Krylov subspace methods for solving Lyapunov equations *revisited*

Daniel Kressner
TU Berlin

kressner@math.tu-berlin.de
`http://www.cs.umu.se/~kressner`

06/15/07, Matrix equations workshop, TU Chemnitz

Lyapunov equation is linear matrix equation

$$AX + XA^T = -C \tag{LYAP}$$

with $n \times n$ coefficient matrices A and C.

Outline

- Basic properties
  - solvability, integral representation
  - Cauchy matrices
  - eigenvalue decay
- Krylov subspace methods
  - basic Arnoldi and Lanczos methods
  - convergence
  - preconditioning
  - finite-precision aspects
  - memory-efficient variants

Basic properties – Solvability

(LYAP) $AX + XA^T = -C$ can be written as $n^2 \times n^2$ linear system:

$$((I_n \otimes A) + (A \otimes I_n))\text{vec}(X) = -\text{vec}(C) \qquad \text{(KRON)}$$

with the Kronecker (tensor) product $\otimes$.

(LYAP) has a unique solution
$\Leftrightarrow$ (KRON) has a unique solution
$\Leftrightarrow$ A has no eigenvalue pairs $(\lambda, -\lambda)$.

In the rest of this talk, we assume A stable and C s.p.sd.

$\Downarrow$

$$X = \int_0^\infty e^{tA} C e^{tA^T}\, dt. \qquad \text{(INT)}$$

Basic properties – Solvability

(LYAP) $AX + XA^T = -C$ can be written as $n^2 \times n^2$ linear system:

$$\left((I_n \otimes A) + (A \otimes I_n)\right)\mathrm{vec}(X) = -\mathrm{vec}(C) \qquad \text{(KRON)}$$

with the Kronecker (tensor) product $\otimes$.

(LYAP) has a unique solution

$\Leftrightarrow$ (KRON) has a unique solution

$\Leftrightarrow$ A has no eigenvalue pairs $(\lambda, -\lambda)$.

In the rest of this talk, we assume A stable and C s.p.sd.

$\Downarrow$

$$X = \int_0^{\infty} e^{tA} C e^{tA^T}\, \mathrm{d}t. \qquad \text{(INT)}$$

Basic properties – Cauchy matrices

Let $A = T\Lambda T^{-1}$ (diagonalization) and $C = bb^T$ with vector b.
Transform (LYAP) to

$$\Lambda\tilde{X} + \tilde{X}\overline{\Lambda} = -(T^{-1}b)(T^{-1}b)^*.$$

Then

$$\tilde{X} = -D_b \left[\frac{1}{\lambda_i + \overline{\lambda_j}}\right]_{i,j=1}^{n} \overline{D_b} \qquad \text{(CAUCHY)}$$

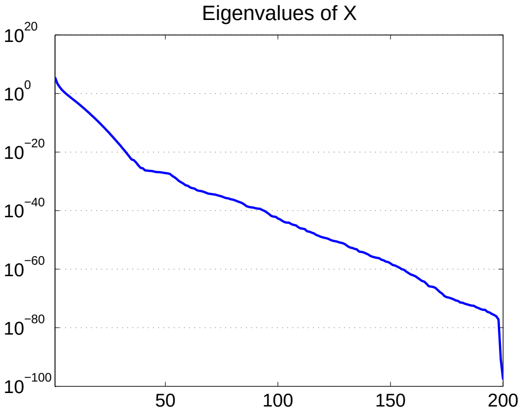
with $D_b = \operatorname{diag}(\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_n)$.

$\left[\frac{1}{\lambda_i + \overline{\lambda_j}}\right]_{i,j=1}^{n}$ is called a Cauchy matrix.

Eigenvalue decay – Example

Consider $AX + XA^T = -BB^T$, where $B \in \mathbb{R}^{n \times m}$ with $m \ll n$. Then the singular values of X often decay very rapidly.

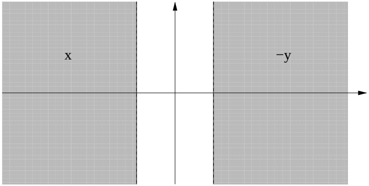
Simple example:
$A \in \mathbb{R}^{200 \times 200}$ second-difference operator and $B = [1, \ldots, 1]^T$.



Eigenvalue decay – Analytic Intuition

$1/(x+y)$ can be well approximated by separated functions

$$\frac{1}{x+y} \approx \sum_{l=1}^{k} f_l(x)\overline{f_l(y)}.$$



Apply to (CAUCHY):

$$\left[\frac{1}{\lambda_i + \overline{\lambda_j}}\right]_{i,j=1}^{n} \approx \left[\sum_{l=1}^{k} f_l(\lambda_i)\overline{f_l(\lambda_j)}\right]_{i,j=1}^{n} = LL^H$$

with

$$L = \begin{bmatrix} f_1(\lambda_1) & \cdots & f_l(\lambda_1) \\ f_1(\lambda_2) & \cdots & f_l(\lambda_2) \\ \vdots & & \vdots \\ f_1(\lambda_n) & \cdots & f_l(\lambda_n) \end{bmatrix}.$$

Basic properties – Eigenvalue decay

Approaches to prove eigenvalue decay

- [Penzl'00]: Apply ADI method to construct explicit low-rank approximations to $X \rightsquigarrow$

$$\frac{\lambda_{mk+1}(X)}{\lambda_1(X)} \leq \left(\prod_{j=0}^{k-1} \frac{\kappa^{(2j+1)/(2k)} - 1}{\kappa^{(2j+1)/(2k)} + 1} \right)^2 = \exp\left(-O\left(\frac{k}{\log \kappa} \right) \right)$$

holds for symmetric A with $\kappa = \kappa(A)$.

