ITERATIVE SOLUTION OF LARGE-SCALE ALGEBRAIC RICCATI EQUATIONS WITH INDEFINITE HESSIAN

Peter Benner

Professur Mathematik in Industrie und Technik
Fakultät für Mathematik
Technische Universität Chemnitz

GAMM Jahrestagung
Bremen, March 31 – April 4, 2008
Goal

Derive numerical algorithms for solving

**(continuous-time) algebraic Riccati equation (ARE)**

with indefinite Hessian,

\[ \mathcal{R}(X) := C^T C + A^T X + XA + X(B_1 B_1^T - B_2 B_2^T)X = 0, \]

where

- \( A \in \mathbb{R}^{n \times n} \) is large and sparse,
- \( B_j \in \mathbb{R}^{n \times m_j} \) (\( j = 1, 2 \)),
- \( C \in \mathbb{R}^{p \times n} \),
- \( n \gg m_j, p \).
Goal

Derive numerical algorithms for solving

(continuous-time) algebraic Riccati equation (ARE)

with indefinite Hessian,

\[ \mathcal{R}(X) := C^T C + A^T X + X A + X \left( B_1 B_1^T - B_2 B_2^T \right) X = 0. \]

### Hessian of \( \mathcal{R}(X) \)

**Frechét derivative** of \( \mathcal{R}(\cdot) \) at \( X \):

\[ \mathcal{R}'_X : Z \rightarrow (A + GX)^T Z + Z(A + GX). \]

**Hessian/2nd order Frechét derivative** of \( \mathcal{R}(\cdot) \) at \( X \):

\[ H : (Z, Y) \rightarrow ZGY + YGZ \]

is indefinite in general unless \( B_1 = 0 \) or \( B_2 = 0 \).
Goal

Derive numerical algorithms for solving

(continuous-time) algebraic Riccati equation (ARE)

with indefinite Hessian,

\[ \mathcal{R}(X) := C^T C + A^T X + X A + X \left( B_1 B_1^T - B_2 B_2^T \right) X = 0. \]

Hessian of \( \mathcal{R}(X) \)

Frechét derivative of \( \mathcal{R}(.) \) at \( X \):

\[ \mathcal{R}_X : Z \rightarrow (A + GX)^T Z + Z (A + GX). \]

Hessian/2nd order Frechét derivative of \( \mathcal{R}(.) \) at \( X \):

\[ \mathcal{H} : (Z, Y) \rightarrow ZGY + YGZ \]

is indefinite in general unless \( B_1 = 0 \) or \( B_2 = 0 \).
Overview

1 Motivation: $H_\infty$-Control

2 Solving Large-Scale Standard AREs
   - Newton’s Method for AREs
   - ADI Method for Lyapunov Equations
   - Low-Rank Newton-ADI for AREs
   - Numerical Results

3 AREs with Indefinite Hessian
   - Lyapunov Iterations/Perturbed Hessian Approach
   - Riccati Iterations
   - Numerical example

4 Conclusions and Open Problems
Motivation: $H_\infty$-Control

Iterative Solution of AREs

Peter Benner

Motivation: $H_\infty$-Control

Large-Scale
Standard AREs

AREs with Indefinite Hessian

Conclusions and Open Problems

Linear time-invariant systems

$$\Sigma : \left\{ \begin{array}{ll} \dot{x} &= A x + B_1 w + B_2 u, \\ z &= C_1 x + D_{11} w + D_{12} u, \\ y &= C_2 x + D_{21} w + D_{22} u, \end{array} \right. \quad \text{where} \ A \in \mathbb{R}^{n \times n}, \ B_k \in \mathbb{R}^{n \times m_k}, \ C_j \in \mathbb{C}^{p_j \times n}, \ D_{jk} \in \mathbb{R}^{p_j \times m_k}. $$

- $x$ – states of the system,
- $w$ – exogenous inputs
- $u$ – control inputs,
- $z$ – performance outputs
- $y$ – measured outputs

\[
x' = A x + B u \\
y = C x + D u \\
v' = E v + F \dot{y} \\
u = H v + K y
\]
Laplace transform $\implies$ transfer function (in frequency domain)

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \equiv \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix}.$$ 

where for $x(0) = 0$, $G_{ij}$ are the rational matrix functions

- $G_{11}(s) = C_1(sI - A)^{-1}B_1 + D_{11},$
- $G_{12}(s) = C_1(sI - A)^{-1}B_2 + D_{12},$
- $G_{21}(s) = C_2(sI - A)^{-1}B_1 + D_{21},$
- $G_{22}(s) = C_2(sI - A)^{-1}B_2 + D_{22},$

describing the transfer from inputs to outputs of $\Sigma$ via

$$z(s) = G_{11}(s)w(s) + G_{12}(s)u(s),$$
$$y(s) = G_{21}(s)w(s) + G_{22}(s)u(s).$$
Consider closed-loop system, where \( K(s) \) is an internally stabilizing controller, i.e., \( K \) stabilizes \( G \) for \( w \equiv 0 \).
Motivation: $H_{\infty}$-Control

The $H_{\infty}$-Optimization Problem

Consider closed-loop system, where $K(s)$ is an **internally stabilizing** controller, i.e., $K$ stabilizes $G$ for $w \equiv 0$.

Goal:

find $K$ that minimize error outputs

$$z = (G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21})w =: \mathcal{F}(G, K)w,$$

where $\mathcal{F}(G, K)$ is the **linear fractional transformation** of $G, K$. 

\[ \text{Diagram:} \quad \begin{array}{ccc}
w & \rightarrow & G(s) \\
& \downarrow & \\
& & K(s) \\
& \uparrow & \\
& & y \\
& \rightarrow & z \\
u & \leftarrow & u
\end{array} \]
Consider closed-loop system, where $K(s)$ is an internally stabilizing controller, i.e., $K$ stabilizes $G$ for $w \equiv 0$.

**Goal:**

find $K$ that minimize error outputs

$$z = (G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21})w =: \mathcal{F}(G, K)w,$$

where $\mathcal{F}(G, K)$ is the linear fractional transformation of $G$, $K$.

**$H_\infty$-optimal control problem:**

$$\min_{K \text{ stabilizing}} \left\| \mathcal{F}(G, K) \right\|_{H_\infty}.$$
Motivation: $H_\infty$-Control

The $H_\infty$-Optimization Problem

Consider closed-loop system, where $K(s)$ is an internally stabilizing controller, i.e., $K$ stabilizes $G$ for $w \equiv 0$.

Goal:

find $K$ that minimize error outputs

$$z = (G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21})w =: \mathcal{F}(G, K)w,$$

where $\mathcal{F}(G, K)$ is the linear fractional transformation of $G, K$.

