The rational Krylov subspace for parameter dependent systems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

Motivation. Model Order Reduction

Given the continuous-time system

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} A & B \\ \hline C & \end{array}\right), \quad A \in \mathbb{R}^{n\times n},\ B, C^T \text{ tall}$$

Analyse the construction of a reduced system

$$\widetilde{\boldsymbol{\Sigma}} = \left(\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \end{array}\right), \qquad \tilde{A} \text{ of size } m \ll n$$

so that all relevant properties are captured by $\widetilde{\boldsymbol{\Sigma}}$

Motivation (cont'd): Linear Dynamical Systems

Time-invariant linear system:

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

- Approximation of parameter-dependent systems
- Approximation of the matrix Transfer function
- Other problems (e.g., Matrix equations)

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

Projection methods: the general idea

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

$$\begin{aligned} A &\rightarrow A_m = V_m^T A V_m \\ B &\rightarrow B_m = V_m^T B \\ C &\rightarrow C_m = C V_m \end{aligned}$$

$$\widetilde{\mathbf{\Sigma}} = \left(\begin{array}{c|c} A_m & B_m \\ \hline C_m & \end{array} \right)$$

Projection methods: the space

Choices of K in the literature:

Projection methods: the space

Choices of K in the literature:

- Standard Krylov subspace:
  $K_m(A, B) = \mathrm{Range}\{[B, AB, \ldots, A^{m-1}B]\}$

Projection methods: the space

Choices of K in the literature:

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

Projection methods: the space

Choices of K in the literature:

- Standard Krylov subspace:
  $K_m(A, B) = \mathrm{Range}\{[B, AB, \ldots, A^{m-1}B]\}$
- Shift-Invert Krylov subspace: $K_m((A - \sigma I)^{-1}, B) = \mathrm{Range}\{[B, (A - \sigma I)^{-1}B, \ldots, (A - \sigma I)^{-(m-1)}B]\}$; 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 space

Choices of K in the literature:

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

  $K_m(A, B, \mathbf{s}) = \mathrm{Range}\{[(A - s_1 I)^{-1}B, (A - s_2 I)^{-1}B, \ldots, (A - s_m I)^{-1}B]\}$

  in the past, $\mathbf{s} = [s_1, \ldots, s_m]$ a-priori

Parameter-dependent linear systems

$$x = f_\sigma(A)b \quad \approx \quad \widetilde{x} \in K$$

Parameter-dependent linear systems

$$x = f_\sigma(A)b \quad \approx \quad \widetilde{x} \in K$$

Here, $f_\sigma(\lambda) = (\lambda + \sigma)^{-1}$. A different perspective:

- Shifted linear systems. Many shifts in a wide range

  (e.g., Structural dynamics, electromagn.)

$$(A + \sigma_j I)x = v, \qquad \sigma_j \in [\alpha, \beta], \quad \text{large interval}, j = O(100)$$

- Few (possibly complex) shifts (e.g., quadrature formulas)

$$z = \sum_{j=1}^{k} \omega_j (A - \sigma_j I)^{-1} v$$

- Transfer function

$$h(\sigma) = c^T (A - i\sigma I)^{-1} b, \quad \sigma \in [\alpha, \beta]$$

Shifted systems

Approximation. Defining $\mathcal{T}_m = V_m^* A V_m$,

$$x \approx x_m = V_m f_\sigma(\mathcal{T}_m) e_1 = V_m(\mathcal{T}_m + \sigma I)^{-1} e_1$$

Galerkin condition: $r_m := b - (A + \sigma I)x_m \perp K$.

(standard Galerkin-type approximation for shifted systems, cf. FOM, CG, ...)

Key fact: A single K for all shifted systems.

