A PRIORI BOUNDS FOR THE RICCATI EQUATION

R. S. BUCY
UNIVERSITY OF SOUTHERN CALIFORNIA

1. Introduction and general results

Since the appearance of [6] and [7], the theory of linear filtering has experienced a renaissance. This theory, although evidently well known to statisticians in terms of "least squares estimates," has found many applications in the early sixties, largely because of the realization and synthesis methods provided in [6] and [7]. For an indication of some of the aerospace applications to guidance of spacecraft, the interested reader may find detailed information in [3]. Although the theory of linear filtering has changed little from that given in [7] for the continuous time problem, the practical realization of the so-called "correlated noise problem" as treated *mathematically* in [6], has recently found a solution in [4]. The full solution of this discrete time filtering problem and its meaning is described in detail in [5]. For readers desirous of a survey of recent results in linear and nonlinear filtering, it is available in [5], while more detailed information can be found in [3].

In this paper, our interest will center on the discrete matrix Riccati equation with emphasis on the study of the asymptotic behavior of its covariance matrix solution. A major tool in this study will be the Duffin parallel resistance of two nonnegative definite matrices A and B denoted by $A:B$. This operation is described in detail in [1] and provides for us a link between the Riccati equation and the classical continued fraction theory described in [9] and [10].

We have undertaken to study the discrete Riccati equation from the point of view of continued fractions because this technique provides considerable generality in that the nonsingular theory becomes a rather special case (see [3], Chapter 5) and much deeper results are obtained for singular problems; also the methods are striking generalizations of classical continued fraction methods.

We will be concerned with the cone C of $d \times d$ real entry symmetric nonnegative definite matrices. The cone C induces a natural partial ordering as for $A \in C$ and $B \in C$, $A \geqq B$ when and only when $A - B \in C$. The object of study will be the map of C;

This research was supported in part by the United States Air Force, Office of Aerospace Research, Applied Mathematics Division, under Grant AF-AF OSR-1244-67.

$$\tau_n(A) = \phi_n H'_{n-1}(H_{n-1}AH'_{n-1}: R_{n-1})H_{n-1}\phi'_n + C_n, \tag{1.1}$$

where with $M_{\ell,m}(R)$ the set of $\ell \times m$ real entry matrices.

The matrices A and ϕ_n belong to $M_{d,d}(R)$, $H_{n-1} \in M_{s,d}(R)$, $R_{n-1} \in M_{s,s}(R)$, and R_{n-1} is positive definite. The matrix $C_n \in M_{d,d}(R)$ with $s \geqq d$ and $H'_{n-1}H_{n-1} = I_d$, the $d \times d$ identity. We will restrict C_n and A to be members of the cone C and will center our attention on iterates of the mapping τ. We will consider the case where $s \geqq d$ and provide motivation for (1.1) in terms of the filtering problem in a later section. We recall that $(A:B) = A(A + B)^{\#}B$ (see [1]), where $^{\#}$ denotes the Moore–Penrose pseudoinverse determined by the axioms:

(a) $AA^{\#}A = A$,
(b) $A^{\#}AA^{\#} = A^{\#}$,
(c) $(AA^{\#})' = AA^{\#}$,
(d) $(A^{\#}A)' = A^{\#}A$,

with $'$ denoting transpose.

The following series of lemmas will prove useful in the sequel.

LEMMA 1.1. *For $A_i \in M_{\ell,k}(R)$,*

$$A_i\left[\sum_{j=1}^{n} A'_jA_j\right]\left[\sum_{j=1}^{n} A'_jA_j\right]^{\#} = A_i \tag{1.2}$$

for $1 \leqq i \leqq n$.

PROOF. See [8].

LEMMA 1.2. *For A and $B \in C$ and $A \geqq B$, $\tau_n(A) \geqq \tau_n(B)$, and in particular $\tau_n(C) \subseteqq C$.*

PROOF. Denoting by $\|\mathbf{x}\|_A^2$ the quadratic form induced by $A \in C$, it is quite easily seen that

$$\|\mathbf{x}\|^2_{\tau_n(A)} = \min_{\mathbf{r} \in R^s} \{\|\phi'_n\mathbf{x} + H'_{n-1}\mathbf{r}\|^2_A + \|\mathbf{r}\|^2_{R_{n-1}} + \|\mathbf{x}\|^2_{C_n}\}, \tag{1.3}$$

so that

$$\tau_n(A) = \phi_n(A - AH'_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}H_{n-1}A)\phi'_n + C_n. \tag{1.4}$$

In particular, if $A \geqq B$, then

$$\begin{aligned}\|\mathbf{x}\|^2_{\tau_n(A)} &= \min_{\mathbf{r} \in R^s} \{\|\phi'_n\mathbf{x} + H'_{n-1}\mathbf{r}\|^2_B + \|\mathbf{r}\|^2_{R_{n-1}} + \|\mathbf{x}\|^2_{C_n} + \|\phi'_n\mathbf{x} + H'_m\mathbf{r}\|^2_{A-B}\};\end{aligned} \tag{1.5}$$

the result follows. Of course, if $S_{n-1}(A) = A - AH'_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}H_{n-1}A$, then

$$\begin{aligned}H_{n-1}S_{n-1}(A)H'_{n-1} &= H_{n-1}AH'_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}R_{n-1} = (H_{n-1}AH_{n-1}):R_{n-1}\end{aligned} \tag{1.6}$$

in view of Lemma 1.1, or

$$S_{n-1}(A) = H'_{n-1}(R_{n-1} - R_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}R_{n-1})H_{n-1} \tag{1.7}$$

since by our assumption, $H'_{n-1}H_{n-1} = I_d$.

