Some recent results concerning algebraic Riccati equations

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

MIMS New Directions Workshop
Functions of Matrices
Manchester, May 15-16, 2008

Università di Pisa

Nonsymmetric Algebraic Riccati Equations

NARE $\longrightarrow$ UQME

- Idea
- Moving the eigenvalues
- From NARE to UQME
- Theoretical and computational consequences

Conclusions and open issues

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

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

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

Assumptions:

Let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with

$$\mathrm{Re}(\lambda_{m+n}) \leq \mathrm{Re}(\lambda_{m+n-1}) \leq \ldots \leq \mathrm{Re}(\lambda_2) \leq \mathrm{Re}(\lambda_1).$$

Assumptions:

Let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with

$$\mathrm{Re}(\lambda_{m+n}) \leq \mathrm{Re}(\lambda_{m+n-1}) \leq \ldots \leq \mathrm{Re}(\lambda_2) \leq \mathrm{Re}(\lambda_1).$$

Assume that

$$\mathrm{Re}(\lambda_{n+1}) < 0 < \mathrm{Re}(\lambda_n) \tag{2}$$

and that there exists a solution S to the NARE such that $\sigma(D - CS) = \{\lambda_1, \ldots, \lambda_n\}$.
The solution S is called the extremal solution.

Our goal is the computation of S

Assumptions:

Let $\sigma(H) = \{\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\}$, with

$$\text{Re}(\lambda_{m+n}) \leq \text{Re}(\lambda_{m+n-1}) \leq \ldots \leq \text{Re}(\lambda_2) \leq \text{Re}(\lambda_1).$$

Assume that

$$\text{Re}(\lambda_{n+1}) < 0 < \text{Re}(\lambda_n) \tag{2}$$

and that there exists a solution S to the NARE such that $\sigma(D - CS) = \{\lambda_1, \ldots, \lambda_n\}$.
The solution S is called the extremal solution.

Our goal is the computation of S

Remark: Most of the result that we will show, are still valid if one of the two "$<$"'s in (2) is replaced by a "$\leq$"; the case $\text{Re}(\lambda_{n+1}) = 0 = \text{Re}(\lambda_n)$ requires some additional assumption.

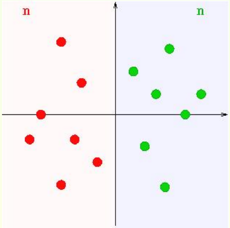
An example: matrix square root

If B is a square matrix having no real negative eigenvalues, then the principal matrix square root of B is $B^{1/2} = -S$, where S is the solution of $X^2 - B = 0$ such that $\sigma(-S) = \{\lambda_1, \ldots, \lambda_n\}$.

An example: matrix square root

If B is a square matrix having no real negative eigenvalues, then the principal matrix square root of B is $B^{1/2} = -S$, where S is the solution of $X^2 - B = 0$ such that $\sigma(-S) = \{\lambda_1, \ldots, \lambda_n\}$.

Location of the eigenvalues of $H = \begin{bmatrix} 0 & I \\ B & 0 \end{bmatrix}$:



Università di Pisa

An example: fluid queues and transport problems

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

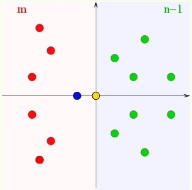
is either a nonsingular M-matrix or a singular irreducible M-matrix.

An example: fluid queues and transport problems

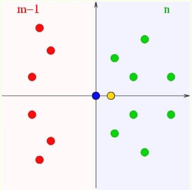
$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}$$

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

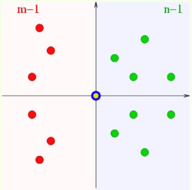
Location of the eigenvalues of H in the singular case:



Positive recurrent



Transient



Null recurrent
(Critical case)

Università di Pisa

From NARE to UQME: idea

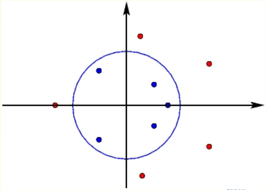
To transform the NARE into a Unilateral Quadratic Matrix Equation (UQME) of the kind

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

with $N \leq m+n$, such that:

1. $\det(A_0 + A_1\lambda + A_2\lambda^2)$ has roots

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



2. there exists a solution G with $\rho(G) = |\xi_N|$
3. from G one may easily recover S.

