Buletinul Ştiinţific al Universităţii „POLITEHNICA" din Timişoara
Seria HIDROTEHNICA
TRANSACTIONS on HYDROTECHNICS
Tom 57(71), Fascicola 2, 2012

Gradient methods in parameter estimation by optimization algorithms in groundwater models

Robert F. Beilicci[1], Erika Beilicci[1], Stefanescu Camelia[1]

Abstract – This application solve analytically and with optimization algorithms a function minimization using the gradient method. To build a model for a real groundwater system it is necessary to solve both the forward problem and a so called inverse problem. Two optimization algorithms methods exist: indirect and direct method. In this paper only the indirect method is used to solve actual problems.

Keywords: Gradient methods, parameter estimation, optimization algorithms, inverse problem

I. INTRODUCTION IN FORWARD AND INVERSE PROBLEMS IN GROUNDWATER STUDIES

To build a model for a real groundwater system it is necessary to solve both the forward problem and a so called inverse problem. So far we have only been dealing with the forward, i.e. simulation problem. In the forward problem we predict the unknown heads by solving either steady-state or transient equations assuming that the parameter values, control variables and boundary conditions are known. In the inverse problem we have to determine unknown physical parameters by fitting the model to observed heads. According to Sun (1994), studies on these two problems are not in balance. The study of forward problem has developed rapidly but the study of inverse problems is still limited to very simple models as compared with the complexity of the models used for forward simulations. The book by Sun (1994) is the first book on the subject of solving the inverse problem in groundwater studies.

The progress of inverse solution techniques is blocked by three main difficulties.

First, the inverse problem is often ill-posed, i.e., its solution may be non-unique and unstable with respect to the observation error.

Second, the quantity and quality of observation data are usually insufficient.

Third, the model structure error, which is difficult to estimate, often dominates the error.

In two-dimensional groundwater models the key parameters to be estimated based on all information available are the transmissivity T for confined aquifer or hydraulic conductivity K for unconfined aquifer, storage coefficient S and recharge R from precipitation. Usually it is not advisable to include R as a fitting parameter but to try to estimate R based on precipitation, evapotranspiration and surface runoff calculations, i.e. calculate the water balance of the soil column above the groundwater level

The very complex and difficult problem of deciding if the inverse solution exists, if it is unique and stable, and if the model is identifiable, are briefly discussed here.

A well-posed mathematical problem must satisfy the following requirements: Existence, uniqueness and stability. Basically, the existence of an inverse solution seems to be no problem at all, since the physical reality must be a solution. However, the observation error of the head values cannot be avoided and therefore, it is possible that an accurate solution of the inverse problem may not exist. However, existence is not the major difficulty in the solution of inverse problems. Even if completely accurate solution may not exist, it is possible to determine the parameters in such a way that the square sum between the calculated and observed values is minimized.

Usually the major difficulty is to find unique set of parameter values. Different combinations of hydro geological conditions may lead to similar observations of water level. In this type of case it is impossible to uniquely determine the parameters of an aquifer only by observing the hydraulic head values, i.e. non-uniqueness of the inverse solution is often observed. This can be formulated in another way by noticing that different type of parameter combinations may lead to completely same values for calculated heads.

An example of the case when no unique solution can be obtained. Consider a one-dimensional state-state problem in confined aquifer by

$$\frac{d}{dx}\left[T(x)\frac{dH}{dx}\right]=0 \qquad x_1 \le x \le x_2 \qquad (1)$$

With Dirichlecht-type boundary conditions $H(x_1) = H_1$ and $H(x_2) = H_2$.

The necessary condition for the solution of the inverse problem is that we have measurements of H(x) and consequently, we can assume that the derivative $dH/dx = H'(x)$ is known.

Now it is straightforward to integrate (1) to yield

$$T(x) = \frac{C}{H'(x)} \qquad (2)$$

Where C is an arbitrary integration constant.

Therefore, T(x) is not unique, although the head is observed at interval (x_1,x_2).

