

Глава 3. ПРИБЛИЖЕНИЕ ФУНКЦИЙ.

Пусть на отрезке $[a, b]$ определена некоторая функция $y = f(x)$, однако полная информация о ней недоступна. Известны лишь ее значения в конечном числе точек x_0, x_1, \dots, x_n , этого отрезка, которые мы будем считать занумерованными в порядке возрастания:

$$a \leq x_0 < x_1 < \dots < x_i < x_{i+1} < \dots < x_n \leq b. \quad (1)$$

Требуется по известным значениям

$$y_i = f(x_i), \quad i = 0, 1, \dots, n \quad (2)$$

«восстановить», хотя бы приближенно, исходную функцию $y = f(x)$, то есть построить на отрезке $[a, b]$ функцию $F(x)$, достаточно близкую к $f(x)$. Функцию $F(x)$ принято называть интерполирующей функцией, точки $x = x_0, x = x_1, \dots, x = x_n$ - узлами интерполяции.

Подобные задачи часто возникают на практике, например, при обработке экспериментальных данных, когда значения переменной y , зависящей от x , измеряется в конечном числе точек x_i : $y_i = f(x_i)$, ($i = 0, 1, \dots, n$) или при работе с табличными функциями, если требуется вычислить $y = f(x)$, при значениях аргумента, не совпадающего ни с одним из табличных x_i .

Поставленный выше в общей форме вопрос о приближении функций является достаточно сложным. Существует не один подход к его решению. Мы ограничимся изложением трех наиболее распространенных методов.

§1. Интерполирование.

1.1. Классическая постановка задачи интерполирования.

Выберем некоторую систему функций $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$, заданных на отрезке $[a, b]$, и будем строить $F(x)$ как их линейную комбинацию:

$$F(x) = \sum_{i=0}^n c_i \varphi_i(x), \quad (3)$$

где числовые коэффициенты c_i ($i = 0, 1, \dots, n$) подлежат определению, согласно условиям:

$$F(x_j) = f(x_j), \quad j = 0, 1, \dots, n. \quad (4)$$

Равенства (4) представляют собой систему линейных алгебраических уравнений относительно коэффициентов c_i :

$$\sum_{i=0}^n c_i \varphi_i(x_j) = f(x_j), \quad j = 0, 1, \dots, n$$

или в развернутом виде:

$$\begin{cases} c_0\varphi_0(x_0) + c_1\varphi_1(x_0) + \dots + c_n\varphi_n(x_0) = f(x_0) \\ c_0\varphi_0(x_1) + c_1\varphi_1(x_1) + \dots + c_n\varphi_n(x_1) = f(x_1) \\ \vdots \\ c_0\varphi_0(x_n) + c_1\varphi_1(x_n) + \dots + c_n\varphi_n(x_n) = f(x_n) \end{cases} \quad (5)$$

Для того, чтобы коэффициенты c_i ($i = 0, 1, \dots, n$) можно было определить и притом единственным образом, необходимо и достаточно, чтобы определитель полученной системы линейных уравнений был отличен от нуля:

$$\Delta = \begin{vmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_n(x_n) \end{vmatrix} \neq 0. \quad (6)$$

Определение.

Система функций $\varphi_i(x)$ ($i = 0, 1, \dots, n$), удовлетворяющая при фиксированных значениях x_j ($j = 0, 1, \dots, n$) условию (6), называется Чебышевской.

Очевидно, что для однозначной разрешимости задачи интерполирования в классической постановке необходимо и достаточно, чтобы система функций $\varphi_i(x)$ ($i = 0, 1, \dots, n$) была Чебышевской. Только такие системы функций мы и будем использовать в этой главе. Необходимым условием принадлежности системы функций $\varphi_i(x)$ ($i = 0, 1, \dots, n$) к Чебышевской является их линейная независимость. Однако это условие не является достаточным. Например, для системы из двух линейно независимых функций $\varphi_0(x) = \sin x$, $\varphi_1(x) = \cos x$, с узлами интерполяции $x_0 = 0$, $x_1 = \pi$, определитель

$$\Delta = \begin{vmatrix} \sin x_0 & \cos x_0 \\ \sin x_1 & \cos x_1 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 0 & -1 \end{vmatrix} = 0$$

и данная система функций при выбранных значениях x_0 и x_1 не является Чебышевской.

1.2. Интерполирование полиномами.

При построении интерполирующей функции $F(x)$ в виде (3) функции $\varphi_i(x)$, естественно, выбираются такими, чтобы их вычисление было простым. В частности, широкое распространение получило интерполирование с помощью степенных функций:

$$\varphi_0(x) = 1; \quad \varphi_1(x) = x; \quad \varphi_2(x) = x^2, \quad \dots \quad \varphi_n(x) = x^n.$$

В этом случае интерполирующая функция представляет собой полином степени n :

$$F(x) = P_n(x) = \sum_{i=0}^n c_i x^i \quad (7)$$

с неизвестными коэффициентами c_i ($i = 0, 1, \dots, n$).

Согласно рассмотренной выше общей схеме построения интерполирующей функции, следует потребовать, чтобы коэффициенты c_i с учетом (7) удовлетворяли системе линейных уравнений:

$$\sum_{i=0}^n c_i x_j^i = f(x_j), \quad j = 0, 1, \dots, n. \quad (8)$$

Определителем этой системы является определитель Ван-дер-Монда:

$$\Delta = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} = \prod_{i>j} (x_i - x_j).$$

В нашем случае этот определитель отличен от нуля, поскольку, согласно (1), все узлы интерполирования различны между собой. Итак, интерполирование с помощью полиномов при сделанных в начале главы предположениях всегда осуществимо и притом единственным образом.

Задача 1.

Построить линейный полином

$$P_1(x) = c_0 + c_1 x$$

по заданным узлам интерполяции $x_0 < x_1$ и соответствующим им значениям функции

$$y_0 = f(x_0) \text{ и } y_1 = f(x_1).$$

Линейная система уравнений для определения c_0 и c_1 в данном случае имеет вид:

$$\begin{aligned} c_0 + c_1 x_0 &= f(x_0), \\ c_0 + c_1 x_1 &= f(x_1). \end{aligned}$$

Определитель этой системы равен $\Delta = x_1 - x_0 \neq 0$. Решив систему, получим:

$$c_0 = \frac{x_1 f(x_0) - x_0 f(x_1)}{x_1 - x_0}; \quad c_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Следовательно,

$$P_1(x) = \frac{x_1 f(x_0) - x_0 f(x_1)}{x_1 - x_0} + \frac{f(x_1) - f(x_0)}{x_1 - x_0} x. \quad (9)$$

Перепишем этот полином в несколько другой форме, выделяя $f(x_0)$ и $f(x_1)$ в качестве множителей

$$P_1(x) = f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0}. \quad (10)$$

Геометрический образ интерполирующей функции $P_1(x)$ - прямая, проходящая на плоскости (x, y) через точки с координатами (x_0, y_0) и (x_1, y_1) . Уравнение этой прямой, наряду с (9) и (10), можно переписать в виде:

$$y = P_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0} (x - x_0). \quad (11)$$

Из данного примера видно, что всегда существуют различные эквивалентные между собой формы записи интерполяционного полинома, удобные в различных ситуациях.

1.3. Построение интерполяционного полинома в форме Лагранжа.

Интерполяционный полином первой степени (9) мы построили, решая напрямую систему двух уравнений с двумя неизвестными - коэффициентами c_0 и c_1 . Однако решить таким же образом систему (8) при произвольном n технически очень сложно. Проще сделать это с помощью специальных методов, учитывающих особенности рассматриваемой задачи. Один из таких методов, принадлежащих Лагранжу, мы и рассмотрим в этом разделе.

