EURASIAN JOURNAL OF MATHEMATICAL AND COMPUTER APPLICATIONS
ISSN 2306-6172 (p); ISSN 2308-9822 (o)
Volume 1, Number 1 (2013) 62 – 77

RESTARTED NONLINEAR CONJUGATE GRADIENT METHOD FOR PARAMETER IDENTIFICATION IN ELLIPTIC SYSTEM

Jingzhi Li, Shanqiang Li[1], Hongyu Liu

Abstract We propose a restarting technique for the nonlinear conjugate gradient (NCG) method in solving the inverse problem of parameter identification in elliptic systems. The technique can enhance the performance of the NCG method in a certain generic situation. We conduct extensive numerical experiments to compare the performance of five well-known conjugate gradient schemes by incorporating the restarting technique. Numerical results show the efficiency and stable performance of the restarted NCG schemes. The smoothing property of the proposed method is also investigated, which could be employed in the multigrid solver for inverse problems.

Key words: Nonlinear conjugate gradient method, inverse problems, parameter identification, finite element methods

AMS Mathematics Subject Classification: 49J20, 49J50, 65N21, 65N30.

1 Introduction

This paper is concerned with the inverse problem of identifying distributed parameters in partial differential equation (PDE) systems. The inverse problem of this kind is of crucial importance in many areas of science and technology, and we refer to [1, 3, 6, 9, 13, 14, 15, 16, 18] and the references therein for related studies. Throughout we shall take the following model problem for our current study. Consider the elliptic PDE system

$$-\nabla \cdot (q(x)\nabla u(x)) = f(x) \qquad \text{in } \Omega, \tag{1.1}$$

$$u(x) = 0 \qquad \text{on } \Gamma, \tag{1.2}$$

where Ω is a bounded Lipschitz domain in $\mathbb{R}^d$, $d = 1, 2$ or 3, $\Gamma := \partial\Omega$, q is a positive measurable function and $f \in H^{-1}(\Omega)$. $q(x)$ represents the distributed parameter, whereas $f(x)$ represents a given source term. The inverse problem is to recover q from measurements z^δ of the state variable u , or ∇z^δ of the state variable gradient ∇u. Here and in the following, δ stands for the noise level in the measurement data. This problem arises in many industrial applications, e.g., in the fluid flow of a one-phase reservoir, it describes the procedure of identifying the inaccessible absolute permeability q of the underground medium by measuring u or ∇u; we refer to [1, 3, 6, 18] for more relevant discussion.

[1]Corresponding Author. Email: lishanqiang@gmail.com

In this paper, we shall mainly consider the numerical reconstruction algorithms for the parameter identification problem described above. Many reconstruction schemes have been developed and investigated in the literature, and among those methods the output least-squares formulation with Tikhonov regularization is one of the most stable and reliable approaches. But the resulting PDE-constrained optimization system is nonlinear and highly ill-conditioned, and hence there is still a lot of ongoing research on how to solve the large-scale system effectively and efficiently. The classical conjugate gradient (CG) method initiated by Hestenes and Stiefel in their seminal paper [11] has been widely used in solving quadratic optimization problems of high dimensions. Later, the CG method was adapted to solve non-quadratic optimization problems, and we would like to mention five different schemes including the Fletcher-Reeves method [8], the Polak-Ribière method [17], the Conjugate Descent method [7] and the Dai-Yuan method [4].

In this paper, we propose an enhancement technique for the nonlinear conjugate gradient method in solving the parameter identification problem within the framework of output least-squares formulation. More specifically, we first reformulate the parameter identification problem into a PDE-constrained optimization problem. Then the Gateaux derivatives of the cost functional at all directions are computed for the gradient direction by solving the direct and adjoint PDEs. Nonlinear CG schemes are used to generate the conjugate gradient direction based on the gradient knowledge. If the conjugate gradient direction is a descent direction, we update the current estimated parameter along the conjugate gradient direction, otherwise we update the current iteration along the negative gradient direction and restart the NCG scheme. Inexact line search using a simple backtracking rule is used in the updating step for this nonlinear problem. The idea of our new algorithms tracks back to the work of Martin Hanke, who discussed in detail how to employ the standard conjugate gradient type methods for the classical linear inverse problem of the Fredholm integral equations of the first kind with non-degenerate kernel functions in [10]. But application of the CG scheme for the nonlinear inverse problem is not so natural as that for its linear counterpart. There is in general no normal equations and one has to apply the CG method in the optimization framework. But as is well known in the quadratic optimization theory that the conjugate gradient direction may not be a descent direction, and in that case, line search along it might fail in the optimization setting, which causes the CG iteration to halt in its midway. The restarting technique we propose in the present work can overcome this halt problem effectively and efficiently.

Finally, we would like to remark that the numerical technique developed in the present paper can also be applied to other inverse problems of identifying conductivities in EIT, and medium parameters in acoustic and electromagnetic scattering etc.; see, e.g. [2, 19, 20].

The rest of the paper is organized as follows. In Section 2, we present the reformulation of the inverse problem as a PDE-constrained optimization problem and then derive the gradient formula for the cost functional. In Section 3, we propose our restarted nonlinear conjugate gradient algorithm (RNCG). Section 4 is devoted to numerical experiments. Our paper is concluded in Section 5.

2 Mathematical formulation

In this section, we briefly describe the PDE-constrained optimization formulations of the parameter identification problem, together with the corresponding finite element discretization. We also derive the explicit formulae for the Gateaux derivatives of the discrete cost functionals on the finite element space. As for the theoretical details behind these formulations including the continuity of the discrete cost functionals, existence of minimizers and weak convergence analysis etc., we refer to [3, 12, 16, 21].

