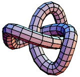
NUMERICAL SOLUTION OF LARGE-SCALE ALGEBRAIC RICCATI EQUATIONS

Peter Benner

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

TECHNISCHE UNIVERSITÄT CHEMNITZ



Université Paul Sabatier, Toulouse
April 28, 2008

Overview

1 Solving Large-Scale AREs
- Motivation
- Basic approaches
- Low-Rank Approximation
- Low-Rank Krylov Subspace Methods

2 Newton's Method for AREs
- ADI Method for Lyapunov Equations
- Low-Rank Newton-ADI for AREs
- Application to LQR Problem
- Numerical Results

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

4 Software

5 Conclusions and Open Problems

6 References

Solving Large-Scale AREs

Algebraic Riccati Equation (ARE)

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 ($A = -M^{-1}L$ 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 AREs

Algebraic Riccati Equation (ARE)

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 ($A = -M^{-1}L$ 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!

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods
Newton's Method for AREs
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

Algebraic Riccati Equation (ARE)

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 ($A = -M^{-1}L$ for FEM),
- G, W low-rank with $G, W \in \{BB^T, C^TC\}$, 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 AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

Conclusions and Open Problems

References

Solving Large-Scale AREs

Algebraic Riccati Equation (ARE)

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 ($A = -M^{-1}L$ 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 AREs

Algebraic Riccati Equation (ARE)

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 ($A = -M^{-1}L$ for FEM),
- G, W low-rank with $G, W \in \{BB^T, C^T C\}$, where $B \in \mathbb{R}^{n \times m}$, $m \ll n$, $\quad C \in \mathbb{R}^{p \times n}$, $p \ll n$.
- Standard (eigenproblem-based) $\mathcal{O}(n^3)$ methods are not applicable!

Large-scale AREs

Peter Benner

Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

Conclusions and Open Problems

References

Numerical solution of linear-quadratic optimal control problem for parabolic PDEs via Galerkin approach, spatial FEM discretization $\rightsquigarrow$

LQR Problem (finite-dimensional)

$$\text{Min } \mathcal{J}(u) = \frac{1}{2}\int_0^\infty \left(y^T Q y + u^T R u\right) dt \quad \text{for} \quad u \in \mathcal{L}_2(0,\infty;\mathbb{R}^m),$$

subject to $M\dot{x} = -Lx + Bu, \quad x(0) = x_0, \quad y = Cx,$
with stiffness $L \in \mathbb{R}^{n\times n}$, mass $M \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$.

Solution of finite-dimensional LQR problem: feedback control

$$u_*(t) = -B^T X_* x(t) =: -K_* x(t),$$

where $X_* = X_*^T \geq 0$ is the unique stabilizing[1] solution of the ARE

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

with $A := -M^{-1}L$, $B := M^{-1}BR^{-\frac{1}{2}}$, $C := CQ^{-\frac{1}{2}}$.

[1] X is stabilizing $\Leftrightarrow \Lambda(A - BB^T X) \subset \mathbb{C}^-$.

Large-scale AREs

Peter Benner

Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

Conclusions and Open Problems

References

Numerical solution of linear-quadratic optimal control problem for parabolic PDEs via Galerkin approach, spatial FEM discretization $\rightsquigarrow$

LQR Problem (finite-dimensional)

$$\text{Min } \mathcal{J}(u) = \frac{1}{2}\int_0^\infty \left(y^TQy + u^TRu\right)dt \quad \text{for} \quad u \in \mathcal{L}_2(0,\infty;\mathbb{R}^m),$$

subject to $\quad M\dot{x} = -Lx + Bu, \quad x(0) = x_0, \quad y = Cx,$
with stiffness $L \in \mathbb{R}^{n\times n}$, mass $M \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$.

Solution of finite-dimensional LQR problem: feedback control

$$u_*(t) = -B^TX_*x(t) =: -K_*x(t),$$

where $X_* = X_*^T \geq 0$ is the unique stabilizing[1] solution of the ARE

$$0 = \mathcal{R}(X) := C^TC + A^TX + XA - XBB^TX,$$

with $A := -M^{-1}L$, $B := M^{-1}BR^{-\frac{1}{2}}$, $C := CQ^{-\frac{1}{2}}$.

[1] X is stabilizing $\Leftrightarrow \Lambda(A - BB^TX) \subset \mathbb{C}^-$.

Treat ARE as nonlinear system of equations

Use Newton's method and its relatives, requires efficient solution of one Lyapunov equation

$$(A - GX_j)^T N_j + N_j(A - GX_j) = -\mathcal{R}(X_j)$$

per iteration. Using Hammarling's method or sign function iteration, this requires $\mathcal{O}(n^3)$ operations.

Sign function becomes feasible using hierarchical matrix structure and formatted arithmetic [Hackbusch/Khoromskij/Grasedyck 2003, Baur/B. 2006].

Large-scale AREs

Peter Benner

Solving Large-Scale AREs
- Motivation
- Basic approaches
- Low-Rank Approximation
- Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

Conclusions and Open Problems

References

Treat ARE as nonlinear system of equations

Solve corresponding Hamiltonian eigenproblem

Let

$$\underbrace{\begin{bmatrix} A & G \\ W & -A^T \end{bmatrix}}_{=:H} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} T,$$

so that colspan$\left(\begin{bmatrix} U \\ V \end{bmatrix}\right) = \mathcal{U} \in \mathbb{R}^{2n \times n}$ is the maximal stable H-invariant subspace (i.e., $\Lambda(T) \subset \mathbb{C}^-$, $\dim \mathcal{U} = n$), then (under mild assumptions),

$$X_* = -VU^{-1}$$

is the stabilizing solution of the ARE.

Can be computed using Schur decomposition of H and related techniques, but: complexity of all methods is $\mathcal{O}(n^3)$!

Several attempts to use partial Schur decomposition and Krylov subspace methods $\rightsquigarrow$ later.

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods
Newton's Method for AREs
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

Treat ARE as nonlinear system of equations

Solve corresponding Hamiltonian eigenproblem

Let

$$\underbrace{\begin{bmatrix} A & G \\ W & -A^T \end{bmatrix}}_{=:H} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} T,$$

so that colspan$\left(\begin{bmatrix} U \\ V \end{bmatrix}\right) = \mathcal{U} \in \mathbb{R}^{2n \times n}$ is the maximal stable H-invariant subspace (i.e., $\Lambda(T) \subset \mathbb{C}^-$, $\dim \mathcal{U} = n$), then (under mild assumptions),

$$X_* = -VU^{-1}$$

is the stabilizing solution of the ARE.

Can be computed using Schur decomposition of H and related techniques, but: complexity of all methods is $\mathcal{O}(n^3)$!

Several attempts to use partial Schur decomposition and Krylov subspace methods $\rightsquigarrow$ later.

