Computational aspects in large scale matrix function approximations

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

The Problem

Given $A \in \mathbb{C}^{n\times n}$ and a sufficiently regular function f, numerically compute or approximate

$$f(A)$$

or, as required in many large-scale applications

$$f(A)v$$

for a given vector v.

The Problem

Given $A \in \mathbb{C}^{n\times n}$ and a sufficiently regular function f, numerically compute or approximate

$$f(A)$$

or, as required in many large-scale applications

$$f(A)v$$

for a given vector v.

Other "matrix" functions:

- Element-wise: $A = (a_{ij})$ then $F = (f(a_{ij}))$
- Matrix polynomials: $P_k(\lambda) = \sum_{i=0}^{k} \lambda^i A_i, \quad A_i \in \mathbb{C}^{n\times n}$
- Functions of matrices: det, norm, trace, $u^T f(A) u$, etc.
- ...

Basic definitions

$A = XJX^{-1}$ Jordan decomposition, with $J = \mathrm{diag}(J_1, \cdots, J_k)$

Assume f is *defined* on the spectrum of A, that is

$$f^{(j)}(\lambda_i) \quad \text{exists}$$

for all i, and j up to the largest Jordan block size of λ_i minus 1

Basic definitions

$A = XJX^{-1}$ Jordan decomposition, with $J = \mathrm{diag}(J_1, \cdots, J_k)$

Assume f is *defined* on the spectrum of A, that is

$$f^{(j)}(\lambda_i) \quad \text{exists}$$

for all i, and j up to the largest Jordan block size of λ_i minus 1

Then $f(A) := X \mathrm{diag}(f(J_1), \ldots, f(J_k)) X^{-1}$ with

$$f(J_\ell) := \begin{bmatrix} f(\lambda_\ell) & f'(\lambda_\ell) & \ldots & \frac{f^{(n_\ell-1)}(\lambda_\ell)}{(n_\ell-1)!} \\ & f(\lambda_\ell) & \ddots & \vdots \\ & & \ddots & f'(\lambda_\ell) \\ & & & f(\lambda_\ell) \end{bmatrix}$$

Basic definitions

$A = XJX^{-1}$ Jordan decomposition, with $J = \text{diag}(J_1, \cdots, J_k)$

Assume f is *defined* on the spectrum of A, that is

$$f^{(j)}(\lambda_i) \quad \text{exists}$$

for all i, and j up to the largest Jordan block size of λ_i minus 1

Then $f(A) := X\text{diag}(f(J_1), \ldots, f(J_k))X^{-1}$ with

$$f(J_\ell) := \begin{bmatrix} f(\lambda_\ell) & f'(\lambda_\ell) & \ldots & \frac{f^{(n_\ell-1)}(\lambda_\ell)}{(n_\ell-1)!} \\ & f(\lambda_\ell) & \ddots & \vdots \\ & & \ddots & f'(\lambda_\ell) \\ & & & f(\lambda_\ell) \end{bmatrix}$$

$\Rightarrow$ No other information on f required

$\Rightarrow$ Any other function g s.t. $g^{(j)}(\lambda_i) = f^{(j)}(\lambda_i)$ will do !

(e.g., (Hermite) interpolating polynomials)

Basic definitions

Cauchy integral representation:

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A)^{-1} \mathrm{dz}, \qquad \Gamma = \partial\Omega, \quad \sigma(A) \subset \Omega$$

where f is analytic in $\overline{\Omega}$

Basic definitions

Cauchy integral representation:

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A)^{-1} \mathrm{dz}, \qquad \Gamma = \partial\Omega, \quad \sigma(A) \subset \Omega$$

where f is analytic in $\overline{\Omega}$

- Possible starting point for approximations by contour integration
- Key formula for theoretical purposes
- Generalizes to operators

Numerical computation and approximation of $f(A)$

- Polynomial approximation: $f(A) \approx p_k(A)$

  (e.g., from Taylor expansion)

Numerical computation and approximation of $f(A)$

- Polynomial approximation: $f(A) \approx p_k(A)$

  (e.g., from Taylor expansion)

- Rational approximation: $f(A) \approx \frac{p_k(A)}{q_m(A)}$

  (e.g., Padè approximations)

  Computation: via continued fractions, or partial fraction expansion

Numerical computation and approximation of $f(A)$

- Polynomial approximation: $f(A) \approx p_k(A)$

  (e.g., from Taylor expansion)

- Rational approximation: $f(A) \approx \frac{p_k(A)}{q_m(A)}$

  (e.g., Padè approximations)

  Computation: via continued fractions, or partial fraction expansion

- Diagonalization and Schur decomposition-based methods

  Computation: $f(A) = Qf(R)Q^*$ and then use $Rf(R) = f(R)R$ to compute $f(R)$ (Parlett, '76)

Numerical computation and approximation of $f(A)$

- Polynomial approximation: $f(A) \approx p_k(A)$

  (e.g., from Taylor expansion)

- Rational approximation: $f(A) \approx \frac{p_k(A)}{q_m(A)}$

  (e.g., Padè approximations)

  Computation: via continued fractions, or partial fraction expansion

- Diagonalization and Schur decomposition-based methods

  Computation: $f(A) = Qf(R)Q^*$ and then use $Rf(R) = f(R)R$ to compute $f(R)$ (Parlett, '76)

- Matrix iterations: $X_{k+1} = \Phi(X_k)$, with $X_k \rightarrow_{k\rightarrow\infty} f(A)$

  Typically for: $\mathrm{sgn}(\lambda), \sqrt{\lambda}, \sqrt[p]{\lambda}, ...$

Higham's book, '08

Numerical computation and approximation of $f(A)v$

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

$$x = f(A)\,v$$

with f regular function such that $f(A)$ is well defined

$$f(A) \qquad \text{vs.} \qquad f(A)v$$

Focus:

- A large dimension
- A symmetric pos. (semi)def., or A *positive real*

Applications

- Numerical solution of time-dependent PDEs
  (e.g. $\exp(\lambda), \sqrt{\lambda^{-1}}, \cos(\lambda), \varphi_k(\lambda)...$)
- Scientific Computing problems (e.g. QCD, $\mathrm{sign}(\lambda)$)
- (Analysis of) reduced Dynamical System Models
  (e.g., through Grammian Matrices)
- Numerical solution of some Inverse Problems ($\exp(\lambda), \cosh(\lambda)$, ...)
- Fluxes on manifolds
- ...

Numerical approximation. I

$$f(A)v \quad \approx \quad \widetilde{x} \qquad \widetilde{x} = ???$$

Two broad paths:

- Substitute f with "simpler" function, $f \approx \mathcal{R}$

$$\|f(A)v - \widetilde{x}\| \leq \|f(A)v - \mathcal{R}(A)v\| + \|\mathcal{R}(A)v - \widetilde{x}\|$$

and $\quad \Rightarrow \quad \widetilde{x} \approx \mathcal{R}(A)v$

(Rational approximation, contour integration, etc.)

Numerical approximation. I

$$f(A)v \quad \approx \quad \tilde{x} \qquad \tilde{x} = ???$$

Two broad paths:

- Substitute f with "simpler" function, $f \approx \mathcal{R}$

$$\|f(A)v - \tilde{x}\| \leq \|f(A)v - \mathcal{R}(A)v\| + \|\mathcal{R}(A)v - \tilde{x}\|$$

  and $\quad \Rightarrow \quad \tilde{x} \approx \mathcal{R}(A)v$

  (Rational approximation, contour integration, etc.)

- Approximation by projection: Find V with $\dim(V) \ll n$ and

$$\tilde{x} \in \mathrm{range}(V)$$

Numerical approximation. II

$$f(A)v \quad \approx \quad \widetilde{x}$$

Important issues:

- $\star$ How does approximation quality depend on properties of f ?
- $\star$ How does approximation quality depend on properties of A ?
- $\star$ Efficiency ?
- $\star$ Measure accuracy of approximation?

Numerical approximation of $f(A)v$

- Rational approximation
- Projection-type approximation

Rational Approximation

$$x = f(A)\,v \quad \approx \quad \mathcal{R}_{\mu,\nu}(A)v$$

$$\mathcal{R}_{\mu,\nu}(\lambda) = \frac{\Phi_\mu(\lambda)}{\Psi_\nu(\lambda)}, \qquad \Phi_\mu(\lambda),\ \Psi_\nu(\lambda) \quad \text{polynomials}$$

- Polynomial Approx., $\nu = 0$: Chebyshev, Leja, Taylor,...
- Rational Approx.: Padé or Chebyshev, e.g. $\mu = \nu$
- Rational Approx with multiple pole
- Quadrature Methods

Rational Approximation

$$x = f(A)\,v \quad \approx \quad \mathcal{R}_{\mu,\nu}(A)v$$

$$\mathcal{R}_{\mu,\nu}(\lambda) = \frac{\Phi_\mu(\lambda)}{\Psi_\nu(\lambda)}, \qquad \Phi_\mu(\lambda),\ \Psi_\nu(\lambda) \quad \text{polynomials}$$

- Polynomial Approx., $\nu = 0$: Chebyshev, Leja, Taylor,...
- Rational Approx.: Padé or Chebyshev, e.g. $\mu = \nu$
- Rational Approx with multiple pole
- Quadrature Methods

