Modern Krylov subspace methods (and applications to parabolic control problems)

Daniel B. Szyld
Temple University, Philadelphia

Scientific and Statistical Computing Seminar
University of Chicago
25 October 2012

Collaborators

Xiuhong Du, Alfred University

Marcus Sarkis, Worcester Poly. Inst., and IMPA, Rio de Janeiro

Christian Schaerer, Universidad Nacional de Asunción

Valeria Simoncini, Università di Bologna

Plan for the talk

- Introduction to Truncated Krylov subspace methods
- Inexact Krylov subspace methods:
  Introduction and Applications
- Special case of parabolic control problems

General Problem Statement

Solve a system

$$Hx = b,$$

H Hermitian or non-Hermitian
using Krylov subspace iterative methods

Krylov subspace methods

$$\mathcal{K}_m(H, r_0) = \text{span}\{r_0, Hr_0, H^2 r_0, \ldots, H^{m-1} r_0\}.$$

Subspaces are nested: $\mathcal{K}_m \subset \mathcal{K}_{m+1}$.
Given x_0, $r_0 = b - Hx_0$, find approximation

$$x_m \in x_0 + \mathcal{K}_m(H, r_0),$$

satisfying some property.

Krylov subspace methods (cont.)

Conditions:

Galerkin, e.g., FOM, CG:

$$b - Hx_m \perp \mathcal{K}_m(H, r_0)$$

Petrov-Galerkin, e.g., GMRES, MINRES:

$$b - Hx_m \perp H\mathcal{K}_m(H, r_0)$$

or equivalently

$$x_m = \text{arg min}\{\|b - Hx\|_2\}, \quad x \in x_0 + \mathcal{K}_m(H, r_0)$$

Krylov subspace methods (cont.)
Some implementation issues

- Methods work by suitably choosing a basis of $\mathcal{K}_m(H, r_0)$
- Let $v_1, v_2, \ldots, v_m$ be such a basis, chosen to be orthonormal.
- One can of course run Gram-Schmidt on $\{r_0, Hr_0, H^2 r_0, \ldots, H^{m-1} r_0\}$ (not advised).

  Instead, Arnoldi[1951] (for general H) said:

  $v_1 = r_0$ normalized

  $v_2 = Hv_1 - \langle v_1, Hv_1 \rangle v_1$ normalized,

  $\text{span}\{v_1, v_2\} = \text{span}\{r_0, Hr_0\}$

  $v_3 = Hv_2 - \langle v_2, Hv_2 \rangle v_2 - \langle v_1, Hv_2 \rangle v_1$ normalized,

  $\text{span}\{v_1, v_2, v_3\} = \text{span}\{r_0, Hr_0, H^2 r_0\}$

  etc.

Arnoldi method

Let $\beta = \|r_0\|$, and $v_1 = r_0/\beta$.

For $k = 1, \ldots$

Compute Hv_k, then $\quad v_{k+1}h_{k+1,k} = Hv_k - \sum_{j=1}^{k} v_j h_{jk}$,

where $h_{jk} = \langle v_j, Hv_k \rangle$, $j \leq k$,

and $h_{k+1,k}$ is positive and such that $\|v_{k+1}\| = 1$.

In practice: Modified Gram-Schmidt or Householder orthogonalization

With $V_m = [v_1, v_2, \ldots, v_m]$, obtain Arnoldi relation:

$$HV_m = V_{m+1}H_{m+1,m}$$

$H_{m+1,m}$ is $(m+1) \times m$ upper Hessenberg

Arnoldi relation (cont.)

$$H_{m+1,m} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1m} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2m} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & h_{m,m-1} & h_{mm} \\ & & & & h_{m+1,m} \end{bmatrix}$$

$$= \begin{bmatrix} H_m \\ h_{m+1,m} e_m^T \end{bmatrix}$$

$$HV_m = V_{m+1} H_{m+1,m} = V_m H_m + h_{m+1,m} v_{m+1} e_m^T$$

Krylov subspace methods (cont.)
More implementation issues

Element in $\mathcal{K}_m(H, v_1)$ is a linear combination of $v_1, v_2, \ldots, v_m$, i.e., of the form $V_m y$, $y \in \mathbb{R}^m$

Each method finds $y = y_m$ and we have $x_m = V_m y_m$
For FOM, or CG, we have the Galerkin condition $r_m \perp \mathcal{K}_m$, i.e.,

$$0 = V_m^T (b - H x_m) = V_m^T b - V_m^T H V_m y_m \quad = \quad \beta e_1 - H_m y_m$$

FOM: Full Orthogonalization Method [Saad, 1978]

Use Arnoldi methods to construct V_m, H_m,

Solve $H_m y_m = \beta e_1$

Approximation is $x_m = V_m y_m$.

Test for convergence: Is $\|r_m\| < \varepsilon$?

Cost: one matrix-vector product per step,
at step m, orthogonalization (m inner-products),
one solution of small $m \times m$ upper Hessenberg matrix (LU with no fill).
($\|r_m\|$ cheap or free - no details here).

Storage: at step m, m vectors $v_1, v_2, \ldots, v_m$.

GMRES implementation

We use Arnoldi methods to obtain V_m, and as before, element in $\mathcal{K}_m(H, v_1)$ is of the form $V_m y$, $y \in \mathbb{R}^m$.

