Structured Preconditioners for Saddle Point Problems

V. Simoncini

Dipartimento di Matematica
Università di Bologna
valeria@dm.unibo.it

Collaborators on this project

Mario Arioli, *RAL, UK*

Michele Benzi, *Emory University (GA)*

Ilaria Perugia, *Università di Pavia*

Miro Rozloznik, *Institute of Computer Science, Academy of Science, Prague*

Constrained minimization problem

$$\text{minimize} \quad J(u) = \frac{1}{2}\langle Au, u\rangle - \langle f, u\rangle$$

subject to $Bu = g$

$A \quad n \times n$ symmetric, $B\ m \times n$, $m \leq n$

$$\Downarrow$$

Lagrange multipliers approach

Karush-Kuhn-Tucker (KKT) system

Application problems

- 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)
- ...

The algebraic Saddle Point Problem

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

- A sym. pos.semidef., B full rank

The algebraic Saddle Point Problem

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

- A sym. pos.semidef., B full rank
- A sym. pos.semidef., B rank defic., C sym. [semi]def.

The algebraic Saddle Point Problem

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

- A sym. pos.semidef., B full rank
- A sym. pos.semidef., B rank defic., C sym. [semi]def.
- A nonsym., B and C as above

The algebraic Saddle Point Problem

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

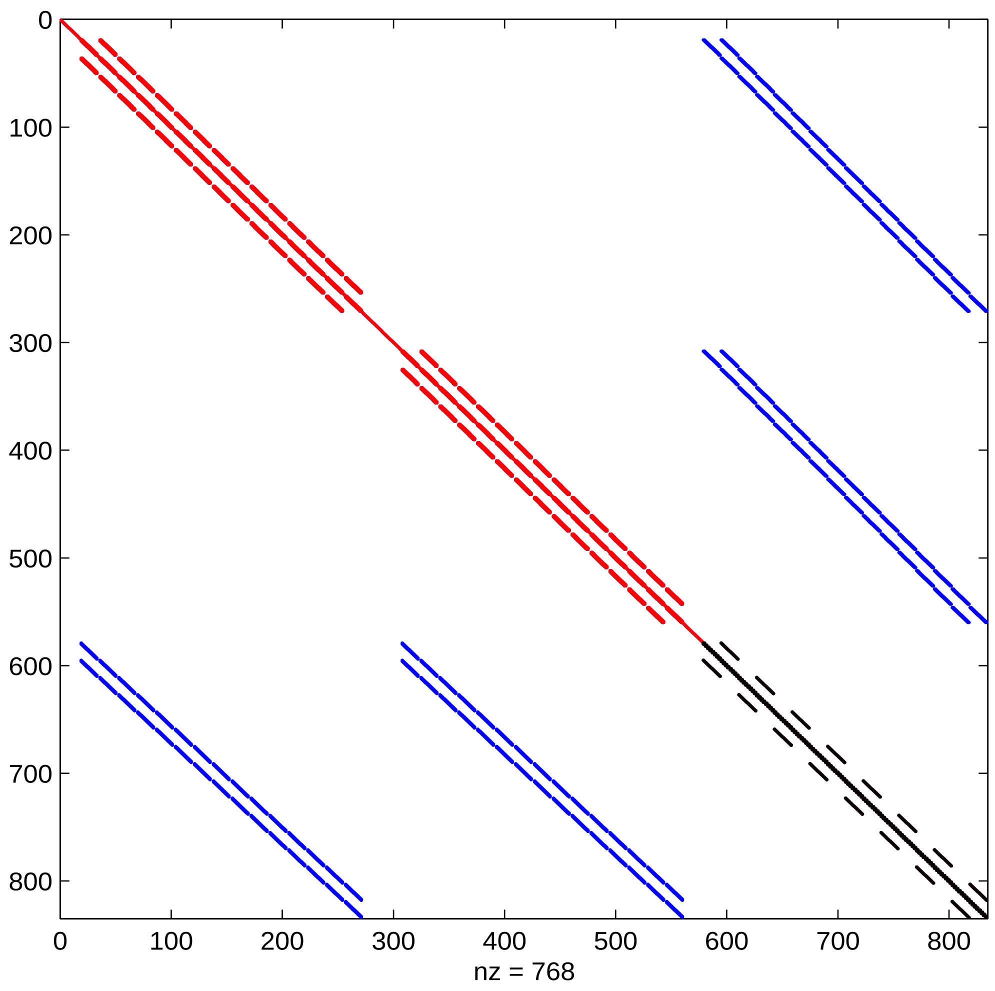
- A sym. pos.semidef., B full rank
- A sym. pos.semidef., B rank defic., C sym. [semi]def.
- A nonsym., B and C as above

