Efficient preconditioning of optimality systems

Kent-Andre Mardal, Bjørn Fredrik Nielsen and Martin S. Alnæs

Center for Biomedical Computing

10th August 2012

Outline

- Abstract framework for preconditioning
- A few examples (elliptic and Stokes problems)
- The problem with inverse problems in an abstract setting
- The solution for inverse problems in an abstract setting
- Two examples, one severly ill-posed and a comparison of methods

(I will also show some FEniCS code during the talk)

Abstract Framework based on Functional Analysis

Let us consider a well posed PDE problem:
Find $u \in V$ such that for $f \in V^*$

$$Au = f,$$

Here $A : V \to V^*$ is well posed in the sense that the solution admit the following:

$$\|u\|_V \leqslant C_1 \|f\|_{V^*}$$

However, a discretization of the problem leads to

$$\mathrm{cond}(A_h) \to \infty \text{ as } h \to 0.$$

The discretization ruins the nice property of the continnuous problem!

Abstract Framework based on Functional Analysis

The problem is that

$$A : V \to V^*$$

is bounded, while the matrix

$$A_h : \mathbb{R}^n \to \mathbb{R}^n$$

is unbounded. However, introducting a preconditioner we

$$B : V^* \to V$$

we obtained that

$$BA : V \to V$$

is bounded and the corresponding preconditioned problem

$$B_h A_h : \mathbb{R}^n \to \mathbb{R}^n$$

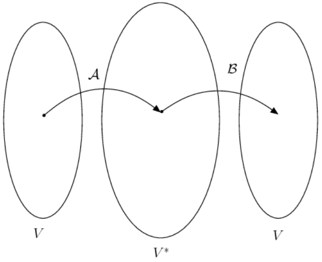
has bounded spectrum.

Abstract Framework based on Functional Analysis

If A is an elliptic operator in either L_2, $H(\text{div})$, $H(\text{curl})$, H^1 or H^2 then norm equivalent preconditioners, i.e.,

$$c_0(Au, u) \leqslant (ABAu, u) \leqslant c_1(Au, u) \quad \forall u$$

are well-known! These preconditioners are norm equivalent to the Riesz mapping (multigrid, domain decomposition etc.).



[Mardal, Winther NLAA 2011, Hiptmair Comp. & Math. with Appl. 2006, Kirby SIAM Review 2011 Arnold, Falk, Winther MMAN and Math. Comp. 1997]

Example: An elliptic problem

Consider an elliptic problem:
Find $u \in H_0^1$ such that for $f \in H^{-1}$

$$Au = -\nabla \cdot (K \nabla u) = f$$

Here, K positive definite and bounded.

The Riesz mapping is $B = \Delta^{-1}$ and the spectrum of BA is bounded by the extreme values of K.

Multigrid and domain decomposition give efficient operators that are equivalent with Δ^{-1}.

Example: Stokes problem

Consider the Stokes problem: Find $u, p \in H^1_0 \times L^2_0$ such that for $f \in H^{-1}$

$$A \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} -\Delta & -\nabla \\ \nabla\cdot & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}$$

The Riesz mapping B taking $H^{-1} \times L^2_0 \to H^1_0 \times L^2_0$ is

$$B = \begin{bmatrix} -\Delta^{-1} & 0 \\ 0 & I \end{bmatrix}$$

The spectrum of BA is bounded (even though B is very different from A)!

Same as Silvester, Wathen 94, Rusten, Winther 92 + many more, but put in a functional analysis setting

It is easy to construct spectrally equivalent and efficient versions of B with multigrid and domain decomposition techniques

Corresponding code in FEniCS

```
v,u = TestFunction(V), TrialFunction(V)
q,p = TestFunction(Q), TrialFunction(Q)

A   = assemble(inner(grad(v), grad(u))*dx)
B   = assemble(div(v)*p*dx)
C   = assemble(div(u)*q*dx)
D = assemble(p*q*dx)

AA = block_mat([[A, B],
                [C, 0]])

BB = block_mat([[ML(A),  0],
                [0, ML(D)]])
# (also create b and enforce bc)

AAinv = MinRes(AA, precond=BB, tolerance=1e-8)
x = AAinv * b
```

The problem with inverse problems

Let us consider an abstract inverse problem: Find $u \in V$ such that for $f \in V^*$

$$Au = f$$

The problem is <u>not</u> well-posed

$$\|A\|_{L(V,V^*)} \leqslant C_1 \;\text{ but }\; \|A^{-1}\|_{L(V^*,V)} \to \infty$$

A has a accumulation point at zero!

A few eigenvalues outside a clustering is not necessarily a bad thing for Krylov solvers (c.f. O. Axelsson and G. Lindskog, Numer. Math. 1986))!

Parameter identification problem

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

  subject to

$$Au = -Bv + g,$$

- Bounded linear operators:

$$\begin{aligned} A &: H_2 \to H_2^*, \quad \text{continuously invertible} \\ B &: H_1 \to H_2^*, \\ T &: H_2 \to H_3 \quad \text{observation operator} \\ L &: H_1 \to H_1^* \quad \text{regularization operator} \end{aligned}$$

Optimality system

- $$\begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} \alpha L v_{\mathrm{prior}} \\ Qd \\ g \end{bmatrix}$$
- $K : H_2 \to H_2^*, \quad u \to (Tu, T\phi)_{H_3} = (T^* Tu, \phi)_{H_2}$
- Typically ill-posed for $\alpha = 0$

Optimality system, cont.

