THE RECOVERY OF AN ANISOTROPIC CONDUCTIVITY
IN GROUNDWATER MODELLING

IAN KNOWLES AND AIMIN YAN

Abstract. In order to model groundwater flow effectively, one is faced inevitably with the problem of suitably estimating the various subsurface parameters in the differential equations governing the flow from measurements, over time, of the piezometric head at various well sites, together with certain ancillary data. A new method is presented that allows for the estimation of an anisotropic hydraulic conductivity as the unique global minimum of a convex functional.

1. Introduction

Flow in a confined groundwater system may be modelled by the parabolic equation

\[ Q(x) \frac{\partial w}{\partial t} = \nabla \cdot (P(x) \nabla w) + R(x, t) \]

over \(x\) in a bounded region \(\Omega \subset \mathbb{R}^n\), \(n = 2\) or \(3\), representing the physical subsurface region, and \(0 \leq t \leq 1\) \(\cite{3, 3.3.17}\). Here, \(\Omega\) is a \(C^2\) domain in \(\mathbb{R}^n\), \(Q > 0\) and \(R\) are assumed to be continuous, and \(P\) is a symmetric and strictly positive \(n \times n\) matrix with entries in \(C^2(\Omega)\). In physical terms, \(Q\) is the specific storage, \(w\) is the piezometric head, \(P\) is the hydraulic conductivity tensor, and \(R\) is a source/sink term for the flow. The hydraulic conductivity, which quantifies the ease with which the water traverses the porous medium, is said to be isotropic if the volumetric flux \(P(x) \nabla w\) (also called the Darcy flow) is everywhere parallel to the pressure gradient \(\nabla w\), i.e. \(P(x) = P(x)I\); otherwise, the conductivity is said to be anisotropic, and corresponds to the situation in which the porous medium possesses different flow properties in different directions, which is typical in layered media, and in fractured media, for example.

Due to the general inaccessibility of the subsurface, it is somewhere between difficult and impossible to measure directly the various parameters appearing in the flow equation. It is easier, though not necessarily easy, to obtain data on the piezometric head \(w(x, t)\) from wells drilled into the aquifer at various locations \(x\). We are interested here in the possibility of recovering parameters in the groundwater model from such measurements, together

\[ Date: \text{March 8, 2002.} \]

\[ \text{Supported in part by US National Science Foundation grants DMS-0107492 and DMS-0079478.} \]
with certain necessary ancillary data, as discussed below. In particular, we note that much has been written on the problem of recovering an isotropic conductivity; see [1, 2, 4, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 20, 22, 23, 24, 25] and the references therein, with no claim as to completeness. In contrast, relatively little is known about recovery in the anisotropic case, even though it is widely believed that anisotropy is the norm in nature, rather than the exception.

In the sections below, we present a method that seems to be effective in anisotropic recovery. The basic idea, which is outlined in the work [14], involves the transformation of the flow equation (1) via

$$u(x, \lambda) = \int_0^1 w(x, t) e^{-\lambda t} dt,$$

(2)

to the selfadjoint elliptic equation

$$-\nabla \cdot [P(x) \nabla u(x, \lambda)] + \lambda Q(x) u(x, \lambda) = R^*(x),$$

(3)

where

$$R^*(x) = \frac{e^{-\lambda} - 1}{\lambda} R(x) + Q(x)[w(x, 0) - w(x, 1) e^{-\lambda}].$$

(4)

Here we assume, for simplicity of exposition, that the recharge term $R$ is a function of $x$ only, and that the functions $Q$ and $R$ are known entities; the general case, which is covered in detail in [16] when $P$ is isotropic, extends easily to the matrix situation discussed below.

For any fixed $\lambda > 0$ it is a relatively simple matter to compute values $u(x, \lambda)$ from the known values for $w(x, t)$. We arrive then at the following modified problem: given $u(x, \lambda)$ for $x$ in $\Omega$ and all $\lambda > 0$ and $P$ on the boundary of $\Omega$, and given also the functions $Q$ and $R$, recover the matrix function $P$.

