An operator theoretical approach to preconditioning optimality systems

Kent-André Mardal and Bjørn Fredrik Nielsen
Simula Research Laboratory
Oslo
Norway

EMG, Ischia, 2010

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
- Some examples (fruitfly and heart infarction)

(This framework is closely related to W. Zulehner's talk yesterday)

Abstract framework for preconditioning

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

$$\mathcal{A}u = f,$$

where $\mathcal{A}$ is a linear operator.

This problem is well-posed if $\mathcal{A}$ is an isomorphism mapping V to V^*, i.e.,

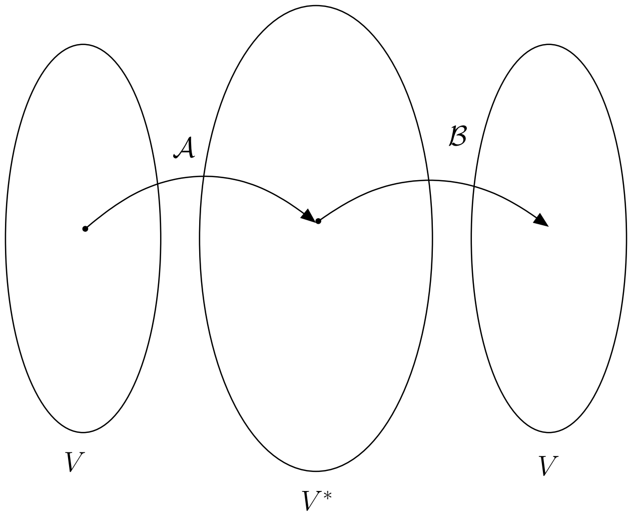
$$\|\mathcal{A}\|_{\mathcal{L}(V,V^*)} \leq C_1 \quad \text{and} \quad \|\mathcal{A}^{-1}\|_{\mathcal{L}(V^*,V)} \leq C_2$$

Note that $\mathcal{A}$ has an unbounded spectrum and this causes problems for iterative solvers (both in the continuous and discrete cases).

Abstract framework for preconditioning

From a mathematical point of view the Riesz mapping $\mathcal{B} : V^* \rightarrow V$ is the perfect preconditioner, since

$$\|\mathcal{B}\|_{\mathcal{L}(V^*,V)} = 1 \quad \text{and} \quad \|\mathcal{B}^{-1}\|_{\mathcal{L}(V,V^*)} = 1$$



Consequently,

$$\|\mathcal{B}\mathcal{A}\|_{\mathcal{L}(V,V)} \leq C_1 \text{ and } \|(\mathcal{B}\mathcal{A})^{-1}\|_{\mathcal{L}(V,V)} \leq C_2$$

Spectral Equivalence

It is well-known how to produce spectrally equivalent and efficient representations of Riesz mappings using multigrid and/or domain decomposition methods in a number of spaces like H^1, $H(\text{div})$, $H(\text{curl})$, and H^2.

In fact, you can view multigrid and domain decomposition methods as Riesz mappings in equivalent Sobolev spaces.

Example: An elliptic problem

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

$$\mathcal{A}u = -\nabla \cdot (K \nabla u) = f$$

Here, K positive definite and bounded.

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

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

Example: Stokes problem

Another example is Stokes problem: Find $u, p \in H_0^1 \times L_0^2$ such that for $f \in H^{-1}$

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

The Riesz mapping $\mathcal{B}$ taking $H^{-1} \times L_0^2 \to H_0^1 \times L_0^2$ is

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

The spectrum of $\mathcal{B}\mathcal{A}$ is bounded!

It is easy to construct spectrally equivalent and efficient versions of $\mathcal{B}$ with multigrid and domain decomposition techniques

The problem with inverse problems

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

$$\mathcal{A}u = f$$

The problem is not well-posed

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

$\mathcal{A}$ has a accumulation point at zero!

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

We will utilize clustering, but we will also construct V carefully (like Zulehner did yesterday).

Weighted Sobolev spaces

Consider the problem: Find $u \in H^1_0$, for $f \in H^{-1}$

$$\mathcal{A}_\alpha u = u - \alpha^2 \Delta u = f$$

Here, $\alpha > 0$

$$\|\mathcal{A}_\alpha^{-1}\|_{\mathcal{L}(H^{-1}, H^1_0)} \rightarrow \infty \text{ as } \alpha \rightarrow 0$$

If we consider $\mathcal{A}_\alpha$ in $V = L_2 \cap \alpha H^1_0$ with inner product

$$(u, v)_{L_2 \cap \alpha H^1_0} = (u, v)_{L_2} + \alpha^2 (\nabla u, \nabla v)$$

Then

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

(Bergh and Löfström, Interpolation Spaces, 1976)

Parameter identification problem

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

  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

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

[simula . research laboratory]

Optimality system, cont.

