

Krylov Subspace Methods for Linear Systems and Matrix Equations

V. Simoncini

Dipartimento di Matematica

Università di Bologna

valeria@dm.unibo.it

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



Projection methods for large-scale problems

Given the system of n equations

$$\mathcal{F}(x) = 0 \quad x \in \mathbb{R}^n$$

- Construct approximation space \mathcal{K}_m ($m = \dim(\mathcal{K}_m)$)
- Find $\tilde{x} \in \mathcal{K}_m$ such that $\tilde{x} \approx x$
- ★ **Projection** onto a much smaller space $m \ll n$



Projection methods for large-scale problems

Given the system of n equations

$$\mathcal{F}(x) = 0 \quad x \in \mathbb{R}^n$$

- Construct approximation space \mathcal{K}_m ($m = \dim(\mathcal{K}_m)$)

- Find $\tilde{x} \in \mathcal{K}_m$ such that $\tilde{x} \approx x$

★ **Projection** onto a much smaller space $m \ll n$

Approximation process:

residual: $r := \mathcal{F}(\tilde{x})$

Construct (left) space \mathcal{L}_m of dimension m and impose

$$r \perp \mathcal{L}_m$$



Challenges

- How to select \mathcal{K}_m



Challenges

- How to select \mathcal{K}_m
- How to derive good \mathcal{L}_m so that it also carries other good properties



Challenges

- How to select \mathcal{K}_m
- How to derive good \mathcal{L}_m so that it also carries other good properties

General idea:

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

$$\tilde{x}_m \in \mathcal{K}_m \quad \text{and} \quad \tilde{x}_m \rightarrow x \quad \text{as} \quad m \rightarrow \infty$$

(in some sense)

Analogously, $\mathcal{L}_m \subset \mathcal{L}_{m+1}$



More specific problems

$$\mathcal{F}(x) = 0$$

$$A \in \mathbb{R}^{n \times n}$$

• $Ax - b = 0$ $(\varphi_k(A)x - b = 0)$ where φ_k polynomial of degree k



More specific problems

$$\mathcal{F}(x) = 0$$

$$A \in \mathbb{R}^{n \times n}$$

- $Ax - b = 0$ $(\varphi_k(A)x - b = 0)$ where φ_k polynomial of degree k
- $Ax - \lambda Mx = 0$ $(\varphi_k(\lambda, A_0, A_1, \dots, A_k)x = 0)$



More specific problems

$$\mathcal{F}(x) = 0$$

$$A \in \mathbb{R}^{n \times n}$$

- $Ax - b = 0$ $(\varphi_k(A)x - b = 0)$ where φ_k polynomial of degree k
- $Ax - \lambda Mx = 0$ $(\varphi_k(\lambda, A_0, A_1, \dots, A_k)x = 0)$
- $AX + XA^T + Q = 0$



More specific problems

$$\mathcal{F}(x) = 0$$

$$A \in \mathbb{R}^{n \times n}$$

- $Ax - b = 0$ $(\varphi_k(A)x - b = 0)$ where φ_k polynomial of degree k
- $Ax - \lambda Mx = 0$ $(\varphi_k(\lambda, A_0, A_1, \dots, A_k)x = 0)$
- $AX + XA^T + Q = 0$
- Evaluation of Transfer and other matrix functions



Choosing \mathcal{L}_m

Focus on linear system:

$$Ax = b, \quad A \text{ nonsing.}$$

$x_m \in \mathcal{K}_m$:



Choosing \mathcal{L}_m

Focus on linear system:

$$Ax = b, \quad A \text{ nonsing.}$$

$x_m \in \mathcal{K}_m$:

★ $\mathcal{L}_m = \mathcal{K}_m$ If A symmetric positive definite,

$$r_m = b - Ax_m \perp \mathcal{L}_m \quad \Leftrightarrow \quad \|x - x_m\|_A = \min_{\tilde{x} \in \mathcal{K}_m} !$$



Choosing \mathcal{L}_m

Focus on linear system:

$$Ax = b, \quad A \text{ nonsing.}$$

$x_m \in \mathcal{K}_m$:

★ $\mathcal{L}_m = \mathcal{K}_m$ If A symmetric positive definite,

$$r_m = b - Ax_m \perp \mathcal{L}_m \Leftrightarrow \|x - x_m\|_A = \min_{\tilde{x} \in \mathcal{K}_m} !$$

★ $\mathcal{L}_m = A\mathcal{K}_m$

$$r_m = b - Ax_m \perp \mathcal{L}_m \Leftrightarrow \|r_m\|_2 = \min_{\tilde{x} \in \mathcal{K}_m} !$$



Choosing \mathcal{L}_m

Focus on linear system:

$$Ax = b, \quad A \text{ nonsing.}$$

$x_m \in \mathcal{K}_m$:

★ $\mathcal{L}_m = \mathcal{K}_m$ If A symmetric positive definite,

$$r_m = b - Ax_m \perp \mathcal{L}_m \quad \Leftrightarrow \quad \|x - x_m\|_A = \min_{\tilde{x} \in \mathcal{K}_m} !$$

★ $\mathcal{L}_m = AK_m$

$$r_m = b - Ax_m \perp \mathcal{L}_m \quad \Leftrightarrow \quad \|r_m\|_2 = \min_{\tilde{x} \in \mathcal{K}_m} !$$

“Optimal” properties hold for any choice of \mathcal{K}_m Eiermann & Ernst A.N. '01

Typically: $\mathcal{L}_m = \mathcal{K}_m \Rightarrow$ FOM, CG $\quad \mathcal{L}_m = AK_m \Rightarrow$ GMRES, MINRES



Relaxing optimality in \mathcal{L}_m . Truncation

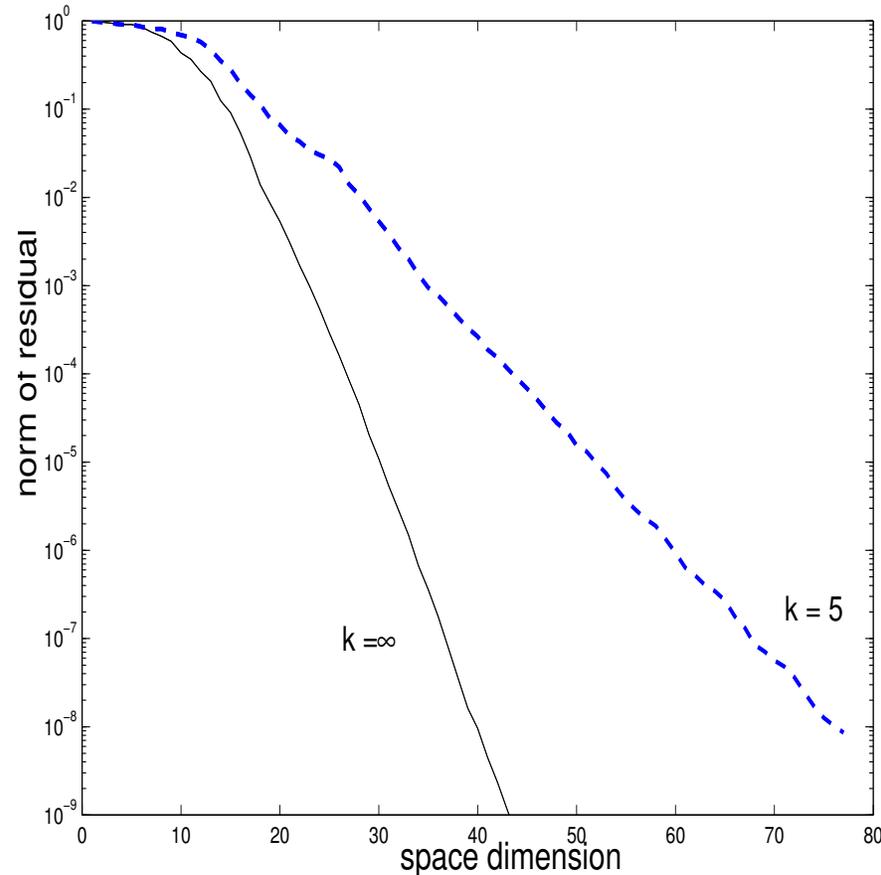
Example: A is non-normal, but “nice” $\mathcal{L}_m = AK_m$



