Statistica Neerlandica (2002) Vol. 56, nr. 3, pp. 376–385

On convex quadratic approximation

Dick den Hertog*

Tilburg University, Department of Econometrics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Etienne de Klerk†

Faculty of Information Technology and Systems, Department TWI/SSOR, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

Kees Roos‡

Faculty of Information Technology and Systems, Department TWI/SSOR, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

Let $n \geqslant 1$ and $f : \mathbb{R}^n \to \mathbb{R}$ a convex function. Given distinct points $z_1, z_2, \ldots, z_N$ in $\mathbb{R}^n$ we consider the problem of finding a quadratic function $g : \mathbb{R}^n \to \mathbb{R}$ such that $\|[f(z_1) - g(z_1), \ldots, f(z_N) - g(z_N)]\|$ is minimal for a given norm $\|\cdot\|$. For the Euclidean norm this is the well-known *quadratic least squares* problem. (If the norm is not specified we will simply refer to g as the *quadratic approximation*.) In this paper we prove the result that the quadratic approximation is not necessarily convex for $n \geq 2$, even though it is convex if $n = 1$. This result has many consequences both for the field of statistics and optimization. We show that the best *convex* quadratic approximation can be obtained in the multivariate case by using semidefinite programming techniques.

Key words and Phrases: convex function, quadratic regression, least squares regression, quadratic interpolation, semidefinite programming.

1 Introduction

Interpolation and approximation are widely used techniques in many research fields; see Box and Draper (1987), Montgomery (1984), and Myers (1999). In this paper we investigate whether the quadratic interpolation and quadratic approximation of a convex function in a finite number of points is convex or not. We call this the convexity preserving property. We will prove that the quadratic approximation is

*D.denHertog@kub.nl
†E.deKlerk@its.tudelft.nl
‡C.Roos@its.tudelft.nl

convexity preserving for the univariate case, but that even the quadratic interpolation function for the multivariate case is not convexity preserving.

To the best of our knowledge these results are not described in the literature. Our conjecture is that the result for the multivariate case has not been discovered since quadratic approximation is mostly used for the univariate case. Also, we could not find a proof in the literature for the convexity preserving property of quadratic approximation for the univariate case.

The consequences of these results are significant, in both the field of statistics and optimization. Several optimization methods use quadratic interpolation or quadratic least squares approximations to (locally) approximate the objective and/or the constraint functions; see BARTHELEMY and HAFTKA (1993), BOOKER *et al.* (1990), CONN and TOINT (1996), CONN *et al.* (1997), DEN HERTOG (1996), DEN HERTOG and STEHOUWER (2000), POWELL (1994), POWELL (1996), SCHOOFS (1987), SOBIESZANSKI-SOBIESKI and HAFTKA (1997), TOROPOV (1992), and TOROPOV *et al.* (1993).

Due to the absence of the convexity preserving property, it may happen that the resulting optimization problem is nonconvex. Such a nonconvex problem is not only difficult to solve, but may also be a bad approximation of the original problem.

We show that convexity can be enforced via semidefinite programming formulations. More precisely, the problem of finding the best convex quadratic approximation in the least squares sense may be formulated as a semidefinite programming problem. Semidefinite programming problems can be solved efficiently nowadays; see ALIZADEH (1991), DE KLERK (1997), NESTEROV and NEMIROVSKII (1992), NESTEROV and NEMIROVSKII (1994), and STURM (1997).

We note that particularly in the field of Computer Aided Design much attention has been given to convexity preserving properties for several interpolation and approximation techniques (KUIJT, 1998, LE MEHAUTE and UTRERAS, 1994). However, this research is mostly restricted to splines and to the univariate and bivariate cases.

This paper is organized as follows. After some preliminaries in Section 2, we treat the univariate case in Section 3. We show that the quadratic approximation is convexity preserving. In Section 4 we give an example for the bivariate case which shows that the quadratic interpolation function is not convexity preserving. We show that requiring convexity of a quadratic approximation leads to a semidefinite programming problem, which can be solved efficiently. In Section 5 we suggest some future research.

