldots, \min(m, n)$).

Совершенно аналогично доказывается

Т е о р е м а 2. *Произвольная прямоугольная многочленная матрица с размерами $m \times n$ при помощи правых элементарных операций всегда может быть приведена к виду*

$$\left\| \begin{array}{cccccc} c_{11}(\lambda) & 0 & \ldots & 0 & 0 \ldots 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \ldots & 0 & 0 \ldots 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \ldots & c_{mm}(\lambda) & 0 \ldots 0 \end{array} \right\|, \quad (m \leqslant n)$$

$$\left\| \begin{array}{cccc} c_{11}(\lambda) & 0 & \ldots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1}(\lambda) & c_{n2}(\lambda) & \ldots & c_{nn}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \ldots & c_{mn}(\lambda) \end{array} \right\|, \quad (m \geqslant n) \qquad (6)$$

где многочлены $c_{k1}(\lambda)$, $c_{k2}(\lambda)$, ..., $c_{k,k-1}(\lambda)$ *имеют меньшую степень, нежели* $c_{kk}(\lambda)$, *если только* $c_{kk}(\lambda) \not\equiv 0$, *и все равны тождественно нулю, если* $c_{kk}(\lambda) = \text{const} \neq 0$ ($k = 2, 3, \ldots, \min(m, n)$).

2. Из теорем 1 и 2 вытекает следующее

С л е д с т в и е. *Если определитель квадратной многочленной матрицы $P(\lambda)$ не зависит от λ и отличен от нуля, то эту матрицу можно представить в виде произведения конечного числа элементарных матриц.*

Действительно, согласно теореме 1 матрицу $P(\lambda)$ при помощи левых элементарных операций можно привести к виду

$$\left\| \begin{array}{cccc} b_{11}(\lambda) & b_{12}(\lambda) & \ldots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \ldots & b_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_{nn}(\lambda) \end{array} \right\|, \qquad (7)$$

где n — порядок матрицы $P(\lambda)$. Так как при применении элементарных операций к квадратной многочленной матрице определитель этой матрицы умножается лишь на постоянный отличный от нуля множитель, то определитель матрицы (7), как и определитель $P(\lambda)$, не зависит от λ и отличен от нуля, т. е.

$$b_{11}(\lambda)\, b_{22}(\lambda) \ldots b_{nn}(\lambda) = \text{const} \neq 0.$$

Отсюда

$$b_{kk}(\lambda) = \text{const} \neq 0 \quad (k = 1, 2, \ldots, n).$$

Но тогда в силу той же теоремы 1 матрица (7) имеет диагональный вид $\| b_{kk}\delta_{ik} \|_1^n$ и потому может быть приведена при помощи левых элементарных операций типа 1° к единичной матрице E. Тогда и обратно, единичную матрицу E можно привести к $P(\lambda)$ при помощи левых элементарных операций с матрицами $S_1, S_2, \ldots, S_p$. Следовательно,

$$P(\lambda) = S_p S_{p-1} \ldots S_1 E = S_p S_{p-1} \ldots S_1.$$

Из доказанного следствия получаем (см. стр. 137—138) равносильность двух определений 2 и 2′ эквивалентности многочленных матриц.

3. Вернемся к нашему примеру системы дифференциальных уравнений (4). Применим теорему 1 к матрице операторных коэффициентов $\| a_{ik}(D) \|$. Тогда, как было указано на стр. 138, система (4) заменится равносильной системой

$$\left.\begin{aligned} b_{11}(D)\,x_1 + b_{12}(D)\,x_2 + \ldots + b_{1s}(D)\,x_s &= \\ &= -b_{1,\,s+1}(D)\,x_{s+1} - \ldots - b_{1n}(D)\,x_n, \\ b_{22}(D)\,x_2 + \ldots + b_{2s}(D)\,x_s &= \\ &= -b_{2,\,s+1}(D)\,x_{s+1} - \ldots - b_{2n}(D)\,x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ b_{ss}(D)\,x_s &= -b_{s,\,s+1}(D)\,x_{s+1} - \ldots - b_{sn}(D)\,x_n, \end{aligned}\right\} \qquad (4')$$

где $s = \min(m, n)$. В этой системе мы функции $x_{s+1}, \ldots, x_n$ можем выбрать произвольно, после чего последовательно определятся функции $x_s, x_{s-1}, \ldots, x_1$, причем на каждом этапе этого определения приходится интегрировать одно дифференциальное уравнение с одной неизвестной функцией.

4. Перейдем теперь к установлению «канонического» вида, к которому можно привести прямоугольную многочленную матрицу $A(\lambda)$, применяя к ней как левые, так и правые элементарные операции.

Среди всех не равных тождественно нулю элементов $a_{ik}(\lambda)$ матрицы $A(\lambda)$ возьмем тот элемент, который имеет наименьшую степень относительно λ, и путем соответствующей перестановки строк и столбцов сделаем его элементом $a_{11}(\lambda)$. После этого найдем частные и остатки от деления многочленов $a_{i1}(\lambda)$ и $a_{1k}(\lambda)$ на $a_{11}(\lambda)$:

$$a_{i1}(\lambda) = a_{11}(\lambda)\,q_{i1}(\lambda) + r_{i1}(\lambda), \qquad a_{1k}(\lambda) = a_{11}(\lambda)\,q_{1k}(\lambda) + r_{1k}(\lambda)$$
$$(i = 2, 3, \ldots, m; \qquad k = 2, 3, \ldots, n).$$

Если хотя бы один из остатков $r_{i1}(\lambda)$, $r_{1k}(\lambda)$ ($i = 2, \ldots, m$; $k = 2, \ldots, n$), например $r_{1k}(\lambda)$, не равен тождественно нулю, то, вычитая из k-го столбца первый столбец, предварительно помноженный на $q_{1k}(\lambda)$, мы заменим элемент $a_{1k}(\lambda)$ остатком $r_{1k}(\lambda)$, который имеет меньшую степень, нежели $a_{11}(\lambda)$. Тогда мы имеем возможность снова уменьшить степень элемента, стоящего в левом верхнем углу матрицы, поместив на это место элемент с наименьшей степенью относительно λ.

Если же все остатки $r_{21}(\lambda), \ldots, r_{m1}(\lambda)$; $r_{12}(\lambda), \ldots, r_{1n}(\lambda)$ равны тождественно нулю, то, вычитая из i-й строки первую, помноженную предварительно на $q_{i1}(\lambda)$ ($i = 2, \ldots, m$), а из k-го столбца — первый, предварительно помноженный на $q_{1k}(\lambda)$ ($k = 2, \ldots, n$), мы приведем нашу многочленную матрицу к виду

$$\begin{Vmatrix} a_{11}(\lambda) & 0 & \ldots & 0 \\ 0 & a_{22}(\lambda) & \ldots & a_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{m2}(0) & \ldots & a_{mn}(\lambda) \end{Vmatrix}.$$

Если при этом хотя бы один из элементов $a_{ik}(\lambda)$ ($i = 2, \ldots, m$; $k = 2, \ldots, n$) не делится без остатка на $a_{11}(\lambda)$, то, прибавляя к первому столбцу тот столбец, который содержит этот элемент, мы придем к предыдущему случаю и, следовательно, снова сможем заменить элемент $a_{11}(\lambda)$ многочленом меньшей степени.

Поскольку первоначальный элемент $a_{11}(\lambda)$ имел определенную степень и процесс уменьшения этой степени не может неограниченно продолжаться, то после конечного числа элементарных операций мы должны получить матрицу вида

$$\begin{Vmatrix} a_1(\lambda) & 0 & \dots & 0 \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & b_{m2}(\lambda) & \dots & b_{mn}(\lambda) \end{Vmatrix}, \tag{8}$$

в которой все элементы $b_{ik}(\lambda)$ делятся без остатка на $a_1(\lambda)$. Если среди этих элементов $b_{ik}(\lambda)$ имеются не равные тождественно нулю, то, продолжая тот же процесс приведения для строк с номерами $2, \dots, m$ и столбцов с номерами $2, \dots, n$, мы матрицу (8) приведем к виду

$$\begin{Vmatrix} a_1(\lambda) & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & 0 & \dots & 0 \\ 0 & 0 & c_{33}(\lambda) & \dots & c_{3n}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & c_{m3}(\lambda) & \dots & c_{mn}(\lambda) \end{Vmatrix},$$

где $a_2(\lambda)$ делится без остатка на $a_1(\lambda)$, а все многочлены $c_{ik}(\lambda)$ делятся без остатка на $a_2(\lambda)$. Продолжая этот процесс далее, мы в конце концов придем к матрице вида

$$\begin{Vmatrix} a_1(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_s(\lambda) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{Vmatrix}, \tag{9}$$

где многочлены $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$ $(s \leqslant m, n)$ не равны тождественно нулю и каждый из них делится без остатка на предыдущий.

Помножая первые s строк на соответствующие отличные от нуля числовые множители, мы сможем добиться того, чтобы старшие коэффициенты многочленов $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$ были равны единице.

Определение 3. Многочленная прямоугольная матрица называется *канонической диагональной*, если она имеет вид (9), где 1) многочлены $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$ не равны тождественно нулю и 2) каждый из многочленов $a_2(\lambda), \dots, a_s(\lambda)$ делится без остатка на предыдущий. При этом предполагается, что старшие коэффициенты всех многочленов $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$ равны единице.

Таким образом, мы доказали, что *произвольная прямоугольная многочленная матрица* $A(\lambda)$ *эквивалентна некоторой канонической диагональной*. В следующем параграфе мы покажем, что *многочлены* $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$ *однозначно определяются заданием матрицы* $A(\lambda)$, и установим формулы, связывающие эти многочлены с элементами матрицы $A(\lambda)$.

§ 3. Инвариантные многочлены и элементарные делители многочленной матрицы

1. Введем понятие об инвариантных многочленах λ-матрицы $A(\lambda)$.

Пусть многочленная матрица $A(\lambda)$ имеет ранг r, т. е. в этой матрице имеются не равные тождественно нулю миноры r-го порядка, в то время как все миноры порядка $>r$ тождественно относительно λ равны нулю. Обозначим через $D_j(\lambda)$ наибольший общий делитель всех миноров j-го порядка матрицы $A(\lambda)$ ($j=1, 2, \ldots, r$)[1]. Тогда, как нетрудно видеть, в ряду

$$D_r(\lambda), \quad D_{r-1}(\lambda), \ldots, D_1(\lambda), \quad D_0(\lambda)\equiv 1$$

каждый многочлен делится без остатка на последующий [2]. Соответствующие частные обозначим через $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$:

$$i_1(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda)=\frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \ \ldots, \ i_r(\lambda)=\frac{D_1(\lambda)}{D_0(\lambda)}=D_1(\lambda). \tag{10}$$

Определение 4. Многочлены $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$, определяемые формулами (10), называются *инвариантными многочленами* прямоугольной матрицы $A(\lambda)$.

Термин «инвариантные многочлены» связан со следующими соображениями. Пусть $A(\lambda)$ и $B(\lambda)$ — две эквивалентные многочленные матрицы. Тогда они получаются друг из друга при помощи элементарных операций. Но нетрудно непосредственно проверить, что элементарные операции не изменяют ни ранга r матрицы $A(\lambda)$, ни самих многочленов $D_1(\lambda)$, $D_2(\lambda)$, $\ldots$, $D_r(\lambda)$. Действительно, применяя к тождеству (3″) формулу, выражающую минор произведения матриц через миноры сомножителей (см. стр. 22), мы для произвольного минора матрицы $B(\lambda)$ получим выражение

$$B\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}=\sum_{\substack{1\leqslant\alpha_1<\alpha_2<\ldots<\alpha_p\leqslant m\\ 1\leqslant\beta_1<\beta_2<\ldots<\beta_p\leqslant m}} P\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ \alpha_1 & \alpha_2 & \ldots & \alpha_p \end{pmatrix} A\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ \beta_1 & \beta_2 & \ldots & \beta_p \end{pmatrix}(\lambda) Q\begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$$

$$(p=1, 2, \ldots, \min(m, n)).$$

Отсюда следует, что все миноры порядка $\geqslant r$ матрицы $B(\lambda)$ равны нулю и, следовательно, для ранга r^* матрицы $B(\lambda)$ имеем:

$$r^*\leqslant r.$$

Кроме того, из этой же формулы вытекает, что $D_p^*(\lambda)$ — наибольший общий делитель всех миноров p-го порядка матрицы $B(\lambda)$ — делится на $D_p(\lambda)$ нацело ($p=1, 2, \ldots, \min(m, n)$). Но матрицы $A(\lambda)$ и $B(\lambda)$ можно поменять ролями. Поэтому $r\leqslant r^*$, и $D_p(\lambda)$ делится без остатка на $D_p^*(\lambda)$ ($p=1, 2, \ldots, \min(m, n)$). Отсюда [3]

$$r=r^*, \quad D_1^*(\lambda)=D_1(\lambda), \quad D_2^*(\lambda)=D_2(\lambda), \ \ldots, \ D_r^*(\lambda)=D_r(\lambda).$$

---

[1]) В каждом $D_j(\lambda)$ берем старший коэффициент равным единице.

[2]) Если к какому-нибудь минору j-го порядка применить разложение Безу по элементам какой-либо строки, то каждое слагаемое в этом разложении будет делиться на $D_{j-1}(\lambda)$; следовательно, любой минор j-го порядка, а значит и $D_j(\lambda)$, делится на $D_{j-1}(\lambda)$ ($j=2, 3, \ldots, r$).

[3]) Старшие коэффициенты у $D_p(\lambda)$ и $D_p^*(\lambda)$ ($p=1, 2, \ldots, r$) равны единице.

Поскольку элементарные операции не меняют многочленов $D_1(\lambda), D_2(\lambda), \ldots, D_r(\lambda)$, то они не изменяют и многочленов $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$, определяемых формулами (10).

Таким образом, многочлены $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$ остаются неизменными, инвариантными при переходе от одной матрицы к другой, ей эквивалентной.

Если многочленная матрица имеет канонический диагональный вид (9), то, как нетрудно видеть, для этой матрицы

$$D_1(\lambda) = a_1(\lambda), \quad D_2(\lambda) = a_1(\lambda)\, a_2(\lambda), \quad \ldots, \quad D_r(\lambda) = a_1(\lambda)\, a_2(\lambda) \ldots a_r(\lambda).$$

Но тогда в силу соотношений (10) диагональные многочлены в (9) $a_1(\lambda)$, $a_2(\lambda)$, $\ldots$, $a_r(\lambda)$ совпадают с инвариантными многочленами

$$i_1(\lambda) = a_r(\lambda), \quad i_2(\lambda) = a_{r-1}(\lambda), \quad \ldots, \quad i_r(\lambda) = a_1(\lambda). \tag{11}$$

Здесь $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$ являются одновременно и инвариантными многочленами исходной матрицы $A(\lambda)$, поскольку эта матрица эквивалентна матрице (9).

Полученные результаты мы можем сформулировать в виде следующей теоремы.

Теорема 3. *Многочленная прямоугольная матрица $A(\lambda)$ всегда эквивалентна канонической диагональной матрице*

$$\begin{Vmatrix} i_r(\lambda) & 0 & \ldots & 0 & 0 \ldots 0 \\ 0 & i_{r-1}(\lambda) & \ldots & 0 & 0 \ldots 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & i_1(\lambda) & 0 \ldots 0 \\ 0 & 0 & \ldots & 0 & 0 \ldots 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 \ldots 0 \end{Vmatrix}. \tag{12}$$

При этом здесь обязательно r — ранг, а $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$ — инвариантные многочлены матрицы $A(\lambda)$, определяемые формулами (10).

Следствие 1. *Для того чтобы две прямоугольные матрицы одинаковых размеров $A(\lambda)$ и $B(\lambda)$ были эквивалентны, необходимо и достаточно, чтобы они имели одни и те же инвариантные многочлены.*

Действительно, необходимость этого условия была выяснена выше. Достаточность следует из того, что две многочленные матрицы, имеющие одни и те же инвариантные многочлены, эквивалентны одной и той же канонической диагональной матрице и, следовательно, эквивалентны между собой.

Таким образом, *инвариантные многочлены образуют полную систему инвариантов λ-матрицы.*

Следствие 2. В ряду инвариантных многочленов

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \ldots, \quad i_r(\lambda) = \frac{D_1(\lambda)}{D_2(\lambda)} \qquad (D_0(\lambda) \equiv 1) \tag{13}$$

каждый многочлен, начиная со второго, является делителем предыдущего.

Это утверждение не вытекает непосредственно из формул (13). Оно следует из того, что многочлены $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_r(\lambda)$ совпадают с многочленами $a_r(\lambda)$, $a_{r-1}(\lambda)$, $\ldots$, $a_1(\lambda)$ канонической диагональной матрицы (9).

2. Укажем методы вычисления инвариантных многочленов для квазидиагональных λ-матриц, если известны инвариантные многочлены матриц, стоящих в диагональных клетках.

Теорема 4. *Если в квазидиагональной прямоугольной матрице*

$$C(\lambda)=\begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

любой инвариантный многочлен матрицы $A(\lambda)$ является делителем любого инвариантного многочлена матрицы $B(\lambda)$, то совокупность инвариантных многочленов матрицы $C(\lambda)$ получается объединением инвариантных многочленов матриц $A(\lambda)$ и $B(\lambda)$.

Доказательство. Обозначим через $i'_1(\lambda)$, $i'_2(\lambda)$, ..., $i'_r(\lambda)$ и $i''_1(\lambda)$, $i''_2(\lambda)$, ..., $i''_q(\lambda)$ соответственно инвариантные многочлены λ-матриц $A(\lambda)$ и $B(\lambda)$. Тогда [1])

$$A(\lambda)\sim\{i'_r(\lambda),\ \ldots,\ i'_1(\lambda),\ 0,\ \ldots,\ 0\},\qquad B(\lambda)\sim\{i''_q(\lambda),\ \ldots,\ i''_1(\lambda),\ 0,\ \ldots,\ 0\}$$

и, следовательно,

$$C(\lambda)\sim\{i'_r(\lambda),\ \ldots,\ i'_1(\lambda),\quad i''_q(\lambda),\ \ldots,\ i''_1(\lambda),\ 0,\ \ldots,\ 0\}. \tag{14}$$

λ-матрица, стоящая в правой части этого соотношения, имеет каноническую диагональную форму. Тогда согласно теореме 3 не равные тождественно нулю диагональные элементы этой матрицы образуют полную систему инвариантных многочленов матрицы $C(\lambda)$. Теорема доказана.

Для того чтобы в общем случае при произвольных инвариантных многочленах матриц $A(\lambda)$ и $B(\lambda)$ определить инвариантные многочлены $C(\lambda)$, мы воспользуемся важным понятием об элементарных делителях.

Разложим инвариантные многочлены $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$ на неприводимые в данном числовом поле K множители:

$$\begin{aligned} i_1(\lambda)&=[\varphi_1(\lambda)]^{c_1}[\varphi_2(\lambda)]^{c_2}\ldots[\varphi_s(\lambda)]^{c_s},\\ i_2(\lambda)&=[\varphi_1(\lambda)]^{d_1}[\varphi_2(\lambda)]^{d_2}\ldots[\varphi_s(\lambda)]^{d_s},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ i_r(\lambda)&=[\varphi_1(\lambda)]^{l_1}[\varphi_2(\lambda)]^{l_2}\ldots[\varphi_s(\lambda)]^{l_s} \end{aligned}\qquad \begin{pmatrix} c_k\geqslant d_k\geqslant\ldots\geqslant l_k\geqslant 0;\\ k=1,\ 2,\ \ldots,\ s \end{pmatrix}\ [2]). \tag{15}$$

Здесь $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, ..., $\varphi_s(\lambda)$ — все различные неприводимые в поле K многочлены (со старшими коэффициентами, равными единице), входящие в состав $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$.

Определение 5. Все отличные от единицы степени среди $[\varphi_1(\lambda)]^{c_1}$, ..., $[\varphi_s(\lambda)]^{l_s}$ в разложении (15) называются *элементарными делителями* матрицы $A(\lambda)$ в поле K [3]).

Теорема 5. *Совокупность элементарных делителей прямоугольной квазидиагональной матрицы*

$$C(\lambda)=\begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

---

[1]) Знаком $\sim$ мы здесь обозначаем эквивалентность матриц, а фигурными скобками $\{\ \}$ — диагональную прямоугольную матрицу вида (12).

[2]) Некоторые из показателей c_k, d_k, ..., l_k ($k=1,\ 2,\ \ldots,\ s$) могут равняться нулю.

[3]) Формулы (15) дают возможность не только по инвариантным многочленам определить элементарные делители $A(\lambda)$ в поле K, но и наоборот, — по элементарным делителям определить инвариантные многочлены.

всегда получается объединением элементарных делителей матрицы $A(\lambda)$ *с элементарными делителями матрицы* $B(\lambda)$.

Доказательство. Разложим инвариантные многочлены матриц $A(\lambda)$ и $B(\lambda)$ на неприводимые в поле K множители[1]):

$$i'_1(\lambda)=[\varphi_1(\lambda)]^{c'_1}[\varphi_2(\lambda)]^{c'_2}\dots[\varphi_s(\lambda)]^{c'_s},\quad i''_1(\lambda)=[\varphi_1(\lambda)]^{c''_1}[\varphi_2(\lambda)]^{c''_2}\dots[\varphi_s(\lambda)]^{c''_s},$$
$$i'_2(\lambda)=[\varphi_1(\lambda)]^{d'_1}[\varphi_2(\lambda)]^{d'_2}\dots[\varphi_s(\lambda)]^{d'_s},\quad i''_2(\lambda)=[\varphi_1(\lambda)]^{d''_1}[\varphi_2(\lambda)]^{d''_2}\dots[\varphi_s(\lambda)]^{d''_s},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$i'_r(\lambda)=[\varphi_1(\lambda)]^{h'_1}[\varphi_2(\lambda)]^{h'_2}\dots[\varphi_s(\lambda)]^{h'_s},\quad i''_q(\lambda)=[\varphi_1(\lambda)]^{g''_1}[\varphi_2(\lambda)]^{g''_2}\dots[\varphi_s(\lambda)]^{g''_s}.$$

Обозначим через

$$c_1\geqslant d_1\geqslant\dots\geqslant l_1>0 \tag{16}$$

все отличные от нуля числа среди $c'_1, d'_1, \dots, h'_1, c''_1, d''_1, \dots, g''_1$.

Тогда матрица $C(\lambda)$ эквивалентна матрице (14), а последняя перестановкой строк и столбцов может быть приведена к «диагональному» виду

$$\{[\varphi_1(\lambda)]^{c_1}\cdot(*),\ [\varphi_1(\lambda)]^{d_1}\cdot(*),\ \dots,\ [\varphi_1(\lambda)[^{l_1}\cdot(*),\ (**),\ \dots,\ (**)\}, \tag{17}$$

где через $(*)$ мы обозначили многочлены, взаимно простые с $\varphi_1(\lambda)$, а через $(**)$ — многочлены, взаимно простые с $\varphi_1(\lambda)$ либо тождественно равные нулю. Из вида матрицы (17) непосредственно вытекают следующие разложения многочленов $D_r(\lambda)$, $D_{r-1}(\lambda)$, ... и $i_1(\lambda)$, $i_2(\lambda)$, ... для матрицы $C(\lambda)$:

$$D_r(\lambda)=[\varphi_1(\lambda)]^{c_1+d_1+\dots+l_1}\cdot(*),\qquad D_{r-1}(\lambda)=[\varphi_1(\lambda)]^{d_1+\dots+l_1}\cdot(*)\ \text{ и т. д.},$$
$$i_1(\lambda)=[\varphi_1(\lambda)]^{c_1}(*),\quad i_2(\lambda)=[\varphi_1(\lambda)]^{d_1}(*)\ \text{ и т. д.}$$

Отсюда следует, что $[\varphi_1(\lambda)]^{c_1}$, $[\varphi_1(\lambda)]^{d_1}$, ..., $[\varphi_1(\lambda)]^{l_1}$, т. е. все не равные единице из степеней

$$[\varphi_1(\lambda)]^{c'_1},\ \dots,\ [\varphi_1(\lambda)]^{h'_1},\quad [\varphi_1(\lambda)]^{c''_1},\ \dots,\ [\varphi_1(\lambda)]^{g''_1},$$

являются элементарными делителями матрицы $C(\lambda)$.

Аналогично определяются элементарные делители матрицы $C(\lambda)$, являющиеся степенями $\varphi_2(\lambda)$ и т. д. Теорема доказана.

Примечание. Совершенно аналогично предыдущему можно построить теорию эквивалентности для целочисленных матриц (т. е. матриц, у которых элементы — целые числа). При этом в 1°, 2° (см. стр. 135) $c=\pm 1$, $b(\lambda)$ заменяется целым числом, а в формулах (3), (3′), (3″) вместо $P(\lambda)$ и $Q(\lambda)$ стоят целочисленные матрицы с определителями, равными ± 1.

3. Пусть дана теперь матрица $A=\|a_{ik}\|_1^n$ с элементами из поля K. Составим для нее характеристическую матрицу

$$\lambda E-A=\begin{Vmatrix}\lambda-a_{11} & -a_{12} & \dots & -a_{1n}\\ -a_{21} & \lambda-a_{12} & \dots & -a_{2n}\\ \dots & \dots & \dots & \dots\\ -a_{n1} & -a_{n2} & \dots & -a_{nn}\end{Vmatrix}. \tag{18}$$

[1]) Если какой-либо неприводимый многочлен $\varphi_k(\lambda)$ входит множителем в одни инвариантные многочлены и не входит в другие, то в эти последние многочлены мы вписываем $\varphi_k(\lambda)$ с нулевым показателем.

Характеристическая матрица является λ-матрицей ранга n. Ее инвариантные многочлены

$$i_1(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}, \quad i_2(\lambda)=\frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \ \ldots, \ i_n(\lambda)=\frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda)\equiv 1) \qquad (19)$$

называются *инвариантными многочленами матрицы* A, а соответствующие элементарные делители в поле K — *элементарными делителями в поле* K *матрицы* A. Первый инвариантный многочлен $i_1(\lambda)$ совпадает с минимальным многочленом $\psi(\lambda)$ матрицы A [1]). Знание инвариантных многочленов (а следовательно, и элементарных делителей) матрицы A позволяет исследовать ее структуру. Поэтому представляют интерес практические способы вычисления инвариантных многочленов матрицы. Сами формулы (19) дают алгоритм для вычисления этих многочленов, но этот алгоритм при больших n очень громоздок.

Теорема 3 дает другой способ вычисления инвариантных многочленов, основанный на приведении характеристической матрицы (18) при помощи элементарных операций к каноническому диагональному виду.

Пример.

$$A=\begin{Vmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{Vmatrix}, \qquad \lambda E-A=\begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14 & 5 & 1 & \lambda \end{Vmatrix}.$$

В характеристической матрице $\lambda E-A$ прибавим к четвертой строке третью, предварительно помноженную на λ:

$$\begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

Теперь, прибавляя к первым трем столбцам четвертый, предварительно помноженный соответственно на -6, -1, $\lambda-2$, получим:

$$\begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

К первому столбцу прибавляем второй, помноженный на $\lambda-3$:

$$\begin{Vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

Прибавим ко второй и четвертой строкам первую, помноженную соответственно на $\lambda+1$ и $5-\lambda$; найдем:

$$\begin{Vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 0 & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

[1]) См. формулу (49) на стр. 99; там $\Delta(\lambda)\equiv D_n(\lambda)$.

Прибавим к четвертой строке вторую, затем умножим первую и третью строки на -1. После перестановки строк и столбцов получим:

$$\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (\lambda-1)^2 & 0 \\ 0 & 0 & 0 & (\lambda-1)^2 \end{Vmatrix}.$$

Матрица A имеет два элементарных делителя: $(\lambda-1)^2$ и $(\lambda-1)^2$.

§ 4. Эквивалентность линейных двучленов

В предыдущих параграфах мы рассматривали прямоугольные λ-матрицы. В этом же параграфе мы рассмотрим две **квадратные** λ-матрицы $A(\lambda)$ и $B(\lambda)$ n-го порядка, у которых все элементы имеют степень не выше единицы относительно λ. Эти многочленные матрицы могут быть представлены в виде матричных двучленов:

$$A(\lambda)=A_0\lambda+A_1, \quad B(\lambda)=B_0\lambda+B_1.$$

Мы будем предполагать, что эти двучлены имеют первую степень и регулярны, т. е. что $|A_0|\neq 0$, $|B_0|\neq 0$ (см. стр. 87).

Следующая теорема устанавливает критерий эквивалентности таких двучленов:

Теорема 6. *Если два регулярных двучлена первой степени $A_0\lambda+A_1$ и $B_0\lambda+B_1$ эквивалентны, то эти двучлены строго эквивалентны, т. е. в тождестве*

$$B_0\lambda+B_1=P(\lambda)(A_0\lambda+A_1)Q(\lambda) \tag{20}$$

можно $P(\lambda)$ и $Q(\lambda)$ — матрицы с постоянными и отличными от нуля определителями — заменить постоянными неособенными матрицами P и Q[1]:

$$B_0\lambda+B_1=P(A_0\lambda+A_1)Q. \tag{21}$$

Доказательство. Так как определитель матрицы $P(\lambda)$ не зависит от λ и отличен от нуля[2]), то обратная матрица $M(\lambda)=P^{-1}(\lambda)$ также будет многочленной. Пользуясь этой матрицей, мы тождество (20) перепишем в виде

$$M(\lambda)(B_0\lambda+B_1)=(A_0\lambda+A_1)Q(\lambda). \tag{22}$$

Рассматривая $M(\lambda)$ и $Q(\lambda)$ как матричные многочлены, разделим $M(\lambda)$ слева на $A_0\lambda+A_1$, а $Q(\lambda)$ — справа на $B_0\lambda+B_1$:

$$M(\lambda)=(A_0\lambda+A_1)S(\lambda)+M, \tag{23}$$

$$Q(\lambda)=T(\lambda)(B_0\lambda+B_1)+Q; \tag{24}$$

здесь M и Q — постоянные (не зависящие от λ) квадратные матрицы

---

[1]) Тождество (21) равносильно двум матричным равенствам: $B_0=PA_0Q$ и $B_1=PA_1Q$.

[2]) Эквивалентность двучленов $A_0\lambda+A_1$ и $B_0\lambda+B_1$ означает наличие тождества (20), в котором $|P(\lambda)|=\text{const}\neq 0$ и $|Q(\lambda)|=\text{const}\neq 0$. Однако последние соотношения в данном случае вытекают из самого тождества (20). Действительно, определители регулярных двучленов первой степени имеют степень n: $|A_0\lambda+A_1|=|A_0|\lambda^n+\ldots$, $|B_0\lambda+B_1|=|B_0|\lambda^n+\ldots$; $|A_0|\neq 0$, $|B_0|\neq 0$. Поэтому из $|B_0\lambda+B_1|=|P(\lambda)||A_0\lambda+A_1||Q(\lambda)|$ следует:

$$|P(\lambda)|=\text{const}\neq 0, \quad |Q(\lambda)|=\text{const}\neq 0.$$

n-го порядка. Полученные выражения для $M(\lambda)$ и $Q(\lambda)$ поставим в (22). После небольших преобразований получим:

$$(A_0\lambda + A_1)[T(\lambda) - S(\lambda)](B_0\lambda + B_1) = M(B_0\lambda + B_1) - (A_0\lambda + A_1)Q. \quad (25)$$

Разность, стоящая в квадратных скобках, должна тождественно равняться нулю, так как в противном случае произведение, стоящее в левой части равенства (25), имело бы степень $\geqslant 2$, в то время как в правой части этого же равенства стоит многочлен не выше первой степени. Поэтому

$$S(\lambda) = T(\lambda); \quad (26)$$

но тогда из (25) получим:

$$M(B_0\lambda + B_1) = (A_0\lambda + A_1)Q. \quad (27)$$

Покажем теперь, что M — неособенная матрица. Для этого разделим $P(\lambda)$ слева на $B_0\lambda + B_1$:

$$P(\lambda) = (B_0\lambda + B_1)U(\lambda) + P. \quad (28)$$

Из (22), (23) и (28) следует:

$$\begin{aligned} E = M(\lambda)P(\lambda) &= M(\lambda)(B_0\lambda + B_1)U(\lambda) + M(\lambda)P = \\ &= (A_0\lambda + A_1)Q(\lambda)U(\lambda) + (A_0\lambda + A_1)S(\lambda)P + MP = \\ &= (A_0\lambda + A_1)[Q(\lambda)U(\lambda) + S(\lambda)P] + MP. \end{aligned} \quad (29)$$

Так как последняя часть этой цепочки равенств должна иметь нулевую степень относительно λ (поскольку она равна E), то выражение в квадратных скобках должно тождественно равняться нулю. Но тогда из (29)

$$MP = E. \quad (30)$$

Отсюда следует: $|M| \neq 0$ и $M^{-1} = P$.

Умножая обе части равенства (27) слева на P, получим:

$$B_0\lambda + B_1 = P(A_0\lambda + B_1)Q.$$

Неособенность матрицы P следует из (30). Но неособенность матриц P и Q следует и из самого тождества (21), так как из этого тождества вытекает равенство

$$B_0 = PA_0Q$$

и потому

$$|P||A_0||Q| = |B_0| \neq 0.$$

Теорема доказана.

Примечание. Из доказательства следует [см. (24) и (28)], что в качестве постоянных матриц P и Q, которыми мы заменяем λ-матрицы $P(\lambda)$ и $Q(\lambda)$ в тождестве (20), можно взять соответственно левый и правый остатки от деления $P(\lambda)$ и $Q(\lambda)$ на $B_0\lambda + B_1$.

§ 5. Критерий подобия матриц

Пусть дана матрица $A = \|a_{ik}\|_1^n$ с числовыми элементами из поля K. Ее характеристическая матрица $\lambda E - A$ является λ-матрицей ранга n и потому имеет n инвариантных многочленов (см. § 3)

$$i_1(\lambda),\ i_2(\lambda),\ \ldots,\ i_n(\lambda).$$

Следующая теорема показывает, что эти инвариантные многочлены определяют исходную матрицу A с точностью до преобразования подобия.

Теорема 7. *Для того чтобы две матрицы* $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ *были подобны* ($B = T^{-1}AT$), *необходимо и достаточно, чтобы они имели одни и те же инвариантные многочлены, или, что то же, одни и те же элементарные делители в поле* K.

Доказательство. Условие необходимо. Действительно, если матрицы A и B подобны, то существует такая неособенная матрица T, что

$$B = T^{-1}AT.$$

Отсюда

$$\lambda E - B = T^{-1}(\lambda E - A)\,T.$$

Это равенство показывает, что характеристические матрицы $\lambda E - A$ и $\lambda E - B$ эквивалентны и потому имеют одни и те же инвариантные многочлены.

Условие достаточно. Пусть характеристические матрицы $\lambda E - A$ и $\lambda E - B$ имеют одни и те же инвариантные многочлены. Тогда эти λ-матрицы эквивалентны (см. следствие 1 из теоремы 3), и, следовательно, существуют две многочленные матрицы $P(\lambda)$ и $Q(\lambda)$ такие, что

$$\lambda E - B = P(\lambda)(\lambda E - A)\,Q(\lambda). \tag{31}$$

Применяя к матричным двучленам $\lambda E - A$ и $\lambda E - B$ теорему 6, мы можем в тождестве (31) заменить λ-матрицы $P(\lambda)$ и $Q(\lambda)$ постоянными матрицами P и Q:

$$\lambda E - B = P(\lambda E - A)\,Q, \tag{32}$$

причем в качестве P и Q можно взять (см. примечание на стр. 124) соответственно левый и правый остатки от деления $P(\lambda)$ и $Q(\lambda)$ на $\lambda E - B$, т. е. на основании обобщенной теоремы Безу можно положить[1]):

$$P = \hat{P}(B), \quad Q = Q(B). \tag{33}$$

Приравнивая в обеих частях равенства (32) коэффициенты при нулевой и при первой степенях λ, получим:

$$B = PAQ, \quad E = PQ,$$

т. е.

$$B = T^{-1}AT,$$

где

$$T = Q = P^{-1}.$$

Теорема доказана.

Замечание. Попутно нами установлено следующее предложение, которое мы сформулируем как

Добавление к теореме 7. *Если* $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ — *две подобные матрицы*

$$B = T^{-1}AT, \tag{34}$$

то в качестве преобразующей матрицы T *можно взять матрицу*

$$T = Q(B) = [\hat{P}(B)]^{-1}, \tag{35}$$

---

[1]) Напоминаем, что $\hat{P}(B)$ — левое значение многочлена $P(\lambda)$, а $Q(B)$ — правое значение многочлена $Q(\lambda)$ после замены λ на B (см. стр. 90—92).

где $P(\lambda)$ *и* $Q(\lambda)$ *— многочленные матрицы в тождестве*

$$\lambda E - B = P(\lambda)(\lambda E - A)Q(\lambda),$$

связывающем эквивалентные характеристические матрицы $\lambda E - A$ *и* $\lambda E - B$; *в формуле* (35) $Q(B)$ *обозначает правое значение матричного многочлена* $Q(\lambda)$, *а* $\hat{P}(B)$ *— левое значение матричного многочлена* $P(\lambda)$ *при замене аргумента* λ *матрицей* B.

§ 6. Нормальные формы матрицы

1. Пусть дан некоторый многочлен с коэффициентами из поля K

$$g(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_{m-1}\lambda + \alpha_m.$$

Рассмотрим квадратную матрицу m-го порядка

$$L = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\alpha_m \\ 1 & 0 & \ldots & 0 & -\alpha_{m-1} \\ 0 & 1 & \ldots & 0 & -\alpha_{m-2} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & \ldots & 1 & -\alpha_1 \end{Vmatrix}. \tag{36}$$

Нетрудно проверить, что многочлен $g(\lambda)$ является характеристическим многочленом матрицы L:

$$|\lambda E - L| = \begin{vmatrix} \lambda & 0 & 0 & \ldots & 0 & \alpha_m \\ -1 & \lambda & 0 & \ldots & 0 & \alpha_{m-1} \\ 0 & -1 & \lambda & \ldots & 0 & \alpha_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & -1 & \alpha_1 + \lambda \end{vmatrix} = g(\lambda).$$

С другой стороны, минор элемента α_m в характеристическом определителе равен ± 1. Поэтому $D_{m-1}(\lambda) = 1$ и $i_1(\lambda) = \dfrac{D_m(\lambda)}{D_{m-1}(\lambda)} = D_m(\lambda) = g(\lambda)$, $i_2(\lambda) = \ldots = i_n(\lambda) = 1$.

Таким образом, матрица L имеет единственный отличный от единицы инвариантный многочлен, равный $g(\lambda)$.

Матрицу L мы будем называть *сопровождающей* матрицей для многочлена $g(\lambda)$.

Пусть дана матрица $A = \|a_{ik}\|_1^n$ с инвариантными многочленами

$$i_1(\lambda),\ i_2(\lambda),\ \ldots,\ i_t(\lambda),\ i_{t+1}(\lambda) = 1,\ \ldots,\ i_n(\lambda) = 1. \tag{37}$$

Здесь все многочлены $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_t(\lambda)$ имеют степень выше нулевой, причем каждый из этих многочленов, начиная со второго, является делителем предыдущего. Сопровождающие матрицы для этих многочленов обозначим через L_1, L_2, $\ldots$, L_t.

Тогда квазидиагональная матрица n-го порядка

$$L_I = \{L_1, L_2, \ldots, L_t\} \tag{38}$$

имеет своими инвариантными многочленами многочлены (37) (см. теорему 4 на стр. 145). Поскольку матрицы A и L_I имеют одни и те же

инвариантные многочлены, они подобны, т. е. существует всегда такая неособенная матрица U ($|U| \neq 0$), что

$$A = UL_I U^{-1}. \tag{I}$$

Матрица L_I называется *первой естественной нормальной формой* для матрицы A. Эта нормальная форма характеризуется: 1) квазидиагональным видом (38), 2) специальной структурой диагональных клеток (36) и 3) дополнительным условием: в ряду характеристических многочленов диагональных клеток каждый многочлен, начиная со второго, является делителем предыдущего [1]).

2. Обозначим теперь через

$$\chi_1(\lambda),\ \chi_2(\lambda),\ \ldots,\ \chi_u(\lambda) \tag{39}$$

элементарные делители матрицы $A = \|a_{ik}\|_1^n$ в числовом поле K. Соответствующие сопровождающие матрицы обозначим через

$$L^{(1)},\ L^{(2)},\ \ldots,\ L^{(u)}.$$

Поскольку $\chi_j(\lambda)$ — единственный элементарный делитель матрицы $L^{(j)}$ ($j = 1, 2, \ldots, u$) [2]), то согласно теореме 5 квазидиагональная матрица

$$L_{II} = \{L^{(1)}, L^{(2)}, \ldots, L^{(u)}\} \tag{40}$$

имеет своими элементарными делителями многочлены (39).

Матрицы A и L_{II} имеют одни и те же элементарные делители в поле K. Поэтому эти матрицы подобны, т. е. существует всегда такая неособенная матрица V ($|V| \neq 0$), что

$$A = VL_{II}V^{-1}. \tag{II}$$

Матрица L_{II} называется *второй естественной нормальной формой* для матрицы A. Эта нормальная форма характеризуется: 1) квазидиагональным видом (40), 2) специальной структурой диагональных клеток (36) и 3) дополнительным условием: характеристический многочлен каждой диагональной клетки представляет собой степень неприводимого в поле K многочлена.

Замечание. Элементарные делители матрицы A в отличие от инвариантных многочленов существенно связаны с данным числовым полем K. Если мы вместо исходного числового поля K возьмем другое числовое поле (которому также принадлежат элементы данной матрицы A), то элементарные делители могут измениться. Вместе с элементарными делителями изменится и вторая естественная нормальная форма матрицы.

Так, например, пусть дана матрица $A = \|a_{ik}\|_1^n$ с вещественными элементами. Характеристический многочлен этой матрицы будет иметь вещественные коэффициенты. В то же время этот многочлен может иметь комплексные корни. Если K — поле вещественных чисел, то среди элементарных делителей могут быть и степени неприводимых квадратных

---

[1]) Из условий 1), 2), 3) автоматически следует, что характеристические многочлены диагональных клеток в L_I являются инвариантными многочленами матрицы L_I и, следовательно, матрицы A.

[2]) $\chi_j(\lambda)$ — единственный инвариантный многочлен матрицы $L^{(j)}$ и в то же время $\chi_j(\lambda)$ есть степень неприводимого в поле K многочлена.

трехчленов с вещественными коэффициентами. Если K — поле комплексных чисел, то каждый элементарный делитель имеет вид $(\lambda-\lambda_0)^p$.

3. Допустим теперь, что числовое поле K содержит не только элементы матрицы A, но и все характеристические числа этой матрицы [1]). Тогда элементарные делители матрицы A имеют вид [2])

$$(\lambda-\lambda_1)^{p_1},\ (\lambda-\lambda_2)^{p_2},\ \ldots,\ (\lambda-\lambda_u)^{p_u} \qquad (p_1+p_2+\ldots+p_u=n). \tag{41}$$

Рассмотрим один из таких элементарных делителей

$$(\lambda-\lambda_0)^p$$

и поставим ему в соответствие следующую матрицу порядка p:

$$\begin{Vmatrix} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{Vmatrix} = \lambda_0 E^{(p)} + H^{(p)}. \tag{42}$$

Нетрудно проверить, что эта матрица имеет только один элементарный делитель $(\lambda-\lambda_0)^p$. Матрицу (42) мы будем называть *жордановой клеткой*, соответствующей элементарному делителю $(\lambda-\lambda_0)^p$.

Жордановы клетки, соответствующие элементарным делителям (41), обозначим через

$$J_1,\ J_2,\ \ldots,\ J_u.$$

Тогда квазидиагональная матрица

$$J=\{J_1,\ J_2,\ \ldots,\ J_u\}$$

имеет своими элементарными делителями степени (41).

Матрицу J можно еще записать так:

$$J=\{\lambda_1E_1+H_1,\ \lambda_2E_2+H_2,\ \ldots,\ \lambda_uE_u+H_u\},$$

где

$$E_k=E^{(p_k)},\quad H_k=H^{(p_k)} \qquad (k=1,\ 2,\ \ldots,\ u).$$

Поскольку матрицы A и J имеют одни и те же элементарные делители, они подобны между собой, т. е. существует такая неособенная матрица T ($|T|\neq 0$), что

$$A=TJT^{-1}=T\{\lambda_1E_1+H_1,\ \lambda_2E_2+H_2,\ \ldots,\ \lambda_uE_u+H_u\}T^{-1}. \tag{III}$$

Матрица J называется *жордановой нормальной формой* или просто *жордановой формой* матрицы A. Жорданова форма характеризуется квазидиагональным видом и специальной структурой (42) диагональных клеток.

---

[1]) Это всегда имеет место для любой матрицы A, если K — поле комплексных чисел.

[2]) Среди чисел $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_u$ могут быть и равные между собой.

На нижеследующей схеме выписана жорданова матрица J при элементарных делителях $(\lambda-\lambda_1)^2$, $(\lambda-\lambda_2)^3$, $\lambda-\lambda_3$, $(\lambda-\lambda_4)^2$:

$$J=\left\|\begin{array}{cc|ccccc|cc} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \end{array}\right\|. \tag{43}$$

Если все элементарные делители матрицы A — первой степени (и только в этом случае), жорданова форма является диагональной матрицей, и в этом случае мы имеем:

$$A=T\{\lambda_1, \lambda_2, \ldots, \lambda_n\}T^{-1}. \tag{44}$$

Таким образом, *матрица A имеет простую структуру* (см. гл. III, § 8) *в том и только в том случае, когда все ее элементарные делители имеют первую степень* [1]).

Иногда вместо жордановой клетки (42) рассматривают *«нижнюю» жорданову клетку* p-го порядка

$$\left\|\begin{array}{ccccc} \lambda_0 & 0 & \ldots & 0 & 0 \\ 1 & \lambda_0 & \ldots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda_0 & 0 \\ 0 & \ldots & 0 & 1 & \lambda_0 \end{array}\right\| = \lambda_0 E^{(p)} + F^{(p)}.$$

Эта матрица также имеет только один элементарный делитель $(\lambda-\lambda_0)^p$. Элементарным делителям (41) соответствует *«нижняя» жорданова матрица* [2])

$$J_{(1)}=\{\lambda_1 E_1+F_1, \lambda_2 E_2+F_2, \ldots, \lambda_u E_u+F_u\}$$
$$(E_k=E^{(p_k)}, F_k=F^{(p_k)}; k=1, 2, \ldots, u).$$

Произвольная матрица A, имеющая элементарные делители (41), всегда подобна матрице $J_{(1)}$, т. е. существует такая неособенная матрица T_1 ($|T_1|\neq 0$), что

$$A=T_1 J_{(1)} T_1^{-1}=T_1\{\lambda_1 E_1+F_1, \lambda_2 E_2+F_2, \ldots, \lambda_n E_n+F_n\}T_1^{-1}. \tag{IV}$$

Заметим еще, что если $\lambda_0\neq 0$, то каждая из матриц

$$\lambda_0(E^{(p)}+H^{(p)}), \quad \lambda_0(E^{(p)}+F^{(p)})$$

---

[1]) Часто вместо «элементарные делители первой степени» говорят «линейные элементарные делители» или «простые элементарные делители».

[2]) В отличие от нижней жордановой матрицы $J_{(1)}$ матрицу J иногда называют *верхней жордановой матрицей*.

имеет только один элементарный делитель: $(\lambda-\lambda_0)^p$. Поэтому для неособенной матрицы A, имеющей элементарные делители (41), наряду с (III) и (IV) имеют место представления

$$A=T_2\{\lambda_1(E_1+H_1),\ \lambda_2(E_2+H_2),\ \ldots,\ \lambda_u(E_u+H_u)\}T_2^{-1}, \tag{V}$$

$$A=T_3\{\lambda_1(E_1+F_1),\ \lambda_2(E_2+F_2),\ \ldots,\ \lambda_u(E_u+F_u)\}T_3^{-1}. \tag{VI}$$

§ 7. Элементарные делители матрицы $f(A)$

1. В настоящем параграфе рассмотрим следующую задачу:

Даны элементарные делители (в поле комплексных чисел) матрицы $A=\|a_{ik}\|_1^n$ *и дана функция* $f(\lambda)$, *определенная на спектре матрицы* A. *Требуется определить элементарные делители (в поле комплексных чисел) матрицы* $f(A)$.

Обозначим через

$$(\lambda-\lambda_1)^{p_1},\quad (\lambda-\lambda_2)^{p_2},\quad \ldots,\quad (\lambda-\lambda_u)^{p_u}$$

элементарные делители матрицы A [1]). Тогда матрица A подобна жордановой матрице J

$$A=TJT^{-1}$$

и, следовательно (см. 2° на стр. 106),

$$f(A)=Tf(J)T^{-1}.$$

При этом

$$J=\{J_1, J_2, \ldots, J_u\},\quad J_i=\lambda_iE^{(p_i)}+H^{(p_i)}\qquad (i=1, 2, \ldots, u),$$

а

$$f(J)=\{f(J_1), f(J_2), \ldots, f(J_u)\}, \tag{45}$$

где (см. пример 2 на стр. 106)

$$f(J_i)=\begin{Vmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \cdots & \frac{f^{(p_i-1)}(\lambda_i)}{(p_i-1)!} \\ 0 & f(\lambda_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \frac{f'(\lambda_i)}{1!} \\ 0 & 0 & \cdots & f(\lambda_i) \end{Vmatrix} (i=1, \ldots, u). \tag{46}$$

Поскольку подобные матрицы $f(A)$ и $f(J)$ имеют одни и те же элементарные делители, то в дальнейшем вместо матрицы $f(A)$ мы будем рассматривать матрицу $f(J)$.

2. Определим сначала дефект d матрицы $f(A)$ или, что то же, матрицы $f(J)$ [2]). Дефект квазидиагональной матрицы равен сумме дефектов отдельных диагональных клеток, а дефект матрицы $f(J_i)$ [см. (46)] равен

---

[1]) Среди чисел $\lambda_1, \lambda_2, \ldots, \lambda_u$ могут быть и равные между собой.

[2]) $d=n-r$, где r—ранг матрицы $f(A)$.

наименьшему из чисел k_i и p_i, где k_i — кратность λ_i как корня $f(\lambda)$ [1]), поскольку

$$f(\lambda_i)=f'(\lambda_i)=\ldots=f^{(k_i-1)}(\lambda_i)=0,\quad f^{(k_i)}(\lambda_i)\neq 0 \qquad (i=1, 2, \ldots, u). \tag{46}$$

Мы пришли к теореме:

Теорема 8. *Дефект матрицы $f(A)$, где матрица A имеет элементарные делители*

$$(\lambda-\lambda_1)^{p_1},\quad (\lambda-\lambda_2)^{p_2},\quad \ldots,\quad (\lambda-\lambda_u)^{p_u}, \tag{47}$$

определяется формулой

$$d=\sum_{i=1}^{u}\min(k_i, p_i); \tag{48}$$

здесь k_i — кратность λ_i как корня $f(\lambda)$ $(i=1, 2, \ldots, u)$.

В качестве приложения доказанной теоремы определим все элементарные делители произвольной матрицы $A=\|a_{ik}\|_1^n$, соответствующие характеристическому числу λ_0:

$$\underbrace{\lambda-\lambda_0,\ \ldots,\ \lambda-\lambda_0}_{g_1};\quad \underbrace{(\lambda-\lambda_0)^2,\ \ldots,\ (\lambda-\lambda_0)^2}_{g_2};\quad \ldots;\quad \underbrace{(\lambda-\lambda_0)^m,\ \ldots,\ (\lambda-\lambda_0)^m}_{g_m},$$

где $g_i\geqslant 0$ $(i=1, 2, \ldots, m-1)$, $g_m>0$, в случае, когда даны дефекты

$$d_1,\ d_2,\ \ldots,\ d_m$$

матриц

$$A-\lambda_0 E,\ (A-\lambda_0 E)^2,\ \ldots,\ (A-\lambda_0 E)^m.$$

Для этого заметим, что $(A-\lambda_0 E)^j=f_j(A)$, где $f_j(\lambda)=(\lambda-\lambda_0)^j$ $(j=1, 2, \ldots, m)$. Поэтому для определения дефекта матрицы $(A-\lambda_0 E)^j$ следует в формуле (48) положить $k_i=j$ для элементарных делителей, соответствующих характеристическому числу λ_0, и $k_i=0$ во всех других слагаемых $(j=1, 2, \ldots, m)$. Таким образом, получим формулы

$$\left.\begin{aligned} &g_1+g_2+g_3+\ldots+g_m=d_1,\\ &g_1+2g_2+2g_3+\ldots+2g_m=d_2,\\ &g_1+2g_2+3g_3+\ldots+3g_m=d_3,\\ &\qquad\ldots\ldots\ldots\ldots\ldots\\ &g_1+2g_2+3g_3+\ldots+mg_m=d_m. \end{aligned}\right\} \tag{49}$$

Отсюда [2])

$$g_j=2d_j-d_{j-1}-d_{j+1} \qquad (j=1, 2, \ldots, m;\ d_0=0,\ d_{m+1}=d_m). \tag{50}$$

3. Вернемся к основному вопросу об определении элементарных делителей матрицы $f(A)$. Как уже отмечалось выше, элементарные делители $f(A)$ совпадают с элементарными делителями $f(J)$, а элементарные делители квазидиагональной матрицы составляются из элементарных делите-

---

[1]) В общем случае, когда $f(\lambda)$ — не многочлен, под кратностью корня λ_i функции $f(\lambda)$ понимают целое число k_i, определяемое из условий (46). k_i может и равняться нулю; в этом случае $f(\lambda_i)\neq 0$.

[2]) Если задан ряд чисел d_1, d_2, d_3 ..., где d_j — дефект степени $(A-\lambda_0 E)^j$ $(j=1, 2, 3\ \ldots)$, то число m — наибольший из показателей степеней элементарных делителей вида $(\lambda-\lambda_0)^v$ — определяется как индекс, при котором $d_{m-1}<d_m=d_{m+1}$.

лей диагональных клеток (см. теорему 5). Поэтому вопрос сводится к разысканию элементарных делителей матрицы C, имеющей правильную треугольную форму:

$$C=\sum_{k=0}^{p-1}a_kH^k=\begin{Vmatrix} a_0 & a_1 & \cdots & a_{p-1}\\ 0 & a_0 & \ddots & \vdots\\ \vdots & \ddots & \ddots & a_1\\ 0 & 0 & \cdots & a_0 \end{Vmatrix}. \tag{51}$$

Рассмотрим отдельно два случая:

1° $a_1\neq 0$. Характеристический многочлен матрицы C, очевидно, равен

$$D_p(\lambda)=(\lambda-a_0)^p.$$

Тогда, поскольку $D_p(\lambda)$ делится на $D_{p-1}(\lambda)$ без остатка, то

$$D_{p-1}(\lambda)=(\lambda-a_0)^g \qquad (g\leqslant p).$$

Здесь через $D_{p-1}(\lambda)$ мы обозначили наибольший общий делитель миноров $(p-1)$-го порядка характеристической матрицы

$$\lambda E-C=\begin{Vmatrix} \lambda-a_0 & -a_1 & \cdots & -a_{p-1}\\ 0 & \lambda-a_0 & \ddots & \vdots\\ \vdots & \ddots & \ddots & -a_1\\ 0 & 0 & \cdots & \lambda-a_0 \end{Vmatrix}.$$

Легко видеть, что минор нулевого элемента, отмеченного значком «+», равен $(-a_1)^{p-1}$ и, следовательно, в нашем случае отличен от нуля. Поэтому здесь $g=0$. Но тогда из

$$D_p(\lambda)=(\lambda-a_0)^p, \quad D_{p-1}(\lambda)=1$$

следует, что матрица C имеет только один элементарный делитель $(\lambda-a_0)^p$.

2° $a_1=\ldots=a_{k-1}=0$, $a_k\neq 0$. В этом случае

$$C=a_0E+a_kH^k+\ldots+a_{p-1}H^{p-1} \quad (H=H^{(p)}).$$

Поэтому при любом целом положительном j дефект матрицы

$$(C-a_0E)^j=a_k^jH^{kj}+\ldots$$

определится равенством

$$d_j=\begin{cases} kj & \text{при } kj\leqslant p,\\ p & \text{при } kj>p. \end{cases}$$

Положим

$$p=qk+h \qquad (0\leqslant h<k). \tag{52}$$

Тогда[1])

$$d_1=k,\ d_2=2k,\ \ldots,\ d_q=qk,\ d_{q+1}=p. \tag{53}$$

[1]) В данном случае число $q+1$ играет роль числа m в формулах (49) и (50) (см. сноску [2]) к стр. 156).

Поэтому согласно формулам (50) имеем:

$$g_1 = \ldots = g_{q-1} = 0, \quad g_q = k-h, \quad g_{q+1} = h.$$

Таким образом, матрица C имеет элементарные делители

$$\underbrace{(\lambda-a_0)^{q+1}, \ \ldots, \ (\lambda-a_0)^{q+1}}_{h}, \quad \underbrace{(\lambda-a_0)^{q}, \ \ldots, \ (\lambda-a_0)^{q}}_{k-h} \tag{54}$$

де целые числа $q>0$ и $h\geqslant 0$ определяются из (52).

4. Теперь мы уже имеем возможность выяснить, какие элементарные делители имеет матрица $f(J)$ [см. формулы (45) и (46)]. Каждому элементарному делителю матрицы A

$$(\lambda-\lambda_0)^p$$

отвечает в матрице $f(J)$ диагональная клетка

$$f(\lambda_0 E+H)=\sum_{i=0}^{p-1}\frac{f^{(i)}(\lambda_0)}{i!}H^i=\begin{Vmatrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdots & \frac{f^{(p-1)}(\lambda_0)}{(p-1)!} \\ 0 & f(\lambda_0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{f'(\lambda_0)}{1!} \\ 0 & 0 & & f(\lambda_0) \end{Vmatrix} \tag{55}$$

Очевидно, вопрос сводится к нахождению элементарных делителей клеток вида (55). Но матрица (55) имеет правильную треугольную форму (51), причем здесь

$$a_0 = f(\lambda_0), \ a_1 = f'(\lambda_0), \ a_2 = \frac{f''(\lambda_0)}{2!}, \ \ldots$$

Таким образом, приходим к теореме:

Теорема 9. *Элементарные делители матрицы $f(A)$ получаются из элементарных делителей матрицы A следующим образом: элементарному делителю*

$$(\lambda-\lambda_0)^p \tag{56}$$

матрицы A при $p=1$ или при $p>1$ и $f'(\lambda_0)\neq 0$ отвечает один элементарный делитель

$$(\lambda-f(\lambda_0))^p \tag{57}$$

матрицы $f(A)$; при $p>1$ $f'(\lambda_0)=\ldots=f^{(k-1)}(\lambda_0)$, $f^{(k)}(\lambda_0)\neq 0$ $(k<p)$ элементарному делителю (56) матрицы A соответствуют следующие элементарные делители матрицы $f(A)$:

$$\underbrace{(\lambda-f(\lambda_0))^{q+1}, \ \ldots, \ (\lambda-f(\lambda_0))^{q+1}}_{h}, \quad \underbrace{(\lambda-f(\lambda_0))^{q}, \ \ldots, \ (\lambda-f(\lambda_0))^{q}}_{k-h} \tag{58}$$

где

$$p=qk+h, \quad 0\leqslant q, \quad 0\leqslant h<k;$$

наконец, при $p>1$, $f'(\lambda_0)=f''(\lambda_0)=\ldots=f^{(p-1)}(\lambda_0)=0$ элементарному делителю (56) соответствуют p элементарных делителей первой степени

матрицы $f(A)$ [1]:

$$\lambda - f(\lambda_0), \ \ldots, \ \lambda - f(\lambda_0). \tag{59}$$

Отметим следующие частные положения, содержащиеся в этой теореме.

1. *Если* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *— характеристические числа матрицы* A, *то* $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ *суть характеристические числа матрицы* $f(A)$ (как в первом, так и во втором рядах чисел каждое характеристическое число повторяется столько раз, какова его кратность как корня характеристического уравнения) [2].

2. *Если производная* $f'(\lambda)$ *не равна нулю на спектре матрицы* A [3], *то при переходе от матрицы* A *к матрице* $f(A)$ *элементарные делители не «расщепляются», т. е. если матрица* A *имеет элементарные делители*

$$(\lambda - \lambda_1)^{p_1}, \ (\lambda - \lambda_2)^{p_2}, \ \ldots, \ (\lambda - \lambda_n)^{p_n},$$

то матрица $f(A)$ *имеет элементарные делители*

$$(\lambda - f(\lambda_1))^{p_1}, \ (\lambda - f(\lambda_2))^{p_2}, \ \ldots, \ (\lambda - f(\lambda_n))^{p_n}.$$

§ 8. Общий метод построения преобразующей матрицы

Во многих вопросах теории матриц и их приложений достаточно знать только нормальную форму, к которой приводится данная матрица $A = \| a_{ik} \|_1^n$ преобразованием подобия. Нормальная форма вполне определяется инвариантными многочленами характеристической матрицы $\lambda E - A$. Для нахождения последних можно воспользоваться определяющими формулами [см. (10) на стр. 143] или приведением характеристической матрицы $\lambda E - A$ при помощи элементарных преобразований к канонической диагональной форме.

В некоторых же вопросах необходимо знать не только нормальную форму $\widetilde{A}$ данной матрицы A, но и преобразующую неособенную матрицу T.

Непосредственный способ определения матрицы T состоит в следующем.

Равенство

$$A = T\widetilde{A}T^{-1}$$

переписывается так:

$$AT - T\widetilde{A} = 0.$$

Это матричное уравнение относительно T равносильно системе n^2 линейных однородных уравнений относительно n^2 неизвестных коэффициентов матрицы T. Определение преобразующей матрицы сводится к решению этой системы из n^2 уравнений. При этом из множества решений необходимо выбрать такое решение, для которого $|T| \neq 0$. Существование такого решения обеспечено тем, что матрицы A и $\widetilde{A}$ имеют одни и те же инвариантные многочлены [4].

Заметим, что в то время как нормальная форма определяется однозначно заданием данной матрицы A [5], для преобразующей матрицы T мы всегда имеем бесчисленное множество значений, охватываемых формулой

$$T = UT_1, \tag{60}$$

---

[1]) (57) получается из (58), если положить $k = 1$; (59) получается из (58), если положить $k = p$ либо $k > p$.

[2]) Предложение 1 было нами отдельно установлено в гл. V, стр. 106.

[3]) То есть $f'(\lambda_i) \neq 0$ для тех λ_i, которые являются кратными корнями минимального многочлена.

[4]) Из этого факта следует подобие матриц A и $\widetilde{A}$.

[5]) Это утверждение безоговорочно справедливо в отношении первой естественной нормальной формы. Что же касается второй нормальной формы или жордановой нормальной формы, то они определяются однозначно с точностью до порядка диагональных клеток.

где T_1 — одна из преобразующих матриц, а U — произвольная матрица, перестановочная с A [1]).

Предложенный выше способ определения преобразующей матрицы T очень прост по своей идее, но практически мало пригоден, так как связан с очень большими вычислениями (так, уже при $n=4$ он требует решения системы из 16 линейных уравнений).

Переходим к изложению более эффективного метода построения преобразующей матрицы T. Этот метод опирается на добавление к теореме 7 (стр. 150). Согласно этому добавлению в качестве преобразующей матрицы можно взять матрицу

$$T=Q(\widetilde{A}), \tag{61}$$

коль скоро

$$\lambda E-\widetilde{A}=P(\lambda)(\lambda E-A)Q(\lambda).$$

Последнее равенство выражает собой эквивалентность характеристических матриц $\lambda E-A$ и $\lambda E-\widetilde{A}$. Здесь $P(\lambda)$ и $Q(\lambda)$ — многочленные матрицы с постоянными отличными от нуля определителями.

Для конкретного нахождения матрицы $Q(\lambda)$ мы приводим к канонической диагональной форме каждую из λ-матриц $\lambda E-A$ и $\lambda E-\widetilde{A}$ при помощи соответствующих элементарных преобразований [2]):

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\}=P_1(\lambda)(\lambda E-A)Q_1(\lambda), \tag{62}$$

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\}=P_2(\lambda)(\lambda E-\widetilde{A})Q_2(\lambda), \tag{63}$$

где

$$Q_1(\lambda)=T_1T_2\ldots T_{p_1},\quad Q_2(\lambda)=T_1^*T_2^*\ldots T_{p_2}^*, \tag{64}$$

а $T_1, \ldots, T_{p_1}, T_1^*, \ldots, T_{p_2}^*$ — элементарные матрицы, соответствующие элементарным операциям над столбцами λ-матриц $\lambda E-A$ и $\lambda E-\widetilde{A}$. Из (62), (63) и (64) следует:

$$\lambda E-\widetilde{A}=P(\lambda)(\lambda E-A)Q(\lambda),$$

где

$$Q(\lambda)=Q_1(\lambda)Q_2^{-1}(\lambda)=T_1T_2\ldots T_{p_1}T_{p_2}^{*-1}T_{p_2-1}^{*-1}\ldots T_1^{*-1}. \tag{65}$$

Матрицу $Q(\lambda)$ вычисляем, применив последовательно к столбцам единичной матрицы E элементарные операции с матрицами $T_1, \ldots, T_{p_1}, T_{p_2}^{*-1}, \ldots, T_1^{*-1}$. После этого [согласно формуле (61)] заменяем в $Q(\lambda)$ аргумент λ матрицей $\widetilde{A}$.

Пример.

$$\left\|\begin{matrix}1 & 0 & 1\\ 0 & 1 & -1\\ -1 & -1 & 1\end{matrix}\right\|.$$

Введем символические обозначения для левых и правых элементарных операций и соответствующих матриц (см. стр. 136—137):

$$S'=\{(c)\,i\},\quad S''=\{i+(b(\lambda))\,j\},\quad S'''=\{ij\},$$
$$T'=[(c)\,i],\quad T''=[i+(b(\lambda))\,j],\quad T'''=[ij].$$

Читатель легко проверит, что характеристическая матрица

$$\lambda E-A=\left\|\begin{matrix}\lambda-1 & 0 & -1\\ 0 & \lambda-1 & 1\\ 1 & 1 & \lambda-1\end{matrix}\right\|$$

---

[1]) Формулу (60) можно заменить формулой

$$T=T_1V,$$

где V — произвольная матрица, перестановочная с $\widetilde{A}$.

[2]) Здесь существенно лишь то, что обе λ-матрицы $\lambda E-A$ и $\lambda E-\widetilde{A}$ приводятся к одному и тому же виду. Мы избрали канонический диагональный вид, поскольку существует алгоритм, обеспечивающий приведение к такому виду.

приводится к каноническому диагональному виду

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)^3 \end{Vmatrix}$$

с помощью следующих последовательно выполненных элементарных операций:

$$\begin{gathered}[1+(\lambda-1)3],\quad \{2+1\},\quad \{3+(\lambda-1)1\},\quad \{(-1)1\},\quad [1-2],\\ [1-(\lambda^2-2\lambda+1)2],\quad \{2-(\lambda-1)1\},\quad \{(-1)2\},\quad [13],\quad \{23\}.\end{gathered} \tag{*}$$

Из канонического диагонального вида матрицы $\lambda E-A$ усматриваем, что матрица A имеет только один элементарный делитель $(\lambda-1)^3$. Поэтому соответствующей жордановой формой будет матрица

$$J=\begin{Vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{Vmatrix}.$$

Нетрудно видеть, что характеристическая матрица $\lambda E-J$ приводится к тому же каноническому диагональному виду с помощью элементарных операций

$$\begin{gathered}\{3+(\lambda-1)2\},\quad \{3+(\lambda^2-2\lambda+1)^3 1\},\quad [2+(\lambda-1)3],\quad [1+(\lambda-1)2],\\ \{(-1)1\}\ \{(-1)2\},\quad [13],\quad \{12\}.\end{gathered} \tag{**}$$

Выбрасывая из (*) и (**) левые элементарные операции, обозначенные символом $\{\ldots\}$, мы в соответствии с формулами (64), (65) получим

$$\begin{aligned}Q(\lambda)&=Q_1(\lambda)\,Q_2^{-1}(\lambda)=\\ &=[1+(\lambda-1)3]\,[1-2]\,[1-(\lambda^2-2\lambda+1)2]\,[13]\,[13]\,[1-(\lambda-1)2]\,[2-(\lambda-1)3]=\\ &=[1+(\lambda-1)3]\,[1-(\lambda^2-\lambda+1)2]\,[2-(\lambda-1)3].\end{aligned}$$

Применим к единичной матрице последовательно эти правые элементарные операции:

$$E=\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}\to\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda-1 & 0 & 1 \end{Vmatrix}\to\begin{Vmatrix} 1 & 0 & 0 \\ -\lambda^2+\lambda-1 & 1 & 0 \\ \lambda-1 & 0 & 1 \end{Vmatrix}\to$$

$$\to\begin{Vmatrix} 1 & 0 & 0 \\ -\lambda^2+\lambda-1 & 1 & 0 \\ \lambda-1 & -\lambda+1 & 1 \end{Vmatrix}=Q(\lambda).$$

Таким образом,

$$Q(\lambda)=\begin{Vmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}\lambda^2+\begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{Vmatrix}\lambda+\begin{Vmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{Vmatrix}.$$

Замечая, что

$$J^2=\begin{Vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{Vmatrix},$$

находим

$$T=Q(J)=\begin{Vmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}\begin{Vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{Vmatrix}+\begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{Vmatrix}\begin{Vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{Vmatrix}+$$

$$+\begin{Vmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{Vmatrix}=\begin{Vmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & +1 & 0 \end{Vmatrix}.$$

Проверка.

$$AT=\begin{Vmatrix} 1 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{Vmatrix},\quad TJ=\begin{Vmatrix} 1 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{Vmatrix},$$

$$|T|=\begin{vmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix}=1.$$

Следовательно, $AT=TJ$ ($|T|\neq 0$), т. е. $A=TJT^{-1}$.

§ 9. Второй метод построения преобразующей матрицы

1. Мы изложим еще один метод построения преобразующей матрицы, который часто приводит к меньшим вычислениям, нежели метод предыдущего параграфа. Однако этот второй метод применим лишь тогда, когда нормальная форма жорданова и известны элементарные делители

$$(\lambda-\lambda_1)^{p_1},\quad (\lambda-\lambda_2)^{p_2},\ldots \tag{66}$$

данной матрицы A.

Пусть $A=TJT^{-1}$, где

$$J=\{\lambda_1E^{(p_1)}+H^{(p_1)},\quad \lambda_2E^{(p_2)}+H^{(p_2)},\ldots\}=$$

$$=\begin{Vmatrix} \overbrace{\lambda_1 & 1 & \ldots & 0}^{p_1} & & & \\ \cdot & \cdot & & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ & & & 1 & & & \\ 0 & \ldots & & \lambda_1 & & & \\ & & & & \overbrace{\lambda_2 & 1 & \ldots & 0}^{p_2} \\ & & & & \cdot & \cdot & & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & & & 1 \\ & & & & 0 & \ldots & & \lambda_2 \\ & & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & \cdot \end{Vmatrix}.$$

Тогда, обозначая через t_k k-й столбец матрицы T ($k=1, 2, \ldots, n$), мы матричное равенство

$$AT=TJ$$

заменим эквивалентной системой равенств

$$At_1=\lambda_1 t_1,\ At_2=\lambda_1 t_2+t_1,\ \ldots,\ At_{p_1}=\lambda_1 t_{p_1}+t_{p_1-1}, \tag{67}$$

$$At_{p_1+1}=\lambda_2 t_{p_1+1},\quad At_{p_1+2}=\lambda_2 t_{p_1+2}+t_{p_1+1},\ \ldots,\quad At_{p_1+p_2}=\lambda_2 t_{p_1+p_2}+t_{p_1+p_2-1}, \tag{68}$$

. . . . . .

которую перепишем еще так:

$$(A-\lambda_1 E)\,t_1=0,\quad (A-\lambda_1 E)\,t_2=t_1,\ \ldots,\ (A-\lambda_1 E)\,t_{p_1}=t_{p_1-1}, \tag{67'}$$

$$(A-\lambda_2 E)\,t_{p_1+1}=0,\quad (A-\lambda_2 E)\,t_{p_1+2}=t_{p_1+1},\ \ldots,\ (A-\lambda_2 E)\,t_{p_1+p_2}=t_{p_1+p_2-1}, \tag{68'}$$

. . . . . .

Таким образом, все столбцы матрицы T разбиваются на «жордановы цепочки» столбцов: $[t_1, t_2, \ldots, t_{p_1}]$, $[t_{p_1+1}, t_{p_1+2}, \ldots, t_{p_1+p_2}]$, ... Каждой жордановой клетке в J [или, что то же, каждому элементарному делителю (66)] соответствует своя жорданова цепочка столбцов. Каждая жорданова цепочка столбцов характеризуется системой уравнений типа (67), (68) и т. п.

Нахождение преобразующей матрицы T сводится к разысканию жордановых цепочек, которые в совокупности давали бы n линейно независимых столбцов.

Мы покажем, что эти жордановы цепочки столбцов можно определить при помощи приведенной матрицы $C(\lambda)$ (см. гл. IV, § 5).

Для матрицы $C(\lambda)$ имеем тождество

$$(\lambda E-A)\,C(\lambda)=\psi(\lambda)\,E, \tag{69}$$

где $\psi(\lambda)$ — минимальный многочлен матрицы A.

Пусть

$$\psi(\lambda)=(\lambda-\lambda_0)^m\chi(\lambda)\qquad(\chi(\lambda_0)\neq 0).$$

Продифференцируем почленно последовательно $m-1$ раз тождество (69):

$$\left.\begin{aligned}&(\lambda E-A)\,C'(\lambda)+C(\lambda)=\psi'(\lambda)\,E,\\&(\lambda E-A)\,C''(\lambda)+2C'(\lambda)=\psi''(\lambda)\,E,\\&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\&(\lambda E-A)\,C^{(m-1)}(\lambda)+(m-1)\,C^{(m-2)}(\lambda)=\psi^{(m-1)}(\lambda)\,E.\end{aligned}\right\} \tag{70}$$

Подставляя λ_0 вместо λ в (69), (70) и замечая, что правые части при этом обращаются в нуль, получим:

$$(A-\lambda_0 E)\,C=0,\quad (A-\lambda_0 E)\,D=C,\quad (A-\lambda_0 E)\,F=D,\ \ldots,\ (A-\lambda_0 E)\,K=G, \tag{71}$$

где

$$\left.\begin{aligned}&C=C(\lambda_0),\quad D=\frac{1}{1!}C'(\lambda_0),\quad F=\frac{1}{2!}C''(\lambda_0),\ \ldots,\ G=\frac{1}{(m-2)!}C^{(m-2)}(\lambda_0),\\&K=\frac{1}{(m-1)!}C^{(m-1)}(\lambda_0).\end{aligned}\right\} \tag{72}$$

Заменим в равенствах (71) матрицы (72) их k-ми столбцами ($k=1, 2, \ldots, n$). Получим:

$$(A-\lambda_0 E)\,C_k=0,\quad (A-\lambda_0 E)\,D_k=C_k,\ \ldots,\ (A-\lambda_0 E)\,K_k=G_k \tag{73}$$

$$(k=1,\ 2,\ \ldots,\ n).$$

Поскольку $C=C(\lambda_0)\neq 0$ [1]), можно выбрать такое $k\ (\leqslant n)$, что

$$C_k\neq 0. \tag{74}$$

Тогда m столбцов

$$C_k,\ D_k,\ F_k,\ \ldots,\ G_k,\ K_k \tag{75}$$

линейно независимы. В самом деле, пусть

$$\gamma C_k+\delta D_k+\ldots+\varkappa K_k=0. \tag{76}$$

[1]) Из $C(\lambda_0)=0$ следовало бы, что все элементы $C(\lambda)$ имеют общий делитель степени выше нулевой, что противоречит определению $C(\lambda)$.

Помножая обе части (76) последовательно на $A-\lambda_0 E, \ldots, (A-\lambda_0 E)^{m-1}$, получим:

$$\delta C_k+\ldots+\varkappa G_k=0, \ \ldots, \ \varkappa C_k=0. \tag{77}$$

Из (76) и (77) в силу (74) находим:

$$\gamma=\delta=\ldots=\varkappa=0.$$

Поскольку линейно независимые столбцы (75) удовлетворяют системе уравнений (73), они образуют жорданову цепочку векторов, отвечающую элементарному делителю $(\lambda-\lambda_0)^m$ [ср. (73) с (67')].

Если при некотором k $C_k=0$, но $D_k\neq 0$, то столбцы $D_k, \ldots, G_k, K_k$ образуют жорданову цепочку из $m-1$ векторов и т. д.

2. Покажем сначала, как построить преобразующую матрицу T в том случае, когда матрица A имеет попарно взаимно простые элементарные делители:

$$(\lambda-\lambda_1)^{m_1}, \ (\lambda-\lambda_2)^{m_2}, \ \ldots, \ (\lambda-\lambda_s)^{m_s} \qquad (\lambda_i\neq\lambda_j \text{ при } i\neq j; \ \ i,j=1, 2, \ldots, s).$$

Элементарному делителю $(\lambda-\lambda_j)^{m_j}$ ставим в соответствие жорданову цепочку столбцов $C^{(j)}$, $D^{(j)}$, $\ldots$, $G^{(j)}$, $K^{(j)}$, построенную по указанному выше способу. Тогда

$$(A-\lambda_j E)\,C^{(j)}=0, \quad (A-\lambda_j E)\,D^{(j)}=C^{(j)}, \ \ldots, \ (A-\lambda_j E)\,K^{(j)}=G^{(j)}. \tag{78}$$

Давая j значения $1, 2, \ldots, s$, получим s жордановых цепочек, содержащих в совокупности n столбцов. Эти столбцы линейно независимы.

Действительно, пусть

$$\sum_{j=1}^{s}[\gamma_j C^{(j)}+\delta_j D^{(j)}+\ldots+\varkappa_j K^{(j)}]=0. \tag{79}$$

Помножим обе части равенства (79) слева на произведение

$$(A-\lambda_1 E)^{m_1}\ldots(A-\lambda_{j-1}E)^{m_{j-1}}(A-\lambda_j E)^{m_j-1}\ (A-\lambda_{j+1}E)^{m_{j+1}}\ldots(A-\lambda_s E)^{m_s}. \tag{80}$$

Получим:

$$\varkappa_j=0.$$

Заменяя в (80) m_j-1 последовательно на m_j-2, m_j-3, $\ldots$, найдем:

$$\gamma_j=\delta_j=\ldots=\varkappa_j=0 \qquad (j=1, 2, \ldots, s),$$

что и требовалось доказать.

Матрицу T определим формулой

$$T=(C^{(1)}, D^{(1)}, \ldots, K^{(1)}; \ C^{(2)}, D^{(2)}, \ldots, K^{(2)}; \ \ldots; \ C^{(s)}, D^{(s)}, \ldots, K^{(s)}). \tag{81}$$

Пример.

$$A=\begin{Vmatrix} 8 & 3 & -10 & -3 \\ 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{Vmatrix};$$

$\psi(\lambda)=\Delta(\lambda)=(\lambda-1)^2(\lambda+1)^2=\lambda^4-2\lambda^2+1$,

элементарные делители: $(\lambda-1)^2$, $(\lambda+1)^2$,

$$\Psi(\lambda, \mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}=\mu^3+\lambda\mu^2+(\lambda^2-2)\,\mu+\lambda^3-2\lambda,$$

$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ 3 & 2 & 2 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 4 & 0 & 2 \end{array}$$

$$C(\lambda)=\Psi(\lambda E, A)=A^3+\lambda A^2+(\lambda^2-2)\,A+(\lambda^3-2\lambda)\,E.$$

Составим первый столбец $C_1(\lambda)$:

$$C_1(\lambda)=[A^3]_1+\lambda[A^2]_1+(\lambda^2-2)\,A_1+(\lambda^3-2\lambda)\,E_1.$$

Для вычисления первого столбца матрицы A^2 мы помножим все строки матрицы A на первый столбец матрицы A. Получим[1]: $[A^2]_1=(1, 4, 0, 2)$. Помножая на этот столбец все строки матрицы A, найдем $[A^3]_1=(3, 6, 2, 3)$. Поэтому

$$C_1(\lambda)=\begin{Vmatrix}3\\6\\2\\3\end{Vmatrix}+\lambda\begin{Vmatrix}1\\4\\0\\2\end{Vmatrix}+(\lambda^2-2)\begin{Vmatrix}3\\2\\2\\1\end{Vmatrix}+(\lambda^3-2\lambda)\begin{Vmatrix}1\\0\\0\\0\end{Vmatrix}=\begin{Vmatrix}\lambda^3+3\lambda^2-\lambda-3\\2\lambda^2+4\lambda+2\\2\lambda^2-2\\\lambda^2+2\lambda+1\end{Vmatrix}.$$

Отсюда $C_1(1)=(0, 8, 0, 4)$ и $C_1'(1)=(8, 8, 4, 4)$. Поскольку $C_1(-1)=(0, 0, 0, 0)$, то переходим ко вторым столбцам и, действуя аналогично предыдущему, находим: $C_2(-1)=(-4, 0, -4, 0)$ и $C_2'(-1)=(4, -4, 4, -4)$. Составляем матрицу:

$$(C_1(1), C_1'(1); C_2(-1), C_2'(-1))=\begin{Vmatrix}0&8&-4&4\\8&8&0&-4\\0&4&-4&4\\4&4&0&-4\end{Vmatrix}.$$

Сокращаем[2] первые два столбца на 4 и вторые два столбца на -4.

$$T=\begin{Vmatrix}0&2&1&-1\\2&2&0&1\\0&1&1&-1\\1&1&0&1\end{Vmatrix}.$$

Предлагаем читателю проверить, что

$$AT=T\cdot\begin{Vmatrix}1&1&0&0\\0&1&0&0\\0&0&-1&1\\0&0&0&-1\end{Vmatrix}.$$

3. Переходя к общему случаю, будем разыскивать жордановы цепочки векторов, отвечающие характеристическому числу λ_0, которому соответствуют p элементарных делителей $(\lambda-\lambda_0)^m$, q элементарных делителей $(\lambda-\lambda_0)^{m-1}$, r делителей $(\lambda-\lambda_0)^{m-2}$ и т. д.

Установим предварительно некоторые свойства матриц

$$C=C(\lambda_0),\quad D=C'(\lambda_0),\quad F=\frac{1}{2!}C''(\lambda_0),\ \ldots,\ K=\frac{1}{(m-1)!}C^{(m-1)}(\lambda_0). \tag{82}$$

1° *Матрицы* (82) *могут быть представлены в виде многочленов от* A:

$$C=h_1(A),\quad D=h_2(A),\ \ldots,\ K=h_m(A), \tag{83}$$

где

$$h_i(\lambda)=\frac{\psi(\lambda)}{(\lambda-\lambda_0)^i}\qquad (i=1, 2, \ldots, m). \tag{84}$$

В самом деле,

$$C(\lambda)=\Psi(\lambda E, A),$$

где

$$\Psi(\lambda, \mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}.$$

---

[1]) Столбец, на который умножаем строки, мы подписываем под строками матрицы A. Курсивом набраны элементы контрольной суммарной строки.

[2]) При умножении всех столбцов жордановой цепочки на число $c\neq 0$ цепочка остается жордановой.

Поэтому

$$\frac{1}{k!}C^{(k)}(\lambda_0)=\frac{1}{k!}\Psi^{(k)}(\lambda_0 E, A), \tag{85}$$

где

$$\frac{1}{k!}\Psi^{(k)}(\lambda_0, \mu)=\frac{1}{k!}\left[\frac{\partial^k}{\partial\lambda^k}\psi(\lambda, \mu)\right]_{\lambda=\lambda_0}=$$

$$=\frac{1}{k!}\left[\frac{\partial^k}{\partial\lambda^k}\frac{\psi(\mu)}{\mu-\lambda}\right]_{\lambda=\lambda_0}=\frac{\psi(\mu)}{(\mu-\lambda_0)^{k+1}}=h_{k+1}(\mu). \tag{86}$$

Из (82), (85) и (86) следует (83).

2° *Матрицы* (82) *имеют соответственно ранги*

$$p,\ 2p+q,\ 3p+2q+r,\ \ldots$$

Это свойство матриц (82) непосредственно получается из 1° и теоремы 8 гл. VI, если положить ранг равным $n-d$ и воспользоваться формулой (48) для дефекта функции от A (стр. 130).

3° *В ряду матриц* (82) *столбец каждой матрицы является линейной комбинацией столбцов любой последующей матрицы.*

Возьмем две матрицы $h_i(A)$ и $h_k(A)$ в ряду (82) (см. 1°). Пусть $i<k$. Тогда из (84) следует:

$$h_i(A)=h_k(A)(A-\lambda_0 E)^{k-i}.$$

Отсюда j-й столбец y_j $(j=1, 2, \ldots, n)$ матрицы $h_i(A)$ линейно выражается через столбцы $z_1, z_2, \ldots, z_n$ матрицы $h_k(A)$:

$$y_j=\sum_{g=1}^{n}\alpha_g z_g,$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n$ — элементы j-го столбца матрицы $(A-\lambda_0 E)^{k-i}$.

4° *Не меняя основных формул* (71), *можно в матрице* C *любой столбец заменить произвольной линейной комбинацией всех столбцов, сделав соответствующую замену в* $D, \ldots, K$.

Теперь перейдем к построению жордановых цепочек столбцов для элементарных делителей

$$\underbrace{(\lambda-\lambda_0)^m, \ldots, (\lambda-\lambda_0)^m}_{p};\quad \underbrace{(\lambda-\lambda_0)^{m-1}, \ldots, (\lambda-\lambda_0)^{m-1}}_{q};\ \ldots$$

Пользуясь свойствами 2° и 4°, мы матрицу C преобразуем к виду

$$C=(C_1, C_2, \ldots, C_p; 0, 0, \ldots, 0), \tag{87}$$

где столбцы $C_1, C_2, \ldots, C_p$ линейно независимы между собой. При этом

$$D=(D_1, D_2, \ldots, D_p; D_{p+1}, \ldots D_n).$$

Согласно 3° для любого i $(1\leqslant i\leqslant p)$ столбец C_i есть линейная комбинация столбцов $D_1, D_2, \ldots, D_n$:

$$C_i=\alpha_1 D_1+\ldots+\alpha_p D_p+\alpha_{p+1}D_{p+1}+\ldots+\alpha_n D_n. \tag{88}$$

Помножим обе части этого равенства на $A-\lambda_0 E$. Тогда, замечая, что [см. (73)]

$$(A-\lambda_0 E)C_i=0\quad (i=1, 2, \ldots, p),\qquad (A-\lambda_0 E)D_j=C_j\quad (j=1, 2, \ldots, n),$$

получим в силу (87):

$$0=\alpha_1 C_1+\alpha_2 C_2+\ldots+\alpha_p C_p,$$

откуда в (88)

$$\alpha_1=\ldots=\alpha_p=0.$$

Поэтому столбцы $C_1, C_2, \ldots, C_p$ представляют собой линейно независимые комбинации столбцов $D_{p+1}, \ldots, D_n$, и потому, согласно 4° и 2°, не меняя матрицы C, можно вместо $D_{p+1}, \ldots, D_{2p}$ взять столбцы $C_1, \ldots, C_p$, а вместо $D_{2p+q+1}, \ldots, D_n$ — нули.

Тогда матрица D примет вид

$$D=(D_1, \ldots, D_p; C_1, C_2, \ldots, C_p; D_{2p+1}, \ldots, D_{2p+q}; 0, 0, \ldots, 0). \tag{89}$$

Таким же образом, сохраняя вид (87) и (89) для матриц C и D, мы следующую матрицу F представим в виде

$$\left.\begin{aligned} F=(&F_1, \ldots, F_p; D_1, \ldots, D_p; F_{2p+1}, \ldots, F_{2p+q}; C_1, \ldots, C_p;\\ &D_{2p+1}, \ldots, D_{2p+q}, F_{3p+2q+1}, \ldots, F_{3p+2q+r}, 0, \ldots, 0)\end{aligned}\right\} \tag{90}$$

и т. д.

Формулы (73) дадут нам жордановы цепочки

$$\left.\begin{aligned} &\underbrace{\overbrace{(C_1, D_1, \ldots, K_1)}^{m}, \ldots, \overbrace{(C_p, D_p, \ldots, K_p)}^{m}}_{p};\\ &\underbrace{\overbrace{(D_{2p+1}, F_{2p+1}, \ldots, K_{2p+1})}^{m-1}, \ldots, \overbrace{(D_{2p+q}, F_{2p+q}, \ldots, K_{2p+q})}^{m-1}; \ldots}_{q}\end{aligned}\right\} \tag{91}$$

Эти жордановы цепочки независимы между собой. Действительно, все столбцы C_i в цепочках (91) независимы, так как они образуют p независимых столбцов матрицы C. Все столбцы C_i, D_j в (91) независимы, так как они образуют $2p+q$ независимых столбцов в матрице D и т. д.; наконец, *все столбцы в* (91) *независимы, так как они образуют* $n_0=mp+(m-1)q+\ldots$ *независимых столбцов в матрице* K. Число столбцов в (91) равно сумме степеней элементарных делителей, соответствующих данному характеристическому числу λ_0.

Пусть матрица $A=\|a_{ik}\|_1^n$ имеет s различных характеристических чисел λ_j $[j=1, 2, \ldots, s$; $\Delta(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\ldots(\lambda-\lambda_s)^{n_s}$; $\psi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots$ $\ldots(\lambda-\lambda_s)^{m_s}]$. Для каждого характеристического числа λ_j составим свою систему независимых жордановых цепочек (91); число столбцов в этой системе будет равно n_j $(j=1, 2, \ldots, s)$. Все полученные таким образом цепочки содержат $n=n_1+n_2+\ldots+n_s$ столбцов.

Эти n столбцов линейно независимы и составляют одну из искомых преобразующих матриц T.

Доказательство линейной независимости полученных n столбцов проводится следующим образом.

Любая линейная комбинация этих n столбцов может быть представлена в виде

$$\sum_{j=1}^{s} H_j=0, \tag{92}$$

где H_j—линейная комбинация столбцов в жордановых цепочках (91), соответствующих характеристическому числу λ_j $(j=1, 2, \ldots, s)$. Но любой столбец в жордановой цепочке, соответствующей характеристическому числу λ_j, удовлетворяет уравнению

$$(A-\lambda_j E)^{m_j} x=0.$$

Поэтому

$$(A-\lambda_j E)^{m_j} H_j=0. \tag{93}$$

Возьмем фиксированное число j $(1\leqslant j\leqslant s)$ и построим интерполяционный многочлен Лагранжа—Сильвестра $r(\lambda)$ (см. гл. V, §§ 1, 2) по следующим значениям на спектре матрицы:

$$r(\lambda_i)=r'(\lambda_i)=\ldots=r^{(m_i-1)}(\lambda_i)=0 \text{ при } i\neq j$$

и

$$r(\lambda_j)=1, \quad r'(\lambda_j)=\ldots=r^{(m_j-1)}(\lambda_j)=0.$$

Тогда при любом $i\neq j$ $r(\lambda)$ делится на $(\lambda-\lambda_i)^{m_i}$ без остатка; поэтому в силу (93)

$$r(A)H_i=0 \qquad (i\neq j). \tag{94}$$

Точно так же разность $r(\lambda)-1$ делится на $(\lambda-\lambda_j)^{m_j}$ без остатка; поэтому

$$r(A)\,H_j=H_j. \tag{95}$$

Помножая обе части (92) на $r(A)$, мы согласно (94) и (95) получим:

$$H_j=0.$$

Это справедливо для любого $j=1, 2, \ldots, s$. Но H_j есть линейная комбинация независимых столбцов, отвечающих одному и тому же характеристическому числу λ_j ($j=1, 2, \ldots, s$). Поэтому все коэффициенты в линейной комбинации H_j ($j=1, 2, \ldots, s$) и, следовательно, все коэффициенты в (92) равны нулю.

Замечание. Укажем на некоторые преобразования над столбцами матрицы T, при которых она остается преобразующей к той же жордановой форме (при том же расположении жордановых диагональных клеток):

I. *Умножение всех столбцов какой-либо жордановой цепочки на произвольное число, отличное от нуля.*

II. *Прибавление к каждому (начиная со второго) столбцу жордановой цепочки предыдущего столбца той же цепочки, предварительно помноженного на одно и то же произвольное число.*

III. *Прибавление ко всем столбцам жордановой цепочки соответствующих столбцов другой цепочки, содержащей такое же или большее число столбцов и отвечающей тому же характеристическому числу.*

Пример 1.

$$A=\begin{Vmatrix} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 & -3 \\ 0 & 0 & -1 & 2 & -2 \\ 1 & -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 & 2 \end{Vmatrix};$$

$$\Delta(\lambda)=(\lambda-1)^4(\lambda+1),$$
$$\psi(\lambda)=(\lambda-1)^2(\lambda+1)=\lambda^3-\lambda^2-\lambda+1.$$

Элементарные делители матрицы A

$$(\lambda-1)^2,\ (\lambda-1)^2,\ \lambda+1.$$

$$J=\begin{Vmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{Vmatrix},$$

$$\Psi(\lambda,\mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}=\mu^2+(\lambda-1)\mu+\lambda^2-\lambda-1,$$

$$C(\lambda)=\Psi(\lambda E, A)=A^2+(\lambda-1)A+(\lambda^2-\lambda-1)E.$$

Вычисляем последовательно столбцы матрицы A^2 и соответствующие столбцы матриц $C(\lambda)$, $C(1)$, $C'(\lambda)$, $C'(1)$, $C(-1)$. Нам нужно получить два линейно независимых столбца матрицы $C(1)$ и один отличный от нуля столбец матрицы $C(-1)$.

$$C(\lambda)=\begin{Vmatrix} 1 & 0 & 0 & 2 & * \\ 0 & 1 & 0 & 2 & * \\ 0 & 0 & 1 & 0 & * \\ 2 & -2 & 2 & -1 & * \\ 2 & -2 & 2 & -2 & * \end{Vmatrix}+(\lambda-1)\begin{Vmatrix} 1 & 0 & 0 & 1 & * \\ 0 & 1 & -2 & 3 & * \\ 0 & 0 & -1 & 2 & * \\ 1 & -1 & 1 & 0 & * \\ 1 & -1 & 1 & -1 & * \end{Vmatrix}+(\lambda^2-\lambda-1)\begin{Vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{Vmatrix},$$

$$C(+1)=\begin{Vmatrix} 0 & 0 & 0 & 2 & * \\ 0 & 0 & 0 & 2 & * \\ 0 & 0 & 0 & 0 & * \\ 2 & -2 & 2 & -2 & * \\ 2 & -2 & 2 & -2 & * \end{Vmatrix},\quad C'(\lambda)=\begin{Vmatrix} 1 & 0 & 0 & 1 & * \\ 0 & 1 & -2 & 3 & * \\ 0 & 0 & -1 & 2 & * \\ 1 & -1 & 1 & 0 & * \\ 1 & -1 & 1 & -1 & * \end{Vmatrix}+(2\lambda-1)\begin{Vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{Vmatrix},$$

$$C'(+1)=\begin{Vmatrix} 2 & * & * & 1 & * \\ 0 & * & * & 3 & * \\ 0 & * & * & 2 & * \\ 1 & * & * & 1 & * \\ 1 & * & * & -1 & * \end{Vmatrix},\quad C(-1)=\begin{Vmatrix} 0 & 0 & 0 & * & * \\ 0 & 0 & 4 & * & * \\ 0 & 0 & 4 & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{Vmatrix}.$$

Поэтому [1])

$$T=(C_1(+1),\ C_1'(+1),\ C_4(+1),\ C_4'(+1),\ C_3(-1))=\begin{Vmatrix}0&2&2&1&0\\0&0&2&3&4\\0&0&0&2&4\\2&1&-2&1&0\\2&1&-2&-1&0\end{Vmatrix}.$$

Матрицу T можно несколько упростить. Последовательно

1) разделим пятый столбец на 4;
2) к третьему столбцу прибавим первый, к четвертому — второй;
3) из четвертого столбца вычтем третий;
4) разделим первый и второй столбцы на 2;
5) вычтем из второго столбца первый, предварительно помноженный на $1/2$.

Получим матрицу

$$T_1=\begin{Vmatrix}0&1&2&1&0\\0&0&2&1&1\\0&0&0&2&1\\1&0&0&2&0\\1&0&0&0&0\end{Vmatrix}.$$

Предлагаем читателю проверить, что $AT_1=T_1J$ и $|T_1|\neq 0$.

Пример 2.

$$A=\begin{Vmatrix}1&-1&1&-1\\-3&3&-5&4\\8&-4&3&-4\\15&-10&11&-11\end{Vmatrix};\qquad \Delta(\lambda)=(\lambda+1)^4,\quad \psi(\lambda)=(\lambda+1)^3.$$

Элементарные делители: $(\lambda+1)^3$, $\lambda+1$.

$$J=\begin{Vmatrix}-1&1&0&0\\0&-1&1&0\\0&0&-1&0\\0&0&0&-1\end{Vmatrix}.$$

Составляем многочлены

$$h_1(\lambda)=\frac{\psi(\lambda)}{\lambda+1}=(\lambda+1)^2,\quad h_2(\lambda)=\frac{\psi(\lambda)}{(\lambda+1)^2}=\lambda+1,\quad h_3(\lambda)=\frac{\psi(\lambda)}{(\lambda+1)^3}=1$$

и матрицы [2])

$$C=h_1(A)=(A+E)^2,\quad D=h_2(A)=A+E,\quad F=E:$$

$$C=\begin{Vmatrix}0&0&0&0\\2&-1&1&-1\\0&0&0&0\\-2&1&-1&1\end{Vmatrix},\quad D=\begin{Vmatrix}2&-1&1&-1\\-3&4&-5&4\\8&-4&4&-4\\15&-10&11&-10\end{Vmatrix},\quad F=\begin{Vmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{Vmatrix}.$$

В качестве первых трех столбцов матрицы T возьмем третьи столбцы этих матриц: $T=(C_3,\ D_3,\ F_3,\ *)$. В матрицах C, D, F из первого столбца вычтем удвоенный третий, а ко второму и четвертому столбцам прибавим третий. Получим:

$$\widetilde{C}=\begin{Vmatrix}0&0&0&0\\0&0&1&0\\0&0&0&0\\0&0&-1&0\end{Vmatrix},\quad \widetilde{D}=\begin{Vmatrix}0&0&1&0\\7&-1&-5&-1\\0&0&4&0\\-7&1&11&1\end{Vmatrix},\quad \widetilde{F}=\begin{Vmatrix}1&0&0&0\\0&1&0&0\\-2&1&1&1\\0&0&0&1\end{Vmatrix}.$$

---

[1]) Здесь нижний индекс означает номер столбца; например, $C_4'(+1)$ обозначает четвертый столбец матрицы $C'(+1)$.

[2]) Так как имеется только один элементарный делитель наивысшей степени, то ранг матрицы C должен быть равен единице. Поэтому достаточно, например, вычислить 7 элементов, стоящих в первом столбце и во второй строке матрицы C. Тогда сразу определятся остальные элементы матрицы C.

В матрицах $\widetilde{D}$, $\widetilde{F}$ к первому столбцу прибавим четвертый столбец, умноженный на 7, а из второго столбца вычтем четвертый. Получим:

$$\widetilde{C}=\begin{Vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{Vmatrix},\quad \widetilde{\widetilde{D}}=\begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -5 & -1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 11 & 1 \end{Vmatrix},\quad \widetilde{\widetilde{F}}=\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 5 & 0 & 1 & 1 \\ 7 & -1 & 0 & 1 \end{Vmatrix}.$$

В качестве последнего столбца в T берем первый столбец в $\widetilde{\widetilde{F}}$.

Имеем:

$$T=(C_3,\ D_3,\ F_3,\ \widetilde{\widetilde{F}}_1)=\begin{Vmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & 0 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 7 \end{Vmatrix}.$$

Для контроля можно проверить, что $AT=TJ$ и $|T|\neq 0$.

ГЛАВА VII

СТРУКТУРА ЛИНЕЙНОГО ОПЕРАТОРА В n-МЕРНОМ ПРОСТРАНСТВЕ (ГЕОМЕТРИЧЕСКАЯ ТЕОРИЯ ЭЛЕМЕНТАРНЫХ ДЕЛИТЕЛЕЙ)

Изложенная в предыдущей главе аналитическая теория элементарных делителей дала нам возможность для любой квадратной матрицы определить подобную ей матрицу, имеющую «нормальную» или «каноническую» форму. С другой стороны, в главе III мы видели, что поведение линейного оператора в n-мерном пространстве в различных базисах задается при помощи класса подобных матриц. Наличие в этом классе матрицы, имеющей нормальную форму, тесно связано с важными и глубокими свойствами линейного оператора в n-мерном пространстве. Изучению этих свойств посвящена настоящая глава. Исследование структуры линейного оператора приводит нас независимо от содержания предыдущей главы к теории преобразования матрицы к нормальной форме. Поэтому содержание настоящей главы может быть названо *геометрической теорией элементарных делителей* [1]).

§ 1. Минимальный многочлен вектора, пространства (относительно заданного линейного оператора)

Рассмотрим n-мерное векторное пространство $\boldsymbol{R}$ над полем K и линейный оператор $\boldsymbol{A}$ в этом пространстве.

Пусть $\boldsymbol{x}$ — произвольный вектор из $\boldsymbol{R}$. Составим ряд векторов

$$\boldsymbol{x},\ \boldsymbol{Ax},\ \boldsymbol{A^2x},\ \ldots \tag{1}$$

В силу конечномерности пространства найдется такое целое число $p\ (0 \leqslant p \leqslant n)$, что векторы $\boldsymbol{x}$, $\boldsymbol{Ax}$, ..., $\boldsymbol{A^{p-1}x}$ линейно независимы, а $\boldsymbol{A^px}$ есть линейная комбинация этих векторов с коэффициентами из поля K:

$$\boldsymbol{A^px} = -\gamma_1\boldsymbol{A^{p-1}x} - \gamma_2\boldsymbol{A^{p-2}x} - \ldots - \gamma_p\boldsymbol{x}. \tag{2}$$

Составим многочлен $\varphi(\lambda) = \lambda^p + \gamma_1\lambda^{p-1} + \ldots + \gamma_{p-1}\lambda + \gamma_p$. Тогда равенство (2) запишется так:

$$\varphi(\boldsymbol{A})\,\boldsymbol{x} = 0. \tag{3}$$

Всякий многочлен $\varphi(\lambda)$, для которого имеет место равенство (3), мы будем называть *аннулирующим многочленом для вектора* $\boldsymbol{x}$ [2]). Но нетрудно

---

[1]) В основу излагаемой здесь геометрической теории элементарных делителей положена статья автора [81а]. Другие геометрические построения теории элементарных делителей см. в [14], §§ 96—99, а также [24] и [42б].

[2]) Конечно, подразумевается: относительно данного оператора $\boldsymbol{A}$. Это обстоятельство мы для краткости в определении не оговариваем, поскольку на протяжении всей этой главы мы будем иметь дело с одним оператором $\boldsymbol{A}$.

видеть, что из всех аннулирующих многочленов для вектора x построенный нами многочлен является аннулирующим многочленом наименьшей степени со старшим коэффициентом 1. Такой многочлен мы будем называть *минимальным аннулирующим многочленом вектора* x или просто *минимальным многочленом вектора* x.

Заметим, что *произвольный аннулирующий многочлен* $\tilde{\varphi}(\lambda)$ *вектора* x *делится нацело на минимальный многочлен* $\varphi(\lambda)$.

В самом деле, пусть

$$\tilde{\varphi}(\lambda) = \varphi(\lambda)\varkappa(\lambda) + \varrho(\lambda), \tag{4}$$

где $\varkappa(\lambda)$, $\varrho(\lambda)$ — частное и остаток от деления $\tilde{\varphi}(\lambda)$ на $\varphi(\lambda)$. Тогда

$$\tilde{\varphi}(A)x = \varkappa(A)\varphi(A)x + \varrho(A)x = \varrho(A)x \tag{5}$$

и, следовательно,

$$\varrho(A)x = 0. \tag{6}$$

Но степень остатка $\varrho(\lambda)$ должна быть меньше степени минимального многочлена $\varphi(\lambda)$. Значит, $\varrho(\lambda) \equiv 0$.

Из доказанного предложения следует, в частности, что каждому вектору x отвечает *только один* минимальный многочлен.

Выберем в пространстве R некоторый базис $e_1, e_2, \ldots, e_n$. Обозначим через $\varphi_1(\lambda), \varphi_2(\lambda), \ldots, \varphi_n(\lambda)$ минимальные многочлены базисных векторов $e_1, e_2, \ldots, e_n$, а через $\psi(\lambda)$ — наименьшее общее кратное этих многочленов ($\psi(\lambda)$ берем со старшим коэффициентом 1). Тогда $\psi(\lambda)$ будет аннулирующим многочленом для всех базисных векторов $e_1, e_2, \ldots, e_n$. Так как произвольный вектор $x \in R$ представляется в виде $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, то

$$\psi(A)x = x_1\psi(A)e_1 + x_2\psi(A)e_2 + \ldots + x_n\psi(A)e_n = 0,$$

т. е.

$$\psi(A) = 0. \tag{7}$$

Многочлен $\psi(\lambda)$ является *аннулирующим многочленом для всего пространства* R. Пусть $\tilde{\psi}(\lambda)$ — произвольный аннулирующий многочлен для всего пространства R. Тогда $\tilde{\psi}(\lambda)$ будет аннулирующим многочленом для базисных векторов $e_1, e_2, \ldots, e_n$. Следовательно, $\tilde{\psi}(\lambda)$ должен быть общим кратным для минимальных многочленов $\varphi_1(\lambda), \varphi_2(\lambda), \ldots, \varphi_n(\lambda)$ этих векторов, и потому многочлен $\tilde{\psi}(\lambda)$ должен делиться на наименьшее общее кратное $\psi(\lambda)$ без остатка. Отсюда следует, что из всех аннулирующих многочленов для всего пространства R построенный нами многочлен $\psi(\lambda)$ имеет наименьшую степень и старший коэффициент 1. Такой многочлен однозначно определяется заданием пространства R и оператора A и называется *минимальным многочленом пространства* R [1]. Единственность минимального многочлена пространства R следует из установленного выше положения: *произвольный аннулирующий многочлен* $\tilde{\psi}(\lambda)$ *пространства* R *делится нацело на минимальный многочлен* $\psi(\lambda)$. Хотя само построение минимального многочлена $\psi(\lambda)$ было связано

[1]) Если оператору A в некотором базисе $e_1, e_2, \ldots, e_n$ соответствует матрица $A = \|a_{ik}\|_1^n$, то аннулирующий или минимальный многочлен пространства R (относительно A) будет аннулирующим или соответственно минимальным многочленом матрицы A и наоборот. Ср. с гл. IV, § 5.

с определенным базисом $e_1, e_2, \ldots, e_n$, однако многочлен $\psi(\lambda)$ не зависит от выбора этого базиса (это вытекает из единственности минимального многочлена для пространства R).

Наконец, отметим еще, что минимальный многочлен пространства R является аннулирующим для любого вектора x из R, и потому *минимальный многочлен пространства делится без остатка на минимальный многочлен любого вектора из этого пространства.*

§ 2. Расщепление на инвариантные подпространства с взаимно простыми минимальными многочленами

Подпространство $R' \subset R$ называется *инвариантным* относительно данного оператора A, если $AR' \subset R'$, т. е. из $x \in R'$ следует $Ax \in R'$. Другими словами, оператор A переводит векторы инвариантного подпространства снова в векторы этого же подпространства.

В дальнейшем мы будем производить расщепление всего пространства (см. гл. III, § 1) на инвариантные относительно A подпространства. Такое расщепление сводит изучение поведения оператора во всем пространстве к изучению его поведения в отдельных составляющих подпространствах.

Докажем теперь следующую теорему:

Теорема 1 (1-я теорема о расщеплении пространства на инвариантные подпространства). *Если для данного линейного оператора A минимальный многочлен пространства $\psi(\lambda)$ представляется в поле K в виде произведения двух взаимно простых многочленов $\psi_1(\lambda)$ и $\psi_2(\lambda)$ (со старшими коэффициентами, равными единице),*

$$\psi(\lambda) = \psi_1(\lambda)\,\psi_2(\lambda), \tag{8}$$

то все пространство R расщепляется на два инвариантных подпространства I_1 и I_2,

$$R = I_1 + I_2, \tag{9}$$

для которых минимальными многочленами служат соответственно множители $\psi_1(\lambda)$ и $\psi_2(\lambda)$.

Доказательство. Обозначим через I_1 совокупность всех векторов x, удовлетворяющих уравнению $\psi_1(A)\,x = 0$. Аналогично определим I_2 с помощью уравнения $\psi_2(A)\,x = 0$. Определенные таким образом I_1 и I_2 суть подпространства в R.

Из взаимной простоты $\psi_1(\lambda)$ и $\psi_2(\lambda)$ вытекает[1]) существование таких многочленов $\chi_1(\lambda)$ и $\chi_2(\lambda)$ (с коэффициентами из K), что имеет место тождество

$$1 = \psi_1(\lambda)\,\chi_1(\lambda) + \psi_2(\lambda)\,\chi_2(\lambda). \tag{10}$$

Пусть теперь x — произвольный вектор из R. Заменим в (10) λ на A и применим обе части полученного операторного равенства к вектору x:

$$x = \psi_1(A)\,\chi_1(A)\,x + \psi_2(A)\,\chi_2(A)\,x, \tag{11}$$

т. е.

$$x = x' + x'', \tag{12}$$

где

$$x' = \psi_2(A)\,\chi_2(A)\,x, \qquad x'' = \psi_1(A)\,\chi_1(A)\,x. \tag{13}$$

---

[1]) См., например, [20], стр. 177.

Далее,

$$\psi_1(A)\,x' = \psi(A)\,\chi_2(A)\,x = 0, \qquad \psi_2(A)\,x'' = \psi(A)\,\chi_1(A)\,x = 0,$$

т. е. $x' \in I_1$ и $x'' \in I_2$.

I_1 и I_2 не имеют общих векторов, отличных от нуля. Действительно, если $x_0 \in I_1$ и $x_0 \in I_2$, т. е. $\psi_1(A)\,x_0 = 0$ и $\psi_2(A)\,x_0 = 0$, то в силу (11)

$$x_0 = \chi_1(A)\,\psi_1(A)\,x_0 + \chi_2(A)\,\psi_2(A)\,x_0 = 0.$$

Таким образом, доказано, что $R = I_1 + I_2$.

Пусть, далее, $x \in I_1$. Тогда $\psi_1(A)\,x = 0$. Помножая обе части этого равенства слева на A и переставляя местами A и $\psi_1(A)$, получим $\psi_1(A)\,Ax = 0$, т. е. $Ax \in I_1$. Этим доказано, что подпространство I_1 инвариантно относительно A. Аналогично доказывается инвариантность подпространства I_2.

Докажем теперь, что $\psi_1(\lambda)$ есть минимальный многочлен для I_1. Пусть $\tilde{\psi}_1(\lambda)$ — произвольный аннулирующий многочлен для I_1, а x — произвольный вектор из R. Используя уже установленное разложение (12), напишем:

$$\tilde{\psi}_1(A)\,\psi_2(A)\,x = \psi_2(A)\,\tilde{\psi}_1(A)\,x' + \tilde{\psi}_1(A)\,\psi_2(A)\,x'' = 0.$$

Поскольку x — произвольный вектор из R, то отсюда вытекает, что произведение $\tilde{\psi}_1(\lambda)\,\psi_2(\lambda)$ есть аннулирующий многочлен для R и потому делится без остатка на $\psi(\lambda) = \psi_1(\lambda)\,\psi_2(\lambda)$; другими словами, $\tilde{\psi}_1(\lambda)$ делится на $\psi_1(\lambda)$. Но $\tilde{\psi}_1(\lambda)$ — произвольный аннулирующий многочлен для I_1, а $\psi_1(\lambda)$ — один из аннулирующих многочленов (в силу определения I_1). Значит, $\psi_1(\lambda)$ есть минимальный многочлен для I_1. Совершенно аналогично доказывается, что $\psi_2(\lambda)$ есть минимальный многочлен для инвариантного подпространства I_2.

Теорема доказана полностью.

Разложим многочлен $\psi(\lambda)$ на неприводимые в поле K множители:

$$\psi(\lambda) = [\varphi_1(\lambda)]^{c_1}\,[\varphi_2(\lambda)]^{c_2} \ldots [\varphi_s(\lambda)]^{c_s} \tag{14}$$

(здесь $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, ..., $\varphi_s(\lambda)$ — различные неприводимые в K многочлены со старшими коэффициентами 1). Тогда на основании доказанной теоремы

$$R = I_1 + I_2 + \ldots + I_s, \tag{15}$$

где I_k — инвариантное подпространство с минимальным многочленом $[\varphi_k(\lambda)]^{c_k}$ $(k = 1, 2, \ldots, s)$.

Таким образом, доказанная теорема сводит изучение поведения линейного оператора в произвольном пространстве к изучению поведения этого оператора в пространстве, где минимальный многочлен есть степень неприводимого в K многочлена. Это обстоятельство будет нами использовано для доказательства следующего важного для нас предложения:

Теорема 2. *В пространстве всегда существует вектор, минимальный многочлен которого совпадает с минимальным многочленом всего пространства.*

Доказательство. Рассмотрим сначала тот частный случай, когда минимальный многочлен пространства R есть степень неприводимого в K многочлена $\varphi(\lambda)$:

$$\psi(\lambda) = [\varphi(\lambda)]^l.$$

Выберем в $\boldsymbol{R}$ базис $e_1, e_2, \ldots, e_n$. Минимальный многочлен вектора e_i является делителем многочлена $\psi(\lambda)$ и поэтому представляется в виде $[\varphi(\lambda)]^{l_i}$, где $l_i \leqslant l$ $(i=1, 2, \ldots, n)$.

Но минимальный многочлен пространства есть наименьшее общее кратное минимальных многочленов базисных векторов, т. е. $\psi(\lambda)$ совпадает с наибольшей из степеней $[\varphi(\lambda)]^{l_i}$ $(i=1, 2, \ldots, n)$. Другими словами, $\psi(\lambda)$ совпадает с минимальным многочленом одного из базисных векторов $e_1, e_2, \ldots, e_n$.

Переходя к общему случаю, докажем предварительно следующую лемму:

Лемма. *Если минимальные многочлены векторов e' и e'' взаимно просты, то минимальный многочлен суммы векторов $e'+e''$ равен произведению минимальных многочленов слагаемых векторов.*

Доказательство. В самом деле, пусть $\chi_1(\lambda)$ и $\chi_2(\lambda)$ — минимальные многочлены векторов e' и e''. По условию $\chi_1(\lambda)$ и $\chi_2(\lambda)$ взаимно просты. Пусть $\chi(\lambda)$ — произвольный аннулирующий многочлен для вектора $e=e'+e''$. Тогда

$$\chi_2(A)\chi(A)e'=\chi_2(A)\chi(A)e-\chi(A)\chi_2(A)e''=0,$$

т. е. $\chi_2(\lambda)\chi(\lambda)$ есть аннулирующий многочлен для e'. Следовательно, $\chi_2(\lambda)\chi(\lambda)$ делится без остатка на $\chi_1(\lambda)$, и так как $\chi_1(\lambda)$ и $\chi_2(\lambda)$ взаимно просты, то $\chi(\lambda)$ делится на $\chi_1(\lambda)$. Аналогично доказывается, что $\chi(\lambda)$ делится на $\chi_2(\lambda)$. Но $\chi_1(\lambda)$ и $\chi_2(\lambda)$ взаимно просты. Следовательно, $\chi(\lambda)$ делится на произведение $\chi_1(\lambda)\chi_2(\lambda)$. Итак, произвольный аннулирующий многочлен вектора e делится на аннулирующий многочлен $\chi_1(\lambda)\chi_2(\lambda)$. Поэтому $\chi_1(\lambda)\chi_2(\lambda)$ и будет минимальным многочленом вектора $e=e'+e''$.

Вернемся к теореме 2. Для доказательства в общем случае используем расщепление (15). Так как минимальные многочлены подпространств $\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots, \boldsymbol{I}_s$ суть степени неприводимого многочлена, то для этих подпространств наше предложение уже доказано. Поэтому существуют такие векторы $e' \in \boldsymbol{I}_1$, $e'' \in \boldsymbol{I}_2, \ldots, e^{(s)} \in \boldsymbol{I}_s$, минимальными многочленами которых будут соответственно $[\varphi_1(\lambda)]^{c_1}, [\varphi_2(\lambda)]^{c_2}, \ldots, [\varphi_s(\lambda)]^{c_s}$. В силу леммы минимальный многочлен вектора $e=e'+e''+\ldots+e^{(s)}$ равен произведению $[\varphi_1(\lambda)]^{c_1}[\varphi_2(\lambda)]^{c_2}\ldots\varphi_s(\lambda)]^{c_s}$, т. е. равен минимальному многочлену пространства $\boldsymbol{R}$.

§ 3. Сравнения. Надпространство

Пусть дано некоторое подпространство $\boldsymbol{I} \subset \boldsymbol{R}$. Мы будем говорить, что два вектора x, y из $\boldsymbol{R}$ сравнимы по $\bmod \boldsymbol{I}$, и будем писать $x \equiv y \pmod{\boldsymbol{I}}$ в том и только в том случае, если $y-x \in \boldsymbol{I}$. Легко проверяется, что введенное таким образом понятие сравнения обладает следующими свойствами: для любых x, y, $z \subset \boldsymbol{R}$

1. $x \equiv x \pmod{\boldsymbol{I}}$ (рефлективность сравнения).
2. Из $x \equiv y \pmod{\boldsymbol{I}}$ следует $y \equiv x \pmod{\boldsymbol{I}}$ (обратимость или симметричность сравнения).
3. Из $x \equiv y$, $y \equiv z \pmod{\boldsymbol{I}}$ следует $x \equiv z \pmod{\boldsymbol{I}}$ (транзитивность сравнения).

Наличие этих трех свойств сравнения дает нам возможность распределить все векторы пространства на классы, относя в каждый класс векторы, попарно сравнимые между собой по $\bmod \boldsymbol{I}$ (векторы из разных классов уже будут несравнимы по $\bmod \boldsymbol{I}$). Класс, содержащий вектор x,

будем обозначать через $\widehat{x}$[1]). Само подпространство $\boldsymbol{I}$ будет одним из этих классов, а именно, классом $\widehat{0}$. Обратим внимание, что каждому сравнению $x \equiv y \pmod{\boldsymbol{I}}$ отвечает равенство[2]) соответствующих классов: $\widehat{x} = \widehat{y}$.

Элементарно доказывается, что сравнения можно почленно складывать и почленно умножать на число из K:

1. Из $x \equiv x'$, $y \equiv y' \pmod{\boldsymbol{I}}$ следует $x+y \equiv x'+y' \pmod{\boldsymbol{I}}$.
2. Из $x \equiv x' \pmod{\boldsymbol{I}}$ следует $\alpha x \equiv \alpha x' \pmod{\boldsymbol{I}}$ $(\alpha \in K)$.

Эти свойства сравнения показывают, что операции сложения и умножения на число из K не «ломают» классов. Если возьмем два класса $\widehat{x}$ и $\widehat{y}$ и будем складывать элементы x, x', ... первого класса с любыми элементами y, y', ... второго класса, то все полученные таким образом суммы будут принадлежать одному и тому же классу, который мы назовем суммой классов $\widehat{x}$ и $\widehat{y}$ и обозначим через $\widehat{x}+\widehat{y}$. Аналогично, если все векторы x, x', ... класса $\widehat{x}$ умножим на число $\alpha \in K$, то полученные произведения будут принадлежать одному классу, который обозначим через $\alpha\widehat{x}$.

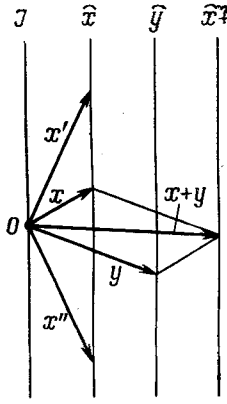


Рис. 4.

Таким образом, в многообразии $\widehat{\boldsymbol{R}}$ всех классов $\widehat{x}$, $\widehat{y}$, ... введены две операции: «сложение» и «умножение» на число из K. Эти операции, как легко проверить, обладают свойствами, сформулированными в определении векторного пространства (гл. III, § 1). Поэтому $\widehat{\boldsymbol{R}}$, как и $\boldsymbol{R}$, есть векторное пространство над полем K. Мы будем называть $\widehat{\boldsymbol{R}}$ надпространством по отношению к $\boldsymbol{R}$. Если n, m, $\widehat{n}$ — числа измерений соответственно пространств $\boldsymbol{R}$, $\boldsymbol{I}$, $\widehat{\boldsymbol{R}}$, то $\widehat{n} = n - m$.

Все введенные в этом параграфе понятия можно очень хорошо проиллюстрировать на следующем примере.

Пример. Пусть $\boldsymbol{R}$ — совокупность всех векторов в трехмерном пространстве, K — поле вещественных чисел. Для большей наглядности будем векторы изображать в виде направленных отрезков с началом в точке O. Пусть $\boldsymbol{I}$ — некоторая прямая, проходящая через O (точнее, совокупность векторов, идущих вдоль некоторой прямой, проходящей через O; рис. 4).

Сравнение $x \equiv x' \pmod{\boldsymbol{I}}$ означает, что векторы x и x' отличаются на вектор из $\boldsymbol{I}$, т. е. отрезок, соединяющий концы x и x', параллелен прямой $\boldsymbol{I}$. Поэтому класс $\widehat{x}$ изобразится прямой, проходящей через конец вектора x и параллельной $\boldsymbol{I}$, точнее, «щеткой» векторов, исходящих из O, концы которых лежат на этой прямой. «Щетки» можно складывать и умножать на вещественное число (складывая и умножая векторы, входящие в эти щетки). Эти «щетки» и являются элементами надпространства $\widehat{\boldsymbol{R}}$. В данном примере $n=3$, $m=1$, $\widehat{n}=2$.

Другой пример получим, если в качестве $\boldsymbol{I}$ возьмем плоскость, проходящую через точку O. В этом примере $n=3$, $m=2$, $\widehat{n}=1$.

Пусть теперь в $\boldsymbol{R}$ задан линейный оператор $\boldsymbol{A}$. Предположим, что $\boldsymbol{I}$ *есть инвариантное подпространство относительно* $\boldsymbol{A}$. Читатель легко

[1]) Так как каждый класс содержит бесчисленное множество векторов, то в силу этого условия он имеет и бесчисленное множество обозначений.

[2]) То есть совпадение.

докажет, что из $x \equiv x' \pmod{I}$ следует $Ax \equiv Ax' \pmod{I}$, т. е. что к обеим частям сравнения можно применять оператор A. Другими словами, если ко всем векторам x, x', ... некоторого класса $\widehat{x}$ применить оператор A, то полученные векторы Ax, Ax', ... также принадлежат к одному классу, который мы обозначим через $A\widehat{x}$. Линейный оператор A переводит класс в класс и, таким образом, является линейным оператором в $\widehat{R}$.

Мы будем говорить, что векторы $x_1, x_2, \ldots, x_p$ линейно зависимы по $\bmod I$, если существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_p$ в K, не равные одновременно нулю, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_q x_p \equiv 0 \pmod{I}. \tag{16}$$

Заметим, что не только понятие о линейной зависимости векторов, но все понятия, все предложения и рассуждения, приведенные в предыдущих параграфах этой главы, могут быть слово в слово повторены с одной лишь заменой всюду знака $=$ знаком $\equiv \pmod{I}$, где I — некоторое фиксированное подпространство, инвариантное относительно A.

Таким образом, вводятся понятия аннулирующий, минимальный многочлен вектора, пространства по $\bmod I$. Все эти понятия мы будем называть «относительными» в отличие от введенных ранее «абсолютных» понятий (имеющих место при знаке $=$).

Обратим внимание читателя на то, что *относительный минимальный многочлен (вектора, пространства) есть делитель абсолютного*. Пусть, например, $\sigma_1(\lambda)$ есть относительный минимальный многочлен вектора x, а $\sigma(\lambda)$ — соответствующий абсолютный минимальный многочлен.

Тогда

$$\sigma(A)\,x = 0;$$

но отсюда следует, что и

$$\sigma(A)\,x \equiv 0 \pmod{I}.$$

Поэтому $\sigma(\lambda)$ является относительным аннулирующим многочленом для вектора x и как таковой делится без остатка на относительный минимальный многочлен $\sigma_1(\lambda)$.

Наряду с «абсолютными» предложениями предыдущих параграфов мы имеем и «относительные» предложения. Так, например, имеем предложение: «В любом пространстве всегда существует вектор, относительный минимальный многочлен которого совпадает с относительным минимальным многочленом всего пространства».

Справедливость всех «относительных» предложений обусловлена тем, что, оперируя со сравнениями по $\bmod I$, мы по существу имеем дело с равенствами, только не в пространстве R, а в надпространстве $\widehat{R}$.

§ 4. Расщепление пространства на циклические инвариантные подпространства

Пусть $\sigma(\lambda) = \lambda^p + \alpha_1\lambda^{p-1} + \ldots + \alpha_{p-1}\lambda + \alpha_p$ — минимальный многочлен вектора e. Тогда векторы

$$e, Ae, \ldots, A^{p-1}e \tag{17}$$

линейно независимы, а

$$A^p e = -\alpha_p e - \alpha_{p-1} Ae - \ldots - \alpha_1 A^{p-1} e. \tag{18}$$

Векторы (17) образуют базис некоторого p-мерного подпространства $\boldsymbol{I}$. Это подпространство мы будем называть *циклическим*, имея в виду специальный характер базиса (17) и равенство (18) [1]). Оператор $\boldsymbol{A}$ переводит первый из векторов (17) во второй, второй — в третий и т. д. Последний же базисный вектор переводится оператором $\boldsymbol{A}$ в линейную комбинацию базисных векторов согласно равенству (18). Таким образом, оператор $\boldsymbol{A}$ переводит любой базисный вектор в вектор из $\boldsymbol{I}$; значит, он и произвольный вектор из $\boldsymbol{I}$ переводит в вектор из $\boldsymbol{I}$. Другими словами, *циклическое подпространство всегда инвариантно относительно* $\boldsymbol{A}$.

Произвольный вектор $\boldsymbol{x} \in \boldsymbol{I}$ представляется в виде линейной комбинации базисных векторов (17), т. е. в виде

$$\boldsymbol{x} = \chi(\boldsymbol{A})\,\boldsymbol{e}, \tag{19}$$

где $\chi(\lambda)$ — многочлен от λ с коэффициентами из K степени $\leqslant p-1$. Перебирая всевозможные многочлены $\chi(\lambda)$ степени $\leqslant p-1$ с коэффициентами из K, мы получим все векторы из $\boldsymbol{I}$, и при этом каждый вектор $\boldsymbol{x} \in \boldsymbol{I}$ только один раз, т. е. только при одном многочлене $\chi(\lambda)$. Имея в виду базис (17) либо формулу (19), мы будем говорить, что вектор $\boldsymbol{e}$ *порождает* подпространство $\boldsymbol{I}$.

Заметим еще, что *минимальный многочлен порождающего вектора* $\boldsymbol{e}$ *будет одновременно и минимальным многочленом всего подпространства* $\boldsymbol{I}$.

Сейчас нам предстоит установить основное предложение всей этой теории, согласно которому пространство $\boldsymbol{R}$ расщепляется на циклические подпространства.

Пусть $\psi_1(\lambda) = \psi(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_m$ есть минимальный многочлен пространства $\boldsymbol{R}$. Тогда в пространстве существует вектор $\boldsymbol{e}$, для которого этот многочлен является минимальным (теорема 2, стр. 174). Пусть $\boldsymbol{I}_1$ обозначает циклическое подпространство с базисом

$$\boldsymbol{e},\ \boldsymbol{Ae},\ \ldots,\ \boldsymbol{A}^{m-1}\boldsymbol{e}. \tag{20}$$

Если $n = m$, то $\boldsymbol{R} = \boldsymbol{I}_1$. Пусть $n > m$ и пусть многочлен

$$\psi_2(\lambda) = \lambda^p + \beta_1\lambda^{p-1} + \ldots + \beta_p$$

будет минимальным многочленом $\boldsymbol{R}$ по $\operatorname{mod} \boldsymbol{I}_1$. Согласно замечанию, сделанному в конце § 3, $\psi_2(\lambda)$ будет делителем $\psi_1(\lambda)$, т. е. существует такой многочлен $\varkappa(\lambda)$, что

$$\psi_1(\lambda) = \psi_2(\lambda)\,\varkappa(\lambda). \tag{21}$$

Далее, в $\boldsymbol{R}$ существует вектор $\boldsymbol{g}^*$, относительный минимальный многочлен которого есть $\psi_2(\lambda)$. Тогда

$$\psi_2(\boldsymbol{A})\,\boldsymbol{g}^* \equiv 0 \qquad (\operatorname{mod} \boldsymbol{I}_1), \tag{22}$$

т. е. существует многочлен $\chi(\lambda)$ степени $\leqslant m-1$ такой, что

$$\psi_2(\boldsymbol{A})\,\boldsymbol{g}^* = \chi(\boldsymbol{A})\,\boldsymbol{e}. \tag{23}$$

Применим к обеим частям этого равенства оператор $\varkappa(\boldsymbol{A})$. Тогда слева в силу (21) получим $\psi_1(\boldsymbol{A})\,\boldsymbol{g}^*$, т. е. нуль, поскольку $\psi_1(\lambda)$ есть абсолютный минимальный многочлен пространства; следовательно,

$$\varkappa(\boldsymbol{A})\,\chi(\boldsymbol{A})\,\boldsymbol{e} = 0. \tag{24}$$

[1]) Правильнее было бы называть это подпространство циклическим относительно линейного оператора A. Но, поскольку вся теория строится для одного оператора $\boldsymbol{A}$, мы для сокращения опускаем слова «относительно линейного оператора $\boldsymbol{A}$» (см. аналогичное замечание в сноске [2]) к стр. 171).

Это равенство показывает, что произведение $\varkappa(\lambda)\chi(\lambda)$ является аннулирующим многочленом для вектора e и потому делится без остатка на минимальный многочлен $\psi_1(\lambda) = \varkappa(\lambda)\psi_2(\lambda)$, т. е. $\chi(\lambda)$ делится на $\psi_2(\lambda)$:

$$\chi(\lambda) = \varkappa_1(\lambda)\psi_2(\lambda), \tag{25}$$

где $\varkappa_1(\lambda)$ — некоторый многочлен. Используя это разложение многочлена $\chi(\lambda)$, мы равенство (23) сможем записать так:

$$\psi_2(A)\,g = 0, \tag{26}$$

где вектор g определяется равенством

$$g = g^* - \varkappa_1(A)\,e. \tag{27}$$

Последнее равенство показывает, что

$$g \equiv g^* \pmod{I_1}. \tag{28}$$

Поэтому $\psi_2(\lambda)$, будучи относительным минимальным многочленом для вектора g^*, будет таковым и для вектора g. Но тогда из равенства (26) следует, что $\psi_2(\lambda)$ является одновременно и абсолютным минимальным многочленом для вектора g.

Из того, что $\psi_2(\lambda)$ есть абсолютный минимальный многочлен вектора g, следует, что подпространство I_2 с базисом

$$g,\ Ag,\ \ldots,\ A^{p-1}g \tag{29}$$

будет циклическим.

Из того, что $\psi_2(\lambda)$ есть относительный минимальный многочлен для g по mod I_1, вытекает, что векторы (29) линейно независимы по mod I, т. е. никакая линейная комбинация векторов (29) с не равными одновременно нулю коэффициентами не может равняться линейной комбинации векторов (20). Так как эти последние сами линейно независимы, то последнее наше утверждение означает линейную независимость $m + p$ векторов

$$e,\ Ae,\ \ldots,\ A^{m-1}e;\quad g,\ Ag,\ \ldots,\ A^{p-1}g. \tag{30}$$

Векторы (30) образуют базис инвариантного подпространства $I_1 + I_2$ с числом измерений $m + p$.

Если $n = m + p$, то $R = I_1 + I_2$. Если же $n > m + p$, то мы рассмотрим R по mod $(I_1 + I_2)$ и продолжим далее наш процесс выделения циклических инвариантных подпространств. Так как все пространство R конечномерно, имеет n измерений, то этот процесс должен приостановиться на некотором подпространстве I_t, где $t \leqslant n$.

Мы приходим к следующей теореме:

Теорема 3 (2-я теорема о расщеплении пространства на инвариантные подпространства). *Пространство всегда можно расщепить на циклические относительно данного линейного оператора A подпространства $I_1, I_2, \ldots, I_t$ с минимальными многочленами $\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_t(\lambda)$,*

$$R = I_1 + I_2 + \ldots + I_t \tag{31}$$

так, чтобы $\psi_1(\lambda)$ совпадало с минимальным многочленом $\psi(\lambda)$ всего пространства и каждое $\psi_i(\lambda)$ было бы делителем $\psi_{i-1}(\lambda)$ $(i = 2, 3, \ldots, t)$.

Отметим теперь некоторые свойства циклических пространств. Пусть R — циклическое n-мерное пространство, $\psi(\lambda) = \lambda^m + \ldots$ — минимальный многочлен этого пространства. Тогда из определения циклического пространства следует, что $m = n$. Обратно, пусть нам дано произвольное

пространство $\boldsymbol{R}$ и известно, что $m = n$. Применяя доказанную теорему о расщеплении, мы представим $\boldsymbol{R}$ в виде (31). Но число измерений циклического подпространства $\boldsymbol{I}_1$ равно m, так как его минимальный многочлен совпадает с минимальным многочленом всего пространства. Так как по условию $m = n$, то $\boldsymbol{R} = \boldsymbol{I}_1$, т. е. $\boldsymbol{R}$ есть циклическое пространство.

Таким образом, установлен следующий критерий цикличности пространства:

Теорема 4. *Пространство циклично тогда и только тогда, когда его число измерений совпадает со степенью его минимального многочлена.*

Пусть теперь мы имеем расщепление циклического пространства $\boldsymbol{R}$ на два инвариантных подпространства $\boldsymbol{I}_1$ и $\boldsymbol{I}_2$:

$$\boldsymbol{R} = \boldsymbol{I}_1 + \boldsymbol{I}_2. \tag{32}$$

Обозначим числа измерений пространств $\boldsymbol{R}$, $\boldsymbol{I}_1$ и $\boldsymbol{I}_2$ соответственно через n, n_1 и n_2, минимальные многочлены этих пространств — через $\psi(\lambda)$, $\psi_1(\lambda)$ и $\psi_2(\lambda)$, степени этих минимальных многочленов — через m, m_1 и m_2. Тогда

$$m_1 \leqslant n_1, \qquad m_2 \leqslant n_2. \tag{33}$$

Сложим почленно эти неравенства:

$$m_1 + m_2 \leqslant n_1 + n_2. \tag{34}$$

Так как $\psi(\lambda)$ есть наименьшее общее кратное многочленов $\psi_1(\lambda))$ и $\psi_2(\lambda)$, то

$$m \leqslant m_1 + m_2. \tag{35}$$

Кроме того, из (32) следует:

$$n = n_1 + n_2. \tag{36}$$

(34), (35) и (36) дают нам цепочку соотношений

$$m \leqslant m_1 + m_2 \leqslant n_1 + n_2 = n. \tag{37}$$

Но в силу цикличности пространства $\boldsymbol{R}$ крайние числа в этой цепочке, числа m и n, равны между собой. Следовательно, имеет место равенство и в промежуточных звеньях этой цепочки, т. е.

$$m = m_1 + m_2 = n_1 + n_2.$$

Из того, что $m = m_1 + m_2$, заключаем, что $\psi_1(\lambda)$ и $\psi_2(\lambda)$ взаимно просты.

Из $m_1 + m_2 = n_1 + n_2$, принимая во внимание (33), находим:

$$m_1 = n_1, \qquad m_2 = n_2. \tag{38}$$

Эти же равенства означают цикличность подпространства $\boldsymbol{I}_1$ и $\boldsymbol{I}_2$.

Таким образом, мы приходим к следующему предложению:

Теорема 5. *Циклическое пространство расщепляется только на такие инвариантные подпространства, которые* 1° *сами циклические и* 2° *имеют взаимно простые минимальные многочлены.*

Те же рассуждения (проведенные в обратном порядке) показывают, что теорема 5 допускает обращение.

Теорема 6. *Если пространство расщепляется на инвариантные подпространства, которые* 1° *являются циклическими и* 2° *имеют взаимно простые минимальные многочлены, то само пространство является циклическим.*

Пусть теперь R — циклическое пространство и минимальный многочлен его есть степень неприводимого в поле K многочлена: $\psi(\lambda) = [\varphi(\lambda)]^c$. В этом случае минимальный многочлен любого инвариантного подпространства в R тоже будет степенью этого неприводимого многочлена $\varphi(\lambda)$. Следовательно, минимальные многочлены любых двух инвариантных подпространств не могут быть взаимно простыми. Но тогда в силу доказанного предложения R не расщепляется на инвариантные подпространства.

Пусть, обратно, известно, что некоторое пространство R не расщепляется на инвариантные подпространства. Тогда R — циклическое пространство, иначе его можно было расщепить с помощью второй теоремы о расщеплении на циклические подпространства; кроме того, минимальный многочлен R должен быть степенью неприводимого многочлена, так как в противном случае R можно было бы расщепить на инвариантные подпространства в силу первой теоремы о расщеплении.

Таким образом, приходим к следующему выводу:

Теорема 7. *Пространство не расщепляется на инвариантные подпространства тогда и только тогда, когда* 1° *оно циклическое и* 2° *минимальный многочлен его есть степень неприводимого в поле K многочлена.*

Вернемся теперь к расщеплению (31) и разложим минимальные многочлены $\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_t(\lambda)$ циклических подпространств $I_1, I_2, \ldots, I_t$ на неприводимые в поле K множители:

$$\left.\begin{aligned}
\psi_1(\lambda) &= [\varphi_1(\lambda)]^{c_1}[\varphi_2(\lambda)]^{c_2}\ldots[\varphi_s(\lambda)]^{c_s},\\
\psi_2(\lambda) &= [\varphi_1(\lambda)]^{d_1}[\varphi_2(\lambda)]^{d_2}\ldots[\varphi_s(\lambda)]^{d_s},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\psi_t(\lambda) &= [\varphi_1(\lambda)]^{l_1}[\varphi_2(\lambda)]^{l_2}\ldots[\varphi_s(\lambda)]^{l_s}
\end{aligned}\right\} \qquad (39)$$

$(c_k \geqslant d_k \geqslant \ldots \geqslant l_k \geqslant 0; \quad k = 1, 2, \ldots, s)$ [1]).

Применим к I_1 первую теорему о расщеплении. Тогда получим:

$$I_1 = I_1' + I_1'' + \ldots + I_1^{(s)},$$

где $I_1', I_1'', \ldots, I_1^{(s)}$ — циклические подпространства с минимальными многочленами $[\varphi_1(\lambda)]^{c_1}, [\varphi_2(\lambda)]^{c_2}, \ldots, [\varphi_s(\lambda)]^{c_s}$. Аналогично расщепим подпространства $I_2, \ldots, I_t$. Тем самым мы получим расщепление всего пространства R на циклические подпространства с минимальными многочленами $[\varphi_k(\lambda)^{c_k}, [\varphi_k(\lambda)]^{d_k}, \ldots, [\varphi_k(\lambda)]^{l_k}$ $(k = 1, 2, \ldots, s)$ (при этом выбрасываются те степени, у которых показатели равны нулю). Из теоремы 7 следует, что эти циклические подпространства уже далее нерасщепимы (на инвариантные подпространства). Приходим к следующей теореме:

Теорема 8 (3-я теорема о расщеплении пространства на инвариантные подпространства). *Пространство всегда можно расщепить на циклические инвариантные подпространства*

$$R = I' + I'' + \ldots + I^{(u)} \qquad (40)$$

так, чтобы минимальный многочлен каждого из этих циклических подпространств был степенью неприводимого многочлена.

[1]) Некоторые из показателей $d_k, \ldots, l_k$ при $k > 1$ могут равняться нулю.

Эта теорема дает расщепление пространства на нерасщепимые далее инвариантные подпространства.

Замечание. Теорему 8 (3-ю теорему о расщеплении) мы получили, применяя первые две теоремы о расщеплении. Однако 3-ю теорему о расщеплении можно получить другим путем, а именно, как непосредственное (почти тривиальное) следствие из теоремы 7.

Действительно, пространство $\boldsymbol{R}$, если оно вообще расщепляется, всегда можно расщепить на нерасщепимые далее инвариантные подпространства:

$$\boldsymbol{R} = \boldsymbol{I}' + \boldsymbol{I}'' + \ldots + \boldsymbol{I}^{(u)}.$$

Согласно теореме 7 каждое из слагаемых подпространств является циклическим и имеет в качестве своего минимального многочлена степень неприводимого в K многочлена.

§ 5. Нормальная форма матрицы

Пусть $\boldsymbol{I}_1$ — m-мерное инвариантное подпространство в $\boldsymbol{R}$. Выберем в $\boldsymbol{I}_1$ произвольно базис $e_1, e_2, \ldots, e_m$ и дополним его до базиса в $\boldsymbol{R}$:

$$e_1, e_2, \ldots, e_m, e_{m+1}, \ldots, e_n$$

Посмотрим, как будет выглядеть матрица A оператора $\boldsymbol{A}$ в этом базисе. Напомним читателю, что k-й столбец матрицы A заполняется координатами вектора $\boldsymbol{A}e_k$ ($k = 1, 2, \ldots, n$). При $k \leqslant m$ вектор $\boldsymbol{A}e_k \in \boldsymbol{I}_1$ (в силу инвариантности $\boldsymbol{I}_1$) и, следовательно, последние $n - m$ координат вектора $\boldsymbol{A}e_k$ равны нулю. Поэтому матрица A имеет такую форму:

$$A = \begin{matrix} m & n-m \\ \overbrace{} & \overbrace{} \end{matrix} \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix}, \tag{41}$$

где A_1 и A_2 — квадратные матрицы порядка m и $n - m$, а A_3 — прямоугольная матрица. Равенство нулю четвертого «блока» и выражает инвариантность подпространства $\boldsymbol{I}_1$. Матрица A_1 задает оператор $\boldsymbol{A}$ в $\boldsymbol{I}_1$ (при базисе $e_1, e_2, \ldots, e_m$).

Допустим теперь, что $e_{m+1}, \ldots, e_n$ тоже есть базис некоторого инвариантного подпространства $\boldsymbol{I}_2$, т. е. $\boldsymbol{R} = \boldsymbol{I}_1 + \boldsymbol{I}_2$ и базис всего пространства составлен из двух частей, которые служат базисами в инвариантных подпространствах $\boldsymbol{I}_1$ и $\boldsymbol{I}_2$. Тогда, очевидно, в (41) и блок A_3 будет равен нулю и матрица A будет иметь квазидиагональный вид:

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = \{A_1, A_2\}, \tag{42}$$

где A_1 и A_2 — квадратные матрицы порядков m и $n - m$, задающие оператор в подпространствах $\boldsymbol{I}_1$ и $\boldsymbol{I}_2$ (при базисах соответственно $e_1, e_2, \ldots, e_m$ и $e_{m+1}, \ldots, e_n$). Нетрудно видеть, что и, обратно, квазидиагональному виду матрицы всегда соответствует расщепление пространства на инвариантные подпространства (при этом базис всего пространства составлен из базисов этих подпространств).

В силу 2-й теоремы о расщеплении мы можем расщепить все пространство $\boldsymbol{R}$ на циклические подпространства $\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots, \boldsymbol{I}_t$:

$$\boldsymbol{R} = \boldsymbol{I}_1 + \boldsymbol{I}_2 + \ldots + \boldsymbol{I}_t. \tag{43}$$

В ряду минимальных многочленов этих подпространств $\psi_1(\lambda)$, $\psi_2(\lambda)$, ..., $\psi_t(\lambda)$ каждый многочлен есть делитель предыдущего (отсюда уже автоматически следует, что первый многочлен есть минимальный многочлен всего пространства).

Пусть

$$\begin{aligned}
&\psi_1(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_m,\\
&\psi_2(\lambda) = \lambda^p + \beta_1\lambda^{p-1} + \ldots + \beta_p, \qquad (m \geqslant p \geqslant \ldots \geqslant v).\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&\psi_t(\lambda) = \lambda^v + \varepsilon_1\lambda^{v-1} + \ldots + \varepsilon_v.
\end{aligned} \tag{44}$$

Обозначим через e, g, ..., l порождающие векторы в подпространствах I_1, I_2, ..., I_t и составим базис всего пространства R из следующих базисов циклических подпространств:

$$e, Ae, \ldots, A^{m-1}e;\ g, Ag, \ldots, A^{p-1}g;\ \ldots;\ l, Al, \ldots, A^{v-1}l. \tag{45}$$

Посмотрим, какова будет матрица L_I, отвечающая оператору A в этом базисе.

Как было выяснено в начале этого параграфа, матрица L_I должна иметь квазидиагональную форму

$$L_I = \begin{pmatrix} L_1 & 0 & \ldots & 0\\ 0 & L_2 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & L_t \end{pmatrix}. \tag{46}$$

Матрица L_1 отвечает оператору A в I_1 при базисе $e_1 = e$, $e_2 = Ae$, ..., $e_m = A^{m-1}e$. Припоминая правило составления матрицы по заданному оператору и заданному базису (гл. III, стр. 71), найдем:

$$L_1 = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\alpha_m\\ 1 & 0 & \ldots & 0 & -\alpha_{m-1}\\ 0 & 1 & \ddots & & \vdots\\ \vdots & \ddots & \ddots & 0 & -\alpha_2\\ 0 & 0 & \ldots & 1 & -\alpha_1 \end{Vmatrix}. \tag{47}$$

Аналогично

$$L_2 = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\beta_p\\ 1 & 0 & \ldots & 0 & -\beta_{p-1}\\ 0 & 1 & \ddots & & \vdots\\ \vdots & \ddots & \ddots & 0 & -\beta_2\\ 0 & 0 & \ldots & 1 & -\beta_1 \end{Vmatrix}, \tag{48}$$

и т. д.

Вычислив характеристические многочлены матриц L_1, L_2, ..., L_t, получим:

$$|\lambda E - L_1| = \psi_1(\lambda), \quad |\lambda E - L_2| = \psi_2(\lambda), \ \ldots, \ |\lambda E - L_t| = \psi_t(\lambda)$$

(для циклических подпространств характеристический многочлен оператора A совпадает с минимальным многочленом подпространства относительно этого оператора).

Матрица L_{I} отвечает оператору A в «каноническом» базисе (45). Если A — матрица, отвечающая оператору A в произвольном базисе, то матрица A подобна матрице L_{I}, т. е. существует такая неособенная матрица T, что

$$A = TL_{\mathrm{I}}T^{-1}. \tag{49}$$

Про матрицу L_{I} мы будем говорить, что она имеет *первую естественную нормальную форму*. Первая естественная нормальная форма характеризуется

1) квазидиагональным видом (46),

2) специальной структурой диагональных клеток (47), (48) и т. п.,

3) дополнительным условием: характеристический многочлен каждой диагональной клетки делится нацело на характеристический многочлен следующей клетки.

Точно так же, если бы мы исходили не из 2-й, а из 3-й теоремы о расщеплении, то в соответствующем базисе оператору A отвечала бы матрица L_{II}, имеющая вторую *естественную нормальную форму*, характеризуемую

1) квазидиагональным видом

$$L_{\mathrm{II}} = \{L_1, L_2, \ldots, L_u\},$$

2) специальной структурой диагональных клеток (47), (48) и т. п.,

3) дополнительным условием: характеристический многочлен каждой клетки является степенью неприводимого в поле K многочлена.

В следующем параграфе мы покажем, что в классе подобных матриц, отвечающих одному и тому же оператору, существует только одна матрица, имеющая первую нормальную форму [1]), и только одна [2]), имеющая вторую нормальную форму. Более того, мы дадим алгоритм для нахождения многочленов $\psi_1(\lambda)$, $\psi_2(\lambda)$, ..., $\psi_t(\lambda)$ по элементам матрицы A. Знание этих многочленов даст нам возможность выписать все элементы матриц L_{I} и L_{II}, подобных матрице A и имеющих соответственно первую и вторую нормальную форму.

§ 6. Инвариантные многочлены. Элементарные делители

1 [3]). Обозначим через $D_p(\lambda)$ наибольший общий делитель всех миноров p-го порядка характеристической матрицы $A_\lambda = \lambda E - A$ ($p = 1, 2, \ldots, n$) [4]). Так как в ряду

$$D_n(\lambda), \quad D_{n-1}(\lambda), \ldots, D_1(\lambda)$$

каждый многочлен делится на последующий без остатка, то формулы

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \ldots, i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1) \tag{50}$$

---

[1]) Это не означает, что существует только один канонический базис вида (45). Канонических базисов может быть много, но всем им отвечает одна и та же матрица L_{I}.

[2]) С точностью до порядка диагональных клеток.

[3]) В п. 1 настоящего параграфа повторяются для характеристической матрицы основные понятия гл. VI, § 3, установленные там для произвольной многочленной матрицы.

[4]) В наибольшем общем делителе всегда выбираем старший коэффициент равным единице.

определяют n многочленов, произведение которых равно характеристическому многочлену

$$\Delta(\lambda) = |\lambda E - A| = D_n(\lambda) = i_1(\lambda)\, i_2(\lambda) \ldots i_n(\lambda). \tag{51}$$

Многочлены $i_p(\lambda)$ $(p = 1, 2, \ldots, n)$ разложим на неприводимые в поле K множители:

$$i_p(\lambda) = [\varphi_1(\lambda)]^{\gamma_p} [\varphi_2(\lambda)]^{\delta_p} \ldots \quad (p = 1, 2, \ldots, n), \tag{52}$$

где $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, $\ldots$ — различные неприводимые в поле K многочлены.

Многочлены $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_n(\lambda)$ называются *инвариантными многочленами*, а все степени, отличные от постоянной, среди $[\varphi_1(\lambda)]^{\gamma_p}$, $[\varphi_2(\lambda)]^{\delta_p}$, $\ldots$, называются *элементарными делителями* характеристической матрицы $A_\lambda = \lambda E - A$ или просто матрицы A.

Произведение всех элементарных делителей, как и произведение всех инвариантных многочленов, равно характеристическому многочлену $\Delta(\lambda) = |\lambda E - A|$.

Название «инвариантные многочлены» оправдано тем, что две подобные матрицы A и $\widetilde{A}$

$$\widetilde{A} = T^{-1} A T \tag{53}$$

всегда имеют одни и те же инвариантные многочлены

$$i_p(\lambda) = \widetilde{i}_p(\lambda) \qquad (p = 1, 2, \ldots, n). \tag{54}$$

Действительно, из (53) следует:

$$\widetilde{A}_\lambda = \lambda E - \widetilde{A} = T^{-1}(\lambda E - A)\, T = T^{-1} A_\lambda T. \tag{55}$$

Отсюда (см. гл. I, § 2) получаем соотношение между минорами подобных матриц A_λ и $\widetilde{A}_\lambda$:

$$\widetilde{A}_\lambda \begin{pmatrix} i_1\, i_2 \ldots i_p \\ k_1 k_2 \ldots k_p \end{pmatrix} = \sum_{\substack{\alpha_1 < \alpha_2 < \ldots < \alpha_p \\ \beta_1 < \beta_2 < \ldots < \beta_p}} T^{-1} \begin{pmatrix} i_1\, i_2 \ldots i_p \\ \alpha_1 \alpha_2 \ldots \alpha_p \end{pmatrix} A_\lambda \begin{pmatrix} \alpha_1 \alpha_2 \ldots \alpha_p \\ \beta_1 \beta_2 \ldots \beta_p \end{pmatrix} T \begin{pmatrix} \beta_1 \beta_2 \ldots \beta_p \\ k_1 k_2 \ldots k_p \end{pmatrix} \tag{56}$$

$$(p = 1, 2, \ldots, n).$$

Это равенство показывает, что каждый общий делитель всех миноров p-го порядка матрицы A_λ является общим делителем всех миноров p-го порядка матрицы $\widetilde{A}_\lambda$ и наоборот (поскольку матрицы A и $\widetilde{A}$ можно поменять местами). Отсюда вытекает: $D_p(\lambda) = \widetilde{D}_p(\lambda)$ $(p = 1, 2, \ldots, n)$ и, следовательно, имеет место (54).

Поскольку все матрицы, представляющие данный оператор A в различных базисах, подобны между собой и потому имеют одни и те же инвариантные многочлены и, следовательно, одни и те же элементарные делители, то можно говорить об инвариантных многочленах и элементарных делителях оператора A.

2. Возьмем теперь в качестве $\widetilde{A}$ матрицу L_{I}, имеющую первую естественную нормальную форму, и вычислим инвариантные многочлены матрицы A, исходя из вида матрицы $\widetilde{A}_\lambda = \lambda E - \widetilde{A}$ (на схеме (57) эта

матрица выписана для случая $m=5$, $p=4$, $q=4$, $r=3$):

$$\left(\begin{array}{ccccc|cccc|cccc|ccc}
\lambda & 0 & 0 & 0 & \alpha_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & \lambda & 0 & 0 & \alpha_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & \lambda & 0 & \alpha_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & \lambda & \alpha_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & \alpha_1+\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & \beta_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & \lambda & 0 & \beta_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \beta_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \beta_1+\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & \gamma_4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & 0 & \gamma_3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \gamma_2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \gamma_1+\lambda & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & \varepsilon_3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \varepsilon_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \varepsilon_1+\lambda
\end{array}\right) \tag{57}$$

Пользуясь теоремой Лапласа, найдем:

$$D_n(\lambda)=|\lambda E-\widetilde{A}|=|\lambda E-L_1||\lambda E-L_2|\ldots|\lambda E-L_t|=\psi_1(\lambda)\,\psi_2(\lambda)\ldots\psi_t(\lambda). \tag{58}$$

Перейдем к отысканию $D_{n-1}(\lambda)$. Обратим внимание на минор элемента α_m. Этот минор с точностью до множителя ± 1 равен

$$|\lambda E-L_2|\ldots|\lambda E-L_t|=\psi_2(\lambda)\ldots\psi_t(\lambda). \tag{59}$$

Мы докажем, что этот минор $(n-1)$-го порядка будет делителем всех прочих миноров $(n-1)$-го порядка и что, следовательно,

$$D_{n-1}(\lambda)=\psi_2(\lambda)\ldots\psi_t(\lambda). \tag{60}$$

Для этого возьмем сначала минор элемента, находящегося вне диагональных клеток, и покажем, что такой минор равен нулю. Для получения этого минора нам придется из матрицы (57) вычеркнуть одну строку и один столбец. Вычеркнутые линии в рассматриваемом случае пересекут две разные диагональные клетки и, следовательно, у каждой из этих двух клеток будет вычеркнуто по одной линии. Пусть, например, у j-й диагональной клетки будет вычеркнута одна из строк. Возьмем в миноре ту вертикальную полосу, в которой содержится эта диагональная клетка. В этой полосе, имеющей s столбцов, все строки, за исключением $s-1$ строк, будут состоять сплошь из нулей (здесь через s мы обозначили порядок матрицы A_j). Разлагая рассматриваемый определитель $(n-1)$-го порядка на основании теоремы Лапласа по минорам s-го порядка, содержащимся в указанной полосе, мы и убедимся в том, что он равен нулю.

Возьмем теперь минор элемента, находящегося внутри одной из диагональных клеток. В этом случае вычеркиваемые линии «искалечат» только одну из диагональных клеток, например j-ю, и матрица минора

снова будет квазидиагональной. Поэтому такой минор будет равен

$$\psi_1(\lambda)\ldots\psi_{j-1}(\lambda)\,\psi_{j+1}(\lambda)\ldots\psi_t(\lambda)\,\chi(\lambda), \tag{61}$$

где $\chi(\lambda)$ — определитель «искалеченной» j-й диагональной клетки. В силу того, что $\psi_i(\lambda)$ делится нацело на $\psi_{i+1}(\lambda)$ ($i = 1, 2, \ldots, t-1$), произведение (61) разделится без остатка на произведение (59). Таким образом, равенство (60) можно считать доказанным. Аналогичными рассуждениями получим:

$$\left.\begin{aligned} D_{n-2}(\lambda) &= \psi_3(\lambda)\ldots\psi_t(\lambda),\\ &\cdots\cdots\cdots\cdots\cdots\\ D_{n-t+1}(\lambda) &= \psi_t(\lambda),\\ D_{n-t}(\lambda) &= \ldots = D_1(\lambda) = 1. \end{aligned}\right\} \tag{62}$$

Из (58), (60) и (62) находим:

$$\left.\begin{aligned} &\varphi_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)} = i_1(\lambda), \quad \psi_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)} = i_2(\lambda), \ldots\\ &\ldots, \psi_t(\lambda) = \frac{D_{n-t+1}(\lambda)}{D_{n-1}(\lambda)} = i_t(\lambda), \quad i_{t+1}(\lambda) = \ldots = i_n(\lambda) = 1. \end{aligned}\right\} \tag{63}$$

Формулы (63) показывают, что многочлены $\psi_1(\lambda)$, $\psi_2(\lambda)$, ..., $\psi_t(\lambda)$ совпадают с отличными от единицы инвариантными многочленами оператора A (либо соответствующей матрицы A). Но тогда отличные от единицы $[\varphi_k(\lambda)]^{c_k}$, $[\varphi_k(\lambda)]^{d_k}$, ..., ($k = 1, 2, \ldots$) в разложении (39) совпадают с элементарными делителями оператора A (либо соответствующей матрицы A). Поэтому задание инвариантных многочленов, или, что то же, задание элементарных делителей в поле K, однозначно определяет элементы нормальных форм L_{I} и L_{II}.

На стр. 185 было установлено, что две подобные матрицы имеют одни и те же инвариантные многочлены. Пусть теперь, обратно, известно, что две матрицы A и B с элементами из K имеют одни и те же инвариантные многочлены. Так как матрица L_{I} однозначно определяется заданием этих многочленов, то обе матрицы A и B подобны одной и той же матрице L_{I} и, следовательно, подобны между собой. Таким образом, приходим к следующему предложению:

Теорема 9. *Для того чтобы две матрицы с элементами из K были подобны, необходимо и достаточно, чтобы у этих матриц были одни и те же инвариантные многочлены* [1]).

Характеристический многочлен $\Delta(\lambda)$ оператора A совпадает с $D_n(\lambda)$ и потому равен произведению всех инвариантных многочленов:

$$\Delta(\lambda) = \psi_1(\lambda)\,\psi_2(\lambda)\ldots\psi_t(\lambda). \tag{64}$$

Но $\psi_1(\lambda)$ есть минимальный многочлен всего пространства относительно A; значит, $\psi_1(A) = 0$ и в силу (64)

$$\Delta(A) = 0. \tag{65}$$

Таким образом, мы попутно получили теорему Гамильтона — Кэли (см. гл. IV, § 3).

[1]) Или одинаковые элементарные делители в поле K.

§ 7. Нормальная жорданова форма матрицы

Пусть все корни характеристического многочлена $\Delta(\lambda)$ оператора $\boldsymbol{A}$ принадлежат полю K. Это, в частности, всегда будет иметь место, если K есть поле всех комплексных чисел.

В рассматриваемом случае разложение инвариантных многочленов на элементарные делители в поле K будет выглядеть так:

$$\left.\begin{aligned} i_1(\lambda) &= (\lambda-\lambda_1)^{c_1}(\lambda-\lambda_2)^{c_2}\ldots(\lambda-\lambda_s)^{c_s},\\ i_2(\lambda) &= (\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\ldots(\lambda-\lambda_s)^{d_s},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ i_t(\lambda) &= (\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\ldots(\lambda-\lambda_s)^{l_s} \end{aligned}\right\} \quad \left.\begin{matrix} c_k \geqslant d_k \geqslant \ldots \geqslant l_k \geqslant 0,\\ c_k > 0;\ k = 1, 2, \ldots, s \end{matrix}\right) . \tag{66}$$

Так как произведение всех инвариантных многочленов равно характеристическому многочлену $\Delta(\lambda)$, то $\lambda_1, \lambda_2, \ldots, \lambda_s$ в (66) суть все различные между собой корни характеристического многочлена $\Delta(\lambda)$.

Возьмем какой-либо элементарный делитель

$$(\lambda-\lambda_0)^p; \tag{67}$$

здесь λ_0 — одно из чисел $\lambda_1, \lambda_2, \ldots \lambda_s$, а p — один из (отличных от нуля) показателей $c_k, d_k, \ldots, l_k$ $(k = 1, 2, \ldots, s)$.

Этому элементарному делителю в расщеплении (40) отвечает определенное циклическое подпространство $\boldsymbol{I}$, порождающий вектор которого обозначим буквой $\boldsymbol{e}$. Для этого вектора $(\lambda-\lambda_0)^p$ будет минимальным многочленом.

Рассмотрим векторы

$$\boldsymbol{e}_1 = (\boldsymbol{A}-\lambda_0\boldsymbol{E})^{p-1}\boldsymbol{e}, \quad \boldsymbol{e}_2 = (\boldsymbol{A}-\lambda_0\boldsymbol{E})^{p-2}\boldsymbol{e}, \ldots, \boldsymbol{e}_p = \boldsymbol{e}. \tag{68}$$

Векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$ линейно независимы, так как в противном случае существовал бы аннулирующий многочлен для вектора $\boldsymbol{e}$ степени $< p$, что невозможно. Теперь заметим, что

$$(\boldsymbol{A}-\lambda_0\boldsymbol{E})\boldsymbol{e}_1 = 0, \quad (\boldsymbol{A}-\lambda_0\boldsymbol{E})\boldsymbol{e}_2 = \boldsymbol{e}_1, \ldots, (\boldsymbol{A}-\lambda_0\boldsymbol{E})\boldsymbol{e}_p = \boldsymbol{e}_{p-1} \tag{69}$$

или

$$\boldsymbol{A}\boldsymbol{e}_1 = \lambda_0\boldsymbol{e}_1, \quad \boldsymbol{A}\boldsymbol{e}_2 = \lambda_0\boldsymbol{e}_2 + \boldsymbol{e}_1, \ldots, \boldsymbol{A}\boldsymbol{e}_p = \lambda_0\boldsymbol{e}_p + \boldsymbol{e}_{p-1}. \tag{70}$$

Имея равенства (70), нетрудно выписать матрицу, отвечающую оператору $\boldsymbol{A}$ в $\boldsymbol{I}$ при базисе (68). Эта матрица будет выглядеть так:

$$\begin{Vmatrix} \lambda_0 & 1 & 0 \ldots 0 \\ 0 & \lambda_0 & 1 \ldots 0 \\ \cdot & & \cdot \quad \cdot \\ \cdot & & \cdot \quad \cdot \\ \cdot & & \cdot \quad 1 \\ 0 & 0 & 0 \ldots \lambda_0 \end{Vmatrix} = \lambda_0 E^{(p)} + H^{(p)}, \tag{71}$$

где $E^{(p)}$ — единичная матрица порядка p, а $H^{(p)}$ — матрица порядка p, у которой элементы первой «наддиагонали» равны единице, а все остальные элементы равны нулю.

Линейно независимые векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$, для которых имеют место равенства (70), образуют так называемую *жорданову цепочку векторов* в $\boldsymbol{I}$. Из жордановых цепочек, взятых в каждом из под-

пространств I', I'', ..., $I^{(u)}$, составляется *жорданов базис* в R. Если минимальные многочлены этих подпространств, т. е. элементарные делители оператора A, обозначим теперь через

$$(\lambda-\lambda_1)^{p_1}, \quad (\lambda-\lambda_2)^{p_2}, \ldots, (\lambda-\lambda_u)^{p_u} \tag{72}$$

(среди чисел $\lambda_1, \lambda_2, \ldots, \lambda_u$ могут быть и равные), то матрица J, отвечающая оператору A в жордановом базисе, будет иметь следующий квазидиагональный вид:

$$J=\{\lambda_1 E^{(p_1)}+H^{(p_1)}, \quad \lambda_2 E^{(p_2)}+H^{(p_2)}, \ldots, \lambda_u E^{(p_u)}+H^{(p_u)}\}. \tag{73}$$

Про матрицу J говорят, что она имеет *нормальную жорданову форму* или просто *жорданову форму*. Матрица J сразу выписывается, если известны элементарные делители оператора A в поле K, содержащем все корни характеристического уравнения $\Delta(\lambda)=0$.

Произвольная матрица A всегда подобна матрице J, имеющей нормальную жорданову форму, т. е. для произвольной матрицы A всегда существует такая неособенная матрица T ($|T|\neq 0$), что

$$A=TJT^{-1}. \tag{74}$$

Если все элементарные делители оператора A первой степени (и только в этом случае), жорданова форма является диагональной матрицей, и в этом случае мы имеем:

$$A=T\{\lambda_1, \lambda_2, \ldots, \lambda_n\}T^{-1}. \tag{75}$$

Таким образом, *линейный оператор A имеет простую структуру* (см. гл. III, § 8) *в том и только в том случае, когда все элементарные делители оператора A линейны.*

Векторы $e_1, e_2, \ldots, e_p$, определяемые равенствами (70), занумеруем в обратном порядке:

$$g_1=e_p=e, \quad g_2=e_{p-1}=(A-\lambda_0 E)e, \ldots, g_p=e_1=(A-\lambda_0 E)^{p-1}e. \tag{76}$$

Тогда

$$(A-\lambda_0 E)g_1=g_2, \quad (A-\lambda_0 E)g_2=g_3, \ldots, (A-\lambda_0 E)g_p=0, \tag{77}$$

откуда

$$Ag_1=\lambda_0 g_1+g_2, \quad Ag_2=\lambda_0 g_2+g_3, \ldots, Ag_p=\lambda_0 g_p. \tag{78}$$

Векторы (76) образуют базис в циклическом инвариантном подпространстве I, соответствующем в расщеплении (40) элементарному делителю $(\lambda-\lambda_0)^p$. В этом базисе, как легко видеть, оператору A будет отвечать матрица

$$\begin{Vmatrix} \lambda_0 & 0 & 0 & \ldots & 0 \\ 1 & \lambda_0 & 0 & \ldots & 0 \\ 0 & 1 & \lambda_0 & & \\ \cdot & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & \ldots & 1 & \lambda_0 \end{Vmatrix} = \lambda_0 E^{(p)}+F^{(p)}. \tag{79}$$

Про векторы (76) говорят, что они образуют *нижнюю жорданову цепочку* векторов. Если мы в каждом из подпространств I', I'', ..., $I^{(u)}$ в расщеплении (40) возьмем нижнюю жорданову цепочку векторов, то из этих цепочек составится *нижний жорданов базис*, в котором оператору A

отвечает квазидиагональная матрица

$$J_1=\{\lambda_1 E^{(p_1)}+F^{(p_1)},\quad \lambda_2 E^{(p_2)}+F^{(p_2)},\ \ldots,\ \lambda_u E^{(p_u)}+F^{(p_u)}\}. \tag{80}$$

Про матрицу J_1 говорят, что она имеет *нижнюю жорданову форму*. В отличие от матрицы (80) матрицу (73) мы иногда будем называть *верхней жордановой матрицей*.

Таким образом, *произвольная матрица A всегда подобна как некоторой верхней, так и некоторой нижней жордановой матрице.*

§ 8. Метод акад. А. Н. Крылова преобразования векового уравнения

1. Если дана матрица $A=\|a_{ik}\|_1^n$, то ее характеристическое (вековое) уравнение записывается в виде

$$\Delta(\lambda)\equiv(-1)^n\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22}-\lambda & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda\end{vmatrix}=0. \tag{81}$$

В левой части этого уравнения стоит характеристический многочлен n-й степени $\Delta(\lambda)$. Для непосредственного вычисления коэффициентов этого многочлена нужно раскрыть характеристический определитель $|A-\lambda E|$, а это связано при больших n с большим объемом вычислительной работы, поскольку λ входит в диагональные элементы определителя [1]).

Академик А. Н. Крылов в 1937 г. [103] предложил преобразование характеристического определителя, в результате которого λ входит только в элементы одного столбца (или строки). Преобразование Крылова существенно упрощает вычисление коэффициентов характеристического уравнения [2]).

В этом параграфе мы дадим алгебраический вывод преобразованного характеристического уравнения, несколько отличающийся от вывода самого Крылова [3]).

Введем в рассмотрение n-мерное векторное пространство $\boldsymbol{R}$ с базисом $e_1, e_2, \ldots, e_n$ и линейный оператор $\boldsymbol{A}$ в $\boldsymbol{R}$, определяемый данной матрицей $A=\|a_{ik}\|_1^n$ при этом базисе. Выберем в $\boldsymbol{R}$ произвольный вектор $x\neq 0$ и составим ряд векторов

$$x,\ Ax,\ A^2x,\ \ldots \tag{82}$$

Пусть первые p векторов этого ряда x, Ax, $\ldots$, $A^{p-1}x$ линейно независимы а $(p+1)$-й вектор A^px есть линейная комбинация этих p векторов:

$$A^px=-\alpha_p x-\alpha_{p-1}Ax-\ldots-\alpha_1 A^{p-1}x, \tag{83}$$

или

$$\varphi(A)\,x=0, \tag{84}$$

где

$$\varphi(\lambda)=\lambda^p+\alpha_1\lambda^{p-1}+\ldots+\alpha_p. \tag{85}$$

Все дальнейшие векторы ряда (82) также линейно выражаются через первые p векторов этого ряда [4]). Таким образом, в ряду (82) имеется p линейно независимых

---

[1]) Напоминаем, что коэффициент при λ^k в $\Delta(\lambda)$ равен (с точностью до знака) сумме всех главных миноров порядка $n-k$ матрицы A $(k=1, 2, \ldots, n)$. Таким образом, уже при $n=6$ для непосредственного определения коэффициента при λ в $\Delta(\lambda)$ нужно вычислить шесть определителей пятого порядка, а для коэффициента при λ^2 нужно вычислить 15 определителей четвертого порядка и т. п.

[2]) Алгебраическому анализу метода А. Н. Крылова преобразования векового уравнения посвящен ряд исследований [109а, б; 133, 81б, 132].

[3]) А. Н. Крылов пришел к своему преобразованному уравнению, исходя из рассмотрения системы n линейных дифференциальных уравнений с постоянными коэффициентами. Вывод Крылова в алгебраизированной форме можно найти, например, в [109а, 81б], а также в книге [33], § 21.

[4]) Применяя к обеим частям равенства (83) оператор A, мы выражаем линейно $A^{p+1}x$ через векторы Ax, $\ldots$, $A^{p-1}x$, A^px. Но A^px в силу (83) линейно выражается через векторы x, Ax, $\ldots$, $A^{p-1}x$. Поэтому мы получаем аналогичное выражение для $A^{p+1}x$. Применяя к этому выражению вектора $A^{p+1}x$ оператор A, мы выразим $A^{p+2}x$ через x, Ax, $\ldots$, $A^{p-1}x$ и т. д.

векторов, и это максимальное число линейно независимых векторов ряда (82) может быть всегда реализовано на первых p векторах ряда.

Многочлен $\varphi(\lambda)$ является минимальным (аннулирующим) многочленом вектора x относительно оператора A (см. § 1). *Метод А. Н. Крылова есть метод эффективного определения минимального многочлена* $\varphi(\lambda)$ *вектора* x.

Мы рассмотрим раздельно два случая: *регулярный* случай, когда $p=n$, и *особый* случай, когда $p<n$.

Многочлен $\varphi(\lambda)$ является делителем минимального многочлена $\psi(\lambda)$ всего пространства R [1]), а $\psi(\lambda)$ в свою очередь является делителем характеристического многочлена $\Delta(\lambda)$. Поэтому $\varphi(\lambda)$ всегда является делителем $\Delta(\lambda)$.

В регулярном случае $\varphi(\lambda)$ и $\Delta(\lambda)$ имеют одну и ту же степень n, и поскольку старшие коэффициенты у них равны, то эти многочлены совпадают. Таким образом, *в регулярном случае*

$$\Delta(\lambda)\equiv\psi(\lambda)\equiv\varphi(\lambda),$$

и потому *метод Крылова в регулярном случае есть метод вычисления коэффициентов характеристического многочлена* $\Delta(\lambda)$.

В особом случае, как мы увидим ниже, метод Крылова не дает возможности определить $\Delta(\lambda)$ и в этом случае он определяет только многочлен $\varphi(\lambda)$, являющийся делителем $\Delta(\lambda)$.

При изложении преобразования Крылова мы будем обозначать координаты вектора x в заданном базисе $e_1, e_2, \ldots, e_n$ через $a, b, \ldots, l$, а координаты вектора $A^k x$ через $a_k, b_k, \ldots, l_k$ $(k=1, 2, \ldots, n)$.

