Indefinite inner products in iterative linear system solvers

V. Simoncini

Dipartimento di Matematica

Università di Bologna

valeria@dm.unibo.it

http://www.dm.unibo.it/~simoncin

The problem

Solve the algebraic linear system

$$Ax = b$$

$A \in \mathbb{C}^{n\times n}$, large dimension ($n \gg 1000$)

Università di Bologna

The problem

Solve the algebraic linear system

$$Ax = b$$

$A \in \mathbb{C}^{n\times n}$, large dimension ($n \gg 1000$)

Approximation process: Given x_0 and $r_0 = b - Ax_0$, then

$$x_m \in x_0 + \mathcal{K}_m(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$$

The problem

Solve the algebraic linear system

$$Ax = b$$

$A \in \mathbb{C}^{n \times n}$, large dimension ($n \gg 1000$)

Approximation process: Given x_0 and $r_0 = b - Ax_0$, then

$$x_m \in x_0 + \mathcal{K}_m(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$$

Construct sequence of approximation spaces $\mathcal{K}_m \subset \mathcal{K}_{m+1}$ such that

$$\widetilde{x}_m \in \mathcal{K}_m \qquad \text{and} \quad \widetilde{x}_m \to x \qquad \text{as} \quad m \to \infty$$

(in some sense)

Projection process

($x_0 = 0 \Rightarrow r_0 = b$)

Let $\{v_1, \ldots, v_m\}$ be basis of $\mathcal{K}_m(A, b)$, $v_1 = b$. Then

$$x \approx x_m = V_m y_m \qquad V_m = [v_1, \ldots, v_m]$$

What choice for y_m?

Università di Bologna

Projection process

$(x_0 = 0 \Rightarrow r_0 = b)$

Let $\{v_1, \ldots, v_m\}$ be basis of $\mathcal{K}_m(A, b)$, $v_1 = b$. Then

$$x \approx x_m = V_m y_m \qquad V_m = [v_1, \ldots, v_m]$$

What choice for y_m?

Residual $r_m := b - AV_m y_m$ satisfies

$$r_m \perp_\star \mathcal{L}_m$$

Selection of $\mathcal{L}_m$ and of orthogonality constraint distinguish among several different methods

Università di Bologna

Typical (classical) approaches

⋆ A Hermitian positive definite. Galerkin condition:

$$x_m \in \mathcal{K}_m(A, b) \quad \text{and} \quad \mathcal{L}_m = \mathcal{K}_m(A, b)$$

Typical (classical) approaches

⋆ A Hermitian positive definite. Galerkin condition:

$$x_m \in \mathcal{K}_m(A, b) \quad \text{and} \quad \mathcal{L}_m = \mathcal{K}_m(A, b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^* r_m = 0$$

Conjugate Gradients (CG): sound implementation of Galerkin condition

Typical (classical) approaches

⋆ A Hermitian positive definite. Galerkin condition:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = \mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^* r_m = 0$$

Conjugate Gradients (CG): sound implementation of Galerkin condition

$$(V_m^* A V_m) y_m = V_m^* b \qquad V_m^* A V_m \quad \text{Hermitian positive definite}$$

Università di Bologna

Typical (classical) approaches

⋆ A Hermitian positive definite. Galerkin condition:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = \mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^* r_m = 0$$

Conjugate Gradients (CG): sound implementation of Galerkin condition

$$(V_m^* A V_m) y_m = V_m^* b \qquad V_m^* A V_m \quad \text{Hermitian positive definite}$$

If $V_m^* V_m = I$, then $\qquad T_m := V_m^* A V_m \qquad$ tridiagonal

Università di Bologna

Typical (classical) approaches

⋆ A Hermitian positive definite. Galerkin condition:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = \mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^* r_m = 0$$

Conjugate Gradients (CG): sound implementation of Galerkin condition

$$(V_m^* A V_m) y_m = V_m^* b \qquad V_m^* A V_m \quad \text{Hermitian positive definite}$$

If $V_m^* V_m = I$, then $\qquad T_m := V_m^* A V_m \qquad$ tridiagonal

Moreover,

$$\|x - x_m\|_A = \min_{\widetilde{x} \in \mathcal{K}_m} !$$

Typical (classical) approaches. II

⋆ A non-Hermitian:

$$x_m \in \mathcal{K}_m(A, b) \quad \text{and} \quad \mathcal{L}_m = A\mathcal{K}_m(A, b)$$

Typical (classical) approaches. II

⋆ A non-Hermitian:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = A\mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad (AV_m)^* r_m = 0$$

Typical (classical) approaches. II

⋆ A non-Hermitian:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = A\mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad (AV_m)^* r_m = 0$$

$$(V_m^* A^* A V_m) y_m = V_m^* A^* b \qquad V_m^* A^* A V_m$$ Hermitian positive definite