We first fix some notations for our subsequent use. In order to derive the approximate discrete problem by the finite element method, we first triangulate the polyhedral domain Ω with a regular triangulation $\mathcal{T}^h$ of simplicial elements, namely intervals in one dimension, triangles in two dimensions and tetrahedra in three dimensions. Then the finite element space V_h is defined to be the space consisting of continuous and piecewise linear functions over the triangulation $\mathcal{T}^h$, and $\overset{\circ}{V}_h$ is a subspace of V_h with all functions vanishing on the boundary $\partial\Omega$. Henceforth, the regularization parameter is always represented as γ and the admissible parameter set is given as

$$K = \{q \in H^1(\Omega);\ \|q\|_{H^1(\Omega)} < \infty \text{ and } 0 < \alpha_1 \le q(x) \le \alpha_2 \text{ a.e. in } \Omega\}, \tag{2.1}$$

where α_1 and α_2 are two positive constants. Let $\{x_i\}_{i=1}^N$ be the set of all the nodal points of the triangulation $\mathcal{T}^h$, then the constrained subset K is approximated by

$$K_h = \{v_h \in V_h; \alpha_1 \le v_h(x) \le \alpha_2, \quad \forall x \in \Omega\}. \tag{2.2}$$

With the above preparations, we are in a position to present the mathematical formulations of the parameter identification problem associated with the PDE system (1.1)–(1.2). We first consider the case that the measured data is ∇z^δ of ∇u on Ω. By the output least-squares method combined with the Tikhonov regularization, the inverse problem can be formulated as the following PDE-constrained minimization problem

$$\min_{q\in K} J(q) = \frac{1}{2}\int_\Omega q|\nabla v - \nabla z^\delta|^2\, dx + \gamma\int_\Omega |\nabla q|^2\, dx \tag{2.3}$$

subject to $q \in K$ and $v \equiv v(q) \in H_0^1(\Omega)$ satisfying

$$\int_\Omega q\nabla v\cdot\nabla\phi\, dx = \int_\Omega f\phi\, dx \quad \text{for all } \phi \in H_0^1(\Omega). \tag{2.4}$$

Next, by applying the finite element discretization, the discrete constrained minimization problem corresponding to (2.3)–(2.4) is given as follows,

$$\min_{q_h\in K_h} J_h(q_h) = \frac{1}{2}\int_\Omega q_h|\nabla v_h - \nabla z^\delta|^2 dx + \gamma\int_\Omega |\nabla q_h|^2\, dx \tag{2.5}$$

subject to $q_h \in K_h$ and $v_h \equiv v_h(q_h) \in \overset{\circ}{V}_h$ satisfying

$$\int_\Omega q_h\nabla v_h\cdot\nabla\phi_h\, dx = \int_\Omega f\phi_h\, dx \quad \text{for all } \phi_h \in \overset{\circ}{V}_h. \tag{2.6}$$

For the subsequent use, we derive the gradient of the functional $J_h(q_h)$ in (2.5). It can be shown that for each $q_h \in K_h$, the Gateaux derivative of $v_h(q_h)$ with respect to any $p_h \in V_h$, namely $v_h'(q_h)p_h \in \overset{\circ}{V}_h$, satisfies

$$\int_\Omega q_h \nabla v_h'(q_h)p_h \cdot \nabla \phi_h \, dx = -\int_\Omega p_h \nabla v_h(q_h) \cdot \nabla \phi_h \, dx \quad \text{for all } \phi_h \in \overset{\circ}{V}_h. \tag{2.7}$$

Next, we introduce an adjoint equation of finding $w_h \in \overset{\circ}{V}_h$ such that

$$\int_\Omega q_h \nabla w_h \cdot \nabla \phi_h \, dx = \int_\Omega q_h \nabla (v_h - z^\delta) \cdot \nabla \phi_h \, dx \quad \text{for all } \phi_h \in \overset{\circ}{V}_h. \tag{2.8}$$

By combining (2.7) and (2.8), the Gateaux derivative of the functional $J_h(q_h)$ with respect to p_h can be further formulated as follows

$$\begin{aligned} & J_h'(q_h)p_h \\ = \; & \frac{1}{2}\int_\Omega p_h |\nabla v_h - \nabla z^\delta|^2 dx + \int_\Omega q_h (\nabla v_h - \nabla z^\delta) \cdot \nabla v_h'(q_h)p_h dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \\ = \; & \frac{1}{2}\int_\Omega p_h |\nabla v_h - \nabla z^\delta|^2 dx + \int_\Omega q_h \nabla v_h'(q_h)p_h \cdot \nabla w_h dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \\ = \; & \frac{1}{2}\int_\Omega p_h (\nabla v_h - \nabla z^\delta)^2 \, dx - \int_\Omega p_h \nabla v_h \cdot \nabla w_h \, dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \,, \end{aligned} \tag{2.9}$$

where we have taken $\phi_h = w_h$ in (2.7) and $\phi_h = v_h'(q_h)p_h$ in (2.8), and then plugged them into the first and second equalities in (2.9).

Next, we consider the case that the measured data is z^δ of u on Ω. Similar to the first case, the inverse problem can be formulated as the following PDE-constrained minimization problem

$$\min_{q \in K} J(q) = \frac{1}{2}\int_\Omega |v - z^\delta|^2 \, dx + \gamma \int_\Omega |\nabla q|^2 \, dx \tag{2.10}$$

subject to $q \in K$ and $v \equiv v(q) \in H_0^1(\Omega)$ satisfying

$$\int_\Omega q \nabla v \cdot \nabla \phi \, dx = \int_\Omega f\phi \, dx \quad \text{for all } \phi \in H_0^1(\Omega) \,. \tag{2.11}$$

