Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 222 (2008) 511–523
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

On derivative estimation and the solution of least squares problems

John A. Belward[a,*], Ian W. Turner[b], Miloš Ilić[b]

[a] *Advanced Computational Modelling Centre, University of Queensland, Australia*
[b] *School of Mathematical Sciences, Queensland University of Technology, Australia*

Received 9 January 2007; received in revised form 17 September 2007

Abstract

Surface interpolation finds application in many aspects of science and technology. Two specific areas of interest are surface reconstruction techniques for plant architecture and approximating cell face fluxes in the finite volume discretisation strategy for solving partial differential equations numerically. An important requirement of both applications is accurate local gradient estimation. In surface reconstruction this gradient information is used to increase the accuracy of the local interpolant, while in the finite volume framework accurate gradient information is essential to ensure second order spatial accuracy of the discretisation.

In this work two different least squares strategies for approximating these local gradients are investigated and the errors associated with each analysed. It is shown that although the two strategies appear different, they produce the same least squares error. Some carefully chosen case studies are used to elucidate this finding.

Keywords: Derivative estimation; Heat transfer and diffusion; Surface approximation; Plant architecture

1. Introduction

The desire to determine accurate estimates of the gradient of a function $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, for some domain of interest D, from a set of scattered function values arises in several different ways. In the applications which motivated the work reported here, the need arose in approximating the diffusive flux in finite volume discretisation methods for solving conservation equations numerically [4,14] and in the representation of the surfaces of leaves [1,2,10,13] where a smooth fit to scattered data may be required.

Taylor expansions relate function values to derivatives and these generate linear relations amongst the derivatives and function values, which provides a mechanism for derivative estimation. When the data points are subject to error it is natural to form overdetermined systems of equations and then obtain gradient estimates by minimising residuals.

A simple approach to derivative estimation is to evaluate difference quotients to approximate directional derivatives. The latter are inner products of direction vectors with the gradient and are thus linear combinations of the components of the gradient derivatives. From these relationships a linear system for the gradients can be obtained. However, at best, such approximations only provide first order spatial accuracy and to determine higher orders of accuracy some form of extrapolation may be used.

* Corresponding author. Tel.: +61 7 33788876.
E-mail addresses: jab@maths.uq.edu.au (J.A. Belward), i.turner@qut.edu.au (I.W. Turner), m.ilic@qut.edu.au (M. Ilić).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2007.11.022

We report two methods based on the extrapolation approach; these are superficially different but they lead to the same least squares solution for full rank problems. We present a proof of this result.

Our result is in the context of bivariate interpolation. Considerably more generality is offered by Moving Least Squares (MLS) methods [5,7]. Such methods may be used in an arbitrary number of dimensions. Interpreted in terms of projection operators these local methods lead to powerful mesh independent surface interpolation algorithms for scattered data [8].

Our methods are appropriate when the scale of the problem is such that the presence of noise in the data may be disregarded. For problems in which this assumption is inappropriate regularisation methods may be used. The MLS algorithm of Reference [8] uses Backus–Gilbert and the work of Ling, [9], uses Tikhonov regularisation on this important class of problems.

The structure of the paper is as follows. In Section 2 we review the literature on surface fitting for plant leaves and outline the least squares approach used for local gradient estimation when constructing a surface interpolant [10]. In Section 3 we give a brief overview of the finite volume method for solving diffusion equations and provide the motivation for accurate gradient estimation to ensure second order spatial accuracy of the discretisation. The least squares strategy is again instrumental in this gradient approximation and the method employed in [4,14] is presented. Although this solution process appears somewhat different to that utilised for surface fitting given in [10], in Section 4 we prove that the two methods give the same least squares solution. In Section 5 numerical case studies are presented exhibiting the second order accuracy of these methods. This section also includes comparative comments on Ling's results and addresses a complexity issue. Finally in Section 6 the conclusions of the research are summarised and we hint at future research directions.

2. Surface fitting for leaves of plants

Surface fitting arises in many different applications. The latest application for the authors is in the measurement of leaves of plants to capture their image for simulation and modelling purposes.

Many representations of surfaces are possible. Here our representation is as a function from $\mathbb{R}^2$ to $\mathbb{R}$, $z = f(x, y)$, thus a reference plane is assumed to exist, with a unique ordinate at each data point in the xy-plane. This is a restriction that will be reasonable in many situations, but will not be valid in situations where leaves curl sufficiently. The placement of points in the reference plane cannot be totally controlled, thus the situation is one of scattered data points.

In order to obtain a surface with a continuously turning normal, i.e. continuous gradient, we chose to represent the function by a set of cubic elements defined on a union of triangular domains obtained by a Delaunay triangulation of the data points in the reference plane. This plane was determined by a least squares fit of a single affine function to the data points. We represented the surface by using Clough–Tocher basis functions (see, for example, [6]). A piecewise cubic surface with continuous gradient may be obtained if the function values and the gradient are known at the original data points and the gradient is known at the mid-points of the edges of the triangulation. In fact the same is true if the gradient at the mid-point of an edge is approximated by the arithmetic mean of its value at the nodes. In the latter case we can only reproduce quadratics exactly, whereas cubics are reproduced if the true values of the gradient at the mid-points are used.

There are many alternative ways of interpolating a scattered data set. Our method is taken from Lancaster and Salkauskas [6]. A survey may be found in the paper in [12] with further references to the literature within.

Since the problem was specified with data values prescribed at the scattered points it was necessary to determine the gradient at these points. The problem becomes: Given the values z_i at the points $(x_i, y_i)^{\mathrm{T}}$ estimate the values of the gradient, $\nabla f(x_i, y_i)$.

An estimate of the directional derivative at a point may be obtained simply by calculating the difference quotient in the direction in question. Since the directional derivative is the inner product of the direction with the gradient the difference quotient provides information for an approximation to the gradient. By evaluating several difference quotients we can obtain a set of linear relations from which a linear system may be obtained whose solution is a first order approximation to the gradient.