To this end, let $\mathcal{D}$ denote the set of all symmetric strictly positive matrix functions $p$ with $C^2(\Omega)$ entries and $p|_{\partial \Omega} = P|_{\partial \Omega}$, and consider the functional $G(p, \lambda)$ defined on $\mathcal{D}$ by

$$G(p, \lambda) = \int_\Omega p(x) \nabla (u - u_{p,\lambda}) \cdot \nabla (u - u_{p,\lambda}) + \lambda Q(x) (u - u_{p,\lambda})^2 dx,$$

(5)

where $v = u_{p,\lambda}$ is the unique solution of the boundary value problem

$$L_p(v) = -\nabla \cdot (p(x) \nabla v(x, \lambda)) + \lambda Q(x) v(x, \lambda) = R^*(x), \quad v|_{\partial \Omega} = u|_{\partial \Omega}.$$

(6)

Notice that, in this setting $u = u_{P,\lambda}$, where $P$ is the matrix function that we wish to recover. Notice also that, except for the next section, the smoothness assumption on $p$ in $\mathcal{D}$ may be relaxed to requiring that $p$ be in $L^\infty(\Omega)$.

The functional, $H$, that we actually minimize to recover the conductivity, is formed by choosing $N$ unequal positive values of the $\lambda$ parameter, $\lambda_1, \lambda_2, \ldots, \lambda_N$, and then setting

$$H(p) = \sum_{i=1}^N G(p, \lambda_i).$$

(7)
In two dimensions, for example, we seek three functions, \( P_{11}, P_{12} = P_{21}, \) and \( P_{22} \); so it is natural to expect that one would need to use at least three of the functions \( u(x, \lambda_i) \) in this process. A more precise discussion of this point leads us directly to the uniqueness question for this inverse problem, which we consider in some detail in the next section.

\section{Uniqueness}

By way of example, we begin by presenting conditions under which solutions of (3), \( u(x, \lambda_1) \) and \( u(x, \lambda_2), \lambda_1 \neq \lambda_2, \) determine the functions \( P_{11} \) and \( P_{22} \) in the two dimensional case.

Set \( n = 2 \) and consider matrices \( P \) and \( \tilde{P} \) in \( \mathcal{D} \) of the form

\[
P = \begin{pmatrix} P_{11} & 0 \\ 0 & P_{22} \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} \tilde{P}_{11} & 0 \\ 0 & \tilde{P}_{22} \end{pmatrix},
\]

and such that \( P = \tilde{P} \) on the boundary of \( \Omega \). Let solutions \( u(x, \lambda_i), i = 1, 2, \) satisfy the equations

\[
-\nabla \cdot [P(x)\nabla u(x, \lambda_i)] + \lambda_i Q(x)u(x, \lambda_i) = R^*(x), \quad i = 1, 2,
\]

\[
-\nabla \cdot [\tilde{P}(x)\nabla u(x, \lambda_i)] + \lambda_i Q(x)u(x, \lambda_i) = R^*(x), \quad i = 1, 2,
\]

so that

\[
-\nabla \cdot [(P(x) - \tilde{P}(x))\nabla u(x, \lambda_i)] = 0,
\]

for \( i = 1, 2 \). Let \( P_{11} - \tilde{P}_{11} = r, P_{22} - \tilde{P}_{22} = s, \) and define \( \mu = (r, s)^T \). Then, assuming \( x = (x_1, x_2) \), the equations (9) can be written in the matrix form

\[
A \mu_{x_1} + B \mu_{x_2} + C \mu = 0,
\]

where

\[
A = \begin{pmatrix} u_{x_1}(x_1, x_2, \lambda_1) & 0 \\ u_{x_1}(x_1, x_2, \lambda_2) & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & u_{x_2}(x_1, x_2, \lambda_1) \\ 0 & u_{x_2}(x_1, x_2, \lambda_2) \end{pmatrix},
\]

and

\[
C = \begin{pmatrix} u_{x_1x_1}(x_1, x_2, \lambda_1) & u_{x_2x_2}(x_1, x_2, \lambda_1) \\ u_{x_1x_1}(x_1, x_2, \lambda_2) & u_{x_2x_2}(x_1, x_2, \lambda_2) \end{pmatrix}.
\]

Notice that

\[
|A - \tau B| = \tau \begin{vmatrix} u_{x_1}(x_1, x_2, \lambda_1) & u_{x_2}(x_1, x_2, \lambda_1) \\ u_{x_1}(x_1, x_2, \lambda_2) & u_{x_2}(x_1, x_2, \lambda_2) \end{vmatrix},
\]

so that by the standard theory of characteristics for the system (10), see for example [5, p. 171], we have that, provided

\[
|A - \tau B| \neq 0
\]

