В.И.Дмитриев, И.В.Дмитриева, Ж.Г. Ингтем

ИНТЕГРАЛЬНАЯ ФОРМА СПЛАЙН-ФУНКЦИИ В ЗАДАЧАХ АППРОКСИМАЦИИ

Ведение

Задачи аппроксимации неточно заданной функции возникают во многих практических проблемах. При математическом моделировании, при экспериментальных работах и при натурных наблюдениях мы получаем приближенные данные на некоторой сетке параметров. По этим данным необходимо построить аппроксимирующую функцию, которая позволяла бы вычислить значение функции для любого параметра из области задания функции. Очень часто кроме функции необходимо определить первую и вторую производные функции. Обычно для решения этой задачи используют аппроксимацию с помощью сплайн-функций второго порядка (параболического сплайна)[1] – [4].

Аппроксимационный параболический сплайн определяется как функция $S(x)\in C_1$, $x\in[0,l]$ и имеющая кусочно-постоянную вторую производную. Сплайн на некоторой сетке $\{x_n\}$, $x_n = nh$, $n\in[0,N]$, $h=\dfrac{l}{N}$ представляется в виде полинома второго порядка на каждом интервале сетки $x\in[x_n, x_{n+1}]$. Коэффициенты полиномов определяются через значения сплайна на сетке $\{S_n = S(x_n)\}$, $n\in[0,N]$ и значения производной сплайна в начальной точке $S'(x=0)=S'_0$. Таким образом, аппроксимационный параболический сплайн можно представить в виде:

$$S(x) = R(x, S_0, \ldots, S_n, S'_0) \tag{1}$$

где под R понимаем алгоритм определения сплайна при заданных $\{S_n\}$ и S'_0. Неизвестные $\{S_n\}$, S'_0 определяются из условия аппроксимации:

$$\min_{\{S_n\},S'_0}\left\{\sum_{m=0}^{M}\left(R(x^{(m)}, S_0, \ldots, S_n, S'_0) - \tilde{f}_m\right)^2\right\} \tag{2}$$

где $\tilde{f}_m = \tilde{f}(x^{(m)})$ – приближенные данные функции $f(x)$ на некоторой сетке $\{x^{(m)}\}, m\in[0,M]$, $x^0 = x_0 = 0$, $x^{(M)} = x_N = l$, $M > N+1$. Из условия минимизации определяются параметры сплайна $\{S_n\}$, S'_0, что полностью определяет сплайн согласно (1). Первая и вторая производные функции вычисляются путем дифференцирования выражения (1). Такой подход позволяет восстановить функцию, при этом даже первая производная определяется с погрешностями, а вторая – с очень большими ошибками. В статье описан новый подход к построению аппроксимационной сплайн-функции, позволяющий с хорошей точностью восстанавливать функцию и её производные.

Постановка задачи

Пусть изучается некоторый процесс, описываемый функцией $f(x)$, $x \in [0, l]$. Известны приближенные значения функции на некоторой произвольной сетке $\{x^{(m)}\}$, $m \in [0, M]$, причем $x^{(0)} = 0$, $x^{(M)} = l$. Обозначим их $\tilde{f}_m \approx f(x^{(m)})$. Необходимо построить аппроксимационный сплайн, с помощью которого можно достаточно точно восстановить как саму функцию, так и её первую и вторую производные.

Основная идея построения сплайна в новом методе состоит в задании сплайна n-го порядка $S_n(x)$ через производную n-го порядка $P_n(x)$, которая считается кусочно-постоянной функцией. Таким образом, сплайн-функция $S_n(x) \in C_{n-1}$ является решением следующей задачи:

$$\frac{d^n S_n(x)}{dx^n} = P_n(x), x \in [0, l]$$

с начальными условиями (3)

$$S_n(x=0) = S_n^0; \quad \left.\frac{dS_n}{dx}\right|_{x=0} = S_n'; \quad \left.\frac{d^m S_n}{dx}\right|_{x=0} = 0, \quad m \in [2, n],$$

где плотность сплайна $P_n(x)$ является кусочно-постоянной функцией на сетке построения сплайна $\{x_k = kh\}$, $h = l/K$, $k \in [0, K]$.

