Analysis of projection-type methods for approximating the matrix exponential operator

V. Simoncini

Dipartimento di Matematica, Università di Bologna
and CIRSA, Ravenna, Italy
valeria@dm.unibo.it

Joint work with Luciano Lopez, Bari

Matrix rational function approximation problem

Determine $x_m \in \mathcal{K}_m$ that approximates the solution x of

$$\Psi_\nu(A)x = \Phi_\mu(A)v$$

A sym. negative semidef. $\mathcal{K}_m$ approx. space, dim($\mathcal{K}_m$)=m

Ψ_ν, Φ_μ polynomials of degree ν, μ resp.

****

Motivation: Approximation to the exponential operator

$$\exp(A)v \approx (\Psi_\nu(A))^{-1}\Phi_\mu(A)v$$

Other functions: $(\Psi_\nu(\lambda))^{-1}\Phi_\mu(\lambda) \approx \lambda^{\frac{1}{2}}, \cos(\lambda), \ldots$

Approximation to the exponential operator

Used in large range of applications
(e.g. within ODEs and time-dependent PDEs)

$$\exp(A)v \approx (\Psi_\nu(A))^{-1}\Phi_\mu(A)v$$

- Polynomial approximation, $\nu = 0$
- Padè (rational f.) approximation, e.g., $\mu = \nu$
- Chebyshev (rational f.) approximation, $\mu = \nu$
- RD (rational f.) approximation
- ...

Focus: Large matrix dimension

Approximation using Krylov subspace

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

V_m s.t. range$(V_m) = \mathcal{K}_m(A, v)$, and $V_m^* V_m = I$

Arnoldi relation

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^*$$

A common approach

$$\exp(A)v \approx V_m \exp(H_m) e_1, \qquad \|v\| = 1$$

Approximation of $\exp(A)v$ in Krylov subspace. I

Typical convergence bounds (Hochbruck & Lubich '97)

$$\begin{aligned}
\|\exp(A)v - V_m \exp(H_m)e_1\| &\leq 10e^{-m^2/(5\rho)}, \qquad \sqrt{4\rho} \leq m \leq 2\rho, \\
\|\exp(A)v - V_m \exp(H_m)e_1\| &\leq \frac{10}{\rho}e^{-\rho}\left(\frac{e\rho}{m}\right)^m, \qquad m \geq 2\rho
\end{aligned}$$

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

see also Druskin & Knizhnerman '89, Stewart & Leyk '96

Predict superlinear convergence

Approximation of $\exp(A)v$ in Krylov subspace. II

Typical a-posteriori estimate (see, e.g., Saad '92)

$$\|\exp(A)v - V_m \exp(H_m)e_1\| \approx O(h_{m+1,m}|e_m^* \exp(H_m)e_1|)$$

Note: for $Ax(t) + x'(t) = 0$, $x(0) = v$

$$h_{m+1,m}|e_m^* \exp(tH_m)e_1| = \|Ax_m(t) + x'_m(t)\|$$

plays role of residual norm
(Druskin & Greenbaum & Knizhnerman '98)

Exploring Krylov subspace approximation

$$\exp(A)v \approx V_m \exp(H_m)e_1, \qquad \|v\| = 1$$

$$\exp(\lambda) \approx \frac{\Phi_\nu(\lambda)}{\Psi_\nu(\lambda)} \qquad \text{Rational function approx}$$

- Increase our understanding of approximation in $\mathcal{K}_m(A, v)$
- Analyze role of "residual" $\quad h_{m+1,m}|e_m^* \exp(H_m)e_1|$
- Set up the stage for acceleration procedures

Projection of Rational functions onto Krylov subspaces

Basic fact: ($\mathcal{R}_\nu = \Phi_\nu/\Psi_\nu$)

$$
\begin{aligned}
&\|\exp(A)v - V_m \exp(H_m)e_1\| \leq \\
&\qquad \|\exp(A)v - \mathcal{R}_\nu(A)v\| + \|\mathcal{R}_\nu(A)v - V_m\mathcal{R}_\nu(H_m)e_1\| \\
&\qquad +\|V_m\left(\mathcal{R}_\nu(H_m)e_1 - \exp(H_m)e_1\right)\|.
\end{aligned}
$$

Focus: $\mathcal{R}_\nu$ Padè and Chebyshev approximation

($\Psi_\nu(A)$ positive definite)

Projection onto Krylov subspace

$$x_\star = (\Psi_\nu(A))^{-1}\Phi_\nu(A)v \quad \Leftrightarrow \quad x_\star \text{ solves } \quad \Psi_\nu(A)x = \Phi_\nu(A)v$$