Moving the eigenvalues

Theorem

Let $f(z)$ be a complex function, analytic in a region containing the eigenvalues of H. If

$$H \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} R,$$

then

$$f(H) \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} f(R).$$

Moving the eigenvalues

Theorem

Let $f(z)$ be a complex function, analytic in a region containing the eigenvalues of H. If

$$H \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} R,$$

then

$$f(H) \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} f(R).$$

Consequence: the matrix $\widehat{H} = f(H)$ defines a new NARE, having the same solutions of the original NARE.

Moving the eigenvalues

Theorem

Let $f(z)$ be a complex function, analytic in a region containing the eigenvalues of H. If

$$H\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} R,$$

then

$$f(H)\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} f(R).$$

Consequence: the matrix $\widehat{H} = f(H)$ defines a new NARE, having the same solutions of the original NARE.

Remark: the theorem is still valid under weaker assumptions on $f(z)$

Cayley transform

The Cayley transform $f(z) = (z - \gamma)/(z + \gamma)$ applied to H yields the matrix

$$H_\gamma = f(H) = \begin{bmatrix} D + \gamma I & -C \\ B & -A + \gamma I \end{bmatrix}^{-1} \begin{bmatrix} D - \gamma I & -C \\ B & -A - \gamma I \end{bmatrix}.$$

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.

Shrink and shift

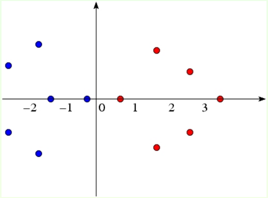
The “shrink-and-shift” function $f(z) = 1 + tz$ applied to H yields the matrix

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

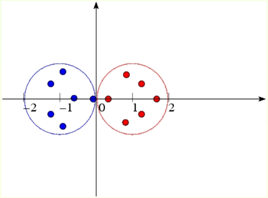
If M is an M-matrix and if $0 < t < 1/\max(a_{i,i}, d_{i,i})$ the eigenvalues of the matrix have a splitting w.r.t. the unit circle.

For a general H other conditions on t may be found.

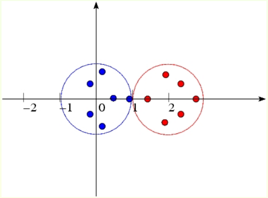
Shrink and shift



Original eigenvalues



Shrink by t



Shift by 1

NARE→UQME: 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

$$\begin{bmatrix} D & 0 \\ B & 0 \end{bmatrix} + \begin{bmatrix} -I & -C \\ 0 & -A \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} Y^2 = 0 \qquad (3)$$

Theorem

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

- m *equal to* 0
- *the* $m+n$ *eigenvalues* $\lambda_1, \dots, \lambda_{m+n}$ *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* (3) *with* m *eigenvalues equal to zero and* n *eigenvalues equal to* $\lambda_1, \dots, \lambda_n$*.*

NARE→UQME: *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

$$L - \lambda U.$$

NARE→UQME: *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

$$\begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I & U_1 \\ L_2 & -I \end{bmatrix} Y + \begin{bmatrix} 0 & 0 \\ 0 & U_2 \end{bmatrix} Y^2 = 0 \qquad (4)$$

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* (4) *with* m *eigenvalues equal to zero and* n *eigenvalues equal to* $\lambda_1, \ldots, \lambda_n$.

NARE→UQME: “Small size” transform

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

$$\begin{bmatrix} I & 0 \\ -U & I \end{bmatrix} H \begin{bmatrix} I & 0 \\ -U & I \end{bmatrix}^{-1} - \lambda I. \qquad (5)$$

If $\det C \neq 0$, by choosing $U = C^{-1}D$, (5) 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

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

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 $\lambda_1, \dots, \lambda_n$*.*

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

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

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

A few remarks

- In all these transformations the $m+n$ eigenvalues of H are roots of $\det A(\lambda)$.
- If we replace H with $f(H)$, then $\det A(\lambda)$ has $m+n$ roots inside the unit disk, and $m+n$ roots outside (including the ones at infinity). Moreover, the solution of the UQME with smallest spectral radius is the one of interest.

A few remarks

- In all these transformations the $m+n$ eigenvalues of H are roots of $\det A(\lambda)$.
- If we replace H with $f(H)$, then $\det A(\lambda)$ has $m+n$ roots inside the unit disk, and $m+n$ roots outside (including the ones at infinity). Moreover, the solution of the UQME with smallest spectral radius is the one of interest.

