On the numerical solution of large scale Riccati equation

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

Joint works with M. Monsalve, Y.Lin, D. Szyld

The problem

Find $X \in \mathbb{R}^{n\times n}$ such that

$$AX + XA^\top - XBB^\top X + C^\top C = 0$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{s\times n}$, $p, s = \mathcal{O}(1)$

A rich literature for numerical methods:

Lancaster-Rodman 1995, Bini-Iannazzo-Meini 2012, ...

The problem

Find $X \in \mathbb{R}^{n\times n}$ such that

$$AX + XA^\top - XBB^\top X + C^\top C = 0$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{s\times n}$, $p, s = \mathcal{O}(1)$

A rich literature for numerical methods:

Lancaster-Rodman 1995, Bini-Iannazzo-Meini 2012, ...

We focus on the large scale case: $n \gg 1000$

- (Inexact) Kleinman iteration (Newton-type method)
- Projection methods
- Invariant subspace iteration
- (Sparse) multilevel methods
- ....

Kleinman iteration

Assume A stable. Compute sequence $\{X_k\}$ with $X_k \rightarrow_{k\rightarrow\infty} X$

1: Given $X_0 \in \mathbb{R}^{n\times n}$ such that $X_0 = X_0^\top$, $A^\top - BB^\top X_0$ is stable

2: **For** $k = 0, 1, \ldots,$ until convergence

3: **Set** $\mathcal{A}_k^\top = A^\top - BB^\top X_k$

4: **Set** $\mathcal{C}_k^\top = [X_k B, \;\; C^\top]$

5: **Solve** $\mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k = 0$

Kleinman iteration

Assume A stable. Compute sequence $\{X_k\}$ with $X_k \to_{k\to\infty} X$

1: Given $X_0 \in \mathbb{R}^{n\times n}$ such that $X_0 = X_0^\top$, $A^\top - BB^\top X_0$ is stable.
2: **For** $k = 0, 1, \ldots,$ until convergence
3: **Set** $\mathcal{A}_k^\top = A^\top - BB^\top X_k$
4: **Set** $\mathcal{C}_k^\top = [X_k B, \;\; C^\top]$
5: **Solve** $\mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k = 0$

Critical issues:

- The full matrix X_k cannot be stored (sparse or low-rank approx)
- Cheap stopping criterion with
  $\mathfrak{R}(X_{k+1}) := AX_{k+1} + X_{k+1}A^\top - X_{k+1}BB^\top X_{k+1} + C^\top C$
- Each iteration k requires the solution of the Lyapunov equation:

$$\mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k = 0$$

(Benner, Feitzinger, Hylla, Saak, Sachs, ...)

Inexact Kleinman iteration. Computation of the residual norm.

• Solve the Lyapunov equation only approximately ($inexactly$). That is, X_{k+1} s.t.

$$\mathfrak{L}(X_{k+1}) := \mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k \approx 0$$

(Lyapunov residual matrix)

It holds that:

$$\mathfrak{R}(X_{k+1}) = \mathfrak{L}(X_{k+1}) - (X_k B - X_{k+1}B)(X_k B - X_{k+1}B)^\top$$

(see also, Saak, previous Workshop, Aachen 2011) so that, in general,

$$\|\mathfrak{R}(X_{k+1})\|_F \leq \|\mathfrak{L}(X_{k+1})\|_F + \|X_k B - X_{k+1}B\|_F^2$$

Inexact Kleinman iteration. Computation of the residual norm.

Assume that for each k, the solution to the Lyapunov equation

$$\mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k = 0$$

is approximated by means of a Galerkin projection method. Then

- The Riccati residual matrix satisfies

$$\|\mathfrak{R}(X_{k+1})\|_F^2 = \|\mathfrak{L}(X_{k+1})\|_F^2 + \|X_k B - X_{k+1} B\|_F^4$$

Inexact Kleinman iteration. Computation of the residual norm.

Assume that for each k, the solution to the Lyapunov equation

$$\mathcal{A}_k X_{k+1} + X_{k+1}\mathcal{A}_k^\top + \mathcal{C}_k^\top \mathcal{C}_k = 0$$

is approximated by means of a Galerkin projection method. Then

- The Riccati residual matrix satisfies

$$\|\mathfrak{R}(X_{k+1})\|_F^2 = \|\mathfrak{L}(X_{k+1})\|_F^2 + \|X_k B - X_{k+1}B\|_F^4$$

- The norm computation is cheap:

$$\|\mathfrak{R}(X_{k+1})\|_F^2 = \|\mathfrak{L}(X_{k+1})\|_F^2 + \|(G^\top - Y_m^{(L)} V_m^\top B)\|_F^4$$

  where V_m basis for Galerkin projection space of dim. $\mathcal{O}(m)$.
  Here: $G^\top = V_m^\top X_k B$
  $Y_m^{(L)}$ sol'n to reduced Lyapunov pb; both are $\mathcal{O}(k)$ matrices

Inexact Kleinman iteration. Algorithmic considerations.