Università di Bologna

Typical (classical) approaches. II

⋆ A non-Hermitian:

$$x_m \in \mathcal{K}_m(A, b) \quad \text{and} \quad \mathcal{L}_m = A\mathcal{K}_m(A, b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad (AV_m)^* r_m = 0$$

$$(V_m^* A^* A V_m) y_m = V_m^* A^* b \qquad V_m^* A^* A V_m \quad \text{Hermitian positive definite}$$

$$\|b - AV_m y_m\|_2 = \min_{\widetilde{y} \in \mathbb{C}_m} !$$

Università di Bologna

Typical (classical) approaches. II

⋆ A non-Hermitian:

$$x_m \in \mathcal{K}_m(A,b) \quad \text{and} \quad \mathcal{L}_m = A\mathcal{K}_m(A,b)$$

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad (AV_m)^* r_m = 0$$

$(V_m^* A^* A V_m) y_m = V_m^* A^* b \qquad V_m^* A^* A V_m$ Hermitian positive definite

$$\|b - AV_m y_m\|_2 = \min_{\widetilde{y} \in \mathbb{C}_m} !$$

Arnoldi relation: $(V_m^* V_m = I)$

$$AV_m = V_m H_m + v_{m+1} h_{m+1} e_m^* \qquad \text{Range}(V_m) \subset \text{Range}(V_{m+1})$$

Other typical orthogonality properties

- $\star \quad e_m = x - x_m \;\perp \mathcal{K}_m(A^*, A^*b)$

Università di Bologna

Other typical orthogonality properties

- ⋆ $e_m = x - x_m \perp \mathcal{K}_m(A^*, A^*b)$
- ⋆ Given M Hermitian and positive definite,

$$r_m \perp_M \mathcal{L}_m$$

i.e., for Range(L_m)=$\mathcal{L}_m$ it holds $L_m^* M r_m = 0$

Università di Bologna

Other typical orthogonality properties

- ⋆ $e_m = x - x_m \perp \mathcal{K}_m(A^*, A^*b)$
- ⋆ Given M Hermitian and positive definite,

$$r_m \perp_M \mathcal{L}_m$$

i.e., for Range(L_m)=$\mathcal{L}_m$ it holds $L_m^* M r_m = 0$

- ⋆ Or, in particular,

$$V_m^* M V_m = I_m, \qquad M \text{ fixed}$$

minimization of $\|r_m\|_M$ or $\|e_m\|_M$

- e.g. $M = A^*A$
- stemming from application

Other typical orthogonality properties

⋆ $e_m = x - x_m \perp \mathcal{K}_m(A^*, A^*b)$

⋆ Given M Hermitian and positive definite,

$$r_m \perp_M \mathcal{L}_m$$

i.e., for Range(L_m)=$\mathcal{L}_m$ it holds $L_m^* M r_m = 0$

⋆ Or, in particular,

$$V_m^* M V_m = I_m, \qquad M \text{ fixed}$$

minimization of $\|r_m\|_M$ or $\|e_m\|_M$

- e.g. $M = A^*A$
- stemming from application

⋆ $M = M(m)$ varies with the subspace dimension

An example with M hpd

A positive real,

$A = S + K$, S Hermitian part of A, $S^{-1}A = I + S^{-1}K$

$$M = S$$

An example with M hpd

A positive real,

$A = S + K$, $\quad S$ Hermitian part of A, $\qquad S^{-1}A = I + S^{-1}K$

$$M = S$$

$$\Downarrow$$

- Generate $\mathcal{K}_m(S^{-1}A, S^{-1}b)$

- Create S-orthogonal V_m, i.e. $V_m^* S V_m = I$

An example with M hpd

A positive real,

$A = S + K$, S Hermitian part of A, $S^{-1}A = I + S^{-1}K$

$$M = S$$

$$\Downarrow$$

- Generate $\mathcal{K}_m(S^{-1}A, S^{-1}b)$
- Create S-orthogonal V_m, i.e. $V_m^* S V_m = I$

$S^{-1}AV_m = V_m H_m + v_{m+1} h_{m+1,m} e_m^*$ gives

$$\begin{aligned} V_m^* S S^{-1} A V_m &= V_m^* S V_m H_m + V_m^* S v_{m+1} h_{m+1,m} e_m^* \\ V_m^* A V_m &= H_m \quad \text{tridiagonal} \end{aligned}$$

Concus & Golub, Widlund,...

