

6. E. A. Artyukhin and A. V. Nenarokomov, *Gagarin Scientific Lectures on Cosmonautics and Aviation for 1985* [in Russian], Moscow (1986), pp. 160-161.
7. E. A. Artyukhin and S. A. Budnik, *Gagarin Scientific Lectures on Cosmonautics and Aviation for 1986* [in Russian], Moscow (1987), pp. 138-139.
8. S. Patankar, *Numerical Methods for Solving Heat Transfer and Fluid Dynamics Problems* [Russian translation], Moscow (1984).

CONSTRUCTION OF SMOOTHING SPLINES BY LINEAR PROGRAMMING

METHODS

A. G. Pogorelov

UDC 517.536.946

The mathematical questions and algorithms for constructing n -th order smoothing splines by means of experimental (kinetic) dependences are elucidated.

1. Let the function $f(x) \in C^Q[X]$, $Q \geq n$ that takes on the approximate values $f(x_1) + \delta_1, \dots, f(x_N) + \delta_N$ be given discretely with the errors $\delta_1, \dots, \delta_N$ at the nodes x_1, \dots, x_N on the segment $X \subset R$. It is required to approximate the function $f(x)$ in each interval $[x_i, x_{i+1})$, $i = \overline{1, N-1}$ by a polynomial of n -th degree, $n \geq 3$:

$$y_i(x) = a_{0i} + a_{1i}x + a_{2i}x^2 + \dots + a_{ni}x^n, \quad x \in [x_i, x_{i+1}) \quad (1)$$

so as to satisfy the requirements [1-6]: I) fusion of the spline derivatives at the mesh nodes $S = \{x_1, \dots, x_N\}$ up to the $(n-1)$ order

$$\begin{cases} a_{0i} + a_{1i}x_i + a_{2i}x_i^2 + \dots + a_{ni}x_i^n = a_{0,i+1}, \\ \dots \\ (n-1)! a_{n-1,i} + n! a_{ni}x_i = a_{n-1,i+1}, \quad i = \overline{1, N-2}; \end{cases} \quad (2)$$

II) the requirement of minimal variation of the $(n-1)$ -derivative of $y_i(x)$ (i.e., $\int_{x_i}^{x_N} (y^{(n-1)}(x))^2 dx \rightarrow \min$), corresponding to condition $|a_{v,i}| \rightarrow \min$, $v = n-1, n$, $i = \overline{1, N-1}$, in order to avoid oscillating behavior of the graph of the spline; III) location of the spline graph within the error corridor:

$$\begin{cases} |f_\delta(x_i) - a_{0i}| \leq \delta_i, \quad i = \overline{1, N-1}, \\ |f_\delta(x_N) - a_{0,N-1} - a_{1,N-1}x_N - \dots - a_{n,N-1}x_N^n| \leq \delta_N. \end{cases} \quad (3)$$

2. Conditions I and III yield the search domain for the interval values of the spline approximation coefficients by the system of constraints

$$\begin{cases} a_{0i} \leq f_\delta(x_i) + \delta_i, \\ -a_{0i} \leq -f_\delta(x_i) + \delta_i, \\ a_{0i} + a_{1i}x_{i+1} + a_{2i}x_{i+1}^2 + \dots + a_{ni}x_{i+1}^n - a_{0,i+1} = 0, \\ \dots \\ (n-1)! a_{n-1,i} + n! a_{ni}x_{i+1} - a_{n-1,i+1} = 0, \quad i = \overline{1, N-2}, \\ a_{0,N-1} \leq f_\delta(x_{N-1}) + \delta_{N-1}, \end{cases} \quad (4)$$

N. D. Zelinskii Institute of Organic Chemistry, Academy of Sciences of the USSR, Moscow. Translated from *Inzhenerno-Fizicheskii Zhurnal*, Vol. 56, No. 3, pp. 471-477, March, 1989. Original article submitted April 18, 1988.

