

Министерство высшего и профессионального образования Российской Федерации
Дальневосточный государственный университет
Сплайн-функции
Методическое пособие
Владивосток
1999 г.

УДК 519.63

Рассматриваются методы построения, исследования и применения сплайн-функций в численном анализе. Подробно рассматриваются вопросы приближения функций, численное дифференцирование и интегрирование.

Рекомендуется студентам университетов и технических вузов, инженерам и научным работникам, применяющих сплайн-функции на практике.

Методическое пособие подготовлено кафедрой компьютерных технологий. Печатается по решению учебно-методического Совета ДВГУ.

Составитель А. Г. Колобов.

© Дальневосточный Государственный Университет, 1999 г.

Введение

Большинство численных методов решения задач математического анализа так или иначе связано с аппроксимацией функций. Это и собственно задачи приближения функций (интерполяция, сглаживание, наилучшие приближения) и задачи, в которых аппроксимация присутствует как промежуточный этап исследования (численное дифференцирование и интегрирование, численное решение дифференциальных и интегральных уравнений).

Типичной задачей приближения является задача интерполяции: по заданной таблице чисел $(x_i, f(x_i))$, $i = 0, \dots, N$, восстановить функцию $f(x)$ с той или иной точностью на отрезке $[a, b]$ действительной оси. Классический метод ее решения состоит в построении интерполяционного многочлена Лагранжа, определяемого равенством [1]:

$$L_N(x) = \sum_{i=0}^N f(x_i) \frac{\omega_N(x)}{(x - x_i) \omega'_N(x_i)}, \quad \omega_N(x) = \prod_{i=0}^N (x - x_i).$$

Хотя, согласно теореме Вейерштрасса всякая непрерывная функция $f(x)$ на отрезке $[a, b]$ может быть как угодно хорошо приближена многочленами, практические возможности применения многочленов Лагранжа ограничены.

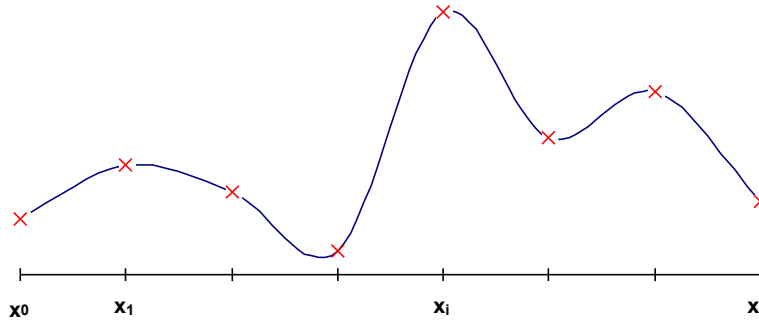
Прежде всего, используя подобный аппарат, мы должны быть уверены, что, выбрав достаточно большое число узлов интерполяции, получим хорошее приближение интерполируемой функции. Однако, как показывает ряд простых примеров, это часто нельзя гарантировать.

С.Н.Берштейном (1916 г.) было установлено, что последовательность интерполяционных многочленов Лагранжа $\{L_N(x)\}$, построенных для непрерывной функции $f(x) = |x|$ на отрезке $[-1, 1]$ по равноотстоящим узлам $(x_0 = -1, x_N = 1)$, с возрастанием N не стремится к $f(x)$. Еще более любопытен другой пример, восходящий к Рунге (1901 г.) и состоящий в том, что указанный интерполяционный процесс не сходится на $[-1, 1]$ даже для гладкой сколь угодно раз дифференцируемой функции $f(x) = (1 + 25x^2)^{-1}$. В обоих случаях $\lim_{N \rightarrow \infty} \max_{-1 \leq x \leq 1} |f(x) - L_N(x)| = \infty$.

Иногда эти трудности удается преодолеть путем специального выбора узлов интерполяции или за счет перехода к каким-либо обобщенным многочленам. Однако такой путь, как правило весьма усложняет вычисления и к тому же не избавляет нас от второй проблемы — быстрого накопления ошибок округления с ростом степени многочлена. Поэтому на практике для того, чтобы достаточно хорошо приблизить функцию, вместо построения интерполяционного многочлена высокой степени используют интерполяцию кусочными многочленами.

Примером такого рода является кусочно-линейная интерполяция. В общем случае отрезок $[a, b]$ точками $a = \xi_0 < \xi_1 < \dots < \xi_n = b$ разбивается на части и на каждом промежутке $[\xi_i, \xi_{i+1}]$, $i = 0, \dots, n-1$ строится свой интерполяционный многочлен. Полученные таким образом многочлены (обычно одной и той же степени) дают интерполяцию функции $f(x)$ на всем отрезке $[a, b]$, которая, вообще говоря, не обеспечивает гладкого перехода от одного звена к другому и может быть даже разрывной, если точки ξ_i , $i = 1, \dots, n-1$, не включаются в число узлов интерполяции. Это допустимо, если не требуется восстанавливать функцию с заданной степенью гладкости. Для гладкого восстановления таблично заданной функции нужно увеличить степень составляющих многочленов, а остающиеся свободными коэффициенты определять из условий гладкого сопряжения многочленов на соседних промежутках. Получающиеся при этом гладкие кусочно-многочленные функции с однородной структурой (составленные из многочленов одной и той же степени) называются сплайн-функциями или просто сплайнами. Простейший и исторически самый старый пример сплайна — ломанная. Термин сплайн произошел от английского spline, что в переводе означает рейка, стержень — название приспособления, которое применяли чертежники для проведения гладких кривых через заданные точки.

Возьмем гибкую стальную линейку, поставим ее на ребро и, закрепив один конец в заданной точке $(x_0, f(x_0))$, поместим между опорами, которые располагаются так, чтобы линейка проходила через заданные точки (рис.1).



Функция $S(x)$, описывающая профиль линейки, является кубическим многочленом между двумя соседними точками опоры и дважды непрерывно дифференцируемой функцией на всем промежутке интегрирования.

В отличие от интерполяционных многочленов Лагранжа, последовательность интерполяционных кубических сплайнов на равномерной сетке узлов всегда сходится к интерполируемой непрерывной функции, причем сходимость повышается с улучшением дифференциальных свойств функции $f(x)$. Так, для функции $f(x) = (1 + 25x^2)^{-1}$ из примера Рунге кубический сплайн на сетке с числом узлов $N=6$ дает погрешность того же порядка, что и многочлен $L_5(x)$.

Алгоритмы построения кубических сплайнов являются весьма простыми и эффективно реализуются на ЭВМ, причем внесение ошибок округления при вычислениях оказывается незначительными.

Опыт применения сплайн-функций как аппарата приближения функций в численном анализе показывает, что во всех известных случаях удавалось добиться ощутимых результатов по сравнению с классическим аппаратом многочленов. В одних задачах переход к сплайнам приводит к повышению точности результатов, в других — к значительному сокращению вычислительных затрат, в третьих — достигаются оба эффекта одновременно. Наконец, с помощью сплайнов удалось решить и такие задачи, которые другим путем решить было бы невозможно.

Среди них на первом месте стоит проблема представления и хранения геометрической информации в самых различных областях знания, будь то естественные науки, техника, архитектура, картография. Традиционно в более или менее сложных ситуациях эта задача решается путем изображения объекта или процесса на плоскости в виде графиков, чертежей и т.п. Вследствие ограниченности масштабов изображений этот способ принципиально не может обеспечить требуемую точность во всех случаях. Применение для данных целей сплайнов позволяет хранить геометрическую информацию в числовой форме и с любой точностью. При обработке информации на ЭВМ использование сплайнов позволяет на единой методологической основе разрабатывать математическое обеспечение средств машинной графики. Особенно широкое применение получили сплайны в технике как аппарат для математического моделирования поверхностей деталей и агрегатов сложной формы, таких, как аэродинамические обводы летательных аппаратов, корпуса судов, лопасти гидротурбин, кузова легковых автомобилей и т.п.. Такие математические модели стали необходимыми при создании систем автоматизации проектирования изделий на основе ЭВМ, технологической подготовки их производства, включая разработку программы для оборудования.

Применение методов Ритца-Галеркина и коллокации решения краевых задач математической физики было весьма ограниченным из-за больших вычислительных трудностей. При переводе этих методов на сплайновую основу прогресс был столь значителен, что новые варианты методов получили специальное название: метод конечных элементов и метод сплайн-коллокации. Область применения сплайнов непрерывно растет.

Определение сплайнов.

Пусть на отрезке $[a, b]$ задано разбиение $\Delta: a = x_0 < x_1 < \dots < x_N = b$. Для целого $k \geq 0$ через $C^k = C^k[a, b]$ обозначим множество k раз непрерывно дифференцируемых на $[a, b]$ функций, а через $C^{-1}[a, b]$ — множество кусочно-непрерывных функций с точками разрыва первого рода.

Определение. Функция $S_{n,v}(x)$ называется **сплайном степени n дефекта v** (v — целое число, $0 \leq v \leq n+1$) с узлами на сетке Δ , если

а) на каждом отрезке $[x_i, x_{i+1}]$ функция $S_{n,v}(x)$ является многочленом степени n , т.е.

$$S_{n,v}(x) = \sum_{\alpha=0}^n a_{\alpha}^i (x - x_i)^{\alpha} \quad \text{для } x \in [x_i, x_{i+1}], \quad i = 0, \dots, N-1 \quad (1)$$

$$b) S_{n,v}(x) \in C^{n-v}[a,b]$$

Определение сплайна имеет смысл и на всей вещественной оси, если положить $a = -\infty$, $b = +\infty$.

Множество сплайнов, удовлетворяющих определению, обозначим через $S_{n,v}(\Delta)$. Ясно, что этому множеству принадлежат и сплайны степени n дефекта $v_1 < v$ и сплайны степени $n_1 < n$ дефекта $v_1 < v$, если $n_1 - v_1 \geq n - v$, в том числе многочлены степени не выше n . Так как обычные операции сложения элементов из $S_{n,v}(\Delta)$ и их умножения на действительные числа не выводят за пределы множества, то оно является линейным множеством или линейным пространством. Размерность этого пространства $n+1+v(N-1)$ [3].

Простейшим примером сплайна является единичная функция Хевисайда

$$\Theta(x) = \begin{cases} 1 & \text{при } x \geq 0, \\ 0 & \text{при } x < 0, \end{cases}$$

с которой естественным образом связана усеченная степенная функция

$$x_+^n = x^n \Theta(x) = \begin{cases} x^n & \text{при } x \geq 0, \\ 0 & \text{при } x < 0. \end{cases}$$

Функции $\Theta(x)$ и x_+^n являются сплайнами соответственно нулевой степени и степени n дефекта 1 с единственным узлом в нулевой точке.

Базисные сплайны с конечными носителями.

В математическом анализе встречаются конструкции, связанные с финитными функциями, то есть гладкими функциями, которые определяются на всей действительной оси, но отличны от нуля лишь на некотором конечном интервале (носителе). Рассмотрим финитные сплайны из пространства $S_{n,1}(\Delta)$.

Расширим сетку Δ , добавив дополнительно точки $x_{-n} < \dots < x_{-1} < a$; $b < x_{N+1} < \dots < x_{N+n}$ (можно положить, например, $x_{-i} = x_0 - i(x_1 - x_0)$, $x_{N+i} = x_N + i(x_N - x_{N-1})$, $i=1, \dots, n$).

Возьмем функцию $\phi_n(x, t) = (-1)^{n+1}(n+1)(x-t)_+^n$ и построим для нее разделенные разности $(n+1)$ -го порядка по значениям аргумента $t = x_i, \dots, x_{i+n+1}$. В результате получаются функции переменной x :

$$\tilde{B}_n^i(x) = \varphi_n[x, x_i, \dots, x_{i+n+1}], i = -n, \dots, N-1. \quad (2)$$

Так как для разделенной разности $(n+1)$ -го порядка от функции $g(t)$ по точкам x_i, \dots, x_{i+n+1} справедливо равенство

$$g[x_i, \dots, x_{i+n+1}] = \sum_{p=i}^{i+n+1} \frac{g(x_p)}{\omega'_{n+1,i}(x_p)}, \quad \omega_{n+1,i}(t) = \prod_{j=i}^{i+n+1} (t - x_j)$$

[1, с.38], то

$$\tilde{B}_n^i(x) = (-1)^{n+1}(n+1) \sum_{p=i}^{i+n+1} \frac{(x - x_p)_+^n}{\omega'_{n+1,i}(x_p)}, i = -n, \dots, N-1.$$

Из определения усеченных степенных функций следует, что функция $\tilde{B}_n^i(x)$ является сплайном степени n дефекта 1 на сетке узлов x_i, \dots, x_{i+n+1} .

Приведем основные свойства функций (2):

$$1) \frac{n}{n+1} \tilde{B}_n^i(x) = \frac{x - x_i}{x_{i+n+1} - x_i} \tilde{B}_{n-1}^i(x) + \frac{x_{i+n+1} - x}{x_{i+n+1} - x_i} \tilde{B}_{n-1}^{i+1}(x),$$

$$2) \tilde{B}_n^i(x) \begin{cases} > 0 & \text{для } x \in (x_i, x_{i+n+1}), \\ \equiv 0 & \text{для } x \notin (x_i, x_{i+n+1}), \end{cases}$$

$$3) \int_{-\infty}^{+\infty} \tilde{B}_n^i(x) dx = 1,$$

4) функции $\tilde{B}_n^i(x)$ являются сплайнами степени n дефекта 1 с конечными носителями минимальной длины, линейно независимы и образуют базис в пространстве сплайнов $S_{n,1}(\Delta)$.

Доказательство этих свойств можно найти в [3].

Функции $\tilde{B}_n^i(x)$, $i = -n, \dots, N-1$, называются **базисными сплайнами с конечными носителями минимальной длины** (В-сплайнами). В силу свойства 4) всякий сплайн $S(x) \in S_{n,1}(\Delta)$ может быть единственным образом записан в виде:

$$S(x) = \sum_{i=-n}^{N-1} b_i \tilde{B}_n^i(x), \quad (3)$$

где b_i - некоторые постоянные коэффициенты. Эту запись сплайна называют его **представлением через В-сплайны**.

При практических вычислениях удобнее использовать не сами В-сплайны, а функции, получающиеся из них умножением на постоянные множители $B_n^i(x) = \frac{x_{i+n+1} - x_i}{n+1} \tilde{B}_n^i(x)$.

Эти функции называются **нормализованными В-сплайнами**. Нормирующий множитель равен среднему арифметическому шагов $h_i = x_{i+1} - x_i$ на отрезке, где В-сплайн отличен от нуля.

Свойство 1) для нормализованных В-сплайнов имеет вид:

$$B_n^i(x) = \frac{x - x_i}{x_{i+n} - x_i} B_{n-1}^i(x) + \frac{x_{i+n+1} - x}{x_{i+n+1} - x_{i+1}} B_{n-1}^{i+1}(x) \quad (4)$$

С его помощью легко можно построить последовательность сплайнов B_0^i, B_1^i, \dots . Приведем первые четыре функции этой последовательности для случая равноудаленных узлов $h_i = h$.

Будем обозначать $\delta_{in} = \frac{1}{h}(x - x_{i+(n+1)/2})$. Точка $x_{i+(n+1)/2}$ - это середина отрезка носителя В-сплайна. Тогда имеем

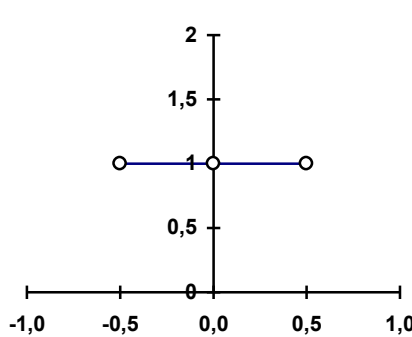
$$B_0^i = \begin{cases} 1 & \text{для } x \in (x_i, x_{i+1}), \\ 0 & \text{для } x \notin (x_i, x_{i+1}); \end{cases}$$

$$B_1^i(\delta) = (1 + \delta_{i1})B_0^i + (1 - \delta_{i1})B_0^{i+1};$$

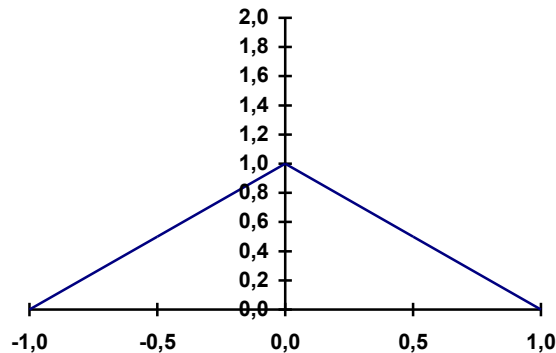
$$B_2^i(\delta) = \frac{1}{2} \left(\frac{3}{2} + \delta_{i2} \right)^2 B_0^i + \left(\frac{3}{4} - \delta_{i2}^2 \right) B_0^{i+1} + \frac{1}{2} \left(\frac{3}{2} - \delta_{i2} \right)^2 B_0^{i+2};$$

$$B_3^i(\delta) = \frac{1}{6} (2 + \delta_{i3})^3 B_0^i + \left(\frac{2}{3} - \delta_{i3}^2 - \frac{\delta_{i3}^3}{2} \right) B_0^{i+1} + \left(\frac{2}{3} - \delta_{i3}^2 + \frac{\delta_{i3}^3}{2} \right) B_0^{i+2} + \frac{1}{6} (2 - \delta_{i3})^3 B_0^{i+3}.$$

Эти В-сплайны изображены на рис.2, а,б,в,г соответственно



а)



б)

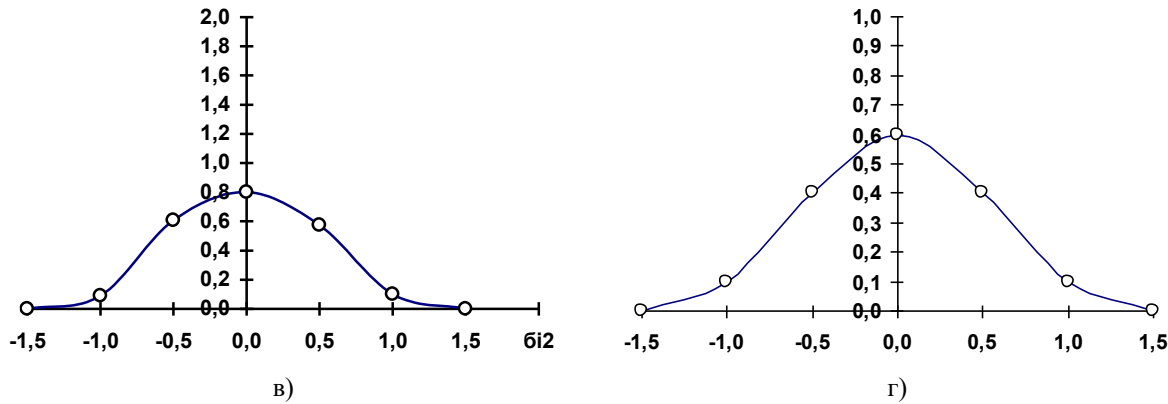


рис . 2

Вычисление сплайнов и их производных.

При вычислении сплайнов первостепенное значение имеет форма их представления. В настоящее время наиболее распространенным является кусочно-многочленное представление сплайнов (1). В этом случае для запоминания сплайна требуется хранить $N + 1$ абсцисс узлов и коэффициенты многочленов, количество которых равно $(n + 1)N$, то есть $(n + 2)N + 1$ чисел. Вычисление значения сплайна в точке состоит в вычислении значения многочлена и при использовании, например, схемы Горнера требует выполнения $2n$ арифметических операций. Коэффициенты составляющих сплайн многочленов связаны $(n - \nu + 1)(N - 1)$ условиями гладкости в узлах сетки. Поэтому запоминать можно не все $(n + 1)N$ коэффициентов, а лишь часть их, но не менее $n + 1 + \nu(N - 1)$ чисел, что равно размерности пространства $S_{n,\nu}(\Delta)$. Остальные коэффициенты находятся непосредственно в процессе вычисления сплайна, что естественно приводит к увеличению числа арифметических операций на этом этапе. Какие из коэффициентов выгодно запоминать, а какие вычислять, зависит от условий конкретной задачи.

Остановимся здесь подробнее на вычислении сплайнов дефекта 1, представленных через нормализованные В-сплайны. На отрезке $[x_i, x_{i+1}]$

$$S(x) = \sum_{p=i-n}^i b_p B_n^p(x) \quad (5)$$

Для запоминания сплайна нужно хранить минимальный объем информации $(2N + n + 1)$ чисел.

Нормализованные В-сплайны выражаются через $(n + 1)$ -е разделенные разности функции $\varphi_n(x, t)$ по аргументу t :

$$B_n^p(x) = \frac{x_{p+n+1} - x_p}{n + 1} \varphi_n[x; x_p, \dots, x_{p+n+1}]$$

или через ее n -е разности

$$B_n^p(x) = \frac{1}{n + 1} \left(\varphi_n[x; x_{p+1}, \dots, x_{p+n+1}] - \varphi_n[x; x_p, \dots, x_{p+n}] \right).$$

Функции $\varphi_n(x, t)$, а вместе с нею и $B_n^p(x)$ имеют по аргументу x непрерывные производные до порядка $n - 1$ и разрывные производные порядка n , при этом $\frac{d}{dx} \varphi_n(x, t) = -(n + 1) \varphi_{n-1}(x, t)$. Нормализованные В-сплайны суть линейные комбинации значений функции $\varphi_n(x, t)$ при $t = x_p, \dots, x_{p+n+1}$, и, очевидно, операции составления разделенных разностей и дифференцирования по аргументу x перестановочны.