Range(V_m) $= \mathcal{K}_m(A, v)$. Galerkin approximation:

$$\text{Solve} \quad V_m^*\Psi_\nu(A)V_m y = V_m^*\Phi_\nu(A)v, \qquad x_m^G = V_m y_m^G$$

Minimization property:

$$\min_{x\in K_m(A,v)} \|x_\star - x\|_{\Psi_\nu(A)} = \|x_\star - x_m^G\|_{\Psi_\nu(A)}$$

Linear bounds for convergence rate

Using Partial Fraction expansion: $\frac{\Phi_\nu(\lambda)}{\Psi_\nu(\lambda)} = \tau_0 + \sum_{j=1}^{\nu} \frac{\tau_j}{\lambda - \xi_j}$

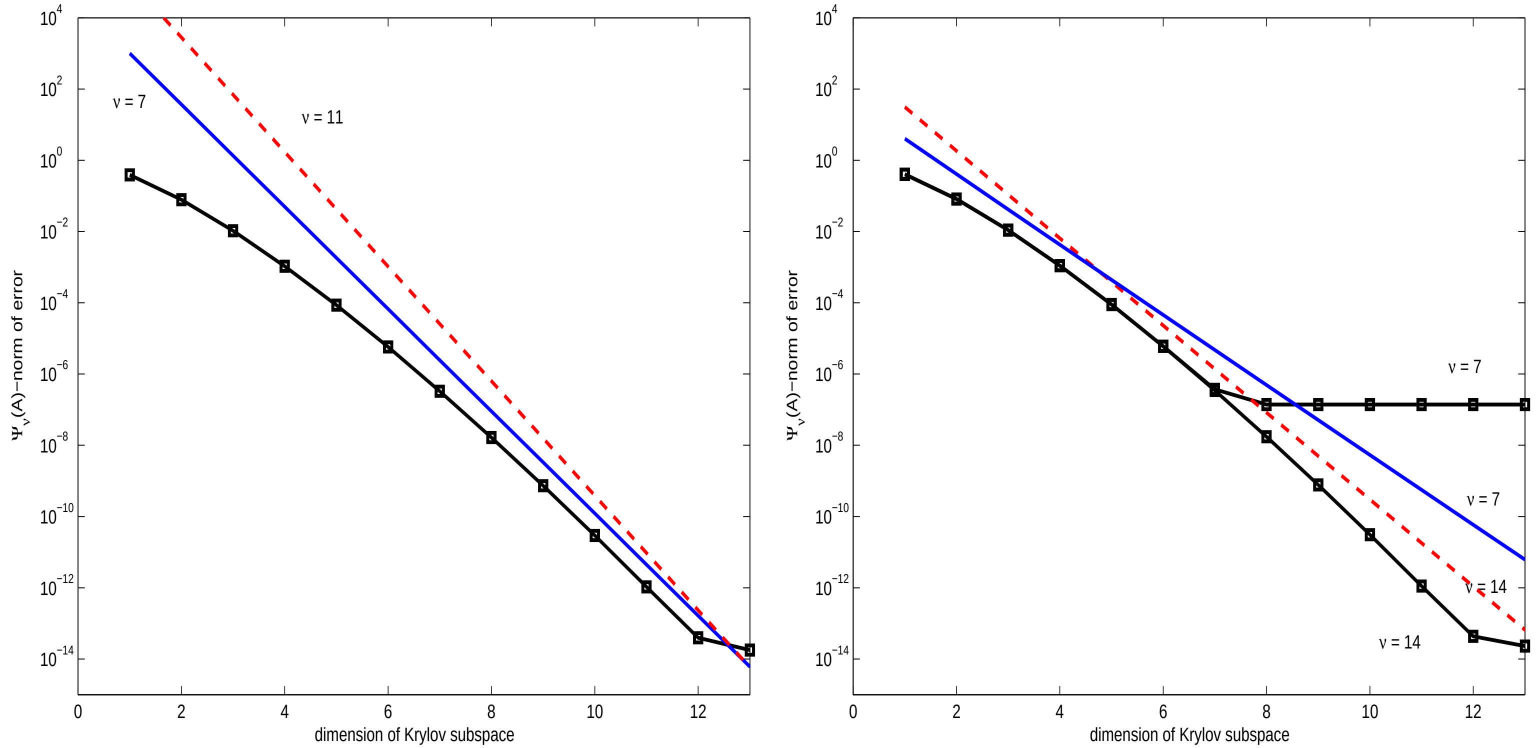
$$x_\star = (\Psi_\nu(A))^{-1}\Phi_\nu(A)_\nu v = \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$$