Recall: $\beta = \|r_0\|$, $\quad HV_m = V_m H_{m+1,m}$

$$\begin{aligned} b - Hx_m = r_0 - HV_m y_m &= \beta v_1 - V_{m+1} H_{m+1,m} y \\ &= V_{m+1}(\beta e_1 - H_{m+1,m} y) \end{aligned}$$

$$\|r_m\| = \min_{x \in \mathcal{K}_m} \|b - Hx\| = \min_{y \in \mathbb{R}^m} \|\beta e_1 - H_{m+1,m} y\|$$

QR factorization $H_{m+1,m} = Q_{m+1} R_{m+1,m}$, $R_{m+1,m} = \begin{bmatrix} R_m \\ 0 \end{bmatrix}$,

$$\|r_m\| = \min_{y \in \mathcal{R}^m} \|Q_{m+1}^T \beta e_1 - R_{m+1,m} y\|.$$

$$\|r_m\| = \min_{y \in \mathcal{R}^m} \|Q_{m+1}^T \beta e_1 - R_{m+1,m} y\|.$$

$$Q_{m+1}^T \beta e_1 = \begin{bmatrix} t_m \\ \rho_{m+1} \end{bmatrix}$$

Then, $y_m = R_m^{-1} t_m$, and $x_m = V_m y_m$.
Furthermore, $\|r_m\| = \|Q_{m+1}^T \beta e_1 - R_{m+1,m} y_m\| = |\rho_{m+1}|$

Use this computed residual for stopping
(may deviate from true residual!)

Use Givens rotations for QR, save rotations from previous steps.
Only two entries per step needed.

GMRES

Cost: one matrix-vector product per step,
at step m, orthogonalization (m inner-products),
rotations for QR, solution of $m \times m$ triangular system.

Storage: at step m, m vectors $v_1, v_2, \ldots, v_m$.

Residual norms monotonically nonincreasing, but stagnation possible.
Superlinear convergence can be observed.

- Main costs:
  1. Matrix-vector product: Hv_k
  2. Orthogonalization
  3. Storage (if there is no recursion)
- One popular alternative: Restarted methods. Hit and miss. Little theory. [Saad, 1996], [Simoncini, 2000]
- It may happen than larger value of restart parameter is less efficient! [Embree, 2003]

This Talk

- Consider the case when one does not fully orthogonalize: Truncated methods.
- Reduce the cost of matrix-vector product when H is either
  - Not known exactly
  - Computationally expensive (e.g., Schur complement, reduced Hessian)
  - Preconditioned with variable matrix (i.e., iteration dependent)
- Apply all this to Parabolic Control Problems
- Use a Parareal-in-time approximation

Truncated Krylov subspace methods

For the same amount of storage (max ℓ vectors), instead of restarting:
Only orthogonalize with respect of the previous ℓ vectors.
In Arnoldi we have then:

$$v_{k+1}h_{k+1,k} = Hv_k - \sum_{j=\max\{1,k-\ell+1\}}^{k} v_j h_{jk},$$

where $h_{jk} = \langle v_j, Hv_k\rangle$, $j \leq k$,
and $h_{k+1,k}$ is positive and such that $\|v_{k+1}\| = 1$.

Only need to store these previous ℓ vectors.

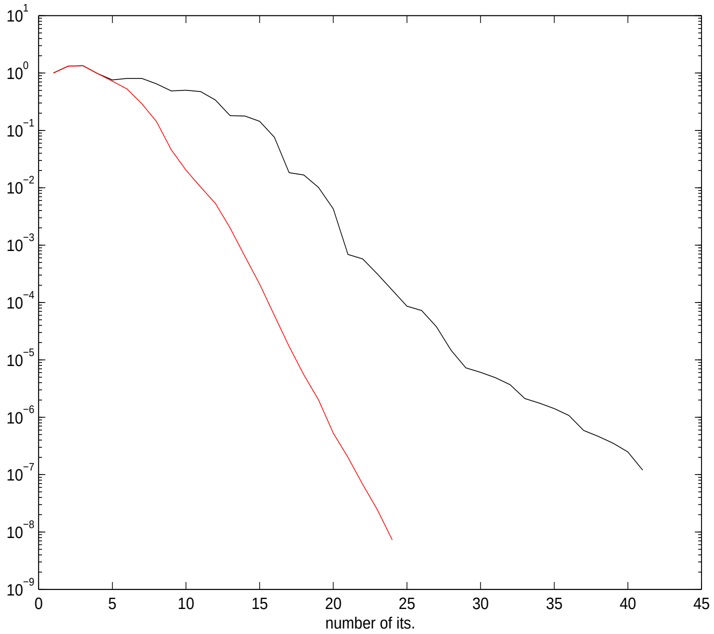
Upper Hessenberg matrix is now banded (upper bandwidth ℓ).

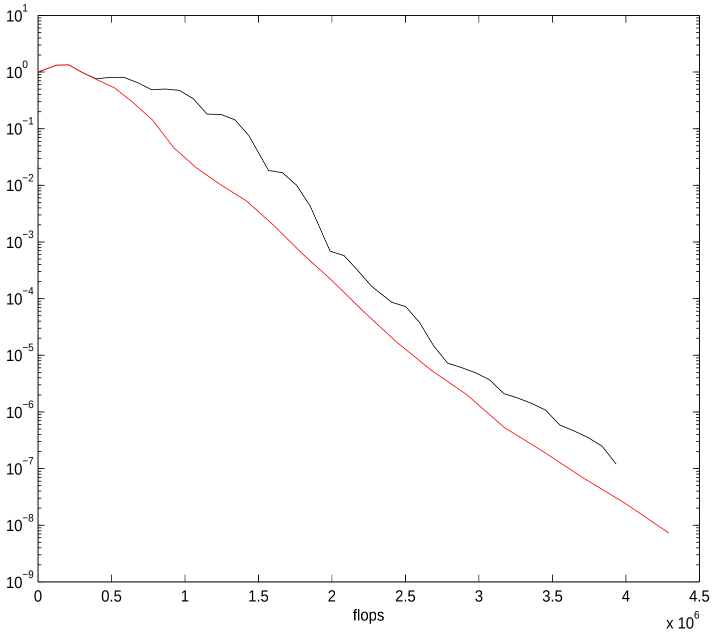
$$H_{m+1,m} \quad = \quad \begin{bmatrix} h_{11} & h_{12} & h_{13} & & & \\ h_{21} & h_{22} & h_{23} & & & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & h_{m,m-1} & h_{mm} \\ & & & & h_{m+1,m} \end{bmatrix}$$