2 Preliminaries

Let $n \geqslant 1$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a convex function. Given distinct points $z_1, z_2, \dots, z_N$ in $\mathbb{R}^n$ we consider the problem of finding a quadratic function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$f(z_i) = g(z_i), \quad i = 1, 2, \dots, N. \tag{1}$$

The function g being quadratic, we can write it as

$$g(z) = z^T Qz + r^T z + \gamma \tag{2}$$

for some suitable symmetric $n \times n$ matrix Q, n-vector r and some scalar γ. Hence, the problem of finding g such that (1) holds amounts to finding Q, r and γ such that

$$z_i^T Qz_i + r^T z_i + \gamma = f(z_i), \; i = 1, 2, \ldots, N. \tag{3}$$

This is a linear system of N equations in the unknown entries of Q, r and γ. The number of unknowns in Q is equal to $n + \frac{1}{2}(n^2 - n)$, hence the total number of unknowns is given by

$$n + \tfrac{1}{2}(n^2 - n) + n + 1 = \tfrac{1}{2}(n + 1)(n + 2).$$

Let us call the points $z_1, z_2, \ldots, z_N$ *quadratically independent* if

$$z_i^T Qz_i + r^T z_i + \gamma = 0, \; i = 1, 2, \ldots, N \quad \Rightarrow \quad Q = 0, \; r = 0, \; \gamma = 0. \tag{4}$$

Note that in this case $N \geq \frac{1}{2}(n + 1)(n + 2)$. Moreover, if $N = \frac{1}{2}(n + 1)(n + 2)$ then system (3) has a unique solution. We conclude that if the given points $z_1, z_2, \ldots, z_N$ are quadratically independent and $N = \frac{1}{2}(n + 1)(n + 2)$ then there exists a unique quadratic function g such that (1) holds. This is the interpolation case. When $N > \frac{1}{2}(n + 1)(n + 2)$, the linear system (3) is overdetermined and we can find a least norm solution:

$$\min_{Q,r,\gamma} \|x\|$$

where

$$x_i := z_i^T Qz_i + r^T z_i + \gamma - f(z_i), \; i = 1, \ldots, N.$$

If the norm is the Euclidean norm, then the function g is the quadratic least squares approximation.

3 Quadratic approximation in the univariate case

In this section we consider the univariate case ($n = 1$), *i.e.* f is a one-dimensional convex function. It is obvious that for any three quadratically independent points z_1, z_2, z_3 the function g will be convex. In other words, the quadratic interpolation function is convexity preserving. We proceed to show that also the quadratic approximation is convexity preserving. More precisely, we show that the quadratic approximation g of f with respect to a set of quadratically independent points

$$\mathcal{Z} := \{z_1, z_2, \ldots, z_N\}$$

is convex for any norm.

THEOREM 1. *Let* $z_1 < z_2 < \cdots z_N$ *be quadratically independent points in* $\mathbb{R}$, *and let* $y_i = f(z_i)$ $(i = 1, \dots, N)$, *where* f *is a given univariate convex function. The quadratic approximation to this data set, i.e.* g, *is a convex quadratic function.*

PROOF. Assume that the quadratic approximation g to the data set is strictly concave; see Figure 1.

Now we distinguish between two possibilities:

(i) the function g intersects f in two points;
(ii) the function g intersects f in at most one point.

Case (i) is illustrated in Figure 1. One can now construct the chord through the two points of intersection. This chord then defines an affine function which is clearly a better approximation to the data set at each data point in $\mathcal{Z}$.
In case (ii) the relative interiors of the epigraph of the function f, namely

$$\text{epi}(f) = \{(z, y)|y \geqslant f(z)\},$$

and the set

$$\{(z, y)|y \leqslant g(z)\}$$

are disjoint. These are convex sets, and therefore there exists a line separating them, by the well-known separation theorem for convex sets (see *e.g.* Theorem 11.3 in ROCKAFELLAR 1970). This line again gives a better approximation to the data than g. □