Note: Solve systems with A to approximate $(A + \sigma I)^{-1} v$

Added feature: restarting made easy

A numerical example

Matrix A discretization of $\quad \mathcal{L}(u) = -\Delta u + 50(x+y)(u_x + u_y)$

$b = \mathbf{1}$, 500 shift in $[0, 5]$

("backslash" takes 2.52 secs for each system of size 160 000)

n	subspace dimension	std Krylov CPU time (#cycles)	Extended Krylov CPU time (#cycles)
2500	10	1.50 (41)	0.65 (4)
2500	20	1.57 (14)	0.63 (2)
10 000	10	3.23 (49)	1.96 (7)
10 000	20	4.35 (27)	1.75 (3)
160 000	10	399.80 (>300)	62.81 (13)
160 000	20	356.63 (97)	109.87 (6)

Sylvester-type equations

In shifted systems, the right-hand side may depend on the parameter:
$B = [b(\sigma_1), \ldots, b(\sigma_s)]$

$$AX + XS = B, \qquad S = \mathrm{diag}(\sigma_1, \ldots, \sigma_s)$$

(special case of more general Sylvester equation)

Generate approximation space starting with $B = V_1\beta_0$

(V_1 orthnormal columns, β_0 possibly rectangular)

A numerical example. Direct frequency analysis in structural dynamics

$$(K^\star - \sigma^2 M)x(\sigma) = b(\sigma), \qquad \text{size } 3627$$

$K^\star$ complex symmetric, $B = [b(\sigma_1), \ldots, b(\sigma_s)]$ of rank 2.

Simultaneous solution:

$$K^\star X + MXS = B, \quad S = -\text{diag}(\sigma_1^2, \ldots, \sigma_s^2).$$

For M real sym nonsingular, $M = LL^T$ and

$$(L^{-1}K^\star L^{-T})\widetilde{X} + \widetilde{X}S = L^{-1}B, \qquad \widetilde{X} = UX$$

with $L^{-1}K^\star L^{-T}$ complex symmetric $\Rightarrow$ bilinear form $x^T y$

A numerical example. Direct frequency analysis in structural dynamics

Comparison with: $Q(\sigma)x(\sigma) = b(\sigma)$, $Q(\sigma) = K^\star - \sigma^2 M$ for each σ,

(complex symmetric CG method, preconditioned with LDL$^\top$ with drop. tol. 10^{-2})

$\#\sigma$	Sparse Direct Solver	Complex sym CG	Extended Krylov Direct solves $K^\star$	Extended Krylov Iter.solves $K^\star$
50	12.57	48.30 (108)	1.59 (36)	8.00 (36)
100	24.88	96.33 (108)	2.12 (40)	9.69 (40)

Transfer function approximation

A with field of values in $\mathbb{C}^-$

$$h(\omega) = c(A - i\omega I)^{-1}b, \qquad \omega \in [\alpha, \beta]$$

Given space K of size m and V_m s.t. $K=\text{range}(V_m)$,

$$h(\omega) \approx cV_m(V_m^T A V_m - i\omega I)^{-1}(V_m^T b)$$

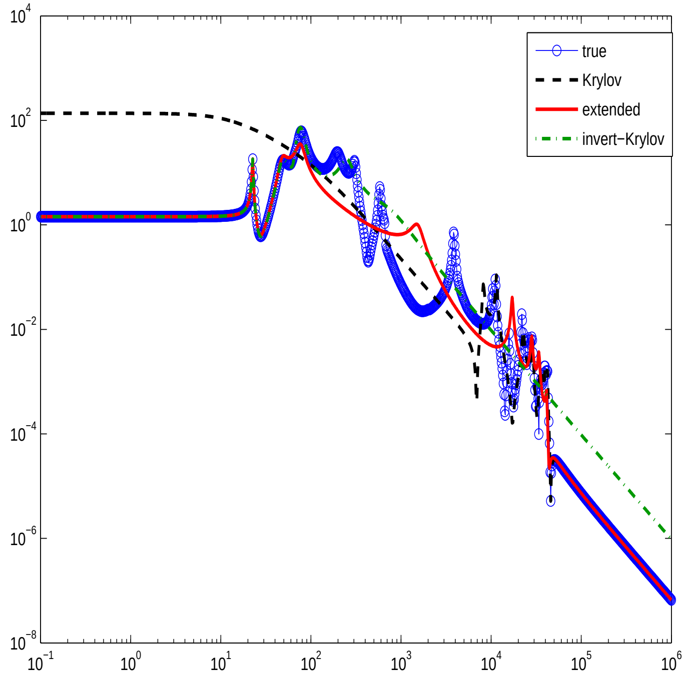
Next:

Classical benchmark experiment with Standard, Shift-invert and Extended Krylov

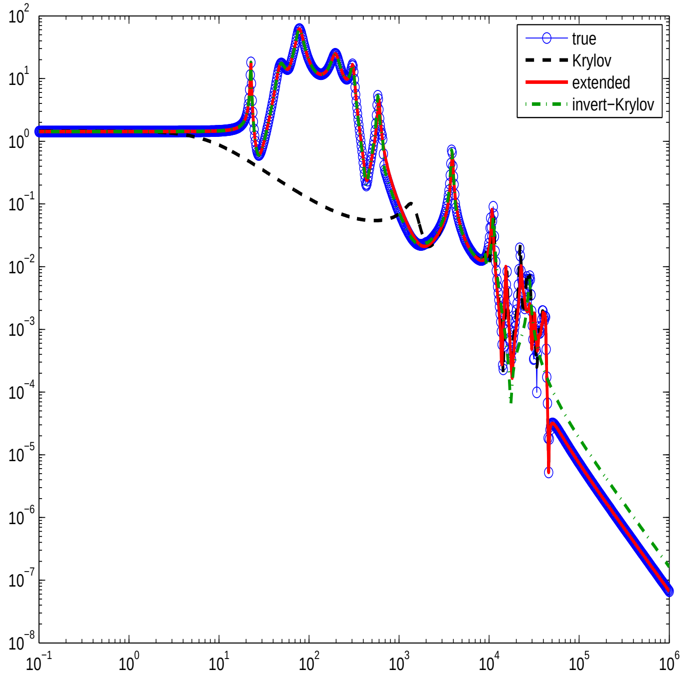
An example: CD Player, $n = 120$

$$|h(\omega)| = |C_{2,:}(A - i\omega I)^{-1}B_{:,1}|$$

$m = 20$



$m = 50$



Rational Krylov Subspace Method. Choice of poles

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

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 RKSM

$$K_m(A,b,\mathbf{s}) = \mathrm{span}\{(A-s_1I)^{-1}b, (A-s_2I)^{-1}b, \ldots, (A-s_mI)^{-1}b\}$$

$\mathbf{s} = [s_1, \ldots, s_m]$ to be chosen sequentially (no fixed m)

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

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

A new adaptive choice of poles for RKSM. 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:

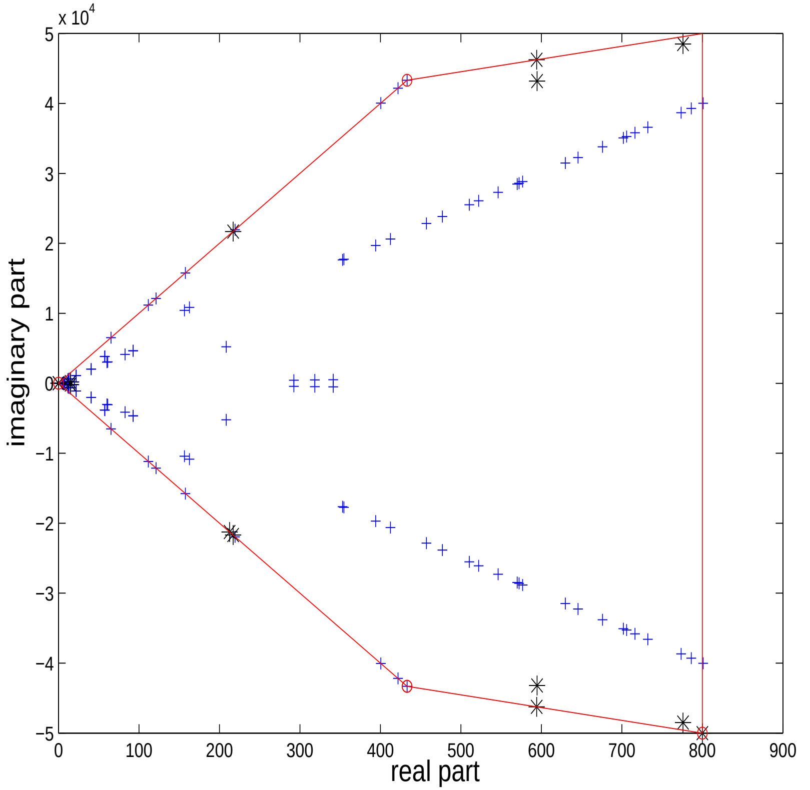
$$s_{m+1} := \arg\left(\max_{s \in \partial \mathcal{S}_m} \frac{1}{|r_m(s)|} \right)$$

