Solution of structured algebraic linear systems in PDE-constrained optimization problems

V. Simoncini

Dipartimento di Matematica, Università di Bologna
and CIRSA, Ravenna, Italy
valeria@dm.unibo.it

The problem

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

- Computational Fluid Dynamics (Elman, Silvester, Wathen 2005)
- Elasticity problems
- Mixed (FE) formulations of II and IV order elliptic PDEs
- Linearly Constrained Programs
- Linear Regression in Statistics
- Image restoration and registration
- ... Survey: Benzi, Golub and Liesen, Acta Num 2005

The problem

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

Hypotheses:

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

More hypotheses later on specific problems...

Application problems for PDE-constrained optimization

Three representative driving settings:

- Constrained magnetostatic problem
- Simplified Optimal Control Problem
- Optimal transport problem

Computational Algebraic Aspects: Battermann, Benzi, Biros, Dollar, Ghattas, Gould, Haber, Heinkenschloss, Herzog, Sachs, Schöberl, Stoll, Wathen, 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.)

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
$$\begin{aligned} \operatorname{div} \mathbf{B} &= 0 \\ \operatorname{curl} \mathbf{H} &= \mathbf{J} \end{aligned}$$

boundary conditions are enforced

constraints: Lagrange multipliers

3D magnetic field-based mixed formulation

$a(\mathbf{u},\mathbf{v}) = (\mathbf{u},\mathbf{v})_\mu \quad \forall \mathbf{u},\mathbf{v} \in V$

$b(\mathbf{u},\mathbf{p}) = (\mathrm{curl}\,\mathbf{u},\mathbf{p})_\mu \quad \forall \mathbf{u} \in V,\ \forall \mathbf{p} \in Q$ (V and Q properly chosen)

$$\begin{cases} a(\mathbf{u},\mathbf{u}^*) + b(\mathbf{u}^*,\mathbf{p}) = 0 & \forall \mathbf{u}^* \in V \\ b(\mathbf{u},\mathbf{p}^*) - c(\mathbf{p},\mathbf{q}) = (\mathbf{g},\mathbf{p}^*) & \forall \mathbf{p}^* \in Q \end{cases}$$

Stabilization term: $c(\mathbf{p},\mathbf{q}) = (\mathrm{div}\,\mathbf{p},\mathrm{div}\,\mathbf{q})_\mu \quad \forall \mathbf{p},\mathbf{q} \in Q$

(cf. Arioli, Simoncini, Perugia, '99)

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: B rank deficient $\qquad$ C semidefinite matrix

Range(C), Range(B) complementary spaces

$BB^T + C$ sym. positive definite

Constrained Optimal Control Problem

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3.$

Given $\hat{u}$ ($desired\ state$) defined in $\hat{\Omega} \subseteq \Omega$, find u:

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

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

$$u = \hat{u} \quad on \quad \partial\Omega$$

- β: regularization parameter
- Typical extra condition: box constraints for f
- Alternative PDE constraints
- Alternative PDE boundary conditions

Constrained Optimal Control Problem. Discretization.

Discrete cost functional

$$\min_{u_h, f_h} \frac{1}{2}\|u_h - \hat{u}\|_2^2 + \beta\|f_h\|_2^2 = \min_{\mathbf{u},\mathbf{f}} \frac{1}{2}\mathbf{u}^T \bar{M}\mathbf{u} - \mathbf{u}^T\mathbf{u} + \alpha + \beta\mathbf{f}^T M\mathbf{f}$$

with $\alpha = \|\hat{u}\|_2^2$, M mass matrix, $\bar{M}$ portion of mass matrix

Constraint:

$$-\nabla^2 u = f \quad in \quad \Omega \Rightarrow K\mathbf{u} = M\mathbf{f} + \mathbf{d}$$

K stiffness matrix (discrete Laplacian)

Solution via Lagrangian:

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

The resulting saddle point problem

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

First order condition on Lagrangian 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 (depends on where $\hat{u}$ is defined)

Parameter identification problem

Model for groundflow: $\mathcal{A}(q)u \equiv -\nabla \cdot (\sigma \nabla u) = f(x)$, with $x \in \Omega$

u: fluid pressure $\sigma(x)$: (spatially dep.) hydraulic conductivity

$f(x)$: in-going or outgoing fluid (incompressible flows)

