A Continuation Method of Parameter Inversion for Non-Equilibrium Convection–Dispersion Equation *

CUI Kai(崔凯)**, YANG Guo-Wei(杨国伟)

Laboratory of High-Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100080

(Received 13 April 2005)

Based on the homotopy mapping, a globally convergent method of parameter inversion for non-equilibrium convection-dispersion equations (CDEs) is developed. Moreover, in order to further improve the computational efficiency of the algorithm, a properly smooth function, which is derived from the sigmoid function, is employed to update the homotopy parameter during iteration. Numerical results show the feature of global convergence and high performance of this method. In addition, even the measurement quantities are heavily contaminated by noises, and a good solution can be found.

PACS: 02.30.Zz, 02.60.Pn, 47.55.Mh

Over recent years, the study of solute transport in porous media has attracted considerable attention.[1,2] The non-equilibrium convection-dispersion equation (CDE) has been successfully used to describe the procedure of many kinds of solute transport in porous media, such as heavy metal ion transport through unsaturated soils.[3] The dimensionless form of the one-dimensional non-equilibrium CDEs can be expressed as follows:[4]

$$K\rho\frac{\partial C}{\partial t}+\frac{\partial(\theta C)}{\partial t}+\rho K_A\gamma(1-S)C-\rho K_B\gamma S =\frac{\partial}{\partial Z}\Big(\alpha q\frac{\partial C}{\partial Z}\Big)-\frac{\partial(qC)}{\partial Z}, \tag{1}$$

$$\frac{\partial S}{\partial t}=K_A(1-S)C-K_BS, \tag{2}$$

where C and S are the dimensionless solute concentrations associated with the solution and solid phases of the soil, K is the sorption distribution coefficient, ρ is the bulk density of the solid phases, θ is the volumetric water content, K_A and K_B represent the precipitating rate and the dissolving rate respectively, γ is the relative maximum sorption capacity, α is the mechanical dispersion coefficient, and q is Darcy's velocity.

Under appropriate initial and boundary conditions, Eqs. (1) and (2) can be solved by some numerical methods, such as the finite element method (FEM) or the finite difference method (FDM), and then the concentration distribution of the solute in porous media can be obtained. Obviously, the capability of the equation to simulate reality deeply depends on the possibility that determines the values of the governing parameters with accuracy. However, some parameters, which include α, K, K_A, K_B, and γ, are in general not directly measurable, so they have to be estimated by parameter inversion on the basis of the available observed data. The discrete form of this inversion problem is generally formulated to the following nonlinear optimization problem:

$$\min J(\boldsymbol{p})=\frac{1}{2}\|\boldsymbol{C}(\boldsymbol{p})-\boldsymbol{C_E}\|^2, \tag{3}$$

in which $\boldsymbol{p}=[\alpha,K,K_A,K_B,\gamma]$ is a vector whose components are unknown parameters, $\boldsymbol{C}(\boldsymbol{p})$ and $\boldsymbol{C_E}$ are the calculated and observed solute concentrations at the observation position respectively.

There are mainly two kinds of methods to solve the inverse problem for Eq. (3). One are the so-called heuristic algorithms and the other are the gradient-based methods. Both methods have their advantages and disadvantages. The former, such as the Genetic Algorithms (GA),[5,6] belong to the random search methods, in which the initial solution need not be carefully selected, and the global solution can also be obtained. However, these methods usually consume much computation time. In contrast, the gradient-based methods, such as the Levenberg–Marquardt method,[7–9] which use the gradient of the objective function during iteration, are generally more effective than the heuristic algorithms. However, they are locally convergent. In other words, an improper initial solution may lead to a divergent iteration process.

The present status naturally motivates us to develop a global convergence method with high performance for this problem. In this Letter, a homotopy-based continuation method is proposed. Homotopy is an important concept of topology. Watson *et al.*[10–12] developed a series of global convergence methods for solving the Brouwer fixed-point problems, nonlinear systems of equations, nonlinear optimization problems and other problems by employing homotopy mappings. To our problem, by applying the optimality

* Supported by the K C Wang Education of Foundation of Hong Kong, and the National Natural Science Foundation of China under Grant Nos 10402043 and 10372106.

