

Continuation Methods for Solving Modified Discrete-Time Algebraic Riccati Equations

Wen-Wei Lin, Chern-Shuh Wang, and Chuan-Hsiang Han

Abstract—It is well known that the continuation methods have been successfully applied to solve polynomial systems and fixed point problems, etc. In this paper, we consider a discrete-time algebraic Riccati equation with an admissible, low rank, and symmetric perturbation. Our attention will be directed primarily to this modified discrete-time algebraic Riccati equation and the numerical method for its solution based on proceeding along the continuation path.

I. INTRODUCTION

Discrete-time algebraic Riccati equations arise in modeling discrete-time systems in many fields of applications. They often play an important role in control theory, e.g., LQG problem and H^∞ control theory. The ordinary discrete-time algebraic Riccati equation (DARE) is given by

$$R(X) \equiv F^T X F - X - F^T X G_1 (G_2 + G_1^T X G_1)^{-1} \times G_1^T X F + K = 0 \quad (1)$$

where $F, X, K = K^T \geq 0 \in \mathbb{R}^{n \times n}$, $G_1 \in \mathbb{R}^{n \times m}$, $G_2 \in \mathbb{R}^{m \times m}$ ($m \leq n$), and $G_2 = G_2^T > 0$. Under stability and detectability assumptions [10], DARE (1) is known to have a unique nonnegative definite symmetric solution [10], [16]. Define $G \equiv G_1 G_2^{-1} G_1^T$. Then, the coefficient matrices of DARE (1) can be grouped to form a $2n \times 2n$ symplectic pencil

$$N - \lambda L \equiv \begin{bmatrix} F & 0 \\ -K & I \end{bmatrix} - \lambda \begin{bmatrix} I & G \\ 0 & F^T \end{bmatrix}.$$

It is known [16] that if the columns of $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in \mathbb{R}^{2n \times n}$ span the invariant subspace of $N - \lambda L$ associated with its eigenvalues inside the unit circle, then the unique positive semidefinite solution X to the DARE (1) is given by $X = X_2 X_1^{-1}$. To solve DARE efficiently, so far, there are several elegant numerical methods had been proposed [1], [2], [6], [7], [13], [14], and [16].

In this paper, we present a method for the numerical solution of modified discrete-time algebraic Riccati equations (MDARE)

$$R(X) + P(X) = 0 \quad (2)$$

where $P: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ is differentiable, symmetric in X (i.e., $P(X^T) = P(X)^T$), and $P(X)$ is of low rank for all $X \in \mathbb{R}^{n \times n}$. Here, we remark that assumptions on MDARE such that it possesses a unique stabilizing solution remain an open question [3]. Examples of perturbations P may be encountered in practice by solving the reduced-order state-estimation problem or reduced-order dynamic-compensation problem [3]. Suppose that the matrix F in (1) is stable and the reduced-order state-estimation problem is solvable. Then, the solution can be described by $n \times n$ nonnegative-definite matrices X , \hat{X} , and \hat{Y} satisfying the following system of equations

$$X = F^T X F - F^T X G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F + K + (I - \tau) \hat{X} (I - \tau^T), \quad (3)$$

$$\hat{X} = F^T \tau \hat{X} \tau^T F + F^T X G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F, \quad (4)$$

$$\hat{Y} = F_1^T \tau^T \hat{Y} \tau F_1 + \Sigma(X) \quad (5)$$

Manuscript received October 29, 1993; revised October 16, 1994.
The authors are with the Institute of Applied Mathematics, National Tsing Hua University, Hsinchu, Taiwan 30043; R.O.C.
IEEE Log Number 9408801.

where

$$\tau \equiv \sum_{i=1}^e \psi E_i \psi^{-1} \quad (\text{unit-rank eigenprojection}) \quad (6)$$

for some nonsingular ψ such that $\psi^{-1} \hat{X} \psi$ is diagonal with rank e , $F_1^T \equiv F - G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F$, $\Sigma(X)$ is a suitable matrix rational polynomial in X (see [3] for details) and E_i is a matrix with 1 in the (i, i) th position and zeros elsewhere. We now assume that I) $\rho(F^T \tau) < 1$, i.e., all eigenvalues of $F^T \tau$ are in the unit circle or II) $(I - \tau)(F^T \tau)^{d_0} = 0$ for some $d_0 \geq 1$. Then we substitute (4) into (3) repeatedly and get the perturbed modified discrete-time algebraic Riccati equation (2) with

$$P(X) = \sum_{k=0}^d (I - \tau)(F^T \tau)^k F^T X G_1 (G_2 + G_1^T X G_1)^{-1} \times G_1^T X F (\tau^T F)^k (I - \tau^T). \quad (7)$$

Here $d = \infty$ for the first assumption and $d = d_0 - 1$ for the second assumption. For the first case, we always take a finite sum approach in practice by truncating the high order terms of $P(X)$. For the reduced-order dynamic-compensation problem [3], it can also be treated in the same way as above.