By applying the finite element discretization, the discrete counterpart of (2.10)–(2.11) is then given by

$$\min_{q_h \in K_h} J_h(q_h) = \frac{1}{2}\int_\Omega |v_h - z^\delta|^2 dx + \gamma \int_\Omega |\nabla q_h|^2 \, dx \tag{2.12}$$

subject to $q_h \in K_h$ and $v_h \equiv v_h(q_h) \in \overset{\circ}{V}_h$ satisfying

$$\int_\Omega q_h \nabla v_h \cdot \nabla \phi_h \, dx = \int_\Omega f\phi_h \, dx \quad \text{for all } \phi_h \in \overset{\circ}{V}_h \,. \tag{2.13}$$

We also introduce the following adjoint equation: Find $w_h \in \overset{\circ}{V}_h$ such that

$$\int_\Omega q_h \nabla w_h \cdot \nabla \phi_h \, dx = \int_\Omega (u_h - z^\delta)\phi_h \, dx \quad \text{for all } \phi_h \in \overset{\circ}{V}_h. \tag{2.14}$$

By using (2.13) and (2.14), the Gateaux derivative of the functional $J_h(q_h)$ in (2.12) with respect to p_h can be given as

$$\begin{aligned} J_h'(q_h)p_h &= \int_\Omega (v_h - z^\delta) v_h'(q_h)p_h \, dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \\ &= \int_\Omega q_h \nabla w_h \cdot \nabla v_h'(q_h)p_h \, dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \\ &= -\int_\Omega p_h \nabla w_h \cdot \nabla v_h \, dx + 2\gamma \int_\Omega \nabla q_h \cdot \nabla p_h \, dx \end{aligned} \tag{2.15}$$

where we have taken $\phi_h = w_h$ in (2.13) and $\phi_h = v_h'(q_h)p_h$ in (2.14), and then plugged them into the first and second equalities in (2.15).

3 RNCG methods for parameter identification

In this section, we present the restarted NCG method in treating the concerned inverse problem. In general, the conjugate gradient type method for a minimization problem with a cost functional $J_h(q)$ has the following form:

$$q_{k+1} = q_k + \alpha_k d_k, \tag{3.1}$$

where α_k is the step length and d_k is the search direction of the form

$$d_k = \begin{cases} -g_k, & \text{for } k = 0; \\ -g_k + \beta_k d_{k-1}, & \text{for } k > 0. \end{cases} \tag{3.2}$$

In (3.2), g_k denotes the gradient direction of the cost functional ∇J_h, and it can be represented by a linear combination of the basis functions of V_h with the weights being the Gateaux derivatives along each basis function. Moreover, in (3.2), β_k is such chosen that d_k becomes the k-th conjugate gradient when the cost functional is quadratic and the line search is exact. Variants of conjugate gradient methods differ in the criterion of how to select β_k ($k \geq 1$). Five well-known schemes for β_k ($k \geq 1$) are given as follows:

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})}, \quad \text{(Hestenes-Stiefel [11])} \tag{3.3}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \text{(Fletcher-Reeves [8])} \tag{3.4}$$

$$\beta_k^{FR} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}, \quad \text{(Polak-Ribière [17])} \tag{3.5}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}}, \quad \text{(Conjugate Descent [7])} \tag{3.6}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})}. \quad \text{(Dai-Yuan [4])} \tag{3.7}$$

In the following we formulate the initial version of the NCG algorithm for the parameter identification problem in Algorithm 1.

Algorithm 1 Nonlinear Conjugate Gradient

$k := 0$;
$q_h^0 :=$ initial value;
$g_0 := \nabla J_h(q_h^0)$;
$d_0 := -g_0$;
while (stopping rule) %NCG iterations
  Choose $\alpha_k > 0$ s.t. $\alpha_k \|\nabla J_h\|_\infty = \kappa \|q_h\|_\infty$;
  while $J(q_h^k + \alpha_k d_k) \geq J(q_h^k)$
    $\alpha_k := \alpha_k/2$;
  endwhile
  $q_h^{k+1} := q_h^k + \alpha_k d_k$;
  $g_{k+1} := \nabla J_h(q_h^{k+1})$;
  $\beta_{k+1} :=$ one of the criteria (3.3)–(3.7);
  $d_{k+1} := -g_{k+1} + \beta_{k+1} d_k$;
  $k := k + 1$;
end

One critical drawback of the **NCG algorithm** is that the conjugate gradient direction may not be a descent direction for the cost functional $J_h(q_h)$, which causes the innermost *while loop* to loop forever. The remedy we propose is to restart the CG algorithm when the step length α_k is smaller than a threshold value $\varepsilon > 0$, i.e., we replace the current conjugate gradient d_k with the current negative gradient $-g_k$ as the initial guess and restart the NCG algorithm. This suggests the restarted nonlinear conjugate gradient algorithm (RNCG), which is formulated in Algorithm 2.

It is remarked that we will terminate the program when it reaches the maximum iteration number or the relative error between the successive two iterates, namely $\|q_k - q_{k-1}\|/\|q_k\|$, is less than 10^{-16} for the purpose of comparison. Other stopping rules, such as monitoring the absolute difference of the consecutive iterates $\|q_k - q_{k-1}\|$, can be also employed in practice. To avoid the denominators in (3.3)–(3.7) tending to zero, we add a small numbers 10^{-6} with the same sign as that of the denominators in the denominators. We observe that Hestenes-Stiefel, Conjugate Descent and Dai-Yuan methods will stagnate during the iteration without this safeguard trick. For the line search step, we take the simple backtracking criterion, i.e., reducing the step size by a half if the cost functional is not sufficiently decreased (i.e., the step size is overemphasized). And the initial step size for α_k is chosen such that the magnitude of $\|\alpha_k \nabla J_h\|_\infty$ is of the same order as that of $\|q_h\|_\infty$. This is to make the decrease comparable to the magnitude of the current guess.

