Verfahren für große dünnbesetzte Matrixgleichungen

Jens Saak

Seminar: Algebraische Matrixgleichungen WS 09/10

Chemnitz, den 18.01.2010

Department of Mathematics
CHEMNITZ UNIVERSITY OF TECHNOLOGY

Outline

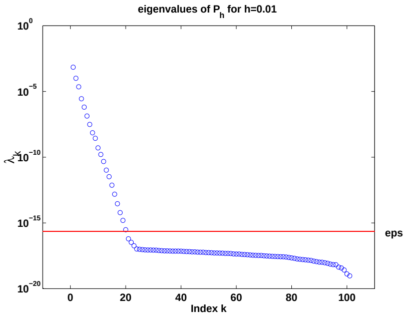
Motivation

Why use Low Rank Approximation?

The spectrum of an AREs solution

Motivating Example

- Linear 1D heat equation with point control.
- $\Omega = [0,1]$.
- FEM discretization using linear B-splines.
- h=0.01.



$$X = X^T \geq 0 \Longrightarrow X = ZZ^T = \sum_{k=1}^{n} \lambda_k z_k z_k^T \approx \sum_{k=1}^{r} \lambda_k z_k z_k^T = Z_{(r)} Z_{(r)}^T.$$

Low Rank Solvers for Lyapunov Equations

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

Consider $$FX + XF^T = -GG^T$$

ADI Iteration for the Lyapunov Equation (LE) [WACHSPRESS '88]

For $j = 1, \ldots, J$

$$\begin{aligned} X_0 &= 0 \\ (F + p_j I)X_{j-\frac{1}{2}} &= -GG^T - X_{j-1}(F^T - p_j I) \\ (F + p_j I)X_j^T &= -GG^T - X_{j-\frac{1}{2}}^T(F^T - p_j I) \end{aligned}$$

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

Consider $$FX + XF^T = -GG^T$$

ADI Iteration for the Lyapunov Equation (LE) [Wachspress '88]

For $j = 1, \dots, J$

$$\begin{aligned} X_0 &= 0 \\ (F + p_j I)X_{j-\frac{1}{2}} &= -GG^T - X_{j-1}(F^T - p_j I) \\ (F + p_j I)X_j^T &= -GG^T - X_{j-\frac{1}{2}}^T(F^T - p_j I) \end{aligned}$$

Rewrite as one step iteration and factorize $X_i = Z_i Z_i^T$, $i = 0, \dots, J$

$$\begin{aligned} Z_0 Z_0^T &= 0 \\ Z_j Z_j^T &= -2p_j (F + p_j I)^{-1} G G^T (F + p_j I)^{-T} \\ &\quad + (F + p_j I)^{-1}(F - p_j I) Z_{j-1} Z_{j-1}^T (F - p_j I)^T (F + p_j I)^{-T} \end{aligned}$$

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

$$Z_j = [\sqrt{-2p_j}(F + p_j I)^{-1}G,\ (F + p_j I)^{-1}(F - p_j I)Z_{j-1}]$$

[Penzl '00]

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

$$Z_j = [\sqrt{-2p_j}(F + p_j I)^{-1} G,\ (F + p_j I)^{-1}(F - p_j I) Z_{j-1}]$$

[Penzl '00]

Observing that $(F - p_i I)$, $(F + p_k I)^{-1}$ commute, we rewrite Z_J as

$$Z_J = [z_J,\ P_{J-1} z_J,\ P_{J-2}(P_{J-1} z_J),\ \ldots,\ P_1(P_2 \cdots P_{J-1} z_J)],$$

[J. R. Li/White '02]

where

$$z_J = \sqrt{-2p_J}(F + p_J I)^{-1} G$$

and

$$P_i := \frac{\sqrt{-2p_i}}{\sqrt{-2p_{i+1}}} \left[I - (p_i + p_{i+1})(F + p_i I)^{-1} \right].$$

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

Algorithm 1 Low-rank Cholesky factor ADI iteration (LRCF-ADI) [Penzl '97/'00, J. R. Li/White '99/'02, Benner/J. R. Li/Penzl '99/'08]

Input: F, G defining $FX + XF^T = -GG^T$ and shifts $\{p_1, \dots, p_{i_{max}}\}$
Output: $Z = Z_{i_{max}} \in \mathbb{C}^{n \times t_{i_{max}}}$, such that $ZZ^H \approx X$