[1] "Politehnica" University of Timisoara, Faculty of Civil Engineering, Department of Hydrotecnnics, Timisoara, George Enescu Str., No. 1/A, Timisoara, Romania, E-mail: robert.beilicci@hidro.upt.ro

This kind of non-uniqueness cannot be removed by increasing the number of observation wells. As a general conclusion it can be stated that if Dirichlecht-type boundary conditions are used in steady-state solution all over the aquifer boundaries, it is usually not possible to obtain a unique solution to inverse problems unless there exist some flux boundaries or additional sinks/sources in the aquifer area. E.g. if the left boundary in the previous example is replaced by a given flux boundary condition T(dH/dx) at $x=x_1$ is -q, it is possible to obtain a solution

$$T(x) = \frac{-q}{H'(x)} \tag{3}$$

Which implies that T(x) is now uniquely determined. Therefore, it is crucial to recognize the influence of the boundary conditions on the existence of a unique solution to inverse problems.

Usually, it is also possible to obtain a unique solution to inverse problems if additional information on the parameter values can be supplemented. In the previous example it would imply a measurement of transmissivity T(x) at any point between x_1 and x_2. Moreover, if dynamic measurements are available, i.e. H(x,t) is known, it is possible to have a unique solution. Unfortunately, inverse solutions in groundwater modeling are often unstable and the main reasons for this are the head observation errors. This type of solution is said to be ill-posed

II. PARAMETER ESTIMATION BY OPTIMIZATION ALGORITHMS

Two main solution methods exist: indirect and direct method. In this paper only the indirect method is used to solve actual problems

The trial-and-error procedure is the simplest way to solve the inverse groundwater problem. In this method we need some hydraulic head measurements, a model that calculates the forward problem (simulation) and a person (e.g. hydro geologist) who is familiar with the considered aquifer. The flow chart of the trial-and-error procedure is given in Fig. 1.

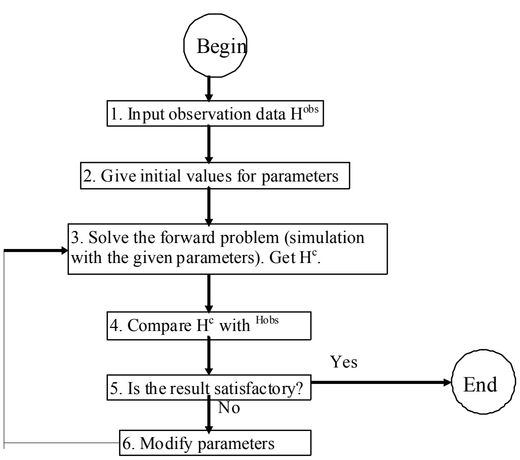

Fig. 1. Flowchart of the trial -and-error procedure

The trial-and-error procedure described in the previous section can be replaced by a computer program which transfers inverse problems into optimization problems.

Referring to Fig. 1, we need to see what steps can be completed by computer. To replace Step 4, it is possible to calculate a criterion to measure the difference between observed and calculated heads.

The most common criterion is the Output Least Squares OLS) defined by

$$E(\bar{p}) = \sum_{m=1}^{M} W_m^2 \left[H_m^c(\bar{p}) - H_m^{obs} \right]^2 \tag{4}$$

Where $H_m^c(\bar{p})$ is the calculated head and H_m^{obs} is the observed head, M is total number of measurements and Wm is a weight associated to measurement m and $\bar{p}$ is the parameter vector and the components $(p1\ p2\ ...\ p_N)^T$ need to be estimated. If $E(\bar{p}_2) < E(\bar{p}_1)$ then parameter vector $\bar{p}_2$ is better than $\bar{p}_1$.

Step 5 of Fig. 1. can be replaced by checking that either the error $E(\bar{p})$ is smaller than a prescribed convergence criteria ε or that the successive parameter values do not differ from each other more than a prescribed amount.

