Spectral Properties of Saddle Point Linear Systems and Relations to Iterative Solvers

Part I: Spectral Properties

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

Outline of the 3-hour Presentation

- Schematic presentation of certain algebraic preconditioners (Today)
- Iterative solvers. Some (hopefully) helpful considerations... (Tomorrow)
- Spectral analysis of nonsymmetric preconditioners (Last Talk)

The problem

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

- 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
- ... Survey: Benzi, Golub and Liesen, Acta Num 2005

The problem. Simplifications

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

To make things simple:

- ⋆ A symmetric positive (semi)definite
- ⋆ B^T tall, possibly rank deficient
- ⋆ C symmetric positive (semi)definite
- ⋆ Warning: we shall use $g = 0$ in some cases

Spectral properties

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

$\sigma(\mathcal{M})$ 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

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

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

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

A nonsingular

Spectral properties

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

$0 = \lambda_n \leq \cdots \leq \lambda_1$ eigs of A

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

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

$$\left[\frac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2})\right] \quad \cup \quad \left[\alpha_0, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

A singular but $\frac{u^T A u}{u^T u} > \alpha_0 > 0$, $u \in \mathrm{Ker}(B)$

Spectral properties

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

$0 < \lambda_n \leq \cdots \leq \lambda_1$ eigs of A

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

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

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

B full rank

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

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

Spectral properties

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

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

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

Good (= slim) spectrum: $\lambda_1 \approx \lambda_n, \quad \sigma_1 \approx \sigma_m$

Spectral properties

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

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

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

Good (= slim) spectrum: $\lambda_1 \approx \lambda_n, \quad \sigma_1 \approx \sigma_m$

e.g.

$$\mathcal{M} = \begin{bmatrix} I & U^T \\ U & O \end{bmatrix}, \quad UU^T = I$$

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

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. $\sigma(\mathcal{P}_0^{-\frac{1}{2}} \mathcal{M} \mathcal{P}_0^{-\frac{1}{2}}) = \{1, 1/2 \pm \sqrt{5}/2\}$

Block diagonal Preconditioner

$\star$ 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. $\quad \sigma(\mathcal{P}_0^{-\frac{1}{2}} \mathcal{M} \mathcal{P}_0^{-\frac{1}{2}}) = \{1, 1/2 \pm \sqrt{5}/2\}$

A more practical choice:

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{S} \end{bmatrix} \qquad \text{spd.} \quad \widetilde{A} \approx A \qquad \widetilde{S} \approx BA^{-1}B^T$$

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

Still an Indefinite Problem

Giving up symmetry ...

- Change the preconditioner: *Mimic the LU factors*

$$\mathcal{M} = \begin{bmatrix} I & O \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix}$$

Giving up symmetry ...

- Change the preconditioner: *Mimic the LU factors*

$$\mathcal{M} = \begin{bmatrix} I & O \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix}$$

- Change the preconditioner: *Mimic the Structure*

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \Rightarrow \mathcal{M} \approx \mathcal{P}$$

Giving up symmetry ...

- Change the preconditioner: *Mimic the LU factors*

$$\mathcal{M} = \begin{bmatrix} I & O \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix}$$

- Change the preconditioner: *Mimic the Structure*

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \mathcal{M}$$

- Change the matrix: *Eliminate indef.* $\quad \mathcal{M}_- = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix}$

Giving up symmetry ...

- Change the preconditioner: *Mimic the LU factors*

$$\mathcal{M} = \begin{bmatrix} I & O \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \begin{bmatrix} A & B^T \\ O & BA^{-1}B^T + C \end{bmatrix}$$

- Change the preconditioner: *Mimic the Structure*

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \Rightarrow \mathcal{P} \approx \mathcal{M}$$

- Change the matrix: *Eliminate indef.* $\quad \mathcal{M}_- = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix}$

- Change the matrix: *Regularize* $\ (C = 0)$

$$\mathcal{M} \ \Rightarrow \ \mathcal{M}_\gamma = \begin{bmatrix} A & B^T \\ B & -\gamma W \end{bmatrix} \text{ or } \mathcal{M}_\gamma = \begin{bmatrix} A + \frac{1}{\gamma} B^T W^{-1} B & B^T \\ B & O \end{bmatrix}$$

... But recovering symmetry in disguise

Nonstandard inner product:

Let $\mathcal{W}$ be any of $\mathcal{M}\mathcal{P}^{-1}, \mathcal{M}_{-}$

For $\sigma(\mathcal{W})$ in $\mathbb{R}^{+}$, find sym matrix H such that

$$\mathcal{W}H = H\mathcal{W}^T$$

(that is, $\mathcal{W}$ is H-symmetric)

... But recovering symmetry in disguise

Nonstandard inner product:

Let $\mathcal{W}$ be any of $\mathcal{M}\mathcal{P}^{-1}, \mathcal{M}_-$

For $\sigma(\mathcal{W})$ in $\mathbb{R}^+$, find sym matrix H such that

$$\mathcal{W}H = H\mathcal{W}^T$$

(that is, $\mathcal{W}$ is H-symmetric)

