On the Riccati Equation of the Optimal Filter with Disturbance Decoupling Property for Linear Stochastic Systems

Akio Tanikawa

Faculty of Information Science and Technology,
Osaka Institute of Technology
Kitayama, Hirakata-shi, 573-0196, Japan
E-mail: tanikawa@is.oit.ac.jp

Abstract

For discrete-time linear stochastic systems with unknown disturbances, we consider the optimal filter with disturbance decoupling property and the equation (i.e., Riccati equation) which the covariance matrices of the estimation errors of the filter satisfy. Assuming that the stochastic processes have constant coefficients, we prove convergence of the Riccati equation and derive a simple equation (called the algebraic Riccati equation (ARE)) which is the limit of the Riccati equation under some conditions similar to those for the Kalman filter. We also prove asymptotic stability of the systems whose optimal gains are determined by the ARE.

Keywords: Linear stochastic system, Optimal filter, Unknown input, Riccati equation.

1 Introduction

We consider discrete-time linear stochastic systems with unknown disturbances and the optimal filter of the systems with disturbance decoupling property. Since modeled systems made by engineers are not always accurate representations of real systems, we should consider systems with unknown inputs.

In this paper, we are concerned with the optimal filtering problem which investigates the optimal estimate \( \hat{x}_t \) of state \( x_t \) at time \( t \) with minimum variance based on the observation \( Y_t \) of the outputs \( \{y_0, y_1, \ldots, y_t\} \), i.e., \( Y_t = \sigma\{y_s, s = 0, 1, \ldots, t\} \) (the smallest \( \sigma \)-field generated by \( \{y_0, y_1, \ldots, y_t\} \) (see e.g., [18], Chapter 4)). Optimal filters for linear stochastic systems with unknown inputs have been investigated by many researchers. We mention some of the well-known works, e.g., [4]-[13], [19] and [20]. This paper was inspired particularly by the optimal disturbance decoupling observer (ODDO) proposed by Chen and Patton (in [5] and [6]). Their ODDO was modified by the author and his colleagues in [24] and [25] as the optimal filter with disturbance decoupling property. Later, this optimal filter was utilized to derive the optimal smoother with disturbance decoupling property in [21] and [22].

In this paper, we consider the optimal filter with disturbance decoupling property and fundamental properties of the equation (i.e., Riccati equation) which the covariance matrices of the estimation errors of the filter satisfy. In Section 2, we review preliminary results and give new formulas which play important roles in Section 3. In Section 3, assuming that the stochastic processes have constant coefficients, we prove convergence of the Riccati equation and derive a simple equation (called the algebraic Riccati equation (ARE)) which is the limit of the Riccati equation under some conditions similar to those for the Kalman filter. We also prove asymptotic stability of the systems whose optimal gains are determined by the ARE.

2 Preliminaries

Consider the following discrete-time linear stochastic system for \( t = 0, 1, 2, \cdots \):

\[
\begin{align*}
x_{t+1} &= A_t x_t + E_t d_t + S_t \zeta_t, \\
y_t &= C_t x_t + \eta_t,
\end{align*}
\]

where

\[
\begin{align*}
x_t &\in \mathbb{R}^n \quad \text{the state vector}, \\
y_t &\in \mathbb{R}^m \quad \text{the output vector}, \\
d_t &\in \mathbb{R}^q \quad \text{the unknown input vector}.
\end{align*}
\]

Suppose that \( \zeta_t \) and \( \eta_t \) are independent zero mean white noise sequences with covariance matrices \( I \) (the identity matrix) and \( R_t \). Let \( A_t, C_t \) and \( E_t \) be known matrices with appropriate dimensions.