Relaxing optimality in \mathcal{L}_m . Truncation

Example: A is non-normal, but “nice”

$$\mathcal{L}_m = AK_m$$



(-): $r \perp \mathcal{L}_m$

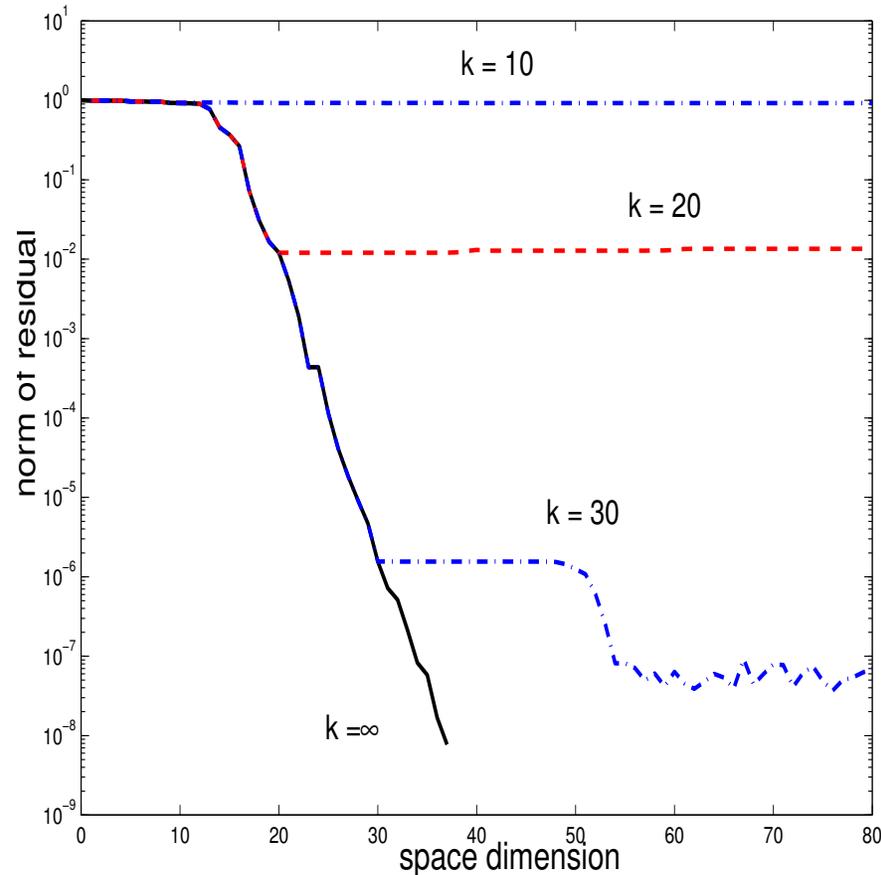
(- -): $r \perp_{loc} \mathcal{L}_m$

k gives “locality”



Relaxing optimality in \mathcal{L}_m . Truncation

Example: A is non-normal, more “nasty” $\mathcal{L}_m = A\mathcal{K}_m$



(-): $r \perp \mathcal{L}_m$

(- -): $r \perp_{loc} \mathcal{L}_m$

Simoncini & Szyld, Num.Math.'05



Choosing \mathcal{K}_m

The “classical” approach

$$\mathcal{K}_m := \mathcal{K}_m(A, v) = \text{span}\{v, Av, \dots, A^{m-1}v\}$$

(e.g., $v = b$ in linear systems)

$$\tilde{x} \in \mathcal{K}_m, \quad \tilde{x} = p_{m-1}(A)v$$



Choosing \mathcal{K}_m

The “classical” approach

$$\mathcal{K}_m := \mathcal{K}_m(A, v) = \text{span}\{v, Av, \dots, A^{m-1}v\}$$

(e.g., $v = b$ in linear systems)

$$\tilde{x} \in \mathcal{K}_m, \quad \tilde{x} = p_{m-1}(A)v$$

Nice Properties:

- For m sufficiently large, $\mathcal{K}_m(A, v)$ invariant for A
- Convergence (analysis) in terms of spectral properties of A
- Variants of “basic” methods by acting on polynomial p_{m-1}



Choosing \mathcal{K}_m

The “classical” approach

$$\mathcal{K}_m := \mathcal{K}_m(A, v) = \text{span}\{v, Av, \dots, A^{m-1}v\}$$

(e.g., $v = b$ in linear systems)

$$\tilde{x} \in \mathcal{K}_m, \quad \tilde{x} = p_{m-1}(A)v$$

Nice Properties:

- For m sufficiently large, $\mathcal{K}_m(A, v)$ invariant for A
- Convergence (analysis) in terms of spectral properties of A
- Variants of “basic” methods by acting on polynomial p_{m-1}

This choice of \mathcal{K}_m is good, but no longer *sufficiently good*



Getting greedy

$$\mathcal{K}_m = \mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k \quad \text{New question: } \mathcal{S}_k?$$



Getting greedy

$$\mathcal{K}_m = \mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k \quad \text{New question: } \mathcal{S}_k?$$

- ★ Faster approximation when spectral a-priori knowledge is available (even for A Hermitian)



Getting greedy

$$\mathcal{K}_m = \mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k \quad \text{New question: } \mathcal{S}_k?$$

★ Faster approximation when spectral a-priori knowledge is available (even for A Hermitian)

★ Main motivating problem:

Large m may be required for accurate approximation x_m



Computational/Memory costs increase **nonlinearly** with m
(A non-normal)



Intermezzo. Restarting procedure

Computational/Memory costs increase **nonlinearly** with m
(A non-normal)

Restarting procedure: x_0 initial approx, $r_0 = b - Ax_0$

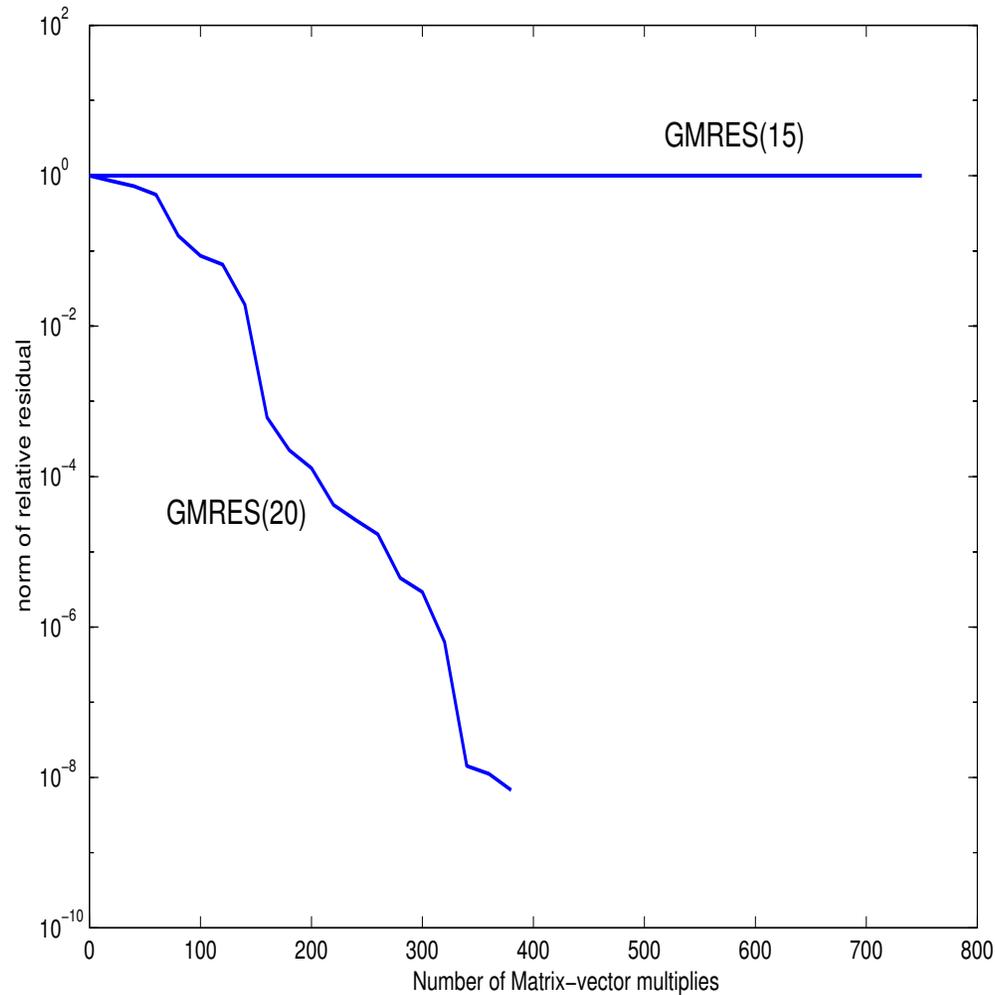
$$\begin{array}{rcl} \mathcal{K}_m(A, r_0) & \rightarrow & x_m^{(1)}, r_m^{(1)} \\ \mathcal{K}_m(A, r_m^{(1)}) & \rightarrow & x_m^{(2)}, r_m^{(2)} \\ & \vdots & \vdots \end{array}$$

