Computing Convex Spline Approximations

David L. Russell
Department of Mathematics, Virginia Tech
Blacksburg, VA 24060, USA

Abstract

After a short abstract discussion of convex approximation we specialize to a study of such approximation on compact intervals of R^1 via splines of even order $2\,m$ with given knots. We see that, from the point of view of computation, the convex spline approximation problem is straightforward, reducing to a quadratic program with linear inequality constraints, for orders 2 and 4, linear and cubic splines, respectively, but is more complicated for higher even orders $2\,m,\ m > 2$, if all convex splines of that order are to be admitted as candidates for the optimization. With restriction to a smaller set of splines of order $2\,m,\ p > 2$, obtained via a convolution process, treatment of the problem via quadratic programming with linear inequality constraints is restored to viability. Computational examples are presented for illustration.

1 A Convex Approximation Problem

We consider a convex spline approximation problem for *discrete data*, $\{z_k \,|k = 1, 2, 3, ..., N\}$ associated with distinct $\{x_k \,|k = 1, 2, 3, ..., N\}$, $\mathcal{R} = [a, b]$, a closed, bounded interval, with

$$a = x_0 < x_1 < x_2 < \ ... \ < x_{N-1} < x_N = b \equiv a + L. \qquad (1.1)$$

To each pair z_k, x_k we assign a "weight" μ_k, $k = 1, 2, 3, ..., N$. We pose the problem of finding $\hat{y} \in \mathcal{C}$, a set of convex spline functions of order $m + 1$ on $\mathcal{R}$ with given knots in that interval, such that $\hat{y}$ solves the problem $\min_{y \in \mathcal{C}} \phi(y)$ where, for $\epsilon > 0$,

$$\phi(y) = \sum_{k=1}^{\infty} \mu_k \left(z_k - y(x_k)\right)^2 + \epsilon \, \|y\|_m^2 ,$$

$\|y\|_m$ the standard Sobolev norm in $H^m(\mathcal{R})$. The proof of existence and uniqueness of $\hat{y}$ is straightforward. For simplicity, in this talk we take the spline knots to coincide with the $\{x_k\}$ and assume these to be equally spaced with

$$x_k - x_{k-1} = h, \ k = 1, 2, 3, \ ... \, N.$$

To this end the most usable and reliable techniques are those of *quadratic programming* with quadratic objective function and linear equality and/or inequality constraints. Consequently we emphasize characterization of convexity in terms of a finite number of such constraints. The most obvious convexity condition $y''(x) \geq 0$ clearly does not meet this objective if it has to be verified at every point $x \in [a, b]$ or for x in any other infinite set.

We begin with *linear* splines; degree 1, order 2. With $y_k \equiv y(x_k)$, $k = 0, 1, 2, \ ... \, N$, and the slope increment from each subinterval $[x_{k-1}, x_k]$ to the next, $[x_k, x_{k+1}]$, indicated by u_k, $k = 1, 2, 3, \ ... \, N - 1$, these splines correspond to solutions of second order linear recursion equations (corresponding to equality constraints)

$$\frac{y_{k+1} - y_k}{h} - \frac{y_k - y_{k-1}}{h} = u_k, \quad k = 1, 2, 3, \ ... \, N - 1.$$

The convexity condition in this case corresponds simply to the linear inequality constraints $u_k \geq 0, \;\; k = 2, 3, \dots N-1$. The "decision variables" for the optimization process are the initial y_0, y_1 and the "controls" $u_k, \;\; k = 1, 2, 3, \dots N - 1.6$

Next we consider *quadratic* splines, degree 2, order 3, we let

$$y_k^0 = y(x_k), \;\; y_k^1 = y_k^{'}(x_k), \;\;\; k = 0, 1, 2, \dots N.$$

With u_k the constant second derivative on $[x_{k-1}, \, x_k], \; k = 1, 2, 3, \dots N$, we have equality constraints corresponding to the two dimensional system of linear recursion equations

$$\begin{pmatrix} y_k^0 \\ y_k^1 \end{pmatrix} = \begin{pmatrix} 1 & h_k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{k-1}^0 \\ y_{k-1}^1 \end{pmatrix} + \begin{pmatrix} (h_k)^2/2 \\ h_k \end{pmatrix} u_k, \qquad k = 1, 2, 3, \dots N.$$

The convexity condition here is $u_k \geq 0, \;\; k = 1, 2, 3, \dots N$; the decision variables are y_0^0, y_0^1 and the indicated u_k. Only the latter are subject to inequality constraints; the first two are unconstrained.

