Computational Mathematics and Modeling, Vol. 24, No. 4, October, 2013

INTEGRAL FORM OF THE SPLINE FUNCTION IN APPROXIMATION PROBLEMS

V. I. Dmitriev,[1] I. V. Dmitrieva,[2] and J. G. Ingtem[3]

UDC 518.12

The article examines a new (integral) approach to the construction of a spline-approximation function. The proposed approach simplifies the process of spline construction. The integral form of the spline yields an analytical representation of the spline function and its derivatives on the entire approximation interval. Unlike with polynomial splines, the number of unknowns to be determined for integral spline construction does not depend on the order of the spline (when constructing parabolic, cubic, and n th order splines, the number of unknowns does not change and depends only on the spline grid). This spline makes it possible to efficiently approximate the function and its derivatives from given function values on both fine and coarse grids.

Keywords: spline, approximation, spline approximation, integral spline

Introduction

Approximation of a function specified with errors is a problem that arises in many applications. In mathematical modeling, in experimental research, and in observations in nature we typically obtain approximate values on some parameter grid. These values are then used to construct an approximating function, which makes it possible to evaluate the function for any parameter from its definition domain. Often, in addition to the function values, it is required to determine the first and second derivatives of the function. This problem is typically solved by approximation with second order spline functions (parabolic spline) [1–4].

The approximating parabolic spline is defined as the function $S(x) \in C_1$, $x \in [0, l]$ with a piecewise-constant second derivative. On some grid $\{x_n\}$, $x_n = nh$, $n \in [0, N]$, $h = \frac{l}{N}$ the spline is represented as a second-order polynomial on each grid interval $x \in [x_n, x_{n+1}]$. The polynomial coefficients are determined from the spline values on the grid $\{S_n = S(x_n)\}$, $n \in [0, N]$ and the value of the spline derivative at the initial point $S'(x=0) = S'_0$. Thus, the approximating parabolic spline may be represented in the form

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

where R is the algorithm that constructs the spline from given $\{S_n\}$ and S'_0. The unknown $\{S_n\}$, S'_0 are determined from the approximation condition:

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

[1] Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia; e-mail: dmitriev@cs.msu.su.
[2] Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia; e-mail: dmitriev@cs.msu.su.
[3] Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia; e-mail: nmail2002@yandex.ru.

Translated from Prikladnaya Matematika i Informatika, No. 41, pp. 27–37, 2012.

 1046–283X/13/2404–0488

where $\tilde{f}_m = \tilde{f}(x^{(m)})$ are the approximate values of the function $f(x)$ on some grid $\{x^{(m)}\}$, $m \in [0, M]$, $x^0 = x_0 = 0$, $x^{(M)} = x_N = l$, $M > N+1$. The spline parameters $\{S_n\}$, S_0' are determined from the minimization condition, and this completely defines the spline according to (1). The first and the second derivatives of the function are evaluated by differentiating (1). This approach reconstructs the function, but the first derivative is determined with errors, while the second derivative is determined with very large errors. In this article we describe a new approach to the construction of an approximating spline function that ensures restoration of both the function and its derivatives with satisfactory accuracy.

Statement of the Problem

Suppose that we are investigating some process described by the function $f(x)$, $x \in [0, l]$. Approximate function values are known on an arbitrary grid $\{x^{(m)}\}$, $m \in [0, M]$, where $x^{(0)} = 0$, $x^{(M)} = l$. Denote the known function values by $\tilde{f}_m \approx f(x^{(m)})$. It is required to find an approximating spline that reconstructs with sufficient accuracy the function with its first and second derivatives.

The main idea of spline construction in the new method relies on the specification of the nth order spline $S_n(x)$ in terms of the nth order derivative $P_n(x)$, which is assumed to be piecewise-constant. Thus, the spline function $S_n(x) \in C_{n-1}$ is the solution of the following problem:

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

with the initial conditions

$$S_n(x=0) = S_n^0; \qquad \left.\frac{dS_n}{dx}\right|_{x=0} = S_n'; \qquad \left.\frac{d^m S_n}{dx}\right|_{x=0} = 0, \qquad m \in [2, n], \tag{3}$$

where the spline density $P_n(x)$ is a piecewise-constant function on the spline-construction grid $\{x_k = kh\}$, $h = l/K$, $k \in [0, K]$.