Assume Galerkin procedure is used for inner (Lyapunov) equation. E.g., projection space

$$K_m(A, C^\top) = \mathrm{range}([C^\top, AC^\top, \ldots, A^{m-1}C^\top])$$

- At iteration $k = 0$, for $X_0 = 0$, solve

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

Approximate solution: $X \approx X_m = V_m Y_m V_m^\top$, where Y_m solves

$$(V_m^\top A V_m)Y + Y(V_m^\top A^\top V_m) + V_m^\top C^\top C V_m = 0$$

$\Rightarrow$ Richer spaces: Extended, Rational, augmented Krylov, etc.

Galerkin projection method for the Riccati equation

Given the basis V_k for an approximation space, determine approx

$$X_k = V_k Y_k V_k^\top$$

to the Riccati solution matrix by orthogonal projection:

$$V_k^\top \mathfrak{R}(X_k) V_k = 0$$

(Galerkin condition), giving

$$(V_k^\top A V_k) Y + Y (V_k^\top A^\top V_k) - Y_k (V_k^\top B B^\top V_k) Y_k + (V_k^\top C^\top)(C V_k) = 0$$

(Heyouni-Jbilou 2009)

Galerkin projection method for the Riccati equation

Given the basis V_k for an approximation space, determine approx

$$X_k = V_k Y_k V_k^\top$$

to the Riccati solution matrix by orthogonal projection:

$$V_k^\top \mathfrak{R}(X_k) V_k = 0$$

(Galerkin condition), giving

$$(V_k^\top A V_k) Y + Y (V_k^\top A^\top V_k) - Y_k (V_k^\top B B^\top V_k) Y_k + (V_k^\top C^\top)(C V_k) = 0$$

(Heyouni-Jbilou 2009)

Key questions:

- Which approximation space?
- What expected performance, compared with Galerkin method on the Lyapunov equation?

Performance of solvers

Problem: A: 3D Laplace operator, B, C `randn` matrices, tol$=10^{-8}$

$(n, p, s) = (125000, 5, 5)$

	its	inner its	time	space dim	rank(X_f)
Newton $X_0 = 0$	15	5, . . . , 5	808	100	95
Newton $X_0 = X_4^{eksm}$	10	5, . . . , 5	706	100	94
GP-EKSM	20		531	200	105
GP-RKSM	25		524	125	105

$(n, p, s) = (125000, 20, 20)$

	its	inner its	time	space dim	rank(X_f)
Newton $X_0 = 0$	19	5, . . . , 5	2332	400	346
Newton $X_0 = X_4^{eksm}$	15	5, . . . , 5	2042	400	347
GP-EKSM	15		622	600	364
GP-RKSM	20		720	400	358

GP=Galerkin projection

Some matrix relations

$X_k^{(R)}$: Galerkin approx to Riccati equation in Range(V_k)

$X_k^{(L)}$: Galerkin approx to Lyapunov equation $(B = 0)$ in Range(V_k)

(here Range(V_k) is a Krylov-type subspace)

- $X_k^{(L)} \geq X_k^{(R)}$
- $\|\mathfrak{R}(X_k^{(R)}) - \mathfrak{L}(X_k^{(L)})\| = \|t_k^\top (Y_k^{(R)} - Y_k^{(L)})\|$, with

$$\|Y_k^{(L)} - Y_k^{(R)}\| \leq \frac{1}{2\alpha} \|(B^\top V_k) Y_k^{(R)}\|^2,$$

where $\alpha = -\lambda_{\max}((V_k^\top A V_k + V_k^\top A^\top V_k)/2)$

Some matrix relations

$X_k^{(R)}$: Galerkin approx to Riccati equation in Range(V_k)

$X_k^{(L)}$: Galerkin approx to Lyapunov equation ($B = 0$) in Range(V_k)

(here Range(V_k) is a Krylov-type subspace)

