# Approximation of functions of large matrices.
# Part II. Application Problems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

and CIRSA, Ravenna, Italy

valeria@dm.unibo.it

## Problem in context

Given $v \in \mathbb{R}^n$ and $A$ symmetric and negative semidefinite, approximate

$$x = \exp(A)v$$

Wide range of applications. Here we focus on

- Flows on constraint manifolds, e.g.,

$$Q_t = H(Q,t)Q, \quad Q(t)|_{t=0} = Q_0 \in V_k(\mathbb{R}^n)$$

  $V_k$ Stiefel manifold

- (Analysis of) Low dimensional models of dynamical systems: approximate solution to Lyapunov equation

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

## Structure preserving approaches

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

$Q_{n+1} = \exp(tA_n)Q_n$ 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 $\mathcal{K}_m(A, Q^{(0)})$  $Q^{(0)} = [q_1^{(0)}, \ldots, q_k^{(0)}]$

$$\mathcal{K}_m(A, Q^{(0)}) = \mathrm{span}\{Q^{(0)}, AQ^{(0)}, \ldots, A^{m-1}Q^{(0)}\}$$

- $\mathcal{V}_m$ 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

## 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:

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

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

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

## An example

Linear Hamiltonian system: $\begin{cases} q' = Aq & \quad A = J^{-1}S \\ q(0) = q_0 & \end{cases}$

with $S \in \mathbb{R}^{400\times 400}$ symmetric (eigs. in [1, 100])

Energy function: $E(Q(t)) = Q(t)^T S Q(t)$, constant for all $t > 0$

---

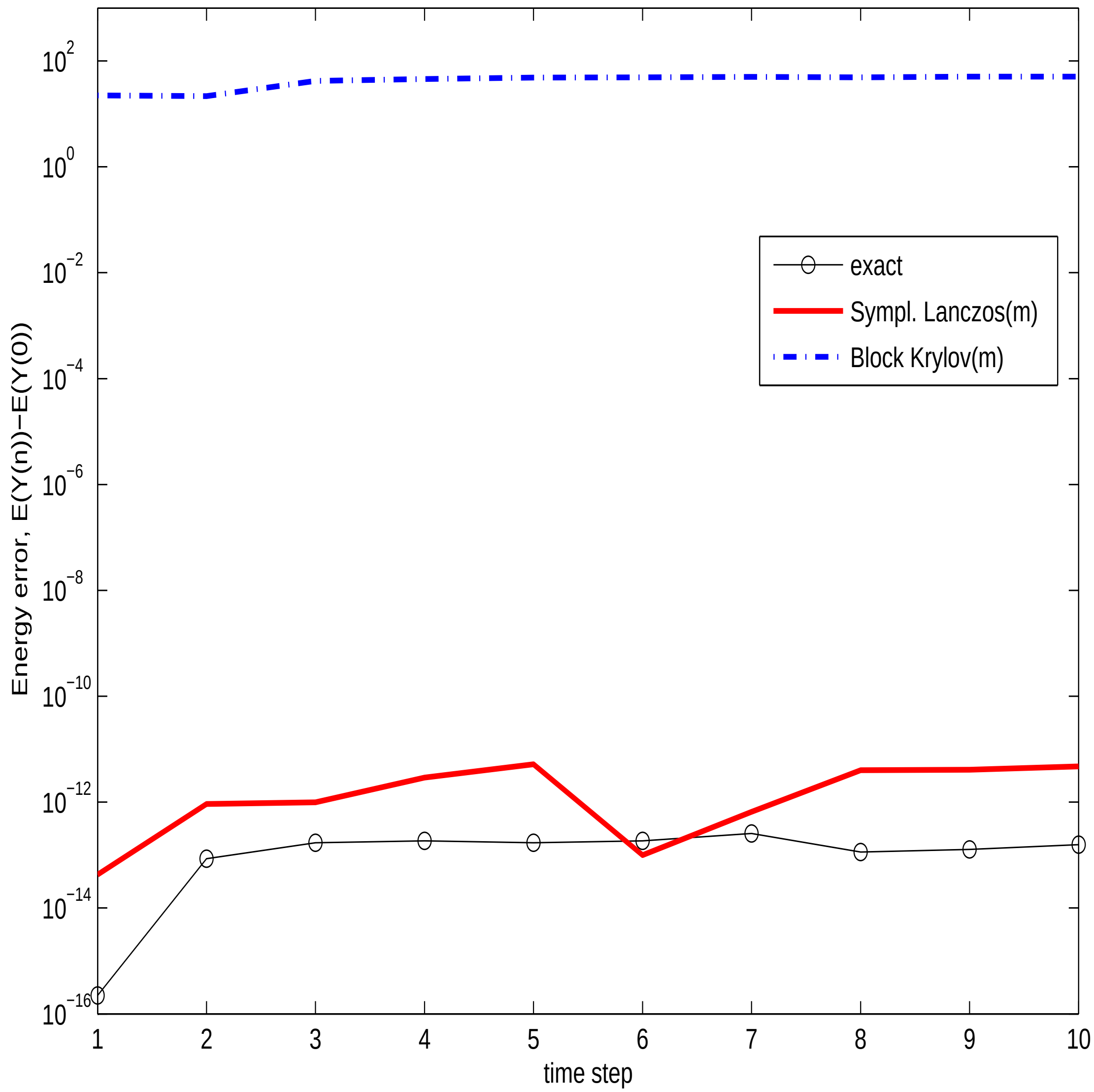
Numerical symplectic integrator: starting with $Q^{(0)} = Q_0$,