Имеем

$$\left[B_n^p(x) \right]' = \varphi_{n-1}[x; x_p, \dots, x_{p+n}] - \varphi_{n-1}[x; x_{p+1}, \dots, x_{p+n+1}] = \frac{n}{x_{p+n} - x_p} B_{n-1}^p(x) - \frac{n}{x_{p+n+1} - x_{p+1}} B_{n-1}^{p+1}(x).$$

Тогда производная сплайна имеет вид

$$S'(x) = n \sum_{p=i-n+1}^i b_p^{(1)} B_{n-1}^p(x),$$

где $b_p^{(1)} = \frac{b_p - b_{p-1}}{x_{n+p} - x_p}$.

Производная сплайна порядка r выражается формулой

$$S^{(r)}(x) = n \dots (n-r+1) \sum_{p=i-n+r}^i b_p^{(r)} B_{n-r}^p(x), \quad (6)$$

если обозначать

$$b_p^{(0)} = b_p, \quad b_p^{(l)} = \frac{b_p^{(l-1)} - b_{p-1}^{(l-1)}}{x_{p+n+1-l} - x_p}, \quad l = 1, \dots, r, \quad (7)$$

Формулы (5),(6) и рекуррентное соотношение (4) дают необходимые средства для построения алгоритмов вычисления сплайнов и их производных.

Алгоритм: Учитывая тождество (4), выразим сплайн (5) через В-сплайн более низких степеней с коэффициентами, являющимися многочленами от x :

$$S(x) = \sum_{p=i-n+1}^i b_p^{[1]}(x) B_{n-1}^p(x),$$

причем

$$b_p^{[1]}(x) = \frac{x_{p+n} - x}{x_{p+n} - x_p} b_{p-1} + \frac{x - x_p}{x_{p+n} - x_p} b_p.$$

Продолжая эту процедуру, получаем

$$S(x) = \sum_{p=i-n+1}^i b_p^{[l]}(x) B_{n-l}^p(x),$$

где

$$b_p^{[l]}(x) = \frac{x_{p+n+1-l} - x}{x_{p+n+1-l} - x_p} b_{p-1}^{[l-1]}(x) + \frac{x - x_p}{x_{p+n+1-l} - x_p} b_p^{[l-1]}(x), \quad l = 1, \dots, n. \quad (8)$$

Так как

$$B_0^i(x) = \begin{cases} 1 & \text{для } x \in [x_i, x_{i+1}), \\ 0 & \text{для } x \notin [x_i, x_{i+1}), \end{cases} \quad (9)$$

то $S(x) = b_i^{[n]}(x)$ для $x \in [x_i, x_{i+1})$. Чтобы вычислить значение сплайна в какой-либо точке x этого промежутка, по формулам (8) построим таблицу:

$$\begin{array}{cccc} b_{i-n}^{[0]} & b_{i-n+1}^{[0]} & \dots & b_i^{[0]} \\ & b_{i-n+1}^{[1]}(x) & \dots & b_i^{[1]}(x) \\ & & \dots & \vdots \\ & & & b_i^{[n]}(x) \end{array} \quad (10)$$

Число $b_i^{[n]}(x)$ и является искомым значением $S(x)$. Такая организация вычислений требует выполнения $4n(n+1)$ арифметических операций.

Вычисление значения производных $S^{(r)}(x)$ производится следующим образом:

1) По формулам (7) находятся коэффициенты $b_p^{(r)}$, $p = i-n+r, \dots, i$, заполняющие таблицу

$$\begin{array}{ccccccc}
b_{i-n}^{(0)} & \cdots & b_{i-n+r}^{(0)} & \cdots & b_i^{(0)} & & \\
& & \vdots & & \vdots & & \\
& & \cdots \vdots \cdots & \cdots & \cdots \vdots & & \\
& & \vdots & & \vdots & & \\
& & b_{i-n+r}^{(r)} & \cdots & b_i^{(r)} & &
\end{array} \quad (11)$$

2) по формулам (8) с исходными данными $\bar{b}_p^{[0]} = b_p^{(r)}$ аналогично таблице (10) строится таблица

$$\begin{array}{ccccccc}
\bar{b}_{i-n+r}^{[0]} & \bar{b}_{i-n+r+1}^{[0]} & \cdots & \bar{b}_i^{[0]} & & & \\
& \bar{b}_{i-n+r+1}^{[1]}(x) & \cdots & \bar{b}_i^{[1]}(x) & & & \\
& & \cdots & \vdots & & & \\
& & & \bar{b}_i^{[n-r]}(x) & & &
\end{array} \quad (12)$$

3) Искомое значение $S^{(r)}(x) = n \dots (n-r+1) \bar{b}_i^{[n-r]}(x)$.

На весь процесс вычислений расходуется $4(n-r)(n-r+1) + \frac{3}{2}(2n-r+\frac{5}{3})r$ арифметических операций. Исследуем устойчивость алгоритма.

Пусть на некотором шаге вместо точных величин $b_p^{[l-1]}(x)$ мы нашли величины $\tilde{b}_p^{[l-1]}(x)$, причем $|\tilde{b}_p^{[l-1]}(x) - b_p^{[l-1]}(x)| \leq \varepsilon$. Тогда на следующем шаге получится $\tilde{b}_p^{[l]}(x)$ и, очевидно,

$$\left| \tilde{b}_p^{[l]}(x) - b_p^{[l]}(x) \right| \leq \frac{x_{p+n+1-l} - x}{x_{p+n+1-l} - x_p} \varepsilon + \frac{x - x_p}{x_{p+n+1-l} - x_p} \varepsilon = \varepsilon.$$

Значит, процесс составления таблицы (10) устойчив.

Хуже обстоит дело с вычислением производных $S^{(r)}(x)$ (6). Если значения $b_p^{(l-1)}$ найдены с ошибкой ε , то погрешность вычисления $b_p^{(l)}$ по формулам (7) уже будет

$$\left| \tilde{b}_p^{(l)} - b_p^{(l)} \right| \leq \frac{2\varepsilon}{x_{p+n+1-l} - x_p}.$$

При густой сетке узлов знаменатель справа мал, что приводит к возрастанию ошибки в определении $b_p^{(l)}$. Так как эти величины в обоих алгоритмах играют роль исходных данных, то все последующие, хотя и устойчивые, вычисления проводятся с той точностью, с какой найдены $b_p^{(l)}$. Следует также учесть возрастание ошибки на последнем шаге вычисления $S^{(r)}(x)$ за счет умножения на число $n \dots (n-r+1)$. Таким образом, при вычислении производной, особое внимание следует уделить выбору сетки, шаг которой должен увязываться с величиной погрешности исходных значений коэффициентов b_p . В противном случае попытка получить более точный результат путем использования густой сетки может дать прямо противоположные ожидаемому эффект. Иногда с целью повышения точности результата разумнее сделать сетку более “редкой”, выбросив некоторые из узлов.

Сплайны первой степени

Сплайны

$$f'_i = f'(x_i)$$

Эрмитовы кубические сплайны

Пусть в узлах сетки $\Delta: a = x_0 < x_1 < \dots < x_n = b$ заданы значения некоторой функции $f(x)$ и её производной $f'(x)$:

$$f_i = f(x_i), f'_i = f'(x_i), i = 0, 1, \dots, N.$$

Кубическим интерполяционным сплайном дефекта 2 (эрмитовым кубическим сплайном) будем называть функцию $S_{3,2}(x)$, удовлетворяющую условиям:

$$1) \text{ на каждом из промежутков } [x_i, x_{i+1}] \quad S_{3,2}(x) = a_{i0} + a_{i1}(x - x_i) + a_{i2}(x - x_i)^2 + a_{i3}(x - x_i)^3;$$

$$2) S_{3,2}(x_i) = f_i, \quad S'_{3,2}(x_i) = f'_i, \quad i = 0, \dots, N.$$

Очевидно, $S_{3,2}(x) \in C^1[a, b]$. Вторая производная эрмитова кубического сплайна, вообще говоря, разрывна в узлах сетки Δ . Для сокращения записи опустим индекс, указывающий дефект сплайна.

Учитывая условия интерполяции, для вычисления коэффициентов $a_{i\alpha}$, $\alpha = 0, 1, 2, 3$, при каждом i имеем систему уравнений $S_3(x_i) = f_i$, $S_3(x_{i+1}) = f_{i+1}$, $S'_3(x_i) = f'_i$, $S'_3(x_{i+1}) = f'_{i+1}$.

Решив эту систему, получаем на $[x_i, x_{i+1}]$

$$S_3(x) = \varphi_1(t)f_i + \varphi_2(t)f_{i+1} + \varphi_3(t)h_i f'_i + \varphi_4(t)h_i f'_{i+1} \quad (16)$$

причем $\varphi_1(t) = (1-t)^2(1+2t)$, $\varphi_2(t) = t^2(3-2t)$, $\varphi_3(t) = t(1-t)^2$, $\varphi_4(t) = -t^2(1-t)$, $h_i = x_{i+1} - x_i$, $t = (x - x_i)/h_i$.

Приведем теорему об оценке погрешности интерполяции эрмитовыми кубическими сплайнами. Доказательство ее опускаем ввиду громоздкости (см., например, [3], стр. 60).

Теорема 2. Если $S_3(x)$ интерполирует на сетке Δ функцию $f(x)$, то имеют место оценки

$$\|S_3^{(r)}(x) - f^{(r)}(x)\|_\infty \leq R_r, \quad r = 0, 1, 2, 3,$$

где R_r даны в таблице 2.

Класс функций	R_0	R_1	R_2	R_3
$C^1[a, b]$	$\frac{3}{8}\bar{h}\omega(f')$	$\frac{3}{2}\omega(f')$	—	—
$W_\infty^2[a, b]$	$\frac{1}{16}\bar{h}^2\ f''(x)\ _\infty$	$0.24750\bar{h}\ f''(x)\ _\infty$	—	—
$C^1C_\Delta^2[a, b]$	$\frac{1}{32}\bar{h}^2\omega(f'')$	$0.12375\bar{h}\omega(f'')$	$\frac{4}{3}\omega(f'')$	—
$C^1W_{\Delta,\infty}^3[a, b]$	$\frac{1}{96}\bar{h}^3\ f'''(x)\ _\infty$	$0.032302\bar{h}^2\ f'''(x)\ _\infty$	$\frac{8}{27}\bar{h}\ f'''(x)\ _\infty$	—
$C^1C_\Delta^3[a, b]$	$\frac{1}{192}\bar{h}^3\omega(f''')$	$0.016151\bar{h}^2\omega(f''')$	$\frac{4}{27}\bar{h}\omega(f''')$	$\omega(f''')$
$C^1W_{\Delta,\infty}^4[a, b]$	$\frac{1}{384}\bar{h}^4\ f^{IV}(x)\ _\infty$	$\frac{\sqrt{3}}{216}\bar{h}^3\ f^{IV}(x)\ _\infty$	$\frac{1}{12}\bar{h}^2\ f^{IV}(x)\ _\infty$	$\frac{1}{2}\bar{h}\ f^{IV}(x)\ _\infty$

Порядки погрешности в классе $C^1W_{\Delta,\infty}^4[a, b]$ являются наивысшими для эрмитовых кубических сплайнов. Дальнейшее увеличение гладкости функции $f(x)$ не приводит к повышению порядка аппроксимации. Более того, при этом не могут быть уменьшены и константы в оценках.

Сплайны двух переменных на нерегулярной сетке

Пусть на плоскости xOy задано множество точек $P_i(x_i, y_i)$, $i = 0, 1, \dots, N$, в которых известны значения некоторой функции $f(P_i) = f_i$. Построим в плоскости xOy многоугольник такой, чтобы, во-первых, его вершины принадлежали множеству заданных точек и, во-вторых, все точки, не попадающие в вершины, лежали внутри него.

Полученный многоугольник разобьем на треугольники с вершинами в точках P_i таким образом, чтобы каждая точка области, ограниченной многоугольником, принадлежала одному, и только одному, из треугольников (рис.4).

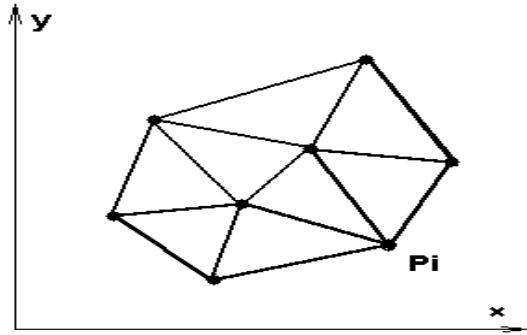


рис. 4

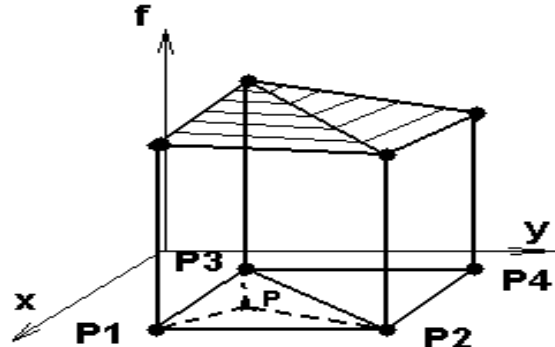


рис. 5

Описанная процедура называется триангуляцией. Заметим, что триангуляция осуществляется не единственным способом.

Предположим, что на плоскости xOy построена некая триангуляция. Интерполяционным сплайном первой степени назовем функцию $S_1(x, y) = S_1(p)$, которая в каждом из треугольников имеет вид $S_1(p) = ax + by + c$ и удовлетворяет условиям $S_1(p_i) = f_i$, $i = 0, 1, \dots, N$.

Геометрически такой сплайн представляет собой поверхность, составленную из кусков плоскостей (рис.5).

В силу локальности сплайна $S_1(P)$ достаточно рассмотреть один из треугольников, например треугольник Ω с вершинами в точках P_1, P_2, P_3 . Очевидно, $S_1(P)$ удовлетворяет уравнению

$$\begin{vmatrix} S_1(P) & x & y & 1 \\ f_1 & x_1 & y_1 & 1 \\ f_2 & x_2 & y_2 & 1 \\ f_3 & x_3 & y_3 & 1 \end{vmatrix} = 0$$

Введем обозначение

$$A(P_1, P_2, P_3) = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

Из аналитической геометрии известно, что величина $\frac{1}{2} |A(P_1, P_2, P_3)|$ равна площади треугольника Ω , причем, если нумерация вершин соответствует обходу треугольника против часовой стрелки, то $A(P_1, P_2, P_3) > 0$.

Разрешая уравнение (2) относительно $S_1(P)$, имеем

$$S_1(P) = \frac{1}{A(P_1, P_2, P_3)} [f_1 A(P, P_2, P_3) + f_2 A(P_1, P, P_3) + f_3 A(P_1, P_2, P)].$$

Условие $A(P_1, P_2, P_3) \neq 0$ является условием существования единственного сплайна $S_1(P)$ и предполагается в дальнейшем выполненным. Из геометрического смысла величины $A(P_1, P_2, P_3)$ следует, что оно эквивалентно требованию невырожденности треугольника Ω . Таким образом, необходимо, чтобы точки P_1, P_2, P не лежали на одной прямой.

Для того чтобы вычислить значение интерполяционного сплайна в некоторой точке P , очевидно, вначале следует выделить тот треугольник, которому эта точка принадлежит. Если вершины всех треугольников упорядочены в направлении против часовой стрелки, то признаком принадлежности точки P треугольнику Ω будет одновременное выполнение условий $A(P, P_2, P_3) > 0$, $A(P_1, P, P_3) > 0$, $A(P_1, P_2, P) > 0$.

На вопрос о погрешности интерполяции отвечает следующая теорема (доказательство см. в [3], стр.89).

Теорема 3. Если $f(x, y) \in W_\infty^2[\Omega]$, то

$$\|S_1(x, y) - f(x, y)\|_c \leq \frac{h^2}{3} \|D^2 f(x, y)\|_\infty, \|D^{r,s}(S_1(x, y) - f(x, y))\|_\infty \leq \frac{2h}{\sin \gamma} \|D^2 f(x, y)\|_\infty, r + s = 1,$$

где h - длина наибольшей стороны треугольника, γ - наименьший из углов,

$$\|D^2 f(x, y)\|_\infty = \max_{r+s=2} \{ \|D^{r,s} f(x, y)\|_\infty \}.$$

$$\text{Здесь } D^{r,s} f(x, y) = \frac{\partial^{r+s} f(x, y)}{\partial x^r \partial y^s}.$$

Существенным моментом в оценке погрешности производных является то, что ее правая часть зависит от минимального угла треугольника. Можно показать, что в общем случае это ограничение убрать нельзя. Однако для одного важного с практической точки зрения типа треугольников можно получить оценки приближения производных, не зависящие от величины его минимального угла.

Теорема 4. Пусть Ω - прямоугольный треугольник, катеты которого параллельны осям координат. Если $f(x, y) \in W_\infty^2[\Omega]$, то

$$\begin{aligned} \|D^{1,0}(S_1(x, y) - f(x, y))\|_\infty &\leq \tilde{h} \left\{ \|D^{1,1} f(x, y)\|_\infty + \frac{1}{2} \|D^{2,0} f(x, y)\|_\infty \right\}, \\ \|D^{0,1}(S_1(x, y) - f(x, y))\|_\infty &\leq \tilde{h} \left\{ \|D^{1,1} f(x, y)\|_\infty + \frac{1}{2} \|D^{0,2} f(x, y)\|_\infty \right\}, \end{aligned}$$

где \tilde{h} - длина наибольшего из катетов треугольника.

На основе этих оценок можно дать некоторые рекомендации по выбору триангуляции. А именно, если узлы интерполяции заданы произвольным нерегулярным образом и необходимо приближать производные интерполируемой функции, то следует стремиться к тому, чтобы не возникало треугольников с очень острыми углами. Однако, если речь идет только о приближении функции, то существенны лишь размеры треугольников, а величина их углов не играет никакой роли.

Кубические сплайны класса C^2

Пусть на отрезке $[a, b]$ в узлах сетки $\Delta: a = x_0 < x_1 < \dots < x_N = b$ заданы значения некоторой функции $f_i = f(x_i)$, $i = 0, \dots, N$. Интерполяционным кубическим сплайном $S(x)$ называется сплайн, удовлетворяющий условиям

$$S(x_i) = f_i, \quad i = 0, \dots, N \quad (17)$$

Сплайн $S(x)$ на каждом из отрезков $[x_i, x_{i+1}]$ определяется четырьмя коэффициентами, и поэтому для его построения на всем промежутке $[a, b]$ необходимо определить $4N$ коэффициентов. Условие $S(x) \in C^2[a, b]$ эквивалентно требованию непрерывности сплайна и его производных $S^{(r)}(x)$, $r = 0, 1, 2$ во всех внутренних узлах x_i , $i = 1, 2, \dots, N-1$ сетки Δ , что дает $3(N-1)$ равенств. Таким образом, вместе с равенствами (17) получается $4N-2$ соотношений. Два дополнительных условия обычно задается в виде ограничений на значения сплайна и его производных на концах промежутка $[a, b]$ (или вблизи концов) и называются краевыми условиями. Существует несколько различных видов краевых условий, из которых наиболее употребительными являются следующие типы:

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

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

$$S^{(r)}(a) = S^{(r)}(b), \quad r = 1, 2.$$

$$S'''(x_p + 0) = S'''(x_p - 0), \quad p = 1, N-1.$$

Условия типа III носят названия периодических. Естественно требовать их выполнения в том случае, когда интерполируемая функция $f(x)$ — периодическая с периодом $b-a$. В дальнейшем мы будем рассматривать только перечисленные краевые условия.

Перейдем к непосредственному описанию алгоритмов построения интерполяционных кубических сплайнов. Введем обозначение

$$S'(x_i) = m_i, \quad i = 0, \dots, N. \quad (18)$$

Очевидно, сплайн $S(x)$ можно рассматривать как эрмитов кубический сплайн, удовлетворяющий условиям (17), (18). В соответствии с формулой (16) для $x \in [x_i, x_{i+1}]$ получаем

$$S(x) = f_i(1-t)^2(1+2t) + f_{i+1}t^2(3-2t) + m_i h_i(1-t)^2 - m_{i+1} h_i t^2(1-t), \quad (19)$$

где $h_i = x_{i+1} - x_i$, $t = (x - x_i)/h_i$.

Кубический сплайн, представленный в таком виде на каждом из промежутков $[x_i, x_{i+1}]$, непрерывен вместе со своей первой производной всюду на $[a, b]$. Выберем величины m_i так, чтобы была непрерывна и вторая производная. Так как

$$S''(x) = (f_{i+1} - f_i)(6-12t)/h_i^2 + m_i(-4+6t)/h_i + m_{i+1}(-2+6t)/h_i$$

и

$$S''(x_i + 0) = 6 \frac{f_{i+1} - f_i}{h_i^2} - \frac{4m_i}{h_i} - \frac{2m_{i+1}}{h_i} \quad (20)$$

$$S''(x_i - 0) = -6 \frac{f_i - f_{i-1}}{h_{i-1}^2} + \frac{2m_{i-1}}{h_{i-1}} + \frac{4m_i}{h_{i-1}}$$

то условие непрерывности второй производной $S''(x_i + 0) = S''(x_i - 0)$ в точках $x_i, i = 1, 2, \dots, N-1$, принимает вид

$$\lambda_i m_{i-1} + 2m_i + \mu_i m_{i+1} = 3 \left(\mu_i \frac{f_{i+1} - f_i}{h_i} + \lambda_i \frac{f_i - f_{i-1}}{h_{i-1}} \right) \quad (21)$$

Здесь $\mu_i = h_{i-1}(h_{i-1} + h_i)^{-1}$, $\lambda_i = 1 - \mu_i$.