2. Регулярный случай: $p=n$. В этом случае векторы $x, Ax, \ldots, A^{n-1}x$ линейно независимы, и равенства (83), (84), (85) принимают вид

$$A^n x=-\alpha_n x-\alpha_{n-1}Ax-\ldots-\alpha_1 A^{n-1}x \tag{86}$$

или

$$\Delta(A)\,x=0, \tag{87}$$

где

$$\Delta(\lambda)=\lambda^n+\alpha_1\lambda^{n-1}+\ldots+\alpha_{n-1}\lambda+\alpha_n. \tag{88}$$

Условие линейной независимости векторов $x, Ax, \ldots, A^{n-1}x$ может быть аналитически записано так (см. гл. III, § 1):

$$M=\begin{vmatrix} a & b & \ldots & l \\ a_1 & b_1 & \ldots & l_1 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & b_{n-1} & \ldots & l_{n-1} \end{vmatrix}\neq 0. \tag{89}$$

Рассмотрим матрицу, составленную из координат векторов $x, Ax, \ldots, A^n x$:

$$\begin{Vmatrix} a & b & \ldots & l \\ a_1 & b_1 & \ldots & l_1 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & b_{n-1} & \ldots & l_{n-1} \\ a_n & b_n & \ldots & l_n \end{Vmatrix}. \tag{90}$$

В регулярном случае ранг этой матрицы равен n. Первые n строк этой матрицы линейно независимы, а последняя, $(n+1)$-я, строка есть линейная комбинация предыдущих n.

Зависимость между строками матрицы (90) получим, заменяя векторное равенство (86) эквивалентной системой n скалярных равенств

$$\left.\begin{aligned} -\alpha_n a-\alpha_{n-1}a_1-\ldots-\alpha_1 a_{n-1}&=a_n,\\ -\alpha_n b-\alpha_{n-1}b_1-\ldots-\alpha_1 b_{n-1}&=b_n,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ -\alpha_n l-\alpha_{n-1}l_1-\ldots-\alpha_1 l_{n-1}&=l_n. \end{aligned}\right\} \tag{91}$$

[1]) $\psi(\lambda)$ — минимальный многочлен матрицы A.

Из этой системы n линейных уравнений мы можем однозначно определить искомые коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ [1]) и подставить полученные значения в (88). Это исключение $\alpha_1, \alpha_2, \ldots, \alpha_n$ из (88) и (91) можно провести в симметричной форме. Для этого перепишем (88) и (91) так:

$$\left.\begin{aligned} a\alpha_n + a_1\alpha_{n-1} + \ldots + a_{n-1}\alpha_1 + a_n\alpha_0 &= 0, \\ b\alpha_n + b_1\alpha_{n-1} + \ldots + b_{n-1}\alpha_1 + b_n\alpha_0 &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ l\alpha_n + l_1\alpha_{n-1} + \ldots + l_{n-1}\alpha_1 + l_n\alpha_0 &= 0, \\ 1\alpha_n + \lambda\alpha_{n-1} + \ldots + \lambda^{n-1}\alpha_1 + [\lambda^n - \Delta(\lambda)]\alpha_0 &= 0 \end{aligned}\right\} \quad (\alpha_0 = 1).$$

Поскольку эта система из $(n+1)$ уравнений с $(n+1)$ неизвестными $\alpha_0, \alpha_1, \ldots, \alpha_n$ имеет ненулевое решение ($\alpha_0 = 1$), то определитель этой системы должен равняться нулю:

$$\begin{vmatrix} a & a_1 & \ldots & a_{n-1} & a_n \\ b & b_1 & \ldots & b_{n-1} & b_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ l & l_1 & \ldots & l_{n-1} & l_n \\ 1 & \lambda & \ldots & \lambda^{n-1} & \lambda^n - \Delta(\lambda) \end{vmatrix} = 0. \tag{92}$$

Отсюда мы определяем $\Delta(\lambda)$, предварительно транспонируя определитель (92) относительно главной диагонали:

$$M\Delta(\lambda) = \begin{vmatrix} a & b & \ldots & l & 1 \\ a_1 & b_1 & \ldots & l_1 & \lambda \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & b_{n-1} & \ldots & l_{n-1} & \lambda^{n-1} \\ a_n & b_n & \ldots & l_n & \lambda^n \end{vmatrix}, \tag{93}$$

где постоянный множитель M определяется формулой (89) и отличен от нуля.

Тождество (93) и представляет собой преобразование Крылова. В определителе Крылова, стоящем в правой части этого тождества, λ входит только в элементы последнего столбца; остальные же элементы этого определителя от λ не зависят.

Замечание. В регулярном случае все пространство $\boldsymbol{R}$ является циклическим (относительно оператора $\boldsymbol{A}$). Если в качестве базиса выбрать векторы $\boldsymbol{x}$, $\boldsymbol{Ax}, \ldots, \boldsymbol{A}^{n-1}\boldsymbol{x}$, то в этом базисе оператору $\boldsymbol{A}$ соответствует матрица $\widetilde{A}$, имеющая естественную нормальную форму

$$\widetilde{A} = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\alpha_n \\ 1 & 0 & \ldots & 0 & -\alpha_{n-1} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & 0 & -\alpha_2 \\ 0 & \ldots & & 1 & -\alpha_1 \end{Vmatrix}. \tag{94}$$

Переход от основного базиса $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots. \boldsymbol{e}_n$ к базису $\boldsymbol{x}, \boldsymbol{Ax}, \ldots, \boldsymbol{A}^{n-1}\boldsymbol{x}$ осуществляется при помощи неособенной преобразующей матрицы

$$T = \begin{Vmatrix} a & a_1 & \ldots & a_{n-1} \\ b & b_1 & \ldots & b_{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ l & l_1 & \ldots & l_{n-1} \end{Vmatrix}. \tag{95}$$

При этом

$$A = T\widetilde{A}T^{-1}. \tag{96}$$

[1]) Определитель этой системы в силу (89) отличен от нуля.

3. Особый случай: $p<n$. В этом случае векторы x, Ax, ..., $A^{n-1}x$ линейно зависимы, и потому

$$M=\begin{vmatrix} a & b & \dots & l \\ a_1 & b_1 & \dots & l_1 \\ \dots & \dots & \dots & \dots \\ a_{n-1} & b_{n-1} & \dots & l_{n-1} \end{vmatrix}=0.$$

Равенство (93) было выведено при условии $M\neq 0$. Но обе части этого равенства представляют собой целые рациональные функции от λ и параметров a, b, ..., l [1]). Поэтому «из соображений непрерывности» следует, что равенство (93) имеет место и при $M=0$. Но тогда в определителе Крылова (93) после его раскрытия все коэффициенты окажутся равными нулю. Таким образом, в особом случае ($p<n$) формула (93) переходит в тривиальное тождество $0=0$.

Рассмотрим матрицу, составленную из координат векторов x, Ax, ..., A^px

$$\left\|\begin{matrix} a & b & \dots & l \\ a_1 & b_1 & \dots & l_1 \\ \dots & \dots & \dots & \dots \\ a_{p-1} & b_{p-1} & \dots & l_{p-1} \\ a_p & b_p & \dots & l_p \end{matrix}\right\|. \tag{97}$$

Эта матрица имеет ранг p и первые p строк в ней линейно независимы, последняя же $(p+1)$-я строка есть линейная комбинация первых p строк с коэффициентами $-\alpha_p$, $-\alpha_{p-1}$, ..., $-\alpha_1$ [см. (83)]. Из n координат a, b, ..., l мы сможем выбрать такие p координат c, f, ..., h, чтобы определитель, составленный из этих координат векторов x, Ax, ..., $A^{p-1}x$, был отличен от нуля:

$$M^*=\begin{vmatrix} c & f & \dots & h \\ c_1 & f_1 & \dots & h_1 \\ \dots & \dots & \dots & \dots \\ c_{p-1} & f_{p-1} & \dots & h_{p-1} \end{vmatrix}\neq 0. \tag{98}$$

Далее, из (83) вытекает:

$$\left.\begin{aligned} -\alpha_p c-\alpha_{p-1}c_1-\dots-\alpha_1 c_{p-1}&=c_p, \\ -\alpha_p f-\alpha_{p-1}f_1-\dots-\alpha_1 f_{p-1}&=f_p, \\ \dots\dots\dots\dots\dots\dots\dots& \\ -\alpha_p h-\alpha_{p-1}h_1-\dots-\alpha_1 h_{p-1}&=h_p. \end{aligned}\right\} \tag{99}$$

Из этой системы уравнений однозначно определяются коэффициенты α_1, α_2, ..., α_p многочлена $\varphi(\lambda)$ (минимального многочлена вектора x). Совершенно аналогично регулярному случаю (лишь с заменой n на p и букв a, b, ..., l буквами c, f, ..., h) мы сможем исключить α_1, α_2, ..., α_p из (85) и (99) и получить следующую формулу для $\varphi(\lambda)$:

$$M^*\varphi(\lambda)=\begin{vmatrix} c & f & \dots & h & 1 \\ c_1 & f_1 & \dots & h_1 & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ c_{p-1} & f_{p-1} & \dots & h_{p-1} & \lambda^{p-1} \\ c_p & f_p & \dots & h_p & \lambda^p \end{vmatrix}. \tag{100}$$

4. Остановимся на выяснении вопроса, для каких матриц $A=\|a_{ik}\|_1^n$ и при каком выборе исходного вектора x или, что то же, при каком выборе исходных параметров a, b, ..., l имеет место регулярный случай.

---

[1]) $a_i=a_{11}^{(i)}a+a_{12}^{(i)}b+\dots+a_{1n}^{(i)}l$, $b^i=a_{21}^{(i)}a+a_{22}^{(i)}b+\dots+a_{2n}^{(i)}l$ и т. д. ($i=1, 2, \dots, n$), где $a_{jk}^{(i)}$ ($j, k=1, 2, \dots, n$) — элементы матрицы A^i ($i=1, 2, \dots, n$).

Мы уже видели, что в регулярном случае

$$\Delta(\lambda) \equiv \psi(\lambda) \equiv \varphi(\lambda).$$

Совпадение характеристического многочлена $\Delta(\lambda)$ с минимальным многочленом $\psi(\lambda)$ означает, что у матрицы $A = \| a_{ik} \|_1^n$ нет двух элементарных делителей с одним и тем же характеристическим числом, т. е. все элементарные делители попарно взаимно просты. В случае, когда A—матрица простой структуры, это требование равносильно условию, что характеристическое уравнение матрицы A не имеет кратных корней.

Совпадение многочлена $\psi(\lambda)$ с $\varphi(\lambda)$ означает, что в качестве вектора x выбран вектор, порождающий (при помощи оператора A) все пространство R. Такой вектор согласно теореме 2 § 2 всегда существует.

Если же условие $\Delta(\lambda) \equiv \psi(\lambda)$ не выполняется, то, как бы ни выбрать вектор $x \neq 0$, мы многочлена $\Delta(\lambda)$ не получим, так как полученный по методу Крылова многочлен $\varphi(\lambda)$ является делителем $\psi(\lambda)$, который в рассматриваемом случае не совпадает с многочленом $\Delta(\lambda)$, а является лишь его делителем. Варьируя вектор x, мы можем в качестве $\varphi(\lambda)$ получить любой делитель $\psi(\lambda)$ [1]).

Полученные выводы мы можем сформулировать в виде следующей теоремы:

Теорема 14. *Преобразование Крылова дает выражение для характеристического многочлена $\Delta(\lambda)$ матрицы $A = \| a_{ik} \|_1^n$ в виде определителя* (93) *в том и только в том случае, когда выполняются два условия:*

1° *Элементарные делители матрицы A попарно взаимно просты.*

2° *Исходные параметры $a, b, \ldots, l$ являются координатами вектора x, порождающего (при помощи оператора A, соответствующего матрице A) все n-мерное пространство* [2]).

В общем же случае преобразование Крылова приводит к некоторому делителю $\varphi(\lambda)$ характеристического многочлена $\Delta(\lambda)$. Этот делитель $\varphi(\lambda)$ является минимальным многочленом для вектора x с координатами $a, b, \ldots, l$ ($a, b, \ldots, l$— исходные параметры в преобразовании Крылова).

5. Покажем, как найти координаты собственного вектора y для любого характеристического числа λ_0, которое является корнем многочлена $\varphi(\lambda)$, получающегося по методу Крылова [3]).

Вектор $y \neq 0$ будем искать в виде

$$y = \xi_1 x + \xi_2 Ax + \ldots + \xi_p A^{p-1} x. \tag{101}$$

Подставляя это выражение для y в векторное равенство

$$Ay = \lambda_0 y$$

и используя (101), мы получим:

$$\xi_1 Ax + \xi_2 A^2 x + \ldots + \xi_{p-1} A^{p-1} x + \xi_p (-\alpha_p x - \alpha_{p-1} Ax - \ldots - \alpha_1 A^{p-1} x) = \\ = \lambda_0 (\xi_1 x + \xi_2 Ax + \ldots + \xi_p A^{p-1} x). \tag{102}$$

Отсюда, между прочим, следует, что $\xi_p \neq 0$, так как равенство $\xi_p = 0$ в силу (102) давало бы линейную зависимость между векторами $x, Ax, \ldots, A^{p-1}x$. В дальнейшем мы полагаем $\xi_p = 1$. Тогда из (102) получаем:

$$\left.\begin{aligned} &\xi_p = 1,\ \xi_{p-1} = \lambda_0 \xi_p + \alpha_1,\ \xi_{p-2} = \lambda_0 \xi_{p-1} + \alpha_2,\ \ldots,\ \xi_1 = \lambda_0 \xi_2 + \alpha_{p-1}, \\ &0 = \lambda_0 \xi_1 + \alpha_p. \end{aligned}\right\} \tag{103}$$

Первые p из этих равенств определяют нам последовательно величины $\xi_p, \xi_{p-1}, \ldots, \xi_1$ (координаты вектора x в «новом» базисе $x, Ax, \ldots, A^{p-1}x$); последнее же равенство является следствием из предыдущих и из соотношения $\lambda_0^p + \alpha_1 \lambda_0^{p-1} + \ldots + \alpha_p = 0$.

---

[1]) См., например, [81б], стр. 48.

[2]) В аналитическом виде это условие означает, что столбцы $x, Ax, \ldots, A^{n-1}x$, где $x = (a, b, \ldots, l)$, линейно независимы.

[3]) Последующие рассуждения имеют место как для регулярного случая $p = n$, так и для особого случая $p < n$.

Координаты $a', b', \ldots, l'$ вектора y в исходном базисе могут быть найдены по следующим формулам, которые вытекают из (101):

$$\left.\begin{aligned} a' &= \xi_1 a + \xi_2 a_1 + \ldots + \xi_p a_{p-1}, \\ b' &= \xi_1 b + \xi_2 b_1 + \ldots + \xi_p b_{p-1}, \\ &\ldots\ldots\ldots\ldots\ldots \\ l' &= \xi_1 l + \xi_2 l_1 + \ldots + \xi_p l_{p-1}. \end{aligned}\right\} \tag{104}$$

Пример 1.

Рекомендуем читателю следующую схему вычислений.

Под данной матрицей A выписываем строку из координат вектора x: $a, b, \ldots, l$. Этими числами задаемся произвольно (при одном лишь ограничении: по крайней мере одно из этих чисел отлично от нуля). Под строкой $a, b, \ldots, l$ пишем строку $a_1, b_1, \ldots, l_1$, т. е. координаты вектора Ax. Числа $a_1, b_1, \ldots, l_1$ получаются путем последовательного умножения строки $a, b, \ldots, l$ на строки данной матрицы A. Так, например, $a_1 = a_{11}a + a_{12}b + \ldots + a_{1n}l$, $b_1 = a_{21}a + a_{22}b + \ldots + a_{2n}l$ и т. д. Под строкой $a_1, b_1, \ldots, l_1$ пишем строку $a_2, b_2, \ldots, l_2$ и т. д. Каждая из приписываемых строк, начиная со второй, определяется путем последовательного умножения предыдущей строки на строки данной матрицы.

Над строками данной матрицы выписываем контрольную суммарную строку

	8	*3*	*−10*	*−3*		
$A=$	3	−1	−4	2		
	2	3	−2	−4		
	2	−1	−3	2		
	1	2	−1	−3	y	z
$x = e_1 + e_2$	1	1	0	0	−1	1
Ax	2	5	1	3	−1	−1
A^2x	3	5	2	2	1	−1
A^3x	0	9	−1	5	1	1
A^4x	5	9	4	4		
y	0	8	0	4		
	0	2	0	1		
z	−4	0	−4	0		
	1	0	1	0		

В данном случае мы имеем регулярный случай, поскольку

$$M = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 5 & 1 & 3 \\ 3 & 5 & 2 & 2 \\ 0 & 9 & -1 & 5 \end{vmatrix} = -16.$$

Определитель Крылова имеет вид

$$-16\Delta(\lambda) = \begin{vmatrix} 1 & 1 & 0 & 0 & 1 \\ 2 & 5 & 1 & 3 & \lambda \\ 3 & 5 & 2 & 2 & \lambda^2 \\ 0 & 9 & -1 & 5 & \lambda^3 \\ 5 & 9 & 4 & 4 & \lambda^4 \end{vmatrix}.$$

Раскрывая этот определитель и сокращая на -16, найдем:

$$\Delta(\lambda) = \lambda^4 - 2\lambda^2 + 1 = (\lambda - 1)^2 (\lambda + 1)^2.$$

Обозначим через $y = \xi_1 x + \xi_2 Ax + \xi_3 A^2x + \xi_4 A^3x$ собственный вектор матрицы A, соответствующий характеристическому числу $\lambda_0 = 1$. Числа $\xi_1, \xi_2, \xi_3, \xi_4$ найдем

по формулам (103):

$$\xi_4=1,\quad \xi_3=1\cdot\lambda_0+0=1,\quad \xi_2=1\cdot\lambda_0-2=-1,\quad \xi_1=-1\cdot\lambda_0+0=-1.$$

Контрольное равенство $-1\cdot\lambda_0+1=0$, конечно, удовлетворяется.

Полученные числа $\xi_1, \xi_2, \xi_3, \xi_4$ располагаем в вертикальном столбце параллельно столбцу векторов x, Ax, A^2x, A^3x. Помножая столбец $\xi_1, \xi_2, \xi_3, \xi_4$ на столбец a_1, a_2, a_3, a_4, мы получим первую координату a' вектора y в исходном базисе e_1, e_2, e_3, e_4; аналогично получаем b', c', d'. Находим координаты (после сокращения на 4) вектора y: 0, 2, 0, 1. Аналогично определяем координаты 1, 0, 1, 0 собственного вектора z для характеристического числа $\lambda_0=-1$.

Далее, согласно (94) и (95)

$$A=T\widetilde{A}T^{-1},$$

где

$$\widetilde{A}=\begin{Vmatrix}0&0&0&-1\\1&0&0&0\\0&1&0&2\\0&0&1&0\end{Vmatrix},\qquad T=\begin{Vmatrix}1&2&3&0\\1&5&5&9\\0&1&2&-1\\0&3&2&5\end{Vmatrix}.$$

Пример 2. Рассмотрим ту же матрицу A, но в качестве исходных параметров возьмем числа $a=1$, $b=0$, $c=0$, $d=0$.

$$A=\begin{Vmatrix}8&3&-10&-3\\3&-1&-4&2\\2&3&-2&-4\\2&-1&-3&2\\1&2&-1&-3\end{Vmatrix}$$

$x=e_1$	1	0	0	0
Ax	3	2	2	1
A^2x	1	4	0	2
A^3x	3	6	2	3

Но в данном случае

$$M=\begin{vmatrix}1&0&0&0\\3&2&2&1\\1&4&0&2\\3&6&2&3\end{vmatrix}=0$$

и $p=3$. Мы имеем дело с особым случаем.

Беря первые три координаты векторов x, Ax, A^2x, A^3x, определитель Крылова записываем в виде

$$\begin{vmatrix}1&0&0&1\\3&2&2&\lambda\\1&4&0&\lambda^2\\3&6&2&\lambda^3\end{vmatrix}.$$

Раскрывая этот определитель и сокращая на -8, получим:

$$\varphi(\lambda)=\lambda^3-\lambda^2-\lambda+1=(\lambda-1)^2(\lambda+1).$$

Отсюда находим три характеристических числа: $\lambda_1=1$, $\lambda_2=1$, $\lambda_3=-1$. Четвертое характеристическое число получим из условия, что сумма всех характеристических чисел равна следу матрицы. Но $\operatorname{Sp}A=0$. Поэтому $\lambda_4=-1$.

Приведенные примеры показывают, что при применении метода Крылова, выписывая последовательно строки матрицы

$$\begin{Vmatrix}a&b&\dots&l\\a_1&b_1&\dots&l_1\\a_2&b_2&\dots&l_2\\\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\end{Vmatrix},\qquad(105)$$

нужно следить за рангом получаемой матрицы с тем, чтобы остановиться на первой [$(p+1)$-й сверху] строке, которая является линейной комбинацией предыдущих. Определение ранга связано с вычислением известных определителей. Кроме того, получив определитель Крылова в виде (93) или (100), для раскрытия его по элементам последнего столбца следует вычислить известное число определителей $(p-1)$-го порядка [в регулярном случае $(n-1)$-го порядка].

Вместо раскрытия определителя Крылова можно определить коэффициенты $\alpha_1, \alpha_2, \ldots$ непосредственно из системы уравнений (91) [или (99)], применяя к этой системе какой-либо эффективный метод решения, например метод исключения. Этот метод можно применить непосредственно к матрице

$$\left\| \begin{array}{ccccc} a & b & \ldots & l & 1 \\ a_1 & b_1 & \ldots & l_1 & \lambda \\ a_2 & b_2 & \ldots & l_2 & \lambda^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\|, \tag{106}$$

пользуясь им параллельно с получением соответствующих строк по методу Крылова. Тогда мы своевременно обнаружим зависимую от предыдущих строку матрицы (105) без вычисления каких-либо определителей.

Поясним это подробнее. В первой строке матрицы (106) выбираем какой-либо элемент $c \neq 0$ и с его помощью обращаем в нуль стоящий под ним элемент c_1, вычитая из второй строки первую, помноженную на $\frac{c_1}{c}$. Затем во второй строке выбираем какой-либо элемент $f_1^* \neq 0$ и с помощью элементов c и f_1^* обращаем в нуль элементы c_2 и f_2 и т. д. [1]). В результате такого преобразования в последнем столбце матрицы (106) степени λ^k заменятся многочленами k-й степени $g_k(\lambda) = \lambda^k + \ldots$ $(k=0, 1, 2, \ldots)$.

Так как при нашем преобразовании при любом k ранг матрицы, образованной первыми k строками и первыми n столбцами матрицы (106), не меняется, то $(p+1)$-я строка этой матрицы после преобразования будем иметь вид

$$0, 0, \ldots, 0, g_p(\lambda).$$

Проведенное нами преобразование не изменяет величины определителя Крылова

$$\left| \begin{array}{ccccc} c & f & \ldots & h & 1 \\ c_1 & f_1 & \ldots & h_1 & \lambda \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{p-1} & f_{p-1} & \ldots & h_{p-1} & \lambda^{p-1} \\ c_p & f_p & \ldots & h_p & \lambda^p \end{array} \right| = M^* \varphi(\lambda).$$

Поэтому

$$M^* \varphi(\lambda) = c f_1^* \ldots g_p(\lambda), \tag{107}$$

т. е. [2]) $g_p(\lambda)$ и будем искомым многочленом $\varphi(\lambda)$, $g_p(\lambda) \equiv \varphi(\lambda)$.

Рекомендуем следующее упрощение. Получив k-ю преобразованную строку в матрице (106)

$$a_{k-1}^*, b_{k-1}^*, \ldots, l_{k-1}^*, g_{k-1}(\lambda), \tag{108}$$

следующую $(k+1)$-ю строку следует получать, умножая ряд $a_{k-1}^*, b_{k-1}^*, \ldots, l_{k-1}^*$ (а не первоначальный ряд $a_{k-1}, b_{k-1}, \ldots, l_{k-1}$) на строки данной матрицы [3]). Тогда мы найдем $(k+1)$-ю строку в виде

$$\tilde{a}_k, \tilde{b}_k \ldots l_k \tilde{\lambda} g_{k-1}(\lambda)$$

---

[1]) Элементы $c_1, f_1^*, \ldots$ не должны принадлежать последнему столбцу, содержащему степени λ.

[2]) Напоминаем, что старшие коэффициенты многочленов $\varphi(\lambda)$ и $g_p(\lambda)$ равны единице.

[3]) Упрощение заключается в том, что в преобразованной строке (108) $k-1$ элементов равны нулю. Поэтому такую строку проще умножать на строки матрицы A.

и после вычитания предыдущих строк получим:

$$a_k^*,\ b_k^*,\ \ldots,\ l_k^*,\ g_k(\lambda).$$

Рекомендуемое нами небольшое видоизменение метода Крылова (соединение его с методом исключения) позволяет сразу получить интересующий нас многочлен $\varphi(\lambda)$ [в регулярном случае $\Delta(\lambda)$] без вычисления каких-либо определителей и решения вспомогательной системы уравнений [1]).

Пример.

$$A=\begin{array}{|ccccc|l}
\multicolumn{5}{c}{\begin{matrix}4 & 4 & 1 & 5 & 0\end{matrix}} & \\
1 & 1 & -1 & 1 & 0 & \\
1 & 2 & -1 & 0 & 1 & \\
-1 & 2 & 3 & -1 & 0 & \\
1 & -2 & 1 & 2 & -1 & \\
2 & 1 & -1 & 3 & 0 & \\
\hline
0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & -1 & 0 & \lambda \\
0 & 2 & 3 & -4 & -2 & \lambda^2 \quad [2-4\lambda] \\
0 & -2 & 3 & 0 & 0 & \lambda^2-4\lambda+2 \\
-5 & -7 & 5 & 7 & -5 & \lambda^3-4\lambda^2+2\lambda \quad [5+7\lambda] \\
-5 & 0 & 5 & 0 & 0 & \lambda^3-4\lambda^2+9\lambda+5 \\
-10 & -10 & 20 & 0 & -15 & \lambda^4-4\lambda^3+9\lambda^2+5\lambda \quad [15-5(\lambda^2-4\lambda+2)- \\
 & & & & & \quad -2(\lambda^3-4\lambda^2+9\lambda+5)] \\
0 & 0 & -5 & 0 & 0 & \lambda^4-6\lambda^3+12\lambda^2+7\lambda-5 \\
5 & 5 & -15 & -5 & 0 & \lambda^5-6\lambda^4+12\lambda^3+7\lambda^2-5\lambda \quad [-5-5\lambda+(\lambda^3- \\
 & & & & & \quad -4\lambda^2+9\lambda+5)-2(\lambda^4-6\lambda^3+12\lambda^2+7\lambda-5)] \\
0 & 0 & 0 & 0 & 0 & \underbrace{\lambda^5-8\lambda^4+25\lambda^3-21\lambda^2-15\lambda+10}_{\Delta(\lambda)}
\end{array}$$

[1]) Наряду с методом акад. А. Н. Крылова мы познакомили читателя в гл. IV с методом Д. К. Фаддеева вычисления коэффициентов характеристического многочлена. Метод Д. К. Фаддеева связан с бо́льшими вычислениями, нежели метод Крылова. но метод Фаддеева является более общим, в нем нет особых случаев. Обращаем внимание читателя еще на весьма эффективный метод А. М. Данилевского [88]; см. также [8], стр. 235—239, обзорную статью [79] и книгу [33], § 24.

ГЛАВА VIII

МАТРИЧНЫЕ УРАВНЕНИЯ

В этой главе мы рассмотрим некоторые типы матричных уравнений, встречающиеся в разнообразных вопросах теории матриц и ее приложений.

§ 1. Уравнение $AX = XB$

Пусть дано уравнение

$$AX = XB, \tag{1}$$

где A и B — две заданные квадратные матрицы (вообще говоря, разных порядков)

$$A = \| a_{ij} \|_1^m, \quad B = \| b_{kl} \|_1^n,$$

а X — искомая прямоугольная матрица размером $m \times n$:

$$X = \| x_{jk} \| \qquad (j = 1, 2, \ldots, m; \ k = 1, 2, \ldots, n).$$

Выпишем элементарные делители матриц A и B (в поле комплексных чисел):

$$(A): (\lambda - \lambda_1)^{p_1}, \ (\lambda - \lambda_2)^{p_2}, \ \ldots, \ (\lambda - \lambda_u)^{p_u} \quad (p_1 + p_2 + \ldots + p_u = m),$$

$$(B): (\lambda - \mu_1)^{q_1}, \ (\lambda - \mu_2)^{q_2}, \ \ldots, \ (\lambda - \mu_v)^{q_v} \quad (q_1 + q_2 + \ldots + q_v = n).$$

В соответствии с этими элементарными делителями приведем матрицы A и B к нормальной жордановой форме

$$A = U\widetilde{A}U^{-1}, \qquad B = V\widetilde{B}V^{-1}, \tag{2}$$

где U и V — квадратные неособенные матрицы соответственно порядка m и n, а $\widetilde{A}$ и $\widetilde{B}$ — жордановы матрицы:

$$\begin{aligned} \widetilde{A} &= \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ \lambda_2 E^{(p_2)} + H^{(p_2)}, \ \ldots, \ \lambda_u E^{(p_u)} + H^{(p_u)}\}, \\ \widetilde{B} &= \{\mu_1 E^{(q_1)} + H^{(q_1)}, \ \mu_2 E^{(q_2)} + H^{(q_2)}, \ \ldots, \ \mu_v E^{(q_v)} + H^{(q_v)}\}. \end{aligned} \tag{3}$$

Подставляя в уравнение (1) вместо A и B их выражения (2), получим:

$$U\widetilde{A}U^{-1}X = XV\widetilde{B}V^{-1}.$$

Умножим обе части этого равенства слева на U^{-1}, а справа — на V:

$$AU^{-1}XV = U^{-1}XV\widetilde{B}. \tag{4}$$

Вводя вместо искомой матрицы X новую искомую матрицу $\widetilde{X}$ (тех же размеров $m \times n$)

$$\widetilde{X} = U^{-1}XV, \tag{5}$$

мы уравнение (4) запишем так:

$$\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B}. \tag{6}$$

Мы заменили матричное уравнение (1) уравнением (6) того же вида, но в котором заданные матрицы имеют нормальную жорданову форму.

В соответствии с квазидиагональным видом матриц $\widetilde{A}$ и $\widetilde{B}$ разобьем матрицу $\widetilde{X}$ на блоки:

$$\widetilde{X} = (X_{\alpha\beta}) \qquad (\alpha = 1,\ 2,\ \ldots,\ u;\quad \beta = 1,\ 2,\ \ldots,\ v)$$

(здесь $X_{\alpha\beta}$ — прямоугольная матрица размером $p_\alpha \times q_\beta$; $\alpha = 1,\ 2,\ \ldots,\ u$; $\beta = 1,\ 2,\ \ldots,\ v$).

Используя правило умножения блочной матрицы на квазидиагональную (см. стр. 56), произведем умножение матриц в левой и правой частях уравнения (6). Тогда это уравнение распадается на uv матричных уравнений

$$[\lambda_\alpha E^{(p_\alpha)} + H^{(p_\alpha)}]\, X_{\alpha\beta} = X_{\alpha\beta}\, [\mu_\beta E^{(q_\beta)} + H^{(q_\beta)}]$$
$$(\alpha = 1,\ 2,\ \ldots,\ u;\quad \beta = 1,\ 2,\ \ldots,\ v),$$

которые перепишем еще так:

$$(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta} = H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta \qquad (\alpha = 1,\ 2,\ \ldots,\ u;\ \beta = 1,\ 2,\ \ldots,\ v); \tag{7}$$

при этом мы ввели сокращенные обозначения

$$H_\alpha = H^{(p_\alpha)},\quad G_\beta = H^{(q_\beta)} \qquad (\alpha = 1,\ 2,\ \ldots,\ u;\quad \beta = 1,\ 2,\ \ldots,\ v). \tag{8}$$

Возьмем какое-нибудь из уравнений (7). Могут представиться два случая:

1. $\lambda_\alpha \neq \mu_\beta$. Проитерируем $r-1$ раз равенство (7)[1]):

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{\sigma+\tau=r} (-1)^\tau H_\alpha^\sigma X_{\alpha\beta} G_\beta^\tau. \tag{9}$$

Заметим, что в силу (8)

$$H_\alpha^{p_\alpha} = G_\beta^{q_\beta} = 0. \tag{10}$$

Если в (9) взять $r \geqslant p_\alpha + q_\beta - 1$, то в каждом члене суммы, стоящей в правой части равенства (9), выполняется по крайней мере одно из соотношений

$$\sigma \geqslant p_\alpha,\quad \tau \geqslant q_\beta,$$

и потому в силу (10) либо $H_\alpha^\sigma = 0$, либо $G_\beta^\tau = 0$. Так как, кроме того, в рассматриваемом случае $\lambda_\alpha \neq \mu_\beta$, то из (9) находим:

$$X_{\alpha\beta} = 0. \tag{11}$$

2. $\lambda_\alpha = \mu_\beta$. В этом случае уравнение (7) принимает вид

$$H_\alpha X_{\alpha\beta} = X_{\alpha\beta} G_\beta. \tag{12}$$

---

[1]) Обе части равенства (7) умножаем на $\mu_\beta - \lambda_\alpha$ и в каждом члене правой части заменяем $(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta}$ на $H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta$. Этот процесс повторяем $r-1$ раз.

В матрицах H_α и G_β элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Учитывая эту специфичную структуру матриц H_α и G_β и полагая

$$[X_{\alpha\beta} = \| \xi_{ik} \| \qquad (i = 1, 2, \ldots, p_\alpha;\ k = 1, 2, \ldots, q_\beta),$$

мы заменим матричное уравнение (12) следующей эквивалентной ему системой скалярных соотношений [1]):

$$\xi_{i+1,\,k} = \xi_{i,\,k-1}\ (\xi_{i0} = \xi_{p_\alpha+1,\,k} = 0;\ i = 1, 2, \ldots, p_\alpha;\ k = 1, 2, \ldots, q_\beta). \quad (13)$$

Равенства (13) означают:

1) В матрице $X_{\alpha\beta}$ на каждой линии, параллельной главной диагонали, стоят равные между собой элементы,

2) $\xi_{21} = \xi_{31} = \ldots = \xi_{p_\alpha 1} = \xi_{p_\alpha 2} = \ldots = \xi_{p_\alpha, q_\beta - 1} = 0$.

Пусть $p_\alpha = q_\beta$. В этом случае $X_{\alpha\beta}$ — квадратная матрица. Из 1), 2) следует, что в матрице $X_{\alpha\beta}$ все элементы, расположенные под главной диагональю, равны нулю, все элементы главной диагонали равны некоторому числу $c_{\alpha\beta}$, все элементы первой наддиагонали равны некоторому числу $c'_{\alpha\beta}$ и т. д., т. е.

$$X_{\alpha\beta} = \begin{Vmatrix} c_{\alpha\beta} & c'_{\alpha\beta} & & \ldots c_{\alpha\beta}^{(p_\alpha-1)} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ & \ddots & \ddots & c'_{\alpha\beta} \\ 0 & \ldots & 0 & c_{\alpha\beta} \end{Vmatrix} = T_{p_\alpha}; \quad (14)$$

$$(p_\alpha = q_\beta)$$

здесь $c_{\alpha\beta}$, $c'_{\alpha\beta}$, $\ldots$, $c_{\alpha\beta}^{(p_\alpha-1)}$ — произвольные параметры (уравнения (12) не накладывают никаких ограничений на значения этих параметров).

Легко видеть, что при $p_\alpha < q_\beta$

$$X_{\alpha\beta} = (\overbrace{0}^{q_\beta - p_\alpha}, T_{p_\alpha}), \quad (15)$$

а при $p_\alpha > q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} T_{\alpha\beta} \\ 0 \end{pmatrix}_{\}p_\alpha - q_\beta}. \quad (16)$$

Про матрицы (14), (15) и (16) мы будем говорить, что они имеют *правильную верхнюю треугольную форму*. Число произвольных параметров в $X_{\alpha\beta}$ равно наименьшему из чисел p_α и q_β. Приведенная ниже схема показывает структуру матрицы $X_{\alpha\beta}$ при $\lambda_\alpha = \mu_\beta$ (произвольные параметры

---

[1]) Из структуры матриц H_α и G_β следует, что произведение $H_\alpha X_{\alpha\beta}$ получается из $X_{\alpha\beta}$ сдвигом всех строк на одно место вверх и заполнением последней строки нулями; аналогично $X_{\alpha\beta}G_\beta$ получается из $X_{\alpha\beta}$ сдвигом всех столбцов на одно место вправо и заполнением первого столбца нулями (см. гл. I, стр. 25).

Для сокращения обозначений мы не пишем при ξ_{ik} дополнительных индексов α, β.

здесь обозначены через a, b, c, d):

$$X_{\alpha\beta}=\begin{Vmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{Vmatrix}, \qquad X_{\alpha\beta}=\begin{Vmatrix} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{Vmatrix}, \qquad X_{\alpha\beta}=\begin{Vmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}.$$

$(p_\alpha=q_\beta=4)$ $\qquad$ $(p_\alpha=3,\ q_\beta=5)$ $\qquad$ $(p_\alpha=5,\ q_\beta=3)$

Для того чтобы при подсчете произвольных параметров в матрице $\widetilde{X}$ охватить и случай 1, обозначим через $d_{\alpha\beta}(\lambda)$ наибольший общий делитель элементарных делителей $(\lambda-\lambda_\alpha)^{p_\alpha}$ и $(\lambda-\mu_\beta)^{q_\beta}$, а через $\delta_{\alpha\beta}$ — степень многочлена $d_{\alpha\beta}(\lambda)$ ($\alpha=1,\ 2,\ \ldots,\ u$; $\beta=1,\ 2,\ \ldots,\ v$). В случае 1 $\delta_{\alpha\beta}=0$; в случае 2 имеем: $\delta_{\alpha\beta}=\min(p_\alpha,\ q_\beta)$. Таким образом, в обоих случаях число произвольных параметров в $X_{\alpha\beta}$ равно $\delta_{\alpha\beta}$. Число произвольных параметров в $\widetilde{X}$ определяется формулой

$$N=\sum_{\alpha=1}^{u}\sum_{\beta=1}^{v}\delta_{\alpha\beta}.$$

В дальнейшем нам удобно будет общее решение уравнения (6) обозначить через $X_{\widetilde{A}\widetilde{B}}$ (до сих пор мы это решение обозначали буквой $\widetilde{X}$).

Полученные в этом параграфе результаты можно сформулировать в виде следующей теоремы:

Теорема 1. *Общее решение матричного уравнения*

$$AX=XB,$$

где

$$A=\|a_{ik}\|_1^m=U\widetilde{A}U^{-1}=U\{\lambda_1E^{(p_1)}+H^{(p_1)},\ \ldots,\ \lambda_uE^{(p_u)}+H^{p_u}\}U^{-1},$$

$$B=\|b_{ik}\|_1^n=V\widetilde{B}V^{-1}=V\{\mu_1E^{(q_1)}+H^{(q_1)},\ \ldots,\ \mu_vE^{(q_v)}+H^{(q_v)}\}V^{-1},$$

задается формулой

$$X=UX_{\widetilde{A}\widetilde{B}}V^{-1}. \tag{17}$$

Здесь $X_{\widetilde{A}\widetilde{B}}$ — общее решение уравнения $\widetilde{A}\widetilde{X}=\widetilde{X}\widetilde{B}$ — имеет следующую структуру: $X_{\widetilde{A}\widetilde{B}}$ разбивается на блоки

$$X_{\widetilde{A}\widetilde{B}}=(\overset{q_\beta}{\widetilde{X}_{\alpha\beta}})\}\,p_\alpha \qquad (\alpha=1,\ 2,\ \ldots,\ u;\ \beta=1,\ 2,\ \ldots,\ v);$$

если $\lambda_\alpha\neq\mu_\beta$, то на месте $X_{\alpha\beta}$ стоит нулевая матрица, если же $\lambda_\alpha=\mu_\beta$, то на месте $X_{\alpha\beta}$ стоит произвольная правильная верхняя треугольная матрица.

$X_{\widetilde{A}\widetilde{B}}$, а следовательно, и X зависят линейно от N произвольных параметров $c_1,\ c_2,\ \ldots,\ c_N$:

$$X=\sum_{i=1}^{N}c_jX_j, \tag{18}$$

где N определяется формулой

$$N=\sum_{\alpha=1}^{u}\sum_{\beta=1}^{v}\delta_{\alpha\beta} \tag{19}$$

[*здесь* $\delta_{\alpha\beta}$ *обозначает степень наибольшего общего делителя* $(\lambda - \lambda_\alpha)^{p_\alpha}$ *и* $(\lambda - \mu_\beta)^{q_\beta}$].

Заметим, что матрицы $X_1, X_2, \ldots, X_N$, фигурирующие в формуле (18), суть решения исходного уравнения (1) (матрица X_j получается из X, если параметру c_j дать значение единицы, а остальным параметрам — нулевые значения; $j = 1, 2, \ldots, N$). Эти решения линейно независимы, так как в противном случае при некоторых значениях параметров $c_1, c_2, \ldots, c_N$, не равных одновременно нулю, матрица X, а следовательно, и $X_{\widetilde{AB}}$ равнялись бы нулю, что невозможно. Таким образом, равенство (19) показывает, что любое решение исходного уравнения представляет собой линейную комбинацию N линейно независимых решений.

Если матрицы A *и* B *не имеют общих характеристических чисел* (характеристические многочлены $|\lambda E - A|$ и $|\lambda E - B|$ взаимно просты), *то* $N = \sum_{\alpha=1}^{u} \sum_{\beta=1}^{v} \delta_{\alpha\beta} = 0$ и, следовательно, $X = 0$, т. е. *в этом случае уравнение* (1) *имеет только тривиальное нулевое решение* $X = 0$.

Замечание. Пусть элементы матриц A и B принадлежат некоторому числовому полю K. Тогда нельзя утверждать, что элементы матриц U, V, $X_{\widetilde{AB}}$, фигурирующих в формуле (17), также принадлежат полю K. Элементы этих матриц можно выбрать в расширенном поле K_1, которое получается из поля K путем приобщения к последнему корней характеристических уравнений $|\lambda E - A| = 0$ и $|\lambda E - B| = 0$. С такого рода расширением основного поля всегда приходится иметь дело, когда пользуются приведением заданных матриц к нормальной жордановой форме.

Однако матричное уравнение (1) эквивалентно системе mn линейных однородных уравнений, где неизвестными служат элементы x_{jk} $(j = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$ искомой матрицы X:

$$\sum_{j=1}^{m} a_{ij} x_{jk} = \sum_{h=1}^{n} x_{ih} b_{hk} \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n). \tag{20}$$

Нами доказано, что эта система имеет N линейно независимых решений, где N определяется формулой (19). Но известно, что базисные линейно независимые решения можно выбрать в основном поле K, которому принадлежат коэффициенты уравнений (20). Таким образом, в формуле (18) матрицы $X_1, X_2, \ldots, X_N$ можно выбрать так, чтобы их элементы принадлежали полю K. Тогда, придавая в формуле (18) произвольным параметрам всевозможные значения из поля K, мы получим все матрицы X с элементами из K, удовлетворяющие уравнению (1)[1]).

§ 2. Частный случай: $A = B$. Перестановочные матрицы

Рассмотрим частный случай уравнения (1) — уравнение

$$AX = XA, \tag{21}$$

где $A = \|a_{ik}\|_1^n$ — заданная, а $X = \|x_{ik}\|_1^n$ — искомая матрица. Мы пришли к задаче Фробениуса: определить все матрицы X, перестановочные с данной матрицей A.

---

[1]) Матрицы $A = \|a_{ij}\|_1^m$ и $B = \|b_{kl}\|_1^n$ определяют линейный оператор $\widehat{F}(X) = AX - XB$ в пространстве прямоугольных матриц X с размерами $m \times n$. Исследование операторов такого типа проведено в работах [85а, б].

Приведем матрицу A к нормальной жордановой форме:

$$A = U\widetilde{A}U^{-1} = U\{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ldots, \lambda_u E^{(p_u)} + H^{(p_u)}\}U^{-1}. \qquad (22)$$

Тогда, полагая в формуле (17) $V = U$, $\widetilde{B} = \widetilde{A}$ и обозначая $X_{\widetilde{A}\widetilde{A}}$ сокращенно через $X_{\widetilde{A}}$, получим все решения уравнения (21), т. е. все матрицы, перестановочные с A, в следующем виде:

$$X = UX_{\widetilde{A}}U^{-1}, \qquad (23)$$

где $X_{\widetilde{A}}$ обозначает произвольную матрицу, перестановочную с $\widetilde{A}$. Как было выяснено в предыдущем параграфе, $X_{\widetilde{A}}$ разбивается на u^2 блоков

$$X_{\widetilde{A}} = (X_{\alpha\beta})_1^u$$

в соответствии с разбиением на блоки жордановой матрицы $\widetilde{A}$; $X_{\alpha\beta}$ — нулевая матрица либо произвольная правильная верхняя треугольная матрица в зависимости от того, $\lambda_\alpha \neq \lambda_\beta$ или $\lambda_\alpha = \lambda_\beta$.

Для примера выпишем элементы матрицы $X_{\widetilde{A}}$ для случая, когда матрица A имеет следующие элементарные делители:

$$(\lambda-\lambda_1)^4, \quad (\lambda-\lambda_1)^3, \quad (\lambda-\lambda_2)^2, \quad \lambda-\lambda_2 \qquad (\lambda_1 \neq \lambda_2).$$

В этом случае $X_{\widetilde{A}}$ имеет такой вид:

$$\left\|\begin{array}{ccccccccc}
a & b & c & d & e & f & g & 0 & 0 & 0 \\
0 & a & b & c & 0 & e & f & 0 & 0 & 0 \\
0 & 0 & a & b & 0 & 0 & e & 0 & 0 & 0 \\
0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & h & k & l & m & p & q & 0 & 0 & 0 \\
0 & 0 & h & k & 0 & m & p & 0 & 0 & 0 \\
0 & 0 & 0 & h & 0 & 0 & m & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & r & s & t \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & z
\end{array}\right\|$$

($a, b, \ldots, z$ — произвольные параметры).

Число параметров в $X_{\widetilde{A}}$ равно N, где $N = \sum_{\alpha,\beta=1}^{u} \delta_{\alpha\beta}$; здесь $\delta_{\alpha\beta}$ обозначает степень наибольшего общего делителя многочленов $(\lambda-\lambda_\alpha)^{p_\alpha}$ и $(\lambda-\lambda_\beta)^{p_\beta}$.

Введем в рассмотрение инвариантные многочлены матрицы A: $i_1(\lambda)$, $i_2(\lambda)$, $\ldots$, $i_t(\lambda)$; $i_{t+1}(\lambda) = \ldots = i_n(\lambda) = 1$. Степени этих многочленов обозначим через $n_1 \geqslant n_2 \geqslant \ldots \geqslant n_t > n_{t+1} = \ldots = 0$. Так как каждый инвариантный многочлен является произведением нескольких попарно взаимно простых элементарных делителей, то формулу для N можно записать и так:

$$N = \sum_{g,j=1}^{t} \varkappa_{gj}, \qquad (24)$$

где $\varkappa_{gj}$ — степень наибольшего общего делителя многочленов $i_g(\lambda)$ и $i_j(\lambda)$

$(g, j = 1, 2, \ldots, t)$. Но наибольшим общим делителем многочленов $i_g(\lambda)$ и $i_j(\lambda)$ является один из этих же многочленов и потому $\varkappa_{gj} = \min(n_g, n_j)$. Отсюда получаем:

$$N = n_1 + 3n_2 + \ldots + (2t-1)\,n_t.$$

Число N является числом линейно независимых матриц, перестановочных с матрицей A (можно считать, что элементы этих матриц принадлежат основному полю K, содержащему элементы матрицы A; см. замечание в конце предыдущего параграфа). Мы пришли к теореме:

Теорема 2. *Число линейно независимых матриц, перестановочных с матрицей* $A = \| a_{ik} \|_1^n$, *определяется формулой*

$$N = n_1 + 3n_2 + \ldots + (2t-1)\,n_t, \tag{25}$$

где $n_1, n_2, \ldots, n_t$ — *степени непостоянных инвариантных многочленов* $i_1(\lambda), i_2(\lambda), \ldots, i_t(\lambda)$ *матрицы* A.

Заметим, что

$$n = n_1 + n_2 + \ldots + n_t. \tag{26}$$

Из (25) и (26) вытекает:

$$N \geqslant n, \tag{27}$$

причем знак $=$ имеет место в том и только в том случае, когда $t = 1$, т. е. когда все элементарные делители матрицы A попарно взаимно просты.

Пусть $g(\lambda)$ — некоторый многочлен от λ. Тогда матрица $g(A)$ перестановочна с A. Возникает обратный вопрос: в каком случае любая матрица, перестановочная с A, может быть представлена как многочлен от A? В этом случае любая матрица, перестановочная с A, была бы линейной комбинацией линейно независимых матриц

$$E, A, A^2, \ldots, A^{n_1-1}.$$

В рассматриваемом случае $N = n_1 \leqslant n$; сопоставляя с (27), получаем: $N = n_1 = n$. Таким образом, мы получили

Следствие 1 из теоремы 2. *Все матрицы, перестановочные с* A, *представляются как многочлены от* A *в том и только в том случае, когда* $n_1 = n$, *т. е. когда все элементарные делители матрицы* A *попарно взаимно просты.*

Многочлены от матрицы, перестановочной с A, также перестановочны с A. Поставим вопрос: в каком случае все матрицы, перестановочные с A, представляются в виде многочленов от некоторой (одной и той же) матрицы C? Допустим, что такой случай имеет место. Тогда, так как матрица C удовлетворяет в силу теоремы Гамильтона — Кэли своему характеристическому уравнению, то любая матрица, перестановочная с C, выразится линейно через матрицы

$$E, C, C^2, \ldots, C^{n-1}.$$

Поэтому в рассматриваемом случае $N \leqslant n$. Сопоставляя с (27), находим $N = n$. Но тогда из (25) и (26) и $n_1 = n$.

Следствие 2 из теоремы 2. *Все матрицы, перестановочные с* A, *представляются в виде многочленов от одной и той же матрицы* C *в том и только в том случае, когда* $n_1 = n$, *т. е. когда все элементарные делители у* $\lambda E - A$ *взаимно просты. В этом случае все матрицы, перестановочные с* A, *представляются и в виде многочленов от* A.

Отметим еще одно очень важное свойство перестановочных матриц.

Теорема 3. *Если две матрицы* $A=\|a_{ik}\|_1^n$, $B=\|b_{ik}\|_1^n$ *перестановочны и одна из них, например* A, *имеет квазидиагональный вид*

$$A=\{\overbrace{A_1}^{s_1},\ \overbrace{A_2}^{s_2}\},\tag{28}$$

где матрицы A_1 *и* A_2 *не имеют общих характеристических чисел, то и другая матрица имеет такой же квазидиагональный вид*

$$B=\{\overbrace{B_1}^{s_1},\ \overbrace{B_2}^{s_2}\}.\tag{29}$$

Доказательство. Разобьем матрицу B на блоки в соответствии с квазидиагональным видом (28):

$$B=\begin{pmatrix}\overbrace{B_1}^{s_1} & \overbrace{X}^{s_2}\\ Y & B_2\end{pmatrix}.$$

Записывая, что $AB=BA$, получим четыре матричных равенства:

$$1.\ A_1B_1=B_1A_1,\quad 2.\ A_1X=XA_2,\quad 3.\ A_2Y=YA_1,\quad 4.\ A_2B_2=B_2A_2.\tag{30}$$

Второе и третье из уравнений (30), как было выяснено в § 1 (стр. 203), имеют только нулевые решения $X=0$, $Y=0$, поскольку матрицы A_1 и A_2 не имеют общих характеристических чисел. Таким образом, наше предложение доказано. Первое и четвертое из равенств (30) выражают перестановочность матриц A_1 и B_1, A_2 и B_2.

Доказанное предложение в геометрической формулировке гласит:

Теорема 3′. *Если*

$$\boldsymbol{R}=\boldsymbol{I}_1+\boldsymbol{I}_2$$

есть расщепление всего пространства $\boldsymbol{R}$ *на инвариантные относительно оператора* $\boldsymbol{A}$ *подпространства* $\boldsymbol{I}_1$ *и* $\boldsymbol{I}_2$ *и минимальные многочлены этих подпространств (относительно* $\boldsymbol{A}$*) взаимно просты, то эти подпространства* $\boldsymbol{I}_1$ *и* $\boldsymbol{I}_2$ *инвариантны относительно любого линейного оператора* $\boldsymbol{B}$, *перестановочного с* $\boldsymbol{A}$.

Из доказанной теоремы вытекает следующее следствие[1]):

Следствие 1. *Если линейные операторы* $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *попарно перестановочны, то можно расщепить все пространство* $\boldsymbol{R}$ *на инвариантные относительно всех операторов* $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *подпространства*

$$\boldsymbol{R}=\boldsymbol{I}_1+\boldsymbol{I}_2+\ldots+\boldsymbol{I}_w$$

так, чтобы минимальный многочлен любого из этих подпространств относительно любого из операторов $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *был степенью неприводимого многочлена.*

Как частный случай отсюда получим

Следствие 2. *Если линейные операторы* $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *попарно перестановочны и все характеристические числа этих операторов принадлежат основному полю* K, *то можно расщепить все пространство* $\boldsymbol{R}$ *на инвариантные относительно всех операторов* $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *подпространства* $\boldsymbol{I}_1$, $\boldsymbol{I}_2$, ..., $\boldsymbol{I}_w$, *в каждом из которых любой из операторов* $\boldsymbol{A}$, $\boldsymbol{B}$, ..., $\boldsymbol{L}$ *имеет равные характеристические числа.*

---

[1]) См. также теорему 1 гл. VII (стр. 173).

И, наконец, отметим уже частный случай этого предложения:

Следствие 3. *Если операторы простой структуры* $\boldsymbol{A}, \boldsymbol{B}, \ldots, \boldsymbol{L}$ (см. гл. III, § 8) *попарно перестановочны, то можно составить базис пространства из общих собственных векторов этих операторов.*

Дадим еще матричную формулировку последнего предложения:

Перестановочные матрицы простой структуры можно одновременно, т. е. одним и тем же преобразованием подобия, привести к диагональному виду.

§ 3. Уравнение $AX-XB=C$

Пусть дано матричное уравнение

$$AX-XB=C, \tag{31}$$

где $A=\|a_{ij}\|_1^m$, $B=\|b_{kl}\|_1^n$ — заданные квадратные матрицы порядков m и n, $C=\|c_{jk}\|$ — заданная, а $X=\|x_{jk}\|$ — искомая прямоугольные матрицы размером $m\times n$. Уравнение (31) эквивалентно системе mn скалярных уравнений относительно элементов матрицы X:

$$\sum_{j=1}^{m} a_{ij}x_{jk}-\sum_{l=1}^{n} x_{il}b_{lk}=c_{ik} \qquad (i=1,\ 2,\ \ldots,\ m;\ k=1,\ 2,\ \ldots,\ n).$$

Соответствующая однородная система уравнений

$$\sum_{j=1}^{m} a_{ij}x_{jk}-\sum_{l=1}^{n} x_{il}b_{lk}=0 \qquad (i=1,\ 2,\ \ldots,\ m;\ k=1,\ 2,\ \ldots,\ n)$$

в матричном виде записывается так:

$$AX-XB=0. \tag{32}$$

Таким образом, если уравнение (32) имеет только нулевое решение, то уравнение (31) имеет одно-единственное решение. Но в § 1 было установлено, что уравнение (32) имеет только нулевое решение тогда и только тогда, когда матрицы A и B не имеют общих характеристических чисел. Следовательно, *если матрицы* A и B *не имеют общих характеристических чисел, то уравнение* (31) *имеет одно и только одно решение; если же матрицы* A и B *имеют общие характеристические числа, то в зависимости от «свободного члена»* C *могут представиться два случая: либо уравнение* (31) *противоречиво, либо оно имеет бесчисленное множество решений, задаваемых формулой*

$$X=X_0+X_1,$$

где X_0 — фиксированное частное решение уравнения (31), X_1 — общее решение однородного уравнения (32) (структура X_1 была выяснена в § 1).

§ 4. Скалярное уравнение $f(X)=0$

Рассмотрим сначала уравнение

$$g(X)=0, \tag{33}$$

где

$$g(\lambda)=(\lambda-\lambda_1)^{a_1}(\lambda-\lambda_2)^{a_2}\ldots(\lambda-\lambda_h)^{a_h}$$

— заданный многочлен переменной λ, а X — искомая квадратная матрица порядка n. Так как минимальный многочлен матрицы X, т. е. первый

инвариантный многочлен $i_1(\lambda)$, должен быть делителем многочлена $g(\lambda)$, то элементарные делители матрицы X должны иметь следующий вид:

$$(\lambda-\lambda_{i_1})^{p_{i_1}},\ (\lambda-\lambda_{i_2})^{p_{i_2}},\ \ldots,\ (\lambda-\lambda_{i_\nu})^{p_{i_\nu}}\left(\begin{array}{l} i_1,\ i_2,\ \ldots,\ i_\nu=1,\ 2,\ \ldots,\ h,\\ p_{i_1}\leqslant a_{i_1},\ p_{i_2}\leqslant a_{i_2},\ \ldots,\ p_{i_h}\leqslant a_{i_h},\\ p_{i_1}+p_{i_2}+\ldots+p_{i_\nu}=n \end{array}\right)$$

(среди индексов $i_1, i_2, \ldots, i_\nu$ могут быть и равные, n — заданный порядок искомой матрицы X).

Искомая матрица X представится в виде

$$X=T\{\lambda_{i_1}E^{(p_{i_1})}+H^{(p_{i_1})},\ \ldots,\ \lambda_{i_\nu}E^{(p_{i_\nu})}+H^{(p_{i_\nu})}\}\,T^{-1}, \tag{34}$$

где T — произвольная неособенная матрица порядка n. Множество решений уравнения (33) с заданным порядком искомой матрицы распадается согласно формуле (34) на конечное число классов подобных между собой матриц.

Пример 1. Дано уравнение

$$X^m=0. \tag{35}$$

Если некоторая степень матрицы равна нулю, то матрица называется *нильпотентной*. Наименьший из показателей, при которых степень матрицы равна нулю, называется *индексом нильпотентности* данной матрицы.

Очевидно, решениями уравнения (35) являются все нильпотентные матрицы с индексом нильпотентности $\mu\leqslant m$. Формула, охватывающая все решения данного порядка n, выглядит так:

$$X=T\{H^{(p_1)},\ H^{(p_2)},\ \ldots,\ H^{(p_\nu)}\}\,T^{-1}\left(\begin{array}{l} p_1,\ p_2,\ \ldots,\ p_\nu\leqslant m,\\ p_1+p_2+\ldots+p_\nu=n \end{array}\right) \tag{36}$$

(T — произвольная неособенная матрица).

Пример 2. Дано уравнение

$$X^2=X. \tag{37}$$

Матрица, удовлетворяющая этому уравнению, называется *идемпотентной*. Элементарными делителями идемпотентной матрицы могут быть только λ либо $\lambda-1$. Поэтому идемпотентную матрицу можно определить как матрицу простой структуры (т. е. приводящуюся к диагональной форме) с характеристическими числами, равными нулю или единице. Формула, охватывающая все идемпотентные матрицы данного порядка, имеет вид

$$X=T\{\underbrace{1,\ 1,\ \ldots,\ 1}_{n},\ 0,\ \ldots,\ 0\}\,T^{-1}, \tag{38}$$

где T — произвольная неособенная матрица порядка n.

Рассмотрим теперь более общее уравнение

$$f(X)=0, \tag{39}$$

где $f(\lambda)$ — регулярная функция в некоторой области G плоскости комплексного аргумента λ. От искомого решения $X=\|x_{ik}\|_1^n$ будем требовать, чтобы характеристические числа его принадлежали области G. Выпишем все нули функции $f(\lambda)$, лежащие в области G, и их кратности:

$$\lambda_1,\ \lambda_2,\ \ldots,$$

$$a_1,\ a_2,\ \ldots$$

Как и в предыдущем случае, каждый элементарный делитель матрицы X должен иметь вид

$$(\lambda-\lambda_i)^{p_i}\ (p_i\leqslant a_i)$$

и потому

$$X = T\{\lambda_{i_1}E^{(p_{i_1})} + H^{(p_{i_1})}, \ldots, \lambda_{i_\nu}E^{(p_{i_\nu})} + H^{(p_{i_\nu})}\}T^{-1} \tag{40}$$

$$(i_1, i_2, \ldots, i_\nu = 1, 2, \ldots;\ p_{i_1} \leqslant a_{i_1},\ p_{i_2} \leqslant a_{i_2}, \ldots, p_{i_\nu} \leqslant a_{i_\nu};$$
$$p_{i_1} + p_{i_2} + \ldots + p_{i_\nu} = n)$$

(T — произвольная неособенная матрица).

§ 5. Матричное многочленное уравнение

Рассмотрим уравнения

$$A_0X^m + A_1X^{m-1} + \ldots + A_m = 0, \tag{41}$$

$$Y^mA_0 + Y^{m-1}A_1 + \ldots + A_m = 0, \tag{42}$$

где $A_0, A_1, \ldots, A_m$ — заданные, а X и Y — искомые квадратные матрицы порядка n. Уравнение (33), рассмотренное в предыдущем параграфе, представляет собой весьма частный (можно сказать, тривиальный) случай уравнений (41), (42) и получается из последних, если положить $A_i = \alpha_i E$, где α_i — число и $i = 1, 2, \ldots, m$.

Следующая теорема устанавливает связь между уравнениями (41), (42) и (33).

Теорема 4. *Каждое решение матричного уравнения*

$$A_0X^m + A_1X^{m-1} + \ldots + A_m = 0$$

удовлетворяет скалярному уравнению

$$g(X) = 0, \tag{43}$$

где

$$g(\lambda) \equiv |A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m|. \tag{44}$$

Этому же скалярному уравнению удовлетворяет и любое решение Y матричного уравнения

$$Y^mA_0 + Y^{m-1}A_1 + \ldots + A_m = 0.$$

Доказательство. Обозначим через $F(\lambda)$ матричный многочлен

$$F(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m.$$

Тогда уравнения (41) и (42) запишутся так (см. стр. 88):

$$F(X) = 0, \quad \widehat{F}(Y) = 0.$$

Согласно обобщенной теореме Безу (гл. IV, § 2), если X и Y — решения этих уравнений, то матричный многочлен $F(\lambda)$ делится справа на $\lambda E - X$ и слева на $\lambda E - Y$:

$$F(\lambda) = Q(\lambda)(\lambda E - X) = (\lambda E - Y)Q_1(\lambda).$$

Отсюда

$$g(\lambda) = |F(\lambda)| = |Q(\lambda)|\Delta(\lambda) = |Q_1(\lambda)|\Delta_1(\lambda), \tag{45}$$

где $\Delta(\lambda) = |\lambda E - X|$ и $\Delta_1(\lambda) = |\lambda E - Y|$ — характеристические многочлены матриц X и Y. По теореме Гамильтона — Кэли (гл. IV, § 3)

$$\Delta(X) = 0, \quad \Delta_1(Y) = 0.$$

Поэтому из (45) вытекает:

$$g(X) = g(Y) = 0.$$

Теорема доказана.

Мы доказали, что каждое решение уравнения (41) удовлетворяет скалярному уравнению (степени $\leqslant mn$)

$$g(\lambda) = 0.$$

Но множество матричных решений этого уравнения с заданным порядком n распадается на конечное число классов подобных между собой матриц (см. § 4). Поэтому все решения уравнения (41) приходится искать среди матриц вида

$$T_i D_i T_i^{-1} \tag{46}$$

здесь D_i — известные матрицы; при желании можно считать, что D_i имеют нормальную жорданову форму; T_i — произвольные неособенные матрицы n-го порядка; $i = 1, 2, \ldots, h$). Подставим в (41) вместо X матрицу (46) и подберем T_i так, чтобы удовлетворялось уравнение (41). Для каждого T_i получим свое линейное уравнение

$$A_0 T_i D_i^m + A_1 T_i D_i^{m-1} + \ldots + A_m T_i = 0 \quad (i = 1, 2, \ldots, h). \tag{47}$$

Единственный способ, который мы можем предложить для нахождения решения T_i уравнения (47), заключается в замене матричного уравнения системой линейных однородных скалярных уравнений относительно элементов искомой матрицы T_i. Каждое неособенное решение T_i уравнения (47), будучи подставлено в (46), дает решение данного уравнения (41). Аналогичные рассуждения могут быть проведены для уравнения (42).

В следующих двух параграфах мы рассмотрим частные случаи уравнения (41), связанные с извлечением корня m-й степени из матрицы.

Заметим, что теорема Гамильтона — Кэли является частным случаем теоремы 4. В самом деле, любая квадратная матрица A, будучи подставлена вместо λ, удовлетворяет уравнению

$$\lambda E - A = 0.$$

Поэтому в силу доказанной теоремы

$$\Delta(A) = 0,$$

где $\Delta(\lambda) = |\lambda E - A|$.

Теорема 4 может быть обобщена следующим образом:

Теорема 5 (Филлипса)[1]. *Если попарно перестановочные между собой квадратные матрицы n-го порядка $X_0, X_1, \ldots, X_m$ удовлетворяют матричному уравнению*

$$A_0 X_0 + A_1 X_1 + \ldots + A_m X_m = 0 \tag{48}$$

($A_0, A_1, \ldots, A_m$ — заданные квадратные матрицы n-го порядка), то эти же матрицы $X_0, X_1, \ldots, X_m$ удовлетворяют скалярному уравнению

$$g(X_0, X_1, \ldots, X_m) = 0, \tag{49}$$

где

$$g(\xi_0, \xi_1, \ldots, \xi_m) = |A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m|. \tag{50}$$

[1] См. [231].

Доказательство. Положим[1])

$$F(\xi_0, \xi_1, \ldots, \xi_m) = \| f_{ik}(\xi_0, \xi_1, \ldots, \xi_m) \|_1^n = A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m;$$

$\xi_0, \xi_1, \ldots, \xi_m$ — скалярные переменные.

Обозначим через $\widehat{F}(\xi_0, \xi_1, \ldots, \xi_m) = \| \widehat{f}_{ik}(\xi_0, \xi_1, \ldots, \xi_m) \|_1^n$ присоединенную матрицу для матрицы F [$\widehat{f}_{ik}$ есть алгебраическое дополнение (адъюнкта) элемента f_{ki} в определителе $|F(\xi_0, \xi_1, \ldots, \xi_m)| = |f_{ik}|_1^n$ $(i, k = 1, 2, \ldots, n)$]. Тогда каждый элемент $\widehat{f}_{ik}$ $(i, k = 1, 2, \ldots, n)$ матрицы $\widehat{F}$ есть однородный многочлен относительно $\xi_0, \xi_1, \ldots, \xi_{m-1}$ степени $m-1$, и потому матрицу $\widehat{F}$ можно представить в виде

$$\widehat{F} = \sum_{j_0+j_1+\ldots+j_m=n-1} F_{j_0 j_1 \ldots j_m} \xi_0^{j_0} \xi_1^{j} \ldots \xi_m^{j_m},$$

где $F_{j_0 j_1 \ldots j_m}$ — некоторые постоянные матрицы порядка n.

Из определения матрицы $\widehat{F}$ следует тождество

$$\widehat{F}F = g(\xi_0, \xi_1, \ldots, \xi_m) E.$$

Запишем это тождество следующим образом:

$$\sum_{j_0+j_1+\ldots+j_{m-1}=n-1} F_{j_0 j_1 \ldots j_m} (A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m) \xi_0^{j_0} \xi_1^{j_1} \ldots \xi_m^{j_m} = g(\xi_0, \xi_1, \ldots, \xi_m) E. \quad (51)$$

Переход от левой части к правой части в тождестве (51) осуществляется путем раскрытия скобок и приведения подобных членов. При этом *приходится переставлять местами переменные* $\xi_0, \xi_1, \ldots, \xi_m$ *между собой и не приходится переставлять местами переменные* $\xi_0, \xi_1, \ldots, \xi_m$ *с матричными коэффициентами* A_i *и* $F_{j_0 j_1 \ldots j_m}$. Поэтому равенство (51) не нарушится, если мы вместо переменных $\xi_0, \xi_1, \ldots, \xi_m$ подставим попарно перестановочные между собой матрицы $X_0, X_1, \ldots, X_m$:

$$\sum_{0+j_1+\ldots+j_m=n-1} F_{j_0 j_1 \ldots j_m} (A_0X_0 + A_1X_1 + \ldots + A_mX_m) X_0^{j_0} X_1^{j_1} \ldots X_m^{j_m} = g(X_0, X_1, \ldots, X_m). \quad (52)$$

Но по условию $A_0X_0 + A_1X_1 + \ldots + A_mX_m = 0$. Тогда из (52) находим: $g(X_0, X_1, \ldots, X_m) = 0$, что и требовалось доказать.

Замечание 1. Теорема 5 сохраняет свою силу, если уравнение (48) заменить уравнением

$$X_0A_0 + X_1A_1 + \ldots + X_mA_m = 0. \quad (53)$$

Действительно, теорему 5 можно применить к уравнению

$$A_0'X_0 + A_1'X_1 + \ldots + A_m'X_m = 0$$

и затем перейти в этом уравнении почленно к транспонированным матрицам.

Замечание 2. Теорема 4 получится как частный случай теоремы 5, если в качестве $X_0, X_1, \ldots, X_m$ взять

$$X^m, X^{m-1}, \ldots, X, E.$$

---

[1]) $f_{ik}(\xi_0, \xi_1, \ldots, \xi_m)$ — линейные формы от $\xi_0, \xi_1, \ldots, \xi_m$ $(i, k = 1, 2, \ldots, n)$.

§ 6. Извлечение корня m-й степени из неособенной матрицы

Этот и следующий параграфы мы посвятим уравнению

$$X^m = A, \tag{54}$$

где A — заданная, а X — искомая матрицы (обе порядка n), m — данное целое положительное число.

В данном параграфе мы рассмотрим случай, когда $|A| \neq 0$ (A — неособенная матрица). В этом случае все характеристические числа матрицы A отличны от нуля (ибо $|A|$ равен произведению этих характеристических чисел).

Обозначим через

$$(\lambda - \lambda_1)^{p_1},\ (\lambda - \lambda_2)^{p_2},\ \ldots,\ (\lambda - \lambda_u)^{p_u} \tag{55}$$

элементарные делители матрицы A и приведем матрицу A к жордановой форме[1]):

$$A = U\widetilde{A}U^{-1} = U\{\lambda_1 E_1 + H_1,\ \ldots,\ \lambda_u E_u + H_u\}U^{-1}. \tag{56}$$

Так как характеристические числа искомой матрицы X при возведении в m-ю степень дают характеристические числа матрицы A, то и у матрицы X все характеристические числа отличны от нуля. Поэтому на этих характеристических числах производная от

$$f(\lambda) = \lambda^m$$

не обращается в нуль. Но в таком случае (см. гл. VI, стр. 159) элементарные делители матрицы X не «расщепляются» при возведении матрицы X в m-ю степень. Из сказанного следует, что элементарными делителями матрицы X будут:

$$(\lambda - \xi_1)^{p_1},\ (\lambda - \xi_2)^{p_2},\ \ldots,\ (\lambda - \xi_u)^{p_u}, \tag{57}$$

где $\xi_j^m = \lambda_j$, т. е. ξ_j является одним из корней m-й степени из λ_j $\left(\xi_j = \sqrt[m]{\lambda_j};\ j = 1, 2, \ldots, u\right)$.

Определим теперь $\sqrt[m]{\lambda_j E_j + H_j}$ следующим образом. Возьмем в λ-плоскости круг с центром в точке λ_j, не захватывающий нуля. В этом круге мы имеем m раздельных ветвей функции $\sqrt[m]{\lambda}$. Эти ветви можно отличать одну от другой по значениям, которые они принимают в центре круга, в точке λ_j. Обозначим через $\sqrt[m]{\lambda}$ ту ветвь, значение которой в точке λ_j совпадает с характеристическим числом ξ_j искомой матрицы X, и, исходя из этой ветви, определим функцию от матрицы $\sqrt[m]{\lambda_j E_j + H_j}$ с помощью обрывающегося ряда

$$\sqrt[m]{\lambda_j E_j + H_j} = \lambda_j^{\frac{1}{m}} E_j + \frac{1}{m}\lambda_j^{\frac{1}{m}-1} H_j + \frac{1}{2!}\frac{1}{m}\left(\frac{1}{m} - 1\right)\lambda_j^{\frac{1}{m}-2} H_j^2 + \ldots \tag{58}$$

Так как производная от рассматриваемой функции $\sqrt[m]{\lambda}$ в точке λ_j не равна нулю, то матрица (58) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$, где $\xi_j = \sqrt[m]{\lambda_j}$ (здесь $j = 1, 2, \ldots, u$). Отсюда следует, что квази-

---

[1]) Здесь $E_j = E^{(p_j)}$, $H_j = H^{(p_j)}$ $(j = 1, 2, \ldots, u)$.

диагональная матрица

$$\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\}$$

имеет элементарные делители (57), т. е. те же элементарные делители, что и искомая матрица X. Поэтому существует такая неособенная матрица T ($|T| \neq 0$), что

$$X = T\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\}\, T^{-1}. \tag{59}$$

Для определения матрицы T заметим, что, подставляя в обе части тождества

$$(\sqrt[m]{\lambda})^m = \lambda$$

вместо λ матрицу $\lambda_j E_j + H_j$ ($j = 1, 2, \ldots, u$), получим:

$$(\sqrt[m]{\lambda_j E_j + H_j})^m = \lambda_j E_j + H_j \qquad (j = 1, 2, \ldots, u).$$

Теперь из (54) и (59) следует:

$$A = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_u E_u + H_u\}\, T^{-1}. \tag{60}$$

Сопоставляя (56) и (60), найдем:

$$T = UX_{\widetilde{A}}, \tag{61}$$

где $X_{\widetilde{A}}$ — произвольная неособенная матрица, перестановочная с $\widetilde{A}$ (структура матрицы $X_{\widetilde{A}}$ детально описана в § 2).

Подставляя в (59) вместо T выражение $UX_{\widetilde{A}}$, получаем формулу, охватывающую все решения уравнения (54):

$$X = UX_{\widetilde{A}}\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\}\, X_{\widetilde{A}}^{-1} U^{-1}. \tag{62}$$

Многозначность правой части этой формулы имеет как дискретный, так и континуальный характер: дискретный (в данном случае и конечный) характер этой многозначности получается за счет выбора различных ветвей функции $\sqrt[m]{\lambda}$ в различных клетках квазидиагональной матрицы (при этом даже при $\lambda_j = \lambda_k$ ветви $\sqrt[m]{\lambda}$ в j-й и в k-й диагональных клетках могут быть различными); континуальный характер многозначности получается за счет произвольных параметров, содержащихся в матрице $X_{\widetilde{A}}$.

Все решения уравнения (54) мы будем называть *корнями m-й степени из матрицы A* и обозначать многозначным символом $\sqrt[m]{A}$. Обратим внимание на то, что $\sqrt[m]{A}$ в общем случае не является функцией от матрицы A (т. е. не представляется в виде многочлена от A).

Замечание. Если все элементарные делители матрицы A попарно взаимно просты, т. е. числа $\lambda_1, \lambda_2, \ldots, \lambda_u$ все различны, то матрица $X_{\widetilde{A}}$ имеет квазидиагональный вид

$$X_{\widetilde{A}} = \{X_1, X_2, \ldots, X_u\},$$

где матрица X_j перестановочна с $\lambda_j E_j + H_j$ и, следовательно, перестановочна с любой функцией от $\lambda_j E_j + H_j$, в частности с $\sqrt[m]{\lambda_j E_j + H_j}$

($j=1, 2, \ldots, u$). Поэтому в рассматриваемом случае формула (62) принимает вид

$$X=U\{\sqrt[m]{\lambda_1 E_1+H_1}, \sqrt[m]{\lambda_2 E_2+H_2}, \ldots, \sqrt[m]{\lambda_u E_u+H_u}\}U^{-1}.$$

Таким образом, если элементарные делители матрицы A попарно взаимно просты, то в формуле для $X=\sqrt[m]{A}$ имеется только дискретная многозначность. В этом случае любое значение $\sqrt[m]{A}$ можно представить как многочлен от A.

Пример. Пусть требуется найти все квадратные корни из матрицы

$$A=\begin{Vmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{Vmatrix},$$

т. е. все решения уравнения

$$X^2=A.$$

В данном случае матрица A уже имеет нормальную жорданову форму. Поэтому в формуле (62) можно положить $A=\widetilde{A}$, $U=E$. Матрица $X_{\widetilde{A}}$ в данном случае выглядит так (см. стр. 204):

$$X_{\widetilde{A}}=\begin{Vmatrix}a & b & c\\ 0 & a & 0\\ 0 & d & e\end{Vmatrix},$$

где a, b, c, d, e — произвольные параметры.

Формула (62), дающая все искомые решения X, в данном случае принимает следующий вид:

$$X=\begin{Vmatrix}a & b & c\\ 0 & a & 0\\ 0 & d & e\end{Vmatrix}\begin{Vmatrix}\varepsilon & \frac{\varepsilon}{2} & 0\\ 0 & \varepsilon & 0\\ 0 & 0 & \eta\end{Vmatrix}\begin{Vmatrix}a & b & c\\ 0 & a & 0\\ 0 & d & e\end{Vmatrix}^{-1} \qquad (\varepsilon^2=\eta^2=1). \tag{63}$$

Не изменяя X, мы можем в формуле (62) помножить $X_{\widetilde{A}}$ на такой скаляр, чтобы $|X_{\widetilde{A}}|=1$. В данном случае это приведет к равенству $a^2e=1$, откуда $e=a^{-2}$.

Вычислим элементы матрицы $X_{\widetilde{A}}^{-1}$. Для этого выпишем линейное преобразование с матрицей коэффициентов $X_{\widetilde{A}}$:

$$\begin{aligned}y_1&=ax_1+bx_2+cx_3,\\ y_2&=ax_2,\\ y_3&=dx_2+a^{-2}x_3.\end{aligned}$$

Разрешим эту систему уравнений относительно x_1, x_2, x_3. Тогда получим преобразование с обратной матрицей $X_{\widetilde{A}}^{-1}$:

$$\begin{aligned}x_1&=a^{-1}y_1-(a^{-2}b-cd)\,y_2-acy_3,\\ x_2&=a^{-1}y_2,\\ x_3&=-a\,dy_2+a^2y_3.\end{aligned}$$

Отсюда находим:

$$X_{\widetilde{A}}^{-1}=\begin{Vmatrix}a & b & c\\ 0 & a & 0\\ 0 & d & a^{-2}\end{Vmatrix}^{-1}=\begin{Vmatrix}a^{-1} & cd-a^{-2}b & -ac\\ 0 & a^{-1} & 0\\ & -ad & a^2\end{Vmatrix}.$$

Формула (63) дает:

$$X=\begin{Vmatrix} \varepsilon & (\varepsilon-\eta)\,a\,cd+\frac{\varepsilon}{2} & a^2c\,(\eta-\varepsilon) \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon-\eta)\,da^{-1} & \eta \end{Vmatrix}=$$

$$=\begin{Vmatrix} \varepsilon & (\varepsilon-\eta)\,vw+\frac{\varepsilon}{2} & (\eta-\varepsilon)\,v \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon-\eta)\,w & \eta \end{Vmatrix} \qquad (v=a^2c;\ w=a^{-1}d). \tag{64}$$

Решение X зависит от двух произвольных параметров v и w и от двух произвольных знаков ε и η.

§ 7. Извлечение корня m-й степени из особенной матрицы

Переходим к разбору случая, когда $|A|=0$ (A — особенная матрица).

Как и в первом случае, приведем матрицу A к нормальной жордановой форме:

$$A=U\{\lambda_1E^{(p_1)}+H^{(p_1)},\ \ldots,\ \lambda_uE^{(p_u)}+H^{(p_u)};\ H^{(q_1)},\ H^{(q_2)},\ \ldots,\ H^{(q_t)}\}\,U^{-1}; \tag{65}$$

здесь мы через $(\lambda-\lambda_1)^{p_1},\ \ldots,\ (\lambda-\lambda_u)^{p_u}$ обозначили элементарные делители матрицы A, отвечающие ненулевым характеристическим числам, а через $\lambda^{q_1},\ \lambda^{q_2},\ \ldots,\ \lambda^{q_t}$ — элементарные делители с нулевыми характеристическими числами.

Тогда

$$A=U\{A_1,\ A_2\}\,U^{-1}, \tag{66}$$

где

$$A_1=\{\lambda_1E^{(p_1)}+H^{(p_1)},\ \ldots,\ \lambda_uE^{(p_u)}+H^{(p_u)}\},\quad A_2=\{H^{(q_1)},\ H^{(q_2)},\ \ldots,\ H^{(q_t)}\}. \tag{67}$$

Заметим, что A_1 — неособенная матрица ($|A_1|\neq 0$), а A_2 — нильпотентная матрица с индексом нильпотентности $\mu=\max(q_1, q_2, \ldots, q_t)$ ($A_2^{\mu}=0$).

Из исходного уравнения (54) следует перестановочность матрицы A с искомой матрицей X, а следовательно, и перестановочность подобных им матриц

$$U^{-1}AU=\{A_1,\ A_2\}\quad\text{и}\quad U^{-1}XU. \tag{68}$$

Как было доказано в § 2 (теорема 3), из перестановочности матриц (68) и из того факта, что матрицы A_1 и A_2 не имеют общих характеристических чисел, вытекает, что и вторая из матриц (68) имеет соответственную квазидиагональную форму

$$U^{-1}XU=\{X_1,\ X_2\}. \tag{69}$$

Заменяя в уравнении (54) матрицы A и X подобными им матрицами

$$\{A_1,\ A_2\}\quad\text{и}\quad\{X_1,\ X_2\},$$

мы заменим уравнение (54) двумя уравнениями:

$$X_1^m=A_1, \tag{70}$$

$$X_2^m=A_2. \tag{71}$$

Так как $|A_1| \neq 0$, то к уравнению (70) применимы результаты предыдущего параграфа. Поэтому X_1 находим по формуле (62):

$$X_1 = X_{A_1}\left\{\sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \ldots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}}\right\} X_{A_1}^{-1}. \tag{72}$$

Таким образом, остается рассмотреть уравнение (71), т. е. заняться нахождением всех корней m-й степени из нильпотентной матрицы A_2, уже имеющей нормальную жорданову форму:

$$A_2 = \{H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}\}. \tag{73}$$

$\mu = \max(q_1, q_2, \ldots, q_t)$ — индекс нильпотентности матрицы A_2. Из $A_2^{\mu} = 0$ и из (71) находим:

$$X_2^{m\mu} = 0.$$

Последнее равенство показывает, что искомая матрица X_2 также является нильпотентной с индексом нильпотентности ν, где $m(\mu - 1) < \nu \leqslant m\mu$. Приведем матрицу X_2 к жордановой форме:

$$X_2 = T\{H^{(\nu_1)}, H^{(\nu_2)}, \ldots, H^{(\nu_s)}\} T^{-1} \tag{74}$$

$$(\nu_1, \nu_2, \ldots, \nu_s \leqslant \nu).$$

Возведем теперь обе части последнего равенства в m-ю степень. Получим:

$$A_2 = X_2^m = T\{[H^{(\nu_1)}]^m, [H^{(\nu_2)}]^m, \ldots, [H^{(\nu_s)}]^m\} T^{-1}. \tag{75}$$

Выясним теперь, какие элементарные делители имеет матрица $[H^{(\nu)}]^m$ [1]). Обозначим через $\boldsymbol{H}$ линейный оператор, задаваемый матрицей $H^{(\nu)}$ в ν-мерном векторном пространстве с базисом $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_\nu$. Тогда из вида матрицы $H^{(\nu)}$ (в матрице $H^{(\nu)}$ все элементы первой наддиагонали равны единице и все остальные элементы равны нулю) следует, что

$$\boldsymbol{H}\boldsymbol{e}_1 = 0, \quad \boldsymbol{H}\boldsymbol{e}_2 = \boldsymbol{e}_1, \quad \ldots, \quad \boldsymbol{H}\boldsymbol{e}_\nu = \boldsymbol{e}_{\nu-1}. \tag{76}$$

Эти равенства показывают, что для оператора $\boldsymbol{H}$ векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_\nu$ образуют жорданову цепочку векторов, соответствующую элементарному делителю λ^ν.

Равенства (76) запишем так:

$$\boldsymbol{H}\boldsymbol{e}_j = \boldsymbol{e}_{j-1} \qquad (j = 1, 2, \ldots, \nu;\ \boldsymbol{e}_0 = 0).$$

Очевидно, что

$$\boldsymbol{H}^m \boldsymbol{e}_j = \boldsymbol{e}_{j-m} \qquad (j = 1, 2, \ldots, \nu;\ \boldsymbol{e}_0 = \boldsymbol{e}_{-1} = \ldots = \boldsymbol{e}_{-m+1} = 0). \tag{77}$$

Представим число ν в виде

$$\nu = km + r \qquad (r < m),$$

---

[1]) Ответ на этот вопрос дает теорема 9 гл. VI (стр. 158—159). Здесь мы вынуждены другим методом исследовать этот вопрос, так как нам нужно отыскать не только элементарные делители матрицы $[H^{(\nu)}]^m$, но и матрицу $\boldsymbol{P}_{\nu, m}$, преобразующую $[H^{(\nu)}]^m$ к жордановой форме.

где k, r — целые неотрицательные числа. Расположим базисные векторы $e_1, e_2, \ldots, e_v$ следующим образом:

$$\begin{array}{llll} e_1, & e_2, & \ldots, & e_m, \\ e_{m+1}, & e_{m+2}, & \ldots, & e_{2m}, \\ \ldots & \ldots & \ldots & \ldots \\ e_{(k-1)m+1}, & e_{(k-1)m+2}, & \ldots, & e_{km}, \\ e_{km+1}, & \ldots, & e_{km+r}. & \end{array} \tag{78}$$

В этой таблице мы имеем m столбцов: первые r содержат по $(k+1)$ векторов в каждом, остальные — по k векторов. Равенство (78) показывает, что векторы каждого столбца образуют жорданову цепочку векторов относительно оператора H^m. Если вместо последовательной нумерации векторов (78) по строкам занумеровать их по столбцам, то в полученном таким образом новом базисе матрица оператора H^m будет иметь следующую нормальную жорданову форму:

$$\{\underbrace{H^{(k+1)}, \ldots, H^{(k+1)}}_{r}, \underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}\}\ ^{1}),$$

и следовательно,

$$[H^{(v)}]^m = P_{v,m}\{\underbrace{H^{(k+1)}, \ldots, H^{(k+1)}}_{r}, \underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}\}P_{v,m}^{-1}, \tag{79}$$

где матрица $P_{v,m}$ (матрица перехода от одного базиса к другому) имеет вид (см. гл. III, § 4)

$$P_{v,m} = \left\| \begin{array}{cccccc} \overbrace{1 & 0 & \ldots & 0}^{m} & 0 & \ldots \\ 0 & 0 & \ldots & 0 & 1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & & \\ 0 & 1 & \ldots & 0 & & \\ \ldots & \ldots & \ldots & \ldots & & \end{array} \right\| \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\right\} m \tag{80}$$

Матрица $H^{(v)}$ имеет один элементарный делитель λ^v. При возведении матрицы $H^{(v)}$ в m-ю степень этот элементарный делитель «расщепляется». Как показывает формула (79), матрица $[H^{(v)}]^m$ имеет элементарные делители:

$$\underbrace{\lambda^{k+1}, \ldots, \lambda^{k+1}}_{r}, \underbrace{\lambda^k, \ldots, \lambda^k}_{m-r}.$$

Возвращаясь теперь к равенству (75), положим:

$$v_i = k_i m + r_i \ (0 \leqslant r_i \leqslant m,\ k_i \geqslant 0;\ i = 1, 2, \ldots, s). \tag{81}$$

---

[1]) В случае $k=0$ клетки $\underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}$ отсутствуют, и эта матрица имеет вид $\underbrace{\{H^{(1)}, \ldots, H^{(1)}\}}_{r}$.

Тогда в силу (79) равенство (75) перепишется так:

$$A_2 = X_2^m = TP\{\underbrace{H^{(k_1+1)}, \ldots, H^{(k_1+1)}}_{r_1}, \underbrace{H^{(k_1)}, \ldots, H^{(k_1)}}_{m-r_1},$$
$$\underbrace{H^{(k_2+1)}, \ldots, H^{(k_2+1)}}_{r_2}, H^{(k_2)}, \ldots\} P^{-1}T^{-1}, \tag{82}$$

где $P = \{P_{v_1, m}, P_{v_2, m}, \ldots, P_{v_s, m}\}$.

Сопоставляя (82) с (73), видим, что клетки

$$H^{(k_1+1)}, \ldots, H^{(k_1+1)}, H^{(k_1)}, \ldots, H^{(k_1)}, H^{(k_2+1)}, \ldots, H^{(k_2+1)}, \ldots \tag{83}$$

с точностью до порядка должны совпасть с клетками

$$H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}. \tag{84}$$

Условимся систему элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \ldots, \lambda^{v_s}$ называть *возможной* для X_2, если после возведения матрицы в m-ю степень эти элементарные делители, расщепляясь, порождают заданную систему элементарных делителей матрицы A_2: $\lambda^{q_1}, \lambda^{q_2}, \ldots, \lambda^{q_t}$. Число возможных систем элементарных делителей всегда конечно, поскольку

$$\max(v_1, v_2, \ldots, v_s) \leqslant m\mu, \quad v_1 + v_2 + \ldots + v_s = n_2 \tag{85}$$
$$(n_2 — \text{степень матрицы } A_2).$$

В каждом конкретном случае возможные системы элементарных делителей для X_2 могут быть легко определены путем конечного числа испытаний.

Покажем, что для каждой возможной системы элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \ldots, \lambda^{v_s}$ существуют соответствующие решения уравнения (71), и определим все эти решения. В этом случае существует преобразующая матрица Q такая, что

$$\{H^{(k_1+1)}, \ldots, H^{(k_1+1)}, H^{(k_1)}, \ldots, H^{(k_1)}, H^{(k_2+1)}, \ldots\} = Q^{-1}A_2Q. \tag{86}$$

Матрица Q осуществляет перестановку клеток в квазидиагональной матрице, что достигается надлежащей перенумерацией базисных векторов. Поэтому матрицу Q можно считать известной. Используя (86), мы из (82) получим:

$$A_2 = TPQ^{-1}A_2QP^{-1}T^{-1}.$$

Отсюда

$$TPQ^{-1} = X_{A_2}$$

или

$$T = X_{A_2}QP^{-1}, \tag{87}$$

где X_{A_2} — произвольная матрица, перестановочная с A_2.

Подставляя выражение (87) для T в (74), будем иметь:

$$X_2 = X_{A_2}QP^{-1}\{H^{(v_1)}, H^{(v_2)}, \ldots, H^{(v_s)}\} PQ^{-1}X_{A_2}^{-1}. \tag{88}$$

Из (69), (72) и (88) получим общую формулу, охватывающую все искомые решения:

$$X = U\{X_{A_1}, X_{A_2}QP^{-1}\}\{\sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \ldots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}},$$
$$H^{(v_1)}, \ldots, H^{(v_s)}\} \cdot \{X_{A_1}^{-1}, PQ^{-1}X_{A_2}^{-1}\} U^{-1}. \tag{89}$$

Обратим внимание читателя на то, что корень m-й степени из особенной матрицы не всегда существует. Его существование связано с существованием системы возможных элементарных делителей для матрицы X_2.

Легко видеть, например, что уравнение

$$X^m = H^{(p)}$$

не имеет решений при $m>1$, $p>1$.

Пример. Требуется извлечь корень квадратный из матрицы

$$A=\begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

т. е. найти все решения уравнения

$$X^2=A.$$

В данном случае $A=A_2$, $X=X_2$, $m=2$, $t=2$, $q_1=2$, $q_2=1$. Матрица X может иметь только один элементарный делитель λ^3. Поэтому $s=1$, $v_1=3$, $k_1=1$, $r_1=1$ и [см. (80)]

$$P=P_{3,2}=\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix}=P^{-1}, \quad Q=E.$$

Кроме того, как и в примере на стр. 214, можно положить в формуле (88):

$$X_{A_2}=\begin{Vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{Vmatrix}, \quad X_{A_2}^{-1}=\begin{Vmatrix} a^{-1} & cd-a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{Vmatrix}.$$

Из этой формулы получим:

$$X=X_2=X_{A_3}P^{-1}H^{(3)}PX_{A_2}^{-1}=\begin{Vmatrix} 0 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & \beta^{-1} & 0 \end{Vmatrix},$$

где $\alpha=ca^{-1}-a^2d$ и $\beta=a^3$ — произвольные параметры.

§ 8. Логарифм матрицы

1. Рассмотрим матричное уравнение

$$e^X=A. \tag{90}$$

Все решения этого уравнения будем называть логарифмами (натуральными) матрицы A и обозначать через $\ln A$.

Характеристические числа λ_j матрицы A связаны с характеристическими числами ξ_j матрицы X формулой $\lambda_j=e^{\xi_j}$; поэтому, если уравнение (90) имеет решение, то все характеристические числа матрицы A отличны от нуля и матрица A является неособенной ($|A|\neq 0$). Таким образом, условие $|A|\neq 0$ является необходимым для существования решения уравнения (90). Ниже мы увидим, что это условие является и достаточным.

Итак, пусть $|A|\neq 0$. Выпишем элементарные делители матрицы A:

$$(\lambda-\lambda_1)^{p_1},\ (\lambda-\lambda_2)^{p_2},\ \ldots,\ (\lambda-\lambda_u)^{p_u} \tag{91}$$
$$(\lambda_1\lambda_2\ldots\lambda_u\neq 0,\ p_1+p_2+\ldots+p_u=n).$$

В соответствии с этими элементарными делителями приведем матрицу A к нормальной жордановой форме:

$$A = U\widetilde{A}U^{-1} = \\ = U\{\lambda_1 E^{(p_1)} + H^{(p_1)},\ \lambda_2 E^{(p_2)} + H^{(p_2)},\ \ldots,\ \lambda_u E^{(p_u)} + H^{(p_u)}\}U^{-1}. \quad (92)$$

Так как производная от функции e^{ξ} отлична от нуля при всех значениях ξ, то (см. гл. VI, стр. 159) при переходе от матрицы X к матрице $A = e^X$ элементарные делители не расщепляются, т. е. матрица X имеет элементарные делители

$$(\lambda - \xi_1)^{p_1},\ (\lambda - \xi_2)^{p_2},\ \ldots,\ (\lambda - \xi_u)^{p_u}, \quad (93)$$

где $e^{\xi_j} = \lambda_j$ $(j = 1, 2, \ldots, u)$, т. е. ξ_j есть одно из значений $\ln \lambda_j$ $(j = 1, 2, \ldots, u)$.

Возьмем в плоскости комплексного переменного λ круг с центром в точке λ_j радиуса $< |\lambda_j|$ и обозначим через $f_j(\lambda) = \ln \lambda$ ту из ветвей функции $\ln \lambda$ в рассматриваемом круге, которая в точке λ_j принимает значение, равное характеристическому числу ξ_j матрицы X $(j = 1, 2, \ldots, u)$. После этого полагаем:

$$\ln(\lambda_j E^{(p_j)} + H^{(p_j)}) = f_j(\lambda_j E^{(p_j)} + H^{(p_j)}) = \ln \lambda_j E^{(p_j)} + \lambda_j^{-1} H^{(p_j)} + \ldots \quad (94)$$

Так как производная от $\ln \lambda$ нигде не обращается в нуль (в конечной части плоскости λ), то матрица (94) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$. В силу этого квазидиагональная матрица

$$\{\ln(\lambda_1 E^{(p_1)} + H^{(p_1)}),\ \ln(\lambda_2 E^{(p_2)} + (\lambda E_j^{(p_2)}),\ \ldots,\ \ln(\lambda_u E^{(p_u)} + H^{(p_u)})\} \quad (95)$$

имеет те же элементарные делители, что и искомая матрица X. Поэтому существует такая матрица T $(|T| \neq 0)$, что

$$X = T\{\ln(\lambda_1 E^{(p_1)} + H^{(p_1)}),\ \ldots,\ \ln(\lambda_u E^{(p_u)} + H^{(p_u)}\}\, T^{-1}. \quad (96)$$

Для определения матрицы T заметим, что

$$A = e^X = T\{\lambda_1 E^{(p_1)} + H^{(p_1)},\ \ldots,\ \lambda_u E^{(p_u)} + H^{(p_u)}\}\, T^{-1}. \quad (97)$$

Сопоставляя (97) с (92), находим:

$$T = UX_{\widetilde{A}}, \quad (98)$$

где $X_{\widetilde{A}}$ — произвольная матрица, перестановочная с матрицей $\widetilde{A}$. Подставляя выражение для T из (98) в (96), получим общую формулу, охватывающую все логарифмы матрицы:

$$X = UX_{\widetilde{A}}\{\ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \\ \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}),\ \ldots,\ \ln(\lambda_u E^{(p_u)} + H^{(p_u)})\}\, X_{\widetilde{A}}^{-1} U^{-1}. \quad (99)$$

Замечание. Если все элементарные делители матрицы A взаимно просты, то в правой части формулы (99) можно выбросить множители $X_{\widetilde{A}}$ и $X_{\widetilde{A}}^{-1}$ (см. аналогичное замечание на стр. 213).

2. Выясним, когда вещественная неособенная матрица A имеет вещественный логарифм X. Пусть искомая матрица имеет несколько элементарных делителей, отвечающих характеристическому числу вида $\varrho + i\pi$: $(\lambda - \varrho - i\pi)^{q_1}, \ldots, (\lambda - \varrho - i\pi)^{q_t}$. Поскольку матрица X вещественна, то она имеет и сопряженные элементарные делители: $(\lambda - \varrho + i\pi)^{q_1}, \ldots, (\lambda - \varrho + i\pi)^{q_t}$. При переходе от матрицы X к матри-

це A элементарные делители не расщепляются, но характеристические числа $\varrho+i\pi$, $\varrho-i\pi$ заменяются в них числами $e^{\varrho+i\pi}=-\mu$, $e^{\varrho-i\pi}=-\mu$, где $\mu=e^{\varrho}>0$. Поэтому в системе элементарных делителей матрицы A каждый элементарный делитель, соответствующий отрицательному характеристическому числу (если таковые существуют), повторяется четное число раз. Докажем теперь, что это необходимое условие является и достаточным, т. е. что *вещественная неособенная матрица A тогда и только тогда имеет вещественный логарифм X, когда у матрицы A либо совсем нет элементарных делителей, соответствующих отрицательным характеристическим числам* [1], *либо каждый такой элементарный делитель повторяется четное число раз* [2].

Действительно, пусть это условие выполнено. Тогда в квазидиагональной матрице (95) в соответствии с формулой (94) в тех клетках, где λ_i вещественно и положительно, возьмем для $\ln\lambda_i$ вещественное значение; если же в какой-либо клетке имеется комплексное λ_n, то найдется другая клетка такого же размера с $\lambda_g=\bar{\lambda}_n$. В этих клетках возьмем комплексно сопряженные значения для $\ln\lambda_n$ и $\ln\lambda_g$. Каждая же клетка по условию повторяется в (98) четное число раз с сохранением размера клетки. Тогда в половине этих клеток положим $\ln\lambda_k=\ln|\lambda_k|+i\pi$, а в другой половине возьмем $\ln\lambda_k=\ln|\lambda_k|-i\pi$. Тогда в квазидиагональной матрице (98) диагональные клетки либо будут вещественными, либо будут попарно комплексно сопряженными. Но *такая квазидиагональная матрица всегда подобна вещественной матрице* [3]. Поэтому существует такая неособенная матрица T_1 ($|T_1|\neq 0$), что матрица

$$X_1=T_1\{\ln(\lambda_1E^{(p_1)}+H^{(p_1)}),\ \ldots,\ \ln(\lambda_nE^{(p_n)}+H^{(p_n)})\}\,T_1^{-1}$$

вещественна. Но тогда будет вещественной и матрица

$$A_1=e^{X_1}=T_1\{\lambda_1E^{(p_1)}+H^{(p_1)},\ \ldots,\ \lambda_nE^{(p_n)}+H^{(p_n)}\}\,T^{-1}. \tag{100}$$

Сопоставляя формулу (100) с формулой (92), заключаем, что матрицы A и A_1 подобны между собой (поскольку они подобны одной и той же жордановой матрице). Но две подобные вещественные матрицы могут быть преобразованы друг в друга с помощью некоторой неособенной вещественной матрицы W ($|W|\neq 0$):

$$A=WA_1W^{-1}=We^{X_1}W^{-1}=e^{WX_1W^{-1}}.$$

Тогда матрица $X=WX_1W^{-1}$ и будет искомым вещественным логарифмом матрицы A.

---

[1]) В этом случае существует вещественный $\ln A=r(A)$, где $r(\lambda)$ — надлежащий интерполяционный многочлен для $\ln_0\lambda$ (см. стр. 104).

[2]) Это условие, в частности, выполняется, когда $A=B^2$, где B — вещественная матрица.

[3]) Для того чтобы убедиться в этом, достаточно показать, что квазидиагональная матрица

$$D=\begin{pmatrix}B & 0\\ 0 & \bar{B}\end{pmatrix}$$

всегда подобна некоторой вещественной матрице. Здесь $B=U+iV$, $\bar{B}=U-iV$, где U и V — вещественные матрицы. Обозначая через E единичную матрицу тех же размеров, что и B, и полагая

$$T=\begin{pmatrix}E & iE\\ E & -iE\end{pmatrix},$$

легко проверим, что

$$T^{-1}=\begin{pmatrix}\frac{1}{2}E & \frac{1}{2}E\\ \frac{1}{2i}E & -\frac{1}{2i}E\end{pmatrix},\quad T^{-1}DT=\begin{pmatrix}U & -V\\ V & U\end{pmatrix}$$

ГЛАВА IX

ЛИНЕЙНЫЕ ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ

§ 1. Общие соображения

В главах III и VII мы изучали линейные операторы в произвольном n-мерном векторном пространстве. Все базисы такого пространства равноправны между собой. Данному линейному оператору в каждом базисе отвечает некоторая матрица. Матрицы, отвечающие одному и тому же оператору в различных базисах, подобны между собой. Таким образом, изучение линейных операторов в n-мерном векторном пространстве давало возможность выявить свойства матрицы, присущие одновременно всему классу подобных между собой матриц.

В начале этой главы мы введем метрику в n-мерное векторное пространство, относя специальным образом каждым двум векторам некоторое число — их «скалярное произведение». С помощью скалярного произведения определяется «длина» вектора и «косинус угла» между двумя векторами. Такая метризация приводит нас к унитарному пространству, если основное поле K — поле всех комплексных чисел, и к евклидову пространству, если K — поле всех вещественных чисел.

В настоящей главе мы будем изучать свойства линейных операторов, связанные с метрикой пространства. По отношению к метрике пространства уже не все базисы равноправны. Однако равноправными являются все ортонормированные базисы. Переход от одного ортонормированного базиса к другому осуществляется в унитарном (соответственно евклидовом) пространстве при помощи специального — унитарного (соответственно ортогонального) — преобразования. Поэтому две матрицы, отвечающие одному и тому же линейному оператору в двух различных базисах унитарного (евклидова) пространства, унитарно-подобны (ортогонально-подобны) между собой. Таким образом, изучая линейные операторы в n-мерном метризованном пространстве, мы изучаем те свойства матрицы, которые остаются инвариантными при переходе от данной матрицы к матрице унитарно- или ортогонально-подобной. Это приводит нас естественным образом к исследованию свойств специальных классов матриц (нормальных, эрмитовых, унитарных, симметрических, кососимметрических, ортогональных).

§ 2. Метризация пространства

Рассмотрим векторное пространство R над полем комплексных чисел. Пусть каждым двум векторам x и y из R, заданным в определенном порядке, отнесено некоторое комплексное число, называемое *скалярным произведением* этих векторов и обозначаемое через (xy) или (x, y). Пусть при этом имеют место следующие свойства «скалярного умножения».

Для любых векторов x, y, z из R и любого комплексного числа α

$$\left.\begin{aligned} &1. && (xy) = \overline{(yx)}\ ^{1}), \\ &2. && (\alpha x,\ y) = \alpha\,(xy), \\ &3. && (x+y,\ z) = (xz)+(yz). \end{aligned}\right\} \qquad (1)$$

В этом случае говорят, что в пространство R внесена *эрмитова метрика*.

Заметим еще, что из 1, 2 и 3 следует для любых x, y, z из R:

$2'$. $(x,\ \alpha y) = \bar{\alpha}(xy)$,

$3'$. $(x,\ y+z) = (xy)+(xz)$.

Из 1 заключаем, что для любого вектора x скалярное произведение (xx) является вещественным числом.

Если для любого вектора x из R

$$4. \qquad (xx) \geqslant 0, \qquad (2)$$

то эрмитова метрика называется *неотрицательной*. Если же при этом

$$5. \qquad (xx) > 0 \quad \text{при } x \neq 0, \qquad (3)$$

то эрмитова метрика называется *положительно определенной*.

Определение 1. Векторное пространство R с положительно определенной эрмитовой метрикой мы будем называть *унитарным* пространством [2]).

В настоящей главе мы будем рассматривать конечномерные унитарные пространства [3]).

Под длиной вектора x понимают [4]) $|x| = \sqrt{(x, x)}$. Из 2 и 5 следует, что каждый вектор, отличный от нуля, имеет положительную длину и лишь вектор-нуль имеет длину, равную нулю. Вектор x называется *нормированным* (также *единичным вектором* или *ортом*), если $|x| = 1$. Для «нормировки» произвольного вектора $x \neq 0$ достаточно умножить этот вектор на любое комплексное число λ, у которого $|\lambda| = \frac{1}{|x|}$.

По аналогии с обычным трехмерным векторным пространством два вектора x и y называются ортогональными (обозначение: $x \perp y$), если $(xy) = 0$. В этом случае из 1, 3, $3'$ следует:

$$(x+y,\ x+y) = (xx)+(yy)$$

т. е. (теорема Пифагора!)

$$|x+y|^2 = |x|^2+|y|^2 \qquad (x \perp y).$$

Пусть унитарное пространство R имеет конечное число измерений n. Рассмотрим в R произвольный базис $e_1, e_2, \ldots, e_n$. Обозначим через x_i и y_i $(i = 1, 2, \ldots, n)$ соответственно координаты векторов x и y в этом базисе:

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{i=1}^{n} y_i e_i.$$

---

[1]) Черта над числом означает переход к комплексно сопряженному числу.

[2]) Исследование n-мерных векторных пространств с произвольной (не положительно определенной) метрикой проведено в статье [117], а также в книге [24], гл. IX и X.

[3]) В §§ 2—7 этой главы во всех случаях, когда конечномерность пространства не будет особо оговорена, все рассуждения сохраняют свою силу и для бесконечномерных пространств.

[4]) Здесь знаком $\sqrt{\ }$ обозначаем неотрицательное (арифметическое) значение корня.

Тогда в силу 2, 3, 2′ и 3′

$$(\boldsymbol{xy}) = \sum_{i,\,k=1}^{n} h_{ik} x_i \bar{y}_k, \tag{4}$$

где

$$h_{ik} = (\boldsymbol{e}_i \boldsymbol{e}_k) \qquad (i,\ k = 1,\ 2,\ \ldots,\ n). \tag{5}$$

В частности,

$$(\boldsymbol{xx}) = \sum_{i,\,k=1}^{n} h_{ik} x_i \bar{x}_k. \tag{6}$$

Из 1 и (5) следует:

$$h_{ki} = \bar{h}_{ik} \qquad (i,\ k = 1,\ 2,\ \ldots,\ n). \tag{7}$$

Форма $\sum\limits_{i,\,k=1}^{n} h_{ik} x_i \bar{x}_k$, где $h_{ki} = \bar{h}_{ik}$ $(i, k = 1, 2, \ldots, n)$, называется *эрмитовой*[1]). Таким образом, квадрат длины вектора представляется в виде эрмитовой формы его координат. Отсюда и название «эрмитова метрика». Форма, стоящая в правой части равенства (6), является в силу 4 *неотрицательной*:

$$\sum_{i,\,k=1}^{n} h_{ik} x_i \bar{x}_k \geqslant 0, \tag{8}$$

при всех значениях переменных $x_1, x_2, \ldots, x_n$. В силу же дополнительного условия 5 эта форма будет *положительно определенной*, т. е. знак $=$ в (8) будет иметь место только при равенстве нулю всех x_i $(i = 1, 2, \ldots, n)$.

Определение 2. Систему векторов $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ будем называть ортонормированной, если

$$(\boldsymbol{e}_i \boldsymbol{e}_k) = \delta_{ik} = \begin{cases} 0 & \text{при } i \neq k, \\ 1 & \text{при } i = k \end{cases} \qquad (i,\ k = 1,\ 2,\ \ldots,\ m). \tag{9}$$

При $m = n$, где n — число измерений пространства, получаем *ортонормированный базис* пространства.

В § 7 будет доказано, что *в каждом n-мерном унитарном пространстве существует ортонормированный базис.*

Пусть x_i и y_i $(i = 1, 2, \ldots, n)$ — соответственно координаты векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в ортонормированном базисе. Тогда в силу (4), (5) и (9)

$$\left.\begin{aligned} (\boldsymbol{xy}) &= \sum_{i=1}^{n} x_i \bar{y}_i, \\ (\boldsymbol{xx}) &= \sum_{i=1}^{n} |x_i|^2. \end{aligned}\right\} \tag{10}$$

Фиксируем произвольно некоторый базис в n-мерном пространстве $\boldsymbol{R}$. При этом базисе каждая метризация пространства связана с некоторой положительно определенной эрмитовой формой $\sum\limits_{i,\,k=1}^{n} h_{ik} x_i \bar{x}_k$, и наоборот, согласно (4) каждая такая форма

[1]) В соответствии с этим выражение, стоящее в правой части равенства (4), называется билинейной эрмитовой формой (относительно величин $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$).

определяет некоторую положительно определенную эрмитову метрику в $\boldsymbol{R}$. Однако *все эти метрики не дают существенно различных унитарных n-мерных пространств.* Действительно, возьмем две такие метрики со скалярным произведением соответственно (xy) и $(xy)'$. По отношению к этим метрикам определим ортонормированные базисы в $\boldsymbol{R}$: e_i и e_i' $(i=1, 2, \ldots, n)$. Отнесем друг другу векторы x и x' из $\boldsymbol{R}$ $(x \to x')$, имеющие в этих базисах одинаковые координаты. Это соответствие является аффинным[1]). Кроме того, в силу (10)

$$(xy)=(x'y')'.$$

Таким образом, *с точностью до аффинного преобразования пространства все положительно определенные эрмитовы метризации n-мерного векторного пространства совпадают друг с другом.*

Если основным числовым полем K является поле вещественных чисел, то метрика, удовлетворяющая постулатам 1, 2, 3, 4 и 5, называется *евклидовой.*

Определение 3. Векторное пространство $\boldsymbol{R}$ над полем вещественных чисел с положительно евклидовой метрикой называется *евклидовым пространством.*

Если x_i и y_i $(i=1, 2, \ldots, n)$ суть координаты векторов x и y в некотором базисе $e_1, e_2, \ldots, e_n$ n-мерного евклидова пространства, то

$$(xy)=\sum_{i,\,k=1}^{n} s_{ik}x_iy_k, \qquad |x|^2=\sum_{i,\,k=1}^{n} s_{ik}x_ix_k.$$

Здесь $s_{ik}=s_{ki}$ $(i,\ k=1,\ 2, \ldots, n)$ — вещественные числа[2]). Выражение $\sum_{i,\,k=1}^{n} s_{ik}x_ix_k$ называется *квадратичной формой* относительно $x_1, x_2, \ldots, x_n$. Из положительной определенности метрики вытекает, что квадратичная форма $\sum_{i,\,k=1}^{n} s_{ik}x_ix_k$, задающая аналитически эту метрику, является *положительно определенной*, т. е. $\sum_{i,\,k=1}^{n} s_{ik}x_ix_k>0$, если $\sum_{i=1}^{n} x_i^2>0$.

При ортонормированном базисе

$$(xy)=\sum_{i=1}^{n} x_iy_i, \qquad |x|^2=\sum_{i=1}^{n} x_i^2. \tag{11}$$

При $n=3$ получаем известные формулы для скалярного произведения двух векторов и для квадрата длины вектора в трехмерном евклидовом пространстве.

§ 3. Критерий Грама линейной зависимости векторов

Пусть векторы $x_1, x_2, \ldots, x_m$ унитарного или евклидова пространства $\boldsymbol{R}$ линейно зависимы, т. е. существуют такие не равные одновременно нулю числа $c_1, c_2, \ldots, c_m$[3]), что

$$c_1x_1+c_2x_2+\ldots+c_mx_m=0. \tag{12}$$

---

[1]) То есть оператор A, относящий вектору x из $\boldsymbol{R}$ вектор x' из $\boldsymbol{R}$, является линейным и неособенным.

[2]) $s_{ik}=(e_ie_k)$ $(i,\ k=1,\ 2, \ldots, n)$.

[3]) В случае евклидова пространства $c_1, c_2, \ldots, c_m$ — вещественные числа.

Умножив последовательно обе части этого равенства слева скалярно на x_1, x_2, ..., x_m, получим:

$$\left.\begin{aligned}
(x_1x_1)\,\bar{c}_1+(x_1x_2)\,\bar{c}_2+\ldots+(x_1x_m)\,\bar{c}_m&=0,\\
(x_2x_1)\,\bar{c}_1+(x_2x_2)\,\bar{c}_2+\ldots+(x_2x_m)\,\bar{c}_m&=0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
(x_mx_1)\,\bar{c}_1+(x_mx_2)\,\bar{c}_2+\ldots+(x_mx_m)\,\bar{c}_m&=0.
\end{aligned}\right\} \tag{13}$$

Рассматривая $\bar{c}_1$, $\bar{c}_2$, ..., $\bar{c}_m$ как ненулевое решение системы линейных однородных уравнений (13) с определителем

$$\Gamma(x_1, x_2, \ldots, x_m)=\begin{vmatrix}
(x_1x_1) & (x_1x_2) & \ldots & (x_1x_m)\\
(x_2x_1) & (x_2x_2) & \ldots & (x_2x_m)\\
\ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots\\
(x_mx_1) & (x_mx_2) & \ldots & (x_mx_m)
\end{vmatrix}, \tag{14}$$

заключаем, что этот определитель равен нулю:

$$\Gamma(x_1, x_2, \ldots, x_m)=0.$$

Определитель $\Gamma(x_1, x_2, \ldots, x_m)$ называется *определителем Грама*, составленным для векторов x_1, x_2, ..., x_m.

Пусть, обратно, определитель Грама (14) равен нулю. Тогда система уравнений (13) имеет ненулевое решение $\bar{c}_1$, $\bar{c}_2$, ..., $\bar{c}_m$. Равенства (13) можно записать так:

$$\left.\begin{aligned}
(x_1,\ c_1x_1+c_2x_2+\ldots+c_mx_m)&=0,\\
(x_2,\ c_1x_1+c_2x_2+\ldots+c_mx_m)&=0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
(x_m,\ c_1x_1+c_2x_2+\ldots+c_mx_m)&=0.
\end{aligned}\right\} \tag{13'}$$

Умножая почленно эти равенства соответственно на c_1, c_2, ..., c_m и складывая, получим:

$$|(c_1x_1+c_2x_2+\ldots+c_mx_m)|^2=0;$$

отсюда в силу положительной определенности метрики

$$c_1x_1+c_2x_2+\ldots+c_mx_m=0,$$

т. е. векторы x_1, x_2, ..., x_m линейно зависимы.

Нами доказана

Теорема 1. *Для того чтобы векторы x_1, x_2, ..., x_m были линейно независимы, необходимо и достаточно, чтобы определитель Грама, составленный из этих векторов, не был равен нулю.*

Отметим следующее свойство определителя Грама.

Если какой-либо главный минор определителя Грама равен нулю, то равен нулю и сам определитель Грама.

Действительно, главный минор является определителем Грама для части векторов. Из равенства нулю этого главного минора следует линейная зависимость между этими векторами, а значит, и между векторами полной системы.

Пример. Даны n комплексных функций $f_1(t)$, $f_2(t)$, ..., $f_n(t)$ вещественного аргумента t, кусочно непрерывных в замкнутом интервале $[\alpha, \beta]$. Требуется определить, при каком условии они будут линейно зависимы. Для этого мы в прост-

ранстве кусочно непрерывных в $[\alpha, \beta]$ функций введем положительно определенную метрику, полагая

$$(f, g)=\int_{\alpha}^{\beta} f(t)\,\overline{g(t)}\,dt.$$

Тогда критерий Грама (теорема 1) в применении к данным функциям даст искомое условие

$$\begin{vmatrix} \int_{\alpha}^{\beta} f_1(t)\,\overline{f_1(t)}\,dt & \dots & \int_{\alpha}^{\beta} f_1(t)\,\overline{f_n(t)}\,dt \\ \dots\dots\dots\dots & & \dots\dots\dots\dots \\ \int_{\alpha}^{\beta} f_n(t)\,\overline{f_1(t)}\,dt & \dots & \int_{\alpha}^{\beta} f_n(t)\,\overline{f_n(t}\,dt \end{vmatrix}=0.$$

§ 4. Ортогональное проектирование

Пусть в унитарном или в евклидовом пространстве $\boldsymbol{R}$ даны произвольный вектор x и некоторое m-мерное подпространство $\boldsymbol{S}$ с базисом $x_1, x_2, \dots, x_m$. Мы покажем, что вектор x можно (и притом единственным способом) представить в виде суммы

$$\left.\begin{aligned} &x=x_S+x_N, \\ &x_S\in\boldsymbol{S},\ \text{а}\ x_N\perp\boldsymbol{S} \end{aligned}\right\} \qquad (15)$$

где

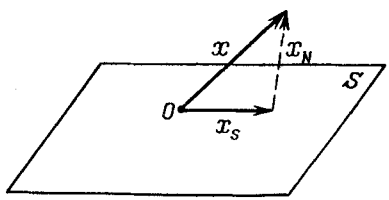


Рис. 5.

(знаком $\perp$ мы обозначаем ортогональность векторов; под ортогональностью к подпространству понимаем ортогональность ко всем векторам из этого подпространства); x_S — *ортогональная проекция вектора* x *на подпространство* $\boldsymbol{S}$, x_N — *проектирующий вектор*[1]).

Пример. $\boldsymbol{R}$ — трехмерное евклидово векторное пространство, а $m=2$. Все векторы будем строить из фиксированной точки O. Тогда S — плоскость, проходящая через O; x_S — ортогональная проекция вектора x на плоскость S; x_N — перпендикуляр, опущенный из конца вектора x на плоскость S (рис. 5); $h=|x_N|$ — расстояние конца вектора x от плоскости S.

Для установления разложения (15) искомое x_S представим в виде

$$x_S=c_1x_1+c_2x_2+\dots+c_mx_m, \qquad (16)$$

где $c_1, c_2, \dots, c_m$ — некоторые комплексные числа[2]).

Для определения этих чисел будем исходить из соотношений

$$(x-x_S, x_k)=0 \qquad (k=1, 2, \dots, m). \qquad (17)$$

Подставляя в (17) вместо x_S его выражение из (16), получим:

$$\left.\begin{aligned} &(x_1x_1)c_1+\dots+(x_mx_1)c_m+(xx_1)\cdot(-1)=0, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &(x_1x_m)c_1+\dots+(x_mx_m)c_m+(xx_m)\cdot(-1)=0, \\ &x_1c_1+\dots+\quad x_mc_m+\quad x_S\cdot(-1)=0. \end{aligned}\right\} \qquad (18)$$

[1]) В данном случае x_S — проекция вектора x на подпространство S параллельно подпространству $\boldsymbol{T}$, состоящему из всех векторов из $\boldsymbol{R}$, ортогональных к $\boldsymbol{S}$ (см. стр. 69).

[2]) В случае евклидова пространства $c_1, c_2, \dots, c_m$ — вещественные числа.

Рассматривая эту систему равенств как систему линейных однородных уравнений, имеющую ненулевое решение $c_1, c_2, \ldots, c_m, -1$, приравниваем определитель этой системы нулю (предварительно транспонировав его относительно главной диагонали)[1]):

$$\begin{vmatrix} (x_1x_1) & \ldots & (x_1x_m) & x_1 \\ \cdot & \cdot & \cdot & \cdot \\ (x_mx_1) & \ldots & (x_mx_m) & x_m \\ (xx_1) & \ldots & (xx_m) & x_S \end{vmatrix} = 0. \tag{19}$$

Выделяя из этого определителя член, содержащий x_S, получим (в легко понятных условных обозначениях):

$$x_S = -\frac{\begin{vmatrix} & & x_1 \\ & \Gamma & \cdot \\ & & x_m \\ (xx_1) & \ldots (xx_m) & 0 \end{vmatrix}}{\Gamma}, \tag{20}$$

где $\Gamma = \Gamma(x_1, x_2, \ldots, x_m)$ — определитель Грама для векторов $x_1, x_2, \ldots, x_m$ (в силу линейной независимости этих векторов $\Gamma \neq 0$). Из (15) и (20) находим:

$$x_N = x - x_S = \frac{\begin{vmatrix} & & x_1 \\ & \Gamma & \cdot \\ & & x_m \\ (xx_1) & \ldots (xx_m) & x \end{vmatrix}}{\Gamma}. \tag{21}$$

Формулы (20) и (21) выражают проекции x_S вектора x на подпространство S и проектирующий вектор x_N через данный вектор x и базис подпространства S.

Обратим внимание еще на одну важную формулу. Обозначим через h длину вектора x_N. Тогда в силу (15) и (21)

$$h^2 = (x_Nx_N) = (x_Nx) = \frac{\begin{vmatrix} & & (x_1x) \\ & \Gamma & \cdot \\ & & (x_mx) \\ (xx_1) & \ldots (xx_m) & (xx) \end{vmatrix}}{\Gamma},$$

т. е.

$$h^2 = \frac{\Gamma(x_1, x_2, \ldots, x_m, x)}{\Gamma(x_1, x_2, \ldots, x_m)}. \tag{22}$$

Величину h можно еще интерпретировать следующим образом:

Построим векторы $x_1, x_2, \ldots, x_m, x$ из одной точки и построим на этих векторах, как на ребрах, $(m+1)$-мерный параллелепипед. h будет

[1]) Определитель, стоящий в левой части равенства (19), представляет собой вектор, i-я координата которого получается, если в последнем столбце все векторы $x_1, \ldots, x_m, x_S$ заменить их i-ми координатами ($i = 1, 2, \ldots, n$); координаты берутся в некотором произвольном базисе. Для оправдания перехода от (18) к (19) достаточно в последнем равенстве (18) и в последнем столбце в (19) заменить векторы $x_1, \ldots, x_m, x_S$ их i-ми координатами.

высотой этого параллелепипеда, опущенной из конца ребра x на основание S, проходящее через ребра $x_1, x_2, \ldots, x_m$.

Пусть y — произвольный вектор в S, а x — произвольный вектор в R. Если все векторы построить из начала координат n-мерного точечного пространства, то $|x-y|$ и $|x-x_S|$ будут соответственно равны величинам наклонной и высоты, проведенным из конца вектора x к гиперплоскости S [1]). Поэтому, записывая, что *высота короче наклонной*, будем иметь [2]):

$$h=|x-x_S|\leqslant|x-y|$$

(знак равенства лишь при $y=x_S$). Таким образом, среди всех векторов $y\in S$ вектор x_S наименее уклоняется от заданного вектора $x\in R$. Величина $h=\sqrt{(x-x_S,\ x-x_S)}$ является *квадратичной погрешностью* при приближении $x\approx x_S$ [3]).

§ 5. Геометрический смысл определителя Грама и некоторые неравенства

1. Рассмотрим произвольные векторы $x_1, x_2, \ldots, x_m$. Допустим сначала, что эти векторы линейно независимы. В этом случае определитель Грама, составленный для любых из этих векторов, будет отличен от нуля. Тогда, полагая согласно (22)

$$\frac{\Gamma(x_1, x_2, \ldots, x_{p+1})}{\Gamma(x_1, x_2, \ldots, x_p)}=h_p^2>0 \quad (p=1, 2, \ldots, m-1) \tag{23}$$

и перемножая почленно эти неравенства и неравенство

$$\Gamma(x_1)=(x_1x_1)>0, \tag{24}$$

получим:

$$\Gamma(x_1, x_2, \ldots, x_m)>0.$$

Таким образом, *определитель Грама для линейно независимых векторов положителен, для линейно зависимых равен нулю. Отрицательным определитель Грама никогда не бывает.*

Обозначим для сокращения $\Gamma_p=\Gamma(x_1, x_2, \ldots, x_p)$ $(p=1, 2, \ldots, m)$. Тогда из (23) и (24)

$$\sqrt{\Gamma_1}=|x_1|=V_1,$$
$$\sqrt{\Gamma_2}=V_1h_1=V_2,$$

где V_2 — площадь параллелограмма, построенного на x_1 и x_2. Далее,

$$\sqrt{\Gamma_3}=V_2h_2=V_3,$$

где V_3 — объем параллелепипеда, построенного на векторах x_1, x_2, x_3. Продолжая далее, найдем:

$$\sqrt{\Gamma_4}=V_3h_3=V_4,$$

и, наконец,

$$\sqrt{\Gamma_m}=V_{m-1}h_{m-1}=V_m. \tag{25}$$

Естественно V_m назвать объемом m-мерного параллелепипеда, построенного на векторах $x_1, x_2, \ldots, x_m$, как на ребрах [4]).

---

[1]) Ср. с примером на стр. 227.

[2]) $|x-y|^2=|x_N+x_S-y|^2=|x_N|^2+|x_S-y|^2\geqslant|x_N|^2=h^2$.

[3]) Относительно использования метризованных функциональных пространств в задачах аппроксимации функций см. [2].

[4]) Формула (25) дает индуктивное определение объема m-мерного параллелепипеда.

Обозначим через $x_{1k}, x_{2k}, \ldots, x_{nk}$ координаты вектора x_k $(k=1, 2, \ldots, m)$ в некотором ортонормированном базисе в R, и пусть

$$X=\|x_{ik}\| \quad (i=1, 2, \ldots, n;\ k=1, 2, \ldots, m).$$

Тогда на основании (14)

$$\Gamma_m=|X'\overline{X}|$$

и потому [см. формулу (25)]

$$V_m^2=\Gamma_m=\sum_{1\leqslant i_1<i_2<\ldots<i_m\leqslant n} \mathrm{mod} \begin{vmatrix} x_{i_1 1} & x_{i_1 2} & \ldots & x_{i_1 m} \\ x_{i_2 1} & x_{i_2 2} & \ldots & x_{i_2 m} \\ \ldots & \ldots & \ldots & \ldots \\ x_{i_m 1} & x_{i_m 2} & \ldots & x_{i_m m} \end{vmatrix}^2. \tag{26}$$

Это равенство имеет следующий геометрический смысл:

Квадрат объема параллелепипеда равен сумме квадратов объемов его проекций на все координатные m-мерные подпространства. В частности, при $m=n$ из (26) следует:

$$V_n=\mathrm{mod} \begin{vmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1} & x_{n2} & \ldots & x_{nn} \end{vmatrix}. \tag{26'}$$

При помощи формул (20), (21), (22), (26), (26') решается ряд основных метрических задач n-мерной унитарной и евклидовой аналитической геометрии.

2. Вернемся к разложению (15). Из него непосредственно следует:

$$(xx)=(x_S+x_N,\ x_S+x_N)=(x_S,\ x_S)+(x_N,\ x_N)\geqslant(x_N x_N)=h^2,$$

что в сочетании с (22) дает неравенство (для произвольных векторов x_1, $x_2, \ldots, x_m, x$)

$$\Gamma(x_1, x_2, \ldots, x_m, x)\leqslant\Gamma(x_1, x_2, \ldots, x_m)\,\Gamma(x); \tag{27}$$

при этом знак равенства имеет место тогда и только тогда, когда вектор x ортогонален к векторам $x_1, x_2, \ldots, x_n$.

Отсюда нетрудно получить так называемое *неравенство Адамара*

$$\Gamma(x_1, x_2, \ldots, x_m)\leqslant\Gamma(x_1)\,\Gamma(x_2)\ldots\Gamma(x_m), \tag{28}$$

где знак равенства имеет место тогда и только тогда, когда векторы x_1, $x_2,, \ldots, x_m$ попарно ортогональны. Неравенство (29) выражает собой следующий геометрически очевидный факт:

Объем параллелепипеда не превосходит произведения длин его ребер и равен этому произведению лишь тогда, когда параллелепипед прямоугольный.

Неравенству Адамара можно придать его обычный вид, полагая в (28) $m=n$ и вводя в рассмотрение определитель Δ, составленный из координат $x_{1k}, x_{2k}, \ldots, x_{nk}$ векторов x_k $(k=1, 2, \ldots, n)$, в некотором ортонормированном базисе:

$$\Delta=\begin{vmatrix} x_{11} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots \\ x_{n1} & \ldots & x_{nn} \end{vmatrix}.$$

Тогда из (26′) и (28) следует:

$$|\Delta|^2 \leqslant \sum_{i=1}^{n} |x_{i1}|^2 \sum_{i=1}^{n} |x_{i2}|^2 \dots \sum_{i=1}^{n} |x_{in}|^2. \tag{28′}$$

3. Установим теперь обобщенное неравенство Адамара, охватывающее как неравенство (27), так и неравенство (28):

$$\Gamma(x_1, x_2, \dots, x_m) \leqslant \Gamma(x_1, \dots, x_p)\,\Gamma(x_{p+1}, \dots, x_m), \tag{29}$$

причем знак равенства имеет место тогда и только тогда, когда каждый из векторов $x_1, x_2, \dots, x_p$ ортогонален к любому из векторов $x_{p+1}, \dots, x_m$ либо один из определителей $\Gamma(x_1, \dots, x_p)$, $\Gamma(x_{p+1}, \dots, x_m)$ равен нулю.

Неравенство (28′) имеет следующий геометрический смысл:

Объем параллелепипеда не превосходит произведения объемов двух дополнительных «граней» и равен этому произведению в том и только в том случае, когда эти «грани» взаимно ортогональны либо хотя бы одна из них имеет нулевой объем.

Справедливость неравенства (29) установим индуктивно относительно числа векторов $x_{p+1}, \dots, x_m$. Неравенство справедливо, когда это число равно 1 [см. формулу (27)].

Введем в рассмотрение два подпространства S и S_1 соответственно с базисами $x_1, \dots, x_{m-1}$ и $x_{p+1}, \dots, x_{m-1}$. Очевидно, $S_1 \subset S$. Рассмотрим ортогональные разложения

$$x_m = x_{S_1} + x_{N_1} \quad (x_{S_1} \in S_1,\ x_{N_1} \perp S_1),$$
$$x_{N_1} = x'_S + x_N \quad (x'_S \in S,\ x_N \perp S).$$

Отсюда

$$x_m = x_S + x_N \quad (x_S = x_{S_1} + x'_S,\quad x_N \perp S).$$

Заменяя квадрат объема параллелепипеда произведением квадрата объема основания на квадрат высоты [см. формулу (22)], найдем:

$$\Gamma(x_1, \dots, x_{m-1}, x_m) = \Gamma(x_1, \dots, x_{m-1})\,\Gamma(x_N), \tag{30}$$
$$\Gamma(x_{p+1}, \dots, x_{m-1}, x_m) = \Gamma(x_{p+1}, \dots, x_{m-1})\,\Gamma(x_{N_1}). \tag{30′}$$

При этом из разложения вектора x_{N_1} следует:

$$\Gamma(x_N) \leqslant \Gamma(x_{N_1}), \tag{31}$$

причем здесь знак $=$ имеет место, лишь когда $x_{N_1} = x_N$.

Используя теперь соотношения (30), (30′), (31) и предположение индукции, получим:

$$\Gamma(x_1, \dots, x_m) = \Gamma(x_1, \dots, x_{m-1})\,\Gamma(x_N) \leqslant$$
$$\leqslant \Gamma(x_1, \dots, x_{m-1})\,\Gamma(x_{N_1}) \leqslant \Gamma(x_1, \dots, x_p)\,\Gamma(x_{p+1}, \dots, x_{m-1})\,\Gamma(x_{N_1}) =$$
$$= \Gamma(x_1, \dots, x_p)\,\Gamma(x_{p+1}, \dots, x_m). \tag{32}$$

Мы получили неравенство (29). Переходя к выяснению, когда в этом неравенстве имеет место знак $=$, примем, что $\Gamma(x_1, \dots, x_p) \neq 0$ и $\Gamma(x_{p+1}, \dots, x_m) \neq 0$. Тогда согласно (30′) также $\Gamma(x_{p+1}, \dots, x_{m-1}) \neq 0$ и $\Gamma(x_{N_1}) \neq 0$. Коль скоро в соотношениях (32) всюду имеет место знак равенства, то $x_{N_1} = x_N$ и, кроме того, по предположению индукции, каждый из векторов $x_{p+1}, \dots, x_{m-1}$ ортогонален к каждому из векторов $x_1, \dots, x_p$. Этим свойством обладает, очевидно, и вектор

$$x_m = x_{S_1} + x_{N_1} = x_{S_1} + x_N.$$

Таким образом, обобщенное неравенство Адамара установлено полностью.

4. Обобщенному неравенству Адамара (29) можно придать и аналитическую форму.

Пусть $\sum_{i,k=1}^{n} h_{ik}x_i\bar{x}_k$ — произвольная положительно определенная эрмитова форма. Рассматривая $x_1, x_2, \ldots, x_n$ как координаты вектора x в n-мерном пространстве $\boldsymbol{R}$ при базисе $e_1, e_2, \ldots, e_n$, примем форму $\sum_{i,k=1}^{n} h_{ik}x_i\bar{x}_k$ за основную метрическую форму в $\boldsymbol{R}$ (см. стр. 224). Тогда $\boldsymbol{R}$ станет унитарным пространством. Применим обобщенное неравенство Адамара к базисным векторам $e_1, e_2, \ldots, e_n$:

$$\Gamma(e_1, e_2, \ldots, e_n) \leqslant \Gamma(e_1, \ldots, e_p)\,\Gamma(e_{p+1}, \ldots, e_n).$$

Полагая $H = \| h_{ik} \|_1^n$ и замечая, что $(e_ie_k) = h_{ik}$ $(i, k = 1, 2, \ldots, n)$, мы последнее неравенство сможем записать так:

$$H\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leqslant H\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix} H\begin{pmatrix} p+1 & \ldots & n \\ p+1 & \ldots & n \end{pmatrix} \quad (p < n); \qquad (33)$$

при этом знак равенства имеет место в том и только в том случае, когда $h_{ik} = h_{ki} = 0$ $(i = 1, 2, \ldots, p;\ k = p+1, \ldots, n)$.

Неравенство (33) имеет место для матрицы коэффициентов $H = \| h_{ik} \|_1^n$ произвольной положительно определенной эрмитовой формы. В частности, неравенство (33) имеет место, если H — вещественная матрица коэффициентов положительно определенной квадратичной формы $\sum_{i,k=1}^{n} h_{ik}x_ix_k$ [1]).

5. Обратим внимание читателя на неравенство Буняковского. Для произвольных векторов x, $y \in \boldsymbol{R}$

$$|(xy)|^2 \leqslant (xx)(yy), \qquad (34)$$

причем знак равенства имеет место лишь тогда, когда векторы x и y отличаются скалярным множителем.

Справедливость неравенства Буняковского сразу вытекает из установленного уже неравенства

$$\Gamma(x, y) = \begin{vmatrix} (xx) & (xy) \\ (yx) & (yy) \end{vmatrix} \geqslant 0.$$

По аналогии со скалярным произведением векторов в трехмерном евклидовом пространстве в n-мерном унитарном пространстве можно ввести «угол» θ между векторами x и y, определив его из соотношения [2])

$$\cos^2\theta = \frac{|(xy)|^2}{(xx)(yy)}.$$

Из неравенства Буняковского следует, что θ имеет вещественное значение.

---

[1]) Аналитический вывод обобщенного неравенства Адамара приведен в книге [7], § 8.

[2]) В случае евклидова пространства угол θ между векторами x и y определяется из формулы

$$\cos\theta = \frac{(xy)}{|x|\,|y|}.$$

§ 6. Ортогонализация ряда векторов

1. Наименьшее подпространство, содержащее векторы $x_1, x_2, \ldots, x_p$, будем обозначать через $[x_1, x_2, \ldots, x_p]$. Это подпространство состоит из всевозможных линейных комбинаций $c_1x_1+c_2x_2+\ldots+c_px_p$ векторов $x_1, x_2, \ldots, x_p$ ($c_1, c_2, \ldots, c_p$ — комплексные числа)[1]. Если векторы $x_1, x_2, \ldots, x_p$ линейно независимы, то они образуют базис подпространства $[x_1, x_2, \ldots, x_p]$. В этом случае это подпространство имеет p измерений.

Два ряда векторов

$$\boldsymbol{X}\colon x_1, x_2, \ldots,$$

$$\boldsymbol{Y}\colon y_1, y_2, \ldots,$$

содержащих одинаковое конечное или оба бесконечное число векторов, назовем *эквивалентными*, если для всех возможных p

$$[x_1, x_2, \ldots, x_p] \equiv [y_1, y_2, \ldots, y_p] \quad (p=1, 2, \ldots).$$

Ряд векторов

$$\boldsymbol{X}\colon x_1, x_2, \ldots$$

назовем *невырожденным*, если при любом возможном p векторы $x_1, x_2, \ldots, x_p$ линейно независимы.

Ряд векторов называется *ортогональным*, если любые два вектора этого ряда взаимно ортогональны.

Под *ортогонализацией* ряда векторов будем понимать замену этого ряда эквивалентным ортогональным рядом.

Теорема 2. *Всякий невырожденный ряд векторов можно проортогонализировать. Процесс ортогонализации приводит к векторам, определенным однозначно с точностью до скалярных множителей.*

Доказательство. 1. Докажем сначала вторую часть этой теоремы. Пусть два ортогональных ряда $\boldsymbol{Y}\colon y_1, y_2, \ldots$ и $\boldsymbol{Z}\colon z_1, z_2, \ldots$ эквивалентны одному и тому же невырожденному ряду $\boldsymbol{X}\colon x_1, x_2, \ldots$ Тогда ряды $\boldsymbol{Y}$ и $\boldsymbol{Z}$ эквивалентны между собой. Поэтому при любом p существуют числа $c_{p1}, c_{p2}, \ldots, c_{pp}$ такие, что

$$z_p = c_{p1}y_1 + c_{p2}y_2 + \ldots + c_{p,\,p-1}y_{p-1} + c_{pp}y_p \quad (p=1, 2, \ldots).$$

Умножая последовательно обе части этого равенства скалярно на $y_1, y_2, \ldots, y_{p-1}$ и учитывая ортогональность ряда $\boldsymbol{Y}$ и соотношения

$$z_p \perp [z_1, z_2, \ldots, z_{p-1}] \equiv [y_1, y_2, \ldots, y_{p-1}],$$

получим: $c_{p1}=c_{p2}=\ldots=c_{p,\,p-1}=0$ и, следовательно,

$$z_p = c_{pp}y_p \quad (p=1, 2, \ldots).$$

2. Конкретное осуществление процесса ортогонализации произвольного невырожденного ряда векторов $\boldsymbol{X}\colon x_1, x_2, \ldots$ дается следующим построением.

Пусть $S_p \equiv [x_1, x_2, \ldots, x_p]$, $\Gamma_p = \Gamma(x_1, x_2, \ldots, x_p)$ $(p=1, 2, \ldots)$. Спроектируем ортогонально вектор x_p на подпространство S_{p-1} $(p=1, 2, \ldots)$[2]:

$$x_p = x_{pS_{p-1}} + x_{pN}, \quad x_{pS_{p-1}} \in S_{p-1}, \quad x_{pN} \perp S_{p-1} \quad (p=1, 2, \ldots).$$

---

[1]) В случае евклидова пространства эти числа вещественны.

[2]) При $p=1$ мы полагаем: $x_{1S_0}=0$, $x_{1N}=x_1$.

Положим

$$y_p = \lambda_p x_{pN} \quad (p = 1, 2, \ldots; \ x_{1N} = x_1),$$

где λ_p $(p = 1, 2, \ldots)$ — произвольные отличные от нуля числа.

Тогда (как легко видеть)

$$Y: \ y_1, \ y_2, \ldots$$

— ортогональный ряд, эквивалентный ряду X. Теорема 2 доказана.

Согласно (21)

$$x_{pN} = \frac{\begin{vmatrix} & & x_1 \\ & \Gamma & \vdots \\ & & x_{p-1} \\ (x_p x_1) & \ldots & (x_p x_{p-1}) \ x_p \end{vmatrix}}{\Gamma_{p-1}} \quad (p = 1, 2, \ldots; \ \Gamma_0 = 1).$$

Полагая $\lambda_p = \Gamma_{p-1}$ $(p = 1, 2, \ldots; \ \Gamma_0 = 1)$, получим для векторов проортогонализированного ряда следующие формулы:

$$y_1 = x_1, \ y_2 = \begin{vmatrix} (x_1 x_1) & x_1 \\ (x_2 x_1) & x_2 \end{vmatrix}, \ \ldots, \ y_p = \begin{vmatrix} (x_1 x_1) & \ldots & (x_1 x_{p-1}) & x_1 \\ \ldots & \ldots & \ldots & \ldots \\ (x_{p-1} x_1) & \ldots & (x_{p-1} x_{p-1}) & x_{p-1} \\ (x_p x_1) & \ldots & (x_p x_{p-1}) & x_p \end{vmatrix}, \ \ldots \quad (35)$$

В силу (22)

$$(y_p y_p) = \Gamma_{p-1}^2 \, |x_{pN}| = \Gamma_{p-1}^2 \cdot \frac{\Gamma_p}{\Gamma_{p-1}} = \Gamma_{p-1} \Gamma_p \quad (p = 1, 2, \ldots; \ \Gamma_0 = 1). \quad (36)$$

Поэтому, полагая

$$z_p = \frac{y_p}{\sqrt{\Gamma_{p-1} \Gamma_p}} \quad (p = 1, 2, \ldots), \quad (37)$$

получим ортонормированный ряд Z, эквивалентный данному ряду X.

Пример. Определим скалярное произведение в пространстве кусочно непрерывных в интервале $[-1, +1]$ вещественных функций равенством

$$(f, g) = \int_{-1}^{+1} f(x) \, g(x) \, dx.$$

Рассмотрим невырожденный ряд «векторов»

$$1, \ x, \ x^2, \ x^3, \ldots$$

Проортогонализируем его по формулам (35):

$$y_0 \equiv 1, \ y_m = \begin{vmatrix} \frac{1}{1} & 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \ldots & 1 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \ldots & x \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \ldots & x^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & x^m \end{vmatrix} \quad (m = 1, 2, \ldots).$$

Эти ортогональные между собой многочлены с точностью до постоянных множителей совпадают с известными многочленами Лежандра [1]):

$$P_0(x)=1, \quad P_m(x)=\frac{1}{2^m m!}\frac{d^m(x^2-1)^m}{dx^m} \quad (m=1, 2, \ldots).$$

Тот же ряд степеней $1, x, x^2, \ldots$ при другой метрике

$$(f, g)=\int_a^b f(x)\, g(x)\, \tau(x)\, dx \quad [\tau(x)\geqslant 0 \text{ при } a\leqslant x\leqslant b]$$

даст другой ряд ортогональных многочленов.

Так, например, при $a=-1$, $b=1$ и $\tau(x)=\frac{1}{\sqrt{1-x^2}}$ получаются многочлены Чебышева:

$$T_n(x)=\frac{1}{2^{n-1}}\cos(n \arccos x).$$

При $a=-\infty$, $b=+\infty$ и $\tau(x)=e^{-x^2}$ получаются многочлены Чебышева—Эрмита и т. д. [2]).

2. Отметим еще так называемое неравенство Бесселя для ортонормированного ряда векторов $\boldsymbol{Z}$: $z_1, z_2, \ldots$ Пусть дан произвольный вектор x. Обозначим через ξ_p проекцию этого вектора на орт ε_p:

$$\xi_p=(x\varepsilon_p) \qquad (p=1, 2, \ldots).$$

Тогда проекция вектора x на подпространство $\boldsymbol{S}_p=[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p]$ представится в виде [см. (20)]

$$x_{s_p}=\xi_1\varepsilon_1+\xi_2\varepsilon_2+\ldots+\xi_p\varepsilon_p \qquad (p=1, 2, \ldots).$$

Но $|x_{s_p}|^2=|\xi_1|^2+|\xi_2|^2+\ldots+|\xi_p|^2\leqslant|x|^2$. Поэтому для произвольного p

$$|\xi_1|^2+|\xi_2|^2+\ldots+|\xi_p|^2\leqslant|x|^2. \tag{38}$$

Это—неравенство Бесселя.

В случае конечномерного пространства n измерений это неравенство имеет совершенно очевидный геометрический смысл. При $p=n$ оно переходит в равенство Пифагора

$$|\xi_1|^2+|\xi_2|^2+\ldots+|\xi_n|^2=|x|^2.$$

В случае бесконечномерного пространства и бесконечного ряда $\boldsymbol{Z}$ из (38) следуют сходимость ряда $\sum_{k=1}^{\infty}|\xi_k|^2$ и неравенство

$$\sum_{k=1}^{\infty}|\xi_k|^2\leqslant|x|^2.$$

Составим ряд

$$\sum_{k=1}^{\infty}\xi_k\varepsilon_k.$$

p-й отрезок этого ряда (при любом p)

$$\xi_1\varepsilon_1+\xi_2\varepsilon_2+\ldots+\xi_p\varepsilon_p$$

---

[1]) См. [19], стр. 77 и далее.

[2]) Более подробно об этом см. [19], гл. II, § 9, а также [9].

равен проекции x_{S_p} вектора x на подпространство $S_p=[z_1, z_2, \ldots, z_p]$ и потому является наилучшим приближением для вектора x в этом подпространстве:

$$\left|(x-\sum_{k=1}^{p}\xi_k z_k)\right|\leqslant\left|(x-\sum_{k=1}^{p}c_k z_k)\right|,$$

где $c_1, c_2, \ldots, c_p$ — произвольные комплексные числа. Вычислим соответствующее квадратичное отклонение δ_p:

$$\delta_p^2=\left|(x-\sum_{k=1}^{p}\xi_k z_k)\right|^2=(x-\sum_{k=1}^{p}\xi_k z_k,\quad x-\sum_{k=1}^{p}\xi_k z_k)=|x|^2-\sum_{k=1}^{p}|\xi_k|^2.$$

Отсюда

$$\lim_{p\to\infty}\delta_p^2=|x|^2-\sum_{k=1}^{\infty}|\xi_k|^2.$$

Если

$$\lim_{p\to\infty}\delta_p=0,$$

то говорят, что ряд $\sum_{k=1}^{\infty}\xi_k z_k$ *сходится в среднем* (сходится по норме) к вектору x.

В этом случае для вектора x из R имеет место равенство (теорема Пифагора в бесконечномерном пространстве!)

$$(xx)=|x|^2=\sum_{k=1}^{\infty}|\xi_k|^2. \tag{39}$$

Если для любого вектора x из R ряд $\sum_{k=1}^{p}\xi_k z_k$ в среднем сходится к вектору x, то ортонормированный ряд векторов $z_1, z_2, \ldots$ называется *полным*. В этом случае, заменяя в (39) x на $x+y$ и используя равенство (39) трижды, для векторов $x+y$, x и y мы легко получим:

$$(xy)=\sum_{k=1}^{\infty}\xi_k\bar{\eta}_k\qquad[\xi_k=(xz_k),\eta_k=(yz_k);\ k=1,\ 2,\ \ldots]. \tag{40}$$

Пример. Рассмотрим пространство всех комплексных функций $f(t)$ (t — вещественный аргумент), кусочно непрерывных в замкнутом интервале $[0, 2\pi]$. Скалярное произведение двух функций $f(t)$ и $g(t)$ определим формулой

$$(f, g)=\int_0^{2\pi}f(t)\,\overline{g(t)}\,dt.$$

В частности,

$$(f, f)=\int_0^{2\pi}|f(t)|^2\,dt.$$

Возьмем бесконечную последовательность функций

$$\frac{1}{\sqrt{2\pi}}e^{ikt}\qquad(k=0, \pm1, \pm2, \ldots).$$

Эти функции образуют ортонормированный ряд, так как

$$\int_0^{2\pi} e^{i\mu t} e^{-i\nu t}\,dt = \int_0^{2\pi} e^{i(\mu-\nu)t}\,dt = \begin{cases} 0 & \text{при } \mu \neq \nu, \\ 2\pi & \text{при } \mu = \nu. \end{cases}$$

Ряд

$$\sum_{k=-\infty}^{\infty} f_k e^{ikt} \qquad \left(f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, e^{-ikt}\,dt; \quad k = 0, \pm 1, \pm 2, \ldots \right)$$

сходится в среднем к функции $f(t)$ в интервале $[0, 2\pi]$. Этот ряд называется *рядом Фурье* для функции $f(t)$, а коэффициенты f_k $(k = 0, \pm 1, \pm 2, \ldots)$ — коэффициентами Фурье для $f(t)$.

В теории рядов Фурье доказывается, что система функций e^{ikt} $(k = 0, \pm 1, \pm 2, \ldots)$ является полной [1]).

Условие полноты дает *равенство Парсеваля* [см. равенство (40)]

$$\int_0^{2\pi} f(t)\,\overline{g(t)}\,dt = \sum_{k=-\infty}^{+\infty} \frac{1}{2\pi} \int_0^{2\pi} f(t)\, e^{-ikt}\,dt \int_0^{2\pi} \overline{g(t)}\, e^{ikt}\,dt.$$

Если $f(t)$ — вещественная функция, то f_0 вещественно, а f_k и f_{-k} — комплексно сопряженные числа $(k = 1, 2, \ldots)$. Полагая

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, e^{-ikt}\,dt = \frac{1}{2}(a_k - ib_k),$$

где

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt\,dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt\,dt \quad (k = 0, 1, 2, \ldots),$$

будем иметь:

$$f_k e^{ikt} + f_{-k} e^{-ikt} = a_k \cos kt + b_k \sin kt \qquad (k = 1, 2, \ldots).$$

Поэтому для вещественной функции $f(t)$ ряд Фурье принимает вид

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) \left(\begin{array}{l} a_k = \dfrac{1}{\pi} \displaystyle\int_0^{2\pi} f(t) \cos kt\,dt, \\ \\ b_k = \dfrac{1}{\pi} \displaystyle\int_0^{2\pi} f(t) \sin kt\,dt, \end{array} \quad k = 0, 1, 2, \ldots \right).$$

§ 7. Ортонормированный базис

Базис любого конечномерного подпространства $\boldsymbol{S}$ в унитарном или евклидовом пространстве $\boldsymbol{R}$ является невырожденным рядом векторов и потому согласно теореме 2 предыдущего параграфа может быть проортогонализирован и пронормирован. Таким образом, *в любом конечномерном подпространстве* $\boldsymbol{S}$ (*и, в частности, во всем пространстве* $\boldsymbol{R}$, *если оно конечномерно*) *существует ортонормированный базис.*

Пусть $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ — ортонормированный базис пространства $\boldsymbol{R}$. Обозначим через $x_1, x_2, \ldots, x_n$ координаты произвольного вектора $\boldsymbol{x}$ в этом базисе:

$$\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k.$$

[1]) См., например, [19], гл. II.

Умножая обе части этого равенства справа на e_k и учитывая ортонормированность базиса, легко найдем:

$$x_k = (xe_k) \qquad (k = 1, 2, \ldots, n),$$

т. е. *в ортонормированном базисе координата вектора равна скалярному произведению его на соответствующий базисный орт*

$$x = \sum_{k=1}^{n} (xe_k)\, e_k. \tag{41}$$

Пусть $x_1, x_2, \ldots, x_n$ и $x'_1, x'_2, \ldots, x'_n$ суть соответственно координаты одного и того же вектора x в двух различных ортонормированных базисах $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ унитарного пространства $\boldsymbol{R}$. Формулы преобразования координат имеют вид

$$x_i = \sum_{k=1}^{n} u_{ik} x'_k \qquad (i = 1, 2, \ldots, n). \tag{42}$$

При этом коэффициенты $u_{1k}, u_{2k}, \ldots, u_{nk}$, образующие k-й столбец матрицы $U = \| u_{ik} \|_1^n$, являются, как нетрудно видеть, координатами вектора e'_k в базисе $e_1, e_2, \ldots, e_n$. Поэтому, записывая в координатах [см. (10)] условия ортонормированности базиса $e'_1, e'_2, \ldots, e'_n$, получим соотношения

$$\sum_{i=1}^{n} u_{ik} \overline{u_{il}} = \delta_{kl} = \begin{cases} 1 & k = l, \\ 0 & k \neq l. \end{cases} \tag{43}$$

Преобразование (42), у которого коэффициенты удовлетворяют условию (43), называется *унитарным*, а соответствующая матрица U — *унитарной матрицей*. Таким образом, *в n-мерном унитарном пространстве переход от одного ортонормированного базиса к другому ортонормированному осуществляется при помощи унитарного преобразования координат.*

Пусть дано n-мерное евклидово пространство $\boldsymbol{R}$. Переход от одного ортонормированного базиса в $\boldsymbol{R}$ к другому осуществляется при помощи преобразования координат

$$x_i = \sum_{k=1}^{n} v_{ik} x'_k \qquad (i = 1, 2, \ldots, n), \tag{44}$$

коэффициенты которого связаны между собой соотношениями

$$\sum_{i=1}^{n} v_{ik} v_{il} = \delta_{kl} \qquad (k, l = 1, 2, \ldots, n). \tag{45}$$

Такое преобразование координат называется *ортогональным*, а соответствующая матрица V — *ортогональной матрицей*.

Отметим интересную матричную запись процесса ортогонализации. Пусть $A = \| a_{ik} \|_1^n$ — произвольная неособенная матрица ($|A| \neq 0$) с комплексными элементами. Рассмотрим унитарное пространство $\boldsymbol{R}$ с ортонормированным базисом $e_1, e_2, \ldots, e_n$ и определим линейно независимые векторы $a_1, a_2, \ldots, a_n$ равенством

$$a_k = \sum_{i=1}^{n} a_{ik} e_i \qquad (k = 1, 2, \ldots, n).$$

Подвергнем векторы $a_1, a_2, \ldots, a_n$ процессу ортогонализации. Полученный ортонормированный базис в $\boldsymbol{R}$ обозначим через $u_1, u_2, \ldots, u_n$. Пусть при этом

$$u_k = \sum_{i=1}^{n} u_{ik} e_i \qquad (k = 1, 2, \ldots, n).$$

Тогда

$$[a_1 a_2 \dots a_p] = [u_1 u_2 \dots u_p] \qquad (p=1, 2, \dots, n),$$

т. е.

$$\begin{aligned} a_1 &= c_{11} u_1, \\ a_2 &= c_{12} u_1 + c_{22} u_2, \\ &\dots\dots\dots \\ a_n &= c_{1n} u_1 + c_{2n} u_2 + \dots + c_{nn} u_n, \end{aligned}$$

где c_{ik} $(i, k=1, 2, \dots, n;\ i \leqslant k)$ — некоторые комплексные числа.

Полагая $c_{ik}=0$ при $i>k$ $(i, k=1, 2, \dots, n)$, будем иметь:

$$a_k = \sum_{p=1}^{n} c_{pk} u_p \qquad (k=1, 2, \dots, n).$$

Переходя здесь к координатам и вводя верхнюю треугольную матрицу $C=\| c_{ik} \|_1^n$ и унитарную матрицу $U=\| u_{ik} \|_1^n$, получим:

$$a_{ik} = \sum_{p=1}^{n} u_{ip} c_{pk} \qquad (i, k=1, 2, \dots, n),$$

или

$$A = UC. \qquad (*)$$

Согласно этой формуле *произвольная неособенная матрица* $A=\| a_{ik} \|_1^n$ *представима в виде произведения унитарной матрицы* U *на верхнюю треугольную* C.

Так как процесс ортогонализации однозначно определяет векторы $u_1, u_2, \dots, u_n$ с точностью до скалярных множителей $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ $(|\varepsilon_i|=1;\ i=1, 2, \dots, n)$, то в формуле (*) множители U и C определяются однозначно с точностью до диагонального множителя $M=\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$:

$$U = U_1 M, \quad C = M^{-1} C_1.$$

В этом можно убедиться и непосредственно.

Замечание 1. Если A — вещественная матрица, то в формуле (*) множители U и C можно выбрать вещественными. В этом случае U — ортогональная матрица.

Замечание 2. Формула (*) сохраняет свою силу и для особенной матрицы A $(|A|=0)$. В этом можно убедиться, полагая $A = \lim\limits_{m\to\infty} A_m$, где $|A_m| \neq 0$ $(m=1, 2, \dots)$.

Тогда $A_m = U_m C_m$ $(m=1, 2, \dots)$. Выделяя из последовательности U_m сходящуюся подпоследовательность U_{m_p} $(\lim\limits_{p\to\infty} U_{m_p} = U)$ и переходя к пределу, из равенства $A_{m_p} = U_{m_p} C_{m_p}$ при $p \to \infty$, получим искомое разложение $A=UC$. Однако в случае $|A|=0$ множители U и C уже не определяются однозначно с точностью до диагонального множителя M.

Замечание 3. Вместо (*) можно получить формулу

$$A = DW, \qquad (**)$$

где D — нижняя треугольная, а W — унитарная матрица. Действительно, применяя установленную ранее формулу (*) к транспонированной матрице A'

$$A' = UC$$

и полагая $W=U'$, $D=C'$, получим (**) [1]).

§ 8. Сопряженный оператор

Пусть в n-мерном унитарном пространстве $\boldsymbol{R}$ задан произвольный линейный оператор.

Определение 4. Линейный оператор $\boldsymbol{A}^*$ называется *сопряженным* по отношению к оператору $\boldsymbol{A}$ в том и только в том случае, если

[1]) Из унитарности матрицы U следует унитарность и матрицы U', так как условия унитарности (43), записанные в матричном виде: $U'\overline{U}=E$, влекут: $U\overline{U}'=E$.

для любых двух векторов x, y из R выполняется равенство

$$(Ax, y) = (x, A^*y). \tag{46}$$

Мы докажем, что для каждого линейного оператора A существует сопряженный оператор A^* и притом только один. Для доказательства выберем в R некоторый ортонормированный базис $e_1, e_2, \ldots, e_n$. Тогда [см. (41)] для искомого оператора A^* и произвольного вектора y из R должно выполняться равенство

$$A^*y = \sum_{k=1}^{n} (A^*y, e_k) e_k.$$

В силу (46) это равенство может быть переписано так:

$$A^*y = \sum_{k=1}^{n} (y, Ae_k) e_k. \tag{47}$$

Примем теперь равенство (47) за определение оператора A^*.

Легко проверить, что определенный таким образом оператор A^* является линейным и удовлетворяет равенству (46) при произвольных векторах x и y из R. Кроме того, равенство (47) однозначно определяет оператор A^*. Таким образом, устанавливаются существование и единственность сопряженного оператора A^*.

Пусть A — линейный оператор в унитарном пространстве, а $A = \|a_{ik}\|_1^n$ — матрица, отвечающая этому оператору в ортонормированном базисе $e_1, e_2, \ldots, e_n$. Тогда, применяя формулу (41) к вектору $Ae_k = \sum_{i=1}^{n} a_{ik}e_i$, получим:

$$a_{ik} = (Ae_k, e_i) \quad (i, k = 1, 2, \ldots, n). \tag{48}$$

Пусть теперь сопряженному оператору A^* в этом же базисе отвечает матрица $A^* = \|a_{ik}^*\|_1^n$. Тогда по формуле (48)

$$a_{ik}^* = (A^*e_k, e_i) \quad (i, k = 1, 2, \ldots, n). \tag{49}$$

Из (48) и (49) в силу (46) следует:

$$a_{ik}^* = \overline{a}_{ki} \quad (i, k = 1, 2, \ldots, n),$$

т. е.

$$A^* = \overline{A}'.$$

Матрица A^* является транспонированной и комплексно сопряженной для A. Такую матрицу принято называть (см. главу I) *сопряженной* по отношению к A.

Таким образом, *в ортонормированном базисе сопряженным операторам отвечают сопряженные матрицы.*

Из определения сопряженного оператора вытекают следующие его свойства:

1° $(A^*)^* = A$,

2° $(A + B)^* = A^* + B^*$,

3° $(\alpha A)^* = \alpha A^*$ (α — скаляр),

4° $(AB)^* = B^*A^*$.

Введем теперь одно важное понятие. Пусть S — произвольное подпространство в R. Обозначим через T совокупность всех векторов y из R, ортогональных к S. Легко видеть, что T есть тоже подпространство в R и что каждый вектор x из R однозначно представляется в виде суммы $x = x_S + x_T$, где $x_S \in S$, $x_T \in T$, т. е. имеет место расщепление

$$R = S + T, \quad S \perp T.$$

Это расщепление получаем, применяя к произвольному вектору x из R разложение (15) предыдущего параграфа. T называется *ортогональным дополнением* к S. Очевидно, S будет ортогональным дополнением к T. Мы пишем $S \perp T$, понимая под этим то, что любой вектор из S ортогонален любому вектору из T.

Теперь мы сможем сформулировать фундаментальное свойство сопряженного оператора:

5° *Если некоторое подпространство S инвариантно относительно A, то ортогональное дополнение T этого подпространства будет инвариантно относительно A^*.*

Действительно, пусть $x \in S$, $y \in T$. Тогда из $Ax \in S$ следует $(Ax, y) = 0$ и отсюда в силу (46) $(x, A^*y) = 0$. Так как x — произвольный вектор из S, то $A^*y \in T$, что и требовалось доказать.

Введем следующее определение:

Определение 5. Две системы векторов $x_1, x_2, \ldots, x_m$ и $y_1, y_2, \ldots, y_m$ назовем *биортонормированными*, если

$$(x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, m), \tag{50}$$

где δ_{ik} — символ Кронекера.

Теперь докажем следующее предложение:

6° *Если A — линейный оператор простой структуры, то сопряженный оператор A^* также имеет простую структуру, причем можно так выбрать полные системы собственных векторов $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$ операторов A и A^*, чтобы они были биортонормированы:*

$$Ax_i = \lambda_i x_i, \quad A^* y_i = \mu_i y_i, \quad (x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n).$$

Действительно, пусть $x_1, x_2, \ldots, x_n$ — полная система собственных векторов оператора A. Введем обозначение

$$S_k = [x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n] \quad (k = 1, 2, \ldots, n).$$

Рассмотрим одномерное ортогональное дополнение $T_k = [y_k]$ к $(n-1)$-мерному подпространству S_k $(k = 1, 2, \ldots, n)$. Тогда T_k инвариантно относительно A^*:

$$A^* y_k = \mu_k y_k, \quad y_k \neq 0 \quad (k = 1, 2, \ldots, n).$$

Из $S_k \perp y_k$ следует: $(x_k y_k) \neq 0$, так как в противном случае вектор y_k должен был бы равняться нулю. Помножая x_k, y_k $(k = 1, 2, \ldots, n)$ на надлежащие числовые множители, получим:

$$(x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n).$$

Из биортонормированности систем векторов $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$ следует, что векторы каждой из этих систем линейно независимы.

Отметим еще такое предложение:

7° *Если операторы A и A^* имеют общий собственный вектор, то характеристические числа этих операторов, отвечающие общему собственному вектору, комплексно сопряжены.*

В самом деле, пусть $Ax = \lambda x$, $A^*x = \mu x$ $(x \neq 0)$. Тогда, полагая в (46) $y = x$, будем иметь $\lambda(x, x) = \bar{\mu}(x, x)$, откуда $\lambda = \bar{\mu}$.

8° Пусть y — собственный вектор оператора A^*, и пусть $S^{(n-1)}$ — ортогональное дополнение к одномерному подпространству $T = [y]$. Поскольку $A = (A^*)^*$, то согласно утверждению 5° подпространство $S^{(n-1)}$ инвариантно относительно оператора A. Таким образом, *у всякого линейного оператора в n-мерном унитарном пространстве существует $(n-1)$-мерное инвариантное подпространство.*

Рассматривая далее оператор A в подпространстве $S^{(n-1)}$, мы сможем указать на основании установленного предложения $(n-2)$-мерное инвариантное подпространство $S^{(n-2)}$ оператора A, принадлежащее $S^{(n-1)}$. Повторяя рассуждение, мы построим цепочку из n последовательно вложенных инвариантных подпространств оператора A (индекс наверху указывает размерность):

$$S^{(1)} \subset S^{(2)} \subset \ldots \subset S^{(n-1)} \subset S^{(n)} = R.$$

Пусть теперь e_1 — нормированный вектор, принадлежащий $S^{(1)}$. Выберем в $S^{(2)}$ нормированный вектор e_2 такой, что $(e_1, e_2) = 0$. В $S^{(3)}$ найдем нормированный вектор e_3 такой, что $(e_1, e_3) = 0$ и $(e_2, e_3) = 0$. Продолжая этот процесс, мы построим ортонормированный базис векторов

$$e_1, e_2, \ldots, e_n,$$

обладающий тем свойством, что каждое подпространство, натянутое на первые k базисных векторов

$$S^{(k)} = [e_1, e_2, \ldots, e_k] \qquad (k = 1, 2, \ldots, n),$$

инвариантно относительно оператора A.

Пусть теперь $\| a_{ij} \|_1^n$ — матрица оператора A в построенном базисе. Мы имеем $Ae_j = \sum_{i=1}^{n} a_{ij} e_i$, где $a_{ij} = (Ae_j, e_i)$. Поскольку Ae_j принадлежит $S^{(j)}$, то при $i > j$ $a_{ij} = (Ae_j, e_i) = 0$ и, следовательно, матрица оператора является верхней треугольной. Мы пришли к следующей теореме:

Для любого линейного оператора A в n-мерном унитарном пространстве можно построить ортонормированный базис, в котором матрица этого оператора является треугольной.

Это предложение принято называть теоремой Шура. Разумеется, привлекая общую теорему о приведении матрицы оператора к жордановой форме, легко доказать теорему Шура последовательной ортогонализацией жорданова базиса. Приведенное доказательство по существу использует лишь существование у линейного оператора, действующего в n-мерном унитарном пространстве, собственного вектора.

§ 9. Нормальные операторы в унитарном пространстве

Определение 6. Линейный оператор A называется *нормальным*, если он перестановочен со своим сопряженным:

$$AA^* = A^*A. \tag{51}$$

Определение 7. Линейный оператор H называется *эрмитовым*, если он равен своему сопряженному:

$$H^* = H. \tag{52}$$

Определение 8. Линейный оператор U называется *унитарным*, если он обратен своему сопряженному:

$$UU^* = E. \tag{53}$$

Заметим, что унитарный оператор можно определить как изометричный оператор в эрмитовом пространстве, т. е. как оператор, сохраняющий метрику.

Действительно, пусть при произвольных векторах x и y из R

$$(Ux, Uy) = (x, y). \tag{54}$$

Тогда согласно (46) $(U^*Ux, y) = (x, y)$ и, следовательно, в силу произвольности вектора y $U^*Ux = x$, т. е. $U^*U = E$ или $U^* = U^{-1}$. Обратно, из (53) следует (54).

Из (53) или (54) вытекает, что 1° произведение двух унитарных операторов есть снова унитарный оператор, 2° единичный оператор E является унитарным и 3° обратный оператор для унитарного есть также унитарный оператор. Поэтому совокупность всех унитарных операторов является группой[1]). Эту группу называют *унитарной группой*.

Эрмитов оператор и унитарный оператор являются частными видами нормального оператора.

Теорема 3. *Произвольный линейный оператор A всегда можно представить в виде*

$$A = H_1 + iH_2, \tag{55}$$

где H_1 и H_2 — эрмитовы операторы («эрмитовы компоненты» оператора A). Эрмитовы компоненты однозначно определяются заданием оператора A. Оператор A нормален тогда и только тогда, когда его эрмитовы компоненты H_1 и H_2 перестановочны между собой.

Доказательство. Пусть имеет место (55). Тогда

$$A^* = H_1 - iH_2. \tag{56}$$

Из (55) и (56) находим:

$$H_1 = \frac{1}{2}(A + A^*), \qquad H_2 = \frac{1}{2i}(A - A^*). \tag{57}$$

Обратно, формулы (57) определяют эрмитовы операторы H_1 и H_2, связанные с A равенством (55).

Пусть теперь A — нормальный оператор: $AA^* = A^*A$. Тогда из (57) следует: $H_1H_2 = H_2H_1$. Обратно, из $H_1H_2 = H_2H_1$ в силу (55) и (56) следует: $AA^* = A^*A$. Теорема доказана.

Представление произвольного линейного оператора A в виде (55) является аналогом представления произвольного комплексного числа z в виде $x_1 + ix_2$, где x_1 и x_2 — вещественные числа.

[1]) См. сноску [5]) к стр. 28.

Пусть в некотором ортонормированном базисе операторам $\boldsymbol{A}$, $\boldsymbol{H}$ и $\boldsymbol{U}$ отвечают соответственно матрицы A, H, U. Тогда операторным равенствам

$$\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A}, \quad \boldsymbol{H}^* = \boldsymbol{H}, \quad \boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{E} \tag{58}$$

будут соответствовать матричные равенства

$$AA^* = A^*A, \quad H^* = H, \quad UU^* = E. \tag{59}$$

Поэтому мы и определяем *нормальную матрицу* как матрицу, перестановочную со своей сопряженной, *эрмитову* как равную своей сопряженной и, наконец, *унитарную* как обратную своей сопряженной.

Тогда *в ортонормированном базисе нормальному (эрмитову, унитарному) оператору отвечает соответственно нормальная (эрмитова, унитарная) матрица.*

Эрмитова матрица $H = \| h_{ik} \|_1^n$ в силу (59) характеризуется следующими соотношениями между элементами:

$$h_{ki} = \overline{h}_{ik} \qquad (i, k = 1, 2, \ldots, n),$$

т. е. эрмитова матрица всегда является матрицей коэффициентов некоторой эрмитовой формы (см. § 1).

Унитарная матрица $U = \| u_{ik} \|_1^n$ в силу (59) характеризуется следующими сотношениями между элементами:

$$\sum_{j=1}^{n} u_{ij}\overline{u}_{kj} = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n). \tag{60}$$

Так как из $UU^* = E$ следует $U^*U = E$, то из (60) следуют эквивалентные соотношения

$$\sum_{j=1}^{n} u_{ji}\overline{u}_{jk} = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n). \tag{61}$$

Равенства (60) выражают собой «ортонормированность» строк, а равенства (61) — ортонормированность столбцов в матрице $U = \| u_{ik} \|_1^n$ [1]).

Унитарная матрица является матрицей коэффициентов некоторого унитарного преобразования (см. § 7).

Оператор $\boldsymbol{P}$, осуществляющий ортогональное проектирование векторов унитарного пространства $\boldsymbol{R}$ на заданное подпространство $\boldsymbol{S}$, является эрмитовым проекционным оператором.

Действительно, этот оператор является проекционным, т. е. $\boldsymbol{P}^2 = \boldsymbol{P}$ (см. гл. III, § 6). Далее, из ортогональности векторов $\boldsymbol{x}_S = \boldsymbol{P}\boldsymbol{x}$ и $\boldsymbol{y} - \boldsymbol{y}_S = (\boldsymbol{E} - \boldsymbol{P})\boldsymbol{y}$ $(\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{R})$ следует:

$$0 = (\boldsymbol{P}\boldsymbol{x}, (\boldsymbol{E} - \boldsymbol{P})\boldsymbol{y}) = ((\boldsymbol{E} - \boldsymbol{P}^*)\boldsymbol{P}\boldsymbol{x}, \boldsymbol{y}).$$

Отсюда в силу произвольности векторов $\boldsymbol{x}$, $\boldsymbol{y}$

$$(\boldsymbol{E} - \boldsymbol{P}^*)\boldsymbol{P} = 0,$$

т. е. $\boldsymbol{P} = \boldsymbol{P}^*\boldsymbol{P}$. Из этого равенства следует, что $\boldsymbol{P}$ — эрмитов оператор, так как $(\boldsymbol{P}^*\boldsymbol{P})^* = \boldsymbol{P}^*\boldsymbol{P}$.

---

[1]) Таким образом, ортонормированность столбцов в матрице U является следствием ортонормированности строк и наоборот.

§ 10. Спектр нормальных, эрмитовых, унитарных операторов

Установим предварительно одно свойство перестановочных операторов, сформулировав его в виде леммы.