throughout the region \( \Omega \), initial values of the vector function \( \mu(x) \) on any curve in \( \Omega \) for which no segment is parallel to a coordinate axis, may be uniquely continued throughout the region. So, noting that \( \mu \) is zero on the boundary of \( \Omega \), if no segment of the boundary of \( \Omega \) is parallel to a coordinate
axis, we see that uniqueness holds. On the other hand, if a part of the boundary takes the form $x_1 = a$ for some constant $a$, we have that $\tau = 0$ and the above theory is not applicable; an exactly similar situation arises if the boundary has the form $x_2 = b$ for some constant $b$, corresponding to $\sigma = \tau^{-1} = 0$. In these cases however, we can assume that $\mu$ is known on and below a line $x_2 = b$ intersecting $x_1 = a$ at $x = (a, b)$; this is so, for example, if the boundary for $x_2 < b$ is not parallel to a coordinate axis. Now, provided that the values $u_{x_1}(a, b, \lambda)$ and $u_{x_2}(a, b, \lambda)$ are non-zero for some $\lambda$, the values of $r_{x_1}$ are known on the line $x_2 = b$ near $(a, b)$; in consequence, as the values of $s_{x_2}$ may be found on $x_2 = b$ near $(a, b)$ from the differential equation, the values of $s$ may be propagated parallel to the $x_2$ axis in a unique fashion. Likewise, the values of $s_{x_2}$ are known on the line $x_1 = a$, so that from the differential equation again, the value of $r_{x_1}$ is known and the values of $r$ may be propagated parallel to the $x_1$ axis in a similar unique fashion. Thus the values of $\mu$ propagate uniquely from such a corner. In summary, we have the following uniqueness result:

**Proposition 1.** In two dimensions the diagonal elements $P_{11}$ and $P_{22}$ of the matrix $P$ in $D$ are uniquely identified by two solutions $u(x, \lambda_1)$ and $u(x, \lambda_2)$, $\lambda_1 \neq \lambda_2$, of (3), provided that for some $\lambda$ the functions $u_{x_1}(x_1, x_2, \lambda)$ and $u_{x_2}(x_1, x_2, \lambda)$ are non-zero in $\Omega$ and the condition (11) holds throughout the region $\Omega$.

**Remarks.**

1. The same argument may be applied to identify any two of the functions $P_{11}$, $P_{12}$, and $P_{22}$, if the third function is known.

2. It is clear from the above proof that boundary data for $P$ is only needed on a part of the boundary; for example, the specification of $P$ on two adjacent sides of a square would usually suffice. For practical data, this corresponds to a knowledge of $P$ on either the “inflow” or “outflow” parts of the boundary; as neither of these parts is usually known a priori, in practice it is usual to require that $P$ be known on all of $\partial \Omega$.

Consider now the identification of all three of the entries in $P$. Here we continue to assume that $n = 2$ and consider matrices $P$ and $\tilde{P}$ in $D$ of the form

\[
P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12} & \tilde{P}_{22} \end{pmatrix},
\]

again assuming that $P = \tilde{P}$ on the boundary of $\Omega$. We assume that solutions $u(x, \lambda_i)$, $1 \leq i \leq 3$, are given with the $\lambda_i$ pairwise unequal. Then (9) holds for $1 \leq i \leq 3$. With $r$ and $s$ as above set $P_{12} - \tilde{P}_{12} = t$, and $\mu = (r, s, t)^T$. 

In this case the equations (9) take the matrix form (10) with

\[
A = \begin{pmatrix}
    u_{x_1}(x_1, x_2, \lambda_1) & 0 & u_{x_2}(x_1, x_2, \lambda_1) \\
    u_{x_1}(x_1, x_2, \lambda_2) & 0 & u_{x_2}(x_1, x_2, \lambda_2) \\
    u_{x_1}(x_1, x_2, \lambda_3) & 0 & u_{x_2}(x_1, x_2, \lambda_3)
\end{pmatrix},
\]

\[
B = \begin{pmatrix}
    0 & u_{x_2}(x_1, x_2, \lambda_1) & u_{x_1}(x_1, x_2, \lambda_1) \\
    0 & u_{x_2}(x_1, x_2, \lambda_2) & u_{x_1}(x_1, x_2, \lambda_2) \\
    0 & u_{x_2}(x_1, x_2, \lambda_3) & u_{x_1}(x_1, x_2, \lambda_3)
\end{pmatrix},
\]

\[
C = \begin{pmatrix}
    u_{x_1x_1}(x_1, x_2, \lambda_1) & u_{x_2x_2}(x_1, x_2, \lambda_1) & 2u_{x_1x_2}(x_1, x_2, \lambda_1) \\
    u_{x_1x_2}(x_1, x_2, \lambda_2) & u_{x_2x_2}(x_1, x_2, \lambda_2) & 2u_{x_1x_2}(x_1, x_2, \lambda_2) \\
    u_{x_1x_1}(x_1, x_2, \lambda_3) & u_{x_2x_2}(x_1, x_2, \lambda_3) & 2u_{x_1x_2}(x_1, x_2, \lambda_3)
\end{pmatrix}.
\]

