May 1979 LIDS-P-908

On the Numerical Solution of the Discrete Time Algebraic Riccati Equation*

by

T. Pappas**

A.J. Laub**

N.R. Sandell, Jr.**

ABSTRACT

In this paper we shall present two new algorithms for solution of the discrete-time algebraic Riccati equation. These algorithms are related to Potter's and to Laub's methods, but are based on solution of a generalized rather than an ordinary eigenvalue problem. The key feature of the new algorithms is that the system transition matrix need not be inverted. Thus the numerical problems associated with an ill-conditioned transition matrix do not arise and, moreover, the algorithm is directly applicable to problems with a singular transition matrix. Such problems arise commonly in practice when a continuous-time system with time delays is sampled.

* Research supported in part by the Department of Energy under Contract ERDA-E(49-18)-2087 and the U.S. Army Research Office under Contract No. DAAG29-79-C-0031.

** Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

Submitted to IEEE Transactions on Automatic Control

1. INTRODUCTION

Most research activity in control theory during the last two decades has been centered around what is usually termed modern control theory. However, it is a fact that the overwhelming majority of practical control systems have been, and continue to be, designed using classical control theory. The reasons for this much discussed gap between theory and practice are too many and complex to discuss here, but we would venture to suggest that one of the reasons for the gap is the inadequate attention that has been paid to the development of numerical algorithms to implement modern control-theoretic design methods.

The objective of this paper is to present a new method for the solution of the discrete-time algebraic Riccati equation. Solution of this equation is basic to design of sampled-data control systems using the solution of the linear-quadratic-Gaussian optimal control problem [1]. Several algorithms have been suggested for direct solution of discrete-time algebraic Riccati equations, for example [2], (in addition to the possibility of iterating the Riccati difference equation to steady-state), but these all require explicit inversion of the state transition matrix. If this matrix is ill-conditioned, numerical difficulties arise. Perhaps even more important is the fact that singular transition matrices cannot be handled. Such transition matrices arise when a continuous system with a time delay is sampled. Since time delays inevitably arise in process control applications due to transportation lags (see, for example [3]), the restriction to nonsingular transition matrices is quite restrictive.

In fact, our original motivation for developing the method proposed in this paper arose from just such an application by Bialkowski [4] in modeling a paper machine.

To illustrate how singular transition matrices arise we give a simple example, taken from Bialkowski's paper [4]. Consider a process whose unit step response is characterized by a pure delay T, a first order time constant τ, and a steady state gain K, and suppose the process is viewed at sample intervals D. (See Figure 1) The discrete equations for a time delay of three unit delays are

$$x(k+1) = \begin{pmatrix} 1-D/\tau & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} x(k) + \begin{pmatrix} KD/\tau \\ 0 \\ 0 \\ 0 \end{pmatrix} u(k)$$

$$y(k) = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} x(k)$$

where

$x(k)$ = vector of current states

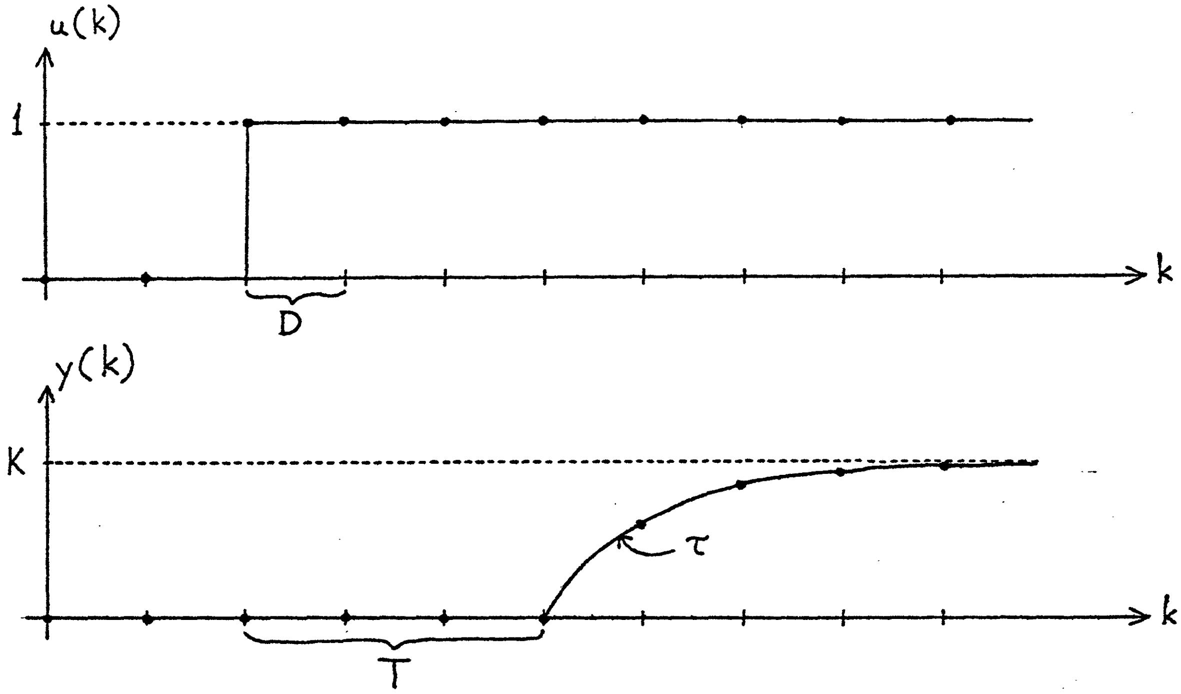
$x(k+1)$ = vector of states one step ahead

$u(k)$ = control input

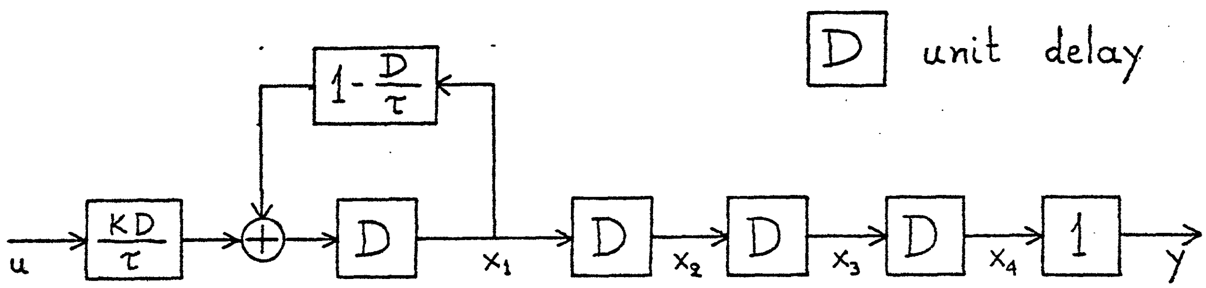
$y(k)$ = output

The state transition matrix of this system is obviously singular.