Cubic splines, corresponding to degree 3, order 4, are the most commonly used in applications. Defining u_k to be the constant third derivative on $[x_{k-1}, \, x_k]$ we have, for $k = 1, 2, 3, \dots N$,

$$\begin{pmatrix} y_k^0 \\ y_k^1 \\ y_k^2 \end{pmatrix} = \begin{pmatrix} 1 & h_k & (h_k)^2/2 \\ 0 & 1 & h_k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{k-1}^0 \\ y_{k-1}^1 \\ y_{k-1}^2 \end{pmatrix} + \begin{pmatrix} (h_k)^3/6 \\ (h_k)^2/2 \\ h_k \end{pmatrix} u_k. \tag{1.2}$$

The basic convexity condition remains $y^{''}(x) \geq 0, \; x \in [a, \, b]$, involving **infinitely many** points x. However, in the cubic spline case $y^{''}(x)$ is **linear** on each $[x_{k-1}, \, x_k]$ so its non-negativity over the whole interval obtains if and only if both $y_{k-1}^2 = y^{''}(x_{k-1})$ and $y_k^2 = y^{''}(x_k)$ are both non-negative. Thus the convexity condition for cubic splines can be stated simply as

$$y_k^2 \geq 0, \;\; k = 0, 1, 2, \dots N. \tag{1.3}$$

The decision variables are y_0^0, y_0^1, y_0^2 and the $u_k, \;\; k = 1, 2, \dots N$.

In contrast to the linear and quadratic cases the condition (1.3) cannot be stated exclusively in terms of the u_k because y_0^2, required to be non-negative, is not determined by the u_k, $k = 1, 2, \ldots N$. These *state constraints* complicate some approaches to optimal convex cubic spline approximation.

For the last reason stated it is convenient to reformulate the construction of cubic splines. At stage $k-1$, y_{k-1}^2 being already determined, we take v_k to be the as yet undetermined value of y_k^2. Using the linearity of the second derivative on each subinterval and assuming $y_{k-1}^2 \geq 0$ we require $v_k \geq 0$ to ensure convexity on $[x_{k-1}, x_k]$. We then find that

$$y_k^0 = y_{k-1}^0 + h_k\, y_{k-1}^1 + \frac{(h_k)^2}{3}\, y_{k-1}^2 + \frac{(h_k)^2}{6}\, v_k,$$

yielding the discrete linear system

$$\begin{pmatrix} y_k^0 \\ y_k^1 \\ y_k^2 \end{pmatrix} = \begin{pmatrix} 1 & h_k/2 & (h_k)^2/3 \\ 0 & 1 & h_k/2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_{k-1}^0 \\ y_{k-1}^1 \\ y_{k-1}^2 \end{pmatrix} + \begin{pmatrix} (h_k)^2/6 \\ h_k/2 \\ 1 \end{pmatrix} v_k. \qquad (1.4)$$

The decision variables here are y_0^0, y_0^1, y_0^2 and the v_k, $k = 1, 2, ..., N$. The constraints are now *control* inequality constraints $v_k \geq 0$ except for the constraint $y_0^2 \geq 0$ on the initial conditions.

Fifth order (fourth degree) splines not being widely used, the natural next step is to examine convex sixth order (fifth degree) spline functions as approximation candidates. When we attempt to carry this out, however, we encounter a significant difficulty.

Proposition 1.1 *The set of all sixth order splines $s_6(x)$ convex on an interval $[x_j, x_k]$, $k > j$, cannot be precisely specified by imposing a finite number of linear inequality constraints on $s_6(x)$ or its derivatives, at any predefined, finite set of points in $[x_j, x_k]$.*

Remark It is possible, via linear inequality constraints, to describe certain *subsets* of convex sixth order splines.

