A STRUCTURE-PRESERVING METHOD FOR GENERALIZED ALGEBRAIC RICCATI EQUATIONS BASED ON PENCIL ARITHMETIC

R. Byers[*], P. Benner[†]

[*] Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA, `byers@math.ukans.edu`
[†] Institut für Mathematik, Sekretariat MA 4-5, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany, `benner@math.uni-bremen.de`

Keywords: algebraic Riccati equation, linear-quadratic regulator, H_2/H_∞-control, sign function method, spectral projection method.

Abstract

This paper describes a numerical method for extracting the stable right deflating subspace of a matrix pencil $Z - \lambda Y$ using a spectral projection method. It has several advantages compared to other spectral projection methods like the sign function method. In particular it avoids the rounding error induced loss of accuracy associated with matrix inversions. The new algorithm is particularly well adapted to solving continuous-time algebraic Riccati equations. In numerical examples, it solves Riccati equations to high accuracy.

1 Introduction

One of the most important computational tool in control design is the numerical solution of (generalized) continuous-time algebraic Riccati equations (CAREs) of the form

$$0 = \mathcal{R}(X) = Q + A^T XE + E^T XA - E^T XGXE, \quad (1)$$

where $A, E, G, Q \in \mathbb{R}^{n\times n}$, $G = G^T$, $Q = Q^T$, and $X = X^T \in \mathbb{R}^{n\times n}$ is the sought-after solution. It arises in the computation of linear-quadratic regulators, optimal H_2- and H_∞-controllers, model reduction based on stochastic or positive real balancing, finding equilibria in differential games, see, e.g., [2, 3, 19, 28, 31, 32]. In all these applications, a particular solution is desired which has the property that $\lambda E - (A - GXE)$ is a stable matrix pencil in the sense that all its eigenvalues lie in the open left half complex plane. (Assuming that E is nonsingular, all eigenvalues of the matrix pencil are finite).

A classical approach to solving the CARE (1) is to compute the stable right deflating subspace of the corresponding Hamiltonian/skew-Hamiltonian matrix pencil

$$H - \lambda K := \begin{bmatrix} A & G \\ Q & -A^T \end{bmatrix} - \lambda \begin{bmatrix} E & 0 \\ 0 & E^T \end{bmatrix}. \quad (2)$$

(The stable right deflating subspace is the right deflating subspace corresponding to eigenvalues in the open left half complex plane.) Under suitable assumptions typically satisfied in the control problems mentioned above, $H - \lambda K$ has exactly n eigenvalues contained in the open left half plane. It is well known that if the columns of $\left[\begin{smallmatrix} U \\ V \end{smallmatrix}\right] \in \mathbb{R}^{2n\times n}$ form a basis for the corresponding n-dimensional stable right deflating subspace, then $X = -VU^{-1}E^{-1}$ is the required stabilizing solution of the CARE (1).

There are many numerical methods for solving (1). Here we will focus on spectral projection methods which have been used successfully for solving many computational problems in control theory. The matrix sign function is a popular method for computing projectors onto the stable invariant subspace of a matrix Z [29] or onto the stable right deflating subspace of a regular matrix pencil $Z - \lambda Y$ [17]. See [22] for a survey of the theoretical and computational aspects of the sign function. The most frequently used iteration employed by the sign function method is the (generalized) sign-Newton iteration [17] given by

$$Z_0 \leftarrow Z \quad (3)$$

$$c_k \leftarrow \left| \frac{\det(Z_k)}{\det(Y)} \right|^{1/n} \quad (4)$$

$$Z_{k+1} \leftarrow \frac{1}{2c_k}\left(Z_k + c_k^2 Y Z_k^{-1} Y\right) \quad (5)$$

(For the matrix case, set $Y = I$ in (5).) If both Y and Z are nonsingular and $Z - \lambda Y$ has no eigenvalues on the imaginary axis, then it can be shown that Z_k and Y_k are nonsingular for all k, $Z_\infty := \lim_{k\to\infty} Z_k$ exists, $(I - Y^{-1}Z_\infty)/2$ is the projector onto the stable right deflating subspace of $Z-\lambda Y$ parallel to the anti-stable right deflating subspace, and $(I + Y^{-1}Z_\infty)/2$ is the projector onto the anti-stable subspace parallel to the stable deflating subspace. Thus, a basis for the stable deflating subspace of $Z - \lambda Y$ as required when solving (1) can be obtained from the null space of $Z_\infty + Y$. The method has been proved both theoretically and numerically efficient and accurate for problems with spectra well separated from the imaginary axis and well conditioned matrices Y and Z [5, 6, 13, 15].