** Email: kcui@imech.ac.cn

condition, an equivalent system of Eq. (3) can be obtained in the form of

$$\left[\frac{\partial \boldsymbol{C}(\boldsymbol{p})}{\partial \boldsymbol{p}}\right]^T [\boldsymbol{C}(\boldsymbol{p}) - \boldsymbol{C}_{\boldsymbol{E}}] = 0. \tag{4}$$

To simplify the representation, let $\boldsymbol{G} = \partial \boldsymbol{C}(\boldsymbol{p})/\partial \boldsymbol{p}$ and $\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{p})$.

The homotopy map of Eq. (4) is constructed as

$$\begin{aligned}\boldsymbol{H}(\boldsymbol{p}, \lambda) &= (1-\lambda)[\boldsymbol{G}^T\boldsymbol{C} - \boldsymbol{G}^T\boldsymbol{C}_{\boldsymbol{E}}] + \lambda(\boldsymbol{p} - \boldsymbol{p}^0) \\ &= 0, \quad \lambda \in [0, 1],\end{aligned} \tag{5}$$

where λ is an embedding parameter, i.e. homotopy parameter; $\boldsymbol{p}^0$ is an initial estimation of solution. Obviously, when $\lambda = 1$, the solution of Eq. (5) can be obtained immediately, i.e. $\boldsymbol{p} = \boldsymbol{p}^0$. When $\lambda = 0$, Eq. (5) transforms to Eq. (4). Therefore, if the homotopy parameter can track through a proper path from one to zero, the solution of original problems of Eq. (4) can be obtained. This is the basic idea of the homotopy methods. As λ changes, by using Taylor series expansion at the nth iteration step, the linearized version of formulation (5) can be obtained to be

$$(1-\lambda^n)\{\boldsymbol{G}^T[\boldsymbol{C}^n + \boldsymbol{G}(\boldsymbol{p}-\boldsymbol{p}^n)] - \boldsymbol{C}_{\boldsymbol{E}}\} + \lambda^n(\boldsymbol{p}-\boldsymbol{p}^n) = 0, \tag{6}$$

where λ^n and $\boldsymbol{C}^n = \boldsymbol{C}(\boldsymbol{p}^n)$ are the homotopy parameter and the calculated concentration at the n-th iteration step respectively. Let $\Delta\boldsymbol{p} = \boldsymbol{p} - \boldsymbol{p}^n$, then Eq. (6) can be rewritten as

$$\begin{aligned}&[(1-\lambda^n)\boldsymbol{G}^T\boldsymbol{G} + \lambda^n\boldsymbol{I}]\Delta\boldsymbol{p} = (1-\lambda^n)\boldsymbol{G}^T(\boldsymbol{C}^n - \boldsymbol{C}_{\boldsymbol{E}}), \\ &\boldsymbol{p}^{n+1} = \boldsymbol{p}^n + \Delta\boldsymbol{p},\end{aligned} \tag{7}$$

in which $\boldsymbol{I}$ is the unit matrix.

The strategy of determining the homotopy parameter λ^n in Eq. (7) is very important. Generally, it can be determined by solving a series of ordinary difference equations (ODEs),[7−10] but this ODE-based method will not be appropriate to our problem. As is well known, the observed quantities will be inevitably contaminated by observation errors; by using the singular value decomposition[13] analysis, we can know if a proper value of λ, but not zero, was chosen at the end of iteration, and the observation noises can be restrained effectively. Hence there are two criteria for updating the homotopy parameter that should be abided by. First, in order to ensure a stable iteration process, λ should be diminished along a smooth path during iteration. Second, to restrain the measurement errors, λ should be terminated at a proper point, which may be close to zero, but should be unequal to zero. Here we propose a function, which is derived from the sigmoid function of the neural networks, as follows:

$$\lambda^n = \frac{1}{1 + e^{\beta(n-N_0)}}, \tag{8}$$

where n is the number of times of iteration, β and N_0 are the adjustable parameters. It is clear that this function which we call the quasi-sigmoid function could naturally meet the two requirements mentioned above. Parameter N_0 is mainly used to ensure the stability of computation at the early stage of iteration, while β can balance the stability and the efficiency of computation. The lower value of N_0 and β can accelerate iteration, but may lead to a divergent process. In our problems, we suggest that N_0 should be an integer from 0 to 5, and β should be chosen from 0.01 to 0.5.

