On some algorithmic issues concerning nonsymmetric algebraic Riccati equations

Beatrice Meini
Joint work with D. Bini, B. Iannazzo, F. Poloni

Università di Pisa

Chemnitz, June 15, 2007

1. Nonsymmetric Algebraic Riccati Equations
   - Preliminaries
   - Outline of SDA
   - Cyclic reduction
2. New class of algorithms
   - From NARE to UQME
   - Eigenvalue transform
3. Applications
   - Singular cases
   - A problem from transport theory

Nonsymmetric Algebraic Riccati Equations

Given $D \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times m}$, $C \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{m\times n}$, find $X \in \mathbb{R}^{m\times n}$ such that

NARE

$$XCX - AX - XD + B = 0 \tag{1}$$

Nonsymmetric Algebraic Riccati Equations

Given $D \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times m}$, $C \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{m\times n}$, find $X \in \mathbb{R}^{m\times n}$ such that

NARE

$$XCX - AX - XD + B = 0 \qquad (1)$$

Remark: Any solution X of (1) is such that

$$\begin{bmatrix} D & -C \\ B & -A \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} (D - CX)$$

The eigenvalues of $D - CX$ are eigenvalues of $H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}$

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with
$\mathrm{Re}(\lambda_{m+n}) \leq \ldots \leq \mathrm{Re}(\lambda_2) \leq \mathrm{Re}(\lambda_1)$.

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with
Re$(\lambda_{m+n}) \leq \ldots \leq$Re$(\lambda_2) \leq$Re$(\lambda_1)$.

- If M is nonsingular then Re$(\lambda_{n+1}) < 0 <$Re(λ_n)

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with Re$(\lambda_{m+n}) \leq \ldots \leq$Re$(\lambda_2) \leq$Re$(\lambda_1)$.

- If M is nonsingular then Re$(\lambda_{n+1}) < 0 <$Re(λ_n)
- If M is singular, then Re$(\lambda_{n+1}) \leq 0 \leq$Re$(\lambda_n)$. Moreover, only one of the following conditions is satisfied:

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \dots, \lambda_{m+n}\}$, with Re$(\lambda_{m+n}) \leq \dots \leq$Re$(\lambda_2) \leq$Re$(\lambda_1)$.

- If M is nonsingular then Re$(\lambda_{n+1}) < 0 <$Re(λ_n)
- If M is singular, then Re$(\lambda_{n+1}) \leq 0 \leq$Re$(\lambda_n)$. Moreover, only one of the following conditions is satisfied:
  - $\lambda_n = 0$ and $\lambda_{n+1} \in \mathbb{R}^-$ (positive recurrent case);

Important case

Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \dots, \lambda_{m+n}\}$, with Re$(\lambda_{m+n}) \leq \dots \leq$Re$(\lambda_2) \leq$Re$(\lambda_1)$.

- If M is nonsingular then Re$(\lambda_{n+1}) < 0 <$Re(λ_n)
- If M is singular, then Re$(\lambda_{n+1}) \leq 0 \leq$Re$(\lambda_n)$. Moreover, only one of the following conditions is satisfied:
  - $\lambda_n = 0$ and $\lambda_{n+1} \in \mathbb{R}^-$ (positive recurrent case);
  - $\lambda_n \in \mathbb{R}^+$ and $\lambda_{n+1} = 0$ (transient case);

Important case

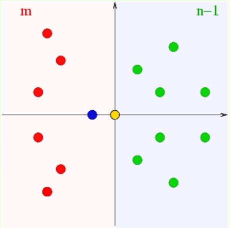
Assumption : assume that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

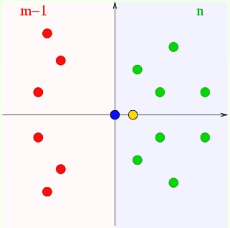
is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with Re$(\lambda_{m+n}) \leq \ldots \leq$Re$(\lambda_2) \leq$Re$(\lambda_1)$.

- If M is nonsingular then Re$(\lambda_{n+1}) < 0 <$Re(λ_n)
- If M is singular, then Re$(\lambda_{n+1}) \leq 0 \leq$Re$(\lambda_n)$. Moreover, only one of the following conditions is satisfied:
  - $\lambda_n = 0$ and $\lambda_{n+1} \in \mathbb{R}^-$ (positive recurrent case);
  - $\lambda_n \in \mathbb{R}^+$ and $\lambda_{n+1} = 0$ (transient case);
  - $\lambda_n = \lambda_{n+1} = 0$ (null recurrent case).

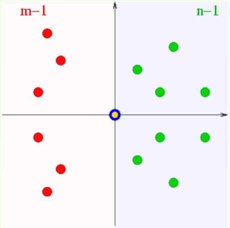
Location of the eigenvalues: singular case



Positive recurrent



Transient



Null recurrent
(Critical case)

Interest

Compute the minimal entrywise nonnegative solution S of the NARE (1)

Interest

Compute the minimal entrywise nonnegative solution S of the NARE (1)

Invariant subspace property:
The seeked solution S is the unique matrix such that

$$H \begin{bmatrix} I \\ S \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix} R, \qquad R = D - CS,$$

and $\sigma(R) = \{\lambda_1, \ldots, \lambda_n\}$.