Università di Bologna

An example with M hpd

A positive real,

$A = S + K$, S Hermitian part of A, $S^{-1}A = I + S^{-1}K$

$$M = S$$

$$\Downarrow$$

- Generate $\mathcal{K}_m(S^{-1}A, S^{-1}b)$

- Create S-orthogonal V_m, i.e. $V_m^* S V_m = I$

$S^{-1}AV_m = V_m H_m + v_{m+1} h_{m+1,m} e_m^*$ gives

$$\begin{aligned} V_m^* S S^{-1} A V_m &= V_m^* S V_m H_m + V_m^* S v_{m+1} h_{m+1,m} e_m^* \\ V_m^* A V_m &= H_m \quad \text{tridiagonal} \end{aligned}$$

Concus & Golub, Widlund,...

This is not the whole story!

Motivations for an indefinite inner product

- Exploit inherent properties of the problem. For instance,

$$A \text{ complex symmetric}$$

Motivations for an indefinite inner product

- Exploit inherent properties of the problem. For instance,

  A complex symmetric

- Exploit matrix structure. E.g.,

$$A = \begin{pmatrix} H & B \\ B^* & 0 \end{pmatrix}$$

(but also A Hamiltonian, Symplectic, etc.)

Motivations for an indefinite inner product

- Exploit inherent properties of the problem. For instance,

  A complex symmetric

- Exploit matrix structure. E.g.,

$$A = \begin{pmatrix} H & B \\ B^* & 0 \end{pmatrix}$$

(but also A Hamiltonian, Symplectic, etc.)

... to gain in efficiency with (hopefully) no loss in reliability

Indefinite inner products and J-symmetry

⋆ Indefinite inner product $((x, y))$ does not satisfy

$$((x, x)) > 0 \quad \forall x$$

⋆ Given J Hermitian and nonsingular,

- A is J-Hermitian if $A^*J = JA$

Indefinite inner products and J-symmetry

$\star$ Indefinite inner product $((x,y))$ does not satisfy

$$((x,x)) > 0 \quad \forall x$$

$\star$ Given J Hermitian and nonsingular,

- A is J-Hermitian if $A^*J = JA$
- J-inner product:

$$\langle x,y\rangle_J = x^*Jy$$

($\langle x,x\rangle_J = 0$ for some x)

Indefinite inner products and J-symmetry

$\star$ Indefinite inner product $((x, y))$ does not satisfy

$$((x,x)) > 0 \quad \forall x$$

$\star$ Given J Hermitian and nonsingular,

- A is J-Hermitian if $A^* J = JA$
- J-inner product:

$$\langle x, y \rangle_J = x^* J y$$

  ($\langle x, x \rangle_J = 0$ for some x)
- Note:

  If $A^* J = JA$ and J hpd then A similar to Hermitian matrix

Complex Symmetric matrices

$$Ax = b$$

$A \in \mathbb{C}^{n \times n}$ complex symmetric, that is, $A = A^T$ (no conjugation)

Alternatives:

Complex Symmetric matrices

$$Ax = b$$

$A \in \mathbb{C}^{n\times n}$ complex symmetric, that is, $A = A^T$ (no conjugation)

Alternatives:

$\star$ A treated as complex non-Hermitian matrix

Complex Symmetric matrices

$$Ax = b$$

$A \in \mathbb{C}^{n \times n}$ complex symmetric, that is, $A = A^T$ (no conjugation)

Alternatives:

$\star$ A treated as complex non-Hermitian matrix

$\star$ Real formulation (twice the size) to exploit symmetry (maybe)

Complex Symmetric matrices

$$Ax = b$$

$A \in \mathbb{C}^{n\times n}$ complex symmetric, that is, $A = A^T$ (no conjugation)

Alternatives:

$\star$ A treated as complex non-Hermitian matrix

$\star$ Real formulation (twice the size) to exploit symmetry (maybe)

$\star$ "Natural" inner product for Krylov subspace methods:

$$((x, y)) := x^T y \qquad \text{no conjugation}$$

♣ $x \neq 0$ isotropic: $\qquad ((x, x)) = 0$

Complex Symmetric matrices

$$Ax = b$$

$A \in \mathbb{C}^{n\times n}$ complex symmetric, that is, $A = A^T$ (no conjugation)

Alternatives:

$\star$ A treated as complex non-Hermitian matrix

$\star$ Real formulation (twice the size) to exploit symmetry (maybe)

$\star$ "Natural" inner product for Krylov subspace methods:

$$((x, y)) := x^T y \qquad \text{no conjugation}$$

♣ $x \neq 0$ isotropic: $\qquad ((x, x)) = 0$

$\star$ Other structure-preserving approaches (e.g., Bunse-Gerstner & Stöver, '99)