Now, the $3 \times 3$ matrix $A - \tau B$ has rank at most 2 for any value of $\tau$; in fact, the rank is at most 2 even if $N > 3$ values of $\lambda$ are introduced, when the matrix is of order $N \times 3$. Thus the standard uniqueness argument fails here. However, one can adapt the “corner” argument used above to show that, under analogous conditions, the vector $\mu$ may be propagated uniquely into the interior as before. Here, the differential equations take, for each $\lambda$, the form

\[
(13) \quad r_{x_1} u_{x_1} + t_{x_1} u_{x_2} + t_{x_2} u_{x_1} + s_{x_2} u_{x_2} + \ldots = 0,
\]

where the dots indicate other terms that are known on the initial curve. Under condition (11) if $\mu$ is known at the boundary of a “corner” $(a, b)$ then $\mu_{x_1}$, and in particular $r_{x_1}$ and $t_{x_1}$ are known on $x_2 = b$, and so from two of the equations (13) and assuming condition (11) the derivatives $s_{x_2}$ and $t_{x_2}$ can be uniquely recovered, and the values of $s$ and $t$ uniquely propagated parallel to $x_1 = a$. In similar fashion, the derivatives $s_{x_2}$ and $t_{x_2}$ are known on $x_1 = a$, and by the same reasoning $r$ and $t$ can be uniquely propagated parallel to $x_2 = b$.

It is worth noting that Proposition 1 can be extended to include unique recovery of the entries in $P(x)$ together with $Q(x)$ and $R(x, t)$, by combining the method above with that of [14]; in this case the condition (11) has to be modified appropriately. Conditions like (11) have a physical origin and reflect the fact that one would not expect to recover flow parameters like $P$, $Q$ and $R$ in regions where there is no movement of the groundwater. In practice conditions like (11) are relatively easy to check computationally for a given dataset. This general type of restriction is embodied in the following uniqueness assumption that is assumed to be in force in the material to follow:

**Uniqueness Assumption.** Recall that $u_{p, \lambda}$ denotes the unique solution of the boundary value problem (6), and assume that $m$ entries of $P$ in $\mathcal{D}$ are to be identified. Then for any matrix functions $p_1$ and $p_2$ in $\mathcal{D}$, $p_1 = p_2$ if and only if $u_{p_1, \lambda} = u_{p_2, \lambda}$ for $m$ distinct values of $\lambda$. 
3. Properties of G and H

Before proceeding to the actual minimization procedures used to compute the conductivity entries, we first list some properties of the functionals \( G \) defined above that are needed later.

**Proposition 2.**

(a) For any matrix \( p \) in \( D \),

\[
G(p, \lambda) = \int_{\Omega} p(x) \nabla u \cdot \nabla u - p(x) \nabla u_{p,\lambda} \cdot \nabla u_{p,\lambda} + \\
\lambda Q(x)(u^2 - u_{p,\lambda}^2) - 2R^*(x)(u - u_{p,\lambda}) \, dx.
\]

(b) For any matrix \( p \) in \( D \) and any matrix \( h \) the Gâteaux derivative of \( G \) at \( p \) in the direction \( h \) is given by

\[
G'(p, \lambda)[h] = \int_{\Omega} h(x) \nabla u \cdot \nabla u - h(x) \nabla u_{p,\lambda} \cdot \nabla u_{p,\lambda} \, dx.
\]

(c) For matrices \( p \) in \( D \), and symmetric matrices \( h \), and \( k \), the second Gâteaux derivative for \( G \) is given by

\[
G''(p, \lambda)[h, k] = 2 \int_{\Omega} L_p^{-1}(e_p(h)) e_p(k)
\]

where \( L_p \) is the operator formed in \( L^2(\Omega) \) from equation (6) and homogeneous Dirichlet boundary conditions, and

\[
e_p(h) = \nabla \cdot h \nabla u_{p,\lambda}.
\]

**Proof.** To establish part (a) first note the identity

\[
G(p, \lambda) = \int_{\Omega} p(x) \nabla u \cdot \nabla u - p(x) \nabla u_{p,\lambda} \cdot \nabla u_{p,\lambda} + \\
+ 2p(x) \nabla u_{p,\lambda} \cdot \nabla (u_{p,\lambda} - u) + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx.
\]

