The Extended Krylov Subspace for Matrix Function Approximations: Analysis and Applications

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

Partially joint work with Leonid Knizhnerman, Moscow

The Problem

Given $A \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^n$ and f sufficiently smooth function, approximate

$$x = f(A)v$$

$\star$ A large dimensions, $\|v\| = 1$

Applications:

- Numerical solution of evolution PDEs (e.g. $\exp(\lambda), \sqrt{\lambda^{-1}}, \cos(\lambda), ...$)
- Inverse Problems $(\exp(\lambda), \cosh(\lambda), ...)$
- Fluxes on manifolds
- Problems in Scientific Computing (e.g. QCD, $\mathrm{sign}(\lambda)$)
- (Analysis of) reduced Dynamical System Models (via Gramians)

The Problem

Given $A \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^n$ and f sufficiently smooth function, approximate

$$x = f(A)v$$

$\star$ A large dimensions, $\|v\| = 1$

Projection-type methods

$\mathcal{K}$ approximation space, $m = \dim(\mathcal{K})$

$V \in \mathbb{R}^{n\times m}$ s.t. $\mathcal{K} = \mathrm{range}(V)$

$$x = f(A)v \quad \approx \quad x_m = Vf(V^\top AV)(V^\top v)$$

The Problem

Given $A \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^n$ and f sufficiently smooth function, approximate

$$x = f(A)v$$

$\star$ A large dimensions, $\|v\| = 1$

Projection-type methods

$\mathcal{K}$ approximation space, $m = \dim(\mathcal{K})$

$V \in \mathbb{R}^{n\times m}$ s.t. $\mathcal{K} = \mathrm{range}(V)$

$$x = f(A)v \quad \approx \quad x_m = Vf(V^\top AV)(V^\top v)$$

Question: Which $\mathcal{K}$?

Some explored alternatives for $\mathcal{K}$

- Krylov subspace, $\mathcal{K} = K_m(A, v)$
- Shift-Invert Krylov subspace, $\mathcal{K} = K_m((I + \gamma A)^{-1}, v)$ for some γ
- Rational Krylov subspace, for some $\omega_1, \omega_2, \ldots$
  $\mathcal{K} = \mathrm{span}\{v, (A - \omega_1 I)^{-1} v, (A - \omega_2 I)^{-1} v, \ldots\}$
- Extended Krylov subspace, $\mathcal{K} = K_m(A, v) + K_m(A^{-1}, A^{-1} v)$
- Restarted Krylov subspace

Note: In all cases, A nonsymmetric.

Theory mostly for field of values of A in $\mathbb{C}^+$

Field of values: $W(A) = \{x^* A x,\ x \in \mathbb{C}^n, \|x\| = 1\}$

Krylov subspace approximation

"Classical" approach:

$$\mathcal{K} = K_m(A, v) = \mathrm{span}\{v, Av, \ldots, A^{m-1}v\}$$

For $H_m = V_m^\top A V_m$, $v = V_m e_1$ and $V_m^\top V_m = I_m$:

$$x_m = V_m f(H_m) e_1$$

Polynomial approximation: $x_m = p_{m-1}(A)v$
(p_{m-1} interpolates f at eigenvalues of H_m)

⋆ Numerical and theoretical results since mid '80s (Saad '92)

Acceleration Procedures: Shift-Invert Krylov

Choose γ s.t. $(I + \gamma A)$ is invertible, and construct

$\mathcal{K} = K_m((I + \gamma A)^{-1}, v)$, Moret-Novati '04, van den Eshof-Hochbruck '06

with $T_m = V_m^\top (I + \gamma A)^{-1} V_m$, $v = V_m e_1$ and $V_m^\top V_m = I_m$

$$x_m = V_m f(\frac{1}{\gamma}(T_m^{-1} - I_m))e_1$$

Rational approximation: $x_m = p_{m-1}((I + \gamma A)^{-1})v$

Choice of γ: A spd, $\gamma = \frac{1}{\sqrt{\lambda_{\min}\lambda_{\max}}}$ (Moret, 2009)

A nonsym, (Beckermann-Reichel tr2008)

Acceleration Procedures: Extended Krylov

For A nonsingular,

$\mathcal{K} = K_{m_1}(A, v) + K_{m_2}(A^{-1}, A^{-1}v),$ Druskin-Knizhnerman 1998, A sym.