Представим искомым полином $P_n(x)$ в виде:

$$P_n(x) = \sum_{i=0}^n f(x_i) Q_{n,i}(x), \quad (12)$$

где $Q_{n,i}(x)$ полиномы степени n , «ориентированные» на точки x_i в том смысле, что

$$Q_{n,i}(x) = \begin{cases} 0, & x = x_j \quad \forall j \neq i, \\ 1, & x = x_i. \end{cases} \quad (13)$$

Такие полиномы легко построить:

$$Q_{n,i}(x) = \prod_{\substack{j=0 \\ j \neq i}}^{j=n} \frac{(x - x_j)}{(x_i - x_j)} \quad (14)$$

или в развернутом виде:

$$\begin{aligned} Q_{n,0}(x) &= \frac{(x - x_1)(x - x_2) \dots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_n)}, \\ Q_{n,i}(x) &= \frac{(x - x_0) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}, \\ Q_{n,n}(x) &= \frac{(x - x_0)(x - x_1) \dots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \dots (x_n - x_{n-1})}. \end{aligned} \quad (15)$$

Иногда нам будет удобно записывать $Q_{n,i}(x)$ в виде:

$$Q_{n,i}(x) = \frac{(x - x_0) \dots [i] \dots (x - x_n)}{(x_i - x_0) \dots [i] \dots (x_i - x_n)}.$$

Из выражения (12) и формул (13) очевидно, что построенный полином $P_n(x)$ действительно является интерполяционным полиномом для функции $y = f(x)$ на сетке с узлами x_0, x_1, \dots, x_n . Его принято называть интерполяционным полиномом в форме Лагранжа. Этим подчеркивается, что возможны и другие эквивалентные представления интерполяционного полинома $P_n(x)$. С одним из них мы познакомимся в следующем разделе.

В заключение отметим, что из трех различных представлений интерполяционного полинома первой степени (9)- (11) формула (10) дает его запись в форме Лагранжа.

Задача 2.

Написать интерполяционный полином второй степени для функции $y = \sin x$ по ее значениям в трех точках: $x_0 = 0$, $x_1 = \pi/6$, $x_2 = \pi/2$. Вычислить с помощью этого полинома приближенное значение синуса в точке $x = \pi/4$, сравнить полученный результат с точным значением синуса и подсчитать погрешность $R_2\left(\frac{\pi}{4}\right)$.

Воспользуемся для записи полинома формулой Лагранжа (12). В рассматриваемом случае $y_0 = \sin x_0 = 0$, так что в формуле останется только два слагаемых соответствующих точкам x_1 и x_2 . В результате получим:

$$P_2(x) = \frac{1}{2} \frac{x\left(x - \frac{\pi}{2}\right)}{\frac{\pi}{6}\left(-\frac{\pi}{3}\right)} + \frac{x\left(x - \frac{\pi}{6}\right)}{\frac{\pi}{2}\left(\frac{\pi}{3}\right)} = \frac{x}{\pi^2} \left(\frac{7\pi}{2} - 3x \right) \quad (16)$$

Перейдем к выполнению второй части задания. Вычислим с помощью интерполяционного полинома (16) приближенные значения синуса в точке $x = \pi/4$ и подсчитаем погрешность:

$$P_2\left(\frac{\pi}{4}\right) = \frac{11}{16}, \quad R_2\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} - \frac{11}{16} = 0.0197 < 0.02. \quad (17)$$

На рис. 1 приведены для сравнения графики функций $\sin x$ (сплошная линия) и $P_2(x)$ (пунктир).

1.4. Интерполяционный полином в форме Ньютона.

Интерполяционный полином в форме Лагранжа, несмотря на своё изящество, неудобен для вычислений тем, что при увеличении числа узлов интерполяции приходится перестраивать весь полином заново.

Перепишем интерполяционный полином Лагранжа в иной, эквивалентной форме

$$P_n(x) = P_0(x) + \sum_{i=1}^n (P_i(x) - P_{i-1}(x)), \quad (18)$$

где $P_i(x)$ - полиномы Лагранжа степени $i \leq n$, соответствующие узлам интерполирования x_0, x_1, \dots, x_i . В частности, $P_0(x) = f(x_0)$ - полином нулевой степени.

Полином

$$Q_i(x) = P_i(x) - P_{i-1}(x) \quad (19)$$

имеет степень i и по построению обращается в ноль при $x = x_0, x = x_1, \dots, x = x_{i-1}$, поэтому его можно представить в виде

$$Q_i(x) = A_i(x - x_0)(x - x_1) \dots (x - x_{i-1}), \quad (20)$$

где A_i - числовой коэффициент при x^i . Поскольку $P_{i-1}(x)$ не содержит степени i , то A_i просто совпадает с коэффициентом при x^i в полиноме $P_i(x)$. Согласно (12) и (15) его можно записать в виде

$$A_i = \sum_{k=0}^i \frac{f(x_k)}{\omega_{k,i}}, \quad (21)$$

где

$$\omega_{k,i} = (x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_i). \quad (22)$$

При этом

$$A_0 = f(x_0). \quad (23)$$

Формулы (19) и (21) позволяют написать рекуррентное соотношение для полинома $P_n(x)$:

$$P_n(x) = P_{n-1}(x) + A_n(x - x_0) \dots (x - x_{n-1}). \quad (24)$$

Выражая аналогичным образом по индукции $P_{n-1}(x)$ через $P_{n-2}(x)$, $P_{n-2}(x)$ через $P_{n-3}(x)$ и т. д., получим окончательную формулу для полинома $P_n(x)$:

$$P_n(x) = A_0 + A_1(x - x_0) + A_2(x - x_0)(x - x_1) + \dots + A_i(x - x_0) \dots (x - x_{i-1}) + \dots + A_n(x - x_0) \dots (x - x_{n-1}). \quad (25)$$

Представление (25) удобно для вычислителя, поскольку увеличение n на единицу требует только добавления к «старому» многочлену одного дополнительного слагаемого. Такое представление интерполяционного полинома $P_n(x)$ называют интерполяционным полиномом в форме Ньютона.

Из трех эквивалентных представлений интерполяционного полинома первой степени (9) - (11) формула (11) дает его запись в форме Ньютона.

Задача 3.

Написать интерполяционный полином второй степени в форме Ньютона для функции $y = \sin x$ по ее значениям в трех точках: $x_0 = 0$, $x_1 = \pi/6$, $x_2 = \pi/2$ (см. задачу 2).

Согласно формуле (25)

$$P_2(x) = A_0 + A_1x + A_2x \left(x - \frac{\pi}{6} \right). \quad (26)$$

Коэффициенты в этом разложении вычисляются по формулам (21) и (23):

$$A_0 = 0, \quad A_1 = \frac{1}{2} \left(\frac{6}{\pi} \right), \quad A_2 = -\frac{18}{2\pi^2} + \frac{6}{\pi^2} = -\frac{3}{\pi^2}. \quad (27)$$

Подставляя найденные значения коэффициентов в формулу (26), получим

$$P_2(x) = \frac{3}{\pi}x - \frac{3x}{\pi^2} \left(x - \frac{\pi}{6} \right) = \frac{x}{\pi^2} \left(\frac{7\pi}{2} - 3x \right). \quad (28)$$

Первоначальные выражения для интерполяционного полинома в форме Лагранжа и Ньютона различны, но окончательные ответы, естественно, совпадают.

1.5. Погрешность интерполирования.

Поставим вопрос о том, насколько хорошо интерполяционный полином $P_n(x)$ приближает функцию $f(x)$ на отрезке $[a, b]$, то есть попытаемся оценить погрешность (остаточный член)

$$R_n(x) = f(x) - P_n(x), \quad x \in [a, b]. \quad (29)$$

Сразу же отметим, что по определению интерполяционного полинома

$$R_n(x_i) = 0 \text{ при } i = 0, 1, \dots, n, \quad (30)$$

поэтому речь идет об оценке $R_n(x)$ при значениях $x \neq x_i$.

Для того, чтобы это сделать, следует ввести дополнительно предположение о гладкости функции $f(x)$. Предположим, что $f(x)$ имеет $(n+1)$ непрерывную производную на отрезке $[a, b]$.