4 Numerical experiments

For numerical illustrations of our proposed methods for the parameter identification problem, we implemented the five schemes of the nonlinear conjugate gradient algorithms in Matlab and conducted experiments on a desktop with dual cores of 3.1GHz

Algorithm 2 Restarted Nonlinear Conjugate Gradient

$k := 0$;
q_h^0 := initial value;
$g_0 := \nabla J_h(q_h^0)$;
$d_0 := -g_0$;
while (stopping rule) %NCG iterations
  Choose $\alpha_k > 0$ s.t. $\alpha_k \|\nabla J_h\|_\infty = \kappa \|q_h\|_\infty$
  while $J(q_h^k + \alpha_k d_k) \geq J(q_h^k)$
    $\alpha_k := \alpha_k/2$;
    if $\alpha_k < \varepsilon$
      $d_k := -g_k$;
      Choose $\alpha_k > 0$ s.t. $\alpha_k \|\nabla J_h\|_\infty = \kappa \|q_h\|_\infty$
      while $J(q_h^k + \alpha_k d_k) \geq J(q_h^k)$
        $\alpha_k := \alpha_k/2$;
      endwhile
      break;
    endif
  endwhile
  $q_h^{k+1} := q_h^k + \alpha_k d_k$;
  $g_{k+1} := \nabla J_h(q_h^{k+1})$;
  β_{k+1} := one of the criteria (3.3)–(3.7);
  $d_{k+1} := -g_{k+1} + \beta_{k+1} d_k$;
  $k := k + 1$;
endwhile

CPUs.

The test problem is a parameter identification problem of elliptic type in two dimensions

$$-\nabla \cdot \Big(q(x,y)\nabla u(x,y)\Big) = f(x,y) \quad (x,y) \in (0,1)\times(0,1). \tag{4.1}$$

Without loss of generality, the boundary condition are always assumed to be of homogeneous Dirichlet type.

The parameters related in the algorithm are chosen as follows. The lower and upper bounds α_1 and α_2 in the constraint set K_h are taken to be 0.5 and 20, respectively. We let the threshold value $\varepsilon = 10^{-5}$. The initial guess q_h^0 is set to be a constant function 10 for both test problems, which is far away from the exact parameters. We add uniform random errors to the exact solution u or its gradient ∇u to generate the observed data z^δ or ∇z^δ in the following way,

$$z^\delta = u(1+\delta r) \quad \text{or} \quad \nabla z^\delta = \nabla u(1+\delta r).$$

where δ represents the random error level and r denotes a uniforma random number between -1 and 1 pointwise. We choose the noise level δ to be 0.01, 0.05 or 0.1, which is common in engineering applications. The regularization parameter γ is chosen following the Morozov discrepancy rule in accordance with the error level and different test problems.

We divide the region $(0,1)\times(0,1)$ into 32×32 equal squares and then further divide these squares through their diagonal into two triangles and the finite element mesh size is $\sqrt{2}/32$.

For the parameter identification problems when ∇z^δ is available, the maximum number of the iterations for all the algorithms is taken to be 100, while for the problems when z^δ is available, the maximum number is taken to be 200.

In the following experiments we illustrate the convergence of the k-th iterate q_k to the true parameter q_{exact} for all the algorithms by plotting the discrete L^2 norm of the error, i.e. $\|q_k - q_{\text{exact}}\|$ for the increasing iteration number. The discrete L^2 norm of current and previous iterates, i.e. $\|q_k - q_{k-1}\|$ for the increasing iteration number are also plotted for comparison.

4.1 RNCG algorithm for parameter identification

Example 1. (H^1-data case) We take the exact coefficient $q(x,y)$ and state variable $u(x,y)$ in (4.1) and observed data $\nabla z^\delta(x,y)$ as follows:

$$\begin{aligned} q(x,y) &= 3 + 32x(1-x)y(1-y), \\ u(x,y) &= \sin(\pi x)\sin(\pi y), \\ \nabla z^\delta(x,y) &= (1+\delta r(x,y))\nabla u(x,y), \end{aligned}$$

and the right hand side function $f(x,y)$ is computed through (4.1) using $u(x,y)$ and $q(x,y)$. Here we use the RNCG algorithm with the Fletcher-Reeves scheme as an example. Figure 1 shows the exact parameter $q_{\text{exact}}(x,y)$ and the identified parameter $q_h(x,y)$ when the error levels are 0.01, 0.05 and 0.10 with the regularization parameter $\gamma = 0.003$, 0.02 and 0.026, respectively, determined by the discrepancy principle. Figure

2 shows the convergence history curve along the iterations with respect to different noise levels. We see that the method converges quite fast and stable for different level of noises and the profile of the parameter can be well approximated within the first 100 iterations except for those points on the boundary or where the gradients of u are nearly zero. The convergence history curve of the L^2 norm of the error between the current iterate q_k and the exact coefficient q_{exact} level off in the end which indicates numerically the convergence of the RNCG algorithm. Even after adding white noise up to ten percents, the identified parameter still has the correct profile as the exact one. We find that restarting is triggered at about 10–20 iterations, and then happens from time to time with increasing frequency in particular when noise level is small, namely $\delta = 0.01$. Similar phenomena are observed for the L^2 measurement case in the next example. At the same time, we implements the other schemes for $\delta = 0.01$ with the same parameter setting, we list the identified parameters in Figure 3. From numerical results, we find that the iterates using Hestenes-Stiefel and Conjugate Descent methods fail to delineate the parameter due to the trap of local minimum. Both Hestenes-Stiefel and Conjugate Descent methods stagnate after the first 10 to 20 iterations and the convergence curves level off in a premature stage. While the iterates using Polak-Ribière and Dai-Yuan methods converge to the exact solution quite stable. We plot the convergence history for the five schemes in Figure 4.

