NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2000; **7**:133–146

A Newton-type algorithm for solving an extremal constrained interpolation problem

Krassimira Vlachkova*,†

Faculty of Mathematics and Informatics, University of Sofia, Blvd. J. Bourchier 5, 1126 Sofia, Bulgaria

SUMMARY

Given convex scattered data in R^3 we consider the constrained interpolation problem of finding a smooth, minimal L_p-norm $(1 < p < \infty)$ interpolation network that is convex along the edges of an associated triangulation. In previous work the problem has been reduced to the solution of a nonlinear system of equations. In this paper we formulate and analyse a Newton-type algorithm for solving the corresponding type of systems. The correctness of the application of the proposed method is proved and its superlinear (in some cases quadratic) convergence is shown. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: convex interpolation; nonlinear system of equation; iterative methods

1. INTRODUCTION

The problem of interpolating data in 3-dimensional Euclidean space R^3 by smooth surfaces has various applications in both theory and practice and has consequently received considerable attention in the last decades, (see, for example, a survey by Böhm *et al.* [1]). The problem can be formulated as follows: given a set of points (data) $(x_i, y_i, z_i) \in R^3$, $i = 1, \ldots, n$, find a bivariate function S, defined in a certain domain D, that possesses continuous first-order partial derivatives and such that $S(x_i, y_i) = z_i$. In many cases, it is important that the interpolant S satisfies some additional constraints or possesses a certain extremal property.

In Reference [2], Nielson presents an interpolation method based on the so-called minimum norm networks. The method works in the following three steps.

(i) *Triangulation*. The projections $V_i := (x_i, y_i, 0)$, $i = 1, \ldots, n$, are used as the vertices of a triangulation of the domain D in the plane $(x, y, 0)$.
(ii) *Minimum norm network*. The interpolant S and its first-order partial derivatives S_x and S_y are defined on the edges of the triangulation so as to satisfy an extremal property.
(iii) *Blending*. The obtained network is extended to S by an appropriate blending method.

* Correspondence to: Krassimira Vlachkova, Faculty of Mathematics and Informatics, University of Sofia, Blvd. J. Bourchier 5, 1126 Sofia, Bulgaria
† E-mail: krassivl@argo.bas.bg

Received 30 November 1997
Revised 23 March 1999

Copyright © 2000 John Wiley & Sons, Ltd.

In Reference [3] Andersson *et al*. pay special attention to the second step of the above method, i.e. the construction of the minimum norm network. By a different approach, the authors give a new proof of the Nielson's result and additionally consider the case of convex data. In this case, it is natural for the theory, and important for the practice, that the constructed interpolant inherits the convexity of the data.

Andersson *et al*. formulate the corresponding extremal constrained interpolation problem of finding a minimum L_2-norm interpolation network that is convex along the edges of the triangulation. The convex minimum L_2-norm network is characterized as a solution of a nonlinear system of equations. The results from Reference [3] are extended in Reference [4] to the class of L_p-norms for $1 < p \leq \infty$.

The problem of finding a numerical method for the solution of the corresponding non-linear systems of equations naturally arises and its investigation is thereby desirable. In the univariate case, Andersson and Elfving [5] proposed a Newton-type method for the solution of the corresponding constrained interpolation problem. Under certain conditions, the method has been shown to be correct and superlinearly convergent.

In this paper, we apply the approach from Reference [5] and, based on results from References [3, 4], formulate a Newton-type algorithm for solving nonlinear systems that arise from the constrained minimum L_p-norm ($1 < p < \infty$) interpolation problem in R^3. The method has been proposed for the case $p = 2$ in Reference [3] without proper investigation of its validity and convergence. Here we prove that under certain conditions the method correctly applies to the corresponding class of nonlinear systems and achieves superlinear (quadratic for $1 < p \leq 2$) convergence.

The paper is organized as follows: in Section 2 we introduce notation, formulate the extremal problem of interest and present some related results from References [3,4]. We explain how the extremal problem reduces to the solution of a nonlinear system of equations and describe its general type. In Section 3, we formulate a Newton-type algorithm for solving this type of system and prove its validity and convergence. Finally, in Section 4 we make concluding remarks and present some numerical experiments and results.

2. NOTATION AND RELATED RESULTS

Let n ($n \geq 3$) be an integer and let $P_i = (x_i, y_i, z_i)$, $i = 1, \dots, n$, be different points in R^3. We call this set of points *data*. We assume that the projections $V_i = (x_i, y_i)$ of the data into the plane $(x, y, 0)$ are different. Stressing the fact that V_i are in a general position, although not necessarily irregularly placed, we call such data *scattered*.

A polygonal domain D that contains all the points V_i and has vertices at some of them is associated with the data. D is not necessarily simple—it may contain holes.

Next, we say that T is a *triangulation* of the points V_i, $i = 1, \dots, n$, in D if T is a collection of non–overlapping, non–degenerate closed triangles T_{ijk} with vertices V_i, V_j and V_k. Each projection point V_i is a vertex of at least one triangle and no vertex of a triangle lies in the interior of an edge of another triangle. The union of all triangles is the domain D.

The set of edges of the triangles in T is denoted by E. If there is an edge between vertices V_i and V_j in E, it will be referred by e_{ij}. The edges in E define an *edge network* in the domain D, which we denote by E_D (or simply by E if no ambiguity arises).