Large-scale AREs

Peter Benner

Solving Large-Scale AREs
- Motivation
- Basic approaches
- Low-Rank Approximation
- Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

Conclusions and Open Problems

References

Treat ARE as nonlinear system of equations

Solve corresponding Hamiltonian eigenproblem

Let

$$\underbrace{\begin{bmatrix} A & G \\ W & -A^T \end{bmatrix}}_{=:H} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} T,$$

so that colspan$\left(\left[\begin{smallmatrix} U \\ V \end{smallmatrix}\right]\right) = \mathcal{U} \in \mathbb{R}^{2n\times n}$ is the maximal stable H-invariant subspace (i.e., $\Lambda(T) \subset \mathbb{C}^-$, $\dim \mathcal{U} = n$), then (under mild assumptions),

$$X_* = -VU^{-1}$$

is the stabilizing solution of the ARE.

Can be computed using Schur decomposition of H and related techniques, but: complexity of all methods is $\mathcal{O}(n^3)$!

Several attempts to use partial Schur decomposition and Krylov subspace methods $\rightsquigarrow$ later.

Low-Rank Approximation

ARE $0 = A^T X + XA - XBB^T X + CC^T$

Large-scale AREs

Peter Benner

Solving Large-Scale AREs
- Motivation
- Basic approaches
- Low-Rank Approximation
- Low-Rank Krylov Subspace Methods

Newton's Method for AREs

AREs with Indefinite Hessian

Software

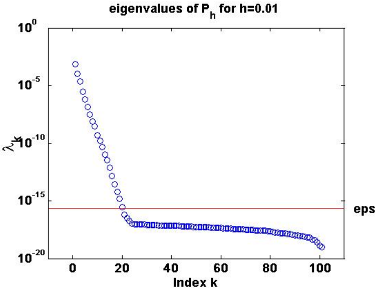
Conclusions and Open Problems

References

Consider spectrum of ARE solution:

Example:

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



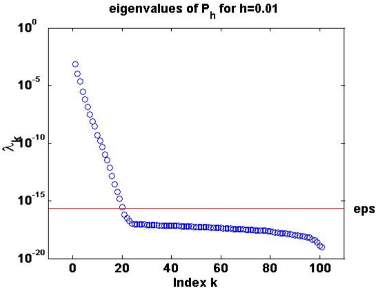
Idea: $X = X^T \geq 0 \implies$

$$X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx \sum_{k=1}^{r} \lambda_k y_k y_k^T =: Y^{(r)}(Y^{(r)})^T.$$

Consider spectrum of ARE solution:

Example:

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



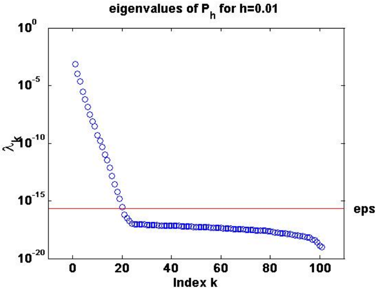
Idea: $X = X^T \geq 0 \implies$

$$X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx \sum_{k=1}^{r} \lambda_k y_k y_k^T =: Y^{(r)}(Y^{(r)})^T.$$

Consider spectrum of ARE solution:

Example:

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



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

$$X = YY^T = \sum_{k=1}^{n} \lambda_k y_k y_k^T \approx \sum_{k=1}^{r} \lambda_k y_k y_k^T =: Y^{(r)}(Y^{(r)})^T.$$

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods
Newton's Method for AREs
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX.$

1. Apply (block-)Arnoldi process to A with start (block-)vector C to generate the Krylov space

$$\mathcal{K}_\ell(A, C) = \operatorname{span}\{C, AC, A^2C, \dots, A^{\ell-1}C\}$$

with orthogonal basis V_ℓ such that

$$AV_\ell = V_\ell A_\ell + W_{\ell+1}A_{\ell+1,\ell}\begin{bmatrix} 0 & I_p \end{bmatrix}$$

and $A_\ell = V_\ell^T A V_\ell$ is block upper-Hessenberg.

2. Set $B_\ell := V_\ell^T B$, $C_\ell := V_\ell^T C$.
3. Find stabilizing solution of the ARE

$$0 = \mathcal{R}_\ell(X_\ell) = C_\ell C_\ell^T + A_\ell X_\ell + X_\ell A_\ell^T - X_\ell B_\ell B_\ell^T X_\ell.$$

4. Set $\tilde{X} := V_\ell X_\ell V_\ell^T$.

Low-Rank Krylov Subspace Methods

Block-Arnoldi method

[Jaimoukha/Kasenally '94]

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX$.

1. Apply (block-)Arnoldi process to A with start (block-)vector C to generate the Krylov space

$$\mathcal{K}_\ell(A, C) = \mathrm{span}\{C, AC, A^2C, \dots, A^{\ell-1}C\}$$

with orthogonal basis V_ℓ such that

$$AV_\ell = V_\ell A_\ell + W_{\ell+1}A_{\ell+1,\ell}\left[\begin{array}{cc} 0 & I_p \end{array}\right]$$

and $A_\ell = V_\ell^T A V_\ell$ is block upper-Hessenberg.

2. Set $B_\ell := V_\ell^T B$, $C_\ell := V_\ell^T C$.
3. Find stabilizing solution of the ARE

$$0 = \mathcal{R}_\ell(X_\ell) = C_\ell C_\ell^T + A_\ell X_\ell + X_\ell A_\ell^T - X_\ell B_\ell B_\ell^T X_\ell.$$

4. Set $\tilde{X} := V_\ell X_\ell V_\ell^T$.

Low-Rank Krylov Subspace Methods
Block-Arnoldi method

[JAIMOUKHA/KASENALLY '94]

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^T X.$

1. Apply (block-)Arnoldi process to A with start (block-)vector C to generate the Krylov space

$$\mathcal{K}_\ell(A, C) = \operatorname{span}\{C, AC, A^2C, \ldots, A^{\ell-1}C\}$$

with orthogonal basis V_ℓ such that

$$AV_\ell = V_\ell A_\ell + W_{\ell+1} A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix}$$

and $A_\ell = V_\ell^T A V_\ell$ is block upper-Hessenberg.

2. Set $B_\ell := V_\ell^T B$, $C_\ell := V_\ell^T C$.
3. Find stabilizing solution of the ARE

$$0 = \mathcal{R}_\ell(X_\ell) = C_\ell C_\ell^T + A_\ell X_\ell + X_\ell A_\ell^T - X_\ell B_\ell B_\ell^T X_\ell.$$

4. Set $\tilde{X} := V_\ell X_\ell V_\ell^T$.

Low-Rank Krylov Subspace Methods

Block-Arnoldi method [Jaimoukha/Kasenally '94]

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX.$

Properties:

+ $\tilde{X}$ satisfies Galerkin-type condition $V_\ell^T \mathcal{R}(\tilde{X}) V_\ell = 0.$
+ Computable residual error norm

$$\|\mathcal{R}(\tilde{X})\|_F = \sqrt{2} \cdot \|A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix} X_\ell\|_F.$$

– Block-Arnoldi, i.e., each step needs p matrix-vector products.
– Stabilizing X_ℓ may not exist as corresponding Hamiltonian matrix

$$H_\ell := \begin{bmatrix} A_\ell^T & B_\ell B_\ell^T \\ C_\ell C_\ell^T & -A_\ell \end{bmatrix}$$

may have purely imaginary eigenvalues!
– No stabilization guarantee for $\tilde{X}$!
– No convergence results for residuals.