Truncated Krylov subspace methods (cont.)

- Truncated GMRES [Saad and Wu, 1996]
  Truncated FOM [Saad, 1981], [Jia, 1996]
- Basis of $\mathcal{K}_m(H, r_0)$, $v_1, v_2, \ldots, v_m$ $(m > \ell)$ is not orthogonal,
  but $x_m \in \mathcal{K}_m(H, r_0)$,
  and minimization (or Galerkin condition) is over the whole space.
- $H_{m+1,m}$ banded with upper semiband $\ell - 2$.
  Matrix with basis vectors V_m not orthogonal.
  Can be implemented so that only O(ℓ) vectors are stored.
- Extreme case, $\ell = 3$, $H_{m+1,m}$ tridiagonal.
  If H is SPD, FOM reduces to CG (and V_m automatically orthogonal).
- Theory for "non-optimal methods" [Simoncini and Szyld, 2005]

Example: $L(u) = -u_{xx} + -u_{yy} + 100(x+y)u_x + 100(x+y)u_y$, on $[0,1]^2$, Dirichlet b.c., centered 5 pts. discretization, $n = 2500$.





GMRES, Truncated $\ell = 3$.

Inexact Krylov subspace methods

- At the kth iteration of the Krylov space method use

$$(H + D_k)v_{k-1} \text{ instead of } Hv_{k-1},$$

  where $\|D_k\|$ can be monitored

- Two examples now:
  - Schur complement, where the inverse is approximated
  - Inexact preconditioning

- [Bouras, Frayssé, and Giraud, CERFACS reports 2000, SIMAX 2005] show experimentally that as k progresses $\|D_k\|$ can be allowed to be larger; see also [Golub and Ye, 1999], [Notay, 1999], [Sleijpen and van der Eshof, 2004] and [Simoncini and Eldén, 2002]

Inexact Krylov (cont.)

We repeat: $\|D_k\|$ small at first, $\|D_k\|$ can be big later. Convergence is maintained!

- Instead of $HV_m = V_{m+1}H_{m+1,m}$ we have now

$$[(H + D_1)v_1, (H + D_2)v_2, \ldots, (H + D_m)v_m] \quad = \quad V_{m+1}H_{m+1,m}$$

- Subspace spanned by $v_1, v_2, \ldots, v_m$ is not a Krylov subspace, but V_m orthogonal (in the full case)

Theorem for Inexact FOM

[Simoncini and Szyld, 2003]

True residual: $r_m = b - Hx_m = r_0 - HV_m y_m$

Computed residual(e.g.): $\tilde{r}_m = r_0 - V_{m+1} H_{m+1,m} y_m = r_0 - W_m y_m$

Let $\varepsilon > 0$. If for every $k \le m$,

$$\|D_k\| \le \frac{\sigma_{min}(H_{m_*})}{m_*} \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon \equiv \ell_m^F \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon \ ,$$

then $\|V_m^T r_m\| \le \varepsilon$ and $\|r_m - \tilde{r}_m\| \le \varepsilon$.

m_* being the maximum number of iterations allowed

Similar results for inexact GMRES

see also [Giraud, Gratton, Langou, 2007]

Theorem for Inexact Truncated FOM

$$\|D_k\| \le \frac{\sigma_{min}(H_{m_*})\sigma_{min}(V_m)}{m_*} \frac{1}{\|\tilde{r}_{k-1}\|}\varepsilon \equiv \ell_m^{TF} \frac{1}{\|\tilde{r}_{k-1}\|}\varepsilon \ ,$$

implies $\|V_m^T r_m\| \le \varepsilon$ and $\delta_m = \|r_m - \tilde{r}_m\| \le \varepsilon$.

Notes:

- This result applies in particular to Inexact CG
- ℓ_m can be estimated from problem, if information is available.

First Experiment

$H = \text{diag}([10^{-4}, 2, 3, \cdots, 100])$ $D_k = \text{symm}\ [\alpha_k \text{randn}(100, 100)]$
$b = \text{randn}(100, 1)$ We chose $\varepsilon = 10^{-8}$