We will be concerned with the iterated composition mapping $\tau_1(\tau_2(\cdots(\tau_n(A))\cdots)$. It is clear that $\tau_1\,\tau_2\cdots\tau_n(0)$ is monotone increasing in C, while $\tau_1\cdots\tau_n(\infty)$ is monotone decreasing in C. We remark that Riesz has shown bounded monotone sequences in C converge, that is, if $A_n \geqq A_{n-1}$ and $A_n \leqq \alpha I_d$ then $\lim_{n\to\infty} A_n$ exists.

For a direct generalization of the scalar continued fraction theory see Wall [10].

THEOREM 1.1. *Let* $t_i(A) = D_1 + A$ *if* $i = 1$, *and*

$$(1.8)\qquad t_i(A) = -\phi_{i-1}H'_{i-2}R_{i-2}\big(H_{i-2}(A + D_i)H'_{i-2} + R_{i-2}\big)^{\#}R_{i-2}H_{i-2}\phi'_{i-1},$$

otherwise then

$$(1.9)\qquad \begin{aligned}&\tau_1\tau_2\cdots\tau_n(A)\\ &\quad= t_1\cdots t_n\big(-\phi_nH'_{n-1}R_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}R_{n-1}H_{n-1}\phi'_n\big),\end{aligned}$$

where $D_i = C_i + \phi_iH'_{i-1}R_{i-1}H_{i-1}\phi'_i$.

PROOF. An induction proof will be used to establish (1.9). For $n = 1$, it follows that,

$$(1.10)\qquad \begin{aligned}\tau_1(A) &= D_1 - \phi_1H'_0R_0(H_0AH'_0 + R_0)^{\#}R_0H_0\phi'_1\\ &= t_1\big(-\phi_1H'_0R_0(H_0AH'_0 + R_0)^{\#}R_0H_0\phi'_1\big).\end{aligned}$$

Now assume (1.9) is true for all $n \leqq k$, then

$$(1.11)\qquad \begin{aligned}&\tau_1\cdots\tau_{k+1}(A)\\ &= \tau_1\cdots\tau_k\big(\tau_{k+1}(A)\big)\\ &= t_1\cdots t_k\big(-\phi_kH'_{k-1}R_{k-1}\big(H_{k-1}\tau_{k+1}(A)H'_{k-1} + R_{k-1}\big)^{\#}R_{k-1}H_{k-1}\phi'_k\big),\end{aligned}$$

in view of the induction hypothesis. Now by the definition of t_{k+1}, (1.11) is equal to $t_1\cdots t_{k+1}\big(\tau_{k+1}(A) - D_{k+1}\big)$; however,

$$(1.12)\qquad \tau_{k+1}(A) - D_{k+1} = -\phi_{k+1}H'_kR_k(H_kAH'_k + R_k)^{\#}R_kH_k\phi'_{k+1},$$

in view of the proof of Lemma 1.2, so that the assertion follows.

COROLLARY 1.1. *For* $A \in C$,

$$(1.13)\qquad \begin{aligned}\tau_1\cdots\tau_n(0) &= t_1\cdots t_n(-\phi_nH'_{n-1}R_{n-1}H_{n-1}\phi'_n)\\ &\leqq \tau_1\cdots\tau_n(A) \leqq t_1\cdots t_n(0) = \tau_1\cdots\tau_n(\infty).\end{aligned}$$

PROOF. The proof is immediate.

COROLLARY 1.2. *Let* $t_i^*(A) = D_i + A$. *The limits* $V_i = \lim_{n\to\infty} t_i^*t_{i+1}\cdots t_n(0)$ *and* $L_i = \lim_{n\to\infty} t_i^*t_{i+1}\cdots t_n(-\phi_nH'_{n-1}R_{n-1}H_{n-1}\phi'_n)$ *exist and* $L_i \leqq V_i$.

PROOF. The map $t_i^*t_{i+1}\cdots t_n(\infty) = \lim_{A\uparrow\infty}\tau_i\cdots\tau_n(A)$ is monotone non-increasing in n and bounded above by D_1 and below by 0 and hence converges. A similar argument demonstrates the existence of the other limit.

REMARK 1.1. Notice that V_i and L_i are equilibrium solutions in that from the relation

$$(1.14)\qquad \tau_i\tau_{i+1}\cdots\tau_n(A) = \tau_i\big(\tau_{i+1}\cdots\tau_n(A)\big)$$

and as τ_i preserves order in C, it follows that $V_i = \tau_i(V_{i+1})$ and $L_i = \tau_i(L_{i+1})$ when $L_i \in \mathring{C}$, $\mathring{C}$ the interior of C. Further, when τ_i is autonomous, L_i and V_i are constant.