Basic properties – Eigenvalue decay

- [Antoulas/Sorensen/Zhou'02]: Apply results on Cholesky factor diagonal of Cauchy matrices [Gohberg/Koltracht'96] to (CAUCHY) $\rightsquigarrow$

$$\frac{\lambda_{mk}(X)}{\lambda_1(X)} \leq \frac{\kappa(T)\Re\lambda_1}{\Re\lambda_k} \prod_{j=1}^{k-1} \left| \frac{\lambda_k - \lambda_j}{\overline{\lambda_k} + \lambda_j} \right|^2 .$$

  holds for nonsymmetric, diagonalizable A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

- [Grasedyck/Hackbusch/Khoromsky'02]: Apply numerical quadrature [Stenger'93] to (INT) $\rightsquigarrow$

$$\frac{\lambda_{mk}(X)}{\lambda_1(X)} \leq O(\exp(-\sqrt{k})).$$

Basic properties – Eigenvalue decay

- [Antoulas/Sorensen/Zhou'02]: Apply results on Cholesky factor diagonal of Cauchy matrices [Gohberg/Koltracht'96] to (CAUCHY) $\rightsquigarrow$

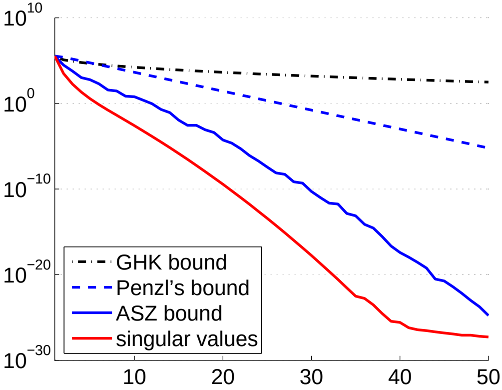
$$\frac{\lambda_{mk}(X)}{\lambda_1(X)} \leq \frac{\kappa(T)\Re\lambda_1}{\Re\lambda_k} \prod_{j=1}^{k-1} \left| \frac{\lambda_k - \lambda_j}{\overline{\lambda_k} + \lambda_j} \right|^2 .$$

  holds for nonsymmetric, diagonalizable A with eigenvalues $\lambda_1, \ldots, \lambda_n$.
- [Grasedyck/Hackbusch/Khoromsky'02]: Apply numerical quadrature [Stenger'93] to (INT) $\rightsquigarrow$

$$\frac{\lambda_{mk}(X)}{\lambda_1(X)} \leq O(\exp(-\sqrt{k})).$$

Basic properties – Eigenvalue decay

Example:
$A \in \mathbb{R}^{200\times 200}$ second-difference operator and $B = [1, \dots, 1]^T$.



Existing solution methods for Lyapunov equations

$n \leq O(10^4)$

Schur form + substitution [Bartels/Stewart'72] [Hammarling'82], [Jonsson/Kågström'02], [Sorensen/Zhou'03], [Quintana-Ortí/van de Geijn'03], [Granat'05], [Kressner'06].

form/store A^{-1}

Sign function iteration [Roberts'71], [Beavers/Denman'75], [Byers'87], [Benner/Quintana-Ortí'99], [Grasedyck et al.'03], [Baur/Benner'04].

apply $(A - \sigma I)^{-1}$

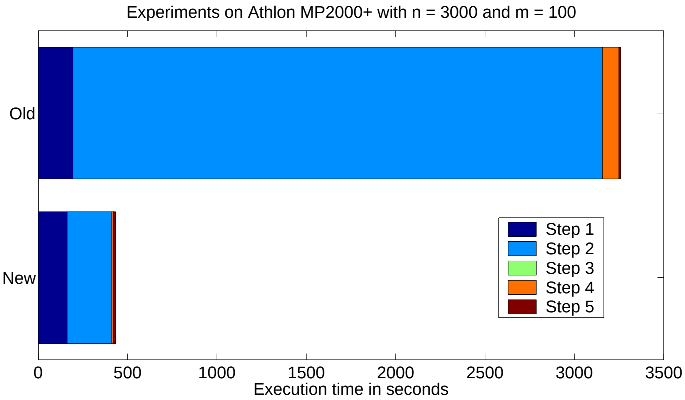
ADI [Wachspress'88], [Penzl'99], [Li/White'02], [Gugercin et al.'03], [Simoncini'06] a.m.o.; API [Hodel et al.'96].

Multigrid methods [Rosen/Wang'95], [Penzl'97], [Grasedyck/Hackbusch'04].

apply A

Krylov subspace methods [Saad'90], [Hu/Reichel'92], [Jaimoukha/Kasenally'94], [Hochbruck/Starke'95], [Kressner/Marković'07].