1: $V_1 = \sqrt{-2\,\mathrm{Re}\,(p_1)}(F + p_1 I)^{-1} G$
2: $Z_1 = V_1$
3: **for** $i = 2, 3, \dots, i_{max}$ **do**
4: $\quad V_i = \sqrt{\mathrm{Re}\,(p_i)/\,\mathrm{Re}\,(p_{i-1})}(V_{i-1} - (p_i + \overline{p_{i-1}})(F + p_i I)^{-1} V_{i-1})$
5: $\quad Z_i = [Z_{i-1} \;\; V_i]$
6: **end for**

Low Rank Solvers for Lyapunov Equations

(G-)LRCF-ADI

Algorithm 2 Generalized Low-rank Cholesky factor ADI iteration (G-LRCF-ADI) [Benner '04, Benner/S. '09, S. '09]

Input: E, F, G defining $FXE^T + EXF^T = -GG^T$ and shifts $\{p_1, \ldots, p_{i_{max}}\}$
Output: $Z = Z_{i_{max}} \in \mathbb{C}^{n \times t_{i_{max}}}$, such that $ZZ^H \approx X$

1: $V_1 = \sqrt{-2\,\mathrm{Re}\,(p_1)}(F + p_1 E)^{-1} G$
2: $Z_1 = V_1$
3: **for** $i = 2, 3, \ldots, i_{max}$ **do**
4: $\quad V_i = \sqrt{\mathrm{Re}\,(p_i) / \mathrm{Re}\,(p_{i-1})}(V_{i-1} - (p_i + \overline{p_{i-1}})(F + p_i E)^{-1} E V_{i-1})$
5: $\quad Z_i = [Z_{i-1} \;\; V_i]$
6: **end for**

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

Consider Schur/singular value decomposition $X = U\Sigma U^T$,
$U \in \mathbb{R}^{n\times n}$, $U^TU = I$, $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$ and $|\sigma_1| \geq |\sigma_2| \geq \dots \geq |\sigma_n|$.
The best rank-m Frobenius-norm approximation to X is thus given by

$$X_m := U \begin{bmatrix} \Sigma_m & 0 \\ 0 & 0 \end{bmatrix} U^T = U_m \Sigma_m U_m^T.$$

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

Consider Schur/singular value decomposition $X = U\Sigma U^T$,
$U \in \mathbb{R}^{n\times n}$, $U^T U = I$, $\Sigma = \mathrm{diag}\,(\sigma_1, \ldots, \sigma_n)$ and $|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_n|$.
The best rank-m Frobenius-norm approximation to X is thus given by

$$X_m := U \begin{bmatrix} \Sigma_m & 0 \\ 0 & 0 \end{bmatrix} U^T = U_m \Sigma_m U_m^T.$$

Krylov Projection Idea [Saad '90, Jaimoukha/Kasenally '94]

Solve

$$(U_m^T F U_m) Y_m + Y_m (U_m^T F^T U_m) = -U_m^T G G^T U_m, \quad (1)$$

on *colspan*(U_m) and get X_m as

$$X_m = U_m Y_m U_m^T.$$

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

Consider Schur/singular value decomposition $X = U\Sigma U^T$,
$U \in \mathbb{R}^{n\times n}$, $U^T U = I$, $\Sigma = \mathrm{diag}\,(\sigma_1, \ldots, \sigma_n)$ and $|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_n|$.
The best rank-m Frobenius-norm approximation to X is thus given by

$$X_m := U \begin{bmatrix} \Sigma_m & 0 \\ 0 & 0 \end{bmatrix} U^T = U_m \Sigma_m U_m^T.$$

Krylov Projection Idea [Saad '90, Jaimoukha/Kasenally '94]

Solve

$$F_m Y_m + Y_m F_m^T = G_m G_m^T, \qquad F_m \in \mathbb{R}^{m\times m},\ m \ll n \tag{1}$$

on *colspan*(U_m) and get X_m as

$$X_m = U_m Y_m U_m^T.$$

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

Consider Schur/singular value decomposition $X = U\Sigma U^T$,
$U \in \mathbb{R}^{n\times n},\ U^T U = I$, $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ and $|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_n|$.
The best rank-m Frobenius-norm approximation to X is thus given by

$$X_m := U \begin{bmatrix} \Sigma_m & 0 \\ 0 & 0 \end{bmatrix} U^T = U_m \Sigma_m U_m^T.$$

Krylov Projection Idea [Saad '90, Jaimoukha/Kasenally '94]

