Iterative methods for solving large scale matrix equation problems

Valeria Simoncini

Dipartimento di Matematica

Alma Mater Studiorum, Università di Bologna

valeria.simoncini@unibo.it

Schedule of the course module

- Iterative methods for large scale linear systems
  (Today, Dec 15, 15-18)
- Stopping criteria and other effective methods + Lab (?)
  (Tomorrow, Thu, Dec 16, 11-13)
- Preconditioning
  (Tomorrow, Thu, Dec 16, 15-18)
- Computational experience
  (Fri, Dec 17, 9-11)

Lectures: see `https://www.dm.unibo.it/~simoncin/corso.html`

Iterative methods for large scale linear systems

Outline

- Projection and polynomial -type methods
- **Coefficient matrix role in tailoring the solution strategy**
  - Real symmetric or complex Hermitian
  - Complex symmetric and H-symmetric
  - Complex/Real non-Hermitian
- Stopping criteria and inexactness

The Problem

$$Ax = b \qquad \text{or} \quad AX = B, \quad B = [b_1, \ldots, b_s]$$

$A \in \mathbb{C}^{n \times n}$, B full column rank, $s \ll n$

- A large and sparse
- A large and structured: blocks, banded, ...
- A functional: $A = CS^{-1}D$, preconditioned, integral, ...
- ....

The Problem

$$Ax = b \qquad \text{or} \quad AX = B, \quad B = [b_1, \ldots, b_s]$$

$A \in \mathbb{C}^{n\times n}$, B full column rank, $s \ll n$

- A large and sparse
- A large and structured: blocks, banded, ...
- A functional: $A = CS^{-1}D$, preconditioned, integral, ...
- ....

The solution approach. Generate sequence of approximate solutions:

$$\{x_0, x_1, x_2, \ldots\}, \qquad x_k \to_{k\to\infty} x$$

Occurrence of the problem

Very broad range of applications in Engineering and Scientific Computing

Original application context:

- Discretization of 2D and 3D PDEs
  (linear steady state, nonlinear, evolutive, etc.)
- Eigenvalue problems
- Approximation of matrix functions
- Workhorses of more advanced techniques
- ...

Relevant Bibliographic Pointers

YOUSEF SAAD

Iterative methods for sparse linear systems

SIAM, Society for Industrial and Applied Mathematics, 2003, 2nd edition.

VALERIA SIMONCINI AND DANIEL B. SZYLD

Recent developments in Krylov Subspace Methods for linear systems

Numerical Linear Algebra with Appl., v. 14, n.1 (2007), pp.1-59.

“Projection” methods (or, reduction methods)

- Approximation vector space K_m. At each iteration m

$$\{x_m\} \text{ such that } x_m \in K_m$$

K_m: dimension[a] m, with the “expansion” property:

$$K_m \subseteq K_{m+1}$$

- Computation of iterate. Galerkin condition:

$$\text{residual} \quad r_m := b - Ax_m \quad \perp \quad K_m$$

$\Rightarrow$ This condition uniquely defines $x_m \in K_m$

[a] At most

Optimality property of Galerkin projection method

A symmetric and positive definite. Let $x^\star$ be the true solution.

Galerkin property: Impose that

$$\text{residual} \quad r_m := b - Ax_m \quad \perp \quad K_m$$

is equivalent to: Find

$$x_m \quad \text{solution to} \quad \min_{x \in K_m} \|x^\star - x\|_A$$

where $\|\cdot\|_A$ is the energy norm, or A-norm, namely $\|v\|_A^2 := v^T A v$

Convergence and spectral properties

- In exact arithmetic (i.e., in theory), finite termination property
- A-priori bound for energy norm of the error:
  If $K_m = \text{span}\{b, Ab, \ldots, A^{m-1}b\}$, then

$$\|x^\star - x_m\|_A \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^m \|x^\star - x_0\|_A$$

  where $\kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ is the condition number of A

(Conjugate Gradients, Hestenes & Stiefel, '52)

Convergence and spectral properties

- In exact arithmetic (i.e., in theory), finite termination property
- A-priori bound for energy norm of the error:
  If $K_m = \mathrm{span}\{b, Ab, \ldots, A^{m-1}b\}$, then

$$\|x^\star - x_m\|_A \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^m \|x^\star - x_0\|_A$$

  where $\kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ is the condition number of A

(Conjugate Gradients, Hestenes & Stiefel, '52)

Consequences:

- Convergence: The closer κ to 1 the faster
- Convergence depends on spectral properties, not directly on problem size!

A well established algorithm

Classical Conjugate Gradient method:

Given x_0. Set $r_0 = b - Ax_0$, $p_0 = r_0$

for $i = 0, 1, \ldots$

$$\alpha_i = \frac{r_i^* r_i}{p_i^* A p_i}$$

$$x_{i+1} = x_i + p_i \alpha_i$$

$$r_{i+1} = r_i - A p_i \alpha_i$$

$$\beta_{i+1} = \frac{r_{i+1}^* A p_i}{p_i^* A p_i}$$

$$p_{i+1} = r_i + p_i \beta_{i+1}$$

end

* At each iteration: 1 `Mxv`, 3 `-axpys`, 2 `-dots`

* Short-term recurrence

* Implicit space generation, no explicit computation of the orthonormal basis!