$H_\infty$-suboptimal control problem:

For given constant $\gamma > 0$, find all internally stabilizing controllers satisfying

$$\|\mathcal{F}(G, K)\|_{H_\infty} < \gamma.$$
Simplifying assumptions

1. $D_{11} = 0$;
2. $D_{22} = 0$;
3. $(A, B_1)$ stabilizable, $(C_1, A)$ detectable;
4. $(A, B_2)$ stabilizable, $(C_2, A)$ detectable ($\Rightarrow$ $\Sigma$ internally stabilizable);
5. $D_{12}^T [C_1 \ D_{12}] = [0 \ I_{m_2}]$;
6. $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^T = \begin{bmatrix} 0 \\ I_{p_2} \end{bmatrix}$.

Remark. 1., 2., 5., 6. only for notational convenience, 3. can be relaxed, but derivations get even more complicated.
Motivation: $H_\infty$-Control

Solution of the $H_\infty$-(Sub-)Optimal Control Problem

Theorem [Doyle/Glover/Khargonekar/Francis ’89]

Given the Assumptions 1.–6., there exists an admissible controller $K(s)$ solving the $H_\infty$-suboptimal control problem $\iff$

(i) There exists a solution $X_\infty = X_\infty^T \geq 0$ to the ARE

$$C_1C_1^T + A^TX + XA + X(\gamma^{-2}B_1B_1^T - B_2B_2^T)X = 0,$$

such that $A_X$ is Hurwitz, where $A_X := A + (\gamma^{-2}B_1B_1^T - B_2B_2^T)X_\infty$.

(ii) There exists a solution $Y_\infty = Y_\infty^T \geq 0$ to the ARE

$$B_1B_1^T + AY + YA^T + Y(\gamma^{-2}C_1C_1^T - C_2C_2^T)Y = 0,$$

such that $A_Y$ is Hurwitz where $A_Y := A + Y_\infty(\gamma^{-2}C_1C_1^T - C_2C_2^T)$.

(iii) $\gamma^2 > \rho(X_\infty Y_\infty)$.

$H_\infty$-optimal control

Find minimal $\gamma$ for which (i)–(iii) are satisfied $\Rightarrow \gamma$-iteration based on solving AREs (1)–(2) repeatedly for different $\gamma$. 
Iteration Solution of AREs

Motivation: $H_\infty$-Control

Solution of the $H_\infty$-(Sub-)Optimal Control Problem

**Theorem** [Doyle/Glover/Khargonekar/Francis '89]

Given the Assumptions 1.–6., there exists an admissible controller $K(s)$ solving the $H_\infty$-suboptimal control problem $\iff$

(i) There exists a solution $X_\infty = X_\infty^T \geq 0$ to the ARE 

$$C_1 C_1^T + A^T X + XA + X(\gamma^{-2} B_1 B_1^T - B_2 B_2^T)X = 0,$$

such that $A_X$ is Hurwitz, where $A_X := A + (\gamma^{-2} B_1 B_1^T - B_2 B_2^T)X_\infty$.

(ii) There exists a solution $Y_\infty = Y_\infty^T \geq 0$ to the ARE 

$$B_1 B_1^T + AY + YA^T + Y(\gamma^{-2} C_1 C_1^T - C_2 C_2^T)Y = 0,$$

such that $A_Y$ is Hurwitz where $A_Y := A + Y_\infty (\gamma^{-2} C_1 C_1^T - C_2 C_2^T)$.

(iii) $\gamma^2 > \rho(X_\infty Y_\infty)$.

$H_\infty$-optimal control

Find minimal $\gamma$ for which (i)–(iii) are satisfied $\implies$ $\gamma$-iteration based on solving AREs (1)–(2) repeatedly for different $\gamma$. 
Motivation: $H_\infty$-Control
Solution of the $H_\infty$-(Sub-)Optimal Control Problem

$H_\infty$-(sub-)optimal controller

If (i)–(iii) hold, a suboptimal controller is given by

$$\hat{K}(s) = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & 0 \end{bmatrix} = \hat{C}(sI_n - \hat{A})^{-1} \hat{B},$$

where for

$$Z_\infty := (I - \gamma^{-2} Y_\infty X_\infty)^{-1},$$

$$\hat{A} := A + (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) X_\infty - Z_\infty Y_\infty C_2^T C_2,$$

$$\hat{B} := Z_\infty Y_\infty C_2^T,$$

$$\hat{C} := -B_2^T X_\infty.$$
Solving Large-Scale Standard AREs

General form for $A, G = G^T, W = W^T \in \mathbb{R}^{n \times n}$ given and $X \in \mathbb{R}^{n \times n}$ unknown:

$$0 = \mathcal{R}(X) := A^T X + XA - XGX + W.$$

Large-scale AREs from semi-discretized PDE control problems:

- $n = 10^3 - 10^6$ ($\longrightarrow 10^6 - 10^{12}$ unknowns!),
- $A$ has sparse representation ($A = -M^{-1}K$ for FEM),
- $G, W$ low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}, m \ll n, C \in \mathbb{R}^{p \times n}, p \ll n$.
- Standard (eigenproblem-based) $O(n^3)$ methods are not applicable!
Solving Large-Scale Standard AREs

General form for $A, G = G^T, W = W^T \in \mathbb{R}^{n \times n}$ given and $X \in \mathbb{R}^{n \times n}$ unknown:

$$0 = \mathcal{R}(X) := A^T X + X A - X G X + W.$$

Large-scale AREs from semi-discretized PDE control problems:

- $n = 10^3 - 10^6 \implies 10^6 - 10^{12}$ unknowns!,
- $A$ has sparse representation ($A = -M^{-1} K$ for FEM),
- $G, W$ low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}, m \ll n, C \in \mathbb{R}^{p \times n}, p \ll n.$
- Standard (eigenproblem-based) $\mathcal{O}(n^3)$ methods are not applicable!
Solving Large-Scale Standard AREs

General form for $A$, $G = G^T$, $W = W^T \in \mathbb{R}^{n \times n}$ given and $X \in \mathbb{R}^{n \times n}$ unknown:

$$0 = \mathcal{R}(X) := A^T X + XA - XGX + W.$$

Large-scale AREs from semi-discretized PDE control problems:

- $n = 10^3 - 10^6 \implies 10^6 - 10^{12} \text{ unknowns!}$,
- $A$ has sparse representation ($A = -M^{-1}K$ for FEM),
- $G, W$ low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}$, $m \ll n$, $C \in \mathbb{R}^{p \times n}$, $p \ll n$.
- Standard (eigenproblem-based) $O(n^3)$ methods are not applicable!
Solving Large-Scale Standard AREs

Iterative Solution of AREs

Peter Benner

Motivation: $H_\infty$ -Control

Large-Scale Standard AREs

Newton’s Method for AREs

ADI for Lyapunov

Low-Rank Newton-ADI

Numerical Results

AREs with Indefinite Hessian

Conclusions and Open Problems

General form for $A, G = G^T, W = W^T \in \mathbb{R}^{n \times n}$ given and $X \in \mathbb{R}^{n \times n}$ unknown:

$$0 = \mathcal{R}(X) := A^T X + XA - XGX + W.$$ 

Large-scale AREs from semi-discretized PDE control problems:

- $n = 10^3 - 10^6$ ($\Longrightarrow 10^6 - 10^{12}$ unknowns!),
- $A$ has sparse representation ($A = -M^{-1}K$ for FEM),
- $G, W$ low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}, m \ll n, C \in \mathbb{R}^{p \times n}, p \ll n$.
- Standard (eigenproblem-based) $O(n^3)$ methods are not applicable!
Solving Large-Scale Standard AREs

Iterative Solution of AREs

Peter Benner

Motivation: $H_\infty$-Control
Large-Scale
Standard AREs
Newton’s
Method for AREs
ADI for Lyapunov
Low-Rank
Newton-ADI
Numerical
Results
AREs with Indefinite Hessian
Conclusions and Open Problems

General form for $A, G = G^T, W = W^T \in \mathbb{R}^{n \times n}$ given and $X \in \mathbb{R}^{n \times n}$ unknown:

$$0 = \mathcal{R}(X) := A^T X + XA - XGX + W.$$

Large-scale AREs from semi-discretized PDE control problems:

- $n = 10^3 - 10^6$ ($\implies 10^6 - 10^{12}$ unknowns!),
- $A$ has sparse representation ($A = -M^{-1}K$ for FEM),
- $G, W$ low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}, m \ll n, C \in \mathbb{R}^{p \times n}, p \ll n$.
- Standard (eigenproblem-based) $O(n^3)$ methods are not applicable!
Consider spectrum of ARE solution (analogous for Lyapunov equations).

Example:
- Linear 1D heat equation with point control,
- \( \Omega = [0, 1] \),
- FEM discretization using linear B-splines,
- \( h = 1/100 \implies n = 101 \).

Idea: \( X = X^T \succeq 0 \implies \)

\[
X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx Y^{(r)} (Y^{(r)})^T = \sum_{k=1}^{r} \lambda_k y_k y_k^T.
\]
Consider spectrum of ARE solution (analogous for Lyapunov equations).

**Example:**

- Linear 1D heat equation with point control,
- \( \Omega = [0, 1] \),
- FEM discretization using linear B-splines,
- \( h = 1/100 \implies n = 101 \).

**Idea:** \( X = X^T \geq 0 \implies \)

\[
X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx Y^{(r)} (Y^{(r)})^T = \sum_{k=1}^{r} \lambda_k y_k y_k^T.
\]
Consider spectrum of ARE solution (analogous for Lyapunov equations).

Example:
- Linear 1D heat equation with point control,
- $\Omega = [0, 1]$,
- FEM discretization using linear B-splines,
- $h = 1/100 \implies n = 101$.

Idea: $X = X^T \geq 0 \implies$

\[
X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx Y^{(r)}(Y^{(r)})^T = \sum_{k=1}^{r} \lambda_k y_k y_k^T.
\]
Consider \( 0 = \mathcal{R}(X) = C^T C + A^T X + XA - XBB^T X \).

Frechét derivative of \( \mathcal{R}(X) \) at \( X \):

\[
\mathcal{R}'_X : Z \rightarrow (A - BB^T X)^T Z + Z(A - BB^T X).
\]

Newton-Kantorovich method:

\[
X_{j+1} = X_j - \left( \mathcal{R}'_X \right)^{-1} \mathcal{R}(X_j), \quad j = 0, 1, 2, \ldots
\]

Newton’s method (with line search) for AREs

FOR \( j = 0, 1, \ldots \)

1. \( A_j \leftarrow A - BB^T X_j =: A - BK_j \).
2. Solve the Lyapunov equation \( A_j^T N_j + N_j A_j = -\mathcal{R}(X_j) \).
3. \( X_{j+1} \leftarrow X_j + t_j N_j \).

END FOR \( j \)
Newton’s Method for AREs

Consider \( 0 = \mathcal{R}(X) = C^T C + A^T X + X A - X B B^T X \).

Frechét derivative of \( \mathcal{R}(X) \) at \( X \):

\[
\mathcal{R}_X' : Z \rightarrow (A - B B^T X)^T Z + Z(A - B B^T X).
\]

Newton-Kantorovich method:

\[
X_{j+1} = X_j - \left( \mathcal{R}_X' \right)^{-1} \mathcal{R}(X_j), \quad j = 0, 1, 2, \ldots
\]

Newton’s method (with line search) for AREs

FOR \( j = 0, 1, \ldots \)

1. \( A_j \leftarrow A - B B^T X_j =: A - B K_j. \)
2. Solve the Lyapunov equation \( A_j^T N_j + N_j A_j = -\mathcal{R}(X_j). \)
3. \( X_{j+1} \leftarrow X_j + t_j N_j. \)

END FOR \( j \)
Newton’s Method for AREs

M克莱因曼 ’68, 梅尔曼 ’91, 兰卡斯特/罗德曼 ’95, B./贝耶斯 ’94/’98, B. ’97, 郭/劳布 ’99

- Consider $0 = \mathcal{R}(X) = C^T C + A^T X + X A - X B B^T X$.
- Frechét derivative of $\mathcal{R}(X)$ at $X$:
  
  $\mathcal{R}'_X : Z \rightarrow (A - B B^T X)^T Z + Z (A - B B^T X)$.

- Newton-Kantorovich method:
  
  $X_{j+1} = X_j - (\mathcal{R}'_X)^{-1} \mathcal{R}(X_j), \quad j = 0, 1, 2, \ldots$

Newton’s method (with line search) for AREs

FOR $j = 0, 1, \ldots$

2. Solve the Lyapunov equation $A_j^T N_j + N_j A_j = -\mathcal{R}(X_j)$.
3. $X_{j+1} \leftarrow X_j + t_j N_j$.

END FOR $j$
**Newton's Method for AREs**

*Kleinman '68, Mehrmann '91, Lancaster/Rodman '95, B./Byers '94/’98, B. ’97, Guo/Laub ’99*

- Consider $0 = R(X) = C^T C + A^T X + XA - XBB^T X$.
- Frechét derivative of $R(X)$ at $X$:


- Newton-Kantorovich method:

  $$X_{j+1} = X_j - \left(R'_X \right)^{-1} R(X_j), \quad j = 0, 1, 2, \ldots$$

---

**Newton's method (with line search) for AREs**

FOR $j = 0, 1, \ldots$

1. $A_j \leftarrow A - BB^T X_j =: A - BK_j$.
2. Solve the Lyapunov equation $A_j^T N_j + N_j A_j = -R(X_j)$.
3. $X_{j+1} \leftarrow X_j + t_j N_j$.

END FOR $j$
Newton’s Method for AREs
Properties and Implementation

- Convergence for $K_0$ stabilizing:
  - $A_j = A - BK_j = A - BB^T X_j$ is stable $\forall j \geq 0$.
  - $\lim_{j \to \infty} \|R(X_j)\|_F = 0$ (monotonically).
  - $\lim_{j \to \infty} X_j = X_* \geq 0$ (locally quadratic).