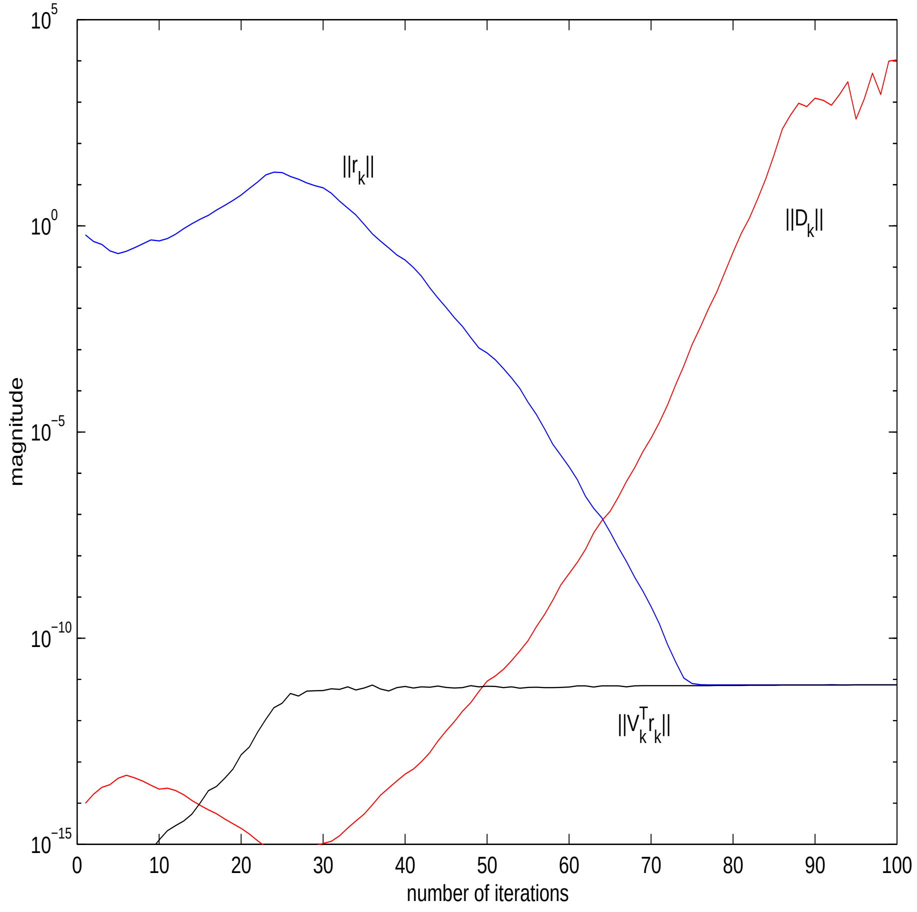
- Our condition (e.g. for FOM)

$$\|D_k\| \leq \frac{\sigma_{min}(H)}{m_*} \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon$$

is very conservative. In most cases it is too strict.

However, $\sigma_{min}(H)$ does play a role.

CG: condition $\|D_k\| \leq \frac{\sigma_{min}(H)}{m_*} \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon$



$\|V_m^T r_m\| \ll \varepsilon$

Applications:
I. Schur complement systems

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix},$$

$$\begin{aligned} B^T A^{-1} B x &= B^T A^{-1} f; \qquad Aw = f - Bx \\ Hx &= b \end{aligned}$$

A^{-1} not exactly (use Krylov method).

Applications: I. Schur complement systems (cont.)

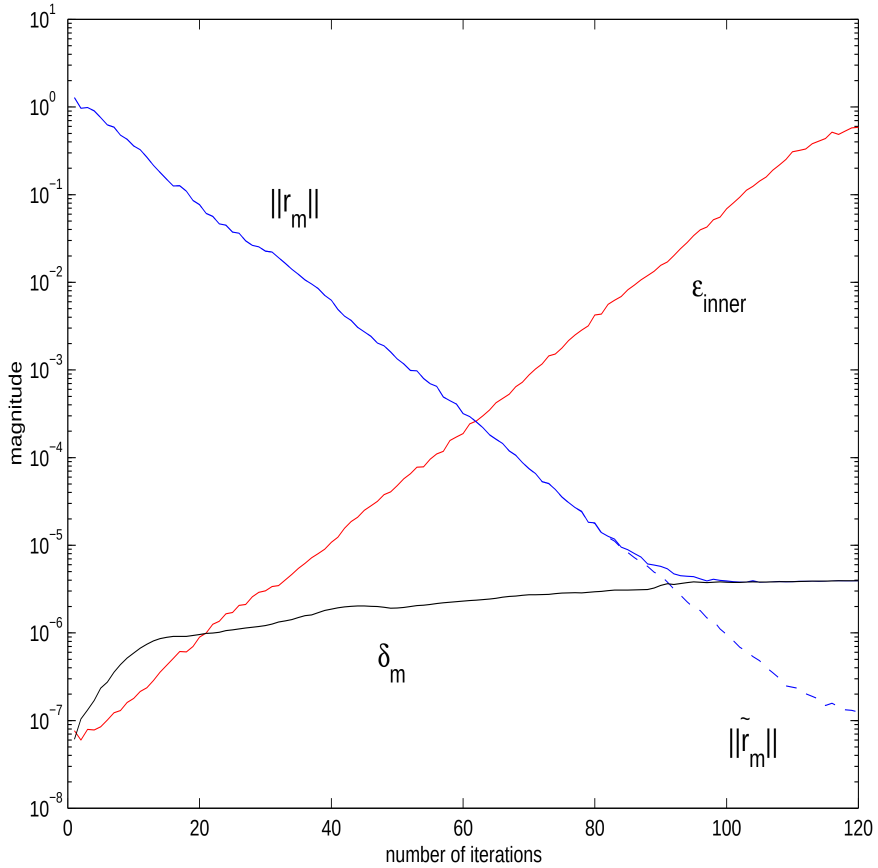
- A^{-1} not exactly (use Krylov method).
- Replace Hv with $\mathcal{H}v := B^T z_j^{(k)}$, where $z_j^{(k)}$ is the approximation obtained at the jth (inner) iteration of the solution to the equation

$$Az = Bv$$

- Question is then: How many inner iterations?
  i.e., at what value of j stop?
  "Translate" conditions on $\|D_k\|$ to conditions on norm of inner residual.