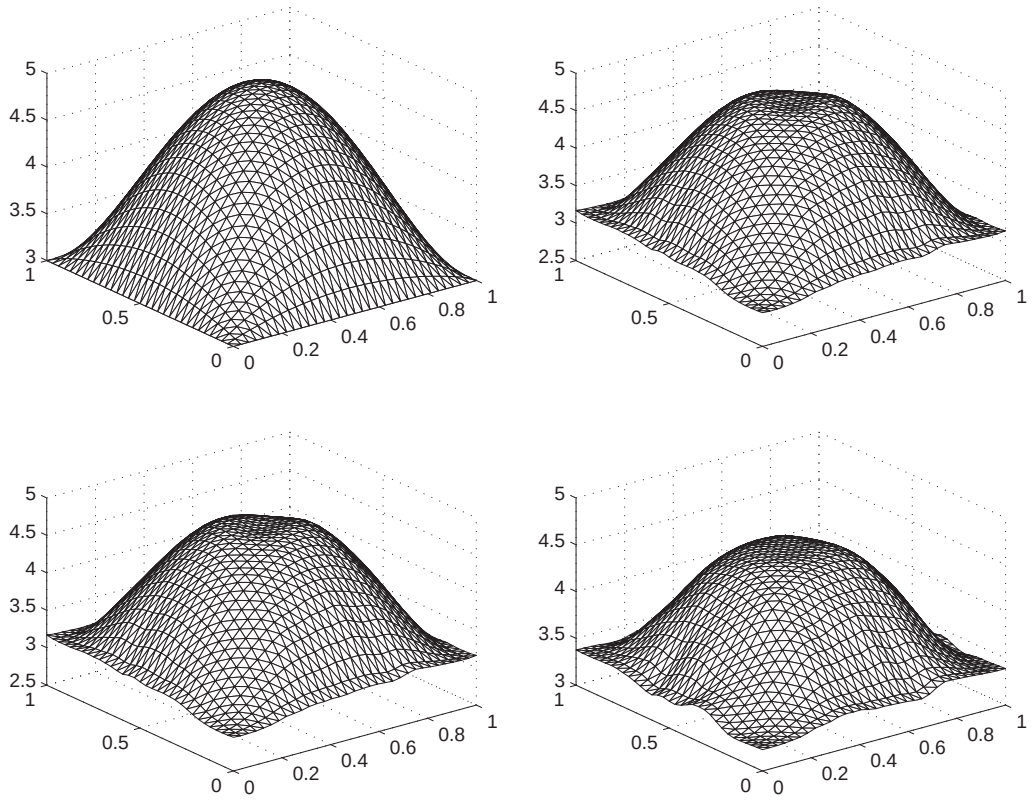

Figure 1: From left to right and then from up to down: the exact parameter q_{exact}, identified parameters q_h when $\delta = 0.01$, 0.05, 0.10.

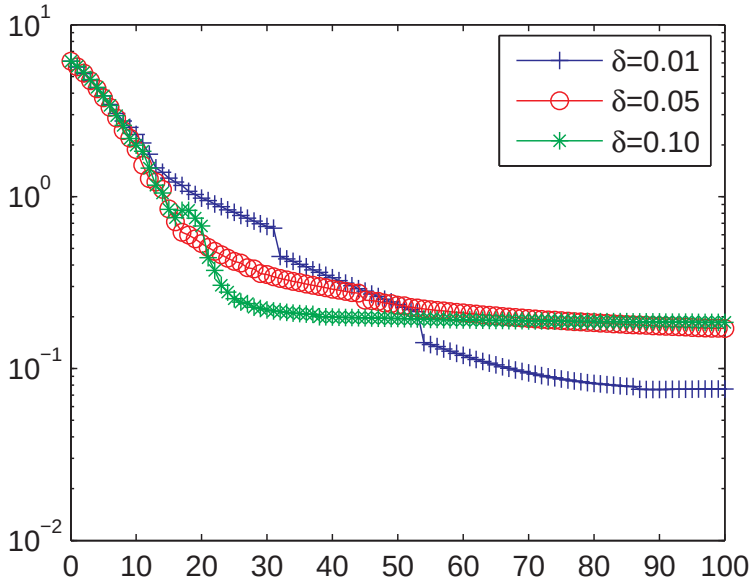

Figure 2: $\|q_k - q_{\text{exact}}\|_{L^2}$ versus iterations with different noise level. Logarithmic scale on the vertical axis.