Warning: Larger m not always implies faster convergence
(Embree, Ernst, ...)



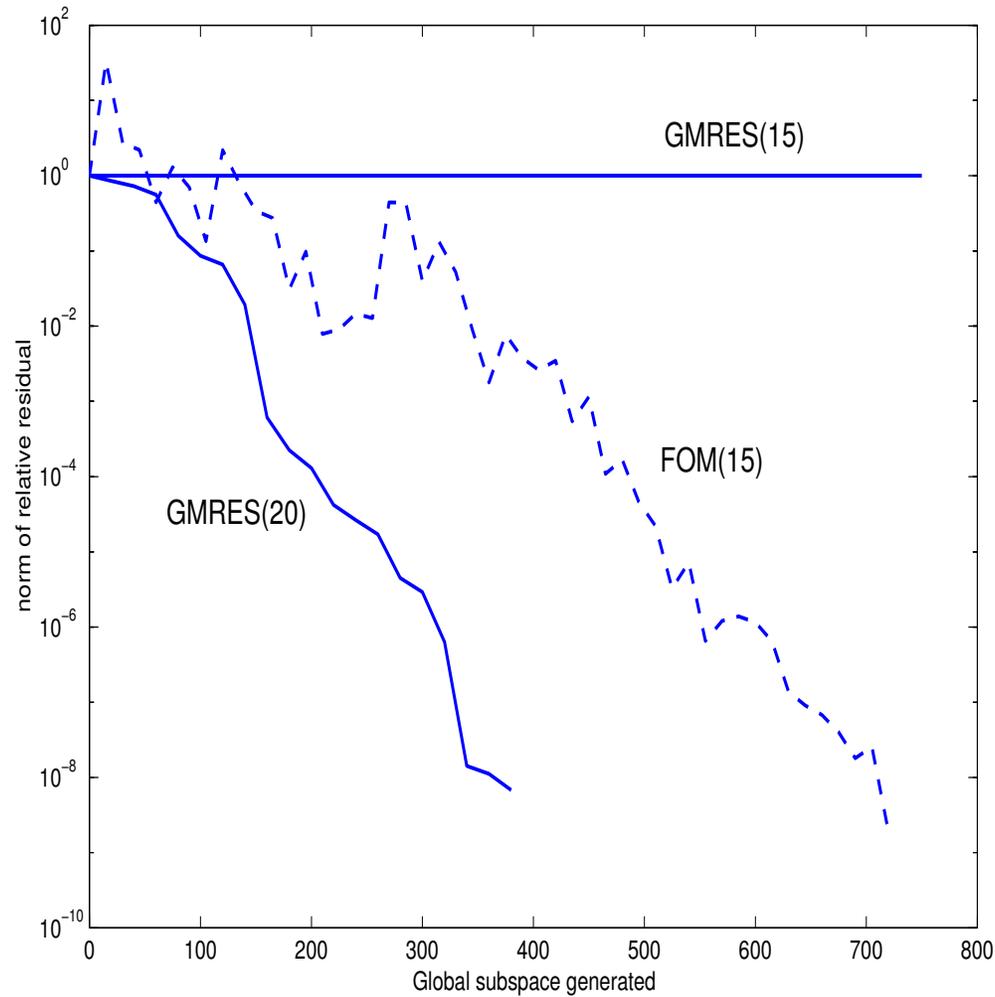
Restarted Methods

Convergence strongly depends on choice of m ...



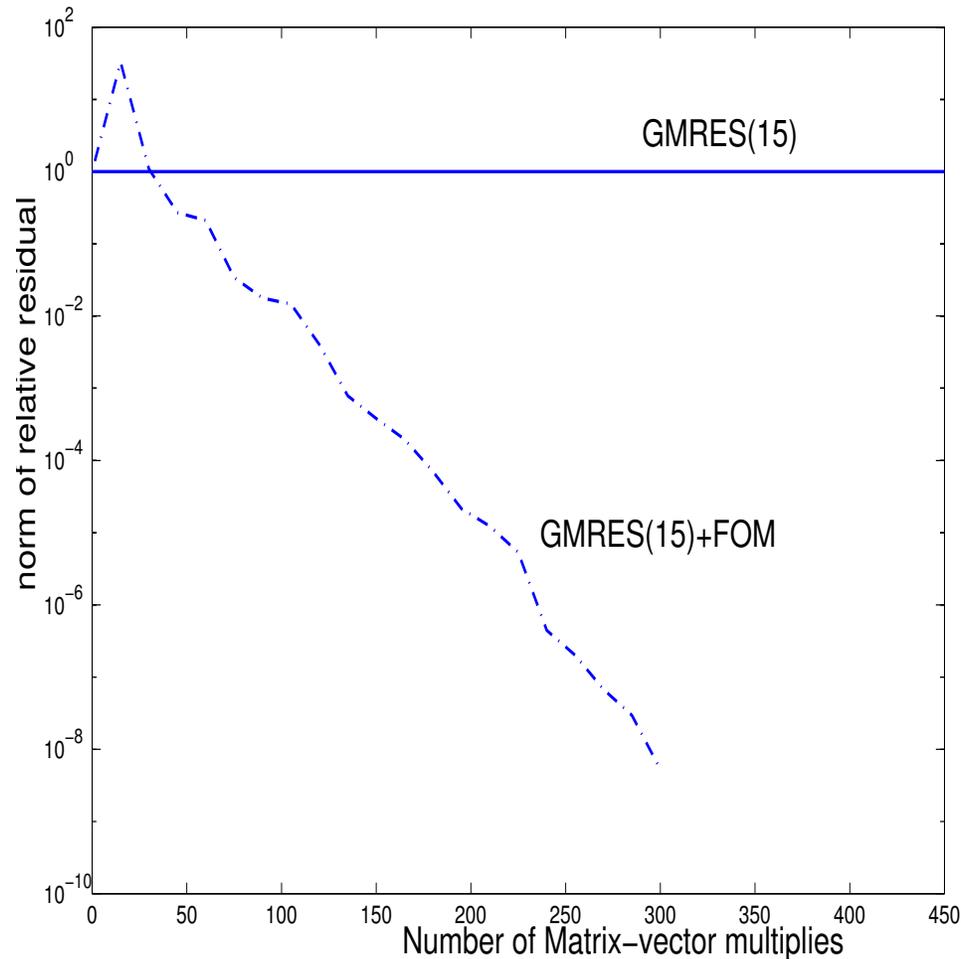
Restarted Methods

Convergence strongly depends on choice of m ... true?