In order to examine the method, transport parameters of cadmium ions through unsaturated soils are estimated. A laboratory soil column experiment method is adopted to collect the experimental data; detailed information about the experiment can be found in Ref. [4]. Equations (1) and (2) are numerically solved by using the FEM. For comparison, the homotopy method and the widely used Levenberg–Marquardt (LM) method are simultaneously employed to solve this problem. Two examples are carried out: the first example is based on the virtual experimental data, while the second is based on practical observation.

The main aim of the first example is to test the convergence of the method. The detailed procedure is as follows: first, the values of parameters are given, then substitute them into Eqs. (1) and (2) and the "true" concentration can be calculated. Next some random noises with different intensity are added to the true concentration for simulating experimental

Table 1. Numerical inversion results for different initial values of parameters. IT is the number of iteration; LM denotes the Levenberg–Marquardt method; HM denotes the homotopy method. The true parameters are [0.6835, 0.1154, 1.9298, 0.00255, 0.6037].

No	Initial values						Inversed values					IT
	α	K_{Cd}	K_A	K_B	γ		α	K_{Cd}	K_A	K_B	γ	
1	0.5	0.5	3.0	0.5	0.5	HM	0.6835	0.1154	1.9298	0.00255	0.6037	10
						LM	0.6835	0.1154	1.9298	0.0026	0.6037	24
2	3.0	0.5	3.0	0.5	0.5	HM	0.6835	0.1154	1.9298	0.00255	0.6037	41
						LM	0.6835	0.1154	1.9298	0.0025	0.6037	18
3	0.5	0.5	6.0	0.5	0.5	HM	0.6835	0.1154	1.9298	0.00255	0.6037	29
						LM	0.6835	0.1154	1.9298	0.00255	0.6037	34
4	3.0	0.8	10.0	0.9	0.1	HM	0.6835	0.1154	1.9298	0.00255	0.6037	37
						LM	divergent					

errors. Finally, we estimate the parameters from different initial values based on the virtual experimental data. In this example, β and N_0 of formula (8) are evaluated as 0.5 and 0 respectively. All the numerical results are presented in Tables 1 and 2.

In Table 1, different initial values of computation are selected: in the cases 1–3, the parameters can be estimated accurately by both methods. In case 4, the start point is far away from the true values of the parameters, and the Levenberg–Marquardt method fails, while the homotopy method works very well. These results show the feature of global convergence of the homotopy method. In Table 2, although the noise is relatively intensive, the distance between the true and the calculated result is not very great. These results show the capacity of restraining noise of the homotopy method.

Table 2. Numerical inversion value for different noise levels. NA is the noise intensity, the initial values of parameters are [0.5, 0.5, 3.0, 0.5, 0.5]. The true parameters are [0.6835, 0.1154, 1.9298, 0.0026, 0.6037].

NA	Inversed results					IT
	α	K_{Cd}	K_A	K_B	γ	
1%	0.6532	0.1130	1.8786	0.0028	0.6021	10
2%	0.6567	0.1151	1.8582	0.0028	0.6027	10
4%	0.5877	0.1092	1.7918	0.0025	0.6003	11
8%	0.6889	0.1199	1.8117	0.0004	0.5960	10

Table 3. Experimental conditions. Here L denotes the height of soil sample, PH represents the pH value of the solution, ΔP is the pressure difference between two ends of the soil sample, C_0 is the initial concentration of cadmium ions in the pollutant solution, T_1 denotes the time to add pollutant solution, T_2 denotes the time to add distilled water, θ denotes the volumetric water content, ρ denotes the bulk density of dry soil, and q is Darcy's velocity.