The scalar c_k in (4) is a parameter chosen to accelerate convergence. The particular choice used in (4) is a generalization proposed in [17] of determinantal scaling [14]. There are many other possibilities for the acceleration parameter [9, 16, 21, 29].

A weakness of the iteration (5) is that inverses have to be formed either explicitly or implicitly by solving linear systems. If any of the Z_k's in (5) is ill-conditioned with respect to inversion, then a severe loss of accuracy is possible. The following example demonstrates this effect.

Example 1 Construct a pencil $Z - \lambda Y_p$ as follows. Let B_p be the 10-by-10 Jordan block with eigenvalue $1/p$; let K be the 10-by-10 matrix with $(1, 1)$ entry equal to one and all other entries equal to zero; let W be the 10-by-10 matrix all of whose entries are one; and let U be the 10-by-10 elementary reflector $U = I - 0.2 \cdot W$. Construct $Z - \lambda Y_p$ as the 20-by-20 Hamiltonian/skew-Hamiltonian pencil pencil

$$Z = \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} I - 2K & K \\ I - K & 2K - I \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} \quad (6)$$

$$Y_p = \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} B_p & 0 \\ 0 & B_p^T \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}. \quad (7)$$

Note that this matrix pencil has exactly the structure of (2) corresponding to a generalized algebraic Riccati equation. The eigenvalues are $\pm 1/p$ each with algebraic multiplicity 10 and geometric multiplicity 2. The stable and unstable deflating subspaces grow increasingly ill-conditioned as p increases from $p = 1$ to $p = 7$. In addition, as p varies from $p = 1$ to $p = 7$, the condition number of Y_p varies from 10^1 to 10^8.

We calculated orthonormal bases of the 10-dimensional stable right deflating subspace using the QZ algorithm [27] and as the null space of $Z + Y_{p,\infty}$ obtaining $Y_{p,\infty}$ from the generalized sign-Newton iteration (3)–(5). (The computations were run under MATLAB version 6 [25] on a workstation with unit round approximately 2.22×10^{-16}.) This produced orthonormal bases $V_{p,\mathrm{qz}}$ and $V_{p,\mathrm{gl}}$ of rounding-error-corrupted approximate deflating subspaces from the QZ algorithm and the generalized sign-Newton iteration (3)–(5) respectively. The example is simple enough to be able to calculate an exact, analytic orthonormal basis V_p of the right stable deflating subspace. The forward or absolute errors are $\|V_{p,\mathrm{qz}} V_{p,\mathrm{qz}}^H - V_p V_p^H\|_F$ and $\|V_{p,\mathrm{gl}} V_{p,\mathrm{gl}}^H - V_p V_p^H\|_F$. If $\sigma_1, \sigma_2, \sigma_3, \ldots \sigma_{20}$ are the 20 singular values of $[ZV_{p,\mathrm{qz}}, Y_p V_{p,\mathrm{qz}}]$ or $[ZV_{p,\mathrm{gl}}, YV_{p,\mathrm{gl}}]$, then the respective backward errors are $(\sigma_{n+1}^2 + \sigma_{n+2}^2 + \ldots + \sigma_{2n}^2)^{1/2}$. The backward error is the magnitude of the smallest Frobenius norm perturbation of Y_p and Z which yields a pencil for which the computed deflating subspace is an exact deflating subspace. Table 1 lists these forward and backward errors for $p = 1, 2, \ldots 7$. The table demonstrates how ill-conditioned Z_k in (5) can adversely affect both forward and backward errors. Note particularly the backward errors in comparison with the expensive but backward stable QZ algorithm. For $p \geq 3$ the iterates Z_k in (5) are so ill-conditioned that our program failed to meet its stopping criterion $\|Z_{k+1} - Z_k\|_F \leq n^2 \varepsilon \|Z_{j+1}\|_F$ where ε is the machine precision 2.22×10^{-16}. In that case, we terminated the program after 50 iterations. For $p \geq 4$, many iterates had condition numbers larger than 10^{14}.