Different combinations of "eigenvalue transformations $f(z)$" with "NARE→UQME reductions" → old and new theoretical/computational properties

A theoretical result

Theorem

Let $\mathcal{Q}(\lambda) = \lambda^{-1}A(\lambda)$, where $A(\lambda)$ is obtained by applying the Cayley transform and the UL based reduction. Then:

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

A theoretical result

Theorem

Let $\mathcal{Q}(\lambda) = \lambda^{-1} A(\lambda)$, where $A(\lambda)$ is obtained by applying the Cayley transform and the UL based reduction. Then:

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

The Structure-preserving Doubling Algorithm is Cyclic Reduction!

SDA applied to the NARE is CR applied to the UQME

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

obtained with "Cayley transform + UL-based reduction".

Different combinations + CR → different algorithms

We may combine the different strategies, and apply CR for solving the UQME:

- “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 + CR → different algorithms

We may combine the different strategies, and apply CR for solving the UQME:

- "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 + CR → different algorithms

We may combine the different strategies, and apply CR for solving the UQME:

- "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 + CR → different algorithms

We may combine the different strategies, and apply CR for solving the UQME:

- "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`).

Numerical results: a random NARE

M is a randomly chosen singular M-matrix M, generated using Matlab's commands `R=rand(2*n)`, `M=diag(R*ones(2*n,1))-R`. The reported values are the average of ten different choices of the random matrix.

n	sda	ss-ul	ss-ram	nodoub
8	0.016927	0.015696	0.017	0.015045
16	0.028276	0.028877	0.032625	0.026565
32	0.083346	0.084624	0.099644	0.07022
64	0.48015	0.48756	0.58651	0.38958
128	4.8408	4.8895	6.1907	3.9967
256	34.036	34.497	40.72	26.933
512	291.47	295.6	354.06	228.18

Table: Running time in seconds

n	sda	ss-ul	ss-ram	nodoub
8	4.3812e-15	3.8386e-15	2.8644e-15	2.8628e-11
16	1.3656e-14	1.0136e-14	6.8251e-15	9.1426e-11
32	3.8594e-14	2.3889e-14	1.8441e-14	3.2387e-10
64	1.1038e-13	6.2969e-14	4.6679e-14	2.5328e-08
128	3.6836e-13	1.5803e-13	1.2221e-13	6.6213e-09
256	1.0805e-12	4.3243e-13	3.3097e-13	5.6768e-10
512	3.2239e-12	1.1668e-12	9.0803e-13	3.8776e-10

Table: Absolute residual

A new iteration for the matrix square root

"Shrink-and-shift + Small-size transform" applied to $H = \begin{bmatrix} 0 & I \\ B & 0 \end{bmatrix}$

lead to the UQME

$$Y^2 + 2Y + (I - t^2 B) = 0$$

such that the minimal solvent is $Y = I + tB^{1/2}$.

CR can be applied to solve the UQME.

Comparisons with the existing methods are still to be performed

A problem from transport theory

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

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

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

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

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

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

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

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

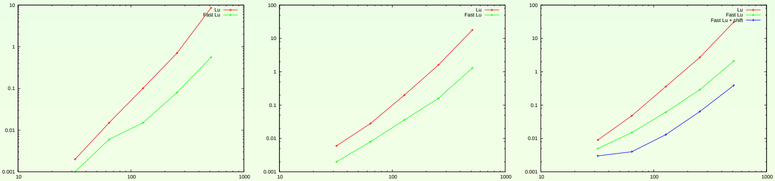
$$\mathcal{D}\psi(\lambda) - \psi(\lambda)\mathcal{D} = \mathbf{u}_1(\lambda)\mathbf{v}_1 + \mathbf{u}_2\mathbf{v}2(\lambda)$$

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 $A_k(\lambda)$ 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.
- 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.

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 $\approx 10^{-8}$)

Conclusions and open issues

- The interpretation provided in this talk casts new light on the relationship between UQMEs and NAREs, and on the SDA algorithm.
- 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.

Conclusions and open issues (cont.)

- An analysis of the accuracy and efficiency of the algorithms obtained with the different combinations is still to be performed.
- Are there other nice functions, besides Cayley and "Shrink-and-shift", to move the eigenvalues?
- It would be important to find more general transformations which map the matrix H to a new one $\widetilde{H}$ where the block $\widetilde{H}_{1,2}$ is not only nonsingular but numerically well conditioned.
- Can we apply our approach to other matrix equations, whose solutions is expressed by means of an invariant subspace?