EFFICIENT PRECONDITIONERS FOR OPTIMALITY SYSTEMS ARISING IN CONNECTION WITH INVERSE PROBLEMS

BJØRN FREDRIK NIELSEN AND KENT-ANDRE MARDAL

Abstract. This paper is devoted to the numerical treatment of linear optimality systems (OS) arising in connection with inverse problems for partial differential equations. If such inverse problems are regularized by Tikhonov regularization, then it follows from standard theory that the associated OS is well-posed, provided that the regularization parameter α is positive and that the involved state equation satisfies suitable assumptions.

We explain and analyze how certain mapping properties of the operators appearing in the OS can be employed to define efficient preconditioners for finite element (FE) approximations of such systems. The key feature of the scheme is that the number of iterations needed to solve the preconditioned problem by the minimal residual method is bounded independently of the mesh parameter h, used in the FE discretization, and only increases moderately as $\alpha \to 0$. More specifically, if the stopping criterion for the iteration process is defined in terms of the associated energy norm, then the number of iterations required (in the severely ill-posed case) cannot grow faster than $O((\ln(\alpha))^2)$. Our analysis is based on a careful study of the involved operators which yields the distribution of the eigenvalues of the preconditioned OS.

Finally, the theoretical results are illuminated by a number of numerical experiments addressing both a model problem studied by Borzi, Kunisch and Kwak [14] and an inverse problem arising in connection with electrocardiography [35].

1. Introduction

Let H_1, H_2 and H_3 be Hilbert spaces with inner products $(\cdot,\cdot)_{H_1}$, $(\cdot,\cdot)_{H_2}$, $(\cdot,\cdot)_{H_3}$, norms $\|\cdot\|_{H_1}$, $\|\cdot\|_{H_2}$, $\|\cdot\|_{H_3}$ and dual spaces H_1', H_2' and H_3'. We will consider parameter identification problems which can be written on the form

$$\min_{v\in H_1} \left\{ \frac{1}{2}\|Tu-d\|^2_{H_3} + \frac{1}{2}\alpha \|v-v_{\text{prior}}\|^2_{H_1} \right\} \tag{1}$$

subject to

$$Au = -Bv + g \quad \text{(state equation)}, \tag{2}$$

Key words and phrases. All-at-once methods, inverse problems, optimality systems, preconditioning, minimal residual method, isolated eigenvalues, Tikhonov regularization, electrocardiography.

An example

$B = \bar{H} \cup T$

Want to use an observation d of $u|_{\partial B}$

and

$$\nabla \cdot (M \nabla u) = \begin{cases} 0 & \text{in } T, \\ -\nabla \cdot (M_i \nabla v) & \text{in } H, \end{cases}$$

to compute v

Output least squares

regularization

$$\min_{v \in H^1(H)} \left\{ \frac{1}{2} \| u - d \|^2_{L^2(\partial B)} + \frac{1}{2} \alpha \| v - v_{prior} \|^2_{H^1(H)} \right\}$$

subject to

$$\int_B (M \nabla u) \cdot \nabla \phi \, dx = - \int_H (M_i \nabla v) \cdot \nabla \phi \, dx \quad \forall \phi \in H^1(B)$$

Recall that:

$$\min_{x,y} f(x,y)$$

Subject to $g(x,y) = 0$.

Lagrangian $L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$

Necessary conditions

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial y} = 0, \quad \frac{\partial L}{\partial \lambda} = 0$$

give

$$f_x(x,y) + \lambda g_x(x,y) = 0$$
$$f_y(x,y) + \lambda g_y(x,y) = 0$$
$$g(x,y) = 0$$

Lagrangian

$$L_\alpha(v, u, w) = \frac{1}{2}\|u-d\|^2_{L^2(\partial B)} + \frac{1}{2}\alpha\|v - v_{prior}\|^2_{H^1(H)}$$

$$+ \int_B (M\nabla u)\cdot\nabla w\,dx + \int_H (M_i\nabla v)\cdot\nabla w\,dx$$

for $v \in H^1(H)$ and $u, w \in H^1(B)$

Necessary conditions for a minimum

$$\frac{\partial L_\alpha}{\partial v} = 0, \quad \frac{\partial L_\alpha}{\partial u} = 0, \quad \frac{\partial L_\alpha}{\partial w} = 0$$

give a 3×3 system of PDEs.