В силу (30) $R_n(x)$ можно представить в виде:

$$R_n(x) = \omega_{n+1}(x)r_n(x), \quad (31)$$

где $\omega_{n+1}(x)$ - полином степени $(n+1)$:

$$\omega_{n+1}(x) = (x - x_0)(x - x_1)\dots(x - x_n). \quad (32)$$

Зафиксируем произвольное значение $x \in [a, b]$ и рассмотрим вспомогательную функцию от переменной t :

$$g(t) = f(t) - P_n(t) - \omega_{n+1}(t)r_n(x),$$

заданную на отрезке $[a, b]$ и содержащую переменную x в качестве параметра. В силу своего определения функция $g(t)$ обязана обращаться в нуль в узлах интерполирования при $t = x_i$ и кроме того при $t = x$, т. е. как функция аргумента t она имеет $(n+2)$ нуля:

$$g(x_i) = 0, \quad i = 0, 1, \dots, n, \quad g(x) = 0. \quad (33)$$

Если $x \in [x_0, x_n]$, то все ее нули также лежат на отрезке $[x_0, x_n]$. Если $x < x_0$, то эти нули, вообще говоря, принадлежат отрезку $[x, x_n]$, а если $x > x_n$, то они находятся на отрезке $[x_0, x]$. Объединяя эти три случая, скажем, что указанные нули функции $g(t)$ принадлежат отрезку $[\alpha, \beta]$, где $\alpha = \min(x_0, x) \geq a$, $\beta = \max(x_n, x) \leq b$.

Согласно известной теореме Ролля можно утверждать, что производная $g'(t)$ имеет по крайней мере $(n+1)$ нуль на отрезке $[\alpha, \beta]$ (эти нули перемежаются с нулями самой функции $g(t)$). Повторяя это рассуждение, заключаем, что $g''(t)$ имеет по крайней мере n нулей на отрезке $[\alpha, \beta]$, $g'''(t)$ - $(n-1)$ нуль и, наконец, $g^{(n+1)}(t)$ обращается хотя бы один раз в нуль в некоторой точке $t = \xi \in [\alpha, \beta]$, то есть

$$g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - P_n^{(n+1)}(\xi) - (n+1)!r_n(x) = 0.$$

Учитывая, что $(n+1)$ производная полинома степени n тождественно равна нулю, получаем, что

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}; \quad \xi \in [\alpha, \beta] \quad (34)$$

и соответственно

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x). \quad (35)$$

Формула (35) не позволяет вычислить погрешность, поскольку точное значение аргумента ξ нам неизвестно. Однако с ее помощью погрешность можно оценить:

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|, \quad (36)$$

где

$$M_{n+1} = \max_{x \in [\alpha, \beta]} |f^{(n+1)}(x)| \leq \max_{x \in [a, b]} |f^{(n+1)}(x)|. \quad (37)$$

Обсудим роль полинома $\omega_{n+1}(x)$ (32) в оценке (36). На отрезке $[x_0, x_n]$ он имеет $(n+1)$ нуль, а его значения между этими нулями сравнительно невелики, но, когда точка x выходит за пределы отрезка $[x_0, x_n]$ и удаляется от точки x_0 влево или от точки x_n вправо, оценка (36) ухудшается из-за быстрого роста функции $|\omega_{n+1}(x)|$. Это хорошо видно на рис. 2, где в качестве примера приведен график функции $\omega_4(x)$ с корнями $x_0 = -3/2$, $x_1 = -1/2$, $x_2 = 1/2$, $x_3 = 3/2$:

$$\omega_4(x) = \left(x^2 - \frac{9}{4}\right) \left(x^2 - \frac{1}{4}\right).$$

Ее наибольшее по модулю значение на отрезке $\left[-\frac{3}{2}, \frac{3}{2}\right]$ равно единице. Однако уже в точках $x = \pm 2$ за пределами отрезка полином $\omega_4(x)$ принимает значение

$$\omega_4(\pm 2) = \frac{105}{16} = 6.5625.$$

Из сказанного можно сделать следующий вывод. Если $x \in [x_0, x_n]$, то множитель $|\omega_{n+1}(x)|$ не обесценивает оценку (36). Такой случай называют собственно интерполяцией $f(x)$. Противоположный случай, когда точка x лежит вне отрезка называют экстраполяцией функции $f(x)$. Отмеченная выше особенность поведения полинома $\omega_{n+1}(x)$ резко ухудшает оценку (36) при экстраполяции. Поэтому на практике экстраполяции избегают или ограничиваются многочленами невысокой степени ($n = 1, 2$), когда рост функции $|\omega_{n+1}(x)|$ не настолько критичен.

Задача 4.

Написать мажорантную оценку для погрешности (36) при вычислении приближенного значения $\sin x$ в точке $x = \pi/4$ с помощью интерполяционного полинома второй степени $P_2(x)$ (16). Сравнить ее с погрешностью (17), подсчитанной непосредственно.

Формула для погрешности (35) принимает в данном случае вид:

$$R_2\left(\frac{\pi}{4}\right) = \frac{1}{6} (-\cos \xi) \omega_3\left(\frac{\pi}{4}\right) = \cos \xi \frac{\pi^3}{1152}, \quad 0 \leq \xi \leq \frac{\pi}{2}.$$

Она правильно определяет знак погрешности, но не позволяет вычислить ее величину, поскольку значение аргумента ξ неизвестно. Чтобы получить мажорантную оценку погрешности (36), нужно заменить $\cos \xi$ на его наибольшее значение – единицу. В результате будем иметь:

$$R_2\left(\frac{\pi}{4}\right) \leq \frac{\pi^3}{1152} < 0.027.$$

Эта оценка согласуется с величиной погрешности (17), вычисленной «в лоб».

1.6. О сходимости интерполяционного процесса.

Поставим вопрос, будут ли сходиться интерполяционные полиномы $P_n(x)$ к интерполируемой функции $f(x)$ на отрезке $[a, b]$ при неограниченном возрастании числа узлов n .

Упорядоченное множество точек $x_i, i = 0, 1, \dots, n$ (1) назовем сеткой на отрезке $[a, b]$ и обозначим для краткости Ω_n . Рассмотрим последовательность сеток с возрастающим числом узлов:

$$\Omega_0 = \{x_0^{(0)}\}, \quad \Omega_1 = \{x_0^{(1)}, x_1^{(1)}\}, \dots, \Omega_n = \{x_0^{(n)}, x_1^{(n)} \dots x_n^{(n)}\}, \dots$$

и отвечающую ей последовательность интерполяционных полиномов $P_n(x)$, построенных для фиксированной непрерывной на отрезке $[a, b]$ функции $f(x)$.

Интерполяционный процесс для функции сходится в точке $x_* \in [a, b]$, если существует предел

$$\lim_{n \rightarrow \infty} P_n(x_*) = f(x_*).$$

Наряду с обычной сходимостью часто рассматривается сходимость в различных нормах. Так, равномерная сходимость на отрезке $[a, b]$ означает, что

$$\max_{x \in [a, b]} |f(x) - P_n(x)| \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Сходимость или расхождение интерполяционного процесса зависит как от выбора последовательности сеток, так и от гладкости функции $f(x)$. Если $f(x)$ - целая аналитическая функция, то при произвольном расположении узлов на отрезке $[a, b]$ интерполяционный многочлен $P_n(x)$ равномерно сходится к $f(x)$ при $n \rightarrow \infty$.

Положение резко меняется, если производные функции разрывны или не существуют в отдельных точках. Например для функции $f(x) = |x|$ на отрезке $[-1, 1]$, покрытом равномерной сеткой узлов, значения $P_n(x)$ между узлами интерполяции неограниченно возрастают при $n \rightarrow \infty$. Вместе с тем, для заданной непрерывной функции $f(x)$ за счет выбора сеток можно добиться сходимости и притом равномерной на $[a, b]$. Однако построение таких сеток довольно сложно и, главное, такие сетки «индивидуальны» для каждой конкретной функции.

Если заметить дополнительно, что объем вычислений при построении интерполяционного полинома быстро нарастает с ростом n , то становится понятно, что на практике вычислители избегают пользоваться интерполяционными полиномами высокой степени. Вместо этого, в случае необходимости, при больших значениях n используется кусочно-полиномиальная интерполяция, которую мы обсудим в следующем параграфе.