Parameter of interest: $q(x) = \log(\sigma(x))$ for given $u_{obs} = Cu$

Set $F(q) = C\mathcal{A}(q)^{-1} f$ (non-linear function) then

$$\min_q \frac{1}{2} ||F(q) - z||^2 + \alpha \mathcal{J}_{reg}(q)$$

where $\mathcal{J}_{reg}$ regularization functional (e.g. total variation)

PDE-constrained formulation

$$\min_{q,u} \frac{1}{2}||Cu - z||^2 + \alpha \mathcal{J}_{reg}(q)$$
$$\mathcal{A}(q)u - f = 0.$$

Space discretization + inexact Newton method provide linear systems:

$$\begin{bmatrix} H & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} v \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} -L_v \\ -L_{\boldsymbol{\lambda}} \end{bmatrix}$$

H: Hessian of the operator

B: Jacobian of the constraint

$v = [q, u]$ variables, $\boldsymbol{\lambda}$: Lagrange multiplier

A similar setting: Monge-Kantorovich mass transfer problem

Pb: Given two density functions u_0 and u_T on the set Ω, find an "optimal" mapping from u_0 to u_T

Formulation (time in $[0, T]$):

$$\min_{u,m} \frac{1}{2}\|u(T,\mathbf{x}) - u_T(\mathbf{x})\|^2 + \frac{1}{2}\alpha T \int_\Omega \int_0^T u\|m\|^2 dt d\mathbf{x}$$

$$s.t. \qquad u_t + \nabla \cdot (um) = 0, \qquad u(0,\mathbf{x}) = u_0$$

$u(t,\mathbf{x})$: density field; $\qquad m(t,\mathbf{x})$: velocity field

Time and space discretiz. $+$ Gauss-Newton approx on the Lagrangian

$$\mathcal{L}(u,m,p) = \frac{1}{2}\|Qu - u_T\|^2 + \frac{1}{2}\alpha T h_t h_x^2 u^T L \mathrm{diag}(m) m + p^T V(A(m)u - q)$$

A preconditioning technique for a class of PDE-constrained optimization problems, M. Benzi, E. Haber and L. Taralli, Adv. in Comput. Math. '10

A similar setting: Monge-Kantorovich mass transfer problem

$$\mathcal{L}(u,m,p) = \frac{1}{2}\|Qu - u_T\|^2 + \frac{1}{2}\alpha T h_t h_x^2 u^T L \mathrm{diag}(m) m + p^T V (A(m)u - q)$$

Time and space discretization + Gauss-Newton approximation provide linear systems:

$$\begin{bmatrix} Q^TQ & 0 & B_1^T \\ 0 & D & B_2^T \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{u}_k \\ \widetilde{m}_k \\ \widetilde{p}_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

with $Q^TQ > 0$ diagonal, D diagonal and highly singular

Solution. CG for $\mathcal{M}x = b$

CG: minimum error method (in energy norm). For $\mathcal{M}$ spd ($x_0 = 0$)

$$\min_{x \in K_k(\mathcal{M},b)} \|x_\star - x\|_{\mathcal{M}} \quad \leq \quad 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|x_\star\|_{\mathcal{M}}$$