Krylov subspace methods using $((\cdot,\cdot))$

Galerkin condition with x^*y replaced by condition with $((x,y)) = x^T y$

$\mathcal{L}_m = \mathcal{K}_m$ and

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^T r_m = 0$$

Krylov subspace methods using $((\cdot, \cdot))$

Galerkin condition with $x^* y$ replaced by condition with $((x, y)) = x^T y$

$\mathcal{L}_m = \mathcal{K}_m$ and

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^T r_m = 0$$

If $V_m^T V_m = I$ then $\qquad T_m := V_m^T A V_m$ tridiagonal

Krylov subspace methods using $((\cdot, \cdot))$

Galerkin condition with x^*y replaced by condition with $((x, y)) = x^T y$

$\mathcal{L}_m = \mathcal{K}_m$ and

$$r_m = b - AV_m y_m \perp \mathcal{L}_m \qquad \Leftrightarrow \qquad V_m^T r_m = 0$$

If $V_m^T V_m = I$ then $\qquad T_m := V_m^T A V_m$ tridiagonal

BUT

no minimization is carried out

An example with A complex symmetric

$A \in \mathbb{C}^{3627 \times 3627}$ $\quad A = K + iC_H$

Stiffness (real) + hysteretic damping matrix

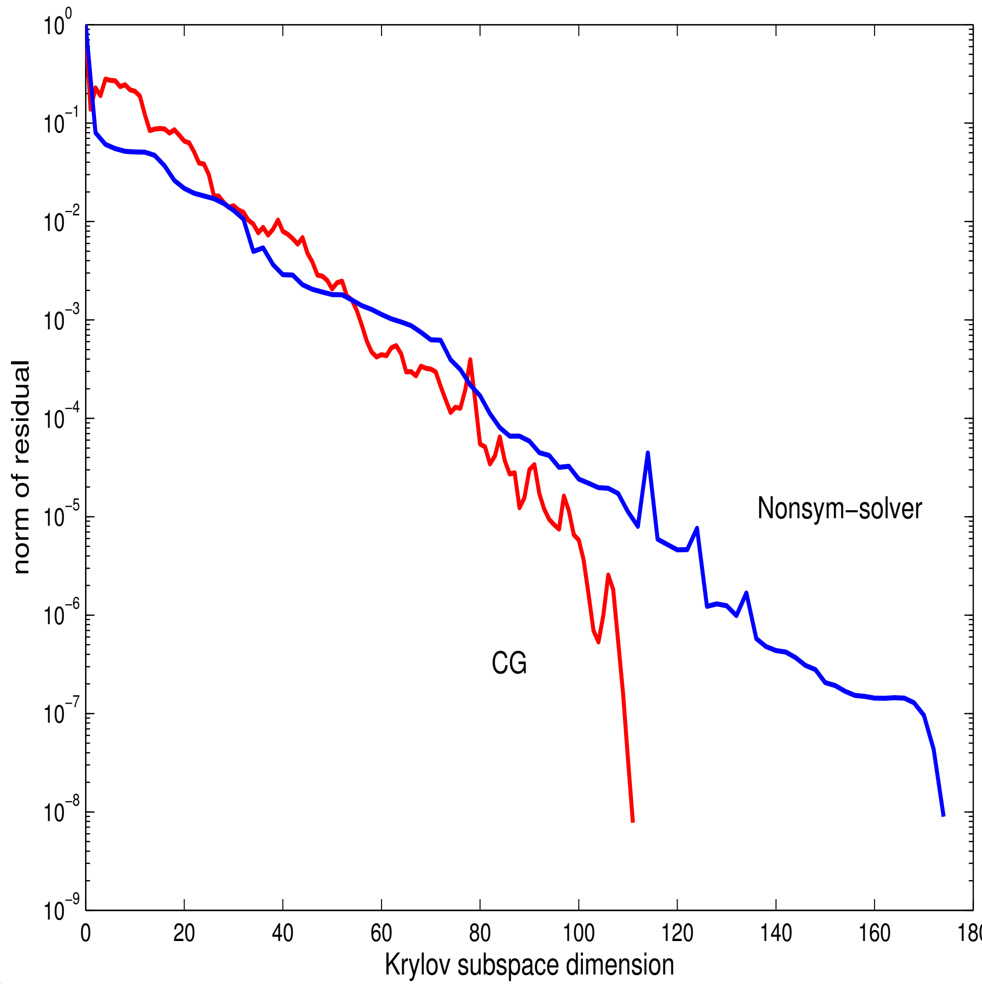
(Structural dynamic problem – ILU Preconditioner)

An example with A complex symmetric

$A \in \mathbb{C}^{3627\times 3627}$ $\quad A = K + iC_H$

Stiffness (real) + hysteretic damping matrix

(Structural dynamic problem – ILU Preconditioner)