Forward Errors

p	QZ	(3)–(5)	Inverse-Free
1	10^{-15}	10^{-15}	10^{-15}
2	10^{-13}	10^{-12}	10^{-13}
3	10^{-9}	10^{-9}	10^{-10}
4	10^{-7}	10^{-7}	10^{-8}
5	10^{-5}	10^{-3}	10^{-7}
6	10^{-4}	10^{-1}	10^{-5}
7	10^{-3}	10^{-1}	10^{-4}

Backward Errors

p	QZ	(3)–(5)	Inverse-Free
1	10^{-15}	10^{-15}	10^{-15}
2	10^{-15}	10^{-13}	10^{-14}
3	10^{-15}	10^{-10}	10^{-11}
4	10^{-15}	10^{-8}	10^{-9}
5	10^{-15}	10^{-5}	10^{-8}
6	10^{-15}	10^{-3}	10^{-7}
7	10^{-15}	10^{-2}	10^{-7}

Table 1: Rounding error induced forward and backward errors in the computed stable deflating subspace of $Z - \lambda Y_p$ given by (7) and (6).

2 Inverse-Free Methods

To overcome the problem with inverses in (5), inverse-free methods have been investigated. In particular the inverse-free spectral divide and conquer method [7, 24] has received some attention in recent years. It can be considered as an instance of the disk function method [10, 11]. The method computes spectral projectors onto the deflating subspaces corresponding to eigenvalues inside and outside the unit circle. Hence it can be used to solve the CARE (1) by applying it to the Cayley-transform of the matrix pencil (2), $Z - \lambda Y = (H - K) - \lambda(H + K)$. Unfortunately, the iteration described in [7, 24] does more than twice the amount of floating point arithmetic than (5).

We propose a new inverse-free iteration scheme that computes the projector onto the stable invariant subspace of a matrix or the stable right deflating subspace of a matrix pencil without the need to compute a Cayley transformation. It also allows the use of scaling to accelerate convergence. The computational cost of the new method is comparable to that of the disk function method [7, 24] described above, though still being somewhat higher than that of (5).

The generalized sign-Newton iteration (3)–(5) preserves both the left and right deflating subspaces of $Z - \lambda Y = Z_0 - \lambda Y$. Most CARE (1) applications require only the right stable deflating subspace. The left deflating subspaces are not needed. This suggests that one might be able to avoid some of the hazards of matrix inversion by replacing the sequence of pencils generated by (5) with another sequence having the same right deflating subspaces but possibly different left deflating subspaces.

Call a sequence of pencils $\hat{Z}_k - \lambda \hat{Y}_k$ a *right handed sign-Newton sequence* (RHSNS) if there is a sequence of nonsingular matrices M_k for which

$$\hat{Z}_k - \lambda\hat{Y}_k = M_k Z_k - \lambda M_k Y \tag{8}$$

where $Z_k - \lambda Y$ satisfies (3)–(5). (Of course, $M_k = \hat{Y}_k Y^{-1} = \hat{Z}_k Z_k^{-1}$.) A RHSNS has the same eigenvalues and right deflating subspaces as $Z_k - \lambda Y$ in (5), but it may have different left deflating subspaces. The eigenvalues and Kronecker canonical form of a RHSNS has the same convergence properties as (5) although the individual matrices $\hat{Z}_k$ and $\hat{Y}_k$ may or may not converge to a limit. If $\hat{Z}_\infty = \lim_{k\to\infty}\hat{Z}_k$ and $\hat{Y}_\infty = \lim_{k\to\infty}\hat{Y}_k$ exist, then the stable right deflating subspace of $Z - \lambda Y$ is the null space of $\hat{Z}_\infty + \hat{Y}_\infty$ and the anti-stable right deflating subspace is the null space of $\hat{Z}_\infty - \hat{Y}_\infty$. (Even if $\hat{Z}_k$ and/or $\hat{Y}_k$ do not converge, a practical numerical procedure might extract a good approximation to the stable right deflating subspace from an n-dimensional approximate null space of $\hat{Z}_k + \hat{Y}_k$ for large enough k.)