Convergence bound: $\sigma(A) \subseteq [\alpha, \beta]$

$$\frac{\|x_\star - x_m^G\|_{\Psi_\nu(A)}}{\|x_\star - x_0^G\|_{\Psi_\nu(A)}} \leq 2 \sum_{j=1}^{\nu} \left(\max_{\lambda \in [\alpha,\beta]} \frac{|\tau_j \Psi_\nu(\lambda)|}{|\Phi_\nu(\lambda)|\,|\lambda - \xi_j|} \right) \frac{1}{\rho_j^m + 1/\rho_j^m}$$

$\rho_j = \rho_j(\alpha, \beta, \xi_j)$

Galerkin approximation



`A = diag(log(linspace(0.2,0.99,100)))`, $v = \mathbf{1}$

Left: Padè and upper bound for $\nu = 7, 11$

Right: Chebyshev and upper bounds for $\nu = 7, 14$

Krylov approximation

$$x_\star = (\Psi_\nu(A))^{-1}\Phi_\nu(A)v \qquad \approx \qquad V_m y_m^K = V_m(\Psi_\nu(H_m))^{-1}\Phi_\nu(H_m)e_1$$

$V_m y_m^K$ is a term-wise Galerkin projection: (van der Vorst, '87)

$$\begin{aligned}
x_\star &= \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v \approx \tau_0 v + \sum_{j=1}^{\nu} \tau_j V_m y_m^{(j)} \\
&= V_m \left(\tau_0 e_1 + \sum_{j=1}^{\nu} \tau_j (H_m - \xi_j I)^{-1} e_1 \right) \\
&= V_m (\Psi_\nu(H_m))^{-1} \Phi_\nu(H_m) e_1 \equiv V_m y_m^K
\end{aligned}$$

A-posteriori estimate and residual

$$x_\star = \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v \approx V_m \left(\tau_0 e_1 + \sum_{j=1}^{\nu} \tau_j (H_m - \xi_j I)^{-1} e_1 \right)$$

Defining $\quad r_m^K := \sum_{j=1}^{\nu} \tau_j r_m^{(j)} \quad$ ($r_m^{(j)}$ single residuals) we have

$$\|r_m^K\| = h_{m+1,m} |e_m^* y_m^K|$$

Comparison with Galerkin approximation

If $m > \nu$, then

$$\|y_m^G - y_m^K\| \leq \gamma \; \|(y_m^K)_{m-\nu+1:m}\|, \qquad \gamma = O(h_{m+1,m}^2)$$

where

$$|e_k^* y_m^K| \leq \sum_{j=1}^{\nu} \frac{|\tau_j|}{\sigma_{\min}(H_m - \xi_j I)} \|r_{k-1}^{(j)}\|, \qquad 1 < k \leq m,$$

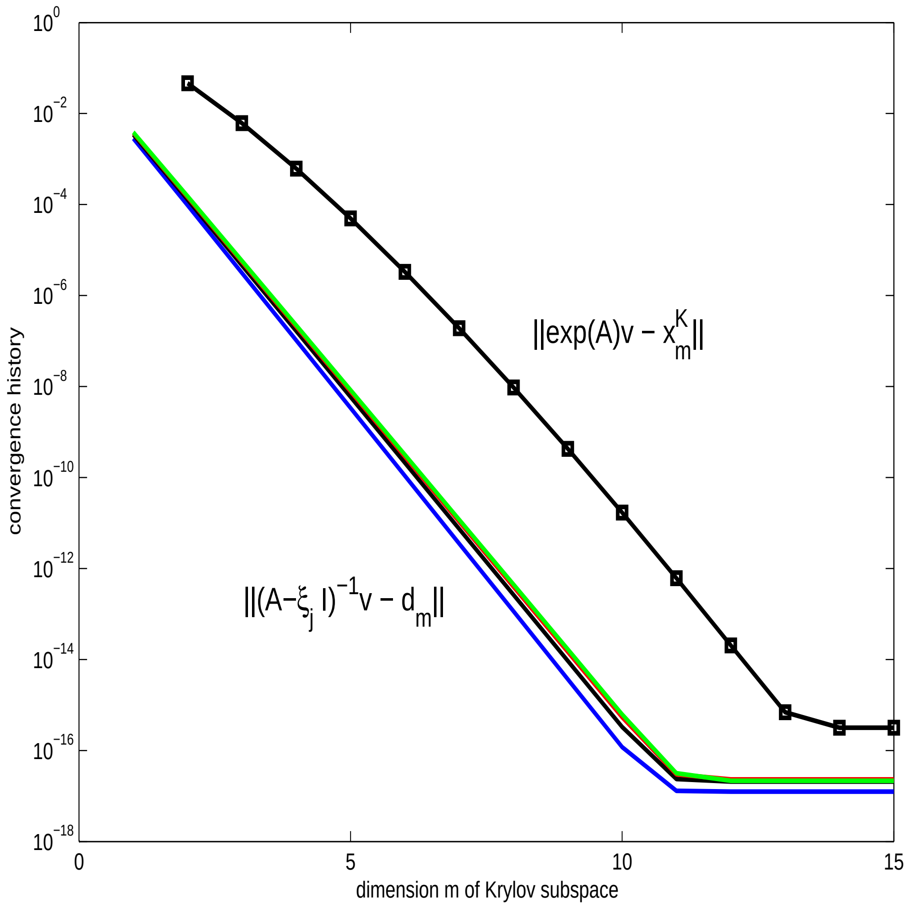
$r_{k-1}^{(j)}$ residual of system $(A - \xi_j I)x = v$ after $k-1$ iterations