Solve

$$F_m Y_m + Y_m F_m^T = G_m G_m^T, \qquad F_m \in \mathbb{R}^{m\times m},\ m \ll n \tag{1}$$

on *colspan*(U_m) and get X_m as

$$X_m = U_m C_m C_m^T U_m^T \quad \text{where } Y_m = C_m C_m^T.$$

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

Algorithm 3 Basic Krylov Subspace Method for the Lyapunov Equation

Input: F,G defining $FX + XF^T = -GG^T$, an initial Krylov subspace $\mathcal{V}$, e.g., $\mathcal{V} = \mathcal{K}_p(F, G)$ with orthogonal basis $V \in \mathbb{C}^{n\times p}$.
Output: $Z \in \mathbb{C}^{n\times t}$, such that $ZZ^H \approx X$

repeat
  if not first step **then**
    increase dimension of $\mathcal{V}$ and update V.
  end if
  Solve the "small" LE for $\tilde{Z}$ with a classical solver:

$$(V^T FV)\tilde{Z}\tilde{Z}^T + \tilde{Z}\tilde{Z}^T(V^T F^T V) = -V^T GG^T V,$$

  Lift $\tilde{Z}$ to the full space: $Z = U\tilde{Z}$
until res(Z)$<$ TOL

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

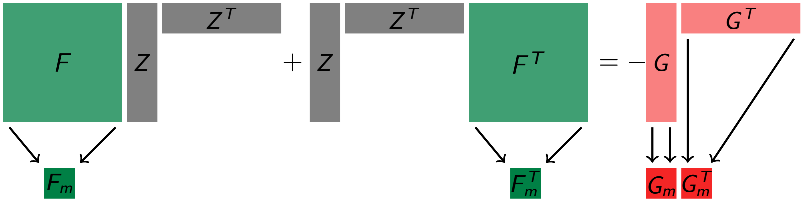
$$F\,Z\,Z^T + Z\,Z^T\,F^T = -G\,G^T$$

Legend:

new factor original matrix projected matrix projected Cholesky factor
old factor original rhs projected rhs

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations



Legend:

new factor original matrix projected matrix projected Cholesky factor
old factor original rhs projected rhs

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations

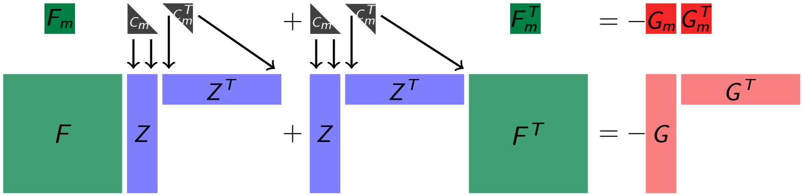
$$F_m \; C_m C_m^T + C_m C_m^T \; F_m^T = -G_m G_m^T$$

Legend:

new factor original matrix projected matrix projected Cholesky factor
old factor original rhs projected rhs

Low Rank Solvers for Lyapunov Equations

Krylov Subspace Based Solvers for Lyapunov Equations



Legend:

new factor | original matrix | projected matrix | projected Cholesky factor

old factor | original rhs | projected rhs

LRCF-NM for the ARE

LRCF-NM for the ARE

Newton's Method for AREs

Consider $\mathfrak{R}(X) := C^T C + A^T X + XA - XBB^T X = 0$

Newton's Iteration for the ARE

$$\mathfrak{R}'|_X(N_\ell) = -\mathfrak{R}(X_\ell), \qquad X_{\ell+1} = X_\ell + N_\ell, \qquad \ell = 0, 1, \dots$$

where the Frechét derivative of $\mathfrak{R}$ at X is the Lyapunov operator

$$\mathfrak{R}'|_X : \quad Q \mapsto (A - BB^T X)^T Q + Q(A - BB^T X),$$

i.e., in every Newton step solve a

Lyapunov Equation [Kleinman '68]

$$(A - BB^T X_\ell)^T X_{\ell+1} + X_{\ell+1}(A - BB^T X_\ell) = -C^T C - X_\ell BB^T X_\ell.$$