Restarted Methods

Switch to FOM residual vector only at the very first restart



Pictures from Simoncini, SIMAX 2000.

Intermezzo ends.



Augmented projection spaces

$$\mathcal{K}_m = \mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k$$

- \mathcal{S}_k from spectral information of A

$$\mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k = \text{span}\{v, Av, \dots, A^{m-k-1}v, y_1, y_2, \dots, y_k\}$$

y_1, \dots, y_k approximate eigenvectors of A associated to cluster

(Baglama, Calvetti, Erhel, Morgan, Nabben, Reichel, Saad, Sorensen, Vuik ...)



Augmented projection spaces

$$\mathcal{K}_m = \mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k$$

- \mathcal{S}_k from spectral information of A

$$\mathcal{K}_{m-k}(A, v) \cup \mathcal{S}_k = \text{span}\{v, Av, \dots, A^{m-k-1}v, y_1, y_2, \dots, y_k\}$$

y_1, \dots, y_k approximate eigenvectors of A associated to cluster

(Baglama, Calvetti, Erhel, Morgan, Nabben, Reichel, Saad, Sorensen, Vuik ...)

- \mathcal{S}_k **important** space from previous restarts

Ranking based on “error” or “effectiveness” of the past spaces

(De Sturler, Baker, Jessup, Manteuffel, ...)



Augmented method. Spectral Information available.

Structural Dynamics problem: $(\mathcal{A}\mathcal{B}^{-1} + \sigma I)x = b$

\mathcal{A}, \mathcal{B} complex sym. $n = 86,000$. Solve for σ in a wide interval

Note: at each iteration solve system with matrix \mathcal{B}

$\mathcal{B} = \mathcal{B}(\kappa)$ κ related to artificial stiff springs at ground boundaries

\mathcal{B} numerically singular as $\kappa \rightarrow 0$



Augmented method. Spectral Information available.

Structural Dynamics problem: $(\mathcal{A}\mathcal{B}^{-1} + \sigma I)x = b$

\mathcal{A}, \mathcal{B} complex sym. $n = 86,000$. Solve for σ in a wide interval

Note: at each iteration solve system with matrix \mathcal{B}

$\mathcal{B} = \mathcal{B}(\kappa)$ κ related to artificial spring stiffness at ground boundaries

\mathcal{B} numerically singular as $\kappa \rightarrow 0$

	Fill-in p=5		Fill-in p=15	
	E. Time [s]	# its (outer/ avg. inner)	E. Time [s]	# its (outer/ avg. inner)
$\kappa = 0$ ACG	14066	296/38	12790	281/33
$\kappa = 100$	14072	117/120	13739	121/102
$\kappa = 1000$	8694	88/96	8724	89/83

$\kappa = 1000$ unrealistic

Perotti & Simoncini, '02

ACG: Augmented CG, Saad & Yeung & Ehrel & Guyomarc'h, 2000

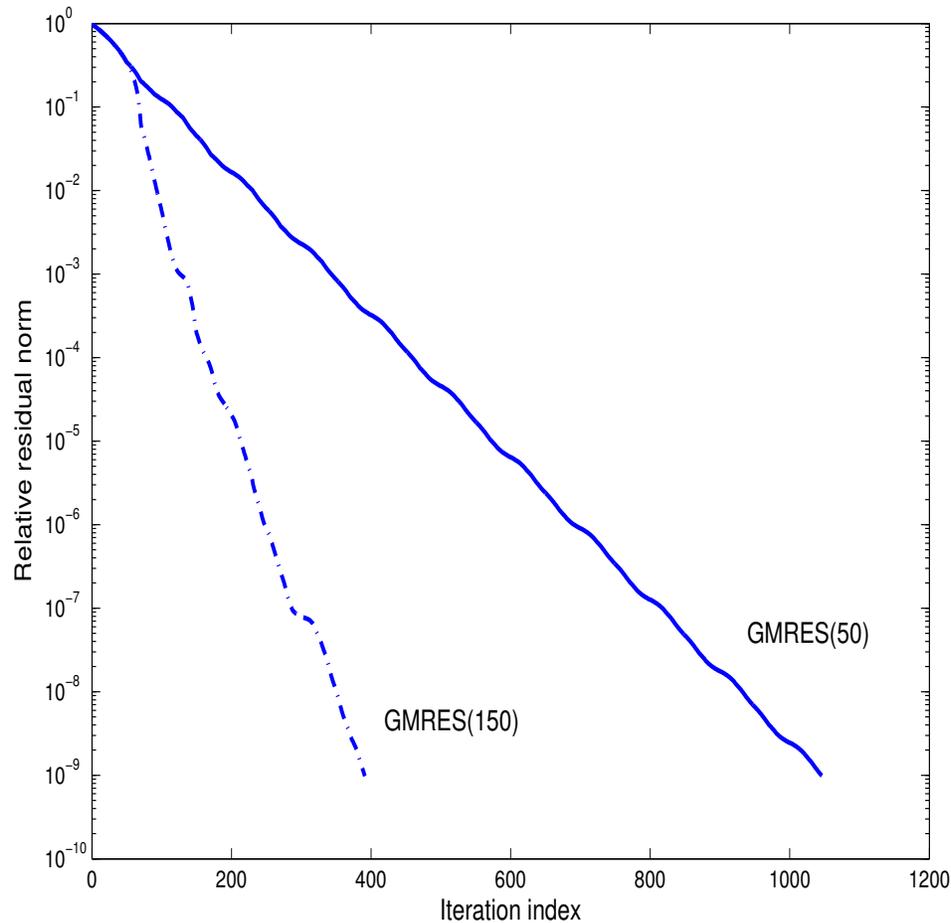


Example of Augmented method. Error Space

Tricky way to enhance approximation space

$$A \quad \text{from } L(u) = -1000\Delta u + 2e^{4(x^2+y^2)}u_x - 2e^{4(x^2+y^2)}u_y$$

$n = 40\,000$

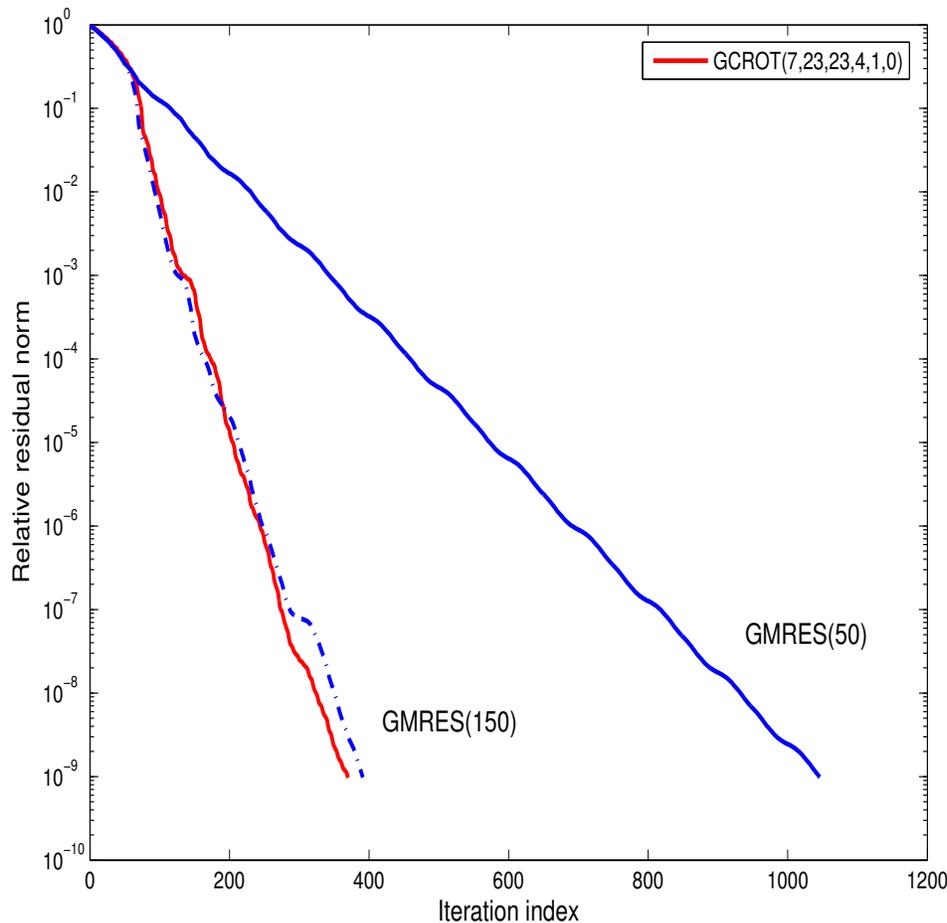


Example of Augmented method. Error Space

Tricky way to enhance approximation space (E. De Sturler)