In order to describe the gradient estimation strategy, let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ be differentiable at $\mathbf{x}_0 \in D$, define Frechet derivative $df_{\mathbf{x}_0} : \mathbb{R}^2 \to \mathbb{R}$ as $df_{\mathbf{x}_0}(\mathbf{h}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h}$, $\mathbf{h} = (h_1, h_2)^{\mathrm{T}}$ and let the scalar function $E(\mathbf{x}_0, \mathbf{h})$ be

such that $\lim_{\mathbf{h}\to\mathbf{0}} E(\mathbf{x}_0, \mathbf{h}) = 0$ then

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \|\mathbf{h}\| E(\mathbf{x}_0, \mathbf{h}). \tag{2.1}$$

Suppose that $\mathbf{x}_0 = (x_0, y_0)^{\mathrm{T}}$ and we are given p scattered data points $\mathbf{x}_i = (x_i, y_i)^{\mathrm{T}} \in \mathcal{B}_\delta(\mathbf{x}_0),\ i = 1, \ldots, p$ with δ suitably chosen and we require an estimate of the gradient $\nabla f(\mathbf{x}_0)$. The relation (2.1) can be used to write an overdetermined system of equations for $\mathbf{g} = (\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y})^{\mathrm{T}}$ in the form:

$$\mathbf{Dg} = \boldsymbol{\delta}\mathbf{f}, \tag{2.2}$$

whose least squares solution, according to (2.1), will provide a first order estimate of the required gradient. The matrix $\mathbf{D} \in \mathbb{R}^{p\times 2}$ is given by $\begin{pmatrix} \delta x_1 & \delta y_1 \\ \vdots & \vdots \\ \delta x_p & \delta y_p \end{pmatrix}$ with $\delta x_i = x_i - x_0$, $\delta y_i = y_i - y_0$ and the right hand side vector $\boldsymbol{\delta}\mathbf{f} \in \mathbb{R}^{p\times 1}$ has its ith component given by $\boldsymbol{\delta}\mathbf{f}_i = f(\mathbf{x}_0 + \delta\mathbf{x}_i) - f(\mathbf{x}_0)$. Supposing now that I is an open interval in $\mathbb{R}$ containing $[0, 1]$, $f \in C^{r+1}$ in D and $\mathbf{x}_0 + t\mathbf{h} \in D,\ \forall t \in I$, then Taylor's Theorem for Several Variables states that $\exists\theta,\ 0 < \theta < 1$, such that

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \frac{(\mathbf{h}\cdot\nabla) f(\mathbf{x}_0)}{1!} + \cdots + \frac{(\mathbf{h}\cdot\nabla)^r f(\mathbf{x}_0)}{r!} + R_r,$$

where the remainder R_r has the Lagrange form

$$R_r = \frac{(\mathbf{h}\cdot\nabla)^{r+1} f(\mathbf{x}_0 + \theta\mathbf{h})}{(r+1)!}.$$

Truncating the Taylor expansion for f after the second order term enables the following overdetermined system to be solved, in the least squares sense, to obtain a second order approximation for the components of $\nabla f(\mathbf{x}_0)$:

$$(\mathbf{D}\dot{:}\mathbf{M}) \begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} = \boldsymbol{\delta}\mathbf{f}, \tag{2.3}$$

where now the matrix $\mathbf{M} \in \mathbb{R}^{p\times 3}$ is given by $\begin{pmatrix} \frac{1}{2}\delta x_1^2 & \delta x_1\delta y_1 & \frac{1}{2}\delta y_1^2 \\ \vdots & \vdots & \\ \frac{1}{2}\delta x_p^2 & \delta x_s\delta y_p & \frac{1}{2}\delta y_p^2 \end{pmatrix}$ and $\mathbf{z} = (\frac{\partial^2 f}{\partial x^2},\ \frac{\partial^2 f}{\partial x\partial y},\ \frac{\partial^2 f}{\partial y^2})^{\mathrm{T}}$. Eq. (2.3) can also be written in the expanded form

$$\mathbf{Dg} = \boldsymbol{\delta}\mathbf{f} - \mathbf{Mz}, \tag{2.4}$$

from which it can be seen that a more accurate estimate of the gradient than that offered by Eq. (2.2) can be obtained if the second order derivative terms are eliminated from the system. This elimination can be performed using an orthogonal reduction of $\mathbf{M}$, namely $\mathbf{Q}^{\mathrm{T}}\mathbf{M} = \mathbf{T}$ with $\mathbf{Q}^{\mathrm{T}} \in \mathbb{R}^{p\times p}$ constructed using the product of elementary reflectors and $\mathbf{T} \in \mathbb{R}^{p\times 3}$ has upper-trapezoidal form.

Applying $\mathbf{Q}^{\mathrm{T}}$ to (2.4) produces

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Dg} = \mathbf{Q}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} - \mathbf{Tz}, \tag{2.5}$$

and discarding the first three equations from system (2.5) ensures that the remaining equations will not contain the unknown second derivatives. The least squares solution of the remaining $p-3$ equations then provides a second order accurate estimate of the gradient $\mathbf{g}$.

Note also that each of the estimates of the directional derivative may be weighted without loss of second order accuracy. This follows since the effect of a weight factor w_i is to introduce a diagonal matrix $\mathbf{W} = \mathrm{diag}(w_1, w_2, \ldots, w_s)$, where typically one would use inverse distance, or inverse distance squared, weights $w_i = \|\boldsymbol{\delta}\mathbf{x}_i\|^{-d}$, $d = 1, 2$ to give more significance to points closer to $\mathbf{x}_0$. In this case the system (2.4) becomes

$$\mathbf{WDg} = \mathbf{W}\boldsymbol{\delta}\mathbf{f} - \mathbf{WMz}$$

and now we apply orthogonal reduction to $\mathbf{WM}$ and then follow the steps of the earlier paragraph.

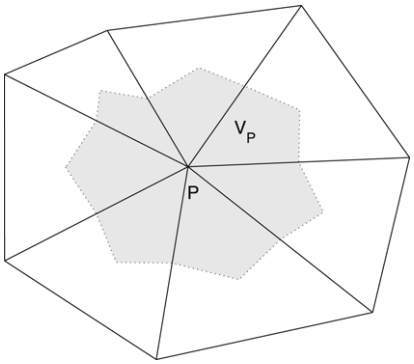

Fig. 1.

Although the initial thought might be that the ordering of the equations would have some impact on the gradient estimation process, this is indeed not the case. To see why, let $\mathbf{R} \in \mathbb{R}^{p\times p}$ be a permutation matrix that permutes the rows of $\mathbf{M}$. Because the orthogonal reduction of $\mathbf{M}$ produces unique matrices $\mathbf{Q}$ and $\mathbf{T}$ such that $\mathbf{Q}^{\mathrm{T}}\mathbf{M} = \mathbf{T}$, it follows that applying orthogonal reduction to the permuted system $\mathbf{RMx} = \mathbf{R}\delta\mathbf{f}$ yields with $\tilde{\mathbf{Q}}^{\mathrm{T}}\mathbf{RM} = \mathbf{T}$ and $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}\tilde{\mathbf{Q}}$ exactly the same system as (2.5).