A sym,

$$\mathcal{M}x = b \qquad \mathcal{M} \quad \text{sym. indef.}$$

With n positive and m negative real eigenvalues

Typical Sparsity pattern (3D problem)



Spectral properties

- $\mathcal{M} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$ $\quad$ $0 < \lambda_n \leq \cdots \leq \lambda_1$ eigs of A

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

Spectral properties

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

- (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

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

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

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

- (Silvester & Wathen 1994), $0 \le \sigma_m \le \cdots \le \sigma_1$

$$\lambda_n - \lambda_{\max}(C) - \sqrt{(\lambda_n + \lambda_{\max}(C))^2 + 4\sigma_1^2}$$

Spectral properties

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

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

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

- (Silvester & Wathen 1994), $0 \le \sigma_m \le \cdots \le \sigma_1$

$$\lambda_n - \lambda_{\max}(C) - \sqrt{(\lambda_n + \lambda_{\max}(C))^2 + 4\sigma_1^2}$$

More results for special cases (e.g. Perugia & S. 2000)

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 $\mathcal{M}u = b$

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

Idealized case:

Idealized case:

$\star$ A nonsing., $C = 0$:

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

MINRES converges in at most 3 iterations

(Murphy, Golub & Wathen, 2002)

Idealized case:

$\star$ A nonsing., $C = 0$:

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

MINRES converges in at most 3 iterations

(Murphy, Golub & Wathen, 2002)

$\star$ A nonsing., $C \neq 0$:

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

GMRES converges in at most 2 iterations

Block diagonal Preconditioner

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

Rusten Winther (1992), Silvester Wathen (1993-1994), Klawonn (1998)
Fischer Ramage Silvester Wathen (1998...), . . .

Block diagonal Preconditioner

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

Rusten Winther (1992), Silvester Wathen (1993-1994), Klawonn (1998)
Fischer Ramage Silvester Wathen (1998...), . . .

$\lambda \neq 0$ eigs of $\mathcal{P}^{-\frac{1}{2}}\mathcal{M}\mathcal{P}^{-\frac{1}{2}}$,

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

Constraint Preconditioner

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

Axelsson (1979), Ewing Lazarov Lu Vassilevski (1990), Braess Sarazin (1997) Golub Wathen (1998) Vassilevski Lazarov (1996), Lukšan Vlček (1998-1999), Perugia S. Arioli (1999), Keller Gould Wathen (2000) Perugia S. (2000), Gould Hribar Nocedal (2001), Rozloznik S. (2002), Durazzi Ruggiero (2003), Axelsson Neytcheva (2003), ...

Constraint Preconditioner

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

Axelsson (1979), Ewing Lazarov Lu Vassilevski (1990), Braess Sarazin (1997) Golub Wathen (1998) Vassilevski Lazarov (1996), Lukšan Vlček (1998-1999), Perugia S. Arioli (1999), Keller Gould Wathen (2000) Perugia S. (2000), Gould Hribar Nocedal (2001), Rozloznik S. (2002), Durazzi Ruggiero (2003), Axelsson Neytcheva (2003), ...