A natural implementation: Two-sided Lanczos

Lanczos method for non-Hermitian matrices:

$$\mathcal{K}_m(A,b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*, \widehat{b}), \quad \widehat{b} \text{ auxiliary vector}$$

$$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m \quad \text{diagonal matrix}$$

$$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m \text{ tridiagonal}$$

Università di Bologna

A natural implementation: Two-sided Lanczos

Lanczos method for non-Hermitian matrices:

$$\mathcal{K}_m(A,b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*, \widehat{b}), \quad \widehat{b} \text{ auxiliary vector}$$

$$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m \quad \text{diagonal matrix}$$

$$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m \text{ tridiagonal}$$

In most implementations now:

$$\mathcal{L}_m = \mathcal{K}_m(A^T, b), \qquad s.t. \quad L_m^T V_m = D_m$$

A complex symmetric $\quad \Rightarrow \quad \mathcal{K}_m(A,b) = \mathcal{L}_m, \qquad L_m^T A V_m$ symmetric

Università di Bologna

A natural implementation: Two-sided Lanczos

Lanczos method for non-Hermitian matrices:

$$\mathcal{K}_m(A, b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*, \widehat{b}), \quad \widehat{b} \text{ auxiliary vector}$$

$$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m \quad \text{diagonal matrix}$$

$$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m \text{ tridiagonal}$$

In most implementations now:

$$\mathcal{L}_m = \mathcal{K}_m(A^T, b), \qquad s.t. \quad L_m^T V_m = D_m$$

A complex symmetric $\quad \Rightarrow \quad \mathcal{K}_m(A, b) = \mathcal{L}_m, \qquad L_m^T A V_m$ symmetric

♣ Two-sided Lanczos provides the setting for convergence analysis

Università di Bologna

Two-sided Lanczos and J-inner product

$\mathcal{K}_m(A, b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*, \widehat{b}), \quad \widehat{b}$ auxiliary vector

$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m$ diagonal matrix

$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m$ tridiagonal

Two-sided Lanczos and J-inner product

$$\mathcal{K}_m(A,b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*,\widehat{b}), \quad \widehat{b} \text{ auxiliary vector}$$

$$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m \quad \text{diagonal matrix}$$

$$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m \text{ tridiagonal}$$

Assume A is J-symmetric (or J-Hermitian). Then

$$\text{for} \qquad \widehat{b} := Jb$$

it holds

$$\mathcal{L}_m = J\mathcal{K}_m(A,b), \qquad L_m = J V_m \Sigma_m$$

(Σ_m diagonal matrix)

Two-sided Lanczos and J-inner product

$$\mathcal{K}_m(A,b), \qquad \mathcal{L}_m = \mathcal{K}_m(A^*, \widehat{b}), \quad \widehat{b} \text{ auxiliary vector}$$

$$V_m, \qquad L_m \quad s.t. \qquad L_m^* V_m = D_m \quad \text{diagonal matrix}$$

$$r_m \perp \mathcal{L}_m \quad \Rightarrow \quad (L_m^* A V_m) y_m = L_m^* b \quad L_m^* A V_m \text{ tridiagonal}$$

Assume A is J-symmetric (or J-Hermitian). Then

$$\text{for} \qquad \widehat{b} := Jb$$

it holds

$$\mathcal{L}_m = J\mathcal{K}_m(A,b), \qquad L_m = J V_m \Sigma_m$$

(Σ_m diagonal matrix)

$$\Downarrow$$

No need to generate space $\mathcal{L}_m$ and its basis! (Simplified Lanczos)

Freund & Nachtigal 1995

Disclaimers

- Still in the two-sided Lanczos framework
- Possible breakdown ($L_m^\star V_m$ singular, $\star = T, *$)
- Stability issues
- Specific convergence analysis: open problem

Università di Bologna

An application

$$(A + \sigma B)x = b$$

A, B symmetric, B nonsingular $\quad \sigma \in [\alpha, \beta] \subset \mathbb{R}$

Problem: Solve for several (a few hundreds, say) values of σ

An application

$$(A + \sigma B)x = b$$

A, B symmetric, B nonsingular $\quad \sigma \in [\alpha, \beta] \subset \mathbb{R}$

Problem: Solve for several (a few hundreds, say) values of σ

$$(AB^{-1} + \sigma I)\widehat{x} = b \qquad x = B^{-1}\widehat{x}$$

Shift-invariance of Krylov space: $\quad \mathcal{K}_m(AB^{-1} + \sigma I, b) = \mathcal{K}_m(AB^{-1}, b)$

Università di Bologna

An application

$$(A + \sigma B)x = b$$

A, B symmetric, B nonsingular $\quad \sigma \in [\alpha, \beta] \subset \mathbb{R}$

