Spectral analysis of saddle point matrices with indefinite leading blocks

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

Partially joint work with Nick Gould, RAL

The problem

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

- 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

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

Hypotheses:

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

More hypotheses later...

Why are we interested in spectral bounds?

- To detect "sensitive" blocks in the coeff. matrix

  (guidelines for preconditioning strategies)

Why are we interested in spectral bounds?

- To detect "sensitive" blocks in the coeff. matrix
  (guidelines for preconditioning strategies)

- To "tune" the stabilization parameter (matrix C)

Why are we interested in spectral bounds?

- To detect “sensitive” blocks in the coeff. matrix
  (guidelines for preconditioning strategies)

- To “tune” the stabilization parameter (matrix C)

- To predict convergence behavior of the iterative solver

Iterative solver. Convergence considerations.

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

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

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

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

Iterative solver. Convergence considerations.

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

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

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

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

Iterative solver. Convergence considerations.

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

$\mathcal{M}$ is nonsymmetric and indefinite $\quad\longrightarrow\quad$ GMRES

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

Iterative solver. Convergence considerations.

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

$\mathcal{M}$ is nonsymmetric and indefinite $\quad \longrightarrow \quad$ GMRES

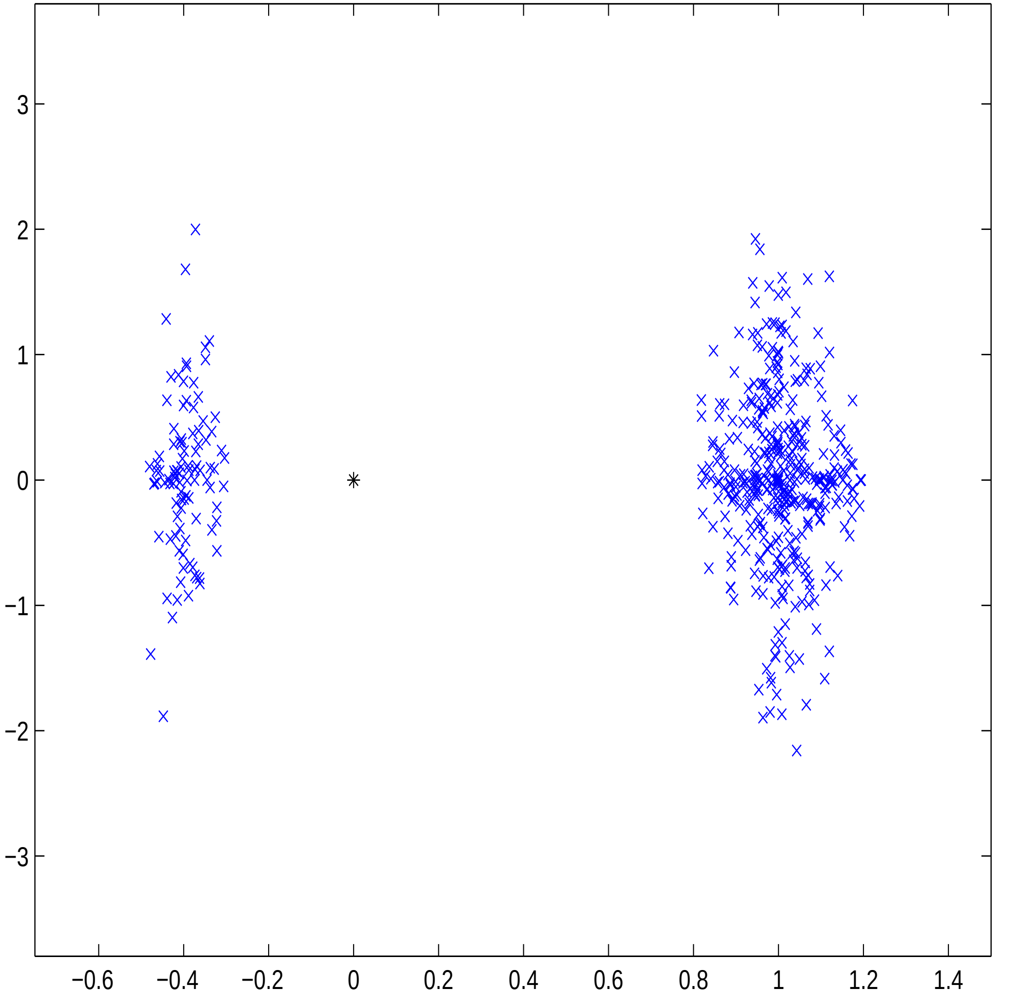
$$x_k \in x_0 + K_k(\mathcal{M}, r_0), \quad \text{s.t.} \qquad \min \|b - \mathcal{M}x_k\|$$