Bartels/Stewart method



Step 1 $A \rightsquigarrow$ Hessenberg H (Old: LAPACK 3, New: [Quintana-Ortí/Van de Geijn'06])

Step 2 $H \rightsquigarrow$ Schur $\widetilde{A}$ (Old: LAPACK 3, New: [Braman/Byers/Mathias'02])

Step 4 Sol. of $\widetilde{A}\widetilde{A}X + \widetilde{X}\widetilde{A}^T = \widetilde{B}\widetilde{B}^T$ (Old: SLICOT, New: [Kressner'06])

Krylov subspace methods

Apply Arnoldi method to generate orthonormal basis U_k of

$$\text{span}\{b, Ab, A^b, \ldots, A^{k-1}b\}$$

$u_1 = b/\|b\|_2$
for j = 1,. . . ,k **do**
  $w = Au_j$
  $h_{1:j,j} = [u_1, \ldots, u_j]^T w$
  $w = w - [u_1, \ldots, u_j]h_{1:j,j}$
  $h_{j+1,j} = \|w\|_2$
  $u_{j+1} = u_j/h_{j+1,j}$
end for

Setting

$$U_k = [u_1, \ldots, u_k], \qquad H_k = [h_{ij}]_{i,j=1}^k$$

yields Arnoldi decomposition

$$AU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^T.$$

Krylov subspace methods

Apply Arnoldi method to generate orthonormal basis U_k of

$$\text{span}\{b, Ab, A^b, \ldots, A^{k-1}b\}$$

$u_1 = b/\|b\|_2$
for j = 1,. . . ,k **do**
  $w = Au_j$
  $h_{1:j,j} = [u_1, \ldots, u_j]^T w$
  $w = w - [u_1, \ldots, u_j]h_{1:j,j}$
  $h_{j+1,j} = \|w\|_2$
  $u_{j+1} = u_j/h_{j+1,j}$
end for

Setting

$$U_k = [u_1, \ldots, u_k], \qquad H_k = [h_{ij}]_{i,j=1}^{k}$$

yields Arnoldi decomposition

$$AU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^T.$$

Arnoldi method applied to Lyapunov equation

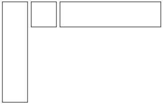
$$U_k^T \quad \rightarrow \quad AX + XA^T = -bb^T \quad \leftarrow \quad U_k$$

This is a *compressed* Lyapunov equation of size $k \times k$.

- Requires stability of H_k to have positive semi-definiteness of Y_k. Sufficient condition: $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.
- Hammarling's method requires $\mathcal{O}(k^3)$ flops and yields factorized solution $Y_k = L_k L_k^T$.

Set

$$X_k = U_k Y_k U_k^T =$$



Then (hopefully) $X_k \rightarrow X$ as $k \rightarrow \infty$.

Convergence can be cheaply monitored:

$$\|AX_k + X_k A^T + bb^T\|_2 = \|e_k^T Y_k\|_2 \, \|b\|_2^2.$$

Arnoldi method applied to Lyapunov equation

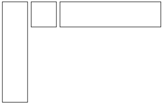
$$U_k^T \;\rightarrow\; AX + XA^T = -bb^T \;\leftarrow\; U_k$$

This is a *compressed* Lyapunov equation of size $k \times k$.

- Requires stability of H_k to have positive semi-definiteness of Y_k. Sufficient condition: $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.
- Hammarling's method requires $\mathcal{O}(k^3)$ flops and yields factorized solution $Y_k = L_k L_k^T$.

Set

$$X_k = U_k Y_k U_k^T =$$



Then (hopefully) $X_k \rightarrow X$ as $k \rightarrow \infty$.

Convergence can be cheaply monitored:

$$\|AX_k + X_k A^T + bb^T\|_2 = \|e_k^T Y_k\|_2 \, \|b\|_2^2.$$

Arnoldi method applied to Lyapunov equation

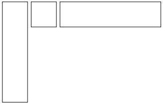
$$H_k Y_k + Y_k H_k^T = -\|b\|_2^2 e_1 e_1^T$$

This is a *compressed* Lyapunov equation of size $k \times k$.

- Requires stability of H_k to have positive semi-definiteness of Y_k. Sufficient condition: $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.
- Hammarling's method requires $\mathcal{O}(k^3)$ flops and yields factorized solution $Y_k = L_k L_k^T$.

Set

$$X_k = U_k Y_k U_k^T =$$



Then (hopefully) $X_k \to X$ as $k \to \infty$.

Convergence can be cheaply monitored:

$$\|AX_k + X_k A^T + bb^T\|_2 = \|e_k^T Y_k\|_2 \, \|b\|_2^2.$$

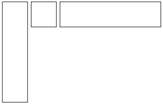
Arnoldi method applied to Lyapunov equation

$$H_k Y_k + Y_k H_k^T = -\|b\|_2^2 e_1 e_1^T$$

This is a *compressed* Lyapunov equation of size $k \times k$.

- Requires stability of H_k to have positive semi-definiteness of Y_k. Sufficient condition: $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.
- Hammarling's method requires $\mathcal{O}(k^3)$ flops and yields factorized solution $Y_k = L_k L_k^T$.