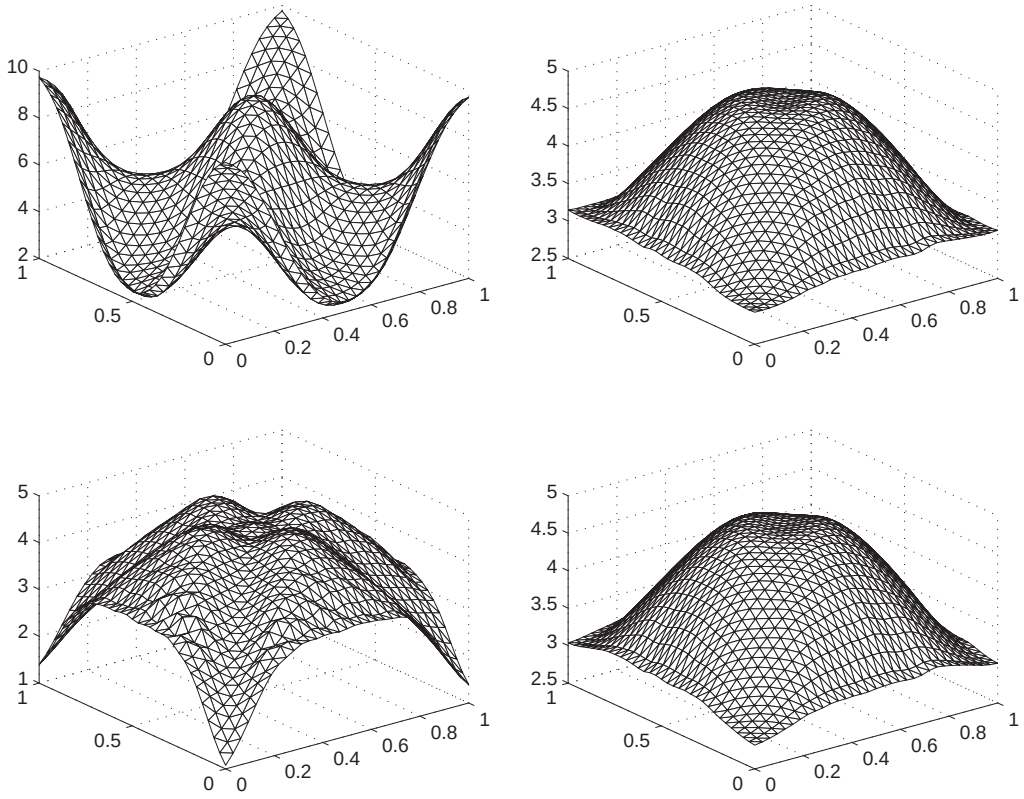

Figure 3: From left to right and then from up to down: identified parameters q_h when $\delta = 0.01$ with Hestenes-Stiefel, Polak-Ribière, Conjugate Descent and Dai-Yuan schemes.

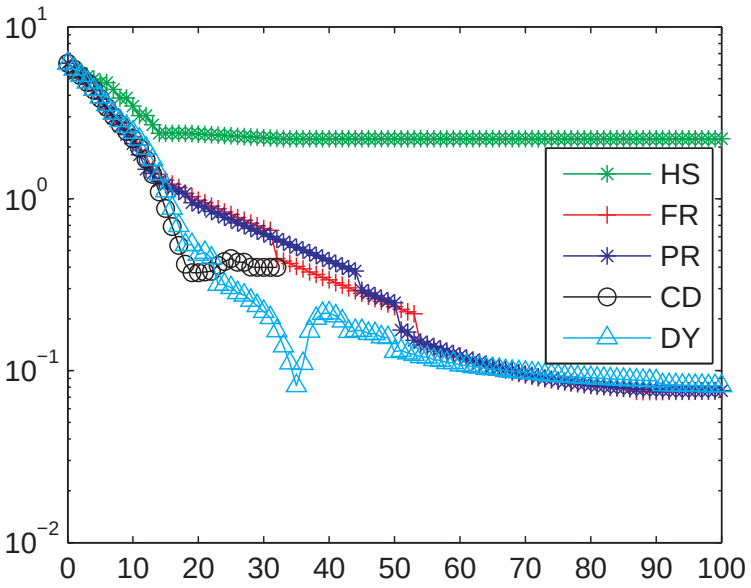

Figure 4: Convergence history for five NCG schemes.

Example 2. **(L^2-data case)** We take the exact coefficient $q(x,y)$ and state variable $u(x,y)$ in (4.1) and observed data $z^\delta(x,y)$ as follows:

$$\begin{aligned} q(x,y) &= 7+6x^2y(1-y), \\ u(x,y) &= \sin(\pi x)\sin(\pi y), \\ z^\delta(x,y) &= (1+\delta r(x,y))u(x,y), \end{aligned}$$

and the right hand side function $f(x,y)$ is computed through (4.1) using $u(x,y)$ and $q(x,y)$. The identification process using the L^2-data and the Fletcher-Reeves method is much slower than that using the H^1-data case, but the identified parameters is still satisfactory, with similar profiles with the exact one as shown in Figure 5. The regularization parameters are chosed 1.1×10^{-7}, 2.3×10^{-7} and 3.8×10^{-7} for $\delta=0.01$, 0.05 and 0.1, respectively, from the discrepancy rule. Here the fidelity term of the cost functional is rather small compared with the regularization term, so the regularization parameters differ little compared with the H^1 data case. Figure 6 shows the convergence history for different noise level.

Note that with the increase of the noise in the measurement data, the ill-posedness of the parameter identification problems could be mitigated due to the noise in the data. In such a way, one could achieve robustness with respect to a wide range of noise levels. For inverse problems with large noise, readers may refer to [5] and the references therein.

4.2 Smoothing property of RNCG algorithm

RNCG methods have similar smoothing property as the Uzawa method, which makes them suitable to be potential smoothers in the multigrid solver for the inverse problems [16].