For a given triangulation T, there is a unique continuous function $L : D \to R^1$ that is linear inside each of the triangles of T and interpolates the data, i.e. $L(V_i) = z_i$ for $i = 1, \dots, n$.

Copyright © 2000 John Wiley & Sons, Ltd.

Next, we define the notion of convex scattered data in a given domain D.

Definition 2.1. *Scattered data in a polygonal domain D are convex if there exists a triangulation T of D such that the corresponding function L is convex. The data are strictly convex if they are convex and the gradient of L has a jump discontinuity across each edge inside D.*

There exist software packages that can test scattered data in a given domain for convexity and construct the corresponding triangulation T. Hereafter, we shall assume that our data in a polygonal domain D are scattered, strictly convex and that T is the corresponding triangulation of D from Definition 2.1.

Definition 2.2. *A curve network is a collection of real-valued univariate functions $\{f_e\}_{e\in E}$, where f_e is defined on the edge $e \in E$.*

With any real-valued bivariate function F defined on D, we naturally associate the curve network defined by the restriction $F_{|E}$ as follows:

$$f_e(t) := F\left(\left(1-\frac{t}{\|e\|}\right)V_i + \frac{t}{\|e\|}V_j\right) = F\left(x_e(t), y_e(t)\right) \tag{2.1}$$

where

$$x_e(t) = \left(1-\frac{t}{\|e\|}\right)x_i + \frac{t}{\|e\|}x_j, \qquad y_e(t) = \left(1-\frac{t}{\|e\|}\right)y_i + \frac{t}{\|e\|}y_j$$

$$0 \le t \le \|e\|, \quad \|e\| := \sqrt{(x_i-x_j)^2+(y_i-y_j)^2}, \quad e \in E$$

In our presentation, according to the context, F will denote either a real-valued bivariate function or a curve network defined by Equation (2.1). For a real p, such that $1 < p < \infty$, we introduce the following class of functions defined on D:

$$\mathscr{F}_p := \{F(x,y) \quad | \quad F(x_i,y_i) = z_i,\ i=1,\ldots,n,\ \partial F/\partial x, \partial F/\partial y \in C(D)$$
$$f_e' \in AC_{[0,\|e\|]},\ f_e'' \in L^p_{[0,\|e\|]},\ e \in E\}$$

and the corresponding class of so-called *smooth interpolation edge-convex curve networks*

$$\mathcal{C}_p(E) := \left\{F_{|E} = \{f_e\}_{\in E} \mid F(x,y) \in \mathscr{F}_p,\ f_e'' \ge 0,\ e \in E\right\}$$

For $F \in \mathcal{C}_p(E)$, we denote the curve network of second derivatives of F by $F'' := \{f_e''\}_{e\in E}$. The L_p-norm of F'' is defined by

$$\|F''\|_p := \left(\int_E |F''|^p \,\mathrm{d}t\right)^{1/p} := \left(\sum_{e\in E}\int_0^{\|e\|} |f_e''(t)|^p \,\mathrm{d}t\right)^{1/p}$$

Now we are ready to formulate the extremal problem of finding the smooth interpolation edge-convex curve network with minimum L_p-norm second derivative that is of interest:

$$(P_p) \quad \text{Find } F^* \in \mathcal{C}_p(E) \text{ such that } \|F^{*\prime\prime}\|_p = \inf_{F\in\mathcal{C}_p(E)} \|F''\|_p$$

Copyright © 2000 John Wiley & Sons, Ltd.

The problem (P_p) has been considered in Reference [3] for $p = 2$ and in Reference [4] for any $p,\ 1 < p \le \infty$. In both cases, it has a unique solution, which has been successfully characterized in References [3,4]. Below we outline the approach for this characterization and formulate the main result of Reference [4], which reduces the solution of (P_p) to the solution of a nonlinear system of equations.

The key observation is that the second derivative $F^{*\prime\prime}$ of the optimal curve network is always the positive part of a curve network that is linear on each of the edges of E. Thereby, the idea is to represent it by simple linear curve networks called *basic curve networks*. As has been proved in Reference [3], these basic curve networks form a basis in a subspace of linear curve networks and can be viewed as a successful generalization of one-dimensional linear B-splines to the case of linear curve networks defined on the edges of a given triangulation.

Recall that e_{ij} denotes the edge between projections V_i and V_j in the triangulation T. For $i = 1, \dots, n$, let m_i be the number of the edges incident to V_i in T. Furthermore, let $\{e_{ii_1}, \dots, e_{ii_{m_i}}\}$ be the edges incident to V_i listed in clockwise order around V_i. The first edge e_{ii_1} is chosen so that the coefficient $\lambda_{1,i}^{(s)}$ defined below is not zero—this is always possible. A *basic curve network* B_{is} is defined for any pair of indices is, such that $i = 1, \dots, n$ and $s = 1, \dots, m_i - 2$. The basic curve network B_{is} is zero on all edges of E except e_{ii_s},$e_{ii_{s+1}}$ and $e_{ii_{s+2}}$, where it is linear. More precisely, B_{is} is defined by

$$B_{is} := \begin{cases} \lambda_{r,i}^{(s)}\left(1 - \frac{t}{\|e_{ii_{s+r-1}}\|}\right) & \text{on } e_{ii_{s+r-1}},\ r = 1, 2, 3, \\ & \quad 0 \le t \le \|e_{ii_{s+r-1}}\| \\ & \\ 0 & \text{on the other edges of } E \end{cases}$$

The coefficients $\lambda_{r,i}^{(s)},\ r = 1, 2, 3,$ are uniquely determined to sum to one and to form a zero linear combination of the three unit vectors along the edges $e_{ii_{s+r-1}}$ starting at V_i.