1.7. Интерполяционный полином Эрмита.

Расширим постановку задачи об интерполяции. Ранее полагалось, что в узлах интерполяции заданы только значения функции $f(x)$. Пусть теперь в узлах $x_k \in [a, b]$, $k = 0, 1, \dots, m$, среди которых нет совпадающих, заданы значения функции $f(x_k)$, и её производных $f^{(i)}(x_k)$, $i = 1, 2, \dots, N_k - 1$ до $(N_k - 1)$ -го порядка включительно. Числа N_k при этом называют кратностью узла x_k . В каждой точке x_k , таким образом, задано N_k величин:

$$f(x_k), f'(x_k), \dots, f^{(N_k-1)}(x_k).$$

В общей сложности на всей совокупности узлов x_0, x_1, \dots, x_m известно $N_0 + N_1 + \dots + N_m$ величин, что дает возможность ставить вопрос о построении полинома $H_n(x)$ степени

$$n = N_0 + \dots + N_m - 1, \tag{38}$$

удовлетворяющего требованиям:

$$H_n^{(i)}(x_k) = f^{(i)}(x_k), \quad k = 0, 1, \dots, m, \quad i = 0, 1, \dots, N_k - 1. \tag{39}$$

Такой полином называется интерполяционным полиномом Эрмита для функции $f(x)$. Рассмотренный ранее вариант построения интерполяционного полинома $P_n(x)$ по известным значениям функции $f(x)$ в узлах интерполяции является частным случаем построения полинома Эрмита при условии, что все узлы простые: $N_k = 1$, $k = 0, 1, \dots, m$.

Докажем, что интерполяционный полином Эрмита существует и является единственным. Представим его в стандартном виде

$$H_n(x) = a_0 + a_1x + \dots + a_nx^n.$$

Наше утверждение будет справедливо, если показать, что коэффициенты a_0, a_1, \dots, a_n определяются из условий (39) и притом единственным образом. Условия (39) представляют собой систему линейных алгебраических уравнений относительно этих коэффициентов, причем число уравнений и число неизвестных равны $N_0 + N_1 + \dots + N_m = n + 1$.

Рассмотрим соответствующую однородную систему

$$\bar{H}_n^{(i)}(x_k) = 0, \quad k = 0, 1, \dots, m, \quad i = 0, 1, \dots, N_k - 1. \tag{40}$$

Уравнения (40) просто указывают на то, что числа x_k являются корнями полинома $\bar{H}_n(x)$ кратности N_k . Мы видим, таким образом, что полином $\bar{H}_n(x)$ имеет, с учетом кратности, не менее $N_0 + N_1 + \dots + N_m = n + 1$ корней. Поскольку его степень равна n , то он должен тождественно равняться нулю. Это означает, что $\bar{a}_0 = \bar{a}_1 = \dots = \bar{a}_n = 0$, т.е. однородная система уравнений (40) имеет только тривиальное решение. Отсюда следует, что неоднородная система (39) при любой правой части разрешима и при том единственным образом.

Исследование погрешности интерполирования полинома Эрмита $R_n(x) = f(x) - H_n(x)$ почти дословно повторяет проведенное ранее исследование для

полинома с простыми узлами x_k , в которых заданы только $f(x_k)$. Достаточно представить $R_n(x)$ в виде

$$R_n(x) = r_n(x)\omega_{n+1}(x), \quad (41)$$

где

$$\omega_{n+1}(x) = (x-x_0)^{N_0} (x-x_1)^{N_1} \dots (x-x_m)^{N_m}, \quad n+1 = N_0 + \dots + N_m \quad (42)$$

и рассмотреть функцию

$$g(t) = f(t) - H_n(t) - r_n(x)\omega_{n+1}(t).$$

Применяя теорему Ролля к функции $g(t)$ и ее производным с учетом кратности корней в узлах $t = x_k$ и условия $g(x) = 0$ придем к формуле

$$f(x) - H_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \quad (43)$$

которая по существу повторяет формулу (35). С ее помощью можно написать оценку типа (36):

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|, \quad (44)$$

где M_{n+1} - максимальное значение модуля функции $f^{(n+1)}(x)$ (37). Здесь полином $\omega_{n+1}(x)$ (42) является обобщением полинома (32) на случай кратных корней.

Построение полинома Эрмита в общем случае при произвольном числе узлов и их кратности приводит к довольно громоздким выражениям и редко используется. Поэтому мы ограничимся двумя примерами, встречающимися на практике.

Пример 1

Построить интерполяционный полином Эрмита для функции $f(x)$ по известным значениям в узлах $f(x_k) = f_k$, $k = 0, 1, \dots, m$ и значению $f'(x_j) = f'_j$ в одном из узлов $x = x_j$.

Степень полинома $H_n(x)$ в данном случае равна $m+1$.

Будем искать $H_{m+1}(x)$ в виде

$$H_{m+1}(x) = \sum_{\substack{k=0 \\ k \neq j}}^m f_k \frac{(x-x_0) \dots [k] \dots (x-x_m)}{(x_k-x_0) \dots [k] \dots (x_k-x_m)} \left(\frac{x-x_j}{x_k-x_j} \right) + \\ + \left[f_j + \alpha_j (x-x_j) \right] \frac{(x-x_0) \dots [j] \dots (x-x_m)}{(x_j-x_0) \dots [j] \dots (x_j-x_m)}.$$

Здесь выражения, стоящие под знаком суммы, суть обычные составляющие полинома в форме Лагранжа в узлах x_k , $k \neq j$, «усиленные» дополнительными множителями $(x-x_j)/(x_k-x_j)$. Слагаемое, отвечающее кратному узлу $x = x_j$, выделено отдельно как особое. Постоянная α_j подлежит определению.

Из структуры $H_{m+1}(x)$ видно, что $H_{m+1}(x_i) = f_i$, $i = 0, 1, \dots, m$. Найдем производную $H'_{m+1}(x_j)$ в узле $x = x_j$. Слагаемые, стоящие под знаком суммы, содержат множители $(x-x_j)^2$ и потому их производные обращаются в нуль при $x = x_j$. Таким образом,

$$H'_{m+1}(x_j) = f_j \left(\frac{1}{x_j - x_0} + \dots + \frac{1}{x_j - x_{j-1}} + \frac{1}{x_j - x_{j+1}} + \dots + \frac{1}{x_j - x_m} \right) + \alpha_j. \quad (45)$$

Для соблюдения требования $H'_{m+1}(x_j) = f'_j$ следует положить

$$\alpha_j = f'_j - f_j A_j,$$

где для краткости обозначено

$$A_j = \frac{1}{x_j - x_0} + \dots + \frac{1}{x_j - x_{j-1}} + \frac{1}{x_j - x_{j+1}} + \dots + \frac{1}{x_j - x_m}. \quad (46)$$

Итак:

$$H_{m+1}(x) = \sum_{\substack{k=0 \\ k \neq j}}^m f_k \frac{(x - x_0) \dots [k] \dots (x - x_m)}{(x_k - x_0) \dots [k] \dots (x_k - x_m)} \left(\frac{x - x_j}{x_k - x_j} \right) + \\ + \left[f_j + (f'_j - f_j A_j)(x - x_j) \right] \frac{(x - x_0) \dots [j] \dots (x - x_m)}{(x_j - x_0) \dots [j] \dots (x_j - x_m)}. \quad (47)$$

Пример 2.

Построить интерполяционный полином Эрмита для функции $f(x)$ в случае, когда во всех узлах интерполяции x_k , $k = 0, 1, \dots, m$ заданы значения функции $f(x_k) = f_k$ и ее первой производной $f'(x_k) = f'_k$.

В данном случае $N_k = 2$, $k = 0, 1, \dots, m$, так что степень полинома $H_n(x)$ равна $2m + 1$.