The method proposed in this paper does <u>not</u> require inversion of the transition matrix. The key idea is to obtain the solution of the discrete-



Unit Step Response



Schematic of system equations

Figure 1.

time algebraic Riccati equation from a basis for the stable eigenspace of a certain generalized eigenvalue problem. Two versions of the method are presented, the first related to Potter's method and utilizing generalized eigenvectors and the second related to Laub's method and utilizing generalized Schur vectors.

The structure of the paper is as follows. Section 2 contains the development of the algorithm using generalized eigenvectors. Section 3 contains the development of the algorithm using generalized Schur vectors. Section 4 discusses some numerical considerations associated with the algorithms, and Section 5 contains some simple examples. Section 6 is the summary and conclusions.

Notation

Throughout the paper $A \in \mathbb{F}^{m \times n}$ will denote an $m \times n$ matrix with coefficients in a field $\mathbb{F}$. The field will usually be the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. The notations A^T and A^H will denote transpose and conjugate transpose, respectively, while A^{-T} will denote $(A^T)^{-1} = (A^{-1})^T$. For $A \in \mathbb{F}^{n \times n}$ its spectrum (set of n eigenvalues) will be denoted $\sigma(A)$. When a matrix $A \in \mathbb{F}^{2n \times 2n}$ is partitioned into four $n \times n$ blocks as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

we shall frequently refer to the individual blocks A_{ij} without further discussion.

2. SOLUTION OF THE DISCRETE-TIME ALGEBRAIC RICCATI EQUATION BY A GENERALIZED EIGENVECTOR APPROACH

2.1 Problem Formulation

In this section we shall be concerned with the discrete-time algebraic Riccati equation

$$F^TXF-X-F^TXG_1(G_2+G_1^TXG_1)^{-1}G_1^TXF+H=0 \tag{1}$$

Here $F, H, X \in \mathbb{R}^{n\times n}$, $G_1 \in \mathbb{R}^{n\times m}$, $G_2 \in \mathbb{R}^{m\times m}$, and $G_2 = G_2^T > 0$, $H = H^T \geq 0$. Also, $m \leq n$. We shall assume that (F, G_1) is a stabilizable pair[1], and that (F, C) is a detectable pair[2], where C is a full-rank factorization of H (i.e., $C^TC = H$ and rank (C) = rank (H)). As discussed in Section 1, we do not need any other assumptions on F. Finally, we define G: $G = G_1G_2^{-1}G_1^T$.

Under the above assumptions (1) is known to have a unique non-negative definite solution; see [5]. Moreover, if (F, C) is completely reconstructible[3], (1) has a unique positive definite solution. There are, of course, many other solutions to (1) but we are interested in computing the non-negative definite one. (Or the positive definite one if it exists.)

---

1. The pair (F, G_1) is stabilizable if $w^HG_1 = 0$ and $w^HF = \lambda w^H$ for some constant λ implies $|\lambda| < 1$ or $w=0$. (See [5] for further characterizations.)
2. The pair (F, C) is detectable if $w^HC^T = 0$ and $w^HF^T = \lambda w^H$ for some constant λ implies $|\lambda| < 1$ or $w=0$. (See [5].)
3. The pair (F, C) is completely reconstructible if $w^HC^T = 0$ and $w^HF^T = \lambda w^H$ for some constant λ implies $w=0$. (See [5].)

The Riccati equation (1) arises in various problems, including the discrete-time linear-quadratic optimal control problem. The algorithms we present are motivated by the relationship between the Riccati equation and the two-point boundary value problem associated with this optimal control problem. If we let x_k denote the state at time t_k and γ_k the corresponding adjoint vector, the Hamiltonian difference equations arising from the discrete maximum principle applied to the linear-quadratic problem are of the form

$$\begin{pmatrix} I & G \\ O & F^T \end{pmatrix} \begin{pmatrix} x_{k+1} \\ \gamma_{k+1} \end{pmatrix} = \begin{pmatrix} F & O \\ -H & I \end{pmatrix} \begin{pmatrix} x_k \\ \gamma_k \end{pmatrix}$$

Note that if F is invertible we can work with the symplectic matrix

$$Z = \begin{pmatrix} I & G \\ O & F^T \end{pmatrix}^{-1} \begin{pmatrix} F & O \\ -H & I \end{pmatrix} = \begin{pmatrix} F + GF^{-T}H & -GF^{-T} \\ -F^{-T}H & F^{-T} \end{pmatrix}$$

using a basis for the stable eigenspace of Z to compute the desired solution to the Riccati equation. Two examples of such bases are the eigenvectors and the Schur vectors corresponding to the stable eigenvalues of Z; see [2] and [6], respectively.

The key idea of this paper is to consider, instead of the standard eigenvalue problem for Z, the generalized eigenvalue problem

$$Mz = \lambda Lz$$

with

$$L = \begin{pmatrix} I & G \\ O & F^T \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} F & O \\ -H & I \end{pmatrix} .$$

We shall show that we can use a basis for the stable generalized eigenspace of the problem to construct the solution of (1). We shall discuss two ways of finding this basis. The first is by computing the generalized eigenvectors corresponding to the stable generalized eigenvalues. The second is by using the generalized Schur vectors of the problem.

2.2 The Generalized Eigenvalue Problem

Definition: Consider the generalized eigenvalue problem:

$$Mz = \lambda Lz \tag{2}$$

The generalized eigenvalues of the problem are the roots of the generalized characteristic equation $\det(M - \lambda L) = 0$. For each generalized eigenvalue λ, a nonzero vector satisfying (2) will be called a generalized eigenvector of the problem corresponding to λ. If λ is a generalized eigenvalue with multiplicity $r > 1$, then the set of vectors $\{z_1,\ldots,z_\ell\}$ satisfying:

$$Mz_1 = \lambda Lz_1$$

$$(M - \lambda L)z_k = Lz_{k-1} \qquad k = 2,3,\ldots,\ell; \quad \ell \le r$$

will be called a chain of generalized principal vectors, and the vector z_k will be called a generalized principal vector of grade k.

From now on we shall drop the adjective "generalized" except where its absence would cause confusion.

Theorem 1: Let $A, B, U, V \in \mathbb{C}^{n\times n}$ with U and V nonsingular. Then the eigenvalues of the problems $Az = \lambda Bz$ and $UAVz = \lambda UBVz$ are the same.

Proof: See [7].

Theorem 2: (a) Let $A, B \in \mathbb{C}^{n\times n}$. Then there exist unitary matrices Q and Z such that QAZ and QBZ are both upper triangular.