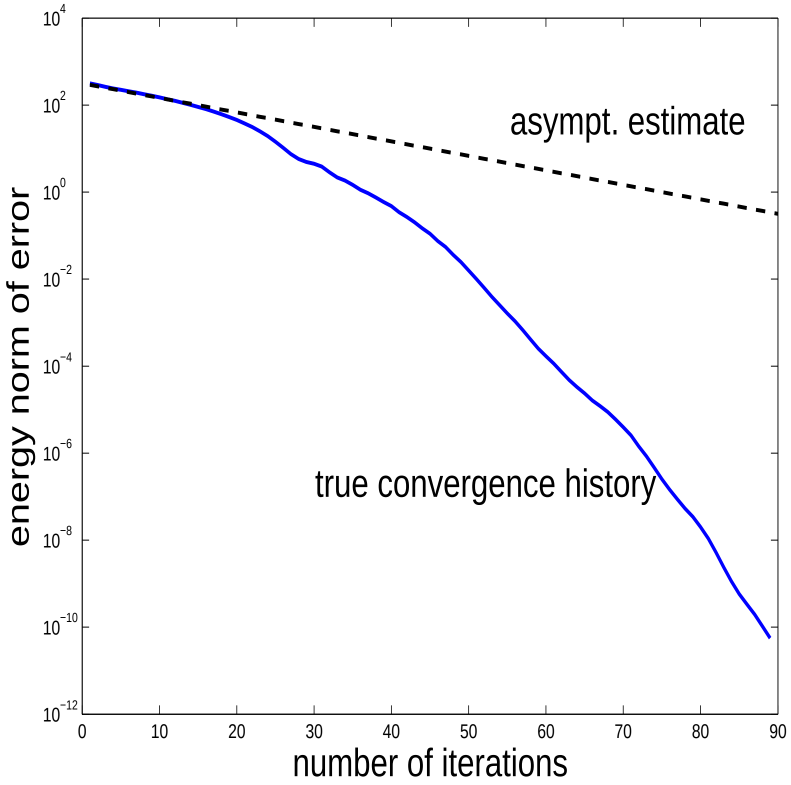
with $\kappa = \lambda_{\max}(\mathcal{M})/\lambda_{\min}(\mathcal{M})$

Solution. CG for $\mathcal{M}x = b$

CG: minimum error method (in energy norm). For $\mathcal{M}$ spd ($x_0 = 0$)

$$\min_{x \in K_k(\mathcal{M},b)} \|x_\star - x\|_{\mathcal{M}} \quad \leq \quad 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|x_\star\|_{\mathcal{M}}$$

with $\kappa = \lambda_{\max}(\mathcal{M})/\lambda_{\min}(\mathcal{M})$ Superlinear Convergence



Symmetry wrto Euclidean inner product

$$\mathcal{M}x = b, \qquad \mathcal{M} \text{ spd}$$

Classical CG: $(u, v) = u^T v$

Given x_0

$r_0 = b - \mathcal{M}x_0$, $p_0 = r_0$

for $i = 0, 1, \ldots$

$\alpha_i = \frac{(r_i, r_i)}{(p_i, \mathcal{M}p_i)}$

$x_{i+1} = x_i + p_i \alpha_i$

$r_{i+1} = r_i - \mathcal{M}p_i \alpha_i$

$\beta_{i+1} = \frac{(r_{i+1}, \mathcal{M}p_i)}{(p_i, \mathcal{M}p_i)}$

$p_{i+1} = r_i + p_i \beta_{i+1}$

Symmetry wrto H-inner product (H spd)

$$\mathcal{M}x = b$$

Assume there exists H spd such that $H\mathcal{M}$ is also spd

H-sym CG: $(u, v)_H = u^T H v$

Given x_0

$r_0 = b - \mathcal{M}x_0$, $p_0 = r_0$

for $i = 0, 1, \ldots$

$\alpha_i = \frac{(r_i, r_i)_H}{(p_i, \mathcal{M}p_i)_H}$

$x_{i+1} = x_i + p_i \alpha_i$

$r_{i+1} = r_i - \mathcal{M}p_i \alpha_i$

$\beta_{i+1} = \frac{(r_{i+1}, \mathcal{M}p_i)_H}{(p_i, \mathcal{M}p_i)_H}$

$p_{i+1} = r_i + p_i \beta_{i+1}$

Iterative solver. Convergence considerations.

$$\mathcal{M}x = b$$

$\mathcal{M}$ is symmetric and indefinite $\quad\longrightarrow\quad$ MINRES

Iterative solver. Convergence considerations.

$$\mathcal{M}x = b$$

$\mathcal{M}$ is symmetric and indefinite $\quad \longrightarrow \quad$ MINRES

$$x_k \in x_0 + K_k(\mathcal{M}, r_0), \quad \text{s.t.} \quad \min \|b - \mathcal{M}x_k\|$$

$r_k = b - \mathcal{M}x_k$, $k = 0, 1, \ldots$, x_0 starting guess

If $\mu(\mathcal{M}) \subset [-a, -b] \cup [c, d]$, with $|b - a| = |d - c|$, then

$$\|b - \mathcal{M}x_{2k}\| \leq 2 \left(\frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \right)^k \|b - \mathcal{M}x_0\|$$