Problem: Solve for several (a few hundreds, say) values of σ

$$(AB^{-1} + \sigma I)\widehat{x} = b \qquad x = B^{-1}\widehat{x}$$

Shift-invariance of Krylov space: $\quad \mathcal{K}_m(AB^{-1} + \sigma I, b) = \mathcal{K}_m(AB^{-1}, b)$

$\star$ AB^{-1} is B^{-1}-symmetric $\qquad$ (that is, $(AB^{-1})^T B^{-1} = B^{-1}(AB^{-1})$)

$$\Downarrow$$

Simplified Lanczos method with $J = B^{-1}$

(Perotti & Simoncini 2002)

Application to Preconditioning

Saddle-point problem:

$$Ax = b \quad \Leftrightarrow \quad \begin{pmatrix} H & B \\ B^* & -C \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

H, C Hermitian pos. (semi-)definite $\Rightarrow$ A Hermitian indef. (nonsing.)

Recent survey: Benzi & Golub & Liesen, 2005

Università di Bologna

Application to Preconditioning

Saddle-point problem:

$$Ax = b \qquad \Leftrightarrow \qquad \begin{pmatrix} H & B \\ B^* & -C \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

H, C Hermitian pos. (semi-)definite $\Rightarrow$ A Hermitian indef. (nonsing.)

Recent survey: Benzi & Golub & Liesen, 2005

⋆ Preconditioning technique: Find nonsingular P s.t.

$$AP^{-1}\widehat{x} = b$$

"easier" to solve, with P cheap to invert (or, $P_1^{-1}AP_2^{-1}\widehat{x} = P_1^{-1}b$)

Various successful choices, mostly problem dependent

For simplicity: $C = 0$

Structured Preconditioners

$$A = \begin{pmatrix} H & B \\ B^* & 0 \end{pmatrix}$$

- Block diagonal (AP^{-1} is P^{-1}-Hermitian – minimization in the "correct" norm)

$$P = \begin{pmatrix} \widetilde{H} & 0 \\ 0^* & S \end{pmatrix} \qquad \widetilde{H} \approx H, \; S \approx B^* H^{-1} B$$

Università di Bologna

Structured Preconditioners

$$A = \begin{pmatrix} H & B \\ B^* & 0 \end{pmatrix}$$

- Block diagonal (AP^{-1} is P^{-1}-Hermitian – minimization in the “correct” norm)

$$P = \begin{pmatrix} \widetilde{H} & 0 \\ 0^* & S \end{pmatrix} \qquad \widetilde{H} \approx H, \; S \approx B^* H^{-1} B$$

- Block indefinite (more in the next few slides)

$$P = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix} \qquad \widetilde{H} \approx H$$

Structured Preconditioners

$$A = \begin{pmatrix} H & B \\ B^* & 0 \end{pmatrix}$$

- Block diagonal (AP^{-1} is P^{-1}-Hermitian – minimization in the "correct" norm)

$$P = \begin{pmatrix} \widetilde{H} & 0 \\ 0^* & S \end{pmatrix} \qquad \widetilde{H} \approx H, \; S \approx B^* H^{-1} B$$

- Block indefinite (more in the next few slides)

$$P = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix} \qquad \widetilde{H} \approx H$$

- Block triangular (AP^{-1} is similar to Hermitian, under conditions on $\widetilde{H}$, S)

$$P = \begin{pmatrix} \widetilde{H} & B \\ 0^* & -S \end{pmatrix} \qquad S \approx B^* H^{-1} B$$

Indefinite (Constraint) Preconditioners

$$P = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix} \qquad\qquad AP^{-1}\widehat{x} = b$$

$\widetilde{H} \approx H$, cheap to solve with

More Indefinite-style preconditioners: Dollar, Gould, Wathen, Schilders, 2005

Indefinite (Constraint) Preconditioners

$$P = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix} \qquad\qquad AP^{-1}\widehat{x} = b$$

$\widetilde{H} \approx H$, cheap to solve with

More Indefinite-style preconditioners: Dollar, Gould, Wathen, Schilders, 2005

* AP^{-1} not symmetrizable!
* However: AP^{-1} is P^{-1}-Hermitian

Use Simplified Lanczos method

Indefinite (Constraint) Preconditioners

$$P = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix} \qquad AP^{-1}\widehat{x} = b$$

$\widetilde{H} \approx H$, cheap to solve with

More Indefinite-style preconditioners: Dollar, Gould, Wathen, Schilders, 2005

* AP^{-1} not symmetrizable!
* However: AP^{-1} is P^{-1}-Hermitian

Use Simplified Lanczos method

What spectral properties?

