Rational Krylov subspace methods for large-scale Lyapunov equations

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

joint work with Vladimir Druskin and Leonid Knizhnerman (Schlumberger-Doll Research)

Model Order Reduction

Given the continuous time-invariant linear system

$$\mathbf{x}'(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \qquad \mathbf{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & \end{array}\right), \; A \in \mathbb{C}^{n\times n}$$

$$\mathbf{y}(t) = C\mathbf{x}(t), \quad \mathbf{x}(0) = x_0$$

Analyse the construction of a reduced system

$$\hat{\mathbf{\Sigma}} = \left(\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \end{array}\right)$$

with $\tilde{A}$ of size $m \ll n$

Projection methods and Linear Dynamical Systems

Time-invariant linear system:

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

Emphasis: A large dimensions, $W(A) \subset \mathbb{C}^-$

Projection methods: the general idea

Given space $K \subset \mathbb{R}^n$ of size m and (orthonormal) basis V_m,

$$A \to A_m = V_m^* A V_m, \quad B \to B_m = V_m^* B, \quad C \to C_m = C V_m$$

Reduced problem uses: A_m, B_m, C_m

The Lyapunov matrix equation

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

Galerkin approximation by projection:

K of dim. m, V_m orthonormal basis. $X \approx X_m = V_m Y V_m^*$

$$R_m \perp K \qquad \Leftrightarrow \qquad V_m^* R_m V_m = 0$$

that is,

$$V_m^* A V_m Y + Y V_m^* A^* V_m + V_m^* B B^* V_m = 0$$

Small size equation. Solved with dense methods.

Projection methods: the general idea

Choices of K in the literature:

- Standard Krylov subspace: $K_m(A, B) = \text{span}\{B, AB, \ldots, A^{m-1}B\}$

Projection methods: the general idea

Choices of K in the literature:

- Standard Krylov subspace: $K_m(A,B) = \mathrm{span}\{B, AB, \ldots, A^{m-1}B\}$
- Shift-Invert Krylov subspace:
  $K_m((A-\sigma I)^{-1}, B) = \mathrm{span}\{B, (A-\sigma I)^{-1}B, \ldots, (A-\sigma I)^{-(m-1)}B\}$;
  often $\sigma = 0$

Projection methods: the general idea

Choices of K in the literature:

- Standard Krylov subspace: $K_m(A,B) = \mathrm{span}\{B, AB, \ldots, A^{m-1}B\}$
- Shift-Invert Krylov subspace:
  $K_m((A-\sigma I)^{-1}, B) = \mathrm{span}\{B, (A-\sigma I)^{-1}B, \ldots, (A-\sigma I)^{-(m-1)}B\}$;
  often $\sigma = 0$
- Extended Krylov subspace:
  $\mathbf{EK}_m(A,B) = K_m(A,B) + K_m(A^{-1}, A^{-1}B)$

Projection methods: the general idea

Choices of K in the literature:

- Standard Krylov subspace: $K_m(A,B) = \mathrm{span}\{B, AB, \ldots, A^{m-1}B\}$
- Shift-Invert Krylov subspace:
  $K_m((A-\sigma I)^{-1}, B) = \mathrm{span}\{B, (A-\sigma I)^{-1}B, \ldots, (A-\sigma I)^{-(m-1)}B\}$;
  often $\sigma = 0$
- Extended Krylov subspace:
  $\mathbf{EK}_m(A,B) = K_m(A,B) + K_m(A^{-1}, A^{-1}B)$
- Rational Krylov subspace:
  $K_m(A, B, \mathbf{s}) = \mathrm{span}\{(A - s_1 I)^{-1}B, (A - s_2 I)^{-1}B, \ldots, (A - s_m I)^{-1}B\}$
  usually $\mathbf{s} = [s_1, \ldots, s_m]$ a-priori

Rational Krylov Subspaces. A long tradition...

$$K_m(A, B, \mathbf{s}) = \text{span}\{(A - s_1 I)^{-1}B, (A - s_2 I)^{-1}B, \ldots, (A - s_m I)^{-1}B\}$$

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

Rational Krylov Subspaces in MOR. Choice of poles.

$$K_m(A,B,\mathbf{s}) = \text{span}\{(A-s_1 I)^{-1}B, (A-s_2 I)^{-1}B, \ldots, (A-s_m I)^{-1}B\}$$

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)

A new adaptive choice of poles for RKS. $B = b \in \mathbb{C}^n$

$$K_m(A, b, \mathbf{s}) = \mathrm{span}\{(A - s_1 I)^{-1} b, (A - s_2 I)^{-1} b, \ldots, (A - s_m I)^{-1} b\}$$