- $$\mathcal{A}_\alpha = \begin{bmatrix} \alpha L & 0 & B' \\ 0 & K & A' \\ B & A & 0 \end{bmatrix} : X \times Y \to (X \times Y)^*$$
- $X = H_1 \times H_2$
  - $\|x\|_X^2 = \alpha\|x_1\|_{H_1}^2 + \alpha\|x_2\|_{H_2}^2 + (T^*Tx_2, x_2)_{H_2}$
- $Y = H_2$
  - $\|y\|_Y^2 = \frac{1}{\alpha}\|y\|_{H_2}^2$

Preconditioning

- Preconditioner, isomorphism

$$\mathcal{B}_\alpha : (X \times Y)^* \to X \times Y$$

- For example

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

(in practice we use multigrid preconditioners)

Example 1

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

subject to

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

(This is the fruitfly example that has been studied by many. Our approach is close Schöberl and Zulehner, SIAM J. Matrix Anal., 2007)

Example 1, cont.

$h \setminus \alpha$	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

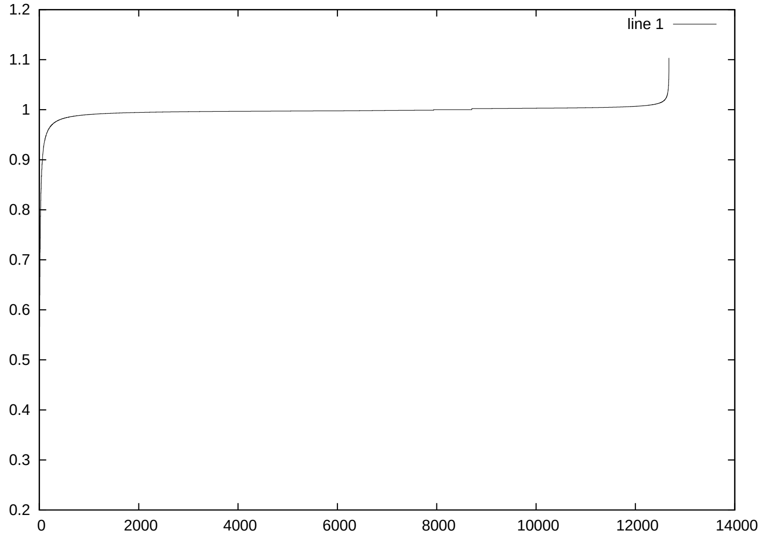
Table 1: Number of iterations

Example 1, cont.

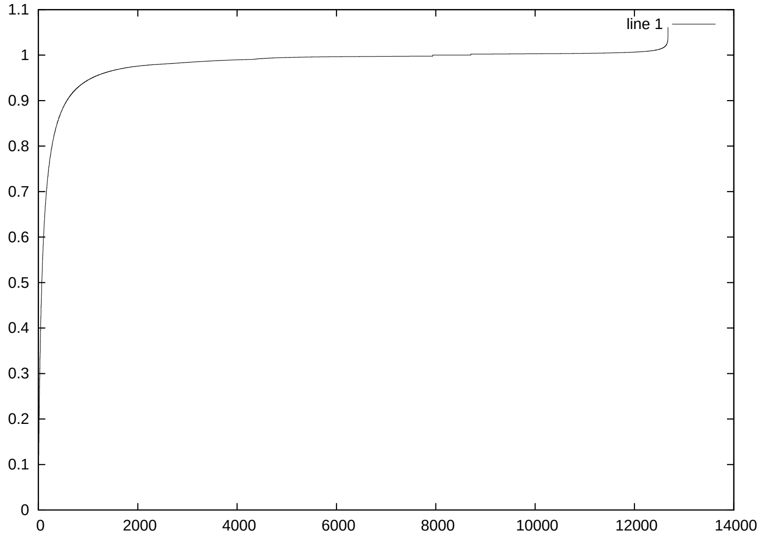
$h \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
2^{-1}	1.28	1.45	4.15	17.6	31.0
2^{-2}	1.34	1.61	5.07	16.9	52.3
2^{-3}	1.36	1.67	5.38	16.3	53.2
2^{-4}	1.37	1.68	5.46	16.2	53.5
2^{-5}	1.37	1.69	5.48	16.3	53.5

Table 2: Condition number $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$

Example 1, cont.