Interest

Compute the minimal entrywise nonnegative solution S of the NARE (1)

Invariant subspace property:

The seeked solution S is the unique matrix such that

$$H \begin{bmatrix} I \\ S \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix} R, \qquad R = D - CS,$$

and $\sigma(R) = \{\lambda_1, \ldots, \lambda_n\}$.

There are many algorithms for solving AREs based on the invariant subspace property

One of the most efficient is the Structure-preserving Doubling Algorithm (SDA) by [Guo, Lin, Wei, 2006]

Outline of SDA

- Assume for simplicity that M is a nonsingular M-matrix. Therefore $\sigma(R) = \{\lambda_1, \ldots, \lambda_n\} \in \mathbb{C}^+$.
- Apply the Cayley transform $z \to (z-\gamma)/(z+\gamma)$ with $\gamma > 0$ to R and obtain

$$(H - \gamma I)\begin{bmatrix} I \\ S \end{bmatrix} = (H + \gamma I)\begin{bmatrix} I \\ S \end{bmatrix} R_\gamma,$$

where $R_\gamma = (R + \gamma I)^{-1}(R - \gamma I)$.

Key property: $\rho(R_\gamma) < 1$

Outline of SDA

SDA generates the matrix sequences

$$L_k = \begin{bmatrix} D_k & 0 \\ -H_k & I \end{bmatrix}, \quad U_k = \begin{bmatrix} I & -G_k \\ 0 & F_k \end{bmatrix}$$

such that

$$L_k \begin{bmatrix} I \\ S \end{bmatrix} = U_k \begin{bmatrix} I \\ S \end{bmatrix} R_\gamma^{2^k}, \quad k = 0, 1, \ldots$$

Since $\rho(R_\gamma) < 1$ then H_k quadratically converges to S

The computation of L_{k+1} and U_{k+1}, given L_k and U_k is performed by means of

$$D_{k+1} = D_k(I - G_k H_k)^{-1} D_k$$

$$F_{k+1} = F_k(I - H_k G_k)^{-1} F_k$$

$$G_{k+1} = G_k + D_k(I - G_k H_k)^{-1} G_k F_k$$

$$H_{k+1} = H_k + F_k(I - H_k G_k)^{-1} H_k D_k$$

for $k = 0, 1, 2, \ldots$, where

$$D_0 = I - 2\gamma V^{-1} \qquad F_0 = I - 2\gamma W^{-1}$$

$$G_0 = 2\gamma(D + \gamma I)^{-1} C W^{-1} \qquad H_0 = 2\gamma W^{-1} B (D + \gamma I)^{-1}$$

$$W = A + \gamma I - B(D + \gamma I)^{-1} C \qquad V = D + \gamma I - C(A + \gamma I)^{-1} B$$

Computational cost and convergence of SDA

- The matrix operations can be arranged in such a way that

  10 matrix products, 2 LU factorizations

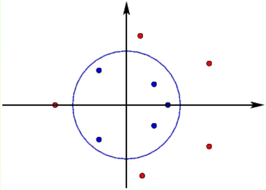
  are sufficient, for the overall cost of $\frac{64}{3}n^3$ ops (where we assume $m = n$).
- If M is nonsingular, or M is singular irreducible and $\lambda_n \neq \lambda_{n+1}$, the convergence is quadratic.
- In the critical case ($\lambda_n = \lambda_{n+1} = 0$) the convergence turns to linear with rate $1/2$.

Cyclic Reduction (CR) for UQME

Given: $A_0, A_1, A_2 \in \mathbb{R}^{N\times N}$ such that the roots of $\varphi(\lambda) = \det(A_0 + \lambda A_1 + \lambda^2 A_2)$ are

$$|\xi_1| \le \cdots \le |\xi_N| \le 1 < |\xi_{N+1}| \le \cdots \le |\xi_{2N}|$$

(including zeros at ∞ if $\deg\varphi(\lambda) < 2N$)

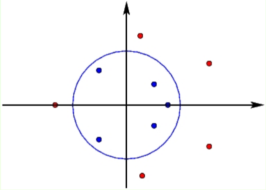


Cyclic Reduction (CR) for UQME

Given: $A_0, A_1, A_2 \in \mathbb{R}^{N\times N}$ such that the roots of $\varphi(\lambda) = \det(A_0 + \lambda A_1 + \lambda^2 A_2)$ are

$$|\xi_1| \le \cdots \le |\xi_N| \le 1 < |\xi_{N+1}| \le \cdots \le |\xi_{2N}|$$

(including zeros at ∞ if $\deg\varphi(\lambda) < 2N$)



Goal: compute the solution G of the Unilateral Quadratic Matrix Equation (UQME)

$$A_0 + A_1 X + A_2 X^2 = 0,$$

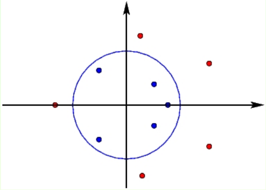
such that $\rho(G) = |\xi_N|$, provided it exists.