COROLLARY 1.3. Suppose $[\Pi_{j=i}^{n-1} \rho_{j,j+1}^n]\|D_n - C_n\|$ tends to zero as $n \to \infty$, where

$$(1.15) \qquad \rho_{i,i+1}^n = \|\phi_i H'_{i-1} R_{i-1}[(H_{i-1} L_{i+1,n} H_{i-1} + R_{i-1})^{\#}]^2 R_{i-1} H_{i-1} \phi'_i\|$$

and $L_{i,n} = \tau_i \cdots \tau_n(0)$, then

$$(1.16) \qquad L_i = \lim_{n \to \infty} \tau_i \cdots \tau_n(0) = \lim_{n \to \infty} \tau_i \cdots \tau_n(A) = \lim_{n \to \infty} \tau_i \cdots \tau_n(\infty) = V_i$$

for all $A \in C$.

PROOF. Let $E_{i,n} = \tau_i \cdots \tau_n(\infty) - \tau_i \cdots \tau_n(0)$; then with $E_{n,n} = D_n - C_n$ and with $L_{i,n} = \tau_i \cdots \tau_n(0)$

$$(1.17) \qquad \begin{aligned} E_{i,n} = \phi_i H'_{i-1}\{(H_{i-1}[E_{i+1,n} + L_{i+1,n}]H'_{i-1} : R_{i-1}) \\ - (H_{i-1} L_{i+1,n} H'_{i-1} : R_{i-1})\} H_{i-1} \phi'_i \end{aligned}$$

or in view of Lemma 27 in [1]

$$(1.18) \qquad E_{i,n} = \phi_i H'_{i-1} R_{i-1} C_{i,n}^{\#}(E_{i+1,n} : C_{i,n}) C_{i,n}^{\#} R_{i-1} H_{i-1} \phi'_i,$$

where $C_{i,n} = H_{i-1} L_{i+1,n} H'_{i-1} + R_{i-1}$. Hence, if $\varepsilon_{i,n} = \|E_{i,n}\|$, then $\varepsilon_{i,n} \leqq \rho_{i,i+1}^n \varepsilon_{i+1,n}$, where

$$(1.19) \qquad \rho_{i,i+1}^n = \|\phi_i H'_{i-1} R_{i-1}[(H_{i-1} L_{i+1,n} H'_{i-1} + R_{i-1})^{\#}]^2 R_{i-1} H_{i-1} \phi'_i\|$$

so that

$$(1.20) \qquad \varepsilon_{i,n} \leqq \left[\prod_{j=i}^{n-1} \rho_{j,j+1}^n\right] \|D_n - C_n\|$$

by Corollary 1.1.

EXAMPLE 1.1. Consider the scalar Riccati equation

$$(1.21) \qquad p_n = \alpha_n^2 p_{n-1} - \frac{\alpha_n^2 p_{n-1}^2}{r_{n-1} + p_{n-1}} + q_n,$$

where $q_n \geqq 0$ and $r_n \geqq 0$. The associated continued fraction is

$$(1.22) \qquad b_0 + \frac{a_1|}{|b_1} + \frac{a_2|}{|b_2} \cdots,$$

where

$$(1.23) \qquad \begin{aligned} b_0 &= q_0 + \alpha_0^2 r_{-1}, \\ b_i &= q_i + \alpha_i^2 r_{i-1} + r_{i-2}, \\ a_i &= \alpha_{i-1}^2 r_{i-2}^2. \end{aligned}$$

In view of Corollary 1.3, it suffices that

$$(1.24) \qquad \left[\prod_{j=1}^{n-1} \frac{\alpha_j^2 r_{j-1}^2}{(q_{j+1} + r_{j-1})^2}\right] \alpha_n^2 r_{n-1}$$

tends to zero as $n \to \infty$. Notice this result is far stronger than those of [2]. In the autonomous case the associated continued fraction is

$$y = \beta_0^* + \frac{\alpha^*|}{|\beta^*} + \frac{\alpha^*|}{|\beta^*} + \cdots, \tag{1.25}$$

where $\beta_0^* = \beta^* - r = q + \alpha^2 r$ and $\alpha^* = -\alpha^2 r^2$ and it is well known that the continued fraction converges to $y = x - r$ with x the positive root of $x^2 - \beta^* x - \alpha^* = 0$. For detailed information see [10], problem 1.4, page 23. Explicitly, y is given by

$$y = \tfrac{1}{2}(q + (\alpha^2 + 1)r + [(q + (\alpha^2 + 1)r)^2 - 4\alpha^2 r]^{1/2}) - r. \tag{1.26}$$

Further, if $\alpha = e^{\Delta f}$, $r = r^*/\Delta$, and $q = q^*\Delta$, it is easily verified that as $\Delta \to 0$

$$y(\Delta) \downarrow fr^* + (f^2(r^*)^2 + r^*q^*)^{1/2}, \tag{1.27}$$

the equilibrium solution of the analogous continuous time Riccati equation (see Example 1, p. 98, [7]).

COROLLARY 1.4. *For* $n \geqq 1$, $C_1 \leqq \tau_1 \cdots \tau_n(A) \leqq D_1$.