К уравнениям (21) следует добавить уравнения, вытекающие из краевых условий. Таким образом, получается система $N+1$ уравнений для определения $N+1$ неизвестных $m_i, i = 0, \dots, N$.

В случае краевых условий типов I и II

$$2m_0 + \mu_0^* m_1 = c_0^*, \quad (22)$$

$$\lambda_i m_{i-1} + 2m_i + \mu_i m_{i+1} = c_i, \quad i = 1, \dots, N-1,$$

$$\lambda_N^* m_{N-1} + 2m_N = c_N^*,$$

причем $c_i = 3 \left(\mu_i \frac{f_{i+1} - f_i}{h_i} + \lambda_i \frac{f_i - f_{i-1}}{h_{i-1}} \right)$.

Здесь для условий типа I

$$\mu_0^* = \lambda_N^* = 0, \quad c_0^* = 2f_0', \quad c_N^* = 2f_N',$$

а для условий типа II согласно (20)

$$\mu_0^* = \lambda_N^* = 1, \quad c_0^* = 3 \frac{f_1 - f_0}{h_0} - \frac{h_0}{2} f_0'', \quad c_N^* = 3 \frac{f_N - f_{N-1}}{h_{N-1}} + \frac{h_{N-1}}{2} f_N''.$$

Чтобы вывести систему уравнений в случае краевых условий типа IV, определим

$$S'''(x) = \frac{6}{h_i^2} \left(m_{i+1} + m_i - 2 \frac{f_{i+1} - f_i}{h_i} \right).$$

Отсюда получаем условия непрерывности функции $S'''(x)$ в точках x_p , $p = 1, N-1$:

$$\frac{1}{h_p^2} \left(m_{p+1} - m_p - 2 \frac{f_{p+1} - f_p}{h_p} \right) = \frac{1}{h_{p-1}^2} \left(m_p + m_{p-1} - 2 \frac{f_p - f_{p-1}}{h_{p-1}} \right),$$

из которых вытекают уравнения, отвечающие краевым условиям типа IV:

$$\begin{aligned} m_0 + (1 - \gamma_0^2) m_1 - \gamma_0^2 m_2 &= 2 \left(\frac{f_1 - f_0}{h_0} - \gamma_0^2 \frac{f_2 - f_1}{h_1} \right), \\ -\gamma_N^2 m_{N-2} + (1 - \gamma_N^2) m_{N-1} + m_N &= 2 \left(\frac{f_N - f_{N-1}}{h_{N-1}} - \gamma_N^2 \frac{f_{N-1} - f_{N-2}}{h_{N-2}} \right), \end{aligned} \quad (23)$$

где $\gamma_0 = h_0/h_1$, $\gamma_N = h_{N-1}/h_{N-2}$. Исключив неизвестные m_0, m_N из уравнений (21) и (23), приходим к системе

$$\begin{cases} (1 + \gamma_0) m_1 + \gamma_0 m_2 = c_1^* \\ \lambda_i m_{i-1} + 2m_i + \mu_i m_{i+1} = c_i, \quad i = 2, \dots, N-2, \\ \gamma_N m_{N-2} + (1 + \gamma_N) m_{N-1} = c_{N-1}^* \end{cases} \quad (24)$$

$$\text{где } c_1^* = \frac{1}{3} c_1 + 2\gamma_0 \frac{f_2 - f_1}{h_1}, \quad c_{N-1}^* = \frac{1}{3} c_{N-1} + 2\gamma_N \frac{f_{N-1} - f_{N-2}}{h_{N-2}}.$$

Если $f(x)$ - периодическая функция, то, продолжая периодическим образом сетку Δ и в соответствии с этим полагая $f_0 = f_N$, $f_{N+1} = f_1$, $m_0 = m_N$, $m_1 = m_{N+1}$, $h_N = h_0$, можем записать условие (21) в точке x_N . Система для определения m_i выглядит следующим образом:

$$\begin{cases} 2m_i + \mu_i m_2 + \lambda_i m_N = c_1, \\ \lambda_i m_{i-1} + 2m_i + \mu_i m_{i+1} = c_i, \quad i = 2, \dots, N-1, \\ \mu_N m_1 + \lambda_N m_{N-1} + 2m_N = c_N. \end{cases} \quad (25)$$

Итак, построение интерполяционного кубического сплайна по формуле (19) сводится к вычислению величин m_i . Для граничных условий типов I и II они находятся из системы (22), для условий типа IV из уравнений (23), (24), а в периодическом случае из (25) в предположении $m_0 = m_N$.

Матрицы систем во всех четырех случаях суть матрицы с диагональным преобладанием (см. Приложение 2). Такие матрицы невырождены, и поэтому системы имеют и притом единственные решения.

Таким образом, справедлива

Теорема 5. Интерполяционный кубический сплайн $S(x)$, удовлетворяющий условиям (17) и одному из типов краевых условий I-IV, существует и единственен.

Решения систем уравнений относительно m_i находятся методом прогонки (см. приложение 3).

В некоторых случаях более удобным является другое представление кубического сплайна, в котором вместо величин m_i присутствует $M_i = S''(x_i)$, $i = 0, \dots, N$. Используя то обстоятельство, что на каждом промежутке $[x_i, x_{i+1}]$ $S(x)$ есть кубический многочлен, а также условия

$$S(x_i) = f_i, \quad S(x_{i+1}) = f_{i+1}, \quad S''(x_i) = M_i, \quad S''(x_{i+1}) = M_{i+1},$$

нетрудно получить для $S(x)$ следующую формулу:

$$S(x) = f_i(1-t) + f_{i+1}t - \frac{h_i^2}{6} t(1-t) [(2-t)M_i + (1+t)M_{i+1}], \quad x \in [x_i, x_{i+1}], \quad i = 0, 1, \dots, N-1. \quad (26)$$

Отсюда

$$S'(x) = \frac{f_{i+1} - f_i}{h_i} - \frac{h_i}{6} [(2-6t+3t^2)M_i + (1-3t^2)M_{i+1}], \quad (27)$$

$$S''(x) = M_i(1-t) + M_{i+1}t, \quad (28)$$

$$S'''(x) = (M_{i+1} - M_i)/h_i. \quad (29)$$

Из (26) очевидно, что функция $S(x)$ непрерывна в точках x_i , $i = 1, \dots, N-1$, а из (28) следует, что непрерывна и ее вторая производная. Далее, согласно (27)

$$S'(x_i + 0) = \frac{f_{i+1} - f_i}{h_i} - \frac{h_i}{6}(2M_i + M_{i+1}), \quad (30)$$

$$S'(x_i - 0) = \frac{f_i - f_{i-1}}{h_{i-1}} + \frac{h_{i-1}}{6}(M_{i-1} + 2M_i),$$

и, следовательно, для того, чтобы была непрерывна первая производная сплайна, необходимо выполнение условий

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = \frac{6}{h_{i-1} + h_i} \left(\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}} \right), \quad i = 1, \dots, N-1.$$

Эти уравнения вместе с краевыми условиями одного из типов I-IV образуют систему относительно неизвестных M_i . Для условий типов I и II она имеет вид

$$\begin{cases} 2M_0 + \lambda_0^* M_1 = d_0^*, \\ \mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i, \quad i = 1, \dots, N-1, \\ \mu_N^* M_{N-1} + 2M_N = d_N^*, \end{cases} \quad (31)$$

где

$$d_i = \frac{6}{h_{i-1} + h_i} \left(\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}} \right).$$

В случае граничных условий типа I согласно (30)

$$\lambda_0^* = \mu_N^* = 1, \quad d_0^* = \frac{6}{h_0} \left(\frac{f_1 - f_0}{h_0} - f_0' \right), \quad d_N^* = \frac{6}{h_{N-1}} \left(f_N' - \frac{f_N - f_{N-1}}{h_{N-1}} \right),$$

а для условий типа II

$$\lambda_0 = \mu_N^* = 0, \quad d_0^* = 2f_0'', \quad d_N^* = 2f_N''.$$

Система уравнений при краевых условиях типа IV имеет вид

$$\begin{cases} (1 + \lambda_1)M_1 + (\lambda_1 - \mu_1)M_2 = \lambda_1 d_1, \\ \mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i, \\ (\mu_{N-1} - \lambda_{N-1})M_{N-2} + (1 + \mu_{N-1})M_{N-1} = \mu_{N-1} d_{N-1}, \\ M_0 = \lambda_1^{-1}(M_1 - \mu_1 M_2), \quad M_N = \mu_{N-1}^{-1}(M_{N-1} - \lambda_{N-1} M_{N-2}). \end{cases} \quad i = 2, \dots, N-2, \quad (32)$$

Наконец, в периодическом случае

$$\begin{cases} 2M_1 + \lambda_1 M_2 + \mu_1 M_N = d_1, \\ \mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i, \\ \lambda_N M_1 + \mu_N M_{N-1} + 2M_N = d_N, \\ M_0 = M_N. \end{cases} \quad i = 2, \dots, N-2, \quad (33)$$

Матрицы систем во всех четырех случаях - снова матрицы с диагональным преобладанием.

Получение оценок погрешности для кубических сплайнов представляет собой гораздо более сложную задачу, чем для эрмитовых кубических сплайнов, что связано в основном с неявным заданием сплайнов. Поэтому приведем здесь только результаты без доказательства (см. [3], стр. 101).

Теорема 6. Если $S(x)$ интерполирует функцию $f(x)$ и удовлетворяет краевым условиям одного из типов I, II, III, то имеют место оценки

$$\|S^{(r)}(x) - f^{(r)}(x)\|_\infty \leq R_r, \quad i = 0, 1, 2, 3, \quad (34)$$

где R_r даются таблицей 3.

В таблице обозначено $\beta = \frac{\bar{h}}{h} = \frac{\max_i h_i}{\min_i h_i}$.

Таблица 3.

Класс функций	R_0	R_1	R_2	R_3
$C^1[a, b]$	$\frac{9}{8} \bar{h} \omega(f')$	$4\omega(f')$	—	—
$W_\infty^2[a, b]$	$\frac{13}{48} \bar{h}^2 \ f''(x)\ _\infty$	$0.86229 \bar{h} \ f''(x)\ _\infty$	—	—
$C^2[a, b]$	$\frac{19}{96} \bar{h}^2 \omega(f'')$	$\frac{2}{3} \bar{h} \omega(f'')$	$4\omega(f'')$	—
$W_\infty^3[a, b]$	$\frac{41}{864} \bar{h}^3 \ f'''(x)\ _\infty$	$\frac{4}{27} \bar{h}^2 \ f'''(x)\ _\infty$	$\left(\frac{1}{2} + \frac{4\sqrt{3}}{9}\right) \bar{h} \ f'''(x)\ _\infty$	—
$C^2 C_\Delta^3[a, b]$	$\frac{41}{1728} \bar{h}^3 \omega(f''')$	$\frac{2}{27} \bar{h}^2 \omega(f''')$	$\left(\frac{1}{2} + \frac{2\sqrt{3}}{9}\right) \bar{h} \omega(f''')$	$\left(1 + \frac{4\sqrt{3}}{9} \beta\right) \omega(f''')$
$C^2 W_{\Delta, \infty}^4[a, b]$	$\frac{5}{384} \bar{h}^4 \ f^{(IV)}(x)\ _\infty$	$\frac{1}{24} \bar{h}^3 \ f^{(IV)}(x)\ _\infty$	$\frac{3}{8} \bar{h}^2 \ f^{(IV)}(x)\ _\infty$	$\frac{1}{2} (\beta^{-1} + \beta) \bar{h} \ f^{(IV)}(x)\ _\infty$

В случае краевых условий IV справедливы оценки погрешности, совпадающие по порядку с соответствующими оценками из теоремы 6.

При практическом использовании кубических сплайнов для целей интерполяции неизбежно приходится иметь дело с проблемой выбора краевых условий. Особую важность этот вопрос приобретает в том случае, когда требуется обеспечить хорошее качество приближения вблизи концов промежутка интерполяции $[a, b]$ (вдали от них, как можно показать, влияние граничных условий незначительно). Выбор краевых условий осуществляется в зависимости от того, какими данными мы располагаем об интерполируемой функции $f(x)$.

Если $f(x)$ периодическая, то следует использовать периодические краевые условия (тип III).

Когда известны значения $f'(x)$ или $f''(x)$ в точках a и b , то естественно воспользоваться краевыми условиями типа I или II. При этом, если существует возможность выбора между ними, то предпочтительнее условия типа I.

Наибольшие трудности возникают, когда заданы только узловые значения f_i . Вначале рассмотрим некоторые из известных рекомендаций. Часто предлагается использовать естественные краевые условия: $S'''(a) = S'''(b) = 0$ (смысл этого термина будет ясен из п.10). Однако такое решение весьма неудачно, так как для точек x , близких к концам отрезка $[a, b]$, в лучшем случае можно получить $|S^{(r)}(x) - f^{(r)}(x)| = O(\bar{h}^{2-r})$, $r = 0, 1$. То есть здесь точность приближения кубическим сплайном снижается до точности приближения сплайном первой степени. О приближении $f''(x)$ говорить вообще не приходится.

Другой подход состоит в том, что используются краевые условия типа I или II, но необходимые значения производных заменяются подходящими разностными аппроксимациями. Недостаток этого способа состоит в том, что приходится использовать односторонние аппроксимации, которые, как известно, имеют невысокую точность.

В большинстве случаев кардинальным решением вопроса является применение краевых условий типа IV. Этот вывод подтверждает и практический опыт расчетов.

Иногда можно рекомендовать краевые условия общего вида: $M_0 + \lambda M_1 = d_0$, $\mu M_{N-1} + M_N = d_N$, где свободные параметры λ, μ, d_0, d_N выбираются из каких-либо соображений. Аналогичные краевые условия можно записать и в терминах величин m_i .

Многие специалисты склонны видеть в необходимости задания краевых условий один из недостатков кубических сплайнов. На самом деле более правильной следует считать противоположную точку зрения. Действительно, по существу

через краевые условия мы включаем в конструкцию сплайна свободные параметры, выбором которых можно в существенной степени управлять поведением сплайна.

Кубические В-сплайны

В ряде приложений более эффективным по сравнению с кусочно-многочленным оказывается представление кубических сплайнов через В-сплайны. В п.2 последние были определены аксиоматическим образом. Здесь представлен другой способ их построения, применимый не только для полиномиальных сплайнов, но и различных их обобщений.

Дополним сетку $\Delta: a = x_0 < x_1 < \dots < x_N = b$ узлами $x_{-3} < x_{-2} < x_{-1} < a$, $b < x_{N+1} < x_{N+2} < x_{N+3}$. Эти узлы вполне произвольны, только в периодическом случае должно выполняться ограничение $h_{N+i} = h_i$.

Построим кубический В-сплайн, отличный от нуля на интервале (x_{i-2}, x_{i+2}) . В приложениях В-сплайны нечетных степеней удобно нумеровать по среднему узлу их интервалов-носителей. Искомый В-сплайн будем обозначать через $B_i(x)$ вместо $\tilde{B}_3^{i-2}(x)$. Положим еще $y_p = B_i(x_p)$, $M_p = B_i''(x_p)$.

Как и для всякого кубического сплайна, для $B_i(x)$ имеют место уравнения:

$$\mu_p M_{p-1} + 2M_p + \lambda_p M_{p+1} = \frac{6}{h_{p-1} + h_p} \left(\frac{y_{p+1} - y_p}{h_p} - \frac{y_p - y_{p-1}}{h_{p-1}} \right), \quad (35)$$

$$p = i-1, i, i+1.$$

Так как $B_i(x) \equiv 0$ при $x \notin (x_{i-2}, x_{i+2})$, то $B_i^{(r)}(x_{i-2}) = B_i^{(r)}(x_{i+2}) = 0$, $r = 0, 1, 2$. Эти условия вследствие (26)-(28) можно представить так

$$\begin{aligned} y_{i-2} = y_{i+2} = 0, \quad M_{i-2} = M_{i+2} = 0, \\ y_{i-1} = \frac{1}{6} h_{i-2}^2 M_{i-1}, \quad y_{i+1} = \frac{1}{6} h_{i+1}^2 M_{i+1} \end{aligned} \quad (36)$$

Найденные параметры исключим из уравнений (35). Последние приводятся к виду

$$\begin{aligned} (h_{i-2} + h_{i-1})(h_{i-2} + 2h_{i-1})M_{i-1} + h_{i-1}^2 M_i &= 6y_i, \\ (h_{i-2} + h_{i-1})M_{i-1} + (h_{i-1} + h_i)M_i + (h_i + h_{i+1})M_{i+1} &= 0, \\ h_i^2 M_i + (h_i + h_{i+1})(2h_i + h_{i+1})M_{i+1} &= 6y_i. \end{aligned} \quad (37)$$

Получилась система трех уравнений для определения четырех параметров: $y_i, M_{i-1}, M_i, M_{i+1}$. Одним из параметров можно распорядиться по своему усмотрению. Положим

$$y_i = \frac{h_{i-1}(h_{i-2} + h_{i-1})(2h_i + h_{i+1}) + h_i(h_i + h_{i+1})(h_{i-2} + 2h_{i-1})}{(h_{i-1} + h_i)(h_{i-2} + h_{i-1} + h_i)(h_{i-1} + h_i + h_{i+1})} \quad (38)$$

Тогда из уравнений (37) находим

$$\begin{aligned} M_{i-1} &= 6(h_{i-2} + h_{i-1})^{-1}(h_{i-2} + h_{i-1} + h_i)^{-1}, \\ M_i &= -6(h_{i-1} + h_i)^{-1} \left[(h_{i-2} + h_{i-1} + h_i)^{-1} + (h_{i-1} + h_i + h_{i+1})^{-1} \right], \\ M_{i+1} &= 6(h_i + h_{i+1})^{-1}(h_{i-1} + h_i + h_{i+1})^{-1}. \end{aligned} \quad (39)$$

Формулы (36), (38), (39) полностью определяют сплайн $B_i(x)$ на промежутке $[x_{i-2}, x_{i+2}]$. В таблице 4 приведены его значения в узлах x_p , $p = i-2, \dots, i+2$.

Таблица 4.

x	$B_i(x)$	$B_i'(x)$	$B_i''(x)$
x_{i-2}	0	0	0

x_{i-1}	$\frac{h_{i-2}^2}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$	$\frac{3h_{i-2}}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$	$\frac{6}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$
x_i	$1 - \frac{1}{x_{i+1} - x_{i-1}} \cdot \left(\frac{h_i^2}{x_{i+1} - x_{i-2}} + \frac{h_{i-1}^2}{x_{i+2} - x_{i-1}} \right)$	$\frac{3}{x_{i+1} - x_{i-1}} \cdot \left(\frac{h_i}{x_{i+1} - x_{i-2}} - \frac{h_{i-1}}{x_{i+2} - x_{i-1}} \right)$	$-\frac{6}{x_{i+1} - x_{i-1}} \left(\frac{1}{x_{i+1} - x_{i-2}} + \frac{1}{x_{i+2} - x_{i-1}} \right)$
x_{i+1}	$\frac{h_{i+1}^2}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$	$-\frac{3h_{i+1}}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$	$\frac{6}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$
x_{i+2}	0	0	0

Нетрудно показать, что выбор параметра y_i по формуле (38) соответствует тому, что построенный В-сплайн оказывается нормализованным.

О применении В-сплайнов для решения задачи интерполяции

Интерполяционный кубический сплайн $S(x)$ можно находить, используя его представление через В-сплайн:

$$S(x) = \sum_{i=-1}^{N+1} b_i B_i(x) \quad (40)$$

Рассмотрим конкретный пример сплайна, удовлетворяющего условиям (17) и краевым условиям типа I. Для определения коэффициентов b_i получаем систему уравнений:

$$b_{-1}B'_{-1}(x_0) + b_0B'_0(x_0) + b_1B'_1(x_0) = f'_0, \quad (41)$$

$$b_{i-1}B_{i-1}(x_i) + b_iB_i(x_i) + b_{i+1}B_{i+1}(x_i) = f_i, i = 0, \dots, N, \quad (42)$$

$$b_{N-1}B'_{N-1}(x_N) + b_NB'_N(x_N) + b_{N+1}B'_{N+1}(x_N) = f'_N. \quad (43)$$

Элементы матрицы системы вычисляются с помощью таблицы 4. То, что в каждой строке матрицы только три элемента отличны от нуля, объясняется финитностью В-сплайнов. Исключим из полученной системы неизвестные b_{-1} и b_{N+1} . Получим

$$\begin{aligned} b_0[B_0(x_0) - k_0B_{-1}(x_0)] + b_i[B_i(x_0) - k_1B_{-1}(x_0)] &= f_0 - k'_0B_{-1}(x_0), \\ b_{i-1}B_{i-1}(x_i) + b_iB_i(x_i) + b_{i+1}B_{i+1}(x_i) &= f_i, i = 1, \dots, N-1, \\ b_{N-1}[B_{N-1}(x_N) - k_{N-1}B_{N+1}(x_N)] + b_N[B_N(x_N) - k_NB_{N+1}(x_N)] &= f_N - k'_NB_{N+1}(x_N), \end{aligned} \quad (44)$$

где

$$\begin{aligned} k_0 &= \frac{B'_0(x_0)}{B'_{-1}(x_0)}, & k_1 &= \frac{B'_1(x_0)}{B'_{-1}(x_0)}, & k'_0 &= \frac{f'_0}{B'_{-1}(x_0)}, \\ k_{N-1} &= \frac{B'_{N-1}(x_N)}{B'_{N+1}(x_N)}, & k_N &= \frac{B'_N(x_N)}{B'_{N+1}(x_N)}, & k'_N &= \frac{f'_N}{B'_{N+1}(x_N)}, \end{aligned}$$

После нахождения из системы (44) величин b_i , $i = 0, \dots, N, b_{-1}$ и b_{N+1} вычисляется из равенств (41), (43). Разрешимость системы (41)-(43) и подобных ей систем при других граничных условиях следует из единственности интерполяционного сплайна (Теорема 5).