3. Flux calculation

The need to solve two-dimensional conservative laws numerically finds application in many fields of science and engineering and a method well suited for the task is the finite volume method (FVM). Typically, this method sees the solution domain tessellated with control volumes, or cells, (see Fig. 1) and the conservation law is then integrated over the cell in a discrete sense. It is well known that the spatial accuracy of the FVM is influenced to a large extent by two factors—the quadrature rule used to approximate the surface integral of the flux and the estimation of the flux itself at a cell face using only the cell and neighbouring nodal information. Since for diffusion equations the flux is proportional to the gradient of the transported quantity, substantial effort is placed on its accurate reconstruction at the cell face. Numerous methods have been investigated for this purpose, including Gauss reconstruction methods, finite element shape function methods and least squares methods. These methods have been explored extensively in [4,14] and the wealth of other references cited there.

We illustrate the FVM discretisation process for a representative diffusion equation of the following form:

$$\frac{\partial\psi(\varphi)}{\partial t} = \nabla\cdot(\mathbf{D}\nabla\varphi) + S(\varphi). \tag{3.1}$$

Eq. (3.1) is transformed into control-volume form by integrating over the control volume V_P:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_P}\psi(\varphi)\,\mathrm{d}V - \int_{V_P}\nabla\cdot(\mathbf{D}\nabla\varphi)\mathrm{d}V - \int_{V_P}S\mathrm{d}V = 0. \tag{3.2}$$

Applying the Divergence theorem to (3.2), and defining

$$\bar{\psi} = \frac{1}{\Delta V_P}\int_{V_P}\psi\mathrm{d}V, \qquad \bar{S} = \frac{1}{\Delta V_P}\int_{V_P}S\mathrm{d}V \tag{3.3}$$

as the volume-averaged values of the accumulation term and source term respectively, we obtain (in a two-dimensional setting):

$$\frac{\mathrm{d}\bar{\psi}}{\mathrm{d}t} - \frac{1}{\Delta A_P}\int_{C_P}\mathbf{D}\nabla\varphi\cdot\hat{\mathbf{n}}\mathrm{d}s - \bar{S} = 0 \tag{3.4}$$

which, since no approximation has been made at this stage, is an exact reformulation of (3.1). In Eq. (3.4) ΔA_P represents the cell area and C_P the line integral path that defines the cell surface (or boundary).

The FVM discretisation process proceeds by approximating the cell averages with their corresponding cell nodal values and using the mid-point quadrature rule to approximate the line integral:

$$\frac{\mathrm{d}\psi_P}{\mathrm{d}t} - \frac{1}{\Delta A_P} \sum_{j \in \mathcal{N}_p} \left(\mathbf{D}\nabla\varphi \cdot \hat{\mathbf{n}}\right)_j - S_p \approx 0, \tag{3.5}$$

where subscript j indicates evaluation at the mid-point of the jth cell face and $\mathcal{N}_p$ is the set of cell faces that defines C_P. One notes from (3.5) that to evaluate this discrete finite volume equation it is necessary to approximate $\nabla\varphi$ at the cell face mid-point.

In this work we analyse the least squares strategy proposed in [4] for this gradient estimation. One obvious difference to the method outlined in Section 2 is that the value of the dependent variable φ (the conserved quantity) is unknown at the cell face and consequently, it is necessary to solve a linear least squares problem with six unknowns, namely the function value, its gradient components and the second derivative values. In [4] the gradient estimation was obtained by solving the normal equations for the least squares solution and extracting the relevant components.

An alternate formulation [3] sees the gradient estimation required at the vertices of the triangles where the value of φ is known and in this case the resulting overdetermined system with five unknowns is identical to the system described by (2.3) in Section 2. If we apply orthogonal reduction directly to this system:

$$\mathbf{P}^{\mathrm{T}}(\mathbf{D}\dot{:}\mathbf{M}) \begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} = \mathbf{P}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} \tag{3.6}$$

an upper triangular system can be solved and the relevant gradient approximation extracted. Denoting $\mathbf{P}^{\mathrm{T}}(\mathbf{D}\dot{:}\mathbf{M}) = \begin{pmatrix} \mathbf{U} \\ \mathbf{0} \end{pmatrix}$ and $\mathbf{P}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}$ the least squares solution $\mathbf{x}$ is obtained by solving $\mathbf{U}\mathbf{x} = \mathbf{c}_1$ with error given by $\|\mathbf{c}_2\|_2^2$. The relevant gradient approximation is then taken as $\frac{\partial f}{\partial x} \approx x_1$ and $\frac{\partial f}{\partial y} \approx x_2$.

We show in the next section that the two least squares methods outlined in Sections 2 and 3 lead to identical gradient approximations.

4. Uniqueness of the gradient estimates

The uniqueness of the gradient estimates presented in Sections 2 and 3 is not a result of the analytical properties of the approximation processes, it is a consequence of the method of linear least squares. There is no loss of generality in demonstrating this uniqueness for a system with five unknowns as arises in the gradient estimation processes discussed in Sections 2 and 3.

Denote the residual for system (2.3) by

$$\mathbf{E} = \mathbf{A}\mathbf{x} - \boldsymbol{\delta}\mathbf{f} \tag{4.1}$$

where $\mathbf{x} = (\mathbf{g}, \mathbf{z})^{\mathrm{T}}$ and $\mathbf{A} = (\mathbf{D}\dot{:}\mathbf{M})$ is assumed to have full rank. $\mathbf{D}$, $\mathbf{M}$ and $\boldsymbol{\delta}\mathbf{f}$ are as defined in Section 2. We wish to find

$$e = \min_{\mathbf{x} \in \mathbb{R}^5} \|\mathbf{E}\|_2^2 = \min_{\mathbf{x} \in \mathbb{R}^5} \|\mathbf{A}\mathbf{x} - \boldsymbol{\delta}\mathbf{f}\|_2^2 . \tag{4.2}$$

Method 1

As described in Section 2, the orthogonal reduction of $\mathbf{M}$ enables the system residual to be written as

$$(\mathbf{Q}^{\mathrm{T}}\mathbf{D}\dot{:}\mathbf{Q}^{\mathrm{T}}\mathbf{M}) \begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} - \mathbf{Q}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \mathbf{Q}^{\mathrm{T}}\mathbf{E}. \tag{4.3}$$

After partitioning $\mathbf{Q}^{\mathrm{T}}\mathbf{D}$ and $\mathbf{Q}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}$ the left hand side of the above equation becomes

$$\begin{pmatrix} \mathbf{D}_{11} & \tilde{\mathbf{T}} \\ \mathbf{D}_{21} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} - \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}, \tag{4.4}$$

where $\mathbf{D}_{11} \in \mathbb{R}^{3\times 2}$, $\mathbf{D}_{21} \in \mathbb{R}^{(p-3)\times 2}$, $\tilde{\mathbf{T}} \in \mathbb{R}^{3\times 3}$, $\mathbf{d}_1 \in \mathbb{R}^{3\times 1}$ and $\mathbf{d}_2 \in \mathbb{R}^{(p-3)\times 1}$, we then expand and take norms in (4.3) to obtain:

$$\|\mathbf{E}\|_2^2 = \left\|\mathbf{Q}^{\mathrm{T}}\mathbf{E}\right\|_2^2 = \left\|\mathbf{D}_{11}\mathbf{g} + \tilde{\mathbf{T}}\mathbf{z} - \mathbf{d}_1\right\|_2^2 + \|\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2\|_2^2 . \tag{4.5}$$

Let $\mathbf{x}_1$ be the minimiser of the second term on the right hand side, namely

$$\mathbf{x}_1 = \arg\min_{\mathbf{g}\in\mathbb{R}^2} \|\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2\|_2^2 . \tag{4.6}$$

Once $\mathbf{x}_1$ is determined, it then can be used to compute $\mathbf{z}$ by solving $\tilde{\mathbf{T}}\mathbf{z} = \mathbf{d}_1 - \mathbf{D}_{11}\mathbf{x}_1$ so that the first term on the right hand side of (4.5) can be made zero.