Note: more general but less tractable bounds available

General preconditioning strategy

- Find $\mathcal{P}$ such that

$$\mathcal{M}\mathcal{P}^{-1}\hat{u} = b \qquad \hat{u} = \mathcal{P}u$$

is easier (faster) to solve than $\mathcal{M}u = b$

- A look at efficiency:

- Dealing with $\mathcal{P}$ should be cheap
- Storage requirements for $\mathcal{P}$ should be low

Possibly zero storage

- Properties (algebraic/functional) should be exploited

Mesh/parameter independence

Structure preserving preconditioners

Spectral properties

A symmetric positive definite

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \qquad \begin{array}{ll} 0 < \lambda_n \le \cdots \le \lambda_1 & \text{eigs of } A \\ 0 < \sigma_m \le \cdots \le \sigma_1 & \text{sing. vals of } B \end{array}$$

$\sigma(\mathcal{M})$ subset of (Rusten & Winther 1992)

$$\left[\frac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2})\right] \quad \cup \quad \left[\lambda_n, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

Spectral properties

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \qquad \begin{array}{ll} 0 < \lambda_n \leq \cdots \leq \lambda_1 & \text{eigs of } A \\ 0 = \sigma_m \leq \cdots \leq \sigma_1 & \text{sing. vals of } B \end{array}$$

$\sigma(\mathcal{M})$ subset of

$$\left[\frac{1}{2}(-\gamma_1 + \lambda_n - \sqrt{(\gamma_1 + \lambda_n)^2 + 4\sigma_1^2})\ , \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\theta})\right] \cup \left[\lambda_n, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

B rank deficient, but $\quad \theta = \lambda_{\min}(BB^T + C)$ full rank

$$\gamma_1 = \lambda_{\max}(C)$$

Block diagonal Preconditioner

⋆ A nonsing., $C = 0$:

$$\mathcal{P}_0 = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix}$$

$$\Rightarrow \qquad \mathcal{P}_0^{-\frac{1}{2}} \mathcal{M} \mathcal{P}_0^{-\frac{1}{2}} = \begin{bmatrix} I & A^{-\frac{1}{2}} B^T (BA^{-1}B^T)^{-\frac{1}{2}} \\ (BA^{-1}B^T)^{-\frac{1}{2}} BA^{-\frac{1}{2}} & 0 \end{bmatrix}$$

MINRES converges in at most 3 iterations. $\sigma(\mathcal{P}_0^{-\frac{1}{2}} \mathcal{M} \mathcal{P}_0^{-\frac{1}{2}}) = \{1, 1/2 \pm \sqrt{5}/2\}$

A more practical choice:

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{S} \end{bmatrix}$$

spd. $\widetilde{A} \approx A \qquad \widetilde{S} \approx BA^{-1}B^T$

eigs in $\qquad [-a, -b] \cup [c, d], \qquad a, b, c, d > 0$

Still an Indefinite Problem

Constraint (Indefinite) Preconditioner

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & B^T \\ B & -C \end{bmatrix} \qquad \mathcal{M}\mathcal{P}^{-1} = \begin{bmatrix} A\widetilde{A}^{-1}(I - \Pi) + \Pi & \star \\ O & I \end{bmatrix}$$

with $\Pi = B(B\widetilde{A}^{-1}B^T + C)^{-1}B\widetilde{A}^{-1}$

- If C nonsing $\Rightarrow$ all eigs real and positive
- If $B^T C = 0$ and $BB^T + C > 0 \Rightarrow$ all eigs real and positive

Special case: $C = 0$ $\quad \Rightarrow$ at most $2m$ unit eigs with Jordan blocks

Constraint (Indefinite) Preconditioner. Generalizations

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

Primal-based: $\widetilde{C} \approx C$ nonsing, $\quad \widetilde{A} \approx A + B^T \widetilde{C}^{-1} B$

- If $A + B^T \widetilde{C}^{-1} B > \widetilde{A}$ and $\widetilde{C} > C \Rightarrow$ all eigs real and positive

Constraint (Indefinite) Preconditioner. Generalizations

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

Primal-based: $\widetilde{C} \approx C$ nonsing, $\quad \widetilde{A} \approx A + B^T \widetilde{C}^{-1} B$