Step 6 is the crucial step in the solution of the inverse method and the purpose is to solve the following optimization problem: find the specific set of parameters $\bar{p}_{OPT}$ such that

$$E(\bar{p}_{OPT}) = \min E(\bar{p}) \tag{5}$$

The summary of the indirect method is given in Fig. 2. As shown in Fig. 2, the forward problem is solved several times. In other words, the inverse problem is solved indirectly through the solution of the forward problem.

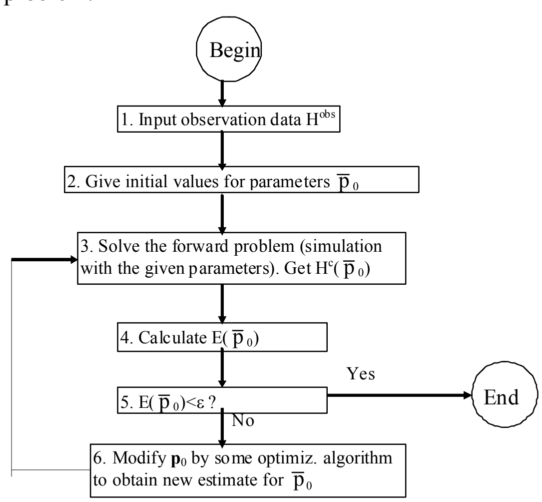

Fig. 2. Flowchart of the indirect method

The advantage of the indirect method is that inverse problems can be solved rapidly by computer without human participation. Moreover, a set of "best fitting" parameters can be found by solving the optimization problem. However, there is always a danger that the optimization method does not find a global but a local optimum.

III. FORMULATION OF THE NON-LINEAR OPTIMIZATION PROBLEM

Consider the N-dimensional optimization problem (number of parameters to be optimized is N):

$$\min E(\overline{p}) \qquad (6)$$

Where the objective function $E(\overline{p})$.

If function $E(\overline{p})$ is second-order differentiable, the following are necessary conditions for $\overline{p}$ being a local minimum of $E(\overline{p})$:

- First derivate (gradient g= $\nabla E(\overline{p})$) vanishes at $\overline{p}$

$$\left.\frac{\partial E}{\partial p_n}\right|_{\hat{p}} = 0 \qquad (n=1,...,N) \qquad (7)$$

- Hessian matrix $G=\nabla^2 E(\overline{p})$ (second derivate of E with respect to $\overline{p}$) is a positive semi-definite matrix

$$G=\begin{bmatrix} \frac{\partial^2 E}{\partial p_1^2} & \frac{\partial^2 E}{\partial p_1 \partial p_2} & \cdots & \frac{\partial^2 E}{\partial p_1 \partial p_N} \\ \frac{\partial^2 E}{\partial p_1 \partial p_2} & \frac{\partial^2 E}{\partial p_2^2} & \cdots & \frac{\partial^2 E}{\partial p_2 \partial p_N} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \frac{\partial^2 E}{\partial p_1 \partial p_N} & \frac{\partial^2 E}{\partial p_2 \partial p_N} & \cdots & \frac{\partial^2 E}{\partial p_N^2} \end{bmatrix} \qquad (8)$$

If $E(\overline{p})$ is a differentiable convex function, then $\nabla E(\overline{p}\hat{})=0$ is the necessary and sufficient condition for $\overline{p}\hat{}$ being a local minimum of $E(\overline{p})$. In the identification of model parameters, objective function $E(\overline{p})$ depends on model output and therefore we cannot get an explicit expression for $\nabla E(\overline{p})$ and equation cannot be solved directly. For practical optimization problems, $\nabla E(\overline{p})$ has to be obtained numerically and the solution of the problem is iterative:

Step 1) Choose initial guess $\overline{p}_0$.

Step 2) Designate a way to generate a search sequence:

$$\overline{p}_0, \overline{p}_1, \overline{p}_2, ..., \overline{p}_k \qquad (9)$$

such that $E(\overline{p}_{k+1}) < E(\overline{p}_k)$ for all iteration index k.