Proof Assuming the contrary, it is equivalent that the set of all convex sixth order splines on $[x_j,\, x_k]$ should be precisely described by imposing a finite number of linear inequality constraints on the local polynomial coefficients of $s_6(x)$, six such real coefficients $c_{\ell,\, i},\ \ i\, =\, 0, 1, 2, 3, 4, 5$, for each of the $M\ \equiv\ k - j$ subintervals $[x_{\ell-1},\, x_\ell]$ of $[x_j,\, x_k]$. This corresponds to defining a polyhedron $\mathcal{P}$ in the $6\, M$ - dimensional space of all such coefficients. ($\mathcal{P}$ has dimension $<\ 6\, M$ because of spline continuity conditions at the knots.) Then the intersection of $\mathcal{P}$ with any other polyhedron, or affine hyperplane, $\mathcal{S} \subset R^{6M}$, would also be a polyhedron - possibly empty, of course. We define a subspace $\mathcal{S}$ by requiring $c_{\ell,\, 5}\ =\ 0$ for all such ℓ as described. Then $\mathcal{P} \cap \mathcal{S}$ corresponds to all convex splines $s_{\,6}(x)$ which reduce, on $[x_j,\, x_k]$, to a single fourth degree polynomial, independent of ℓ, with coefficients c_i multiplying $(x)^i,\ \ i\, =\, 0, 1, 2, 3, 4$; $\mathcal{P} \cap \mathcal{S}\ \equiv\ \mathcal{P}_{\mathcal{S}}$ is clearly not empty. The corresponding second derivatives are the quadratic polynomials

$$p(x)\ =\ 12\, c_4\, x^2\ +\ 6\, c_3\, x\ +\ 2\, c_2. \tag{1.5}$$

The polyhedron $\mathcal{P}_{\mathcal{S}}$ must therefore correspond precisely to all fourth degree polynomials for which $p(x)$ is non-negative on the interval $[x_j,\, x_k]$. Making the change of variable $x\ =\ x_j\ +\ (x\ -\ x_j)\ \equiv\ x_j\ +\ \xi$ and substituting into $p(x)$ we can re-express that polynomial as

$$\hat{p}(\xi)\ =\ 12\, c_4\, \xi^2\ +\ \hat{c}_3\, \xi\ +\ \hat{c}_2, \tag{1.6}$$

where $\hat{c}_2,\ \hat{c}_3$, are linear functions of $c_2,\ c_3,\ c_4$. Further intersecting $\mathcal{P}_{\mathcal{S}}$ with the polyhedral set corresponding to $\hat{c}_2,\ \hat{c}_3,\ c_4$ such that $\hat{p}(0)\ =\ \hat{c}_2\ =\ 1,\ \ c_4\ \geq\ 1$, we obtain a polyhedral set $\mathcal{Q}$. Now the set of all second degree polynomials $\hat{p}(\xi)$ with coefficients restricted as indicated and such that $\hat{p}(\xi)\ \geq\ 0,\ \ \xi\ \in\ [0,\, x_k - x_j]$ is the complement of the set of such polynomials whose roots are real and distinct, the lesser root lying in the interval $(0,\, x_k - x_j)$. The latter corresponds, of course, to the assumed inequality $c_4\ \geq\ 1$ and

$$\hat{c}_3^{\,2}\ >\ 48\, c_4, \qquad -\sqrt{\hat{c}_3^{\,2}\ -\ 48\, c_4}\ <\ 24\, c_4\, (x_k\ -\ x_j)\ +\ \hat{c}_3,$$

a set which is non-empty and clearly not a polyhedron. Then $\mathcal{Q}$ is not a polyhedron $\ \Rightarrow\ \ \mathcal{P}$ is not a polyhedron; the proposition follows.

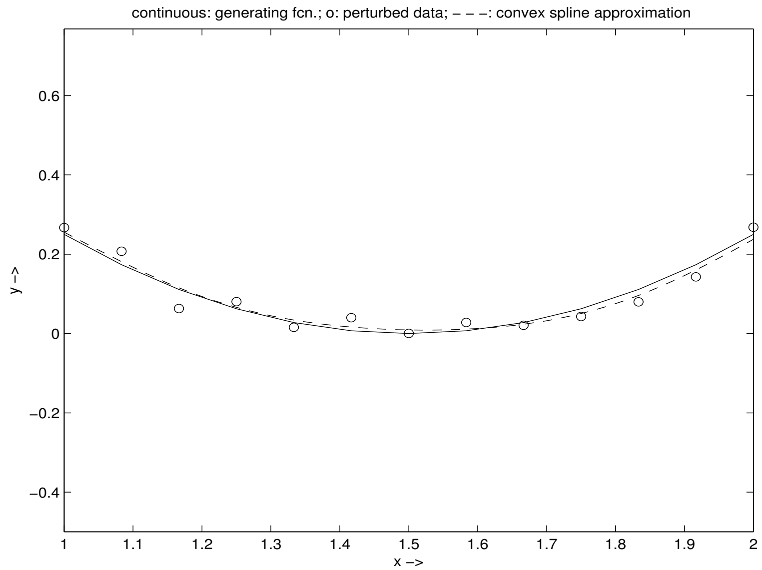

Figure 1: convex cubic approximation.

