Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Acad. Sci. Paris, Ser. I 336 (2003) 691–696
COMPTES RENDUS MATHEMATIQUE

Mathematical Problems in Mechanics/Numerical Analysis

Gradient computation in a nonlinear inverse problem
Calcul du gradient dans un problème inverse non linéaire

Dan-Gabriel Calugaru [a], Jean-Marie Crolet [b]

[a] *Université Claude Bernard Lyon 1, ISTIL, MCS/CDCSP, 15, boulevard Latarjet, 69622 Villeurbanne cedex, France*
[b] *Université de Franche-Comté, Équipe de calcul scientifique, 16, route de Gray, La Bouloie, 25030 Besançon cedex, France*

Received 27 November 2002; accepted 4 March 2003

Presented by Charles-Michel Marle

Abstract

This paper deals with a nonlinear inverse problem to determine the Neumann condition on the boundary $\Gamma_L \subset \partial\Omega$, from measurements in the domain Ω. This condition is characterised by the width of Γ_L and by the constant value of the flux on this boundary. The direct problem is the Laplacian problem corresponding to flow modelling in a confined aquifer and Γ_L corresponds to the contact with a fault. Some properties of associated direct application are given and in particular, we show how one can compute its gradient by some explicit formulas. ***To cite this article: D.-G. Calugaru, J.-M. Crolet, C. R. Acad. Sci. Paris, Ser. I 336 (2003).***

Résumé

On s'intéresse à un problème inverse non linéaire d'identification de la condition Neumann sur la frontière $\Gamma_L \subset \partial\Omega$, à partir de mesures dans le domaine Ω. Cette condition est caractérisée par la largeur de Γ_L et par la valeur constante du flux sur cette frontière. Le problème direct est celui du laplacien et correspond à la modélisation de l'écoulement dans un aquifère captif en contact avec une faille. On étudie quelques propriétés de l'application directe associée et, en particulier, nous donnons des formules explicites pour calculer son gradient. ***Pour citer cet article : D.-G. Calugaru, J.-M. Crolet, C. R. Acad. Sci. Paris, Ser. I 336 (2003).***

Version française abrégée

Depuis quelques années, l'étude de problèmes inverses a reçu de plus en plus d'attention. La technique habituelle pour la résolution d'un problème inverse est la méthode des moindres carrés qui consiste à formuler et résoudre un problème d'optimisation pour une fonction objectif [4]. Pour résoudre de tels problèmes, il existe un nombre important d'algorithmes numériques, mais la plupart nécessitent, à chaque itération, le calcul d'un gradient

E-mail addresses: calugaru@cdcsp.univ-lyon1.fr, calugaru@math.univ-fcomte.fr (D.-G. Calugaru), jmcrolet@univ-fcomte.fr (J.-M. Crolet).

1631-073X/03/$ – see front matter
doi:10.1016/S1631-073X(03)00130-4

(voire d'un hessien). Généralement, ce calcul revient à déterminer la matrice de sensibilité dont les composantes expriment la variation de la variable d'état aux points d'observation par rapport à la variation de chaque paramètre. Pour ceci, il y a quelques méthodes usuelles (par différences finies, par équations de sensibilité ou par état adjoint), mais ces méthodes possèdent toutes des inconvénients [3].

Dans cette Note, on considère un problème inverse à deux inconnues scalaires. L'origine du problème est liée à la sismique et plus précisément au comportement d'une faille lors de l'augmentation des contraintes tectoniques. On restreint le problème physique complexe (écoulement dans un système aquifère multicouche traversé par une faille verticale), à l'étude de l'écoulement dans l'aquifère supérieur, milieu poreux continu, saturé, homogène, isotrope et incompressible. De plus, on suppose que cet aquifère n'est pas traversé par la faille mais que celle-ci reste toujours en-dessous, l'interface aquifère/faille (notée Γ_L) restant toujours localisée sur la frontière inférieure de l'aquifère (Γ_{inf}) comme à la Fig. 1. On suppose aussi que le flux F est constant sur Γ_L et que l'interface aquifère/faille est symétrique par rapport à l'axe de la faille : $\Gamma_L = [-L, L] \times \{0\}$. En dehors de Γ_L, sur la frontière inférieure du domaine, le flux est nul.