Запишем исходный полином в виде:

$$H_{2m+1}(x) = \sum_{k=0}^m \left[f_k + \alpha_k (x - x_k) \right] \frac{(x - x_0)^2 \dots [k] \dots (x - x_m)^2}{(x_k - x_0)^2 \dots [k] \dots (x_k - x_m)^2}. \quad (48)$$

Представление (48) удобно тем, что автоматически выполняются условия

$$H_{2m+1}(x_k) = f_k.$$

При вычислении производной полинома (48) в узле $x = x_k$ следует учесть, что все слагаемые суммы, кроме слагаемого, отвечающему самому узлу x_k , дают нулевой вклад в производную в этой точке, поэтому

$$H'_{2m+1}(x_k) = f_k \left(\frac{2}{x_k - x_0} + \dots + \frac{2}{x_k - x_{k-1}} + \frac{2}{x_k - x_{k+1}} + \dots + \frac{2}{x_k - x_m} \right) + \alpha_k = f'_k.$$

Отсюда

$$\alpha_k = f'_k - 2f_k A_k,$$

где, числа A_k определяются формулой (46). Таким образом, решением данной задачи является полиномом Эрмита

$$H_{2m+1}(x) = \sum_{k=0}^m \left[f_k + (f'_k - 2f_k A_k)(x - x_k) \right] \frac{(x - x_0)^2 \dots [k] \dots (x - x_m)^2}{(x_k - x_0)^2 \dots [k] \dots (x_k - x_m)^2}. \quad (49)$$

Задача 5

Построить полином Эрмита второй степени $H_2(x)$ для функции $\sin x$ по следующим данным:

$$\sin 0 = 0, \sin \frac{\pi}{2} = 1, \sin' \frac{\pi}{2} = 0.$$

Вычислить с помощью этого полинома приближенное значение синуса в точке $x = \pi/4$. Найти погрешность, сравнить ее с погрешностью, которую дает интерполяционный полином $P_2(x)$ (16) задачи 2 и с теоретической оценкой погрешности (44).

Здесь мы имеем задачу, которая в общем виде была разобрана в примере 1: согласно (49) узел $x_0 = 0$ является простым, а узел $x_1 = \pi/2$ - двукратным. В этом случае в формуле (47) сумма, соответствующая простым узлам, сводится к одному слагаемому, которое в силу нулевого значения синуса в точке $x_0 = 0$ обращается в ноль. Второй член в формуле (47) соответствует кратному корню $x_1 = \pi/2$. Подставляя сюда соответствующее значение синуса и его производной в этой точке, а также значение коэффициента $A_1 = 2/\pi$, будем иметь:

$$H_2(x) = \left[1 - \frac{2}{\pi} \left(x - \frac{\pi}{2} \right) \right] \frac{2x}{\pi} = \frac{4}{\pi^2} x(\pi - x). \quad (50)$$

Вычислим значение полинома $H_2(x)$ в точке $x = \pi/4$ и подсчитаем погрешность

$$H_2\left(\frac{\pi}{4}\right) = \frac{3}{4}, R_2\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} - \frac{3}{4} = -0.04282. \quad (51)$$

Теоретическая формула для погрешности (43) принимает в данном случае вид:

$$R_2\left(\frac{\pi}{4}\right) = \frac{1}{6} (-\cos \xi) \omega_3\left(\frac{\pi}{4}\right) = -\cos \xi \frac{\pi^3}{384}, \quad 0 \leq \xi \leq \frac{\pi}{2}. \quad (52)$$

Она правильно определяет знак погрешности и позволяет написать для нее мажорантную оценку

$$\left| R_2\left(\frac{\pi}{4}\right) \right| \leq \frac{\pi^3}{384} < 0.081. \quad (53)$$

Данная оценка согласуется с величиной погрешности (51), подсчитанной «в лоб».

При подсчете приближенного значения $\sin x$ с помощью полинома Эрмита $H_2(x)$ (50) в точке $x = \pi/4$ мы получили погрешность (51), модуль которой в два с лишним раза превышает погрешность (17) полинома $P_2(x)$ (16). Чтобы понять причину такого расхождения, рассмотрим рис. 3, на котором приведены графики функций $\sin x$ (сплошная линия) и $H_2(x)$ (пунктир). Сравним его с рис. 1, на котором изображены графики функции $\sin x$ и полинома $P_2(x)$. Из-за нулевого значения производной $H_2'(x)$ в точке $x = \pi/2$ график полинома $H_2(x)$ качественно больше похож на график синуса, чем график полинома $P_2(x)$. Однако равенство полинома

$P_2(x)$ синусу не только в граничных точках отрезка $\left[0, \frac{\pi}{2}\right]$, но и во внутренней точке

$x = \pi/6$ приводит к тому, что полином $P_2(x)$ приближает синус на отрезке $\left[0, \frac{\pi}{2}\right]$

лучше чем полином $H_2(x)$. Это хорошо видно при сравнении рис. 1 и рис. 3. Подсчет погрешностей (17) и (51) в точке $x = \pi/4$ является дополнительным тому подтверждением.

§2. Интерполирование сплайнами.

Увеличение степени интерполяционного полинома может оказаться невыгодным из-за быстрого роста объема вычислений. К тому же далеко не всегда оно приводит к повышению точности. Во второй половине XX века с появлением компьютеров и развитием современной вычислительной математики при обработке больших таблиц получила развитие новая идея – строить приближение функций с помощью кусочно-полиномиальной интерполяции с использованием полиномов сравнительно невысоких степеней. Наиболее удобными оказались полиномы третьей степени. Такие конструкции получили название кубических сплайнов.

2.1. Определение кубического сплайна.

Пусть на отрезке $[a, b]$ задана функция $y = f(x)$. Рассмотрим сетку узлов

$$a = x_0 < x_1 < x_2 < \dots < x_n = b \quad (54)$$

и обозначим через h_i расстояние между смежными узлами

$$h_i = x_i - x_{i-1}, \quad i = 1, \dots, n \quad (55)$$

Определение:

Назовем кубическим сплайном функции $y = f(x)$, $x \in [a, b]$ на сетке (54) функцию $S(x)$ удовлетворяющую условиям:

S1. На каждом отрезке $[x_{i-1}, x_i]$ функция $S(x)$ является полиномом третьей степени.

S2. Функция $S(x)$, её первая $S'(x)$ и вторая $S''(x)$ производные непрерывны на сегменте $[a, b]$.

S3. $S(x_i) = f(x_i) = f_i, \quad i = 0, \dots, n$

S4. На концах сегмента $[a, b]$ функция $S''(x)$ удовлетворяет условиям $S''(a) = S''(b) = 0$.

Замечание. На концах сегмента $[a, b]$ могут быть заданы в принципе и другие условия, например:

$$S''(a) = A, \quad S''(b) = B.$$

Справедлива следующая теорема.

Теорема.

Существует единственный сплайн $S(x)$, удовлетворяющий требованиям (S1) – (S4).

Мы проведем конструктивное доказательство этой теоремы.

2.2. Формулировка системы уравнений для коэффициентов кубического сплайна.