Note: $\mathcal{K} = A^{-m_2} K_{m_1+m_2}(A, v)$

Algorithm (augmentation-style)

- Fix $m_2 \ll m_1$
- Run m_2 steps of Inverted Lanczos
- Run m_1 steps of Standard Lanczos + orth.

Extended Krylov: an effective implementation

$m_1 = m_2 = m$ not fixed a priori

$$\begin{aligned}\mathcal{K} &= K_m(A, v) + K_m(A^{-1}, A^{-1}v) \\ &= \text{span}\{v, A^{-1}v, Av, A^{-2}v, A^2v, \ldots\}\end{aligned}$$

⋆ *Block* Arnoldi-type recurrence:

- $U_1 \leftarrow \text{orth}([v, A^{-1}v])$
- $U_{j+1} \leftarrow [AU_j(:,1), A^{-1}U_j(:,2)] + \text{orth} \qquad j = 1, 2, \ldots$

⋆ Recurrence to cheaply compute $\mathcal{T}_m = \mathcal{U}_m^\top A \mathcal{U}_m, \quad \mathcal{U}_m = [U_1, \ldots, U_m]$

⋆ Compute $x_m = \mathcal{U}_m f(\mathcal{T}_m) e_1$

Simoncini, 2007

Extended Krylov: Convergence theory I

f satisfying $$f(z) = \int_{-\infty}^{0} \frac{1}{z-\zeta} d\mu(\zeta), \qquad z \in \mathbb{C} \backslash]-\infty, 0]$$

(with convenient measure $d\mu(\zeta)$)

Extended Krylov: Convergence theory II

Nonsingular A, with $0 \notin W(A)$.

Let Φ_1, Φ_2 be the conformal maps for $W(A)$ and $W(A)^{-1}$

There exists $a > 0$ s.t. $|\Phi_1(-a)| = |\Phi_2(-\frac{1}{a})|$ so that

$$\|f(A)v - \mathcal{U}_m f(\mathcal{T}_m)e_1\| \leq \frac{c}{|\Phi_1(-a)|^m}$$

Extended Krylov: Convergence theory II

Nonsingular A, with $0 \notin W(A)$

Let Φ_1, Φ_2 be the conformal maps for $W(A)$ and $W(A)^{-1}$

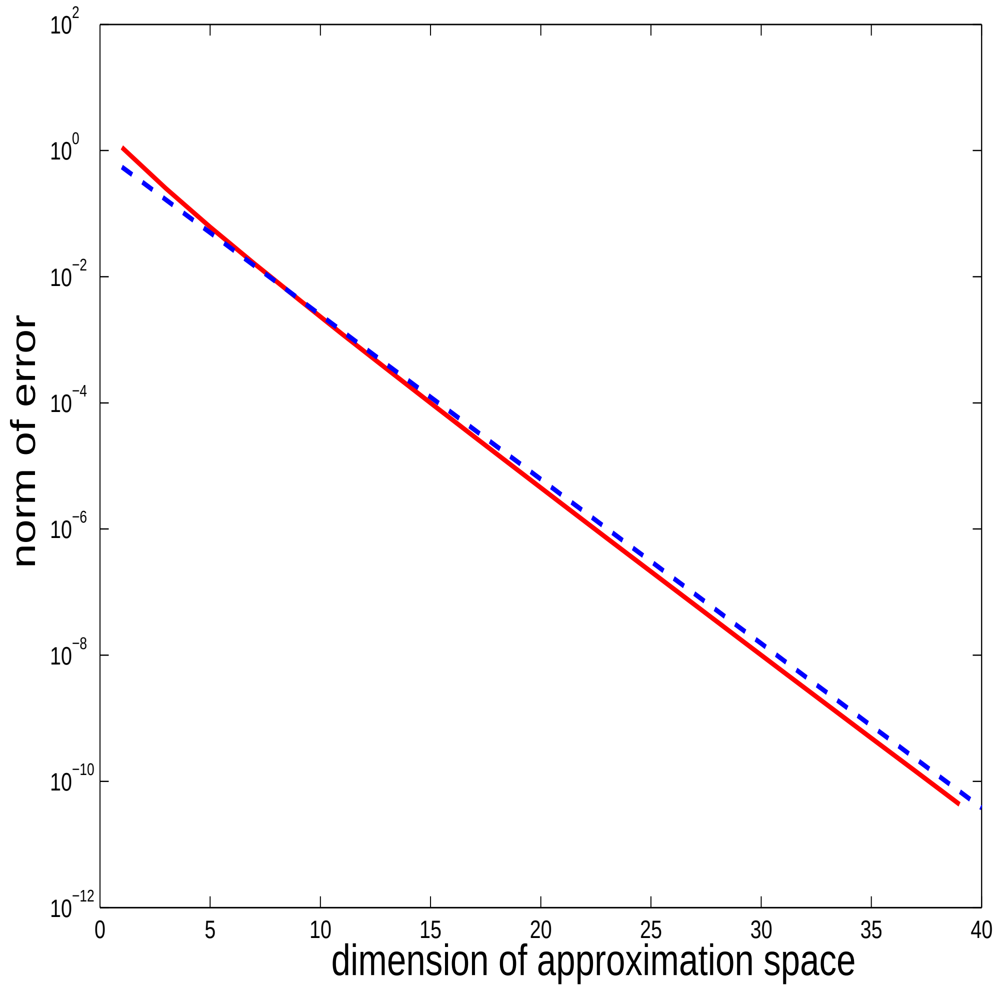
There exists $a > 0$ s.t. $|\Phi_1(-a)| = |\Phi_2(-\frac{1}{a})|$ so that