Les conséquences d'une activité sismique pour l'aquifère supérieur seraient l'augmentation du flux F entrant par la partie de frontière Γ_L, ainsi que l'élargissement de cette interface [1]. L'objectif du problème inverse est alors d'identifier la condition de Neumann sur Γ_L, i.e. de déterminer le flux entrant par Γ_L dans l'aquifère ainsi que la largeur de cette interface, $2L$, à partir de la connaissance du niveau piézométrique dans un ensemble Ω^{obs} de points d'observation, de cardinal M.

Le problème direct consiste à résoudre l'écoulement dans l'aquifère pour un couple donné de paramètres (F, L) et il est décrit par le système (1). Le problème inverse consistant à déterminer le couple (F, L) vérifiant les observations est formulé dans (2) et admet une formulation par moindres carrés classique.

Pour ce problème inverse, on peut déterminer des formulations explicites pour les coefficients de sensibilité : les dérivées premières de l'application directe associée par rapport aux deux paramètres sont exprimées par Éq. (4) pour la dérivée en F et par Éq. (5) pour la dérivée en L. La prise en compte de ces formules améliore considérablement le temps de calcul et la précision des algorithmes numériques d'optimisation utilisés pour résoudre le problème inverse.

1. Introduction

In the last decade, the interest for inverses problems dealing with estimation of physical parameters, initial state or boundary conditions in various physical systems has increased. In these problems, one uses some supplementary data resulting from measurements in accessible regions.

Conventional method for solving such problems is to use the output least squares (OLS) formulation, i.e., to solve an optimization problem for an objective function, which has the form of a sum of square functions [4]. Generally speaking, noting $\mathbf{p} = \{p_1, \ldots, p_N\}$ the set of unknown parameters, $\mathbf{u}_\mathbf{p}$ the state variable corresponding to $\mathbf{p}$ model parameters, and $\mathbf{u}^{\text{obs}} = \{u_1^{\text{obs}}, \ldots, u_M^{\text{obs}}\}$ the observations set corresponding to measurements in points P_i $(i = 1, \ldots, M)$, then the objective function is $J(\mathbf{p}) = \frac{1}{2}\sum_{i=1}^{M} \Phi_i(\mathbf{p})^2$, where $\Phi_i(\mathbf{p}) = u_\mathbf{p}(P_i) - u_i^{\text{obs}}$. Let us denote $\boldsymbol{\Phi} = (\Phi_1(\mathbf{p}), \ldots, \Phi_M(\mathbf{p}))$ the residual vector.

The major advantage of this formulation is the possibility to use numerical optimization methods to minimise the objective function on an admissible set of parameters, while the practical disadvantage is the computation of gradient vector (and eventually of Hessian matrix) of objective function. For instance, if Gauss–Newton method or one of its alternatives is used, the gradient and the Hessian are used in the form $\nabla J = \mathbf{A}^{\mathrm{T}}\boldsymbol{\Phi}$, $\mathbf{H} \simeq \mathbf{A}^{\mathrm{T}}\mathbf{A}$, where $A_{ij} = \partial\Phi_i/\partial p_j$, for $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N\}$. There are several techniques to realise the computation of sensitivity coefficients A_{ij} (by finite differences, by sensitivity equations or by adjoint state), but there are some practical drawbacks for each method [3].

In this paper, an inverse problem with only two scalar unknowns is considered and we show how to find and use explicit formulas to compute sensitivity coefficients. The outline is as follows: the next section deals with the physical problem and with the practical interest to identify the Neumann condition. The mathematical model that governs the direct problem as well as the formulation of the inverse problem to identify this condition are given in Section 3. Some properties of direct application (nonlinearity, continuity, differentiability) are the object of Section 4 where some explicit formulas for sensitivity coefficients are also deduced. Numerical implementation of this formulas is finally discussed.