Cyclic Reduction (CR) for UQME

Given: $A_0, A_1, A_2 \in \mathbb{R}^{N\times N}$ such that the roots of $\varphi(\lambda) = \det(A_0 + \lambda A_1 + \lambda^2 A_2)$ are

$$|\xi_1| \le \cdots \le |\xi_N| \le 1 < |\xi_{N+1}| \le \cdots \le |\xi_{2N}|$$

(including zeros at ∞ if $\deg\varphi(\lambda) < 2N$)



Goal: compute the solution G of the Unilateral Quadratic Matrix Equation (UQME)

$$A_0 + A_1X + A_2X^2 = 0,$$

such that $\rho(G) = |\xi_N|$, provided it exists.

CR is a quadratically convergent, fast and numerically stable algorithm for computing G

Cyclic Reduction

CR is a versatile algorithm invented by G. Golub [Buzbee, Golub, Nielson 1970] for the f.d. Poisson equation.

Details on this algorithm can be found in the book
D. Bini, G. Latouche, B. Meini, "Numerical Solution of Structured Markov Chains", Oxford Univ. Press 2005.

Cyclic Reduction

CR is a versatile algorithm invented by G. Golub [Buzbee, Golub, Nielson 1970] for the f.d. Poisson equation.

Details on this algorithm can be found in the book
D. Bini, G. Latouche, B. Meini, "Numerical Solution of Structured Markov Chains", Oxford Univ. Press 2005.

- Rediscovered by Latouche and Ramaswami (1993) for QBDs

Cyclic Reduction

CR is a versatile algorithm invented by G. Golub [Buzbee, Golub, Nielson 1970] for the f.d. Poisson equation.

Details on this algorithm can be found in the book
D. Bini, G. Latouche, B. Meini, "Numerical Solution of Structured Markov Chains", Oxford Univ. Press 2005.

- Rediscovered by Latouche and Ramaswami (1993) for QBDs
- Revisited by Bini and Meini (1996ff), applied to UQMEs and extended to equations of the kind $X = \sum_{i=0}^{+\infty} A_i X^i$

Cyclic Reduction

CR is a versatile algorithm invented by G. Golub [Buzbee, Golub, Nielson 1970] for the f.d. Poisson equation.

Details on this algorithm can be found in the book
D. Bini, G. Latouche, B. Meini, "Numerical Solution of Structured Markov Chains", Oxford Univ. Press 2005.

- Rediscovered by Latouche and Ramaswami (1993) for QBDs
- Revisited by Bini and Meini (1996ff), applied to UQMEs and extended to equations of the kind $X = \sum_{i=0}^{+\infty} A_i X^i$
- Applied to the following matrix equations: $X = A \pm BX^{-1}C$ (Meini 2002); matrix square and pth root (Bini, Higham, Meini 2005); NARE (Ramaswami 1999).

Few words about CR for UQME

CR generates the matrix sequences

$$\begin{aligned}
A_0^{(k+1)} &= -A_0^{(k)} S^{(k)} A_0^{(k)}, \quad S^{(k)} = (A_1^{(k)})^{-1} \\
A_2^{(k+1)} &= -A_2^{(k)} S^{(k)} A_2^{(k)}, \\
A_1^{(k+1)} &= A_1^{(k)} - A_0^{(k)} S^{(k)} A_2^{(k)} - A_2^{(k)} S^{(k)} A_0^{(k)}, \\
\widehat{A}^{(k+1)} &= \widehat{A}^{(k)} - A_0^{(k)} S^{(k)} A_2^{(k)}, \quad k \geq 0
\end{aligned}$$

starting from $A_i^{(0)} = A_i$, $i = 1, 2, 3$, $\widehat{A}^{(0)} = A_1$, such that

$$\begin{aligned}
A_0^{(k)} + A_1^{(k)} G^{2^k} + A_2^{(k)} G^{2 \cdot 2^k} = 0 \\
A_0 + \widehat{A}^{(k)} G + A_2^{(k)} G^{2^k + 1} = 0
\end{aligned}$$

Few words about CR for UQME

Convergence property: the convergence is quadratic, more specifically:

$$||(\widehat{A}^{(k)})^{-1}A_0 - G|| = O(|\xi_N/\xi_{N+1}|^{2^k})$$

Few words about CR for UQME

Convergence property: the convergence is quadratic, more specifically:

$$||(\widehat{A}^{(k)})^{-1}A_0 - G|| = O(|\xi_N/\xi_{N+1}|^{2^k})$$

Cost: 6 matrix products, one PLU factorization: $\frac{38}{3}N^3$ ops

Few words about CR for UQME

Convergence property: the convergence is quadratic, more specifically:

$$||(\widehat{A}^{(k)})^{-1}A_0 - G|| = O(|\xi_N/\xi_{N+1}|^{2^k})$$

Cost: 6 matrix products, one PLU factorization: $\frac{38}{3}N^3$ ops

Applicability: under mild conditions the matrices $A_0^{(k)}$ are invertible

Few words about CR for UQME

Convergence property: the convergence is quadratic, more specifically:

$$||(\widehat{A}^{(k)})^{-1}A_0 - G|| = O(|\xi_N/\xi_{N+1}|^{2^k})$$

Cost: 6 matrix products, one PLU factorization: $\frac{38}{3}N^3$ ops

Applicability: under mild conditions the matrices $A_0^{(k)}$ are invertible

Critical case: If $|\xi_N| = |\xi_{N+1}| = 1$ convergence turns to linear with rate $1/2$. Quadratic convergence can be recovered by means of the shift technique.

