Indefinite Preconditioners for PDE-constrained optimization problems

V. Simoncini

Dipartimento di Matematica, Università di Bologna, Italy

valeria.simoncini@unibo.it

Partly joint work with Debora Sesana, Università del Piemonte Orientale

The problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad \mathcal{A}x = b$$

Hypotheses:

- ⋆ $A \in \mathbb{R}^{n \times n}$ symmetric
- ⋆ $B^T \in \mathbb{R}^{n \times m}$ tall, $m \leq n$
- ⋆ C symmetric positive (semi)definite

More hypotheses later on specific problems...

Computational Algebraic Aspects:

Elman, Silvester, Wathen 2005 (book)

Benzi, Golub and Liesen, Acta Num 2005

Constraint (Indefinite) Preconditioner

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ \widetilde{B} & -\widetilde{C} \end{bmatrix} = \begin{bmatrix} I & 0 \\ \widetilde{B}\widetilde{A}^{-1} & I \end{bmatrix} \begin{bmatrix} \widetilde{A} & 0 \\ 0 & -S \end{bmatrix} \begin{bmatrix} I & \widetilde{A}^{-1}\widetilde{B} \\ 0 & I \end{bmatrix}$$

with $\widetilde{C} = S - B\widetilde{A}^{-1}B^T$ for some S.

Assume $\widetilde{B} = B$.

For particular choices of $\widetilde{A}, \widetilde{C}$, all eigs of $\mathcal{A}\mathcal{P}^{-1}$ are real and positive

(under certain conditions, variants of the CG method can be used)

Many contributions (Bai, Bergamaschi, Cao, Dollar, Durazzi, Ewing, Gondzio, Gould, Herzog, Keller, Lazarov, Lu, Lukšan, Ng, Perugia, Rozložník, Ruggiero, Sachs, Schilders, Schöberl, Vassilevski, Venturin, Vlček, Wang, Wathen, Zilli, Zulehner, ...)

The Magnetostatic problem

(3D) Maxwell equations: $\operatorname{div}\mathbf{B} = 0 \quad \operatorname{curl}\mathbf{H} = \mathbf{J}$

Constitutive law: $\mathbf{B} = \mu\mathbf{H}$

($\mathbf{B}$ displ. field; $\mathbf{H}$ magn. field; μ magn. perm.; $\mathbf{J}$ current dens.)

Constrained quadratic programming formulation:

$$\min \frac{1}{2}\int_{\Omega} \mu^{-1}\,|\mathbf{B} - \mu\mathbf{H}|^2\,dx$$

with $\begin{aligned} \mathbf{B}\cdot\mathbf{n} &= f_B \quad \text{on } \Gamma_B \\ \mathbf{H}\wedge\mathbf{n} &= \mathbf{f}_H \quad \text{on } \Gamma_H \end{aligned}$ and $\begin{aligned} \operatorname{div}\mathbf{B} &= 0 \\ \operatorname{curl}\mathbf{H} &= \mathbf{J} \end{aligned}$

Magnetostatic problem: Algebraic Saddle-Point problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

2D: A pos.def. on Ker(B)
B full row rank, $C = 0$

3D: A pos.def. on Ker(B), B rank deficient C semidefinite matrix
Range(C), Range(B) complementary spaces
$BB^T + C$ sym. positive definite

A zero-order operator, B first-order operator

Magnetostatic problem: Indefinite Preconditioning

$C = 0$. After scaling, Exact preconditioner:

$$\mathcal{P} = \begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ B & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -H \end{bmatrix} \begin{bmatrix} I & B^T \\ 0 & I \end{bmatrix} \qquad H = BB^T$$

Weyr canonical form $\qquad (\widehat{B}^T = B^T H^{-\frac{1}{2}})$

$$\mathcal{A}\mathcal{P}^{-1}\mathcal{X} = \mathcal{X} \begin{bmatrix} I_{n-m} + \Theta & & \\ & I_m & I_m \\ & & I_m \end{bmatrix}, \quad \mathcal{X} = \left[\begin{array}{c|c|c} \widehat{X} & \widehat{B}^T & (A-I)^{-1}\widehat{B}^T \\ 0 & 0_m & B(A-I)^{-1}\widehat{B}^T \end{array} \right]$$

where $(A - I)(I - \widehat{B}^T \widehat{B})\widehat{X} = \widehat{X}\Theta$ partial eigenvalue decomposition, associated with its nonzero eigenvalues