THEOREM 1.2. *If* $S_\alpha^+ = \{A \in C \mid \tau_n(A) \geqq A$ *for all* $n \geqq \alpha\}$ *and* $S_\alpha^- = \{A \in C \mid \tau_n(A) \leqq A$ *for all* $n \geqq \alpha\}$, *then* $\tau_\alpha \cdots \tau_n(A)$ *is monotone nondecreasing for* $A \in S^+$ *and monotone nonincreasing for* $A \in S^-$. *Further, if there exist real* γ_α *and* β_α *such that* $D_n \leqq \gamma_\alpha I$, $C_n \geqq \beta_\alpha I$, *for all* $n \geqq \alpha$, *then*

$$\begin{aligned} S_\alpha^+ &\supseteq \{A \in C \mid A \leqq \beta_\alpha I\}, \\ S_\alpha^- &\supseteq \{A \in C \mid A \geqq \gamma_\alpha I\}. \end{aligned} \tag{1.28}$$

PROOF. Since τ_n preserves the ordering of C, if $\tau_n(A) \geqq A$, then $\tau_{n-1}\tau_n(A) \geqq \tau_{n-1}(A)$, and so forth, so that $\tau_\alpha \cdots \tau_n(A) \geqq \tau_\alpha \cdots \tau_{n-1}(A)$ for $A \in S_\alpha^+$. The other assertions are obvious.

REMARK 1.2. In view of the double argument convexity of $(A : B)$, S_α^+ is convex (see [1], Theorem 24).

We will now specialize further and assume that ϕ_i, C_i, and R_i are invertible for all i; in view of Corollary 1.4, it suffices to consider $\tau_n(A)$ for $A \in C$ the interior of C.

Now

$$\begin{aligned} \tau_n(A) &= \phi_n(A^{-1} + O_n^{-1})^{-1}\phi_n' + C_n \\ &= \phi_n(A : O_n)\phi_n' + C_n \end{aligned} \tag{1.29}$$

with $O_n = (H_{n-1}' R_{n-1}^{-1} H_{n-1})^{-1}$. This relation follows since

$$\begin{aligned} &(H_{n-1}AH_{n-1}' : R_{n-1}) \\ &\quad = H_{n-1}AH_{n-1}'(H_{n-1}AH_{n-1}' + R_{n-1})^{-1}R_{n-1} \\ &\quad = H_{n-1}AH_{n-1} - H_{n-1}AH_{n-1}'(H_{n-1}AH_{n-1}' + R_{n-1})^{-1}(H_{n-1}AH_{n-1}'), \end{aligned} \tag{1.30}$$

so that

$$H'_{n-1}(H_{n-1}AH'_{n-1} : R_{n-1})H_{n-1} \tag{1.31}$$
$$= A - AH'_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{-1}H_{n-1}A$$
$$= (A^{-1} + O_n^{-1})^{-1} = (A : O_n)$$

by the Schur lemma (see [3], Theorem (8.6)). Hence, (1.1) demonstrates (1.29) in view of (1.30). Again, it is of interest to study $\tau_1 \cdots \tau_n(A)$, however, now for $A \geqq C_0 > 0$.

In order to simplify this study, we introduce the following transformation

$$\begin{aligned} \sigma_i^*(A) &= +A + S_i, \\ \sigma_i(A) &= -\phi_{i-1}O_{i-1}(A + S_i + O_{i-1})^{-1}O_{i-1}\phi'_{i-1} \end{aligned} \tag{1.32}$$

with $S_i = C_i + \phi_i O_i \phi'_i$.

It will be convenient to rewrite (1.29) as

$$\tau_n(A) = S_n - \phi_n O_n(O_n + A)^{-1}O_n\phi_n. \tag{1.33}$$

THEOREM 1.3. *For $A \in C$,*

$$\sigma_i^*\sigma_{i+1} \cdots \sigma_n\big(-\phi_n O_n(O_n + A)^{-1}O_n\phi'_n\big) = \tau_i \cdots \tau_n(A). \tag{1.34}$$

PROOF. We prove this by induction. For $n = i$,

$$\sigma_i^*\big(-\phi_i O_i(O_i + A)^{-1}O_i\phi'_i\big) = S_i - \phi_i O_i(O_i + A)^{-1}O_i\phi'_i = \tau_i(A) \tag{1.35}$$

in view of (1.33). Suppose (1.34) is valid for all $n \leqq k$, then

$$\tau_i \cdots \tau_{k+1}(A) = \tau_i \cdots \tau_k\big(\tau_{k+1}(A)\big), \tag{1.36}$$

which equals by the induction hypothesis

$$\sigma_i^* \cdots \sigma_k\big(-\phi_k O_k(O_k + \tau_{k+1}(A))^{-1}O_k\phi'_k\big) \tag{1.37}$$
$$= \sigma_i^*\sigma_{i+1} \cdots \sigma_{k+1}\big(\tau_{k+1}(A) - S_{k+1}\big)$$
$$= \sigma_i^*\sigma_{i+1} \cdots \sigma_{k+1}\big(-\phi_{k+1}O_{k+1}(O_{k+1} + A)^{-1}O_{k+1}\phi'_{k+1}\big),$$

demonstrating the assertion for invertible A. Since (1.33) holds for all $A \in C$, the theorem is valid.