Когда коэффициенты b_i в (40) найдены, для вычислений значений сплайна и его производных можно воспользоваться алгоритмом, описанном в п.3. Однако в некоторых случаях, например, при вычислении сплайнов нескольких точках из промежутка $[x_i, x_{i+1}]$, более эффективным является другой способ. Из (40) следует, что

$$f_p = \frac{1}{6}(b_{p-1} - b_p)B''_{p-1}(x_p)h_p^2 + b_p + \frac{1}{6}(b_{p+1} - b_p)B''_{p+1}(x_p)h_{p-1}^2,$$

$$M_p = (b_{p-1} - b_p)B''_{p-1}(x_p) + (b_{p+1} - b_p)B''_{p+1}(x_p).$$

По этим формулам находим $f_i, f_{i+1}, M_i, M_{i+1}$, что позволяет в дальнейшем использовать формулу (26).

Применение В-сплайнов для вычислений значений интерполяционных кубических сплайнов приводит к увеличению необходимого числа операций по сравнению с алгоритмами из п.7. Но зато количество хранимой информации здесь меньше, а именно: координаты узлов x_i и коэффициенты b_i , то есть всего $2(N+5)$ чисел. При использовании алгоритмов из п.7 нужно хранить $3(N+1)$ величин x_i, y_i, m_i или M_i .

Рассмотрим вопрос о решении системы (44). В общем случае матрица системы не обладает свойством диагонального преобладания. Чтобы оно имело место, необходимо, прежде всего, чтобы

$$r_i = B_i(x_i) - B_{i-1}(x_i) - B_{i+1}(x_i) > 0, \quad i = 1, \dots, N-1.$$

Используя данные таблицы 4, получаем

$$r_i = 1 - \frac{2}{h_{i-1} + h_i} \left(\frac{h_i^2}{h_{i-2} + h_{i-1} + h_i} + \frac{h_{i-1}^2}{h_{i-1} + h_i + h_{i+1}} \right) > 0, \quad i = 1, \dots, N-1.$$

Сетка Δ характеризуется числом $\rho = \max_{|i-j|=1} \frac{h_i}{h_j}$, $\rho \geq 1$. Очевидно, r_i принимает минимальное значение, если

$$h_{i-2} = \rho^{-1}h_{i-1}, \quad h_{i+1} = \rho^{-1}h_i, \text{ которое равно}$$

$$\tilde{r}_i = 1 - 2\rho \frac{(1+\rho)(h_{i-1}^2 + h_i^2) - h_i h_{i-1}}{[(1+\rho)h_{i-1} + \rho h_i][h_{i-1}\rho + (1+\rho)h_i]}.$$

Полагая $h_{i-1} = \rho_i h_i$, где $\rho^{-1} \leq \rho_i \leq \rho$, имеем

$$r_i = 1 - 2\rho \frac{(1+\rho)(1+\rho_i^2) - \rho_i}{[(1+\rho)\rho_i + \rho][\rho\rho_i + 1 + \rho]} = \varphi(\rho_i).$$

Нетрудно вычислить

$$\varphi'(\rho_i) = \frac{2\rho(1+\rho)^2(1+2\rho)(1-\rho_i^2)}{[(1+\rho)\rho_i + \rho]^2[\rho\rho_i + 1 + \rho]^2}.$$

В точке $\rho_i = 1$ функция $\varphi(\rho_i)$ достигает максимума. Поэтому

$$\min_{\rho^{-1} \leq \rho_i \leq \rho} \varphi(\rho_i) = \varphi(\rho) = \varphi(\rho^{-1}) = \frac{\rho(3+\rho-\rho^2)}{(2+\rho)(1+\rho+\rho^2)}.$$

Так как $r_i \geq \varphi(\rho)$, то r_i будет больше нуля, если $3+\rho-\rho^2 > 0$. Отсюда следует ограничение $\rho < \frac{1+\sqrt{13}}{2}$.

Неравенство $r_0 > 0$ и $r_N > 0$ для первого и последнего уравнений системы выполняются, если, например, выбирать дополнительные узлы сетки так, чтобы $h_{-3} = h_{-2} = h_{-1} = h_0$ и $h_{N-1} = h_N = h_{N+1} = h_{N+2}$. Этого же достаточно и в случае граничных условий типа II и IV.

Итак, при условии $\rho < 1 + \sqrt{13}/2$ матрица системы (5) есть матрица с диагональным преобладанием и для решения системы рекомендуется применять метод прогонки. Если же эти условия не выполняются, то следует прибегнуть к методу немонотонной прогонки.

Экстремальное свойство сплайнов

Рассмотрим экстремальные свойства интерполяционных сплайнов, что позволит глубже понять сущность сплайновой интерполяции и естественным образом приведет к практически важной задаче приближения сплайнами со сглаживанием исходных данных.

Исследуем задачу о минимизации функционала

$$J(f) = \int_a^b |f''(x)|^2 dx. \quad (45)$$

В качестве множеств допустимых функций возьмем класс $W_2^2[a, b]$ и его подклассы $\tilde{W}_2^2[a, b]$, $\overline{W}_2^2[a, b]$, где $\tilde{W}_2^2[a, b]$ состоит из функций, периодических с периодом $b - a$, и $\overline{W}_2^2[a, b]$ - из функций $f(x)$, удовлетворяющих условиям $f'(a) = z'_0$, $f'(b) = z'_N$ (z'_0, z'_N - заданные числа).

Пусть в узлах сетки $\Delta: a = x_0 < \dots < x_N = b$ заданы значения $z_i^0, i = 0, \dots, N$.

Теорема 7. Среди всех функций $f(x) \in W_2^2[a, b]$, интерполирующих значения z_i^0 , кубический сплайн $S(x)$ с краевыми условиями

$$S''(a) = S''(b) = 0, \quad (46)$$

минимизирует функционал $J(f)$.

Если функция $f(x)$ принадлежит одному из классов $\tilde{W}_2^2[a, b]$, $\overline{W}_2^2[a, b]$, то минимум функционалу доставляет кубический сплайн из того же множества.

Во всех случаях функция, дающая минимум, единственна.

Доказательство. Пусть $f(x)$ и $S(x)$ принимают одинаковые значения z_i^0 . Запишем тождество

$$J(f - S) = J(f) - J(S) - 2I, \quad (47)$$

$$\text{где } I = \int_a^b [f''(x) - S''(x)]S''(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} [f''(x) - S''(x)]S''(x)dx.$$

Интегрированием по частям получаем

$$\begin{aligned} \int_{x_i}^{x_{i+1}} [f''(x) - S''(x)]S''(x)dx &= \\ [f'(x) - S'(x)]S''(x) \Big|_{x_i}^{x_{i+1}} - \int_{x_i}^{x_{i+1}} [f'(x) - S'(x)]S'''(x)dx &= \\ [f'(x) - S'(x)]S''(x) \Big|_{x_i}^{x_{i+1}} - S'''(x_i +) [f(x) - S(x)] \Big|_{x_i}^{x_{i+1}}. \end{aligned}$$

При выводе этой формулы учтено, что $S'''(x)$ постоянна для $x \in [x_i, x_{i+1}]$. Так как $[f'(x) - S'(x)]S''(x)$ - непрерывная функция, то

$$I = - \sum_{i=0}^{N-1} S'''(x_i +) [f(x) - S(x)] \Big|_{x_i}^{x_{i+1}} + [f'(x) - S'(x)]S''(x) \Big|_a^b.$$

Пусть $S(x)$ удовлетворяет условиям (46). Тогда второй член в выражении I равен нулю. Если $f(x)$, $S(x) \in \tilde{W}_2^2[a, b]$, то это имеет место в силу $f'(a) - S'(a) = f'(b) - S'(b)$ и $S''(a) = S''(b)$. Если $f(x)$, $S(x) \in \overline{W}_2^2[a, b]$, то равенство нулю достигается за счет того, что $f'(a) = S'(a)$ и $f'(b) = S'(b)$. Следовательно, можем записать в неперiodическом случае

$$I = (f_0 - S_0)S'''(x_0 +) + \sum_{i=1}^{N-1} (f_i - S_i)[S'''(x_i +) - S'''(x_{i-1} +)] - (f_N - S_N)S'''(x_{N+1} +), \quad (48)$$

а в периодическом

$$I = \sum_{i=1}^N (f_i - S_i) [S'''(x_i +) - S'''(x_{i-1} +)].$$

Так как $f(x_i) = S(x_i)$, $i = 0, \dots, N$, то $I = 0$ и из (47) вытекает соотношение $J(f) = J(S) + J(f - S)$.

Очевидно, $J(f - S) \geq 0$, и, следовательно, $J(S) \leq J(f)$, то есть сплайн $S(x)$ сообщает минимум функционалу $J(f)$. Функция $S(x)$ единственна. В самом деле, если $f(x)$ - другая минимизирующая функция, то $J(f) = J(S)$ и, значит, $J(f - S) = 0$. Последнее означает, что почти всюду на $[a, b]$ $f''(x) = S''(x)$, откуда $f(x) = S(x) + \alpha x + \beta$. Но постоянно α, β равны нулю, так как $f(x_i) = S(x_i)$, $i = 0, \dots, N$. Иногда установленное экстремальное свойство берут в качестве определения кубического интерполяционного сплайна как функции, интерполирующей заданные значения z_i^0 и минимизирующей функционал $J(f)$. Однако при этом сужается множество рассматриваемых сплайнов, в частности, исключаются сплайны с граничными условиями типов II, IV.

Сглаживание экспериментальных данных

На практике часто приходится иметь дело со случаем, когда значение z_i в узлах сетки Δ заданы с некоторой погрешностью. Например, они могут быть результатами каких-либо измерений. Если погрешность исходных данных относительно велика, то это крайне неблагоприятно влияет на поведение интерполяционного сплайна и особенно его производных. В частности, график сплайна обычно имеет резко выраженные осцилляции. Поэтому возникает вопрос, нельзя ли построить сплайн, проходящий вблизи заданных значений, но более "гладкий", чем интерполяционный (рис.6). Такие сплайны называются сглаживающими, а процедура их построения сглаживанием.

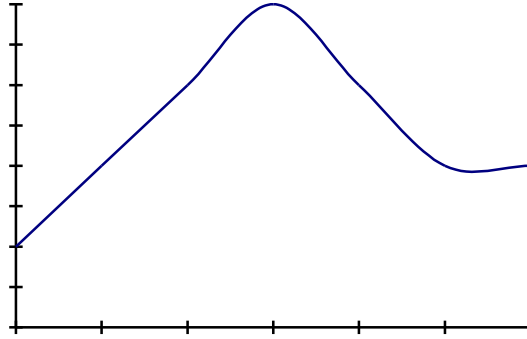


Рис. 6

Покажем, что сглаживающие сплайны возникают при решении задачи о минимизации функционала

$$J_1(f) = \int_a^b |f''(x)|^2 dx + \sum_{i=0}^N \rho_i^{-1} (f_i - z_i^0)^2, \quad (49)$$

где z_i^0 и $\rho_i > 0$ — заданные величины. Очевидно, чем меньше коэффициенты ρ_i , тем ближе проходит функция, минимизирующая функционал $J_1(f)$, к заданным значениям z_i^0 . В качестве множеств допустимых функций будем рассматривать те же классы функций, что и в п.10.

Теорема 8. Среди функций $f(x) \in W_2^2[a, b]$ кубический сплайн $S(x)$, удовлетворяющий условиям (46), минимизирует функционал $J_1(f)$.

Если функция $f(x)$ принадлежит одному из классов $\tilde{W}_2^2[a, b]$ или $\overline{W}_2^2[a, b]$, то минимум функционалу доставляет кубический сплайн из того же класса.

В каждом случае функция, дающая минимум, единственна.

Доказательство. Рассмотрим класс $W_2^2[a, b]$. Выясним вначале необходимые условия минимума. Пусть функция $\varphi(x) \in W_2^2[a, b]$ минимизирует $J_1(f)$. Покажем, что $\varphi(x)$ должна быть кубическим сплайном с краевыми условиями (46). Действительно, предположим, что это не так. Возьмем интерполяционный кубический сплайн $S(x)$ с граничными условиями (46). Второе слагаемое в выражении $J_1(f)$ одинаково для функций $\varphi(x)$ и $S(x)$. Но в силу теоремы 7

$$\int_a^b |S'''(x)|^2 dx < \int_a^b |\varphi'''(x)|^2 dx$$

Поэтому $J_1(S) < J_1(\varphi)$, а это противоречит предположению, что $\varphi(x)$ минимизирует $J_1(f)$.

Итак, пусть $S(x)$ — кубический сплайн, минимизирующий $J_1(f)$. Возьмем сплайн $\tilde{S}(x) = S(x) + \alpha F_k(x)$, где $F_k(x)$, $0 \leq k \leq N$, — фундаментальный кубический сплайн, удовлетворяющий условиям

$$F_k(x_i) = \delta_{ki}, i = 0, \dots, N; F_k''(a) = F_k''(b) = 0.$$

Из (49) следует

$$J_1(\tilde{S}) = \int_a^b |S'''(x) + \alpha F_k''(x)|^2 dx + \sum_{i=0}^N \rho_i^{-1} (S_i - z_i^0)^2 + 2\alpha \rho_k^{-1} (S_k - z_k^0) + \alpha^2 \rho_k^{-1}$$

Тогда

$$J_1(\tilde{S}) - J_1(S) = \alpha^2 a_k + 2\alpha b_k,$$

где

$$a_k = \int_a^b |F_k''(x)|^2 dx + \rho_k^{-1}, \quad b_k = \int_a^b F_k''(x) S'''(x) dx + \rho_k^{-1} (S_k - z_k^0).$$

Здесь $a_k > 0$. Предположим, что $b_k \neq 0$. Выберем α так, чтобы $|\alpha| < 2|b_k|a_k^{-1}$, $\text{sgn } \alpha = -\text{sgn } b_k$. Тогда $J_1(\tilde{S}) - J_1(S) < 0$, то есть $J_1(\tilde{S}) < J_1(S)$. Но это противоречит предположению, что сплайн $S(x)$ минимизирует функционал $J_1(f)$. Следовательно, $b_k = 0$. Так как k — любой индекс от 0 до N , то соотношения

$$b_k = 0, \quad k = 0, \dots, N, \quad (50)$$

суть необходимые условия минимума.

Преобразуем интеграл, входящий в b_k , таким же образом, как интеграл I в п.10. По аналогии с (48) можно записать

$$\int_a^b F_k''(x) S'''(x) dx = F_k(x_0) S'''(x_0 +) + \sum_{i=1}^{N-1} F_k(x_i) [S'''(x_i +) - S'''(x_{i-1} +)] - F_k(x_N) S'''(x_{N-1} +) = D_k.$$

По свойствам функции $F_k(x)$ отсюда следуют формулы

$$D_k = \begin{cases} S'''(x_0 +), & k = 0, \\ S'''(x_k +) - S'''(x_k -), & k = 1, \dots, N-1, \\ -S'''(x_N -), & k = N. \end{cases}$$

Окончательно необходимые условия минимума (50) запишем в виде

$$S_i + \rho_i D_i = z_i^0, \quad i = 0, \dots, N. \quad (51)$$

Покажем далее, что соотношения (51) и краевые условия (46) являются достаточными условиями минимума. Действительно, пусть существует удовлетворяющий им сплайн $S(x)$. Для любой функции $f(x) \in W_2^2[a, b]$ справедливо тождество

$$\tilde{J}_1(f - S) = J_1(f) - J_1(S) - 2 \left[I + \sum_{i=0}^N \rho_i^{-1} (f_i - S_i) (S_i - z_i^0) \right],$$

где

$$\tilde{J}_1(f - S) = \int_a^b [f''(x) - S''(x)]^2 dx + \sum_{i=0}^N \rho_i^{-1} (f_i - S_i)^2,$$

а I определяется формулой (48). Легко видеть, что при выполнении условий (51) выражение в квадратных скобках обращается в нуль. Поэтому

$$J_1(f) = J_1(S) + \tilde{J}_1(f - S).$$

Так как $\tilde{J}_1(f - S) \geq 0$, то сплайн $S(x)$ сообщает минимум функционалу $J_1(f)$.

Нам осталось доказать, что удовлетворяющий условиям (46) и (51) сплайн существует и единственен.

Обозначим $z_i = S(x_i)$, $M_i = S''(x_i)$ и введем векторы и матрицы:

$$z^0 = (z_0^0, \dots, z_N^0)^T, \quad z = (z_0, \dots, z_N)^T, \quad M = (M_0, \dots, M_N)^T;$$

$$A = \begin{bmatrix} 1 & 0 & & & & \\ 0 & u_1 & h_1 & & & 0 \\ & h_1 & u_2 & h_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & h_{N-3} & u_{N-2} & h_{N-2} \\ 0 & & & & h_{N-2} & u_{N-1} & 0 \\ & & & & & 0 & 1 \end{bmatrix},$$

где $u_i = 2(h_{i-1} + h_i)$.

$$H = \begin{bmatrix} 0 & 0 & 0 & & 0 \\ h_0^{-1} & -(h_0^{-1} + h_1^{-1}) & h_1^{-1} & & \\ & \ddots & \ddots & \ddots & \\ & & h_{N-2}^{-1} & -(h_{N-2}^{-1} + h_{N-1}^{-1}) & h_{N-1}^{-1} \\ 0 & & 0 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} \rho_0 & & 0 \\ & \ddots & \\ 0 & & \rho_N \end{bmatrix}.$$

Величины z_i , M_i связаны уравнениями (31). Запишем их в матричной форме:

$$AM = 6Hz \tag{52}$$

Учитывая, что $S'''(x_i) = (M_{i+1} - M_i)/h_i$ и $M_0 = M_N = 0$, представим равенства (51) в виде

$$z + RH^T M = z^0 \tag{53}$$

Равенства (52), (53) образуют систему $2(N+1)$ уравнений для определения параметров сплайна z_i , M_i , $i = 0, \dots, N$. Покажем ее разрешимость.

Из (7) и (8) следует, что

$$(A + 6HRH^T)M = 6Hz^0. \tag{54}$$

Матрица A в системе (52) симметричная с диагональным преобладанием, причем диагональные элементы положительны. Поэтому все её собственные числа по теореме Гершгорина [2] положительны, а сама она положительно определенная. Матрица HRH^T — тоже симметричная. Кроме того, она положительно полуопределенная. В самом деле, если обозначить $\tau = H^T t$, то $t^T (HRH^T) t = \tau^T R \tau \geq 0$, так как матрица R — положительно полуопределенная.

Матрица $A + 6HRH^T$ системы (54) будет также положительно определенной, так как соответствующая ей сумма квадратичных форм положительна. Следовательно, она невырожденная. Тем самым доказано существование и единственность кубического сглаживающего сплайна в классе $W_2^2[a, b]$.

Доказательство утверждения теоремы для классов $\tilde{W}_2^2[a, b]$, $\bar{W}_2^2[a, b]$ проводится аналогично. При этом в классе $\bar{W}_2^2[a, b]$ необходимые условия минимума (51) остаются без изменений, а в периодическом случае следует учесть, что $D_0 = D_N = S'''(x_0 +) - S'''(x_N -)$, и достаточно рассматривать условия (51) при $i = 1, \dots, N$.

Выпишем в развернутом виде системы уравнений, которые необходимо решать при построении сглаживающего сплайна. В непериодическом случае они имеют пятидиагональную структуру

$$\begin{aligned} a_0 M_0 + b_0 M_1 + c_0 M_2 &= g_0, \\ b_0 M_0 + a_1 M_1 + b_1 M_2 + c_1 M_3 &= g_1, \\ c_{i-2} M_{i-2} + b_{i-1} M_{i-1} + a_i M_i + b_i M_{i+1} + c_i M_{i+2} &= g_i, \quad i = 2, \dots, N-2 \\ c_{N-3} M_{N-3} + b_{N-2} M_{N-2} + a_{N-1} M_{N-1} + b_{N-1} M_N &= g_{N-1}, \\ c_{N-2} M_{N-2} + b_{N-1} M_{N-1} + a_N M_N &= g_N. \end{aligned} \quad (55)$$

Коэффициенты этой системы определяются формулами

$$\begin{aligned} a_i &= \frac{1}{3} (h_{i-1} + h_i) + \frac{1}{h_{i-1}^2} \rho_{i-1} + \left(\frac{1}{h_{i-1}} + \frac{1}{h_i} \right)^2 \rho_i + \frac{1}{h_i^2} \rho_{i+1}, \quad i = 1, 2, \dots, N-1, \\ b_i &= \frac{1}{6} h_i - \frac{1}{h_i} \left[\left(\frac{1}{h_{i-1}} + \frac{1}{h_i} \right) \rho_i + \left(\frac{1}{h_i} + \frac{1}{h_{i+1}} \right) \rho_{i+1} \right], \quad i = 1, \dots, N-2, \\ c_i &= \frac{1}{h_i h_{i+1}} \rho_{i+1}, \quad i = 1, \dots, N-3, \\ g_i &= \frac{z_{i+1}^0 - z_i^0}{h_i} - \frac{z_i^0 - z_{i-1}^0}{h_{i-1}}, \quad i = 1, \dots, N-1 \end{aligned} \quad (56)$$

Если сглаживающий сплайн $S(x)$ удовлетворяет условиям $S''(a) = S''(b) = 0$, то $a_0 = a_N = 1$, $b_0 = c_0 = c_{N-2} = b_{N-1} = g_0 = g_N = 0$.