2. Physical problem

The origin of the inverse problem studied in this paper is related with the seismical research and more precisely with the behavior of a fault during tectonic constraints. Indeed, one of physical systems which interest the seismologists is a multi-layer system (several superposed horizontal aquifers separated by impermeable layers) crossed by a vertical fault.

During the increase of tectonic constraints, there are some modifications in the deep aquifers, like increases of groundwater pressure. Moreover, the width and the permeability of the fault increase (other cracks are generated and older fissures become larger). These phenomena lead to a supplementary ascendant flow, essentially through the fault and then, the residual state of the flow in the subsurface aquifer can be strongly perturbed [1]. Consequently, a seismic activity in deeper layers can be reflected by measurements of piezometric level in the shallow aquifer. From a seismical point of view, the problem to characterise the increase of seismic constraints from such measurements data is open.

One restricts the present study to this subsurface aquifer and, for an easier presentation, we deal with 2D flow problem in a vertical section of aquifer. The aquifer is supposed continuous, saturated, homogeneous, isotropic and incompressible. Moreover, one supposes that it is not crossed by the fault, but that the fault always remains below it, in such a way that the aquifer/fault interface (noted Γ_L) is always located on the lower border of the aquifer (Γ_{inf}) as shown in Fig. 1.

In agreement with the preceding scenario, the consequences of a deeper seismic activity for the flow in the upper aquifer would be the increase of flux entering by the part Γ_L of inferior border, as well as the widening of this interface [1]. The flux value is supposed constant on Γ_L and the aquifer/fault interface is supposed symmetrical with respect to the vertical fault axis: $\Gamma_L = [-L, L] \times \{0\}$. Apart from Γ_L, the flux is null on the lower border of the domain.

The superior boundary condition is defined by assuming that the aquifer is confined (null flux is imposed). Let us suppose that horizontal dimension of the aquifer is sufficiently large so that the piezometric level has a constant and known value "far" from the fault (this value, noted h_D, defines Dirichlet condition on lateral borders Γ_D).

Then the inverse problem consists in determining the flux of fluid entering in aquifer by Γ_L, as well as the interface width, from the knowledge of the piezometric level in a subset $\Omega^{\text{obs}} = \{P_i,\ i = 1, \ldots, M\}$ of measurement points in the aquifer domain Ω. The observation is denoted by $\mathbf{h}^{\text{obs}} = \{h_1^{\text{obs}}, \ldots, h_M^{\text{obs}}\}$.

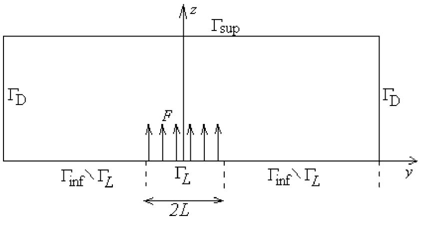

Fig. 1. Neumann condition reflecting the presence of a fault below the subsurface aquifer.

Fig. 1. Condition de Neumann reflétant la présence d'une faille sous l'aquifère de surface.

3. Direct and inverse problems

The direct problem consists in solving the flow problem in Ω, for boundary conditions explained above. We recall that the piezometric level h is the variable which defines the flow in the aquifer [2] and it is related to the intrinsic pressure of the fluid by $P = \rho g(h - z)$, where ρ is the fluid density and g is gravitational acceleration.

