On some recent algorithms for solving nonsymmetric algebraic Riccati equations

Beatrice Meini
Joint work with D. Bini and F. Poloni

Due giorni di algebra lineare numerica
Bologna, 6-7 Marzo, 2008

Università di Pisa

Nonsymmetric Algebraic Riccati Equations
- Preliminaries
- Outline of SDA
- Outline of Cyclic Reduction

NARE $\longrightarrow$ UQME
- Ramaswami's transform
- UL based transform
- "Small size" transform

Eigenvalues transform
- Shrink and shift
- Cayley transform

Numerical results and conclusions

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

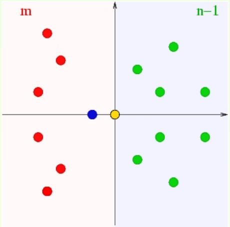
Assumption : assume that

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

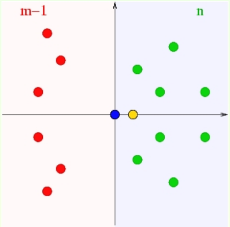
is either a nonsingular M-matrix or a singular irreducible M-matrix.
Spectral properties: let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with $\mathrm{Re}(\lambda_{m+n}) \leq \ldots \leq \mathrm{Re}(\lambda_2) \leq \mathrm{Re}(\lambda_1)$.

- If M is nonsingular then $\mathrm{Re}(\lambda_{n+1}) < 0 < \mathrm{Re}(\lambda_n)$
- If M is singular, then $\mathrm{Re}(\lambda_{n+1}) \leq 0 \leq \mathrm{Re}(\lambda_n)$. 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).

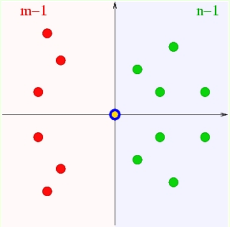
Location of the eigenvalues: singular case



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 $\sigma(R) = \{\lambda_1, \ldots, \lambda_n\}$. The solution S is called the extremal solution.

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 $\sigma(R) = \{\lambda_1, \dots, \lambda_n\} \in \mathbb{C}^+$.
- 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

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

Cost: $\frac{64}{3}n^3$ ops per step (where we assume $m = n$).

Remark: The convergence is still quadratic if M is singular irreducible and $\lambda_n \neq \lambda_{n+1}$. If $\lambda_n = \lambda_{n+1} = 0$ the convergence is linear with rate $1/2$. The convergence turns to quadratic by applying a shift technique to the null eigenvalues of H [Guo, [illegible], Meini, 07]

UNIVERSITÀ DI PISA

Cyclic Reduction (CR)

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

- 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].

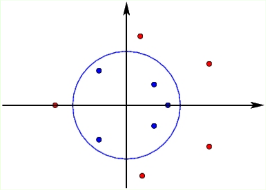
Details on this algorithm can be found in the book
D. Bini, G. Latouche, B. Meini, "Numerical Solution of Structured Markov Chains", Oxford Univ. Press 2005.

Few words about CR for UQME

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

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

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

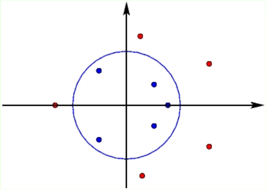


Few words about CR for UQME

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

Università di Pisa

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

$$A_0 + \widehat{A}^{(k)} G + A_2^{(k)} G^{2^k+1} = 0$$

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

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 [He, Meini, Rhee, 01].

New class of algorithms

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

$$A_0 + A_1 Y + A_2 Y^2 = 0, \quad A_0, A_1, A_2 \in \mathbb{R}^{N \times N}$$

with $N \leq m + n$, such that $\det(A_0 + A_1\lambda + A_2\lambda^2)$ has roots

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

and apply cyclic reduction.

New class of algorithms

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

$$A_0 + A_1 Y + A_2 Y^2 = 0, \quad A_0, A_1, A_2 \in \mathbb{R}^{N\times N}$$

with $N \le m + n$, such that $\det(A_0 + A_1\lambda + A_2\lambda^2)$ has roots

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

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.