$\mathbf{s} = [s_1, \ldots, s_m]$ to be chosen sequentially

The fundamental idea: Assume you wish to solve

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

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

$$b - (A - sI)x_m = \frac{r_m(A)b}{r_m(s)}, \qquad r_m(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{z - s_j}$$

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

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

A new adaptive choice of poles for RKS. Cont'd

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

For A symmetric:

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

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

For A nonsymmetric:

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

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

Approximation by projection

$X \approx X_m = V_m Y V_m^*$ with

$$(V_m^* A V_m) Y + Y (V_m^* A V_m)^* + (V_m^* b)(V_m^* b)^* = 0$$

Some technical issues:

- $K_m(A, b, \mathbf{s}) = \mathrm{span}\{b, (A - s_2 I)^{-1} b, \ldots, \prod_{j=2}^{m} (A - s_j I)^{-1} b\}$ (includes b)
- All real poles (all real arithmetic work)
- Cheap computation of $V_m^* A V_m$ at each iteration m $(K_m(A, b, \mathbf{s}) \subseteq K_{m+1}(A, b, \mathbf{s}))$
- Cheap computation of the residual norm

$$\|R_m\| = \|A X_m + X_m A^* + b b^*\|$$

- Cheap factorized form of solution $X_m = X_{\hat{m}} := Z_{\hat{m}} Z_{\hat{m}}^*$, $\hat{m} \le m$

Some numerical experiments

Closest competitors:

- ADI – problem: computation of parameters
- Extended Krylov Subspace method – outperforms ADI in general

$$\mathbf{EK}_m(A, b) = K_m(A, b) + K_m(A^{-1}, A^{-1}b)$$

Comparison measures:

- Efficiency (CPU time)
- Memory (space dimension)
- Rank of solution

The RAIL (symmetric) data set

n		Rational space direct	Extended space direct		
1357	CPU time (s)	0.84	0.36		
	dim. Approx. Space	21	64		
	Rank of Solution	21	47		
20209	CPU time (s)	11.19	10.97		
	dim. Approx. Space	25	124		
	Rank of Solution	25	75		
79841	CPU time (s)	51.54	73.03		
	dim. Approx. Space	26	168		
	Rank of Solution	26	103		

The RAIL (symmetric) data set

n		Rational space direct	Extended space direct	Rational space iterative	Extended space iterative
1357	CPU time (s)	0.84	0.36	0.96	1.60
	dim. Approx. Space	21	64	21	68
	Rank of Solution	21	47	21	45
20209	CPU time (s)	11.19	10.97	25.31	201.94
	dim. Approx. Space	25	124	25	126
	Rank of Solution	25	75	25	75
79841	CPU time (s)	51.54	73.03	189.48	2779.95
	dim. Approx. Space	26	168	26	170
	Rank of Solution	26	103	26	98

Inner solves: PCG with IC(10^{-2})

More Tests: two nonsymmetric problems

n		Rational space direct	Extended space direct	Rational space iterative	Extended space iterative
9669	CPU time (s)	3.16	3.06	3.01	9.95
	dim. Approx. Space	16	36	16	36
	Rank of Solution	16	24	16	24
20082	CPU time (s)	59.99	45.84	13.01	25.28
	dim. Approx. Space	15	26	15	26
	Rank of Solution	15	22	15	22

Convective thermal flow problems (FLOW, CHIP data sets)

⋆ All real shifts used

ADI and Rational Krylov subspaces

Main consideration (see, e.g., Li, Wright 2000)

$$X_m^{(ADI)} \in K_m(A, B, \mathbf{s})$$

and also, for $U_m = [(A - s_1 I)^{-1} b, \ldots, (A - s_m I)^{-1} b]$,

$$X_m^{(ADI)} = U_m \boldsymbol{\alpha}^{-1} U_m^*$$

with $\boldsymbol{\alpha}$ Cauchy matrix (Druskin, Knizhnerman, Simoncini 2010)

ADI and Rational Krylov subspaces

Main consideration (see, e.g., Li, Wright 2000)

$$X_m^{(ADI)} \in K_m(A, B, \mathbf{s})$$

and also, for $U_m = [(A - s_1 I)^{-1} b, \ldots, (A - s_m I)^{-1} b]$,

$$X_m^{(ADI)} = U_m \boldsymbol{\alpha}^{-1} U_m^*$$

with $\boldsymbol{\alpha}$ Cauchy matrix (Druskin, Knizhnerman, Simoncini 2010)