Using this strategy, we denote the least squares error for Method 1 as

$$e_1 = \|\mathbf{D}_{21}\mathbf{x}_1 - \mathbf{d}_2\|_2^2 . \tag{4.7}$$

The gradient approximation is then given by $\frac{\partial f}{\partial x} \approx (\mathbf{x}_1)_1$ and $\frac{\partial f}{\partial y} \approx (\mathbf{x}_1)_2$.

Method 2

Suppose we now consider the orthogonal reduction discussed for the flux estimation problem discussed in Section 3. In this case we have

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{x} - \mathbf{P}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \mathbf{P}^{\mathrm{T}}\mathbf{E} \tag{4.8}$$

with $\mathbf{P}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix}$ where $\mathbf{c}_1 \in \mathbb{R}^{5\times 1}$, $\mathbf{c}_2 \in \mathbb{R}^{(p-5)\times 1}$ and $\mathbf{P}^{\mathrm{T}}\mathbf{A} = \begin{pmatrix}\mathbf{U}\\ \mathbf{0}\end{pmatrix}$ where $\mathbf{U} \in \mathbb{R}^{5\times 5}$ is upper triangular, and

$$\|\mathbf{E}\|_2^2 = \|\mathbf{U}\mathbf{x} - \mathbf{c}_1\|_2^2 + \|\mathbf{c}_2\|_2^2 \tag{4.9}$$

so that we can write the least squares error given in (4.2) as

$$e = \min_{\mathbf{x}\in\mathbb{R}^5} \|\mathbf{E}\|_2^2 = \|\mathbf{c_2}\|_2^2 . \tag{4.10}$$

The gradient approximation in this case is obtained by extracting the first two components of the least squares solution $\mathbf{x} = \mathbf{U}^{-1}\mathbf{c}_1$.

It is interesting to note that although the gradient approximations for Method 1 and Method 2 are determined using quite different approaches the two methods produce the same error, which we now formally state and prove.

Proposition 1. *The errors e defined in (4.10) and e_1 defined in (4.7) are equal.*

Proof. First observe that

$$e = \min_{\mathbf{x}\in\mathbb{R}^5} \|\mathbf{E}\|_2^2 \le \|\mathbf{A}\mathbf{x} - \boldsymbol{\delta}\mathbf{f}\|_2^2 , \quad \forall\, \mathbf{x} \in \mathbb{R}^5 \tag{4.11}$$

or, according to (4.5)

$$e \le \left\|\mathbf{D}_{11}\mathbf{g} + \tilde{\mathbf{T}}\mathbf{z} - \mathbf{d}_1\right\|_2^2 + \|\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2\|_2^2 , \ \mathbf{x} = (\mathbf{g}, \mathbf{z})^{\mathrm{T}}. \tag{4.12}$$

In particular, with the choice $\mathbf{x} = (\mathbf{x}_1, \mathbf{z})^{\mathrm{T}}$, with $\mathbf{x}_1$ and $\mathbf{z}$ as given above, we obtain

$$e \le \|\mathbf{D}_{21}\mathbf{x}_1 - \mathbf{d}_2\|_2^2 = e_1. \tag{4.13}$$

Our objective to complete the proof is to show that $e \ge e_1$ implying that $e = e_1$. First partition $\mathbf{Q} = (\mathbf{Q}_1 \vdots \mathbf{Q}_2)$ with the matrices $\mathbf{Q}_1 \in \mathbb{R}^{p\times 3}$ and $\mathbf{Q}_2 \in \mathbb{R}^{p\times(p-3)}$. Then

$$\mathbf{Q}^{\mathrm{T}}\mathbf{A} = \begin{pmatrix}\mathbf{Q}_1^{\mathrm{T}}\mathbf{D} & \mathbf{Q}_1^{\mathrm{T}}\mathbf{M}\\ \mathbf{Q}_2^{\mathrm{T}}\mathbf{D} & \mathbf{Q}_2^{\mathrm{T}}\mathbf{M}\end{pmatrix}$$

so that in (4.4), $\mathbf{D}_{11} = \mathbf{Q}_1^{\mathrm{T}}\mathbf{D}$, $\mathbf{D}_{21} = \mathbf{Q}_2^{\mathrm{T}}\mathbf{D}$, $\mathbf{Q}_1^{\mathrm{T}}\mathbf{M} = \tilde{\mathbf{T}}$ and $\mathbf{Q}_2^{\mathrm{T}}\mathbf{M} = \mathbf{0}$.

Using $\mathbf{d}_2 = \mathbf{Q}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}$ we see that

$$\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2 = \mathbf{Q}_2^{\mathrm{T}}(\mathbf{D}\mathbf{g} - \boldsymbol{\delta}\mathbf{f}). \tag{4.14}$$

Now partition $\mathbf{P} = (\mathbf{P}_1 \vdots \mathbf{P}_2)$, where $\mathbf{P}_1 \in \mathbb{R}^{p\times 5}$,

$$\mathbf{P}^{\mathrm{T}}\mathbf{A} = \begin{pmatrix}\mathbf{P}_1^{\mathrm{T}}\\ \mathbf{P}_2^{\mathrm{T}}\end{pmatrix}\mathbf{A} = \begin{pmatrix}\mathbf{U}\\ \mathbf{0}\end{pmatrix} \quad \text{and} \quad \mathbf{P}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \begin{pmatrix}\mathbf{P}_1^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}\\ \mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}\end{pmatrix} = \begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix}.$$

From this partitioning we deduce that $\mathbf{c}_2 = \mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}$, $\mathbf{P}_2^{\mathrm{T}}\mathbf{A} = \mathbf{0}$ and hence, $\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}$ is a projector onto $\mathcal{N}(\mathbf{A}^{\mathrm{T}})$, the left hand null space of $\mathbf{A}$.