We consider the case of partial fraction expansion:

$$\mathcal{R}_{\mu,\nu}(\lambda) = q(\lambda) + \sum_{k=1}^{\nu} \frac{\omega_k}{\lambda - \xi_k} \qquad (\mathcal{R}_\nu = \mathcal{R}_{\nu,\nu})$$

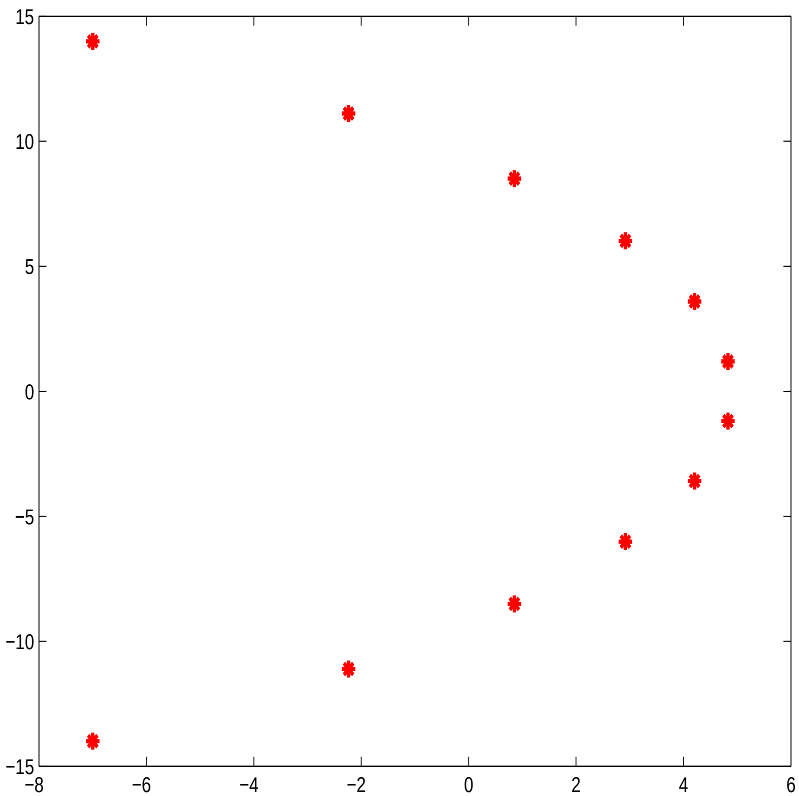
Rational Approximation: poles

$f(\lambda) = \exp(-\lambda)$

$\mathcal{R}_\nu$: ℓ_∞ best approx

in $[0, \infty)$, Chebyshev

$\|f - \mathcal{R}_\nu\|_\infty \approx 10^{-\nu}$

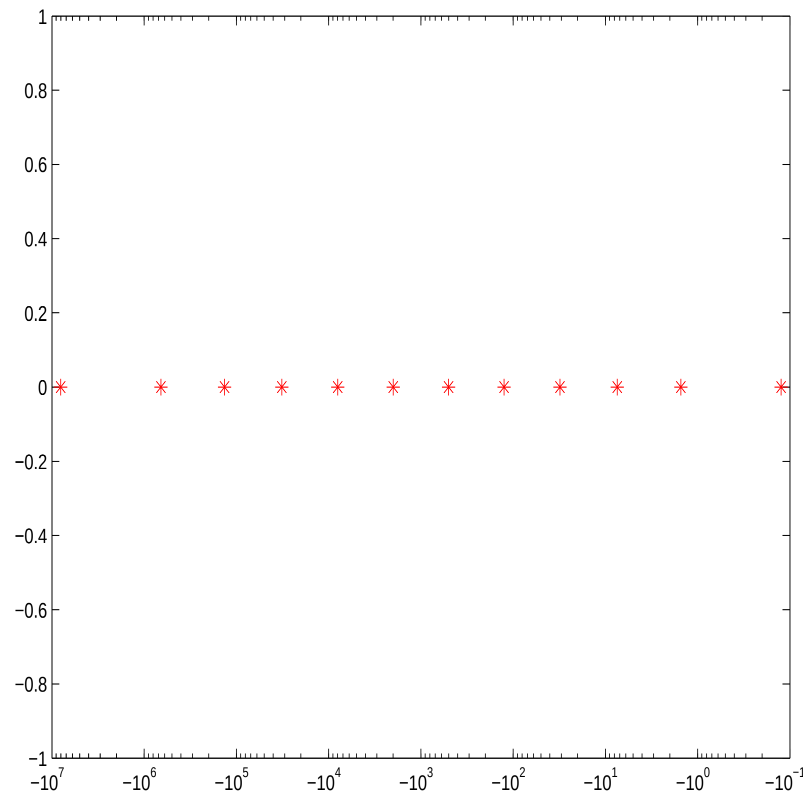


$f(\lambda) = \lambda^{-\frac{1}{2}}$

$\mathcal{R}_\nu$: Zolotarev approx

in $[a, b] \subseteq (0, \infty)$

$\|f - \mathcal{R}_\nu\| \approx e^{-\pi\sqrt{2\nu}}$



Contour integration for $f(A)$ and $f(A)v$.

From the Cauchy formula:

$$f(A) = \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}\mathrm{dz}, \qquad \Gamma = \partial\Omega, \quad \sigma(A) \subset \Omega$$

Used by Hale, Higham & Trefethen for functions with singularities in $(-\infty, 0]$ and $\sigma(A) \approx (0, \infty)$:

$$f(A) \approx c_0 \Im \left(\sum_{j=1}^{N} \frac{f(z_j)}{z_j}(z_j I - A)^{-1} d\omega_j \right)$$

$z_j = z(t_j)$ nodes in a Möbius transf. for mid-point rule integration.

$d\omega_j$ measure associated with the conformal mapping to an annulus.

Tight connection with rational approximation in some cases (e.g. $\lambda^{1/2}$)

Matrix Rational approximation. Computational aspects.

$$f(A)v \approx \mathcal{R}_\nu(A)v = \sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v$$

$$\approx \sum_{k=1}^{\nu} \omega_k \widetilde{x}_k$$

- $\forall k$, $(A - \xi_k I)$ "Shifted" matrix , $\xi_k \in \mathbb{C}$
- $\xi_{2j-1} = \bar{\xi}_{2j}$, $j = 1, \ldots, \lfloor \nu/2 \rfloor$
- $\forall k$, $\widetilde{x}_k$ approximate solution

Matrix Rational approximation. Computational aspects.

$$
\begin{aligned}
f(A)v \quad \approx \quad \mathcal{R}_\nu(A)v \quad &= \quad \sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v \\
&\approx \quad \sum_{k=1}^{\nu} \omega_k \widetilde{x}_k
\end{aligned}
$$

- $\forall k$, $(A - \xi_k I)$ "Shifted" matrix , $\xi_k \in \mathbb{C}$
- $\xi_{2j-1} = \bar{\xi}_{2j}$, $j = 1, \ldots, \lfloor \nu/2 \rfloor$
- $\forall k$, $\widetilde{x}_k$ approximate solution

$\Rightarrow$ Direct Methods (complex $A - \xi_k I$ of modest size)

$\Rightarrow$ Iterative Methods for **shifted** linear systems

Iterative Methods for **shifted** linear systems

$$f(A)v \quad \approx \quad \mathcal{R}_\nu(A)v \quad = \quad \sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v$$

- ♣ Krylov space methods for shifted systems with simultaneous shifts
- ♣ Preconditioners for sequences of shifted systems
- ♣ Ad-hoc stopping strategies for partial fraction expansions
- ♣ Embarassingly parallel

Error estimates for iterative solvers

$\widetilde{x}_i$: Krylov subspace methods: $$\sum_{i=1}^{\nu} \omega_i (A - \xi_i I)^{-1} v \approx \widetilde{z} := \sum_{i=1}^{\nu} \omega_i \widetilde{x}_i$$

$$\|\mathcal{R}_\nu(A)v - \sum_{i=1}^{\nu} \omega_i \widetilde{x}_i\| = ??$$

Error estimate during iteration :

- Estimate for real symmetric A and complex poles
- Lower estimate for A spd and real negative poles

(Borici etal, '05, Frommer & S., '08, Frommer etal., '12)

Error estimates for iterative solvers

$\widetilde{x}_i$: Krylov subspace methods: $$\sum_{i=1}^{\nu} \omega_i (A - \xi_i I)^{-1} v \approx \widetilde{z} := \sum_{i=1}^{\nu} \omega_i \widetilde{x}_i$$

$$\|\mathcal{R}_\nu(A)v - \sum_{i=1}^{\nu} \omega_i \widetilde{x}_i\| = ??$$

Error estimate during iteration :