$$\|f(A)v - \mathcal{U}_m f(\mathcal{T}_m)e_1\| \le \frac{c}{|\Phi_1(-a)|^m}$$

e.g. for A symmetric (Φ_1, Φ_2 known, $a = \sqrt{\lambda_{\min}\lambda_{\max}}$) :

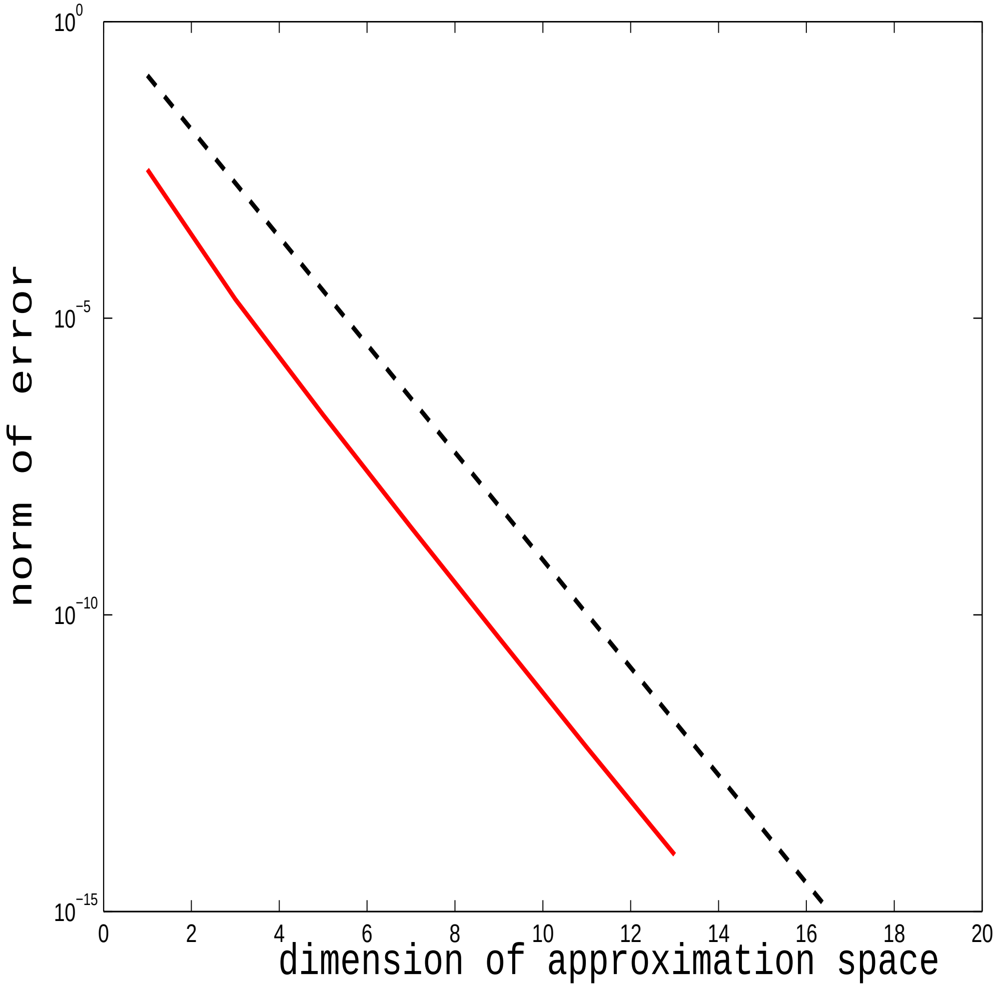
$\|x - x_m\| = O(\exp(-2m\sqrt[4]{\frac{\lambda_{\min}}{\lambda_{\max}}}))$