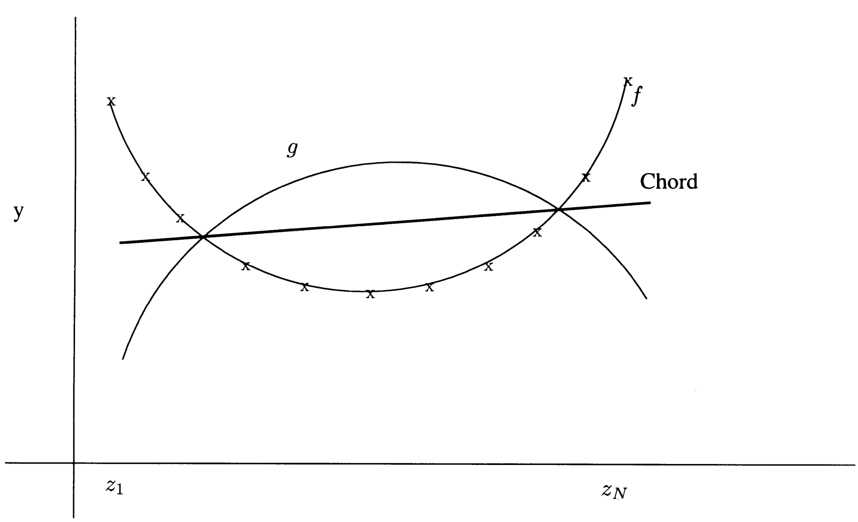

Fig. 1. Illustration of the proof of Theorem 1.

4 Quadratic approximation for the multivariate case

As already stated in the previous section, it is obvious that if $n = 1$ (univariate case) then for any three quadratically independent points z_1, z_2, z_3 the function g will be convex. Surprisingly enough the analogous property does not hold if n is larger than 1 (multivariate case). This means that quadratic interpolation in the multivariate case is not convexity preserving. Consequently, also quadratic approximation in all norms (including 1-norm, 2-norm (least squares), ∞-norm) is not convexity preserving. In this section we will first give a bivariate example for which the quadratic interpolation is not convexity preserving. Then we will show that convexity can be preserved by using semidefinite programming techniques.

A counter-example for the bivariate case

The following (bivariate) example shows that quadratic interpolation is not convexity preserving in multivariate cases.

EXAMPLE 1. Consider the case where f is given by

$$f(x) = -\ln x_1 x_2, \quad x_1 > 0,\ x_2 > 0,$$

which is clearly a convex function, and the points are the 6 columns of the matrix Z given by

$$Z = \begin{pmatrix} 1 & 2 & 3 & 2 & 4 & 6 \\ 2 & 1 & 2 & 3 & 4 & 6 \end{pmatrix}.$$

These points are quadratically independent since the coefficient matrix of the linear system (4), and hence also of (3), is given by

$$\begin{pmatrix} 1 & 2 & 4 & 1 & 2 & 1 \\ 4 & 2 & 1 & 2 & 1 & 1 \\ 9 & 6 & 4 & 3 & 2 & 1 \\ 4 & 6 & 9 & 2 & 3 & 1 \\ 16 & 16 & 16 & 4 & 4 & 1 \\ 36 & 36 & 36 & 6 & 6 & 1 \end{pmatrix}.$$

and this matrix is nonsingular. The (unique, but rounded) solution of (3) is given by

$$Q = \begin{pmatrix} -0.2050 & 0.2628 \\ 0.2628 & -0.2050 \end{pmatrix}, \quad r = \begin{pmatrix} -0.7804 \\ -0.7804 \end{pmatrix}, \quad \gamma = 1.6219.$$

The eigenvalues of Q are -0.4677 and 0.0578, showing that Q is indefinite. Hence the quadratic approximation g of f determined by the given points $z_1, z_2, \ldots, z_6$, is not convex. Figure 4.1 shows some of the level curves of f (dashed) and g (solid) as well as the points $z_i, i = 1, 2 \ldots, 6$.

