http://dx.doi.org/10.1090/conm/218/03052

Contemporary Mathematics
Volume **218**, 1998
B 0-8218-0988-1-03052-1

Overlapping Domain Decomposition and Multigrid Methods for Inverse Problems

Xue-Cheng Tai, Johnny Frøyen, Magne S. Espedal, and Tony F. Chan

1. Introduction

This work continues our earlier investigations [**2**], [**3**] and [**8**]. The intention is to develop efficient numerical solvers to recover the diffusion coefficient, using observations of the solution u, from the elliptic equation

$$-\nabla \cdot (q(x)\nabla u) = f(x) \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega. \tag{1}$$

Our emphasis is on the numerical treatment of discontinuous coefficient and efficiency of the numerical methods. It is well known that such an inverse problem is illposed. Its numerical solution often suffers from undesirable numerical oscillation and very slow convergence. When the coefficient is smooth, successful numerical methods have been developed in [**5**] [**7**]. When the coefficient has large jumps, the numerical problem is much more difficult and some techniques have been proposed in [**2**] and [**3**]. See also [**9**], [**4**] and [**6**] for some related numerical results in identifying some discontinuous coefficients.

The two fundamental tools we use in [**2**] and [**3**] are the total variation (TV) regularization technique and the augmented Lagrangian technique. The TV regularization allows the coefficient to have large jumps and at the same time it will discourage the oscillations that normally appear in the computations. The augmented Lagrangian method enforces the equation constraint in an H^{-1} norm and was studied in detail in [**7**]. Due to the bilinear structure of the equation constraint, the augmented Lagrangian reduces the output-least-squares (OLS) minimization to a system of coupled algebraic equations. How to solve these algebraic equations is of great importance in speeding up the solution procedure. The contribution of the present work is to propose an overlapping domain decomposition (DD) and a multigrid (MG) technique to evaluate the H^{-1} norm and at the same time use them as a preconditioner for one of the algebraic equations. Numerical tests will be given to show the speed-up using these techniques.

1991 *Mathematics Subject Classification.* Primary 65N55; Secondary 35R30.

The fourth author was supported in part by NSF Grant ASC 9720257.

2. The augmented Lagrangian method

Let u_d be an observation for the solution u and $\vec{u}_g$ be an observation for the gradient ∇u, both may contain random observation errors. Due to the illposedness of the inverse problem, it is often preferable to use the OLS minimization to recover $q(x)$. Let us define $K = \{q| \quad q \in L^\infty(\Omega), 0 < k_1 \le q(x) \le k_2 < \infty\}$, with k_1 an k_2 known *a priori*, to be the admissible set for the coefficient. Let the mapping $e : K \times H_0^1(\Omega) \to H^{-1}(\Omega)$ be $e(q,u) = -\nabla \cdot (\nabla q \nabla u) - f$, which is the equation constraint. We shall use $R(q) = \int_\Omega \sqrt{|\nabla q|^2 + \epsilon dx}$, which approximate the TV-norm of $q(x)$, as the regularization term. In our experiments, the value of ϵ is always taken in $\epsilon \in [0.001, 0.01]$. The OLS minimization can be written

$$(2) \qquad \min_{e(q,u)=0,\ q\in K} \frac{1}{2}\|u-u_d\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla u - \vec{u}_g\|^2_{\mathbf{L}^2(\Omega)} + \beta R(q).$$

For more details on numerical approximations of the TV-norms, we refer to Chan and Tai [3]. As the inverse problem is illposed, its numerical solution is very sensitive to the observation errors. When the observation errors are very large, we must use proper noise removal procedure, see section 4 of [3] for the detailed algorithms that remove noise from the observations.

The Lagrangian method is often used for minimization problems with equality constraint. However, the augmented Lagrangian method is better when the minimization problem is illposed or the Hessian matrix of the cost functional has very small positive eigenvalues. For minimization (2), the associated augmented Lagrangian functional is

$$(3) \qquad \begin{aligned} L_r(q,u,\lambda) &= \frac{1}{2}\|u-u_d\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla u - \vec{u}_g\|^2_{\mathbf{L}^2(\Omega)} + \beta R(q) \\ &\quad + \frac{r}{2}\|e(q,u)\|^2_{H^{-1}(\Omega)} + (\lambda, e(q,u))_{H^{-1}(\Omega)}, \\ &\qquad \forall q \in K,\ u \in H_0^1(\Omega),\ \lambda \in H^{-1}(\Omega). \end{aligned}$$

The following algorithm is used to find a saddle point for $L_r(q,u,\lambda)$:

Algorithm

Step 1 Choose $u_0 \in H_0^1(\Omega), \lambda_0 \in H^{-1}(\Omega)$ and $r > 0$.