The following theorem shows how to generate a RHSNS without necessarily using an explicit inverse. It is based on an arithmetic for matrix pencils introduced and discussed in detail in [12].

Theorem 1 *If Y, $Z \in \mathbb{C}^{n\times n}$ are nonsingular then the following generates a RHSNS.*

$$\begin{aligned}
\hat{Z}_0 - \lambda\hat{Y}_0 &= (M_0 Z) - \lambda(M_0 Y) && (9)\\
\hat{c}_k &= \left|\frac{\det(\hat{Z}_k)}{\det(\hat{Y}_k)}\right|^{1/n} && (10)\\
\hat{Z}_{k+1} - \lambda\hat{Y}_{k+1} &= \alpha_k\left(\tilde{Y}_k\hat{Z}_k\right) && (11)\\
&\quad - \lambda\left(\frac{\alpha_k}{2}\right)\left(\hat{c}_k\tilde{Y}_k\hat{Y}_k + \hat{c}_k^{-1}\tilde{Z}_k\hat{Z}_k\right)
\end{aligned}$$

where $M_0 \in \mathbb{C}^{n\times n}$ is any nonsingular matrix, $\alpha_k \in \mathbb{C}$ is any nonzero scalar, and $\tilde{Y}_k$, $\tilde{Z}_k \in \mathbb{C}^{n\times n}$ are any matrices such that $\operatorname{rank}[\tilde{Y}_k,\ \tilde{Z}_k] = n$ and

$$\begin{bmatrix} \tilde{Y}_k & \tilde{Z}_k \end{bmatrix}\begin{bmatrix} -\hat{Z}_k \\ \hat{Y}_k \end{bmatrix} = 0. \tag{12}$$

Proof. Let $Z_k - \lambda Y$ be determined by the sign-Newton iteration (3)–(5) and let $\hat{Z}_k - \lambda\hat{Y}_k$ be any sequence of pencils satisfying (9)–(12). We will show that for all k, there exists a nonsingular matrix $M_k \in \mathbb{C}^{n\times n}$ satisfying (8). The proof is by induction on k.

Equation (8) holds for $k = 0$ by hypothesis (9). Assume that for some integer k, there exists a nonsingular matrix $M_k \in \mathbb{C}^{n\times n}$ satisfying (8). Observe first that in (10)

$$\begin{aligned}
\hat{c}_k &= \left|\frac{\det(\hat{Z}_k)}{\det(\hat{Y}_k)}\right|^{1/n}\\
&= \left|\frac{\det(M_k Z_k)}{\det(M_k Y)}\right|^{1/n}\\
&= \left|\frac{\det(Z_k)}{\det(Y)}\right|^{1/n}\\
&= c_k.
\end{aligned}$$

So $\hat{c}_k$ in (10) is equal to c_k in (4).

By induction hypothesis $\hat{Z}_k = M_k Z_k$ is a product of nonsingular matrices, so $\hat{Z}_k$ is nonsingular. It follows that $\operatorname{rank}\left[\begin{smallmatrix} -\hat{Z}_k \\ \hat{Y}_k \end{smallmatrix}\right] = n$, and all bases of the left null space of $\left[\begin{smallmatrix} -\hat{Z}_k \\ \hat{Y}_k \end{smallmatrix}\right]$ take the form $[\tilde{Y}_k,\ \tilde{Z}_k] = W_k[\hat{Y}_k\hat{Z}_k^{-1},\ I]$ for some nonsingular matrix $W_k \in \mathbb{C}^{n\times n}$. Hence,

$$\begin{aligned}
\hat{Y}_{k+1} &= \alpha_k\tilde{Y}_k\hat{Z}_k\\
&= \alpha_k W_k\hat{Y}_k\hat{Z}_k^{-1}\hat{Z}_k\\
&= \alpha_k W_k M_k Y
\end{aligned}$$

and

$$\begin{aligned}
\hat{Z}_{k+1} &= \frac{\alpha_k}{2}\left(\hat{c}_k\tilde{Y}_k\hat{Y}_k + \hat{c}_k^{-1}\tilde{Z}_k\hat{Z}_k\right)\\
&= \frac{\alpha_k}{2}\left(c_k W_k\hat{Y}_k\hat{Z}_k^{-1}\hat{Y}_k + c_k^{-1}W_k\hat{Z}_k\right)\\
&= \frac{\alpha_k}{2}\left(c_k W_k M_k Y Z_k^{-1}M_k^{-1}M_k Y + c_k^{-1}W_k M_k Z\right)\\
&= (\alpha_k W_k M_k)\left(\frac{1}{2c_k}\right)(c_k^2 Y Z_k^{-1} Y + Z).
\end{aligned}$$