Now, using integration by parts and equation (6) together with the fact that the solutions \( u \) and \( u_{p,\lambda} \) share the same boundary data, we have, after some rearrangement

\[
\int_{\Omega} 2p(x) \nabla u_{p,\lambda} \cdot \nabla (u_{p,\lambda} - u) + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx = \\
\int_{\Omega} -2(u_{p,\lambda} - u) \nabla \cdot p(x) \nabla u_{p,\lambda} + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx = \\
\int_{\Omega} 2(u_{p,\lambda} - u)[R^*(x) - \lambda Q(x)u_{p,\lambda}] + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx = \\
\int_{\Omega} \lambda Q(x)(u^2 - u_{p,\lambda}^2) - 2R^*(x)(u - u_{p,\lambda}) \, dx
\]

and (a) now follows.
For any $\mathbf{p}$ in $\mathcal{D}$ and any function matrix $\mathbf{h}$, and any real $\epsilon$, we have, using part (a) together with some straightforward manipulation,

$$G(\mathbf{p} + \epsilon \mathbf{h}, \lambda) - G(\mathbf{p}, \lambda)$$

$$= \int_{\Omega} \left\{ (\mathbf{p} + \epsilon \mathbf{h}) \nabla u \cdot \nabla u - (\mathbf{p} + \epsilon \mathbf{h}) \nabla u_{p+\epsilon h,\lambda} \cdot \nabla u_{p+\epsilon h,\lambda} + \lambda Q(x)(u^2 - u_{p+\epsilon h,\lambda}^2) - 2R^*(x)(u - u_{p+\epsilon h,\lambda}) - p(x) \nabla u \cdot \nabla u \ight. \nabla u_{p+\epsilon h,\lambda} \cdot \nabla u_{p+\epsilon h,\lambda} - \lambda Q(x)(u^2 - u_{p,\lambda}^2) + 2R^*(x)(u - u_{p,\lambda}) \} \, dx$$

$$= \epsilon \int_{\Omega} \{ \mathbf{h}(x) \nabla u \cdot \nabla u - \mathbf{h}(x) \nabla u_{p+\epsilon h,\lambda} \cdot \nabla u_{p+\epsilon h,\lambda} \} \, dx$$

Next, we integrate by parts the $\mathbf{p}$-terms in the second integral, observing that $u_{p+\epsilon h,\lambda}$ and $u_{p,\lambda}$ have the same boundary data, and use the differential equation (6) again to obtain

$$G(\mathbf{p} + \epsilon \mathbf{h}, \lambda) - G(\mathbf{p}, \lambda)$$

$$= \epsilon \int_{\Omega} \{ \mathbf{h}(x) \nabla u \cdot \nabla u - \mathbf{h}(x) \nabla u_{p+\epsilon h,\lambda} \cdot \nabla u_{p+\epsilon h,\lambda} \} \, dx$$

The formula in (b) now follows.

Finally, using (b), we have for $\mathbf{p}$ in $\mathcal{D}$ and $\mathbf{h}$, $\mathbf{k}$ any symmetric function matrices that

$$\mathbf{h}(x) \nabla u_{p,\lambda} \cdot \nabla u_{p+\epsilon k,\lambda} = \mathbf{h}(x) \nabla u_{p+\epsilon k,\lambda} \cdot \nabla u_{p,\lambda},$$
and hence

\[ G'(p + \epsilon k, \lambda)[h] - G'(p, \lambda)[h] \]

\[ = \int_{\Omega} h(x) \nabla u_{p,\lambda} \cdot \nabla u_{p,\lambda} - h(x) \nabla u_{p+\epsilon k,\lambda} \cdot \nabla u_{p+\epsilon k,\lambda} \, dx \]

\[ = -\int_{\Omega} h(x) \nabla u_{p,\lambda} \cdot \nabla (u_{p+\epsilon k,\lambda} - u_{p,\lambda}) \]

\[ + h(x) \nabla u_{p+\epsilon k,\lambda} \cdot \nabla (u_{p+\epsilon k,\lambda} - u_{p,\lambda}) \, dx \]

\[ = \int_{\Omega} (u_{p+\epsilon k,\lambda} - u_{p,\lambda}) \nabla \cdot h(x) \nabla (u_{p+\epsilon k,\lambda} + u_{p,\lambda}) \, dx. \]

From the identity

\[ -\nabla \cdot p(x) \nabla (u_{p+\epsilon k,\lambda} - u_{p,\lambda}) + \lambda Q(x)(u_{p+\epsilon k,\lambda} - u_{p,\lambda}) = \epsilon \nabla \cdot k(x) \nabla u_{p+\epsilon k,\lambda}, \]

it follows that

\[ u_{p+\epsilon k,\lambda} - u_{p,\lambda} = \epsilon L^{-1}_p (\nabla \cdot k(x) \nabla u_{p+\epsilon k,\lambda}). \]

The formula (c) now follows directly from the above identities. \(\square\)

**Remark.** By slightly modifying the above arguments one can show that the derivatives of \(G\) listed above are actually Fréchet derivatives; in fact \(G\) is an analytic functional in the sense of [6, Ch. 15].