Find v, u and w such that

$$\alpha (v, \phi)_{H^1(H)} + \int_H \nabla w \cdot (M_i \nabla \phi)\, dx = \alpha (v_{prior}, \phi)_{H^1(H)} \quad \forall \phi \in H^1(H)$$

$$(u, \phi)_{L^2(\partial B)} + \int_B \nabla w \cdot (M \nabla \phi) dx = (d, \phi)_{L^2(\partial B)} \quad \forall \phi \in H^1(B)$$

$$\int_H \nabla \phi \cdot (M_i \nabla v)\, dx + \int_B \nabla \phi \cdot (M \nabla u)\, dx = 0 \qquad \forall \phi \in H^1(B)$$

Optimality system, saddle point problem

$$\underbrace{\begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix}}_{\mathcal{A}_\alpha} \underbrace{\begin{bmatrix} v \\ u \\ w \end{bmatrix}}_{P} = b_\alpha$$

$$\mathcal{A}_\alpha : H^1(H) \times H^1(B) \times H^1(B) \longrightarrow \left(H^1(H) \times H^1(B) \times H^1(B)\right)'$$

Preconditioner

$$\mathcal{B}_\alpha : \left(H^1(H) \times H^1(B) \times H^1(B)\right)' \longrightarrow H^1(H) \times H^1(B) \times H^1(B).$$

well-behaved.

For example

$$\mathcal{B}_\alpha^{-1} = \begin{bmatrix} \alpha L & 0 & 0 \\ 0 & \alpha A + K & 0 \\ 0 & 0 & \frac{1}{\alpha} A \end{bmatrix}$$

$$\mathcal{B}_\alpha \mathcal{A}_\alpha p = \mathcal{B}_\alpha b_\alpha \qquad (1)$$

$$\mathcal{B}_\alpha \mathcal{A}_\alpha : \left(H^1(H) \times H^1(B) \times H^1(B)\right) \to \left(H^1(H) \times H^1(B) \times H^1(B)\right)$$

can solve (1) with an <u>iterative scheme</u>

- $\mathcal{B}_\alpha A_\alpha$ is well-behaved as the mesh parameter $h \to 0$

- Condition number

$$\kappa(\mathcal{B}_\alpha A_\alpha) = \mathcal{O}\left(\frac{1}{\alpha}\right),$$

but a <u>very</u> <u>limited</u> ~~auto~~ ~~ei~~ number of eigenvalues are close to zero.

Main result

The number of iterations needed by the minimal residual method cannot grow faster than $O([\ln(\alpha)]^2)$ as $\alpha \to 0$.

subject to: Find $\bar{u} \in \bar{H}^1(B)$ such that

$$\int_B (\mathbf{M}\nabla \bar{u}) \cdot \nabla \phi \, dx = \int_H (\mathbf{M}_i \nabla v) \cdot \nabla \phi \quad \text{for all } \phi \in \bar{H}^1(B) \, dx, \tag{38}$$

where

$$\bar{H}^1(B) = \left\{ \psi \in H^1(B); \int_{\partial B} T\psi \, dx = 0 \right\}.$$

More precisely, $(v, \bar{u}) \in H^1(H) \times \bar{H}^1(B)$ solves (37)-(38) if and only if $(v, u) = (v, \bar{u} + \frac{1}{|\partial B|} \int_{\partial B} d \, dx) \in H^1(H) \times H^1(B)$ solves (32)-(33). A similar connection between the optimality systems associated with (37)-(38) and (32)-(33) can, of course, also be established. We will not dwell any further upon this issue.

3.2.1. *Numerical results.* The shapes of the heart H and the body B are not simple. Consequently, we used nonuniform meshes in the FE discretization procedure. In all the tables presented in connection with (32)-(33), l represents the refinement level of the grid. More precisely, as l increases the mesh size h decreases. (l is the number of times an initial coarse mesh has been refined).

