Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 415 (2006) 359–372
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Singular Riccati equations stabilizing large-scale systems

K. Gallivan [a*], X. Rao [a], P. Van Dooren [b]

[a] *Florida State University, 476 Direc. Science Lib., Tallahassee, FL 32306, USA*
[b] *Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium*

Received 27 July 2003; accepted 13 December 2004
Available online 26 February 2005
Submitted by D.C. Sorensen

Abstract

In this paper we discuss the convergence of a stabilization algorithm based on a singular version of the discrete Riccati difference equation. This method is particularly appealing for large scale linear time invariant dynamical systems since one can nicely exploit the sparsity of such systems in order to reduce the complexity of the algorithm.

Keywords: Stabilization; Linear time-invariant system; Riccati difference equation

1. Introduction

In this paper, we focus on the stabilization of a discrete-time system

$$x_{i+1} = Ax_i + Bu_i, \tag{1}$$

where A and B are $n \times n$ and $n \times p$ real matrices which are known, and x_i and u_i are vectors of dimension n and p respectively. The stabilization of the system requires the computation of a $p \times n$ feedback matrix F such that all eigenvalues of $A - BF$ are inside the unit circle and therefore the system defined by replacing A with $A - BF$ is stable. For small and moderate values of n, F can be computed via

* Corresponding author.
E-mail address: gallivan@csit.fsci.edu (K. Gallivan).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.12.031

pole placement or the solution of a matrix equation, e.g., Riccati or Lyapunov equations. The computational requirements for standard algorithms for these approaches, however, is prohibitive for large values of n. Fortunately, when n is large and $p \ll n$, the system matrix A and/or input matrix B are typically very sparse. Algorithms for such problems must therefore exploit this structure in order to efficiently compute a stabilizing feedback.

An important contribution to solving large scale stabilization problems with a few unstable eigenvalues is Saad's projection method [8]. In this algorithm, stabilization or eigenvalue assignment is only imposed on a small invariant subspace that contains the anti-stable invariant subspace of A. Such an approach is often effective, but it can have convergence difficulties and the need for a basis of the invariant subspace can cause excess space requirements for very large systems.

In Saad's projection method, a left invariant subspace V^{T} of A (with presumably small dimension), that contains the left anti-stable invariant subspace of A is computed. In order to exploit the possible sparsity of the matrix A one often chooses to compute the basis directly by a subspace iteration like method. The low-order projected system $(V^{\mathrm{T}}AV, V^{\mathrm{T}}B)$ is then stabilized and the reduced feedback F_v is lifted back to form a stabilizing feedback $F = F_v V^{\mathrm{T}}$ of the original system (A, B). Subspace iteration like methods as proposed by Saad, generate a sequence of approximations to a particular invariant subspace V starting from an initial subspace V_0. The convergence of such methods depends on the separation between eigenvalues of A "contained" in V and the eigenvalues of A not "contained" in V. This is the so-called *gap* of A with respect to V and if it is too small, one should try to compute a larger space instead (see [6]).

In this paper, we discuss an efficient alternative that addresses this convergence difficulty. We also prove that this algorithm converges under very mild conditions and we show that it avoids the need for an explicitly formed basis of the invariant subspace.

2. Discrete Riccati equation stabilization

The major results of this paper are based on the discrete-time Riccati equation (DRE) and the discrete-time Riccati difference equation (DRDE)

$$P = A^{\mathrm{T}}(P - PB(R + B^{\mathrm{T}}PB)^{-1}B^{\mathrm{T}}P)A + Q, \tag{2}$$

$$P_{i+1} = A^{\mathrm{T}}(P_i - P_iB(R + B^{\mathrm{T}}P_iB)^{-1}B^{\mathrm{T}}P_i)A + Q, \tag{3}$$

where R and Q are $p \times p$ and $n \times n$ non-negative matrices and Q is usually decomposed into $L_Q \cdot L_Q^{\mathrm{T}}$. The most general results about DRE and DRDE convergence are given in [2]. It is shown there that under the condition of stabilizability of (A, B), a stabilizer and non-negative solution P_s of DRE (2) exists and a stabilizing feedback F can be computed by

$$F := \hat{R}^{-1}B^{\mathrm{T}}P_sA, \quad \hat{R} := (R + B^{\mathrm{T}}P_sB).$$

Whether the solution of DRDE (3) converges to the stabilizing solution of DRE depends on properties of (A^{T}, L_Q) and the initial condition P_0. We establish in this paper that this algorithm converges to the stabilizing solution under more general conditions than those reported in [2].