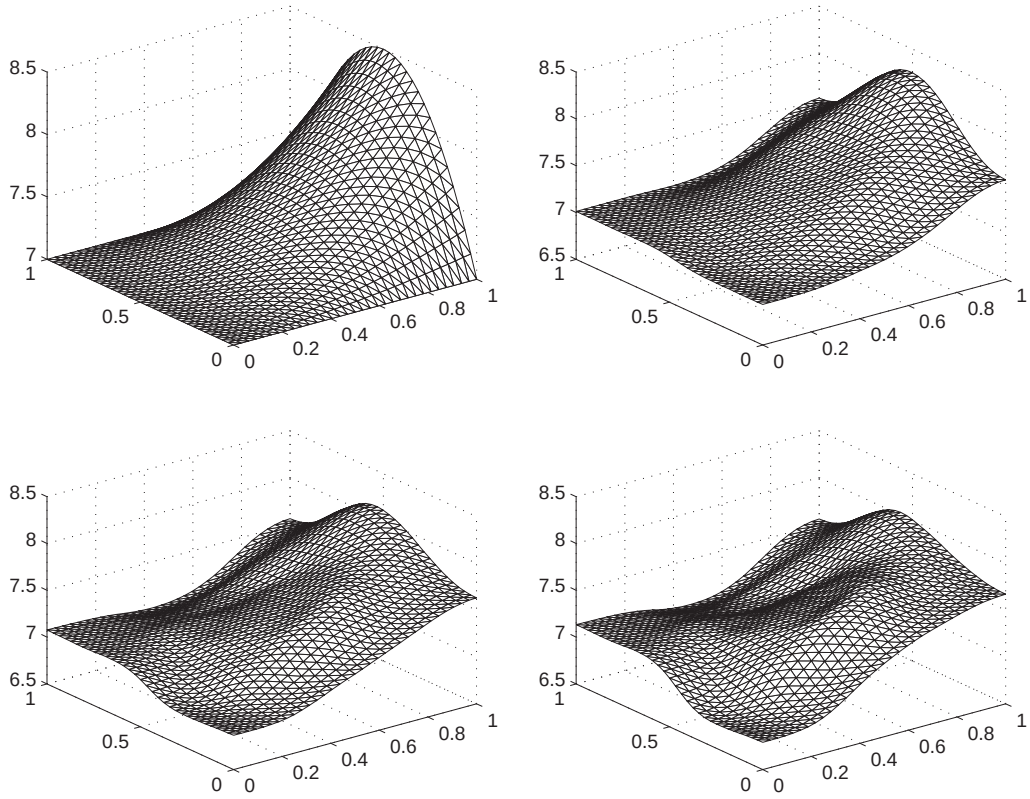

Figure 5: From left to right and then from up to down: the exact parameter q_{exact}, identified parameters q_h when $\delta = 0.01, 0.05, 0.10$.

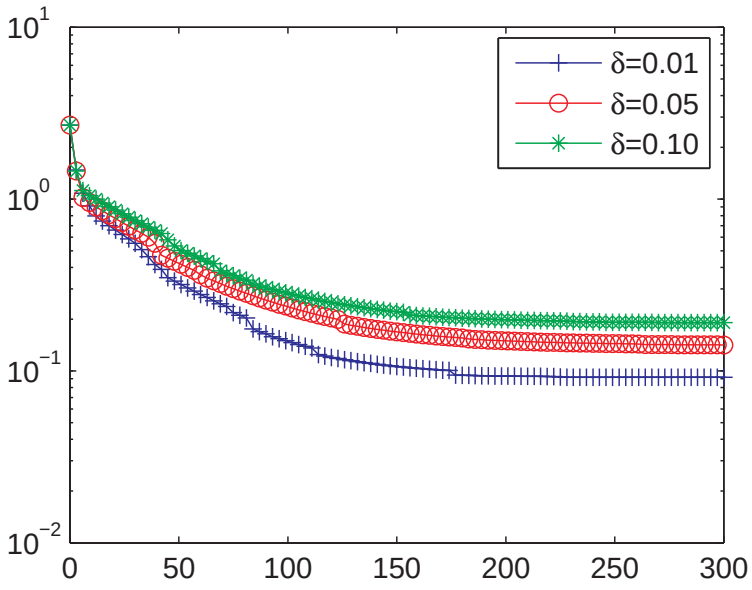

Figure 6: $\|q_k - q_{\text{exact}}\|_{L^2}$ versus iterations with different noise level. Logarithmic scale on the vertical axis.

As we will see, the smoothing property of the inverse problem solver depends highly on the regularization term and the regularization parameter. Here we use the H^1-seminorm as the regularization term in our example, which acts as the role of the inverse of a Laplace operator to force the error to be smooth. But on the other hand, the regularization parameter plays a role on determining how strong the smoothing property is. More precisely, the smaller the regularization parameter, the less smoothing property the gradient methods have. If the regularization parameter is too small, the smoothing property can be totally superseded by the nonlinearity of the cost functional and thus the resulting errors may still oscillate wildly.

Here we demonstrate the smoothing property of the RNCG algorithm for the parameter identification problems by an example. We take the test problem (4.1) with the exact parameter $q_e(x,y) = 7 + 6x^2y(1-y)$.

We deliberately set initial guesses

$$\begin{aligned} q_3^{(0)}(x,y) &= q_e(x,y) + 2\sin(2^j x)\sin(2^k y), \\ u(x,y) &= \sin(\pi x)\sin(\pi y) \end{aligned}$$

where j, k are used to control the high or low frequency mode errors in the initial guess. The observed data $\nabla z = \nabla u$, i.e., with no noise in the data. We add a small regularization parameter $\gamma = 10^{-4}$. Figure 7 presents the damping process of the high frequency modes when $j = 7$ and $k = 7$ by using the RNCG algorithm with the Fletcher-Reeves updating scheme, and the iteration processes of the damping process of the low-median frequency modes when $j = 3$ and $k = 3$ by the same method. We see clearly the high frequency mode in the initial guess could be quickly damped in the first two iterations while the low frequency mode decays much slower. Therefore, the RNCG method is shown numerically a good candidate for the multi grid solver for inverse problems.

5 Concluding remarks

Efficient RNCG methods are proposed in this work to solve a model nonlinear inverse problem of parameter identification. The proposed methods can also be used to the identification of the source term, initial condition, or radiative coefficient and the corresponding inverse problems in parabolic systems, which will be studied in the future. The smoothing property of the conjugate gradient method suggests us to utilize the NCG algorithm as a smoother component in the multigrid setting [16]. The optimal convergence rate of the multigrid-like method will provide an efficient and robust numerical method for the parameter identification problem.

Acknowledgement

The work of Jingzhi Li is supported by the NSF of China (No. 11201453 and 91130022). The authors would like to thank the anonymous referee for many constructive comments, which significantly improve the results of this paper.