The solution of problem (3) can be obtained in integral form:

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

Since $P_n(x) = P_n^{(k)}$ for $x \in [x_{k-1}, x_k]$, $k \in [1, K]$, the integral in (4) can be evaluated analytically.

As a result, we obtain an analytical representation of the spline in integral form for $x \in [x_m, x_{m+1}]$

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

Integrating, we obtain

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

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

We have thus obtained an analytical representation of the nth order spline function that depends on $(K+2)$ parameters (S_n^0, S_n', $P_n^{(k)}$, $k \in [1, K]$), where K is the number of intervals on which the spline is constructed. Therefore, the spline function may be written in the form

$$S_n(x) = A\left(x, S_n^0, S_n', P_m^{(1)}, \dots, P_n^{(K)}\right),$$

where A is the algorithm for the evaluation of the spline (5) depending on the parameters $\left(S_n^0, S_n', P_m^{(1)}, \dots, P_n^{(K)}\right)$. High-order splines are usually not applied in practice. A parabolic spline (a spline of order 2) is used to approximate the function and its first and second derivatives. If the third derivative is required, a cubic spline (a spline of order 3) is used.

Approximating the Function and Its Derivatives

Consider the construction of an integral form for a second-order (parabolic) spline. In this case, we have by (5)

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

for $x \in [x_m, x_{m+1}]$, $m \in [0, K-1]$. Assume that we know approximate values of the function $f(x)$ on some grid $\{x^{(n)}\}$, $n \in [0, N]$, i.e., $f(x^{(n)}) \approx \tilde{f}_n$. The parameter vector $\bar{\nu} = \{S_2^0, S_2^1, \bar{P}\}$, where $\bar{P} = \{P_2^{(k)}\}$, $k \in [1, K]$, determining by (6) the parabolic spline $S_2(x, \bar{\nu})$ is obtained by minimizing the mean square error:

$$\min_{\bar{\nu}} \sum_{n=0}^{N}\left(S_2(x^{(n)}, \bar{\nu}) - \tilde{f}_n\right)^2 . \tag{7}$$

If $N > K+2$, then from (7) we uniquely determine the parameter vector $\bar{\nu}$ and thus obtain the spline function that approximates the given function. If $N \le K+2$, then $\bar{\nu}$ is unstable and the minimization problem needs to be stabilized. This is achieved using the stabilizer

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

By the theory of regularization of ill-posed problems [5], the problem reduces to minimization of the stabilizing functional:

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

where α is the regularization parameter dependent on the observation error in $\tilde{f}_n$.

Problem (8) reduces to solving a system of linear algebraic equations for $\overline{\mathrm{v}}$. The vector of spline values on the grid $\{x^{(n)}\}$, $n\in[0,N]$ may be written in matrix form:

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

Then from (8) we obtain the system

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

where $\hat{S}^T$ is the transpose of the matrix $\hat{S}$, $\overline{f}=\{\tilde{f}_n\}$, and $\hat{\Omega}$ is the regularization matrix:

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

The first two rows and two columns of $\hat{\Omega}$ are all zeros, because regularization requires gradual variation only of $P_2^{(k)}$, $k\in[1,K]$, i.e., the second derivative of the spline function. The matrix $\hat{S}$ is easily determined from the expression of the spline on the grid $\{x^{(n)}\}$: by (6) we have for $x^{(n)}\in[x_m,x_{m+1}]$, $m\in[0,K-1]$:

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

Hence we obtain the matrix $\hat{S}$ in the form

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

The matrix elements c_k^n, $k \in [1, K]$, $n \in [0, N]$ are evaluated from (11); they depend on the relationship between the spline grid $\{x_k\}$, $k \in [0, K]$ and the grid on which the function is specified $\{x^{(n)}\}$, $n \in [0, N]$. $c_k^{(0)} = 0$, $k \in [1, K]$, because $S_2(x^{(0)} = x_0 = 0) = S_2^0$. The simplest expressions for $c_m^{(n)}$ are obtained when the two grids are identical: $K = N$, $x_k = kh$, $k \in [0, K]$, $x^{(n)} = nh$, $n \in [0, N = K]$, $h = l/N$. In this case, we have $x^{(n)} = x_m$: $c_k^{(0)} = 0$, $k \in [1, K]$ and

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

As a result we obtain that the $(N+2)$ unknowns $\bar{\nu} = (S_2^0, S_2^1, P^{(1)}, \cdots, P^{(N)})$ must be determined when specifying the $(N+1)$ function values $\tilde{f}_n$, $n \in [0, N]$. Thanks to regularization, the vector of spline parameters $\bar{\nu}$ determined from (10) is stable.

