On Characterization of Quadratic Splines

Bintong Chen [*] Kaj Madsen [†] Shuzhong Zhang [‡]

Submitted Oct. 5, 2002
(revised Sept. 18 and Oct. 23, 2003)

Abstract

A quadratic spline is a differentiable piecewise quadratic function. Many problems in numerical analysis and optimization literature can be reformulated as unconstrained minimizations of quadratic splines. However, only special cases of quadratic splines are studied in the existing literature, and algorithms are developed on a case by case basis. There lacks an analytical representation of a general or even a convex quadratic spline. The current paper fills this gap by providing an analytical representation of a general quadratic spline. Furthermore, for convex quadratic spline, it is shown that the representation can be refined in a neighborhood of a non-degenerate point and a set of non-degenerate minimizers. Based on these characterizations, many existing algorithms for specific convex quadratic splines are also finitely convergent for a general convex quadratic spline. Finally, we study the relationship between the convexity of a quadratic spline function and the monotonicity of the corresponding LCP problem. It is shown that, although both conditions lead to easy solvability of the problem, they are different in general.

Keywords. Quadratic Spline, Convexity, LCP, Monotonicity.

[*]Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, U.S.A. (chenbi@wsu.edu).

[†]Department of Mathematical Modelling, Technical University of Denmark, 2800, Lyngby, Denmark. (km@imm.dtu.dk).

[‡]Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Shatin, Hong Kong. (zhang@se.cuhk.edu.hk).

1 Introduction

A function $F : R^n \to R$ is called a *quadratic spline* if it is continuously differentiable and piecewise quadratic, i.e., there exist finitely many quadratic functions

$$F_j(x) = \frac{1}{2}x^T P_j x + p_j^T x + c_j, \quad j = 1, \ldots, m \tag{1}$$

for some integer $m > 0$ such that

$$F(x) \in \{F_1(x), F_2(x), \ldots, F_m(x)\} \text{ for } x \in \Re^n, \tag{2}$$

where $P_1, \ldots, P_m$ are $n \times n$ symmetric matrices, $p_1, \ldots p_m \in \Re^n$, and $c_1, \ldots, c_m \in \Re$. Many problems in numerical analysis and optimization literature can be reformulated as unconstrained minimizations of quadratic splines.

In a paper by Li and Swetits [7], a special case of quadratic spline with the following structure is considered:

$$F(x) := V(x) + \frac{1}{2}\|(Ax + b)_+\|^2, \tag{3}$$

where A is an $m \times n$ matrix, $b \in \Re^m$, $V(x)$ is a convex quadratic function, and $F(x)$ is assumed to be bounded below. We mention here that throughout the paper the Euclidean norm will be used and z_+ will be used to denote the nonnegative part of $z \in \Re^n$, component-wise. It has been shown that the quadratic spline of structure (3) is general enough to include the reformulation of the following problems as special cases:

- least two norm solution of a system of linear inequalities [2];
- least distance problem [11];
- convex quadratic programming with simple bound constraints [5, 6];
- strictly convex quadratic programming [5, 6].

On the other hand, there are some important quadratic spline reformulations that do not fit into the structure (3). Consider the reformulation of Huber's M-estimator, which is considered as the most popular estimator for robust linear regression. It minimizes a linear residual vector $r(x)$ by using the Huber function:

$$\rho(t) = \begin{cases} \frac{1}{2}t^2 & \text{if } |t| \le \gamma \\ \gamma|t| - \frac{1}{2}\gamma^2 & \text{otherwise,} \end{cases}$$

where $\gamma > 0$ is a tuning constant. After some algebraic manipulation, the Huber's M-estimator problem can be reformulated as a minimization of the following convex quadratic spline:

$$F(x) = -\frac{1}{2}\|r(x)\|^2 - \gamma^2 + \frac{1}{2}\|(r(x) + \gamma e)_+\|^2 + \frac{1}{2}\|(-r(x) + \gamma e)_+\|^2, \tag{4}$$

where $e \in \Re^n$ is the vector with all of its elements equal to 1. Clearly, the F in (4) does not fit into the structure of (3), even though it is a convex quadratic spline. See [1, 8] for finitely convergent algorithms of minimizing the function F in (4).

Minimization of a quadratic spline is also related to a linear complementarity problem, i.e., to find an x such that

$$x \geq 0, \quad G(x) \geq 0, \quad x^T G(x) = 0, \tag{5}$$

where $G(x) = Mx + q$ is a linear function for some matrix $M \in \Re^n \times \Re^n$ and vector $q \in \Re^n$. Mangasarian and Solodov [10] proved that x is a solution of the LCP (5) if and only if it is a global minimizer of the following quadratic spline:

$$F(x) = x^T G(x) + \frac{1}{2\alpha}[\|(x - \alpha G(x))_+\|^2 - \|x\|^2 + \|(G(x) - \alpha x)_+\|^2 - \|G(x)\|^2]$$

for all $\alpha > 0$. If M is a symmetric matrix and $0 < \alpha\|M\| < 1$, then Li [4] showed further that x is a solution of the above LCP if and only if it is a stationary point of the following quadratic spline:

$$F(x) = x^T G(x) - \|G(x)\|^2 + \|(G(x) - \frac{1}{\alpha}x)_+\|^2$$

Clearly, both quadratic splines in the above reformulations do not fit into the structure of (3).