$$A \quad \text{from } L(u) = -1000\Delta u + 2e^{4(x^2+y^2)}u_x - 2e^{4(x^2+y^2)}u_y$$

$n = 40\,000$



Code: courtesy of Oliver Ernst

see also Baker, Jessup, Manteuffel '05



Università di Bologna

Changing \mathcal{K}_m

Modify \mathcal{K}_m instead of augmenting it!

- Increase flexibility
- Cope with “involved” operators
- ...



Changing \mathcal{K}_m . Flexible methods

Original problem

$$AP^{-1}x = b \quad P \text{ preconditioner}$$

$$\mathcal{K}_m(AP^{-1}, b) = \text{span}\{b, AP^{-1}b, \dots, (AP^{-1})^{m-1}b\}$$

at each iteration i : $z_i = P^{-1}v_i$



Changing \mathcal{K}_m . Flexible methods

Original problem

$$AP^{-1}x = b \quad P \text{ preconditioner}$$

$$\mathcal{K}_m(AP^{-1}, b) = \text{span}\{b, AP^{-1}b, \dots, (AP^{-1})^{m-1}b\}$$

at each iteration i : $z_i = P^{-1}v_i$

Flexible variant:

Saad, '93

$$\text{Iteration } i: \quad z_i = P^{-1}v_i \quad \Rightarrow \quad z_i = P_i^{-1}v_i$$

$$\tilde{\mathcal{K}}_m \in \text{span}\{b, z_1, z_2, \dots, z_{m-1}\} \neq \mathcal{K}_m(AP^{-1}, b)$$



Flexible methods. An example

Flexible method may be used as a Truncated/Augmented method.

$$z = P^{-1}v \quad \Leftrightarrow \quad z \approx A^{-1}v$$

$$\text{span}\{b, z_1, z_2, \dots, z_{m-1}\}$$



Flexible methods. An example

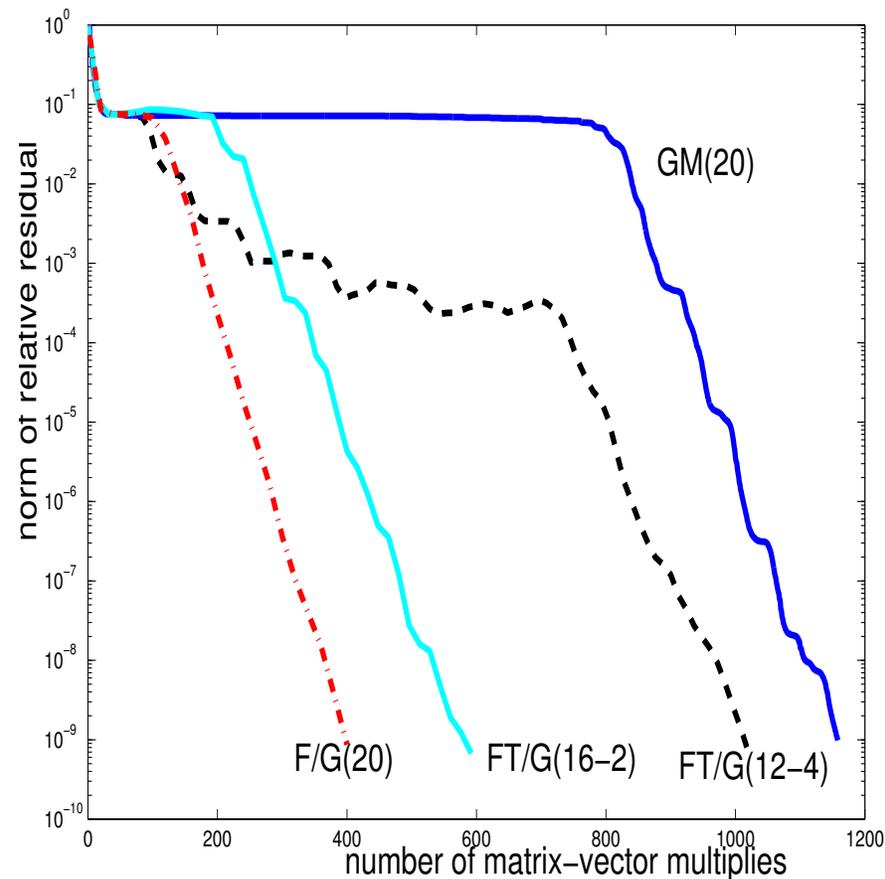
Flexible method may be used as a Truncated/Augmented method.

$$z = P^{-1}v \quad \Leftrightarrow z \approx A^{-1}v$$

$$\text{span}\{b, z_1, z_2, \dots, z_{m-1}\}$$

A from $L(u) = -\Delta u + 1000xu_x$

$$n = 900$$



Simoncini & Szyld, SINUM '03

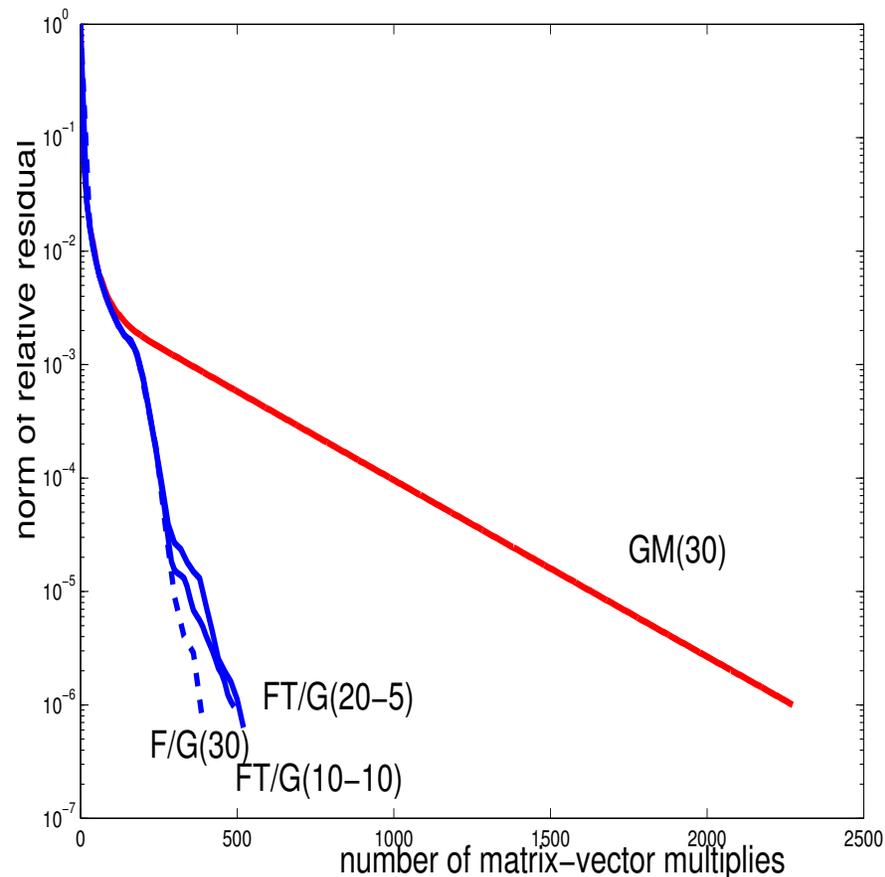


Flexible methods. A second example

Flexible method may be used as a Truncated/Augmented method.

$$z = P^{-1}v \quad \Leftrightarrow \quad z \approx A^{-1}v \quad \text{span}\{b, z_1, z_2, \dots, z_{m-1}\}$$

$$A \quad \text{from } L(u) = -1000\Delta u + 2e^{4(x^2+y^2)}u_x - 2e^{4(x^2+y^2)}u_y \quad n = 40\,000$$



Simoncini & Szyld, SINUM '03



Changing \mathcal{K}_m . Inexact methods.