Set

$$X_k = U_k Y_k U_k^T =$$ 

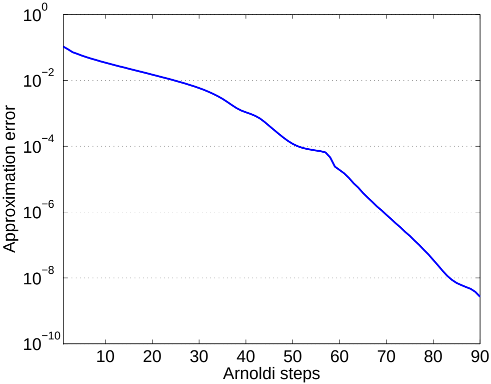
Then (hopefully) $X_k \to X$ as $k \to \infty$.

Convergence can be cheaply monitored:

$$\|AX_k + X_k A^T + bb^T\|_2 = \|e_k^T Y_k\|_2 \, \|b\|_2^2.$$

Convergence of Arnoldi for Lyapunov

2D instationary heat equation on the unit square with homogeneous 1st kind boundary conditions; 30×30 finite difference discretization.



Linear convergence observed.

Convergence analysis – basic ideas

Solution of original (LYAP) can be written as (INT):

$$X = \int_0^\infty e^{At} bb^T e^{A^T t}\, \mathrm{d}t.$$

Similarly for the compressed equation:

$$\begin{aligned} X_k = U_k Y_k U_k^T &= \|b\|_2^2\, U_k \left(\int_0^\infty e^{H_k t} e_1 e_1^T e^{H_k^T t}\, \mathrm{d}t \right) U_k^T \\ &= \int_0^\infty e^{A_k t} bb^T e^{A_k^T t}\, \mathrm{d}t, \end{aligned}$$

where $A_k = U_k H_k U_k^T = U_k U_k^T A U_k U_k^T$.

$e^{A_k t} b =$ conventional Krylov subspace approximation to $e^{At} b$.

Convergence analysis – basic ideas

Solution of original (LYAP) can be written as (INT):

$$X = \int_0^\infty e^{At} bb^T e^{A^T t} \, \mathrm{d}t.$$

Similarly for the compressed equation:

$$\begin{aligned} X_k = U_k Y_k U_k^T &= \|b\|_2^2 \, U_k \left(\int_0^\infty e^{H_k t} e_1 e_1^T e^{H_k^T t} \, \mathrm{d}t \right) U_k^T \\ &= \int_0^\infty e^{A_k t} bb^T e^{A_k^T t} \, \mathrm{d}t, \end{aligned}$$

where $A_k = U_k H_k U_k^T = U_k U_k^T A U_k U_k^T$.

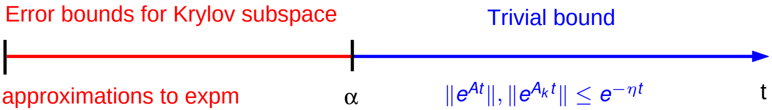
$e^{A_k t} b =$ conventional Krylov subspace approximation to $e^{At} b$.

Convergence analysis – basic ideas

Enables the use of convergence results for expm:

$$
\begin{aligned}
\|X - X_k\| &\leq \int_0^\infty \left\| e^{At} bb^T e^{A^T t} - e^{A_k t} bb^T e^{A_k^T t} \right\| \, \mathrm{d}t \\
&\leq 2 \int_0^\infty \left\| \left(e^{At} - e^{A_k t} \right) bb^T \left(e^{A^T t} + e^{A_k^T t} \right) \right\| \, \mathrm{d}t
\end{aligned}
$$

Bounds for red term:



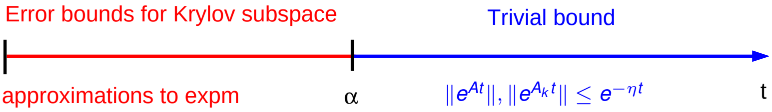
with $\eta = -\lambda_{\max}(A + A^T)/2 > 0$. To obtain convergence for $k \to \infty$, the cut point α has to be chosen dynamically (α increases monotonically with k).

Convergence analysis – basic ideas

Enables the use of convergence results for expm:

$$
\begin{aligned}
\|X - X_k\| &\leq \int_0^\infty \left\| e^{At} bb^T e^{A^T t} - e^{A_k t} bb^T e^{A_k^T t} \right\| \, \mathrm{d}t \\
&\leq 2 \int_0^\infty \left\| \left(e^{At} - e^{A_k t} \right) bb^T \left(e^{A^T t} + e^{A_k^T t} \right) \right\| \, \mathrm{d}t
\end{aligned}
$$

Bounds for red term:



with $\eta = -\lambda_{\max}(A + A^T)/2 > 0$. To obtain convergence for $k \to \infty$, the cut point α has to be chosen dynamically (α increases monotonically with k).

Most general result

[Gallopoulos/Saad'92]: Bounds for expm from Taylor approx. of exp:

$$\| \left(e^{At} - e^{A_k t}\right) b\| \leq 2\|b\| \frac{\|At\|^k}{k!}$$

Optimal cut point $\alpha = O(k)$ $\rightsquigarrow$ Loss of superlinear convergence!