It is then natural to ask if we can represent a general (convex) quadratic spline analytically. This paper attempts to provide a positive answer to the above questions. The paper is organized as follows: In Section 2, we provide an analytical representation of a general quadratic spline function. We show that any quadratic spline can be cast into a proposed structure. In Section 3, we show that for convex quadratic spline, the representation can be refined in a neighborhood of a nondegenerate point or a set of nondegenerate minimizers. Based on this refinement, some existing finitely convergent Newton type algorithms for specific convex quadratic splines (e.g., [7]), are also finitely convergent for general convex quadratic splines. In Section 4, we discuss two useful conditions under which the problem can be solved efficiently. We conclude the paper in Section 5.

2 Representation of a General Quadratic Spline

The definition of quadratic spline, (1) and (2), given in Section 1, implies a subdivision of $\Re^n$ into m closed regions S_j, $j = 1, \dots, m$, such that

$$\cup_{j=1}^m S_j = \Re^n \text{ and } F(x) = F_j(x) \text{ for } x \in S_j, j = 1, \dots, m.$$

If $S_i \cap S_j \neq \emptyset$ for some $i, j \in \{1, \dots, m\}$, S_i and S_j are called neighborhood regions. Without loss of generality, we assume that every region S_j is of dimension n and none of the intersections of neighborhood regions $S_i \cap S_j$, $i \neq j$, is of dimension n. Here, the dimension of S_i should be interpreted as the dimension of the affine hull of set S_i. See [13] for justifications of these assumptions.

The following properties about a quadratic spline was given by Sun (Lemma 1 in [13]).

Lemma 1 *Let F be a quadratic spline defined by (1) and (2). Suppose S_i and S_j are two neighborhood regions such that the intersection $S_i \cap S_j$ is of dimension $n-1$. Then there exists a vector $a \in \Re^n$ such that $a^T(x - y) = 0$ for any $x, y \in S_i \cap S_j$ and $P_i - P_j = \omega a a^T$, where ω is either 1 or -1.*

In addition, it was further demonstrated in [12, 13] that each region S_j associated with the quadratic spline F is a polyhedron.

We extend Lemma 1 by following a similar proof techniques established in [13]. The extension is instrumental to an analytical representation of a general quadratic spline.

Lemma 2 *Let F be a quadratic spline. Suppose S_i and S_j are two neighborhood regions such that $S_i \cap S_j$ is of dimension $n-1$. Then there exist a vector $a \in \Re^n$ and a scalar $b \in \Re$ such that*

1. $a^T x + b = 0$ *for all* $x \in S_i \cap S_j$ *and* $F_i(x) = F_j(x) + \omega(a^T x + b)^2$.
2. $F(x) = F_j(x) + \omega[(a^T x + b)_+]^2$ *for all* $x \in S_i \cup S_j$,

where ω is either 1 or -1.

Proof: Since $S_i \cap S_j$ is of dimension $n-1$, there exist n point $x^1, \ldots, x^n \in S_i \cap S_j$ such that $x^2 - x^1, \ldots, x^n - x^1$ are linearly independent. By Lemma 1, there exists a vector $a \in \Re^n$ such that

$$P_i - P_j = \omega a a^T. \tag{6}$$

and

$$a^T(x^k - x^1) = 0 \quad \text{for } k = 2, \ldots, n. \tag{7}$$

Since F is continuous at $x^1, \ldots, x^n$, we have, using (6),

$$\frac{1}{2}(a^T x^k)^2 + p_i^T x^k + c_i = p_j^T x^k + c_j \quad \text{for } k = 1, \ldots, n.$$

Subtracting the above identities with $k = 2, \ldots, n$ from that with $k = 1$ and using property (7), we obtain

$$(p_i - p_j)^T (x^k - x^1) = 0 \quad \text{for } k = 2, \ldots, n.$$

Using property (7) again, we get

$$p_i - p_j = \omega b a \tag{8}$$

for some scalar $b \in \Re$. We next show that

$$c_i - c_j = \frac{1}{2}\omega b^2. \tag{9}$$

Let $x \in S_i \cap S_j$. Since F is C^1 at x,

$$F_i'(x) = F_j'(x), \quad F_i(x) = F_j(x).$$

Applying relations (6) and (8) to the above equalities, we have

$$\omega(a^T x + b)a = 0 \tag{10}$$

$$\frac{1}{2}\alpha(a^T x)^2 + \omega b a^T x + c_i - c_j = 0. \tag{11}$$

If $a = 0$ equalities (8) and (9) hold with $b = 0$. Otherwise, equality (10) implies

$$a^T x + b = 0 \tag{12}$$

Substituting (12) to equality (11), we obtain equality (9). That

$$F_i(x) = F_j(x) + \omega(a^T x + b)^2 \tag{13}$$

is then followed by combining equalities (6), (8), and (9).

