Structured Preconditioners for Symmetric Saddle Point Problems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

Partly joint work with M. Benzi, Emory University

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
- Elasticity problems
- Mixed (FE) formulations of II and IV order elliptic PDEs
- Linearly Constrained Programs
- Linear Regression in Statistics
- Weighted Least Squares (Image restoration)
- ... Survey: Benzi, Golub and Liesen, Acta Num 2005

Spectral properties. 1

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

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

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

Spectral properties. 2

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

(Silvester & Wathen 1994)

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

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

Additional results for more general cases (e.g., A sing., A nonsym, ...)

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 for $\mathcal{P}$ should be low
  - Properties (algebraic/functional) exploited

Block diagonal, block triangular, block "constraint" preconditioners

Structure preserving preconditioning: Ideal cases

⋆ A nonsing., $C = 0$:

$$\mathcal{P} = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix} \quad \Rightarrow \quad \mathcal{M}\mathcal{P}^{-1} \quad \text{eigs: } 1, \frac{1}{2} \pm \frac{\sqrt{5}}{2}$$

MINRES converges in at most 3 iterations (Murphy, Golub & Wathen '02)

⋆ A nonsing., $C \neq 0$:

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

GMRES converges in at most 2 iterations

Ideal preconditioners are not feasible.

Approximations to ideal precs. are feasible - Perturbed behavior

Block diagonal Preconditioner

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{C} \end{bmatrix} \qquad \text{sym. pos. def.}$$

$$\widetilde{A} \approx A \qquad \widetilde{C} \approx BA^{-1}B^T + C$$

$\lambda \neq 0$ eigs of $\mathcal{P}^{-\frac{1}{2}}\mathcal{M}\mathcal{P}^{-\frac{1}{2}}$ (sym. indef.)

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

Rusten Winther (1992), Silvester Wathen (1993-1994), ...

Triangular preconditioner

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

(Bramble & Pasciak, Elman, Klawonn, Axelsson & Neytcheva, Simoncini)

Spectrum of $\mathcal{M}\mathcal{P}^{-1}$:

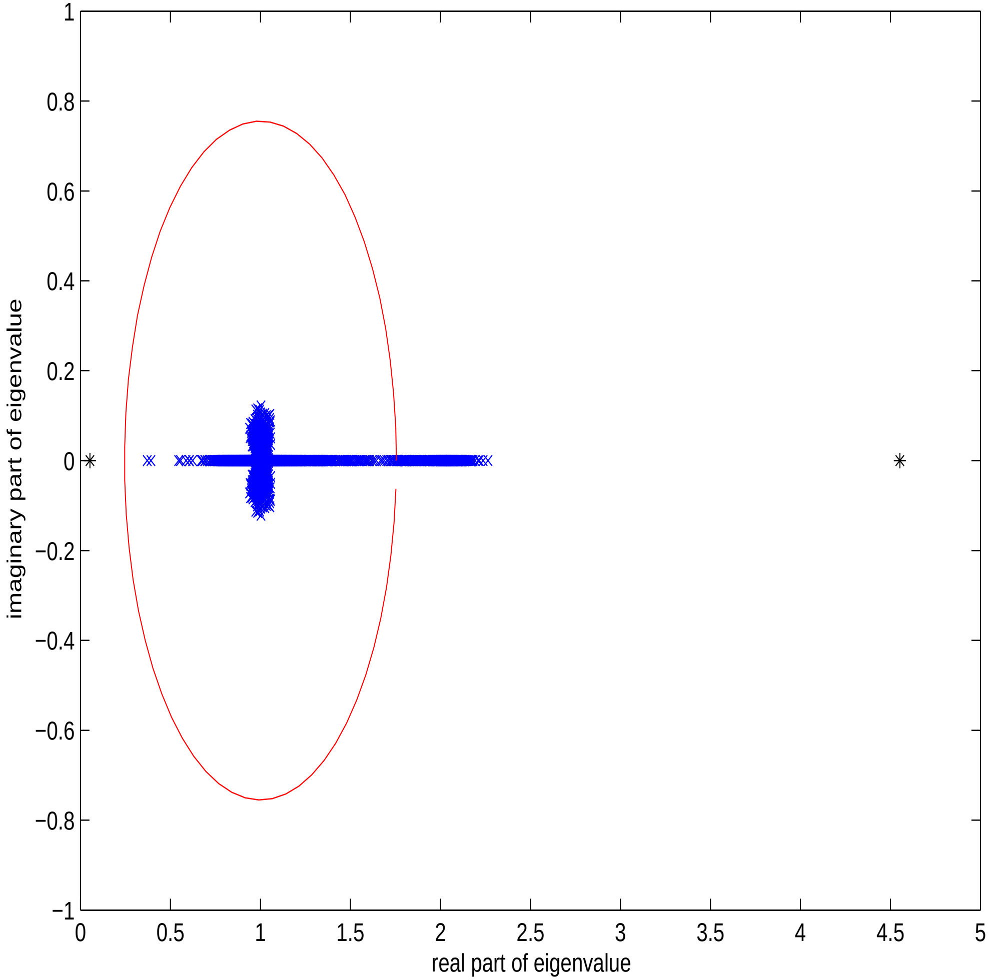
small complex cluster around 1 $\cup$ real interval :