(b) Let $A, B \in \mathbb{R}^{n\times n}$. Then there exist orthogonal matrices Q and Z such that QAZ is quasi-upper triangular (real Schur form) and QBZ is upper triangular.

Proof: See [8].

For the remainder of this paper, we shall consider only the generalized eigenvalue problem defined in Section 2.1.

Theorem 3: None of the eigenvalues of the generalized eigenvalue problem $Mz = \lambda Lz$ lies on the unit circle.

Proof: Suppose that $|\lambda| = 1$ is an eigenvalue. Then $Mz = \lambda Lz$ for some $z \neq 0$, and this can be written as

$$\begin{pmatrix} F & 0 \\ -H & I \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \lambda \begin{pmatrix} I & G \\ 0 & F^T \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

This implies

$$Fz_1 = \lambda z_1 + \lambda G z_2 \qquad (3)$$

and $$-Hz_1 + z_2 = \lambda F^T z_2 \tag{4}$$

Premultiplying (3) by $\lambda^* z_2^H$ and postmultiplying the conjugate transpose of (4) by z_1 we get

$$\lambda^* z_2^H F z_1 = |\lambda|^2 z_2^H z_1 + |\lambda|^2 z_2^H G z_2 \tag{3a}$$

$$z_2^H z_1 = \lambda^* z_2^H F z_1 + z_1^H H z_1 \tag{4a}$$

Adding (3a) and (4a) and noting that $|\lambda|^2 = 1$, we get

$$z_2^H G z_2 + z_1^H H z_1 = 0$$

Now recall that $G = G_1 G_2^{-1} G_1^T$ with $G_2 > 0$ and $H = C^T C$. Then $(z_2^H G_1) G_2^{-1} (G_1^T z_2) + (z_1^H C^T) C z_1 = 0$ implies

$$G_1^T z_2 = 0 \tag{5}$$

and $$Cz_1 = 0 \tag{6}$$

Combining (5) and (6) with (3) and (4) we get

$$Fz_1 = \lambda z_1 \tag{7}$$

$$F^T z_2 = \frac{1}{\lambda} z_2 \tag{8}$$

Using (5) and (8) we see that $z_2^H G_1 = 0$ and $z_2^H F = \frac{1}{\lambda} z_2^H$. By stabilizability this implies $z_2 = 0$, since $|\frac{1}{\lambda}| > 1$. Also from (6) and (7) we have $z_1^H C^T = 0$ and $z_1^H F^T = \lambda z_1^H$. By detectability this implies $z_1 = 0$, since $|\frac{1}{\lambda}| > 1$. Therefore $z = 0$. But this is a contradiction, so none of the eigenvalues lies on the unit circle.

Theorem 4: Consider the generalized eigenvalue problem $Mz = \lambda Lz$. If $\lambda \neq 0$ is an eigenvalue, then $\frac{1}{\lambda}$ is also an eigenvalue with the same multiplicity.

Proof: M and L have the following property

$$LJL^T = MJM^T = \begin{pmatrix} 0 & F \\ -F^T & 0 \end{pmatrix} \quad \text{where } J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

It is very easy to show also that $\det(M - \lambda L) \not\equiv 0$. To see this, note that if the determinant were identically 0, it would, in particular, be 0 for $|\lambda| = 1$ which would contradict Theorem 3.

Now consider the following two generalized eigenvalue problems:

$$Mz = \lambda Lz \tag{A}$$

$$L^T x = \mu M^T x \tag{B}$$

Looking at the corresonding characteristic polynomials

$$\det(M - \lambda L) = 0$$

$$\det(L^T - \mu M^T) \equiv \det(L - \mu M) = 0$$

we see that if λ is an eigenvalue of (A), then $\frac{1}{\lambda}$ is an eigenvalue of (B) with the same multiplicity.

Now let λ_1 be an eigenvalue of (A) with multiplicity r. Then $\mu_1 = \frac{1}{\lambda_1}$ is an eigenvalue of (B) with multiplicity r and we shall show that μ_1 is also an eigenvalue of (A) with multiplicity r.

For problem (B), consider any chain of principal vectors $\{x_1,\ldots,x_\ell\}$, The following Theorem completes the description of the symplectic nature of the eigenvalues of the problem $Mz = \lambda Lz$.

<u>Theorem 5</u>: Consider the generalized eigenvalue problem $Mz = \lambda Lz$. If $\lambda = 0$ is an eigenvalue with multiplicity r, then there are only 2n-r finite eigenvalues for this problem. We may say that the r missing eigenvalues are "infinite" eigenvalues (or "reciprocals of 0").

<u>Proof</u>: Using Theorems 1 and 2a, we may assume that L and M are both upper triangular, with diagonal elements α_i and β_i respectively. Then

$$\det(M - \lambda L) = \prod_{i=1}^{2n} (\alpha_i - \lambda\beta_i)$$

If $\lambda_1 = 0$ is an eigenvalue with multiplicity r, then without loss of generality $\alpha_1 = \ldots = \alpha_r = 0$. With an argument identical to the one used in Theorem 4 we can then show that $\mu_1 = 0$ is an eigenvalue, with multiplicity r, of the problem $L^T x = \mu M^T x$. But

$$\det(L^T - \mu M^T) = \det(L - \mu M) = \prod_{i=1}^{2n} (\beta_i - \mu\alpha_i)$$

Thus $\beta_{2n-r} = \ldots = \beta_{2n} = 0$. Note that $\alpha_i = \beta_i = 0$ is not possible because $\det(M - \lambda L) \not\equiv 0$. Therefore,

$$\det(M - \lambda L) = \prod_{i=1}^{r} (-\lambda\beta_i) \prod_{i=r+1}^{2n-r-1} (\alpha_i - \lambda\beta_i) \prod_{i=2n-r}^{2n} \alpha_i .$$

If we adopt the convention that the reciprocal of zero is infinity, then we may say that the missing eigenvalues are "infinite."

To summarize, we can now use Theorems 3, 4, and 5 to arrange the eigenvalues of our problem in the following way:

$\ell \leq r$, corresponding to μ_1:

$$L^T x_1 = \mu_1 M^T x_1$$

$$(L^T - \mu_1 M^T) x_k = M^T x_{k-1} \qquad k = 2,3,\ldots,\ell$$

Then $\quad MJL^T x_1 = \mu_1 MJM^T x_1 = \mu_1 LJL^T x_1.$

Also $\quad (MJL^T - \mu_1 MJM^T) x_k = MJM^T x_{k-1}$

or $\quad (MJL^T - \mu_1 LJL^T) x_k = LJL^T x_{k-1}.$

If we set $z_k = JL^T x_k$ we get

$$Mz_1 = \mu_1 Lz_1$$

$$(M - \mu_1 L) z_k = Lz_{k-1}$$

The independence of the z_k for all chains corresponding to μ_1 follows easily from $\det(M - \lambda L) \not\equiv 0$. So μ_1 is also an eigenvalue of (A) with multiplicity r.