Convergence rate. $A \in \mathbb{R}^{400\times 400}$ normal. $f(\lambda) = \lambda^{-1/2}$



$\sigma(A)$ on an elliptic curve in $\mathbb{C}^+$ with center on real axis

Rate. $A \in \mathbb{R}^{200\times 200}$ Jordan block, $\sigma(A) = \{4\}$. $f(\lambda) = \lambda^{1/2}$



$W(A)$ disk centered at 4 and unit radius

Large-scale numerical experiments

A from FD discretization of

$$\begin{aligned}
\mathcal{L}_1(u) &= -100u_{x_1x_1} - u_{x_2x_2} + 10x_1u_{x_1}, \\
\mathcal{L}_2(u) &= -100u_{x_1x_1} - u_{x_2x_2} - u_{x_3x_3} + 10x_1u_{x_1}, \\
\mathcal{L}_3(u) &= -e^{-x_1x_2}u_{x_1x_1} - e^{x_1x_2}u_{x_2x_2} + \frac{1}{x_1+x_2}u_{x_1}, \\
\mathcal{L}_4(u) &= -\mathrm{div}(e^{3x_1x_2}\mathrm{grad}u) + \frac{1}{x_1+x_2}u_{x_1}
\end{aligned}$$

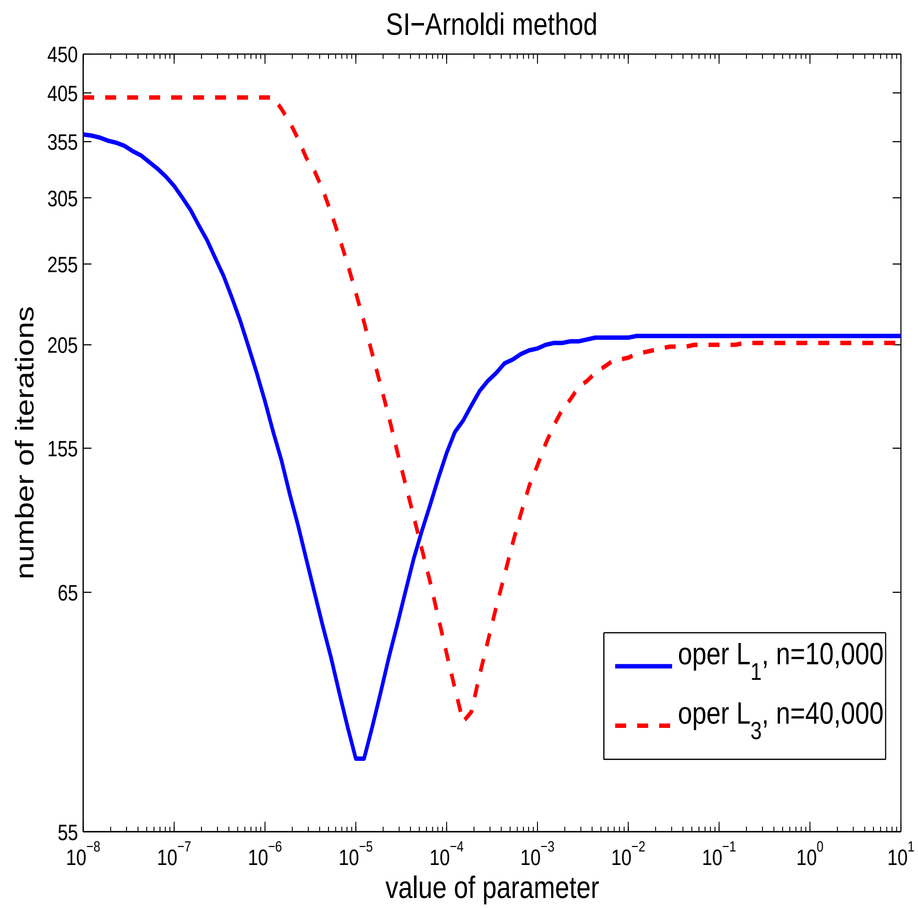
on unit square/cube, Dirichlet hom. bc.