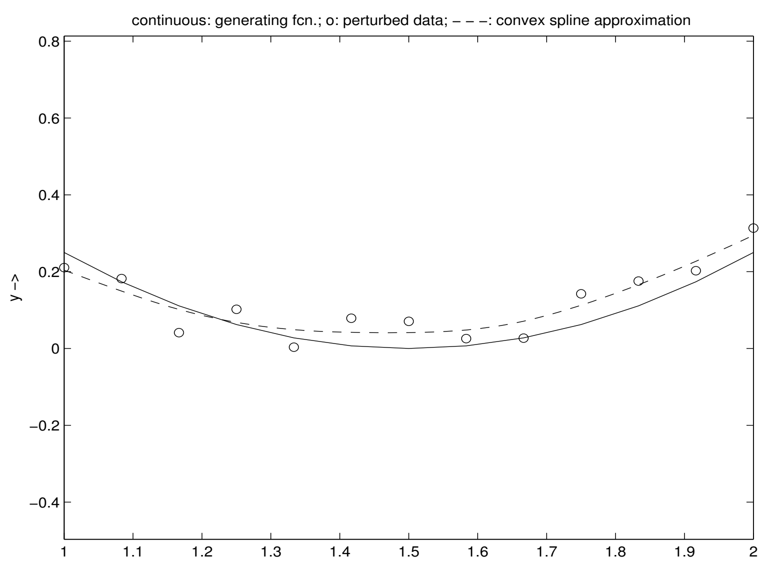

Figure 2: convex cubic approximation.

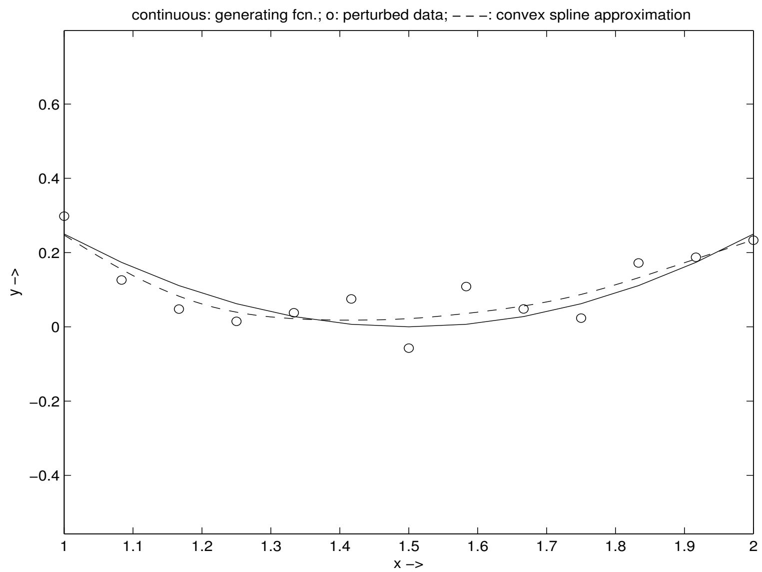

Figure 3: convex cubic approximation.

2 Convex Spline Approximation via Convolution

We have noted the difficulty in imposing linear constraints on higher order splines in order to precisely characterize convexity. Now we indicate how this can be at least partially overcome through construction of splines via *convolution*.