When F is nonsingular, all of the eigenvalues of the problem $Mz = \lambda Lz$ are nonzero. This can be proved very easily by contradiction. For if $\lambda = 0$ were an eigenvalue we would have $Mz = 0$, whence

$$\begin{pmatrix} F & O \\ -H & I \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0 \tag{9}$$

implies $z_1 = 0$ and $z_2 = 0$.

When F is singular, however, there must be at least one eigenvalue equal to zero, because equation (9) has a nontrivial solution.

$$\underbrace{0,\ldots,0}_{r},\ \underbrace{\lambda_{r+1},\ldots,\lambda_n}_{n-r},\ \underbrace{\frac{1}{\lambda_n},\ldots,\frac{1}{\lambda_{r+1}}}_{n-r},\ \underbrace{\infty,\ldots,\infty}_{r}$$

with $0 < |\lambda_i| < 1,\ i = r+1,\ldots,n$

2.3 Main Theorem

For the problem $Mz = \lambda Lz$, let U be the $2n\times n$ matrix of the generalized eigenvectors and generalized principal vectors corresponding to the n stable eigenvalues. The matrix U, a basis for the stable eigenspace, can be partitioned into two $n\times n$ submatrices

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

Then $\quad MU = LUS \qquad (10)$

where S is the $n\times n$ "Jordan Canonical Form," corresponding to all λ_i with $|\lambda_i| < 1$. We can rewrite (10) as

$$M \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = L \begin{pmatrix} U_1S \\ U_2S \end{pmatrix} \qquad (10a)$$

or

$$\begin{pmatrix} F & O \\ -H & I \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} I & G \\ O & F^T \end{pmatrix} \begin{pmatrix} U_1S \\ U_2S \end{pmatrix}$$

which implies the following two relations:

$$FU_1 = U_1S + GU_2S \qquad (11)$$

$$-HU_1 + U_2 = F^TU_2S \qquad (12)$$

If U_1 is invertible (as we shall prove in the next Theorem), then (11) can be written as

$$F = U_1SU_1^{-1} + GU_2SU_1^{-1} \tag{13}$$

or

$$F^T = U_1^{-H}S^HU_1^H + U_1^{-H}S^HU_2^HG \tag{13^H}$$

and (12) can be written as

$$U_2U_1^{-1} = F^TU_2SU_1^{-1} + H \tag{14}$$

Observe that relations (10) through (14) actually hold for any set of n generalized eigenvectors and generalized principal vectors $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$, and a corresponding matrix $\hat{S}$ in "Jordan Canonical Form", and this observation will be used in the proof of the Lemmas and Theorem that follow.

Lemma 1: All solutions of the Riccati equation (1) are of the form $X = PQ^{-1}$, where $\begin{pmatrix} Q \\ P \end{pmatrix}$ is a set of n generalized eigenvectors and generalized principal vectors of the problem $Mz = \lambda Lz$.

Proof: Suppose that X is a solution of the Riccati equation

$$F^TXF - X - F^TXG_1(G_2 + G_1^TXG_1)^{-1} G_1^TXF + H = 0 \tag{1}$$

Let

$$E = F - G_1(G_2 + G_1^TXG_1)^{-1} G_1^TXF \tag{15}$$

Then

$$F^TXE = F^TXF - F^TXG_1(G_2 + G_1^TXG_1)^{-1}G_1^TXF = X - H \tag{16}$$

There exists a nonsingular Q such that $Q^{-1}EQ = \hat{S}$, where $\hat{S}$ is in Jordan Canonical Form. Let $P = XQ$ or $X = PQ^{-1}$. From (16) we have $F^T PQ^{-1}EQ = XQ - HQ$ whence

$$F^T P\hat{S} = P - HQ \tag{17}$$

From (15) we have

$$Q\hat{S}Q^{-1} = F - G_1(G_2 + G_1^T XG_1)^{-1}G_1^T XF \tag{18}$$

This implies

$$G_1^T XQ\hat{S}Q^{-1} = G_1^T XF - G_1^T XG_1(G_2 + G_1^T XG_1)^{-1}G_1^T XF \quad .$$

Notice also that

$$G_2(G_2 + G_1^T XG_1)^{-1}G_1^T XF = G_1^T XF - G_1^T XG_1(G_2 + G_1^T XG_1)^{-1}G_1^T XF$$

so that

$$G_2(G_2 + G_1^T XG_1)^{-1}G_1^T XF = G_1^T XQ\hat{S}Q^{-1}$$

or

$$(G_2 + G_1^T XG_1)^{-1}G_1^T XF = G_2^{-1}G_1^T XQ\hat{S}Q^{-1}$$

whence

$$G_1(G_2 + G_1^T XG_1)^{-1}G_1^T XF = GXQ\hat{S}Q^{-1}$$

where $G = G_1G_2^{-1}G_1^T$. Now we use (18) again to obtain

$$F - Q\hat{S}Q^{-1} = GXQ\hat{S}Q^{-1}$$

or

$$FQ = Q\hat{S} + GP\hat{S} \tag{19}$$

From (17) and (19) we now have

$$\begin{pmatrix} F & 0 \\ -H & I \end{pmatrix} \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} I & G \\ 0 & F^T \end{pmatrix} \begin{pmatrix} Q \\ P \end{pmatrix} \hat{S}$$

<u>Lemma 2</u>: Consider any set of n generalized eigenvectors and generalized principal vectors $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$, and a corresponding matrix $\hat{S}$ in "Jordan Canonical Form". Assume further that V_1 is an invertible matrix. Then

(a) $X = V_2V_1^{-1}$ solves (1) assuming $(G_2 + G_1^TXG_1)$ is invertible.

(b) $X = X^T \geq 0$ if and only if $\hat{S}$ is stable.