- $X_k^{(L)} \geq X_k^{(R)}$
- $\|\mathfrak{R}(X_k^{(R)}) - \mathfrak{L}(X_k^{(L)})\| = \|t_k^\top (Y_k^{(R)} - Y_k^{(L)})\|$, with

$$\|Y_k^{(L)} - Y_k^{(R)}\| \leq \frac{1}{2\alpha} \|(B^\top V_k) Y_k^{(R)}\|^2,$$

  where $\alpha = -\lambda_{\max}((V_k^\top A V_k + V_k^\top A^\top V_k)/2)$

- Residual norms: $X^{(R)}$ as Riccati vs Lyapunov solution

$$\|\mathfrak{R}(X_k^{(R)})\|_F^2 = \|\mathfrak{L}(X_k^{(R)})\|_F^2 - \|X_k^{(R)} B B^\top X_k^{(R)}\|_F^2$$

Some matrix relations

$X_k^{(R)}$: Galerkin approx to Riccati equation in Range(V_k)

$X_k^{(L)}$: Galerkin approx to Lyapunov equation ($B = 0$) in Range(V_k)

(here Range(V_k) is a Krylov-type subspace)

- $X_k^{(L)} \geq X_k^{(R)}$
- $\|\mathfrak{R}(X_k^{(R)}) - \mathfrak{L}(X_k^{(L)})\| = \|t_k^\top (Y_k^{(R)} - Y_k^{(L)})\|$, with

$$\|Y_k^{(L)} - Y_k^{(R)}\| \leq \frac{1}{2\alpha} \|(B^\top V_k) Y_k^{(R)}\|^2,$$

where $\alpha = -\lambda_{\max}((V_k^\top A V_k + V_k^\top A^\top V_k)/2)$

- Residual norms: $X^{(R)}$ as Riccati vs Lyapunov solution

$$\|\mathfrak{R}(X_k^{(R)})\|_F^2 = \|\mathfrak{L}(X_k^{(R)})\|_F^2 - \|X_k^{(R)} B B^\top X_k^{(R)}\|_F^2$$

But: Why does it work (well) ?

A low-rank subspace iteration for the Riccati equation

Consider $A^\top X + XA - XFX + G = 0$. Let

$$\mathcal{H} = \begin{bmatrix} A & -F \\ -G & -A^\top \end{bmatrix}$$

with eigenvalues satisfying

$$\Re(\lambda_1) \leq \Re(\lambda_2) \leq \ldots \leq \Re(\lambda_n) < 0 < \Re(\lambda_{n+1}) \leq \Re(\lambda_{n+2}) \leq \ldots \leq \Re(\lambda_{2n})$$

Then range($\begin{bmatrix} I_n \\ X \end{bmatrix}$) is an invariant subspace of $\mathcal{H}$ (X stabilizing soln)

A low-rank subspace iteration for the Riccati equation

Consider $A^\top X + XA - XFX + G = 0$. Let

$$\mathcal{H} = \begin{bmatrix} A & -F \\ -G & -A^\top \end{bmatrix}$$

with eigenvalues satisfying

$$\Re(\lambda_1) \leq \Re(\lambda_2) \leq \ldots \leq \Re(\lambda_n) < 0 < \Re(\lambda_{n+1}) \leq \Re(\lambda_{n+2}) \leq \ldots \leq \Re(\lambda_{2n})$$

Then range($\begin{bmatrix} I_n \\ X \end{bmatrix}$) is an invariant subspace of $\mathcal{H}$ (X stabilizing soln)

$\Rightarrow$ Subspace iteration with the Cayley transformation matrix

$$\mathcal{S}(\alpha) = (\mathcal{H} + \alpha I)^{-1}(\mathcal{H} - \overline{\alpha} I)$$

with eigs: $|\sigma_1| \geq \ldots \geq |\sigma_n| > 1 > |\sigma_{n+1}| \geq \ldots \geq |\sigma_{2n}|$

(as for acceleration procedures in QR iteration)

Subspace iteration with Cayley transformation

Given $X_0 \in \mathbb{R}^{n\times n}$ and α_k, $k = 1, 2, \ldots$ with $\Re(\alpha_k) > 0$

For $k = 1, 2, \ldots$

Compute

$$\begin{bmatrix} M_k \\ N_k \end{bmatrix} := \mathcal{S}(\alpha_k) \begin{bmatrix} I \\ X_{k-1} \end{bmatrix} \qquad \text{(with } \mathcal{S}(\alpha) = (\mathcal{H} + \alpha I)^{-1}(\mathcal{H} - \overline{\alpha} I)\text{)}$$

$$X_k := N_k M_k^{-1}$$

End

Subspace iteration with Cayley transformation

Given $X_0 \in \mathbb{R}^{n\times n}$ and α_k, $k = 1, 2, \ldots$ with $\Re(\alpha_k) > 0$