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

Constraint Preconditioner

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

Axelsson (1979), Ewing Lazarov Lu Vassilevski (1990), Braess Sarazin (1997) Golub Wathen (1998) Vassilevski Lazarov (1996), Lukšan Vlček (1998-1999), Perugia S. Arioli (1999), Keller Gould Wathen (2000) Perugia S. (2000), Gould Hribar Nocedal (2001), Rozloznik S. (2002), Durazzi Ruggiero (2003), Axelsson Neytcheva (2003), ...

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

Remark: $B^T x_k = 0$ for all iterates x_k

Constraint Preconditioner

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

Axelsson (1979), Ewing Lazarov Lu Vassilevski (1990), Braess Sarazin (1997) Golub Wathen (1998) Vassilevski Lazarov (1996), Lukšan Vlček (1998-1999), Perugia S. Arioli (1999), Keller Gould Wathen (2000) Perugia S. (2000), Gould Hribar Nocedal (2001), Rozloznik S. (2002), Durazzi Ruggiero (2003), Axelsson Neytcheva (2003), ...

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

Remark: $B^T x_k = 0$ for all iterates x_k

$$\mathcal{Q}^{-1} = \begin{bmatrix} I & -B^T \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -(\mathbf{BB^T} + \mathbf{C})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -B & I \end{bmatrix}$$

A “different” perspective

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

Polyak 1970, ..., Fischer & Ramage & Silvester & Wathen 1997, Bai & Golub & Ng 2003, Sidi 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004, ...

$\Lambda(\mathcal{M}_-)$ in $\mathbb{C}^+$

A “different” perspective

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

Polyak 1970, ..., Fischer & Ramage & Silvester & Wathen 1997, Bai & Golub & Ng 2003, Sidi 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004, ...

$\Lambda(\mathcal{M}_-)$ in $\mathbb{C}^+$

- More refined spectral information possible

A "different" perspective

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

Polyak 1970, ..., Fischer & Ramage & Silvester & Wathen 1997, Bai & Golub & Ng 2003, Sidi 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004, ...

$\Lambda(\mathcal{M}_-)$ in $\mathbb{C}^+$

- More refined spectral information possible
- New classes of preconditioners

A “different” perspective

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

Polyak 1970, ..., Fischer & Ramage & Silvester & Wathen 1997, Bai & Golub & Ng 2003, Sidi 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004, ...

$\Lambda(\mathcal{M}_-)$ in $\mathbb{C}^+$

- More refined spectral information possible
- New classes of preconditioners
- General framework for spectral analysis of some inexact preconditioners

Spectral properties of $\mathcal{M}_-$

A $n \times n$ sym. semidef. matrix, B $m \times n$, $m \leq n$

- $\mathcal{M}_-$ has at least $n - m$ real eigenvalues

Spectral properties of $\mathcal{M}_-$

A $n \times n$ sym. semidef. matrix, B $m \times n$, $m \leq n$

- $\mathcal{M}_-$ has at least $n - m$ real eigenvalues
- Let $C = \beta I$. If $\lambda_{\min}(A + \beta I) \geq 4\,\lambda_{\max}(B^T A^{-1} B + \beta I)$, then all eigenvalues of $\mathcal{M}_-$ are real.

Spectral properties of $\mathcal{M}_-$

A $n \times n$ sym. semidef. matrix, B $m \times n$, $m \leq n$

- $\mathcal{M}_-$ has at least $n - m$ real eigenvalues

- Let $C = \beta I$. If $\lambda_{\min}(A + \beta I) \geq 4\,\lambda_{\max}(B^T A^{-1} B + \beta I)$, then all eigenvalues of $\mathcal{M}_-$ are real.