Step 3) Check the convergence and if it is satisfied, then end the search procedure, and a local minimum is approximately achieved.

Search sequence has a general form:

$$\overline{p}_{k+1}=\overline{p}_k+s_k d_k \qquad (10)$$

Where s_k is a step size along a direction that is called displacement direction d_k.

The key problem of the parameter optimization procedure is how to determine s_k and d_k.

Three different methods for solving this problem various optimization algorithms can be divided into three main categories:

1) An optimization is called a search method if it only utilizes values of the objective functions (downhill simplex method);
2) An optimization algorithm is called a gradient method if it utilizes gradients of objective functions;
3) An optimization algorithm is called a second order method if it utilizes second derivatives of objective functions;

IV. GRADIENT METHODS

The basic principle of the gradient methods is to use the negative gradient direction as the search direction in each iteration. In other words: the function $E(\overline{p})$ to be minimized is most rapidly reduced in the direction of its negative gradient. This implies that d_k of Eq. (10) is replaced by

$$d_k = -g_k \qquad (11)$$

where $g_k = \nabla E(\overline{p}k)$.

The optimal step s_k of Eq. (10) could be obtained by performing a line search in one direction The gradient method and a simple method to determine s_k is described in detail in Example.

IV. EXAMPLE

The problem is to minimize a two-variable function

$$F(x,y) = x^2 + 2y^2 - 4x - 4y +6 \qquad (12)$$

The derivatives of F(x,y) with respect to $F_x(x,y)$ and $F_y(x,y)$ are

$$F_x(x,y) = 2x - 4 \qquad (13)$$

$$F_y(x,y) = 4y - 4 \qquad (14)$$

The Hessian matrix is now

$$G=\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \qquad (15)$$

Analytically it is easy to solve $\nabla F(x\hat{},y\hat{}) = 0$ to yield $x\hat{} = 2$ and $y\hat{} = 1$ and calculate that $\min F(x\hat{},y\hat{})=0$.

In this Example the goal is to use iterative gradient method for solving the same problem. As initial guess for x and y the following values are used: $x_0 = y_0 = -1$. The optimization procedure is shown in the file in Table 1.

The procedure is described step by step in Frame 1.

The procedure shown in Frame 1 and Table 1 converges to the exact solution in 10 iterations if the initial step size s_0 is 2.

	A	B	C	D	E	F	G	H	I	J	K
1	Iter	Xk	Yk	F	F'x	F'y	dF	Sk	Xk+1	Yk+1	Fk+1
2	0	-1	-1	17	-6	-8	10	2	0.2	0.6	3.56
3	1	0.2	0.6	3.56	-3.6	-1.6	3.93954	2	2.028	1.41228	0.340708
4	2	2.0276	1.412	0.34071	0.055246	1.649108	1.65003	0.5	2.011	0.91256	0.015411
5	3	2.0109	0.913	0.01541	0.021764	-0.349771	0.35045	0.125	2.003	1.03732	0.002795
6	4	2.0031	1.037	0.00279	0.006238	0.149264	0.14939	0.06	2.001	0.97737	0.001025
7	5	2.0006	0.977	0.00102	0.001227	-0.090527	0.09054	0.03	2	1.00737	0.000109
8	6	2.0002	1.007	0.00011	0.000414	0.029462	0.02947	0.008	2	0.99987	4.61E-08
9	7	2.0001	1	4.6E-08	0.000203	-0.000535	0.00057	2E-04	2	1.00009	1.49E-08
10	8	2	1	1.5E-08	3.67E-05	0.000342	0.00034	1E-04	2	0.99997	1.97E-09