Решение задачи (3) можно получить в интегральном виде:

$$S_n(x) = S_n^0 + S_n' x + \frac{1}{(n-1)!}\int_0^x (x-\xi)^{n-1} P_n(\xi) d\xi. \tag{4}$$

Так как $P_n(x) = P_n^{(k)}$ при $x \in [x_{k-1}, x_k]$, $k \in [1, K]$, то интеграл в (4) вычисляется аналитически.

В результате получаем аналитическое представление сплайна в интегральной форме при $x \in [x_m, x_{m+1}]$

$$S_n(x) = S_n^0 + S_n' x + \sum_{k=1}^{m} \frac{P_n^{(k)}}{(n-1)!}\int_{x_{k-1}}^{x_k} (x-\xi)^{n-1} d\xi + \frac{P_n^{(m+1)}}{(n-1)!}\int_{x_m}^{x} (x-\xi)^{n-1} d\xi$$

или, после интегрирования имеем

$$S_n(x) = S_n^0 + S_n' x + \frac{1}{n!}\sum_{k=1}^{m} P_n^{(k)}\Big((x - x_{k-1})^n - (x - x_k)^n\Big) + \frac{P_n^{(m+1)}}{n!}(x - x_m)^n, \qquad (5)$$

при $x \in [x_m, x_{m+1}]$, $m \in [0, K-1]$.

Таким образом мы получили аналитическое представление сплайн-функции n-го порядка, зависящего от $(K+2)$ параметров (S_n^0, S_n', $P_n^{(k)}$, $k \in [1, K]$), где K – число интервалов, на которых строится сплайн. Следовательно, сплайн-функцию можно записать в виде:

$$S_n(x) = A\Big(x, S_n^0, S_n', P_m^{(1)}, \ldots, P_n^{(K)}\Big),$$

где A описывает алгоритм расчета сплайна (5), зависящий от параметров $\Big(S_n^0, S_n', P_m^{(1)}, \ldots, P_n^{(K)}\Big)$. На практике сплайны высокого порядка обычно не применяются. Для аппроксимации функции и её первой и второй производных используют параболический сплайн (сплайн второго порядка). Если необходимо вычислить и третью производную применяют кубический сплайн (сплайн третьего порядка).

Аппроксимация функции и её производных.

Рассмотрим построение интегральной формы для сплайна второго порядка (параболический сплайн). В этом случае, согласно (5), имеем

$$S_2(x) = S_2^0 + S_2^1 x + \frac{1}{2}\sum_{k=1}^{m} P_2^{(k)}\Big((x - x_{k-1})^2 - (x - x_k)^2\Big) + \frac{P_2^{(m+1)}}{2}(x - x_m)^2 \qquad (6)$$

при $x \in [x_m, x_{m+1}]$, $m \in [0, K-1]$. Пусть нам известны приближенные значения аппроксимируемой функции $f(x)$ на некоторой сетке $\{x^{(n)}\}$, $n \in [0, N]$, т.е. $f(x^{(n)}) \approx \tilde{f}_n$. Вектор параметров $\overline{v} = \{S_2^0, S_2^1, \overline{P}\}$, где $\overline{P} = \{P_2^{(k)}\}$, $k \in [1, K]$, определяющий, согласно (6), параболический сплайн $S_2(x, \overline{v})$, находится из минимизации среднеквадратичной ошибки:

$$\min_{\overline{v}} \sum_{n=0}^{N} \Big(S_2(x^{(n)}, \overline{v}) - \tilde{f}_n\Big)^2. \qquad (7)$$

Если $N > K + 2$, то из (7) однозначно определяется вектор параметров $\overline{v}$, а, следовательно, находится сплайн функция, аппроксимирующая заданную функцию. Если $N \le K + 2$, то $\overline{v}$ определяется неустойчиво и требуется стабилизировать задачу минимизации. Это делается с помощью стабилизатора

$$\Omega(P) = \sum_{k=1}^{K-1} (P^{(k+1)} - P^{(k)})^2.$$

Согласно теории регуляризации некорректно поставленных задач [5] задача сводится к минимизации стабилизирующего функционала:

$$\min_{\overline{v}} \left\{ \sum_{n=0}^{N} \Big(S_2(x^{(n)}, \overline{v}) - \bar{f}_n\Big)^2 + \alpha \sum_{k=1}^{K-1} \Big(P^{(k+1)} - P^{(k)}\Big)^2 \right\}, \qquad (8)$$

где α – параметр регуляризации, который зависит от погрешности, с которой измеряется $\tilde{f}_n$.