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

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

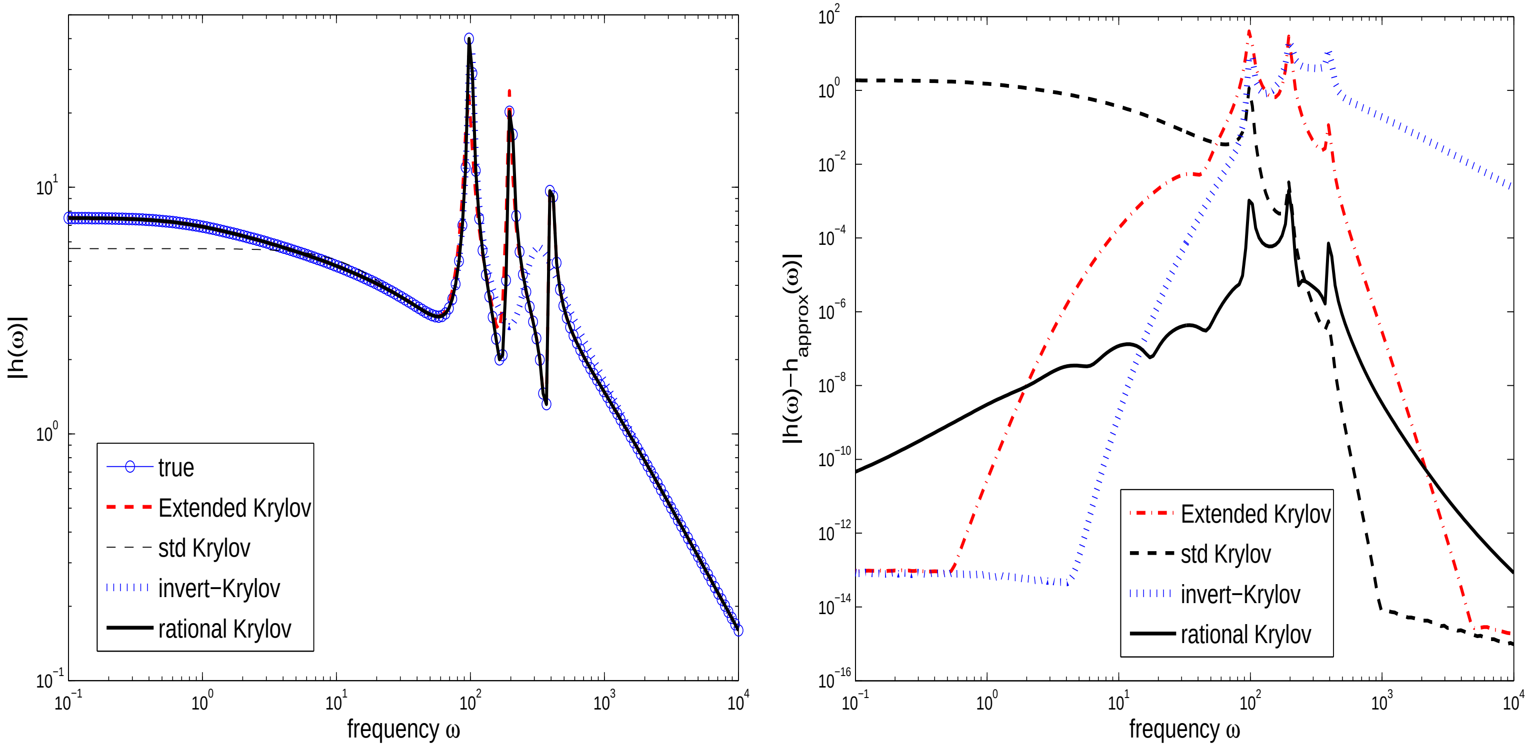
$\mathcal{S}_m$: encloses mirrored current Ritz values: -eigs$(V_m^T A V_m)$

