

Acta Universitatis Palackianae Olomucensis. Facultas Rerum
Naturalium. Mathematica

Jiří Kobza

Quartic splines with minimal norms

Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica, Vol. 40 (2001), No. 1, 103--124

Persistent URL: <http://dml.cz/dmlcz/120425>

Terms of use:

© Palacký University Olomouc, Faculty of Science, 2001

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>



Quartic Splines with Minimal Norms ^{*}

JIŘÍ KOBZA

*Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kobza@risc.upol.cz*

(Received January 4, 2001)

Abstract

Function values interpolating splines of the odd degree $2n - 1$ with special boundary conditions are known to minimize the L_2 -norm of the n -th derivative on some wide class of interpolants. Similar extremal property have special even degree splines interpolating the values of the derivative or the mean values. When we restrict the minimization to the linear space of quartic splines on the given knotset only, we can use the spline free parameters to find the interpolating spline with the minimal value of the norm of the user's interest. It could be sometimes more easy for the user to determine the proper spline norm from the geometry of the problem solved rather than to find the corresponding boundary conditions for the interpolant he searches.

Key words: Quartic spline, optimal spline, interpolating spline with minimal norm.

2000 Mathematics Subject Classification: 41A15, 65D05

1 Introduction

Let us have given the spline knotset $\mathbf{x} = \{x_i, i = 0(1)n+1\}$ on the real axis with stepsizes $h_i = x_{i+1} - x_i$ and the prescribed values $\mathbf{g} = \{g_i, i = 0(1)n\}$, which can be the function values prescribed in points of interpolation (FVI problem)

^{*}Supported by the Council of Czech Government, J 14/98:153100011.

$\mathbf{t} = \{t_i, x_i < t_i < x_{i+1}, i = 0(1)n\}$ or the mean values prescribed on the intervals $[x_i, x_{i+1}]$, $i = 0(1)n$ (MVI problem). The set of quartic splines $s(x) \in C^3$ on a given knotset \mathbf{x} forms a linear space. The subspace of quartic splines which interpolate given function or mean values \mathbf{g} has four free parameters, which can be used to fulfil some boundary conditions (see e.g. [2], [7]). Sometimes it is not easy for the user to find the boundary conditions corresponding to his problem—he usually searches for some satisfactory smooth interpolating curve. Such geometric criterion is used e.g. in the known notion of natural, periodic, complete cubic spline, (see [2]) minimizing the L_2 -norm of the spline second derivative. Similar results are known for quadratic and quartic splines interpolating mean values (see [7], [9]). The minimization is considered here on some W_1^2 or W_2^2 classes of interpolating functions. When we restrict ourself to the subsets of interpolating splines only in the problem considered, we can use spline free parameters to find interpolating spline with minimal norm, which can be chosen according to the geometry or physical meaning of the problem (errors or stability problems—norm of function values, minimal energy—norm of the first derivative, smoothness of the process—norm of the second derivative; corresponding vector norms on sufficiently fine knotsets can approximate above mentioned norms). In the following we will give some overview of many possibilities how to compute quartic splines with optimal parameters. We shall consider in the FVI problem the case $t_i \neq x_i$ only, which seems to us as the most suitable in the practice (symmetry, stability, four free parameters).

For the functionals optimized we will therefore choose the different norms of the spline k -th derivative ($k = 0, 1, 2, 3$) or the vectors of their discrete values in knots. We will use then as the most appropriate the local representations containing the interpolated values g_i and the values $m_j = s'(x_j)$, $M_j = s''(x_j)$ in the spline knots—on the boundaries of the local intervals $[x_i, x_{i+1}]$. Such a local representation (denoted as $[\mathbf{g}, \mathbf{m}, \mathbf{M}]$) we can write (see [6]) as

$$s(x) = \psi(u)g_i + h_i[\varphi_0^1(u)m_i + \varphi_1^1(u)m_{i+1}] + h_i^2[\varphi_0^2(u)M_i + \varphi_1^2(u)M_{i+1}] \quad (1)$$

where the local variable $u = (x - x_i)/h_i$, local stepsize $h_i = x_{i+1} - x_i$ are used together with cardinal basis interpolatory functions ψ, φ_i^j , which for function values interpolation (FVI) problem with $d_i = (t_i - x_i)/h_i$ are

$$\begin{aligned} \psi(u) &= 1, \\ \varphi_0^1(u) &= u - d_i - (u^3 - d_i^3) + (u^4 - d_i^4)/2, \\ \varphi_1^1(u) &= u^3 - d_i^3 - (u^4 - d_i^4)/2, \\ \varphi_0^2(u) &= (u^2 - d_i^2)/2 - 2(u^3 - d_i^3)/3 + (u^4 - d_i^4)/4, \\ \varphi_1^2(u) &= -(u^3 - d_i^3)/3 + (u^4 - d_i^4)/4. \end{aligned} \quad (2)$$

For the mean value interpolation (MVI) problem we obtain

$$\begin{aligned} \psi(u) &= 1, & \varphi_0^1(u) &= -\frac{7}{20} + u - u^3 + \frac{1}{2}u^4, \\ \varphi_1^1(u) &= -\frac{3}{20} + u^3 - \frac{1}{2}u^4, & \varphi_1^2(u) &= \frac{1}{30} - \frac{1}{3}u^3 + \frac{1}{4}u^4, \\ \varphi_0^2(u) &= -\frac{1}{20} + \frac{1}{2}u^2 - \frac{2}{3}u^3 + \frac{1}{4}u^4. \end{aligned} \quad (3)$$

Such basis functions we can compute also for another spline local representations with parameters $[\mathbf{g}, \mathbf{s}, \mathbf{m}]$, $[\mathbf{g}, \mathbf{s}, \mathbf{M}]$, $[\mathbf{g}, \mathbf{s}, \mathbf{T}]$, $[\mathbf{g}, \mathbf{m}, \mathbf{T}]$ (see [6]). The $[\mathbf{g}, \mathbf{m}, \mathbf{T}]$ local representation will be used in the proof of Theorem 2.

Given the interpolated values g_i , the local parameters m_i, M_i in knots x_i have to satisfy the *spline continuity conditions* (CC) $s(x) \in C^3[x_0, x_{n+1}]$ in case of maximal smoothness $s \in C^3$, which will be considered in this contribution. Such CC can be expressed in various forms—as recurrences in terms of one local parameter (see [4] and section 2), or (more easily) in terms of two local parameters used in local representations mentioned (section 3). The first possibility can be used when we want to minimize the norm of the vector of discrete values of one local parameter in knots. We shall use in this case the notation

$$J_{kd}(s) = \|[s^{(k)}(x_i)]\|_2^2 = \sum_{i=0}^{n+1} (s^{(k)}(x_i))^2, \quad k = 0, 1, 2, 3. \quad (4)$$

The second possibility is more appropriate when we want to minimize the L_2 -norm of some spline derivative—we shall denote the corresponding functionals as

$$J_k(s) = \int_{x_0}^{x_{n+1}} [s^{(k)}(x)]^2 dx, \quad k = 0, 1, 2, 3, \quad (5)$$

which can be (using proper local representation) expressed as quadratic form in the values of the two local parameters used. We will show in the following that the problem of computing optimal local parameters of the quartic spline can be expressed as the quadratic programming problem with equality constraints given by CC. In the most simple cases we can use for its solution pseudoinverse matrix approach, in some another cases (weighted discrete norms) the approach for optimal solutions of difference equations (given by CC) described in [8]; in general cases we can use standard algorithms of quadratic programming.