Hence, with $M_{k+1} = \alpha_k W_k M_k$ the identity (8) is satisfied. □

There are many possible choices of M_0 in (9), $\tilde{Y}_k$, $\tilde{Z}_k$ in (12) and α_k in (11). As suggested in the proof, if $M_0 = I$, $\tilde{Y}_k = \hat{Y}_k\hat{Z}_k^{-1}$, $\tilde{Z}_k = I$ and $\alpha_k \equiv 1$, then (9)–(11) reduce to the generalized sign-Newton iteration (3)–(5).

If

$$\begin{bmatrix} -\hat{Z}_k \\ \hat{Y}_k \end{bmatrix} = \begin{bmatrix} Q_{11,k} & Q_{12,k} \\ Q_{21,k} & Q_{22,k} \end{bmatrix}\begin{bmatrix} R_k \\ 0 \end{bmatrix} \tag{13}$$

is a QR (unitary-triangular) factorization partitioned into n-by-n blocks, then a possible choice of $\tilde{Y}_k$ and $\tilde{Z}_k$ in (12) is $\tilde{Y}_k = Q_{12,k}^H$ and $\tilde{Z}_k = Q_{22,k}^H$. This choice does not require an explicit inverse. It is used in [7] for the disk function inverse-free algorithm. It is not clear whether it is possible to improve on (13) as a way to choose $\tilde{Y}_k$ and $\tilde{Z}_k$. A possible alternative appears in [12].

The choice of the scalar α_k is subtle. A poor choice of α_k leads to a RHSNS in which $\hat{Z}_k$ and/or $\hat{Y}_k$ diverge or converge to zero. For example, if $Y = Z = I$ and $\alpha_k \equiv 1$, then (13) may give $\tilde{Y}_k = \tilde{Z}_k = 2^{-1/2}I$. With these choices, for $k = 1, 2, 3, \ldots \hat{Y}_k = (\sqrt{2})^k I$, $\hat{Z}_k = (\sqrt{2})^k I$. Hence, $\lim_{k\to\infty} Y_k = \lim_{k\to\infty} Z_k = \infty$. If $\alpha_k \equiv 2$, then $\lim_{j\to\infty} Y_k = \lim_{j\to\infty} Z_k = 0$. Converging to zero is at least as problematic as diverging to ∞. Note that in this example, the sign-Newton iteration (3)–(5) is stationary. The Kronecker structure of $\hat{Z}_k - \lambda\hat{Y}_k$ is stationary, so a numerical procedure may stop immediately and obtain the stable right deflating subspace as the null space of $\hat{Z}_0 + \hat{Y}_0$ This is an extreme case. In more typical examples, the Kronecker structure (including eigenvalues) of $\hat{Z}_k - \lambda\hat{Y}_k$ converge so rapidly that one can stop a numerical procedure before a less-than-optimal choice of α_k causes numerical instability. If $\tilde{Y}_k$ and $\tilde{Z}_k$ are obtained from (13), then the example above shows that a necessary condition for convergence of the sequences $\hat{Z}_k$ and $\hat{Y}_k$ is $\alpha_k = \sqrt{2}$, see [12] for details.

Note that (13) determines $\tilde{Y}_k = Q_{12,k}^H$ and $\tilde{Z}_k = Q_{22,k}^H$ only up to right multiplication by an arbitrary n-by-n unitary factor. In order to assure that $\hat{Y}_k$ and $\hat{Z}_k$ converge, one can require that $\tilde{Y}_k = Q_{12,k}$ be triangular with positive diagonal entries. This choice leads to a particularly efficient numerical algorithm implementation which is described in detail in [12].

In summary the inverse-free sign function iteration can be described as follows.