Functional formulation of CR

Let

$$\varphi_k(\lambda) = A_0^{(k)} + \lambda A_1^{(k)} + \lambda^2 A_2^{(k)}, \quad \psi_k(\lambda) = \lambda \varphi_k(\lambda)^{-1}, \quad k = 0, 1, \ldots$$

Theorem

For any $k \geq 0$,

$$\varphi_{k+1}(\lambda^2) = \varphi_k(\lambda) \left(A_1^{(k)} \right)^{-1} \varphi_k(-\lambda) \quad \text{(Graeffe iteration)}$$

$$\psi_{k+1}(\lambda^2) = \tfrac{1}{2}(\psi_k(\lambda) + \psi_k(-\lambda))$$

New class of algorithms

Idea: To transform the NARE into a UQME of the kind

$$A_0 + A_1 Y + A_2 Y^2 = 0, \quad A_0, A_1, A_2 \in \mathbb{R}^{(m+n)\times(m+n)}$$

such that $\det(A_0 + \lambda A_1 + \lambda^2 A_2)$ has roots

$$|\xi_1| \le \cdots \le |\xi_{m+n}| \le 1 \le |\xi_{m+n+1}| \le \cdots \le |\xi_{2(m+n)}|$$

and apply cyclic reduction.

New class of algorithms

Idea: To transform the NARE into a UQME of the kind

$$A_0 + A_1 Y + A_2 Y^2 = 0, \quad A_0, A_1, A_2 \in \mathbb{R}^{(m+n)\times(m+n)}$$

such that $\det(A_0 + \lambda A_1 + \lambda^2 A_2)$ has roots

$$|\xi_1| \le \cdots \le |\xi_{m+n}| \le 1 \le |\xi_{m+n+1}| \le \cdots \le |\xi_{2(m+n)}|$$

and apply cyclic reduction.

H.-G. Xu and L.-Z. Lu (1995) reduced an ARE to an equation $Y^2 - M^2 = 0$ but with no splitting property.

First transform of AREs to UQMEs

The linear matrix pencil

$$H - \lambda I = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix} - \lambda I$$

can be transformed into a quadratic matrix polynomial by multiplying the second block column by λ

$$\begin{bmatrix} D & 0 \\ B & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & -C \\ 0 & -A \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix}$$

This matrix polynomial defines a UQME

First transform of AREs to UQMEs

Theorem

The matrix

$$Y = \begin{bmatrix} D - CX & 0 \\ X & 0 \end{bmatrix}$$

solves the UQME

$$\begin{bmatrix} D & 0 \\ B & 0 \end{bmatrix} + \begin{bmatrix} -I & -C \\ 0 & -A \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} Y^2 = 0$$

if and only if X solves the NARE (1). Moreover, the eigenvalues of $D - CX$ are eigenvalues of the matrix pencil

First transform of AREs to UQMEs

Theorem

The matrix

$$Y = \begin{bmatrix} D - CX & 0 \\ X & 0 \end{bmatrix}$$

solves the UQME

$$\begin{bmatrix} D & 0 \\ B & 0 \end{bmatrix} + \begin{bmatrix} -I & -C \\ 0 & -A \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} Y^2 = 0$$

if and only if X *solves the NARE (1). Moreover, the eigenvalues of* $D - CX$ *are eigenvalues of the matrix pencil*

Remark: The seeked solution S of the NARE is obtained from the solution Y of the UQME with eigenvalues with the largest real parts

Second transform of AREs to UQMEs

Consider the block LU factorization $H = LU$ and transform the pencil

$$H - \lambda I = LU - \lambda I$$

into the new pencil

$$U - \lambda L^{-1} = \begin{bmatrix} D & C(A - BD^{-1}C)^{-1} \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ -BD^{-1} & -A + BD^{-1}C \end{bmatrix}$$

Multiply the second block column by λ and get the quadratic matrix polynomial

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & C(A - BD^{-1}C)^{-1} \\ BD^{-1} & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & A - BD^{-1}C \end{bmatrix}$$

Second transform of AREs to UQMEs

Theorem

The matrix

$$Y = \begin{bmatrix} D - CX & 0 \\ X & 0 \end{bmatrix}$$

solves the UQME

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I & C(A - BD^{-1}C)^{-1} \\ BD^{-1} & I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & A - BD^{-1}C \end{bmatrix} Y^2 = 0$$

if and only if X solves the NARE (1). Moreover, the eigenvalues of $D - CX$ are eigenvalues of the matrix pencil