The Conjugate Gradient method. Geometric properties

$$\boxed{x_{i+1} = x_i + p_i\alpha_i} \qquad \boxed{r_{i+1} = r_i - Ap_i\alpha_i} \qquad \boxed{p_{i+1} = r_i + p_i\beta_{i+1}}$$

$$K_k = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

For simplicity, assume $x_0 = 0$.

⋆ Using $p_0 = r_0 = b$, we have

$$r_0 = b, \quad r_1 \in \text{span}\{r_0, Ar_0\}, \quad r_2 \in \text{span}\{r_0, Ar_0, A^2r_0\}, \ldots,$$

$\Rightarrow r_k \in K_{k+1}(A, r_0)$

$\Rightarrow x_{k+1} \in K_{k+1}(A, r_0), \quad x_{k+1} = p_0\alpha_0 + p_1\alpha_1 + \ldots + p_k\alpha_k$

The Conjugate Gradient method. Geometric properties

$\boxed{x_{i+1} = x_i + p_i \alpha_i}$ $\boxed{r_{i+1} = r_i - A p_i \alpha_i}$ $\boxed{p_{i+1} = r_i + p_i \beta_{i+1}}$

$K_k = \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1} r_0\}$

For simplicity, assume $x_0 = 0$.

$\star$ Using $p_0 = r_0 = b$, we have

$$r_0 = b, \quad r_1 \in \mathrm{span}\{r_0, Ar_0\}, \quad r_2 \in \mathrm{span}\{r_0, Ar_0, A^2 r_0\}, \ldots,$$

$\Rightarrow r_k \in K_k(A, r_0)$

$\Rightarrow x_{k+1} \in K_{k+1}(A, r_0), \quad x_{k+1} = p_0\alpha_0 + p_1\alpha_1 + \ldots + p_k\alpha_k$

$\star$ It holds

$$r_i^T r_j = 0, \quad p_i^T A p_j = 0, \quad \text{for all } i \neq j$$

The **Block** Conjugate Gradient

$$AX = B, \qquad B = [b_1, b_2, \ldots, b_s]$$

$R_0 = B - AX_0$, $P_0 = R_0 \in \mathbb{C}^{n \times s}$

for $k = 0, 1, \ldots$

$\quad \boldsymbol{\alpha}_k = (P_k^* A P_k)^{-1} (R_k^* R_k) \in \mathbb{C}^{s \times s}$

$\quad X_{k+1} = X_k + P_k \boldsymbol{\alpha}_k$

$\quad R_{k+1} = R_k - A P_k \boldsymbol{\alpha}_k$

$\quad \boldsymbol{\beta}_{k+1} = (P_k^* A P_k)^{-1} (R_{k+1}^* A P_k) \in \mathbb{C}^{s \times s}$

$\quad P_{k+1} = R_k + P_k \boldsymbol{\beta}_{k+1}$

end

A more general picture. Nonsymmetric problems

- A normal, $AA^* = A^*A$
- A (highly) non-normal, $\|AA^* - A^*A\| \gg 0$
- A “Hermitian” in disguise:

A more general picture. Nonsymmetric problems

- A normal, $AA^* = A^*A$
- A (highly) non-normal, $\|AA^* - A^*A\| \gg 0$
- A "Hermitian" in disguise:
  - $\star$ $A = M + \sigma I,\ \sigma \in \mathbb{C}$, $M \in \mathbb{R}^{n\times n}$ symmetric

A more general picture. Nonsymmetric problems

- A normal, $AA^* = A^*A$
- A (highly) non-normal, $\|AA^* - A^*A\| \gg 0$
- A "Hermitian" in disguise:
  - $\star$ $A = M + \sigma I,\ \sigma \in \mathbb{C}$, $M \in \mathbb{R}^{n\times n}$ symmetric
  - $\star$ $A = M + \sigma H,\ \sigma \in \mathbb{C}$, $M, H \in \mathbb{R}^{n\times n}$ symmetric

A more general picture. Nonsymmetric problems

- A normal, $AA^* = A^*A$
- A (highly) non-normal, $\|AA^* - A^*A\| \gg 0$
- A "Hermitian" in disguise:
  - ⋆ $A = M + \sigma I,\ \sigma \in \mathbb{C},\ M \in \mathbb{R}^{n\times n}$ symmetric
  - ⋆ $A = M + \sigma H,\ \sigma \in \mathbb{C},\ M, H \in \mathbb{R}^{n\times n}$ symmetric
  - ⋆ There exists nonsing. Herm. $H \in \mathbb{C}^{n\times n}$ such that $HA = A^*H$, e.g. M, C Hermitian

$$A = \begin{bmatrix} M & B \\ -B^* & C \end{bmatrix}, \quad H = \begin{bmatrix} I & \\ & -I \end{bmatrix},$$

A more general picture. Nonsymmetric problems

- A normal, $AA^* = A^*A$
- A (highly) non-normal, $\|AA^* - A^*A\| \gg 0$
- A "Hermitian" in disguise:
  - ⋆ $A = M + \sigma I, \; \sigma \in \mathbb{C}, \, M \in \mathbb{R}^{n\times n}$ symmetric
  - ⋆ $A = M + \sigma H, \; \sigma \in \mathbb{C}, \, M, H \in \mathbb{R}^{n\times n}$ symmetric
  - ⋆ There exists nonsing. Herm. $H \in \mathbb{C}^{n\times n}$ such that $HA = A^*H$, e.g. M, C Hermitian

$$A = \begin{bmatrix} M & B \\ -B^* & C \end{bmatrix}, \quad H = \begin{bmatrix} I & \\ & -I \end{bmatrix},$$

  - ⋆ $Ax = b \quad \Leftrightarrow \quad A^*Ax = A^*b$ (not recommended in general...)