Inner system solves:

- Extended Krylov: systems with A solved with GMRES/AMG
- SI-Arnoldi: systems with $I + \gamma A$ solved with IDR(s)/ILU

An intermezzo

SI-Arnoldi requires getting the parameter γ:



Number of SI-Arnoldi iterations as a function of the parameter for $f(\lambda) = \lambda^{1/2}$

Comparisons: CPU Time in Matlab (space dim.)

f	Oper.	n	SI-Arnoldi	EKSM	Std Krylov
$\lambda^{1/2}$	$\mathcal{L}_1$	2500	0.9 (*59*)	0.6 (*48*)	7 (*193*)
		10000	4.0 (*66*)	3.6 (*68*)	*46 (*300*)
		160000	642.9(*246*)	219.7(*122*)	*458(*300*)
	$\mathcal{L}_2$	27000	10.8 (*55*)	7.4 (*40*)	6.7(*119*)
		125000	86.7 (*60*)	65.3 (*52*)	138.7(*196*)
	$\mathcal{L}_3$	40000	26.3 (*75*)	21.1 (*72*)	*87 (*300*)
		160000	318.5(*144*)	173.3 (*96*)	*442(*300*)
	$\mathcal{L}_4$	40000	41.1(*117*)	25.4(*106*)	*89 (*300*)
		160000	580.2(*442*)	231.2(*144*)	*461 (*300*)

Comparisons: CPU Time in Matlab (space dim.)

f	Oper.	n	SI-Arnoldi	EKSM	Std Krylov
$\lambda^{-1/3}$	$\mathcal{L}_1$	2500	0.6 (*43*)	0.4 (*30*)	2.2(*131*)
		10000	2.6 (*46*)	1.8 (*38*)	26.2(*252*)
		160000	79.3 (*48*)	99.7 (*64*)	*460(*300*)
	$\mathcal{L}_2$	27000	7.8 (*41*)	4.8 (*26*)	3.1 (*82*)
		125000	64.8 (*45*)	38.9 (*32*)	67.5(*138*)
	$\mathcal{L}_3$	40000	20.7 (*61*)	13.7 (*48*)	*88 (*300*)
		160000	116.5 (*62*)	105.2 (*62*)	*460 (*300*)
	$\mathcal{L}_4$	40000	35.8(*104*)	14.2 (*66*)	*88 (*300*)
		160000	208.1(*104*)	112.2 (*84*)	*461 (*300*)

Stopping criterion

Unlike linear systems: no equation $\Rightarrow$ no residual

Estimate of the error:

(first suggested for $f(\lambda) = e^{-\lambda}$ by van den Eshof-Hochbruck '06)

$$\frac{\|x - x_m\|}{\|x_m\|} \approx \frac{\delta_{m+j}}{1 - \delta_{m+j}}, \qquad \delta_{m+j} = \frac{\|x_{m+j} - x_m\|}{\|x_m\|}$$

Stopping criterion:

$$\texttt{if } \frac{\delta_{m+j}}{1-\delta_{m+j}} \leq \texttt{tol then stop}$$

Computational costs awareness: inexact solves in EKSM

systems with A: GMRES with relaxed inner tolerance

$$\epsilon_m^{(\text{inner})} = \frac{\texttt{tolin}}{\|x - x_{m-1}\|}.$$

Final outer error (# outer its / # inner its)

tolin	fixed inner tol	relaxed inner tol
1e-10	6.97e-11 (24/901)	6.58e-11 (24/559)
1e-12	6.48e-11 (24/1052)	6.48e-11 (24/716)

$$\mathcal{L}(u) = -u_{xx} - u_{yy} - u_{zz} + 50(x+y)u_x$$

$f(\lambda) = \lambda^{-1/3}$ $\qquad \epsilon^{(\text{outer})} = 10^{-10}$

Conclusions

- Efficient generation of the Extended Krylov subspace
- Complete theory for EKSM for a large class of functions
- Performance:

  - Competitive with respect to available methods

  (when solving with A can be made cheap)

  - Does not depend on parameters

  - Projection-type method: wide applicability (work in progress)

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