For $\mathcal{M}$ non-normal indefinite :

- In theory, complete stagnation is possible;
- Rule of thumb: tight spectral clusters help

Note: $\mathcal{M}$ indefinite $\Rightarrow$ Elman's bound not applicable

Rule of thumb: clustering helps



GMRES: Nonstagnation condition (Simoncini & Szyld, '08)

Let $H = \frac{1}{2}(\mathcal{M} + \mathcal{M}^\top)$, $S = \frac{1}{2}(\mathcal{M} - \mathcal{M}^\top)$. If

$$H \text{ nonsingular and } \quad \|SH^{-1}\| < 1$$

Then

$$\|r_2\| \leq \left(1 - \frac{\theta_{\min}^2}{\|\mathcal{M}^2\|^2}\right)^{\frac{1}{2}} \|r_0\| \quad \theta_{\min} = \lambda_{\min}(\tfrac{1}{2}(\mathcal{M}^2 + (\mathcal{M}^2)^\top)) > 0$$

The same relation holds at every other iteration

$\mathcal{M}$ symmetric indefinite. Well-exercised spectral properties

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

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

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

$\mathcal{M}$ symmetric indefinite. Well-exercised spectral properties

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

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

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

A positive definite

$\mathcal{M}$ symmetric indefinite. Well-exercised spectral properties

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

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

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

A semidefinite but $\frac{u^\top A u}{u^\top u} > \alpha_0 > 0$, $u \in \text{Ker}(B)$ Perugia & S., '00

$\mathcal{M}$ symmetric indefinite. Well-exercised spectral properties

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

$\mu(\mathcal{M})$ 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

$\mathcal{M}$ symmetric indefinite. Well-exercised spectral properties

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

$\mu(\mathcal{M})$ subset of (Silvester & Wathen 1994)

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

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

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

Spectral properties. Interpretation.

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

$\mu(\mathcal{M})$ 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. Interpretation.

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

$\mu(\mathcal{M})$ 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.

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

Block diagonal Preconditioner

⋆ A spd, $C = 0$:

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

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

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

Block diagonal Preconditioner

⋆ A spd, $C = 0$:

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

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

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

A more practical choice:

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

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

Still an Indefinite Problem, but possibly much easier to solve

Indefinite A

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

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

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

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

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

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

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

Indefinite A, $C = 0$. Cont'd

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

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

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

Sharpness of the bounds

Ex.1. $A = \begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}, \; B^\top = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} \; \mu(\mathcal{M}) = \{-1.5441, 0.0014257, 4.5427\}$

Ex.2. $A = \begin{bmatrix} 0.01 & 3 \\ 3 & -0.01 \end{bmatrix}, \; B = [0, \; 3] \; \mu(\mathcal{M}) = \{-4.2452, 5.0 \cdot 10^{-3}, 4.2402\}$

Ex.3. $A = \begin{bmatrix} 1 & -4 & 0 \\ -4 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, B^\top = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \; \mu(\mathcal{M}) = \begin{matrix} \{-4.3528, & & -0.22974, \\ 0.22974, & 2, & 4.3528\} \end{matrix}$

case	λ_n	λ_1	α_0	σ_m, σ_1	$\mathcal{I}^-$	$\mathcal{I}^+$
Ex.1	-1.5414	4.5414	1.0	0.1	[-1.5478, -0.0022]	[0.0004, 4.5436]
Ex.2	-3.0000	3.0000	0.01	3	[-4.8541, -1.8541]	[$4.9917 \cdot 10^{-3}$, 4.8541]
Ex.3	-4.1231	4.1231	2.0	1	[-4.3528, -0.22974]	[0.0762, 4.3528]