Table 1 Function minimization using the gradient method

0) Set iteration index k=0 to cell A2 and initial estimate for x to cell B2 and for y to cell C2, respectively.
1) Use the initial estimate shown in cells B2 and C2 to calculate $F(x_k,y_k)$ in cell D2
2) Calculate derivative of F(x,y) with respect to x, $F_x(x_k,y_k)$, and with respect to y, $F_y(x_k,y_k)$. The results are shown in cells E2 and F2.
3) Calculate the lenght of the gradient vector to cell G2:

$$|F| = \sqrt{\left[F_x(x_k, y_k)\right]^2 + \left[F_x(x_k, y_k)\right]^2}$$

4) Choose estimate for step size s_m in Eq. (5-7) (cell H2).
5) Calculate new estimate for unknown parameters x and y using Eq. (9):

$$x_{k+1} = x_k - s_k \frac{F_x(x_k, y_k)}{|F|} \quad \text{or in EXCEL: I2 = B2 - H2*(E2/G2)}$$

$$y_{k+1} = y_k - s_k \frac{F_y(x_k, y_k)}{|F|} \quad \text{or in EXCEL: J2 = C2 - H2*(F2/G2)}$$

6) Calculate function value $F(x_{k+1},yk_{+1})$ to cell K2
7) Compare if $F(x_{k+1},yk_{+1}) < F(x_k,yk)$, i.e. if K2 < D2? If this is true then accept (x_{k+1},yk_{+1}) and go to step 8). If $F(x_{k+1},yk_{+1}) > F(x_k,yk)$ then reduce step in H2 by 50 % and go back to step 5).
8) Test if $F(x_{k+1},yk_{+1})$ smaller than iteration stopping criteria ε (in this example $\varepsilon = 10^{-8}$)?
   If $F(x_{k+1},yk_{+1}) < \varepsilon$, accept (x_{k+1},yk_{+1}) as the final solution of the problem.
   If $F(x_{k+1},yk_{+1}) > \varepsilon$, use (x_{k+1},yk_{+1}) as the initial value for the next iteration, set k = k+1 and go to 1). In EXCEL-solution (x_{k+1},yk_{+1}) from cells I2 and J2 and $F((x_{k+1},yk_{+1}))$ from cell K2 are copied to the next row to cells B3, C3 and D3, respectively.

Frame 1 Step by step solution of function F(x,y) minimization using the gradient method. EXCEL-cells shown in Table 1

The drawback of the gradient method is that if the initial estimate is too far from the true solution, the method terminates far from the solution due to round-off errors. However, the basic principles of the optimization methods are very well presented in this example and the reader is encouraged to go through the calculations very carefully.

V. CONLUSIONS

The trial-and-error procedure is the simplest way to solve the inverse groundwater problem. In this method we need some hydraulic head measurements, a model that calculates the forward. The trial-and-error procedure can be replaced by a computer program which transfers inverse problems into optimization problems. The advantage of the indirect method is that inverse problems can be solved rapidly by computer without human participation.
The basic principle of the gradient methods is to use the negative gradient direction as the search direction in each iteration.

REFERENCES

[1] Nelder, J.A. and Mead, R. (1965). Computer Journal, vol 7., p. 308.
O'Neill, R. (1971). Function minimization using the simplex procedure.

[2] Pinder, G.F. and Gray, W.G. (1977). Finite element simulation in surface and subsurface hydrology. Academic Press, New York.

[3] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988). Numerical recipes in C. The art of scientific compuring. Cambridge University Press, Cambridge.

[4] Remson, I., Hornberger, G.M. and Molz, F.J. (1971). Numerical methods in subsurface hydrology. Wiley-Interscience, New York.

[5] Rushton, K.R. and Redshaw, S.C. (1979). Seepage and groundwater flow. Wiley, New York.

[6] Sun Ne-Zheng (1994). Inverse problems in groundwater modeling. Kluwer Academic Publishers. London.

[7] Toth, J. (1962). A theory of groundwater motion in in small drainage basins in Central Alberta, Canada. Journal of Geophysical Research, 67(11): 4375-4387.

[8] Wang, H.F. and Anderson, M.P. (1982). Introduction to groundwater modeling: Finite difference and finite element methods. W.H. Freeman and Company, San Fransisco.