In [25], we considered the optimal estimate \( \hat{x}_{t+1} \) of the state \( x_{t+1} \) which was proposed by Chen and Patton ([5] and [6]) with the following structure:

\[
\begin{align*}
z_{t+1} &= F_{t+1} z_t + K_{t+1} y_t, \\
\hat{x}_{t+1} &= z_{t+1} + H_{t+1} y_t + \zeta_t, \tag{4}
\end{align*}
\]

for \( t = 0, 1, 2, \cdots \). Here, \( \zeta_0 \) is chosen to be \( z_0 \) for a fixed \( z_0 \). Denote the state estimation error and its covariance matrix respectively by \( e_t \) and \( P_t \). Namely, we use the notations \( e_t = x_t - \hat{x}_t \) and \( P_t = E\{e_t e_t^T\} \) for \( t = 0, 1, 2, \cdots \).
0, 1, 2, \cdots. Here, \( E \) denotes expectation and \( T \) denotes transposition of a matrix. We assume in this paper that random variables \( x_0, \{y_t\}, \{z_t\} \) are independent. As in [5], [6] and [25], we consider state estimate (3)-(4) with the matrices \( F_{t+1}, T_{t+1}, H_{t+1} \) and \( K_{t+1} \) of the forms:

\[
K_{t+1} = K_{t+1}^1 + K_{t+1}^2,
\]

(5)

\[
E_t = H_{t+1} C_{t+1} E_t,
\]

(6)

\[
T_{t+1} = I - H_{t+1} C_{t+1},
\]

(7)

\[
F_{t+1} = A_{t} - H_{t+1} C_{t+1} A_{t} - K_{t+1}^1 C_{t},
\]

(8)

\[
K_{t+1}^2 = F_{t+1} H_{t}.
\]

(9)

The next lemma on equality (6) was obtained and used by Chen and Patton ([5] and [6]). Before stating it, we assume that \( E_0 \) is a full column rank matrix. Notice that this assumption is not an essential restriction.

\begin{lemma}
Equality (6) holds if and only if
\[
\text{rank}(C_{t+1} E_t) = \text{rank}(E_t).
\]
(10)

When this condition holds, matrix \( H_{t+1} \) which satisfies (6) must have the form

\[
H_{t+1} = E_t \left\{ (C_{t+1} E_t)^T (C_{t+1} E_t) \right\}^{-1} (C_{t+1} E_t)^T.
\]

(11)

Hence, we have

\[
C_{t+1} H_{t+1} = C_{t+1} E_t \left\{ (C_{t+1} E_t)^T (C_{t+1} E_t) \right\}^{-1} (C_{t+1} E_t)^T,
\]

(12)

which is a non-negative definite symmetric matrix.
\end{lemma}

When the matrix \( K_{t+1}^1 \) has the form

\[
K_{t+1}^1 = A_{t+1}^1 (P_t C_t^T - H_t R_t) (C_t P_t C_t^T + R_t)^{-1},
\]

(13)

\[
A_{t+1}^1 = A_{t} - H_{t+1} C_{t+1} A_{t},
\]

(14)

we obtained the following result (Theorem 2.7 in [25]) on the optimal filtering algorithm under the next condition which is supposed throughout the paper.

\begin{condition}
The matrices \( C_t H_t \) and \( R_t \) are commutative, i.e.,
\[
C_t H_t R_t = R_t C_t H_t,
\]
(15)
\end{condition}

\begin{proposition}
The optimal gain matrix \( K_{t+1}^1 \) which makes the variance of the state estimation error \( e_{t+1} \) minimum is determined by (13). Hence, we obtain the optimal filtering algorithm:

\[
\hat{x}_{t+1} = A_{t+1}^1 \{ \hat{x}_t + G_t (y_t - C_t \hat{x}_t) \} + H_{t+1} y_{t+1},
\]

(16)

\[
P_{t+1} = A_{t+1}^1 M_t A_{t+1}^1 + T_{t+1} S_t S_t^T T_{t+1}^T + H_{t+1} R_{t+1} H_{t+1}^T,
\]

(17)

where

\[
G_t = (P_t C_t^T - H_t R_t) (C_t P_t C_t^T + R_t)^{-1},
\]

(18)

and

\[
M_t = P_t - G_t (C_t P_t - R_t H_t^T).
\]

(19)

Here, we note that \( H_0 = O \) and that the equation (17) is called the Riccati equation.
\end{proposition}

\begin{remark}
If the matrix \( R_t \) has the form
\[
R_t = r_t I
\]
with some positive number \( r_t \) for each \( t = 1, 2, \cdots \), then it is obvious that condition (15) holds.
\end{remark}

