Applied Mathematical Sciences, Vol. 2, 2008, no. 28, 1365 - 1376

The Best Gradient Approximation to an Observed Function

Phil Howlett

Centre for Industrial and Applied Mathematics (CIAM)
School of Mathematics and Statistics
University of South Australia, Australia
Phil.Howlett@unisa.edu.au

Abstract

We wish to find the potential function whose gradient best approximates an observed square integrable function. With an appropriate choice of topology we show that the gradient operator is a bounded linear operator and that the best approximation is obtained by solving a self-adjoint partial differential equation.

Mathematics Subject Classification: 41A65, 47A05, 47A50

Keywords: gradient operator, best approximation, Sobolev spaces

1 Introduction

Write $\boldsymbol{r}$ to denote the general position vector in $\mathbb{R}^3$. Let $\mathcal{U} = [0,1]^3$ denote the unit cube in $\mathbb{R}^3$ with boundary $\partial\mathcal{U}$. Suppose an observed square integrable function $\boldsymbol{g} : \mathcal{U} \mapsto \mathbb{R}^3$ is given. We wish to find the potential function $u : \mathcal{U} \mapsto \mathbb{R}$ with $u(\boldsymbol{r}) = 0$ when $\boldsymbol{r} \in \partial\mathcal{U}$, which minimises the total residual error

$$\iiint_{\mathcal{U}} \|\boldsymbol{\nabla} u(\boldsymbol{r}) - \boldsymbol{g}(\boldsymbol{r})\|^2 \, dV.$$

2 Preliminary Notes

We need to establish some basic facts about the gradient operator. There are no specific references but a general background can be found in the books by Aubin [1], Treves [2] and Yosida [3]. Define the Hilbert space H of functions $u : \mathcal{U} \mapsto \mathbb{R}$ such that

$$\iiint_{\mathcal{U}} \left[\, |u(\boldsymbol{r})|^2 + \|\boldsymbol{\nabla} u(\boldsymbol{r})\|^2 \, \right] \, dV < \infty$$

with $u(\boldsymbol{r}) = 0$ when $\boldsymbol{r} \in \partial\mathcal{U}$ and inner product

$$\langle u, v\rangle_H = \iiint_{\mathcal{U}} \Big[\ u(\boldsymbol{r})v(\boldsymbol{r}) + \langle \boldsymbol{\nabla} u(\boldsymbol{r}), \boldsymbol{\nabla} v(\boldsymbol{r})\rangle\ \Big]\, dV$$

for each $u, v \in H$ and the Hilbert space K of functions $\boldsymbol{g} : \mathcal{U} \mapsto \mathbb{R}^3$ such that

$$\iiint_{\mathcal{U}} \|\boldsymbol{g}(\boldsymbol{r})\|^2\ dV < \infty$$

with inner product

$$\langle \boldsymbol{g}, \boldsymbol{h}\rangle_K = \iiint_{\mathcal{U}} \langle \boldsymbol{g}(\boldsymbol{r}), \boldsymbol{h}(\boldsymbol{r})\rangle\ dV$$

for each $\boldsymbol{g}, \boldsymbol{h} \in K$. The mapping $\boldsymbol{\nabla} : H \mapsto K$ defined by

$$\boldsymbol{\nabla} u = \begin{bmatrix} \dfrac{\partial u}{\partial x} \\[2ex] \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial u}{\partial z} \end{bmatrix}$$

is a bounded linear map. We call $\boldsymbol{\nabla}$ the gradient operator. Let $\mathcal{D}$ denote the set of all infinitely differentiable functions $\varphi : \mathcal{U} \mapsto \mathbb{R}$ with compact support $\mathrm{spt}(\varphi) \subset \mathrm{int}(\mathcal{U})$ and for each $m = 2, 3, \ldots$ let $\mathcal{D}^m$ be the set of infinitely differentiable functions $\boldsymbol{\psi} : \mathcal{U} \mapsto \mathbb{R}^m$ with compact support $\mathrm{spt}(\boldsymbol{\psi}) \subset \mathrm{int}(\mathcal{U})$. The adjoint map $\boldsymbol{\nabla}^* : K \mapsto H$ is given by $\boldsymbol{\nabla}^*\boldsymbol{k} = w$ where $w \in H$ is the uniquely defined function with

$$\langle w, \varphi\rangle_H = \langle \boldsymbol{k}, \boldsymbol{\nabla}\varphi\rangle_K \tag{1}$$

for all $\varphi \in \mathcal{D}$. To find a suitable formula for the adjoint operator we digress for a moment to consider the divergence operator. We have the following elementary result.