- If $A + B^T \widetilde{C}^{-1} B > \widetilde{A}$ and $\widetilde{C} > C$ $\Rightarrow$ all eigs real and positive

Dual-based: $(C = O) \quad \widetilde{A} \approx A, \quad \widetilde{C} = S - B\widetilde{A}^{-1}B^T$ for some S

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

Constraint (Indefinite) Preconditioner. Generalizations

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

Primal-based: $\widetilde{C} \approx C$ nonsing, $\quad \widetilde{A} \approx A + B^T \widetilde{C}^{-1} B$

- If $A + B^T \widetilde{C}^{-1} B > \widetilde{A}$ and $\widetilde{C} > C \Rightarrow$ all eigs real and positive

Dual-based: $(C = O) \quad \widetilde{A} \approx A, \quad \widetilde{C} = S - B\widetilde{A}^{-1}B^T$ for some S

- If $\widetilde{A} > A$ and $\widetilde{C} < 0 \Rightarrow$ all eigs real and positive

  $\mathcal{M}\mathcal{P}^{-1}$ is H-symmetric with $H = \mathrm{blkdiag}(\widetilde{A} - A, B\widetilde{A}^{-1}B^T - S)$

Symmetric and indefinite A

$$\mathcal{M} = \begin{bmatrix} A & B^\top \\ B & O \end{bmatrix}$$

$\lambda_n \leq \cdots \leq \lambda_1$ eigs of A

$0 < \sigma_m \leq \cdots \leq \sigma_1$ sing. vals of B

A pos.def. on $\mathrm{Ker}(B)$

$\mu(\mathcal{M})$ subset of

$$\left[\frac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2})\right] \quad \cup \quad \left[\Gamma, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

If $m = n$, $\Gamma = \frac{1}{2}(\lambda_n + \sqrt{\lambda_n^2 + 4\sigma_m^2})$

Gould & Simoncini, '09

Indefinite A, $C = 0$. Cont'd. $m < n$

$$\left[\frac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2})\right] \quad \cup \quad \left[\Gamma, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

Letting $\alpha_0 > 0$ be s.t. $\frac{u^\top A u}{u^\top u} > \alpha_0$, $u \in \text{Ker}(B)$

$$\Gamma \geq \begin{cases} \dfrac{\alpha_0 \sigma_m^2}{|\alpha_0 \lambda_n - \|A\|^2 - \sigma_m^2|} & \text{if } \alpha_0 + \lambda_n \leq 0 \\ \dfrac{\alpha_0 \lambda_n - \|A\|^2 - \sigma_m^2}{2(\alpha_0 + \lambda_n)} + \sqrt{\left(\dfrac{\alpha_0 \lambda_n - \|A\|^2 - \sigma_m^2}{2(\alpha_0 + \lambda_n)}\right)^2 + \dfrac{\alpha_0 \sigma_m^2}{\alpha_0 + \lambda_n}} & \text{otherwise.} \end{cases}$$

Augmenting the (1,1) block

A sym and indefinite. Equivalent formulation ($C = 0$):

$$\begin{bmatrix} A + \tau B^\top B & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f + \tau B^\top g \\ g \end{bmatrix}, \quad \tau \in \mathbb{R}$$

coefficient matrix: $\mathcal{M}(\tau)$

Augmenting the (1,1) block

A sym and indefinite. Equivalent formulation ($C=0$):

$$\begin{bmatrix} A+\tau B^\top B & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f+\tau B^\top g \\ g \end{bmatrix}, \qquad \tau \in \mathbb{R}$$

coefficient matrix: $\mathcal{M}(\tau)$

Condition on τ for definiteness of $A+\tau B^\top B$:

$$\tau > \frac{1}{\sigma_m^2}\left(\frac{\|A\|^2}{\alpha_0} - \lambda_n\right)$$

Ex.2. $A = \begin{bmatrix} 0.01 & 3 \\ 3 & -0.01 \end{bmatrix}$, $\mu(\mathcal{M}) = \{-4.2452, 5.0\cdot 10^{-3}, 4.2402\}$