<u>Proof of (a)</u>: Suppose $(G_2 + G_1^TXG_1)$ is invertible. Using the analogs of equations (13) and (14) we have

$$F^TXF - X - F^TXG_1(G_2 + G_1^TXG_1)^{-1}G_1^TXF + H = F^TV_2V_1^{-1}(V_1\hat{S}V_1^{-1} + GV_2\hat{S}V_1^{-1}) - V_2V_1^{-1}$$

$$- F^TV_2V_1^{-1}G_1(G_2 + G_1^TV_2V_1^{-1}G_1)^{-1}G_1^TV_2V_1^{-1}F + V_2V_1^{-1} - F^TV_2\hat{S}V_1^{-1}$$

$$= F^TV_2V_1^{-1}GV_2\hat{S}V_1^{-1} - F^TV_2V_1^{-1}G_1(G_2 + G_1^TV_2V_1^{-1}G_1)^{-1}G_1^TV_2V_1^{-1}F$$

$$= F^TV_2V_1^{-1}G_1G_2^{-1}G_1^TV_2\hat{S}V_1^{-1} - F^TV_2V_1^{-1}G_1(G_2 + G_1^TV_2V_1^{-1}G_1)^{-1}G_1^TV_2\hat{S}V_1^{-1} -$$

$$F^TV_2V_1^{-1}G_1(G_2 + G_1^TV_2V_1^{-1}G_1)^{-1}G_1^TV_2V_1^{-1}G_1G_2^{-1}G_1^TV_2\hat{S}V_1^{-1} \text{ by (13)}$$

$$= KW^{-1}Z - K(W + Y)^{-1}Z - K(W + Y)^{-1}YW^{-1}Z$$

$$= K[W^{-1} - (W + Y)^{-1} - (W + Y)^{-1}YW^{-1}]Z$$

$$\equiv K(W + Y)^{-1}[(W + Y)W^{-1} - I - YW^{-1}]Z$$

$$\equiv 0$$

where $K = F^T V_2 V_1^{-1} G_1$

$W = G_2$

$Y = G_1^T V_2 V_1^{-1} G_1$

and $Z = G_1^T V_2 \hat{S} V_1^{-1}$

So $X = V_2 V_1^{-1}$ satisfies equation (1).

<u>Proof of (b):</u> From the analogs of equations (13^H) and (14) we have

$$V_2 V_1^{-1} = (V_1^{-H}\hat{S}^H V_1^H + V_1^{-H}\hat{S}^H V_2^H G) V_2 \hat{S} V_1^{-1} + H$$

$$= V_1^{-H}\hat{S}^H V_1^H V_2 \hat{S} V_1^{-1} + V_1^{-H}\hat{S}^H V_2^H G V_2 \hat{S} V_1^{-1} + H$$

This is equivalent to

$$V_2 V_1^{-1} - (V_1 \hat{S} V_1^{-1})^H V_2 V_1^{-1} (V_1 \hat{S} V_1^{-1}) = (V_2 \hat{S} V_1^{-1})^H G (V_2 \hat{S} V_1^{-1}) + H.$$

We have

$$X = V_2 V_1^{-1} .$$

Let

$$A = V_1 \hat{S} V_1^{-1}$$

and

$$Q = (V_2 \hat{S} V_1^{-1})^H G (V_2 \hat{S} V_1^{-1}) + H$$

Note that $X \in \mathbb{R}^{n\times n}$ since we may always choose real vectors in $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ to span the eigenspace corresponding to complex conjugate pairs of eigenvalues. Note also that $Q = Q^H \geq 0$ and A has the same spectrum as $\hat{S}$. Now there exists a unique symmetric non-negative definite solution of the Lyapunov equation

$$X - A^H XA = Q$$

if and only if $\hat{S}$ is stable. (see Anderson and Moore [5], p. 67). So $X = X^T \geq 0$ if and only if $\hat{S}$ is stable.

<u>Theorem 6:</u> With respect to the notation and assumptions above the following results hold:

(a) U_1 is invertible and
$X = U_2 U_1^{-1}$ is a solution of the Riccati equation (1)
with $X = X^T \geq 0$ ($X = X^T > 0$ if (F,C) completely reconstructible)

(b) $\sigma(S)$ = "closed-loop" spectrum
$= \sigma(F - G_1(G_2 + G_1^T X G_1)^{-1} G_1^T XF)$
$= \sigma(F - G(X^{-1} + G)^{-1} F)$ when X is invertible
$= \sigma(F - GF^{-T}(X-H))$ when F is invertible.

<u>Proof of (a):</u> Assume that U_1 is invertible. Then $X = X^T = U_2 U_1^{-1} \geq 0$ by Lemma 2(b). Also $X \geq 0$ implies $G_2 + G_1^T X G_1 > 0$, and thus $X = U_2 U_1^{-1}$ solves (1) by Lemma 2(a). So if U_1 is invertible we are done. Suppose then that U_1 is not invertible. Now we know that the Riccati equation

has a unique non-negative definite solution [5], which by Lemma 1 must be of the form $X = V_2V_1^{-1} \geq 0$, where $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ is a set of eigenvectors of the problem $Mz = \lambda Lz$, different from the desired set $\begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$. But this implies that the corresponding matrix $\hat{S}$ in "Jordan Canonical Form" is unstable, which by Lemma 2(b) implies that $X = V_2V_1^{-1}$ is not non-negative definite and this is a contradiction.

If (F,C) is completely reconstructible, then we know that (1) has a unique positive definite solution [5], and so the unique non-negative definite solution X must also be positive definite.

<u>Proof of (b)</u>: We must show that

$$\sigma(S) = \sigma(F - G_1(G_2 + G_1^TXG_1)^{-1}G_1^TXF) = \text{"closed loop" spectrum.}$$

That is, we must show that

$$U_1S = (F - G_1(G_2 + G_1^TXG_1)^{-1}G_1^TXF)U_1 \ .$$

Recalling (11), the above equality is equivalent to

$$FU_1 - GU_2S = FU_1 - G_1(G_2 + G_1^TXG_1)^{-1}G_1^TXFU_1.$$

Using (13) we see that

$$\begin{aligned} G_1^TXF &= G_1^TX(U_1SU_1^{-1} + GU_2SU_1^{-1}) \\ &= G_1^TU_2SU_1^{-1} + G_1^TXG_1G_2^{-1}G_1^TU_2SU_1^{-1} \\ &= (G_2 + G_1^TXG_1)G_2^{-1}G_1^TU_2SU_1^{-1}. \end{aligned}$$

This implies that

$$(G_2 + G_1^T X G_1)^{-1} G_1^T XF = G_2^{-1} G_1^T U_2 S U_1^{-1}$$

or

$$G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T XF = G_1 G_2^{-1} G_1^T U_2 S U_1^{-1}$$

or

$$G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T XFU_1 = GU_2 S.$$

Thus

$$FU_1 - GU_2 S = FU_1 - G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T XFU_1$$

or

$$U_1 S U_1^{-1} = F - G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T XF$$

and so $\sigma(S) = \sigma(F - G_1(G_2 + G_1^T X G_1)^{-1} G_1^T XF)$ = "closed-loop" spectrum.

The other equalities follow from well-known identities.