Lemma 1 *Let $\boldsymbol{\psi} \in \mathcal{D}^3$. For all $\varphi \in \mathcal{D}$ we have*

$$\iiint_{\mathcal{U}} \langle \boldsymbol{\nabla}, \boldsymbol{\psi}(\boldsymbol{r})\rangle\ \varphi(\boldsymbol{r})\ dV = (-1) \iiint_{\mathcal{U}} \langle \boldsymbol{\psi}(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r})\rangle\ dV.$$

Proof of Lemma 1 It can be seen that

$$\langle \boldsymbol{\nabla}, \boldsymbol{\psi}\varphi \rangle = \langle \boldsymbol{\nabla}, \boldsymbol{\psi} \rangle \, \varphi + \langle \boldsymbol{\psi}, \boldsymbol{\nabla}\varphi \rangle$$

and so

$$\iiint_{\mathcal{U}} \langle \boldsymbol{\nabla}, \boldsymbol{\psi}(\boldsymbol{r})\varphi(\boldsymbol{r}) \rangle dV = \iiint_{\mathcal{U}} \Big[\langle \boldsymbol{\nabla}, \boldsymbol{\psi}(\boldsymbol{r}) \rangle \, \varphi(\boldsymbol{r}) + \langle \boldsymbol{\psi}(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r}) \rangle \Big] \, dV.$$

If $\boldsymbol{n}(\boldsymbol{r})$ denotes the unit outward normal to the surface $\partial\mathcal{U}$ at the point $\boldsymbol{r}$ then by Gauss's theorem

$$\iiint_{\mathcal{U}} \langle \boldsymbol{\nabla}, \boldsymbol{\psi}(\boldsymbol{r})\varphi(\boldsymbol{r}) \rangle \, dV = \iint_{\partial\mathcal{U}} \langle \boldsymbol{\psi}(\boldsymbol{r})\varphi(\boldsymbol{r}), \boldsymbol{n} \rangle \, dS = 0$$

since $\varphi(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \in \partial\mathcal{U}$. This establishes the desired result. □

This basic result is used to define the generalised divergence.

Definition 1 *For each $\boldsymbol{k} \in K$ the generalised divergence operator $\langle \boldsymbol{\nabla}, \boldsymbol{k} \rangle : \mathcal{D} \mapsto \mathbb{R}$ is defined by the formula*

$$\langle \boldsymbol{\nabla}, \boldsymbol{k} \rangle(\varphi) = (-1) \iiint_{\mathcal{U}} \langle \boldsymbol{k}(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r}) \rangle \, dV.$$

Remark 1 *For $\boldsymbol{\psi} \in \mathcal{D}^3$ Lemma 1 shows that*

$$\langle \boldsymbol{\nabla}, \boldsymbol{\psi} \rangle(\varphi) = \iiint_{\mathcal{U}} \langle \boldsymbol{\nabla}, \boldsymbol{\psi}(\boldsymbol{r}) \rangle \, \varphi(\boldsymbol{r}) \, dV$$

for all $\varphi \in \mathcal{D}$. Note also that if $\boldsymbol{k} \in K$ and $\{\boldsymbol{\psi}_n\}_{n\in\mathbb{N}} \subset \mathcal{D}^3$ with $\|\boldsymbol{\psi}_n - \boldsymbol{k}\|_K \to 0$ as $n \to \infty$ then

$$\langle \boldsymbol{\nabla}, \boldsymbol{\psi}_n \rangle(\varphi) \to \langle \boldsymbol{\nabla}, \boldsymbol{k} \rangle(\varphi)$$

for all $\varphi \in \mathcal{D}$ as $n \to \infty$.

Rewrite equation (1) as

$$\iiint_{\mathcal{U}} \Big[\; w(\boldsymbol{r})\varphi(\boldsymbol{r}) + \langle \boldsymbol{\nabla} w(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r}) \rangle \; \Big] \, dV = \iiint_{\mathcal{U}} \langle \boldsymbol{k}(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r}) \rangle \, dV$$

from which it follows that

$$\iiint_{\mathcal{U}} \Big[\; \langle \boldsymbol{k}(\boldsymbol{r}) - \boldsymbol{\nabla} w(\boldsymbol{r}), \boldsymbol{\nabla}\varphi(\boldsymbol{r}) \rangle \; - w(\boldsymbol{r})\varphi(\boldsymbol{r}) \; \Big] \, dV = 0. \tag{2}$$