- Need large-scale Lyapunov solver; here, ADI iteration:
  linear systems with dense, but “sparse+low rank” coefficient
  matrix $A_j$:
  \[
  A_j = \begin{bmatrix}
  A & -B \\
  \text{sparse} & \text{m} \\
  \end{bmatrix} \cdot K_j
  \]

- $m \ll n \implies$ efficient “inversion” using
  Sherman-Morrison-Woodbury formula:
  \[
  (A - BK_j)^{-1} = (I_n + A^{-1}B(I_m - K_jA^{-1}B)^{-1}K_j)A^{-1}.
  \]

- BUT: $X = X^T \in \mathbb{R}^{n \times n} \implies n(n + 1)/2$ unknowns!
Iterative Solution

Newton’s Method for AREs

Properties and Implementation

Motivation: $H_\infty$-Control
Large-Scale
Standard AREs

Newton’s Method for AREs
ADI for Lyapunov
Low-Rank Newton-ADI
Numerical Results

AREs with Indefinite Hessian
Conclusions and Open Problems

- Convergence for $K_0$ stabilizing:
  - $A_j = A - BK_j = A - BB^T X_j$ is stable $\forall j \geq 0$.
  - $\lim_{j \to \infty} \|R(X_j)\|_F = 0$ (monotonically).
  - $\lim_{j \to \infty} X_j = X^*_\geq 0$ (locally quadratic).

- Need large-scale Lyapunov solver; here, ADI iteration:
  linear systems with dense, but “sparse+low rank” coefficient matrix $A_j$:

\[
A_j = \begin{pmatrix} A & -B \cdot K_j \\ \text{sparse} & m \cdot \end{pmatrix}
\]

- $m \ll n \implies$ efficient “inversion” using
  Sherman-Morrison-Woodbury formula:

\[
(A - BK_j)^{-1} = (I_n + A^{-1}B(I_m - K_jA^{-1}B)^{-1}K_j)A^{-1}.
\]

- BUT: $X = X^T \in \mathbb{R}^{n \times n} \implies n(n+1)/2$ unknowns!
Newton's Method for AREs
Properties and Implementation

Convergence for $K_0$ stabilizing:

- $A_j = A - BK_j = A - BB^T X_j$ is stable $\forall j \geq 0$.
- $\lim_{j \to \infty} ||R(X_j)||_F = 0$ (monotonically).
- $\lim_{j \to \infty} X_j = X_* \geq 0$ (locally quadratic).

Need large-scale Lyapunov solver; here, ADI iteration:
linear systems with dense, but “sparse+low rank” coefficient matrix $A_j$:

\[
A_j = A - BK_j = \begin{bmatrix} A_{\text{sparse}} & -m \end{bmatrix} \cdot K_j
\]

$m \ll n \implies$ efficient “inversion” using
Sherman-Morrison-Woodbury formula:

\[
(A - BK_j)^{-1} = (I_n + A^{-1} B (I_m - K_j A^{-1} B)^{-1} K_j) A^{-1}.
\]

BUT: $X = X^T \in \mathbb{R}^{n \times n} \implies n(n + 1)/2$ unknowns!
Newton’s Method for AREs
Properties and Implementation

**Iterative Solution**

- **Motivation:** $H_\infty$-Control
- **Large-Scale**
  - Standard AREs
  - Newton’s Method for AREs
  - ADI for Lyapunov
  - Low-Rank
  - Newton-ADI
- **Numerical Results**
- **AREs with Indefinite Hessian**
- **Conclusions and Open Problems**

---

**Newton’s Method for AREs**

- **Properties and Implementation**

- **Convergence for $K_0$ stabilizing:**
  - $A_j = A - BK_j = A - BB^T X_j$ is stable $\forall j \geq 0$.
  - $\lim_{j \to \infty} \|R(X_j)\|_F = 0$ (monotonically).
  - $\lim_{j \to \infty} X_j = X_* \geq 0$ (locally quadratic).

- **Need large-scale Lyapunov solver; here, ADI iteration:**
  - linear systems with dense, but “sparse+low rank” coefficient matrix $A_j$:
    
    \[
    A_j = \begin{bmatrix}
    A & -B \\
    \text{sparse} & -m
    \end{bmatrix} \cdot K_j
    \]

    $n \ll m \implies$ efficient “inversion” using Sherman-Morrison-Woodbury formula:

    \[
    (A - BK_j)^{-1} = (I_n + A^{-1}B(I_m - K_jA^{-1}B)^{-1}K_j)A^{-1}.
    \]

- **BUT:** $X = X^T \in \mathbb{R}^{n \times n} \implies n(n + 1)/2$ unknowns!
For $A \in \mathbb{R}^{n \times n}$ stable, $B \in \mathbb{R}^{n \times m}$ ($w \ll n$), consider Lyapunov equation

$$AX +XA^T = -BB^T.$$ 

- **ADI Iteration:** [WACHSPRESS 1988]

  \[
  (A + p_k I)X_{(k-1)/2} = -BB^T - X_{k-1}(A^T - p_k I) \\
  (A + \overline{p_k} I)X_k^T = -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)
  \]

  with parameters $p_k \in \mathbb{C}^-$ and $p_{k+1} = \overline{p_k}$ if $p_k \notin \mathbb{R}$.

- For $X_0 = 0$ and proper choice of $p_k$: $\lim_{k \to \infty} X_k = X$ superlinear.

- Re-formulation using $X_k = Y_k Y_k^T$ yields iteration for $Y_k$...
For $A \in \mathbb{R}^{n \times n}$ stable, $B \in \mathbb{R}^{n \times m}$ ($w \ll n$), consider Lyapunov equation

$$AX + XA^T = -BB^T.$$

**ADI Iteration:**

$$\begin{align*}
(A + p_k I)X_{(k-1)/2} &= -BB^T - X_{k-1}(A^T - p_k I) \\
(A + \overline{p_k} I)X_k^T &= -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)
\end{align*}$$

with parameters $p_k \in \mathbb{C}^-$ and $p_{k+1} = \overline{p_k}$ if $p_k \not\in \mathbb{R}$.

For $X_0 = 0$ and proper choice of $p_k$: $\lim_{k \to \infty} X_k = X$ superlinear.

Re-formulation using $X_k = Y_k Y_k^T$ yields iteration for $Y_k$...
Iterative Solution of AREs

Peter Benner

Motivation: $\mathcal{H}_\infty$-Control

Large-Scale Standard AREs

Newton’s Method for AREs

ADI for Lyapunov

Low-Rank Newton-ADI

Numerical Results

AREs with Indefinite Hessian

Conclusions and Open Problems

ADIMethod for Lyapunov Equations

- For $A \in \mathbb{R}^{n \times n}$ stable, $B \in \mathbb{R}^{n \times m}$ ($w \ll n$), consider Lyapunov equation

  $$AX + XA^T = -BB^T.$$  

- ADI Iteration: $[\text{WACHSPRESS 1988}]$

  $$(A + p_k I)X_{(k-1)/2} = -BB^T - X_{k-1}(A^T - p_k I)$$
  $$(A + \overline{p_k} I)X_k^T = -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)$$

  with parameters $p_k \in \mathbb{C}^-$ and $p_{k+1} = \overline{p_k}$ if $p_k \notin \mathbb{R}$.

- For $X_0 = 0$ and proper choice of $p_k$: $\lim_{k \to \infty} X_k = X$ superlinear.