The Riccati difference equation (3) has several equivalent formulations. First, one can rewrite it as the Schur complement (with respect to the (1, 1) block) of the compound matrix

$$M = \begin{bmatrix} R + B^{\mathrm{T}} P_i B & B^{\mathrm{T}} P_i A \\ A^{\mathrm{T}} P_i B & A^{\mathrm{T}} P_i A + Q \end{bmatrix}. \tag{4}$$

From this one easily derives a factorized form of the algorithm [4]. One needs to assume that the Cholesky factorizations of the positive semi-definite matrices R, Q and P_i, are given:

$$R := L_R \cdot L_R^{\mathrm{T}}, \quad Q := L_Q \cdot L_Q^{\mathrm{T}}, \quad P_i := S_i \cdot S_i^{\mathrm{T}}. \tag{5}$$

Using these one obtains trivially the following non-square factorization of M:

$$M = \begin{bmatrix} L_R & B^{\mathrm{T}} S_i & 0 \\ 0 & A^{\mathrm{T}} S_i & L_Q \end{bmatrix} \cdot \begin{bmatrix} L_R^{\mathrm{T}} & 0 \\ S_i^{\mathrm{T}} B & S_i^{\mathrm{T}} A \\ 0 & L_Q^{\mathrm{T}} \end{bmatrix}. \tag{6}$$

The so-called square root form of the Riccati difference iteration is then obtained from a lower triangular reduction of the left factor ([4]):

$$\begin{bmatrix} L_R & B^{\mathrm{T}} S_i & 0 \\ 0 & A^{\mathrm{T}} S_i & L_Q \end{bmatrix} \cdot U_i = \begin{bmatrix} \hat{L}_i & 0 & 0 \\ \hat{K}_i & S_{i+1} & 0 \end{bmatrix}, \tag{7}$$

where U_i is orthogonal. We will assume in this paper that $R > 0$, which implies that $\hat{R}_i := R + B P_i B^{\mathrm{T}} > 0$ as well. As a consequence, we obtain a decomposition of M:

$$M = \begin{bmatrix} \hat{L}_i & 0 \\ \hat{K}_i & S_{i+1} \end{bmatrix} \cdot \begin{bmatrix} \hat{L}_i^{\mathrm{T}} & \hat{K}_i^{\mathrm{T}} \\ 0 & S_{i+1}^{\mathrm{T}} \end{bmatrix}, \tag{8}$$

from which it follows that the Schur complement with respect to the (1, 1) block equals $P_{i+1} = S_{i+1} \cdot S_{i+1}^{\mathrm{T}}$. Notice that this holds even if P_{i+1} is not of full rank.

Another formulation of (3) follows from the underlying two-point boundary value problem [1,10]:

$$\begin{bmatrix} A & 0 \\ -Q & I_n \end{bmatrix} \begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} = \begin{bmatrix} I_n & B R^{-1} B^{\mathrm{T}} \\ 0 & A^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \end{bmatrix},$$

where $P_i = Y_i X_i^{-1}$ implies $P_i = Y_{i+1} X_{i+1}^{-1}$ and vice versa (this implies of course that both X_i and X_{i+1} must be invertible). We rederive this formulation below in a more explicit form.

Lemma 1. *If $R > 0$ the DRDE (3) can be rewritten as follows*:

$$\begin{bmatrix} A & 0 \\ -Q & I_n \end{bmatrix} \begin{bmatrix} I_n \\ P_{i+1} \end{bmatrix} = \begin{bmatrix} I_n & BR^{-1}B^{\mathrm{T}} \\ 0 & A^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} I_n \\ P_i \end{bmatrix} A_{F_i}, \tag{9}$$

where

$$A_{F_i} := A - B \cdot F_i, \quad F_i := \hat{R}_i^{-1} B^{\mathrm{T}} P_i A, \quad \hat{R}_i := \hat{R} + B^{\mathrm{T}} P_i B.$$

Proof. We need to show the following two identities:

$$A = (I + BR^{-1}B^{\mathrm{T}}P_i)A_{F_i}, \quad P_{i+1} - Q = A^{\mathrm{T}} P_i A_{F_i}.$$

Using the definition of the matrices involved, the second equation becomes

$$\begin{aligned} P_{i+1} &= A^{\mathrm{T}} P_i A - A^{\mathrm{T}} P_i B F_i + Q \\ &= A^{\mathrm{T}} P_i A - A^{\mathrm{T}} P_i B \hat{R}_i^{-1} B^{\mathrm{T}} P_i A + Q, \end{aligned}$$

which is the DRDE. The first equation becomes

$$A = A + BR^{-1}B^{\mathrm{T}} P_i A - BF_i - BR^{-1}B^{\mathrm{T}} P_i B F_i,$$

which is equivalent to

$$0 = B[R^{-1}\hat{R}_i - I - R^{-1}B^{\mathrm{T}} P_i B]F_i$$

and is clearly an identity. □

3. Convergence of the DRDE

If one wants to study the convergence of the DRDE, the above lemma plays a crucial role. It is clear from (9) that the generalized eigenvalue problem

$$\lambda M_1 - M_2 := \lambda \begin{bmatrix} A & 0 \\ -Q & I_n \end{bmatrix} - \begin{bmatrix} I_n & BR^{-1}B^{\mathrm{T}} \\ 0 & A^{\mathrm{T}} \end{bmatrix} \tag{10}$$

will determine the convergence of the DRDE. For simplicity we assume A to be invertible here but it can be shown that this assumption does not affect our results. Iteration (9) is then a subspace iteration with a space of dimension n:

$$\begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} = M_1^{-1} M_2 \begin{bmatrix} X_i \\ Y_i \end{bmatrix}.$$