3. SOLUTION OF THE DISCRETE-TIME ALGEBRAIC RICCATI EQUATION BY A SCHUR VECTOR APPROACH

In this section we consider the real Schur vector approach for the solution of (1). The advantage of this approach is that it is not necessary to calculate the eigenvectors corresponding to the stable eigenspace of the problem $Mz = \lambda Lz$. The calculation of the eigenvectors has potentially severe numerical difficulties, especially in the case of multiple eigenvalues; see, for example, [9], and [10]. By analogy with [6] we thus instead determine a basis for the stable eigenspace by means of generalized Schur vectors.

Theorem 7: Let A be a quasi-upper triangular matrix and B an upper triangular matrix. To each block on the diagonal of A and the respective entries on the diagonal of B, there corresponds a real generalized eigenvalue or a pair of complex generalized eigenvalues of the generalized eigenvalue problem $Az = \lambda Bz$. Furthermore, there exist orthogonal matrices Q and Z, such that QAZ is quasi-upper triangular, QBZ is upper triangular and the blocks on their diagonals are arranged so that the blocks corresponding to the stable generalized eigenvalues are all in the upper left quarter of QAZ and QBZ.

Proof: To prove the above theorem we must prove that we can interchange the order of appearance on the diagonals

(i) of two adjacent 1x1 entries corresponding to two distinct real values

(ii) of a 2x2 block corresponding to a pair of complex eigenvalues and an adjacent 1x1 entry corresponding to a real eigenvalue

(iii) of two adjacent 2x2 blocks corresponding to two distinct pairs of complex eigenvalues.

by means of orthogonal matrices Q and Z, while retaining the respective quasi-upper triangular and upper triangular structures.

We will prove only case (i) here, which is the simplest. It becomes clear that (ii) and (iii) are also possible, but quite complicated. Let

$$A_o = \begin{pmatrix} a & x \\ 0 & b \end{pmatrix} \quad \text{and } B_o = \begin{pmatrix} c & y \\ 0 & d \end{pmatrix}$$

Note that $\frac{a}{c} \neq \frac{b}{d}$, since, without loss of generality, the two eigenvalues are distinct. We want to find orthogonal Q and Z such that

$$QA_oZ = \begin{pmatrix} a' & x' \\ 0 & b' \end{pmatrix} \quad \text{and } QB_oZ = \begin{pmatrix} c' & y' \\ 0 & d' \end{pmatrix}$$

with $\frac{a'}{c'} = \frac{b}{d}$ and $\frac{b'}{d'} = \frac{a}{c}$.

It can be shown by direct computation that

$$Q = \begin{pmatrix} \frac{B}{\sqrt{1+B^2}} & -\frac{1}{\sqrt{1+B^2}} \\ \frac{1}{\sqrt{1+B^2}} & \frac{B}{\sqrt{1+B^2}} \end{pmatrix}$$

$$Z = \begin{pmatrix} \frac{A}{\sqrt{1+A^2}} & \frac{1}{\sqrt{1+A^2}} \\ \frac{1}{\sqrt{1+A^2}} & -\frac{A}{\sqrt{1+A^2}} \end{pmatrix}$$

with $A = \frac{yb - xd}{ad - bc}$ and $B = \frac{xc - ya}{ad - bc}$

are orthogonal matrices which effect the desired transformations.

Consider now our generalized eigenvalue problem $Mz = \lambda Lz$. By Theorem 2(b) there exist orthogonal matrices Q_1 and Z_1 such that Q_1MZ_1 is quasi-upper triangular and Q_1LZ_1 is upper triangular. The real generalized eigenvalues are obtained by dividing the entries of the 1×1 blocks on the diagonal of Q_1MZ_1 by the corresponding entries of Q_1LZ_1. The complex generalized eigenvalues are obtained by some more complicated calculations, involving the 2×2 blocks on the diagonal of Q_1MZ_1 and corresponding entries on the diagonal of Q_1LZ_1.

By Theorem 7 there exist orthogonal matrices Q_2 and Z_2 such that $Q_2Q_1MZ_1Z_2$ is quasi-upper triangular and $Q_2Q_1LZ_1Z_2$ is upper triangular and, moreover, the diagonal blocks corresponding to the stable generalized eigenvalues are in the upper left quarter of the matrices. Note that $Q_2Q_1LZ_1Z_2$ has no zero entries on the upper half of its diagonal, for such entries would correspond to infinite and so unstable eigenvalues.

Let $Q = Q_2Q_1$ and $Z = Z_1Z_2$. Then:

$$QMZ = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \text{and} \quad QLZ = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}.$$

As noted above, B_{11} has nonzero diagonal entries and is therefore invertible. Note that A_{21} is 0 because of the symplectic nature of the problem.

Let $S = B_{11}^{-1}A_{11}$. Then:

$$\begin{bmatrix} A_{11} \\ 0 \end{bmatrix} = \begin{bmatrix} B_{11} \\ 0 \end{bmatrix} B_{11}^{-1}A_{11} \tag{17}$$

Define $\Psi := \begin{pmatrix} I_n \\ O_n \end{pmatrix}$.

Then (17) is equivalent to

$$QMZ\Psi = QLZ\Psi S$$

And since Q is orthogonal we have

$$MZ\Psi = LZ\Psi S$$

Let $Z\Psi = U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$.

Then we have

$$M \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = L \begin{pmatrix} U_1 S \\ U_2 S \end{pmatrix} . \qquad (18)$$

Note the similarity of (18) and (10a). Their only difference is that S is now real quasi-upper triangular, ($S = B_{11}^{-1}A_{11}$ = (upper triang.) × (quasi-upper triang.) = quasi-upper triang.) whereas before it was complex upper triangular (actually in "Jordan Canonical Form").

We now have Theorem 8 which is the analogue of Theorem 6.

Theorem 8: With respect to the notation and assumptions of this section, the following results hold:

(a) U_1 is invertible and
$X = U_2U_1^{-1}$ is a solution of the Riccati equation (1)
with $X = X^T \geq 0$ ($X = X^T > 0$ if (F,C) completely reconstructible)

(b) $\sigma(S)$ = "closed-loop" spectrum

$$= \sigma(F - G_1(G_2 + G_1^T X G_1)^{-1} G_1^T X F)$$

$$= \sigma(F - G(X^{-1} + G)^{-1} F) \text{ when } X \text{ is invertible}$$

$$= \sigma(F - GF^{-T}(X - H)) \text{ when } F \text{ is invertible.}$$

Proof: The proof is essentially the same as in Theorem 6. The only difference lies in the fact that we are now dealing with generalized Schur vectors and the matrix S is in "Real Schur Form". We omit the details.

4. NUMERICAL CONSIDERATIONS

In this section we describe how the two approaches for the solution of the discrete-time algebraic Riccati equation can be implemented. We are also going to compare the two approaches and discuss their advantages over other methods for solution of the same equation.