A standard iterative method for the solution of the system of (3)–(5) proposed by [3] iterates τ , \hat{X} , \hat{Y} , and X until convergence. In some practical cases the matrix τ will converge faster than \hat{X} and \hat{Y} . Thus, for this case, to solve MDARE (2) directly is more efficient than to solve (3)–(5). In general, the matrix τ in (6) is an oblique (nonsymmetric) projection matrix (i.e., $\tau = \tau^2$). Throughout this paper we assume that $(I - \tau)$ is of low rank, then $P(X)$ is also of low rank.

In this paper, we construct a new continuation equation for the solution of the MDARE. A standard approach for the continuation method is to employ a predictor-corrector method [18] for following the continuous curve constructed by the continuation equation $H(X, s) = 0$ from $s = 0$ to $s = 1$. Furthermore, each step of the continuation method requires the solution of a sequence of discrete-time Lyapunov equations [8], [9] to obtain the solution of a modified discrete-time Lyapunov equation of the form (see Section III for details)

$$A^T X A - X + \sum_{k=0}^d (B_k X C_k + (B_k X C_k)^T) + V = 0. \quad (8)$$

Briefly, the rest of this paper is organized as follows. In Section II, we construct a continuation mapping with MDARE as the desired problem. Many of computations for the predict-correct process to proceed along the continuation curve are presented in Section III. The numerical algorithm and its implements are given in Section IV.

II. CONSTRUCTING THE CONTINUATION EQUATION FOR SOLVING THE MODIFIED DISCRETE-TIME ALGEBRAIC RICCATI EQUATION

In this section, we give a continuation equation $H(X, s) = 0$ to construct a continuation curve monotone in parameter s such that $H(X, 0) \equiv H_0(X) = 0$ is an easy problem and $H(X, 1) \equiv H_1(X) = 0$ is the desired problem MDARE, i.e.,

$$H(X, s) = (1 - s)H_0(X) + sH_1(X). \quad (9)$$

In general, a desired solution of MDARE is symmetric [3]. If we want to find a desired solution of MDARE, then for any fixed $s \in [0, 1]$, the continuation map $H(\cdot, s)$ can be a function mapping from $S^{n \times n}$ into $S^{n \times n}$, where $S^{n \times n} = \{X \in \mathbb{R}^{n \times n} \mid X = X^T\}$.

To ensure the existence of continuation curve $H(X, s) = 0$ in (9) by applying Sard's Theorem [15], the regularity of $H(X, s) = 0$ must be considered here. From (1), (3), and the Cramer's rule, it is obvious that the equation found at the bottom of the page is a rational polynomial function in X . Here $\text{adj}(G_2 + G_1^T X G_1)$ is the adjoint matrix of $(G_2 + G_1^T X G_1)$. Thus, we can not ensure the existence of the resulting continuation curve. To avoid this flaw, we first consider the following defective set

$$W_1 = \{X \in S^{n \times n} \mid \det(G_2 + G_1^T X G_1) = 0\}.$$

It is observed that the set W_1 is closed and has measure zero. Now, to guarantee that the constructed continuation curve can be parameterized in the parameter s , we consider

$$W_2 = \{X \in S^{n \times n} \mid \det[D_X(\det(G_2 + G_1^T X G_1) \times (R(X) + P(X)))] = 0\}.$$

Here D_X denotes the differentiation of the variables X . Also, W_2 is closed and measure zero. Now, let $U = S^{n \times n} \setminus (W_1 \cup W_2)$ be an open set in $S^{n \times n}$. We can construct the desired continuation mapping monotone in the parameter s which is defined in the domain $U \times [0, 1]$ (see [12] for details). Before the construction, we consider an auxiliary continuation equation. Define $\Phi: U \times [0, 1] \rightarrow S^{n \times n}$ by

$$\Phi(X, s) = \det(G_2 + G_1^T X G_1) \{ (R(X) + P(X)) + (s-1)(P(X_0) - Z) \} \quad (10)$$

where $Z \in S^{n \times n}$ is an arbitrary nonnegative definite matrix and $X_0 \in S^{n \times n}$ is the unique, nonnegative definite solution of the new DARE

$$R(X) + Z = 0. \quad (11)$$

Since Φ is defined on $U \times [0, 1]$, we have $D_X \Phi = D_X(\det(G_2 + G_1^T X G_1)(R(X) + P(X)))$ is nonsingular. This implies that $\Phi(X, s) = 0$ can be parameterized in s , i.e., $\Phi(X, s) = 0$ represents a continuation curve monotone in s .