Denoting $\mathbf{x}$ as the least squares solution of Eq. (4.1), the residual is simply the projection of $\boldsymbol{\delta}\mathbf{f}$ onto $\mathcal{N}(\mathbf{A}^{\mathrm{T}})$, namely

$$\mathbf{A}\mathbf{x} - \boldsymbol{\delta}\mathbf{f} = \mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f},$$

which implies that $e = \left\|\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}\right\|_2^2$.

Furthermore, we can write with $\mathbf{x} = (\bar{\mathbf{x}}_1, \bar{\mathbf{x}_2})^{\mathrm{T}}$ so that

$$\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \mathbf{D}\bar{\mathbf{x}}_1 + \mathbf{M}\bar{\mathbf{x}}_\mathbf{2} - \boldsymbol{\delta}\mathbf{f}. \tag{4.15}$$

Applying $\mathbf{Q}_2^{\mathrm{T}}$ to (4.15) gives:

$$\mathbf{Q}_2^{\mathrm{T}}\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \mathbf{Q}_2^{\mathrm{T}}(\mathbf{D}\bar{\mathbf{x}}_1 - \boldsymbol{\delta}\mathbf{f}) + \mathbf{Q}_2^{\mathrm{T}}\mathbf{M}\bar{\mathbf{x}}_2$$

and since $\mathbf{Q}_2^{\mathrm{T}}\mathbf{M} = \mathbf{0}$, we obtain

$$\mathbf{Q}_2^{\mathrm{T}}\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f} = \mathbf{Q}_2^{\mathrm{T}}\mathbf{r}_1 \quad \text{where } \mathbf{r}_1 = \mathbf{D}\bar{\mathbf{x}}_1 - \boldsymbol{\delta}\mathbf{f}.$$

Taking norms and using $\mathbf{Q}_2^{\mathrm{T}}\mathbf{Q}_2 = \mathbf{I}$ we have:

$$\left\|\mathbf{Q}_2\mathbf{Q}_2^{\mathrm{T}}\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}\right\| = \left\|\mathbf{Q}_2^{\mathrm{T}}\mathbf{r}_1\right\|.$$

Noting that $\mathbf{Q}_2\mathbf{Q}_2^{\mathrm{T}}$ is an orthogonal projector onto null space $\mathcal{N}(\mathbf{M}^{\mathrm{T}})$ and since $\mathcal{N}(\mathbf{A}^{\mathrm{T}}) \subset \mathcal{N}(\mathbf{M}^{\mathrm{T}})$ we deduce that

$$\mathbf{Q}_2\mathbf{Q}_2^{\mathrm{T}}(\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}) = \mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}. \tag{4.16}$$

Finally, $e = \left\|\mathbf{P}_2\mathbf{P}_2^{\mathrm{T}}\boldsymbol{\delta}\mathbf{f}\right\|_2^2 = \|\mathbf{D}_{21}\bar{\mathbf{x}}_1 - \mathbf{d}_2\|_2^2 \geq \|\mathbf{D}_{21}\mathbf{x}_1 - \mathbf{d}_2\|_2^2 = e_1$. Thus, we have the main result $e = e_1$ as asserted. □

It should be noted that the result $e \geq e_1$ in the above proof was obtained from a linear algebra perspective. It provides us with the insight to derive an alternative proof based on analytic considerations, which we state as follows:

$$\begin{aligned} e &= \min_{\mathbf{x}\in\mathbb{R}^5} \|\mathbf{E}\|_2^2 = \min_{\mathbf{x}\in\mathbb{R}^5}\left\{\left\|\mathbf{D}_{11}\mathbf{g} + \tilde{\mathbf{T}}\mathbf{z} - \mathbf{d}_1\right\|_2^2 + \|\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2\|_2^2\right\}\\ &\geq \min_{\mathbf{x}\in\mathbb{R}^5}\left\|\mathbf{D}_{11}\mathbf{g} + \tilde{\mathbf{T}}\mathbf{z} - \mathbf{d}_1\right\|_2^2 + \min_{\mathbf{x}\in\mathbb{R}^5}\|\mathbf{D}_{21}\mathbf{g} - \mathbf{d}_2\|_2^2\\ &= 0 + e_1 = e_1 \end{aligned} \tag{4.17}$$

where we have partitioned $\mathbf{x} = (\mathbf{g}, \mathbf{z})^{\mathrm{T}}$.

In fact the inequality relation between e and the terms on the right hand side of Eq. (4.17) can be sharpened to equality, for the minima of those terms are taken at the same value of $\mathbf{x}$, $(\mathbf{g}, \mathbf{z})^{\mathrm{T}}$.

In the light of this remark we note that the solution of Method 1, $(\mathbf{g}, \mathbf{z})^{\mathrm{T}}$, is the solution of

$$\min_{\mathbf{g},\mathbf{z}}\left\|Q^{\mathrm{T}}A\begin{bmatrix}\mathbf{g}\\ \mathbf{z}\end{bmatrix} - Q^{\mathrm{T}}\delta f\right\|_2^2. \tag{4.18}$$

where Q is orthogonal. The normal equations for this problem are

$$(Q^{\mathrm{T}}A)^{\mathrm{T}}(Q^{\mathrm{T}}A)\begin{bmatrix}\mathbf{g}\\ \mathbf{z}\end{bmatrix} = (Q^{\mathrm{T}}A)^{\mathrm{T}}Q^{\mathrm{T}}\delta f \tag{4.19}$$

which under the assumption of full rank for A are equivalent to

$$A^{\mathrm{T}}A\begin{bmatrix}\mathbf{g}\\ \mathbf{z}\end{bmatrix} = A^{\mathrm{T}}\delta f. \tag{4.20}$$

The solution of Method 2 is

$$\min_{\mathbf{g},\mathbf{z}} \left\| P^{\mathrm{T}}A\begin{bmatrix}\mathbf{g}\\ \mathbf{z}\end{bmatrix} - P^{\mathrm{T}}\delta f \right\|_2^2 \tag{4.21}$$

where P is orthogonal. Following the same manipulation as was done for Q we find that this, too, is equivalent to Eq. (4.20). So we conclude for the third time that the solutions of Method 1 and Method 2 are identical.

Finally, returning to the problem of gradient estimation, Method 1 does not directly involve the higher order derivatives. Thus it appears that the higher order terms may be eliminated from any subset of the equations, leading to apparently different estimates. The equivalence of Methods 1 and 2 shows that in fact the solution obtained will be identical in all cases.

5. Experimental results

All results presented in this section were performed using Matlab version 7.0.1 run on an Apple PowerBook G4.