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



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

Transfer function evaluation. FOM Data set

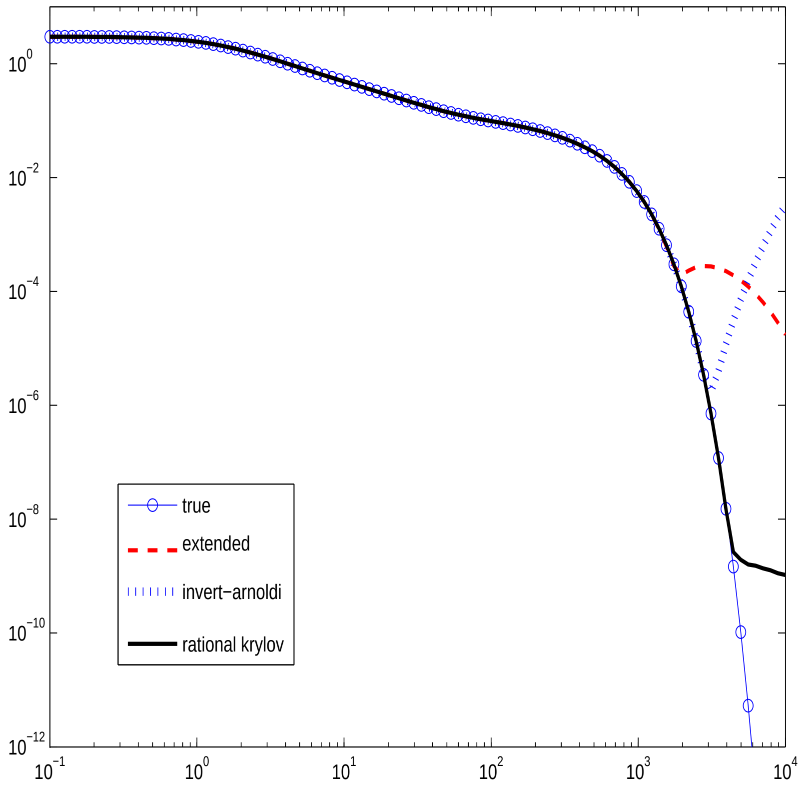


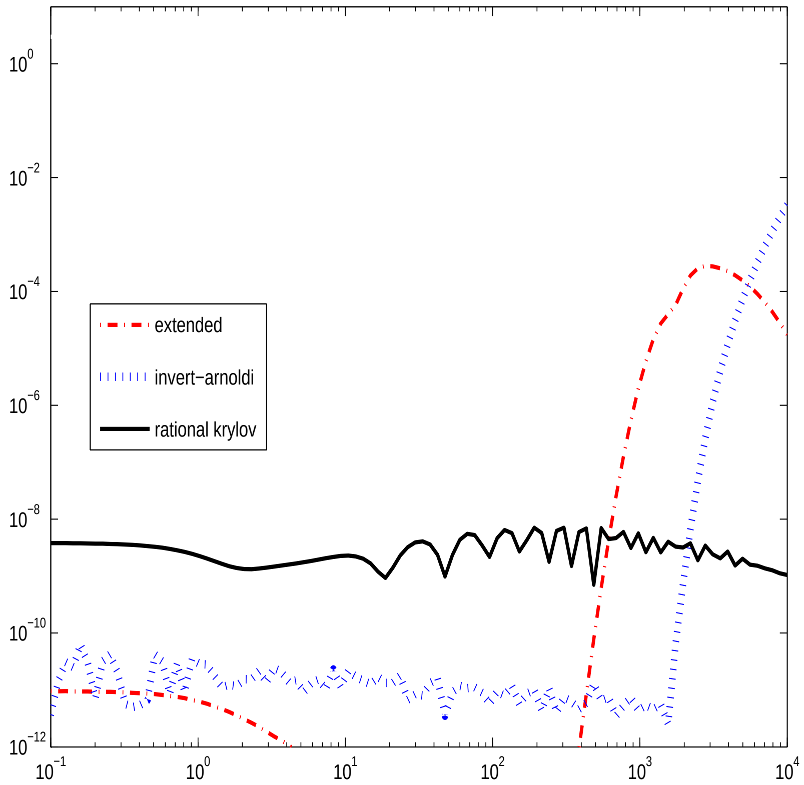
$m = 20$. A of size 1006 (normal matrix),

