Optimization methods for inverse problems with energy norm

Hrtus, Rostislav

Optimization methods for inverse problems with energy norm

Rostislav Hrtus[1,2], Jaroslav Haslinger[1], and Radim Blaheta[1]

[1]Institute of Geonics AS CR, Department IT4Innovations
[2]Department of Applied Mathematics, FEECS,
VŠB – Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava – Poruba
{hrtus, blaheta}@ugn.cas.cz

Abstract. The paper deals with the application of the energy norm in the inverse problem which is based on a minimization of a nonlinear least squares function. Three different parameter distribution examples are solved by the steepest descent, nonlinear conjugate gradient and Newton methods. Efficiency comparison is also presented.

1 Setting of the problem

We consider a **state problem**

$$\begin{cases} -\operatorname{div}(p\nabla u) = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases} \tag{1}$$

where p is a parameter vector and $\Omega = \langle 0,1\rangle \times \langle 0,1\rangle$ is a unit square which is decomposed into subdomains $\Omega_i, i = 1,\dots,r$ such that $\bar{\Omega} = \bigcup_{i=1}^{r} \bar{\Omega}_i$, $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$.
The set $\mathcal{U}_{ad}$ of admissible parameters is defined as follows:

$$\mathcal{U}_{ad} = \left\{ p \in L^{\infty}(\Omega) \mid p_{min} \le p \le p_{max}, p_{|\Omega_i} \in P_0(\Omega_i), i = 1,\dots,r \right\},$$

i.e. $\mathcal{U}_{ad}$ is the set of piecewise constant functions on the partition of Ω into $\{\Omega_i\}_{i=1}^{r}$, $0 < p_{min} < p_{max}$ are given and $p = (p_1,\dots,p_r), p_i = p_{|\Omega_i}$. In addition, material interfaces align the decomposition of Ω.

The state problem (1) is solved using the Finite Element Method. After the discretization by piecewise linear elements we obtain a linear system of algebraic equations

$$A(p)u(p) = b, \tag{2}$$

where A is a symmetric positive definite matrix, $p \in \mathbb{R}^r$ $b \in \mathbb{R}^n$ Inverse problem will be based on the minimization of a cost function $\mathcal{J}$ which depends on the solution of the discretized state problem $u(p)$ and measurements $\hat{u} \in R^n$ of the state variable. We suppose that the measurements are at our disposal at each node of the discretization, therefore we are able to use the energy norm (3). Now

© M. Krátký, J. Dvorský, P. Moravec (Eds.): WOFEX 2014, pp. 484–489.
VŠB – Technical University of Ostrava, FEECS, 2014, ISBN 978-80-248-3458-0.

we are ready to formulate the inverse problem in the discrete form which reads as follows:

$$(\mathbb{P})_h \begin{cases} \text{find } p^* \in \mathcal{U}_{ad} \text{ such that} \\ \mathcal{J}(p^*) \ \leq \ \mathcal{J}(p) \quad \forall p \in \mathcal{U}_{ad}, \end{cases}$$

where

$$\mathcal{J}(p) = \frac{1}{2} \langle A(p)(u(p) - \hat{u}), u(p) - \hat{u} \rangle \tag{3}$$

For the existence of a solution to the inverse problem $(\mathbb{P})_h$, we refer to [1]. Function $\mathcal{J}$ is the energy norm and its application in identification problems is thoroughly studied in [3] and [4].

2 Numerical methods

In what follows optimization Algorithm 1 will be exploited. There are two important steps: sensitivity analysis which relies on a differentiation of the cost function and enables to find a descend direction d^k in the $k - th$ iteration and finding a step length with a line search method in order to obtain a proper parameter α (in Newton method this step is omitted).