1. Set $\hat{Z}_0 := Z$, $\hat{Y}_0 := Y$.
2. FOR $k = 0, 1, 2, \ldots$ until convergence
   i) Calculate matrices $Q_{12,k}$ and $Q_{22,k}$ satisfying (13). Set $\tilde{Y}_k := Q_{12,k}^H$ and $\tilde{Z}_k := Q_{22,k}^H$.
   ii) Set $\hat{c}_k := |\det(\hat{Z}_k)/\det(\hat{Y}_k)|^{1/n}$.
   iii) Set $\begin{array}{rcl} \hat{Z}_{k+1} & := & \frac{1}{\sqrt{2}}\left(\hat{c}_k^{-1}\tilde{Z}_k\hat{Z}_k + \hat{c}_k\tilde{Y}_k\hat{Y}_k\right), \\ \hat{Y}_{k+1} & := & \sqrt{2}\tilde{Z}_k\hat{Y}_k \end{array}.$

Like the sign-Newton algorithm (3)–(5), the right deflating subspaces of $Z - \lambda Y$ are preserved throughout the iteration. In particular, both the sign-Newton algorithm and the inverse-free sign function algorithm preserves any special structure that the right deflating subspaces may have. Linear quadratic and H_∞ optimal control problems [23] along with quadratic eigenvalue linear damping models [18] lead to invariant subspace problems whose right deflating subspaces are Lagrangian. This special structure is preserved in the inverse-free sign function algorithm.

3 Numerical Results

Example 1 continued. We implemented the inverse-free sign function algorithm in the same environment as describe above for the sign-Newton algorithm (3)–(5). When applied to Example 1 in the introduction, we obtain the results that appear in Table 1. The stable right deflating subspace computed using the inverse-free sign function algorithm has much smaller backward errors than when computed using the sign-Newton algorithm, but larger backward errors than when computed using the expensive but backward stable QZ algorithm. The inverse-free sign function forward errors are significantly smaller than the sign-Newton algorithm (3)–(5) forward errors. The inverse-free sign function forward errors are even slightly smaller than the QZ algorithm forward errors.

Example 2 We generated a CARE of the form (1) from the finite element semi-discretization of a point control problem for a heat equation described in [30] and summarized as Example 4.2 in the benchmark collection [1]. Here, $n = 200$ and the other parameters take the default values given in [1]. In contrast to [1], in order to obtain a generalized CARE with $E \neq I_n$, we did not invert the Gramian (or mass) matrix. Both the sign-Newton iteration and the inverse-free sign function iteration required 17 iterations to convergence. The convergence history ($\|A_{k+1} - A_k\|_F / \|A_{k+1}\|_F$ for the sign-Newton iteration (3)–(5) and $\|R_{k+1} - R_k\|_F / \|R_{k+1}\|_F$ from (13) for the inverse-free sign function iteration) is shown in Figure 1. The figure

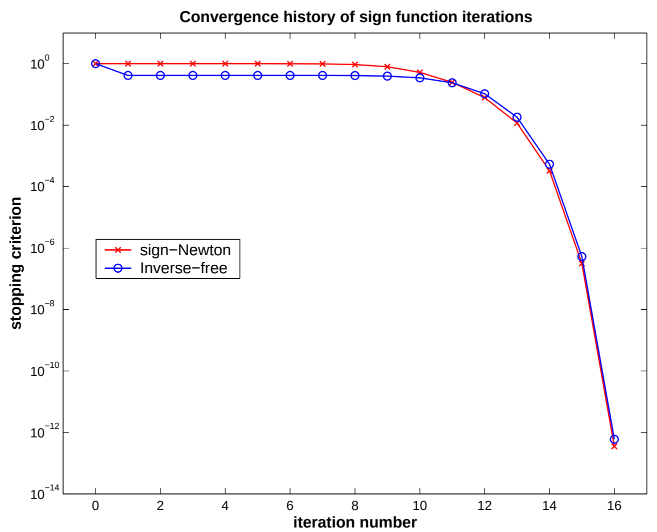

Figure 1: Example 2, Convergence history.

shows the similar convergence behavior expected from mathematically equivalent iterations. The locally quadratic convergence rate is evident for both iterations.