$$Q^{(r+1)} = \exp(hA)Q^{(r)}, \qquad r \geq 0 \qquad h = \frac{1}{40}$$

$\star$ $X_m = \exp(hA)Q^{(r)}$ standard Krylov subspace approximation

$\Rightarrow$ energy function is **not** constant, unless $X_m$ is accurate

## Conservation of energy. Error: $|E(Q^{(r)}) - E(Q_0)|$



## Outlook

Implementation: For $A$ Hamiltonian,

⋆ Control Stability

⋆ Robustness: (quasi) breakdown ?

Theoretical issues:

- Convergence properties
- Further generalizations

L. Lopez & Simoncini, *Preserving Geometric Properties of the Exponential matrix by block Krylov subspace methods*, BIT, 2006.

## Solving the Lyapunov equation. The problem

Approximate $X$ in:

$$AX + XA^\top + BB^\top = 0$$

$A \in \mathbb{R}^{n\times n}$ pos.real $\qquad B \in \mathbb{R}^{n\times s} \qquad$ here: $B = b$ $(s = 1)$

## Solving the Lyapunov equation. The problem

Approximate $X$ in:

$$AX + XA^\top + BB^\top = 0$$

$A \in \mathbb{R}^{n\times n}$ pos.real $\qquad B \in \mathbb{R}^{n\times s} \qquad$ here: $B = b$ $(s = 1)$

---

Applications: signal processing, system and control theory

Time-invariant linear system:

$$\mathbf{x}'(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \qquad \mathbf{x}(0) = x_0$$

Analytic solution:

$$X = \int_0^\infty e^{-tA}BB^\top e^{-tA^\top} dt = \int_0^\infty xx^\top dt \quad \text{with} x = \exp(-tA)B.$$

see, e.g., Antoulas 2005, Benner 2006

## Standard Krylov subspace projection

$$X \approx X_m \qquad X_m \in \mathcal{K}$$

Galerkin condition: $R := AX_m + X_m A^\top + bb^\top \quad \perp \quad \mathcal{K}$

$$V_m^\top R V_m = 0 \qquad\qquad \mathcal{K} = \text{range}(V_m)$$

## Standard Krylov subspace projection

$$X \approx X_m \qquad X_m \in \mathcal{K}$$

Galerkin condition: $R := AX_m + X_m A^\top + bb^\top \quad \perp \quad \mathcal{K}$

$$V_m^\top R V_m = 0 \qquad\qquad \mathcal{K} = \mathrm{range}(V_m)$$

---

Assume $V_m^\top V_m = I_m$ and let $X_m := V_m Y_m V_m^\top$.