Therefore, we construct the desired continuation equation as

$$H(X, s) \equiv \Phi(X, s) / \det(G_2 + G_1^T X G_1) = (R(X(s)) + P(X(s))) + (s-1) \times (P(X_0) - Z) = 0. \quad (12)$$

Note that in (12), it is observed that the easy problem $H_0(X) = R(X) + P(X) - (P(X_0) - Z)$ has a solution X_0 satisfying DARE (11).

Remark 2.1: From the similar results proposed by [5], [19], we know that by utilizing an arbitrary choice Z , the continuation curve $\Phi(X, s) = 0$ in (10) exists with probability one when Φ is defined on $U \times [0, 1]$.

Remark 2.2: Although the set $W_1 \cup W_2$ is measure zero. In practice, we may encounter ill conditioned matrices which are singular to working precision. In such case, the continuation curve possibly blows up as s approaches to one. In [12], however, a restarting strategy is developed to obtain a bounded continuation curve.

III. NUMERICAL APPROACH FOR CONTINUATION CURVE

One of the most effective ways to follow the continuation curve is known as Euler-Newton continuation. That is, if (X_i, s_i) is considered as an initial point of the continuation curve for some $s_i \in [0, 1]$, where $X_i \equiv X(s_i)$, then the next point $(X_{i+1}, s_{i+1}) \equiv (X(s_i + \Delta s), s_i + \Delta s)$ is obtained by

- i) Euler step (Predict step): $X_{i+1,0} = X_i + \dot{X}_i \cdot \Delta s$;
- ii) Newton step (Correct step): $X_{i+1,j+1} = X_{i+1,j} + \Delta_{i+1,j+1}$, for $j = 0, 1, \dots$.

Here $\dot{X}_i \equiv \frac{dX(s_i)}{ds}$ and $\Delta_{i+1,j+1}$ is the $j+1$ th Newton correction. As the sequence $\{X_{i+1,j}\}$ converges, we set $X_{i+1} = \lim_{j \rightarrow \infty} X_{i+1,j}$. Repeat the process, if each Euler-Newton step is convergent, then we obtain a desired solution $X(1)$ as s approaches one. The remaining question is how to compute the predict factor \dot{X}_i and the correction factor $\Delta_{i,j+1}$ at each Euler-Newton step.

A. Calculation of Predict Factor \dot{X}_i

Because the following calculation is independent of the subscript of \dot{X}_i , we omit the subscript "i". We first derive the formula for \dot{X} by differentiating (12) with respect to s . This yields

$$(F^T X G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T - F^T) \dot{X} \times (G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F - F) - \dot{X} + \frac{d}{ds} P(X(s)) + P(X_0) - Z = 0. \quad (13)$$

To find $\frac{d}{ds} P(X(s))$ explicitly, we have

$$\frac{d}{ds} P(X(s)) = (I - \tau) \left[\sum_{k=0}^d L_k \dot{X} \Gamma_k + \sum_{k=0}^d (L_k \dot{X} \Gamma_k)^T \right] (I - \tau^T) \quad (14)$$

where

$$L_k = (F^T \tau)^k \left(F^T - \frac{1}{2} F^T X G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T \right), \quad (15)$$

$$\Gamma_k = G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F (\tau^T F)^k, \quad \text{for } k = 0, 1, \dots, d. \quad (16)$$

Let $\hat{F} = G_1 (G_2 + G_1^T X G_1)^{-1} G_1^T X F - F$. Combining (13) and (14), we have

$$\hat{F}^T \dot{X} \hat{F} - \dot{X} + (I - \tau) \left[\sum_{k=0}^d L_k \dot{X} \Gamma_k + \sum_{k=0}^d (L_k \dot{X} \Gamma_k)^T \right] \times (I - \tau^T) + P(X_0) - Z = 0. \quad (17)$$