Ramaswami's transform

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

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

This matrix polynomial defines a UQME

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

Ramaswami's transform

Theorem

The roots of the matrix polynomial $A(\lambda)$ *are:*

- m *equal to* 0
- *the* $m+n$ *eigenvalues* $\lambda_1, \ldots, \lambda_{m+n}$ *of* H
- n *at infinity.*

Moreover

$$V = \begin{bmatrix} D - CS & 0 \\ S & 0 \end{bmatrix},$$

where S *is the extremal solution of* (1)*, is the unique solution of the UQME* (2) *with* m *eigenvalues equal to zero and* n *eigenvalues equal to* $\lambda_1, \ldots, \lambda_n$*.*

UL based transform

Consider the block UL factorization

$$H = U^{-1}L, \quad U = \begin{bmatrix} I & -U_1 \\ 0 & U_2 \end{bmatrix}, \quad L = \begin{bmatrix} L_1 & 0 \\ -L_2 & I \end{bmatrix},$$

and transform the pencil $H - \lambda I$ into the new pencil

$$L - \lambda U.$$

UL based transform

Consider the block UL factorization

$$H = U^{-1}L, \quad U = \begin{bmatrix} I & -U_1 \\ 0 & U_2 \end{bmatrix}, \quad L = \begin{bmatrix} L_1 & 0 \\ -L_2 & I \end{bmatrix},$$

and transform the pencil $H - \lambda I$ into the new pencil

$$L - \lambda U.$$

Now multiply the second block row by $-\lambda$ and get

$$A(\lambda) = \begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I & U_1 \\ L_2 & -I \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 \\ 0 & U_2 \end{bmatrix} \lambda^2,$$

which defines the UQME

$$\begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I & U_1 \\ L_2 & -I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & U_2 \end{bmatrix} Y^2 = 0 \qquad (3)$$

UL based transform

Theorem

The roots of the matrix polynomial $A(\lambda)$ are:

- *m equal to 0*
- *the $m+n$ eigenvalues $\lambda_1, \ldots, \lambda_{m+n}$ of H*
- *n at infinity.*

Moreover

$$V = \begin{bmatrix} D - CS & 0 \\ S(D - CS) & 0 \end{bmatrix},$$

where S is the extremal solution of (1), is the unique solution of the UQME (3) with m eigenvalues equal to zero and n eigenvalues equal to $\lambda_1, \ldots, \lambda_n$.

"Small size" transform

The matrix pencil $H - \lambda I$ is transformed into

$$\begin{bmatrix} I & 0 \\ -U & I \end{bmatrix} H \begin{bmatrix} I & 0 \\ -U & I \end{bmatrix}^{-1} - \lambda I. \qquad (4)$$

If $\det C \neq 0$, by choosing $U = C^{-1}D$, (4) becomes

$$\begin{bmatrix} 0 & I \\ R(C^{-1}D) & A - C^{-1}DC \end{bmatrix} - \lambda I,$$

where $R(U) = UCU - AU - UD + B$, which defines the UQME

$$(B - AC^{-1}D)C + (C^{-1}DC - A)Y + Y^2 = 0$$

"Small size" transform

Theorem

The roots of

$$A(\lambda) = (B - AC^{-1}D)C + (C^{-1}DC - A)\lambda + I\lambda^2$$

are the eigenvalues of H.
Moreover, $Y = C^{-1}(D - CS)C$ *is the unique solution of the UQME*

$$(B - AC^{-1}D)C + (C^{-1}DC - A)Y + Y^2 = 0$$

with eigenvalues $\lambda_1, \ldots, \lambda_n$.