- Let C be sym. Let $\lambda \in \Lambda(\mathcal{M}_-)$.
  If $\Im(\lambda) \neq 0$, then

$$
\begin{aligned}
\frac{1}{2}(\lambda_{\min}(A) + \lambda_{\min}(C)) \leq \quad \Re(\lambda) \quad &\leq \frac{1}{2}(\lambda_{\max}(A) + \lambda_{\max}(C)) \\
|\Im(\lambda)| \quad &\leq \sigma_{\max}(B).
\end{aligned}
$$

  If $\Im(\lambda) = 0$ then

$$2\min\{\lambda_{\min}(A), \lambda_{\min}(C)\} \leq \lambda \leq (\lambda_{\max}(A) + \lambda_{\max}(C)).$$

cf. Sidi 2003 for $C = 0$

Hermitian Skew-Hermitian Preconditioning

$$\mathcal{M}_- = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} + \begin{bmatrix} 0 & B^T \\ -B & 0 \end{bmatrix} = \mathcal{H} + \mathcal{S}$$

Hermitian Skew-Hermitian Preconditioning

$$\mathcal{M}_- = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} + \begin{bmatrix} 0 & B^T \\ -B & 0 \end{bmatrix} = \mathcal{H} + \mathcal{S}$$

Use the preconditioner

$$\mathcal{R} = \frac{1}{2\alpha}(\mathcal{H} + \alpha I)(\mathcal{S} + \alpha I) \qquad \alpha \in \mathbb{R},\ \alpha > 0$$

Bai & Golub & Ng 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004

Hermitian Skew-Hermitian Preconditioning

$$\mathcal{M}_- = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} + \begin{bmatrix} 0 & B^T \\ -B & 0 \end{bmatrix} = \mathcal{H} + \mathcal{S}$$

Use the preconditioner

$$\mathcal{R} = \frac{1}{2\alpha}(\mathcal{H} + \alpha I)(\mathcal{S} + \alpha I) \qquad \alpha \in \mathbb{R},\ \alpha > 0$$

Bai & Golub & Ng 2003, Benzi & Gander & Golub 2003, Benzi & Golub 2004, S. & Benzi 2004

Sharp spectral bounds for $C = 0$ (S. & Benzi 2004)

General framework

$$\mathcal{P}^{-1}\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ -\widetilde{B} & \widetilde{C} \end{bmatrix} \quad \widetilde{A} \geq 0,\ \widetilde{C} \geq 0$$

We can provide a spectral analysis when $\mathcal{P}$ is:

$$\mathcal{P}^{-1}\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ -\widetilde{B} & \widetilde{C} \end{bmatrix} \quad \widetilde{A} \geq 0,\ \widetilde{C} \geq 0$$

We can provide a spectral analysis when $\mathcal{P}$ is:

- Inexact Constraint Preconditioner

General framework

$$\mathcal{P}^{-1}\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ -\widetilde{B} & \widetilde{C} \end{bmatrix} \quad \widetilde{A} \geq 0,\ \widetilde{C} \geq 0$$

We can provide a spectral analysis when $\mathcal{P}$ is:

- Inexact Constraint Preconditioner
- Hermitian Skew-Hermitian Preconditioner ($C = \beta I$)

$$\mathcal{P}^{-1}\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \widetilde{A} & \widetilde{B}^T \\ -\widetilde{B} & \widetilde{C} \end{bmatrix} \quad \widetilde{A} \geq 0, \; \widetilde{C} \geq 0$$

We can provide a spectral analysis when $\mathcal{P}$ is:

- Inexact Constraint Preconditioner
- Hermitian Skew-Hermitian Preconditioner ($C = \beta I$)
- Indefinite Block diagonal Preconditioner

$$\begin{bmatrix} \widehat{A} & 0 \\ 0 & -\widehat{C} \end{bmatrix}$$

cf. Fischer et al. 1997

Final Considerations

- Performance of preconditioners is problem dependent

Final Considerations

- Performance of preconditioners is problem dependent
- Ad-Hoc preconditioners usually designed (information from application problem)

Final Considerations

- Performance of preconditioners is problem dependent
- Ad-Hoc preconditioners usually designed (information from application problem)
- Non-Hermitian problem not fully understood

Final Considerations

- Performance of preconditioners is problem dependent
- Ad-Hoc preconditioners usually designed (information from application problem)
- Non-Hermitian problem not fully understood
- Visit `http://www.dm.unibo.it/~simoncin`