Лемма 1. *Перестановочные операторы* A *и* B ($AB=BA$) *всегда имеет общий собственный вектор.*

Доказательство. Пусть x есть собственный вектор оператора A: $Ax=\lambda x$, $x\neq 0$. Тогда в силу перестановочности операторов A и B

$$AB^k x=\lambda B^k x \qquad (k=0,\ 1,\ 2,\ \ldots). \tag{62}$$

Пусть в ряду векторов

$$x,\ Bx,\ B^2x,\ \ldots$$

первые p векторов линейно независимы, в то время как $(p+1)$-й вектор B^px является уже линейной комбинацией предыдущих. Тогда подпространство $S\equiv[x,\ Bx,\ \ldots,\ B^{p-1}x]$ будет инвариантно относительно B и потому в этом подпространстве S будет существовать собственный вектор y оператора B: $By=\mu y$, $y\neq 0$. С другой стороны, равенства (62) показывают, что векторы x, Bx, $\ldots$, $B^{p-1}x$ являются собственными векторами оператора A, отвечающими одному и тому же характеристическому числу λ. Поэтому и любая линейная комбинация этих векторов, в частности вектор y, будет собственным вектором оператора A, отвечающим характеристическому числу λ. Таким образом, доказано существование общего собственного вектора операторов A и B.

Пусть A — произвольный нормальный оператор в n-мерном эрмитовом пространстве R. В этом случае операторы A и A^* перестановочны между собой и потому имеют общий собственный вектор x_1. Тогда (см. § 8, 7°)

$$Ax_1=\lambda_1 x_1,\quad A^*x_1=\bar{\lambda}_1 x_1 \quad (x_1\neq 0).$$

Обозначим через S_1 одномерное подпространство, содержащее вектор x_1 ($S_1=[x_1]$), а через T_1 — ортогональное дополнение для S_1 в R:

$$R=S_1+T_1,\quad S_1\perp T_1.$$

Так как S_1 инвариантно относительно A и A^*, то (см. § 8, 5°) T_1 также инвариантно относительно этих операторов. Поэтому перестановочные операторы A и A^* имеют согласно лемме 1 общий собственный вектор x_2 в T_1:

$$Ax_2=\lambda_2 x_2,\quad A^*x_2=\bar{\lambda}_2 x_2 \quad (x_2\neq 0).$$

Очевидно, $x_1\perp x_2$. Полагая $S_2=[x_1,\ x_2]$ и

$$R=S_2+T_2,\quad S_2\perp T_2,$$

мы аналогичными соображениями установим существование в T_2 общего собственного вектора x_3 операторов A и A^*. Очевидно, $x_1\perp x_3$ и $x_2\perp x_3$. Продолжая этот процесс далее, мы получим n попарно ортогональных общих собственных векторов x_1, x_2, $\ldots$, x_n операторов A и A^*:

$$\left.\begin{aligned} &Ax_k=\lambda_k x_k,\quad A^*x_k=\bar{\lambda}_k x_k \quad (x_k\neq 0),\\ &(x_i x_k)=0 \quad \text{при} \quad i\neq k \end{aligned}\right\} \qquad (i,\ k=1,\ 2,\ \ldots,\ n). \tag{63}$$

Векторы x_1, x_2, $\ldots$, x_n можно пронормировать. При этом равенства (63) сохранятся.

Таким образом, мы доказали, что нормальный оператор всегда имеет полную ортонормированную[1]) систему собственных векторов.

Так как из $\lambda_k = \lambda_l$ всегда следует $\overline{\lambda}_k = \overline{\lambda}_l$, то из равенств (63) вытекает:

1° *Если оператор A нормален, то каждый собственный вектор оператора A является собственным вектором сопряженного оператора A^**, т. е. если оператор A нормален, то операторы A и A^* имеют одни и те же собственные векторы.

Пусть теперь, обратно, дано, что линейный оператор A имеет полную ортонормированную систему собственных векторов:

$$Ax_k = \lambda_k x_k, \quad (x_i x_k) = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n).$$

Докажем, что в этом случае A является нормальным оператором. Действительно, положим:

$$y_l = A^* x_l - \overline{\lambda}_l x_l.$$

Тогда

$$(x_k y_l) = (x_k, A^* x_l) - \lambda_l (x_k x_l) = (Ax_k, x_l) - \lambda_l (x_k x_l) = \\ = (\lambda_k - \lambda_l)\,\delta_{kl} = 0 \qquad (k, l = 1, 2, \ldots, n).$$

Отсюда следует:

$$y_l = A^* x_l - \overline{\lambda}_l x_l = 0 \qquad (l = 1, 2, \ldots, n),$$

т. е. имеют место все равенства (63).

Но тогда

$$AA^* x_k = \lambda_k \overline{\lambda}_k x_k \quad \text{и} \quad A^* A x_k = \lambda_k \overline{\lambda}_k x_k \qquad (k = 1, 2, \ldots, n),$$

откуда

$$AA^* = A^* A.$$

Таким образом, мы получили следующую «внутреннюю» (спектральную) характеристику нормального оператора A (наряду с «внешней»: $AA^* = A^* A$):

Теорема 4. *Линейный оператор тогда и только тогда является нормальным, когда этот оператор имеет полную ортонормированную систему собственных векторов.*

В частности, нами доказано, что нормальный оператор всегда является оператором простой структуры.

Пусть A — нормальный оператор с характеристическими числами $\lambda_1, \lambda_2, \ldots, \lambda_n$. По интерполяционной формуле Лагранжа определим два многочлена $p(\lambda)$ и $q(\lambda)$ из условий

$$p(\lambda_k) = \overline{\lambda}_k, \quad q(\overline{\lambda}_k) = \lambda_k \qquad (k = 1, 2, \ldots, n).$$

Тогда в силу (63)

$$A^* = p(A), \quad A = q(A^*), \tag{64}$$

т. е.

2° *Для нормального оператора A каждый из операторов A и A^* представим в виде многочлена от другого из операторов; при этом эти два многочлена определяются заданием характеристических чисел оператора A.*

---

[1]) Под полной ортонормированной системой векторов мы здесь и в дальнейшем понимаем ортонормированную систему из n векторов, где n — число измерений пространства.

Пусть S — инвариантное подпространство в R для нормального оператора A и $R=S+T$, $S \perp T$. Тогда согласно § 8, 5° (стр. 241) подпространство T инвариантно относительно A^*. Но $A=q(A^*)$, где $q(\lambda)$ — многочлен. Поэтому T инвариантно и относительно данного оператора A. Таким образом,

3° *Если S — инвариантное подпространство относительно нормального оператора T, а T — ортогональное дополнение к S, то и T является инвариантным подпространством для A.*

Остановимся теперь на спектре эрмитова оператора. Так как эрмитов оператор H является частным видом нормального оператора, то по доказанному он имеет полную ортонормированную систему собственных векторов:

$$Hx_k=\lambda_k x_k, \quad (x_k x_l)=\delta_{kl} \qquad (k, l=1, 2, \ldots, n). \tag{65}$$

Из $H^*=H$ следует:

$$\overline{\lambda}_k=\lambda_k \qquad (k=1, 2, \ldots, n), \tag{66}$$

т. е. все характеристические числа эрмитова оператора H вещественны.

Нетрудно видеть, что и обратно, нормальный оператор с вещественными характеристическими числами всегда эрмитов. В самом деле, из (65), (66) и

$$H^*x_k=\lambda_k x_k \qquad (k=1, 2, \ldots, n)$$

следует:

$$H^*x_k=Hx_k \qquad (k=1, 2, \ldots, n),$$

т. е.

$$H^*=H.$$

Таким образом, мы получили следующую «внутреннюю» характеристику эрмитова оператора (наряду с «внешней»: $H^*=H$):

Теорема 5. *Линейный оператор H является эрмитовым тогда и только тогда, когда он имеет полную ортонормированную систему собственных векторов с вещественными характеристическими числами.*

Остановимся теперь на спектре унитарного оператора. Поскольку унитарный оператор U является нормальным, то он имеет полную ортонормированную систему собственных векторов

$$Ux_k=\lambda_k x_k, \quad (x_k x_l)=\delta_{kl} \qquad (k, l=1, 2, \ldots, n). \tag{67}$$

При этом

$$U^*x_k=\overline{\lambda}_k x_k \qquad (k=1, 2, \ldots, n). \tag{68}$$

Из $UU^*=E$ находим:

$$\lambda_k\overline{\lambda}_k=1. \tag{69}$$

Обратно, из (67), (68), (69) следует: $UU^*=E$. Таким образом, среди нормальных операторов унитарный оператор выделяется тем, что у него все характеристические числа по модулю равны единице.

Мы получили следующую «внутреннюю» характеристику унитарного оператора (наряду с «внешней»: $UU^*=E$):

Теорема 6. *Линейный оператор тогда и только тогда является унитарным, когда он имеет полную ортонормированную систему собственных векторов с характеристическими числами, по модулю равными единице.*

Так как в ортонормированном базисе нормальная (эрмитова, унитарная матрица соответственно определяет нормальный (эрмитов, унитарный) оператор, то получаем следующие предложения:

Теорема 4′. *Матрица A является нормальной тогда и только тогда, когда она унитарно-подобна диагональной матрице*:

$$A = U \| \lambda_i \delta_{ik} \|_1^n U^{-1} \quad (U^* = U^{-1}). \tag{70}$$

Теорема 5′. *Матрица H является эрмитовой тогда и только тогда, когда она унитарно-подобна диагональной матрице с вещественными числами на диагонали*:

$$H = U \| \lambda_i \delta_{ik} \|_1^n U^{-1} \qquad (U^* = U^{-1};\ \lambda_i = \bar{\lambda}_i;\ i = 1, 2, \ldots, n). \tag{71}$$

Теорема 6′. *Матрица U является унитарной тогда и только тогда, если она унитарно-подобна диагональной матрице с диагональными элементами, по модулю равными единице*:

$$U = U_1 \| \lambda_i \delta_{ik} \|_1^n U_1^{-1} \qquad (U_1^* = U_1^{-1};\ |\lambda_i| = 1;\ i = 1, 2, \ldots, n). \tag{72}$$

§ 11. Неотрицательные и положительно определенные эрмитовы операторы

Введем следующее определение.

Определение 9. Эрмитов оператор $\boldsymbol{H}$ называется *неотрицательным*, если для любого вектора x из $\boldsymbol{R}$

$$(\boldsymbol{H}x, x) \geqslant 0,$$

и *положительно определенным*, если для любого вектора $x \neq 0$ из $\boldsymbol{R}$

$$(\boldsymbol{H}x, x) > 0.$$

Если задать вектор x его координатами $x_1, x_2, \ldots, x_n$ в произвольном ортонормированном базисе, то $(\boldsymbol{H}x, x)$, как легко видеть, представится в виде эрмитовой формы от переменных $x_1, x_2, \ldots, x_n$, причем неотрицательному (соответственно положительно определенному) оператору будет отвечать неотрицательная (соответственно положительно определенная) эрмитова форма (см. § 1).

Выберем ортонормированный базис $x_1, x_2, \ldots, x_n$ из собственных векторов оператора $\boldsymbol{H}$:

$$\boldsymbol{H}x_k = \lambda_k x_k, \quad (x_k x_l) = \delta_{kl} \qquad (k, l = 1, 2, \ldots, n). \tag{73}$$

Тогда, полагая $x = \sum_{k=1}^{n} \xi_k x_k$, будем иметь:

$$(\boldsymbol{H}x, x) = \sum_{k=1}^{n} \lambda_k |\xi_k|^2 \qquad (k = 1, 2, \ldots, n).$$

Отсюда сразу следует «внутренняя» характеристика неотрицательного и положительно определенного оператора:

Теорема 7. *Эрмитов оператор тогда и только тогда является неотрицательным (соответственно положительно определенным), если все его характеристические числа неотрицательны (соответственно положительны).*

Из сказанного вытекает, что положительно определенный эрмитов оператор есть неособенный неотрицательный эрмитов оператор.

Пусть $\boldsymbol{H}$ — неотрицательный эрмитов оператор. Для него имеют место равенства (73) с $\lambda_k \geqslant 0$ $(k = 1, 2, \ldots, n)$. Положим $\varrho_k = \sqrt{\lambda_k} \geqslant 0$ $(k =$

$= 1, 2, \ldots, n$) и определим линейный оператор $\boldsymbol{F}$ равенствами

$$\boldsymbol{F}x_k = \varrho_k x_k \qquad (k = 1, 2, \ldots, n). \tag{74}$$

Тогда $\boldsymbol{F}$ будет также неотрицательным оператором, причем

$$\boldsymbol{F}^2 = \boldsymbol{H}. \tag{75}$$

Неотрицательный эрмитов оператор $\boldsymbol{F}$, связанный с $\boldsymbol{H}$ равенством (75), будем называть арифметическим корнем квадратным из оператора $\boldsymbol{H}$ и будем обозначать так:

$$\boldsymbol{F} = \sqrt{\boldsymbol{H}}.$$

Если $\boldsymbol{H}$ — положительно определенный оператор, то и $\boldsymbol{F}$ будет положительно определенным.

Определим интерполяционный многочлен Лагранжа $g(\lambda)$ равенствами

$$g(\lambda_k) = \varrho_k \left(= \sqrt{\lambda_k}\right) \qquad (k = 1, 2, \ldots, n). \tag{76}$$

Тогда из (73), (74) и (76) следует:

$$\boldsymbol{F} = g(\boldsymbol{H}). \tag{77}$$

Последнее равенство показывает, что $\sqrt{\boldsymbol{H}}$ *является многочленом от* $\boldsymbol{H}$ *и однозначно определяется заданием неотрицательного эрмитова оператора* $\boldsymbol{H}$ (коэффициенты многочлена $g(\lambda)$ зависят от характеристических чисел оператора $\boldsymbol{H}$).

Примерами неотрицательных эрмитовых операторов являются операторы $\boldsymbol{A}\boldsymbol{A}^*$ и $\boldsymbol{A}^*\boldsymbol{A}$, где $\boldsymbol{A}$ — произвольный линейный оператор в данном пространстве. Действительно, при произвольном векторе x

$$(\boldsymbol{A}\boldsymbol{A}^* x, x) = (\boldsymbol{A}^* x, \boldsymbol{A}^* x) \geqslant 0,$$
$$(\boldsymbol{A}^*\boldsymbol{A} x, x) = (\boldsymbol{A} x, \boldsymbol{A} x) \geqslant 0.$$

Если оператор $\boldsymbol{A}$ неособенный, то $\boldsymbol{A}\boldsymbol{A}^*$ и $\boldsymbol{A}^*\boldsymbol{A}$ — положительно определенные эрмитовы операторы.

Операторы $\sqrt{\boldsymbol{A}\boldsymbol{A}^*}$ и $\sqrt{\boldsymbol{A}^*\boldsymbol{A}}$ мы будем называть *левым и правым модулями* оператора $\boldsymbol{A}$.

У нормального оператора левый и правый модули равны между собой [1]).

§ 12. Полярное разложение линейного оператора в унитарном пространстве. Формулы Кэли

Докажем следующую теорему [2]):

Теорема 8. *Произвольный линейный оператор* $\boldsymbol{A}$ *в унитарном пространстве всегда представим в виде*

$$\boldsymbol{A} = \boldsymbol{H}\boldsymbol{U}, \tag{78}$$

$$\boldsymbol{A} = \boldsymbol{U}_1\boldsymbol{H}_1, \tag{79}$$

где $\boldsymbol{H}$, $\boldsymbol{H}_1$ — *неотрицательные эрмитовы, а* $\boldsymbol{U}$, $\boldsymbol{U}_1$ — *унитарные операторы. Оператор* $\boldsymbol{A}$ *нормален тогда и только тогда, когда*

---

[1]) Относительно подробного исследования нормальных операторов см. [82б]. В этой работе устанавливается необходимое и достаточное условие для того, чтобы произведение двух нормальных операторов было также нормальным оператором.

[2]) См. [82б], стр. 77.

в разложении (78) [*или в* (79)] *множители* H *и* U (*соответственно* H_1 *и* U_1) *перестановочны между собой.*

Доказательство. Из разложений (78) и (79) следует, что H и H_1 являются соответственно левым и правым модулями оператора A.

Действительно,

$$AA^* = HUU^*H = H^2, \quad A^*A = HU_1^*U_1H_1 = H_1^2.$$

Заметим, что достаточно установить разложение (78), так как, применяя это разложение к оператору A^*, получим $A^* = HU$ и, следовательно,

$$A = U^{-1}H,$$

т. е. разложение (79) для оператора A.

Установим сначала разложение (78) для частного случая, когда A — неособенный оператор ($|A| \neq 0$). Полагаем:

$$H = \sqrt{AA^*} \text{ (при этом } |H|^2 = |A|^2 \neq 0), \quad U = H^{-1}A$$

и проверяем унитарность оператора U:

$$UU^* = H^{-1}AA^*H^{-1} = H^{-1}H^2H = E.$$

Заметим, что в рассматриваемом случае в разложении (78) не только первый множитель H, но и второй U однозначно определяются заданием неособенного оператора A.

Рассмотрим теперь общий случай, когда A может быть и особенным оператором. Заметим прежде всего, что полная ортонормированная система собственных векторов оператора A^*A всегда преобразуется оператором A снова в ортогональную же систему векторов. Действительно, пусть

$$A^*Ax_k = \varrho_k^2 x_k \qquad [(x_k x_l) = \delta_{kl},\ \varrho_k \geqslant 0;\ k, l = 1, 2, \ldots, n].$$

Тогда

$$(Ax_k, Ax_l) = (A^*Ax_k, x_l) = \varrho_k^2 \cdot (x_k x_l) = 0 \qquad (k \neq l).$$

При этом

$$|Ax_k|^2 = (Ax_k, Ax_k) = \varrho_k^2 \qquad (k = 1, 2, \ldots, n).$$

Поэтому существует такая ортонормированная система векторов $z_1, z_2, \ldots, z_n$, что

$$Ax_k = \varrho_k z_k \qquad [(z_k z_l) = \delta_{kl};\ k, l = 1, 2, \ldots, n]. \tag{80}$$

Определим линейные операторы H и U равенствами

$$Ux_k = z_k, \quad Hz_k = \varrho_k z_k. \tag{81}$$

Из (80) и (81) находим:

$$A = HU.$$

При этом в силу (81) H — неотрицательный эрмитов оператор, поскольку он имеет полную ортонормированную систему собственных векторов $z_1, z_2, \ldots, z_n$ и неотрицательные характеристические числа $\varrho_1, \varrho_2, \ldots, \varrho_n$, а U — унитарный оператор, ибо он переводит ортонормированную систему векторов $x_1, x_2, \ldots, x_n$ снова в ортонормированную $z_1, z_2, \ldots, z_n$.

Таким образом, можно считать доказанным, что для произвольного линейного оператора A имеют место разложения (78) и (79), причем эрмитовы множители H и H_1 всегда однозначно определяются заданием

оператора A (они суть соответственно левый и правый модули оператора A), а унитарные множители U и U_1 определяются однозначно лишь в случае неособенного A.

Из (78) легко находим:

$$AA^* = H^2, \quad A^*A = U^{-1}H^2U. \tag{82}$$

Если A — нормальный оператор ($AA^* = A^*A$), то из (82) вытекает:

$$H^2U = UH^2. \tag{83}$$

Поскольку $H = \sqrt{H^2} = g(H^2)$ (см. § 11), то из (83) следует перестановочность U с H. Обратно, если H и U перестановочны между собой, то из (82) вытекает, что A — нормальный оператор. Теорема доказана.

Вряд ли необходимо особо отмечать то, что наряду с операторными равенствами (78) и (79) имеют место соответствующие матричные равенства.

Характеристические числа оператора $H = \sqrt{AA^*}$ (которые в силу (82) являются также характеристическими числами оператора $H_1 = \sqrt{A^*A}$), называют иногда сингулярными числами оператора A [1]).

Разложения (78) и (79) являются аналогом представления комплексного числа z в виде $z = ru$, где $r = |z|$, а $|u| = 1$.

Пусть теперь $x_1, x_2, \ldots, x_n$ — полная ортонормированная система собственных векторов произвольного унитарного оператора U. Тогда

$$Ux_k = e^{if_k}x_k, \quad (x_kx_l) = \delta_{kl} \qquad (k, l = 1, 2, \ldots, n), \tag{84}$$

где f_k ($k = 1, 2, \ldots, n$) — вещественные числа. Определим эрмитов оператор F равенствами

$$Fx_k = f_kx_k \qquad (k = 1, 2, \ldots, n). \tag{85}$$

Тогда [2])

$$e^{iF}x_k = e^{if_k}x_k \qquad (k = 1, \ldots, n). \tag{85'}$$

Из (84) и (85′) следует

$$U = e^{iF}. \tag{86}$$

Таким образом, унитарный оператор U всегда представим в виде (86), где F — эрмитов оператор. Обратно, если F — эрмитов оператор, то $U = e^{iF}$ — унитарный оператор.

Разложения (78) и (79) вместе с (86) дают следующие равенства:

$$A = He^{iF}, \tag{87}$$

$$A = e^{iF_1}H_1, \tag{88}$$

---

[1]) Если характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ и сингулярные числа $\varrho_1, \varrho_2, \ldots, \varrho_n$ линейного оператора A занумеровать так, чтобы

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n|, \quad \varrho_1 \geqslant \varrho_2 \geqslant \ldots \geqslant \varrho_n,$$

то имеют место неравенства Вейля

$$|\lambda_1| \leqslant \varrho_1, \quad |\lambda_1| + |\lambda_2| \leqslant \varrho_1 + \varrho_2, \ \ldots \ |\lambda_1| + \ldots + |\lambda_n| \leqslant \varrho_1 + \ldots + \varrho_n.$$

Более подробно об этом см. в Добавлении на стр. 542. (*Прим. ред.*)

[2]) $e^{iF} = r(F)$, где $r(\lambda)$ — интерполяционный многочлен Лагранжа для функции e^{if} в точках $f_1, f_2, \ldots, f_n$.

где $\boldsymbol{H}$, $\boldsymbol{F}$, $\boldsymbol{H}_1$, $\boldsymbol{F}_1$ — эрмитовы операторы и притом $\boldsymbol{H}$ и $\boldsymbol{H}_1$ неотрицательны.

Разложения (87) и (88) являются аналогом представления комплексного числа z в виде $z=re^{i\varphi}$, где $r\geqslant 0$ и φ — вещественные числа.

Замечание. В равенстве (86) оператор $\boldsymbol{F}$ не определяется однозначно заданием оператора $\boldsymbol{U}$. Действительно, оператор $\boldsymbol{F}$ определяется при помощи чисел $f_k\,(k=1, 2, \ldots, n)$, а к каждому из этих чисел можно прибавить произвольную кратность 2π, не изменяя исходных равенств (84). Выбирая надлежащим образом эти слагаемые, кратные 2π, мы можем достичь того, чтобы из $e^{if_k}=e^{if_l}$ всегда следовало: $f_k=f_l$ $(1\leqslant k,\ l\leqslant n)$. Тогда можно определить интерполяционный многочлен $g(\lambda)$ равенствами

$$g(e^{if_k})=f_k \qquad (k=1, 2, \ldots, n). \tag{89}$$

Из (84), (85) и (89) будет следовать:

$$\boldsymbol{F}=g(\boldsymbol{U})=g(e^{i\boldsymbol{F}}). \tag{90}$$

Совершенно аналогично можно нормировать выбор $\boldsymbol{F}_1$ так, чтобы

$$\boldsymbol{F}_1=h(\boldsymbol{U}_1)=h(e^{i\boldsymbol{F}}), \tag{91}$$

где $h(\lambda)$ — некоторый многочлен.

В силу (90) и (91) перестановочность $\boldsymbol{H}$ и $\boldsymbol{U}$ ($\boldsymbol{H}_1$ и $\boldsymbol{U}_1$) влечет перестановочность $\boldsymbol{H}$ и $\boldsymbol{F}$ (соответственно $\boldsymbol{H}_1$ и $\boldsymbol{F}_1$) и наоборот. Поэтому согласно теореме 8 оператор $\boldsymbol{A}$ будет нормальным тогда и только тогда, когда в формуле (87) $\boldsymbol{H}$ и $\boldsymbol{F}$ (или в формуле (88) $\boldsymbol{H}_1$ и $\boldsymbol{F}_1$) перестановочны между собой, если только характеристические числа оператора $\boldsymbol{F}$ (соответственно $\boldsymbol{F}_1$) надлежащим образом нормированы.

В основе формулы (86) лежит тот факт, что функциональная зависимость

$$\mu=e^{if} \tag{92}$$

переводит n произвольных чисел на вещественной оси $f_1, f_2, \ldots, f_n$ в некоторые числа $\mu_1, \mu_2, \ldots, \mu_n$, лежащие на окружности $|\mu|=1$, и наоборот.

Трансцендентную зависимость (92) можно заменить рациональной зависимостью

$$\mu=\frac{1+if}{1-if}, \tag{93}$$

которая переводит вещественную ось $f=\overline{f}$ в окружность $|\mu|=1$; при этом бесконечно удаленная точка на вещественной оси переходит в точку $\mu=-1$. Из (93) находим:

$$f=i\frac{1-\mu}{1+\mu}. \tag{94}$$

Повторяя рассуждения, которые привели нас к формуле (86), мы из (93) и (94) получим две взаимно обратные формулы:

$$\left.\begin{aligned}\boldsymbol{U}&=(\boldsymbol{E}+i\boldsymbol{F})(\boldsymbol{E}-i\boldsymbol{F})^{-1},\\ \boldsymbol{F}&=i(\boldsymbol{E}-\boldsymbol{U})(\boldsymbol{E}+\boldsymbol{U})^{-1}.\end{aligned}\right\} \tag{95}$$

Мы получили формулы Кэли. Эти формулы устанавливают взаимно однозначное соответствие между произвольными эрмитовыми опе-

раторами F и теми унитарными операторами U, у которых среди характеристических чисел нет -1 [1]).

Формулы (86), (87), (88) и (95), конечно, будут верны и тогда, когда мы в них все операторы заменим соответствующими матрицами.

Пользуясь полярным разложением матрицы A ранга r:

$$A=U_1H_1 \qquad (H_1=\sqrt{A^*A},\ U^*U=E) \tag{96}$$

и формулой (71)

$$H_1=V^{-1}\|\mu_i\delta_{ik}\|_1^n V \qquad (V^*V=E,\ \mu_1>0,\ \ldots,\ \mu_r>0,\ \mu_{r+1}=\ldots=\mu_n=0), \tag{97}$$

можно произвольную квадратную матрицу A ранга r представить в виде произведения

$$A=UMV, \tag{98}$$

где $U=U_1V^{-1}$ и V—унитарные матрицы ($U^*U=V^*V=E$), а M—диагональная матрица

$$M=\{\mu_1,\ \ldots,\ \mu_r,\ 0,\ \ldots,\ 0\} \qquad (\mu_1>0,\ \ldots,\ \mu_r>0), \tag{98'}$$

в которой диагональные элементы являются характеристическими числами правого модуля $H_1=\sqrt{A^*A}$ (а следовательно, и левого модуля $H=\sqrt{AA^*}$) матрицы A.

Формулу (98) можно записать в виде

$$A=X\Delta Y^*, \tag{99}$$

где X и Y—$n\times r$-матрицы, образованные первыми r столбцами унитарных матриц U и V^*, а Δ—диагональная матрица r-го порядка:

$$\Delta=\{\mu_1,\ \ldots,\ \mu_r\} \qquad (\mu_1>0,\ \ldots,\ \mu_r>0). \tag{100}$$

Пусть теперь A—произвольная прямоугольная $m\times n$-матрица ранга r. Примем сначала, что $m\leqslant n$. Дополним матрицу A нулевыми строками до квадратной матрицы A_1, после чего применим формулу

$$A_1=\begin{pmatrix}A\\0\end{pmatrix}=X_1\Delta Y^*. \tag{101}$$

Представим $n\times r$-матрицу X_1 в виде

$$X_1=\begin{matrix}\overbrace{}^{r}\\\begin{pmatrix}X\\\widehat{X}\end{pmatrix}\begin{matrix}\}\,m\\\}\,n-m\end{matrix}\end{matrix}.$$

Тогда из равенства (101) найдем:

$$A=X\Delta Y^* \tag{102}$$

и

$$\widehat{X}\Delta Y^*=0. \tag{103}$$

Умножим обе части этого равенства справа на Y. Тогда, поскольку $Y^*Y=E$, получим: $\widehat{X}\Delta=0$, т. е. $\widehat{X}=0$. Но тогда столбцы матрицы X, как и столбцы матрицы Y, унитарно-ортогональны между собой и нормированы.

Случай $m\geqslant n$ сводится к случаю $m\leqslant n$, если применить сначала формулу к матрице A^*, а затем из полученного равенства определить матрицу A. Мы установили следующую теорему [2]):

Теорема 9. *Произвольная прямоугольная $m\times n$-матрица ранга r всегда представима в виде произведения*

$$A=X\Delta Y^*, \tag{104}$$

---

[1]) Особую точку -1 можно заменить любым числом μ_0 ($|\mu_0|=1$). Для этого вместо (93) надо взять дробно-линейную функцию, отображающую вещественную ось $f=\bar{f}$ на окружность $|\mu|=1$ и переводящую точку $f=\infty$ в точку $\mu=\mu_0$. При этом соответствующим образом видоизменятся формулы (94) и (95).

[2]) См. Lanzoc C., Linear systems in selfadjoint form, Amer. Math. Monthly, v. 65, 1958, стр. 665—779; Schwerdtfeger H., Direct proof of Lanzoc's decomposition theorem; там же, v. 67, 1960, стр. 855—860.

где X и Y — унитарные по отношению к столбцам прямоугольные матрицы соответственно размеров $m\times r$ и $n\times r$, а Δ — диагональная матрица r-го порядка с положительными диагональными элементами $\mu_1, \ldots, \mu_r$ [1]).

Полагая $B=X$, $C=\Delta Y$, мы приходим к установленному в главе I (стр. 32) разложению

$$A=BC, \tag{105}$$

где матрицы B и C имеют соответствующие размеры $m\times r$ и $r\times n$. Однако доказанная теорема дает уточнение этого разложения. Она утверждает, что множители B и C могут быть выбраны так, чтобы в матрице B все столбцы, а в матрице C все строки были унитарно-ортогональны между собой.

§ 13. Линейные операторы в евклидовом пространстве

Рассмотрим n-мерное евклидово пространство $\boldsymbol{R}$. Пусть дан произвольный линейный оператор $\boldsymbol{A}$ в $\boldsymbol{R}$.

Определение 10. Линейный оператор $\boldsymbol{A}'$ называется *транспонированным* оператором для оператора $\boldsymbol{A}$, если для любых векторов x и y из $\boldsymbol{R}$:

$$(\boldsymbol{A}x,\ y)=(x,\ \boldsymbol{A}'y). \tag{106}$$

Существование и единственность транспонированного оператора устанавливаются совершенно аналогично тому, как это делалось в § 8 для сопряженного оператора в унитарном пространстве.

Транспонированный оператор обладает следующими свойствами:

1° $(\boldsymbol{A}')'=\boldsymbol{A}$,

2° $(\boldsymbol{A}+\boldsymbol{B})'=\boldsymbol{A}'+\boldsymbol{B}'$,

3° $(\alpha\boldsymbol{A})'=\alpha\boldsymbol{A}'$ (α — вещественное число),

4° $(\boldsymbol{AB})'=\boldsymbol{B}'\boldsymbol{A}'$.

Введем ряд определений.

Определение 11. Линейный оператор $\boldsymbol{A}$ называется *нормальным*, если

$$\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{A}'\boldsymbol{A}.$$

Определение 12. Линейный оператор $\boldsymbol{S}$ называется *симметрическим*, если

$$\boldsymbol{S}'=\boldsymbol{S}.$$

Определение 13. Симметрический оператор $\boldsymbol{S}$ называется *неотрицательным*, если для любого вектора x из $\boldsymbol{R}$

$$(\boldsymbol{S}x,\ x)\geqslant 0.$$

Определение 14. Симметрический оператор $\boldsymbol{S}$ называется *положительно определенным*, если для любого вектора $x\neq 0$ из $\boldsymbol{R}$

$$(\boldsymbol{S}x,\ x)>0.$$

Определение 15. Линейный оператор $\boldsymbol{K}$ называется *кососимметрическим*, если

$$\boldsymbol{K}'=-\boldsymbol{K}.$$

---

[1]) $\mu_1, \ldots, \mu_r$ — отличные от нуля характеристические числа матрицы $\sqrt{AA^*}$ (или $\sqrt{A^*A}$).

Произвольный линейный оператор A всегда представим, и притом однозначно, в виде

$$A = S + K, \tag{107}$$

где S — симметрический, а K — кососимметрический оператор.

Действительно, из (107) следует:

$$A' = S - K. \tag{108}$$

Из (107) и (108) вытекает:

$$S = \frac{1}{2}(A + A'), \quad K = \frac{1}{2}(A - A'). \tag{109}$$

Обратно, формулы (109) всегда определяют симметрический оператор S и кососимметрический K, для которых имеет место равенство (107).

S и K носят название *симметрической* и *кососимметрической компонент* оператора A.

Определение 16. Оператор O называется *ортогональным*, если он сохраняет метрику пространства, т. е. если для любых векторов x, y из R

$$(Ox, Oy) = (x, y). \tag{110}$$

Равенство (110) в силу (106) можно переписать так: $(x, O'Oy) = (x, y)$. Отсюда следует:

$$O'O = E. \tag{111}$$

Обратно, из (111) вытекает (110) (при произвольных векторах x, y)[1]. Из (111) следует: $|O|^2 = 1$, т. е.

$$|O| = \pm 1.$$

Мы будем ортогональный оператор O называть *оператором первого рода*, если $|O| = 1$, и *второго рода*, если $|O| = -1$.

Симметрический, кососимметрический, ортогональный операторы суть частные виды нормального оператора.

Рассмотрим произвольный ортонормированный базис в данном евклидовом пространстве. Пусть линейному оператору A в этом базисе соответствует матрица $A = \|a_{ik}\|_1^n$ (здесь все a_{ik} — вещественные числа). Читатель без труда покажет, что транспонированному оператору A' отвечает в этом же базисе транспонированная матрица $A' = \|a'_{ik}\|_1^n$, где $a'_{ik} = a_{ki}$ ($i, k = 1, 2, \ldots, n$). Отсюда вытекает, что в ортонормированном базисе нормальному оператору A отвечает нормальная матрица A ($AA' = A'A$), симметрическому оператору S отвечает симметрическая матрица $S = \|s_{ik}\|_1^n$ ($S' = S$), кососимметрическому оператору K — кососимметрическая матрица $K = \|k_{ij}\|_1^n$ ($K' = -K$) и, наконец, ортогональному оператору O — ортогональная матрица O ($OO' = E$)[2].

Аналогично тому, как это делалось в § 8 для сопряженного оператора, здесь устанавливается следующее предложение:

Если некоторое подпространство S в R инвариантно относительно линейного оператора A, то ортогональное дополнение T к S в R инвариантно относительно оператора A'.

---

[1]) Ортогональные операторы в евклидовом пространстве образуют группу (эту группу называют ортогональной).

[2]) Исследованию структуры ортогональных матриц посвящены работы [91б, 108а, 82а]. Ортогональную матрицу, как и ортогональный оператор, мы будем называть матрицей первого или второго рода в зависимости от того, $|O| = +1$ или $|O| = -1$.

Для исследования линейных операторов в евклидовом пространстве R мы расширим евклидово пространство R до некоторого унитарного пространства $\widetilde{R}$. Это расширение проведем следующим образом:

1. Векторы из R будем называть «вещественными» векторами.

2. Введем в рассмотрение «комплексные» векторы $z=x+iy$, где x и y — вещественные векторы, т. е. $x\in R$, $y\in R$.

3. Естественным образом определяются операции сложения комплексных векторов и умножения на комплексное число. Тогда совокупность всех комплексных векторов образует n-мерное векторное пространство $\widetilde{R}$ над полем комплексных чисел, содержащее в себе как часть R.

4. В $\widetilde{R}$ вводится эрмитова метрика так, чтобы в R она совпадала с имеющейся там евклидовой метрикой. Читатель легко проверит, что искомая эрмитова метрика задается следующим образом:

Если $z=x+iy$, $w=u+iv$ и $(x, y, u, v\in R)$, то

$$(zw)=(xu)+(yv)+i[(yu)-(xv)].$$

Полагая при этом $\bar{z}=x-iy$ и $\bar{w}=\bar{u}-iv$, будем иметь:

$$(\bar{z}\bar{w})=\overline{(zw)}.$$

Если выбрать вещественный базис, т. е. базис в R, то $\widetilde{R}$ будет представлять собой совокупность всех векторов с комплексными, а R — с вещественными координатами в этом базисе.

Всякий линейный оператор A в R однозначно расширяется до линейного оператора в $\widetilde{R}$:

$$A(x+iy)=Ax+iAy.$$

Среди всех линейных операторов в $\widetilde{R}$ операторы, получившиеся в результате такого расширения из операторов в R, характеризуются тем, что переводят R в R $(AR\subset R)$. Такие операторы будем называть *вещественными*.

В вещественном базисе вещественные операторы определяются вещественными матрицами, т. е. матрицами с вещественными элементами.

Вещественный оператор A переводит комплексно сопряженные векторы $z=x+iy$ и $\bar{z}=x-iy$ $(x, y\in R)$ снова в комплексно сопряженные:

$$Az=Ax+iAy,\quad A\bar{z}=Ax-iAy\quad (Ax,\ Ay\in R).$$

У вещественного оператора вековое уравнение имеет вещественные коэффициенты, поэтому вместе с корнем p-й кратности λ оно имеет и корень p-й кратности $\bar{\lambda}$. Из $Az=\lambda z$ следует: $A\bar{z}=\bar{\lambda}\bar{z}$, т. е. сопряженным характеристическим числам соответствуют сопряженные собственные векторы[1]).

Двумерное подпространство $[z, \bar{z}]$ имеет вещественный базис: $x=\frac{1}{2}(z+\bar{z})$, $y=\frac{1}{2i}(z-\bar{z})$. Плоскость в R с этим базисом будем называть инвариантной плоскостью оператора A, отвечающей паре характеристических чисел λ, $\bar{\lambda}$. Пусть $\lambda=\mu+i\nu$.

---

[1]) Если характеристическому числу λ вещественного оператора A отвечают линейно независимые собственные векторы $z_1, z_2, \ldots, z_p$, то характеристическому числу $\bar{\lambda}$ отвечают линейно независимые собственные векторы $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_p$.

Тогда, как легко видеть,

$$\boldsymbol{Ax} = \mu\boldsymbol{x} - \nu\boldsymbol{y},$$
$$\boldsymbol{Ay} = \nu\boldsymbol{x} + \mu\boldsymbol{y}.$$

Рассмотрим вещественный оператор $\boldsymbol{A}$ простой структуры с характеристическими числами:

$\lambda_{2k-1} = \mu_k + i\nu_k$, $\lambda_{2k} = \mu_k - i\nu_k$, $\lambda_l = \mu_l$ $(k = 1, 2, \ldots, q;\ l = 2q+1, \ldots, n)$,

где μ_k, ν_k, μ_l — вещественные числа, причем $\nu_k \neq 0$ $(k = 1, 2, \ldots, q)$.

Тогда соответствующие этим характеристическим числам собственные векторы $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n$ можно выбирать так, чтобы

$$\boldsymbol{z}_{2k-1} = \boldsymbol{x}_k + i\boldsymbol{y}_k,\ \boldsymbol{z}_{2k} = \boldsymbol{x}_k - i\boldsymbol{y}_k,\ \boldsymbol{z}_l = \boldsymbol{x}_l \quad (k = 1, 2, \ldots, q); \qquad (112)$$
$$(l = 2q+1, \ldots, n).$$

Векторы

$$\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{x}_2, \boldsymbol{y}_2, \ldots, \boldsymbol{x}_q, \boldsymbol{y}_q, \boldsymbol{x}_{2q+1}, \ldots, \boldsymbol{x}_n \qquad (113)$$

образуют базис в евклидовом пространстве $\boldsymbol{R}$. При этом

$$\left.\begin{aligned} \boldsymbol{Ax}_k &= \mu_k\boldsymbol{x}_k - \nu_k\boldsymbol{y}_k, \\ \boldsymbol{Ay}_k &= \nu_k\boldsymbol{x}_k + \mu_k\boldsymbol{y}_k, \\ \boldsymbol{Ax}_l &= \mu_l\boldsymbol{x}_l \end{aligned}\right\} \quad \begin{pmatrix} k = 1, 2, \ldots, q, \\ l = 2q+1, \ldots, n \end{pmatrix}. \qquad (114)$$

В базисе (113) оператору $\boldsymbol{A}$ соответствует вещественная квазидиагональная матрица

$$\left\{\begin{Vmatrix} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{Vmatrix}, \ldots, \begin{Vmatrix} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{Vmatrix}, \mu_{2q+1}, \ldots, \mu_n\right\}. \qquad (115)$$

Таким образом, *для каждого оператора $\boldsymbol{A}$ простой структуры в евклидовом пространстве существует такой базис, в котором оператору $\boldsymbol{A}$ соответствует матрица вида* (115). *Отсюда следует, что всякая вещественная матрица простой структуры вещественно-подобна канонической матрице вида* (115):

$$A = T\left\{\begin{Vmatrix} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{Vmatrix}, \ldots, \begin{Vmatrix} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{Vmatrix}, \mu_{2q+1}, \ldots, \mu_n\right\}T^{-1} \quad (T = \overline{T}). \qquad (116)$$

Транспонированный оператор $\boldsymbol{A}'$ для $\boldsymbol{A}$ в $\boldsymbol{R}$ после расширения становится сопряженным оператором $\boldsymbol{A}^*$ для $\boldsymbol{A}$ в $\widetilde{\boldsymbol{R}}$. Следовательно, *нормальный, симметрический, кососимметрический, ортогональный операторы в $\boldsymbol{R}$ после расширения становятся соответственно нормальным, эрмитовым, умноженным на i эрмитовым, унитарным вещественным операторами в $\boldsymbol{R}$.*

Нетрудно показать, что *для нормального оператора $\boldsymbol{A}$ в евклидовом пространстве можно выбрать канонический базис — ортонормированный базис* (113), *для которого имеют место равенства* (114)[1]. *Поэтому вещественная нормальная матрица всегда вещественно- и ортогонально-подобна матрице вида* (115):

$$A = O\left\{\begin{Vmatrix} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{Vmatrix}, \ldots, \begin{Vmatrix} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{Vmatrix}, \mu_{2q+1}, \ldots, \mu_n\right\}O^{-1} \qquad (117)$$
$$(O = O'^{-1} = \overline{O}).$$

[1]) Из ортонормированности базиса (112) в эрмитовой метрике следует ортонормированность базиса (113) в соответствующей евклидовой метрике.

У симметрического оператора $\boldsymbol{S}$ в евклидовом пространстве все характеристические числа вещественны, так как после расширения этот оператор становится эрмитовым. Для симметрического оператора $\boldsymbol{S}$ в формулах (114) следует положить $q=0$. Тогда получим:

$$\boldsymbol{S}\boldsymbol{x}_l=\mu_l\boldsymbol{x}_l \quad [(\boldsymbol{x}_k\boldsymbol{x}_l)=\delta_{kl};\ k,\ l=1,\ 2,\ \ldots,\ n]. \tag{118}$$

Симметрический оператор $\boldsymbol{S}$ в евклидовом пространстве всегда имеет ортонормированную систему собственных векторов с вещественными характеристическими числами[1]). Поэтому *вещественная симметрическая матрица всегда вещественно- и ортогонально-подобна диагональной матрице:*

$$S=O\{\mu_1,\ \mu_2,\ \ldots,\ \mu_n\}O^{-1} \quad (O=O'^{-1}=\overline{O}). \tag{119}$$

У кососимметрического оператора $\boldsymbol{K}$ в евклидовом пространстве все характеристические числа чисто мнимы (после расширения этот оператор равен произведению i на эрмитов оператор). Для кососимметрического оператора в формулах (114) следует положить:

$$\mu_1=\mu_2=\ldots=\mu_q=\mu_{2q+1}=\ldots=\mu_n=0,$$

после чего эти формулы принимают вид

$$\begin{aligned}&\boldsymbol{K}\boldsymbol{x}_k=-\nu_k\boldsymbol{y}_k,\\ &\boldsymbol{K}\boldsymbol{y}_k=\ \ \nu_k\boldsymbol{x}_k,\quad (k=1,\ 2,\ \ldots,\ q;\ l=2q+1,\ \ldots,\ n).\\ &\boldsymbol{K}\boldsymbol{x}_l=0\end{aligned} \tag{120}$$

Поскольку $\boldsymbol{K}$ является нормальным оператором, базис (113) можно считать ортонормированным. Таким образом, *всякая вещественная кососимметрическая матрица вещественно- и ортогонально-подобна канонической кососимметрической матрице*:

$$K=O\left\{\begin{Vmatrix}0 & \nu_1\\ -\nu_1 & 0\end{Vmatrix},\ \ldots,\ \begin{Vmatrix}0 & \nu_q\\ -\nu_q & 0\end{Vmatrix},\ 0,\ \ldots,\ 0\right\}O^{-1}\ (O=O'^{-1}=\overline{O}). \tag{121}$$

У ортогонального оператора $\boldsymbol{O}$ в евклидовом пространстве все характеристические числа по модулю равны единице (после расширения такой оператор становится унитарным). Поэтому в случае ортогонального оператора в формулах (114) следует положить:

$$\mu_k^2+\nu_k^2=1,\ \mu_l=\pm 1 \quad (k=1,\ 2,\ \ldots,\ q;\ l=2q+1,\ \ldots,\ n).$$

При этом базис (113) можно считать ортонормированным. Формулы (114) можно представить в виде

$$\begin{aligned}&\boldsymbol{O}\boldsymbol{x}_k=\cos\varphi_k\boldsymbol{x}_k-\sin\varphi_k\boldsymbol{y}_k,\\ &\boldsymbol{O}\boldsymbol{y}_k=\sin\varphi_k\boldsymbol{x}_k+\cos\varphi_k\boldsymbol{y}_k,\\ &\boldsymbol{O}\boldsymbol{x}_l=\pm\boldsymbol{x}_l\end{aligned}\quad \begin{pmatrix}k=1,\ 2,\ \ldots,\ q,\\ l=2q+1,\ \ldots,\ n\end{pmatrix}. \tag{122}$$

Из сказанного следует, что *всякая вещественная ортогональная матрица вещественно- и ортогонально-подобна канонической ортогональной*:

$$O=O_1\left\{\begin{Vmatrix}\cos\varphi_1 & \sin\varphi_1\\ -\sin\varphi_1 & \cos\varphi_1\end{Vmatrix},\ \ldots,\ \begin{Vmatrix}\cos\varphi_q & \sin\varphi_q\\ -\sin\varphi_q & \cos\varphi_q\end{Vmatrix},\ \pm 1,\ \ldots,\ \pm 1\right\}O_1^{-1} \tag{123}$$

$$(O_1=O_1'^{-1}=\overline{O_1}).$$

[1]) Симметрический оператор S является неотрицательным, если в (118) все $\mu_l\geqslant 0$, и положительно определенным, если в (118) все $\mu_l>0$.

Пример. Рассмотрим произвольное конечное вращение вокруг точки O в трехмерном евклидовом пространстве. Оно переводит направленный отрезок $\overrightarrow{OA}$ в направленный отрезок $\overrightarrow{OB}$ и потому может быть рассматриваемо как оператор $\boldsymbol{O}$ в трехмерном векторном пространстве (образованном всевозможными отрезками $\overrightarrow{OA}$). Этот оператор линейный и притом ортогональный. Определитель этого оператора равен единице, так как оператор $\boldsymbol{O}$ не изменяет ориентации в пространстве.

Итак, $\boldsymbol{O}$—ортогональный оператор первого рода. Для него формулы (122) будут выглядеть так:

$$\boldsymbol{O}x_1=\cos\varphi x_1-\sin\varphi y_1,$$
$$\boldsymbol{O}y_1=\sin\varphi x_1+\cos\varphi y_1,$$
$$\boldsymbol{O}x_2=\pm x_2.$$

Из равенства $|\boldsymbol{O}|=1$ следует, что $\boldsymbol{O}x_2=x_2$. Это означает, что все точки прямой, проходящей через точку O в направлении вектора x_2, неподвижны. Таким образом, мы видим, что имеет место утверждение:

Произвольное конечное вращение твердого тела вокруг неподвижной точки может быть осуществлено конечным поворотом на угол φ вокруг некоторой неподвижной оси, проходящей через эту точку.

Рассмотрим теперь произвольное конечное движение в трехмерном евклидовом пространстве, переводящее точку x в точку

$$x'=c+\boldsymbol{O}x. \qquad (*)$$

Движение складывается из поворота $\boldsymbol{O}$ вокруг некоторой оси, проходящей через начало координат, и параллельного сдвига на вектор c. Обозначим через u, z_1, z_2 собственные векторы $\boldsymbol{O}$, соответствующие характеристическим числам $\lambda=1$, λ_1, λ_2 (при этом $\lambda_2=\bar{\lambda}_1$, $z_2=\bar{z}_1$):

$$\boldsymbol{O}u=u,\ \boldsymbol{O}z_1=\lambda_1 z_1,\ \boldsymbol{O}z_2=\lambda_2 z^2.$$

Докажем существование такой точки x_0, перемещение которой $x_0'-x_0$ параллельно вектору u (т. е. параллельно оси конечного поворота $\boldsymbol{O}$). Для этого положим

$$c=\gamma u+\gamma_1 z_1+\gamma_2 z_2,\ x_0=\xi u+\xi_1 z_1+\xi_2 z_2\ (\gamma_2=\bar{\gamma}_1,\ \xi_2=\bar{\xi}_1)$$

и найдем, что

$$x_0'-x_0=c+(\boldsymbol{O}-E)\,x_0=\gamma u+[\gamma_1+(\lambda_1-1)\,\xi_1]\,z_1+[\gamma_2+(\lambda_2-1)\,\xi_2]\,z_2.$$

Поэтому, определив координаты ξ_1 и ξ_2 искомой точки x_0 из равенств

$$\xi_1=\frac{\gamma_1}{1-\lambda_1},\quad \xi_2=\frac{\gamma_2}{1-\lambda_2}=\bar{\xi}_1,$$

получим для перемещения точки x_0 требуемую формулу

$$x_0'-x_0=\gamma u.$$

Складывая почленно это равенство с вытекающим из (*) равенством

$$x'-x_0'=\boldsymbol{O}(x-x_0),$$

получим

$$x'-x_0=\boldsymbol{O}(x-x_0)+\gamma u. \qquad (**)$$

Эта формула показывает, что при рассматриваемом конечном движении радиус-вектор точки, проведенный из x_0, поворачивается вокруг некоторой оси на фиксированный угол; затем к нему прибавляется параллельный оси вектор γu. Другими словами, движение представляет собой винтовой сдвиг вокруг оси, проходящей через точку x_0 параллельно вектору u. Нами доказана теорема Эйлера—Даламбера:

Произвольное конечное движение в трехмерном евклидовом пространстве представляет собой винтовое перемещение вокруг некоторой неподвижной оси.

§ 14. Полярное разложение оператора и формулы Кэли в евклидовом пространстве

1. В § 12 было установлено полярное разложение линейного оператора в унитарном пространстве. Совершенно аналогично получается полярное разложение линейного оператора в евклидовом пространстве.

Теорема 9. *Линейный оператор* A *всегда представим в виде произведений*

$$A = SO, \tag{124}$$

$$A = O_1 S_1, \tag{125}$$

где S *и* S_1 *— неотрицательные симметрические, а* O *и* O_1 *— ортогональные операторы; при этом* $S = \sqrt{AA'} = g(AA')$, $S_1 = \sqrt{A'A} = h(A'A)$, *где* $g(\lambda)$, $h(\lambda)$ *— вещественные многочлены.*

В том и только в том случае, когда A *— нормальный оператор, множители* S *и* O (*множители* S_1 *и* O_1) *между собой перестановочны* [1]).

Аналогичное предложение имеет место для матриц.

Отметим геометрическое содержание формул (124) и (125). Будем откладывать векторы n-мерного точечного евклидова пространства из начала координат. Тогда каждый вектор будет радиус-вектором некоторой точки пространства. Ортогональное преобразование, осуществляемое оператором O (или O_1), является «вращением» в этом пространстве, поскольку оно сохраняет евклидову метрику и оставляет на месте начало координат [2]). Симметрический же оператор S (или S_1) осуществляет «дилатацию» n-мерного пространства (т. е. «растяжение» вдоль n взаимно перпендикулярных направлений с различными в общем случае коэффициентами растяжения $\varrho_1, \varrho_2, \ldots, \varrho_n$ ($\varrho_1, \varrho_2, \ldots, \varrho_n$ — произвольные неотрицательные числа). Согласно формулам (124) и (125) произвольное линейное однородное преобразование n-мерного евклидова пространства можно получить, осуществляя последовательно некоторое вращение и некоторую дилатацию (в любом порядке).

2. Подобно тому как это было сделано в предыдущем параграфе для унитарного оператора, рассмотрим теперь некоторые представления для ортогонального оператора в евклидовом пространстве R.

Пусть K — произвольный кососимметрический оператор ($K' = -K$) и

$$O = e^K. \tag{126}$$

Тогда O — ортогональный оператор первого рода. Действительно,

$$O' = e^{K'} = e^{-K} = O^{-1}$$

и

$$|O| = 1\ ^{3}).$$

---

[1]) Как и в теореме 8, операторы S и S_1 определяются однозначно заданием A. Если A — неособенный оператор, то однозначно определяются и ортогональные множители O и O_1.

[2]) В случае $|O| = 1$ это будет собственно вращением; в случае же $|O| = -1$ это будет соединение вращения с зеркальным отображением относительно некоторой координатной плоскости.

[3]) Если $k_1, k_2, \ldots, k_n$ — характеристические числа оператора K, то $\mu_1 = e^{k_1}$, $\mu_2 = e^{k_2}, \ldots, \mu_n = e^{k_n}$ — характеристические числа оператора $O = e^K$; при этом

$$|O| = \mu_1 \mu_2 \ldots \mu_n = e^{\sum_{i=1}^{n} k_i} = 1, \text{ поскольку } \sum_{i=1}^{n} k_i = 0.$$

Покажем, что любой ортогональный оператор первого рода представим в виде (126). Для этого возьмем соответствующую ортогональную матрицу O. Поскольку $|O|=1$, то согласно формуле (123) [1])

$$O=O_1\left\{\left\|\begin{matrix}\cos\varphi_1 & \sin\varphi_1\\ -\sin\varphi_1 & \cos\varphi_1\end{matrix}\right\|,\ \ldots,\ \left\|\begin{matrix}\cos\varphi_q & \sin\varphi_q\\ -\sin\varphi_q & \cos\varphi_q\end{matrix}\right\|,\ +1,\ \ldots,\ +1\right\}O_1^{-1} \quad (127)$$

$$(O_1=O_1'^{-1}=\overline{O}_1).$$

Определим кососимметрическую матрицу K равенством

$$K=O_1\left\{\left\|\begin{matrix}0 & \varphi_1\\ -\varphi_1 & 0\end{matrix}\right\|,\ \ldots,\ \left\|\begin{matrix}0 & \varphi_q\\ -\varphi_q & 0\end{matrix}\right\|,\ 0\ \ldots,\ 0\right\}O_1^{-1}. \quad (128)$$

Поскольку

$$e^{\left\|\begin{matrix}0 & \varphi\\ -\varphi & 0\end{matrix}\right\|}=\left\|\begin{matrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{matrix}\right\|,$$

то из (127) и (128) следует:

$$O=e^K. \quad (129)$$

Матричное равенство (129) влечет операторное равенство (126).

Для представления ортогонального оператора второго рода введем в рассмотрение специальный оператор $\boldsymbol{W}$, определив его в некотором ортонормированном базисе $e_1, e_2, \ldots, e_n$ равенствами

$$\boldsymbol{W}e_1=e_1,\ \ldots,\ \boldsymbol{W}e_{n-1}=e_{n-1},\ \boldsymbol{W}e_n=-e_n. \quad (130)$$

$\boldsymbol{W}$ — ортогональный оператор второго рода. Если $\boldsymbol{O}$ — произвольный ортогональный оператор второго рода, то $\boldsymbol{W}^{-1}\boldsymbol{O}$ и $\boldsymbol{O}\boldsymbol{W}^{-1}$ суть операторы первого рода и потому представимы в виде $e^{\boldsymbol{K}}$ и $e^{\boldsymbol{K}_1}$, где $\boldsymbol{K}$ и $\boldsymbol{K}_1$ — кососимметрические операторы. Отсюда получим формулы для ортогонального оператора второго рода

$$\boldsymbol{O}=\boldsymbol{W}e^{\boldsymbol{K}}=e^{\boldsymbol{K}_1}\boldsymbol{W}. \quad (131)$$

Базис $e_1, e_2, \ldots, e_n$ в формулах (130) можно выбрать так, чтобы он совпадал с базисом x_k, y_k, x_l ($k=1, 2, \ldots, q$; $l=2q+1, \ldots, n$) в формулах (120) и (122). Определенный таким образом оператор $\boldsymbol{W}$ будет перестановочен с $\boldsymbol{K}$; поэтому две формулы (131) сольются в одну:

$$\boldsymbol{O}=\boldsymbol{W}e^{\boldsymbol{K}}\quad(\boldsymbol{W}=\boldsymbol{W}'=\boldsymbol{W}^{-1};\ \boldsymbol{K}'=-\boldsymbol{K},\ \boldsymbol{W}\boldsymbol{K}=\boldsymbol{K}\boldsymbol{W}). \quad (132)$$

Остановимся еще на формулах Кэли, устанавливающих связь между ортогональными и кососимметрическими операторами в евклидовом пространстве. Формула

$$\boldsymbol{O}=(\boldsymbol{E}-\boldsymbol{K})(\boldsymbol{E}+\boldsymbol{K})^{-1}, \quad (133)$$

как легко проверить, переводит кососимметрический оператор $\boldsymbol{K}$ в ортогональный $\boldsymbol{O}$. Из (133) можно выразить $\boldsymbol{K}$ через $\boldsymbol{O}$:

$$\boldsymbol{K}=(\boldsymbol{E}-\boldsymbol{O})(\boldsymbol{E}+\boldsymbol{O})^{-1}. \quad (134)$$

---

[1]) Среди характеристических чисел ортогональной матрицы O первого рода имеется четное число равных -1. Диагональная матрица $\left\|\begin{matrix}-1 & 0\\ 0 & -1\end{matrix}\right\|$ может быть записана в виде $\left\|\begin{matrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{matrix}\right\|$ при $\varphi=\pi$.

Формулы (133) и (134) устанавливают взаимно однозначное соответствие между кососимметрическими операторами и теми ортогональными операторами, которые не имеют характеристического числа -1. Вместо (133) и (134) можно взять формулы

$$\boldsymbol{O} = -(\boldsymbol{E}-\boldsymbol{K})(\boldsymbol{E}+\boldsymbol{K})^{-1}, \tag{135}$$

$$\boldsymbol{K} = (\boldsymbol{E}+\boldsymbol{O})(\boldsymbol{E}-\boldsymbol{O})^{-1}. \tag{136}$$

В этом случае роль особой точки будет играть число $+1$.

3. Полярное разложение вещественной матрицы в соответствии с теоремой 9 позволяет получить основные формулы (117), (119), (121), (123), не прибегая к включению евклидова пространства в унитарное так, как это было сделано ранее. Второй вывод основных формул опирается на следующую теорему:

Теорема 10. *Если две вещественные нормальные матрицы подобны*:

$$B = T^{-1}AT \quad (AA' = A'A, \quad BB' = B'B, \quad A = \overline{A}, \quad B = \overline{B}), \tag{137}$$

то эти матрицы вещественно- и ортогонально-подобны:

$$B = O^{-1}AO \quad (O = \overline{O} = O'^{-1}). \tag{138}$$

Доказательство. Поскольку нормальные матрицы A и B имеют одни и те же характеристические числа, то (см. 2° на стр. 246) существует такой многочлен $g(\lambda)$, что

$$A' = g(A), \quad B' = g(B).$$

Поэтому вытекающее из (137) равенство

$$g(B) = T^{-1}g(A)\,T$$

может быть записано так:

$$B' = T^{-1}A'T. \tag{139}$$

Переходя в этом равенстве к транспонированным матрицам, получим:

$$B = T'AT'^{-1}. \tag{140}$$

Сопоставление (137) с (140) дает:

$$TT'A = ATT'. \tag{141}$$

Воспользуемся теперь полярным разложением матрицы T:

$$T = SO, \tag{142}$$

где $S = \sqrt{TT'} = h(TT')$ [$h(\lambda)$ — многочлен] — симметрическая, а O — вещественная ортогональная матрица. Поскольку согласно (141) матрица A перестановочна с TT', то она же перестановочна с матрицей $S = h(TT')$. Поэтому подставляя в (137) выражение для T из (142), будем иметь:

$$B = O^{-1}S^{-1}ASO = O^{-1}AO.$$

Теорема доказана.

Рассмотрим вещественную каноническую матрицу

$$\left\{\begin{Vmatrix} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{Vmatrix}, \ldots, \begin{Vmatrix} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{Vmatrix}, \quad \mu_{2q+1}, \ldots, \mu_n\right\}. \tag{143}$$

Матрица (143) нормальна и имеет характеристические числа $\mu_1 \pm i\nu_1, \ldots$ $\ldots, \mu_q \pm i\nu_q, \mu_{2q+1}, \ldots, \mu_n$. Так как нормальные матрицы имеют простую

структуру, то любая нормальная матрица, имеющая те же характеристические числа, подобна (а в силу теоремы 10 вещественно- и ортогонально-подобна) матрице (143). Таким образом, приходим к формуле (117).

Совершенно так же получаются формулы (119), (121), (123).

§ 15. Коммутирующие нормальные операторы

В § 10 мы доказали, что два коммутирующих оператора A и B в n-мерном пространстве R всегда имеют общий собственный вектор. Методом индукции можно показать, что это положение справедливо не только для двух, но для любого конечного числа коммутирующих операторов. Действительно, если даны m попарно коммутирующих операторов $A_1, A_2, \ldots, A_m$, среди которых первые $m-1$ имеют общий собственный вектор x, то, повторяя дословно рассуждения леммы 1 (стр. 245) [в качестве A берем любое A_i $(i=1, 2, \ldots, m)$, а в качестве B — оператор A_m], мы получаем вектор y, который является общим собственным вектором операторов $A_1, A_2, \ldots, A_m$.

Доказанное положение справедливо и для бесконечного множества коммутирующих операторов, поскольку такое множество может содержать только конечное число ($\leqslant n^2$) линейно независимых операторов, а общий собственный вектор последних будет общим собственным вектором всех операторов из данного множества.

Пусть теперь дано произвольное конечное или бесконечное множество попарно коммутирующих нормальных операторов $A, B, C, \ldots$ Все они имеют общий собственный вектор x_1. Обозначим $(n-1)$-мерное подпространство, состоящее из всех векторов из R, ортогональных к x_1, через T_1. Согласно § 10, 3° (стр. 247) подпространство T_1 инвариантно относительно операторов $A, B, C, \ldots$ Поэтому все эти операторы имеют общий собственный вектор x_2 в T_1. Рассматривая ортогональное дополнение T_2 к плоскости $[x_1, x_2]$, выделим в нем вектор x_3 и т. д. Таким образом, мы получим ортогональную систему $x_1, x_2, \ldots, x_n$ общих собственных векторов для операторов $A, B, C, \ldots$ Эти векторы можно пронормировать. Нами доказана

Теорема 11. *Если дано конечное или бесконечное множество попарно коммутирующих нормальных операторов $A, B, C, \ldots$ в унитарном пространстве R, то все эти операторы имеют полную ортонормированную систему общих собственных векторов $z_1, z_2, \ldots, z_n$:*

$$Az_i=\lambda_i z_i,\ Bz_i=\lambda_i' z_i,\ Cz_i=\lambda_i'' z_i,\ \ldots\ [(z_i z_k)=\delta_{ik};\ i, k=1, 2, \ldots, n]. \quad (144)$$

В матричной формулировке эта теорема гласит:

Теорема 11′. *Если дано конечное или бесконечное множество попарно коммутирующих нормальных матриц, то все эти матрицы одним и тем же унитарным преобразованием могут быть приведены к диагональному виду, т. е. существует такая унитарная матрица U, что*

$$\left.\begin{aligned} &A=U\{\lambda_1, \ldots, \lambda_n\}U^{-1},\ B=U\{\lambda_1', \ldots, \lambda_n'\}U^{-1},\\ &C=U\{\lambda_1'', \ldots, \lambda_n''\}U^{-1},\ \ldots\ (U=U^{*-1}). \end{aligned}\right\} \quad (145)$$

Пусть теперь даны коммутирующие нормальные операторы в евклидовом пространстве R. Обозначим через $A, B, C, \ldots$ линейно независимые среди них (их конечное число). Включим (с сохранением метрики) R в унитарное пространство $\tilde{R}$, как это было сделано в § 13. Тогда согласно теореме 11 операторы $A, B, C, \ldots$ будут иметь в $\tilde{R}$ полную общую

ортонормированную систему собственных векторов $z_1, z_2, \ldots, z_n$, т. е. будут выполняться равенства (144).

Рассмотрим произвольную линейную комбинацию операторов $A, B, C, \ldots$:

$$P = \alpha A + \beta B + \gamma C + \ldots$$

При любых вещественных значениях $\alpha, \beta, \gamma, \ldots$ оператор P является вещественным ($AR \subset R$) нормальным оператором в $\tilde{R}$ и

$$\left.\begin{aligned} &Pz_j = \Lambda_j z_j, \quad \Lambda_j = \alpha\lambda_j + \beta\lambda_j' + \gamma\lambda_j'' + \ldots \\ &[(z_j z_k) = \delta_{jk}; \quad j, k = 1, 2, \ldots, n]. \end{aligned}\right\} \tag{146}$$

Характеристические числа $\Lambda_j (j = 1, 2, \ldots, n)$ оператора P являются линейными формами относительно $\alpha, \beta, \gamma, \ldots$ В силу вещественности оператора P эти формы можно разбить на попарно комплексно сопряженные и вещественные; при надлежащей нумерации собственных векторов будем иметь:

$$\Lambda_{2k-1} = M_k + iN_k, \quad \Lambda_{2k} = M_k - iN_k, \quad \Lambda_l = M_l \tag{147}$$

$$(k = 1, 2, \ldots, q; \quad l = 2q+1, \ldots, n),$$

где M_k, N_k, M_l — вещественные линейные формы от $\alpha, \beta, \gamma, \ldots$

В соответствии с этим мы можем в (146) считать векторы z_{2k-1} и z_{2k} комплексно сопряженными, а z_l — вещественными:

$$z_{2k-1} = x_k + iy_k, \quad z_{2k} = x_k - iy_k, \quad z_l = x_l \tag{148}$$

$$(k = 1, 2, \ldots, q; \quad l = 2q+1, \ldots, n).$$

Тогда, как легко видеть, вещественные векторы

$$x_k, \; y_k, \; x_l \quad (k = 1, 2, \ldots, q; \; l = 2q+1, \ldots, n) \tag{149}$$

образуют ортонормированный базис в R. В этом каноническом базисе имеем [1]):

$$\left.\begin{aligned} &Px_k = M_k x_k - N_k y_k, \\ &Py_k = N_k x_k + M_k y_k, \\ &Px_l = M_l x_l \end{aligned}\right\} \quad \begin{pmatrix} k = 1, 2, \ldots, q, \\ l = 2q+1, \ldots, n \end{pmatrix}. \tag{150}$$

Поскольку все операторы данного множества получаются из P при частных значениях $\alpha, \beta, \gamma, \ldots$, то базис (149), не зависящий от этих параметров, является общим каноническим базисом для всех данных операторов.

Нами доказана

Теорема 12. *Если дано любое множество коммутирующих нормальных линейных операторов в евклидовом пространстве R, то все эти операторы имеют общий ортонормированный канонический базис x_k, y_k, x_l.*

$$\left.\begin{aligned} &Ax_k = \mu_k x_k - \nu_k y_k, \qquad Bx_k = \mu_k' x_k - \nu_k' y_k, \ldots, \\ &Ay_k = \nu_k x_k + \mu_k y_k, \qquad By_k = \nu_k' x_k + \mu_k' y_k, \ldots, \\ &Ax_l = \mu_l x_l; \qquad\qquad Bx_l = \mu_l' x_l, \ldots \end{aligned}\right\} \tag{151}$$

[1]) Равенства (150) следуют из равенств (146), (147) и (148).

Приведем матричную формулировку теоремы 12:

Теорема 12′. *Любое множество коммутирующих вещественных нормальных матриц A, B, C, ... при помощи одного и того же вещественного ортогонального преобразования O может быть приведено к каноническому виду*

$$\left.\begin{aligned} A &= O\left\{\begin{vmatrix}\mu_1 & \nu_1\\ -\nu_1 & \mu_1\end{vmatrix}, \ldots, \begin{vmatrix}\mu_q & \nu_q\\ -\nu_q & \mu_q\end{vmatrix}, \mu_{2q+1}, \ldots, \mu_n\right\}O^{-1},\\ B &= O\left\{\begin{vmatrix}\mu_1' & \nu_1'\\ -\nu_1' & \mu_1'\end{vmatrix}, \ldots, \begin{vmatrix}\mu_q' & \nu_q'\\ -\nu_q' & \mu_q'\end{vmatrix}, \mu_{2q+1}', \ldots, \mu_n'\right\}O^{-1},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\end{aligned}\right\} \quad (152)$$

Примечание. Если какой-либо из операторов $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, ... (какая-либо из матриц A, B, C, ...), например $\boldsymbol{A}$ (A), является симметрическим (симметрической), то в соответствующих формулах (151) [соответственно (в 152)] все ν равны нулю. В случае косой симметрии все μ равны нулю. В случае, если $\boldsymbol{A}$ — ортогональный оператор (A — ортогональная матрица), то $\mu_k = \cos\varphi_k$, $\nu_k = \sin\varphi_k$, $\mu_l = \pm 1$ $(k = 1, 2, \ldots, q;\quad l = 2q+1, \ldots, n)$.

§ 16. Псевдообратный оператор

Пусть дан произвольный линейный оператор $\boldsymbol{A}$, отображающий n-мерное унитарное пространство $\boldsymbol{R}$ в m-мерное унитарное пространство S (см. гл. III, § 2). Обозначим через r ранг оператора A, т. е. число измерений подпространства $\boldsymbol{AR}$. Рассмотрим два ортогональных расщепления пространств $\boldsymbol{R}$ и S:

$$\boldsymbol{R} = \boldsymbol{R}_1 + \boldsymbol{R}_2, \quad \boldsymbol{R}_1 \perp \boldsymbol{R}_2, \quad \boldsymbol{R}_2 = \boldsymbol{N}_{\boldsymbol{A}}, \quad (153)$$

$$S = S_1 + S_2, \quad S_1 \perp S_2, \quad S_1 = \boldsymbol{AR}. \quad (154)$$

Здесь подпространство $\boldsymbol{R}_2 = \boldsymbol{N}_{\boldsymbol{A}}$ состоит из всех векторов $x \in \boldsymbol{R}$, удовлетворяющих уравнению $\boldsymbol{A}x = 0$. Поэтому число измерений подпространства $\boldsymbol{R}_2$ равно $d = n - r$ (см. стр. 78). Следовательно, число измерений ортогонального дополнения $\boldsymbol{R}_1$ равно r.

С другой стороны, $\boldsymbol{AR}_2 \equiv 0$ и $\boldsymbol{AR}_1 \equiv \boldsymbol{AR} \equiv S_1$. Поскольку подпространства $\boldsymbol{R}_1$ и S_1 имеют одно и то же число измерений r, то линейный оператор $\boldsymbol{A}$ устанавливает взаимно однозначное соответствие между векторами подпространств $\boldsymbol{R}_1$ и S_1. Поэтому однозначно определяется обратный оператор $\boldsymbol{A}^{-1}$, отображающий S_1 в $\boldsymbol{R}_1$.

Псевдообратным оператором $\boldsymbol{A}^+$ для оператора $\boldsymbol{A}$ назовем линейный оператор, отображающий S в $\boldsymbol{R}$ и определяемый равенствами

$$\left.\begin{aligned} \boldsymbol{A}^+ y &= \boldsymbol{A}^{-1} y \qquad (y \in S_1),\\ \boldsymbol{A}^+ y &= 0 \qquad\quad (y \in S_2).\end{aligned}\right\} \quad (155)$$

Псевдообратный оператор $\boldsymbol{A}^+$ однозначно определяется заданием линейного оператора $\boldsymbol{A}$, отображающего пространство $\boldsymbol{R}$ в S и заданием метрики в пространствах $\boldsymbol{R}$ и S. При изменении метрики в пространствах $\boldsymbol{R}$ и S изменяется и псевдообратный оператор $\boldsymbol{A}^+$ [1]).

Роль псевдообратного оператора выясняется из следующей геометрической интерпретации.

Уравнение

$$\boldsymbol{A}x = y \quad (156)$$

[1]) В этом отличие от обратного оператора $\boldsymbol{A}^{-1}$, определение которого не связано с метрикой. Но зато псевдообратный оператор $\boldsymbol{A}^+$ определяется в общем случае при любых m, n, r, а обратный оператор $\boldsymbol{A}^{-1}$ может быть определен лишь в частном случае, когда линейный оператор $\boldsymbol{A}$ устанавливает взаимно однозначное соответствие между векторами пространств $\boldsymbol{R}$ и S, т. е. когда $m = n = r$. В этом частном случае оператор $\boldsymbol{A}^+$ не зависит от метрики пространств $\boldsymbol{R}$ и S и совпадает с обратным оператором $\boldsymbol{A}^{-1}$.

при заданном $y \in S$ либо не имеет решений в R (если y не принадлежит подпространству $S=AR$), либо имеет решения (если $y \in AR$). В последнем случае все решения уравнения (156) получаются из одного решения x^0 прибавлением произвольного вектора $x_2 \in R_2 = N_A$.

Докажем, что вектор

$$x^0 = A^+ y \tag{157}$$

представляет собой *наилучшее приближенное решение* уравнения (156), т. е.

$$|Ax^0 - y| = \min_{x \in R} |Ax - y| \tag{158}$$

и из всех векторов $x \in R$, для которых этот минимум реализуется, вектор x^0 имеет наименьшую длину $|x^0|$.

Действительно, пусть $y = y_1 + y_2$ ($y_1 \in S_1$, $y_2 \in S_2$) и $x^0 = A^+ y = A^+ y_1$. Тогда $y_1 = Ax^0$ представляет собой ортогональную проекцию вектора y на подпространство $S = AR$, состоящее из всех векторов вида Ax, где $x \in R$. Поэтому имеет место равенство (158). С другой стороны, пусть $x' \in R$ — какой-либо другой вектор ($x' \neq x^0$), для которого реализуется минимум (158). Тогда

$$Ax' = Ax^0 = y_1 \tag{159}$$

и, следовательно,

$$A(x' - x^0) = 0, \tag{160}$$

т. е. $x' - x^0 \in R_2$. Поэтому, поскольку $x^0 \perp (x'_0 - x^0)$, то по теореме Пифагора из равенства $x' = x^0 + (x' - x^0)$ находим:

$$|x'|^2 = |x^0|^2 + |x' - x^0|^2 > |x^0|^2. \tag{161}$$

Таким образом, *существует только одно наилучшее приближенное решение уравнения* (156) *и это решение определяется формулой* (157).

Выберем в пространствах R и S ортонормированные базисы. В этих базисах квадрат длины векторов $x \in R$ и $y \in S$ определяется формулами

$$|x|^2 = \sum_{i=1}^{n} |x_i|^2, \qquad |y|^2 = \sum_{i=1}^{n} |y_i|^2 \tag{162}$$

и векторные равенства

$$Ax = y, \qquad x^0 = A^+ y$$

переходят в матричные

$$Ax = y, \qquad x^0 = A^+ y. \tag{163}$$

Поскольку x^0 при любом y представляет собой наилучшее приближенное решение [в смысле метрики (162)] системы линейных уравнений, то A^+ — псевдообратная матрица для прямоугольной матрицы A (см. гл. I, § 4). Таким образом, если в пространствах R и S выбраны ортонормированные базисы, то операторам A и A^+ в этих базисах соответствуют взаимно псевдообратные матрицы A и A^+.

ГЛАВА X

КВАДРАТИЧНЫЕ И ЭРМИТОВЫ ФОРМЫ

§ 1. Преобразование переменных в квадратичной форме

1. *Квадратичной формой* называется однородный многочлен второй степени относительно n переменных $x_1, x_2, \ldots, x_n$. Квадратичную форму всегда можно представить в виде

$$\sum_{i,\,k=1}^{n} a_{ik}x_ix_k \quad (a_{ik}=a_{ki};\ i,\ k=1, 2, \ldots, n),$$

где $A=\|a_{ik}\|_1^n$ — симметрическая матрица.

Обозначая через x столбцевую матрицу $(x_1, x_2, \ldots, x_n)$ и пользуясь сокращенным обозначением для квадратичной формы

$$A(x,\ x)=\sum_{i,\,k=1}^{n} a_{ik}x_ix_k, \tag{1}$$

мы можем написать [1]):

$$A(x, x)=x'Ax. \tag{2}$$

Если $A=\|a_{ik}\|_1^n$ — вещественная симметрическая матрица, то форма (1) называется *вещественной*. В этой главе мы будем в основном рассматривать вещественные квадратичные формы.

Определитель $|A|=|a_{ik}|_1^n$ называется *дискриминантом* квадратичной формы $A(x, x)$. Форма называется *сингулярной*, если ее дискриминант равен нулю.

Каждой квадратичной форме соответствует *билинейная форма*

$$A(x, y)=\sum_{i,\,k=1}^{n} a_{ik}x_iy_k \tag{3}$$

или

$$A(x,\ y)=x'Ay \quad [x=(x_1, \ldots, x_n),\ y=(y_1, \ldots, y_n)]. \tag{4}$$

Если $x^1, x^2, \ldots, x^l, y^1, y^2, \ldots, y^m$ — столбцевые матрицы, а $c_1, c_2, \ldots, c_l, d_1, d_2, \ldots, d_m$ — скаляры, то в силу билинейности выражения $A(x, y)$ [см. (4)]

$$A\Big(\sum_{i=1}^{l} c_ix^i,\ \sum_{j=1}^{m} d_jy^j\Big)=\sum_{i=1}^{l}\sum_{j=1}^{m} c_i\,d_jA(x^i,\ y^j). \tag{5}$$

---

[1]) Значок $'$ означает транспонирование. В формуле (2) квадратичная форма представлена в виде произведения трех матриц: строчной x', квадратной A и столбцевой x.

Если задан некоторый симметрический оператор A в n-мерном евклидовом пространстве и этому оператору в некотором ортонормированном базисе $e_1, e_2, \ldots, e_n$ соответствует матрица $A = \| a_{ik} \|_1^n$ [1]), то для любых векторов

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{i=1}^{n} y_i e_i$$

имеет место тождество

$$A(x, y) = (Ax, y) = (x, Ay)\ ^{2}).$$

В частности,

$$A(x, x) = (Ax, x) = (x, Ax).$$

При этом

$$a_{ik} = (Ae_i, e_k) \qquad (i, k = 1, 2, \ldots, n).$$

2. Посмотрим, как изменяется матрица коэффициентов формы при преобразовании переменных:

$$x_i = \sum_{k=1}^{n} t_{ik} \xi_k \qquad i = (1, 2, \ldots, n). \tag{6}$$

В матричной записи это преобразование выглядит так:

$$x = T\xi. \tag{6'}$$

Здесь x и ξ — столбцевые матрицы: $x = (x_1, x_2, \ldots, x_n)$ и $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$, а T — *преобразующая матрица*: $T = \| t_{ik} \|_1^n$.

Подставляя в (2) выражение для x, из (6′) получим:

$$A(x, x) = \xi' T' A T \xi = \xi' \widetilde{A} \xi = \widetilde{A}(\xi, \xi),$$

где

$$\widetilde{A} = T' A T. \tag{7}$$

Формула (7) выражает матрицу $\widetilde{A} = \| \widetilde{a}_{ik} \|_1^n$ коэффициентов преобразованной формы $\widetilde{A}(\xi, \xi) = \sum_{i,\, k=1}^{n} \widetilde{a}_{ik} \xi_i \xi_k$ через матрицу коэффициентов первоначальной формы $A = \| a_{ik} \|_1^n$ и матрицу преобразования $T = \| t_{ik} \|_1^n$.

Из формулы (7) следует, что при преобразовании формы ее дискриминант умножается на квадрат определителя преобразования:

$$|\widetilde{A}| = |A|\,|T|^2. \tag{8}$$

В дальнейшем мы будем пользоваться исключительно неособенными преобразованиями переменных ($|T| \neq 0$). При таких преобразованиях, как видно из формулы (7), ранг матрицы коэффициентов остается неизменным (ранг матрицы A равен рангу матрицы $\widetilde{A}$) [3]). Ранг матрицы коэффициентов обычно называют *рангом формы*.

---

[1]) $Ae_k = \sum_{i=1}^{n} a_{ik} e_i$ $(k = 1, \ldots, n)$; см. стр. 70.

[2]) В $A(x, y)$ скобки составляют часть условного обозначения; в (Ax, y) и в (x, Ay) скобки обозначают скалярное произведение.

[3]) См. стр. 27.

Определение 1. Две симметрические матрицы A и $\widetilde{A}$, связанные равенством (7), в котором $|T| \neq 0$, называются *конгруэнтными*.

Таким образом, с каждой квадратичной формой связан целый класс попарно конгруэнтных симметрических матриц. Как было уже отмечено выше, все эти матрицы имеют один и тот же ранг — ранг формы. Ранг является инвариантом для данного класса матриц. В случае вещественной квадратичной формы вторым инвариантом является так называемая «сигнатура» квадратичной формы. К введению этого понятия мы и переходим.

§ 2. Приведение квадратичной формы к сумме квадратов. Закон инерции

Вещественную квадратичную форму $A(x, x)$ можно бесчисленным множеством способов представить в виде

$$A(x, x) = \sum_{i=1}^{r} a_i X_i^2, \tag{9}$$

где $a_i \neq 0$ $(i = 1, 2, \ldots, r)$ и

$$X_i = \sum_{k=1}^{n} \alpha_{ik} x_k \quad (i = 1, 2, \ldots, r)$$

— независимые вещественные линейные формы от переменных $x_1, x_2, \ldots, x_n$ (отсюда $r \leqslant n$).

Рассмотрим неособенное преобразование переменных, при котором первые r из новых переменных $\xi_1, \xi_2, \ldots, \xi_n$ связаны с $x_1, x_2, \ldots, x_n$ формулами [1])

$$\xi_i = X_i \quad (i = 1, 2, \ldots, r).$$

Тогда в новых переменных

$$A(x, x) = \widetilde{A}(\xi, \xi) = \sum_{i=1}^{r} a_i \xi_i^2$$

и, следовательно, матрица $\widetilde{A}$ имеет диагональный вид:

$$\widetilde{A} = \{a_1, a_2, \ldots, a_r, 0, \ldots, 0\}.$$

Но ранг матрицы $\widetilde{A}$ равен r. Следовательно, *число квадратов в представлении* (9) *всегда равно рангу формы*.

Мы покажем, что неизменным при различных представлениях формы $A(x, x)$ в виде (9) является не только число всех квадратов, но и число положительных [2]) (а значит, и число всех отрицательных) квадратов.

Теорема 1 (закон инерции квадратичных форм). *При представлении вещественной квадратичной формы $A(x, x)$ в виде суммы*

---

[1]) Нужное преобразование получаем, дополняя систему линейных форм $X_1, \ldots, X_r$ линейными формами $X_{r+1}, \ldots, X_n$ так, чтобы n форм X_j $(j = 1, 2, \ldots, n)$ были независимы, и полагая $\xi_j = X_j$ $(j = 1, 2, \ldots, n)$.

[2]) Под числом положительных (отрицательных) квадратов в представлении (9) мы понимаем число положительных (соответственно отрицательных) коэффициентов a_i.

независимых квадратов [1])

$$A(x, x) = \sum_{i=1}^{r} a_i X_i^2$$

число положительных квадратов и число отрицательных квадратов не зависят от способа представления формы в указанном виде.

Доказательство. Пусть наряду с представлением (9) имеет место другое представление формы $A(x, x)$ в виде суммы независимых квадратов

$$A(x, x) = \sum_{i=1}^{r} b_i Y_i^2,$$

и пусть

$$a_1 > 0,\ a_2 > 0,\ \ldots,\ a_h > 0,\ a_{h+1} < 0,\ \ldots,\ a_r < 0,$$
$$b_1 > 0,\ b_2 > 0,\ \ldots,\ b_g > 0,\ b_{g+1} < 0,\ \ldots,\ b_r < 0.$$

Допустим, что $h \neq g$, например $h < g$. Тогда в тождестве

$$\sum_{i=1}^{r} a_i X_i^2 = \sum_{i=1}^{r} b_i Y_i^2 \tag{10}$$

дадим переменным $x_1, x_2, \ldots, x_n$ значения, удовлетворяющие системе $r-(g-h)$ уравнений

$$X_1 = 0,\ X_2 = 0,\ \ldots,\ X_h = 0,\ Y_{g+1} = 0,\ \ldots,\ Y_r = 0 \tag{11}$$

и не обращающие в нуль хотя бы одну из форм $X_{h+1}, \ldots, X_r$ [2]). При этих значениях переменных левая часть тождества (10) равна

$$\sum_{j=h+1}^{r} a_j X_j^2 < 0,$$

а правая равна

$$\sum_{k=1}^{g} b_k Y_k^2 \geqslant 0.$$

Таким образом, допущение $h \neq g$ привело нас к противоречию. Теорема доказана.

Определение 2. Разность σ между числом π положительных и числом ν отрицательных квадратов в представлении формы $A(x, x)$ называют *сигнатурой* формы $A(x, x)$.

Ранг r и сигнатура σ определяют однозначно числа π и ν, так как

$$r = \pi + \nu, \quad \sigma = \pi - \nu.$$

Заметим еще, что в формуле (9) положительный множитель $\sqrt{|a_i|}$ можно отнести к форме $X_i\ (i = 1, 2, \ldots, r)$. Тогда формула (9) принимает вид

$$A(x, x) = X_1^2 + X_2^2 + \ldots + X_\pi^2 - X_{\pi+1}^2 - \ldots - X_r^2. \tag{12}$$

---

[1]) Под суммой независимых квадратов мы понимаем сумму вида (9), в которой все $a_i \neq 0$ и формы $X_1, X_2, \ldots, X_r$ линейно независимы.

[2]) Такие значения существуют, так как в противном случае уравнения $X_{h+1} = 0, \ldots, X_r = 0$, а значит, и все r уравнений $X_1 = 0, \ldots, X_r = 0$ были бы следствием $r-(g-h)$ уравнений (11). Это невозможно, поскольку линейные формы $X_1, X_2, \ldots, X_r$ независимы.

Полагая $\xi_i = X_i (i = 1, 2, \ldots, r)$[1]), мы форму $A(x, x)$ приводим к каноническому виду

$$\widetilde{A}(\xi, \xi) = \xi_1^2 + \xi_2^2 + \ldots + \xi_\pi^2 - \xi_{\pi+1}^2 - \ldots - \xi_r^2. \tag{13}$$

Отсюда на основании теоремы 1 заключаем, что *всякая вещественная симметрическая матрица A конгруэнтна диагональной матрице, у которой диагональные элементы равны $+1$, -1 или 0:*

$$A = T' \{\underbrace{+1, \ldots, +1}_{\pi}, \underbrace{-1, \ldots, -1}_{\nu}, 0, \ldots, 0\} T. \tag{14}$$

В следующем параграфе будет дано правило для определения сигнатуры по коэффициентам квадратичной формы.

§ 3. Метод Лагранжа приведения квадратичной формы к сумме квадратов. Формула Якоби

Из предыдущего параграфа вытекает, что для определения ранга и сигнатуры формы достаточно каким-либо способом привести эту форму к сумме независимых квадратов.

Мы изложим здесь метод приведения Лагранжа.

1. Метод Лагранжа. Пусть дана квадратичная форма

$$A(x, x) = \sum_{i, k=1}^{n} a_{ik} x_i x_k.$$

Рассмотрим два случая:

1) При некотором $g\ (1 \leqslant g \leqslant n)$ диагональный коэффициент $a_{gg} \neq 0$. Тогда, полагая

$$A(x, x) = \frac{1}{a_{gg}} \Big(\sum_{k=1}^{n} a_{gk} x_k \Big)^2 + A_1(x, x), \tag{15}$$

непосредственной проверкой убеждаемся в том, что квадратичная форма $A_1(x, x)$ уже не содержит переменной x_g. Этот способ выделения квадрата из квадратичной формы применим всегда, когда в матрице $A = \| a_{ik} \|_1^n$ имеются диагональные элементы, отличные от нуля.

2) Коэффициенты $a_{gg} = 0$, $a_{hh} = 0$, но $a_{gh} \neq 0$. В этом случае полагаем:

$$A(x, x) = \frac{1}{2a_{hg}} \Big[\sum_{k=1}^{n} (a_{gk} + a_{hk}) x_k \Big]^2 - \frac{1}{2a_{hg}} \Big[\sum_{k=1}^{n} (a_{gk} - a_{hk}) x_k \Big]^2 + A_2(x, x). \tag{16}$$

Формы

$$\sum_{k=1}^{n} a_{gk} x_k, \quad \sum_{k=1}^{n} a_{hk} x_k \tag{17}$$

линейно независимы, так как первая содержит x_h, но не содержит x_g, а вторая, наоборот, содержит x_g, но не содержит x_h. Поэтому и формы, стоящие под знаком квадрата в (16), линейно независимы [как сумма и разность независимых линейных форм (17)].

Таким образом, мы выделили в $A(x, x)$ два независимых квадрата. Каждый из этих квадратов содержит x_g и x_h, в то время как форма $A_2(x, x)$, как легко проверить, этих переменных не содержит.

[1]) См. сноску [1]) на стр. 269.

Последовательным комбинированием приемов 1) и 2) можно всегда с помощью рациональных операций привести форму $A(x, x)$ к сумме квадратов. При этом полученные квадраты будут независимы, так как на каждом этапе выделяемые квадраты содержат переменные, которые отсутствуют в последующих квадратах.

Заметим еще, что основные формулы (15) и (16) могут быть записаны так:

$$A(x, x)=\frac{1}{4a_{gg}}\left(\frac{\partial A}{\partial x_g}\right)^2+A_1(x, x), \tag{15'}$$

$$A(x, x)=\frac{1}{8a_{gh}}\left[\left(\frac{\partial A}{\partial x_g}+\frac{\partial A}{\partial x_h}\right)^2-\left(\frac{\partial A}{\partial x_g}-\frac{\partial A}{\partial x_h}\right)^2\right]+A_2(x, x). \tag{16'}$$

Пример.

$$A(x, x)=4x_1^2+x_2^2+x_3^2+x_4^2-4x_1x_2-4x_1x_3+4x_1x_4+4x_2x_3-4x_3x_4.$$

Применяем формулу (15') ($g=1$):

$$A(x, x)=\frac{1}{16}(8x_1-4x_2-4x_3+4x_4)^2+A_1(x, x)=(2x_1-x_2-x_3+x_4)^2+A_1(x, x),$$

где

$$A_1(x, x)=2x_2x_3+2x_2x_4-2x_3x_4.$$

Применяем формулу (16') ($g=2$, $h=3$):

$$A_1(x, x)=\frac{1}{8}(2x_2+2x_3)^3-\frac{1}{8}(2x_3-2x_2+4x_4)^2+A_2(x, x)=$$

$$=\frac{1}{2}(x_2+x_3)^2-\frac{1}{2}(x_3-x_2+2x_4)^2+A_2(x, x),$$

где

$$A_2(x, x)=2x_4^2.$$

Окончательно

$$A(x, x)=(2x_1-x_2-x_3+x_4)^2+\frac{1}{2}(x_2+x_3)^2-\frac{1}{2}(x_3-x_2+2x_4)^2+2x_4^2,$$

$$r=4, \quad \sigma=2.$$

2. Формула Якоби. Обозначим через r ранг квадратичной формы $A(x, x)=\sum\limits_{i, k=1}^{n} a_{ik}x_ix_k$ и допустим, что

$$D_k=A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}\neq 0 \qquad (k=1, 2, \dots, r). \tag{18}$$

Поскольку $a_{11}=D_1\neq 0$, то, выделяя по методу Лагранжа из формы $A(x, x)$ один квадрат, получим:

$$A(x, x)=\frac{1}{a_{11}}(a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n)^2+A_1(x, x), \tag{19}$$

где квадратичная форма

$$A_1(x, x)=\sum_{i, k=2}^{n} a_{ik}^{(1)}x_ix_k \qquad (a_{ik}^{(1)}=a_{ki}^{(1)}, \quad i, k=2, \dots, n) \tag{20}$$

не содержит переменной x_1. Из тождества (19) следует, что коэффициенты формы $A_1(x, x)$ определяются формулами

$$a_{ik}^{(1)}=a_{ik}-\frac{a_{1i}a_{1k}}{a_{11}} \qquad (i, k=2, \dots, n). \tag{21}$$

Но тогда эти коэффициенты совпадают с соответствующими элементами

матрицы

$$G_1 = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{Vmatrix},$$

которая получается из симметрической матрицы $A = \| a_{ik} \|_1^n$ после применения к ней первого этапа алгоритма исключения Гаусса [1]) (см. гл. II, § 1).

Таким образом, процесс выделения одного квадрата по методу Лагранжа, по существу, совпадает с первым этапом алгоритма Гаусса. Элементы первой строки матрицы G_1 являются коэффициентами в выделяемом квадрате; величина, обратная элементу a_{11}, является множителем при квадрате. Остальные элементы матрицы G_1 определяют коэффициенты формы $A_1(x, x)$. Для выделения второго квадрата следует выполнить второй этап алгоритма Гаусса и т. д. Применяя к симметрической матрице $A = \| a_{ik} \|_1^n$ полный алгоритм Гаусса, состоящий из r этапов [2]), получим матрицу

$$G_r = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{Vmatrix}$$

и соответственно представление квадратичной формы $A(x, x)$ в виде суммы квадратов

$$A(x, x) = \sum_{k=1}^{r} \frac{1}{a_{kk}^{(k-1)}} (a_{kk}^{(k-1)} x_k + a_{k,k+1}^{(k-1)} x_{k+1} + \dots + a_{kn}^{(k-1)} x_n)^2 \qquad (22)$$

$$(a_{1j}^{(0)} = a_{1j}, \quad j = 1, \dots, n).$$

Введем сокращенные обозначения для независимых линейных форм

$$X_k = a_{kk}^{(k-1)} x_k + a_{k,k+1}^{(k-1)} x_{k+1} + \dots + a_{kn}^{(k-1)} x_n \quad (a_{1k}^{(0)} = a_{1k};\ k = 1, \dots, r). \qquad (23)$$

Замечая, что [3])

$$a_{kk}^{(k-1)} = \frac{D_k}{D_{k-1}} \qquad (k = 1, \dots, r;\ D_0 = 1,\ a_{11}^{(0)} = a_{11}). \qquad (24)$$

можно тождество (22) записать в виде

$$A(x, x) = \sum_{k=1}^{r} \frac{D_{k-1}}{D_k} X_k^2 \qquad (D_0 = 1) \qquad (25)$$

---

[1]) Из формулы (21) и симметричности матрицы $A = \| a_{ik} \|_1^n$ следует симметричность матрицы $A_1 = \| a_{ik} \|_2^n$.

[2]) Выполнение алгоритма возможно благодаря неравенствам (18). Из этих неравенств следует, что $a_{11} \neq 0$, $a_{22}^{(1)} \neq 0$, ..., $a_{rr}^{(r-1)} \neq 0$ (см. стр. 44).

[3]) См. стр. 44; формулы (24) получаются приравниванием последовательных главных миноров D_{k-1} и D_k в матрицах A и G_r; при этом получаем $D_k = a_{11} a_{22}^{(1)} \dots a_{kk}^{(k-1)}$ $(k = 1, 2, \dots, r)$.

Эта формула, дающая представление квадратичной формы $A(x, x)$ в виде суммы независимых квадратов, носит название формулы Якоби[1]).

Для коэффициентов, фигурирующих в формуле Якоби линейных форм X_k, имеют место равенства[2])

$$a_{kq}^{(k-1)} = \frac{A\begin{pmatrix} 1 & \dots & k-1 & k \\ 1 & \dots & k-1 & q \end{pmatrix}}{A\begin{pmatrix} 1 & \dots & k-1 \\ 1 & \dots & k-1 \end{pmatrix}} \qquad (k=1, \dots, r). \tag{26}$$

Если через G обозначить произвольную верхнюю треугольную матрицу, у которой первые r строк совпадают с соответствующими строками матрицы G_r, то на основе формулы Якоби можно утверждать, что преобразование переменных $\xi = Gx$ ($\xi = (\xi_1, \xi_2, \dots, \xi_n)$) переводит квадратичную форму $\sum_{k=1}^{n} D_{k-1}/D_k \xi_k^2$ с диагональной матрицей коэффициентов $\hat{D} = \{1/D_1, D_1/D_2, \dots, D_{r-1}/D_r, 0, \dots, 0\}$ в квадратичную форму $A(x, x)$. Но тогда [см. (7)] справедливо равенство

$$A = G'\hat{D}G.$$

Эта формула устанавливает разложение симметрической матрицы A на треугольные множители и совпадает с формулой (55) на стр. 55.

Формулу Якоби часто представляют в другом виде.

Вместо X_k ($k=1, 2, \dots, r$) вводят линейно независимые формы

$$Y_k = D_{k-1}X_k \qquad (k=1, 2, \dots, r;\ D_0=1). \tag{27}$$

Тогда формула Якоби (25) запишется так:

$$A(x, x) = \sum_{k=1}^{r} \frac{Y_k^2}{D_{k-1}D_k}. \tag{28}$$

Здесь

$$Y_k = c_{kk}x_k + c_{k,k+1}x_{k+1} + \dots + c_{kn}x_n \qquad (k=1, 2, \dots, r), \tag{29}$$

где

$$c_{kq} = A\begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & q \end{pmatrix} \quad (q=k, k+1, \dots, n;\ k=1, 2, \dots, r). \tag{30}$$

Пример.

$$A(x, x) = x_1^2 + 3x_2^2 - 3x_3^2 - 4x_1x_2 + 2x_1x_3 - 2x_1x_4 - 6x_2x_3 + 8x_2x_4 + 2x_3x_4.$$

Приводим матрицу

$$A = \begin{Vmatrix} 1 & -2 & 1 & -1 \\ -2 & 3 & -3 & 4 \\ 1 & -3 & 0 & 1 \\ -1 & 4 & 1 & -3 \end{Vmatrix}$$

к гауссовой форме

$$G = \begin{Vmatrix} 1 & -2 & 1 & -1 \\ 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Vmatrix}.$$

Отсюда $r=2$, $a_{11}=1$, $a_{22}^{(1)}=-1$.

[1]) Другой вывод формулы Якоби, не использующий алгоритма Гаусса, можно найти, например, в [7], стр. 43—44.

[2]) См. формулу (13) на стр. 43.

Формула (22) дает:

$$A(x,\ x)=(x_1-2x_2+x_3-x_4)^2-(-x_2-x_3+2x_4)^2.$$

Из формулы Якоби (28) вытекает

Теорема 2 (Якоби). *Если для квадратичной формы*

$$A(x,\ x)=\sum_{i,\ k=1}^{n}a_{ik}x_ix_k$$

ранга r имеют место неравенства

$$D_k=A\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}\neq 0\qquad (k=1,\ 2,\ \dots,\ r),\tag{31}$$

то число положительных квадратов π и число отрицательных квадратов ν формы $A(x,\ x)$ совпадают соответственно с числом знакопостоянств P и с числом знакоперемен V в ряду чисел

$$1,\ D_1,\ D_2,\ \dots,\ D_r,\tag{32}$$

т. е. $\pi=P(1,\ D_1,\ D_2,\ \dots,\ D_r)$, $\nu=V(1,\ D_1,\ D_2,\ \dots,\ D_r)$ *и сигнатура*

$$\sigma=r-2V(1,\ D_1,\ D_2,\ \dots,\ D_r).\tag{33}$$

Замечание 1. *В случае, когда в ряду чисел* $1,\ D_1,\ \dots,\ D_r\neq 0$ *имеются нули, но нет трех подряд идущих нулей, для определения сигнатуры можно пользоваться формулой*

$$\sigma=r-2V(1,\ D_1,\ D_2,\ \dots,\ D_r),$$

опуская нулевое D_k, если $D_{k-1}D_{k+1}\neq 0$, и полагая в случае $D_k=D_{k+1}=0$

$$V(D_{k-1},\ D_k,\ D_{k+1},\ D_{k+2})=\begin{cases}1 \text{ при } \dfrac{D_{k+2}}{D_{k-1}}<0,\\ 2 \text{ при } \dfrac{D_{k+2}}{D_{k-1}}>0.\end{cases}\tag{34}$$

Мы приводим здесь это правило без обоснования [1]).

Замечание 2. *При наличии трех подряд идущих нулей в ряду $D_1,\ D_2,\ \dots,\ D_{r-1}$ сигнатура квадратичной формы не может быть непосредственно определена при помощи теоремы Якоби.* В этом случае знаки ненулевых D_k не определяют сигнатуру формы. Следующий пример убеждает нас в этом:

$$A(x,\ x)=2a_1x_1x_4+a_2x_2^2+a_3x_3^2\qquad (a_1a_2a_3\neq 0).$$

Здесь

$$D_1=D_2=D_3=0,\qquad D_4=-a_1^2a_2a_3\neq 0.$$

В то же время

$$\nu=\begin{cases}1 \text{ при } a_2>0,\ \ a_3>0,\\ 3 \text{ при } a_2<0,\ \ a_3<0.\end{cases}$$

В обоих случаях $D_4<0$.

Замечание 3. *Если $D_1\neq 0,\ \dots,\ D_{r-1}\neq 0$, а $D_r=0$, то знаки $D_1,\ D_2,\ \dots,\ D_{r-1}$ не определяют сигнатуру формы.* В качестве подтверждающего примера можно привести форму

$$ax_1^2+ax_2^2+bx_3^2+2ax_1x_2+2ax_2x_3+2ax_1x_3=a(x_1+x_2+x_3)^2+(b-a)x_3^2.$$

Однако в последнем случае перенумерацией переменных можно достичь того, чтобы имело место и неравенство $D_r\neq 0$. Действительно, пусть s-я строка ($s\geqslant r$) линейно независима по отношению к первым $r-1$ строкам. Поменяем между собой

[1]) Это правило было установлено для случая одного нулевого D_k Гундельфингером и для двух подряд идущих нулевых D_k Фробениусом [182f].

номера переменных x_r и x_s. После этого в новой матрице коэффициентов A первые r строк, а значит (в силу симметричности матрицы) и первые r столбцов, линейно независимы. Тогда в произвольном миноре r-го порядка Δ_r каждую строку представим в виде линейной комбинаций первых r строк, а затем каждый столбец — в виде линейной комбинации первых r столбцов. В соответствии с этим, расщепляя минор Δ_r на сумму определителей r-го порядка, мы в конце концов получим, что минор Δ_r равен произведению главного минора D_r на некоторый числовой множитель: $\Delta_r = cD_r$. Но среди миноров Δ_r имеются отличные от нуля миноры, так как r — ранг матрицы A. Поэтому $D_r \neq 0$.

§ 4. Положительные квадратичные формы

В этом параграфе мы остановимся на специальном, но важном классе положительных квадратичных форм.

Определение 3. Вещественная квадратичная форма $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ называется *неотрицательной* (*неположительной*), если при любых вещественных значениях переменных

$$A(x, x) \geqslant 0 \qquad (\leqslant 0). \tag{35}$$

В этом случае симметрическая матрица коэффициентов A называется положительно полуопределенной (отрицательно полуопределенной).

Определение 4. Вещественная квадратичная форма $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ называется *положительно определенной* (*отрицательно определенной*), если при любых не равных одновременно нулю вещественных значениях переменных ($x \neq 0$)

$$A(x, x) > 0 \qquad (< 0). \tag{36}$$

В этом случае матрица A также называется положительно определенной (отрицательно определенной).

Класс положительно определенных (отрицательно определенных) форм является частью класса неотрицательных (соответственно неположительных) форм.

Пусть дана неотрицательная форма $A(x, x)$. Представим ее в виде суммы независимых квадратов:

$$A(x, x) = \sum_{i=1}^{r} a_iX_i^2. \tag{37}$$

В этом представлении все квадраты должны быть положительными:

$$a_i > 0 \qquad (i = 1, 2, \ldots, r). \tag{38}$$

Действительно, если бы какое-либо a_i было <0, то можно было бы подобрать такие значения $x_1, x_2, \ldots, x_n$, при которых

$$X_1 = \ldots = X_{i-1} = X_{i+1} = \ldots = X_r = 0, \qquad X_i \neq 0.$$

Но тогда при этих значениях переменных форма $A(x, x)$ имела бы отрицательное значение, что по условию невозможно. Очевидно, что и обратно, из (37) и (38) следует положительность формы $A(x, x)$.

Таким образом, неотрицательная квадратичная форма характеризуется равенствами $\sigma = r$ ($\pi = r$, $\nu = 0$).

Пусть теперь $A(x, x)$ — положительно определенная форма. Тогда $A(x, x)$ и неотрицательная форма. Поэтому она представима в виде (37),

где все a_i $(i = 1, 2, \ldots, r)$ положительны. Из положительной определенности формы следует, что $r = n$. Действительно, в случае $r < n$ можно подобрать такие не равные одновременно нулю значения $x_1, x_2, \ldots, x_n$, при которых все X_i обращались бы в нуль. Но тогда в силу (37) $A(x, x) = 0$ при $x \neq 0$, что противоречит условию (36).

Легко видеть, что и обратно, если в (37) $r = n$ и все $a_1, a_2, \ldots, a_n$ положительны, то $A(x, x)$ — положительно определенная форма.

Другими словами, *неотрицательная форма тогда и только тогда является положительно определенной, когда она не сингулярна.*

Следующая теорема дает критерий положительной определенности формы в виде неравенств, которым должны удовлетворять коэффициенты формы. При этом используются уже встречавшиеся в предыдущих параграфах обозначения для последовательных главных миноров матрицы A:

$$D_1 = a_{11}, \quad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots, D_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Теорема 3. *Для того чтобы квадратичная форма была положительно определенной, необходимо и достаточно, чтобы выполнялись неравенства*

$$D_1 > 0, \quad D_2 > 0, \ldots, D_n > 0. \tag{39}$$

Доказательство. Достаточность условий (39) следует непосредственно из формулы Якоби (28). Необходимость условий (39) устанавливается следующим образом. Из положительной определенности формы $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ следует положительная определенность «урезанных» форм [1])

$$A_p(x, x) = \sum_{i,k=1}^{p} a_{ik}x_ix_k \qquad (p = 1, 2, \ldots, n).$$

Но тогда все эти формы должны быть несингулярны, т. е.

$$D_p = |A_p| \neq 0 \qquad (p = 1, 2, \ldots, n).$$

Теперь мы имеем возможность воспользоваться формулой Якоби (28) (при $r = n$). Поскольку в правой части этой формулы все квадраты должны быть положительными, то

$$D_1 > 0, \quad D_1D_2 > 0, \quad D_2D_3 > 0, \ldots, D_{n-1}D_n > 0.$$

Отсюда следуют неравенства (39). Теорема доказана.

Поскольку любой главный минор матрицы A при надлежащей перенумерации переменных можно поместить в левый верхний угол, то имеет место

Следствие. *В положительно определенной квадратичной форме* $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ *все главные миноры матрицы коэффициентов*

[1]) Форма $A_p(x, x)$ получается из формы $A(x, x)$, если в последней положить $x_{p+1} = \ldots = x_n = 0$ $(p = 1, 2, \ldots, n)$.

положительны [1]):

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1, 2, \dots, n). \tag{40}$$

Замечание. Из неотрицательности последовательных главных миноров

$$D_1 \geqslant 0,\ D_2 \geqslant 0,\ \dots,\ D_n \geqslant 0$$

не следует неотрицательность формы $A(x, x)$. Действительно, форма

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2,$$

в которой $a_{11} = a_{12} = 0$, $a_{22} < 0$, удовлетворяет условиям $D_1 \geqslant 0$, $D_2 \geqslant 0$, но не является неотрицательной.

Однако имеет место следующая

Теорема 4. *Для того чтобы квадратичная форма* $A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k$ *была неотрицательной, необходимо и достаточно, чтобы все главные миноры ее матрицы коэффициентов были неотрицательны:*

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \geqslant 0 \quad (1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1, 2, \dots, n).$$

Доказательство. Введем вспомогательную форму

$$A_\varepsilon(x, x) = A(x, x) + \varepsilon \sum_{i=1}^{n} x_i^2 \quad (\varepsilon > 0).$$

Очевидно, $\lim\limits_{\varepsilon \to 0} A_\varepsilon(x, x) = A(x, x)$.

Из неотрицательности формы $A(x, x)$ следует положительная определенность формы $A_\varepsilon(x, x)$ и, следовательно, неравенства (см. следствие из теоремы 3)

$$A_\varepsilon\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1, 2, \dots, n).$$

Переходя к пределу при $\varepsilon \to 0$, получаем условия (40).

Пусть, наоборот, даны условия (40). Из этих условий следует:

$$A_\varepsilon\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} = \varepsilon^p + \dots \geqslant \varepsilon^p > 0 \quad (1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;$$

$$p = 1, 2, \dots, n).$$

Но тогда (согласно теореме 3) $A_\varepsilon(x, x)$ — положительно определенная форма

$$A_\varepsilon(x, x) > 0 \quad (x \neq 0).$$

Переходя к пределу при $\varepsilon \to 0$, получаем отсюда:

$$A(x, x) \geqslant 0.$$

Теорема доказана.

---

[1]) Таким образом, из положительности последовательных главных миноров вещественной симметрической матрицы следует положительность всех остальных главных миноров.

Условия неположительности и отрицательной определенности формы получаются соответственно из неравенств (39) и (40), если эти неравенства применить к форме $-A(x, x)$.

Теорема 5. *Для того чтобы квадратичная форма $A(x, x)$ была отрицательно определенной, необходимо и достаточно, чтобы имели место неравенства*

$$D_1<0,\ D_2>0,\ D_3<0,\ \ldots,\ (-1)^n D_n>0, \tag{39'}$$

Теорема 6. *Для того чтобы квадратичная форма $A(x, x)$ была неположительной, необходимо и достаточно, чтобы имели место неравенства*

$$(-1)^p A\begin{pmatrix} i_1\ i_2\ \ldots\ i_p \\ i_1\ i_2\ \ldots\ i_p\end{pmatrix}\geqslant 0\ (1\leqslant i_1<i_2<\ldots<i_p\leqslant n;\ p=1,\ 2,\ \ldots,\ n). \tag{40'}$$

§ 5. Приведение квадратичной формы к главным осям

Рассмотрим произвольную вещественную квадратичную форму

$$A(x,\ x)=\sum_{i,\ k=1}^{n} a_{ik}x_ix_k.$$

Ее матрица коэффициентов $A=\|a_{ik}\|_1^n$ является вещественной симметрической. Поэтому (см. гл. IX, § 13) она ортогонально-подобна некоторой вещественной диагональной матрице Λ, т. е. существует такая вещественная ортогональная матрица O, что

$$\Lambda=O^{-1}AO\quad(\Lambda=\|\lambda_i\delta_{ik}\|_1^n,\quad OO'=E). \tag{41}$$

Здесь $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$ — характеристические числа матрицы A.

Поскольку для ортогональной матрицы $O^{-1}=O'$, то из (41) следует, что форма $A(x, x)$ при ортогональном преобразовании переменных

$$x=O\xi\qquad(OO'=E) \tag{42}$$

или в более подробной записи

$$x_i=\sum_{k=1}^{n} o_{ik}\xi_k\quad\Big(\sum_{j=1}^{n} o_{ij}o_{kj}=\delta_{ik};\ i,\ k=1,\ 2,\ \ldots,\ n\Big) \tag{42'}$$

переходит в форму

$$\Lambda(\xi,\ \xi)=\sum_{i=1}^{n}\lambda_i\xi_i^2. \tag{43}$$

Теорема 7. *Вещественная квадратичная форма $A(x, x)=\sum_{i,\ k=1}^{n} a_{ik}x_ix_k$ всегда может быть приведена при помощи ортогонального преобразования к канонической форме* (43); *при этом $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$ — характеристические числа матрицы $A=\|a_{ik}\|_1^n$.*

Приведение квадратичной формы $A(x, x)$ при помощи ортогонального преобразования к канонической форме (43) называется *приведением к главным осям*. Это название связано с тем, что уравнение центральной гиперповерхности второго порядка,

$$\sum_{i,\ k=1}^{n} a_{ik}x_ix_k=c\qquad(c=\text{const}\neq 0), \tag{44}$$

при ортогональном преобразовании переменных (42) принимает канонический вид

$$\sum_{i=1}^{n} \varepsilon_i \frac{\xi_i^2}{a_i^2} = 1 \left(\frac{\varepsilon_i}{a_i^2} = \frac{\lambda_i}{c}; \quad \varepsilon_i = \pm 1; \quad i = 1, 2, \ldots, n \right). \tag{45}$$

Если мы будем рассматривать $x_1, x_2, \ldots, x_n$ как координаты в некотором ортонормированном базисе n-мерного евклидова пространства, то $\xi_1, \xi_2, \ldots, \xi_n$ будут координатами в новом ортонормированном базисе того же пространства, причем «поворот»[1]) осей осуществляется ортогональным преобразованием (42). Новые оси координат являются осями симметрии центральной поверхности (44) и обычно называются *главными осями* этой поверхности.

Из формулы (43) следует, что *ранг r формы $A(x, x)$ равен числу отличных от нуля характеристических чисел матрицы A, а сигнатура σ равна разности между числом положительных и числом отрицательных характеристических чисел матрицы A.*

Отсюда, в частности, вытекает и такое предложение:

Если при непрерывном изменении коэффициентов квадратичной формы остается неизменным ее ранг, то при этом изменении коэффициентов остается неизменной и ее сигнатура.

При этом мы исходим из того, что непрерывное изменение коэффициентов влечет непрерывное изменение характеристических чисел. Сигнатура может измениться лишь тогда, когда какое-либо характеристическое число поменяет знак. Но тогда в какой-то промежуточный момент рассматриваемое характеристическое число обратится в нуль, что влечет изменение ранга формы.

Из формулы (43) также следует, что вещественная симметрическая матрица A является положительно полуопределенной (положительно определенной) в том и только в том случае, когда все характеристические числа матрицы A неотрицательны (положительны)[2]), т. е. когда она представима в виде

$$A = O \| \lambda_i \delta_{ik} \|_1^n O^{-1} \quad [\lambda_i \geqslant 0 \ (>0); \ i = 1, \ldots, n]. \tag{46}$$

Положительно полуопределенная (определенная) матрица

$$F = O \| \sqrt{\lambda_i} \delta_{ik} \|_1^n O^{-1} \tag{47}$$

является корнем квадратным из положительно полуопределенной (определенной) матрицы A:

$$F = \sqrt{A}. \tag{48}$$

---

[1]) Если $|O| = -1$, то преобразование (45) представляет собой соединение вращения с зеркальным отображением (см. стр. 260). Однако приведение к главным осям можно всегда осуществить при помощи ортогональной матрицы O первого рода ($|O| = 1$). Это следует из того, что, не меняя канонической формы, мы можем сделать дополнительное преобразование

$$\xi_i = \xi_i' \ (i = 1, 2, \ldots, n-1), \quad \xi_n = -\xi_n'.$$

[2]) Отсюда сразу следует, что в ортонормированном базисе евклидова пространства неотрицательному (положительно определенному) оператору A отвечает положительно полуопределенная (положительно определенная) матрица A. В этом можно убедиться и непосредственно, сопоставляя определения 3 и 4 из § 11 гл. IX.

§ 6. Пучок квадратичных форм

В теории малых колебаний приходится одновременно рассматривать две квадратичные формы, из которых одна задает потенциальную, а вторая — кинетическую энергию системы. Вторая форма всегда является положительно определенной.

Изучению системы двух таких форм мы посвящаем этот параграф.

Две вещественные квадратичные формы

$$A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k, \qquad B(x, x) = \sum_{i,k=1}^{n} b_{ik}x_ix_k$$

определяют *пучок* форм $A(x, x) - \lambda B(x, x)$ (λ — параметр).

Если форма $B(x, x)$ — положительно определенная, то пучок $A(x, x) - \lambda B(x, x)$ называют *регулярным.*

Уравнение

$$|A - \lambda B| = 0$$

называется характеристическим уравнением пучка форм $A(x, x) - \lambda B(x, x)$.

Обозначим через λ_0 какой-либо корень этого уравнения. Поскольку матрица $A - \lambda_0 B$ особенная, то существует столбец $z = (z_1, z_2, \ldots, z_n) \neq 0$ такой, что $(A - \lambda_0 B)z = 0$ или

$$Az = \lambda_0 Bz \qquad (z \neq 0).$$

Число λ_0 мы будем называть *характеристическим числом пучка* $A(x, x) - \lambda B(x, x)$, а z — соответствующим *главным столбцом* или «*главным вектором*» этого пучка. Имеет место следующая

Теорема 8. *Характеристическое уравнение*

$$|A - \lambda B| = 0$$

регулярного пучка форм $A(x, x) - \lambda B(x, x)$ *всегда имеет* n *вещественных корней* λ_k ($k = 1, 2, \ldots, n$), *которым соответствуют главные векторы* $z^k = (z_{1k}, z_{2k}, \ldots, z_{nk})$ ($k = 1, 2, \ldots, n$):

$$Az^k = \lambda_k Bz^k \qquad (k = 1, 2, \ldots, n). \tag{49}$$

Эти главные векторы z^k *могут быть выбраны так, чтобы выполнялись соотношения*[1])

$$B(z^i, z^k) = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n). \tag{50}$$

Доказательство. Заметим, что равенства (49) могут быть записаны так:

$$B^{-1}Az^k = \lambda_k z^k \qquad (k = 1, 2, \ldots, n). \tag{49'}$$

Таким образом, наша теорема утверждает, что матрица

$$D = B^{-1}A \tag{51}$$

имеет: 1° простую структуру, 2° вещественные характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ и 3° собственные столбцы (векторы) $z^1, z^2, \ldots, z^n$, соответствующие этим характеристическим числам и удовлетворяющие соотношениям (50).

---

[1]) Иногда говорят, что равенства (50) выражают собой ортонормированность векторов $z^1, \ldots, z^n$ в B-метрике.

Матрица D, являясь произведением двух симметрических матриц B^{-1} и A, не обязательно сама должна быть симметрической, поскольку $D=B^{-1}A$, а $D'=AB^{-1}$. Однако, полагая $F=\sqrt{B}$ [1]), из равенства (51) легко получаем:

$$D=F^{-1}SF, \tag{52}$$

где

$$S=F^{-1}AF^{-1} \tag{52'}$$

— симметрическая матрица. Из того, что *матрица D подобна симметрической матрице S*, сразу следуют утверждения 1° и 2°. Обозначая через $u^k\,(k=1, \ldots, n)$ пронормированную систему собственных векторов симметрической матрицы S:

$$Su^k=\lambda_k u^k\ (k=1, \ldots, n), \qquad (u^k)'\,u^l=\delta_{kl}\ (k,\ l=1, \ldots, n) \tag{53}$$

и полагая

$$u^k=Fz^k \quad (k=1, \ldots, n), \tag{54}$$

мы из равенств (52), (52'), (53), (54) найдем:

$$Dz^k=\lambda_k z^k, \quad B(z^k,\ z^l)=(z^k)'\,Bz^l=\delta_{kl},$$

где $k,\ l=1, \ldots, n$, т. е. доказано утверждение 3°, и теорема 8 доказана полностью.

Заметим, что из (50) следует, что столбцы $z^1, z^2, \ldots, z^n$ линейно независимы. В самом деле, пусть

$$\sum_{k=1}^{n} c_k z^k=0. \tag{55}$$

Тогда при любом $i\ (1\leqslant i\leqslant n)$ согласно (50)

$$0=B\Big(z^i,\ \sum_{k=1}^{n} c_k z_k\Big)=\sum_{k=1}^{n} c_k B(z_i,\ z_k)=c_i.$$

Таким образом, в (55) все $c_i\,(i=1,\ 2, \ldots, n)$ равны нулю, и никакой линейной зависимости между столбцами $z^1, z^2, \ldots, z^n$ не существует.

Квадратную матрицу, составленную из главных столбцов $z^1, z^2, \ldots, z^n$, удовлетворяющих соотношениям (50),

$$Z=(z^1,\ z^2, \ldots, z^n)=\|\,z_{ik}\,\|_1^n$$

мы будем называть *главной матрицей* для пучка форм $A(x, x)-\lambda B(x, x)$. Главная матрица Z — неособенная ($|Z|\neq 0$), поскольку ее столбцы линейно независимы.

Равенства (50) могут быть записаны так:

$$z^{i'}Bz^k=\delta_{ik} \qquad (i,\ k=1,\ 2, \ldots, n). \tag{56}$$

Кроме того, помножив обе части равенств (49) слева на строчную матрицу $z^{i'}$, получим:

$$z^{i'}Az^k=\lambda_k z^{i'}Bz^k=\lambda_k\delta_{ik} \qquad (i,\ k=1,\ 2, \ldots, n). \tag{57}$$

Вводя главную матрицу $Z=(z^1, z^2, \ldots, z^n)$, мы равенства (56) и (57) можем представить в виде

$$Z'AZ=\|\,\lambda_k\delta_{ik}\,\|_1^n, \quad Z'BZ=E. \tag{58}$$

[1]) F — положительно определенная матрица (см. стр. 280). Поэтому $|F|\neq 0$.

Формулы (58) показывают, что неособенное преобразование

$$x = Z\xi \tag{59}$$

одновременно приводит квадратичные формы $A(x, x)$ и $B(x, x)$ к суммам квадратов:

$$\sum_{k=1}^{n} \lambda_k \xi_k^2 \quad \text{и} \quad \sum_{k=1}^{n} \xi_k^2. \tag{60}$$

Это свойство преобразования (59) характеризует главную матрицу Z. Действительно, пусть преобразование (59) одновременно приводит формы $A(x, x)$ и $B(x, x)$ к каноническим видам (60). Тогда имеют место равенства (58), а следовательно, для столбцов матрицы Z, (56) и (57). Из (58) следует неособенность матрицы Z ($|Z| \neq 0$). Равенства же (57) перепишем так:

$$z^{i'}(Az^k - \lambda_k Bz^k) = 0 \qquad (i = 1, 2, \ldots, n); \tag{61}$$

здесь k имеет произвольное фиксированное значение ($1 \leqslant k \leqslant n$). Систему равенств (61) можно объединить в одно равенство:

$$Z'(Az^k - \lambda_k Bz^k) = 0,$$

откуда, поскольку Z' — неособенная матрица,

$$Az^k - \lambda_k Bz^k = 0,$$

т. е. при любом k получаем (49). Следовательно, Z — главная матрица. Нами доказана

Теорема 9. *Если* $Z = \| z_{ik} \|_1^n$ *— главная матрица регулярного пучка форм* $A(x, x) - \lambda B(x, x)$, *то преобразование*

$$x = Z\xi \tag{62}$$

приводит одновременно формы $A(x, x)$ *и* $B(x, x)$ *соответственно к суммам квадратов*

$$\sum_{k=1}^{n} \lambda_k \xi_k^2, \qquad \sum_{k=1}^{n} \xi_k^2, \tag{63}$$

где (63) $\lambda_1, \lambda_2, \ldots, \lambda_n$ *— характеристические числа пучка* $A(x, x) - \lambda B(x, x)$, *соответствующие столбцам* $z^1, z^2, \ldots, z^n$ *матрицы* Z.

Обратно, если некоторое преобразование (62) *одновременно переводит формы* $A(x, x)$ *и* $B(x, x)$ *к виду* (63), *то* $Z = \| z_{ik} \|_1^n$ *— главная матрица регулярного пучка форм* $A(x, x) - \lambda B(x, x)$.

Иногда характерное свойство преобразования (62), сформулированное в теореме 9, используется для построения главной матрицы и доказательства теоремы 8[1]). Для этого сначала совершают преобразование переменных $x = Ty$, приводящее форму $B(x, x)$ к единичной сумме квадратов $\sum_{k=1}^{n} y_k^2$ (что всегда возможно, поскольку $B(x, x)$ — положительно определенная форма). При этом форма $A(x, x)$ переходит в некоторую форму $A_1(y, y)$. Теперь форму $A_1(y, y)$ приводят к виду $\sum_{k=1}^{n} \lambda_k \xi_k^2$ при помощи ортогонального преобразования $y = O\xi$ (приведение к главным осям!).

[1]) См. [7], стр. 56—57.

При этом, очевидно[1]), $\sum_{k=1}^{n} y_k^2 = \sum_{k=1}^{n} \xi_k^2$. Таким образом, преобразование $x = Z\xi$, где $Z = TO$, приводит данные две формы к виду (63). После этого показывают (как это было сделано на стр. 282), что столбцы $z^1, z^2, \ldots, z^n$ матрицы Z удовлетворяют соотношениям (49) и (50).

В частном случае, когда $B(x, x)$ — единичная форма, т. е. $B(x, x) = \sum_{k=1}^{n} x_k^2$ и, следовательно, $B = E$, характеристическое уравнение пучка $A(x, x) - \lambda B(x, x)$ совпадает с характеристическим уравнением матрицы A, а главные векторы пучка становятся собственными векторами матрицы A. В этом случае соотношения (50) записываются так: $z^{i\prime} z^k = \delta_{ik}$ $(i, k = 1, 2, \ldots, n)$, и выражают ортонормированность столбцов $z^1, z^2, \ldots, z^n$.

Теоремы 8 и 9 допускают наглядную геометрическую интерпретацию. Введем евклидово пространство $\boldsymbol{R}$ с базисом $e_1, e_2, \ldots, e_n$ и основной метрической формой $B(x, x)$.

Рассмотрим в $\boldsymbol{R}$ центральную гиперповерхность второго порядка, уравнение которой

$$A(x, x) \equiv \sum_{i, k=1}^{n} a_{ik} x_i x_k = c. \tag{64}$$

После преобразования координат $x = Z\xi$, где $Z = \| z_{ik} \|_1^n$ — главная матрица пучка $A(x, x) - \lambda B(x, x)$, новыми базисными векторами являются векторы $z^1, z^2, \ldots, z^n$, координаты которых в старом базисе составляют столбцы матрицы Z, т. е. главные векторы пучка. Эти векторы образуют ортонормированный базис, в котором уравнение гиперповерхности (64) имеет вид

$$\sum_{k=1}^{n} \lambda_k \xi_k^2 = c. \tag{65}$$

Следовательно, главные векторы пучка $z^1, z^2, \ldots, z^n$ совпадают по направлению с главными осями гиперповерхности (64), а характеристические числа пучка $\lambda_1, \lambda_2, \ldots, \lambda_n$ определяют величины полуосей: $\lambda_k = \pm \frac{c}{a_k^2}$ $(k = 1, 2, \ldots, n)$.

Таким образом, задача определения характеристических чисел и главных векторов регулярного пучка форм $A(x, x) - \lambda B(x, x)$ эквивалентна задаче приведения к главным осям уравнения (64) центральной гиперповерхности второго порядка в том случае, когда уравнение гиперповерхности задано в обобщенной косоугольной системе координат[2]), в которой «единичная сфера» имеет уравнение $B(x, x) = 1$.

Пример.

Дано уравнение поверхности второго порядка

$$2x^2 - 2y^2 - 3z^2 - 10yz + 2xz - 4 = 0 \tag{66}$$

в обобщенной косоугольной системе координат, в которой уравнение единичной сферы

$$2x^2 + 3y^2 + 2z^2 + 2xz = 1. \tag{67}$$

Требуется уравнение (66) привести к главным осям.

---

[1]) Ортогональные преобразования не меняют суммы квадратов переменных, поскольку $(Ox)'\, Ox = x'x$.

[2]) То есть косоугольная система координат с различными масштабами длин вдоль осей.

В данном случае

$$A=\begin{Vmatrix}2 & 0 & 1\\ 0 & -2 & -5\\ 1 & -5 & -3\end{Vmatrix},\qquad B=\begin{Vmatrix}2 & 0 & 1\\ 0 & 3 & 0\\ 1 & 0 & 2\end{Vmatrix}.$$

Характеристическое уравнение пучка $|A-\lambda B|=0$ имеет вид

$$\begin{vmatrix}2-2\lambda & 0 & 1-\lambda\\ 0 & -2-3\lambda & -5\\ 1-\lambda & -5 & -3-2\lambda\end{vmatrix}=0. \tag{68}$$

Это уравнение имеет три корня: $\lambda_1=1$, $\lambda_2=1$, $\lambda_3=-4$.

Координаты главного вектора, соответствующего характеристическому числу 1, обозначим через u, v, w. Величины u, v, w определяются из системы однородных уравнений, коэффициенты которых совпадают с элементами определителя (68) при $\lambda=1$:

$$\begin{aligned}0\cdot u+0\cdot v+0\cdot w&=0,\\ 0\cdot u-5v-5w&=0,\\ 0\cdot u-5v-5w&=0.\end{aligned}$$

Фактически мы здесь имеем лишь одно соотношение

$$v+w=0.$$

Характеристическому числу $\lambda=1$ должны отвечать два ортонормированных главных вектора. Координаты первого можем выбрать произвольно, лишь бы выполнялось условие $v+w=0$. Выберем их так:

$$u=0,\ v,\ w,\ \text{где } w=-v.$$

Координаты второго главного вектора возьмем в виде

$$u',\ v',\ w',\ \text{где } w'=-v',$$

и запишем условие ортогональности $[B(z^1, z^2)=0]$:

$$2uu'+3vv'+2ww'+uw'+u'w=0.$$

Отсюда найдем: $u'=5v'$. Таким образом, координаты второго главного вектора

$$u'=5v',\ v',\ w'=-v'.$$

Аналогично, полагая в характеристическом определителе $\lambda=-4$, найдем для соответствующего главного вектора:

$$u'',\ v''=-u'',\ w''=-2u''.$$

Величины v, v' и u'' определяются из условия: координаты главного вектора должны удовлетворять уравнению единичной сферы $[B(x, x)=1]$, т. е. уравнению (67). Отсюда находим:

$$v=\frac{1}{\sqrt{5}},\qquad v'=\frac{1}{3\sqrt{5}},\qquad u''=-\frac{1}{3}.$$

Поэтому главная матрица имеет вид

$$Z=\begin{Vmatrix}0 & \frac{\sqrt{5}}{3} & -\frac{1}{3}\\ \frac{1}{\sqrt{5}} & \frac{1}{3\sqrt{5}} & \frac{1}{3}\\ -\frac{1}{\sqrt{5}} & -\frac{1}{3\sqrt{5}} & \frac{2}{3}\end{Vmatrix},$$

и соответствующее преобразование координат ($x=Z\xi$) приводит уравнения (66) и (67) к каноническому виду

$$\xi_1^2+\xi_2^2-4\xi_3^2-4=0,\quad \xi_1^2+\xi_2^2+\xi_3^2=1.$$

Первое уравнение может быть еще записано так:

$$\frac{\xi_1^2}{4}+\frac{\xi_2^2}{4}-\frac{\xi_3^2}{1}=1.$$

Это — уравнение однополостного гиперболоида вращения с вещественной полуосью, равной 2, и мнимой, равной 1. Координаты орта оси вращения определяются третьим столбцом матрицы Z, т. е. равны $\frac{1}{3}$, $\frac{1}{3}$, $\frac{2}{3}$. Координаты ортов других двух ортогональных осей задаются первым и вторым столбцами.

§ 7. Экстремальные свойства характеристических чисел регулярного пучка форм [1])

1. Пусть даны две квадратичные формы:

$$A(x,\ x)=\sum_{i,\ k=1}^{n} a_{ik}x_ix_k \quad \text{и} \quad B(x,\ x)=\sum_{i,\ k=1}^{n} b_{ik}x_ix_k,$$

причем форма $B(x,\ x)$ — положительно определенная. Характеристические числа регулярного пучка форм $A(x,\ x)-\lambda B(x,\ x)$ занумеруем так, чтобы они шли в неубывающем порядке:

$$\lambda_1\leqslant\lambda_2\leqslant\ldots\leqslant\lambda_n. \tag{69}$$

Соответствующие этим характеристическим числам главные векторы [2]) по-прежнему будем обозначать через $z^1,\ z^2,\ \ldots,\ z^n$:

$$z^k=(z_{1k},\ z_{2k},\ \ldots,\ z_{nk});\quad z=1,\ 2,\ \ldots,\ n.$$

Определим наименьшее значение (минимум) отношения форм $\frac{A(x,\ x)}{B(x,\ x)}$, рассматривая все возможные значения переменных, не равные одновременно нулю ($x\neq 0$). Для этого удобно перейти к новым переменным $\xi_1,\ \xi_2,\ \ldots,\ \xi_n$ при помощи преобразования

$$x=Z\xi \qquad \left(x_i=\sum_{k=1}^{n} z_{ik}\xi_k;\ i=1,\ 2,\ \ldots,\ n\right),$$

где $Z=\|z_{ik}\|_1^n$ — главная матрица пучка $A(x,\ x)-\lambda B(x,\ x)$. В новых переменных отношение форм представится в виде [см. (63)]

$$\frac{A(x,\ x)}{B(x,\ x)}=\frac{\lambda_1\xi_1^2+\lambda_2\xi_2^2+\ldots+\lambda_n\xi_n^2}{\xi_1^2+\xi_2^2+\ldots+\xi_n^2}. \tag{70}$$

Возьмем на числовой оси n точек $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$. Припишем этим точкам соответственно неотрицательные массы $m_1=\xi_1^2,\ m_2=\xi_2^2,\ \ldots,\ m_n=\xi_n^2$. Тогда согласно формуле (70) отношение $\frac{A(x,\ x)}{B(x,\ x)}$ будет числовой координатой центра этих масс. Поэтому

$$\lambda_1\leqslant\frac{A(x,\ x)}{B(x,\ x)}\leqslant\lambda_n.$$

Отбрасывая временно вторую часть неравенства, выясним, когда в первой части имеет место знак равенства. Для этого выделим в (69)

---

[1]) При изложении этого параграфа мы следуем книге [7], § 10.

[2]) Здесь мы употребляем термин «главный вектор» в смысле главного столбца пучка. См. стр. 281. Вообще в этом параграфе, имея в виду геометрическую интерпретацию (см. стр. 284), мы часто столбец будем называть вектором.

группы равных характеристических чисел:

$$\lambda_1 = \ldots = \lambda_{p_1} < \lambda_{p_1+1} = \ldots = \lambda_{p_1+p_2} < \ldots \tag{71}$$

Центр масс может совпадать с крайней точкой λ_1 лишь в том случае, когда все массы вне этой точки равны нулю, т. е. когда

$$\xi_{p_1+1} = \ldots = \xi_n = 0.$$

В этом случае соответствующее x будет линейной комбинацией главных столбцов $z^1, z^2, \ldots, z^{p_1}$[1]). Поскольку все эти столбцы отвечают характеристическим числам, равным λ_1, то и x будет главным столбцом (вектором) для $\lambda = \lambda_1$.

Нами доказана

Теорема 10. *Наименьшее характеристическое число регулярного пучка $A(x, x) - \lambda B(x, x)$ является минимумом отношения форм $A(x, x)$ и $B(x, x)$*

$$\lambda_1 = \min \frac{A(x, x)}{B(x, x)}, \tag{72}$$

причем этот минимум достигается только на векторах, являющихся главными для характеристического числа λ_1.

2. Для того чтобы дать аналогичную «минимальную» характеристику для следующего характеристического числа λ_2, ограничимся рассмотрением всех векторов x, ортогональных к z^1, т. е. удовлетворяющих уравнению[2])

$$B(z^1, x) = 0.$$

Для этих векторов

$$\frac{A(x, x)}{B(x, x)} = \frac{\lambda_2 \xi_2^2 + \ldots + \lambda_n \xi_n^2}{\xi_2^2 + \ldots + \xi_n^2}$$

и, следовательно,

$$\min \frac{A(x, x)}{B(x, x)} = \lambda_2 \qquad [B(z^1, x) = 0].$$

При этом знак равенства достигается только на тех векторах, ортогональных к z^1, которые являются главными для характеристического числа λ_2.

Переходя к дальнейшим характеристическим числам, мы в конце концов получим следующую теорему:

Теорема 11. *При любом $p\,(1 \leqslant p \leqslant n)$ p-е по величине характеристическое число λ_p в ряду* (69) *является минимумом отношения*

---

[1]) Из $x = Z\xi$ следует: $x = \sum_{k=1}^{n} \xi_k z^k$.

[2]) Здесь и дальше под ортогональностью двух векторов (столбцов) x, y мы будем понимать ортогональность в B-метрике, т. е. равенство $B(x, y) = 0$. Это находится в полном соответствии с геометрической интерпретацией, данной в предыдущем параграфе. Мы рассматриваем величины $x_1, x_2, \ldots, x_n$ как координаты вектора x в некотором базисе евклидова пространства, в котором квадрат длины (норма) вектора задается положительно определенной формой $B(x, x) = \sum_{i,k=1}^{n} b_{ik} x_i x_k$. В этой метрике векторы $z^1, z^2, \ldots, z^n$ образуют ортонормированный базис. Поэтому если вектор $x = \sum_{k=1}^{n} \xi_k z^k$ ортогонален к одному из z^k, то соответствующее $\xi_k = 0$.

форм

$$\lambda_p = \min \frac{A(x, x)}{B(x, x)} \tag{73}$$

при условии, что варьируемый вектор x ортогонален к первым $p-1$ ортонормированным главным векторам $z^1, z^2, \ldots, z^{p-1}$:

$$B(z^1, x) = 0, \ldots, B(z^{p-1}, x) = 0. \tag{74}$$

При этом минимум достигается только на тех векторах, которые удовлетворяют условию (74) *и являются одновременно главными векторами для характеристического числа λ_p.*

3. Характеристика числа λ_p, данная в теореме 11, имеет то неудобство, что она связана с предыдущими главными векторами $z^1, z^2, \ldots, z^{p-1}$ и, следовательно, может быть использована только тогда, когда эти векторы известны. Кроме того, в выборе этих векторов имеется известный произвол.

Для того чтобы дать характеристику числа λ_p ($p = 1, 2, \ldots, n$), свободную от указанных недостатков, мы введем понятие о связях, наложенных на переменные $x_1, x_2, \ldots, x_n$.

Пусть даны линейные формы от переменных $x_1, x_2, \ldots, x_n$:

$$L_k(x) = l_{1k}x_1 + l_{2k}x_2 + \ldots + l_{nk}x_n \quad (k = 1, 2, \ldots, h). \tag{74'}$$

Мы будем говорить, что на переменные $x_1, x_2, \ldots, x_n$ или (что то же) на вектор x наложены h связей $L_1, L_2, \ldots, L_h$, если рассматриваются лишь значения переменных, удовлетворяющие системе уравнений

$$L_k(x) = 0 \qquad (k = 1, 2, \ldots, h). \tag{74''}$$

Сохраняя обозначения (74') для произвольных линейных форм, мы введем специализированные обозначения для «скалярных произведений» вектора x на главные векторы $z^1, z^2, \ldots, z^n$:

$$\widetilde{L}_k(x) = B(z^k, x) \qquad (k = 1, 2, \ldots, n)^1). \tag{75}$$

Кроме того, в случае, когда на варьируемый вектор наложены связи (74''), будем обозначать $\min \frac{A(x, x)}{B(x, x)}$ так:

$$\mu\left(\frac{A}{B}; L_1, L_2, \ldots, L_h\right).$$

В этих обозначениях равенство (73) запишется так:

$$\lambda_p = \mu\left(\frac{A}{B}; \widetilde{L}_1, \widetilde{L}_2, \ldots, \widetilde{L}_{p-1}\right) \qquad (p = 1, 2, \ldots, n). \tag{76}$$

Рассмотрим связи

$$L_1(x) = 0, \ldots, L_{p-1}(x) = 0 \tag{77}$$

и

$$\widetilde{L}_{p+1}(x) = 0, \ldots, \widetilde{L}_n(x) = 0. \tag{78}$$

Поскольку число связей (77) и (78) меньше n, то существует вектор $x^{(1)} \neq 0$, удовлетворяющий одновременно всем этим связям. Так как связи (78) выражают ортогональность вектора x к главным векторам

---

1) $\widetilde{L}_k(x) = z^{k'}Bx = \widetilde{l}_{1k}x_1 + \widetilde{l}_{2k}x_2 + \ldots + \widetilde{l}_{nk}x_n$, где $\widetilde{l}_{1k}, \widetilde{l}_{2k}, \ldots, \widetilde{l}_{nk}$ — элементы строчной матрицы $z^{k'}B$ ($k = 1, 2, \ldots, n$).

$z^{p+1}, \ldots, z^n$, то соответствующие вектору $x^{(1)}$ координаты $\xi_{p+1} = \ldots = \xi_n = 0$. Поэтому согласно (70) 65

$$\frac{A(x^{(1)}, x^{(1)})}{B(x^{(1)}, x^{(1)})} = \frac{\lambda_1 \xi_1^2 + \ldots + \lambda_p \xi_p^2}{\xi_1^2 + \ldots + \xi_p^2} \leqslant \lambda_p.$$

Но тогда

$$\mu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_{p-1}\right) \leqslant \frac{A(x^{(1)}, x^{(1)})}{B(x^{(1)}, x^{(1)})} \leqslant \lambda_p.$$

Это неравенство в соединении с (76) показывает, что при варьировании связей $L_1, L_2, \ldots, L_{p-1}$ величина μ остается $\leqslant \lambda_p$ и достигает λ_p, если взять специализированные связи $\widetilde{L}_1, \widetilde{L}_2, \ldots, \widetilde{L}_{p-1}$.

Нами доказана

Теорема 12. *Если мы рассмотрим минимум отношения двух форм* $\frac{A(x, x)}{B(x, x)}$ *при произвольных* $p-1$ *связях* $L_1, L_2, \ldots, L_{p-1}$ *и будем варьировать связи, то максимум этих минимумов будет равен* λ_p:

$$\lambda_p = \max \mu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_{p-1}\right) \qquad (p = 1, \ldots, n). \tag{79}$$

Теорема 12 дает «максимально-минимальную» характеристику числам $\lambda_1, \lambda_2, \ldots, \lambda_n$ в отличие от «минимальной» характеристики, о которой идет речь в теореме 11.

4. Заметим, что при замене формы $A(x, x)$ в пучке $A(x, x) - \lambda B(x, x)$ на форму $-A(x, x)$ все характеристические числа пучка меняют знак, а соответствующие главные векторы остаются неизменными. Таким образом, характеристическими числами пучка $-A(x, x) - \lambda B(x, x)$ являются числа $-\lambda_n \leqslant -\lambda_{n-1} \leqslant \ldots \leqslant -\lambda_1$.

Кроме того, обозначая через

$$\nu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_h\right) = \max \frac{A(x, x)}{B(x, x)} \tag{80}$$

в случае, когда на варьируемый вектор наложены связи $L_1, L_2, \ldots, L_h$, мы сможем написать:

$$\mu\left(-\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_h\right) = -\nu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_h\right)$$

и

$$\max \mu\left(-\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_h\right) = -\min \nu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_h\right).$$

Поэтому, применяя к отношению $-\frac{A(x, x)}{B(x, x)}$ теоремы 10, 11, 12, мы вместо формул (72), (76), (79) получим формулы

$$\begin{aligned} \lambda_n &= \max \frac{A(x, x)}{B(x, x)}, \\ \lambda_{n-p+1} &= \nu\left(\frac{A}{B};\ \widetilde{L}_n,\ \widetilde{L}_{n-1},\ \ldots,\ \widetilde{L}_{n-p+2}\right), \\ \lambda_{n-p+1} &= \min \nu\left(\frac{A}{B};\ L_1,\ L_2,\ \ldots,\ L_{p-1}\right) \end{aligned} \qquad (p = 2, \ldots, n).$$

Эти формулы устанавливают соответственно «максимальные» и «минимально-максимальные» свойства чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$, которые мы сформулируем в виде следующей теоремы:

Теорема 13. *Пусть характеристическим числам*

$$\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$$

регулярного пучка форм $A(x, x) - \lambda B(x, x)$ *соответствуют линейно независимые главные векторы пучка* $z^1, z^2, \ldots, z^n$. *Тогда*

1) *Наибольшее характеристическое число* λ_n *является максимумом отношения форм*

$$\lambda_n = \max \frac{A(x, x)}{B(x, x)}, \tag{81}$$

причем этот максимум достигается только на главных векторах пучка, соответствующих характеристическому числу λ_n.

2) *p-е (с конца) характеристическое число* λ_{n-p+1} $(2 \leqslant p \leqslant n)$ *является максимумом того же отношения форм*

$$\lambda_{n-p+1} = \max \frac{A(x, x)}{B(x, x)} \tag{82}$$

при условии, что на варьируемый вектор x наложены связи:

$$B(z^n, x) = 0, \; B(z^{n-1}, x) = 0, \; \ldots, \; B(z^{n-p+2}, x) = 0, \tag{83}$$

т. е.

$$\lambda_{n-p+1} = \nu\left(\frac{A}{B}; \widetilde{L}_n, \widetilde{L}_{n-1}, \ldots, \widetilde{L}_{n-p+2}\right); \tag{84}$$

этот максимум достигается только на главных векторах пучка, соответствующих характеристическому числу λ_{n-p+1} *и удовлетворяющих связям* (83).

3) *Если в максимуме отношения форм* $\frac{A(x, x)}{B(x, x)}$ *при связях*

$$L_1(x) = 0, \; \ldots, \; L_{p-1}(x) = 0$$

$(2 \leqslant p \leqslant n)$ *варьировать связи, то наименьшее значение (минимум) этого максимума равно* λ_{n-p+1}:

$$\lambda_{n-p+1} = \min \nu\left(\frac{A}{B}; L_1, L_2, \ldots, L_{p-1}\right). \tag{85}$$

5. Пусть даны h *независимых* связей[1])

$$L_1^0(x) = 0, \; L_2^0(x) = 0, \; \ldots, \; L_h^0(x) = 0. \tag{86}$$

Тогда из них можно выразить h из переменных $x_1, x_2, \ldots, x_n$ через остальные переменные, которые мы обозначим буквами $v_1, v_2, \ldots, v_{n-h}$. Поэтому при наложении связей (86) регулярный пучок форм $A(x, x) - \lambda B(x, x)$ переходит в пучок $A^0(v, v) - \lambda B^0(v, v)$, причем $B^0(v, v)$ — снова положительно определенная форма (только от $n-h$ переменных). Полученный таким образом регулярный пучок имеет $n-h$ вещественных характеристических чисел

$$\lambda_1^0 \leqslant \lambda_2^0 \leqslant \ldots \leqslant \lambda_{n-h}^0. \tag{87}$$

При наложении связей (86) можно по-разному выразить все переменные через $n-h$ независимых $v_1, v_2, \ldots, v_{n-h}$. Однако характеристические числа (87) не зависят от этого произвола и имеют вполне определенные значения. Это следует хотя бы из максимально-минимальных свойств

[1]) Связи (86) являются независимыми, когда независимы линейные формы $L_1^0(x), L_2^0(x), \ldots, L_h^0(x)$, стоящие в левых частях уравнений связей.

характеристических чисел

$$\lambda_1^0 = \min \frac{A^0(v, v)}{B^0(v, v)} = \mu\left(\frac{A}{B};\ L_1^0,\ L_2^0,\ \ldots,\ L_h^0\right) \tag{88}$$

и вообще

$$\lambda_p^0 = \max \mu\left(\frac{A^0}{B^0};\ L_1,\ L_2,\ \ldots,\ L_{p-1}\right) =$$

$$= \max \mu\left(\frac{A}{B};\ L_1^0,\ \ldots,\ L_h^0;\ L_1,\ \ldots,\ L_{p-1}\right), \tag{89}$$

при этом в формуле (89) варьируются только связи $L_1, L_2, \ldots, L_{p-1}$.

Имеет место

Теорема 14. *Если* $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ — *характеристические числа регулярного пучка форм* $A(x, x) - \lambda B(x, x)$, *а* $\lambda_1^0 \leqslant \lambda_2^0 \leqslant \ldots \leqslant \lambda_{n-h}^0$ — *характеристические числа того же пучка при наложении* h *независимых связей, то*

$$\lambda_p \leqslant \lambda_p^0 \leqslant \lambda_{p+h} \quad (p = 1, 2, \ldots, n-h). \tag{90}$$

Доказательство. Неравенства $\lambda_p \leqslant \lambda_p^0$ $(p = 1, 2, \ldots, n-h)$ сразу следуют из формул (79) и (89). Действительно, при добавлении новых связей величина минимума $\mu\left(\frac{A}{B};\ L_1, \ldots, L_{p-1}\right)$ увеличивается или остается прежней. Поэтому

$$\mu\left(\frac{A}{B};\ L_1, \ldots, L_{p-1}\right) \leqslant \mu\left(\frac{A}{B};\ L_1^0, \ldots, L_h^0;\ L_1, \ldots, L_{p-1}\right).$$

Отсюда

$$\lambda_p = \max \mu\left(\frac{A}{B};\ L_1, \ldots, L_{p-1}\right) \leqslant \lambda_p^0 =$$

$$= \max \mu\left(\frac{A}{B};\ L_1^0, \ldots, L_h^0;\ L_1, \ldots, L_{p-1}\right).$$

Вторые части неравенств (90) имеют место в силу соотношений

$$\lambda_p^0 = \max \mu\left(\frac{A}{B};\ L_1^0, \ldots, L_h^0;\ L_1, \ldots, L_{p-1}\right) \leqslant$$

$$\leqslant \max \mu\left(\frac{A}{B};\ L_1, \ldots, L_{p-1},\ L_p, \ldots, L_{p+h-1}\right) = \lambda_{p+h}.$$

Здесь в правой части варьируются не только связи $L_1, \ldots, L_{p-1}$, но и связи $L_p, \ldots, L_{p+h-1}$; в левой же части последние связи заменены фиксированными связями $L_1^0, \ldots, L_h^0$.

Теорема доказана.

6. Пусть даны два регулярных пучка форм

$$A(x, x) - \lambda B(x, x), \quad \widetilde{A}(x, x) - \lambda \widetilde{B}(x, x) \tag{91}$$

и пусть при любом $x \neq 0$

$$\frac{A(x, x)}{B(x, x)} \leqslant \frac{\widetilde{A}(x, x)}{\widetilde{B}(x, x)}.$$

Тогда, очевидно,

$$\max \mu\left(\frac{A}{B};\ L_1, L_2, \ldots, L_{p-1}\right) \leqslant \max \mu\left(\frac{\widetilde{A}}{\widetilde{B}};\ L_1, L_2, \ldots, L_{p-1}\right)$$

$$(p = 1, 2, \ldots, n).$$

Поэтому, обозначая через $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ и $\widetilde{\lambda}_1 \leqslant \widetilde{\lambda}_2 \leqslant \ldots \leqslant \widetilde{\lambda}_n$ соответственно характеристические числа пучков (91), будем иметь:

$$\lambda_p \leqslant \widetilde{\lambda}_p \quad (p = 1, 2, \ldots, n).$$

Таким образом, доказана

Теорема 15. *Если даны два регулярных пучка форм* $A(x, x) - \lambda B(x, x)$ *и* $\widetilde{A}(x, x) - \lambda \widetilde{B}(x, x)$ *с характеристическими числами соответственно* $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ *и* $\widetilde{\lambda}_1 \leqslant \widetilde{\lambda}_2 \leqslant \ldots \leqslant \widetilde{\lambda}_n$, *то из тождественного соотношения*

$$\frac{A(x, x)}{B(x, x)} \leqslant \frac{\widetilde{A}(x, x)}{\widetilde{B}(x, x)} \tag{92}$$

следует:

$$\lambda_p \leqslant \widetilde{\lambda}_p \quad (p = 1, 2, \ldots, n). \tag{93}$$

Рассмотрим частный случай, когда в неравенстве (92) $B(x, x) \equiv \widetilde{B}(x, x)$. В этом случае разность $\widetilde{A}(x, x) - A(x, x)$ является неотрицательной квадратичной формой и поэтому может быть представлена в виде суммы независимых положительных квадратов:

$$\widetilde{A}(x, x) = A(x, x) + \sum_{i=1}^{r} [X_i(x)]^2.$$

Тогда при наложении r независимых связей

$$X_1(x) = 0, \quad X_2(x) = 0, \ldots, X_r(x) = 0$$

формы $A(x, x)$ и $\widetilde{A}(x, x)$ совпадают, и пучки $A(x, x) - \lambda B(x, x)$ и $\widetilde{A}(x, x) - \lambda B(x, x)$ имеют одни и те же характеристические числа

$$\lambda_1^0 \leqslant \lambda_2^0 \leqslant \ldots \leqslant \lambda_{n-r}^0.$$

Применяя теорему 14 к каждому из пучков $A(x, x) - \lambda B(x, x)$ и $\widetilde{A}(x, x) - \lambda B(x, x)$, будем иметь:

$$\widetilde{\lambda}_p \leqslant \lambda_p^0 \leqslant \lambda_{p+r} \quad (p = 1, 2, \ldots, n - r).$$

Присоединяя сюда неравенство (93), приходим к теореме:

Теорема 16. *Если* $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ *и* $\widetilde{\lambda}_1 \leqslant \widetilde{\lambda}_2 \leqslant \ldots \leqslant \widetilde{\lambda}_n$ *— характеристические числа двух регулярных пучков форм* $A(x, x) - \lambda B(x, x)$ *и* $\widetilde{A}(x, x) - \lambda B(x, x)$, *где*

$$\widetilde{A}(x, x) = A(x, x) + \sum_{i=1}^{r} [X_i(x)]^2,$$

а $X_i(x)$ $(i = 1, 2, \ldots, r)$ *— независимые линейные формы, то имеют место неравенства*[1])

$$\lambda_p \leqslant \widetilde{\lambda}_p \leqslant \lambda_{p+r} \quad (p = 1, 2, \ldots, n). \tag{94}$$

Совершенно аналогично доказывается

Теорема 17. *Если* $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ *и* $\widetilde{\lambda}_1 \leqslant \widetilde{\lambda}_2 \leqslant \ldots \leqslant \widetilde{\lambda}_n$ *— характеристические числа регулярных пучков форм* $A(x, x) - \lambda B(x, x)$, *и* $A(x, x) -$

---

[1]) Вторые части этих неравенств имеют место только при $p \leqslant n - r$.

$-\lambda\widetilde{B}(x, x)$, где форма $\widetilde{B}(x, x)$ получается из $B(x, x)$ прибавлением r положительных квадратов, то имеют место неравенства[1])

$$\lambda_{p-r} \leqslant \widetilde{\lambda}_p \leqslant \lambda_p \quad (p = 1, 2, \ldots, n). \tag{95}$$

Замечание. В теоремах 16 и 17 можно утверждать, что при некотором p имеем $\lambda_p < \widetilde{\lambda}_p$ (соответственно $\widetilde{\lambda}_p < \lambda_p$), если, конечно, $r \neq 0$[2]).

§ 8. Малые колебания системы с n степенями свободы

Результаты предыдущих двух параграфов имеют важные приложения в теории малых колебаний механической системы с n степенями свободы.

Рассмотрим свободные колебания консервативной механической системы с n степенями свободы вблизи ее устойчивого положения равновесия. Отклонение системы от положения равновесия будем задавать при помощи независимых обобщенных координат $q_1, q_2, \ldots, q_n$. Само положение равновесия при этом соответствует нулевым значениям этих координат: $q_1 = 0$, $q_2 = 0, \ldots, q_n = 0$. Тогда кинетическая энергия системы представится в виде квадратичной формы относительно обобщенных скоростей $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$[3]):

$$T = \sum_{i,\,k=1}^{n} b_{ik}(q_1, q_2, \ldots, q_n)\dot{q}_i\dot{q}_k.$$

Разлагая коэффициенты $b_{ik}(q_1, q_2, \ldots, q_n)$ в ряд по степеням $q_1, q_2, \ldots, q_n$,

$$b_{ik}(q_1, q_2, \ldots, q_n) = b_{ik} + \ldots \quad (i, k = 1, 2, \ldots, n)$$

и сохраняя (ввиду малости отклонений $q_1, q_2, \ldots, q_n$) только постоянные члены b_{ik}, будем иметь:

$$T = \sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k \quad (b_{ik} = b_{ki};\ i, k = 1, 2, \ldots, n).$$

Кинетическая энергия всегда положительна и обращается в нуль только при нулевых скоростях $\dot{q}_1 = \dot{q}_2 = \ldots = \dot{q}_n = 0$. Поэтому $\sum_{i,\,k=1}^{n} b_{ik}\dot{q}_i\dot{q}_k$ — положительно определенная форма.

Потенциальная энергия системы является функцией от координат: $\Pi(q_1, q_2, \ldots, q_n)$. Не нарушая общности, можем принять $\Pi_0 = \Pi(0, 0, \ldots, 0) = 0$. Тогда, разлагая потенциальную энергию в ряд по степеням $q_1, q_2, \ldots, q_n$, получим:

$$\Pi = \sum_{i=1}^{n} a_i q_i + \sum_{i,\,k=1}^{n} a_{ik} q_i q_k + \ldots$$

Поскольку в положении равновесия потенциальная энергия всегда имеет стационарное значение, то

$$a_i = \left(\frac{\partial \Pi}{\partial q_i}\right)_0 = 0 \quad (i = 1, 2, \ldots, n).$$

---

[1]) Первые части неравенств имеют место при $p > r$.

[2]) См. [7], стр. 71—73.

[3]) Точкой мы обозначаем производную по времени.

Сохраняя только члены второго порядка относительно $q_1, q_2, \ldots, q_n$, мы будем иметь:

$$\Pi = \sum_{i,\,k=1}^{n} a_{ik} q_i q_k \quad (a_{ik} = a_{ki};\quad i, k = 1, 2, \ldots, n).$$

Таким образом, потенциальная энергия Π и кинетическая энергия T определяются двумя квадратичными формами:

$$\Pi = \sum_{i,\,k=1}^{n} a_{ik} q_i q_k, \quad T = \sum_{i,\,k=1}^{n} b_{ik} \dot{q}_i \dot{q}_k, \tag{96}$$

причем вторая форма — положительно определенная.

Напишем теперь дифференциальные уравнения движения в форме уравнений Лагранжа второго рода [1]):

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = -\frac{\partial \Pi}{\partial q_i} \quad (i = 1, 2, \ldots, n). \tag{97}$$

Подставляя сюда вместо T и Π их выражения из (96), получаем:

$$\sum_{k=1}^{n} b_{ik} \ddot{q}_k + \sum_{k=1}^{n} a_{ik} q_k = 0 \quad (k = 1, 2, \ldots, n). \tag{98}$$

Вводя в рассмотрение вещественные симметрические матрицы

$$A = \| a_{ik} \|_1^n \quad \text{и} \quad B = \| b_{ik} \|_1^n$$

и столбцевую матрицу $q = (q_1, q_2, \ldots, q_n)$, мы систему уравнений (98) можем записать в следующей матричной форме:

$$B\ddot{q} + Aq = 0. \tag{98'}$$

Будем искать решения системы (98) в виде гармонических колебаний

$$q_1 = v_1 \sin(\omega t + \alpha), \quad q_2 = v_2 \sin(\omega t + \alpha), \quad \ldots, \quad q_n = v_n \sin(\omega t + \alpha);$$

в матричной записи

$$q = v \sin(\omega t + \alpha). \tag{99}$$

Здесь $v = (v_1, v_2, \ldots, v_n)$ — постоянный амплитудный столбец («вектор»), ω — частота и α — начальная фаза колебаний.

Подставляя выражение (99) для q в (98'), получим после сокращения на $\sin(\omega t + \alpha)$:

$$Av = \lambda B v \quad (\lambda = \omega^2).$$

Но это уравнение совпадает с уравнением (49). Следовательно, искомый амплитудный вектор является главным вектором, а квадрат частоты $\lambda = p^2$ — соответствующим характеристическим числом регулярного пучка форм $A(x, x) - \lambda B(x, x)$.

Мы наложим на потенциальную энергию дополнительное ограничение, потребовав, чтобы функция $\Pi(q_1, q_2, \ldots, q_n)$ в положении равновесия имела строгий минимум [2]).

---

[1]) См., например, Г. К. Суслов, Теоретическая механика, § 191 или Ф. Р. Гантмахер, Лекции по аналитической механике, § 6.

[2]) То есть чтобы значение Π_0 в положении равновесия было меньше всех других значений функции в некоторой окрестности положения равновесия.

Тогда на основании теоремы Лежен-Дирихле [1]) положение равновесия системы будет устойчивым. С другой стороны, сделанное нами допущение означает, что квадратичная форма $\Pi = A(q, q)$ также является положительно определенной.

Согласно теореме 8 регулярный пучок форм $A(x, x) - \lambda B(x, x)$ имеет n вещественных характеристических чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ и n соответствующих этим числам главных векторов $v^1, v^2, \ldots, v^n$ [$v^k = (v_{1k}, v_{2k}, \ldots, v_{nk})$; $k = 1, 2, \ldots, n$], удовлетворяющих условиям

$$B(v^i, v^k) = \sum_{\mu, \nu=1}^{n} b_{\mu\nu} v_{\mu i} v_{\nu k} = \delta_{ik} \quad (i, k = 1, 2, \ldots, n). \tag{100}$$

Из положительной определенности формы $A(x, x)$ следует, что все характеристические числа пучка $A(x, x) - \lambda B(x, x)$ положительны [2]):

$$\lambda_k > 0 \quad (k = 1, 2, \ldots, n).$$

Но тогда существует n гармонических колебаний [3])

$$v^k \sin(\omega_k t + \alpha_k) \quad (\omega_k^2 = \lambda_k, \ k = 1, 2, \ldots, n), \tag{101}$$

амплитудные векторы которых $v^k = (v_{1k}, v_{2k}, \ldots, v_{nk})$ $(k = 1, 2, \ldots, n)$ удовлетворяют условиям «ортонормированности» (100).

В силу линейности уравнения (98′) произвольное колебание может быть получено наложением гармонических колебаний (101):

$$q = \sum_{k=1}^{n} A_k \sin(\omega_k t + \alpha_k) v^k, \tag{102}$$

где A_k, α_k $(k = 1, 2, \ldots, n)$ — произвольные постоянные. Действительно, при любых значениях этих постоянных выражение (102) является решением уравнения (98′). С другой стороны, за счет произвольных постоянных можно удовлетворить любым начальным условиям:

$$q|_{t=0} = q_0, \quad \dot{q}|_{t=0} = \dot{q}_0.$$

В самом деле, из (102) находим:

$$q_0 = \sum_{k=1}^{n} A_k \sin \alpha_k v^k, \quad \dot{q}_0 = \sum_{k=1}^{n} \omega_k A_k \cos \alpha_k v^k. \tag{103}$$

Поскольку главные столбцы $v^1, v^2, \ldots, v^n$ всегда линейно независимы, то из равенств (103) однозначно определяются величины $A_k \sin \alpha_k$ и $\omega_k A_k \cos \alpha_k$ $(k = 1, 2, \ldots, n)$ и, следовательно, произвольные постоянные A_k и α_k $(k = 1, 2, \ldots, n)$.

Решение (102) нашей системы дифференциальных уравнений (98) может быть более подробно записано так:

$$q_i = \sum_{k=1}^{n} A_k \sin(\omega_k t + \alpha_k) v_{ik}. \tag{104}$$

Заметим, что к тем же формулам (102), (104) можно прийти, исходя из теоремы 9. Действительно, рассмотрим неособенное преобразование

---

[1]) См. Г. К. Суслов, Теоретическая механика, § 210; Ф. Р. Гантмахер, Лекции по аналитической механике, § 33.

[2]) Это следует хотя бы из представления (63).

[3]) Здесь начальные фазы α_k $(k = 1, 2, \ldots, n)$ — произвольные постоянные.

переменных с матрицей $V = \| v_{ik} \|_1^n$, приводящее одновременно обе формы $A(x, x)$ и $B(x, x)$ к каноническому виду (63). Полагая

$$q_i = \sum_{k=1}^{n} v_{ik}\theta_k \quad (i = 1, 2, \ldots, n) \tag{105}$$

или в сокращенной записи

$$q = V\theta \quad [\theta = (\theta_1, \theta_2, \ldots, \theta_n)] \tag{106}$$

и замечая, что $\dot{q} = V\dot{\theta}$, будем иметь:

$$\Pi = A(q, q) = \sum_{i=1}^{n} \lambda_k \theta_k^2, \qquad T = B(q, q) = \sum_{k=1}^{n} \dot{\theta}_k^2. \tag{107}$$

Координаты $\theta_1, \theta_2, \ldots, \theta_n$, в которых потенциальная и кинетическая энергии представляются в виде (107), называются *нормальными координатами*.

Воспользуемся уравнениями Лагранжа второго рода (98), подставив в них вместо Π и T их выражения (107). Получим:

$$\ddot{\theta}_k + \lambda_k \theta_k = 0 \quad (k = 1, 2, \ldots, n). \tag{108}$$

Поскольку форма $A(q, q)$ — положительно определенная, то все числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ положительны и могут быть представлены в виде

$$\lambda_k = \omega_k^2 \quad (\omega_k > 0;\ k = 1, 2, \ldots, n). \tag{109}$$

Из (107) и (108) находим:

$$\theta_k = A_k \sin(\omega_k t + \alpha_k) \quad (k = 1, 2, \ldots, n). \tag{110}$$

Подставляя эти выражения для θ_k в равенства (105), получим снова формулы (104) и, следовательно, (102). Величины v_{ik} $(i, k = 1, 2, \ldots, n)$ при обоих выводах будут одни и те же, поскольку согласно теореме 9 матрица $V = \| v_{ik} \|_1^n$ в (106) есть главная матрица регулярного пучка форм $A(x, x) - \lambda B(x, x)$.

Отметим еще механическую интерпретацию теорем 14 и 15.

Занумеруем частоты $\omega_1, \omega_2, \ldots, \omega_n$ данной механической системы в порядке неубывания:

$$0 < \omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n.$$

Этим определится и расположение соответствующих характеристических чисел $\lambda_k = \omega_k^2$ $(k = 1, 2, \ldots, n)$ пучка $A(x, x) - \lambda B(x, x)$:

$$\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n.$$

Наложим на данную систему h независимых конечных стационарных связей[1]). Поскольку отклонения $q_1, q_2, \ldots, q_n$ считаются малыми величинами, то эти связи можно считать линейными относительно $q_1, q_2, \ldots, q_n$:

$$L_1(q) = 0, \quad L_2(q) = 0, \quad \ldots, \quad L_h(q) = 0.$$

После наложения связей наша система будет иметь $n - h$ степеней свободы. Частоты этой системы

$$\omega_1^0 \leqslant \omega_2^0 \leqslant \ldots \leqslant \omega_{n-h}^0$$

[1]) Конечная стационарная связь выражается уравнением $f(q_1, q_2, \ldots, q_n) = 0$, где $f(q_1, q_2, \ldots, q_n)$ — некоторая функция от обобщенных координат.

связаны с характеристическими числами $\lambda_1^0 \leqslant \lambda_2^0 \leqslant \ldots \leqslant \lambda_{n-h}^0$ пучка $A(x, x) - \lambda B(x, x)$ при наложении связей $L_1, L_2, \ldots, L_h$ соотношениями $\lambda_j^0 = \omega_j^{02}$ $(j = 1, 2, \ldots, n-h)$. Поэтому из теоремы 14 непосредственно следует:

$$\omega_j \leqslant \omega_j^0 \leqslant \omega_{j+h} \quad (j = 1, 2, \ldots, n-h).$$

Таким образом, *при наложении h связей частоты системы могут только увеличиться, однако при этом величина новой j-й частоты ω_j^0 не может превзойти величины прежней $(j+h)$-й частоты ω_{j+h}.*

Точно так же на основании теоремы 15 можно утверждать, что *при увеличении жесткости системы, т. е. при увеличении формы $A(q, q)$ для потенциальной энергии [без изменения формы $B(q, q)$], частоты могут только увеличиться, а при увеличении инерции системы, т. е. при увеличении формы $B(\dot{q}, \dot{q})$ для кинетической энергии [без изменения формы $A(q, q)$], частоты могут только уменьшиться.*

Теоремы 16 и 17 вносят дополнительное уточнение в это положение.

§ 9. Эрмитовы формы [1])

Все результаты §§ 1—7 этой главы, установленные для вещественных квадратичных форм, могут быть перенесены на эрмитовы формы.

Напомним [2]), что *эрмитовой формой* называется выражение

$$H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k \quad (h_{ik} = \bar{h}_{ki}; \quad k = 1, 2, \ldots, n). \tag{111}$$

Эрмитовой форме (111) соответствует следующая *билинейная эрмитова форма*:

$$H(x, y) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{y}_k; \tag{112}$$

при этом

$$H(y, x) = \overline{H(x, y)} \tag{113}$$

и, в частности,

$$H(x, x) = \overline{H(x, x)}, \tag{113'}$$

т. е. форма $H(x, x)$ принимает только вещественные значения.

Матрица коэффициентов эрмитовой формы $H = \| h_{ik} \|_1^n$ является эрмитовой, т. е. $H^* = H$ [3]).

Пользуясь матрицей $H = \| h_{ik} \|_1^n$, можно представить $H(x, y)$ и, в частности, $H(x, x)$ в виде произведения трех матриц — строчной, квадратной и столбцевой:

$$H(x, y) = x' H \bar{y}, \quad H(x, x) = x' H \bar{x} \text{ [4])}. \tag{114}$$

---

[1]) В предыдущих параграфах все числа и переменные были вещественными. В этом же параграфе все числа и переменные принимают комплексные значения.

[2]) См. гл. IX, § 2.

[3]) Звездочкой * мы отмечаем переход к сопряженной матрице (см. гл. I, § 3).

[4]) Здесь $x = x(x_1, x_2, \ldots, x_n)$, $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$, $y = (y_1, y_2, \ldots, y_n)$, $\bar{y} = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$; значок $'$ означает транспонирование.

Если

$$x=\sum_{i=1}^{m} c_i u^i, \quad y=\sum_{k=1}^{p} d_k v^k, \tag{115}$$

где u^i, v^k — столбцевые матрицы, c_i, d_k — комплексные числа ($i=1, 2, \ldots, m$; $k=1, 2, \ldots, p$), то

$$H(x, y)=\sum_{i=1}^{m}\sum_{k=1}^{p} c_i\bar{d}_k H(u^i, v^k). \tag{116}$$

Подвергнем переменные $x_1, x_2, \ldots, x_n$ линейному преобразованию

$$x_i=\sum_{k=1}^{n} t_{ik}\xi_k \quad (i=1, 2, \ldots, n), \tag{117}$$

или в матричной записи

$$x=T\xi \quad (T=\|t_{ik}\|_1^n). \tag{117'}$$

После преобразования эрмитова форма $H(x, x)$ примет вид

$$\tilde{H}(\xi, \xi)=\sum_{i, k=1}^{n} \tilde{h}_{ik}\xi_i\bar{\xi}_k,$$

где матрица новых коэффициентов $\tilde{H}=\|\tilde{h}_{ik}\|_1^n$ связана с матрицей старых коэффициентов $H=\|h_{ik}\|_1^n$ формулой

$$\tilde{H}=T'H\bar{T}. \tag{118}$$

В этом непосредственно убеждаемся после замены во второй формуле (114) x на $T\xi$.

Если положить $T=\overline{W}$, то формулу (118) можно еще переписать так:

$$\tilde{H}=W^*HW. \tag{119}$$

Из формулы (118) следует, что ранги матриц H и $\tilde{H}$ равны, если преобразование (117) — неособенное ($|T|\neq 0$). Ранг матрицы H называется *рангом формы* $H(x, x)$.

Определитель $|H|$ называется *дискриминантом* эрмитовой формы $H(x, x)$. Из (118) следует формула преобразования дискриминанта при переходе к новым переменным:

$$|\tilde{H}|=|H||T||\bar{T}|.$$

Эрмитова форма называется *сингулярной*, если ее дискриминант равен нулю. Очевидно, сингулярная форма остается сингулярной при любом преобразовании переменных (117).

Эрмитову форму $H(x, x)$ можно бесчисленным множеством способов представить в виде

$$H(x, x)=\sum_{i=1}^{r} a_i X_i\overline{X}_i, \tag{120}$$

где $a_i\neq 0$ $(i=1, 2, \ldots, r)$ — вещественные числа, а

$$X_i=\sum_{k=1}^{n} a_{ik}x_k \quad (i=1, 2, \ldots, r)$$

— независимые комплексные линейные формы от переменных $x_1, x_2, \ldots, x_n$ [1]).

[1]) Следовательно, $r\leqslant n$.

Правую часть в (120) будем называть *суммой независимых квадратов*[1]), а каждое слагаемое в этой сумме — *положительным* или *отрицательным квадратом* в зависимости от того, соответствующее $a_i > 0$ или < 0. Как и для квадратичных форм, число r в (120) равно рангу формы $H(x, x)$.