Low-Rank Krylov Subspace Methods

Block-Arnoldi method

[Jaimoukha/Kasenally '94]

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX.$

Properties:

+ $\tilde{X}$ satisfies Galerkin-type condition $V_\ell^T \mathcal{R}(\tilde{X}) V_\ell = 0.$
+ Computable residual error norm

$$\|\mathcal{R}(\tilde{X})\|_F = \sqrt{2} \cdot \|A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix} X_\ell\|_F.$$

– Block-Arnoldi, i.e., each step needs p matrix-vector products.

– Stabilizing X_ℓ may not exist as corresponding Hamiltonian matrix

$$H_\ell := \begin{bmatrix} A_\ell^T & B_\ell B_\ell^T \\ C_\ell C_\ell^T & -A_\ell \end{bmatrix}$$

may have purely imaginary eigenvalues!

– No stabilization guarantee for $\tilde{X}$!

– No convergence results for residuals.

Low-Rank Krylov Subspace Methods

Block-Arnoldi method

[Jaimoukha/Kasenally '94]

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX$.

Properties:

+ $\tilde{X}$ satisfies Galerkin-type condition $V_\ell^T \mathcal{R}(\tilde{X}) V_\ell = 0$.
+ Computable residual error norm

$$\|\mathcal{R}(\tilde{X})\|_F = \sqrt{2} \cdot \|A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix} X_\ell\|_F.$$

– Block-Arnoldi, i.e., each step needs p matrix-vector products.

– Stabilizing X_ℓ may not exist as corresponding Hamiltonian matrix

$$H_\ell := \begin{bmatrix} A_\ell^T & B_\ell B_\ell^T \\ C_\ell C_\ell^T & -A_\ell \end{bmatrix}$$

may have purely imaginary eigenvalues!

– No stabilization guarantee for $\tilde{X}$!

– No convergence results for residuals.

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^T X.$

Properties:

+ $\tilde{X}$ satisfies Galerkin-type condition $V_\ell^T \mathcal{R}(\tilde{X}) V_\ell = 0.$
+ Computable residual error norm

$$\|\mathcal{R}(\tilde{X})\|_F = \sqrt{2} \cdot \|A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix} X_\ell\|_F.$$

– Block-Arnoldi, i.e., each step needs p matrix-vector products.
– Stabilizing X_ℓ may not exist as corresponding Hamiltonian matrix

$$H_\ell := \begin{bmatrix} A_\ell^T & B_\ell B_\ell^T \\ C_\ell C_\ell^T & -A_\ell \end{bmatrix}$$

may have purely imaginary eigenvalues!
– No stabilization guarantee for $\tilde{X}$!
– No convergence results for residuals.

Consider $0 = \mathcal{R}(X) = CC^T + AX + XA^T - XBB^TX.$

Properties:

+ $\tilde{X}$ satisfies Galerkin-type condition $V_\ell^T \mathcal{R}(\tilde{X}) V_\ell = 0.$
+ Computable residual error norm

$$\|\mathcal{R}(\tilde{X})\|_F = \sqrt{2} \cdot \|A_{\ell+1,\ell} \begin{bmatrix} 0 & I_p \end{bmatrix} X_\ell\|_F.$$

– Block-Arnoldi, i.e., each step needs p matrix-vector products.
– Stabilizing X_ℓ may not exist as corresponding Hamiltonian matrix

$$H_\ell := \begin{bmatrix} A_\ell^T & B_\ell B_\ell^T \\ C_\ell C_\ell^T & -A_\ell \end{bmatrix}$$

may have purely imaginary eigenvalues!
– No stabilization guarantee for $\tilde{X}$!
– No convergence results for residuals.

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods
Newton's Method for AREs
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

Consider

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

1. Apply symplectic Lanczos method to Hamiltonian matrix $\begin{bmatrix} A & BB^T \\ C^T C & -A^T \end{bmatrix}$ to generate Krylov subspace

$$\mathcal{K}_{2\ell}(H, v_1) = \text{span}\{v_1, Hv_1, H^2 v_1, \ldots, H^{2\ell-1} v_1\},$$

with symplectic basis $S_\ell = [v_1, w_1, \ldots, v_\ell, w_\ell] \in \mathbb{R}^{2n,2\ell}$, i.e.,

$$S_\ell^T \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} S_\ell = \begin{bmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{bmatrix}$$

such that

$$HS_\ell = S_\ell H_\ell + \zeta_{\ell+1} v_{\ell+1} e_{2\ell}^T.$$

2. Compute ARE solution corresponding to H_ℓ, prolongate to $\mathbb{R}^{n \times n}$ to obtain approximate rank-ℓ solution $\tilde{X}$.

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Motivation
Basic approaches
Low-Rank Approximation
Low-Rank Krylov Subspace Methods
Newton's Method for AREs
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

Consider

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

1. Apply symplectic Lanczos method to Hamiltonian matrix $\begin{bmatrix} A & BB^T \\ C^T C & -A^T \end{bmatrix}$ to generate Krylov subspace

$$\mathcal{K}_{2\ell}(H, v_1) = \operatorname{span}\{v_1, Hv_1, H^2 v_1, \ldots, H^{2\ell-1} v_1\},$$

with symplectic basis $S_\ell = [v_1, w_1, \ldots, v_\ell, w_\ell] \in \mathbb{R}^{2n,2\ell}$, i.e.,

$$S_\ell^T \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} S_\ell = \begin{bmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{bmatrix}$$

such that

$$HS_\ell \quad = \quad S_\ell H_\ell + \zeta_{\ell+1} v_{\ell+1} e_{2\ell}^T.$$

2. Compute ARE solution corresponding to H_ℓ, prolongate to $\mathbb{R}^{n\times n}$ to obtain approximate rank-ℓ solution $\tilde{X}$.

Properties

+ $\tilde{X}$ satisfies Galerkin-type condition $\quad Z_\ell^T \mathcal{R}(\tilde{X}) Z_\ell = 0.$
+ In general, less (but twice as expensive) matrix-vector products than for block-Arnoldi.
+ Purely imaginary eigenvalues of small Hamiltonian matrix H_ℓ can be removed by implicit restarts, i.e., can always get stable H_ℓ-invariant subspace.
+ Stabilization property for projected feedback matrix

$$Z_\ell^T (A - BB^T \tilde{X}) Z_\ell.$$

for sufficiently small Lanczos residual (can be achieved by implicit restarts).