$\frac{1}{\sigma_m^2}\left(\frac{\|A\|^2}{\alpha_0} - \lambda_n\right) = 100.33$

for $\tau = 100 \rightarrow A+\tau B^\top B$ is indefinite

Application to **practical** block diagonal preconditioners

Indefinite A (and $C = 0$). Indefinite preconditioner:

Let $\mathcal{P}_{\pm} = \mathrm{blkdiag}(A, \pm\widetilde{S})$ with A, $\widetilde{S}$ nonsingular. Then

$$\mu(\mathcal{P}_{\pm}^{-1}\mathcal{M}) \subset \left\{1, \frac{1}{2}(1+\sqrt{1+4\xi}), \frac{1}{2}(1-\sqrt{1+4\xi})\right\} \subset \mathbb{C},$$

ξ : (possibly complex) eigenvalues of $(BA^{-1}B^{\top}, \pm\widetilde{S})$

Application to ideal block diagonal preconditioners

Indefinite A, $C \neq 0$. Indefinite preconditioner:

Let $\mathcal{P}_+ = \mathrm{blkdiag}(A, C + BA^{-1}B^\top)$. Then

$$\mu(\mathcal{P}_+^{-1}\mathcal{M}) \subset \left\{1, \frac{1}{2}(1 \pm \sqrt{5}), \frac{1}{2\theta}(\theta - 1 \pm \sqrt{(1-\theta)^2 + 4\theta^2})\right\} \subset \mathbb{R}.$$

θ finite eigs of $(C + BA^{-1}B^\top, C)$

Similar (but complex) results for $\mathcal{P}_- = \mathrm{blkdiag}(A, -C - BA^{-1}B^\top)$

Application to ideal block diagonal preconditioners

Indefinite A, Definite preconditioner, $C = 0$:

$$\mathcal{P}(\tau) = \begin{bmatrix} P_A & \\ & P_C \end{bmatrix}, \qquad \begin{array}{l} P_A \approx P_A(\tau) = A + \tau B^\top B \\ P_C \approx P_C(\tau) = B(A + \tau B^\top B)^{-1} B^\top \end{array}$$

- Definite preconditioner on definite problem:

$\mathcal{P}(\tau)^{-1}\mathcal{M}(\tau)$ has eigenvalues

$$1,\ \tfrac{1}{2}(1+\sqrt{5}),\ \tfrac{1}{2}(1-\sqrt{5})$$

with multiplicity $n - m$, m and m, respectively.

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: B rank deficient $\qquad$ C semidefinite matrix
Range(C), Range(B) complementary spaces
$BB^T + C$ sym. positive definite

Magnetostic problem: solution strategies

$$BB^T + C \approx C_0$$

Geometric/Algebraic Multigrid, Incomplete Fact., Inner-Outer, etc.

$$P_d = \begin{bmatrix} I & 0 \\ 0 & C_0 \end{bmatrix}$$

$$P_{in} = \begin{bmatrix} I & 0 \\ B & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -C_0 \end{bmatrix} \begin{bmatrix} I & B^T \\ 0 & I \end{bmatrix}$$

$A \approx I$ (or its diagonal)

Magnetostic problem: Numerical example

Computational cost 3D Magnetostatic problem

C_0: Incomplete Cholesky fact. (ICT package, Saad & Chow)

Elapsed Time

Size	MA27	QMR		QMR
		$P_d(it)$	$P_{in}(it)$	ILDLT(10)
2208	**2.3**	**4.9**(53)	**3.1**(18)	**1.5**
4371	**10.2**	**11.0**(67)	**8.4**(20)	**5.2**
8622	**83.4**	**24.3**(86)	**18.3**(29)	**31.0**
22675	**753.5**	**128.2**(179)	**63.2**(45)	**246.0**

(Old results)

Constrained Optimal Control Problem
The resulting saddle point problem

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

First order condition on Lagrangian 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 (this depends 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

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

Solution strategies

Ideal cases:

$$P_d = \begin{bmatrix} KD^{-1}K^T & 0 \\ 0 & D \end{bmatrix}, \qquad D = \frac{1}{2\beta}M$$