COROLLARY 1.5. *The maps σ and τ satisfy the relations*

$$\sigma_1^*\sigma_2 \cdots \sigma_n\big(-\phi_n O_n(O_n + C_{n+1})^{-1}O_n\phi'_n\big) \tag{1.38}$$
$$= \tau_1 \cdots \tau_n(C_{n+1}) = \tau_1 \cdots \tau_{n+1}(0)$$
$$\leqq \tau_1 \cdots \tau_{n+1}(A) \leqq \tau_1 \cdots \tau_{n+1}(\infty) = \tau_1 \cdots \tau_n(S_{n+1})$$
$$= \sigma^*\sigma_2 \cdots \sigma_n\big(-\phi_n O_n(O_n + S_{n+1})^{-1}O_n\phi'_n\big).$$

PROOF. Since $\tau_1 \cdots \tau_{n+1}(0) \leqq \tau_1 \cdots \tau_{n+1}(A) \leqq \tau_1 \cdots \tau_{n+1}(\infty)$, the result follows from the theorem and the obvious relations

$$\begin{aligned} \tau_1 \cdots \tau_{n+1}(0) &= \tau_1 \cdots \tau_n(C_{n+1}), \\ \tau_1 \cdots \tau_{n+1}(\infty) &= \tau_1 \cdots \tau_n(S_{n+1}). \end{aligned} \tag{1.39}$$

Now for $n \geqq 1$, it is clear that

$$\sigma_1^*\sigma_2 \cdots \sigma_{n+1}(\infty) = \sigma_1^*\sigma_2 \cdots \sigma_n(0) \equiv A_n B_n^{-1}, \tag{1.40}$$

so that we will suppose that

$$\sigma_1^*\sigma_2 \cdots \sigma_{n+1}(\Gamma) = (A_{n+1} + A_n K_n \Gamma)(B_{n+1} + B_n K_n \Gamma)^{-1}. \tag{1.41}$$

In fact, if the following relations are satisfied

$$\begin{aligned} K_n &= \phi_n'^{-1} O_n^{-1}, \\ A_{n+1} &= A_n K_n (S_{n+1} + O_n) - A_{n-1} K_{n-1} \phi_n O_n, \\ B_{n+1} &= B_n K_n (S_{n+1} + O_n) - B_{n-1} K_{n-1} \phi_n O_n, \end{aligned} \tag{1.42}$$

with $A_0 = K_0^{-1}$, $A_1 = S_1$, $B_0 = O$, $B_1 = I$, the relationship (1.41) can be proved by induction. The only point here to be verified is that $B_{n+1} + B_n K_n \Gamma$ is invertible, and this follows by a simple disconjugacy argument analogous to that given in [3], Theorem (5.1).

REMARK 1.3. Equation (1.41) is the analogy of the symplectic system

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} = \begin{bmatrix} \phi_n'^{-1} & \phi_n'^{-1} O_n^{-1} \\ C_n \phi_n'^{-1} & \phi_n - C_n \phi_n'^{-1} O_n^{-1} \end{bmatrix} \begin{bmatrix} X_{n-1} \\ Y_{n-1} \end{bmatrix} \tag{1.43}$$

associated with the mapping $\tau_k(\Gamma)$ given by (1.29) as $\tau_n(\Gamma) = Y_n X_n^{-1}$ with $X_{n-1} = I$ and $Y_{n-1} = \Gamma$.

Another point of interest is that of investigation of the equilibrium solutions of (1.33). Previously, we have shown the existence of two equilibrium solutions L_i and V_i. Let E_i denote general equilibrium solution, then if $\Sigma_i = E_i + O_{i-1}$ it follows from (1.33) that

$$\Sigma_i = S_i + O_{i-1} - \phi_i O_i (\Sigma_{i+1})^{-1} O_i \phi_i' \tag{1.44}$$

and in the autonomous case

$$\Sigma = S + O - \phi O (\Sigma)^{-1} O \phi'. \tag{1.45}$$

These equations generalize the quadratic fixed points of scalar continued fractions.

EXAMPLE 1.2. If $\Phi = I$ in (1.45), then

$$[O^{-1/2}\Sigma O^{-1/2}] = O^{-1/2} C O^{-1/2} + 2I - [O^{-1/2}\Sigma O^{-1/2}]^{-1} \tag{1.46}$$

and

$$\begin{aligned} O^{-1/2}\bar{P}O^{-1/2} &= -I + T\{1 + \tfrac{1}{2}\mu_i + [\mu_i + (\tfrac{1}{2}\mu_i)^2]^{1/2}\}T' \\ &= T\{\tfrac{1}{2}\mu_i + [\mu_i + (\tfrac{1}{2}\mu_i)^2]^{1/2}\}T, \end{aligned} \tag{1.47}$$

where $T'O^{-1/2}CO^{-1/2}T = \{\mu_i\}$, the notation $\{\lambda_i\}$ denotes the diagonal matrix

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}. \tag{1.48}$$

The equilibrium solution of $\tau(A)$ is unique and equals $\bar{P}$ with $\bar{P} = L = V$, if O and C are positive definite.