Data: p^0
Result: p^*
$k = 0$;
while $\left\|p^k - p^*\right\| > precision$ **do**
  sensitivity analysis $\rightarrow$ set the direction d^k;
  sufficient decrease of $\mathcal{J}\left(p^k + \alpha d^k\right)$ subject to $\alpha \in (0, 1\rangle$ (line search) ;
  set $p^{k+1} = p^k + \alpha d^k$;
  $k = k + 1$;
end
$p^k = p^*$;

Algorithm 1: The optimization algorithm

In case of the energy norm defining $\mathcal{J}$ one can derive the gradient of $\mathcal{J}$ for steepest descent method (SDM) so that $d^k = -\nabla\mathcal{J}\left(p^k\right)$, see e.g. [5].

$$\nabla\mathcal{J}\left(p^k\right) = \left(\frac{\partial\mathcal{J}}{\partial p_1}(p^k), \ldots, \frac{\partial\mathcal{J}}{\partial p_m}(p^k)\right)^T,$$

where

$$\frac{\partial\mathcal{J}}{\partial p_j}(p^k) = -\frac{1}{2}\left\langle \frac{\partial A}{\partial p_j}(p^k)(u(p^k) + \hat{u}), u(p^k) - \hat{u} \right\rangle.$$

In the case of Newton method, the Hessian matrix $H^k := H\left(p^k\right)$ at p^k of $\mathcal{J}$ has the following form:

$$H^k = \frac{1}{2} W^T A(p^k)^{-1} W \quad W = [w_1, \ldots w_m] \in R^{n \times m}, \ w_j = \frac{\partial A}{\partial p_j}(p^k) u(p^k)$$

The Newton step is given by a solution of the linear system: $H^k d^k = -\nabla \mathcal{J}\left(p^k\right)$ which is solved by the preconditioned conjugate gradient method with incomplete Cholesky factorization of the matrix A.
Nonlinear conjugate gradient (NCG) methods can take various forms, in our case Fletcher-Reeves formula was used, see e.g. [5]. NCG generates orthogonalized directions d^k.
Wolfe line search was chosen in case of SDM and NCG, for $\alpha^k \in \langle 0, 1\rangle$ and $c_1 = 10^{-4}$, $c_2 = 0.4$ (similar values as are suggested in [5]):

$$\mathcal{J}\left(p^k + \alpha^k d^k\right) \leq \mathcal{J}\left(p^k\right) + c_1 \alpha^k \nabla \mathcal{J}\left(p^k\right)^T d^k, \; \nabla \mathcal{J}\left(p^k + \alpha^k d^k\right) \geq c_2 \nabla \mathcal{J}\left(p^k\right)^T d^k$$

3 Numerical results

Three cases of a decomposition of $\Omega = \langle 0, 1\rangle \times \langle 0, 1\rangle$ marked by different colours can be seen in Fig 1. 1) domain A on the left corresponds to two parameters $p = (p_1, p_2)$, where p_1 is for the blue part and p_2 for the red part. 2) Checkerboard partition of the middle domain B. 3) The partition of C in addition contains small L-shaped regions of p_3 material. We present numerical results corresponding to these decompositions using a regular discretization 64×64 with 8192 elements.

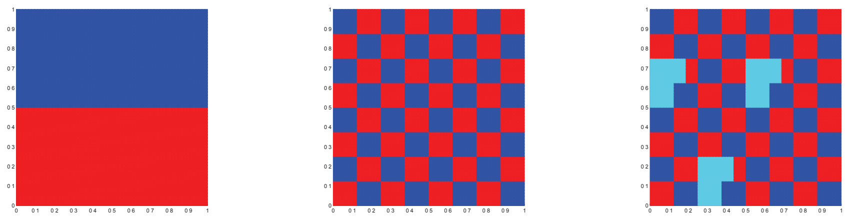
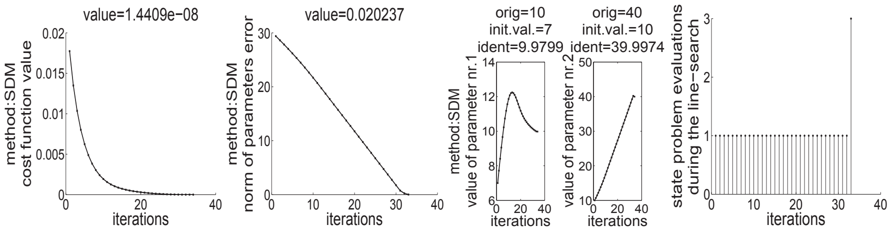