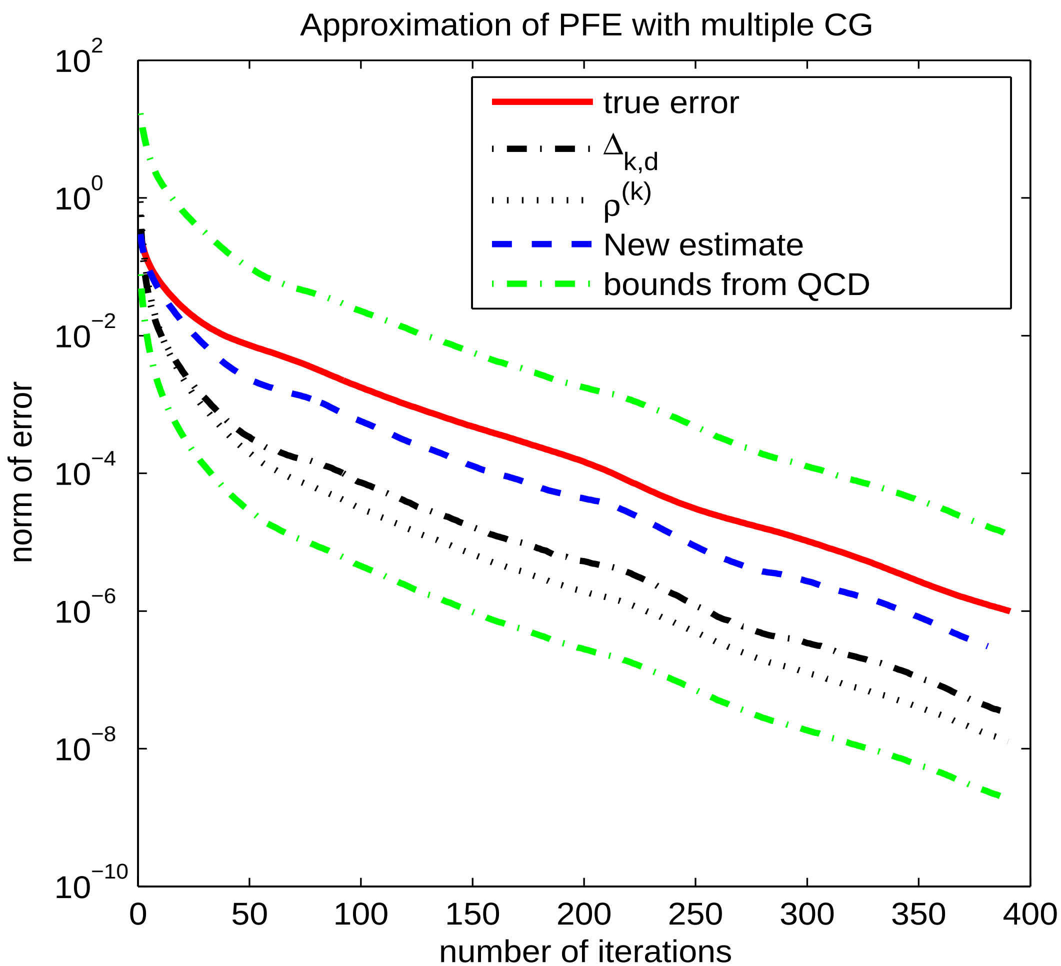
- Estimate for real symmetric A and complex poles
- Lower estimate for A spd and real negative poles

(Borici etal, '05, Frommer & S., '08, Frommer etal., '12)

⋆ Does not require spectral info

⋆ Computational cost only 3-5 additional iterations

CG for A spd and $f(\lambda) = \mathrm{sign}(\lambda) = (\lambda^2)^{-\frac{1}{2}}\lambda$: $\mathrm{sign}(A)v$



11 poles used in the expansion.

$\Delta_{k,d} := \|z^{(k+d)} - z^{(k)}\|$, $\rho^{(k)}$: lin.comb. CG residuals

Numerical approximation of $f(A)v$

- Rational approximation
- Projection-type approximation

Projection-type methods.

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

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

$$x = f(A)v \quad \approx \quad x_m = Vf(V^TAV)(V^Tv)$$

Question: Which $\mathcal{K}$?

Some explored choices for $\mathcal{K}$. Arnoldi-type methods.

- Krylov subspace, $\mathcal{K} = K_m(A, v)$
- Restarted Krylov subspace
- Alternatives: Semi-iterative approaches
  (PAIN method, Chebyshev iteration)

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

Note: In all cases, A nonsymmetric.

Some explored choices for $\mathcal{K}$. Arnoldi-type methods.

- Krylov subspace, $\mathcal{K} = K_m(A, v)$
- Restarted Krylov subspace
- Alternatives: Semi-iterative approaches
  (PAIN method, Chebyshev iteration)

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

Note: In all cases, A nonsymmetric.

Large body of recent literature

Krylov subspace approximation

"Classical" approach:

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

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

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

Krylov subspace approximation

"Classical" approach:

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

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

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

Polynomial approximation: $\qquad x_m = p_{m-1}(A)v$

(p_{m-1} interpolates f at eigenvalues of H_m)

$\star$ Numerical and theoretical results since mid '80s

(van der Vorst'87, Saad'92, Hochbruck & Lubich'97, ...)

Krylov vs. rational approximation

$f \to \mathcal{R}_\nu$

CG-type approximation (Galerkin) of linear systems:

$$
\begin{aligned}
\mathcal{R}_\nu(A)v &= \omega_0 v + \sum_{j=1}^{\nu} \omega_j (A - \xi_j I)^{-1} v \\
&\approx \omega_0 v + \sum_{j=1}^{\nu} \omega_j V_m (H_m - \xi_j I)^{-1} e_1 \\
&= V_m \left(\omega_0 e_1 + \sum_{j=1}^{\nu} \omega_j (H_m - \xi_j I)^{-1} e_1 \right) \equiv V_m \mathcal{R}_\nu(H_m) e_1
\end{aligned}
$$

Krylov vs. rational approximation

$f \to \mathcal{R}_\nu$

CG-type approximation (Galerkin) of linear systems:

$$
\begin{aligned}
\mathcal{R}_\nu(A)v &= \omega_0 v + \sum_{j=1}^{\nu} \omega_j (A - \xi_j I)^{-1} v \\
&\approx \omega_0 v + \sum_{j=1}^{\nu} \omega_j V_m (H_m - \xi_j I)^{-1} e_1 \\
&= V_m \left(\omega_0 e_1 + \sum_{j=1}^{\nu} \omega_j (H_m - \xi_j I)^{-1} e_1 \right) \equiv V_m \mathcal{R}_\nu(H_m) e_1
\end{aligned}
$$

Krylov approximation: $x \approx V_m f(H_m) e_1$

Application. Evolution Problem

$$\begin{cases} \frac{\partial u(x,y,t)}{\partial t} = \Delta u, \quad (x,y) \in (0,1)^2 \quad t \in [0, 0.1] \\ u(x,y,t) = 0, \quad (x,y) \in \partial([0,1]^2) \\ u(x,y,0) = 1, \quad (x,y) \in [0,1]^2 \end{cases}$$

Implicit Euler: $u_{i+1} = (I + \delta t A)^{-1} u_i, \quad i = 0, 1, ...$

Exponential Integrator: $u(t) = \exp(-tA) u_0 \quad t = 0.1$

Application. Evolution Problem

$$\begin{cases} \frac{\partial u(x,y,t)}{\partial t} = \Delta u, & (x,y) \in (0,1)^2 \quad t \in [0, 0.1] \\ u(x,y,t) = 0, & (x,y) \in \partial([0,1]^2) \\ u(x,y,0) = 1, & (x,y) \in [0,1]^2 \end{cases}$$

Implicit Euler: $u_{i+1} = (I + \delta t A)^{-1} u_i, \quad i = 0, 1, \ldots$

Exponential Integrator: $u(t) = \exp(-tA)u_0 \quad t = 0.1$

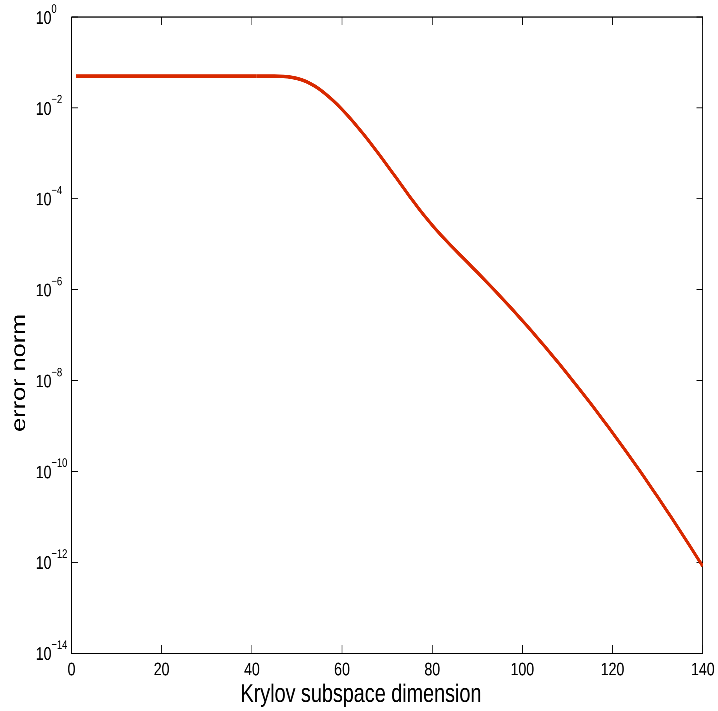
	Euler		Exp	
step δt	CPU	error	CPU	error (#its*)
0.001	1.9	$2 \cdot 10^{-3}$	0.09	$9 \cdot 10^{-4}$(37)
0.005	0.4	$1 \cdot 10^{-2}$	0.07	$4 \cdot 10^{-3}$(28)
0.01	0.2	$2 \cdot 10^{-2}$	0.05	$1 \cdot 10^{-2}$(25)