Original problem

$$Ax = b \quad A \quad \text{Possibly not available exactly}$$

$$\mathcal{K}_m(A, b) = \text{span}\{b, Ab, \dots, A^{m-1}b\}$$



Changing \mathcal{K}_m . Inexact methods.

Original problem

$$Ax = b \quad A \quad \text{Possibly not available exactly}$$

$$\mathcal{K}_m(A, b) = \text{span}\{b, Ab, \dots, A^{m-1}b\}$$

Inexact (relaxed) variant:

$$\text{Iteration } i: \quad z_i = Av_i \quad \Rightarrow \quad z_i = Av_i + f_i$$

$$\tilde{x}_m \in \text{span}\{b, z_1, z_2, \dots, z_{m-1}\} \neq \mathcal{K}_m(A, b)$$



Changing \mathcal{K}_m . Inexact methods.

Original problem

$$Ax = b \quad A \quad \text{Possibly not available exactly}$$

$$\mathcal{K}_m(A, b) = \text{span}\{b, Ab, \dots, A^{m-1}b\}$$

Inexact (relaxed) variant:

$$\text{Iteration } i: \quad z_i = Av_i \quad \Rightarrow \quad z_i = Av_i + f_i$$

$$\tilde{x}_m \in \text{span}\{b, z_1, z_2, \dots, z_{m-1}\} \neq \mathcal{K}_m(A, b)$$

♣ “Worse” than for Flexible methods:

$$r_m = b - Ax_m \quad \text{not available!}$$

Available: \tilde{r}_m computable residual



Inexact Methods. An example

if $\|f_i\| = O\left(\frac{\varepsilon}{\|\tilde{r}_{i-1}\|}\right) \quad \forall i \quad \Rightarrow \quad \|b - Ax_m\| \leq \varepsilon \quad \text{for } m \text{ large enough}$



Inexact Methods. An example

if $\|f_i\| = O\left(\frac{\varepsilon}{\|\tilde{r}_{i-1}\|}\right) \quad \forall i \quad \Rightarrow \quad \|b - Ax_m\| \leq \varepsilon \quad \text{for } m \text{ large enough}$

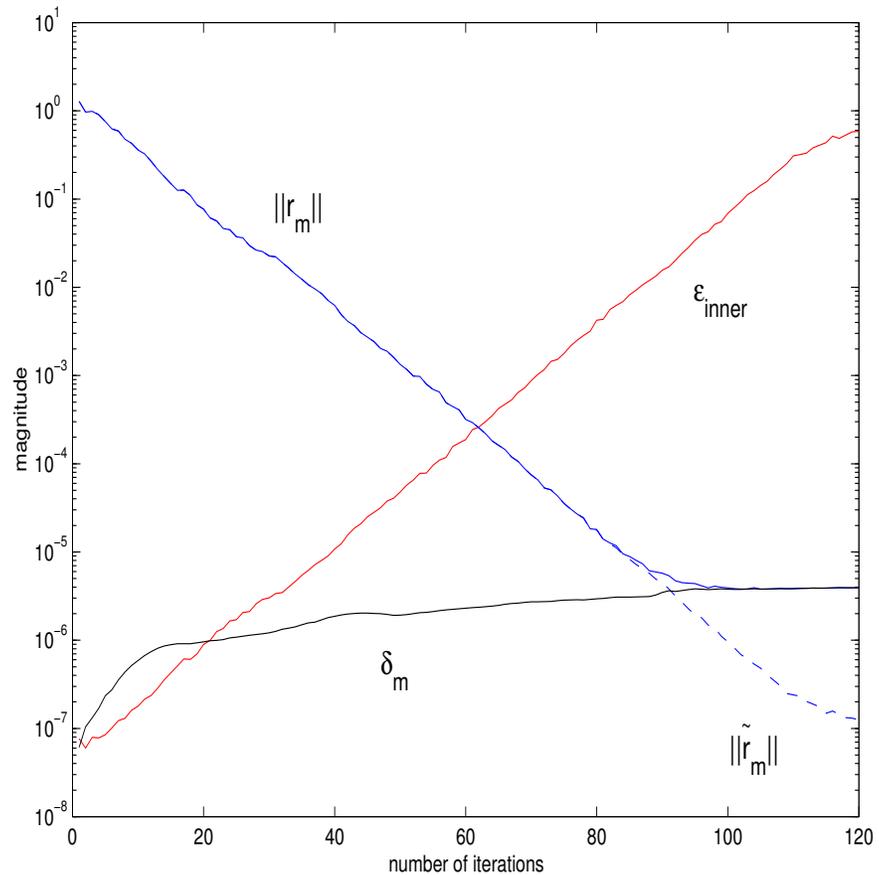
$$\underbrace{B^T S^{-1} B}_A x = b$$

At each it. i solve:

$$Sw_i = Bv_i$$

$$\|Sw_i - Bv_i\| \leq \epsilon_{\text{inner}}(\|f_i\|)$$

$$\delta_m = \|r_m - \tilde{r}_m\|$$



Simoncini & Szyld, SISC '03



A different application. The Lyapunov equation

$$AX + XA^T + Q = 0$$

with A dissipative, $Q = BB^T$ of low rank $X \approx \tilde{X}$ low rank



A different application. The Lyapunov equation

$$AX + XA^T + Q = 0$$

with A dissipative, $Q = BB^T$ of low rank $X \approx \tilde{X}$ low rank

• Standard Krylov approach: $\tilde{X} \in \mathcal{K}_m(A, B)$ and

$$R = A\tilde{X} + \tilde{X}A^T + Q \perp \mathcal{L}_m = \mathcal{K}_m(A, B)$$

$\tilde{X} = V_m Y_m V_m^T$ for some Y_m $\text{Range}(V_m) = \mathcal{K}_m(A, B)$ Saad, '90



A different application. The Lyapunov equation

$$AX + XA^T + Q = 0$$

with A dissipative, $Q = BB^T$ of low rank $X \approx \tilde{X}$ low rank

● Standard Krylov approach: $\tilde{X} \in \mathcal{K}_m(A, B)$ and

$$R = A\tilde{X} + \tilde{X}A^T + Q \quad \perp \quad \mathcal{L}_m = \mathcal{K}_m(A, B)$$

$\tilde{X} = V_m Y_m V_m^T$ for some Y_m $\text{Range}(V_m) = \mathcal{K}_m(A, B)$ Saad, '90

● New “Enhanced” approach: $\tilde{X} \in \mathcal{K}_k(A, B) \cup \mathcal{K}_k(A^{-1}, B)$ and

$$R = A\tilde{X} + \tilde{X}A^T + Q \quad \perp \quad \mathcal{L}_m = \mathcal{K}_k(A, B) \cup \mathcal{K}_k(A^{-1}, B)$$