By choosing the signs of a and b properly, we may set

$$a^T x + b \geq 0 \quad \forall x \in S_i \quad \text{and} \quad a^T x + b \leq 0 \quad \forall x \in S_j.$$

Result 2 then follows from (13) immediately. □

Based on Lemma 2, we are able to provide a representation of any general quadratic spline function F.

Theorem 1 *Any quadratic spline $F : \Re^n \to \Re$ can be written as:*

$$F(x) = Q(x) + \frac{1}{2}\|(Ax+b)_+\|^2, \tag{14}$$

where $Q : \Re^n \to \Re$ is a quadratic function, A is an $m \times n$ matrix, $b \in \Re^m$, and $m > 0$ is an integer.

Proof: Let $S_1 \cup S_2 \cup \ldots \cup S_t = \Re^n$ be the subdivision associated with the quadratic spline F. We prove the result in a recursive way. We start by choosing any region, say S_1, and setting $S = S_1$ and $G(x) = F_1(x)$. Now consider a recursive step. If $S = \Re^n$, $F(x) = G(x)$ is a representation of the quadratic spline and we proved the result. Otherwise, there exists a region that shares the border with S and their intersection is of dimension $n-1$, since S is a union of polyhedral regions. Suppose the intersection is defined by a hyperplane, say $a^T x + b = 0$. Let $\bar{S}$ be a union of all neighborhood regions of S such that their intersections are given by the above hyperplane. Let $S = S \cup \bar{S}$. By result 2 of Lemma 2, for some properly scaled a and b, we have

$$F(x) = G(x) + \frac{1}{2}\omega[(a^T x + b)_+]^2 \quad \forall x \in S,$$

where $\omega = 1$ or -1. If $\omega = -1$, using the fact $t^2 = (t_+)^2 + [(-t)_+]^2$, we may rewrite the above expression as follows:

$$F(x) = G(x) - \frac{1}{2}(a^T x + b)^2 + \frac{1}{2}[(-a^T x - b)_+]^2 \quad \forall x \in S.$$

We then set $G(x) := G(x) - \frac{1}{2}(a^T x + b)^2$. The above process is repeated until $S = \Re^n$. Notice that m obtained by the above process is finite since the number of subdivisions S_j is finite. It is also clear that $F(x)$ obtained from the above process has the structure (14) with $Q(x) = G(x)$. □

3 Representation of Convex Quadratic Splines

Since many quadratic splines resulting from reformulations are convex, it would be interesting to see if the representation in (14) can be refined for a convex quadratic spline. A reasonable question would be: whether or not F being convex implies Q being convex in (14)? In other words, can all convex quadratic spline be represented by (3) instead? Unfortunately, the answer to this question is negative. This is demonstrated by the convex quadratic spline given in (4), where the corresponding function $Q(x)$ is not convex. Naturally, one may be led to ask whether a local representation in the form of (3) would be possible, i.e. what if we consider only those x which are in a neighborhood of minimizers of F. Although this seems to be a plausible conjecture, it is not true in general either. Before we discuss a counter-example we shall first note a lemma.

Lemma 3 *A quadratic spline function is globally convex if and only if it is convex in each open region where it is quadratic.*

Proof: The 'only if' part of the statement is merely a matter of definition. What needs to be shown is that if a quadratic spline is convex in each open region, then it must be globally convex. To this end, obviously we need only to show that such a quadratic spline function is convex in every possible directions. This means that we need only to show the desired lemma for an arbitrary one-dimensional quadratic spline.

In this case, we may partition the entire domain $\Re^1$ by a finite number of intervals $I_i = (l_i, r_i]$ with $l_1 = -\infty$, $r_i = l_{i+1}$, $i = 1, ..., m-1$, and $I_m = (l_m, \infty)$.

A smooth function F is convex in an interval I if and only if F' is a nondecreasing function in I. In particular, in our case F' is a nondecreasing function within each I_i, $i = 1, ..., m$. Moreover, $F'(r_i) = F'(l_{i+1})$ for $i = 1, ..., m-1$, due to the continuous differentiability. Hence, F' is a nondecreasing function in $\Re^1$, and thus F is a globally convex function. □

Let us consider now a quadratic spline function given as follows

$$F(x_1, x_2) = -\epsilon(x_1 + x_2)^2 + \frac{3\epsilon}{2}(x_1^2 + x_2^2) + (x_1)_+^2 + (x_2)_+^2 + (-x_1 - x_2)_+^2 \tag{15}$$

where $0 < \epsilon < 0.5$. The regions in which this function remains quadratic is shown in Figure 1:

It is easy to verify that the Hessian matrix of the function is positive semidefinite in each of these six open regions: I, II, III, IV, V and VI. For example, in region I,

$$F''(x) = 2 \begin{bmatrix} 1 + \frac{3\epsilon}{2} - \epsilon, & -\epsilon \\ -\epsilon, & 1 + \frac{3\epsilon}{2} - \epsilon \end{bmatrix} \succeq 0.$$

Similar relations can be verified for each of these six regions, and so the function F is convex in each of these regions. By Lemma 3, F is a convex quadratic spline. Moreover, the representation of (15) cannot be reduced any further. However, the quadratic part of the representation,

$$Q(x_1, x_2) = -\epsilon(x_1 + x_2)^2 + \frac{3\epsilon}{2}(x_1^2 + x_2^2) = \epsilon(0.5x_1^2 - 2x_1x_2 + 0.5x_2^2)$$

is *not* convex.

If we examine this example more closely, then we may observe that the trouble is caused by three lines crossing at the origin simultaneously. Following the terminology from linear programming, we call this situation *degeneracy*. We now formally introduce this concept below. Based on Theorem 1, it is sufficient for us to consider the quadratic spline in the form of (14). For simplicity, denote $r(x) := Ax + b$, where $A = [a_1, a_2, ..., a_m]^T$ is an $m \times n$ matrix and $b \in \Re^m$. We may rewrite (14) as

$$F(x) := Q(x) + \frac{1}{2}\|r(x)_+\|^2, \tag{16}$$

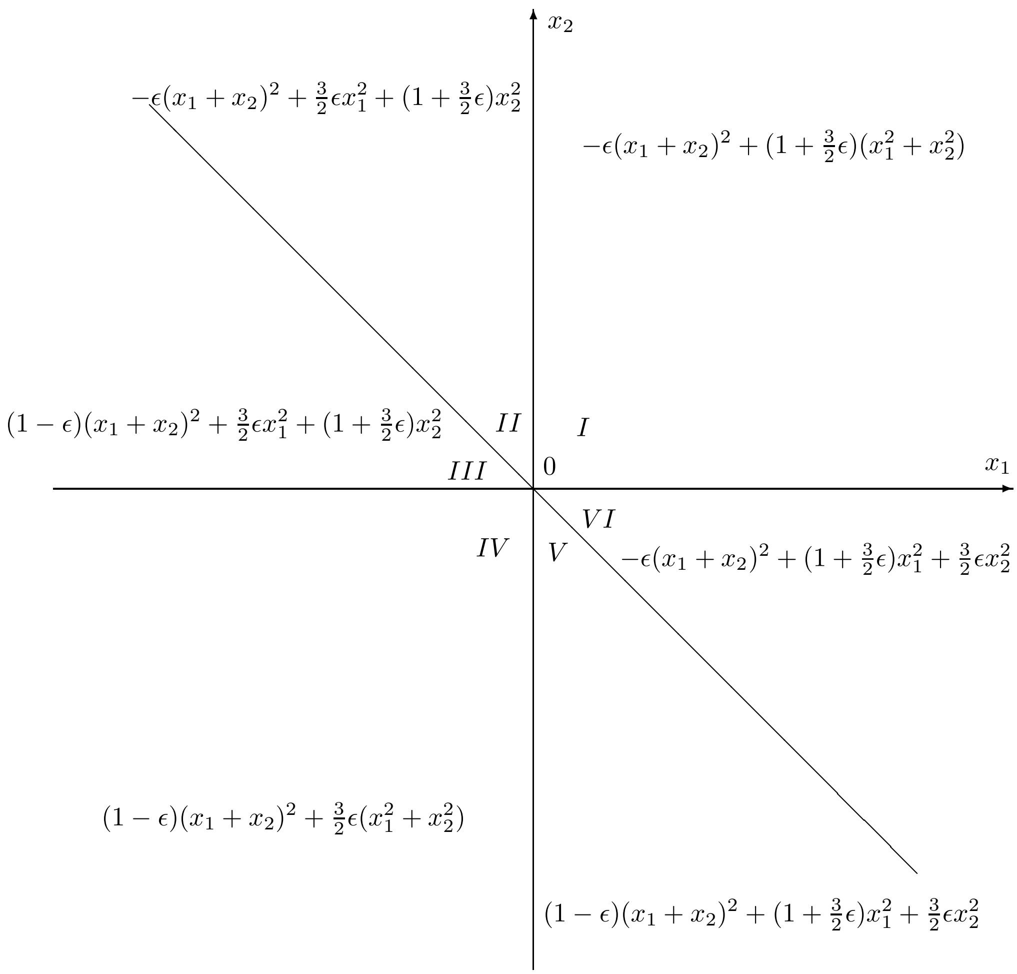

Figure 1: Quadratic Regions of F

where $Q(x)$ is a quadratic function. For each given $x \in \Re^n$, we define the following index sets:

$$\begin{aligned} I_+(x) &:= \{j \mid r_j(x) \geq 0\}; \\ I_{++}(x) &:= \{j \mid r_j(x) > 0\}; \\ I_0(x) &:= \{j \mid r_j(x) = 0\}; \\ I_-(x) &:= \{j \mid r_j(x) \leq 0\}; \\ I_{--}(x) &:= \{j \mid r_j(x) < 0\}. \end{aligned}$$

Definition 1 *The quadratic spline $F(x)$ is called nondegenerate at x if* rank $\{a_i \mid i \in I_0(x)\} = |I_0(x)|$.