Using Darcy's law for homogeneous and isotropic media, and mass conservation equation for incompressible fluids, one obtains a Laplace equation for h. Therefore, taking into account boundary conditions, the piezometric level in Ω is given by the boundary problem:

$$\begin{cases} \Delta h = 0, & \Omega, \\ h(y,z) = h_D, & (y,z) \in \Gamma_D, \\ (\nabla h)(y,z) \cdot \mathbf{n} = 0, & (y,z) \in \Gamma_{\text{sup}}, \\ (\nabla h)(y,z) \cdot \mathbf{n} = F \times \chi_{\Gamma_L}(y,z), & (y,z) \in \Gamma_{\text{inf}}, \end{cases} \tag{1}$$

where χ_{Γ_L} is the characteristic function of Γ_L in Γ_{inf}.

Obviously, the direct problem (1) is well posed if values of F and L are given and then, to indicate that the solution of the problem (1) is associated to these values, it is noted $h_{F,L}$. Using this notation, we define the inverse problem which has as unknown the couple of parameters (F, L) and which uses as additional data the piezometric level observation $\mathbf{h}^{\text{obs}}$. The inverse problem is formulated as follows:

$$\text{To find the couple } (F^*, L^*) \in \mathbf{R}^2 \text{ such as} h_{F^*,L^*}|_{\Omega^{\text{obs}}} = \mathbf{h}^{\text{obs}}. \tag{2}$$

From a physical point of view, the flow is generally directed from fault towards the aquifer and it is limited between a minimal value and a maximum value: $0 \leqslant F_m \leqslant F \leqslant F_M$. In the same way, L can vary between a minimal value $L_m > 0$ and a maximum value $L_M < a$, where $2a$ denotes the width of Ω. Then, we define the set of admissible parameters as: $P_{\text{ad}} = [F_m, F_M] \times [L_m, L_M]$. Therefore, an objective function can be introduced in the same manner than in Section 1 and the OLS formulation corresponding to (2) consists in minimizing the objective function on P_{ad}.

4. Properties of direct application

Firstly, we study linearity properties of direct application, which is given by:

$$D : P_{\text{ad}} \to \mathbf{R}^M, \quad D(F, L) = \big(h_{F,L}(P_1), \ldots, h_{F,L}(P_M)\big).$$

Proposition 4.1. *The application D is linear with respect to F, but is nonlinear with respect to L.*

Proof. Let us introduce the functions A_L and B, defined as solutions of problems (P_{A_L}) and (P_B) :

$$(P_{A_L}) \quad \begin{cases} \Delta A_L = 0 & \text{in } \Omega, \\ A_L = 0 & \text{on } \Gamma_D, \\ \nabla A_L \cdot \mathbf{n} = 0 & \text{on } \Gamma_{\text{sup}}, \\ \nabla A_L \cdot \mathbf{n} = \chi_{\Gamma_L} & \text{on } \Gamma_{\text{inf}}, \end{cases} \qquad (P_B) \quad \begin{cases} \Delta B = 0 & \text{in } \Omega, \\ B = h_D & \text{on } \Gamma_D, \\ \nabla B \cdot \mathbf{n} = 0 & \text{on } \Gamma_{\text{sup}} \cup \Gamma_{\text{inf}}. \end{cases}$$

Then, functions $h_{F,L} - B$ and $A_L \cdot F$ are solutions of the problem (1) with homogeneous Dirichlet condition. Because this problem has an unique solution, one obtains the linearity relation:

$$h_{F,L}(y,z) = A_L(y,z) \cdot F + B(y,z). \tag{3}$$

In addition, by applying a maximum principle for the problem (P_{A_L}), one can deduce the positivity of A_L and then D is an increasing function with respect to F.