Here, we remark that the standard Kalman filter is a special case of the optimal filter proposed in this section (see e.g., Theorem 5.2 (page 90) in [18]).

\begin{proposition}
Suppose that \( E_t \equiv O \) holds for all \( t \) (i.e., the unknown input term is zero). Then, Lemma 2.1 cannot be applied directly. But, we can choose \( H_t \equiv O \) for all \( t \) in this case, and the optimal filter given in Proposition 2.2 reduces to the standard Kalman filter.
\end{proposition}

We now note that some matrices are projection matrices which will play important roles later and can be proved by simple computations (see [21] also).

\begin{lemma}
Matrices \( C_t^T H_t T = T_t^T = I - C_t^T H_t T \) are projection matrices which have the following properties:
\[
(C_t^T H_t T) (C_t^T H_t T) = C_t^T H_t T,
\]

(20)

\[
(I - C_t^T H_t T) (I - C_t^T H_t T) = I - C_t^T H_t T,
\]

(21)

\[
C_t^T H_t T (I - C_t^T H_t T) = O,
\]

(22)

and moreover
\[
H_t^T (I - C_t^T H_t T) = O.
\]

(23)
\end{lemma}

We need the following lemma and proposition to rewrite the Riccati equation (17).

\begin{lemma}
For \( t = 0, 1, 2, \cdots, \) we have
\[
G_t R_t H_t^T = O \quad \text{and} \quad H_t R_t G_t^T = O.
\]
(24)
\end{lemma}

\begin{proposition}
For \( t = 0, 1, 2, \cdots, \) we have
\[
M_t = (I - G_t C_t) P_t (I - G_t C_t)^T + G_t R_t G_t^T
\]
and the Riccati equation
\[
P_{t+1} = A_{t+1}^1 M_t A_{t+1}^1 + T_{t+1} S_t S_t^T T_{t+1}^T + H_{t+1} R_{t+1} H_{t+1}^T
\]

(25)

\[
= A_{t+1}^1 (I - G_t C_t) P_t (I - G_t C_t)^T A_{t+1}^1 + T_{t+1} S_t S_t^T T_{t+1}^T
\]

(26)
\end{proposition}
3 Stochastic systems with constant coefficients

From now on, we consider the following discrete-time linear stochastic system with constant coefficients for 

\[ x_{t+1} = Ax_t + Ed + S \zeta_t, \quad t \geq 0 \]

\[ y_t = Cx_t + \eta_t. \]

Namely, the matrices \( A_t, C_t, E_t, S_t, R_t \) and the vector \( d_t \) do not depend on \( t \) and so the suffix \( t \) is dropped. We also drop the suffix \( t \) from \( H_t, T_t \) and \( A^1_t \). \( \zeta_t \) and \( \eta_t \) are supposed to be independent zero mean white noise sequences with covariance matrices \( I \) (the identity matrix) and \( R \). However, \( P_t, K_t, F_t, G_t \) and \( M_t \) still depend on \( t \).

In order to prove convergence of the sequence \( P_t \), we need some lemmas. For \( \alpha \) (with \( 0 \leq \alpha \leq 1 \)), we set

\[
\Phi(U) = A^1 \left\{ U - (UC^T - \alpha HR) (CU^T + R)^{-1} \times (CU^T - \alpha RH^T) \right\} (A^1)^T + TSS^TT^T + HRH^T.
\]

Then, it is easy to observe that \( P_t = \Phi(P_0) \) holds with \( \alpha = 0 \) and that \( P_{t+1} = \Phi(P_t) \) holds with \( \alpha = 1 \) for \( t \geq 1 \). We can prove monotonicity of \( P_t \) as follows.

Lemma 3.1 If the matrices \( Q_1 \) and \( Q_2 \) are both non-negative definite and symmetric with \( Q_2 \geq Q_1 \), then \( \Phi(Q_2) \geq \Phi(Q_1) \).