Equivalence of ADI with generic RKSM:

ADI coincides with the Galerkin solution X_m in Rational Krylov space if and only if

$$s_j = -\bar{\lambda}_j$$

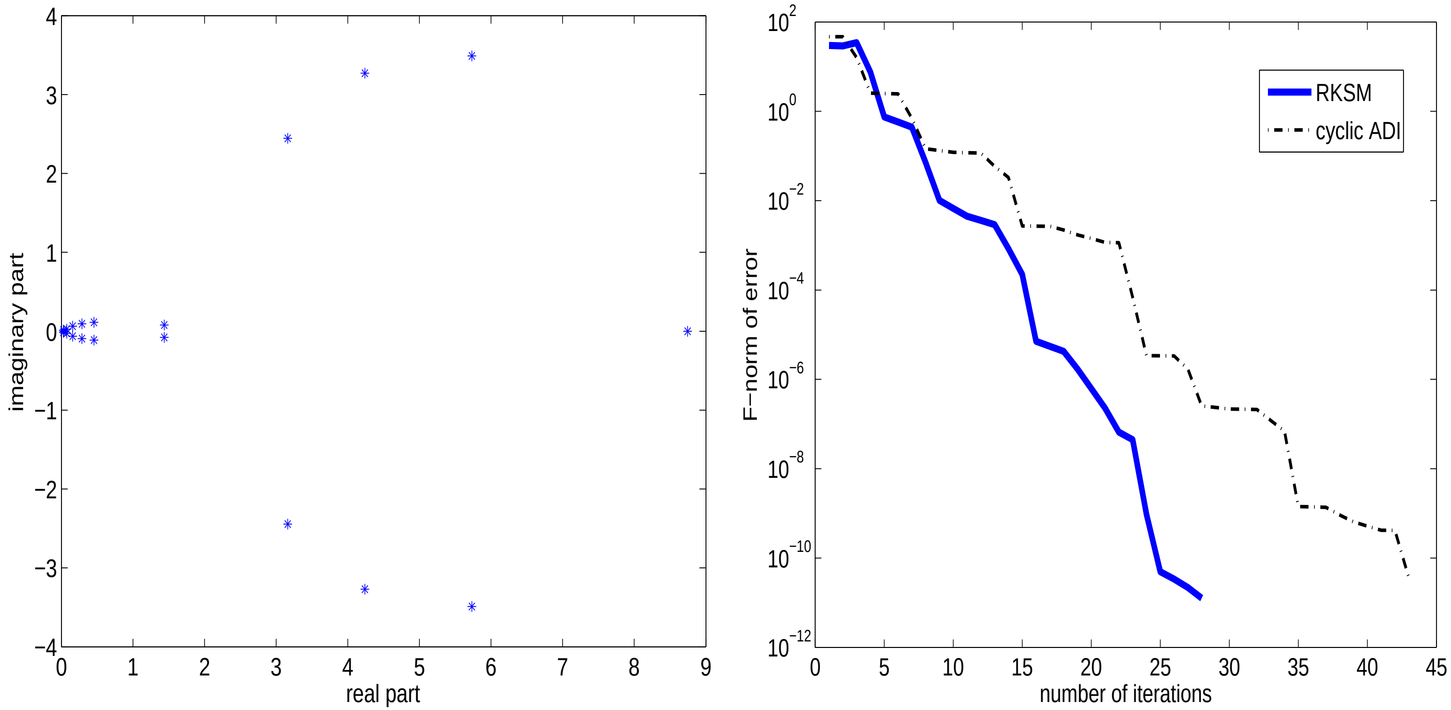
where $\lambda_j = \mathrm{eigs}(V_m^* A V_m)$ Ritz values (suitably ordered)