To show that D is nonlinear with respect of L, let us suppose that, for a fixed value L_0, there are two functions A_1 and B_1 that do not depend of L, so that $h_{F,L}$ can be written:

$$h_{F,L}(y,z) = A_1(y,z) \cdot (L - L_0) + B_1(y,z).$$

One obtains immediately $B_1 = h_{F,L_0}$. To set one's ideas let us consider a value $L > L_0$. Then A_1 must be solution of a boundary problem similar with (P_{A_L}), but with Neumann data $\frac{F}{L-L_0}\chi_{\Gamma_L\setminus\Gamma_{L_0}}$ on Γ_{inf}. Then a contradiction is obtained since the solution of this problem still depends of L. □

Dealing with the properties of regularity of the direct application, the continuity with respect to F can be directly deduced from linearity relation (3). For the continuity with respect to L, it is easy to see that $h_{F,L+\varepsilon} \to h_{F,L}$ when $\varepsilon \to 0$. To study the differentiability of the direct application, consider $P \in \Omega^{\mathrm{obs}}$ and analyse the direct application restricted at this point, defined by $D_P : P_{\mathrm{ad}} \to \mathbf{R}$, $D_P(F, L) = h_{F,L}(P)$.

Theorem 4.2. *The application D_P is C^1.*

Proof. The linearity relation (3) also signifies the derivability of D_P with respect to F. In addition, it gives the expression of partial derivative:

$$\frac{\partial D_P}{\partial F}(F, L) = A_L(P). \tag{4}$$

F being given, consider the application $L \to f(L) \equiv D_P(F, L)$. If (y_P, z_P) are coordinates of point P, the fundamental solution of Laplace equation, corresponding to this point is given by:

$$\varphi : \mathbf{R}^2 \setminus \{P\} \to \mathbf{R}, \quad \varphi(y, z) = -\frac{1}{4\pi} \ln\big((y - y_P)^2 + (z - z_P)^2\big).$$

Using this function, which is $C^\infty(\mathbf{R}^2 \setminus \{P\})$ and verifies $\Delta\varphi = 0$ in this domain, we define g as the solution of the boundary problem:

$$\begin{cases} \Delta g = 0 & \text{in } \Omega, \\ g = -\varphi & \text{on } \Gamma_D, \\ \nabla g \cdot \mathbf{n} = -\nabla\varphi \cdot \mathbf{n} & \text{on } \Gamma_{\mathrm{inf}} \cup \Gamma_{\mathrm{sup}}. \end{cases}$$

Finally, we define the function $G : \Omega \setminus \{P\} \to \mathbf{R}$, by $G = g + \varphi$. Obviously, functions φ, g, G depend of P, but, for the sake of simplicity, this dependance is not written. One can observe that, by construction, the function G verifies: $\Delta G = 0$ in $\Omega \setminus \{P\}$, $G = 0$ on Γ_D, $\nabla G \cdot \mathbf{n} = 0$ on $\Gamma_{\mathrm{inf}} \cup \Gamma_{\mathrm{sup}}$.

Let us consider the ball $B_{P,\varepsilon} \subset \Omega$ with center in P and with radius $\varepsilon > 0$ and write Green's identity for functions $h_{F,L}$ and G in $\Omega \setminus B_{P,\varepsilon}$, denoting by n the normal unit vector to this domain :

$$\int_{\Omega\setminus B_{P,\varepsilon}} (h_{F,L} \cdot \Delta G - G \cdot \Delta h_{F,L})\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega\cup\partial B_{P,\varepsilon}} \left(h_{F,L} \cdot \frac{\partial G}{\partial n} - G \cdot \frac{\partial h_{F,L}}{\partial n}\right)\mathrm{d}\sigma.$$

Since G and $h_{F,L}$ are harmonic functions, left-hand term is null. As usual for boundary integral representations with Greens's functions, one uses properties of functions $h_{F,L}$ and G to express the right-hand term. Finally, after passing to the limit for $\varepsilon \to 0$, we deduce:

$$f(L) \equiv h_{F,L}(P) = F \int_{-L}^{+L} G(y, 0)\,\mathrm{d}y - \int_{\Gamma_D} h_D \cdot \frac{\partial G}{\partial n}\,\mathrm{d}\sigma.$$

To study the derivability of f, one uses previous equation and then, for all $h > 0$, one obtains:

$$\frac{f(L+h) - f(L)}{h} = F \cdot \left[\frac{1}{h} \int_{-(L+h)}^{-L} G(y, 0)\,\mathrm{d}y + \frac{1}{h} \int_{L}^{L+h} G(y, 0)\,\mathrm{d}y\right].$$

Passing to the limit for $h \to 0$, we deduce:

$$\lim_{h\to 0,\ h>0} \frac{f(L+h)-f(L)}{h} = F \cdot \bigl[G(-L,0)+G(L,0)\bigr]$$

the same limit being also obtained for $h < 0$.