Second transform of AREs to UQMEs

Theorem

The matrix

$$Y = \begin{bmatrix} D - CX & 0 \\ X & 0 \end{bmatrix}$$

solves the UQME

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I & C(A - BD^{-1}C)^{-1} \\ BD^{-1} & I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & A - BD^{-1}C \end{bmatrix} Y^2 = 0$$

if and only if X solves the NARE (1). Moreover, the eigenvalues of $D - CX$ are eigenvalues of the matrix pencil

Remark: The seeked solution S of the NARE is obtained from the solution Y of the UQME with eigenvalues with the largest real parts

A few remarks

- The UQMEs encountered in the previous slides are associated with matrix polynomials of the kind $\varphi(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$

$$\varphi(\lambda) = \begin{cases} \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ 0 & * \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} \\ \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ * & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & * \end{bmatrix} \end{cases}$$

A few remarks

- The UQMEs encountered in the previous slides are associated with matrix polynomials of the kind $\varphi(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$

$$\varphi(\lambda) = \begin{cases} \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ 0 & * \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} \\ \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ * & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & * \end{bmatrix} \end{cases}$$

- the nonzero roots of $\det \varphi(\lambda)$ are eigenvalues of $H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}$

A few remarks

- The UQMEs encountered in the previous slides are associated with matrix polynomials of the kind $\varphi(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$

$$\varphi(\lambda) = \begin{cases} \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ 0 & * \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} \\ \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ * & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & * \end{bmatrix} \end{cases}$$

- the nonzero roots of $\det \varphi(\lambda)$ are eigenvalues of $H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}$
- The nonzero roots of $\det \varphi(\lambda)$ have a splitting w.r.t. the imaginary axis

A few remarks

- The UQMEs encountered in the previous slides are associated with matrix polynomials of the kind $\varphi(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$

$$\varphi(\lambda) = \begin{cases} \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ 0 & * \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} \\ \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -I & * \\ * & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & * \end{bmatrix} \end{cases}$$

- the nonzero roots of $\det \varphi(\lambda)$ are eigenvalues of $H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}$
- The nonzero roots of $\det \varphi(\lambda)$ have a splitting w.r.t. the imaginary axis
- The solution of the UQME associated with the eigenvalues with the largest real part is the one to be computed

A few remarks

- Algorithms for UQME reach the highest efficiency for eigenvalues split w.r.t. the unit circle where the solution with eigenvalues of modulus less than 1 is seeked
- Therefore it is important to transform a splitting w.r.t the imaginary axis into a splitting w.r.t. the unit circle

A few remarks

- Algorithms for UQME reach the highest efficiency for eigenvalues split w.r.t. the unit circle where the solution with eigenvalues of modulus less than 1 is seeked
- Therefore it is important to transform a splitting w.r.t the imaginary axis into a splitting w.r.t. the unit circle

  Three approaches:

  - shrink and shift (Ramaswami 1999)
  - Cayley transform applied to the pencil (Guo, Lin, Wei, 2006)
  - Cayley transform applied to the UQME (Bini, Latouche, Meini, 2006)

Shrink and shift

Multiply the Riccati equation by t,

$$tXCX - tAX - tXD + tB = 0,$$

Shrink and shift

Multiply the Riccati equation by t,

$$tXCX - tAX - tXD + tB = 0,$$

add I to $-tA$ and subtract I from $-tD$ and get:

$$tXCX - (tA - I)X - X(tD + I) + tB = 0 \qquad (2)$$

The associated matrix is

$$H_t = \begin{bmatrix} I + tD & -tC \\ tB & I - tA \end{bmatrix}$$

Shrink and shift

Multiply the Riccati equation by t,

$$tXCX - tAX - tXD + tB = 0,$$

add I to $-tA$ and subtract I from $-tD$ and get:

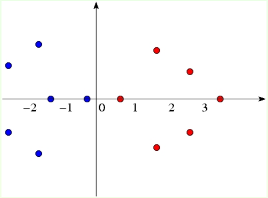
$$tXCX - (tA - I)X - X(tD + I) + tB = 0 \tag{2}$$

The associated matrix is

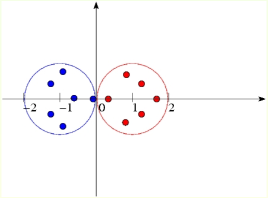
$$H_t = \begin{bmatrix} I + tD & -tC \\ tB & I - tA \end{bmatrix}$$

If $0 < t < 1/\max(a_{i,i}, d_{i,i})$ the eigenvalues of H_t have a splitting w.r.t. the unit circle

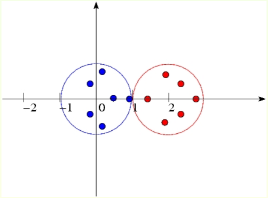
Transformation of the eigenvalues



Original eigenvalues



Shrink by t



Shift by 1

Combination with the UQME tranforms

- The new NARE (2) is transformed into a UQME, by transforming the pencil $H_t - \lambda I$ into a quadratic matrix polynomial.