"Small size" transform

Remark: The condition $\det C \neq 0$ is not restrictive. Indeed, X solves (1) if and only if $\widetilde{X} = X(I - MX)^{-1}$ solves

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

where M is any matrix such that $\det(I - MX) \neq 0$, and

$$\begin{aligned} &\widetilde{A} = A - BM, \quad \widetilde{B} = B, \\ &\widetilde{C} = \widetilde{R}(M), \quad \widetilde{D} = D - MB, \\ &\widetilde{R}(M) = MBM - DM - MA + C. \end{aligned}$$

Open issue: Find M such that $\widetilde{R}(M)$ is well-conditioned.

A few remarks

- The UQMEs of size $m + n$ are associated with matrix polynomials of the kind

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

A few remarks

- The UQMEs of size $m+n$ are associated with matrix polynomials of the kind

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

- The eigenvalues of H are roots of $\det A(\lambda)$.

A few remarks

- The UQMEs of size $m+n$ are associated with matrix polynomials of the kind

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

- The eigenvalues of H are roots of $\det A(\lambda)$.
- The nonzero roots of $\det A(\lambda)$ have a splitting w.r.t. the imaginary axis

A few remarks

- The UQMEs of size $m+n$ are associated with matrix polynomials of the kind

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

- The eigenvalues of H are roots of $\det A(\lambda)$.
- The nonzero roots of $\det A(\lambda)$ 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

Eigenvalues transform

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.

Three approaches to transform a splitting w.r.t the imaginary axis into a splitting w.r.t. the unit circle:

- 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:

$$tXCX - (tA - I)X - X(tD + I) + tB = 0 \tag{5}$$

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:

$$tXCX - (tA - I)X - X(tD + I) + tB = 0 \qquad (5)$$

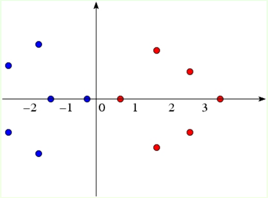
The associated matrix is

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

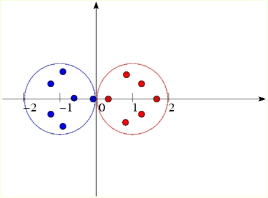
If $0 < t < 1/\max(a_{i,i}, d_{i,i})$ the eigenvalues of H_t have a splitting w.r.t. the unit circle

Università di Pisa

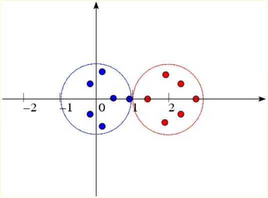
Transformation of the eigenvalues



Original eigenvalues



Shrink by t



Shift by 1

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

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

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.

Cayley transform applied to the pencil

- The Cayley transform $z \rightarrow (z-\gamma)/(z+\gamma)$ 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.
- Three UQMEs can be obtained from the pencil $H_\gamma - \lambda I$. The UL-based transform yields

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

SDA is CR!

Theorem

Cyclic Reduction applied to

$$\begin{bmatrix} -D_\gamma & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} I & -G_\gamma \\ -H_\gamma & I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & -F_\gamma \end{bmatrix} Y^2 = 0 \quad (6)$$

coincides with SDA. Moreover, the spectral minimal solution of (6) is $\begin{bmatrix} R_\gamma & 0 \\ SR_\gamma & 0 \end{bmatrix}$*, where* $R_\gamma = (R + \gamma I)^{-1}(R - \gamma I)$.

Università di Pisa

Some theoretical results

Theorem

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

Different combinations → different algorithms

We may combine the different strategies, for instance:

- "Shrink and shift" + "Ramaswami transform" lead to an algorithm similar to that of Ramaswami (1999) of cost $(68/3)n^3$ ops per step (`ss-ram`).

Different combinations → different algorithms

We may combine the different strategies, for instance:

- "Shrink and shift" + "Ramaswami transform" lead to an algorithm similar to that of Ramaswami (1999) of cost $(68/3)n^3$ ops per step (`ss-ram`).
- "Cayley transform" + "UL-based reduction" lead to SDA, having a cost $(64/3)n^3$ per step (`sda`).