Intuitively, if x is nondegenerate point of a quadratic spline F, then it should not be an intersection point of 'too many' regions S_j. The main result of this section is to show that in the absence of degeneracy, any convex quadratic spline can always be locally expressed in the form of (3). In essence, this implies that the counter-example stated above can exist only in the presence of degeneracy.

Theorem 2 *Suppose that a convex quadratic spline function $F(x)$ is not degenerate at x. Then in a neighborhood of x, there exist a convex quadratic function $\tilde{Q}$, a matrix $\tilde{A}$, and a vector $\tilde{b}$, such that*

$$F(x) = \tilde{Q}(x) + \frac{1}{2}\|(\tilde{A}x + \tilde{b})_+\|^2.$$

Proof: From Theorem 1, the quadratic spline $F(x)$ can be written as (16):

$$Q(x) + \frac{1}{2}\|r(x)_+\|^2.$$

Let

$$Q(x) = \frac{1}{2}x^T Q x + q^T x.$$

Based on our early discussion, Q does not have to be positive semidefinite even if F is a convex quadratic spline. Consider the following system of linear equations:

$$a_i^T d = -1, \text{ for } i \in I_0(x).$$

Since F is nondegenerate at x, the above system admits at least one solution. Let d be such a solution and define $x(t) = x + td$ with $t \geq 0$. Clearly for sufficiently small $t > 0$ we have

$$\begin{aligned} I_{++}(x(t)) &= I_{++}(x) \\ I_{--}(x(t)) &= I_{--}(x) \cup I_0(x) \\ I_0(x(t)) &= \emptyset. \end{aligned}$$

Therefore,

$$F(x(t)) = \tilde{Q}(x(t)) =: \frac{1}{2}x(t)^T \tilde{Q} x(t) + q^T x(t),$$

where

$$\tilde{Q} := Q + \sum_{i \in I_{++}(x(t))} a_i a_i^T.$$

By Lemma 3, $x(t)$ resides in an open domain in which F is convex quadratic. It follows that $\tilde{Q} \succeq 0$. Let y be in a neighborhood of x. Then,

$$I_{++}(y) = I_{++}(x) = I_{++}(x(t))$$

where $t > 0$ is sufficiently small. Hence,

$$F(y) = \tilde{Q}(y) + \frac{1}{2} \sum_{i \in I_{--}(x) \cup I_0(x)} (r_i(y))_+^2$$

where $\tilde{Q}(y)$ is a convex quadratic function as required. □

Next we shall establish that the index set $I_{++}(x)$ is in fact invariant for any nondegenerate minimizer x of a convex quadratic spline. A similar result was established in [9] for a specific convex quadratic spline.

Lemma 4 *Let X be the set of all minimizers of a convex quadratic spline function F. Suppose that F is not degenerate at any $x \in X$. Then it must hold that*

$$I_{++}(x) \equiv I$$

for all $x \in X$ where I is a certain fixed index set.

Proof: For a fixed $x \in X$, we need to show that $I_{++}(y) = I_{++}(x)$ for any other $y \in X$.

Suppose that such is not the case. Then, without loss of generality we may assume that there is $y \in X$ with the property that $j \in I_{++}(x)$ but $j \notin I_{++}(y)$. In particular we may further assume $j \in I_0(y)$.

Let

$$F_j(z) = Q(z) + \frac{1}{2} \sum_{i \neq j} (r_i(z))_+^2.$$

Due to the nondegeneracy assumption, $F_j(z)$ is locally convex around y. Moreover,

$$\nabla F_j(y) = \nabla F(y) = 0$$

implying that y is a local minimizer of F_j.