Note that basic curve networks are associated with points that have at least three edges incident to them. Thus, if a point is incident to two edges only (this might happen on the boundary of D) then no basic curve network is associated with that point. We denote by N_B the set of pairs of indices is for which a basic curve network is defined, i.e.

$$N_B := \{is \mid m_i \ge 3,\ i = 1, \dots, n,\ s = 1, \dots, m_i - 2\}$$

With each basic curve network B_{is} for $is \in N_B$ we associate a number d_{is} defined by

$$d_{is} = \frac{\lambda_{1,i}^{(s)}}{\|e_{ii_s}\|}(z_{i_s} - z_i) + \frac{\lambda_{2,i}^{(s)}}{\|e_{ii_{s+1}}\|}(z_{i_{s+1}} - z_i) + \frac{\lambda_{3,i}^{(s)}}{\|e_{ii_{s+2}}\|}(z_{i_{s+2}} - z_i)$$

which reflects the position of the data in the supporting set of B_{is}. The numbers d_{is} possess interesting properties and can be viewed as a generalization of the second-order divided differences in the univariate case. For us it is important that these numbers are positive when the data are strictly convex. The next theorem characterizes the solution of the problem (P_p).

Theorem 2.1. (Reference [4]). *In the case of strictly convex data the problem (P_p) has a unique*

Copyright © 2000 John Wiley & Sons, Ltd.

solution F^*. The second derivative of the solution $F^{*\prime\prime}$ is of the form

$$F^{*\prime\prime} = \left(\sum_{is \in N_B} \alpha_{is} B_{is} \right)_+^{q-1}$$

where $1/p + 1/q = 1$, $(x)_+ := \max(x, 0)$ and the coefficients α_{is} satisfy the following nonlinear system of equations:

$$\int_E \left(\sum_{is \in N_B} \alpha_{is} B_{is} \right)_+^{q-1} B_{kl} \, \mathrm{d}t = d_{kl}, \qquad \text{for} \quad kl \in N_B. \tag{2.2}$$

3. NEWTON-TYPE ALGORITHM

According to Theorem 2.1, the solution of the extremal constrained interpolation problem (P_p) is reduced to the solution of the system (2.2). When the data are strictly convex, the system (2.2) always has a solution and it is unique (see Reference [3] for the case $p = 2$). The proof of this result in the case $1 < p < \infty$ is similar and we omit it here.

We begin this section by defining a Newton-type algorithm for the solution of the system (2.2). Next, we prove the correctness of the presented algorithm and investigate its convergence.

Let N be the size of the system (2.2) or, which is the same, the number of basic functions B_{is} for $is \in N_B$. We denote by R^N the corresponding linear vector space. It will be convenient to view the vectors in R^N as column vectors. To keep our notation from the previous section we use pairs of indices for the elements α_{is} of the vectors $\alpha \in R^N$. Furthermore, we assume that N_B is lexicographically ordered and thereby

$$\alpha = (\alpha_{11}, \ldots, \alpha_{1m_1-2}, \alpha_{21}, \ldots, \alpha_{n1}, \ldots, \alpha_{nm_n-2})^{\mathrm{T}}$$

The same enumeration is used for the equations of the system (2.2). So, the vector on the right-hand side of system (2.2) becomes

$$d = (d_{11}, \ldots, d_{1m_1-2}, d_{21}, \ldots, d_{n1}, \ldots, d_{nm_n-2})^{\mathrm{T}}$$

Linear combinations of basic curve networks B_{is} that correspond to the vectors $\alpha \in R^N$ are denoted by

$$\varphi(\alpha) := \sum_{is \in N_B} \alpha_{is} B_{is}$$

Definition 3.1. *Strictly convex scattered data in R^3 are degenerate if the solution of the corresponding extremal problem (P_p) has the property that*

$$\varphi(\alpha^*) := \sum_{is \in N_B} \alpha^*_{is} B_{is}$$

is zero on a whole edge of the triangulation T.

Hereafter, we assume that our data are non-degenerate and α is in a neighbourhood of the solution α^*.