4.1 Algorithm Implementations

In both approaches there are two steps. The first is to find a matrix U, containing the generalized eigenvectors or the generalized Schur vectors, such that

$$MU = LUS$$

where S is a stable nxn matrix in Jordan form and real Schur form, respectively. The second step is the solution of the n^{th} order linear matrix equation

$$XU_1 = U_2 .$$

Equivalently, since X is symmetric, we can solve the equation

$$U_1^T X = U_2^T$$

to find $X = U_1^{-T} U_2^T = U_2 U_1^{-1}$. Any good linear equation solver can be used for this purpose.

Now we discuss the first step in each of the two methods.

Generalized Eigenvector Approach

One obvious way to find the generalized eigenvectors is to first find the generalized eigenvalues of the problem $Mz = \lambda Lz$ and then to compute the eigenvectors and principal vectors corresponding to the stable eigenvalues using the defining relationships

$$Mz_1 = \lambda L z_1$$

$$(M - \lambda L) z_k = L z_{k-1}, \qquad k = 1, \ldots, n_\lambda ,$$

where n_λ is the multiplicity of the eigenvalue λ. This method of finding the eigenvectors is only generally reliable for hand calculation, however. A numerical computation of the eigenvectors can be attempted using the following sequence of subroutines: QZHES, QZIT, QZVAL, and QZVEC. The above subroutines are available in [11].

The above computation works well, however, only when we do not have multiple or near multiple eigenvalues. In the case of multiple eigenvalues there is no reliable method for the machine computation of the generalized principal vectors. Since this case arises frequently, it is preferable to use the real Schur vector approach.

Real Schur Vector Approach

The implementation of the real Schur vector approach consists of the computation of orthogonal matrices which transform the matrix M to real Schur form and L to upper triangular form, in such a way that the diagonal blocks corresponding to the stable eigenvalues are in the upper left quarters of the matrices.

The following sequence of subroutines:

QZHES, QZIT, and QZVAL

reduces M to real Schur form and L to upper triangular form, but the order in which the eigenvalues (i.e., the ratios which determine them) appear is arbitrary. The matrix Z and the reduced matrices QMZ and QLZ, as well as the quantities whose ratios give the generalized eigenvalues, are included in the output of the above sequence of subroutines. Note that the matrix Q is not available, but it is not needed.

Our task is now to effect the reordering of the eigenvalues. We need subroutines to find orthogonal transformations that exchange adjacent 1x1 with 1x1, 1x1 with 2x2, and 2x2 with 2x2 blocks. The 1x1 blocks correspond to real eigenvalues and the 2x2 blocks to complex conjugate eigenvalues. Each of these subroutines takes, as input, the improperly ordered matrices QMZ and QLZ from the previous subroutines and also the accumulated transformations Z. It then effects the appropriate orthogonal transformations Q' and Z' that do the necessary exchange and produces, as output, the new matrices Q'QMZZ', Q'QLZZ', and ZZ'.

Once we have these subroutines the reordering can be done in the following way. Check if the last eigenvalue in the upper left quarter is stable. If it is not move it to the last position in the lower right quarter. Check the next eigenvalue in the upper left quarter and if it is unstable move it to the next position in the lower right quarter. Continue this process until n eigenvalues have been moved, or n stable eigenvalues have been found in the upper left quarter. This process requires at most n^2 exchanges.

When the reordering is finished, the desired U matrix with the real Schur vectors is the matrix consisting of the first n columns of Z.

4.2 Comparisons

From the implementation of the two approaches it becomes apparent that the generalized eigenvector approach is only of theoretical interest, and is convenient only for hand calculation, when the order of the matrices M and L is small.

The real Schur vector approach is numerically very attractive, because it uses only orthogonal transformations which are numerically stable. Moreover, it is not affected by the existence of multiple eigenvalues, which causes difficulties in the generalized eigenvector approach.

The two approaches considered here for the solution of the discrete-time algebraic Riccati equation are direct generalizations of the eigenvector and the Schur vector approach when the matrix F is restricted to be nonsingular. Hence they have the same advantages over other methods of solution of the discrete-time algebraic Riccati equation [6]. The additional advantage of the methods presented here is that they work when F is singular or nearly singular. Even in the case that F is nonsingular the Schur vector approach presented here may be more attractive numerically then other approaches because it avoids a matrix inversion and a few matrix multiplications. The necessity to form F^{-1} when F is nonsingular but badly conditioned with respect to inversion may cause severe difficulties. The above remarks are analogous to caveats associated with attempting to convert the generalized eigenvalue problem $Ax = \lambda Bx$ into the eigenvalue problem $B^{-1}Ax = \lambda x$ even when B is nonsingular; see [8].

5. EXAMPLES

In this section was give a few examples to illustrate the application of both the eigenvector and the Schur vector approach. All computing was done on an IBM 370/168 in double precision arithmetic.

Example 1: This is a very simple example to illustrate the generalized eigenvector approach. We want to solve equation (1) when

$$F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad G_1 = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}, \quad G_2 = 1 .$$

Then G becomes

$$G = G_1 G_2^{-1} G_1^T = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} .$$

We solve the equation $\det(M-\lambda L) = 0$ to find the finite eigenvalues $\lambda_1 = 0$, $\lambda_2 = \frac{1}{2}$, $\lambda_3 = 2$. The fourth eigenvalue is an infinite one, in accordance with the results of section 2.2. The eigenvectors corresponding to the two stable eigenvalues can be put in a matrix U.

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 0 & 2 \\ 1 & 2 \\ -1 & -1 \end{pmatrix}$$

We can now find the solution

$$X = U_2 U_1^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 3/2 \end{pmatrix} .$$

It can easily be verified that this solution is positive definite and satisfies equation (1). It can also be checked that the pair (F,G_1) is stabilizable (controllable, in fact) and the pair (F,C) is completely reconstructible which explains the positive definiteness of the solution.

This problem can also be solved using the Schur vector approach, which gives the same solution with an accuracy of 15 significant digits. (For the details of the implementation see example 2.)