- Re-formulation using $X_k = Y_k Y_k^T$ yields iteration for $Y_k$...
For $A \in \mathbb{R}^{n \times n}$ stable, $B \in \mathbb{R}^{n \times m}$ ($w \ll n$), consider Lyapunov equation

$$AX + XA^T = -BB^T.$$ 

ADI Iteration: \cite{WACHSPRESS 1988}

\[
(A + p_k I)X_{(k-1)/2} = -BB^T - X_{k-1}(A^T - p_k I)
\]

\[
(A + \overline{p_k} I)X_k^T = -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)
\]

with parameters $p_k \in \mathbb{C}^-$ and $p_{k+1} = \overline{p_k}$ if $p_k \notin \mathbb{R}$.

For $X_0 = 0$ and proper choice of $p_k$: $\lim_{k \to \infty} X_k = X$ superlinear.

Re-formulation using $X_k = Y_k Y_k^T$ yields iteration for $Y_k$...
Factored ADI Iteration
Lyapunov equation $0 = AX + XA^T = -BB^T$.

Setting $X_k = Y_k Y_k^T$, some algebraic manipulations $\implies$


\begin{align*}
V_1 & \leftarrow \sqrt{-2\text{Re}(p_1)}(A + p_1I)^{-1}B, \quad Y_1 \leftarrow V_1 \\
\text{FOR } j = 2, 3, \ldots \\
V_k & \leftarrow \sqrt{\frac{\text{Re}(p_k)}{\text{Re}(p_{k-1})}} (V_{k-1} - (p_k + p_{k-1})(A + p_kI)^{-1}V_{k-1}), \\
Y_k & \leftarrow \begin{bmatrix} Y_{k-1} & V_k \end{bmatrix} \\
Y_k & \leftarrow \text{rrqr}(Y_k, \tau) \quad \% \text{ column compression}
\end{align*}

At convergence, $Y_{k_{\text{max}}} Y_{k_{\text{max}}}^T \approx X$, where

\[
\text{range } (Y_{k_{\text{max}}}) = \text{range } \left( \begin{bmatrix} V_1 & \ldots & V_{k_{\text{max}}} \end{bmatrix} \right), \quad V_k = \square \in \mathbb{C}^{n \times m}.
\]

Note: Implementation in real arithmetic possible by combining two steps.
Iterative Solution of AREs

Peter Benner

Motivation: $H_\infty$-Control
Large-Scale Standard AREs
Newton's Method for AREs
ADI for Lyapunov
Low-Rank Newton-ADI
Numerical Results
AREs with Indefinite Hessian
Conclusions and Open Problems

Factored ADI Iteration

Lyapunov equation $0 = AX + XA^T = -BB^T$.

Setting $X_k = Y_k Y_k^T$, some algebraic manipulations $\implies$


\[
V_1 \leftarrow \sqrt{-2\text{Re}(p_1)}(A + p_1 I)^{-1}B, \quad Y_1 \leftarrow V_1
\]

\[
\text{FOR } j = 2, 3, \ldots \]

\[
V_k \leftarrow \sqrt{\frac{\text{Re}(p_k)}{\text{Re}(p_{k-1})}} \left(V_{k-1} - (p_k + \overline{p_{k-1}})(A + p_k I)^{-1}V_{k-1}\right),
\]

\[
Y_k \leftarrow \begin{bmatrix} Y_{k-1} & V_k \end{bmatrix}
\]

\[
Y_k \leftarrow \text{rrqr}(Y_k, \tau) \quad \% \text{ column compression}
\]

At convergence, $Y_{k_{\text{max}}} Y_{k_{\text{max}}}^T \approx X$, where

\[
\text{range}(Y_{k_{\text{max}}}) = \text{range} \left( \begin{bmatrix} V_1 & \ldots & V_{k_{\text{max}}} \end{bmatrix} \right), \quad V_k = \begin{bmatrix} \vdots \end{bmatrix} \in \mathbb{C}^{n \times m}.
\]

Note: Implementation in real arithmetic possible by combining two steps.
Low-Rank Newton-ADI for AREs

Re-write Newton’s method for AREs

\[
A_j^T N_j + N_j A_j = -R(X_j)
\]

\[
\iff
\]

\[
A_j^T (X_j + N_j) + (X_j + N_j) A_j = -C^T C - X_j B B^T X_j
\]

Set \(X_j = Z_j Z_j^T\) for rank \((Z_j) \ll n\)

\[
A_j^T (Z_{j+1} Z_{j+1}^T) + (Z_{j+1} Z_{j+1}^T) A_j = -W_j W_j^T
\]

Factored Newton Iteration [B./LI/PENZL 1999/2008]

Solve Lyapunov equations for \(Z_{j+1}\) directly by factored ADI iteration and use ‘sparse + low-rank’ structure of \(A_j\).
Re-write Newton’s method for AREs

\[ A_j^T N_j + N_j A_j = -R(X_j) \]

\[ \iff \]

\[ A_j^T (X_j + N_j) + (X_j + N_j) A_j = -C^T C - X_j BB^T X_j \]

Set \( X_j = Z_j Z_j^T \) for rank \( (Z_j) \ll n \) \( \implies \)

\[ A_j^T (Z_{j+1} Z_{j+1}^T) + (Z_{j+1} Z_{j+1}^T) A_j = -W_j W_j^T \]

Factored Newton Iteration \[\text{[B./Li/Penzl 1999/2008]}\]

Solve Lyapunov equations for \( Z_{j+1} \) directly by factored ADI iteration and use ‘sparse + low-rank’ structure of \( A_j \).
- Linear 2D heat equation with homogeneous Dirichlet boundary and point control/observation.
- FD discretization on uniform $150 \times 150$ grid.
- $n = 22.500$, $m = p = 1$, 10 shifts for ADI iterations.
- Convergence of large-scale matrix equation solvers:
Solving Large-Scale Standard AREs
Performance of matrix equation solvers

Iterative Solution of AREs
Peter Benner

Motivation: $H_\infty$-Control
Large-Scale Standard AREs
Newton’s Method for AREs
ADI for Lyapunov
Low-Rank Newton-ADI
Numerical Results
AREs with Indefinite Hessian
Conclusions and Open Problems

Performance of Newton’s method for accuracy $\sim 1/n$

<table>
<thead>
<tr>
<th>grid</th>
<th>unknowns</th>
<th>$\frac{|R(X)|_F}{|X|_F}$</th>
<th>it. (ADI it.)</th>
<th>CPU (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$8 \times 8$</td>
<td>2,080</td>
<td>4.7e-7</td>
<td>2 (8)</td>
<td>0.47</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>32,896</td>
<td>1.6e-6</td>
<td>2 (10)</td>
<td>0.49</td>
</tr>
<tr>
<td>$32 \times 32$</td>
<td>524,800</td>
<td>1.8e-5</td>
<td>2 (11)</td>
<td>0.91</td>
</tr>
<tr>
<td>$64 \times 64$</td>
<td>8,390,656</td>
<td>1.8e-5</td>
<td>3 (14)</td>
<td>7.98</td>
</tr>
<tr>
<td>$128 \times 128$</td>
<td>134,225,920</td>
<td>3.7e-6</td>
<td>3 (19)</td>
<td>79.46</td>
</tr>
</tbody>
</table>