Augmenting the (1,1) block

Equivalent formulation ($C = 0$):

$$\begin{bmatrix} A + \tau B^\top B & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a + \tau B^\top b \\ b \end{bmatrix}, \quad \tau \in \mathbb{R}$$

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

Augmenting the (1,1) block

Equivalent formulation ($C = 0$):

$$\begin{bmatrix} A + \tau B^\top B & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a + \tau B^\top b \\ b \end{bmatrix}, \quad \tau \in \mathbb{R}$$

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

Augmenting the (1,1) block

Assume "good" τ is taken.

$$\begin{bmatrix} A + \tau B^\top B & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a + \tau B^\top b \\ b \end{bmatrix}, \quad \tau \in \mathbb{R}$$

Spectral intervals for (1,1) spd may be obtained

“Regularized” problem

$$\begin{bmatrix} A & B^\top \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a + \tau B^\top b \\ b \end{bmatrix}, \qquad \tau \in \mathbb{R}$$

Coefficient matrix: $\mathcal{M}_C$

Warning: for A indefinite, conditions on C required:

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \text{singular!}$$

Note: Perturbation results yield spectral bounds assuming $\lambda_{\max}^C < \Gamma$

“Regularized” problem

More accurate result:

If $\lambda_{\max}^{C} < \frac{\alpha_0 \sigma_m^2}{\|A\|^2 - \lambda_n \alpha_0}$, then $\mu(\mathcal{M}_C) \subset \mathcal{I}^- \cup \mathcal{I}^+$ with

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

$$\mathcal{I}^+ = \left[\Gamma_C, \frac{1}{2}\left(\lambda_1 + \sqrt{(\lambda_1)^2 + 4\sigma_1^2}\right)\right] \subset \mathbb{R}^+,$$

For $m = n$, $\quad \Gamma_C = \frac{1}{2}\left(\lambda_n - \lambda_{\max}^C + \sqrt{(\lambda_n + \lambda_{\max}^C)^2 + 4\sigma_m^2}\right)$

more complicated (but explicit!) estimate for $m < n$

“Regularized” problem

An example:

$$\mathcal{M}_C = \begin{bmatrix} \lambda_n & 0 & \sigma \\ 0 & \lambda_1 & 0 \\ \sigma & 0 & -\gamma^C \end{bmatrix},$$

with $\lambda_n < 0, \lambda_1 > 0, \sigma > 0$. If $\gamma^C = -\sigma^2/\lambda_n$ then $\mathcal{M}_C$ is singular.

“Regularized” problem

An example:

$$\mathcal{M}_C = \begin{bmatrix} \lambda_n & 0 & \sigma \\ 0 & \lambda_1 & 0 \\ \sigma & 0 & -\gamma^C \end{bmatrix},$$

with $\lambda_n < 0, \lambda_1 > 0, \sigma > 0$. If $\gamma^C = -\sigma^2/\lambda_n$ then $\mathcal{M}_C$ is singular.

Our estimate requires (for $\|A\| = \alpha_0 = -\lambda_n$): $0 \leq \gamma^C \leq \frac{1}{2} \, \frac{-\sigma^2}{\lambda_n}$

(half the value from singularity!)

“Regularized” problem

An example:

$$\mathcal{M}_C = \begin{bmatrix} \lambda_n & 0 & \sigma \\ 0 & \lambda_1 & 0 \\ \sigma & 0 & -\gamma^C \end{bmatrix},$$

with $\lambda_n < 0, \lambda_1 > 0, \sigma > 0$. If $\gamma^C = -\sigma^2/\lambda_n$ then $\mathcal{M}_C$ is singular.