Explicitly, in terms of the unique positive semidefinite square root, we obtain

$$\bar{P} = \tfrac{1}{2}C + O^{1/2}(\tfrac{1}{4}O^{-1/2}CO^{-1}CO^{-1/2} + O^{-1/2}CO^{-1/2})^{1/2}O^{1/2}. \tag{1.49}$$

The continuous version of (1.49) when $C = A\Delta$ and $O = B(1/\Delta)$ and Δ tends to zero is

$$\bar{P} = B^{1/2}(B^{-1/2}AB^{-1/2})^{1/2}B^{1/2}. \tag{1.49'}$$

This equilibrium solution of $\dot{P} = -PB^{-1}P + A$ has been obtained by Reid, Bellman, and others and has lead to the comment that in this particular case the Riccati equation is a continuous "square rooter."

We rewrite (1.29) in the more convenient form for L_i and V_i as

$$L_i = \phi_i^* L_{i+1}\phi_i^{*\prime} + \phi_i(L_{i+1}: O_i)O_i^{-1}(L_{i+1}: O_i)\phi_i' + C_i \tag{1.50}$$

with $\phi_i^{*\prime} = (L_{i+1} + O_i)^{-1}O_i\phi_i'$ and

$$V_i = \phi_i^{**}V_{i+1}\phi_i'^{**} + \phi_i(V_{i+1}: O_i)O_i^{-1}(V_{i+1}: O_i)\phi_i' + C_i \tag{1.51}$$

with $\phi_i^{**\prime} = (V_{i+1} + O_i)^{-1}O_i\phi_i'$.

It also follows easily that $E_i = V_i - L_i$ satisfies

$$E_i = \phi_i^* E_{i+1}\phi_i^{**\prime} \tag{1.52}$$

(see [1], Lemma 27). Now the following theorem establishes the asymptotic theory for iterates of τ_i.

THEOREM 1.4. *Suppose ϕ_i, R_i and, O_i are invertible and further that there exist α and β real positive numbers such that*

(i) $0 < \alpha I \leqq C_i$,
(ii) $V_i \leqq \beta I$,

and that $\|\phi_i - I\| \leqq C_1$, then $V_i = L_i = \bar{P}_i$ and

$$\lim_{n\to\infty} \tau_i \cdots \tau_n(A) = \bar{P}_i, \tag{1.53}$$

the convergence being exponential.

Further (1.44) *has a unique supnorm bound solution with $\Sigma_i \geqq O_{i-1}$ given by $\Sigma_i = \bar{P}_i + O_{i+1}$ and $\bar{P}_i \in C$.*

PROOF. From (1.50) and (1.51) and our assumptions, it is easily checked that the quadratic forms $\|\mathbf{x}\|^2_{L_i}$ and $\|\mathbf{x}\|^2_{V_i}$ are Liapounov functions for $\mathbf{x}_{(i+1)} = \phi_i^{*\prime}\mathbf{x}_i$ and $\mathbf{y}_{(i+1)} = \phi_i^{**\prime}\mathbf{y}_i$, respectively. In particular, this implies

$$\|\phi_i^* \cdots \phi_n^*\| \leqq c_* \exp\{-\gamma_*(n - i)\} \tag{1.54}$$

for $n \geqq i$ for some c_* and γ_* positive and analogous relations for $\phi_i^{**\prime}$ hold with parameters c_{**} and γ_{**}. But since $V_i - L_i$ satisfies (1.52), it follows that

$$E_i = \phi_i^* \cdots \phi_k^* E_{k+1} \phi_k^{**\prime} \cdots \phi_i^{**\prime}, \tag{1.55}$$

and hence,

$$\|E_i\|^2 \leqq c_* c_{**} \exp\{-(\gamma_* + \gamma_{**})(k - i)\} \|E_{k+1}\|^2. \tag{1.56}$$

However, as k is arbitrary and $\|E_{k+1}\|^2$ is uniformly bounded, E_i must be zero. The validity of equation (1.53) follows as $L_{i,n} \leqq \tau_i \cdots \tau_n(A) \leqq V_{i,n}$, since $L_{i,n}$ and $V_{i,n}$ converge to $\bar{P}_i = V_i = L_i$, as $E_i = 0$. Now suppose (1.44) has two solutions Σ_i and T_i, then $x_i = \Sigma_i - O_{i-1}$ and $y_i = T_i - O_{i-1}$ are equilibrium solutions of (1.29) both supnorm bounded and members of C. Now

$$\begin{aligned} \tau_i \cdots \tau_n(x_{n+1}) &= x_i, \\ \tau_i \cdots \tau_n(y_{n+1}) &= y_i, \end{aligned} \tag{1.57}$$

and hence, as x_{n+1} and y_{n+1} are members of C,

$$\begin{aligned} L_{i,n} &\leqq x_i \leqq V_{i,n}, \\ L_{i,n} &\leqq y_i \leqq V_{i,n}, \end{aligned} \tag{1.58}$$

but the last equation implies $x_i = y_i = \bar{P}_i$.

REMARK 1.4. Theorem 1.4 is essentially unchanged when R_i is singular and provides stability and a unique equilibrium solution for (1.1) when L_i is uniformly positive definite and V_i is uniformly bounded in C.