Proof To prove this lemma, we use the formula on the matrix-valued function \( V(s) \)

\[
\frac{d}{ds} V^{-1}(s) = -V^{-1}(s) \left[ \frac{d}{ds} V(s) \right] V^{-1}(s).
\]

Denoting \( U(s) = Q_1 + s(Q_2 - Q_1) \), we have

\[
\Phi(Q_2) - \Phi(Q_1) = \int_0^1 \frac{d}{ds} [\Phi(U(s))] \, ds
\]

\[
= A^1 \left\{ \int_0^1 \{ (I - (UC^T - \alpha HR) (CU^T + R)^{-1} \times (Q_2 - Q_1) (I - C^T (CU^T + R)^{-1} \times (CU^T - \alpha RH^T) \} \frac{ds}{(A^1)^T} + TSS^TT^T + HRH^T
\]

\[
\geq 0.
\]

Let us choose \( P_0 = O \). Then, we have \( P_t = \Phi(P_0) = TSS^TT^T + HRH^T \geq O \). It then follows from (3.1) that \( P_2 = \Phi(P_1) \geq \Phi(P_0) = P_t \). Thus, we have

\[
P_0 = O \leq P_1 \leq P_2 \leq P_3 \leq \cdots
\]

We now give two definitions to discuss convergence of the sequence of matrices \( \{P_t\} \). \( (A^1, S) \) is said to be stabilizable if there is a matrix \( L \) such that \( A^1 + SL \) is asymptotically stable. \( (C, A^1) \) is said to be detectable if there is a matrix \( L \) such that \( A^1 + LC \) is asymptotically stable.

Lemma 3.2 If \( (C, A^1) \) is detectable, then the sequence of matrices \( \{P_t\} \) is bounded for any initial matrix \( P_0 \geq O \).

Proof Since \( (C, A^1) \) is detectable, there exists a matrix \( L \in \mathbb{R}^n \times \mathbb{R}^m \) such that \( A := A^1 + LC \) is asymptotically stable. Instead of the dynamical system (16), we consider the following filter by substituting \( L \) into \( A^1 G_t \)

\[
\dot{x}_{t+1} = (A^1 - LC) x_t + Ly_t + H y_{t+1}
\]

\[
= (A^1 - LC) x_t + L(C x_t + \eta_t) + H(C x_{t+1} + \eta_{t+1})
\]

\[
= (A - HCA - LC) x_t + L x_t + L \eta + H x_{t+1} + L \eta + H \eta_{t+1}
\]

with \( \dot{x}_0 = \bar{x}_0 \). Hence, we have

\[
x_{t+1} - \dot{x}_{t+1} = (A - HCA - LC) (x_t - \dot{x}_t)
\]

\[
+ (I - HCA) E d + (I - HCA) S \zeta = L \eta - H \eta_{t+1}
\]

\[
= (A - HCA - LC) (x_t - \dot{x}_t) + T S \zeta = L \eta - H \eta_{t+1}
\]

Using the notations \( \dot{e}_t := x_t - \dot{x}_t \), \( t \dot= E \{ \dot{e}_t \dot{e}_t^T \} \), we have

\[
\dot{P}_{t+1} = A^1 \Sigma_0 - A^1 \Sigma_0 A^1 + \sum_{k=1}^{\infty} \dot{A}^k (LLR^T + HHR^T + TSS^TT^T + \dot{A}HRL^T + LHRH^T \dot{A}^T)^k.
\]

Since \( \dot{P}_t \) is identical to \( P_t \) (the optimal covariance matrix) when we choose \( L = A^1 G_t \), we note that \( \dot{P}_t \leq \dot{P}_t \). Due to asymptotic stability of \( A \), the right hand side of (32) converges as \( t \to \infty \). Hence, we have

\[
\dot{P}_t \leq \Sigma_0 + \sum_{k=0}^{\infty} \dot{A}^k (LLR^T + HHR^T + TSS^TT^T + \dot{A}HRL^T + LHRH^T \dot{A}^T)^k
\]

and boundedness of \( \{P_t\} \).

In view of (30) and Lemma 3.2, we can obtain the following convergence results.