− $\tilde{X} \neq \tilde{X}^T$ for $\ell < n$.
− No stabilization guarantee for $\tilde{X}$.
− No convergence results for residuals.

Properties

+ $\tilde{X}$ satisfies Galerkin-type condition $Z_\ell^T \mathcal{R}(\tilde{X}) Z_\ell = 0.$
+ In general, less (but twice as expensive) matrix-vector products than for block-Arnoldi.
+ Purely imaginary eigenvalues of small Hamiltonian matrix H_ℓ can be removed by implicit restarts, i.e., can always get stable H_ℓ-invariant subspace.
+ Stabilization property for projected feedback matrix

$$Z_\ell^T (A - BB^T \tilde{X}) Z_\ell.$$

for sufficiently small Lanczos residual (can be achieved by implicit restarts).

– $\tilde{X} \neq \tilde{X}^T$ for $\ell < n$.
– No stabilization guarantee for $\tilde{X}$.
– No convergence results for residuals.

Properties

+ $\tilde{X}$ satisfies Galerkin-type condition $\quad Z_\ell^T \mathcal{R}(\tilde{X}) Z_\ell = 0.$
+ In general, less (but twice as expensive) matrix-vector products than for block-Arnoldi.
+ Purely imaginary eigenvalues of small Hamiltonian matrix H_ℓ can be removed by implicit restarts, i.e., can always get stable H_ℓ-invariant subspace.
+ Stabilization property for projected feedback matrix

$$Z_\ell^T (A - BB^T \tilde{X}) Z_\ell.$$

  for sufficiently small Lanczos residual (can be achieved by implicit restarts).

– $\tilde{X} \neq \tilde{X}^T$ for $\ell < n$.
– No stabilization guarantee for $\tilde{X}$.
– No convergence results for residuals.

Properties

+ $\tilde{X}$ satisfies Galerkin-type condition $\quad Z_\ell^T \mathcal{R}(\tilde{X}) Z_\ell = 0.$
+ In general, less (but twice as expensive) matrix-vector products than for block-Arnoldi.
+ Purely imaginary eigenvalues of small Hamiltonian matrix H_ℓ can be removed by implicit restarts, i.e., can always get stable H_ℓ-invariant subspace.
+ Stabilization property for projected feedback matrix

$$Z_\ell^T (A - BB^T \tilde{X}) Z_\ell.$$

for sufficiently small Lanczos residual (can be achieved by implicit restarts).

– $\tilde{X} \neq \tilde{X}^T$ for $\ell < n$.
– No stabilization guarantee for $\tilde{X}$.
– No convergence results for residuals.

Related work

- Two-sided structured Arnoldi method for Hamiltonian matrices based on symplectic URV decomposition/product eigenvalue problem.
  [B./Mehrmann/Kressner/Xu in progress]

- Improved generation of low-rank approximate solution from r-dimensional ($r \ll n$) H-invariant subspace $\mathrm{range}(W)$ via

$$
\begin{aligned}
&W_1 := W(1:n,:), \qquad W_2 = W(n+1:2n,:)\\
&[U,\Sigma,V] = \mathrm{svd}(W_1^T W_2),\\
&\tilde{X} = W_2 U S^{-\frac{1}{2}} \quad (= \tilde{X}^T).
\end{aligned}
$$

  [Amodei '03, B./Mehrmann/Sorensen '03]

- Variants of Arnoldi method using Frobenius inner product $\rightsquigarrow$ global Arnoldi. [Jbilou '03–'06]

Related work

- Two-sided structured Arnoldi method for Hamiltonian matrices based on symplectic URV decomposition/product eigenvalue problem.

  [B./Mehrmann/Kressner/Xu in progress]

- Improved generation of low-rank approximate solution from r-dimensional ($r \ll n$) H-invariant subspace $\mathrm{range}(W)$ via

$$W_1 := W(1:n,:), \qquad W_2 = W(n+1:2n,:)$$
$$[U, \Sigma, V] = \mathrm{svd}(W_1^T W_2),$$
$$\tilde{X} = W_2 U S^{-\frac{1}{2}} \quad (= \tilde{X}^T).$$

  [Amodei '03, B./Mehrmann/Sorensen '03]

- Variants of Arnoldi method using Frobenius inner product ⇝ global Arnoldi. [Jbilou '03–'06]

Related work

- Two-sided structured Arnoldi method for Hamiltonian matrices based on symplectic URV decomposition/product eigenvalue problem.

  [B./Mehrmann/Kressner/Xu in progress]

- Improved generation of low-rank approximate solution from r-dimensional ($r \ll n$) H-invariant subspace $\mathrm{range}(W)$ via

$$W_1 := W(1:n,:), \qquad W_2 = W(n+1:2n,:)$$
$$[U, \Sigma, V] = \mathrm{svd}(W_1^T W_2),$$
$$\tilde{X} = W_2 U S^{-\frac{1}{2}} \quad (= \tilde{X}^T).$$

  [Amodei '03, B./Mehrmann/Sorensen '03]

- Variants of Arnoldi method using Frobenius inner product $\rightsquigarrow$ global Arnoldi. [Jbilou '03–'06]

Newton's Method for AREs

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

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

$$\mathcal{R}'_X : Z \to (A - BB^T X)^T Z + Z(A - BB^T X).$$

- Newton-Kantorovich method:

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

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

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

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

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

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 $\quad A_j^T N_j + N_j A_j = -\mathcal{R}(X_j)$.
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 $\quad 0 = \mathcal{R}(X) = C^T C + A^T X + XA - XBB^T X.$
- Frechét derivative of $\mathcal{R}(X)$ at X:

$$\mathcal{R}'_X : Z \to (A - BB^T X)^T Z + Z(A - BB^T X).$$

- Newton-Kantorovich method:

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

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 $\quad A_j^T N_j + N_j A_j = -\mathcal{R}(X_j).$
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 $0 = \mathcal{R}(X) = C^T C + A^T X + XA - XBB^T X$.
- Frechét derivative of $\mathcal{R}(X)$ at X:

$$\mathcal{R}'_X : Z \to (A - BB^T X)^T Z + Z(A - BB^T X).$$

- Newton-Kantorovich method:

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

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

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

$$A_j \;=\; A \;-\; B \;\cdot\; K_j$$

$$=\; \text{sparse} \;-\; m \;\cdot\; \square$$

- $m \ll n \implies$ 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!

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

$$A_j \;=\; A \;-\; B \;\cdot\; K_j$$

$=$ sparse $-$ m $\cdot$ [ ]

- $m \ll n \implies$ 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!