Spectral properties. I

⋆ Eigenvalues of AP^{-1} are real and positive, in $\lambda \in \{1\} \cup [\alpha_0, \alpha_1]$

⋆ Presence of Jordan blocks

Spectral properties. I

⋆ Eigenvalues of AP^{-1} are real and positive, in $\lambda \in \{1\} \cup [\alpha_0, \alpha_1]$

⋆ Presence of Jordan blocks

However

Jordan blocks do not influence convergence

(with appropriate starting approximate solution)

$$AP^{-1} = \begin{pmatrix} H(I-\Pi)+\Pi & (H-I)B(B^*B)^{-1} \\ 0 & I \end{pmatrix} \qquad r_0 = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$\Pi = B(B^*B)^{-1}B^*$ Projector

Spectral properties. I

⋆ Eigenvalues of AP^{-1} are real and positive, in $\lambda \in \{1\} \cup [\alpha_0, \alpha_1]$

⋆ Presence of Jordan blocks

However

Jordan blocks do not influence convergence

(with appropriate starting approximate solution)

$$AP^{-1} = \begin{pmatrix} H(I - \Pi) + \Pi & (H - I)B(B^*B)^{-1} \\ 0 & I \end{pmatrix} \qquad r_0 = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

$\Pi = B(B^*B)^{-1}B^*$ Projector

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), Dollar, Gould, Wathen, Schilders (2005),...

Computational Considerations: Exact P vs Inexact P

$$P^{-1} = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix}^{-1} = \begin{pmatrix} I & -B^T \\ O & I \end{pmatrix} \begin{pmatrix} I & O \\ O & -(\mathbf{BB^*})^{-1} \end{pmatrix} \begin{pmatrix} I & O \\ -B & I \end{pmatrix}$$

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

Computational Considerations: Exact P vs Inexact P

$$P^{-1} = \begin{pmatrix} \widetilde{H} & B \\ B^* & 0 \end{pmatrix}^{-1} = \begin{pmatrix} I & -B^T \\ O & I \end{pmatrix} \begin{pmatrix} I & O \\ O & -(\mathbf{BB}^*)^{-1} \end{pmatrix} \begin{pmatrix} I & O \\ -B & I \end{pmatrix}$$

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

3D Magnetostatic problem. Elapsed Time

	SIMPLIFIED LANCZOS		SIMPLIFIED LANCZOS
SIZE	P	$\widehat{P}(2)(it)$	ILDLT(10)
1119	**3.0**(15)	**1.7**(18)	**0.7**
2208	**11.7** (13)	**3.1**(18)	**1.5**
4371	**64.6** (17)	**8.4**(20)	**5.2**
8622	**466.0**(16)	**18.3**(29)	**31.0**
22675	**3745.5**(25)	**63.2**(45)	**246.0**

$BB^* \approx S$ Incomplete Cholesky fact. $\Rightarrow$ $\widehat{P}$ $\quad$ $A\widehat{P}^{-1}\widehat{x} = b$

Università di Bologna

Spectral properties. II

- $A\widehat{P}^{-1}$ still $\widehat{P}^{-1}$-symmetric

Spectral properties. II

- $A\widehat{P}^{-1}$ still $\widehat{P}^{-1}$-symmetric
- Eigenvalue bounds: Let $\widehat{C} = B^*(2I - H)BS^{-1}$, $S \approx B^*B$
  - $\star$ If $\Im(\lambda) \neq 0$ then

$$(\lambda_{\min}(H) + \lambda_{\min}(\widehat{C})) \leq \Re(\lambda) \leq \frac{1}{2}(\lambda_{\max}(H) + \lambda_{\max}(\widehat{C}))$$

$$|\Im(\lambda)| \leq \sigma_{\max}((I - H)BS^{-\frac{1}{2}}).$$

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

$$\min\{\lambda_{\min}(H), \lambda_{\min}(\widehat{C})\} \leq \lambda \leq \max\{\lambda_{\max}(H), \lambda_{\max}(\widehat{C})\}$$

Spectral properties. II

- $A\widehat{P}^{-1}$ still $\widehat{P}^{-1}$-symmetric
- Eigenvalue bounds: Let $\widehat{C} = B^*(2I - H)BS^{-1}$, $S \approx B^*B$

  ⋆ If $\Im(\lambda) \neq 0$ then

$$(\lambda_{\min}(H) + \lambda_{\min}(\widehat{C})) \leq \Re(\lambda) \leq \frac{1}{2}(\lambda_{\max}(H) + \lambda_{\max}(\widehat{C}))$$