* : Stopping criterion tolerance related to Euler timestep

$\Rightarrow$ General exponential integrators (Hochbruck,Ostermann, Acta Num'10)

...When things are not so easy

$$\|\exp(-A)v - V_m \exp(-H_m)e_1\| \qquad A \in \mathbb{R}^{400\times 400}, \|A\| = 10^5$$



$$\|\exp(-A)v - V_m \exp(-H_m)e_1\| \le 10e^{-m^2/(5\rho)}, \qquad \sqrt{4\rho} \le m \le 2\rho$$

$$\|\exp(-A)v - V_m \exp(-H_m)e_1\| \le \frac{10}{\rho}e^{-\rho}\left(\frac{e\rho}{m}\right)^m, \qquad m \ge 2\rho$$

where $\sigma(A) \subseteq [0, 4\rho]$

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

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

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, '09)

A nonsym, (Beckermann & Reichel'10)

Acceleration Procedures: Restarted Krylov

$$AV_m^{(1)} = V_m^{(1)} H_m^{(1)} + v_{m+1}^{(1)} h_{m+1,m}^{(1)} e_m^T \quad (V_m^{(1)})^T V_m^{(1)} = I$$

$$AV_m^{(2)} = V_m^{(2)} H_m^{(2)} + v_{m+1}^{(2)} h_{m+1,m}^{(2)} e_m^T \quad (V_m^{(2)})^T V_m^{(2)} = I$$

with $V_m^{(2)} e_1 = v_{m+1}^{(1)}$.

Acceleration Procedures: Restarted Krylov

$$AV_m^{(1)} = V_m^{(1)} H_m^{(1)} + v_{m+1}^{(1)} h_{m+1,m}^{(1)} e_m^T \quad (V_m^{(1)})^T V_m^{(1)} = I$$

$$AV_m^{(2)} = V_m^{(2)} H_m^{(2)} + v_{m+1}^{(2)} h_{m+1,m}^{(2)} e_m^T \quad (V_m^{(2)})^T V_m^{(2)} = I$$

with $V_m^{(2)} e_1 = v_{m+1}^{(1)}$. Then

$$A[V_m^{(1)}, V_m^{(2)}] = [V_m^{(1)}, V_m^{(2)}]\widehat{H}_{2m} + v_{m+1}^{(2)} h_{m+1,m}^{(2)} e_{2m}^T,$$

with

$$\widehat{H}_{2m} = \begin{bmatrix} H_m^{(1)} & 0 \\ e_1 h_{m+1,m}^{(1)} e_m^T & H_m^{(2)} \end{bmatrix}.$$

Acceleration Procedures: Restarted Krylov

$$AV_m^{(1)} = V_m^{(1)} H_m^{(1)} + v_{m+1}^{(1)} h_{m+1,m}^{(1)} e_m^T \quad (V_m^{(1)})^T V_m^{(1)} = I$$

$$AV_m^{(2)} = V_m^{(2)} H_m^{(2)} + v_{m+1}^{(2)} h_{m+1,m}^{(2)} e_m^T \quad (V_m^{(2)})^T V_m^{(2)} = I$$

with $V_m^{(2)} e_1 = v_{m+1}^{(1)}$. Then

$$A[V_m^{(1)}, V_m^{(2)}] = [V_m^{(1)}, V_m^{(2)}] \widehat{H}_{2m} + v_{m+1}^{(2)} h_{m+1,m}^{(2)} e_{2m}^T,$$

with

$$\widehat{H}_{2m} = \begin{bmatrix} H_m^{(1)} & 0 \\ e_1 h_{m+1,m}^{(1)} e_m^T & H_m^{(2)} \end{bmatrix}.$$

Therefore (Eiermann-Ernst, '06)

$$\begin{aligned} f(A)v &\approx x_m^{(1)} = V_m^{(1)} f(H_m^{(1)}) \\ &\approx x_m^{(2)} = V_m^{(1)} f(H_m^{(1)}) e_1 + V_m^{(2)} f(\widehat{H}_{2m}) e_1|_{(2)} \\ & \quad\; x_m^{(2)} = x_m^{(1)} + V_m^{(2)} f(\widehat{H}_{2m}) e_1|_{(2)} \end{aligned}$$

Acceleration Procedures: Extended Krylov

For A nonsingular,

$\mathcal{K} = K_{m_1}(A, v) + K_{m_2}(A^{-1}, A^{-1}v),$ Druskin-Knizhnerman'98, 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:

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

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

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

Simoncini '07, Reichel etal, '10, Knizhnerman-S., '10, ...

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:

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

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

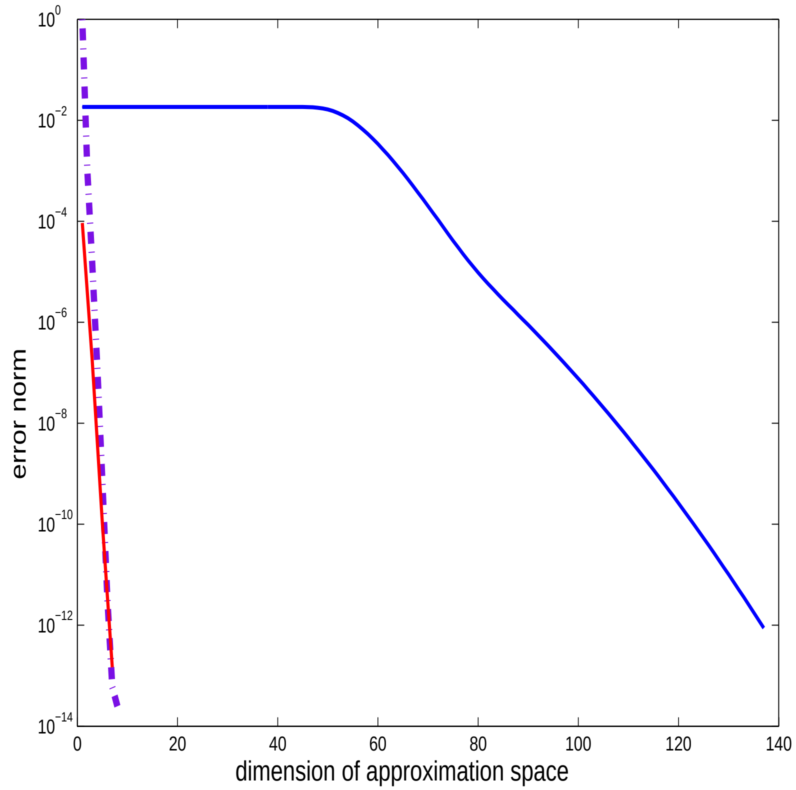
⋆ Compute $x_m = \mathcal{U}_m f(\mathcal{T}_m) e_1$

Simoncini '07, Reichel etal, '10, Knizhnerman-S., '10, ...