Our estimate requires (for $\|A\| = \alpha_0 = -\lambda_n$): $0 \le \gamma^C \le \frac{1}{2} \frac{-\sigma^2}{\lambda_n}$
(half the value from singularity!)

Related results: Bai, Ng, Wang ('09) qualitatively similar bound based on $B^\top C^{-1} B$, $A + B^\top C^{-1} B$ (no full rank hyp. on B)
Bai (tech.rep.'09)

Full rank assumption of B

In some optimization problems:

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} C_1 & 0 \\ 0 & 0 \end{bmatrix},$$

with positive definite C_1

Natural assumption: $A + B_1^\top C_1^{-1} B_1$ definite on the null space of the full-rank B_2. In this case,

$$\mathcal{M}_C = \begin{bmatrix} \begin{pmatrix} A & B_1^\top \\ B_1 & -C_1 \end{pmatrix} & \begin{pmatrix} B_2^\top \\ 0 \end{pmatrix} \\ \begin{pmatrix} B_2 & 0 \end{pmatrix} & 0 \end{bmatrix}.$$

Spectral analysis: Use Bai, Ng, Wang result to get spectral intervals for the "(1,1)" block, and then apply our bounds for $\mathcal{M}_C$

Application to ideal block diagonal preconditioners

Indefinite preconditioner, $C = 0$:

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

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

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

$$\mu(\mathcal{P}_-^{-1}\mathcal{M}) \subset \left\{1, \frac{1}{2}(1+i\sqrt{3}), \frac{1}{2}(1-i\sqrt{3})\right\} \subset \mathbb{C}^+.$$

Application to **practical** block diagonal preconditioners

Indefinite preconditioner, $C = 0$:

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 preconditioner, $C \neq 0$:

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

Application to ideal block diagonal preconditioners

Definite preconditioner, $C = 0$:

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

- 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.

General nonsymmetric problem

$$\mathcal{M} = \begin{bmatrix} F & B^\top \\ B & -\beta C \end{bmatrix} \qquad F \quad \text{nonsymmetric}$$

Preconditioning strategies (other alternatives are possible):

$$\mathcal{P}_{tr} = \begin{bmatrix} \widetilde{F} & B \\ & \pm\widetilde{C} \end{bmatrix} \qquad \mathcal{P}_d = \begin{bmatrix} \widetilde{F} & \\ & \pm\widetilde{C} \end{bmatrix} \text{ with } \widetilde{C} > 0$$

- $\widetilde{F} \approx F$
- $\widetilde{F} \approx F + B^\top \widetilde{C}^{-1} B$ (augmentation block precond.)

For $+\widetilde{C}$: $\qquad \mathcal{P}^{-1}\mathcal{M}$ indefinite

Augmentation block preconditioning

⋆ Appealing for F singular

For $+\widetilde{C}$:

$$\mathcal{P}_d^{-1}\mathcal{M},\ \mathcal{P}_{tr}^{-1}\mathcal{M} \text{ have clusters in } \mathbb{C}^- \text{ and } \mathbb{C}^+$$

$\Rightarrow$ Indefinite matrix $\Rightarrow$ Elman's bound not applicable

Analysis of clusters:

- Schötzau & Greif '06 (F sym)
- Cao '07

Nonstagnation condition revisited. Grcar tech.rep'89

Let ϕ_k be polynomial with $\phi_k(0) = 0$. If $\frac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top) > 0$ then

$$\|r_k\| \le \left(1 - \frac{\theta_{\min}^2}{\|\phi_k(\mathcal{M})\|^2}\right)^{\frac{1}{2}} \|r_0\| \quad \theta_{\min} = \lambda_{\min}(\tfrac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top))$$

Elman's bound: $k = 1$

Nonstagnation condition revisited. Grcar tech.rep'89

Let ϕ_k be polynomial with $\phi_k(0) = 0$. If $\frac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top) > 0$ then

$$\|r_k\| \le \left(1 - \frac{\theta_{\min}^2}{\|\phi_k(\mathcal{M})\|^2}\right)^{\frac{1}{2}} \|r_0\| \quad \theta_{\min} = \lambda_{\min}(\tfrac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top))$$