2 One parameter Continuity Conditions used

2.1 Function values interpolation problem

Using the technique of divided differences (see e.g. [4], [7]) or symbolic computing devices (as Mathematica—see e.g. [5]), we can obtain the CC expressed as recurrences in one local parameter with the coefficients depending on the geometry of the set of spline knots and points of interpolation and with right-hand side coefficients dependent also on prescribed values g_i . In the most frequently used case of equidistant knotset \mathbf{x} with $h_i = h$ and FVI problem with points of interpolation $t_i = (x_i + x_{i+1})/2$, the CC with local parameters

$$g_i = s(t_i), \quad s_i = s(x_i), \quad m_i = s'(x_i), \quad M_i = s''(x_i), \quad T_i = s'''(x_i)$$

can be written for knots x_i , $i = 2(1)n - 1$ as (see [5])

$$\begin{aligned} s_{i-2} + 76s_{i-1} + 230s_i + 76s_{i+1} + s_{i+2} &= \\ &= 16(g_{i-2} + 11g_{i-1} + 11g_i + g_{i+1}), \end{aligned} \quad (6)$$

$$\begin{aligned} m_{i-2} + 76m_{i-1} + 230m_i + 76m_{i+1} + m_{i+2} &= \\ &= \frac{384}{6h}(-g_{i-2} - 3g_{i-1} + 3g_i + g_{i+1}), \end{aligned} \quad (7)$$

$$\begin{aligned} M_{i-2} + 76M_{i-1} + 230M_i + 76M_{i+1} + M_{i+2} &= \\ &= \frac{384}{2h^2}(g_{i-2} - g_{i-1} - g_i + g_{i+1}), \end{aligned} \quad (8)$$

$$\begin{aligned} T_{i-2} + 76T_{i-1} + 230T_i + 76T_{i+1} + T_{i+2} &= \\ &= \frac{384}{h^3}(-g_{i-2} + 3g_{i-1} - 3g_i + g_{i+1}). \end{aligned} \quad (9)$$

Mention please the known fact of identical coefficients on the left sides of these independent recurrences (valid in the case of equidistant set only). When we want to find e.g. the interpolating spline $s(x)$ with minimal value of $J_{1d}(s)$, we can use the pseudoinverse approach (see e.g. [1]) to underdetermined system of linear equations (7) with four free parameters and full row rank matrix to find the solution—the vector $\mathbf{m} = [m_i]$ with minimal value of $J_{1d}(s)$. For computing of the corresponding components of the vector $\mathbf{M} = [M_i]$ in the spline local representation (1) we cannot use then similarly the system (8) (because we have used all free parameters yet in computing parameters m_i). The recurrences (6)–(9) have been obtained as result of some elimination processes and give values s_i, m_i, M_i, T_i relatively independent (we could recognize it on the discontinuity of the spline with parameters computed in such an independent way). When we have computed optimal parameters m_i from the system (7), we have to use for computing the values M_i the formulas from the middle part of the mentioned elimination process—in the equidistant case we obtain e.g.

$$\begin{aligned} M_0 &= [-1216g_0 + 1344g_1 - 128g_3 \\ &\quad + h(-109m_0 - 1006m_1 + 7m_2 + 146m_3 + 2m_4)]/30h^2; \\ M_1 &= [320g_0 - 384g_1 + 64g_3 \\ &\quad + h(5m_0 + 269m_1 - 8m_2 - 73m_3 - m_4)]/30h^2, \\ M_j &= [-64g_{j-2} + 192g_{j-1} - 128g_{j+1} \qquad j = 2(1)n - 1, \\ &\quad + h(-m_{j-2} - 70m_{j-1} + 115m_j + 146m_{j+1} + 2m_{j+2})]/30h^2, \\ M_n &= [-64g_{n-3} - 384g_{n-2} + 448g_n \\ &\quad - h(m_{n-3} + 79m_{n-2} + 452m_{n-1} + 421m_n + 7m_{n+1})]/30h^2, \\ M_{n+1} &= [320g_{n-3} + 1344g_{n-2} - 1664g_n \\ &\quad + h(5m_{n-3} + 386m_{n-2} + 1603m_{n-1} + 1538m_n + 116m_{n+1})]/30h^2. \end{aligned} \quad (10)$$

Similarly, when we want to find spline with minimal value of the functional $J_{2d}(s)$, we have to compute with the pseudoinverse the optimal solution \mathbf{M} of the system (8) and then use these values for computing local parameters m_i

from the explicit relations

$$\begin{aligned}
 m_0 &= \frac{1}{1920h} [3072g_0 - 5568g_1 + 2496g_3 \\
 &\quad - h^2(656M_0 + 3763M_1 + 3975M_2 + 1001M_3 + 13M_4)]; \\
 m_1 &= \frac{1}{1920h} [-3072g_0 + 3648g_1 - 576g_3 \\
 &\quad + h^2(16M_0 + 573M_1 + 905M_2 + 231M_3 + 3M_4)], \\
 m_j &= \frac{1}{1920h} [3072g_{j-2} - 5568g_{j-1} + 2496g_{j+1} \quad j = 2(1)n - 1 \\
 &\quad - h^2(16M_{j-2} + 1203M_{j-1} + 3335M_j + 1001M_{j+1} + 13M_{j+2})], \\
 m_n &= \frac{1}{1920h} [-3072g_{n-3} + 3648g_{n-2} - 576g_n \\
 &\quad + h^2(16M_{n-3} + 1213M_{n-2} + 3465M_{n-1} + 871M_n + 3M_{n+1})], \\
 m_{n+1} &= \frac{1}{1920h} [3072g_{n-3} - 5568g_{n-2} + 2496g_n \\
 &\quad + h^2(-16M_{n-3} - 1203M_{n-2} + 2695M_{n-1} + 1559M_n + 627M_{n+1})] \tag{11}
 \end{aligned}$$

Similar approach we can use in case of minimization of $J_{0d}(s)$, $J_{3d}(s)$. When we have computed e.g. optimal values \mathbf{s} from (6), then the remaining local parameters \mathbf{m} for the local representation $[\mathbf{g}, \mathbf{s}, \mathbf{m}]$ we have to compute from the formula

$$\begin{aligned}
 m_j &= \frac{1}{990h} [-240g_{j-2} - 3344g_{j-1} - 64g_{j+1} \\
 &\quad + 15s_{j-2} + 1184s_{j-1} + 2585s_j + 260s_{j+1} + 4s_{j+2}], \quad j = 2(1)n - 1
 \end{aligned}$$

and the four boundary values from some special formulas.

In case of more general norms

$$J_{kd}^w(s) = \sum_{i=0}^{n+1} w_i [s^{(k)}(x_i)]^2, \quad \|\mathbf{m}\|^2 = \mathbf{m}^T \mathbf{R}_1 \mathbf{m}, \quad \|M\|^2 = M^T \mathbf{R}_2 M \tag{12}$$

with positive weighting coefficients w_i or positive definite matrices $\mathbf{R}_1, \mathbf{R}_2$ we can use the quadratic programming technique or least squares approach described in [8] to find minimum of such functionals under CC conditions (interpreted as difference equations now).