To simplify the ideas we assume a set of points $x_k,\ -\infty < k < \infty$, uniformly separated so that $x_k - x_{k-1} = h > 0$ for all k. We associate with each point x_k, a numerical value w_k. We define the "tent function" (second order spline basis function)

$$\tau(x) = \begin{cases} h^{-2}(x+h), & -h \le x < 0, \\ h^{-2}(h-x), & 0 \le x < h, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\int_{-\infty}^{\infty} \tau(x)\,dx = 1$. It is easy to see that

$$s_2(x) = \sum_k w_k\,\tau(x - x_k)$$

is the linear (second order) spline function with values w_k at the points x_k. The spline function $s_2(x)$ is convex on an interval $[x_{-\nu},\, x_{N+\nu}]$ if and only if

$$u_k \;\equiv\; h^{-2}\,(w_{k+1} \;-\; 2\,w_k \;+\; w_{k-1}) \;\geq\; 0, -\nu{+}1 \;\leq\; k \;\leq\; N{+}\nu{-}1. \quad (2.1)$$

Denoting the m-fold self-convolution of $\tau(x)$ by $\tau^{*m}(x)$), we may then define even order splines, for integral $p \geq 0$,

$$s_{2\,p+2}(x) \;=\; (\tau^{*p} \;*\; s_2)(x) \;=\; \sum_k w_k\, \tau^{*(p+1)}(x \;-\; x_k); \qquad (2.2)$$

that is, one may generate successive even order spline functions by repeated self-convolution of the tent function $\tau(x)$. It may then be shown that

$$s_{2p+2}(x_k) \;=\; \sum_j \beta_{2p+2,\,j}\, w_{k+j}, \qquad (2.3)$$

where the $\beta_{2p+2,\,j}$ are the nodal values of the standard B-spline $b_{2p+2}(x)$ of order $2p+2$ (center value corresponding to $\beta_{2p+2,0}$). One further sees that

$$s_{2p+2}(x) \;=\; \int_{-\infty}^{\infty} b_{2p}(x-\xi)\, w(\xi)\, d\xi. \qquad (2.4)$$

If one wants $s_{2p+2}(x)$ to assume given values $\hat{w}_k$ at the points x_k one solves

$$\hat{w}_k \;=\; \sum_j \beta_{2p+2,\,j}\, w_{k+j}$$

for the w_k; a unique solution exists in standard contexts. Since the spline basis function $b_{2p}(x)$ has only non-negative values, $s_{2p+2}(x)$ is convex if the piecewise linear spline $w(x) \;=\; s_2(x)$ is convex, for which we have the criterion: $s_2(x)$ is convex on $[x_{-\nu},\, x_{N+\nu}]$ if and only if the second order differences

$$w_{k+1} \;-\; 2 * w_k \;+\; w_{k-1}, \;\; k \;=\; -\nu + 1, ..., N + \nu - 1$$

are all non-negative. In the cubic spline case $2\,p + 2 \;=\; 4$ this condition is also necessary but for larger values of p one can show that $s_{2p+2}(x)$ may be convex even if this condition is not satisfied for all k.

To summarize: it is not possible to describe *all* $2p+2 \geq 6$ order splines convex on an interval $\mathcal{R}$ by means of linear inequalities and/or equations and therefore the most general problem of optimal convex $2p+2 \geq 6$ order spline approximation of given data on that interval relative to some quadratic, or quadratic plus linear, cost criterion cannot be reduced to a quadratic programming problem with linear inequality constraints. It is possible to do so if we restrict the set of candidate splines $s_{2p+2}(x)$ constructed via the convolution process described above. By contrast the set of all convex fourth order (cubic) splines is the same as the set of splines $s_4(x)$ constructed via the convolution process.