To solve (17), we have to extend some calculated results proposed Richter *et al.* [17] to the discrete-times case. By spectrum decomposition, we have $I - \tau = T \cdot \text{block.diag} \{0, \dots, 0, I_{l \times l}\} \cdot T^{-1}$, where $l = \text{rank}(I - \tau)$. It is easily seen that $l \ll n$. Let $m = \frac{l(l+1)}{2}$. Now, we consider the following definitions: Define $\pi: S^{n \times n} \rightarrow \mathfrak{R}^m$ such

$$H_1(X) = R(X) + P(X) = \left[F^T X F - X - \left(\frac{F^T X G_1 \text{adj}(G_2 + G_1^T X G_1) G_1^T X F}{\det(G_2 + G_1^T X G_1)} \right) + K \right] + \sum_{i=0}^k \left[\frac{(I - \tau)(F^T \tau)^i F^T X G_1 \text{adj}(G_2 + G_1^T X G_1) G_1^T X F (\tau^T F)^i (I - \tau^T)}{\det(G_2 + G_1^T X G_1)} \right]$$

that $\pi(W) = \text{vec}(W(n-l+1:n, n-l+1:n))$ with $W \in S^{n \times n}$, and $\Pi: \mathfrak{R}^m \rightarrow S^{n \times n}$ such that

$$\Pi(q) \equiv T \cdot \text{block.diag} \left\{ 0, \dots, 0, \begin{bmatrix} q_1 & \dots & q_l \\ \vdots & & \vdots \\ q_l & \dots & q_m \end{bmatrix} \right\} \cdot T^{-1},$$

$$q \equiv \begin{bmatrix} q_1 \\ \vdots \\ q_m \end{bmatrix} \in \mathfrak{R}^m.$$

As in [17] it is first necessary to solve discrete-time Lyapunov equations

$$\begin{aligned} \hat{F}^T X^{(j)} \hat{F} - X^{(j)} + V^{(j)} &= 0 \text{ with } V^{(j)} \\ &\equiv \begin{cases} V & j = 0 \\ \Pi(e_j) & j = 1, \dots, m. \end{cases} \end{aligned} \quad (18)$$

Consequently, we construct a small order $m (\ll n^2)$ linear equation

$$(I - M)y = d \quad (19)$$

with $\text{col}_j(M) = \pi(X^{(j)})$, $j = 1, \dots, m$, and $d = \pi(X^{(0)})$. Once the solution of this linear equation is computed, a desired solution \hat{X} of (17) is found by solving a standard discrete-time Lyapunov equation

$$\hat{F}^T X \hat{F} - X + \Pi(y) + V = 0.$$

To solve $m+1$ Lyapunov equations in (18) by using a standard solution method [8], it requires about $(m+16)n^3$ flops totally. In addition, to solve the linear system (19) it requires about $\frac{1}{2}m^3$ flops. This shows that we can obtain \hat{X} in an economical way as $\frac{d}{ds}P(X(s))$ is low rank with $m \ll n$.

Remark 3.1: For large sparse problems, the use of the above solution method is not feasible in our continuation algorithm, because the triangularization of \hat{F} causes excessive computational requirements. As the rapid numerical solution of the continuous-time Lyapunov equation proposed by Richter *et al.* [18] we can use FRPI [9] iteration or block-diagonalization methods with parallel implementation techniques to solve the discrete-time Lyapunov equation (18).

B. Calculation of Correction Factor $\Delta_{i,j+1}$

In this subsection, we find the $j+1$ th Newton correction factor $\Delta_{i,j+1}$ so that $X_{i,j+1} = X_{i,j} + \Delta_{i,j+1}$ at the i th step Euler-Newton iteration. For convenience, we omit the subscripts, and set $\bar{X} = X_i = \lim_{j \rightarrow \infty} X_{i,j}$, $X = X_{i,j}$, $\bar{\Delta} = \Delta_{i,j}$ and $\Delta = \bar{X} - X$. Since $\bar{X} = X + \Delta$ can be considered as an exact solution of (12), we have

$$\begin{aligned} &F^T(X + \Delta)F - (X + \Delta) - F^T(X + \Delta)G_1 \\ &\times (G_2 + G_1^T(X + \Delta)G_1)^{-1}G_1^T(X + \Delta)F \\ &+ K + P(X + \Delta) + (s-1)(P(X_0) - Z) = 0. \end{aligned} \quad (20)$$