Задача (8) сводится к решению системы линейных алгебраических уравнений относительно $\bar{v}$. Вектор значений сплайна на сетке $\{x^{(n)}\}$, $n \in [0, N]$ можно записать в матричном виде:

$$\bar{S} = \{S_2(x^{(n)}, \bar{v})\} = \hat{S} \cdot \bar{v}. \tag{9}$$

Тогда из (8) получаем систему

$$\left(\alpha\hat{\Omega} + \hat{S}^T \cdot \hat{S}\right) \cdot \bar{v} = \hat{S}^T \cdot \bar{f}, \tag{10}$$

где $\hat{S}^T$ – транспонированная матрица $\hat{S}$, $\bar{f} = \{\tilde{f}_n\}$, а $\hat{\Omega}$ – матрица регуляризации:

$$\Omega = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & & 0 \\ 0 & 0 & 0 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & -1 & 0 & & \vdots \\ 0 & 0 & -1 & 2 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & & 0 & -1 & 2 & -1 \\ 0 & 0 & \cdots & & 0 & -1 & 1 \end{bmatrix}$$

Первые две строки и два столбца $\hat{\Omega}$ –нулевые, так как регуляризация требует плавности изменения только $P_2^{(k)}$, $k \in [1, K]$, то есть второй производной сплайн-функции. Матрица $\hat{S}$ легко определяется из выражения сплайна на сетке $\{x^{(n)}\}$: согласно (6) имеем при $x^{(n)} \in [x_m, x_{m+1}]$, $m \in [0, K-1]$:

$$S_2(x^{(n)}) = S_2^0 + S_2^1 x^{(n)} + \sum_{k=1}^{m} h\left(x^{(n)} - x_k + \frac{h}{2}\right) P_2^{(k)} + \frac{(x^{(n)} - x_m)^2}{2} P_2^{(m+1)}. \tag{11}$$

Откуда получаем матрицу $\hat{S}$ в виде:

$$S = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x^{(1)} & c_1^1 & \cdots & c_K^1 \\ 1 & x^{(2)} & c_1^2 & \cdots & c_K^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x^{(N)} & c_1^N & \cdots & c_K^N \end{bmatrix}$$

Элементы матрицы c_k^n, $k \in [1, K]$, $n \in [0, N]$ вычисляются из (11) и зависят от соотношения сетки сплайна $\{x_k\}$, $k \in [0, K]$ и сетки задания функции $\{x^{(n)}\}$, $n \in [0, N]$. $c_k^{(0)} = 0$, $k \in [1, K]$, так как $S_2(x^{(0)} = x_0 = 0) = S_2^0$. Наиболее просто $c_m^{(n)}$ выписываются при совпадении сеток $K = N$ $x_k = kh$,

$k \in [0, K]$, $x^{(n)} = nh$, $n \in [0, N = K]$, $h = l/N$. В этом случае имеем $x^{(n)} = x_m$: $c_k^{(0)} = 0$, $k \in [1, K]$ и

$$c_k^{(n)} = \begin{cases} \left(n - k + \frac{1}{2}\right) \cdot h^2, & k \in [1, n-1] \\ \dfrac{h^2}{2}, & k = n \\ 0, & k \in [n+1, N = K] \end{cases}$$

В результате получаем, что $(N+2)$ неизвестные $\bar{v} = (S_2^0, S_2^1, P^{(1)}, \cdots, P^{(N)})$ необходимо определить при задании $(N+1)$ значения функции $\tilde{f}_n$, $n \in [0, N]$. Благодаря регуляризации вектор параметров сплайна $\bar{v}$ из системы (10) определяется устойчиво.