Let $r_k^{inner} = Az_j^{(k)} - Bv$ be the inner residual

Take $$\|r_k^{inner}\| < \frac{\sigma_{m_\star}(H_{m_\star})}{\|B^T A^{-1}\| m_\star} \frac{1}{\|\tilde{r}_{k-1}^{fom}\|}\varepsilon \equiv \varepsilon_{\text{inner}}$$



- Two-dim. saddle point magnetostatic problem from [Perugia, Simoncini, Arioli, 1999], A is 1272×1272
- Inexact FOM, $m_\star = 120$, $\varepsilon = 10^{-4}$

Applications:
II. Inexact Preconditioning

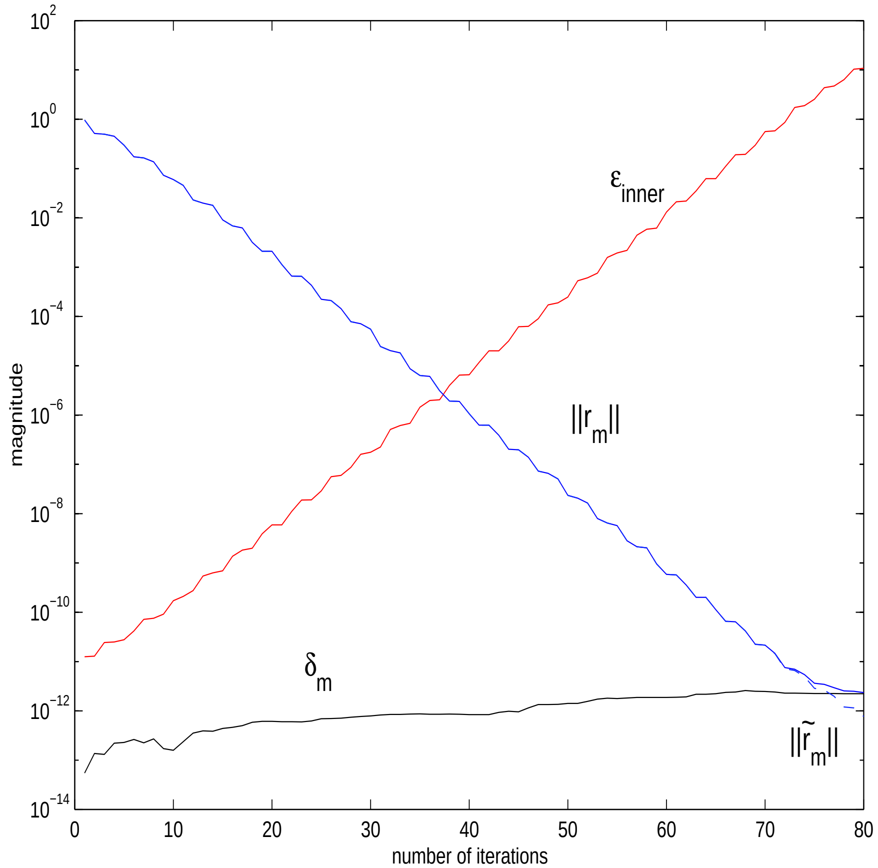
$$Hx = b \quad \longrightarrow \quad H\mathcal{P}^{-1}\bar{x} = b, \quad x = \mathcal{P}^{-1}\bar{x}$$

$\mathcal{P}^{-1}$ not performed exactly (use Krylov method)
$H\mathcal{P}^{-1}v_k$ replaced with $H\tilde{z}_k, \quad \tilde{z}_k \approx \mathcal{P}^{-1}v_k$
Arnoldi relation $\qquad H\mathcal{P}^{-1}V_m = V_{m+1}H_{m+1,m} \qquad$ is transformed into

$$H[\tilde{z}_1, \cdots, \tilde{z}_m] = V_{m+1}H_{m+1,m}.$$

Use Flexible Krylov subspace method
$r_k^{inner} = v_k - \mathcal{P}\tilde{z}_k$ inner residual

$$\|r_k^{inner}\| \le \frac{\sigma_{m_\star}(H_{m_\star})}{\|H\mathcal{P}^{-1}\| m_\star} \frac{1}{\|\tilde{r}_{k-1}^{gm}\|} \varepsilon \equiv \varepsilon_{inner}$$



For same 2D saddle point, use $\mathcal{P} = \begin{bmatrix} I & 0 \\ 0 & B^T B \end{bmatrix}$. Solve $B^T B p_k = rhs$ iteratively, $m_\star = 80$, $\varepsilon = 10^{-9}$, tolerance ε_{inner}

Some CPU Times: Same Magnetostatic 2D Problem

Outer tolerance: $\varepsilon = 10^{-8}$

$$\|r_k^{inner}\| \leq \frac{c_0}{\|r_{m-1}^{outer}\|}\varepsilon \equiv \varepsilon_{inner}$$

c_0: Constant estimated a–priori: Here we use 10^{-2} and 10^{-4}.

Elapsed Time

CPU in seconds of a Sun Enterprise 4500 (Fortran code)
(4 CPU 400MHertz, 2GBytes RAM) CG iterations.