Resulting approximation result for Lyapunov equation:

$$\|X - X_k\| \leq \frac{2\|b\|^2}{\eta} \gamma^k$$

with the convergence rate

$$\gamma = \exp W\left(-\frac{2\eta}{e\|A\|}\right) \approx \exp\left(-\frac{2\eta}{e\|A\|}\right);$$

W is Lambert-W function ($=$ inverse function of $f(W) = We^W$)
and $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.

$\gamma \approx 0.998$ for numerical example!

Most general result

[Gallopoulos/Saad'92]: Bounds for expm from Taylor approx. of exp:

$$\| \left(e^{At} - e^{A_k t}\right) b\| \leq 2\|b\| \frac{\|At\|^k}{k!}$$

Optimal cut point $\alpha = O(k)$ $\rightsquigarrow$ Loss of superlinear convergence!

Resulting approximation result for Lyapunov equation:

$$\|X - X_k\| \leq \frac{2\|b\|^2}{\eta}\gamma^k$$

with the convergence rate

$$\gamma = \exp W\left(-\frac{2\eta}{e\|A\|}\right) \approx \exp\left(-\frac{2\eta}{e\|A\|}\right);$$

W is Lambert-W function ($=$ inverse function of $f(W) = We^W$) and $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.

$\gamma \approx 0.998$ for numerical example!

Most general result

[Gallopoulos/Saad'92]: Bounds for expm from Taylor approx. of exp:

$$\| \left(e^{At} - e^{A_k t}\right) b\| \leq 2\|b\| \frac{\|At\|^k}{k!}$$

Optimal cut point $\alpha = O(k)$ $\rightsquigarrow$ Loss of superlinear convergence!

Resulting approximation result for Lyapunov equation:

$$\|X - X_k\| \leq \frac{2\|b\|^2}{\eta} \gamma^k$$

with the convergence rate

$$\gamma = \exp W\left(-\frac{2\eta}{e\|A\|}\right) \approx \exp\left(-\frac{2\eta}{e\|A\|}\right);$$

W is Lambert-W function ($=$ inverse function of $f(W) = We^W$)
and $\eta = -\lambda_{\max}(A + A^T)/2 > 0$.

$\gamma \approx 0.998$ for numerical example!

Symmetric case

[Hochbruck/Lubich'97] provides improved bounds on Krylov subspace approx. to expm; especially for symmetric A:

$$\left\|e^{At}b - e^{A_k t}b\right\| \leq \begin{cases} 10e^{-\eta t}e^{-\rho t}\frac{e^k \rho^{k-1} t^{k-1}}{k^k}, & t \in [0, k/(2\rho)], \\ 10e^{-\eta t}e^{-k^2/(5\rho t)}, & t \in [k/(2\rho), k^2/(4\rho)], \end{cases}$$

where $\rho = (\lambda_{\max}(A) - \lambda_{\min}(A))/4$.

Application to (LYAP) with nearly optimal cut point $\alpha = k/\sqrt{10\,\rho\eta}$ yields

$$\|X - X_k\| \leq \frac{80\,\|b\|^2}{\rho k}\left(\frac{\sqrt{e}}{2}\right)^k + \frac{\|b\|^2}{\eta}\left(20\,e^{-k\eta/\rho} + 1\right)\gamma^k$$

for all $k \geq 4\sqrt{\rho/(10\eta)}$, where $\gamma = e^{-\sqrt{2\eta/(5\rho)}}$.

$\gamma \approx 0.938$ (instead of $\gamma \approx 0.998$) for numerical example.

Symmetric case

[Hochbruck/Lubich'97] provides improved bounds on Krylov subspace approx. to expm; especially for symmetric A:

$$\|e^{At}b - e^{A_k t}b\| \le \begin{cases} 10e^{-\eta t} e^{-\rho t} \frac{e^k \rho^{k-1} t^{k-1}}{k^k}, & t \in [0, k/(2\rho)], \\ 10e^{-\eta t} e^{-k^2/(5\rho t)}, & t \in [k/(2\rho), k^2/(4\rho)], \end{cases}$$

where $\rho = (\lambda_{\max}(A) - \lambda_{\min}(A))/4$.

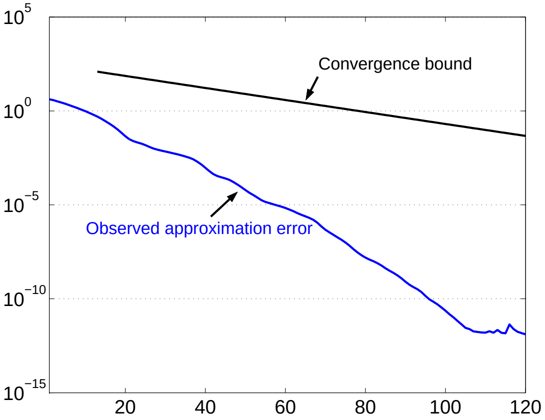
Application to (LYAP) with nearly optimal cut point $\alpha = k/\sqrt{10\,\rho\eta}$ yields

$$\|X - X_k\| \le \frac{80\,\|b\|^2}{\rho k}\left(\frac{\sqrt{e}}{2}\right)^k + \frac{\|b\|^2}{\eta}(20\,e^{-k\eta/\rho} + 1)\gamma^k$$

for all $k \ge 4\sqrt{\rho/(10\eta)}$, where $\gamma = e^{-\sqrt{2\eta/(5\rho)}}$.