If H is spd then

- $\mathcal{W}$ is diagonalizable
- Use PCG on $\mathcal{W}$ with H-inner product

Triangular preconditioner

$$A \text{ spd}, \quad \mathcal{P} = \begin{bmatrix} \widetilde{A} & B^T \\ 0 & -\widetilde{C} \end{bmatrix} \qquad \widetilde{A} \approx A, \quad \widetilde{C} \approx BA^{-1}B^T + C$$

Ideal case: $\widetilde{A} = A,\ \widetilde{C} = BA^{-1}B^T + C \quad \Rightarrow \quad \mathcal{M}\mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix}$

Triangular preconditioner

$$A \text{ spd}, \quad \mathcal{P} = \begin{bmatrix} \widetilde{A} & B^T \\ 0 & -\widetilde{C} \end{bmatrix} \qquad \widetilde{A} \approx A, \quad \widetilde{C} \approx BA^{-1}B^T + C$$

Ideal case: $\widetilde{A} = A,\ \widetilde{C} = BA^{-1}B^T + C \quad \Rightarrow \quad \mathcal{M}\mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix}$

Recovering symmetry?

- If $\widetilde{C} = C$ nonsing., then $\sigma(\mathcal{M}\mathcal{P}^{-1})$ in $\mathbb{R}^+$
- If $\widetilde{A} < A$ then $\sigma(\mathcal{M}\mathcal{P}^{-1})$ in $\mathbb{R}^+$ with

$$\lambda \in [\chi_1, \chi_2] \ni 1, \qquad \chi_j = \chi_j((B^T\widetilde{A}^{-1}B + C)\widetilde{C}^{-1}, \widetilde{A}^{-1}A)$$

Constraint (Indefinite) Preconditioner

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

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 $B^TC = 0$ 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, $\widetilde{A} \approx A + B^T \widetilde{C}^{-1} B$

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

Dual-based: $(C = O)$ $\widetilde{A} \approx A$, $\widetilde{C} = S - B\widetilde{A}^{-1}B^T$ 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 $H = \mathrm{blkdiag}(\widetilde{A} - A, B\widetilde{A}^{-1}B^T - S)$

The "minus-signed" Problem

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

B full rank $\Rightarrow \mathcal{M}_-$ positive stable $\Rightarrow$ eigs in $\mathbb{C}^+$

B

The "minus-signed" Problem

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

B full rank $\Rightarrow \mathcal{M}_-$ positive stable $\Rightarrow$ eigs in $\mathbb{C}^+$

Important facts

- $\mathcal{M}_-$ has always at least $n - m$ real eigs
- If $2\|B\| < \lambda_{\min}(A) - \lambda_{\max}(C) \Rightarrow$ all eigs real and positive
  for $C = 0$, condition simplifies: $\lambda_{\min}(A) > 4\lambda_{\max}(B^T A^{-1} B)$
- If B full rank and $\lambda_{\min}(A) > \lambda_{\max}(C)$
  apon scaling all eigs real and positive

The "minus-signed" Problem

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

B full rank $\Rightarrow \mathcal{M}_-$ positive stable $\Rightarrow$ eigs in $\mathbb{C}^+$

Important facts

- $\mathcal{M}_-$ has always at least $n - m$ real eigs
- If $2\|B\| < \lambda_{\min}(A) - \lambda_{\max}(C) \Rightarrow$ all eigs real and positive
  for $C = 0$, condition simplifies: $\lambda_{\min}(A) > 4\lambda_{\max}(B^T A^{-1} B)$
- If B full rank and $\lambda_{\min}(A) > \lambda_{\max}(C)$
  apon scaling all eigs real and positive

Hermitian-Skew-Hermitian preconditioners

Block diagonal preconditioner + nonsym solver

Regularized Problem

Augmented Lagrangian approach:

$$\mathcal{M}_\gamma = \begin{bmatrix} A + \frac{1}{\gamma} B^T W^{-1} B & B^T \\ B & O \end{bmatrix}$$

Particularly interesting for A indefinite or singular

$\star$ Any of the above preconditioners may be used.

Regularized Problem

Augmented Lagrangian approach:

$$\mathcal{M}_\gamma = \begin{bmatrix} A + \frac{1}{\gamma} B^T W^{-1} B & B^T \\ B & O \end{bmatrix}$$

Particularly interesting for A indefinite or singular

⋆ Any of the above preconditioners may be used.

Somehow related preconditioner for $\mathcal{M} = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}$:

$$\mathcal{P} = \begin{bmatrix} A + B^T W^{-1} B & B^T \\ O & W \end{bmatrix}$$

Questions (to be answered)

- ⋆ Which formulation/preconditioner for which iterative solver?

Questions (to be answered)

- ⋆ Which formulation/preconditioner for which iterative solver?
- ⋆ Is the theoretical spectral information useful in practice?

Questions (to be answered)

- ⋆ Which formulation/preconditioner for which iterative solver?
- ⋆ Is the theoretical spectral information useful in practice?
- ⋆ Are the imposed “constraints” needed?