$$\begin{cases} -a_{0,N-1} \leq -f_{\delta}(x_{N-1}) + \delta_{N-1}, \\ a_{0,N-1} + a_{1,N-1}x_N + a_{2,N-1}x_N^2 + \dots + a_{n,N-1}x_N^n \leq f_{\delta}(x_N) + \delta_N, \\ -(a_{0,N-1} + a_{1,N-1}x_N + a_{2,N-1}x_N^2 + \dots + a_{n,N-1}x_N^n) \leq -f_{\delta}(x_N) + \delta_N. \end{cases} \quad (4)$$

Since the coefficients a_{ij} , $i = \overline{0, n}$; $j = \overline{1, N-1}$ can have different signs and the standard linear programming problem to which obtaining the interval estimates for a_{ij} reduces has just nonnegative solutions, we set $a_{ij} = a'_{ij} - a''_{ij}$, where $a'_{ij}, a''_{ij} \geq 0$. Then condition II results in the requirement of minimum of the absolute value $|a_{vj}| = |a'_{vj} - a''_{vj}|$, $v = n-1, n$, in each interval $[x_j, x_{j+1})$, $j = \overline{1, N-1}$. Obtaining the interval estimates for a_{ij} with requirements I-III taken into account can be realized by different means, for instance: A) determination of the coefficients a_{vj} , $v = n-1, n$, $j = \overline{1, N-1}$, initially and then all the rest; a_{ij} , $i = \overline{0, n-2}$, $j = \overline{1, N-1}$; B) simultaneous determination of all coefficients by using the multiparametric regularization method [7] for the components of the solution a_{vj} , $v = n-1, n$, $j = \overline{1, N-1}$.

MODIFICATION A

In each interval $[x_j, x_{j+1})$ we find minimal values of the coefficients a_{vj}^0 , $v = n-1, n$, $j = \overline{1, N-1}$, in absolute value, for which we solve two problems: Maximize $z_{1vj} = a_{vj}$ under the constraints (4) and maximize $z_{2vj} = -a_{vj}$ also under the constraints (4). Then taken as a_{vj}^0 , $v = n-1, n$, should be the minimal value in absolute value of $\{|z_{1vj\max}|, |z_{2vj\max}|\}$, i.e., $a_{vj}^0 = \text{sign } a_{vj}^0 |a_{vj}^0|$, where $|a_{vj}^0| = \min\{|z_{1vj\max}|, |z_{2vj\max}|\}$, $v = n-1, n$. Furthermore, we find the uniformly minimal value $a_{vj\min}$, $v = n-1, n$, in absolute value in the segment $[x_1, x_N]$ as $a_{vj\min} = \alpha a_{vj}^0$, where the proportionality factor α is determined from the solution of the problem to maximize $z_3 = -\alpha$ under the constraints (4) but in which all the monomials $a_{vj}x_j^m$ are replaced by $\alpha a_{vj}x_j^m$, $v = n-1, n$, $j = \overline{1, N-1}$, $m = \overline{0, n}$.

Afterwards we proceed to obtain interval estimates for all the other coefficients a_{ij} , $i = \overline{0, n-2}$, $j = \overline{1, N-1}$ for which $\bar{z}_{ij} = a_{ij}$, $i = \overline{0, n-2}$, $j = \overline{1, N-1}$ must be maximized under the constraints (4) but in which all the monomials $a_{vj}x_j^m$, $v = n-1, n$, are replaced by the quantities $z_{3\max} a_{vj}^0 x_j^m$ already known and transposed, respectively, into the right sides of the constraints, and also to maximize $\bar{z}_{ij} = -a_{ij}$ under the same constraints. Then the desired interval estimates for a_{ij} , $i = \overline{0, n-2}$, $j = \overline{1, N-1}$ are determined as

$$a_{ij} = \begin{cases} \leq z_{ij\max}^+ & \text{for } z_{ij\max}^+ > 0, \leq \bar{z}_{ij\max}^- & \text{for } \\ \bar{z}_{ij\max}^- < 0, \geq z_{ij\max}^- & \text{for } \bar{z}_{ij\max}^- > 0, \\ \geq z_{ij\max}^+ & \text{for } z_{ij\max}^+ < 0, \end{cases} \quad i = \overline{1, n-2}, j = \overline{1, N-1}. \quad (5)$$