2.2 Mean values interpolation problem

In case of mean values interpolation (MVI) the prescribed values are

$$g_i = \frac{1}{h_i} \int_{x_i}^{x_{i+1}} s(x) dx, \quad h_i = x_{i+1} - x_i, \quad i = 0(1)n. \tag{13}$$

The corresponding CC written with local parameters $\mathbf{s}, \mathbf{m}, \mathbf{M}, \mathbf{T}$ in case of equidistant knotset \mathbf{x} we can choose from the following recursions with $i = 2(1)n - 1$ (see [5] for the general case)

$$s_{i-2} + 26s_{i-1} + 66s_i + 26s_{i+1} + s_{i+2} = 5(g_{i-2} + 11g_{i-1} + 11g_i + g_{i+1}), \quad (14)$$

$$m_{i-2} + 26m_{i-1} + 66m_i + 26m_{i+1} + m_{i+2} = \frac{20}{h}(-g_{i-2} - 3g_{i-1} + 3g_i + g_{i+1}), \quad (15)$$

$$M_{i-2} + 26M_{i-1} + 66M_i + 26M_{i+1} + M_{i+2} = \frac{60}{h^2}(g_{i-2} - g_{i-1} - g_i + g_{i+1}), \quad (16)$$

$$T_{i-2} + 26T_{i-1} + 66T_i + 26T_{i+1} + T_{i+2} = \frac{120}{h^3}(-g_{i-2} + 3g_{i-1} - 3g_i + g_{i+1}). \quad (17)$$

We can again compute one of the vectors \mathbf{s} , \mathbf{m} , \mathbf{M} , \mathbf{T} from corresponding underdetermined system from (14)–(17) using pseudoinverse of the full row rank matrix of the system and then to compute the remaining local parameters from the corresponding explicit relations. When we have computed vector \mathbf{m} with minimal norm and we use local representation $[\mathbf{g}, \mathbf{m}, \mathbf{M}]$, then we can compute \mathbf{M} using explicit formulas

$$\begin{aligned} M_0 &= [-380g_0 + 420g_1 - 40g_3 \\ &\quad + h(-55m_0 - 272m_1 - 21m_2 + 46m_3 + 2m_4)]/12h^2, \\ M_1 &= [100g_0 - 120g_1 + 20g_3 \\ &\quad + h(5m_0 + 73m_1 + 6m_2 - 23m_3 - m_4)]/12h^2, \\ M_j &= [-20g_{j-2} + 60g_{j-1} - 40g_{j+1} \qquad \qquad \qquad j = 2(1)n - 1 \\ &\quad + h(-m_{j-2} - 20m_{j-1} + 33m_j + 46m_{j+1} + 2m_{j+2})]/12h^2, \\ M_n &= [-20g_{n-3} - 120g_{n-2} + 140g_n \\ &\quad - h(m_{n-3} + 29m_{n-2} + 138m_{n-1} + 125m_n + 7m_{n+1})]/12h^2, \\ M_{n+1} &= [100g_{n-3} + 420g_{n-2} - 520g_n \\ &\quad + h(5m_{n-3} + 136m_{n-2} + 483m_{n-1} + 454m_n + 62m_{n+1})]/12h^2. \end{aligned} \quad (18)$$

When we have computed the vector \mathbf{M} with minimal norm, then the corresponding components of the vector \mathbf{m} we can compute as

$$\begin{aligned} m_0 &= [600g_0 - 1020g_1 + 420g_3 \\ &\quad + h^2(-90M_0 - 573M_1 - 641M_2 - 189M_3 - 7M_4)]/240h, \\ m_1 &= [-600g_0 + 780g_1 - 180g_3 \\ &\quad + h^2(10M_0 + 177M_1 + 269M_2 + 81M_3 + 3M_4)]/240h, \\ m_j &= [600g_{j-2} - 1020g_{j-1} + 420g_{j+1} \qquad \qquad \qquad j = 2(1)n - 1 \\ &\quad - h^2(10M_{j-2} + 253M_{j-1} + 561M_j + 189M_{j+1} + 7M_{j+2})]/240h, \\ m_n &= [-600g_{n-3} + 780g_{n-2} - 180g_n \\ &\quad + h^2(10M_{n-3} + 257M_{n-2} + 589M_{n-1} + 161M_n + 3M_{n+1})]/240h, \\ m_{n+1} &= [600g_{n-3} - 1020g_{n-2} + 420g_n \\ &\quad - h^2(10M_{n-3} + 253M_{n-2} + 481M_{n-1} - 131M_n - 73M_{n+1})]/240h. \end{aligned} \quad (19)$$

More generally, when we have computed some of vectors $\mathbf{s}, \mathbf{m}, \mathbf{M}, \mathbf{T}$ with minimal norm, we can compute the remaining parameters from explicit formulas

$$\begin{aligned}
 m_j &= [-70g_{j-2} - 935g_{j-1} - 15g_{j+1} \\
 &\quad + 14s_{j-2} + 397s_{j-1} + 2185s_j + 561s_{j+1} + 45s_{j+2}]/176h, \\
 m_j &= (-g_{j-2} + g_{j+1})/3h - h^2(T_{j-2} + 29T_{j-1} + 90T_j + 29T_{j+1} + T_{j+2})/360, \\
 s_j &= [260g_{j-2} + 300g_{j-1} + 160g_{j+1} \\
 &\quad + h(13m_{j-2} + 314m_{j-1} + 165m_j - 184m_{j+1} - 8m_{j+2})]/720, \\
 s_j &= (g_{j-2} + g_{j+1}) - h^2(M_{j-2} + 27M_{j-1} + 84M_j + 27M_{j+1} + M_{j+2})/120, \\
 s_j &= [-600g_{j-2} + 2520g_{j-1} \\
 &\quad - h^3(5T_{j-2} + 124T_{j-1} + 231T_j + 58T_{j+1} + 2T_{j+2})]/2160, \\
 M_j &= (2g_{j-2} - 3g_{j-1} + g_{j+1})/3h^2 \\
 &\quad + h(2T_{j-2} + 55T_{j-1} + 33T_j - 29T_{j+1} - T_{j+2})/360.
 \end{aligned} \tag{20}$$

($j = 2(1)n - 1$; additional special formulas are needed for boundary values.)

2.3 Existence, uniqueness of the optimal solution

The quadratic functionals we minimize are nonnegative and so there exist their minima. In all cases mentioned here (FVI, MVI problems on equidistant knotset) the two matrices appearing in CC for FVI and MVI problems with various local parameters have full row rank. The pseudoinverse solution with minimal norm for each local parameter is then unique in such problems (see [1]). These uniqueness and the existence of simple explicit formulas for computing the values of the remaining local parameters give us then the proof of the uniqueness of such solution. Even in the general case our functionals $J_{kd}(s)$ are convex functions which are minimized over convex set determined by spline CC for the local parameter chosen—such a problem is known to have a unique minimizer. Thus we have proven the following theorem.

Theorem 1 *In the problems FVI, MVI on the general spline knotset there exists for each functional $J_{kd}(s)$, $k = 0, 1, 2, 3$ the unique quartic interpolatory spline with the minimal value of such functional. The vector of optimal values of one kind of the local parameters of such spline can be computed using pseudoinverse solution of the system of corresponding continuity conditions; the values of the second unknown local parameter used in the local representation has to be computed from complementary formulas as given for equidistant case in (10), (11), (18), (19), (20). We can use some special LSQ approach described in [5] for minimization of functionals of the type (12); we can use the algorithms for the problem of quadratic programming with equality constraints in such or more general cases.*

Remarks The expressions for the general case of CC and functionals for FVI problem on the general knotset are too lengthy to be written and discussed here.