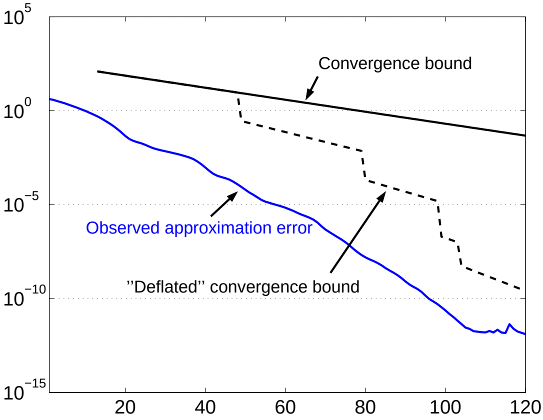
$\gamma \approx 0.938$ (instead of $\gamma \approx 0.998$) for numerical example.

Numerical example, ctd.



The bound can be considerably improved by neglecting converged eigenvalues of A during the Arnoldi process.

Numerical example, ctd.



Open: Genuine a priori bounds matching the convergence behavior better.

Outlook: Direct bounds

Symmetric case: Krylov subspace approximation to X can be related to

$$\min_{X_k \in \mathbb{R}^{k\times k}} \|AU_kX_kU_k + U_kX_kU_kA + bb^T\|_2$$

$$= \min_{p_1,\ldots,p_k \in \mathcal{P}_{k-1}} \max_{\lambda,\mu \in \Lambda(A)} \left| (\lambda+\mu) \sum_{l=1}^{k} p_l(\lambda)p_l(\mu) + 1 \right| |b_\lambda b_\mu|,$$

where $\mathcal{P}_{k-1}$ space of all polynomials of degree at most $k-1$.

Replacing $\Lambda(A)$ by $[\lambda_{\min}(A), \lambda_{\max}(A)]$ leads to approx. problem:

$$\min_{p_1,\ldots,p_k \in \mathcal{P}_{k-1}} \max_{x,y \in [\lambda_{\min}(A),\lambda_{\max}(A)]} \left| \frac{1}{x+y} + \sum_{l=1}^{k} p_l(x)p_l(y) \right|.$$

Outlook: Direct bounds

Symmetric case: Krylov subspace approximation to X can be related to

$$\min_{X_k \in \mathbb{R}^{k\times k}} \|AU_kX_kU_k + U_kX_kU_kA + bb^T\|_2$$

$$= \min_{p_1,\dots,p_k \in \mathcal{P}_{k-1}} \max_{\lambda,\mu \in \Lambda(A)} \left|(\lambda+\mu)\sum_{l=1}^{k} p_l(\lambda)p_l(\mu) + 1\right| |b_\lambda b_\mu|,$$

where $\mathcal{P}_{k-1}$ space of all polynomials of degree at most $k-1$.

Replacing $\Lambda(A)$ by $[\lambda_{\min}(A), \lambda_{\max}(A)]$ leads to approx. problem:

$$\min_{p_1,\dots,p_k \in \mathcal{P}_{k-1}} \max_{x,y \in [\lambda_{\min}(A),\lambda_{\max}(A)]} \left|\frac{1}{x+y} + \sum_{l=1}^{k} p_l(x)p_l(y)\right|.$$

Outlook: Preconditioning

$$AX + XA^T = -BB^T \quad \rightsquigarrow \quad XA^{-T} + A^{-1}X = -(A^{-1}B)(A^{-1}B)^T$$

Weighted sum of both equations:

$$(A + \alpha A^{-1})X + X(A + \alpha A^{-1})^T = -[B, A^{-1}B][B, A^{-1}B]^T$$

A symmetric and optimal choice of α:

$$\kappa(A + \alpha A^{-1}) \leq \sqrt{\kappa(A)}.$$

- Krylov subspace method based on A and A^{-1} [Simoncini'06] converges much faster.
- Effect of inexactness $P \approx A^{-1}$ not clear yet.

Outlook: Preconditioning

$$AX + XA^T = -BB^T \quad \rightsquigarrow \quad XA^{-T} + A^{-1}X = -(A^{-1}B)(A^{-1}B)^T$$

Weighted sum of both equations:

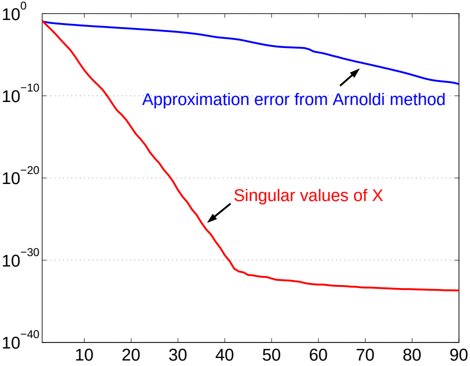
$$(A + \alpha A^{-1})X + X(A + \alpha A^{-1})^T = -[B, A^{-1}B][B, A^{-1}B]^T$$

A symmetric and optimal choice of α:

$$\kappa(A + \alpha A^{-1}) \leq \sqrt{\kappa(A)}.$$

- Krylov subspace method based on A and A^{-1} [Simoncini'06] converges much faster.
- Effect of inexactness $P \approx A^{-1}$ not clear yet.