If $w \in H$ then it is well known that Poisson's equation

$$\nabla^2 \zeta = w \quad \Leftrightarrow \quad \langle \boldsymbol{\nabla}, \boldsymbol{\nabla}\zeta \rangle = w$$

with $\zeta(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \in \partial\mathcal{U}$ has a unique classical solution

$$\zeta(\boldsymbol{r}) = -\frac{1}{4\pi}\iiint_{\mathcal{U}} \frac{w(\boldsymbol{s})}{\|\boldsymbol{s}-\boldsymbol{r}\|} dV.$$

If $\boldsymbol{\xi} = \boldsymbol{\nabla}\zeta$ then $w = \langle \boldsymbol{\nabla}, \boldsymbol{\xi}\rangle$ and equation (2) becomes

$$\iiint_{\mathcal{U}} \Big[\ \langle \boldsymbol{k}(\boldsymbol{r}) - \boldsymbol{\nabla}\langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\rangle, \boldsymbol{\nabla}\varphi(\boldsymbol{r})\rangle \ - \langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\rangle \ \varphi(\boldsymbol{r}) \ \Big] dV = 0.$$

Now use the standard identity

$$\langle \boldsymbol{\nabla}, \boldsymbol{\xi}\,\varphi\rangle = \langle \boldsymbol{\nabla}, \boldsymbol{\xi}\rangle\varphi + \langle \boldsymbol{\xi}, \boldsymbol{\nabla}\varphi\rangle$$

to deduce that

$$\iiint_{\mathcal{U}} \left[\langle \ [\boldsymbol{k}(\boldsymbol{r}) + \boldsymbol{\xi}(\boldsymbol{r}) - \boldsymbol{\nabla}\langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\rangle] \ , \boldsymbol{\nabla}\varphi(\boldsymbol{r})\rangle - \langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\,\varphi(\boldsymbol{r})\rangle \right] dV = 0.$$

By Gauss's theorem

$$\iiint_{\mathcal{U}} \langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\,\varphi(\boldsymbol{r})\rangle dV \ = \ \iint_{\partial\mathcal{U}} \langle \boldsymbol{\xi}(\boldsymbol{r})\,\varphi(\boldsymbol{r}), \boldsymbol{n}(\boldsymbol{r})\rangle \ dS \ = \ 0$$

since $\varphi(\boldsymbol{r}) = 0$ when $\boldsymbol{r} \in \partial\mathcal{U}$. Thus the above equation can be rewritten as

$$\iiint_{\mathcal{U}} \langle \ [\boldsymbol{k}(\boldsymbol{r}) + \boldsymbol{\xi}(\boldsymbol{r}) - \boldsymbol{\nabla}\langle \boldsymbol{\nabla}, \boldsymbol{\xi}(\boldsymbol{r})\rangle] \ , \boldsymbol{\nabla}\varphi(\boldsymbol{r})\rangle \ dV = 0 \tag{3}$$

for all $\varphi \in \mathcal{D}$. From Definition 1 of the generalised divergence it follows that

$$\langle \boldsymbol{\nabla}, \ [\boldsymbol{k} + \boldsymbol{\xi} - \boldsymbol{\nabla}\langle \boldsymbol{\nabla}, \boldsymbol{\xi}\rangle] \ \rangle = 0$$

and since $w = \langle \boldsymbol{\nabla}, \boldsymbol{\xi}\rangle \in H$ we deduce that

$$w + \langle \boldsymbol{\nabla}, \boldsymbol{k} - \boldsymbol{\nabla} w\rangle = 0 \quad \Leftrightarrow \quad w = (-1)(I - \nabla^2)^{-1}\langle \boldsymbol{\nabla}, \boldsymbol{k}\rangle.$$

Remark 2 *If* $\boldsymbol{k} \in K$ *then* $w = \boldsymbol{\nabla}^*\boldsymbol{k} \in H$ *is defined in terms of the generalised divergence by the formula*

$$\boldsymbol{\nabla}^*\boldsymbol{k} = (-1)(I - \nabla^2)^{-1}\langle \boldsymbol{\nabla}, \boldsymbol{k}\rangle. \tag{4}$$

3 The main result

We state the main result as a formal theorem. Detailed arguments are presented in subsequent sections.