For the user the algorithm described needs only to choose the order of the derivative minimized and to know corresponding continuity conditions—not to

search for complementary boundary conditions. As we can see on the examples below, the difference in the results is to be seen on the boundary intervals mainly.

Example 1 For the monotone (staircase) function values on the equidistant knotset $t = 1 : 2 : 29$, $x = 0 : 2 : 30$,

$$\mathbf{g} = [1, 10, 10, 10, 11, 13, 14, 15, 15, 15, 16, 17, 18, 20, 20]$$

the quartic spline with minimal norm $\|\mathbf{m}\|_2 = 2.43$ ($\|\mathbf{M}\|_2 = 84.3$) is plotted with dashed line on Fig.1; with dash-dot line is here plotted the spline with minimal norm $\|\mathbf{M}\|_2 = 2.29$ ($\|\mathbf{m}\|_2 = 6.33$) and with dotted line the spline with minimal norm of the vector $[\mathbf{m}, \mathbf{M}]$ equal to 6.3 (see Section 4.1).

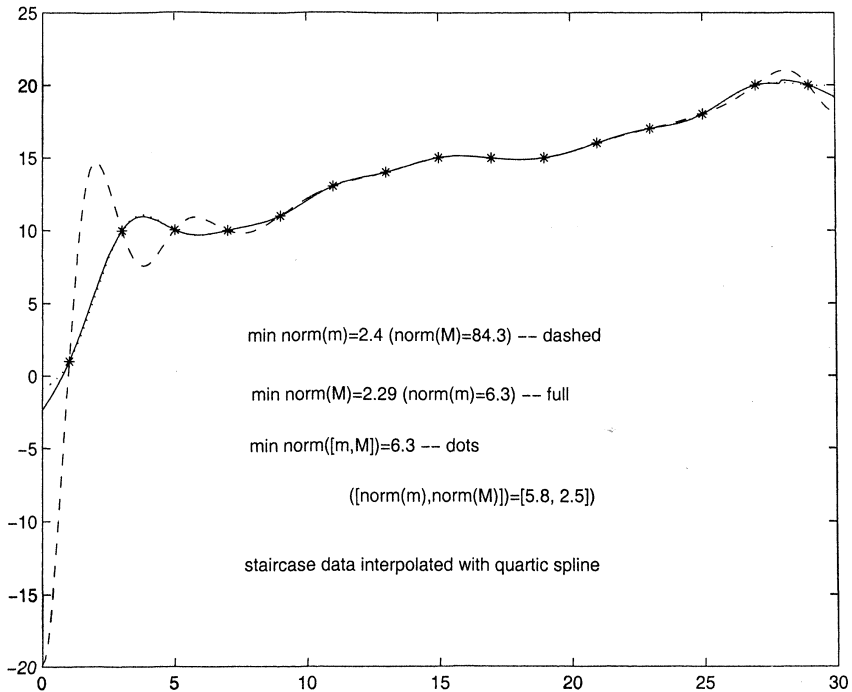


Fig. 1

Example 2 On the equidistant knotset $x = 0 : 2 : 20$ the mean values

$$\mathbf{g} = [5, 1, 3, 4, 7, 13, 8, 11, 15, 9]$$

are prescribed.

The quartic MVI spline with minimal value of the norm $\|\mathbf{m}\|_2 = 7.75$ (and $\|\mathbf{M}\|_2 = 53.1$) is plotted in dots on Fig. 2. Interpolating spline with minimal norm $\|\mathbf{M}\|_2 = 6.47$ ($\|\mathbf{m}\|_2 = 10.2$) is here plotted in dash. With full line is plotted here the spline with minimal norm $\|[\mathbf{m}, \mathbf{M}]\|_2 = 11.5$.

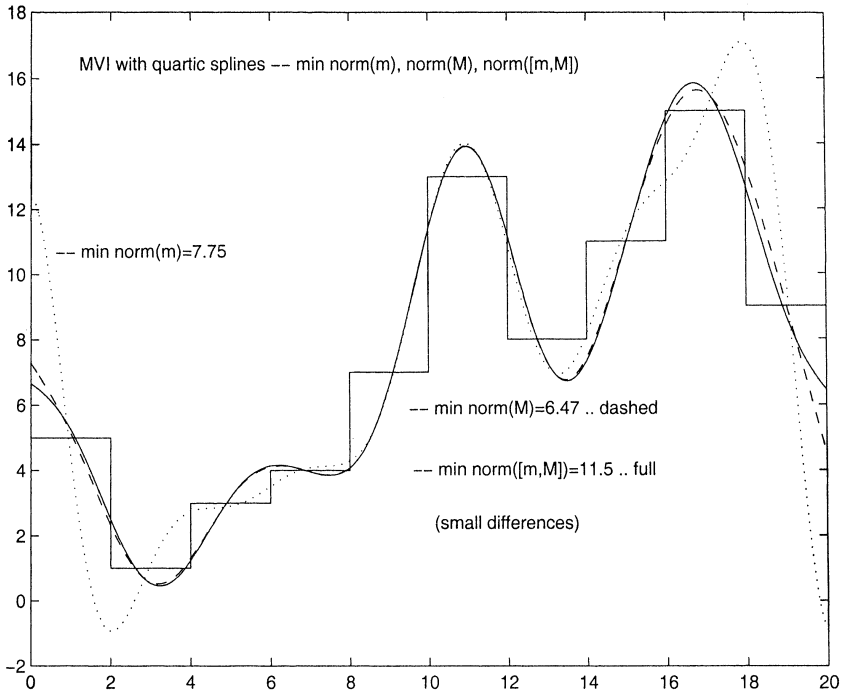


Fig. 2

In both cases we can see the substantial differences near the boundaries only (which is caused by damped error propagation known at these splines—connected with diagonal dominance in CC). More details about the norms of the vectors considered in Examples 1,2 are given in the Table 1.

	FVI – Ex. 1			MVI – Ex. 2		
min	$\ \mathbf{m}\ $	$\ M\ $	$\ [m, M]\ $	$\ m\ $	$\ M\ $	$\ [m, M]\ $
$\ m\ $	2.4255	84.3086	84.3435	7.7538	53.1378	53.7005
$\ M\ $	6.3270	2.2873	6.7278	10.1722	6.4710	12.0560
$\ [m, M]\ $	5.7936	2.4591	6.2939	9.4004	6.6338	11.5055

Table 1.

3 Optimization with two local parameters

3.1 Continuity conditions with two parameters

It is not an easy matter to obtain the continuity conditions (6)–(9), (14)–(17) and completing formulas for the second parameter even for equidistant knotset. We obtain very lengthy formulas in the general case. More simple approach we can obtain when we use the spline local representation (1)–(2) and express

the CC for $s, s^{(3)}$ in each inner spline knot. We obtain then the CC as linear recursions between values of the local parameters \mathbf{m} , \mathbf{M} , with coefficients depending on the geometry of the spline knots and points of interpolation, with components of the right-hand side depending also on prescribed values g_i (for the general case see [5]).