Elman's bound: $k = 1$

The simplest case: $k = 2$

If $\phi_2(H) > 0$, then $\phi_2(\mathcal{M}) > 0$ iff $\|S\phi_2(H)^{-1/2}\| < 1$

Nonstagnation condition revisited. Grcar tech.rep'89

Let ϕ_k be polynomial with $\phi_k(0) = 0$. If $\frac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top) > 0$ then

$$\|r_k\| \le \left(1 - \frac{\theta_{\min}^2}{\|\phi_k(\mathcal{M})\|^2}\right)^{\frac{1}{2}} \|r_0\| \qquad \theta_{\min} = \lambda_{\min}(\tfrac{1}{2}(\phi_k(\mathcal{M}) + \phi_k(\mathcal{M})^\top))$$

Elman's bound: $k = 1$

The simplest case: $k = 2$

If $\phi_2(H) > 0$, then $\phi_2(\mathcal{M}) > 0$ iff $\|S\phi_2(H)^{-1/2}\| < 1$

$\Rightarrow$ In Simoncini & Szyld '08: $\phi_2(\lambda) = \lambda^2$

$\Rightarrow$ Here: $\phi_2(\lambda) = \lambda(\lambda - \alpha)$, $\quad \alpha = \max\{0, \lambda_+(H) + \lambda_-(H)\}$

($\lambda_+(H), \lambda_-(H)$: closest pos/neg eigs to zero)

Example. Navier-Stokes problem

IFISS Package (Elman, Ramage, Silvester)

"Flow over a step". Uniform grid, Q1-P0 elements

Prec	blocks	$\lambda_{\min}(H)$	$\lambda_{\max}(S^\top S, \phi_2(H))$	α	# its
$P_{d,aug}$	$\widetilde{C}(0)$	-3.5512	0.9906	0.3951	16
	$\widetilde{C}(10^{-1})$	-2.7567	0.9724	0.4252	19
	Q	-4.2339	1.5620	0.3558	29
$P_{tr,aug}$	$\widetilde{C}(0)$	-3.8091	0.9672	0	14
	$\widetilde{C}(10^{-1})$	-3.0814	1.1063	0.0216	21
	$\widetilde{C}(10^{-2})$	-3.7450	0.97097	0	16
P_{tr}	$\widehat{F}, W(1)$	-7.3000	0.9923	0	11
	$\widehat{F}, W(0)$	-13.818	0.9924	0	17

$\widetilde{C}(\mathrm{tol}) = B\widetilde{F}^{-1}B^\top + \beta C \qquad \widetilde{F} = \mathrm{luinc}(F, \mathrm{tol})$

(2,2) block: $W(s_1) = B\widehat{F}^{-1}B^\top + s_1\beta C \qquad \widehat{F} = \mathrm{luinc}(F, 10^{-2})$

Example. Navier-Stokes problem

IFISS Package (Elman, Ramage, Silvester)