Outline

- What is the added difficulty with A non-Hermitian ?
- How to handle "Symmetry in disguise"
- Non-normal (non-Hermitian) case
  - ⋆ Long-term recurrences and their problems
  - ⋆ Coping with them ⇒ Restarted, truncated, flexible
  - ⋆ Making it without ⇒ short-term recurrences
- Tricks for all trades

What goes "wrong" with A non-Hermitian. I

$$\{x_k\}, \quad \text{with} \quad x_k \in x_0 + K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

Let $V_k = [v_1, \ldots, v_k]$ be a (orthogonal) basis of $K_k(A, r_0)$. Then

$$x_k = x_0 + V_k y_k, \qquad y_k \in \mathbb{C}^k$$

A condition is required to specify y_k.

What goes "wrong" with A non-Hermitian. I

$$\{x_k\}, \quad \text{with} \quad x_k \in x_0 + K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

Let $V_k = [v_1, \ldots, v_k]$ be a (orthogonal) basis of $K_k(A, r_0)$. Then

$$x_k = x_0 + V_k y_k, \qquad y_k \in \mathbb{C}^k$$

A condition is required to specify y_k. For instance:

$$r_k := b - Ax_k = r_0 - AV_k y_k \quad \perp \quad K_k(A, r_0) \qquad V_k^* r_k = 0$$

(Galerkin condition, again!) so that

$$0 = V_k^* r_k = V_k^* r_0 - V_k^* A V_k y_k \quad \Leftrightarrow \quad y_k \ s.t. \ (V_k^* A V_k) y_k = V_k^* r_0$$

What goes "wrong" with A non-Hermitian. I

$$\{x_k\}, \quad \text{with} \quad x_k \in x_0 + K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

Let $V_k = [v_1, \ldots, v_k]$ be a (orthogonal) basis of $K_k(A, r_0)$. Then

$$x_k = x_0 + V_k y_k, \qquad y_k \in \mathbb{C}^k$$

A condition is required to specify y_k. For instance:

$$r_k := b - Ax_k = r_0 - AV_k y_k \quad \perp \quad K_k(A, r_0) \qquad V_k^* r_k = 0$$

(Galerkin condition, again!) so that

$$0 = V_k^* r_k = V_k^* r_0 - V_k^* A V_k y_k \quad \Leftrightarrow \quad y_k \ s.t. \ (V_k^* A V_k) y_k = V_k^* r_0$$

Hence

$$x_k = x_0 + V_k (V_k^* A V_k)^{-1} V_k^* r_0 \qquad \text{with} \quad V_k^* r_0 = e_1 \|r_0\|$$

And: $V_k^* A V_k$ upper Hessenberg (Gram-Schmidt procedure to build V_k)

What goes "wrong" with A non-Hermitian. II

If A were Hpd $\quad \Rightarrow \quad V_k^* A V_k$ also Hpd $\quad \Rightarrow \quad$ tridiagonal

$$V_k^* A V_k = L_k L_k^* \qquad L_k \text{ bidiagonal}$$

$$\begin{aligned} x_k &= x_0 + V_k L_k^{-*} L_k^{-1} e_1 \|r_0\| \\ &= x_0 + V_{k-1} L_{k-1}^{-*} L_{k-1}^{-1} e_1 \|r_0\| + p_k \alpha_k \\ &= x_{k-1} + p_k \alpha_k \end{aligned}$$

with $p_k \in \text{span}\{v_{k-1}, v_k\}$

(development underlying Conjugate Gradients)

What goes “wrong” with A non-Hermitian. II

If A were Hpd $\quad \Rightarrow \quad V_k^* A V_k$ also Hpd $\quad \Rightarrow \quad$ tridiagonal

$$V_k^* A V_k = L_k L_k^* \qquad L_k \text{ bidiagonal}$$

$$\begin{aligned} x_k &= x_0 + V_k L_k^{-*} L_k^{-1} e_1 \|r_0\| \\ &= x_0 + V_{k-1} L_{k-1}^{-*} L_{k-1}^{-1} e_1 \|r_0\| + p_k \alpha_k \\ &= x_{k-1} + p_k \alpha_k \end{aligned}$$

with $p_k \in \mathrm{span}\{v_{k-1}, v_k\}$

(development underlying Conjugate Gradient)

A non-Hermitian $\Rightarrow V_k^* A V_k$ only upper Hessenberg

$$p_k \in \mathrm{span}\{v_1, \ldots, v_k\}$$

What goes “wrong” with A non-Hermitian. III

$p_k \in \text{span}\{v_1, \dots, v_k\}$, with $\{v_1, \dots, v_k\}$ orthogonal basis

Alternatives

- Give up orthogonal basis, $V_k^* V_k = I_k$
- Give up optimality condition, e.g. $r_k \perp K_k(A, r_0)$
- Resume symmetry

Symmetry in disguise. Complex symmetric shifted systems. 1.