In conclusion, for fixed F, application $L \to D_P(F,L)$ is derivable and we get:

$$\frac{\partial D_P}{\partial L}(F,L) = F \cdot \bigl[G(-L,0)+G(L,0)\bigr]. \tag{5}$$

In addition, the relation (5) shows that for a fixed value of $F > 0$, the function $L \to h_{F,L}(P)$ is also increasing (by construction, G is strictly positive on its definition set except Γ_D). □

5. Numerical implementation

As it is shown in Section 1, to solve minimization problem for OLS criterion by Gauss–Newton method, only the sensitivity matrix corresponding to (F^k, L^k) is needed (apart of calculation of residual vector $\boldsymbol{\Phi}(F^k, L^k)$) to construct a new approximation (F^{k+1}, L^{k+1}) of solution (F^*, L^*), at iteration k.

The derivatives with respect to F can be computed by solving problems (P_{A_L}) for $L = L^k$ and (P_B) and then using linearity relation (3). We note that, since (P_B) does not depend of F and L, it can be solved before starting the iterative algorithm. The same remark yields for calculating the derivatives with respect to L, using formula (5). Indeed, the function G constructed in the demonstration of Theorem 4.2 does not depend of F and L, but it depends of observation point and then we must construct functions G for each point $P_i,\ i = 1, \ldots, M$, before starting the optimization algorithm.

Consequently, in this algorithm, an important part of gradient computation cost is reduced since, at each iteration, a single problem (P_{A_L}) for $L = L^k$ is to be solved. In an other hand, we note that using (3), A_{L^K} can be computed as soon as h_{F^k,L^k} is known (B being already computed).

In conclusion, for each iteration of inverse problem algorithm, the computation cost is only the one of single direct problem, that is obviously optimal. However, we must note that if CPU time for such an algorithm is minimal, it requires some computer memory to stock values of function B restricted to Ω^{obs} and especially, values of M functions G (or g) restricted to $\{(y,0),\ y \in [-L_M, -L_m] \cup [L_m, L_M]\}$. Indeed, a good value of derivatives with respect with L can be determined only if the values of G functions for current L^k are available (this means that the values of g functions have been previously determined and stored with low errors, by numerically solving the corresponding boundary problems).

6. Conclusions

A physical problem whose origin is the seismic has been described. In such a problem one is interested to identify two scalar parameters (F and L) which define a Neumann condition. It is shown that derivatives of associated direct application with respect to the two parameters can be expressed by explicit formulas (Eq. (4) for the derivative in F and Eq. (5) for the derivative in L). The use of these formulas can considerably improve the computational cost and precision of numerical optimization algorithm used to solve the inverse problem. We note that a similar analysis can be made for a constant piecewise flux on Γ_L instead of constant flux.

References

[1] D.G. Calugaru, J.M. Crolet, A. Chambaudet, Radon transport as an indicator of seismic activity. An algorithm for inverse problems, Developm. Water Sci. 47 (2002) 631–638.

[2] G. De Marsily, Hydrogéologie quantitative, Masson, Paris, 1981.

[3] N. Point, A.V. Wouwer, M. Remy, Practical issues in distributed parameter estimation: gradient computation and optimal experiment design, Control Eng. Practice 4 (11) (1996) 1553–1562.

[4] N.Z. Sun, Inverse Problemes in Groundwater Modeling, Kluwer Academic, 1999.