Also for this example the iteration counts obtained with the multigrid preconditioner $\widehat{\mathcal{B}_\alpha}$ and the standard stopping criterion (29) are well-behaved, see Table 4. Indeed, the number of iterations needed seems to be bounded independently of both h and α.

$l \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
0	32	40	55	42	25
1	28	36	49	52	24
2	26	30	41	51	26
3	28	28	36	47	32
4	29	28	32	41	41

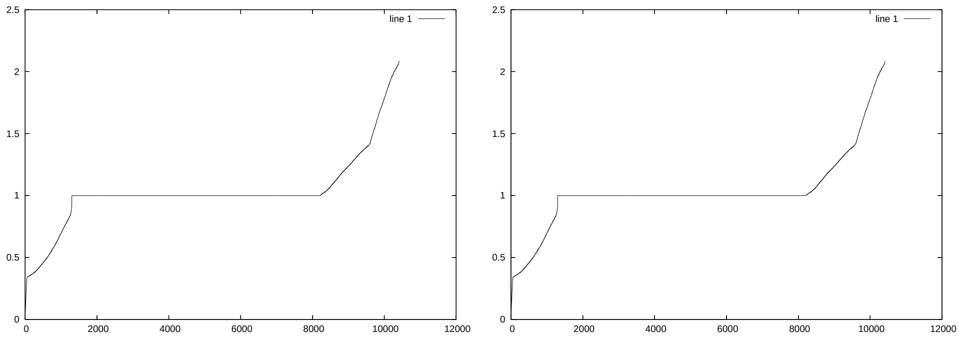
TABLE 4. The number of preconditioned minimal residual iterations needed to solve the model problem studied in Example 2. These results were generated with the energy stopping criterion (29). Here, l is the refinement level of the grid, i.e. the mesh size h decreases as l increases.

According to Table 5, the condition number $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$ of $\mathcal{B}_\alpha \mathcal{A}_\alpha$ seems to be approximately of order $O(\alpha^{-1})$. Hence, the results reported in Table 4 are far better than what one would expect from the standard estimate (30). As observed in Example 1, almost all of the eigenvalues of $\mathcal{B}_\alpha \mathcal{A}_\alpha$ are of order $O(1)$, see Figure 3. We will return to this issue in the next section.

If the α independent stopping rule (31) is used, then the workload increases as α decreases, see Table 6. Nevertheless, the number of iterations needed does not "explode" for small values of α. (The results presented in Table 6 were generated with $p^* = 0$ and a random initial guess p_0 for the minimal residual method).

$l \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
1	16	108	672	5000	29729
2	16	109	680	5076	40157

Table 5. This table contains numerical results obtained in Example 2. More precisely, the condition number $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$ of $\mathcal{B}_\alpha \mathcal{A}_\alpha$ for various grid refinement levels l and $\alpha = 1,\ 10^{-1},\ 10^{-2},\ 10^{-3},\ 10^{-4}$. (The mesh size h decreases as l increases).



(a) $\alpha = 10^{-1}$ (b) $\alpha = 10^{-2}$

Figure 3. The absolute value of the eigenvalues of $\mathcal{B}_\alpha \mathcal{A}_\alpha$ on mesh level $l = 2$ with regularization parameter $\alpha = 10^{-1},\ 10^{-2}$. These results were obtained for the model problem studied in Example 2.

$l \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}
0	32	100	358	588
1	28	71	237	771
2	26	54	188	895
3	28	53	179	688
4	29	46	150	494

Table 6. The number of preconditioned minimal residual iterations needed to solve the model problem studied in Example 2. These results were generated with the α independent stopping criterion (31). For $\alpha \leq 10^{-4}$ instabilities occurred. Here, l is the refinement level of the grid, i.e. the mesh size h decreases as l increases.

4. Theoretical considerations

This section is devoted to a theoretical study of the preconditioning strategy proposed and tested above. We have seen that the standard estimate (30) cannot explain the results presented in tables 1 and 4. In order to analyze these observations, we will show that $\mathcal{A}_\alpha$ is bounded, that the Babuška-Brezzi conditions hold and characterize the eigenvalues of the preconditioned