Consider now the operator $\Phi : R^N \to R^N$ defined by

$$(\Phi(\alpha))_{kl} = \int_E (\varphi(\alpha))_+^{q-1} B_{kl}\, \mathrm{d}t \tag{3.1}$$

where $1/p + 1/q = 1$. For $q > 1$, the integrals along the edges in definition (3.1) exist and therefore the operator $\Phi(\alpha)$ is well defined in R^N. In our analysis below it will be helpful to represent the positive part of $\varphi(\alpha)$ by its corresponding characteristic curve network $\chi(\alpha)$ defined on the edges $e \in E$ by

$$\chi_e(\alpha) = \chi_e(\alpha; t) = \begin{cases} 1 & \text{if } \varphi_e(\alpha; t) > 0 \\ 0 & \text{if } \varphi_e(\alpha; t) \le 0 \end{cases}$$

where $\varphi_e(\alpha; t)$ is the restriction of $\varphi(\alpha)$ on e and $t \in [0, \|e\|]$. Then definition (3.1) becomes

$$(\Phi(\alpha))_{kl} = \int_E (\varphi(\alpha))^{q-1}\, \chi(\alpha) B_{kl}\, \mathrm{d}t \tag{3.2}$$

In the formulation of our Newton-type algorithm for Equations (2.2) we need the Gâteaux (weak) derivative $\Phi'(\alpha)$ of $\Phi(\alpha)$. First, assuming that $\Phi'(\alpha)$ exists, we find it in an explicit form. Then we shall prove that our computations are valid.

For any fixed $\alpha \in R^N$, the Gâteaux derivative $\Phi'(\alpha)$ (if it exists) is a linear continuous operator from R^N to R^N. The linear operator $\Phi'(\alpha)$ is an $N \times N$ matrix whose elements in our case (see Reference [6], pp. 489–490) are given by

$$\left(\Phi'(\alpha)\right)_{kl,is} = \frac{\partial}{\partial \alpha_{is}} (\Phi(\alpha))_{kl}\,, \qquad \text{where} \quad kl, is \in N_B$$

We formally differentiate in definition (3.2) and obtain

$$\frac{\partial}{\partial \alpha_{is}} (\Phi(\alpha))_{kl} = (q-1) \int_E (\varphi(\alpha))^{q-2}\, \chi(\alpha) B_{kl} B_{is}\, \mathrm{d}t \tag{3.3}$$

To verify the correctness of the differentiation we investigate the continuity with respect to α of the functions in the integrals in Equations (3.2) and (3.3). Then we prove that for any fixed α the operator defined by Equation (3.3) is a linear continuous operator from R^N to R^N.

The functions $(\varphi(\alpha))^{q-1}\, \chi(\alpha)$ in the integrals in definition (3.2) and $(\varphi(\alpha))^{q-2}\, \chi(\alpha)$ in Equation (3.3) for $q \ge 2$ are always continuous with respect to α.‡ In the case $1 < q < 2$, since the data are non-degenerate, the functions in Equation (3.3) are continuous with respect to α.

Now we fix α and consider Equation (3.3). The integrals on the edges in Equation (3.3) readily exist in the case $q \ge 2$. In the case $1 < q < 2$, we observe that the restriction $\varphi_e(\alpha; t)$ is a linear function in $[0, \|e\|]$. Since the data are non-degenerate, $\varphi_e(\alpha; t)$ can have no more than one simple zero in the

‡ Note that this is true even if the data are degenerate. Let $\bar{\alpha}$ be a vector such that $\varphi(\bar{\alpha})$ is zero on a whole edge $e \in E$. Then the characteristic function $\chi(\alpha)$ is not continuous in $\bar{\alpha}$ but since it is bounded and in addition the restriction $\varphi_e(\bar{\alpha}; t) \equiv 0$ for $t \in [0, \|e\|]$, then $(\varphi(\alpha))^{q-1}\, \chi(\alpha)$ and $(\varphi(\alpha))^{q-1}\, \chi(\alpha)$ for $q \ge 2$ are continuous in $\bar{\alpha}$.

Copyright © 2000 John Wiley & Sons, Ltd.

interval $[0, \|e\|]$. Thus, the integral

$$\int_0^{\|e\|} (\varphi_e(\alpha; t))^{q-2}\, \chi_e(\alpha; t) B_{kl}(t) B_{is}(t)\, \mathrm{d}t$$

converges. For any fixed α, the linear operator defined by Equation (3.3) is obviously bounded and therefore continuous. Thus, the Gâteaux derivative exists for any fixed α in a neighbourhood of α^* and is represented by Equation (3.3).

Now, in our notation the system (2.2) becomes

$$\Phi(\alpha) = d \tag{3.4}$$

and the Newton-type algorithm for solving Equation (3.4) follows from the equation

$$\Phi'(\alpha^\nu)(\alpha^{\nu+1} - \alpha^\nu) = d - \Phi(\alpha^\nu)$$

where α^0 is initially chosen and α^ν for $\nu = 1, 2, \ldots$ are consecutive approximations of the solution α^* of Equation (3.4). More precisely, we shall apply the following procedure.

Algorithm 3.1. (Newton-type algorithm)

Step 1. Choose an appropriate initial vector $\alpha^0 \in R^N$, and set an accuracy parameter $\varepsilon > 0$.
Step 2. For $\nu = 0, 1, \ldots$ find the next approximation vector $\alpha^{\nu+1}$ from the following linear system

$$\Phi'(\alpha^\nu)(\alpha^{\nu+1} - \alpha^\nu) = d - \Phi(\alpha^\nu) \tag{3.5}$$

Stop if $\|\alpha^{\nu+1} - \alpha^\nu\| < \varepsilon$.

To prove the correctness and the convergence of this algorithm, we apply a theorem from Reference [7], which in our case states the following.