Projected Lyapunov equation:

$$(V_m^\top A V_m) Y_m + Y_m (V_m^\top A^\top V_m) \quad + \quad V_m^\top bb^\top V_m = 0$$

$$\Updownarrow$$

$$T_m Y_m + Y_m T_m^\top \quad + \quad e_1 e_1 = 0$$

with $b = V_m e_1$ (Saad, 1990, for $\mathcal{K} = \mathcal{K}_m(A, b)$; Jaimoukha & Kasenally, 1994)

## Extended Krylov subspace method

Galerkin condition: $\mathcal{X}_m \in \mathcal{K}$ s.t.

$$R := A\mathcal{X}_m + \mathcal{X}_m A^\top + bb^\top \quad \perp \quad \mathcal{K}$$

$$\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B), \qquad \text{range}(\mathcal{V}_m) = \mathcal{K}$$

(Druskin-Knizhnerman 1998, S., 2007)

Projected Lyapunov equation:

$$(\mathcal{V}_m^\top A \mathcal{V}_m)Y_m + Y_m(\mathcal{V}_m^\top A^\top \mathcal{V}_m) \quad + \quad \mathcal{V}_m^\top bb^\top \mathcal{V}_m = 0$$

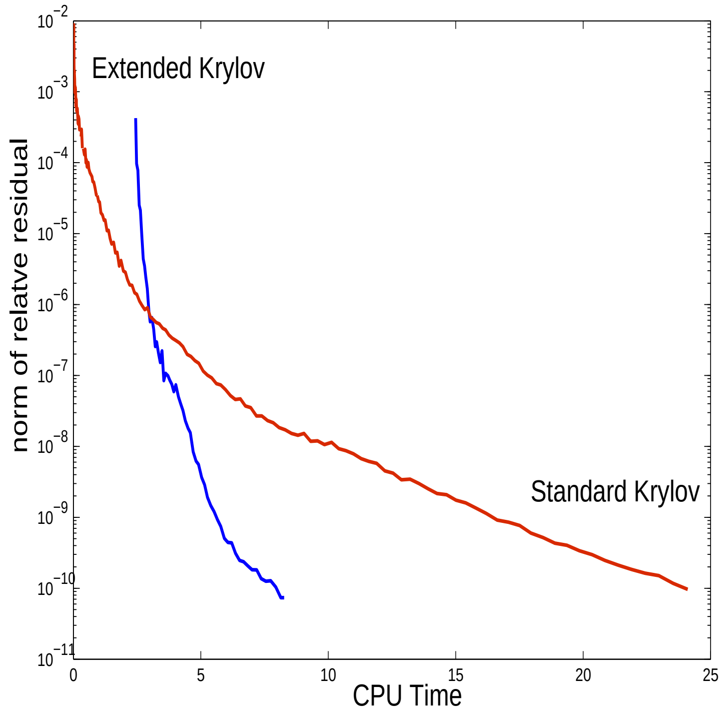
$$\Updownarrow$$

$$\mathcal{T}_m Y_m + Y_m \mathcal{T}_m^\top \quad + \quad e_1 e_1 = 0$$

## Performance evaluation

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

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

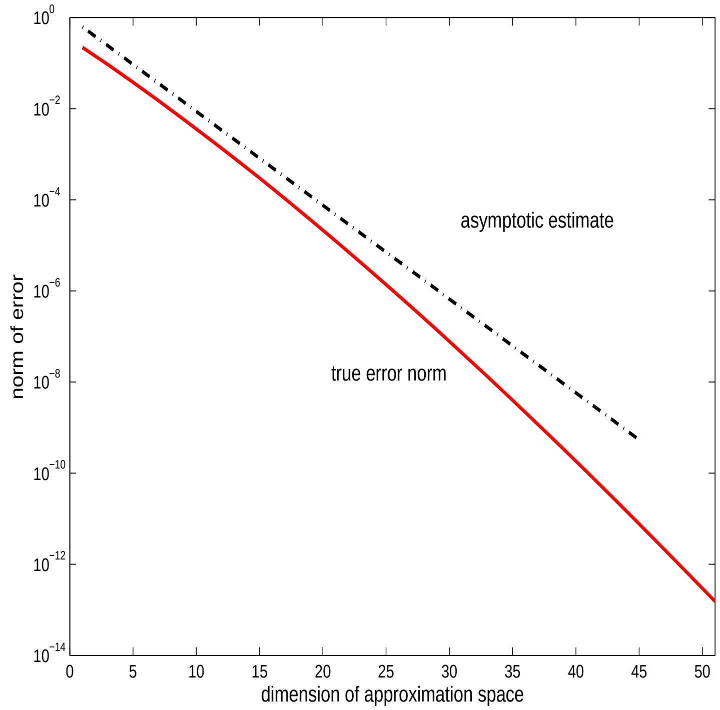



approximation space dim.: 146 (Standard Krylov) 112 (Extended Krylov)