LRCF-NM for the ARE

Newton's Method for AREs

Consider $\mathfrak{R}(X) := C^T C + A^T X + XA - XBB^T X = 0$

Newton's Iteration for the ARE

$$\mathfrak{R}'|_X(N_\ell) = -\mathfrak{R}(X_\ell), \qquad X_{\ell+1} = X_\ell + N_\ell, \qquad \ell = 0, 1, \dots$$

where the Frechét derivative of $\mathfrak{R}$ at X is the Lyapunov operator

$$\mathfrak{R}'|_X : \quad Q \mapsto (A - BB^T X)^T Q + Q(A - BB^T X),$$

i.e., in every Newton step solve a

Lyapunov Equation [Kleinman '68]

$$F_\ell^T X_{\ell+1} + X_{\ell+1} F_\ell = -G_\ell G_\ell^T.$$

LRCF-NM for the ARE

Newton's Method for AREs

Consider $\mathfrak{R}(X) := C^TC + A^TXE + E^TXA - E^TXBB^TXE = 0$

Newton's Iteration for the ARE

$$\mathfrak{R}'|_X(N_\ell) = -\mathfrak{R}(X_\ell), \qquad X_{\ell+1} = X_\ell + N_\ell, \qquad \ell = 0, 1, \ldots$$

where the Frechét derivative of $\mathfrak{R}$ at X is the Lyapunov operator

$$\mathfrak{R}'|_X : \quad Q \mapsto (A - BB^TX)^TQE + E^TQ(A - BB^TX),$$

i.e., in every Newton step solve a

Lyapunov Equation [KLEINMAN '68]

$$F_\ell^TX_{\ell+1}E + E^TX_{\ell+1}F_\ell = -\tilde{G}_\ell\tilde{G}_\ell^T.$$

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [Benner/J. R. Li/Penzl '99/'08]

$$F_\ell = A - BB^T X_\ell =: A - BK_\ell$$ is "sparse + low rank"

$$G_\ell = [C^T \; K_\ell^T]$$ is low rank factor

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [Benner/J. R. Li/Penzl '99/'08]

$F_\ell = A - BB^T X_\ell =: A - BK_\ell$ is "sparse + low rank"

$G_\ell = [C^T \;\; K_\ell^T]$ is low rank factor

- apply LRCF-ADI in every Newton step
- exploit structure of F_ℓ using Sherman-Morrison-Woodbury formula

$$(A - BK_\ell + p_k^{(\ell)} I_n)^{-1} = (I_n + (A + p_k^{(\ell)} I_n)^{-1} B (I_m - K_\ell (A + p_k^{(\ell)} I_n)^{-1} B)^{-1} K_\ell)(A + p_k^{(\ell)} I_n)^{-1}$$

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [Benner/J. R. Li/Penzl '99/'08]

$$F_\ell = A - BB^T X_\ell =: A - BK_\ell \quad \text{is "sparse + low rank"}$$
$$G_\ell = [C^T \;\; K_\ell^T] \quad \text{is low rank factor}$$

- apply LRCF-ADI in every Newton step
- exploit structure of F_ℓ using Sherman-Morrison-Woodbury formula

$$(A - BK_\ell + p_k^{(\ell)} I_n)^{-1} =$$
$$(I_n + (A + p_k^{(\ell)} I_n)^{-1} B (I_m - K_\ell (A + p_k^{(\ell)} I_n)^{-1} B)^{-1} K_\ell)(A + p_k^{(\ell)} I_n)^{-1}$$

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [Benner/J. R. Li/Penzl '99/'08]

$$F_\ell = A - BB^T X_\ell =: A - BK_\ell \qquad \text{is "sparse + low rank"}$$
$$G_\ell = [C^T \ \ K_\ell^T] \qquad \text{is low rank factor}$$

- apply LRCF-ADI in every Newton step
- exploit structure of F_ℓ using Sherman-Morrison-Woodbury formula

$$(A - BK_\ell + p_k^{(\ell)} I_n)^{-1} =$$
$$(I_n + (A + p_k^{(\ell)} I_n)^{-1} B(I_m - K_\ell(A + p_k^{(\ell)} I_n)^{-1} B)^{-1} K_\ell)(A + p_k^{(\ell)} I_n)^{-1}$$

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [BENNER/J. R. LI/PENZL '99/'08]

$$F_\ell = A - BB^T X_\ell =: A - BK_\ell \quad \text{is "sparse + low rank"}$$
$$G_\ell = [C^T \;\; K_\ell^T] \quad \text{is low rank factor}$$

- apply LRCF-ADI in every Newton step
- exploit structure of F_ℓ using Sherman-Morrison-Woodbury formula

$$(A - BK_\ell + p_k^{(\ell)} I_n)^{-1} = (I_n + (A + p_k^{(\ell)} I_n)^{-1} B (I_m - K_\ell (A + p_k^{(\ell)} I_n)^{-1} B)^{-1} K_\ell)(A + p_k^{(\ell)} I_n)^{-1}$$