Having found $\bar{\nu}$, we obtain an analytical expression for the spline function (6) approximating the function $f(x)$. The first derivative of the function is approximately given by

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

and the second derivative is determined at the grid midpoints:

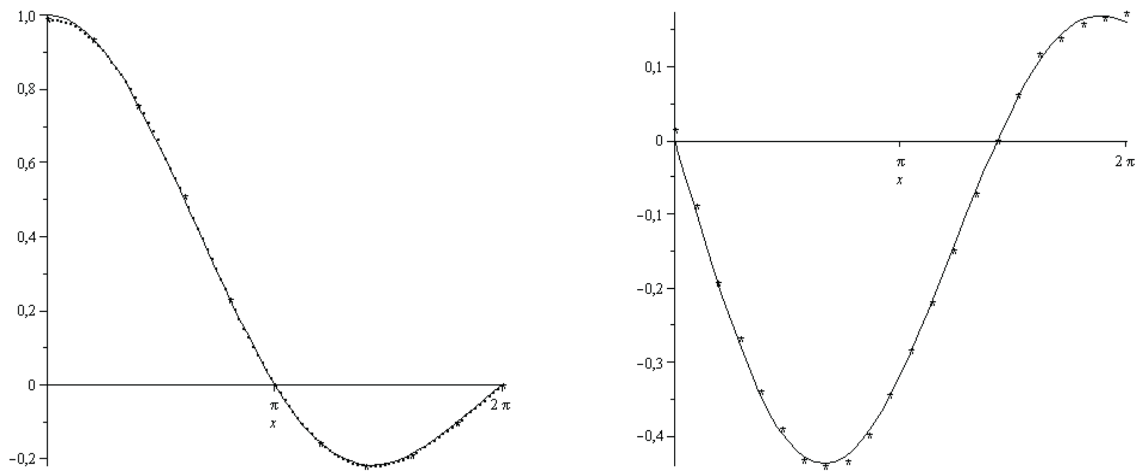
$$f''\left(x_m + \frac{h}{2}\right) \approx S_2' + P_2^{(m+1)}, \qquad m \in [0, K-1]. \tag{13}$$

If the derivative values $f'(x)$ and $f''(x)$ are required on a denser grid, repeated approximation may be applied to these results.

Numerical Results

To estimate the efficiency of the proposed method, consider the Dirichlet function

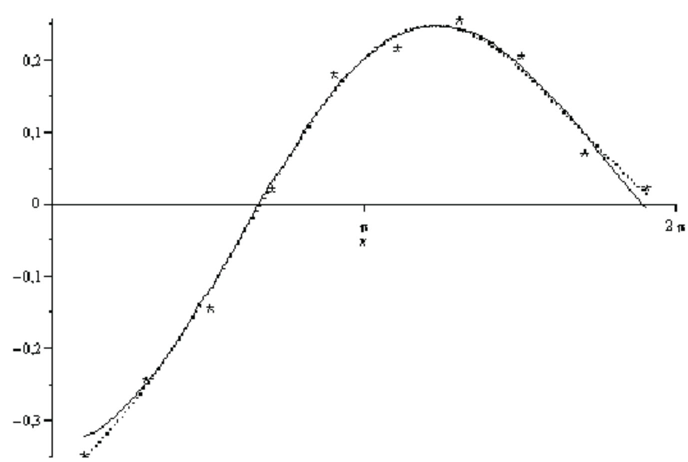
$$D(x) = \frac{\sin(x)}{x}, \qquad x \in [0, 2\pi]. \tag{14}$$



*given approximate function values *values of the first derivative of the spline

_____ exact function, ... integral spline

Fig. 1. Approximation of $D(x)$ and $D'(x)$ by an integral spline.



*values of P_i computed at midpoints of the grid $\{x_i\}_1^{11}$, _____ exact function, ... integral spline

Fig. 2. Approximation of the second derivative.

The function is specified at 11 points $x_n = \frac{2\pi}{10}$, $n \in [1,11]$ with error $\delta = 10^{-2}$. The spline grid is identical with the function grid.