Here,

- Convection-diffusion equation,
- $m = 1$ input and $p = 2$ outputs,
- $X = X^T \in \mathbb{R}^{n \times n} \Rightarrow \frac{n(n+1)}{2}$ unknowns.
Back to

\[ \mathcal{R}(X) := C^T C + A^T X + X A + X (B_1 B_1^T - B_2 B_2^T) X = 0. \]
Iterative Solution of AREs

Peter Benner

Motivation: $H_\infty$-Control

Large-Scale Standard AREs

AREs with Indefinite Hessian

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Numerical example

Conclusions and Open Problems

AREs with Indefinite Hessian

Back to

\[ \mathcal{R}(X) := C^T C + A^T X + X A + X (B_1 B_1^T - B_2 B_2^T) X = 0. \]

Problems

- For large-scale problems, resulting, e.g., from $H_\infty$ control, standard methods based on Hamiltonian/even eigenvalue problem can not be used due to $O(n^3)$ complexity/dense matrix algebra.

- Krylov subspace methods might be employed, but so far no convergence results, and in case of convergence, no guarantee that stabilizing solution is computed.

- Newton/Newton-ADI method will in general diverge/converge to a non-stabilizing solution.
Iterative Solution of AREs

Peter Benner

Motivation: $H_{\infty}$-Control

Large-Scale Standard AREs

AREs with Indefinite Hessian

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Numerical example

Conclusions and Open Problems

Back to

$$\mathcal{R}(X) := C^T C + A^T X + XA + X (B_1 B_1^T - B_2 B_2^T) X = 0.$$ 

Problems

- For large-scale problems, resulting, e.g., from $H_{\infty}$ control, standard methods based on Hamiltonian/even eigenvalue problem can not be used due to $O(n^3)$ complexity/dense matrix algebra.

- Krylov subspace methods might be employed, but so far no convergence results, and in case of convergence, no guarantee that stabilizing solution is computed.

- Newton/Newton-ADI method will in general diverge/converge to a non-stabilizing solution.
Back to

\[ R(X) := C^T C + A^T X + XA + X(B_1 B_1^T - B_2 B_2^T)X = 0. \]

**Problems**

- For large-scale problems, resulting, e.g., from $H_\infty$ control, standard methods based on Hamiltonian/even eigenvalue problem can not be used due to $O(n^3)$ complexity/dense matrix algebra.

- Krylov subspace methods might be employed, but so far no convergence results, and in case of convergence, no guarantee that stabilizing solution is computed.

- Newton/Newton-ADI method will in general diverge/converge to a non-stabilizing solution.
Iterative Solution of AREs

Peter Benner

Motivation: $H_{\infty}$-Control

Large-Scale Standard AREs

AREs with Indefinite Hessian

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations Numerical example

Conclusions and Open Problems

AREs with Indefinite Hessian

Back to

$$\mathcal{R}(X) := C^T C + A^T X +XA + X (B_1 B_1^T - B_2 B_2^T)X = 0.$$ 

Problems

Quick-and-dirty solution: consider $X^{-1} \mathcal{R}(X)X^{-1} = 0 \quad [\text{DAMM 2002}]$

$\Rightarrow$ standard ARE for $\tilde{X} \equiv X^{-1}$

$$\tilde{\mathcal{R}}(\tilde{X}) := (B_1 B_1^T - B_2 B_2^T) + \tilde{X}A^T + A\tilde{X} + \tilde{X}C^T C\tilde{X} = 0.$$ 

Newton’s method will converge to stabilizing solution, Newton-ADI can be employed (with modification for indefinite constant term).

But: low-rank approximation of $\tilde{X}$ will not yield good approximation of $X \Rightarrow$ not feasible for large-scale problems!
Idea

Perturb Hessian to enforce semi-definiteness: write

\[0 = A^T X +XA + Q - XG X = A^T X +XA + Q - XDX + X(D - G)X,\]

where \(D = G + \alpha I \geq 0\) with \(\alpha \geq \min\{0, -\lambda_{\text{max}}(G)\} \).
Lyapunov Iterations/Perturbed Hessian Approach
[Cherfi/Abou-Kandil/Bourles 2005 (Proc. ACSE 2005)]

Idea

Perturb Hessian to enforce semi-definiteness: write

\[ 0 = A^T X + XA + Q - XGX = A^T X + XA + Q - XDX + X(D - G)X, \]

where \( D = G + \alpha I \geq 0 \) with \( \alpha \geq \min\{0, -\lambda_{\max}(G)\}. \)

Here: \( G = B_2B_2^T - B_1B_1^T \)

\[ \Rightarrow \text{use } \alpha = \|B_1\|^2 \text{ for spectral/Frobenius norm or } \]

\[ \alpha = \|B_1\|_1 \cdot \|B_1\|_\infty. \]

Remark

\( W \geq -G \) can be used instead of \( \alpha I \), e.g., \( W = \beta B_1B_1^T \) with \( \beta \geq 1. \)
Lyapunov Iterations/Perturbed Hessian Approach
[Cherfi/Abou-Kandil/Bourles 2005 (Proc. ACSE 2005)]

Idea

Perturb Hessian to enforce semi-definiteness: write

\[
0 = A^T X + XA + Q - XGX = A^T X + XA + Q - XDX + X(D - G)X,
\]

where \( D = G + \alpha I \geq 0 \) with \( \alpha \geq \min\{0, -\lambda_{\text{max}}(G)\} \).

Lyapunov iteration

Based on

\[
(A - DX)^T X + X(A - DX) = -Q - XDX - \alpha X^2,
\]

iterate

\[
\text{FOR } k = 0, 1, \ldots, \text{ solve Lyapunov equation }
\]

\[
(A - DX_k)^T X_{k+1} + X_{k+1}(A - DX_k) = -Q - X_k DX_k - \alpha X_k^2.
\]
Lyapunov Iterations/Perturbed Hessian Approach [Cherfi/Abou-Kandil/Bourles 2005 (Proc. ACSE 2005)]

Lyapunov iteration

FOR $k = 0, 1, \ldots$, solve Lyapunov equation

$$(A - DX_k)^T X_{k+1} + X_{k+1}(A - DX_k) = -Q - X_k DX_k - \alpha X_k^2.$$  

Easy to convert to low-rank iteration employing low-rank ADI for Lyapunov equations, e.g. with $W = B_1 B_1^T$ instead of $\alpha I$: the Lyapunov equation becomes

$$(A - B_2 B_2^T Y_k Y_k)^T Y_{k+1} Y_{k+1}^T + Y_{k+1} Y_{k+1}^T (A - B_2 B_2^T Y_k Y_k) = -CC^T - Y_k Y_k^T B_1 B_1^T Y_k Y_k^T - Y_k Y_k^T B_2 B_2^T Y_k Y_k^T$$

$$= -[C, Y_k Y_k^T B_1, Y_k Y_k^T B_2]
\begin{bmatrix}
C^T \\
B_1^T Y_k Y_k^T \\
B_2^T Y_k Y_k^T
\end{bmatrix}.$$
Lyapunov Iterations/Perturbed Hessian Approach