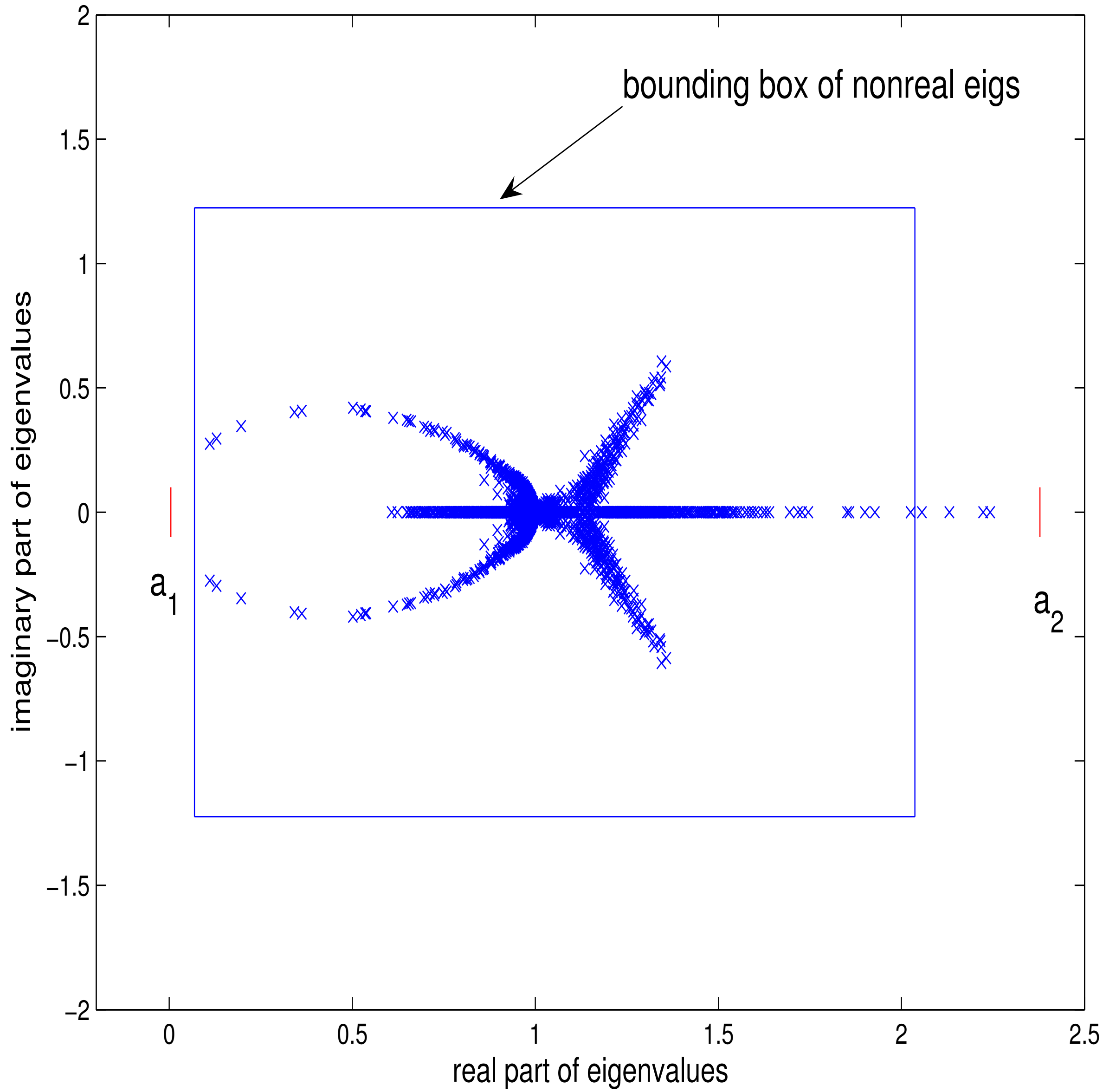
$$|\Im(\lambda)| \leq \sigma_{\max}((I - H)BS^{-\frac{1}{2}}).$$

  ⋆ If $\Im(\lambda) = 0$ then

$$\min\{\lambda_{\min}(H), \lambda_{\min}(\widehat{C})\} \leq \lambda \leq \max\{\lambda_{\max}(H), \lambda_{\max}(\widehat{C})\}$$

- Eigenvectors: open problem

Spectral bounds



Benzi & Simoncini, 2006

Final Considerations

- Indefinite inner product appropriate
  to exploit inherent problem properties
- Many computational issues still open
- Convergence analysis still very challenging

Final Considerations

- Indefinite inner product appropriate
  to exploit inherent problem properties
- Many computational issues still open
- Convergence analysis still very challenging

valeria@dm.unibo.it
http://www.dm.unibo.it/~simoncin

“ Recent computational developments in Krylov Subspace Methods for linear systems”

with Daniel Szyld, (Temple University)
To appear J. Numerical Linear Algebra w/Appl.