Теорема 18 (закон инерции эрмитовых форм). *При представлении эрмитовой формы $H(x, x)$ в виде суммы независимых квадратов*

$$H(x, x) = \sum_{i=1}^{r} a_i X_i \overline{X}_i$$

число положительных и число отрицательных квадратов не зависят от способа представления.

Доказательство совершенно аналогично доказательству теоремы 1 (стр. 270).

Разность σ между числом π положительных и числом ν отрицательных квадратов в (120) называется сигнатурой эрмитовой формы $H(x, x)$: $\sigma = \pi - \nu$.

Метод Лагранжа приведения квадратичных форм к сумме квадратов может быть использован и для эрмитовых форм, только при этом основные формулы (15) и (16) на стр. 271 должны быть заменены формулами[2])

$$H(x, x) = \frac{1}{h_{gg}} \left| \sum_{k=1}^{n} h_{kg} x_k \right|^2 + H_1(x, x), \tag{121}$$

$$H(x, x) = \frac{1}{2} \left\{ \left| \sum_{k=1}^{n} \left(h_{kf} + \frac{h_{kg}}{h_{fg}} \right) x_k \right|^2 - \left| \sum_{k=1}^{n} \left(h_{kf} - \frac{h_{kg}}{h_{fg}} \right) x_k \right|^2 \right\} + H_2(x, x). \tag{122}$$

Пусть теперь для эрмитовой формы $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i x_k$ ранга r выполняются неравенства

$$D_k = H \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0. \tag{123}$$

Тогда совершенно так же, как и для квадратичной формы (см. стр. 273—274), получаем формулу Якоби в двух видах:

$$H(x, x) = \sum_{k=1}^{n} \frac{D_{k-1}}{D_k} X_k \overline{X}_k, \qquad H(x, x) = \sum_{k=1}^{n} \frac{Y_k \overline{Y}_k}{D_{k-1} D_k} \quad (D_0 = 1), \tag{124}$$

где

$$X_k = \frac{1}{D_k} Y_k, \quad Y_k = c_{kk} x_k + c_{k, k+1} x_k + \dots + c_{kn} x_n \quad (k = 1, \dots, r), \tag{125}$$

а

$$c_{kq} = H \begin{pmatrix} 1 & \dots & k-1 & q \\ 1 & \dots & k-1 & k \end{pmatrix} \qquad (q = k, k+1, \dots, n; \ k = 1, \dots, r). \tag{126}$$

---

[1]) Эта терминология связана с тем, что произведение $X_i \overline{X}_i$ равно квадрату модуля X_i ($X_i \overline{X}_i = |X_i|^2$).

[2]) Формула (121) применяется в случае, когда $h_{gg} \neq 0$, а формула (122) — в случае, когда $h_{ff} = h_{gg} = 0$, а $h_{fg} \neq 0$.

В соответствии с формулой Якоби (124) число отрицательных квадратов в представлении формы $H(x, x)$ равно числу знакоперемен в ряду $1, D_1, D_2, \ldots, D_r$

$$\nu = V(1, D_1, D_2, \ldots, D_r) \tag{127}$$

и, следовательно, сигнатура σ эрмитовой формы $H(x, x)$ определится формулой

$$\sigma = r - 2V(1, D_1, D_2, \ldots, D_r). \tag{128}$$

Все замечания относительно особых случаев, которые могут здесь представиться, сделанные для квадратичных форм (§ 3), автоматически переносятся на эрмитовы формы.

Определение 5. Эрмитова форма $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ называется *неотрицательной* (*неположительной*), если при любых значениях переменных

$$H(x, x) \geqslant 0 \quad (\text{соответственно } \leqslant 0).$$

Определение 6. Эрмитова форма $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ называется *положительно определенной* (*отрицательно определенной*), если при любых значениях переменных $x_1, x_2, \ldots, x_n$, не равных одновременно нулю,

$$H(x, x) > 0 \quad (\text{соответственно } < 0).$$

Теорема 19. *Для того чтобы эрмитова форма* $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ *была положительно определенной, необходимо и достаточно, чтобы имели место неравенства*

$$D_k = H\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix} > 0 \quad (k = 1, 2, \ldots, n). \tag{129}$$

Теорема 20. *Для того чтобы эрмитова форма* $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ *была неотрицательной, необходимо и достаточно, чтобы все главные миноры матрицы* $H = \| h_{ik} \|_1^n$ *были неотрицательны:*

$$H\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \geqslant 0 \qquad (i_1, i_2, \ldots, i_p = 1, 2, \ldots, n;\ p = 1, 2, \ldots, n). \tag{130}$$

Доказательство теорем 19 и 20 совершенно аналогично доказательству теорем 3 и 4 для квадратичных форм.

Условия отрицательной определенности и неположительности эрмитовой формы $H(x, x)$ получаются соответственно из условий (129) и (130), если последние применить к форме $-H(x, x)$.

Из теоремы 5′ главы IX (стр. 248) следует теорема о приведении эрмитовой формы к главным осям:

Теорема 21. *Эрмитова форма* $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ *всегда может быть приведена при помощи унитарного преобразования переменных*

$$x = U\xi \qquad (UU^* = E) \tag{131}$$

к канонической форме

$$\Lambda(\xi, \bar{\xi}) = \sum_{i=1}^{n} \lambda_i \xi_i \bar{\xi}_i; \tag{132}$$

при этом $\lambda_1, \lambda_2, \ldots, \lambda_n$ — *характеристические числа матрицы* $H = \| h_{ik} \|_1^n$.

Справедливость теоремы 21 вытекает из формулы

$$H = U \| \lambda_i \delta_{ik} \| U^{-1} = T' \| \lambda_i \delta_{ik} \| \bar{T} \quad (U' = \bar{U}^{-1} = T). \tag{133}$$

Пусть даны две эрмитовы формы $H(x, x) = \sum_{i, k=1}^{n} h_{ik} x_i \bar{x}_k$ и $G(x, x) = \sum_{i, k=1}^{n} g_{ik} x_i \bar{x}_k$. Рассмотрим пучок эрмитовых форм $H(x, x) - \lambda G(x, x)$ (λ — вещественный параметр). Этот пучок называется *регулярным*, если форма $G(x, x)$ — положительно определенная. При помощи эрмитовых матриц $H = \| h_{ik} \|_1^n$ и $G = \| g_{ik} \|_1^n$ составим уравнение

$$| H - \lambda G | = 0.$$

Это уравнение называется *характеристическим уравнением пучка эрмитовых форм.* Корни этого уравнения называются *характеристическими числами пучка.*

Если λ_0 — характеристическое число пучка, то существует столбец $z = (z_1, z_2, \ldots, z_n) \neq 0$ такой, что

$$Hz = \lambda_0 z.$$

Столбец z мы будем называть *главным столбцом* или *главным вектором пучка* $H(x, x) - \lambda G(x, x)$, соответствующим характеристическому числу λ_0.

Имеет место

Теорема 22. *Характеристическое уравнение регулярного пучка эрмитовых форм* $H(x, x) - \lambda G(x, x)$ *имеет* n *вещественных корней* $\lambda_1, \lambda_2, \ldots, \lambda_n$. *Этим корням соответствуют* n *главных векторов* $z^1, z^2, \ldots, z^n$, *удовлетворяющих условиям «ортонормированности»*:

$$G(z^i, z^k) = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n).$$

Доказательство совершенно аналогично доказательству теоремы 8.

Все экстремальные свойства характеристических чисел регулярного пучка квадратичных форм сохраняют свою силу и для эрмитовых форм.

Теоремы 10 — 17 сохраняют свою силу, если в этих теоремах термин «квадратичные формы» заменить везде термином «эрмитовы формы». Доказательства теорем остаются при этом неизменными.

§ 10. Ганкелевы формы

Пусть даны $2n - 1$ чисел $s_0, s_1, \ldots, s_{2n-2}$. При помощи этих чисел составим квадратичную форму от n переменных

$$S(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k. \tag{134}$$

Квадратичная форма (134) называется *ганкелевой.* Соответствующая ей симметрическая матрица $S = \| s_{i+k} \|_0^{n-1}$ также называется *ганкелевой.*

Эта матрица имеет вид

$$S=\begin{Vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{Vmatrix}.$$

Последовательные главные миноры матрицы S будем обозначать через $D_1, D_2, \dots, D_n$:

$$D_p=|s_{i+k}|_0^{p-1} \qquad (p=1, 2, \dots, n).$$

В настоящем параграфе мы установим основные результаты Фробениуса относительно ранга и сигнатуры вещественных ганкелевых форм[1]).

Предварительно докажем две леммы.

Лемма 1. *Если в ганкелевой матрице* $S=\|s_{i+k}\|_0^{n-1}$ *первые* h *строк линейно независимы, а первые* $h+1$ *строк линейно зависимы, то*

$$D_h\neq 0.$$

Доказательство. Обозначим через $\Gamma_1, \Gamma_2, \dots, \Gamma_h, \Gamma_{h+1}$ первые $h+1$ строк матрицы S. По условию теоремы строки $\Gamma_1, \Gamma_2, \dots, \Gamma_h$ линейно независимы, а строка Γ_{h+1} линейно выражается через эти строки:

$$\Gamma_{h+1}=\sum_{j=1}^{h}\alpha_j\Gamma_{h-j+1}$$

или

$$s_q=\sum_{j=1}^{h}\alpha_j s_{q-j} \quad (q=h,\ h+1,\ \dots,\ h+n-1). \tag{135}$$

Выпишем матрицу, состоящую из первых h строк $\Gamma_1, \Gamma_2, \dots, \Gamma_h$ матрицы S:

$$\begin{Vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ \dots & \dots & \dots & \dots & \dots \\ s_{h-1} & s_h & s_{h+1} & \dots & s_{h+n-2} \end{Vmatrix}. \tag{136}$$

Эта матрица имеет ранг h. С другой стороны, в силу (135) любой столбец этой матрицы выражается линейно через h предыдущих столбцов. Следовательно, любой столбец матрицы выражается линейно через h первых столбцов. Но тогда, поскольку ранг матрицы (136) равен h, эти первые h столбцов матрицы (136) должны быть линейно независимы, т. е.

$$D_h\neq 0.$$

Лемма доказана.

Лемма 2. *Если для матрицы* $S=\|s_{i+k}\|_0^{n-1}$ *при некотором* $h\,(<n)$

$$D_h\neq 0, \quad D_{h+1}=\dots=D_n=0 \tag{137}$$

и

$$t_{ik}=\frac{S\begin{pmatrix}1 & \dots & h & h+i+1\\ 1 & \dots & h & h+k+1\end{pmatrix}}{S\begin{pmatrix}1 & \dots & h\\ 1 & \dots & h\end{pmatrix}}=\frac{1}{D_h}\begin{vmatrix} D_h & \begin{matrix}s_{h+k}\\ \cdot\\ \cdot\\ \cdot\\ s_{2h+k-1}\end{matrix} \\ s_{h+i}\ \dots\ s_{2h+i-1} & s_{2h+i+k}\end{vmatrix} \tag{138}$$

$$(i, k=0, 1, \dots, n-h-1)$$

[1]) См. [182f].

то матрица $T=\|t_{ik}\|_0^{n-h-1}$ *также ганкелева и все ее элементы, расположенные над второй диагональю, равны нулю, т. е. существуют такие числа* $t_{n-h-1}, \ldots, t_{2n-2h-2}$, *что*

$$t_{ik}=t_{i+k} \quad (i, k=0, 1, \ldots, n-h-1; \quad t_0=t_1=\ldots=t_{n-h-2}=0).$$

Доказательство. Введем в рассмотрение матрицы

$$T_p=\|t_{ik}\|_0^{p-1} \qquad (p=1, 2, \ldots, n-h).$$

В этих обозначениях $T=T_{n-h}$.

Мы докажем, что любая из матриц $T_p (p=1, 2, \ldots, n-h)$ является ганкелевой и что в ней $t_{ik}=0$ при $i+k\leqslant p-2$. Доказательство будем вести индуктивно относительно p.

Для матрицы T_1 наше утверждение тривиально, для матрицы T_2 оно очевидно, так как

$$T_2=\begin{Vmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{Vmatrix}, \quad t_{01}=t_{10} \text{ (в силу симметрии } S) \text{ и } t_{00}=\frac{D_{h+1}}{D_h}=0.$$

Допустим, что наше утверждение справедливо для матрицы $T_p (p<n-h)$, и докажем его справедливость для матрицы $T_{p+1}=\|t_{ik}\|_0^p$. Из допущения следует существование таких чисел $t_{p-1}, t_p, \ldots, t_{2p-2}$, что при $t_0=\ldots=t_{p-2}=0$

$$T_p=\|t_{i+k}\|_0^{p-1}.$$

При этом

$$|T_p|=\pm t_{p-1}^p. \tag{139}$$

С другой стороны, пользуясь детерминантным тождеством Сильвестра [см. (28) на стр. 48], найдем:

$$|T_p|=\frac{D_{h+p}}{D_h}=0. \tag{140}$$

Из сопоставления (139) с (140) получаем:

$$t_{p-1}=0. \tag{141}$$

Далее из (138):

$$t_{ik}=s_{2h+i+k}+\frac{1}{D_h}\begin{vmatrix} & & s_{h+k} \\ & D_h & \vdots \\ & & s_{2h+k-1} \\ s_{h+i} \ \ldots \ s_{2h+i-1} & & 0 \end{vmatrix}. \tag{142}$$

На основании предыдущей леммы из (137) следует, что $(h+1)$-я строка матрицы $S=\|s_{i+k}\|_0^{n-1}$ линейно зависит от первых h строк:

$$s_q=\sum_{g=1}^{h}\alpha_g s_{q-g} \quad (q=h, h+1, \ldots, h+n-1). \tag{143}$$

Пусть $i, k\leqslant p\leqslant i+k\leqslant 2p-1$. При этом одно из чисел i и k меньше p. Не нарушая общности рассуждений, примем, что $i<p$. Тогда, разлагая с помощью (143) последний столбец в определителе, стоящем в правой

части равенства (142), и снова используя соотношения (142), будем иметь:

$$t_{ik} = s_{2h+i+k} + \sum_{g=1}^{h} \frac{\alpha_g}{D_h} \begin{vmatrix} & & s_{h+k-g} \\ & D_h & \cdot \\ & & \cdot \\ & & s_{2h+k-g-1} \\ s_{h+i} \ \ldots \ s_{2h+i-1} & & 0 \end{vmatrix} =$$

$$= s_{2h+i+k} + \sum_{g=1}^{h} \alpha_g (t_{i,\,k-g} - s_{2h+i+k-g}). \qquad (144)$$

Но в силу допущения индукции имеет место (141), и поскольку в (144) $i < p$, $k-g < p$ и $i+k-g \leqslant 2p-2$, то $t_{i,\,k-g} = t_{i+k-g}$. Следовательно, при $i+k < p$ все $t_{ik} = 0$, а при $p \leqslant i+k \leqslant 2p-1$ величина t_{ik} в силу (144) зависит только от $i+k$.

Таким образом, T_{p+1} — ганкелева матрица и в этой матрице все элементы $t_0, t_1, \ldots, t_{p-1}$, стоящие над второй диагональю, равны нулю.

Лемма доказана.

Пользуясь леммой 2, докажем следующую теорему:

Теорема 23. *Если ганкелева матрица* $S = \| s_{i+k} \|_0^{n-1}$ *имеет ранг* r *и при некотором* $h\,(<r)$

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_r = 0,$$

то главный минор r*-го порядка, образованный первыми* h *и последними* $r-h$ *линиями матрицы* S*, не равен нулю:*

$$D^{(r)} = S \begin{pmatrix} 1 \ \ldots \ h, & n-r+h+1, & n-r+h+2, & \ldots, & n \\ 1 \ \ldots \ h, & n-r+h+1, & n-r+h+2, & \ldots, & n \end{pmatrix} \neq 0.$$

Доказательство. На основании предыдущей леммы матрица

$$T = \| t_{ik} \|_0^{n-h-1} \left[t_{ik} = \frac{S \begin{pmatrix} 1 \ \ldots \ h & h+i+1 \\ 1 \ \ldots \ h & h+k+1 \end{pmatrix}}{S \begin{pmatrix} 1 \ \ldots \ h \\ 1 \ \ldots \ h \end{pmatrix}} (i,\ k = 0,\ 1,\ \ldots,\ n-h-1) \right]$$

есть ганкелева матрица, в которой все элементы над второй диагональю равны нулю. Поэтому

$$|T| = \pm t_{0,\,n-h-1}^{n-h}.$$

С другой стороны [1]), $|T| = \frac{D_n}{D_h} = 0$. Следовательно, $t_{0,\,n-h-1} = 0$, и матрица T имеет вид

$$T = \begin{Vmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & u_{n-h-1} \\ \vdots & & & \iddots & \vdots \\ \vdots & & \iddots & & \vdots \\ \vdots & \iddots & & & \vdots \\ 0 & u_{n-h-1} & \cdots & u_2 & u_1 \end{Vmatrix}.$$

[1]) На основании детерминантного тождества Сильвестра [см. (28) на стр. 48].

Матрица T должна иметь ранг $r-h$[1]). Поэтому при $r<n-1$ в матрице T элементы $u_{r-h+1}=\ldots=u_{n-h+1}=0$ и матрица T всегда имеет вид

$$T=\begin{Vmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & \iddots & \vdots \\ \vdots & & \iddots & & 0 \\ \vdots & \iddots & & \iddots & \vdots \\ \vdots & & \iddots & & u_{r-h} \\ \vdots & \iddots & & \iddots & \vdots \\ \vdots & & \iddots & & \vdots \\ 0 & \ldots & 0 & u_{r-h} \ldots & u_1 \end{Vmatrix} \quad (u_{r-h}\neq 0).$$

Но тогда в силу тождества Сильвестра (см. стр. 48)

$$D^{(r)}=D_h T\begin{pmatrix} n-r+1 & \ldots & n-h \\ n-r+1 & \ldots & n-h \end{pmatrix}=\pm D_h u_{r-h}^{r-h}\neq 0,$$

что и требовалось доказать.

Рассмотрим вещественную[2]) ганкелеву форму

$$S(x, x)=\sum_{i,\,k=0}^{\infty} s_{i+k}x_i x_k$$

ранга r. Обозначим через π, ν, σ соответственно число положительных квадратов, число отрицательных квадратов и сигнатуру этой формы:

$$\pi+\nu=r,$$
$$\sigma=\pi-\nu=r-2\nu.$$

Согласно теореме Якоби (стр. 275) эти величины могут быть определены из рассмотрения знаков последовательных миноров

$$D_0=1,\ D_1,\ D_2,\ \ldots,\ D_{r-1},\ D_r \tag{145}$$

при помощи формул

$$\left.\begin{aligned} &\pi=P(1, D_1, \ldots, D_r),\ \ \nu=V(1, D_1, \ldots, D_r), \\ &\sigma=P(1, D_1, \ldots, D_r)-V(1, D_1, \ldots, D_r)=r-2V(1, D_1, \ldots, D_r). \end{aligned}\right\} \tag{146}$$

Эти формулы становятся неприменимыми в случае, когда последний член в ряду (145) либо три подряд идущих промежуточных члена равны нулю (см. § 3). Однако для ганкелевой формы, как показал Фробениус,

---

[1]) Из тождества Сильвестра следует, что все миноры матрицы T, у которых порядок $>r-h$, равны нулю. С другой стороны, матрица S содержит некоторый окаймляющий D_h минор r-го порядка, отличный от нуля. Отсюда следует, что и соответствующий минор порядка $r-h$ матрицы T отличен от нуля.

[2]) В предыдущих леммах 1, 2 и теореме 23 в качестве основного поля можно было брать произвольное числовое поле и, в частности, поле всех комплексных чисел или поле всех вещественных чисел.

можно дать правило, позволяющее использовать формулы (146) в самом общем случае:

Теорема 24 (Фробениуса). *Для вещественной ганкелевой формы*

$$S(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k$$

ранга r величины π, ν, σ могут быть определены из формул (146), *если*

1) *при*

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_r = 0 \quad (h < r) \tag{147}$$

заменить в этих формулах D_r на $D^{(r)}$, где

$$D^{(r)} = S \begin{pmatrix} 1 & \ldots & h & n-r+h+1 & \ldots & n \\ 1 & \ldots & h & n-r+h+1 & \ldots & n \end{pmatrix} \neq 0;$$

2) *в любой группе из p промежуточных нулевых определителей*

$$(D_h \neq 0)\, D_{h+1} = D_{h+2} = \ldots = D_{h+p} = 0 \quad (D_{h+p+1} \neq 0) \tag{148}$$

нулевым определителям приписать знаки по формуле

$$\operatorname{sign} D_{h+j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} D_h. \tag{149}$$

При этом величины P, V, $P-V$, соответствующие группе (148), *получат значения*[1])

	p нечетно	p четно
$P_{h,p} = P(D_h, D_{h+1}, \ldots, D_{h+p+1})$	$\frac{p+1}{2}$	$\frac{p+1+\varepsilon}{2}$
$V_{h,p} = V(D_h, D_{h+1}, \ldots, D_{h+p+1})$	$\frac{p+1}{2}$	$\frac{p+1-\varepsilon}{2}$
$P_{h,p} - V_{h,p}$	0	ε

; (150)

$$\varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \frac{D_{h+p+1}}{D_h}.$$

Доказательство. Рассмотрим сначала случай, когда $D_r \neq 0$. В этом случае формы

$$S(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k \quad \text{и} \quad S_r(x, x) = \sum_{i, k=0}^{r-1} s_{i+k} x_i x_k$$

имеют не только один и тот же ранг r, но и одну и ту же сигнатуру σ.

[1]) Формулы (149) и (150) применимы и к случаю (147), только здесь нужно положить $p = r-h-1$ и под D_{h+p+1} понимать не $D_r = 0$, а $D^{(r)} \neq 0$.

Действительно, пусть

$$S(x, x) = \sum_{i=1}^{r} \varepsilon_i Z_i^2$$

где Z_i — вещественные линейные формы, а $\varepsilon_i = \pm 1$ $(i = 1, 2, \ldots, r)$. Положим $x_{r+1} = \ldots = x_n = 0$ Тогда формы $S(x, x)$, Z_i перейдут соответственно в $S_r(x, x)$, $\widehat{Z}_i (i = 1, 2, \ldots, r)$, причем $S_r(x, x) = \sum_{i=1}^{r} \varepsilon_i \widehat{Z}_i^2$, т. е. $S_r(x, x)$ имеет такое же число положительных (отрицательных) независимых квадратов, как и форма $S(x, x)$ [1]). Таким образом, σ есть сигнатура формы $S_r(x, x)$.

Проварьируем непрерывно параметры $s_0, s_1, \ldots, s_{2r-2}$ так, чтобы при новых значениях параметров $s_0^*, s_1^*, \ldots, s_{2r-2}^*$ [2]) все члены ряда

$$1, D_1^*, D_2^*, \ldots, D_r^* \quad (D_q^* = |s_{i+k}^*|_0^{q-1};\ q = 1, 2, \ldots, r)$$

были отличны от нуля и чтобы в процессе варьирования ни один из отличных от нуля определителей (145) не обратился в нуль [3]).

Так как при варьировании не изменялся ранг формы $S_r(x, x)$, то не изменялась и ее сигнатура (см. стр. 280). Поэтому

$$\sigma = P(1, D_1^*, \ldots, D_r^*) - V(1, D_1^*, \ldots, D_r^*). \tag{151}$$

Если при некотором i $D_i \neq 0$, то

$$\operatorname{sign} D_i^* = \operatorname{sign} D_i.$$

Поэтому весь вопрос сводится к определению перемен знака между теми D_i^*, которым соответствуют $D_i = 0$. Точнее, для каждой группы вида (148) требуется определить

$$P(D_h^*, D_{h+1}^*, \ldots, D_{h+p+1}^*) - V(D_h^*, D_{h+1}^*, \ldots, D_{h+p}^*, D_{h+p+1}^*).$$

Для этого положим:

$$t_{ik} = \frac{1}{D_h} \begin{vmatrix} & & s_{h+k} \\ & D_h & \cdot \\ & & \cdot \\ & & s_{2h+k-1} \\ s_{h+i} & \ldots\ s_{2h+i-1} & s_{2h+i+k} \end{vmatrix} \quad (i, k = 0, 1, \ldots, p).$$

Согласно лемме 2 матрица

$$T = \| t_{ik} \|_0^p$$

ганкелева и все элементы ее, стоящие над второй диагональю, равны

---

[1]) Линейные формы $\widehat{Z}_1, \widehat{Z}_2, \ldots, Z_r$ линейно независимы, поскольку квадратичная форма $S_r(x, x) = \sum_{i=1}^{r} \varepsilon_i \widehat{Z}_i^2$ имеет ранг r $(D_r \neq 0)$.

[2]) В этом параграфе значок* не означает перехода к сопряженной матрице.

[3]) Такую вариацию всегда можно осуществить, поскольку в пространстве параметров $s_0, s_1, \ldots, s_{2r-2}$ уравнение вида $D_i = 0$ определяет некоторую алгебраическую гиперповерхность. Если точка принадлежит нескольким таким гиперповерхностям, то она может быть всегда аппроксимирована сколь угодно близкими точками, лежащими вне этих гиперповерхностей.

нулю, т. е. матрица T имеет вид

$$T=\begin{Vmatrix} 0 & \dots & 0 & t_p \\ \cdot & \iddots & \iddots & * \\ \cdot & \iddots & \iddots & \cdot \\ \cdot & \iddots & & \cdot \\ 0 & \iddots & & \cdot \\ t_p & * & \dots & * \end{Vmatrix}. \tag{152}$$

Обозначим последовательные миноры матрицы T через $\widehat{D}_1, \widehat{D}_2, \dots, \widehat{D}_{p+1}$:

$$\widehat{D}_q=|t_{ik}|_0^{q-1} \quad (q=1, 2, \dots, p+1).$$

Наряду с матрицей T введем в рассмотрение матрицу

$$T^*=\|t_{ik}^*\|_0^p,$$

где

$$t_{ik}^*=\frac{1}{D_h^*}\begin{vmatrix} & & s_{h+k}^* \\ & D_h^* & \cdot \\ & & \cdot \\ & & s_{2h+k-1}^* \\ s_{h+i} & \dots\ s_{2h+i-1}^* & s_{2h+i+k}^* \end{vmatrix} \quad (i, k=0, 1, \dots, p),$$

и соответственные определители

$$\widehat{D}_q^*=|t_{ik}^*|_0^{q-1} \quad (q=1, 2, \dots, p+1).$$

Согласно детерминантному тождеству Сильвестра

$$D_{h+q}^*=D_h^*\widehat{D}_q^* \quad (q=1, 2, \dots, p+1).$$

Поэтому

$$P(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*)-V(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*)=$$

$$=\widehat{P}(1, \widehat{D}_1^*, \dots, \widehat{D}_{p+1}^*)-V(1, \widehat{D}_1^*, \dots, \widehat{D}_{p+1}^*)=\widehat{\sigma}^*, \tag{153}$$

где $\widehat{\sigma}^*$ — сигнатура формы $T^*(x, x)=\sum\limits_{i, k=0}^{p} t_{ik}^* x_i x_k$.

Наряду с формой $T^*(x, x)$ рассмотрим формы

$$T(x, x)=\sum_{i, k=0}^{p} t_{i+k}x_i x_k \text{ и } T^{**}(x, x)=t_p(x_0x_p+x_1x_{p-1}+\dots+x_px_0).$$

Матрица T^{**} получается из матрицы T [см. (152)], если в последней заменить нулями все элементы, стоящие под второй диагональю. Сигнатуры форм $T(x, x)$ и $T^{**}(x, x)$ обозначим соответственно через $\widehat{\sigma}$ и $\widehat{\sigma}^{**}$.

Так как формы $T^*(x, x)$ и $T^{**}(x, x)$ получаются из формы $T(x, x)$ таким варьированием коэффициентов, в процессе которого ранг формы не меняется ($|T^{**}| = |T| = \frac{D_{h+p+1}}{D_h} \neq 0$, $|T^*| = \frac{D^*_{h+p+1}}{D^*_h} \neq 0$), то и сигнатуры форм $T(x, x)$, $T^*(x, x)$ и $T^{**}(x, x)$ должны быть одинаковы:

$$\hat{\sigma} = \hat{\sigma}^* = \hat{\sigma}^{**}. \tag{154}$$

Но

$$T^{**}(x, x) \begin{cases} 2t_p(x_0x_{2k-1} + \ldots + x_{k-1}x_k) & \text{при } p = 2k-1, \\ t_p[2(x_0x_{2k} + \ldots + x_{k-1}x_{k+1}) + x_k^2] & \text{при } p = 2k. \end{cases}$$

Так как каждое произведение вида $x_\alpha x_\beta$ при $\alpha \neq \beta$ можно заменить разностью квадратов $\left(\frac{x_\alpha + x_\beta}{2}\right)^2 - \left(\frac{x_\alpha - x_\beta}{2}\right)^2$ и таким образом получить разложение $T^{**}(x, x)$ независимые вещественные квадраты, то

$$\hat{\sigma}^{**} = \begin{cases} 0 & \text{при } p \text{ нечетном}, \\ \operatorname{sign} t_p & \text{при } p \text{ четном}. \end{cases} \tag{155}$$

С другой стороны, из (152)

$$\frac{D_{h+p+1}}{D_h} = |T| = (-1)^{\frac{p(p+1)}{2}} t_p^{p+1}. \tag{156}$$

Из (153), (154), (155) и (156) следует:

$$P(D^*_h, D^*_{h+1}, \ldots, D^*_{h+p+1}) - V(D^*_h, D^*_{h+1}, \ldots, D^*_{h+p+1}) = \begin{cases} 0 \text{ при } p \text{ нечетном}, \\ \varepsilon \text{ при } p \text{ четном}, \end{cases} \tag{157}$$

где

$$\varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \frac{D_{h+p+1}}{D_h}.$$

Так как

$$P(D^*_{h+1}, D^*_{h+2}, \ldots, D^*_{h+p+1}) + V(D^*_{h+1}, D^*_{h+2}, \ldots, D^*_{h+p+1}) = p+1, \tag{158}$$

то из (157) и (158) вытекает таблица (150).

Пусть теперь $D_r = 0$. Тогда при некотором $h < r$

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_r = 0.$$

В этом случае согласно теореме 23

$$D^{(r)} = S\begin{pmatrix} 1 & \ldots & h & n-r+h+1 & \ldots & n \\ 1 & \ldots & h & n-r+h+1 & \ldots & n \end{pmatrix} \neq 0.$$

Рассматриваемый случай сводится к предыдущему перенумерацией переменных в квадратичной форме $S(x, x) = \sum_{i,k=0}^{n-1} s_{i+k}x_ix_k$. Полагаем:

$$\begin{gathered} \tilde{x}_0 = x_0, \ \ldots, \ \tilde{x}_{h-1} = x_{h-1}, \ \tilde{x}_h = x_{n-r+h}, \ \ldots, \ \tilde{x}_{r-1} = x_{n-1}, \\ \tilde{x}_r = x_h, \ \ldots, \ \tilde{x}_{n-1} = x_{n-r+h-1}. \end{gathered} \tag{159}$$

При этом $S(x, x) = \sum_{i,k=0}^{n-1} \tilde{s}_{i+k}x_ix_k$.

Исходя из структуры матрицы T на стр. 305 и пользуясь полученными из детерминантного тождества Сильвестра соотношениями

$$\widehat{D}_j = \frac{D_{h+j}}{D_h}, \quad \widehat{\widetilde{D}}_j = \frac{\widetilde{D}_{h+j}}{D_h} \quad (j = 1, 2, \ldots, n-h),$$

найдем, что ряд 1, $\widetilde{D}_1$, $\widetilde{D}_2$, ..., $\widetilde{D}_n$ получается из ряда 1, D_1, D_2, ..., D_n заменой одного элемента D_r на $D^{(r)}$.

Таким образом, показано, что во всех случаях можно пользоваться таблицей (150).

Заметим, что при p нечетном [p — число нулевых определителей в группе (148)] из формулы (156) следует:

$$\operatorname{sign} \frac{D_{h+p+1}}{D_h} = (-1)^{\frac{p+1}{2}}. \tag{160}$$

Пользуясь этим равенством, читатель легко проверит, что таблице (150) соответствует то приписывание знаков нулевым определителям, которое дается формулой (149).

Теорема доказана полностью [1]).

Примечание. При $p=1$ из формулы (160) следует: $D_h D_{h+2} < 0$. Поэтому имеет место правило Гундельфингера, т. е. при подсчете $v(1, D_1, \ldots, D_r)$ можно D_{h+1} опустить. При $p=2$ из таблицы (150) вытекает правило Фробениуса (см. стр. 275).

---

[1] Нетрудно убедиться, что теорема 23, а с ней и теорема 24, сохраняют силу также при $h=0$, если считать, как мы условились на стр. 305, $D_0 \equiv 1$ (см. Фробениус, [182 f]).

ЧАСТЬ ВТОРАЯ

СПЕЦИАЛЬНЫЕ ВОПРОСЫ И ПРИЛОЖЕНИЯ

ГЛАВА XI

КОМПЛЕКСНЫЕ СИММЕТРИЧЕСКИЕ, КОСОСИММЕТРИЧЕСКИЕ И ОРТОГОНАЛЬНЫЕ МАТРИЦЫ

В главе IX в связи с изучением линейных операторов в евклидовом пространстве были исследованы вещественные симметрические, кососимметрические и ортогональные матрицы, т. е. вещественные квадратные матрицы, характеризуемые соответственно соотношениями

$$S' = S, \quad K' = -K, \quad O' = O^{-1}$$

(здесь $'$ означает переход к транспонированной матрице). Было выяснено, что в поле комплексных чисел все эти матрицы имеют линейные элементарные делители, и были установлены нормальные формы для этих матриц, т. е. «простейшие» вещественные симметрические, кососимметрические и ортогональные матрицы, которым вещественно- и ортогонально-подобны произвольные матрицы рассматриваемых типов.

Настоящая глава посвящена исследованию комплексных симметрических, кососимметрических и ортогональных матриц. Выясняется, какие элементарные делители могут иметь эти матрицы, и устанавливаются для них нормальные формы. Эти формы имеют значительно более сложную структуру, нежели соответствующие нормальные формы в вещественном случае. Предварительно в первом параграфе устанавливаются интересные связи между комплексными ортогональными, унитарными и вещественными симметрическими, кососимметрическими и ортогональными матрицами.

§ 1. Некоторые формулы для комплексных ортогональных и унитарных матриц

Начнем с леммы.

Лемма 1 [1]. 1. *Если матрица G одновременно является и эрмитовой и ортогональной ($G' = \overline{G} = G^{-1}$), то она представима в виде*

$$G = Ie^{iK}, \tag{1}$$

где I — вещественная симметрическая инволютивная матрица, а K — перестановочная с нею вещественная кососимметрическая матрица:

$$I = \overline{I} = I', \quad I^2 = E, \quad K = \overline{K} = -K'. \tag{2}$$

2. *Если дополнительно G является положительно определенной эрмитовой матрицей* [2], *то в формуле (1) $I = E$ и*

$$G = e^{iK}. \tag{3}$$

---

[1] [81в], стр. 223—225.

[2] То есть G — матрица коэффициентов положительно определенной эрмитовой формы (см. гл. X, § 9).

Доказательство. 1. Пусть

$$G = S + iT, \tag{4}$$

где S и T — вещественные матрицы. Тогда

$$\overline{G} = S - iT \quad \text{и} \quad G' = S' + iT'. \tag{5}$$

Поэтому равенство $\overline{G} = G'$ влечет: $S = S'$, $T = -T'$, т. е. S — симметрическая, а T — кососимметрическая матрица.

Далее, комплексное равенство $G\overline{G} = E$ после подстановки в него выражений для G и $\overline{G}$ из (4) и (5) распадается на два вещественных равенства:

$$S^2 + T^2 = E, \; ST = TS. \tag{6}$$

Второе из этих равенств показывает, что S и T коммутируют.

Согласно теореме 12′ главы IX (стр. 265) коммутирующие нормальные матрицы S и T можно одним и тем же вещественным ортогональным преобразованием привести к квазидиагональной канонической форме. Поэтому [1])

$$S = O\{s_1, s_1, s_2, s_2, \ldots, s_q, s_q, s_{2q+1}, \ldots, s_n\}O^{-1} \qquad (O = \overline{O} = O'^{-1}) \tag{7}$$

$$T = O\left\{\begin{Vmatrix} 0 & t_1 \\ -t_1 & 0 \end{Vmatrix}, \begin{Vmatrix} 0 & t_2 \\ -t_2 & 0 \end{Vmatrix}, \ldots, \begin{Vmatrix} 0 & t_q \\ -t_q & 0 \end{Vmatrix}, 0, \ldots, 0\right\}O^{-1}$$

(числа s_i и t_i вещественны). Отсюда

$$G = S + iT = O\left\{\begin{Vmatrix} s_1 & it_1 \\ -it_1 & s_1 \end{Vmatrix}, \begin{Vmatrix} s_2 & it_2 \\ -it_2 & s_2 \end{Vmatrix}, \ldots, \begin{Vmatrix} s_q & it_q \\ -it_q & s_q \end{Vmatrix}, s_{2q+1}, \ldots, s_n\right\}O^{-1}. \tag{8}$$

С другой стороны, подставляя выражения (7) для S и T в первое из равенств (6), найдем:

$$s_1^2 - t_1^2 = 1, \; s_2^2 - t_2^2 = 1, \; \ldots, \; s_q^2 - t_q^2 = 1, \; s_{2q+1} = \pm 1, \; \ldots, \; s_n = \pm 1. \tag{9}$$

Теперь нетрудно проверить, что матрица типа $\begin{Vmatrix} s & it \\ -it & s \end{Vmatrix}$ при $s^2 - t^2 = 1$ всегда представима в виде

$$\begin{Vmatrix} s & it \\ -it & s \end{Vmatrix} = \varepsilon e^{i\begin{Vmatrix} 0 & \varphi \\ -\varphi & 0 \end{Vmatrix}},$$

где

$$|s| = \operatorname{ch}\varphi, \quad \varepsilon t = \operatorname{sh}\varphi, \quad \varepsilon = \operatorname{sign} s.$$

Поэтому в силу (8) и (9) имеем:

$$G = O\{\pm e^{i\begin{Vmatrix} 0 & \varphi_1 \\ -\varphi_1 & 0 \end{Vmatrix}}, \pm e^{i\begin{Vmatrix} 0 & \varphi_2 \\ -\varphi_2 & 0 \end{Vmatrix}}, \ldots, \pm e^{i\begin{Vmatrix} 0 & \varphi_q \\ -\varphi_q & 0 \end{Vmatrix}}, \pm 1, \ldots, \pm 1\}O^{-1}, \tag{10}$$

т. е.

$$G = Ie^{iK},$$

---

[1]) См. также примечание к теореме 12′ гл. IX (стр. 265).

где

$$\left.\begin{aligned} I &= O\{\pm 1, \pm 1, \ldots, \pm 1\}O^{-1}, \\ K &= O\left\{\begin{Vmatrix} 0 & \varphi_1 \\ -\varphi_1 & 0 \end{Vmatrix}, \ldots, \begin{Vmatrix} 0 & \varphi_q \\ -\varphi_q & 0 \end{Vmatrix}, 0, \ldots, 0\right\}O^{-1} \end{aligned}\right\} \qquad (11)$$

и

$$IK = KI.$$

Из (11) вытекают равенства (2).

2. Если дополнительно известно, что G — положительно определенная эрмитова матрица, то можно утверждать, что все характеристические числа матрицы G положительны (гл. IX, стр. 248). Но в силу формулы (10) этими характеристическими числами являются числа

$$\pm e^{\varphi_1}, \ \pm e^{-\varphi_1}, \ \pm e^{\varphi_2}, \ \pm e^{-\varphi_2}, \ \ldots, \ \pm e^{\varphi_q}, \ \pm e^{-\varphi_q}, \ \pm 1, \ \ldots, \ \pm 1$$

[здесь знаки соответствуют знакам в формуле (10)].

Поэтому в формуле (10) и в последующей формуле (11) всюду, где стоит $\pm$, сохраняется знак $+$. Следовательно,

$$I = O\{1, 1, \ldots, 1\}O^{-1} = E,$$

что и требовалось доказать.

Лемма доказана полностью.

С помощью леммы мы докажем следующую теорему:

Теорема 1. *Комплексная ортогональная матрица O всегда представима в виде*

$$O = Re^{iK}, \qquad (12)$$

где R — вещественная ортогональная, а K — вещественная кососимметрическая матрица

$$R = \overline{R} = R'^{-1}, \quad K = \overline{K} = -K'. \qquad (13)$$

Доказательство. Допустим, что формула (12) имеет место. Тогда

$$O^* = \overline{O}' = e^{iK}R'$$

и

$$O^*O = e^{iK}R'Re^{iK} = e^{2iK}.$$

Теперь в силу предыдущей леммы искомую вещественную кососимметрическую матрицу K можно определить из равенства

$$O^*O = e^{2iK}, \qquad (14)$$

поскольку матрица O^*O — положительно определенная эрмитова и ортогональная матрица[1]). После того как матрица K определена из (14), мы находим R из (12):

$$R = Oe^{-iK}. \qquad (15)$$

Тогда

$$R^*R = e^{-iK}O^*Oe^{iK} = E,$$

т. е. R — унитарная матрица. С другой стороны, из (15) следует, что матрица R как произведение двух ортогональных матриц сама

[1]) Комплексная ортогональная матрица O является неособенной, так как из равенства $OO' = E$ следует, что $|O| = \pm 1$.

ортогональна: $R'R=E$. Таким образом, R является одновременно унитарной и ортогональной и, следовательно, вещественной ортогональной. Формулу (15) можно записать в виде (12).

Теорема доказана [1]).

Установим теперь следующую лемму:

Лемма 2. *Если матрица D является одновременно симметрической и унитарной ($D=D'=\overline{D}^{-1}$), то она всегда представима в виде*

$$D=e^{iS}, \tag{16}$$

где S — вещественная симметрическая матрица ($S=\overline{S}=S'$).

Доказательство. Положим

$$D=U+iV \quad (U=\overline{U},\ V=\overline{V}). \tag{17}$$

Тогда

$$\overline{D}=U-iV, \quad D'=U'+iV'.$$

Комплексное равенство $D=D'$ распадается на два вещественных:

$$U=U', \quad V=V'.$$

Таким образом, U и V — вещественные симметрические матрицы.

Равенство $D\overline{D}=E$ влечет:

$$U^2+V^2=E, \quad UV=VU. \tag{18}$$

Согласно второму из этих равенств матрицы U и V коммутируют. Применяя к ним теорему 12′ (вместе с примечанием) главы IX (стр. 265), получим:

$$U=O\{s_1, s_2, \ldots, s_n\}O^{-1}, \quad V=O\{t_1, t_2, \ldots, t_n\}O^{-1}. \tag{19}$$

Здесь $O=\overline{O}=O'^{-1}$, а s_k и t_k ($k=1, 2, \ldots, n$) — вещественные числа. Теперь первое из равенств (18) дает:

$$s_k^2+t_k^2=1 \qquad (k=1, 2, \ldots, n).$$

Поэтому существуют такие вещественные числа φ_k ($k=1, 2, \ldots, n$), что

$$s_k=\cos\varphi_k, \quad t_k=\sin\varphi_k \quad (k=1, 2, \ldots, n).$$

Подставляя эти выражения для s_k и t_k в (19) и пользуясь (17), найдем:

$$D=O\{e^{i\varphi_1}, e^{i\varphi_2}, \ldots, e^{i\varphi_n}\}O^{-1}=e^{iS},$$

где

$$S=O\{\varphi_1, \varphi_2, \ldots, \varphi_n\}O^{-1}. \tag{20}$$

Из (20) следует: $S=\overline{S}=S'$.

Лемма доказана.

Пользуясь этой леммой, докажем следующую теорему:

Теорема 2. *Унитарная матрица U всегда представима в виде*

$$U=Re^{iS}, \tag{21}$$

где R — вещественная ортогональная, а S — вещественная симметрическая матрица:

$$R=\overline{R}=R'^{-1}, \quad S=\overline{S}=S'. \tag{22}$$

[1]) Формула (12), как и полярное разложение комплексной матрицы [в соответствии с формулами (87), (88) на стр. 249], имеет тесную связь с важной теоремой Картана, устанавливающей известные представления для автоморфизмов полупростых комплексных групп Ли [81в], стр. 232—233.

Доказательство. Из формулы (21) следует:

$$U' = e^{iS}R'. \tag{23}$$

Перемножая почленно (21) и (23), получим в силу (22):

$$U'U = e^{iS}R'Re^{iS} = e^{2iS}.$$

Согласно лемме 2 вещественную симметрическую матрицу S можно определить из уравнения

$$U'U = e^{2iS}, \tag{24}$$

поскольку матрица $U'U$ является симметрической унитарной. После того как матрица S определена, мы определим матрицу R равенством

$$R = Ue^{-iS}. \tag{25}$$

Тогда

$$R' = e^{-iS}U', \tag{26}$$

и потому из (24), (25) и (26) вытекает

$$R'R = e^{-iS}U'Ue^{-iS} = E,$$

т. е. R — ортогональная матрица.

С другой стороны, согласно (25) R есть произведение двух унитарных матриц и, следовательно, R — унитарная матрица. Поскольку R одновременно является ортогональной и унитарной, R — вещественная матрица. Формулу (25) можно переписать в виде (21).

Теорема доказана.

§ 2. Полярное разложение комплексной матрицы

Докажем следующую теорему:

Теорема 3. *Если* $A = \| a_{ik} \|_1^n$ — *неособенная матрица с комплексными элементами, то*

$$A = SO \tag{27}$$

и

$$A = O_1S_1, \tag{28}$$

где S *и* S_1 — *комплексные симметрические, а* O *и* O_1 — *комплексные ортогональные матрицы. При этом*

$$S = \sqrt{AA'} = f(AA'), \qquad S_1 = \sqrt{A'A} = f_1(A'A),$$

где $f(\lambda)$, $f_1(\lambda)$ — *некоторые многочлены относительно* λ.

Как в разложении (27), *так и в разложении* (28) *сомножители* S *и* O (*соответственно* O_1 *и* S_1) *перестановочны между собой в том и только в тот случае, когда матрицы* A *и* A' *перестановочны между собой.*

Доказательство. Достаточно установить разложение (27), так как, применив это разложение к матрице A' и определив из полученной формулы матрицу A, мы придем к разложению (28).

Если имеет место формула (27), то

$$A = SO, \qquad A' = O^{-1}S$$

и потому

$$AA' = S^2. \tag{29}$$

Обратно, поскольку AA' — неособенная матрица ($|AA'| = |A|^2 \neq 0$), то функция $\sqrt{\lambda}$ определена на спектре этой матрицы [1]), и, следовательно, существует такой интерполяционный многочлен $f(\lambda)$, что

$$\sqrt{AA'} = f(AA'). \tag{30}$$

Симметрическую матрицу (30) обозначим через

$$S = \sqrt{AA'}.$$

Тогда имеет место (29) и, следовательно, $|S| \neq 0$. Определяя матрицу O из равенства (27)

$$O = S^{-1}A,$$

легко проверяем, что эта матрица является ортогональной. Таким образом разложение (27) установлено.

Если в разложении (27) множители S и O перестановочны между собой, то перестановочны и матрицы

$$A = SO \text{ и } A' = O^{-1}S,$$

так как

$$AA' = S^2, \quad A'A = O^{-1}S^2O.$$

Обратно, если $AA' = A'A$, то

$$S^2 = O^{-1}S^2O,$$

т. е. матрица O перестановочна с $S^2 = AA'$. Но тогда матрица O перестановочна и с матрицей $S = f(AA')$.

Таким образом теорема доказана полностью.

Пользуясь полярным разложением, докажем теорему:

Теорема 4. *Если две комплексные симметрические, либо кососимметрические, либо ортогональные матрицы подобны*

$$B = T^{-1}AT, \tag{31}$$

то эти матрицы ортогонально-подобны, т. е. существует такая ортогональная матрица O, что

$$B = O^{-1}AO. \tag{32}$$

Доказательство. Из условия теоремы следует существование такого многочлена $q(\lambda)$, что

$$A' = q(A), \quad B' = q(B). \tag{33}$$

Этот многочлен $q(\lambda)$ в случае симметрических матриц тождественно равен λ, а в случае кососимметрических матриц тождественно равен $-\lambda$. Если же A и B — ортогональные матрицы, то $q(\lambda)$ — интерполяционный многочлен для $\frac{1}{\lambda}$ на общем спектре матриц A и B.

Пользуясь равенствами (33), мы проведем доказательство данной теоремы совершенно аналогично доказательству соответствующей теоремы 10 главы IX для вещественного случая (см. стр. 262). Из (31) следует:

$$q(B) = T^{-1}q(A)T$$

---

[1]) См. гл. V, § 1. Мы берем однозначную ветвь функции $\sqrt{\lambda}$ в односвязной области, содержащей все характеристические числа матрицы AA' и не содержащей числа 0.

или в силу (33)

$$B' = T^{-1}A'T.$$

Отсюда

$$B = T'AT'^{-1}.$$

Сопоставляя это равенство с (31), легко находим:

$$TT'A = ATT'. \tag{34}$$

Применим к неособенной матрице T полярное разложение

$$T = SO \quad (S = S' = f(TT'), \quad O' = O^{-1}).$$

Поскольку согласно (34) матрица TT' перестановочна с A, то и матрица $S = f(TT')$ также перестановочна с A. Поэтому, подставляя в (31) вместо T произведение SO, будем иметь:

$$B = O^{-1}S^{-1}ASO = O^{-1}AO.$$

Теорема доказана.

§ 3. Нормальная форма комплексной симметрической матрицы

Докажем следующую теорему.

Теорема 5. *Существует комплексная симметрическая матрица с любыми наперед заданными элементарными делителями* [1]).

Доказательство. Рассмотрим матрицу H n-го порядка, у которой элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Докажем, что существует симметрическая матрица S, подобная матрице H:

$$S = THT^{-1}. \tag{35}$$

Преобразующую матрицу T будем искать, исходя из условия:

$$S = THT^{-1} = S' = T'^{-1}H'T'.$$

Это условие можно переписать так:

$$VH = H'V, \tag{36}$$

где V—симметрическая матрица, связанная с T равенством

$$T'T = -2iV \text{ [2])}. \tag{37}$$

Вспоминая свойства матриц H и $F = H'$ (стр. 25—26), мы найдем, что любое решение V матричного уравнения (36) имеет следующий вид:

$$V = \begin{Vmatrix} 0 & \cdots & 0 & a_0 \\ \vdots & & a_0 & a_1 \\ \vdots & \iddots & \iddots & \\ 0 & \iddots & & \vdots \\ a_0 & a_1 & \cdots & a_{n-1} \end{Vmatrix}, \tag{38}$$

где $a_0, a_1, \ldots, a_{n-1}$ — произвольные комплексные числа.

---

[1]) Относительно содержания настоящего параграфа, а также последующих §§ 4 и 5 см. [256].

[2]) Для упрощения дальнейших формул нам удобно здесь ввести множитель $-2i$.

Поскольку нам достаточно отыскать одну преобразующую матрицу T, то мы в этой формуле положим $a_0=1$, $a_1=\ldots=a_{n-1}=0$ и определим матрицу V равенством [1])

$$V=\begin{Vmatrix} 0 & \ldots & 0 & 1 \\ 0 & \ldots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \ldots & 0 & 0 \end{Vmatrix}. \tag{39}$$

Кроме того, преобразующую матрицу T будем искать в виде симметрической матрицы:

$$T=T'. \tag{40}$$

Тогда уравнение (37) для T перепишется так:

$$T^2=-2iV. \tag{41}$$

Теперь неизвестную матрицу T будем искать в виде многочлена от V. Поскольку $V^2=E$, в качестве такого многочлена можно взять многочлен первой степени: $T=\alpha E+\beta V$. Из уравнения (41), учитывая равенство $V^2=E$, найдем: $\alpha^2+\beta^2=0$, $2\alpha\beta=-2i$. Этим соотношениям мы удовлетворим, полагая $\alpha=1$, $\beta=-i$. Тогда

$$T=E-iV. \tag{42}$$

T — неособенная симметрическая матрица [2]). В то же время из (41): $T^{-1}=\frac{1}{2}iV^{-1}T=\frac{1}{2}iVT$, т. е.

$$T^{-1}=\frac{1}{2}(E+iV). \tag{43}$$

Таким образом, симметрическая форма S матрицы H определится равенством

$$S=THT^{-1}=\frac{1}{2}(E-iV)H(E+iV),\quad V=\begin{Vmatrix} 0 & \ldots & 0 & 1 \\ 0 & \ldots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \ldots & 0 & 0 \end{Vmatrix}. \tag{44}$$

Поскольку матрица S удовлетворяет уравнению (36) и $V^2=E$, то равенство (44) может быть переписано еще так:

$$2S=(H+H')+i\,(HV-VH)=H+H'+i\,(H-H')V=$$

$$=\begin{Vmatrix} 0 & 1 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ 0 & \ldots & 1 & 0 \end{Vmatrix}+i\begin{Vmatrix} 0 & \ldots & 1 & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ 0 & -1 & \ldots & 0 \end{Vmatrix}. \tag{45}$$

Формула (45) определяет симметрическую форму S матрицы H.

[1]) Матрица V является одновременно симметрической и ортогональной.

[2]) Неособенность матрицы T следует, в частности, из (41), поскольку V — неособенная матрица.

В дальнейшем, если n — порядок матрицы H, $H=H^{(n)}$, то соответствующие матрицы T, V и S будем еще обозначать и так: $T^{(n)}$, $V^{(n)}$ и $S^{(n)}$.

Пусть даны произвольные элементарные делители:

$$(\lambda-\lambda_1)^{p_1}, \quad (\lambda-\lambda_2)^{p_2}, \ \ldots, \ (\lambda-\lambda_u)^{p_u}. \tag{46}$$

Составим соответствующую жорданову матрицу

$$J=\{\lambda_1 E^{(p_1)}+H^{(p_1)}, \quad \lambda_2 E^{(p_2)}+H^{(p_2)}, \ \ldots, \ \lambda_u E^{(p_u)}+H^{(p_u)}\}.$$

Для каждой матрицы $H^{(p_j)}$ введем соответствующую симметрическую форму $S^{(p_j)}$. Из

$$S^{(p_j)}=T^{(p_j)}H^{(p_j)}[T^{(p_j)}]^{-1} \quad (j=1, \ 2, \ \ldots, \ u)$$

следует:

$$\lambda_j E^{(p_j)}+S^{(p_j)}=T^{(p_j)}[\lambda_j E^{(p_j)}+H^{(p_j)}][T^{(p_j)}]^{-1}.$$

Поэтому, полагая

$$\widetilde{S}=\{\lambda_1 E^{(p_1)}+S^{(p_1)}, \ \lambda_2 E^{(p_2)}+S^{(p_2)}, \ \ldots, \ \lambda_u E^{(p_u)}+S^{(p_u)}\}, \tag{47}$$

$$T=\{T^{(p_1)}, \ T^{(p_2)}, \ \ldots, \ T^{(p_u)}\}, \tag{48}$$

будем иметь:

$$\widetilde{S}=TJT^{-1}.$$

$\widetilde{S}$ — симметрическая форма жордановой матрицы J. Матрица $\widetilde{S}$ подобна матрице J и имеет те же элементарные делители (46), что и матрица J. Теорема доказана.

Следствие 1. *Произвольная квадратная комплексная матрица* $A=\|a_{ik}\|_1^n$ *подобна симметрической матрице.*

Привлекая теорему 4, получим:

Следствие 2. *Произвольная комплексная симметрическая матрица* $S=\|s_{ik}\|_1^n$ *ортогонально-подобна симметрической матрице, имеющей нормальную форму* $\widetilde{S}$, *т. е. существует такая ортогональная матрица* O, *что*

$$S=O\widetilde{S}O^{-1}. \tag{49}$$

Нормальная форма комплексной симметрической матрицы имеет квазидиагональный вид

$$\widetilde{S}=\{\lambda_1 E^{(p_1)}+S^{(p_1)}, \ \lambda_2 E^{(p_2)}+S^{(p_2)}, \ \ldots, \ \lambda_u E^{(p_u)}+S^{(p_u)}\}, \tag{50}$$

где клетки $S^{(p)}$ определяются так [см. (44), (45)]:

$$2S^{(p)}=[E^{(p)}-iV^{(p)}]H^{(p)}[E^{(p)}+iV^{(p)}]=$$

$$=[H^{(p)}+H^{(p)'}+i(H^{(p)}-H^{(p)'})V^{(p)}]=$$

$$=\begin{Vmatrix} 0 & 1 & \ldots & 0 \\ \cdot & \cdot & & \\ 1 & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ 0 & \ldots & 1 & 0 \end{Vmatrix} + i \begin{Vmatrix} 0 & \ldots & 1 & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & -1 & \ldots & 0 \end{Vmatrix}. \tag{51}$$

§ 4. Нормальная форма комплексной кососимметрической матрицы

Выясним, какие ограничения на элементарные делители накладывает косая симметрия матрицы. При этом мы будем опираться на следующую теорему:

Теорема 6. *Кососимметрическая матрица всегда имеет четный ранг.*

Доказательство. Пусть кососимметрическая матрица K имеет ранг r. Тогда среди строк матрицы K имеется r линейно независимых с номерами $i_1, i_2, \ldots, i_r$; все остальные строки являются линейными комбинациями этих строк. Поскольку столбцы матрицы K получаются из соответствующих строк, если элементы последних помножить на -1, то и любой столбец матрицы K есть линейная комбинация столбцов с номерами $i_1, i_2, \ldots, i_r$. Поэтому произвольный минор r-го порядка матрицы K может быть представлен в виде

$$\alpha K \begin{pmatrix} i_1 & i_2 \ldots i_r \\ i_1 & i_2 \ldots i_r \end{pmatrix},$$

где α — число.

Отсюда вытекает, что

$$K \begin{pmatrix} i_1 & i_2 \ldots i_r \\ i_1 & i_2 \ldots i_r \end{pmatrix} \neq 0.$$

Но кососимметрический определитель нечетного порядка всегда равен нулю. Следовательно, r — четное число.

Теорема доказана.

Теорема 7. 1° *Если λ_0 — характеристическое число кососимметрической матрицы K и ему соответствуют элементарные делители*

$$(\lambda-\lambda_0)^{l_1}, \quad (\lambda-\lambda_0)^{l_2}, \ldots, (\lambda-\lambda_0)^{l_t},$$

то $-\lambda_0$ также является характеристическим числом матрицы K и этому числу соответствуют элементарные делители матрицы K в том же числе и тех же степеней

$$(\lambda+\lambda_0)^{l_1}, \quad (\lambda+\lambda_0)^{l_2}, \ldots, (\lambda+\lambda_0)^{l_t}.$$

2° *Если число нуль является характеристическим числом кососимметрической матрицы K*[1]*, то в системе элементарных делителей матрицы K все элементарные делители четной степени, соответствующие характеристическому числу нуль, повторяются четное число раз.*

Доказательство. 1° Транспонированная матрица K' имеет те же элементарные делители, что и матрица K. Но $K' = -K$, а элементарные делители матрицы $-K$ получаются из элементарных делителей матрицы K, если в последних все характеристические числа $\lambda_1, \lambda_2, \ldots$ заменить на $-\lambda_1, -\lambda_2, \ldots$ Отсюда следует первая часть нашей теоремы.

2° Пусть характеристическому числу нуль матрицы K отвечает δ_1 элементарных делителей вида λ, δ_2 — вида λ^2 и т. д. Вообще мы через δ_p обозначим число элементарных делителей вида λ^p ($p=1, 2, \ldots$). Мы докажем, что $\delta_2, \delta_4, \ldots$ — четные числа.

---

[1]) То есть если $|K|=0$. При n нечетном всегда $|K|=0$.

Дефект d матрицы K равен числу линейно независимых собственных векторов, соответствующих характеристическому числу нуль или, что то же, числу элементарных делителей вида λ, λ^2, λ^3, ... Поэтому

$$d = \delta_1 + \delta_2 + \delta_3 + \ldots \tag{52}$$

Поскольку согласно теореме 6 ранг матрицы K — четное число, а $d = n - r$, то число d имеет ту же четность, что и число n. Такое же утверждение можно сделать относительно дефектов d_3, d_5, ... матриц K^3, K^5, ..., поскольку нечетные степени кососимметрической матрицы снова являются кососимметрическими матрицами. Поэтому все числа $d_1 = d$, d_3, d_5, ... имеют одну и ту же четность.

С другой стороны, при возведении матрицы K в степень m каждый элементарный делитель λ^p этой матрицы при $p < m$ расщепляется на p элементарных делителей (первой степени), а при $p \geqslant m$ — на m элементарных делителей [1]). Поэтому число элементарных делителей матриц K, K^3, ..., являющихся степенями λ, определится по формулам [2])

$$\begin{aligned} d_3 &= \delta_1 + 2\delta_2 + 3(\delta_3 + \delta_4 + \ldots), \\ d_5 &= \delta_1 + 2\delta_2 + 3\delta_3 + 4\delta_4 + 5(\delta_5 + \delta_6 + \ldots), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{53}$$

Сопоставляя (52) с (53) и помня, что все числа $d_1 = d$, d_3, d_5, ... имеют одну и ту же четность, легко заключаем, что δ_2, δ_4, ... — четные числа.

Теорема доказана полностью.

Теорема 8. *Существует кососимметрическая матрица с любыми наперед заданными элементарными делителями, удовлетворяющими ограничениям* 1°, 2° *предыдущей теоремы.*

Доказательство. Найдем сначала кососимметрическую форму для квазидиагональной матрицы порядка $2p$:

$$J_{\lambda_0}^{(pp)} = \{\lambda_0 E + H,\ -\lambda_0 E - H\}, \tag{54}$$

имеющей два элементарных делителя $(\lambda - \lambda_0)^p$ и $(\lambda + \lambda_0)^p$; здесь $E = E^{(p)}$, $H = H^{(p)}$.

Будем искать такую преобразующую матрицу T, чтобы матрица

$$TJ_{\lambda_0}^{(pp)}T^{-1}$$

была кососимметрической, т. е. чтобы имело место равенство

$$TJ_{\lambda_0}^{(pp)}T^{-1} + T'^{-1}[J_{\lambda_0}^{(pp)}]'\,T' = 0$$

или

$$WJ_{\lambda_0}^{(pp)} + [J_{\lambda_0}^{(pp)}]'\,W = 0, \tag{55}$$

где W — симметрическая матрица, связанная с матрицей T равенством [3])

$$T'T = -2iW. \tag{56}$$

---

[1]) См. гл. VI, теорема 9 стр. 158.

[2]) Эти формулы были выведены (без ссылки на теорему 9) в гл. VI [см. формулы (49) на стр. 156].

[3]) См. сноску [2]) к стр. 319.

Разобьем матрицу W на четыре квадратных блока, каждый порядка p:

$$W=\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}.$$

Тогда (55) можно переписать так:

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}\begin{pmatrix} \lambda_0 E+H & 0 \\ 0 & -\lambda_0 E-H \end{pmatrix}+ \\ +\begin{pmatrix} \lambda_0 E+H' & 0 \\ 0 & -\lambda_0 E-H' \end{pmatrix}\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}=0. \quad (57)$$

Выполняя указанные действия над блочными матрицами в левой части матричного уравнения (57), мы заменим это уравнение системой четырех матричных уравнений:

$$\left.\begin{array}{ll} 1)\ H'W_{11}+W_{11}(2\lambda_0 E+H)=0, & 2)\ H'W_{12}-W_{12}H=0, \\ 3)\ H'W_{21}-W_{21}H=0, & 4)\ H'W_{22}+W_{22}(2\lambda_0 E+H)=0. \end{array}\right\} \quad (58)$$

Уравнение $AX-XB=0$, где A и B — квадратные матрицы без общих характеристических чисел, имеет только нулевое решение $X=0$ [1]). Поэтому первое и четвертое уравнения (58) дают: $W_{11}=W_{22}=0$ [2]). Что же касается второго из этих уравнений, то, как мы видели при доказательстве теоремы 5, этому уравнению можно удовлетворить, полагая

$$W_{12}=V=\begin{Vmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 1 & \dots & 0 & 0 \end{Vmatrix}, \quad (59)$$

поскольку [см. (36)]

$$VH-H'V=0.$$

Из симметрии матрицы W следует, что

$$W_{21}=W_{12}'=V.$$

Тогда автоматически удовлетворяется и уравнение 3).

Таким образом,

$$W=\begin{pmatrix} 0 & V \\ V & 0 \end{pmatrix}=V^{(2p)}. \quad (60)$$

Но тогда, как уже было выяснено на стр. 320, уравнение (56) удовлетворится, если положить:

$$T=E^{(2p)}-iV^{(2p)}. \quad (61)$$

При этом

$$T^{-1}=\frac{1}{2}(E^{(2p)}+iV^{(2p)}). \quad (62)$$

---

[1]) См. гл. VIII, § 1.

[2]) При $\lambda_0\neq 0$ уравнения 1) и 4), кроме нулевых, других решений не имеют. При $\lambda_0=0$ существуют и другие решения, но мы выбираем нулевые решения.

Следовательно, искомая кососимметрическая матрица найдется по формуле

$$K_{\lambda_0}^{(pp)} = \frac{1}{2}[E^{(2p)} - iV^{(2p)}]\, J_{\lambda_0}^{(pp)}\, [E^{(2p)} + iV^{(2p)}] =$$

$$= \frac{1}{2}[J_{\lambda_0}^{(pp)} - J_{\lambda_0}^{(pp)'} + i(J_{\lambda_0}^{(pp)} V^{(2p)} - V^{(2p)} J_{\lambda_0}^{(pp)})]\ ^{1}). \qquad (63)$$

Подставляя вместо $J_{\lambda_0}^{(pp)}$ и $V^{(2p)}$ соответствующие блочные матрицы из (54) и (60), найдем:

$$2K_{\lambda_0}^{(pp)} = \begin{pmatrix} H-H' & 0 \\ 0 & H'-H \end{pmatrix} + i \begin{pmatrix} H+H'+2\lambda_0 E & 0 \\ 0 & -H-H'-2\lambda E \end{pmatrix} \begin{pmatrix} 0 & V \\ V & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} H-H' & i(2\lambda_0 V + HV + VH) \\ -i(2\lambda_0 V + HV + VH) & H'-H \end{pmatrix}, \qquad (64)$$

т. е.

$$K_{\lambda_0}^{(pp)} = \frac{1}{2} \left\| \begin{array}{cccc|cccc} 0 & 1 & \ldots & 0 & 0 & \ldots & i & 2\lambda_0 \\ -1 & 0 & \ddots & \vdots & \vdots & \iddots & 2\lambda_0 & i \\ \vdots & \ddots & \ddots & 1 & i & \iddots & \iddots & \vdots \\ 0 & \ldots & -1 & 0 & 2\lambda_0 & i & \ldots & 0 \\ \hline 0 & \ldots & -i & -2\lambda_0 & 0 & -1 & \ldots & 0 \\ \vdots & \iddots & -2\lambda_0 & -i & 1 & 0 & \ddots & \vdots \\ \vdots & \iddots & \iddots & \vdots & \vdots & \iddots & \ddots & -1 \\ -i & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ -2\lambda_0 & -i & \ldots & 0 & 0 & \ldots & 1 & 0 \end{array} \right\|. \qquad (65)$$

Построим теперь кососимметрическую матрицу q-го порядка $K^{(q)}$, имеющую один элементарный делитель λ^q, где q — нечетное число. Очевидно, что искомая кососимметрическая матрица будет подобна матрице

$$J^{(q)} = \left\| \begin{array}{ccccccc} 0 & 1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 1 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & -1 & 0 \\ \vdots & & & & 0 & 0 & -1 \\ 0 & \ldots & \ldots & \ldots & 0 & 0 & 0 \end{array} \right\|. \qquad (66)$$

---

1) Здесь мы используем равенства (55), (60) и равенство $[V^{(2p)}]^2 = E^{(2p)}$. Из этих равенств следует, что $V^{(2p)} J_{\lambda_0}^{(pp)} = -J_{\lambda_0}^{(pp)'} V^{(2p)}$ и $V^{(2p)} J_{\lambda_0}^{(pp)} V^{(2p)} = -J_{\lambda_0}^{(pp)'}$.

В этой матрице все элементы вне первой наддиагонали равны нулю, а вдоль первой наддиагонали сначала идут $\frac{q-1}{2}$ единиц, а затем $\frac{q-1}{2}$ элементов, равных -1. Полагая

$$K^{(q)} = TJ^{(q)}T^{-1}, \tag{67}$$

из условия косой симметрии найдем:

$$W_1 J^{(q)} + J^{(q)\prime} W_1 = 0, \tag{68}$$

где

$$T'T = -2iW_1. \tag{69}$$

Непосредственной проверкой убеждаемся в том, что матрица

$$W_1 = V^{(q)} = \begin{Vmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 1 & \dots & 0 & 0 \end{Vmatrix}$$

удовлетворяет уравнению (68). Принимая это значение для W_1, мы из (69), как и ранее, находим

$$T = E^{(q)} - iV^{(q)}, \quad T^{-1} = \frac{1}{2}[E^{(q)} + iV^{(q)}], \tag{70}$$

$$2K^{(q)} = [E^{(q)} - iV^{(q)}]\,(q)\,[E^{(q)} + iV^{(q)}]J = J^{(q)} - J^{(q)\prime} + i\,(J^{(q)} + J^{(q)\prime})\,V^{(q)}. \tag{71}$$

Произведя соответствующие вычисления, найдем:

$$2K^{(q)} = \begin{Vmatrix} 0 & 1 & \dots & \dots & 0 \\ -1 & 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \dots & \dots & 1 & 0 \end{Vmatrix} + i \begin{Vmatrix} 0 & \dots & \dots & 1 & 0 \\ \vdots & & \iddots & \iddots & 1 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ -1 & \iddots & \iddots & & \vdots \\ 0 & -1 & \dots & \dots & 0 \end{Vmatrix}. \tag{72}$$

Пусть даны произвольные элементарные делители, удовлетворяющие условиям теоремы 7:

$$\left.\begin{array}{l} (\lambda - \lambda_j)^{p_j},\ (\lambda + \lambda_j)^{p_j} \qquad (j = 1,\ 2,\ \dots,\ u), \\ \lambda^{q_k} \qquad (k = 1,\ 2,\ \dots,\ v;\ q_1,\ q_2,\ \dots,\ q_v - \text{нечетные числа}). \end{array}\right\}^{1)} \tag{73}$$

Тогда квазидиагональная кососимметрическая матрица

$$\widetilde{K} = \{K_{\lambda_1}^{(p_1 p_1)},\ \dots,\ K_{\lambda_u}^{(p_u p_u)};\ K^{(q_1)},\ \dots,\ K^{(q_v)}\} \tag{74}$$

или элементарные делители (73).

Теорема доказана.

Следствие. *Произвольная комплексная кососимметрическая матрица K ортогонально-подобна кососимметрической матрице, имеющей нормальную форму K, определяемую формулами* (74), (65), (72), т. е. *существует такая* (*комплексная*) *ортогональная матрица O, что*

$$K = O\widetilde{K}O^{-1}. \tag{75}$$

---

[1]) Среди чисел $\lambda_1, \lambda_2, \dots, \lambda_u$ могут быть и равные нулю. Кроме того, одно из чисел u и v может равняться нулю, т. е. в частном случае могут быть только элементарные делители одного типа.

Замечание. Если K — вещественная кососимметрическая матрица, то она имеет линейные элементарные делители (см. гл. IX, § 13)

$$\lambda - i\varphi_1,\ \lambda + i\varphi_1,\ \ldots,\ \lambda - i\varphi_u,\ \lambda + i\varphi_u,\ \underbrace{\lambda, \ldots, \lambda}_{v \text{ раз}}\quad (\varphi_j \text{ — вещественные числа}).$$

В этом случае, полагая в (74) все $p_j = 1$ и все $q_k = 1$, получим нормальную форму вещественной кососимметрической матрицы

$$\widetilde{K} = \left\{ \begin{Vmatrix} 0 & \varphi_1 \\ -\varphi_1 & 0 \end{Vmatrix}, \ldots, \begin{Vmatrix} 0 & \varphi_u \\ -\varphi_u & 0 \end{Vmatrix}, 0, \ldots, 0 \right\}.$$

§ 5. Нормальная форма комплексной ортогональной матрицы

Начнем с выяснения, какие ограничения на элементарные делители накладывает ортогональность матрицы.

Теорема 9. 1. *Если $\lambda_0\ (\lambda_0^2 \neq 1)$ — характеристическое число ортогональной матрицы O и этому характеристическому числу соответствуют элементарные делители*

$$(\lambda - \lambda_0)^{f_1},\ (\lambda - \lambda_0)^{f_2},\ \ldots,\ (\lambda - \lambda_0)^{f_t},$$

то $\dfrac{1}{\lambda_0}$ также является характеристическим числом матрицы O и этому характеристическому числу соответствуют такие же элементарные делители, как и числу λ_0:

$$(\lambda - \lambda_0^{-1})^{f_1},\ (\lambda - \lambda_0^{-1})^{f_2},\ \ldots,\ (\lambda - \lambda_0^{-1})^{f_t}.$$

2. *Если $\lambda_0 = \pm 1$ является характеристическим числом ортогональной матрицы O, то элементарные делители четной степени, соответствующие этому характеристическому числу λ_0, повторяются четное число раз.*

Доказательство. 1. Для любой неособенной матрицы O при переходе от O к O^{-1} каждый элементарный делитель $(\lambda - \lambda_0)^f$ заменяется элементарным делителем $(\lambda - \lambda_0^{-1})^f$ [1]). С другой стороны, матрицы O и O' всегда имеют одни и те же элементарные делители. Поэтому из условия ортогональности $O' = O^{-1}$ сразу следует первая часть нашей теоремы.

2. Допустим, что число 1 является характеристическим числом матрицы O, а число -1 не является таковым ($|E - O| = 0$, $|E + O| \neq 0$). Тогда воспользуемся формулами Кэли (см. гл. IX, § 14), которые сохраняют свою силу и для комплексных матриц. Определим матрицу K равенством

$$K = (E - O)(E + O)^{-1}. \tag{76}$$

Непосредственной проверкой убеждаемся в том, что $K' = -K$, т. е. что K — кососимметрическая матрица. Разрешая уравнение (76) относительно O, находим [2]):

$$O = (E - K)(E + K)^{-1}.$$

Полагая $f(\lambda) = \dfrac{1-\lambda}{1+\lambda}$, имеем $f'(\lambda) = -\dfrac{2}{(1+\lambda)^2} \neq 0$. Следовательно, при переходе от матрицы K к матрице $O = f(K)$ элементарные делители

---

[1]) См. гл. VI, § 7. Полагая $f(\lambda) = \dfrac{1}{\lambda}$, имеем $f'(\lambda) = -\dfrac{1}{\lambda^2} \neq 0$. Отсюда следует, что при переходе от матрицы O к матрице O^{-1} элементарные делители не расщепляются (см. стр. 158).

[2]) Заметим, что из (76) следует: $E + K = 2(E + O)^{-1}$ и, следовательно, $|E + K| = 2^n |E + O|^{-1} \neq 0$.

не расщепляются [1]). Поэтому в системе элементарных делителей матрицы O элементарные делители вида $(\lambda-1)^{2p}$ повторяются четное число раз, поскольку это имеет место для элементарных делителей вида λ^{2p} матрицы K (см. теорему 7).

Случай, когда ортогональная матрица O имеет характеристическое число -1, но не имеет характеристического числа $+1$, сразу сводится к разобранному случаю путем рассмотрения ортогональной матрицы $-O$.

Переходим к наиболее сложному случаю, когда матрица O одновременно имеет характеристическое число $+1$ и характеристическое число -1. Обозначим через $\psi(\lambda)$ минимальный многочлен матрицы O. Используя доказанную первую часть теоремы, мы сможем записать $\psi(\lambda)$ в виде

$$\psi(\lambda)=(\lambda-1)^{m_1}(\lambda+1)^{m_2}\prod_{j=1}^{u}(\lambda-\lambda_j)^{p_j}(\lambda-\lambda_j^{-1})^{p_j} \qquad (\lambda_j^2\neq 1;\ j=1,\ 2,\ \ldots,\ u).$$

Рассмотрим многочлен $g(\lambda)$ степени $<m$ [m — степень $\psi(\lambda)$], у которого $g(1)=1$, а все остальные $m-1$ значений на спектре матрицы O равны нулю, и положим [2]):

$$P=g(O). \tag{77}$$

Заметим, что функции $[g(\lambda)]^2$ и $g\left(\frac{1}{\lambda}\right)$ принимают те же значения на спектре матрицы O, что и функция $g(\lambda)$. Поэтому

$$P^2=P, \quad P'=g(O')=g(O^{-1})=P, \tag{78}$$

т. е. P — симметрическая проекционная матрица [3]).

Определим многочлен $h(\lambda)$ и матрицу Q равенствами

$$h(\lambda)=(\lambda-1)\,g(\lambda), \tag{79}$$

$$Q=h(O)=(O-E)\,P. \tag{80}$$

Поскольку степень $[h(\lambda)]^{m_1}$ обращается в нуль на спектре матрицы O, эта степень делится на $\psi(\lambda)$ без остатка. Отсюда следует:

$$Q^{m_1}=0,$$

т. е. Q — нильпотентная матрица с индексом нильпотентности m_1.

Из (80) находим [4]):

$$Q'=(O'-E)\,P. \tag{81}$$

Рассмотрим матрицу

$$R=Q(Q'+2E). \tag{82}$$

Из (78), (80) и (81) следует:

$$R=QQ'+2Q=(O-O')\,P.$$

Из этого представления матрицы R видно, что R — к о с о с и м м е т р и ч е с к а я м а т р и ц а.

---

[1]) См. стр. 158.

[2]) Из основной формулы (см. стр. 111)

$$g(A)=\sum_{k=1}^{s}[g(\lambda_k)\,Z_{k1}+g'(\lambda_k)\,Z_{k2}+\ldots]$$

следует, что

$$P=Z_{11}.$$

[3]) См. гл. III, § 6 (стр. 81).

[4]) Все фигурирующие здесь матрицы P, Q, Q', $O'=O^{-1}$ перестановочны между собой и с O, поскольку все они являются функциями от O.

С другой стороны, из (82)

$$R^k = Q^k (Q' + 2E)^k \qquad (k = 1, 2, \ldots). \tag{83}$$

Но Q', как и Q, — нильпотентная матрица и, следовательно,

$$|Q' + 2E| \neq 0.$$

Поэтому из (83) вытекает, что при любом k матрицы R^k и Q^k имеют один и тот же ранг.

Но при k нечетном матрица R^k является кососимметрической и потому (см. стр. 322) имеет четный ранг. Следовательно, каждая из матриц

$$Q,\ Q^3,\ Q^5,\ \ldots$$

имеет четный ранг.

Поэтому, повторяя дословно для матрицы Q рассуждения, проведенные на стр. 322—323 для матрицы K, мы сможем утверждать, что среди элементарных делителей матрицы Q делители вида λ^{2p} повторяются четное число раз. Но каждому элементарному делителю λ^{2p} матрицы Q соответствует элементарный делитель $(\lambda-1)^{2p}$ матрицы O и наоборот[1]). Отсюда следует, что среди элементарных делителей матрицы O делители вида $(\lambda-1)^{2p}$ повторяются четное число раз.

Аналогичное утверждение для элементарных делителей вида $(\lambda+1)^{2p}$ мы получим, применяя доказанное уже положение к матрице $-O$.

Таким образом, теорема доказана полностью.

Докажем теперь обратную теорему.

Теорема 10. *Любая система степеней вида*

$$\left.\begin{array}{l} (\lambda-\lambda_j)^{p_j},\ (\lambda-\lambda_j^{-1})^{p_j} \quad (\lambda_j \neq 0;\ j = 1, 2, \ldots, u), \\ (\lambda-1)^{q_1},\ (\lambda-1)^{q_2},\ \ldots,\ (\lambda-1)^{q_v}, \\ (\lambda+1)^{t_1},\ (\lambda+1)^{t_2},\ \ldots,\ (\lambda+1)^{t_w} \\ (q_1, \ldots, q_v, t_1, \ldots, t_w — \text{нечетные числа}) \end{array}\right\} \tag{84}$$

является системой элементарных делителей некоторой комплексной ортогональной матрицы O[2]).

Доказательство. Обозначим через μ_j числа, связанные с числами $\lambda_j\ (j = 1, 2, \ldots, u)$ равенствами

$$\lambda_j = e^{\mu_j} \qquad (j = 1, 2, \ldots, u).$$

Введем в рассмотрение «канонические» кососимметрические матрицы (см. предыдущий параграф)

$$K_{\mu_j}^{(p_j p_j)} \quad (j = 1, 2, \ldots, u);\quad K^{(q_1)}, \ldots, K^{(q_v)}; K^{(t_1)}, \ldots, K^{(t_w)},$$

имеющие соответственно элементарные делители

$$(\lambda-\mu_j)^{p_j},\ (\lambda+\mu_j)^{p_j} \quad (j = 1, 2, \ldots, u);\quad \lambda^{q_1}, \ldots, \lambda^{q_v}; \lambda^{t_1}, \ldots, \lambda^{t_w}.$$

Если K — кососимметрическая матрица, то

$$O = e^K$$

---

[1]) Поскольку $h(1) = 0$, $h'(1) \neq 0$, то при переходе от матрицы O к матрице $Q = h(O)$ элементарные делители вида $(\lambda-1)^{2p}$ матрицы O, не расщепляясь, заменяются элементарными делителями λ^{2p} (см. гл. VI, § 7).

[2]) Некоторые (или даже все) из чисел λ_j могут равняться ± 1. Одно число или два из чисел u, v, w могут равняться нулю. Тогда элементарные делители соответствующего вида отсутствуют у матрицы O.

является ортогональной ($O' = e^{K'} = e^{-K} = O^{-1}$). При этом каждому элементарному делителю $(\lambda - \mu)^p$ матрицы K отвечает элементарный делитель $(\lambda - e^{\mu})^p$ матрицы O [1]).

Поэтому квазидиагональная матрица

$$\widetilde{O} = \{e^{K^{(p_1 p_1)}_{\mu_1}}, \ldots, e^{K^{(p_u p_u)}_{\mu_u}}; e^{K^{(q_1)}}, \ldots, e^{K^{(q_v)}}; -e^{K^{(t_1)}}, \ldots, -e^{K^{(t_w)}}\} \quad (85)$$

является ортогональной и имеет элементарные делители (84).

Теорема доказана.

Из теорем 4, 9 и 10 вытекает

Следствие. *Произвольная (комплексная) ортогональная матрица O всегда ортогонально-подобна ортогональной матрице, имеющей нормальную форму $\widetilde{O}$, т. е. существует такая ортогональная матрица O_1, что*

$$O = O_1 \widetilde{O} O_1^{-1}. \quad (86)$$

Примечание. Подобно тому как это было сделано для кососимметрической матрицы $\widetilde{K}$, можно конкретизировать форму диагональных клеток в нормальной форме $\widetilde{O}$ [2]).

---

[1]) Это следует из того, что при $f(\lambda) = e^{\lambda}$ имеем: $f'(\lambda) = e^{\lambda} \neq 0$ при любом λ

[2]) См. [256].

ГЛАВА XII

СИНГУЛЯРНЫЕ ПУЧКИ МАТРИЦ

§ 1. Введение

1. Настоящая глава посвящена следующей задаче:

Даны четыре матрицы A, B; A_1, B_1 *одинаковых размеров* $m \times n$ *с элементами из числового поля* K. *Требуется найти, при каких условиях существуют две квадратные неособенные матрицы* P *и* Q *соответственно порядков* m *и* n *такие, что одновременно*

$$PAQ = A_1, \quad PBQ = B_1 \text{ [1]}. \tag{1}$$

Вводя в рассмотрение пучки матриц $A + \lambda B$ и $A_1 + \lambda B_1$, два матричных равенства (1) можно заменить одним равенством

$$P(A + \lambda B)Q = A_1 + \lambda B_1. \tag{2}$$

Определение 1. Два пучка прямоугольных матриц $A + \lambda B$ и $A_1 + \lambda B_1$ одних и тех же размеров $m \times n$, связанные равенством (2), в котором P и Q — постоянные (т. е. не зависящие от λ) квадратные неособенные матрицы соответственно порядков m и n, мы будем называть *строго эквивалентными* [2]).

Согласно общему определению эквивалентности λ-матриц (см. гл. VI, стр. 138) пучки $A + \lambda B$ и $A_1 + \lambda B_1$ являются эквивалентными, если имеет место равенство вида (2), в котором P и Q — две квадратные λ-матрицы с постоянными и отличными от нуля определителями. При строгой же эквивалентности требуется дополнительно, чтобы матрицы P и Q не зависели от λ [3]).

Критерий эквивалентности пучков $A + \lambda B$ и $A_1 + \lambda B_1$ следует из общего критерия эквивалентности λ-матриц и состоит в совпадении инвариантных многочленов или, что то же, элементарных делителей пучков $A + \lambda B$ и $A_1 + \lambda B_1$ (см. гл. VI, стр. 144).

В настоящей главе будет установлен критерий строгой эквивалентности двух пучков матриц и для каждого пучка будет определена строго эквивалентная ему каноническая форма.

---

[1]) Если такие матрицы P и Q существуют, то их элементы могут быть выбраны из поля K. Это вытекает из того, что равенства (1) могут быть переписаны в виде $PA = A_1Q^{-1}$, $PB = B_1Q^{-1}$ и потому равносильны некоторой системе линейных однородных уравнений с коэффициентами из поля K относительно элементов матриц P и Q^{-1}.

[2]) См. гл. VI, стр. 138.

[3]) Мы заменили встречающийся в литературе термин «эквивалентные пучки» термином «строго эквивалентные пучки» для того, чтобы резко разграничить определение 1 от определения эквивалентности из гл. VI.

2. Поставленная задача допускает естественную геометрическую интерпретацию. Рассмотрим пучок линейных операторов $A + \lambda B$, отображающих R_n в R_m. При определенном выборе базисов в этих пространствах пучку операторов $A + \lambda B$ отвечает пучок прямоугольных матриц $A + \lambda B$ (размером $m \times n$); при изменении базисов в R_n и R_m пучок $A + \lambda B$ заменяется строго эквивалентным пучком $P(A + \lambda B)Q$, где P и Q — квадратные неособенные матрицы порядков m и n (см. гл. III, §§ 2 и 4). Таким образом, критерий строгой эквивалентности дает характеристику того класса пучков матриц $A + \lambda B$ (размером $m \times n$), которые описывают один и тот же пучок операторов $A + \lambda B$, отображающих R_n в R_m, при различных выборах базисов в этих пространствах.

Для получения канонической формы пучка нужно найти те базисы в R_n и R_m, в которых пучок операторов $A + \lambda B$ описывается возможно более простой матрицей.

Поскольку пучок операторов задается двумя операторами A и B, то можно еще сказать, что *настоящая глава посвящена одновременному изучению двух операторов A и B, отображающих R_n в R_m.*

3. Все пучки матриц $A + \lambda B$ размером $m \times n$ подразделяются на два основных типа: на *регулярные* и *сингулярные* пучки.

Определение 2. Пучок матриц $A + \lambda B$ называется *регулярным*, если 1) A и B — квадратные матрицы одного и того же порядка n и 2) определитель $|A + \lambda B|$ не равен тождественно нулю. Во всех остальных случаях ($m \neq n$ или $m = n$, но $|A + \lambda B| \equiv 0$) пучок называется *сингулярным*.

Критерий строгой эквивалентности, а также каноническая форма для регулярных пучков матриц были установлены К. Вейерштрассом в 1867 г. [150] на основе его теории элементарных делителей, изложенной нами в главах VI и VII. Аналогичные вопросы для сингулярных пучков получили свое разрешение позже, в 1890 г., в исследованиях Л. Кронекера [133] [1]). Результаты Кронекера и составляют основное содержание этой главы.

§ 2. Регулярный пучок матриц

1. Рассмотрим частный случай, когда пучки $A + \lambda B$ и $A_1 + \lambda B_1$ состоят из квадратных матриц ($m = n$) и $|B| \neq 0$, $|B_1| \neq 0$. В этом случае, как было доказано в главе VI (стр. 140), два понятия «эквивалентность» и «строгая эквивалентность» пучков совпадают. Поэтому, применяя к пучкам общий критерий эквивалентности λ-матриц (стр. 148), приходим к теореме:

Теорема 1. *Два пучка квадратных матриц одного и того же порядка $A + \lambda B$ и $A_1 + \lambda B_1$, у которых $|B| \neq 0$ и $|B_1| \neq 0$, являются строго эквивалентными в том и только в том случае, когда эти пучки имеют одни и те же элементарные делители в поле K.*

Пучок квадратных матриц $A + \lambda B$ с $|B| \neq 0$ в главе VI называется регулярным, поскольку он представляет собой частный случай регулярного матричного многочлена относительно λ (см. гл. IV, стр. 87). В предыдущем параграфе настоящей главы мы дали более широкое определение регулярного пучка. Согласно этому определению в регулярном пучке возможно равенство $|B| = 0$ (и даже $|A| = |B| = 0$).

---

[1]) Из дальнейших исследований, в которых по-иному трактуются сингулярные пучки матриц, укажем на [99б, 251, 207а].

Для того чтобы выяснить, сохранится ли теорема 1 для регулярных пучков (при расширенном определении 1), рассмотрим следующий пример:

$$A+\lambda B=\begin{Vmatrix}2&1&3\\3&2&5\\3&2&6\end{Vmatrix}+\lambda\begin{Vmatrix}1&1&2\\1&1&2\\1&1&3\end{Vmatrix},\quad A_1+\lambda B_1=\begin{Vmatrix}2&1&1\\1&2&1\\1&1&1\end{Vmatrix}+\lambda\begin{Vmatrix}1&1&1\\1&1&1\\1&1&1\end{Vmatrix}. \tag{3}$$

Нетрудно видеть, что здесь каждый из пучков $A+\lambda B$ и $A_1+\lambda B_1$ имеет только один элементарный делитель $\lambda+1$. В то же время эти пучки не являются строго эквивалентными, так как матрицы B и B_1 имеют соответственно ранги 2 и 1, а из равенства (2), если бы оно имело место, следовало бы, что ранги матриц B и B_1 равны между собой. При этом пучки (3) являются регулярными согласно определению 1, так как

$$|A+\lambda B|\equiv|A_1+\lambda B_1|\equiv\lambda+1.$$

Разобранный пример показывает, что теорема 1 неверна при расширенном определении регулярного пучка.

2. Для того чтобы сохранить теорему 1, нам придется ввести понятие о «бесконечных» элементарных делителях пучка. Будем пучок $A+\lambda B$ задавать при помощи «однородных» параметров λ, μ : $\mu A+\lambda B$. Тогда определитель $\Delta(\lambda,\mu)\equiv|\mu A+\lambda B|$ будет однородной функцией от λ, μ. Определяя наибольший общий делитель $D_k(\lambda,\mu)$ всех миноров k-го порядка матрицы $\mu A+\lambda B$ ($k=1,2,\ldots,n$), получим инвариантные многочлены по известным формулам

$$i_1(\lambda,\mu)=\frac{D_n(\lambda,\mu)}{D_{n-1}(\lambda,\mu)},\quad i_2(\lambda,\mu)=\frac{D_{n-1}(\lambda,\mu)}{D_{n-2}(\lambda,\mu)},\quad\ldots;$$

при этом все $D_k(\lambda,\mu)$ и $i_j(\lambda,\mu)$ — однородные относительно λ и μ многочлены. Разлагая инвариантные многочлены на степени неприводимых в поле K однородных многочленов, получим элементарные делители $e_\alpha(\lambda,\mu)$ ($\alpha=1,2,\ldots$) пучка $\mu A+\lambda B$ в поле K.

Совершенно очевидно, что, полагая $\mu=1$ в $e_\alpha(\lambda,\mu)$, мы вернемся к элементарным делителям $e_\alpha(\lambda)$ пучка $A+\lambda B$. Обратно, из каждого элементарного делителя $e_\alpha(\lambda)$ степени q пучка $A+\lambda B$ мы получим соответствующий элементарный делитель $e_\alpha(\lambda,\mu)$ по формуле $e_\alpha(\lambda,\mu)=\mu^q e_\alpha\left(\frac{\lambda}{\mu}\right)$. Таким способом могут быть получены все элементарные делители пучка $\mu A+\lambda B$ за исключением элементарных делителей вида μ^q.

Элементарные делители вида μ^q существуют в том и только в том случае, когда $|B|=0$, и носят название «бесконечных» элементарных делителей для пучка $A+\lambda B$.

Поскольку из строгой эквивалентности пучков $A+\lambda B$ и $A_1+\lambda B_1$ следует строгая эквивалентность пучков $\mu A+\lambda B$ и $\mu A_1+\lambda B_1$, то у строго эквивалентных пучков $A+\lambda B$ и $A_1+\lambda B_1$ должны совпадать не только «конечные», но и «бесконечные» элементарные делители.

Пусть теперь даны два регулярных пучка $A+\lambda B$ и $A_1+\lambda B_1$, у которых соответственно совпадают все (в том числе и бесконечные) элементарные делители. Введем однородные параметры: $\mu A+\lambda B$, $\mu A_1+\lambda B_1$. Преобразуем параметры:

$$\begin{aligned}\lambda&=\alpha_1\tilde\lambda+\alpha_2\tilde\mu,\\ \mu&=\beta_1\tilde\lambda+\beta_2\tilde\mu\end{aligned}\qquad(\alpha_1\beta_2-\alpha_2\beta_1\neq0).$$

В новых параметрах пучки запишутся так: $\tilde\mu\tilde A+\tilde\lambda\tilde B$, $\tilde\mu\tilde A_1+\tilde\lambda\tilde B_1$, где

$\widetilde{B} = \beta_1 A + \alpha_1 B$, $\widetilde{B}_1 = \beta_1 A_1 + \alpha_1 B_1$. Из регулярности пучков $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ вытекает, что можно подобрать числа α_1 и β_1 так, чтобы $|\widetilde{B}| \neq 0$ и $|\widetilde{B}_1| \neq 0$.

Поэтому согласно теореме 1 пучки $\widetilde{\mu}\widetilde{A} + \widetilde{\lambda}\widetilde{B}$ и $\widetilde{\mu}\widetilde{A}_1 + \widetilde{\lambda}\widetilde{B}_1$, а следовательно, и исходные пучки $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ (или, что то же, $A + \lambda B$ и $A_1 + \lambda B_1$) строго эквивалентны. Таким образом, мы пришли к следующему обобщению теоремы 1.

Теорема 2. *Для того чтобы два регулярных пучка $A + \lambda B$ и $A_1 + \lambda B_1$ были строго эквивалентны, необходимо и достаточно, чтобы эти пучки имели одни и те же («конечные» и «бесконечные») элементарные делители.*

В разобранном ранее примере пучки (3) имели один и тот же «конечный» элементарный делитель $\lambda + 1$, но отличались «бесконечными» элементарными делителями (первый пучок имеет один «бесконечный» элементарный делитель μ^2, а второй — два: μ, μ). Поэтому эти пучки и не оказались строго эквивалентными.

3. Пусть теперь дан произвольный регулярный пучок $A + \lambda B$. Тогда существует такое число c, что $|A + cB| \neq 0$. Данный пучок представим в виде $A_1 + (\lambda - c)B$, где $A_1 = A + cB$ и потому $|A_1| \neq 0$. Умножим пучок слева на A_1^{-1}: $E + (\lambda - c)A_1^{-1}B$. Преобразованием подобия приводим этот пучок к виду

$$E + (\lambda - c)\{J_0, J_1\} = \{E - cJ_0 + \lambda J_0, E - cJ_1 + \lambda J_1\}\ ^{1}), \tag{4}$$

где $\{J_0, J_1\}$ — квазидиагональная нормальная форма матрицы $A_1^{-1}B$, J_0 — жорданова нильпотентная [2]) матрица, а $|J_1| \neq 0$.

Первый диагональный блок правой части (4) умножим на $(E - cJ_0)^{-1}$. Получим: $E + \lambda(E - cJ_0)^{-1}J_0$. Здесь коэффициент при λ — нильпотентная матрица [3]). Поэтому преобразованием подобия этот пучок можно привести к виду

$$E + \widehat{\lambda J}_0 = \{N^{(u_1)}, N^{(u_2)}, \ldots, N^{(u_s)}\}$$
$$(N^{(u)} = E^{(u)} + \lambda H^{(u)})\ ^{4}). \tag{5}$$

Второй диагональный блок в правой части (4) умножением на J_1^{-1}, а затем преобразованием подобия может быть приведен к виду $J + \lambda E$, где J — матрица, имеющая нормальную форму [5]), а E — единичная матрица. Мы пришли к теореме

Теорема 3. *Произвольный регулярный пучок $A + \lambda B$ может быть приведен к (строго эквивалентному) каноническому квазидиагональному виду*

$$\{N^{(u_1)}, N^{(u_2)}, \ldots, N^{(u_s)}, J + \lambda E\}$$
$$(N^{(u)} = E^{(u)} + \lambda H^{(u)}), \tag{6}$$

---

[1]) Единичные матрицы E в диагональных блоках правой части (4) имеют соответственно те же порядки, что J_0 и J_1.

[2]) То есть $J_0^l = 0$ при некотором целом $l > 0$.

[3]) Из $J_0^l = 0$ следует: $[(E - cJ_0)^{-1}J_0]^l = 0$.

[4]) Здесь $E^{(u)}$ — единичная матрица порядка u, а $H^{(u)}$ — матрица порядка u, у которой элементы первой наддиагонали равны единице, а остальные элементы равны нулю.

[5]) Поскольку здесь матрицу J можно заменить любой матрицей, ей подобной, то можно считать, что J имеет любую нормальную форму [например, естественную первого рода или второго рода или жорданову (см. гл. VI, § 6)].

где первые s диагональных блоков соответствуют бесконечным элементарным делителям $\mu^{u_1}, \mu^{u_2}, \ldots, \mu^{u_s}$ пучка $A+\lambda B$, а нормальная форма последнего диагонального блока $J+\lambda E$ однозначно определяется конечными элементарными делителями данного пучка.

§ 3. Сингулярные пучки. Теорема о приведении

Переходим к рассмотрению сингулярного пучка матриц $A+\lambda B$ с размерами $m \times n$. Обозначим через r *ранг пучка*, т. е. наибольший из порядков миноров, не равных тождественно нулю. Из сингулярности пучка следует, что всегда имеет место по крайней мере одно из неравенств $r<n$ или $r<m$. Пусть $r<n$. Тогда столбцы λ-матрицы $A+\lambda B$ линейно зависимы, т. е. уравнение

$$(A+\lambda B)\,x=0, \tag{7}$$

где x — искомый столбец, имеет ненулевое решение. Каждое ненулевое решение этого уравнения определяет некоторую линейную зависимость между столбцами λ-матрицы $A+\lambda B$. Мы ограничимся только теми решениями $x(\lambda)$ уравнения (7), которые являются многочленами относительно λ [1]), и среди этих решений возьмем решение наименьшей степени ε

$$x(\lambda)=x_0-\lambda x_1+\lambda^2 x_2-\ldots+(-1)^{\varepsilon}\lambda^{\varepsilon}x_{\varepsilon} \qquad (x_{\varepsilon}\neq 0). \tag{8}$$

Подставляя это решение в (7) и приравнивая нулю коэффициенты при степенях λ, получим:

$$Ax_0=0,\ Bx_0-Ax_1=0,\ Bx_1-Ax_2=0,\ \ldots,\ Bx_{\varepsilon-1}-Ax_{\varepsilon}=0,\ Bx_{\varepsilon}=0. \tag{9}$$

Рассматривая эту систему равенств как систему линейных однородных уравнений относительно элементов столбцов $x_0, -x_1, +x_2, \ldots, (-1)^{\varepsilon}x_{\varepsilon}$, заключаем, что матрица коэффициентов этой системы

$$M_{\varepsilon}=M_{\varepsilon}[A+\lambda B]=\left\|\begin{matrix} \overbrace{\begin{matrix} A & 0 & \ldots & 0 \\ B & A & & \cdot \\ 0 & B & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdot & A \\ 0 & 0 & \ldots & B \end{matrix}}^{\varepsilon+1} \end{matrix}\right\| \tag{10}$$

имеет ранг $\varrho_{\varepsilon}<(\varepsilon+1)n$. В то же время в силу минимального свойства числа ε для рангов $\varrho_1, \ldots, \varrho_{\varepsilon-1}$ матриц

$$M_0=\left\|\begin{matrix} A \\ B \end{matrix}\right\|,\quad M_1=\left\|\begin{matrix} A & 0 \\ B & A \\ 0 & B \end{matrix}\right\|,\ \ldots,\ M_{\varepsilon-1}=\left\|\begin{matrix} \overbrace{\begin{matrix} A & 0 & \ldots & 0 \\ B & A & & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & & \ddots & A \\ 0 & \ldots & & B \end{matrix}}^{\varepsilon} \end{matrix}\right\| \tag{10'}$$

имеют место равенства $\varrho_0=n,\ \varrho_1=2n,\ \ldots,\ \varrho_{\varepsilon-1}=\varepsilon n$.

[1]) Для определения элементов столбца x, удовлетворяющего уравнению (7), приходится решать систему линейных однородных уравнений, у которых коэффициенты при неизвестных линейно зависят от λ. Базисные линейно независимые решения x всегда могут быть выбраны так, чтобы их элементами были многочлены от λ.

Таким образом, *число* ε *есть наименьшее значение индекса* k, *при котором в соотношении* $\varrho_k \leqslant (k+1)n$ *имеет место знак* $<$.

Теперь мы сформулируем и докажем следующую фундаментальную теорему:

Теорема 4. *Если уравнение* (7) *имеет решение минимальной степени* ε *и* $\varepsilon > 0$, *то данный пучок* $A+\lambda B$ *строго эквивалентен пучку вида*

$$\left\| \begin{matrix} L_\varepsilon & 0 \\ 0 & \widehat{A}+\lambda\widehat{B} \end{matrix} \right\|, \tag{11}$$

где

$$L_\varepsilon = \left\| \begin{matrix} \overbrace{\lambda & 1 & 0 & \dots & 0 & 0}^{\varepsilon+1} \\ 0 & \lambda & 1 & & . & . \\ & & \ddots & \ddots & . & . \\ . & . & & \ddots & \ddots & . \\ . & . & & & \ddots & 0 \\ . & . & & & & 1 \\ 0 & 0 & & \dots & \lambda & 1 \end{matrix} \right\} \varepsilon, \tag{12}$$

а $\widehat{A}+\lambda\widehat{B}$ *— пучок матриц, для которого уравнение, аналогичное* (7), *не имеет решений степени* $< \varepsilon$.

Доказательство теоремы разобьем на три этапа. Сначала докажем, что данный пучок $A+\lambda B$ строго эквивалентен пучку вида

$$\left\| \begin{matrix} L_\varepsilon & D+\lambda F \\ 0 & \widehat{A}+\lambda\widehat{B} \end{matrix} \right\|, \tag{13}$$

где D, F, $\widehat{A}$, $\widehat{B}$ — постоянные прямоугольные матрицы соответственных размеров. Затем установим, что уравнение $(\widehat{A}+\lambda\widehat{B})\,\widehat{x}=0$ не имеет решений $\widehat{x}(\lambda)$ степени $< \varepsilon$. После этого мы покажем, что дальнейшими преобразованиями пучок (13) может быть приведен к квазидиагональному виду (11).

1. Первую часть доказательства облечем в геометрическую форму. Вместо пучка матриц $A+\lambda B$ рассмотрим пучок операторов $\boldsymbol{A}+\lambda\boldsymbol{B}$, отображающих $\boldsymbol{R}_n$ в $\boldsymbol{R}_m$, и покажем, что при надлежащем выборе базисов в этих пространствах матрица, соответствующая оператору $\boldsymbol{A}+\lambda\boldsymbol{B}$, будет иметь форму (13).

Вместо уравнения (7) возьмем векторное уравнение

$$(\boldsymbol{A}+\lambda\boldsymbol{B})\,\boldsymbol{x}=0 \tag{14}$$

с векторным решением

$$\boldsymbol{x}(\lambda)=\boldsymbol{x}_0-\lambda\boldsymbol{x}_1+\lambda^2\boldsymbol{x}_2-\dots+(-1)^\varepsilon\lambda^\varepsilon\boldsymbol{x}_\varepsilon; \tag{15}$$

равенства (9) заменятся векторными равенствами

$$\boldsymbol{A}\boldsymbol{x}_0=0,\quad \boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{B}\boldsymbol{x}_0,\quad \boldsymbol{A}\boldsymbol{x}_2=\boldsymbol{B}\boldsymbol{x}_1,\ \dots,\ \boldsymbol{A}\boldsymbol{x}_\varepsilon=\boldsymbol{B}\boldsymbol{x}_{\varepsilon-1},\quad \boldsymbol{B}\boldsymbol{x}_\varepsilon=0. \tag{16}$$

Ниже мы докажем, что векторы

$$\boldsymbol{A}\boldsymbol{x}_1,\ \boldsymbol{A}\boldsymbol{x}_2,\ \dots,\ \boldsymbol{A}\boldsymbol{x}_\varepsilon \tag{17}$$

линейно независимы. Отсюда легко будет следовать линейная независимость векторов

$$x_0,\ x_1,\ \ldots,\ x_\varepsilon. \tag{18}$$

Действительно, поскольку $Ax_0 = 0$, из $\alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_s x_s = 0$ находим: $\alpha_1 Ax_1 + \ldots + \alpha_s Ax_\varepsilon = 0$, откуда в силу линейной независимости векторов (17) $\alpha_1 = \alpha_2 = \ldots = \alpha_\varepsilon = 0$. Но $x_0 \neq 0$, поскольку в противном случае $\frac{1}{\lambda} x(\lambda)$ было бы решением уравнения (14) степени $\varepsilon - 1$, что невозможно. Поэтому и $\alpha_0 = 0$.

Если теперь принять векторы (17) и (18) в качестве первых базисных векторов для новых базисов соответственно в R_m и R_n, то в новых базисах операторам A и B в силу (16) будут соответствовать матрицы

$$\widetilde{A} = \left\| \begin{array}{ccccccccc} \overbrace{0 & 1 & & \ldots & 0}^{\varepsilon+1} & * & \ldots & * \\ 0 & 0 & 1 & \ldots & 0 & * & \ldots & * \\ \cdot & \cdot & \cdot & & \cdot & & & \\ \cdot & \cdot & & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & & & \\ 0 & 0 & & \ldots & 1 & * & \ldots & * \\ 0 & 0 & & \ldots & 0 & * & \ldots & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & \ldots & 0 & * & \ldots & * \end{array} \right\|, \quad \widetilde{B} = \left\| \begin{array}{ccccccccc} \overbrace{1 & 0 & \ldots & 0 & 0}^{\varepsilon+1} & * & \ldots & * \\ 0 & 1 & \ldots & 0 & 0 & * & \ldots & * \\ \cdot & & \cdot & & \cdot & & & \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & & & \\ 0 & 0 & \ldots & & 1 & 0 & * & \ldots & * \\ 0 & 0 & \ldots & & 0 & 0 & * & \ldots & * \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & & 0 & 0 & * & \ldots & * \end{array} \right\|;$$

тогда λ-матрица $\widetilde{A} + \lambda\widetilde{B}$ будет иметь вид (13). Все предыдущие рассуждения будут обоснованными, если мы докажем, что векторы (17) линейно независимы. Допустим противное, и пусть Ax_h $(h \geqslant 1)$ — первый в ряду (17) вектор, линейно зависящий от предыдущих векторов:

$$Ax_h = \alpha_1 Ax_{h-1} + \alpha_2 Ax_{h-2} + \ldots + \alpha_{h-1} Ax_1.$$

В силу (16) это равенство может быть переписано так:

$$Bx_{h-1} = \alpha_1 Bx_{h-2} + \alpha_2 Bx_{h-3} + \ldots + \alpha_{h-1} Bx_0,$$

т. е.

$$Bx^*_{h-1} = 0,$$

где

$$x^*_{h-1} = x_{h-1} - \alpha_1 x_{h-2} - \alpha_2 x_{h-3} - \ldots - \alpha_{h-1} x_0.$$

Далее, опять в силу (16)

$$Ax^*_{h-1} = B(x_{h-2} - \alpha_1 x_{h-3} - \ldots - \alpha_{h-2} x_0) = Bx^*_{h-2},$$

где

$$x^*_{h-2} = x_{h-2} - \alpha_1 x_{h-3} - \ldots - \alpha_{h-2} x_0.$$

Продолжая этот процесс далее и вводя еще векторы

$$x^*_{h-3} = x_{h-3} - \alpha_1 x_{h-4} - \ldots - \alpha_{h-3} x_0,\ \ldots,\ x^*_1 = x_1 - \alpha_1 x_0,\ x^*_0 = x_0,$$

мы получим цепочку равенств

$$\boldsymbol{B}x_{h-1}^{*}=0,\quad \boldsymbol{A}x_{h-1}^{*}=\boldsymbol{B}x_{h-2}^{*},\ \ldots,\ \boldsymbol{A}x_{1}^{*}=\boldsymbol{B}x_{0}^{*},\quad \boldsymbol{A}x_{0}^{*}=0. \tag{19}$$

Из (19) следует, что

$$x^{*}(\lambda)=x_{0}^{*}-\lambda x_{1}^{*}+\ldots+(-1)^{h-1}x_{h-1}^{*}\qquad (x_{0}^{*}=x_{0}\neq 0)$$

есть ненулевое решение уравнения (14) степени $\leqslant h-1<\varepsilon$, что невозможно. Таким образом, векторы (17) линейно независимы.

2. Докажем теперь, что уравнение $(\widehat{A}+\lambda\widehat{B})\widehat{x}=0$ не имеет решений степени $<\varepsilon$. Сначала обратим внимание на то, что уравнение $L_{\varepsilon}y=0$, как и уравнение (7), имеет ненулевое решение наименьшей степени ε. В этом можно убедиться непосредственно, если матричное уравнение $L_{\varepsilon}y=0$ заменить системой обыкновенных уравнений

$$\lambda y_{1}+y_{2}=0,\quad \lambda y_{2}+y_{3}=0,\ \ldots,\ \lambda y_{\varepsilon}+y_{\varepsilon+1}=0,$$

$[y=(y_{1},\ y_{2},\ \ldots,\ y_{\varepsilon+1})]$, откуда $y_{k}=(-1)^{k-1}y_{1}\lambda^{k-1}\qquad (k=1,\ 2,\ \ldots,\ \varepsilon+1)$.

С другой стороны, если пучок имеет «треугольный» вид (13), то соответствующие этому пучку матрицы $M_{k}(k=0,\ 1,\ \ldots,\ \varepsilon)$ [см. (10) и (10′) на стр. 335] после надлежащей перестановки строк и столбцов также могут быть приведены к треугольному виду

$$\left\|\begin{matrix} M_{k}[L_{\varepsilon}] & M_{k}[D+\lambda F] \\ 0 & M_{k}[\widehat{A}+\lambda\widehat{B}] \end{matrix}\right\|. \tag{20}$$

При $k=\varepsilon-1$ все столбцы этой матрицы, а значит, и столбцы матрицы $M_{\varepsilon-1}[L_{\varepsilon}]$, линейно независимы[1]). Но $M_{\varepsilon-1}[L_{\varepsilon}]$ — квадратная матрица порядка $\varepsilon(\varepsilon+1)$. Поэтому и в матрице $M_{\varepsilon-1}[\widehat{A}+\lambda\widehat{B}]$ все столбцы линейно независимы, а это, как было выяснено в начале параграфа, означает, что уравнение $(\widehat{A}+\lambda\widehat{B})\widehat{x}=0$ не имеет решений степени $\leqslant\varepsilon-1$, что и требовалось доказать.

3. Заменим пучок (13) строго эквивалентным ему пучком

$$\left\|\begin{matrix} E_{1} & Y \\ 0 & E_{2} \end{matrix}\right\|\left\|\begin{matrix} L_{\varepsilon} & D+\lambda F \\ 0 & \widehat{A}+\lambda\widehat{B} \end{matrix}\right\|\left\|\begin{matrix} E_{3} & -X \\ 0 & E_{4} \end{matrix}\right\|=\left\|\begin{matrix} L_{\varepsilon} & D+\lambda F+Y(\widehat{A}+\lambda\widehat{B})-L_{\varepsilon}X \\ 0 & \widehat{A}+\lambda\widehat{B} \end{matrix}\right\|, \tag{21}$$

где E_{1}, E_{2}, E_{3}, E_{4} — квадратные единичные матрицы соответственно порядков ε, $m-\varepsilon$, $\varepsilon+1$ и $n-\varepsilon-1$, а X, Y — произвольные постоянные прямоугольные матрицы соответствующих размеров. Наша теорема будет полностью доказана, если мы покажем, что матрицы X и Y могут быть выбраны так, чтобы имело место матричное равенство

$$L_{\varepsilon}X=D+\lambda F+Y(\widehat{A}+\lambda\widehat{B}). \tag{22}$$

---

[1]) Это следует из того, что ранг матрицы (20) при $k=\varepsilon-1$ равен εn; аналогичное равенство имеет место для ранга матрицы $M_{\varepsilon-1}[L_{\varepsilon}]$.

Введем обозначения для элементов матриц D, F, X, а также для строк матрицы Y и для столбцов матриц $\widehat{A}$, $\widehat{B}$:

$$D=\|d_{ik}\|, \quad F=\|f_{ik}\|, \quad X=\|x_{jk}\|$$

$$(i=1, 2, \ldots, \varepsilon;\ k=1, 2, \ldots, n-\varepsilon-1;\ j=1, 2, \ldots, \varepsilon+1),$$

$$Y=\begin{Vmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_\varepsilon \end{Vmatrix}, \quad \widehat{A}=(a_1, a_2, \ldots, a_{n-\varepsilon-1}), \quad \widehat{B}=(b_1, b_2, \ldots, b_{n-\varepsilon-1}).$$

Тогда матричное уравнение (22) можно заменить системой скалярных уравнений, записывая, что элементы k-го столбца в левой и правой частях равенства (22) соответственно равны друг другу ($k=1, 2, \ldots, n-\varepsilon-1$):

$$\left.\begin{aligned} x_{2k}+\lambda x_{1k}&=d_{1k}+\lambda f_{1k}+y_1a_k+\lambda y_1b_k,\\ x_{3k}+\lambda x_{2k}&=d_{2k}+\lambda f_{2k}+y_2a_k+\lambda y_2b_k,\\ x_{4k}+\lambda x_{3k}&=d_{3k}+\lambda f_{3k}+y_3a_k+\lambda y_3b_k,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ x_{\varepsilon+1,\,k}+\lambda x_{\varepsilon k}&=d_{\varepsilon k}+\lambda f_{\varepsilon k}+y_\varepsilon a_k+\lambda y_\varepsilon b_k\\ &(k=1, 2, \ldots, n-\varepsilon-1). \end{aligned}\right\} \tag{23}$$

В левых частях этих равенств стоят линейные двучлены относительно λ. Свободный член каждого из первых $\varepsilon-1$ этих двучленов равен коэффициенту при λ в следующем двучлене. Но тогда и правые части должны удовлетворять этому условию. Поэтому

$$\begin{aligned} y_1a_k-y_2b_k&=f_{2k}-d_{1k},\\ y_2a_k-y_3b_k&=f_{3k}-d_{2k},\\ &\cdots\cdots\cdots\cdots\\ y_{\varepsilon-1}a_k-y_\varepsilon b_k&=f_{\varepsilon k}-d_{\varepsilon-1,\,k}\\ &(k=1, 2, \ldots, n-\varepsilon-1). \end{aligned} \tag{24}$$

Если равенства (24) имеют место, то, очевидно, из (23) можно определить искомые элементы матрицы X.

Теперь осталось показать, что система уравнений (24) относительно элементов матрицы Y всегда имеет решение при любых d_{ik} и f_{ik} ($i=1, 2, \ldots, \varepsilon$; $k=1, 2, \ldots, n-\varepsilon-1$). Действительно, матрица, составленная из коэффициентов при неизвестных элементах строк y_1, $-y_2$, $+y_3$, $-y_4$, ..., может быть записана после транспонирования в виде

$$\begin{Vmatrix} \overbrace{\begin{matrix} \widehat{A} & 0 & \ldots & 0 \\ \widehat{B} & \widehat{A} & \ddots & \\ 0 & \widehat{B} & \ddots & \\ \cdot & \cdot & \ddots & \\ \cdot & & & \widehat{A} \\ 0 & 0 & \ldots & \widehat{B} \end{matrix}}^{\varepsilon-1} \end{Vmatrix}$$

Но эта матрица является матрицей $M_{\varepsilon-2}$ для пучка прямоугольных матриц $\widehat{A}+\lambda\widehat{B}$ [см. (10′) на стр. 335]. Ранг же этой матрицы равен $(\varepsilon-1)\,(n-\varepsilon-1)$, поскольку по доказанному уравнение $(\widehat{A}+\lambda\widehat{B})\,\widehat{x}=0$ не имеет решений степени $<\varepsilon$. Таким образом, ранг системы уравнений (24) равен числу уравнений, а такая система при любых свободных членах является совместной (непротиворечивой).

Теорема доказана полностью.

§ 4. Каноническая форма сингулярного пучка матриц

Пусть дан произвольный сингулярный пучок матриц $A+\lambda B$ размеров $m\times n$. Допустим сначала, что как между столбцами, так и между строками этого пучка нет линейной зависимости с постоянными коэффициентами.

Пусть $r<n$, где r — ранг пучка, т. е. столбцы пучка $A+\lambda B$ линейно зависимы между собой. В этом случае уравнение $(A+\lambda B)\,x=0$ имеет ненулевое решение минимальной степени ε_1. Из принятого в начале этого параграфа ограничения следует, что $\varepsilon_1>0$. Поэтому согласно теореме 4 данный пучок можно преобразовать к виду

$$\begin{pmatrix} L_{\varepsilon_1} & 0 \\ 0 & A_1+\lambda B_1 \end{pmatrix},$$

где уравнение $(A_1+\lambda B_1)\,x^{(1)}=0$ не имеет решений $x^{(1)}$ степени $<\varepsilon_1$.

Если это уравнение имеет ненулевое решение минимальной степени ε_2 (при этом непременно $\varepsilon_2\geqslant\varepsilon_1$), то, применяя к пучку $A_1+\lambda B_1$ теорему 4, мы данный пучок преобразуем к виду

$$\begin{pmatrix} I_{\varepsilon_1} & 0 & 0 \\ 0 & L_{\varepsilon_2} & 0 \\ 0 & 0 & A_2+\lambda B_2 \end{pmatrix}.$$

Продолжая этот процесс далее, мы приведем данный пучок к квазидиагональному виду

$$\left\| \begin{matrix} L_{\varepsilon_1} & & & & 0 \\ & L_{\varepsilon_2} & & & \\ & & \ddots & & \\ & & & L_{\varepsilon_p} & \\ 0 & & & & A_p+\lambda B_p \end{matrix} \right\|, \tag{25}$$

где $0<\varepsilon_1\leqslant\varepsilon_2\leqslant\ldots\leqslant\varepsilon_p$, а уравнение $(A_p+\lambda B_p)\,x^{(p)}=0$ не имеет ненулевых решений, т. е. столбцы матрицы $A_p+\lambda B_p$ линейно независимы[1]).

Если строки пучка $A_p+\lambda B_p$ линейно зависимы, то транспонированный пучок $A_p'=\lambda B_p'$ может быть приведен к виду (25), где вместо чисел $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_p$

[1]) Очевидно, что $\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_p\leqslant m$, $\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_p+p\leqslant n$. Эти соотношения могут стать равенствами лишь одновременно. В этом случае блок $A_p+\lambda B_p$ будет отсутствовать.

будут фигурировать числа $(0<)\,\eta_1\leqslant\eta_2\leqslant\ldots\leqslant\eta_q$ [1]). Но тогда данный пучок $A+\lambda B$ окажется преобразованным к квазидиагональному виду

$$\left\|\begin{matrix} L_{\varepsilon_1} & & & & & & & & 0 \\ & L_{\varepsilon_2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & L_{\varepsilon_p} & & & & & \\ & & & & L'_{\eta_1} & & & & \\ & & & & & L'_{\eta_2} & & & \\ & & & & & & \ddots & & \\ & & & & & & & L'_{\eta_q} & \\ 0 & & & & & & & & A_0+\lambda B_0 \end{matrix}\right\|, \tag{26}$$

$$(0<\varepsilon_1\leqslant\varepsilon_2\leqslant\ldots\leqslant\varepsilon_p,\quad 0<\eta_1\leqslant\eta_2\leqslant\ldots\leqslant\eta_q),$$

где у пучка $A_0+\lambda B_0$ как столбцы, так и строки линейно независимы, т. е. $A_0+\lambda B_0$ — регулярный пучок [2]).

Рассмотрим теперь общий случай, когда строки и столбцы данного пучка могут быть связаны линейными зависимостями с постоянными коэффициентами. Обозначим максимальное число постоянных независимых решений уравнений

$$(A+\lambda B)\,x=0 \quad\text{и}\quad (A'+\lambda B')\,y=0$$

соответственно через g и h. Вместо первого из этих уравнений, подобно тому как мы это делали при доказательстве теоремы 4, рассмотрим соответствующее векторное уравнение $(\boldsymbol{A}+\lambda\boldsymbol{B})\,\boldsymbol{x}=0$ ($\boldsymbol{A}$ и $\boldsymbol{B}$ — операторы, отображающие $\boldsymbol{R}_n$ в $\boldsymbol{R}_m$). Линейно независимые постоянные решения этого уравнения обозначим через $e_1, e_2, \ldots, e_g$ и примем за первые базисные векторы в $\boldsymbol{R}_n$. Тогда в соответствующей матрице $\widetilde{A}+\lambda\widetilde{B}$ первые g столбцов будут состоять из нулей

$$\widetilde{A}+\lambda\widetilde{B}=(\overbrace{0,\ldots,0}^{g},\ \widetilde{A}_1+\lambda\widetilde{B}_1). \tag{27}$$

Совершенно так же в пучке $\widetilde{A}_1+\lambda\widetilde{B}_1$ первые h строк можно сделать нулевыми. Тогда данный пучок примет вид

$$\left\|\begin{matrix} \overbrace{0,\ldots,0}^{g} & & 0 \\ 0,\ \ldots,\ 0 & & A^0+\lambda B^0 \end{matrix}\right\}\begin{matrix} h \\ \\ \end{matrix}, \tag{28}$$

где строки и столбцы пучка $A^0+\lambda B^0$ уже не связаны линейными зависимостями с постоянными коэффициентами. К пучку $A^0+\lambda B^0$ применимо представление типа (26). Таким образом, в самом общем случае пучок

---

[1]) Так как между строками пучка $A+\lambda B$, а следовательно, и пучка $A_p+\lambda B_p$ нет линейной зависимости с постоянными коэффициентами, то $\eta_1>0$.

[2]) Если в данном пучке $r=n$, т. е. столбцы пучка линейно независимы, то в (26) будут отсутствовать первые p диагональных блоков вида L_ε ($p=0$). Точно так же, если $r=m$, т. е. в $A+\lambda B$ строки линейно независимы, то в (26) будут отсутствовать диагональные блоки вида L'_η ($q=0$).

$A+\lambda B$ всегда может быть приведен к каноническому квазидиагональному виду

$$\{h[\overset{g}{\overline{0}},\ L_{\varepsilon_{g+1}},\ \ldots,\ L_{\varepsilon_p},\ L'_{\eta_{h+1}},\ \ldots,\ L'_{\eta_g},\ A_0+\lambda B_0\}. \tag{29}$$

Выбор индексов при ε и η связан с тем, что нам удобно здесь считать $\varepsilon_1=\varepsilon_2=\ldots=\varepsilon_g=0$ и $\eta_1=\eta_2=\ldots=\eta_h=0$.

Заменяя фигурирующий в (29) регулярный пучок $A_0+\lambda B_0$ его канонической формой (6) (см. § 2, стр. 334), получим окончательно следующую квазидиагональную матрицу:

$$\{h[\overset{g}{\overline{0}};\ L_{\varepsilon_{g+1}},\ \ldots,\ L_{\varepsilon_p};\ L'_{\eta_{h+1}},\ \ldots,\ L'_{\eta_g};\ N^{(u_1)},\ \ldots,\ N^{(u_s)};\ J+\lambda E\}, \tag{30}$$

где матрица J имеет жорданову или естественную нормальную форму, а $N^{(u)}=E^{(u)}+\lambda H^{(u)}$.

Матрица (30) *представляет собой каноническую форму пучка* $A+\lambda B$ *в самом общем случае.*

Для того чтобы по данному пучку непосредственно определить его каноническую форму (30), не осуществляя последовательно процесс приведения, мы, следуя Кронекеру, в следующем параграфе введем понятие о минимальных индексах пучка.

§ 5. Минимальные индексы пучка. Критерий строгой эквивалентности пучков

Пусть дан произвольный сингулярный пучок прямоугольных матриц $A+\lambda B$. Тогда k многочленных столбцов $x_1(\lambda)$, $x_2(\lambda)$, ..., $x_k(\lambda)$, являющихся решениями уравнения

$$(A+\lambda B)\,x=0, \tag{31}$$

будут линейно зависимыми, если ранг многочленной матрицы, составленной из этих столбцов, $X=[x_1(\lambda),\ x_2(\lambda),\ \ldots,\ x_k(\lambda)]$ меньше k. В этом случае существует k многочленов $p_1(\lambda)$, $p_2(\lambda)$, ..., $p_k(\lambda)$, не равных одновременно тождественно нулю, таких, что

$$p_1(\lambda)\,x_1(\lambda)+p_2(\lambda)\,x_2(\lambda)+\ldots+p_k(\lambda)\,x_k(\lambda)\equiv 0.$$

Если же ранг матрицы X равен k, то подобной зависимости не существует, и решения $x_1(\lambda)$, $x_2(\lambda)$, ..., $x_k(\lambda)$ линейно независимы.

Среди всех решений уравнения (31) возьмем ненулевое решение $x_1(\lambda)$ наименьшей степени ε_1. Среди всех решений того же уравнения, линейно независимых от $x_1(\lambda)$, выберем решение $x_2(\lambda)$ наименьшей степени ε_2. Очевидно, что $\varepsilon_1\leqslant\varepsilon_2$. Этот процесс продолжим, выбирая среди решений, линейно независимых от $x_1(\lambda)$ и $x_k(\lambda)$, решение $x_3(\lambda)$ минимальной степени ε_3 и т. д. Так как число линейно независимых решений уравнения (31) всегда $\leqslant n$, то этот процесс должен закончиться. Мы получим *фундаментальный ряд решений* уравнения (31)

$$x_1(\lambda),\ x_2(\lambda),\ \ldots,\ x_p(\lambda) \tag{32}$$

со степенями

$$\varepsilon_1\leqslant\varepsilon_2\leqslant\ldots\leqslant\varepsilon_p. \tag{33}$$

В общем случае фундаментальный ряд решений не определяется однозначно (с точностью до скалярных множителей) заданием пучка $A+\lambda B$.

Однако *два различных фундаментальных ряда решений имеют всегда один и тот же ряд степеней* $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p$. Действительно, рассмотрим наряду с (32) второй фундаментальный ряд решений $\widetilde{x}_1(\lambda), \widetilde{x}_2(\lambda), \ldots$ со степенями $\widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2, \ldots$ Пусть среди степеней (33)

$$\varepsilon_1 = \ldots = \varepsilon_{n_1} < \varepsilon_{n_1+1} = \ldots = \varepsilon_{n_2} < \ldots$$

и аналогично в ряду $\widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2, \ldots$

$$\widetilde{\varepsilon}_1 = \ldots = \widetilde{\varepsilon}_{\widetilde{n}_1} < \widetilde{\varepsilon}_{\widetilde{n}_1+1} = \ldots = \widetilde{\varepsilon}_{\widetilde{n}_2} < \ldots$$

Очевидно, что $\varepsilon_1 = \widetilde{\varepsilon}_1$. Любой столбец $\widetilde{x}_i(\lambda)$ $(i = 1, 2, \ldots, \widetilde{n}_1)$ есть линейная комбинация столбцов $x_1(\lambda), x_2(\lambda), \ldots, x_{n_1}(\lambda)$, так как в противном случае в ряду (32) можно было бы решение $x_{n_1+1}(\lambda)$ заменить решением $\widetilde{x}_i(\lambda)$ с меньшей степенью. Очевидно, что и наоборот, любой столбец $x_i(\lambda)$ $(i = 1, 2, \ldots, n_1)$ является линейной комбинацией столбцов $\widetilde{x}_1(\lambda), \widetilde{x}_2(\lambda), \ldots, \widetilde{x}_{\widetilde{n}_1+1}(\lambda)$. Поэтому $n_1 = \widetilde{n}_1$ и $\varepsilon_{n_1+1} = \widetilde{\varepsilon}_{\widetilde{n}_1+1}$. Теперь аналогичными рассуждениями убеждаемся в том, что $n_2 = \widetilde{n}_2$ и $\varepsilon_{n_2+1} = \widetilde{\varepsilon}_{\widetilde{n}_2+1}$ и т. д.

Каждое решение $x_k(\lambda)$ фундаментального ряда (32) дает линейную зависимость степени ε_k между столбцами матрицы $A + \lambda B$ $(k = 1, 2, \ldots, p)$. Поэтому числа $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p$ называются *минимальными индексами для столбцов* пучка $A + \lambda B$.

Аналогично вводятся *минимальные индексы* $\eta_1, \eta_2, \ldots, \eta_q$ *для строк* пучка $A + \lambda B$. При этом уравнение $(A + \lambda B)x = 0$ заменяется уравнением $(A' + \lambda B')y = 0$ и числа $\eta_1, \eta_2, \ldots, \eta_q$ определяются как минимальные индексы для столбцов транспонированного пучка $A' + \lambda B'$.

Строго эквивалентные пучки имеют одни и те же минимальные индексы. Действительно, пусть даны два таких пучка: $A + \lambda B$ и $P(A + \lambda B)Q$ (P и Q — квадратные неособенные матрицы). Тогда уравнение (30) для первого пучка после почленного умножения слева на P может быть записано так:

$$P(A + \lambda B)Q \cdot Q^{-1}x = 0.$$

Отсюда видно, что все решения уравнения (30) после умножения слева на Q^{-1} дают полную систему решений уравнения

$$P(A + \lambda B)Qz = 0.$$

Поэтому пучки $A + \lambda B$ и $P(A + \lambda B)Q$ имеют одни и те же минимальные индексы для столбцов. Совпадение минимальных индексов для строк устанавливается переходом к транспонированным пучкам.

Вычислим минимальные индексы для канонической квазидиагональной матрицы

$$\{\underbrace{0, \ldots, 0}_{h}, L_{\varepsilon_{g+1}}, \ldots, L_{\varepsilon_p}; L'_{\eta_{h+1}}, \ldots, L'_{\eta_q}, A_0 + \lambda B_0\} \qquad (34)$$

[$A_0 + \lambda B_0$ — регулярный пучок, имеющий нормальную форму (6)].

Заметим предварительно, что *полная система минимальных индексов для столбцов* (строк) *квазидиагональной матрицы получается соединением из соответствующих систем минимальных индексов отдельных диагональных блоков.* Матрица L_ε имеет только один индекс ε для столбцов, а строки этой матрицы линейно независимы. Точно так же матрица L'_η имеет только один индекс η для строк, а столбцы этой матрицы линейно независимы. Регулярный пучок $A_0 + \lambda B_0$ совсем не имеет минимальных индексов.

Поэтому матрица (34) имеет минимальные индексы для столбцов

$$\varepsilon_1 = \ldots = \varepsilon_g = 0, \quad \varepsilon_{g+1}, \ \ldots, \ \varepsilon_p,$$

а для строк

$$\eta_1 = \ldots = \eta_h = 0, \quad \eta_{h+1}, \ \ldots, \ \eta_q.$$

Заметим еще, что матрица L_ε не имеет элементарных делителей, так как среди ее миноров максимального порядка ε имеется минор, равный единице, и минор, равный λ^ε. Это же положение, разумеется, верно и для транспонированной матрицы L'_ε. Так как элементарные делители квазидиагональной матрицы получаются путем соединения элементарных делителей отдельных диагональных блоков (см. гл. VI, стр. 146), то *элементарные делители λ-матрицы* (34) *совпадают с элементарными делителями ее регулярного «ядра»* $A_0 + \lambda B_0$.

Каноническая форма пучка (34) *вполне определяется заданием минимальных индексов* $\varepsilon_1, \ldots, \varepsilon_p, \eta_1, \ldots, \eta_q$ *и элементарных делителей этого пучка или (что то же) строго эквивалентного ему пучка* $A + \lambda B$. Так как два пучка, имеющих одну и ту же каноническую форму, строго эквивалентны между собой, то мы доказали следующую теорему:

Теорема 5 (Кронекера). *Для того чтобы два произвольных пучка прямоугольных матриц* $A + \lambda B$ *и* $A_1 + \lambda B_1$ *одних и тех же размеров* $m \times n$ *были строго эквивалентны, необходимо и достаточно, чтобы эти пучки имели одни и те же минимальные индексы и одни и те же («конечные» и «бесконечные») элементарные делители.*

В заключение для наглядности выпишем каноническую форму пучка $A + \lambda B$, имеющего минимальные индексы $\varepsilon_1 = 0$, $\varepsilon_2 = 1$, $\varepsilon_3 = 2$, $\eta_1 = 0$, $\eta_2 = 0$, $\eta_3 = 2$ и элементарные делители λ^2, $(\lambda + 2)^2$, μ^3:

$$\left\|\begin{array}{cccccccccc}
0 & & & & & & & & & \\
0 & & & & & & & & & \\
 & \lambda & 1 & & & & & & & \\
 & & \lambda & 1 & 0 & & & & & \\
 & & 0 & \lambda & 1 & & & & & \\
 & & & & \lambda & 0 & & & & \\
 & & & & 1 & \lambda & & & & \\
 & & & & 0 & 1 & & & & \\
 & & & & & & 1 & \lambda & 0 & \\
 & & & & & & 0 & 1 & \lambda & \\
 & & & & & & & 0 & 0 & 1 \\
 & & & & & & & & \lambda & 1 \\
 & & & & & & & & 0 & \lambda \\
 & & & & & & & & & \lambda+2 \quad 1 \\
 & & & & & & & & & 0 \quad \lambda+2
\end{array}\right\|^{1)} \qquad (35)$$

[1]) Все неотмеченные элементы этой матрицы равны нулю.

§ 6. Сингулярные пучки квадратичных форм

Пусть даны две комплексные квадратичные формы:

$$A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_i x_k, \qquad B(x, x) = \sum_{i,k=1}^{n} b_{ik}x_i x_k; \tag{36}$$

они порождают пучок квадратичных форм $A(x, x)+\lambda B(x, x)$. Этому пучку форм соответствует пучок симметричных матриц $A+\lambda B$ ($A'=A$, $B'=B$). Если мы в пучке форм $A(x, x)+\lambda B(x, x)$ переменные подвергнем неособенному линейному преобразованию $x=Tz$ ($|T|\neq 0$), то преобразованному пучку форм $\widetilde{A}(z, z)+\lambda\widetilde{B}(z, z)$ будет соответствовать пучок матриц

$$\widetilde{A}+\lambda\widetilde{B}=T'(A+\lambda B)T; \tag{37}$$

здесь T — постоянная (т. е. не зависящая от λ) неособенная квадратная матрица n-го порядка.

Два пучка матриц $A+\lambda B$ и $\widetilde{A}+\lambda\widetilde{B}$, связанные тождеством (37), называются *конгруэнтными* (ср. с определением 1 гл. X, стр. 269).

Очевидно, что конгруэнтность представляет собой специальный частный случай строгой эквивалентности пучков матриц. Однако в тех случаях, когда рассматривается конгруэнтность двух пучков симметрических (или кососимметрических) матриц, понятие конгруэнтности совпадает с понятием строгой эквивалентности. Это утверждает

Теорема 6. *Два строго эквивалентных пучка комплексных симметрических (или кососимметрических) матриц всегда конгруэнтны между собой.*

Доказательство. Пусть даны два строго эквивалентных пучка симметрических (кососимметрических) матриц $\Lambda\equiv A+\lambda B$ и $\widetilde{\Lambda}\equiv\widetilde{A}+\lambda\widetilde{B}$:

$$\widetilde{\Lambda}=P\Lambda Q \quad (\Lambda'=\pm\Lambda,\ \widetilde{\Lambda}'=\pm\widetilde{\Lambda};\ |P|\neq 0,\ |Q|\neq 0). \tag{38}$$

Переходя к транспонированным матрицам, получаем:

$$\widetilde{\Lambda}=Q'\Lambda P'. \tag{39}$$

Из (38) и (39) найдем:

$$\Lambda QP'^{-1}=P^{-1}Q'\Lambda. \tag{40}$$

Полагая

$$U=QP'^{-1}, \tag{41}$$

равенство (40) перепишем так:

$$\Lambda U=U'\Lambda. \tag{42}$$

Из (42) легко следует:

$$\Lambda U^k=U'^k\Lambda \qquad (k=0,\ 1,\ 2,\ \ldots),$$

и вообще

$$\Lambda S=S'\Lambda, \tag{43}$$

где

$$S=f(U), \tag{44}$$

а $f(\lambda)$ — произвольный многочлен относительно λ. Допустим, что этот многочлен так выбран, что $|S|\neq 0$. Тогда из (43) найдем:

$$\Lambda=S'\Lambda S^{-1}. \tag{45}$$

Подставляя полученное выражение для Λ в (38), будем иметь:

$$\widetilde{\Lambda} = PS'\Lambda S^{-1}Q. \tag{46}$$

Для того чтобы это соотношение было преобразованием конгруэнтности, нужно, чтобы выполнялось равенство

$$(PS')' = S^{-1}Q,$$

которое может быть переписано так:

$$S^2 = QP'^{-1} = U.$$

Но матрица $S = f(U)$ удовлетворит этому уравнению, если в качестве $f(\lambda)$ взять интерполяционный многочлен $\sqrt{\lambda}$ на спектре матрицы U. Это возможно сделать, поскольку многозначная функция $\sqrt{\lambda}$ имеет однозначную ветвь, определенную на спектре матрицы U, так как $|U| \neq 0$.

После этого равенство (46) станет условием конгруэнтности

$$\widetilde{\Lambda} = T'\Lambda T \quad \left(T = SQ = \sqrt{QP'^{-1}}Q\right). \tag{47}$$

Из доказанной теоремы и из теоремы 5 вытекает

Следствие. *Два пучка квадратичных форм*

$$A(x, x) + \lambda B(x, x) \text{ и } \widetilde{A}(z, z) + \lambda\widetilde{B}(z, z)$$

могут быть переведены друг в друга преобразованием $x = Tz$ ($|T| \neq 0$) *тогда и только тогда, когда пучки симметрических матриц* $A + \lambda B$ *и* $\widetilde{A} + \lambda\widetilde{B}$ *имеют одни и те же элементарные делители («конечные» и «бесконечные») и одни и те же минимальные индексы.*

Примечание. Для пучка симметрических матриц строки и столбцы имеют одни и те же минимальные индексы:

$$p = q; \quad \varepsilon_1 = \eta_1, \ \ldots, \ \varepsilon_p = \eta_p. \tag{48}$$

Поставим следующий вопрос. *Даны две произвольные комплексные квадратичные формы*

$$A(x, x) = \sum_{i,\,k=1}^{n} a_{ik}x_i x_k, \quad B(x, x) = \sum_{i,\,k=1}^{n} b_{ik}x_i x_k.$$

При каких условиях неособенным преобразованием переменных $x = Tz$ ($|T| \neq 0$) *можно одновременно привести эти формы к суммам квадратов*

$$\sum_{i=1}^{n} a_i z_i^2 \text{ и } \sum_{i=1}^{n} b_i z_i^2? \tag{49}$$

Аналогичный вопрос возникает для двух эрмитовых форм $A(x, x)$ и $B(x, x)$, но в этом случае вместо (49) следует писать:

$$\sum_{i=1}^{n} a_i z_i \bar{z}_i \text{ и } \sum_{i=1}^{n} b_i z_i \bar{z}_i, \tag{50}$$

причем здесь a_i и b_i ($i = 1, 2, \ldots, n$) — вещественные числа.

Допустим, что квадратичные формы $A(x, x)$ и $B(x, x)$ обладают указанным свойством. Тогда пучок матриц $A + \lambda B$ будет конгруэнтен пучку диагональных матриц

$$\{a_1 + \lambda b_1, \ a_2 + \lambda b_2, \ \ldots, \ a_n + \lambda b_n\}. \tag{51}$$

Пусть среди диагональных двучленов $a_i+\lambda b_i$ имеется ровно $r\,(r\leqslant n)$ не равных тождественно нулю. Не нарушая общности, можем считать, что

$$a_1=b_1=0,\ \ldots,\ a_{n-r}=b_{n-r}=0,\quad a_i+\lambda b_i\not\equiv 0\qquad (i=n-r+1,\ \ldots,\ n).$$

Полагая

$$A_0+\lambda B_0=\{a_{n-r+1}+\lambda b_{n-r+1},\ \ldots,\ a_n+\lambda b_n\},$$

мы матрицу (51) представим в виде

$$\{\overbrace{0,\ldots,0}^{n-r},\ A_0+\lambda B_0\}. \tag{52}$$

Сопоставляя (52) с (34) (стр. 343), мы видим, что в данном случае все минимальные индексы равны нулю. Кроме того, все элементарные делители имеют первую степень. Мы пришли к теореме

Теорема 7. *Две квадратичные формы $A(x, x)$ и $B(x, x)$ одновременным преобразованием переменных могут быть приведены к суммам квадратов* [(49) *или* (50)] *в том и только в том случае, когда у пучка матриц $A+\lambda B$ все элементарные делители (конечные и бесконечные) первой степени, а все минимальные индексы равны нулю.*

Для того чтобы в общем случае одновременно привести две квадратичные формы $A(x, x)$ и $B(x, x)$ к некоторому каноническому виду, нужно заменить пучок матриц $A+\lambda B$ строго эквивалентным ему «каноническим» пучком симметрических матриц.

Пусть пучок симметрических матриц $A+\lambda B$ имеет минимальные индексы $\varepsilon_1=\ldots=\varepsilon_g=0$, $\varepsilon_{g+1}\not\equiv 0,\ \ldots,\ \varepsilon_p\not\equiv 0$ и элементарные делители бесконечные $\mu^{u_1},\ \mu^{u_2},\ \ldots,\ \mu^{u_s}$ и конечные $(\lambda+\lambda_1)^{c_1},\ (\lambda+\lambda_2)^{c_2},\ \ldots\ (\lambda+\lambda_t)^{c_t}$. Тогда в канонической форме (30) $g=h$, $p=q$ и $\varepsilon_{g+1}=\eta_{g+1},\ \ldots,\ \varepsilon_p=\eta_p$. Заменим в (30) каждые два диагональных блока вида L_ε и L'_ε одним диагональным блоком $\begin{pmatrix}0 & L'_\varepsilon\\ L_\varepsilon & 0\end{pmatrix}$, а каждый блок вида $N^{(u)}=E^{(u)}+\lambda H^{(u)}$ заменим строго эквивалентным симметрическим блоком

$$\widetilde{N}^{(u)}=V^{(u)}N^{(u)}=\begin{Vmatrix}0 & 0 & \ldots & 0 & 1\\ 0 & 0 & \ldots & 1 & \lambda\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 1 & \lambda & \ldots & 0 & 0\end{Vmatrix}\left(V^{(u)}=\begin{Vmatrix}0 & 0 & \ldots & 0 & 1\\ 0 & 0 & \ldots & 1 & 0\\ \cdot & & \cdot & & \cdot\\ \cdot & \cdot & & & \cdot\\ \cdot & 1 & & & \cdot\\ 1 & 0 & \ldots & 0 & 0\end{Vmatrix}\right). \tag{53}$$

Кроме того, вместо регулярного диагонального блока $J+\lambda E$ в (30) (J — жорданова матрица)

$$J+\lambda E=\{(\lambda+\lambda_1)E^{(c_1)}+H^{(c_1)},\ \ldots,\ (\lambda+\lambda_t)E^{(c_t)}+H^{(c_t)}\} \tag{54}$$

возьмем строго эквивалентный ему пучок

$$\{Z_{\lambda_1}^{(c_1)},\ \ldots,\ Z_{\lambda_t}^{(c_t)}\}, \tag{55}$$

где

$$Z_{\lambda_i}^{(c_i)} = V^{(c_i)}[(\lambda+\lambda_i)E^{(c_i)} + H^{(c_i)}] =$$

$$= \begin{Vmatrix} 0 & \dots & 0 & \lambda+\lambda_i \\ 0 & \dots & \lambda+\lambda_i & 1 \\ \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \lambda+\lambda_i & 1 & \dots & 0 \end{Vmatrix} \qquad (i=1, 2, \dots, t). \tag{56}$$

Пучок $A+\lambda B$ строго эквивалентен симметрическому пучку

$$\widetilde{A}+\lambda\widetilde{B} = \left\{0, \begin{Vmatrix} 0 & L'_{\varepsilon_{g+1}} \\ L_{\varepsilon_{g+1}} & 0 \end{Vmatrix}, \dots, \begin{Vmatrix} 0 & L'_{\varepsilon_p} \\ L_{\varepsilon_p} & 0 \end{Vmatrix}; \widetilde{N}^{(u_1)}, \dots, \widetilde{N}^{(u_s)}; Z_{\lambda_1}^{(c_1)}, \dots, Z_{\lambda_t}^{(c_t)}\right\}. \tag{57}$$

Две квадратичные формы с комплексными коэффициентами $A(x, x)$ и $B(x, x)$ преобразованием переменных $x = Tz$ ($|T| \neq 0$) могут быть одновременно приведены к каноническому виду $\widetilde{A}(z, z)$ и $\widetilde{B}(z, z)$, определяемому равенством (57).

§ 7. Приложения к дифференциальным уравнениям

Рассмотрим приложения полученных результатов к интегрированию системы m линейных дифференциальных уравнений первого порядка с n неизвестными функциями с постоянными коэффициентами[1])

$$\sum_{k=1}^{n} a_{ik}x_k + \sum_{k=1}^{n} b_{ik}\frac{dx_k}{dt} = f_i(t) \qquad (i=1, 2, \dots, m) \tag{58}$$

или в матричной записи:

$$Ax + B\frac{dx}{dt} = f(t); \tag{59}$$

здесь

$$A = \|a_{ik}\|, \quad B = \|b_{ik}\| \qquad (i=1, 2, \dots, m;\ k=1, 2, \dots, n),$$

$$x = (x_1, x_2, \dots, x_n), \quad f = (f_1, f_2, \dots, f_m)^{2}).$$

Введем новые неизвестные функции $z_1, z_2, \dots, z_n$, связанные со старыми $x_1, x_2, \dots, x_n$ линейным неособенным преобразованием с

[1]) Частный случай, когда $m=n$ и система (58) разрешена относительно производных, был подробно исследован в гл. V, § 5.

Как известно, система линейных дифференциальных уравнений с постоянными коэффициентами любого s-го порядка может быть приведена к виду (58), если все производные от неизвестных функций до $(s-1)$-го порядка включительно дополнительно ввести в качестве новых неизвестных функций.

[2]) Напоминаем, что круглыми скобками обозначается столбцевая матрица. Так, $x=(x_1, x_2, \dots, x_n)$—столбец с элементами $x_1, x_2, \dots, x_n$.

постоянными коэффициентами:

$$x = Qz \qquad [z = (z_1, z_2, \ldots, z_n);\ |Q| \neq 0]. \tag{60}$$

Кроме того, вместо уравнений (58) можно взять любые m независимых линейных комбинаций их, что равносильно умножению матриц A, B, f слева на квадратную неособенную матрицу m-го порядка P. Подставляя Qz вместо x в (59) и умножая (59) почленно слева на P, получим:

$$\widetilde{A}z + \widetilde{B}\frac{dz}{dt} = \widetilde{f}(t), \tag{61}$$

где

$$\widetilde{A} = PAQ, \quad \widetilde{B} = PBQ, \quad \widetilde{f} = Pf = (\widetilde{f}_1, \widetilde{f}_2, \ldots, \widetilde{f}_n). \tag{62}$$

При этом пучки матриц $A + \lambda B$ и $\widetilde{A} + \lambda\widetilde{B}$ строго эквивалентны друг другу:

$$\widetilde{A} + \lambda\widetilde{B} = P(A + \lambda B)Q. \tag{63}$$

Выберем матрицы P и Q так, чтобы пучок $\widetilde{A} + \lambda\widetilde{B}$ имел каноническую квазидиагональную форму

$$\widetilde{A} + \lambda\widetilde{B} = \{0, L_{\varepsilon_{g+1}}, \ldots, L_{\varepsilon_p}, L'_{\eta_{h+1}}, \ldots, L'_{\eta_q}, N^{(u_1)}, \ldots, N^{(u_s)}, J + \lambda E\}. \tag{64}$$

В соответствии с диагональными блоками в (64) система дифференциальных уравнений распадается на $\nu = p - g + q - h + s + 2$ отдельных систем вида

$$0 \cdot \overset{1}{z} = \overset{1}{\widetilde{f}}, \tag{65}$$

$$L_{\varepsilon_{g+i}}\left(\frac{d}{dt}\right)\overset{1+i}{z} = \overset{1+i}{\widetilde{f}} \qquad (i = 1, 2, \ldots, p-g), \tag{66}$$

$$L'_{\eta_{h+j}}\left(\frac{d}{dt}\right)\overset{p-g+1+j}{z} = \overset{p-g+1+j}{\widetilde{f}} \qquad (j = 1, 2, \ldots, q-h), \tag{67}$$

$$N^{(u_k)}\left(\frac{d}{dt}\right)\overset{p-g+q-h+1+k}{z} = \overset{p-g+q-h+1+k}{\widetilde{f}} \qquad (k = 1, 2, \ldots, s), \tag{68}$$

$$\left(J + \frac{d}{dt}\right)\overset{\nu}{z} = \overset{\nu}{\widetilde{f}}, \tag{69}$$

где

$$z = \left\| \begin{matrix} \overset{1}{z} \\ \overset{2}{z} \\ \cdot \\ \cdot \\ \cdot \\ \overset{\nu}{z} \end{matrix} \right\|, \quad \widetilde{f} = \left\| \begin{matrix} \overset{1}{\widetilde{f}} \\ \overset{2}{\widetilde{f}} \\ \cdot \\ \cdot \\ \cdot \\ \overset{\nu}{\widetilde{f}} \end{matrix} \right\|, \tag{70}$$

$$\overset{1}{z} = (z_1, \ldots, z_g), \quad \overset{1}{\widetilde{f}} = (\widetilde{f}_1, \ldots, \widetilde{f}_h), \quad \overset{2}{z} = (z_{g+1}, \ldots), \quad \overset{2}{\widetilde{f}} = (\widetilde{f}_{h+1}, \ldots) \text{ и т. д.}, \tag{71}$$

$$\Lambda\left(\frac{d}{dt}\right) = A + B\frac{d}{dt}, \text{ если } \Lambda(\lambda) \equiv A + \lambda B. \tag{72}$$

Таким образом, интегрирование системы (59) в самом общем случае сведено к интегрированию частных систем (65)—(69) такого же типа. В этих системах пучок матриц $A+\lambda B$ имеет соответственно вид 0, L_ε, L'_η, $N^{(u)}$, $J+\lambda E$.

1) Для того чтобы система (65) не была противоречивой, необходимо и достаточно, чтобы

$$\overset{1}{\widetilde{f}} \equiv 0,$$

т. е.

$$\widetilde{f}_1 \equiv 0, \quad \ldots, \quad \widetilde{f}_h \equiv 0. \tag{73}$$

В этом случае в качестве неизвестных функций $z_1, z_2, \ldots, z_g$, составляющих столбец $\overset{1}{z}$, могут быть взяты произвольные функции от t.

2) Система (66) представляет собой систему вида

$$L_\varepsilon\left(\frac{d}{dt}\right) z = \widetilde{f} \tag{74}$$

или в подробной записи

$$\frac{dz_1}{dt}+z_2=\widetilde{f}_1(t), \quad \frac{dz_2}{dt}+z_3=\widetilde{f}_2(t), \quad \ldots, \quad \frac{dz_\varepsilon}{dt}+z_{\varepsilon+1}=\widetilde{f}_\varepsilon(t)\,^{1}). \tag{75}$$

Такая система всегда совместна. Если в качестве $z_{\varepsilon+1}(t)$ взять произвольную функцию от t, то последовательными квадратурами из (75) определятся все остальные неизвестные функции $z_\varepsilon, z_{\varepsilon-1}, \ldots, z_1$.

3) Система (67) представляет собой систему вида

$$L'_\eta\left(\frac{d}{dt}\right) z = \widetilde{f} \tag{76}$$

или в подробной записи

$$\frac{dz_1}{dt}=\widetilde{f}_1(t), \quad \frac{dz_2}{dt}+z_1=\widetilde{f}_2(t), \quad \ldots, \quad \frac{dz_\eta}{dt}+z_{\eta-1}=\widetilde{f}_\eta(t), \quad \widetilde{z}_\eta=\widetilde{f}_{\eta+1}(t)\,^{2}). \tag{77}$$

Из всех уравнений (77), кроме первого, мы однозначно определяем z_η, $z_{\eta-1}, \ldots, z_1$:

$$z_\eta=\widetilde{f}_{\eta+1}, \quad z_{\eta-1}=\widetilde{f}_\eta-\frac{d\widetilde{f}_{\eta+1}}{dt}, \quad \ldots, \quad z_1=$$

$$=\widetilde{f}_2-\frac{d\widetilde{f}_3}{dt}+\ldots+(-1)^{\eta-1}\frac{d^{\eta-1}\widetilde{f}_{\eta+1}}{dt^{\eta-1}}. \tag{78}$$

Подставляя полученное выражение для z_1 в первое уравнение, получаем условие совместности:

$$\widetilde{f}_1-\frac{d\widetilde{f}_2}{dt}+\frac{d^2\widetilde{f}_3}{dt^2}-\ldots+(-1)^\eta\frac{d^\eta\widetilde{f}_{\eta+1}}{dt^\eta}=0. \tag{79}$$

---

[1]) Мы изменили индексы при z и $\widetilde{f}$ для упрощения обозначений. Для того чтобы от системы (75) вернуться к системе (66), нужно ε заменить на ε_i и к каждому индексу при z прибавить $g+\varepsilon_{g+1}+\ldots+\varepsilon_{g+i-1}+i-1$, а к каждому индексу при $\widetilde{f}$ следует прибавить $h+\varepsilon_{g+1}+\ldots+\varepsilon_{g+i-1}$.

[2]) Здесь, как и в предыдущем случае, мы изменили обозначения. См. предыдущую сноску.

4) Система (68) представляет собой систему вида

$$N^{(u)}\left(\frac{d}{dt}\right)z=\widetilde{f} \tag{80}$$

или в подробной записи

$$\frac{dz_2}{dt}+z_1=\widetilde{f}_1,\quad \frac{dz_3}{dt}+z_2=\widetilde{f}_2,\quad \ldots,\quad \frac{dz_u}{dt}+z_{u-1}=\widetilde{f}_{u-1},\quad z_u=\widetilde{f}_u. \tag{81}$$

Отсюда последовательно однозначно определяем решение

$$z_u=\widetilde{f}_u,\quad z_{u-1}=\widetilde{f}_{u-1}-\frac{d\widetilde{f}_u}{dt},\quad \ldots,\quad z_1=\widetilde{f}_1-\frac{d\widetilde{f}_2}{dt}+\frac{d^2f_3}{dt}-\ldots+(-1)^{u-1}\frac{d^{u-1}f_u}{dt^{u-1}}. \tag{82}$$

5) Система (69) представляет собой систему вида

$$Jz+\frac{dz}{dt}=\widetilde{f}. \tag{83}$$

Как было показано в главе V, § 5, общее решение такой системы имеет вид

$$z=e^{-Jt}z_0+\int_0^t e^{-J(t-\tau)}f(\tau)\,d\tau; \tag{84}$$

здесь z_0 — столбец с произвольными элементами (начальными значениями неизвестных функций при $t=0$).

Обратный переход от системы (61) к системе (59) осуществляется формулами (60) и (62), согласно которым каждая из функций $x_1, \ldots, x_n$ является линейной комбинацией функций $z_1, \ldots, z_n$, а каждая из функций $\widetilde{f}_1(t), \ldots, \widetilde{f}_m(t)$ линейно (с постоянными коэффициентами) выражается через функции $f_1(t), \ldots, f_m(t)$.

Проведенный анализ показывает, что *для совместности системы* (58) *в общем случае должны выполняться некоторые определенные линейные конечные и дифференциальные зависимости (с постоянными коэффициентами) между правыми частями уравнений.*

Если эти условия выполнены, то общее решение системы содержит (в общем случае) линейно как произвольные постоянные, так и произвольные функции.

Характер условий совместности и характер решений (в частности количество произвольных постоянных и произвольных функций) определяются минимальными индексами и элементарными делителями пучка $A+\lambda B$, *поскольку от этих индексов и делителей зависит каноническая форма системы дифференциальных уравнений* (65) — (69).

ГЛАВА XIII

МАТРИЦЫ С НЕОТРИЦАТЕЛЬНЫМИ ЭЛЕМЕНТАМИ

В этой главе изучаются свойства вещественных матриц с неотрицательными элементами. Эти матрицы находят себе существенное применение в теории вероятностей при исследовании цепей Маркова («стохастические матрицы», см. [25]) и в теории малых колебаний упругих систем («осцилляционные матрицы», см. [7]).

§ 1. Общие свойства

Начнем с определения.

Определение 1. Прямоугольную матрицу A с вещественными элементами

$$A = \| a_{ik} \| \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

мы будем называть *неотрицательной* (обозначение: $A \geqslant 0$) или *положительной* (обозначение: $A > 0$), если все элементы матрицы A неотрицательны (соответственно положительны): $a_{ik} \geqslant 0$ (соответственно > 0).

Определение 2. Квадратная матрица $A = \| a_{ik} \|_1^n$ называется *разложимой*, если при некотором разбиении всех индексов $1, 2, \ldots, n$ на две дополнительные системы (без общих индексов) $i_1, i_2, \ldots, i_\mu$; $k_1, k_2, \ldots, k_\nu$ $(\mu + \nu = n)$

$$a_{i_\alpha k_\beta} = 0 \qquad (\alpha = 1, 2, \ldots, \mu;\ \beta = 1, 2, \ldots, \nu).$$

В противном случае матрицу A будем называть *неразложимой*.

Под *перестановкой рядов* в квадратной матрице $A = \| a_{ik} \|_1^n$ мы будем понимать соединение перестановки строк с такой же перестановкой столбцов матрицы A.

Определение разложимой и неразложимой матриц может быть сформулировано так:

Определение 2′. Матрица $A = \| a_{ik} \|_1^n$ называется *разложимой*, если перестановкой рядов она может быть приведена к виду

$$\widetilde{A} = \begin{Vmatrix} B & 0 \\ C & D \end{Vmatrix},$$

где B и D — квадратные матрицы. В противном случае матрица A называется *неразложимой*.

Пусть матрица $A = \| a_{ik} \|_1^n$ соответствует линейному оператору $\boldsymbol{A}$ в n-мерном векторном пространстве $\boldsymbol{R}$ с базисом $e_1, e_2, \ldots, e_n$. Пере-

становке рядов в матрице A соответствует перенумерация базисных векторов, т. е. переход от базиса $e_1, e_2, \ldots, e_n$ к новому базису $e'_1 = e_{j_1}$, $e'_2 = e_{j_2}, \ldots, e'_n = e_{j_n}$, где $(j_1, j_2, \ldots, j_n)$ — некоторая перестановка индексов $1, 2, \ldots, n$. При этом матрица A переходит в подобную ей матрицу $\widetilde{A} = T^{-1}AT$ (в каждой строке и в каждом столбце преобразующей матрицы T один элемент равен единице, а все остальные элементы равны нулю).

Под ν-мерным *координатным подпространством* в $\boldsymbol{R}$ мы будем понимать любое подпространство в $\boldsymbol{R}$ с базисом $e_{k_1}, e_{k_2}, \ldots, e_{k_\nu}$ $(1 \leqslant k_1 < k_2 < \ldots < k_\nu \leqslant n)$. С каждым базисом $e_1, e_2, \ldots, e_n$ пространства $\boldsymbol{R}$ связаны C_n^ν ν-мерных координатных подпространств. Определение разложимой матрицы может быть еще дано в следующей форме:

Определение 2″. Матрица $A = \| a_{ik} \|_1^n$ называется *разложимой* в том и только в том случае, если соответствующий этой матрице оператор $\boldsymbol{A}$ имеет ν-мерное инвариантное координатное подпространство с $\nu < n$.

Докажем следующую лемму:

Лемма 1. *Если $A \geqslant 0$ — неразложимая матрица и n — порядок матрицы A, то*

$$(E + A)^{n-1} > 0. \tag{1}$$

Доказательство. Для доказательства леммы достаточно показать, что для любого вектора (столбца)[1] $y \geqslant 0$ $(y \neq 0)$ имеет место неравенство

$$(E + A)^{n-1} y > 0.$$

Это же неравенство будет установлено, если мы только покажем, *что при условии $y \geqslant 0$ и $y \neq 0$ вектор $z = (E + A)y$ всегда имеет меньше нулевых координат, нежели вектор y.* Допустим противное. Тогда векторы y и z имеют одни и те же нулевые координаты[2]. Не нарушая общности рассуждений, можно принять, что столбцы y и z имеют вид[3]

$$y = \begin{Vmatrix} u \\ 0 \end{Vmatrix}, \quad z = \begin{Vmatrix} v \\ 0 \end{Vmatrix} \qquad (u > 0, v > 0),$$

где столбцы u и v имеют один и тот же размер.

Полагая соответственно

$$A = \begin{Vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Vmatrix},$$

будем иметь:

$$\begin{Vmatrix} u \\ 0 \end{Vmatrix} + \begin{Vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Vmatrix} \begin{Vmatrix} u \\ 0 \end{Vmatrix} = \begin{Vmatrix} v \\ 0 \end{Vmatrix},$$

откуда

$$A_{21} u = 0.$$

---

[1]) Здесь и далее в этой главе мы под вектором будем понимать столбец из n чисел. Этим самым мы как бы отождествляем вектор со столбцом его координат в том базисе, в котором данная матрица $A = \| a_{ik} \|_1^n$ задает некоторый линейный оператор.

[2]) Здесь мы исходим из того, что $z = y + Ay$ и $Ay \geqslant 0$; поэтому положительным координатам вектора y соответствуют положительные координаты вектора z.

[3]) К такому виду можно привести столбцы y и z при помощи некоторой (одной и той же для y и z) перенумерации координат.

Поскольку $u > 0$, то отсюда вытекает:

$$A_{21} = 0.$$

Это равенство противоречит неразложимости матрицы A.

Таким образом, лемма доказана.

Введем в рассмотрение степени матрицы A:

$$A^q = \| a_{ik}^{(q)} \|_1^n \quad (q = 1, 2, \ldots).$$

Тогда из леммы вытекает

Следствие. *Если $A \geqslant 0$ — неразложимая матрица, то для любой пары индексов $(1 \leqslant) i, k (\leqslant n)$ существует целое положительное число q такое, что*

$$a_{ik}^{(q)} > 0. \tag{2}$$

При этом число q всегда можно выбрать в пределах

$$\left.\begin{aligned} q &\leqslant m-1, \text{ если } i \neq k, \\ q &\leqslant m, \text{ если } i = k, \end{aligned}\right\} \tag{3}$$

где m — степень минимального многочлена $\psi(\lambda)$ матрицы A.

Действительно, обозначим через $r(\lambda)$ остаток от деления $(\lambda + 1)^{n-1}$ на $\psi(\lambda)$. Тогда в силу (1) $r(A) > 0$. Так как степень $r(\lambda)$ меньше m, то из полученного неравенства вытекает, что при любых $(1 \leqslant) i, k (\leqslant n)$ по крайней мере одно из неотрицательных чисел

$$\delta_{ik}, \quad a_{ik}, \quad a_{ik}^{(2)}, \quad \ldots, \quad a_{ik}^{(m-1)}$$

не равно нулю. Поскольку $\delta_{ik} = 0$ при $i \neq k$, то отсюда следует первое из соотношений (3). Второе соотношение (для $i = k$) получается аналогично, если неравенство $r(A) > 0$ заменить неравенством $Ar(A) > 0$[1]).

Замечание. Это следствие из леммы показывает, что в неравенстве (1) можно заменить число $n - 1$ числом $m - 1$, где m — степень минимального многочлена матрицы A.

§ 2. Спектральные свойства неразложимых неотрицательных матриц

1. Перрон в 1907 г. установил замечательные свойства спектра (т. е. совокупности характеристических чисел и собственных векторов) положительных матриц [2]).

Теорема 1 (Перрона). *Положительная матрица $A = \| a_{ik} \|_1^n$ всегда имеет вещественное и притом положительное характеристическое число r, которое является простым корнем характеристического уравнения и превосходит модули всех других характеристических чисел. Этому «максимальному» характеристическому числу r соответствует собственный вектор $z = (z_1, z_2, \ldots, z_n)$ матрицы A с положительными координатами $z_i > 0$ $(i = 1, 2, \ldots, n)$* [3]).

---

[1]) Произведение неразложимой неотрицательной матрицы на положительную всегда представляет собой положительную матрицу.

[2]) См. [230 a, b], а также [7], стр. 100.

[3]) Поскольку число r является простым характеристическим числом, то собственный вектор z, отвечающий этому числу, определяется с точностью до скалярного множителя. По теореме Перрона все координаты вектора z отличны от нуля, вещественны и одного знака. Умножением вектора z на ± 1 можно сделать все его координаты положительными. В этом последнем случае вектор (столбец) $z = (z_1, z_2, \ldots, z_n)$ будем называть *положительным* (сравни с определением 1).

Положительная матрица является частным видом неразложимой неотрицательной матрицы. Фробениус [1]) обобщил теорему Перрона, исследовав спектральные свойства неразложимых неотрицательных матриц.

Теорема 2 (Фробениуса). *Неразложимая неотрицательная матрица* $A = \| a_{ik} \|_1^n$ *всегда имеет положительное характеристическое число* r, *которое является простым корнем характеристического уравнения. Модули всех других характеристических чисел не превосходят числа* r. *«Максимальному» характеристическому числу* r *соответствует собственный вектор* z *с положительными координатами.*

Если при этом A *имеет* h *характеристических чисел* $\lambda_0 = r, \lambda_1, \ldots, \lambda_{n-1}$, *по модулю равных* r, *то эти числа все различны между собой и являются корнями уравнения*

$$\lambda^h - r^h = 0, \tag{4}$$

и вообще совокупность всех характеристических чисел $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$ *матрицы* $A = \| a_{ik} \|_1^n$, *рассматриваемая как система точек в комплексной* λ*-плоскости, переходит сама в себя при повороте этой плоскости на угол* $\frac{2\pi}{h}$. *При* $h > 1$ *перестановкой рядов можно матрицу* A *привести к следующему «циклическому» виду:*

$$A = \begin{Vmatrix} 0 & A_{12} & 0 & \ldots & 0 \\ 0 & 0 & A_{23} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & A_{h-1,\,h} \\ A_{h1} & 0 & 0 & \ldots & 0 \end{Vmatrix}, \tag{5}$$

где вдоль диагонали стоят квадратные блоки.

Поскольку теорема Перрона следует как частный случай из теоремы Фробениуса, то мы будем доказывать только последнюю [2]). Предварительно условимся относительно некоторых обозначений.

Мы будем писать:

$$C \leqslant D \text{ или } D \geqslant C,$$

где C и D — вещественные прямоугольные матрицы одинаковых размеров $m \times n$

$$C = \| c_{ik} \|, \quad D = \| d_{ik} \| \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n),$$

в том и только в том случае, когда

$$c_{ik} \leqslant d_{ik} \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n). \tag{6}$$

Если во всех неравенствах (6) можно отбросить знак равенства, то мы будем писать:

$$C < D \text{ или } D > C.$$

В частности, $C \geqslant 0$ ($C > 0$) обозначает, что все элементы матрицы C неотрицательны (соответственно положительны).

Кроме того, через C^+ мы будем обозначать $\operatorname{mod} C$, т. е. матрицу, которая получается, если все элементы матрицы C заменить их модулями.

---

[1]) См. [182 d, e].

[2]) Непосредственное доказательство теоремы Перрона см. [7], стр. 100 и далее.

2. Доказательство теоремы Фробениуса[1]. Для фиксированного вещественного вектора $x = (x_1, x_2, \ldots, x_n) \geqslant 0$ $(x \neq 0)$ полагаем:

$$r_x = \min_{1 \leqslant i \leqslant n} \frac{(Ax)_i}{x_i} \quad \left((Ax)_i = \sum_{k=1}^{n} a_{ik} x_k;\ i = 1, 2, \ldots, n \right);$$

при этом при определении минимума исключаются те значения индекса i, для которых $x_i = 0$. Очевидно, $r_x \geqslant 0$ и r_x есть наибольшее из вещественных чисел ϱ, для которых имеет место неравенство

$$\varrho x \leqslant Ax.$$

Мы докажем, что функция r_x достигает своего наибольшего значения r на некотором векторе $z \geqslant 0$:

$$r = r_z = \max_{(x \geqslant 0)} r_x = \max_{(x \geqslant 0)} \min_{1 \leqslant i \leqslant n} \frac{(Ax)_i}{x_i}. \tag{7}$$

Из определения r_x следует, что при умножении вектора $x \geqslant 0$ $(x \neq 0)$ на число $\lambda > 0$ величина r_x не меняется. Поэтому при разыскании максимума функции r_x можно ограничиться замкнутым множеством M, состоящим из векторов x, для которых

$$x \geqslant 0, \quad (xx) \equiv \sum_{i=1}^{n} x_i^2 = 1.$$

Если бы функция r_x была непрерывна на множестве M, то существование максимума было бы обеспечено. Однако функция r_x непрерывна в любой «точке» $x > 0$, но в граничных точках множества M, где одна из координат обращается в нуль, может испытывать разрывы. Поэтому мы вместо множества M введем множество N, состоящее из всех векторов y вида

$$y = (E + A)^{n-1} x \qquad (x \in M).$$

Множество N, как и M, ограничено и замкнуто и состоит согласно лемме 1 из положительных векторов.

Кроме того, помножая обе части неравенства

$$r_x x \leqslant Ax$$

на $(E + A)^{n-1} > 0$, получаем:

$$r_x y \leqslant Ay \quad [y = (E + A)^{n-1} x].$$

Отсюда, исходя из определения r_y, находим:

$$x \leqslant r_y.$$

Поэтому при разыскании максимума r_x мы можем множество M заменить множеством N, состоящим только из положительных векторов. На ограниченном замкнутом множестве N функция r_x непрерывна и поэтому достигает своего наибольшего значения на некотором векторе $z > 0$.

Любой вектор $z \geqslant 0$, для которого

$$r_z = r, \tag{8}$$

будем называть экстремальным.

---

[1]) Приведенное здесь доказательство принадлежит Виландту [260].

Докажем теперь, что: 1) *определенное равенством* (7) *число* r *положительно и является характеристическим числом матрицы* A *и* 2) *любой экстремальный вектор* z *положителен и является собственным вектором матрицы* A *для характеристического числа* r, *т. е.*

$$r>0, \quad z>0, \quad Az=rz. \tag{9}$$

Действительно, если $u=(\underbrace{1, 1, \ldots, 1}_{n})$, то $r_u=\min\limits_{1\leqslant i\leqslant n}\sum\limits_{k=1}^{n} a_{ik}$. Но тогда $r_u>0$, поскольку ни одна строка неразложимой матрицы не может состоять сплошь из нулей. Следовательно, и $r>0$, так как $r\geqslant r_u$. Далее, пусть

$$x=(E+A)^{n-1}z. \tag{10}$$

Тогда согласно лемме 1 $x>0$. Допустим теперь, что $Az-rz\neq 0$. Тогда в силу (1), (8) и (10) получаем последовательно:

$$Az-rz\geqslant 0, \quad (E+A)^{n-1}(Az-rz)>0, \quad Ax-rx>0.$$

Последнее же неравенство противоречит определению числа r, так как из этого неравенства следовало бы $Ax-(r+\varepsilon)x>0$ при достаточно малом $\varepsilon>0$, т. е. $r_x\geqslant r+\varepsilon>r$. Следовательно, $Az=rz$. Но тогда

$$0<x=(E+A)^{n-1}z=(1+r)^{n-1}z,$$

откуда вытекает: $z>0$.