**Proposition 3.**

(a) For \(p\) in \(D\) and \(\lambda > 0\)

\[ G(p, \lambda) = 0 \iff G'(p, \lambda) = 0 \iff u_{p,\lambda} = u. \]

(b) Assume that \(p\) lies in \(D\) and \(G''(p, \lambda)[h, h] = 0\) for some symmetric matrix \(h\). Then \(u_{p+\epsilon h,\lambda} = u_{p,\lambda}\) for all \(\epsilon\) small enough.

**Proof.** The assertion in (a) that \(G(p, \lambda) = 0\) if and only if \(u_{p,\lambda} = u\) follows immediately from the definition of \(G\), and one direction of the remaining assertion is obvious. If \(G'(p)[h] = 0\) for all \(h\) then the \(L^2\) gradient of \(G\), \(\nabla G\), satisfies \(\nabla G(p) = (\gamma_{jk}) = 0\), where, for \(1 \leq j, k \leq n\),

\[ (17) \quad \gamma_{jk} = u_x^j u_{x_k} - u_{x_j,x_k} + u_{x_k,x_j}. \]

Consequently, from Proposition 2 part (a),

\[ \int_{\Omega} p(x) \nabla (u - u_{p,\lambda}) \cdot \nabla (u - u_{p,\lambda}) + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx \]

\[ = \int_{\Omega} Q(x)(u^2 - u_{p,\lambda}^2) - 2R^*(x)(u - u_{p,\lambda}) \, dx. \]

If we interchange \(p\) and \(P\) in this formula and note that \(u = u_{P,\lambda}\) we obtain

\[ \int_{\Omega} P(x) \nabla (u - u_{p,\lambda}) \cdot \nabla (u - u_{p,\lambda}) + \lambda Q(x)(u - u_{p,\lambda})^2 \, dx \]

\[ = \int_{\Omega} Q(x)(u_{p,\lambda}^2 - u^2) - 2R^*(x)(u_{p,\lambda} - u) \, dx. \]
Adding, we find that

\[ \int_{\Omega} (p + P)(x) \nabla(u - u_{p,\lambda}) \cdot \nabla(u - u_{p,\lambda}) + 2\lambda Q(x)(u - u_{p,\lambda})^2 \, dx = 0, \]

from which it follows that \( u - u_{p,\lambda} = 0 \).

Part (b) is a consequence of Proposition 2 part (c) in that, if we assume that \( G''(p, \lambda)[h, h] = 0 \), then \( e_p(h) = \nabla \cdot h \nabla u_{p,\lambda} = 0 \), so that from (6), for all \( \epsilon \) small enough \( p + \epsilon h \) is strictly positive and

\[ -\nabla \cdot (p(x) + \epsilon h) \nabla u_{p,\lambda} + \lambda Q(x) u_{p,\lambda} = R^*(x). \]

But \( u_{p+\epsilon h,\lambda} \) is the unique solution of this equation with the boundary data \( u|_{\partial \Omega} \); it follows immediately that \( u_{p+\epsilon h,\lambda} = u_{p,\lambda} \).

We now combine the previous results into the main objective of this section:

**Theorem 4.** Assume that the uniqueness assumption of \( \S 1 \) holds, and in equation (7) set \( N = m \). Then, for \( p \) in \( D \),

\[ H(p) = 0 \iff H'(p) = 0 \iff p = P, \]

and the functional \( H \) is strictly convex on \( D \).

**Proof.** Noting that \( H(p) = 0 \) if and only if \( G(p, \lambda_i) = 0 \) for \( 1 \leq i \leq m \), the first assertion follows from Proposition 3 part (a), and the uniqueness assumption. Next, if \( H'(p) = 0 \), the same proof that was used for \( G \) shows that \( H(p) = 0 \), and the rest follows from the statements above. Finally, let \( H''(p)[h, h] = 0 \). As the functionals \( G(p, \lambda_i), 1 \leq i \leq m \), are convex, it follows that \( G''(p, \lambda_i)[h, h] = 0 \) for \( 1 \leq i \leq m \). By Proposition 3 part (b), for all \( \epsilon \) small enough and \( 1 \leq i \leq m \), \( u_{p+\epsilon h,\lambda_i} = u_{p,\lambda_i} \). The uniqueness assumption now dictates that \( h = 0 \). \( \square \)