Figure 1 shows the result of approximating the function and its first derivative. It is easy to see that the function and its first derivative are determined fairly accurately. Figure 2 shows the result of approximating the second derivative of the Dirichlet function $D''(x)$.

The solid curve in Fig. 2 plots the exact values of the second derivative; dots mark the grid values of $D''(x_n)$, and the broken curve approximates the second derivative by the spline. We clearly see that the second derivative is approximated with high accuracy on the entire interval, except the first and the last points. This is attributable to the coarseness of the spline. As the spline increment is reduced, the accuracy at the first and last points is improved.

Cubic Spline

A parabolic integral spline is sufficient to approximate the function and its first derivative from known values. Although Fig. 2 shows that this also produces a good approximation of the second derivative, it is preferable to construct a cubic spline for this purpose.

By (5) we have the following representation for the cubic spline:

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

for $x \in [x_m, x_{m+1}]$, $m \in [0, K-1]$. Note that unlike in polynomial splines, the number of unknowns in an integral spline is independent of the spline order. Given K intervals, a second-order spline requires finding $3K$ unknowns and a third-order spline $4K$ unknowns. For integral splines, on the other hand, the number of unknowns is $K+2$ in either case.

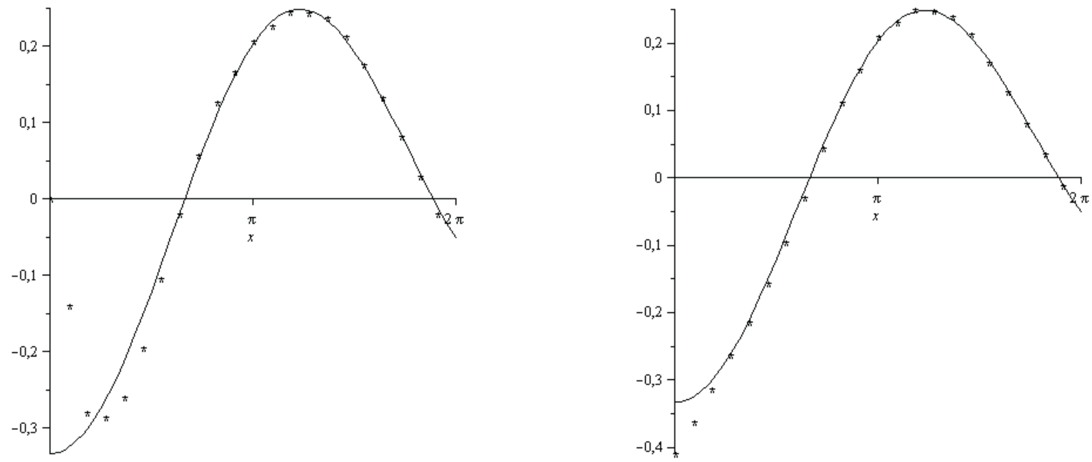
Assume that, similarly to the construction of a parabolic integral spline, we have approximate values of the approximated function $f(x)$ on some grid $\{x^{(n)}\}$, $n \in [0, N]$, $f(x^{(n)}) \approx \tilde{f}_n$. The parameter vector $\bar{\nu} = \{S_3^0, S_3^1, \bar{P}\}$, where $\bar{P} = \{P_3^{(k)}\}$, $k \in [1, K]$, defining the cubic spline $S_3(x, \bar{\nu})$ (15) is obtained by minimizing the mean square error. The spline density $P_3(x)$ corresponds to the third derivative of the cubic spline. Thus, by (3), $S_3''(0) = 0$, and as a result, at the beginning of the interval, the second derivative of the cubic spline is not identical with the second derivative of the function (for functions whose derivatives do not vanish at the beginning of the interval). The vector $\bar{\nu}$ for a cubic integral spline is determined similarly to the parabolic spline. To eliminate this shortcoming, we can add a term $S_3'' \cdot \frac{x^2}{2}$ in formula (4):

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

and the number of unknowns is correspondingly increased by 1, i.e., becomes $K+3$. Thus, in addition to the function values, the values of the first derivative, and $\bar{P} = \{P_3^{(k)}\}$, $k \in [1, K]$ we find also the value of the second derivative at the beginning of the interval, i.e., the mean square error is minimized over the vector $\bar{\nu} = \{S_3^0, S_3', S_3'', \bar{P}\}$.