Low-Rank Krylov Subspace Methods

Properties and Implementation

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

  $A_j \quad = \quad A \quad - \quad B \quad \cdot \quad K_j$

  $\quad\;\; = \quad$ sparse $\quad - \quad m \quad \cdot \quad$ [ ]

- $m \ll n \Longrightarrow$ 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: $X = X^T \in \mathbb{R}^{n\times n} \Longrightarrow n(n+1)/2$ unknowns!

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

$$A_j \quad = \quad A \quad - \quad B \quad \cdot \quad K_j$$

$$= \quad \text{sparse} \quad - \quad m \quad \cdot$$

- $m \ll n \implies$ 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!

ADI Method 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: [Wachspress 1988]
$$\begin{aligned}(A + p_k I)X_{(k-1)/2} &= -BB^T - X_{k-1}(A^T - p_k I)\\ (A + \overline{p_k} I){X_k}^T &= -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)\end{aligned}$$
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...

ADI Method 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: [Wachspress 1988]
$$\begin{aligned}(A + p_k I)X_{(k-1)/2} &= -BB^T - X_{k-1}(A^T - p_k I)\\ (A + \overline{p_k} I){X_k}^T &= -BB^T - X_{(k-1)/2}(A^T - \overline{p_k} I)\end{aligned}$$
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...

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
ADI for Lyapunov
Low-Rank Newton-ADI
Application to LQR Problem
Numerical Results
AREs with Indefinite Hessian
Software
Conclusions and Open Problems
References

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

ADI Method 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: [Wachspress 1988]

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

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 $\Longrightarrow$

Algorithm [Penzl 1997, Li/White 2002, B./Li/Penzl 1999/2008]

$$V_1 \leftarrow \sqrt{-2\mathrm{Re}\,(p_1)}(A + p_1 I)^{-1} B, \qquad Y_1 \leftarrow V_1$$

FOR $j = 2, 3, \ldots$

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

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

$$Y_k \leftarrow \mathrm{rrqr}(Y_k, \tau) \qquad \text{\% column compression}$$

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

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

Note: Implementation in real arithmetic possible by combining two steps.

Factored ADI Iteration

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

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

Algorithm [Penzl 1997, Li/White 2002, B./Li/Penzl 1999/2008]

$$V_1 \leftarrow \sqrt{-2\mathrm{Re}\,(p_1)}(A + p_1 I)^{-1}B, \qquad Y_1 \leftarrow V_1$$

FOR $j = 2, 3, \dots$

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

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

$$Y_k \leftarrow \mathrm{rrqr}(Y_k, \tau) \qquad \text{\% column compression}$$

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

$$\mathrm{range}\,(Y_{k_{\max}}) = \mathrm{range}\left(\begin{bmatrix} V_1 & \dots & V_{k_{\max}} \end{bmatrix}\right), \qquad V_k = \Big[\;\Big] \in \mathbb{C}^{n \times m}.$$

Note: Implementation in real arithmetic possible by combining two steps.

Low-Rank Newton-ADI for AREs

Re-write Newton's method for AREs

$$A_j^T N_j + N_j A_j = -\mathcal{R}(X_j)$$

$$\Longleftrightarrow$$

$$A_j^T \underbrace{(X_j + N_j)}_{=X_{j+1}} + \underbrace{(X_j + N_j)}_{=X_{j+1}} A_j = \underbrace{-C^T C - X_j B B^T X_j}_{=:-W_j W_j^T}$$

Set $X_j = Z_j Z_j^T$ for $\mathrm{rank}\,(Z_j) \ll n \Longrightarrow$

$$A_j^T \left(Z_{j+1} Z_{j+1}^T\right) + \left(Z_{j+1} Z_{j+1}^T\right) A_j = -W_j W_j^T$$

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.

Low-Rank Newton-ADI for AREs

Re-write Newton's method for AREs

$$A_j^T N_j + N_j A_j = -\mathcal{R}(X_j)$$

$$\Longleftrightarrow$$

$$A_j^T \underbrace{(X_j + N_j)}_{=X_{j+1}} + \underbrace{(X_j + N_j)}_{=X_{j+1}} A_j = \underbrace{-C^T C - X_j B B^T X_j}_{=:-W_j W_j^T}$$

Set $X_j = Z_j Z_j^T$ for $\operatorname{rank}(Z_j) \ll n \Longrightarrow$

$$A_j^T \left(Z_{j+1} Z_{j+1}^T\right) + \left(Z_{j+1} Z_{j+1}^T\right) A_j = -W_j W_j^T$$

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.

Optimal feedback

$$K_* = B^T X_* = B^T Z_* Z_*^T$$

can be computed by direct feedback iteration:

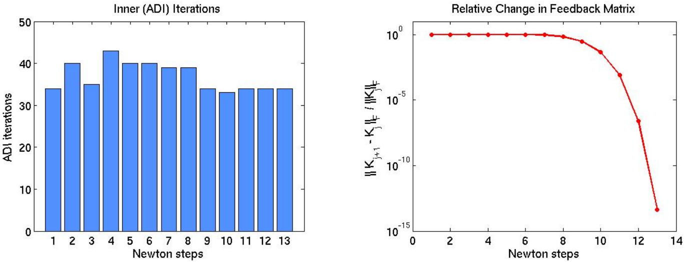
- jth Newton iteration:

$$K_j = B^T Z_j Z_j^T = \sum_{k=1}^{k_{\max}} (B^T V_{j,k}) V_{j,k}^T \quad \xrightarrow{j\to\infty} \quad K_* = B^T Z_* Z_*^T$$

- K_j can be updated in ADI iteration, no need to even form Z_j, need only fixed workspace for $K_j \in \mathbb{R}^{m\times n}$!

Related to earlier work by [Banks/Ito 1991].

- Linear 2D heat equation with homogeneous Dirichlet boundary and point control/observation.
- FD discretization on uniform 150×150 grid.
- $n = 22.500$, $m = p = 1$, 10 shifts for ADI iterations.
- Convergence of large-scale matrix equation solvers:



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

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

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.

Confirms mesh independence principle for Newton-Kleinman [Burns/Sachs/Zietsmann 2006].

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

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

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.

Confirms mesh independence principle for Newton-Kleinman [Burns/Sachs/Zietsmann 2006].