The level sets of g are clearly not convex and differ very much from the corresponding level sets of f.

In many cases it is important to have a convex quadratic approximation of f. In the next section we show how this can be achieved.

Convex quadratic approximations for the multivariate case

Our aim is to obtain a good convex quadratic approximation g of f on the points in the finite set

$$\mathcal{Z} := \{z_1, z_2, \ldots, z_N\}.$$

Convexity of g is equivalent to the matrix Q in (2) being positive semidefinite, yielding the condition

$$Q \succeq 0. \tag{5}$$

It is clear from the above example that it is impossible to guarantee convexity if we want g to coincide with f on $\mathcal{Z}$. Therefore, to achieve a convex quadratic approximation we need to relax the condition (1). This can be done in several ways. Here we will treat the infinity norm, the 1-norm and the 2-norm.

First one may want to minimize the infinity norm of $f - g$ at $\mathcal{Z}$, yielding the objective

$$\min \max_{z \in \mathcal{Z}} |f(z) - g(z)|. \tag{6}$$

It will be convenient to use the notation

$$s(z) = f(z) - z^T Q z - r^T z - \gamma, \quad z \in \mathcal{Z}.$$

With the above objective we can find g by solving the problem

$$\min(t : -t \leq s(z) \leq t \ (\forall\, z \in \mathcal{Z}),\ Q \succeq 0). \tag{7}$$

One also might minimize the 1-norm of $f - g$ at $\mathcal{Z}$, yielding the objective

$$\min \sum_{z \in \mathcal{Z}} |f(z) - g(z)|. \tag{8}$$

Then g can be found by solving

$$\min\left(\sum_{z \in \mathcal{Z}} t_z : -t_z \leqslant s(z) \leqslant t_z \ (\forall\, z \in \mathcal{Z}),\ Q \succeq 0\right). \tag{9}$$

Finally, we can minimize the 2-norm of $f - g$ at $\mathcal{Z}$ (least squares), yielding the objective

$$\min \sum_{z \in \mathcal{Z}} (f(z) - g(z))^2, \tag{10}$$

and then g can be found by solving

$$\min\left(t : \sqrt{\sum_{z \in \mathcal{Z}} s(z)^2} \leqslant t,\ Q \succeq 0\right). \tag{11}$$

For the first two cases the resulting problems (7) and (9) have linear constraints and a semidefinite constraint $Q \succeq 0$. Such a semidefinite programming problem can efficiently be solved (ALIZADEH, 1991, DE KLERK, 1997, NESTEROV and NEMIROVSKII, 1992, NESTEROV and NEMIROVSKII, 1994, STURM, 1997, VANDENBERGHE and BOYD, 1996). The third resulting problem (11) again can be efficiently solved, since the new constraint is a second order cone (Lorentz cone) constraint (STURM, 1997).

In practice one sometimes wants to add the condition that the approximation is exact or an upper- or underestimate in several points in $\mathcal{Z}$. Observe that such additional properties that $f(z) \geq g(z), z \in \mathcal{Z}$ (or $f(z) \leq g(z), z \in \mathcal{Z}$) then we simply add the constraints $s(z) \geq 0$ (respectively $s(z) \leq 0$) to the above minimization problems. The resulting problems can still be formulated as semidefinite programming problems.

EXAMPLE 2. For the bivariate example given above we calculated the least squares solution while preserving convexity. Using SeDuMi (STURM, 1999) we solved problem (11). We obtained the following (rounded) solution:

$$Q = 0.02750\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad r = -0.7287\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \gamma = 1.2196.$$

The eigenvalues of Q are 0.55 and 0, showing that Q is positive semidefinite. Hence the quadratic approximation g of f determined by the given points $z_1, z_2, \ldots, z_6$, is convex, but degenerate. Note that Q is not positive definite because the constraint $Q \succeq 0$ is binding at the optimal solution of problem (11). (If we remove the constraint $Q \succeq 0$, then we get the non-convex interpolation function of the previous example.)