Сведем задачу построения сплайна к отысканию коэффициентов упомянутых полиномов третьей степени на каждом из отрезков $[x_{i-1}, x_i]$. Для этого сопоставим отрезку $[x_{i-1}, x_i]$ полином $S_i(x)$, для удобства записанный в виде:

$$S_i(x) = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{6}(x - x_i)^3, \quad x \in [x_{i-1}, x_i], \quad i = 1, \dots, n. \quad (56)$$

При этом, очевидно:

$$S_i'(x) = b_i + c_i(x - x_i) + \frac{d_i}{2}(x - x_i)^2, \quad (57)$$

$$S_i''(x) = c_i + d_i(x - x_i), \quad (58)$$

так, что

$$S_i(x_i) = a_i, \quad S_i'(x_i) = b_i, \quad S_i''(x_i) = c_i. \quad (59)$$

Для выполнения требований (S3) в узлах интерполяции с номерами $i = 1, \dots, n$ следует положить:

$$a_i = f(x_i) = f_i, \quad i = 1, \dots, n \quad (60)$$

Требую непрерывности сплайна в узлах x_i ($i = 1, \dots, n-1$) и выполнения условия (S3) при $i = 0$, получим:

$$S_i(x_{i-1}) = f_{i-1}, \quad i = 1, \dots, n \quad (61)$$

или

$$f_i + b_i(x_{i-1} - x_i) + \frac{c_i}{2}(x_{i-1} - x_i)^2 + \frac{d_i}{6}(x_{i-1} - x_i)^3 = f_{i-1}, \quad i = 1, \dots, n.$$

Это равенство можно переписать следующим образом:

$$b_i h_i - \frac{c_i}{2} h_i^2 + \frac{d_i}{6} h_i^3 = f_i - f_{i-1}, \quad i = 1, \dots, n. \quad (62)$$

Условие (S2) непрерывности первой производной $S'(x)$ в узлах x_i ($i = 1, \dots, n-1$) принимает вид:

$$S_i'(x_{i-1}) = S_{i-1}'(x_{i-1}) = b_{i-1}, \quad i = 2, \dots, n \quad (63)$$

и приводит к соотношениям

$$b_i - c_i h_i + \frac{d_i}{2} h_i^2 = b_{i-1}, \quad i = 2, \dots, n$$

или

$$c_i h_i - \frac{d_i}{2} h_i^2 = b_i - b_{i-1}, \quad i = 2, \dots, n. \quad (64)$$

Аналогичным образом условия непрерывности второй производной $S''(x)$ в тех же узлах:

$$S_i''(x_{i-1}) = S_{i-1}''(x_{i-1}) = c_{i-1}, \quad i = 2, \dots, n \quad (65)$$

означают, что

$$d_i h_i = c_i - c_{i-1}, \quad i = 2, \dots, n. \quad (66)$$

Наконец, дополнительные граничные условия (S4) дают еще два уравнения

$$\begin{cases} S_1''(x_0) = S_1''(a) = c_1 - d_1 h_1 = 0 \\ S_n''(x_n) = S_n''(b) = c_n = 0 \end{cases} \quad (67)$$

В итоге мы получили замкнутую систему (62), (64), (66), (67), содержащую в сумме $3n$ линейных уравнений для отыскания $3n$ неизвестных: $b_i, c_i, d_i, i = 1, 2, \dots, n$

2.3. Редукция системы.

Удобно формально ввести ещё одно неизвестное c_0 , положив при этом $c_0 = 0$, и первое уравнение в (67) переписать в виде:

$$d_1 h_1 = c_1 - c_0,$$

то есть в форме аналогичной (66).

Теперь уравнения (66) и (67) естественно представить в единообразном виде

$$d_i h_i = c_i - c_{i-1}, \quad i = 1, 2, \dots, n \quad (68)$$

$$c_0 = 0, \quad c_n = 0. \quad (69)$$

Обратим внимание на то, что из системы (68) можно выразить все коэффициенты d_i через разности $c_i - c_{i-1}$, а затем из системы (62) выразить через c_i и c_{i-1} коэффициенты b_i . Подставляя полученные выражения в (64), приходим к системе линейных уравнений для c_i :

$$\frac{1}{3} c_{i-2} h_{i-1} + \frac{2}{3} c_{i-1} (h_{i-1} + h_i) + \frac{1}{3} c_i h_i = 2 \left(\frac{f_i - f_{i-1}}{h_i} - \frac{f_{i-1} - f_{i-2}}{h_{i-1}} \right), \quad i = 2, 3, \dots, n. \quad (70)$$

Сдвигая индекс i на единицу, получим симметричную форму записи уравнений (70):

$$h_i c_{i-1} + 2(h_i + h_{i+1}) c_i + h_{i+1} c_{i+1} = 6 \left(\frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i} \right), \quad i = 1, \dots, n-1. \quad (71)$$

Кроме того, согласно (69)

$$c_0 = c_n = 0. \quad (72)$$

Система (71) содержит $n-1$ уравнение с $(n-1)$ -ой неизвестной: c_1, c_2, \dots, c_{n-1} . Величины c_0 и c_n определены дополнительными соотношениями (72). Если сетка (54) равномерная, т. е. $h_i = h = const$, то уравнения (71) принимают особенно простой вид:

$$c_{i-1} + 4c_i + c_{i+1} = 6 \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2}. \quad (73)$$

Для уравнений системы (71) выполнено условие диагонального преобладания. Отсюда следует существование и единственность решения задачи (71), (72). По найденным величинам c_i можно рассчитать остальные коэффициенты сплайна по формулам

$$d_i = \frac{c_i - c_{i-1}}{h_i}, \quad i = 1, \dots, n \quad (74)$$

и

$$b_i = \frac{1}{2} h_i c_i - \frac{1}{6} h_i^2 d_i + \frac{f_i - f_{i-1}}{h}, \quad i = 1, \dots, n, \quad (75)$$

завершив тем самым построение сплайна. Теорема доказана.

2.4. Замечание о решении системы.

Уравнения (71) имеют так называемую трехточечную структуру, общий вид таких систем

$$A_i y_{i-1} + C_i y_i + B_i y_{i+1} = F_i, \quad i = 1, 2, \dots, n-1, \quad (76)$$

$$y_0 = 0, \quad y_n = 0 \quad (77)$$

соответствует системе линейных уравнений с трехдиагональной матрицей T для определения вектора неизвестных $y = (y_1, y_2, \dots, y_{n-1})$:

$$Ty = F,$$

где

$$T = \begin{vmatrix} C_1 & B_1 & 0 & 0 & 0 & 0 \\ A_2 & C_2 & B_2 & 0 & 0 & 0 \\ 0 & A_3 & C_3 & B_3 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & A_{n-1} & C_{n-1} \end{vmatrix}, \quad F = \begin{vmatrix} F_1 \\ F_2 \\ F_3 \\ \dots \\ \dots \\ F_{n-1} \end{vmatrix}. \quad (78)$$

При этом легко видеть, что в нашем случае

$$|C_i| > |A_i| + |B_i|, \quad i = 1, \dots, n-1, \quad (79)$$

поскольку

$$C_i = 2(h_i + h_{i+1}), \quad A_i = h_i, \quad B_i = h_{i+1}. \quad (80)$$

Как было показано в главе 1, решение подобных систем эффективно осуществляется методом прогонки.

Задача 6.

Рассмотреть функцию $y = f(x) = 3^x$ на отрезке $[-1, 1]$ с узлами интерполяции $x_0 = -1, x_1 = 0, x_2 = 1$. Построить кубический сплайн. Найти его значение при $x = 1/2$, т. е. вычислить приближенно $\sqrt{3}$. Подсчитать погрешность.

В рассматриваемом случае мы имеем равномерную сетку с шагом $h = 1$. У нее одна внутренняя точка x_1 и две граничные - x_0 и x_2 . Система (73) сводится к одному уравнению относительно коэффициента c_1 , которое с учетом дополнительных соотношений (70), определяющих нулевые значения коэффициентов c_0 и c_2 , принимает вид:

$$4c_1 = 6 \left(\frac{1}{3} - 2 + 3 \right). \quad (81)$$

Таким образом, в нашей задаче:

$$c_0 = 0, \quad c_1 = 2, \quad c_2 = 0.$$

Остальные коэффициенты сплайна находятся по формулам (60), (74), (75):

$$a_1 = 1, a_2 = 3; d_1 = 2, d_2 = -2; b_1 = 4/3, b_2 = 7/3.$$

Теперь можно выписать кубические полиномы, определяющие сплайн:

$$S(x) = \begin{cases} S_1(x) = 1 + \frac{4}{3}x + x^2 + \frac{1}{3}x^3, & -1 \leq x \leq 0, \\ S_2(x) = 3 + \frac{7}{3}(x-1) - \frac{1}{3}(x-1)^3, & 0 \leq x \leq 1. \end{cases} \quad (82)$$

Легко проверить, что построенная таким образом функция $S(x)$ непрерывна вместе с первой и второй производной во внутренней узловой точке $x = 0$.