For sufficiently small $t > 0$ we have

$$\begin{aligned}
F(y) &= F(y + t(x-y)) = F_j(y + t(x-y)) + (a_j^T y + b_j + t a_j^T (x-y))_+^2 \\
&= F_j(y + t(x-y)) + t^2 (a_j^T (x-y))_+^2 \\
&= F_j(y + t(x-y)) + t^2 (a_j^T (x-y))^2 \\
&> F_j(y + t(x-y)) \\
&\geq F_j(y) = F(y)
\end{aligned}$$

which is a contradiction. □

Based on the above lemma, both $I_{++}(x)$ and $I_{-}(x)$ are unique for all $x \in X$. For simplicity, we denote them by $I_{++}(X)$ and $I_{-}(X)$, respectively. It follows that the set of local minimizers X belongs to only a single region. More precisely, we have

$$X \subseteq \{x \mid r_i(x) \leq 0,\ i \in I_{-}(X)\}$$

and

$$F(x) = Q(x) + \frac{1}{2} \sum_{j \in I_{++}(X)} r_j^2(x) \quad \forall x \in X.$$

Denote

$$\tilde{Q}(x) := Q(x) + \frac{1}{2} \sum_{j \in I_{++}(X)} r_j^2(x)$$

and let $\tilde{r}(x)$ be a vector whose components are given by $r_i(x) : i \in I_-(X)$. Then $\tilde{Q}(x)$ is a convex function since $F(x)$ is convex in a neighborhood of X. Summarizing the above discussion, we have the following key result of the section:

Theorem 3 *Let F be a convex quadratic spline function. Let X be the set of minimizers of $F(x)$. Assume that F is nondegenerate at any $x \in X$. Then*

$$F(x) = \tilde{Q}(x) + \frac{1}{2} \|(\tilde{r}(x))_+\|^2$$

for all x in a neighborhood of X, where $\tilde{Q}(x)$ is a convex quadratic function.

Notice that the set $I_0(x)$ may not be invariant for all $x \in X$. This can be seen from an example with only one variable $F(x) = (-x-1)_+^2 + (x-1)_+^2$. The optimal set is $X = [-1, +1]$. Clearly $I_0(-1) \neq I_0(+1) \neq I_0(0)$.

If Q is a convex function, then we can characterize X further based on Lemma 4.

Corollary 1 *Consider a quadratic spline in the form of (16). If $Q(x)$ is a convex function and F is nondegenerate at the set of minimizers, then $r_i(x) = \text{constant} > 0$ for all $x \in X$ and $i \in I_{++}(X)$.*

Proof: Recall that $F(x) = \tilde{Q}(x)$ for all $x \in X$. Let $x, z \in X$ be arbitrary. Then

$$\tilde{Q}''(x)(x - z) = (Q''(x) + \sum_{i \in I_{++}(X)} a_i^T a_i)(x - z) = 0.$$

The result then follows from the assumption that $Q''(x)$ is positive semi-definite. □

Notice that a similar characterization has been obtained for the solution set of the Huber's M-estimator problem [1, 8]. In addition, the assumption that requires Q to be convex also seems to be necessary. This can be seen from the following example of one variable convex quadratic spline:

$$F(x) = -x^2 + 2x + \|(1 - x)_+\|^2 + \|x_+\|^2.$$

In this example, $Q(x) = -x^2 + 2x$ is a concave function, $r_1(x) = 1 - x$, and $r_2(x) = x$. By observation, $X = \{x \mid x \leq 0\}$, $I_{++}(X) = 1$, and $I_-(X) = 2$. It is clear that $r_1(x) = 1 - x$ is not a constant in X.

Before we conclude this section, we would like to pose the following question of which we are not aware of an easy answer.

For a given quadratic spline

$$F(x) = Q(x) + \frac{1}{2}\|(Ax+b)_+\|^2,$$

determine if F is a globally convex function.

The data of the problem include the Hessian matrix of $Q(x)$ and the coefficients of the linear term in $Q(x)$, as well as the matrix A and the vector b. If $F(x)$ is indeed *not* convex, then a certificate for the nonconvexity needs to involve only two points, say y and z, with the property that $F(\frac{y+z}{2}) > \frac{F(y)+F(z)}{2}$. Therefore, the problem is clearly in the co-NP class. However, in case $F(x)$ is in fact a convex function, then it is unclear if there also is a polynomially sized certificate (in terms of the input length of the original data) to confirm this fact. Namely, we wonder whether or not the problem is in the class NP. If it turns out that the problem is indeed in NP and co-NP at the same time, then the next natural question would be: is there a polynomial time algorithm to find such certificates to verify the global convex/nonconvex property of $F(x)$?

4 Convexity versus Monotonicity

It is well known that minimizing a convex quadratic function, subject to polyhedral constraints, is equivalent to solving a linear complementarity problem satisfying the so-called monotonicity property. In this section, we will investigate if a similar equivalence holds for minimization of a convex quadratic spline. In particular, we compare the conditions to ensure the convexity of a quadratic spline with the conditions to ensure the monotonicity of the corresponding linear complementarity problem.

Recall that we are interested in finding the minimum of the following quadratic spline function

$$F(x) = \frac{1}{2}x^TQx + q^Tx + \frac{1}{2}\sum_{i=1}^{m}(a_i^Tx+b_i)_+^2.$$

The first order optimality condition for minimizing F is

$$\nabla F(x) = Qx + q + \sum_{i=1}^{m}(a_i^Tx+b_i)_+a_i = 0$$

which is equivalent to the following (horizontal) linear complementarity problem

$$(LCP)\left[\begin{array}{lcl} A^Tz+Qx+q & = & 0 \\ z-t & = & Ax+b \\ z^Tt & = & 0,\ z\geq 0, t\geq 0. \end{array}\right.$$

Such a linear complementarity problem is called *monotone* if for any displacements Δz, Δx and Δt satisfying

$$\left[\begin{array}{lcl} A^T\Delta z+Q\Delta x & = & 0 \\ \Delta z-\Delta t & = & A\Delta x \end{array}\right.$$

it must follow that $\langle \Delta z, \Delta t\rangle \geq 0$.