Покажем теперь, что *модули всех характеристических чисел не превосходят* r. Пусть

$$Ay=\alpha y \ (y\neq 0). \tag{11}$$

Переходя к модулям в левой и правой частях равенства (11), получим [1]):

$$|\alpha|\,y^{+}\leqslant Ay^{+}. \tag{12}$$

Откуда

$$|\alpha|\leqslant r_{y^{+}}\leqslant r.$$

Допустим, что характеристическому числу r соответствует какой-либо собственный вектор y:

$$Ay=ry \quad (y\neq 0).$$

Тогда, полагая в (11) и (12) $\alpha=r$, заключаем, что y^{+} — экстремальный вектор и, следовательно, $y^{+}>0$, т. е. $y=(y_1, y_2, \ldots, y_n)$, где $y_i\neq 0$ $(i=1, 2, \ldots, n)$. Отсюда следует, что *характеристическому числу* r *соответствует только одно собственное направление*, так как при наличии двух линейно независимых собственных векторов z и z_1 мы смогли бы подобрать числа c и d так, чтобы собственный вектор $y=cz+dz_1$ имел нулевую координату, а это по доказанному невозможно.

Введем в рассмотрение присоединенную матрицу для характеристической матрицы $\lambda E-A$:

$$B(\lambda)=\|B_{ik}(\lambda)\|_1^n=\Delta(\lambda)(\lambda E-A)^{-1},$$

где $\Delta(\lambda)$ — характеристический многочлен матрицы A, а $B_{ik}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ki}-a_{ki}$ в определителе $\Delta(\lambda)$. Из того, что характеристическому числу r соответствует (с точностью до множителя) только один собственный вектор $z=(z_1, z_2, \ldots, z_n)$, где

[1]) Относительно обозначения y^{+} см. стр. 355.

$z_1 > 0$, $z_2 > 0$, . . ., $z_n > 0$, вытекает, что $B(r) \neq 0$ и что в любом ненулевом столбце матрицы $B(r)$ все элементы отличны от нуля и одного знака. То же положение имеет место и для строк матрицы $B(r)$, поскольку в предыдущих рассуждениях можно матрицу A заменить транспонированной матрицей A'. Из отмеченных свойств строк и столбцов матрицы A вытекает, что все $B_{ik}(r)$ $(i, k = 1, 2, \ldots, n)$ отличны от нуля и одного знака σ. Поэтому

$$\sigma\Delta'(r) = \sigma \sum_{i=1}^{n} B_{ii}(r) > 0,$$

т. е. $\Delta'(r) \neq 0$ и r — *простой корень характеристического уравнения* $\Delta(\lambda) = 0$.

Так как r — максимальный корень многочлена $\Delta(\lambda) = \lambda^n + \ldots$, то $\Delta(\lambda)$ возрастает при $\lambda = r$. Поэтому $\Delta'(r) > 0$ и $\sigma = 1$, т. е.

$$B_{ik}(r) > 0 \quad (i, k = 1, 2, \ldots, n). \tag{13}$$

3. Переходя к доказательству второй части теоремы Фробениуса, мы воспользуемся следующей интересной леммой [1]):

Лемма 2. *Если* $A = \| a_{ik} \|_1^n$ *и* $C = \| c_{ik} \|_1^n$ — *две квадратные матрицы одного и того же порядка* n, *причем* A — *неразложимая матрица и*

$$C^+ \leqslant A \text{ [2])}, \tag{14}$$

то между любым характеристическим числом γ *матрицы* C *и максимальным характеристическим числом* r *матрицы* A *имеет место неравенство*

$$|\gamma| \leqslant r. \tag{15}$$

В соотношении (15) *знак равенства может иметь место в том и только в том случае, когда*

$$C = e^{i\varphi} DAD^{-1}, \tag{16}$$

где $e^{i\varphi} = \frac{\gamma}{r}$, а D — *диагональная матрица, у которой диагональные элементы по модулю равны единице* $(D^+ = E)$.

Доказательство леммы. Обозначим через y собственный вектор матрицы C, отвечающей характеристическому числу γ:

$$Cy = \gamma y \quad (\gamma \neq 0). \tag{17}$$

Из (14) и (17) находим:

$$|\gamma|\, y^+ \leqslant C^+ y^+ \leqslant A y^+. \tag{18}$$

Поэтому

$$|\gamma| \leqslant r_{y^+} \leqslant r.$$

Разберем теперь подробно случай $|\gamma| = r$. В этом случае из (18) следует, что y^+ — экстремальный вектор для матрицы A и, следовательно, $y^+ > 0$ и y^+ — собственный вектор матрицы A для характеристического числа r. Поэтому соотношение (18) принимает вид

$$Ay^+ = C^+ y^+ = r y^+, \quad y^+ > 0. \tag{19}$$

---

[1]) См. [260].

[2]) C — комплексная матрица, а $A \geqslant 0$.

Отсюда в силу (14)

$$C^+ = A. \tag{20}$$

Пусть $y = (y_1, y_2, \ldots, y_n)$, где

$$y_j = |y_j| e^{i\psi_j} \quad (j = 1, 2, \ldots, n).$$

Определим диагональную матрицу D равенством

$$D = \{e^{i\psi_1}, e^{i\psi_2}, \ldots, e^{i\psi_n}\}.$$

Тогда

$$y = Dy^+.$$

Подставляя это выражение для y в (17) и полагая там $\gamma = re^{i\varphi}$, легко найдем:

$$Fy^+ = ry^+, \tag{21}$$

где

$$F = e^{-i\varphi} D^{-1} C D. \tag{22}$$

Сопоставляя (19) с (21), получим:

$$Fy^+ = C^+ y^+ = Ay^+. \tag{23}$$

Но в силу (22) и (20)

$$F^+ = C^+ = A.$$

Поэтому из (23) находим:

$$Fy^+ = F^+ y^+.$$

Поскольку $y^+ > 0$, то это равенство может иметь место лишь тогда, когда

$$F = F^+,$$

т. е.

$$e^{-i\varphi} D^{-1} C D = A.$$

Отсюда

$$C = e^{i\varphi} D A D^{-1}.$$

Лемма доказана.

4. Вернемся к теореме Фробениуса и применим доказанную лемму к неразложимой матрице $A \geqslant 0$, имеющей ровно h различных характеристических чисел с максимальным модулем r:

$$\lambda_0 = re^{i\varphi_0}, \ \lambda_1 = re^{i\varphi_1}, \ \ldots, \ \lambda_{h-1} = re^{i\varphi_{h-1}} \ (0 = \varphi_0 < \varphi_1 < \varphi_2 < \ldots < \varphi_{h-1} < 2\pi).$$

Тогда, полагая в лемме $C = A$ и $\gamma = \lambda_k$, для любого $k = 0, 1, \ldots, h-1$ будем иметь:

$$A = e^{i\varphi_k} D_k A D_k^{-1}, \tag{24}$$

где D_k — диагональная матрица с $D_k^+ = E$.

Пусть снова z — положительный собственный вектор матрицы A, соответствующий максимальному характеристическому числу r:

$$Az = rz \quad (z > 0). \tag{25}$$

Тогда, полагая

$$\overset{k}{y} = D_k z \quad (\overset{k}{y^+} = z > 0), \tag{26}$$

из (25) и (26) найдем:

$$A\overset{k}{y}=\lambda_k\overset{k}{y} \quad (\lambda_k=re^{i\varphi_k};\quad k=0,\ 1,\ \ldots,\ h-1). \tag{27}$$

Последние равенства показывают, что векторы $\overset{0}{y}, \overset{1}{y}, \ldots, \overset{h-1}{y}$, определяемые из (26), являются собственными векторами матрицы A для характеристических чисел $\lambda_0, \lambda_1, \ldots, \lambda_{h-1}$.

Из (24) следует, что не только $\lambda_0=r$, но и *каждое из характеристических чисел* $\lambda_1, \ldots, \lambda_{h-1}$ *матрицы* A *является простым*. Поэтому собственные векторы $\overset{k}{y}$, а значит, и матрицы D_k $(k=0,\ 1,\ \ldots,\ h-1)$ определяются с точностью до скалярных множителей. Для однозначного определения матриц $D_0, D_1, \ldots, D_{h-1}$ мы будем выбирать первые диагональные элементы этих матриц равными единице. Тогда $D_0=E$ и $\overset{0}{y}=z>0$.

Далее, из (24) следует:

$$A=e^{i(\varphi_j\pm\varphi_k)}D_jD_k^{\pm1}AD_k^{\mp1}D_j^{-1} \quad (j,\ k=0,\ 1,\ \ldots,\ h-1).$$

Отсюда аналогично предыдущему заключаем, что вектор

$$D_jD_k^{\pm1}z$$

есть собственный вектор матрицы A, соответствующий характеристическому числу $re^{i(\varphi_j\pm\varphi_k)}$.

Поэтому $e^{i(\varphi_j\pm\varphi_k)}$ совпадает с одним из чисел $e^{i\varphi_l}$, а матрица $D_jD_k^{\pm1}$ — с соответствующей матрицей D_l, т. е. при некоторых $(0\leqslant)\, l_1,\ l_2\,(\leqslant h-1)$

$$e^{i(\varphi_j+\varphi_k)}=e^{i\varphi_{l_1}},\quad e^{i(\varphi_j-\varphi_k)}=e^{i\varphi_{l_2}},$$

$$D_jD_k=D_{l_1},\quad D_jD_k^{-1}=D_{l_2}.$$

Таким образом, *числа* $e^{i\varphi_0}, e^{i\varphi_1}, \ldots, e^{i\varphi_{h-1}}$ *и соответствующие диагональные матрицы* $D_0, D_1, \ldots, D_{h-1}$ *образуют две изоморфные между собой мультипликативные абелевы группы.*

В каждой конечной группе, состоящей из h различных элементов, h-я степень любого элемента равна единице группы [1]). Поэтому $e^{i\varphi_0}$, $e^{i\varphi_1}, \ldots, e^{i\varphi_{h-1}}$ — корни h-й степени из единицы. Так как существует всего h различных корней из единицы и $\varphi_0=0<\varphi_1<\varphi_2<\ldots<\varphi_{h-1}<2\pi$, то

$$\varphi_k=\frac{2k\pi}{h} \quad (k=0,\ 1,\ 2,\ \ldots,\ h-1)$$

и

$$e^{i\varphi_k}=\varepsilon^k \quad (\varepsilon=e^{i\varphi_1}=e^{\frac{2\pi}{h}i} \quad k=0,\ 1.\ \ldots,\ h-1), \tag{28}$$

$$\lambda_k=r\varepsilon^k \quad (k=0,\ 1,\ \ldots,\ h-1). \tag{29}$$

Числа $\lambda_0, \lambda_1, \ldots, \lambda_{h-1}$ образуют полную систему корней уравнения (4).

В соответствии с (28) будем иметь [2]):

$$D_k=D^k \quad (D=D_1;\ k=0,\ 1,\ \ldots,\ h-1). \tag{30}$$

---

[1]) См., например, [39], стр. 324.

[2]) Здесь мы опираемся на изоморфность мультипликативных групп $e^{i\varphi_0}, e^{i\varphi_1}, \ldots, e^{i\varphi_{h-1}}$ и $D_0, D_1, \ldots, D_{h-1}$.

Теперь равенство (24) нам дает (при $k=1$):

$$A=e^{i\frac{2\pi}{h}}DAD^{-1}. \tag{31}$$

Отсюда следует, что матрица A при умножении на $e^{\frac{2\pi i}{h}}$ переходит в подобную матрицу, и, следовательно, *вся система n характеристических чисел матрицы A при умножении на $e^{i\frac{2\pi}{h}}$ переходит в самое себя*[1]).

Далее,

$$D^h=E,$$

поэтому все диагональные элементы в D — корни h-й степени из единицы. Перестановкой рядов в A (и соответственно в D) можно добиться того, чтобы матрица D имела следующий квазидиагональный вид:

$$D=\{\eta_0E_0,\ \eta_1E_1,\ \ldots,\ \eta_{s-1}E_{s-1}\}, \tag{32}$$

где $E_0,\ E_1,\ \ldots,\ E_{s-1}$ — единичные матрицы, а

$$\eta_p=e^{i\psi_p},\ \psi_p=n_p\frac{2\pi}{h}$$

(n_p — целое; $p=0,\ 1,\ \ldots,\ s-1;\ 0=n_0<n_1<\ldots<n_{s-1}<h$).

Очевидно, что $s\leqslant h$.

Записывая A в блочном виде [в соответствии с (32)]

$$A=\left\|\begin{array}{cccc} A_{11} & A_{12} & \ldots & A_{1s}\\ A_{21} & A_{22} & \ldots & A_{2s}\\ \ldots & \ldots & \ldots & \ldots\\ A_{s1} & A_{s2} & \ldots & A_{ss}\end{array}\right\|, \tag{33}$$

мы равенство (31) заменим системой равенств

$$\varepsilon A_{pq}=\frac{\eta_{q-1}}{\eta_{p-1}}A_{pq}\quad (p,\ q=1,\ 2,\ \ldots,\ s;\ \varepsilon=e^{\frac{2\pi}{h}i}). \tag{34}$$

Отсюда при любых p и q либо $\frac{\eta_{q-1}}{\eta_{p-1}}=\varepsilon$, либо $A_{pq}=0$.

Возьмем $p=1$. Поскольку все матрицы $A_{12},\ A_{13},\ \ldots,\ A_{1s}$ не могут одновременно обратиться в нуль, то одно из чисел $\frac{\eta_1}{\eta_0},\ \frac{\eta_2}{\eta_0},\ \ldots,\ \frac{\eta_{s-1}}{\eta_0}$ ($\eta_0=1$) должно равняться ε. Это возможно лишь при $n_1=1$. Тогда $\frac{\eta_1}{\eta_0}=\varepsilon$ и $A_{11}=A_{13}=\ldots=A_{1s}=0$. Полагая в (34) $p=2$, аналогично найдем, что $n_2=2$ и что $A_{21}=A_{22}=A_{24}=\ldots=A_{2s}=0$ и т. д. В результате получим:

$$A=\left\|\begin{array}{ccccc} 0 & A_{12} & 0 & \ldots & 0\\ 0 & 0 & A_{23} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & A_{s-1,\,s}\\ A_{s1} & A_{s2} & A_{s3} & \ldots & A_{ss}\end{array}\right\|.$$

[1]) Число h есть наибольшее целое число, обладающее этим свойством, поскольку матрица A имеет ровно h характеристических чисел с максимальным модулем r. Кроме того, из (31) вытекает, что все характеристические числа матрицы разбиваются на системы (по h чисел в каждой) вида $\mu_0,\ \mu_0\varepsilon,\ \ldots,\ \mu_0\varepsilon^{h-1}$ и что в пределах каждой такой системы любым двум характеристическим числам отвечают элементарные делители соответственно одинаковых степеней. Одну из таких систем образуют корни уравнения (4): $\lambda_0,\ \lambda_1,\ \ldots,\ \lambda_{h-1}$.

При этом $n_1=1$, $n_2=2$, ..., $n_{s-1}=s-1$. Но тогда при $p=s$ в правых частях равенства (34) стоят множители

$$\frac{\eta_{q-1}}{\eta_{s-1}}=e^{(q-s)\frac{2\pi}{h}i} \qquad (q=1,\ 2,\ \ldots,\ s).$$

Одно из этих чисел должно равняться $\varepsilon=e^{\frac{2\pi}{h}i}$. Это возможно лишь, когда $s=h$ и $q=1$ и, следовательно, $A_{s2}=\ldots=A_{ss}=0$.

Таким образом,

$$D=\{E_0,\ \varepsilon E_1,\ \varepsilon^2 E_2,\ \ldots,\ \varepsilon^{h-1}E_{h-1}\}$$

и матрица A имеет вид (5).

Теорема Фробениуса доказана полностью.

5. Сделаем несколько замечаний к теореме Фробениуса.

Замечание 1. При доказательстве теоремы Фробениуса мы попутно установили, что для неразложимой матрицы $A\geqslant 0$, имеющей максимальное характеристическое число r, присоединенная матрица $B(\lambda)$ при $\lambda=r$ положительна:

$$B(r)>0, \tag{35}$$

т. е.

$$B_{ik}(r)>0 \quad (i,\ k=1,\ 2,\ \ldots,\ n), \tag{35'}$$

где $B_{ik}(r)$ — алгебраическое дополнение элемента $r\delta_{ki}-a_{ki}$ в определителе $|rE-A|$.

Рассмотрим теперь приведенную присоединенную матрицу (см. гл. IV, § 6)

$$C(\lambda)=\frac{B(\lambda)}{D_{n-1}(\lambda)},$$

где $D_{n-1}(\lambda)$ — наибольший общий делитель (со старшим коэффициентом единица) всех многочленов $B_{ik}(\lambda)$ $(i,\ k=1,\ 2,\ \ldots,\ n)$. При этом из (35') следует, что $D_{n-1}(r)\neq 0$. Все корни многочлена $D_{n-1}(\lambda)$ являются характеристическими числами[1]), отличными от r. Поэтому все корни $D_{n-1}(\lambda)$ либо комплексны, либо вещественны, но меньше r. Отсюда $D_{n-1}(r)>0$, что в соединении с (35) дает:

$$C(r)=\frac{B(r)}{D_{n-1}(r)}>0\ ^{2}). \tag{36}$$

Замечание 2. Неравенства (35') позволяют определить границы для величины максимального характеристического числа r.

Введем обозначения

$$s_i=\sum_{k=1}^{n}a_{ik} \quad (i=1,\ 2,\ \ldots,\ n), \quad s=\min_{1\leqslant i\leqslant n}s_i, \quad S=\max_{1\leqslant i\leqslant n}s_i.$$

Тогда *для неразложимой матрицы* $A\geqslant 0$

$$s\leqslant r\leqslant S, \tag{37}$$

причем знак равенства слева или справа от r *имеет место лишь при*

---

[1]) $D_{n-1}(\lambda)$ является делителем характеристического многочлена $D_n(\lambda)\equiv|\lambda E-A|$.

[2]) В следующем параграфе будет доказано, что для неразложимой матрицы $B(\lambda)>0$, $C(\lambda)>0$ при любом вещественном $\lambda\geqslant r$.

$s=S$, т. е. когда все «строчные суммы» $s_1, s_2, \ldots, s_n$ равны между собой[1]).

Действительно, прибавим к последнему столбцу характеристического определителя

$$\Delta(r)=\begin{vmatrix} r-a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & r-a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & r-a_{nn} \end{vmatrix}$$

все предыдущие столбцы и разложим после этого определитель по элементам последнего столбца. Получим:

$$\sum_{k=1}^{n}(r-s_s)\,B_{nk}(r)=0.$$

Отсюда в силу (35′) вытекает неравенство (37).

Замечание 3. *Неразложимая матрица $A\geqslant 0$ не может иметь двух линейно независимых неотрицательных собственных векторов.* Действительно, пусть помимо положительного собственного вектора $z>0$, соответствующего максимальному характеристическому числу r, матрица A имеет еще собственный вектор $y\geqslant 0$ (линейно независимый от z) для характеристического числа α:

$$Ay=\alpha y \quad (y\neq 0;\ y\geqslant 0).$$

Поскольку r — простой корень характеристического уравнения $|\lambda E-A|=0$, то

$$\alpha\neq r.$$

Обозначим через u положительный собственный вектор транспонированной матрицы A' для $\lambda=r$:

$$A'u=ru \quad (u>0).$$

Тогда[2])

$$r(y,\ u)=(y,\ A'u)=(Ay,\ u)=\alpha(yu);$$

отсюда, поскольку $\alpha\neq r$,

$$(y,\ u)=0,$$

а это невозможно при $u>0$, $y\geqslant 0$, $y\neq 0$.

Замечание 4. При доказательстве теоремы Фробениуса мы установили следующую характеристику максимального собственного числа r неразложимой матрицы $A\geqslant 0$:

$$r=\max_{(x\geqslant 0)} r_x,$$

---

[1]) Установлению для r интервала, более узкого, нежели (s, S), посвящены работы [207], [222а] и [161, IV].

[2]) Если $y=(y_1, y_2, \ldots, y_n)$ и $u=(u_1, u_2, \ldots, u_n)$, то под (y, u) мы понимаем «скалярное произведение» $y'u=\sum_{i=1}^{n} y_iu_i$. Тогда $(y, A'u)=y'A'u$ и $(Ay, u)=(Ay)'u=y'A'u$.

где r_x — наибольшее из чисел ρ, для которых имеет место неравенство $\rho x \leqslant Ax$. Другими словами, поскольку $r_x = \min\limits_{1\leqslant i\leqslant n} \frac{(Ax)_i}{x_i}$, то и

$$r = \max_{(x\geqslant 0)} \min_{1\leqslant i\leqslant n} \frac{(Ax)_i}{x_i}.$$

Совершенно аналогично можно для любого вектора $x \geqslant 0\,(x \neq 0)$ определить r^x как наименьшее из чисел σ, удовлетворяющих неравенству

$$\sigma x \geqslant Ax,$$

т. е. положить:

$$r^x = \max_{1\leqslant i\leqslant n} \frac{(Ax)_i}{x_i}.$$

При этом, если при некотором i имеют место соотношения $x_i = 0$, $(Ax)_i \neq 0$, то следует считать $r^x = +\infty$.

Совершенно так же, как и для функции r_x, доказывается, что функция r^x достигает своего наименьшего значения $\widehat{r}$ на некотором векторе $v > 0$.

Покажем, что число $\widehat{r}$, определяемое равенством

$$\widehat{r} = \min_{(x\geqslant 0)} r^x = \min_{(x\geqslant 0)} \max_{1\leqslant i\leqslant n} \frac{(Ax)_i}{x_i}, \tag{38}$$

совпадает с числом r, а вектор $v \geqslant 0$ $(v \neq 0)$, на котором достигается этот минимум, является собственным вектором матрицы A при $\lambda = r$.

Действительно,

$$\widehat{r}v - Av \geqslant 0 \quad (v \geqslant 0,\ v \neq 0).$$

Допустим, что здесь знак $\geqslant$ нельзя заменить знаком равенства. Тогда согласно лемме 1

$$(E + A)^{n-1}(\widehat{r}v - Av) > 0, \quad (E + A)^{n-1} v > 0. \tag{39}$$

Полагая

$$u = (E + A)^{n-1} v > 0,$$

будем иметь:

$$\widehat{r}u > Au$$

и, следовательно, при достаточно малом $\varepsilon > 0$

$$(\widehat{r} - \varepsilon)\, u > Au \quad (u > 0),$$

что противоречит определению числа $\widehat{r}$. Итак,

$$Av = \widehat{r}v.$$

Но тогда

$$u = (E + A)^{n-1} v = (1 + \widehat{r})^{n-1} v.$$

Поэтому из $u > 0$ следует $v > 0$.

В силу замечания 3 отсюда

$$\widehat{r} = r.$$

Таким образом, мы имеем для числа r двойную характеристику:

$$r=\max_{(x\geqslant 0)}\min_{1\leqslant i\leqslant n}\frac{(Ax)_i}{x_i}=\min_{(x\geqslant 0)}\max_{1\leqslant i\leqslant n}\frac{(Ax)_i}{x_i}, \tag{40}$$

причем доказано, что $\max\limits_{(x\geqslant 0)}$ или $\min\limits_{(x\geqslant 0)}$ достигается только на положительном собственном векторе для $\lambda=r$.

Из установленной характеристики числа r вытекают неравенства [1])

$$\min_{1\leqslant i\leqslant n}\frac{(Ax)_i}{x_i}\leqslant r\leqslant\max_{1\leqslant i\leqslant n}\frac{(Ax)_i}{x_i}\quad(x\geqslant 0,\ x\neq 0). \tag{41}$$

Замечание 5. Поскольку в (40) $\max\limits_{(x\geqslant 0)}$ и $\min\limits_{(x\geqslant 0)}$ всегда достигаются только на положительном собственном векторе неразложимой матрицы $A\geqslant 0$, то из неравенств

$$rz\leqslant Az,\quad z\geqslant 0,\ z\neq 0$$

или

$$rz\geqslant Az,\quad z\geqslant 0,\quad z\neq 0$$

всегда следует:

$$Az=rz,\quad z>0.$$

§ 3. Разложимые матрицы

1. Установленные в предыдущем параграфе спектральные свойства неразложимых неотрицательных матриц не сохраняются при переходе к разложимым матрицам. Однако, поскольку произвольная неотрицательная матрица $A\geqslant 0$ всегда может быть представлена как предел последовательности неразложимых и даже положительных матриц A_m

$$A=\lim_{m\to\infty}A_m\ (A_m>0,\ m=1,\ 2,\ \ldots), \tag{42}$$

то некоторые из спектральных свойств неразложимых матриц в ослабленной форме имеют место и для разложимых матриц.

Для произвольной неотрицательной матрицы $A=\|a_{ik}\|_1^n$ мы докажем следующую теорему:

Теорема 3. *Неотрицательная матрица $A=\|a_{ik}\|_1^n$ всегда имеет неотрицательное характеристическое число r такое, что модули всех характеристических чисел матрицы A не превосходят r. Этому «максимальному» характеристическому числу r соответствует неотрицательный собственный вектор*

$$Ay=ry\quad(y\geqslant 0,\ y\neq 0).$$

Доказательство. Пусть для матрицы A имеет место представление (42). Обозначим через $r^{(m)}$ и $y^{(m)}$ максимальное характеристическое число положительной матрицы A_m и соответствующий этому числу нормированный [2]) положительный собственный вектор:

$$A_m y^{(m)}=r^{(m)}y^{(m)}\quad[(y^{(m)}y^{(m)})=1,\ y^{(m)}>0;\ m=1,\ 2,\ \ldots]. \tag{43}$$

---

[1]) См. [167], а также [7], стр. 325 и далее.

[2]) Под нормированным вектором мы понимаем столбец $y=(y_1, y_2, \ldots, y_n)$, для которого $(yy)\equiv\sum\limits_{i=1}^{n}y_i^2=1$.

Тогда из (42) следует, что существует предел

$$\lim r^{(m)} = r,$$

где r — характеристическое число матрицы A. Из того, что $r^{(m)} > 0$ и $r^{(m)} > |\lambda_0^{(m)}|$, где $\lambda_0^{(m)}$ — любое характеристическое число матрицы A_m ($m = 1, 2, \ldots$), предельным переходом получаем:

$$r \geqslant 0, \quad r \geqslant |\lambda_0|, \tag{44}$$

где λ_0 — любое характеристическое число матрицы A. Этот же предельный переход дает нам вместо (35)

$$B(r) \geqslant 0. \tag{45}$$

Далее, из последовательности нормированных собственных векторов $y^{(m)}$ ($m = 1, 2, \ldots$) можно выделить подпоследовательность $y^{(m_p)}$ ($p = 1, 2, \ldots$), сходящуюся к некоторому нормированному (и, следовательно, не равному нулю) вектору y. Перейдем к пределу в обеих частях равенства (43), давая m последовательность значений m_p ($p = 1, 2, \ldots$). Мы получим:

$$Ay = ry \quad (y \geqslant 0, \ y \neq 0).$$

Теорема доказана.

Замечание. При предельном переходе (42) неравенства (37) сохраняются. Поэтому эти неравенства имеют место для произвольной неотрицательной матрицы. Однако условие, при котором в (37) имеет место знак равенства, для разложимой матрицы уже неверно.

2. Установим ряд важных предложений для матриц с неотрицательными элементами:

1° *Если* $A = \| a_{ik} \|_1^n$ — *неотрицательная матрица с максимальным характеристическим числом* r, *то*

$$(\lambda E - A)^{-1} \geqslant 0, \ \frac{d}{d\lambda}(\lambda E - A)^{-1} \leqslant 0 \ \textit{при} \ \lambda > r. \tag{46}$$

Действительно, при $\lambda > r \geqslant 0$ справедливо разложение

$$(\lambda E - A)^{-1} = \sum_{j=0}^{\infty} \frac{A^j}{\lambda^{j+1}} \geqslant 0 \tag{47}$$

и, следовательно,

$$\frac{d}{d\lambda}(\lambda E - A)^{-1} = -\sum_{j=0}^{\infty} \frac{(j+1)A^j}{\lambda^{j+2}} \leqslant 0. \tag{48}$$

2° *Если* $A = \| a_{ik} \|_1^n$ — *неотрицательная матрица с максимальным характеристическим числом* r, *а* $B(\lambda)$ *и* $C(\lambda)$ — *ее присоединенная и приведенная присоединенная матрицы, то*

$$B(\lambda) \geqslant 0, \quad C(\lambda) \geqslant 0 \ \textit{при} \ \lambda \geqslant r. \tag{49}$$

Поскольку $B(\lambda) = (\lambda E - A)^{-1} \Delta(\lambda)$, $C(\lambda) = (\lambda E - A)^{-1} \psi(\lambda)$ и

$$\Delta(\lambda) > 0, \ \psi(\lambda) > 0 \ \text{при} \ \lambda > r, \tag{50}$$

то из первого неравенства (46) сразу вытекают соотношения (49).

3° *Если* $A=\|a_{ik}\|_1^n$ *— неразложимая матрица с максимальным характеристическим числом* r, *то*

$$(\lambda E-A)^{-1}>0,\quad \frac{d}{d\lambda}(\lambda E-A)^{-1}<0 \quad при\ \lambda>r, \tag{51}$$

$$B(\lambda)>0,\quad C(\lambda)>0 \quad при\ \lambda\geqslant r. \tag{52}$$

Действительно, согласно следствию из леммы 1 (стр. 354), в случае неразложимой матрицы $A\geqslant 0$ в соотношениях (47) и (48) можно знак равенства отбросить. Тогда и $B(\lambda)>0$, $C(\lambda)>0$ при $\lambda>r$. Но, как было показано в § 2, для неразложимой матрицы $B(r)>0$, $C(r)>0$. Следовательно, справедливы неравенства (52).

4° *Максимальное характеристическое число* r' *любого главного минора (порядка* $<n$*) неотрицательной матрицы* $A=\|a_{ik}\|_1^n$ *не превосходит максимального характеристического числа* r *матрицы* A:

$$r'\leqslant r. \tag{53}$$

Если для главного минора $(n-1)$*-го порядка* $r'<r$, *то для характеристического определителя* $\Delta(\lambda)=|\lambda E-A|$ *имеем неравенство*

$$\Delta(\lambda)<0 \qquad \text{при } r'<\lambda<r. \tag{54}$$

Если A *— неразложимая матрица, то в* (53) *всегда знак равенства отпадает.*

Если A *— разложимая матрица, то по крайней мере для одного главного минора имеет место в* (53) *знак равенства.*

Действительно, пусть для конкретности r' — максимальное характеристическое число матрицы $A_1=\|a_{ik}\|_{i,\,k=1}^{n-1}$, имеющей характеристический многочлен $\Delta_1(\lambda)=B_{nn}(\lambda)$. Тогда $B_{nn}(r')=0$ и в случае неразложимой матрицы A согласно (52) $B_{nn}(\lambda)>0$ при $\lambda\geqslant r$. Следовательно, $r'<r$. Отсюда в случае разложимой матрицы предельным переходом получаем неравенство (53).

Пусть $r'<\lambda<r$. Тогда, разлагая определитель $\Delta(\lambda)$ по элементам последней строки и последнего столбца, получаем:

$$\Delta(\lambda)=\Delta_1(\lambda)(\lambda-a_{nn})-\sum_{i,\,k=1}^{n-1}A_{ik}^{(1)}(\lambda)\,a_{in}a_{nk}, \tag{55}$$

где $A_{ik}^{(1)}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ik}-a_{ik}$ в определителе $\Delta_1(\lambda)=B_{nn}(\lambda)$ $(i,\ k=1,\ 2,\ \ldots,\ n-1)$. Разделим обе части тождества (55) на $\Delta_1(\lambda)$:

$$\frac{\Delta(\lambda)}{\Delta_1(\lambda)}=\lambda-a_{nn}-\sum_{i,\,k=1}^{n-1}\{(\lambda E-A_1)^{-1}\}_{ik}\,a_{in}a_{nk}. \tag{56}$$

Используя второе неравенство (46) для матрицы A, замечаем, что при $\lambda>r'$ первый член $\lambda-a_{nn}$ в правой части (56) монотонно возрастает, а второй — не убывает. Следовательно, отношение $\Delta(\lambda)/\Delta_1(\lambda)$ монотонно возрастает при $\lambda>r'$. Но тогда это отношение отрицательно при $r'<\lambda<r$, поскольку $\Delta(r)=0$. Но $\Delta_1(\lambda)>0$ при $\lambda>r'$. Следовательно, имеет место неравенство (54).

Мы установили справедливость неравенства (53) для миноров $(n-1)$-го порядка. Постепенным переходом от $n-1$ к $n-2$, от $n-2$ к $n-3$ и т. д. мы докажем справедливость неравенства (53) (без знака $=$ в случае неразложимой матрицы) для главного минора любого порядка.

Если A — разложимая матрица, то перестановкой рядов она может быть представлена в виде

$$A = \begin{Vmatrix} B & 0 \\ C & D \end{Vmatrix}.$$

Тогда число r должно быть характеристическим числом одного из двух главных миноров B и D. Предложение 4° доказано.

Из 4° вытекает

5° *Если* $A \geqslant 0$ *и в характеристическом определителе*

$$\Delta(r) = \begin{vmatrix} r - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & r - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & r - a_{nn} \end{vmatrix}$$

какой-либо из главных миноров обращается в нуль (*матрица* A *разложима*!), *то обращается в нуль любой «объемлющий» главный минор и, в частности, один из главных миноров* $(n-1)$-го *порядка*

$$B_{11}(\lambda),\ B_{22}(\lambda),\ \dots,\ B_{nn}(\lambda).$$

Из 4° и 5° следует:

6° *Матрица* $A \geqslant 0$ *является разложимой в том и только в том случае, когда в одном из соотношений*

$$B_{ii}(r) \geqslant 0 \qquad (i = 1,\ 2,\ \dots,\ n)$$

имеет место знак равенства.

Из 4° вытекает также

7° *Если* r — *максимальное характеристическое число матрицы* $A \geqslant 0$, *то при любом* $\lambda > r$ *все главные миноры характеристической матрицы* $A_\lambda \equiv \lambda E - A$ *положительны*:

$$A_\lambda \begin{pmatrix} i_1\, i_2 \dots i_p \\ i_1\, i_2 \dots i_p \end{pmatrix} > 0 \quad (\lambda > r;\ 1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1,\ 2,\ \dots,\ n). \tag{57}$$

Нетрудно видеть, что, и обратно, из неравенств (57) следует: $\lambda > r$. Действительно,

$$\Delta(\lambda + \mu) = |(\lambda + \mu)E - A| = |A_\lambda + \mu E| = \sum_{k=0}^{n} S_k \mu^{n-k},$$

где S_k — сумма всех главных миноров k-го порядка характеристической матрицы $A_\lambda \equiv \lambda E - A$ $(k = 1,\ 2,\ \dots,\ n)$[1]). Поэтому, если при некотором вещественном λ все главные миноры характеристической матрицы A_λ положительны, то при любом $\mu \geqslant 0$

$$\Delta(\lambda + \mu) \neq 0,$$

т. е. всякое число $\geqslant \lambda$ не является характеристическим числом матрицы A. Следовательно,

$$r < \lambda.$$

Таким образом, неравенства (57) представляют собой необходимые и достаточные условия для того, чтобы число λ было верхней границей

[1]) См. стр. 83.

для модулей характеристических чисел матрицы A [1]). Однако не все неравенства (57) независимы между собой.

Матрица $\lambda E - A$ представляет собой матрицу с неположительными недиагональными элементами [2]). Д. М. Котелянский показал [3]), что для таких матриц, как и для симметрических матриц, положительность всех главных миноров вытекает из положительности последовательных главных миноров.

Лемма 3 (Котелянского). *Если в вещественной матрице* $G = \| g_{ik} \|_1^n$ *все недиагональные элементы отрицательны или равны нулю*

$$g_{ik} \leqslant 0 \qquad (i \neq k;\ i,\ k = 1, 2, \ldots, n), \tag{58}$$

а последовательные главные миноры положительны

$$g_{11} = G\begin{pmatrix} 1 \\ 1 \end{pmatrix} > 0, \quad G\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} > 0, \ \ldots, \ G\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} > 0, \tag{59}$$

то все главные миноры матрицы G *положительны:*

$$G\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} > 0 \quad (1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n;\ \ p = 1, 2, \ldots, n).$$

Доказательство. Будем доказывать лемму индуктивно относительно порядка матрицы n. При $n = 2$ лемма имеет место, так как из

$$g_{12} \leqslant 0 \quad g_{21} \leqslant 0, \quad g_{11} > 0, \quad g_{11}g_{22} - g_{12}g_{21} > 0$$

следует: $g_{22} > 0$. Пусть лемма справедлива для матриц порядка $< n$; докажем ее для матрицы $G = \| g_{ik} \|_1^n$. Введем в рассмотрение окаймляющие определители

$$t_{ik} = G\begin{pmatrix} 1 & i \\ 1 & k \end{pmatrix} = g_{11}g_{ik} - g_{1k}g_{i1} \qquad (i,\ k = 2, \ldots, n).$$

Из (58) и (59) следует:

$$t_{ik} \leqslant 0 \qquad (i \neq k;\ i,\ k = 2, \ldots, n).$$

С другой стороны, применяя к матрице $T = \| t_{ik} \|_2^n$ тождество Сильвестра [гл. II, равенство (30), стр. 48], получаем:

$$T\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = \\ = (g_{11})^{p-1} G\begin{pmatrix} 1 & i_1 & i_2 & \ldots & i_p \\ 1 & i_1 & i_2 & \ldots & i_p \end{pmatrix} \quad \begin{pmatrix} 2 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n, \\ p = 1, 2, \ldots, n-1 \end{pmatrix}. \tag{60}$$

Отсюда в силу (59) следует, что последовательные главные миноры матрицы $T = \| t_{ik} \|_2^n$ положительны:

$$t_{22} = T\begin{pmatrix} 2 \\ 2 \end{pmatrix} > 0, \quad T\begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} > 0, \ \ldots, \ T\begin{pmatrix} 2 & 3 & \ldots & n \\ 2 & 3 & \ldots & n \end{pmatrix} > 0.$$

Таким образом, матрица $(n-1)$-го порядка $T = \| t_{ik} \|_2^n$ удовлетворяет условиям леммы. Поэтому согласно допущению индукции все главные

[1]) См. [123].

[2]) Нетрудно видеть, что, и наоборот, всякая матрица с отрицательными или нулевыми недиагональными элементами всегда может быть представлена в виде $\lambda E - A$, где A — неотрицательная матрица, а λ — вещественное число.

[3]) См. [97в]. Эта работа содержит ряд результатов, относящихся к матрицам, у которых все недиагональные элементы одного знака.

миноры матрицы T положительны:

$$T\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (2 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1,\ 2,\ \dots\ n-1).$$

Но тогда из (60) вытекает, что положительны все главные миноры матрицы G, содержащие первую строку:

$$G\begin{pmatrix} 1 & i_1 & i_2 & \dots & i_p \\ 1 & i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \ (2 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1,\ 2,\ \dots,\ n-1). \quad (61)$$

Возьмем фиксированные индексы $(1 <) i_1 < i_2 < \dots < i_{n-2} (\leqslant n)$ и составим матрицу $(n-1)$-го порядка:

$$\| g_{\alpha\beta} \| \quad (\alpha,\ \beta = 1,\ i_1,\ i_2,\ \dots,\ i_{n-2}). \quad (62)$$

Последовательные главные миноры этой матрицы в силу (61) положительны:

$$g_{11} > 0, \quad G\begin{pmatrix} 1 & i_1 \\ 1 & i_1 \end{pmatrix} > 0, \dots, G\begin{pmatrix} 1 & i_1 & i_2 & \dots & i_{n-2} \\ 1 & i_1 & i_2 & \dots & i_{n-2} \end{pmatrix} > 0,$$

а недиагональные элементы неположительны:

$$g_{\alpha\beta} \leqslant 0 \quad (\alpha \neq \beta;\ \alpha,\ \beta = 1,\ i_1\ i_2,\ \dots,\ i_{n-2}).$$

Но порядок матрицы (62) равен $n-1$. Поэтому согласно допущению индукции все главные миноры этой матрицы положительны; в частности,

$$G\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (63)$$

$$(2 \leqslant i_1 < i_2 < \dots < i_p \leqslant n;\ p = 1,\ 2,\ \dots,\ n-2).$$

Таким образом, все миноры порядка $\leqslant n-2$ матрицы G положительны.

Поскольку в силу (63) $g_{22} > 0$, то мы можем теперь ввести в рассмотрение определители второго порядка, окаймляющие элемент g_{22} (а не g_{11}, как раньше):

$$t^*_{ik} = G\begin{pmatrix} 2 & i \\ 2 & k \end{pmatrix} \quad (i,\ k = 1,\ 3,\ \dots,\ n).$$

Оперируя с матрицей $T^* = \| t^*_{ik} \|$ так, как мы ранее оперировали с матрицей T, мы получим неравенства, аналогичные неравенствам (61):

$$G\begin{pmatrix} 2 & i_1 & \dots & i_p \\ 2 & i_1 & \dots & i_p \end{pmatrix} > 0 \quad (64)$$

$$(i_1 < i_2 < \dots < i_p;\ i_1,\ \dots,\ i_p = 1,\ 3,\ \dots,\ n;\ p = 1,\ 2,\ \dots,\ n-1).$$

Так как любой главный минор матрицы $G = \| g_{ik} \|_1^n$ либо содержит первую строку, либо содержит вторую строку, либо имеет порядок $\leqslant n-2$, то из неравенств (61), (63) и (64) следует, что все главные миноры матрицы A положительны. Лемма доказана.

Доказанная лемма позволяет в условиях (57) сохранить лишь последовательные главные миноры и сформулировать следующую теорему:

Теорема 4. *Для того чтобы вещественное число λ было больше максимального характеристического числа r матрицы $A = \| a_{ik} \|_1^n \geqslant 0$,*

$$r < \lambda,$$

необходимо и достаточно, чтобы при этом значении λ все последователь-

ные главные миноры характеристической матрицы $A_\lambda \equiv \lambda E - A$ *были положительны:*

$$\lambda - a_{11} > 0,$$

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{vmatrix} > 0, \ \ldots, \ \begin{vmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{vmatrix} > 0. \qquad (65)$$

Рассмотрим одно приложение этой теоремы. Пусть у матрицы $C = \| c_{ik} \|_1^n$ все недиагональные элементы неотрицательны. Тогда при некотором $\lambda > 0$ матрица $A = C + \lambda E \geqslant 0$. Характеристические числа $\lambda_i \, (i = 1, 2, \ldots, n)$ матрицы C расположим в порядке возрастания вещественных частей:

$$\operatorname{Re} \lambda_1 \leqslant \operatorname{Re} \lambda_2 \leqslant \ldots \leqslant \operatorname{Re} \lambda_n.$$

Обозначим через r максимальное характеристическое число матрицы A. Поскольку характеристическими числами матрицы A являются суммы $\lambda_i + \lambda$ $(i = 1, 2, \ldots, n)$, то

$$\lambda_n + \lambda = r.$$

В данном случае неравенство $r < \lambda$ имеет место лишь при $\lambda_n < 0$ и означает, что у матрицы C все характеристические числа имеют отрицательные вещественные части. Записывая неравенства (65) для матрицы $-C = \lambda E - A$, мы получим теорему [1]):

Теорема 5. *Для того чтобы у вещественной матрицы* $C = \| c_{ik} \|_1^n$ *с неотрицательными недиагональными элементами*

$$c_{ik} \geqslant 0 \qquad (i \neq k; \ i, k = 1, 2, \ldots, n)$$

все характеристические числа имели отрицательные вещественные части, необходимо и достаточно, чтобы выполнялись неравенства

$$c_{11} < 0, \quad \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} > 0, \ \ldots, \ (-1)^n \begin{vmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{vmatrix} > 0. \qquad (66)$$

Пусть снова A — произвольная неразложимая неотрицательная матрица, а $x \geqslant 0$ $(x \neq 0)$ — некоторый вектор [2]), не являющийся собственным для максимального характеристического числа r. Тогда в силу замечания 5 на стр. 365 существуют индексы i и j $(1 \leqslant i, j \leqslant n)$ такие, что выполняются неравенства

$$(Ax)_i > r x_i \quad \text{и} \quad (Ax)_j < r x_j. \qquad (67)$$

Если же x — собственный вектор матрицы A для характеристического числа r, то в соотношениях (67) знаки неравенств следует заменить знаком равенства. Поэтому для любого вектора $x \geqslant 0$ существуют

---

[1]) См. [123, 97в]. Поскольку $C = A - \lambda E$, $A \geqslant 0$, то λ_n вещественно (это следует из равенства $\lambda_n + \lambda = r$), и этому характеристическому числу соответствует неотрицательный собственный вектор матрицы C: $Cy = \lambda_n y$ $(y \geqslant 0, \ y \neq 0)$.

[2]) Обозначение $x \geqslant 0$ означает, что столбцевая матрица x неотрицательна.

индексы i и k $(1 \leqslant i,\ k \leqslant n)$, при которых

$$(Ax)_i \geqslant rx_i \quad \text{и} \quad (Ax)_j \leqslant rx_j. \tag{67'}$$

В таком ослабленном виде соотношения (67′) остаются в силе и для разложимой матрицы $A \geqslant 0$, поскольку она может быть представлена в виде предела последовательности неразложимых матриц.

Соотношения (67′) позволяют установить следующую теорему:

Теорема 6. *При увеличении любого элемента неотрицательной матрицы A максимальное характеристическое число не убывает. Оно строго возрастает, если A — неразложимая матрица.*

Эта теорема допускает эквивалентную формулировку.

Теорема 6′. *Если даны две неотрицательные матрицы A и A_1 с максимальными характеристическими числами r и r_1, то из неравенства $A \leqslant A_1$ $(A \neq A_1)$ следует неравенство $r \leqslant r_1$. Если же A — неразложимая матрица, то $r < r_1$.*

Доказательство. Пусть A — неразложимая матрица. Тогда A_1 — также неразложимая матрица. Обозначим через x собственный вектор матрицы A_1 для характеристического числа r_1:

$$A_1 x = r_1 x \quad (x > 0).$$

Отсюда

$$(r_1 - r)\, x = Ax - rx + (A_1 - A)\, x. \tag{67''}$$

Но $(A_1 - A)\, x \geqslant 0$. Поэтому, если x не является собственным вектором матрицы A для характеристического числа r, то в силу (67) при некотором индексе i $(1 \leqslant i \leqslant n)$

$$(r_1 - r)\, x_i \geqslant (Ax)_i - rx_i > 0,$$

откуда $r_1 - r > 0$, т. е. $r < r_1$.

Если же x — собственный вектор матрицы A при характеристическом числе r, то $Ax - rx = 0$, и поскольку при некотором индексе i $[(A_1 - A)\, x]_i > 0$, то из равенства (67″) следует:

$$(r_1 - r)\, x_i = (A_1 - A)\, x_i > 0,$$

т. е. снова $r < r_1$.

В случае разложимой матрицы введем в рассмотрение матрицы $A_\varepsilon = A + \varepsilon B$ и $A_{1\varepsilon} = A_1 + \varepsilon B$, где $B > 0$, $\varepsilon > 0$. Тогда $A_\varepsilon \leqslant A_{1\varepsilon}$ и $A_\varepsilon > 0$. Поэтому $r_\varepsilon < r_{1\varepsilon}$, где r_ε и $r_{1\varepsilon}$ — максимальные характеристические числа матриц A_ε и $A_{1\varepsilon}$. В пределе при $\varepsilon \to 0$ матрицы A_ε и $A_{1\varepsilon}$ переходят в матрицы A и A_1, а неравенство $r_\varepsilon < r_{1\varepsilon}$ — в соотношение $r \leqslant r_1$.

Теорема доказана.

§ 4. Нормальная форма разложимой матрицы

Рассмотрим произвольную разложимую матрицу $A = \| a_{ik} \|_1^n$. Перестановкой рядов ее можно представить в виде

$$A = \begin{Vmatrix} B & 0 \\ C & D \end{Vmatrix}, \tag{68}$$

где B, D — квадратные матрицы.

Если какая-либо из матриц B и D разложима, то ее можно также представить в виде, аналогичном (68), после чего матрица A примет вид

$$A=\begin{Vmatrix} K & 0 & 0 \\ H & L & 0 \\ F & G & M \end{Vmatrix}.$$

Если какая-либо из матриц K, L, M разложима, то этот процесс можно продолжить. В результате надлежащей перестановкой рядов мы матрице A придадим треугольную блочную форму:

$$A=\begin{Vmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{ss} \end{Vmatrix}, \tag{69}$$

где диагональные блоки — квадратные неразложимые матрицы.

Диагональный блок A_{ii} $(1\leqslant i\leqslant s)$ будем называть *изолированным*, если

$$A_{ik}=0 \quad (k=1, 2, \dots, i-1, i+1, \dots, s).$$

Перестановкой блочных рядов (см. стр. 352) в матрице (69) можно все изолированные блоки поставить на первые места вдоль главной диагонали, после чего матрица A примет вид

$$A=\begin{Vmatrix} A_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_g & 0 & \dots & 0 \\ A_{g+1,1} & A_{g+1,2} & \dots & A_{g+1,g} & A_{g+1} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{sg} & A_{s,g+1} & \dots & A_s \end{Vmatrix}; \tag{70}$$

здесь A_1, A_2, ..., A_s — неразложимые матрицы, а в каждом ряду

$$A_{f1},\ A_{f2},\ \dots,\ A_{f,f-1} \quad (f=g+1, \dots, s)$$

по крайней мере одна из матриц не равна нулю.

Матрицу (70) будем называть *нормальной формой* разложимой матрицы A.

Покажем, что *нормальная форма матрицы A определяется однозначно с точностью до перестановки блочных рядов*[1]). Для этого рассмотрим линейный оператор $\boldsymbol{A}$, соответствующий матрице A в n-мерном векторном пространстве $\boldsymbol{R}$. Представлению матрицы A в виде (70) соответствует расщепление пространства $\boldsymbol{R}$ на координатные подпространства

$$\boldsymbol{R}=\boldsymbol{R}_1+\boldsymbol{R}_2+\dots+\boldsymbol{R}_g+\boldsymbol{R}_{g+1}+\dots+\boldsymbol{R}_s; \tag{71}$$

при этом всегда $\boldsymbol{R}_s$, $\boldsymbol{R}_{s-1}+\boldsymbol{R}_s$, $\boldsymbol{R}_{s-2}+\boldsymbol{R}_{s-1}+\boldsymbol{R}_s$, ... — инвариантные координатные подпространства для оператора A, причем между любыми двумя

[1]) Не нарушая нормальной формы, можно произвольно переставить между собой первые g блочных рядов. Кроме того, иногда возможны некоторые перестановки между последними $s-g$ блочными рядами, сохраняющие нормальность формы.

соседними из этих подпространств не существует промежуточного инвариантного подпространства.

Допустим, что наряду с нормальной формой (70) данной матрицы имеется другая нормальная форма, которая соответствует другому расщеплению R на координатные подпространства:

$$R = \widehat{R}_1 + \widehat{R}_2 + \ldots + \widehat{R}_g + \widehat{R}_{g+1} + \ldots + \widehat{R}_t. \qquad (71')$$

Однозначность нормальной формы будет показана, если мы докажем совпадение расщеплений (71) и (71′) с точностью до порядка слагаемых.

Пусть инвариантное подпространство $\widehat{R}_t$ имеет общие координатные векторы с R_k и не имеет таковых с $R_{k+1}, \ldots, R_s$. Тогда $\widehat{R}_t$ должно целиком содержаться в R_k, так как в противном случае $\widehat{R}_t$ содержало бы «меньшее» инвариантное подпространство — пересечение $\widehat{R}_t$ с $R_k + R_{k+1} + \ldots + R_s$. Далее, $\widehat{R}_t$ должно совпасть с R_k, так как в противном случае инвариантное подпространство $\widehat{R}_t + R_{k+1} + \ldots + R_s$ было бы промежуточным между инвариантными подпространствами $R_k + R_{k+1} + \ldots + R_s$ и $R_{k+1} + \ldots + R_s$. Поскольку R_k совпадает с $\widehat{R}_t$, R_k — инвариантное подпространство. Поэтому без нарушения нормальной формы матрицы R_k может быть поставлено на место R_s. Таким образом, мы можем считать в расщеплениях (71) и (71′): $R_s \equiv \widehat{R}_t$.

Рассмотрим теперь координатное подпространство $\widehat{R}_{t-1}$. Пусть оно имеет общие координатные векторы с R_l $(l < s)$, но не имеет таковых с $R_{l+1} + \ldots + R_s$. Тогда инвариантное подпространство $\widehat{R}_{t-1} + \widehat{R}_t$ должно целиком содержаться в $R_l + R_{l+1} + \ldots + R_s$, так как в противном случае существовало бы промежуточное инвариантное координатное подпространство между $\widehat{R}_t$ и $\widehat{R}_{t-1} + \widehat{R}_t$. Поэтому $\widehat{R}_{t-1} \subset R_l$. Далее, $\widehat{R}_{t-1} \equiv R_l$, поскольку в противном случае $\widehat{R}_{t-1} + R_{l+1} + \ldots + R_s$ было бы промежуточным инвариантным подпространством между $R_l + R_{l+1} + \ldots + R_s$ и $R_{l+1} + \ldots + R_s$. Из $\widehat{R}_{t-1} \equiv R_l$ следует, что $R_l + R_s$ — инвариантное подпространство. Поэтому R_l можно поставить на место R_{s-1}, после чего будем иметь:

$$\widehat{R}_{t-1} \equiv R_{s-1}, \qquad \widehat{R}_t \equiv R_s.$$

Продолжая этот процесс далее, мы в конце концов придем к тому, что $s = t$ и что расщепления (71) и (71′) совпадают с точностью до порядка слагаемых. Тогда с точностью до перестановки блочных рядов совпадают и соответствующие нормальные формы.

Из однозначности нормальной формы следует, что числа g и s являются некоторыми инвариантами для неотрицательной матрицы A [1]).

Пользуясь нормальной формой матрицы, докажем теорему:

Теорема 7. *Максимальному характеристическому числу r матрицы $A \geqslant 0$ соответствует положительный собственный вектор в том и только в том случае, когда в нормальной форме* (69) *матрицы A:* 1° *каждая из матриц $A_1, A_2, \ldots, A_g$ имеет число r своим характристи-*

[1]) Для неразложимой матрицы $g = s = 1$.

ческим числом и (при $g<s$) 2° ни одна из матриц $A_{g+1}, \ldots, A_s$ этим свойством не обладает.

Доказательство. 1. Пусть максимальному характеристическому числу r соответствует положительный собственный вектор $z>0$. В соответствии с разбиениями на блоки в (70) мы столбец z разобьем на части z^k $(k=1, 2, \ldots, s)$. Тогда равенство

$$Az=rz \qquad (z>0) \tag{72}$$

заменится двумя системами равенств:

$$A_iz^i=rz^i \qquad (i=1, 2, \ldots, g), \tag{72'}$$

$$\sum_{h=1}^{j-1} A_{jh}z^h+A_jz^j=rz^j \qquad (j=+1, \ldots, s). \tag{72''}$$

Из (72′) следует, что число r является характеристическим числом каждой из матриц $A_1, A_2, \ldots, A_g$. Из (72″) находим:

$$A_jz^j\leqslant rz^j, \qquad A_jz^j\neq rz^j \qquad (j=g+1, \ldots, s). \tag{73}$$

Обозначим через r_j максимальное характеристическое число матрицы A_j $(j=g+1, \ldots, s)$. Тогда [см. (41) на стр. 365] из (73) находим:

$$r_j\leqslant\max_i\frac{(A_jz^j)_i}{z_i^j}\leqslant r \qquad (j=g+1, \ldots, s).$$

С другой стороны, равенство $r_j=r$ противоречит вторым соотношениям (73) (см. замечание 5 на стр. 365). Поэтому

$$r_j<r \qquad (j=g+1, \ldots, s). \tag{74}$$

2. Пусть, теперь, обратно, дано, что максимальные характеристические числа матриц A_i $(i=1, 2, \ldots, g)$ равны r, а для матриц A_j $(j=g+1, \ldots, s)$ имеют место неравенства (74). Тогда, заменяя искомое равенство (72) системой равенств (72′), (72″), мы сможем из (72′) определить положительные собственные столбцы z^i матрицы A_i $(i=1, 2, \ldots, g)$. После этого из (72″) найдем столбцы z^j $(j=g+1, \ldots, s)$:

$$z^j=(rE_j-A_j)^{-1}\sum_{h=1}^{j-1}A_{jh}z^h \qquad (j=g+1, \ldots, s), \tag{75}$$

где E_j — единичная матрица того же порядка, что и матрица A_j $(j=g+1, \ldots, s)$.

Поскольку $r_j<r$ $(j=g+1, \ldots, s)$, то [см. (51) на стр. 367]

$$(rE_j-A_j)^{-1}>0 \qquad (j=g+1, \ldots, s). \tag{76}$$

Докажем индуктивно, что определенные по формулам (75) столбцы $z^{g+1}, \ldots, z^s$ положительны. Мы покажем, что при любом j $(g+1\leqslant j\leqslant s)$ из положительности столбцов $z^1, z^2, \ldots, z^{j-1}$ следует: $z^j>0$. Действительно, в этом случае

$$\sum_{h=1}^{j-1}A_{jh}z^h\geqslant 0, \qquad \sum_{h=1}^{j-1}A_{jh}z^h\neq 0,$$

что в соединении с (76) на основании формулы (75) дает:

$$z^j > 0.$$

Таким образом, положительный столбец $z = \begin{Vmatrix} z^1 \\ \cdot \\ \cdot \\ \cdot \\ z^s \end{Vmatrix}$ будет собственным вектором для матрицы A при характеристическом числе r. Теорема доказана.

Следующая теорема дает нам характеристику матрицы $A \geqslant 0$, которая вместе со своей транспонированной матрицей A' обладает тем свойством, что максимальному характеристическому числу отвечает положительный собственный вектор.

Теорема 7'[1]). *Максимальному характеристическому числу r матрицы $A \geqslant 0$ отвечает положительный собственный вектор матрицы A и положительный собственный вектор транспонированной матрицы A' в том и только в том случае, когда матрица A может быть перестановкой рядов представлена в квазидиагональном виде*

$$A = \{A_1, A_2, \ldots, A_s\}, \tag{77}$$

где $A_1, A_2, \ldots, A_s$ — неразложимые матрицы, каждая из которых имеет число r своим максимальным характеристическим числом.

Доказательство. Пусть матрицы A и A' имеют положительные собственные векторы для $\lambda = r$. Тогда по теореме 7 матрица A представима в нормальной форме (69), где матрицы $A_1, A_2, \ldots, A_g$ имеют максимальное характеристическое число r и (при $g < s$) матрицы $A_{g+1}, \ldots, A_s$ имеют максимальные характеристические числа $< r$. Тогда

$$A' = \begin{Vmatrix} A'_1 & \ldots & 0 & A'_{g+1,\,1} & \ldots & A'_{s1} \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & \ldots & A'_g & A'_{g+1,\,g} & \ldots & A'_{sg} \\ 0 & \ldots & 0 & A'_{g+1} & & \\ \cdot & & & \cdot & & \\ \cdot & & & \cdot & & \\ \cdot & & & \cdot & & \\ 0 & \ldots & 0 & 0 & \ldots & A'_s \end{Vmatrix}.$$

Поменяем здесь порядок блочных рядов на обратный:

$$\begin{Vmatrix} A'_s & 0 & 0 & \ldots & 0 \\ A'_{s,\,s-1} & A'_{s-1} & 0 & \ldots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ A'_{s1} & A'_{s-1,\,1} & & \ldots & A'_1 \end{Vmatrix}. \tag{78}$$

Поскольку матрицы $A'_s, A'_{s-1}, \ldots, A'_1$ неразложимы, то из матрицы (78) перестановкой блочных рядов мы получим нормальную форму, поставив

[1]) См. [182e].

на первые места вдоль главной диагонали изолированные блоки. Одним из таких изолированных блоков является блок A'_s. Поскольку нормальная форма матрицы A' должна удовлетворять условию предыдущей теоремы, то максимальное характеристическое число матрицы A'_s должно равняться r. Это возможно лишь при $g=s$. В этом случае нормальная форма (69) переходит в (77).

Если, обратно, дано представление (77) для матрицы A, то тогда

$$A'=\{A'_1, A'_2, \ldots, A'_s\}. \tag{79}$$

Тогда из (77) и (79) в силу предыдущей теоремы заключаем, что матрицы A и A' имеют положительные собственные векторы для максимального характеристического числа r.

Теорема доказана.

Следствие. *Если максимальное характеристическое число r матрицы $A\geqslant 0$ является простым и ему соответствуют положительные собственные векторы матриц A и A', то A — неразложимая матрица.*

Поскольку, обратно, всякая неразложимая матрица обладает свойствами, указанными в этом следствии, то эти свойства представляют собой спектральную характеристику неразложимой неотрицательной матрицы.

§ 5. Примитивные и импримитивные матрицы

Начнем с некоторой классификации неразложимых матриц.

Определение 3. Если неразложимая матрица $A\geqslant 0$ имеет всего h характеристических чисел $\lambda_1, \lambda_2, \ldots, \lambda_h$ с максимальным модулем r ($\lambda_1 = |\lambda_2| = \ldots = |\lambda_h| = r$), то матрица A называется *примитивной* при $h=1$ *и импримитивной* при $h>1$. Число h называется *индексом импримитивности* матрицы A.

Индекс импримитивности h сразу определяется, если известны коэффициенты характеристического уравнения

$$\Delta(\lambda)\equiv\lambda^n+a_1\lambda^{n_1}+a_2\lambda^{n_2}+\ldots+a_t\lambda^{n_t}=0$$
$$(n>n_1>\ldots>n_t;\quad a_1\neq 0,\quad a_2\neq 0,\quad \ldots,\quad a_t\neq 0)$$

матрицы, а именно *число h равно наибольшему общему делителю разностей*

$$n-n_1,\quad n_1-n_2,\quad \ldots,\quad n_{t-1}-n_t. \tag{80}$$

Действительно, согласно теореме Фробениуса спектр матрицы A в комплексной λ-плоскости переходит в себя при повороте на угол $\frac{2\pi}{h}$ вокруг точки $\lambda=0$. Поэтому многочлен $\Delta(\lambda)$ должен получиться из некоторого многочлена $g(\mu)$ по формуле

$$\Delta(\lambda)=g(\lambda^h)\lambda^{n'}.$$

Отсюда следует, что h — общий делитель разностей (80). Наконец, h равняется наибольшему общему делителю d этих разностей, так как спектр не изменяется при повороте на угол $\frac{2\pi}{d}$, а это невозможно при $h<d$.

Следующая теорема устанавливает важное свойство примитивной матрицы:

Теорема 8. *Матрица $A\geqslant 0$ является примитивной в том и только в том случае, когда некоторая степень матрицы A*

положительна:

$$A^p > 0 \quad (p \geqslant 1). \tag{81}$$

Доказательство. Если $A^p > 0$, то матрица A неразложима, так как из разложимости матрицы A следует разложимость матрицы A^p. Далее, для матрицы A число $h = 1$, так как в противном случае положительная матрица A^p имела бы $h(>1)$ характеристических чисел $\lambda_1^p, \lambda_2^p, \ldots, \lambda_h^p$ с максимальным модулем r^p, что противоречит теореме Перрона.

Пусть теперь, обратно, дано, что A — примитивная матрица. Воспользуемся для степени A^p формулой (24) главы V (стр. 113)

$$A^p = \sum_{k=1}^{s} \frac{1}{(m_k-1)!} \left[\frac{C(\lambda)\lambda^p}{\overset{k}{\psi}(\lambda)} \right]_{\lambda=\lambda_k}^{(m_k-1)}, \tag{82}$$

где

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2} \ldots (\lambda-\lambda_s)^{m_s} \quad (\lambda_j \neq \lambda_f \text{ при } j \neq f)$$

— минимальный многочлен матрицы A, $\overset{k}{\psi}(\lambda) = \dfrac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}$ $(k = 1, 2, \ldots, s)$, а $C(\lambda)$ — приведенная присоединенная матрица $C(\lambda) = (\lambda E - A)^{-1}\psi(\lambda)$.

В данном случае можно положить:

$$\lambda_1 = r > |\lambda_2| \geqslant \ldots \geqslant |\lambda_s|, \quad m_1 = 1. \tag{83}$$

Тогда формула (82) примет вид

$$A^p = \frac{C(r)}{\psi'(r)} r^p + \sum_{k=2}^{s} \frac{1}{(m_k-1)!} \left[\frac{C(\lambda)\lambda^p}{\overset{k}{\psi}(\lambda)} \right]_{\lambda=\lambda_k}^{(m_k-1)}.$$

Отсюда легко заключить в силу (83)

$$\lim_{p\to\infty} \frac{A^p}{r^p} = \frac{C(r)}{\psi'(r)}. \tag{84}$$

С другой стороны, $C(r)$ [см. (53)] и $\psi'(r) > 0$ в силу (83). Поэтому

$$\lim_{p\to\infty} \frac{A^p}{r^p} > 0,$$

и, следовательно, начиная с некоторого p, имеет место неравенство (81).

Теорема доказана.

Замечание. *Если матрица A примитивна и $A^p > 0$, то $A^m > 0$ для всех $m > p$, так как матрица A не содержит нулевых рядов.*

Следствие. *Степень примитивной матрицы всегда неразложима и притом примитивна.*

Для наименьшего номера $p = p_A$, начиная с которого выполняется неравенство (81), Фробениус [1]) указал верхнюю оценку, зависящую только от порядка n матрицы A:

$$p_A \leqslant 2n^2 - 2n.$$

Виландт [2]) отметил (без доказательства), что на самом деле

$$p_A \leqslant n^2 - 2n + 2, \tag{85}$$

---

[1]) См. [182e].

[2]) См. [260].

и эта оценка точна. Она достигается на матрице

$$A=\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Приводимое ниже доказательство неравенства (85) по существу совпадает с доказательством, принадлежащим Седлачеку[1]).

Лемма. *Если A — примитивная матрица, то для любых двух (не обязательно различных) индексов i, k существует такая цепочка индексов i, i_1, i_2, ..., i_s, k ($s \geqslant 0$), что*

$$a_{ii_1}>0,\ a_{i_1i_2}>0,\ \dots,\ a_{i_sk}>0.$$

О такой цепочке будем говорить, что она ведет в матрице A из i в k. Число $s+1$ назовем длиной цепочки. Очевидно, в кратчайшей цепочке, ведущей из i в k, все индексы попарно различны.

Для доказательства леммы достаточно взять $s \geqslant 0$ так, чтобы было

$$A^{s+1} \equiv \| a_{ik}^{(s+1)} \|_1^n > 0.$$

Тогда

$$\sum_{i_1,\, i_2,\, \dots,\, i_s=1}^{n} a_{ii_1}a_{i_1i_2} \dots a_{i_sk} = a_{ik}^{(s+1)} > 0,$$

и так как все слагаемые здесь неотрицательны, то по крайней мере одно из них положительно. Оно и дает требуемую цепочку индексов.

Перейдем к доказательству неравенства (85).

Обозначим через l_i наименьшую из длин цепочек, ведущих в матрице A из i в i ($i=1, 2, \dots, n$), и положим[2])

$$l = \min_{1 \leqslant i \leqslant n} l_i.$$

Пусть для определенности

$$a_{12}>0,\ a_{23}>0,\ \dots,\ a_{l1}>0. \tag{86}$$

Тогда в матрице A^l будут положительными первые l диагональных элементов:

$$a_{11}^{(l)}>0,\ a_{22}^{(l)}>0,\ \dots,\ a_{ll}^{(l)}>0. \tag{87}$$

Возьмем произвольный индекс i. Кратчайшая цепочка, ведущая в матрице A из i в какой-нибудь из индексов 1, 2, ..., l, имеет, очевидно, длину, не большую $n-l$. В силу (86) эту цепочку можно продолжить за счет индексов 1, 2, ..., l до цепочки с длиной ровно $n-l$. Получится некоторая цепочка

$$i,\ i_1,\ i_2,\ \dots,\ i_{n-l-1},\ j,$$

где $1 \leqslant j \leqslant l$.

---

[1]) См. [245]. Другие доказательства были даны в статьях [190], [119], [229]. В работе [110] неравенство Виландта послужило отправным пунктом для обобщений, относящихся к любым матрицам $A \geqslant 0$.

[2]) l можно определить как наименьший показатель, для которого матрица A^l имеет положительный диагональный элемент.

Возьмем еще один произвольный индекс k (не обязательно отличный от i). Так как матрица A^l, согласно следствию теоремы 8, тоже примитивна, то найдется цепочка индексов длины, не большей $n-1$, ведущая в матрице A^l из j в k. В силу (87) эту цепочку можно продолжить за счет индекса j до цепочки с длиной ровно $n-1$. Получится некоторая цепочка

$$j,\ j_1,\ j_2,\ \ldots,\ j_{n-2},\ k.$$

Итак,

$$a_{ii_1}>0,\ a_{i_1i_2}>0,\ \ldots,\ a_{i_{n-l-1}j}>0$$

и

$$a^{(l)}_{jj_1}>0,\ a^{(l)}_{j_1j_2}>0,\ \ldots,\ a^{(l)}_{j_{n-2}k}>0.$$

Отсюда

$$a_{ii_1}a_{i_1i_2}\ldots a_{i_{n-l-1}j}a^{(l)}_{jj_1}a^{(l)}_{j_1j_2}\ldots a^{(l)}_{j_{n-2}k}>0.$$

Следовательно,

$$a^{(n-l+l(n-1))}_{ik}>0.$$

Ввиду произвольности индексов i, k оказывается

$$A^{n-l+l(n-1)}>0.$$

Тем самым

$$p_A\leqslant n-l+l(n-1)=(n-2)\,l+n.$$

Заметим теперь, что $l\leqslant n-1$. Действительно, в противном случае $l=n$ и в силу определения числа l матрица A оказывается циклической:

$$A=\begin{pmatrix}0 & a_{12} & 0 & \ldots & 0\\ 0 & 0 & a_{23} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & a_{n-1,\,n}\\ a_{n1} & 0 & 0 & \ldots & 0\end{pmatrix},$$

т. е. импримитивной.

Таким образом,

$$p_A\leqslant(n-2)\,(n-1)+n=n^2-2n+2,$$

что и требовалось доказать.

Докажем теперь следующую теорему:

Т е о р е м а 9. *Если $A\geqslant 0$ — неразложимая матрица и некоторая степень A^q этой матрицы разложима, то степень A^q вполне разложима, т. е. перестановкой рядов A^q может быть представлена в виде*

$$A^q=\{A_1,\ A_2,\ \ldots,\ A_d\},\tag{88}$$

где $A_1,\ A_2,\ \ldots,\ A_d$ — неразложимые матрицы. Эти матрицы имеют одно и то же максимальное характеристическое число. При этом число d есть наибольший общий делитель чисел q и h, где h — индекс импримитивности матрицы A.

Д о к а з а т е л ь с т в о. Поскольку матрица A является неразложимой, то согласно теореме Фробениуса максимальному характеристическому числу r отвечают положительные собственные векторы матриц A и A'. Но тогда эти же положительные векторы являются собственными для неотрицательных матриц A^q и $(A^q)'$ при характеристическом числе

$\lambda = r^q$. Поэтому, применяя к степени A^q теорему 7′, мы представим (после надлежащей перестановки рядов) эту степень в виде (88), где $A_1, A_2, \ldots$ $\ldots, A_d$ — неразложимые матрицы с одним и тем же максимальным характеристическим числом r^q. Но матрица A имеет h характеристических чисел с максимальным модулем r:

$$r,\ r\varepsilon,\ \ldots,\ r\varepsilon^{h-1} \qquad (\varepsilon = e^{\frac{2\pi i}{h}}).$$

Поэтому и матрица A^q имеет h характеристических чисел с максимальным модулем

$$r^q,\ r^q\varepsilon^q,\ \ldots,\ r^q\varepsilon^{q(h-1)},$$

среди которых d чисел равны r^q. Это возможно лишь тогда, когда d — наибольший общий делитель чисел q и h. Теорема доказана.

Если в формулировке теоремы положить $q = h$, то получим

Следствие. *Если A — импримитивная матрица с индексом импримитивности h, то степень A^h разлагается на h примитивных матриц, которые имеют одно и то же максимальное характеристическое число.*

§ 6. Стохастические матрицы

Рассмотрим n возможных состояний некоторой системы

$$S_1,\ S_2,\ \ldots,\ S_n \tag{89}$$

и последовательность моментов времени

$$t_0,\ t_1,\ t_2,\ \ldots$$

Пусть в каждый из этих моментов времени система находится в одном и только в одном из состояний (89), причем p_{ij} обозначает вероятность нахождения системы в состоянии S_j в момент времени t_k, если известно, что в предыдущий момент времени t_{k-1} система находилась в состоянии T_i ($i, j = 1, 2, \ldots, n$; $k = 1, 2, \ldots$). Мы будем предполагать, что *переходные вероятности* p_{ij} ($i, j = 1, 2, \ldots$ $\ldots, n$) не зависят от индекса k (номера момента времени t_k).

Если матрица переходных вероятностей

$$P = \| p_{ij} \|_1^n$$

задана, то говорят, что задана *однородная цепь Маркова с конечным числом состояний* [1]). При этом очевидно, что

$$p_{ij} \geqslant 0, \quad \sum_{j=1}^{n} p_{ij} = 1 \qquad (i, j = 1, 2, \ldots, n). \tag{90}$$

Определение 4. Квадратная матрица $P = \| p_{ij} \|_1^n$ называется *стохастической*, если матрица P неотрицательна и сумма элементов каждой строки матрицы P равна единице, т. е. имеют место соотношения (90) [2]).

Таким образом, для каждой однородной цепи Маркова матрица переходных вероятностей является стохастической и, наоборот, любая

---

[1]) См. [96], а также [29], стр. 9—12.

[2]) Иногда в определение стохастической матрицы включают дополнительное требование $\sum_{i=1}^{n} p_{ij} \neq 0$ ($j = 1, 2, \ldots, n$). См. [29], стр. 13.

стохастическая матрица может быть рассматриваема как матрица переходных вероятностей некоторой однородной цепи Маркова. На этом основывается матричный метод исследования однородных цепей Маркова [1]).

Стохастическая матрица является частным видом неотрицательной матрицы. Поэтому к ней применимы все понятия и положения предыдущих параграфов.

Отметим некоторые специфические свойства стохастической матрицы. Из определения стохастической матрицы следует, что эта матрица имеет характеристическое число 1 с положительным собственным вектором $z = (1, 1, \ldots, 1)$. Легко видеть, что, и обратно, всякая матрица $P \geqslant 0$, имеющая собственный вектор $(1, 1, \ldots, 1)$ при характеристическом числе 1, является стохастической. При этом единица является максимальным характеристическим числом стохастической матрицы, поскольку максимальное характеристическое число всегда заключено между наибольшей и наименьшей из строчных сумм [2]), а для стохастической матрицы все строчные суммы равны единице. Таким образом, нами доказано предложение:

1° *Неотрицательная матрица* $P \geqslant 0$ *является стохастической тогда и только тогда, когда она имеет собственный вектор* $(1, 1, \ldots, 1)$ *при характеристическом числе* 1. *Характеристическое число* 1 *является максимальным для стохастической матрицы.*

Пусть теперь дана неотрицательная матрица $A = \| a_{ik} \|_1^n$, имеющая положительное максимальное характеристическое число $r > 0$ и соответствующий этому числу положительный собственный вектор $z = (z_1, z_2, \ldots, z_n) > 0$:

$$\sum_{j=1}^{n} a_{ij} z_j = r z_i \qquad (i = 1, 2, \ldots, n). \tag{91}$$

Введем в рассмотрение диагональную матрицу $Z = \{z_1, z_2, \ldots, z_n\}$ и матрицу $P = \| p_{ij} \|_1^n$

$$P = \frac{1}{r} Z^{-1} A Z.$$

Тогда

$$p_{ij} = \frac{1}{r} z_i^{-1} a_{ij} z_j \geqslant 0 \qquad (i, j = 1, 2, \ldots, n),$$

и в силу (91)

$$\sum_{j=1}^{n} p_{ij} = 1 \qquad (i = 1, 2, \ldots, n).$$

Таким образом,

2° *Неотрицательная матрица* $A \geqslant 0$, *имеющая положительное максимальное характеристическое число* $r > 0$ *и соответствующий этому числу положительный собственный вектор* $z = (z_1, z_2, \ldots, z_n) > 0$, *всегда подобна произведению числа* r *на некоторую стохастическую матрицу*:

$$A = ZrPZ^{-1} \qquad (Z = \{z_1, z_2, \ldots, z_n\} > 0)\ [3]). \tag{92}$$

[1]) Теория однородных цепей Маркова с конечным (и со счетным) числом состояний была разработана акад. А. Н. Колмогоровым (см. [96]). Последовательное проведение и развитие матричного метода в применении к однородным цепям Маркова читатель найдет в мемуаре [120б] и монографии [29] В. И. Романовского (см. также [4], дополнение 5).

[2]) См. неравенства (37) и замечание на стр. 366.

[3]) Предложение 2° имеет место и при $r = 0$, так как из $A \geqslant 0$, $z > 0$ следует $A = 0$.

В предыдущем параграфе была установлена (см. теорему 7) характеристика класса неотрицательных матриц, имеющих положительный собственный вектор при $\lambda = r$. Формула (92) устанавливает тесную связь этого класса матриц с классом стохастических матриц.

Докажем теперь следующую теорему:

Т е о р е м а 10. *Характеристическому числу* 1 *стохастической матрицы всегда соответствуют только элементарные делители первой степени.*

Д о к а з а т е л ь с т в о. Применим к стохастической матрице $P = \| p_{ij} \|_1^n$ разложение (69) § 4:

$$P = \begin{Vmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & A_g\ 0 & \dots\ 0 \\ A_{g+1,1} & \dots & A_{g+1,g}\ A_{g+1} & \dots\ 0 \\ \dots & \dots & \dots & \dots \\ A_{s1} & \dots & A_{sg} & \dots\ A_s \end{Vmatrix},$$

где $A_1, A_2, \dots, A_s$ — неразложимые матрицы и

$$A_{f1} + A_{f2} + \dots + A_{f,f-1} \neq 0 \qquad (f = g+1, \dots, s).$$

Здесь $A_1, A_2, \dots, A_g$ — неразложимые с т о х а с т и ч е с к и е матрицы, и потому каждая из этих матриц имеет простое характеристическое число 1. Что же касается остальных неразложимых матриц $A_{g+1}, \dots, A_s$, то согласно замечанию 2 на стр. 362 их максимальные характеристические числа < 1, поскольку в каждой из этих матриц хотя бы одна строчная сумма меньше единицы [1]).

Таким образом, матрица P представима в виде

$$P = \begin{Vmatrix} Q_1 & 0 \\ S & Q_2 \end{Vmatrix},$$

где у матрицы Q_1 характеристическому числу 1 соответствуют элементарные делители первой степени, а для матрицы Q_2 число 1 не является характеристическим числом. После этого справедливость теоремы непосредственно вытекает из следующей леммы:

Л е м м а 4. *Если матрица A имеет вид*

$$A = \begin{Vmatrix} Q_1 & 0 \\ S & Q_2 \end{Vmatrix}, \tag{93}$$

где Q_1 и Q_2 — квадратные матрицы, и характеристическое число λ_0 матрицы A является характеристическим числом матрицы Q_1 и не является таковым для матрицы Q_2,

$$|Q_1 - \lambda_0 E| = 0, \quad |Q_2 - \lambda_0 E| \neq 0,$$

то элементарные делители матриц A и Q_1, соответствующие характеристическому числу λ_0, одинаковы.

[1]) Эти свойства матриц $A_1, \dots, A_s$ вытекают также из теоремы 7.

Доказательство. 1. Рассмотрим сначала случай, когда Q_1 и Q_2 не имеют общих характеристических чисел. Покажем, что в этом случае элементарные делители матриц Q_1 и Q_2 в совокупности образуют систему элементарных делителей матрицы A, т. е. что при некоторой матрице T ($|T| \neq 0$)

$$TAT^{-1} = \begin{Vmatrix} Q_1 & 0 \\ 0 & Q_2 \end{Vmatrix}. \tag{94}$$

Матрицу T будем искать в виде

$$T = \begin{Vmatrix} E_1 & 0 \\ U & E_2 \end{Vmatrix}$$

(разбиение на блоки в T соответствует разбиению в A; E_1 и E_2 — единичные матрицы). Тогда

$$TAT^{-1} = \begin{Vmatrix} E_1 & 0 \\ U & E_2 \end{Vmatrix} \begin{Vmatrix} Q_1 & 0 \\ S & Q_2 \end{Vmatrix} \begin{Vmatrix} E_1 & 0 \\ -U & E_2 \end{Vmatrix} = \begin{Vmatrix} Q_1 & 0 \\ UQ_1 - Q_2U + S, & Q_2 \end{Vmatrix}. \tag{94'}$$

Равенство (94′) перейдет в равенство (94), если прямоугольную матрицу U подберем так, чтобы она удовлетворяла матричному уравнению

$$Q_2U - UQ_1 = S.$$

В случае, когда Q_1 и Q_2 не имеют общих характеристических чисел, это уравнение при любой правой части S всегда имеет одно определенное решение (см. гл. VIII, § 3).

2. В случае, когда матрицы Q_1 и Q_2 могут иметь и общие характеристические числа, мы заменим в (93) матрицу Q_1 ее жордановой формой J (в результате этого матрица A заменится матрицей, ей подобной). При этом $J = \{J_1, J_2\}$, где в J_1 собраны все жордановы клетки с характеристическим числом λ_0. Тогда

$$A = \left\|\begin{array}{cccc} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ S_{11} & S_{12} & \multicolumn{2}{c}{\multirow{2}{*}{Q_2}} \\ S_{21} & S_{22} & & \end{array}\right\| = \left\|\begin{array}{cccc} J_1 & 0 & 0 & 0 \\ 0 & & \multicolumn{2}{c}{\multirow{3}{*}{$\hat{Q}_2$}} \\ S_{11} & & & \\ S_{21} & & & \end{array}\right\|$$

Эта матрица подходит под разобранный уже первый случай, поскольку матрицы J_1 и $\hat{Q}_2$ не имеют общих характеристических чисел. Отсюда следует, что элементарные делители вида $(\lambda - \lambda_0)^p$ одинаковы у матриц A и J_1 и, следовательно, одинаковы у матриц A и Q_1. Лемма доказана.

Если неразложимая стохастическая матрица P имеет комплексное характеристическое число λ_0 с $|\lambda_0| = 1$, то матрица $\lambda_0 P$ подобна матрице P [см. (16)] и потому из теоремы 10 вытекает, что числу λ_0 отвечают только элементарные делители первой степени. Пользуясь нормальной формой матрицы и леммой 4, легко распространить это утверждение и на разложимые стохастические матрицы. Таким образом, получаем

Следствие 1. *Если λ_0 — характеристическое число стохастической матрицы и $|\lambda_0| = 1$, то этому числу λ_0 соответствуют элементарные делители первой степени матрицы P.*

Из теоремы 10 в силу 2° (стр. 382) вытекает также

Следствие 2. *Если максимальному характеристическому числу r неотрицательной матрицы A отвечает положительный собственный*

вектор, то все элементарные делители матрицы A, соответствующие любому характеристическому числу λ_0 с $|\lambda_0| = r$, имеют первую степень.

Укажем на некоторые работы, связанные с расположением характеристических чисел стохастической матрицы.

Характеристическое число стохастической матрицы P всегда лежит в круге $|\lambda| \leqslant 1$ λ-плоскости. Совокупность всех точек этого круга, являющихся характеристическими числами каких-либо стохастических матриц n-го порядка, обозначим через M_n.

В 1938 г. в связи с исследованием цепей Маркова акад. А. Н. Колмогоров поставил задачу определения структуры области M_n. Эта задача была частично решена в 1945 г. Н. А. Дмитриевым и Е. Б. Дынкиным [67а, б] и полностью решена в 1951 г. в работе Ф. И. Карпелевича [72]. Оказалось, что граница M_n состоит из конечного числа точек на окружности $|\lambda| = 1$ и определенных криволинейных дуг, соединяющих в круговом порядке эти точки.

Заметим, что в силу предложения 2° (стр. 382) характеристические числа матриц $A = \| a_{ik} \|_1^n \geqslant 0$, имеющих положительный собственный вектор при $\lambda = r$, при фиксированном r образуют множество $r \cdot M_n$ [1]). Поскольку произвольная матрица $A = \| a_{ik} \|_1^n \geqslant 0$ может быть рассматриваема как предел последовательности неотрицательных матриц указанного типа, а множество $r \cdot M_n$ замкнуто, то характеристические числа произвольных матриц $A = \| a_{ik} \|_1^n > 0$ с данным максимальным характеристическим числом r заполняют множество $r \cdot M_n$ [2]).

К этому кругу вопросов относится и работа Х. Р. Сулеймановой [100], в которой устанавливаются некоторые достаточные критерии для того, чтобы n заданных вещественных чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ были характеристическими числами некоторой стохастической матрицы $P = \| p_{ij} \|_1^n$ [3]).

§ 7. Предельные вероятности для однородной цепи Маркова с конечным числом состояний

1. Пусть

$$S_1, S_2, \ldots, S_n$$

— все возможные состояния системы в однородной цепи Маркова, а $P = \| p_{ij} \|_1^n$ — определяющая эту цепь стохастическая матрица, составленная из переходных вероятностей p_{ij} $(i, j = 1, 2, \ldots, n)$ (см. стр. 381).

Обозначим через $p_{ij}^{(q)}$ вероятность нахождения системы в состоянии S_j в момент времени t_k, если известно, что в момент времени t_{k-q} система находилась в состоянии S_i $(i, j = 1, 2, \ldots, n;\ q = 1, 2, \ldots)$. Очевидно, $p_{ij}^{(1)} = p_{ij}$ $(i, j = 1, 2, \ldots, n)$. Пользуясь теоремами о сложении и умножении вероятностей, мы легко найдем:

$$p_{ij}^{(q+1)} = \sum_{h=1}^{n} p_{ih}^{(q)} p_{hj} \qquad (i, j = 1, 2, \ldots, n),$$

или в матричной записи

$$\| p_{ij}^{(q+1)} \| = \| p_{ij}^{(q)} \|_1^n \| p_{ij} \|_1^n.$$

---

[1]) $r \cdot M_n$ есть совокупность точек λ-плоскости вида $r\mu$, где $\mu \in M_n$.

[2]) На возможность сведения указанной задачи для произвольной матрицы $A \geqslant 0$ к аналогичной задаче для стохастической матрицы было указано А. Н. Колмогоровым (см. [89б], дополнение).

[3]) См. также [228b].

Отсюда, давая q последовательно значения 1, 2, ..., получим важную формулу [1])

$$\| p_{ij}^{(q)} \| = P^q \qquad (q = 1, 2, \ldots).$$

Если существуют пределы

$$\lim_{q\to\infty} p_{ij}^{(q)} = p_{ij}^{\infty} \qquad (i, j = 1, 2, \ldots, n)$$

или в матричной записи

$$\lim_{q\to\infty} P^q = P^{\infty} = \| p_{ij}^{\infty} \|_1^n,$$

то величины p_{ij}^{∞} $(i, j = 1, 2, \ldots, n)$ называются *предельными или финальными переходными вероятностями* [2]).

Для выяснения, в каких случаях существуют предельные переходные вероятности, и для вывода соответствующих формул введем следующую терминологию.

Мы будем стохастическую матрицу P и соответствующую ей однородную цепь Маркова называть *правильной*, если у матрицы P нет характеристических чисел, отличных от единицы и равных по модулю единице, и *регулярной*, если дополнительно единица является простым корнем характеристического уравнения матрицы P.

Правильная матрица P характеризуется тем, что в ее нормальной форме (69) (стр. 373) матрицы $A_1, A_2, \ldots, A_g$ являются примитивными. Для регулярной матрицы дополнительно $g = 1$.

Кроме того, однородная цепь Маркова называется *неразложимой*, *разложимой*, *ациклической*, *циклической*, если для этой цепи стохастическая матрица P является соответственно неразложимой, разложимой, примитивной, импримитивной.

Поскольку примитивная стохастическая матрица является частным видом правильной матрицы, постольку ациклическая цепь Маркова является частным видом правильной цепи.

Мы покажем, что *предельные переходные вероятности существуют только у правильных однородных цепей Маркова.*

Действительно, пусть $\psi(\lambda)$ — минимальный многочлен правильной матрицы $P = \| p_{ij} \|_1^n$. Тогда

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_u)^{m_u} \quad (\lambda_i \neq \lambda_k;\ i, k = 1, 2, \ldots, u). \tag{95}$$

Согласно теореме 10 можно принять, что

$$\lambda_1 = 1, \quad m_1 = 1. \tag{95'}$$

На основании формулы (24) гл. V (стр. 113)

$$P^q = \frac{C(1)}{\psi(1)} + \sum_{k=2}^{u} \frac{1}{(m_k - 1)!} \left[\frac{C(\lambda)}{\psi(\lambda)} \lambda^q \right]_{\lambda=\lambda_k}^{(m_k-1)}, \tag{96}$$

---

[1]) Из этой формулы следует, что вероятности $p_{ij}^{(q)}$, так же как и p_{ij} $(i, j = 1, 2, \ldots, n;\ q = 1, 2, \ldots)$, не зависят от номера k исходного момента времени t_k.

[2]) Матрица P^{∞} как предел стохастических матриц также является стохастической.

где $C(\lambda)=(\lambda E-P)^{-1}\psi(\lambda)$ — приведенная присоединенная матрица и

$$\overset{k}{\psi}(\lambda)=\frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}} \qquad (k=1, 2, \ldots, u);$$

при этом

$$\overset{1}{\psi}(\lambda)=\frac{\psi(\lambda)}{\lambda-1} \quad \text{и} \quad \overset{1}{\psi}(1)=\psi'(1).$$

Если P — правильная матрица, то

$$|\lambda_k|<1 \qquad (k=2, 3, \ldots, u),$$

и потому в правой части формулы (96) все слагаемые, кроме первого, при $q\to\infty$ стремится к нулю. Поэтому для правильной матрицы P существует матрица P^∞, составленная из предельных переходных вероятностей, и

$$P^\infty=\frac{C(1)}{\psi'(1)}. \tag{97}$$

Обратное положение очевидно. Если существует предел

$$P^\infty=\lim_{q\to\infty}P^q, \tag{97'}$$

то матрица P не может иметь характеристического числа λ_k, для которого $\lambda_k\neq 1$, а $|\lambda_k|=1$, так как тогда не существовал бы предел $\lim\limits_{q\to\infty}\lambda_k^q$. [Этот же предел должен существовать в силу существования предела (97').]

Мы доказали, что для правильной (и только для правильной) однородной цепи Маркова существует матрица P^∞. Эта матрица определяется формулой (97).

Покажем, как можно выразить матрицу P^∞ через характеристический многочлен

$$\Delta(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\ldots(\lambda-\lambda_u)^{n_u} \tag{98}$$

и присоединенную матрицу $B(\lambda)=(\lambda E-P)^{-1}\Delta(\lambda)$.

Из тождества

$$\frac{B(\lambda)}{\Delta(\lambda)}=\frac{C(\lambda)}{\psi(\lambda)}$$

в силу (95), (95') и (98) вытекает:

$$\frac{n_1B^{(n_1-1)}(1)}{\Delta^{(n_1)}(1)}=\frac{C(1)}{\psi'(1)}.$$

Поэтому формулу (97) можно заменить формулой

$$P^\infty=\frac{n_1B^{(n_1-1)}(1)}{\Delta^{(n_1)}(1)}. \tag{97}$$

Для регулярной цепи Маркова, поскольку она является частным видом правильной цепи, матрица P^∞ существует и определяется любой из формул (97), (97'). В этом случае $n_1=1$ и формула (97') имеет вид

$$P^\infty=\frac{B(1)}{\Delta'(1)}. \tag{99}$$

2. Рассмотрим правильную цепь общего типа (нерегулярную). Соответствующую матрицу P запишем в нормальной форме

$$P=\begin{Vmatrix} Q_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & Q_g & 0 & \cdots & 0 \\ U_{g+1,1} & \cdots & U_{g+1,g} & Q_{g+1} & \cdots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ U_{s1} & \cdots & U_{sg} & U_{s,s-1} & & Q_s \end{Vmatrix}, \tag{100}$$

где $Q_1, \ldots, Q_g$ — примитивные стохастические матрицы, а у неразложимых матриц $Q_{g+1}, \ldots, Q_s$ максимальные характеристические числа <1. Полагая

$$U=\begin{Vmatrix} U_{g+1,1} & \cdots & U_{g+1,g} \\ \cdots & \cdots & \cdots \\ U_{s1} & \cdots & U_{sg} \end{Vmatrix}, \quad W=\begin{Vmatrix} Q_{g+1} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ U_{s,g+1} & \cdots & Q_s \end{Vmatrix},$$

запишем P в виде

$$P=\begin{Vmatrix} Q_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & Q_g & 0 \\ U & & & W \end{Vmatrix}.$$

Тогда

$$P^q=\begin{Vmatrix} Q_1^q & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & Q_g^q & 0 \\ U_q & & & W^q \end{Vmatrix} \tag{101}$$

и

$$P^\infty=\lim_{q\to\infty} P^q=\begin{Vmatrix} Q_1^\infty & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & Q_g^\infty & 0 \\ U_\infty & & & W^\infty \end{Vmatrix}.$$

Но $W^\infty = \lim_{q\to\infty} W^q = 0$, поскольку все характеристические числа матрицы W по модулю меньше единицы. Поэтому

$$P^\infty = \begin{Vmatrix} Q_1^\infty & \dots & 0 & 0 \\ \cdot\cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot\cdot & \cdot \\ 0 & \dots & Q_g^\infty & 0 \\ & U_\infty & & 0 \end{Vmatrix}. \tag{102}$$

Поскольку $Q_1, \dots, Q_g$ — примитивные стохастические матрицы, то матрицы $Q_1^\infty, \dots, Q_g^\infty$ согласно формулам (99) и (35) (стр. 362) положительны:

$$Q_1^\infty > 0, \dots, Q_g^\infty > 0,$$

и в каждом столбце любой из этих матриц все элементы равны между собой:

$$Q_h^\infty = \| q_{*j}^{(h)} \|_{i,\,j=1}^{n} \qquad (h = 1, 2, \dots, g).$$

Заметим, что нормальному виду (100) стохастической матрицы P соответствует разбиение состояний системы $S_1, S_2, \dots, S_n$ на группы:

$$\sum_1, \sum_2, \dots, \sum_g, \sum_{g+1}, \dots, \sum_s. \tag{103}$$

Каждой группе $\sum$ в (104) соответствует своя группа рядов в (101). По терминологии А. Н. Колмогорова[1]) состояния системы, входящие в $\sum_1, \sum_2, \dots, \sum_g$, называются *существенными*, а состояния, входящие в остальные группы $\sum_{g+1}, \dots, \sum_s$ — *несущественными*.

Из вида (101) матрицы P^q следует, что при любом конечном числе шагов q (от момента t_{h-q} к моменту t_h) возможен только переход системы а) из существенного состояния в существенное состояние той же группы, б) из несущественного состояния в существенное состояние и в) из несущественного состояния в несущественное состояние той же или предшествующей группы.

Из вида (102) матрицы P^∞ следует, что в пределе при $q \to \infty$ *переход возможен только из любого состояния в существенное состояние*, т. е. вероятность перехода в любое несущественное состояние при числе шагов $q \to \infty$ стремится к нулю. Поэтому существенные состояния иногда называются и *предельными состояниями*.

3. Из формулы (97) следует:

$$(E - P)\, P^\infty = 0\ ^2).$$

Отсюда видно, что *каждый столбец матрицы P^∞ является собственным вектором стохастической матрицы P для характеристического числа* $\lambda = 1$.

Для регулярной матрицы P число 1 является простым корнем характеристического уравнения и этому числу соответствует только один (с точностью до скалярного множителя) собственный вектор $(1, 1, \dots, 1)$

[1]) См. [96], а также [29], стр. 37—39.

[2]) Эта формула имеет место для произвольной правильной цепи и может быть получена из очевидного равенства $P^q - P \cdot P^{q-1} = 0$ предельным переходом $q \to \infty$.

матрицы P. Поэтому в любом j-м столбце матрицы P^∞ все элементы равны одному и тому же неотрицательному числу p^∞_{*j}:

$$p^\infty_{ij} = p^\infty_{*j} \geqslant 0 \qquad \left(j = 1, 2, \ldots, n; \sum_{j=1}^{n} p_{*j} = 1\right). \tag{104}$$

Таким образом, в регулярной цепи предельные переходные вероятности не зависят от начального состояния.

Обратно, если в некоторой правильной однородной цепи Маркова предельные переходные вероятности не зависят от начального состояния, т. е. имеют место формулы (104), то в схеме (102) для матрицы P^∞ обязательно $g = 1$. Но тогда $n_1 = 1$ и цепь является регулярной.

Для ациклической цепи, которая является частным случаем регулярной цепи, P — примитивная матрица. Поэтому при некотором $q > 0$ $P^q > 0$ (см. теорему 8 на стр. 377). Но тогда и $P^\infty = P^\infty P^q > 0$ [1]).

Обратно, из $P^\infty > 0$ следует, что при некотором $q > 0$ $P^q > 0$, а это по теореме 8 означает примитивность матрицы P и, следовательно, ациклличность данной однородной цепи Маркова.

Полученные результаты мы сформулируем в виде следующей теоремы:

Т е о р е м а 11. 1. *Для того чтобы в однородной цепи Маркова существовали все предельные переходные вероятности, необходимо и достаточно, чтобы цепь была правильной. В этом случае матрица* P^∞, *составленная из предельных переходных вероятностей, определяется формулой* (95) *или* (98).

2. *Для того чтобы в правильной однородной цепи Маркова предельные переходные вероятности не зависели от начального состояния, необходимо и достаточно, чтобы цепь была регулярной. В этом случае матрица* P^∞ *определяется формулой* (99).

3. *Для того чтобы в правильной однородной цепи Маркова все предельные переходные вероятности были отличны от нуля, необходимо и достаточно, чтобы цепь была ациклической* [2]).

4. Введем в рассмотрение столбцы из *абсолютных вероятностей*

$$\overset{k}{p} = (\overset{k}{p_1}, \overset{k}{p_2}, \ldots, \overset{k}{p_n}) \qquad (k = 0, 1, 2, \ldots), \tag{105}$$

где $\overset{k}{p_i}$ — вероятность нахождения системы в момент t_k в состоянии S_i ($i = 1, 2, \ldots, n$; $k = 0, 1, 2, \ldots$). Пользуясь теоремами сложения и умножения вероятностей, найдем:

$$\overset{k}{p_i} = \sum_{h=1}^{n} \overset{0}{p_h} p^{(k)}_{hi} \qquad (i = 1, 2, \ldots, n; \; k = 1, 2, \ldots),$$

или в матричной записи

$$\overset{k}{p} = P'^k \overset{0}{p} \quad (k = 1, 2, \ldots), \tag{106}$$

где P' — транспонированная матрица для матрицы P.

---

[1]) Это матричное равенство получается предельным переходом $m \to \infty$ из равенства $P^m = P^{m-q} \cdot P^q$ $(m > q)$. P^∞ — стохастическая матрица; поэтому $P^\infty \geqslant 0$, и в любой строке матрицы P^∞ имеются ненулевые элементы. Отсюда $P^\infty P^q > 0$. Вместо теоремы 8 можно здесь воспользоваться формулой (99) и неравенством (35) (стр. 362).

[2]) Заметим, что из $P^\infty > 0$ вытекает ациклличность, а следовательно, и регулярность цепи. Поэтому из $P^\infty > 0$ автоматически следует, что предельные переходные вероятности не зависят от начального состояния, т. е. имеют место формулы (104).

Все абсолютные вероятности (105) определяются из формулы (106), если известны начальные вероятности $\overset{0}{p}_1, \overset{0}{p}_2, \ldots, \overset{0}{p}_n$ и матрица переходных вероятностей $P = \| p_{ij} \|_1^n$.

Введем в рассмотрение *предельные абсолютные вероятности*

$$\overset{\infty}{p}_i = \lim_{k\to\infty} \overset{k}{p}_i \qquad (i = 1, 2, \ldots, n)$$

или

$$\overset{\infty}{p} = (\overset{\infty}{p}_1, \overset{\infty}{p}_2, \ldots, \overset{\infty}{p}_n) = \lim_{k\to\infty} \overset{k}{p}.$$

Переходя в обеих частях равенства (106) к пределу при $k \to \infty$, получим:

$$\overset{\infty}{p} = P^{\infty\prime}\, \overset{0}{p}. \tag{107}$$

Заметим, что существование матрицы предельных переходных вероятностей P^∞ влечет существование предельных абсолютных вероятностей $\overset{\infty}{p} = (\overset{\infty}{p}_1, \overset{\infty}{p}_2, \ldots \overset{\infty}{p}_n)$ при любых начальных вероятностях $\overset{0}{p} = (\overset{0}{p}_1, \overset{0}{p}_2, \ldots, \overset{0}{p}_n)$ и наоборот.

Из формулы (107) и из вида (102) матрицы P^∞ вытекает, что *предельные абсолютные вероятности, соответствующие несущественным состояниям, равны нулю.*

Умножая обе части матричного равенства

$$P' \cdot P^{\infty\prime} = P^{\infty\prime}$$

справа на $\overset{0}{p}$, мы в силу (107) получим:

$$P' \overset{\infty}{p} = \overset{\infty}{p}, \tag{108}$$

т. е. столбец предельных абсолютных вероятностей $\overset{\infty}{p}$ *является собственным вектором матрицы* P' *для характеристического числа* $\lambda = 1$.

Если данная цепь Маркова регулярна, то $\lambda = 1$ является простым корнем характеристического уравнения матрицы P'. В этом случае столбец предельных абсолютных вероятностей однозначно определяется из (108) (поскольку $\overset{\infty}{p}_j \geqslant 0$ $(j = 1, 2, \ldots, n)$ и $\sum_{j=1}^{n} \overset{\infty}{p}_j = 1$).

Пусть дана регулярная цепь Маркова. Тогда из (104) и из (107) следует:

$$\overset{\infty}{p}_j = \sum_{h=1}^{n} \overset{0}{p}_h p_{hj}^{\infty} = p_{*j}^{\infty} \sum_{h=1}^{n} \overset{0}{p}_h = p_{*j}^{\infty} \qquad (j = 1, 2, \ldots, n). \tag{109}$$

В этом случае предельные абсолютные вероятности $\overset{\infty}{p}_1, \overset{\infty}{p}_2, \ldots, \overset{\infty}{p}_n$ не зависят от начальных вероятностей $\overset{0}{p}_1, \overset{0}{p}_2, \ldots, \overset{0}{p}_n$.

Обратно, $\overset{\infty}{p}$ может не зависеть от $\overset{0}{p}$ при наличии формулы (107) тогда и только тогда, когда все строки матрицы P^∞ одинаковы, т. е.

$$p_{hj}^{\infty} = p_{*j}^{\infty} \qquad (h, j = 1, 2, \ldots, n),$$

и потому (согласно теореме 11) P — регулярная матрица.

Если P — примитивная матрица, то $P^\infty > 0$, а отсюда в силу (109)

$$\overset{\infty}{p}_j > 0 \qquad (j = 1, 2, \ldots, n).$$

Наоборот, если все $\overset{\infty}{p}_j > 0$ ($j = 1, 2, \ldots, n$) и не зависят от начальных вероятностей, то в каждом столбце матрицы P^∞ все элементы одинаковы и согласно (109) $P^\infty > 0$, а это по теореме 11 означает, что P — примитивная матрица, т. е. данная цепь ациклична.

Из изложенного вытекает, что теорему 11 можно сформулировать так:

Теорема 11′. 1. *Для того чтобы в однородной цепи Маркова существовали все предельные абсолютные вероятности при любых начальных вероятностях, необходимо и достаточно, чтобы цепь была правильной.*

2. *Для того чтобы в однородной цепи Маркова существовали предельные абсолютные вероятности при любых начальных вероятностях и не зависели от этих начальных вероятностей, необходимо и достаточно, чтобы цепь была регулярной.*

3. *Для того чтобы в однородной цепи Маркова при любых начальных вероятностях существовали положительные предельные абсолютные вероятности и эти предельные вероятности не зависели от начальных, необходимо и достаточно, чтобы цепь была ациклической* [1]).

5. Рассмотрим теперь однородную цепь Маркова общего типа с матрицей переходных вероятностей P.

Возьмем нормальную форму (69) матрицы P и обозначим через $h_1, h_2, \ldots, h_g$ индексы imprimitивности матриц $A_1, A_2, \ldots, A_g$ в (69). Пусть h — наименьшее общее кратное целых чисел $h_1, h_2, \ldots, h_g$. Тогда матрица P^h не имеет характеристических чисел, равных по модулю единице, но отличных от единицы, т. е. P^h — правильная матрица; при этом h — наименьший показатель, при котором P^h — правильная матрица. Число h назовем *периодом* данной однородной цепи Маркова.

Поскольку P^h — правильная матрица, то существуют предел

$$\lim_{q\to\infty} P^{hq} = (P^h)^\infty,$$

а значит, и пределы

$$\lim_{q\to\infty} P^{r+qh} = P_r = P^r (P^h)^\infty \qquad (r = 0, 1, \ldots, h-1).$$

Таким образом, в общем случае последовательность матриц

$$P, P^2, P^3, \ldots$$

разбивается на h подпоследовательностей с пределами $P_r = P^r (P^h)^\infty$ ($r = 0, 1, \ldots, h-1$).

Переходя от переходных вероятностей к абсолютным при помощи формулы (106), мы получим, что последовательность

$$\overset{1}{p}, \overset{2}{p}, \overset{3}{p}, \ldots$$

распадается на h подпоследовательностей с пределами

$$\lim_{q\to\infty} \overset{r+qh}{p} = (P'^h)^\infty p^r \qquad (r = 0, 1, 2, \ldots, h-1).$$

---

[1]) Вторую часть теоремы 11′ иногда называют *эргодической*, а первую — *общей квазиэргодической теоремой* для однородных цепей Маркова (см. [4], стр. 473 и 476).

Для произвольной однородной цепи Маркова с конечным числом состояний всегда существуют пределы средних арифметических

$$\widetilde{P} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k = \frac{1}{h} (E + P + \ldots + P^{h-1}) (P^h)^{\infty} \tag{110}$$

и

$$\widetilde{p} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} p^k = \widetilde{P}' p^0. \tag{110'}$$

Здесь $\widetilde{P} = \| \widetilde{p}_{ij} \|_1^n$ и $\widetilde{p} = (\widetilde{p}_1, \widetilde{p}_2, \ldots, \widetilde{p}_n)$. Величины $\widetilde{p}_{ij}$ $(i, j = 1, 2, \ldots, n)$ и $\widetilde{p}_j$ $(j = 1, 2, \ldots, n)$ называются соответственно *средними предельными переходными и средними предельными абсолютными вероятностями.*

Поскольку

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=2}^{N+1} P^k = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k,$$

то

$$\widetilde{P} P = \widetilde{P}$$

и, следовательно, в силу (110')

$$P' \widetilde{p} = \widetilde{p}, \tag{111}$$

т. е. $\widetilde{p}$ — собственный вектор матрицы P' для $\lambda = 1$.

Заметим, что по формулам (69) и (110) мы можем матрицу $\widetilde{P}$ представить в виде

$$\widetilde{P} = \left\| \begin{array}{cccc} \widetilde{A}_1 & 0 & \ldots & 0 \\ 0 & \widetilde{A}_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \widetilde{A}_g \\ & \widetilde{U} & & \end{array} \quad \begin{array}{c} \\ \\ \\ \\ \widetilde{W} \end{array} \right\|,$$

где

$$\widetilde{A}_i = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} A_i^k \quad (i = 1, 2, \ldots, g), \quad \widetilde{W} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} W^k,$$

$$W = \left\| \begin{array}{cccc} A_{g+1} & 0 & \ldots & 0 \\ * & A_{g+2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ * & * & \ldots & A_s \end{array} \right\|.$$

Поскольку все характеристические числа матрицы W по модулю меньше единицы, то

$$\lim_{k\to\infty} W^k = 0$$

и, следовательно, $\widetilde{W} = 0$.

Поэтому

$$\widetilde{P} = \left\| \begin{array}{cccc|c} \widetilde{A}_1 & 0 & \dots & 0 & \\ 0 & \widetilde{A}_2 & \dots & 0 & \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \widetilde{A}_g & \\ \hline & & \widetilde{U} & & 0 \end{array} \right\|. \tag{112}$$

Поскольку $\widetilde{P}$ — стохастическая матрица, то стохастическими являются здесь и матрицы $\widetilde{A}_1, \widetilde{A}_2, \dots, \widetilde{A}_g$.

Из полученного представления для $\widetilde{P}$ и из (107) следует, что *средние предельные абсолютные вероятности, соответствующие несущественным состояниям, всегда равны нулю.*

Если в нормальной форме матрицы P число $g = 1$, то для матрицы P' число $\lambda = 1$ является простым характеристическим числом.

В этом случае $\widetilde{p}$ однозначно определяется из (111), и средние предельные вероятности $\widetilde{p}_1, \widetilde{p}_2, \dots, \widetilde{p}_n$ не зависят от начальных вероятностей $\overset{0}{p}_1, \overset{0}{p}_2, \dots, \overset{0}{p}_n$. Обратно, если $\widetilde{p}$ не зависит от $\overset{0}{p}$, то в силу (110′) матрица $\widetilde{P}$ имеет ранг 1. Но матрица (112) может иметь ранг 1 только тогда, когда $g = 1$.

Полученные результаты мы сформулируем в виде следующей теоремы [1]):

Теорема 12. *Для произвольной однородной цепи Маркова с периодом h матрицы вероятностей P^k и $\overset{k}{p}$ при $k \to \infty$ стремятся к периодическому повторению с периодом h; при этом всегда существуют средние предельные переходные и абсолютные вероятности $\widetilde{P} = \|\widetilde{p}_{ij}\|_1^n$ и $\widetilde{p} = (\widetilde{p}_1, \widetilde{p}_2, \dots, \widetilde{p}_n)$, определяемые формулами* (110) *и* (110′).

Средние предельные абсолютные вероятности, соответствующие несущественным состояниям, всегда равны нулю.

Если в нормальной форме матрицы P число $g = 1$ (и только в этом случае), средние предельные абсолютные вероятности $\widetilde{p}_1, \widetilde{p}_2, \dots, \widetilde{p}_n$ не зависят от начальных вероятностей $\overset{0}{p}_1, \overset{0}{p}_2, \dots, \overset{0}{p}_n$ и однозначно определяются из уравнения (111).

§ 8. Вполне неотрицательные матрицы

В этом и следующем параграфах мы рассмотрим вещественные матрицы, у которых не только элементы, но и все миноры любых порядков неотрицательны. Такие матрицы имеют важные применения в теории малых колебаний упругих систем. Подробное исследование этих матриц и их приложений читатель найдет в книге [7]. Здесь же будут даны только некоторые основные свойства этих матриц.

1. Начнем с определения:

Определение 5. Прямоугольная матрица

$$A = \|a_{ik}\| \quad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, n)$$

[1]) Эту теорему иногда называют *асимптотической теоремой* для однородных цепей Маркова. См. [4], стр. 479—482.

называется *вполне неотрицательной* (*вполне положительной*), если все миноры любых порядков этой матрицы неотрицательны (соответственно положительны):

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \geqslant 0 \text{ (соответственно } > 0)$$

$$\left(1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n;\ p = 1,\ 2,\ \dots,\ \min(m, n)\right).$$

В дальнейшем мы ограничимся рассмотрением только квадратных вполне неотрицательных и вполне положительных матриц.

Пример 1. *Обобщенная матрица Вандермонда*

$$V = \| a_i^{\alpha_k} \|_1^n \quad (0 < a_1 < a_2 < \dots < a_n;\ \alpha_1 < \alpha_2 < \dots < \alpha_n)$$

является вполне положительной. Докажем сначала, что $|V| \neq 0$. В самом деле, из равенства $|V| = 0$ следовало бы, что можно так определить не равные одновременно нулю вещественные числа $c_1, c_2, \dots, c_n$, чтобы функция

$$f(x) = \sum_{k=1}^{n} c_k x^{\alpha_k}$$

имела бы n нулей в точках $x_i = a_i$ $(i = 1, 2, \dots, n)$, где n—число степенных слагаемых. При $n = 1$ это невозможно. Примем индуктивное допущение, что это невозможно для суммы меньшего, нежели n, числа степенных слагаемых, и докажем, что это невозможно и для данной функции $f(x)$. Допустим противное. Тогда по теореме Ролля функция $f_1(x) = [x^{-\alpha_1} f(x)]'$, состоящая из $n-1$ степенных слагаемых, имела бы $n-1$ положительных нулей, а это противоречит допущению индукции.

Итак, $|V| \neq 0$. Но при $\alpha_1 = 0$, $\alpha_2 = 1$, ..., $\alpha_n = n-1$ определитель $|V|$ переходит в обычный определитель Вандермонда $|a_i^{k-1}|_1^n$, который положителен. Так как переход от этого определителя Вандермонда к обобщенному можно осуществить за счет непрерывного изменения показателей $\alpha_1, \alpha_2, \dots, \alpha_n$ с сохранением неравенств между ними $(\alpha_1 < \alpha_2 < \dots < \alpha_n)$, и так как по доказанному определитель при этом не обратится в нуль, то и $|V| > 0$ при любых $(0 <) \alpha_1 < \alpha_2 < \dots < \alpha_n$.

Поскольку любой минор матрицы V может быть рассматриваем как определитель некоторой обобщенной матрицы Вандермонда, то все миноры матрицы V положительны.

Пример 2. Рассмотрим *якобиеву матрицу*

$$J = \begin{Vmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ c_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & c_2 & a_3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{Vmatrix}, \tag{113}$$

т. е. матрицу, в которой все элементы, не принадлежащие главной диагонали, первой наддиагонали и первой поддиагонали, равны нулю. Установим формулу, выражающую произвольный минор этой матрицы через главные миноры и элементы b, c. Пусть

$$1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n$$

и

$$i_1 = k_1,\ i_2 = k_2,\ \dots,\ i_{\nu_1} = k_{\nu_1};\ i_{\nu_1+1} \neq k_{\nu_1+1},\ \dots,\ i_{\nu_2} \neq k_{\nu_2};\ i_{\nu_2+1} = k_{\nu_2+1},\ \dots,\ i_{\nu_3} = k_{\nu_3};\ \dots$$

Тогда

$$J\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = J\begin{pmatrix} i_1 & \dots & i_{\nu_1} \\ k_1 & \dots & k_{\nu_1} \end{pmatrix} J\begin{pmatrix} i_{\nu_1+1} \\ k_{\nu_1+1} \end{pmatrix} \dots J\begin{pmatrix} i_{\nu_2} \\ k_{\nu_2} \end{pmatrix} J\begin{pmatrix} i_{\nu_2+1} & \dots & i_{\nu_3} \\ k_{\nu_2+1} & \dots & k_{\nu_3} \end{pmatrix} \dots \quad (114)$$

Справедливость этой формулы вытекает из легко проверяемого равенства:

$$J\begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix} = J\begin{pmatrix} i_1 & \dots & i_{\nu-1} \\ k_1 & \dots & k_{\nu-1} \end{pmatrix} J\begin{pmatrix} i_\nu \\ k_\nu \end{pmatrix} J\begin{pmatrix} i_{\nu+1} & \dots & i_p \\ k_{\nu+1} & \dots & k_p \end{pmatrix} \text{ (при } l_\nu \neq k_\nu). \quad (115)$$

Из формулы (114) следует, что любой минор равен произведению некоторых главных миноров и некоторых элементов матрицы J. Таким образом, *для того чтобы матрица J была вполне неотрицательной, необходимо и достаточно, чтобы все ее главные миноры и элементы b, с были неотрицательны.*

2. Для вполне неотрицательной матрицы $A = \| a_{ik} \|_1^n$ всегда имеет место следующее важное детерминантное неравенство[1]):

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \leqslant A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} \qquad (p < n). \quad (116)$$

Выводу этого неравенства предпошлем следующую лемму:

Лемма 5. *Если во вполне неотрицательной матрице $A = \| a_{ik} \|_1^n$ какой-либо главный минор равен нулю, то равен нулю любой «объемлющий» главный минор.*

Доказательство. Лемма будет доказана, если мы покажем, что для вполне неотрицательной матрицы $A = \| a_{ik} \|_1^n$ из

$$A\begin{pmatrix} 1 & 2 & \dots & q \\ 1 & 2 & \dots & q \end{pmatrix} = 0 \quad (q < n) \quad (117)$$

всегда следует:

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = 0. \quad (118)$$

При этом рассмотрим два случая:

1) $a_{11} = 0$. Поскольку $\begin{vmatrix} a_{11} & a_{1k} \\ a_{i1} & a_{ik} \end{vmatrix} = -a_{i1}a_{1k} \geqslant 0$, $a_{i1} \geqslant 0$, $a_{1k} \geqslant 0$ $(i, k = 2, \dots, n)$, то либо все $a_{i1} = 0$ $(i = 2, \dots, n)$, либо все $a_{1k} = 0$ $(k = 2, \dots, n)$. Из этих равенств и из $a_{11} = 0$ следует (118).

2) $a_{11} \neq 0$. Тогда при некотором p $(1 \leqslant p \leqslant q)$

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} \neq 0, \quad A\begin{pmatrix} 1 & 2 & \dots & p-1 & p \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} = 0. \quad (119)$$

Введем окаймляющие определители

$$d_{ik} = A\begin{pmatrix} 1 & 2 & \dots & p-1 & i \\ 1 & 2 & \dots & p-1 & k \end{pmatrix} \quad (i, k = p, p+1, \dots, n). \quad (120)$$

---

[1]) См. [82в], а также [7], стр. 111 и далее. Там же выяснено, что знак равенства в (116) может иметь место только лишь в следующих очевидных случаях:

1) один из множителей в правой части (116) равен нулю,

2) все элементы a_{ik} $(i = 1, 2, \dots, p;\ k = p+1, \dots, n)$ либо все элементы a_{ik} $(i = p+1, \dots, n;\ k = 1, 2, \dots, p)$ равны нулю.

Неравенство (116) имеет тот же внешний вид, что и обобщенное неравенство Адамара [см. (29) на стр. 231] для положительно определенной эрмитовой или квадратичной формы.

Из них составим матрицу $D = \| d_{ik} \|_p^n$.

Согласно тождеству Сильвестра (гл. II, § 3)

$$D\begin{pmatrix} i_1 & i_2 & \dots & i_g \\ k_1 & k_2 & \dots & k_g \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{g-1} A\begin{pmatrix} 1 & 2 & \dots & p-1 & i_1 & i_2 & \dots & i_g \\ 1 & 2 & \dots & p-1 & k_1 & k_2 & \dots & k_g \end{pmatrix} \geqslant 0 \quad (121)$$

$$\left(p \leqslant \begin{matrix} i_1 < i_2 < \dots < i_g \\ k_1 < k_2 < \dots < k_g \end{matrix} \leqslant n;\ g = 1, 2, \dots, n-p+1\right)$$

и поэтому D — вполне неотрицательная матрица.

Поскольку в силу (119)

$$d_{pp} = A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = 0,$$

то матрица $D = \| d_{ik} \|_p^n$ подходит под уже разобранный случай 1) и

$$D\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p} A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = 0.$$

Отсюда, поскольку $A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} \neq 0$, следует (118). Лемма доказана.

3. Теперь мы имеем возможность при выводе неравенства (116) предполагать, что все главные миноры матрицы A отличны от нуля, так как согласнно лемме 5 равенство нулю одного из главных миноров возможно лишь тогда, когда $|A| = 0$, а в этом случае неравенство (116) очевидно.

При $n = 2$ справедливость неравенства (116) проверяется непосредственно:

$$A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \leqslant a_{11}a_{22},$$

поскольку $a_{12} \geqslant 0$, $a_{21} \geqslant 0$. Будем устанавливать неравенство (116) для $n > 2$, предполагая, что оно уже справедливо для матриц порядка $< n$. Кроме того, не нарушая общности, можем считать, что $p > 1$, так как в противном случае за счет обратной нумерации строк и столбцов мы поменяли бы ролями числа p и $n-p$.

Вводя снова в рассмотрение матрицу $D = \| d_{ik} \|_p^n$, где d_{ik} $(i, k = p, p+1, \dots, n)$ определяются формулами (120), используя дважды тождество Сильвестра и основное неравенство (116) для матриц порядка $< n$, имеем

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = \frac{D\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix}}{\left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p}} \leqslant \frac{d_{pp} D\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}}{\left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p}} =$$

$$= \frac{A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} A\begin{pmatrix} 1 & 2 & \dots & p-1 & p+1 & \dots & n \\ 1 & 2 & \dots & p-1 & p+1 & \dots & n \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}} \leqslant$$

$$\leqslant A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}. \quad (122)$$

Таким образом, неравенство (116) можно считать установленным.

Введем следующее

Определение 6. Минор

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}\left(1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n\right) \tag{123}$$

матрицы $A = \| a_{ik} \|_1^n$ будем называть *почти главным*, если среди разностей $i_1 - k_1$, $i_2 - k_2$, ..., $i_p - k_p$ только одна разность не равна нулю.

Обращаем внимание на то, что весь вывод неравенства (116) (в том числе и доказательство вспомогательной леммы) сохраняет свою силу, если условие «A — вполне неотрицательная матрица» заменить более слабым условием «в матрице A неотрицательны все главные и почти главные миноры»[1]).

§ 9. Осцилляционные матрицы

1. Характеристические числа и собственные векторы вполне положительных матриц обладают целым рядом замечательных свойств. Однако класс вполне положительных матриц недостаточно широк с точки зрения приложений к малым колебаниям упругих систем. В этом отношении класс вполне неотрицательных матриц имеет уже достаточный объем. Но не для всех вполне неотрицательных матриц имеют место нужные нам спектральные свойства. Однако существует промежуточный класс (между классами вполне положительных и вполне неотрицательных матриц), в котором сохраняются спектральные свойства вполне положительных матриц и который достаточно широк для охвата приложений. Матрицы этого промежуточного класса получили название «осцилляционных». Это название связано с тем, что осцилляционные матрицы образуют математический аппарат для исследования осцилляционных свойств малых колебаний линейных упругих систем[2]).

Определение 7. Матрица $A = \| a_{il} \|_1^n$ называется *осцилляционной*, если A—вполне неотрицательная матрица и существует такое целое число $q > 0$, что A^q—вполне положительная матрица.

Пример. Якобиева матрица J [см. (113)] является осцилляционной в том и только в том случае, когда 1° все числа b, c положительны и 2° последовательные главные миноры положительны:

$$a_1 > 0,\ \begin{vmatrix} a_1 & b_1 \\ c_1 & a_2 \end{vmatrix} > 0,\ \begin{vmatrix} a_1 & b_1 & 0 \\ c_1 & a_2 & b_2 \\ 0 & c_2 & a_3 \end{vmatrix} > 0,\ \dots,\ \begin{vmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ c_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & c_2 & a_3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{vmatrix} > 0. \tag{124}$$

[1]) См. [97б]. Мы пользуемся случаем отметить, что в книге Ф. Р. Гантмахера и М. Г. Крейна [7] (во втором издании) в этом пункте допущена ошибка, на которую первый обратил внимание авторов Д. М. Котелянский. В книге [7], стр. 111 почти главный минор (123) определяется равенством

$$\sum_{\nu=1}^{p} | i_\nu - k_\nu | = 1.$$

При таком определении почти главного минора из неотрицательности главных и почти главных миноров не следует еще неравенство (116). Однако все формулировки и доказательства § 6 гл. II в [7], посвященного основному неравенству, становятся верными, если для почти главного минора дать приведенное здесь определение, взятое нами из работы [97б].

[2]) См. [7], Введение, а также гл. III, IV.

Необходимость условий 1°, 2°. Числа b, c неотрицательны, поскольку матрица $J \geqslant 0$. При этом ни одно из чисел b, c не может равняться нулю, так как в противном случае матрица была бы разложимой, и тогда при любом $q > 0$ неравенство $J^q > 0$ не соблюдалось бы. Следовательно, все числа b, c положительны. Все главные миноры (124) положительны согласно лемме 5, поскольку из $|J| \geqslant 0$ и $|J^q| > 0$ следует: $|J| > 0$.

Достаточность условий 1°, 2°. Раскрывая $|J|$, легко убеждаемся в том, что числа b, c входят в состав $|J|$ только произведениями b_1c_1, b_2c_2, ..., $b_{n-1}c_{n-1}$. Это же относится к любому главному минору «нулевой плотности», т. е. минору, образованному подряд идущими (без пропусков) строками и столбцами. Но любой главный минор матрицы J распадается в произведение главных миноров нулевой плотности. Поэтому *в любой главный минор матрицы J числа b и c входят только произведениями b_1c_1, b_2c_2, ..., $b_{n-1}c_{n-1}$.*

Составим симметрическую якобиеву матрицу

$$\widetilde{J} = \begin{Vmatrix} a_1 & \widetilde{b}_1 & & & 0 \\ \widetilde{b}_1 & a_2 & \widetilde{b}_2 & & \\ & \widetilde{b}_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \widetilde{b}_{n-1} \\ 0 & & & \widetilde{b}_{n-1} & a_n \end{Vmatrix}, \quad \widetilde{b}_i = \sqrt{b_i c_i} > 0 \quad (i = 1, 2, \ldots, n). \tag{125}$$

Из установленного выше свойства главных миноров якобиевой матрицы следует, что соответствующие главные миноры матриц J и $\widetilde{J}$ равны друг другу. Но тогда условия (124) означают, что квадратичная форма

$$\widetilde{J}(x, x)$$

является положительно определенной (см. гл. X, теорема 3, стр. 277). Но у положительно определенной квадратичной формы все главные миноры положительны. Следовательно, и в матрице J все главные миноры положительны. Поскольку по условию 1° все числа b, c положительны, то по формуле (114) все миноры матрицы J неотрицательны, т. е. J — вполне неотрицательная матрица.

Осцилляционность вполне неотрицательной матрицы J, для которой выполняются условия 1°, 2°, вытекает непосредственно из следующего критерия осцилляционности.

Для того чтобы вполне неотрицательная матрица $A = \| a_{ik} \|_1^n$ была осцилляционной, необходимо и достаточно, чтобы выполнялись условия:

1) A — неособенная матрица ($|A| > 0$),

2) все элементы матрицы A, расположенные на главной диагонали, на первой наддиагонали и на первой поддиагонали, отличны от нуля ($a_{ik} > 0$ при $|i - k| \leqslant 1$).

Доказательство этого предложения читатель найдет в [7], глава II, § 7.

2. Для того чтобы сформулировать свойства характеристических чисел и собственных векторов осцилляционной матрицы, введем некоторые предварительные понятия и обозначения.

Рассмотрим вектор (столбец)

$$u = (u_1, u_2, \ldots, u_n).$$

Будем подсчитывать число перемен знака в ряду координат $u_1, u_2, \ldots u_n$ вектора u, приписывая нулевым координатам (если таковые имеются) произвольные знаки. В зависимости от того, какие мы знаки припишем нулевым координатам, число перемен знака будет колебаться в известных пределах. Получающиеся при этом *максимальное* и *минимальное* числа перемен знака будем обозначать соответственно через S_u^+ и S_u^-. В том случае, когда $S_u^- = S_u^+$, мы будем говорить о *точном* числе перемен знака и обозначать его через S_u. Очевидно, $S_u^- = S_u^+$ тогда и только

тогда, когда 1° крайние координаты u_1 и u_n вектора u отличны от нуля, и 2° равенство $u_i=0$ $(1<i<n)$ всегда сопровождается неравенством $u_{i-1}u_{i+1}<0$.

Теперь мы докажем следующую основную теорему:

Теорема 13.1. *Осцилляционная матрица* $A=\|a_{ik}\|_1^n$ *всегда имеет* n *различных положительных характеристических чисел*

$$\lambda_1>\lambda_2>\ldots>\lambda_n>0. \tag{126}$$

2. *У собственного вектора* $\overset{1}{u}=(u_{11}, u_{21}, \ldots, u_{n1})$ *матрицы* A, *отвечающего наибольшему характеристическому числу* λ_1, *все координаты отличны от нуля и одного знака; у собственного вектора* $\overset{2}{u}=(u_{12}, u_{22}, \ldots, u_{n2})$, *отвечающего второму по величине характеристическому числу* λ_2, *в ряду координат имеется точно одна перемена знака и вообще в ряду координат собственного вектора* $\overset{k}{u}=(u_{1k}, u_{2k} \ldots, u_{nk})$, *соответствующего характеристическому числу* λ_k, *имеется точно* $k-1$ *перемен знака* $(k=1, 2, \ldots, n)$.

3. *При любых вещественных числах* $c_g, c_{g+1}, \ldots, c_h$ $\left(1\leqslant g\leqslant h\leqslant n; \sum\limits_{k=g}^{h} c_k^2>0\right)$ *в ряду координат вектора*

$$u=\sum_{k=g}^{h} c_k \overset{k}{u} \tag{127}$$

число перемен знака заключается между $g-1$ *и* $h-1$

$$g-1\leqslant S_u^-\leqslant S_u^+\leqslant h-1. \tag{128}$$

Доказательство. 1. Занумеруем характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ матрицы A так, чтобы

$$|\lambda_1|\geqslant|\lambda_2|\geqslant\ldots\geqslant|\lambda_n|,$$

и введем в рассмотрение p-ю ассоциированную матрицу $\mathfrak{A}_p$ $(p=1, 2, \ldots, n$; см. гл. I, § 4). Характеристическими числами матрицы $\mathfrak{A}_p$ являются всевозможные произведения по p из характеристических чисел матрицы A (см. стр. 86), т. е. произведения

$$\lambda_1\lambda_2\ldots\lambda_p,\ \lambda_1\lambda_2\ldots\lambda_{p-1}\lambda_{p+1},\ \ldots$$

Из условий теоремы следует, что при некотором целом q степень A^q — вполне положительная матрица. Но тогда $\mathfrak{A}_p\geqslant 0$, $\mathfrak{A}_p^q>0$[1]), т. е. $\mathfrak{A}_p$ — неразложимая неотрицательная и притом примитивная матрица. Применяя теорему Фробениуса (см. § 2, стр. 355) к примитивной матрице $\mathfrak{A}_p$ $(p=1, 2, \ldots, n)$, получим:

$$\lambda_1\lambda_2\ldots\lambda_p>0 \quad (p=1, 2, \ldots, n),$$

$$\lambda_1\lambda_2\ldots\lambda_p>\lambda_1\lambda_2\ldots\lambda_{p-1}\lambda_{p+1} \quad (p=1, 2, \ldots, n-1).$$

Отсюда вытекают неравенства (125).

2. Из установленных неравенств (125) вытекает, что $A=\|a_{ik}\|_1^n$ — матрица простой структуры. Тогда и все ассоциированные матрицы $\mathfrak{A}_p$ $(p=1, 2, \ldots, n)$ будут иметь простую структуру (см. стр. 86).

[1]) Матрица $\mathfrak{A}_p^q$ является p-й ассоциированной матрицей для A^q (см. гл. I, стр. 31).

Введем в рассмотрение фундаментальную матрицу $U=\|u_{ik}\|_1^n$ для матрицы A ($\underset{k}{u}$ в k-м столбце матрицы U стоят координаты k-го собственного вектора u матрицы A; $k=1, 2, \ldots, n$). Тогда (см. гл. III, стр. 86) характеристическому числу $\lambda_1\lambda_2 \ldots \lambda_p$ матрицы $\mathfrak{A}_p$ будет соответствовать собственный вектор с координатами

$$U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} \quad (1\leqslant i_1<i_2<\ldots<i_p\leqslant n). \tag{129}$$

По теореме Фробениуса все числа (129) отличны от нуля и одного знака. Умножая векторы $\underset{1}{u}, \underset{2}{u}, \ldots, \underset{n}{u}$ на ± 1, можно сделать все миноры (129) положительными:

$$U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix}>0 \quad \begin{pmatrix} 1\leqslant i_1<i_2<\ldots<i_p\leqslant n; \\ p=1, 2, \ldots, n \end{pmatrix}. \tag{130}$$

Фундаментальная матрица $U=\|u_{ik}\|_1^n$ для матрицы A связана с матрицей A равенством

$$A=U\{\lambda_1, \lambda_2, \ldots, \lambda_n\}U^{-1}. \tag{131}$$

Но тогда

$$A'=U'^{-1}\{\lambda_1,\lambda_2, \ldots, \lambda_n\}U'. \tag{132}$$

Сопоставляя (131) с (132), мы видим, что матрица

$$V=U'^{-1} \tag{133}$$

является фундаментальной для транспонированной матрицы A' при тех же характеристических числах $\lambda_1, \lambda_2, \ldots, \lambda_n$. Но из осцилляционности матрицы A следует осцилляционность транспонированной матрицы A'. Поэтому и для матрицы V при любом $p=1, 2, \ldots, n$ все миноры

$$V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} \quad (1\leqslant i_1<i_2<\ldots<i_p\leqslant n) \tag{134}$$

отличны от нуля и имеют один и тот же знак.

С другой стороны, согласно (133) матрицы U и V связаны равенством

$$U'V=E.$$

Переходя к p-м ассоциированным матрицам (см. гл. I, § 4), будем иметь:

$$\mathfrak{U}_p\mathfrak{V}_p=\mathfrak{E}_p.$$

Отсюда, в частности, записывая, что диагональный элемент матрицы $\mathfrak{E}_p$ равен единице, получим:

$$\sum_{1\leqslant i_1<i_2<\ldots<i_p\leqslant n} U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix}=1. \tag{135}$$

В левой части этого равенства первые множители в слагаемых положительны, а вторые — отличны от нуля и одного знака. Тогда очевидно, что и вторые сомножители положительны, т. е.

$$V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix}>0 \quad \begin{pmatrix} 1\leqslant i_1<i_2<\ldots<i_p\leqslant n; \\ p=1, 2, \ldots, n \end{pmatrix}. \tag{136}$$

Таким образом, для матриц $U=\|u_{ik}\|_1^n$ и $V=U'^{-1}$ одновременно имеют место неравенства (130) и (136).

Выражая миноры матрицы V через миноры обратной матрицы $V^{-1}=U'$ по известным формулам [см. стр. 31], получим:

$$V\begin{pmatrix} j_1\ j_2\ \dots\ j_{n-p} \\ 1\ 2\ \dots\ n-p \end{pmatrix}=\frac{(-1)^{np+\sum_{\nu=1}^{p} i_\nu}}{|U|}U\begin{pmatrix} i_1\ i_2\ \dots\ i_p \\ n\ n-1\ \dots\ n-p+1 \end{pmatrix}, \quad (137)$$

где $i_1<i_2<\dots<i_p$ и $j_1<j_2<\dots<j_{n-p}$ вместе дают полную систему индексов $1, 2, \dots, n$. Так как в силу (130) $|U|>0$, то из (136) и (137) вытекает:

$$(-1)^{np+\sum_{\nu=1}^{p} i_\nu}U\begin{pmatrix} i_1\ i_2\ \dots\ i_p \\ 1\ 2\ \dots\ p \end{pmatrix}>0 \quad \begin{pmatrix} 1\leqslant i_1<i_2<\dots<i_p\leqslant n; \\ p=1, 2, \dots, n \end{pmatrix}. \quad (138)$$

Пусть теперь $u=\sum_{k=g}^{h} c_k{}^k u \left(\sum_{k=g}^{h} c_k^2>0\right)$. Мы покажем, что из неравенств (130) следует вторая часть неравенства (128):

$$S_u^+\leqslant h-1, \quad (139)$$

а из неравенств (138) — первая:

$$S_u^-\geqslant g-1. \quad (140)$$

Допустим, что $S_u^+>h-1$. Тогда можно указать такие $h+1$ координат вектора u

$$u_{i_1}, u_{i_2}, \dots, u_{i_{h+1}} (1\leqslant i_1<i_2<\dots<i_{h+1}\leqslant n), \quad (141)$$

что

$$u_{i_\alpha}u_{i_{\alpha+1}}\leqslant 0 \quad (\alpha=1, 2, \dots, h).$$

При этом координаты (141) не могут все одновременно равняться нулю, так как тогда, приравнивая нулю соответствующие координаты вектора $u=\sum_{k=1}^{h} c_k{}^k u \left(c_1=\dots=c_{g-1}=0; \sum_{k=1}^{h} c_k^2>0\right)$, мы получили бы систему однородных уравнений

$$\sum_{k=1}^{h} c_k u_{i_\alpha{}^k}=0 \quad (\alpha=1, 2, \dots, h)$$

с ненулевым решением $c_1, c_2, \dots, c_h$; в то же время определитель этой системы

$$U\begin{pmatrix} i_1\ i_2\ \dots\ i_h \\ 1\ 2\ \dots\ h \end{pmatrix}$$

согласно (130) отличен от нуля.

Рассмотрим теперь равный нулю определитель

$$\begin{vmatrix} u_{i_1 1} & \dots & u_{i_1 h} & u_{i_1} \\ u_{i_2 1} & \dots & u_{i_2 h} & u_{i_2} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ u_{i_{h+1} 1} & \dots & u_{i_{h+1} h} & u_{i_{h+1}} \end{vmatrix}=0.$$

Раскроем его по элементам последней вертикали:

$$\sum_{\alpha=1}^{h+1}(-1)^{h+\alpha+1}u_{i_\alpha}U\begin{pmatrix}i_1 & \dots & i_{\alpha-1} & i_{\alpha+1} & \dots & i_{h+1};\\ 1 & \dots & \dots & \dots & \dots & h\end{pmatrix}=0.$$

Но такое равенство не может иметь места, так как в левой части нет двух слагаемых разных знаков и по крайней мере одно слагаемое отлично от нуля. Таким образом, допущение $S_u^+ > h-1$ привело нас к противоречию, и неравенство (139) можно считать установленным.

Введем в рассмотрение векторы

$$\overset{k}{u^*}=(u_{1k}^*,\ u_{2k}^*,\ \dots,\ u_{nk}^*) \qquad (k=1,\ 2,\ \dots,\ n),$$

где

$$u_{ik}^*=(-1)^{n+i+k}u_{ik} \qquad (i,\ k=1,\ 2,\ \dots,\ n);$$

тогда для матрицы $U^*=\|u_{ik}^*\|_1^n$ в силу (138) будем иметь:

$$U^*\begin{pmatrix}i_1 & i_2 & \dots & i_p\\ n & n-1 & \dots & n-p+1\end{pmatrix}>0\begin{pmatrix}1\leqslant i_1<i_2<\dots<i_p\leqslant n;\\ p=1,\ 2,\ \dots,\ n\end{pmatrix}. \tag{142}$$

Но неравенства (142) аналогичны неравенствам (130). Поэтому, полагая

$$u^*=\sum_{k=g}^{h}(-1)^k c_k \overset{k}{u^*}, \tag{143}$$

будем иметь неравенство, аналогичное неравенству (139)[1]:

$$S_{u^*}^+\leqslant n-g. \tag{144}$$

Пусть $u=(u_1,\ u_2,\ \dots,\ u_n)$, а $u^*=(u_1^*,\ u_2^*,\ \dots,\ u_n^*)$. Легко видеть, что

$$u_i^*=(-1)^i u_i \qquad (i=1,\ 2,\ \dots,\ n).$$

Поэтому

$$S_{u^*}^+ + S_u^- = n-1$$

и, следовательно, в силу (144) имеет место соотношение (140).

Неравенство (128) установлено. Поскольку из него получается при $g=h=k$ утверждение 2 теоремы, то теорема доказана полностью.

3. Рассмотрим применение доказанной теоремы к исследованию малых колебаний n масс $m_1,\ m_2,\ \dots,\ m_n$, сосредоточенных в n подвижных точках $x_1<x_2<\dots<x_n$ сегментного упругого континуума (струна или стержень конечной длины), простирающегося (в состоянии равновесия) вдоль отрезка $0\leqslant x\leqslant l$ оси x.

Обозначим через $K(x,\ s)$ $(0\leqslant x,\ s\leqslant l)$ функцию влияния этого континуума [$K(x,\ s)$ — прогиб в точке x под действием единичной силы, приложенной в точке s], а через k_{ij} — коэффициенты влияния для данных n масс:

$$k_{ij}=K(x_i,\ x_j) \qquad (i,\ j=1,\ 2,\ \dots,\ n).$$

---

[1]) В неравенствах (142) в векторы $\overset{k}{u}$ $(k=1,\ 2,\ \dots,\ n)$ идут в обратном порядке $u,\ \overset{n-1}{u},\ \dots,\ \overset{g}{u}$. Вектору $\overset{g}{u}$ предшествует $n-g$ векторов этого ряда.

Если в точках $x_1, x_2, \ldots, x_n$ приложены n сил $F_1, F_2, \ldots, F_n$, то соответствующий статический прогиб $y(x)$ $(0 \leqslant x \leqslant l)$ в силу линейного наложения прогибов выразится формулой

$$y(x) = \sum_{j=1}^{n} K(x, x_j) F_j.$$

Заменяя здесь силы F_j силами инерции $-m_j \frac{\partial^2}{\partial t^2} y(x_j; t)$ $(j = 1, 2, \ldots, n)$, получим уравнение свободных колебаний

$$y(x) = -\sum_{j=1}^{n} m_j K(x, x_j) \frac{\partial^2}{\partial t^2} y(x_j; t). \tag{145}$$

Будем искать гармонические колебания континуума в виде

$$y(x) = u(x) \sin(\omega t + \alpha) \qquad (0 \leqslant x \leqslant l). \tag{146}$$

Здесь $u(x)$ — амплитудная функция, ω — частота, α — начальная фаза. Подставляя это выражение для $y(x)$ в (145) и сокращая на $\sin(\omega t + \alpha)$, получим:

$$u(x) = \omega^2 \sum_{j=1}^{n} m_j K(x, x_j) u(x_j). \tag{147}$$

Введем обозначения для переменных прогибов и для амплитудных прогибов в точках расположения масс:

$$y_i = y(x_i, t), \quad u_i = u(x_i) \qquad (i = 1, 2, \ldots, n).$$

Тогда

$$y_i = u_i \sin(\omega t + \alpha) \qquad (i = 1, 2, \ldots, n).$$

Введем еще *приведенные амплитудные прогибы и приведенные коэффициенты влияния*

$$\tilde{u}_i = \sqrt{m_i}\, u_i, \quad a_{ij} = \sqrt{m_i m_j}\, k_{ij} \qquad (i, j = 1, 2, \ldots, n). \tag{148}$$

Заменяя в (147) x последовательно на x_i $(i = 1, 2, \ldots, n)$, получим систему уравнений для амплитудных прогибов:

$$\sum_{j=1}^{n} a_{ij} \tilde{u}_j = \lambda \tilde{u}_i \qquad \left(\lambda = \frac{1}{\omega^2};\ i = 1, 2, \ldots, n\right). \tag{149}$$

Отсюда видно, что амплитудный вектор $\tilde{u} = (\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_n)$ есть собственный вектор матрицы $A = \| a_{ij} \|_1^n = \| \sqrt{m_i m_j}\, k_{ij} \|_1^n$ при $\lambda = \frac{1}{\omega^2}$ (ср. с гл. X, § 8).

В результате подробного анализа[1]) устанавливается, что *матрица коэффициентов влияния* $\| k_{ij} \|_1^n$ *сегментного континуума всегда является осцилляционной матрицей*. Но тогда и матрица $A = \| a_{ij} \|_1^n = \| \sqrt{m_i m_j} k_{ij} \|_1^n$ является осцилляционной! Поэтому матрица A (согласно теореме 13) имеет n положительных характеристических чисел

$$\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0,$$

[1]) См. [10д, е] и [7], гл. III.

т. е. существует n гармонических колебаний континуума с *различными* частотами:

$$(0<)\,\omega_1<\omega_2<\ldots<\omega_n \qquad \left(\lambda_i=\frac{1}{\omega_i^2};\quad i=1,\ 2,\ \ldots,\ n\right).$$

В силу той же теоремы основному тону с частотой ω_1 соответствуют амплитудные прогибы, отличные от нуля и одного знака. В ряду амплитудных прогибов, отвечающих первому обертону с частотой ω_2, имеется точно одна перемена знака и вообще в ряду амплитудных прогибов для обертона с частотой ω_j имеется точно $j-1$ перемен знака ($j=1,\ 2,\ \ldots,\ n$).

Из того факта, что матрица коэффициентов влияния $\|k_{ij}/\|_1^n$ осцилляционна, вытекают и другие осцилляционные свойства континуума: 1) при $\omega=\omega_1$ амплитудная функция $u(x)$, связанная с амплитудными прогибами формулой (147), не имеет узлов; и вообще при $\omega=\omega_j$ эта функция имеет $j-1$ узлов ($j=1,\ 2,\ \ldots,\ n$); 2) узлы двух смежных тонов перемежаются и т. д.

На обосновании этих свойств мы не можем здесь останавливаться[1]).

[1]) См. [7], гл. III, IV.

ГЛАВА XIV

РАЗЛИЧНЫЕ КРИТЕРИИ РЕГУЛЯРНОСТИ И ЛОКАЛИЗАЦИЯ СОБСТВЕННЫХ ЗНАЧЕНИЙ

§ 1. Критерий регулярности Адамара и его обобщения

Пусть $A=\|a_{ik}\|_1^n$ — произвольная $n\times n$-матрица с комплексными элементами. Допустим, что эта матрица вырождена, т. е. $|A|=0$. Тогда существуют такие числа $x_1, x_2, \ldots, x_n$ с максимальным $|x_k|>0$, что[1])

$$\sum_{j=1}^{n} a_{kj}x_j=0. \tag{1}$$

Но тогда

$$|a_{kk}|\,|x_k|\leqslant \sum_{\substack{j=1\\ j\neq k}}^{n}|a_{kj}|\,|x_j|\leqslant |x_k|\sum_{\substack{j=1\\ j\neq k}}^{n}|a_{kj}|.$$

Сокращая на $|x_k|$, получаем:

$$|a_{kk}|\leqslant \sum_{\substack{j=1\\ j\neq k}}^{n}|a_{kj}|. \tag{2}$$

Поэтому, если выполняются *условия Адамара*

$$H_i\equiv |a_{ii}|-\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|>0 \qquad (i=1,\ \ldots,\ n), \tag{3}$$

то неравенство типа (2) невозможно и, следовательно, матрица A является *регулярной* (невырожденной), т. е. $|A|\neq 0$.

Таким образом, справедлива следующая

Теорема 1 (Адамара). *Если для матрицы* $A=\|a_{ik}\|_1^n$ *выполняются* n *неравенств* (3), *то матрица* A *является невырожденной.*

Условие $H_i>0$ означает, что модуль диагонального элемента a_{ii} превосходит (строго!) сумму модулей всех остальных элементов i-й строки. Такой элемент a_{ii} называется доминирующим (для своей строчки). Условия Адамара требуют, чтобы все диагональные элементы матрицы A были *доминирующими* (для своих строк).

Замечание 1. Если выполняются условия Адамара (3), то для $\operatorname{mod}|A|$ справедлива следующая оценка снизу:

$$\operatorname{mod}|A|\geqslant H_1H_2\ldots H_n>0. \tag{4}$$

---

[1]) Равенства (1) выполняются при любом $k=1, \ldots, n$. Мы же берем только то значение k, при котором $|x_k|$ является максимальным.

Для того чтобы убедиться в справедливости неравенства (4), введем вспомогательную матрицу $F = \| f_{ij} \|_1^n$, где

$$f_{ij} = \frac{a_{ij}}{H_i} \qquad (i,\ j = 1,\ \ldots,\ n), \tag{5}$$

для которой, очевидно,

$$|f_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^{n} |f_{ij}| = 1 \qquad (i = 1,\ \ldots,\ n). \tag{6}$$

Обозначим через λ_0 какое-либо характеристическое число этой матрицы. Числу λ_0 соответствует собственный вектор $(x_1,\ x_2,\ \ldots,\ x_n)$ с максимальным $|x_k| > 0$. Тогда

$$\lambda_0 x_k = \sum_{j=1}^{n} f_{kj} x_j. \tag{7}$$

Из этого равенства с учетом соотношений (6) получаем

$$|\lambda_0|\,|x_k| \geqslant |f_{kk}|\,|x_k| - \sum_{\substack{j=1 \\ j \neq k}}^{n} |f_{kj}|\,|x_j| \geqslant |x_k| \left(|f_{kk}| - \sum_{\substack{j=1 \\ j \neq k}}^{n} |f_{kj}| \right) = |x_k|.$$

Сокращая на $|x_k|$, найдем:

$$|\lambda_0| \geqslant 1.$$

Но определитель $|F|$ равен произведению характеристических чисел матрицы F. Каждое из них по модулю $\geqslant 1$. Поэтому и

$$\operatorname{mod} |F| \geqslant 1. \tag{8}$$

С другой же стороны,

$$|F| = \frac{|A|}{H_1 H_2 \ldots H_n}. \tag{9}$$

Из (8) и (9) сразу следует искомое неравенство (4).

Заметим еще, что для всего класса матриц, удовлетворяющих условиям Адамара с заданными значениями $H_1,\ \ldots,\ H_n$, оценка (4) не может быть улучшена, так как неравенство (4) переходит в равенство, если в качестве матрицы A взять матрицу $\| H_i \delta_{ik} \|$.

Замечание 2. Поскольку $|A| = |A'|$, то, заменяя матрицу A транспонированной матрицей A', получаем достаточные условия невырожденности матрицы A в виде условий Адамара для столбцов

$$G_i \equiv |a_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}| > 0 \qquad (i = 1,\ \ldots,\ n). \tag{10}$$

При выполнении этих условий вместо (4) будем иметь

$$\operatorname{mod} |A| \geqslant G_1 G_2 \ldots G_n. \tag{11}$$

Пусть C — произвольная невырожденная $n \times n$-матрица. Тогда матрицы A и AC одновременно являются невырожденными. Поэтому в условиях (3), (10), а также в оценках (4) и (11) можно матрицу A заменить на AC. Варьируя матрицу C, будем получать различные (неэквивалентные между собой) достаточные условия невырожденности, а также оценки для $|A|$, аналогичные (4) и (11). В частности, за счет подбора надлежащей матрицы C можно осуществить произвольную

перестановку столбцов. Тогда вместо условий (3) получим условия

$$H_i \equiv |a_{i\mu_i}| - \sum_{\substack{j=1 \\ j\neq\mu_i}}^{n} |a_{ij}| > 0 \qquad (i=1, \ldots, n), \tag{12}$$

где $(\mu_1, \ldots, \mu_n)$ — фиксированная, но произвольная перестановка индексов $1, 2, \ldots, n$.

Другими словами, матрица $A = \|a_{ij}\|_1^n$ будет невырожденной, если в каждой ее строке имеется доминирующий (не обязательно диагональный!) элемент и эти n доминирующих элементов расположены в различных столбцах.

Аналогичное предложение имеет место для столбцов.

Пусть теперь выполняются ослабленные условия Адамара

$$H_i = |a_{ii}| - \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \geqslant 0 \qquad (i=1, \ldots, n). \tag{13}$$

В этом случае каждый диагональный элемент является *слабо доминирующим* для своей строки.

Допустим, что матрица A вырождена и $Ax=0$, вектор-столбец $x=(x_1, \ldots, x_n) \neq 0$ имеет ровно p элементов x_k с максимальным модулем $|x_k|$, и пусть сначала $p<n$. Перенумеруем координаты вектора x так, чтобы этими максимальными по модулю были первые p координат:

$$|x_1| = \ldots = |x_p| > |x_j| \qquad (j=p+1, \ldots, n).$$

При этом равенство $Ax=0$ сохранится, если мы совершим некоторую (но одну и ту же) перестановку строк и столбцов матрицы A. После этого можно написать:

$$a_{kk}x_k = -\sum_{\substack{j=1 \\ j\neq k}}^{n} a_{kj}x_j \qquad (k=1, \ldots, p),$$

откуда

$$|a_{kk}|\,|x_k| \leqslant \Big(\sum_{\substack{j=1 \\ j\neq k}}^{p} |a_{kj}|\Big)|x_k| + \sum_{j=p+1}^{n} |a_{kj}|\,|x_j| \leqslant$$

$$\leqslant |x_k| \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}| \qquad (k=1, \ldots, n). \tag{14}$$

Сокращая на $|x_k|$, получим

$$|a_{kk}| \leqslant \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}| \qquad (k=1, \ldots, n). \tag{15}$$

Сопоставляя эти соотношения с ослабленными условиями Адамара (13), которые имеют место по условию, заключаем, что во всех соотношениях (15), а значит и в (14), имеет место знак равенства. А это возможно лишь тогда, когда

$$\sum_{j=p+1}^{n} |a_{kj}| = 0 \qquad (k=p+1, \ldots, n),$$

т. е. матрица A имеет вид

$$A=\begin{pmatrix} A_1 & \overbrace{0}^{p} \\ A_3 & A_4 \end{pmatrix}\Big\}^{p}. \tag{16}$$

Но матрица, которая одной и той же перестановкой строк и столбцов приводится к виду (16), называется разложимой (см. гл. 13, § 1). Таким образом, при $p<n$ A — разложимая матрица.

Если же $p=n$, то во всех соотношениях (15) и, следовательно, во всех n ослабленных условиях Адамара (13) имеет место знак равенства.

К этим выводам мы пришли, допустив, что A — вырожденная матрица.

Таким образом, нами доказана следующая теорема, представляющая собой уточнение теоремы Адамара.

Теорема 2 (Ольги Тауски). *Если для неразложимой матрицы A выполняются ослабленные условия Адамара* (13) *и по крайней мере в одном из этих условий имеет место знак* $>$, *то матрица A невырождена.* Само собой разумеется, что и в этой теореме условия $H_i\geqslant 0$ $(i=1, \ldots, n)$ могут быть заменены условиями $G_i\geqslant 0$ $(i=1, \ldots, n)$.

Для дальнейших обобщений теоремы Адамара нам понадобится понятие о норме прямоугольной матрицы. Этому понятию посвящается следующий параграф.

§ 2. Норма матрицы

В n-мерном пространстве R векторов-столбцов x введем понятие о норме вектора. Каждому вектору $x\in R$ ставим в соответствие некоторое вещественное неотрицательное число $\|x\|_R$ или просто $\|x\|$ так, чтобы для произвольных векторов x, y из R и произвольного скаляра λ выполнялись следующие условия:

1° $\|x+y\|\leqslant\|x\|+\|y\|$,

2° $\|\lambda x\|=|\lambda|\,\|x\|$,

3° $\|x\|>0$, если $x\neq 0$.

Полагая в 2° $\lambda=0$, получим, что $\|x\|=0$, если $x=0$. Кроме того, из 2° сразу следует для любых векторов x, $y\in R$: $\|x-y\|\geqslant\|x\|-\|y\|$.

Так, например, можно ввести «кубическую» норму вектора

$$\|x\|_{\mathrm{I}}=\max_{1\leqslant i\leqslant n}|x_i| \tag{17}$$

или «октаэдрическую» норму

$$\|x\|_{\mathrm{II}}=\sum_{i=1}^{n}|x_i|. \tag{17'}$$

«Эрмитову» (в случае вещественного пространства R «евклидову») норму $\|x\|_{\mathrm{III}}$ определяют равенством[1])

$$\|x\|_{\mathrm{III}}=\sqrt{\sum_{i=1}^{n}|x_i|^2}. \tag{17''}$$

[1]) См. гл. IX, § 2.

Легко проверяется, что все эти нормы удовлетворяют постулатам 1°, 2° и 3°.

Рассмотрим теперь произвольную прямоугольную $m \times n$-матрицу A и связанное с нею линейное преобразование $y = Ax$. x — n-мерный вектор-столбец из n-мерного пространства R, а y — m-мерный вектор-столбец из m-мерного пространства S.

Введем в этих пространствах нормы векторов $\|x\|_R = \|x\|$ и $\|y\|_S = \|y\|$. После этого норму прямоугольной матрицы A определим равенством

$$\|A\| = \sup_{\substack{x \in R \\ x \neq 0}} \frac{\|Ax\|_S}{\|x\|_R}. \tag{18}$$

Норма $m \times n$-матрицы A определяется как самой матрицей A, так и теми векторными нормами, которые введены в пространствах R и S. При изменении этих норм изменяется и норма матрицы.

Из определения нормы следует очевидное соотношение

$$\|Ax\|_S \leqslant \|A\| \, \|x\|_R. \tag{18'}$$

Для двух $m \times n$-матриц A и B при одном и том же определении векторных норм имеем соотношение

$$\|A + B\| \leqslant \|A\| + \|B\|. \tag{19}$$

Кроме того, очевидно, что

$$\|\lambda A\| = |\lambda| \, \|A\|. \tag{19'}$$

Пусть $p \times n$-матрица B отображает n-мерное пространство R в p-мерное S, а $m \times p$-матрица A отображает p-мерное пространство S в m-мерное T. Очевидно, матрица AB отображает R в T. Вводя в пространствах R, S и T векторные нормы и определяя с их помощью нормы матриц $\|A\|$, $\|B\|$, $\|AB\|$, легко приходим к неравенству

$$\|AB\| \leqslant \|A\| \, \|B\|. \tag{19''}$$

Так, например, если исходить из «кубических» векторных норм $\|x\|_I = \max\limits_{1 \leqslant k \leqslant n} |x_k|$, $\|y\|_I = \max\limits_{1 \leqslant i \leqslant m} |y_i|$, то норма матрицы $A = \|a_{ik}\|$ ($i = 1, \ldots, m$; $k = 1, \ldots, n$) определяется формулой

$$\|A\| = \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|. \tag{20}$$

Действительно, в этом случае

$$\|Ax\|_I = \max_{1 \leqslant i \leqslant m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| \leqslant \max_{1 \leqslant k \leqslant n} |x_k| \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|,$$

и потому

$$\frac{\|Ax\|_I}{\|x\|} \leqslant \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|.$$

В то же время здесь знак $=$ имеет место, если выбрать координаты $x_1, x_2, \ldots, x_n$ вектора x так, чтобы $|x_1| = |x_2| = \ldots = |x_n|$ и $a_{ph} x_k = |a_{ph} x_k|$ ($k = 1, \ldots, n$), где p — то значение i, при котором достигается

максимум в правой части соотношения (20). Таким образом, эта правая часть равна $\sup\limits_{x\neq 0}\frac{\|Ax\|}{\|x\|}$ и имеет место формула (20)[1]).

Если же исходить из октаэдрических векторных норм

$$\|x\|_{\text{II}}=\sum_{k=1}^{n}|x_k|, \qquad \|y\|_{\text{II}}=\sum_{i=1}^{m}|y_i|,$$

то, как нетрудно показать,

$$\|A\|=\max_{1\leqslant k\leqslant n}\sum_{i=1}^{m}|a_{ik}|. \tag{20'}$$

Рассмотрим теперь эрмитовы векторные нормы: $\|x\|^2=\sum\limits_{k=1}^{n}|x_k|^2$ и $\|y\|^2=\sum\limits_{i=1}^{n}|y_i|^2$. Тогда, вводя в рассмотрение положительную эрмитову матрицу $S=A^*A$, будем иметь:

$$\|Ax\|^2=y^*y=x^*A^*Ax=x^*Sx, \quad \|x\|^2=x^*x.$$

Но тогда (см. гл. X, § 7)

$$\|A\|^2=\max_{x\neq 0}\frac{x^*Sx}{x^*x}=\varrho,$$

где ϱ — максимальное характеристическое число матрицы AA^*. В этом случае

$$\|A\|=\sqrt{\varrho}. \tag{20''}$$

Введем теперь различные нормы для векторных столбцов x и y. Пусть, например,

$$\|x\|_{\text{II}}=\sum_{k=1}^{n}|x_k|, \qquad \|y\|_{\text{I}}=\max_{1\leqslant i\leqslant n}|y_i|.$$

Тогда

$$\|Ax\|_{\text{I}}=\max_{1\leqslant i\leqslant m}\Big|\sum_{k=1}^{n}a_{ik}x_k\Big|\leqslant a\sum_{k=1}^{n}|x_k|=a\|x\|_{\text{II}},$$

---

[1]) Иногда норму квадратной $n\times n$-матрицы A вводят аксиоматически (независимо от векторной нормы): каждой $n\times n$-матрице A ставится в соответствие неотрицательное действительное число $\|A\|$ так, что:

1° $\|A\|>0$, если $A\neq 0$, и $\|A\|\neq 0$, если $A=0$.

2° $\|A+B\|\leqslant\|A\|+\|B\|$.

3° $\|\lambda A\|=|\lambda|\,\|A\|$ (λ — скаляр).

4° $\|AB\|\leqslant\|A\|\,\|B\|$.

Например, можно положить

$$\|A\|=\sqrt{\sum_{i,\,k}|a_{ik}|^2} \quad\text{или}\quad \|A\|=n\max_{i,\,j}|a_{ij}|.$$

Вводят одну и ту же векторную норму как для вектора x, так и для вектора $y=Ax$ и называют эту норму согласованной с нормой $\|A\|$, если всегда выполняется соотношение $\|Ax\|\leqslant\|A\|\,\|x\|$. Наше определение нормы в этом частном случае ($m=n$ и $R\equiv S$) удовлетворяет требованиям 1°—4° и согласуется с векторной нормой. В отличие от произвольной нормы, определяемой аксиоматически, норму матрицы, определяемую с помощью формулы (18), называют операторной нормой, подчиненной данной норме векторов.

где $a = \max\limits_{1 \leqslant i,\ k \leqslant n} |a_{ik}|$. С другой стороны, если $a = a_{pq}$, то, выбирая x_q так, чтобы $a_{pq}x_q = a|x_q|$, и полагая $x_j = 0$ при $j \neq q$, будем иметь равенство $\|Ax\|_{\mathrm{I}} = a\|x\|_{\mathrm{II}}$.

Таким образом, в этом случае

$$\|Ax\| = \max_{1 \leqslant i,\ k \leqslant n} |a_{ik}|. \tag{20''}$$

§ 3. Распространение критерия Адамара на блочные матрицы

Пусть $n \times n$-матрица A разбита на s^2 блоков $A_{\alpha\beta}$ с размерами соответственно $n_\alpha \times n_\beta$ $(\alpha, \beta = 1, \ldots, s)$:

$$A = \begin{pmatrix} \overbrace{A_{11}}^{n_1} & \overbrace{A_{12}}^{n_2} & \ldots & \overbrace{A_{1s}}^{n_s} \\ A_{21} & A_{22} & \ldots & A_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{pmatrix} \begin{matrix} \} n_1 \\ \} n_2 \\ \\ \} n_s \end{matrix}. \tag{21}$$

При этом n-мерное пространство R автоматически расщепляется на s подпространств R_α с числом измерений n_α $(i = 1, \ldots, s)$. Для любого вектора $x \in R$ имеет место разложение

$$x = \sum_{\alpha=1}^{s} x_\alpha \qquad (x_\alpha \in R_\alpha,\ \alpha = 1, \ldots, s). \tag{21'}$$

Введем векторные нормы в пространствах R_α. Поскольку блок-матрица $A_{\alpha\beta}$ отображает R_β в R_α, то тем самым определится и норма

$$\|A_{\alpha\beta}\| = \sup_{\substack{x_\beta \in R_\beta \\ x_\beta \neq 0}} \frac{\|A_{\alpha\beta}x_\beta\|_{R_\alpha}}{\|x_\beta\|_{R_\beta}}. \tag{22}$$

В частности, определяется и норма квадратных матриц $A_{\alpha\alpha}$:

$$\|A_{\alpha\alpha}\| = \sup_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|A_{\alpha\alpha}x_\alpha\|}{\|x_\alpha\|}. \tag{22'}$$

Если $|A_{\alpha\alpha}| \neq 0$, то $\|A_{\alpha\alpha}\| > 0$. В этом случае из (22′) легко следует, что

$$\|A_{\alpha\alpha}^{-1}\| = \sup_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|x_\alpha\|}{\|A_{\alpha\alpha}x_\alpha\|}$$

и, следовательно,

$$\|A_{\alpha\alpha}^{-1}\|^{-1} = \inf_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|A_{\alpha\alpha}x_\alpha\|}{\|x_\alpha\|}. \tag{23}$$

Правая часть этого равенства имеет смысл и в случае, когда $A_{\alpha\alpha}$ — вырожденная матрица. (В этом случае справа стоит нуль.) Исходя из этого и из соображений непрерывности, будем считать, что $\|A_{\alpha\alpha}^{-1}\|^{-1}$ определено и в случае $|A_{\alpha\alpha}| = 0$ равно нулю.

Пусть теперь $|A| = 0$ и имеет место равенство $Ax = 0$ при $x \neq 0$. Исходя из представлений (21) и (21′), раскрывая блочное произведение

Ax, мы сможем написать

$$-A_{\alpha\alpha}x_\alpha = \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} A_{\alpha\beta}x_\beta \qquad (\alpha = 1, \ldots, s). \tag{24}$$

Отсюда в силу установленных ранее свойств нормы матрицы (см. (18′) и (19))

$$\| A_{\alpha\alpha}x_\alpha \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta}x_\beta \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| \, \| x_\beta \| \qquad (\alpha = 1, \ldots, s). \tag{25}$$

С другой стороны, из (23) следует

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \| x_\alpha \| \leqslant \| A_{\alpha\alpha}x_\alpha \| \qquad (\alpha = 1, \ldots, s),$$

что в сочетании с предыдущими неравенствами (25) дает

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \| x_\alpha \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| \, \| x_\beta \| \qquad (\alpha = 1, \ldots, s). \tag{26}$$

Как и в § 1, выберем индекс α так, чтобы $\| x_\alpha \|$ имел наибольшее значение (по сравнению с $\| x_\beta \|$, где $\beta \neq \alpha$), и заменим в правой части (26) все $\| x_\beta \|$ на $\| x_\alpha \|$. После сокращения на $\| x_\alpha \| > 0$ получим

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \|. \tag{27}$$

Поэтому при выполнении «блочных условий Адамара»

$$\| A_{\alpha\alpha}^{-1} \|^{-1} - \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| > 0 \qquad (\alpha = 1, \ldots, s) \tag{28}$$

соотношение (27) невозможно и матрица A не может быть вырожденной.

Мы пришли к теореме:

Теорема 3. *Если выполняются блочные условия Адамара* (28), *то A — невырожденная матрица.*

В частном случае $n_1 = n_2 = \ldots = n_s = 1$ эта теорема переходит в теорему Адамара, если в одномерных подпространствах R_α определить норму так: $\| x_\alpha \| = | x_\alpha |$ $(\alpha = 1, \ldots, s)$.

Само собой разумеется, что, записывая условие невырожденности транспонированной матрицы A', можно в теореме 3 условия Адамара для блочных строк заменить условиями Адамара для блочных столбцов:

$$\| A_{\alpha\alpha}^{-1} \|^{-1} - \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A'_{\beta\alpha} \| > 0 \qquad (\alpha = 1, \ldots, s). \tag{28′}$$

На блочные матрицы легко распространяется и теорема Ольги Тауски, если только в этой теореме потребовать «блочную» неприводимость матрицы A и выполнение ослабленных блочных условий Адамара со строгим знаком $>$ хотя бы в одном из них.

§ 4. Критерий регулярности Фидлера

Пусть снова $n \times n$-матрица A представлена в блочном виде (21). Составим для нее числовую $s \times s$-матрицу с вещественными элементами:

$$G = \begin{pmatrix} \|A_{11}^{-1}\|^{-1} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \|A_{22}^{-1}\|^{-1} & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \|A_{ss}^{-1}\|^{-1} \end{pmatrix}. \tag{29}$$

У этой матрицы все недиагональные элементы $\leqslant 0$, а диагональные $\geqslant 0$. Напомним читателю, что матрица с вещественными элементами называется M-матрицей, если у нее все недиагональные элементы $\leqslant 0$, т. е. неположительные и все главные миноры положительны[1]). Имеет место

Теорема 4 (Фидлера). *Если $s \times s$-матрица G является M-матрицей, то $n \times n$-матрица A является регулярной.*

Доказательство. Допустим, что $|A| = 0$. Тогда $Ax = 0$, где $x \neq 0$. Исходя из представлений (21) и (21′), как и ранее на стр. 413, получаем неравенства (26), которые теперь перепишем так:

$$\|A_{\alpha\alpha}^{-1}\|^{-1}\|x_\alpha\| - \sum_{\substack{\beta=1 \\ \beta\neq\alpha}}^{s} \|A_{\alpha\beta}\|\|x_\beta\| \leqslant 0 \qquad (\alpha = 1, \dots, s). \tag{30}$$

1. Пусть сначала все $\|x_\alpha\| > 0$. Тогда, увеличивая надлежащим образом в (30) коэффициент при $\|x_\alpha\|$, т. е. заменяя $\|A_{\alpha\alpha}^{-1}\|^{-1}$ на некоторое число $\tilde{g}_{\alpha\alpha} \geqslant \|A_{\alpha\alpha}^{-1}\|^{-1}$, мы из неравенств (30) получим систему равенств

$$\tilde{g}_{\alpha\alpha}\|x_\alpha\| - \sum_{\substack{\beta=1 \\ \beta\neq\alpha}}^{s} \|A_{\alpha\beta}\|\|x_\beta\| = 0 \qquad (\alpha = 1, \dots, s),$$

которые в матричной символике запишутся так:

$$\tilde{G}\xi = 0,$$

где

$$\tilde{G} = \begin{pmatrix} \tilde{g}_{11} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \tilde{g}_{22} & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \tilde{g}_{ss} \end{pmatrix},$$

а $\xi \neq 0$ — s-мерный вектор-столбец с элементами $\|x_1\|, \dots, \|x_s\|$. Отсюда сразу следует, что $|\tilde{G}| = 0$. С другой стороны, из определения M-матрицы следует, что $|\tilde{G}| \geqslant |G| > 0$. Мы пришли к противоречию, допустив, что $|A| = 0$.

Если некоторые $\|x_\alpha\| = 0$, то мы возьмем лишь те из соотношений (30), которые соответствуют значениям α, при которых $\|x_\alpha\| > 0$.

---

[1]) Согласно лемме Котелянского (см. стр. 369), для этого достаточно, чтобы n последовательных главных миноров были положительны.

Повторяя дословно предыдущие рассуждения и оперируя вместо $|G|$ некоторым главным минором матрицы G, мы снова придем к противоречию.

Теорема доказана полностью.

§ 5. Круги Гершгорина и другие области локализации

Пусть $A=\|a_{ik}\|_1^n$ — произвольная $n\times n$-матрица с комплексными элементами и λ — некоторое ее характеристическое число. Тогда $A-\lambda E$ — вырожденная матрица, и потому для нее не могут выполняться все условия Адамара, т. е. должно иметь место хотя бы одно из соотношений

$$|a_{ii}-\lambda|\leqslant\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}| \qquad (i=1,\ \ldots,\ n). \tag{31}$$

Каждое из соотношений (31) определяет некоторый круг в комплексной λ-плоскости с центром в точке a_{ii} радиуса $\sum\limits_{j=1,\ j\neq i}^{n}|a_{ij}|$. Мы пришли к теореме, установленной Гершгориным в 1931 г.

Теорема 5 (Гершгорина). *Каждое характеристическое число λ матрицы $A=\|a_{ik}\|_1^n$ всегда расположено в одном из кругов* (31).

Таким образом, объединение всех точек кругов Гершгорина (31) дает некоторую *область локализации* характеристических чисел матрицы A, т. е. область, в которой заведомо лежат все характеристические числа матрицы A. Каждый критерий регулярности приводит к своей области локализации характеристических чисел. Так, исходя из условий Адамара для столбцов, мы получили область локализации в виде объединения n кругов

$$|\lambda-a_{ii}|\leqslant\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}| \qquad (i=1,\ \ldots,\ n). \tag{31'}$$

Блочные условия Адамара приводят нас сразу к теореме:

Теорема 6. *Каждое характеристическое число λ $n\times n$-матрицы A, представленной в блочном виде $A=\{A_{\alpha\beta}\}_1^s$, принадлежит по крайней мере одной из областей*:

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}\leqslant\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s}\|A_{\alpha\beta}\| \qquad (\alpha=1,\ \ldots,\ s), \tag{32}$$

а также по крайней мере одной из областей

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}\leqslant\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s}\|A_{\beta\alpha}\| \qquad (\alpha=1,\ \ldots,\ s) \tag{32'}$$

(здесь E_α — единичная матрица того же порядка, что и $A_{\alpha\alpha}$; $\alpha=1,\ \ldots,\ s$).

Выясним, какую область локализации можно получить, исходя из критерия Фидлера. Пусть неотрицательные числа $c_1,\ c_2,\ \ldots,\ c_s$

выбраны так, чтобы матрица

$$G=\begin{pmatrix} c_1 & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & c_2 & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & c_s \end{pmatrix} \tag{33}$$

была *ослабленной M-матрицей*, т. е. чтобы все главные миноры этой матрицы были неотрицательны (недиагональные элементы в этой матрице заведомо все $\leqslant 0$). Допустим теперь, что при некотором числе λ выполняются s неравенств:

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1} > c_\alpha \qquad (\alpha=1, \dots, s). \tag{34}$$

Тогда, заменяя в матрице (33) c_α на $\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}$ $(\alpha=1, \dots, s)$, мы строго увеличим все диагональные элементы и получим уже M-матрицу (неослабленную!)

$$\begin{pmatrix} \|(A_{11}-\lambda E_1)^{-1}\|^{-1} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \|(A_{22}-\lambda E_2)^{-1}\|^{-1} & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \|(A_{ss}-\lambda E_s)^{-1}\|^{-1} \end{pmatrix}.$$

Но тогда по теореме Фидлера $|A-\lambda E|\neq 0$ и число λ не является характеристическим числом матрицы A.