Theorem 1 *Let $\boldsymbol{g} \in K$ be an observed function and let $\boldsymbol{g}_0 \in K$ be the modified observed function defined by*

$$\boldsymbol{g}_0 = \boldsymbol{g} - \begin{bmatrix} \int_0^1 g_1(x,y,z)dx \\ \int_0^1 g_2(x,y,z)dy \\ \int_0^1 g_3(x,y,z)dz \end{bmatrix}.$$

The problem

$$\min_{u \in H} \|\boldsymbol{\nabla} u - \boldsymbol{g}\|_K$$

is solved by the function $u_0 \in H$ defined as the unique solution to the equation

$$\boldsymbol{\nabla}^* \boldsymbol{\nabla} u_0 = \boldsymbol{\nabla}^* \boldsymbol{g} \quad \Leftrightarrow \quad (-1)(I - \nabla^2)^{-1} \nabla^2 u_0 = (-1)(I - \nabla^2)^{-1} \langle \boldsymbol{\nabla}, \boldsymbol{g} \rangle.$$

If the observed function is written in the form

$$\boldsymbol{g} = \begin{bmatrix} \sum_{m=1}^\infty \sum_{n=1}^\infty [\ \delta_{0,m,n} + \sum_{\ell=1}^\infty \delta_{\ell,m,n} \cos \ell\pi x\ ] \sin m\pi y \cdot \sin n\pi z \\ \sum_{\ell=1}^\infty \sum_{n=1}^\infty [\ \epsilon_{\ell,0,n} + \sum_{m=1}^\infty \epsilon_{\ell,m,n} \cos m\pi y\ ] \sin \ell\pi x \cdot \sin n\pi z \\ \sum_{\ell=1}^\infty \sum_{m=1}^\infty [\ \theta_{\ell,m,0} + \sum_{n=1}^\infty \theta_{\ell,m,n} \cos n\pi z\ ] \sin \ell\pi x \cdot \sin m\pi y \end{bmatrix}.$$

then the solution is given by

$$u_0 = \sum_{\ell=1}^\infty \sum_{m=1}^\infty \sum_{n=1}^\infty \gamma_{\ell,m,n}\ \sin \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z$$

where

$$\gamma_{\ell,m,n} = \frac{\ell\delta_{\ell,m,n} + m\epsilon_{\ell,m,n} + n\theta_{\ell,m,n}}{\pi(\ell^2 + m^2 + n^2)}$$

for all $\ell, m, n \in \mathbb{N}$. The function u_0 also solves the modified problem

$$\min_{u \in H} \|\boldsymbol{\nabla} u - \boldsymbol{g}_0\|_K$$

and if the coefficients in the series expansions for $\boldsymbol{g}$ and $\boldsymbol{g}_0$ satisfy the conditions

$$n\epsilon_{\ell,m,n} = m\theta_{\ell,m,n}, \quad n\delta_{\ell,m,n} = \ell\theta_{\ell,m,n} \quad \text{and} \quad m\delta_{\ell,m,n} = \ell\epsilon_{\ell,m,n}$$

for all $\ell, m, n \in \mathbb{N}$ then the best approximate gradient u_0 becomes an exact gradient for the modified observed function. That is $\boldsymbol{\nabla} u_0 = \boldsymbol{g}_0$.

4 The best approximate gradient

Suppose an observed function $\boldsymbol{g} \in K$ is given and we wish to find $u \in H$ such that

$$\|\boldsymbol{\nabla} u - \boldsymbol{g}\|_K$$

is minimised. Since $\boldsymbol{\nabla} \in \mathcal{L}(H, K)$ is a bounded linear operator it follows that the best approximation $u_0 \in H$ is given by

$$\boldsymbol{\nabla}^*(\boldsymbol{\nabla} u_0 - \boldsymbol{g}) = 0 \quad \Leftrightarrow \quad (-1)(I - \nabla^2)^{-1}\langle \boldsymbol{\nabla}, \boldsymbol{\nabla} u_0 - \boldsymbol{g}\rangle = 0.$$

Of course it is often convenient to think of this equation in the form

$$\boldsymbol{\nabla}^*\boldsymbol{\nabla} u_0 = \boldsymbol{\nabla}^*\boldsymbol{g} \quad \Leftrightarrow \quad (-1)(I - \nabla^2)^{-1}\nabla^2 u_0 = (-1)(I - \nabla^2)^{-1}\langle \boldsymbol{\nabla}, \boldsymbol{g}\rangle. \quad (5)$$

This equation is a standard self-adjoint differential equation[1]. The orthonormal eigenfunctions in H for the self-adjoint operator $\boldsymbol{\nabla}^*\boldsymbol{\nabla} : H \mapsto H$ are given by

$$p_{\ell,m,n} = \frac{2\sqrt{2}}{\sqrt{1 + \pi^2(\ell^2 + m^2 + n^2)}} \, \sin \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z$$

and hence

$$\boldsymbol{\nabla}^*\boldsymbol{\nabla} p_{\ell,m,n} = \frac{\pi^2(\ell^2 + m^2 + n^2)}{1 + \pi^2(\ell^2 + m^2 + n^2)} \, p_{\ell,m,n}$$

for each $\ell, m, n \in \mathbb{N}$. The standard mathematical procedure for solution of equation (5) uses eigenfunction expansions in both H and K but we reject it for the moment to pursue a more intuitive approach.