В заключение вычислим значение сплайна в точке $x = 1/2$, т. е. подсчитаем приближенно $\sqrt{3}$:

$$\sqrt{3} \approx S_2\left(\frac{1}{2}\right) = \frac{15}{8}, \quad \varepsilon = \sqrt{3} - \frac{15}{8} = -0,142949. \quad (83)$$

Значительная погрешность обусловлена прежде всего большим шагом $h = 1$. Определенную роль играют также условия S4:

$$S''(-1) = S''(1) = 0. \quad (84)$$

Вторая производная рассматриваемой функции $f(x) = 3^x$ в точках $x = \pm 1$ в ноль не обращается, т. е. условие (84) дает о ней искаженную информацию. Если учесть при построении сплайна истинные значения функции $f''(x)$ в точках ± 1 , то точность аппроксимации улучшится.

2.5. Сходимость и точность интерполирования сплайнами.

При обсуждении эффективности численного метода в первую очередь обращают внимание на две характеристики:

1. Условие сходимости метода (сходимость).

Речь идет о минимальных по возможности ограничениях, при которых приближенное решение задачи стремится к точному решению задачи.

Сходимость означает, что данный метод в принципе позволяет найти решение задачи с любой степенью точности.

2. Скорость сходимости (точность).

Это характеристика близости приближенного решения к точному (характеристика скорости убывания погрешности) при некоторых дополнительных ограничениях.

Посмотрим как решаются эти вопросы в теории сплайнов.

Итак, на сегменте $[a, b]$ задана функция $f(x)$ и построена сетка

$$a = x_0 < x_1 < x_2 < \dots < x_n = b; \quad h_i = x_i - x_{i-1} > 0.$$

Введем в рассмотрение величину

$$h = \max_{1 \leq i \leq n} h_i. \quad (85)$$

Приведем без доказательства две теоремы.

Теорема 1. Пусть $f(x)$ непрерывна на сегменте $[a, b]$, тогда для любого $\varepsilon > 0$ можно указать $\delta(\varepsilon) > 0$ такое, что при любой сетке, удовлетворяющей условию $h < \delta$ справедливо неравенство

$$|f(x) - S(x)| < \varepsilon \quad \forall x \in [a, b], \quad (86)$$

ными словами $S_h(x)$ при $h \rightarrow 0$ равномерно сходится к непрерывной функции $f(x)$.

Теорема 2. Пусть $f(x)$ имеет на сегменте $[a, b]$ четыре непрерывных производных и дополнительно удовлетворяет условию $f''(a) = f''(b) = 0$. Тогда имеют место неравенства (оценки):

$$|f(x) - S(x)| \leq M_4 h^4 \quad \forall x \in [a, b], \quad (87)$$

$$|f'(x) - S'(x)| \leq M_4 h^3 \quad \forall x \in [a, b], \quad (88)$$

$$|f''(x) - S''(x)| \leq M_4 h^2 \quad \forall x \in [a, b], \quad (89)$$

$$M_4 = \max_{[a,b]} |f^{(4)}(x)|. \quad (90)$$

§3. Метод наименьших квадратов.

Метод наименьших квадратов был предложен Гауссом и Лежандром в конце XVIII - начале XIX веков в связи с проблемой обработки экспериментальных данных. В этом случае задача построения функции непрерывного аргумента по дискретной информации (1), (2) характеризуется двумя особенностями:

1. Число точек x_i , в которых проводятся измерения, обычно бывает достаточно большим.

2. Значения функции y_i (2) в точках сетки x_i (1) определяются приближенно в связи с неизбежными ошибками измерения.

С учетом этих обстоятельств строить функцию $y(x)$ в виде суммы большого числа слагаемых (3) и добиваться ее точного равенства в точках сетки величинам y_i , как это делалось при интерполировании, становится нецелесообразным.

В методе наименьших квадратов аппроксимирующая функция $y(x)$ ищется в виде суммы, аналогичной (3), но содержащей сравнительно небольшое число слагаемых

$$F(x) = \sum_{k=0}^m a_k \varphi_k(x), \quad m < n, \quad (91)$$

в частности, возможен вариант $m \ll n$.

Предположим, что мы каким-то образом выбрали коэффициенты a_k , тогда в каждой точке сетки x_i , можно подсчитать погрешность

$$\delta_i = y_i - F(x_i) = y_i - \sum_{k=0}^m a_k \varphi_k(x_i), \quad i = 0, 1, 2, \dots, n. \quad (92)$$

Сумма квадратов этих величин называется суммарной квадратичной погрешностью

$$J = \sum_{i=0}^n \delta_i^2 = \sum_{i=0}^n \left(y_i - \sum_{k=0}^m a_k \varphi_k(x_i) \right)^2. \quad (93)$$

Она дает количественную оценку того, насколько близки значения функции $F(x)$ (91) в точках сетки к величинам y_i .

Меняя значения коэффициентов a_k , мы будем менять погрешность J , которая является их функцией. В результате естественно возникает задача:

Найти такой, набор коэффициентов a_k , при которых суммарная квадратичная погрешность J оказывается минимальной.

Функцию $F(x)$ (91) с набором коэффициентов, удовлетворяющих этому требованию, называют наилучшим приближением по методу наименьших квадратов.

Построение наилучшего приближения сводится к классической задаче математического анализа об экстремуме функции нескольких переменных. Метод решения этой задачи известен. Необходимым условием экстремума является равенство нулю в экстремальной точке всех первых частных производных рассматриваемой функции. В случае (93) это дает

$$\frac{\partial J}{\partial a_l} = -2 \sum_{i=0}^n (y_i - \sum_{k=0}^m a_k \varphi_k(x_i)) \varphi_l(x_i) = 0 \quad l = 0, 1, \dots, m. \quad (94)$$

Оставим члены, содержащие a_k , слева и поменяем в них порядок суммирования по индексам i и k . Члены, содержащие y_i , перенесем направо. В результате уравнения (94) примут вид

$$\sum_{k=0}^m \gamma_{lk} a_k = b_l, \quad l = 0, 1, \dots, m, \quad (95)$$

где

$$\gamma_{lk} = \sum_{i=0}^n \varphi_l(x_i) \varphi_k(x_i), \quad (96)$$

$$b_l = \sum_{i=0}^n \varphi_l(x_i) y_i. \quad (97)$$

Мы получили систему линейных алгебраических уравнений (95), в которой роль неизвестных играют искомые коэффициенты разложения a_0, a_1, \dots, a_m . Число уравнений и число неизвестных в этой системе совпадает и равно $m+1$. Матрица коэффициентов системы Γ состоит из элементов γ_{lk} , которые определяются формулой (96). Ее называют матрицей Грама для системы функций $\varphi_0(x), \varphi_1(x), \dots, \varphi_m(x)$ на сетке (1). Отметим, что матрица Грама является симметричной: для ее элементов, согласно (96), справедливо равенство $\gamma_{lk} = \gamma_{kl}$. Числа b_l , стоящие в правой части уравнений (95), вычисляются по формуле (97) через значения y_i сеточной функции (2).

Предположим, что функции $\varphi_0(x), \varphi_1(x), \dots, \varphi_m(x)$ выбраны такими, что определитель матрицы Грама, отличен от нуля:

$$\Delta = \det \Gamma \neq 0. \quad (98)$$

В этом случае при любой правой части система (95) имеет единственное решение

$$\bar{a}_0, \bar{a}_1, \dots, \bar{a}_m. \quad (99)$$

Рассмотрим наряду с набором коэффициентов (99), полученных в результате решения системы (95), любой другой набор коэффициентов a_0, a_1, \dots, a_m . Представим числа a_k в виде

$$\begin{aligned} a_0 &= \bar{a}_0 + \Delta a_0, a_1 = \bar{a}_1 + \Delta a_1, \dots, a_m = \bar{a}_m + \Delta a_m, \\ (\Delta a_0)^2 + (\Delta a_1)^2 + \dots + (\Delta a_m)^2 &> 0 \end{aligned} \quad (100)$$

и сравним значения суммарной квадратичной погрешности J для функций $F(x)$ (91), построенных с помощью коэффициентов (99) и (100).