Let $\{\lambda_i : 1 \leqslant i \leqslant 2n\}$ be the set of eigenvalues of $M_1^{-1}M_2$ and assume they are ordered by decreasing magnitude $|\lambda_i|$. If $|\lambda_n|$ is strictly larger than $|\lambda_{n+1}|$ then the above recurrence is known to converge for almost all initial conditions X_0, Y_0, to the so-called dominant invariant subspace of $M_1^{-1}M_2$. If, on the other hand, $|\lambda_n| = |\lambda_{n+1}|$ then the iteration never converges: there exist fixed points but they correspond to very special initial conditions [3]. It turns out that $M_1^{-1}M_2$ is simplectic and therefore has a special eigenvalue pattern: the eigenvalues which are not on the unit circle come in pairs that are mirror images of each other with respect to the unit circle. Therefore the condition $|\lambda_n| > |\lambda_{n+1}|$ is satisfied iff $M_1^{-1}M_2$ has no eigenvalues on the unit circle. We make this assumption in the rest of the paper. This is a classical assumption in the DRDE literature since it is closely linked to the existence of stabilizing solutions of the corresponding feedback problem [2]. We recall in this context the following results proved in [2].

Theorem 2. *A stabilizing solution P_s of the DRE exists and is unique if and only if either of the following two conditions is satisfied*:

(1) (A, B) *is stabilizable and* (A^{T}, Q) *has no unobservable eigenvalues on the unit circle*,
(2) (A, B) *is stabilizable and the pencil* $\lambda M_1 - M_2$ *has no generalized eigenvalues on the unit circle*.

The simplectic structure of the pencil implies that all eigenvalues are then mirror images of each other with respect to the unit circle, and the following result then holds ([2]).

Theorem 3. *Let the simplectic pencil $\lambda M_1 - M_2$ have no generalized eigenvalues on the unit circle. Then there exist invertible matrices S and T such that*

$$\lambda M_1 - M_2 = T \begin{bmatrix} \lambda A_F - I & 0 \\ 0 & \lambda I - A_F^{\mathrm{T}} \end{bmatrix} S,$$

where A_F is stable and depends on the stabilizing solution P_s as follows:

$$A_F := A - B \cdot F, \quad F := \hat{R}^{-1} B^{\mathrm{T}} P_s A, \quad \hat{R} := R + B^{\mathrm{T}} P_s B.$$

Under these conditions, the power method converges, provided the initial matrix $\begin{bmatrix} I_n \\ P_0 \end{bmatrix}$ has a "non-degenerate" component in the direction of the invariant subspace $\begin{bmatrix} I_n \\ P_s \end{bmatrix}$. When expressing the initial matrix as a linear combination of both invariant spaces (spanned by the block columns of S^{-1}):

$$\begin{bmatrix} I_n \\ P_0 \end{bmatrix} = S^{-1} \begin{bmatrix} V \\ W \end{bmatrix},$$

the non-degeneracy implies that V must be invertible. Since

$$V = \begin{bmatrix} I_n & 0_n \end{bmatrix} S \begin{bmatrix} I_n \\ P_0 \end{bmatrix} = S_{11} + S_{12} P_0, \tag{11}$$

it is easy to see that for random initial matrices P_0 the matrix V is generically invertible (i.e. the condition does not hold on a set of matrices P_0 of measure 0).

The DRDE thus almost always converges to the stabilizing solution of the DRE since the corresponding simplectic pencil $\lambda M_1 - M_2$ has no unit circle eigenvalues.

Another way to rewrite the condition that V is invertible is to use the fact that one can choose

$$S^{-1} = \begin{bmatrix} I_n & X \\ P_s & Y \end{bmatrix},$$

where $\begin{bmatrix} I_n \\ P_s \end{bmatrix}$ and $\begin{bmatrix} X \\ Y \end{bmatrix}$ span deflating subspaces of $\lambda M_1 - M_2$ corresponding to the generalized eigenvalues outside and inside the unit circle, respectively. Since S is invertible, V is the Schur complement of

$$Z := \left[\begin{array}{c|cc} 0 & -I_n & 0 \\ \hline I_n & I_n & X \\ P_0 & P_s & Y \end{array}\right]$$

and hence V is invertible iff the above matrix is Z is invertible. By taking appropriate Schur complements of Z this also implies that

$$V \text{ invertible} \iff \begin{bmatrix} I_n & X \\ P_0 & Y \end{bmatrix} \text{invertible} \iff Y - P_0 X \text{ invertible.} \tag{12}$$

Notice that if X is invertible, then $P_{\bar{s}} := YX^{-1}$ is the anti-stabilizing solution of the algebraic Riccati equation, and then V is invertible iff $P_{\bar{s}} - P_0$ is invertible.