More precisely: $\theta \in \Lambda(\mathcal{M}\mathcal{P}^{-1})$

$$\Im(\theta) \neq 0 \quad \Rightarrow |\theta - 1| \leq \sqrt{1 - \lambda_{\min}(A\widetilde{A}^{-1})} \qquad \text{(if } 1 - \lambda_{\min}(A\widetilde{A}^{-1}) \geq 0\text{)}$$

$$\Im(\theta) = 0 \quad \Rightarrow \theta \in [\chi_1, \chi_2] \text{ with } 1 \in [\chi_1, \chi_2]$$

An example



size: 2208, $\widetilde{A} = \texttt{cholinc}(A, 10^{-2})$, $\widetilde{C} = \texttt{cholinc}(C, 10^{-2})$

More to show? An open problem

$$\begin{bmatrix} A & B^T \\ 0 & -C \end{bmatrix} \quad \Rightarrow \quad \mathcal{P} = \begin{bmatrix} \widetilde{A} & B^T \\ 0 & -\widetilde{C} \end{bmatrix}$$

⋆ Pretty clear understandanding of perturbed spectrum

Eigenvectors???

Constraint Preconditioner

$$\mathcal{M}\mathcal{Q}^{-1} \qquad \mathcal{Q} = \begin{bmatrix} \widetilde{A} & B^T \\ B & -C \end{bmatrix}$$

Axelsson ('79), Ewing Lazarov Lu Vassilevski ('90), many papers after '97

Remark: $[B, -C]x_k = 0$ for all iterates x_k (constraint)

$\lambda \neq 0$ eigs of $\mathcal{M}\mathcal{Q}^{-1}$: $\qquad \lambda \in \mathbb{R}^+, \qquad \lambda \in \{1\} \cup [\alpha_0, \alpha_1]$

Detailed results on eigen/principal vectors (Dollar tr05)

More general factorizations (Dollar& Gould & Schilders & Wathen 06)