Step 2 Set $u_n^0 = u_{n-1}$. For $k = 1, 2, \ldots, k_{max}$, do:

Step 2.1 Find $q_n^k \in K$ such that

$$(4) \qquad q_n^k = \arg\min_{q\in K} L_r(q, u_n^{k-1}, \lambda_{n-1}).$$

Step 2.2 Find $u_n^k \in H_0^1(\Omega)$ such that

$$(5) \qquad u_n^k = \arg\min_{u\in H_0^1(\Omega)} L_r(q_n^k, u, \lambda_{n-1}).$$

Step 3 Set $u_n = u_n^k, q_n = q_n^k$, and update λ_n as $\lambda_n = \lambda_{n-1} + re(q_n, u_n)$.

In our simulations, we take $k_{max} = 2$. The above algorithm has a linear rate of convergence, see [7]. Second order scheme can also be used to search for a saddle point, see [7, p.98]. If (q^*, u^*, λ^*) is a saddle point of L_r, then (q^*, u^*) is a minimizer of (2). For fixed u_n^{k-1} and q_n^k, the minimization problems of Step 2.1 and Step 2.2 are equivalent to two algebraic equations. See [3] for the detailed matrix representation of the corresponding algebraic equations. Problems (4) and (5) are solved by direct solver in [2] and [3]. The numerical accuracy and the executing time is superior to earlier literature results. However, we must improve the efficiency and use some iterative solvers in order to be able to solve real life large size problems. Without using iterative solvers, the memory limit will prevent us from doing simulations for real life three dimensional problems. To use an iterative solver, the rate of convergence of the iterative solver is of great concern for the efficiency of the whole algorithm. Moreover, the way that the H^{-1}-norm is evaluated is very critical in avoiding the solving of large size sparse matrices in the iterative procedure. Let Δ denote the Laplace operator, which is a homeomorphism from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$. It is true that $\|\Delta^{-1}f\|_{H_0^1(\Omega)} = \|f\|_{H^{-1}(\Omega)}$. Thus, we need to invert a sparse matrix Δ in order to compute the H^{-1}-norm. However, we can obtain the H^{-1}-norm of f by inverting smaller size matrices with the domain decomposition methods or using multigrid type methods to avoid inverting any matrices by using the theorem of the next section.

3. Space decomposition methods

Recent research reveals that both domain decomposition and multigrid type methods can be analysed using the frame work of space decomposition and subspace correction, see Chan and Sharapov [1], Tai and Espedal [11] [12], Tai [10] and Xu [14]. In this section, we show that we can use them to evaluate the H^{-1}-norm.

We present the results for a general Hilbert space and for general space decomposition techniques. For a given Hilbert space V, we denote V^* as its dual space and use $(\cdot,\cdot)$ to denote its inner product. Notation $\langle\cdot,\cdot\rangle$ is used to denote the duality pairing between V and V^*. We consider the case where V can be decomposed as a sum of subspaces:

$$V = V_1 + V_2 + \cdots + V_m.$$

Moreover, we assume that there is a constant $C_1 > 0$ such that $\forall v \in V$, we can find $v_i \in V_i$ that satisfy:

$$v = \sum_{i=1}^{m} v_i \ , \qquad \text{and} \qquad \left(\sum_{i=1}^{m} \|v_i\|_V^2\right)^{\frac{1}{2}} \le C_1 \|v\|_V \tag{6}$$

and there is an $C_2 > 0$ such that

$$\sum_{i=1}^{m}\sum_{j=1}^{m}(v_i, v_j) \le C_2\left(\sum_{i=1}^{m} \|v_i\|_V^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{m} \|v_i\|_V^2\right)^{\frac{1}{2}} ,\forall v_i \in V_i \quad \text{and} \quad \forall v_j \in V_j \ . \tag{7}$$

THEOREM 1. *Assume the decomposed spaces satisfy (6) and (7), then*

$$\frac{\|f\|_{V^*}}{C_1} \le \left(\sum_{i=1}^{m} \|f\|_{V_i^*}^2\right)^{\frac{1}{2}} \le C_2 \|f\|_{V^*} \qquad \forall f \in V^* \subset V_i^*. \tag{8}$$

Details of the proof of the above theorem will be given in a forthcoming paper. The theorem shows that in order to get $\|f\|_{V^*}$, we just need to use some parallel processors to compute $\|f\|_{V_i^*}^2$ and sum them together. For domain decomposition methods, we need to invert some smaller size matrices to get $\|f\|_{V_i^*}^2$. For multigrid methods, no matrices need to be inverted.