All real and positive eigenvalues: $\{1\} \cup \{1 + \theta_i\}$

Magnetostatic problem: Indefinite Preconditioning

Inexact Indefinite preconditioning:

$$\mathcal{P}_{\text{inex}} = \begin{bmatrix} I_n & 0 \\ B & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -H_{\text{inex}} \end{bmatrix} \begin{bmatrix} I_n & B^T \\ 0 & I_m \end{bmatrix}, \quad BB^T + C \approx H_{\text{inex}} \; spd$$

$$\mathcal{A}\mathcal{P}_{\text{inex}}^{-1} = \mathcal{A}\mathcal{P}^{-1} + \mathcal{E}, \qquad \mathcal{E} \text{ rank-}m$$

with

$$\|\mathcal{E}\| \leq \|\mathcal{A}\| \left\| \begin{bmatrix} B^T \\ -I_m \end{bmatrix} H^{-\frac{1}{2}} \right\| \max_{i=1,\ldots,m} |\lambda_i(HH_{\text{inex}}^{-1}) - 1|$$

Inexact Indefinite Preconditioning. On the choice of $H_{\rm inex}$

- If $H_{\rm inex} > 0$ is such that $H - H_{\rm inex}$ has $k \le m$ zero eigenvalues, then $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ retains $2k$ unit eigenvalues with geometric multiplicity k.

- First order perturbation of (multiple) unit eigenvalue:

$$\lambda(\mathcal{A}\mathcal{P}_{\rm inex}^{-1}) \approx \lambda(\mathcal{A}\mathcal{P}^{-1}) + \xi^{\frac{1}{2}}$$

Assume $A - I < 0$. Then ξ real. If $H - H_{\rm inex} \le 0$ then $\xi \le 0$

$$\Downarrow$$

1) Spectral approximation matters

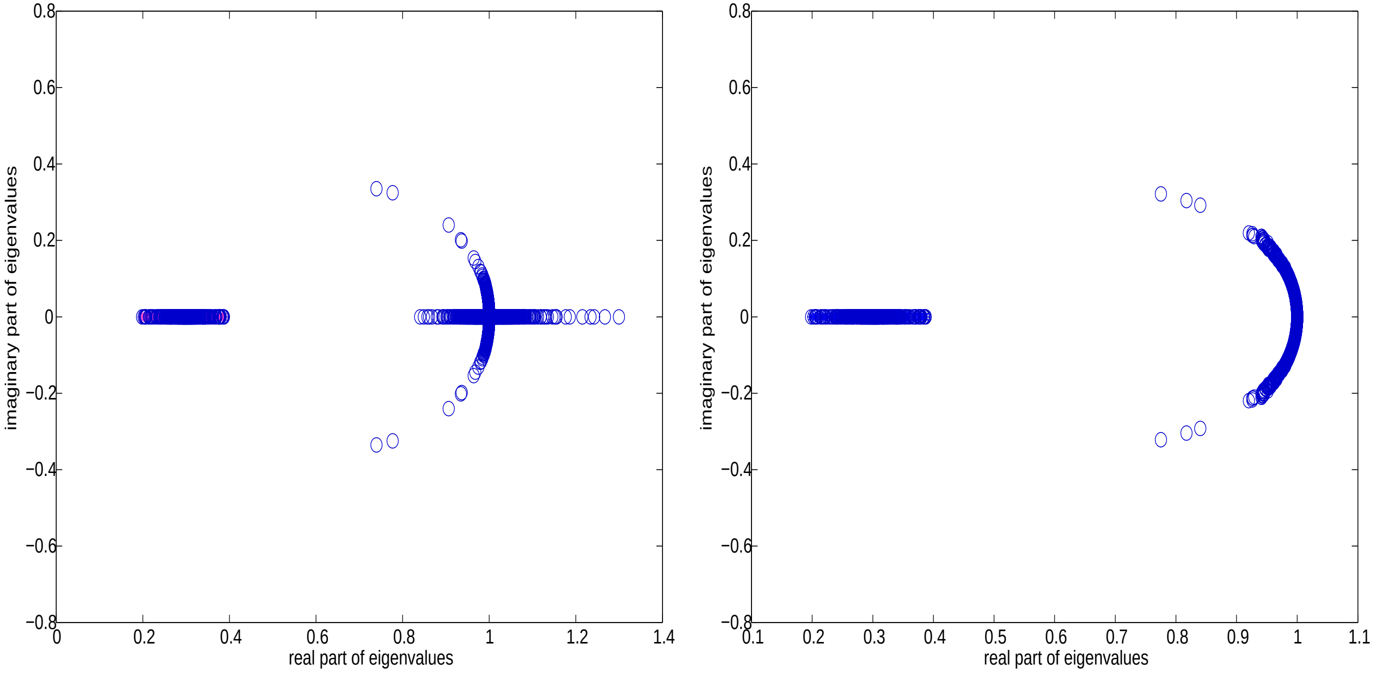
2) Sign of approximation matters

$\Rightarrow$ ξ independent of meshsize

Inexact Indefinite Preconditioning. On the choice of H_{inex}

Spectrum of $\mathcal{A}\mathcal{P}_{\text{inex}}^{-1}$

Incomplete Choleski (`tol`=1e-3) AMG preconditioning



The Stokes problem

Minimize

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx - \int_\Omega f \cdot u dx$$