Theorem 3.1. *(Reference [7]). The approximation sequence $\{\alpha^\nu\}_{\nu=1}^\infty$ is well defined and superlinearly convergent provided*

(i) The operator $\Phi'(\alpha)$ continuously depends on α in a neighborhood of the solution α^*.
(ii) The matrix $\Phi'(\alpha^*)$ is invertible.
(iii) The initial vector α^0 is chosen close enough to α^*.

Additionally, the convergence is at least quadratic if

(iv) The second Gâteaux derivative $\Phi''(\alpha)$ is bounded in a neighbourhood of α^*.

Next until the end of the section, we verify and comment upon the validity of the conditions (*i*), (*ii*), and (*iv*).

The condition (*i*) follows from the continuity with respect to α of the elements of the matrix $\Phi'(\alpha)$, which has been proved already. Note that from the validity of (i) it follows that in a neighbourhood of α^* there exists the Fréchet (strong) derivative of $\Phi(\alpha)$ and both derivatives (weak and strong) coincide.

We establish the condition (*ii*) in the next lemma by showing that the matrix $\Phi'(\alpha^*)$ is positive definite.

Copyright © 2000 John Wiley & Sons, Ltd.

Lemma 3.1. *The matrix $\Phi'(\alpha^*)$ is positive definite.*

Proof
We shall prove that $\beta^{\mathrm{T}}\Phi'(\alpha^*)\beta \geq 0$ for any $\beta \in R^N$. Since $\alpha^* = \{\alpha^*_{is}\}_{is \in N_B}$ solves the system (2.2), we have

$$\int_E \left(\sum_{is \in N_B} \alpha^*_{is} B_{is}(t) \right)_+^{q-1} B_{kl}(t)\,\mathrm{d}t = d_{kl}, \quad kl \in N_B$$

or

$$\int_{\mathrm{supp}(B_{kl})} \left(\sum_{is \in N_B} \alpha^*_{is} B_{is}(t) \right)_+^{q-1} B_{kl}(t)\,\mathrm{d}t = d_{kl}, \quad kl \in N_B \tag{3.6}$$

By the definition, $\mathrm{supp}(B_{kl})$ consists of the three edges $e_{kl},\ e_{kl+1},\ e_{kl+2}$ and since $d_{kl} > 0$ for each $kl \in N_B$, Equation (3.6) implies

$$\max_{\mathrm{supp}(B_{kl})} \left(\sum_{is \in N_B} \alpha^*_{is} B_{is}(t) \right) > 0 \quad \text{for} \quad kl \in N_B \tag{3.7}$$

Let $\beta \in R^N$ be an arbitrary vector. We have

$$\beta^{\mathrm{T}}\Phi'(\alpha^*)\beta = (q-1) \int_E \left(\sum_{is \in N_B} \alpha^*_{is} B_{is}(t) \right)^{q-2} \chi(\alpha^*) \left(\sum_{is \in N_B} \beta_{is} B_{is}(t) \right)^2 \mathrm{d}t \tag{3.8}$$

and consequently $\beta^{\mathrm{T}}\Phi'(\alpha^*)\beta \geq 0$.

Now, we shall prove that

$$\beta^{\mathrm{T}}\Phi'(\alpha^*)\beta = 0 \text{ iff } \beta = 0 \tag{3.9}$$

The sufficiency is obvious. For the necessity, we assume that, for some β, $\beta^{\mathrm{T}}\Phi'(\alpha^*)\beta = 0$. From (3.8) it follows that

$$\chi(\alpha^*) \left(\sum_{is \in N_B} \beta_{is} B_{is}(t) \right) = \chi(\alpha^*)\varphi(\beta) = 0 \tag{3.10}$$

Consider an arbitrary vertex $V_k,\ 1 \leq k \leq n$, of the triangulation and let B_{kl} be the basic curve network for some l, $kl \in N_B$. From (3.7) it follows that on some of the three edges e_{kl}, e_{kl+1}, e_{kl+2}, say e_{kl}, there exists a subinterval (t^1_{kl}, t^2_{kl}) such that the restriction $\chi_{|(t^1_{kl}, t^2_{kl})}(\alpha^*) = 1$. Then from (3.10) we have that

$$\varphi(\beta) = \sum_{is \in N_B} \beta_{is} B_{is}(t) = 0 \quad \text{for every} \quad t \in (t^1_{kl}, t^2_{kl})$$

The basic curve networks B_{is} for $is \in N_B$ are linearly independent and therefore $\beta_{is} = 0$ for those pairs is for which $\mathrm{supp}(B_{is}) \cap (t^1_{kl}, t^2_{kl}) \neq \emptyset$. In particular $\beta_{kl} = 0$ and since kl was chosen arbitrary we have $\beta = 0$. ■

Copyright © 2000 John Wiley & Sons, Ltd.