Some Computational Examples; Order 6

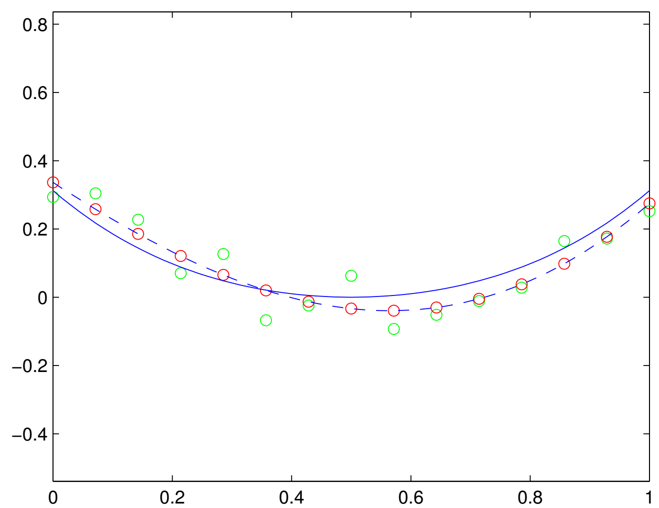

Figure 4

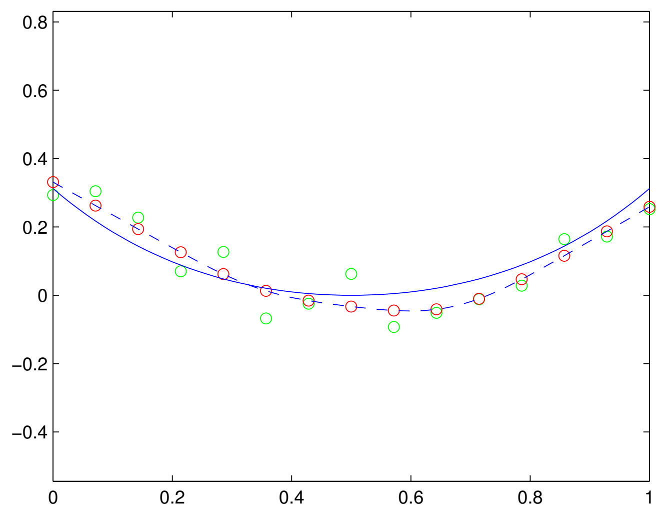

Figure 5

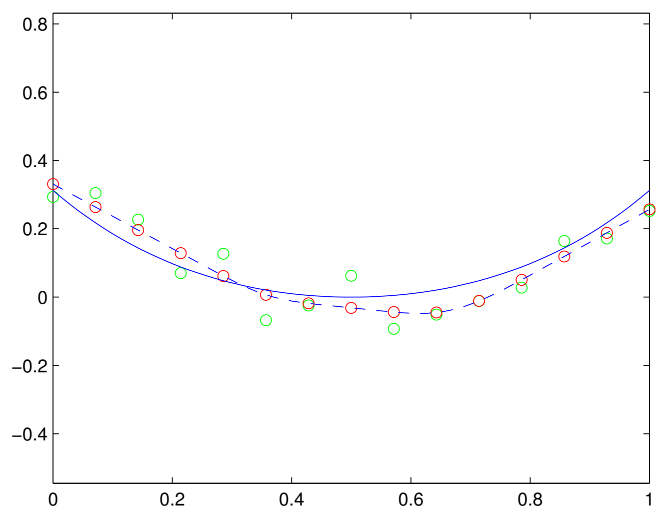

Figure 6

Acknowledgment: The author wishes to acknowledge conversations with Prof. Kenneth L. Judd of the Hoover Institution, Stanford University, which initiated this research effort.

This article appeared in *International Journal of Information, Systems & Sciences*, Vol. 5 (2009), pp. 83 ff.. Related references follow.

References

[1] Bojanic, R., and J. Roulier: *Approximation of convex functions by convex splines and convexity preserving continuous linear operators*, Rev. Anal. Numér. Théorie Approx. 3 (1974), pp. 143 - 150.

[2] Burchard, H.: *Extremal positive splines with applications to interpolation and approximation by generalized convex functions*, Bull. Amer. Math. Soc., 79 (1973), pp. 959 - 963.

[3] De Boor, C.: *A Practical Guide to Splines*, Springer-Verlag, New York, 1978.

[4] Gifford, D., and R. Huotari: *Best L^1 approximation by convex functions*, Num. Alg. 9 (1995), pp. 107 - 111.

[5] Huotari, R., D. Legg and D. Townsend: *Existence of best N-convex approximants in L^1*, Approx. Thy. & Appl., 5 (1989), pp. 51 - 57.

[6] Messac, A., and A. Sivanandan: *A new family of convex splines for data interpolation*, Comput. Aided Geom. Design 15 (1997), pp. 39 - 59.

[7] Renardy, M., and D. Russell *A discrete systems approach to cardinal spline Hermite interpolation*, Lin. Alg. Appl., 406 (2005), pp. 77-98.

[8] Shevchuk, I.: *One construction of cubic convex spline*, in *Approximation and Optimization*, Vol. I, Cluj - Napoca, 1997, pp. 357 - 368.

[9] Wahba, G.: *Spline Models for Observational Data*, CBMS - NSF regional conference series in applied mathematics 59, SIAM Pub., Philadelphia, 1990.