REMARK 1.5. Notice that conditions (i) and (ii) of the theorem depend in general on the existence of uniform *a priori* bounds, which we developed previously.

2. Applications to the theory of filtering

In order to apply the theory of iterates of the mapping $\tau_n(A)$ of the last section to problems in linear filtering theory, we must study the mapping

$$\Delta_n(A) = \phi_n A \phi_n' - K_n(H_{n-1} A H_{n-1}' + R_{n-1})^{\#} K_n' + G_n G_n' \tag{2.1}$$

with $K_n = \phi_n A H_{n-1}' + G_n L_{n-1}'$, $R_{n-1} \geqq L_{n-1} L_{n-1}'$ for $A \in C$. This mapping is similar to τ_n except that $H \in M_{s,d}(R)$ and $G \in M_{d,r}(R)$ with r and s for many problems *less* than d. Note also the more general nonlinear term which arises from the observation noise being correlated with the signal process. Iterates of Δ determine the error covariance matrix of the optimal filter (see [3], Chapters 4 and 9). The following lemma relates $\Delta_n(A)$ to the simpler Riccati mapping T_n.

LEMMA 2.1. *For $A \in C$, $\Delta_n(A) = T_n(A)$, where*

$$T_n(A) = \psi_n S_n(A) \psi_n' + G_n(I - L_{n-1}' R_{n-1}^{\#} L_{n-1}) G_n' \tag{2.2}$$

and

$$\begin{aligned} S_n(A) &= A - A H_{n-1}'(H_{n-1} A H_{n-1}' + R_{n-1})^{\#} H_{n-1} A, \\ \psi_n &= \phi_n - G_n L_{n-1}' R_{n-1}^{\#} H_{n-1}. \end{aligned} \tag{2.3}$$

PROOF. From the definition, $\Delta_n(A)$ can be written as

$$\begin{aligned}(2.4)\qquad \Delta_n(A) = \phi_n S_n(A)\phi'_n - G_n L'_{n-1}(H_{n-1}AH_{n-1} + R_{n-1})^{\#}H_{n-1}A\phi'_n \\ + G_n L'_{n-1}(H_{n-1}AH_{n-1} + R_{n-1})^{\#}L_{n-1}G'_n \\ - \phi_n AH_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}L_{n-1}G'_n + G_n G'_n.\end{aligned}$$

But by the definition of $S_n(A)$ and Lemma 1.1, it follows that

$$\begin{aligned}(2.5)\qquad & S_n(A)H'_{n-1}R^{\#}_{n-1}L_{n-1}G'_n \\ & = AH_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}R_{n-1}R^{\#}_{n-1}L_{n-1}G'_n \\ & = AH_{n-1}(H_{n-1}AH'_{n-1} + R_{n-1})^{\#}L_{n-1}G'_n,\end{aligned}$$

and hence,

$$\begin{aligned}(2.6)\qquad & G_n L'_{n-1}R^{\#}_{n-1}H_{n-1}S_n(A)H'_{n-1}R^{\#}_{n-1}L_{n-1}G'_n \\ & = G_n L'_{n-1}R^{\#}_{n-1}L_{n-1}G'_n - G_n L'_{n-1}(H_{n-1}AH_{n-1} + R_{n-1})^{\#}L_{n-1}G'_n.\end{aligned}$$

In view of the above equalities and (2.4), the lemma follows.

REMARK 2.1. Lemma 2.1 is the discrete time *generalization* of continuous time equivalence of Riccati equations (see [3], especially page 90).

REMARK 2.2. Notice that if $R_{n-1} > 0$, then Q controllability and R observability of (2.1) hold when and only when (2.2) is Q^* controllable and R^* observable. In fact, in the general case, it seems appropriate to call (2.1) controllable and observable when these conditions hold for (2.2).

In order to overcome difficulty that H has in general less sensors than the state dimension, we process the observations in blocks of k corresponding to k sequential time observations. In other words, if $\pi(n, \Gamma, n_0)$ represents solution of the Riccati equation, we find the recursion for $\pi(n_0 + kv, \Gamma, n_0)$ in terms of $\pi(n_0 + k(v - 1), \Gamma, n_0)$. This recursion equation is of the general form of (2.1) with $H_n \in M_{ks,d}(R)$, and hence for $sk \geqq d$ using Lemma 2.1, the results of Section 1 are applicable.

As an example, we consider the mapping

$$(2.7)\qquad T(A) = \phi_*\{A - AH'_*\{H_*AH'_* + R_*\}^{\#}H_*A\}\phi'_* + G_*G'_*.$$

Then $T^k(A) = \phi^k\{A - (AH' + \phi^{-k}GL')(HAH' + R + LL')^{\#}(HA + LG'\phi'^{-k})\phi'^k + GG'$, where

$$(2.8)\qquad L = \begin{Bmatrix} 0 & 0 & \cdots & & & \\ u_0 & 0 & \cdots & & & \\ u_1 & u_0 & 0 & \cdots & & \\ u_2 & u_1 & u_0 & 0 & \cdots & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{k-2} & u_{k-3} & \cdots & u_0 & 0 & \cdots \end{Bmatrix}, \qquad u_i = H_*\phi^i G_*,$$

$$(2.9) \qquad R = \begin{bmatrix} R_* & 0 & 0 & 0 \\ 0 & R_* & 0 & \cdots \\ 0 & 0 & R_* & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0R_* \end{bmatrix},$$

$$(2.10) \qquad H = \begin{bmatrix} H_* \\ H_*\phi_* \\ \vdots \\ H_*\phi_*^{k-1} \end{bmatrix}, \qquad G = [\phi_*^{k-1}G_*, \cdots, G_*].$$