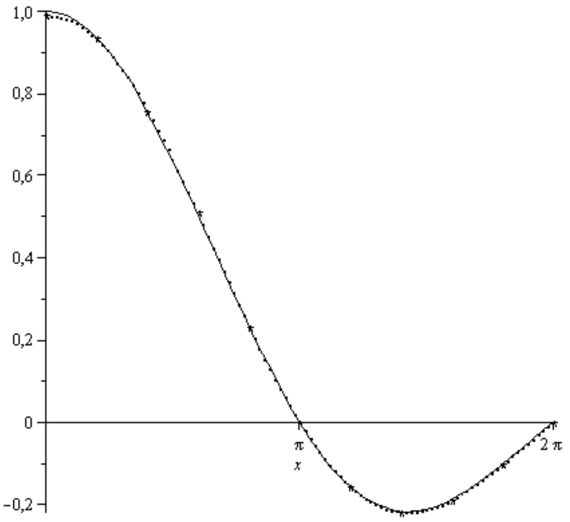
Найдя $\bar{v}$, получаем аналитическое выражение для сплайн-функции (6), которая приближает функцию $f(x)$. Первая производная функции приближенно равна:

$$f'(x_m) \approx S_2' + h \sum_{k=1}^{m} P_2^{(k)}, \qquad m \in [0, K], \tag{12}$$

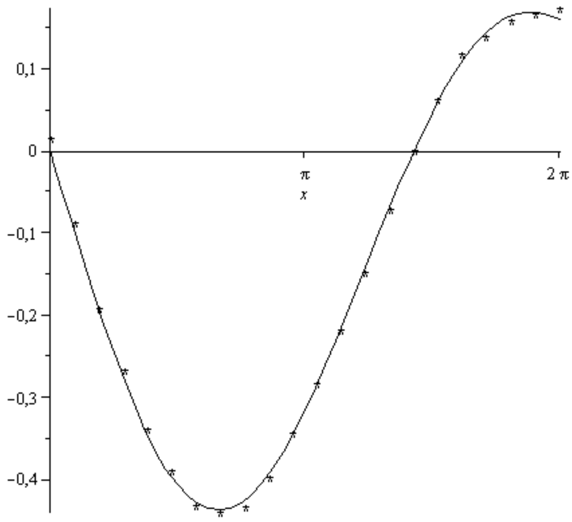
а вторая производная определяется в средних точках сетки:

$$f''\left(x_m + \frac{h}{2}\right) \approx S_2' + P_2^{(m+1)}, \qquad m \in [0, K-1]. \tag{13}$$

Если значения $f'(x)$ и $f''(x)$ нужны на более мелкой сетке, можно применить к полученным значениям повторную аппроксимацию.



******заданные приближенные значения функции*



******значения первой производной сплайна*

___*точная функция*, ….. *интегральный сплайн*

Рис 1

Аппроксимация $D(x)$ и $D'(x)$ с помощью интегрального сплайна.

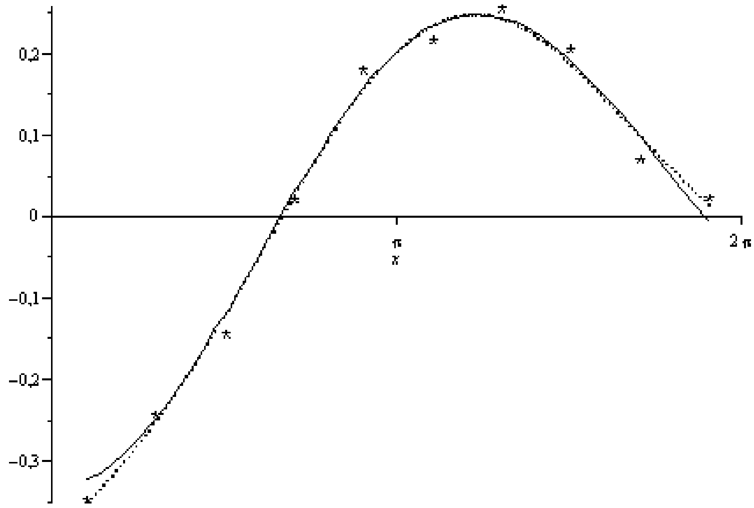
Численные результаты

Для оценки эффективности предложенного метода рассмотрим функцию Дирихле

$$D(x) = \frac{\sin(x)}{x}, \qquad x \in [0, 2\pi]. \tag{14}$$

Функция задается в 11 точках $x_n = \frac{2\pi}{10}$, $n \in [1,11]$ с погрешностью $\delta = 10^{-2}$. Сетка сплайна совпадает с сеткой задания функции. На рисунке 1 приведен результат аппроксимации функции и её первой производной.