Next, we investigate the validity of (*iv*). Consider the second Gâteaux (weak) derivative $\Phi''(\alpha)$. For any fixed $\alpha \in R^N$, $\Phi''(\alpha)$ (if it exists) is a linear continuous operator from R^N to the space of the linear continuous operators in R^N and thus it can be considered as a bi-linear operator from $R^N \times R^N$ to R^N (see Reference [6], pp. 646–652, for details). In our case, $\Phi''(\alpha)$ (if it exists) can be defined by N^3 elements

$$\left(\Phi''(\alpha)\right)_{kl,is,jm} = \frac{\partial^2}{\partial\alpha_{is}\partial\alpha_{jm}} (\Phi(\alpha))_{kl}, \quad \text{where} \quad kl, is, jm \in N_B$$

or by the mapping

$$R^N \times R^N \ni \beta^1 \times \beta^2 \longmapsto \beta \in R^N,$$

where, given β^1 and β^2, the vector β is defined by

$$\beta_{jm} = \sum_{kl \in N_B} \sum_{is \in N_B} \frac{\partial^2}{\partial\alpha_{is}\partial\alpha_{jm}} (\Phi(\alpha))_{kl}\, \beta^1_{kl}\beta^2_{is} \tag{3.11}$$

The next lemma shows that the condition (*iv*) is valid in the case $p \leq 2$ (i.e. $q \geq 2$, since $1/p + 1/q = 1$).

Lemma 3.2. *Let $q \geq 2$. The second Gâteaux derivative $\Phi''(\alpha)$ exists and is bounded in a neighbourhood of the solution α^*.*

Proof

Consider the case $q > 2$ first. For $is, kl, jm \in N_B$ from Equation (3.3) we obtain

$$\begin{aligned} \frac{\partial^2}{\partial\alpha_{is}\partial\alpha_{jm}} (\Phi(\alpha))_{kl} &= \frac{\partial}{\partial\alpha_{jm}} \left(\Phi'(\alpha)\right)_{kl,is} \\ &= (q-1)(q-2) \int_E (\varphi(\alpha))^{q-3} \chi(\alpha) B_{kl} B_{is} B_{jm}\, \mathrm{d}t. \end{aligned} \tag{3.12}$$

Similar to the integrals in Equation (3.3), the integrals in Equation (3.12) readily exist for $q \geq 3$. In the case $2 < q < 3$, the existence is ensured by the non-degeneracy of the data and the integrals are convergent since the restrictions of $\varphi(\alpha)$ on the edges of E are linear functions. The operator defined by Equation (3.11) is bi-linear for any fixed $\alpha \in R^N$. Thus, if $q > 2$ the second Gâteaux derivative $\Phi''(\alpha)$ exists for any fixed $\alpha \in R^N$ and is represented by Equation (3.12). Obviously $\Phi''(\alpha)$ is bounded in a neighbourhood of α^*.

Now, consider the case $q = 2$. In this case, the formula (3.12) becomes

$$\left(\Phi''(\alpha)\right)_{kl,is,jm} = \int_E \frac{\partial\chi(\alpha)}{\partial\alpha_{jm}} B_{kl} B_{is}\, \mathrm{d}t \tag{3.13}$$

Let us consider again the restriction $\varphi_e(\alpha; t)$, $t \in [0, \|e\|]$, of the function $\varphi(\alpha)$ on an arbitrary edge $e \in E$. If $\varphi_e(\alpha; t)$ does not change its sign, then from Equation (3.13) the corresponding integral on e is zero. Otherwise, $\varphi_e(\alpha; t)$ has a simple zero $t_0(\alpha) \in (0, \|e\|)$ for which

$$\varphi_e(\alpha; t_0(\alpha)) = 0 \quad \text{and} \quad \varphi'_e(\alpha; t_0(\alpha)) \neq 0 \tag{3.14}$$

Copyright © 2000 John Wiley & Sons, Ltd.

Let $\delta(t)$ denote Dirac's delta function. Although the characteristic function is not differentiable it might be considered as a distribution and the next formal calculations could be easily verified. We have

$$\int_0^{\|e\|} \frac{\partial \chi(\alpha)}{\partial \alpha_{jm}} B_{kl} B_{is}\, \mathrm{d}t = \int_0^{\|e\|} \delta(t - t_0(\alpha)) \frac{\partial t_0(\alpha)}{\partial \alpha_{jm}} B_{kl} B_{is}\, \mathrm{d}t \tag{3.15}$$

We find the derivative $\partial t_0(\alpha)/\partial \alpha_{jm}$ from (3.14) by differentiation:

$$0 = \frac{\partial \varphi_e(\alpha; t_0(\alpha))}{\partial \alpha_{jm}} = B_{jm}(t_0(\alpha)) + \varphi_e'(\alpha; t_0(\alpha)) \frac{\partial t_0(\alpha)}{\partial \alpha_{jm}}$$

and after a substitution in Equation (3.15) we obtain

$$\begin{aligned} \left(\Phi''(\alpha)\right)_{kl,is,jm} &= -\frac{B_{jm}(t_0(\alpha))}{\varphi'(\alpha; t_0(\alpha))} \int_0^{\|e\|} \delta(t - t_0(\alpha)) B_{kl} B_{is}\, \mathrm{d}t \\ &= -\frac{B_{jm}(t_0(\alpha))}{\varphi'(\alpha; t_0(\alpha))} B_{kl}(t_0(\alpha)) B_{is}(t_0(\alpha)) \end{aligned}$$

The last expression is a bounded function of α in a neighbourhood of the solution α^*.

The edge e was chosen arbitrary and thus, in the case $q = 2$, the derivative $\Phi''(\alpha)$ exists for any fixed α and is bounded in a neighbourhood of the solution α^*. ■

Finally, we summarize our results. For non-degenerate data, the algorithm based on Equation (3.5) is well defined and we can construct a sequence of vectors

$$\alpha^0, \alpha^1, \ldots, \alpha^\nu, \ldots \tag{3.16}$$

for which the following theorem applies.