For $k = 1, 2, \ldots$

Compute

$$\begin{bmatrix} M_k \\ N_k \end{bmatrix} := \mathcal{S}(\alpha_k) \begin{bmatrix} I \\ X_{k-1} \end{bmatrix} \qquad \text{(with } \mathcal{S}(\alpha) = (\mathcal{H} + \alpha I)^{-1}(\mathcal{H} - \overline{\alpha} I)\text{)}$$

$$X_k := N_k M_k^{-1}$$

End

Corresponding to the following fixed point iteration:

$$\begin{aligned} X_k \quad = \quad & [-2\mathfrak{a}_k S_1^{-1} G(A + \alpha_k I)^{-1} + (I - 2\mathfrak{a}_k S_1^{-1}) X_{k-1}] \cdot \\ & \qquad [I - 2\mathfrak{a}_k S_2^{-1} - 2\mathfrak{a}_k S_2^{-1} F(-A^* + \alpha_k I)^{-1} X_{k-1}]^{-1} \end{aligned}$$

(here $\mathfrak{a}_k = \Re(\alpha_k)$)

On the convergence of subspace iteration

* Schur decomposition: $\mathcal{H} = QTQ^*$

* Schur decomposition: $\mathcal{S}_k = QT_{(k)}Q^*$, where

$T_{(k)} := \begin{bmatrix} T_{11(k)} & T_{12(k)} \\ 0 & T_{22(k)} \end{bmatrix}$ with $T_{jj(k)} = (T_{jj} + \alpha_k I)^{-1}(T_{jj} - \bar{\alpha}_k I)$

* $\begin{bmatrix} I \\ X_0 \end{bmatrix} = U_0 R_0$ skinny QR, X_0 s.t. $d = \mathrm{dist}(D_n(\mathcal{H}^*), \mathrm{Range}(\begin{bmatrix} I \\ X_0 \end{bmatrix})) < 1$

If the matrix M_k is nonsingular $\forall k$, then

$$\mathrm{dist}\left(\mathrm{Range}(\begin{bmatrix} I \\ X_+ \end{bmatrix}), \mathrm{Range}(\begin{bmatrix} I \\ X_k \end{bmatrix})\right) \le \gamma \left\| \prod_{i=k}^{1} T_{22(i)} \right\|_2 \left\| \prod_{i=1}^{k} T_{11(i)}^{-1} \right\|_2$$

where $\gamma = \frac{\|R_0^{-1}\|_2}{\sqrt{1-d^2}} \left(1 + \frac{\|T_{12}\|_F}{\mathrm{sep}(T_{11}, T_{22})}\right)$

Incremental low rank Subspace Iteration algorithm

If $F = BB^\top$ and $G = C^\top C$, then low-rank recurrence possible

1: INPUT $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{s\times n}$, $B \in \mathbb{R}^{n\times p}$, α_k, $k = 1, 2, \ldots$, with $\mathfrak{a}_k = \Re(\alpha_k)$

2: $V_1 := -2\mathfrak{a}_1(-A^\top + \alpha_1 I)^{-1}C^\top$, $T_1 = 2\mathfrak{a}_1(I + C(-A + \bar{\alpha}_1 I)^{-1}BB^\top(-A^\top + \alpha_1 I)^{-1}C^*)$

3: **for** $k = 2, 3, \ldots$

4: $\quad v_k := \frac{\alpha_k}{\alpha_{k-1}}(v_{k-1} - (\alpha_{k-1} + \bar{\alpha}_k)(-A^\top + \alpha_k I)^{-1}v_{k-1})$, $V_k = \begin{bmatrix} V_{k-1}, & v_k \end{bmatrix}$

5: $\quad Q_k :=$

$$\begin{bmatrix} & I \\ 1 & \end{bmatrix} \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ & \ddots & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} \begin{bmatrix} \frac{\bar{\alpha}_1+\alpha_k}{2\mathfrak{a}_k} & & & \\ \frac{\alpha_1-\alpha_k}{2\mathfrak{a}_k} & \frac{\bar{\alpha}_2+\alpha_k}{2\mathfrak{a}_k} & & \\ & \frac{\alpha_2-\alpha_k}{2\mathfrak{a}_k} & \ddots & & \\ & & \ddots & \frac{\bar{\alpha}_{k-1}+\alpha_k}{2\mathfrak{a}_k} & \\ & & & \frac{\alpha_{k-1}-\alpha_k}{2\mathfrak{a}_k} & \frac{\bar{\alpha}_k+\alpha_k}{2\mathfrak{a}_k} \end{bmatrix} \otimes I_s$$