LRCF-NM for the ARE

Low-Rank Newton-ADI (LRCF-NM) for AREs

Factored Newton-Kleinman Iteration [Benner/J. R. Li/Penzl '99/'08]

$$F_\ell = A - BB^T X_\ell =: A - BK_\ell \quad \text{is "sparse + low rank"}$$
$$G_\ell = [C^T \ K_\ell^T] \quad \text{is low rank factor}$$

- apply LRCF-ADI in every Newton step
- exploit structure of F_ℓ using Sherman-Morrison-Woodbury formula

$$(A - BK_\ell + p_k^{(\ell)} E)^{-1} = (I_n + (A + p_k^{(\ell)} E)^{-1} B (I_m - K_\ell (A + p_k^{(\ell)} E)^{-1} B)^{-1} K_\ell)(A + p_k^{(\ell)} E)^{-1}$$

LRCF-NM for the ARE

QADI

$$(A^T - K_{j-1}^T B^T + p_j I) X_{j-\frac{1}{2}}^T = -Q - X_{j-1}^T (A - p_j I),$$
$$(A^T - K_{j-\frac{1}{2}}^T B^T + p_j I) X_j = -Q - X_{j-\frac{1}{2}} (A - p_j I).$$

LRCF-NM for the ARE

QADI

$$
\begin{aligned}
(A^T - K_{j-1}^T B^T + p_j I) X_{j-\frac{1}{2}}^T &= -Q - K_{j-1}^T R K_{j-1} - X_{j-1}^T (A - B K_{j-1} - p_j I),\\
(A^T - K_{j-\frac{1}{2}}^T B^T + p_j I) X_j &= -Q - K_{j-\frac{1}{2}}^T R K_{j-\frac{1}{2}} - X_{j-\frac{1}{2}} (A - B K_{j-\frac{1}{2}} - p_j I).
\end{aligned}
$$

LRCF-NM for the ARE

QADI

$$
\begin{aligned}
(A^T - K_{j-1}^T B^T + p_j I) X_{j-\frac{1}{2}}^T &= -Q - K_{j-1}^T R K_{j-1} - X_{j-1}^T (A - B K_{j-1} - p_j I),\\
(A^T - K_{j-\frac{1}{2}}^T B^T + p_j I) X_j &= -Q - K_{j-\frac{1}{2}}^T R K_{j-\frac{1}{2}} - X_{j-\frac{1}{2}} (A - B K_{j-\frac{1}{2}} - p_j I).
\end{aligned}
$$

NK-ADI

$$
\begin{aligned}
(A^T - K_{j-1}^T B^T + p_k I) X_{k-\frac{1}{2}}^T &= -Q - K_{j-1}^T R K_{j-1} - X_{k-1}^T (A - K_{j-1}^T B^T - p_k I),\\
(A^T - K_{j-1}^T B^T + p_k I) X_k &= -Q - K_{j-1}^T R K_{j-1} - X_{k-\frac{1}{2}} (A - B K_{j-1} - p_k I).
\end{aligned}
$$

ADI Shift Parameter Choice

Min-Max-Problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j \in \mathbb{R} | j=1,\ldots,J\} \subset \mathbb{R}} \quad \max_{\lambda \in \sigma(H), \gamma \in \sigma(V)} \left| \prod_{j=1}^{J} \frac{(p_j - \lambda)(p_j - \gamma)}{(p_j + \lambda)(p_j + \gamma)} \right| .$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different strategies to compute suboptimal shifts for the complex case around 1990.

ADI Shift Parameter Choice

Min-Max-Problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j\in\mathbb{R}|j=1,\ldots,J\}\subset\mathbb{R}} \quad \max_{\lambda\in\sigma(F),} \quad \left|\prod_{j=1}^{J}\frac{(p_j-\lambda)}{(p_j+\lambda)}\right|.$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different strategies to compute suboptimal shifts for the complex case around 1990.

ADI Shift Parameter Choice

Common ADI Shift Parameter Choices

We will discuss three possible parameter choices here:

1. (sub-) optimal parameters [WACHSPRESS: ADI model problem '95] :
   - optimal for real spectra
   - Solve the min-max-problem on an elliptic functions region.
   - Computation needs knowledge of the complete spectrum of F.
   - parameters real if imaginary parts are "small"
2. heuristic parameters [PENZL: Lyapack '99]
   - use Ritz values as shifts
   - suboptimal $\Rightarrow$ convergence might be slow
   - in general complex for complex spectra
3. semi-optimal parameters: Idea: combine the advantages of these methods [BENNER/ MENSA/ S. '08: ALA2006]
   - use Arnoldi's method to approximate the outer spectrum
   - compute the (sub-)optimal real parameters for this approximation