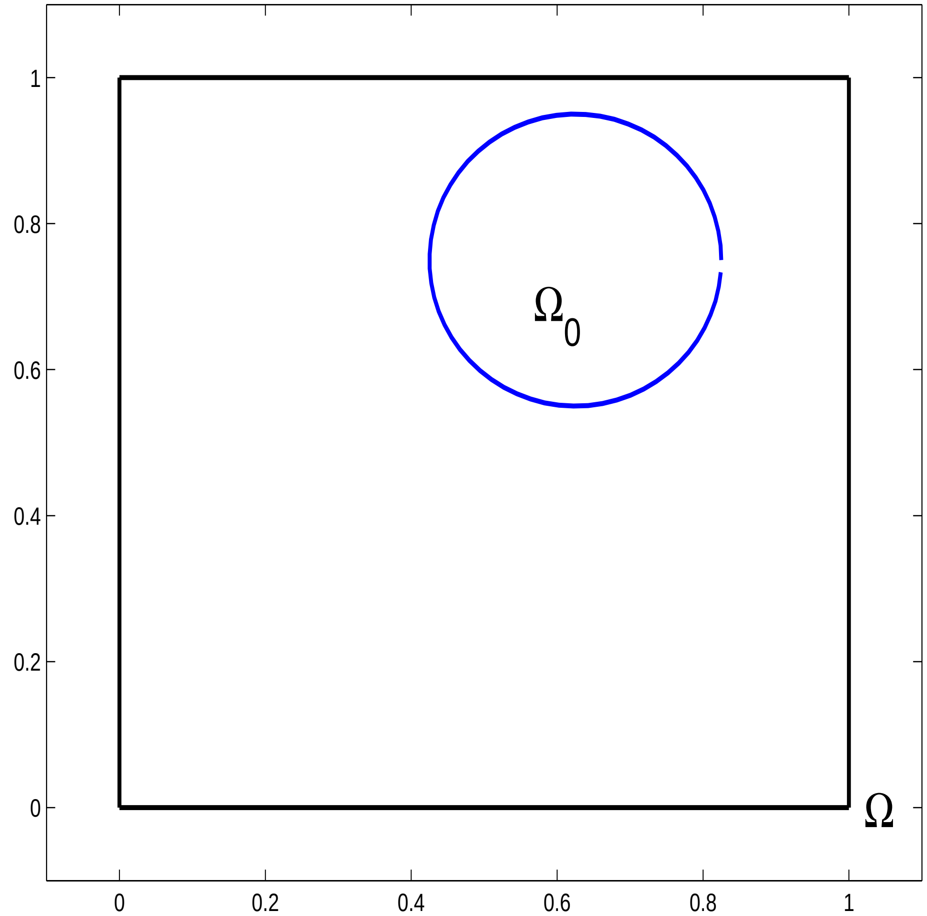
$$P_{in} = \begin{bmatrix} I & -K^T D^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} KD^{-1}K^T & 0 \\ 0 & -D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}K & I \end{bmatrix}$$

Practical cases:

$\widetilde{K} \approx K$, e.g., (algebraic) multigrid method

$\widetilde{D} \approx D$, e.g., $\widetilde{D} =$diag$(\frac{1}{2\beta}M)$

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. Minimal Residual method. No of iterations.

$\bar{M}$ singular, $K \in \mathbb{R}^{n \times n}$ Laplace operator

2D:

	$\beta = 10^{-5}$		$\beta = 10^{-2}$	
n	P_d	P_{in}	P_d	P_{in}
225	47	17	36	6
3969	43	17	33	6
65025	40	17	33	6

Numerical results. Minimal Residual method. No of iterations.

$\bar{M}$ singular, $K \in \mathbb{R}^{n \times n}$ Laplace operator

2D:

	$\beta = 10^{-5}$		$\beta = 10^{-2}$	
n	P_d	P_{in}	P_d	P_{in}
225	47	17	36	6
3969	43	17	33	6
65025	40	17	33	6

3D:

	$\beta = 10^{-5}$		$\beta = 10^{-2}$	
n	P_d	P_{in}	P_d	P_{in}
343	60	16	65	5
3375	62	16	71	5
29791	62	16	72	5

mesh independence can be proved

Conclusions

- Solution to Saddle Point systems largely expanding topic

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

- Structured preconditioning allows one to attack very large dimension pbs

- Practical solution optimality is very problem dependent

- Interplay between Solvers and Preconditioners is crucial