Квадрат погрешности в точке $x = x_i$ для функции $F(x)$ (91) с коэффициентами (100) можно записать в виде

$$\begin{aligned} \delta_i^2 &= \left\{ y_i - \sum_{k=0}^m (\bar{a}_k + \Delta a_k) \varphi_k(x_i) \right\}^2 = \\ &= \left\{ \left(y_i - \sum_{k=0}^m \bar{a}_k \varphi_k(x_i) \right) - \sum_{k=0}^m \Delta a_k \varphi_k(x_i) \right\}^2 = \left(y_i - \sum_{k=0}^m \bar{a}_k \varphi_k(x_i) \right)^2 - \\ &\quad - 2 \left(y_i - \sum_{k=0}^m \bar{a}_k \varphi_k(x_i) \right) \left(\sum_{l=0}^m \Delta a_l \varphi_l(x_i) \right) + \left(\sum_{k=0}^m \Delta a_k \varphi_k(x_i) \right)^2. \end{aligned} \quad (101)$$

Здесь в среднем слагаемом мы заменили в одной из сумм индекс суммирования k на l , чтобы не использовать один и тот же индекс в двух разных суммах и иметь возможность перемножить их почленно.

Чтобы получить суммарную квадратичную погрешность, нужно просуммировать выражения (101) для δ_i^2 по индексу i . Первые слагаемые не содержат Δa_k . Их сумма дает погрешность J , вычисленную для функции (91) с коэффициентами (99) \bar{a}_k .

Рассмотрим теперь сумму вторых слагаемых, которые зависят от Δa_l линейно:

$$\begin{aligned} &-2 \sum_{i=0}^n \left\{ \left(y_i - \sum_{k=0}^m \bar{a}_k \varphi_k(x_i) \right) \cdot \left(\sum_{l=0}^m \Delta a_l \varphi_l(x_i) \right) \right\} = \\ &= -2 \sum_{l=0}^m \Delta a_l \left\{ \sum_{i=0}^n y_i \varphi_l(x_i) - \sum_{k=0}^m \bar{a}_k \sum_{i=0}^n \varphi_k(x_i) \varphi_l(x_i) \right\} = \\ &= -2 \sum_{l=0}^m \Delta a_l \left\{ b_l - \sum_{k=0}^m \bar{a}_k \gamma_{lk} \right\} = 0. \end{aligned} \quad (102)$$

Здесь мы поменяли местами порядок суммирования и воспользовались тем, что коэффициенты \bar{a}_k , удовлетворяют системе уравнений (95).

С учетом (102) будем иметь

$$\begin{aligned} J(\bar{a}_0 + \Delta a_0, \bar{a}_1 + \Delta a_1, \dots, \bar{a}_m + \Delta a_m) &= \\ &= J(\bar{a}_0, \bar{a}_1, \dots, \bar{a}_m) + \sum_{i=0}^n \left(\sum_{k=0}^m \Delta a_k \varphi_k(x_i) \right)^2 > J(\bar{a}_0, \bar{a}_1, \dots, \bar{a}_m). \end{aligned} \quad (103)$$

Формула (103) показывает, что функция $F(x)$ (91) с коэффициентами \bar{a}_k (100), полученными в результате решения уравнений (95), действительно минимизирует суммарную квадратичную погрешность J . Если мы возьмем любой другой набор коэффициентов (100), отличный от (99), то согласно формуле (103) к погрешности $J(\bar{a}_0, \bar{a}_1, \dots, \bar{a}_m)$ добавится положительное слагаемое и она увеличится.

Итак, чтобы построить наилучшее приближение (91) сеточной функции (1), (2) по методу наименьших квадратов, нужно взять в качестве коэффициентов разложения a_k решение системы линейных уравнений (95).

Задача 7

Сеточная функция задана таблицей 1

Таблица 1:

i	x_i	y_i
0	0,0	0,95
1	0,5	1,54
2	1,0	2,04
3	1,5	2,46
4	2,0	2,95

Построить линейную функцию

$$F(x) = a_0 + a_1x, \quad (104)$$

которая дает для нее наилучшее приближение по методу наименьших квадратов.

В рассматриваемом случае имеем:

$$n = 4, \quad m = 1, \quad \varphi_0(x) = 1, \quad \varphi_1(x) = x.$$

Для определения коэффициентов a_0 и a_1 составим систему уравнений (95). Элементы γ_{lk} ($l = 0, 1, \quad k = 0, 1$) матрицы Грама вычисляются по формуле (96)

$$\gamma_{0,0} = \sum_{i=0}^4 \varphi_0(x_i) \varphi_0(x_i) = 5,$$

$$\gamma_{0,1} = \gamma_{1,0} = \sum_{i=0}^4 \varphi_0(x_i) \varphi_1(x_i) = 5,$$

$$\gamma_{1,1} = \sum_{i=0}^4 \varphi_1(x_i) \varphi_1(x_i) = 7.5.$$

Числа b_0 и b_1 , стоящие в правой части уравнений (95), находим по формуле (97)

$$b_0 = \sum_{i=0}^4 \varphi_0(x_i) y_i = 9,94,$$

$$b_1 = \sum_{i=0}^4 \varphi_1(x_i) y_i = 12,40.$$

В результате система (95) принимает в рассматриваемом случае вид

$$\begin{aligned} 5a_0 + 5a_1 &= 9.94 \\ 5a_0 + 7,5a_1 &= 12.40. \end{aligned} \quad (105)$$

Определитель системы (105) $\Delta = 12,5 \neq 0$, так что система имеет единственное решение

$$\bar{a}_0 = 1,004, \quad \bar{a}_1 = 0,984.$$

В результате мы получаем следующую линейную аппроксимацию рассматриваемой табличной функции

$$F(x) = 1,004 + 0,984x. \quad (106)$$

Теперь, когда функция (106) построена, можно подсчитать погрешность аппроксимации в точках сетки:

$$\delta_i = y_i - (1,004 + 0,984x_i), \quad i = 0,1,2,3,4.$$

В результате получаем

$$\delta_0 = -0,054, \quad \delta_1 = -0,044, \quad \delta_2 = 0,052, \quad \delta_3 = -0,020, \quad \delta_4 = -0,022. \quad (107)$$

Отметим, что наибольшая по модулю погрешность достигается в точке $x_0 = 0$: $|\delta_0| = 0,054 > |\delta_i|$, $i = 1,2,3,4$.

В заключение сделаем важное замечание. Обычно бывает известна точность ε , с которой задаются значения функции y_i . Например, если речь идет об экспериментальных данных, то ошибка в определении y_i зависит от методики проведения измерений и точности приборов.

Предположим, что в разобранный примере числа y_i заданы с точностью $\varepsilon = 0,1$. В этом случае построенная линейная функция согласуется с доступной нам информацией о функции $y(x)$: погрешности (107) по модулю не превышают ε . В результате мы можем утверждать, что в пределах точности задания таблицы зависимость y от x можно принять линейной.

Это видно на рис.4. На нем показаны точки (x_i, y_i) , соответствующие рассматриваемой таблице. Для каждой из них указан доверительный интервал $y_i - 0,1 \leq y \leq y_i + 0,1$, в пределах которого может реально находиться значение функции $y(x_i)$ с учетом точности задания величины y_i . Прямая (106) везде проходит внутри доверительных интервалов, что подтверждает сделанный выше вывод.

Рассмотрим теперь противоположный случай: будем считать, что величины y_i заданы с более высокой точностью $\varepsilon = 0,01$. При такой точности построенная линейная аппроксимация (106) не согласуется с данными таблицы: погрешность аппроксимации (107) превышает по модулю ε . В этом случае нужно либо увеличить число членов в разложении функции $F(x)$, добавив к линейной функции квадратичный член a_2x^2 , либо заменить систему функций $\varphi_k(x)$, по которым ведется разложение, на какую-нибудь другую.