International Journal of Innovative Computing, Information and Control
Volume **15**, Number **3**, June **2019**

pp. 811–824

ANALYSIS OF 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
1-79-1 Kitayama, Hirakata City, Osaka 573-0196, Japan
akio.tanikawa@oit.ac.jp

Received August 2018; revised December 2018

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 is satisfied by the covariance matrices of the estimation errors of the filter. In this paper, we assume that the stochastic processes have constant coefficients. We then 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. Moreover, we also prove asymptotic stability of the systems whose optimal gains are determined by the ARE.*
Keywords: Stochastic systems, Optimal filter, Unknown inputs, Riccati equation

1. Introduction.

For discrete-time linear stochastic systems with unknown disturbances, we consider the optimal filter with disturbance decoupling property. If modeled systems made by engineers are very accurate representations of real systems, we do not need to consider systems with unknown inputs. However, most of modeled systems are not free from modeling errors in practice, and we should often 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 $\mathbf{Y}_t$ of the outputs $\{y_0, y_1, \ldots, y_t\}$, i.e., $\mathbf{Y}_t = \sigma\{y_s, s = 0, 1, \ldots, t\}$ (the smallest σ-field generated by $\{y_0, y_1, \ldots, y_t\}$ (see, e.g., [17], Chapter 4)). The problem of investigating optimal (or sub-optimal) filters for systems with noise and modelling uncertainty (including unknown disturbances and modelling errors) did not attract enough research attention up to 1980's. This is partly due to a lack of techniques for designing optimal (minimum estimation error variance) filters with disturbance decoupling property for systems with both noise and unknown disturbances. However, in 1990's, remarkable progress has been made in designing optimal filters for stochastic systems with unknown disturbances. Darouach et al. [7, 8] proposed optimal observers with unknown input decoupling property by transforming a standard (time-invariant) system with unknown inputs into a descriptor (singular) system without unknown inputs. Chang and Hsu [4] also studied optimal observers with unknown input decoupling property for time-invariant systems. Hou and Müller [10] investigated the unknown input decoupled filtering for descriptor (singular) sytems with unknown inputs. They utilized some transformations to make the original systems with unknown inputs into descriptor (singular) systems without unknown inputs. In 1996, Chen and Patton [5] proposed ODDO (Optimal Disturbance Decoupling

DOI: 10.24507/ijicic.15.03.811

Observer) for time varying systems with both noise and unknown disturbances. Their ODDO is more straightforward than other works given above and was easily implemented to robust fault diagnosis problem for jet engine systems. Hou and Patton [12] derived a new optimal filtering formula for a linear discrete-time stochastic system with unknown inputs which is slightly different from the stochastic system in this paper. They derived a filtering formula by a new observer design method (innovations filtering technique) in [11] for deterministic continuous-time system with unknown inputs. This paper is closely related to the paper by Chen and Patton [5] and is a continuation of the proceeding papers [26, 27] where the recursive procedure ODDO was modified by the author and his colleagues as the optimal filter with disturbance decoupling property. Later, this optimal filter was utilized to derive the optimal smoothers with disturbance decoupling property in [23, 24]. Moreover, these unknown input decoupled filtering techniques were applied to some specific problems (see, e.g., [3, 18]).

In this paper, we consider the optimal filter with disturbance decoupling property for linear stochastic systems with unknown inputs and fundamental properties of the equation (i.e., Riccati equation) which is satisfied by the covariance matrices of the estimation errors of the filter. This is the first paper to the author's knowledge which is mainly concerned with fundamental properties (e.g., boundedness, monotone convergence and asymptotic stability) of the Riccati equation for the optimal filter with disturbance decoupling property). 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. Moreover, we also prove asymptotic stability of the systems whose optimal gains are determined by the ARE. Finally, in Section 4, we apply the optimal filter proposed in Section 2 to an illustrative example and present the numerical experiments which show that our optimal filter gives better state estimation compared to the standard Kalman filter for the systems with unknown inputs.

2. **Preliminaries.**

Consider the following discrete-time linear stochastic system for $t = 0, 1, 2, \ldots$:

$$x_{t+1} = A_t x_t + B_t u_t + E_t d_t + S_t \zeta_t, \tag{1}$$

$$y_t = C_t x_t + \eta_t, \tag{2}$$

where

$x_t \in \mathbf{R}^n$	the state vector,
$y_t \in \mathbf{R}^m$	the output vector,
$u_t \in \mathbf{R}^r$	the known input vector,
$d_t \in \mathbf{R}^q$	the unknown input vector.