AREs with Indefinite Hessian

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

Now:

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

AREs with Indefinite Hessian

Now:

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

Problems

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

AREs with Indefinite Hessian

Now:

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

Problems

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

AREs with Indefinite Hessian

Now:

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

Problems

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

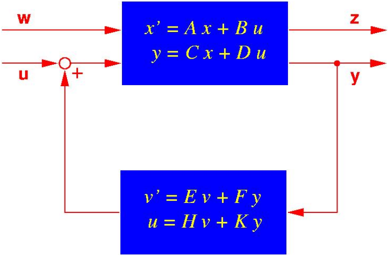
Motivation: H_∞-Control

Linear time-invariant systems

$$\Sigma : \begin{cases} \dot{x} = Ax + B_1 w + B_2 u, \\ z = C_1 x + D_{11} w + D_{12} u, \\ y = C_2 x + D_{21} w + D_{22} u, \end{cases}$$

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



Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

Laplace transform $\Longrightarrow$ transfer function (in frequency domain)

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \equiv \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right].$$

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 Σ via

$$\begin{aligned} z(s) &= G_{11}(s)w(s) + G_{12}(s)u(s), \\ y(s) &= G_{21}(s)w(s) + G_{22}(s)u(s). \end{aligned}$$

Large-scale AREs

Peter Benner

Solving Large-Scale AREs

Newton's Method for AREs

AREs with Indefinite Hessian

H_∞-Control

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

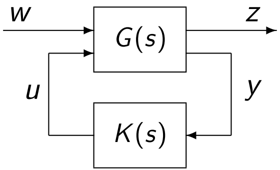
Numerical example

Software

Conclusions and Open Problems

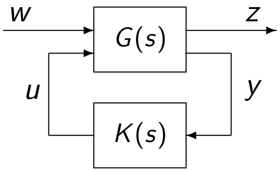
References

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



Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

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



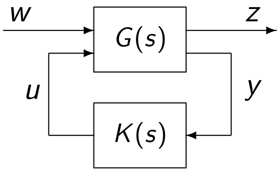
Goal:

find K that minimize error outputs

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

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

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

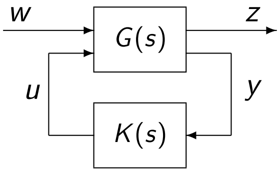
$$z = \left(G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21}\right) w =: \mathcal{F}(G, K)w,$$

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

H_∞-optimal control problem:

$$\min_{K \text{ stabilizing}} \|\mathcal{F}(G, K)\|_{\mathcal{H}_\infty}.$$

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

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

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

H_∞-suboptimal control problem:

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

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

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 ($\Longrightarrow \Sigma$ 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.

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

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

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

$$C_1C_1^T + A^TX + XA + X(\gamma^{-2}B_1B_1^T - B_2B_2^T)X = 0, \qquad (1)$$

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

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

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_∞-optimal control

Find minimal γ for which (i)–(iii) are satisfied $\rightsquigarrow$ γ-iteration based on solving AREs (1)–(2) repeatedly for different γ.

Large-scale AREs

Peter Benner

Solving Large-Scale AREs

Newton's Method for AREs

AREs with Indefinite Hessian

H_∞-Control

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Numerical example

Software

Conclusions and Open Problems

References

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

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

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

$$C_1C_1^T + A^TX + XA + X(\gamma^{-2}B_1B_1^T - B_2B_2^T)X = 0, \tag{1}$$

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

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

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_∞-optimal control

Find minimal γ for which (i)–(iii) are satisfied $\rightsquigarrow$ γ-iteration based on solving AREs (1)–(2) repeatedly for different γ.

H_∞-(sub-)optimal controller

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

$$\hat{K}(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & 0 \end{array}\right] = \hat{C}(sI_n - \hat{A})^{-1}\hat{B},$$

where for

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

$$\begin{aligned} \hat{A} &:= A + (\gamma^{-2} B_1 B_1^T - B_2 B_2^T) X_\infty - Z_\infty Y_\infty C_2^T C_2, \\ \hat{B} &:= Z_\infty Y_\infty C_2^T, \\ \hat{C} &:= -B_2^T X_\infty. \end{aligned}$$

$\hat{K}(s)$ is the central or minimum entropy controller.

Numerical Solution of AREs with Indefinite Hessian

A quick-and-dirty solution [DAMM 2002/04]

ARE with indefinite Hessian

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

Consider $X^{-1}\mathcal{R}(X)X^{-1} = 0$

$\rightsquigarrow$ 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!

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 $\alpha \geq \min\{0, -\lambda_{\max}(G)\}$.

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: $G = B_2 B_2^T - B_1 B_1^T$

$\Rightarrow$ use $\alpha = \|B_1\|^2$ for spectral/Frobenius norm or

$$\alpha = \|B_1\|_1 \cdot \|B_1\|_\infty.$$

Remark

$W \geq -G$ can be used instead of αI, e.g., $W = \beta B_1 B_1^T$ with $\beta \geq 1$.

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

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 $\alpha \geq \min\{0, -\lambda_{\max}(G)\}$.

Lyapunov iteration

Based on

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

iterate

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

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

Lyapunov Iterations/Perturbed Hessian Approach

[Cherfi/Abou-Kandil/Bourles 2005 (Proc. ACSE 2005)]

Lyapunov iteration

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

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

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 αI: the Lyapunov equation becomes

$$\begin{aligned}(A - B_2 B_2^T Y_k Y_k)^T Y_{k+1} Y_{k+1}^T + Y_{k+1} Y_{k+1}^T (A - B_2 B_2^T Y_k Y_k) \\ = \quad -CC^T - Y_k Y_k^T B_1 B_1^T Y_k Y_k^T - Y_k Y_k^T B_2 B_2^T Y_k Y_k^T \\ = \quad -[\, C,\ Y_k Y_k^T B_1,\ Y_k Y_k^T B_2 \,] \begin{bmatrix} C^T \\ B_1^T Y_k Y_k^T \\ B_2^T Y_k Y_k^T \end{bmatrix}.\end{aligned}$$

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

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

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

Main problems

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

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

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

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

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

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

Main problems

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

Large-scale AREs
Peter Benner
Solving Large-Scale AREs
Newton's Method for AREs
AREs with Indefinite Hessian
H_∞-Control
Lyapunov Iterations/Perturbed Hessian Approach
Riccati Iterations
Numerical example
Software
Conclusions and Open Problems
References

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

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

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

Main problems

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

Large-scale AREs

Peter Benner

Solving Large-Scale AREs

Newton's Method for AREs

AREs with Indefinite Hessian

H_∞-Control

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Numerical example

Software

Conclusions and Open Problems

References

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

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

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

Main problems

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

Riccati Iterations

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

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

$$\mathcal{R}(X+Z) = \mathcal{R}(X) + (\underbrace{A + (B_1 B_1^T - B_2 B_2^T)X}_{=:\widehat{A}})^T Z + Z\widehat{A} + Z(B_1 B_1^T - B_2 B_2^T)Z.$$

Furthermore, if $X = X^T$, $Z = Z^T$ solve the standard ARE

$$0 = \mathcal{R}(X) + \widehat{A}^T Z + Z\widehat{A} - ZB_2 B_2^T Z,$$

then

$$\begin{aligned} \mathcal{R}(X+Z) &= ZB_1 B_1^T Z, \\ \|\mathcal{R}(X)\|_2 &= \|B_1^T Z\|_2. \end{aligned}$$