Expanding the matrix $(G_2 + G_1^T(X + \Delta)G_1)^{-1}$, we get

$$\begin{aligned} &(G_2 + G_1^T(X + \Delta)G_1)^{-1} \\ &= \hat{G}_2^{-1} - \hat{G}_2^{-1}G_1^T \Delta G_1 \hat{G}_2^{-1} + \text{H.O.T.} \end{aligned} \quad (21)$$

where $\hat{G}_2 \equiv G_2 + G_1^T X G_1$ and H.O.T. denotes a function in higher order ($n \geq 2$) terms of Δ . Substitute (21) into $P(X + \Delta)$, then

$$\begin{aligned} &P(X + \Delta) \\ &= \sum_{k=0}^d (I - \tau)(F^T \tau)^k F^T \\ &\times [X G_1 \hat{G}_2^{-1} G_1^T X + \Delta G_1 \hat{G}_2^{-1} G_1^T X \\ &+ X G_1 \hat{G}_2^{-1} G_1^T \Delta - X G_1 \hat{G}_2^{-1} G_1^T \Delta G_1 \hat{G}_2^{-1} G_1^T X] \\ &\times F(\tau^T F)^k (I - \tau^T) + \text{H.O.T.} \end{aligned} \quad (22)$$

Substitute (21) and (22) into (20) and omit the higher order terms of Δ , then the desired correction factor $\bar{\Delta}$ solves the following equation

$$\begin{aligned} &\bar{F}^T \bar{\Delta} \bar{F} - \bar{\Delta} + (I - \tau) \cdot \left[\sum_{k=0}^d \bar{L}_k \bar{\Delta} \bar{\Gamma}_k + \sum_{k=0}^d (\bar{L}_k \bar{\Delta} \bar{\Gamma}_k)^T \right] \\ &\cdot (I - \tau^T) + \Theta(X, s) = 0 \end{aligned} \quad (23)$$

where $\bar{F} = F - G_1 \hat{G}_2^{-1} G_1^T X F$, $\bar{L}_k = (F^T \tau)^k F^T X G_1 \hat{G}_2^{-1} G_1^T$, and $\bar{\Gamma}_k = (F - 1/2 G_1 \hat{G}_2^{-1} G_1^T X F)(\tau^T F)^k$, for $k = 0, 1, \dots, d$, and $\Theta(X, s) = F^T X F - X - F^T X G_1 \hat{G}_2^{-1} G_1^T X F + \sum_{k=0}^d (I - \tau)(F^T \tau)^k F^T X G_1 \hat{G}_2^{-1} G_1^T X F (\tau^T F)^k (I - \tau^T) + K + (s-1)(P(X_0) - Z)$. Obviously, (23) is also a modified Lyapunov equation and the method described in subsection 3.1 can be utilized to solve for $\bar{\Delta}$.

IV. ALGORITHM AND NUMERICAL RESULTS

Based on the discussion of the last section, we give an algorithm for solving low-rank, symmetric perturbed MDARE.

Algorithm 4.1:

Step 0) (Input) Given two criterion bounds *eps1*, *eps2*, a positive number k_0 (the maximum iteration number) and an arbitrary nonnegative definite symmetric matrix Z . Set $i = 0$ and $s_0 = 0$.

Step 1) (Find the solution of the easy problem of continuation equation) Solve DARE (11) by using the QZ method proposed by Pappas *et al.* [16].

Step 2) (Euler-Newton step) Repeat:

2.1: (Euler step)

- Compute predict factor \hat{X}_i in (17).
- Find $h_i > 0$ such that $\frac{\|H(X_i + h_i \hat{X}_i, s_i + h_i)\|_F}{\|X_i + h_i \hat{X}_i\|_F} \leq \text{eps1}$. (Here $\|\cdot\|_F$ denotes the Frobenius norm.)
- Set $X_{i+1,0} = X_i + h_i \hat{X}_i$, $j = 0$.

2.2: (Newton step) Iterate until convergence criterions hold.

- Compute correction factor $\Delta_{i+1,j+1}$ in (23).
- Set $X_{i+1,j+1} = X_{i+1,j} + \Delta_{i+1,j+1}$.
- If $\frac{\|H(X_{i+1,j+1}, s_i + h_i)\|_F}{\|X_{i+1,j+1}\|_F} \leq \text{eps2}$, then set $s_{i+1} = s_i + h_i$;
If $s_{i+1} = 1$, then accept $X(1) = X_{i+1,j+1}$ to be the solution of MDARE (2), and stop.
Else set $X_{i+1} = X_{i+1,j+1}$, $i = i + 1$, go to **Repeat**.
- Set $X_{i+1,j} = X_{i+1,j+1}$, $j = j + 1$.
If $j \geq k_0$, then find $h_i > 0$ such that $\frac{\|H(X_i + h_i \hat{X}_i, s_i + h_i)\|_F}{\|X_i + h_i \hat{X}_i\|_F} \leq \text{eps1}$ and go to \odot .
- Goto 2.2.