Suppose that ζ_t and η_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 [27], we considered the optimal estimate $\hat{x}_{t+1}$ of the state x_{t+1} which was proposed by Chen and Patton [5, 6] with the following structure:

$$z_{t+1} = F_{t+1}\, z_t + T_{t+1} B_t u_t + K_{t+1} y_t, \tag{3}$$

$$\hat{x}_{t+1} = z_{t+1} + H_{t+1} y_{t+1}, \tag{4}$$

for $t = 0, 1, 2, \ldots$. Here, $\hat{x}_0$ 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 $P_t = \mathbf{E}\{e_t\, e_t{}^T\}$ for $t = 0, 1, 2, \ldots$. Here, $\mathbf{E}$ denotes expectation and T denotes transposition of a matrix. We assume in this paper that random variables e_0, $\{\eta_t\}$, $\{\zeta_t\}$ are independent. As in [5, 6, 27], we consider state estimate (3) and (4) with the matrices $F_{t+1}, T_{t+1}, H_{t+1}$ and K_{t+1} of the forms:

$$K_{t+1} = K_{t+1}^1 + K_{t+1}^2, \tag{5}$$

$$E_t = H_{t+1} C_{t+1} E_t, \tag{6}$$

$$T_{t+1} = I - H_{t+1} C_{t+1}, \tag{7}$$

$$F_{t+1} = A_t - H_{t+1} C_{t+1}\, A_t - K_{t+1}^1 C_t, \tag{8}$$

$$K_{t+1}^2 = F_{t+1} H_t. \tag{9}$$

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

Lemma 2.1. *Equality (6) holds if and only if*

$$rank\,(C_{t+1} E_t) = rank\,(E_t)\,. \tag{10}$$

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

$$H_{t+1} = E_t \left\{ (C_{t+1} E_t)^T (C_{t+1} E_t) \right\}^{-1} (C_{t+1} E_t)^T\,. \tag{11}$$

Hence, we have

$$C_{t+1} H_{t+1} = C_{t+1} E_t \left\{ (C_{t+1} E_t)^T (C_{t+1} E_t) \right\}^{-1} (C_{t+1} E_t)^T \tag{12}$$

which is a non-negative definite symmetric matrix. □

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

$$K_{t+1}^1 = A_{t+1}^1 \left(P_t C_t{}^T - H_t R_t \right) \left(C_t P_t C_t{}^T + R_t \right)^{-1}, \tag{13}$$

$$A_{t+1}^1 = A_t - H_{t+1} C_{t+1} A_t, \tag{14}$$

we obtained the following result (Theorem 2.7 in [27]) on the optimal filtering algorithm under the next condition which is supposed throughout the paper.

Condition A. The matrices $C_t H_t$ and R_t are commutative, i.e.,

$$C_t H_t R_t = R_t C_t H_t, \tag{15}$$

□

Proposition 2.1. *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:*

$$\hat{x}_{t+1} = A_{t+1}^1 \left\{ \hat{x}_t + G_t \left(y_t - C_t \hat{x}_t \right) \right\} + H_{t+1} y_{t+1} + T_{t+1} B_t u_t, \tag{16}$$

$$P_{t+1} = A_{t+1}^1 M_t A_{t+1}^1{}^T + T_{t+1} S_t S_t{}^T T_{t+1}{}^T + H_{t+1} R_{t+1} H_{t+1}{}^T, \tag{17}$$

where

$$G_t = \left(P_t C_t{}^T - H_t R_t \right) \left(C_t P_t C_t{}^T + R_t \right)^{-1}, \tag{18}$$

and

$$M_t = P_t - G_t \left(C_t P_t - R_t H_t{}^T \right). \tag{19}$$

Here, we note that $H_0 = O$ and that Equation (17) is called the Riccati equation. □

Remark 2.1. *If the matrix R_t has the form*

$$R_t = r_t I$$

with some positive number r_t for each $t = 1, 2, \ldots$, then it is obvious to see that condition (15) holds. □

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 [17]).

Proposition 2.2. *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. However, we can choose $H_t \equiv O$ for all t in this case, and the optimal filter given in Proposition 2.1 reduces to the standard Kalman filter.* □

The optimal filter proposed in Proposition 2.1 is applied to an illustrative example in Section 4 to show its effectiveness. The numerical experiments there indicate that our optimal filter gives better state estimation compared to the standard Kalman filter for the system with unknown inputs. In the next section, we consider the case where all coefficient matrices of the system (1)-(2) are independent of time. In this case, we expect convergence of the matrices $\{P_t\}$ in Proposition 2.1. In fact, we can prove its convergence under some reasonable conditions in Section 3. As a preparation, we need to rewrite the Riccati equation (17) since it is somewhat complicated.

First, we note that some matrices are projection matrices which will play important roles later and can be proved by simple computations (see [23] also).

Lemma 2.2. *Matrices ${C_t}^T {H_t}^T$ and ${T_t}^T = I - {C_t}^T {H_t}^T$ are projection matrices which have the following properties:*

$$\left({C_t}^T {H_t}^T\right)\left({C_t}^T {H_t}^T\right) = {C_t}^T {H_t}^T, \tag{20}$$

$$\left(I - {C_t}^T {H_t}^T\right)\left(I - {C_t}^T {H_t}^T\right) = I - {C_t}^T {H_t}^T, \tag{21}$$

$${C_t}^T {H_t}^T \left(I - {C_t}^T {H_t}^T\right) = O, \tag{22}$$

and moreover

$${H_t}^T \left(I - {C_t}^T {H_t}^T\right) = O. \tag{23}$$

□

Second, by using the equalities in Lemma 2.2, we can prove the following equalities.