Let us consider the same example with the Dirichlet function, replacing the previous parabolic spline with a cubic integral spline to approximate the second derivative from given function values. We note that with spline (15) (Fig. 3a) the second derivative at the beginning of the interval deviates from the approximated function because of the influence of the forced condition $S''(0) = 0$. If we construct an integral cubic spline from formula (16), then the second derivative is determined fairly accurately (Fig. 3b).

Our examples show that the integral spline leads to an efficient approximation of values defined on a coarse grid. Let us consider the case of a difficult function, when the function values have to be specified on a dense grid. We will show that in this case also the integral spline (contrary to the ordinary parabolic spline) approximates the derivatives with good accuracy.



(a) (b)

*second derivative values, _____ exact function, ... integral spline

Fig. 3. Approximation of the second derivative by a cubic integral spline.

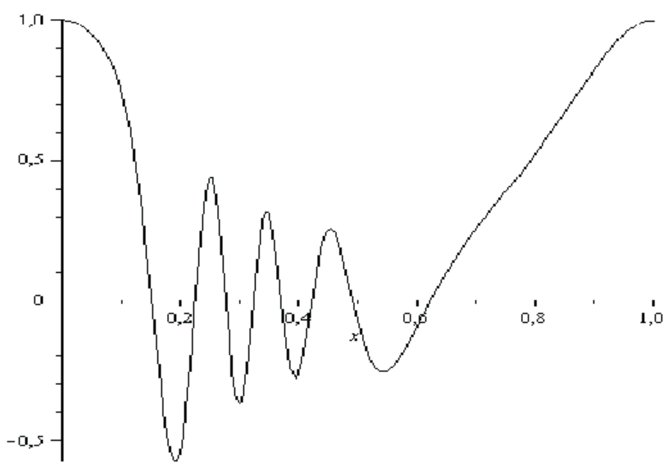


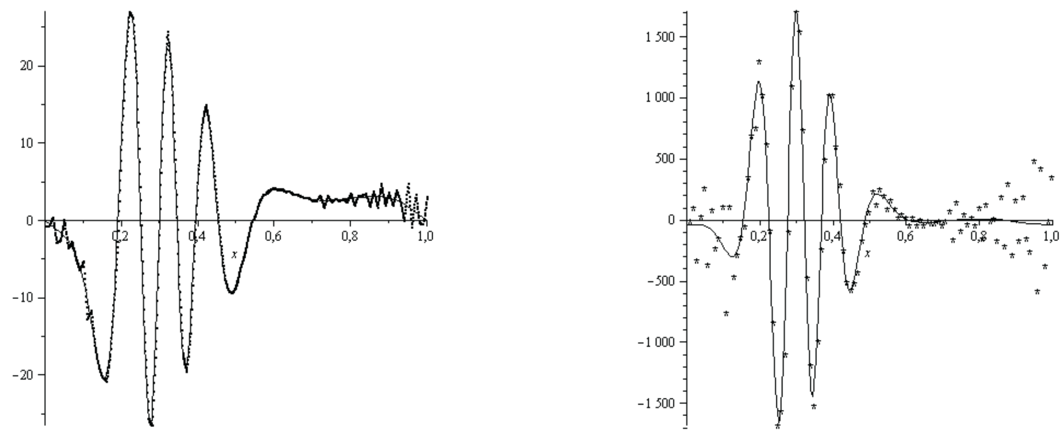
Fig. 4. The graph of the function to be approximated.

Consider the function

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

which is shown in Fig. 4. Assume that on the interval [0, 1] we have 151 approximate function values $\{\tilde{f}_n\}$, $n \in [0, M = 150]$ given with error $\delta \approx 10^{-2}$. It is required to approximate the function and its first and second derivatives.