Note that to determine the step size h_i in Algorithm 4.1, there are many efficient strategies proposed for the adaptive control [11]. For example, at i th step Euler-Newton iteration, we give the initial step size $h_i = 1 - s_i$. If the quantity $r_i \equiv \frac{\|H(X_i + h_i \hat{X}_i, s_i + h_i)\|_F}{\|X_i + h_i \hat{X}_i\|_F}$ is larger than *eps1*, then we cut h_i by half. Repeat this check-updating process until $r_i \leq \text{eps1}$ holds (see [11] for details).

Three examples are given to illustrate the numerical behavior by performing Algorithm 4.1. The first two examples are given by Pappas *et al.* [16] but with some slight modifications of the coefficient matrix F so that the first condition I) $\rho(F^T \tau) < 1$ holds in Example 4.1 and the second condition II) $(I - \tau)(F^T \tau)^k = 0$ with $k = 2$ holds in Example 4.2. In the third example, the coefficient matrices are randomly generated and the second condition II) $(I - \tau)(F^T \tau)^k = 0$ with $k = 1$ holds. All computations are done in MATLAB on a PC-486 computer in double precision arithmetic.

TABLE I

i-th step	$\sum_{k=0}^{i-1} h_k$	iteration no.	$d(X_i)$	$\lambda_{\min}(X_i)$	$\frac{\ H(X_i, s_i)\ _E}{\ X_i\ _F}$
0	0.0	0	1.88e-15	2.43e-07	1.0525e-15
1	1.0	2	1.18e-16	1.86e-03	1.3942e-16

TABLE II

i-th step	$\sum_{k=0}^{i-1} h_k$	iteration no.	$d(X_i)$	$\lambda_{\min}(X_i)$	$\frac{\ H(X_i, s_i)\ _E}{\ X_i\ _F}$
0	0.0	0	1.44e-13	1.0	2.0155e-15
1	1.0	2	6.67e-15	1.0	4.0877e-16

TABLE III

i-th step	$\sum_{k=0}^{i-1} h_k$	iteration no.	$d(X_i)$	$\lambda_{\min}(X_i)$	$\frac{\ H(X_i, s_i)\ _E}{\ X_i\ _F}$
0	0.0	0	4.08e-11	2.57e-01	8.2842e-14
1	0.125	3	7.21e-12	3.40e-01	8.7100e-13
2	0.5625	3	8.21e-12	4.56e-01	7.3968e-13
3	1.0	2	8.95e-12	4.90e-01	1.5861e-09

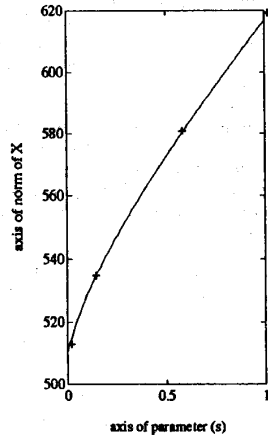


Fig. 1. Norm of X versus s .

Example 4.1: Let $G_1 = [0, 0, 0, 1]^T$, $G_2 = 1/4$, $G := G_1 G_2^{-1} G_1^T$

$$F = Q \cdot \begin{bmatrix} 0.5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot Q^T,$$

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\tau = Q \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot Q^T$$

where

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

Here, two criterion bounds $eps1$ and $eps2$ are 0.1 and 10^{-13} , respectively. Table I shows that the computed solution of MDARE (2) is obtained with a small residual ($< 10^{-14}$) when the continuation

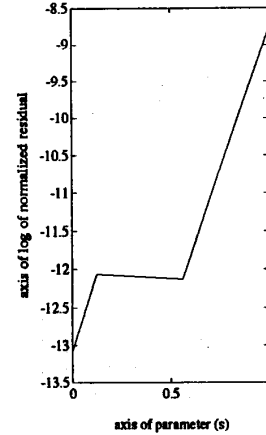


Fig. 2. Log of residual versus s .

parameter s is attained to one. The first column of Table I records the i th Euler step of the Euler-Newton iteration. The second column of Table I records the total length $\sum_{k=0}^{i-1} h_k$ of all Euler steps before the i th Euler-Newton iteration. The third column records the iterative number of the Newton step in the i th Euler-Newton iteration. The fourth and fifth columns denote the symmetry $d(X_i) \equiv \|X_i - X_i^T\|_F$ of the convergent solution X_i , and the algebraically smallest eigenvalue of X_i , respectively. The last column denotes the normalized residual $\frac{\|H(X_i, s_i)\|_E}{\|X_i\|_F}$ of X_i . From Table I we see that the numerical solution X_1 of MDARE is fairly symmetric and positive definite.