Theorem 3.2. *If the given data are non-degenerate and the initial vector α^0 is appropriately chosen, then the sequence* (3.16) *converges to the solution α^* of the system* (3.4) *and the corresponding sequence of curve networks $(\varphi(\alpha^\nu))_+^{q-1}$ converges to the second derivative of the solution of (P_p), $1 < p < \infty$. The convergence is at least quadratic in the case $1 < p \leq 2$ and superlinear otherwise.*

The results in this section have been obtained by applying and adopting the approach and analysis proposed in Reference [5]. The authors there, among others, considered the problem (P_p) for functions defined on a one-dimensional interval and proved a result analogous to Theorem 3.2 in the univariate case.

4. NUMERICAL RESULTS

In this section, we report briefly the results of some experiments done with Algorithm 3.2 above. We have considered scattered convex data generated from convex bivariate functions with Delaunay or regular triangulation associated with the projection points. The solution of the corresponding nonlinear system (2.2) has been approximated by a sequence $\alpha^1, \alpha^2, \ldots$ computed by iterations (3.5). The particular

Copyright © 2000 John Wiley & Sons, Ltd.

calculations can be simplified by observing that

$$\begin{aligned}
\left(\Phi(\alpha^\nu)\right)_{kl} &= \int_E \left(\sum_{is \in N_B} \alpha^\nu_{is} B_{is} \right)_+^{q-1} B_{kl}\, \mathrm{d}t \\
&= \int_E \left(\sum_{is \in N_B} \alpha^\nu_{is} B_{is} \right)_+^{q-2} \left(\sum_{is \in N_B} \alpha^\nu_{is} B_{is} \right) B_{kl}\, \mathrm{d}t \\
&= \sum_{is \in N_B} \alpha^\nu_{is} \int_E \left(\sum_{is \in N_B} \alpha^\nu_{is} B_{is} \right)_+^{q-2} B_{is} B_{kl}\, \mathrm{d}t \\
&= \frac{1}{q-1} \left(\Phi'(\alpha^\nu)\alpha^\nu\right)_{kl}
\end{aligned}$$

Thus, we have $\Phi(\alpha^\nu) = (1/(q-1))\Phi'(\alpha^\nu)\alpha^\nu$ and the system (3.5) becomes

$$\Phi'(\alpha^\nu)\alpha^{\nu+1} = d + (q-2)\Phi(\alpha^\nu) \tag{4.1}$$

The formula (4.1) is particularly simple in the case $p = q = 2$ and here we deal with this case only. It becomes

$$\Phi'(\alpha^\nu)\alpha^{\nu+1} = d$$

which in our notation is

$$\int_E \left(\sum_{is \in N_B} \alpha^\nu_{is} B_{is}(t) \right)_+^0 \left(\sum_{is \in N_B} \alpha^{\nu+1}_{is} B_{is}(t) \right) B_{kl}(t)\, \mathrm{d}t = d_{kl}, \quad kl \in N_B$$

To choose an initial approximation vector α^0 we use the related unconstrained minimum L_2-norm interpolation problem:

$$(\tilde{P}_2) \quad \text{Find } F^* \in \tilde{\mathcal{C}}_2(E) \text{ such that } \|F^{*\prime\prime}\|_2 = \inf_{F \in \tilde{\mathcal{C}}_2(E)} \|F''\|_2$$

where

$$\tilde{\mathcal{C}}_2(E) := \left\{ F_{|E} = \{f_e\}_{e \in E} \mid F(x, y) \in \mathscr{F}_2 \right\}$$

is the class of *smooth interpolation curve networks* (not necessarily convex along the edges of T). In Reference [3] it is shown that the solution of the problem $(\tilde{P}_2)$ reduces to the solution of the following linear system:

$$\sum_{is \in N_B} \alpha_{is} \int_E B_{is}(t) B_{kl}(t)\, \mathrm{d}t = d_{kl}, \quad kl \in N_B \tag{4.2}$$

So we have used as an initial approximation vector α^0 the solution of the linear system (4.2), i.e. the solution of the unconstrained problem. Other initial approximation vectors we have tested produced

Copyright © 2000 John Wiley & Sons, Ltd.

Table I. Data 1.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	0.00	0.00	0.16	0.20	0.40	0.32	0.48	0.52	0.60	0.68	0.00	0.00
y_i	0.00	1.0	0.32	0.80	0.40	0.76	0.48	0.64	0.40	0.56	0.00	1.0
i	13	14	15	16	17	18	19	20	21	22	23	
x_i	0.00	0.24	0.20	0.20	0.36	0.40	0.52	0.60	0.60	0.80	0.96	
y_i	0.60	0.04	0.60	1.0	0.56	1.0	0.52	0.08	0.80	0.20	0.60	

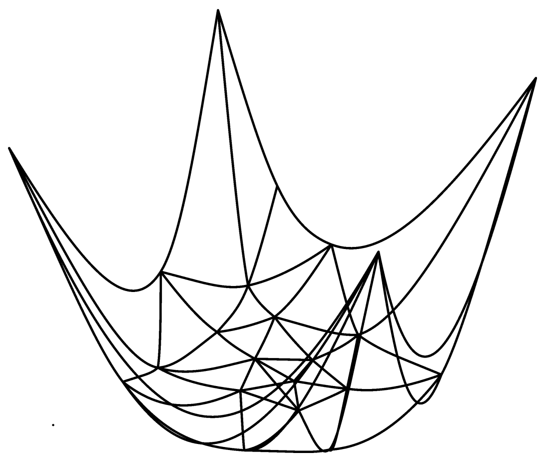

Figure 1. Edge convex minimum norm network based on Data 1.

almost identical convergence patterns. Our experiments suggest that the method converges globally.