Large-scale AREs

Peter Benner

Solving Large-Scale AREs

Newton's Method for AREs

AREs with Indefinite Hessian

H_∞-Control

Lyapunov Iterations/Perturbed Hessian Approach

Riccati Iterations

Numerical example

Software

Conclusions and Open Problems

References

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

$$\mathcal{R}(X+Z) = \mathcal{R}(X) + (\underbrace{A + (B_1 B_1^T - B_2 B_2^T)X}_{=:\widehat{A}})^T Z + Z\widehat{A} + Z(B_1 B_1^T - B_2 B_2^T)Z.$$

Furthermore, if $X = X^T$, $Z = Z^T$ solve the standard ARE

$$0 = \mathcal{R}(X) + \widehat{A}^T Z + Z\widehat{A} - ZB_2 B_2^T Z,$$

then

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

$$\|\mathcal{R}(X)\|_2 = \|B_1^T Z\|_2.$$

Riccati Iterations

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

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

$$\begin{aligned}\mathcal{R}(X+Z) &= \mathcal{R}(X) + (\underbrace{A + (B_1 B_1^T - B_2 B_2^T)X}_{=:\widehat{A}})^T Z + Z\widehat{A} \\ &\quad + Z(B_1 B_1^T - B_2 B_2^T)Z.\end{aligned}$$

Furthermore, if $X = X^T$, $Z = Z^T$ solve the standard ARE

$$0 = \mathcal{R}(X) + \widehat{A}^T Z + Z\widehat{A} - ZB_2 B_2^T Z,$$

then

$$\begin{aligned}\mathcal{R}(X+Z) &= ZB_1 B_1^T Z, \\ \|\mathcal{R}(X)\|_2 &= \|B_1^T Z\|_2.\end{aligned}$$

Riccati Iterations

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

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
   $$\mathcal{R}(X_k) + A_k^T Z_k + Z_k A_k - Z_k B_2 B_2^T Z_k = 0.$$
   - (iii) Set $X_{k+1} := X_k + Z_k$.
   - (iv) IF $\|B_1^T Z_k\|_2 <$ tol THEN Stop.

Remark. ARE for $k = 0$ is the standard LQR/H_2 ARE.

Riccati Iterations

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

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_-,

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

e) $X_- \geq \ldots \geq X_{k+1} \geq X_k \geq \ldots \geq 0.$

f) If $\exists \lim_{k\to\infty} X_k =: \underline{X}$, then $\underline{X} = X_-$, 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
$$C^TC + A^TZ_0 + Z_0A - Z_0B_2B_2^TZ_0 = 0$$
using Newton-ADI, yielding Y_0 with $Z_0 \approx Y_0Y_0^T$.
2. Set $R_1 := Y_0$. $\{\%\ R_1R_1^T \approx X_1.\}$
3. FOR $k = 1, 2, \ldots,$
   - (i) Set $A_k := A + B_1(B_1^TR_k)R_k^T - B_2(B_2^TR_k)R_k^T$.
   - (ii) Solve the ARE
   $$Y_{k-1}(Y_{k-1}^TB_1)(B_1^TY_{k-1})Y_{k-1}^T + A_k^TZ_k + Z_kA_k - Z_kB_2B_2^TZ_k = 0$$
   using Newton-ADI, yielding Y_k with $Z_k \approx Y_kY_k^T$.
   - (iii) Set $R_{k+1} := \texttt{rrqr}([R_k,\ Y_k], \tau)$. $\{\%\ R_{k+1}R_{k+1}^T \approx X_{k+1}\}$
   - (iv) IF $\|(B_1^TY_k)Y_k^T\|_2 <$ tol THEN Stop.

AREs with Indefinite Hessian

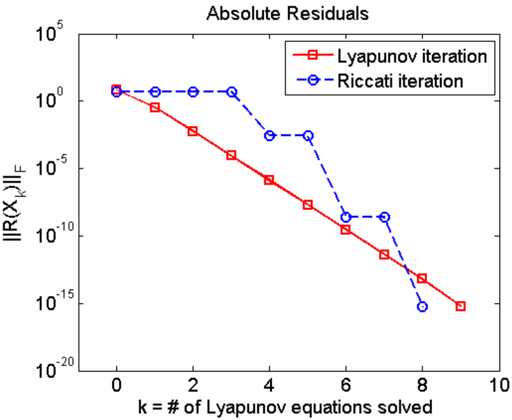
Numerical example

- 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.

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.



Software

Lyapack [Penzl 2000]

MATLAB toolbox for solving

- Lyapunov equations and algebraic Riccati equations,
- model reduction and LQR problems.

Main work horse: Low-rank ADI and Newton-ADI iterations.

Software

Lyapack [Penzl 2000]

MATLAB toolbox for solving

- Lyapunov equations and algebraic Riccati equations,
- model reduction and LQR problems.

Main work horse: Low-rank ADI and Newton-ADI iterations.

MESS – Matrix Equations Sparse Solvers [Saak/Mena/B. 2008]

- Extended and revised version of Lyapack.
- Includes solvers for large-scale differential Riccati equations (based on Rosenbrock and BDF methods).
- Many algorithmic improvements:
  - ADI new parameter selection,
  - column compression based on RRQR,
  - more efficient use of direct solvers,
  - treatment of generalized systems without factorization of the mass matrix.

Software

Lyapack [Penzl 2000]

MATLAB toolbox for solving

- Lyapunov equations and algebraic Riccati equations,
- model reduction and LQR problems.

Main work horse: Low-rank ADI and Newton-ADI iterations.

MESS – Matrix Equations Sparse Solvers [Saak/Mena/B. 2008]

- Extended and revised version of Lyapack.
- Includes solvers for large-scale differential Riccati equations (based on Rosenbrock and BDF methods).
- Many algorithmic improvements:
  - ADI new parameter selection,
  - column compression based on RRQR,
  - more efficient use of direct solvers,
  - treatment of generalized systems without factorization of the mass matrix.

Software

Lyapack [Penzl 2000]

MATLAB toolbox for solving

- Lyapunov equations and algebraic Riccati equations,
- model reduction and LQR problems.

Main work horse: Low-rank ADI and Newton-ADI iterations.

MESS – Matrix Equations Sparse Solvers [Saak/Mena/B. 2008]

- Extended and revised version of Lyapack.
- Includes solvers for large-scale differential Riccati equations (based on Rosenbrock and BDF methods).
- Many algorithmic improvements:
  - ADI new parameter selection,
  - column compression based on RRQR,
  - more efficient use of direct solvers,
  - treatment of generalized systems without factorization of the mass matrix.

Conclusions and Open Problems

- Low-rank Newton-ADI is a powerful and reliable method for solving large-scale AREs with semidefinite Hessian.
- Software is available in MATLAB toolbox Lyapack and its successor MESS.
- LQR problems on finite-time horizons can be solved using large-scale DRE solvers in MESS.
- Low-rank Riccati iteration yields a reliable and efficient method for large-scale AREs with indefinite Hessian, useful, e.g., for H_∞ 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.