5.1. Test problem 1

A set of numerical experiments were performed with the purpose of testing the accuracy of the estimates of the gradient and to verify their asymptotic behaviour. We chose a random set of 20 points, uniformly distributed with respect to their polar distance and angles to lie within a circle of unit radius. We then made the transformation $\mathbf{x} \to \rho\mathbf{x} + \mathbf{c}$, $\mathbf{x} = (x, y)^{\mathrm{T}}$ for various choices of $\mathbf{c} = (a, b)^{\mathrm{T}}$ and ρ. We chose a small, fixed number of points, being those points nearest the test point $\mathbf{c}$. We took, in turn, 6, 8, 10, 12 and 14 points. The weighting of the data points was progressively increased by applying scale factors of d^{α} for $\alpha = 0, -1, -2$. Finally the seed for the random number generator was given several different values. The tests were carried out on the function

$$\frac{\sin r}{r}, \qquad r^2 = x^2 + y^2, \tag{5.1}$$

which is smooth, but mildly oscillatory. The error measures we use to interpret the results given in the Tables 1 and 2 are the norms of the difference between the approximate and exact values. Thus,

$$\mathrm{ge} = \frac{\left\|(\nabla f)_{\mathrm{exact}} - (\nabla f)_{\mathrm{approx}}\right\|}{\|(\nabla f)_{\mathrm{exact}}\|} \tag{5.2}$$

and

$$\mathrm{sde} = \frac{\left\|(f_{xx}, f_{xy}, f_{yy})_{error}\right\|}{\left\|(f_{xx}, f_{xy}, f_{yy})_{\mathrm{exact}}\right\|}.$$

The results depend, therefore, on five distinct combinations of parameters. The most important of these is the dependence on ρ which quantifies the convergence rate of the approximation ; this is the focus of Table 1. It shows the accuracy of the approximation to the gradient for data within distances which are successively reduced by an order of magnitude. The quadratic dependence of the error is evidenced by the values of ge, as is the first order accuracy of the second derivative estimates, evidenced by the values of sde. The test point was (3, 4), the data points were equally weighted, and the 6 nearest points were used.

Table 1
Errors in gradient and Hessian estimates for the function (5.1), scaled relative to the magnitudes of their exact values

Radius	2.5000e−01	2.5000e−02	2.5000e−03	2.5000e−04	2.5000e−05	Exact ∇f
$\frac{\partial f}{\partial x}$	5.4717e−02	5.7029e−02	5.7053e−02	5.7054e−02	5.7054e−02	5.7054e−02
$\frac{\partial f}{\partial y}$	7.3850e−02	7.6050e−02	7.6071e−02	7.6072e−02	7.6072e−02	7.6072e−02
ge	3.3907e−02	3.3753e−04	3.3725e−06	3.3724e−08	3.6708e−10	
Radius	2.5000e−01	2.5000e−02	2.5000e−03	2.5000e−04	2.5000e−05	Exact 2nd order derivatives
$\frac{\partial^2 f}{\partial x^2}$	4.9594e−02	6.5700e−02	6.7339e−02	6.7503e−02	6.7521e−02	6.7521e−02
$\frac{\partial^2 f}{\partial x \partial y}$	5.8292e−02	6.4103e−02	6.4615e−02	6.4665e−02	6.4671e−02	6.4671e−02
$\frac{\partial^2 f}{\partial y^2}$	1.1236e−01	1.0594e−01	1.0532e−01	1.0525e−01	1.0525e−01	1.0525e−01
sde	1.4431e−01	1.4422e−02	1.4414e−03	1.4413e−04	1.6062e−05	

Table 2
Errors in gradient estimates for the function (5.1), for various field points, weightings, number of field points and random seeds

(a, b)	(3 4)*1e−2	(3 4)*1e−1	(3 4)	(3 4)*1e1	(3 4)*1e2
ge	9.0412e−07	8.9218e−07	3.3725e−06	1.9317e−06	1.9043e−06
scale index	0	−1	−2	−3	−4
ge	3.3725e−06	3.3706e−06	3.3689e−06	3.3674e−06	3.3661e−06
number of field pts	6	8	10	12	14
ge	3.3725e−06	4.5109e−06	1.1857e−06	1.1906e−06	1.8288e−06
seed number	1	2	3	4	5
ge	3.8766e−07	2.1210e−06	2.6842e−06	2.3917e−06	1.0448e−06

Table 2 gives an indication of how the errors vary with the remaining parameter combinations. These results show that there is little change in the error with these parameters, certainly within an order of magnitude over the variation of any single parameter. The values of the errors in the second derivative exhibit similar insensitivity to these parameters and are not given here.

Finally we permuted the choice of rows for elimination and observed the machine accurate identical results for all permutations which the observations and theorems of the earlier sections predicted.

5.2. Test problem 2

A similar set of numerical experiments were performed on the solution ϕ of the following classical steady-state heat diffusion equation

$$\nabla \cdot (\mathbf{D}\nabla \mathbf{u}) = -g_0 \tag{5.3}$$

on the square domain $0 \le x \le 1, 0 \le y \le 1$. For the first sets of numerical experiments, a constant diffusion tensor $D = \mathrm{diag}(D_{xx}, D_{yy})$ is used, together with a constant source g_0. The boundaries $x = 0$ and $y = 0$ are taken to be insulated, and the boundaries $x = 1$ and $y = 1$ are subjected to Newtonian cooling with external temperature φ_∞. This problem is solvable analytically, and its solution is given by, [11],

$$\varphi(x, y) = \sum_{n=1}^{\infty} \frac{2({\lambda_n}^2 + {H_x}^2) F_n \cos(\lambda_n x) \cosh(\lambda_n \eta y)}{({\lambda_n}^2 + H_x + {H_x}^2)(\lambda_n \eta \sinh(\lambda_n \eta) + H_y \cosh(\lambda_n \eta))} + g_0 \left[\frac{1 - x^2}{2D_{xx}} + \frac{1}{h} \right] + \varphi_\infty \tag{5.4}$$

where

$$\eta^2 = \frac{D_{xx}}{D_{yy}},$$

$$H_x = \frac{h}{D_{xx}},$$

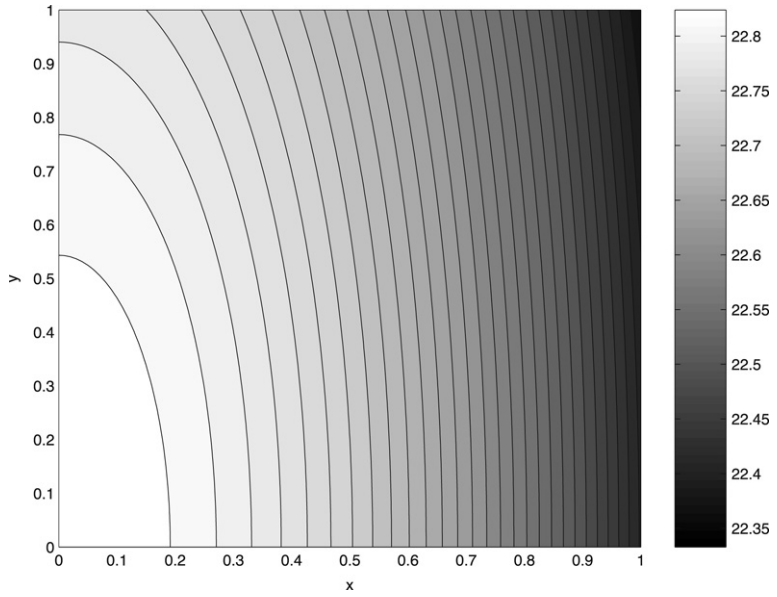

Fig. 2.