We now return to the case where A is singular. If this is the case we consider a perturbed matrix $A_\epsilon := A - \epsilon E$ which has the same Jordan decomposition as A, except for the zero eigenvalue of A which now gets changed to ϵ. The assumptions of Theorem 2 are not affected by this since (i) stabilizability of (A_ϵ, B) and of (A, B) are equivalent and (ii) (A_ϵ, Q) has no unobservable modes on the unit circle provided ϵ is sufficiently small. The stabilizing solution $P_{s,\epsilon}$ of the corresponding perturbed DRE is then well-defined. Moreover,

$$\lim_{\epsilon \to 0} P_{s,\epsilon} = P_s$$

since the corresponding invariant subspaces

$$\begin{bmatrix} I_n \\ P_{s,\epsilon} \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} I_n \\ P_s \end{bmatrix}$$

are well defined and ϵ-close to each other [3]. By continuity, one then sees that the invertibility of A is not needed to prove the convergence of the DRDE. In this section we thus proved the following theorem.

Theorem 4. *Let the simplectic pencil $\lambda M_1 - M_2$ have no generalized eigenvalues on the unit circle and let the initial matrix P_0 satisfy the non-degeneracy condition rank $(S_{11} + S_{12} P_0) = n$. Then the iterates $P_i := Y_i X_i^{-1}$ converge linearly to the stabilizing solution P_s of the DRDE*:

$$\lim_{i \to \infty} P_i = P_s, \quad \lim_{i \to \infty} \|P_{i+1} - P_s\| / \|P_i - P_s\| = c < 1.$$

Moreover, the non-degeneracy condition is satisfied for almost all initial conditions P_0.

Remark. The result of the above Theorem 4 relates the assumptions that were needed to prove convergence of the DRDE so far. In [2] it is shown that under the assumptions of Theorem 2, the DRDE converges for any initial matrix P_0 which is either positive definite (i.e. $P_0 > 0$), or larger than the stabilizing solution (i.e. $P_0 > P_s$). The economical SQR algorithm described in Section 4 requires a *singular* matrix P_0 of rank larger than or equal to P_s. Both assumption required in [2] therefore do not hold then. This is why the above theorem is so crucial for the rest of the paper.

We already know that the invariant subspace computed at each iteration i converges to the stable invariant subspace we are interested in, but one typically wants to know this in terms of the matrix P_i as well. Although it is normal to expect linear convergence here as well, we analyze this in more detail in this section.

The following simple lemma follows by straightforward error analysis of the inverse of a matrix and can be found in slightly modified form in [9].

Lemma 5. *Let A be a square invertible matrix with smallest singular value* $\sigma_{\min}$ *and let E be a perturbation of norm* $\delta := \|E\|_2 < \sigma_{\min}$. *Then*

$$(A+E)^{-1} = A^{-1} - A^{-1}EA^{-1} + \varDelta,$$

$$\varDelta = (A+E)^{-1}EA^{-1}EA^{-1} = A^{-1}EA^{-1}E(A+E)^{-1},$$

$$\|\varDelta\|_2 \approx \|A^{-1}EA^{-1}EA^{-1}\|_2 < \delta^2/\sigma_{\min}^3.$$

Defining the convergence error $E_i := P_i - P_s$ and applying the above lemma to the expressions

$$\hat{R}_i^{-1} = (R + B^{\mathrm{T}}P_iB)^{-1},$$

$$P_{i+1} = A^{\mathrm{T}}[P_i - P_iB\hat{R}_i^{-1}B^{\mathrm{T}}P_i]A + Q,$$

we obtain

$$\hat{R}_i^{-1} = R_i^{-1} - R_i^{-1}B^{\mathrm{T}}E_iBR_i^{-1} + O(\|E_i\|_2^2)$$

and

$$\begin{aligned}E_{i+1} &= A^{\mathrm{T}}[E_i - E_iB\hat{R}_i^{-1}B^{\mathrm{T}}P_i - P_iB\hat{R}_i^{-1}B^{\mathrm{T}}E_i \\ &\quad + P_iB\hat{R}_i^{-1}B^{\mathrm{T}}E_iB\hat{R}_i^{-1}B^{\mathrm{T}}P_i]A + O(\|E_i\|_2^2), \\ &= (A - BF_i)^{\mathrm{T}}E_i(A - BF_i) + O(\|E_i\|_2^2),\end{aligned}$$

where $F_i := \hat{R}_i^{-1}B^{\mathrm{T}}P_iA$.