6: $\quad P_k := Q_k^{-1} + \begin{bmatrix} I & \\ & 0 \end{bmatrix} \otimes I_s$

7: $\quad T_k := P_k^{-*}\left\{\begin{bmatrix} T_{k-1} & 0 \\ 0 & 2\mathfrak{a}_k I \end{bmatrix} + \frac{1}{2\mathfrak{a}_k} Q_k^{-*}V_k^* BB^\top V_k Q_k^{-1}\right\} P_k^{-1}$

8: OUTPUT: V_k, T_k s.t. $X_k = V_k T_k^{-1} V_k^* \approx X_+$

Properties of incremental low rank Subspace Iteration

- Low-rank approximate solution
- One solve per iteration (s if $C^\top$ has s columns)
- The generated space is the Rational Krylov space with poles α_j's

Properties of incremental low rank Subspace Iteration

- Low-rank approximate solution
- One solve per iteration (s if $C^\top$ has s columns)
- The generated space is the Rational Krylov space with poles α_j's
- CF-ADI-like basis V_k (algorithm coincides with ADI for $B = 0$)
- Theory motivates the parameter selection as

$$\{\alpha_1, \ldots, \alpha_k\} \quad = \quad \arg \min_{\alpha_1,\ldots,\alpha_k > 0} \max_{\lambda \in \lambda_+(\mathcal{H})} \prod_{i=1}^{k} \left| \frac{\lambda - \overline{\alpha}_i}{\lambda + \alpha_i} \right|$$

Subspace iteration vs Galerkin RKS method

Assume the field of values of A is in $\mathbb{C}^-$

$X_k = V_k T_k^{-1} V_k^\top$: Subspace iteration approx

$V_k = \mathrm{range}([(-A^\top + \alpha_1 I)^{-1} C^\top, \ldots, (-A^\top + \alpha_k I)^{-1} C^\top])$

$R_k = C^\top C + A^\top X_k + X_k A - X_k B B^\top X_k$: residual matrix. Then

$$V_k^\top R_k V_k = 0 \qquad \Leftrightarrow \qquad (V_k^\top V_k)^{-1} V_k^\top C^\top = T_k^{-1} \mathbf{1}$$

Moreover, the parameters α_j are the mirrored Ritz values of $A^\top - X_k B B^\top$:

$$\alpha_j = -\bar{\lambda}_j, \; j = 1, \ldots, k$$

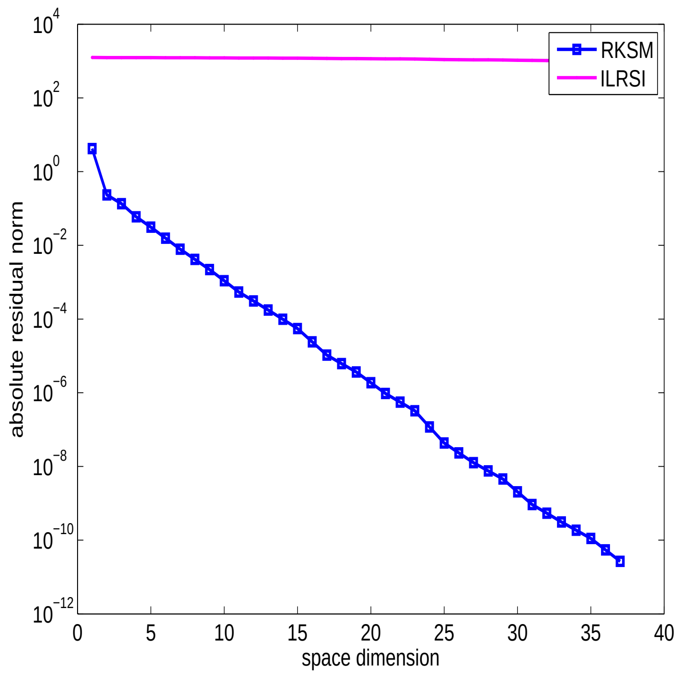
where $\lambda_j = \mathrm{eig}((V_k^\top V_k)^{-1} V_k^\top (A^\top - X_k B B^\top) V_k)$