In case of the equidistant knotset and $d_i = (t_i - x_i)/h_i = 1/2$ in *FVI problem* we obtain CC as system of linear recurrences

$$\begin{aligned} \frac{1}{32}(3m_{i-1} + 26m_i + 3m_{i+1}) + \frac{5}{192}h(M_{i-1} - M_{i+1}) &= \frac{1}{h}(g_i - g_{i-1}), \\ \frac{1}{2h}(m_{i-1} - m_{i+1}) + \frac{1}{6}(M_{i-1} + 4M_i + M_{i+1}) &= 0, \quad i = 1(1)n. \end{aligned} \quad (21)$$

These CC form now the system of $2n$ equations with $2n+4$ parameters, with the matrix consisting of four $(n, n+2)$ -matrices with tridiagonal structure. With the notation

$$\mathbf{A} = \begin{bmatrix} 3 & 26 & 3 & & \\ & \ddots & \ddots & \ddots & \\ & & 3 & 26 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & -1 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & 1 \end{bmatrix}, \quad \mathbf{b} = [b_i] = \left[\frac{192}{5h}(g_{i+1} - g_i) \right]$$

we can write these CC in matrix form as

$$\begin{bmatrix} \frac{6}{5}\mathbf{A} & \mathbf{C} \\ \mathbf{C} & \frac{1}{3}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ h\mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}. \quad (22)$$

Let us mention that the matrices \mathbf{A} , \mathbf{B} , \mathbf{C} have full row rank and are constant for different stepsizes h and values \mathbf{g} . The block matrix of the system (22) is also of the full row rank.

In the *MVI problem* we can use the local representation (1), (3) to obtain the CC on the general knotset with $p_i = h_{i-1}/h_i$, $i = 1(1)n$ as

$$\begin{aligned} \frac{9}{2}h_{i-1}m_{i-1} + \frac{21}{2}(h_{i-1} + h_i)m_i + \frac{9}{2}h_{i+1}m_{i+1} \\ + h_{i-1}^2M_{i-1} + \frac{3}{2}(h_i^2 - h_{i-1}^2)M_i - h_i^2M_{i+1} &= 30(g_i - g_{i-1}), \\ m_{i-1} + (p_i^2 - 1)m_i - p_i^2m_{i+1} \\ + \frac{1}{3}h_{i-1}[M_{i-1} + 2(1 + p_i)M_i + p_iM_{i+1}] &= 0. \end{aligned} \quad (23)$$

Using the matrix notation

$$\mathbf{A} = \begin{bmatrix} 9h_0 & 21(h_0 + h_1) & & & & & \\ & 9h_1 & & & & & \\ & & \ddots & & & & \\ & & & 9h_{n-1} & & & \\ & & & & 21(h_{n-1} + h_n) & & \\ & & & & & 9h_n & \end{bmatrix},$$

$$\begin{aligned}
 \mathbf{C} &= \begin{bmatrix} 2h_0^2 & 3(h_1^2 - h_0^2) & -2h_1^2 & & -2h_2^2 \\ & 2h_1^2 & 3(h_2^2 - h_1^2) & & \\ & & \ddots & \ddots & \\ & & & 2h_{n-1}^2 & 3(h_n^2 - h_{n-1}^2) & -2h_n^2 \\ & & & & & \ddots \\ & & & & & & -2h_n^2 \end{bmatrix}, \\
 \mathbf{D} &= \begin{bmatrix} 1 & p_1^2 - 1 & -p_1^2 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & p_n^2 - 1 & -p_n^2 \end{bmatrix}, \\
 \mathbf{B} &= \begin{bmatrix} h_0 & 2h_0(1 + p_1) & & h_0 p_1 & & \\ & \ddots & & \ddots & & \ddots \\ & & h_{n-1} & 2h_n(1 + p_n) & h_{n-1} p_n & \\ & & & & & \end{bmatrix}, \\
 \mathbf{b}^1 &= [b_i^1] = [60(g_i - g_{i-1})], \quad \mathbf{b}^2 = [b_i^2] = \mathbf{0}, \quad i = 1(1)n,
 \end{aligned}$$

we can write with this four $(n, n+2)$ -matrices the CC in the matrix form

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ 3\mathbf{D} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{b}^1 \\ \mathbf{b}^2 \end{bmatrix}. \tag{24}$$

In the equidistant case we can rearrange the CC to simpler and more symmetric form as in the FVI problem; the full row rank of the system is then better to be seen.

3.2 Functionals $J_k(S)$ minimized—FVI problem

Using the local representations (1) we can compute expressions for functionals $J_k(s)$. In case of the FVI problem with $d_i = (t_i - x_i)/h_i = 1/2$ we obtain

$$\begin{aligned}
 J_0(s) &= \int_{x_0}^{x_{n+1}} [s(x)]^2 dx = \sum_{i=0}^n h_i g_i^2 + \frac{9}{80} \sum_{i=0}^n h_i^2 g_i (m_{i+1} - m_i) \\
 &- \frac{7}{480} \sum_{i=0}^n h_i^3 g_i (M_i + M_{i+1}) + \frac{1}{322560} \sum_{i=0}^n h_i^3 (8483m_i^2 + 9914m_i m_{i+1} + 8483m_{i+1}^2) \\
 &\quad + \frac{1}{258048} \sum_{i=0}^n h_i^5 (239M_i^2 - \frac{1962}{5} M_i M_{i+1} + 239M_{i+1}^2) \\
 &\quad + \frac{1}{322560} \sum_{i=0}^n h_i^4 (3119m_i M_i - 2257m_i M_{i+1} + 2257m_{i+1} M_i - 3119m_{i+1} M_{i+1}) \\
 &= \sum_{i=0}^n h_i g_i^2 + \mathbf{p}^T \begin{bmatrix} \mathbf{m} \\ \mathbf{M} \end{bmatrix} + [\mathbf{m}^T, \mathbf{M}^T] \begin{bmatrix} \mathbf{R}_1 & \frac{1}{2}\mathbf{Q} \\ \frac{1}{2}\mathbf{Q}^T & \mathbf{R}_2 \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{M} \end{bmatrix} \tag{25}
 \end{aligned}$$

with the vector \mathbf{p} and tridiagonal matrices $\mathbf{R}_1, \mathbf{R}_2, \mathbf{Q}$ which can be recognized from the explicit expression written above ($\mathbf{R}_1, \mathbf{R}_2$ are symmetric, positive definite—SPD matrices).

with tridiagonal $(n+1, n+1)$ -matrices

$$\mathbf{R}_1 = \begin{bmatrix} h_0^{-1} & -h_0^{-1} & & & & \\ -h_0^{-1} & h_0^{-1} + h_1^{-1} & -h_1^{-1} & & & \\ & \ddots & \ddots & \ddots & & \\ & & -h_{n-1}^{-1} & h_{n-1}^{-1} + h_n^{-1} & -h_n^{-1} & \\ & & & -h_n^{-1} & h_n^{-1} & \end{bmatrix},$$

$$\mathbf{R}_2 = \begin{bmatrix} 4h_0 & -h_0 & & & & \\ -h_0 & 4(h_0 + h_1) & -h_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -h_{n-1} & 4(h_{n-1} + h_n) & -h_n & \\ & & & -h_n & 4h_n & \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & & & & \\ -1 & 0 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 0 & 1 & \\ & & & -1 & -1 & \end{bmatrix}.$$

Let us mention that the matrices \mathbf{R}_1, \mathbf{Q} are singular (one, two eigenvalues equal to zero—row sum is equal to zero vector). So also the compound matrix is singular (it corresponds to the fact that $J_2(s) = 0$ for the data g_i from some linear function).