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



(b) $\alpha = 10^{-3}$

Figure 1: Absolute value of the eigenvalues of $\mathcal{B}_\alpha \mathcal{A}_\alpha$

Example 2

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

subject to

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

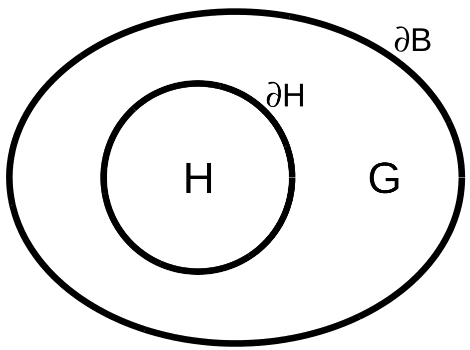


Figure 2: Body $P = \overline{H} \cup G$, heart H, torso G

Example 2, cont.

$l \setminus \alpha$	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

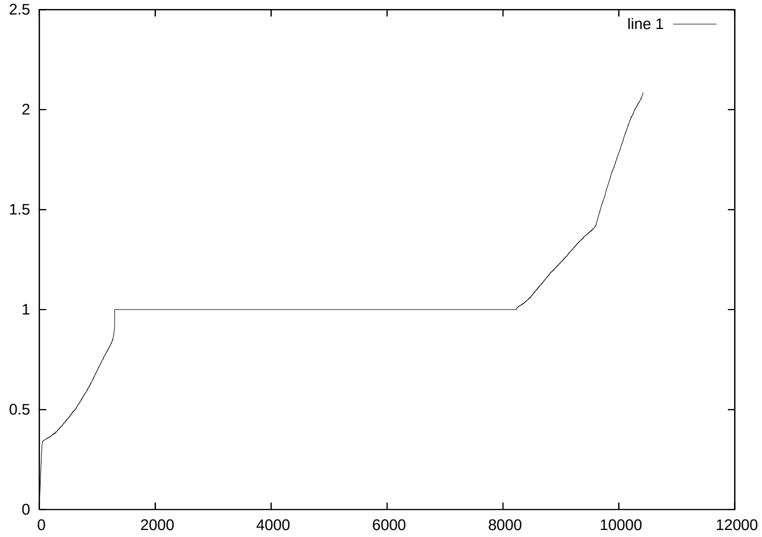
Table 3: Number of iterations

Example 2, cont.

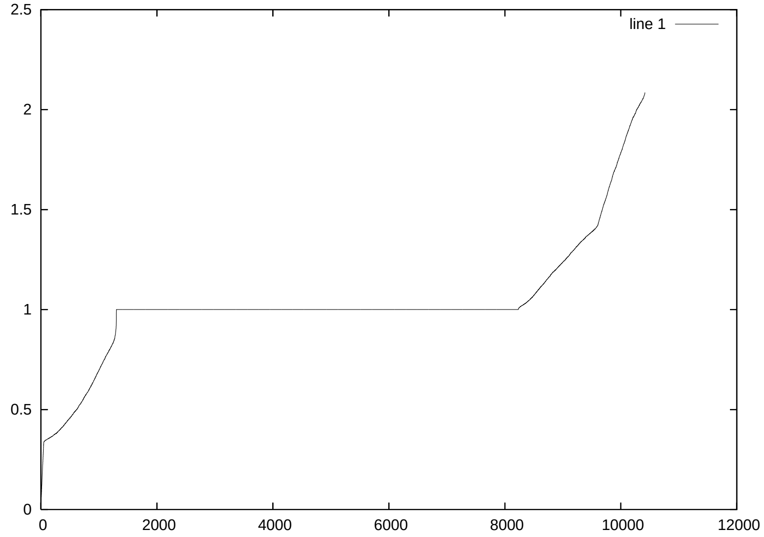
$l \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
1	16	108	672	5000	29729
2	16	109	680	5076	40157

Table 4: Condition number $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$ of $\mathcal{B}_\alpha \mathcal{A}_\alpha$

Example 2, cont.



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



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

Figure 3: Absolute value of the eigenvalues of $\mathcal{B}_\alpha \mathcal{A}_\alpha$

Theoretical considerations

We have:

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

and show that

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

Theoretical considerations, cont.

We use an auxiliary operator:

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

and show that

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

Theoretical considerations, cont.

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

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

- $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$ is bounded independently of h
- $\kappa(\mathcal{B}_\alpha \mathcal{A}_\alpha)$ 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., 2010

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

Mardal, Automated Scientific Computing, Springer, 2011

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

Further reading:

Mardal and Winther, Numer. Linear Algebra Appl., 2010

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

Mardal, Automated Scientific Computing, Springer, 2011

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

Questions?