Если $S(x)$ удовлетворяет условиям $S'(a) = z'_0$, $S'(b) = z'_N$, то

$$\begin{aligned} a_0 &= \frac{h_0}{3} + \frac{1}{h_0^2} (\rho_0 + \rho_1), \quad g_0 = \frac{z_1^0 - z_0^0}{h_0} - z'_0, \\ b_0 &= \frac{h_0}{6} - \frac{1}{h_0} \left(\frac{1}{h_0} + \frac{1}{h_1} \right) \rho_1 - \frac{1}{h_0^2} \rho_0, \quad c_0 = \frac{1}{h_0 h_1} \rho_1, \\ a_N &= \frac{h_{N-1}}{3} + \frac{1}{h_{N-1}^2} (\rho_{N-1} + \rho_N), \quad g_N = z'_N - \frac{z_N^0 - z_{N-1}^0}{h_{N-1}}, \\ b_{N-1} &= \frac{h_{N-1}}{6} - \frac{1}{h_{N-1}} \left(\frac{1}{h_{N-1}} + \frac{1}{h_{N-2}} \right) \rho_{N-1} - \frac{1}{h_{N-1}^2} \rho_N, \quad c_{N-2} = \frac{\rho_{N-1}}{h_{N-1} h_{N-2}} \end{aligned}$$

В периодическом случае система состоит из уравнений

$$\begin{aligned} c_{i-2} M_{i-2} + b_{i-1} M_{i-1} + a_i M_i + b_i M_{i+1} + c_i M_{i+2} &= g_i, \\ i &= 1, 2, \dots, N, \end{aligned} \quad (57)$$

где для всех i коэффициенты определяются формулами (56). При этом величины с индексами $N+k$ и k считаются равными: $M_0 = M_N$, $h_0 = h_N$, $a_0 = a_N, \dots$. Для решения систем (55), (57) наиболее подходят методы, основанные на разложении матриц систем в виде LWL^T , где L — нижняя треугольная матрица с единичной диагональю, а W — диагональная матрица с положительными элементами (см. приложение 4). Кроме того, для системы (55) можно использовать метод немонотонной прогонки.

После вычисления M_i из системы (55) или (57) величины z_i определяются соотношениями (51). Перепишем их в виде $z_i - z_i^0 = -\rho_i D_i$, $i = 0, \dots, N$, причем

$$\begin{aligned}
D_0 &= \frac{1}{h_0}(M_1 - M_0), \\
D_i &= \frac{1}{h_i}(M_{i+1} - M_i) - \frac{1}{h_{i-1}}(M_i - M_{i-1}), \quad i = 1, \dots, N-1 \\
D_N &= -\frac{1}{h_{N-1}}(M_N - M_{N-1}).
\end{aligned} \tag{58}$$

В периодическом случае $h_N = h_0$, $M_N = M_0$, $M_{N+1} = M_1$ и все величины D_i формулами (58), где $i = 1, \dots, N$.

Наиболее важным моментом при построении сглаживающего сплайна является выбор весовых множителей ρ_i . Заметим, что если все $\rho_i = 0$, то $z_i = z_i^0$, $i = 0, \dots, N$, и сглаживающий сплайн превращается в интерполяционный. Отсюда следует, что чем точнее заданы z_i^0 в узлах сетки, тем меньше должны быть весовые множители ρ_i . В частности, если возникает необходимость закрепить точку с номером l , то нужно положить $\rho_l = 0$.

По сравнению с интерполяционным сплайном построение сглаживающего сплайна требует значительного большего объема вычислений.

При решении вопроса о том, каким сплайном пользоваться, нужно учитывать в первую очередь погрешность исходных данных, а также те требования, которые предъявляются к сплайну. Практически бесполезно применение сглаживающих сплайнов, когда исходная информация задана, например, с точностью чисел в компьютере. В то же время неизбежно приходится строить сглаживающие сплайны, когда накладываются жесткие ограничения на качественные характеристики сплайна или когда сплайн используется для вычисления производных по экспериментальной информации.

Обобщенные кубические сплайны

Отличной чертой сплайнов, рассматриваемых в данном параграфе, является присутствие в их формулах большого числа свободных параметров. В одних случаях эти параметры используются для изменения уравнений которыми описываются звенья сплайна, в других — для изменения условий стыковки звеньев между собой. Выбор термина “обобщенные кубические сплайны” обусловлен тем, что при определенных значениях параметров “обобщенный” сплайн переходит в обычный кубический. Для каждого из обобщенных сплайнов существуют задачи, где их применение приводит к лучшим результатам по сравнению с кубическими сплайнами. При этом сохраняется одно из важнейших свойств кубического сплайна — простота и эффективность реализации алгоритмов на ЭВМ. Рассмотрим два обобщенных кубических сплайна — рациональный и дискретный.

1. Рациональные сплайны. Пусть на отрезке $[a, b]$ задана сетка $\Delta: a = x_0 < x_1 < \dots < x_N = b$.

Рациональным сплайном называется функция $S_R(x)$, которая:

а) на каждом промежутке $[x_i, x_{i+1}]$ имеет вид

$$S_R(x) = A_i t + B_i(1-t) + \frac{C_i t^3}{1 + p_i(1-t)} + \frac{D_i(1-t)^3}{1 + q_i t}, \tag{59}$$

где $t = (x - x_i)/h_i$, $h_i = x_{i+1} - x_i$, p_i, q_i — заданные числа, $-1 < p_i, q_i < \infty$;

б) $S_R(x) \in C^2[a, b]$.

Таким образом, рациональный сплайн отличается от кубического лишь более сложным аналитическим выражением.

Пусть в узлах сетки Δ заданы значения f_i , $i = 0, \dots, N$. Рациональный сплайн называется интерполяционным, если

$$S_R(x_i) = f_i, \quad i = 0, \dots, N. \tag{60}$$

Как и в случае кубических сплайнов, для построения интерполяционного рационального сплайна необходимы краевые условия:

$$S'_R(x_k) = f'_k, \quad k = 0, N$$

$$S''_R(x_k) = f''_k, \quad k = 0, N$$

Условие периодичности:

$$S_R^{(r)}(x_0) = S_R^{(r)}(x_N), \quad r = 1, 2.$$

Рассмотрим алгоритм построения интерполяционного рационального сплайна.

Из условий интерполяции следует $B_i + D_i = f_i$, $A_i + C_i = f_{i+1}$.

Подставляем отсюда выражения для A_i, B_i в (59), получаем

$$S_R(x) = f_i(1-t) + f_{i+1}t + C_i \left[\frac{t^3}{1+p_i(1-t)} - t \right] + D_i \left[\frac{(1-t)^3}{1+q_i t} - (1-t) \right] \quad (61)$$

Сплайн, записанный в виде (61), непрерывен в узлах сетки Δ и удовлетворяет условиям (60). Определим коэффициенты C_i, D_i так, чтобы были непрерывны его первая и вторая производные.

Так как

$$S'_R(x) = \frac{f_{i+1} - f_i}{h_i} + \frac{C_i}{h_i} \left\{ \frac{3t^2(1+p_i) - 2t^3 p_i}{[1+p_i(1-t)]^2} - 1 \right\} + \frac{D_i}{h_i} \left\{ \frac{-3(1-t)^2(1+q_i) + 2(1-t)^3 q_i}{(1+q_i t)^2} + 1 \right\}, \quad (62)$$

то

$$S'_R(x_i +) = \frac{f_{i+1} - f_i}{h_i} - \frac{(2+q_i)D_i}{h_i} - \frac{C_i}{h_i},$$

$$S'_R(x_i -) = \frac{f_{i+1} - f_i}{h_i} + \frac{(2+p_i)C_i}{h_i} + \frac{D_i}{h_i}.$$

Обозначается $m_i = S'_R(x_i)$, $i = 0, \dots, N$, получаем уравнения

$$C_i + (2+q_i)D_i = (f_{i+1} - f_i) - m_i h_i,$$

$$(2+p_i)C_i + D_i = -(f_{i+1} - f_i) + m_{i+1} h_i.$$

Из них вычисляем

$$C_i = \frac{-(3+q_i)(f_{i+1} - f_i) + h_i m_i + (2+q_i)h_i m_{i+1}}{(2+q_i)(2+p_i) - 1}, \quad (63)$$

$$D_i = \frac{(3+p_i)(f_{i+1} - f_i) - h_i m_{i+1} - (2+p_i)h_i m_i}{(2+q_i)(2+p_i) - 1}.$$

Формулы (63) являются следствием непрерывности $S'_R(x)$.

Далее, из (62)

$$S''_R(x) = C_i \frac{2p_i^2 t^3 - 6p_i(1+p_i)t^2 + 6(1+p_i)^2 t}{[1+p_i(1-t)]^3 h_i^2} + D_i \frac{2q_i^2(1-t)^3 - 6q_i(1+q_i)(1-t)^2 + 6(1+q_i)^2(1-t)}{(1+q_i t)^3 h_i^2}. \quad (64)$$

Отсюда

$$S''_R(x_i +) = \frac{2D_i}{h_i^2} (3 + 3q_i + q_i^2),$$

$$S''_R(x_i -) = \frac{2C_{i-1}}{h_{i-1}^2} (3 + 3p_{i-1} + p_{i-1}^2),$$

и условие непрерывности $S''_R(x)$ в точке x_i имеет вид

$$h_i^2 C_{i-1} (3 + 3p_{i-1} + p_{i-1}^2) = h_{i-1}^2 D_i (3 + 3q_i + q_i^2).$$

Подставляя сюда D_i и C_{i-1} из (63), получаем

$$\begin{aligned} & \lambda_i P_{i-1} m_{i-1} + [\lambda_i P_{i-1} (2+q_{i-1}) + \mu_i Q_i (2+p_i)] m_i + \mu_i Q_i m_{i+1} = \\ & = \lambda_i P_{i-1} (3+q_{i-1}) \frac{f_i - f_{i-1}}{h_{i-1}} + \mu_i Q_i (3+p_i) \frac{f_{i+1} - f_i}{h_i}, \end{aligned} \quad (65)$$

$$i = 1, 2, \dots, N-1,$$

$$\text{где } \lambda_i = h_i(h_{i-1} + h_i)^{-1}, \mu_i = 1 - \lambda_i, P_{i-1} = \frac{3 + 3p_{i-1} + p_{i-1}^2}{(2 + q_{i-1})(2 + p_{i-1}) - 1}, Q_i = \frac{3 + 3q_i + q_i^2}{(2 + q_i)(2 + p_i) - 1}.$$

Обозначим левую и правую части (65) соответственно через $\varphi(m_{i-1}, m_i, m_{i+1})$ и c_i .

К уравнениям (65) необходимо присоединить уравнения, вытекающие из граничных условий. Опуская несложные выкладки, выпишем системы уравнений относительно неизвестных m_i для каждого вида краевых условий.

Типы I и II.

$$\begin{aligned} Q_0(2 + p_0)m_0 + \lambda_0^* Q_0 m_1 &= C_0^*, \\ \varphi(m_{i-1}, m_i, m_{i+1}) &= c_i, \quad i = 1, 2, \dots, N-1, \\ \mu_N^* P_{N-1} m_{N-1} + P_{N-1}(2 + q_{N-1})m_N &= C_N^* \end{aligned}$$

где для типа I

$$\lambda_0^* = \mu_N^* = 0, \quad C_0^* = Q_0(3 + p_0)f_0', \quad C_N^* = P_{N-1}(2 + q_{N-1})f_N';$$

для типа II

$$\lambda_0^* = \mu_N^* = 1, \quad C_0^* = Q_0(3 + p_0)\frac{f_1 - f_0}{h_0} - \frac{h_0}{2}f_0'', \quad C_N^* = P_{N-1}(3 + q_{N-1})\frac{f_N - f_{N-1}}{h_{N-1}} + \frac{h_{N-1}}{2}f_N''.$$

Тип III.

$$\varphi(m_{i-1}, m_i, m_{i+1}) = C_i, \quad i = 1, 2, \dots, N, \quad (66)$$

причем $f_{N+k} = f_k$, $m_{N+k} = m_k$, $\lambda_{N+k} = \lambda_k$, $p_{N+k} = p_k$, $q_{N+k} = q_k$.

Матрицы всех полученных систем с диагональным преобладанием. Это позволяет использовать для их решения метод прогонки и обеспечивает существование и единственность интерполяционного рационального сплайна $S_R(x)$ при любом типе краевых условий. После вычисления значений m_i , $i = 0, 1, \dots, N$, по формулам (63) определяются коэффициенты C_i, D_i . Вычисление сплайна и его первой и второй производных производится по формулам (61), (62), (64).

При $p_i = q_i = 0$, $i = 0, 1, \dots, N-1$, рациональный сплайн превращается в кубический. Действительно, из (59) видно, что в этом случае на каждом промежутке $[x_i, x_{i+1}]$ $S_R(x)$ является кубическим полиномом. Кроме того, приведенные выше системы относительно величин m_i переходят в соответствующие системы для кубического сплайна (гл.7).

Можно считать, что сплайн первой степени также является частным случаем рационального сплайна. Действительно, исследуем поведение $S_R(x)$ на некотором промежутке $[x_k, x_{k+1}]$ в случае, если $p_k, q_k \rightarrow \infty$, $p_i, q_i \geq 0$, $i \neq k$. Сетку Δ считаем фиксированной. Ограничимся рассмотрением периодического случая, так как для других типов граничных условий рассуждения совершенно аналогичны. Используя следствие теоремы приложения 2, из системы (66) имеем

$$\|m\| \leq \max_i \frac{\lambda_i P_{i-1}(3 + q_{i-1}) + \mu_i Q_i(3 + p_i)}{\lambda_i P_{i-1}(1 + q_{i-1}) + \mu_i Q_i(1 + p_i)} \max_i \left| \frac{f_i - f_{i-1}}{h_{i-1}} \right| \leq 3 \max_i \left| \frac{f_i - f_{i-1}}{h_{i-1}} \right| = 3F.$$

Теперь из (63) получаем

$$\begin{aligned} |C_k| &\leq \frac{4Fh_k(3 + q_k)}{(2 + q_k)(2 + p_k) - 1} \xrightarrow{p_k \rightarrow \infty} 0, \\ |D_k| &\leq \frac{4Fh_k(3 + p_k)}{(2 + q_k)(2 + p_k) - 1} \xrightarrow{q_k \rightarrow \infty} 0. \end{aligned}$$

Обращаясь к формуле (61), находим при $x \in [x_k, x_{k+1}]$

$$\lim_{p_k, q_k \rightarrow \infty} S_R(x) = f_k(1 - t) + tf_{k+1}.$$

В частности, если для всех i одновременно $p_i, q_i \rightarrow \infty$, то $S_R(x) \rightarrow S_1(x)$, то есть сплайн первой степени является предельным случаем рационального сплайна.

Рациональные сплайны сочетают в себе свойства наиболее распространенных на практике сплайнов — первой степени и кубических. Кубические сплайны дают, как правило, высокую точность приближения гладких функций. Однако при этом трудно удовлетворить требованиям качественного характера. Например, если интерполируемая функция выпукла, то в ряде случаев необходимо, чтобы сплайн был выпуклым. Кубический сплайн удовлетворяет этому требованию не всегда. В тоже время сплайн первой степени в указанной ситуации будет выпуклым, но зато здесь трудно обеспечить необходимую точность приближения. К тому же сплайн первой степени не является гладкой функцией. Особенно значительные трудности возникают при приближении функции с большими градиентами. Применение в этом случае как кубических сплайнов, так и сплайнов первой степени обычно связано с большим числом узлов интерполяции.

При использовании рациональных сплайнов путем надлежащего выбора свободных параметров p_i, q_i , как правило, всегда удается одновременно удовлетворить требованиям и количественного и качественного характера, в том числе при интерполяции функций с большими градиентами. Это обстоятельство делает рациональные сплайны практически универсальным средством приближения функций.

Выбор параметров p_i, q_i обычно не вызывает особых затруднений. Небольшой практический опыт работы с рациональными сплайнами позволяет, как правило, с “первого взгляда” достаточно точно задавать их значение. В большинстве случаев можно полагать $p_i = q_i$. Конечно, представляет интерес разработка алгоритмов автоматического выбора величин p_i, q_i .

2. Дискретные кубические сплайны. Пусть на отрезке $[a, b]$ задана сетка $\Delta: a = x_0 < x_1 < \dots < x_N = b$, в узлах которой известны значения f_i $i = 0, \dots, N$. Дискретным интерполяционным кубическим сплайном назовем функцию $S_D(x)$, которая:

на каждом из интервалов $[x_i, x_{i+1}]$ является кубическим многочленом

$$S_D(x) \equiv P^i(x) = \sum_{j=0}^3 a_{ij} (x - x_i)^j, \quad x \in [x_i, x_{i+1}];$$

при заданных $\varepsilon_i \geq 0$, $i = 1, \dots, N-1$,

$$\frac{P^i(x_i + \varepsilon_i) - P^i(x_i - \varepsilon_i)}{2\varepsilon_i} = \frac{P^{i-1}(x_i + \varepsilon_i) - P^{i-1}(x_i - \varepsilon_i)}{2\varepsilon_i} \quad (67)$$

$$\frac{P^i(x_i + \varepsilon_i) - 2P^i(x_i) + P^i(x_i - \varepsilon_i))}{\varepsilon_i^2} = \frac{P^{i-1}(x_i + \varepsilon_i) - 2P^{i-1}(x_i) + P^{i-1}(x_i - \varepsilon_i))}{\varepsilon_i^2} \quad (68)$$

$$S_D(x_i) = f_i, \quad i = 0, \dots, N.$$

Условия (67), (68) можно рассматривать как дискретный аналог условий непрерывности первой и второй производных. Этим и объясняется употребление термина “дискретный сплайн”. Очевидно, при $\varepsilon_i \rightarrow 0$ дискретный кубический сплайн переходит в кубический сплайн класса C^2 .

Для сплайна $S_D(x)$ необходимы граничные условия, в качестве которых будем рассматривать условия типов I - IV, сформулированные в гл.7.

Рассмотрим алгоритм построения дискретного сплайна. Введем обозначение

$$M_i = \frac{1}{\varepsilon_i^2} (P^i(x_i - \varepsilon_i) - 2P^i(x_i) + P^i(x_i + \varepsilon_i)), \quad i = 0, \dots, N,$$

$$P^N(x) = P^{N-1}(x).$$

Нетрудно видеть, что в таком случае сплайн $S_D(x)$ при $x \in [x_i, x_{i+1}]$ записывается в виде

$$S_D(x) = f_i(1-t) + f_{i+1}t - \frac{t(1-t)}{6} h_i^2 [(2-t)M_i + (1+t)M_{i+1}], \quad (69)$$

где $h_i = x_{i+1} - x_i$, $t = (x - x_i)/h_i$.

Формула (69) по виду совпадает с формулой (26) для кубического сплайна класса C^2 . Различный смысл имеют лишь величины M_i .

Отметим, что из (69) следуют соотношения

$$S_D''(x_i +) = S_D''(x_i -) = M_i, \quad i = 0, \dots, N. \quad (70)$$

Сплайн $S_D(x)$, записанный в виде (69), удовлетворяет условиям интерполяции и условиям (68).

Подставляя в (67) вычисление из (69) значения сплайна, получаем

$$\mu_i \left(1 - \frac{\varepsilon_i^2}{h_{i-1}^2}\right) M_{i-1} + \left(2 + \frac{\varepsilon_i^2}{h_{i-1} h_i}\right) M_i + \lambda_i \left(1 - \frac{\varepsilon_i^2}{h_i^2}\right) M_{i+1} = \frac{6}{h_{i-1} + h_i} \left(\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}} \right), \quad (71)$$

$$i = 1, \dots, N-1,$$

где, как обычно, $\mu_i = \frac{h_{i-1}}{h_{i-1} + h_i}$, $\lambda_i = 1 - \mu_i$.

К этим соотношениям необходимо добавить уравнения, вытекающие из граничных условий. В периодическом случае добавляется уравнение (71) с индексами $i = N$ и полагается $M_0 = M_N$. В непараметрическом случае, как это следует из (70), граничные уравнения из гл.7. Например, для условий типа II имеем

$$M_0 = f_0'', \quad M_N = f_N''.$$

Обозначим

$$r_i = \left(2 + \frac{\varepsilon_i^2}{h_i h_{i-1}}\right) - \mu_i \left|1 - \frac{\varepsilon_i^2}{h_{i-1}^2}\right| - \lambda_i \left|1 - \frac{\varepsilon_i^2}{h_i^2}\right|.$$

Легко проверить, что

$$r_i \geq 1, \text{ если } \varepsilon_i < \max\{h_{i-1}, h_i\},$$

$$r_i = 3, \text{ если } \varepsilon_i \geq \max\{h_{i-1}, h_i\}.$$

В силу этих соотношений матрицы систем относительно неизвестных M_i будут с диагональным преобладанием при любых ε_i и для всех рассматриваемых типов краевых условий. Тем самым обеспечивая существование и единственность дискретного кубического сплайна при произвольных ε_i .