Table 3
Errors in gradient estimates for the function (5.4), scaled relative to the magnitudes of their exact values

Radius	2.5000e−01	2.5000e−02	2.5000e−03	2.5000e−04	2.5000e−05	Exact ∇f
$\frac{\partial f}{\partial x}$	−2.6947e−01	−2.7125e−01	−2.7127e−01	−2.7127e−01	−2.7127e−01	−2.7127e−01
$\frac{\partial f}{\partial y}$	−8.6722e−02	−8.6513e−02	−8.6512e−02	−8.6512e−02	−8.6512e−02	−8.6512e−02
ge	6.3577e−03	6.7051e−05	6.7436e−07	6.7395e−09	4.1078e−10	

$$F_n = \int_0^1 \frac{g_0}{D_{yy}} \left[\frac{h(x^2-1)}{2D_{xx}} - 1 \right] \cos(\lambda_n x) \mathrm{d}x$$

and the λ_n are the solutions to

$$\frac{H_x}{\lambda} = \tan(\lambda).$$

The complete set of parameter values used for this test problem are $\varphi_\infty = 20$, $D = \mathrm{diag}(5, 50)$, $g_0 = 10$ and $h = 2$ with the first 30 terms of the series for φ used. The exact solution of (5.3) is shown in Fig. 2.

The test data was generated almost exactly as described for problem 1, the only difference being the test points at which the gradient was estimated. They were chosen so that they lay within the unit square $[0, 1] \times [0, 1]$. The first test point chosen was $(.3, .4)$ and this was varied with four perturbations $(\pm.2, \pm.1)$. As before there were five combinations of parameters used in the numerical experiments, these followed the same pattern as for problem 1.

The results shown in Table 3 again affirm the quadratic dependence on the test region radius as exhibited for test problem 1. The results in Table 4 record the insensitivity of the gradient error to the remaining parameter values. In problem 2 we did not test the second derivative estimates.

Figs. 3–5 are log–log plots of the data for the variation of the error with the radius of the test region. These are plots of the gradient and second derivative errors for problem 1 and for the gradient for problem 2. In all three cases the points are, almost precisely, collinear. The slopes of the lines, are 2, 1 and 2, confirming the quadratic dependence of the gradient error and the linear dependence for the second derivative error.

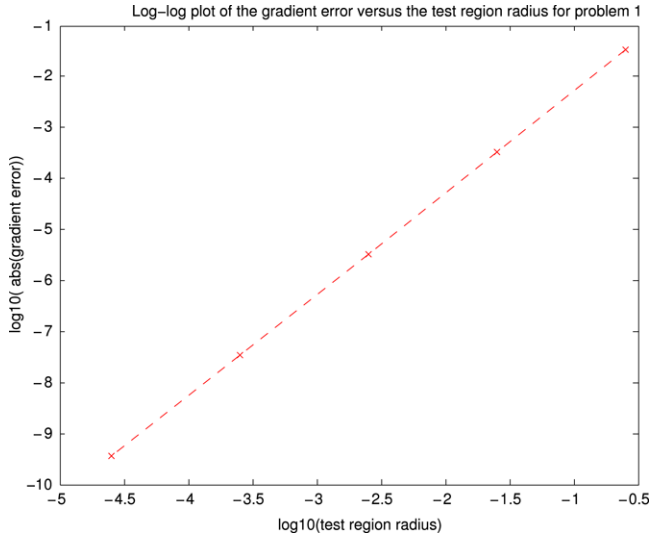

Fig. 3.

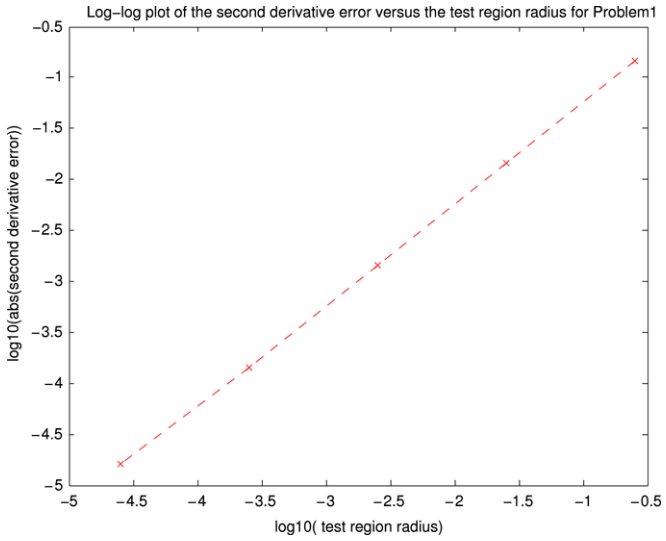

Fig. 4.

Table 4
Errors in gradient estimates for the function (5.4), for various field points, weightings, number of field points and random seeds

(a, b)	(.3.4)	(.1.3)	(.2.6)	(.5.5)	(.4.2)
ge	6.7436e−07	9.5189e−07	9.5785e−07	6.0470e−07	4.7791e−07
scale index	0	−1	−2	−3	−4
ge	6.7436e−07	6.7531e−07	6.7605e−07	6.7663e−07	6.7709e−07
number of field pts	6	8	10	12	14
ge	3.3725e−06	4.5109e−06	1.1857e−06	1.1906e−06	1.8288e−06
seed number	1	2	3	4	5
ge	8.6164e−08	2.7465e−08	7.7532e−07	2.8179e−07	3.6752e−07

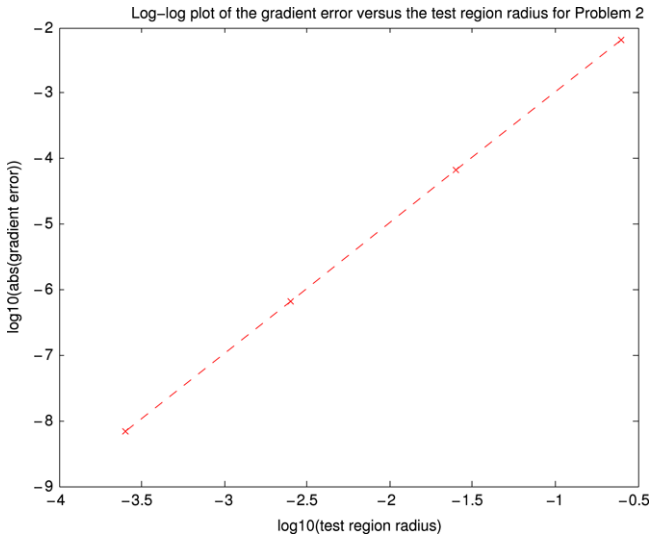

Fig. 5.