Corollary 6. *Let A_{F_i} be the closed loop matrix $A + BF_i$ and let the error $E_i :=P_i - P_s$ between the ith iterate of the DRDE and its steady state value P_s be small, then this error converges linearly and is in first order equal to*

$$E_{i+1} = A_{F_i}^{\mathrm{T}} E_i A_{F_i} + O(\|E_i\|_2^2).$$

Remark. The convergence ratio of Theorem 4 is therefore approximately equal to $\rho(A_F)^2$(the square of the spectral radius of A_F), since A_{F_i} tends to A_F. Notice that this is smaller than 1 since A_F is the stabilized closed loop matrix.

4. The singular SQR algorithm

The square root algorithm (SQR) of this paper is based on the DRDE with $Q = 0$. In the previous section we showed that the DRDE equation converges under very mild conditions to the stabilizing solution provided the corresponding pencil $\lambda M_1 - M_2$ has no unit circle eigenvalues. For $Q = 0$ this pencil has a spectrum that is the union of the spectrum of A and that of A^{-1} since

$$\lambda M_1 - M_2 = \lambda \begin{bmatrix} A & 0 \\ 0 & I_n \end{bmatrix} - \begin{bmatrix} I_n & BR^{-1}B^{\mathrm{T}} \\ 0 & A^{\mathrm{T}} \end{bmatrix}. \tag{13}$$

Therefore the feedback F generated in the limit moves the unstable eigenvalues of A, λ to their unit circle mirror images, $1/\bar{\lambda}$ and leaves the stable eigenvalues unchanged. As a special case of the square root form of DRDE, the SQR stabilization algorithm (developed in [6]) has the form

$$\begin{bmatrix} L_R & B^{\mathrm{T}}S_i \\ 0 & A^{\mathrm{T}}S_i \end{bmatrix} U_i = \begin{bmatrix} \hat{L}_i & 0 \\ \hat{K}_i & S_{i+1} \end{bmatrix}, \tag{14}$$

where U_i is orthogonal and the dimension of S_i is $n \times l$, the same as S_o. Note that the QR decomposition is computed for a small matrix with size $(p + l) \times p$ (the first row of (7)) and the feedback F_i can be computed from $\hat{L}_i$ and $\hat{K}_i$ as follows:

$$F_i = \hat{L}_i^{-T} \hat{K}_i^{\mathrm{T}}.$$

Moreover, if A and B are sparse, the construction of the left factor in the left-hand side of (14) is cheap as well (see [6]).

The SQR iteration can produce the same sequence of subspaces as Saad's subspace iteration method with only an additional economical QR decomposition of S_i since the updating of S_i has the form $S_{i+1} = A^{\mathrm{T}} S_i U_i^{22}$. If S_0 is taken to be the same initial subspace basis as used for Saad's method, SQR will converge. Moreover convergence is easier to check as was pointed out in [6].

It is also useful to point out that for $Q = 0$ the DRDE can be rewritten in a very compact manner:

$$P_{i+1} = A^{\mathrm{T}} P_i A_{F_i}, \tag{15}$$

or equivalently

$$S_{i+1}S_{i+1}^{\mathrm{T}} = A^{\mathrm{T}}S_iS_i^{\mathrm{T}}A_{F_i}. \tag{16}$$

In the limit we also have that P_s satisfies the discrete-time Sylvester equation

$$P_s = A^{\mathrm{T}}P_sA_F.$$

Saad's subspace iteration method essentially performs the QR factorization of $A^{\mathrm{T}}V_i$ where V_i is the previously computed orthogonal base:

$$A^{\mathrm{T}}V_i = V_{i+1}R_{i+1}. \tag{17}$$

Comparing this with

$$A^{\mathrm{T}}S_iU_i^{22} = S_{i+1}, \tag{18}$$

it is obvious that both methods compute the same spaces. Because of (17) and (18),

$$\mathrm{Im}V_0 = \mathrm{Im}S_0 \Rightarrow \mathrm{Im}V_i = \mathrm{Im}S_i \quad \forall i,$$

as long as U_i^{22} and R_{i+1} are invertible. Multiplying (16) by the right inverse of S_{i+1}^{T} we obtain:

$$U_i^{22} = S_i^{\mathrm{T}}A_{F_i}S_{i+1}^{+}.$$

Upon convergence, S_{i+1} and S_i are close to each other, and one shows that $S_i^{\mathrm{T}}A_{F_i}S_{i+1}^{+}$ is then a matrix whose spectrum is a subset of that of A_{F_i} and hence is stable. The effect of such a multiplication is to dampen out the components along the smallest eigenvalues of $S_i^{\mathrm{T}}A_{F_i}S_{i+1}^{+}$, and the iterates S_i may converge to a smaller rank matrix. This is actually what happens in practice if S_0 has dimension larger than the number of unstable eigenvalues of A.