Combination with the UQME tranforms

- The new NARE (2) is transformed into a UQME, by transforming the pencil $H_t - \lambda I$ into a quadratic matrix polynomial.
- The eigenvalues of the UQME are split with respect to the unit circle.

Combination with the UQME tranforms

- The new NARE (2) is transformed into a UQME, by transforming the pencil $H_t - \lambda I$ into a quadratic matrix polynomial.
- The eigenvalues of the UQME are split with respect to the unit circle.
- CR applied to the UQME deriving from the first transform leads to the algorithm of Ramaswami (1999) of cost $(68/3)n^3$ ops per step.

Combination with the UQME tranforms

- The new NARE (2) is transformed into a UQME, by transforming the pencil $H_t - \lambda I$ into a quadratic matrix polynomial.
- The eigenvalues of the UQME are split with respect to the unit circle.
- CR applied to the UQME deriving from the first transform leads to the algorithm of Ramaswami (1999) of cost $(68/3)n^3$ ops per step.
- CR applied to the UQME deriving from the second transform leads to a new algorithm with the same cost of SDA, i.e. $(64/3)n^3$ per step. Formally, this algorithm differs from SDA only for the initial values of the matrix sequences. The iteration is the same. The initial values are simpler than the ones of SDA

Cayley transform applied to the pencil

- The Cayley transform $z \longrightarrow (z-\gamma)/(z+\gamma)$ applied to the pencil $H - \lambda I$ yields the pencil

$$H_\gamma - \lambda I, \quad H_\gamma = (H+\gamma I)^{-1}(H-\gamma I).$$

Cayley transform applied to the pencil

- The Cayley transform $z \longrightarrow (z-\gamma)/(z+\gamma)$ applied to the pencil $H - \lambda I$ yields the pencil

$$H_\gamma - \lambda I, \quad H_\gamma = (H+\gamma I)^{-1}(H-\gamma I).$$

- Since $\mu = \frac{\gamma-\lambda}{\gamma+\lambda}$ is eigenvalue of H_γ iff λ is eigenvalue of H, the eigenvalues of H_γ are split w.r.t. the unit circle.

Cayley transform applied to the pencil

- The Cayley transform $z \to (z-\gamma)/(z+\gamma)$ applied to the pencil $H - \lambda I$ yields the pencil

$$H_\gamma - \lambda I, \quad H_\gamma = (H+\gamma I)^{-1}(H-\gamma I).$$

- Since $\mu = \frac{\gamma-\lambda}{\gamma+\lambda}$ is eigenvalue of H_γ iff λ is eigenvalue of H, the eigenvalues of H_γ are split w.r.t. the unit circle.

- Two UQMEs are obtained by applying the two transforms to the pencil $H_\gamma - \lambda I$. The second transform yields

$$\varphi(\lambda) = \begin{bmatrix} -D_\gamma & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} I & -G_\gamma \\ -H_\gamma & I \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -F_\gamma \end{bmatrix}$$

Some theoretical results

Theorem

Cyclic Reduction applied to

$$\begin{bmatrix} -D_\gamma & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} I & -G_\gamma \\ -H_\gamma & I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & -F_\gamma \end{bmatrix} Y^2 = 0 \qquad (3)$$

coincides with SDA. Moreover, the spectral minimal solution of (3) is $\begin{bmatrix} R_\gamma & 0 \\ SR_\gamma & 0 \end{bmatrix}$*, where* $R_\gamma = (R + \gamma I)^{-1}(R - \gamma I)$.

Some theoretical results

Theorem

Assume that M *is nonsingular and let* $\mathcal{Q}(\lambda) = \lambda^{-1}\varphi(\lambda)$. *Then:*

- *The matrix function* $\mathcal{Q}(\lambda)$ *is analytic for* $|\xi| < |z| < |\eta|$, *where* $\xi = (\lambda_n - \gamma)/(\lambda_n + \gamma)$, $\eta = (\lambda_{n+1} - \gamma)/(\lambda_{n+1} + \gamma)$.

Some theoretical results

Theorem

Assume that M *is nonsingular and let* $\mathcal{Q}(\lambda) = \lambda^{-1}\varphi(\lambda)$. *Then:*

- *The matrix function* $\mathcal{Q}(\lambda)$ *is analytic for* $|\xi| < |z| < |\eta|$, *where* $\xi = (\lambda_n - \gamma)/(\lambda_n + \gamma)$, $\eta = (\lambda_{n+1} - \gamma)/(\lambda_{n+1} + \gamma)$.
- $\mathcal{Q}(\lambda)$ *has the canonical factorization*

$$\mathcal{Q}(\lambda) = \left(I - \lambda \begin{bmatrix} 0 & 0 \\ W & WS \end{bmatrix}\right) \begin{bmatrix} I & -G_\gamma \\ -S & I \end{bmatrix} \left(I - \lambda^{-1} \begin{bmatrix} R_\gamma & 0 \\ SR_\gamma & 0 \end{bmatrix}\right)$$