Теорема 9. Если $S_D(x)$ интерполирует функцию $f(x) \in W_\infty^4[a, b]$ и удовлетворяет краевым условиям одного из типов I, II, III, то

$$\|S_D^{(r)}(x) - f^{(r)}(x)\|_\infty \leq \left\{ K_{0,r} \bar{h}^{4-r} + K_{1,r} \bar{h}^{2-r} \max_i \varepsilon_i^2 \right\} \|f^{(4)}(x)\|_\infty, \quad r = 0, 1, 2, 3,$$

где $K_{0,0} = 17/384$, $K_{0,1} = 1/6$, $K_{0,2} = 3/8$, $K_{0,3} = 1/2 + \beta/2$, $K_{1,2} = 0$, $r = 0, 1, 2, 3$, $\beta = \bar{h} / \min_i h_i$

если $\varepsilon_i < \max\{h_{i-1}, h_i\}$, $i = 1, \dots, N-1$;

$$K_{0,0} = 5/384, \quad K_{0,1} = 1/24, \quad K_{0,2} = 1/8, \quad K_{0,3} = 1/2,$$

$$K_{1,0} = 9/384, \quad K_{1,1} = 1/12, \quad K_{1,2} = 1/6, \quad K_{1,3} = \beta/3,$$

если $\varepsilon_i \geq \max\{h_{i-1}, h_i\}$, $i = 1, \dots, N-1$.

Доказательство данной теоремы опускаем (см.[3]).

Требования к гладкости интерполируемой функции здесь выше, чем для кубических сплайнов C^2 . Очевидно, не следует выбирать параметры ε_i чрезмерно большими. Целесообразно ограничиваться значениями $\varepsilon_i \leq \max\{h_{i-1}, h_i\}$.

Отметим, что для конкретных значений ε_i постоянные в оценках могут быть уменьшены.

Большой практический интерес представляет следующая теорема.

Теорема 10. Если $S_D(x)$ интерполирует функцию $f(x) \in W_\infty^6[a, b]$, удовлетворяет краевым условиям типа II или III и

$$\varepsilon_i^2 = \frac{1}{2} (h_{i-1}^2 - h_{i-1} h_i + h_i^2), \quad i = 1, \dots, N-1,$$

то

$$|M_i - f_i''| \leq \frac{1}{60} \left(\bar{h}^3 \|f^{(v)}(x)\|_{\infty} + \frac{5}{8} \bar{h}^4 \|f^{(vi)}(x)\|_{\infty} \right), i = 0, \dots, N.$$

Если сетка Δ равномерная, то

$$|M_i - f_i''| \leq \frac{1}{160} h^4 \|f^{(vi)}(x)\|_{\infty}, i = 0, \dots, N.$$

Параметрические сплайны

С помощью рассмотренных в предыдущих пунктах сплайнов одной переменной можно приближать лишь такие плоские кривые, которые в выбранной системе координат (не обязательно декартовой) описываются функциональной зависимостью вида $y = f(x)$. Однако не все кривые могут быть представлены подобным образом. Более универсальным способом является параметрическое задание их координат в виде двух функций $x = x(u)$, $y = y(u)$ некоторого параметра u .

При интерполяции кривой, заданной параметрически, естественно ввести разбиение на промежутке измерения параметра u : $u_0 < u_1 < \dots < u_N$, затем вычислить соответствующие значения координат точек на кривой, $x_i = x(u_i)$, $y_i = y(u_i)$, и построить для функций $x(u)$, $y(u)$ интерполяционные сплайны $S(x; u)$, $S(y; u)$. Совокупность этих двух сплайнов называется интерполяционным параметрическим сплайном. В зависимости от вида функций $S(x; u)$, $S(y; u)$ будем говорить о параметрических сплайнах первой степени, эрмитовых параметрических сплайнах и т.д..

В качестве погрешности приближения кривой параметрическим сплайном проще всего взять величину

$$R(u) = \sqrt{|S(x; u) - x(u)|^2 + |S(y; u) - y(u)|^2}.$$

Очевидно, при выполнении определенных требований к гладкости функций $x(u)$, $y(u)$ можно оценить выражения $|S(x; u) - x(u)|$, $|S(y; u) - y(u)|$ по какой-либо норме, например, пространства C . Тогда

$$\|R(u)\|_C \leq \sqrt{\|S(x; u) - x(u)\|_C^2 + \|S(y; u) - y(u)\|_C^2} \quad (72)$$

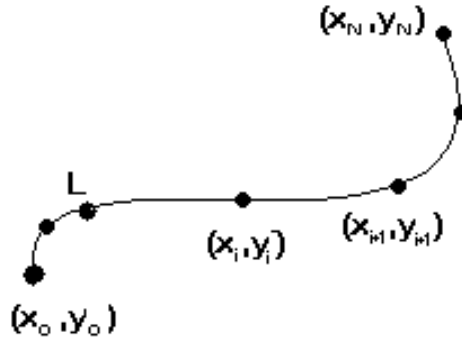


Рис. 7.

Пусть на некоторой кривой L задана последовательность точек $P_i(x_i, y_i)$, $i = 0, 1, \dots, N$ (рис.7). Введем на ней естественную параметризацию $x(s)$, $y(s)$, взяв в качестве параметра длину дуги S кривой, отсчитываемую от точки P_0 . Узлу P_i соответствует значение параметра, равное S_i .

Интерполяционный параметрический сплайн первой степени

на промежутке между точками P_i и P_{i+1} задается соотношениями

$$\begin{aligned} S_1(x; s) &= (1-t)x_i + tx_{i+1} \\ S_1(y; s) &= (1-t)y_i + ty_{i+1} \end{aligned} \quad (73)$$

где $t = (s - s_i)/l_i$, $l_i = s_{i+1} - s_i$.

Геометрически параметрический сплайн первой степени представляет собой ломаную, состоящую из отрезков прямых, соединяющих точки P_i .

Из (73) следует равенство

$$\frac{S'_1(y; s)}{S'_1(x; s)} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad x_i \neq x_{i+1} \quad (74)$$

которое используется для приближенного вычисления наклона касательной к кривой L между точками P_i и P_{i+1} . Если $x_i = x_{i+1}$, то это означает, что данное звено сплайна расположено параллельно оси y .

Отметим интересное свойство формул (73), (74), а именно: координаты точек интерполяционного параметрического сплайна первой степени и угол наклона касательной к нему можно вычислить, не зная фактически ничего о длине дуги интерполируемой кривой. В самом деле, по формулам (73) эти точки вычисляются путем задания величины $t \in [0, 1]$, а в формуле (74) правая часть вообще не зависит от s .

Обозначим через $R_1(s)$ погрешность интерполяции параметрическим сплайном первой степени. Оценить ее можно по формуле (72), учитывая оценки, установленные в гл.4.

Отметим два результата, вытекающие из факта использования естественной параметризации кривой.

Теорема 10. Если $x(s), y(s) \in W^1_\infty[S_0, S_N]$, то

$$\|R_1(s)\|_c \leq \sqrt{2} \bar{l} / 2 \quad (75)$$

где $\bar{l} = \max_i l_i$.

Доказательство. По теореме 1

$$\begin{aligned} \|S_1(x; s) - x(s)\|_c &\leq \frac{1}{2} \bar{l} \|x'(s)\|_\infty, \\ \|S_1(y; s) - y(s)\|_c &\leq \frac{1}{2} \bar{l} \|y'(s)\|_\infty. \end{aligned}$$

Так как величины $x'(s)$, $y'(s)$ представляют собой направляющие косинусы вектора касательной к кривой в точке $(x(s), y(s))$, то $\|x'(s)\|_c \leq 1$, $\|y'(s)\|_c \leq 1$ и из (72) следует (75).

Теорема 11. Если $x(s), y(s) \in CW^2_{\Delta, \infty}[S_0, S_N]$, то

$$\|R_1(s)\|_c \leq \frac{\sqrt{2}}{8} \bar{l}^2 \|K(s)\|_\infty \quad (76)$$

где $K(s)$ — кривизна кривой L .

Доказательство. Согласно теореме 1 имеем

$$\begin{aligned} \|S_1(x; s) - x(s)\|_c &\leq \frac{1}{8} \bar{l}^2 \|x''(s)\|_\infty, \\ \|S_1(y; s) - y(s)\|_c &\leq \frac{1}{8} \bar{l}^2 \|y''(s)\|_\infty. \end{aligned}$$

Кривизна кривой в точке $(x(s), y(s))$ определяется равенством $K(s) = \sqrt{[x''(s)]^2 + [y''(s)]^2}$, и, значит

$$\|x''(s)\|_\infty \leq \|K(s)\|_\infty, \quad \|y''(s)\|_\infty \leq \|K(s)\|_\infty.$$

Для завершения доказательства теперь достаточно обратиться к формуле (72).

Практическое применение оценок (75), (76) затруднено присутствием в них величины \bar{l} . Если обозначить через d_i длину хорды, стягивающей концы дуги l_i , то при достаточно густом расположении узлов интерполяции на кривой L имеем $d_i \approx l_i$ и вместо \bar{l} можно использовать реально вычисляемую величину $\bar{d} = \max_i d_i$.

Параметрическим эрмитовым кубическим сплайном

будем называть совокупность двух эрмитовых сплайнов $S_{3,2}(x; s)$, $S_{3,2}(y; s)$.

Погрешность интерполяции $R_3(s)$ для такого сплайна можно оценить обычным способом с помощью формулы (72), используя результаты гл.5. В частности имеет место

Теорема 12. Если $x(s), y(s) \in W_\infty^2[S_0, S_N]$, то

$$\|R_3(s)\|_c \leq \frac{\sqrt{2}}{16} \bar{l}^2 \|K(s)\|_\infty.$$

Если $x(s), y(s) \in C^1 C_\Delta^2[S_0, S_N]$, то

$$\|R_3(s)\|_c \leq \frac{1}{32} \bar{l}^2 (\omega(x'') + \omega(y'')).$$

Доказательство очевидно.

Отметим, что максимальный порядок приближения параметрическим эрмитовым сплайном есть $O(\bar{l}^4)$. Этот порядок достигается, если $x(s), y(s) \in C^1 W_{\Delta, \infty}^4[S_0, S_N]$.

На практике применение параметрических эрмитовых сплайнов, построенных на основе естественной параметризации кривой, обычно оказывается невозможным. Действительно, чтобы вычислять, например значения сплайна

$$S_{3,2}(x; s) = x_i(1-t)^2(1+2t) + x_{i+1}t^2(3-2t) + x'_i t(1-t)^2 l_i - x'_{i+1} t^2(1-t) l_i$$

(здесь x'_i означает производную по параметру S в точке S_i), необходимо значение величин l_i , x'_i , x'_{i+1} . Такая информация, как правило, отсутствует. Поэтому вместо сплайна $S_{3,2}(x; s)$ будем строить в некотором смысле близкий к нему сплайн $\tilde{S}_{3,2}(x; \tilde{s})$.

Во-первых, точные значения производных x'_i заменим значениями \tilde{x}'_i по формулам

$$\tilde{x}'_0 = (1 + \tilde{\mu}_1)(x_1 - x_0)d_0^{-1} - \tilde{\mu}_1(x_2 - x_1)d_1^{-1}, \quad (77)$$

$$\tilde{x}'_i = \tilde{\mu}_i(x_{i+1} - x_i)d_i^{-1} + \tilde{\lambda}_i(x_i - x_{i-1})d_{i-1}^{-1}, \quad i = 1, \dots, N-1 \quad (78)$$

$$\tilde{x}'_N = (1 + \tilde{\lambda}_{N-1})(x_N - x_{N-1})d_{N-1}^{-1} - \tilde{\lambda}_{N-1}(x_{N-1} - x_{N-2})d_{N-2}^{-1} \quad (79)$$

$$\tilde{\mu}_i = d_{i-1}(d_{i-1} + d_i)^{-1}, \quad \tilde{\lambda}_i = 1 - \tilde{\mu}_i$$

Если кривая замкнута, то есть точки P_0 и P_N совпадают, вместо (77), (79) следует использовать формулу

$$\tilde{x}'_0 = \tilde{x}'_N = \frac{d_{N-1}}{d_0 + d_{N-1}} \cdot \frac{x_1 - x_0}{d_0} + \frac{d_0}{d_0 + d_{N-1}} \cdot \frac{x_N - x_{N-1}}{d_{N-1}}.$$

Во вторых, для описания сплайна введем параметризацию по суммарной длине хорд d_i . Если обозначить новый параметр через \tilde{S} , то сетка узлов будет $\tilde{\Delta}: 0 = \tilde{S}_0 < \tilde{S}_1 < \dots < \tilde{S}_N$, где

$$\tilde{S}_i = \sum_{k=0}^{i-1} d_k, \quad d_k = \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2}.$$

Параметр \tilde{S} при этом пробегает отрезок $[0, \tilde{S}_N]$.

Тогда Эрмитов кубический сплайн принимает вид

$$\tilde{S}_{3,2}(x; \tilde{s}) = x_i(1-t)^2(1+2t) + x_{i+1}t^2(3-2t) + \tilde{x}'_i t(1-t)^2 d_i - \tilde{x}'_{i+1} t^2(1-t) d_i, \quad (80)$$

где $t = (\tilde{S} - \tilde{S}_i)/d_i$. Все величины, входящие выражение для $\tilde{S}_{3,2}(x; \tilde{S})$, реально вычислимы. Аналогичные формулы получаются и для $\tilde{S}_{3,2}(y; \tilde{S})$.

Нетрудно видеть, что параметрический Эрмитов сплайн не меняется при переходе к новому параметру $\bar{S} = \gamma \tilde{S}$, где γ произвольное положительное число.

Действительно, в этом случае значения \tilde{S}_i переходят $\gamma\tilde{S}_i$, величины $d_i = \tilde{S}_{i+1} - \tilde{S}_i$ — в γd_i и из (77) — (79) следует, что \tilde{x}_i' переходят в $\gamma^{-1}\tilde{x}_i'$. Обращаясь к формуле (80), убеждаемся, что x - координаты точек параметрического сплайна не изменятся. То же самое справедливо и для y - координат. В некоторых случаях удобно полагать $\gamma = \tilde{S}_N^{-1}$. При этом $\bar{S} \in [0,1]$. Будем называть такую параметризацию нормированной по суммарной длине хорд.

Рассмотрим вопрос об оценке погрешности интерполяции

$$\|R(t)\|_C = \max_i \|R'(t)\|_C, \quad (81)$$

где величина $\|R'(t)\|_C = \sqrt{\|\tilde{S}_{3,2}(x; \tilde{S}) - x(S)\|_C^2 + \|\tilde{S}_{3,2}(y; \tilde{S}) - y(S)\|_C^2}$ характеризует отклонение сплайна от кривой на промежутке между точками P_i и P_{i+1} . Здесь $S = S_i + l_i t$, $\tilde{S} = \tilde{S}_i + d_i t$, и сравнение сплайнов и функций $x(S)$, $y(S)$ делается при одинаковых значениях промежуточного параметра t .

Теорема 13. Если $x(s), y(s) \in W_\infty^3[S_0, S_N]$, то

$$\|R(t)\|_C = O(\bar{l}^3).$$

Доказательство опускаем (см. [3]).

Порядок погрешности $O(\bar{l}^3)$, гарантируемый теоремой 13, наивысший. Дальнейшее повышение гладкости функций $x(S)$, $y(S)$ не изменяет его.

Интерполяционный параметрический кубический сплайн

есть совокупность двух кубических сплайнов $S(x; u)$, $S(y; u)$ класса C^2 , интерполирующих соответственно координаты x_i , y_i точек P_i , $i = 0, \dots, N$, кривой L на сетке $\Delta: u_0 < u_1 < \dots < u_N$. Для однозначного определения каждого из этих сплайнов должны быть заданы краевые условия. Будем полагать, что для обоих сплайнов они одинакового типа. Если кривая незамкнутая, то удобно применить краевые условия типа IV, а для интерполяции замкнутой кривой естественно использовать периодические краевые условия (тип III).

В качестве параметра снова берем суммарную длину хорд \tilde{S} . В этом случае, согласно формуле (26), сплайн $S(x; \tilde{S})$ может быть записан в виде

$$S(x; \tilde{S}) = x_i(1-t) + x_{i+1}t - \frac{t(1-t)}{6} d_i^2 [(2-t)\tilde{M}_i + (1+t)\tilde{M}_{i+1}],$$

где $t = (\tilde{S} - \tilde{S}_i)/d_i$, $d_i = \tilde{S}_{i+1} - \tilde{S}_i$.

Величины \tilde{M}_i определяется из систем уравнений (31)-(33), в которых h_i заменяется на d_i . Например, для периодических сплайнов система имеет форму

$$\tilde{\mu}_i \tilde{M}_{i-1} + 2\tilde{M}_i + \tilde{\lambda}_i \tilde{M}_{i+1} = 3\tilde{E}_i, \quad i = 1, \dots, N,$$

где $\tilde{\mu}_i = d_{i-1}(d_{i-1} + d_i)^{-1}$, $\tilde{\lambda}_i = 1 - \tilde{\mu}_i$,

$$\tilde{E}_i = \frac{2}{d_{i-1} + d_i} \left(\frac{x_{i+1} - x_i}{d_i} - \frac{x_i - x_{i-1}}{d_{i-1}} \right).$$

Аналогичные формулы имеют место и для сплайна $S(y; \tilde{S})$.

Погрешность интерполяции параметрическим кубическим сплайном определим формулой (81), где положим

$$\|R'(t)\|_C = \sqrt{\|S(x; \tilde{S}) - x(S)\|_C^2 + \|S(y; \tilde{S}) - y(S)\|_C^2}.$$

Сравнение сплайнов и интерполируемых функций делается при одинаковых значениях параметра t . Сформулируем теорему об оценке погрешности интерполяции в случае замкнутой кривой L .

Наряду со сплайном $S(x; \tilde{S})$ введем сплайн $S(x; S)$, построенный с использованием параметризации по длине дуги:

$$S(x; S) = x_i(1-t) + tx_{i+1} - \frac{t(1-t)}{6} l_i^2 [(2-t)M_i + (1+t)M_{i+1}].$$

Здесь M_i вычисляются из системы

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = 3E_i, \quad i = 1, \dots, N,$$

где $\mu_i = l_{i-1}/(l_{i-1} + l_i)^{-1}$, $\lambda_i = 1 - \mu_i$,

$$E_i = \frac{2}{l_{i-1} + l_i} \left(\frac{x_{i+1} - x_i}{l_i} - \frac{x_i - x_{i-1}}{l_{i-1}} \right).$$

Теорема 14. Если $x(S), y(S) \in C^2 W_{\Delta, \infty}^4 [S_0, S_N]$, то

$$\|R(t)\|_C = O(\tilde{t}^3).$$

Если, кроме того, выполняется условие

$$l_{i-1} = l_i + O(l_i^2), \quad i = 0, \dots, N-1, \quad (82)$$

То

$$\|R(t)\|_C = O(\tilde{t}^4).$$

Доказательство теоремы см. в [1].

Как видно, наивысший порядок погрешности $O(\tilde{t}^4)$ обеспечивается при условии (82), то есть когда соседние звенья кривой отличаются на малые второго порядка. На практике более полезны другие условия, равносильные (82), а именно

$$d_{i-1} = d_i + O(d_i^2), \quad i = 1, \dots, N-1.$$

В частности, при малых l_i можно выбрать $d_i = d = \text{const}$ и тем самым добиться наибольшей точности приближения.

Среди всех сплайнов, используемых при аппроксимации кривых, наиболее универсальными свойствами обладают **параметрические рациональные сплайны**. Они представляют собой совокупность пар рациональных сплайнов $S_R(x; \tilde{S})$, $S_R(y; \tilde{S})$. В качестве параметра будем брать по-прежнему суммарную длину хорд. Согласно (59) каждый из сплайнов $S_R(x; \tilde{S})$ и $S_R(y; \tilde{S})$ представляется на участке между точками P_i, P_{i+1} в виде

$$S_R(x; \tilde{S}) = A_i t + B_i(1-t) + \frac{C_i t^3}{1 + p_i(1-t)} + \frac{D_i(1-t)^3}{1 + q_i t},$$

где p_i, q_i — заданные числа $(-1 < p_i, q_i < \infty)$.

Величины p_i, q_i целесообразно выбирать одинаковыми для обоих сплайнов, при этом обычно удается при относительно малом числе узлов интерполяции обеспечить и высокую точность приближения, и хорошие качественные характеристики. Например, выбирая их достаточно большими, можно полностью устранить осцилляции.

Параметрические кубические и рациональные сплайны не изменяются при переходе к нормированной параметризации по суммарной длине хорд.

В заключение отметим, что построение интерполяционных параметрических сплайнов для пространственных кривых осуществляется аналогичным образом. Разница лишь в том, что приходится оперировать с совокупностью трех сплайнов одной переменной.

Численное дифференцирование

Самый простой способ приближенного вычисления производных функций $f(x)$ состоит в замене их производными интерполяционного сплайна, построенного по значениям $f_i = f(x_i)$, $i = 0, \dots, N$, заданным на сетке $\Delta: a = x_0 < x_1 < \dots < x_N = b$. С этой целью можно использовать различные кубические сплайны, что позволяет вычислять производные до третьего порядка включительно. Например, из выражений (19), (26) вытекает следующие формулы численного дифференцирования.

$$\begin{aligned}
S'(x) &= 6t(1-t) \frac{f_{i+1} - f_i}{h_i} + (1-4t+3t^2)m_i - (2t-3t^2)m_{i+1}, \\
S''(x) &= \frac{1}{h_i} \left[6(1-2t) \frac{f_{i+1} - f_i}{h_i} - (4-6t)m_i - (2-6t)m_{i+1} \right], \\
S'''(x) &= \frac{6}{h_i^2} \left(m_{i+1} + m_i - 2 \frac{f_{i+1} - f_i}{h_i} \right), \\
S'(x) &= \frac{f_{i+1} - f_i}{h_i} - \frac{h_i}{6} \left[(2-6t+3t^2)M_i + (1-3t^2)M_{i+1} \right], \\
S''(x) &= M_i(1-t) + M_{i+1}t, \\
S'''(x) &= \frac{M_{i+1} - M_i}{h_i}.
\end{aligned}$$

Напомним, что алгоритм вычисления величин m_i и M_i рассмотрены в гл.7.