4.1 An intuitive solution procedure

It is easy to see that

$$\boldsymbol{\nabla} p_{\ell,m,n} = \frac{2\sqrt{2}\pi}{\sqrt{1 + \pi^2(\ell^2 + m^2 + n^2)}} \begin{bmatrix} \ell \cos \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z \\ m \sin \ell\pi x \cdot \cos m\pi y \cdot \sin n\pi z \\ n \sin \ell\pi x \cdot \sin m\pi y \cdot \cos n\pi z \end{bmatrix}$$

and by looking at the individual components of these vectors it seems reasonable to use them as a basis for our representation of the function $\boldsymbol{g} \in K$. Thus we write

$$\boldsymbol{g} = \begin{bmatrix} \sum_{m=1}^\infty \sum_{n=1}^\infty [\ \delta_{0,m,n} + \sum_{\ell=1}^\infty \delta_{\ell,m,n} \cos \ell\pi x\ ] \sin m\pi y \cdot \sin n\pi z \\ \\ \sum_{\ell=1}^\infty \sum_{n=1}^\infty [\ \epsilon_{\ell,0,n} + \sum_{m=1}^\infty \epsilon_{\ell,m,n} \cos m\pi y\ ] \sin \ell\pi x \cdot \sin n\pi z \\ \\ \sum_{\ell=1}^\infty \sum_{m=1}^\infty [\ \theta_{\ell,m,0} + \sum_{n=1}^\infty \theta_{\ell,m,n} \cos n\pi z\ ] \sin \ell\pi x \cdot \sin m\pi y \end{bmatrix}.$$

[1] One may be tempted to rewrite equation (5) in the *intuitively* simpler form $\nabla^2 u_0 = \langle \boldsymbol{\nabla}, \boldsymbol{g}\rangle$ but in doing so one should realise that the *simplified* equation is an operator equation and is not an equation between functions. This is most easily seen by considering the possibly divergent eigenfunction expansions.

A little thought should convince us that this representation is complete in K. The coefficients of the series can be calculated using the usual Fourier integral formulae. Some elementary algebra now gives

$$\boldsymbol{\nabla}^* \boldsymbol{g} = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{\pi(\ell \delta_{\ell,m,n} + m \epsilon_{\ell,m,n} + n \theta_{\ell,m,n})}{1 + \pi^2(\ell^2 + m^2 + n^2)} \ \sin \ell \pi x \cdot \sin m \pi y \cdot \sin n \pi z. \tag{6}$$

We seek a solution to equation (5) in the form

$$u_0 = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \gamma_{\ell,m,n} \ \sin \ell \pi x \cdot \sin m \pi y \cdot \sin n \pi z$$

from which we calculate

$$\boldsymbol{\nabla} u_0 = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \pi \gamma_{\ell,m,n} \begin{bmatrix} \ell \sin \ell \pi x \cdot \sin m \pi y \cdot \sin n \pi z \\ m \sin \ell \pi x \cdot \cos m \pi y \cdot \sin n \pi z \\ n \sin \ell \pi x \cdot \sin m \pi y \cdot \cos n \pi z \end{bmatrix}$$

and

$$\boldsymbol{\nabla}^* \boldsymbol{\nabla} u_0 = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{\pi^2(\ell^2 + m^2 + n^2) \gamma_{\ell,m,n}}{1 + \pi^2(\ell^2 + m^2 + n^2)} \ \sin \ell \pi x \cdot \sin m \pi y \cdot \sin n \pi z. \tag{7}$$

By equating coefficients in expressions (6) and (7) we obtain the solution

$$\gamma_{\ell,m,n} = \frac{\ell \delta_{\ell,m,n} + m \epsilon_{\ell,m,n} + n \theta_{\ell,m,n}}{\pi(\ell^2 + m^2 + n^2)} \tag{8}$$

for all $\ell, m, n \in \mathbb{N}$.