<u>Example 2</u>: This is a more complicated example, which corresponds to a real world problem; see [4]. We want to solve equation (1) when

$$F = \begin{pmatrix} \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$G_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad G_2 = \frac{1}{4}, \quad \text{and } G := G_1 G_2^{-1} G_1^T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

First we solve equation (1) by the eigenvector approach. Solving the equation $\det(M-\lambda L) = 0$ we find the eight eigenvalues of the problem:

$$\lambda_1 = \lambda_2 = \lambda_3 = 0, \quad \lambda_4 = \frac{21-5\sqrt{17}}{4}, \quad \lambda_5 = \frac{21+5\sqrt{17}}{4}, \quad \lambda_6 = \lambda_7 = \lambda_8 = \infty .$$

The matrix containing the four eigenvectors corresponding to the stable eigenvalues is:

$$U = \begin{pmatrix} U_1 \\ \\ U_2 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 2 & 1 \\ -1 & -1 & 1 & (19-5\sqrt{17})/4 \\ 0 & 1 & 1 & (103-25\sqrt{17})/2 \\ 0 & 0 & -1 & (536-130\sqrt{17}) \\ 2 & -3 & 7/2 & (13+5\sqrt{17})/32 \\ 0 & -2 & 3 & (-19+5\sqrt{17})/16 \\ 0 & 0 & 2 & (-103+25\sqrt{17})/8 \\ 0 & 0 & 0 & (-268+65\sqrt{17})/2 \end{pmatrix}$$

Note that the eigenvectors satisfy

$$\begin{aligned} (M-\lambda_1 L)u_1 &= 0 \\ (M-\lambda_2 L)u_2 &= Lu_1 \\ (M-\lambda_3 L)u_3 &= Lu_2 \\ (M-\lambda_4 L)u_4 &= 0 \quad . \end{aligned}$$

We can now find $X = U_2U_1^{-1}$ by solving the equation $XU_1 = U_2$, or equivalently, since X is symmetric, $U_1^TX = U_2^T$.

The solution (rounded to 10 significant figures) is:

$$X = \begin{pmatrix} 1.328913832 & 0.6578276642 & 0.3156553284 & 0.1313106567 \\ & 1.315655328 & 0.6313106567 & 0.2626213135 \\ & & 1.262621313 & 0.5252426270 \\ & \text{symmetric} & & 1.050485254 \end{pmatrix}$$

This solution gives a residual matrix $R = (r_{ij})$ (when substituted into the Riccati equation) whose elements vary between $r_{11} \approx 2.60 \times 10^{-13}$ and $r_{44} \approx 1.21 \times 10^{-11}$. Obviously the hand calculation of the eigenvectors for a higher dimensional problem would be too complicated to attempt.

Now, we present the solution of (1) for this example by the Schur vector approach. First we use the sequence of subroutines

QZHES, QZIT, and QZVAL

to reduce M and L to upper triangular forms. (Note that all the eigenvalues are real and so the reduced M is upper triangular.) Also we get as a by-product the closed loop eigenvalues. Then we do the reordering of the eigenvalues. Finally we obtain the matrix U as the first n columns of Z and solve the system $U_1^T X = U_2^T$ for X. This computed solution equals, up to 10 significant digits, the solution obtained by the generalized eigenvector approach.

<u>Example 3</u>: This is an example which has a very simple closed-form solution for arbitrary dimension n. Numerical accuracy of our method can thus easily be checked.

We want to solve (1) when

$$F = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad H = I,$$

$$G_1 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad G_2 = 1, \quad \text{and } G := G_1 G_2^{-1} G_1^T = \begin{pmatrix} 0 & & & & \\ & \ddots & & \text{\huge 0} & \\ & & \ddots & & \\ & \text{\huge 0} & & 0 & \\ & & & & 1 \end{pmatrix}$$

Solving equation (1) by the eigenvector approach we find that $\lambda_1 = \ldots = \lambda_n = 0$, $\lambda_{n+1} = \ldots = \lambda_{2n} = \infty$. The matrix of the stable eigenvectors can be explicitly determined as

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{pmatrix}$$

Note that the eigenvectors satisfy:

$$\begin{aligned} Mu_1 &= 0 \\ Mu_2 &= Lu_1 \\ &\vdots \\ Mu_n &= Lu_{n-1} \quad . \end{aligned}$$

We now easily find that

$$X = U_2U_1^{-1} = \begin{pmatrix} 1 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & n \end{pmatrix} .$$

When $n = 10$ the Schur vector approach gives a numerical solution determined accurately to at least 13 correct decimal places.

6. CONCLUSIONS

We have presented two versions of a new method for the solution of discrete-time algebraic Riccati equations. The main feature of this new method is that it does not require inversion of the state transition matrix of the corresponding LQG problem. Thus it is directly applicable to problems with singular transition matrices and elminates the numerical problems that other algorithms may have when the transition matrix is ill-conditioned with respect to inversion.

The first version of the new method is a generalization of the classical eigenvector approach and is mainly of theoretical interest because of the numerical hazards associated with the calculation of the generalized eigenvectors, and especially the generalized principal vectors in the case of multiple generalized eigenvalues. The second version is a generalization of the Schur vector approach and is numerically very attractive because it uses orthogonal transformations which are numerically stable.

The implementation of the two approaches was discussed in some detail. What remains to be done is the implementation of the orthogonal transformations that do the reordering in the case of complex conjugate pairs of eigenvalues.

ACKNOWLEDGMENT

This paper was motivated by discussions with Mr. W.L. Bialkowski of Domtar, Inc., Canada; see [4].

REFERENCES

[1] Athans, M., The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design, IEEE Trans. Aut. Contr., AC-16(1971), 529-552.

[2] Vaughn, D.R., A Nonrecursive Algebraic Solution for the Discrete Riccati Equation, IEEE Trans. Aut. Contr., AC-15(1970), 597-599.

[3] Gould, L.A., Chemical Process Control: Theory and Applications, Addison-Wesley, Reading, MA, 1969.

[4] Bialkowski, W.L., Application of Steady State Kalman Filters - Theory with Field Results, Proc. JACC, Philadelphia, Oct. 1978.

[5] Anderson, B.D.O., and J.B. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, 1979.

[6] Laub, A.J., A Schur Method for Solving Algebraic Riccati Equations, LIDS Report No. LIDS-R-859, Lab. for Info. and Decis. Syst., Cambridge, MA, Oct. 1978.

[7] Stewart, G.W., Introduction to Matrix Computations, Academic Press, New York, 1973.

[8] Moler, C.B., and G.W. Stewart, An Algorithm for Generalized Matrix Eigenvalue Problems, SIAM J. Numer. Anal., 10(1973), 241-256.

[9] Golub, G.H., and J.H. Wilkinson, Ill-Conditioned Eigensystems and the Computation of the Jordan Canonical Form, SIAM Rev., 18(1976), 578-619.

[10] Stewart, G.W., On the Sensitivity of the Eigenvalue Problem $Ax = \lambda Bx$, SIAM J. Numer. Anal., (1972), 669-686.

[11] Garbow, B.S., et. al., Matrix Eigensystem Routines - EISPACK Guide Extension, Lect. Notes in Comp. Sci., Vol. 51, Springer-Verlag, New York, 1977.