Convergence does not capture singular value decay



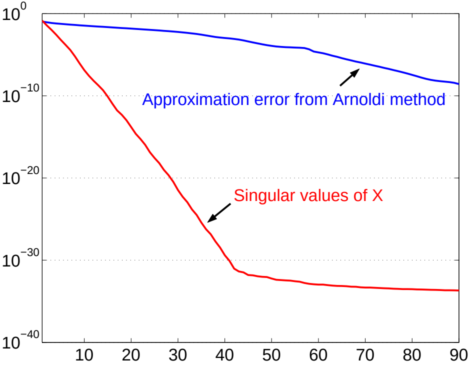
Example:

- svd(X) tells there is a rank 12 approximation with error 10^{-8};
- Arnoldi yields only a rank 84 approximation with error 10^{-8}.

Achieve less storage requirements (allow more mv-products):

- perform restarts (open problem);
- perform two identical passes of Arnoldi method.

Convergence does not capture singular value decay



Example:

- svd(X) tells there is a rank 12 approximation with error 10^{-8};
- Arnoldi yields only a rank 84 approximation with error 10^{-8}.

Achieve less storage requirements (allow more mv-products):

- perform restarts (open problem);
- perform two identical passes of Arnoldi method.

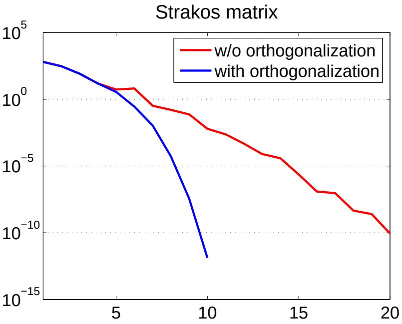
Lanczos method for symmetric matrix

If A is symmetric, then

Arnoldi process can be implemented using $O(n)$ memory.

But:

- no reorthogonalization possible;
- we need U_k to construct X_k.



Paige's classical results can be extended to show:

- loss of orthogonality only has mild effect on convergence *bounds*;
- residual norm can still be estimated using $\|e_k^T Y_k\|_2$.

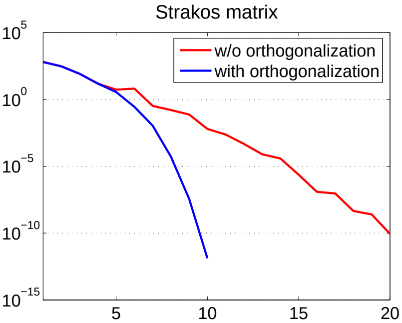
Lanczos method for symmetric matrix

If A is symmetric, then

Arnoldi process can be implemented using $O(n)$ memory.

But:

- no reorthogonalization possible;
- we need U_k to construct X_k.



Paige's classical results can be extended to show:

- loss of orthogonality only has mild effect on convergence *bounds*;
- residual norm can still be estimated using $\|e_k^T Y_k\|_2$.

Two-pass Lanczos method

Assume A is symmetric.

1st p. Apply Lanczos method without reorthogonalization!
Solve projected Lyapunov equation $H_k Y_k + Y_k H_k = -\beta^2 e_1 e_1^T$
and compute low rank approximation $Y_k \approx L_k L_k^T$ with $L_k \in \mathbb{R}^{k \times l}$.

Often, $l \ll k$!

2nd p. Repeat identical copy of Lanczos method to compute $V = U_k L_k$.
Yields approximate solution $X \approx VV^T$.

Reduces memory requirements from $n \times k$ to $n \times l$.
Doubles #mv-products.

Two-pass Lanczos method

Assume A is symmetric.

1st p. Apply Lanczos method without reorthogonalization!
Solve projected Lyapunov equation $H_k Y_k + Y_k H_k = -\beta^2 e_1 e_1^T$
and compute low rank approximation $Y_k \approx L_k L_k^T$ with $L_k \in \mathbb{R}^{k \times l}$.

Often, $l \ll k$!

2nd p. Repeat identical copy of Lanczos method to compute $V = U_k L_k$.
Yields approximate solution $X \approx VV^T$.

Reduces memory requirements from $n \times k$ to $n \times l$.
Doubles #mv-products.

Two-pass Lanczos method

Assume A is symmetric.

1st p. Apply Lanczos method without reorthogonalization!
Solve projected Lyapunov equation $H_k Y_k + Y_k H_k = -\beta^2 e_1 e_1^T$
and compute low rank approximation $Y_k \approx L_k L_k^T$ with $L_k \in \mathbb{R}^{k \times l}$.