Figure 3 shows some of the level curves of f (dashed) and g (solid) as well as the points z_i, $i = 1, 2 \ldots, 6$. Comparing this with Figure 2 we see that the convex approximation approximates f much better within the convex hull of the six specified points if the measure of quality is the maximum error or integral of the error function

$$\text{err}(z) = |f(z) - g(z)|$$

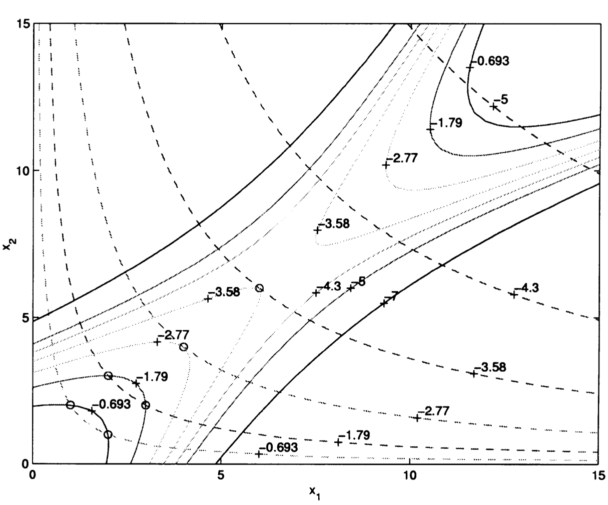

Fig. 2. Level curves of f and g and the points where they coincide.

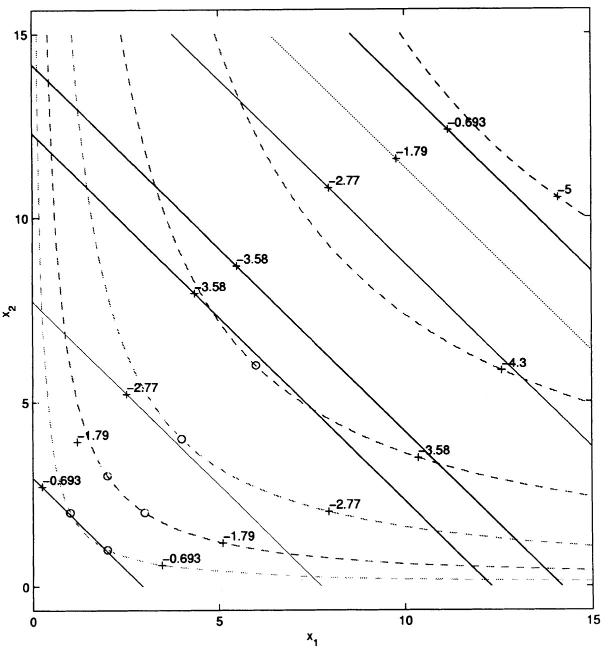

Fig. 3. Level curves of f and g.

over the convex hull. (The convex hull defines a natural trust region for the approximation.)

5 Future research

As already mentioned in the introduction, several optimization methods for solving problems with expensive function evaluations use quadratic interpolation or approximation. A consequence of this paper is that for convex problems the interpolation or approximation may be nonconvex, which may increase the number of iterations of such optimization methods. In the near future we will investigate how we can improve these methods by exploiting the convex structure.

Acknowledgements

We would like to thank Ruud van Lieshout, Kees Vuik and Piet Wesseling for fruitful discussions about Theorem 1, and Wim van Beers for providing us with the references Kuijt (1998) and Le Mehaute and Utreras (1994). We are also indebted to Professor Mike Powell for suggesting a simple proof of Theorem 1.

References

ALIZADEH, F. (1991), *Combinatorial optimization with interior point methods and semi-definite matrices*, Ph.D. Thesis, University of Minnesota, Minneapolis.

BARTHELEMY, J.-F. M. and R. T. HAFTKA (1993), Approximation concepts for optimum structural design – a review, *Structural Optimization* **5**, 129–144.

