Two numerical methods for the solution of Lur'e equations

Federico Poloni[1] Timo Reis[2]

[1]Technische Universität Berlin, Institut für Mathematik

Supported by a postdoctoral grant of the Alexander von Humboldt Foundation

[2]Technische Universität Hamburg–Harburg, Institut für Numerische Simulation

International conference on Matrix Methods in Mathematics and Applications
Moscow, 22–25 June 2011

Control problems and even matrix pencils

(Continous-time) control problems can be naturally expressed as deflating subspace problems for

Even matrix pencils

$$\mathcal{A} - s\mathcal{E} = \begin{bmatrix} 0 & A - sI & B \\ A^* + sI & Q & S \\ B^* & S^* & R \end{bmatrix} \qquad \mathcal{A}, \mathcal{E} \in \mathbb{R}^{n+n+m,n+n+m}$$

$\mathcal{A} - s\mathcal{E}$ is even, i.e., $\mathcal{A} = \mathcal{A}^*$, $\mathcal{E} = -\mathcal{E}^*$

We are looking for a maximal $\mathcal{E}$-neutral deflating subspace, i.e.,

$$\mathcal{A}U = V\widehat{\mathcal{A}} \qquad \mathcal{E}U = V\widehat{\mathcal{E}} \qquad U, V \in \mathbb{C}^{2n+m,k} \qquad U^*\mathcal{E}U = 0$$

Moreover, $\widehat{\mathcal{A}} - s\widehat{\mathcal{E}}$ semi-stable (or semi-unstable).

From control problems to Riccati equations (sometimes)

When R nonsingular, eliminate the last block $\Longrightarrow$ invariant subspace problem for a Hamiltonian matrix

$$\begin{bmatrix} A_R & -G_R \\ -Q_R & -A_R^* \end{bmatrix} - sI = \begin{bmatrix} A - BR^{-1}S^* & -BR^{-1}B^* \\ -Q + SR^{-1}S^* & -(A - BR^{-1}S^*)^* \end{bmatrix} - sI$$

Associated to an algebraic Riccati equation via

$$\begin{bmatrix} A_R & -G_R \\ -Q_R & -A_R^* \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} M \qquad M = A_R - G_R X$$

$\mathcal{E}$-neutrality condition becomes simplecticity (i.e., $X = X^*$, in this form)

Question

What if R is singular?

What if R is singular?

The singular R case has been treated stepmotherly (*T. Reis*)

- numerical problems: nontrivial Jordan blocks at infinity and/or singular pencil
- the Riccati equation cannot be formed
- in engineering practice, often solved by perturbing+inverting R

ARE must be replaced by a system

Lur'e equations

$$
\begin{aligned}
A^T X + XA + Q &= Y^T Y \\
XB + S &= Y^T Z \\
R &= Z^T Z
\end{aligned}
$$

(only X needed in practice)

Lur'e equations and deflating subspaces

Deflating subspace formulation

$$\begin{bmatrix} 0 & -sI + A & B \\ sI + A^* & Q & S \\ B^* & S^* & R \end{bmatrix} \begin{bmatrix} X & 0 \\ I_n & 0 \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -X & Y^* \\ 0 & Z^* \end{bmatrix} \begin{bmatrix} -sI + A & B \\ Y & Z \end{bmatrix}$$

Since $\mathcal{E} = \begin{bmatrix} 0 & I_n & 0 \\ -I_n & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\ker \mathcal{E} = \begin{bmatrix} 0 \\ 0 \\ I_m \end{bmatrix}$ contains the eigenvectors at ∞

When R is singular, some of them start nontrivial Jordan chains
the other vectors in the chains aren't as easy to find

Even Kronecker canonical form

Even Kronecker canonical form [Thompson, '76 & '91] , a powerful tool to analyze Lur'e equations theoretically [Reis, '11]

Canonical form under transformations of the kind $M^T \mathcal{A} M$, $M^T \mathcal{E} M$ (for any M nonsingular)