Problem Size	Fixed Inner Tol	Var. Inner Tol.	Var. Inner Tol.
	$\varepsilon_{inner} = 10^{-10}$	$10^{-10}/\|r\|$	$10^{-12}/\|r\|$
3810	17.0 (54)	11.4 (54)	14.7 (54)
9102	82.9 (58)	62.8 (58)	70.7 (58)
14880	198.4 (54)	156.5 (54)	170.1 (54)

Applications:
III. Parabolic Control Problems

Inverse problem: Recover control $v(x)$ based on field (state) $z(x)$ related by the forward problem

$$z_t + \mathcal{A}z = v, \qquad x \epsilon \Omega$$
$$z = g, \qquad x \epsilon \partial\Omega$$
$$z = z_0, \qquad x \epsilon \Omega / \partial\Omega, \qquad \text{for } \; t = 0$$

$\mathcal{A}$ elliptic, e.g., $\mathcal{A} = -\triangle$

This is a distributed control problem. Similar techniques for boundary control problems (control g).

Associated variational problem

$$J(z(v),v) := \frac{\alpha}{2}\int_{t_0}^{t_f} \|z(v)(t,\cdot)-\tilde{y}(t,\cdot)\|^2_{L^2(\Omega)}$$
$$+ \frac{\beta}{2}\|z(v)(t_f,\cdot)-\tilde{y}(t_f,\cdot)\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\int_{t_0}^{t_f}\|v(t,\cdot)\|^2_{L^2(\Omega)}.$$

discretized as

$$J_h^\tau(\mathbf{z},\mathbf{v}) = \frac{1}{2}(\mathbf{z}-\tilde{\mathbf{y}})^T\mathbf{K}(\mathbf{z}-\tilde{\mathbf{y}}) + \frac{1}{2}\mathbf{v}^T\mathbf{G}\mathbf{v} + (\mathbf{z}-\tilde{\mathbf{y}})^T\mathbf{g}.$$

Discretized forward problem (FD)

$$\mathbf{E}\mathbf{z} + \mathbf{N}\mathbf{v} = \mathbf{f}$$

$$\mathbf{E} = \begin{bmatrix} F_1 & & & \\ -F_0 & F_1 & & \\ & \ddots & \ddots & \\ & & -F_0 & F_1 \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix}$$

Optimization problem

$$\begin{aligned} \min \quad & J_h^\tau(\mathbf{z}, \mathbf{v}) \\ \text{subject to} \quad & \mathbf{E}\mathbf{z} + \mathbf{N}\mathbf{v} = \mathbf{f} \end{aligned}$$

Lagrangian:

$$J_h^\tau(\mathbf{z}, \mathbf{v}) + \mathbf{q}^T(\mathbf{E}\mathbf{z} + \mathbf{N}\mathbf{v} - \mathbf{f})$$

Linearize to obtain

$$\begin{bmatrix} \mathbf{K} & \mathbf{0} & \mathbf{E}^T \\ \mathbf{0} & \mathbf{G} & \mathbf{N}^T \\ \mathbf{E} & \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \\ \mathbf{f} \end{bmatrix}$$

After elimination one obtains

$$\mathbf{Hu} := (\mathbf{G} + \mathbf{N}^T \mathbf{E}^{-T} \mathbf{K} \mathbf{E}^{-1} \mathbf{N})\mathbf{u} = \mathbf{b}$$

H being the spd reduced Hessian

$$\mathbf{H}\mathbf{u} = (\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} = \mathbf{b}$$

We use inexact FOM, approximating each of the the systems with E and E^T with a Parareal method with varying (increasing) tolerance.

MVP $\mathbf{H}\mathbf{v}$

1. Multiply $\mathbf{N}\mathbf{v}$
2. Solve $\mathbf{E}\mathbf{z} = \mathbf{N}\mathbf{v}$ by solving $\mathbf{E}\mathbf{z} = \mathbf{N}\mathbf{v}$ with an inner tolerance ϵ_{in_1}
3. Multiply $\mathbf{K}\mathbf{z}$
4. Solve $\mathbf{E}^T\mathbf{w} = \mathbf{K}\mathbf{z}$ by solving with an inner tolerance ϵ_{in_2}
5. Compute $\mathbf{N}^T\mathbf{w}$

Approximate solutions of systems with $\mathbf{E}$ or $\mathbf{E}^T$

$$\mathbf{Hu} = (\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} = \mathbf{b}$$

We prove that with $\epsilon_{in_1} = \epsilon_{in_2}$, we have spectral equivalence of the form

$$(\mathbf{v}, \mathbf{Gv}) \leq (\mathbf{v}, \mathbf{Hv}) \leq \mu(\mathbf{v}, \mathbf{Gv})$$

We choose FOM, since it reduces to CG when we have full symmetry.