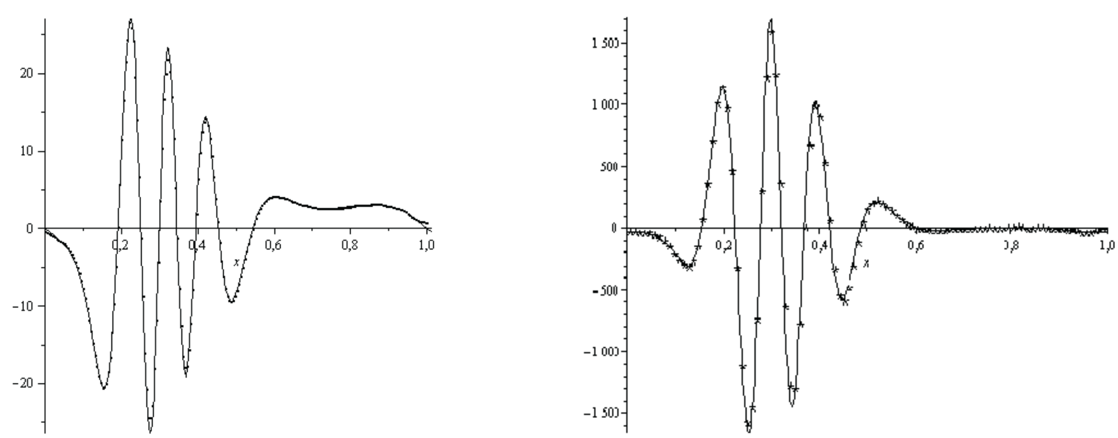
First let us apply the ordinary second-order spline to this problem [6]. In this case, the spline is defined as a second-order polynomial on each interval $[x_{n-1}, x_n]$, $n \in [1, M]$ with smooth matching conditions at the grid



(a) **(b)**

____ $f'(x)$, ... spline first derivative ____ $f''(x)$, * values of spline second derivative

Fig. 5. Approximation of the first and second derivatives by second-order polynomial spline.



(a) **(b)**

____ $f'(x)$, ... spline first derivative ____ $f''(x)$, * values of spline second derivative

Fig. 6. Approximation by an integral spline on a large set of values.

points x_n , $n \in [1, M-1]$, i.e.,

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

We thus have 3M unknowns and $(2M-2)$ equations. The spline coefficients a_n , b_n , c_n are usually expressed in terms of the spline values at the grid points and the derivative on the boundary. We have constructed the usual parabolic spline approximation for our function defined with error $\delta \approx 10^{-2}$ at 151 points on the interval $[0,1]$. The approximating spline is constructed at 101 points. The function values were obtained by minimizing the Tikhonov functional with regularization coefficient $\alpha = 10^{-3}$ and a difference derivative was used for the derivative at the initial point. The computation results for the first and second derivatives obtained from this spline are shown in Fig. 5. We see that the first derivative is determined with large errors at the beginning and the end of the interval, while the second derivative is determined with very large errors.

Let us consider the same problem using approximation by integral parabolic spline. Note that the construction of the integral spline requires second derivative values. In this way we not only construct the spline, but also obtain a stable approximation of the derivatives.

The results obtained for the approximation of the first and second derivatives by integral spline are shown in Fig. 6. It is easy to see that even with very dense grids the integral spline produces much better stable approximations of the derivatives than the ordinary polynomial spline.

We have thus described a new (integral) approach to the construction of approximating splines, which efficiently approximates a function and its derivatives from given function values on both coarse and dense grids. The integral spline representation yields an analytical description of the spline function and all its derivatives on the entire approximation interval. The construction of the integral spline requires solving a single system of equations to find the necessary unknowns.

REFERENCES

1. V. I. Dmitriev and J. G. Ingtem, "Numerical differentiation using spline functions," *Comp. Math. Model.,* **23**, No. 3, 312–318 (2012).
2. V. I. Dmitriev and J. G. Ingtem, "A two-dimensional minimum-derivative spline," *Comp. Math. Model.,* **21**, No. 2, 206–211 (2010).
3. J. G. Ingtem, "Minimum-derivative spline function in interpolation and approximation problems," *Vestn. MGU, Vychisl. Matem. Kibern.,* No. 4, 16–27 (2008).
4. V. I. Dmitriev and J. G. Ingtem, "Solving an integral equation of the first kind by spline approximation," *Comp. Math. Model.,* **15**, No. 2, (2004).
5. A. N. Tikhonov, "Ill-posed problems and methods of their solution," in: *Methods of Solution of Ill-Posed Problems and Their Application: Proc. All-Union School of Young Scientists* [in Russian], MGU, Moscow (19740, pp. 6–11.
6. Yu. S. Volkov and V. L. Miroshnichenko, "Approximation of derivatives by an interpolation spline with a jump," *Mat. Zametki,* 89, No. 1, 127–130 (2011).