Theorem 3.3 Suppose that \( (C, A^1) \) is detectable and that \( P_0 = O \). Then, the solution \( P_t \) of (17) converges to the non-negative definite matrix \( \dot{P} \) as \( t \to \infty \) and \( \dot{P} \) satisfies the equation

\[
\dot{P} = A^1 \left\{ \dot{P} - (P \dot{C}^T - H \dot{R}) (C \dot{P} \dot{C}^T + R)^{-1} (C \dot{P} - RH^T) \right\} (A^1)^T + TSS^TT^T + HHR^T.
\]

(33)
which is called algebraic Riccati equation (ARE). Moreover, using the definitions
\[
\begin{align*}
\hat{G} & := (\hat{P} C^T - H R) \left( C P C^T + R \right)^{-1}, \\
M & := \hat{P} - G \left( C P - R H^T \right),
\end{align*}
\]
we also have
\[
G_t \rightarrow \hat{G}, \quad M_t \rightarrow \hat{M} \quad \text{(as } t \rightarrow \infty)\]
where \(G_t\) and \(M_t\) are defined in Proposition 2.2. \(\blacksquare\)

**Remark 3.4** In view of Lemma 2.6 and Proposition 2.7, we have the following forms of ARE:
\[
P = A^1 \left\{ (I - \hat{G} C) \hat{P} (I - \hat{G} C)^T + \hat{G} \hat{R} C^T + \hat{G} R H^T + H R G T \right\} (A^1)^T + T S S^T T^T + H R H^T,
\]

we now turn to discuss some basic properties of the solutions of ARE. In order to show its uniqueness, we need the following simple formula which can be shown by a simple computation.

**Lemma 3.5** Using the notation \(\psi\) defined by
\[
\psi(P, G) = (I - G C) P (I - G C)^T + G R G^T + H R H^T,
\]
we have
\[
\psi(P^{(1)}, G^{(1)}) - \psi(P^{(2)}, G^{(2)}) = \left( I - G^{(1)} C \right) \left( P^{(1)} - P^{(2)} \right) \left( I - G^{(1)} C \right)^T + \left( G^{(1)} - G^{(2)} \right) \left( C P^{(2)} C^T + R \right) \left( G^{(1)} - G^{(2)} \right)^T.
\]
where \(G^{(i)} = (P^{(i)} C^T - H R) \left( C P^{(i)} C^T + R \right)^{-1}\).

For a solution \(P\) of ARE, we put \(G = (P C^T - H R) \left( C P C^T + R \right)^{-1}\) and call \(P\) a stabilizing solution of ARE if \(\hat{A} := A^1 (I - G C)\) is asymptotically stable.

**Theorem 3.6** Suppose that \((A^1, S)\) is stabilizable and \((C, A^1)\) is detectable. Then, there exists a unique non-negative definite solution \(P\) of ARE (i.e., the equation (39)). Moreover, \(P\) is a stabilizing solution of ARE.

**Proof** Existence of a non-negative definite solution of ARE has been shown in Theorem 3.3 (see \(\hat{P}\) in the theorem). From now on, we use the notation \(P\) instead of \(\hat{P}\). We similarly use the notations \(G\) and \(M\) respectively instead of \(\hat{G}\) and \(\hat{M}\).

We now prove asymptotic stability of the matrix \(\hat{A} := A^1 (I - G C)\). Suppose that \(A^1 (I - G C)\) is not asymptotically stable. Then, there exist \(v \in \mathbb{C}^n\) and \(\lambda \in \mathbb{C} \, (|\lambda| \geq 1)\) such that
\[
(I - G C)^T (A^1)^T v = \lambda v \quad (39)
\]
Since \(P\) is a solution of ARE, we have
\[
P = A^1 \left\{ (I - \hat{G} C) P (I - \hat{G} C)^T + \hat{G} \hat{R} C^T + \hat{G} R H^T \right\} (A^1)^T + T S S^T T^T + H R H^T.
\]
By virtue of (39) and (40), the equality
\[
\left( 1 - |\lambda|^2 \right) v^* P v = v^* A^1 G R G^T (A^1)^T v + v^* T S S^T T^T v + v^* H R H^T v \quad (41)
\]
holds. Here, \(v^* \) denotes the complex conjugate of the transpose of the vector \(v\). In view of \(|\lambda| \geq 1\) and \(R > O\), we have
\[
\begin{align*}
\text{with } \lambda & \neq 0, \quad v^* T S = 0 \quad \text{and } \quad v^* H = 0.
\end{align*}
\]
Notice that the last two equalities imply
\[
S^T \tilde{v} = S^T (T^T \tilde{v} + C^T H^T \tilde{v}) = 0,
\]
where \(\tilde{v}\) is the complex conjugate of \(v\). From these equalities and (39), we have
\[
\begin{align*}
\lambda & \neq 0, \quad v^* S = 0, \quad |\lambda| \geq 1.
\end{align*}
\]
This means that \((A^1, S)\) is not stabilizable. This contradicts our assumption. Thus, \(\hat{A} := A^1 (I - G C)\) is asymptotically stable.