Легко видеть, что функция и её первая производная определяются достаточно хорошо. На рисунке 2 приведен результат аппроксимации второй производной функции Дирихле $D''(x)$.



**полученные значения* P_i *взятые в средних точках сетки* $\{x_i\}_1^{11}$,

____*точная функция*, …….. *интегральный сплайн*

Рис 2

Аппроксимация второй производной

На рисунке сплошной линией обозначены точные значения второй производной, точками значения $D''(x_n)$ на сетке, а пунктирной кривой аппроксимация второй производной с помощью сплайна. Хорошо видно, что вторая производная приближается с высокой точностью на всем отрезке за исключением начальной и конечной точек. Это связано с крупным шагом сплайна. При уменьшении шага точность в начальной и конечной точках повышается.

Кубический сплайн

Для аппроксимации самой функции или её первой производной по известным значениям достаточно построение параболического интегрального сплайна. Несмотря на то, что на рисунке 2 было показано, что можно также построить хорошую аппроксимацию второй производной, целесообразнее для этой цели строить кубический сплайн.

По формуле (5) имеем следующее представление для кубического сплайна:

$$S_3(x) = S_3^0 + S_3' x + \frac{1}{6}\sum_{k=1}^{m} P_3^{(k)}\left((x - x_{k-1})^3 - (x - x_k)^3\right) + \frac{P_3^{(m+1)}}{6}(x - x_m)^3, \tag{15}$$

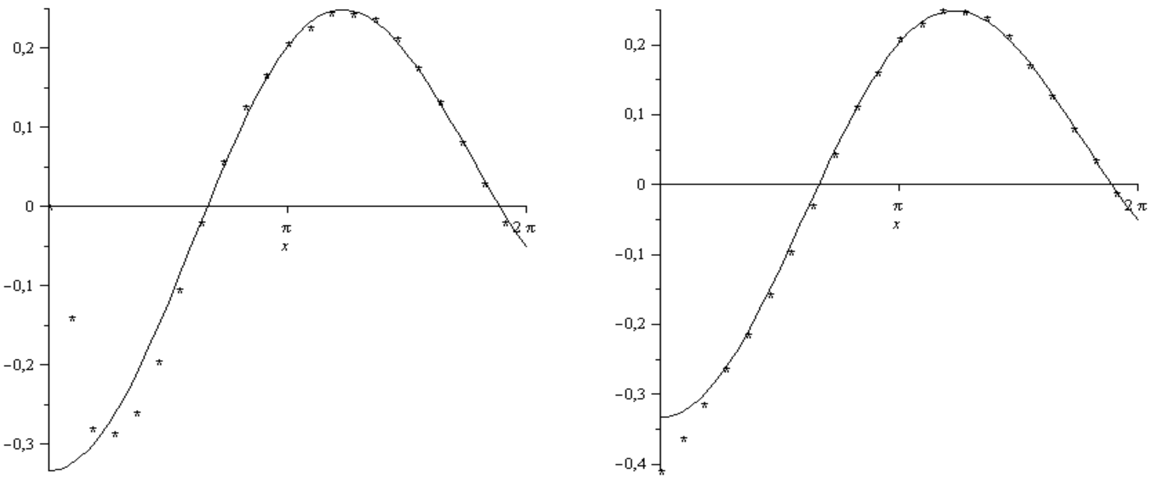
при $x \in [x_m, x_{m+1}]$, $m \in [0, K-1]$. Надо заметить, что в отличие от полиномиальных сплайнов количество неизвестных в интегральном сплайне не зависит от порядка сплайна. То есть исходя из того, что заданно K отрезков для полиномиального сплайна второго порядка необходимо найти $3K$ неизвестных, а для третьего порядка $4K$. Для интегрального сплайна количество неизвестных в том и другом случае будет неизменно равно $K+2$.