A symmetric $\Rightarrow$ (block) short-term recurrence

Acceleration

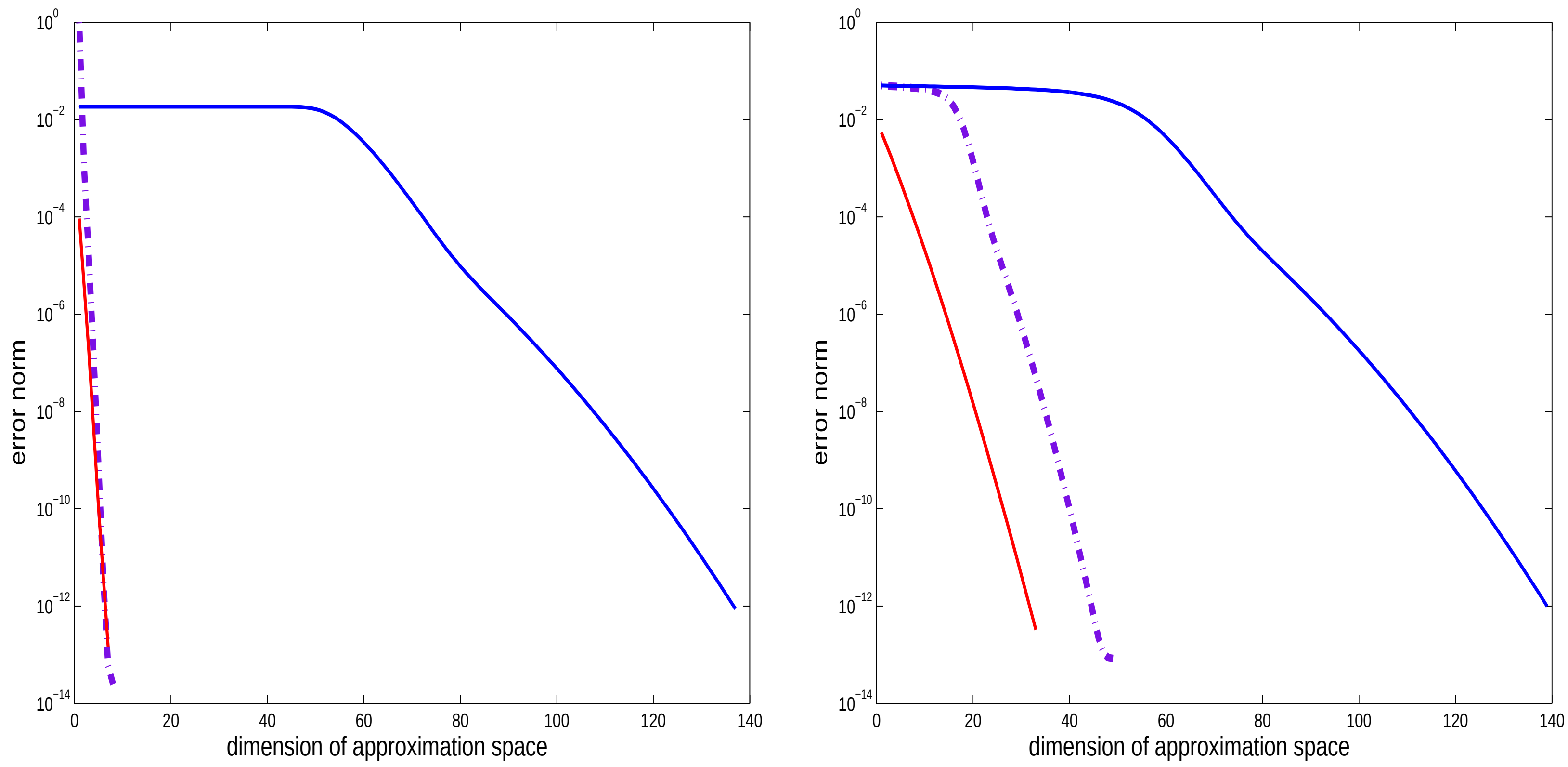
$f(\lambda) = \exp(-\lambda)$



-: std Krylov -.: Spectral accel. -: “extended” space

Acceleration

$f(\lambda) = \exp(-\lambda)$ $\qquad$ $f(\lambda) = \lambda^{-1/2}$



-: std Krylov $\quad$ -.: Spectral accel. $\quad$ -: "extended" space

Large-scale numerical experiments

A from FD discretization of

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

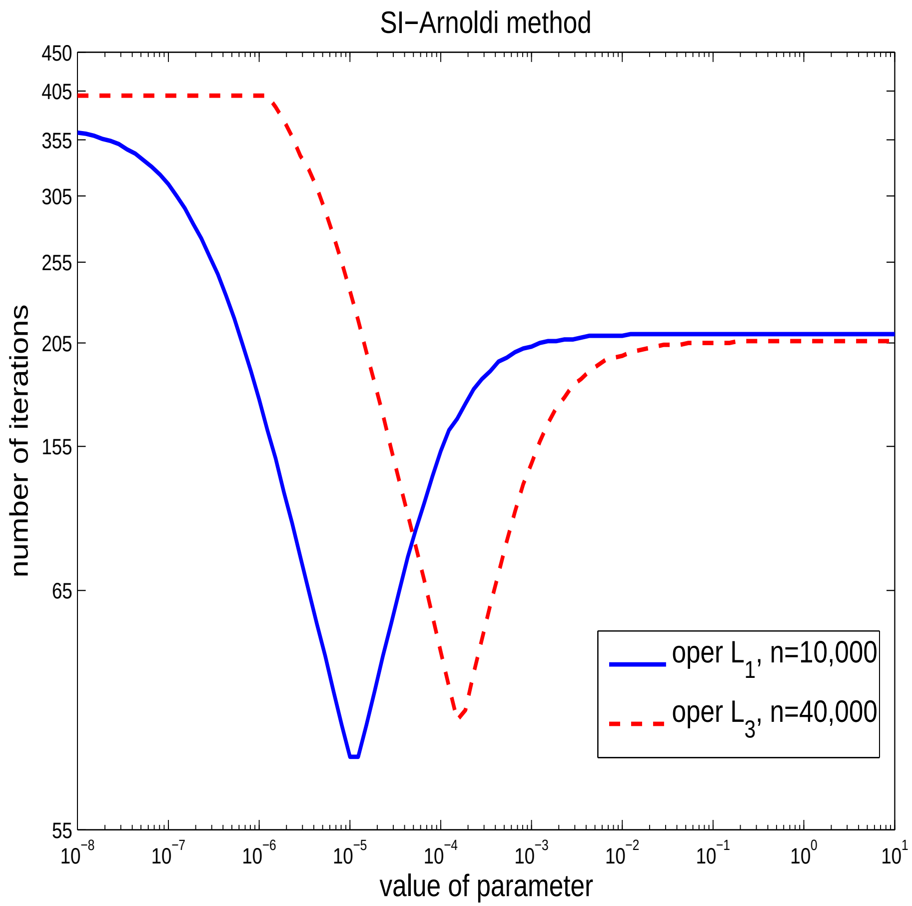
on unit square/cube, Dirichlet hom. bc.

Inner system solves:

- Extended Krylov: systems with A solved with GMRES/AMG
- SI-Arnoldi: span$\{v, (I+\gamma A)^{-1}v, (I+\gamma A)^{-2}v, \ldots\}$
  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 $f(\lambda) = \lambda^{\frac{1}{2}}$

Comparisons: CPU Time in Matlab (space dim.)

f	Oper.	n	SI-Arnoldi	EKSM	Std Krylov
$\lambda^{\frac{1}{2}}$	$\mathcal{L}_1$	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
$\lambda^{-\frac{1}{3}}$	$\mathcal{L}_1$	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 $\|x - x_m\|$ with x_m by projection:

(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:

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

Stopping criterion. The exponential function.

Consider the ODE problem:

$$y' = -Ay, \qquad y(t_0) = y_0$$

with exact solution: $\qquad y(t) = e^{-tA} y_0$

Stopping criterion. The exponential function.

Consider the ODE problem:

$$y' = -Ay, \qquad y(t_0) = y_0$$

with exact solution: $\qquad y(t) = e^{-tA}y_0$

ODE Residual: Given $y_k(t) \approx y(t)$,

$$r_k(t) := -Ay_k - y_k'$$

If $y_k = V_k u_k$ with $u_k = e^{-t(V_k^T A V_k)} V_k^T e_1$, then

$$\|r_k\| = |h_{k+1,k}(e_k^T u_k)|, \quad \text{with } h_{k+1,k} = v_{k+1}^T A v_k$$

(Celledoni,Moret '97, Druskin-Greenbaum-Knizhnerman '98,...,

Botchev-Grimm-Hochbruck '13)

Rational Krylov subspaces (RKS)

$$\mathcal{K} = \text{span}\{(A-\omega_1 I)^{-1}v, \ldots, (A-\omega_m I)^{-1}v\}$$

$$f(A)v \approx V_m f(H_m) V_m^T v, \qquad \mathcal{K} = \text{range}(V_m), \quad H_m = V_m^T A V_m$$

Key properties (cf. Güttel, GAMM Mitteilungen '13):

- The approximation interpolates f at $\sigma(H_m)$ ($\Rightarrow$ exactness for low degree rational functions)
- Relation to rational approximation
- Near-optimality property among all rational approximations of equal degree

Rational Krylov Subspaces (RKS). A long tradition...

$$\mathcal{K}_m(A, v, \boldsymbol{\omega}) = \mathrm{range}([(A - \omega_1 I)^{-1} v, \ldots, (A - \omega_m I)^{-1} v])$$

- Eigenvalue problems (Ruhe, 1984)
- Model Order Reduction (transfer function evaluation)
- ADI for linear matrix equations

Rational Krylov subspaces (RKS)

An alternative characterization :

(appropriate for certain matrix functions and matrix equations)

$$\mathcal{Q}_m(A, v, \boldsymbol{\omega}) = \text{span}\{\frac{p_{m-1}(A)}{q_{m-1}(A)} v \ : \ p_{m-1} \text{polynomial of degree } \leq m-1\}$$

with $q_{m-1}(z) := \prod_{i=1}^{m-1}(z - \omega_i)$ fixed, $\omega_i \neq \infty$.

Recurrence for generating the space (Ruhe, '84):

1. Compute $\hat{v}_{j+1} = (I - \frac{1}{\omega_j} A)^{-1} A v_j$
2. Orthogonalize $\hat{v}_{j+1}$ wrto previous v_i, $i = 1, \ldots, j$

Rational Krylov Subspaces in MOR. Choice of poles.

$$\mathcal{K}_m(A, v, \boldsymbol{\omega}) = \mathrm{range}([(A - \omega_1 I)^{-1}v, \ldots, (A - \omega_m I)^{-1}v])$$

cf. General discussion in Antoulas, 2005.

Various attempts:

- Gallivan, Grimme, Van Dooren (1996–, ad-hoc poles)
- Penzl (1999-2000, ADI shifts - preprocessing, Ritz values)
- ....
- Sabino (2006 - tuning within preprocessing)

- IRKA – Gugercin, Antoulas, Beattie (2008)

$\Rightarrow$ Asymptotically optimal poles, Beckermann-Reichel, '10

Adaptive choice of poles for RKS.

$$K_m(A, v, \boldsymbol{\omega}) = \text{range}([(A-\omega_1 I)^{-1}v, (A-\omega_2 I)^{-1}v, \ldots, (A-\omega_m I)^{-1}v])$$

$\boldsymbol{\omega} = [\omega_1, \ldots, \omega_m]$ to be chosen sequentially

Explored for different functions $\Rightarrow$ different probing sets.