Conclusions and Open Problems

- Low-rank Newton-ADI is a powerful and reliable method for solving large-scale AREs with semidefinite Hessian.
- Software is available in MATLAB toolbox Lyapack and its successor MESS.
- LQR problems on finite-time horizons can be solved using large-scale DRE solvers in MESS.
- Low-rank Riccati iteration yields a reliable and efficient method for large-scale AREs with indefinite Hessian, useful, e.g., for H_∞ 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.

Conclusions and Open Problems

- To-Do list:
  ... for AREs with semidefinite Hessian:
  - computation of stabilizing initial guess.
    (If hierarchical grid structure is available, a multigrid approach is possible, other approaches based on "cheaper" matrix equations under development.)
  - Implementation of coupled Riccati solvers for LQG controller design and balancing-related model reduction.
  - Implementation of ARE and Lyapunov solvers for Oseen/Stokes systems.

Conclusions and Open Problems

- To-Do list:

  ...for AREs with semidefinite Hessian:

  - computation of stabilizing initial guess.
    (If hierarchical grid structure is available, a multigrid approach is possible, other approaches based on "cheaper" matrix equations under development.)
  - Implementation of coupled Riccati solvers for LQG controller design and balancing-related model reduction.
  - Implementation of ARE and Lyapunov solvers for Oseen/Stokes systems.

  ...for AREs with indefinite Hessian:

  - 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?

References

1 L. Amodei.
An oblique projection method for large-scale algebraic Riccati equations.
Manuscript, 2003.

2 H.T. Banks and K. Ito.
A numerical algorithm for optimal feedback gains in high dimensional linear quadratic regulator problems.
SIAM J. Cont. Optim., 29(3):499–515, 1991.

3 U. Baur and P. Benner.
Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic.
Computing, 78(3):211–234, 2006.

4 P. Benner.
Solving large-scale control problems.
IEEE Control Systems Magazine, 14(1):44–59, 2004.

5 P. Benner.
Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems.
Logos–Verlag, Berlin, Germany, 1997.
Also: Dissertation, Fakultät für Mathematik, TU Chemnitz–Zwickau, 1997.

6 P. Benner and R. Byers.
Step size control for Newton's method applied to algebraic Riccati equations.
In J.G. Lewis, editor, *Proc. Fifth SIAM Conf. Appl. Lin. Alg., Snowbird, UT*, pages 177–181. SIAM, Philadelphia, PA, 1994.

7 P. Benner and R. Byers.
An exact line search method for solving generalized continuous-time algebraic Riccati equations.
IEEE Trans. Automat. Control, 43(1):101–107, 1998.

8 P. Benner and H. Faßbender.
An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem.
Linear Algebra Appl., 263:75–111, 1997.

9 P. Benner, J.-R. Li, and T. Penzl.
Numerical solution of large Lyapunov equations, Riccati equations, and linear-quadratic control problems.
Numer. Lin. Alg. Appl., 2008.
Reprint of unpublished manuscript, December 1999.

References

[10] P. Benner, V. Mehrmann, and D.C. Sorensen.
Approximate low rank solution of large scale algebraic Riccati equations.
Manuscript, 2003.

[11] J. Burns, E. Sachs, and L. Zietsmann.
Mesh independence of Kleinman-Newton iterations for Riccati equations in Hilbert space.
Preprint, 2006.

[12] L. Cherfi, H. Abou-Kandil, and H. Bourles.
Iterative method for general algebraic Riccati equation.
Proc. ICGST Intl. Conf. Automatic Control and System Engineering, ACSE 2005, 19–21 December 2005, Cairo, Egypt.

[13] W.R. Ferng, W.-W. Lin, and C.-S. Wang.
The shift-inverted J-Lanczos algorithm for the numerical solutions of large sparse algebraic Riccati equations.
Comput. Math. Appl., 33(10):23–40, 1997.

[14] T. Damm.
Rational Matrix Equations in Stochastic Control.
No. 297 in Lecture Notes in Control and Information Sciences.
Springer-Verlag, Berlin/Heidelberg, FRG, 2004.

[15] J. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis.
State-space solutions to standard H_2 and H_∞ control problems.
IEEE Trans. Automat. Control, 34:831–847, 1989.

[16] R. Freund and V. Mehrmann.
A symplectic look ahead Lanczos method for Hamiltonian matrices.
Manuscript, 1992.

[17] L. Grasedyck, W. Hackbusch, and B.N. Khoromskij.
Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices.
Computing, 70:121–165, 2003.

[18] C.-H. Guo and A.J. Laub.
On a Newton-like method for solving algebraic Riccati equations.
SIAM J. Matrix Anal. Appl., 21(2):694–698, 2000.

References

19 I.M. Jaimoukha and E.M. Kasenally.
Krylov subspace methods for solving large Lyapunov equations.
SIAM J. Numer. Anal., 31:227–251, 1994.

20 K. Jbilou.
Block Krylov subspace methods for large continuous-time algebraic Riccati equations.
Numer. Algorithms, 34:339–353, 2003.

21 K. Jbilou.
An Arnoldi based algorithm for large algebraic Riccati equations.
Appl. Math. Lett., 19(5):437–444, 2006.

22 D.L. Kleinman.
On an iterative technique for Riccati equation computations.
IEEE Trans. Automat. Control, AC-13:114–115, 1968.

23 P. Lancaster and L. Rodman.
The Algebraic Riccati Equation.
Oxford University Press, Oxford, 1995.

24 A. Lanzon, Y. Feng, and B.D.O. Anderson.
An iterative algorithm to solve algebraic Riccati equations with an indefinite quadratic term.
Proc. European Control Conf., ECC 2007, July 2–5, 2007, Kos, Greece, Paper WeA13.4.

25 J.-R. Li and J. White.
Low rank solution of Lyapunov equations.
SIAM J. Matrix Anal. Appl., 24(1):260–280, 2002.

26 V. Mehrmann.
The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.
Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

27 T. Penzl.
A cyclic low rank Smith method for large sparse Lyapunov equations.
SIAM J. Sci. Comput., 21(4):1401–1418, 2000.

References

28 T. Penzl.
LYAPACK Users Guide.
Technical Report SFB393/00-33, Sonderforschungsbereich 393 *Numerische Simulation auf massiv parallelen Rechnern*, TU Chemnitz, 09107 Chemnitz, FRG, 2000.
Available from `http://www.tu-chemnitz.de/sfb393/sfb00pr.html`.

29 J. Saak, H. Mena, and P. Benner.
Matrix Equation Sparse Solvers (MESS): a MATLAB Toolbox for the Solution of Sparse Large-Scale Matrix Equations.
In preparation, 2008.

30 E.L. Wachspress.
Iterative solution of the Lyapunov matrix equation.
Appl. Math. Letters, 107:87–90, 1988.