Пусть по аналогии построения параболического интегрального сплайна известны приближенные значения аппроксимируемой функции $f(x)$ на некоторой сетке $\{x^{(n)}\}$, $n \in [0, N]$, $f(x^{(n)}) \approx \tilde{f}_n$. Вектор параметров $\overline{v} = \{S_3^0, S_3^1, \overline{P}\}$, где $\overline{P} = \{P_3^{(k)}\}$, $k \in [1, K]$, определяющий кубический сплайн $S_3(x, \overline{v})$ (15), находится из минимизации среднеквадратичной ошибки. Плотность сплайна $P_3(x)$ соответствует производной третьего порядка кубического сплайна. Таким образом, в соответствии с (3) $S_3''(0) = 0$, это условие приводит к тому, что в начале отрезка вторая производная кубического сплайна не будет совпадать со второй производной функции (для тех функций, производные которых отличны от нуля в начале отрезка). Нахождение вектора $\overline{v}$ для кубического интегрального сплайна происходит по аналогии с параболическим. Для устранения этого недостатка можно в формуле (4) добавить слагаемое $S_3'' \cdot \frac{x^2}{2}$:

$$S_3(x) = S_3^0 + S_3' x + S_3'' \frac{x^2}{2} + \frac{1}{2}\int_0^x (x - \xi)^2 P_3(\xi) d\xi \tag{16}$$

и соответственно количество неизвестных становиться на одно больше, то есть $K+3$. Таким образом, мы находим помимо значений самой функции, её первой производной и $\overline{P} = \{P_3^{(k)}\}$, $k \in [1, K]$ ещё и значение второй производной в начале отрезка, т.е. минимизация среднеквадратичной ошибки, берется по вектору $\overline{v} = \{S_3^0, S_3', S_3'', \overline{P}\}$.

Рассмотрим тот же пример с функцией Дирихле, однако вместо параболического построим кубический интегральный сплайн для приближения второй производной по заданным значениям функции. Как можно заметить, в случае сплайна (15) (рис 3а) вторая производная в начале отрезка отклоняется от аппроксимируемой функции за счет того, что на неё действует принудительное условие $S''(0) = 0$. Но если построить интегральный кубический сплайн по формуле (16), то вторая производная определяется достаточно хорошо (рис 3б).



а б

* *значения второй производной,* ____*точная функция*, …….. *интегральный сплайн*

Рис 3

Аппроксимация второй производной кубическим интегральным сплайном.

Приведенные примеры показывают, что интегральный сплайн позволяет эффективно аппроксимировать данные, заданные на сетке с большим шагом. Рассмотрим случай сложной функции, когда приходится брать значение функции на мелком шаге. Покажем, что и в случае мелкого шага интегральный сплайн (в отличие от обычного параболического сплайна) позволяет построить аппроксимацию производных с хорошей точностью.

Рассмотрим функцию

$$f(x) = (1+\sin^2(\pi x))^{-2}\cos\left(\pi\left(1/2 + \pi x^3\right)^{-3}\right), \qquad (17)$$

график которой приведен на рисунке 4. Пусть мы имеем на отрезке [0,1] 151 приближенное значение функции $\{\tilde{f}_n\}$, $n \in [0, M = 150]$ с погрешностью $\delta \approx 10^{-2}$. Требуется аппроксимировать функцию и её первую и вторую производные. Рассмотрим вначале применение в этой задаче обычного сплайна второго порядка [6].