No	L (cm)	PH	ΔP (atm)	C_0 (mg/m^3)	T_1 (h)	T_2 (h)	θ	ρ (cm/g^3)	q (cm/h)
1	8.0	2.0	0.2	2550	6.0	6.0	0.2375	1.7809	1.5649
2	8.0	6.0	0.2	12406	8.0	8.0	0.2141	1.7924	0.9796
3	10.0	6.0	0.05	6939.6	3.0	3.0	0.2572	1.8190	2.9426
4	7.5	6.0	0.3	23043	9.0	9.0	0.2203	1.8673	0.9717

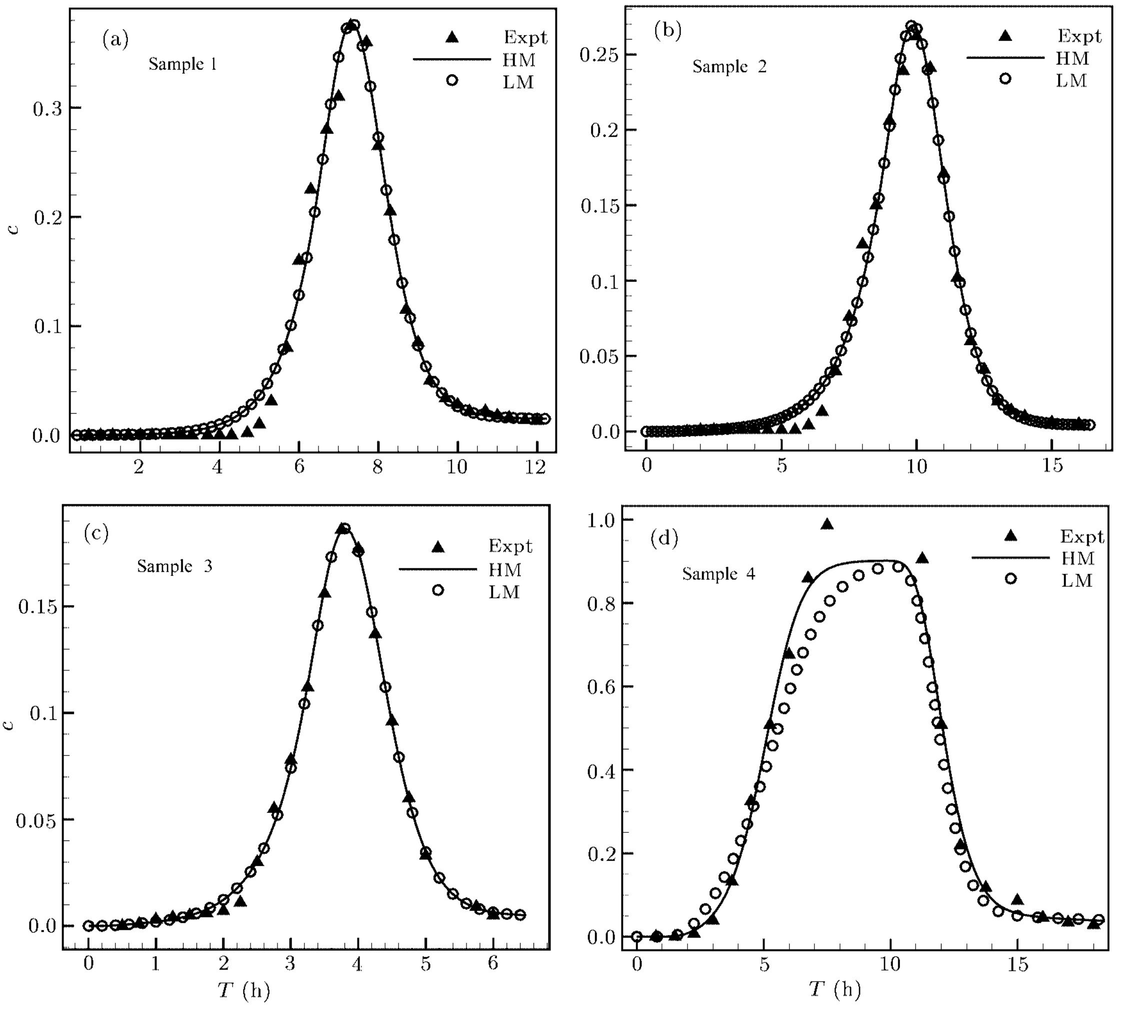

Fig. 1. Experimental and fitted breakthrough curves.