$s_0^{(1)} = 1$, $s_0^{(2)} = 1000$

Transfer function evaluation. Flow Data set. $h(\omega) = c(A - i\omega E)^{-1}b$

$|h(\omega)|$ $\qquad$ $|h(\omega) - h_{approx}(\omega)|$





$m = 20$. A, E of size 9669 (nonsym) $s_0^{(1)} = \|A\|/(\text{condest(A)}\|E\|_F)$,
$s_0^{(1)} \approx \arg(\max \Re(\text{eigs}(A)))$ (all real poles)

Understanding the Rational Krylov subspace. State of the art.

- Shift selection.

  - Historically, a-priori selection (possibly with pre-processing)

  - Given A symmetric, asymptotically optimal rational space for $i\mathbb{R}$ by using an equidistributed nested sequence of real shifts

  (from classical Zolotaryov sol'n)

  Druskin, Knizhnerman, Zaslavsky (SISC 2009)

  - For $f(A)b$, variants of adaptive selection (Güttel-Knizhnerman, tr 2012)

- Convergence Analysis. Mostly very recent, based on classical potential theory.

  (Beckermann, Druskin, Eiermann, Ernst, Güttel, Knizhnerman, Lieberman, Reichel, Simoncini, Vandebril, Zaslavsky)

Projection methods for matrix equations

$$\mathcal{F}(X) = 0$$

linear or quadratic equations, such as:

- $AX + XA^T + BB^T = 0$ Lyapunov equation
- $AX + XB + C = 0$ Sylvester equation
- $AX + XA^T - XBB^TX + C^TC = 0$ Riccati equation

With an approximation in the space K with basis $V_m \otimes V_m$,

(Use the Kronecker formulation for the derivation only)

$$\text{residual} \perp K \qquad \Leftrightarrow \qquad V_m^T \mathcal{F}(X_m) V_m = 0$$

where X_m is the matrix approximation

The Lyapunov matrix equation

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

Approximation by projection: K of dim. m, V_m orthonormal basis.
$X \approx X_m = V_m Y V_m^T$

$$(V_m^T A V_m)Y + Y(V_m^T A^T V_m) + V_m^T BB^T V_m = 0$$

With K being the Krylov subspace (Saad, '90, Jaimoukha & Kasenally, '94), the Extended Krylov subspace (Simoncini, '07), the adaptive Rational Krylov subspace (Druskin & Simoncini, '11)

The Lyapunov equation. Cont'd.

Long term Competitor:

- Alternating Directions Implicit iteration (ADI) (large number of contributions, see, e.g., Benner, Penzl, etc.)

  problem: computation of parameters

⋆ Extended Krylov Subspace method numerically outperforms ADI in general

⋆ Rational Krylov Subspace is at least as good as ADI

(Druskin, Knizhnerman, Simoncini, '11, Beckermann '11, theoretical arguments)

Convergence Analysis

Much is now known on convergence (in the last 5 years!)

Conclusions and outlook

- Projection-type methods still have great potential
- "Second" generation Krylov spaces can attack harder problem than standard linear systems
- Quadratic problems (Heyouni, Jbilou, '09).

  New very promising results for Riccati eqn (Simoncini, Szyld)

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

References:

⋆ *V. Simoncini, A new iterative method for solving large-scale Lyapunov matrix equations. SIAM J. Scient. Comput., v.29, n.3 (2007), pp. 1268-1288*

⋆ *V. Simoncini, The Extended Krylov subspace for parameter dependent systems. Applied Num. Math. v.60 n.5 (2010) 550-560.*

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