Поэтому для любого характеристического числа λ матрицы A по крайней мере одно из неравенств (34) не выполняется, т. е. выполняется одно из соотношений

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1} \leqslant c_\alpha \qquad (\alpha=1, \dots, s). \tag{35}$$

Объединение s областей (35) и образует область локализации Фидлера, зависящую от специально выбираемых неотрицательных параметров $c_1, c_2, \dots, c_s$.

Теорема 7 (Фидлера). *Если неотрицательные числа $c_1, c_2, \dots, c_s$ выбраны так, чтобы матрица* (33) *была ослабленной M-матрицей, то каждое характеристическое число λ матрицы A принадлежит по крайней мере одной из s замкнутых областей* (35).

Рассмотрим в качестве примера симметрическую матрицу 4-го порядка

$$A=\begin{Vmatrix} 0 & 4 & 1 & -1 \\ 4 & 0 & -1 & 1 \\ 1 & -1 & -1 & 15 \\ -1 & 1 & 15 & -1 \end{Vmatrix}.$$

Поскольку матрица A симметрическая, то у нее все характеристические числа вещественны. Поэтому вместо областей локализации в комплексной λ-плоскости можно рассматривать отрезки, высекаемые этими областями на вещественной λ-оси.

I. Область Гершгорина состоит из одного сегмента

$$-18 \leqslant \lambda \leqslant 16,$$

который перекрывают остальные сегменты Гершгорина.

II. Разобьем матрицу A на 4 блока:

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad A_{11}=\begin{Vmatrix} 0 & 4 \\ 4 & 0 \end{Vmatrix}, \quad A_{22}=\begin{Vmatrix} -1 & 15 \\ 15 & -1 \end{Vmatrix}, \quad A_{12}=A_{21}=\begin{Vmatrix} 1 & -1 \\ -1 & 1 \end{Vmatrix}.$$

В данном случае

$$(A_{11}-\lambda E_1)=\frac{1}{\lambda^2-16}\begin{Vmatrix}-\lambda & -4\\ -4 & -\lambda\end{Vmatrix},$$

$$(A_{22}-\lambda E_2)^{-1}=\frac{1}{(\lambda+1)^2-15^2}\begin{Vmatrix}-1-\lambda & -1-\lambda\\ -15 & -15\end{Vmatrix}.$$

Рассмотрим три варианта нормировки подпространств R_1 и R_2:

а) в R_1 и R_2 — кубические нормы;

б) в R_1 кубическая, а в R_2 — октаэдральная норма;

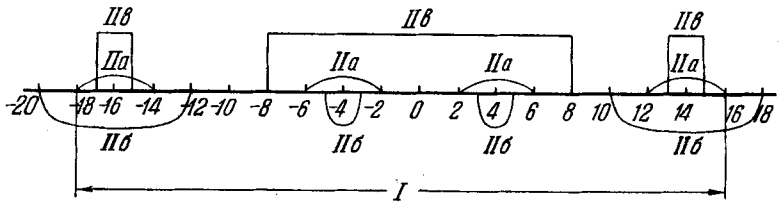


Рис. 6.

в) в R_1 октаэдральная, а в R_2 — кубическая норма.

а) Нормы всех блоков определяются по формуле (20′):

$\|A_{12}\|=2$, $\|A_{21}\|=2$, $\|(A_{11}-\lambda E_1)^{-1}\|=||\lambda|-4|$, $\|(A_{22}-\lambda E_2)^{-1}\|^{-1}=|\lambda+1|-15|$.

Блочные области Гершгорина:

$$||\lambda|-4|\leqslant 2, \qquad ||\lambda+1|-15|\leqslant 2$$

представляют собой совокупность 4 интервалов:

$$-18\leqslant\lambda\leqslant-16, \quad -6\leqslant\lambda\leqslant-2, \quad 2\leqslant\lambda\leqslant 6, \quad 12\leqslant\lambda\leqslant 16. \qquad \text{(IIа)}$$

б) В этом случае выражения для $\|(A_{11}-\lambda E_1)^{-1}\|^{-1}$ и $\|(A_{22}-\lambda E_2)^{-1}\|^{-1}$ остаются прежними, но

$$\|A_{12}\|=\max_x\frac{|x_1-x_2|}{|x_1|+|x_2|}=1, \quad \|A_{21}\|=\max_x\frac{2|x_1-x_2|}{\max\limits_{1\leqslant i\leqslant 2}|x_i|}=4.$$

Блочные области Гершгорина:

$$||\lambda|-4|\leqslant 1, \quad ||\lambda+1|-15|\leqslant 4$$

распадаются на 4 интервала

$$-20\leqslant\lambda\leqslant-12, \quad -5\leqslant\lambda\leqslant-3, \quad 3\leqslant\lambda\leqslant 5, \quad 10\leqslant\lambda\leqslant 18. \qquad \text{(IIб)}$$

в) Отличие от предыдущего случая заключается лишь в том, что здесь

$$\|A_{21}\|=4, \quad \|A_{12}\|=1.$$

Поэтому блочные области Гершгорина

$$||\lambda|-4|\leqslant 4, \quad ||\lambda+1|-15|\leqslant 1$$

распадаются на 3 интервала:

$$-17\leqslant\lambda\leqslant-15, \quad -8\leqslant\lambda\leqslant 8, \quad 13\leqslant\lambda\leqslant 15. \qquad \text{(IIв)}$$

На схеме (рис. 6) изображены области I, IIа, IIб, IIв. Их пересечение дает области локализации:

$$-17\leqslant\lambda\leqslant-15, \quad -5\leqslant\lambda\leqslant-3, \quad 3\leqslant\lambda\leqslant 5, \quad 13\leqslant\lambda\leqslant 15.$$

III. При применении критерия Фидлера будем снова исходить из нормировок а), б), в):

$$G=\begin{pmatrix} c_1 & -\|A_{12}\| \\ -\|A_{21}\| & c_2 \end{pmatrix}=\begin{Vmatrix} c_1 & -2 \\ -2 & c_2 \end{Vmatrix}, \quad |G|=c_1c_2-4\geqslant C$$

Желая получить наименьшие значения c_1 и c_2, полагаем

$$c_1c_2=4.$$

Область Фидлера

$$||\lambda|-4|\leqslant c_1, \quad ||\lambda+1|-15|\leqslant c_2$$

совпадает с областью IIа при $c_1=2$, $c_2=2$, с областью IIб — при $c_1=1$,

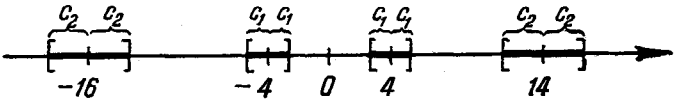


Рис. 7.

$c_2=4$ и с областью IIв при $c_1=4$, $c_2=1$. Область Фидлера состоит из 4 интервалов:

$$\left.\begin{array}{ll} -16-c_2\leqslant\lambda\leqslant -16+c_2, & -4-c_1\leqslant\lambda\leqslant -4+c_1, \\ 4-c_1\leqslant\lambda\leqslant 4+c_1, & 14-c_2\leqslant\lambda\leqslant 14+c_2 \end{array}\right\} \qquad \text{(III)}$$

и зависит от одного положительного параметра, поскольку $c_1=4/c_2$. Можно определить пересечение всех этих областей Фидлера.

Для этого (рис. 7) приравняем величины:

1) $-16-c_2=-4-c_1$, 2) $-16+c_2=-4-c_1$,

3) $-16+c_2=-4+c_1$, 4) $4-c_1=14-c_2$,

5) $4+c_1=14-c_2$, 6) $4+c_1=14+c_2$.

Используя равенство $c_1c_2=4$, получим 6 квадратных уравнений с наименьшими положительными корнями:

1) $c_2^2-12c_2-4=0$, $z_1=-6+\sqrt{40}=0{,}3246\ldots$,

2) $c_2^2-12c_2+4=0$, $z_2=6-\sqrt{32}=0{,}3431\ldots$,

3) $c_1^2+12c_1-4=0$, $z_3=z_1=-6+\sqrt{40}=0{,}3246\ldots$,

4) $c_1^2+10c_1-4=0$, $z_4=-5+\sqrt{29}=0{,}3852\ldots$,

5) $c_1^2-10c_1+4=0$, $z_5=5-\sqrt{21}=0{,}4174\ldots$,

6) $c_2^2+10c_2-4=0$, $z_6=z_4=-5+\sqrt{29}=0{,}3852\ldots$

Нетрудно уяснить себе, что область локализации, состоящая из пересечения всех областей Фидлера, состоит из следующих 4 сегментов:

$$-16-z_1\leqslant\lambda\leqslant -16+z_2, \qquad -4-z_2\leqslant\lambda\leqslant -4+z_1,$$

$$4-z_4\leqslant\lambda\leqslant 4+z_5, \qquad 14-z_5\leqslant\lambda\leqslant 14+z_4.$$

ГЛАВА XV

ПРИЛОЖЕНИЯ ТЕОРИИ МАТРИЦ К ИССЛЕДОВАНИЮ СИСТЕМ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

§ 1. Системы линейных дифференциальных уравнений с переменными коэффициентами. Общие понятия

Пусть дана система линейных однородных дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k \qquad (i = 1, 2, \ldots, n), \tag{1}$$

где $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ — комплексные функции вещественного аргумента t, непрерывные в некотором (конечном или бесконечном) интервале изменения t [1]).

Полагая $P(t) = \| p_{ik}(t) \|_1^n$ и $x = (x_1, x_2, \ldots, x_n)$, мы систему (1) запишем так:

$$\frac{dx}{dt} = P(t)\, x. \tag{2}$$

Интегральной матрицей системы (1) мы будем называть квадратную матрицу $X(t) = \| x_{ik}(t) \|_1^n$, столбцами которой являются n линейно независимых решений системы.

Так как каждый столбец матрицы X удовлетворяет уравнению (2), то и интегральная матрица X удовлетворяет уравнению

$$\frac{dX}{dt} = P(t)\, X. \tag{3}$$

В дальнейшем мы вместо системы (1) будем рассматривать матричное уравнение (3).

Из теоремы о существовании и единственности решения системы линейных дифференциальных уравнений [2]) следует, что интегральная матрица $X(t)$ однозначно определяется, если задано значение этой матрицы при некотором («начальном») значении $t = t_0$ [3]), $X(t_0) = X_0$. В качестве матрицы X_0 можно взять любую неособенную квадратную матрицу n-го порядка. В частном случае, когда $X(t_0) = E$, интегральную матрицу $X(t)$ будем называть *нормированной*.

---

[1]) Все соотношения этого параграфа, в которые входят функции от t, имеют место для данного интервала изменения t.

[2]) Доказательство этой теоремы приведено далее в § 5. См. также И. Г. Петровский, Лекции по теории обыкновенных дифференциальных уравнений, «Наука», 1964.

[3]) Предполагается, что t_0 принадлежит данному интервалу изменения t.

Продифференцируем определитель матрицы X, дифференцируя последовательно строки определителя и используя при этом дифференциальные соотношения

$$\frac{dx_{ij}}{dt} = \sum_{k=1}^{n} p_{ik}x_{kj} \qquad (i, j = 1, 2, \ldots, n).$$

Тогда получим:

$$\frac{d|X|}{dt} = (p_{11} + p_{22} + \ldots + p_{nn})|X|.$$

Отсюда следует известное тождество Якоби

$$|X| = ce^{\int_{t_0}^{t} \operatorname{Sp} P\, dt}, \tag{4}$$

где c — постоянная, а

$$\operatorname{Sp} P = p_{11} + p_{22} + \ldots + p_{nn}$$

— след матрицы $P(t)$.

Так как определитель $|X|$ не может тождественно равняться нулю, то $c \neq 0$. Но тогда из тождества Якоби следует, что определитель $|X|$ при любом значении аргумента отличен от нуля

$$|X| \neq 0,$$

т. е. *интегральная матрица при любом значении аргумента является неособенной.*

Если $\widetilde{X}(t)$ — неособенное ($|\widetilde{X}| \neq 0$) частное решение уравнения (3), то общее решение этого уравнения определяется формулой

$$X = \widetilde{X}C, \tag{5}$$

где C — произвольная постоянная матрица.

Действительно, умножая обе части равенства

$$\frac{d\widetilde{X}}{dt} = P\widetilde{X} \tag{6}$$

справа на C, убеждаемся, что и матрица $\widetilde{X}C$ удовлетворяет уравнению (3). С другой стороны, если X — произвольное решение уравнения (3), то из (6) следует:

$$\frac{dX}{dt} = \frac{d}{dt}(\widetilde{X}\cdot\widetilde{X}^{-1}X) = \frac{d\widetilde{X}}{dt}\widetilde{X}^{-1}X + \widetilde{X}\frac{d}{dt}(\widetilde{X}^{-1}X) = PX + \widetilde{X}\frac{d}{dt}(\widetilde{X}^{-1}X),$$

откуда в силу (3)

$$\frac{d}{dt}(\widetilde{X}^{-1}X) = 0$$

и

$$\widetilde{X}^{-1}X = \text{const} = C,$$

т. е. имеет место (5).

Все интегральные матрицы X системы (1) получаются по формуле (5) при $|C| \neq 0$.

Рассмотрим частный случай:

$$\frac{dX}{dt} = AX, \tag{7}$$

где A — постоянная матрица. При этом $\tilde{X}=e^{At}$ есть частное неособенное решение уравнения (7)[1]) и потому общее решение этого уравнения имеет вид

$$X=e^{At}C, \tag{8}$$

где C — произвольная постоянная матрица.

Полагая в (8) $t=t_0$, найдем: $X_0=e^{At_0}C$. Отсюда $C=e^{-At_0}X_0$ и потому формулу (8) можно представить в виде

$$X=e^{A(t-t_0)}X_0. \tag{9}$$

Эта формула эквивалентна выведенной ранее формуле (46) главы V (стр. 125).

Рассмотрим еще так называемую с и с т е м у К о ш и:

$$\frac{dX}{dt}=\frac{A}{t-a}X \qquad (A\text{ — постоянная матрица}). \tag{10}$$

Этот случай сводится к предыдущему заменой аргумента:

$$u=\ln(t-a).$$

Поэтому общее решение системы (10) выглядит так:

$$X=e^{A\ln(t-a)}C=(t-a)^A C. \tag{11}$$

Функции e^{At} и $(t-a)^A$, встречающиеся в формулах (8) и (11), могут быть представлены в виде (стр. 125)

$$e^{At}=\sum_{k=1}^{s}(Z_{k1}+Z_{k2}t+\ldots+Z_{km_k}t^{m_k-1})e^{\lambda_k t}, \tag{12}$$

$$(t-a)^A=\sum_{k=1}^{s}(Z_{k1}+Z_{k2}\ln(t-a)+\ldots+Z_{km_k}[\ln(t-a)]^{m_k-1})(t-a)^{\lambda_k}. \tag{13}$$

Здесь

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s}$$

$$(\lambda_i\neq\lambda_k \text{ при } i\neq k;\ i,\ k=1,\ 2,\ \ldots,\ s)$$

— минимальный многочлен матрицы A, а Z_{kj} $(j=1,\ 2,\ \ldots,\ m_k;\ k=1,\ 2,\ \ldots,\ s)$ — линейно независимые постоянные матрицы, являющиеся многочленами от A[2]).

З а м е ч а н и е. Иногда в качестве интегральной матрицы системы дифференциальных уравнений (1) берут матрицу W, у которой с т р о к и являются линейно независимыми решениями системы. Очевидно, матрица W будет транспонированной матрицей для X:

$$W=X'.$$

---

[1]) Почленно дифференцируя ряд $e^{At}=\sum_{k=0}^{\infty}\frac{A^k}{k!}t^k$, находим: $\frac{d}{dt}e^{At}=Ae^{At}$.

[2]) В правой части формулы (12) каждое слагаемое $X_k=(Z_{k1}+Z_{k2}t+\ldots+Z_{km_k}t^{m_k-1})e^{\lambda_k t}$ $(k=1,\ 2,\ \ldots,\ s)$ является решением уравнения (7). Действительно, этому уравнению удовлетворяет произведение $g(A)e^{At}$ при любой функции $g(\lambda)$. Но $X_k=f(A)=g(A)e^{At}$, если $f(\lambda)=g(\lambda)e^{\lambda t}$ и $g(\lambda_k)=1$, а все остальные $m-1$ значений функции $g(\lambda)$ на спектре матрицы A равны нулю [см. (16) на стр. 111].

Переходя в обеих частях равенства (3) к транспонированным матрицам, мы вместо (3) получим следующее уравнение для W:

$$\frac{dW}{dt} = W P(t). \tag{3'}$$

В правой части этого уравнения матрица W стоит первым множителем, а не вторым, как X в уравнении (3).

§ 2. Преобразование Ляпунова

Допустим теперь, что в системе (1) [и в уравнении (3)] матрица коэффициентов $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная ограниченная функция от t в интервале $[t_0, \infty)$ [1]).

Введем вместо неизвестных функций $x_1, x_2, \ldots, x_n$ новые неизвестные функции $y_1, y_2, \ldots, y_n$ при помощи преобразования

$$x_i = \sum_{k=1}^{n} l_{ik}(t)\, y_k \qquad (i = 1, 2, \ldots, n). \tag{14}$$

На матрицу преобразования $L(t) = \| l_{ik}(t) \|_1^n$ наложим следующие ограничения:

1° $L(t)$ имеет непрерывную производную $\frac{dL}{dt}$ в интервале $[t_0, \infty)$;

2° $L(t)$ и $\frac{dL}{dt}$ ограничены в интервале $[t_0, \infty)$;

3° существует постоянная m такая, что

$$0 < m < \operatorname{mod} |L(t)| \qquad (t \geqslant t_0),$$

т. е. определитель $|L(t)|$ ограничен по модулю снизу положительной постоянной m.

Преобразование (14), в котором матрица коэффициентов $L(t) = \| l_{ik}(t) \|_1^n$ удовлетворяет условиям 1°—3°, мы будем называть *преобразованием Ляпунова*, а соответствующую матрицу $L(t)$ — *матрицей Ляпунова*.

Такие преобразования рассматривал А. М. Ляпунов в своем знаменитом мемуаре «Общая задача об устойчивости движения» [19].

Примеры. 1. Если $L = \text{const}$ и $|L| \neq 0$, то матрица L удовлетворяет условиям 1°—3°. Следовательно, неособенное преобразование с постоянными коэффициентами всегда является преобразованием Ляпунова.

2. Если $D = \| d_{ik} \|_1^n$ — матрица простой структуры с чисто мнимыми характеристическими числами, то матрица

$$L(t) = e^{Dt}$$

удовлетворяет условиям 1°—3° и потому является матрицей Ляпунова [2]).

Легко проверить, что из свойств 1°—3° матрицы $L(t)$ следует, что существует обратная матрица $L^{-1}(t)$ и что она удовлетворяет тем же условиям 1°—3°, т. е. обратное преобразование для преобразования Ляпунова снова является преобразованием Ляпунова. Точно так же проверяется, что два последовательных преобразования Ляпунова в результате снова дают преобразование Ляпунова. Таким образом, преобразования Ляпунова образуют группу. Преобразования Ляпунова обладают следующим важным свойством:

---

[1]) Это означает, что каждая из функций $p_{ik}(t)$ ($i, k = 1, 2, \ldots, n$) непрерывна и ограничена в интервале $[t_0, \infty)$, т. е. при $t \geqslant t_0$.

[2]) При этом в формуле (12) все $m_k = 1$, а $\lambda_k = i\varphi_k$ (φ_k вещественны, $k = 1, 2, \ldots, s$).

Если при преобразовании (14) *система уравнений* (1) *переходит в систему*

$$\frac{dy_i}{dt} = \sum_{k=1}^{n} q_{ik}(t)\, y_k, \tag{15}$$

нулевое решение которой является устойчивым, асимптотически устойчивым или неустойчивым по Ляпунову (см. гл. V, § 7), *то таким же свойством обладает и нулевое решение исходной системы* (1).

Другими словами, преобразования Ляпунова не изменяют характеристики нулевого решения (в отношении устойчивости). Поэтому эти преобразования могут быть использованы при исследовании устойчивости для упрощения исходной системы уравнений.

Преобразование Ляпунова устанавливает одно-однозначное соответствие между решениями систем (1) и (15), при этом линейно независимые решения остаются таковыми и после преобразования. Поэтому преобразование Ляпунова переводит интегральную матрицу X системы (1) в некоторую интегральную матрицу Y системы (15), при этом

$$X = L(t)\, Y. \tag{16}$$

В матричной записи система (15) имеет вид

$$\frac{dY}{dt} = Q(t)\, Y, \tag{17}$$

где $Q(t) = \| q_{ik}(t) \|_1^n$ — матрица коэффициентов системы (15).

Подставляя в (3) вместо X произведение LY и сопоставляя полученное уравнение с (17), легко найдем следующую формулу, выражающую матрицу Q через матрицы P и L:

$$Q = L^{-1}PL - L^{-1}\frac{dL}{dt}. \tag{18}$$

Две системы (1) и (15) или, что то же, (3) и (17) мы будем называть *эквивалентными* (в смысле Ляпунова), если они переводятся друг в друга преобразованием Ляпунова. Матрицы коэффициентов P и Q эквивалентных систем всегда связаны между собой формулой (18), в которой матрица L удовлетворяет условиям 1°—3°.

§ 3. Приводимые системы

Среди систем линейных дифференциальных уравнений первого порядка наиболее простыми и наиболее изученными являются системы с постоянными коэффициентами. Поэтому представляют интерес системы, которые при помощи преобразования Ляпунова могут быть приведены к системам с постоянными коэффициентами. Такие системы А. М. Ляпунов называл *приводимыми*.

Пусть дана приводимая система

$$\frac{dX}{dt} = PX. \tag{19}$$

Тогда некоторое преобразование Ляпунова

$$X = L(t)\, Y \tag{20}$$

переводит ее в систему

$$\frac{dY}{dt} = AY, \tag{21}$$

где A — постоянная матрица. Поэтому система (19) имеет частное решение

$$\widetilde{X} = L(t)\, e^{At}. \tag{22}$$

Легко видеть, что, и обратно, всякая система (19), имеющая частное решение вида (22), где $L(t)$ — матрица Ляпунова, а A — постоянная матрица, является приводимой и при этом она приводится к виду (21) при помощи преобразования Ляпунова (20).

Следуя А. М. Ляпунову, покажем, что *всякая система* (19) *с периодическими коэффициентами приводима* [1]).

Пусть в данной системе (19) $P(t)$ — непрерывная функция в интервале $(-\infty, +\infty)$ с периодом τ:

$$P(t+\tau) = P(t). \tag{23}$$

Заменяя в (19) t на $t+\tau$ и используя (23), получим:

$$\frac{dX(t+\tau)}{dt} = P(t)\, X(t+\tau).$$

Таким образом, $X(t+\tau)$, как и $X(t)$, является интегральной матрицей системы (19). Поэтому $X(t+\tau) = X(t)\, V$, где V — некоторая постоянная неособенная матрица. Поскольку $|V| \neq 0$, то можно определить [2])

$$V^{\frac{t}{\tau}} = e^{\frac{t}{\tau}\ln V}.$$

Эта матричная функция от t, как и $X(t)$, умножается справа на V, если к аргументу прибавить τ. Поэтому «частное»

$$L(t) = X(t)\, V^{-\frac{t}{\tau}} = X(t)\, e^{-\frac{t}{\tau}\ln V}$$

является непрерывной периодической функцией с периодом τ:

$$L(t+\tau) = L(t)$$

и с определителем $|L(t)| \neq 0$. Матрица $L(t)$ удовлетворяет условиям 1°—3° предыдущего параграфа и, следовательно, является матрицей Ляпунова.

С другой стороны, поскольку решение X системы (19) представимо в виде

$$X = L(t)\, e^{\frac{\ln V}{\tau} t},$$

то система (19) приводима.

В данном случае преобразование Ляпунова

$$X = L(t)\, Y,$$

приводящее систему (19) к виду

$$\frac{dY}{dt} = \frac{1}{\tau}\ln V \cdot Y,$$

имеет периодические коэффициенты с периодом τ.

А. М. Ляпуновым был установлен [3]) весьма важный критерий устойчивости и неустойчивости по первому линейному приближению для

---

[1]) См. [22], § 47.

[2]) Здесь $\ln V = f(V)$, где $f(\lambda)$ — какая-либо однозначная ветвь функции $\ln \lambda$ в односвязной области G, содержащей все характеристические числа матрицы V и не содержащей числа 0. См. гл. V.

[3]) См. [22], § 24.

нелинейных систем дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik}x_k + (**) \qquad (i = 1, 2, \ldots, n), \tag{24}$$

где в правых частях стоят сходящиеся степенные ряды относительно $x_1, x_2, \ldots, x_n$, а $(**)$ обозначает сумму членов этих рядов второго порядка и выше относительно $x_1, x_2, \ldots, x_n$; коэффициенты a_{ik} $(i, k = 1, 2, \ldots, n)$ в линейных членах постоянны [1]).

Критерий Ляпунова. *Нулевое решение системы* (24) *будет устойчивым* (*и притом асимптотически*), *если матрица коэффициентов первого линейного приближения* $A = \| a_{ik} \|_1^n$ *имеет все характеристические числа с отрицательными вещественными частями, и неустойчивым, если хотя бы одно из этих характеристических чисел имеет положительную вещественную часть.*

Приведенные выше рассуждения позволяют использовать этот критерий для систем с периодическими коэффициентами в линейных членах:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k + (**) \qquad (i = 1, \ldots, n). \tag{25}$$

Действительно, на основании предыдущих рассуждений можно при помощи преобразования Ляпунова систему (25) привести к виду (24), где

$$A = \| a_{ik} \|_1^n = \frac{1}{\tau} \ln V,$$

а V — постоянная матрица, на которую умножается интегральная матрица соответствующей линейной системы (19) при сдвиге аргумента на τ. Не нарушая общности, можем считать $\tau > 0$. В силу свойств преобразования Ляпунова нулевое решение исходной системы и нулевое решение преобразованной одновременно являются устойчивыми, асимптотически устойчивыми или неустойчивыми. Но характеристические числа λ_i и ν_i $(i = 1, 2, \ldots, n)$ матриц A и V связаны между собой формулой

$$\lambda_i = \frac{1}{\tau} \ln \nu_i \qquad (i = 1, 2, \ldots, n).$$

Поэтому, применяя критерий Ляпунова к приведенной системе, найдем [2]):

Нулевое решение системы (25) *будет асимптотически устойчивым, если все характеристические числа* $\nu_1, \nu_2, \ldots, \nu_n$ *матрицы* V *по модулю* <1, *и неустойчивым, если хотя бы одно из этих чисел по модулю* >1.

А. М. Ляпунов установил свой критерий устойчивости по линейному приближению для значительно более широкого класса систем, а именно для систем вида (24), у которых система линейного приближения не обязательно система с постоянными коэффициентами, но принадлежит к классу систем, названных Ляпуновым правильными [3]).

Класс правильных линейных систем содержит в себе как часть все приводимые системы.

Критерий неустойчивости для случая, когда первое линейное приближение является правильной системой, был установлен Н. Г. Четаевым [4]).

---

[1]) Коэффициенты при нелинейных членах могут зависеть от t. На эти функциональные коэффициенты налагаются известные ограничения (см. [22], § 11).

[2]) См. там же, § 55.

[3]) Там же, § 9.

[4]) См. [38], стр. 181.

§ 4. Каноническая форма приводимой системы. Теорема Еругина

Пусть даны приводимая система (19) и эквивалентная ей (в смысле Ляпунова) система

$$\frac{dY}{dt} = AY,$$

где A — постоянная матрица.

Нас будет интересовать вопрос, *в какой степени матрица A определяется данной системой* (19). Этот вопрос можно еще сформулировать так:

В каком случае две системы

$$\frac{dY}{dt} = AY \text{ и } \frac{dZ}{dt} = BZ,$$

где A и B — постоянные матрицы, являются эквивалентными по Ляпунову, т. е. переводятся друг в друга преобразованием Ляпунова?

Для того чтобы ответить на этот вопрос, введем понятие о матрицах, имеющих одну и ту же вещественную часть спектра.

Мы будем говорить, что две матрицы A и B n-го порядка *имеют одну и ту же вещественную часть спектра* в том и только в том случае, когда элементарные делители матриц A и B имеют соответственно вид

$$(\lambda-\lambda_1)^{m_1},\ (\lambda-\lambda_2)^{m_2},\ \ldots,\ (\lambda-\lambda_s)^{m_s} \text{ и } (\lambda-\mu_1)^{m_1}, (\lambda-\mu_2)^{m_2},\ \ldots,\ (\lambda-\mu_s)^{m_s},$$

где

$$\operatorname{Re}\lambda_k = \operatorname{Re}\mu_k \quad (k=1,\ 2,\ \ldots,\ s).$$

Имеет место следующая теорема, установленная Н. П. Еругиным[1]):

Теорема 1 (Еругина). *Две системы*

$$\frac{dY}{dt} = AY \text{ и } \frac{dZ}{dt} = BZ \tag{26}$$

(*A и B — постоянные матрицы n-го порядка*) *эквивалентны в смысле Ляпунова в том и только в том случае, если матрицы A и B имеют одну и ту же вещественную часть спектра.*

Доказательство. Пусть даны системы (26). Приведем матрицу A к нормальной жордановой форме[2]) (см. гл. VI, § 7)

$$A = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_s E_s + H_s\}\, T^{-1}, \tag{27}$$

где

$$\lambda_k = \alpha_k + i\beta_k \quad (\alpha_k,\ \beta_k \text{ — вещественные числа; } k=1,\ 2,\ \ldots,\ s). \tag{28}$$

В соответствии с (27) и (28) положим:

$$\left.\begin{aligned} A_1 &= T\{\alpha_1 E_1 + H_1,\ \alpha_2 E_2 + H_2,\ \ldots,\ \alpha_s E_s + H_s\}\, T^{-1}, \\ A_2 &= T\{i\beta_1 E_1,\ i\beta_2 E_2,\ \ldots,\ i\beta_s E_s\}\, T^{-1}. \end{aligned}\right\} \tag{29}$$

---

[1]) См. [11], стр. 9—15. Приведенное здесь доказательство теоремы отличается от доказательства Н. П. Еругина.

[2]) E_k — единичная матрица, в H_k элементы первой наддиагонали равны единице, а остальные элементы равны нулю; порядок E_k, H_k равен степени k-го элементарного делителя матрицы A, т. е. m_k $(k=1,\ 2,\ \ldots,\ s)$.

Тогда

$$A = A_1 + A_2, \quad A_1 A_2 = A_2 A_1. \tag{30}$$

Определим матрицу $L(t)$ равенством $A(t) = e^{A_2 t}$. $L(t)$ — матрица Ляпунова (см. пример 2 на стр. 422).

Но частное решение первой из систем (26) в силу (30) имеет вид

$$e^{At} = e^{A_2 t} e^{A_1 t} = L(t)\, e^{A_1 (t)}.$$

Отсюда следует, что первая из систем (26) эквивалентна системе

$$\frac{dU}{dt} = A_1 U, \tag{31}$$

где согласно (29) матрица A_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы A.

Аналогично вторую из систем (26) заменим эквивалентной

$$\frac{dV}{dt} = B_1 V, \tag{32}$$

где матрица B_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы B.

Наша теорема будет доказана, если мы покажем, что *две системы* (31) *и* (32), *в которых матрицы* A_1 *и* B_1 *— постоянные матрицы с вещественными характеристическими числами, могут быть эквивалентны лишь в том случае, когда матрицы* A_1 *и* B_1 *подобны*[1]).

Пусть преобразование Ляпунова $U = L_1 V$ переводит (31) в (32). Тогда матрица L_1 удовлетворяет уравнению

$$\frac{dL_1}{dt} = A_1 L_1 - L_1 B_1. \tag{33}$$

Это матричное уравнение относительно L_1 эквивалентно системе n^2 дифференциальных уравнений относительно n^2 элементов матрицы L_1. Правая часть в (33) представляет собой линейную операцию над «вектором» L_1 в пространстве n^2 измерений

$$\frac{dL_1}{dt} = \widehat{F}(L_1) \quad [\widehat{F}(L_1) = A_1 L_1 - L_1 B_1]. \tag{33'}$$

Любое характеристическое число линейного оператора $\widehat{F}$ (и соответствующей ему матрицы порядка n^2) представляется в виде разности $\gamma - \delta$, где γ — характеристическое число матрицы A_1, а δ — характеристическое число матрицы B_1[2]). Отсюда следует, что оператор $\widehat{F}$ имеет только вещественные характеристические числа.

---

[1]) Из этого положения следует теорема 1, поскольку эквивалентность систем (31) и (32) означает эквивалентность систем (26), а подобие матриц A_1 и B_1 означает, что эти матрицы имеют одинаковые элементарные делители и потому матрицы A и B имеют одну и ту же вещественную часть спектра.

[2]) В самом деле, пусть Λ_0 — какое-либо характеристическое число оператора $\widehat{F}$. Тогда существует матрица $L \neq 0$ такая, что $\widehat{F}(L) = \Lambda_0 L$ или

$$(A_1 - \Lambda_0 E)\, L = L B_1. \tag{*}$$

Матрицы $A_1 - \Lambda_0 E$ и B_1 имеют хотя бы одно общее характеристическое число, так как в противном случае существовал бы такой многочлен $g(\lambda)$, что

$$g(A_1 - \Lambda_0 E) = 0, \quad g(B_1) = E,$$

а это невозможно, поскольку из (*) следует: $g(A_1 - \Lambda_0 E) \cdot L = L \cdot g(B_1)$ и $L \neq 0$. Но

Обозначим через $\widehat{\psi}(\lambda)=(\lambda-\widehat{\lambda}_1)^{\widehat{m}_1}(\lambda-\widehat{\lambda}_2)^{\widehat{m}_2}\ldots(\lambda-\widehat{\lambda}_u)^{\widehat{m}_u}$ ($\widehat{\lambda}_i$ вещественны; $\widehat{\lambda}_i\neq\widehat{\lambda}_j$ при $i\neq j$; $i,\ j=1,\ 2,\ \ldots,\ u$) минимальный многочлен для $\widehat{F}$. Тогда решение $L_1(t)=e^{\widehat{F}t}L^{(0)}$ системы (33′) в силу формулы (12) (стр. 421) запишется так:

$$L_1(t)=\sum_{k=1}^{u}\sum_{j=0}^{\widehat{m}_k-1}L_{kj}t^j e^{\widehat{\lambda}_k t}, \tag{34}$$

где L_{kj} — постоянные матрицы n-го порядка. Поскольку матрица $L_1(t)$ ограничена в интервале $(t_0,\ \infty)$, то как для любого $\widehat{\lambda}_k>0$, так и при $\widehat{\lambda}_k=0$ и $j>0$ соответствующие матрицы $L_{kj}=0$. Обозначим через $L_-(t)$ сумму всех слагаемых в (34), в которых $\widehat{\lambda}_k<0$. Тогда

$$L_1(t)=L_-(t)+L_0, \tag{35}$$

где

$$\lim_{t\to+\infty}L_-(t)=0,\qquad \lim_{t\to+\infty}\frac{dL_-(t)}{dt}=0,\ L_0=\text{const}. \tag{35′}$$

Тогда согласно (35) и (35′)

$$\lim_{t\to+\infty}L_1(t)=L_0,$$

откуда следует, что

$$|L_0|\neq 0,$$

поскольку определитель $|L_1(t)|$ ограничен по модулю снизу.

Подставляя в (33) вместо $L_1(t)$ сумму $L_-(t)+L_0$, получим:

$$\frac{dL_-(t)}{dt}-A_1L_-(t)+B_1L_-(t)=A_1L_0-B_1L_0,$$

откуда в силу (35′)

$$A_1L_0-L_0B_1=0$$

и, следовательно,

$$B_1=L_0^{-1}A_1L_0. \tag{36}$$

Обратно, если имеет место (36), то преобразование Ляпунова

$$U=L_0V$$

переводит систему (31) в систему (32). Теорема доказана.

Из доказанной теоремы вытекает, что *всякая приводимая система* (19) *при помощи преобразования Ляпунова* $X=LY$ *может быть приведена к виду*

$$\frac{dY}{dt}=JY,$$

если матрицы $A_1-\Lambda_0E$ и B_1 имеют общее характеристическое число, то $\Lambda_0=\gamma-\delta$ где γ и δ — характеристические числа соответственно матриц A_1 и B_1. Подробное исследование оператора $\widehat{F}$ можно найти в работах Ф. Голубчикова [85а, б].

где J — жорданова матрица с вещественными характеристическими числами. Эта каноническая форма системы заданием матрицы $P(t)$ определяется однозначно с точностью до порядка диагональных клеток в J.

§ 5. Матрицант

Рассмотрим систему дифференциальных уравнений

$$\frac{dX}{dt} = P(t)\,X, \tag{37}$$

где $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная матричная функция в некотором интервале (a, b) изменения аргумента t [1]).

Воспользуемся методом последовательных приближений для определения нормированного решения системы (37), т. е. решения, обращающегося в единичную матрицу при $t = t_0$ [t_0 — фиксированное число из интервала (a, b)]. Последовательные приближения X_k $(k = 0, 1, 2, \ldots)$ будем находить из рекуррентных соотношений

$$\frac{dX_k}{dt} = P(t)\,X_{k-1} \qquad (k = 1, 2, \ldots),$$

выбирая в качестве приближения X_0 единичную матрицу E.

Полагая $X_k(t_0) = E$ $(k = 0, 1, 2, \ldots)$, мы сможем X_k представить в виде

$$X_k = E + \int_{t_0}^{t} P(t)\,X_{k-1}\,dt \qquad (k = 1, 2, \ldots).$$

Таким образом,

$$X_0 = E,\ X_1 = E + \int_{t_0}^{t} P(t)\,dt,\ X_2 = E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt,\ \ldots,$$

т. е. X_k $(k = 0, 1, 2, \ldots)$ есть сумма первых $k+1$ членов матричного ряда

$$E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt + \ldots \tag{38}$$

Для того чтобы доказать, что этот ряд абсолютно и равномерно сходится в любой замкнутой части интервала (a, b) и определяет искомое решение уравнения (37), мы построим мажорантный ряд.

Определим неотрицательные функции $g(t)$ и $h(t)$ в интервале (a, b) равенствами [2])

$$g(t) = \max\left[\,|p_{11}(t)|,\ |p_{12}(t)|,\ \ldots,\ |p_{nn}(t)|\,\right],\quad h(t) = \left|\int_{t_0}^{t} g(t)\,dt\right|.$$

---

[1]) (a, b) — произвольный интервал (конечный или бесконечный). Все элементы $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ матрицы $P(t)$ — комплексные функции вещественного аргумента t, непрерывные в интервале (a, b). Все последующее сохраняет силу, если вместо непрерывности потребовать лишь ограниченность и интегрируемость по Риману [в любом конечном подинтервале интервала (a, b)] всех функций $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$.

[2]) По определению значение функции $g(t)$ при каком-либо из значений t равно наибольшему из n^2 модулей значений $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ при том же значении t.

Легко проверяется, что функции $g(t)$, а следовательно, и $h(t)$ непрерывны в интервале (a, b)[1]).

Каждый из n^2 скалярных рядов, на которые распадается матричный ряд (38), мажорируется рядом

$$1+h(t)+\frac{nh^2(t)}{2!}+\frac{n^2h^3(t)}{3!}+\ldots \tag{39}$$

Действительно,

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right| = \left|\int_{t_0}^{t} p_{ik}(t)\,dt\right| \leqslant \left|\int_{t_0}^{t} g(t)\,dt\right| = h(t),$$

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right| = \left|\sum_{j=1}^{n} \int_{t_0}^{t} p_{ij}\,dt \int_{t_0}^{t} p_{jk}(t)\,dt\right| \leqslant$$

$$\leqslant n\left|\int_{t_0}^{t} g(t)\,dt \int_{t_0}^{t} g(t)\,dt\right| = \frac{nh^2(t)}{2},$$

и т. д.

Ряд (39) сходится в интервале (a, b), причем сходится равномерно в любой замкнутой части этого интервала. Отсюда вытекает, что и матричный ряд (38) сходится в (a, b) и притом абсолютно и равномерно в любом замкнутом интервале, входящем в (a, b).

Почленным дифференцированием проверяем, что сумма ряда (38) представляет собой решение уравнения (37); это решение обращается в E при $t=t_0$. Почленное дифференцирование ряда (38) допустимо, поскольку ряд, получающийся после дифференцирования, отличается множителем P от ряда (38) и, следовательно, как и ряд (38), является равномерно сходящимся в любой замкнутой части интервала (a, b).

Таким образом, нами доказана теорема о существовании нормированного решения уравнения (37). Это решение будем обозначать через $\Omega_{t_0}^{t}(P)$ или просто $\Omega_{t_0}^{t}$. Любое другое решение, как было показано § 1, имеет вид

$$X=\Omega_{t_0}^{t}C,$$

где C — произвольная постоянная матрица. Из этой формулы следует, что любое решение, и в частности нормированное, однозначно определяется своим значением при $t=t_0$.

Нормированное решение $\Omega_{t_0}^{t}$ уравнения (37) часто называют *матрицантом*.

Мы показали, что матрицант представим в виде ряда[2])

$$\Omega_{t_0}^{t}=E+\int_{t_0}^{t} P(t)\,dt+\int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt+\ldots, \tag{40}$$

который сходится абсолютно и равномерно в любом замкнутом интервале, в котором функция $P(t)$ непрерывна.

---

[1]) Непрерывность функции $g(t)$ в любой точке t_1 интервала (a, b) следует из того, что разность $g(t)-g(t_1)$ при t, достаточно близком к t_1, всегда совпадает с одной из n^2 разностей $|p_{ik}(t)|-|p_{ik}(t_1)|$ $(i, k=1, 2, \ldots, n)$.

[2]) Представление матрицанта в виде такого ряда было впервые получено Пеано [226].

Отметим некоторые формулы для матрицанта.

1° $$\Omega_{t_0}^{t} = \Omega_{t_1}^{t}\Omega_{t_0}^{t_1} \quad (t_0,\ t_1,\ t \subset (a,\ b)).$$

Действительно, поскольку $\Omega_{t_0}^{t}$ и $\Omega_{t_1}^{t}$ — два решения уравнения (37), то

$$\Omega_{t_0}^{t} = \Omega_{t_1}^{t} C \quad (C \text{ — постоянная матрица}).$$

Полагая здесь $t = t_1$, получим $C = \Omega_{t_0}^{t_1}$.

2° $\Omega_{t_0}^{t}(P+Q) = \Omega_{t_0}^{t}(P)\,\Omega_{t_0}^{t}(S)$, где $S = [\Omega_{t_0}^{t}(P)]^{-1} Q \Omega_{t_0}^{t}(P)$.

Для вывода этой формулы положим:

$$X = \Omega_{t_0}^{t}(P), \quad Y = \Omega_{t_0}^{t}(P+Q)$$

и

$$Y = XZ. \tag{41}$$

Дифференцируя почленно (41), найдем:

$$(P+Q)\,XZ = PXZ + X\frac{dZ}{dt}.$$

Отсюда

$$\frac{dZ}{dt} = X^{-1}QXZ$$

и, следовательно, поскольку из (41) следует, что $Z(t_0) = E$,

$$Z = \Omega_{t_0}^{t}(X^{-1}QX).$$

Подставляя в (41) вместо X, Y, Z соответствующие матрицанты, получаем формулу 2°.

3° $$\ln|\Omega_{t_0}^{t}(P)| = \int_{t_0}^{t} \operatorname{Sp} P\,dt.$$

Эта формула следует из тождества Якоби (4) (стр. 420), если в него вместо $X(t)$ подставить $\Omega_{t_0}^{t}(P)$.

4° *Если* $A = \|a_{ik}\|_1^n = \text{const}$, *то*

$$\Omega_{t_0}^{t}(A) = e^{A(t-t_0)}.$$

Введем следующие обозначения. Если $P = \|p_{ik}\|_1^n$, то через $\operatorname{mod} P$ будем обозначать матрицу

$$\operatorname{mod} P = \|\,|p_{ik}|\,\|_1^n.$$

Кроме того, если $A = \|a_{ik}\|_1^n$ и $B = \|b_{ik}\|_1^n$ — две вещественные матрицы и

$$a_{ik} \leqslant b_{ik} \quad (i,\ k = 1,\ 2,\ \ldots,\ n),$$

то мы будем писать:

$$A \leqslant B.$$

Тогда из представления (40) следует:

5° *Если* $\operatorname{mod} P(t) \leqslant Q(t)$, *то*

$$\operatorname{mod} \Omega_{t_0}^{t}(P) \leqslant \Omega_{t_0}^{t}(Q) \quad (t > t_0).$$

В дальнейшем матрицу n-го порядка, у которой все элементы равны единице, будем обозначать через I:

$$I = \|1\|.$$

Рассмотрим функцию $g(t)$, определенную на стр. 429. Тогда

$$\operatorname{mod} P(t) \leqslant g(t) I.$$

Отсюда в силу 5°

$$\operatorname{mod} \Omega_{t_0}^t(P) \leqslant \Omega_{t_0}^t(g(t) I) \quad (t > t_0). \tag{42}$$

Но $\Omega_{t_0}^t(g(t) I)$ есть нормированное решение уравнения

$$\frac{dX}{dt} = g(t) IX.$$

Следовательно, в силу 4°[1])

$$\Omega_{t_0}^t(g(t) I) = e^{h(t) I} \leqslant \left(1 + h(t) + \frac{nh^2(t)}{2!} + \frac{n^2 h^3(t)}{3!} + \dots\right) I,$$

где

$$h(t) = \int_{t_0}^{t} g(t)\, dt.$$

Поэтому из (42) следует:

$$6° \qquad \operatorname{mod} \Omega_{t_0}^t(P) \leqslant \left(\frac{1}{n} e^{nh(t)} + \frac{n-1}{n}\right) I \leqslant e^{nh(t)} I \quad (t > t_0),$$

где

$$h(t) = \int_{t_0}^{t} g(t)\, dt, \quad g(t) = \max_{1 \leqslant i,\, k \leqslant n} \{|p_{ik}(t)|\}.$$

Покажем теперь, как при помощи матрицанта выражается общее решение системы линейных дифференциальных уравнений с правыми частями:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t) x_k + f_i(t) \quad (i = 1, 2, \dots, n); \tag{43}$$

$p_{ik}(t)$, $f_i(t)$ $(i, k = 1, 2, \dots, n)$ — непрерывные функции в интервале изменения аргумента t.

Вводя столбцевые матрицы («векторы») $x = (x_1, x_2, \dots, x_n)$ и $f = (f_1, f_2, \dots, f_n)$ и квадратную матрицу $P = \| p_{ik} \|_1^n$, запишем эту систему так:

$$\frac{dx}{dt} = P(t) x + f(t). \tag{43'}$$

Будем искать решение этого уравнения в виде

$$x = \Omega_{t_0}^t(P) z, \tag{44}$$

где z — неизвестный столбец, зависящий от t. Подставим это выражение для x в (43'), получим:

$$p\Omega_{t_0}^t(P) z + \Omega_{t_0}^t(P) \frac{dz}{dt} = P\Omega_{t_0}^t(P) z + f(t),$$

откуда

$$\frac{dz}{dt} = [\Omega_{t_0}^t(P)]^{-1} f(t).$$

---

[1]) Используя замену независимой переменной t переменной $h = \int_{t_0}^{t} g(t)\, dt$.

Интегрируя, находим:

$$z = \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + c,$$

где c — произвольный постоянный вектор. Подставим это выражение в (44), получим:

$$x = \Omega_{t_0}^{t}(P) \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + \Omega_{t_0}^{t}(P)\, c. \tag{45}$$

Давая t значение t_0, найдем: $x(t_0) = c$. Поэтому формула (45) принимает вид

$$x = \Omega_{t_0}^{t}(P)\, x(t_0) + \int_{t_0}^{t} K(t, \tau) f(\tau)\, d\tau, \tag{45'}$$

где

$$K(t, \tau) = \Omega_{t_0}^{t}(P) [\Omega_{t_0}^{\tau}(P)]^{-1}$$

— так называемая матрица Коши.

§ 6. Мультипликативный интеграл. Инфинитезимальное исчисление Вольтерра

Рассмотрим матрицант $\Omega_{t_0}^{t}(P)$. Разобьем основной интервал (t_0, t) на n частей, введя промежуточные точки $t_1, t_2, \ldots, t_{n-1}$, и положим $\Delta t_k = t_k - t_{k-1}$ $(k = 1, 2, \ldots, n;\ t_n = t)$. Тогда на основании свойства 1° матрицанта (см. предыдущий параграф)

$$\Omega_{t_0}^{t} = \Omega_{t_{n-1}}^{t} \ldots \Omega_{t_1}^{t_2} \Omega_{t_0}^{t_1}. \tag{46}$$

Выберем в интервале (t_{k-1}, t_k) промежуточную точку τ_k $(k = 1, 2, \ldots, n)$. Тогда, считая Δt_k малыми величинами первого порядка, при вычислении $\Omega_{t_{k-1}}^{t_k}$ с точностью до малых второго порядка можно принять $P(t) \approx \text{const} = P(\tau_k)$. Тогда

$$\Omega_{t_{k-1}}^{t_k} = e^{P(\tau_k)\Delta t_k} + (**) = E + P(\tau_k)\Delta t_k + (**); \tag{47}$$

здесь символом $(**)$ мы обозначаем сумму членов, начиная со второго порядка малости.

Из (46) и (47) находим:

$$\Omega_{t_0}^{t} = e^{P(\tau_n)\Delta t_n} \ldots e^{P(\tau_2)\Delta t_2} e^{P(\tau_1)\Delta t_1} + (*) \tag{48}$$

и

$$\Omega_{t_0}^{t} = [E + P(\tau_n)\Delta t_n] \ldots [E + P(\tau_2)\Delta t_2]\, [E + P(\tau_1)\Delta t_1] + (*). \tag{49}$$

Переходя к пределу при неограниченном увеличении числа интервалов разбиения и стремлении к нулю длин этих интервалов (при предельном переходе малые члены $(*)$ исчезают [1]), получаем точные предельные

[1] Эти рассуждения могут быть уточнены путем оценки членов, обозначенных нами через $(*)$.

формулы

$$\Omega_{t_0}^{t}(P)=\lim_{\Delta t_k\to 0}\left[e^{P(\tau_n)\Delta t_n}\ldots e^{P(\tau_2)\Delta t_2}e^{P(\tau_1)\Delta t_1}\right] \tag{48'}$$

и

$$\Omega_{t_0}^{t}(P)=\lim_{\Delta t_k\to 0}\left[E+P(\tau_n)\Delta t_n\right]\ldots\left[E+P(\tau_2)\Delta t_2\right]\left[E+P(\tau_1)\Delta t_1\right]. \tag{49'}$$

Выражение, стоящее под знаком предела в правой части последнего равенства, представляет собой *интегральное произведение*[1]). Предел его мы назовем *мультипликативным интегралом* и обозначим символом

$$\widehat{\int_{t_0}^{t}}\left[E+P(t)\,dt\right]=\lim_{\Delta t_k\to 0}\left[E+P(\tau_n)\Delta t_n\right]\ldots\left[E+P(\tau_1)\Delta t_1\right]. \tag{50}$$

Формула (49') дает представление матрицанта в виде мультипликативного интеграла

$$\Omega_{t_0}^{t}(P)=\widehat{\int_{t_0}^{t}}(E+P\,dt), \tag{51}$$

а равенства (48) и (49) могут быть использованы для приближенного вычисления матрицанта.

Мультипликативный интеграл ввел впервые Вольтерра в 1887 г. На базе этого понятия Вольтерра построил своеобразное инфинитезимальное исчисление для матричных функций (см. [49])[2]).

Вся специфика мультипликативного интеграла связана с неперестановочностью между собой различных значений подинтегральной матричной функции $P(t)$. В том же весьма частном случае, когда все эти значения перестановочны между собой

$$P(t')P(t'')=P(t'')P(t') \qquad (t',\ t''\in(t_0,\ t)),$$

мультипликативный интеграл, как это видно из (48') и (51), сводится к матрице

$$e^{\int_{t_0}^{t}P(t)\,dt}.$$

Введем теперь *мультипликативную производную*

$$D_tX=\frac{dX}{dt}X^{-1}. \tag{52}$$

Операции D_t и $\widehat{\int_{t_0}^{t}}$ взаимно обратны:

Если

$$D_tX=P,$$

---

[1]) Аналог интегральной суммы для обычного интеграла.

[2]) Мультипликативный интеграл (по-немецки «Produkt-Integral») был использован Шлезингером при исследовании систем линейных дифференциальных уравнений с аналитическими коэффициентами [61a, b], см. также [235].

Мультипликативный интеграл (50) существует не только для функции $P(t)$, непрерывной в интервале интегрирования, но и при значительно более общих предположениях (см. [154]).

то

$$X=\int_{t_0}^{t}(E+P\,dt)\cdot C \qquad (C=X(t_0))\ ^{1})$$

и наоборот. Последняя формула может быть еще записана так [2]):

$$\int_{t_0}^{t}(E+P\,dt)=X(t)\,X(t_0)^{-1}. \tag{53}$$

Предлагаем читателю проверить справедливость следующих дифференциальных и интегральных формул [3]):

Дифференциальные формулы

I. $D_t(XY)=D_t(X)+XD_t(Y)\,X^{-1}$,
$D_t(XC)=D_t(X)$,
$D_t(CY)=CD_t(Y)\,C^{-1}$ $\left(\begin{matrix}C-\text{постоянная}\\ \text{матрица}\end{matrix}\right)$.

II. $D_t(X')=X'\,(D_tX)'X'^{-1}$ [4]).

III. $D_t(X^{-1})=-X^{-1}D_t(X)\,X=-(D_t(X'))'$,
$D_t(X'^{-1})=-(D_t(X))'$.

Интегральные формулы

IV. $$\int_{t_0}^{t}(E+P\,dt)=\int_{t_1}^{t}(E+P\,dt)\int_{t_0}^{t_1}(E+P\,dt).$$

V. $$\int_{t_0}^{t}(E+P\,dt)=\left[\int_{t}^{t_0}(E+P\,dt)\right]^{-1}.$$

VI. $$\int_{t_0}^{t}(E+CPC^{-1}\,dt)=C\int_{t_0}^{t}(E+P\,dt)\,C^{-1}$$ $\left(\begin{matrix}C-\text{постоянная}\\ \text{матрица}\end{matrix}\right)$.

VII. $$\int_{t_0}^{t}[E+(Q+D_tX)\,dt]=X(t)\int_{t_0}^{t}(E+X^{-1}QX\,dt)\,X(t_0)^{-1}$$ [5]).

---

[1]) Здесь произвольная постоянная матрица C является аналогом аддитивной произвольной постоянной в обычном неопределенном интеграле.

[2]) Аналог формулы $\int_{t_0}^{t}P\,dt=X(t)-X(t_0)$ в случае, когда $\frac{dX}{dt}=P$.

[3]) Эти формулы могут быть выведены непосредственно из определения мультипликативных производной и интеграла (см. [67]). Однако интегральные формулы получаются быстрее и проще, если рассматривать мультипликативный интеграл как матрицант и воспользоваться свойствами матрицанта, изложенными в предыдущем параграфе (см. [61a]).

[4]) Значок $'$ обозначает переход к транспонированной матрице.

[5]) Формула VII может быть рассматриваема в известном смысле как аналог формулы интегрирования по частям для обычных (немультипликативных) интегралов. Формула VII следует из формулы 2° § 5.

Выведем еще важную формулу, дающую оценку модуля[1]) разности между двумя мультипликативными интегралами:

$$\text{VIII. } \operatorname{mod}\left[\int_{t_0}^{t}(E+P\,dt)-\int_{t_0}^{t}(E+Q\,dt)\right]\leqslant\frac{1}{n}\,e^{nq(t-t_0)}(e^{nd(t-t_0)}-1)I(t>t_0),$$

если

$$\operatorname{mod}Q\leqslant qI,\qquad \operatorname{mod}(P-Q)\leqslant d\cdot I,\qquad I=\|1\|$$

(q, d — неотрицательные числа, n — порядок матриц P и Q).

Обозначим через D разность $P-Q$. Тогда

$$P=Q+D,\qquad \operatorname{mod}D\leqslant d\cdot I.$$

Рассматривая мультипликативный интеграл как матрицант и пользуясь разложением (40) матрицанта в ряд, найдем:

$$\int_{t_0}^{t}[E+(Q+D)\,dt]-\int_{t_0}^{t}(E+Q\,dt)=$$
$$=\int_{t_0}^{t}D\,dt+\int_{t_0}^{t}D\,dt\int_{t_0}^{t}Q\,dt+\int_{t_0}^{t}Q\,dt\int_{t_0}^{t}D\,dt+\int_{t_0}^{t}D\,dt\int_{t_0}^{t}D\,dt+\ldots$$

Из этого разложения видно, что

$$\operatorname{mod}\left\{\int_{t_0}^{t}[E+(Q+D)\,dt]-\int_{t_0}^{t}(E+Q\,dt)\right\}\leqslant$$
$$\leqslant\int_{t_0}^{t}E+(\operatorname{mod}Q+\operatorname{mod}D)\,dt-\int_{t_0}^{t}(E+\operatorname{mod}Q)\,dt\leqslant$$
$$\leqslant\int_{t_0}^{t}[E+(q+d)\,I\,dt]-\int_{t_0}^{t}[E+qI]\,dt=e^{(q+d)I(t-t_0)}-e^{qI(t-t_0)}=$$
$$=e^{qI(t-t_0)}\,(e^{d\cdot I(t-t_0)}-E)\leqslant\frac{1}{n}\,e^{nq(t-t_0)}[e^{nd(t-t_0)}-1]\,I.$$

Пусть теперь матрицы P и Q зависят от некоторого параметра α:

$$P=P(t,\ \alpha),\quad Q=Q(t,\ \alpha),$$

и пусть

$$\lim_{\alpha\to\alpha_0}P(t,\ \alpha)=\lim_{\alpha\to\alpha_0}Q(t,\ \alpha)=P_0(t),$$

причем стремление к пределу равномерно относительно t в рассматриваемом интервале $(t_0,\ t)$. Допустим, что, кроме того, при $\alpha\to\alpha_0$ матрица $Q(t,\ \alpha)$ по модулю ограничена матрицей qI, где q — положительная постоянная. Тогда, полагая

$$d(\alpha)=\max_{\substack{1\leqslant i,\ k\leqslant n\\ t_0\leqslant\tau\leqslant t}}|p_{ik}(\tau,\ \alpha)-q_{ik}(\tau,\ \alpha)|,$$

будем иметь:

$$\lim_{\alpha\to\alpha_0}d(\alpha)=0.$$

---

[1]) Относительно определения модуля матрицы, а также соотношения $\leqslant$ между матрицами см. стр. 431.

Поэтому из формулы VIII следует:

$$\lim_{\alpha\to\alpha_0}\left[\int_{t_0}^{t}(E+P\,dt)-\int_{t_0}^{t}(E+Q\,dt)\right]=0.$$

В частности, если Q не зависит от $\alpha\,[Q(t,\alpha)=P_0(t)]$, мы получаем:

$$\lim_{\alpha\to\alpha_0}\int_{t_0}^{t}[E+P(t,\alpha)\,dt]=\int_{t_0}^{t}[E+P_0(t)\,dt],$$

где

$$P_0(t)=\lim_{\alpha\to\alpha_0}P(t,\alpha).$$

§ 7. Дифференциальные системы в комплексной области. Общие свойства

Рассмотрим систему дифференциальных уравнений

$$\frac{dx_i}{dz}=\sum_{k=1}^{n}p_{ik}(z)\,x_k. \tag{54}$$

Здесь данные функции $p_{ik}(z)$ и искомые функции $x_i(z)$ $(i,\,k=1,\,2,\,\ldots,\,n)$ предполагаются однозначными аналитическими функциями комплексного аргумента z, регулярными в некоторой области G комплексной z-плоскости.

Вводя квадратную матрицу $P(z)=\|\,p_{ik}(z)\,\|_1^n$ и столбцевую матрицу $x=(x_1,\,x_2,\,\ldots,\,x_n)$, мы, как и в случае вещественного аргумента (§ 1), можем записать систему (54) в виде

$$\frac{dx}{dz}=P(z)\,x. \tag{54'}$$

Обозначая через X интегральную матрицу, т. е. матрицу, столбцами которой являются n линейно независимых решений системы (54), мы вместо (54') можем написать:

$$\frac{dX}{dz}=P(z)\,X. \tag{55}$$

Формула Якоби имеет место и при комплексном аргументе z:

$$|X|=ce^{\int_{z_0}^{z}\mathrm{Sp}\,P\,dz} \tag{56}$$

При этом предполагается, что z_0 и все точки пути, вдоль которого берется $\int_{z_0}^{z}$, являются регулярными точками для однозначной аналитической функции $\mathrm{Sp}\,P(z)=p_{11}(z)+p_{22}(z)+\ldots+p_{nn}(z)$ [1]).

Специфичность рассматриваемого случая комплексного аргумента заключается в том, что при однозначной функции $P(z)$ интегральная матрица $X(z)$ может быть многозначной функцией от z.

---

[1]) Здесь и в дальнейшем в качестве путей интегрирования берутся кусочно гладкие кривые.

В качестве примера рассмотрим систему Коши

$$\frac{dX}{dz} = \frac{U}{z-a} X \quad (U \text{ — постоянная матрица}). \tag{57}$$

Одним из решений этой системы, как и в случае вещественного аргумента, является (см. стр. 421) интегральная матрица

$$X = e^{U \ln(z-a)} = (z-a)^U. \tag{58}$$

В качестве области G возьмем всю z-плоскость за исключением точки $z=a$. Все точки этой области являются регулярными точками матрицы коэффициентов

$$P(z) = \frac{U}{z-a}.$$

Если $U \neq 0$, то точка $z=a$ является особой точкой (полюсом первого порядка) для матричной функции $P(z) = \frac{U}{z-a}$.

Элемент интегральной матрицы (58) при однократном обходе в положительном направлении точки $z=a$ возвращается с новым значением, которое получается из старого умножением справа на постоянную матрицу

$$V = e^{2\pi i U}.$$

Для общей системы (55) теми же рассуждениями, что и в случае вещественного аргумента, убеждаемся в том, что два однозначных решения X и $\widetilde{X}$ в некоторой части области G всегда связаны формулой

$$X = \widetilde{X} C,$$

где C — некоторая постоянная матрица. Эта формула сохранится при любом аналитическом продолжении функций $X(z)$ и $\widetilde{X}(z)$ в области G.

Теорема о существовании и единственности (при заданных начальных значениях) решения системы (54) может быть доказана аналогично вещественному случаю.

Рассмотрим односвязную и притом звездообразную относительно точки z_0 [1]) область G_1, составляющую часть области G, и пусть матричная функция $P(z)$ регулярна [2]) в области G_1. Составим ряд

$$E + \int_{z_0}^{z} P\,dz + \int_{z_0}^{z} P\,dz \int_{z_0}^{z} P\,dz + \ldots \tag{59}$$

Из односвязности области G_1 следует, что каждый встречающийся в ряду (59) интеграл не зависит от пути интегрирования и представляет собой регулярную функцию в области G_1. Поскольку область G_1 звездообразна относительно z_0, то при оценке модулей этих интегралов мы можем считать, что все интегралы берутся вдоль прямолинейного отрезка, соединяющего точки z_0 и z.

Абсолютная и равномерная в любой замкнутой части области G_1, содержащей точку z_0, сходимость ряда (59) вытекает из сходимости

---

[1]) Область называется *звездообразной относительно точки* z_0, если любой отрезок, соединяющий произвольную точку z области с точкой z_0, целиком лежит в данной области.

[2]) То есть все элементы $p_{ik}(z)$ $(i, k = 1, 2, \ldots n)$ матрицы $P(z)$ являются регулярными функциями в области G_1.

мажорантного ряда

$$1+lM+\frac{n}{2!}l^2M^2+\frac{n^2}{3!}l^3M^3+\dots$$

Здесь M — верхняя граница для модуля матрицы $P(z)$, а l — верхняя граница расстояний точки z от точки z_0, причем обе границы относятся к рассматриваемой замкнутой части области G_1.

Путем почленного дифференцирования проверяется, что сумма ряда (59) представляет собой решение уравнения (55). Это решение нормировано, поскольку оно при $z=z_0$ обращается в единичную матрицу E. Однозначное нормированное решение системы (55), как и в вещественном случае, будем называть матрицантом и будем обозначать через $\Omega_{z_0}^{z}(P)$. Таким образом, мы получили представление матрицанта в области G_1 в виде ряда[1])

$$\Omega_{z_0}^{z}(P)=E+\int_{z_0}^{z}P\,dz+\int_{z_0}^{z}P\,dz\int_{z_0}^{z}P\,dz+\dots \tag{60}$$

Свойства 1°—4° *матрицанта, установленные в* § 5, *автоматически переносятся и на случай комплексного аргумента.*

Произвольное решение уравнения (55), регулярное в области G и обращающееся при $z=z_0$ в матрицу X_0, представится в виде

$$X=\Omega_{z_0}^{z}(P)\cdot C \qquad (C=X_0). \tag{61}$$

Формула (61) охватывает все однозначные решения, регулярные в окрестности точки z_0 [z_0 — регулярная точка для матрицы коэффициентов $P(z)$]. Эти решения, будучи аналитически продолжены в область G, дадут все решения уравнения (55), т. е. уравнение (55) не может иметь решений, для которых z_0 была бы особой точкой.

Для аналитического продолжения матрицанта в область G удобно пользоваться мультипликативным интегралом.

§ 8. Мультипликативный интеграл в комплексной области

Мультипликативный интеграл вдоль некоторой кривой в комплексной плоскости определяется следующим образом.

Пусть даны некоторый путь L и матричная функция $P(z)$, непрерывная на кривой L. Разобьем путь L на n частей $(z_0, z_1)(z_1, z_2)\dots(z_{n-1}, z)$; здесь z_0 — начало, $z_n=z$ — конец пути, а $z_1, z_2, \dots, z_{n-1}$ — промежуточные точки разбиения. На отрезке пути (z_{k-1}, z_k) выберем произвольную точку ζ_k и введем обозначения $\Delta z_k=z_k-z_{k-1}$; $k=1, 2, \dots, n$. Тогда по определению

$$\overset{\frown}{\int_L}[E+P(z)\,dz]=\lim_{\Delta z_k\to 0}[E+P(\zeta_n)\Delta z_n]\dots[E+P(\zeta_1)\Delta z_1].$$

Сопоставляя это определение с определением на стр. 434, мы видим, что новое определение совпадает с прежним в том частном случае, когда путь L является отрезком вещественной оси. Однако и в общем случае,

[1]) Приведенное доказательство существования нормированного решения и представления его в области G_1 рядом (60) сохраняет свою силу, если вместо звездообразности области сделать более общее допущение: для любой замкнутой части области G_1 существует такое положительное число l, что любую точку z этой замкнутой части можно соединить с z_0 путем, длина которого $\leqslant l$.

когда путь L произвольно расположен в комплексной плоскости, новое определение может быть сведено к старому при помощи замены переменной интегрирования.

Если

$$z = z(t)$$

есть параметрическое уравнение пути, причем $z(t)$ — непрерывная функция в интервале (t_0, t), имеющая в этом интервале кусочно непрерывную производную $\frac{dz}{dt}$, то, как легко видеть,

$$\int_L [E + P(z)\,dz] = \int_{t_0}^{t} \left\{E + P[z(t)]\frac{dz}{dt}\,dt\right\}.$$

Эта формула показывает, что мультипликативный интеграл вдоль произвольного пути существует, если подинтегральная матрица $P(z)$ непрерывна вдоль этого пути[1]).

Мультипликативная производная определяется прежней формулой

$$D_z X = \frac{dX}{dz} X^{-1}.$$

При этом предполагается, что $X(z)$ — аналитическая функция.

Все дифференциальные формулы (I — III) предыдущего параграфа переносятся без изменения на случай комплексного аргумента. Что же касается интегральных формул IV — VI, то несколько видоизменяется их внешняя запись:

IV′. $$\int_{(L'+L'')} (E + P\,dz) = \int_{L''} (E + P\,dz) \int_{L'} (E + P\,dz).$$

V′. $$\int_{-L} (E + P\,dz) = \left[\int_{L} (E + P\,dz)\right]^{-1}.$$

VI′. $$\int_L (E + CPC^{-1}\,dz) = C \int_L (E + P\,dz)\,C^{-1} \left(\begin{matrix} C \text{ — постоянная} \\ \text{матрица}\end{matrix}\right).$$

В формуле IV′ мы через $L' + L''$ обозначили составной путь, получающийся при прохождении сначала пути L', а затем пути L''. В формуле V′ $-L$ обозначает путь, отличающийся от пути L только направлением.

Формула VII принимает теперь вид

VII′. $$\int_L [E + (Q + D_z X)\,dz] = X(z) \int_L (E + X^{-1}QX\,dz)\,X(z_0)^{-1}.$$

Здесь $X(z_0)$ и $X(z)$ в правой части обозначают соответственно значения $X(z)$ в начале и в конце пути L.

---

[1]) См. сноску [2]) на стр. 434. Даже в случае, когда $P(z)$ — непрерывная функция вдоль L, функция $P[z(t)]\frac{dz}{dt}$ может быть кусочно непрерывной. В этом случае мы можем интервал (t_0, t) разбить на частичные интервалы, в каждом из которых производная $\frac{dz}{dt}$ непрерывна, и под интегралом от t_0 до t понимать сумму интегралов вдоль этих частичных интервалов.

Формула VIII заменится теперь формулой

$$\text{VIII}'. \quad \operatorname{mod}\left[\int_L (E+P\,dz)-\int_L (E+Q\,dz)\right]\leqslant \frac{1}{n}\,e^{nql}\left(e^{nd\cdot l}-1\right)I,$$

где $\operatorname{mod} Q\leqslant qI$, $\operatorname{mod}(P-Q)\leqslant d\cdot I$, $I=\|1\|$, а l — длина пути L.

Формула VIII′ получается сразу из формулы VIII, если в последней сделать преобразование переменной, выбрав в качестве новой переменной интегрирования длину дуги s вдоль пути L $\left(\text{при этом } \left|\frac{dz}{ds}\right|=1\right)$.

Как и в случае вещественного аргумента, существует тесная связь мультипликативного интеграла с матрицантом.

Пусть дана однозначная аналитическая матричная функция $P(z)$, регулярная в области G, и пусть G_0 — односвязная область, содержащая точку z_0 и составляющая часть области G. Тогда матрицант $\Omega_{z_0}^{z}(P)$ будет регулярной функцией от z в области G_0.

Соединим точки z_0 и z произвольным путем L, целиком лежащим в G_0, и выберем на L промежуточные точки $z_1, z_2, \ldots, z_{n-1}$. Тогда, пользуясь равенством

$$\Omega_{z_0}^{z}=\Omega_{z_{n-1}}^{z}\ldots\Omega_{z_1}^{z_2}\Omega_{z_0}^{z_1},$$

совершенно так же, как в § 6 (стр. 433), предельным переходом получим:

$$\Omega_{z_0}^{z}(P)=\int_L (E+P\,dz)=\int_{z_0}^{z} (E+P\,dz). \tag{62}$$

Из этой формулы видно, что мультипликативный интеграл не зависит от формы пути, а зависит только от начала и конца пути, если весь путь интегрирования лежит в односвязной области G_0, в которой подинтегральная функция $P(z)$ регулярна. В частности, для замкнутого контура L, лежащего в односвязной области G_0, имеем:

$$\oint (E+P\,dz)=E. \tag{63}$$

Эта формула представляет собой аналог известной теоремы Коши, согласно которой обычный (немультипликативный) интеграл вдоль замкнутого контура равен нулю, если этот контур лежит в односвязной области, в которой подинтегральная функция регулярна.

Представление матрицанта в виде мультипликативного интеграла (62) может быть использовано для аналитического продолжения матрицанта вдоль произвольного пути L в области G. В этом случае формула

$$X=\int_{z_0}^{z} (E+P\,dz)\,X_0 \tag{64}$$

задает все ветви многозначной интегральной матрицы X дифференциального уравнения $\frac{dX}{dz}=PX$, обращающейся в X_0 на одной из ветвей при $z=z_0$. Различные ветви получаются за счет различных путей, соединяющих точки z_0 и z.

Согласно формуле Якоби (56)

$$|X|=|X_0|\,e^{\int_{z_0}^{z}\operatorname{Sp}P\,dz}$$

и, в частности, при $X_0 = E$

$$\left| \int_{z_0}^{z} (E + P\,dz) \right| = e^{\int_{z_0}^{z} \operatorname{Sp} P\,dz}. \tag{65}$$

Из этой формулы следует, что мультипликативный интеграл всегда представляет собой неособенную матрицу, если только путь интегрирования целиком лежит в области, в которой функция $P(z)$ регулярна.

Если L — произвольный замкнутый путь в G и G — неодносвязная область, то равенство (63) может и не иметь места. Более того, в этом случае значение интеграла

$$\oint (E + P\,dz)$$

не определяется заданием подинтегральной функции и замкнутого пути интегрирования L, а зависит еще и от выбора начальной точки интегрирования z_0 на кривой L. Действительно, выберем на замкнутой кривой L две точки z_0 и z_1 и обозначим участки пути от z_0 до z_1 и от z_1 до z_0 (в направлении интегрирования) соответственно через L_1 и L_2. Тогда согласно формуле IV′ [1])

$$\oint_{z_0} = \int_{L_2} \cdot \int_{L_1}, \quad \oint_{z_1} = \int_{L_1} \cdot \int_{L_2}$$

и, следовательно,

$$\oint_{z_1} = \int_{L_1} \cdot \oint_{z_0} \cdot \int_{L_1}^{-1}. \tag{66}$$

Формула (66) показывает, что символ $\oint (E + P\,dz)$ определяет некоторую матрицу с точностью до преобразования подобия, т. е. определяет только элементарные делители некоторой матрицы.

Рассмотрим элемент $X(z)$ решения (64) в окрестности точки z_0. Пусть L — произвольный замкнутый путь в G, начинающийся и кончающийся в точке z_0. После аналитического продолжения вдоль L элемент $X(z)$ перейдет в некоторый элемент $\widetilde{X}(z)$. При этом новый элемент $\widetilde{X}(z)$ будет удовлетворять тому же дифференциальному уравнению (55), поскольку $P(z)$ — однозначная функция в G. Поэтому $\widetilde{X} = XV$, где V — некоторая неособенная постоянная матрица. Из формулы (64) следует, что

$$\widetilde{X}(z_0) = \oint_{z_0} (E + P\,dz)\, X_0.$$

Сопоставляя это равенство с предыдущим, найдем:

$$V = X_0^{-1} \oint_{z_0} (E + P\,dz)\, X_0. \tag{67}$$

В частности, для матрицанта $X = \Omega_z^z$ имеем $X_0 = E$, и тогда

$$V = \oint_{z_0} (E + P\,dz). \tag{68}$$

[1]) Здесь мы для сокращения обозначений опускаем подинтегральное выражение $E + P\,dz$, одно и то же во всех интегралах.

§ 9. Изолированная особая точка

Займемся исследованием поведения решения (интегральной матрицы) в окрестности изолированной особой точки a.

Пусть матричная функция $P(z)$ регулярна для значений z, удовлетворяющих неравенствам

$$0<|z-a|<R.$$

Совокупность этих значений образует двусвязную область G. Матричная функция $P(z)$ в области G разлагается в ряд Лорана

$$P(z)=\sum_{n=-\infty}^{+\infty} P_n(z-a)^n. \tag{69}$$

Элемент $X(z)$ интегральной матрицы после однократного обхода в положительном направлении вокруг a вдоль пути L перейдет в элемент

$$X^+(z)=X(z)V,$$

где V — некоторая постоянная неособенная матрица.

Пусть U — постоянная матрица, связанная с матрицей V равенством

$$V=e^{2\pi iU}. \tag{70}$$

Тогда матричная функция $(z-a)^U$ после обхода вдоль L также переходит в $(z-a)^UV$. Поэтому аналитическая в области G матричная функция

$$F(z)=X(z)\,(z-a)^{-U} \tag{71}$$

при аналитическом продолжении вдоль L переходит сама в себя (остается неизменной)[1]. Поэтому матричная функция $F(z)$ регулярна в G и разлагается в G в ряд Лорана

$$F(z)=\sum_{n=-\infty}^{+\infty} F_n(z-a)^n. \tag{72}$$

Из (71) следует:

$$X(z)=F(z)\,(z-a)^U. \tag{73}$$

Таким образом, каждая интегральная матрица $X(z)$ может быть представлена в виде (73), где однозначная функция $F(z)$ и постоянная матрица U зависят от матрицы коэффициентов $P(z)$. Однако алгорифмическое определение матрицы U и коэффициентов F_n ряда (72) по коэффициентам P_n ряда (69) в общем случае представляет собой сложную задачу.

Частный случай этой задачи, когда

$$P(z)=\sum_{n=-1}^{\infty} P_n(z-a)^n,$$

---

[1]) Отсюда уже следует, что функция $F(z)$ при обходе вдоль любого другого замкнутого пути в G возвращается к исходному значению.

будет нами разобран полностью в § 10. В этом случае точка a называется *регулярной особой точкой* системы (55).

Если разложение (69) имеет вид

$$P(z)=\sum_{n=-q}^{\infty} P_n(z-a)^n \qquad (q>1;\ P_{-q}\neq 0),$$

то точка a называется *иррегулярной особой точкой типа полюса*. Наконец, если в ряду (69) имеется бесчисленное множество отличных от нуля матричных коэффициентов P_n при отрицательных степенях $z-a$, то точка a называется *существенной особой точкой* данной дифференциальной системы.

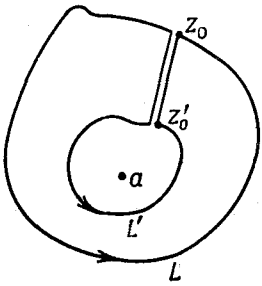


Рис. 8.

Из формулы (73) следует, что интегральная матрица $X(z)$ при любом однократном обходе в положительном направлении (вдоль некоторого замкнутого пути L) умножается справа на одну и ту же матрицу

$$V=e^{2\pi i U}.$$

Если этот обход начинается (и кончается) в точке z_0, то согласно (67)

$$V=X(z_0)^{-1}\oint_{z_0}(E+P\,dz)\,X(z_0). \tag{74}$$

Если вместо интегральной матрицы $X(z)$ мы рассмотрим любую другую интегральную матрицу $\widehat{X}(z)=X(z)C$ (C — постоянная матрица, $|C|\neq 0$), то, как видно из (74), матрица V заменится подобной матрицей

$$\widehat{V}=C^{-1}VC.$$

Таким образом, «интегральные подстановки» V данной системы образуют класс подобных между собой матриц.

Из формулы (74) также следует, что интеграл

$$\oint_{z_0}(E+P\,dz) \tag{75}$$

определяется начальной точкой обхода z_0 и не зависит от формы кривой обхода[1]). Если же мы меняем точку z_0, то получающиеся при этом различные значения интеграла (75) подобны между собой[2]).

В этих свойствах интеграла (75) можно убедиться и непосредственно. Действительно, пусть L и L' — два замкнутых пути в G вокруг точки $z=a$ с начальными точками обхода z_0 и z_0' (см. рис. 8).

Двусвязная область, заключенная между L и L', может быть сделана односвязной, если провести разрез от z_0 до z_0'. Интеграл вдоль разреза мы обозначим через

$$T=\int_{z_0}^{z_0'}(E+P\,dz).$$

---

[1]) Конечно, при условии, что путь интегрирования обходит точку a однократно в положительном направлении.

[2]) Это вытекает из формулы (74), а также из формулы (66).

Поскольку мультипликативный интеграл вдоль замкнутого контура односвязной области равен E, то

$$\overset{\frown}{\int_{L'}} T \overset{\frown}{\int_{L}^{-1}} T^{-1} = E,$$

откуда

$$\overset{\frown}{\int_{L'}} = T \overset{\frown}{\int_{L}} T^{-1}.$$

Таким образом, как и V, интеграл $\oint (E + P\,dz)$ определен с точностью до подобия, и равенство (74) мы иногда будем записывать так:

$$V = \widehat{\oint} (E + P\,dz),$$

понимая под этим совпадение элементарных делителей у матриц, стоящих в левой и правой частях равенства.

Рассмотрим для примера систему с регулярной особой точкой

$$\frac{dX}{dz} = P(z)\,X,$$

где

$$P(z) = \frac{P_{-1}}{z-a} + \sum_{n=0}^{+\infty} P_n (z-a)^n.$$

Пусть

$$Q(z) = \frac{P_{-1}}{z-a}.$$

Пользуясь формулой VIII′ предыдущего параграфа, дадим оценку модуля разности

$$D = \widehat{\oint} (E + P\,dz) - \widehat{\oint} (E + Q\,dz), \tag{76}$$

выбрав в качестве пути интегрирования окружность радиуса $r\,(r < R)$ с положительным направлением обхода. Тогда при

$$\operatorname{mod} P_{-1} \leqslant p_{-1} I, \quad \operatorname*{mod}_{|z-a|=r} \sum_{n=0}^{\infty} P_n (z-a)^n \leqslant d(r)\,I, \quad I = \|1\|$$

мы можем в формуле VIII′ положить:

$$q = \frac{p_{-1}}{r}, \quad d = d(r), \quad l = 2\pi r,$$

после чего получим:

$$\operatorname{mod} D \leqslant \frac{1}{n} e^{2\pi p_{-1}} \left(e^{2\pi n r d(r)} - 1\right) I.$$

Отсюда видно, что [1])

$$\lim_{r \to 0} D = 0. \tag{77}$$

---

[1]) При этом мы используем то, что при надлежащем выборе $d(r)$

$$\lim_{r \to 0} d(r) = d_0,$$

где d_0 — наибольший из модулей элементов матрицы P_0.

С другой стороны, система

$$\frac{dY}{dz} = QY$$

является системой Коши, и в этом случае при любом выборе начальной точки обхода z_0 и при любом $r < R$

$$\oint_{z_0} (E + Q\,dz) = e^{2\pi i P_{-1}}.$$

Поэтому из (76) и (77) следует:

$$\lim_{r \to 0} \oint_{z_0} (E + P\,dz) = e^{2\pi i P_{-1}}. \tag{78}$$

Но элементарные делители интеграла $\oint_{z_0} (E + P\,dz)$ не зависят от z_0 и r и совпадают с элементарными делителями интегральной подстановки V.

Отсюда Вольтерра в своем известном мемуаре (см. [149]), а также в книге [49] (стр. 117—120) делает вывод, что матрицы V и $e^{2\pi i P_{-1}}$ подобны, и потому интегральная подстановка V с точностью до подобия определяется матрицей «вычетов» P_{-1}.

Это утверждение Вольтерра ошибочно.

Из (74) и (78) можно лишь сделать вывод, что *характеристические числа интегральной подстановки V совпадают с характеристическими числами матрицы* $e^{2\pi i P_{-1}}$. Однако элементарные делители у этих матриц могут быть различными. Так, например, матрица

$$\begin{Vmatrix} \alpha & r \\ 0 & \alpha \end{Vmatrix}$$

при любом $r \neq 0$ имеет один элементарный делитель $(\lambda - \alpha)^2$, а предел этой матрицы при $r \to 0$, т. е. матрица $\begin{Vmatrix} \alpha & 0 \\ 0 & \alpha \end{Vmatrix}$, имеет два элементарных делителя $\lambda - \alpha$, $\lambda - \alpha$.

Таким образом, утверждение Вольтерра не вытекает из формул (74) и (78). Но оно и вообще неверно, как показывает следующий пример.

Пусть

$$P(z) = \begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix} \frac{1}{z} + \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}.$$

Соответствующая система дифференциальных уравнений имеет вид:

$$\frac{dx_1}{dz} = x_2, \quad \frac{dx_2}{dz} = -\frac{x_2}{z}.$$

Интегрируя эту систему, находим:

$$x_1 = c \ln z + d, \quad x_2 = \frac{c}{z}.$$

Интегральная матрица

$$X(z) = \begin{Vmatrix} \ln z & 1 \\ z^{-1} & 0 \end{Vmatrix}$$

при однократном положительном обходе вокруг особой точки $z=0$ умножается справа на матрицу

$$V=\begin{Vmatrix} 1 & 0 \\ 2\pi i & 1 \end{Vmatrix}.$$

Эта матрица имеет один элементарный делитель $(\lambda-1)^2$. В то же время матрица

$$e^{2\pi i P_{-1}}=e^{2\pi i \begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}}=\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}=E$$

имеет два элементарных делителя $\lambda-1$, $\lambda-1$.

Рассмотрим теперь случай, когда матрица $P(z)$ имеет конечное число отрицательных степеней $z-a$ (a — регулярная или иррегулярная особая точка типа полюса):

$$P(z)=\frac{P_{-q}}{(z-a)^q}+\ldots+\frac{P_{-1}}{z-a}+\sum_{n=0}^{\infty}P_n(z-a)^n \quad (q\geqslant 1;\ P_{-q}\neq 0).$$

Преобразуем данную систему

$$\frac{dX}{dz}=PX, \tag{79}$$

положив

$$X=A(z)Y, \tag{80}$$

где $A(z)$ — матричная функция, регулярная в точке $z=0$ и принимающая в этой точке значение E:

$$A(z)=E+A_1(z-a)+A_2(z-a)^2+\ldots;$$

степенной ряд в правой части сходится при $|z-a|<r_1$.

Известный американский математик Г. Биркгофф в 1913 г. опубликовал теорему (см. [119]), согласно которой всегда можно подобрать преобразование (80) так, чтобы матрица коэффициентов преобразованной системы

$$\frac{dY}{dz}=P^*(z)Y \tag{79'}$$

содержала только отрицательные степени $z-a$:

$$P^*(z)=\frac{P^*_{-q}}{(z-a)^q}+\ldots+\frac{P^*_{-1}}{z-a}.$$

Теорема Биркгоффа вместе с полным ее доказательством приведена в книге Э. Л. Айнса «Обыкновенные дифференциальные уравнения»[1]). Там же на основе рассмотрения «канонических» систем (79′) проводится исследование поведения решения произвольной системы в окрестности особой точки.

---

[1]) См. [1], стр. 632—641. Биркгофф и Айнс формулируют теорему для особой точки $z=\infty$. Это не является каким-либо ограничением, поскольку любая особая точка $z=a$ преобразованием $z'=\frac{1}{z-a}$ может быть переведена в $z'=\infty$.

Между тем доказательство Биркгоффа содержит ошибку, а сама теорема неверна. В качестве опровергающего примера можно взять хотя бы пример, приведенный выше для опровержения утверждения Вольтерра[1]).

В этом примере $q=1$, $a=0$ и

$$P_{-1}=\begin{Vmatrix}0 & 0\\ 0 & -1\end{Vmatrix},\quad P_0=\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix},\quad P_n=0 \text{ при } n=1,\ 2,\ \ldots$$

Применяя теорему Биркгоффа и подставляя в (79) вместо X произведение AY, мы после замены $\frac{dY}{dz}$ на $\frac{P^*_{-1}}{z}Y$ и сокращения на Y получим:

$$A\frac{P^*_{-1}}{z}+\frac{dA}{dz}=PA.$$

Приравнивая коэффициенты при $\frac{1}{z}$ и свободные члены, найдем:

$$P^*_{-1}=P_{-1},\quad A_1P_{-1}-P_{-1}A_1+A_1=P_0.$$

Полагая $A_1=\begin{Vmatrix}a & b\\ c & d\end{Vmatrix}$, получим:

$$\begin{Vmatrix}a & 0\\ c & 0\end{Vmatrix}-\begin{Vmatrix}0 & 0\\ -c & -d\end{Vmatrix}=\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix}.$$

Это — противоречивое равенство.

В следующем параграфе мы выясним, к какому каноническому виду может быть преобразована система (79) при помощи преобразования (80) в случае регулярной особой точки.

§ 10. Регулярная особая точка

Исследуя поведение решения в окрестности особой точки, мы без нарушения общности рассуждения можем принять, что особой точкой является точка $z=0$[2]).

1. Пусть дана система

$$\frac{dX}{dz}=P(z)\,X, \tag{81}$$

где

$$P(z)=\frac{P_{-1}}{z}+\sum_{m=0}^{\infty}P_mz^m \tag{82}$$

и ряд $\sum_{m=0}^{\infty}P_mz^m$ сходится внутри круга $|z|<r$.

---

[1]) В случае $q=1$ ошибочное утверждение Биркгоффа по существу совпадает с ошибкой Вольтерра (см. стр. 446).

[2]) Преобразованием $z'=z-a$ или $z'=\frac{1}{z}$ можно соответственно любую конечную точку $z=a$ или $z=\infty$ перевести в точку $z'=0$.

Положим

$$X = A(z)\,Y, \tag{83}$$

где

$$A(z) = E + A_1 z + A_2 z^2 + \dots \tag{84}$$

Оставляя пока в стороне вопрос о сходимости ряда (84), постараемся так определить матричные коэффициенты A_m этого ряда, чтобы преобразованная система

$$\frac{dY}{dz} = P^*(z)\,Y, \tag{85}$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m, \tag{86}$$

имела возможно более простой («канонический») вид [1]).

Подставляя в (81) вместо X произведение AY и используя (85), мы получим:

$$A(z)\,P^*(z)\,Y + \frac{dA}{dz}\,Y = P(z)\,A(z)\,Y.$$

Умножая обе части этого равенства справа на Y^{-1}, найдем:

$$P(z)\,A(z) - A(z)\,P^*(z) = \frac{dA}{dz}.$$

Заменяя здесь $P(z)$, $A(z)$, $P^*(z)$ рядами (82), (84), (86) и приравнивая в левой и правой частях равенства коэффициенты при одинаковых степенях z, получим бесконечную систему матричных уравнений для искомых коэффициентов A_1, A_2, ... [2]):

$$\left.\begin{aligned}
&1)\ P_{-1} = P^*_{-1},\\
&2)\ P_{-1}A_1 - A_1(P_{-1}+E) + P_0 = P^*_0,\\
&3)\ P_{-1}A_2 - A_2(P_{-1}+2E) + P_0A_1 - A_1P^*_0 + P_1 = P^*_1,\\
&\qquad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&(m+2)\,P_{-1}A_{m+1} - A_{m+1}[P_{-1}+(m+1)E] +\\
&\quad + P_0A_m - A_mP^*_0 + P_1A_{m-1} - A_{m-1}P^*_1 + \dots + P_m = P^*_m.
\end{aligned}\right\} \tag{87}$$

2. Рассмотрим отдельно несколько случаев:

1° Матрица P_{-1} не имеет различных характеристических чисел, отличающихся друг от друга на целое число.

В этом случае при любом $k = 1, 2, 3, \dots$ матрицы P_{-1} и $P_{-1}+kE$ не имеют общих характеристических чисел, и потому (см. гл. VIII, § 3) [3])

---

[1]) Мы будем стремиться к тому, чтобы в ряду (86) было конечное (и притом возможно меньшее) число коэффициентов P^*_m, отличных от нуля.

[2]) Во всех уравнениях, начиная со второго, мы заменим в силу первого уравнения матрицу P^*_{-1} на P_{-1}.

[3]) Можно, впрочем, доказать это и не опираясь на гл. VIII. Интересующее нас положение равносильно утверждению, что матричное уравнение

$$P_{-1}U = U(P_{-1}+kE) \tag{*}$$

имеет только нулевое решение $U = 0$. Поскольку матрицы P_{-1} и $P_{-1}+kE$ не имеют

матричное уравнение

$$P_{-1}U - U(P_{-1} + kE) = T$$

при любой правой части T имеет одно и только одно решение.

Это решение будем обозначать через

$$\Phi_k(P_{-1}, T).$$

Поэтому в уравнениях (87) можно положить все матрицы P_m^* $(m=0, 1, 2, \ldots)$ равными нулю и последовательно определить A_1, A_2. ... при помощи равенств

$$A_1 = \Phi_1(P_{-1}, -P_0), \quad A_2 = \Phi_2(P_{-1}, -P_1 - P_0A_1), \ldots$$

Тогда преобразованная система является системой Коши

$$\frac{dY}{dz} = \frac{P_{-1}}{z} Y,$$

и потому решение X исходной системы (81) имеет вид [1])

$$X = A(z)\, z^{P_{-1}}. \tag{88}$$

2° Среди различных характеристических чисел матрицы P_{-1} имеются числа, разность между которыми является целым числом; при этом матрица P_{-1} имеет простую структуру.

Обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$ характеристические числа матрицы P_{-1}, расположенные так, чтобы имели место неравенства

$$\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n. \tag{89}$$

Не нарушая общности рассуждений, мы можем матрицу P_{-1} заменить любой, ей подобной. Это следует из того, что, умножая обе части уравнения (81) слева на неособенную матрицу T, а справа — на T^{-1}, мы фактически заменяем все P_m на TP_mT^{-1} $(m=-1, 0, 1, 2, \ldots)$ (при этом и X заменяется на TXT^{-1}). Поэтому мы будем считать, что в рассматриваемом случае P_{-1} — диагональная матрица:

$$P_{-1} = \|\lambda_i\delta_{ik}\|_1^n. \tag{90}$$

Введем обозначения для элементов матриц P_m, P_m^* и A_m:

$$P_m = \|p_{ik}^{(m)}\|_1^n, \quad P_m^* = \|p_{ik}^{(m*)}\|_1^n, \quad A_m = \|x_{ik}^{(m)}\|_1^n. \tag{91}$$

Для определения A_1 мы воспользуемся вторым из уравнений (87). Это матричное уравнение можно заменить скалярными уравнениями

$$(\lambda_i - \lambda_k - 1)\, x_{ik}^{(1)} + p_{ik}^{(0)} = p_{ik}^{(0*)} \qquad (i, k = 1, 2, \ldots, n). \tag{92}$$

Если ни одна из разностей $\lambda_i - \lambda_k$ не равна единице, то мы можем положить $P_0^* = 0$. Тогда из (87_2) [2]) $A_1 = \Phi_1(P_{-1}, -P_0)$.

общих характеристических чисел, то существует такой многочлен $f(\lambda)$, для которого

$$f(P_{-1}) = 0, \quad f(P_{-1} + kE) = E.$$

Но из (*) следует:

$$f(P_{-1})\, U = Uf(P_{-1} + kE).$$

Отсюда $U = 0$.

[1]) Формула (88) определяет одну интегральную матрицу системы (81). Произвольная интегральная матрица получается из (88) умножением справа на произвольную постоянную неособенную матрицу C.

[2]) Мы пользуемся обозначениями, введенными при разборе случая 1°.

В этом случае элементы матрицы A_1 однозначно определятся из уравнений (92):

$$x_{ik}^{(1)} = -\frac{p_{ik}^{(0)}}{\lambda_i - \lambda_k - 1} \quad (i,\ k = 1,\ 2,\ \ldots,\ n). \tag{93}$$

Если же при некоторых i, k [1])

$$\lambda_i - \lambda_k = 1,$$

то соответствующее $p_{ik}^{(0*)}$ определяется из (92):

$$p_{ik}^{(0*)} = p_{ik}^{(0)},$$

а соответствующее $x_{ik}^{(1)}$ выбирается совершенно произвольно.

При тех же i и k, при которых $\lambda_i - \lambda_k \neq 1$, мы полагаем:

$$p_{ik}^{(0*)} = 0,$$

а соответствующее $x_{ik}^{(1)}$ находим по формуле (93).

Определив A_1, мы переходим к определению A_2 из третьего уравнения (87). Заменим это матричное уравнение системой n^2 скалярных уравнений:

$$(\lambda_i - \lambda_k - 2)\, x_{ik}^{(2)} = p_{ik}^{(1*)} - p_{ik}^{(1)} - (P_0 A_1 - A_1 P_0^*)_{ik} \tag{94}$$
$$(i,\ k = 1,\ 2,\ \ldots,\ n).$$

Здесь мы поступаем так же, как и при определении A_1.

Если $\lambda_i - \lambda_k \neq 2$, то мы полагаем:

$$p_{ik}^{(1*)} = 0,$$

и тогда из (94) находим:

$$x_{ik}^{(2)} = -\frac{1}{\lambda_1 - \lambda_2 - 2}\,[p_{ik}^{(1)} - (P_0 A_1 - A_1 P_0^*)_{ik}].$$

Если же $\lambda_i - \lambda_k = 2$, то при этих i и k из (94) следует, что

$$p_{ik}^{(1*)} = p_{ik}^{(1)} + (P_0 A_1 - A_1 P_0^*)_{ik}.$$

В этом случае $x_{ik}^{(2)}$ выбирается произвольно.

Продолжая этот процесс далее, мы последовательно определим все матрицы P_{-1}^*, P^*, P_1^*, ... и A_1, A_2, ...

При этом только конечное число из матриц P_m^* будет отлично от нуля, и, как нетрудно видеть, матрица $P^*(z)$ будет иметь вид [2])

$$P^*(z) = \left\| \begin{array}{cccc} \frac{\lambda_1}{z} & a_{12} z^{\lambda_1 - \lambda_2 - 1} & \ldots & a_{1n} z^{\lambda_1 - \lambda_n - 1} \\ 0 & \frac{\lambda_2}{z} & \ldots & a_{2n} z^{\lambda_2 - \lambda_n - 1} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \frac{\lambda_n}{z} \end{array} \right\|, \tag{95}$$

[1]) В силу (89) это возможно лишь при $i < k$.

[2]) P_m^* $(m \geqslant 0)$ может быть отличным от нуля лишь тогда, когда существуют характеристические числа λ_i и λ_k матрицы P_{-1} такие, что $\lambda_i - \lambda_k - 1 = m$ (при этом в силу (89) $i < k$). При данном m каждому такому равенству соответствует элемент $p_{ik}^{(m*)} = a_{ik}$ матрицы P_m^*; этот элемент может быть отличен от нуля. Все остальные элементы матрицы P_m^* равны нулю.

где $a_{ik}=0$, если $\lambda_i-\lambda_k$ не есть целое положительное число, и $a_{ik}=p_{ik}^{(\lambda_i-\lambda_k-1^*)}$, если $\lambda_i-\lambda_k$ является целым положительным числом.

Обозначим через m_i наибольшую целую часть числа $\operatorname{Re}\lambda_i$ [1]):

$$m_i=[\operatorname{Re}\lambda_i] \qquad (i=1,\ 2,\ \ldots,\ n). \tag{96}$$

Тогда в силу (89)

$$m_1\geqslant m_2\geqslant\ldots\geqslant m_n.$$

При этом если $\lambda_i-\lambda_k$ есть целое число, то

$$\lambda_i-\lambda_k=m_i-m_k.$$

Поэтому в выражении (95) канонической матрицы $P^*(z)$ мы можем все разности $\lambda_i-\lambda_k$ заменить на m_i-m_k. Кроме того, мы положим:

$$\widetilde{\lambda}_i=\lambda_i-m_i \qquad (i=1,\ 2,\ \ldots,\ m), \tag{91'}$$

$$M=\|\,m_i\delta_{ik}\,\|_1^n, \qquad U=\begin{Vmatrix}\widetilde{\lambda}_1 & a_{12} & \ldots & a_{1n}\\ 0 & \widetilde{\lambda}_2 & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & \widetilde{\lambda}_n\end{Vmatrix}. \tag{97}$$

Тогда из (95) следует (см. формулу I на стр. 435):

$$P^*(z)=z^M\frac{U}{z}z^{-M}+\frac{M}{z}=D_z(z^Mz^U).$$

Отсюда вытекает, что $Y=z^Mz^U$ представляет собой решение уравнения (85), а

$$X=A(z)\,z^Mz^U \tag{98}$$

является решением уравнения (81)[2]).

3° Переходим к общему случаю. Как было выяснено выше, мы можем, не нарушая общности, матрицу P_{-1} заменить любой матрицей, ей подобной. Мы примем, что матрица P_{-1} имеет нормальную жорданову форму [3])

$$P_{-1}=\{\lambda_1E_1+H_1,\ \lambda_2E_2+H_2,\ \ldots,\ \lambda_uE_u+H_u\}, \tag{99}$$

причем

$$\operatorname{Re}\lambda_1\geqslant\operatorname{Re}\lambda_2\geqslant\ldots\geqslant\operatorname{Re}\lambda_u. \tag{100}$$

Здесь E обозначает единичную матрицу, а H — матрицу, у которой элементы первой «наддиагонали» равны единице, а остальные элементы равны нулю. Порядки матриц E_i и H_i в различных диагональных клетках будут, вообще говоря, различными; эти порядки совпадают со степенями соответствующих элементарных делителей матрицы P_{-1} [4]).

В соответствии с представлением (99) матрицы P_{-1} разобьем все матрицы P_m, P_m^*, A_m на блоки:

$$P_m=(P_{ik}^{(m)})_1^u,\quad P_m^*=(P_{ik}^{(m^*)})_1^u,\quad A_m=(X_{ik}^{(m)})_1^u.$$

---

[1]) То есть m_i — наибольшее целое число, не превосходящее $\operatorname{Re}\lambda_i$ ($i=1,\ 2,\ \ldots,\ n$).

[2]) Специальный вид матриц (97) соответствует каноническому виду матрицы P_{-1}. Если матрица P_{-1} не имеет каноническую форму, то матрицы M и U в (98) подобны матрицам (97).

[3]) См. гл. VI, § 6.

[4]) Для сокращения обозначений мы не пишем при E_i и H_i индекса, указывающего порядок этих матриц.

Тогда второе из уравнений (87) может быть заменено системой уравнений

$$(\lambda_i E_i + H_i) X_{ik}^{(1)} - X_{ik}^{(1)} [(\lambda_k + 1) E_k + H_k] + P_{ik}^{(0)} = P_{ik}^{(0*)} \qquad (101)$$
$$(i, k = 1, 2, \ldots, u),$$

которые могут быть еще переписаны так:

$$(\lambda_i - \lambda_k - 1) X_{ik}^{(1)} + H_i X_{ik}^{(1)} - X_{ik}^{(1)} H_k + P_{ik}^{(0)} = P_{ik}^{(0*)} \quad (i, k = 1, 2, \ldots, u). \qquad (102)$$

Пусть [1])

$$X_{ik}^{(1)} = \begin{Vmatrix} x_{11} & x_{12} \ldots \\ x_{21} & x_{22} \ldots \\ \ldots & \ldots \\ \ldots & \ldots \end{Vmatrix} = \| x_{st} \|, \quad P_{ik}^{(0)} = \| p_{st}^{(0)} \|, \quad P_{ik}^{(0*)} = \| p_{st}^{(0*)} \|.$$

Тогда матричное уравнение (102) (при фиксированных i и k) можно заменить системой скалярных уравнений вида [2])

$$(\lambda_i - \lambda_k - 1) x_{st} + x_{s+1,\,t} - x_{s,\,t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \qquad (103)$$
$$(x_{v+1,\,t} = x_{s,\,0} = 0; \quad s = 1, 2, \ldots, v; \quad t = 1, 2, \ldots, w),$$

где v и w — порядки матриц $\lambda_i E_i + H_i$ и $\lambda_k E_k + H_k$ в (99).

Если $\lambda_i - \lambda_k \neq 1$, то в системе (103) можно положить все $p_{st}^{(0*)} = 0$ и однозначно определить все x_{st} из рекуррентных соотношений (103). Это означает, что в матричном уравнении (102) мы полагаем

$$P_{ik}^{(0*)} = 0$$

и однозначно определяем $X_{ik}^{(1)}$.

Если $\lambda_i - \lambda_k = 1$, то соотношения (103) принимают вид

$$x_{s+1,\,t} - x_{s,\,t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \qquad (104)$$
$$(s = 1, 2, \ldots, v; \; t = 1, 2, \ldots, w; \; x_{v+1,\,t} = x_{s,\,0} = 0).$$

Нетрудно показать, что из уравнений (104) можно так определить элементы x_{st} матрицы $X_{ik}^{(1)}$, чтобы матрица $P_{ik}^{(0*)}$ имела в соответствии со своими размерами $(v \times w)$ вид

$$\begin{Vmatrix} a_0 & 0 & \ldots & 0 \\ a_1 & a_0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{v-1} & a_{v-2} & \ldots & a_0 \end{Vmatrix}, \quad (v = w) \qquad \begin{Vmatrix} a_0 & 0 & \ldots & 0 & 0 \ldots 0 \\ a_1 & a_0 & \ldots & 0 & 0 \ldots 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{v-1} & & & a_1 & a_0 \; 0 \ldots 0 \end{Vmatrix}, \quad (v < w)$$

[1]) Для сокращения обозначений мы опускаем индексы i, k при обозначении элементов матриц X_{ik}, $P_{ik}^{(0)}$, $P_{ik}^{(0*)}$.

[2]) Рекомендуем читателю вспомнить свойства матрицы H, разобранные на стр. 24—25.

$$\left\|\begin{array}{ccccc} 0 & 0 & \cdots & & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & & 0 \\ a_0 & 0 & \cdots & & 0 \\ a_1 & a_0 & \cdots & & 0 \\ \cdots & \cdots & & & \\ \cdots & \cdots & \cdots & & \cdots \\ \cdots & & \cdots & \cdots & \cdots \\ \cdots & & & \cdots & \cdots \\ & & & \cdots & \cdots \\ a_{w-1} & \cdots & & a_1 & a_0 \end{array}\right\| \quad (v > w) \tag{105}$$

Про матрицы (105) будем говорить, что они имеют *правильную нижнюю треугольную форму*[1]).

Из третьего уравнения (87) мы определяем матрицу A_2. Это уравнение можно заменить системой уравнений

$$(\lambda_i - \lambda_k - 2)\, X_{ik}^{(2)} + H_i X_{ik}^{(2)} - X_{ik}^{(2)} H_k + \{P_0 A_1 - A_1 P_0\}_{ik} + P_{ik}^{(1)} = P_{ik}^{(1*)} \tag{106}$$
$$(i,\ k = 1,\ 2,\ \ldots,\ u).$$

Аналогично тому, как это было при определении A_1, если $\lambda_i - \lambda_k \neq 2$, то из соответствующего уравнения (106) матрица $X_{ik}^{(2)}$ определяется однозначно при $P_{ik}^{(1*)} = 0$. Если же $\lambda_i - \lambda_k = 2$, то можно так определить матрицу $X_{ik}^{(2)}$, чтобы матрица $P_{ik}^{(1*)}$ имела правильную нижнюю треугольную форму.

Продолжая этот процесс, мы определим последовательно все матричные коэффициенты A_1, A_2, ... и P_{-1}^*, P_0^*, P_1^*, ... При этом только конечное число из коэффициентов P_m^* будет отлично от нуля, и матрица $P^*(z)$ будет иметь следующий блочный вид[2]):

$$P^*(z) = \begin{pmatrix} \dfrac{\lambda_1 E_1 + H_1}{z} & B_{12} z^{\lambda_1 - \lambda_2 - 1} & \cdots & B_{1u} z^{\lambda_1 - \lambda_u - 1} \\ 0 & \dfrac{\lambda_2 E_2 + H_2}{z} & \cdots & B_{2u} z^{\lambda_2 - \lambda_u - 1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \dfrac{\lambda_u E_u + H_u}{z} \end{pmatrix}, \tag{107}$$

где

$$B_{ik} = \begin{cases} 0, & \text{если } \lambda_i - \lambda_k \text{ не есть целое положительное число;} \\ P_{ik}^{(\lambda_i - \lambda_k - 1^*)}, & \text{если } \lambda_i - \lambda_k = \text{целому положительному числу} \\ & (i,\ k = 1,\ 2,\ \ldots,\ u). \end{cases}$$

[1]) Аналогично определяются верхние правильные треугольные матрицы.

Из уравнений (104) все элементы матрицы $X_{ik}^{(1)}$ не определяются однозначно; имеется некоторый произвол в выборе элементов x_{st}. Это видно и непосредственно из уравнения (102): при $\lambda_i - \lambda_k = 1$ к матрице $X_{ik}^{(1)}$ можно прибавить произвольную матрицу, перестановочную с H, т. е. произвольную правильную верхнюю треугольную матрицу.

[2]) Размеры квадратных матриц E_i, H_i и прямоугольных матриц B_{ik} определяются размерами диагональных клеток в жордановой матрице P_{-1}, т. е. степенями элементарных делителей матрицы P_{-1}.

Все матрицы B_{ik} $(i, k=1, 2, \ldots, u;\ i<k)$ имеют правильную нижнюю треугольную форму.

Как и в предыдущем случае, обозначим через m_i целую часть $\operatorname{Re}\lambda_i$

$$m_i=[\operatorname{Re}\lambda_i] \qquad (i=1, 2, \ldots, u) \tag{108}$$

и положим

$$\lambda_i=m_i+\widetilde{\lambda}_i \qquad (i=1, 2, \ldots, u). \tag{108'}$$

Тогда снова в выражении (107) для $P^*(z)$ мы можем всюду разность $\lambda_i-\lambda_k$ заменить разностью m_i-m_k. Вводя диагональную матрицу с целыми элементами M и верхнюю треугольную матрицу U при помощи равенств [1])

$$M=(m_iE_i\delta_{ik})_1^u,\quad U=\begin{pmatrix}\widetilde{\lambda}_1E_1+H_1 & B_{12} & \ldots & B_{1u}\\ 0 & \widetilde{\lambda}_2E_2+H_2 & \ldots & B_{2u}\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & \widetilde{\lambda}_uE_u+H_u\end{pmatrix}, \tag{109}$$

мы, исходя из (107), легко получим следующее представление для матрицы $P^*(z)$:

$$P^*(z)=z^M\frac{U}{z}\cdot z^{-M}+\frac{M}{z}=D_z(z^Mz^U).$$

Отсюда следует, что решение уравнения (85) может быть задано в виде

$$Y=z^Mz^U,$$

а решение уравнения (81) может быть представлено так:

$$X=A(z)\,z^Mz^U. \tag{110}$$

Здесь $A(z)$ — матричный ряд (84), M — диагональная матрица с постоянными целыми элементами, U — постоянная треугольная матрица. Матрицы M и U определяются равенствами (108), (108') и (109) [2]).

3. Переходим теперь к доказательству сходимости ряда

$$A(z)=E+A_1z+A_2z^2+\ldots$$

Воспользуемся леммой, которая представляет и самостоятельный интерес.

Лемма. *Если ряд*

$$x=a_0+a_1z+a_2z^2+\ldots \tag{111}$$

формально удовлетворяет системе [3])

$$\frac{dx}{dz}=P(z)\,x, \tag{112}$$

для которой $z=0$ является регулярной особой точкой, то ряд (111) *сходится в любой окрестности точки $z=0$, в которой сходится разложение в ряд* (82) *для матрицы коэффициентов $P(z)$.*

---

[1]) Здесь разбиение на блоки соответствует разбиению матриц P_{-1} и $P^*(z)$.

[2]) См. сноску [2]) к стр. 452.

[3]) Здесь $x=(x_1, x_2, \ldots, x_n)$ — столбец из неизвестных функций; $a_0, a_1, a_2, \ldots$ — постоянные столбцы; $P(z)$ — квадратная матрица коэффициентов.

Доказательство. Пусть

$$P(z)=\frac{P_{-1}}{z}+\sum_{q=0}^{\infty}P_q z^q,$$

где ряд $\sum_{m=0}^{\infty}P_m z^m$ сходится при $|z|<r$. Тогда существуют такие положительные постоянные p_{-1} и p, что [1])

$$\operatorname{mod}P_{-1}\leqslant p_{-1}I,\ \operatorname{mod}P_m\leqslant\frac{p}{r^m}I,\ I=\|1\|\quad(m=0,\ 1,\ 2,\ \ldots).\tag{113}$$

Подставляя в (112) вместо x ряд (111) и приравнивая между собой коэффициенты при одинаковых степенях в обеих частях равенства (112), получим бесконечную систему векторных (столбцевых) равенств

$$\left.\begin{aligned}
P_{-1}a_0&=0,\\
(E-P_{-1})\,a_1&=P_0a_0,\\
(2E-P_{-1})\,a_2&=P_0a_1+P_1a_0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(mE-P_{-1})\,a_m&=P_0a_{m-1}+P_1a_{m-2}+\ldots+P_{m-1}a_0\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\}\tag{114}$$

Нам достаточно доказать, что какой-либо остаток ряда (111)

$$x^{(k)}=a_kz^k+a_{k+1}z^{k+1}+\ldots\tag{115}$$

сходится в окрестности точки $z=0$. Число k подчиним неравенству

$$k>np_{-1}.$$

Тогда число k будет превосходить модули всех характеристических чисел матрицы P_{-1} [2]) и потому при $m\geqslant k$ будем иметь $|mE-P_{-1}|\neq 0$ и

$$(mE-P_{-1})^{-1}=\frac{1}{m}\left(E-\frac{P_{-1}}{m}\right)^{-1}=\frac{1}{m}E+\frac{1}{m^2}P_{-1}+\frac{1}{m^3}P_{-1}^2+\ldots\tag{116}$$

$$(m=k,\ k+1,\ \ldots).$$

В последней части этого равенства стоит сходящийся матричный ряд. Пользуясь этим рядом, мы из (114) можем все коэффициенты ряда (115) выразить однозначно через $a_0,\ a_1,\ \ldots,\ a_{k-1}$ при помощи рекуррентных

---

[1]) Относительно определения модуля матрицы см. стр. 431.

[2]) Если λ_0 — характеристическое число матрицы $A=\|a_{ik}\|_1^n$, то $|\lambda_0|\leqslant n\max_{1\leqslant i,\ k\leqslant n}|a_{ik}|$. Действительно, пусть $Ax=\lambda_0x$, где $x=(x_1,\ x_2,\ \ldots,\ x_n)\neq 0$. Тогда

$$\lambda_0x_i=\sum_{k=1}^{n}a_{ik}x_k\qquad(i=1,\ 2,\ \ldots,\ n).$$

Пусть $|x_j|=\max\{|x_1|,\ |x_2|,\ \ldots,\ |x_n|\}$. Тогда

$$|\lambda_0|\,|x_j|\leqslant\sum_{k=1}^{n}|a_{jk}|\,|x_k|\leqslant|x_j|\,n\max_{1\leqslant i,\ k\leqslant n}|a_{ik}|.$$

Сокращая на $|x_j|$, получим нужное неравенство.

соотношений

$$a_m = \left(\frac{1}{m}E + \frac{1}{m^2}P_{-1} + \frac{1}{m^3}P_{-1}^2 + \ldots\right)(f_{m-1} + P_0 a_{m-1} + \ldots + P_{m-k-1}a_k)$$
$$(m = k,\ k+1,\ \ldots), \tag{117}$$

где

$$f_{m-1} = P_{m-k}a_{k-1} + \ldots + P_{m-1}a_0 \quad (m = k,\ k+1,\ \ldots). \tag{118}$$

Заметим, что ряд (115) формально удовлетворяет дифференциальному уравнению

$$\frac{dx^{(k)}}{dz} = P(z)\,x^{(k)} + f(z), \tag{119}$$

где

$$f(z) = \sum_{m=k-1}^{\infty} f_m z^m = P(z)\,(a_0 + a_1 z + \ldots + a_{k-1}z^{k-1}) -$$
$$- a_1 - 2a_2 z - \ldots - (k-1)\,a_{k-1}z^{k-2}. \tag{120}$$

Из (120) вытекает, что ряд

$$\sum_{m=k-1}^{\infty} f_m z^m$$

сходится при $|z| < r$, и потому существует такое число $N > 0$, что[1]

$$\operatorname{mod} f_m \leqslant \left\| \frac{N}{r^m} \right\| \quad (m = k-1,\ k,\ \ldots). \tag{121}$$

Из вида рекуррентных соотношений (117) следует, что, заменив в них матрицы P_{-1}, P_q, f_{m-1} мажорантными матрицами $p_{-1}I$, $\frac{p}{r^q}I$, $\left\| \frac{N}{r^{m-1}} \right\|$, а столбец a_m столбцом $\| \alpha_m \|$[2] $(m = k,\ k+1,\ \ldots;\ q = 0,\ 1,\ 2,\ \ldots)$, мы получим соотношения, определяющие верхние границы $\| \alpha_m \|$ для $\operatorname{mod} a_m$:

$$\operatorname{mod} a_m \leqslant \| \alpha_m \|. \tag{122}$$

Следовательно, ряд

$$\xi^{(k)} = \alpha_k z^k + a_{k+1}z^{k+1} + \ldots \tag{123}$$

после почленного умножения на столбец $\| 1 \|$ будет мажорантным рядом для ряда (115).

Заменив в (119) матричные коэффициенты P_{-1}, P_q, f_m рядов

$$P(z) = \frac{P_{-1}}{z} + \sum_{q=0}^{\infty} P_q z^q, \qquad f(z) = \sum_{m=k-1}^{\infty} f_m z^m$$

соответствующими мажорантными матрицами $p_{-1}I$, $\frac{p}{r^q}I$, $\left\| \frac{N}{r^m} \right\|$, а также

---

[1]) Здесь $\left\| \frac{N}{r^m} \right\|$ обозначает столбец, у которого все элементы равны одному и тому же числу $\frac{N}{r^m}$.

[2]) Здесь $\| \alpha_m \|$ обозначает столбец $(\alpha_m, \alpha_m, \ldots, \alpha_m)$ (α_m — число; $m = k,\ k+1,\ \ldots$)

заменив $x^{(k)}$ на $\|\xi^{(k)}\|$, мы получим дифференциальное уравнение для $\xi^{(k)}$:

$$\frac{d\xi^{(k)}}{dz} = n\left(\frac{p_{-1}}{z} + \frac{p}{1-\frac{z}{r}}\right)\xi^{(k)} + \frac{N\frac{z^{k-1}}{r^{k-1}}}{1-\frac{z}{r}}. \tag{124}$$

Это линейное дифференциальное уравнение имеет частное решение

$$\xi^{(k)} = \frac{N}{r^{k-1}} \frac{z^{np-1}}{\left(1-\frac{z}{r}\right)^{npr}} \int_0^z z^{k-np-1}\left(1-\frac{z}{r}\right)^{npr-1} dz, \tag{125}$$

которое регулярно в точке $z=0$ и в окрестности этой точки разлагается в сходящийся при $|z|<r$ степенной ряд (123).

Из сходимости мажорантного ряда (123) следует сходимость ряда (115) при $|z|<r$. Лемма доказана.

Замечание 1. Приведенное доказательство позволяет определить все регулярные в особой точке решения дифференциальной системы (112), если таковые существуют.

Для существования регулярных решений (не равных тождественно нулю) необходимо и достаточно, чтобы матрица вычетов P_{-1} имела целое неотрицательное характеристическое число. Если s — наибольшее такое целое характеристическое число, то из первых $s+1$ уравнений (114) можно определить не обращающиеся одновременно в нуль столбцы a_0, a_1, ..., a_s, поскольку определитель соответствующей системы линейных однородных уравнений равен нулю:

$$\Delta = |P_{-1}|\,|E-P_{-1}|\,|\ldots|\,|sE-P_{-1}| = 0.$$

Из остальных уравнений (114) столбцы a_{s+1}, a_{s+2}, ... однозначно выразятся через a_0, a_1, ..., a_s. Полученный ряд (111) сходится согласно лемме. Таким образом, линейно независимые решения первых $s+1$ уравнений (114) определяют все линейно независимые регулярные в особой точке $z=0$ решения системы (112).

Если $z=0$ есть особая точка, то задание начального значения a_0 для регулярного в этой точке решения (111) (если таковое существует) не определяет однозначно этого решения. Однако решение, регулярное в регулярной особой точке, определяется однозначно, если заданы a_0, a_1, ..., a_s, т. е. если заданы начальные значения при $z=0$ самого решения и его первых s производных (s — наибольшее неотрицательное целое характеристическое число матрицы вычетов P_{-1}).

Замечание 2. Доказанная лемма сохраняет свою силу и при $P_{-1}=0$. В этом случае в доказательстве леммы в качестве p_{-1} можно взять любое положительное число. При $P_{-1}=0$ лемма утверждает известное положение о существовании регулярного решения в окрестности регулярной точки системы. В этом случае решение однозначно определяется заданием a_0.

4. Пусть дана система

$$\frac{dX}{dz} = P(z)\,X, \tag{126}$$

где

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m$$

и ряд, стоящий в правой части, сходится при $|z|<r$.

Пусть, далее, полагая

$$X = A(z)\,Y \tag{127}$$

и подставляя вместо $A(z)$ ряд

$$A(z) = A_0 + A_1 z + A_2 z^2 + \ldots, \tag{128}$$

мы после формальных преобразований получаем:

$$\frac{dY}{dz} = P^*(z)\,Y, \tag{129}$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m,$$

причем здесь, как и в выражении для $P(z)$, ряд в правой части сходится при $|z| < r$.

Докажем, что и ряд (128) сходится в окрестности $|z| < r$ точки $z = 0$.

Действительно, из (126), (127) и (129) следует, что ряд (128) формально удовлетворяет следующему дифференциальному матричному уравнению

$$\frac{dA}{dz} = P(z)\,A - AP^*(z). \tag{130}$$

Мы будем рассматривать A как вектор (столбец) в пространстве всех матриц n-го порядка, т. е. в пространстве n^2 измерений. Если мы в этом пространстве определим линейный оператор $\widehat{P}(z)$ над матрицей A, аналитически зависящий от параметра z, при помощи равенства

$$\widehat{P}(z)\,[A] = P(z)\,A - AP^*(z), \tag{131}$$

то дифференциальное уравнение (130) можно будет записать в виде

$$\frac{dA}{dz} = \widehat{P}(z)\,[A]. \tag{132}$$

Правую часть этого уравнения можно рассматривать как произведение матрицы $\widehat{P}(z)$ порядка n^2 на столбец A из n^2 элементов. Из формулы (131) видно, что точка $z = 0$ является регулярной особой точкой для системы (132). Ряд (128) формально удовлетворяет этой системе. Поэтому, применяя лемму, заключаем, что ряд (128) сходится в окрестности $|z| < r$ точки $z = 0$.

В частности, сходится и ряд для $A(z)$ в формуле (110).

Таким образом, нами доказана следующая

Теорема 2. *Всякая система*

$$\frac{dX}{dz} = P(z)\,X \tag{133}$$

с регулярной особой точкой $z = 0$:

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m$$

имеет решение вида

$$X = A(z)\, z^M z^U, \tag{134}$$

где $A(z)$ — *матричная функция, регулярная при* $z=0$ *и обращающаяся в этой точке в единичную матрицу* E, *а* M *и* U — *постоянные матрицы, причем* M *имеет простую структуру и целые характеристические числа, и разность между любыми двумя различными характеристическими числами матрицы* U *не есть целое число.*

Если матрица P_{-1} *приводится к жордановой форме при помощи неособенной матрицы* T

$$P_{-1} = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_s E_s + H_s\} T^{-1} \tag{135}$$
$$(\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_s),$$

то можно взять M *и* U *в виде*

$$M = T\{m_1 E_1,\ m_2 E_2,\ \ldots,\ m_s E_s\} T^{-1}, \tag{136}$$

$$U = T \begin{pmatrix} \widetilde{\lambda}_1 E_1 + H_1 & B_{12} & \ldots & B_{1s} \\ 0 & \widetilde{\lambda}_2 E_2 + H_2 & \ldots & B_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \widetilde{\lambda}_s E_s + H_s \end{pmatrix} T^{-1}, \tag{137}$$

где

$$m_i = [\lambda_i], \quad \widetilde{\lambda}_i = \lambda_i - m_i \quad (i = 1, 2, \ldots, s), \tag{138}$$

B_{ik} — *правильные нижние треугольные матрицы* $(i,\ k = 1,\ 2,\ \ldots,\ s)$, *причем* $B_{ik} = 0$, *если* $\lambda_i - \lambda_k$ *не есть целое положительное число* $(i,\ k = 1, 2, \ldots, s)$.

В частном случае, когда ни одна из разностей $\lambda_i - \lambda_k$ $(i,\ k = 1, 2, \ldots, s)$ *не равна целому положительному числу, в формуле* (134) *можно положить* $M = 0$, $U = P_{-1}$, *т. е. в этом случае решение представимо в виде*

$$X = A(z)\, z^{P_{-1}}. \tag{139}$$

Замечание 1. Обращаем внимание на то, что в настоящем параграфе был установлен алгоритм для определения коэффициентов ряда $A(z) = \sum_{m=0}^{\infty} A_m z^m$ $(A_0 = E)$ через коэффициенты P_m ряда для $P(z)$. Кроме того, доказанная теорема определяет и интегральную подстановку V, на которую умножается решение (134) при однократном обходе особой точки $z=0$ в положительном направлении:

$$V = e^{2\pi i U}.$$

Замечание 2. Из формулировки теоремы следует, что

$$B_{ik} = 0 \text{ при } \widetilde{\lambda}_i \neq \widetilde{\lambda}_k \quad (i,\ k = 1, 2, \ldots, s).$$

Поэтому матрицы

$$\widetilde{\Lambda} = T\{\widetilde{\lambda}_1 E_1,\ \widetilde{\lambda}_2 E_2,\ \ldots,\ \widetilde{\lambda}_s E_s\} T^{-1} \text{ и } \widetilde{U} = T \begin{pmatrix} 0 & B_{12} & \ldots & B_{1s} \\ 0 & 0 & \ldots & B_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{pmatrix} T^{-1} \tag{140}$$

перестановочны между собой:

$$\widetilde{\Lambda}\widetilde{U} = \widetilde{U}\widetilde{\Lambda}.$$

Отсюда

$$z^M z^U = z^M z^{\widetilde{\Lambda}+\widetilde{U}} = z^M z^{\widetilde{\Lambda}} z^{\widetilde{U}} = z^{\Lambda} z^{\widetilde{U}}, \tag{141}$$

где

$$\Lambda = M + \widetilde{\Lambda} = T\{\lambda_1, \lambda_2, \ldots, \lambda_n\} T^{-1}, \tag{142}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все характеристические числа матрицы P_{-1}, расположенные в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$.

С другой стороны,

$$z^{\widetilde{U}} = h(\widetilde{U}),$$

где $h(\lambda)$ — интерполяционный многочлен Лагранжа-Сильвестра для функции $f(\lambda) = z^{\lambda}$.

Поскольку все характеристические числа матрицы $\widetilde{U}$ равны нулю, то $h(\lambda)$ линейно зависит от $f(0), f'(0), \ldots, f^{(g-1)}(0)$, т. е. от $1, \ln z, \ldots, (\ln z)^{g-1}$ (g — наименьший показатель, при котором $\widetilde{U}^g = 0$). Поэтому

$$h(\lambda) = \sum_{j=0}^{g-1} h_j(\lambda)(\ln z)^j$$

и потому

$$z^{\widetilde{U}} = h(\widetilde{U}) = \sum_{j=0}^{g-1} h_j(\widetilde{U})(\ln z)^j = T \begin{pmatrix} 1 & q_{12} & \ldots & q_{1n} \\ 0 & 1 & \ldots & q_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix} T^{-1}, \tag{143}$$

где q_{ij} ($i, j = 1, 2, \ldots, n;\ i < j$) — многочлены от $\ln z$ степени ниже g.

В силу (134), (141), (142) и (143) частное решение системы (126) можно взять в виде

$$X = A(z) \begin{Vmatrix} z^{\lambda_1} & 0 & \ldots & 0 \\ 0 & z^{\lambda_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & z^{\lambda_n} \end{Vmatrix} \begin{Vmatrix} 1 & q_{12} & \ldots & q_{1n} \\ 0 & 1 & \ldots & q_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{Vmatrix}. \tag{144}$$

Здесь $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы P_{-1}, расположенные в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$, а $q_{ij}(i, j = 1, 2, \ldots, n;\ i < j)$ — многочлены от $\ln z$ степени не выше $g-1$, где g — максимальное количество характеристических чисел λ_i, отличающихся между собой на целое число; $A(z)$ — матричная функция, регулярная в точке $z = 0$, причем $A(0) = T$ ($|T| \neq 0$). Если матрица P_{-1} имеет жорданову форму, то $T = E$.

§ 11. Приводимые аналитические системы

В качестве приложения теоремы предыдущего параграфа выясним, в каких случаях система

$$\frac{dX}{dt} = Q(t)X, \tag{145}$$

где

$$Q(t)=\sum_{m=1}^{\infty}\frac{Q_m}{t^m} \tag{146}$$

— сходящийся ряд при $t>t_0$, является приводимой (по Ляпунову), т. е. в каких случаях существует решение системы вида

$$X=L(t)\,e^{Bt}, \tag{147}$$

где $L(t)$ — матрица Ляпунова (т. е. $L(t)$ удовлетворяет условиям 1° — 3° на стр. 422), а B — постоянная матрица [1]). Здесь X, Q — матрицы с комплексными элементами, а t — вещественный аргумент.

Сделаем преобразование

$$z=\frac{1}{t}.$$

Тогда система (145) перепишется в виде

$$\frac{dX}{dz}=P(z)\,X, \tag{148}$$

где

$$P(z)=-z^{-2}Q\left(\frac{1}{z}\right)=-\frac{Q_1}{z}-\sum_{m=0}^{\infty}Q_{m+2}z^m. \tag{149}$$

Ряд, стоящий в правой части выражения для $P(z)$, сходится при $|z|<\frac{1}{t_0}$. Могут представиться два случая:

1) $Q_1=0$. В этом случае точка $z=0$ не является особой для системы (148). Эта система имеет решение, регулярное и нормированное в точке $z=0$. Это решение задается сходящимся степенным рядом

$$X(z)=E+X_1z+X_2z^2+\ldots \qquad \left(|z|<\frac{1}{t_0}\right).$$

Полагая

$$L(t)=X\left(\frac{1}{t}\right), \quad B=0,$$

получим искомое представление (147). Система приводима.

2) $Q_1\neq 0$. В этом случае система (148) имеет регулярную особую точку в точке $z=0$.

Не нарушая общности рассуждений, можно считать матрицу вычетов $P_{-1}=-Q_1$ приведенной к жордановой форме, в которой диагональные элементы $\lambda_1, \lambda_2, \ldots, \lambda_n$ расположены в порядке $\operatorname{Re}\lambda_1\geqslant\operatorname{Re}\lambda_2\geqslant\ldots\geqslant\operatorname{Re}\lambda_n$.

Тогда в формуле (144) $T=E$, и потому система (148) имеет решение

$$X=A(z)\begin{Vmatrix} z^{\lambda_1} & 0 & \ldots & 0\\ 0 & z^{\lambda_2} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & z^{\lambda_n}\end{Vmatrix}\begin{Vmatrix} 1 & q_{12} & \ldots & q_{1n}\\ 0 & 1 & \ldots & q_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & 1\end{Vmatrix},$$

[1]) Если имеет место равенство (147), то преобразование Ляпунова $X=L(t)\,Y$ переводит систему (145) в систему $\frac{dY}{dt}=BY$.

где функция $A(z)$ регулярна при $z=0$ и принимает в этой точке значение E, а q_{ik} $(i, k=1, 2, \ldots, n;\ i<k)$ — многочлены от $\ln z$. Заменяя здесь z на $\frac{1}{t}$, будем иметь:

$$X = A\left(\frac{1}{t}\right)\begin{Vmatrix} \left(\frac{1}{t}\right)^{\lambda_1} & 0 & \ldots & 0 \\ 0 & \left(\frac{1}{t}\right)^{\lambda_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \left(\frac{1}{t}\right)^{\lambda_n} \end{Vmatrix}\begin{Vmatrix} 1 & q_{12}\left(\ln\frac{1}{t}\right) & \ldots & q_{1n}\left(\ln\frac{1}{t}\right) \\ 0 & 1 & \ldots & q_{2n}\left(\ln\frac{1}{t}\right) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{Vmatrix}. \tag{150}$$

Так как преобразование $X = A\left(\frac{1}{t}\right)Y$ является преобразованием Ляпунова, то система (145) будет приводимой к некоторой системе с постоянными коэффициентами в том и только в том случае, когда произведение

$$L_1(t) = \begin{Vmatrix} t^{-\lambda_1} & 0 & \ldots & 0 \\ 0 & t^{-\lambda_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & t^{-\lambda_n} \end{Vmatrix}\begin{Vmatrix} 1 & q_{12}\left(\ln\frac{1}{t}\right) & \ldots & q_{1n}\left(\ln\frac{1}{t}\right) \\ 0 & 1 & \ldots & q_{2n}\left(\ln\frac{1}{t}\right) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{Vmatrix} e^{-Bt}, \tag{151}$$

где B — некоторая постоянная матрица будет матрицей Ляпунова, т. е. когда матрицы $L_1(t)$, $\frac{dL_1}{dt}$ и $L_1^{-1}(t)$ будут ограничены. При этом, как следует из теоремы Еругина (§ 4), матрицу B можно считать матрицей с вещественными характеристическими числами.

Из ограниченности матриц $L_1(t)$ и $L_1^{-1}(t)$ при $t>t_0$ вытекает, что все характеристические числа матрицы B должны равняться нулю. Это следует из выражения для e^{Bt} и e^{-Bt}, получаемого из (151). Кроме того, все числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ должны быть чисто мнимыми, поскольку согласно (151) из ограниченности элементов последней строки в $L_1(t)$ и первого столбца в $L_1^{-1}(t)$ вытекает, что $\operatorname{Re}\lambda_n \geqslant 0$ и $\operatorname{Re}\lambda_1 \leqslant 0$.

Но если все характеристические числа матрицы P_{-1} чисто мнимы, то разность между любыми двумя различными характеристическими числами матрицы P_{-1} не равна целому числу. Поэтому имеет место формула (139):

$$X = A(z)\,z^{P-1} = A\left(\frac{1}{t}\right)t^{Q_1},$$

и для приводимости системы необходимо и достаточно, чтобы матрица

$$L_2(t) = t^{Q_1}e^{-Bt} \tag{152}$$

вместе со своей обратной была бы ограничена при $t>t_0$.

Поскольку все характеристические числа матрицы B должны равняться нулю, то минимальный многочлен для матрицы B имеет вид λ^d. Обозначим через

$$\psi(\lambda) = (\lambda-\mu_1)^{c_1}(\lambda-\mu_2)^{c_2}\ \ldots\ (\lambda-\mu_u)^{c_u} \qquad (\mu_i \neq \mu_k \quad \text{при} \quad i \neq k)$$

минимальный многочлен матрицы Q_1. Поскольку $Q_1 = -P_{-1}$, то числа μ_1, μ_2, ..., μ_u отличаются знаком от соответствующих чисел λ_l и потому все они — чисто мнимые числа. Тогда [см. формулы (12), (13) на стр. 421]

$$t^{Q_1} = \sum_{k=1}^{u} [U_{k0} + U_{k1} \ln t + \ldots + U_{k,\, c_k - 1} (\ln t)^{c_k - 1}]\, t^{\mu_k}, \tag{153}$$

$$e^{Bt} = V_0 + V_1 t + \ldots + V_{d-1} t^{d-1}. \tag{154}$$

Подставляя эти выражения в равенство

$$L_2(t)\, e^{Bt} = t^{Q_1},$$

получим:

$$[L_2(t)\, V_{d-1} + (*)]\, t^{d-1} = Z_0(t)\, (\ln t)^{c-1}, \tag{155}$$

где c — наибольшее из чисел $c_1, c_2, \ldots, c_n$, $(*)$ обозначает матрицу, стремящуюся к нулю при $t \to \infty$, а $Z_0(t)$ — ограниченная матрица при $t > t_0$.

Так как матрицы, стоящие в левой и правой частях равенства (155), должны иметь одинаковый порядок роста при $t \to \infty$, то

$$d = c = 1,$$

т. е.

$$B = 0,$$

и матрица Q_1 имеет простые элементарные делители.

Обратно, если матрица Q_1 имеет простые элементарные делители и чисто мнимые характеристические числа $\mu_1, \mu_2, \ldots, \mu_n$, то

$$X = A(z)\, z^{-Q_1} = A(z) \,\|\, z^{-\mu_i \delta_{ik}} \,\|_1^n$$

есть решение системы (149). Полагая здесь $z = \frac{1}{t}$, найдем:

$$X = A\left(\frac{1}{t}\right) \| \, t^{\mu_i \delta_{ik}} \, \|_1^n.$$

Функция $X(t)$ вместе с $\frac{dX(t)}{dt}$ и обратной матрицей $X^{-1}(t)$ ограничена при $t > t_0$. Поэтому система приводима ($B = 0$). Нами доказана[1])

Теорема 3. *Система*

$$\frac{dX}{dt} = Q(t)\, X,$$

где матрица $Q(t)$ представима сходящимся при $t > t_0$ рядом

$$Q(t) = \frac{Q_1}{t} + \frac{Q_2}{t^2} + \ldots,$$

является приводимой в том и только в том случае, если у матрицы вычетов Q_1 все элементарные делители простые и все характеристические числа чисто мнимы.

[1]) См. работу Еругина [11], стр. 21—23. Там эта теорема доказана для случая, когда матрица Q_1 не имеет различных характеристических чисел, отличающихся между собой на целое число.

§ 12. Аналитические функции от многих матриц и их применение к исследованию дифференциальных систем. Работы И. А. Лаппо-Данилевского

Аналитическая функция от m матриц n-го порядка $X_1, X_2, \ldots, X_m$ может быть задана при помощи ряда

$$F(X_1, X_2, \ldots, X_m) = \alpha_0 + \sum_{\nu=1}^{\infty} \sum_{j_1, j_2, \ldots, j_\nu}^{(1 \ldots m)} \alpha_{j_1 j_2 \ldots j_\nu} X_{j_1} X_{j_2} \ldots X_{j_\nu}, \qquad (156)$$

сходящегося для всех матриц m-го порядка X_j, удовлетворяющих неравенствам

$$\operatorname{mod} X_j < R_j \qquad (j = 1, 2, \ldots, m). \qquad (157)$$

Здесь коэффициенты

$$\alpha_0, \ \alpha_{j_1 j_2 \ldots j_\nu} \qquad (j_1, j_2, \ldots, j_\nu = 1, 2, \ldots, m; \ \nu = 1, 2, 3, \ldots)$$

— комплексные числа, $R_j (j = 1, 2, \ldots, m)$ — постоянные матрицы n-го порядка с положительными элементами и $X_j (j = 1, 2, \ldots, m)$ — матрицы того же порядка, но переменные и с комплексными элементами.

Теория аналитических функций от нескольких матриц была развита И. А. Лаппо-Данилевским. На основе этой теории И. А. Лаппо-Данилевский провел фундаментальные исследования систем линейных дифференциальных уравнений с рациональными коэффициентами.

Система с рациональными коэффициентами путем надлежащего преобразования независимой переменной всегда может быть приведена к виду

$$\frac{dV}{dz} = \sum_{j=1}^{m} \left\{ \frac{U_{j0}}{(z-a_j)^{s_j}} + \frac{U_{j1}}{(z-a_j)^{s_j-1}} + \ldots + \frac{U_{j, s_j-1}}{z-a_j} \right\} X, \qquad (158)$$

где U_{jk} — постоянные матрицы n-го порядка, a_j — комплексные числа, s_j — целые положительные числа $(k = 0, 1, \ldots, s_j - 1; \ j = 1, 2, \ldots, m)$ [1].

Некоторые результаты Лаппо-Данилевского мы проиллюстрируем на частном случае так называемых *регулярных* систем. Последние характеризуются условием $s_1 = s_2 = \ldots = s_m = 1$ и записываются в виде

$$\frac{dX}{dz} = \sum_{j=1}^{m} \frac{U_j}{z-a_j} X. \qquad (159)$$

Следуя Лаппо-Данилевскому, введем в рассмотрение специальные аналитические функции — гиперлогарифмы, — определяемые следующими рекуррентными соотношениями:

$$l_b(z; a_{j_1}) = \int_b^z \frac{dz}{z-a_{j_1}},$$

$$l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int_b^z \frac{l_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z-a_{j_1}} dz.$$

[1] В системе (158) все коэффициенты — правильные рациональные дроби относительно z. К такому виду приводятся любые рациональные коэффициенты, если при помощи дробно-линейного преобразования над переменной z перевести регулярную (для всех коэффициентов) конечную точку $z = c$ в точку $z = \infty$.

Рассматривая точки $a_1, a_2, \ldots, a_m, \infty$ как точки разветвления логарифмического типа, построим соответствующую риманову поверхность $S(a_1, a_2, \ldots, a_m; \infty)$. Каждый гиперлогарифм будет однозначной функцией на этой поверхности. С другой стороны, матрицант системы (159) Ω_b^z (т. е. нормированное в точке $z=b$ решение), будучи аналитически продолжен, также может быть рассматриваем как однозначная функция на $S(a_1, a_2, \ldots, a_m; \infty)$; при этом в качестве b может быть выбрана любая конечная точка на S, отличная от $a_1, a_2, \ldots, a_m$.

Для нормированного решения Ω_b^z Лаппо-Данилевский дает явное выражение через определяющие матрицы $U_1, U_2, \ldots, U_m$ системы (159) в виде ряда

$$\Omega_b^z = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1 \ldots m)} l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) U_{j_1} U_{j_2} \ldots U_{j_\nu}. \qquad (160)$$

Это разложение сходится равномерно относительно z при любых $U_1, U_2, \ldots, U_m$ и представляет Ω_b^z в любой конечной области на поверхности $S(a_1, a_2, \ldots, a_m; \infty)$, если только эта область не содержит внутри и на границе точек $a_1, \ldots, a_m$.

Если ряд (156) сходится при любых матрицах $X_1, X_2, \ldots, X_m$, то соответствующая функция $F(X_1, X_2, \ldots, X_m)$ называется *целой*. Ω_b^z представляет собой целую функцию от матриц $U_1, U_2, \ldots, U_m$.

Заставляя в формуле (160) аргумент z обойти точку a_j в положительном направлении один раз так, чтобы контур обхода не захватывал других точек a_i (при $i \neq j$), мы получим выражение для *интегральной подстановки* V_j, соответствующей точке $z=a_j$:

$$V_j = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1 \ldots m)} p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) U_{j_1} U_{j_2} \ldots U_{j_\nu} \qquad (161)$$

$$(j = 1, 2, \ldots, m),$$

где в понятных обозначениях

$$p_j(b: a_{j_1}) = \int_{(a_j)} \frac{\partial z}{z - a_{j_1}},$$

$$p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int_{(a_j)} \frac{L_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z - a_{j_1}} dz$$

$$\left(\begin{matrix} j_1, j_2, \ldots, j_\nu, \; j = 1, 2, \ldots, m; \\ \nu = 1, 2, 3, \ldots \end{matrix} \right).$$

Ряд (161), как и ряд (160), представляет собой целую функцию от $U_1, U_2, \ldots, U_m$.

Обобщив теорию аналитических функций на случай бесконечного, но счетного множества матриц-аргументов $X_1, X_2, X_3, \ldots$,[1]). Лаппо-Данилевский использовал эту теорию для исследования поведения решения системы в окрестности иррегулярной особой точки[2]). Мы приведем основной результат.

---

[1]) См. [21б], т. I, Мемуар 1.

[2]) См. [21б], т. I, Мемуар 3, см. также [92], [104а, б].

Нормированное решение Ω_b^z системы

$$\frac{dX}{dz} = \sum_{j=-q}^{+\infty} P_j z^j X,$$

где степенной ряд в правой части сходится при $|z| < r\,(r > 1)$ [1]), может быть представлено рядом

$$\Omega_b^z = E + \sum_{\nu=1}^{\infty} \sum_{j_1, j_2, \ldots, j_\nu = -q}^{\infty} P_{j_1} \ldots$$

$$\ldots P_{j_\nu} \sum_{\mu=0}^{\nu} b^{j_{\mu+1}+\ldots+j_\nu+\nu-\mu} z^{j_1+\ldots+j_\mu+\mu} \sum_{\lambda=0}^{n-\mu} \alpha^{*(\lambda)}_{j_{\mu+1},\ldots,j_\nu} \ln^\lambda b \sum_{\varkappa=0}^{\mu} \alpha^{(\varkappa)}_{j_1,\ldots,j_\mu} \ln^\varkappa z. \tag{162}$$

Здесь $\alpha^{*(\lambda)}_{j_{\mu+1},\ldots,j_\nu}$ и $\alpha^{(\varkappa)}_{j_1,\ldots,j_\mu}$ — скалярные коэффициенты, определяемые по специальным формулам. Ряд (162) сходится при любых матрицах P_1, P_2, ... в кольце

$$\varrho < |z| < r$$

(ϱ — произвольное положительное число, меньшее r). Этому кольцу должна принадлежать и точка $b\,(\varrho < |b| < r)$.

Не имея возможности в какой бы то ни было степени подробно изложить содержание работ Лаппо-Данилевского в настоящей книге, мы вынуждены ограничиться приведенными выше формулировками некоторых основных результатов и отослать читателей к соответствующей литературе.

Все относящиеся к дифференциальным уравнениям работы Лаппо-Данилевского изданы посмертно Академией наук СССР в трех томах в 1934—1936 гг. Кроме того, основные результаты автора изложены в статьях [82] и небольшой книге [18а]. Сокращенное изложение некоторых результатов можно найти и в книге В. И. Смирнова «Курс высшей математики», т. III.

[1]) Ограничение $r > 1$ несущественно, так как это условие всегда можно получить заменой z на αz, где α — надлежащим образом выбранное положительное число.

ГЛАВА XVI

ПРОБЛЕМА РАУСА — ГУРВИЦА И СМЕЖНЫЕ ВОПРОСЫ

§ 1. Введение

В главе XII, § 2 мы выяснили, что согласно теореме Ляпунова нулевое решение системы дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik}x_k + (**) \qquad (i = 1, \ldots, n) \tag{1}$$

[a_{ik} ($i, k = 1, 2, \ldots, n$) — постоянные коэффициенты] при любых членах ($**$) второго порядка и выше относительно $x_1, x_2, \ldots, x_n$ является устойчивым, если все характеристические числа матрицы $A = \| a_{ik} \|_1^n$, т. е. все корни векового уравнения $\Delta(\lambda) \equiv |\lambda E - A| = 0$, имеют отрицательные вещественные части.

Поэтому задача установления необходимых и достаточных условий, при которых все корни данного алгебраического уравнения расположены в левой полуплоскости, имеет фундаментальное значение в ряде прикладных областей, в которых исследуется устойчивость механических и электрических систем.

Важность этой алгебраической задачи была ясна основоположникам теории регулирования машин, английскому физику Д. К. Максвеллу и русскому инженеру-исследователю И. А. Вышнеградскому, которые в своих работах, посвященных регуляторам, установили и широко использовали упомянутые алгебраические условия для уравнений не выше третьей степени [1]).

В 1868 г. Максвелл выдвинул математическую задачу об отыскании соответствующих условий для алгебраического уравнения любой степени. Между тем эта задача по существу была решена в опубликованной в 1856 г. работе французского математика Эрмита [128]. В этой работе была установлена тесная связь между числом корней комплексного многочлена $f(z)$, расположенных внутри какой-либо полуплоскости (или даже внутри какого-либо прямоугольника), и сигнатурой некоторой квадратичной формы. Однако результаты Эрмита не были доведены до такого состояния, чтобы они могли быть использованы специалистами, работающими в прикладных областях. Поэтому эта работа Эрмита и не получила соответствующего распространения.

---

[1]) Д. К. Максвелл, О регуляторах (1868); И. А. Вышнеградский, О регуляторах прямого действия (1876). Эти работы опубликованы в сборнике «Теория автоматического регулирования» (издание АН СССР, 1949). См. там же статью А. А. Андронова и И. Н. Вознесенского, О работах Д. К. Максвелла, И. А. Вышнеградского и А. Стодолы в области теории регулирования машин.

В 1875 г. английский механик Раус [45], пользуясь теоремой Штурма и теорией индексов Коши, установил алгоритм для определения числа k корней вещественного многочлена, расположенных в правой полуплоскости ($\mathrm{Re}\, z > 0$). В частном случае $k = 0$ этот алгоритм и дает критерий устойчивости.

В конце XIX в. крупнейший словацкий инженер-исследователь, создатель теории паровых и газовых турбин, А. Стодола, не зная работы Рауса, снова поставил задачу об отыскании условий того, чтобы все корни алгебраического уравнения имели отрицательные вещественные части, и в 1895 г. А. Гурвиц [129], опираясь на работы Эрмита, дает второе (независимое от Рауса) решение той же задачи. Полученные Гурвицем детерминантные неравенства известны в настоящее время под названием условий Рауса — Гурвица.

Однако еще до появления в свет работы Гурвица основатель современной теории устойчивости А. М. Ляпунов в своей знаменитой диссертации («Общая задача об устойчивости движения», Харьков, 1892) установил [1]) теорему, из которой вытекают необходимые и достаточные условия для того, чтобы все корни характеристического уравнения вещественной матрицы $A = \| a_{ik} \|_1^n$ имели отрицательные вещественные части. Эти условия используются в ряде работ по теории регулирования [2]).

Новый критерий устойчивости был установлен в 1914 г. французскими математиками Льенаром и Шипаром [135].

Используя специальные квадратичные формы, эти авторы получили критерий устойчивости, имеющий некоторые преимущества перед критерием Рауса — Гурвица (число детерминантных неравенств в критерии Льенара — Шипара примерно вдвое меньше, нежели в критерии Рауса — Гурвица).

Знаменитые русские математики П. Л. Чебышев и А. А. Марков установили две замечательные теоремы в связи с разложением в ряды непрерывных дробей специального типа. Эти теоремы, как будет показано в § 16, имеют непосредственное отношение к проблеме Рауса — Гурвица.

В очерченном круге вопросов, как увидит читатель, находят себе существенное применение теория квадратичных форм (гл. X) и, в частности, теория ганкелевых форм (гл. X, § 10).

§ 2. Индексы Коши

Начнем с рассмотрения так называемых индексов Коши [3]).

О п р е д е л е н и е 1. *Индексом Коши* вещественной рациональной функции $R(x)$ в пределах от a до b (обозначение $I_a^b R(x)$; a, b — вещественные числа, либо $\pm\infty$) будем называть разность между числом разрывов $R(x)$ с переходом от $-\infty$ к $+\infty$ и числом разрывов с переходом от $+\infty$ к $-\infty$ при изменении аргумента от a к b [4]).

Согласно этому определению, если

$$R(x) = \sum_{i=1}^{p} \frac{A_i}{x - a_i} + R_1(x),$$

[1]) См. [22], § 20.
[2]) См., например, [73].
[3]) См. [10], стр. 419—425.
[4]) При подсчете числа разрывов крайние значения x — пределы a и b — не включаются.

где A_i, α_i $(i = 1, 2, \ldots, p)$ — вещественные числа, а $R_1(x)$ — рациональная функция, не имеющая вещественных полюсов [1]), то

$$I_{-\infty}^{+\infty} R(x) = \sum_{i=1}^{p} \operatorname{sign} A_i \text{ [2]} \tag{2}$$

и вообще

$$I_a^b R(x) = \sum_{a<\alpha_i<b} \operatorname{sign} A_i \quad (a<b). \tag{2'}$$

В частности, если $f(x) = a_0(x-\alpha_1)^{n_1} \ldots (x-\alpha_m)^{n_m}$ — вещественный многочлен ($\alpha_i \neq \alpha_k$ при $i \neq k$; $i, k = 1, 2, \ldots, m$) и среди корней $\alpha_1, \alpha_2, \ldots, \alpha_m$ этого многочлена только первые p вещественны, то

$$\frac{f'(x)}{f(x)} = \sum_{j=1}^{m} \frac{n_j}{x-\alpha_j} = \sum_{i=1}^{p} \frac{n_i}{x-\alpha_i} + R_1(x),$$

где $R_1(x)$ — вещественная рациональная функция, не имеющая вещественных полюсов.

Поэтому *индекс*

$$I_a^b \frac{f'(x)}{f(x)} \quad (a<b)$$

равен числу различных вещественных корней многочлена $f(x)$, *находящихся внутри интервала* (a, b).

Произвольная вещественная рациональная функция $R(x)$ всегда представима в виде

$$R(x) = \sum_{i=1}^{p} \left\{ \frac{A_1^{(i)}}{x-\alpha_i} + \ldots + \frac{A_{n_i}^{(i)}}{(x-\alpha_i)^{n_i}} \right\} + R_1(x),$$

где все α и A — вещественные числа ($A_{n_i}^{(i)} \neq 0$; $i = 1, 2, \ldots, p$) и $R_1(x)$ не имеет вещественных полюсов.

Тогда

$$I_{-\infty}^{+\infty} R(x) = \sum_{(n_i \text{ нечетно})} \operatorname{sign} A_{n_i}^{(i)} \tag{3}$$

и вообще

$$I_a^b R(x) = \sum_{\substack{a<\alpha_i<b \\ n_i \text{ нечетно}}} \operatorname{sign} A_{n_i}^{(i)} \quad (a<b) \text{ [3]}. \tag{3'}$$

Если $R(a) = R(b) = 0$, то индекс $I_a^b R(x)$ выражается через приращение непрерывной функции $\operatorname{arctg} R(x)$:

$$I_a^b R(x) = -\Delta_a^b \operatorname{arctg} R(x) \quad (a<b) \text{ [4]}. \tag{4}$$

---

[1]) Полюсами рациональной функции являются те значения аргумента, при которых эта функция обращается в бесконечность.

[2]) Под sign a (a — вещественное число) мы понимаем $+1$, -1 или 0 в зависимости от того, $a > 0$, $a < 0$ или $a = 0$.

[3]) В (3) сумма распространяется на все те значения i, для которых соответствующее n_i нечетно. В (3') сумма распространяется на все те i, для которых n_i нечетно и $a < \alpha_i < b$.

[4]) Если $a = -\infty$, а $b = +\infty$, то формула (4) справедлива для любой правильной рациональной дроби $R(x)$, поскольку в этом случае $R(-\infty) = R(+\infty) = 0$.

Один из методов вычисления индекса $I_a^b R(x)$ основан на классической теореме Штурма.

Рассмотрим ряд вещественных многочленов

$$f_1(x),\ f_2(x),\ \ldots,\ f_m(x), \tag{5}$$

обладающий двумя свойствами по отношению к интервалу $(a,\ b)$ [1]:

1° При любом значении $x\ (a<x<b)$, обращающем в нуль какую-либо из функций $f_k(x)$, две смежные функции $f_{k-1}(x)$ и $f_{k+1}(x)$ имеют значения, отличные от нуля и разных знаков, т. е. из $f_k(x)=0$ при $a<x<b$ следует: $f_{k-1}(x)\,f_{k+1}(x)<0$.

2° Последняя функция $f_m(x)$ в ряду (5) не обращается в нуль внутри $(a,\ b)$, т. е. $f_m(x)\neq 0$ при $a<x<b$.

Такой ряд (4) многочленов называется *рядом Штурма в интервале* $(a,\ b)$.

Обозначим через $V(x)$ число перемен знака в ряду (5) при фиксированном значении x [2]. Тогда при изменении x от a до b величина $V(x)$ может измениться лишь при переходе через нуль какой-либо из функций ряда (5). Но в силу 1° при переходе через нуль функции $f_k(x)$ $(k=2,\ \ldots\ \ldots,\ m-1)$ величина $V(x)$ не изменяется. При переходе же через нуль функции $f_1(x)$ теряется или приобретается одна перемена знака в ряду (5) в зависимости от того, переходит при этом отношение $\frac{f_2(x)}{f_1(x)}$ от $-\infty$ к $+\infty$ или наоборот. Поэтому имеет место

Теорема 1 (Штурма). *Если* $f_1(x),\ f_2(x),\ \ldots,\ f_m(x)$ — *ряд Штурма в* $(a,\ b)$, а $V(x)$ — *число перемен знака в этом ряду, то*

$$I_a^b \frac{f_2(x)}{f_1(x)} = V(a) - V(b). \tag{6}$$

Примечание. Помножим все члены ряда Штурма на один и тот же произвольный многочлен $d(x)$. Полученный таким образом ряд многочленов назовем *обобщенным рядом Штурма*. Так как умножение всех членов ряда (5) на один и тот же многочлен не меняет ни левой, ни правой части равенства (6), то теорема Штурма сохраняет свою силу и для обобщенного ряда Штурма.

Заметим, что если даны два произвольных многочлена $f(x)$ и $g(x)$ [степень $f(x)\geqslant$ степени $g(x)$], то всегда можно при помощи алгоритма Евклида построить обобщенный ряд Штурма, который начинался бы с функций $f_1(x)\equiv f(x)$, $f_2(x)\equiv g(x)$.

Действительно, обозначая через $-f_3(x)$ остаток от деления $f_1(x)$ на $f_2(x)$, через $-f_4(x)$ — остаток от деления $f_2(x)$ на $f_3(x)$ и т. д., будем иметь цепочку тождеств

$$\begin{aligned} f_1(x) &= q_1(x)\,f_2(x) - f_3(x),\ \ldots,\ f_{k-1}(x) = \\ &= q_{k-1}(x)\,f_k(x) - f_{k+1}(x),\ \ldots,\ f_{m-1}(x) = q_{m-1}(x)\,f_m(x), \end{aligned} \tag{7}$$

где последний не равный тождественно нулю остаток $f_m(x)$ является наибольшим общим делителем $f(x)$ и $g(x)$, а также наибольшим общим

[1] При этом a может равняться $-\infty$, а b $+\infty$.

[2] Если $a<x<b$ и $f_1(x)\neq 0$, то в силу 1° при определении $V(x)$ нулевые значения в ряду (4) можно опустить либо этим значениям можно приписать произвольные знаки. Если a конечно, то под $V(a)$ следует понимать $V(a+\varepsilon)$, где ε — столь малое положительное число, что в полузамкнутом интервале $(a,\ a+\varepsilon]$ ни одна из функций $f_i(x)$ $(i=1,\ 2,\ \ldots,\ m)$ не обращается в нуль. Точно так же, если b конечно, то под $V(b)$ следует понимать $V(b-\varepsilon)$, где число ε определяется аналогично.

делителем всех функций построенного таким образом ряда (5). Если $f_m(x) \neq 0$ $(a < x < b)$, то полученный ряд (5) в силу (7) удовлетворяет условиям 1°, 2° и является рядом Штурма. Если же многочлен $f_m(x)$ имеет корни внутри интервала (a, b), то ряд (5) является обобщенным рядом Штурма, поскольку он становится рядом Штурма после деления всех его членов на $f_m(x)$.

Из сказанного следует, что индекс любой рациональной функции $R(x)$ может быть определен при помощи теоремы Штурма. Для этого достаточно представить $R(x)$ в виде $Q(x) + \frac{g(x)}{f(x)}$, где $Q(x)$, $f(x)$, $g(x)$ — многочлены и степень $g(x) \leqslant$ степени $f(x)$. Тогда, если построить обобщенный ряд Штурма для $f(x)$, $g(x)$, то

$$I_a^b R(x) = I_a^b \frac{g(x)}{f(x)} = V(a) - V(b).$$

При помощи теоремы Штурма можно определить число различных вещественных корней многочлена $f(x)$ внутри интервала (a, b), поскольку это число, как мы видели, равно $I_a^b \frac{f'(x)}{f(x)}$.

§ 3. Алгоритм Рауса

1. Задача Рауса состоит в определении числа k корней вещественного многочлена $f(z)$, расположенных в правой полуплоскости $(\operatorname{Re} z > 0)$.

Рассмотрим сначала случай, когда $f(z)$ не имеет нулей на мнимой оси. В правой полуплоскости построим полуокружность радиуса R с центром в нуле и рассмотрим область, ограниченную этой полуокружностью и отрезком мнимой оси (рис. 9). При достаточно большом R все k нулей многочлена $f(z)$ с положительными вещественными частями будут находиться внутри этой области. Поэтому $\arg f(z)$ при положительном обходе контура области получит приращение $2k\pi$ [1]). С другой стороны, приращение $\arg f(z)$ вдоль полуокружности радиуса R при $R \to \infty$ определяется приращением аргумента старшего члена $a_0 z^n$ и потому равно $n\pi$. Поэтому для приращения $\arg f(z)$ вдоль мнимой оси $(R \to \infty)$ получаем выражение

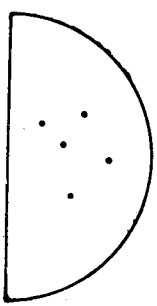


Рис. 9.

$$\Delta_{-\infty}^{+\infty} \arg f(i\omega) = (n - 2k)\pi. \tag{8}$$

Введем не совсем обычные обозначения для коэффициентов многочлена $f(z)$, пусть

$$f(z) = a_0 z^n + b_0 z^{n-1} + a_1 z^{n-2} + b_1 z^{n-3} + \ldots \quad (a_0 \neq 0).$$

Тогда, замечая, что приращение $\Delta \arg f(i\omega)$ в формуле (8) не изменится, если многочлен $f(z)$ помножить на произвольное комплексное число, положим:

$$\frac{1}{i^n} f(i\omega) = f_1(\omega) - i f_2(\omega), \tag{9}$$

---

[1]) В самом деле, если $f(z) = a_0 \prod_{i=1}^{n} (z - z_i)$, то $\Delta \arg f(z) = \sum_{i=1}^{n} \Delta \arg (z - z_i)$. Если точка z_i находится внутри рассматриваемой области, то $\Delta \arg (z - z_i) = 2\pi$; если z_i вне этой области, то $\Delta \arg (z - z_i) = 0$.

где

$$\left.\begin{aligned} f_1(\omega) &= a_0\omega^n - a_1\omega^{n-2} + a_3\omega^{n-4} - \dots, \\ f_2(\omega) &= b_0\omega^{n-1} - b_1\omega^{n-3} + b_3\omega^{n-5} - \dots \end{aligned}\right\} \tag{10}$$

Следуя Раусу, воспользуемся индексом Коши. Из формул (4) и (9) находим:

$$\frac{1}{\pi}\Delta_{-\infty}^{+\infty}\arg f(i\omega) = -\frac{1}{\pi}\Delta_{-\infty}^{+\infty}\operatorname{arctg}\frac{f_2(\omega)}{f_1(\omega)} = I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}.$$

Поэтому из формулы (8) следует, что[1])

$$I_{-\infty}^{+\infty}\frac{b_0\omega^{n-1} - b_1\omega^{n-3} + \dots}{a_0\omega^n - a_1\omega^{n-2} + \dots} = n - 2k. \tag{11}$$

2. Для определения индекса, стоящего в левой части равенства (11), используем теорему Штурма (см. предыдущий параграф). Исходя из функций $f_1(\omega)$ и $f_2(\omega)$, определяемых равенствами (10), построим, следуя Раусу, при помощи алгоритма Евклида обобщенный ряд Штурма (см. стр. 471)

$$f_1(\omega),\ f_2(\omega),\ f_3(\omega),\ \dots,\ f_m(\omega). \tag{12}$$

Рассмотрим теперь регулярный случай: $m = n+1$. В этом случае в ряду (12) степень каждой функции на единицу меньше степени предыдущей, и последняя функция $f_m(\omega)$ имеет нулевую степень[2]).

Из алгоритма Евклида [см. (7)] следует, что

$$f_3(\omega) = \frac{a_0}{b_0}\omega f_2(\omega) - f_1(\omega) = c_0\omega^{n-2} - c_1\omega^{n-4} + c_2\omega^{n-6} - \dots,$$

где

$$c_0 = a_1 - \frac{a_0}{b_0}b_1 = \frac{b_0a_1 - a_0b_1}{b_0},\quad c_1 = a_2 - \frac{a_0}{b_0}b_2 = \frac{b_0a_2 - a_0b_2}{b_0},\ \dots \tag{13}$$

Точно так же

$$f_4(\omega) = \frac{b_0}{c_0}\omega f_3(\omega) - f_2(\omega) = d_0\omega^{n-3} - d_1\omega^{n-5} + \dots,$$

где

$$d_0 = b_1 - \frac{b_0}{c_0}c_1 = \frac{c_0b_1 - b_0c_1}{c_0},\quad d_1 = b_2 - \frac{b_0}{c_0}c_2 = \frac{c_0b_2 - b_0c_2}{c_0},\ \dots \tag{13'}$$

Аналогично определяются коэффициенты остальных многочленов $f_5(\omega),\ \dots,\ f_{n+1}(\omega)$.

При этом каждый из многочленов

$$f_1(\omega),\ f_2(\omega),\ \dots,\ f_{n+1}(\omega) \tag{14}$$

является четной или нечетной функцией, причем степени смежных многочленов всегда имеют разную четность.

---

[1]) Напомним, что формула (10) выведена в предположении, что многочлен $f(z)$ не имеет корней на мнимой оси.

[2]) В регулярном случае ряд (12) является обычным (необобщенным) рядом Штурма.

Составим схему Рауса:

$$\left.\begin{array}{llll} a_0, & a_1, & a_2, & \ldots, \\ b_0, & b_1, & b_2, & \ldots, \\ c_0, & c_1, & c_2, & \ldots, \\ d_0, & d_1, & d_2, & \ldots, \\ \cdots & \cdots & \cdots & \end{array}\right\} \tag{15}$$

В этой схеме, как показывают формулы (13), (13′), каждая строка определяется из двух предыдущих по следующему правилу:

Из чисел верхней строки вычитаются соответствующие числа нижней, предварительно помноженные на такое число, чтобы первая разность обращалась в нуль. Отбрасывая эту нулевую разность, получаем искомую строку.

Регулярный случай, очевидно, характеризуется тем, что при последовательном применении этого правила мы в ряду

$$b_0,\ c_0,\ d_0,\ \ldots$$

не встречаем числа, равного нулю, и этот ряд состоит из $n+1$ чисел.

На рис. 10 и 11 показан скелет регулярной схемы Рауса при n четном ($n=6$) и n нечетном ($n=7$). Здесь элементы схемы отмечены точками.

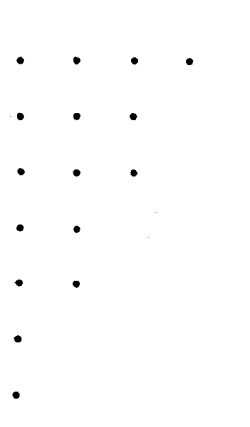


Рис. 10.

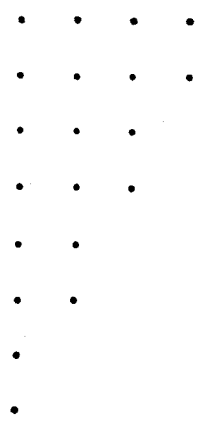


Рис. 11.

В регулярном случае многочлены $f_1(\omega)$ и $f_2(\omega)$ имеют наибольший общий делитель $f_{n+1}(\omega)=\text{const}\neq 0$. Поэтому эти многочлены не обращаются одновременно в нуль, т. е. $f(i\omega)=f_1(\omega)-if_2(\omega)\neq 0$ при ω вещественном. Поэтому *в регулярном случае имеет место формула* (11).

Применяя к левой части этой формулы теорему Штурма в интервале $(-\infty,\ +\infty)$ и используя при этом ряд (14), получаем согласно (11):

$$V(-\infty)-V(+\infty)=n-2k. \tag{16}$$

В данном случае[1])

$$V(+\infty)=V(a_0,\ b_0,\ c_0,\ d_0,\ \ldots),$$

а

$$V(-\infty)=V(a_0,\ -b_0,\ c_0,\ -d_0,\ \ldots).$$

Отсюда

$$V(-\infty)+V(+\infty)=n. \tag{17}$$

Из равенств (16) и (17) находим:

$$k=V(a_0,\ b_0,\ c_0,\ d_0,\ \ldots). \tag{18}$$

Нами доказана для регулярного случая

---

[1]) Знак $f_k(\omega)$ при $\omega=+\infty$ совпадает со знаком старшего коэффициента, а при $\omega=-\infty$ отличается от этого знака множителем $(-1)^{n-k+1}$ ($k=1,\ 2,\ \ldots,\ n+1$).

Теорема 2 (Рауса). *Число корней вещественного многочлена $f(z)$, лежащих в правой полуплоскости $\operatorname{Re} z>0$, равно числу перемен знака в первом столбце схемы Рауса.*

3. Рассмотрим важный частный случай, когда все корни $f(z)$ имеют отрицательные вещественные части («случай устойчивости»). В этом случае многочлен $f(z)$ не имеет чисто мнимых корней, и потому имеет место формула (11), а следовательно, и формула (16). Поскольку $k=0$, формула (16) перепишется так:

$$V(-\infty)-V(+\infty)=n. \tag{19}$$

Но $0\leqslant V(-\infty)\leqslant m-1\leqslant n$ и $0\leqslant V(+\infty)\leqslant m-1\leqslant n$. Поэтому равенство (19) возможно лишь тогда, когда $m=n+1$ (регулярный случай!) и $V(+\infty)=0$, $V(-\infty)=m-1=n$. Тогда из формулы (18) следует:

Критерий Рауса. *Для того чтобы все корни вещественного многочлена $f(z)$ имели отрицательные вещественные части, необходимо и достаточно, чтобы при выполнении алгоритма Рауса все элементы первого столбца схемы Рауса получались отличными от нуля и одного знака.*

4. При установлении теоремы Рауса мы опирались на формулу (11). В дальнейшем нам понадобится обобщение этой формулы. Формула (11) была выведена в предположении, что многочлен $f(z)$ не имеет корней на мнимой оси. Мы покажем, что в общем случае, когда многочлен $f(z)=a_0z^n+b_0z^{n-1}+a_1z^{n-2}+\ldots$ $(a_0\neq 0)$ имеет k корней в правой полуплоскости и s корней на мнимой оси, формула (11) заменяется формулой

$$I_{-\infty}^{+\infty}\frac{b_0\omega^{n-1}-b_1\omega^{n-3}+b_2\omega^{n-5}-\ldots}{a_0\omega^n-a_1\omega^{n-2}+a_2\omega^{n-4}-\ldots}=n-2k-s. \tag{20}$$

В самом деле,

$$f(z)=d(z)\,f^*(z),$$

где вещественный многочлен $d(z)=z^s+\ldots$ имеет s корней на мнимой оси, а многочлен $f^*(z)$ степени $n^*=n-s$ таких корней не имеет. Пусть

$$\frac{1}{i^n}f(i\omega)=f_1(\omega)-if_2(\omega),\qquad \frac{1}{i^{n-s}}f^*(i\omega)=f_1^*(i\omega)-if_2^*(i\omega).$$

Тогда

$$f_1(\omega)-if_2(\omega)=\frac{1}{i^s}\,d(i\omega)\,[f_1^*(i\omega)-if_2^*(i\omega)].$$

Поскольку $\frac{1}{i^s}d(i\omega)$ — вещественный многочлен относительно ω, то

$$\frac{f_2(\omega)}{f_1(\omega)}=\frac{f_2^*(\omega)}{f_1^*(\omega)}.$$

К многочлену $f^*(z)$ применима формула (11). Поэтому

$$I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}=I_{-\infty}^{+\infty}\frac{f_2^*(\omega)}{f_1^*(\omega)}=n^*-2k=n-2k-s,$$

что и требовалось доказать[1]).

---

[1]) Обращаем внимание читателя на интересное обобщение критерия Рауса, содержащееся в работе Фаэдо [176]. Здесь устанавливаются достаточные условия того, что корни всех многочленов $f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n$ с коэффициентами a_i, меняющимися в заданных интервалах

$$\underline{a}_i\leqslant a_i\leqslant\bar{a}_i\qquad (i=0,\ 1,\ \ldots,\ n;\quad \underline{a}_0<0),$$

одновременно все имеют отрицательные вещественные части.

§ 4. Особые случаи. Примеры

1. В предыдущем параграфе мы разобрали регулярный случай, когда при заполнении схемы Рауса ни одно из чисел b_0, c_0, d_0, ... не оказывается равным нулю.

Переходим теперь к рассмотрению особых случаев, когда в ряду чисел b_0, c_0, ... мы встречаемся с числом $h_0=0$. Алгоритм Рауса останавливается на той строке, где находится h_0, так как для получения чисел следующей строки нужно делить на h_0.

Особые случаи могут быть двух типов:

1) В той же строке, где находится h_0, имеются числа, не равные нулю. Это означает, что в каком-то месте ряда (12) произошло понижение степени больше чем на единицу.

2) Одновременно все числа строки, содержащей h_0, оказываются равными нулю. Тогда эта строка является $(m+1)$-й, где m — число членов в обобщенном ряду Штурма (5). В этом случае в ряду (12) степени функций все время понижаются на единицу, но степень последней функции $f_m(\omega)$ больше нуля. В обоих случаях в ряду (12) число функций $m<n+1$.

Поскольку обычный алгоритм Рауса в особых случаях приостанавливается, Раус дает специальные правила для продолжения схемы в случаях 1), 2).

2. В случае 1) следует по Раусу вместо $h_0=0$ подставить «малую» величину ε определенного (но произвольного) знака и продолжать заполнение схемы. При этом последующие элементы первого столбца схемы будут рациональными функциями от ε. Знаки этих элементов определятся, исходя из «малости» и знака ε. Если же какой-либо из этих элементов окажется тождественным нулем относительно ε, то мы этот элемент заменим другой малой величиной η и продолжим алгоритм.