Example 4.2: Let $F = N^{-T} \cdot \begin{bmatrix} J_{6 \times 6} & O_{6 \times 4} \\ O_{4 \times 6} & O_{4 \times 4} \end{bmatrix} \cdot N^T$, $K = I_{10}$, $G_2 = I_2$

$$J_{6 \times 6} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$$G_1 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad G := G_1 G_2^{-1} G_1^T$$

and

$$\tau = N \cdot \begin{bmatrix} I_{6 \times 6} & O_{6 \times 4} \\ \tilde{I}_{4 \times 6} & O_{4 \times 4} \end{bmatrix} \cdot N^{-1},$$

$$\tilde{I}_{4 \times 6} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where $N \in \mathbb{R}^{10 \times 10}$ is a randomly chosen nonsingular matrix. Table II shows the numerical results by performing Algorithm 4.1. The numerical behavior of this example is similar to Example 4.1. The corresponding continuation curves of both two examples seem to be a straight line or a smooth curve with small curvature.

Example 4.3: Let $G_1 \in \mathbb{R}^{20 \times 18}$, $G_2 \in \mathbb{R}^{18 \times 18}$ and $K = K^T \in \mathbb{R}^{20 \times 20}$ be randomly chosen matrices with G_2 positive definite, $\text{rank}(G_1) = 18$ and $\text{rank}(K) = 17$. Let $F = N^{-T} \begin{bmatrix} F_{11} & F_{12} \\ O_{18 \times 2} & F_{22} \end{bmatrix} N^T$ and $\tau = N \begin{bmatrix} O & O \\ O & I_{18 \times 18} \end{bmatrix} N^{-1}$, where $F_{11} \in \mathbb{R}^{2 \times 2}$, $F_{12} \in \mathbb{R}^{2 \times 18}$, $F_{22} \in \mathbb{R}^{18 \times 18}$ and $N \in \mathbb{R}^{20 \times 20}$ are

randomly chosen matrices with N nonsingular. By direct calculation we get

$$(I - \tau)(F^T \tau) = N \begin{bmatrix} I_{2 \times 2} & O \\ O & O \end{bmatrix} \begin{bmatrix} F_{11}^T & O_{2 \times 18} \\ F_{12}^T & F_{22}^T \end{bmatrix} \\ \times \begin{bmatrix} O & O \\ O & I_{18 \times 18} \end{bmatrix} N^{-1} = 0.$$

Here, condition II) is satisfied. The criteria are given by $\text{eps1} = 0.1$ for the step size control and $\text{eps2} = 10^{-6}$ for the convergence of the Newton step, respectively. Table III lists the numerical results. In this example we see that the Newton iteration converges with 2.7 iterations per Euler step on average. Fig. 1 gives the plot of the norms of $X(s)$ v.s. parameter s and denotes the norms of the convergent solutions of Euler-Newton iterations by "+." Fig. 2 gives the plot of logarithm of the normalized residual v.s. parameter s .

As discussion in Section III, to solve (17) or (23) is equivalent to solve $m + 1$ Lyapunov equations and a linear system with order $m = \frac{l(l+1)}{2}$. Hence, it is easily seen that the developed Algorithm 4.1 is more efficient than the direct method by utilizing an enlarged linear system of order n^2 .