MODIFICATION B

To find the interval estimates by the method of linear programming with the requirements I-III taken into account, we apply multiparametric regularization to obtain solutions with minimal projection norm in the solution subspace defined by the coefficients a_{vj} , $v = n-1, n$, $j = \overline{1, N-1}$. Seeking the solution $a_{(r)} \in \mathbb{R}^p$ with minimal projection norm in the subspace \mathbb{R}^p , $r \leq p$ (norm of the vector $(a_{(r)})^T = (a_{k+1}, \dots, a_{k+r})$, $0 \leq k \leq p, 1 \leq r \leq p-k$), by the multiparametric regularization method for the linear systems $X_{(N-p) \times (p-1)} a_{(p-1)} = y_{(N-1)}$ or the linear programming problem $\max_a Ca$ under the constraints $Xa \leq y$ (the dimensionalities of X , a and y are the same)

reduces by analogy with [7] to the solution, respectively, of systems $XW_{(r)}u = y$ or $\max_u CW_{(r)}u$

under the constraints $XW_{(r)}u \leq y$, where $W_{(r)}$ is the matrix, $0 \leq k \leq p, 1 \leq r \leq p-k$, of form

$$W_{(r)} = \begin{bmatrix} E_{(k)} & 0_{(k \times (\max\{N, p\} - k))} \\ \hline X_{k+1, k+r} & 0_{(r \times (p-N)_+)} \\ \hline 0_{(p-k-r) \times (k+r)} & E_{(p-k-r)} & 0_{(p-k-r) \times (p-N)_+} \end{bmatrix}, \quad (6)$$

As in modification A the values a_i can here be estimated as $a_i = (a_{i_{\max}} + a_{i_{\min}})/2$.

3. It is required to determine the contribution of each node x_j with the value $f_\delta(x_j) + \delta_j$ from the network of nodes S as well as the errors δ_j , $j = \overline{1, N}$ in the values of the upper and lower bounds of the interval estimates of the coefficients a_{ij} in order to construct the optimal network $S^* \subset [x_1, x_N]$ from the condition of minimum length of the interval estimate $\Delta a_l = |a_{l_{\max}} - a_{l_{\min}}|$ of the given coefficient a_l , $l = \overline{1, (n+1)(N-1)}$, i.e.,

$$\Delta a_l \rightarrow \min, \quad (10)$$

or from the condition of minimum sum of the lengths of the interval estimates for several or

all the coefficients $\sum_{i=l}^{l+q} \Delta a_i$, $l \geq 1$, $l+q \leq (n+1)(N-1)$, i.e.,

$$\sum_{i=l}^{l+q} \Delta a_i \rightarrow \min. \quad (11)$$

To estimate these contributions as well as the contributions of the conditions for fusion of the derivatives (2) on the boundaries of the interval estimates for a_{ij} it is required to solve problems dual to (4) and (5) (modification A) or to (8) and (9) (modification B).

Modification A. We obtain the contributions mentioned from solutions of the problem:

Minimize $z_l = y^T B_l$, where

$$\begin{aligned} y^T B_l = & (f_\delta(x_1) + \delta_1) \bar{b}'_1 - (f_\delta(x_1) - \delta_1) \bar{b}'_1 - (a_{n-1,1} x_2^{n-1} + a_{n1} x_2^n) b'_{0,1} - ((n-1) a_{n-1,1} x_2^{n-2} + n a_{n1} x_2^{n-1}) b'_{1,1} + \dots + \\ & + (a_{n-1,2} - (n-1) a_{n-1,1} - n a_{n1} x_2) b'_{n-1,1} + \dots + (f_\delta(x_{N-2}) + \delta_{N-2}) \bar{b}'_{N-2} - (f_\delta(x_{N-2}) - \delta_{N-2}) \bar{b}'_{N-2} - \\ & - (a_{n-1,N-2} x_{N-1}^{n-1} + a_{n,N-2} x_{N-1}^n) b'_{0,N-2} - ((n-1) a_{n-1,N-2} x_{N-1}^{n-2} + n a_{n,N-2} x_{N-1}^{n-1}) b'_{1,N-2} + \dots + \\ & + (a_{n-1,N-1} - (n-1) a_{n-1,N-2} - n a_{n,N-2} x_{N-1}) b'_{n-1,N-2} + (f_\delta(x_{N-1}) + \delta_{N-1}) \bar{b}'_{N-1} - (f_\delta(x_{N-1}) - \delta_{N-1}) \bar{b}'_{N-1} + \\ & + (f_\delta(x_N) + \delta_N - a_{n-1,N-1} x_N^{n-1} - a_{n,N-1} x_N^n) \bar{b}'_N - (f_\delta(x_N) - \delta_N - a_{n-1,N-1} x_N^{n-1} - a_{n,N-1} x_N^n) \bar{b}'_N, \end{aligned}$$

under the constraints

$$(X_{(\overline{1, n-1, n})})^T B_l \geq e_l, \quad l = \overline{1, (n-1)(N-1)}. \quad (12)$$

where $B_l^T = (\bar{b}'_1, \bar{b}'_1, b'_{0,1}, \dots, b'_{n-1,1}, \bar{b}'_N)$ is the vector of the contributions, $e_l = (\overline{0 \dots 0 1 0 \dots 0})$; $X_{(\overline{1, n-1, n})}$ denotes the matrix X without columns corresponding to the coefficients a_{vj} , $v = n-1, n$,

$j = \overline{1, N-1}$. Then the components of the vector $\hat{B}_{l_{\min}}: \hat{b}'_{l_{\min}}, \hat{b}'_{l_{\min}}$ are contributions of the quantities $f_\delta(x_j) + \delta_j$ and $f_\delta(x_j) - \delta_j$ at the upper bound of values of the component $a_{l_{\max}}$ of the coefficient vector a (the coefficients $a_{n-1,1}, a_{n1}, \dots, a_{n-1,N-1}, a_{n,N-1}$ are not components of a) and \hat{b}'_{ij} is the contribution of the condition for fusion of the i -th derivative at the j -th node of the network S . Hence, the contribution $f_\delta(x_j)$ to $a_{l_{\max}}$ is determined as $(\hat{b}'_{i_{\min}} + \hat{b}'_{j_{\min}})/2$, while the values of the errors δ_j are as $(\hat{b}'_{i_{\min}} - \hat{b}'_{j_{\min}})/2$. The contributions of these same quantities are estimated analogously at the lower bound of the component $a_{l_{\min}}$ of the coefficients vector of the spline a : Minimize $\bar{z}_l = y^T \bar{B}_l$ under the constraints

$$(X_{(\overline{1, n-1, n})})^T \bar{B}_l \geq -e_l, \quad l = \overline{1, (n-1)(N-1)}. \quad (13)$$

Then the components of the vector of the solution $\check{B}_{l_{\min}}: \check{b}'_{l_{\min}}, \check{b}'_{l_{\min}}, \check{b}'_{ij_{\min}}$, $l = \overline{1, (n-1)(N-1)}$, $i = \overline{0, n-1}$, $j = \overline{1, N}$, are contributions, respectively, of $f_\delta(x_j) + \delta_j$, $f_\delta(x_j) - \delta_j$ and the fusion condition for the i -derivative at the j -node of the mesh S at the lower value of the component $a_{l_{\min}}$ of the coefficients vector of the spline a . Then the contributions of the quantities $f_\delta(x_j)$, δ_j to value $a_{l_{\min}}$ are determined as $(\check{b}'_{i_{\min}} + \check{b}'_{j_{\min}})/2$ and $(\check{b}'_{i_{\min}} - \check{b}'_{j_{\min}})/2$ respectively.