The second example is based on the practical experimental data, totally four kinds of soil samples are considered under different conditions, correlated experimental conditions are listed in Table 3.

In cases 1–3, β and N_0 of formula (8) are evaluated 0.5 and 0; in case 4, they are evaluated 0.01 and 0 respectively. The computational results are listed in Table 4. The experimental and the fitted breakthrough curves are shown in Figs. 1(a)–1(d).

Table 4. Parameters by different inversion methods with the initial values [0.5, 0.5, 0.5, 0.5, 0.5].

Num.		α	K	K_A	K_B	γ	IT
1	LM	0.4624	0.1071	2.6904	0.0040	0.5700	18
	HM	0.4623	0.1071	2.6900	0.0040	0.5700	18
2	LM	0.4574	0.0949	1.5555	0.0010	0.4953	24
	HM	0.4562	0.0949	1.5528	0.0010	0.4953	16
3	LM	0.6142	0.0907	3.6400	0.0031	0.4744	33
	HM	0.6143	0.0907	3.6407	0.0031	0.4744	23
4	LM	0.2310	0.1033	1.2558	0.0290	0.1460	300
	HM	0.2716	0.1184	2.5473	0.0419	0.1131	69

As the results show in Table 4 and Fig. 1, in cases 1–3, the good solution points are obtained by both methods, while the homotopy method is generally more effective than the LM method. In case 4, as is shown in Fig. 1(d), apparently the measurement errors are heavier than the other three cases. Even after 300 iterations, a satisfied solution cannot be obtained by using the LM method; while by using the homotopy method, a good solution can be found after 69 iterations.

Based on our investigation, we can conclude that the continuation method presented is globally convergent and highly efficient. In addition, the measurement noises can be effectively restrained. In addition, this method can also be directly extended to solve other physical inverse problems.

The authors wish to thank Professor LI Xingsi and Professor LI Baoyuan of Dalian University of Technology for their kind help.

References

[1] Xu Y S, Zhong Y J and Huang G X 2004 *Chin. Phys. Lett.* **21** 1298
[2] Yue W Z, Tao G and Zhu K Q 2004 *Chin. Phys. Lett.* **21** 2059
[3] Nielsen D R van Genuchten M T and Biggar J W 1986 *Water Resour. Res.* **22** 89S
[4] Cui K, Li B Y and Li X S 2003 *Proc. Int. Symp. Water Resour. urban Environment* ed Wang Y X (Beijing: China Environmental Science Press) p 690
[5] Zhang B X, Lu L Y and Wang C H 2004 *Chin. Phys. Lett.* **21** 682
[6] Mayer A S and Huang C L 1999 *Adv. Water Resour.* **22** 841
[7] Schwartza R C, Juo A S R and McInnes K J 2000 *J. Hydrology* **229** 149
[8] Inoue M, Simunek J, Shiozawa S and Hopmans J W 2000 *Adv. Water Resour.* **23** 677
[9] Pang L and Close M E 1999 *J. Contaminant Hydrology* **36** 185
[10] Watson L T 1989 *Appl. Math. Comput.* **31BK** 369
[11] Watson L T and Haftka R T 1989 *Comp. Methods Appl. Mech. Eng.* **74** 289
[12] Watson L T, Sosonkina M, Melville R C, Morgan A P and Walker H F 1997 *ACM Trans. Math. Software* **23** 514
[13] Aster R, Borchers B and Thurber C 2004 *Parameter Estimation and Inverse Problems* (New York: Elsevier) p 59