Similarly we can obtain the functional $J_3(s)$ for the L_2 -norm of the spline third derivative as the quadratic form

$$\begin{aligned} J_3(s) &= \int_{x_0}^{x_{n+1}} [s'''(x)]^2 dx = \sum_{i=0}^n \frac{4}{h_i^3} [3(m_i - m_{i+1})^2 \\ &+ h_i^2 (M_i^2 + M_i M_{i+1} + M_{i+1}^2) + 3h_i (m_i - m_{i+1})(M_i + M_{i+1})] \\ &= 4 [\mathbf{m}^T, \mathbf{M}^T] \begin{bmatrix} 3\mathbf{R}_1 & \frac{3}{2}\mathbf{Q} \\ \frac{3}{2}\mathbf{Q}^T & \frac{1}{2}\mathbf{R}_2 \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{M} \end{bmatrix}, \end{aligned} \tag{28}$$

with matrices

$$\mathbf{R}_1 = \begin{bmatrix} h_0^{-3} & -h_0^{-3} & & & & \\ -h_0^{-3} & h_0^{-3} + h_1^{-3} & -h_1^{-3} & & & \\ & \ddots & \ddots & \ddots & & \\ & & -h_{n-1}^{-3} & h_{n-1}^{-3} + h_n^{-3} & -h_n^{-3} & \\ & & & -h_n^{-3} & h_n^{-3} & \end{bmatrix},$$

$$\mathbf{R}_2 = \begin{bmatrix} 2h_0^{-1} & h_0^{-1} & & & & \\ h_0^{-1} & 2(h_0^{-1} + h_1^{-1}) & h_1^{-1} & & & \\ & \ddots & \ddots & \ddots & & \\ & & h_{n-1}^{-1} & 2(h_{n-1}^{-1} + h_n^{-1}) & h_n^{-1} & \\ & & & h_n^{-1} & 2h_n^{-1} & \end{bmatrix},$$

functions for the FVI and MVI problems in (2), (3), we will find that the corresponding functions differ with additive constants only in general and so they have identical derivatives. This fact results then in identical structure of matrices in matrix forms of functionals mentioned. As a consequence we obtain then similar results in questions of existence and uniqueness of optimal FVI and MVI quartic splines. We can use this fact also in computational algorithms. Let us remark that the mentioned property of cardinal basis functions does not mean that the corresponding splines differ with an additive constant.

Lemma 1 *The quadratic forms $J_k(s)$, $k = 1, 2, 3$ for quartic FVI and MVI splines on given spline knotset \mathbf{x} are equal for given index k .*

4 Existence and uniqueness of optimal splines

4.1 Minimal norm of vector $[\mathbf{m}; \mathbf{M}]$

We can search also for the spline with minimal norm of the vector $[\mathbf{m}; \mathbf{M}]$. In the case of minimization of its l_2 -norm we can find the pseudoinverse solution to the system of equations (21) for FVI problem, (23) for MVI problem. The full row rank of matrices of this systems ensures us the existence and uniqueness of optimal values for components of the vectors \mathbf{m} , \mathbf{M} . We can consider similarly the minimization of such functionals as $J_1(s) + J_2(s)$.

Lemma 2 *There exist the unique quartic FVI, MVI interpolatory splines on the general knotset with minimal l_2 -norms of the vector $[\mathbf{m}; \mathbf{M}]$. They can be computed with pseudoinverse approach to the CC (21) or (23).*

The results for such optimal splines from the Examples 1, 2 we can see in the Table 1.

4.2 Functionals $J_k(s)$

Following the definition, all mentioned functionals are nonnegative quadratic forms and therefore there exists their minimum. Positive definiteness of the matrix of the quadratic form is known to be sufficient condition for the uniqueness of minima. The functionals $J_0(s)$ in both FVI and MVI problems are equal to zero in the case $s(x) \equiv 0$ only—it proves their positive definiteness and uniqueness of the minima. For the data not allowing constant FVI ($g_i = \text{const.}$) or MVI spline ($g_i/h_i = g_{i-1}/h_{i-1}$) we obtain also the positive value of the functionals $J_1(s)$, which proves the positive definiteness of corresponding matrices and uniqueness of the minima in such cases (we have controlled it numerically on many examples). We have mentioned yet the singularity of the matrices of quadratic forms in the functionals $J_2(s)$, $J_3(s)$. We cannot prove the uniqueness of the minima with similar simple arguments now. All the problems mentioned are some quadratic programming problems with equality constraints—in some simple cases we have proved the uniqueness yet in Section 2.3. We have now to use some more detailed technique from optimization theory. The conditions

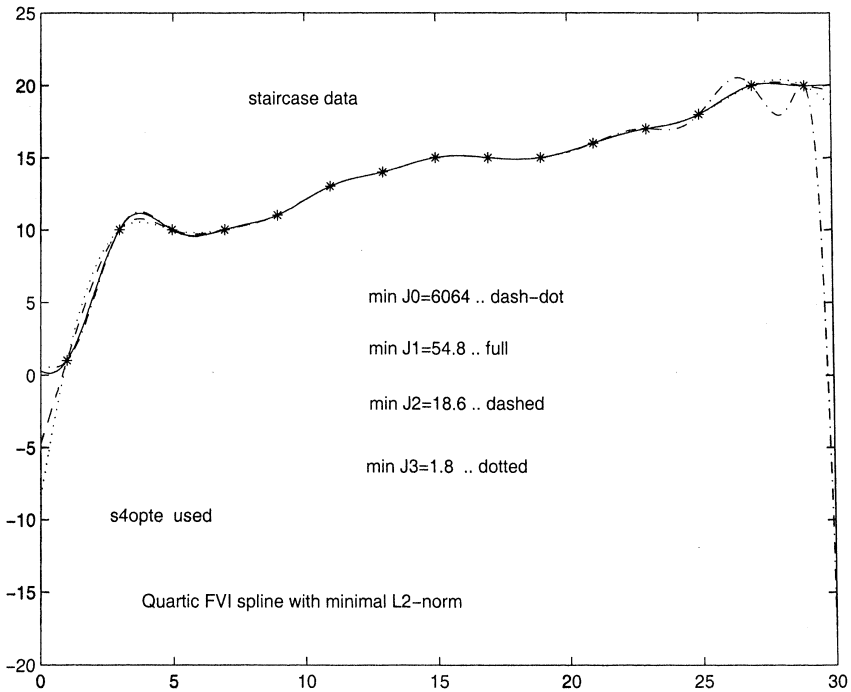


Fig. 3

for the existence and uniqueness of the general quadratic programming problem with equality constraints

$$\min \left\{ \frac{1}{2} \mathbf{p}^T \mathbf{H} \mathbf{p} + \mathbf{g}^T \mathbf{p}; \mathbf{A} \mathbf{p} = -\mathbf{d}, \mathbf{p} \in R^m, \mathbf{A} = (n, m) \right\} \quad (30)$$

were discussed in details in [3], where (Theorem 1.1, Theorem 2.1) the necessary and sufficient conditions for the strong minimizer (unique solution) are stated. The corresponding Kuhn-Tucker matrix

$$\mathbf{K} = \begin{bmatrix} \mathbf{H} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix}$$

with left diagonal block \mathbf{H} taken from minimized quadratic form and the full rank matrix \mathbf{A} from equality constraints is here used and the existence of the strong minimizer (unique minimum) proved under three types of conditions which can be numerically controlled:

- the matrix \mathbf{H} is symmetric, positive definite;
- for matrix \mathbf{Z} of the basis of the null space $N(\mathbf{A})$ of the matrix \mathbf{A} the matrix $\mathbf{Z}^T \mathbf{H} \mathbf{Z}$ is positive definite;

- c) $N(\mathbf{A}) \cap N(\mathbf{H}) = \{0\}$;
- d) the number of negative eigenvalues of \mathbf{K} is equal to the rank of \mathbf{A} and \mathbf{K} is regular.