4. Numerical Tests

Let $\Omega = [0,1] \times [0,1]$. For a given f and piecewise smooth q, we compute the true solution from (1) and get its gradient ∇u. Let R_d and $\vec{R}_g$ be vectors of random numbers between $[-1/2, 1/2]$. The observations are generated by

$$u_d = u + \delta R_d \|u_d\|_{L^2(\Omega)}, \qquad \vec{u}_g = \nabla u + \delta \vec{R}_g \|\vec{u}_g\|_{\mathbf{L}^2(\Omega)} \ . \tag{9}$$

We shall use finite element (FE) approximations. The domain Ω is first divided into subdomains $\Omega_i, i = 1, 2, \cdots, m$ with diameters of the size H, which will also be used as the coarse mesh elements. Each subdomain is refined to form a fine mesh division for Ω of mesh parameter h $(h << H)$. Each subdomain Ω_i is extended by a size $\delta = cH$ $(0 < c < 1)$ to get overlapping subdomains Ω_i^δ. Let $S_0^h(\Omega)$, $S_0^h(\Omega_i^\delta)$ and $S_0^H(\Omega)$ be the bilinear FE spaces with zero traces on the corresponding boundaries on the fine mesh, subdomain Ω_i^δ and coarse mesh respectively. It is true that

$$S_0^h(\Omega) = S_0^H(\Omega) + \sum_{i=1}^{m} S_0^h(\Omega_i^\delta).$$

For the above decomposition, the constants C_1 and C_2 do not depend on the mesh parameters h and H, see [**14**]. Estimate (8) shows that we only need to invert the matrices associated with the subdomains and the coarse mesh to get the H^{-1} norms.

In order to use multigrid type techniques, we take Ω as the coarsest mesh and use rectangular elements. At a given level, we refine each element into four elements by connecting the midpoints of the edges of the rectangles of a coarser grid. Starting from Ω and repeating the above procedure J times, we will get J levels of meshes. Let $V_k, k = 1, 2, \cdots, J$ be the bilinear FE spaces over the levels and denote $\{\phi_i^k\}_{i=1}^{n_k}$ the interior nodal basses for the k^{th} level FE space, it is easy to see that

$$V = \sum_{k=1}^{J} \sum_{i=1}^{n_k} V_i^k \quad \text{with} \quad V = V_J, V_i^k = \text{span}(\phi_i^k).$$

For the multigrid decomposition, the subspaces V_i^k are one dimensional and the constants C_1, C_2 are independent of the mesh parameters and the number of levels. No matrix need to be inverted to get the H^{-1} norms.

The bilinear FE spaces introduced above will be used as the approximation spaces for u and λ. Piecewise constant FE functions on the fine mesh are used to approximate the coefficient q. For a given q, u and λ, let $B(q, u, \lambda)$ and $A(q, \lambda)$ be the matrices that satisfy

$$\partial L_r(q, u, \lambda)/\partial q = B(q, u, \lambda)q, \qquad \partial L_r(q, u, \lambda)/\partial u = A(q, \lambda)u.$$

Let $B_n^k = B(q_n^k, u_n^{k-1}, \lambda_{n-1})$ and $A_n^k = A(q_n^k, \lambda_{n-1})$. Assume that the solution of (4) is in the interior of K, then (4) and (5) are equivalent to solving

$$a) \;\; B_n^k q_n^k = \alpha_n^k, \qquad b) \;\; A_n^k u_n^k = \beta_n^k \tag{10}$$

with some known vector α_n^k and β_n^k. Due to the regularization term $R(q)$, the matrix B_n^k depends on q_n^k. A simple linearization procedure is employed to deal with the nonlinearity, see [**3, 13**]. If we use conjugate gradient (CG) method to solve the equations (10), it is not necessary to know the matrices, we just need to calculate the product of the matrices with given vectors. It is easy to see that the equations in (10) have symmetric and positive define matrices. We use CG without preconditioner to solve (10.a) and use a preconditioned CG to solve (10.b). The stopping criteria for the CG iterations is that the residual has been reduced by a factor of 10^{-10} or the iteration number has reached 300 (In Table 2 the maximum iteration number is 5000 in order to see the CPU time usage for bad β). The