We obtained the following residuals for the CARE solutions X_{gl}, X_{if}, and X_{qz} computed by the sign-Newton iteration, the inverse-free sign function iteration, and the MATLAB Control Toolbox function `care` [8, 26] implementing the generalized Schur method [4].

$$\begin{array}{rcl} \mathcal{R}(X_{\mathrm{gl}}) & = & 1.5 \cdot 10^{-15}, \\ \mathcal{R}(X_{\mathrm{if}}) & = & 3.6 \cdot 10^{-14}, \\ \mathcal{R}(X_{\mathrm{qz}}) & = & 2.1 \cdot 10^{-13}, \end{array}$$

In this example, both iterative methods yield smaller residuals than the MATLAB function.

Example 3 Here, the CARE comes from a linear-quadratic control problem for a second-order linear system described in [20] and summarized as Example 4.3 in the benchmark collection [1]. Here, $n = 60$ and the other parameters take the default values given in [1]. In order to obtain a generalized CARE with $E \neq I_n$, we did not invert the mass matrix. Using the same notation as in Example 2 we get residuals

$$\begin{array}{rcl} \mathcal{R}(X_{\mathrm{gl}}) & = & 3.4 \cdot 10^{-11}, \\ \mathcal{R}(X_{\mathrm{if}}) & = & 4.0 \cdot 10^{-12}, \\ \mathcal{R}(X_{\mathrm{qz}}) & = & 1.3 \cdot 10^{-10}. \end{array}$$

Here, the residual from the inverse-free sign function iteration is almost two orders of magnitude smaller than the residual from the generalized Schur method.

Figure 3 shows the convergence history of the sign-Newton and inverse-free sign function iterations.

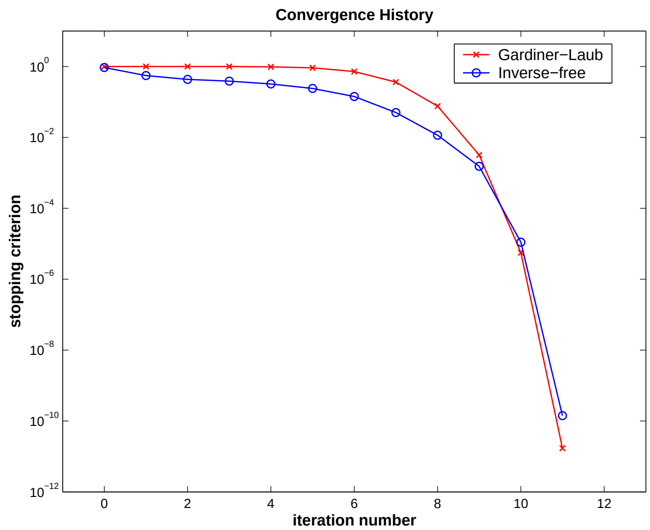

Figure 2: Example 3, convergence history.

4 Conclusions

We have described a new inverse-free sign function iteration method for solving generalized continuous-time algebraic Riccati equations. It is closely related to the generalized sign-Newton iteration and can be considered as an instance of the sign function method. It attains improved numerical backward stability compared to the generalized sign-Newton iteration by avoiding explicit matrix inverses.

Example 1 suggests that inverse-free sign function algorithm is more robust in the presence of rounding errors than the sign-Newton iteration (3)–(5). However, the example also demonstrates that the inverse-free sign function is not backward numerically stable in the conventional sense that rounding errors are equivalent to making a rounding-error-small perturbation of the data. Nevertheless, contrary to expectations, in all three examples, rounding error induced forward errors using the inverse-free sign function are slightly smaller than when using the backward stable QZ algorithm. An understanding of the effects of rounding errors remains an open question.

Acknowledgments

Ralph Byers was partially supported by the University of Kansas General Research Fund allocation 2301062-003 and by the National Science Foundation under awards 0098150, 0112375, and 9977352.

Peter Benner was supported by the DFG Research Center "Mathematics for Key Technologies" in Berlin and the *Deutsche Forschungsgemeinschaft* Research Grant Bu 687/12-1.

References

[1] J. Abels and P. Benner. CAREX – a collection of benchmark examples for continuous-time algebraic Riccati equations (version 2.0). SLICOT Working Note 1999-14, November 1999. Available from `http://www.win.tue.nl/niconet/NIC2/reports.html`.

[2] B.D.O. Anderson and J.B. Moore. *Optimal Control – Linear Quadratic Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[3] A.C. Antoulas. *Lectures on the Approximation of Large-Scale Dynamical Systems*. SIAM Publications, Philadelphia, PA, to appear.