In order to analyze this we put ourselves in a special coordinate system, where

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix},$$

where A_{11} is anti-stable and A_{22} is stable. Such a coordinate system exists since the condition that the pencil (13) has no eigenvalues on the unit circle also implies that A has no eigenvalues on the unit circle. Also, the stabilizability of (A, B) implies then that (A_{11}, B_1) is controllable, whereas the controllability of (A_{22}, B_2) is not guaranteed. This representation can be obtained under a state-space transformation of the system $\{T^{-1}AT, T^{-1}B\}$ which also transforms all matrices P_i to $T^{\mathrm{T}}P_iT$. In this coordinate system the pencil (13) becomes:

$$\lambda M_1 - M_2 = \lambda \begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} - \begin{bmatrix} I & 0 & W_{11} & W_{12} \\ 0 & I & W_{21} & W_{22} \\ 0 & 0 & A_{11}^{\mathrm{T}} & 0 \\ 0 & 0 & 0 & A_{22}^{\mathrm{T}} \end{bmatrix}, \tag{19}$$

where

$$\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} := \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} R^{-1} \begin{bmatrix} B_1^{\mathrm{T}} & B_2^{\mathrm{T}} \end{bmatrix}.$$

Theorem 7. *Let* (A, B) *be in the coordinate system* (19). *Then the solution to the DRE has rank k equal to the dimension of the anti-stable subspace of* A. *The invariant subspaces corresponding to the anti-stable and stable generalized eigenvalues are then respectively spanned by*

$$\begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \\ P_{11} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} I_k & 0 \\ 0 & X_{22} \\ 0 & 0 \\ 0 & I_{n-k} \end{bmatrix}.$$

The matrices P_s *and* A_F *in this coordinate system are given by*

$$P_s := \begin{bmatrix} P_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad A_F := \begin{bmatrix} A_{11} - B_1\hat{R}^{-1}B_1^{\mathrm{T}}P_{11}A_{11} & 0 \\ -B_2\hat{R}^{-1}B_1^{\mathrm{T}}P_{11}A_{11} & A_{22} \end{bmatrix},$$

where $\hat{R} := R + B_1^{\mathrm{T}}P_{11}B_1$. *Moreover,* P_{11} *has rank equal to the number k of unstable eigenvalues of* A.

Proof. Let P_{11} solve the following DRE of smaller dimension:

$$P_{11} = A_{11}^{\mathrm{T}}\left(P_{11} - P_{11}B_1(R + B_1^{\mathrm{T}}P_{11}B_1)^{-1}B_1^{\mathrm{T}}P_{11}\right)A_{11},$$

then it is easy to see that P_s given above solves the DRE by verifying that $P_s = A^{\mathrm{T}}P_sA_F$. Moreover A_F given above is stable since A_{22} is already stable. Since there is a unique stabilizing solution to the DRE, P_s must be that solution. One then also obtains the equality

$$\begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & I \\ P_{11} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I & 0 & W_{11} & W_{12} \\ 0 & I & W_{21} & W_{22} \\ 0 & 0 & A_{11}^{\mathrm{T}} & 0 \\ 0 & 0 & 0 & A_{22}^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & I \\ P_{11} & 0 \\ 0 & 0 \end{bmatrix}A_F,$$

which proves the result of the anti-stable invariant subspace.

The negative semi-definite matrix $X_{22} := -\sum_{i=0}^{\infty} A_{22}^k W_{22} A_{22}^{\mathrm{T}^k}$ obviously solves the reduced order Sylvester equation:

$$X_{22} = A_{22}X_{22}A_{22}^{\mathrm{T}} - W_{22}.$$

Using this it follows that

$$\begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X_{22} \\ 0 & 0 \\ 0 & I \end{bmatrix}\hat{A}_F^{\mathrm{T}} = \begin{bmatrix} I & 0 & W_{11} & W_{22} \\ 0 & I & W_{21} & W_{22} \\ 0 & 0 & A_{11}^{\mathrm{T}} & 0 \\ 0 & 0 & 0 & A_{22}^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X_{22} \\ 0 & 0 \\ 0 & I \end{bmatrix}$$

for

$$\hat{A}_F = \begin{bmatrix} P_{11} & 0 \\ 0 & I \end{bmatrix}A_F\begin{bmatrix} P_{11}^{-1} & 0 \\ 0 & I \end{bmatrix}$$

and this proves the result for the stable invariant subspace. □

This theorem implies that the image of P_s is also the desired anti-stable left invariant subspace of A, which explains that when S_0 has rank larger than the number of unstable eigenvalues of A, some components of S_i have to be damped out in the iteration. When we overestimate the dimension of the anti-stable invariant subspace, we therefore nevertheless converge to a subspace of correct dimension. Moreover, Corollary 6 implies that the spectrum of A_F determines the convergence ratio of P_i towards the stabilizing solution P_s. Convergence will occur provided the initial matrix P_0 satisfies the non-degeneracy condition (11) or (12). A test for checking whether convergence has occurred was presented in [6], where several numerical experiments are also reported. But the following lemma says this condition is almost always satisfied provided the initial condition P_0 is of sufficiently high rank.