A different application. The Lyapunov equation

$$AX + XA^T + Q = 0$$

with A dissipative, $Q = BB^T$ of low rank $X \approx \tilde{X}$ low rank

● Standard Krylov approach: $\tilde{X} \in \mathcal{K}_m(A, B)$ and

$$R = A\tilde{X} + \tilde{X}A^T + Q \perp \mathcal{L}_m = \mathcal{K}_m(A, B)$$

$\tilde{X} = V_m Y_m V_m^T$ for some Y_m $\text{Range}(V_m) = \mathcal{K}_m(A, B)$ Saad, '90

● New “Enhanced” approach: $\tilde{X} \in \mathcal{K}_k(A, B) \cup \mathcal{K}_k(A^{-1}, B)$ and

$$R = A\tilde{X} + \tilde{X}A^T + Q \perp \mathcal{L}_m = \mathcal{K}_k(A, B) \cup \mathcal{K}_k(A^{-1}, B)$$

★ Very competitive w.r.to Cyclic ADI method

Simoncini, tr. 2006



An example. Time-invariant linear system

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10z\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t)$$

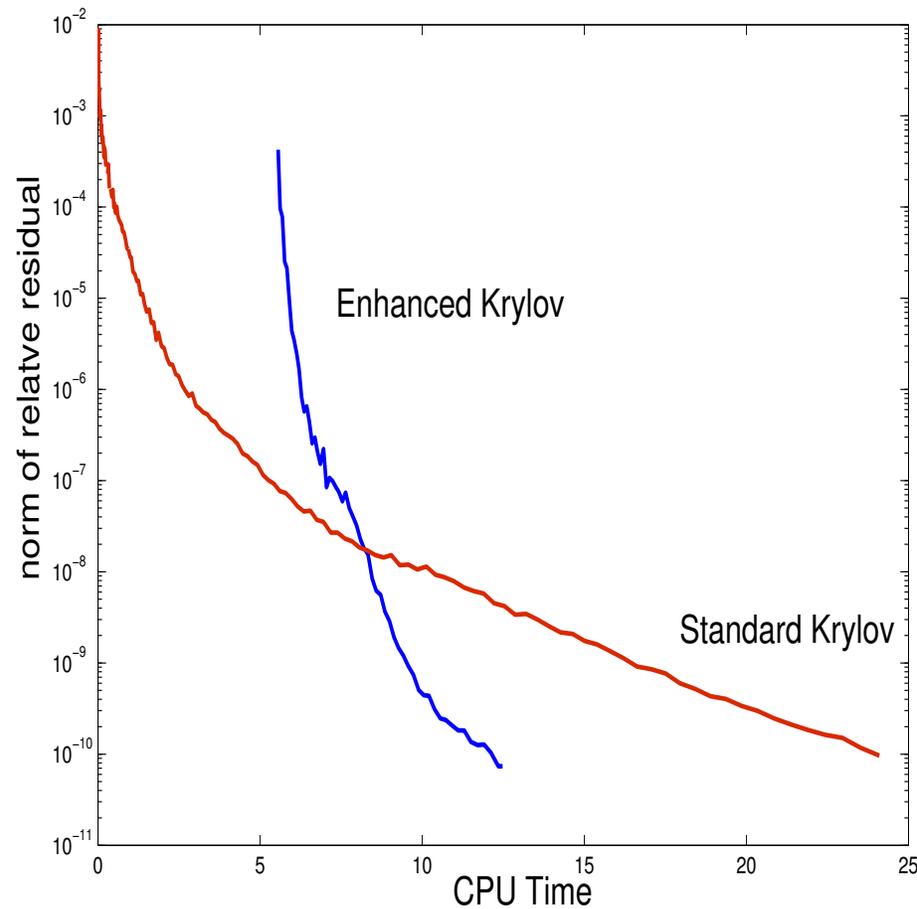
A matrix $18^3 \times 18^3$



An example. Time-invariant linear system

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10z\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t)$$

A matrix $18^3 \times 18^3$



approximation space dim.: 146 (Standard Krylov)

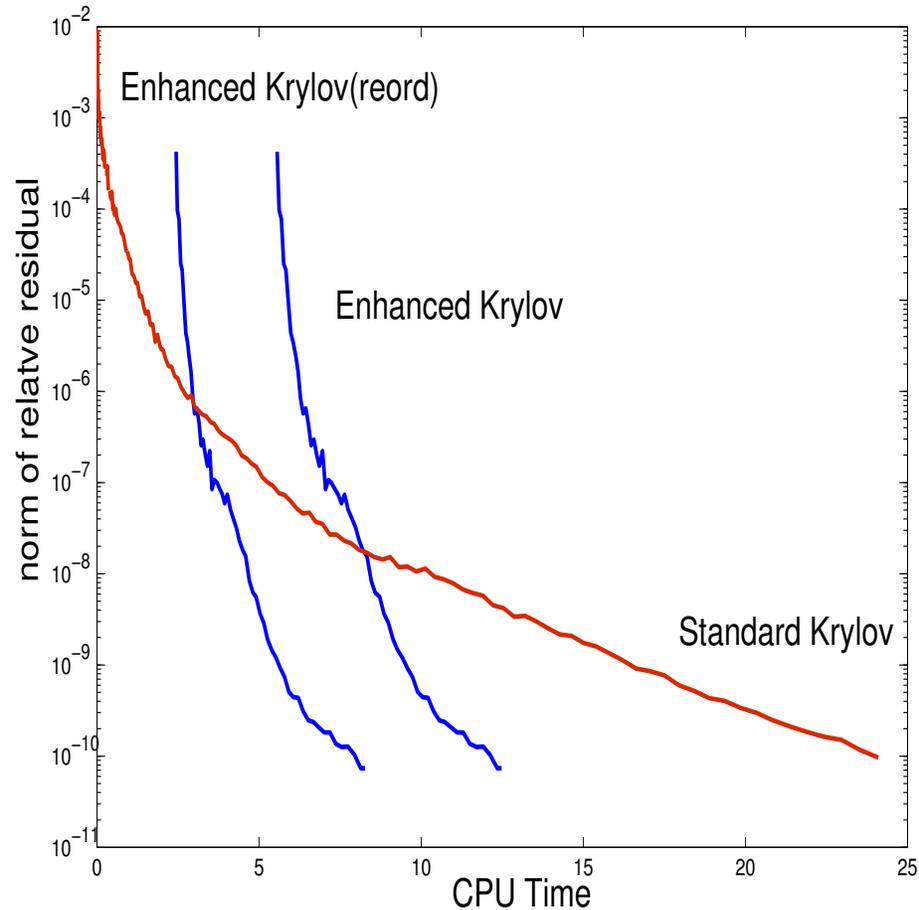
112 (Enhanced Krylov)



An example. Time-invariant linear system

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t)$$

A matrix $18^3 \times 18^3$



approximation space dim.: 146 (Standard Krylov)

112 (Enhanced Krylov)



Conclusions and Pointers

- Projection is a versatile tool

Lots of room for improvements on hard problems



Conclusions and Pointers

- Projection is a versatile tool

Lots of room for improvements on hard problems

- Survey

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

with Daniel Szyld, Temple University

To appear in J. Numerical Linear Algebra w/Appl. (352 refs)

This and other papers at

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



Structural Dynamics Problem

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}(t) \quad n \text{ d.o.f.}$$