Some theoretical results

Theorem

Assume that M *is nonsingular and let* $\mathcal{Q}(\lambda) = \lambda^{-1}\varphi(\lambda)$. *Then:*

- *The matrix function* $\mathcal{Q}(\lambda)$ *is analytic for* $|\xi| < |z| < |\eta|$, *where* $\xi = (\lambda_n - \gamma)/(\lambda_n + \gamma)$, $\eta = (\lambda_{n+1} - \gamma)/(\lambda_{n+1} + \gamma)$.
- $\mathcal{Q}(\lambda)$ *has the canonical factorization*

$$\mathcal{Q}(\lambda) = \left(I - \lambda \begin{bmatrix} 0 & 0 \\ W & WS \end{bmatrix}\right) \begin{bmatrix} I & -G_\gamma \\ -S & I \end{bmatrix} \left(I - \lambda^{-1} \begin{bmatrix} R_\gamma & 0 \\ SR_\gamma & 0 \end{bmatrix}\right)$$

- *The series* $\psi(\lambda) = \mathcal{Q}(\lambda)^{-1}$, $\psi(\lambda) = \sum_{k=-\infty}^{+\infty} \lambda^k \psi_k$ *is such that*

$$\psi_0^{-1} = \begin{bmatrix} I & -T \\ -S & I \end{bmatrix}$$

where T *is the solution of the dual NARE of (1).*

Computational approaches

Remark: $\psi_0^{-1} = \lim_k A_0^{(k)}$

Other possible approaches are:

- Using numerical integration and the Cauchy integral theorem for computing ψ_0
- Using functional iterations borrowed from stochastic processes (QBD) for solving the UQME
- Using Newton's iteration applied to the UQME trying to exploit the specific matrix structure

The shift technique

Assume that M is a singular irreducible M-matrix such that $M\mathbf{e} = 0$.

Idea: To transform the original NARE into a new NARE

$$X\widetilde{C}X - \widetilde{A}X - X\widetilde{D} + \widetilde{B} = 0 \tag{4}$$

such that

$$\widetilde{M} = \begin{bmatrix} \widetilde{D} & -\widetilde{C} \\ -\widetilde{B} & \widetilde{A} \end{bmatrix}$$

is nonsingular and S solves (4)

Shift applied to NARE

- Let $\mathbf{p} \in \mathbb{R}^{m+n}$ such that $\mathbf{p}^T\mathbf{e} = 1$.
- If $\lambda_n = 0$ choose $\eta > 0$; if $\lambda_{n+1} = 0$ choose $\eta < 0$.
- Set $\widetilde{H} = H + \eta \mathbf{e}\mathbf{p}^T$, where

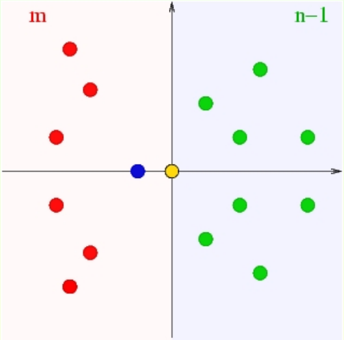
$$H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}, \quad \widetilde{H} = \begin{bmatrix} \widetilde{D} & -\widetilde{C} \\ \widetilde{B} & -\widetilde{A} \end{bmatrix}$$

- Consider the NARE

$$X\widetilde{C}X - \widetilde{A}X - X\widetilde{D} + \widetilde{B} = 0$$

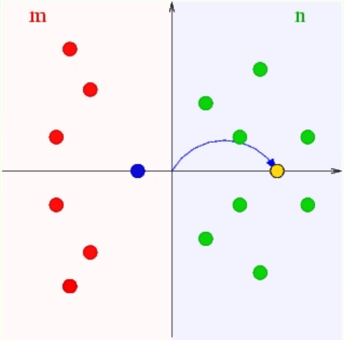
Eigenvalues properties

Positive recurrent case



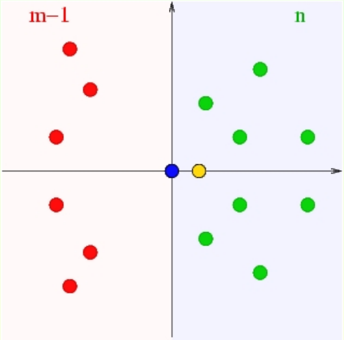
Eigenvalues properties

Positive recurrent case



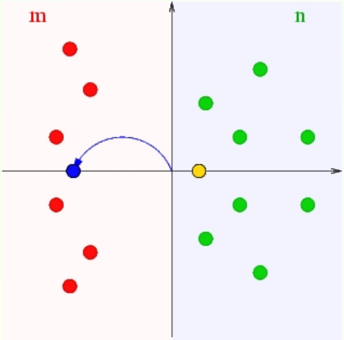
Eigenvalues properties

Transient case



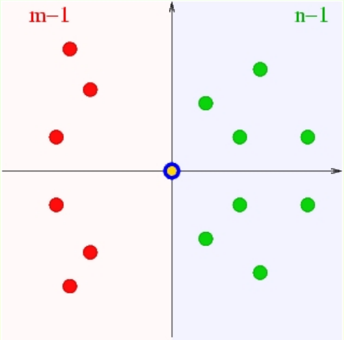
Eigenvalues/vectors properties

Transient case



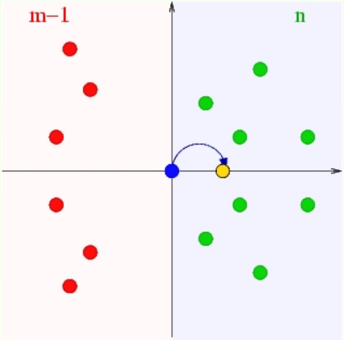
Eigenvalues/vectors properties

Null recurrent case (critical case)