Convergence

Theorem [Cherfi/Abou-Kandil/Bourles 2005]

If

- \( \exists \hat{X} \) such that \( \mathcal{R}(\hat{X}) \geq 0 \),
- \( \exists X_0 = X_0^T \geq \hat{X} \) such that \( \mathcal{R}(X_0) \leq 0 \) and \( A - DX_0 \) is Hurwitz,

then

a) \( X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X} \),
b) \( \mathcal{R}(X_k) \leq 0 \) for all \( k = 0, 1, \ldots \),
c) \( A - DX_k \) is Hurwitz for all \( k = 0, 1, \ldots \),
d) \( \exists \lim_{k \rightarrow \infty} X_k =: X \geq \hat{X} \),
e) \( X \) is semi-stabilizing.

Main problems

- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Lyapunov Iterations/Perturbed Hessian Approach

Convergence

**Theorem [Cherfi/Abou-Kandil/Bourles 2005]**

If

- $\exists \hat{X}$ such that $\mathcal{R}(\hat{X}) \geq 0$,
- $\exists X_0 = X_0^T \geq \hat{X}$ such that $\mathcal{R}(X_0) \leq 0$ and $A - DX_0$ is Hurwitz,

then

a) $X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X}$,

b) $\mathcal{R}(X_k) \leq 0$ for all $k = 0, 1, \ldots$,

c) $A - DX_k$ is Hurwitz for all $k = 0, 1, \ldots$,

d) $\exists \lim_{k \to \infty} X_k =: \underline{X} \geq \hat{X}$,

e) $\underline{X}$ is semi-stabilizing.

**Main problems**

- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Theorem [Cherfi/Abou-Kandil/Bourles 2005]

If
- ∃ \( \hat{X} \) such that \( R(\hat{X}) \geq 0 \),
- ∃ \( X_0 = X_0^T \geq \hat{X} \) such that \( R(X_0) \leq 0 \) and \( A - DX_0 \) is Hurwitz,

then

a) \( X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X} \),
b) \( R(X_k) \leq 0 \) for all \( k = 0, 1, \ldots \),
c) \( A - DX_k \) is Hurwitz for all \( k = 0, 1, \ldots \),
d) \( \exists \lim_{k \to \infty} X_k =: \underline{X} \geq \hat{X} \),
e) \( \underline{X} \) is semi-stabilizing.

Main problems
- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Lyapunov Iterations/Perturbed Hessian Approach

Convergence

**Theorem** [Cherfi/Abou-Kandil/Bourles 2005]

If

- \( \exists \hat{X} \) such that \( R(\hat{X}) \geq 0 \),
- \( \exists X_0 = X_0^T \geq \hat{X} \) such that \( R(X_0) \leq 0 \) and \( A - DX_0 \) is Hurwitz,

then

\( a) \ X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X}, \)
\( b) \ R(X_k) \leq 0 \) for all \( k = 0, 1, \ldots, \)
\( c) \ A - DX_k \) is Hurwitz for all \( k = 0, 1, \ldots, \)
\( d) \ \exists \ \lim_{k \to \infty} X_k =: \hat{X} \)
\( e) \ \hat{X} \) is semi-stabilizing.

**Main problems**

- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Lyapunov Iterations/Perturbed Hessian Approach

**Theorem** [Cherfi/Abou-Kandil/Bourles 2005]

If

- $\exists \hat{X}$ such that $R(\hat{X}) \geq 0$,
- $\exists X_0 = X_0^T \geq \hat{X}$ such that $R(X_0) \leq 0$ and $A - DX_0$ is Hurwitz,

then

a) $X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X}$,

b) $R(X_k) \leq 0$ for all $k = 0, 1, \ldots$,

c) $A - DX_k$ is Hurwitz for all $k = 0, 1, \ldots$,

d) $\exists \lim_{k \to \infty} X_k =: X \geq \hat{X}$,

e) $X$ is semi-stabilizing.

**Main problems**

- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Lyapunov Iterations/Perturbed Hessian Approach

Convergence

**Theorem [Cherfi/Abou-Kandil/Bourles 2005]**

If

- $\exists \hat{X}$ such that $\mathcal{R}(\hat{X}) \geq 0$,
- $\exists X_0 = X_0^T \geq \hat{X}$ such that $\mathcal{R}(X_0) \leq 0$ and $A - DX_0$ is Hurwitz,

then

a) $X_0 \geq \ldots \geq X_k \geq X_{k+1} \geq \ldots \geq \hat{X}$,

b) $\mathcal{R}(X_k) \leq 0$ for all $k = 0, 1, \ldots$,

c) $A - DX_k$ is Hurwitz for all $k = 0, 1, \ldots$,

d) $\exists \lim_{k \to \infty} X_k =: X \geq \hat{X}$,

e) $X$ is semi-stabilizing.

**Main problems**

- Conditions for initial guess make its computation difficult.
- Observed convergence is linear.
Riccati Iterations

[LaNzon/Feng/B.D.O. Anderson 2007 (Proc. ECC 2007)]

Idea

Consider

\[ A^T X + XA + C^T C + X(B_1B_1^T - B_2B_2^T)X =: \mathcal{R}(X). \]

Then

\[ \mathcal{R}(X + Z) = \mathcal{R}(X) + (A + (B_1B_1^T - B_2B_2^T)X)^T Z + Z\hat{A} \]

\[ + Z(B_1B_1^T - B_2B_2^T)Z. \]

Furthermore, if \( X = X^T \), \( Z = Z^T \) solve the standard ARE

\[ 0 = \mathcal{R}(X) + \hat{A}^T Z + Z\hat{A} - ZB_2B_2^TZ, \]

then

\[ \mathcal{R}(X + Z) = ZB_1B_1^TZ, \]

\[ \|\mathcal{R}(X)\|_2 = \|B_1^TZ\|_2. \]
Riccati Iterations

\[\text{[Lanzon/Feng/B.D.O. Anderson 2007 (Proc. ECC 2007)]}\]

**Idea**

Consider

\[A^T X + XA + C^T C + X(B_1B_1^T - B_2B_2^T)X =: \mathcal{R}(X).\]

Then

\[
\mathcal{R}(X + Z) = \mathcal{R}(X) + \left( A + (B_1B_1^T - B_2B_2^T)X \right)^T Z + Z\hat{A} =: \hat{A} \\
+ Z(B_1B_1^T - B_2B_2^T)Z.
\]

Furthermore, if \( X = X^T, Z = Z^T \) solve the standard ARE

\[0 = \mathcal{R}(X) + \hat{A}^T Z + Z\hat{A} - ZB_2B_2^T Z,\]

then

\[
\mathcal{R}(X + Z) = ZB_1B_1^T Z, \\
\|\mathcal{R}(X)\|_2 = \|B_1^T Z\|_2.
\]
Idea