Case 1: $A = M + \sigma I, \quad M \in \mathbb{R}^{n\times n}, \sigma \in \mathbb{C}$

E.g.: Helmholtz equation (wave problems such as vibrating strings and membranes)

Trick: replace $*$ (conj. transp.) with $\top$ (transp.)

$$A = A^\top \qquad \text{complex symmetric}$$

Apply CG with $\top$

Given x_0. Set $r_0 = b - Ax_0$, $p_0 = r_0$

for $i = 0, 1, \dots$

$$\alpha_i = \frac{r_i^\top r_i}{p_i^\top A p_i}$$

$$x_{i+1} = x_i + p_i \alpha_i$$

$$r_{i+1} = r_i - A p_i \alpha_i$$

$$\beta_{i+1} = \frac{r_{i+1}^\top A p_i}{p_i^\top A p_i}$$

$$p_{i+1} = r_i + p_i \beta_{i+1}$$

end

Symmetry in disguise. Complex symmetric shifted systems. 2.

$A = M + \sigma I$: Apply CG with $\top$

Properties:

- V_k real: $K_k(A, r_0) = K_k(A + \sigma I, r_0)$
- $\top$ does not define an inner product!
- $V_k^\top A V_k = V_k^\top M V_k + \sigma I$

  If $\Im(\sigma) \neq 0$ then $V_k^\top A V_k$ is nonsingular $\Rightarrow$ No breakdown

The same code applies in case of any A complex symmetric ($A = A^\top$)

H-symmetry

A is H-Hermitian if there exists $H \in \mathbb{C}^{n \times n}$ Hermitian, nonsingular s.t.

$$HA = A^* H$$

(H-symmetric if $HA = A^\top H$ with H is symmetric)

H-symmetry

A is H-Hermitian if there exists $H \in \mathbb{C}^{n\times n}$ Hermitian, nonsingular s.t.

$$HA = A^* H$$

(H-symmetric if $HA = A^\top H$ with H is symmetric)

If H is Hpd (and HA is also Hpd), use CG in the H-inner product:

Given x_0. Set $r_0 = b - Ax_0$, $p_0 = r_0$

for $i = 0, 1, \ldots$

$$\alpha_i = \frac{r_i^* H r_i}{p_i^* H A p_i}$$

$$x_{i+1} = x_i + p_i \alpha_i$$

$$r_{i+1} = r_i - A p_i \alpha_i$$

$$\beta_{i+1} = \frac{r_{i+1}^* H A p_i}{p_i^* H A p_i}$$

$$p_{i+1} = r_i + p_i \beta_{i+1}$$

end

(H not Hpd $\Rightarrow$ see later)

First Summary

Symmetry in disguise:

- Shifted matrices, $A = M + \sigma I$, M real symmetric
- Complex symmetric matrices
- H-symmetric or H-Hermitian matrices

Long-term recurrences

$$K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}, \quad V_k \text{ orth. basis}$$

1. Arnoldi process : $v_{k+1} \leftarrow Av_k - \sum_{j=1}^{k} v_j h_{j,k}$, that is

$$AV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k^T = V_{k+1} \underline{H}_k \qquad (H_k = V_k^* A V_k)$$

2. $x_k = x_0 + V_k y_k$

Long-term recurrences

$$K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}, \quad V_k \text{ orth. basis}$$

1. Arnoldi process : $v_{k+1} \leftarrow Av_k - \sum_{j=1}^{k} v_j h_{j,k}$, that is

$$AV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k^T = V_{k+1} \underline{H}_k \qquad (H_k = V_k^* A V_k)$$

2. $x_k = x_0 + V_k y_k$

- GMRES. Particular Petrov-Galerkin condition:

$$r_k \perp AK_k \Rightarrow \quad y_k \quad \text{s.t.} \qquad \min_y \|r_0 - AV_k y\|$$

- FOM. Galerkin condition: (H_k nonsingular)

$$r_k \perp K_k \Rightarrow \quad y_k \quad \text{s.t.} \qquad H_k y = e_1 \|r_0\|$$

GMRES

$AV_k = V_{k+1}\underline{H}_k, \qquad r_0 = V_{k+1}e_1\beta_0$

Crucial property:

$$\begin{aligned} \min_y \|r_0 - AV_k y\| &= \min_y \|V_{k+1}(e_1\beta_0 - \underline{H}_k y)\| \\ &= \min_y \|e_1\beta_0 - \underline{H}_k y\| \end{aligned}$$

Least squares problem expands at each iteration.

QR decomposition of $\underline{H}_k$ only updated, not recomputed from scratch.