- $A = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix} : X \times Y \to (X \times Y)^*$
- $X = H_1 \times H_2$
  - $\|x\|_X^2 = \alpha \|x_1\|_{H_1}^2 + \alpha \|x_2\|_{H_2}^2 + (T^* T x_2, x_2)_{H_2}$
- $Y = H_2$
  - $\|y\|_Y^2 = \frac{1}{\alpha} \|y\|_{H_2}^2$

Schöberl and Zulehner, SIAM J. Matrix Anal., 2007 added observations to the space of Lagrange multipliers (Y).

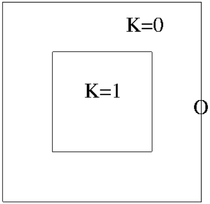
Example 1

$$\min_{v \in L^2(\Omega)} \left\{ \frac{1}{2} \|u - d\|^2_{L^2(\Omega)} + \frac{1}{2}\alpha \|v\|^2_{L^2(\Omega)} \right\}$$

subject to

$$\begin{aligned} u - \Delta u &= v \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Example 2



$$\min_{v \in L^2(\Omega)} \left\{ \frac{1}{2} \|u - d\|^2_{L^2(O)} + \frac{1}{2} \alpha \|v\|^2_{L^2(\Omega)} \right\}$$

subject to

$$\begin{aligned} u - \Delta u &= Kv \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Extensions for parameter dependent problems: Weighted Sobolev spaces

Consider the problem: Find $u \in H_0^1$, for $f \in H^{-1}$

$$A_\alpha u = u - \alpha^2 \Delta u = f$$

Here, $\alpha > 0$

$$\|A_\alpha^{-1}\|_{L(H^{-1}, H_0^1)} \to \infty \text{ as } \alpha \to 0$$

If we consider A_α in $V = L_2 \cap \alpha H_0^1$ with inner product

$$(u, v)_{L_2 \cap \alpha H_0^1} = (u, v)_{L_2} + \alpha^2 (\nabla u, \nabla v)$$

Then

$$\|A_\alpha\|_{L(V, V^*)} = 1 \;\text{ and }\; \|A_\alpha^{-1}\|_{L(V^*, V)} = 1$$

Hence, A_α is the Riesz mapping between these weighted spaces.

(Bergh and Löfström, Interpolation Spaces, 1976)

The Preconditioner

- The preconditioner should be an isomorphism

$$B : (X \times Y)^* \to X \times Y$$

- For example

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

If A and L are Riesz mappings in H_1 and H_2 then B is a Riesz mapping these weighted spaces

In practice we use multigrid preconditioners

The preconditioner

$$B^{-1} = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & \alpha\Delta + I & 0 \\ 0 & 0 & \frac{1}{\alpha}\Delta \end{bmatrix}$$

The components of the preconditioner are simple!

They consist of weighted sums of mass and stiffness matrices.

Standard preconditioners work!

The number of iterations is bounded

$h \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
2^{-1}	4	4	4	4	4
2^{-2}	5	8	11	12	8
2^{-3}	7	8	12	17	14
2^{-4}	7	8	12	18	20
2^{-5}	9	10	12	17	21
2^{-6}	9	10	13	17	18
2^{-7}	8	10	13	15	16
2^{-8}	8	10	11	13	13
2^{-9}	8	8	9	11	12

Table: Number of iterations required for preconditioned MinRes, where the preconditioned residual is reduced by a factor 1000

The condition number increases with α

$h \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
2^{-1}	1.28	1.45	4.15	17.6	31.0
2^{-2}	1.34	1.61	5.07	16.9	52.3
2^{-3}	1.36	1.67	5.38	16.3	53.2
2^{-4}	1.37	1.68	5.46	16.2	53.5
2^{-5}	1.37	1.69	5.48	16.3	53.5

Table: (Exact) Condition number $\kappa(BA)$

Example 2

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

subject to

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

The number of iterations is bounded

$l \setminus \alpha$	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: Number of iterations

The condition number grows

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

Table: (Exact) Condition number $\kappa(BA)$ of BA

Almost all eigenvalues are of unit size

Theoretical considerations

We have:

$$A_\alpha = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix}$$

and show that

$$\|A_\alpha\|_{L(V,V^*)} \leqslant C_1 \text{ and } \|A_\alpha^{-1}\|_{L(V^*,V)} \leqslant C_2/\alpha$$

Theoretical considerations, cont.

We use an auxiliary operator:

$$\hat{A}_\alpha = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' + \frac{1}{\alpha}K' \\ B & A + \frac{1}{\alpha}K & 0 \end{bmatrix}$$

and show that

$$\|\hat{A}_\alpha\|_{L(V,V^*)} \leqslant C_1/\alpha \quad \text{and} \quad \|\hat{A}_\alpha^{-1}\|_{L(V^*,V)} \leqslant C_2$$

Norwegian Centre of Excellence

Theoretical considerations, cont.

$$\hat{A}_\alpha - A_\alpha = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\alpha}K' \\ 0 & \frac{1}{\alpha}K & 0 \end{bmatrix}$$

By using and eigenvalue result of composed hermitian operator in terms of its components from H. Weyl. Mathematische Annalen, 1912 we show that only very few eigenvalues are close to zero

Theoretical considerations

- $\kappa(BA)$ is bounded independently of h
- $\kappa(BA)$ increases as $\alpha \to 0$:
  - Almost all eigenvalues are of order $O(1)$
  - Limited number of eigenvalues close to zero $(O(ln(\alpha)^2)$

Further reading:

Mardal and Winther, Numer. Linear Algebra Appl., 2011

Nielsen and Mardal, SIAM J. Control Optim., 2010

Mardal and Haga, Chapter 37 in Automated Solution of Differential Equations By the Finite Element Method, Springer, to be published soon (look at launchpad.net/fenics-book)

(papers can also be found at http://simula.no/people/kent-and/)