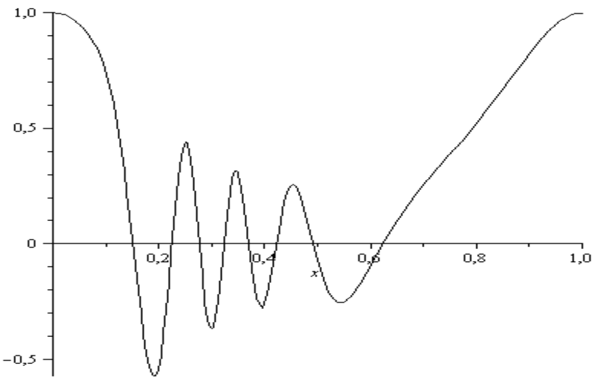
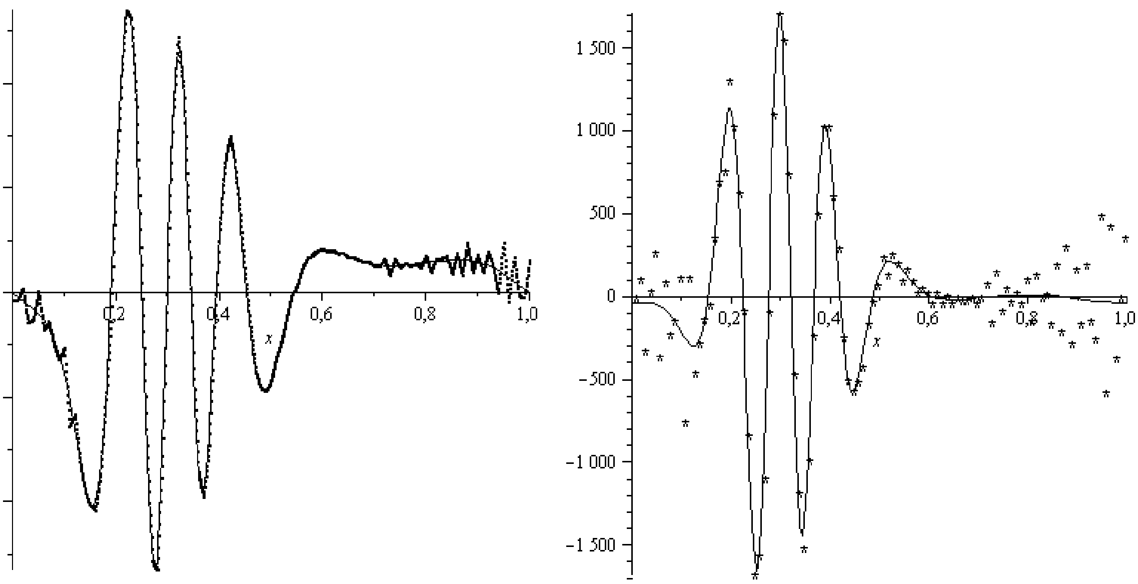


Рис.4 *График аппроксимируемой функции.*

В этом случае сплайн задается в виде полинома второго порядка на каждом участке $[x_{n-1}, x_n]$, $n \in [1, M]$ и условиями гладкой склейки в точках x_n, $n \in [1, M-1]$ сетки т.е.

$$\begin{gathered}\tilde{S}_n(x) = a_n x^2 + b_n x + c_n,\ x \in [x_{n-1}, x_n],\ n \in [1, M],\\ \tilde{S}'_n(x_n) = \tilde{S}'_{n+1}(x_n),\ \tilde{S}_n(x_n) = \tilde{S}_{n+1}(x_n),\ n \in [1, M-1].\end{gathered} \tag{18}$$



а

б

______ $f'(x)$, ……. *первая производная сплайна*

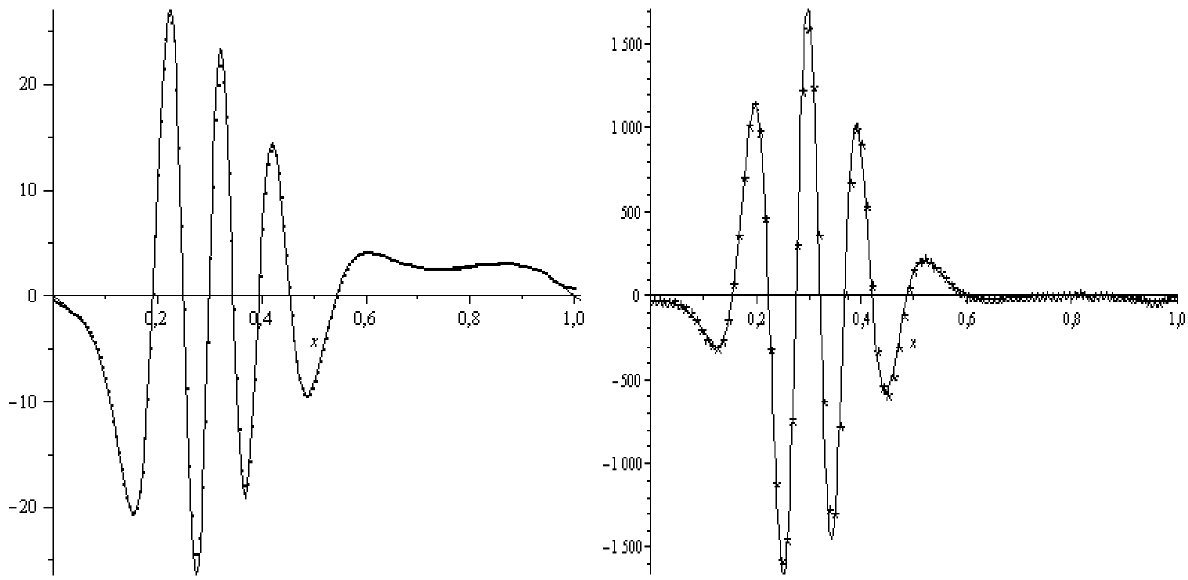
______ $f''(x)$ * *значения второй производной сплайна*