Eliminating Δt from the above statement, we may rewrite this condition as:

$$A^T \Delta z + Q\Delta x = 0 \Longrightarrow \langle \Delta z, \Delta z \rangle + \langle Q\Delta x, \Delta x \rangle \geq 0.$$

A monotone LCP can be solved in polynomial-time using, e.g., interior point methods; see [3]. On the other side, if F is convex then efficient minimization methods exist as well. A natural question is: Are these two conditions state the same fact?

Of course, in the case that $Q \succeq 0$, then obviously F is convex and (LCP) is monotone. However, when Q is indefinite, it is not immediately clear whether or not these two conditions are still equivalent.

We start by investigating the implications of the monotonicity of (LCP). Under the condition that A has full-column rank, it is indeed possible to characterize this property. For the monotonicity to hold, it is both necessary and sufficient that for any given Δx the solution Δz^* of the following problem

$$\begin{array}{ll} \text{minimize} & \|\Delta z\|^2 \\ \text{subject to} & A^T \Delta z = -Q\Delta x \end{array}$$

satisfies $\langle \Delta z^*, \Delta z^* \rangle + \langle Q\Delta x, \Delta x \rangle \geq 0$. If A has full-column rank, then

$$\Delta z^* = -A(A^T A)^{-1} Q\Delta x.$$

This means that the monotonicity of (LCP) is equivalent to

$$Q + Q(A^T A)^{-1} Q \succeq 0.$$

Now we shall see the implications of F being convex. First of all, remark that F is convex then $\langle \nabla F(x) - \nabla F(x'), x' - x \rangle \geq 0$ for any x and x'. Let $\Delta x = x' - x$. The above stated condition is equivalent to

$$\langle Q\Delta x + A^T [(Ax' + b)_+ - (Ax + b)_+], \Delta x \rangle \geq 0$$

for all Δx. Or, equivalently,

$$\langle (Ax' + b)_+ - (Ax + b)_+, A\Delta x \rangle + \langle Q\Delta x, \Delta x \rangle \geq 0$$

for all Δx. The following lemma is readily seen by elementary arguments.

Lemma 5 *For any vectors $a, b \in \Re^n$, it holds that*

$$\langle a_+ - b_+, a_+ - b_+ \rangle \leq \langle a_+ - b_+, a - b \rangle \leq \langle a - b, a - b \rangle.$$

As a consequence of Lemma 5, if F is convex, then

$$\langle A\Delta x, A\Delta x \rangle + \langle Q\Delta x, \Delta x \rangle \geq 0$$

for all Δx. That is,

$$Q + A^T A \succeq 0. \tag{17}$$

Now we are ready to present the main result of this section.

Theorem 4 *In general, the monotonicity of (LCP) does not imply the convexity of F, nor does the convexity of F imply the monotonicity of (LCP).*

Proof: First we shall see that the monotonicity does not imply the convexity in general. It is easy to construct Q and A such that

$$Q + Q(A^T A)^{-1} Q \succeq 0 \text{ but } Q + A^T A \not\succeq 0.$$

Now we choose $b > 0$ sufficiently large in all components. For x and x' small compared to b, we have

$$(Ax' + b)_+ = Ax' + b \text{ and } (Ax + b)_+ = Ax + b.$$

Thus, F cannot be convex in this region since

$$\langle (Ax' + b)_+ - (Ax + b)_+, A\Delta x\rangle + \langle Q\Delta x, \Delta x\rangle = \Delta x^T (A^T A + Q)\Delta x,$$

which can be negative if we appropriately choose x and x'.

To see that the convexity of F does not necessarily imply the monotonicity of (LCP) either, we recall the example which we used in Section 3. Consider

$$F_\epsilon(x_1, x_2) = -\epsilon(x_1 + x_2)^2 + \frac{3\epsilon}{2}(x_1^2 + x_2^2) + (x_1)_+^2 + (x_2)_+^2 + (-x_1 - x_2)_+^2$$

where $0 < \epsilon < 0.5$.

As we showed before, this is a convex quadratic spline function. For this function, we have

$$Q = \epsilon \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \text{ and } A = 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}.$$

Clearly the linear complementarity problem (LCP) associated with this function cannot be monotone if ϵ is chosen to be sufficiently small. To see this, observe that

$$Q + Q(A^T A)^{-1} Q = \epsilon \begin{bmatrix} 1 + 1.1667\epsilon & -2 - 1.0833\epsilon \\ -2 - 1.0833\epsilon & 1 + 1.1667\epsilon \end{bmatrix}$$

is not positive semidefinite for all $0 < \epsilon < 0.5$. □

To check whether the (LCP) is monotone or not is simple. It requires to verify the positive semidefiniteness of a certain matrix. However, the above theorem suggests that checking the global convexity of F is a different matter. The complexity status of the latter problem remains unknown.