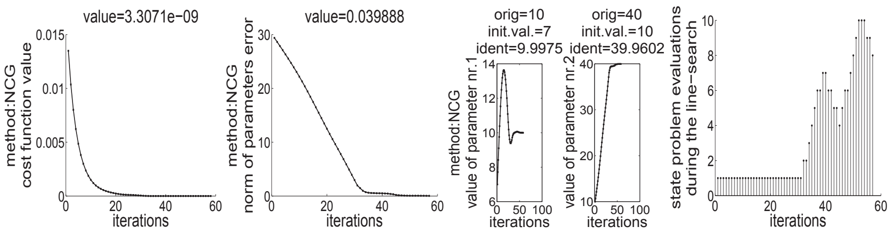
Fig. 1. Domain A on the left, domain B in the middle and domain C on the right side

Inverse problem in our examples consists of computation of the state problem in A and B with $p^* = (10, 40)$ and for C with $p^* = (10, 40, 17)$. The measurements $\hat{u}$ were taken in every node in the domain. The aim is to find the original values p^* based on $\hat{u}$ with initial values $p^0 = (7, 10)$ for the domains A and B, and $p^0 = (7, 10, 36)$ for the domain C.
Presented results are computed with the stopping criterion: $\left\|p^k - p^*\right\| \leq 4 * 10^{-2}$ to make the graphs visually clearer. Moreover, this accuracy turns out to be sufficient for more complicated numerical simulations.
Results in each figure are arranged in a way that all x-axes represent iterations and illustrate (a) the cost function minimization with its value at the last iteration above, (b) convergence of $\left\|p^k - p^*\right\|$ with the value at the last iteration above, (c) visualizes convergence of parameters p_i^k, $\forall i$, and (d) displays cost function evaluations needed at every iteration to satisfy Wolfe line search conditions.



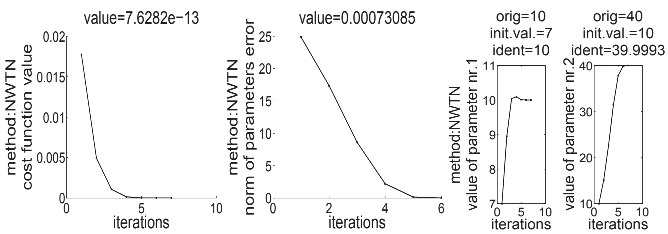
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 2. Steepest descent method on Domain A



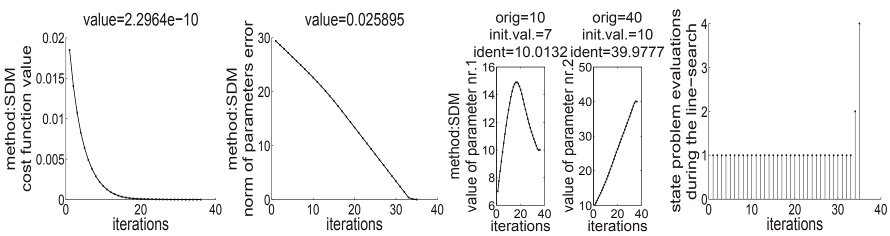
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 3. Nonlinear conjugate gradient method on Domain A



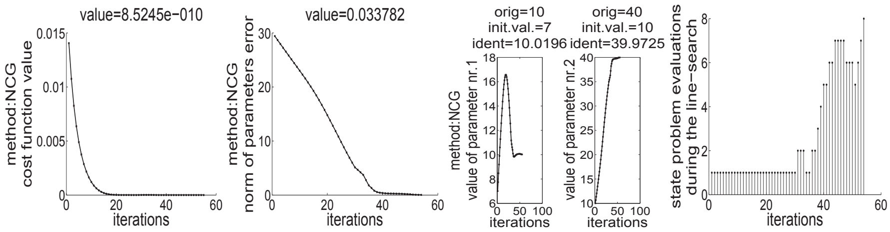
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv.