Next, we prove uniqueness of the solutions of ARE. Let \(P^{(1)}\) and \(P^{(2)}\) be two non-negative definite solutions of ARE. Then, the equality
\[
P^{(i)} = A^1 \psi(P^{(i)}, G^{(i)})(A^1)^T + T S S^T T^T + H R H^T
\]
holds, where
\[
G^{(i)} = \left( P^{(i)} C^T - H R \right) \left( C P^{(i)} C^T + R \right)^{-1}
\]
for \(i = 1, 2\). It then follows from Lemma 3.5 that the equality
\[
P^{(1)} - P^{(2)} = A^1 \left\{ \psi(P^{(1)}, G^{(1)}) - \psi(P^{(2)}, G^{(2)}) \right\} (A^1)^T = A^1 \left( I - G^{(1)} C \right) \left( P^{(1)} - P^{(2)} \right) \left( I - G^{(1)} C \right)^T (A^1)^T + A^1 \left[ G^{(1)} - G^{(2)} \right] \left( C P^{(2)} C^T + R \right) \left( G^{(1)} - G^{(2)} \right)^T (A^1)^T
\]
Suppose that $P_e$ is unique.

Since $P_e$ is unique, we have $P_o = \frac{C}{2} \sum \frac{d}{d^2} + R$. Thus, we have $P_o = \frac{C}{2} \sum \frac{d}{d^2} + R$ also. Hence, we have $P^{(1)} - P^{(2)} = O$. Namely, the solution of ARE is unique.

In the next theorem, we prove the convergence results of the sequence $\{P_t\}$ for any $P_0 (\geq O)$ instead of $P_0 = O$ in Theorem 3.3.

**Theorem 3.7** Suppose that $(A^1, S)$ is stabilizable and $(C, A^1)$ is detectable. Then, for any $P_0 \geq O$, the sequence $\{P_t\}$ given by (17) converges to $P$ (the solution of ARE (i.e., the equation (33))). Moreover, $P$ is a unique non-negative definite stabilizing solution of ARE.

**Proof** Let $\{P_t\}$ be the solution of the Riccati equation (17) with $P_0 = O$. Then, due to Theorem 3.3 and Theorem 3.6, we have that $\lim_{t \to \infty} P_t = P$. We denote by $\tilde{P}_t$ the solution of the Riccati equation (17) with any $P_0$ satisfying $P_0 \geq O$. Namely, we suppose that $\tilde{P}_0$ is an arbitrary non-negative definite matrix (see Remark 3.8). We will prove $\lim_{t \to \infty} \tilde{P}_t = P$.

Since $P_t$ and $\tilde{P}_t$ are solutions of the Riccati equation (17), we have

$$\begin{align*}
\tilde{P}_{t+1} - P_{t+1} &= A^1 \left\{ \psi(\tilde{P}_t, \tilde{G}_t) - \psi(P_t, P_t) \right\} (A^1)^T \\
&= A^1 (I - \tilde{G}_t) (\tilde{P}_t - P_t) (I - \tilde{G}_t) (A^1)^T \\
&+ A^1 (\tilde{G}_t - G_t) (C P C^T + R) (\tilde{G}_t - G_t) (A^1)^T \quad (42),
\end{align*}$$

where $G_t = (\tilde{P}_t C^T - H R) (C \tilde{P} C^T + R)^{-1}$.