Druskin, Knizhnerman, Simoncini 2010,

see also Flagg 2009 for $\Leftarrow$, and Beckermann 2011

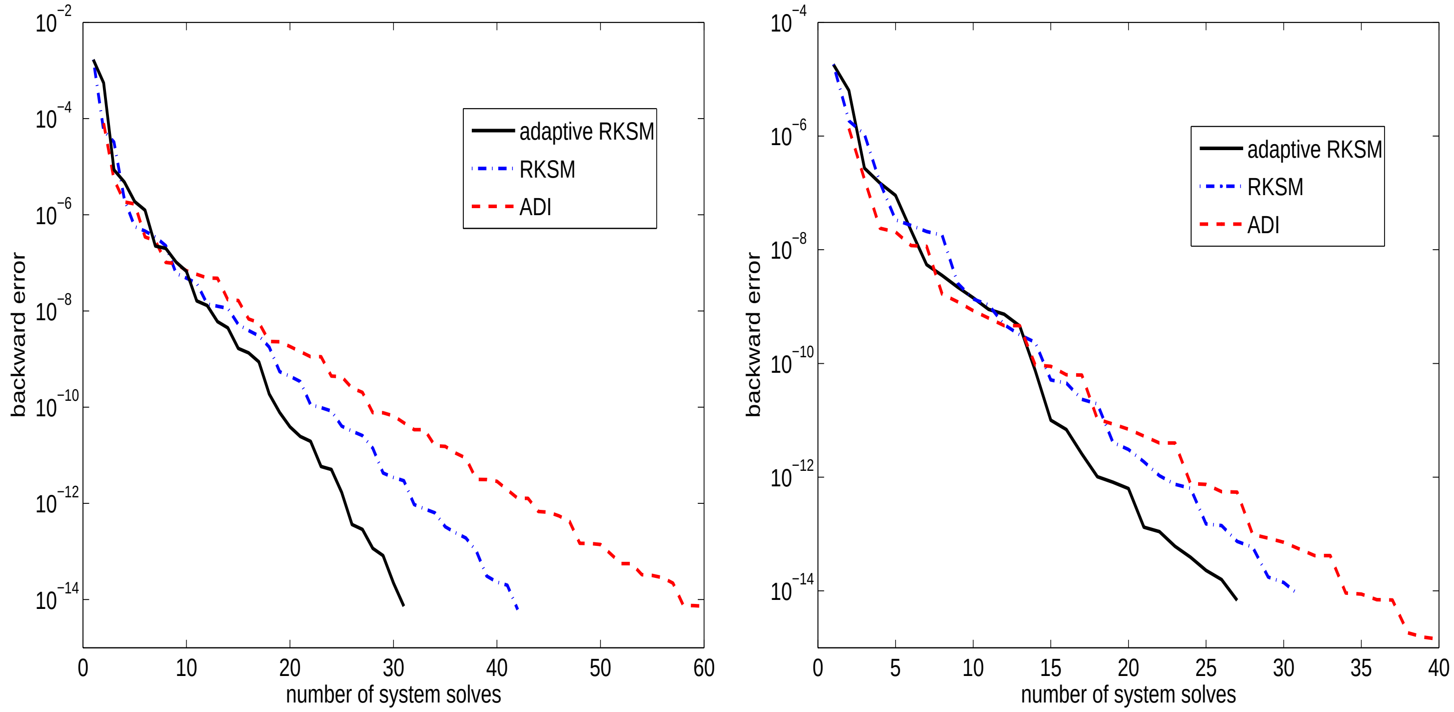
Typical behavior of ADI and generic RKSM for the same poles

Operator: $L(u) = -\Delta u + (50xu_x)_x + (50yu_y)_y$ on $[0,1]^2$



Same non-optimal 20 poles, repeated cyclically.

Expected performance (from Oberwolfach Collection)



Left: `rail` problem, A symmetric.

Right: `flow_meter_model_v0.5` problem, A nonsymmetric.

ADI and RKSM use 10 non-optimal poles cyclically (computed a-priori with `lyapack`, Penzl 2000)

Other recent approaches and convergence results

- Galerkin-Projection Accelerated ADI (Benner, Saak, tr 2010)

Different aims:

- IRKA (Gugercin, Antoulas, Beattie, 2008)
- Riemann optimization approach (Vandereycken, Vandewalle, 2010)

****

Convergence analysis:

- For the Lyapunov equation (Knizhnerman 2010)
- Improved, and for the Sylvester equation (Beckermann 2011)

Conclusions and outlook

- Adaptive procedure makes Rational Krylov subspace appealing
- Competitive in terms of both reduction space and CPU time
- MOR: $b = B$? Balanced Truncation?

`www.dm.unibo.it/~simoncin` `valeria@dm.unibo.it`

References:

⋆ *V. Simoncini, A new iterative method for solving large-scale Lyapunov matrix equations. SIAM J. Scient. Computing (2007)*

⋆ *V. Druskin and V. Simoncini, Adaptive rational Krylov subspaces for large-scale dynamical systems. Systems & Control Letters, (2011)*

⋆ *V. Druskin, L. Knizhnerman and V. Simoncini, Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation. (tr 2010)*