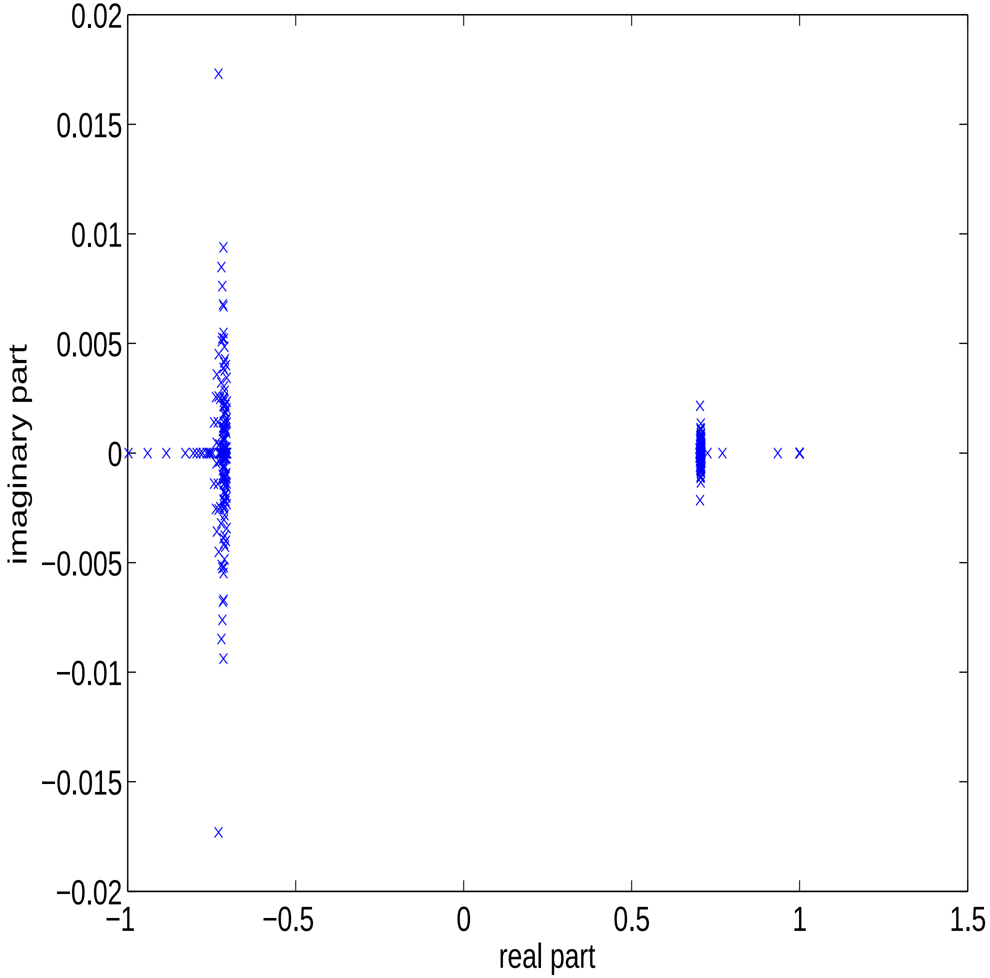
"Flow over a step". Uniform grid, Q1-P0 elements

Prec	blocks	$\lambda_{\min}(H)$	$\lambda_{\max}(S^\top S, \phi_2(H))$	α	# its
$P_{d,aug}$	$\widetilde{C}(0)$	-3.5512	0.9906	0.3951	16
	$\widetilde{C}(10^{-1})$	-2.7567	0.9724	0.4252	19
	Q	-4.2339	1.5620	0.3558	29
$P_{tr,aug}$	$\widetilde{C}(0)$	-3.8091	0.9672	0	14
	$\widetilde{C}(10^{-1})$	-3.0814	1.1063	0.0216	21
	$\widetilde{C}(10^{-2})$	-3.7450	0.97097	0	16
P_{tr}	$\widehat{F}, W(1)$	-7.3000	0.9923	0	11
	$\widehat{F}, W(0)$	-13.818	0.9924	0	17