Notice that $\tilde{P}_0 \geq O (\geq$ the initial matrix of $\{P_t\}$).

Suppose that $P_1 \leq P_2 \leq \cdots \leq P_t \leq \tilde{P}_t$. Then, (42) implies that $P_{t+1} \leq \tilde{P}_{t+1}$. Thus, we obtain

$$P_t \leq \tilde{P}_t, \quad t = 0, 1, \cdots. \quad (43)$$

Since $\{\tilde{P}_t\}$ is monotone non-decreasing (Lemma 3.1) and bounded (Lemma 3.2), we have $\tilde{P}_t \to \tilde{P}$ (as $t \to \infty$), where $\tilde{P}$ is a solution of the ARE (33).

Next, we choose the matrix $L = A^1 \tilde{G} = A^1 (\tilde{P} C^T - H R) (C \tilde{P} C^T + R)^{-1}$ as the asymptotically stable filter in Lemma 3.2. Defining the matrix $\tilde{A}$ by $\tilde{A} := A^1 (I - \tilde{G} C)$, the error covariance matrix $\tilde{P}$ by this filter can be written as

$$\begin{align*}
\tilde{P}_t &= \tilde{A} P_0 (\tilde{A}^T)^T + \sum_{k=0}^{t-1} \tilde{A}^k \left\{ A^1 \tilde{G} R \tilde{G}^T (A^1)^T + A^1 \tilde{G} R H^T \tilde{A}^T + T S S^T T^T + H R H^T \right\} (\tilde{A}^T)^k, \\
&= \tilde{A} P_0 (\tilde{A}^T)^T + \sum_{k=0}^{t-1} \tilde{A}^k \left\{ A^1 \tilde{G} R \tilde{G}^T (A^1)^T + T S S^T T^T + H R H^T \right\} (\tilde{A}^T)^k. \quad (44)
\end{align*}$$

Since this gain $A^1 \tilde{G}$ does not minimize the error covariance matrix, we have $P_t \leq \tilde{P}_t$. Notice that $\tilde{A}$ is asymptotically stable due to Theorem 3.6. By letting $t \to \infty$, we have

$$\lim_{t \to \infty} \tilde{P}_t = \sum_{k=0}^{\infty} \tilde{A}^k \left\{ A^1 \tilde{G} R \tilde{G}^T (A^1)^T + T S S^T T^T + H R H^T \right\} (\tilde{A}^T)^k,$$

where the right hand side is a solution of the ARE (36). Thus, from (43), we obtain

$$P = \lim_{t \to \infty} P_t \leq \lim_{t \to \infty} \tilde{P}_t \leq \lim_{t \to \infty} \tilde{P}_t = P.$$

Hence, we have proved $\lim_{t \to \infty} \tilde{P}_t = P$.

**Remark 3.8** In view of Lemma 3.1, we suppose that $P_0$ in Theorem 3.7 (and $\tilde{P}_0$ in its proof) need to satisfy $0 \leq P_0 \leq P$ and $0 \leq P_0 \leq P$.

### 4 Conclusion

In this paper, we considered discrete-time linear stochastic systems with unknown inputs (or disturbances) and discussed the optimal filter with disturbance decoupling property and fundamental properties of the equation (i.e., Riccati equation) which the covariance matrices of the estimation errors of the filter satisfy. Assuming that the stochastic processes have constant coefficients, we proved convergence of the Riccati equation and derived a simple equation (called the algebraic Riccati equation (ARE)) which is the limit of the Riccati equation. We also proved asymptotic stability of the systems whose optimal gains are determined by the ARE.

### References

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

[2] Bryson, A. E. Jr. and Y. C. Ho, *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, Massachusetts, 1969.
[3] Caliskan, F., H. Mukai, N. Katz and A. Tanikawa, Game estimators for air combat games with unknown enemy inputs, *Proc. American Control Conference*, Denver, Colorado, pp. 5381–5387, 2003.

[4] Chang, S. and P. Hsu, State estimation using general structured observers for linear systems with unknown input, *Proc. 2nd European Control Conference: ECC’93*, Groningen, Holland, pp. 1794–1799, 1993.