Рис 5

Аппроксимация первой и второй производных полиномиальным сплайном второй степени

Таким образом, имеется $3M$ неизвестных и $(2M-2)$ уравнений. Чаще всего коэффициенты сплайна a_n, b_n, c_n выражают через значения сплайна в узлах сетки и через производную на границе. Была построена обычная параболическая сплайн-аппроксимация для нашей функции, заданной с погрешностью $\delta \approx 10^{-2}$ в 151 точке на отрезке $[0,1]$. Аппроксимационный сплайн строился в 101 точке. Значения функции находились из минимизации функционала Тихонова с коэффициентом регуляризации $\alpha = 10^{-3}$, а в качестве производной в начальной точке бралась разностная производная. Результаты вычисления первой и второй производных функции, полученные с помощью найденного сплайна, приведены на рисунке 5. Видно, что первая производная определяется с большими погрешностями в начале и в конце интервала, а вторая производная определяется с очень большими погрешностями.

Рассмотрим применение к этой же задаче аппроксимацию интегральным параболическим сплайном. Заметим, что для построения интегрального сплайна искомыми величинами являются значения второй производной. Это позволяет построить не только сам сплайн, но и устойчиво определить аппроксимацию производных.



а — ______ $f'(x)$, ……. *первая производная сплайна*

б — ______ $f''(x)$ * *значения второй производной сплайна*

Рис 6

Аппроксимация интегральным сплайном на большом количестве данных

Полученные результаты аппроксимации первой и второй производной интегральным сплайном приведены на рисунке 6. Легко видеть, что даже при очень мелком шаге интегральный сплайн устойчиво определяет производные значительно лучше по сравнению с обычным полиномиальным сплайном.

Таким образом, в настоящей статье был описан новый (интегральный) подход к построению сплайн аппроксимационной функций, позволяющий эффективно приближать функцию и её производные по заданным значениям функции на сетке, как с крупным, так и с мелким шагом. Интегральное представление сплайна позволяет аналитически описать сплайн функцию и все её производные на всем отрезке аппроксимации. Построение такого сплайна требует всего лишь решение одной системы уравнения для нахождения необходимых неизвестных.

Литература

1. V.I.Dmitriev, J.G. Ingtem Numerical differentiation using spline functions// Computational mathematics and modeling vol.23 № 3 pp 312-318// Springer 2012.
2. V.I.Dmitriev J.G.Ingtem A two-dimensional minimum-derivative spline// Computational mathematics and modeling vol.21 № 2 pp 206-211// Springer 2010.
3. Ингтем Ж.Г. Сплайн функция с минимальной нормой производной в задачах интерполяции и аппроксимации// Вестник Московского Университета Вычислительная математика и кибернетика №4, 2008, с.16-27.
4. V.I.Dmitriev, J.G. Ingtem Solving an integral equation of the first kind by spline approximation// Computational mathematics and modeling vol.15 №2 April-June 2004//Kluwer academic consultants bureau.
5. Тихонов А.Н. Некорректно поставленные задачи и методы их решения// Методы решения некорректных задач и их применение. Тр. всесоюзной школы молодых ученных. М.: МГУ 1974 с. 6-11.
6. Ю.С. Волков, В.Л. Мирошниченко О приближении производных скачком интерполяционного сплайна //Математические заметки, 2011, т.89, №1, стр127-130.