- Exponential and resolvent functions $f(\lambda) = (\lambda - i\sigma)^{-1}$
- Markov functions, those that can be defined via

$$f(z) = \int_\Gamma \frac{1}{z-\zeta} d\gamma(\zeta)$$

($z^{-1/2}$, $(e^{-\tau\sqrt{z}} - 1)/z$, etc.)

(Güttel-Knizhnerman, '13)

Adaptive choice of poles for RKS. Exponential and resolvent functions

The fundamental idea: Assume you wish to solve

$$(A - sI)x = v$$

with a Galerkin procedure in $K_m(A, v, \boldsymbol{\omega})$. Let V_m be orth. basis. The residual satisfies:

$$v - (A - zI)x_m = \frac{r_m(A)b}{r_m(z)}, \qquad r_m(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{z - \omega_j}$$

with $\lambda_j = \mathrm{eigs}(V_m^T A V_m)$. Moreover,

$$\|r_m(A)b\| = \min_{\theta_1,\ldots,\theta_m} \| \prod_{j=1}^{m} (A - \theta_j I)(A - \omega_j I)^{-1} v \|$$

Adaptive choice of poles for RKS. Cont'd

$$r_m(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{z - \omega_j}, \qquad \lambda_j = \mathrm{eigs}(V_m^T A V_m)$$

For A symmetric:

$$\omega_{m+1} := \arg \left(\max_{s \in [-\lambda_{\max}, -\lambda_{\min}]} \frac{1}{|r_m(s)|} \right)$$

$[\lambda_{\min}, \lambda_{\max}] \approx \mathrm{spec}(A)$ (Druskin, Lieberman, Zaslavski (SISC 2010))

Adaptive choice of poles for RKS. Cont'd

$$r_m(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{z - \omega_j}, \qquad \lambda_j = \mathrm{eigs}(V_m^T A V_m)$$

For A symmetric:

$$\omega_{m+1} := \arg \left(\max_{s \in [-\lambda_{\max}, -\lambda_{\min}]} \frac{1}{|r_m(s)|} \right)$$

$[\lambda_{\min}, \lambda_{\max}] \approx \mathrm{spec}(A)$ (Druskin, Lieberman, Zaslavski (SISC 2010))

For A nonsymmetric:

$$\omega_{m+1} := \arg \left(\max_{s \in \partial \mathcal{S}_m} \frac{1}{|r_m(s)|} \right)$$

where $\mathcal{S}_m \subset \mathbb{C}^+$ approximately encloses the eigenvalues of $-A$

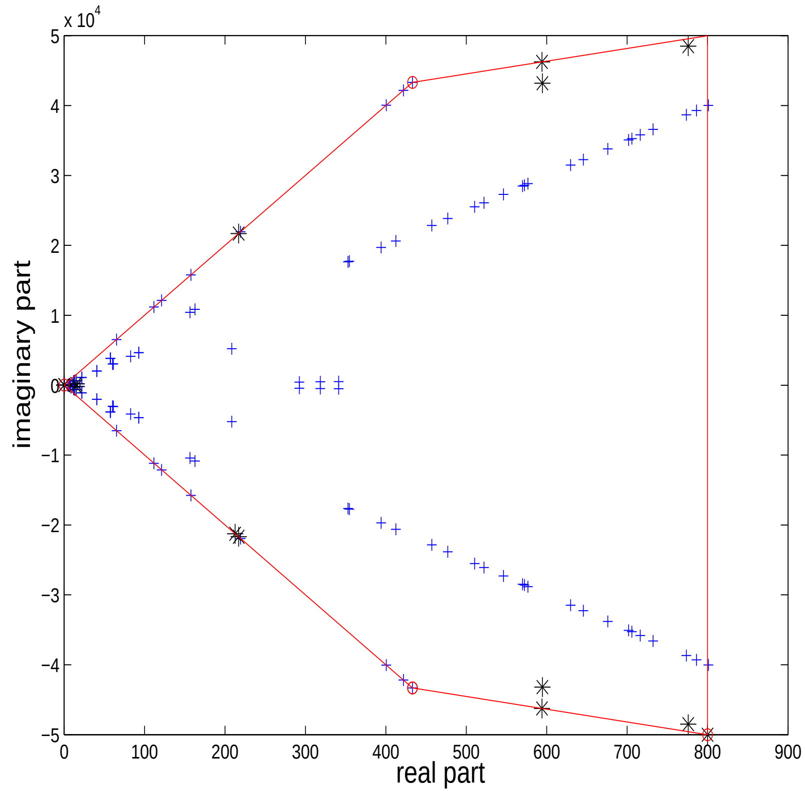
(Druskin-Simoncini '11)

Motivated by potential theory arguments...

Example of $\mathcal{S}_m$. CD Player, $m = 12$

$\mathcal{S}_m$: encloses mirrored current Ritz values: -eigs($V_m^T A V_m$)

and initial estimates: $s_1^{(0)} = 0.1$, $s_{2,3}^{(0)} = 900 \pm i5 \cdot 10^4$

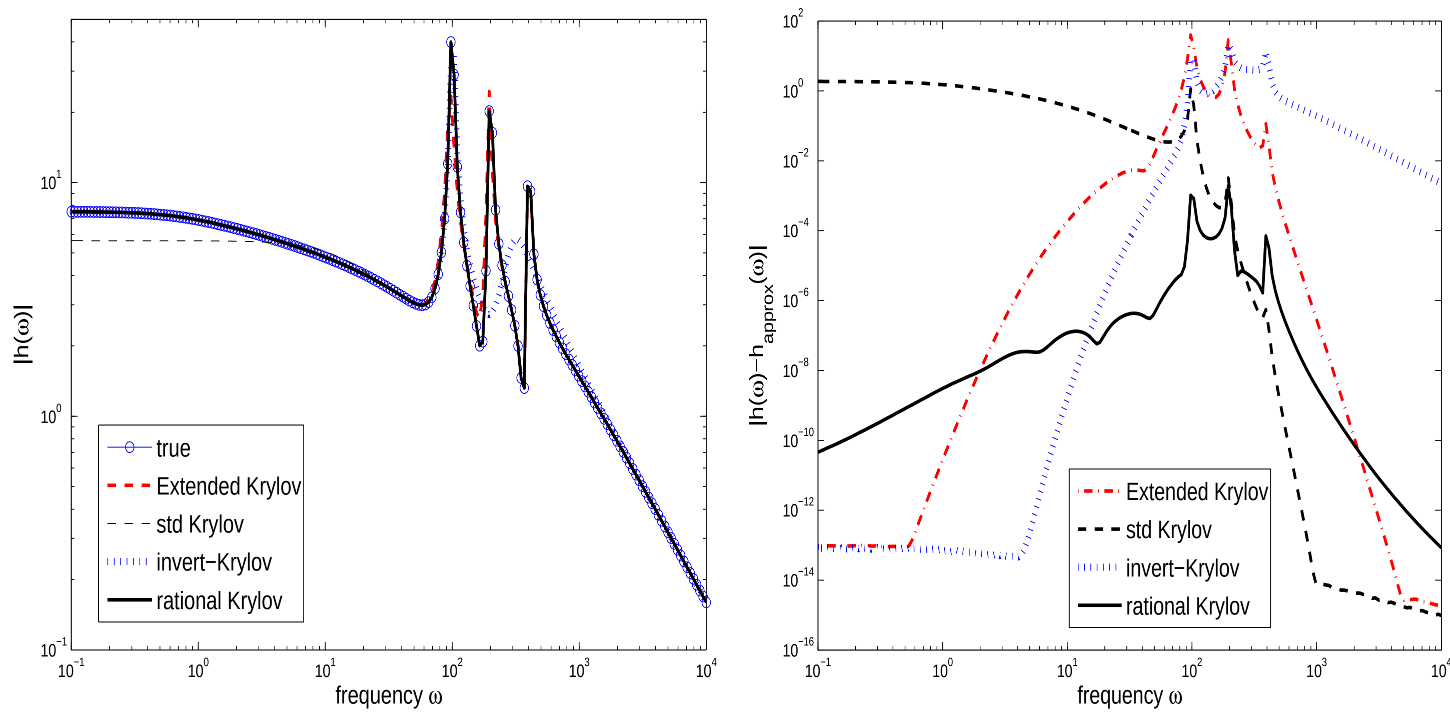


$*$ poles　　$+$ -eigs(A)　　—○— $\partial\mathcal{S}_m$

Approximation of the transfer function

$$h(\omega) = c^T(A - \omega I)^{-1}b, \;\Rightarrow\; h_m(\omega) = c^T V_m (V_m^T A V_m - \omega I)^{-1} V_m^T b$$

$\omega \in i\mathbb{R}$ $\qquad h(\omega) \qquad\qquad |h(\omega) - h_{approx}(\omega)|$



FOM matrix, $n = 1006$, space dim. $m = 20$

The multiple vector case.

Some applications require $f(A)\mathbf{v}$, with $\mathbf{v} \in \mathbb{C}^{n\times p}$:

- Structure preserving functions
- Multiple input/output transfer functions
- ODEs with multiple evaluation for different initial approx

⇒ All projection spaces generalize to multiple vectors

e.g., $K_m(A,\mathbf{v}) = \mathrm{range}([\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}])$ (dim$(K_m(A,\mathbf{v})) \le mp$)

The multiple vector case.