Table 5
Errors in gradient estimates for the function (5.5), scaled relative to the magnitudes of their exact values with ge as defined in Eq. (5.2)

Radius	2.5000e−01	2.5000e−02	2.5000e−03	2.5000e−04	2.5000e−05	Exact ∇f
$\frac{\partial f}{\partial x}$	−1.9783e−03	−2.1224e−03	−2.1309e−03	−2.1310e−03	−2.1310e−03	−2.1310e−03
$\frac{\partial f}{\partial y}$	−3.9333e−02	−3.8148e−02	−3.8141e−02	−3.8140e−02	−3.8140e−02	−3.8140e−02
ge	3.1483e−02	3.0482e−04	3.1120e−06	3.1193e−08	3.1755e−10	

5.3. A comparison and some remarks on complexity

Ling, [9], applies his algorithm to the function

$$\sin(\pi x)\sin(\pi y)\exp(-x^2 - y^2) \tag{5.5}$$

on the square $[-2, 2] \times [-2, 2]$. For 1609 points with minimum separation distance not less than 5.092e−02, the root mean square errors based on a 100×100 grid lie between 1.5e−03 for zero noise and 3.84e−1 for 10% noise.

Performing the same tests as those described in Section 5.1 on this function we obtained the results shown in Table 5. The test point was at (1.3, 1.7); the results are typical, taken from a grid of points with spacing of 0.1 in both coordinate directions. The second order dependence of the error may be observed as in the earlier examples. In comparison with Ling's results [9] those given for a radius of .025 are perhaps the most relevant. The errors for Ling's noise free results [9] are of a similar order of magnitude to those given in Table 5. Results for differing spacing over a range of magnitudes and statements regarding complexity are not given by Ling [9].

The complexity of an algorithm using function values and gradients at the data points is of the same order as one using function values only. To determine a gradient estimate at a particular node, a set of data points is located. Either of the least squares problems described in Section 4, (Method 1 or Method 2) is set up and solved. The coefficient matrix will be $p \times 5$, where p is the number of points nearest the node. Thus the cost at each node is a fixed number, the maximum number of flops needed for the least squares solution of the $p \times 5$ system (Numerical evidence, as presented below, suggests that the six closest points give sufficient accuracy.) The interpolant is evaluated using a Lagrange formula based on the cardinal functions (cubic polynomials) given in Reference [6]. Thus if there are n data points the computation needed to compute the interpolant will be, at most, a fixed multiple of that needed to compute a counterpart based on function values only. Since the complexity of a Delaunay triangulation is $O(n \log n)$, the method, *to an order of magnitude* has the same complexity as a piecewise linear approximation.

6. Conclusions

We have introduced two methods for estimating gradients for use in scattered data surface fitting. Although these methods appeared superficially different, we have proven that they produce the same least squares errors and derivative estimates. Furthermore, the results indicate that the methods are robust in the context of the location of the gradient estimation point $\mathbf{x}_0$, the number of points used for generating the least squares system and the weightings based on the inverse distance of these points from $\mathbf{x}_0$. The results are consistent over a range of randomly generated data sets. The chosen applications highlight the importance of the technique.

The analysis which proves the uniqueness of the gradient estimate is applicable for any full column rank overdetermined linear least squares problem. If offers the opportunity of solving a least squares by partitioning or divide and conquer methods in the following way. First factorise a subset of the columns, reducing the dimension of the least squares problem. Follow this by solving the remaining triangular system. Clearly the process can be continued recursively and any structure such as sparsity within a subset of the original problem could be handled in a similar fashion.

Acknowledgment

The authors wish to thank S.J. Wright, University of Wisconsin, for his comments on an earlier draft of the manuscript and his observations which were the basis for the argument presented in Eqs. (4.17)–(4.21).

References

[1] M. Breslin, Spatial interpolation and fractal analysis applied to rainfall data, Ph.D. Thesis, Department of Mathematics, University of Queensland, 2001.
[2] C. Godin, Representing and encoding plant architecture: A review, Ann. For. Sci. 57 (2000) 413–438.
[3] P.A. Jayantha, I.W. Turner, A comparison of gradient approximations for use in finite volume computational models for two dimensional differential equations, Numer. Heat Transfer Part B: Fundamentals 40 (2001) 367–390.
[4] P.A. Jayantha, I.W. Turner, A second order finite volume technique for simulating transport in anisotropic media, The Int. J. Numer. Methods Heat Fluid Flow 13 (2003) 31–56.
[5] P. Lancaster, K. Salkauskas, Surfaces generated by moving least squares methods fitting, Math. Comp. 37 (1981) 141–158.
[6] P. Lancaster, K. Salkauskas, Curve and Surface Fitting, An Introduction, Academic Press, San Diego, 1986.
[7] David Levin, The power of moving least-squares, Math. Comp. 67 (1998) 1517–1531.
[8] David Levin, in: Brunnett, Hamann, Mueller (Eds.), Geometric Modeling for Scientific Visualization, Springer-Verlag, 2003, pp. 37–49.
[9] Leevan Ling, Finding numerical derivatives for unstructured and noisy data by multiscale kernels, SIAM J. Numer. Anal. 44 (2006) 1780–1800.
[10] B.I. Loch, Surface fitting for the modelling of plant leaves, Ph.D. Thesis, Department of Mathematics, University of Queensland, July, 2004.
[11] M.N. Özişik, Heat Conduction, Wiley, New York, 1980.
[12] M.J.D. Powell, A review of methods for multivariable interpolation at scattered data points, in: I.S. Duff, et al. (Eds.), The State of the Art in Numerical Analysis, Clarendon Press, Oxford, 1997, pp. 283–310.
[13] P. Prusinkiewicz, L. Mundermann, R. Karwowski, B. Lane, The use of positional information in the modeling of plants, in: Proceedings of SIGGRAPH, 2001.
[14] S. Truscott, I.W. Turner, An investigation of the accuracy of the control volume finite element method based on triangular prismatic finite elements for simulating diffusion in anisotropic media, Numer. Heat Transfer Part B: Fundamentals 46 (2004) 243–268.