[4] W.F. Arnold, III and A.J. Laub. Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proc. IEEE*, 72:1746–1754, 1984.

[5] Z. Bai and J. Demmel. Design of a parallel nonsymmetric eigenroutine toolbox, Part I. In R.F. Sincovec et al., editor, *Proceedings of the Sixth SIAM Conference on Parallel Processing for Scientific Computing*, pages 391–398. SIAM, Philadelphia, PA, 1993. *See also:* Tech. Report CSD-92-718, Computer Science Division, University of California, Berkeley, CA 94720.

[6] Z. Bai and J. Demmel. Using the matrix sign function to compute invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 19(1):205–225, 1998.

[7] Z. Bai, J. Demmel, and M. Gu. An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. *Numer. Math.*, 76(3):279–308, 1997.

[8] G.J. Balas, J.C. Doyle, K. Glover, A. Packard, and R. Smith. *μ-Analysis and Synthesis Toolbox. For Use with* MATLAB*, Version 4*. The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass, 01760, 2001.

[9] L. Balzer. Accelerated convergence of the matrix sign function. *Internat. J. Control*, 32:1057–1078, 1980.

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

[11] P. Benner and R. Byers. Disk functions and their relationship to the matrix sign function. In *Proc. European Control Conf. ECC 97*, Paper 936. BELWARE Information Technology, Waterloo, Belgium, 1997. CD-ROM.

[12] P. Benner and R. Byers. An arithmetic for matrix pencils: Theory and new algorithms. Technical report, University of Kansas, Department of Mathematics, 405 Snow Hall, 1460 Jayhawk Blvd, Lawrence, KS 66045-7523, 2003.

[13] R. Byers. Numerical stability and instability in matrix sign function based algorithms. In C.I. Byrnes and A. Lindquist, editors, *Computational and Combinatorial Methods in Systems Theory*, pages 185–200. Elsevier (North-Holland), New York, 1986.

[14] R. Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85:267–279, 1987.

[15] R. Byers, C. He, and V. Mehrmann. The matrix sign function method and the computation of invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 18(3):615–632, 1997.

[16] A. A. Dubrulle. An optimum iteration for the matrix polar decomposition. *Electron. Trans. Numer. Anal.*, 8:21–25 (electronic), 1999.

[17] J.D. Gardiner and A.J. Laub. A generalization of the matrix-sign-function solution for algebraic Riccati equations. *Internat. J. Control*, 44:823–832, 1986.

[18] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[19] M. Green and D.J.N Limebeer. *Linear Robust Control*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

[20] J.J. Hench, C. He, V. Kučera, and V. Mehrmann. Dampening controllers via a Riccati equation approach. *IEEE Trans. Automat. Control*, 43:1280–1284, 1998.

[21] C. Kenney and A.J. Laub. On scaling Newton's method for polar decomposition and the matrix sign function. *SIAM J. Matrix Anal. Appl.*, 13:688–706, 1992.

[22] C. Kenney and A.J. Laub. The matrix sign function. *IEEE Trans. Automat. Control*, 40(8):1330–1348, 1995.

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

[24] A.N. Malyshev. Parallel algorithm for solving some spectral problems of linear algebra. *Linear Algebra Appl.*, 188/189:489–520, 1993.

[25] The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass, 01760. MATLAB, *Version 6*, 2000.

[26] The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass, 01760. *The* MATLAB *Control Toolbox, Version 5*, 2000.

[27] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973.

[28] I.R. Petersen, V.A. Ugrinovskii, and A.V.Savkin. *Robust Control Design Using H^∞ Methods*. Springer-Verlag, London, UK, 2000.

[29] J.D. Roberts. Linear model reduction and solution of the algebraic Riccati equation by use of the sign function. *Internat. J. Control*, 32:677–687, 1980. (Reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971).

[30] I.G. Rosen and C. Wang. A multi–level technique for the approximate solution of operator Lyapunov and algebraic Riccati equations. *SIAM J. Numer. Anal.*, 32(2):514–541, 1995.

[31] A. Saberi, P. Sannuti, and B.M. Chen. *H_2 Optimal Control*. Prentice-Hall, Hertfordshire, UK, 1995.

[32] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.