If ϕ_* is invertible, define M via the equation $H\phi_*^{-k}G = M + L$, using properties of the pseudoinverse

$$(2.11) \qquad \|\mathbf{x}\|^2_{T^k(A)} = \min_{\mathbf{y}\in R^s} \{\|\psi'\mathbf{x} + H'\mathbf{y}\|^2_A + \|\mathbf{y}\|^2_{R+LL'} + \|G'x\|^2_{I-L'(R+LL')^{\#}L}\}$$

for all $x \in R^d$.

An interesting problem is that of characterizing $J_\ell = \{\underline{x} \in R^d \mid T^j(A)\underline{x} = T^\ell(0)\underline{x}$, all $A \in C$, all $j \geqq \ell\}$. This problem has been solved for H_* a vector and $R_* = 0$ (see [4]). In [5], the general solution has been given for $R = 0$. From (2.11) the following theorem determines J_k in general.

THEOREM 2.1. *With matrices given by* (2.8), (2.9), *and* (2.10),

$$(2.12) \qquad \begin{aligned} J_k &= \{\mathbf{x} \in R^d \mid T^k(A)\mathbf{x} = T^k(0)\mathbf{x}, \quad \textit{for all} \quad A \in C\} \\ &= \{\mathbf{x} \in R^d \mid \psi'\mathbf{x} = H'\ell, \ell \in R^s, \|\ell\|^2_{(R+LL')} = 0\}. \end{aligned}$$

PROOF. It is clear that $\{\mathbf{x} \in R^d \mid T^k(A)\mathbf{x} = T^k(0)\mathbf{x}$, for all $A \in C\} \supseteqq J_k$. Since for $A = T^\ell(B)$ and x such that $T^k(0)\mathbf{x} = T^k(A)\mathbf{x}$, it follows that $T^k(0)\mathbf{x} = T^{k+\ell}(B)\mathbf{x}$ for arbitrary $B \in C$, so that first set equality is valid. The second set equality follows from (2.11) by considering $T^k(0)$ and $T^k(A)$ for $A \in \dot{C}$.

REMARK 2.3. Notice that the *invertibility* of ϕ is unnecessary for the validity of Theorem 2.1.

The general technique of enlarging the sensor by block processing is valid in the time dependent case and leads to a structure analogous to (2.8), (2.9), and (2.10). Because of this the *a priori* bounds of Section 1 as well as the asymptotic results apply in general with the only restriction being that there exists a k such that rank $[H'_{n_0}, \cdots \phi'(n_0, n_0 + k)H'_{n_0+k}] = d$ for all n_0.

3. Conclusions

We have shown that the theory of the Riccati equation which arises in the discrete time linear filtering problem can be easily obtained by considering the temporal evolution of k fold iterates. A generalized theory of continued fractions in semidefinite matrices has been given, which provides best possible upper and lower *a priori* bounds for the Riccati equation solutions. It would seem that the upper and lower approximates would provide interesting ways to compute suboptimal filters in environments where the prior variance is unknown.

In a future paper, we will study the analogous continuous time situation.

REFERENCES

[1] W. N. ANDERSON and R. J. DUFFIN, "Series and parallel addition of matrices," *J. Math. Anal. Appl.*, Vol. 26 (1969), pp. 576–593.

[2] R. S. BUCY, "Global theory of the Riccati equations," *J. Comput. Systems Sci.*, Vol. 1 (1967), pp. 349–361.

[3] R. S. BUCY and P. D. JOSEPH, *Filtering for Stochastic Processes with Application to Guidance*, New York, Interscience, 1968.

[4] R. S. BUCY, D. RAPPAPORT, and L. M. SILVERMAN, "Correlated noise filtering and invariant directions for the Riccati equation," *IEEE Trans. Automatic Control*, Vol. AC-15 (1970), pp. 535–540.

[5] R. S. BUCY, "Linear and non-linear filtering," *Proc. IEEE*, Vol. 58 (1970), pp. 854–864.

[6] R. E. KALMAN, "A new approach to linear filtering and prediction problems," *ASME J. Basic Eng.*, Vol. 82 (1960), pp. 35–45.

[7] R. E. KALMAN and R. S. BUCY, "New results in linear filtering and prediction theory," *ASME J. Basic Eng.*, Vol. 83 (1961), pp. 95–108.

[8] R. E. KALMAN and T. S. ENGLAR, "A user's manual for A.S.P.-*C*," NASA CR-475, Ames Research Center, Moffet Field, California, 1965.

[9] O. PERRON, *Die Lehre von den Kettenbruecken*, Stuttgart, Teubner, 1954.

[10] H. S. WALL, *Analytic Theory of Continued Fractions*, New York, Chelsea, 1967.