Modification B. We obtain estimates of the desired contributions from the solutions of the problems dual to (8) and (9): minimize $z_l^+ = y^T B_l$ under the constraints

$$W_{(n-1,n)}^T X^T B_l \geq W_{(n-1,n)}^T C_l, C_l = \overbrace{(0 \dots 0 1 0 \dots 0)}^l, l = \overline{1, (n+1)(N-1)}, \quad (14)$$

and also minimize $\bar{z}_l = y^T B_l$ under the constraints

$$W_{(n-1,n)}^T X^T B_l \geq -W_{(n-1,n)}^T C_l, l = \overline{1, (n+1)(N-1)}. \quad (15)$$

Let $\hat{B}_{l_{\min}}$ and $\check{B}_{l_{\min}}$ denote the solutions of the problems (14) and (15). Then the desired contributions of the nodes of the network S, the errors δ_j , and the conditions for fusion of the derivatives at the boundaries of the interval estimates for the spline coefficients, including the coefficients $a_{n-1,j}$, $j = \overline{1, N-1}$ in this case, are determined by the components of the vectors $\hat{B}_{l_{\min}}$ and $\check{B}_{l_{\min}}$.

In conclusion, we note that the algorithms considered are general in nature and can be applied for the construction of splines of different orders and defects on the basis of other basis functions; questions of the existence and uniqueness of the appropriate splines do not here enter within the framework of this report.

NOTATION

δ_i , error of giving a function at the i -node; x_i , coordinate of the argument at the i -node; $X = [x_1, x_N]$, segment on which the function is given discretely; R , a one-dimensional axis; $f(x) \in C^Q[X]$, a Q times differentiable function in the segment X ; $f(x_i) + \delta_i$, $f_\delta(x_i)$, values of the function in the i -node aggravated by errors; n , order of the polynomial spline; $y_i(x)$, running value of the approximating polynomial between two nodes; a_{ij} , $i = \overline{0, n}$, $j = \overline{1, N-1}$, $a_{\nu j}$, coefficients of the approximating polynomial between two nodes; S , node network; a'_{ij} , a''_{ij} , terms of the difference representation of the spline coefficients; $a_{\nu j}^0$, minimal value, in absolute value, of the coefficients $a_{\nu j}$; $z_{1\nu j}$, $z_{2\nu j}$, z_{ij} , z_{ij} , $z_{3\max}$, z_i , z_i , z_i , z_i are target functions of the appropriate linear programming problems; α , constant factor; $a_{\nu j_{\min}}$, uniformly minimal, in absolute value, values of the coefficients $a_{\nu j}$ in the coefficients X ; RP , RF , Euclidean spaces of dimensionality p and r ; X , matrix of the left side of the system of linear algebraic equations; a , vector of the desired unknowns of the system of linear equations; y , vector of the free terms (the right sides) of the system of linear equation; $W(r)$, matrix of the mapping of the space of solutions of the system of linear equations into control space in the multiparametric regularization procedure; u , vector of the control (regularization) parameters; $B_l = \{B_l^+, B_l^-\}$, vector of contributions of the quantities $f_\delta(x_i) + \delta_i$, $f_\delta(x_i) - \delta_i$ and fusion conditions for derivatives in the interval estimates (their upper and lower bounds) of the smoothing spline coefficients.

LITERATURE CITED

1. C. H. Reinsch, Numer. Math., 10, 177-183 (1967).
2. R. Varga, Functional Analysis and Theory of Approximation in Numerical Analysis [Russian translation], Moscow (1974).
3. S. V. Stechkin and Yu. N. Subbotin, Splines in Computational Mathematics [in Russian], Moscow (1976).
4. V. A. Morozov, Zh. Vychisl. Mat. Mat. Fiz., 11, No. 3, 545-558 (1971).
5. A. I. Grebennikov, Method of Splines and Solution of Incorrect Problems of the Theory of Approximations [in Russian], Moscow (1983).
6. V. A. Vasilenko, Spline-Functions: Theory, Algorithms, Programs [in Russian], Novosibirsk (1983).
7. A. V. Chechkin, Dokl. Akad. Nauk SSSR, 252, No. 4, 807-810 (1980).