4. Implementation and Results

Theorem 4 shows that, under computationally verifiable conditions on the data functions \( u_i = u_{P,\lambda_i}, 1 \leq i \leq m \), an anisotropic conductivity \( P \) can be recovered by minimization of the functional \( H \) defined by

\[ H(p) = \sum_{i=1}^{m} G(p, \lambda_i) \]

We use a version of the steepest descent method. At the outset we note that the \( L^2 \) gradient of \( H \), which is the matrix with entries

\[ (\nabla H)_{jk} = \sum_{i=1}^{m} \gamma_{ijk}, \]

where \( \gamma_{ijk} \) are the entries in \( \nabla G(p, \lambda_i) \), is not suitable as it stands for use in the descent process because in general this gradient is not zero on the
boundary of \( \Omega \). This is a critically important consideration here, as it is an integral part of the uniqueness theory that \( P \) be specified on \( \partial \Omega \).

To circumvent this difficulty we use an alternative gradient due to Neuberger [19]. For each \( \lambda_i \) the Neuberger gradient, \( \eta_i = \nabla N G(p, \lambda_i) = (\eta_{ijk}) \), is defined for \( 1 \leq j, k \leq n \) by

\[
-\Delta \eta_{ijk} + \eta_{ijk} = (\nabla G)_{ijk}, \quad \eta_{ijk}\big|_{\partial \Omega} = 0.
\]

(20)

Notice that if we define the \( \mathcal{L}^2 \) inner product for \( n \times n \) matrix functions \( h = (h_{jk}) \) and \( g = (g_{jk}) \) to be

\[
(h, g)_{L^2} = \sum_{j,k=1}^{n} \int_{\Omega} h_{jk}(x)g_{jk}(x) \, dx,
\]

and \( \gamma_i = (\gamma_{ijk}) \) then \( G'(p, \lambda_i)[h] = (\gamma_i, h)_{L^2} \). In the same fashion, if

\[
(h, g)_{H^1} = \sum_{j,k=1}^{n} \int_{\Omega} \nabla h_{jk} \cdot \nabla g_{jk} + h_{jk}g_{jk} \, dx,
\]

is the corresponding Sobolev inner product, then \( G'(p, \lambda_i)[h] = (\eta_i, h)_{H^1} \), and so not only does the matrix \( \eta_i \) provide a descent direction for \( G(p, \lambda_i) \) (and hence, over \( 1 \leq i \leq m \), for \( H \)) but, most importantly, these matrices vanish on the boundary of \( \Omega \), so that the boundary data for an initial matrix \( P_0 \) is preserved throughout the descent process.

In the numerical work we assume that the region \( \Omega = [-1, 1] \times [-1, 1] \) is overlaid with a \( 30 \times 30 \) discretization grid. The code was written in PGI Fortran 90 and run on a Dell PowerEdge 2450 with Intel 733MHz Pentium III processors and RedHat Linux 6.2.

Since the recovery is very sensitive to the error of the numerical solution, a solver is needed that can accurately and efficiently solve elliptic boundary value problems of the type

\[
-(P_{11}u_x)_x - (P_{12}u_x)_y - (P_{21}u_y)_x - (P_{22}u_y)_y + Qu = F
\]

\[
u(x, y) = B(x, y), \quad (x, y) \in \partial \Omega,
\]

with minimal error. We use the five point difference method with

\[
(P_{11}u_x)_x(i, j)h_x^2 = P_{11}(i + \frac{1}{2}, j)(U_{i+1,j} - U_{i,j}) + P_{11}(i - \frac{1}{2}, j)(U_{i,j} - U_{i-1,j})
\]

\[
(P_{22}u_y)_y(i, j)h_y^2 = P_{22}(i, j + \frac{1}{2})(U_{i,j+1} - U_{i,j}) + P_{22}(i, j - \frac{1}{2})(U_{i,j} - U_{i,j-1})
\]

\[
(P_{12}u_x)_y(i, j)4h_xh_y = (P_{12}(i + 1, j) - P_{12}(i - 1, j))(U_{i,j+1} - U_{i,j-1})
\]

\[
(P_{21}u_y)_x(i, j)4h_xh_y = (P_{21}(i, j + 1) - P_{21}(i, j - 1))(U_{i+1,j} - U_{i-1,j})
\]

Then we employ the band solving subroutine BANDEC adapted from [21] to solve the resulting system of linear equations. This solver was used in conjunction with the true matrix \( P \) to produce the solutions \( u(x, \lambda) \) that
formed our synthetic dataset. It was also called upon to determine the various solutions $u_{p,\lambda}$ that arise during the descent procedure. The numerical derivatives arising from (14) were computed by central differences. This is accurate and efficient here, because the solutions being differentiated are sufficiently smooth functions. For data with noise, one must apply more sophisticated numerical differentiation techniques (c.f. [15, §5]).