The feasible preconditioner

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

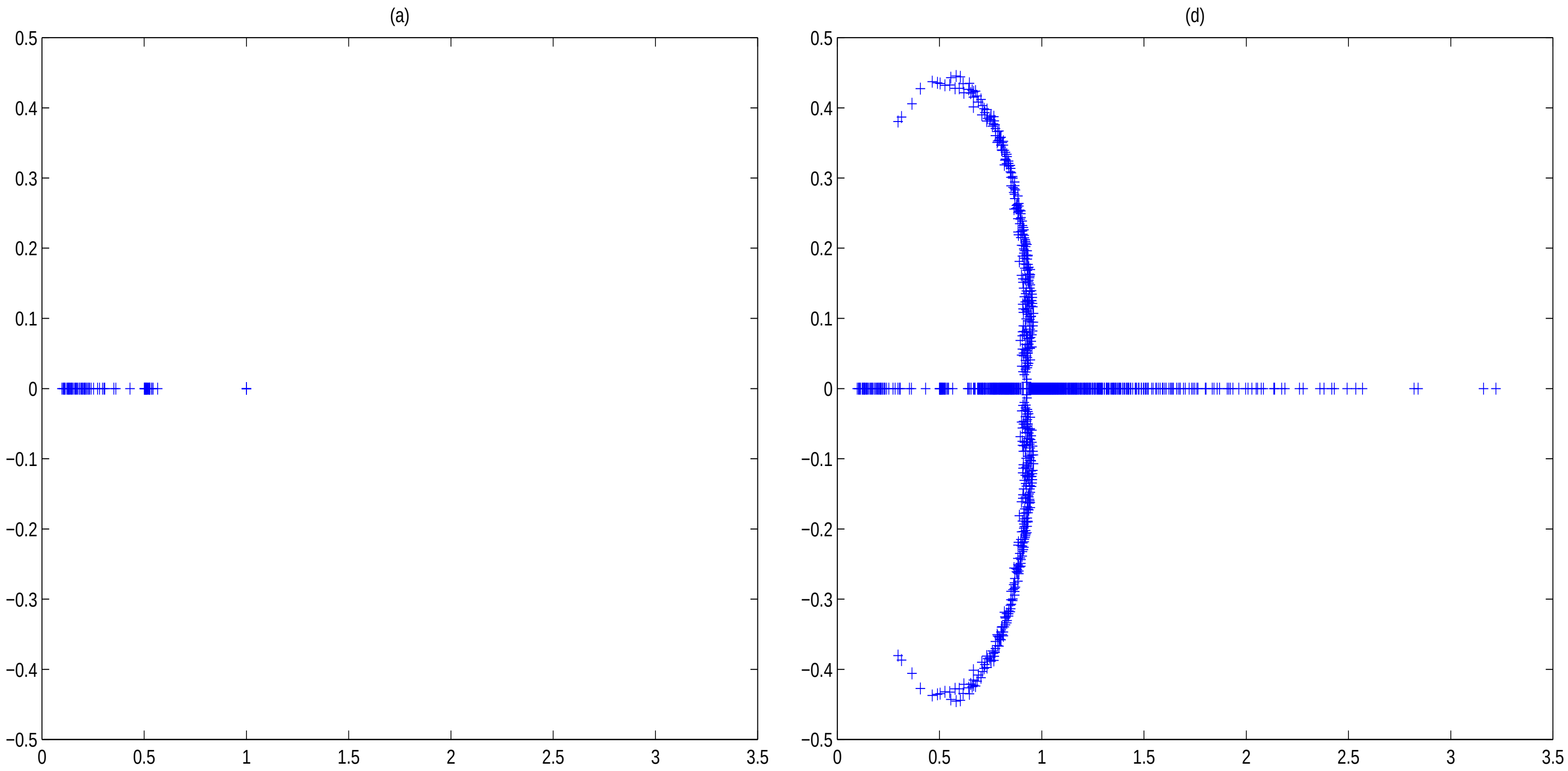
$$\mathcal{Q}^{-1} = \begin{bmatrix} I & -B^T \\ O & I \end{bmatrix} \begin{bmatrix} I & O \\ O & -(\mathbf{BB^T} + \mathbf{C})^{-1} \end{bmatrix} \begin{bmatrix} I & O \\ -B & I \end{bmatrix}$$

($\widetilde{A} = I$ if prescaling used)

$BB^T + C \approx H$ e.g., $H =$ `cholinc`$(BB^T + C$,`tol`$)$

Question: How much does H affect the preconditioner ?

Spectrum of perturbed matrix



$$\|(BB^T + C) - H\|_\infty \approx 2 \cdot 10^{-1} \|BB^T + C\|_\infty$$

Eigenvalue bounds

$C = 0 \Rightarrow H \approx BB^T$, A spd.

Let $\widehat{S} = B(2I - A)B^T H^{-1}$ (scale A so that $\widehat{S} \geq 0$)