Block GMRES

$$R_0 = B - AX_0, \qquad K_k(A, R_0) = \text{span}\{R_0, AR_0, \ldots, A^{k-1}R_0\},$$

$\mathcal{U}_k$ orth. basis, $\mathcal{U}_k = [U_1, U_2, \ldots, U_k] \in \mathbb{C}^{n\times ks}$

Block Arnoldi process (s MxV + Gram-Schmidt)

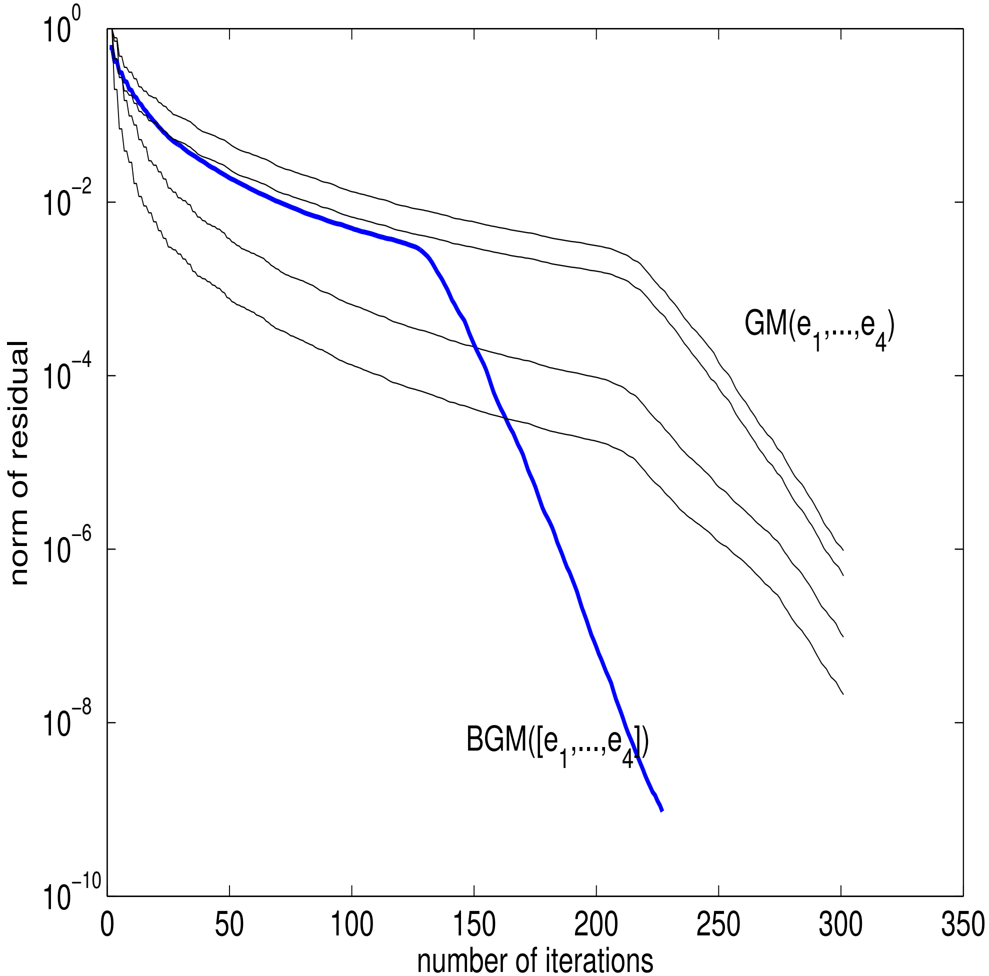
$$\Rightarrow \; A\mathcal{U}_k = \mathcal{U}_k \mathcal{H}_k + U_{k+1}\chi_{k+1,k}E_k^* = \mathcal{U}_{k+1}\underline{\mathcal{H}}_k \qquad (\mathcal{H}_k = \mathcal{U}_k^* A \mathcal{U}_k)$$

$$\min_Y \|R_0 - A\mathcal{U}_k Y\| = \min_Y \|\, E_1\boldsymbol{\rho} - \underline{\mathcal{H}}_k Y \,\| \quad R_0 = U_1\boldsymbol{\rho}$$

$$\underline{\mathcal{H}}_k = \begin{bmatrix} \square & \square & \cdots & \square \\ \square & \square & \cdots & \square \\ O & \square & \cdots & \square \\ O & O & \ddots & \square \\ O & O & O & \square \end{bmatrix}$$

Block GMRES

$A \in \mathbb{R}^{6400 \times 6400}$: FD discretiz. of $\mathcal{L}(u) = -\Delta u + \frac{1000}{x+y} u_x$ in $[-1, 1]^2$



Coping with long-term recurrences

Restarted, Truncated, etc variants.

Coping with long-term recurrences

Restarted, Truncated, etc variants.

Restarted: Choose $m_{\max}$.

Set $x = x_0$, $r_0 = b - Ax_0$

for $i = 1, 2, \ldots$

$z \leftarrow$ GMRES$(A, r_0, m_{\max})$ (or other method)