(cf. H_2-optimal MOR)

A numerical example

Consider the 500×500 Toeplitz matrix

$$A = \text{toeplitz}(-1, \underline{2.5}, 1, 1, 1), \quad C = [1, -2, 1, -2, \ldots], B = \mathbf{1}$$



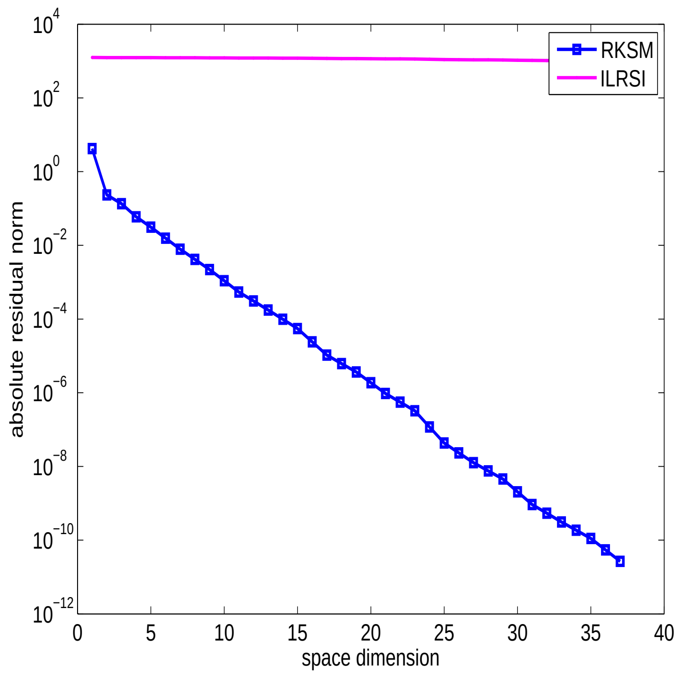
Parameter computation:

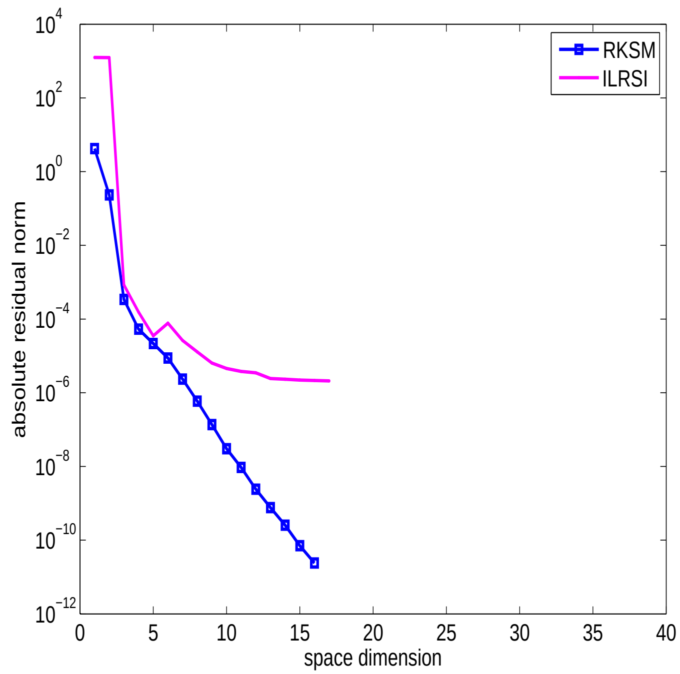
Left: adaptive RKSM on A

A numerical example

Consider the 500×500 Toeplitz matrix

$$A = \text{toeplitz}(-1, \underline{2.5}, 1, 1, 1), \quad C = [1, -2, 1, -2, \ldots], B = \mathbf{1}$$





Parameter computation:

Left: adaptive RKSM on A **Right**: adaptive RKSM on $A^\top - X_k^{(R)} BB^\top$

Conclusions and open problems

- Projection methods a good alternative to Newton iteration (numerical evidence)
- Derivation of an ad-hoc adaptive RKSM for Riccati equations
- Subspace iteration provides a new framework to analyze RKSM

References

- V. Simoncini, Daniel B. Szyld and Marlliny Monsalve,
  On two numerical methods for the solution of large-scale algebraic Riccati equations
  To appear in IMA Journal of Numerical Analysis. (available on-line)

- Yiding Lin and V. Simoncini,
  A new subspace iteration method for the algebraic Riccati equation
  pp.1-25, July 2013. arXiv:1307.3843. Submitted