Consider

\[ A^T X + XA + C^T C + X(B_1 B_1^T - B_2 B_2^T)X =: \mathcal{R}(X). \]

Then

\[
\mathcal{R}(X + Z) = \mathcal{R}(X) + (A + (B_1 B_1^T - B_2 B_2^T)X)^T Z + Z \hat{A}
\]

\[ =: \hat{A} \]

\[ + Z(B_1 B_1^T - B_2 B_2^T)Z. \]

Furthermore, if \( X = X^T \), \( Z = Z^T \) solve the standard ARE

\[ 0 = \mathcal{R}(X) + \hat{A}^T Z + Z \hat{A} - ZB_2 B_2^T Z, \]

then

\[
\mathcal{R}(X + Z) = ZB_1 B_1^T Z,
\]

\[ \|\mathcal{R}(X)\|_2 = \|B_1^T Z\|_2. \]
Riccati Iterations

[Ranzon/Feng/B.D.O. Anderson 2007 (Proc. ECC 2007)]

Iterative Solution of AREs

Motivation:
$H_\infty$-Control
Large-Scale Standard AREs
AREs with Indefinite Hessian
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Conclusions and Open Problems

Riccati iteration

1. Set $X_0 = 0$.
2. FOR $k = 1, 2, \ldots$,
   
   (i) Set $A_k := A + B_1(B_1^T X_k) - B_2(B_2^T X_k)$.
   
   (ii) Solve the ARE
   \[ R(X_k) + A_k^T Z_k + Z_k A_k - Z_k B_2 B_2^T Z_k = 0. \]
   
   (iii) Set $X_{k+1} := X_k + Z_k$.
   
   (iv) IF $\|B_1^T Z_k\|_2 < \text{tol}$ THEN Stop.

Remark. ARE for $k = 0$ is the standard LQR/$H_2$ ARE.
Theorem [Lanzon/Feng/B.D.O. Anderson 2007]

If

- \((A, B_2)\) stabilizable,
- \((A, C)\) has no unobservable purely imaginary modes, and
- \(\exists\) stabilizing solution \(X_\sim\),

then

a) \((A + B_1 B_1^T X_k, B_2)\) stabilizable for all \(k = 0, 1, \ldots\),

b) \(Z_k \geq 0\) for all \(k = 0, 1, \ldots\),

c) \(A + B_1 B_1^T X_k - B_2 B_2^T X_{k+1}\) is Hurwitz for all \(k = 0, 1, \ldots\),

d) \(R(X_{k+1}) = Z_k B_1 B_1^T Z_k\) for all \(k = 0, 1, \ldots\),

e) \(X_\sim \geq \ldots \geq X_{k+1} \geq X_k \geq \ldots \geq 0\).

f) If \(\exists \lim_{k \to \infty} X_k =: X\), then \(X = X_\sim\), and

g) convergence is locally quadratic.
Riccati Iterations
[LANZON/FENG/B.D.O. ANDERSON 2007 (Proc. ECC 2007)]

Riccati iteration – low-rank version [B. 2008]

1. Solve the ARE

\[ C^T C + A^T Z_0 + Z_0 A - Z_0 B_2 B_2^T Z_0 = 0 \]

using Newton-ADI, yielding \( Y_0 \) with \( Z_0 \approx Y_0 Y_0^T \).

2. Set \( R_1 := Y_0 \). \( \{ \% R_1 R_1^T \approx X_1. \} \)

3. FOR \( k = 1, 2, \ldots \),

   (i) Set \( A_k := A + B_1 (B_1^T R_k) R_k^T - B_2 (B_2^T R_k) R_k^T \).

   (ii) Solve the ARE

   \[ Y_{k-1} (Y_{k-1}^T B_1) (B_1^T Y_{k-1}) Y_{k-1}^T + A_k^T Z_k + Z_k A_k - Z_k B_2 B_2^T Z_k = 0 \]

   using Newton-ADI, yielding \( Y_k \) with \( Z_k \approx Y_k Y_k^T \).

   (iii) Set \( R_{k+1} := \text{rrqr} ([R_k, Y_k], \tau). \) \( \{ \% R_{k+1} R_{k+1}^T \approx X_{k+1} \} \)

   (iv) IF \( \| (B_1^T Y_k) Y_k^T \|_2 < \text{tol} \) THEN Stop.
**AREs with Indefinite Hessian**

**Numerical example**

- Trivial example \((n = 2)\) from [Cherfi/Abou-Kandil/Bourles 2005].
- Compare convergence of Lyapunov and Riccati iterations.
- Solution of standard AREs with Newton’s method.
Iterative Solution of AREs

Motivation: $H_{\infty}$-Control

Large-Scale Standard AREs

AREs with Indefinite Hessian

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Conclusions and Open Problems

- Trivial example ($n = 2$) from [Cherfi/Abou-Kandil/Bourles 2005].
- Compare convergence of Lyapunov and Riccati iterations.
- Solution of standard AREs with Newton’s method.

![Absolute Residuals](chart)

- $\|R(X_k)\|_F$
- $k = \#$ of Lyapunov equations solved
Conclusions and Open Problems

- Low-rank Riccati iteration yields a reliable and efficient method for large-scale AREs with indefinite Hessian, useful, e.g., for $H_\infty$ optimization of PDE control problems.

- Low-rank Lyapunov iteration is an extremely simple variant for large-scale problems, but exhibits slower convergence and requires difficult-to-compute initial value.

- To-Do list:
  - Implement Riccati iteration in LyaPack/MESS style.
  - More numerical tests.
  - Re-write Riccati iteration as feedback iteration.
  - Efficient computation of initial value for Lyapunov iterations?
  - $\exists$ perturbed Hessian so that Lyapunov iteration quadratically convergent?
Conclusions and Open Problems

- Low-rank Riccati iteration yields a reliable and efficient method for large-scale AREs with indefinite Hessian, useful, e.g., for $H_\infty$ optimization of PDE control problems.

- Low-rank Lyapunov iteration is an extremely simple variant for large-scale problems, but exhibits slower convergence and requires difficult-to-compute initial value.

- To-Do list:
  - Implement Riccati iteration in LyaPack/MESS style.
  - More numerical tests.
  - Re-write Riccati iteration as feedback iteration.
  - Efficient computation of initial value for Lyapunov iterations?
  - ∃ perturbed Hessian so that Lyapunov iteration quadratically convergent?
Low-rank Riccati iteration yields a reliable and efficient method for large-scale AREs with indefinite Hessian, useful, e.g., for $H_\infty$ optimization of PDE control problems.

Low-rank Lyapunov iteration is an extremely simple variant for large-scale problems, but exhibits slower convergence and requires difficult-to-compute initial value.

To-Do list:
- Implement Riccati iteration in LyaPack/MESS style.
- More numerical tests.
- Re-write Riccati iteration as feedback iteration.
- Efficient computation of initial value for Lyapunov iterations?
- $\exists$ perturbed Hessian so that Lyapunov iteration quadratically convergent?

Fin.