$x \leftarrow x + z, \quad r_0 = b - Ax$

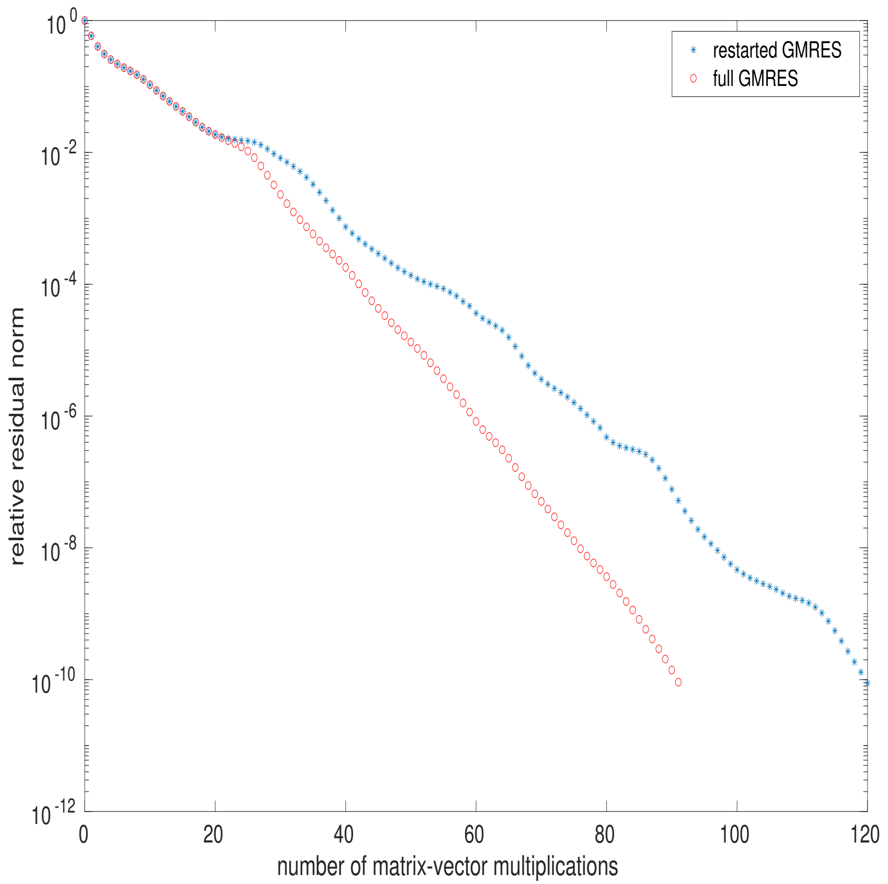
Check Convergence

Restarted GMRES, an example

$$\mathcal{L}(u) = -u_{xx} - u_{yy} - u_{zz} + 100xu_x, \quad \Omega = (0,1)^3$$

$A \in \mathbb{R}^{n\times n}$, $n = 1,000$. GMRES(20) ($m = 20$)

restart	res. norm
1	0.0186592
2	0.000743465
3	3.63848e-05
4	4.77843e-07
5	4.65117e-09
8	.87182e-11



Pros and Cons

Pros:

- Shorter dependencies
- Lower and fixed memory requirements

Pros and Cons

Pros:

- Shorter dependencies
- Lower and fixed memory requirements

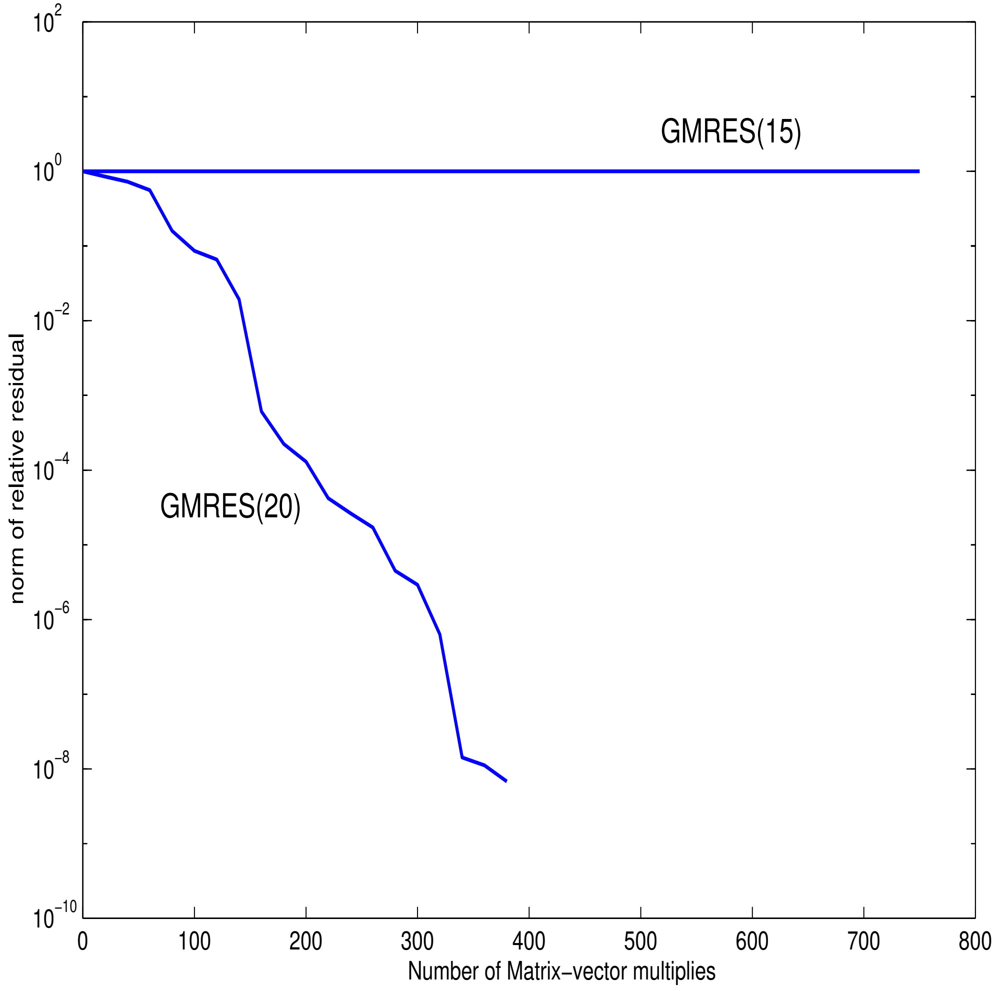
Cons:

- All optimality properties are lost

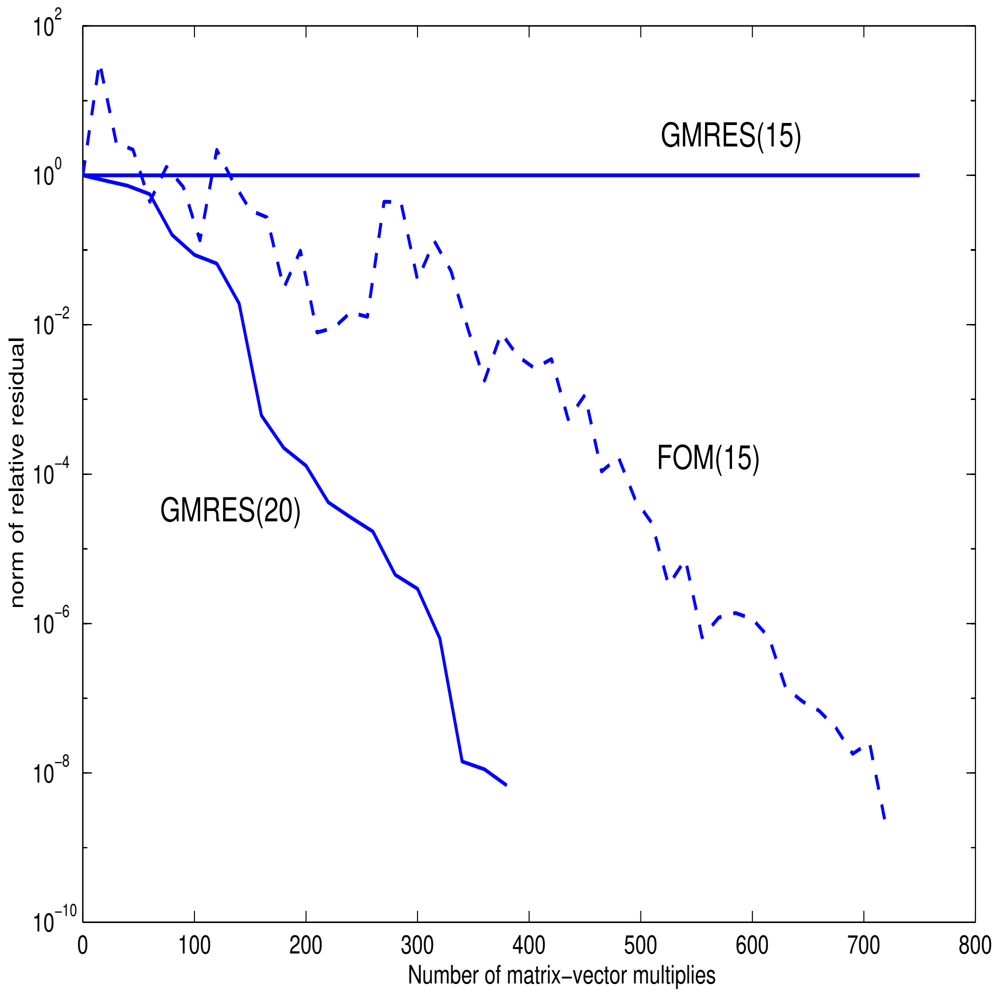
$$K_{m_{\max}}(A, r_0^{(0)}) + K_{m_{\max}}(A, r_0^{(1)}) + \ldots K_{m_{\max}}(A, r_0^{(k)}) + \ldots$$

- Additional parameter. What value for $m_{\max}$??

A problem with the restarting parameter? ...



A problem with the restarting parameter? ... or with the method?



Explanation

$$K_{m_{\max}}(A, r_0^{(0)}) + K_{m_{\max}}(A, r_0^{(1)}) + \ldots K_{m_{\max}}(A, r_0^{(k)}) + \ldots$$

GMRES: $r_0^{(k)} \in \mathrm{range}(V_{m_{\max}+1}^{(k-1)})$. Almost stagnation: $\rightarrow r_0^{(k)} \propto v_1^{(k-1)}$

Explanation

$$K_{m_{\max}}(A, r_0^{(0)}) + K_{m_{\max}}(A, r_0^{(1)}) + \ldots K_{m_{\max}}(A, r_0^{(k)}) + \ldots$$

GMRES: $r_0^{(k)} \in \mathrm{range}(V_{m_{\max}+1}^{(k-1)})$. Almost stagnation: $\rightarrow r_0^{(k)} \propto v_1^{(k-1)}$