Some applications require $f(A)\mathbf{v}$, with $\mathbf{v} \in \mathbb{C}^{n\times p}$:

- Structure preserving functions
- Multiple input/output transfer functions
- ODEs with multiple evaluation for different initial approx

$\Rightarrow$ All projection spaces generalize to multiple vectors

e.g., $K_m(A, \mathbf{v}) = \mathrm{range}([\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}])$ (dim$(K_m(A, \mathbf{v})) \le mp$)

- Easy to implement
- Appropriate for preserving structure
- Redundant in certain applications

The multiple vector case

Block Krylov-type subspaces:

$$K_m(A, \mathbf{v}) = \text{range}([\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}])$$

- $\mathcal{V}$, $\mathcal{V}^T\mathcal{V} = I$ such that

$$\text{Range}(\mathcal{V}) = K_m(A, \mathbf{v})$$

- $\mathcal{H} = \mathcal{V}^T A \mathcal{V}$

Then

$$f(A)\mathbf{v} \quad \approx \quad \mathcal{V} f(\mathcal{H})(\mathcal{V}^T\mathbf{v})$$

Flows on constraint manifolds

Motivational problem:

Approximate k largest Lyapunov exponents of

$$x'(t) = \mathcal{A}(t)x, \quad x \in \mathbb{R}^n$$

This can be accomplished by using the associated system

$$Q_t = A(Q,t)Q, \quad Q \in \mathbb{R}^{n\times k} \qquad A \text{ skew-sym}$$

Q orthonormal columns (Stiefel manifold)

Goal: Numer. method preserving orthogonality for long time intervals

⋆ $A_n = A(Q_n, t_n)$ skew-sym. $\Rightarrow$ $\exp(A_n)$ unitary and

$$Q_{n+1} = \exp(tA_n)Q_n \text{ orthogonal}$$

Preserving orthogonality in Krylov subspace

Let $Q^{(0)} = [q_1^{(0)}, \ldots, q_k^{(0)}]$

Regular Krylov subspaces $\mathcal{K}_m(A, q_i^{(0)})$, $i = 1, \ldots, k$

$$A \text{ skew-sym} \quad \Rightarrow \quad H_{m,i} \text{ skew-sym} \quad \Rightarrow \exp(tH_{m,i}) \text{ unitary}$$

This is not enough:

$$\exp(tA)q_i^{(0)} \quad \approx \quad q_i = V_{m,i} \exp(tH_{m,i})e_1$$

$\{q_1, \ldots, q_k\}$ not orthogonal (though unit norm)

Block Krylov methods come to rescue

Block Krylov subspace $\mathcal{K}_m(A, Q^{(0)})$ $\qquad Q^{(0)} = [q_1^{(0)}, \ldots, q_k^{(0)}]$

$$\mathcal{K}_m(A, Q^{(0)}) = \text{range}([Q^{(0)}, AQ^{(0)}, \ldots, A^{m-1}Q^{(0)}])$$

- $\mathcal{V}_m$ orthonormal columns,

  $\mathcal{H}_m = \mathcal{V}_m^T A \mathcal{V}_m$ skew-sym

- $\mathcal{V}_m \exp(t\mathcal{H}_m)E_1$ orthonormal columns

(Lopez-S.,'06)

Further generalizations: A skew-symmetric and Hamiltonian

- $\exp(tA)$ ortho-symplectic - w.r.to $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$
- $Q^{(0)}$ ortho-symplectic then $\exp(tA)Q^{(0)}$ ortho-symplectic

---

Block Krylov approximation:

- Choose *some* of the columns $\widetilde{Q}^{(0)}$ of $Q^{(0)}$,

$$V = \begin{pmatrix} \widetilde{Q}_1^{(0)} & \widetilde{Q}_2^{(0)} \\ \widetilde{Q}_2^{(0)} & -\widetilde{Q}_1^{(0)} \end{pmatrix} \qquad \mathcal{K}_m(A, V) \qquad \mathcal{H}_m = \mathcal{V}_m^T A \mathcal{V}_m$$

- $\mathcal{V}_m \exp(t\mathcal{H}_m)E_1$ columns of an ortho-symplectic matrix

X ortho-symplectic if $X^T J X = J$ and $X^T X = I$

Further generalizations. A Hamiltonian

$Q^{(0)}$ symplectic then $\exp(A)Q^{(0)}$ symplectic

Construct symplectic basis $\mathcal{V}_m$ and (logically) Hamiltonian $\mathcal{H}_m$:

Block Lanczos procedure in the block J-inner product:

$$[X, Y]_J = J_2^T X^T J Y \qquad X, Y \in \mathbb{R}^{2n\times 2}$$

$$J_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

single vector case: Benner & Faßbender '97-'00, Watkins '04, Salam '06

Multiple vectors. Tangential projection

$$K_m(A, \mathbf{v}, \boldsymbol{\omega}) = \mathrm{range}([(A - \omega_1 I)^{-1}\mathbf{v}, \ldots, (A - \omega_m I)^{-1}\mathbf{v}])$$

- Generates possibly redundant information
- Memory/Computational costs inefficient

An alternative:

$$\mathbf{T}_m = \mathrm{range}([(A - \omega_1 I)^{-1}\mathbf{v}d_1, \ldots, (A - \omega_m I)^{-1}\mathbf{v}d_m])$$

with an adaptive choice of (ω_i, d_i)

(Druskin-Zaslavski-S.,'12)

Tangential rational Krylov subspace

$$\mathbf{T}_m = \text{range}([(A-\omega_1 I)^{-1}\mathbf{v}d_1, \ldots, (A-\omega_m I)^{-1}\mathbf{v}d_m]) = range(V_m)$$

Some properties:

- $\mathcal{H}(\omega_i)d_i = \mathcal{H}_m(\omega_i)d_i, \quad i = 1, \ldots, m$
- For A symmetric,

$$d_i^* \frac{d}{ds}\, \mathcal{H}(s)|_{s=\omega_i}\, d_i = d_i^* \frac{d}{ds}\, \mathcal{H}_m(s)|_{s=\omega_i}\, d_i, \quad i = 1, \ldots, m$$

- Let $v_{m+1} = (A - \omega_{m+1} I)^{-1}\mathbf{v}d_{m+1}$ and $R_m(z) = (A - zI)V_m(H_m - zI)^{-1}V_m^*\mathbf{v} - \mathbf{v}$, then

$$\begin{aligned} \mathbf{T}_{m+1} &:= \text{range}([V_m, v_{m+1}]) \\ &= \text{range}([V_m, (A-\omega_{m+1}I)^{-1}R_m(\omega_{m+1})d_{m+1}]), \end{aligned}$$

and $\dim(\mathbf{T}_{m+1}) = m+1$ if and only if $R_m(\omega_{m+1})d_{m+1} \neq 0$

Adaptive choice of poles and directions.

$$\mathbf{T}_m = \text{range}([(A - \omega_1 I)^{-1}\mathbf{v}d_1, \ldots, (A - \omega_m I)^{-1}\mathbf{v}d_m])$$

Single direction:

$$(d_{m+1}, \omega_{m+1}) = \arg \max_{\substack{s \in \mathcal{S}_m \\ d \in \mathbb{R}^p, \|d\|=1}} \|R_m(s)d\|$$

In fact:

1. Compute $\omega_{m+1} := s$ where $\|R_m(s)\|$ is largest
2. Compute d_{m+1} as principal SVD direction of $R_m(\omega_{m+1})$

Adaptive choice of poles and directions.

$$\mathbf{T}_m = \mathrm{range}([(A - \omega_1 I)^{-1}\mathbf{v}d_1, \ldots, (A - \omega_m I)^{-1}\mathbf{v}d_m])$$

Single direction:

$$(d_{m+1}, \omega_{m+1}) = \arg \max_{\substack{s \in \mathcal{S}_m \\ d \in \mathbb{R}^p, \|d\|=1}} \|R_m(s)d\|$$

In fact:

1. Compute $\omega_{m+1} := s$ where $\|R_m(s)\|$ is largest
2. Compute d_{m+1} as principal SVD direction of $R_m(\omega_{m+1})$

Multiple directions:

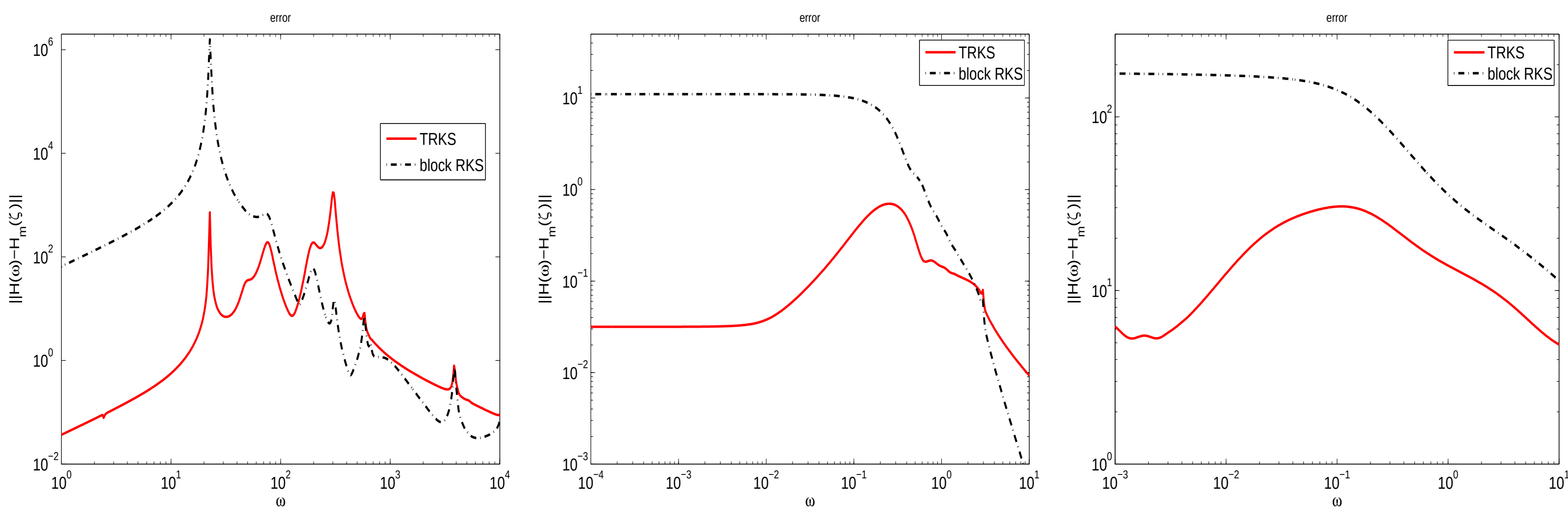
2'. Compute $d_{m+1} \in \mathbb{R}^{p \times \ell}$ as ℓ principal SVD directions of $R_m(s_{m+1})$, e.g.,

$$\ell \quad s.t. \quad \sigma^{(i)} > \frac{1}{10}\sigma^{(1)}, \quad i = 1, \ldots, \ell$$

where $\sigma^{(k)}$, $k = 1, \ldots, p$ are the sing.values of $R_m(s_{m+1})$

Some numerical examples. Approximation of $\mathcal{H}(z) = C(A - zI)^{-1}B$

$$\|\mathcal{H}(z) - \mathcal{H}_m(z)\|, \qquad z \in i[\alpha, \beta]$$



Data from Oberwolfach collection: CD Player, EADY, FLOW

Original block RKSM vs Tangential approach (TRKS)

Final space dimension = 10 ($p = 2, 10, 5$ in the three cases, resp.)

Real poles

Applications. Ill-posed problem. I

$$
\begin{cases}
u_{zz} - Lu = 0, & (x,y,z) \in \Omega \times [0, z_1] \\
u(x,y,z) = 0, & (x,y,z) \in \partial\Omega \times [0, z_1] \\
u(x,y,0) = g(x,y), & (x,y) \in \Omega \\
u_z(x,y,0) = 0, & (x,y) \in \Omega
\end{cases}
$$

L elliptic oper., linear, self-adjoint, positive def.

Pb: determine u for $z = z_1$: $f(x,y) = u(x,y,z_1),\ (x,y) \in \Omega$.

Applications. Ill-posed problem. I

$$\begin{cases} u_{zz} - Lu = 0, & (x,y,z) \in \Omega \times [0, z_1] \\ u(x,y,z) = 0, & (x,y,z) \in \partial\Omega \times [0, z_1] \\ u(x,y,0) = g(x,y), & (x,y) \in \Omega \\ u_z(x,y,0) = 0, & (x,y) \in \Omega \end{cases}$$

L elliptic oper., linear, self-adjoint, positive def.

Pb: determine u for $z = z_1$: $f(x,y) = u(x,y,z_1),\ (x,y) \in \Omega$.

Separation of variables: $u(x,y,z) = \cosh(z\sqrt{L})g$

$\star$: L unbounded $\Rightarrow \cosh(z\sqrt{L})g$ unstable (wrto data perturbations)

Applications. Ill-posed problem. II

Regularization: $\widetilde{g}$ perturbed data

$$u(x,y,z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z)\,\langle s_k, g\rangle\, s_k(x,y)$$

$$\Rightarrow\ v(x,y,z) = \sum_{\lambda_k \le \lambda_c} \cosh(\lambda_k z)\,\langle s_k, \widetilde{g}\rangle\, s_k(x,y)$$

(λ_k^2, s_k) eigenpairs of L

Applications. Ill-posed problem. II

Regularization: $\widetilde{g}$ perturbed data

$$u(x,y,z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z)\,\langle s_k, g\rangle\, s_k(x,y)$$

$$\Rightarrow\ v(x,y,z) = \sum_{\lambda_k \le \lambda_c} \cosh(\lambda_k z)\,\langle s_k, \widetilde{g}\rangle\, s_k(x,y)$$

(λ_k^2, s_k) eigenpairs of L

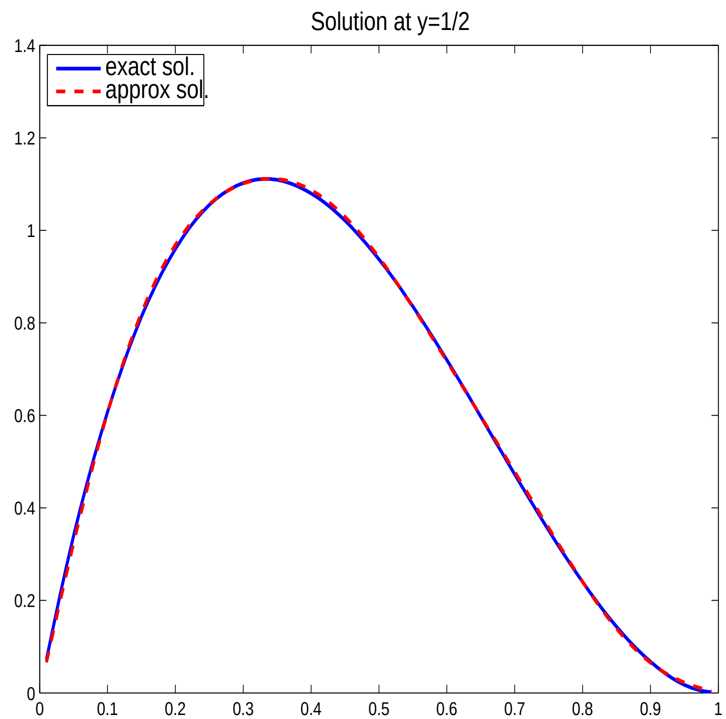
Approx, for instance, in Krylov subspace $\mathcal{K}_m(L, \widetilde{g})$:

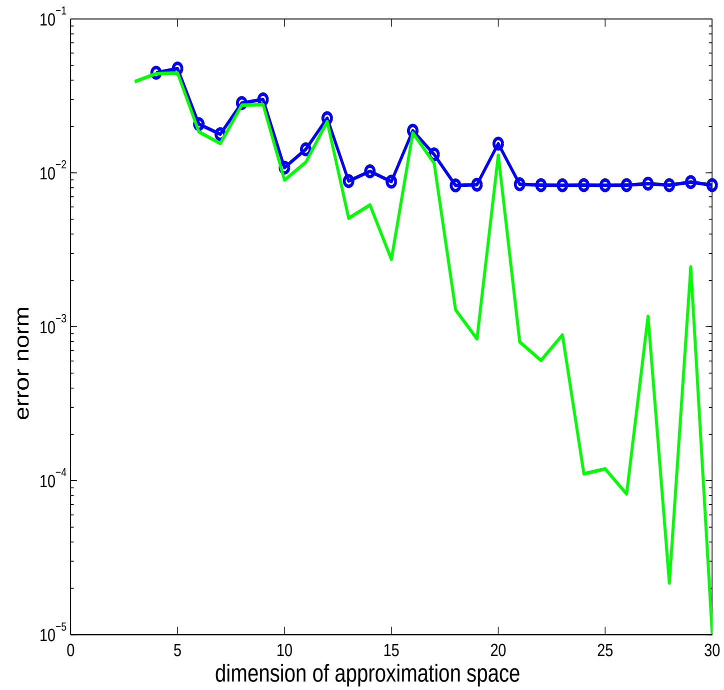
$$u^{(m)}(z) = V_m \cosh(z\sqrt{H_m})e_1\|g\| \quad \Rightarrow$$

$$\Rightarrow\ v^{(m)}(z) = V_m \sum_{\theta_j^{(m)} \le \lambda_c} y_j^{(m)} \cosh(z\theta_j^{(m)})(y_j^{(m)})^T e_1\|\ \|$$

$((\theta_j^{(m)})^2, y_j^{(m)})$ eigenpairs of H_m

Applications. Ill-posed problem. III





Functional error: - $\|v(z) - v^{(m)}(z)\|$ $z = 0.1$

Perturbation error : - $\|u(z) - v^{(m)}(z)\|$ $z = 0.1$

(Eldèn-S.,'09)

Some references as starting point

1. A. C. Antoulas, *Approximation of large-scale Dynamical Systems*, SIAM, 2005
2. S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, To appear in GAMM Mitteilungen, 2013.
3. N. J. Higham, *Matrix Functions – Theory and Applications*, SIAM, 2008
4. M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numerica, 2010, 19, 209-286.