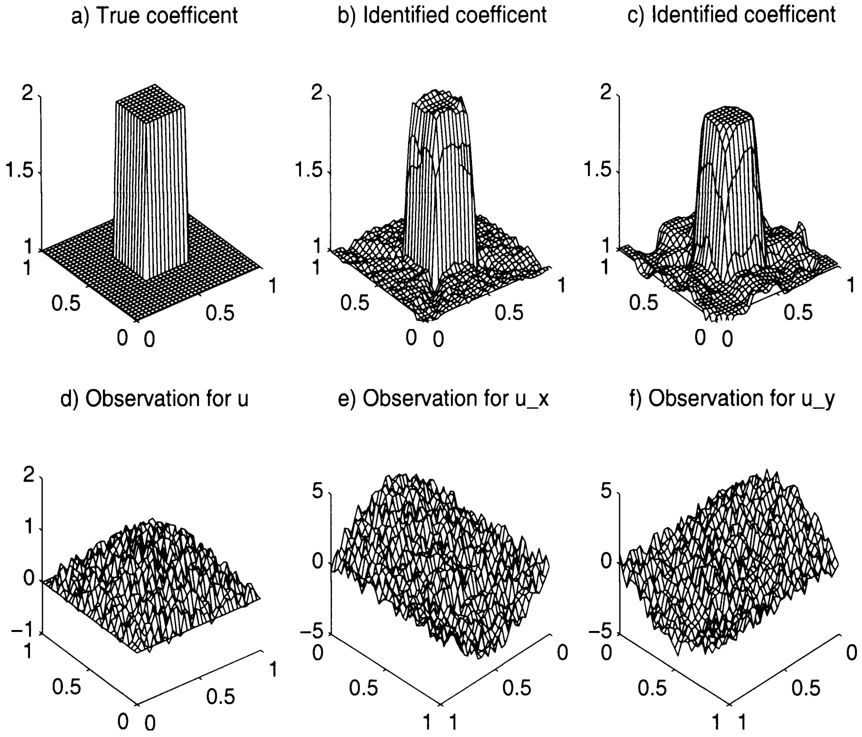

FIGURE 1. The identified coefficients and the observation data.

constants k_1 and k_2 used for defining the admissible coefficient set are taken to be $k_1 = 0,\ k_2 = \infty$. The simulations are tested with a sequential machine in programming language C++. Three different methods are used for the preconditioner and for evaluating the H^{-1} norms:

1. Do LU decomposition (LU-D) for the Laplace operator and use it as the preconditioner and also for the H^{-1} norms.
2. Use domain decomposition for the preconditioning and also for the H^{-1} norms.
3. Use multigrid for the preconditioning and also for the H^{-1} norms.

As the tests are done with a sequential machine, the multiplicative version of the MG and DD methods are used. In Table 1, the CPU time for different iteration numbers is given. We have used $\beta = 0.00125$ and $h = /128$. Here and later h and nx denote the mesh size and number of elements used both for the x- and y-directions, respectively. We observe that the domain decomposition approach and the multigrid approach are much faster than the LU-decomposition. The multigrid method is slightly better than the domain decomposition method. The identified coefficients and the observations are shown in Figure 1. In identifying the coefficient in subfigure 1.b), we have added 10% of noise and used $h = 1/128, \beta = 0.00025, r = 100$. At iteration 20, $\|q_n - q\|_{L^2(\Omega)} = 0.0631$ and $\|e(q_n, u_n)\|_{L^2(\Omega)} = 1.5 \times 10^{-7}$. In subfigure 1.c), the identified coefficient is with noise level $\delta = 100\%$ and we have used $h = 1/128, \beta = 0.03125, r = 100$. At iteration 20, $\|q_n - q\|_{L^2(\Omega)} = 0.1013$ and $\|e(q_n, u_n)\|_{L^2(\Omega)} = 1.1 \times 10^{-6}$.

TABLE 1. CPU time (in sec.) versus iteration with $\beta = 0.00125$, $h = 1/128$.

iter =	MG	DD	LU-D
1	31.87	94.04	942.76
3	121.27	370.15	3350.17
5	221.68	668.06	5999.82
7	327.14	954.52	8749.58
9	433.74	1244.67	11509.80
11	541.47	1538.97	14227.90
12	653.80	1836.38	16935.80
15	777.22	2137.01	19628.30
17	885.29	2439.20	22309.40
20	1043.60	2887.72	26300.10

TABLE 2. CPU time (in Sec.) versus β with $h = 1/64$, *iteration* = 20.

$\beta =$	MG	DD	LU-D
0.00001	1309.67	3270.60	5009.87
0.00005	544.13	1320.36	2146.69
0.00025	256.89	758.76	1107.20
0.00125	173.72	431.31	747.76
0.00625	172.15	459.70	728.08
0.03125	224.22	566.91	922.79

TABLE 3. CPU time (in Sec.) versus $h = 1/nx$ with $\beta = 0.00125$, *iteration* = 20.