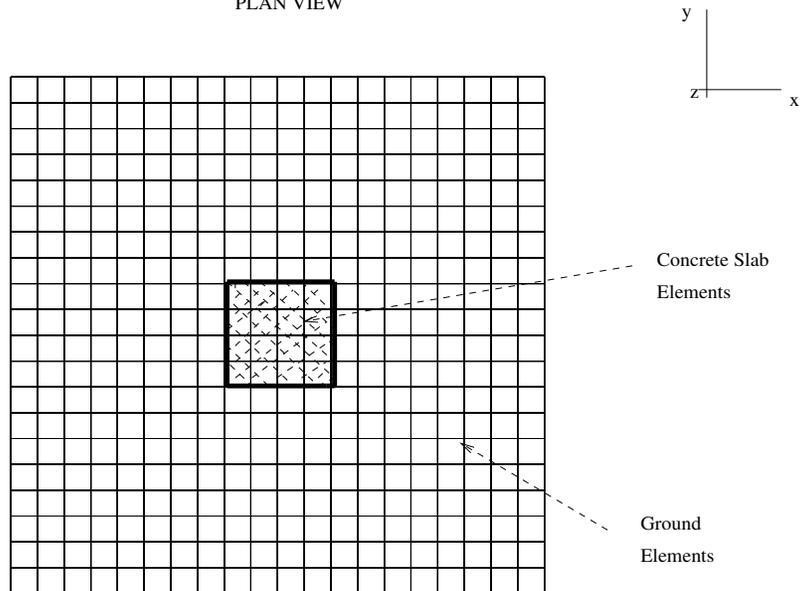
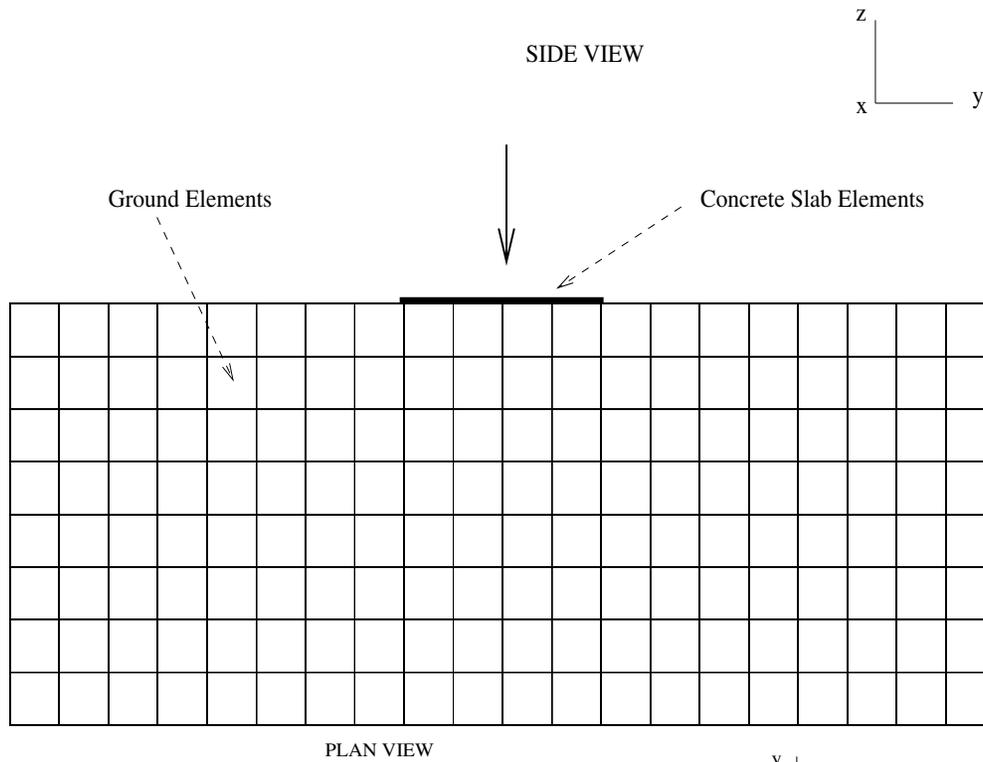
- Linearization under small deformations
- Direct frequency analysis for damping modeling
(as opposed to modal or time analysis)
 - ★ ideal for mechanical properties depending on frequency
 - influence of deformation velocity on materials
 - presence of hysteretic damping, ...

In the frequency domain: ($\bar{\mathbf{a}} = -(2\pi f)^2 \bar{\mathbf{q}}(f)$):

$$\left(\frac{1}{-(2\pi f)^2} \mathbf{K}_* + \frac{1}{i(2\pi f)} \mathbf{C}_V + \mathbf{M} \right) \bar{\mathbf{a}} = \mathbf{b}$$

$$\mathbf{C} = \mathbf{C}_V + \frac{1}{2\pi f} \mathbf{C}_H, \quad \Rightarrow \quad \mathbf{K}_* := \mathbf{K} + i\mathbf{C}_H$$





Model Problem: 3D Soil-Structure interaction

★ viscous damping at the boundary $\Rightarrow \mathbf{K}_*$ singular

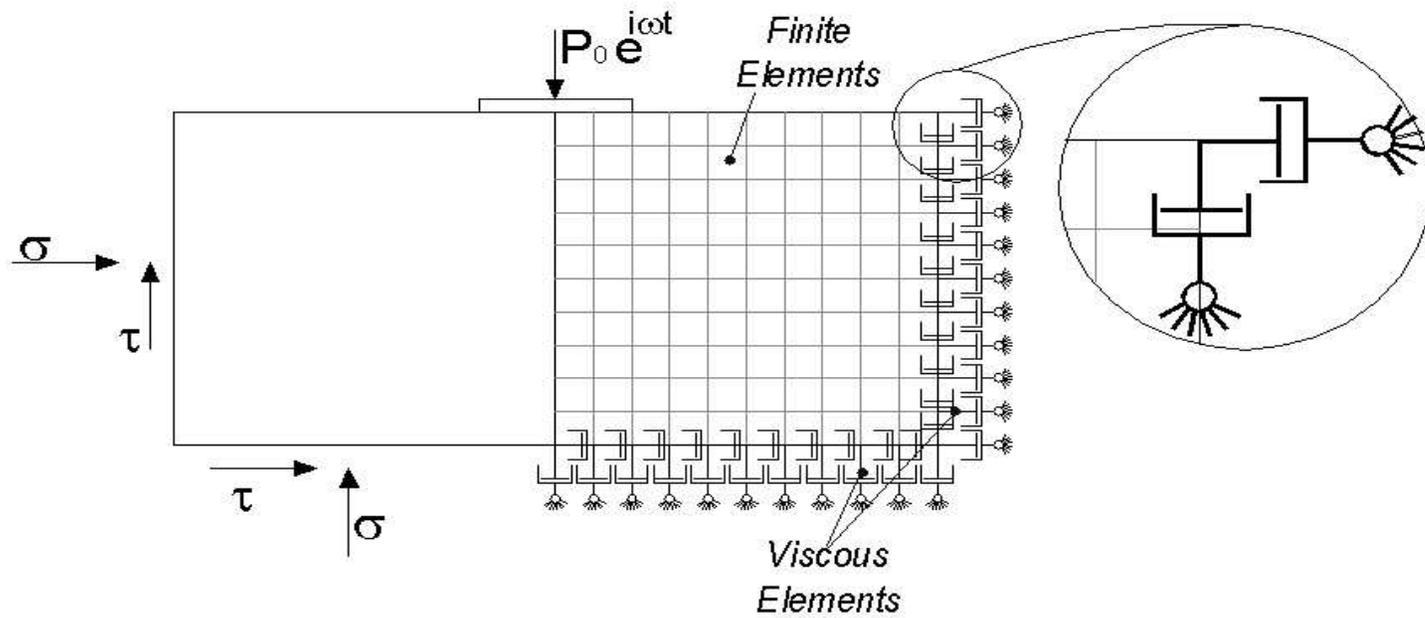
\Rightarrow FE discretization, \mathbf{K}_* almost singular

Algebraic Linearization: $(\sigma^2 \mathbf{K}_* + \sigma i \mathbf{C}_V - \mathbf{M}) \mathbf{x} = \mathbf{b}$, $\mathbf{x} = \mathbf{x}(\sigma)$
equivalent to

$$\left(\underbrace{\begin{bmatrix} i\mathbf{C}_V & -\mathbf{M} \\ -\mathbf{M} & 0 \end{bmatrix}}_A + \sigma \underbrace{\begin{bmatrix} \mathbf{K}_* & 0 \\ 0 & \mathbf{M} \end{bmatrix}}_B \right) \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}$$

with $\mathbf{y} = \sigma \mathbf{x}$





+