4.2 A modified problem with reduced error

The error $\boldsymbol{\varepsilon} = \boldsymbol{\nabla} u_0 - \boldsymbol{g}$ in the solution is given by

$$\boldsymbol{\varepsilon} = \begin{bmatrix} -\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [\delta_{0,m,n} + \sum_{\ell=1}^{\infty} (\delta_{\ell,m,n} - \pi \ell \gamma_{\ell,m,n}) \cos \ell \pi x] \sin m \pi y \cdot \sin n \pi z \\ -\sum_{\ell=1}^{\infty} \sum_{n=1}^{\infty} [\epsilon_{\ell,0,n} + \sum_{m=1}^{\infty} (\epsilon_{\ell,m,n} - \pi m \gamma_{\ell,m,n}) \cos m \pi y] \sin \ell \pi x \cdot \sin n \pi z \\ -\sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} [\theta_{\ell,m,0} + \sum_{n=1}^{\infty} (\theta_{\ell,m,n} - \pi n \gamma_{\ell,m,n}) \cos n \pi z] \sin \ell \pi x \cdot \sin m \pi y \end{bmatrix}$$

but this is not as bad as it might seem. On the one hand

$$\int_0^1 \frac{\partial u_0}{\partial x}(x, y, z) dx \ = \ [u_0(1, y, z) - u_0(0, y, z) \ = \ 0$$

$$\begin{aligned}
\int_0^1 \frac{\partial u_0}{\partial y}(x,y,z)dy &= [u_0(x,1,z) - u_0(x,0,z)] \;=\; 0 \\
\int_0^1 \frac{\partial u_0}{\partial z}(x,y,z)dz &= [u_0(x,y,1) - u_0(x,y,0)] \;=\; 0
\end{aligned}$$

since $u_0(\boldsymbol{r}) = 0$ when $\boldsymbol{r} \in \partial\mathcal{U}$ and on the other hand

$$\begin{aligned}
\int_0^1 g_1(x,y,z)dx &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \delta_{0,m,n} \sin m\pi y \cdot \sin n\pi z \\
\int_0^1 g_2(x,y,z)dy &= \sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} \epsilon_{\ell,0,n} \sin \ell\pi x \cdot \sin n\pi z \\
\int_0^1 g_3(x,y,z)dz &= \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty} \theta_{\ell,m,0} \sin \ell\pi x \cdot \sin m\pi y
\end{aligned}$$

for all $(y,z),(x,z),(x,y) \in [0,1]^2$. If the observed function $\boldsymbol{g} \in K$ is replaced by a modified observed function

$$\boldsymbol{g}_0 = \boldsymbol{g} - \begin{bmatrix} \int_0^1 g_1(x,y,z)dx \\ \\ \int_0^1 g_2(x,y,z)dy \\ \\ \int_0^1 g_3(x,y,z)dz \end{bmatrix}$$

then $\boldsymbol{g}_0 \in K$ and

$$\boldsymbol{g}_0 = \begin{bmatrix} \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \delta_{\ell,m,n} \cos \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z \\ \\ \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \epsilon_{\ell,m,n} \sin \ell\pi x \cdot \cos m\pi y \cdot \sin n\pi z \\ \\ \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \theta_{\ell,m,n} \sin \ell\pi x \cdot \sin m\pi y \cdot \cos n\pi z \end{bmatrix}.$$

The modified problem

$$\min_{u \in H} \|\boldsymbol{\nabla} u - \boldsymbol{g}_0\|_K$$

has the same solution $u_0 \in H$ but the error $\boldsymbol{\varepsilon}_0 = \boldsymbol{\nabla} u_0 - \boldsymbol{g}_0$ is reduced to

$$\boldsymbol{\varepsilon}_0 = \begin{bmatrix} -\sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} (\delta_{\ell,m,n} - \pi\ell\gamma_{\ell,m,n}) \cos \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z \\ \\ -\sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} (\epsilon_{\ell,m,n} - \pi m\gamma_{\ell,m,n}) \sin \ell\pi x \cdot \cos m\pi y \cdot \sin n\pi z \\ \\ -\sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} (\theta_{\ell,m,n} - \pi n\gamma_{\ell,m,n}) \sin \ell\pi x \cdot \sin m\pi y \cdot \cos n\pi z \end{bmatrix}.$$

4.3 Conditions for an exact gradient

Remark 3 *It is well known that for* $\boldsymbol{\psi} \in \mathcal{D}^3$ *the condition* $\boldsymbol{\nabla} \times \boldsymbol{\psi} = 0$ *is sufficient for the existence of a function* $\varphi \in \mathcal{D}$ *with* $\boldsymbol{\nabla}\varphi = \boldsymbol{\psi}$.