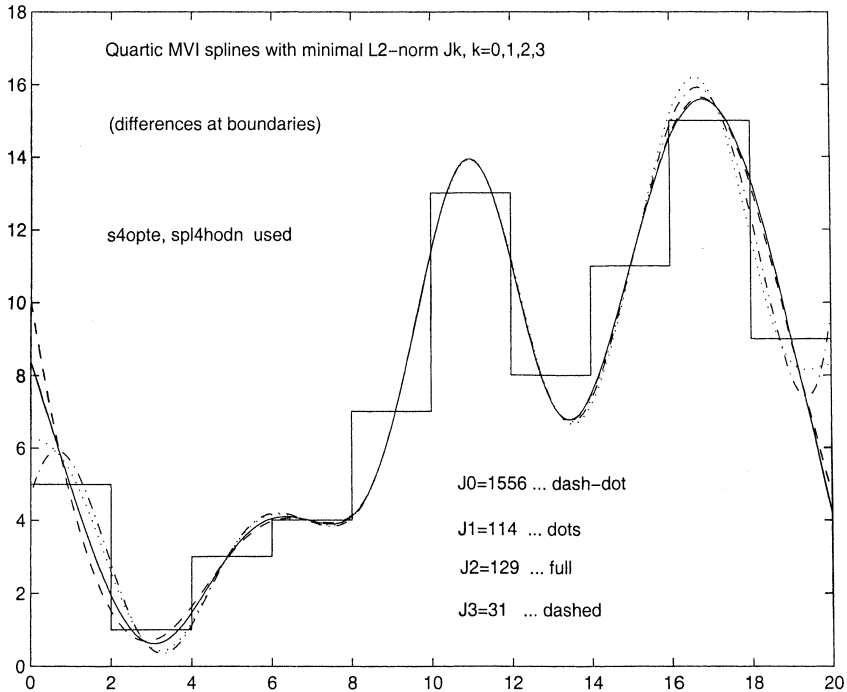


Fig. 4

We have done many computations with Matlab for both FVI and MVI problems with n ranging from 5 to 50 and stepsizes $h = 2, 1, 0.1, 0.01, 0.001$ (the Matlab function $null(A)$ can help to find the matrix \mathbf{Z}). In all cases we have obtained with approaches b), c) the positive results in the question of the uniqueness of the solution minimizing functionals $J_2(s)$, $J_3(s)$ in FVI and MVI problems with positive semidefinite matrices. But it was difficult to find the rigorous proof in the general case. The matrices \mathbf{H} (identical for FVI and MVI problems as stated above) have more simple structure when we instead of the local parameters $[\mathbf{g}, \mathbf{m}, \mathbf{M}]$ use another triplet of parameters $[\mathbf{g}, \mathbf{m}, \mathbf{T}]$ with the vector \mathbf{T} of the third spline derivatives $T_i = s'''(x_i)$ (see [6]). With this local parameters we can write the spline local representation

$$s(x) = \psi(u)g_i + h_i[\varphi_0^1(u)m_i + \varphi_1^1(u)m_{i+1}] + h_i^3[\varphi_0^3(u)T_i + \varphi_1^3(u)T_{i+1}] \quad (31)$$

In the FVI problem on equidistant knotset, $d_i = 1/2$ the $(n, n+2)$ - submatrices are determined as (see [5])

$$\mathbf{A}_{11} = \begin{bmatrix} 1 & 6 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 6 & 1 \end{bmatrix}, \quad \mathbf{A}_{12} = \begin{bmatrix} 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix}$$

$$\mathbf{A}_{21} = -\frac{h^2}{48} \begin{bmatrix} 7 & 18 & 7 & & \\ & \ddots & \ddots & \ddots & \\ & & 7 & 18 & 7 \end{bmatrix}, \quad \mathbf{A}_{22} = -\frac{h^2}{6} \begin{bmatrix} 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \end{bmatrix}.$$

Similar structure of these matrices for MVI problem is described in [5]. We are able to prove now the following Theorem.

Theorem 2 *The quadratic programming problem to find minimum of $J_k(s)$, $k = 2, 3$ under continuity conditions has unique solution in FVI and MVI problems on equidistant knotsets.*

Proof From the description of the coefficients of the matrices \mathbf{A}_{ij} we can see the full row rank of the block matrix of continuity conditions. We shall use the proposition c) and we prove, that the vector from the nullspace of the matrix \mathbf{H} cannot belong to the nullspace of the matrix of the continuity conditions. Let us start with the functional $J_3(s)$, where we have $\mathbf{R}_1 = \mathbf{O}$. The system $\mathbf{R}_2\mathbf{T} = \mathbf{O}$ with regular matrix has only trivial solution $\mathbf{T} = \mathbf{O}$. The nontrivial vector $[\mathbf{m}, \mathbf{T}]$ from the nullspace of the matrix of CC so has to be obtained as the nonzero solution of the overdetermined system of equations $\mathbf{A}_{11}\mathbf{m} = \mathbf{O}$, $\mathbf{A}_{21}\mathbf{m} = \mathbf{O}$. The solutions of such systems are also the solutions of homogeneous linear difference equations—with the roots of characteristic polynomials equal to $-3 + 2\sqrt{2}$, $-3 - 2\sqrt{2}$ in the first part and double root equal to one in the second part of equations. With the exception of the case of two knots such equations have not the common nontrivial (zero) solution.

Let us now consider the case with the functional $J_2(s)$. The system $\mathbf{R}_2\mathbf{T} = \mathbf{O}$ has trivial solution only. If we have $\mathbf{m} \neq \mathbf{O}$, then the solution of $\mathbf{R}_1\mathbf{m} = \mathbf{O}$ is the vector of constants ($m_i = c$) only (proof by induction) with $\mathbf{A}_{21}\mathbf{m} = \mathbf{O}$. The part \mathbf{T} of the nullspace of the matrix of CC with such a vector \mathbf{m} has to be a solution of the system of overdetermined equations

$$\mathbf{A}_{11}\mathbf{T} = -\mathbf{A}_{11}\mathbf{m}, \quad \mathbf{A}_{22}\mathbf{T} = \mathbf{O}.$$

The first system has no trivial solution $\mathbf{T}=\mathbf{O}$. (When we present the solutions of block systems as the solutions of the difference equations with the roots of characteristic polynomials which are again different for each block, we can conclude that such a system has no solution.)

Quite similarly we can prove the uniqueness in the MVI problem. With more detailed technique it is possible to prove in such a way the uniqueness of the solution also with local parameters $[\mathbf{g}, \mathbf{m}, \mathbf{M}]$.