Sample Experiment: 2D heat equation

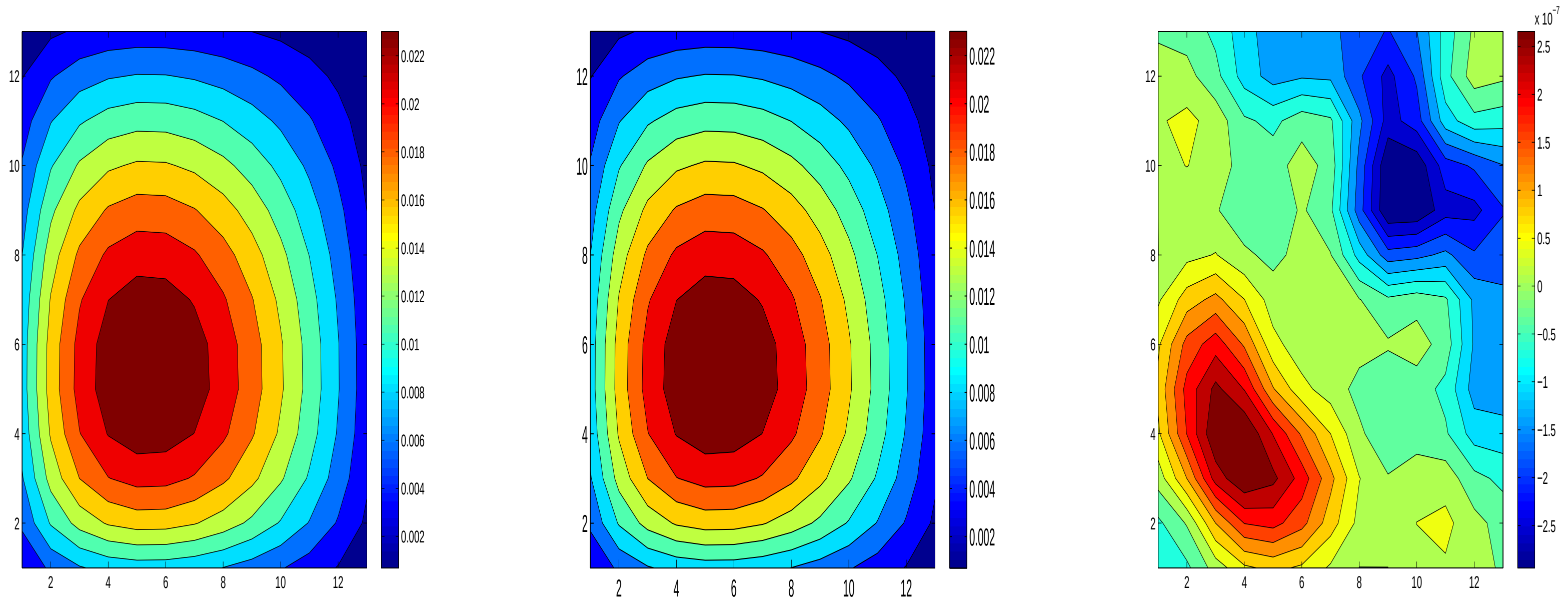
15×15 grid. $\tau = 1/512$, control $\mathbf{u}$ of order 115200

Same relative tolerance for both Parareal systems, outer $\varepsilon = 10^{-6}$

$\ell_m^{(i)} \epsilon$	IFOM	TIFOM			
		$m_T = 2$	$m_T = 4$	$m_T = 8$	$m_T = 12$
10^{-12}	15(576)	16(610)	16(608)	15(576)	15(576)
10^{-10}	15(482)	17(532)	17(528)	16(504)	15(482)
10^{-8}	15(388)	17(426)	17(426)	17(420)	15(388)
10^{-7}	15(340)	18(394)	17(374)	17(368)	15(340)
10^{-6}	15(288)	19(340)	19(338)	18(320)	16(298)
10^{-5}	17(238)	24(298)	21(266)	19(258)	19(242)
10^{-4}	17(180)	n.c.	22(210)	28(254)	20(192)

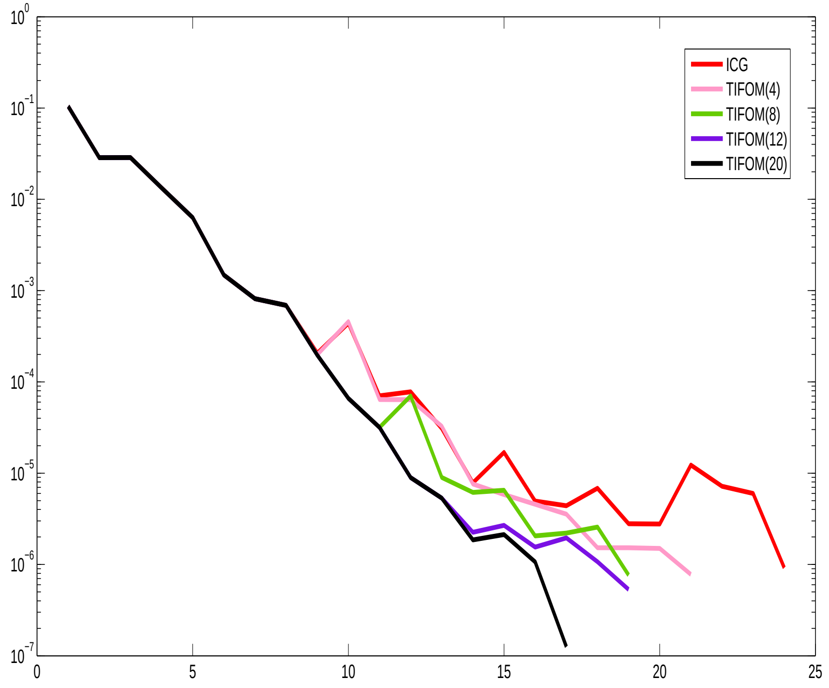
One surface of true and recovered model, and their difference

increasing $\varepsilon_{inner} = 10^{-5}/\|\tilde{r}_{k-1}\|$, $m_T = 8$



True　　Computed　　error O(10^{-7})