Theorem 8. *Let P_0 and P_s be of respective ranks r_0 and k. Then*

(1) *the DRDE converges to P_s only if $r_0 \geqslant k$,*
(2) *the DRDE converges to P_s for almost all P_0 with $r_0 \geqslant k$.*

Proof. The first point trivially follows from the recurrence relation (15). For the second point we consider the coordinate system (19) since the condition of the lemma does not depend on it. Write $P_0 = \hat{S}\hat{S}^{\mathrm{T}}$ with $\hat{S}^{\mathrm{T}} = [\hat{S}_1^{\mathrm{T}} \quad \hat{S}_2^{\mathrm{T}}]$ in this coordinate system, then the non-degeneracy condition (12) implies that the following matrix must have full rank:

$$\begin{bmatrix} I_k & 0 & I_k & 0 \\ 0 & I_{n-k} & 0 & X_{11} \\ \hat{S}_1\hat{S}_1^{\mathrm{T}} & \hat{S}_1\hat{S}_2^{\mathrm{T}} & 0 & 0 \\ \hat{S}_2\hat{S}_1^{\mathrm{T}} & \hat{S}_2\hat{S}_2^{\mathrm{T}} & 0 & I_{n-k} \end{bmatrix}.$$

One easily checks that this is the case iff the $k \times k$ matrix $\hat{S}_1\hat{S}_1^{\mathrm{T}}$ is invertible, which holds for almost all matrices P_0 of rank $r_0 > k$. $\square$

This result is very reassuring since it says that random initial conditions of sufficiently high rank will yield a convergent sequence!

5. Numerical experiments

The results of this paper give a theoretical explanation of the convergence behavior observed in [6]. The analysis also give a proof that the DRDE converges to a stabilizing solution of the DRE under milder conditions than those of [2], provided an asymptotically stabilizing solution exists. The number of iteration steps needed to obtain a stabilized system will depend on several factors and it does not seem possible to give upper bounds on this. Results are nevertheless encouraging, in the sense that one can expect stabilization in very few steps. We quote an example from [6] to exemplify this.

The system matrix A is constructed by randomly generating a 100×100 matrix with the MATLAB RANDN function and scaling it to $\tilde{A}$ so that the spectral radius of $\tilde{A}$ is 0.9. The 100×1 matrix B is generated randomly with RANDN as is a 1×100 matrix $\tilde{F}$. Construct $A = \tilde{A} + B\tilde{F}$. All eigenvalues of A (dots in Fig. 1) are well-separated from the unit circle and only two are unstable. The norm $\|\tilde{F}\|_2 = 9.9734$ and eigen-condition number of $\tilde{A}$ is $k_2(X, \tilde{A}) = 110.8009$. So the system (A, B) should be well-conditioned and easy to stabilize. Fig. 2 show the results of rank 2 SQR. Fig. 3 show the result of rank 3 SQR. $P_0^{1/2}$ is randomly generated with RAND. The spectral radius $A - BF_i$ from both rank 2 and rank 3 SQR converges within 7 iterations, which is shown in Figs. 2 and 3. Furthermore, Fig. 1 shows the spectrum of $A - BF_5$ (the + symbols) converges to a stable configuration in five iterations. We also see that the only eigenvalues moved are the two unstable ones of A.

This example illustrates that for a well-conditioned stabilization problem where A has only few unstable eigenvalues and all eigenvalues of A are well-separated from the unit circle, SQR is very efficient and stabilization is reached within only a few steps. We can relax the condition that all eigenvalues of A are well-separated from the unit circle to that all unstable eigenvalues of A are well-separated from all stable eigenvalues of A. If the unstable eigenvalues of A are well-separated from the unit circle, fast stabilization with SQR is expected and feedback convergence with SQR depends on the choice of the rank of $P_0^{1/2}$, with the worst case when some stable eigenvalues of A are very close to the unit circle and we choose an incorrect rank of $P_0^{1/2}$ (larger than the number of unstable eigenvalues of A). In this case, we can monitor the eigenvalue convergence of $V_i^{\mathrm{T}}(A - BF_i)V_i$ to catch the stability of $A - BF_i$ or modify the rank of $P_i^{1/2}$ during the iteration. If some unstable eigenvalues