subject to $\nabla \cdot u = 0$ in Ω

Lagrangian: $\mathcal{L}(u,p) = J(u) + \int_\Omega p\nabla \cdot u dx$

Optimality condition on discretized Lagrangian leads to:

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}$$

A second-order operator, B first-order operator, C zero-order operator

Thanks to Walter Zulehner

The Stokes problem. Inexact contraint preconditioning

$$\mathcal{P}_{\text{inex}} = \begin{bmatrix} I_n & 0 \\ B\widetilde{A}^{-1} & I_m \end{bmatrix} \begin{bmatrix} \widetilde{A} & 0 \\ 0 & -H_{\text{inex}} \end{bmatrix} \begin{bmatrix} I_n & \widetilde{A}^{-1}B^T \\ 0 & I_m \end{bmatrix}$$

with $H = B\widetilde{A}^{-1}B^T + C \approx H_{\text{inex}}\ spd$

First order spectral perturbation of simple eigenvalues:

$$|\lambda(\mathcal{A}\mathcal{P}_{\text{inex}}) - \lambda(\mathcal{A}\mathcal{P}^{-1})| \approx c\kappa(\widetilde{A}^{-1}A - I)^{\frac{1}{2}} \max_{j=1,\dots,m} |\lambda_j(HH_{\text{inex}}^{-1}) - 1|$$

(for $\widetilde{A}^{-1}A - I$ definite)

$\Rightarrow$ Spectrum independent of mesh parameter

(for judicious choices of $\widetilde{A}, H_{\text{inex}}$)

The Stokes problem. Inexact contraint preconditioning

Selection of $\widetilde{A}$, H_{inex}:

$$\widetilde{A} = \text{AMG}(A), \qquad H_{\text{inex}} = Q \text{ (pressure mass matrix)}$$

IFISS 3.1 (Elman, Ramage, Silvester):

Flow over a backward facing step

Stable Q2-Q1 approximation

$(C = 0)$

stopping tolerance: 10^{-6}

n	m	# it.
1538	209	18
5890	769	18
23042	2945	18
91138	11521	17
362498	45569	17

Constrained Optimal Control Problem. A "toy" problem.

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$. Given $\hat{u}$ (*desired state*) in $\hat{\Omega} \subseteq \Omega$, find u:

$$\min_{u,f} \frac{1}{2}\|u - \hat{u}\|^2_{L_2(\hat{\Omega})} + \beta\|f\|^2_{L_2(\Omega)}$$

$$\text{s.t.} \qquad -\nabla^2 u = f \quad in \quad \Omega$$

with $u = \hat{u}$ on $\partial\Omega$. Lagrangian of discretized problem:

$$\mathcal{L}(\mathbf{f}, \mathbf{u}, \lambda) = \frac{1}{2}\mathbf{u}^T \bar{M}\mathbf{u} - \mathbf{u}^T M\hat{\mathbf{u}} + \frac{1}{2}\|\hat{\mathbf{u}}\|^2 + \beta \mathbf{f}^T M\mathbf{f} + \lambda^T(K\mathbf{u} - M\mathbf{f} - \mathbf{d})$$

K stiffness matrix. First order optimality condition yields:

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

$\bar{M}$ could be singular (depending on where $\hat{u}$ is defined)

Dimension reduction

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

that is, $2\beta\mathbf{f} = \lambda$. Therefore

$$\begin{bmatrix} \bar{M} & K^T \\ K & -\frac{1}{2\beta}M \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