TIFOM Convergence curves

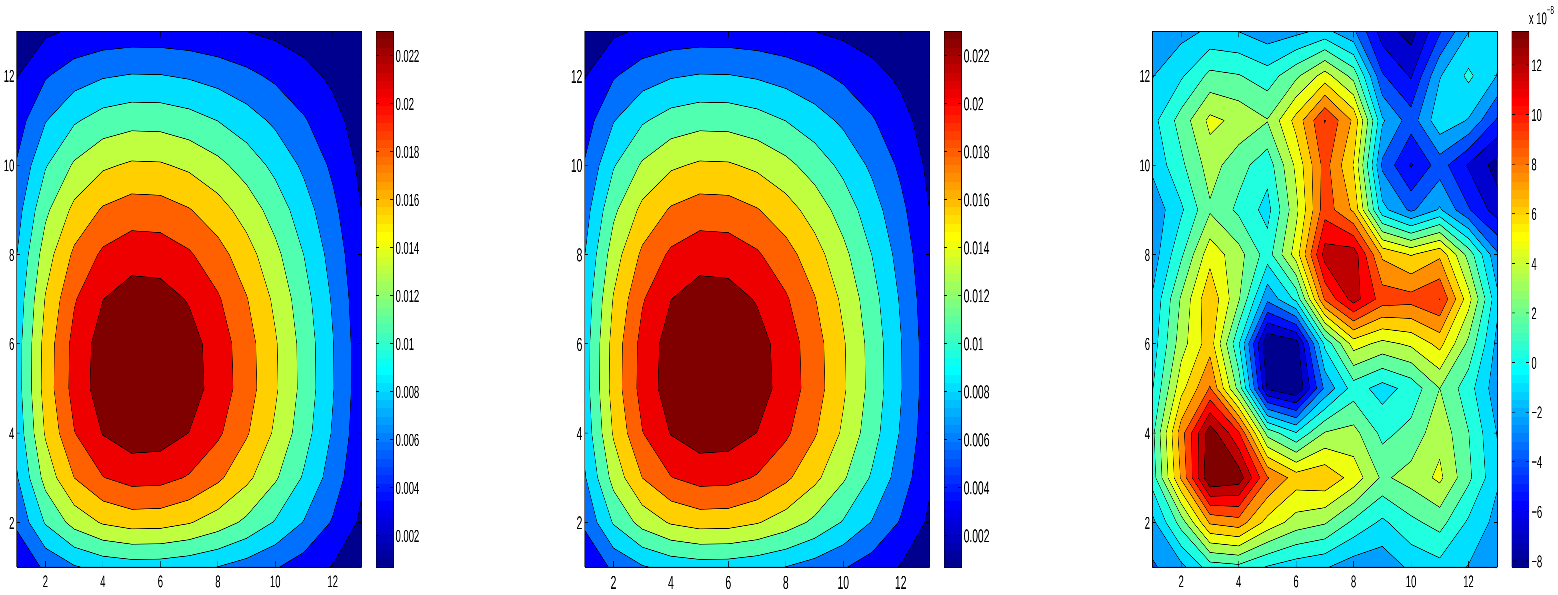


outer $\varepsilon = 10^{-6}$, inner factor $= 10^{-5}$

Similar results for more non-symmetric

$\ell_m^{(1)}\varepsilon$	$\ell_m^{(2)}\varepsilon$	IFOM	TIFOM o-iter. (i-iter.)			
		o-iter. (i-iter)	$m_T = 2$	$m_T = 4$	$m_T = 8$	$m_T = 12$
10^{-6}	10^{-7}	15(288)	18(330)	19(338)	17(310)	16(298)
10^{-6}	10^{-6}	15(288)	19(340)	19(338)	18(320)	16(298)
10^{-6}	10^{-5}	16(300)	23(390)	20(346)	19(334)	17(310)
10^{-6}	10^{-4}	16(300)	51(710)	37(520)	20(352)	17(310)
10^{-5}	10^{-7}	15(234)	19(270)	27(280)	18(254)	21(246)
10^{-5}	10^{-6}	15(234)	19(270)	29(284)	18(254)	19(242)
10^{-5}	10^{-5}	17(238)	24(298)	21(266)	19(258)	19(242)
10^{-5}	10^{-4}	17(240)	30(334)	31(356)	20(270)	19(244)

One surface of true and recovered model, and their difference, outer tolerance $\varepsilon = 10^{-6}$ $\varepsilon_{in1} = 10^{-6}/\|\tilde{r}_{k-1}\|$, $\varepsilon_{in4} = 10^{-4}/\|\tilde{r}_{k-1}\|$, $m_T = 8$



True | Computed | error $O(10^{-6})$

Conclusions

- Inexact matrix-vector product (or inexact preconditioning) might be worth trying for your problem
- Truncated methods might be worth trying for your problem
- New work on inexact and truncated for parabolic control problems (with two inner criteria)

With Valeria Simoncini:

Theory of Inexact Krylov Subspace Methods and
Applications to Scientific Computing
SIAM J. Scientific Computing, v. 25 (2003) 454–477.

On the Occurrence of Superlinear Convergence of Exact
and Inexact Krylov Subspace Methods
SIAM Review, v. 47 (2005) 247–272.

The Effect of Non-Optimal Bases on the Convergence of Krylov
Subspace Methods
Numerische Mathematik, v. 100 (2005) 711-733.

Recent computational developments
in Krylov Subspace Methods for linear systems
Numerical Linear Algebra with Applications, v. 14 (2007) 1-59.

All available at: `http://www.math.temple.edu/szyld`

With Xiuhond Du, Marcus Sarkis and Christian E. Schaerer

Inexact and truncated Parareal-in-time Krylov subspace methods for parabolic optimal control problems
Research Report 12-02-06, February 2012

Also available at: `http://www.math.temple.edu/szyld`