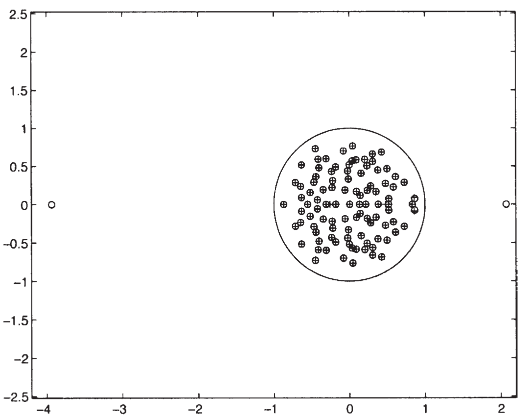

Fig. 1. SQR: Spectrum.

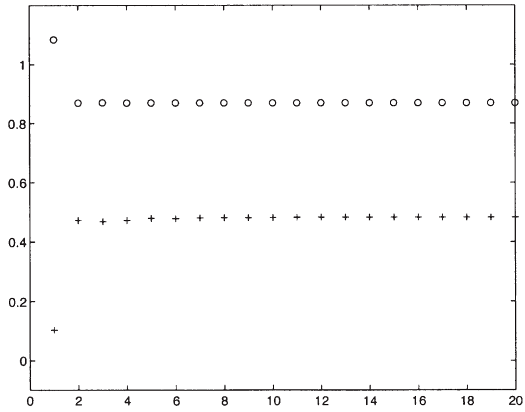

Fig. 2. SQR: $A - BF_i$ (∘) and $V_i'(A - BF_i)V_i$ (+).

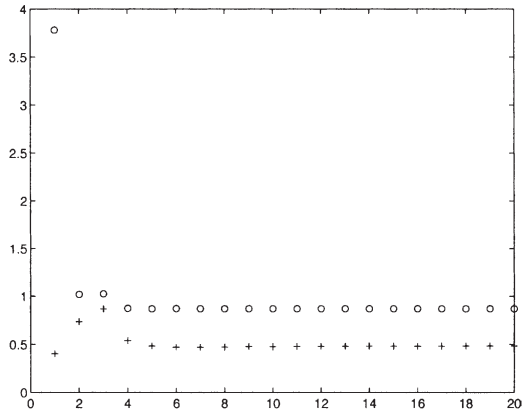

Fig. 3. SQR: $A - BF_i$ (∘) and $V_i'(A - BF_i)V_i$ (+).

of A are very close to the unit circle and stable eigenvalues of A are well-separated from the unit circle, some scaling on A can help to accelerate both stabilization and feedback convergence (see [5,6] for more details).

6. Conclusion

The results of this paper give a theoretical explanation of the convergence behavior observed in [6]. The analysis also give a proof that the DRDE converges to a stabilizing solution of the DRE under milder conditions than those of [2], provided an asymptotically stabilizing solution exists. This paper is an extension of the conference paper [7] treating the same problem.

Acknowledgements

This work was supported by the National Science Foundation under Grant No. CCR-20003050 and by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture.

References

[1] B.D.O. Anderson, Second order convergent algorithms for the steady-state Riccati equation, Int. J. Control, 28 (1978) 295–306.
[2] C.E. De Souza, M.R. Gevers, G.C. Goodwin, Riccati equations in optimal filtering of non-stabilizable systems having singular transition matrices, IEEE Trans. Automat. Control 31 (1986) 831–838.
[3] G. Golub, C. Van Loan, Matrix Computations, second ed., John Hopkins University Press, 1990.
[4] M. Morf, T. Kailath, Square-root algorithms for least-squares estimation, IEEE Trans. Automat. Control AC-20 (1975) 487–497.
[5] X. Rao, Large scale stabilization with linear feedback, M.S. thesis, Dept. Comp. Sc., Florida State University, 1999.
[6] X. Rao, K. Gallivan, P. Van Dooren, Stabilization of large scale dynamical systems, in: Proceedings of the 14th Int. Symp. Math. Th. Netw. & Syst., MTNS2000 Perpignan, June 2000, 1999.
[7] X. Rao, K. Gallivan, P. Van Dooren, Convergence analysis of a Riccati-based stabilization method, in: CD-Rom Proceedings, European Control Conference, Porto, 2001.
[8] Y. Saad, Projection and deflation methods for partial pole assignment in linear state feedback, IEEE Trans. Automat. Control AC-33 (1988) 290–297.
[9] M. Verhaegen, P. Van Dooren, Numerical aspects of different Kalman filter implementations, IEEE Trans. Automat. Control AC-31 (1986) 907–917.
[10] A. White, C. Sage, Optimal Systems Control, second ed., Prentice Hall, Englewood Cliffs, 1977.