⋆ If $\Im(\lambda) \neq 0$ then

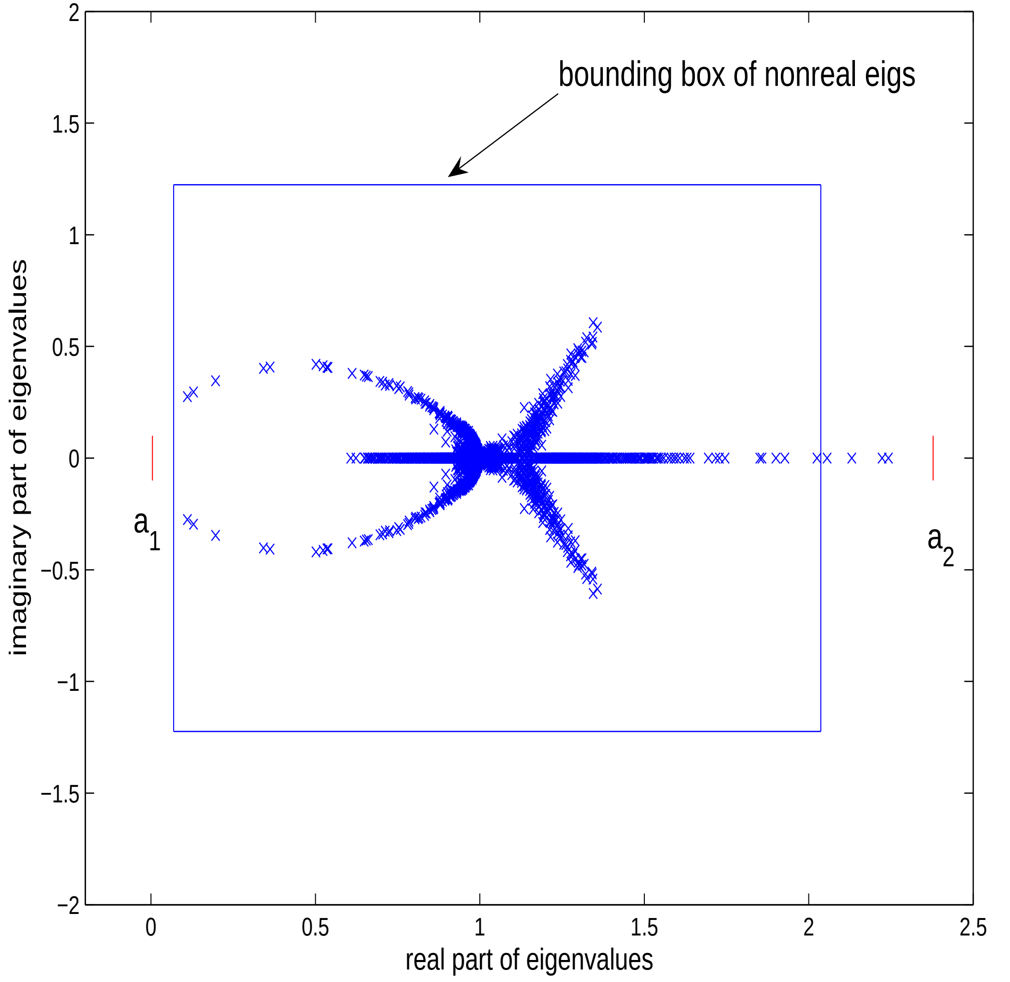
$$\begin{aligned} \frac{1}{2}(\lambda_{\min}(A) + \lambda_{\min}(\widehat{S})) \leq \quad & \Re(\lambda) & \leq \frac{1}{2}(\lambda_{\max}(A) + \lambda_{\max}(\widehat{S})) \\ & |\Im(\lambda)| & \leq \sigma_{\max}((I - A)B^T H^{-\frac{1}{2}}). \end{aligned}$$

⋆ If $\Im(\lambda) = 0$ then

$$\min\{\lambda_{\min}(A), \lambda_{\min}(\widehat{S})\} \leq \lambda \leq \max\{\lambda_{\max}(A), \lambda_{\max}(\widehat{S})\}$$

Benzi Simoncini '06

Spectral bounds



$$H = \texttt{cholinc}(BB^T,\, 10^{-2})$$

Some algebraic details for $\mathcal{M}\widehat{\mathcal{Q}}^{-1}$

$$\begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \widehat{\mathcal{Q}} \begin{bmatrix} x \\ y \end{bmatrix} \qquad \widehat{\mathcal{Q}} = \begin{bmatrix} I & O \\ B & I \end{bmatrix} \begin{bmatrix} I & O \\ O & -H \end{bmatrix} \begin{bmatrix} I & B^T \\ O & I \end{bmatrix}$$

can be written as

$$\begin{bmatrix} I & O \\ -B & I \end{bmatrix} \begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \begin{bmatrix} I & -B^T \\ O & I \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} I & O \\ O & -H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$\begin{bmatrix} A & (I-A)B^T \\ B(I-A) & -B(2I-A)B^T \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} I & O \\ O & -H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Some algebraic details for $\mathcal{M}\widehat{\mathcal{Q}}^{-1}$. cont'ed

$$\begin{bmatrix} A & (I-A)B^T \\ B(I-A) & -B(2I-A)B^T \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} I & O \\ O & -H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Using $BB^T = H + E$, this can be written as

$$\left(\begin{bmatrix} I_n \\ B \end{bmatrix} (A - I_n) \left[I_n, -B^T \right] + \begin{bmatrix} O & O \\ O & E \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix} = (\lambda - 1) \begin{bmatrix} I_n & O \\ O & H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

⋆ For $E = 0$ (i.e. $H = BB^T$) we readily recover the known exact case

⋆ For $E \neq O$, all eigenvector blocks u with $Bu \neq 0$ may give rise to nonreal eigenvalues

⋆ $E \neq 0$ affects the null space of the low rank matrix ($\lambda = 1$)

More to show? A second open problem

$$\widehat{\mathcal{Q}} = \begin{bmatrix} I & O \\ B & I \end{bmatrix} \begin{bmatrix} I & O \\ O & -H \end{bmatrix} \begin{bmatrix} I & B^T \\ O & I \end{bmatrix}$$

⋆ Pretty clear understandanding of perturbed spectrum

Eigenvectors???