REFERENCES

- [1] G. S. Ammar and V. Mehrmann, "On Hamiltonian and symplectic Hessenberg forms," *Lin. Alg. Appl.*, vol. 149, pp. 55-72, 1991.
- [2] A. Y. Barraud, "Investigation autour de la fonction signe d'une matrice, application à l'équation de Riccati," *RAIRO Automat.*, vol. 13, pp. 335-368, 1979.
- [3] D. S. Bernstein, L. D. Davis, and D. C. Hyland, "The optimal projection equations for the reduced-order discrete-time modeling, estimation, and control," *J. Guidance*, vol. 9, no. 3, pp. 288-293, 1986.
- [4] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circ. Syst.*, vol. CAS-25, no. 9, pp. 772-781, 1978.
- [5] S.-N. Chow, J. Mallet-Paret, and J. A. Yorke, "Finding zeros of maps: Homotopy methods that are constructive with probability one," *Math. Comp.*, vol. 31, pp. 887-899, 1978.
- [6] J. D. Gardiner and A. J. Laub, "A generalization of the matrix sign function solution for algebraic Riccati equations," *Int. J. Contr.*, vol. 44, pp. 823-832, 1986.
- [7] A. Bunse-Gerstner, R. Byers, and V. Mehrmann, "A chart of numerical methods for structured eigenvalue problems," *SIAM Matrix Anal. Appl.*, vol. 13, no. 2, pp. 419-453, 1992.
- [8] S. J. Hammaring, "Numerical solution of the stable, nonnegative definite Lyapunov equation," *IMA J. Numer. Anal.*, vol. 2, no. 3, pp. 303-323, 1982.
- [9] A. S. Hodel and K. Poolla, "Parallel solution of large Lyapunov equations," *SIAM Matrix Anal. Appl.*, vol. 13, no. 4, pp. 1189-1203, 1992.
- [10] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley-Interscience, 1972.
- [11] T.-Y. Li and J. A. Yorke, "A simple reliable numerical algorithm for following homotopy paths," in *Analysis and Computation of Fixed Points*. New York: Academic, 1980.
- [12] C.-H. Han, W.-W. Lin, and C.-S. Wang, "Existence of homotopy methods for solving modified algebraic Riccati equations," *Nonlinear Anal. Theory, Methods and Appl.*, submitted.
- [13] L.-Z. Lu and W.-W. Lin, "An iterative algorithm for the solution of the discrete-time algebraic Riccati equation," *Lin. Alg. Appl.*, vol. 188, pp. 465-488, 1993.
- [14] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem*. Berlin: Springer-Verlag, 1991.
- [15] J. W. Milnor, *Topology from the Differentiable Viewpoint*. Charlottesville, VA: Univ. Press of Virginia, 1965.
- [16] T. Pappas, A. J. Laub, and N. R. Sandell, Jr., "On the numerical solution of the discrete-time algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 4, pp. 631-641, 1980.
- [17] S. Richter, L. D. Davis, E. G. Collins, Jr., "Efficient computation of the solutions to modified Lyapunov equations," *SIAM Matrix Anal. Appl.* vol. 14, no. 2, pp. 420-431, 1993.
- [18] S. Richter, A. S. Hodel, and P. G. Pruet, "Homotopy methods for the solution of general modified algebraic Riccati equations," *IEE Proc.-D*, vol. 140, no. 6, pp. 449-454, 1993.
- [19] L. T. Watson and R. Hafika, "Modern homotopy methods in optimization," *Comp. Methods Appl. Mech. Engin.*, vol. 74, pp. 289-305, 1989.

Hierarchical Production and Setup Scheduling in Stochastic Manufacturing Systems

Suresh P. Sethi and Qing Zhang

Abstract—This paper is concerned with an asymptotic analysis of hierarchical production and setup scheduling in a stochastic manufacturing system consisting of a single failure-prone machine and facing constant demands for a number of products. At any given time the system can only produce one type of product, and the system requires a setup if production is to be switched from one type of product to another. A setup may involve setup time or setup cost or both. The objective of the problem is to minimize the total costs of setup, production, and surplus. The control variables are a sequence of setups and a production plan. An asymptotic analysis with respect to increasing rates of change in machine states gives rise to a deterministic limiting optimal control problem in which there is a control variable associated with each of the machine states and the production rate is obtained by weighting these controls with the stationary probabilities of the corresponding states. It is shown that the value function for the original problem converges to the value function of the limiting problem. Asymptotic optimal controls for the original problem from optimal or near-optimal controls for the limiting problem are constructed. A simple illustrative example is provided.

I. INTRODUCTION

The recognition of the complexity of the production planning problems in stochastic manufacturing systems has resulted in various attempts to obtain suboptimal or near-optimal controls. Of particular importance to us is the so-called hierarchical controls approach based on the reduction of a given complex problem into simpler approximate problems or subproblems and to construct a satisfactory solution for the given problem from the solutions of the simpler problems. Moreover, in cases of stochastic systems in which fluctuation rates or frequencies of some processes are much faster than the frequencies associated with other processes, the hierarchical approach may provide us with solutions that are asymptotically optimal as the frequencies of the faster processes become large. The reader is referred to Lehoczky *et al.* [5], Sethi and Zhang [8], and Sethi *et al.* [10] for further details.

A crucial assumption in [5], [8], and [10] has been that the machines are completely flexible and thus do not require any setup for switching the production from one product to another. Ideally, the assumption amounts to the possibility of simultaneous production of different products.

An important class of manufacturing systems consists, however, of systems that have machines which involve setup costs and/or setup times, when switching from production of one product to that

Manuscript received November 19, 1993; revised November 17, 1994. This work was supported in part by NSERC Grant A4619.

S. P. Sethi is with Faculty of Management, University of Toronto, Toronto, Ontario, M5S 1V4 Canada.

Q. Zhang is with Department of Mathematics, University of Georgia, Athens, GA 30602 USA.

IEEE Log Number 9409693.