In terms of the series expansion the corresponding condition on $\boldsymbol{g}_0$ is given by

$$n\epsilon_{\ell,m,n} = m\theta_{\ell,m,n}, \quad n\delta_{\ell,m,n} = \ell\theta_{\ell,m,n} \quad \text{and} \quad m\delta_{\ell,m,n} = \ell\epsilon_{\ell,m,n}$$

for all $\ell, m, n \in \mathbb{N}$. In this case it follows that

$$\begin{aligned} \delta_{\ell,m,n} - \pi\ell\gamma_{\ell,m,n} &= \delta_{\ell,m,n} - \frac{\ell^2\delta_{\ell,m,n} + m\ell\epsilon_{\ell,m,n} + n\ell\theta_{\ell,m,n}}{(\ell^2 + m^2 + n^2)} \\ &= \frac{(m^2+n^2)\delta_{\ell,m,n} - m\ell\epsilon_{\ell,m,n} - n\ell\theta_{\ell,m,n})}{(\ell^2 + m^2 + n^2)} \\ &= 0 \end{aligned}$$

since $n\delta_{\ell,m,n} = \ell\theta_{\ell,m,n}$ and $m\delta_{\ell,m,n} = \ell\epsilon_{\ell,m,n}$. Similar arguments show that

$$\epsilon_{\ell,m,n} - \pi m\gamma_{\ell,m,n} = 0 \quad \text{and} \quad \theta_{\ell,m,n} - \pi n\gamma_{\ell,m,n} = 0$$

for all $\ell, m, n \in \mathbb{N}$. Hence

$$\boldsymbol{\varepsilon}_0 = 0 \quad \Leftrightarrow \quad \boldsymbol{\nabla} u_0 = \boldsymbol{g}_0$$

and the best approximate gradient becomes an exact gradient.

4.4 The solution using eigenfunction expansions

It has already been noted that the orthonormal eigenfunctions in H for the self-adjoint operator $\boldsymbol{\nabla}^*\boldsymbol{\nabla} : H \mapsto H$ are given by

$$p_{\ell,m,n} = \frac{2\sqrt{2}}{\sqrt{1 + \pi^2(\ell^2 + m^2 + n^2)}} \sin \ell\pi x \cdot \sin m\pi y \cdot \sin n\pi z$$

and hence

$$\boldsymbol{\nabla}^*\boldsymbol{\nabla} p_{\ell,m,n} = \frac{\pi^2(\ell^2 + m^2 + n^2)}{1 + \pi^2(\ell^2 + m^2 + n^2)}\, p_{\ell,m,n}$$

for each $\ell, m, n \in \mathbb{N}$. These vectors form a complete set in H. Thus our solution $u_0 \in H$ can be written in the form

$$u_0 = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{\ell,m,n} p_{\ell,m,n} \tag{9}$$

where the coefficients $c_{\ell,m,n}$ are to be determined. The corresponding orthonormal eigenvectors in K for the self-adjoint operator $\boldsymbol{\nabla}\boldsymbol{\nabla}^* : K \mapsto K$ are

$$\boldsymbol{q}_{\ell,m,n} = \frac{2\sqrt{2}}{\sqrt{\ell^2+m^2+n^2}} \begin{bmatrix} \ell\cos\ell\pi x\cdot\sin m\pi y\cdot\sin n\pi z \\ m\sin\ell\pi x\cdot\cos m\pi y\cdot\sin n\pi z \\ n\sin\ell\pi x\cdot\sin m\pi y\cdot\cos n\pi z \end{bmatrix}$$

with

$$\boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{q}_{\ell,m,n} = \frac{\pi^2(\ell^2+m^2+n^2)}{1+\pi^2(\ell^2+m^2+n^2)\pi^2}\ \boldsymbol{q}_{\ell,m,n}$$

for each $\ell, m, n \in \mathbb{N}$. Note that

$$\boldsymbol{\nabla} p_{\ell,m,n} = \frac{\pi\sqrt{\ell^2+m^2+n^2}}{\sqrt{1+\pi^2(\ell^2+m^2+n^2)}}\boldsymbol{q}_{\ell,m,n}$$

and

$$\boldsymbol{\nabla}^*\boldsymbol{q}_{\ell,m,n} = \frac{\pi\sqrt{\ell^2+m^2+n^2}}{\sqrt{1+\pi^2(\ell^2+m^2+n^2)}}p_{\ell,m,n}\ .$$