Eigenvalues/vectors properties

Null recurrent case (critical case)



Properties of the solutions: case $\lambda_n = 0$

Let S be the solution of the original NARE such that

$$\sigma(D - CS) = \{\lambda_1, \ldots, \lambda_n\}$$

Then S is the solution of the new NARE such that

$$\sigma(\widetilde{D} - \widetilde{C}S) = \{\lambda_1, \ldots, \lambda_{n-1}, \eta\}$$

Therefore S can be obtained by solving the new NARE

A similar property holds if $\lambda_{n+1} = 0$.

Consequences

1. Algorithms which require M to be nonsingular can be applied to solve the new NARE.
2. The new NARE is generally better conditioned than the original NARE.
3. CR, SDA, Newton's method converge faster when applied to the new ARE, with respect to the same method applied to the original NARE. Moreover, in the critical case, convergence of CR, SDA and Newton's method turns to quadratic if applied to new NARE.

A problem from transport theory

We consider a specific instance of an ARE encountered in a problem of transport theory where

$$A = \Delta_1 - eq^T, \quad B = ee^T, \quad C = qq^T, \quad D = \Delta_2 - qe^T$$

$$\begin{array}{lll} \Delta_1 = \mathrm{diag}(\delta_1^{(1)}, \delta_2^{(1)}, \ldots, \delta_n^{(1)}), & \text{where} & \delta_i^{(1)} > 0, \\ \Delta_2 = \mathrm{diag}(\delta_1^{(2)}, \delta_2^{(2)}, \ldots, \delta_n^{(2)}), & \text{where} & \delta_i^{(2)} > 0, \\ e = (1, 1, \ldots, 1)^T, & & \\ q = (q_1, q_2, \ldots, q_n)^T, & \text{where} & q_i > 0. \end{array}$$

The Riccati equation is associated with a diagonal plus rank-one M-matrix [Guo01]

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix} = \begin{bmatrix} \Delta_2 & 0 \\ 0 & \Delta_1 \end{bmatrix} - \begin{bmatrix} q \\ e \end{bmatrix} \begin{bmatrix} e^T & q^T \end{bmatrix}$$

Remark: the matrix polynomials $\varphi(\lambda)$ obtained by transforming the NARE into a UQME are Cauchy-like, i.e.,

$$\mathcal{D}\varphi(\lambda) - \varphi(\lambda)\mathcal{D} = \text{rank } 2 \quad \text{where} \quad \mathcal{D} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

Consequence: $\psi(\lambda) = \lambda^{-1}\varphi(\lambda)^{-1}$ is Cauchy-like moreover

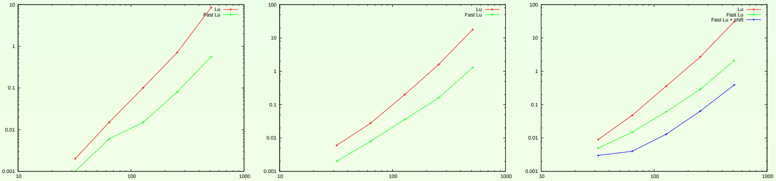
$$\mathcal{D}\psi(\lambda) - \psi(\lambda)\mathcal{D} = \mathbf{u}_1(\lambda)\mathbf{v}_1 + \mathbf{u}_2\mathbf{v}2(\lambda)$$

where $\mathbf{v}_1$ and $\mathbf{u}_2$ are independent of λ

This implies that the matrix functions $\psi_k(\lambda)$ generated by CR are Cauchy-like and $\varphi_k(\lambda)$ are Cauchy-like.

- The Cauchy-like structure can be exploited for designing an implementation of CR and SDA, based on the GKO algorithm at a cost of $O(n^2)$ ops per step.
- The shift technique can be implemented at the same cost
- Similar techniques can be applied for the implementation of Newton's iteration.
- Lu's quadratical convergent iteration can be implemented with $O(n^2)$ ops per step even if complemented with the shift technique.

Some numerical experiments



- $\alpha = .5$, $c = .5$ (noncritical case)
- $\alpha = 10^{-8}$, $c = 1 - 10^{-6}$ (close to critical case)
- $\alpha = 0$, $c = 1$ (critical case)

Noncritical case for $n = 512$, 15 times faster

Critical case with shift, 80 times faster

Accuracy in the critical case, relative error $\approx 10^{-15}$ (instead of $\approx 10^{-8}$)