BOOKER, A. J., J. E. DENNIS, P. D. FRANK, D. B. SERAFINI, V. J. TORCZON and M. W. TROSSET (1999), A rigorous framework for optimization of expensive functions by surrogates, *Structural Optimization* **17**, 1–13.

BOX, G. E. P. and N. R. DRAPER (1987), *Empirical model-building and response surfaces*, Wiley, New York.

CONN, A. R. and Ph.L. TOINT (1996), An algorithm using quadratic interpolation for unconstrained derivative free optimization, in: G. DI PILLO and F. GIANNES (eds.), *Nonlinear Optimization and Applications*, Plenum Publishing, 27–47.

CONN, A. R., K. SCHEINBERG and Ph.L. TOINT (1997), On the convergence of derivative-free methods for unconstrained optimization, in: M. D. BUHMANN and A. ISERLES (eds.), *Approximation Theory and Optimization, tributes to M. J. D. Powell*, Cambridge University Press, 83–108.

HERTOG, D. den (1996), Sequential regression programming, Optix User Guide (confidential).

HERTOG, D. den and H. P. STEHOUWER (2000), Optimizing color picture tubes by high-cost nonlinear programming, to be published in *European Journal on Operational Research*.

KLERK, E. de (1997), *Interior point methods for semidefinite programming*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands.

KUIJT, F. (1998), *Convexity preserving interpolation; stationary nonlinear subdivision and splines*, Ph.D. Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands.

LE MEHAUTE, A. and F. I. UTRERAS (1994), Convexity-preserving interpolatory subdivision, *Computer Aided Geometric Design* **11**, 17–37.

MONTGOMERY, D. C. (1984), *Design and analysis of experiments*, 2nd Edition, Wiley, New York.

MYERS, R. H. (1999), Response surface methodology – current status and future directions, *Journal of Quality Technology* **31**, 30–44.

NESTEROV, Y. and A. A. NEMIROVSKII (1992), Conic formulation of a convex programming problem and duality, *Optimization Methods and Software* **1**(2), 95–115.

NESTEROV, Y. and A. A. NEMIROVSKII (1994), *Interior point polynomial algorithms in convex programming: theory and algorithms*, SIAM Publications, SIAM, Philadelphia.

POWELL, M. J. D. (1994), A direct search optimization method that models the objective and constraint functions by linear interpolation, in: S. GOMEZ and J.-P. HENNART (eds.), *Advances in optimization and numerical analyses, Proceedings of the sixth workshop on optimization and numerical analyses, Oaxaca, Mexico*, Mathematics and its applications, Vol. 275, Kluwer Academic Publishers, Dordrecht, 51–67.

POWELL, M. J. D. (1996), A direct search optimization method that models the objective and constraint functions by linear interpolation, Presentation at the SIAM conference 1996, Virginia.

ROCKAFELLAR, R. T. (1970), *Convex analysis*, Princeton University Press, Princeton, New Jersey.

SCHOOFS, A. J. G. (1987), *Experimental design and structural optimization*, Ph.D. Thesis, Technical University of Eindhoven, Eindhoven, The Netherlands.

SOBIESZANSKI-SOBIESKI, J. and R. T. HAFTKA (1997), Multidisciplinary aerospace design optimization: survey of recent developments, *Structural Optimization* **14**, 1–23.

STURM, J. F. (1997), *Primal-dual interior point approach to semidefinite programming*, Ph.D. Thesis, Tinbergen Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands.

STURM, J. F. (1999), Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optimization Methods and Software*, **11-12**, 625–653.

Toropov, V. V. (1992), Multipoint approximation method in optimization problems with expensive function values, in: A. Sydow (ed.), *Computational System Analysis*, Elsevier, 207–212.

Toropov, V. V., A. A. Filatov and A. A. Polynkine (1993), Multiparameter structural optimization using FEM and multipoint explicit approximations, *Structural Optimization* **6**, 7–14.

Vandenberghe, L. and S. Boyd (1996), Semidefinite programming, *SIAM Review* **38**, 49–95.

Received: June 2001. Revised: January 2002.