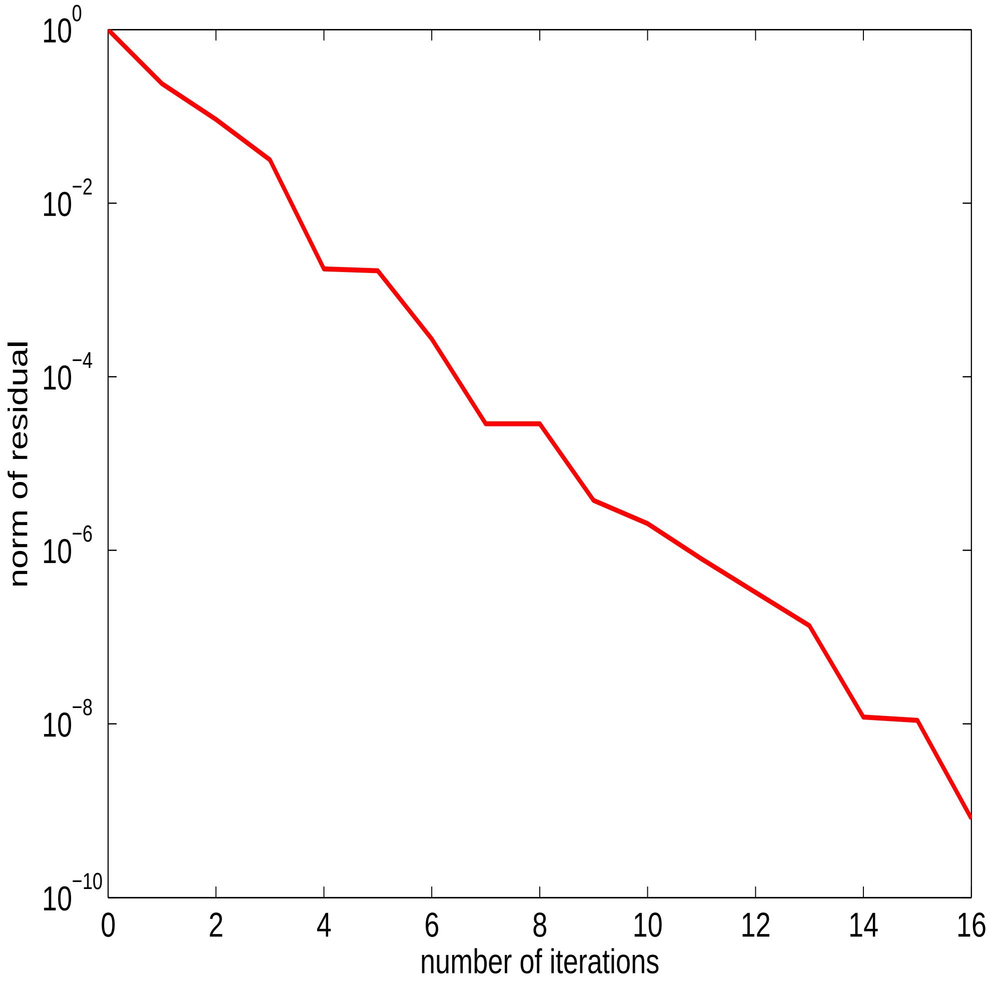
$\widetilde{C}(\text{tol}) = B\widetilde{F}^{-1}B^\top + \beta C \qquad \widetilde{F} = \text{luinc}(F, \text{tol})$

(2,2) block: $W(s_1) = B\widehat{F}^{-1}B^\top + s_1\beta C \qquad \widehat{F} = \text{luinc}(F, 10^{-2})$

Spectrum of $\mathcal{M}P_{tr,aug}^{-1}$



GMRES Convergence history



Mesh independence

$P_{tr,aug}$. (2,2) block: $\widetilde{C} = \beta C + BF^{-1}B^\top$,

n	m	$\lambda_{\min}(H)$	$\lambda_{\max}(S^\top S, \phi_2(H))$	α	# its
418	176	-3.8091	0.9672	0	14
1538	704	-3.7057	0.9662	0	15
5890	2816	-3.6710	0.9660	0	13

Final considerations and outlook

Symmetric case:

- Sharp bounds obtained for symmetric indefinite (1,1) block
- Future work: exploit this knowledge to devise and analyze effective preconditioners

Final considerations and outlook

Symmetric case:

- Sharp bounds obtained for symmetric indefinite (1,1) block
- Future work: exploit this knowledge to devise and analyze effective preconditioners

Nonsymmetric case:

- First attempt to provide convergence information on indefinite problem
- Future work: devise more complete convergence analysis

References

1. V. S. and Daniel B. Szyld , *New conditions for non-stagnation of minimal residual methods*. Numerische Mathematik, v. 109, n.3 (2008), pp. 477-487.

2. Nick Gould and V. S., *Spectral Analysis of saddle point matrices with indefinite leading blocks*. August 2008, To appear in SIAM J. Matrix Analysis Appl.

3. V.S., *On the non-stagnation condition for GMRES and application to saddle point matrices*, In preparation.