We iterate the Simpson’s rule function qsimp in [21] to perform the required quadrature in the formula (5) for $G(p, \lambda)$. Given that parts of the integrand lack smoothness, Simpson’s rule is an effective choice here.

Finally, with regard to the minimization code, we adapted the bracketing and line minimization approach in [21]. The idea is as follows. First, check to see if $H$ can be minimized along a given direction by testing the value of $H(p)$ defined by (7) at the given point and at a nearby point along the given direction. If so, we use our own (somewhat primitive) bracketing method to find the bracketing points. This consists of stepping along the chosen direction using a predetermined increment until a bracketing is found, or a preset stepping limit is encountered; in the latter case the original $p$ is reset to the new $p$ at the stepping limit and a new gradient is computed.

Due to the uncertainties associated with inverse problem computations, this ultra-reliable approach to bracketing has been found to be appropriate, at least for use in test code. Finally, we use the Brent method (adapting the brent function in [21]) to find the minimum. We tested the recovery of $P_{11}$, $P_{12}$ and $P_{22}$ in different combinations and found finally that we get the fastest convergence when we regard $H$ as a function of three variables and use the Powell method by adapting the subroutine powell in [21].

Since the elliptic solver is sensitive to a loss of positive definiteness for $p$, the program tended to crash when non-positive eigenvalues for $p$ were encountered. Noting that $p$ is positive definite if and only if

$$p_{11} > 0, \quad p_{22} > 0, \quad p_{11}p_{22} - p_{12}^2 > 0,$$

we argue that it is reasonable to set lower bounds on the functions $p_{11}$ and $p_{22}$ consistent with local knowledge of a particular aquifer and with the knowledge (c.f. [15, §5]) that the insertion of additional information tends to have a stabilizing effect on an ill-conditioned computation. It is not clear from the physical problem how one might constrain $p_{12}$; we chose to bound the absolute value of $p_{12}$ by the product of the lower bounds for $p_{11}$ and $p_{22}$, so that $p_{11}p_{22} > p_{12}^2$ is always true. Whenever the computed values of $p$ are under the lower bound in the descent search for $p_{11}$ and $p_{22}$, or above the bound for $|p_{12}|$, we set them equal to the bound. With this arrangement, the algorithm became extremely stable with respect to allowing a large number of the descent iterations.

As described in §2, we need at least three different $\lambda$ values to recover $P_{11}$, $P_{12}$ and $P_{22}$. In practice, we found that increasing the number of $\lambda$ values used substantially improved the images. In our tests, we chose the number of $\lambda$ values to be 20. This is consistent with the view that the ill-posedness
Figure 1. Recovery of $P$. 

(a) True $P_{11}$  
(b) Computed $P_{11}$  

(c) True $P_{12}$  
(d) Computed $P_{12}$  

(e) True $P_{22}$  
(f) Computed $P_{22}$
in the computational problem corresponds to a certain loss of information in the data, and, as noted above, the most natural way to offset this is to add as much ancillary information as possible.

We deliberately selected non-smooth $P_{11}$, $P_{12}$ and $P_{22}$ for testing, because the recovery of non-smooth functions is more difficult than the recovery of smooth ones, and also because one cannot assume a priori that the parameters in a real groundwater system are smooth functions. As can be seen, the recovery of $P_{11}$ and $P_{22}$ in Figure 1 is quite good, as the discontinuities are quite clear, and the height is reasonably accurate; this is consistent with $\mathcal{L}^1$ convergence results for isotropic conductivities given in [15]. The recovery of $P_{12}$ in Figure 1 is also very reasonable, except for some small (in the $\mathcal{L}^1$ sense) artifacts. The artifacts are side effects of the recovery of $P_{11}$ and $P_{22}$.  

**Figure 2.** $P_{22}$ constant
They are relatively inconspicuous, for example, when we take $P_{22}$ to be a constant, as may be seen in Figure 2.

5. ACKNOWLEDGEMENT

We are indebted to the referee for a careful review of the original manuscript, and for a number of suggestions greatly improving the readability of the paper.

REFERENCES


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALABAMA AT BIRMINGHAM, BIRMINGHAM, ALABAMA, 35294