Plays well with

- deflating subspaces $M^T(\mathcal{A} - s\mathcal{E})M\, M^{-1}U = M^T V(\widehat{\mathcal{A}} - \widehat{\mathcal{E}})$
- $\mathcal{E}$-neutrality $U^T M^{-T} M^T \mathcal{E} M\, M^{-1} U = 0$ (and similar relations)

Even Kronecker canonical form [Thompson, '76 & '91]

Every even matrix pencil (i.e., $\mathcal{A} = \mathcal{A}^*$, $\mathcal{E} = -\mathcal{E}^*$) can be reduced to a direct sum of the following block types. . .

Even Kronecker canonical form

$$\left[\begin{array}{ccc|ccc} & & & \lambda - s & 1 & \\ & & & & \lambda - s & 1 \\ & & & & & \lambda - s \\ \hline \bar{\lambda} + s & & & & & \\ 1 & \bar{\lambda} + s & & & & \\ & 1 & \bar{\lambda} + s & & & \end{array}\right]$$

paired eigenvalues $(\lambda, -\bar{\lambda})$

$$\left[\begin{array}{cc|ccc} & & & s & 1 \\ & & s & 1 & \\ \hline & s & & & \\ s & 1 & & & \\ 1 & & & & \end{array}\right]$$

singular blocks

$$\left[\begin{array}{cccc} & & & i\mu - s \\ & & i\mu - s & 1 \\ & i\mu - s & 1 & \\ i\mu - s & 1 & & \end{array}\right]$$

imaginary eigenvalues $i\mu$

$$\left[\begin{array}{ccccc} & & & s & 1 \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{array}\right]$$

eigenvalues at ∞

What comes out of the EKCF

Theorem [Reis '11]

Lur'e eqns solvable iff

- all imaginary blocks have even size
- all infinite blocks have odd size

Moreover, third block of $\begin{bmatrix} 0 & A - sI & B \\ A^* + sI & Q & S \\ B^* & S^* & R \end{bmatrix} \Leftrightarrow \ker \mathcal{E} \Leftrightarrow$

$$\begin{bmatrix} & & & s & 1 \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{bmatrix}$$

∞ block

$$\left[\begin{array}{cc|cc} & & s & 1 \\ & & s\ \ 1 & \\ \hline & s & & \\ s & 1 & & \\ 1 & & & \end{array}\right]$$

singular block

Method I: Cayley transform

Some methods for solving Riccati equations go like this:

1. input: generic control problem with nonsingular R
2. obtain Hamiltonian pencil $\mathcal{H} - sI$
3. Cayley transform $(\mathcal{H} + I) - s(\mathcal{H} - I)$
4. left-multiply by a suitable M to enforce $\begin{bmatrix} * & 0 \\ * & I \end{bmatrix} - s \begin{bmatrix} I & * \\ 0 & * \end{bmatrix}$

   Besides, the $*$ blocks form a symmetric matrix (symplectic pencil)
5. solve symplectic subspace problem with method of choice

   e.g. [Fassbender '00 book] , [Chu, Fan, Lin '05]

Method I: Cayley transform

We can modify the workflow like this: [P. Reis, preprint]

1. input: generic control problem, R may be singular
2. Cayley transform $\begin{bmatrix} 0 & A+I & B \\ A^*+I & Q & S \\ B^* & S^* & R \end{bmatrix} - s \begin{bmatrix} 0 & A-I & B \\ A^*+I & Q & S \\ B^* & S^* & R \end{bmatrix}$
3. enforce $\begin{bmatrix} * & 0 & 0 \\ * & I & 0 \\ * & 0 & I \end{bmatrix} - s \begin{bmatrix} I & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$, in fact $\begin{bmatrix} * & 0 & 0 \\ * & I & 0 \\ * & 0 & I \end{bmatrix} - s \begin{bmatrix} I & * & 0 \\ 0 & * & 0 \\ 0 & * & I \end{bmatrix}$
4. The system is now block triangular, we may ignore the third block

   what's left is again a symplectic pencil $\begin{bmatrix} * & 0 \\ * & I \end{bmatrix} - s \begin{bmatrix} I & * \\ 0 & * \end{bmatrix}$
5. Symplectic solver of choice

Method I: where do the eigenvalues end up?