Пример.

$$f(z)=z^4+z^3+2z^2+2z+1.$$

Схема Рауса (с малым параметром ε)

$$\begin{array}{llll} 1, & 2, & 1 & \\ 1, & 2 & & \\ \varepsilon, & 1 & & \\ 2-\dfrac{1}{\varepsilon} & & & \\ 1 & & & \end{array} \qquad k=V\left(1,\ 1,\ \varepsilon,\ 2-\frac{1}{\varepsilon},\ 1\right)=2.$$

Обоснование этого своеобразного метода варьирования элементов схемы заключается в следующем:

Поскольку мы предполагаем отсутствие особенностей второго типа, то функции $f_1(\omega)$ и $f_2(\omega)$ взаимно просты. Отсюда следует, что многочлен $f(z)$ не имеет корней на мнимой оси.

В схеме Рауса все элементы выражаются рационально через элементы первых двух строк, т. е. через коэффициенты данного многочлена. Но нетрудно усмотреть из формул (13), (13′) и аналогичных формул для последующих строк, что, задавшись произвольными значениями для элементов двух любых подряд идущих строк схемы Рауса и для первых элементов предыдущих строк, мы можем целым рациональным образом выразить через эти элементы все числа, стоящие в первых двух строках, т. е. коэффициенты исходного многочлена. Так, например, все числа a, b можно представить в виде целых рациональных функций от

$$a_0,\ b_0,\ c_0,\ \ldots,\ h_0,\ h_1,\ h_2,\ \ldots,\ g_0,\ g_1,\ g_2,\ \ldots$$

Поэтому, заменяя $h_0=0$ на ε, мы фактически видоизменяем наш исходный многочлен. Вместо схемы для $f(z)$ мы имеем схему Рауса для многочлена $F(z, \varepsilon)$, где $F(z, \varepsilon)$ — целая рациональная функция от z и ε, обращающаяся в $f(z)$ при $\varepsilon=0$. Так как корни многочлена $F(z, \varepsilon)$ непрерывно меняются с изменением параметра ε и при $\varepsilon=0$ нет корней на мнимой оси, то при малых по модулю значениях ε число k корней в правой полуплоскости у многочленов $F(z, \varepsilon)$ и $F(z, 0)=f(z)$ одинаково.

3. Переходим к рассмотрению особенностей второго типа. Пусть в схеме Рауса

$$a_0\neq 0,\ b_0\neq 0,\ \ldots,\ e_0\neq 0,\ h_0=0,\ h_1=0,\ h_2=0,\ \ldots$$

В этом случае в обобщенном ряду Штурма (16) последний многочлен имеет вид:

$$f_m(\omega)=e_0\omega^{n-m+1}-e_1\omega^{n-m-1}+\dots$$

Раус предлагает заменить нулевое $f_{m+1}(\omega)$ на $f'_m(\omega)$, т. е. вместо нулевых h_0, h_1, ... написать соответственно коэффициенты

$$(n-m+1)\,e_0,\quad (n-m-1)\,e_1,\ \dots$$

и продолжать алгоритм.

Обоснование этого правила заключается в следующем:

Согласно формуле (20)

$$I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}=n-2k-s$$

s корней многочлена $f(z)$ на мнимой оси совпадают с вещественными корнями многочлена $f_m(\omega)$. Поэтому, если эти вещественные корни простые, то (см. стр. 470)

$$I_{-\infty}^{+\infty}\frac{f'_m(\omega)}{f_m(\omega)}=s,$$

и, следовательно,

$$I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}+I_{-\infty}^{+\infty}\frac{f'_m(\omega)}{f_m(\omega)}=n-2k.$$

Эта формула показывает, что недостающую часть схемы Рауса следует заполнить схемой Рауса для многочленов $f_m(\omega)$ и $f'_m(\omega)$. Коэффициенты многочлена $f'_m(\omega)$ и используются для замены элементов нулевой строки в схеме Рауса.

Если же корни $f_m(\omega)$ не простые, то, обозначая через $d(\omega)$ наибольший общий делитель $f_m(\omega)$ и $f'_m(\omega)$, через $e(\omega)$ наибольший общий делитель $d(\omega)$ и $d'(\omega)$ и т. д., мы будем иметь:

$$I_{-\infty}^{+\infty}\frac{f'(\omega)}{f(\omega)}+I_{-\infty}^{+\infty}\frac{d'(\omega)}{d(\omega)}+I_{-\infty}^{+\infty}\frac{e'(\omega)}{e(\omega)}+\dots=s.$$

Таким образом, искомое число k можно получить, если недостающую часть схемы Рауса дополнить схемами Рауса для $f_m(\omega)$ и $f'_m(\omega)$, $d(\omega)$ и $d'(\omega)$, $e(\omega)$ и $e'(\omega)$ и т. д., т. е. несколько раз применять правило Рауса для ликвидации особенностей 2-го типа.

Пример.

$$f(z)=z^{10}+z^9-z^8-2z^7+z^6+3z^5+z^4-2z^3-z^2+z+1.$$

Схема

ω^{10}	1	−1	1	1	−1	1
ω^9	1	−2	3	−2	1	
ω^8	1	−2	3	−2	1	
ω^7	8	−12	12	−4		
	2	−3	3	−1		
ω^6	−1	3	−3	2		
ω^5	3	−3	3			
	1	−1	1			
ω^4	2	−2	2			
	1	−1	1			
ω^3	4	−2				
	2	−1				
ω^2	−1	2				
ω	1					
ω^0	2					
	1					

$$k=V(1,\ 1,\ 1,\ 2,\ -1,\ 1,\ 1,\ 2,\ -1,\ 1,\ 1)=4.$$

Примечание. Не изменяя знаков элементов первого столбца, можно все элементы какой-либо строки помножить на одно и то же число. Это замечание было использовано при построении схемы.

4. Однако применение обоих правил Рауса не дает возможности во всех случаях определить число k. Применение первого правила (введение малых параметров ε, η, ...) обосновано лишь в том случае, когда многочлен $f(z)$ не имеет корней на мнимой оси.

Если многочлен $f(z)$ имеет корни на мнимой оси, то при варьировании параметра ε некоторые из этих корней могут перейти в правую полуплоскость и изменить число k.

Пример.

$$f(z)=z^6+z^5+3z^4+3z^3+3z^2+2z+1.$$

Схема

ω^6	1	3	3	1
ω^5	1	3	2	
ω^4	ε	1	1	
ω^3	$3-\frac{1}{\varepsilon}$	$2-\frac{1}{\varepsilon}$		$\left(2-\frac{1}{\varepsilon}-\frac{3-\frac{1}{\varepsilon}}{1-\frac{2\varepsilon-1}{3-\frac{1}{\varepsilon}}}=-\varepsilon+\ldots\right)$
ω^2	$1-\frac{2\varepsilon-1}{3-\frac{1}{\varepsilon}}$	1		
ω	u			$V\left(1,\ 1,\ \varepsilon,\ 3-\frac{1}{\varepsilon},\ 1,\ -\varepsilon,\ 1\right)=\begin{cases}4 \text{ при } \varepsilon>0,\\ 2 \text{ при } \varepsilon<0.\end{cases}$
ω^0	1			

Вопрос, чему равно k, остается открытым.

В общем случае, когда $f(z)$ имеет корни на мнимой оси, следует поступать так: Полагая $f(z)=F_1(z)+F_2(z)$, где

$$F_1(z)=a_0z^n+a_1z^{n-2}+\ldots,\quad F_2(z)=b_0z^{n-1}+b_1z^{n-3}+\ldots,$$

следует найти наибольший общий делитель $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Тогда $f(z)=d(z)f^*(z)$.

Если $f(z)$ имеет корень z, для которого $-z$ снова является корнем (этим свойством обладают и все корни на мнимой оси), то из $f(z)=0$ и $f(-z)=0$ следует: $F_1(z)=0$ и $F_2(z)=0$, т. е. z является корнем $d(z)$. Поэтому многочлен $f^*(z)$ не имеет корней z, для которых $-z$ является корнем $f^*(z)$ [1]).

Тогда

$$k=k_1+k_2,$$

где k_1 и k_2 — числа корней в правой полуплоскости многочленов $f^*(z)$ и $d(z)$; k_1 определяется по алгоритму Рауса, а $k_2=\frac{q-s}{2}$, где q — степень $d(z)$, а s — число вещественных корней многочлена $d(i\omega)$ [2]).

В последнем примере

$$d(z)=z^2+1,\quad f^*(z)=z^4+z^3+2z^2+2z+1.$$

Поэтому (см. пример на стр. 476) здесь $k_2=0$, $k_1=2$ и, следовательно,

$$k=2.$$

---

[1]) При определении многочлена $d(z)$ можно исходить не из функций $F_1(z)$ и $F_2(z)$, а из введенных ранее (стр. 472) функций $f_1(z)$ и $f_2(z)$. (Подробнее об этом см. сноску к стр. 490.)

[2]) $d(i\omega)$ — вещественный многочлен или становится таковым после сокращения на i. Число вещественных корней его можно определить при помощи теоремы Штурма.

§ 5. Теорема Ляпунова

Из исследований А. М. Ляпунова, опубликованных в 1892 г. в его монографии «Общая задача об устойчивости движения», вытекает теорема [1]), дающая необходимые и достаточные условия для того, чтобы все корни характеристического уравнения $|\lambda E - A| = 0$ вещественной матрицы $A = \| a_{ik} \|_1^n$ имели отрицательные вещественные части. Поскольку любой многочлен $f(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_n$ $(a_0 \neq 0)$ может быть представлен в виде характеристического определителя $|\lambda E - A|$ [2]), то теорема Ляпунова носит общий характер и относится к любому алгебраическому уравнению $f(\lambda) = 0$.

Пусть дана вещественная матрица $A = \| a_{ik} \|_1^n$ и однородный многочлен m-го измерения относительно переменных $x_1, x_2, \ldots, x_n$

$$V(\underbrace{x, x, \ldots, x}_{m}) \qquad [x = (x_1, x_2, \ldots, x_n)].$$

Найдем полную производную по t от функции $V(x, x, \ldots, x)$ в предположении, что x есть решение дифференциальной системы $\frac{dx}{dt} = Ax$. Тогда

$$\frac{d}{dt} V(x, x, \ldots, x) = V(Ax, x, \ldots, x) +$$

$$+ V(x, Ax, \ldots, x) + \ldots + V(x, x, \ldots, Ax) = W(x, x, \ldots, x), \quad (21)$$

где $W(x, x, \ldots, x)$ — снова однородный многочлен m-го измерения относительно $x_1, x_2, \ldots, x_n$. Равенство (21) определяет линейный оператор $\widehat{A}$, относящий каждому однородному многочлену m-го измерения $V(x, x, \ldots, x)$ некоторый однородный многочлен $W(x, x, \ldots, x)$ того же измерения m

$$W = \widehat{A}(V).$$

Мы ограничимся случаем $m = 2$ [3]). В этом случае $V(x, x)$ и $W(x, x)$ — квадратичные формы от переменных $x_1, x_2, \ldots, x_n$, связанные равенством

$$\frac{d}{dt} V(x, x) = V(Ax, x) + V(x, Ax) = W(x, x), \quad (22)$$

---

[1]) См. [19], § 20.

[2]) Для этого достаточно, например, положить:

$$A = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\frac{a_n}{a_0} \\ 1 & 0 & \ldots & 0 & -\frac{a_{n-1}}{a_0} \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & \ddots & \\ & & & \ddots & \\ 0 & 0 & \ldots & 1 & -\frac{a_1}{a_0} \end{Vmatrix}.$$

[3]) А. М. Ляпунов установил свою теорему (см. ниже теорему 3) при любом целом и положительном m.

откуда [1])

$$W = \widehat{A}(V) = A'V + VA. \tag{23}$$

Здесь $V = \|v_{ik}\|_1^n$ и $W = \|w_{ik}\|_1^n$ — симметрические матрицы, составленные соответственно из коэффициентов форм $V(x, x)$ и $W(x, x)$. Линейный оператор $\widehat{A}$ в пространстве симметрических матриц n-го порядка V целиком определяется заданием матрицы $A = \|a_{ik}\|_1^n$.

Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы A, то каждое характеристическое число оператора $\widehat{A}$ представляется в виде $\lambda_i + \lambda_k$ $(1 \leqslant i \leqslant k \leqslant n)$.

Действительно, пусть u_k — собственный вектор-столбец матрицы A', соответствующий характеристическому числу λ_k, т. е. $Au_k = \lambda_k u_k$ $(u_k \neq 0)$, и пусть $v_{ik} = u_i u_k'$. Тогда

$$\widehat{A}v_{ik} = A'u_iu_k' + u_iu_k'A = (A'u_i)\,u_k' + u_i\,(A'u_k)' =$$

$$= (\lambda_i + \lambda_k)\,u_iu_k' = (\lambda_i + \lambda_k)\,v_{ik} \qquad (i,\ k = 1, \ldots, n). \tag{23'}$$

Если все величины $(\lambda_i + \lambda_k)$ $(i, k = 1, \ldots, n)$ различны между собой, то из равенств (23') следует, что эти величины образуют полную систему характеристических чисел оператора $\widehat{A}$. Общий случай, когда среди сумм $\lambda_i + \lambda_k$ имеются равные между собой, получается из рассмотренного случая с помощью соображений непрерывности.

Из доказанного предложения следует, что оператор $\widehat{A}$ является неособенным, матрица $A = \|a_{ik}\|_1^n$ не имеет нулевых и двух противоположных характеристических чисел. В этом случае задание матрицы W однозначно определяет матрицы V в (23).

Таким образом, *если матрица* $A = \|a_{ik}\|_1^n$ *не имеет нулевых и двух противоположных характеристических чисел, то каждой квадратичной форме* $W(x, x)$ *отвечает одна и только одна квадратичная форма* $V(x, x)$, *связанная с* $V(x, x)$ *равенством* (22).

Теперь сформулируем теорему Ляпунова.

Теорема 3 (Ляпунова). *Если все характеристические числа вещественной матрицы* $A = \|a_{ik}\|_1^n$ *имеют отрицательные вещественные части, то любой отрицательно определенной квадратичной форме* $W(x, x)$ *отвечает положительно определенная квадратичная форма* $V(x, x)$, *связанная с формой* $W(x, x)$ *в силу уравнения*

$$\frac{dx}{dt} = Ax \tag{24}$$

равенством

$$\frac{d}{dt}V(x, x) = W(x, x). \tag{25}$$

Обратно, если для некоторой отрицательно определенной формы $W(x, x)$ *существует положительно определенная форма* $V(x, x)$, *связанная с* $W(x, x)$ *равенством* (25) *в силу уравнения* (24), *то все характеристические числа матрицы* $A = \|a_{ik}\|_1^n$ *имеют отрицательные вещественные части.*

---

[1]) Поскольку $V(x, y) = x'Vy$.

Доказательство. 1. Пусть все характеристические числа матрицы A имеют отрицательные вещественные части. Тогда для любого решения $x = e^{At}x_0$ системы (24) имеем: $\lim\limits_{t\to+\infty} x = 0$ [1]). Пусть формы $V(x, x)$ и $W(x, x)$ связаны формулой (25) и $W(x, x) < 0$ $(x \neq 0)$ [2]).

Допустим, что при некотором $x_0 \neq 0$

$$V_0 = V(x_0, x_0) \leqslant 0.$$

Но $\frac{d}{dt} V(x, x) = W(x, x) < 0 (x = e^{At}x_0)$. Поэтому при $t > 0$ величина $V(x, x)$ отрицательна и убывает при $t \to +\infty$, что находится в противоречии с равенством $\lim\limits_{t\to+\infty} V(x, x) = \lim\limits_{x\to 0} V(x, x) = 0$. Следовательно, $V(x, x) > 0$ при $x \neq 0$, т. е. $V(x, x)$ — положительно определенная квадратичная форма.

2. Пусть, обратно, дано, что в равенстве (25)

$$W(x, x) < 0, \quad V(x, x) > 0 \qquad (x \neq 0).$$

Из (25) следует:

$$V(x, x) = V(x_0, x_0) + \int_0 W(x, x)\, dt \qquad (x = e^{At}x_0). \tag{25'}$$

Докажем, что при произвольном $x_0 \neq 0$ столбец $x = e^{At}x_0$ как угодно близко подходит к нулю при некоторых сколь угодно больших значениях $t > 0$. Допустим противное. Тогда существует число $\nu > 0$ такое, что

$$W(x, x) < -\nu < 0 \qquad (x = e^{At}x_0, \ x_0 \neq 0, \ t > 0).$$

Но тогда из (25')

$$V(x, x) < V(x_0, x_0) - \nu t,$$

и, следовательно, при некоторых достаточно больших значениях t справедливо неравенство $V(x, x) < 0$, что противоречит условию.

Из доказанного следует, что при некоторых достаточно больших значениях t величина $V(x, x)$ $(x = e^{At}x_0, \ x_0 \neq 0)$ будет как угодно близка к нулю. Но $V(x, x)$ монотонно убывает при $t > 0$, поскольку $\frac{d}{dt} V(x, x) = W(x, x) < 0$. Поэтому $\lim\limits_{t\to\infty} V(x, x) = 0$.

Отсюда вытекает, что при любом $x_0 \neq 0$ имеет место равенство $\lim\limits_{t\to+\infty} e^{At}x_0 = 0$, т. е. $\lim\limits_{t\to+\infty} e^{At} = 0$. Это возможно лишь тогда, когда все характеристические числа матрицы A имеют отрицательные вещественные части (см. гл. V, § 6).

Теорема доказана полностью.

В качестве формы $W(x, x)$ в теореме Ляпунова можно взять любую отрицательно определенную форму и, в частности, форму $-\sum\limits_{i=1}^{n} x_i^2$. В этом случае теорема допускает следующую матричную формулировку:

---

[1]) См. гл. V, § 6.

[2]) Форма $W(x, x)$ нам произвольно задана. Форма $V(x, x)$ однозначно определяется из условия (25), поскольку в данном случае матрица A не имеет нулевых и двух противоположных характеристических чисел.

Теорема 3'. *Для того чтобы все характеристические числа вещественной матрицы* $A=\|a_{ik}\|_1^n$ *имели отрицательные вещественные части, необходимо и достаточно, чтобы матричное уравнение*

$$A'V+VA=-E \tag{26}$$

имело в качестве решения V *матрицу коэффициентов некоторой положительно определенной квадратичной формы* $V(x,\ x)>0$.

Из доказанной теоремы вытекает критерий для определения устойчивости нелинейной системы по ее линейному приближению [1]).

Пусть требуется доказать асимптотическую устойчивость нулевого решения нелинейной системы дифференциальных уравнений (1) (стр. 419) в том случае, когда коэффициенты $a_{ik}\,(i,\ k=1,\ 2,\ \ldots,\ n)$ в линейных членах правых частей уравнений образуют матрицу $A=\|A_{ik}\|_1^n$, имеющую только характеристические числа с отрицательными вещественными частями. Тогда, определяя положительно определенную форму $V(x,\ x)$ при помощи матричного уравнения (26) и вычисляя ее полную производную по времени в предположении, что $x=(x_1,\ x_2,\ \ldots,\ x_n)$ есть решение системы (1), будем иметь:

$$\frac{d}{dt}V(x,\ x)=-\sum_{i=1}^{n}x_i^2+R(x_1,\ x_2,\ \ldots,\ x_n),$$

где $R(x_1,\ x_2,\ \ldots,\ x_n)$ — ряд, содержащий члены третьего и более высоких измерений относительно $x_1,\ x_2,\ \ldots,\ x_n$. Поэтому в некоторой достаточно малой окрестности точки $(0,\ 0,\ \ldots,\ 0)$ для любого $x\neq 0$ одновременно

$$V(x,\ x)>0,\qquad \frac{d}{dt}V(x,\ x)<0.$$

Согласно общему критерию устойчивости Ляпунова [2]) это и означает асимптотическую устойчивость нулевого решения системы дифференциальных уравнений.

Если из матричного уравнения (26) выразить элементы матрицы V через элементы матрицы A и полученные выражения подставить в неравенства

$$v_{11}>0,\quad \begin{vmatrix} v_{11} & v_{12}\\ v_{21} & v_{22}\end{vmatrix}>0,\quad \ldots,\quad \begin{vmatrix} v_{11} & v_{12} & \ldots & v_{1n}\\ v_{21} & v_{22} & \ldots & v_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ v_{n1} & v_{n2} & \ldots & v_{nn}\end{vmatrix}>0,$$

то мы получим неравенства, которым должны удовлетворять элементы матрицы $A=\ a_{ik}\|_1^n$ для того, чтобы все характеристические числа матрицы имели отрицательные вещественные части. Однако в значительно более простом виде эти неравенства могут быть получены из критерия Рауса — Гурвица, которому посвящается следующий параграф.

---

[1]) См. [22], § 26; [38], стр. 113 и далее; [23], стр. 66 и далее.
[2]) См. [22], § 16; [38], стр. 19—21 и 31—33; [23], стр. 32—34.

Примечание. Теорема Ляпунова (3) или (3′) непосредственно обобщается на случай произвольной комплексной матрицы $A = \|a_{ik}\|_1^n$. В этом случае квадратичные формы $V(x, x)$ и $W(x, x)$ заменяются эрмитовыми

$$V(x, x) = \sum_{i,k=1}^{n} v_{ik}\bar{x}_i x_k, \qquad W(x, x) = \sum_{i,k=1}^{n} w_{ik}\bar{x}_i x_k.$$

В соответствии с этим матричное уравнение (26) заменится уравнением

$$A^*V + VA = -E \quad (A^* = \bar{A}').$$

§ 6. Теорема Рауса — Гурвица

В предыдущих параграфах был изложен непревзойденный по своей простоте метод Рауса для определения числа k корней в правой полуплоскости вещественного многочлена, коэффициенты которого заданы как конкретные числа. Если же коэффициенты многочлена зависят от параметров и требуется определить, при каких значениях параметров число k будет иметь то или другое значение и, в частности, значение 0 (область устойчивости!)[1]), то желательно иметь конкретные выражения для величин c_0, d_0, ... через коэффициенты данного многочлена. Разрешив эту задачу, мы получим метод определения числа k и, в частности, критерий устойчивости в том виде, в каком он был установлен Гурвицем [129].

Рассмотрим снова многочлен

$$f(z) = a_0 z^n + b_0 z^{n-1} + a_1 z^{n-2} + b_1 z^{n-3} + \ldots \quad (a_0 \neq 0).$$

Назовем *матрицей Гурвица* квадратную матрицу n-го порядка

$$H = \begin{Vmatrix} b_0 & b_1 & b_2 & \ldots & b_{n-1} \\ a_0 & a_1 & a_2 & \ldots & a_{n-1} \\ 0 & b_0 & b_1 & \ldots & b_{n-2} \\ 0 & a_0 & a_1 & \ldots & a_{n-2} \\ 0 & 0 & b_0 & \ldots & b_{n-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{Vmatrix} \quad \left(\begin{matrix} a_k = 0 \text{ при } k > \left[\frac{n}{2}\right], \\ b_k = 0 \text{ при } k > \left[\frac{n-1}{2}\right] \end{matrix} \right). \tag{27}$$

Преобразуем эту матрицу, вычитая из второй, четвертой, ... соответственно первую, третью, ... строки, предварительно помноженные на $\frac{a_0}{b_0}$[2]). Получим матрицу

$$\begin{Vmatrix} b_0 & b_1 & b_2 & \ldots & b_{n-1} \\ 0 & c_0 & c_1 & \ldots & c_{n-2} \\ 0 & b_0 & b_1 & \ldots & b_{n-2} \\ 0 & 0 & c_0 & \ldots & c_{n-3} \\ 0 & 0 & b_0 & \ldots & b_{n-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{Vmatrix}.$$

[1]) Так именно и обстоит дело при проектировании новых механических или электрических систем регулирования.

[2]) Сначала рассматривается регулярный случай, когда $b_0 \neq 0$, $c_0 \neq 0$, $d_0 \neq 0$, ...

Здесь $c_0, c_1, \ldots$ — третья строка схемы Рауса, дополненная нулями $\left(c_k = 0 \text{ при } k > \left[\frac{n}{2}\right] - 1\right)$.

Полученную матрицу снова преобразуем, вычитая из третьей, пятой, ... строки соответственно вторую, четвертую, ... строки, предварительно помноженные на $\frac{b_0}{c_0}$:

$$\left\| \begin{matrix} b_0 & b_1 & b_2 & b_3 & \ldots \\ 0 & c_0 & c_1 & c_2 & \ldots \\ 0 & 0 & d_0 & d_1 & \ldots \\ 0 & 0 & c_0 & c_1 & \ldots \\ 0 & 0 & 0 & d_0 & \ldots \\ 0 & 0 & 0 & c_0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{matrix} \right\|.$$

Продолжая этот процесс далее, мы придем в конце концов к треугольной матрице n-го порядка

$$\ldots R = \left\| \begin{matrix} b_0 & b_1 & b_2 & \ldots \\ 0 & c_0 & c_1 & \ldots \\ 0 & 0 & d_0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{matrix} \right\|,$$

которую назовем *матрицей Рауса*. Она получается из схемы Рауса [см. (15)] 1) отбрасыванием первой строки, 2) сдвигом строк вправо так, чтобы их первые элементы пришлись на главную диагональ, и 3) пополнением нулями до квадратной матрицы n-го порядка.

Определение 2. Две матрицы $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ назовем *равносильными* в том и только в том случае, если при любом $p \leqslant n$ в первых p строках этих матриц соответствующие миноры p-го порядка равны между собой:

$$A \begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = B \begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \quad \begin{pmatrix} i_1, & i_2, & \ldots, & i_p = 1, & 2, & \ldots, & n, \\ & & & p = 1, & 2, & \ldots, & n \end{pmatrix}.$$

Так как при вычитании из какой-либо строки матрицы какой-либо предыдущей строки, помноженной предварительно на произвольное число, миноры p-го порядка в первых p строках ($p = 1, 2, \ldots, n$) не меняют своей величины, то согласно определению 2 матрицы Гурвица и Рауса H и R равносильны:

$$H \begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = R \begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \quad \begin{pmatrix} i_1, & i_2, & \ldots, & i_p = 1, & 2, & \ldots, & n, \\ & & & p = 1, & 2, & \ldots, & n. \end{pmatrix} \quad (28)$$

Равносильность матриц H и R позволяет выразить все элементы матрицы R, т. е. элементы схемы Рауса, через миноры матрицы Гурвица H и, следовательно, через коэффициенты данного многочлена. Действи-

тельно, давая p в (28) последовательно значения 1, 2, 3, ..., получим:

$$\left.\begin{array}{lll} H\begin{pmatrix}1\\1\end{pmatrix}=b_0, & H\begin{pmatrix}1\\2\end{pmatrix}=b_1, & H\begin{pmatrix}1\\3\end{pmatrix}=b_2,\ \ldots, \\ H\begin{pmatrix}1&2\\1&2\end{pmatrix}=b_0c_0, & H\begin{pmatrix}1&2\\1&3\end{pmatrix}=b_0c_1, & H\begin{pmatrix}1&2\\1&4\end{pmatrix}=b_0c_2,\ \ldots, \\ H\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=b_0c_0d_0, & H\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}=b_0c_0d_1, & H\begin{pmatrix}1&2&3\\1&2&5\end{pmatrix}=b_0c_0d_2,\ \ldots \end{array}\right\} \quad (29)$$

и т. д.

Отсюда находим следующие выражения для элементов схемы Рауса:

$$\left.\begin{array}{lll} b_0=H\begin{pmatrix}1\\1\end{pmatrix}, & b_1=H\begin{pmatrix}1\\2\end{pmatrix}, & b_2=H\begin{pmatrix}1\\3\end{pmatrix},\ \ldots, \\ c_0=\dfrac{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}}, & c_1=\dfrac{H\begin{pmatrix}1&2\\1&3\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}}, & c_2=\dfrac{H\begin{pmatrix}1&2\\1&4\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}},\ \ldots, \\ d_0=\dfrac{H\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}, & d_1=\dfrac{H\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}, & d_2=\dfrac{H\begin{pmatrix}1&2&3\\1&2&5\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}},\ \ldots \\ \ldots\ldots\ldots\

алгоритма Рауса. Так, например, если $b_0 = H\begin{pmatrix}1\\1\end{pmatrix} = 0$, то алгоритм Рауса нам не дает ничего, кроме первых двух строк, составленных из коэффициентов данного многочлена. Однако, если при $\Delta_1 = 0$ остальные определители $\Delta_2, \Delta_3, \ldots$ отличны от нуля, мы при помощи формул (30), минуя строку из c, можем определить все последующие строки схемы Рауса.

Согласно формулам (32)

$$b_0 = \Delta_1, \quad c_0 = \frac{\Delta_2}{\Delta_1}, \quad d_0 = \frac{\Delta_3}{\Delta_2}, \quad \ldots,$$

и потому

$$V(a_0, b_0, c_0, \ldots) = V\left(a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \ldots, \frac{\Delta_n}{\Delta_1}\right) =$$

$$= V(a_0, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots).$$

Поэтому теорема Рауса может быть сформулирована так:

Теорема 4 (Рауса—Гурвица). *Число k корней вещественного многочлена $f(z) = a_0 z^n + \ldots$, расположенных в правой полуплоскости, определяется формулой*

$$k = V\left(a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \frac{\Delta_3}{\Delta_2}, \ldots, \frac{\Delta_n}{\Delta_{n-1}}\right) \tag{33}$$

или (что то же)

$$k = V(a_0, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots). \tag{33'}$$

Примечание. Приведенная формулировка теоремы Рауса—Гурвица предполагает, что имеет место регулярный случай

$$\Delta_1 \neq 0, \quad \Delta_2 \neq 0, \quad \ldots, \quad \Delta_n \neq 0.$$

В § 8 мы покажем, как пользоваться этой формулой в особых случаях, когда некоторые из определителей Гурвица Δ_i равны нулю.

Рассмотрим теперь тот частный случай, когда все корни многочлена $f(z)$ расположены в левой полуплоскости $\operatorname{Re} z < 0$. В этом случае согласно критерию Рауса все $a_0, b_0, c_0, d_0, \ldots$ должны быть отличны от нуля и одного знака. Так как здесь мы имеем дело с регулярным случаем, то получаем из (33) при $k = 0$ следующий критерий:

Теорема 5 (Критерий Рауса—Гурвица). *Для того чтобы у вещественного многочлена $f(z) = a_0 z^n + \ldots$ ($a_0 \neq 0$) все корни имели отрицательные вещественные части, необходимо и достаточно, чтобы имели место неравенства*

$$a_0\Delta_1 > 0, \ \Delta_2 > 0, \ a_0\Delta_3 > 0, \ \Delta_4 > 0, \ \ldots, \ \begin{matrix} a_0\Delta_n > 0 \text{ (при } n \text{ нечетном)}, \\ \Delta_n > 0 \text{ (при } n \text{ четном)}. \end{matrix} \tag{34}$$

Примечание. Если $a_0 > 0$, то эти условия записываются так:

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \ldots, \quad \Delta_n > 0. \tag{35}$$

Если принять обычные обозначения для коэффициентов многочлена $f(z) = a_0 z^n - a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_{n-1} z + a_n$, то при $a_0 > 0$ условия Рауса—Гурвица (35) записываются в виде следующих детерминантных

неравенств:

$$a_1>0,\quad \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}>0,\quad \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a & a_3 \end{vmatrix}>0,\quad \ldots,\quad \begin{vmatrix} a_1 & a_3 & a_5 \ldots 0 \\ a_0 & a_2 & a_4 \ldots 0 \\ 0 & a_1 & a_3 \ldots 0 \\ 0 & a_0 & a_2 \ldots 0 \\ \cdot & \cdot & \cdot \ldots \cdot \\ \cdot & \cdot & \cdot \ldots a_n \end{vmatrix}>0. \quad (35')$$

Вещественный многочлен $f(z)=a_0z^n+\ldots$, коэффициенты которого удовлетворяют условиям (34), т. е. вещественный многочлен, у которого все корни имеют отрицательные вещественные части, обычно называют *многочленом Гурвица.*

Отметим два замечательных свойства схемы Рауса:

1. Обозначим элементы $p+1$-й строки схемы Рауса через α_{p0}, α_{p1}, α_{p2}, ...; тогда $\alpha_{pj}=\dfrac{\Delta_p^{(p+j)}}{\Delta_{p-1}}$ ($p, j=0, 1, \ldots$). Здесь $\Delta_p^{(p)}=\Delta_p=H\begin{pmatrix}1 \ldots p\\ 1 \ldots p\end{pmatrix}$ — определитель Гурвица, а $\Delta_p^{(p+j)}=H\begin{pmatrix}1 \ldots p-1\ p\\ 1 \ldots p-1\ p+j\end{pmatrix}$ при $j\geqslant 1$ — «побочный» определитель Гурвица p-го порядка. Между элементами схемы Рауса имеет место основная зависимость [см. формулы (13), (13') на стр. 473]:

$$\alpha_{pj}=\frac{\alpha_{p0}}{\alpha_{p+1,0}}\alpha_{p+1,j+1}+\alpha_{p+2,j-1} \qquad (p, j=0, 1, \ldots;\ \alpha_{kl}=0, \text{ если } k>n \text{ либо } j<0).$$

Элементы любой p-й строки схемы Рауса получаются из элементов двух последующих строк с помощью двух операций — умножения на отношение $\alpha_{p0}/\alpha_{p+1,0}$ и сложения. Поэтому (в регулярном случае!) элементы произвольной p-й строки схемы Рауса могут быть выражены с помощью операций сложения и умножения через элементы последних двух строк $\alpha_{n-1,0}$ и $\alpha_{n,0}$ и через отношения $\dfrac{\alpha_{p0}}{\alpha_{p+1,0}}, \ldots, \dfrac{\alpha_{n-2,0}}{\alpha_{n-1,0}}$ и представлены в виде

$$\alpha_{pj}=\frac{\varphi_{pj}(\alpha_{p0}, \alpha_{p+1,0}, \ldots, \alpha_{n0})}{\alpha_{p+1,0}\ldots\alpha_{n0}} \qquad (p, j=0, 1, \ldots), \qquad (36)$$

где $\varphi_{pj}(\alpha_{p0}, \alpha_{p+1,0}, \ldots, \alpha_{n0})$ — многочлены с целыми положительными коэффициентами.

С помощью формул (36) все элементы схемы Рауса и, в частности (при $p=0, 1$), коэффициенты исходного многочлена $f(z)$ рационально выражаются (и притом с положительными коэффициентами) через элементы первого столбца схемы Рауса.

Если выполняется критерий Рауса, т. е. все элементы первого столбца схемы Рауса положительны, то из формулы (36) непосредственно следует, что в этом случае все элементы схемы Рауса и, в частности, коэффициенты основного многочлена, положительны.

Заметим еще, что, заменяя в формулах (36) величины α_{pj} через отношения $\Delta_p^{(p+j)}/\Delta_{p-1}$, можно рационально (с положительными коэффициентами) выразить побочные определители Гурвица $\Delta_p^{(p+j)}$ через основные.

2. Пусть $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$ — $(m+1)$-я и $(m+2)$-я строки схемы $\left(f_0=\dfrac{\Delta_m}{\Delta_{m-1}},\ g_0=\dfrac{\Delta_{m+1}}{\Delta_m}\right)$. Так как эти две строки вместе с последующими образуют самостоятельную схему Рауса, то элементы $(m+p+1)$-й строки (в первоначальной схеме) выражаются через элементы $(m+1)$-й и $(m+2)$-й строк, $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$ — по тем же формулам, по каким элементы $(p+1)$-й строки выражаются через элементы первых

двух строк $a_0, a_1, \ldots$ и $b_0, b_1, \ldots$, т. е., полагая

$$\widetilde{H}=\begin{Vmatrix} g_0 & g_1 & g_2 \ldots \\ f_0 & f_1 & f_2 \ldots \\ 0 & g_0 & g_1 \ldots \\ 0 & f_0 & f_1 \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Vmatrix},$$

будем иметь:

$$\frac{\Delta_{m+p}^{(m+j)}}{\Delta_{m+p-1}}=\frac{\widetilde{\Delta}_p^{(j)}}{\widetilde{\Delta}_{p-1}} \qquad (j=p,\ p+1,\ \ldots). \tag{37}$$

Определитель Гурвица Δ_{m+p} равен произведению первых $m+p$ чисел в ряду $b_0, c_0, \ldots$:

$$\Delta_{m+p}=b_0c_0 \ldots f_0g_0 \ldots l_0.$$

Но

$$\Delta_m=b_0c_0 \ldots f_0, \quad \widetilde{\Delta}_p=g_0 \ldots l_0.$$

Поэтому имеет место следующее важное соотношение:

$$\Delta_{m+p}=\Delta_m\widetilde{\Delta}_p\ ^{1}). \tag{38}$$

Формула (38) имеет место всегда, если только определены числа $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$, т. е. при условии $\Delta_{m-1}\neq 0$, $\Delta_m\neq 0$.

Формулы (37) имеют смысл, если дополнительно к условиям $\Delta_{m-1}\neq 0$, $\Delta_m\neq 0$ выполняется и условие $\Delta_{m+p-1}\neq 0$. Из этого условия уже следует, что и знаменатель дроби, стоящей в правой части равенства (37), не равен нулю: $\widetilde{\Delta}_{p-1}\neq 0$.

§ 7. Формула Орландо

При рассмотрении случаев, когда некоторые из определителей Гурвица равны нулю, нам понадобится следующая формула Орландо [137], выражающая определитель Δ_{n-1} через старший коэффициент a_0 и корни $z_1, z_2, \ldots, z_n$ многочлена $f(z)$ [2])

$$\Delta_{n-1}=(-1)^{\frac{n(n-1)}{2}}a_0^{n-1}\prod_{i<k}^{1\ldots n}(z_i+z_k). \tag{39}$$

При $n=2$ эта формула сводится к известной формуле для коэффициента b_0 в квадратном уравнении $a_0z^2+b_0z+a_1=0$:

$$\Delta_1=b_0=-a_0(z_1+z_2).$$

Допустим теперь, что формула (39) справедлива для многочлена n-й степени $f(z)=a_0z^n+b_0z^{n-1}+\ldots$, и покажем, что она справедлива для многочлена $(n+1)$-й степени

$$F(z)=(z+h)f(z)=$$
$$=a_0z^{n+1}+(b_0+ha_0)z^n+(a_1+hb_0)z^{n-1}+\ldots \quad (h=-z_{n+1}).$$

---

[1]) Здесь $\widetilde{\Delta}_p$ — минор p-го порядка, стоящий в левом верхнем углу матрицы $\widetilde{H}$.

[2]) При этом коэффициентами многочлена $f(z)$ могут быть произвольные комплексные числа.

Для этого составим вспомогательный определитель $(n+1)$-го порядка

$$D=\begin{vmatrix} b_0 & b_1 & \dots & b_{n-1} & h^n \\ a_0 & a_1 & \dots & a_{n-1} & -h^{n-1} \\ 0 & b_0 & \dots & b_{n-2} & h^{n-2} \\ 0 & a_0 & \dots & a_{n-2} & -h^{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & (-1)^n \end{vmatrix} \left(\begin{matrix} a_k=0 & \text{при} & k>\left[\frac{n}{2}\right], \\ b_k=0 & \text{при} & k>\left[\frac{n-1}{1}\right] \end{matrix} \right).$$

Помножим первую строку D на a_0 и прибавим к ней вторую, помноженную на $-b_0$, третью, помноженную на a_1, четвертую — на $-b_1$ и т. д. Тогда все элементы первой строки, кроме последнего, обратятся в нуль, а последний элемент будет равен $f(h)$. Отсюда легко заключаем, что

$$D=(-1)^n\Delta_{n-1}f(h).$$

С другой стороны, прибавляя к каждой (кроме последней) строке определителя D последующую, помноженную на h, мы получим помноженный на $(-1)^n$ определитель Гурвица Δ_n^* n-го порядка для многочлена $F(z)$:

$$D=(-1)^n\begin{vmatrix} b_0+ha_0 & b_1+ha_1 & \dots \\ a_0 & a_1+hb_0 & \dots \\ 0 & b_0+ha_0 & \dots \\ 0 & a_0 & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix}=(-1)^n\Delta_n^*.$$

Таким образом,

$$\Delta_n^*=\Delta_{n-1}f(h)=a_0\Delta_{n-1}\prod_{i=1}^{n}(h-z_i).$$

Заменяя здесь Δ_{n-1} на его выражение из (39) и полагая $h=-z_{n+1}$, получаем:

$$\Delta_n^*=(-1)^{\frac{(n+1)n}{2}}a_0^n\prod_{i<k}^{1\dots n+1}(z_i+z_k).$$

Таким образом, методом математической индукции установлена справедливость формулы Орландо для многочлена любой степени.

Из формулы Орландо следует, что $\Delta_{n-1}=0$ *тогда и только тогда, когда сумма двух корней многочлена* $f(z)$ *равна нулю*[1]).

Так как $\Delta_n=c\Delta_{n-1}$, где c — свободный член многочлена $f(z)$ $(c=(-1)^n a_0z_1z_2\dots z_n)$, то из (39) следует:

$$\Delta_n=(-1)^{\frac{n(n+1)}{2}}a_0^nz_1z_2\dots z_n\prod_{i<k}^{1\dots n}(z_i+z_k). \tag{40}$$

Последняя формула показывает, что Δ_n *обращается в нуль тогда и только тогда, когда у* $f(z)$ *существует такой корень* z, *что и* $-z$ *является корнем.*

[1]) В частности, $\Delta_{n-1}=0$, когда $f(z)$ имеет хотя бы одну пару сопряженных чисто мнимых корней или кратный нулевой корень.

§ 8. Особые случаи в теореме Рауса—Гурвица

При рассмотрении особых случаев, когда некоторые из определителей Гурвица равны нулю, мы можем предполагать, что $\Delta_n \neq 0$ (и, следовательно, $\Delta_{n-1} \neq 0$).

Действительно, если $\Delta_n=0$, то, как было выяснено в конце предыдущего параграфа, вещественный многочлен $f(z)$ имеет такой корень z', для которого $-z'$ также является корнем $f(z)$. Если положить $f(z)=F_1(z)+F_2(z)$, где

$$F_1(z)=a_0z^n+a_1z^{n-1}+\ldots, \quad F_2(z)=b_0z^{n-1}+b_1z^{n-1}+\ldots,$$

то из равенств $f(z')=f(-z')=0$ можно заключить, что $F_1(z')=F_2(z')=0$. Следовательно, z' будет корнем наибольшего общего делителя $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Полагая $f(z)=d(z)f^*(z)$, мы сведем задачу Рауса—Гурвица для $f(z)$ к такой же задаче для многочлена $f^*(z)$, для которого уже последний определитель Гурвица отличен от нуля[1]).

1. Рассмотрим сначала тот случай, когда

$$\Delta_1=\ldots=\Delta_p=0, \quad \Delta_{p+1}\neq 0, \ \ldots, \ \Delta_n\neq 0. \tag{41}$$

Из $\Delta_1=0$ следует: $b_0=0$; из $\Delta_2=\begin{vmatrix} 0 & b_1 \\ a_0 & a_1 \end{vmatrix}=-a_0b_1=0$ вытекает: $b_1=0$. Но тогда автоматически

$$\Delta_3=\begin{vmatrix} 0 & b_1 & b_2 \\ a_0 & a_1 & a_2 \\ 0 & 0 & b_1 \end{vmatrix}=-a_0b_1^2=0.$$

Из

$$\Delta_4=\begin{vmatrix} 0 & 0 & b_2 & b_3 \\ a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & b_2 \\ 0 & a_0 & a_1 & a_2 \end{vmatrix}=-a_0^2b_2^2=0$$

следует: $b_2=0$, а тогда $\Delta_5=-a_0^2b_2^3=0$ и т. д.

Приведенные рассуждения показывают, что в (41) всегда p—*нечетное* число: $p=2h-1$. При этом $b_0=b_1=b_2=\ldots=b_{h-1}=0$, $b_h\neq 0$ и [2])

$$\Delta_{p+1}=\Delta_{2h}=(-1)^{\frac{h(h+1)}{2}}a_0^hb_h^h, \quad \Delta_{p+2}=\Delta_{2h+1}=(-1)^{\frac{h(h+1)}{2}}a_0^hb_h^{h+1}=\Delta_{p+1}b_h. \tag{42}$$

---

[1]) В случае, когда

$$\Delta_n=\Delta_{n-1}=\ldots=\Delta_{m+1}=0, \quad \Delta_m\neq 0, \quad \Delta_{m-1}\neq 0, \ \ldots, \ \Delta_1\neq 0,$$

можно в явном виде записать уравнения $d(z)=0$. Действительно, функции $F_1(z)$ и $F_2(z)$ связаны с функциями $f_1(\omega)$ и $f_2(\omega)$ [см. формулу (9) на стр. 472] соотношениями

$$F_1(z)=i^nf_1(-iz), \qquad F_2(z)=i^{n-1}f_2(-iz).$$

Поэтому уравнение $d(z)=0$ совпадает с уравнением $f_{m+1}(-iz)=0$, где многочлен $f_{m+1}(\omega)$—наибольший общий делитель многочленов $f_1(z)$ и $f_2(z)$—определяется последней строкой схемы Рауса. Следовательно, в силу формул (31) уравнение $d(z)=0$ может быть записано в виде

$$\sum_{j=0}^{\left[\frac{n-m}{2}\right]}\Delta_m^{(m+j)}z^{n-m-2j}=0,$$

где $\Delta_m^{(m+j)}=H\begin{pmatrix} 1 \ldots m-1 & m \\ 1 \ldots m-1 & m+j \end{pmatrix}$ $(j=0, 1, \ldots)$—побочные определители Гурвица.

[2]) Из (42) следует, что при h нечетном $\operatorname{sign}\Delta_{p+2}=(-1)^{\frac{h+1}{2}}\operatorname{sign}a_0$, а при h четном $\operatorname{sign}\Delta_{p+1}=(-1)^{\frac{h}{2}}$.

Проварьируем, т. е. изменим немного, коэффициенты $b_0, b_1, \ldots, b_{h-1}$ так, чтобы при новых проварьированных значениях $b_0^*, b_1^*, \ldots, b_{h-1}^*$ все определители Гурвица $\Delta_1^*, \Delta_2^*, \ldots, \Delta_n^*$ стали отличными от нуля и чтобы при этом определители $\Delta_{p+1}^*, \ldots, \Delta_n^*$ сохранили свои прежние знаки. Мы будем считать $b_0^*, b_1^*, \ldots, b_{h+1}^*$ «малыми» величинами разного порядка «малости», а именно мы примем, что каждое b_{j-1}^* по абсолютной величине «значительно» меньше b_j^* ($j=1, 2, \ldots, h$; $b_h^*=b_h$). Последнее означает, что при вычислении знака целого алгебраического выражения относительно b_i^* мы можем пренебрегать членами, в которых некоторые b_i^* имеют индекс $<j$ по сравнению с членами, где все b_i^* имеют индекс $\geqslant j$. После этого мы легко найдем «знакоопределяющие» члены в $\Delta_1^*, \Delta_2^*, \ldots, \Delta_p^*$ ($p=2h-1$) [1]:

$$\Delta_1^*=b_0^*,\ \Delta_2^*=-a_0b_1^*+\ldots,\ \Delta_3^*=-a_0b_1^{*2}+\ldots,\ \Delta_4^*=-a_0^2b_2^{*2}+\ldots,$$

$$\Delta_5^*=-a_0^2b_2^{*3}+\ldots,\ \Delta_6^*=a_0^3b_3^{*3}+\ldots$$

и т. д.; вообще

$$\left.\begin{aligned}\Delta_{2j}^*&=(-1)^{\frac{j(j+1)}{2}}a_0^jb_j^{*j}+\ldots \qquad (j=1, 2, \ldots, h-1),\\ \Delta_{2j+1}^*&=(-1)^{\frac{j(j+1)}{2}}a_0^jb_j^{*j+1}+\ldots \qquad (j=0, 1, \ldots, h-1).\end{aligned}\right\} \tag{43}$$

Выберем $b_0^*, b_1^*, \ldots, b_{2h-1}^*$ положительными; тогда знаки Δ_i^* определятся из формулы

$$\operatorname{sign}\Delta_i^*=(-1)^{\frac{j(j+1)}{2}}\operatorname{sign}a_0^j \qquad \left(j=\left[\frac{i}{2}\right],\ i=1, 2, \ldots, p\right). \tag{44}$$

При любом малом варьировании коэффициентов многочлена число k остается неизменным, поскольку многочлен $f(z)$ не имеет корней на мнимой оси. Поэтому, исходя из (44), мы определяем число корней в правой полуплоскости по формуле

$$k=V\left(a_0,\ \Delta_1^*,\ \frac{\Delta_2^*}{\Delta_2^*},\ \ldots,\ \frac{\Delta_{p+1}}{\Delta_p^*},\ \frac{\Delta_{p+2}}{\Delta_{p+1}}\right)+V\left(\frac{\Delta_{p+2}}{\Delta_{p+1}},\ \ldots,\ \frac{\Delta_n}{\Delta_{n-1}}\right). \tag{45}$$

Элементарный подсчет, проведенный на основании формул (42) и (44), показывает, что

$$V\left(a_0,\ \Delta_1^*,\ \frac{\Delta_2^*}{\Delta_1^*},\ \ldots,\ \frac{\Delta_{p+1}}{\Delta_p^*},\ \frac{\Delta_{p+2}}{\Delta_{p+1}}\right)=h+\frac{1-(-1)^h\varepsilon}{2}\left(\begin{matrix}p=2h-1,\\ \varepsilon=\operatorname{sign}\left(a_0\dfrac{\Delta_{p+2}}{\Delta_{p+1}}\right)\end{matrix}\right). \tag{46}$$

Заметим, что величина, стоящая в левой части равенства (46), не зависит от способа варьирования коэффициентов и при любых малых варьированиях сохраняет одно и то же значение. Это следует из формулы (45), поскольку k не меняет своего значения при малом варьировании коэффициентов.

2. Пусть теперь при $s>0$

$$\Delta_{s+1}=\ldots=\Delta_{s+p}=0, \tag{47}$$

а все остальные определители Гурвица отличны от нуля.

Обозначим через $\widetilde{a}_0, \widetilde{a}_1, \ldots$ и $\widetilde{b}_0, \widetilde{b}_1, \ldots$ элементы $(s+1)$-й и $(s+2)$-й строки в схеме Рауса $\left(\widetilde{a}_0=\dfrac{\Delta_s}{\Delta_{s-1}},\ \widetilde{b}_0=\dfrac{\Delta_{s+1}}{\Delta_s}\right)$. Соответствующие определители Гурвица

[1] По существу аналогичные члены уже были вычислены выше для $\Delta_1, \Delta_2, \ldots, \Delta_p$.

§ 8. Особые случаи в теореме Рауса — Гурвица

При рассмотрении особых случаев, когда некоторые из определителей Гурвица равны нулю, мы можем предполагать, что $\Delta_n \neq 0$ (и, следовательно, $\Delta_{n-1} \neq 0$).

Действительно, если $\Delta_n = 0$, то, как было выяснено в конце предыдущего параграфа, вещественный многочлен $f(z)$ имеет такой корень z', для которого $-z'$ также является корнем $f(z)$. Если положить $f(z) = F_1(z) + F_2(z)$, где

$$F_1(z) = a_0 z^n + a_1 z^{n-1} + \ldots, \quad F_2(z) = b_0 z^{n-1} + b_1 z^{n-1} + \ldots,$$

то из равенств $f(z') = f(-z') = 0$ можно заключить, что $F_1(z') = F_2(z') = 0$. Следовательно, z' будет корнем наибольшего общего делителя $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Полагая $f(z) = d(z) f^*(z)$, мы сведем задачу Рауса—Гурвица для $f(z)$ к такой же задаче для многочлена $f^*(z)$, для которого уже последний определитель Гурвица отличен от нуля [1]).

1. Рассмотрим сначала тот случай, когда

$$\Delta_1 = \ldots = \Delta_p = 0, \quad \Delta_{p+1} \neq 0, \ \ldots, \ \Delta_n \neq 0. \tag{41}$$

Из $\Delta_1 = 0$ следует: $b_0 = 0$; из $\Delta_2 = \begin{vmatrix} 0 & b_1 \\ a_0 & a_1 \end{vmatrix} = -a_0 b_1 = 0$ вытекает: $b_1 = 0$. Но тогда автоматически

$$\Delta_3 = \begin{vmatrix} 0 & b_1 & b_2 \\ a_0 & a_1 & a_2 \\ 0 & 0 & b_1 \end{vmatrix} = -a_0 b_1^2 = 0.$$

Из

$$\Delta_4 = \begin{vmatrix} 0 & 0 & b_2 & b_3 \\ a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & b_2 \\ 0 & a_0 & a_1 & a_2 \end{vmatrix} = -a_0^2 b_2^2 = 0$$

следует: $b_2 = 0$, а тогда $\Delta_5 = -a_0^2 b_2^3 = 0$ и т. д.

Приведенные рассуждения показывают, что в (41) всегда p — *нечетное* число: $p = 2h - 1$. При этом $b_0 = b_1 = b_2 = \ldots = b_{h-1} = 0$, $b_h \neq 0$ и [2])

$$\Delta_{p+1} = \Delta_{2h} = (-1)^{\frac{h(h+1)}{2}} a_0^h b_h^h, \quad \Delta_{p+2} = \Delta_{2h+1} = (-1)^{\frac{h(h+1)}{2}} a_0^h b_h^{h+1} = \Delta_{p+1} b_h. \tag{42}$$

---

[1]) В случае, когда

$$\Delta_n = \Delta_{n-1} = \ldots = \Delta_{m+1} = 0, \quad \Delta_m \neq 0, \quad \Delta_{m-1} \neq 0, \ \ldots, \ \Delta_1 \neq 0,$$

можно в явном виде записать уравнения $d(z) = 0$. Действительно, функции $F_1(z)$ и $F_2(z)$ связаны с функциями $f_1(\omega)$ и $f_2(\omega)$ [см. формулу (9) на стр. 472] соотношениями

$$F_1(z) = i^n f_1(-iz), \qquad F_2(z) = i^{n-1} f_2(-iz).$$

Поэтому уравнение $d(z) = 0$ совпадает с уравнением $f_{m+1}(-iz) = 0$, где многочлен $f_{m+1}(\omega)$ — наибольший общий делитель многочленов $f_1(z)$ и $f_2(z)$ — определяется последней строкой схемы Рауса. Следовательно, в силу формул (31) уравнение $d(z) = 0$ может быть записано в виде

$$\sum_{j=0}^{\left[\frac{n-m}{2}\right]} \Delta_m^{(m+j)} z^{n-m-2j} = 0,$$

где $\Delta_m^{(m+j)} = H\begin{pmatrix} 1 & \ldots & m-1 & m \\ 1 & \ldots & m-1 & m+j \end{pmatrix}$ $(j = 0, 1, \ldots)$ — побочные определители Гурвица.

[2]) Из (42) следует, что при h нечетном $\operatorname{sign} \Delta_{p+2} = (-1)^{\frac{h+1}{2}} \operatorname{sign} a_0$, а при h четном $\operatorname{sign} \Delta_{p+1} = (-1)^{\frac{h}{2}}$.

в которой при подсчете величины V для каждой группы подряд идущих p нулевых определителей (p — всегда нечетное число!)

$$(\Delta_s \neq 0) \quad \Delta_{s+1} = \ldots = \Delta_{s+p} = 0 \quad (\Delta_{s+p+1} \neq 0)$$

следует положить:

$$V\left(\frac{\Delta_s}{\Delta_{s-1}}, \frac{\Delta_{s+1}}{\Delta_s}, \ldots, \frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}\right) = h + \frac{1-(-1)^h \varepsilon}{2}, \tag{50}$$

где

$$p = 2h-1 \quad \text{и} \quad \varepsilon = \operatorname{sign}\left(\frac{\Delta_s}{\Delta_{s-1}} \frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}\right)^{1)}.$$

§ 9. Метод квадратичных форм. Определение числа различных вещественных корней многочлена

Раус получил свой алгоритм, применяя теорему Штурма к вычислению индекса Коши правильной рациональной дроби специального типа [см. формулу (11) на стр. 473]. У этой дроби из двух многочленов — числителя и знаменателя — один содержит только четные, а другой только нечетные степени аргумента z.

В настоящем параграфе и в последующих параграфах мы изложим более глубокий и более перспективный метод квадратичных форм Эрмита в применении к проблеме Рауса — Гурвица. При помощи этого метода мы получим выражение для индекса произвольной рациональной дроби через коэффициенты числителя и знаменателя. Метод квадратичных форм позволяет применить к проблеме Рауса — Гурвица результаты тонких исследований Фробениуса по теории ганкелевых форм (гл. X, § 10) и установить тесную связь некоторых замечательных теорем П. Л. Чебышева и А. А. Маркова с задачей устойчивости.

Мы познакомим читателя с методом квадратичных форм сначала на сравнительно простой задаче определения числа различных вещественных корней многочлена.

При решении этой задачи мы можем ограничиться случаем, когда $f(z)$ — вещественный многочлен. Действительно, пусть дан комплексный многочлен $f(z) = u(z) + iv(z)$ [$u(z)$ и $v(z)$ — вещественные многочлены]. Каждый вещественный корень многочлена $f(z)$ обращает в нуль одновременно и $u(z)$ и $v(z)$. Поэтому комплексный многочлен $f(z)$ имеет те же вещественные корни, что и вещественный многочлен $d(z)$, являющийся наибольшим общим делителем многочленов $u(z)$ и $v(z)$.

Итак, пусть $f(z)$ — вещественный многочлен с различными корнями $\alpha_1, \alpha_2, \ldots, \alpha_q$ соответственно кратностей $n_1, n_2, \ldots, n_q$:

$$f(z) = a_0 (z-\alpha_1)^{n_1} (z-\alpha_2)^{n_2} \ldots$$

$$\ldots (z-\alpha_q)^{n_q} [a_0 \neq 0; \quad \alpha_i \neq \alpha_k \quad \text{при } i \neq k \quad (i, k = 1, 2, \ldots, q)].$$

Введем в рассмотрение суммы Ньютона

$$s_p = \sum_{j=1}^{q} n_j \alpha_j^p \qquad (p = 0, 1, 2, \ldots).$$

---

1) При $s=1$ отношение $\frac{\Delta_s}{\Delta_{s-1}}$ следует заменить на Δ_1, а при $s=0$ — на a_0.

При помощи этих сумм составим ганкелеву форму

$$S_n(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k,$$

где n — любое целое число $\geqslant q$.

Тогда имеет место следующая

Теорема 6. *Число всех различных корней многочлена $f(z)$ равно рангу, а число всех различных вещественных корней равно сигнатуре формы $S_n(x, x)$.*

Доказательство. Из определения формы $S_n(x, x)$ непосредственно вытекает следующее ее представление:

$$S_n(x, x) = \sum_{j=1}^{q} n_j (x_0 + \alpha_j x_1 + \alpha_j^2 x_2 + \ldots + \alpha_j^{n-1} x_{n-1})^2. \qquad (51)$$

Здесь каждому корню α_j многочлена $f(z)$ соответствует квадрат линейной формы $Z_j = x_0 + \alpha_j x_1 + \ldots + \alpha_j^{n-1} x_{n-1}$ $(j = 1, 2, \ldots, q)$. Формы Z_1, Z_2, $\ldots$, Z_q линейно независимы между собой, так как коэффициенты этих линейных форм образуют матрицу Вандермонда $\| \alpha_j^h \|$, ранг которой равен числу различных α_j, т. е. q. Следовательно (см. стр. 269), ранг r формы $S_n(x, x)$ равен q.

В представлении (51) каждому вещественному корню α_j отвечает положительный квадрат. Каждой паре комплексно сопряженных корней α_j и $\bar{\alpha}_j$ отвечают две комплексно сопряженные формы:

$$Z_j = P_j + iQ_j, \qquad \bar{Z}_j = P_j - iQ_j;$$

соответствующие слагаемые в (51) в сумме дают один положительный и один отрицательный квадрат:

$$n_j Z_j^2 + n_j \bar{Z}_j^2 = 2n_j P_j^2 - 2n_j Q_j^2.$$

Отсюда легко усмотреть[1]), что сигнатура формы $S_n(x, x)$, т. е. разность между числом положительных и числом отрицательных квадратов, равна числу различных вещественных α_j.

Теорема доказана.

Из доказанной теоремы вытекает, что *все формы*

$$S_n(x, x) \qquad (n = q, q+1, \ldots)$$

имеет один и тот же ранг и одну и ту же сигнатуру.

Применяя теорему 6 к определению числа различных вещественных корней, возьмем в качестве n степень многочлена $f(z)$. Используя установленное в главе X (стр. 275) правило определения сигнатуры квадратичной формы, мы получаем

Следствие. *Число различных вещественных корней вещественного многочлена $f(z)$ равно избытку числа постоянств знака над числом перемен знака в ряду чисел*

$$1, \; s_0, \; \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix}, \; \ldots, \; \begin{vmatrix} s_0 & s_1 & \ldots & s_{r-1} \\ s_1 & s_2 & \ldots & s_r \\ \ldots & \ldots & \ldots & \ldots \\ s_{r-1} & s_r & \ldots & s_{2r-2} \end{vmatrix}, \qquad (52)$$

[1]) Квадратичная форма $S_n(x, x)$ представлена в виде суммы (алгебраической) q квадратов вещественных форм Z_j (для вещественных λ_j), P_j и Q_j (для комплексных λ_j). Эти формы независимы, так как число q равняется рангу r формы $S_n(x, x)$.

где s_p $(p=0, 1, 2, \ldots)$ — *суммы Ньютона для многочлена* $f(z)$, *а* r — ранг ганкелевой формы $S_n(x, x)=\sum_{i, k=0}^{n-1} s_{i+k}x_ix_k$ [n — *степень многочлена* $f(z)$].

Сформулированное таким образом правило для определения числа различных вещественных корней непосредственно применимо лишь в случае, когда все числа в ряду (52) отличны от нуля. Однако, поскольку здесь идет речь о вычислении сигнатуры ганкелевой квадратичной формы, то на основе результатов главы X, § 10 это правило с надлежащими уточнениями применяется в самом общем случае (более подробно об этом см. § 11 этой главы).

Число различных вещественных корней вещественного многочлена $f(z)$ равно индексу $I_{-\infty}^{+\infty}\frac{f'(z)}{f(z)}$ (см. стр. 470). Поэтому следствие из теоремы 6 дает нам формулу

$$I_{-\infty}^{+\infty}\frac{f'(z)}{f(z)}=r-2V\left(1,\ s_0,\ \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix},\ \ldots,\ \begin{vmatrix} s_0 & s_1 & \ldots & s_{r-1} \\ s_1 & s_2 & \ldots & s_r \\ \ldots & \ldots & \ldots & \ldots \\ s_{r-1} & s_r & \ldots & s_{2r-2} \end{vmatrix}\right).$$

В § 11 мы установим аналогичную формулу для индекса произвольной рациональной дроби. Необходимые для этого сведения о бесконечных ганкелевых матрицах будут даны в следующем параграфе.

§ 10. Бесконечные ганкелевы матрицы конечного ранга

1. Пусть дана последовательность комплексных чисел

$$s_0,\ s_1,\ s_2,\ \ldots$$

Эта последовательность чисел определяет бесконечную симметрическую матрицу

$$S=\begin{Vmatrix} s_0 & s_1 & s_2 & \ldots \\ s_1 & s_2 & s_3 & \ldots \\ s_2 & s_3 & s_4 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{Vmatrix},$$

которую называют обычно *ганкелевой*. Наряду с бесконечными ганкелевыми матрицами рассматриваются конечные ганкелевы матрицы $S_n=\|s_{i+k}\|_0^{n-1}$ и связанные с ними ганкелевы формы

$$S_n(x, x)=\sum_{i, k=0}^{n-1} s_{i+k}x_ix_k.$$

Последовательные главные миноры матрицы S будем обозначать через $D_1,\ D_2,\ D_3,\ \ldots$:

$$D_p=|s_{i+k}|_0^{p-1} \qquad (p=1, 2, \ldots).$$

Бесконечные матрицы могут быть конечного и бесконечного ранга. В последнем случае в этих матрицах существуют отличные от нуля миноры сколь угодно большого порядка. Следующая теорема дает необходимое и достаточное условие, которому должна удовлетворять последовательность чисел $s_0,\ s_1,\ s_2,\ \ldots$ для того, чтобы порождаемая ею бесконечная ганкелева матрица $S=\|s_{i+k}\|_0^{\infty}$ имела конечный ранг.

Теорема 7. *Бесконечная матрица* $S=\|s_{i+k}\|_0^\infty$ *имеет конечный ранг* r *тогда и только тогда, когда существует* r *чисел* $\alpha_1, \alpha_2, \ldots, \alpha_r$ *таких, что*

$$s_q=\sum_{g=1}^{r}\alpha_g s_{q-g} \quad (q=r,\ r+1,\ \ldots), \tag{53}$$

и r *есть наименьшее число, обладающее этим свойством.*

Доказательство. Если матрица $S=\|s_{i+k}\|_0^\infty$ имеет конечный ранг r, то первые $r+1$ строк $\Gamma_1, \Gamma_2, \ldots, \Gamma_{r+1}$ этой матрицы линейно зависимы. Поэтому существует число $h\leqslant r$ такое, что строки $\Gamma_1, \Gamma_2, \ldots, \Gamma_h$ линейно назависимы, а строка Γ_{h+1} есть линейная комбинация этих строк:

$$\Gamma_{h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{h-g+1}.$$

Рассмотрим строки $\Gamma_{q+1}, \Gamma_{q+2}, \ldots, \Gamma_{q+h+1}$, где q — любое целое неотрицательное число. Из структуры матрицы S непосредственно видно, что строки $\Gamma_{q+1}, \Gamma_{q+2}, \ldots, \Gamma_{q+h+1}$ получаются из строк $\Gamma_1, \Gamma_2, \ldots, \Gamma_{h+1}$ «укорочением», отбрасыванием элементов, стоящих в первых q столбцах. Поэтому

$$\Gamma_{q+h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{q+h-g+1} \quad (q=0,\ 1,\ 2,\ \ldots).$$

Таким образом, в матрице S любая строка, начиная с $(h+1)$-й, выражается линейно через h предыдущих и, следовательно, выражается линейно через h линейно независимых первых строк. Отсюда следует, что *для матрицы* S ранг $r=h$[1]). Линейная зависимость $\Gamma_{q+h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{q+h-g+1}$ после замены h на r в более подробной записи дает (53).

Обратно, если выполняется условие (53), то в матрице S любая строка (столбец) является линейной комбинацией первых r строк (столбцов). Поэтому все миноры матрицы S, порядок которых $>r$, равны нулю, и матрица S имеет конечный ранг $\leqslant r$. Но этот ранг не может быть $<r$, так как тогда, как было уже показано, имели бы место соотношения вида (53) при меньшем значении r, а это противоречит условию 2). Таким образом, теорема доказана полностью.

Следствие. *Если бесконечная ганкелева матрица* $S=\|s_{i+k}\|_0^\infty$ *имеет конечный ранг* r, *то*

$$D_r=|s_{i+k}|_0^{r-1}\neq 0.$$

Действительно, из соотношений (53) следует, что любая строка (столбец) матрицы S есть линейная комбинация первых r строк (столбцов). Поэтому любой минор r-го порядка матрицы S может быть представлен в виде αD_r, где α — некоторое число. Отсюда следует неравенство $D_r\neq 0$.

[1]) Положение «число линейно независимых строк в прямоугольной матрице равно рангу этой матрицы» справедливо не только для конечных, но и для бесконечных строк.

Примечание. Для конечных ганкелевых матриц ранга r неравенство $D_r \neq 0$ может не иметь места. Так, например, матрица $S_2 = \begin{Vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{Vmatrix}$ при $s_0 = s_1 = 0$, $s_2 \neq 0$ имеет ранг 1, в то время как $D_1 = s_0 = 0$.

2. Выясним замечательные взаимные связи между бесконечными ганкелевыми матрицами и рациональным функциями.

Пусть дана правильная рациональная дробная функция

$$R(z) = \frac{g(z)}{h(z)},$$

где

$$h(z) = a_0 z^m + \ldots + a_m \,(a_0 \neq 0), \quad g(z) = b_1 z^{m-1} + b_2 z^{m-2} + \ldots + b_m.$$

Напишем разложение $R(z)$ в степенной ряд по отрицательным степеням z:

$$R(z) = \frac{g(z)}{h(z)} = \frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \ldots$$

Если все полюсы функции $R(z)$, т. е. все значения z, при которых $R(z)$ обращается в бесконечность, лежат в круге $|z| \leqslant a$, то ряд, стоящий в правой части разложения, сходится при $|z| > a$. Обе части последнего равенства помножим на знаменатель $h(z)$:

$$(a_0 z^m + a_1 z^{m-1} + \ldots + a_m)\left(\frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \ldots\right) = b_1 z^{m-1} + b_2 z^{m-2} + \ldots + b_m.$$

Приравнивая между собой коэффициенты при одинаковых степенях z в обеих частях этого тождества, получим следующую систему соотношений:

$$\left.\begin{aligned} a_0 s_0 &= b_1, \\ a_0 s_1 + a_1 s_0 &= b_2, \\ &\cdots\cdots\cdots \\ a_0 s_{m-1} + a_1 s_{m-2} + \ldots + a_{m-1} s_0 &= b_m, \end{aligned}\right\} \tag{54}$$

$$a_0 s_q + a_1 s_{q-1} + \ldots + a_m s_{q-m} = 0 \qquad (q = m,\ m+1,\ \ldots). \tag{54'}$$

Полагая

$$\alpha_g = -\frac{a_g}{a_0} \qquad (g = 1,\ 2,\ \ldots,\ m),$$

мы можем соотношения (54′) записать в виде (53) (при $r = m$). Следовательно, согласно теореме 7 построенная при помощи коэффициентов s_0, s_1, s_2, ... бесконечная ганкелева матрица

$$S = \| s_{i+k} \|_0^\infty$$

имеет конечный ранг ($\leqslant m$).

Обратно, если матрица $S = \| s_{i+k} \|_0^\infty$ имеет конечный ранг r, то имеют место соотношения (53), которые могут быть переписаны в виде (54′) (при $m = r$). Тогда, определяя числа b_1, b_2, ..., b_m равенствами (54), будем иметь разложение

$$\frac{b_1 z^{m-1} + \ldots + b_m}{a_0 z^m + a_1 z^{m-1} + \ldots + a_m} = \frac{s_0}{z} + \frac{s_1}{z^2} + \ldots \tag{54''}$$

Наименьшая степень знаменателя m, при которой имеет место это разложение, совпадает с наименьшим числом m, при котором имеют место соотношения (53). По теореме 7 это наименьшее значение m равно рангу матрицы $S = \| s_{i+k} \|_0^\infty$.

При этом значении m рациональная дробь, стоящая в левой части равенства (54″), является несократимой.

Таким образом, нами доказана следующая теорема:

Теорема 8. *Матрица* $S = \| s_{i+k} \|_0^\infty$ *имеет конечный ранг в том и только в том случае, когда сумма ряда*

$$R(z) = \frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \dots$$

есть рациональная функция переменной z. *В этом случае ранг матрицы* S *совпадает с числом полюсов функции* $R(z)$, *считая каждый полюс столько раз, какова его кратность.*

§ 11. Определение индекса произвольной рациональной дроби через коэффициенты числителя и знаменателя

1. Пусть дана произвольная рациональная функция. Напишем ее разложение в ряд по нисходящим степеням z [1]):

$$R(z) = s_{-u-1}z^u + \dots + s_{-2}z + s_{-1} + \frac{s_0}{z} + \frac{s_1}{z^2} + \dots \qquad (55)$$

Последовательность коэффициентов при отрицательных степенях z

$$s_0,\ s_1,\ s_2,\ \dots$$

определяет бесконечную ганкелеву матрицу $S = \| s_{i+k} \|_0^\infty$.

Таким образом, устанавливается соответствие

$$R(z) \sim S.$$

Очевидно, что двум рациональным функциям, разность между которыми есть целая функция, отвечает одна и та же матрица S. Однако не всякая матрица $S = \| s_{i+k} \|_0^\infty$ соответствует рациональной функции. В предыдущем параграфе было установлено, что матрица S тогда и только тогда соответствует рациональной функции, когда эта бесконечная матрица имеет конечный ранг. Этот ранг равен числу полюсов (с учетом кратностей) функции $R(z)$, т. е. равен степени знаменателя $f(z)$ в несократимой дроби $\frac{g(z)}{f(z)} = R(z)$. С помощью разложения (55) устанавливается взаимно однозначное соответствие между правильными рациональными функциями $R(z)$ и ганкелевыми матрицами $S = \| s_{i+k} \|_0^\infty$ конечного ранга.

Отметим некоторые свойства соответствия:

1° Если $R_1(z) \sim S_1$, $R_2(z) \sim S_2$, то при любых числах c_1, c_2

$$c_1R_1(z) + c_2R_2(z) \sim c_1S_1 + c_2S_2.$$

В дальнейшем нам придется встретиться со случаем, когда коэффициенты числителя и знаменателя $R(z)$ будут целыми рациональными функциями параметра α; тогда и R будет рациональной функцией от z и α. Из разложения (54) следует, что в этом случае и числа $s_0, s_1, s_2, \dots$,

[1]) Ряд (55) сходится вне любого круга (с центром в точке $z = 0$), содержащего все полюсы функции $R(z)$.

т. е. элементы матрицы S, будут рационально зависеть от α. Дифференцируя по α почленно разложение (55), получим:

2° Если $R(z, \alpha) \sim S(\alpha)$, то $\frac{\partial R}{\partial \alpha} \sim \frac{\partial S}{\partial \alpha}$ [1]).

2. Напишем разложение $R(z)$ на простейшие дроби:

$$R(z) = Q(z) + \sum_{i=1}^{q} \left\{ \frac{A_1^{(j)}}{z-\alpha_j} + \frac{A_2^{(j)}}{(z-\alpha_j)^2} + \ldots + \frac{A_{\nu_j}^{(j)}}{(z-\alpha_j)^{\nu_j}} \right\}, \tag{56}$$

где $Q(z)$ — многочлен, и покажем, как по числам α и A построить матрицу S, соответствующую рациональной функции $R(z)$.

Для этого рассмотрим сначала простейшую рациональную дробь

$$\frac{1}{z-\alpha} = \sum_{p=0}^{\infty} \frac{\alpha^p}{z^{p+1}}.$$

Ей отвечает матрица

$$S_\alpha = \| \alpha^{i+k} \|_0^\infty.$$

Соответствующая этой матрице форма $S_{\alpha n}(x, x)$ имеет вид

$$S_{\alpha n}(x, x) = \sum_{i, k=0}^{n-1} \alpha^{i+k} x_i x_k = (x_0 + \alpha x_1 + \ldots + \alpha^{n-1} x_{n-1})^2.$$

Если

$$R(z) = Q(z) + \sum_{j=1}^{q} \frac{A^{(j)}}{z-\alpha_j},$$

то в силу 1° соответствующая матрица S определится по формуле

$$S = \sum_{j=1}^{q} A^{(j)} S_{\alpha_j} = \| \sum_{j=1}^{q} A^{(j)} \alpha_j^{i+k} \|_0^\infty,$$

а соответствующие квадратичные формы имеют вид

$$S_n(x, x) = \sum_{j=1}^{q} A^{(j)} (x_0 + \alpha_j x_1 + \ldots + \alpha_j^{n-1} x_{n-1})^2.$$

Для того чтобы перейти к общему случаю (56), мы предварительно $h-1$ раз продифференцируем почленно соотношение

$$\frac{1}{z-\alpha} \sim S_\alpha = \| \alpha^{i+k} \|_0^\infty.$$

Получим согласно 1° и 2° [2]):

$$\frac{1}{(z-\alpha)^h} \sim \frac{1}{(h-1)!} \frac{\partial^{h-1} S_\alpha}{\partial \alpha^{h-1}} = \| C_{i+k}^{h-1} \alpha^{i+k-h+1} \|_0^\infty \quad (C_{i+k}^{h-1} = 0 \text{ при } i+k < h-1).$$

Поэтому, пользуясь снова правилом 1°, в общем случае, когда для $R(z)$ имеет место разложение (56), находим:

$$R(z) \sim S = \sum_{j=1}^{q} \left(A_1^{(j)} + A_2^{(j)} \frac{\partial}{\partial \alpha_j} + \ldots + \frac{1}{(\nu_j - 1)!} A_{\nu_j}^{(j)} \frac{\partial^{\nu_j - 1}}{\partial \alpha_j^{\nu_j - 1}} \right) S_{\alpha_j}. \tag{57}$$

[1]) Если $S = \| s_{i+k} \|_0^\infty$, то $\frac{\partial S}{\partial \alpha} = \left\| \frac{\partial s_{i+k}}{\partial \alpha} \right\|_0^\infty$.

[2]) Здесь и ниже C_d^h означает число сочетаний из d по h.

Выполняя дифференцирование, получим:

$$S=\Big\|\sum_{j=1}^{q} A_1^{(j)}\alpha_j^{i+k}+A_2^{(j)}C_{i+k}^{1}\alpha_j^{i+k-1}+\ldots+A_{\nu_j}^{(j)}C_{i+k}^{\nu_j-1}\alpha_j^{i+k-\nu_j+1}\Big\|_0^{\infty}. \quad (57')$$

Соответствующая ганкелева форма $S_n(x, x)=\sum\limits_{i,k=0}^{n-1} s_{i+k}x_ix_k$ будет равна

$$S_n(x, x)=\sum_{j=1}^{q}\left(A_1^{(j)}+A_2^{(j)}\frac{\partial}{\partial\alpha_j}+\ldots+\frac{1}{(\nu_j-1)!}A_{\nu_j}^{(j)}\frac{\partial^{\nu_j-1}}{\partial\alpha_j^{\nu_j-1}}\right)\times$$

$$\times(x_0+\alpha_jx_1+\ldots+\alpha_j^{n-1}x_{n-1})^2. \quad (57'')$$

3. Теперь мы имеем возможность сформулировать и доказать основную теорему [1]):

Теорема 9. *Если*

$$R(z)\sim S$$

и m — ранг матрицы S [2]), то индекс Коши $I_{-\infty}^{+\infty}R(z)$ равен сигнатуре [3]) формы $S_n(x, x)$ при любом $n\geqslant m$:

$$I_{-\infty}^{+\infty}R(z)=\sigma[S_n(x, x)].$$

Доказательство. Пусть имеет место разложение (56). Тогда согласно (57)

$$S=\sum_{j=1}^{q}T_{\alpha_j},$$

где каждое слагаемое имеет вид

$$T_\alpha=\left(A_1+A_2\frac{\partial}{\partial\alpha}+\ldots+\frac{1}{(\nu-1)!}A_\nu\frac{\partial^{\nu-1}}{\partial\alpha^{\nu-1}}\right)S_\alpha,\quad S_\alpha=\|\alpha^{i+k}\|_0^{\infty} \quad (58)$$

и

$$S_n(x, x)=\sum_{j=1}^{q}T_{\alpha_j}(x, x)=\sum_{\alpha_j\text{вещ.}}T_{\alpha_j}(x, x)+\sum_{\alpha_j\text{компл.}}[T_{\alpha_j}(x, x)+T_{\bar\alpha_j}(x, x)].$$

Согласно теореме 8 ранг матрицы T_{α_j} и, следовательно, ранг формы $T_{\alpha_j}(x, x)$ равен $\nu_j\,(j=1, 2, \ldots, q)$, а ранг $S_n(x, x)$ равен $m=\sum\limits_{j=1}^{q}\nu_j$. Но если ранг суммы нескольких вещественных квадратичных форм равен сумме рангов слагаемых форм, то такое же соотношение имеет место и для сигнатур:

$$\sigma[S_n(x, x)]=\sum_{\alpha_j\text{вещ.}}\sigma[T_{\alpha_j}(x, x)]+\sum_{\alpha_j\text{компл.}}\sigma[T_{\alpha_j}(x, x)+T_{\bar\alpha_j}(x, x)]. \quad (59)$$

Рассмотрим раздельно два случая:

1) α вещественно. При любой вариации параметров $A_1, A_2, \ldots, A_{\nu-1}$ и α в

$$\frac{A_1}{z-\alpha}+\frac{A_2}{(z-\alpha)^2}+\ldots+\frac{A_\nu}{(z-\alpha)^\nu}\qquad (A_\nu\neq 0) \quad (60)$$

---

[1]) Эта теорема была доказана Эрмитом в 1856 г. для простейшего случая, когда $R(z)$ не имеет кратных полюсов [195]. В общем случае эта теорема была доказана Гурвицем [195] (см. также [16], стр. 17—19). Приведенное в тексте доказательство отличается от доказательства Гурвица.

[2]) Как мы уже отмечали, m равно степени знаменателя в несократимом представлении рациональной дроби $R(z)$ (см. теорему 8 на стр. 498).

[3]) Сигнатуру формы $S_n(x, x)$ будем обозначать через $\sigma[S_n(x, x)]$.

ранг соответствующей матрицы T_α будет оставаться неизменным ($=\nu$); следовательно, будет оставаться неизменной и сигнатура формы $T_\alpha(x, x)$ (см. стр. 280). Поэтому $\sigma[T_\alpha(x, x)]$ не изменится, если мы в (59) и (60) положим: $A_1 = \ldots = A_{\nu-1} = 0$ и $\alpha = 0$, т. е. вместо T_α возьмем матрицу

$$\left[\frac{1}{(\nu-1)!}\frac{\partial^{\nu-1}S_\alpha}{\partial\alpha^{\nu-1}}\right]_{\alpha=0} = \left\| \begin{array}{ccccccc} & & & \overbrace{\quad}^{\nu-1} & & & \\ 0 & 0 & \ldots & 0 & A_\nu & 0 & 0 \ldots \\ 0 & & & & \ddots & & \\ \cdot & & & & \ddots & & \\ \cdot & & & & \ddots & & \\ \cdot & & & & \ddots & & \\ 0 & & & & \ddots & & \\ A_\nu & & & & \ddots & & \\ 0 & & & & \ddots & & \\ 0 & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \end{array} \right\|.$$

Соответствующая квадратичная форма равна

$$\begin{array}{ll} 2A_\nu(x_0x_{\nu-1} + x_1x_{\nu-2} + \ldots + x_{s-1}x_s) & \text{при } \nu = 2s, \\ A_\nu[2(x_0x_{\nu-1} + \ldots + x_{s-2}x_s) + x_{s-1}^2] & \text{при } \nu = 2s-1 \end{array} \qquad (s = 1, 2, 3, \ldots).$$

Но сигнатура верхней формы всегда равна нулю, а сигнатура *нижней* формы равна $\operatorname{sign} A_\nu$[1]). Таким образом, если α вещественно, то

$$\sigma[T_\alpha(x, x)] = \begin{cases} 0 & \text{при } \nu \text{ четном}, \\ \operatorname{sign} A_\nu & \text{при } \nu \text{ нечетном}. \end{cases} \qquad (61)$$

2) α — комплексное число. Пусть

$$T_\alpha(x, x) = \sum_{k=1}^{\nu}(P_k + iQ_k)^2, \qquad T_{\bar\alpha}(x, x) = \sum_{k=1}^{\nu}(P_k - iQ_k)^2,$$

где P_k, Q_k ($k = 1, 2, \ldots, \nu$) — вещественные линейные формы переменных $x_0, x_1, \ldots, x_{n-1}$. Тогда

$$T_\alpha(x, x) + T_{\bar\alpha}(x, x) = 2\sum_{k=1}^{\nu}P_k^2 - 2\sum_{k=1}^{\nu}Q_k^2. \qquad (62)$$

Так как ранг этой квадратичной формы равен 2ν, то P_k, Q_k ($k = 1, 2, \ldots, \nu$) линейно независимы, и потому согласно (62) при невещественном α

$$\sigma[T_\alpha(x, x) + T_{\bar\alpha}(x, x)] = 0. \qquad (63)$$

Из (59), (61) и (63) вытекает:

$$\sigma[S_n(x, x)] = \sum_{\left(\substack{\alpha_j \text{вещ.} \\ \nu \text{ нечетно}}\right)} \operatorname{sign} A_\nu^{(j)}.$$

---

[1]) Каждое произведение $x_0x_{\nu-1}$, $x_1x_{\nu-2}$, ... можно заменить соответственно на разности квадратов $\left(\frac{x_0+x_{\nu-1}}{2}\right)^2 - \left(\frac{x_0-x_{\nu-1}}{2}\right)^2$, $\left(\frac{x_1+x_{\nu-2}}{2}\right)^2 - \left(\frac{x_1-x_{\nu-2}}{2}\right)^2$, ... Все получающиеся при этом квадраты независимы между собой.

Но на стр. 470 было выяснено, что сумма, стоящая в правой части этого равенства, равна $I_{-\infty}^{+\infty}R(z)$. Таким образом, теорема доказана.

Из доказанной теоремы вытекает:

Следствие 1. *Если* $R(z) \sim S = \| s_{i+k} \|_0^{\infty}$ *и* m *— ранг матрицы* S, *то все квадратичные формы* $S_n(x, x) = \sum_{i,k=0}^{n-1} s_{i+k}x_ix_k$ $(n = m, m+1, \ldots)$ *имеют одну и ту же сигнатуру.*

В главе X, § 10 (стр. 305, 306) было установлено правило вычисления сигнатуры ганкелевой квадратичной формы, причем исследования Фробениуса дали возможность сформулировать правило с охватом всех особых случаев. Согласно доказанной теореме этим правилом можно пользоваться для вычисления индекса Коши. Таким образом, получаем

Следствие 2. *Индекс произвольной рациональной функции* $R(z)$, *которой соответствует матрица* $S = \| s_{i+k} \|_0^{\infty}$ *ранга* m, *определяется по формуле*

$$I_{-\infty}^{+\infty}R(z) = m - 2V(1, D_1, D_2, \ldots, D_m), \tag{64}$$

где

$$D_f = |\, s_{i+k} \,|_0^{f-1} = \begin{vmatrix} s_0 & s_1 & \ldots & s_{f-1} \\ s_1 & s_2 & \ldots & s_f \\ \ldots & \ldots & \ldots & \ldots \\ s_{f-1} & s_f & \ldots & s_{2f-2} \end{vmatrix} \qquad (f = 1, 2, \ldots, m); \tag{65}$$

если среди определителей $D_1, D_2, \ldots, D_m$ *имеется группа подряд идущих определителей, равных нулю*[1])

$$(D_h \neq 0) \quad D_{h+1} = \ldots = D_{h+p} = 0 \quad (D_{h+p+1} \neq 0),$$

то при вычислении $V(D_h, D_{h+1}, \ldots, D_{h+p+1})$ *можно принять*:

$$\operatorname{sign} D_{h+j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} D_h \qquad (j = 1, 2, \ldots, p),$$

что дает:

$$V(D_h, D_{h+1}, \ldots, D_{h+p+1}) = \begin{cases} \dfrac{p+1}{2} \text{ при } p \text{ нечетном}, \\ \dfrac{p+1-\varepsilon}{2} \text{ при } p \text{ четном}, \ \varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \dfrac{D_{h+p+1}}{D_h}. \end{cases} \tag{66}$$

Для того чтобы выразить индекс рациональной функции через коэффициенты числителя и знаменателя, нам понадобятся некоторые вспомогательные соотношения.

Прежде всего всегда можно представить $R(z)$ в виде[2])

$$R(z) = Q(z) + \frac{g(z)}{h(z)},$$

---

[1]) При этом всегда $D_m \neq 0$ (стр. 496) и при p нечетном $\operatorname{sign} \dfrac{D_{h+p+1}}{D_h} = (-1)^{\frac{p+1}{2}}$ (стр. 309).

[2]) У нас нет необходимости заменять $R(z)$ правильной рациональной дробью. Для дальнейшего достаточно, чтобы степень $g(z)$ не превосходила степени $h(z)$.

где $Q(z)$, $g(z)$, $h(z)$ — многочлены, причем

$$h(z)=a_0z^m+a_1z^{m-1}+\ldots+a_m\,(a_0\neq 0),\quad g(z)=b_0z^m+b_1z^{m-1}+\ldots+b_m.$$

Очевидно,

$$I_{-\infty}^{+\infty}R(z)=I_{-\infty}^{+\infty}\frac{g(z)}{h(z)}.$$

Пусть

$$\frac{g(z)}{h(z)}=s_{-1}+\frac{s_0}{z}+\frac{s_1}{z^2}+\ldots$$

Тогда, освобождаясь здесь от знаменателя и после этого приравнивая друг другу коэффициенты при одинаковых степенях z в левой и правой частях равенства, получаем:

$$\begin{aligned}
&a_0s_{-1}=b_0,\\
&a_0s_0+a_1s_{-1}=b_1,\\
&\cdots\cdots\cdots\cdots\\
&a_0s_{m-1}+a_1s_{m-2}+\ldots+a_ms_{-1}=b_m,\\
&a_0s_t+a_1s_{t-1}+\ldots+a_ms_{t-m}=0\quad(t=m,\,m+1,\,\ldots).
\end{aligned}\tag{67}$$

Пользуясь соотношениями (67), находим выражение для следующего определителя $2p$-го порядка, в котором кладем $a_j=0$, $b_j=0$ при $j>m$:

$$\begin{vmatrix}
a_0 & a_1 & a_2 & \ldots & a_{2p-1}\\
b_0 & b_1 & b_2 & \ldots & b_{2p-1}\\
0 & a_0 & a_1 & \ldots & a_{2p-2}\\
0 & b_0 & b_1 & \ldots & b_{2p-2}\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{vmatrix}=
\begin{vmatrix}
1 & 0 & 0 & \ldots & 0\\
s_{-1} & s_0 & s_1 & \ldots & s_{2p-2}\\
0 & 1 & 0 & \ldots & 0\\
0 & s_{-1} & s_0 & \ldots & s_{2p-3}\\
 & \cdots & \cdots & \cdots & \\
 & \cdots & \cdots & \cdots &
\end{vmatrix}\cdot
\begin{vmatrix}
a_0 & a_1 & a_2 & \ldots & a_{2p-1}\\
0 & a_0 & a_1 & \ldots & a_{2p-2}\\
0 & 0 & a_0 & \ldots & a_{2p-3}\\
 & \cdots & \cdots & \cdots & \\
 & \cdots & \cdots & \cdots & \\
0 & 0 & 0 & \ldots & a_0
\end{vmatrix}=$$

$$=(-1)^{\frac{p(p-1)}{2}}a_0^{2p}\begin{vmatrix}
s_{p-1} & s_p & \ldots & s_{2p-2}\\
s_{p-2} & s_{p-1} & \ldots & s_{2p-3}\\
\cdots & \cdots & \cdots & \cdots\\
s_0 & s_1 & \ldots & s_{p-1}
\end{vmatrix}=a_0^{2p}\begin{vmatrix}
s_0 & s_1 & \ldots & s_{p-1}\\
s_1 & s_2 & \ldots & s_p\\
\cdots & \cdots & \cdots & \cdots\\
s_{p-1} & s_p & \ldots & s_{2p-2}
\end{vmatrix}=a_0^{2p}D_p.\tag{68}$$

Введем сокращенное обозначение

$$\nabla_{2p}=\begin{vmatrix}
a_0 & a_1 & \ldots & a_{2p-1}\\
b_0 & b_1 & \ldots & b_{2p-1}\\
0 & a_0 & \ldots & a_{2p-2}\\
0 & b_0 & \ldots & b_{2p-2}\\
\cdots & \cdots & \cdots & \cdots
\end{vmatrix}\qquad(p=1,\,2,\,\ldots;\ a_j=b_j=0\ \text{при}\ j>m).\tag{69}$$

Тогда формула (68) запишется так:

$$\nabla_{2p}=a_0^{2p}D_p\qquad(p=1,\,2,\,\ldots).\tag{68'}$$

В силу этой формулы следствие 2 на стр. 502 приводит нас к следующей теореме:

Теорема 10. *Если* $\nabla_{2m} \neq 0$ [1]), *то*

$$I_{-\infty}^{+\infty} \frac{b_0 z^m + b_1 z^{m-1} + \ldots + b_m}{a_0 z^m + a_1 z^{m-1} + \ldots + a_m} = m - 2V(1, \nabla_2, \nabla_4, \ldots, \nabla_{2m}) \quad (a_0 \neq 0), \quad (70)$$

где ∇_{2p} $(p = 1, 2, \ldots, m)$ *определяется формулой* (69); *если при этом имеются подряд идущие нулевые определители*

$$(\nabla_{2h} \neq 0) \quad \nabla_{2h+2} = \ldots = \nabla_{2h+2p} = 0 \quad (\nabla_{2h+2p+2} \neq 0),$$

то в формуле (70) *при подсчете* $V(\nabla_{2h}, \nabla_{2h+2}, \ldots, \nabla_{2h+2p+2})$ *следует положить*:

$$\operatorname{sign} \nabla_{2h+2j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} \nabla_{2h} \qquad (j = 1, 2, \ldots, p),$$

или, что то же [2]),

$$V(\nabla_{2h}, \ldots, \nabla_{2h+2p+2}) = \begin{cases} \dfrac{p+1}{2} \text{ при } p \text{ нечетном}, \\ \dfrac{p+1-\varepsilon}{2} \text{ при } p \text{ четном}, \ \varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \dfrac{\nabla_{2h+2p+2}}{\nabla_{2h}} \end{cases}$$

Замечание. *Если* $\nabla_{2m} = 0$, *т. е. дробь, стоящая под знаком индекса в формуле* (70), *сократима, то формулу* (70) *следует заменить формулой*

$$I_{-\infty}^{+\infty} \frac{b_0 z^m + b_1 z^{m-1} + \ldots + b_m}{a_0 z^m + a_1 z^{m-1} + \ldots + a_m} = r - 2V(1, \nabla_2, \nabla_4, \ldots, \nabla_{2r}), \quad (70')$$

где r — *число полюсов (с учетом кратностей) рациональной дроби, стоящей под знаком индекса (т. е.* r — *степень знаменателя после сокращения дроби). Здесь* $\nabla_{2r} \neq 0$.

Действительно, если $\nabla_{2m} = 0$, то интересующий нас индекс равен

$$r - 2V(1, D_1, D_2, \ldots, D_r),$$

так как число r является рангом соответствующей матрицы $S = \| s_{i+k} \|_0^\infty$. Но равенство (68′) имеет формальный характер и оно справедливо и для сократимой дроби [3]). Поэтому

$$V(1, D_1, D_2, \ldots, D_r) = V(1, \nabla_2, \nabla_4, \ldots, \nabla_{2r}).$$

и мы приходим к формуле (70′).

Формула (70′) дает возможность выразить индекс любой рациональной дроби, у которой степень числителя не превышает степени знаменателя, через коэффициенты числителя и знаменателя.

§ 12. Второе доказательство теоремы Рауса—Гурвица

В § 6 мы доказали теорему Рауса—Гурвица, опираясь на теорему Штурма и алгоритм Рауса. В этом параграфе мы дадим доказательство теоремы Рауса—Гурвица, основанное на теореме 10 § 11 и на свойствах индексов Коши.

---

[1]) Условие $\nabla_{2m} \neq 0$ означает, что $D_m \neq 0$ и что, следовательно, дробь, стоящая в (70) под знаком индекса, несократима.

[2]) При p нечетном $\dfrac{\nabla_{2h+2p+2}}{\nabla_{2h}} = \operatorname{sign} \dfrac{D_{h+p+1}}{D_h} = (-1)^{\frac{p+1}{2}}$ (см. сноску [1]) на стр. 502.

[3]) Из равенства (68′) следует, что значения ∇_{2p} определяются рациональной дробью $R(z)$ (точнее, ее правильной частью), а не числителем и знаменателем в отдельности. Поэтому при сокращении дроби $g(z)/h(z)$ меняются элементы в каждом определителе ∇_{2p}, а величина его остается неизменной.

Отметим некоторые свойства индексов Коши, которые нам понадобятся в дальнейшем.

1° $I_a^b R(x) = -I_b^a R(x)$ [1]).

2° $I_a^b R_1(x)\, R(x) = \operatorname{sign} R_1(x)\, I_a^b R(x)$, *если* $R_1(x) \neq \begin{cases} 0 \\ \infty \end{cases}$ *внутри* (a, b).

3°. *Если* $a < c < b$, *то* $I_a^b R(x) = I_a^c R(x) + I_c^b R(x) + \eta_c$, *где* $\eta_c = 0$, *если* $R(c)$ *— конечная величина, и* $\eta_c = \pm 1$, *если в точке* c *функция* $R(x)$ *обращается в бесконечность; при этом* $\eta_c = +1$ *соответствует переходу в точке* c *от* $-\infty$ *к* $+\infty$ (*при возрастании* x), *а* $\eta_c = -1$ *— переходу от* $+\infty$ *к* $-\infty$.

4° *Если* $R(-x) = -R(x)$, *то* $I_{-a}^{0} R(x) = I_0^a R(x)$. *Если* $R(-x) = R(x)$, *то* $I_{-a}^{0} R(x) = -I_0^a R(x)$.

5° $I_a^b R(x) + I_a^b \frac{1}{R(x)} = \frac{\varepsilon_b - \varepsilon_a}{2}$, *где* ε_a *— знак* $R(x)$ *внутри* (a, b) *вблизи* a, ε_b *— знак* $R(x)$ *внутри* (a, b) *вблизи* b.

Первые четыре свойства непосредственно следуют из определения индекса Коши (см. § 2). Свойство 5° вытекает из того, что сумма индексов $I_a^b R(x) + I_a^b \frac{1}{R(x)}$ равна разности $n_1 - n_2$, где n_1 — число перемен знака $R(x)$ с переходом от отрицательных значений к положительным при изменении x от a до b, а n_2 — число перемен знака $R(x)$ с переходом от положительных к отрицательным значениям.

Рассмотрим вещественный многочлен [2])

$$f(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_{n-1} z + a_n \quad (a_0 > 0).$$

Мы его можем представить в виде

$$f(z) = h(z^2) + z g(z^2),$$

где

$$h(u) = a_n + a_{n-2} u + \ldots, \quad g(u) = a_{n-1} + a_{n-3} u + \ldots$$

Введем обозначение

$$\varrho = I_{-\infty}^{+\infty} \frac{a_1 z^{n-1} - a_3 z^{n-3} + \ldots}{a_0 z^n - a_2 z^{n-2} + \ldots}. \tag{71}$$

В § 3 мы показали [см. (20) на стр. 475], что

$$\varrho = n - 2k - s, \tag{72}$$

где k — число корней многочлена $f(z)$ с положительными вещественными частями, а s — число корней $f(z)$, расположенных на мнимой оси.

Преобразуем выражение (71) для ϱ.

Рассмотрим сначала случай четного n. Пусть $n = 2m$. Тогда

$$h(u) = a_0 u^m + a_2 u^{m-1} + \ldots + a_n, \quad g(u) = a_1 u^{m-1} + a_3 u^{m-2} + \ldots + a_{n-1}.$$

---

[1]) Здесь и далее нижний предел при индексе может равняться $-\infty$, а верхний предел $+\infty$.

[2]) Здесь мы, в отличие от § 3, возвращаемся к обычным обозначениям для коэффициентов многочлена.

Пользуясь свойствами 1°—4° и полагая $\eta = \pm 1$, если соответственно $\lim\limits_{u\to-0} \frac{g(u)}{h(u)} = \pm\infty$ и $\eta = 0$ в остальных случаях, будем иметь:

$$\varrho = -I_{-\infty}^{+\infty} \frac{zg(-z^2)}{h(-z^2)} = -(I_{-\infty}^{0} + I_{0}^{+\infty} + \eta) = -2I_{-\infty}^{0} \frac{zg(-z^2)}{h(-z^2)} - \eta =$$

$$= 2I_{-\infty}^{0} \frac{g(-z^2)}{h(-z^2)} - \eta = 2I_{-\infty}^{0} \frac{g(u)}{h(u)} - \eta = I_{-\infty}^{0} \frac{g(u)}{h(u)} - I_{-\infty}^{0} \frac{u\,g(u)}{h(u)} - \eta =$$

$$= I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} - I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)}.$$

Точно так же при n нечетном, $n = 2m+1$, имеем:

$$h(u) = a_1u^m + a_3u^{m-1} + \ldots + a_n, \quad g(u) = a_0u^m + a_2u^{m-1} + \ldots + a_{n-1}.$$

Полагая $\zeta = \operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{u=-0}$ [1]), если $\lim\limits_{u\to 0} \frac{g(u)}{h(u)} = 0$ и $\zeta = 0$ в остальных случаях, найдем:

$$\varrho = I_{-\infty}^{+\infty} \frac{h(-z^2)}{zg(-z^2)} = I_{-\infty}^{0} + I_{0}^{+\infty} + \zeta = 2I_{-\infty}^{0} \frac{h(-z^2)}{zg(-z^2)} + \zeta = 2I_{-\infty}^{0} \frac{h(u)}{ug(u)} + \zeta =$$

$$= I_{-\infty}^{0} \frac{h(u)}{ug(u)} - I_{-\infty}^{0} \frac{h(u)}{g(u)} + \zeta = I_{-\infty}^{+\infty} \frac{h(u)}{ug(u)} - I_{-\infty}^{+\infty} \frac{h(u)}{g(u)}.$$

Таким образом [2]),

$$\varrho = I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} - I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} \quad (n = 2m), \tag{73'}$$

$$\varrho = I_{-\infty}^{+\infty} \frac{h(u)}{ug(u)} - I_{-\infty}^{+\infty} \frac{h(u)}{g(u)} \quad (n = 2m+1). \tag{73''}$$

По-прежнему через $\Delta_1, \Delta_2, \ldots, \Delta_n$ будем обозначать определители Гурвица для данного многочлена $f(z)$. Примем, что $\Delta_n \neq 0$ [3]).

1) $n = 2m$. По формуле (70) [4])

$$I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} = m - 2V(1,\ \Delta_1,\ \Delta_3,\ \ldots,\ \Delta_{n-1}), \tag{74}$$

$$I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} = m - 2V(1,\ -\Delta_2,\ +\Delta_4,\ -\Delta_6,\ \ldots) =$$

$$= -m + 2V(1,\ \Delta_2,\ \Delta_4,\ \ldots,\ \Delta_n). \tag{75}$$

---

[1]) Здесь под $\operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{u=-0}$ мы понимаем знак $\frac{g(u)}{h(u)}$ при очень малом по абсолютной величине отрицательном u.

[2]) Если $a_1 \neq 0$, то две формулы (73') и (73'') можно объединить в одну формулу

$$\varrho = I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} + I_{-\infty}^{+\infty} \frac{h(u)}{ug(u)}. \tag{73'''}$$

[3]) В этом случае $s = 0$ и, следовательно, $\varrho = n - 2k$. Кроме того, $\Delta_n \neq 0$ означает, что дроби, стоящие под знаком индексов, в формулах (73'), (73'') несократимы.

[4]) При вычислении $\nabla_2, \nabla_4, \ldots, \nabla_{2m}$ величины $a_0, a_1, \ldots, a_m$ и $b_0, b_1, \ldots, b_m$ следует соответственно заменить при вычислении первого индекса на $a_0, a_2, \ldots, a_{2m}$ и $0, a_1, a_3, \ldots, a_{2m-1}$, а при вычислении второго индекса — на $a_0, a_2, \ldots, a_{2m}$ и $a_1, a_3, \ldots, a_{2m-1}, 0$.

Но тогда согласно (73′)

$$\varrho = n - 2V(1, \Delta_1, \Delta_3, \ldots, \Delta_{n-1}) - 2V(1, \Delta_2, \Delta_4, \ldots, \Delta_n),$$

что в соединении с равенством $\varrho = n - 2k$ дает:

$$k = V(1, \Delta_1, \Delta_3, \ldots, \Delta_{n-1}) + V(1, \Delta_2, \Delta_4, \ldots, \Delta_n). \tag{76}$$

2) $n = 2m + 1$. По формуле (70) [1]),

$$I_{-\infty}^{+\infty} \frac{h(u)}{ug(u)} = m + 1 - 2V(1, \Delta_1, \Delta_3, \ldots, \Delta_n), \tag{77}$$

$$I_{-\infty}^{+\infty} \frac{h(u)}{g(u)} = m - 2V(1, -\Delta_2, +\Delta_4, -\ldots) =$$

$$= -m + 2V(1, \Delta_2, \Delta_4, \ldots, \Delta_{n-1}). \tag{78}$$

Равенство $\varrho = 2m + 1 - 2k$ вместе с равенствами (73″), (77) и (78) дает нам снова формулу (76).

Теорема Рауса—Гурвица доказана (см. стр. 486).

Замечание 1. Если в формуле

$$k = V(1, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots)$$

некоторые промежуточные определители Гурвица равны нулю, то формула сохраняет силу и в этом случае, только в каждой группе подряд идущих нулевых определителей

$$(\Delta_l \neq 0)\ \Delta_{l+2} = \Delta_{l+4} = \ldots = \Delta_{l+2p} = 0\ (\Delta_{l+2p+2} \neq 0)$$

следует приписать этим определителям (в соответствии с теоремой 7) знаки

$$\operatorname{sign} \Delta_{l+2j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} \Delta_l \quad (j = 1, 2, \ldots, p),$$

что дает:

$$V(\Delta_l, \Delta_{l+2}, \ldots, \Delta_{l+2p+2}) = \begin{cases} \dfrac{p+1}{2} \text{ при } p \text{ нечетном,} \\ \dfrac{p+1-\varepsilon}{2} \text{ при } p \text{ четном, } \varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \dfrac{\Delta_{l+2p+2}}{\Delta_l}. \end{cases} \tag{79}$$

Внимательное сопоставление этого правила вычисления k при наличии нулевых определителей Гурвица с правилом, данным в теореме 5 (стр. 492), показывает, что оба правила совпадают [2]).

Замечание 2. Если $\Delta_n = 0$, то многочлены $ug(u)$ и $h(u)$ не являются взаимно простыми. Обозначим через $d(u)$ наибольший общий делитель многочленов $g(u)$ и $h(u)$, а через $u^\gamma d(u)$ — наибольший общий делитель $ug(u)$ и $h(u)$ ($\gamma = 0$ или 1). Степень $d(u)$ обозначим через δ и положим $h(u) = d(u) h_1(u)$ и $g(u) = d(u) g_1(u)$.

Несократимой рациональной дроби $\frac{g_1(u)}{h_1(u)}$ всегда соответствует некоторая бесконечная ганкелева матрица $S = \| s_{i+k} \|_0^\infty$ ранга r, где r — степень

---

[1]) Здесь при вычислении первого индекса в формуле (70) вместо $a_0, a_1, \ldots, a_{m+1}$ и $b_0, b_1, \ldots, b_{m+1}$ берем соответственно $a_0, a_2, \ldots, a_{2m}$, 0 и 0, $a_1, a_3, \ldots, a_{2m+1}$, а при вычислении второго индекса вместо $a_0, a_1, \ldots, a_m$ и $b_0, b_1, \ldots, b_m$ ставим: $a_1, a_3, \ldots, a_{2m+1}$ и $a_0, a_2, \ldots, a_{2m}$.

[2]) При этом следует учесть замечания, сделанные в сносках [1]), [2]) на стр. 492.

$h_1(u)$. При этом соответствующий определитель $D_r \neq 0$, а $D_{r+1} = D_{r+2} = \ldots = 0$. В силу формулы (68′) $\nabla_{2r} \neq 0$, $\nabla_{2r+2} = \nabla_{2r+4} = \ldots = 0$. Кроме того,

$$I_{-\infty}^{+\infty} \frac{g_1(u)}{h_1(u)} = r - 2V(1, \nabla_2, \ldots, \nabla_{2r}).$$

Применяя все это к дробям, стоящим под знаком индекса в (74), (75), (77) и (78), мы легко найдем, что при любом n (четном и нечетном) и $\varkappa = 2\delta + \gamma$

$$\Delta_{n-\varkappa-1} \neq 0, \quad \Delta_{n-\varkappa} \neq 0, \quad \overbrace{\Delta_{n-\varkappa+1} = \ldots = \Delta_n}^{\varkappa} = 0$$

и что все формулы (74), (75), (77) и (78) сохраняют свою силу и в рассматриваемом случае, если в правых частях этих формул опустить все Δ_i при $i > n - \varkappa$ и заменить число m [а в формуле (77) число $m+1$] на степень соответствующего знаменателя подиндексной дроби после ее сокращения. Тогда мы получим с учетом (73′) и (73″):

$$\varrho = n - \varkappa - 2V(1, \Delta_1, \Delta_3, \ldots) - 2V(1, \Delta_2, \Delta_4, \ldots).$$

Вместе с формулой $\varrho = n - 2k - s$ это дает:

$$k_1 = V(1, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots), \tag{80}$$

где $k_1 = k + \frac{s}{2} - \frac{\varkappa}{2}$ — число всех корней $f(z)$, лежащих в правой полуплоскости, за исключением тех, которые одновременно являются корнями и многочлена $f(-z)$ [1]).

§ 13. Некоторые дополнения к теореме Рауса—Гурвица. Критерий устойчивости Льенара и Шипара

Пусть дан многочлен с вещественными коэффициентами

$$f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n \quad (a_0 > 0).$$

Тогда условия Рауса—Гурвица, необходимые и достаточные для того, чтобы все корни многочлена $f(z)$ имели отрицательные действительные части, записываются в виде неравенств

$$\Delta_1 > 0, \quad \Delta_2 > 0, \ \ldots, \ \Delta_n > 0, \tag{81}$$

где

$$\Delta_i = \begin{Vmatrix} a_1 & a_3 & a_5 & \ldots \\ a_0 & a_2 & a_4 & \ldots \\ 0 & a_1 & a_3 & \ldots \\ 0 & a_0 & a_2 & a_4 \\ & & & \ddots \\ & & & \ddots \\ & & & \ddots \\ & & & a_i \end{Vmatrix} \quad (a_k = 0 \text{ при } k > n)$$

— определитель Гурвица i-го порядка ($i = 1, 2, \ldots, n$).

[1]) Это следует из того, что $\varkappa$ — степень наибольшего общего делителя многочленов $h(u)$ и $ug(u)$; $\varkappa$ — число «особых» корней многочлена $f(z)$, т. е. тех корней z^*, для которых $-z^*$ также является корнем $f(z)$. Число этих особых корней равно числу последних (включая Δ_n) подряд идущих нулевых определителей Гурвица

$$\Delta_{n-\varkappa+1} = \ldots = \Delta_n = 0.$$

Если условия (81) выполнены, то многочлен $f(z)$ представляется в виде произведения a_0 на множители вида $z+u$, z^2+vz+w ($u>0$, $v>0$, $w>0$), и потому все коэффициенты многочлена $f(z)$ положительны[1]):

$$a_1>0,\ a_2>0,\ \ldots,\ a_n>0. \tag{82}$$

В отличие от условий (81) условия (82) являются необходимыми, но отнюдь не достаточными для расположения всех корней $f(z)$ в левой полуплоскости $\operatorname{Re} z<0$.

Однако при выполнении условий (82) неравенства (81) уже не являются независимыми. Так, например, при $n=4$ условия Рауса—Гурвица приводятся к одному неравенству $\Delta_3>0$, при $n=5$—к двум: $\Delta_2>0$, $\Delta_4>0$, при $n=6$—к двум; $\Delta_3>0$, $\Delta_5>0$[2]).

Это обстоятельство было исследовано французскими математиками Льенаром и Шипаром и дало возможность им в 1914 г.[3]) установить критерий устойчивости, отличный от критерия Рауса—Гурвица.

Теорема 11 (Критерий Льенара и Шипара). *Необходимые и достаточные условия для того, чтобы вещественный многочлен* $f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n$ $(a_0>0)$ *имел все корни с отрицательными вещественными частями, могут быть записаны в любом из следующих четырех видов*[4]):

1) $a_n>0,\ a_{n-2}>0,\ \ldots;\ \Delta_1>0,\ \Delta_3>0,\ \ldots,$

2) $a_n>0,\ a_{n-2}>0,\ \ldots;\ \Delta_2>0,\ \Delta_4>0,\ \ldots,$

3) $a_n>0;\ a_{n-1}>0,\ a_{n-3}>0,\ \ldots;\ \Delta_1>0,\ \Delta_3>0,\ \ldots,$

4) $a_n>0;\ a_{n-1}>0,\ a_{n-3}>0,\ \ldots;\ \Delta_2>0,\ \Delta_4>0,\ \ldots.$

Из теоремы 11 вытекает, что для вещественного многочлена $f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n$ $(a_0>0)$, у которого все коэффициенты (или даже только часть a_n, a_{n-2}, ... или a_n, a_{n-1}, b_{n-3}, ...) положительны, детерминантные неравенства Гурвица (81) не являются независимыми, а именно: *из положительности определителей Гурвица нечетного порядка следует положительность определителей Гурвица четного порядка и наоборот.*

Условия 1) были получены Льенаром и Шипаром в работе [135] при помощи специальных квадратичных форм. Мы дадим более простой вывод условий 1) [а также условий 2), 3), 4)], опирающийся на теорему 10 § 11 и теорию индексов Коши, получив эти условия как частный случай значительно более общей теоремы, к изложению которой мы и переходим.

Введем снова в рассмотрение многочлены $h(u)$ и $g(u)$, связанные с $f(z)$ тождеством

$$f(z)=h(z^2)+zg(z^2).$$

---

[1]) $a_0>0$ по условию.

[2]) Это обстоятельство для первых значений n было установлено в ряде работ по теории регулирования независимо от общего критерия Льенара и Шипара, с которым авторы этих работ, очевидно, не были знакомы.

[3]) См. [208]. Изложение некоторых основных результатов Льенара и Шипара можно найти в фундаментальном обзоре М. Г. Крейна и М. А. Наймарка [16].

[4]) Условия 1), 2), 3), 4) имеют известное преимущество перед условиями Гурвица, поскольку они содержат примерно вдвое меньше детерминантных неравенств, нежели условия Гурвица. Из двух серий детерминантных неравенств $\Delta_1>0$, $\Delta_3>0$, ... и $\Delta_2>0$, $\Delta_3>0$, ... практически лучше та, которая представляется в виде $\Delta_{n-1}>0$, $\Delta_{n-3}>0$, ..., так как она содержит определители меньшего порядка.

Если n четно, $n=2m$, то

$$h(u)=a_0u^m+a_2u^{m-1}+\ldots+a_n,\quad g(u)=a_1u^{m-1}+a_3u^{m-2}+\ldots+a_{n-1});$$

если же n нечетно, $n=2m+1$, то

$$h(u)=a_1u^m+a_3u^{m-1}+\ldots+a_n,\quad g(u)=a_0u^m+a_2u^{m-1}+\ldots+a_{n-1}.$$

Тогда условия $a_n>0$, $a_{n-2}>0$, ... (соответственно $a_{n-1}>0$, $a_{n-2}>0$, ...) можно заменить более общими условиями: $h(u)$ [соответственно $g(u)$] не меняет знака при $u>0$ [1]).

При этих условиях можно вывести формулы для числа корней многочлена $f(z)$ в правой полуплоскости, используя только определители Гурвица нечетного порядка или только определители четного порядка.

Теорема 12. *Если для вещественного многочлена*

$$f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n=h(z^2)+zg(z^2)\quad(a_0>0)$$

выполняется условие: $h(u)$ [или $g(u)$] не меняет знака при $u>0$ и последний определитель Гурвица $\Delta_n\neq0$, то число k корней многочлена $f(z)$, расположенных в правой полуплоскости, определяется по формулам

	$n=2m$	$n=2m+1$
$h(u)$ *не меняет знака при* $u>0$	$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots,\ \Delta_{n-1})=$ $=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots,\ \Delta_n)$	$k=2V(1,\Delta_1,\Delta_3,\ldots,\Delta_n)-\frac{1-\varepsilon_\infty}{2}=$ $=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots,\ \Delta_{n-1})+\frac{1-\varepsilon_\infty}{2}$
$g(u)$ *не меняет знака при* $u>0$	$k=2V(1,\Delta_1,\Delta_3,\ldots,\Delta_{n-1})+\frac{\varepsilon_\infty-\varepsilon_0}{2}=$ $=2V(1,\Delta_2,\Delta_4,\ldots,\Delta_n)-\frac{\varepsilon_\infty-\varepsilon_0}{2}$	$k=2V(1,\Delta_1,\Delta_3,\ldots,\Delta_n)-\frac{1-\varepsilon_0}{2}=$ $=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots,\ \Delta_{n-1})+\frac{1-\varepsilon_0}{2}$

(83)

где

$$\varepsilon_\infty=\operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{u=+\infty},\quad \varepsilon_0=\operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{u=+0}\text{ [2]}.\qquad(83')$$

Доказательство. Снова введем обозначение [3])

$$\varrho=I_{-\infty}^{+\infty}\frac{a_1z^{n-1}-a_3z^{n-3}+\ldots}{a_0z^n-a_2z^{n-2}+\ldots}=n-2k.\qquad(84)$$

---

[1]) То есть $h(u)\geqslant0$ или $h(u)\leqslant0$ при $u>0$ [соответственно $g(u)\geqslant0$ или $g(u)\leqslant0$ при $u>0$].

[2]) Если $a_1\neq0$, то $\varepsilon_\infty=\operatorname{sign}a_1$, и вообще если $a_1=a_3=\ldots=a_{2\mu-1}=0$, $a_{2\mu+1}\neq0$, то $\varepsilon_\infty=\operatorname{sign}a_{2\mu+1}$; если $a_{n-1}\neq0$, то $\varepsilon_0=\operatorname{sign}\frac{a_{n-1}}{a_n}$, и вообще если $a_{n-1}=a_{n-3}=\ldots=a_{n-2\mu+1}=0$, $a_{n-2\mu-1}\neq0$, то $\varepsilon_0=\operatorname{sign}\frac{a_{n-2\mu-1}}{a_n}$.

[3]) См. (71), (72); в нашем случае $s=0$.

Рассмотрим в соответствии с таблицей (83) четыре случая:

1) $n=2m$; $h(u)$ не меняет знака при $u>0$. Тогда [1])

$$I_0^{+\infty}\frac{g(u)}{h(u)}=I_0^{+\infty}\frac{ug(u)}{h(u)}=0,$$

и потому из очевидного равенства

$$I_{-\infty}^{0}\frac{g(u)}{h(u)}=-I_{-\infty}^{0}\frac{ug(u)}{h(u)}$$

следует [2]):

$$I_{-\infty}^{+\infty}\frac{g(u)}{h(u)}=-I_{-\infty}^{+\infty}\frac{ug(u)}{h(u)}.$$

Но тогда из (73′), (74) и (84) находим:

$$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots)$$

Аналогично из формул (73), (75) и (84) следует:

$$k=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots,\ \Delta_n).$$

2) $n=2m$; $g(u)$ не меняет знака при $u>0$. В этом случае

$$I_0^{+\infty}\frac{h(u)}{g(u)}=I_0^{+\infty}\frac{h(u)}{ug(u)}=0,$$

$$I_{-\infty}^{0}\frac{h(u)}{g(u)}+I_{-\infty}^{0}\frac{h(u)}{ug(u)}=0$$

и, следовательно, пользуясь обозначениями (83′), найдем:

$$I_{-\infty}^{+\infty}\frac{h(u)}{g(u)}+I_{-\infty}^{+\infty}\frac{h(u)}{ug(u)}-\varepsilon_0=0. \tag{85}$$

Заменяя функции, стоящие под знаком индекса, их обратными величинами, мы в силу 5° (см. стр. 505) получим:

$$I_{-\infty}^{+\infty}\frac{g(u)}{h(u)}+I_{-\infty}^{+\infty}\frac{ug(u)}{h(u)}=\varepsilon_\infty-\varepsilon_0.$$

Но это в силу (73′), (74) и (84) дает:

$$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots)+\frac{\varepsilon_\infty-\varepsilon_0}{2}.$$

Аналогично из (73′), (75) и (84) находим:

$$k=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots)-\frac{\varepsilon_\infty-\varepsilon_0}{2}.$$

3) $n=2m+1$, $g(u)$ не меняет знака при $u>0$.

В этом случае, как и в предыдущем, имеет место формула (85). Из равенств (73″), (74), (78), (84) и (85) легко получаем:

$$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots)-\frac{1-\varepsilon_0}{2},\quad k=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots)+\frac{1-\varepsilon_0}{2}.$$

4) $n=2m+1$, $h(u)$ не меняет знака при $u>0$.

---

[1]) Если $h(u_1)=0$ $(u_1>0)$, то $g(u_1)\neq 0$, поскольку $\Delta_n\neq 0$. Поэтому из $h(u)\geqslant 0$ $(u>0)$ следует, что $\dfrac{g(u)}{h(u)}$ не меняет знака при переходе через $u=u_1$.

[2]) Из $\Delta_n=a_n\Delta_{n-1}\neq 0$ вытекает, что $h(0)=a_n\neq 0$.

Из равенств $I_0^{+\infty}\frac{g(u)}{h(u)}=I_0^{\infty}\frac{ug(u)}{h(u)}=0$, $I_{-\infty}^{0}\frac{g(u)}{h(u)}+I_{-\infty}^{0}\frac{ug(u)}{h(u)}=0$ заключаем:

$$I_{-\infty}^{+\infty}\frac{g(u)}{h(u)}+I_{-\infty}^{+\infty}\frac{ug(u)}{h(u)}=0.$$

Обращая функции, стоящие под знаком индекса, получаем:

$$I_{-\infty}^{+\infty}\frac{h(u)}{g(u)}+I_{-\infty}^{+\infty}\frac{h(u)}{ug(u)}=\varepsilon_{\infty}.$$

Но тогда формулы (73″), (77) и (84) дают:

$$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots)-\frac{1-\varepsilon_{\infty}}{2},\quad k=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots)+\frac{1-\varepsilon_{\infty}}{2}.$$

Теорема 12 доказана полностью.

Из этой теоремы как частный случай получается теорема 11.

Следствие из теоремы 12. *Если вещественный многочлен* $f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n\,(a_0>0)$ *имеет положительные коэффициенты*

$$a_0>0,\ a_1>0,\ a_2>0,\ \ldots,\ a_n>0$$

и $\Delta_n\neq 0$, *то число* k *корней этого многочлена, расположенных в правой полуплоскости* $\operatorname{Re} z>0$, *определяется формулой*

$$k=2V(1,\ \Delta_1,\ \Delta_3,\ \ldots)=2V(1,\ \Delta_2,\ \Delta_4,\ \ldots).$$

Замечание. Если $\Delta_n\neq 0$, но в последней формуле или в формулах (83) некоторые из промежуточных определителей Гурвица равны нулю, то формулы остаются верными, но при вычислении величин $V(1, \Delta_1, \Delta_3, \ldots)$ и $V(1, \Delta_2, \Delta_4, \ldots)$ следует руководствоваться правилом, изложенным в замечании 1 на стр. 507.

Если же $\Delta_n=\Delta_{n-1}=\ldots=\Delta_{n-\varkappa+1}=0$, $\Delta_{n-\varkappa}\neq 0$, то, отбрасывая в формулах (83) определители $\Delta_{n-\varkappa+1}, \ldots, \Delta_n$ [1]), мы определим по этим формулам число k_1 «неособых» корней $f(z)$, расположенных в правой полуплоскости $\operatorname{Re} z>0$, если соответствующий из многочленов $h_1(u)$ и $g_1(u)$, получающихся из $h(u)$ и $g(u)$ после деления на их наибольший общий делитель $d(u)$, удовлетворяет условиям теоремы 12 [2]).

§ 14. Некоторые свойства многочлена Гурвица. Теорема Стильтьеса. Представление многочленов Гурвица при помощи непрерывных дробей

1. Пусть дан вещественный многочлен

$$f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n\qquad(a_0\neq 0).$$

Представим в виде

$$f(z)=h(z^2)+zg(z^2).$$

Выясним, какие условия должны быть наложены на многочлены $h(u)$ и $g(u)$ для того, чтобы многочлен $f(z)$ был многочленом Гурвица.

Полагая в формуле (20) (стр. 475) $k=s=0$, мы получим необходимое и достаточное условие для того, чтобы $f(z)$ был многочленом Гурвица в виде равенства

$$\varrho=n,$$

---

[1]) См. стр. 508.

[2]) Это условие выполняется, если $h(u)\not\equiv 0$ при $u>0$ или $g(u)\not\equiv 0$ при $u>0$.

где, как и в предыдущих параграфах,

$$\varrho = I_{-\infty}^{+\infty} \frac{a_1 z^{n-1} - a_3 z^{n-3} + \dots}{a_0 z^n - a_2 z^{n-2} + \dots}.$$

Пусть $n = 2m$. Согласно формуле (73′) (стр. 506) это условие может быть записано так:

$$n = 2m = I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} - I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)}. \tag{86}$$

Так как абсолютная величина индекса рациональной дроби не может превосходить степени знаменателя (в данном случае m), то равенство (86) может иметь место тогда и только тогда, когда одновременно

$$I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} = m, \qquad I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} = -m. \tag{87}$$

При $n = 2m+1$ равенство (73″) (поскольку $\varrho = n$) дает:

$$n = I_{-\infty}^{+\infty} \frac{h(u)}{ug(u)} - I_{-\infty}^{+\infty} \frac{h(u)}{g(u)}.$$

Заменяя здесь дроби, стоящие под знаком индекса, их обратными величинами (см. 5° на стр. 505) и замечая при этом, что $h(u)$ и $g(u)$ имеют одну и ту же m-ю степень, получаем [1]):

$$n = 2m+1 = I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} - I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} + \varepsilon_\infty. \tag{88}$$

Исходя снова из того, что абсолютная величина индекса дроби не может превосходить степени знаменателя, заключаем, что равенство (88) имеет место тогда и только тогда, когда одновременно

$$I_{-\infty}^{+\infty} \frac{g(u)}{h(u)} = m, \qquad I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} = -m, \qquad \varepsilon_\infty = 1. \tag{89}$$

Если $n = 2m$, то первое из равенств (87) означает, что многочлен $h(u)$ имеет m различных вещественных корней $u_1 < u_2 < \dots < u_m$ и что правильная дробь $\frac{g(u)}{h(u)}$ представима в виде

$$\frac{g(u)}{h(u)} = \sum_{i=1}^{m} \frac{R_i}{u - u_i}, \tag{90}$$

где

$$R_i = \frac{g(u_i)}{h'(u_i)} > 0 \qquad (i = 1, 2, \dots, m). \tag{90′}$$

Из этого представления дроби $\frac{g(u)}{h(u)}$ следует, что между любыми двумя корнями u_i, u_{i+1} многочлена $h(u)$ лежит вещественный корень u_i' многочлена $g(u)$ $(i = 1, 2, \dots, m-1)$ и что старшие коэффициенты многочленов $h(u)$ и $g(u)$ имеют одинаковые знаки, т. е.

$$h(u) = a_0(u-u_1)\dots(u-u_m), \qquad g(u) = a_1(u-u_1')\dots(u-u_{m-1}'),$$
$$u_1 < u_1' < u_2 < u_2' < \dots < u_{m-1} < u_{m-1}' < u_m; \qquad a_0 a_1 > 0.$$

---

[1]) Как и в предыдущем параграфе, $\varepsilon_\infty = \operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{u=+\infty}$

Второе из равенств (87) вносит лишь одно дополнительное условие

$$u_m < 0.$$

Согласно этому условию все корни $h(u)$ и $g(u)$ должны быть отрицательными. Если $n = 2m+1$, то из первого равенства (89) следует, что $h(u)$ имеет m различных вещественных корней $u_1, u_2, \ldots, u_m$ и

$$\frac{g(u)}{h(u)} = s_{-1} + \sum_{i=1}^{m} \frac{R_i}{u-u_i} \qquad (s_{-1} \neq 0), \tag{91}$$

где

$$R_1 = \frac{g(u_i)}{h'(u_i)} > 0 \qquad (i = 1, 2, \ldots, m). \tag{91'}$$

Из третьего равенства (89) вытекает, что

$$s_{-1} > 0, \tag{92}$$

т. е. что старшие коэффициенты a_0 и a_1 имеют одинаковые знаки. Кроме того, из (91), (91') и (92) следует, что $g(u)$ имеет m вещественных корней $u_1' < u_2' < \ldots < u_m'$, лежащих внутри интервалов $(-\infty, u_1)$, (u_1, u_2), $\ldots$, (u_{m-1}, u_m). Другими словами,

$$h(u) = a_1(u-u_1)\ldots(u-u_m), \qquad g(u) = a_0(u-u_1')\ldots(u-u_m'),$$
$$u_1' < u_1 < u_2' < u_2 < \ldots < u_m' < u_m; \qquad a_0 a_1 > 0.$$

Второе из равенств (89), как и при $n = 2m$, вносит лишь одно дополнительное неравенство

$$u_m < 0.$$

Определение 3. Мы будем говорить, что два многочлена $h(u)$ и $g(u)$ m-й степени [или первый m-й, а второй — $(m-1)$-й степени] образуют *положительную пару*[1]), если корни этих многочленов $u_1, u_2, \ldots, u_m$ и $u_1', u_2', \ldots, u_m'$ (соответственно $u_1', u_2', \ldots, u_{m-1}'$) все различны, вещественны, отрицательны и перемежаются следующим образом:

$$u_1' < u_1 < u_2' < u_2 < \ldots < u_m' < u_m < 0$$
$$(\text{соответственно } u_1 < u_2' < u_2 < \ldots < u_{m-1}' < u_m < 0),$$

а старшие коэффициенты этих многочленов имеют одинаковые знаки[2]).

Вводя положительные числа $v_i = -u_i$, $v_i' = -u_i'$ и помножая оба многочлена $h(u)$ и $g(u)$, образующих положительную пару, на ± 1 так, чтобы старшие коэффициенты этих многочленов стали положительными, мы эти многочлены сможем представить в виде

$$h(u) = a_1 \prod_{i=1}^{m} (u+v_i), \qquad g(u) = a_0 \prod_{i=1}^{m} (u+v_1'), \tag{93}$$

где

$$a_1 > 0,\ a_0 > 0,\ 0 < v_m < v_m' < v_{m-1} < v_{m-1}' < \ldots < v_1 < v_1',$$

---

[1]) См. [7], стр. 333. Определение положительной пары многочленов, проведенное здесь, несколько отличается от определения, данного в книге [7].

[2]) Если мы отбросим требование об отрицательности корней, мы получим вещественную пару многочленов. Относительно использования этого понятия в задаче Рауса—Гурвица см. [23].

если оба многочлена $h(u)$ и $g(u)$ имеют степень m, и в виде

$$h(u)=a_0\prod_{i=1}^{m}(u+v_i),\qquad g(u)=a_1\prod_{i=1}^{m-1}(u+v_i'),\tag{93'}$$

где

$$a_0>0,\ a_1>0,\ 0<v_m<v_{m-1}'<v_{m-1}<\ldots<v_1'<v_1,$$

если $h(u)$ имеет степень m, а $g(u)$ — степень $m-1$.

Приведенные ранее рассуждения доказывают следующие две теоремы:

Теорема 13. *Для того чтобы многочлен $f(z)=h(z^2)+zg(z^2)$ был многочленом Гурвица, необходимо и достаточно, чтобы многочлены $h(u)$ и $g(u)$ составляли положительную пару*[1]).

Теорема 14. *Для того чтобы два многочлена $h(u)$ и $g(u)$, из которых первый имеет степень m, а второй имеет степень m или $m-1$, составляли положительную пару, необходимо и достаточно, чтобы имели место равенства*

$$I_{-\infty}^{+\infty}\frac{g(u)}{h(u)}=m,\quad I_{+\infty}^{+\infty}\frac{ug(u)}{h(u)}=-m\tag{94}$$

и в случае, когда степени $h(u)$ и $g(u)$ одинаковы, дополнительное условие

$$\varepsilon_\infty=\operatorname{sign}\left[\frac{g(u)}{h(u)}\right]_{+\infty}=1.\tag{95}$$

2. Из последней теоремы, используя свойства индексов Коши, мы легко получим теорему Стильтьеса о представлении дроби $\frac{g(u)}{h(u)}$ в виде непрерывной дроби специального типа в случае, когда многочлены $h(u)$ и $g(u)$ образуют положительную пару многочленов.

Доказательство теоремы Стильтьеса опирается на следующую лемму:

Лемма. *Если многочлены $h(u)$, $g(u)$ [степень $h(u)$ равна m] составляют положительную пару и*

$$\frac{g(u)}{h(u)}=c+\cfrac{1}{du+\cfrac{h_1(u)}{g_1(u)}},\tag{96}$$

где c, d — постоянные, а $h_1(u)$, $g_1(u)$ — многочлены степени $\leqslant m-1$, то

1° $c\geqslant 0$, $d>0$,

2° *многочлены $h_1(u)$, $g_1(u)$ имеют степень $m-1$,*

3° *многочлены $h_1(u)$, $g_1(u)$ составляют положительную пару.*

Задание $h(u)$ и $g(u)$ однозначно определяет многочлены $h_1(u)$, $g_1(u)$ (с точностью до общего постоянного множителя) и постоянные c и d.

Обратно, из (96) и 1°, 2°, 3° следует, что многочлены $h(u)$ и $g(u)$ образуют положительную пару, причем $h(u)$ имеет степень m, а $g(u)$ — степень m или $m-1$ в зависимости от того, $c>0$ или $c=0$.

Доказательство. Пусть $h(u)$, $g(u)$ — положительная пара. Тогда из (94) и (96) следует:

$$m=I_{-\infty}^{+\infty}\frac{g(u)}{h(u)}=I_{+\infty}^{+\infty}\cfrac{1}{du+\cfrac{h_1(u)}{g_1(u)}},\tag{97}$$

Из этого равенства следует, что степень $g_1(u)$ равна $m-1$ и что $d\neq 0$.

[1]) Эта теорема представляет собой частный случай так называемой теоремы Эрмита—Билера (см. [37], стр. 21).

Далее из (97) находим:

$$m = -I_{-\infty}^{+\infty}\left[du + \frac{h_1(u)}{g_1(u)}\right] + \operatorname{sign} d = -I_{-\infty}^{+\infty}\frac{h_1(u)}{g_1(u)} + \operatorname{sign} d.$$

Отсюда следует, что $d > 0$ и что

$$I_{-\infty}^{+\infty}\frac{h_1(u)}{g_1(u)} = -(m-1). \tag{98}$$

Теперь второе равенство (94) дает:

$$-m = I_{-\infty}^{+\infty}\frac{ug(u)}{h(u)} = I_{-\infty}^{+\infty}\left[cu + \frac{1}{d + \dfrac{h_1(u)}{ug_1(u)}}\right] =$$

$$= I_{-\infty}^{+\infty}\frac{1}{d + \dfrac{h_1(u)}{ug_1(u)}} = -I_{-\infty}^{+\infty}\left[d + \frac{h_1(h)}{ug_1(u)}\right] = -I_{-\infty}^{+\infty}\frac{h_1(u)}{ug_1(u)}. \tag{99}$$

Отсюда следует, что $h_1(u)$ имеет степень $m-1$ [1]).

Условие (95) в силу (96) дает: $c > 0$. Если же степень $g(u)$ меньше степени $h(u)$, то из (96) вытекает: $c = 0$.

Из (98) и (99) следует:

$$I_{-\infty}^{+\infty}\frac{g_1(u)}{h_1(u)} = m-1, \qquad I_{-\infty}^{+\infty}\frac{ug_1(u)}{h_1(u)} = -m + \varepsilon_{\infty}^{(1)}, \tag{100}$$

где

$$\varepsilon_{\infty}^{(1)} = \operatorname{sign}\left[\frac{g_1(u)}{h_1(u)}\right]_{u=+\infty}.$$

Так как второй из индексов (100) по абсолютной величине $\leqslant m-1$, то

$$\varepsilon_{\infty}^{(1)} = 1, \tag{101}$$

а тогда из (100) и (101) на основании теоремы 12 заключаем, что многочлены $h_1(u)$ и $g_1(u)$ образуют положительную пару.

Из (96) следует:

$$c = \lim_{u\to\infty}\frac{g(u)}{h(u)}, \qquad \lim_{u\to\infty}\left[\frac{g(u)}{h(u)} - c\right]u = \frac{1}{d}.$$

После того как c и d определены, из (96) определяется отношение $\dfrac{h_1(u)}{g_1(u)}$.

Соотношения (97), (98), (99), (100), (101), использованные в обратном порядке, устанавливают вторую часть леммы. Таким образом, лемма доказана полностью.

Пусть нам дана положительная пара многочленов $h(u)$, $g(u)$ и m есть степень многочлена $h(u)$. Тогда, разделив $g(u)$ на $h(u)$ и обозначив частное через c_0, а остаток через $g_1(u)$, получим:

$$\frac{g(u)}{h(u)} = c_0 + \frac{g_1(u)}{h(u)} = c_0 + \frac{1}{\dfrac{h(u)}{g_1(u)}}.$$

$\dfrac{h(u)}{g_1(u)}$ можно представить в виде $d_0u + \dfrac{h_1(u)}{g_1(u)}$, где степень $h_1(u)$, как и степень $g_1(u)$, меньше m. Отсюда

$$\frac{g(u)}{h(u)} = c_0 + \frac{1}{d_0u + \dfrac{h_1(u)}{g_1(u)}}. \tag{102}$$

---

[1]) Из равенства (99) следует, что знаменатель $ug_1(u)$ имеет степень m и что между любыми двумя корнями знаменателя $ug_1(u)$ содержится корень числителя $h_1(u)$.

Таким образом, для положительной пары $h(u)$ и $g(u)$ всегда имеет место представление (96). Согласно лемме

$$c_0 \geqslant 0, \quad d_0 > 0,$$

а многочлены $h_1(u)$ и $g_1(u)$ имеют степень $m-1$ и образуют положительную пару.

Применяя эти же рассуждения к положительной паре $h_1(u)$, $g_1(u)$, получим равенство

$$\frac{g_1(u)}{h_1(u)} = c_1 + \cfrac{1}{d_1 u + \cfrac{h_2(u)}{g_2(u)}}, \tag{102'}$$

где

$$c_1 > 0, \quad d_1 > 0,$$

а многочлены $h_2(u)$ и $g_2(u)$ имеют степень $m-2$ и образуют положительную пару. Продолжая этот процесс далее, мы в конце концов придем к положительной паре h_m и g_m, где h_m и g_m — постоянные одного знака. Мы положим:

$$\frac{g_m}{h_m} = c_m. \tag{102$^{(m)}$}$$

Тогда из (102), (102'), ..., (102$^{(m)}$) вытекает:

$$\frac{g(u)}{h(u)} = c_0 + \cfrac{1}{d_0 u + \cfrac{1}{c_1 + \cfrac{1}{d_1 u + \cfrac{1}{c_2 + \ddots + \cfrac{1}{d_{m-1} u + \cfrac{1}{c_m}}}}}}.$$

Пользуясь второй частью леммы, мы аналогично покажем, что при любых $c_0 \geqslant 0$, $c_1 > 0, \ldots, c_m > 0$, $d_0 > 0$, $d_1 > 0, \ldots, d_{m-1} > 0$ написанная непрерывная дробь однозначно (с точностью до общего постоянного множителя) всегда определяет положительную пару многочленов $h(u)$ и $g(u)$, причем $h(u)$ имеет степень m, а $g(u)$ имеет степень m при $c_0 > 0$ и степень $m-1$ при $c_0 = 0$.

Таким образом, нами доказана [1])

Теорема 15 (Стильтьеса). *Если $h(u)$, $g(u)$ — положительная пара многочленов и $h(u)$ имеет степень m, то*

$$\frac{g(u)}{h(u)} = c_0 + \cfrac{1}{d_0 u + \cfrac{1}{c_1 + \cfrac{1}{d_1 u + \cfrac{1}{c_2 + \ddots + \cfrac{1}{d_{m-1} u + \cfrac{1}{c_m}}}}}}, \tag{103}$$

[1]) Доказательство теоремы Стильтьеса, не опирающееся на теорию индексов Коши, можно найти в книге [7], стр. 333—337.

где

$$c_0 \geqslant 0,\ c_1 > 0, \ldots, c_m > 0,\ d_0 > 0, \ldots, d_{m-1} > 0.$$

При этом $c_0 = 0$, *если* $g(u)$ *имеет степень* $m-1$, *и* $c_0 > 0$, *если* $g(u)$ *имеет степень* m. *Постоянные* c_i, d_k *однозначно определяются заданием* $h(u)$, $g(u)$.

Обратно, при любом $c_0 \geqslant 0$ *и любых положительных* $c_1, \ldots, c_m$, $d_0, \ldots, d_{m-1}$ *непрерывная дробь* (103) *определяет положительную пару многочленов* $h(u)$, $g(u)$, *где* $h(u)$ *имеет степень* m.

Из теоремы 13 и теоремы Стильтьеса следует:

Т е о р е м а 16. *Вещественный многочлен* n*-й степени* $f(z) = h(z^2) + zg(z^2)$ *в том и только в том случае является многочленом Гурвица, если имеет место формула* (103) *при неотрицательном* c_0 *и положительных* $c_1, \ldots, c_m$, $d_0, \ldots, d_{m-1}$. *При этом* $c_0 > 0$, *когда* n *нечетно, и* $c = 0$, *когда* n *четно.*

§ 15. Область устойчивости. Параметры Маркова

Каждому вещественному многочлену n-й степени можно отнести точку n-мерного пространства, координаты которой равны частным от деления на старший коэффициент всех остальных коэффициентов. В таком «пространстве коэффициентов» все многочлены Гурвица образуют некоторую n-мерную область, которая определяется [1]) неравенствами Гурвица $\Delta_1 > 0$, $\Delta_2 > 0, \ldots, \Delta_n > 0$ или, например, неравенствами Льенара — Шипара $a_n > 0$, $a_{n-2} > 0, \ldots, \Delta_1 > 0$, $\Delta_3 > 0, \ldots$ Эту область будем называть *областью устойчивости*. Если коэффициенты уравнения заданы как функции p параметров, то область устойчивости строится в пространстве этих параметров.

Исследование области устойчивости представляет большой практический интерес; так, например, такое исследование существенно при проектировании новых систем регулирования [2]).

В § 16 мы покажем, что две замечательные теоремы, установленные А. А. Марковым и П. Л. Чебышевым в связи с разложением непрерывных дробей в степенные ряды по отрицательным степеням аргумента, имеют тесное отношение к исследованию области устойчивости. При формулировке и доказательстве этих теорем нам удобно будет задавать многочлен не его коэффициентами, а специальными параметрами, которые мы назовем *параметрами Маркова*.

Пусть дан вещественный многочлен

$$f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n \qquad (a_0 \neq 0).$$

Представим его в виде

$$f(z) = h(z^2) + zg(z^2).$$

Примем, что многочлены $h(u)$ и $ug(u)$ взаимно просты ($\Delta_n \neq 0$). Несократимую рациональную дробь $\frac{g(u)}{h(u)}$ разложим в ряд по убывающим

[1]) При $a_0 = 1$.

[2]) Исследованию области устойчивости, а также областей, соответствующих различным значениям k (k — число корней в правой полуплоскости), посвящен ряд работ Ю. И. Наймарка (см. монографию [27]).

степеням u [1]):

$$\frac{g(u)}{h(u)} = s_{-1} + \frac{s_0}{u} - \frac{s_1}{u^2} + \frac{s_2}{u^3} - \frac{s_3}{u^4} + \dots \tag{104}$$

Если n нечетно, то для получения этой формулы необходимо добавочно предположить, что $a_1 \neq 0$ (в противном случме $s_{-1} = \infty$).

Последовательность чисел $s_0, s_1, s_2, \dots$ определяет бесконечную ганкелеву матрицу $S = \| s_{i+k} \|_0^\infty$. Определим рациональную функцию $R(v)$ равенством

$$R(v) = -\frac{g(-v)}{h(-v)}. \tag{105}$$

Тогда

$$R(v) = -s_{-1} + \frac{s_0}{v} + \frac{s_1}{v^2} + \frac{s_2}{v^3} + \dots, \tag{106}$$

и потому имеет место соответствие (см. стр. 498)

$$R(v) \sim S. \tag{107}$$

Отсюда следует [2]), что матрица S имеет ранг $m = \left[\frac{n}{2}\right]$, поскольку m — степень многочлена $h(u)$ и, следовательно, число полюсов функции $R(v)$.

При $n = 2m$ (в этом случае $s_{-1} = 0$) задание матрицы S однозначно определяет несократимую дробь $\frac{g(u)}{h(u)}$ и, следовательно, с точностью до постоянного множителя однозначно определяет $f(z)$. При $n = 2m + 1$ для задания $f(z)$, помимо матрицы S, необходимо еще знать коэффициент s_{-1}.

С другой стороны, для задания бесконечной ганкелевой матрицы m-го ранга S достаточно задать лишь первые $2m$ чисел $s_0, s_1, \dots, s_{2m-1}$. Числа $s_0, s_1, \dots, s_{2m-1}$ могут быть выбраны произвольно при одном лишь ограничении

$$D_m = |s_{i+k}|_0^m \neq 0; \tag{108}$$

все последующие коэффициенты разложения (104) $s_{2m}, s_{2m+1}, \dots$ однозначно (и даже рационально) выражаются через первые $2m$: $s_0, s_1, \dots, s_{2m-1}$. Действительно, у бесконечной ганкелевой матрицы m-го ранга S элементы связаны между собой рекуррентными соотношениями (см. теорему 7 на стр. 496)

$$s_q = \sum_{g=1}^{m} \alpha_g s_{q-g} \qquad (q = m, m+1, \dots). \tag{109}$$

Если числа $s_0, s_1, \dots, s_{2m-1}$ удовлетворяют неравенству (108), то после задания этих чисел из первых m соотношений (109) однозначно определяются коэффициенты $\alpha_1, \alpha_2, \dots, \alpha_m$; тогда последующие соотношения (109) определяют $s_{2m}, s_{5m+1}, \dots$

Таким образом, вещественный многочлен $f(z)$ степени $n = 2m$ при $\Delta_n \neq 0$ может быть однозначно [3]) задан при помощи $2m$ чисел $s_0, s_1, \dots, s_{2m-1}$, удовлетворяющих неравенству (108). При $n = 2m + 1$ к этим числам следует прибавить еще s_{-1}.

---

[1]) Для дальнейшего нам удобно коэффициенты при четных отрицательных степенях u обозначать через $-s_1, -s_3$ и т. д.

[2]) См. теорему 8 (стр. 498).

[3]) С точностью до постоянного множителя.

n величин $s_0, s_1, \ldots, s_{2m-1}$ (при $n = 2m$) или $s_{-1}, s_0, \ldots, s_{2m-1}$ (при $n = 2m+1$) мы будем называть *параметрами Маркова* для многочлена $f(z)$. В n-мерном пространстве эти параметры могут быть рассматриваемы как координаты точки, изображающей данный многочлен $f(z)$.

Выясним, какие условия должны быть наложены на параметры Маркова для того, чтобы соответствующий многочлен $f(z)$ был многочленом Гурвица. Этим самым мы определим область устойчивости в пространстве параметров Маркова.

Многочлен Гурвица характеризуется условиями (94) и дополнительным условием (95) при $n = 2m+1$. Вводя функцию $R(v)$ [см. (105)], мы равенства (94) запишем так:

$$I_{-\infty}^{+\infty} R(v) = m, \qquad I_{-\infty}^{+\infty} vR(v) = m. \tag{110}$$

Дополнительное же условие (95) для $n = 2m+1$ дает:

$$s_{-1} > 0.$$

Введем наряду с матрицей $S = \| s_{i+k} \|_0^\infty$ бесконечную ганкелеву матрицу $S^{(1)} = \| s_{i+k+1} \|_0^\infty$. Тогда поскольку из (106)

$$vR(v) = -s_{-1}v + s_0 + \frac{s_1}{v} + \frac{s_2}{v^2} + \ldots,$$

то имеет место соответствие

$$vR(v) \sim S^{(1)}. \tag{111}$$

Матрица $S^{(1)}$, как и матрица S, имеет конечный ранг m, так как функция $vR(v)$, как и $R(v)$, имеет m полюсов. Поэтому и формы

$$S_m(x, x) = \sum_{i,k=0}^{m-1} s_{i+k} x_i x_k, \qquad S_m^{(1)}(x, x) = \sum_{i,k=0}^{m-1} s_{i+k+1} x_i x_k$$

имеют ранг m. Но согласно теореме 9 (стр. 500) сигнатуры этих форм в силу соответствий (107), (111) равны индексам (110) и, следовательно, также равны m. Таким образом, условия (110) означают положительную определенность квадратичных форм $S_m(x, x)$ и $S_m^{(1)}(x, x)$. Нами установлена

Теорема 17. *Для того чтобы вещественный многочлен* $f(z) = h(z^2) + zg(z^2)$ *степени* $n = 2m$ или $n = 2m+1$ *был многочленом Гурвица, необходимо и достаточно*[1]), *чтобы*

---

[1]) Мы не оговариваем специально неравенства $\Delta_n \neq 0$, поскольку это неравенство автоматически следует из условий теоремы. В самом деле, если $f(z)$ — многочлен Гурвица, то, как известно, $\Delta_n \neq 0$. Если же даны условия 1°, 2°, то из положительной определенности формы $S_m^{(1)}(x, x)$ вытекает равенство

$$-I_{-\infty}^{+\infty} \frac{ug(u)}{h(u)} = I_{-\infty}^{+\infty} vR(v) = m,$$

из которого следует несократимость дроби $\frac{ug(u)}{h(u)}$, что выражается неравенством $\Delta_n \neq 0$.

Точно так же из условий теоремы автоматически следует, что $D_m = |s_{i+k}|_0^{m-1} \neq 0$, т. е. что числа $s_0, s_1, \ldots, s_{2m-1}$ и (при $n = 2m+1$) число s_{-1} являются параметрами Маркова для многочлена $f(z)$.

1° *квадратичные формы*

$$S_m(x,\ x)=\sum_{i,\ k=0}^{m-1} s_{i+k}x_ix_k, \qquad S_m^{(1)}(x,\ x)=\sum_{i,\ k=0}^{m-1} s_{i+k+1}x_ix_k \tag{112}$$

были положительно определенными и

2° (*при* $n=2m+1$)

$$s_{-1}>0. \tag{113}$$

Здесь $s_{-1},\ s_0,\ s_1,\ \ldots,\ s_{2m-1}$ — *коэффициенты в разложении*

$$\frac{g(u)}{h(u)}=s_{-1}+\frac{s_0}{u}-\frac{s_1}{u^2}+\frac{s_2}{u^3}-\frac{s_3}{u^4}+\ldots$$

Введем обозначения для определителей:

$$D_p=|\,s_{i+k}\,|_0^{p-1}, \qquad D_p^{(1)}=|\,s_{i+k+1}\,|_0^{p-1} \qquad (p=1,\ 2,\ \ldots,\ m). \tag{114}$$

Тогда условие 1° эквивалентно системе детерминантных неравенств

$$\left.\begin{aligned} &D_1=s_0>0,\ D_2=\begin{vmatrix} s_0 & s_1\\ s_1 & s_2\end{vmatrix}>0,\ \ldots,\ D_m=\begin{vmatrix} s_0 & s_1 & \ldots & s_{m-1}\\ s_1 & s_2 & \ldots & s_m\\ \ldots & \ldots & \ldots & \ldots\\ s_{m-1} & s_m & \ldots & s_{2m-2}\end{vmatrix}>0,\\ &D_1^{(1)}=s_1>0,\ D_2^{(1)}=\begin{vmatrix} s_1 & s_2\\ s_2 & s_3\end{vmatrix}>0,\ \ldots,\ D_m^{(1)}=\begin{vmatrix} s_1 & s_2 & \ldots & s_m\\ s_2 & s_3 & \ldots & s_{m+1}\\ \ldots & \ldots & \ldots & \ldots\\ s_m & s_{m+1} & \ldots & s_{2m-1}\end{vmatrix}>0.\end{aligned}\right\} \tag{115}$$

В случае $n=2m$ неравенства (115) определяют область устойчивости в пространстве параметров Маркова. При $n=2m+1$ к этим неравенствам следует прибавить еще одно:

$$s_{-1}>0. \tag{116}$$

В следующем параграфе мы выясним, какие свойства матрицы S вытекают из неравенств (115) и тем самым выделим специальный класс бесконечных ганкелевых матриц S, которые соответствуют многочленам Гурвица.

§ 16. Связь с проблемой моментов

1. Сформулируем следующую проблему моментов из положительной полуоси[1]) $0<v<\infty$:

Дана последовательность вещественных чисел $s_0,\ s_1,\ \ldots$ *Требуется определить положительные числа*

$$\mu_1>0,\quad \mu_2>0,\ \ldots,\ \mu_m>0;\quad 0<v_1<v_2<\ldots<v_m \tag{117}$$

так, чтобы имели место равенства

$$s_p=\sum_{j=1}^{m}\mu_j v_j^p \qquad (p=0,\ 1,\ 2,\ \ldots) \tag{118}$$

---

[1]) Эту проблему моментов следовало бы называть дискретной в отличие от общей степенной проблемы моментов, в которой суммы $\sum_{j=1}^{m}\mu_j v_j^p$ заменены интегралами Стильтьеса $\int_0^\infty v^p\,d\mu(v)$ (см. [3]).

Нетрудно усмотреть, что система равенств (118) равносильна следующему разложению в ряд по отрицательным степеням u:

$$\sum_{j=1}^{m} \frac{\mu_j}{u+v_j} = \frac{s_0}{u} - \frac{s_1}{u^2} + \frac{s_2}{u^3} - \dots \tag{119}$$

В этом случае бесконечная ганкелева матрица $S = \| s_{i+k} \|_0^\infty$ имеет конечный ранг m, и в силу неравенств (117) в несократимой правильной рациональной дроби

$$\frac{g(u)}{h(u)} = \sum_{j=1}^{m} \frac{\mu_j}{u+v_j} \tag{120}$$

[старшие коэффициенты в $h(u)$ и $g(u)$ выбираем положительными] многочлены $h(u)$ и $g(u)$ образуют положительную пару [см. (91) и (91′)].

Поэтому (см. теорему 14) сформулированная нами проблема моментов имеет решение в том и только в том случае, когда последовательность чисел s_0, s_1, s_2, ... при помощи равенств (119) и (120) определяет многочлен Гурвица $f(z) = h(z^2) + zg(z^2)$ степени $2m$.

Решение проблемы моментов единственно, поскольку из разложения (119) однозначно определяются положительные числа v_j и μ_j $(j = 1, 2, \dots, m)$.

Наряду с «бесконечной» проблемой моментов (118) рассмотрим и «конечную» проблему моментов, задаваемую первыми $2m$ из уравнений (118):

$$s_p = \sum_{j=1}^{m} \mu_j v_j^p \qquad (p = 0, 1, \dots, 2m-1). \tag{121}$$

Из этих соотношений уже вытекают следующие выражения для ганкелевых квадратичных форм:

$$\begin{aligned} \sum_{i,\,k=0}^{m-1} s_{i+k} x_i x_k &= \sum_{j=1}^{m} \mu_j (x_0 + x_1 v_j + \dots + x_{m-1} v_j^{m-1})^2, \\ \sum_{i,\,k=0}^{m-1} s_{i+k+1} x_i x_k &= \sum_{j=1}^{m} \mu_j v_j (x_0 + x_1 v_j + \dots + x_{m-1} v^{m-1})^2. \end{aligned} \tag{122}$$

Поскольку линейные формы относительно переменных x_0, x_1, ..., x_{m-1}

$$x_0 + x_1 v_j + \dots + x_{m-1} v_j^{m-1} \qquad (j = 1, 2, \dots, m)$$

независимы (коэффициенты этих форм образуют отличный от нуля определитель Вандермонда!) и $v_j > 0$, $\mu_j > 0$ $(j = 1, 2 \dots, m)$, то квадратичные формы (122) являются положительно определенными. Тогда согласно теореме 17 числа s_0, s_1, ..., s_{2m-1} являются параметрами Маркова некоторого многочлена Гурвица $f(z)$. Эти числа являются первыми $2m$ коэффициентами разложения (119). Вместе с остальными коэффициентами s_{2m}, s_{2m+1}, ... они определяют бесконечную разрешимую проблему моментов (118), которая имеет то же решение, что и конечная проблема (121).

Таким образом, нами доказана следующая

Теорема 18. 1. *Для того чтобы конечная проблема моментов*

$$s_p = \sum_{j=1}^{m} \mu_j v_j^p \tag{123}$$

($p=0, 1, \ldots, 2m-1$; $\mu_1>0, \ldots, \mu_m>0$, $0<v_1<v_2<\ldots<v_m$), *где* s_p — *заданные, а* v_j *и* μ_j — *искомые вещественные числа* ($p=0, 1, \ldots, 2m-1$; $j=1, 2, \ldots, m$), *имела решение, необходимо и достаточно, чтобы квадратичные формы*

$$\sum_{i,\,k=0}^{m-1} s_{i+k}x_ix_k, \qquad \sum_{i,\,k=0}^{m-1} s_{i+k+1}x_ix_k \tag{124}$$

были положительно определенными, т. е. чтобы числа $s_0, s_1 \ldots, s_{2m-1}$ *были параметрами Маркова некоторого многочлена Гурвица степени* $2m$.

2. *Для того чтобы бесконечная проблема моментов*

$$s_p=\sum_{j=1}^{m}\mu_j v_j^p \tag{125}$$

($p=0, 1, 2, \ldots$; $\mu_1>0, \ldots, \mu_m>0$; $0<v_1<v_2<\ldots<v_m$), *где* s_p — *заданные, а* v_j *и* μ_j — *искомые вещественные числа* ($p=0, 1, 2, \ldots$; $j=1, 2, \ldots, m$), *имела решение, необходимо и достаточно, чтобы* 1° *квадратичные формы* (124) *были положительно определенными и* 2° *бесконечная ганкелева матрица* $S=\|s_{i+k}\|_0^\infty$ *имела ранг* m, *т. е. чтобы ряд*

$$\frac{s_0}{u}-\frac{s_1}{u^2}+\frac{s_2}{u^3}-\ldots=\frac{g(u)}{h(u)} \tag{126}$$

опребелял многочлен Гурвица $f(z)=h(z^2)+zg(z^2)$ *степени* $2m$.

3. *Решение проблемы моментов, как конечной* (123), *так и бесконечной* (124), *всегда единственно.*

2. Доказанную теорему мы используем для исследования миноров бесконечной ганкелевой матрицы $S=\|s_{i+k}\|_0^\infty$ ранга m, соответствующей некоторому многочлену Гурвица, т. е. матрицы $S=\|s_{i+k}\|_0^\infty$, для которой квадратичные формы (124) являются положительно определенными. В этом случае порождающие матрицу S числа s_0, s_1 s_2, ... могут быть представлены в виде (123), и потому для произвольного минора матрицы S порядка $h\leqslant m$ имеем:

$$\begin{Vmatrix} s_{i_1+k_1}\ldots s_{i_1k_h} \\ \ldots\ldots\ldots \\ s_{i_h+k_1}\ldots s_{i_h+k_h} \end{Vmatrix} = \begin{Vmatrix} \mu_1v_1^{i_1}\mu_2v_2^{i_1}\ldots\mu_mv_m^{i_1} \\ \ldots\ldots\ldots\ldots \\ \mu_1v_1^{i_h}\mu_2v_2^{i_h}\ldots\mu_mv_m^{i_h} \end{Vmatrix} \cdot \begin{Vmatrix} v_1^{k_1}\ldots v_1^{k_h} \\ v_2^{k_1}\ldots v_2^{k_h} \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ v_m^{k_1}\ldots v_m^{k_h} \end{Vmatrix},$$

и, следовательно,

$$S\begin{pmatrix} i_1 & i_2\ldots i_h \\ k_1 & k_2\ldots k_h \end{pmatrix} = \sum_{1\leqslant a_1<a_2<\ldots<a_h\leqslant m} \mu_{a_1}\mu_{a_2}\ldots\mu_{a_h} \begin{vmatrix} v_{a_1}^{i_1} & v_{a_2}^{i_1}\ldots v_{a_h}^{i_1} \\ v_{a_1}^{i_2} & v_{a_2}^{i_2}\ldots v_{a_h}^{i_2} \\ \ldots\ldots\ldots\ldots \\ v_{a_1}^{i_h} & v_{a_2}^{i_h}\ldots v_{a_h}^{i_h} \end{vmatrix} \begin{vmatrix} v_{a_1}^{k_1} & v_{a_1}^{k_2}\ldots v_{a_1}^{k_h} \\ v_{a_2}^{k_1} & v_{a_2}^{k_2}\ldots v_{a_2}^{k_h} \\ \ldots\ldots\ldots\ldots \\ v_{a_h}^{k_1} & v_{a_h}^{k_2}\ldots v_{a_h}^{k_h} \end{vmatrix}. \tag{127}$$

Но из неравенств

$$0<v_1<v_2<\ldots<v_m, \quad i_1<i_2<\ldots<i_h, \quad k_1<k_2<\ldots<k_h$$

следует положительность обобщенных определителей Вандермонда[1])

$$\begin{vmatrix} v_{a_1}^{i_1} & v_{a_2}^{i_1}\ldots v_{a_h}^{i_1} \\ v_{a_1}^{i_2} & v_{a_2}^{i_2}\ldots v_{a_h}^{i_2} \\ \ldots\ldots\ldots \\ v_{a_1}^{i_h} & v_{a_2}^{i_h}\ldots v_{a_h}^{i_h} \end{vmatrix}>0, \quad \begin{vmatrix} v_{a_1}^{k_1} & v_{a_1}^{k_2}\ldots v_{a_1}^{k_h} \\ v_{a_2}^{k_1} & v_{a_2}^{k_2}\ldots v_{a_2}^{k_h} \\ \ldots\ldots\ldots \\ v_{a_h}^{k_1} & v_{a_h}^{k_2}\ldots v_{a_h}^{k_h} \end{vmatrix}>0.$$

Поэтому, поскольку и числа $\mu_j>0$ $(j=1, 2, \ldots, m)$, то из (127) вытекает

$$S\begin{pmatrix} i_1 & i_2\ldots i_h \\ k_1 & k_2\ldots k_h \end{pmatrix}>0 \quad \left(0\leqslant \begin{matrix} i_1<i_2<\ldots<i_h \\ k_1<k_2<\ldots<k_h, \end{matrix} \quad h=1, 2, \ldots, m\right). \tag{128}$$

Обратно, если в бесконечной ганкелевой матрице $S=\| s_{i+k} \|_0^\infty$ ранга m все миноры любого порядка $h\leqslant m$ положительны, то квадратичные формы (127) являются положительно определенными.

Введем

Определение 4. Бесконечную матрицу $A=\| a_{ik} \|_0^\infty$ будем называть *вполне положительной ранга* m в том и только в том случае, когда все миноры матрицы A порядка $h\leqslant m$ положительны, а все миноры порядка $h>m$ равны нулю.

Теперь сформулируем установленные свойства матрицы S[2]).

Теорема 19. *Для того чтобы бесконечная ганкелева матрица* $S=\| s_{i+k} \|_0^\infty$ *была вполне положительной ранга* m, *необходимо и достаточно, чтобы* 1) *матрица* S *имела ранг* m *и* 2) *квадратичные формы*

$$\sum_{i,\,k=0}^{m-1} s_{i+k}x_ix_k, \qquad \sum_{i,\,k=0}^{m-1} s_{i+k+1}x_ix_k$$

были положительно определенными.

Из этой теоремы и теоремы 17 следует

Теорема 20. *Вещественный многочлен* $f(z)$ *степени* n *является многочленом Гурвица в том и только в том случае, когда соответствующая этому многочлену бесконечная ганкелева матрица* $S=\| s_{i+k} \|_0^\infty$ *является вполне положительной ранга* $m=\left[\frac{n}{2}\right]$ *и в случае* n *нечетного дополнительно* $s_{-1}>0$.

При этом элементы $s_0, s_1, s_2, \ldots$ *матрицы* S *и число* s_{-1} *определяются из разложения*

$$\frac{g(u)}{h(u)}=s_{-1}+\frac{s_0}{u}-\frac{s_1}{u^2}+\frac{s_2}{u^3}-\ldots, \tag{129}$$

где

$$f(z)=h(z^2)+zg(z^2).$$

---

[1]) См. стр. 395, пример 1.

[2]) См. [82д].

§ 17. Связь между определителями Гурвица и определителями Маркова[1])

Рассмотрим сначала случай четного $n=2m$. Тогда

$$\frac{g(u)}{h(u)}=\frac{a_1u^{m-1}+a_3u^{m-2}+\dots}{a_0u^m+a_2u^{m-1}+\dots}=\frac{s_0}{u}-\frac{s_1}{u^2}+\frac{s_2}{u^3}-\dots \tag{130}$$

Согласно формуле (68′) на стр. 503 имеем:

$$\nabla_{2p}=\begin{vmatrix} a_0 & a_2 & a_4 & a_6\dots \\ 0 & a_1 & a_3 & a_5\dots \\ 0 & a_0 & a_2 & a_4\dots \\ 0 & 0 & a_1 & a_3\dots \\ \dots & \dots & \dots & \dots \end{vmatrix}=a_0^{2p}D_p \qquad (p=1,\dots,m).$$

С другой стороны, $\nabla_{2p}=a_0\Delta_{2p-1}$, где Δ_{2p-1} — определитель Гурвица порядка $2p-1$. Поэтому

$$\Delta_{2p-1}=a_0^{2p-1}D_p \qquad (p=1,\dots,m). \tag{131}$$

Помножим обе части равенства (130) на u и снова применим формулу (68′) со стр. 503. Получим:

$$(-1)^p\Delta_{2p}=\nabla^1_{2p}=\begin{vmatrix} a_0 & a_2 & a_4\dots \\ a_1 & a_3 & a_5\dots \\ 0 & a_0 & a_2\dots \\ \dots & \dots & \dots \end{vmatrix}=a_0^{2p}\begin{vmatrix} -s_1 & s_2 & -s_3\dots(-1)^ps_p \\ s_2 & -s_3 & s_4\dots \end{vmatrix}=$$

$$=(-1)^pa_0^{2p}D_p^{(1)},$$

откуда

$$\Delta_{2p}=a_0^{2p}D_p \qquad (p=1,\ 2,\ \dots,\ m). \tag{131′}$$

При нечетном $n=2m+1$ имеем:

$$\frac{g(u)}{h(u)}=\frac{a_0u^m+a_2u^{m-1}+\dots}{a_1u^m+a_3u^{m-1}+\dots}=s_{-1}+\frac{s_0}{u}-\frac{s_2}{u^2}+\dots, \tag{132}$$

откуда снова по формуле (68′) стр. 503 имеем:

$$\Delta_{2p}=\begin{vmatrix} a_1 & a_3\dots \\ a_0 & a_2\dots \\ 0 & a_1\dots \\ 0 & a_0\dots \\ \dots & \dots \end{vmatrix}=a_1^{2p}D_p \qquad (p=1,\dots,m).$$

С другой стороны, из (132) находим:

$$\left(\frac{g(u)}{h(u)}-s_{-1}\right)u=\frac{a_2'u^m+a_4'u^{m-1}+\dots}{a_1u^m+a_3u^{m-1}+\dots}=s_0-\frac{s_1}{u}+\frac{s_2}{u^2}-\dots, \tag{132′}$$

где

$$a_{2p}'=a_{2p}-s_{-1}-a_{2p+1} \quad (p=1,\dots,m); \qquad a_0'=0.$$

Но тогда для $p=1,\dots,m$

$$\nabla'_{2p}=\begin{vmatrix} a_1 & a_3 & a_5\dots \\ a_2' & a_4' & a_6'\dots \\ 0 & a_1 & a_3\dots \\ 0 & a_2' & a_4'\dots \\ \dots & \dots & \dots \end{vmatrix}=a_1^{2p}\begin{vmatrix} -s_1+s_2+\dots(-1)^ps_p \\ s_2-s_3\dots \\ \dots \end{vmatrix}=(-1)^pa_1^{2p}D_p^{(1)}. \tag{133}$$

---

[1]) Определители Гурвица Δ_k введены на стр. 485; определители Маркова задаются формулами (114). — *Прим. ред.*

С другой же стороны,

$$\nabla'_{2p}=(-1)^p\begin{vmatrix} a'_2 & a'_4 & \dots \\ a_1 & a_3 & \dots \\ 0 & a'_2 & \dots \\ 0 & a_1 & \dots \\ \dots & \dots & \dots \end{vmatrix}=\frac{(-1)^p}{a_1}\begin{vmatrix} a_1 & a_3 & a_5 & \dots \\ a'_0 & a'_2 & a'_4 & \dots \\ 0 & a_1 & a_3 & \dots \\ 0 & a'_0 & a'_2 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}. \qquad (133')$$

В полученном определителе $(2p+1)$-го порядка к каждой строке с четным номером прибавим предыдущую строку, помноженную предварительно на s_{-1}. Тогда этот определитель перейдет в Δ_{2p+1}. Поэтому из (133) и (133') получаем:

$$\Delta_{2p+1}=a_1^{2p+1}D_p^{(1)}.$$

Таким образом имеет место следующая связь между определителями Гурвица и Маркова:

а) при $n=2m$:

$$\Delta_{2p-1}=a_0^{2p-1}D_p \qquad (p=1, \dots, m);$$
$$\Delta_{2p}=a_0^{2p}D_p^{(1)}$$

б) при $n=2m+1$ (заменяя a_1 на a_0s_{-1}):

$$\Delta_{2p}=(a_0s_{-1})^{2p}D_p \qquad (p=1, \dots, m),$$
$$\Delta_{2p+1}=(a_0s_{-1})^{2p+1}D_p^{(1)} \qquad (p=0, 1, \dots, m).$$

Эти формулы показывают, как неравенства Маркова (115) переходят в неравенства Гурвица и наоборот. Кроме того, эти неравенства в соединении с критерием Льенара—Шипара дают нам следующую теорему:

Для того чтобы вещественный многочлен $f(z)=h(z^2)+zg(z^2)$ *со старшим коэффициентом* $a_0>0$ *был гурвицевым, необходимо и достаточно, чтобы*:

1) *все коэффициенты этого многочлена были положительны и*

2) *одна из квадратичных форм* (112) была положительно определенной.

§ 18. Теоремы Маркова и Чебышева

В своем известном мемуаре «О функциях, получаемых при обращении рядов в непрерывные дроби», напечатанном в Записках Петербургской Академии наук за 1894 г.[1]), покойный акад. А. А. Марков доказал две теоремы, из которых вторая иными методами и в не столь общей формулировке была в 1892 г. установлена П. Л. Чебышевым[2]).

В этом параграфе мы покажем, что эти теоремы имеют непосредственное отношение к исследованию области устойчивости в параметрах Маркова, и дадим сравнительно простое доказательство (не связанное с непрерывными дробями) этих теорем, опирающееся на теорему 19 предыдущего параграфа.

Переходя к формулировке первой теоремы, процитируем соответствующее место упомянутого выше мемуара А. А. Маркова[3]):

«Основываясь на предыдущем, нетрудно уже доказать две замечательные теоремы, которыми мы закончим нашу статью.

Одна касается определителей[4])

$$\Delta_1, \Delta_2, \dots, \Delta_m, \Delta^{(1)}, \Delta^{(2)}, \dots, \Delta^{(m)},$$

---

[1]) См. также [22], стр. 78—105.

[2]) Эта теорема была опубликована в работе П. Л. Чебышева «О разложении в непрерывную дробь рядов, расположенных по нисходящим степеням переменной». См. [31], стр. 307—362.

[3]) [25], стр. 95, третья строка снизу и далее.

[4]) В наших обозначениях $D_1, D_2, \dots, D_m, D_1^{(1)}, D_2^{(1)}, \dots, D_m^{(1)}$ (см. стр. 521).