## Convergence analysis of Extended Krylov: $A$ symmetric

Conjecture: $\|X - \mathcal{X}_m\| \approx \left(\frac{\hat\kappa^{1/4}-1}{\hat\kappa^{1/4}+1}\right)^m$, $m = 1, 2, \ldots,$

$\hat\kappa =$cond$(A + \lambda_{\min} I)$




Idea: $[\hat\lambda_{\min}, \hat\lambda_{\max}] = [\hat\lambda_{\min}, \chi] \cup [\chi, \hat\lambda_{\max}]$

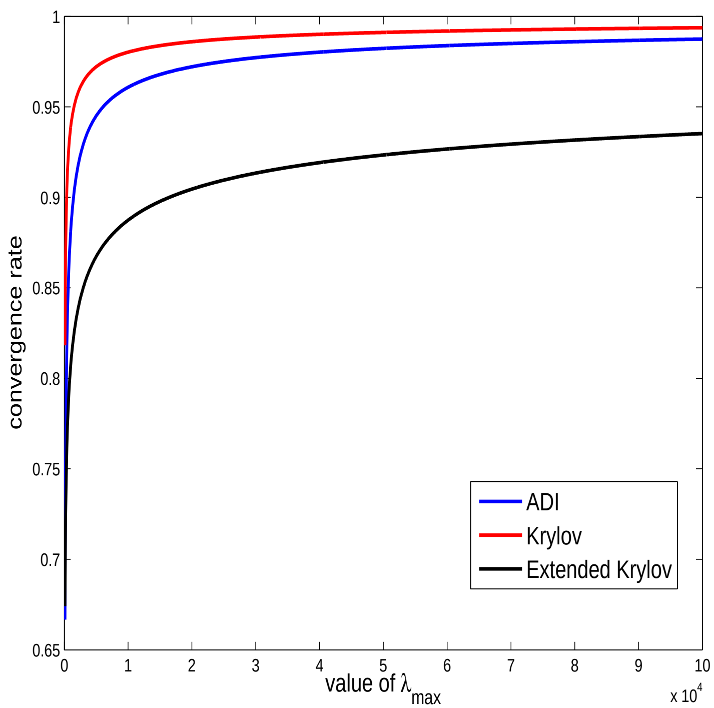
$\chi$ s.t. cond$([\hat\lambda_{\min}, \chi])$ = cond$([\chi, \hat\lambda_{\max}])$

# Comparison of convergence rates: $A$ symmetric

ADI iteration: $\varepsilon_{adi,j} \approx \left(\frac{\sqrt{\kappa_{adi}}-2}{\sqrt{\kappa_{adi}}+2}\right)^j$

Standard Krylov: $\varepsilon_{kr,j} \approx \left(\frac{\sqrt{\hat{\kappa}}-1}{\sqrt{\hat{\kappa}}-1}\right)^j$

Extended Krylov: $\varepsilon_{ek,\ell} \approx \left(\left(\frac{\sqrt[4]{\hat{\kappa}}-1}{\sqrt[4]{\hat{\kappa}}-1}\right)^{1/2}\right)^{\ell}$




$\lambda_{\min} = 1,\ \lambda_{\max} \in [10^2, 10^5]$

## Conclusions and future work

- Good understanding of Standard Krylov method
- Better understanding of convergence

⋆ Extended Krylov: $\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B)$ prove conjectures

⋆ Connection with the convergence theory of other methods

⋆ New acceleration procedures

V.Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*. *SIAM J. Sci. Comput.*, 29(3):1268–1288, 2007.

V. Simoncini and V. Druskin, *Convergence analysis of projection methods for the numerical solution of large Lyapunov equations*. August 2007.

Available at `www.dm.unibo.it/~simoncin`