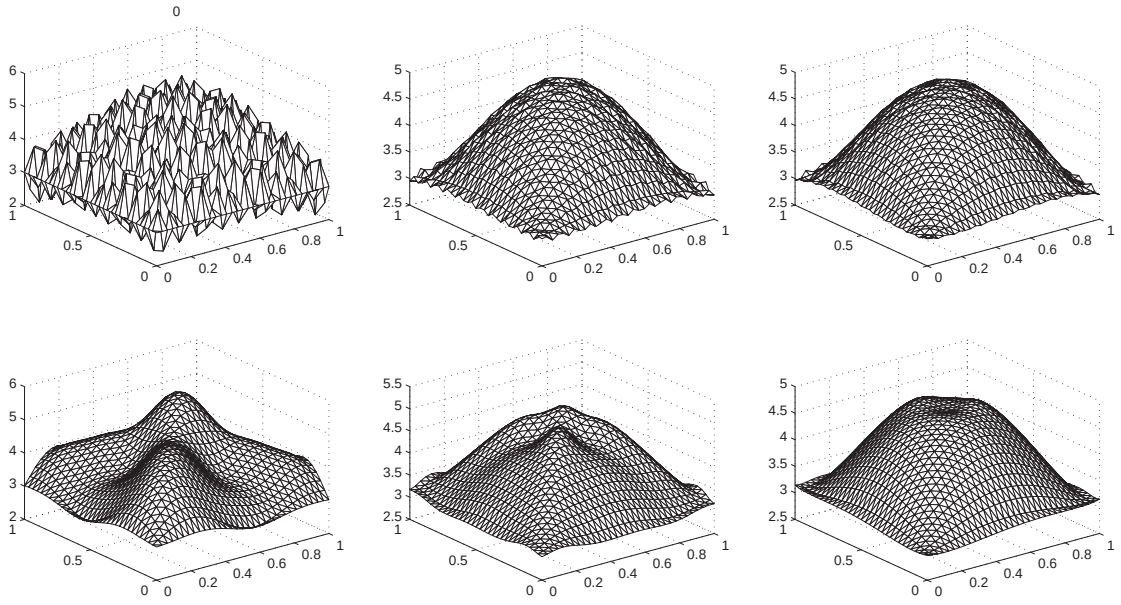

Figure 7: Illustration of the smoothing property of the RNCG algorithm in two dimension. Top from left to right: initial guess and the first two RNCG iterates for high frequency mode $j = k = 7$. Bottom from left to right: the initial guess and the tenth, twentieth RNCG iterates for low-median frequency mode $j = k = 3$.

References

[1] H. T. Banks and K. Kunisch, *Estimation techniques for distributed parameter systems*, Boston : Birkhäuser, 1989.

[2] G. Bao and P. Li, *Inverse medium scattering problems for electromagnetic waves*, SIAM J. Appl. Math., 65 (2005), pp. 2049–2066.

[3] Z. Chen and J. Zou, *An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems*, SIAM J. Control Optim., 37 (1999), pp. 892–910.

[4] Y. Dai and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optimization, 10 (1999), pp. 177–182.

[5] H. Egger, *Regularization of inverse problems with large noise*, Journal of Physics: Conference Series, 124 (2008), p. 012022.

[6] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Dordrecht ; Boston : Kluwer Academic Publishers, 1996.

[7] R. Fletcher, *Practical Method of Optimization, Vol I: Unconstrained Optimization*, Wiley, New York, 2 ed., 1987.

[8] R. Fletcher and C. Reeves, *Function minimization by conjugate gradients*, Comput. J., 7 (1964), pp. 149–154.

[9] R. B. Guenther, R. Hudspeth, W. McDougal, and J. Gerlach, *Remarks on parameter identification. I*, Numer. Math., 47 (1985), pp. 355–361.

[10] M. Hanke, *Conjugate gradient type methods for ill-posed problems*, Harlow, Essex, England : Longman Scientific & Technical ; New York : Wiley, 1995.

[11] M. Hestenes and E. Stiefel, *Method of conjugate gradient for solving linear system*, J. Res. Nat. Bur. Stand., 49 (1952), pp. 409–436.

[12] Y. L. Keung and J. Zou, *Numerical identifications of parameters in parabolic systems*, Inverse Problems, 14 (1998), pp. 83–100.

[13] K. Kunisch and W. Ring, *Regularization of nonlinear ill-posed problems with closed operators*, Num. Funct. Anal. and Optimiz., 14 (1993), pp. 389–404.

[14] J. Li, H. Liu, and J. Zou, *Multilevel linear sampling method for inverse scattering problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1228–1250.

[15] J. Li, J. Xie, and J. Zou, *An adaptive finite element reconstruction of distributed fluxes*, Inverse Problems, 27 (2011), p. 075009 (25pp).

[16] J. Li and J. Zou, *A multilevel model correction method for parameter identification*, Inverse Problems, 23 (2007), pp. 1759–1786.

[17] B. Polak and G. Ribiére, *Note sur la convergence des méthodes de directions conjuguées*, Rev. Fran. Informat. Rech. Opér., 16 (1969), pp. 35–43.

[18] W. W.-G.Yeh, *Review of parameter identification procedures in groundwater hydrology: The inverse problem*, Water Resources Research, 22 (1986), pp. 95–108.

[19] G. Uhlmann, *Electrical impedance tomography and Calderon's problem*, Inverse Problems, 2009, 25, p. 123011.

[20] G. Uhlmann, *Inside Out: Inverse Problems and Applications*, Cambridge University Press, Cambridge, 2003

[21] J. Zou, *Numerical methods for elliptic inverse problems*, Int. J. Comput. Math., 70 (1998), pp. 211–232.

Jingzhi Li
Faculty of Science,
South University of Science
and Technology of China,
Shenzhen 518055, P. R. China,
Email: li.jz@sustc.edu.cn;

Hongyu Liu
Department of Mathematics,
Harbin University of Science and Technology,
Harbin 150080, Heilongjiang, P. R. China,
and School of Mathematics and Statistics,
Central South University, Changsha,
Hunan 410075, P. R. China.

Shanqiang Li
College of Automation,
Harbin Engineering University,
Harbin 150001, Heilongjiang, P. R. China
and Department of Mathematics,
Harbin University of Science and Technology,
Harbin 150080, Heilongjiang, P. R. China
Email: lishanqiang@gmail.com

Received 11 Feb 2013, in final form 25 Mar 2013