5 Concluding Remarks

In this paper we consider the characterization of quadratic spline functions and some related properties. We show that any smooth quadratic spline function can be exclusively represented in the form of the sum of a quadratic function and several squared terms, each of these terms is expressed as the nonnegative part of a certain linear function. It is therefore sufficient to consider the minimization of quadratic splines in such a form. If the first quadratic term is convex, then the whole function is convex, in which case efficient minimization algorithms exist (e.g. [7]). However, it is possible that the first quadratic term is indefinite, while the whole quadratic spline function is globally convex. In this case, it is desirable to express the spline function in a neighborhood of the solution set in the form where the first quadratic term is convex. The reason for demanding a convex quadratic term is that if this is the case then finite convergence is established for some existing minimization algorithms (e.g. [7]). It turns out however, that such an expression may not always be possible in the presence of degeneracy. We further show that in fact the degeneracy is the only cause of complication. Speaking more explicitly, if we assume the function to be nondegenerate at the solution set, then such an expression is always possible. Naturally, this leads to a way to resolve the problem even in the presence of degeneracy, that is, we may resort to the perturbation technique, or more generally, the lexicographic method. To implement the latter method, for example, we need only to introduce a set of conceptual perturbation parameters with different magnitudes. This set of small numbers are added to the vector b componentwise. We then proceed with any Newton-type method with appropriate step-lengths, taking the perturbation parameters into account. Because in this setting no degeneracy persists, we then obtain finite convergence.

Interestingly, we may reformulate the minimization problem in the form of a linear complementarity problem, using the first order optimality condition. It is well known that if an LCP is monotone, then there exist efficient polynomial-time algorithms. This leads us to consider conditions under which such a reformulated LCP problem is indeed monotone. A characterization for this property, in terms of the original data Q and A, is obtained. This condition is compared with the conditions under which the original quadratic spline function is convex. It turns out that these two conditions, both leading to efficient solution methods, do not necessarily imply each other. That is to say, they are essentially quite different conditions.

One problem still remains: How can we efficiently check whether or not a given quadratic spline function is globally convex?

Acknowledgment

The project was initiated when the first author was visiting Technical University of Denmark and Erasmus University. The visit was partially funded by the Danish Natural Science Research Council.

References

[1] B. Chen and M. Pinar, "On Newton's method for Huber's robust M-estimation in linear regression", *BIT Numerical Mathematics 38* (1998), 674-684.

[2] S.-P. Han, "Least-squares solution of linear inequalities", Tech. Report TR-2141, Mathematics Research Center, University of Wisconsin-Madison, (1980).

[3] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Computer Science, Vol 538, Springer Verlag, (1991).

[4] W. Li, "A Merit function and a Newton-type method for symmetric linear complementarity problems", Proceedings of International Conference on Complementarity Problems, held at the Johns Hopkins University, November 1-4, 1995, edited by M. Ferris and J.-S. Pang, SIAM, Philadelphia, 181-203, (1997).

[5] W. Li and J. Swetits, "A Newton method for convex regression, data smoothing and quadratic programming with bounded constraints", *SIAM J. Optim. 3* (1993), 466-488.

[6] W. Li and J. Swetits, "A new algorithm for solving strictly convex quadratic programs", *SIAM J. Optim. 7* (1997), 466-488.

[7] W. Li and J. Swetits, "Regularized Newton methods for minimization of convex quadratic splines with singular Hessians", in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, edited by M. Fukushima and L. Qi, Kluker Academic Publishers, 235-257, (1999).

[8] K. Madsen and H.B. Nielsen, "Finite Algorithms for Robust Linear Regression", *BIT Numerical Mathematics 30* (1990), 682–699.

[9] K. Madsen and H.B. Nielsen, "A Finite Smoothing Algorithm for Linear ℓ_1 Estimation", *SIAM J. on Optimization 3* (1993), 68–80.

[10] O.L. Mangasarian and M.V. Solodov, "Nonlinear complementarity as unconstrained minimization", *Math. Programming 62* (1993), 277-297.

[11] C. Micchelli, P. Smith, J. Swetits, and J. Ward, "Constrained L_p-approximation", *Constr. Approx. 1* (1985), 93-102.

[12] S. Scholtes, "Introduction to piecewise differentiable equations", Preprint No. 53/1994, Institut fur Statistik und Mathematische Wirtschaftstheorie, University of Karlsruhe (1994).

[13] J. Sun, "On piecewise quadratic Newton and trust region problems", *Math. Programming 76* (1997), 451-467.