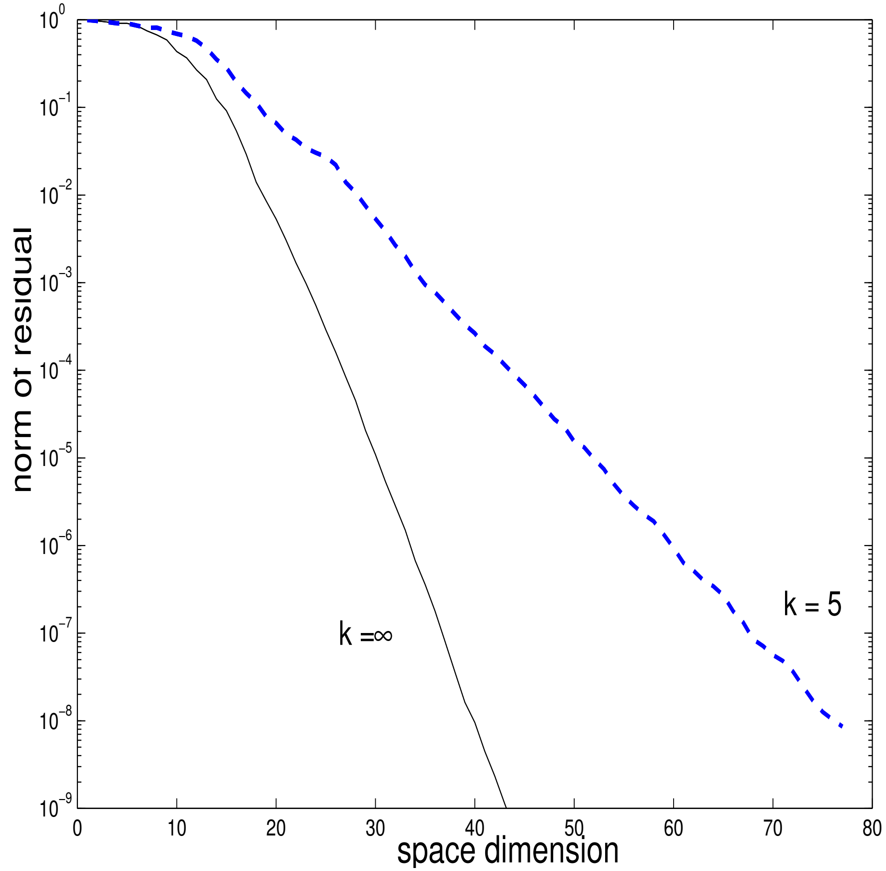
FOM: $r_0^{(k)} \propto v_{m_{\max}+1}^{(k-1)}$ Subspace keeps growing

Truncating

Only local orthogonalization (k-term recurrence, H_m banded)

Truncating

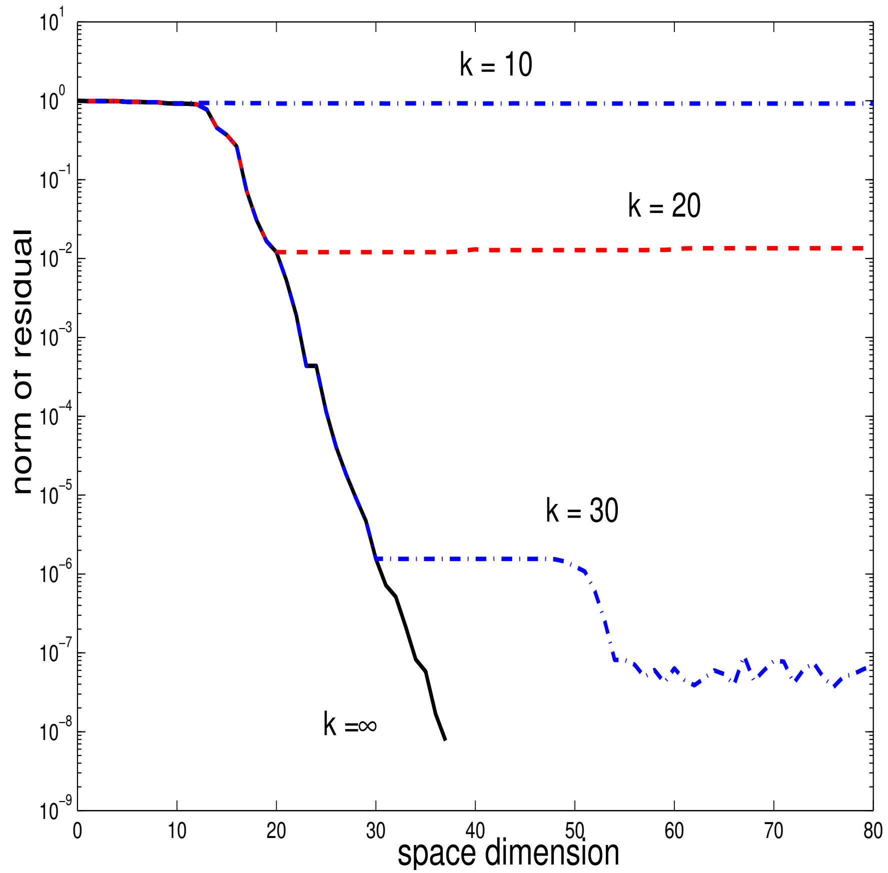
Only local orthogonalization (k-term recurrence, H_m banded)



a reasonable strategy

Truncating

...but not always good



Making it without long-term recurrences: short-term recurrences for A non-Hermitian

- Non-Hermitian Lanczos
- BiCGStab(ℓ): ℓ iterations of GMRES at every step
- IDR(s): $r_k \in \mathcal{G}_k$, where $\mathcal{G}_{k+1} \subset \mathcal{G}_k$