We know that third block $\Leftrightarrow \ker \mathcal{E} \Leftrightarrow$

$$\begin{bmatrix} & & & s & \boxed{1} \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{bmatrix}$$

- After Cayley, this goes to a $\lambda = 1$ Jordan block of the same size
- Deflation of the second triangular block: shorten every block by 1.

Method I: block sizes

After the Cayley transform,

- even size imaginary blocks $\mapsto$ even size unimodular blocks
- odd size $\lambda = \infty$ blocks $\mapsto$ odd size $\mathcal{C}(\lambda) = 1$ blocks
  but we reduce dimension by 1 for each, so they become even size

Even-size unimodular blocks $\Rightarrow$ solution of the symplectic problem exist, algorithms work fine

Method I in a nutshell

- Cayley-then-deflate, not deflate-then-Cayley!
- only for dense problems
- not as easy to handle singular pencils

Algorithm I — numerical experiments

Some examples from the CAREX benchmark set [Benner, Laub, Mehrmann '95] modified to get a singular R. Competitors:

- Method I + SDA [Chu, Fan, Lin '05]
- Regularization with different ε + SDA
- Regularization + Newton-Kleinman

CAREX #	I+SDA	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-8}$+N
3	6E-15	5E-2	5E-2	5E-2	9E-10
4	4E-15	6E-7	5E-9	1E-7	5E-9
5	2E-10	3E-7	1E-9	3E-8	1E-9
6	2E-15	6E-12	2E-13	1E-12	2E-13

Table: Final accuracy attained (lower=better)

Method II: compute and deflate

We wish to compute and deflate the subspace at infinity/singular

Relation defining chains at infinity/singular

$$\mathcal{E} v_1 = 0 \qquad \mathcal{A} v_k = \mathcal{E} v_{k+1}$$

First vectors of every chain: spanned by $\ker \mathcal{E} = \operatorname{span} \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}$

We'd like to extend the chains by computing the next vectors $v_2, v_3, \dots$ from v_1

But we cannot take $v_{k+1} = \mathcal{E}^{-1} \mathcal{A} v_k$ as $\mathcal{E}$ is singular

All we can get is $\mathcal{E}^{\dagger} \mathcal{A} v_k = v_{k+1} + w$, $w \in \ker \mathcal{E}$

Compute and deflate — infinite blocks

What we'd like to compute

$$\begin{array}{ccccc} & & \color{red}{v_3} & \color{blue}{v_2} & \color{green}{v_1} \\ & & \downarrow & \downarrow & \downarrow \\ \end{array}$$
$$\begin{bmatrix} & & & s & 1 \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{bmatrix}$$

What we actually get

$$\begin{array}{ccccc} & & \color{red}{w_3} & \color{blue}{w_2} & \color{green}{v_1} \\ & & \downarrow & \downarrow & \downarrow \\ \end{array}$$
$$\begin{bmatrix} & & & s & 1 \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{bmatrix}$$

$w_2 \in \operatorname{span}\{v_1, v_2\}$

$w_3 \in \operatorname{span}\{v_1, v_2, v_3\}$

Compute and deflate — when to stop

For both infinite and singular chain, we can stop computing at half of the chain length (only this is needed for the solution)

$$
\begin{array}{c}
\begin{array}{ccccc} & & \color{red}{w_3} & \color{blue}{w_2} & \\ & & \color{red}{\downarrow} & \color{blue}{\downarrow} & \end{array} \\
\begin{bmatrix} & & & s & 1 \\ & & s & 1 & \\ & s & 1 & & \\ s & 1 & & & \\ 1 & & & & \end{bmatrix}
\end{array}
$$

idea to find out when:

- $w_k^T \mathcal{A} w_k$ is preserved by canonical form
- $w_k^T \mathcal{A} w_k = 0$ until we hit the first half of the chain, $\neq 0$ afterwards

Compute and deflate — interaction between chains

$$
\begin{array}{c}
\begin{array}{ccccc} & & v_3 & v_2 & v_1 \end{array} \\
\begin{bmatrix}
 & & & s & 1 \\
 & & s & 1 & \\
 & s & 1 & & \\
s & 1 & & & \\
1 & & & &
\end{bmatrix}
\end{array}
\qquad\qquad
\begin{array}{c}
\begin{array}{ccc} y_3 & y_2 & y_1 \end{array} \\
\begin{bmatrix}
 & s & 1 \\
s & 1 & \\
1 & &
\end{bmatrix}
\end{array}
$$

We get instead

$$w_2 \in \operatorname{span}(v_1, v_2, y_1), \qquad\qquad z_2 \in \operatorname{span}(y_1, y_2, v_1)$$

blocks get mixed because recursion is $\mathcal{E}^\dagger \mathcal{A} v_k = v_{k+1} + w$, $w \in \ker \mathcal{E}$

Moreover, we don't get w_2 and z_2, but $\operatorname{span}(w_2, z_2)$ and so on

Further tricks

Need further tricks (won't mention them here) to handle singular blocks

$$\left[\begin{array}{cc|ccc} & & & s & 1 \\ & & s & 1 & \\ \hline & s & & & \\ s & 1 & & & \\ 1 & & & & \end{array}\right] \qquad v_1 \in \ker \mathcal{E}, \quad \mathcal{E} v_2 = \mathcal{A} v_1, \quad \mathcal{E} v_3 = \mathcal{A} v_2, \quad 0 = \mathcal{A} v_3$$

and to make sure that infinite chains that have already reached half-chain won't creep back in after deflation (due to terms in $\ker \mathcal{E}$)

But this procedure can be carried on with success

Crucial point

We have to take rank decisions: what happens if the singular values drop smoothly?

Sparse Lur'e equations

The procedure can be carried on for sparse matrices (assuming the infinite/singular space is small) yielding a tall skinny W with a basis for the infinite/singular space

Now, let $\Pi = I - WW^T$ complementary projector

$$\Pi \begin{bmatrix} 0 & A - sI & B \\ A^* + sI & Q & S \\ B^* & S^* & R \end{bmatrix} \Pi^T \cong \begin{bmatrix} 0 & A_1 - sI & B_1 & 0 \\ A_1^* + sI & Q_1 & S_1 & 0 \\ B_1^* & S_1^* & R_1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

R_1 nonsingular, can turn this (implicitly) into a Riccati equation

Key point

Keeping its coefficients in a form that is sparsely representable

I heard you like iterations, so...

Nested iterations

- projected Riccati equation is solved using Newton
- Lyapunov equations in Newton are solved using ADI (Lyapack [Penzl, '99])
- singular linear systems in ADI are solved using iterative methods
  - preconditioning isn't straightforward, matrices are represented as products

We only get good results if all these iterations behave reasonably

Method II — numerical experiments

Lur'e equations derived from a test problem in Lyapack

	`demo-r3`
n	821
m	6
rank decisions accuracy	6.5×10^{-16}
infinite chains	$6\times$ length 3
singular chains	0
no. of Newton steps taken	7
avg. ADI itns per Newton step	35
relative residual	5.5×10^{-15}
deviation from stability	-8.3×10^{-09}

Method II — numerical experiments

Lur'e equations derived from a test problem in Lyapack

	`demo-r3`
n	821
m	6
rank decisions accuracy	6.5×10^{-16}
infinite chains	$6\times$ length 3
singular chains	0
no. of Newton steps taken	7
avg. ADI itns per Newton step	35
relative residual	5.5×10^{-15}
deviation from stability	-8.3×10^{-09}

Thanks for your attention