Different combinations → different algorithms

We may combine the different strategies, for instance:

- "Shrink and shift" + "Ramaswami transform" lead to an algorithm similar to that of Ramaswami (1999) of cost $(68/3)n^3$ ops per step (`ss-ram`).
- "Cayley transform" + "UL-based reduction" lead to SDA, having a cost $(64/3)n^3$ per step (`sda`).
- "Shrink and shift" + "UL-based reduction" lead to a new algorithm with the same cost of SDA. Formally, this algorithm differs from SDA only for the initial values, which are simpler (`ss-ul`).

Different combinations → different algorithms

We may combine the different strategies, for instance:

- "Shrink and shift" + "Ramaswami transform" lead to an algorithm similar to that of Ramaswami (1999) of cost $(68/3)n^3$ ops per step (`ss-ram`).
- "Cayley transform" + "UL-based reduction" lead to SDA, having a cost $(64/3)n^3$ per step (`sda`).
- "Shrink and shift" + "UL-based reduction" lead to a new algorithm with the same cost of SDA. Formally, this algorithm differs from SDA only for the initial values, which are simpler (`ss-ul`).
- "Cayley transform" + "Small-size transform" lead to a new algorithm, having a cost $(38/3)n^3$ (`nodoub`).

UNIVERSITÀ DI PISA

NARE deriving from a problem in neutron transport theory

$$A = \widehat{\Delta} - eq^T, \quad B = ee^T, \quad C = qq^T, \quad D = \Delta - qe^T$$

with

$$\begin{aligned} \Delta &= \mathrm{diag}(\delta_1, \ldots, \delta_n), & \widehat{\Delta} &= \mathrm{diag}(\hat{\delta}_1, \ldots, \hat{\delta}_n), \\ \delta_i &= \frac{1}{cx_i(1-\alpha)}, & \hat{\delta}_i &= \frac{1}{cx_i(1+\alpha)}, \quad i = 1, \ldots, n, \\ e &= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T, & q_i &= \frac{w_i}{2x_i}, \quad i = 1, \ldots, n, \end{aligned}$$

$(x_i)_{i=1}^n$ and $(w_i)_{i=1}^n$ being the nodes and weights of a Gaussian discretization. Here we have chosen $\alpha = 10^{-8}$, $c = 1 - 10^{-6}$, which yields a close-to-null-recurrent Riccati equation.

Running time in seconds

n	sda	ss-ul	ss-ram	nodoub
8	0.045209	0.02735	0.030078	0.027061
16	0.039896	0.041282	0.046027	0.03845
32	0.14559	0.14666	0.18047	0.13432
64	0.92806	0.93415	1.1707	0.8448
128	7.2632	7.3491	9.1974	6.6499
256	60.841	61.926	76.835	55.03
512	499.95	504.37	625.06	448.46

Residual errors

n	sda	ss-ul	ss-ram	nodoub
8	1.654e-13	5.8367e-14	6.6482e-14	1.4294e-11
16	1.328e-12	2.4418e-13	2.7769e-13	1.6405e-10
32	3.4631e-12	1.964e-12	1.7786e-12	7.8717e-10
64	2.2679e-11	1.3598e-11	8.2769e-12	7.8282e-09
128	1.3316e-10	8.1521e-11	6.4269e-11	5.4047e-08
256	1.0096e-09	5.6852e-10	3.7115e-10	4.5315e-07
512	6.7923e-09	4.2861e-09	1.7767e-09	5.4083e-06

Conclusions and open issues

- The interpretation provided in this talk casts new light on the SDA algorithm and on the relationship between UQMEs and NAREs.
- Several other approaches to the solution of the NARE can be developed with this new setting. Among the possible ideas:
  - using numerical integration and the Cauchy integral theorem for computing the matrix 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.
- It would be important to find for more general transformations which map a Hamiltonian matrix H to a new one $\widetilde{H}$ where the block $\widetilde{H}_{1,2}$ is not only nonsingular but numerically well conditioned.