а другая — корней уравнения [1])

$$\psi_m(x) = 0.$$

Теорема об определителях. *Если для чисел*

$$s_0,\ s_1,\ s_2,\ \ldots,\ s_{2m-2},\ s_{2m-1}$$

мы имеем две системы значений

1) $s_0 = a_0,\ s_1 = a_1,\ s_2 = a_2,\ \ldots,\ s_{2m-2} = a_{2m-2},\ s_{2m-1} = a_{2m-1},$

2) $s_0 = b_0,\ s_1 = b_1,\ s_2 = b_2,\ \ldots,\ s_{2m-2} = b_{2m-2},\ s_{2m-1} = b_{2m-1},$

при которых все определители

$$\Delta_1 = s_0,\quad \Delta_2 = \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix},\ \ldots,\ \Delta_m = \begin{vmatrix} s_0 & s_1 & \ldots & s_{m-1} \\ s_1 & s_2 & \ldots & s_m \\ \ldots & \ldots & \ldots & \ldots \\ s_{m-1} & s_m & & s_{2m-2} \end{vmatrix},$$

$$\Delta^{(1)} = s_1,\ \Delta^{(2)} = \begin{vmatrix} s_1 & s_2 \\ s_2 & s_3 \end{vmatrix},\ \ldots\ \Delta^{(m)} = \begin{vmatrix} s_1 & s_2 & \ldots & s_m \\ s_2 & s_3 & \ldots & s_{m+1} \\ \ldots & \ldots & \ldots & \ldots \\ s_m & s_{m+1} & & s_{2m-1} \end{vmatrix}$$

оказываются числами положительными и удовлетворены неравенства

$$a_0 \geqslant b_0,\ b_1 \geqslant a_1,\ a_2 \geqslant b_2,\ b_3 \geqslant a_3,\ \ldots,\ a_{2m-2} \geqslant b_{2m-2},\ b_{2m-1} \geqslant a_{2m-1},$$

то наши определители

$$\Delta_1,\ \Delta_2,\ \ldots,\ \Delta_m;\quad \Delta^{(1)},\ \Delta^{(2)},\ \ldots\ \Delta^{(m)}$$

должны оставаться числами положительными при всех значениях

$$s_0,\ s_1,\ s_2\ \ldots,\ s_{2m-1},$$

удовлетворяющих неравенствам

$$a_0 \geqslant s_0 \geqslant b_0,\ b_1 \geqslant s_1 \geqslant a_1,\ a_2 \geqslant s_2 \geqslant b_2,\ \ldots,$$
$$a_{2m-2} \geqslant s_{2m-2} \geqslant b_{2m-2},\ b_{2m-1} \geqslant s_{2m-1} \geqslant a_{2m-1}.$$

При тех же условиях должно быть

$$\begin{vmatrix} a_0 & a_1 & \ldots & a_{k-1} \\ a_1 & a_2 & \ldots & a_k \\ \ldots & \ldots & \ldots & \ldots \\ a_{k-1} & a_k & \ldots & a_{2k-2} \end{vmatrix} \geqslant \begin{vmatrix} s_0 & s_1 & \ldots & s_{k-1} \\ s_1 & s_2 & \ldots & s_k \\ \ldots & \ldots & \ldots & \ldots \\ s_{k-1} & s_k & \ldots & s_{2k-2} \end{vmatrix} \geqslant \begin{vmatrix} b_0 & b_1 & \ldots & b_{k-1} \\ b_1 & b_1 & \ldots & b_k \\ \ldots & \ldots & \ldots & \ldots \\ b_{k-1} & b_k & \ldots & b_{2k-2} \end{vmatrix}$$

и

$$\begin{vmatrix} b_1 & b_2 & \ldots & b_k \\ b_2 & b_3 & \ldots & b_{k+1} \\ \ldots & \ldots & \ldots & \ldots \\ b_k & b_{k+1} & \ldots & b_{2k-1} \end{vmatrix} \geqslant \begin{vmatrix} s_1 & s_2 & \ldots & s_k \\ s_2 & s_3 & \ldots & s_{k+1} \\ \ldots & \ldots & \ldots & \ldots \\ s_k & s_{k+1} & \ldots & s_{2k-1} \end{vmatrix} \geqslant \begin{vmatrix} a_1 & a_2 & \ldots & a_k \\ a_2 & a_3 & \ldots & a_{k+1} \\ \ldots & \ldots & \ldots & \ldots \\ a_k & a_{k+1} & \ldots & a_{2k-1} \end{vmatrix}$$

при $k = 1,\ 2,\ \ldots,\ m$».

Для того чтобы дать иную формулировку этой теоремы, связанную с задачей устойчивости, введем некоторые понятия и обозначения.

---

[1]) В наших обозначениях $h(-x) = 0$.

Параметры Маркова $s_0, s_1, \ldots, s_{2m-1}$ (при $n=2m$) или $s_{-1}, s_0, s_1, \ldots$ $\ldots, s_{2m-1}$ (при $n=2m+1$) будем рассматривать как координаты некоторой точки P n-мерного пространства. Область устойчивости в этом пространстве будем обозначать через G. Область G характеризуется неравенствами (115), (116) (стр. 521).

Мы будем говорить, что точка $P=\{s_i\}$ «предшествует» точке $P^*=\{s_i^*\}$, и писать $P \prec P^*$, если

$$\left.\begin{array}{c} s_0 \leqslant s_0^*,\ s_1^* \leqslant s_1,\ s_2 \leqslant s_2^*,\ s_3^* \leqslant s_3,\ \ldots,\ s_{2m-1}^* \leqslant s_{2m-1} \\ \text{и (при } n=2m+1) \\ s_{-1} \leqslant s_{-1}^* \end{array}\right\} \qquad (134)$$

и хотя бы в одном из этих соотношений имеет место знак $<$.

Если имеют место только соотношения (134) без последней оговорки, то мы будем писать:

$$P \preceq P^*.$$

Мы будем говорить, что *точка Q лежит «между» точками P и R*, если $R \prec Q \prec R$.

Каждой точке P соответствует бесконечная ганкелева матрица ранга m: $S=\| s_{i+k} \|_0^\infty$. Эту матрицу будем обозначать еще так: S_P.

Теперь мы дадим теореме Маркова следующую формулировку:

Теорема 21 (Маркова). *Если две точки P и R принадлежат области устойчивости G, то и любая точка Q, расположенная между точками P и R, также принадлежит области G, т. е.*

$$\text{из } P,\ R \in G,\ P \preceq Q \preceq R \text{ следует: } Q \in G.$$

Доказательство. Из $P \preceq Q \preceq R$ следует, что две точки P и R можно соединить отрезком кривой

$$s_i = (-1)_i\, \varphi_i(t)\ [\alpha \leqslant t \leqslant \gamma;\ i=0, 1, \ldots, 2m-1 \text{ и (при } n=2m+1)\, i=-1], \qquad (135)$$

содержащим точку Q так, чтобы 1) функции $\varphi_i(t)$ были непрерывными, монотонно возрастающими и дифференцируемыми при изменении t от $t=\alpha$ до $t=\gamma$ и 2) чтобы значениям α, β, γ $(\alpha<\beta<\gamma)$ аргумента t соответствовали точки P, Q, R на кривой.

При помощи величин (135) составим бесконечную ганкелеву матрицу $S=S(t)=\| s_{i+k} \|_0^\infty$ ранга m. Рассмотрим часть этой матрицы, а именно прямоугольную матрицу

$$\left\| \begin{array}{cccc} s_0 & s_1 & \ldots & s_m \\ s_1 & s_2 & \ldots & s_{m+1} \\ \ldots & \ldots & \ldots & \ldots \\ s_{m-1} & s_m & \ldots & s_{2m-1} \end{array} \right\|. \qquad (136)$$

Согласно условию теоремы при $t=\alpha$ и при $t=\gamma$ матрица $S(t)$ вполне положительна ранга m, и потому все миноры матрицы (136) порядка $p=1, 2, \ldots, m$ положительны.

Мы докажем, что это свойство сохраняется при любом промежуточном значении t $(\alpha<t<\gamma)$.

Для $p=1$ это очевидно. Докажем это утверждение для миноров p-го порядка в предположении, что оно верно для миноров $(p-1)$-го порядка. Рассмотрим произвольный минор p-го порядка, образованный

подряд идущими строками и столбцами матрицы (132):

$$D_p^{(q)} = \begin{vmatrix} s_q & s_{q+1} & \dots & s_{q+p-1} \\ s_{q+1} & s_{q+2} & \dots & s_{q+p} \\ \dots & \dots & \dots & \dots \\ s_{q+p-1} & s_{q+p} & \dots & s_{q+2p-2} \end{vmatrix} \qquad [q = 0,\ 1,\ \dots,\ 2(m-p)+1].$$

Вычислим производную от этого минора

$$\frac{d}{dt} D_p^{(q)} = \sum_{i,\,k=0}^{p-1} \frac{\partial D_p^{(q)}}{\partial s_{q+i+k}} \frac{ds_{q+i+k}}{dt}.$$

$\frac{\partial D_p^{(q)}}{\partial s_{q+i+k}}$ $(i,\ k = 0,\ 1,\ \dots,\ p-1)$ — алгебраические дополнения (адъюнкты) элементов определителя $D_p^{(q)}$. Поскольку согласно допущению все миноры этого определителя положительны, то

$$(-1)^{i+k} \frac{\partial D_p^{(q)}}{\partial s_{q+i+k}} > 0 \quad (i,\ k = 0, 1,\ \dots,\ p-1). \tag{137}$$

С другой стороны, из (131) находим:

$$(-1)^{q+i+k} \frac{ds_{q+i+k}}{dt} = \frac{d\varphi_{q+i+k}}{dt} \geqslant 0 \qquad (i,\ k = 0,\ 1,\ \dots,\ p-1). \tag{137'}$$

Из (134), (135) и (136) следует:

$$(-1)^q \frac{d}{dt} D_p^{(q)} \geqslant 0 \quad \begin{pmatrix} q = 0,\ 1,\ \dots,\ 2(m-p)+1, \\ p = 1,\ 2,\ \dots,\ m, \\ \alpha \leqslant t \leqslant \gamma \end{pmatrix}. \tag{137''}$$

Таким образом, каждый минор $D_p^{(q)}$ при q четном монотонно возрастает (точнее, не убывает), а при q нечетном — монотонно убывает (точнее, не возрастает) при возрастании аргумента t от значения $t = \alpha$ до значения $t = \gamma$, и поскольку при $t = \alpha$ и $t = \gamma$ этот минор положителен, то он будет положительным при любом промежуточном значении t $(\alpha < t < \gamma)$.

Из того, что положительны миноры матрицы (136) порядка $p-1$ и миноры p-го порядка, образованные подряд идущими строками и столбцами, уже следует, что все миноры p-го порядка матрицы (136) положительны[1]).

Из доказанного следует, что при любом t $(\alpha \leqslant t \leqslant \gamma)$ величины $s_0,\ s_1,\ \dots,\ s_{2m-1}$ и (при $n = 2m+1$) s_{-1} удовлетворяют неравенствам (115) и (116), т. е. при любом t эти величины являются параметрами Маркова для некоторого многочлена Гурвица. Другими словами, вся кривая (135), а значит, и точка Q лежат в области устойчивости G.

Теорема Маркова доказана.

Примечание. Поскольку доказано, что каждая точка кривой (135) принадлежит области G, то при любом t $(\alpha \leqslant t \leqslant \gamma)$ величины (135) определяют вполне положительную ранга m матрицу $S(t) = \| s_{i+k}(t) \|_0^\infty$. Поэтому неравенства (137), а следовательно, и (137'') имеют место при любом t $(\alpha \leqslant t \leqslant \beta)$, т. е. с возрастанием t любое $D_p^{(q)}$ возрастает, если q — четное число, и убывает, если q нечетно $[q = 0, 1,\ \dots,\ 2(m-p)+1;\ p = 1,\ \dots,\ m]$.

[1]) Это следует из детерминантного тождества Фекете (см. [7], стр. 306—307).

Другими словами, из $\boldsymbol{P} \preceq \boldsymbol{Q} \preceq \boldsymbol{R}$ следует:

$$(-1)^q D_p^{(q)}(\boldsymbol{P}) \leqslant (-1)^q D_p^{(q)}(\boldsymbol{Q}) \leqslant (-1)^q D_p^{(q)}(\boldsymbol{R})$$
$$[q = 0,\ 1,\ \ldots,\ 2(m-p)+1;\ p = 1,\ \ldots,\ m].$$

Эти неравенства при $q = 0,\ 1$ дают неравенства Маркова (стр. 527).

Рассмотрим теперь упомянутую в начале этого параграфа теорему Чебышева—Маркова. Снова приведем цитату из мемуара А. А. Маркова[1]):

«Теорема о корнях. *Если числа*

$$a_0,\ a_1,\ a_2,\ \ldots,\ a_{2m-2},\ a_{2m-1},$$
$$s_0,\ s_1,\ s_2,\ \ldots,\ s_{2m-2},\ s_{2m-1},$$
$$b_0,\ b_1,\ b_2,\ \ldots,\ b_{2m-2},\ b_{2m-1}$$

удовлетворяют всем условиям предыдущей теоремы[2]), *то уравнения*

$$\begin{vmatrix} a_0 & a_1 & \ldots & a_{m-1} & 1 \\ a_1 & a_2 & \ldots & a_m & x \\ a_2 & a_3 & \ldots & a_{m+1} & x^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_m & a_{m+1} & \ldots & a_{2m-1} & x^m \end{vmatrix} = 0,$$

$$\begin{vmatrix} s_0 & s_1 & \ldots & s_{m-1} & 1 \\ s_1 & s_2 & \ldots & s_m & x \\ s_2 & s_3 & \ldots & s_{m+1} & x^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ s_m & s_{m+1} & \ldots & s_{2m-1} & x^m \end{vmatrix} = 0,$$

$$\begin{vmatrix} b_0 & b_1 & \ldots & b_{m-1} & 1 \\ b_1 & b_2 & \ldots & b_m & x \\ b_2 & b_3 & \ldots & b_{m+1} & x^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ b_m & b_{m+1} & \ldots & b_{2m-1} & x^m \end{vmatrix} = 0$$

степени m относительно неизвестной x не имеют ни кратных, ни мнимых, ни отрицательных корней.

И корни второго уравнения больше соответственных корней первого и меньше соответственных корней последнего уравнения».

Выясним, в какой связи эта теорема находится с областью устойчивости в пространстве параметров Маркова. Полагая $f(z) = h(z^2) + zg(z^2)$ и

$$h(-v) = c_0 v^m + c_1 v^{m-1} + \ldots + c_m \qquad (c_0 \neq 0),$$

мы из разложения (105)

$$R(v) = -\frac{g(-v)}{h(-v)} = -s_{-1} + \frac{s_0}{v} + \frac{s_1}{v^2} + \ldots$$

получим тождество

$$-g(-v) = \left(-s_{-1} + \frac{s_0}{v} + \frac{s_1}{v^2} + \ldots\right)(c_0 v^m + c_1 v^{m-1} + \ldots + c_m).$$

---

[1]) См. [25], стр. 103, 5-я строка сверху и далее.

[2]) Имеется в виду предыдущая теорема Маркова—теорема об определителях (стр. 527).

Приравнивая нулю коэффициенты при степенях $v^{-1}, v^{-2}, \ldots, v^{-m}$, найдем:

$$\left.\begin{aligned} &s_0c_m + s_1c_{m-1} + \ldots + s_mc_0 = 0,\\ &s_1c_m + s_2c_{m-1} + \ldots + s_{m+1}c_0 = 0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ &s_{m-1}c_m + s_mc_{m-1} + \ldots + s_{2m-1}c_0 = 0; \end{aligned}\right\} \tag{138}$$

к этим соотношениям добавим уравнение

$$h(-v) = 0, \tag{139}$$

записанное так:

$$c_m + vc_{m-1} + \ldots + v^mc_0 = 0. \tag{139'}$$

Исключая из (138) и (139′) коэффициенты $c_0, c_1, \ldots, c_m$, мы представим уравнение (139) в виде

$$\begin{vmatrix} s_0 & s_1 & \ldots & s_{m-1} & 1\\ s_1 & s_2 & \ldots & s_m & v\\ s_2 & s_3 & \ldots & s_{m+1} & v^2\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ s_m & s_{m+1} & \ldots & s_{2m-1} & v^m \end{vmatrix} = 0. \tag{139''}$$

Таким образом, алгебраическое уравнение в теореме Чебышева—Маркова совпадает с уравнением (139), а неравенства, налагаемые на величины $s_0, s_1, \ldots, s_{2m-1}$, совпадают с неравенствами (115), определяющими область устойчивости в пространстве параметров Маркова.

Теорема Чебышева—Маркова выясняет, как меняются корни $u_1 = -v_1$, $u_2 = -v_2$, $\ldots$, $u_m = -v_m$ многочлена $h(u)$, когда соответствующие параметры Маркова $s_0, s_1, \ldots, s_{2m-1}$ изменяются в области устойчивости.

Первая часть теоремы утверждает известные нам факты: при выполнении неравенств (115) все корни $u_1, u_2, \ldots, u_m$ многочлена $h(u)$ простые, вещественные и отрицательные [1]). Мы эти корни будем обозначать так:

$$u_1(\boldsymbol{P}),\ u_2(\boldsymbol{P}),\ \ldots,\ u_m(\boldsymbol{P}),$$

где $\boldsymbol{P}$ — соответствующая точка области G.

Тогда вторая (основная) часть теоремы Чебышева—Маркова может быть сформулирована так:

Теорема 22 (Чебышева—Маркова). *Если $\boldsymbol{P}$ и $\boldsymbol{Q}$ — две точки области G и точка $\boldsymbol{P}$ «предшествует» точке $\boldsymbol{Q}$,*

$$\boldsymbol{P} \prec \boldsymbol{Q}, \tag{140}$$

то

$$u_1(\boldsymbol{P}) < u_1(\boldsymbol{Q}),\ u_2(\boldsymbol{P}) < u_2(\boldsymbol{Q}),\ \ldots,\ u_m(\boldsymbol{P}) < u_m(\boldsymbol{Q})\ ^{2}). \tag{141}$$

Доказательство. Коэффициенты многочлена $h(u)$ можно выразить рационально через параметры $s_0, s_1, \ldots, s_{2m-1}$ [3]). Тогда из

$$h(u_i) = 0 \qquad (i = 1, 2, \ldots, m)$$

---

[1]) См. теорему 13 на стр. 515.

[2]) Другими словами, корни $u_1, u_2, \ldots, u_m$ возрастают при возрастании $s_0, s_2, \ldots, s_{2m-2}$ и убывании $s_1, s_3, \ldots, s_{2m-1}$.

[3]) Хотя бы из уравнений (138), положив в них для конкретности $c_0 = 1$.

следует [1]):

$$\frac{\partial h(u_i)}{\partial s_l} + h'(u_i)\frac{du_i}{ds_l} = 0 \qquad (i = 1, 2, \ldots, m;\ l = 0, 1, \ldots, 2m-1). \tag{142}$$

С другой стороны, дифференцируя почленно разложение

$$\frac{g(u)}{h(u)} = s_{-1} + \frac{s_0}{u} - \frac{s_1}{u} + \frac{s_2}{u^3} - \ldots$$

по параметру s_l, найдем:

$$\frac{h(u)\dfrac{\partial g(u)}{\partial s_l} - g(u)\dfrac{\partial h(u)}{\partial s_l}}{h^2(u)} = \frac{(-1)^l}{u^{l+1}} + \frac{1}{u^{2m+1}}(*). \tag{143}$$

Помножая обе части этого равенства на многочлен $\dfrac{h^2(u)}{u-u_i}$ и обозначая через C_{il} коэффициент этого многочлена при степени u^l, получим:

$$\frac{h(u)}{u-u_i}\frac{\partial g(u)}{\partial s_l} - \frac{g(u)\dfrac{\partial h(u)}{\partial s_l}}{u-u_i} = \frac{(-1)^l C_{il}}{u} + \ldots \tag{144}$$

Приравнивая коэффициенты при $\dfrac{1}{u}$ (вычеты) в левой и правой частях равенства (144), найдем:

$$(-1)^{l-1} g(u_i)\frac{\partial h(u_i)}{\partial s_l} = C_{il}, \tag{145}$$

что в сочетании с (142) дает:

$$\frac{du_i}{ds_l} = \frac{(-1)^l C_{il}}{g(u_i)\,h'(u_i)}.$$

Вводя величины

$$R_i = \frac{g(u_i)}{h'(u_i)} \qquad (i = 1, 2, \ldots, m), \tag{146}$$

мы получим формулу Чебышева—Маркова

$$\frac{du_i}{ds_l} = \frac{(-1)^l C_{il}}{R_i\,[h'(u_i)]^2} \qquad (i = 1, 2, \ldots, m;\ l = 0, 1, \ldots, 2m-1). \tag{147}$$

Но в области устойчивости величины R_i $(i = 1, 2, \ldots, m)$ положительны [см. (90′) на стр. 513]. То же можно сказать и о коэффициентах C_{il}. Действительно,

$$\frac{h^2(u)}{u-u_i} = c_0^2 (u+v_1)^2 \ldots (u+v_{i-1})^2 (u+v_i)(u+v_{i+1})^2 \ldots (u+v_m)^2, \tag{148}$$

где

$$v_i = -u_i > 0 \qquad (i = 1, 2, \ldots, m).$$

Из (148) видно, что все коэффициенты C_{il} в разложении $\dfrac{h^2(u)}{u-u_i}$ по степени u положительны. Таким образом, из формулы Чебышева—Маркова получаем:

$$(-1)^l \frac{du_i}{ds_l} > 0. \tag{149}$$

---

[1]) Здесь $\dfrac{\partial h(u_i)}{\partial s_l} = \left[\dfrac{\partial h(u)}{\partial s_l}\right]_{u=u_i}$.

При доказательстве теоремы Маркова мы показали, что любые две точки $P \prec Q$ области G можно соединить дугой кривой $s_l = (-1)^l \varphi_l(t)$ $(l = 0, 1, \ldots, 2m-1)$, где $\varphi_l(t)$ — монотонно возрастающие дифференцируемые функции от t [t изменяется в пределах от α до β $(\alpha < \beta)$, причем $t = \alpha$ соответствует точка P, а $t = \beta$ — точка Q]. Тогда вдоль этой кривой в силу (149) [1])

$$\frac{du_l}{dt} = \sum_{l=0}^{2m-1} \frac{\partial u_i}{\partial s_l} \frac{ds_l}{dt} \geqslant 0, \qquad \frac{du_i}{dt} \not\equiv 0 \quad (\alpha \leqslant t \leqslant \beta). \tag{150}$$

Отсюда, интегрируя, получаем:

$$u_{i_{(t=\alpha)}} = u_i(P) < u_{i_{(t=\beta)}} = u_i(Q) \qquad (i = 1, 2, \ldots, m).$$

Теорема Чебышева — Маркова доказана.

§ 19. Обобщенная задача Рауса—Гурвица

В этом параграфе мы дадим правило определения числа корней в правой полуплоскости для многочлена $f(z)$ с комплексными коэффициентами. Пусть

$$f(iz) = b_0 z^n + b_1 z^{n-1} + \ldots + b_n + i(a_0 z^n + a_1 z^{n-1} + \ldots + a_n), \tag{151}$$

где $a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_n$ — вещественные числа. Если n есть степень многочлена $f(z)$, то $b_0 + ia_0 \not\equiv 0$. Не нарушая общности, можем считать, что $a_0 \not\equiv 0$ [в противном случае мы бы заменили многочлен $f(z)$ на $if(z)$].

Мы будем предполагать, что вещественные многочлены

$$a_0 z^n + a_1 z^{n-1} + \ldots + a_n \text{ и } b_0 z^n + b_1 z^{n-1} + \ldots + b_n \tag{152}$$

взаимно просты, т. е. что результант этих многочленов отличен от нуля [2]):

$$\nabla_{2n} = \begin{vmatrix} a_0 & a_1 & \ldots & a_n & 0 & \ldots & 0 \\ b_0 & b_1 & \ldots & b_n & 0 & \ldots & 0 \\ 0 & a_0 & \ldots & a_{n-1} & a_n & \ldots & 0 \\ 0 & b_0 & \ldots & b_{n-1} & b_n & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix} \not\equiv 0. \tag{153}$$

Отсюда следует, в частности, что многочлены (152) не имеют общих вещественных корней и что, следовательно, многочлен $f(z)$ не имеет корней на мнимой оси.

Обозначим через k число корней $f(z)$, имеющих положительные вещественные части. Рассматривая область в правой полуплоскости, ограниченную мнимой осью и полуокружностью радиуса R $(R \to \infty)$, и повторяя дословно рассуждения, приведенные на стр. 472 для вещественного многочлена $f(z)$, получим формулу для приращения $\arg f(z)$ вдоль мнимой оси

$$\Delta_{-\infty}^{+\infty} \arg f(z) = (n - 2k)\pi. \tag{154}$$

---

[1]) Поскольку $(-1)^l \frac{ds_l}{dt} = \frac{d\varphi_l}{dt} \geqslant 0$ $(\alpha \leqslant t \leqslant \beta)$, причем хотя бы для одного l существуют такие значения t, при которых $(-1)^l \frac{ds_l}{dt} > 0$.

[2]) ∇_{2n} — определитель порядка $2n$.

Отсюда в силу (151) и условия $a_0 \neq 0$ получаем:

$$I_{-\infty}^{+\infty} \frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + z^n} = n - 2k. \tag{155}$$

Пользуясь теоремой 10 § 11 (стр. 504), отсюда получаем:

$$k = V(1, \nabla_2, \nabla_4, \ldots, \nabla_{2n}), \tag{156}$$

где

$$\nabla_{2p} = \begin{vmatrix} a_0 & a_1 & \ldots & a_{2p-1} \\ b_0 & b_1 & \ldots & b_{2p-1} \\ 0 & a_0 & \ldots & a_{2p-2} \\ 0 & b_0 & \ldots & b_{2p-2} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix} \quad (p = 1, 2, \ldots, n;\ a_k = b_k = 0 \text{ при } k > n). \tag{157}$$

Мы пришли к теореме:

Теорема 23. *Если дан комплексный многочлен $f(z)$, для которого $f(iz) = b_0 z^n + b_1 z^{n-1} + \ldots + b_n + i(a_0 z^n + a_1 z^{n-1} + \ldots + a_n)$ $(a_0 \neq 0)$, причем многочлены $a_0 z^n + \ldots + a_n$ и $b_0 z^n + \ldots + b_n$ взаимно просты $(\nabla_{2n} \neq 0)$, то число корней многочлена $f(z)$, расположенных в правой полуплоскости, определяется формулами* (156), (157).

При этом [1]), если среди определителей (157) имеются равные нулю, то для каждой группы подряд идущих нулей

$$(\nabla_{2h} \neq 0)\ \nabla_{2h+2} = \ldots = \nabla_{2h+2p} = 0\ (\nabla_{2h+2p+2} \neq 0) \tag{158}$$

при подсчете $V(1, \nabla_2, \nabla_4, \ldots, \nabla_{2n})$ следует положить:

$$\operatorname{sign} \nabla_{2h+2j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} \nabla_{2h} \qquad (j = 1, 2, \ldots, p), \tag{159}$$

или, что то же,

$$V(\nabla_{2h}, \nabla_{2h+2}, \ldots, \nabla_{2h+2p}, \nabla_{2h+2p+2}) = \begin{cases} \dfrac{p+1}{2} \text{ при } p \text{ нечетном}, \\ \dfrac{p+1-\varepsilon}{2} \text{ при } p \text{ четном}, \ \varepsilon = (-1)^{\frac{p}{2}} \operatorname{sign} \dfrac{\nabla_{2h+1p+2}}{\nabla_{2h}}. \end{cases} \tag{160}$$

Предоставляем самому читателю проверить, что в частном случае, когда $f(z)$ — вещественный многочлен из теоремы 23, можно получить теорему Рауса — Гурвица (см. § 6) [2]).

В заключение отметим, что в этой главе были рассмотрены приложения квадратичных форм (в частности, ганкелевых форм) к одной задаче распределения корней многочлена в комплексной плоскости, к задаче Рауса — Гурвица. Между тем квадратичные и эрмитовы формы имеют интересные приложения и к другим задачам распределения корней. Читателя, интересующегося этими вопросами, мы отошлем к цитированному нами уже обзору М. Г. Крейна и М. А. Наймарка «Метод симметрических и эрмитовых форм в теории отделения корней алгебраических уравнений», Харьков, 1936.

[1]) См. стр. 503.

[2]) Удобные алгоритмы для решения обобщенной задачи Рауса—Гурвица можно найти в монографии [27] и статье [26]. См. также [37].

ДОБАВЛЕНИЕ

НЕРАВЕНСТВА ДЛЯ СОБСТВЕННЫХ И СИНГУЛЯРНЫХ ЧИСЕЛ

В. Б. Лидский

Ниже рассматриваются неравенства, которым удовлетворяют собственные и сингулярные числа линейных операторов в n-мерном унитарном пространстве.

Основное внимание уделяется неравенствам Неймана — Хорна и Вейля (§§ 2 и 3), которые позволяют оценивать собственные числа оператора посредством его сингулярных чисел.

В § 4 устанавливается максимально-минимальное свойство сумм и произведений собственных чисел эрмитовых операторов, обнаруженное Виландтом и Амир-Моэзом.

Результаты § 4 используются далее в § 5 для доказательства неравенств, содержащих оценку собственных и сингулярных чисел операторов $\boldsymbol{A}+\boldsymbol{B}$ и $\boldsymbol{AB}$.

В § 6 рассматривается задача о собственных числах суммы и произведения эрмитовых операторов в постановке И. М. Гельфанда *).

§ 1. Мажорирующие последовательности

В этом параграфе мы остановимся на ряде вспомогательных вопросов, связанных с конечными числовыми последовательностями. Рассмотрим две убывающие последовательности чисел, по n элементов в каждой:

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n, \tag{1}$$

$$\alpha_1' \geqslant \alpha_2' \geqslant \ldots \geqslant \alpha_n'. \tag{2}$$

Принято говорить, что последовательность (2) мажорируется последовательностью (1), если

$$\alpha_1'+\alpha_2'+\ldots+\alpha_m' \leqslant \alpha_1+\alpha_2+\ldots+\alpha_m \quad (1 \leqslant m \leqslant n-1) \tag{3}$$

и

$$\alpha_1'+\alpha_2'+\ldots+\alpha_n' = \alpha_1+\alpha_2+\ldots+\alpha_n. \tag{3'}$$

При выполнении условий (3) и (3′) пишут

$$\alpha' < \alpha. \tag{4}$$

*) Автор приносит благодарность А. С. Маркусу, прочитавшему рукопись настоящей статьи и сделавшему ряд полезных замечаний.

Квадратную матрицу $T=\|t_{ij}\|_1^n$ мы будем в дальнейшем называть двояко стохастической, если матрицы T и T' являются стохастическими, другими словами, если $t_{ij}\geqslant 0$,

$$\sum_{j=1}^{n} t_{ij}=1, \quad 1\leqslant i\leqslant n, \tag{5}$$

и

$$\sum_{i=1}^{n} t_{ij}=1, \quad 1\leqslant j\leqslant n. \tag{5'}$$

Справедливо следующее утверждение (см. [35], стр. 63).

Лемма 1. *Последовательность* α' *мажорируется последовательностью* α *тогда и только тогда, когда существует двояко стохастическая матрица* T *такая, что*

$$\alpha'=T\alpha. \tag{6}$$

Достаточность условия (6) доказывается легко*). В самом деле,

$$\sum_{k=1}^{m}\alpha_k'=\sum_{k=1}^{m}\sum_{j=1}^{n}t_{kj}\alpha_j=\sum_{j=1}^{n}\Big(\sum_{k=1}^{m}t_{kj}\Big)\alpha_j=\sum_{j=1}^{n}\omega_j\alpha_j. \tag{7}$$

Мы положили

$$\omega_j=\sum_{k=1}^{m}t_{kj} \quad (1\leqslant j\leqslant n). \tag{8}$$

Легко видеть, что $0\leqslant\omega_j\leqslant 1$ и

$$\sum_{j=1}^{n}\omega_j=\sum_{k=1}^{m}\Big(\sum_{j=1}^{n}t_{kj}\Big)=m. \tag{9}$$

Имеем на основании равенства (7)

$$\sum_{j=1}^{m}\alpha_j-\sum_{k=1}^{m}\alpha_k'=\sum_{j=1}^{m}\alpha_j-\sum_{j=1}^{n}\omega_j\alpha_j=$$

$$=\alpha_1(1-\omega_1)+\ldots+\alpha_m(1-\omega_m)-\omega_{m+1}\alpha_{m+1}-\ldots-\omega_n\alpha_n. \tag{10}$$

Уменьшая слагаемые в правой части, получаем

$$\sum_{j=1}^{m}\alpha_j-\sum_{j=1}^{m}\alpha_j'\geqslant\alpha_m(1-\omega_1)+\ldots+\alpha_m(1-\omega_m)-\omega_{m+1}\alpha_m-\ldots$$

$$\ldots-\omega_n\alpha_m=\alpha_m(m-m)=0. \tag{10'}$$

Следовательно, неравенства (3) имеют место. Так как, далее, при $m=n$ согласно (8) $\omega_j=1$ $(j=1, 2, \ldots, n)$, то в силу (7) справедливо и равенство (3').

Таким образом, достаточность условия (6) установлена. Доказательство необходимости этого условия требует известных усилий. Мы проведем его по индукции **). В случае $n=1$ последовательности содержат по

*) В формуле (6) под α и α' следует понимать столбцевые матрицы с элементами (1) и (2).

**) В книге Харди, Литтлвуда и Пойа «Неравенства» приводится доказательство, основанное на другой идее. Настоящее, более короткое, доказательство принадлежит А. С. Маркусу (см. [113]).

одному элементу, $\alpha_1' = \alpha_1$, и матрица T, очевидно, существует. Предположим, что утверждение справедливо для случая последовательностей из $n-1$ элементов и рассмотрим две последовательности α' и α, которые связаны соотношением $\alpha' < \alpha$ и состоят из n элементов.

Из условия $\alpha_1' \leqslant \alpha_1$ и равенства (3') следует, что $\alpha_n \leqslant \alpha_1' \leqslant \alpha_1$. Поэтому найдется такое k $(1 \leqslant k \leqslant n-1)$, при котором

$$\alpha_{k+1} \leqslant \alpha_1' \leqslant \alpha_k. \tag{11}$$

Следовательно, при некотором τ, $0 \leqslant \tau \leqslant 1$, мы имеем

$$\alpha_1' = \tau\alpha_k + (1-\tau)\,\alpha_{k+1}. \tag{12}$$

Наряду с α' и α рассмотрим две последовательности по $n-1$ элементов в каждой:

$$\alpha_2', \alpha_3', \ldots, \alpha_k', \alpha_{k+1}', \alpha_{k+2}', \ldots, \alpha_n' \tag{13}$$

и

$$\alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \alpha_k + \alpha_{k+1} - \alpha_1', \alpha_{k+2}, \ldots, \alpha_n. \tag{13'}$$

Обозначим эти последовательности через $\widetilde{\alpha}'$ и $\widetilde{\alpha}$ соответственно.

Учитывая (11), легко заключить, что элементы последовательности $\widetilde{\alpha}$ расположены в порядке убывания. Без труда проверяется также соотношение $\widetilde{\alpha}' < \widetilde{\alpha}$. Поэтому в силу индуктивного предположения существует такая двояко стохастическая матрица $\widetilde{T} = \| t_{ij} \|_1^{n-1}$, что $\widetilde{\alpha}' = \widetilde{T}\widetilde{\alpha}$, или в развернутой записи:

$$\alpha_{s+1}' = t_{s1}\alpha_1 + \ldots + t_{s,k-1}\alpha_{k-1} + t_{sk}(\alpha_k + \alpha_{k+1} - \alpha_1') + \\ + t_{s,k+1}\alpha_{k+2} + \ldots + t_{s,n-1}\alpha_n \qquad (1 \leqslant s \leqslant n-1).$$

Подставив сюда α_1' из равенства (12), получим при $1 \leqslant s \leqslant n-1$:

$$\alpha_{s+1}' = t_{s1}\alpha_1 + \ldots + t_{sk}(1-\tau)\,\alpha_k + t_{sk}\tau\alpha_{k+1} + t_{s,k+1}\alpha_{k+2} + \ldots + t_{s,n-1}\alpha_n.$$

Добавляя сюда равенство $\alpha_1' = \tau\alpha_k + (1-\tau)\,\alpha_{k+1}$, легко убеждаемся в том, что последовательности α' и α связаны двояко стохастической матрицей

$$T = \begin{Vmatrix} 0 & 0 & 0 & \ldots & \tau & 1-\tau & \ldots & 0 \\ t_{11} & t_{12} & t_{13} & \ldots & t_{1k}(1-\tau) & t_{1k}\tau & \ldots & t_{1,n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ t_{n-1,1} & t_{n-1,2} & t_{n-1,3} & \ldots & t_{n-1,k}(1-\tau) & t_{n-1,k}\tau & \ldots & t_{n-1,n-1} \end{Vmatrix}.$$

Лемма доказана полностью.

Нам понадобится ниже также следующее предложение (см. [233]):

Лемма 2. *Пусть $\varphi(t)$ — непрерывная выпуклая*) монотонно возрастающая функция. Пусть*

$$\alpha_1' \geqslant \alpha_2' \geqslant \ldots \geqslant \alpha_p', \tag{14}$$

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_p \tag{15}$$

и

$$\alpha_1' + \alpha_2' + \ldots + \alpha_m' \leqslant \alpha_1 + \alpha_2 + \ldots + \alpha_m \quad (1 \leqslant m \leqslant p). \tag{16}$$

*) Функция $\varphi(t)$ называется выпуклой на интервале, если для любых точек этого интервала $\varphi\left(\frac{x+y}{2}\right) \leqslant \frac{1}{2}[\varphi(x) + \varphi(y)]$.

Тогда

$$\varphi(\alpha'_1)+\varphi(\alpha'_2)+\ldots+\varphi(\alpha'_p)\leqslant\varphi(\alpha_1)+\varphi(\alpha_2)+\ldots+\varphi(\alpha_p). \qquad (17)$$

Доказательство. Предположим сначала, что при $m=p$ в соотношении (16) имеет место равенство. Тогда последовательность α' мажорируется последовательностью α и согласно лемме 1

$$\alpha'_s=\sum_{j=1}^{p} t_{sj}\alpha_j, \qquad 1\leqslant s\leqslant p, \qquad (18)$$

где t_{sj} — элементы двояко стохастической матрицы. В силу выпуклости $\varphi(t)$ из равенства (18) следует *), что

$$\varphi(\alpha'_s)\leqslant\sum_{j=1}^{p} t_{sj}\varphi(\alpha_j). \qquad (19)$$

Суммируя неравенства (19), получаем

$$\sum_{s=1}^{p}\varphi(\alpha'_s)\leqslant\sum_{j=1}^{p}\Big(\sum_{s=1}^{p} t_{sj}\Big)\varphi(\alpha_j)=\sum_{j=1}^{p}\varphi(\alpha_j). \qquad (20)$$

Таким образом, в указанном случае неравенство (17) выполняется.

Рассмотрим теперь общий случай. Пусть в соотношении (16) при $m=p$ имеет место знак $<$. Положим

$$\sum_{j=1}^{p}\alpha_j-\sum_{j=1}^{p}\alpha'_j=c>0.$$

Наряду с последовательностями (14) и (15) рассмотрим две последовательности:

$$\alpha'_1\geqslant\alpha'_2\geqslant\ldots\geqslant\alpha'_p\geqslant\alpha'_{p+1} \qquad (21)$$

и

$$\alpha_1\geqslant\alpha_2\geqslant\ldots\geqslant\alpha_p\geqslant\alpha_{p+1}, \qquad (22)$$

где α'_{p+1} и α_{p+1} — произвольные два числа, удовлетворяющие неравенствам (21) и (22) и соотношению

$$\alpha_{p+1}=\alpha'_{p+1}-c. \qquad (23)$$

Легко видеть, что при таком выборе α'_{p+1} и α_{p+1} последовательность (21) мажорируется последовательностью (22), и по доказанному имеем

$$\varphi(\alpha'_1)+\varphi(\alpha'_2)+\ldots+\varphi(\alpha'_p)+\varphi(\alpha'_{p+1})\leqslant$$
$$\leqslant\varphi(\alpha_1)+\varphi(\alpha_2)+\ldots+\varphi(\alpha_p)+\varphi(\alpha_{p+1}). \qquad (24)$$

Так как, далее, $\varphi(t)$ — монотонно возрастающая функция и $\alpha'_{p+1}>\alpha_{p+1}$, то $\varphi(\alpha'_{p+1})\geqslant\varphi(\alpha_{p+1})$ и из (24) снова следует неравенство (17).

Лемма доказана полностью.

Замечание. Из наших рассуждений следует, что в том случае, когда последовательность (14) мажорируется последовательностью (15) [т. е. при $m=p$ в (16) достигается равенство], то неравенство (17) справедливо для любой непрерывной выпуклой функции $\varphi(t)$ (возрастание является излишним требованием).

---

*) Доказательство неравенства (19) для непрерывных выпуклых функций проводится по индукции (см. [35], стр. 93).

§ 2. Неравенства Неймана — Хорна

Пусть A — линейный оператор, действующий в n-мерном унитарном пространстве R. Собственные числа неотрицательного эрмитова оператора *) $\sqrt{A^*A}$ (см. стр. 249) принято называть *сингулярными* числами оператора A.

В настоящем параграфе мы установим неравенства, связывающие сингулярные числа произведения двух операторов с сингулярными числами сомножителей.

Пусть $y_1, y_2, \ldots, y_m$ и $z_1, z_2, \ldots, z_m$ — два набора векторов из R. Введем сокращенное обозначение для определителя порядка m, связанного с данными наборами:

$$[(y_i, z_j)] = \begin{vmatrix} (y_1, z_1) & (y_2, z_1) & \ldots & (y_m, z_1) \\ (y_1, z_2) & (y_2, z_2) & \ldots & (y_m, z_2) \\ \ldots & \ldots & \ldots & \ldots \\ (y_1, z_m) & (y_2, z_m) & \ldots & (y_m, z_m) \end{vmatrix}. \tag{25}$$

Рассмотрим далее неотрицательный эрмитов оператор H, действующий в R. Собственные значения оператора H занумеруем в убывающем порядке:

$$h_1 \geqslant h_2 \geqslant \ldots \geqslant h_n \geqslant 0. \tag{26}$$

Справедливо следующее предложение, принадлежащее А. Хорну ([191b]):

Лемма 3. *Пусть*

$$x_1, x_2, \ldots, x_m \qquad (m \leqslant n) \tag{27}$$

— произвольный набор векторов из R. Тогда **)

$$[(Hx_i, x_j)] \leqslant h_1 h_2 \ldots h_m [(x_i, x_j)]. \tag{28}$$

Для доказательства рассмотрим ортонормированный базис собственных векторов оператора H:

$$e_1, e_2, \ldots, e_n \tag{29}$$

и разложим каждый из векторов x_i $(i = 1, 2, \ldots, m)$ по базису (29). Вычисляя скалярное произведение, получаем:

$$(Hx_i, x_j) = \sum_{s=1}^{n} h_s (x_i, e_s) \overline{(x_j, e_s)} = \sum_{s=1}^{n} h_s (x_i, e_s)(e_s, x_j), \tag{30}$$
$$i, j = 1, 2, \ldots, m.$$

Равенство (30) позволяет рассматривать матрицу определителя $[(Hx_i, x_j)]$ как результат умножения двух прямоугольных матриц размеров $m \times n$ и $n \times m$.

Разлагая определитель по формуле Бине — Коши (см. стр. 20), получаем в принятых обозначениях для определителей:

$$[(Hx_i, x_j)] = \sum_{1 \leqslant s_1 < s_2 < \ldots < s_m \leqslant n} [(x_i, h_s e_s)] [(e_s, x_j)]. \tag{31}$$

---

*) Мы будем пользоваться также обозначением $(A^*A)^{1/2}$.

**) Определитель, стоящий в левой части (28), неотрицателен. Действительно, полагая $H^{1/2}x_i = y_i$, получаем

$$[(Hx_i, x_j)] = [(H^{1/2}x_i, H^{1/2}x_j)] = [(y_i, y_j)] \geqslant 0.$$

Здесь

$$[(x_i,\ h_s e_s)] = \begin{vmatrix} (x_1,\ h_{s_1}e_{s_1}) & \dots & (x_m,\ h_{s_1}e_{s_1}) \\ (x_1,\ h_{s_2}e_{s_2}) & \dots & (x_m,\ h_{s_2}e_{s_2}) \\ \dots\dots\dots\dots\dots & & \\ (x_1,\ h_{s_m}e_{s_m}) & \dots & (x_m,\ h_{s_m}e_{s_m}) \end{vmatrix}, \qquad (31')$$

а суммирование ведется по всевозможным наборам натуральных чисел $1 \leqslant s_1 < s_2 < \dots < s_m \leqslant n$.

Оценив правую часть (31) по неравенству Коши—Буняковского, получим:

$$[(\boldsymbol{H}x_i, x_j)]^2 \leqslant \Big(\sum_{1\leqslant s_1<\dots<s_m\leqslant n} |[(x_i, h_s e_s)]|^2\Big)\Big(\sum_{1\leqslant s_1<\dots<s_m\leqslant n} |[e_s,\ x_j]|^2\Big). \qquad (32)$$

Вторая сумма в правой части неравенства (32) равна определителю Грама $[(x_i,\ x_j)]$. В этом легко убедиться, положив в формуле (31) $\boldsymbol{H} = \boldsymbol{E}$, где $\boldsymbol{E}$ — единичный оператор. Впрочем, соответствующее равенство отдельно доказано на стр. 230 (формула (26)).

В первой сумме правой части (32) вынесем из каждого определителя (31′) произведение $h_{s_1}h_{s_2}\dots h_{s_m}$ и заменим его бо́льшим $h_1h_2\dots h_m$. В результате получим:

$$[(\boldsymbol{H}x_i,\ x_j)]^2 \leqslant h_1^2h_2^2\dots h_m^2\,[(x_i,\ x_j)]^2.$$

Извлекая из обеих частей этого неравенства квадратные корни, мы устанавливаем справедливость неравенства (28). Докажем далее следующий факт.

Лемма 4. *Пусть $\boldsymbol{K}$ — произвольный оператор в $\boldsymbol{R}$ и*

$$\varkappa_1 \geqslant \varkappa_2 \geqslant \dots \geqslant \varkappa_n \qquad (33)$$

— его сингулярные числа. Тогда для произвольного набора векторов x_1, x_2 ..., x_m $(m \leqslant n)$ справедливо неравенство

$$[(\boldsymbol{K}x_i,\ \boldsymbol{K}x_j)] \leqslant \varkappa_1^2\varkappa_2^2\dots\varkappa_m^2\,[(x_i,\ x_j)]. \qquad (34)$$

Неравенство (34) немедленно следует из леммы 3 при $\boldsymbol{H} = \boldsymbol{K}^*\boldsymbol{K}$.

Установим, наконец, еще одно вспомогательное предложение.

Лемма 5. *Пусть $\boldsymbol{A}$ и $\boldsymbol{B}$ — линейные операторы в $\boldsymbol{R}$, $\boldsymbol{C} = \boldsymbol{AB}$, и пусть α_s, β_s и γ_s $(s = 1, 2, \dots, n)$ — сингулярные числа соответственно $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$, занумерованные в убывающем порядке. Тогда при любом $m \leqslant n$ справедливы неравенства* *)

$$\gamma_1\gamma_2\dots\gamma_m \leqslant \alpha_1\alpha_2\dots\alpha_m\beta_1\beta_2\dots\beta_m. \qquad (35)$$

Для доказательства рассмотрим ортонормированный базис $e_1, e_2, \dots, e_n$ собственных векторов оператора $\boldsymbol{C}^*\boldsymbol{C}$. Последовательно применяя (34), получаем

$$[(\boldsymbol{C}e_i,\ \boldsymbol{C}e_j)] = [(\boldsymbol{AB}e_i,\ \boldsymbol{AB}e_j)] \leqslant \alpha_1^2\alpha_2^2\dots\alpha_m^2\,[(\boldsymbol{B}e_i,\ \boldsymbol{B}e_j)] \leqslant$$
$$\leqslant \alpha_1^2\alpha_2^2\dots\alpha_m^2\beta_1^2\beta_2^2\dots\beta_m^2\,[(e_i,\ e_j)]. \qquad (36)$$

---

*) При $m = n$ в формуле (35) достигается равенство. Действительно, имеем $\boldsymbol{C}^*\boldsymbol{C} = \boldsymbol{B}^*\boldsymbol{A}^*\boldsymbol{AB}$, отсюда $|\boldsymbol{C}^*\boldsymbol{C}| = |\boldsymbol{A}^*\boldsymbol{A}|\,|\boldsymbol{B}^*\boldsymbol{B}|$.

Поскольку определитель матрицы оператора равен произведению его собственных чисел, то $\gamma_1^2\gamma_2^2\dots\gamma_n^2 = \alpha_1^2\alpha_2^2\dots\alpha_n^2\beta_1^2\beta_2^2\dots\beta_n^2$.

С другой стороны, поскольку e_i $(1 \leqslant i \leqslant m)$ — собственные векторы $\boldsymbol{C}^*\boldsymbol{C}$, мы имеем:

$$[(\boldsymbol{C}e_i,\ \boldsymbol{C}e_j)] = [(\boldsymbol{C}^*\boldsymbol{C}e_i,\ e_j)] = \gamma_1^2\gamma_2^2 \ldots \gamma_n^2 [(e_i,\ e_j)]. \tag{37}$$

Следовательно, (35) имеет место.

Мы в состоянии теперь доказать следующую теорему, которая является основной целью настоящего параграфа.

Теорема 1 (Нейман — Хорн [218b, 191b]). *Пусть $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$ — линейные операторы в n-мерном унитарном пространстве $\boldsymbol{R}$. Пусть $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ и пусть α_s, β_s и γ_s $(s = 1, 2, \ldots, n)$ — сингулярные числа операторов $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$, занумерованные в порядке убывания.*

Пусть $f(x)$ — непрерывная при $x \geqslant 0$ функция такая, что $\varphi(t) = f(e^t)$ — монотонно возрастающая выпуклая функция параметра t. Тогда при всех $m \leqslant n$ справедливы неравенства *)

$$\sum_{s=1}^{m} f(\gamma_s) \leqslant \sum_{s=1}^{m} f(\alpha_s\beta_s). \tag{38}$$

Доказательство. Пусть сначала операторы $\boldsymbol{A}$ и $\boldsymbol{B}$ невырождены, тогда все числа α_s, β_s и γ_s положительны. Логарифмируя неравенства (35), получаем

$$\sum_{s=1}^{m} \ln \gamma_s \leqslant \sum_{s=1}^{m} \ln(\alpha_s\beta_s), \qquad 1 \leqslant m \leqslant n. \tag{39}$$

На основании леммы 2 имеем

$$\sum_{s=1}^{m} \varphi(\ln \gamma_s) \leqslant \sum_{s=1}^{m} \varphi(\ln \alpha_s\beta_s), \qquad 1 \leqslant m \leqslant n. \tag{40}$$

Так как $\varphi(t) = f(e^t)$, то отсюда следует (38). В случае вырожденных операторов неравенства (38) устанавливаются по непрерывности.

Замечание 1°. В случае $f(x) = x^\sigma$ $(\sigma \geqslant 0)$ получаем

$$\sum_{s=1}^{m} \gamma_s^\sigma \leqslant \sum_{s=1}^{m} \alpha_s^\sigma\beta_s^\sigma \qquad (1 \leqslant m \leqslant n). \tag{41}$$

В таком виде неравенства (38) встречаются в приложениях чаще всего.

Замечание 2°. При $m = n$ неравенство (39) превращается в равенство (см. сноску на стр. 540). Поэтому при $m = n$ неравенство (40) справедливо для любой непрерывной выпуклой функции $\varphi(t)$ (см. замечание к лемме 2).

В частности, неравенство (41) при $m = n$ справедливо и для $\sigma < 0$.

Замечание 3°. Пусть

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n$$

— сингулярные числа оператора $\boldsymbol{A}$ и пусть

$$\alpha_{l,1} \geqslant \alpha_{l,2} \geqslant \ldots \geqslant \alpha_{l,n}$$

*) Неравенства (38) установлены А. Хорном в работе [191b]. В работе Дж. Неймана [218b] доказано лишь, что $\sum_{s=1}^{n} \gamma_s^2 \leqslant \sum_{s=1}^{n} \alpha_s^2\beta_s^2$, однако развитый там метод позволяет доказать неравенства (38) в общем виде.

— сингулярные числа оператора A^l (l — натуральное число). Тогда при любом $\sigma \geqslant 0$ и любом $1 \leqslant m \leqslant n$

$$\sum_{s=1}^{m} \alpha_{l,s}^{\sigma} \leqslant \sum_{s=1}^{m} \alpha_s^{\sigma l}. \tag{42}$$

Неравенства (42) докажем индукцией по l. При $l=1$ соотношение (42) очевидно; пусть оно выполняется для $l-1$. Так как $A^l = A^{l-1} \cdot A$, то согласно (41)

$$\sum_{s=1}^{m} \alpha_{l,s}^{\sigma} \leqslant \sum_{s=1}^{m} \alpha_{l-1,s}^{\sigma} \alpha_s^{\sigma}, \qquad 1 \leqslant m \leqslant n. \tag{43}$$

Применяя к правой части (43) неравенство Гёльдера*) с

$$p = \frac{l}{l-1} \text{ и } q = l \quad (p^{-1} + q^{-1} = 1),$$

получаем

$$\sum_{s=1}^{m} \alpha_{l,s}^{\sigma} \leqslant \Big(\sum_{s=1}^{m} \alpha_{l-1,s}^{p\sigma}\Big)^{1/p} \Big(\sum_{s=1}^{m} \alpha_s^{q\sigma}\Big)^{1/q}. \tag{44}$$

По предположению индукции, имеем для первой суммы в правой части (44):

$$\Big(\sum_{s=1}^{m} \alpha_{l-1,s}^{p\sigma}\Big)^{1/p} \leqslant \Big(\sum_{s=1}^{m} \alpha_s^{p\sigma(l-1)}\Big)^{1/p} = \Big(\sum_{s=1}^{m} \alpha_s^{l\sigma}\Big)^{1-\frac{1}{l}}.$$

Учитывая, что во второй сумме правой части (44) $q=l$, легко получаем из (44):

$$\sum_{s=1}^{m} \alpha_{l,s}^{\sigma} \leqslant \sum_{s=1}^{m} \alpha_s^{l\sigma},$$

что и требовалось доказать.

В частности, при $\sigma = 2$ и $m = n$ из формулы (42) следует, что

$$\mathrm{Sp}\,(A^{*l} A^l) \leqslant \mathrm{Sp}\,(A^* A)^l. \tag{45}$$

§ 3. Неравенства Вейля

В настоящем параграфе мы выведем принадлежащие Г. Вейлю неравенства, которые позволяют оценивать собственные числа линейного оператора A посредством его сингулярных чисел**). Нам понадобится следующее важное предложение:

Лемма 6. *Пусть* λ_s $(s = 1, 2, \ldots, n)$ *— собственные числа линейного оператора* A, *занумерованные так, что*

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n|, \tag{46}$$

и пусть

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n \tag{47}$$

— сингулярные числа этого оператора. Тогда при любом $m \leqslant n$ *справедливы неравенства*

$$|\lambda_1| |\lambda_2| \ldots |\lambda_m| \leqslant \alpha_1 \alpha_2 \ldots \alpha_m. \tag{48}$$

*) См. [35], стр. 40.

**) Определение см. в § 2.

Для доказательства рассмотрим ортонормированный базис

$$e_1, e_2, \ldots, e_n, \tag{49}$$

в котором матрица оператора A имеет треугольный вид. Существование такого базиса устанавливается теоремой Шура (см. гл. IX, стр. 242).

Мы воспользуемся леммой 4 и двумя способами оценим определитель

$$[(Ae_k, Ae_l)]_1^m. \tag{50}$$

Пусть a_{ij} — элементы матрицы оператора A в базисе (49). Имеем $Ae_j = \sum_{i=1}^{n} a_{ij}e_i$. Поскольку $a_{ij}=0$ при $i>j$, то

$$Ae_j = \sum_{i=1}^{j} a_{ij}e_j \tag{51}$$

и

$$(Ae_k, Ae_l) = \Big(\sum_{i=1}^{k} a_{ik}e_i, \sum_{i=1}^{l} a_{il}\ e_i\Big) = \sum_{i=1}^{q} a_{ik}\bar{a}_{il} \qquad (q=\min(k, l)). \tag{52}$$

Формула (52) позволяет записать определитель (50) в виде следующего произведения двух определителей:

$$[(Ae_k, Ae_l)]_1^m = \begin{vmatrix} \bar{a}_{11} & 0 & 0 & \ldots & 0 \\ \bar{a}_{12} & \bar{a}_{22} & 0 & \ldots & 0 \\ \bar{a}_{13} & \bar{a}_{23} & \bar{a}_{33} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \bar{a}_{1m} & \bar{a}_{2m} & \bar{a}_{3m} & \ldots & \bar{a}_{mm} \end{vmatrix} \cdot \begin{vmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1m} \\ 0 & a_{22} & a_{23} & \ldots & a_{2m} \\ 0 & 0 & a_{33} & \ldots & a_{3m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{mm} \end{vmatrix}. \tag{53}$$

Поскольку $a_{ii}=\lambda_i$ и оба определителя в правой части (53) равны произведению диагональных элементов, то

$$[(Ae_k, Ae_l)]_1^m = |\lambda_1|^2 |\lambda_2|^2 \ldots |\lambda_m|^2. \tag{54}$$

С другой стороны, в силу леммы 4

$$[(Ae_k, Ae_l)]_1^m \leqslant \alpha_1^2\alpha_2^2 \ldots \alpha_m^2, \tag{55}$$

так как $[(e_k, e_l)]_1^m = 1$.

Неравенства (48) следуют теперь из соотношений (54) и (55). Лемма 6 доказана *).

---

*) Эту лемму можно доказать иначе, если воспользоваться теоремой 4 гл. III, согласно которой произведение $\lambda_1\lambda_2 \ldots \lambda_m$ является собственным числом ассоциированной матрицы $\mathfrak{A}_m$ для матрицы A. Пусть x — соответствующий собственный вектор $\mathfrak{A}_m$; умножая равенство $\mathfrak{A}_m x = \lambda_1\lambda_2 \ldots \lambda_m x$ на сопряженное, получаем

$$|\lambda_1\lambda_2 \ldots \lambda_m|^2 = \frac{x^*\mathfrak{A}_m^*\mathfrak{A}_m x}{x^*x}.$$

Матрица $\mathfrak{A}_m^*\mathfrak{A}_m$ является ассоциированной для A^*A и, следовательно, наибольшее собственное число $\mathfrak{A}_m^*\mathfrak{A}_m$ равно $\alpha_1^2\alpha_2^2 \ldots \alpha_m^2$. Поскольку дробь в правой части последнего равенства не превосходит $\alpha_1^2\alpha_2^2 \ldots \alpha_m^2$, то мы приходим к неравенствам (48).

Используя неравенства (48), мы сейчас докажем следующую теорему:

Теорема 2 (см. [257]). *Пусть A — линейный оператор и пусть λ_s и α_s $(s=1, 2, \ldots, n)$ — его собственные и сингулярные числа, занумерованные так же, как и в лемме 6. Пусть $f(x)$ — непрерывная при $x \geqslant 0$ функция такая, что $\varphi(t)=f(e^t)$ — монотонно возрастающая выпуклая функция параметра t. Тогда при любом $m \leqslant n$ справедливы неравенства*

$$\sum_{s=1}^{m} f(|\lambda_s|) \leqslant \sum_{s=1}^{m} f(\alpha_s). \tag{56}$$

Доказательство. Если оператор A невырожден, то согласно (48) получаем:

$$\sum_{s=1}^{m} \ln|\lambda_s| \leqslant \sum_{s=1}^{m} \ln \alpha_s \tag{57}$$

при всех $m \leqslant n$. Отсюда на основании леммы 2 уже следуют неравенства (56). Если оператор A вырожден, то неравенства (56) могут быть получены по непрерывности. Теорема доказана.

Замечание 1°. При $m=n$ неравенство (48) превращается в равенство, так как в этом случае равенство имеет место в формуле (55). Следовательно, при $m=n$ равенство достигается и в (57). Используя замечания к лемме 2, мы можем заключить, что $\sum_{s=1}^{n} f(|\lambda_s|) \leqslant \sum_{s=1}^{n} f(\alpha_s)$ для любой функции $f(x)$, если только функция $\varphi(t)=f(e^t)$ выпукла. Возрастание оказывается излишним требованием. Например, при любом вещественном σ

$$\sum_{s=1}^{n} |\lambda_s|^{\sigma} \leqslant \sum_{s=1}^{n} \alpha_s^{\sigma}. \tag{58}$$

Замечание 2°. Рассмотрим функцию

$$f(x)=\ln(1+xz), \qquad x \geqslant 0, \tag{59}$$

где z фиксировано и положительно. Легко проверить, что функция (59) удовлетворяет всем условиям теоремы 2. Поэтому при любом m, $1 \leqslant m \leqslant n$, имеем:

$$\sum_{s=1}^{m} \ln(1+|\lambda_s|z) \leqslant \sum_{s=1}^{m} \ln(1+\alpha_s z).$$

Потенцируя, получаем неравенство

$$\prod_{s=1}^{m} (1+|\lambda_s|z) \leqslant \prod_{s=1}^{m} (1+\alpha_s z), \tag{60}$$

которое используется в теории интегральных операторов.

§ 4. Максимально-минимальные свойства сумм и произведений собственных чисел эрмитовых операторов

В настоящем параграфе мы получим обобщение ряда результатов, относящихся к экстремальным свойствам собственных чисел эрмитовых операторов (см. § 7 гл. X).

1°. Для дальнейшего удобно придать теореме 12 гл. X, устанавливающей основное максимально-минимальное свойство собственных чисел, несколько иную формулировку.

Пусть A — эрмитов оператор, действующий в n-мерном унитарном пространстве R, и пусть

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n \tag{61}$$

— его собственные числа. Обозначим через R_q q-мерное подпространство пространства R.

Справедлива следующая формула:

$$\lambda_q = \max_{R_q} \min_{\substack{x \in R_q \\ (x, x)=1}} (Ax, x). \tag{62}$$

В этой формуле минимум берется по всем нормированным векторам x, принадлежащим некоторому фиксированному подпространству R_q, а затем берется максимум по всем q-мерным подпространствам. Равенство (62) составляет содержание теоремы 12 гл. X и представляет собой видоизмененную запись формулы (79). Действительно, всякое q-мерное подпространство R_q может быть рассмотрено как совокупность векторов, удовлетворяющих некоторой системе $n-q$ независимых линейных уравнений

$$L_k(x) = 0, \quad k = 1, 2, \ldots, n-q \tag{63}$$

(см. обозначения на стр. 288). Если ввести эрмитову форму $B(x, x) \equiv \equiv \sum_{s=1}^{n} |x_s|^2$ и выбирать лишь нормированные векторы x, то тогда $B(x, x) = 1$ и

$$\mu\left(\frac{A}{B}, L_1, L_2, \ldots, L_{n-q}\right) = \min_{\substack{x \in R_q \\ (x, x)=1}} (Ax, x). \tag{64}$$

В силу (79) гл. X мы получаем:

$$\lambda_{(n-q+1)} = \max_{R_q} \min_{\substack{x \in R_q \\ (x, x)=1}} (Ax, x), \tag{65}$$

где $n-q+1$ — номер собственного числа оператора A при нумерации, принятой в гл. X (в возрастающем порядке). Легко сообразить, что при новой нумерации $\lambda_{(n-q+1)} = \lambda_q$. Таким образом, соотношение (62) действительно имеет место. Легко также видеть, что согласно (62)

$$\lambda_1 = \max_{(x, x)=1} (Ax, x), \tag{66}$$

$$\lambda_n = \min_{(x, x)=1} (Ax, x). \tag{66'}$$

Непосредственным обобщением равенства (66) является следующее предложение, принадлежащее Фаню [177а].

Теорема 3. *Для любого эрмитова оператора A с собственными числами* (61) *и любого $m \leqslant n$ справедлива формула*

$$\lambda_1 + \lambda_2 + \ldots + \lambda_m = \max_{(x_i, x_j)=\delta_{ij}} (Ax_i, x_i). \tag{67}$$

Максимум в правой части (67) берется по всем системам взаимно ортогональных нормированных векторов

$$x_1, x_2, \ldots, x_m. \tag{68}$$

Для доказательства рассмотрим ортонормированный базис собственных векторов оператора A:

$$e_1, e_2, \ldots, e_n.$$

Полагая $x_i = \sum_{s=1}^{n} (x_i, e_s) e_s$, легко найдем:

$$(Ax_i, x_i) = \sum_{s=1}^{n} \lambda_s |(x_s, e_s)|^2. \tag{69}$$

Расширим систему (68) до ортонормированного базиса в R и заметим, что

$$\sum_{s=1}^{n} |(x_i, e_s)|^2 = (x_i, x_i) = 1, \tag{70}$$

$$\sum_{i=1}^{n} |(x_i, e_s)|^2 = (e_s, e_s) = 1. \tag{71}$$

Матрица $\| |(x_i, e_s)|^2 \|_1^n$, таким образом, является двояко стохастической. Из равенств (69) следует, что последовательность

$$(Ax_i, x_i), \qquad i = 1, 2, \ldots, n, \tag{72}$$

связана с последовательностью (61) двояко стохастической матрицей. На основании леммы 1 (см. (10′)) мы заключаем, что

$$\sum_{i=1}^{m} (Ax_i, x_i) \leqslant \sum_{i=1}^{m} \lambda_i. \tag{73}$$

Так как, далее, при $x_i = e_i$ $(1 \leqslant i \leqslant m)$ в формуле (73) достигается равенство, то теорема доказана.

Если наряду с оператором A ввести оператор $-A$, собственные числа которого, очевидно, равны $-\lambda_s$ $(s = 1, 2, \ldots, n)$, то на основании доказанной теоремы можно легко заключить, что

$$\lambda_n + \lambda_{n-1} + \ldots + \lambda_{n-m+1} = \min_{(x_i, x_j) = \delta_{ij}} \sum_{i=1}^{m} (Ax_i, x_i). \tag{74}$$

Эта формула обобщает равенство (66′).

Замечание. Из неравенств (73) на основании леммы 2 заключаем, что для любой непрерывной выпуклой возрастающей функции $\varphi(t)$ и любого m имеет место неравенство

$$\sum_{i=1}^{m} \varphi(a_{ii}) \leqslant \sum_{i=1}^{m} \varphi(\lambda_i),$$

где $a_{ii} = (Ax_i, x_i)$.

2°. Дальнейшее обобщение формулы (62) связано с установлением максимально-минимальных свойств сумм собственных чисел эрмитова оператора A вида

$$\lambda_{i_1} + \lambda_{i_2} + \ldots + \lambda_{i_l}, \tag{75}$$

где $1 \leqslant i_1 < i_2 < \ldots < i_l \leqslant n$ — некоторый набор натуральных чисел. Соответствующая теорема принадлежит Г. Виландту [260e].

Сохраняя в основном ход рассуждений Виландта, мы докажем более общее предложение, установленное Амир Моэзом [149]. Этим предложением мы воспользуемся и при оценке произведений собственных и сингулярных чисел.

Предварительно введем некоторые обозначения. Пусть

$$1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n \tag{76}$$

— фиксированный набор m натуральных чисел. Рассмотрим некоторую цепочку последовательно вложенных подпространств пространства $\boldsymbol{R}$:

$$\boldsymbol{R}_{i_1} \subset \boldsymbol{R}_{i_2} \subset \ldots \subset \boldsymbol{R}_{i_m}, \tag{77}$$

где индекс указывает размерность подпространства. Пусть, далее,

$$\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_m} \tag{78}$$

— система m взаимно ортогональных и нормированных векторов:

$$(\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_{k'}}) = \delta_{kk'}, \tag{79}$$

таких, что

$$\boldsymbol{x}_{i_k} \in \boldsymbol{R}_{i_k} \qquad (k = 1, 2, \ldots, m). \tag{80}$$

Условимся называть систему векторов (78) системой, *подчиненной* цепочке (77).

Сформулируем и докажем теперь следующую лемму:

Лемма 7 (см. [149]).

Пусть $\boldsymbol{A}$ — эрмитов оператор с собственными числами

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n, \tag{61}$$

и пусть

$$\varphi(t_1, t_2, \ldots, t_m) \tag{81}$$

— монотонно возрастающая по каждому аргументу функция m вещественных переменных ($m \leqslant n$)). Пусть (77) — некоторая цепочка подпространств, построенная по фиксированному набору (76), и пусть*

$$\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_m} \tag{82}$$

— подчиненная этой цепочке система векторов. Обозначим через $\boldsymbol{P}_m$ оператор проектирования на подпространство

$$[\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_m}], \tag{83}$$

натянутое на векторы (82), и пусть $\widetilde{\lambda}_1 \geqslant \widetilde{\lambda}_2 \geqslant \ldots \geqslant \widetilde{\lambda}_m$ — собственные числа эрмитова оператора

$$\boldsymbol{P}_m \boldsymbol{A} \boldsymbol{P}_m, \tag{83'}$$

*рассматриваемого в подпространстве (83)**). Тогда*

$$\varphi(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_m}) = \max_{\boldsymbol{R}_{i_1} \subset \boldsymbol{R}_{i_2} \subset \ldots \subset \boldsymbol{R}_{i_m}} \min_{\boldsymbol{x}_{i_k} \in \boldsymbol{R}_{i_k}} \varphi(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_m). \tag{84}$$

Поясним, что в формуле (84) сначала выбирается некоторая цепочка подпространств и находится минимум по всем подчиненным ей системам векторов; после этого берется максимум по всевозможным цепочкам.

Доказательство формулы (84), очевидно, сводится к доказательству следующих двух утверждений.

---

*) Предполагается, что область определения функции $\varphi(t_1, t_2, \ldots, t_m)$ содержит куб

$$\alpha \leqslant t_s \leqslant \beta, \qquad s = 1, 2, \ldots, m,$$

где α и β — границы спектра оператора $\boldsymbol{A}$.

**) Определение оператора проектирования дано на стр. 244.

А. *К любой цепочке* (77) *всегда можно подобрать такую подчиненную систему* (78), *что*

$$\varphi(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_m) \leqslant \varphi(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_m}). \tag{85}$$

В. *Существует такая цепочка* (77), *что для любой подчиненной ей системы векторов выполняется неравенство*

$$\varphi(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_m}) \leqslant \varphi(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_m). \tag{86}$$

Докажем сначала утверждение В. Пусть

$$e_1, e_2, \ldots, e_n \tag{87}$$

— базис собственных векторов оператора A, соответствующих собственным числам (61).

Выберем следующую цепочку подпространств:

$$R_{i_k} = [e_1, e_2, \ldots, e_{i_k}], \qquad k = 1, 2, \ldots, m, \tag{88}$$

и покажем, что в этом случае неравенство (86) выполняется всегда. Пусть (78) — система векторов, подчиненная цепочке (88), и пусть S_l — некоторое l-мерное подпространство, принадлежащее оболочке (83). Как мы видели [см. (62)], при любом выборе S_l

$$\widetilde{\lambda}_l \geqslant \min_{\substack{x \in S_l \\ (x, x)=1}} (P_m A P_m x, x). \tag{89}$$

Заметив, что при $x \in S_l$ $(P_m A P_m x, x) = (A P_m x, P_m x) = (Ax, x)$, положим

$$S_l = [x_{i_1}, x_{i_2}, \ldots, x_{i_l}].$$

Легко видеть, что при этом $S_l \subset R_l$, вследствие чего мы получаем:

$$\min_{\substack{x \in S_l \\ (x, x)=1}} (Ax, x) \geqslant \min_{\substack{x \in R_l \\ (x, x)=1}} (Ax, x). \tag{90}$$

Но минимум в правой части (90) достигается на собственном векторе e_{i_l} и равен λ_{i_l}. Сравнивая (90) и (89), мы заключаем, что

$$\widetilde{\lambda}_l \geqslant \lambda_{i_l}.$$

Отсюда в силу возрастания функции (81) следует (86). Таким образом, утверждение В доказано.

Утверждение А доказывается труднее. Мы проведем его по индукции, предполагая, что для операторов, действующих в $(n-1)$-мерном пространстве, это утверждение справедливо. Заметим, что в случае $n = 1$ (пространство одномерно) $\lambda_{i_1} = \widetilde{\lambda}_{i_1}$ и неравенство (85) имеет место для любой функции $\varphi(t)$.

Переходя к случаю n-мерного пространства, мы можем считать, что $m < n$. Действительно, при $m = n$ подпространство (83) совпадает со всем пространством, $\widetilde{\lambda}_s = \lambda_s$ $(s = 1, 2, \ldots, n)$, и, следовательно, неравенство (85) справедливо.

При $m < n$ мы разберем два подслучая:

1) Пусть $i_m < n$, тогда существует некоторое $(n-1)$-мерное подпространство $\widetilde{R}_{n-1}$, содержащее все подпространства цепочки (77). Пусть P_{n-1} — оператор проектирования на подпространство $\widetilde{R}_{n-1}$. Введем в $\widetilde{R}_{n-1}$

эрмитов оператор

$$A_{n-1} = P_{n-1} A P_{n-1}. \tag{91}$$

Ясно, что для всех $x \in R_{n-1}$ имеет место равенство $(A_{n-1}x, x) = (Ax, x)$. Если λ'_s $(s = 1, 2, \ldots, n-1)$ — собственные числа оператора A_{n-1}, то в силу теоремы 14 (стр. 291) *)

$$\lambda_s \geqslant \lambda'_s \qquad (s = 1, 2, \ldots, n-1). \tag{92}$$

По индуктивному предположению, для любой цепочки (77) в $\tilde{R}_{n-1}$ найдется подчиненная ей система векторов (78) таких, что

$$\varphi(\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_m) \leqslant \varphi(\lambda'_{i_1}, \lambda'_{i_2}, \ldots, \lambda'_{i_m}).$$

Отсюда в силу (92) сразу следует, что неравенство (85) в рассматриваемом случае действительно имеет место.

2) Рассмотрим теперь случай $i_m = n$ $(m < n)$. Пусть

$$i_m = n, \; i_{m+1} = n-1, \; \ldots, \; i_{m-p} = n-p \qquad (p \geqslant 0) \tag{93}$$

— последние элементы набора (76), и пусть число $n-p-1$ уже не принадлежит набору (76).

Обозначим через $i_{m'}$ наибольший из оставшихся номеров набора (76). Очевидно **),

$$i_{m'} \leqslant n-p-2. \tag{94}$$

Цепочка подпространств (77) в рассматриваемом случае может быть записана следующим образом:

$$R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_{m'}} \subset R_{n-p} \subset R_{n-p+1} \subset \ldots \subset R_n. \tag{95}$$

Пусть

$$e_{n-p}, \; e_{n-p+1}, \; \ldots, \; e_n \tag{96}$$

— собственные векторы оператора A, занумерованные так же, как и в (87). Обозначим через R_{n-1} $(n-1)$-мерное подпространство, содержащее векторы (96) (их всего $p+1$) и подпространство $R_{i_{m'}}$.

Такое подпространство действительно существует, ибо

$$p+1+i_{m'} \leqslant p+1+n-p-2 = n-1.$$

Наряду с цепочкой (95) рассмотрим цепочку

$$R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_{m'}} \subset \tilde{R}_{n-p-1} \subset \tilde{R}_{n-p} \subset \ldots \subset \tilde{R}_{n-1}, \tag{97}$$

которая получается из цепочки (95), если каждое подпространство в (95) заменить его пересечением с подпространством $\tilde{R}_{n-1}$. Ясно, что первые m' подпространств цепочки (95) при этом не изменяются, так как они содержатся в $\tilde{R}_{n-1}$; последующие подпространства цепочки уменьшат свою размерность на единицу ***). Введем снова оператор A_{n-1} на

*) Напомним, что в § 7 гл. X собственные числа занумерованы в порядке возрастания. Неравенства (92) легко также следуют из формулы (62).

**) Может случиться, что номера (93) исчерпывают весь набор (76), тогда для сохранения единообразия мы будем считать $i_{m'} = 0$, а соответствующее подпространство $R_{i_{m'}}$ — состоящим из нуль-вектора.

***) Если размерность пересечения не уменьшается, то, очевидно, пересечение всегда можно сузить так, чтобы подпространства в цепочке (97) имели указанные размерности.

подпространстве $\widetilde{R}_{n-1}$ по формуле (91). По индуктивному предположению, к цепочке (97) можно подобрать подчиненную систему векторов

$$x_{i_1},\ x_{i_2},\ \ldots,\ x_{i_{m'}},\ x_{n-p-1},\ \ldots,\ x_{n-1} \tag{98}$$

таких, что

$$\varphi(\widetilde{\lambda}_1,\ \widetilde{\lambda}_2,\ \ldots,\ \widetilde{\lambda}_m) \leqslant \varphi(\lambda'_{i_1},\ \lambda'_{i_2},\ \ldots,\ \lambda'_{i_{m'}},\ \lambda'_{n-p-1},\ \ldots,\ \lambda'_{n-1}). \tag{99}$$

В правой части штрихами обозначены собственные числа оператора A_{n-1}. Согласно (92) имеем:

$$\lambda_{i_1} \geqslant \lambda'_{i_1},\ \lambda_{i_2} \geqslant \lambda'_{i_2},\ \ldots,\ \lambda_{i_m} \geqslant \lambda'_{i_{m'}}. \tag{100}$$

Разумеется, на основании (92) мы не можем утверждать, что

$$\lambda_{n-p} \geqslant \lambda'_{n-p-1},\ \lambda_{n-p+1} \geqslant \lambda'_{n-p},\ \ldots,\ \lambda_n \geqslant \lambda'_{n-1}. \tag{100'}$$

Однако поскольку векторы (96) лежат в подпространстве $\widetilde{R}_{n-1}$, они являются собственными векторами оператора A_{n-1}. Соответствующие им собственные числа $\lambda_{n-p} \geqslant \lambda_{n-p+1} \geqslant \ldots \geqslant \lambda_n$ во всяком случае не меньше, чем числа $\lambda'_{n-p+1} \geqslant \lambda'_{n-p} \geqslant \ldots \geqslant \lambda'_{n-p}$, которые являются наименьшими собственными числами оператора A_{n-1}. Таким образом, (100') имеет место, и в силу (100) и (100') мы заключаем, что

$$\varphi(\lambda'_{i_1},\ \lambda'_{i_2},\ \ldots,\ \lambda'_{i_m},\ \lambda'_{n-p-1},\ \lambda'_{n-p}) \leqslant \varphi(\lambda_{i_1},\ \lambda_{i_2},\ \ldots,\ \lambda_{i_{m'}},\ \lambda_{n-p},\ \ldots,\ \lambda_n). \tag{101}$$

Это неравенство вместе с (99) приводит к доказательству утверждения A, поскольку система векторов (98) подчинена не только цепочке (97), но и исходной цепочке (95). Таким образом, лемма 7 доказана полностью.

Из доказанной леммы вытекает следующая теорема:

Теорема 4 (см. [260e]). *Пусть A — эрмитов оператор с собственными числами $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$. Пусть*

$$1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n \tag{102}$$

— фиксированный набор n натуральных чисел; тогда

$$\lambda_{i_1} + \lambda_{i_2} + \ldots + \lambda_{i_m} = \max_{R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_m}} \min_{x_{i_k} \in R_{i_k}} \sum_{k=1}^{m} (Ax_{i_k},\ x_{i_k}), \tag{103}$$

*где минимум берется по всем системам векторов x_{i_k} $(k = 1, 2, \ldots, m)$, подчиненным цепочке**) $R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_m}$.

Для доказательства заметим, что матрица оператора $P_m A P_m$ (см. лемму 7) в ортонормированном базисе (82) имеет вид

$$\| (Ax_{i_k},\ x_{i_{k'}}) \|_{k,\ k'=1}^{m},$$

так как

$$(P_m A P_m x_{i_k},\ x_{i_{k'}}) = (Ax_{i_k},\ x_{i_{k'}}).$$

Поэтому сумма, стоящая в правой части (103), есть след оператора $P_m A P_m$ и, значит, равна $\widetilde{\lambda}_1 + \widetilde{\lambda}_2 + \ldots + \widetilde{\lambda}_m$. Формула (103) после этих замечаний следует из леммы 7 при

$$\varphi(t_1,\ t_2,\ \ldots,\ t_m) \equiv t_1 + t_2 + \ldots + t_m.$$

Теорема 4 доказана.

---

*) Определение см. на стр. 547.

Укажем, что формула (62), равно как и (67) и (74), является частным случаем теоремы 4.

В заключение установим следующее утверждение, тоже непосредственно вытекающее из леммы 7.

Теорема 5 (см. [149]). *Пусть A — неотрицательно определенный эрмитов оператор, и пусть*

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n \geqslant 0$$

— его собственные значения. Пусть $1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n$ — фиксированный набор m натуральных чисел. Тогда

$$\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_m} = \max_{R_{i_1}\subset R_{i_2}\subset\ldots\subset R_{i_m}} \min_{x_{i_k}\in R_{i_k}} \operatorname{Det} \| (Ax_{i_k}, x_{i_{k'}}) \|_{k,\,k'=1}^{m}, \tag{104}$$

где минимум берется по всем системам векторов, подчиненным цепочке $R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_m}$.

Для доказательства теоремы достаточно в формуле (84) положить

$$\varphi(t_1, t_2, \ldots, t_m) = t_1 t_2 \ldots t_m \qquad (t_s \geqslant 0,\ s = 1, 2, \ldots, m)$$

и заметить, что стоящий в правой части (104) определитель равен $\widetilde{\lambda}_1\widetilde{\lambda}_2\ldots\widetilde{\lambda}_m$.

Теоремами 4 и 5 мы воспользуемся в следующем параграфе при выводе неравенств для сумм и произведений собственных и сингулярных чисел.

§ 5. Неравенства для собственных и сингулярных чисел сумм и произведений операторов

Пусть A и B — два эрмитовых оператора в n-мерном унитарном пространстве R, собственные числа которых известны. Теоремы предыдущего параграфа дают возможность оценить суммы вида (75) собственных чисел оператора $A+B$. Близкие по характеру оценки мы получим и в случае произведения операторов. Мы начнем со следующего предложения.

Теорема 6 (см. [260e]). *Пусть A, B и C — эрмитовы операторы такие, что $C = A + B$; пусть λ_s, μ_s и ν_s $(s = 1, 2, \ldots, n)$ — собственные числа операторов A, B и C соответственно, занумерованные в порядке убывания.*

Тогда для любого набора m натуральных чисел

$$1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n \tag{105}$$

справедливо неравенство

$$\nu_{i_1} + \nu_{i_2} + \ldots + \nu_{i_m} \leqslant \lambda_{i_1} + \lambda_{i_2} + \ldots + \lambda_{i_m} + \mu_1 + \mu_2 + \ldots + \mu_m. \tag{106}$$

При $m = n$ в формуле (106) *достигается равенство.*

Доказательство. По заданному набору (105) выберем такую цепочку подпространств

$$R_{i_1} \subset R_{i_2} \subset \ldots \subset R_{i_m}, \tag{107}$$

чтобы для любой системы векторов

$$x_{i_1}, x_{i_2}, \ldots, x_{i_m}, \tag{108}$$

подчиненной цепочке (107), выполнялось неравенство

$$\nu_{i_1} + \nu_{i_2} + \ldots + \nu_{i_m} \leqslant \sum_{k=1}^{m} (Cx_{i_k}, x_{i_k}). \tag{109}$$

Такая цепочка (107) найдется в силу теоремы 4 (ср. лемму 7

Заметив, что

$$\sum_{k=1}^{m}(\boldsymbol{C}\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_k})=\sum_{k=1}^{m}(\boldsymbol{A}\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_k})+\sum_{k=1}^{m}(\boldsymbol{B}\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_k}), \tag{110}$$

подберем такую систему векторов, подчиненную *) цепочке (107), чтобы

$$\sum_{k=1}^{m}(\boldsymbol{A}\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_k})\leqslant\lambda_{i_1}+\lambda_{i_2}+\ldots+\lambda_{i_m}. \tag{111}$$

Такую систему векторов можно найти также на основании теоремы 4 (ср. лемму 7, А). Так как, далее, в силу теоремы 3 для любой ортогональной нормированной системы $\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_k}$

$$\sum_{k=1}^{m}(\boldsymbol{B}\boldsymbol{x}_{i_k}, \boldsymbol{x}_{i_k})\leqslant\mu_1+\mu_2+\ldots+\mu_m, \tag{112}$$

то неравенство (106) следует из (109), (110), (111) и (112). При $m=n$ (106) превращается в равенство, поскольку $\operatorname{Sp}\boldsymbol{C}=\operatorname{Sp}\boldsymbol{A}+\operatorname{Sp}\boldsymbol{B}$.

Теорема 6 доказана полностью.

Следствие. *Для любой непрерывной выпуклой функции* $\varphi(t)$ *справедливо неравенство*

$$\sum_{s=1}^{n}\varphi(\nu_s-\lambda_s)\leqslant\sum_{s=1}^{n}\varphi(\mu_s). \tag{113}$$

Этот результат следует из неравенств (106) на основании замечания к лемме 2.

Оказывается, неравенства типа (106) справедливы для сингулярных чисел произвольных линейных операторов. Для доказательства соответствующего предложения мы используем следующее замечание: пусть A — матрица некоторого линейного оператора $\boldsymbol{A}$ в ортонормированном базисе, и пусть $\alpha_1\geqslant\alpha_2\geqslant\ldots\geqslant\alpha_n$ — сингулярные числа $\boldsymbol{A}$. Рассмотрим квадратную матрицу

$$\widehat{A}=\begin{pmatrix}0 & A\\ A^* & 0\end{pmatrix}$$

порядка $2n$. Мы сейчас покажем, что собственные числа матрицы $\widehat{A}$ равны $\pm\alpha_s$ ($s=1, 2, \ldots, n$). Действительно, раскрыв характеристический определитель $\Delta(\lambda)=|\widehat{A}-\lambda E_{2n}|$ по формуле (1б) стр. 59, легко найдем:

$$\Delta(\lambda)=|\lambda^2E_n-A^*A|.$$

Отсюда сразу следует сформулированное утверждение.

Сопоставив каждому оператору $\boldsymbol{A}$ оператор $\widehat{\boldsymbol{A}}$, действующий в $2n$-мерном унитарном пространстве, мы на основании теоремы 6 установим следующее предложение:

Теорема 7 (см. [149]). *Пусть* $\boldsymbol{A}$, $\boldsymbol{B}$ *и* $\boldsymbol{C}$ *— линейные операторы в* n*-мерном унитарном пространстве. Пусть* $\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}$, *и пусть* α_s, β_s *и* γ_s ($s=1, 2, \ldots, n$) *— сингулярные числа операторов* $\boldsymbol{A}$, $\boldsymbol{B}$ *и* $\boldsymbol{C}$, *занумерованные в порядке убывания. Тогда для любого набора натуральных*

*) Определение см. на стр. 547.

чисел (105) *справедливо неравенство*

$$\gamma_{i_1}+\gamma_{i_2}+\ldots+\gamma_{i_m}\leqslant\alpha_{i_1}+\alpha_{i_2}+\ldots+\alpha_{i_m}+\beta_1+\beta_2+\ldots+\beta_m. \quad (114)$$

Замечание. Так как собственные числа операторов $\widehat{A}$, $\widehat{B}$ и $\widehat{C}$ располагаются относительно начала координат симметричными парами, неравенства (114) на основании теоремы 6 легко обобщаются следующим образом:

$$\pm(\gamma_{i_1}-\alpha_{i_1})\pm(\gamma_{i_2}-\alpha_{i_2})\pm\ldots\pm(\gamma_{i_m}-\alpha_{i_m})\leqslant\beta_1+\beta_2+\ldots+\beta_m.$$

Поскольку выбор знака перед каждой скобкой произволен, то

$$|\gamma_{i_1}-\alpha_{i_1}|+|\gamma_{i_2}-\alpha_{i_2}|+\ldots+|\gamma_{i_m}-\alpha_{i_m}|\leqslant\beta_1+\beta_2+\ldots+\beta_m. \quad (114')$$

Используя лемму 2, мы можем на основании неравенств (114) заключить, что для любой непрерывной возрастающей выпуклой функции $\varphi(t)$, $t\geqslant 0$, и любого $m\leqslant n$

$$\sum_{s=1}^{m}\varphi(|\gamma_s-\alpha_s|)\leqslant\sum_{s=1}^{m}\varphi(\beta_s)\ *).$$

Перейдем теперь к оценке сингулярных и собственных чисел произведений двух операторов. Основной в этом направлении является следующая теорема, обобщающая неравенство (35) (см. лемму 5).

Теорема 8. *Пусть* $\boldsymbol{A}$, $\boldsymbol{B}$ *и* $\boldsymbol{C}$ *— линейные операторы в* n*-мерном унитарном пространстве. Пусть* $\boldsymbol{C}=\boldsymbol{AB}$ *и пусть* α_s, β_s *и* γ_s $(s=1, 2, \ldots, n)$ *— сингулярные числа операторов* $\boldsymbol{A}$, $\boldsymbol{B}$ *и* $\boldsymbol{C}$ *соответственно, занумерованные в порядке убывания; тогда для любого набора натуральных чисел*

$$1\leqslant i_1<i_2<\ldots<i_m\leqslant n \quad (115)$$

справедливы неравенства

$$\gamma_{i_1}\gamma_{i_2}\ldots\gamma_{i_m}\leqslant\alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_m}\beta_1\beta_2\ldots\beta_m, \quad (116)$$

$$\gamma_{i_1}\gamma_{i_2}\ldots\gamma_{i_m}\leqslant\alpha_1\alpha_2\ldots\alpha_m\beta_{i_1}\beta_{i_2}\ldots\beta_{i_m}. \quad (116')$$

Доказательство этой теоремы вполне аналогично доказательству теоремы 6.

Докажем сначала неравенства (116'). На основании леммы 4 § 2 получаем:

$$[(\boldsymbol{C}\boldsymbol{x}_{i_k},\ \boldsymbol{C}\boldsymbol{x}_{i_k})]=[(\boldsymbol{AB}\boldsymbol{x}_{i_k},\ \boldsymbol{AB}\boldsymbol{x}_{i_k})]\leqslant\alpha_1^2\alpha_2^2\ldots\alpha_m^2[(\boldsymbol{B}\boldsymbol{x}_{i_k},\ \boldsymbol{B}\boldsymbol{x}_{i_k})] \quad (117)$$

для любой системы векторов $\boldsymbol{x}_{i_k}$ $(k=1, 2, \ldots, m)$. По данному набору натуральных чисел (115) найдем на основании теоремы 5 такую цепочку подпространств, чтобы для любой подчиненной ей системы векторов выполнялось неравенство

$$\gamma_{i_1}^2\gamma_{i_2}^2\ldots\gamma_{i_m}^2\leqslant[(\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{x}_{i_k},\ \boldsymbol{x}_{i_k})], \quad (118)$$

после чего на основании той же теоремы 5 найдем такую подчиненную выбранной цепочке систему векторов, чтобы

$$[(\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{x}_{i_k},\ \boldsymbol{x}_{i_k})]\leqslant\beta_{i_1}^2\beta_{i_2}^2\ldots\beta_{i_m}^2. \quad (119)$$

Очевидно, неравенство (116) уже следует из неравенств (117), (118) и (119).

*) По поводу этих неравенств см. [212e].

Для доказательства неравенства (116′) следует повторить рассуждение применительно к оператору $C^* = B^*A^*$ и воспользоваться при этом тем фактом, что сингулярные числа у сопряженных операторов равны [см. (82), стр. 251]. Теорема 8 доказана полностью.

Заметим попутно, что сингулярные числа операторов AB и BA в общем случае не совпадают.

Отметим еще следующий факт, вытекающий из теоремы 8:

Теорема 9. *Пусть A и B — два положительно определенных эрмитовых оператора с собственными числами λ_s и μ_s ($s = 1, 2, \ldots, n$), занумерованными в убывающем порядке, и пусть*

$$\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_n \tag{120}$$

— собственные числа оператора AB.

Тогда для любого набора (115) *выполняется неравенство*

$$\nu_{i_1}\nu_{i_2} \ldots \nu_{i_m} \leqslant \lambda_{i_1}\lambda_{i_2} \ldots \lambda_{i_m}\mu_1\mu_2 \ldots \mu_m. \tag{121}$$

Доказательство. Так как оператор B невырожден, то

$$AB = B^{-1/2}(B^{1/2}AB^{1/2})\,B^{1/2} = B^{-1/2}\{(A^{1/2}B^{1/2})^*(A^{1/2}B^{1/2})\}\,B^{1/2} \tag{122}$$

и, следовательно, собственные числа (120) являются квадратами сингулярных чисел оператора $A^{1/2}\cdot B^{1/2}$.

Применяя к произведению $A^{1/2}\cdot B^{1/2}$ неравенство (116), получаем (121).

§ 6. Другая постановка задачи о спектре суммы и произведения эрмитовых операторов

В настоящем параграфе мы сопоставим набору собственных чисел $\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_n$ суммы эрмитовых операторов A и B точку в n-мерном координатном пространстве и рассмотрим множество точек, получающихся при сложении всевозможных операторов A и B с данными спектрами. Аналогичную задачу мы рассмотрим и в случае произведения операторов.

Постановка задач на собственные значения в указанной геометрической форме принадлежит И. М. Гельфанду (см. по этому поводу [76], [106], [113], [191c].

Мы приведем здесь лишь аналоги теорем 6 и 9, первоначально полученные другими методами.

1° Нам понадобится геометрическое описание всех последовательностей α', которые мажорируются данной последовательностью α. Соответствующую лемму мы установим, используя некоторые результаты Фробениуса, Кёнига и Биркгофа. Начнем со следующего замечания.

Пусть $T = \| t_{ij} \|^n$ — квадратная матрица. *Нормальным набором* элементов матрицы T назовем набор n элементов этой матрицы, взятых по одному из каждой строки и каждого столбца, т. е. набор вида

$$t_{1j_1},\ t_{2j_2},\ \ldots,\ t_{nj_n}, \tag{123}$$

где $j_1, j_2, \ldots, j_n$ — некоторая перестановка индексов $1, 2, \ldots, n$.

Справедливо следующее предложение, имеющее самостоятельное значение в ряде разделов математики.

Лемма 8 (Фробениус — Кёниг [182e], [203]). *Пусть $T = \| t_{ij} \|_1^n$ — квадратная матрица порядка n с неотрицательными элементами, и пусть каждый нормальный набор матрицы T содержит нулевой элемент; тогда существует состоящий из нулей минор матрицы T размеров $p \times q$ такой, что $p + q = n + 1$.*

Доказательство *) мы проведем по индукции, предположив, что для матриц всех порядков $k < n$ лемма справедлива. Случай $n = 1$ тривиален.

Обращаясь к матрице T порядка n, мы можем, очевидно, считать, что не все ее элементы равны нулю. Пусть для определенности $t_{nn} \neq 0$ (этого всегда можно добиться перестановками строк и столбцов, так как при этом условия леммы не нарушаются).

Для матрицы $T_1 = \| t_{ij} \|_1^{n-1}$, очевидно, выполняются условия леммы, и по индуктивному предположению найдется состоящий из нулей минор M_1 матрицы T_1 размеров $p_1 \times q_1$:

$$p_1 + q_1 = n. \tag{124}$$

Не ограничивая общности, можно считать, что минор M_1 расположен на пересечении первых p_1 строк и первых q_1 столбцов матрицы T.

Разобьем матрицу T следующим образом на блоки:

$$T = \begin{array}{c} \\ p_1 \\ q_1 \end{array} \overbrace{\left\| \begin{array}{c|c} M_1 & T_2 \\ \hline T_3 & \end{array} \right\|}^{\begin{array}{cc} q_1 & p_1 \end{array}}$$

и рассмотрим квадратные матрицы T_2 и T_3 размеров $p_1 \times p_1$ и $q_1 \times q_1$.

Хотя бы одна из матриц T_2 или T_3 обладает тем свойством, что каждый нормальный набор ее элементов содержит нулевой элемент (в противном случае можно было бы образовать нормальный набор положительных элементов всей матрицы T). Пусть указанным свойством обладает матрица T_2. По предположению индукции, T_2 обладает состоящим из одних нулей минором M_2 размеров $p_2 \times q_2$:

$$p_2 + q_2 = p_1 + 1. \tag{125}$$

Очевидно, можно считать, что минор M_2 расположен в строках матрицы T с номерами $1, 2, \ldots, p_2$ и в столбцах с номерами $q_1 + 1, q_1 + 2, \ldots, q_1 + q_2$. Легко видеть, что при этом минор матрицы T, расположенный в строках с номерами $1, 2, \ldots, p_2$ и в столбцах с номерами $1, 2, \ldots, q_1, q_1 + 1, \ldots, q_1 + q_2$, состоит из одних нулей, причем в силу (124) и (125)

$$p_2 + (q_1 + q_2) = q_1 + p_1 + 1 = n + 1.$$

Лемма 8 доказана.

Следствие. *Пусть элементы квадратной матрицы T неотрицательны и сумма элементов в каждой строке и в каждом столбце равна $\omega > 0$. Тогда матрица T обладает нормальным набором положительных элементов.*

В самом деле, допустив противное, мы сможем согласно лемме 8 указать минор матрицы T размеров $p \times q$, $p + q = n + 1$, состоящий из одних нулей. Легко видеть, что сумма элементов матрицы T, расположенных в тех p строках и тех q столбцах, на пересечении которых расположен данный минор, равна $p\omega + q\omega = (n + 1)\omega$. Последнее, однако, невозможно, так как сумма всех элементов матрицы T равна $n\omega$.

Условимся в дальнейшем матрицу P порядка n называть матрицей перестановки (permutation matrix), если она обладает нормальным набо-

*) См. [172].

ром элементов, каждый из которых равен единице, а все остальные элементы матрицы равны нулю. Ясно, что умножение матрицы P на столбцевую матрицу x приводит к некоторой перестановке элементов матрицы x.

Так как и, обратно, каждая перестановка элементов x порождает некоторую матрицу P, то всего существует $n!$ различных матриц перестановок.

Докажем теперь следующее предложение, принадлежащее Биркгофу [153].

Лемма 9. *Множество всех двояко стохастических матриц совпадает с выпуклой оболочкой матриц перестановок. Другими словами, любая двояко стохастическая матрица T может быть представлена в виде*

$$T = \sum_{s=1}^{n!} \tau_s P_s, \tag{126}$$

где

$$\tau_s \geqslant 0, \quad \sum_{s=1}^{n!} \tau_s = 1, \tag{127}$$

а P_s — матрицы перестановок. Обратно, правая часть (126) *при условии* (127) *является двояко стохастической матрицей.*

Последняя часть утверждения почти очевидна. В самом деле, сумма элементов i-го столбца матрицы $\tau_s P_s$ равна τ_s. Поэтому сумма элементов i-го столбца правой части (126) равна $\sum_{s=1}^{n!} \tau_s = 1$. Аналогично в случае строк.

Доказательство первой части леммы существенно использует предыдущую лемму 8.

Пусть T — двояко стохастическая матрица. Тогда по следствию из леммы 8 существует положительный набор элементов

$$t_{1j_1}, \ t_{2j_2}, \ \ldots, \ t_{nj_n} \tag{128}$$

этой матрицы. Пусть

$$\min_s t_{sj_s} = \tau_1 \qquad (\tau_1 > 0), \tag{129}$$

и пусть P_1 — матрица перестановки, у которой на местах элементов набора (128) стоят единицы. Рассмотрим матрицу

$$B_1 = T - \tau_1 P_1. \tag{130}$$

В силу (129) элементы матрицы B неотрицательны, а сумма элементов в каждой строке и в каждом столбце B_1 равна

$$1 - \tau_1 = \omega_1 \geqslant 0.$$

Заметим, что число нулевых элементов B_1 во всяком случае на единицу больше, чем у матрицы T. Если $\omega_1 = 0$, то $B_1 = 0$, и лемма доказана. Если $\omega_1 > 0$, то B_1 обладает нормальным набором положительных элементов, и, повторив рассуждение, мы придем к неотрицательной матрице $B_2 = T - \tau_1 P_1 - \tau_2 P_2$, число нулевых элементов которой уже на два больше, чем у T. Суммы элементов в столбцах и строках B_2 равны $1 - \tau_1 - \tau_2 = \omega_2 \geqslant 0$. Ясно, что на некотором k-м шаге ($k < n^2 - n + 1$) этот процесс приведет к числу $\omega_k = 1 - \tau_1 - \tau_2 - \ldots - \tau_k = 0$ и, следовательно, к матрице

$$B_k = T - \tau_1 P_1 - \tau_2 P_2 - \ldots - \tau_k P_k = 0.$$

Действительно, при $k = n^2 - n + 1$ матрица B_k уже не имеет нормального набора положительных элементов (у нее $n^2 - n + 1$ нулевых элементов), и, следовательно, ω_k не может быть положительным числом. Лемма 9 доказана.

2° Условимся каждой числовой последовательности $(x_1, x_2, \ldots, x_n)$ ставить в соответствие точку в n-мерном координатном пространстве $\boldsymbol{D}_n$.

Пусть

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n \tag{131}$$

— некоторая последовательность. Рассмотрим $n!$ последовательностей, получающихся из последовательности (131) всевозможными перестановками ее элементов

$$\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}. \tag{131'}$$

Сопоставив каждой последовательности (131') точку в $\boldsymbol{D}_n$, обозначим через $K(\alpha)$ линейную выпуклую оболочку, натянутую на эти точки.

Легко видеть, что множество $K(\alpha)$ состоит из всех точек

$$x = \sum_{s=1}^{n!} \tau_s P_s \alpha, \tag{132}$$

где $\tau_s \geqslant 0$, $\sum_{s=1}^{n!} \tau_s = 1$, P_s — матрицы перестановок и α — столбцевая матрица с координатами (131).

Заметим попутно, что каждая точка принадлежит множеству $K(\alpha)$ вместе с $n!$ точками, получающимися перестановками ее координат. Для доказательства достаточно умножить равенство (132) на матрицу перестановки и воспользоваться тем, что произведение матриц перестановок есть матрица перестановки.

Мы теперь без труда докажем следующее предложение:

Л е м м а 10. *Для того чтобы последовательность*

$$\alpha'_1 \geqslant \alpha'_2 \geqslant \ldots \geqslant \alpha'_n \tag{133}$$

мажорировалась последовательностью $\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n$, *необходимо и достаточно, чтобы точка* α' *с координатами* (133) *принадлежала выпуклой линейной оболочке* $K(\alpha)$ *).

Д о к а з а т е л ь с т в о л е м м ы. Пусть сначала $\alpha' \in K(\alpha)$, тогда

$$\alpha' = \sum_{s=1}^{n!} \tau_s P_s \alpha, \tag{134}$$

где $\tau_s \geqslant 0$, $\sum_{s=1}^{n!} \tau_s = 1$, а P_s — матрицы перестановок. Следовательно, по лемме

$$\alpha' = T\alpha, \tag{135}$$

где T — двояко стохастическая матрица, и значит, согласно лемме 1 последовательность (133) мажорируется последовательностью (131).

---

*) Это утверждение было установлено Радо [234], использовавшего при доказательстве теорему об отделении выпуклых множеств плоскостями. Другое доказательство, основанное на теореме о крайних точках выпуклых множеств, дано в работе [113], содержащей обстоятельный обзор литературы.

Идея приводимого здесь доказательства, опирающегося на лемму Биркгофа, принадлежит А. Хорну [191d].

Пусть теперь, наоборот, известно, что $\alpha' < \alpha$, тогда по лемме 1 найдется дважды стохастическая матрица T, с которой выполняется равенство (135). Этому равенству согласно лемме 9 можно придать вид (134), откуда уже следует, что $\alpha' \in K$ (α).

3° Перейдем теперь к теоремам, составляющим цель настоящего параграфа.

Теорема 10 (см. [106а]). *Пусть* $\boldsymbol{A}$ *и* $\boldsymbol{B}$ *— эрмитовы операторы в n-мерном унитарном пространстве с собственными значениями*

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n \tag{136}$$

и

$$\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n. \tag{137}$$

Пусть $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ *и пусть*

$$\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_n \tag{138}$$

— собственные числа $\boldsymbol{C}$. *Обозначим через* K_1 *выпуклую линейную оболочку точек*

$$(\lambda_1 + \mu_{j_1}, \lambda_2 + \mu_{j_2}, \ldots, \lambda_n + \mu_{j_n}) \tag{139}$$

и через K_2 *выпуклую линейную оболочку точек*

$$(\mu_1 + \lambda_{j_1}, \mu_2 + \lambda_{j_2}, \ldots, \mu_n + \lambda_{j_n}) \tag{140}$$

(*берутся всевозможные перестановки* $j_1, j_2, \ldots, j_n$ *чисел* 1, 2, ..., n).

Тогда точка $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$ *принадлежит пересечению оболочек* K_1 *и* K_2.

Доказательство. Рассмотрим точку

$$(\nu_1 - \lambda_1, \nu_2 - \lambda_2, \ldots, \nu_n - \lambda_n). \tag{141}$$

Согласно теореме 6 [см. неравенства (106)] последовательность, полученная упорядочением координат точки (141), мажорируется последовательностью $\mu_1, \mu_2, \ldots, \mu_n$. На основании леммы 10 мы заключаем, что точка (141) принадлежит выпуклой линейной оболочке точек $(\mu_{j_1}, \mu_{j_2}, \ldots, \mu_{j_n})$. Отсюда следует, что $\nu = (\nu_1, \nu_2, \ldots, \nu_n) \in K_1$. Поменяв ролями $\boldsymbol{A}$ и $\boldsymbol{B}$, легко получим, что $\nu \in K_2$. Таким образом, теорема 10 доказана.

Мы вывели теорему 10 из теоремы 6. Легко видеть, что и обратно, ввиду леммы 10 теорема 6 следует из теоремы 10*).

В связи с теоремой 10 сделаем несколько замечаний.

Обозначим через M множество точек (138), отвечающих спектрам операторов $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$, где $\boldsymbol{A}$ и $\boldsymbol{B}$ — всевозможные эрмитовы операторы с данными спектрами (136) и (150). Утверждение теоремы 10 состоит в том, что $M \subset K_1 \cap K_2$.

Полное описание множества M до сих пор не получено, несмотря на имеющиеся в этой области важные исследования ([76], [191с]).

В частности, в работе [191с] найдено полное описание множества M для случая $n \leqslant 4$.

Приведем без доказательства относящийся к этому вопросу следующий просто формулируемый результат [106а]:

Пусть при всех $i = 1, 2, \ldots, n-1$

$$\mu_1 - \mu_n < \lambda_{i+1} - \lambda_i \tag{142}$$

*) В [106а] теорема 10 доказана другим методом.

или

$$\lambda_1-\lambda_n<\mu_{i+1}-\mu_i. \tag{142'}$$

Тогда

$$M=K_1\cap K_2.$$

Отметим здесь же, что при условии (142) или (142') среди собственных значений оператора $C=A+B$ нет кратных. В самом деле, пусть, например, выполнено условие (142). Согласно теореме 10

$$\nu_{i+1}-\nu_i=\lambda_{i+1}-\lambda_i+\sum_{s=1}^{n!}\tau_s(\mu_{(s)}-\mu_{(s')}),$$

где

$$\tau_s\geqslant 0 \quad \text{и} \quad \sum_{s=1}^{n!}\tau_s=1.$$

Поэтому $\nu_{i+1}-\nu_i\geqslant\lambda_{i+1}-\lambda_i-\sum_{s=1}^{n!}\tau_s(\mu_1-\mu_n)=\lambda_{i+1}-\lambda_i-(\mu_1-\mu_n)>0$ и, следовательно, $\nu_{i+1}\neq\nu_i$.

Сформулируем в геометрических терминах теорему, эквивалентную теореме 9.

Теорема 11 ([106]). *Пусть A и B — положительно определенные эрмитовы операторы с собственными числами λ_s и μ_s $(s=1, 2, \ldots, n)$, занумерованными в убывающем порядке, и пусть*

$$\nu_1\geqslant\nu_2\geqslant\ldots\geqslant\nu_n$$

— собственные числа оператора $C=AB$. Пусть K_1 — выпуклая линейная оболочка точек

$$(\ln\lambda_1+\ln\mu_{j_1},\ \ln\lambda_2+\ln\mu_{j_2},\ \ldots,\ \ln\lambda_n+\ln\mu_{j_n})$$

и K_2 — выпуклая линейная оболочка точек

$$(\ln\mu_1+\ln\lambda_{j_1},\ \ln\mu_2+\ln\lambda_{j_2},\ \ldots,\ \ln\mu_n+\ln\lambda_{j_n}).$$

Тогда точка с координатами $(\ln\nu_1, \ln\nu_2, \ldots, \ln\nu_n)$ *принадлежит пересечению оболочек K_1 и K_2.*

Доказательство. Прологарифмировав неравенства (121), получим:

$$(\ln\nu_{i_1}-\ln\lambda_{i_1})+(\ln\nu_{i_2}-\ln\lambda_{i_2})+\ldots+(\ln\nu_{i_m}-\ln\lambda_{i_m})\leqslant$$

$$\leqslant\ln\mu_1+\ln\mu_2+\ldots+\ln\mu_m.$$

И в этом случае при $m=n$ достигается равенство, ибо $|C|=|A||B|$.

Далее поступаем так же, как в теореме 10.

Отметим также следующую теорему о спектре произведения унитарных матриц, принадлежащую А. А. Нудельману и П. А. Шварцман (УМН, т. 13, вып. 6 (84), 1958).

Пусть $0\leqslant\varphi_1\leqslant\varphi_2\leqslant\ldots\leqslant\varphi_n<2\pi$ — аргументы собственных чисел унитарной матрицы U, и пусть $0\leqslant\psi_1\leqslant\psi_2\leqslant\ldots\leqslant\psi_n<2\pi$ — аргументы собственных чисел унитарной матрицы V. Пусть, кроме того, $\varphi_n+\psi_n-\varphi_1-\psi_1<2\pi$. Обозначим через N_1 выпуклую оболочку, натянутую на $n!$ векторов $(\varphi_1+\psi_{i_1}, \varphi_2+\psi_{i_2}, \ldots, \varphi_n+\psi_{i_n})$, а через N_2 — выпуклую оболочку векторов $(\psi_1+\varphi_{i_1}, \psi_2+\varphi_{i_2}, \ldots, \psi_n+\varphi_{i_n})$. Пусть, наконец, $0\leqslant\omega_1\leqslant\omega_2\leqslant\ldots\leqslant\omega_n<2\pi$ — аргументы собственных чисел матрицы UV; тогда точка с координатами $(\omega_1, \omega_2, \ldots, \omega_n)$ принадлежит пересечению выпуклых оболочек N_1 и N_2.

ЛИТЕРАТУРА

А. МОНОГРАФИИ, ОБЗОРЫ, УЧЕБНИКИ

1. Айнс Э. Л., Обыкновенные дифференциальные уравнения, ДНТВУ, 1939, гл. XIX.
2. Ахиезер Н. И., Лекции по теории аппроксимации, Гостехиздат, 1948, гл. I.
3. Ахиезер Н. И. и Крейн М. Г., О некоторых вопросах теории моментов, ДНТВУ, 1938.
4. Бернштейн С. Н., Теория вероятностей, 4-е изд., Гостехиздат, 1946.
5. Бохер М., Введение в высшую алгебру, ОНТИ, 1933.
6. Булгаков Б. В., Колебания, т. I, Гостехиздат, 1949, гл. I.
7. Гантмахер Ф. Р. и Крейн М. Г., Осцилляционные матрицы и ядра и малые колебания механических систем, 2-е изд., Гостехиздат, 1950.
8. Гельфанд И. М., Лекции по линейной алгебре, 2-е изд., Гостехиздат, 1951.
9. Геронимус Я. Л., Теория ортогональных многочленов, Гостехиздат, 1950, гл. II.
10. Граев Д. А., Элементы высшей алгебры, Киев, 1914.
11. Еругин Н. П., Приводимые системы, Труды Матем. ин-та им. В. А. Стеклова, т. XIII (1946).
12. Еругин Н. П., Метод Лаппо-Данилевского в теории линейных дифференциальных уравнений, Изд-во Ленингр. ун-та, 1956.
13. Каган В. Ф., Основания теории определителей, Гос. изд-во Украины, 1922.
14. Клейн Ф., Высшая геометрия, ГОНТИ, 1939, §§ 96—99.
15. Крейн М. Г., Основные положения теории λ-зон устойчивости канонической системы линейных дифференциальных уравнений с периодическими коэффициентами, Сб. «Памяти А. А. Андронова», Изд-во Академии наук СССР, 1955.
16. Крейн М. Г. и Наймарк М. А., Метод симметрических и эрмитовых форм в теории отделения корней алгебраических уравнений, Харьков, 1936.
17. Крейн М. Г. и Рутман М. А., Линейные операторы, оставляющие инвариантным конус в пространстве Банаха, УМН 3, № 1 (23) (1948).
18. Кудрявцев Л. Д., О некоторых математических вопросах теории электрических цепей, УМН 3, № 4 (26) (1948).
19. Курант Р. и Гильберт Д., Методы математической физики, т. I, 3-е изд., Гостехиздат, 1951, гл. I, II.
20. Курош А. Г., Курс высшей алгебры, 3-е изд., Гостехиздат, 1952.
21. Лаппо-Данилевский И. А., а) Теория функций от матриц и систем линейных дифференциальных уравнений, Гостехиздат, 1934.
    б) Mémoires sur la théorie des systèmes des équations différentielles linéaires, т. I, II, III, Труды Физико-матем. ин-та им. В. А. Стеклова, VI—VIII (1934—1936).
22. Ляпунов А. М., Общая задача об устойчивости движения, Гостехиздат, 1950.
23. Малкин И. Г., Теория устойчивости движения, Гостехиздат, 1952.
24. Мальцев А. И., Основы линейной алгебры, Гостехиздат, 1948.
25. Марков А. А., Избранные труды, Гостехиздат, 1948.
26. Мейман Н. Н., Некоторые вопросы расположения корней полиномов, УМН 4, № 6 (34) (1949).
27. Наймарк Ю. И., Устойчивость линеаризованных систем, Изд-во Ленингр. Военно-воздушной инж. академии, 1949.
28. Потапов В. П., Мультипликативная структура J-нерастягивающих матриц функций, Труды Моск. матем. о-ва 4 (1955).
29. Романовский В. И., Дискретные цепи Маркова, Гостехиздат, 1948.
30. Смирнов В. И., Курс высшей математики, т. III, Физматгиз, 1958.
31. Стильтьес Т. И., Исследование о непрерывных дробях, ОНТИ, Харьков, 1936.

32. Фаддеев Д. К. и Соминский И. С., Сборник задач по высшей алгебре, 2-е изд., Гостехиздат, 1949.
33. Фаддеева В. Н., Вычислительные методы линейной алгебры, Гостехиздат, 1950.
34. Фрезер Р., Дункан Б., Коллар А., Теория матриц и ее приложения к дифференциальным уравнениям и динамике, ИЛ, 1950.
35. Харди Г., Литтлвуд Дж., Пойа Г., Неравенства, ИЛ, 1948.
36. Чебышев П. Л., Полное собрание сочинений, т. III, Изд-во АН СССР, 1948.
37. Чеботарев Н. Г. и Мейман Н. Н., Проблема Рауса — Гурвица для полиномов и целых функций, Труды Матем. ин-та им. В. А. Стеклова, т. XXVI, Изд-во АН СССР (1949).
38. Четаев Н. Г., Устойчивость движения, Гостехиздат, 1946.
39. Шапиро Г. М. Высшая алгебра, 4-е изд., Учпедгиз, 1938.
40. Шилов Г. Е., Введение в теорию линейных пространств, Гостехиздат, 1952.
41. Широков П. А., Тензорное исчисление, ГТТИ, 1934.
42. Шрейер О. и Шпернер Е., а) Введение в линейную алгебру в геометрическом изложении, ОНТИ, 1934;
б) Теория матриц, ОНТИ, 1936.
43. Aitken A. C., Determinants and matrices Edinbourg, 5-е изд., 1948.
44. Bodewig E., Matrix calculus, Amsterdam, 1956; New York, Interscience publ., 1959.
45. Cahen G., Eléments de calcul matriciel, Paris, 1955.
46. Collatz L., Eigenwertaufgaben mit technischen Anwendungen, Leipzig, 1949.
47. Cullis C. E., Matrices and determinoids, т. I—III, Cambridge, 1913—1925.
48. Jung H., Matrizen und Determinanten. Eine Einführung, Leipzig, 1953.
49. Gröbner W., Matrizenrechnung, München, 1956.
50. Kowalewski G., Einführung in die Determinantentheorie, Leipzig, 1909.
51. Lichnerowicz A., Algèbre et analyse linéaires, 2 ed., Paris, 1956.
52. Mac Duffee C. C., a) The theory of matrices, Berlin, 1933;
b) Vectors and matrices, New York, 1943.
53. Marden M., The geometry of the zeros of a polynomial in a complex variable, New York, 1949.
54. Mirsky L., An introduction to linear algebra, Oxford, 1955.
55. Muir, Sir Thomas, The theory of determinants, t. I—III, Cambridge, 1906—1923.
56. Muth P., Theorie und Anwendung der Elementarteilertheorie, Leipzig, 1899.
57. Parodi M., Sur quelques propriétés des valeurs caractéristiques des matrices carrées, Memorial des Sciences Mathématiques, № 118, Paris, 1952.
58. Perlis S., Theory of matrices, Cambridge, 1952.
59. Pickert G., Normalformen von Matrize, Enzykl. Math. Wissensch. I. 1.7. Band 1. Algebra und Zahlentheorie, 1. Teil B. Algebra. Heft 3, Teil 1. Leipzig, 1953.
60. Routh E. J., a) Stability of a given state of motion, London, 1877;
b) The advanced part of a treatise on the dynamics of a system of rigid bodies, ч. II, 5-е изд., London, 1892.
61. Schlesinger L., a) Vorlesungen über lineare Differentialgleichungen, Berlin, 1908,
b) Einführung in die Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage, Berlin, 1922.
62. Schmeidler W., Vorträge über Determinanten und Matrizen mit Anwendungen in Physik und Technik, Berlin, 1949.
63. Schwerdtfeger H., Introduction to linear algebra and the theory of matrices, Groningen, 1950.
64. Thrall R. and Tornheim L., Vector spaces and matrices, New York — London, 1957.
65. Turnbull H. W. and Aitken A. C., An introduction to the theory of canonical matrices, London, 1932.
66. Turnbull H. W., The theory of determinants, matrices and invariants, London, 1929.
67. Volterra V. et Hostinsky B., Opérations infinitésimales linéaires, Paris, 1938.
68. Wedderburn J. H. M., Lectures on matrices, New York, 1934.
69. Weyl H., Mathematische Analyse des Raumproblems, Berlin, 1923.
70. Winter A., Spektraltheorie der unendlichen Matrizen, Leipzig, 1929.
71. Zurmühl R., Matrizen und ihre technischen Anwendungen, 3-te Auf., Berlin, 1961.

Б. СПЕЦИАЛЬНЫЕ СТАТЬИ

72. Азбелев Н. и Виноград Р., Процесс последовательных приближений для отыскания собственных чисел и собственных векторов, ДАН **83** (1952), 173—174.
73. Айзерман М. А., Об учете нелинейных функций от нескольких аргументов при исследовании устойчивости системы автоматического регулирования, Автоматика и телемеханика, т. VIII, № 1 (1947).
74. Аржаных И. С., Распространение метода А. Н. Крылова на полиномиальные матрицы, ДАН **81** (1951), 749—752.
75. Атрашенок П. В., Определение произвола в выборе матрицы, приводящей систему линейных дифференциальных уравнений к системе с постоянными коэффициентами, Вестник Ленингр. ун-та, сер. матем., физ. и хим. **2** (1953), 17—29.
76. Березин Ф. А. и Гельфанд И. М., Несколько замечаний к теории сферических функций на симметрических римановых многообразиях, Труды Московского матем. о-ва **5** (1956), 311—351.
77. Булгаков Б. В., Деление прямоугольных матриц, ДАН **85** (1952), 21—24.
78. Вержбицкий Б. Д., Некоторые вопросы теории рядов композиций нескольких матриц, Матем. сб. **5** (47) (1939), 505—512.
79. Вейленд Г., Представление векового уравнения в виде многочлена, перевод с англ., УМН **2**, № 4 (20) (1947), 128—158.
80. Виленкин Н. Я., Об одной оценке максимального собственного значения матрицы, Ученые записки Моск. гос. педагогического ин-та им. В. И. Ленина **CVIII**, № 2 (1957), 55—57.
81. Гантмахер Ф. Р., а) Геометрическая теория элементарных делителей по Круллю, Труды Одесского гос. ун-та, Математика, т. I (1935), 89—108.
    б) К алгебраическому анализу метода акад. А. Н. Крылова преобразования векового уравнения. Труды II Всесоюзного матем. съезда, т. II (1934), 45—48.
    в) On the classification of real simple Lie groups, Матем. сб. **5** (47) (1939), 217—250.
82. Гантмахер Ф. Р. и Крейн М. Г., а) К структуре ортогональной матрицы. Труды Физ.-мат. отдела ВУАН, Киев (1929), 1—8.
    б) О нормальных операторах в эрмитовом пространстве, Известия Физ.-мат. об-ва при Казанском университете, т. IV, вып. 1, сер. 3 (1929—1930), 71—84.
    в) Об одном специальном классе детерминантов в связи с интегральными ядрами Келлога, Матем. сб. **42** (1935), 501—508.
    г) Sur les matrices oscillatoires, Comptes Rendus de l'Acad. des Sciences, Paris, 201 (1935), 577—579.
    д) Sur les matrices oscillatoires et complètement non-négatives, Compositio Mathematica **4** (1937), 445—476.
83. Гельфанд И. М., Лидский В. Б., О структуре областей устойчивости линейных канонических систем дифференциальных уравнений с периодическими коэффициентами, УМН **10**, № 1 (63) (1955).
84. Гершгорин С. А., Ueber die Abgrenzung der Eigenwerte einer Matrix, ИАН, сер. физ.-матем. (1931), 749—754.
85. Голубчиков А. Ф., а) Об одном матричном уравнении, Учен. зап. Сталингр. пед. ин-та, № 3 (1953), 71—82.
    б) О структуре автоморфизмов комплексных простых групп Ли, ДАН **77**, 1 (1951), 7—9.
86. Граве Д. А., Малые колебания и некоторые предложения алгебры, ИАН, сер. физ.-матем. (1929), 563—570.
87. Гроссман Д. П., К задаче численного решения систем совместных линейных алгебраических уравнений, УМН **5**, № 3 (37) (1950), 87—103.
88. Данилевский А. М., О численном решении векового уравнения, Матем. сб. **2** (44) (1937), 169—172.
89. Дмитриев Н. А. и Дынкин Е. Б., а) О характеристических числах стохастических матриц, ДАН **49** (1945), 159—162.
    б) Характеристические корни стохастических матриц, ИАН, матем. серия, **10** (1946), 167—194.
90. Донская Л. И., Построения решения линейной системы в окрестности регулярной особой точки в особых случаях, Вестник Ленингр. ун-та (1952), № 6.
    б) О структуре решений системы линейных дифференциальных уравнений в окрестности регулярной особой точки, Вестник Ленингр. ун-та (1954), № 8, 55—64.

91. Дубнов Я. С., а) О совместных инвариантах системы аффиноров, Труды Всесоюзн. съезда матем. в Москве (1927), 236—237.
б) О симметрично сдвоенных ортогональных матрицах, М., Изв. асс. ин-тов ун-та (1927), 33—35.
в) О матрицах Дирака, М., Ученые зап. ун-та **2 : 2** (1934), 43—48.
91′. Дубнов Я. С. и Иванов В. К., О понижении степени аффинорных полиномов, ДАН СССР **41** (1943), 99—102.
92. Еругин Н. П., а) Sur la substitution exposante pour quelques systèmes irrégulières, Матем. сб. **42** (1935), 745—753.
б) Показательная подстановка иррегулярной системы линейных дифференциальных уравнений, ДАН СССР **17** (1937), 235—236.
в) О проблеме Римана для системы Гаусса. Л., Ученые зап. пед. ин-та **28** (1939), 293—304.
93. Ершов А. П., Об одном методе обращения матриц, ДАН СССР **100** (1955), 209—211.
93′. Каган В. Ф., О некоторых системах чисел, к которым приводят лоренцовы преобразования. М., Изв. асс. ин-тов ун-та (1927), 3—31.
94. Карпелевич Ф. И., О характеристических корнях матрицы с неотрицательными элементами, ИАН, матем. сер., **15** (1951), 361—383.
95. Коваленко К. Р. и Крейн М. Г., О некоторых исследованиях А. М. Ляпунова по дифференциальным уравнениям с периодическими коэффициентами ДАН **75** (1950), 495—499.
96. Колмогоров А. Н., Цепи Маркова со счетным множеством возможных состояний. М., Бюлл. ун-та (А) **1 : 3** (1937).
97. Котелянский Д. М., а) Про монотонні n-го порядка функції від матриць. Труды Одеського державного ун-ту, Збірник матем. відділу, т. 3 (1941), 103—114.
б) К теории неотрицательных и осцилляционных матриц., Укр. матем. журнал, **2**, № 2 (1950), 94—101.
в) О некоторых свойствах матриц с положительными элементами, Матем. сб. **31** (**73**) (1952), 497—506.
г) Об одном свойстве знакосимметрических матриц, УМН 8, № 4 (56) (1953), 163—167.
д) О некоторых достаточных признаках вещественности и простоты матричного спектра, Матем. сб. **36**, № 1 (1955), 163—168.
е) О влиянии преобразования Гаусса на спектры матриц, УМН **10**, № 1 (1955), 117—121.
ж) О расположении точек матричного спектра, Укр. матем. журнал **7**, № 2 (1955), 131—133.
з) Оценки для определителей матриц с преобладающей главной диагональю, ИАН СССР, сер. матем., **20**, № 1 (1956), 137—144.
98. Кравчук М. Ф., а) До теорії перемінних матриць, Киев, Зап. физ.-матем. отд. АН УССР **1 : 2** (1924), 28—33.
б) До загальної теорії білінійних форм, Київ, Изв. Политехн. с.-х. ин-та **19** (1924), 17—18.
в) Про одне перетворення квадратичних форм, Київ. Зап. физ.-матем. отд. АН УССР **1 : 2** (1924), 87—90.
г) Про квадратичні форми та лінійні перетворення, Київ, Зап. физ.-матем. отд. АН УССР **1 : 3** (1924), 1—89.
д) Перемінні множини лінійних перетворень. Київ, Зап. с.-госп. ин-ту **1** (1926), 25—58.
е) Ueber vertauschbare Matrizen. Rend. circ. mat., Palermo **51**, (1927), 126—130.
ж) О структуре перестановочных групп матриц. Труды II Всесоюзн. матем. съезда, II (1934), 11—12.
99. Кравчук М. Ф. и Гольдбаум Я. С., а) Про групи коммутативних матриць, Київ, Труды авиац. ин-та (1929), 73—98 (1936), 12—23.
б) Об эквивалентности особенных пучков матриц Киев, Труды авиац. ин-та (1936), 5—27.
100. Красносельский М. А. и Крейн С. Г., Итерационный процесс с минимальными невязками, Матем. сб. **31** (**73**) (1952), 315—334.
101. Крейн М. Г., а) Добавление к работе «К структуре ортогональной матрицы», Труды физ.-мат. отдела ВУАН, Киев (1931), 103—107.
б) О спектре якобиевой формы в связи с теорией крутильных колебаний валов, Матем. сб. **40** (1933), 455—466.
в) Об одном новом классе эрмитовых форм, ИАН, сер. физ.-мат. (1933), 1259—1275.

г) Об узлах гармонических колебаний механических систем некоторого специального типа, Матем. сб. **41** (1934), 339—348.

д) Sur quelques applications des noyaux de Kellog aux problèmes d'oscillations. Сообщ. Харьк. матем. об-ва (4) **11** (1935), 3—19.

е) Sur les vibrations propres des tiges dont l'une des extrémités est encastrée et l'autre libre, там же (4) **12** (1935), 3—11.

ж) Обобщение некоторых исследований А. М. Ляпунова о линейных дифференциальных уравнениях с периодическими коэффициентами, ДАН **73** (1950), 445—448.

з) Об одном применении принципа неподвижной точки в теории линейных преобразований пространств с индефинитной метрикой, УМН **5**, № 2 (36) (1950), 180—190.

и) О применении одного алгебраического предложения в теории матриц монодромии, УМН **6**, № 1 (41) (1951), 171—177.

к) О некоторых вопросах, связанных с кругом идей Ляпунова в теории устойчивости, УМН **3**, № 3 (25) (1948), 166—169.

л) К теории целых матриц функций экспоненциального типа, Украинск. матем. ж. **3**, № 1 (1951), 164—173.

м) О некоторых задачах теории колебаний штурмовых систем, Прикл. матем. и механ. **16**, № 5 (1952), 555—568.

102. Крейн М. Г. и Наймарк М. А., а) Об одном преобразовании безутианты, приводящем к теореме Штурма, Харьков, Зап. матем о-ва (4) **10** (1933), 33—40.

б) О применении безутианты к вопросам отделения корней алгебраических уравнений, Труды Одесского гос. ун-та, Математика, т. I (1935), 51—69.

103. Крылов А. Н., О численном решении уравнения, которым в технических вопросах определяются частоты колебаний материальных систем, ИАН, серия физ.-мат. (1931), 491—539.

104. Лаппо-Данилевский И. А., а) Основные задачи теории систем линейных дифференциальных уравнений с произвольными рациональными коэффициентами, Труды 1-го Всесоюзн. съезда матем., ОНТИ (1936), 254—262.

б) Résolution algorithmique des problèmes réguliers de Poincaré et de Riemann, I, Журн. физ.-матем. о-ва, Л., **2 : 1** (1928), 94—120; II, там же, 121—154.

в) Théorie des matrices satisfaisantes à des systèmes des équations différentielles linéaires à coefficients rationnels arbitraires. Журн. физ.-матем. об-ва, Л., **2 : 2** (1928), 41—80.

105. Лившиц М. С. и Потапов В. П., Теорема умножения характеристических матриц-функций, ДАН СССР **72** (1950), 164—173.

106. Лидский В. Б., а) О собственных значениях суммы и произведения симметрических матриц, ДАН **75** (1950), 769—772.

б) Осцилляционные теоремы для канонической системы дифференциальных уравнений, ДАН **102** (1955), 877—880.

107. Липин Н. В., О регулярных матрицах. Л., Труды ин-та инж. ж.-д. транспорта **9** (1934), 105.

108. Лопшиц А. М., а) Векторное решение задачи о симметрически сдвоенных матрицах, Труды Всеросс. съезда матем. в Москве (1927), 186—187.

б) Характеристическое уравнение наинизшей степени для аффинора и применение его к интегрированию дифференциальных уравнений, Труды сем. по вект. и тенз. исчислению, вып. II—III (1935).

в) Численный метод нахождения собственных значений и собственных плоскостей линейного оператора, там же, т. VII, 233—259.

г) Экстремальная теорема для гиперэллипсоида и ее применение к решению системы линейных алгебраических уравнений, там же, т. IX (1952), 183—197.

109. Лузин Н. Н., а) О методе акад. А. Н. Крылова составления векового уравнения, ИАН, сер. физ.-матем. (1931), 903—958.

б) О некоторых свойствах перемещающего множителя в методе акад. А. Н. Крылова, ИАН, сер. физ.-матем. (1932), I, 596—638; II, 735—762; III, 1065—1102.

в) К изучению матричной теории дифференциальных уравнений, Атом. и телемех. 5 (1940), 3—66.

110. Любич Ю. И., Оценки для оптимальной детерминизации недетерминированных автономных автоматов, Сиб. матем ж. 5 (1964), 337—355.

111. Люстерник Л. А., а) Нахождение собственных значений функций на электрической схеме, Электричество **11** (1946), 67—68.

б) Об электрическом моделировании симметрических матриц, УМН 4 : 2 (30) (1949), 198—200.

Люстерник Л. А. и Прохоров А. М., Определение собственных значений и функций некоторых операторов с помощью электрической цепи, ДАН 55 (1947), 579—582; ИАН, ОФН 11 (1947), 141—145.
113. Маркус А. С., Собственные и сингулярные числа суммы и произведения линейных операторов, УМН, т. XIX, вып. 4 (118) (1965), стр. 93—123.
114. Маянц Л. С., Метод уточнения корней вековых уравнений высоких степеней и численного анализа их зависимости от параметров соответствующих матриц, ДАН 50 (1945), 121—124.
115. Нейгауз М. Г. и Лидский В. Б., Об ограниченности решений линейных систем дифференциальных уравнений с периодическими коэффициентами, ДАН 77 (1951), 183—193.
116. Папкович П. Ф., Об одном методе разыскания корней характеристического определителя, Прикл. матем. и мех. 1 (1933), 314—318.
117. Понтрягин Л. С., Эрмитовы операторы в пространстве с индефинитной метрикой, ИАН, сер. матем. 8 (1944), 243—280.
118. Потапов В. П., О голоморфных ограниченных в единичном круге матрицах-функциях, ДАН 72 (1950), 849—853.
119. Птак В., Об одной комбинаторной теореме и ее применении к неотрицательным матрицам, Чехосл. матем. ж. 8 (1958), 487—495.
120. Романовский В. И., а) Un théorème sur les zéros des matrices non-négatives, Bull. Soc. Math. France 61 (1933), 213—219.
б) Recherches sur les chaînes de Markoff, Acta Math. 66 (1935), 147—251.
121. Рехтман-Ольшанская П. Г., Об одном утверждении академика А. А. Маркова, УМН, т. XII, № 3 (75) (1959), 181—187.
122. Сарымсаков Т. А., О последовательности стохастических матриц, ДАН СССР 47 (1945), 331—333.
123. Севастьянов Б. А., Теория ветвящихся случайных процессов, УМН, т. VI, вып. 6 (1951), 46—99.
124. Семендяев К. А., О нахождении собственных значений и инвариантных многообразий матриц посредством итераций, Прикл. матем. и мех. 3 (1943), 193—221.
125. Смогоржевский А. С., По унітарні типи циркулянтів, Киев, Журн. матем. цикла АН УССР 1 (1932), 89—91.
126. Смогоржевский А. С. и Кравчук М. Ф., Про ортогональні перетворення, Киев, Зап. ин-та нар. просв. 2 (1927), 151—156.
127. Сулейманова Х. Р., а) Стохастические матрицы с действительными характеристическими числами, ДАН 66 (1949), 343—345.
б) О характеристических числах стохастических матриц, Учен. записки Моск. гос. пед. ин-та, сер. матем., 71, вып. 1 (1953), 167—197.
128. Султанов Р. М., Некоторые свойства матриц с элементами из некоммутативного кольца, Баку, Труды сектора матем. АН АзССР 2 (1946), 11—17.
129. Сушкевич А. К., Про деякі типи особливих матриць, Учені записки Харьківського держ. ун-ту 10 (1937), 1—16.
130. Турчанинов А. С., О некоторых приложениях исчисления матриц к линейным дифференциальным уравнениям, Одесса, Учен. зап. высш. шк. 1 (1921), 41—48.
131. Фаге М. К., а) Обобщение неравенства Адамара об определителях, ДАН 54 (1946), 765—768.
б) О симметризуемых матрицах, УМН 6, № 3 (43) (1951), 153—156.
132. Фаддеев Д. К., О преобразовании векового уравнения матрицы, Л., Труды Ин-та инж. пром. строит. 4 (1937), 78—86.
133. Хлодовский И. Н., К теории общего случая преобразования векового уравнения методом академика А. Н. Крылова, ИАН, сер. физ.-мат., 8 (1933), 1077—1102.
134. Хуа Ло-ген, а) Геометрия симметрических матриц над полем действительных чисел, ДАН 53 (1946), I, 99—102; II, 199—200.
б) Автоморфизмы действительной симплектической группы, ДАН 53, (1946), 307—310.
135. Хуа Ло-ген, Б. А. Розенфельд, Геометрия прямоугольных матриц и ее применения к вещественной проективной и неевклидовой геометрии, Известия высших учебных заведений СССР, Математика, 1 (1957), 233—246.
136. Цейтлин М. Л., Применение матричного исчисления к синтезу релейно-контактных схем, ДАН 86 (1952), 525—528.
137. Циммерман Г. К., Разложение нормы матрицы по произведениям норм ее строк, Научные зап. Николаевск. гос. пед. ин-та, 4 (1953), 130—135.

138. Шварцман А. П., О матрицах Грина самосопряженных конечноразностных операторов, Труды Одесского гос. ун-та, Математика, т. III (1941), 35—77.
139. Шиффнер Л. М., а) Разложение интегралов системы дифференциальных уравнений с правильными особыми точками в ряды по степеням элементов дифференциальных подстановок, Труды физ.-мат. ин-та им. Стеклова **9** (1935), 235—266.
б) О степени матрицы, Матем. сб. **42** : **3** (1935), 385—394.
140. Шостак Р. Я., О признаке условной определенности квадратичной формы переменных, подчиненных линейным связям, и о достаточном признаке условного экстремума функции переменных, УМН **9**, № 2 (1954), 199—206.
141. Шрейдер Ю. А., Решение систем линейных совместных алгебраических уравнений, ДАН **76** (1951), 651—655.
142. Штаерман И. Я., Новый метод решения некоторых алгебраических уравнений, которые имеют применения в математической физике и технике, Киев, Журн. ин-та матем. АН УССР **1** (1934), 83—89; **4** (1934), 9—20.
143. Штаерман И. Я. и Ахиезер Н. И., К теории квадратичных форм, Киев, Изв. политехн. с.-х. ин-та **19** (1934), 116—123.
144. Шура-Бура М. Р., Оценка ошибок при численном обращении матриц высокого порядка, УМН **6**, № 4 (44) (1951), 121—150.
145. Яглом И. М., Квадратичные и кососимметрические билинейные формы в вещественном симплектическом пространстве, Труды семинара по векторному и тензорному анализу (Москва, Гостехиздат) **8** (1950), 364—381.
146. Якубович В. А., Некоторые критерии приводимости системы дифференциальных уравнений, ДАН **66** (1949), 577—580.
147. Afriat S., Composite matrices, Quart. J. Math. **5**, N 18, 1954, 81—98.
148. Aitken A. C., Studies in practical mathematics. The evaluation with applications of a certain triple product matrix, Proc. Roy. Soc., Edinbourgh **57**, 1936—1937.
149. Amir Moéz Ali R., Extreme properties of eigenvalues of a hermitian transformation and singular values of the sum and product of linear transformations. Duke Math. J. **23** (1956), 463—476.
150. Baker H. F., On the integration of linear differential equations, Proc. Lond. Math. Soc. **35** (1903), 333—378.
151. Barankin E. W., Bounds for characteristic roots of a matrix, Bull. Amer. Math. Soc. **51** (1945), 767—770.
152. Bartsch H., Abschätzungen für die kleinste charakteristische Zahl einer positiv-definiten Matrix, ZAM **34**, № 1—2 (1954), 72—74.
153. Birkhoff G. D., a) Equivalent singular points of ordinary linear differential equations, Math. Ann. **74** (1913), 134—139.
b) Tres observations sobre el algebra lineal, Revista Universidad Nacional Tucuman, ser. A, 5 (1946), 147 — 151.
154. Birkhoff Garrett, On product integration, Journ. of Math. and Phys., т. XVI (1937), 104—132.
155. Bellman R., Notes on matrix theory, Amer. Math. Monthly **60** (1953), 173—175; **62**, № 3 (1955), 172—173; № 8 (1955), 571—572; № 9 (1955), 647—648; **64**, № 3 (1957), 189—191.
156. Bellman R., Hoffman A., On a theorem of Ostrowski, Arch. Math. **5**, № 1—3 (1954), 123—127.
157. Bendat J., Scherman S., Monotone and convex operator functions, Trans. Amer. Math. Soc. **9**, № 1 (1955), 58—71.
158. Berge C., Sur une propriété des matrices doublement stochastiques, C. r. Acad. Sci. **241**, № 3 (1955), 269—271.
159. Bjerhammar A., Rectangular reciprocal matrices, with specially references to geodesic calculus, Bull. geod. int. (1951), 188—220.
160. Bott R., Duffin R., On the algebra of networks, Trans. Amer. Math. Soc. **74**, № 1 (1953), 99—109.
161. Brauer A., Limits for the characteristic roots of a matrix, Duke Math. J. I, **13** (1946), 387—395; II, **14** (1947), 21—26; III, **15** (1948), 871—877; IV, V, **19** (1952), 73—91, 553—563; VI, **22**, 1955, 387—395.
b) Über die Lage der charakteristischen Wurzeln einer Matrix, J. reine und angew. Math. **192**, № 2 (1953), 113—116.
c) Bounds for the rations of the coordinates of the characteristic vectors of a matrix, Proc. Nat. Acad. Sci. USA **41**, № 3 (1955), 162—164.
d) The theorems of Ledermann and Ostrowski on positive matrices Duke, Math. J. **24**, № 2 (1957), 265—274.

162. Brenner J., Bounds for determinants, Proc. Nat. Acad. Sci. USA **40** (1954), 452—454; Proc. Amer. Math. Soc. **8** (1957), 532—534; C. R. Acad. Sci. **238** (1954), 555—556.
163. Bruijn N., Inequalities concerning minors and eigenvalues, Nieuw. arch. wiskunde 4, № 1 (1956), 18—35.
164. Bruijn N., Szekeres C., On some exponential and polar representations of matrices, Nieuw arch. wiskunde **3**, № 1 (1955), 20—32.
165. Cayley A., A memoire on the theory of matrices, London, Phil. Trans. **148** (1857), 17—37; Coll Works **2**, 476—496.
166. Coheen H. E., On a lemma of Stiltjes on matrices, Amer. Math. Monthly **56** (1949), 328—329.
167. Collatz L., a) Einschließungssatz für die charakteristischen Zahlen von Matrizen, Math. Zs. **48** (1942), 221—226.
b) Über monotone Systeme linearer Ungleichungen, J. reine und angew. Math. **194**, № 1—4 (1955), 193—194.
168. Cremer H., Die Verringerung der Zahl der Stabilitätskriterien bei Voraussetzung positiver Koeffizienten der charakteristischen Gleichung, ZAM **33**, № 7 (1953), 222—227.
169. Cremer H., Effertz F. H., Über die algebraischen Kriterien für die Stabilitat von Regelungssystemen. Math. Ann. **137**, 4 (1959), 328—350.
170. Diliberto S., On system of ordinary differential equations. Contributions to the theory of non-linear oscilations, Princeton (1950), 1—38.
171. Dobsch O., Matrixfunktionen beschränkter Schwankung, Math. Zs. **43** (1937), 353—388.
172. Dulmage L. and Halperin I., On a theorem of Frobenius — König and J. von Neumann's game of hide and seek, Trans. Roy. Soc. Canada, ser. III, **49** (1955), 23—29.
173. Duncan W., Reciprocation of triply-partitioned matrices, J. Roy. Aeronaut. Soc. **60**, № 542 (1956), 131—132.
174. Egervary E., a) On hypermatrices whose blocks are commetable in pairs and their application in lattice-dynamics, Acta scient. math. **15**, № 3—4 (1954), 211—222.
b) On a lemma of Stieltjes on matrices, Acta Scient. Math. **15**, № 2 (1954), 99—103.
175. Epstein M., Flanders H., On the reduction of a matrix to diagonal form, Amer. Math. Monthly **62**, № 3 (1955), 168—171.
176. Faedo S., Un nuovo problema di stabilita per le equazioni algebriche a coefficienti reali, Aun. Scuola notm. super. Pisa Sci. fis. e. mat. **7**, № 1—2 (1953), 53—63.
177. Fan Ky, a) On a theorem of Weyl concerning eigenvalues of linear transformations, Proc. Nat. Acad. Sc. I, **35** (1949), 652—655; II, 36 (1950), 31—35.
b) Maximum properties and inequalities for the eigenvalues of completely continuous operators, Proc. Nat. Acad. Sci. USA **37** (1951), 760—766.
c) A comparision theorem for eigenvalues of normal matrices, Pacific J. Math. **5** (1955), 911—913.
d) Some inequalities concerning positive-definite Hermitian matrices, Proc. Cambridge Philos. Soc. **51**, № 3 (1955), 414—421.
178. Fan Ky, Hoffman A., Some metric inequalities in the space of matrices, Proc. Amer. Math. Soc. **6**, № 1 (1955), 111—116.
179. Fan Ky, Housenholder A. S., A note concerning positive matrices and M-matrices. Monatsh. Math. **63**, № 3 (1959), 265—270.
180. Fan Ky, Paul Gordon, Imbedding conditions for Hermitian and normal matrices, Canad. J. Math. **9** (1957), 298—304.
181. Fan Ky, Todd J., A determinantial inequality, J. London Math. Soc. **30**, № 1 (1955), 58—64.
182. Frobenius G., a) Ueber lineare Substitutionen und bilineare Formen, J. reine und angew. Math. **84** (1877), 1—63.
b) Ueber die cogredienten Transformationen der bilinearen Formen, Sitz.-Ber. Akad. Wiss. Phys.-math. Klasse, Berlin (1896), 7—16.
c) Ueber die vertauschbaren Matrizen, ibidem (1896), 601—604.
d) Ueber Matrizen aus positiven Elementen, ibidem (1908), 471—476; (1909), 514—518.
e) Ueber Matrizen aus nicht negativen Elementen, ibidem (1912), 456—477.
f) Ueber das Trägheitsgesetz der quadratischen Formen, ibidem (1894), 241—256; 407—431.

183. Gautschi W., Bounds of matrices with regard to an hermitian metric, Compositio math. 12, № 1 (1954), 1—16.
184. Goddard L., An extension of a matrix theorem of A. Brauer, Proc. Intern Congr. Math., 1954, 2, Amsterdam (1954), 22—23.
185. Haynsworth E., Bounds for determinants with dominant main diagonal, Duke Math. J. 20, № 2 (1953), 199—209.
186. Hellman O., Die Anwendung des Matrizanten bei Eigenwertaufgaben, ZAM 35, № 8 (1955), 300—315.
187. Hjelmslev J., Introduction à la théorie des suites monotones, Oversight over Kgl. Danske Videnskaberns Selbskabs Forhandlinger, 1914, 1—74.
188. Hoffman A., Taussky Olga, A characterisation of normal matrices, J. Res. Nat. Bur. Standards 52, № 1 (1954), 17—19.
189. Hoffman A., Wielandt H., The variation of the spectrum of a normal matrix, Duke Math. J. 20, № 1 (1953), 37—39.
190. Holladay I., Varga R., On powers of non-negative matrices, Proc. Amer. Math. Soc. 9 (1958), 631—634.
191. Horn A., a) On the eigenvalues of a matrix with prescribed singular values, Proc. Amer. Math. Soc. 5, № 1 (1954), 4.
b) On the singular values of product of completely continuous operators, Proc. Nat. Acad. USA 36, № 7 (1950), 374—375.
c) Eigenvalues of sums of Hermitian matrices, Pacific Journal of Math. 12, № 1 (1962), 225—241.
d) Doubly stochastic matrices and the diagonal of a rotation matrix, Amer. Journ. Math. 76, № 3 (1954), 620—630.
192. Hsu P. L., a) On symmetric, orthogonal, and skew symmetric matrices, Proc. Edinburgh Math. Soc., ser. 2, 10 (1953), 37—44.
b) On a kind of transformations of matrices, Acta Math. Sinica 5, № 3, 333—347.
193. Hoteling H., Some new methods in matrix calculation, Ann. Math. Stat. 14, № 1 (1943).
194. Hua Loo-keng, a) On the theory of automorphic functions of a matrix variable, Amer. J. Math. 66 (1944), I, 470—488; II, 531—563.
b) Geometries of matrices, Trans. Amer. Math. Soc. 57 (1945), 441—490.
c) Orthogonal classification of Hermitian matrices, Trans. Amer. Math. Soc. 59 (1946), 508—523.
d) Inequalities involving determinants, Acta Math. Sinica 5 (1955), 463—470.
195. Hermite C., Sur le nombre des racines d'une équation algébrique comprise entre des limites données, J. für die reine und angew.Math. 52 (1856), 39—51.
196. Hurwitz A., Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, Math. Ann. 46 (1895), 273—284.
197. Ingraham M. H., On the reduction of a matrix to its rational canonical form, Bull. Amer. Math. Soc. 39 (1933), 379—382.
198. Jonescu D., O identitate importantá ai descompunere a unei forme bilineare intro sumá de produse, Gaz., mat. si fiz. A7, № 7 (1955), 303—312.
199. Jongmans F., Problemes matriciels lies an ring, Bull, de la Soc. Roy. sci de Liege 29 (1960), 3—4, 51—60.
200. Ishak M., Sur les spectres des matrices, Sémin. P. Dubreil et Ch. Pisot. Fac. sci. Paris 9, № 14 (1955—1956), 1—14.
201. Khan N. A., The characteristic roots of the product of matrices, Quart. J. Math. 7, № 26 (1956), 138—143.
202. Kowalewski G., Natürliche Normalformen linearer Transformationen, Leipz. Ber. 69 (1917), 325—335.
203. König D., Ueber Graphen und ihre Anvendungen, Math. Ann. (1916), 453—465.
204. Kraus F., Ueber konvexe Matrixfunktionen, Math. Zs. 41 (1936), 18—42.
205. Kronecker L., Algebraische Reduktion der Scharen bilinearer Formen, Sitzungsber. Akademie Berlin (1890), 763—776.
206. Krull W., Theorie und Anwendung der verallgemeinerten Abelschen Gruppen, Sitzb. Heidl. Akad. (1926), 1.
207. Ledermann W., a) Reduction of singular pencils of matrices, Proc. Edinb. Math. Soc., cep. 2, 4 (1935), 92—105.
b) Bounds for the greatest latent root of positive matrix, J. of Lond. Math. Soc. 25 (1950), 265—268.
208. Liénard et Chipart, Sur la signe de la partie réelle des racines d'une équation algébrique, J. de Math. pure et appl. (6) 10 (1914), 291—346.
209. Löwner K., a) Ueber monotone Matrixfunktionen, Math. Zs. 38 (1933), 177—216.

b) Some classes of functions defined by difference or differential inequalities, Bull. Amer. Math. Soc. **56** (1950), 308—319.

210. Marcus M., a) A remark on a norm inequality for square matrices, Proc. Amer. Math. Soc. **6**, № 1 (1955), 117—119.
b) An eigenvalue inequality for the product of normal matrices, Amer. Math. Monthly **63**, № 3 (1956), 173—174.
211. Marcus M., McGregor J. L., Extremal properties of Hermitian matrices, Canad. J. Math. **8** (1956), 524—531.
212. Mirsky L., a) An inequality for positive definite matrices, Amer. Math. Monthly **62**, № 6 (1955), 428—430.
b) The norm of adjugate and inverse matrices, Arch. Math. **7** (1956), 276—277.
c) The spread of a matrix, Mathematica **3** (1956), 127—130.
d) Inequalities for normal and Hermitian matrices, Duke Math. J. **24**, № 4 (1957), 591—599.
e) Symmetric gange functions and unitarily invariant norms, Quart. J. Math. **11** (1960), 50—59.
213. Mitrović D., Conditions graphiques pour que toutes les racines d'une équation algébrique soient à parties réelles négatives, C. r. Acad. Sci. **240**, № 11 (1955), 1177—1179.
214. Moore, Bull. Amer. Math. Soc. **26** (1920), 394—395.
215. Morgenstern D., Eine Verschärfung der Ostrowski'schen Determinantenabschätzung, Math. Zs. **66** (1956), 143—146.
216. Motzkin T., Taussky Olga, Pairs of matrices with property W, Trans. Amer. Math. Soc. I, **73**, № 1 (1952), 108—114; II, **80**, № 2 (1955), 387—401.
217. Nagy Bela, Remark on S. N. Roy's paper «A useful theorem in matrix theory», Proc. Amer. Math. Soc. **7**, № 1 (1956).
218. Neumann J., a) Approximative of matrices of high order, Portug. Math. **3** (1942), 1—62.
b) Some matrix-inequalities and metrization of matrix-space, Известия Научно-исслéдовательского ин-та математики и механики при Томском государств. ун-те им. В. В. Куйбышева, т. 1, вып. 3 (1937).
219. Okamoto M., On a certain type of matrices with an application to experimental design., Osaka Math. J. **6**, № 1 (1954), 73—82.
220. Oppenheim A., Inequalities connected with definite Hermitian forms, Amer. Math. Monthly **61**, № 7 (1954), 463—466.
221. Orlando L., Sul problema di Hurwitz relativo alle parti realli delle radici di un'equazione algebrica, Math. Ann. **71** (1911), 233—245.
222. Ostrowski A., a) Bounds for the greatest latent root of a positive matrix, J. of London Math. Soc. **27** (1952), 253—256.
b) Sur quelques applications des fonctions convexes et concaves au sens de J. Schur, J. Math. pures et appl. **31** (1952), 253—292.
223. Ostrowski A., c) On nearly triangular matrices, J. Res. Nat. Standards **52**, № 6 (1954), 344—345.
d) On the spektrum of one-parametric family of matrices, J. reine und angew. Math. **193**, № 3/4 (1954), 143—160.
e) Sur les determinants à diagonale dominante, Bul. Soc. math. Belgique **7**, № 1 (1955), 46—51.
f) Note on bounds for some determinants, Duke Math. J. **22**, № 1 (1955), 95—102.
g) Über Normen von Matrizen, Math. Zs. **63**, № 1 (1955), 2—18.
224. Papulis A., Limits on the zeros of a network determinant, Quart. Appl. Math. **15**, № 2 (1957), 193—194.
225. Parodi M., a) Remarques sur la stabilité, C. R. Acad. Sci. Paris **228** (1949), 51—52, ibidem, 807—808, 1198—1200.
b) Sur une propriété des racines d'une equation qui intervient en mécanique, C. R. Acad. Sci. **241**, № 16 (1955), 1019—1021.
c) Sur la localisation des valeurs caracteristiques des matrices dans le plan complexe, C. R. Acad. sci. **242**, № 22 (1956), 2617—2618.
226. Peano G., Intégration par séries des équations différentielles linéaires, Math. Ann. **32** (1888), 450—456.
227. Penrose R., a) A generalized inverse for matrices, Proc. Cambridge Phil. Soc. **51**, № 3 (1955), 406—413.
b) On best approximate solutions of linear matrix equations, Proc. Cambridge Phil. Soc. **52**, № 1 (1956), 17—19.
228. Perfect H., a) On matrices with positive elements, Quart. J. Math. **2** (1951), 286—290.

b) On positive stochastic matrices with real characteristic roots, Proc. Cambridge Phil. Soc. 48 (1952), 271—276.
c) Methods of constructing certain stochastic matrices, I, Duke Math. J. 20, № 3 (1953), 395—404; II, 22, № 2 (1955), 303—311.
d) A lower bound for the diagonal elements of a non-negative matrix, J. London Math. Soc. 31 (1956), 491—493.
229. Perkins P., A theorem on regular matrices, Pacif. J. Math. 11 (1961), 1529—1533.
230. Perron O., a) Jacobischer Kettenbruchalgorithmus, Math. Ann. 64 (1907), 1—76.
b) Ueber Matrizen, ibidem, 64 (1907), 248—263.
231. Phillips H. B., Functions of matrices, Amer. J. Math. 41 (1919), 266—278.
232. Pignani, On certain matrix equations, Amer. Math. Monthly (1957), 573—576.
233. Polya G., Remark on Weyl's note, Proc. Nat. Acad. USA 36 (1950), 49—50.
234. Rado R., An ineguality, J. London Math. Soc. 27 (1952), 1—6.
235. Rasch G., Zur Theorie und Anwendung des Produktintegrals, J. für die reine und angew. Math. 171 (1934), 65—119.
236. Rham G. de, Sur un théorème de Stieltjes relatif à certain matrices, Acad. Sci. Serbe, Publ. Inst. Math., 1952, 133—154.
237. Richter H., a) Über Matrixfunktionen, Math. Ann. 122 (1950), 16—35.
b) Bemerkung zur Norm der inversen einer Matrix, Arch. Math. 5, № 4—6 (1954), 447—448.
c) Zur Abschätzung von Matrizennormen , Math. Nachrichten 18 (1959), 178—187.
238. Roth W., a) On the characteristic polynomial of the product of two matrices, Proc. Amer. Math. Soc. 5, № 1 (1954), 1—3.
b) On the characteristic polynomial of the product of several matrices, Proc. Amer. Math. Soc. 7, № 4 (1956), 578—582.
239. Roy S., A useful theorem in matrix theory, Proc. Amer. Math. Soc. 5, № 4 (1954), 635—638.
240. Schneider H., a) An inequality for latent rosts applied to determinants with dominant principal diagonal, J. London Math. Soc. 28, 1, № 109 (1953), 8—20.
b) A pair of matrices with property W, Amer. Math. Monthly 62, № 4 (1955), 247—249.
c) A matrix problem concerning projections, Proc. Edinbourgh Math. Soc. 10, № 3 (1956), 129—130.
d) The elementary divisors, associated with 0, of a singular I-matrix, Proc. Edingbourgh Math. Soc. 10, № 3 (1956), 108—122.
241. Schoda K., Über mit einer Matrix vertauschbaren Matrizen, Math. Zs. 29 (1929), 696—712.
242. Schoenberg J., a) Ueber variationsvermindernde lineare Transformationen, Math. Zs. 32 (1930), 321—328.
b) Zur Abzählung der reellen Wurzeln algebraischer Gleichungen, Math. Zs. 38 (1933), 546.
243. Schoenberg T., Whitney A., A theorem on polygons in dimensions with application to variation-diminishing linear transformations, Comp. Math. 9 (1951), 141—160.
244. Schur J., Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen, Math. Ann. 66 (1909), 488—510.
245. Sedláček I., O incidenčnich maticich orientovaných grafů, Časop. pěst. mat. 84 (1959), 303—316.
246. Siegel C. L., Symplectic Geometry, Am. J. Math. 65 (1943), 1—86.
247. Stenzel H., Ueber die Darstellbarkeit einer Matrix als Produkt von zwei symmetrischen Matrizen, Math. Zs. 15 (1922).
248. Stöhr A., Oszillationstheoreme für die Eigenvectoren spezieller Matrizen, J. reine und angew. Math., B. 185, № 3 (1943).
249. Taussky Olga, a) Bounds for characteristic roots of matrices, Duke Math. J. 15 (1948), 1043—1044.
b) A determinantal inequality of H. P. Robertson, J. Washington Acad. of Science 47, № 8 (1957), 263—264.
250. Toeplitz O., Das algebraische Analogen zu einem Satz von Fejer, Math. Zs. 2 (1918), 187—197.
251. Turnbull H. W., On the reduction of singular matrix pencils, Proc. Edinbourgh Math. Soc., cep. 2, 4 (1935), 67—76.

252. Vivier M., Note sur les structure unitaires et paraunitaires, C. R. Acad. sci. 240, № 10 (1955), 1039—1041.
253. Volterra V., Sui fondamenti della teoria delle equazioni differenziali lineari, Mem. Soc. Ital., Sci. (3) 6 (1887), 1—104; (3) 12 (1902), 3—68.
254. Walker A., Weston J., a) Inclusion theorems for the eigenvalues of a normal matrix, J. London Math. Soc. 24 (1944), 28—31.
b) Ein Einschliesungssatz für charakteristische Wurzeln normaler Matrizen, Arch. der Math. 1, 1948/49, 348—352.
c) Die Einschliesung von Eigenwerten normaler Matrizen, Math. Ann. 121 (1949), 234—241.
255. Weierstrass K., Zur Theorie der bilinearen und quadratischen Formen, Monatsh. Akad. Wissenschaft, Berlin (1867), 310—338.
256. Wellstein J., Ueber symmetrische, alternierende und orthogonale Normalformen von Matrizen, J. für die reine und angew. Math. 163 (1930), 166—182.
257. Weyl H., Inequalities between the two kinds of eigenvalues of a linear transformation, Proc. Nat. Acad. Sci. 35 (1949), 408—411.
258. Weyr E., Zur Theorie der bilinearen Formen, Monatsh. für Math. und Phys. (1890), 163—236.
259. Whitney A., A reduction theorem for totaly positive matrices, J. Analyse Math. 2 (1952), 88—92.
260. Wielandt H., a) Unzerlegbare, nicht negative Matrizen, Math. Zs. 52 (1950), 642—648.
b) Lineare Scharen von Matrizen mit reellen Eigenwerten, Math. Zs. 53 (1950), 219—225.
c) Pairs of normal matrices with property W, J. Res. Nat. Bur. Standards 51, № 2 (1953), 89—90.
d) Inclusions theorems for eigenvalues, Nat. Bur. Standards, Appl. Math. № 29 (1953), 75—78.
e) An extremum property of sums of eigenvalues, Proc. Amer. Math. Soc. 6, № 1 (1955), 106—110.
f) On eigenvalues of sums of normal matrices, Pacif. J. Math. 5, № 4 (1955), 633—638.
261. Wong Y., a) An inequality for Minkowski matrices, Proc. Amer. Math. Soc. 4, № 1 (1953), 137—141.
b) On non-negativ-valued matrices, Proc. Nat. Acad. Sci. USA 40, № 2 (1954), 121—124.
262. Winter A., On criteria for linear stability, J. Math. Mech. 6 (1957), 301—309.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Алгоритм Гаусса 41
— — обобщенный 55
— —, механическая интерпретация 45
— Рауса 472

Базис векторного пространства 66
— ортонормированный 224

Вектор 65
— главный пучка квадратичных форм 281
— — — эрмитовых форм 301
— собственный линейного оператора 82
— — матрицы 82
—, аннулирующий многочлен 171
—, координаты 67
—, минимальный многочлен 172
—, норма 409
—, относительный аннулирующий многочлен 177
—, — минимальный многочлен 177
Векторы линейно зависимые и линейно независимые см. Зависимость линейных векторов
Вероятность абсолютная 390
— — предельная 391
— — — средняя 393
— переходная 381
— — предельная 386
— — — средняя 393
Вычитание матриц 16

Делитель элементарный квадратной матрицы 147
— — λ-матрицы 145
Дефект линейного оператора 77
— матрицы 77
Дискриминант квадратичной формы 267
— эрмитовой формы 298
Дополнение ортогональное 241

Жорданова форма см. Форма нормальная квадратной матрицы жорданова (верхняя)

Зависимость линейная векторов 66
— — —, критерий Грама 225
Задача Рауса — Гурвица обобщенная 533
Закон инерции квадратичных форм 269
— — эрмитовых форм 299

Индекс imprimitивности 377
— Коши 469
Интеграл мультипликативный 434
— — в комплексной области 439
— от матрицы 127

Компонента линейного оператора кососимметрическая 255
— — — симметрическая 255
Конгруэнтность пучков симметрических матриц 345
— симметрических матриц 269
Координаты вектора 67
Корень из матрицы 212
Критерий Адамара 406
— — для блочных матриц 413
— Грама линейной зависимости векторов 225
— Льенара и Шипара 508
— Ляпунова 425
— подобия матриц 150
— Рауса 475
— Рауса — Гурвица 486
— Фидлера 416
— эквивалентности λ-матриц 144
— — строгой регулярных пучков матриц 332
— — — сингулярных пучков матриц 344

Линейный оператор см. Оператор линейный
Логарифм матрицы 219
λ-матрица 135
—, инвариантные многочлены 143
—, ранг 143
—, элементарные делители 145
—, — операции (левые и правые) 135 и 136

Матрица 13
— ассоциированная p-я 30
— бесконечная вполне положительная 524
— — ганкелева 495
— блочная 55
— Вандермонда обобщенная 395
— вполне неотрицательная 395
— — положительная 395
— ганкелева 301
— главная пучка квадратичных форм 282
— Гурвица 483
— двояко стохастическая 536
— диагональная 14
— единичная 24
— жорданова (верхняя и нижняя) 153
— идемпотентная 208
— импримитивная 377
— инволютивная 81
— интегральная 418
— квадратная 13
— —, аннулирующий многочлен 98
— —, дефект 77
— —, естественная нормальная форма (первая и вторая) 152
— —, жорданова нормальная форма (верхняя и нижняя) 153, 189 и 154, 190
— —, инвариантные многочлены 147
— —, компоненты 111
— —, минимальный многочлен 98
— —, перестановка рядов 352
— —, порядок 13
— —, разложение на треугольные множители 49
— —, след 96
— —, функция от нее 103
— —, характеристический многочлен 92
— —, элементарные делители 145
— квазидиагональная 56
— квазитреугольная (верхняя и нижняя) 56
— комплексная кососимметрическая 313
— — —, нормальная форма 322
— — неособенная, полярное разложение 317
— — ортогональная 313
— — —, нормальная форма 327
— — симметрическая 313
— — —, нормальная форма 319
— кососимметрическая 29
— Ляпунова 422
— многочленная см. λ-матрица
— — квадратная 87
— неособенная 26
— неотрицательная 352
— неразложимая 352
— —, индекс импримитивности 377
— нильпотентная 208
— нормальная 244
— обратная 26
— ограниченная 132
— ортогональная 238
— особенная 26
— осцилляционная 398
Матрица положительная 352
— преобразующая 150
— примитивная 377
— присоединенная 92
— — приведенная 99
— проекционная 81
— простой структуры 85
— прямоугольная 13
— псевдообратная 32
— разложимая 352
— —, нормальная форма 372
— Рауса 484
— симметрическая 29
— сопряженная 29
— столбцевая 14
— стохастическая 381
— строчная 14
— транспонированная 29
— треугольная (верхняя и нижняя) 28
— унитарная 238, 244
— фундаментальная 86
— характеристическая 92
— эрмитова 244
— якобиева 395
—, интеграл 127
—, корень из нее 212
—, логарифм 219
—, минор 14
—, многочлен от нее 24
—, производная 124
—, ранг 14
—, степень 24
—, функция от нее см. Функция от матрицы
Матрицант 430
Матрицы перестановочные (коммутирующие) 18
— подобные 80
—, вычитание 16
—, сложение 15
—, умножение 17
—, — на число 16
Метод Гревилля нахождения псевдообратной матрицы 39
— Крылова 190
— Лагранжа приведения к сумме квадратов квадратичной формы 271
— — — — — — эрмитовой формы 299
— построения преобразующей матрицы 159
— Фаддеева 96
— Якоби приведения к сумме квадратов квадратичной формы 272
— — — — — — эрмитовой формы 299
Метрика 223
— евклидова 225
— эрмитова 223
— — неотрицательная 223
— — положительно определенная 223
Минор 14
— главный 14
— почти главный 398
Многочлен аннулирующий вектора 171
— — матрицы 98
— — пространства 172

Многочлен инвариантный квадратной матрицы 145
— — λ-матрицы 143
— интерполяционный Лагранжа — Сильвестра 104
— матричный 87
— — регулярный 87
— —, левое и правое деление 87
— минимальный вектора 172
— — — относительный 177
— — матрицы 98
— — пространства 172
— от матрицы 24
— характеристический 82, 92
Модуль линейного оператора левый 249
— — — правый 249

Неравенства Вейля 251, 542
— Неймана — Хорна 539
— Сильвестра 78
Неравенство Адамара 230
— — обобщенное 231
— Бесселя 235
— Буняковского 232
— детерминантное для вполне неотрицательных матриц 396
Норма вектора 409
— линейного оператора 410
— матрицы 410

Оператор линейный 70
— — в $\boldsymbol{R}$ 79
— — вещественный 256
— — инволютивный 81
— — кососимметрический 255
— — неособенный 81
— — нормальный в евклидовом пространстве 254
— — — в унитарном пространстве 242
— — ортогональный 255
— — — второго рода 255
— — — первого рода 255
— — проекционный 81
— — простой структуры 84
— — псевдообратный 265
— — симметрический 254
— — — неотрицательный 254
— — — положительно определенный 254
— — сопряженный 240
— — транспонированный 254
— — унитарный 243
— — эрмитов 243
— — — неотрицательный 248
— — — положительно определенный 248
— —, дефект 77
— —, инвариантное подпространство 173
— —, кососимметрическая и симметрическая компоненты 255
— —, левый и правый модули 249
— —, полярное разложение в евклидовом пространстве 260
— —, полярное разложение в унитарном пространстве 249
Оператор линейный, ранг 77
— —, собственные векторы 82
— —, характеристические (собственные) числа 82
Операторы линейные, сложение 71
— —, умножение 71
— —, умножение на число 73
Операция элементарная 135
— — левая 136
— — правая 136
Определитель Грама 226
— —, геометрический смысл 229
— Гурвица 485
— Маркова 521, 525
Ортогонализация ряда векторов 233

Параметры Маркова 518
Перестановка рядов в квадратной матрице 352
Подпространство 68
— инвариантное 173
— координатное 353
— циклическое 178
Порядок матрицы 13
Предел последовательности матриц 49
Преобразование координат 70
— — ортогональное 238
— — унитарное 238
— линейное 15
— Ляпунова 428
Приведение квадратичной формы к главным осям 279
— — — к сумме квадратов 271
— эрмитовой формы к главным осям 300
Проблема Рауса — Гурвица 468
Проектирование ортогональное 227
Проекционная матрица 81
Проекционный оператор 81
Производная мультипликативная 434
— от матрицы 124
Пространство 65
— бесконечномерное 66
— векторное см. Пространство
— евклидово 225
— конечномерное 66
— унитарное 223
— —, ортонормированный базис 224
—, аннулирующий многочлен 171
—, базис 66
—, метрика см. Метрика
—, минимальный многочлен 172
Псевдообратная матрица 32
Пучок квадратичных форм 281
— — — регулярный 281
— — —, главная матрица 282
— — —, главный вектор 281
— — —, характеристические числа 281
— — —, характеристическое уравнение 281
— матриц 331
— — регулярный 332
— — —, каноническая форма 334
— — сингулярный 332
— — —, каноническая форма 340

Пучок матриц сингулярный, ранг 335
— эрмитовых форм 301
— — — регулярный 301
— — —, главный вектор 301
— — —, характеристические числа 301
— — —, характеристическое уравнение 301

Равенство Парсеваля 237
Разложение квадратной матрицы на треугольные множители 49
— полярное комплексной матрицы 317
— — линейного оператора в евклидовом пространстве 260
— — — — в унитарном пространстве 249
Ранг квадратичной формы 268
— линейного оператора 77
— матрицы 14
— — многочленной 143
— сингулярного пучка матриц 335
— эрмитовой формы 298
Ряд векторов ортогональный 233
— — полный 236
— —, ортогонализация 233

Сигнатура квадратичной формы 270
— эрмитовой формы 299
Система линейных дифференциальных уравнений с переменными коэффициентами 419
— — — — — — — приводимая 423
— — — — — — — — — аналитическая 461
— — — — — — —, интегральная матрица 419
— — — — — — —, преобразование Ляпунова 422
— — — — с постоянными коэффициентами 124
Системы векторов биортонормированные 241
Скелетное разложение матрицы 32
След матрицы 96
Сложение линейных операторов 71
— матриц 15
Собственное число см. Число собственное
Степень матрицы 24
Схема Рауса 474
Сходимость в среднем 236

Теорема Безу обобщенная 92
— Гамильтона — Кэли 93
— Гершгорина 415
— Еругина 426
— Кронекера об ассоциированных матрицах 86
— Ляпунова 479
— Маркова 528
— Неймана — Хорна 541
— Ольги Тауски 409
— о расщеплении вторая 179
— — — первая 173
Теорема о расщеплении третья 181
— Перрона 354
— Рауса 475
— Рауса — Гурвица 483
— Стильтьеса 512
— Фробениуса о ганкелевых формах 305
— — о неотрицательных матрицах 355
— Чебышева — Маркова 531
— Штурма 471
— Шура 242
Тождество Сильвестра детерминантное 47

Умножение линейных операторов 71
— матриц 17
— на число линейного оператора 73
— — — матрицы 16
Уравнение матричное многочленное 209
— характеристическое (вековое) матрицы 82
— — пучка квадратичных форм 281
— — — эрмитовых форм 301

Форма билинейная 267
— квадратичная 225, 267
— — вещественная 267
— — ганкелева 301
— — неотрицательная 276
— — неположительная 276
— — отрицательно определенная 276
— — положительно определенная 225, 276
— — сингулярная 267
— —, дискриминант 267
— —, закон инерции 269
— —, приведение к главным осям 279
— —, приведение к сумме квадратов 271
— —, ранг 268
— —, сигнатура 270
— —, формула Якоби 272
— нормальная квадратной матрицы естественная первая 152
— — — — — вторая 152
— — — — жорданова верхняя 153, 189
— — — — — нижняя 154, 190
— — комплексной кососимметрической матрицы 322
— — — ортогональной матрицы 327
— — — симметрической матрицы 319
— — разложимой матрицы 372
— эрмитова 224, 297
— — билинейная 297
— — неотрицательная 224, 300
— — неположительная 300
— — отрицательно определенная 300
— — положительно определенная 224, 300
— — сингулярная 298
— —, дискриминант 298
— —, закон инерции 299
— —, приведение к главным осям 300
— —, формула Якоби 299
Формула Бине — Коши 20
— Орландо 488

Формула основная для функции от матрицы 111
— Чебышева — Маркова 532
— Якоби для квадратичной формы 272
— — для эрмитовой формы 299
Формулы Кэли 252
— — в евклидовом пространстве 261
Функция от матрицы 103
— — —, основная формула 111
— от многих матриц аналитическая 465

Характеристическое число см. Число характеристическое

Цепочка векторов жорданова 188
— — — нижняя 189
Цепь Маркова однородная 381
— — — ациклическая 386
— — — неразложимая 386
— — — правильная 386
— — — разложимая 386
— — — регулярная 386
Цепь Маркова однородная циклическая 386

Число характеристическое (собственное) линейного оператора 82
— — пучка квадратичных форм 281
— — — — —, экстремальные свойства 286
— — — эрмитовых форм 301
— — — — —, экстремальные свойства 301

Эквивалентность линейных двучленов 148
— λ-матриц 137
— — левая 137
— — правая 137
— —, критерий 144
— рядов векторов 233
— строгая пучков матриц 331
— — — — регулярных, критерий 332
— — — — сингулярных, критерий 344
Элементарный делитель см. Делитель элементарный