The vectors $\boldsymbol{q}_{\ell,m,n}$ do not span K. The orthonormal set of eigenvectors in K can be completed by adding the vectors

$$\boldsymbol{q}_{0,m,n} = 2\begin{bmatrix} \sin m\pi y\cdot\sin n\pi z \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{q}_{\ell,0,n} = 2\begin{bmatrix} 0 \\ \sin\ell\pi x\cdot\sin n\pi z \\ 0 \end{bmatrix}$$

$$\text{and}\quad \boldsymbol{q}_{\ell,m,0} = 2\begin{bmatrix} 0 \\ 0 \\ \sin\ell\pi x\cdot\sin m\pi y \end{bmatrix}$$

for which we note that

$$\boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{q}_{0,m,n} = 0, \quad \boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{q}_{\ell,0,n} = 0 \quad\text{and}\quad \boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{q}_{\ell,m,0} = 0$$

for each $\ell, m, n \in \mathbb{N}$ and also the vectors

$$\boldsymbol{r}_{\ell,m,n} = \frac{2\sqrt{2}}{\sqrt{m^2+n^2}} \begin{bmatrix} 0 \\ -n\sin\ell\pi x\cdot\cos m\pi y\cdot\sin n\pi z \\ m\sin\ell\pi x\cdot\sin m\pi y\cdot\cos n\pi z \end{bmatrix}$$

and

$$\boldsymbol{s}_{\ell,m,n} = \frac{2\sqrt{2}}{\sqrt{m^2+n^2}\sqrt{\ell^2+m^2+n^2}} \begin{bmatrix} (m^2+n^2)\cos\ell\pi x\cdot\sin m\pi y\cdot\sin n\pi z \\ -\ell m\sin\ell\pi x\cdot\cos m\pi y\cdot\sin n\pi z \\ -\ell n\sin\ell\pi x\cdot\sin m\pi y\cdot\cos n\pi z \end{bmatrix}$$

for which

$$\boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{r}_{\ell,m,n} = 0 \quad\text{and}\quad \boldsymbol{\nabla}\boldsymbol{\nabla}^*\boldsymbol{s}_{\ell,m,n} = 0$$

for each $\ell, m, n \in \mathbb{N}$. If we write

$$\boldsymbol{g} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} d_{0,m,n}\boldsymbol{q}_{0,m,n} + \sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} d_{\ell,0,n}\boldsymbol{q}_{\ell,0,n} + \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty} d_{\ell,m,0}\boldsymbol{q}_{\ell,m,0} + \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[d_{\ell,m,n}\boldsymbol{q}_{\ell,m,n} + e_{\ell,m,n}\boldsymbol{r}_{\ell,m,n} + f_{\ell,m,n}\boldsymbol{s}_{\ell,m,n}\right]$$

where the coefficients are calculated using the usual inner product formulae then it follows that $\boldsymbol{\nabla}^*\boldsymbol{g}$ is represented by the series

$$\boldsymbol{\nabla}^*\boldsymbol{g} = \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{\pi\sqrt{\ell^2+m^2+n^2}}{\sqrt{1+\pi^2(\ell^2+m^2+n^2)}} d_{\ell,m,n}\boldsymbol{p}_{\ell,m,n}\,. \tag{10}$$

Equating coefficients in the representations (9) and (10) gives

$$c_{\ell,m,n} = \frac{\sqrt{1+\pi^2(\ell^2+m^2+n^2)}}{\pi\sqrt{\ell^2+m^2+n^2}}\, d_{\ell,m,n}$$

for all $\ell, m, n \in \mathbb{N}$. Therefore the minimum mean square error for the observed function is

$$\|\boldsymbol{\varepsilon}\|_K^2 = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} d_{0,m,n}^2 + \sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} d_{\ell,0,n}^2 + \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty} d_{\ell,m,0}^2 + \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[e_{\ell,m,n}^2 + f_{\ell,m,n}^2\right]$$

while for the modified observed function the minimum mean square error is

$$\|\boldsymbol{\varepsilon}_0\|_K^2 = \sum_{\ell=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[e_{\ell,m,n}^2 + f_{\ell,m,n}^2\right].$$

5 Conclusions

We have shown that the best gradient approximation to an observed square integrable function on the unit cube satisfies a standard self-adjoint differential equation. The method relies on formulation of the gradient operator as a bounded linear operator on an appropriate Sobolev space. By using different coordinate systems with a range of symmetry conditions and introducing appropriate weight functions a similar analysis can be applied to a variety of regions both bounded and unbounded. The corresponding eigenfunction basis will normally involve the *so-called* special functions of mathematical physics.

In practical applications this procedure could be used to eliminate measurement errors when constructing a legitimate potential function to match an observed gradient.

ACKNOWLEDGEMENTS. The author wishes to thank Professor Peter Taylor from the University of Melbourne for inviting him to visit Melbourne University in 2007 when much of this work was completed.

References

[1] Jean-Pierre Aubin, (2000) *Applied Functional Analysis*, Second Edition, John Wiley & Sons, Inc.

[2] François Treves, (1967) *Topological Vector Spaces, Distributions and Kernels*, Academic Press.

[3] Kôsaku Yosida, (1980) *Functional Analysis*, Sixth Edition, Springer-Verlag.

Received: November 11, 2007