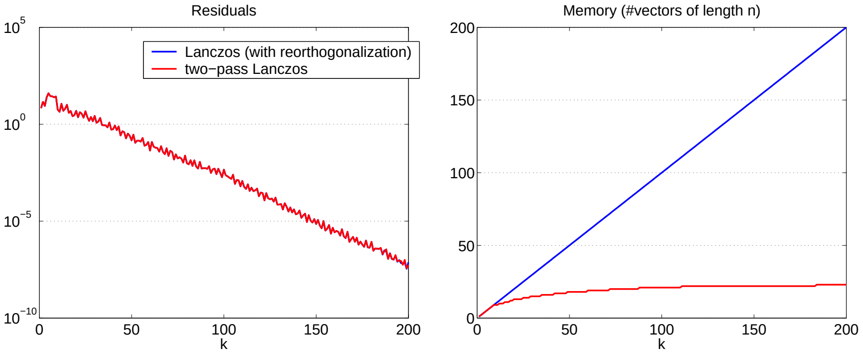
Often, $l \ll k$!

2nd p. Repeat identical copy of Lanczos method to compute $V = U_k L_k$.
Yields approximate solution $X \approx VV^T$.

Reduces memory requirements from $n \times k$ to $n \times l$.
Doubles #mv-products.

Spiral Inductor PEEC Model [Li/Kamon'05]

A is 1434×1434 symmetric negative definite matrix and $m = 1$.



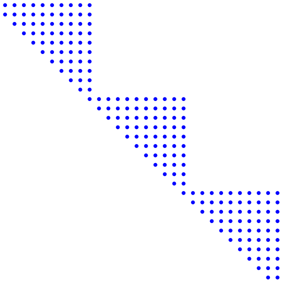
Two-pass "restarted" Arnoldi

[Eiermann/Ernst'06]: "Restarted" Arnoldi for computing $e^{At}b$ (and other matrix functions).

Rough idea: Restart Arnoldi method every k iterations with last column of Krylov subspace basis $\rightsquigarrow$ Krylov decomposition

$$A[U_k^{(1)}, \ldots, U_k^{(p)}] = [U_k^{(1)}, \ldots, U_k^{(p)}]H_{pk} + u_{k+1}^{(p)} h_{pk+1,pk} e_{pk}^T,$$

where H_{pk} has structure



Two-pass restarted Arnoldi

Advantages:

- solution of projected Lyapunov equation requires $\mathcal{O}(pk^3)$ instead of $\mathcal{O}(p^3k^3)$ flops;
- only k basis vectors need to be stored when using a two-pass approach.

Disadvantages:

- more likely to break down if $\lambda_{\max}(A + A^T) \geq 0$;
- at least doubles #mv-products (significantly more are needed for small k).

Two-pass restarted Arnoldi

Advantages:

- solution of projected Lyapunov equation requires $\mathcal{O}(pk^3)$ instead of $\mathcal{O}(p^3k^3)$ flops;
- only k basis vectors need to be stored when using a two-pass approach.

Disadvantages:

- more likely to break down if $\lambda_{\max}(A + A^T) \geq 0$;
- at least doubles #mv-products (significantly more are needed for small k).

LYAPACK benchmark example [Penzl'99]

A is finite difference discretization ($n = 10000$)

$$\triangle u(x,y) - 10x\frac{\partial u(x,y)}{\partial x} - 20y\frac{\partial u(x,y)}{\partial y} \qquad \text{on} \;\; \Omega = [0,1] \times [0,1].$$

Number of Arnoldi iterations to attain a residual of norm $\leq 10^{-5}$.

k	kp
∞	300
100	400
50	450
20	540
10	> 1000

Two-pass approach reduces memory requirements from $300n$ to $\approx 50n$ (however, #mv-products is tripled).

Conclusions

Novel results on Krylov subspace methods for Lyapunov equations:

- convergence bounds;
- efficient solution of compressed equation (Gohberg-Koltracht);
- analysis of loss of orthogonality in symmetric Lanczos.

(Joint work with Darija Marković and Ninoslav Truhar, both U Osijek, Croatia.)

Still to do:

- analysis of preconditioning:

$$(A + \alpha P)X + X(A + \alpha P)^T = BB^T + \alpha PBB^T P^T,$$

  where $P \approx A^{-1}$;
- analysis of nonsymmetric restarted Arnoldi;
- production code implementation.

Thanks for your attention!

Conclusions

Novel results on Krylov subspace methods for Lyapunov equations:

- convergence bounds;
- efficient solution of compressed equation (Gohberg-Koltracht);
- analysis of loss of orthogonality in symmetric Lanczos.

(Joint work with Darija Marković and Ninoslav Truhar, both U Osijek, Croatia.)

Still to do:

- analysis of preconditioning:

$$(A + \alpha P)X + X(A + \alpha P)^T = BB^T + \alpha PBB^T P^T,$$

  where $P \approx A^{-1}$;
- analysis of nonsymmetric restarted Arnoldi;
- production code implementation.

Thanks for your attention!

Conclusions

Novel results on Krylov subspace methods for Lyapunov equations:

- convergence bounds;
- efficient solution of compressed equation (Gohberg-Koltracht);
- analysis of loss of orthogonality in symmetric Lanczos.

(Joint work with Darija Marković and Ninoslav Truhar, both U Osijek, Croatia.)

Still to do:

- analysis of preconditioning:

$$(A + \alpha P)X + X(A + \alpha P)^T = BB^T + \alpha PBB^T P^T,$$

  where $P \approx A^{-1}$;
- analysis of nonsymmetric restarted Arnoldi;
- production code implementation.

Thanks for your attention!