Если использовать представление сплайна $S(x)$ через кубические В-сплайны

$$S(x) = \sum_{i=-1}^{N+1} b_i B_i(x),$$

то, очевидно,

$$S^{(r)}(x) = \sum_{i=-1}^{N+1} b_i B_i^{(r)}(x), \quad r = 1, 2, 3,$$

и дело сводится к вычислению производных от В-сплайнов.

Асимптотические формулы для кубических сплайнов класса C^2 .

Пусть кубический сплайн $S(x)$ интерполирует функцию $f(x)$ на равномерной сетке $\Delta: x_i = a + ih, i = 0, \dots, N, x_N = b$. В дальнейшем не будем оговаривать требования к гладкости $f(x)$, считая, что она имеет столько непрерывных производных, сколько нужно для выполнения преобразований. Вначале рассмотрим случай, когда $f(x)$ — периодическая функция с периодом $b - a$.

Для величин M_i имеем систему

$$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2} (f_{i-1} - 2f_i + f_{i+1}), \quad i = 1, \dots, N, \quad (83)$$

где полагается $M_k = M_{N+k}, f_k = f_{N+k}, k = 0, 1$.

Будем искать решение системы (83) в виде

$$M_i = f_i'' + \alpha_i h^2 f_i^{IV}. \quad (84)$$

Подставляем (84) в (83) и разлагая обе части i -го уравнения по формуле Тейлора в точке x_i , находим

$$6f_i'' + h^2 f_i^{IV} + (\alpha_{i-1} + 4\alpha_i + \alpha_{i+1})h^2 f_i^{IV} = f_i'' + \frac{h^2}{2} f_i^{IV} + O(h^4), \quad i = 1, \dots, N.$$

Сравнивая коэффициенты при одинаковых по порядку производных, получаем систему

$$\begin{aligned}
4\alpha_1 + \alpha_2 + \alpha_N &= -\frac{1}{2}, \\
\alpha_{i-1} + 4\alpha_i + \alpha_{i+1} &= -\frac{1}{2}, \quad i = 2, \dots, N-1, \\
\alpha_1 + \alpha_{N-1} + 4\alpha_N &= -\frac{1}{2},
\end{aligned}$$

которая имеет единственное решение $\alpha_i = -\frac{1}{12}, i = 1, \dots, N$.

При таких α_i величин M_i , определяемые формулой (84) удовлетворяют системе (83) с точностью $O(h^4)$. В силу теоремы 2 приложения норма матрицы, обратной к матрице этой системы, не превышает $\frac{1}{2}$, и поэтому внесение в правую часть погрешности порядка $O(h^4)$ приводит к изменению решения на величину того же порядка. Следовательно,

$$M_i = f_i'' - \frac{1}{2} h^2 f_i^{IV} + O(h^4), \quad i = 0, \dots, N. \quad (85)$$

Точно так же можно получить выражения для M_i , содержащие производные функций $f(x)$ более высоких порядков. Приведем одно из них

$$M_i = f_i'' - \frac{h^2}{12} f_i^{IV} + \frac{h^4}{360} f_i^{VI} + \frac{17h^6}{60480} f_i^{VIII} + O(h^8). \quad (86)$$

Используя (85), (26) и разложение Тейлора в точке $x = x_i + th_i$, находим

$$S(x) = f(x) - \frac{u^2 h^4}{24} f^{IV}(x) - \frac{(1-2t)(1+3u)}{180} h^5 f^V(x) + O(h^6),$$

где обозначено $u = t(1-t)$.

В непериодическом случае все полученные выводы сохраняют силу, если в качестве граничных условий задать величины M_0 и M_N по формулам (86).

Кроме того, формула (86) верна при любых граничных условиях, если только рассматривается интервал $[x_i, x_{i+1}]$, расположенный достаточно далеко от концов отрезка $[a, b]$.

Отметим, что асимптотические разложения для величин M_i можно получить и для неравномерных сеток. Но в этом случае коэффициенты при производных, в отличие от (86), будут зависеть от сетки и их вычисления, за исключением некоторых частных случаев, требует применения компьютера.

Получим разложение для коэффициентов b_i в интерполяционной формуле (40), что позволит выяснить их смысл. Учитывая выражения для узловых значений нормализованных В-сплайнов (таблица 4), в периодическом случае на равномерной сетке приходим к следующей системе относительно неизвестных b_i :

$$b_{i-1} + 4b_i + b_{i+1} = 6f_i, \quad i = 1, \dots, N.$$

Отыскивая ее решение в виде $b_i = f_i + \gamma_i h^2 f_i''$, получаем $\gamma_i = -\frac{1}{6}$, $i = 1, \dots, N$, и

$$b_i = f_i - \frac{h^2}{6} f_i'' + O(h^4), \quad i = 0, \dots, N. \quad (87)$$

Следовательно, на равномерной сетке коэффициент b_i с точностью $O(h^2)$ равен узловому значению f_i .

Если в формуле (40) заменить b_i на \tilde{b}_i , где $|\tilde{b}_i - b_i| = O(h^4)$, то в силу свойств нормализованных В-сплайнов сплайн

$$\tilde{S}(x) = \sum_{i=-1}^{N+1} \tilde{b}_i B_i(x)$$

будет удовлетворять условию $|\tilde{S}(x) - S(x)| = O(h^4)$.

Одно из приложений формулы (87) состоит в том, что на ее основе можно указать эффективный способ определения таких \tilde{b}_i , не требующий решения системы уравнений.

Действительно,

$$f_i'' = \frac{(f_{i-1} - 2f_i + f_{i+1}))}{h^2} + O(h^2).$$

Подставляя это выражение в (87), получаем

$$b_i = -\frac{1}{6} f_{i-1} + \frac{4}{3} f_i - \frac{1}{6} f_{i+1} + O(h^4),$$

то есть можно положить

$$\tilde{b}_i = \frac{1}{6} (-f_{i-1} + 8f_i - f_{i+1}), \quad i = 0, \dots, N.$$

В итоге получены явные формулы для $\tilde{S}(x)$. Естественно, этот сплайн не интерполирует заданные значения f_i , но по порядку обеспечивает такую же точность приближения, что и интерполяционный. Такие сплайны называются локально аппроксимационными.

На основе полученных асимптотических разложений получается ряд интересных формул. Возьмем разностную аппроксимацию второй производной

$$\frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} = f_i'' + \frac{h^2}{12} f_i^{IV} + O(h^4).$$

Комбинируя ее со сплайновой аппроксимацией, в силу (85) получаем

$$\frac{1}{2} \left(M_i + \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} \right) = f_i'' + O(h^4). \quad (88)$$

Таким образом, среднее арифметическое сплайновой и разностной аппроксимаций второй производной дает точность $O(h^4)$, хотя в отдельности каждая из них имеет точность $O(h^2)$.

Учитывая соотношения (83), формулу (88) можно записать в другом виде:

$$\frac{1}{12} (M_{i-1} + 10M_i + M_{i+1}) = f_i'' + O(h^4).$$

Пожалуй, наиболее неожиданным является следующий результат:

$$\frac{M_{i-1} - 2M_i + M_{i+1}}{h^2} = f_i^{IV} + O(h^4),$$

который является следствием разложения (86).

Несмотря на то, что $S^{IV}(x) = 0$ почти всюду на $[a, b]$ и, казалось бы, не имеет смысла составлять вторую разнесенную разность от величин M_i , мы получаем аппроксимацию четвертой производной функции $f(x)$ с очень высокой точностью.

Численное интегрирование

Наиболее простой способ получения формул численного интегрирования для интеграла

$$\int_a^b f(x) dx \quad (89)$$

основан на применении аппарата интерполирования. При этом $f(x)$ заменяется некоторым интерполяционным сплайном $S(x)$ и в качестве приближенного значения интеграла (89) берется величина

$$\int_a^b S(x) dx \quad (90)$$

Погрешность вычисления интеграла (89) можно оценить, например, следующим образом:

$$\left| \int_a^b f(x) dx - \int_a^b S(x) dx \right| \leq \int_a^b |S(x) - f(x)| dx \leq (b-a) \|S(x) - f(x)\|_C.$$

Следовательно, достаточно иметь оценку погрешности приближения функции $f(x)$ сплайном $S(x)$.

Получим формулы вычисления интеграла (90) для конкретных сплайнов. Рассмотрим сплайн первой степени $S_1(x)$ и различные виды кубических сплайнов.

Пусть на $[a, b]$ задана сетка $\Delta: a = x_0 < \dots < x_N = b$. Сплайн $S_1(x)$ на промежутке $[x_i, x_{i+1}]$ записывается в виде

$$S_1(x) = f_i(1-t) + f_{i+1}t.$$

Поэтому

$$\int_a^b S_1(x) dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} S_1(x) dx = \sum_{i=0}^{N-1} h_i \int_0^1 [f_i(1-t) + f_{i+1}t] dt.$$

Отсюда

$$\int_a^b S_1(x)dx = \frac{1}{2} \sum_{i=0}^{N-1} h_i (f_i + f_{i+1}). \quad (91)$$

Формула (91) есть не что иное, как всем известная составная (обобщенная) формула трапеций. Особенно часто она используется на равномерной сетке:

$$\int_a^b S_1(x)dx = \frac{h}{2} f_0 + h \sum_{i=1}^{N-1} f_i + \frac{h}{2} f_N. \quad (92)$$

Аналогичным образом выводятся формулы интегрирования для кубических сплайнов. Если $S(x)$ — кубический сплайн класса C^2 , то при использовании представления (19) получаем

$$\int_a^b S(x)dx = \frac{1}{2} \sum_{i=0}^{N-1} h_i (f_i + f_{i+1}) + \frac{1}{12} \sum_{i=0}^{N-1} (m_i - m_{i+1}) h_i^2.$$

На равномерной сетке вторая сумма в правой части упрощается и формула приобретает вид

$$\int_a^b S(x)dx = \frac{h}{2} f_0 + h \sum_{i=1}^{N-1} f_i + \frac{h}{2} f_N + \frac{h^2}{12} (m_0 - m_N) \quad (93)$$

Фактически применение этой формулы не требует предварительного построения сплайна. Например, если $S(x)$ удовлетворяет краевым условиям типа I, то $m_0 = f'_0$, $m_N = f'_N$. В других случаях можно заменить m_0 и m_N разностными аппроксимациями.

Совсем просто выглядит формула (93), когда $f(x)$ — периодическая функция с периодом $b-a$. В этом случае $f_0 = f_N$, $m_0 = m_N$ и

$$\int_a^b S(x)dx = h \sum_{i=0}^{N-1} f_i.$$

Эта формула совпадает с формулой трапеций (92). Полученный результат является частным случаем утверждения о том, что использование сплайнов произвольной степени дефекта 1 для численного интегрирования периодических функций на равномерной сетке приводит к формуле трапеций. Довольно просто это утверждение доказывается, если использовать аппарат В-сплайнов.

Если для $S(x)$ используется представление (26), то

$$\int_a^b S(x)dx = \frac{1}{2} \sum_{i=0}^{N-1} h_i (f_i + f_{i+1}) - \frac{1}{24} \sum_{i=0}^{N-1} h_i^3 (M_i + M_{i+1}) \quad (94)$$

Аналог формулы (94) для равномерной сетки проще всего получить из (93), используя соотношения

$$m_0 = \frac{f_1 - f_0}{h} - \frac{h}{6} (2M_0 + M_1), \quad m_N = \frac{f_N - f_{N-1}}{h} - \frac{h}{6} (M_{N-1} + 2M_N).$$

Находим

$$\int_a^b S(x)dx = \frac{5h}{12} (f_0 + f_N) + \frac{13h}{12} (f_1 + f_{N-1}) + h \sum_{i=2}^{N-2} f_i - \frac{h^3}{72} (2M_0 + M_1 + M_{N-1} + 2M_N).$$

Использование В-сплайнов при численном интегрировании приводит к более громоздким формулам по сравнению с приведенными выше.

Для сплайна S_D формулы интегрирования совпадают с формулами (93), (94). Если используется эрмитов сплайн $S_{3,2}(x)$, то в формулах нужно заменить величины m_i на f'_i , $i = 0, \dots, N$.

Приложение.

Некоторые нормированные пространства и классы.

Рассмотрим наиболее распространенные в вычислительной математике нормированные пространства и классы.

Прежде всего, это пространство $C[a, b]$ — непрерывных на $[a, b]$ функций с нормой

$$\|f(x)\|_{C[a,b]} = \max_{x \in [a,b]} |f(x)|.$$

На сетке $\Delta: a = x_0 < x_1 < \dots < x_N = b$ эти функции характеризуются их колебаниями на струнах $[x_i, x_{i+1}]$

$$\omega(f) = \max_{0 \leq i \leq N-1} \omega_i(f).$$

Характеристикой функции, не зависящей от сетки Δ , является модуль непрерывности

$$\omega(f; h) = \max_{\substack{x', x'' \in [a,b] \\ |x'' - x'| \leq h}} |f(x'') - f(x')|, \quad h \leq b - a.$$

Если обозначить $\bar{h} = \max_i h_i$, то очевидны неравенства

$$\omega_i(f) \leq \omega(f) \leq \omega(f; \bar{h}).$$

Теорема о среднем для непрерывных функций: если $f(x) \in C[a, b]$ и величины α, β имеют одинаковые знаки, то

$$\alpha f(a) + \beta f(b) = (\alpha + \beta) f(\zeta), \quad a \leq \zeta \leq b.$$

Теорема о среднем для интегралов

$$\int_a^b f(x) g(x) dx = f(\zeta) \int_a^b g(x) dx, \quad a \leq \zeta \leq b,$$

имеет место, если $f(x)$ и $g(x)$ непрерывны на $[a, b]$ и $g(x)$, кроме того, знакопостоянна.

Формула Тейлора с остаточным членом в форме Лагранжа для функций, имеющих на $[a, b]$ непрерывную производную k -го порядка, имеет вид

$$f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \dots + \frac{f^{(k-1)}(a)(x-a)^{k-1}}{(k-1)!} + \frac{f^{(k)}(\zeta)(x-a)^k}{k!},$$

где ζ — некоторая точка из промежутка $[a, x]$.

$L_p[a, b]$ — пространство измеримых на $[a, b]$ функций $f(x)$, для которых функция $|f(x)|^p$ интегрируема по Лебегу на $[a, b]$. Норма в этом пространстве есть

$$\|f(x)\|_{L_p[a,b]} = \left(\int_a^b |f(x)|^p dx \right)^{1/p}.$$

Говорят, что измеримая функция $f(x)$ ограничена в существенном на $[a, b]$, если существует такое число μ , что мера множества точек, для которых $|f(x)| > \mu$, равна нулю. Минимальное из чисел μ , обладающее этим свойством, обозначается символом $\text{ess sup}_{x \in [a,b]} |f(x)|$. Введем пространство $L_\infty[a, b]$, измеримых и ограниченных в существенном функций с нормой

$$\|f(x)\|_{L_\infty[a,b]} = \text{ess sup}_{x \in [a,b]} |f(x)|.$$

Обозначение L_∞ оправдано тем, что для $f(x) \in L_\infty[a, b]$ при конечных a и b

$$\lim_{p \rightarrow \infty} \|f(x)\|_{L_p[a,b]} = \text{ess sup}_{x \in [a,b]} |f(x)|.$$

Заметим, что пространству $L_\infty[a, b]$ принадлежат, в частности, кусочно-непрерывные на $[a, b]$ функции с разрывами первого рода. При этом, если $f(x) \in C[a, b]$, то $\|f(x)\|_{C[a,b]} = \|f(x)\|_{L_\infty[a,b]}$. Иногда, если это не приводит к путанице, вместо обозначений $\|f(x)\|_{C[a,b]}$ и $\|f(x)\|_{L_\infty[a,b]}$ используется $\|f(x)\|_C$ и $\|f(x)\|_\infty$ соответственно.

Класс функций $f(x)$, имеющих на $[a, b]$ абсолютно непрерывную производную порядка $l-1$ и l -ю производную из $L_p[a, b]$, $1 \leq p \leq \infty$, обозначается через $W_p^l[a, b]$. Формула Тейлора для функций из этого класса применяется с остаточным членом в интегральной форме:

$$f(x) = f(a) + \dots + \frac{f^{(l-1)}(a)(x-a)^{l-1}}{(l-1)!} + \frac{1}{(l-1)!} \int_a^x (x-v)^{l-1} f^{(l)}(v) dv.$$

Если $f(x) \in L_p[a, b]$ и $g(x) \in L_q[a, b]$, $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p \leq \infty$, то справедливо неравенство Гельдера

$$\int_a^b |f(x)g(x)| dx \leq \|f(x)\|_{L_p[a, b]} \cdot \|g(x)\|_{L_q[a, b]}.$$

Через $C^k C_\Delta^l[a, b]$, $l > k$, обозначается класс функций таких, что

$$f(x) \in C^k[a, b] \text{ и } f(x) \in C^l[x_i, x_{i+1}], \quad i = 0, \dots, N-1.$$

Обозначение $C^k W_{\Delta, p}^l[a, b]$, $l > k$, $1 \leq p \leq \infty$, используется для класса функций $f(x)$, удовлетворяющих условиям

$$f(x) \in C^k[a, b] \text{ и } f(x) \in W_p^l[x_i, x_{i+1}], \quad i = 0, \dots, N-1.$$

Матрицы с диагональным преобладанием.

Квадратная матрица $A = [a_{ij}]$ порядка n называется **матрицей с диагональным преобладанием**, если выполняются условия

$$|a_{ii}| - \sum_{j \neq i} |a_{ij}| = r_i > 0, \quad i, j = 1, \dots, N \quad (1)$$

Теорема 1. Матрица с диагональным преобладанием невырожденна.

Доказательство. Действительно, если предположить, что матрица вырожденна, то есть $|A| = 0$, то однородная система уравнений $Ay = 0$ имеет нетривиальное решение $y = (y_1, \dots, y_N)^T$,

$$\sum_j a_{ij} y_j = 0, \quad i = 1, \dots, N.$$

Пусть k таково, что $|y_k| \geq |y_i|$, $i = 1, \dots, N$. Тогда из k -го уравнения следует

$$|a_{kk} y_k| \leq \sum_{j \neq k} |a_{kj} y_j| \leq |y_k| \sum_{j \neq k} |a_{kj}|.$$

Отсюда

$$|a_{kk}| \leq \sum_{j \neq k} |a_{kj}|,$$

что противоречит предположению (1).

Оценка нормы обратной матрицы.

Если в пространстве векторов y норма введена соотношением

$$\|y\| = \max_{1 \leq i \leq N} |y_i|,$$

то согласованной с ней нормой в пространстве матриц $A = [a_{ij}]$, $i, j = 1, \dots, N$, является норма

$$\|A\| = \sup_{y \neq 0} \frac{\|Ay\|}{\|y\|} = \max_{1 \leq i \leq N} \sum_j |a_{ij}|.$$

Пусть задана система уравнений $Ay = d$ с невырожденной матрицей A . Её решение можно представить в виде $y = A^{-1}d$. Норма обратной матрицы A^{-1} есть

$$\|A^{-1}\| = \sup_{d \neq 0} \frac{\|y\|}{\|d\|} \quad (2)$$

Отсюда $\|y\| \leq \|A^{-1}\| \cdot \|d\|$. Следовательно, чтобы оценить норму решения системы через норму её правой части, достаточно вычислить или оценить $\|A^{-1}\|$.

Теорема 2. Если A — матрица с диагональным преобладанием, то справедлива оценка

$$\|A^{-1}\| \leq \max_{1 \leq i \leq N} \frac{1}{r_i} \quad (3)$$

Доказательство. Пусть y — произвольный ненулевой вектор. Если k таково, что $\|y\| = |y_k|$, то

$$\|d\| = \|Ay\| = \max_{1 \leq i \leq N} \left| \sum_j a_{ij} y_j \right| \geq \left| \sum_j a_{kj} y_j \right| \geq |a_{kk}| \cdot |y_k| - \sum_{j \neq k} |a_{kj} y_j| \geq |y_k| \left(|a_{kk}| - \sum_{j \neq k} |a_{kj}| \right) = \|y\| r_k \geq \|y\| \min_{1 \leq i \leq N} r_i.$$

Следовательно,

$$\|y\| \leq \|d\| \left(\min_{1 \leq i \leq N} r_i \right)^{-1} = \|d\| \max_{1 \leq i \leq N} \frac{1}{r_i}.$$

Учитывая эту оценку в равенстве (2), получаем (3).

Следствие. Если матрица системы $Ay = d$ с диагональным преобладанием, то

$$\|y\| \leq \max_{1 \leq i \leq N} \frac{|d_i|}{r_i}.$$

Доказательство. Разделив каждую из строк системы на r_i , получим систему $\bar{A}y = \bar{d}$, где $\bar{d}_i = d_i/r_i$. Согласно теореме 2 $\|\bar{A}^{-1}\| \leq 1$, и поэтому $\|y\| \leq \|\bar{d}\|$, что и требовалось доказать.

Метод прогонки для решения систем уравнений с трехдиагональными матрицами.

Многие задачи теории сплайнов приводят к решению систем линейных уравнений с ленточными матрицами $Ay = d$.

Величина $\mu(A) = \|A^{-1}\| \cdot \|A\|$ называется мерой (числом) обусловленности системы или матрицы. Системы уравнений и матрицы с небольшими значениями мер обусловленности принято называть хорошо обусловленными. Возможность эффективного решения системы зависит от того, является ли она хорошо обусловленной или нет.

Если A — матрица с диагональным преобладанием, то из теоремы 2 следует оценка

$$\mu(A) = \|A\| \max_i \frac{1}{r_i},$$

где r_i определяются формулой 1.