[5] Chen, J. and R. J. Patton, Optimal filtering and robust fault diagnosis of stochastic systems with unknown disturbances, *IEE Proc. of Control Theory Applications* Vol. 143, No. 1, pp. 31–36, 1996.

[6] Chen, J. and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic Publishers, Norwell, Massachusetts, 1999.

[7] Chen, J., R. J. Patton and H. -Y. Zhang, Design of unknown input observers and robust fault detection filters, *Int. J. Control* Vol. 63, No. 1, pp. 85–105, 1996.

[8] Darouach, M., M. Zasadzinski and J. Y. Keller, State estimation for discrete systems with unknown inputs using state estimation of singular systems, *Proc. American Control Conference*, pp. 3014–3015, 1992.

[9] Darouach, M., M. Zasadzinski, O. A. Bassang and S. Nowakowski, Kalman filtering with unknown inputs via optimal state estimation of singular systems, *Int. J. Systems Science* Vol. 26, pp. 2015–2028, 1995.

[10] Frank, P. M., Fault diagnosis in dynamic system using analytical and knowledge based redundancy: a survey and some new results, *Automatica* Vol. 26, No. 3, pp. 459–474, 1990.

[11] Hou, M. and P. C. Müller, Unknown input decoupled Kalman filter for time-varying systems, *Proc. 2nd European Control Conference: ECC’93*, Groningen, Holland, pp. 2266–2270, 1993.

[12] Hou, M. and P. C. Müller, Disturbance decoupled observer design: a unified viewpoint, *IEEE Trans. Automatic Control* Vol. 39, No. 6, pp. 1338–1341, 1994.

[13] Hou, M. and R. J. Patton, Optimal filtering for systems with unknown inputs, *IEEE Trans. Automatic Control* Vol. 43, No. 3, pp. 445–449, 1998.

[14] Kailath, T., A view of three decades of linear filtering theory, *IEEE Trans. Inform. Theory Vol. 20*, No. 2, pp. 146–181, 1974.

[15] Kailath, T., *Lectures on Linear Least-Squares Estimation*, Springer, 1976.

[16] Kalman, R. E., A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng. Vol. 82D*, No. 1, pp. 34–45, 1960.

[17] Kalman, R. E., New methods in Wiener filtering theory, in *Proc. of First Symp. Eng. Appl. of Random Function Theory and Probability* (J. L. Bogdanoff and F. Kozin, eds.), Wiley, pp. 270-388, 1963.

[18] Katayama, T., *Applied Kalman Filtering, New Edition*, in Japanese, Asakura-Shoten, Tokyo, Japan, 2000.

[19] Patton, R. J., P. M. Frank and R. N. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Application*, Prentice Hall, 1996.

[20] Sawada, Y. and A. Tanikawa, Optimal filtering and robust fault diagnosis of stochastic systems with unknown inputs and colored observation noises, in *Proc. 5th IASTED Conf. Decision and Control*, Tsukuba, Japan, pp. 149-154, 2002.

[21] Tanikawa, A., On a smoother for discrete-time linear stochastic systems with unknown disturbances, *Int. J. Innovative Computing, Information and Control* Vol. 2, No. 5, pp. 907–916, 2006.

[22] Tanikawa, A., On new smoothing algorithms for discrete-time linear stochastic systems with unknown disturbances, *Int. J. Innovative Computing, Information and Control* Vol. 4, No. 1, pp. 17–27, 2008.

[23] Tanikawa, A., Optimal filters with disturbance decoupling property for nonlinear stochastic systems with unknown disturbances, in *Proc. of the 49th ISCIE Int. Symp. on Stochastic Systems Theory and Its Appl.*, Hiroshima, Japan, pp. 21-26, 2017.

[24] Tanikawa, A. and H. Mukai, Minimum variance state estimators with disturbance decoupling property for optimal filtering problems with unknown inputs, unpublished, 2006.

[25] Tanikawa, A. and Y. Sawada, Minimum variance state estimators with disturbance decoupling property for optimal filtering problems with unknown inputs, in *Proc. of the 35th ISCIE Int. Symp. on Stochastic Systems Theory and Its Appl.*, Ube, Japan, pp. 96-99, 2003.