$nx =$	MG	DD	LU-D
16	6.06	16.83	9.88
32	24.56	60.89	61.71
64	171.76	406.91	720.45
128	1393.95	3859.8	35258.20

The regularization parameter β is introduced to prevent numerical oscillations. If it is chosen to be big, the discontinuity is smeared out and large errors are introduced. If it is chosen to be too small, it can not control the numerical oscillations and so prevent us from getting accurate numerical solutions. From our numerical tests, we find that the value of β is also of critical importance for the rate of convergence for the CG method. In Table 2, the CPU time in seconds for different values of β is compared for the three different approaches. It is clear that very small or very large β increases the computing time.

Table 3 is used to show the CPU time usage for different mesh sizes h. Let us note that the finest mesh is of size $h = 1/128$ with a total number of grid points $128 \times 128 \approx 2 \times 10^4$. For inverse problems we considered here, there are not many numerical approaches that can handle such a large number of unknowns. It shall also be noted that 100% of observation errors are added to the observations, i.e. $\delta = 100\%$ in (9), see d) e) f) of Figure 1.

References

1. T. F. Chan and I. Sharapov, *Subspace correction multilevel methods for elliptic eigenvalue problems*, Proceedings of the 9th international domain decomposition methods (P. Bjørstad, M. Espedal, and D. Keyes, eds.), John Wiley and Sons, To appear.
2. T. F. Chan and X.-C. Tai, *Augmented Lagrangian and total variation methods for recovering discontinuous coefficients from elliptic equations*, Computational Science for the 21st Century (M. Bristeau, G. Etgen, W. Fitzgibbon, J. L. Lions, J. Periaux, and M. F. Wheeler, eds.), John Wiley & Sons, 1997, pp. 597–607.
3. ———, *Identification of discontinuous coefficient from elliptic problems using total varaition regularization*, Tech. Report CAM-97-35, University of Carlifornia at Los Angeles, Department of Mathematics, 1997.
4. Z. Chen and J. Zou, *An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems*, Tech. Report Report 97-06, The Chinese University of Hong Kong, 1997.
5. K. Ito and K. Kunisch, *The augmented Lagrangian method for parameter estimation in elliptic systems*, SIAM J. Control Optim. **28** (1990), 113–136.
6. Y. L. Keung and J. Zou, *Numerical identifications of parameters in parabolic systems*, Tech. Report Report 97-16, The Chinese University of Hong Kong, 1997.
7. K. Kunisch and X.-C. Tai, *Sequential and parallel splitting methods for bilinear control problems in Hilbert spaces*, SIAM J. Numer. Anal. **34** (1997), 91–118.
8. ———, *Nonoverlapping domain decomposition methods for inverse problems*, Proceedings of the 9th international domain decomposition methods (P. Bjørstad, M. Espedal, and D. Keyes, eds.), John Wiley and Sons, To appear.
9. T. Lin and E. Ramirez, *A numerical method for parameter identification of a two point boundary value problem*, 1997.
10. X.-C. Tai, *Parallel function and space decomposition methods*, The Finite element Method: Fifty years of the Courant Element (P. Neittaanmaki, ed.), Lecture notes in pure and applied mathematics, vol. 164, Marcel Dekker, 1994, pp. 421–432.
11. X.-C. Tai and M. Espedal, *Rate of convergence of a space decomposition method for linear and nonlinear problems*, SIAM J. Numer. Anal. (1998), to appear.
12. ———, *A space decomposition method for minimization problems*, Proceedings of the 9th international domain decomposition methods (P. Bjørstad, M. Espedal, and D. Keyes, eds.), John Wiley and Sons, To appear.
13. C. Vogel and M. Oman, *Iterative methods for total variation dennoising*, SIAM J. Sci. Comp. **17** (1996), 227–238.
14. J. C. Xu, *Iteration methods by space decomposition and subspace correction*, SIAM Rev. **34** (1992), 581–613.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BERGEN, JOHANNES BRUNSGATE 12, 5008, BERGEN, NORWAY
E-mail address: Xue-Cheng.Tai@mi.uib.no
URL: http://www.mi.uib.no/~tai

RF-ROGALAND RESEARCH, THORMØLLENSGT. 55, 5008 BERGEN, NORWAY
E-mail address: Johnny.Froyen@rf.no

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BERGEN, JOHANNES BRUNSGATE 12, 5008, BERGEN, NORWAY
E-mail address: Magne.Espedal@mi.uib.no
URL: http://www.mi.uib.no/~resme

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, 405 HILGARD AVENUE, LOS ANGELES, CA 90095-1555
E-mail address: chan@math.ucla.edu
URL: http://www.math.ucla.edu/~chan