τ_j partial fraction coeff's

$\sigma_{\min}(\cdot)$ smallest singular value

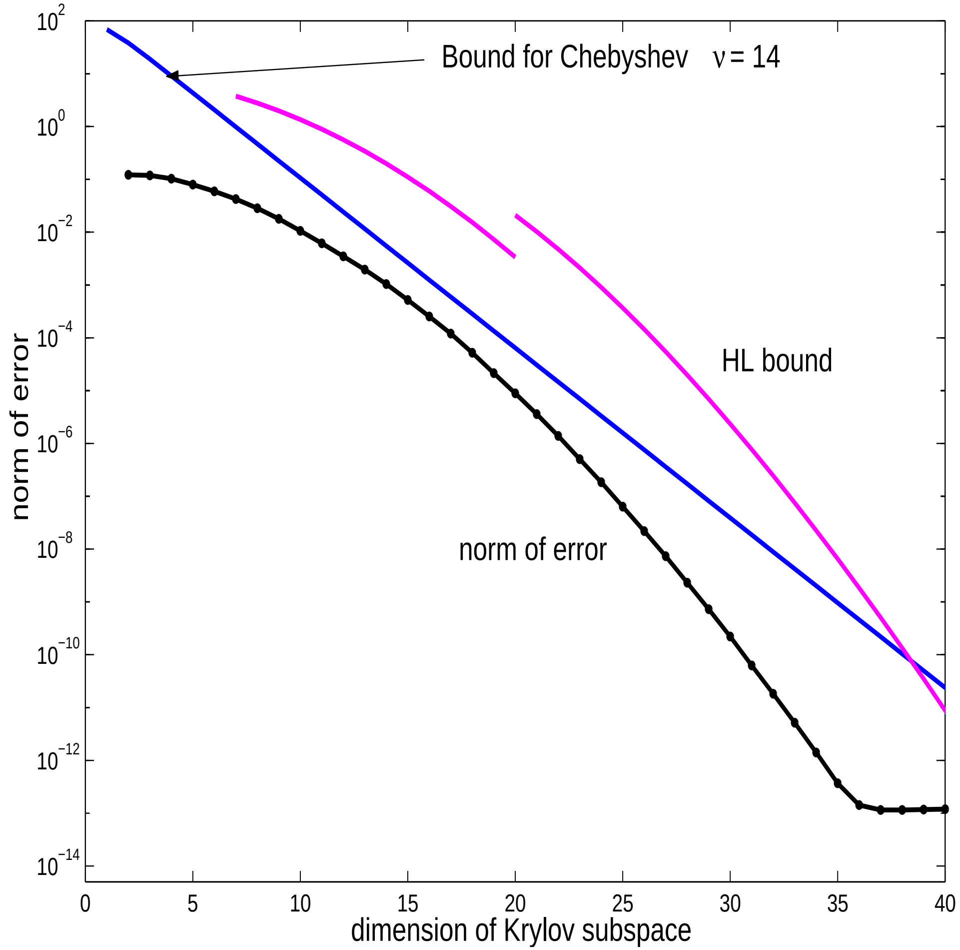
⋆ Similar convergence estimates as for Galerkin

Relation to convergence of systems $(A - \xi_j I)x = v, j = 1, \ldots, \nu$



(Padè, $\nu = 7$)

One more example



$A \in \mathbb{R}^{1001 \times 1001}$, diagonal, uniform random distr. in $[-40, 0]$

Conclusions and Outlook

1. Convergence of $(A - \xi_j I)x = v$ plays a role
2. Preconditioning strategies
3. Generalization to non-Hermitian case

Paper available at: `www.dm.unibo.it/~simoncin`

e-mail: `valeria@dm.unibo.it`