The program code is written in FORTRAN and run, in double precision arithmetic, on a PC with standard Pentium processor. For the experiments presented here we used accuracy parameter $\varepsilon = 10^{-8}$. The graphs of the edge-convex minimum L_2-norm curve networks are drawn using the *Mathematica* 2.2 package.

Below we present two of our experiments. In the first example $n = 23$ and the projections V_i of the data are shown in Table I. The associated triangulation in this case is the Delaunay triangulation. The heights z_i are obtained from the bivariate convex function

$$F(x, y) = 2.5e^{(x-0.5)^2+(y-0.5)^2}((x-0.5)^2 + (y-0.5)^2)$$

The stopping criterion $\|\alpha^{\nu+1} - \alpha^{\nu}\| < 10^{-8}$ was satisfied after three iterations and the 3rd and 4th iterations were identical. The corresponding edge convex curve network is shown in Figure 1.

In the second example $n = 30$ and the associated triangulation is a regular mesh with vertices V_i as shown in Table II. The heights z_i are obtained from the function

$$F(x, y) = -\ln^2(16 - (x-3)^2 - (y - 3\sqrt{3}/2)^2)$$

Copyright © 2000 John Wiley & Sons, Ltd.

Table II. Data 2.

i	1	2	3	4	5	6	7	8
x_i	1.5	2.5	3.5	1.0	2.0	3.0	4.0	0.5
y_i	$3\sqrt{3}$	$3\sqrt{3}$	$3\sqrt{3}$	$2.5\sqrt{3}$	$2.5\sqrt{3}$	$2.5\sqrt{3}$	$2.5\sqrt{3}$	$2\sqrt{3}$
i	9	10	11	12	13	14	15	16
x_i	1.5	2.5	3.5	4.5	0.	1.0	2.0	3.0
y_i	$2\sqrt{3}$	$2\sqrt{3}$	$2\sqrt{3}$	$2\sqrt{3}$	$1.5\sqrt{3}$	$1.5\sqrt{3}$	$1.5\sqrt{3}$	$1.5\sqrt{3}$
i	17	18	19	20	21	22	23	24
x_i	4.0	5.	1.5	1.5	3.5	4.5	0.	1.0
y_i	$1.5\sqrt{3}$	$1.5\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$0.5\sqrt{3}$
i	25	26	27	28	29	30		
x_i	2.0	3.0	4.0	1.5	2.5	3.5		
y_i	$0.5\sqrt{3}$	$0.5\sqrt{3}$	$0.5\sqrt{3}$	0.0	0.0	0.0		

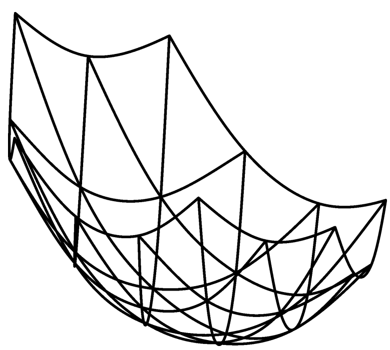

Figure 2. Edge convex minimum norm network based on Data 2.

The stopping criterion was satisfied after three iterations and the 3rd and 4th iterations were identical. The corresponding edge-convex minimum L_2-norm curve network is shown in Figure 2.

ACKNOWLEDGEMENTS

The author thanks the two anonymous referees, whose valuable comments and remarks helped to improve the presentation of this paper.

REFERENCES

1. Böhm W, Farin G, Kahmann J. A survey of curve and surface methods in CAGD, *Computer Aided Geometric Design* 1984; **1**:1–60.

Copyright © 2000 John Wiley & Sons, Ltd.

2. Nielson GM. A method for interpolating scattered data based upon a minimum norm network. *Mathematics of Computation* 1983; **40**:253–271.
3. Andersson L-E, Elfving T, Iliev G, Vlachkova K. Interpolation of convex scattered data in R^3 based upon an edge convex minimum norm network. *Journal of Approximation Theory* 1995; **80**(3):299–320.
4. Vlachkova K. Interpolation of convex scattered data in R^3 based upon a convex minimum L_p norm network, $1 < p \leq \infty$. *Comptes Rendues de l'Academie Bulgare des Sciences* 1992; **45**(12):13–15.
5. Andersson L-E, Elfving T. An algorithm for constrained interpolation, *SIAM Journal on Scientific and Statistical Computing* 1987; **8**(6):1012–1025.
6. Kantorovich LV, Akilov GP. *Functional Analysis*. Nauka: Moscow, 1977 (in Russian).
7. Ortega JM, Rheinboldt WC. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press: New York, 1970.

Copyright © 2000 John Wiley & Sons, Ltd.