Метод прогонки в случае матрицы с диагональным преобладанием.

Пусть имеется система

$$\begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 & c_1 \\ c_2 & a_2 & b_2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & c_{N-1} & a_{N-1} & b_{N-1} \\ b_N & 0 & 0 & \dots & 0 & c_N & a_N \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \\ y_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N-1} \\ d_N \end{bmatrix} \quad (4)$$

Рассмотрим сначала случай, когда $c_1 = b_N = 0$. Решение ищется в виде

$$y_i = v_i y_{i+1} + u_i, \quad i = 1, \dots, N-1. \quad (5)$$

Используя выражение для y_{i-1} из (5), исключим это неизвестное из i -го уравнения системы. Получаем

$$(a_i + c_i v_{i-1}) y_i + b_i y_{i+1} = d_i - c_i u_{i-1}.$$

Сравнивая это соотношение с (5), выводим рекуррентные формулы для прогоночных коэффициентов v_i , u_i (прямая прогонка):

$$v_0 = u_0 = 0, \quad (6)$$

$$v_i = -\frac{b_i}{a_i + c_i v_{i-1}}, \quad u_i = \frac{d_i - c_i u_{i-1}}{a_i + c_i v_{i-1}}, \quad i = 1, \dots, N.$$

Очевидно, $y_N = u_N$. Все остальные неизвестные находятся по формулам (2) (обратная прогонка). Для реализации алгоритма требуется выполнить $8N$ арифметических операций.

На практике часто приходится иметь дело со случаем, когда требуется решить несколько систем, отличающихся лишь правыми частями. Нетрудно видеть, что в формулах (6) величины v_i и $a_i + c_i v_{i-1}$ не зависят от правой части системы. Поэтому, если вычислить их и запомнить, то для решения каждой из систем потребуется $5N$ арифметических операций.

Пусть теперь c_1 и b_N отличны от нуля. Будем искать y_i в виде

$$y_i = v_i y_{i+1} + w_i y_N + u_i, \quad i = 1, \dots, N-1. \quad (7)$$

Подставив выражение для y_{i-1} из (7) в i -е уравнение системы, получаем

$$(a_i + c_i v_{i-1}) y_i + b_i y_{i+1} + c_i w_{i-1} y_N = d_i - c_i u_{i-1}.$$

Отсюда приходим к формулам (6) для коэффициентов v_i , u_i , $i = 1, \dots, N-1$. Величины w_i определяются соотношениями

$$w_0 = 1, \quad w_i = -\frac{c_i w_{i-1}}{a_i + c_i v_{i-1}}, \quad i = 1, \dots, N-1. \quad (8)$$

Теперь выразим все y_i через y_N в виде

$$y_i = s_i y_N + t_i, \quad i = 1, \dots, N-1 \quad (9)$$

Подставляя выражение для y_{i+1} в (7), получаем

$$y_i = (v_i s_{i+1} + w_i) y_N + (v_i t_{i+1} + u_i).$$

Сравнивая это соотношение с (9), находим рекуррентные формулы для величин s_i , t_i

$$s_N = 1, \quad t_N = 0, \quad (10)$$

$$s_i = v_i s_{i+1} + w_i, \quad t_i = v_i t_{i+1} + u_i, \quad i = N-1, \dots, 1.$$

Подставляя y_1 и y_{N-1} из (9) в последнее уравнение системы (1), находим

$$y_N = \frac{d_N - b_N t_1 - c_N t_{N-1}}{a_N + b_N s_1 + c_N s_{N-1}} \quad (11)$$

после чего решение вычисляется по формулам (9). Всего необходимо выполнить $14N$ арифметических операций.

При решении серии систем с одинаковой матрицей предварительно следует вычислить и запомнить не зависящие от правой части величины v_i , w_i , s_i , $a_i + c_i v_{i-1}$. Тогда для решения каждой из систем потребуется $7N$ арифметических операций.

Алгоритм называется **корректным**, если все действия, необходимые для его реализации, выполнимы. Исследование корректности алгоритма прогонки сводится к выяснению условий, при которых знаменатели в формулах (6),(8),(11) не равен нулю.

Будем называть алгоритм прогонки **устойчивым**, если выполняются неравенства

$$|v_i| \leq 1, \quad |w_i| \leq 1, \quad i = 1, \dots, N-1. \quad (12)$$

В этом случае при счете по формулам (6), (8)-(10) не происходит прогрессивного накопления погрешностей округления за счет операций умножения.

Покажем, что если система (4) с диагональным преобладанием, то алгоритм прогонки корректен и устойчив. Вначале установим неравенства

$$|v_i| + |w_i| \leq 1, \quad i = 1, \dots, N-1, \quad (13)$$

из которых вытекают неравенства (12). Имеем $|v_0| + |w_0| = 1$. Предположим, что (13) выполняется при $i-1$. Тогда

$$|v_i| + |w_i| = \frac{|b_i| + |c_i| \cdot |w_{i-1}|}{|a_i + c_i v_{i-1}|} \leq \frac{|b_i| + |c_i|(1 - |v_{i-1}|)}{|a_i| - |c_i| \cdot |v_{i-1}|} = \frac{(|b_i| + |c_i|) - |c_i| \cdot |v_{i-1}|}{|a_i| - |c_i| \cdot |v_{i-1}|}.$$

Так как $|a_i| - |c_i| \cdot |v_{i-1}| \geq |a_i| - |c_i| > 0$ и $|a_i| > |b_i| + |c_i|$, то отсюда следует, что $|v_i| + |w_i| \leq 1$. Неравенства (13) доказаны. Попутно установлено, что знаменатель в формулах (6), (8) отличен от нуля. Кроме того, из (10) вытекает, что $|s_i| \leq |v_i| |s_{i+1}| + |w_i|$. Учитывая (13) и то, что $s_N = 1$, отсюда по индукции получаем $|s_i| \leq 1, i = 1, \dots, N-1$. Теперь очевидно, что знаменатель в формуле (11) не обращается в нуль, так как $|a_N + b_N s_1 + c_N s_{N-1}| \geq |a_N| - |b_N| |s_1| - |c_N| |s_{N-1}| > 0$. Пусть система (4) хорошо обусловлено, но без диагонального преобладания. В этом случае изложенный алгоритм не корректен, ибо не исключена возможность обращения в нуль знаменателей в формулах (6).

В случае $c_1 = b_N = 0$ опишем алгоритм, свободный от этого недостатка. Единственное условие его корректности — невырожденность матрицы системы (4).

Метод немонотонной прогонки.

Алгоритм прогонки, описанный в предыдущем параграфе, есть реализация алгоритма исключения Гаусса для систем с трехдиагональными матрицами. При этом в случае матриц с диагональным преобладанием автоматически используется главный элемент, что обеспечивает устойчивость вычислений. Алгоритм, который рассматривается ниже, основан на методе Гаусса с выбором главного элемента по строке для матриц без диагонального преобладания.

Пусть в результате i -го шага процесса исключения получена система:

$$H y_{k_i} + b_i y_{i+1} = F, \quad (14)$$

$$B y_{k_i} + a_{i+1} y_{i+1} + b_{i+1} y_{i+2} = G, \quad (15)$$

$$c_{i+2} y_{i+1} + a_{i+2} y_{i+2} + b_{i+2} y_{i+3} = d_{i+2}, \quad (16)$$

где $k_i \leq i$. В частности, при $i=1$ полагаем

$$k_i = 1, \quad H = a_1, \quad B = c_2, \quad F = d_1, \quad G = d_2.$$

Опишем $(i+1)$ -й шаг процесса исключения. Возможны два случая.

$|H| \geq |b_i|$. Уравнение (14) запишем в виде

$$y_{k_i} - v_i y_{i+1} = u_i,$$

где $v_i = -b_i/H$, $u_i = F/H$, причем $|v_i| \leq 1$.

С помощью этого соотношения исключим из (15) неизвестное y_{k_i} . Находим

$$(v_i B + a_{i+1}) y_{i+1} + b_{i+1} y_{i+2} = G - u_i B. \quad (17)$$

Объединяя уравнение (17) с оставшимися уравнениями исходной системы и обозначая $H = v_i B + a_{i+1}$, $F = G - u_i B$, $k_{i+1} = i+1$, $B = c_{i+2}$, $G = d_{i+2}$, снова приходим к системе вида (14)-(16), в которой i заменено на $i+1$.

$|H| < |b_i|$. Это означает, что b_i — главный элемент в строке. Преобразуем (14) к виду

$$y_{i+1} - v_i y_{k_i} = u_i,$$

где $v_i = -H/b_i$, $u_i = F/b_i$, причем опять $|v_i| \leq 1$.

Используя это уравнение, исключаем y_{i+1} из (15) и (16). Находим

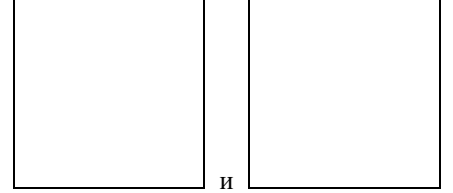
$$(B + v_i a_{i+1}) y_{k_i} + b_{i+1} y_{i+2} = G - u_i a_{i+1},$$

$$v_i c_{i+2} y_{k_i} + a_{i+2} y_{i+2} + b_{i+2} y_{i+3} = d_{i+2} - u_i c_{i+2}.$$

Объединяя полученные уравнения с оставшимися уравнениями системы и вводя обозначения $k_{i+1} = k_i$,

$$H = B + v_i a_{i+1}, \quad B = v_i c_{i+2}, \quad F = G - u_i a_{i+1}, \quad G = d_{i+2} - c_{i+2} u_i,$$

снова приходим к системе вида (14) - (16).



Итак, описан один шаг алгоритма исключения. Этот шаг выполним, если одновременно не обращаются в нуль. Но это действительно так, потому, что в противном случае система (4) была бы вырожденной. Следовательно, описанный алгоритм исключения корректен.

В результате исходная система (4) заменяется уравнениями

$$y_{\alpha_i} = v_i y_{\beta_i} + u_i, \quad i = 1, \dots, N-1, \quad (18)$$

$$y_{\alpha_N} = u_N.$$

При этом каждое неизвестное войдет хотя бы в одно уравнение.

Суммируя проделанные рассуждения, приведем расчетные формулы для вычисления коэффициентов v_i , u_i и индексов α_i , β_i .

Знак (=) используется как символ операции присвоения.

Полагаем $H = a_1$, $B = c_2$, $F = d_1$, $G = d_2$, $\beta_0 = 1$, $a_j = b_j = c_j = d_j = 0$, $j = N+1, N+2$.

Для $i = 1, 2, \dots, N$ выполняются описанные выше действия:

а) если $|H| \geq |b_i|$, то $v_i = -b_i/H$, $u_i = F/H$, $H = v_i B + a_{i+1}$, $B = c_{i+2}$, $F = G - u_i B$, $G = d_{i+2}$, $\alpha_i = \beta_{i-1}$, $\beta_i = i+1$;

б) если $|H| < |b_i|$, то $v_i = -H/b_i$, $u_i = F/b_i$, $H = B + v_i a_{i+1}$, $B = v_i c_{i+2}$, $F = G - u_i a_{i+1}$, $G = d_{i+2} - u_i c_{i+2}$, $\alpha_i = i+1$, $\beta_i = \beta_{i-1}$.

На обратном ходе алгоритма из (18) вычисляются неизвестные y_i . Порядок вычисления определяется индексами α_i , β_i и может иметь немонотонный характер. Отсюда название алгоритма. Однако всегда $|v_i| \leq 1$, и поэтому обратный ход устойчив по отношению к накоплению погрешностей округления.

Объем арифметических операций в методе немонотонной прогонки не превышает $12N$. При решении систем с одинаковой матрицей на каждую из них тратится не более $7N$ операций.

Алгоритм решения систем уравнений с пятидиагональными матрицами.

1. Вначале рассмотрим алгоритм для решения систем с симметричными положительно определенными матрицами

$$Ay = g \quad (19)$$

или

$$\begin{bmatrix} a_0 & b_0 & c_0 & 0 & \dots & 0 & c_{N-1} & b_N \\ b_0 & a_1 & b_1 & c_1 & \dots & 0 & 0 & c_N \\ c_0 & b_1 & a_2 & b_2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{N-1} & 0 & 0 & 0 & \dots & b_{N-2} & a_{N-1} & b_{N-1} \\ b_N & c_N & 0 & 0 & \dots & c_{N-2} & b_{N-1} & a_N \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ y_{N-1} \\ y_N \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \dots \\ g_{N-1} \\ g_N \end{bmatrix} \quad (20)$$

Матрица A представима в виде

где U — нижняя треугольная матрица, а W — диагональная матрица с диагональными элементами, отличными от

$$U = \begin{bmatrix} 1 & & & & & \\ u_0 & 1 & & & & 0 \\ v_0 & u_1 & 1 & & & \\ & & & & & \\ & 0 & & & 1 & \\ & & & v_{N-2} & u_{N-1} & 1 \end{bmatrix}, W = \begin{bmatrix} w_0 & & & & 0 \\ & w_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & w_N \end{bmatrix}.$$
$$\begin{aligned} s_i &= b_i - u_{i-1}c_{i-1}, \\ w_i &= a_i - v_{i-2}c_{i-2} - u_{i-1}s_{i-1}, \\ u_i &= s_i/w_i, v_i = c_i/w_i \\ i &= 0, 1, \dots, N \end{aligned} \quad (22)$$

Обозначая

запишем (19) в виде

Отсюда находим неизвестные z_i :

Здесь формально введены неизвестные z_{-1}, z_{-2} .

$$y_i = z_i/w_i - u_i y_{i+1} - v_i y_{i+2}, i = N, \dots, 0.$$

Общее количество операций равно $17N$.

Перейдем к случаю, когда в системе (20) хотя бы один из элементов c_{N-1}, b_N, c_N отличен от нуля. Снова используем разложение (21), но на этот раз матрицу U будем искать в виде

$$U = \begin{bmatrix} 1 & & & & & & & & & & \\ u_0 & 1 & & & & & & & & & 0 \\ v_0 & u_1 & 1 & & & & & & & & \\ 0 & v_1 & u_2 & 1 & & & & & & & \\ \hline 0 & 0 & 0 & 0 & \dots & v_{N-5} & u_{N-4} & 1 & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & v_{N-4} & u_{N-3} & 1 & & \\ p_0 & p_1 & p_2 & p_3 & \dots & p_{N-5} & p_{N-4} & p_{N-3} & p_{N-2} & 1 & \\ q_0 & q_1 & q_2 & q_3 & \dots & q_{N-5} & q_{N-4} & q_{N-3} & q_{N-2} & q_{N-1} & 1 \end{bmatrix}.$$

$$\begin{aligned} r_0 &= c_{N-1}, r_1 = -u_0 r_0, r_i = -v_{i-2} r_{i-2} - u_{i-1} r_{i-1}, \quad i = 2, \dots, N-4, \\ r_{N-3} &= c_{N-3} - v_{N-5} r_{N-5} - u_{N-4} r_{N-4}, \\ t &= b_N, t_1 = c_N - u_0 t_0, t_i = -v_{i-2} t_{i-2} - u_{i-1} t_{i-1}, \quad r = 2, \dots, N-3, \\ t_{N-2} &= c_{N-2} - v_{N-4} t_{N-4} - u_{N-3} t_{N-3}, \end{aligned}$$

$$w_{N-1} = a_{N-1} - \sum_{i=1}^{N-2} p_i r_i, \quad t_{N-1} = b_{N-1} - \sum_{i=1}^{N-2} p_i t_i, \quad w_N = a_N - \sum_{i=1}^{N-1} q_i t_i,$$

$$p_i = r_i / w_i, \quad i = 0, 1, \dots, N-2,$$

$$q_i = t_i / w_i, \quad i = 0, 1, \dots, N-1.$$

По формулам (24) определяем $z_i, i = 0, \dots, N-2$. Кроме того,

$$z_{N-1} = g_{N-1} - \sum_{i=0}^{N-2} p_i z_i, \quad z_N = g_N - \sum_{i=0}^{N-1} q_i z_i.$$

Наконец, из системы (23) находим неизвестные y_i :

$$y_N = z_N / w_N, \quad y_{N-1} = z_{N-1} / w_{N-1} - q_{N-1} y_N,$$

$$y_{N-2} = z_{N-2} / w_{N-2} - p_{N-2} y_{N-1} - q_{N-2} y_N,$$

$$y_{N-3} = z_{N-3} / w_{N-3} - u_{N-3} y_{N-2} - p_{N-3} y_{N-1} - q_{N-3} y_N,$$

$$y_i = z_i / w_i - u_i y_{i+1} - v_i y_{i+2} - p_i y_{N-1} - q_i y_N, \quad i = N-4, \dots, 0.$$

Общее число арифметических операций - 39N.

Рассмотрим систему

$$\begin{bmatrix} a_0 & b_0 & c_0 & & & & & & & & \\ d_1 & a_1 & b_1 & c_1 & & & & & & & \\ e_2 & d_2 & a_2 & b_2 & c_2 & & & & & & \\ & e_3 & d_3 & a_3 & c_3 & & & & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ & & & & & e_{N-2} & d_{N-2} & a_{N-2} & b_{N-2} & c_{N-2} & \\ & & & & & & e_{N-1} & d_{N-1} & a_{N-1} & b_{N-1} & \\ & & & & & & & e_N & d_N & a_N & \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ \dots \\ y_{N-2} \\ y_{N-1} \\ y_N \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \\ \dots \\ g_{N-2} \\ g_{N-1} \\ g_N \end{bmatrix} \quad (25)$$

Для ее решения можно использовать метод немонотонной прогонки. Алгоритм строится аналогично случаю системы с трехдиагональной матрицей.

Приведем расчетные формулы алгоритма.

Знак (=) в них означает операцию присвоения.

Положим

$$H = a_0, \quad B = b_0, \quad D = g_0, \quad c = d_1, \quad E = a_1, \quad F = g_1, \quad Q = e_2, \quad R = d_2, \quad S = g_2, \quad L = 0, \quad T = e_3, \quad K = g_3, \quad \beta_{-1} = 0,$$

$$\gamma_{-1} = 1, \quad b_N = c_N = c_{N-1} = 0, \quad a_i = b_i = d_i = e_i = g_i = 0, \quad i = N+1, \dots, N+4.$$

Последовательно для $i = 0, 1, \dots, N$ вычисляем прогоночные коэффициенты u_i, v_i, w_i и индексы $\alpha_i, \beta_i, \gamma_i$. При этом в зависимости от соотношений между расчетными величинами на $(i+1)$ -м шаге выполняется одна из трех групп операций:

1) если $|H| \geq |B|$ и $|H| \geq |c_i|$, то

$$\begin{aligned}
u_i &= -B/H, & v_i &= -c_i/H, & w_i &= D/H, \\
\alpha_i &= \beta_{i-1}, & \beta_i &= \gamma_{i-1}, & \gamma_i &= i+2, \\
H &= E + u_i C, & B &= b_{i+1} + v_i C, & D &= F - w_i C, \\
C &= R + u_i E, & E &= a_{i+2} + v_i Q, & F &= S - w_i Q, \\
Q &= T + u_i L, & R &= d_{i+3} + v_i L, & S &= K - w_i L, \\
L &= 0, & T &= e_{i+4}, & K &= g_{i+4},
\end{aligned}$$

2) если $|B| > |H|$ и $|B| \geq |c_i|$, то

$$\begin{aligned}
u_i &= -H/B, & v_i &= -c_i/B, & w_i &= D/B, \\
\alpha_i &= \gamma_{i-1}, & \beta_i &= \beta_{i-1}, & \gamma_i &= i+2, \\
H &= C + u_i E, & B &= b_{i+1} + v_i E, & D &= F - w_i E, \\
C &= Q + u_i R, & E &= a_{i+2} + v_i R, & F &= S - w_i R, \\
Q &= L + u_i T, & R &= d_{i+3} + v_i T, & S &= K - w_i T, \\
L &= 0, & T &= e_{i+4}, & K &= g_{i+4},
\end{aligned}$$

3) если $|c_i| > |H|$ и $|c_i| > |B|$, то

$$\begin{aligned}
u_i &= -H/c_i, & v_i &= -B/c_i, & w_i &= D/c_i, \\
\alpha_i &= i+2, & \beta_i &= \beta_{i-1}, & \gamma_i &= \gamma_{i-1}, \\
H &= C + u_i b_{i+1}, & B &= E + v_i b_{i+1}, & D &= F - w_i b_{i+1}, \\
C &= Q + u_i a_{i+2}, & E &= R + v_i a_{i+2}, & F &= S - w_i a_{i+2}, \\
Q &= L + u_i d_{i+3}, & R &= T + v_i d_{i+3}, & S &= K - w_i d_{i+3}, \\
L &= u_i e_{i+4}, & T &= v_i e_{i+4}, & K &= g_{i+4} - w_i e_{i+4},
\end{aligned}$$

Теперь, полагая $y_{N+1} = y_{N+2} = 0$, вычисляем неизвестные y_i по формулам

$$y_{\alpha_i} = u_i y_{\beta_i} + v_i y_{\gamma_i} + w_i, \quad i = N, \dots, 0.$$

Объем арифметических операций — $30N$. Если система (25) имеет диагональное преобладание, то в описанном алгоритме на каждом шаге выполняется только группа операций а). В этом случае достаточно $19N$ операций.

Литература:

Бахвалов Н.С. Численные методы, т. I - М.: Наука, 1973.

Гантмахер Ф.Р. Теория матриц. - М.: Наука, 1966.

Завьялов Ю.С., Квалов Б.И., Мирошниченко В.Л. Методы сплайн-функций. - М.: Наука, 1980.