Fig. 4. Newton method on Domain A



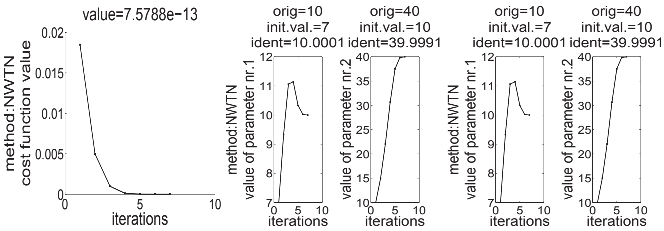
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 5. Steepest descent method on Domain B



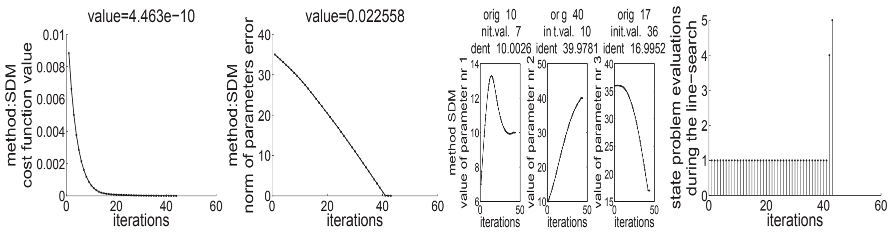
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 6. Nonlinear conjugate gradient method on Domain B



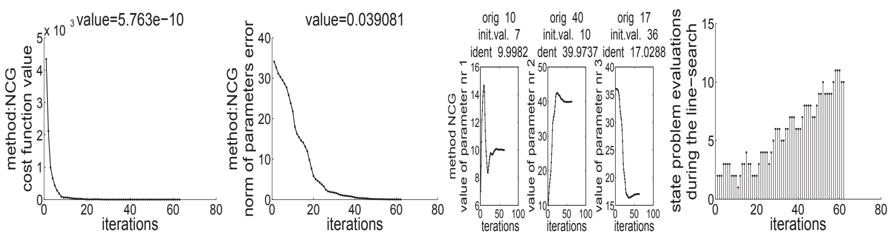
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv.

Fig. 7. Newton method on Domain B



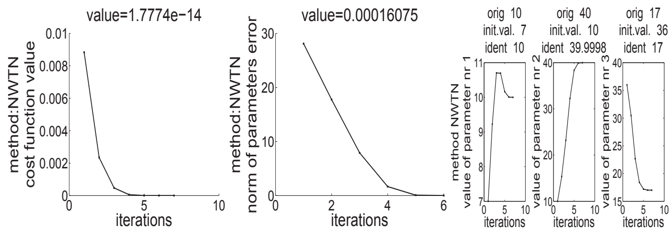
(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 8. Steepest descent method on Domain C



(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv. (d) fcn eval.

Fig. 9. Nonlinear conjugate gradient method on Domain C



(a) cost fcn val. (b) $\|p^k - p^*\|$ (c) param. conv.

Fig. 10. Newton method on Domain C

Time to convergence [s]	domain A	domain B	domain C
number of elements	8192	8192	8192
SDM	151	162.3	269.6
NCG	617.7	750	1163
Newton	16	15.3	21

3.1 Conclusions

The contribution is focused on the solution of inverse problems based on the energy norm and compares different optimization methods. the energy norm enables cheap computations of the gradient and Hessian of the cost function. SDM gave uniform parameter error convergence, smooth parameter convergence without significant oscillations and low number of line search rule violations. NCG is not so stable, it is more line search settings dependent. Both SDM and NCG methods deteriorate their efficiency when more complicated structure containing three parameters is computed. Newton method gave fast convergence which according to the theory seems to be quadratic even if the initial guess is not close to the solution.

Acknowledgement

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) funded by Structural Funds of the European Union and state budget of the Czech Republic.

References

1. R. Hrtus, R. Blaheta, J. Haslinger: Identification problems with a priori given material interfaces (2013)
2. Haslinger J.: Identification Methods. Manuscript (2014)
3. T.F. Chan, X. Tai: Identification of discontinuous coefficients in elliptic problems using total variation regularization (2003)
4. M. S. Gockenbach, A. A. Khan: An Abstract Framework for Elliptic Inverse Problems: Part 1. An Output Least-Squares Approach (2005)
5. J. Nocedal, S.J. Wright, Numerical Optimization, Springer 2006