Let us mention that the zero value of the functional $J_k(s)$ we obtain when the data \mathbf{g} correspond to some polynomial of the degree $k - 1$.

We can summarize the results obtained in the subsection 4.2 in the following Theorem.

Theorem 3 *The problem of finding a quartic FVI or MVI interpolatory spline on equidistant knotset with minimal value of the functional $J_k(s)$ has the unique solution for each $k \in \{0, 1, 2, 3\}$. We can compute it with quadratic programming techniques or with some special techniques mentioned above.*

Remark The uniqueness of such optimal splines will hold also for slightly nonequidistant knotsets.

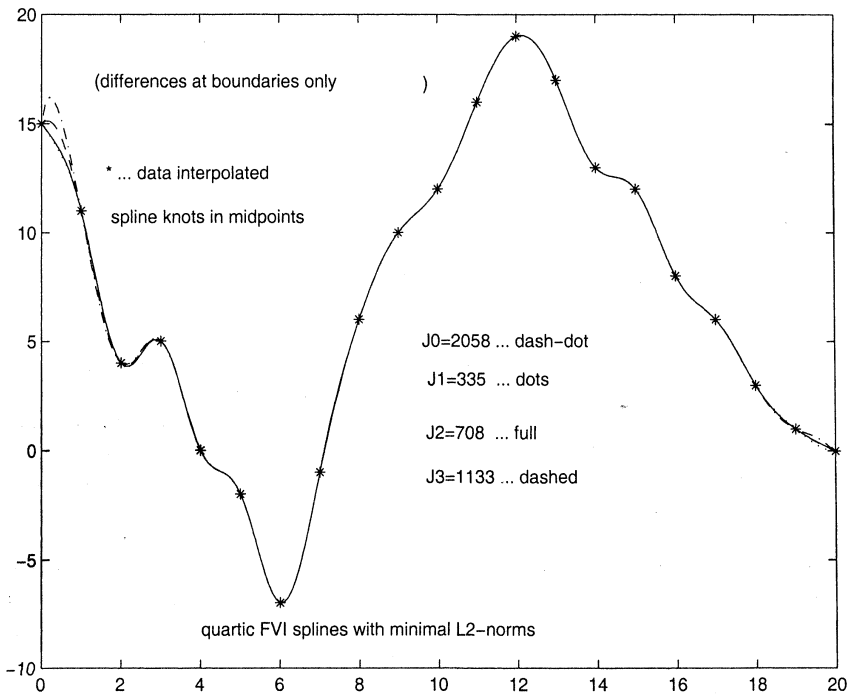


Fig. 5

Examples For the monotone data from the Example 1 (FVI) we have computed the local parameters of the quartic splines with minimal values of the functionals $J_k(s)$, $k = 0, 1, 2, 3$. We can see their plots on Fig. 3 and recognize significant differences in the boundary intervals only (especially in the case of $J_0(s)$).

We have computed also the optimal MVI quartic splines for the mean values given in the Example 2 with minimal values of functionals mentioned above.

The results are to be seen in the Fig. 4—with the significant differences on the boundary only.

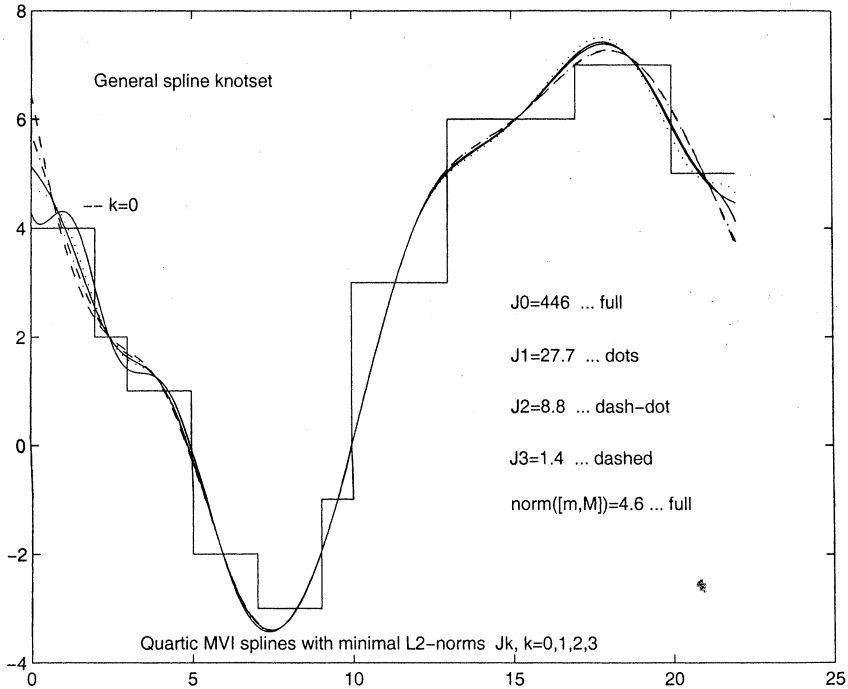


Fig. 6

Example 3 For the equidistant points of interpolation $t = 0 : 1 : 20$, spline knots $x = -0.5 : 1 : 20.5$ and the following general data

$$g = [15, 11, 4, 5, 0, -2, -7, -1, 6, 10, 12, 16, 19, 17, 13, 12, 8, 6, 3, 1, 0]$$

the plots of quartic FVI splines are given in Fig. 5. We can see again the visible differences near the boundaries only. Similar result we can obtain for the spline with minimal norm of the vector $[m, M]$.

Example 4 For the general knotset and MVI problem with

$$x = [0, 2, 3, 5, 7, 9, 10, 13, 17, 20, 22], \quad g = [4, 2, 1, -2, -3, -1, 3, 6, 7, 5]$$

we can find the plots of quartic MVI splines with minimal values of $J_k(s)$, $k = 0, 1, 2, 3$ and minimal value of norm $[m, M]$ in Fig. 6.

All the examples have been computed with special MATLAB M-files **s4opte**, **spl4hodn** worked out by the author. These examples show the visually nice properties of splines with minimal norm of the first or second derivative.

References

- [1] Bjorck, A.: Numerical Methods for Least Squares Problems. *SIAM, Philadelphia*, 1996.
- [2] Boor, C.: A Practical Guide to Splines. *Springer*, 1978.
- [3] Gould, I. M.: *On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem*. *Mathematical Programming* **32** (1985) 90–99.
- [4] Kobza, J.: *Quartic interpolatory splines*. *Studia Univ. Babes-Bolyai, Math.* (1996).
- [5] Kobza, J.: *Spline recurrences for quartic splines*. *Acta Univ. Palacki. Olomuc., Fac. rer. nat.* **34**, *Math.* 63 (1995), 229–236.
- [6] Kobza, J.: *Local representation of quartic splines*. *Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math.* **36** (1997), 63–78.
- [7] Kobza, J.: *Splajny*. *Vydavatelství UP, Olomouc*, 1993 (textbook in Czech).
- [8] Kobza, J.: *Computing solutions of linear difference equations*. Dept. Math. Anal. and Appl. Math., Fac. Sci., Palacki Univ., Olomouc, Preprint series 21, 1999; Proceedings SANM XIII, Nectiny 1999, 157–172.
- [9] Kobza, J., Ženčák, P.: *Some algorithms for quartic smoothing splines*. *Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math.* **36** (1997), 79–94.