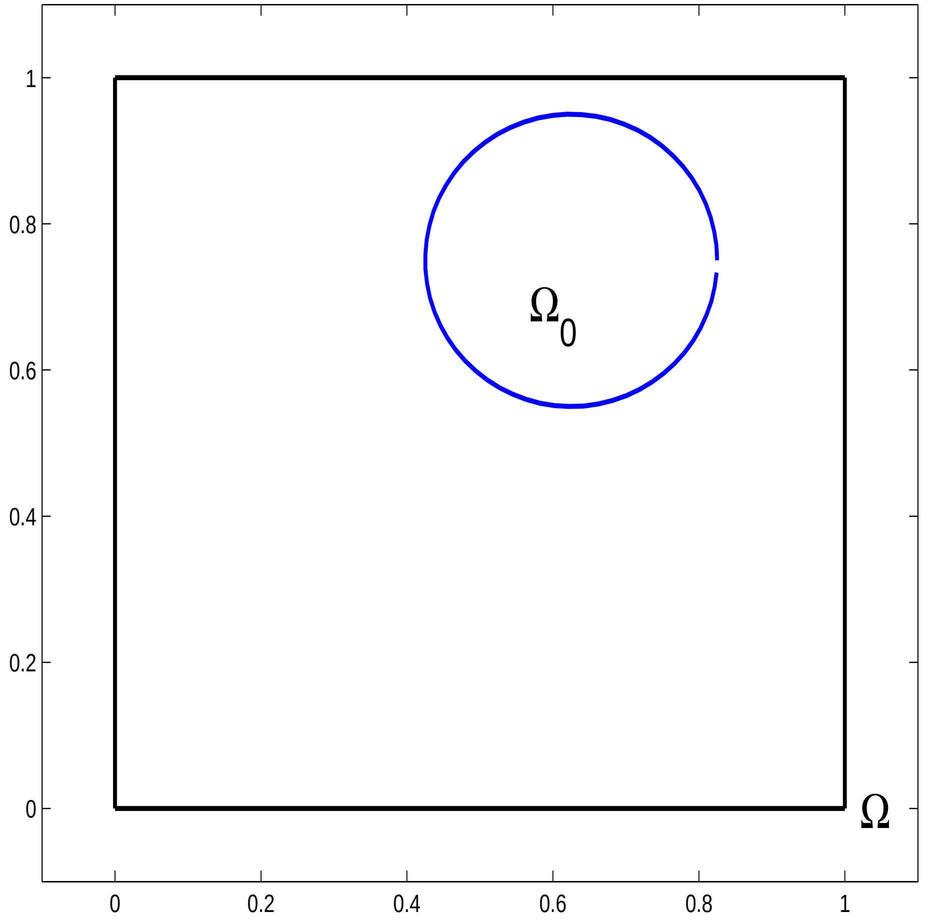
with $\bar{M} = \bar{M}^T \geq 0$, $K = K^T$ square, $M = M^T > 0$

Indefinite Preconditioning strategy

$$\mathcal{P} = \begin{bmatrix} 0 & \widetilde{K} \\ \widetilde{K} & -\frac{1}{2\beta}M \end{bmatrix}, \qquad \mathcal{P}^{-1} = \begin{bmatrix} \widetilde{K}^{-1}C\widetilde{K}^{-1} & \widetilde{K}^{-1} \\ \widetilde{K}^{-1} & 0 \end{bmatrix}, \quad \widetilde{K} \approx K$$

- If $\widetilde{K} = K$, then $\lambda_i(\mathcal{A}\mathcal{P}^{-1}) = 1 + \eta, \;\; 0 \le \eta \le \frac{c}{\beta}$ (independent of meshsize)
- If $\widetilde{K}$ spectrally equivalent to K, still independence of meshsize

Numerical results: 2D and 3D



2D: $\hat{u}(x,y) = 2$ in Ω_0 and $\hat{u}(x,y) = 0$ on $\partial\Omega$ (undefined elsewhere)

Data thanks to Sue H. Thorne, RAL, UK

Numerical results

$\bar{M}$ singular, $\widetilde{K}$ =AMG(K)

2D:

	$\beta = 10^{-5}$	$\beta = 10^{-2}$
n	# it.	# it.
961	10	3
3969	10	3
16129	10	3
65025	10	3
261121	10	4

3D:

	$\beta = 10^{-5}$	$\beta = 10^{-2}$
n	# it.	# it.
343	8	3
3375	9	3
29791	9	3
250047	9	3

Final considerations

- "Plain" use of Indefinite (constraint) preconditioning should **not** be discouraged
- Interplay between Solvers and Preconditioners is crucial
- Preconditioning strategies for Saddle Point systems largely expanding topic

  (also: block diagonal/triangular, augmented, projected CG, etc...)

References for this talk:

V.Simoncini, *Reduced order solution of structured linear systems arising in certain PDE-constrained optimization problems*, to appear in COAP.

D. Sesana and V. Simoncini, *Spectral analysis of inexact constraint preconditioning for symmetric saddle point matrices*, Submitted, Jan.2012.