Lemma 2.3. *For $t = 0, 1, 2, \ldots$, we have*

$$G_t R_t {H_t}^T = O \quad \text{and} \quad H_t R_t {G_t}^T = O. \tag{24}$$

□

Proof: For $t = 0$, the equalities obviously hold. For $t \geq 1$, we note that P_t has the form

$$P_t = T_t \Gamma_t {T_t}^T + H_t R_t {H_t}^T \tag{25}$$

with some matrix Γ_t in view of (17). Due to Lemma 3.5 and Remark 3.7 in [27], we have

$$\begin{aligned}
G_t R_t {H_t}^T &= \left(P_t {C_t}^T - H_t R_t\right)\left(C_t P_t {C_t}^T + R_t\right)^{-1} R_t {H_t}^T \\
&= \left(P_t {C_t}^T - H_t R_t\right) \times \Big\{ {H_t}^T {C_t}^T \left(C_t P_t {C_t}^T + R_t\right)^{-1} \\
&\quad + \left(I - {H_t}^T {C_t}^T\right)\left(C_t P_t {C_t}^T + R_t\right)^{-1} \Big\} R_t {H_t}^T \\
&= \left(P_t {C_t}^T {H_t}^T - H_t R_t {H_t}^T\right) {C_t}^T \left(C_t P_t {C_t}^T + R_t\right)^{-1} R_t {H_t}^T
\end{aligned}$$

$$+\left(P_t{C_t}^T - H_tR_t\right)\left(C_tP_t{C_t}^T + R_t\right)^{-1}\left(I - {H_t}^T{C_t}^T\right)R_t{H_t}^T. \tag{26}$$

Noting that

$$P_t{C_t}^T{H_t}^T - H_tR_t{H_t}^T = T_t\Gamma_t{T_t}^T{C_t}^T{H_t}^T + H_tR_t{H_t}^T{C_t}^T{H_t}^T - H_tR_t{H_t}^T = O$$

holds due to (22), (23) and (25) and that

$$\left(I - {H_t}^T{C_t}^T\right)R_t{H_t}^T = R_t\left(I - {H_t}^T{C_t}^T\right){H_t}^T = O$$

holds due to (23), we have that the right hand side of (26) is equal to O. By taking matrix transposition, we can obtain the second equality from the first one. □

Then, we can obtain the following simplified form of the Riccati equation (17).

Proposition 2.3. *For $t = 0, 1, 2, \ldots$, we have*

$$M_t = \left(I - G_tC_t\right)P_t\left(I - G_tC_t\right)^T + G_tR_t{G_t}^T \tag{27}$$

and the Riccati equation

$$\begin{aligned} P_{t+1} &= A^1_{t+1}M_t{A^1_{t+1}}^T + T_{t+1}S_t{S_t}^T{T_{t+1}}^T + H_{t+1}R_{t+1}{H_{t+1}}^T \\ &= A^1_{t+1}\left(I - G_tC_t\right)P_t\left(I - G_tC_t\right)^T{A^1_{t+1}}^T + A^1_{t+1}G_tR_t{G_t}^T{A^1_{t+1}}^T \\ &\quad + T_{t+1}S_t{S_t}^T{T_{t+1}}^T + H_{t+1}R_{t+1}{H_{t+1}}^T. \end{aligned} \tag{28}$$

□

Proof: We have

$$\begin{aligned} &\left(I - G_tC_t\right)P_t\left(I - G_tC_t\right)^T + G_tR_t{G_t}^T \\ = {}& P_t - G_tC_tP_t - P_t{C_t}^T{G_t}^T + G_tC_tP_t{C_t}^T{G_t}^T + G_tR_t{G_t}^T \\ = {}& P_t - G_t\left(P_t{C_t}^T - H_tR_t\right)^T - \left(P_t{C_t}^T - H_tR_t\right){G_t}^T + G_t\left(C_tP_t{C_t}^T + R_t\right){G_t}^T \\ &- G_tR_t{H_t}^T - H_tR_t{G_t}^T \\ = {}& P_t - \left(P_t{C_t}^T - H_tR_t\right)\left(C_tP_t{C_t}^T + R_t\right)^{-1}\left(C_tP_t - R_t{H_t}^T\right) - G_tR_t{H_t}^T - H_tR_t{G_t}^T \\ = {}& M_t - G_tR_t{H_t}^T - H_tR_t{G_t}^T. \end{aligned} \tag{29}$$

It then follows from Lemma 2.3 that Equality (28) holds. □

3. **Stochastic Systems with Constant Coefficients.** In this section, we consider the case where all coefficient matrices of the system (1)-(2) are independent of time. In this case, we can prove that the sequence of matrices $\{P_t\}$ in Proposition 2.1 converges to a matrix $\bar{P}$ as $t \to \infty$. Then, by solving the matrix equation of $\bar{P}$ derived from the Riccati equation (17), we can easily obtain the optimal filter by using $\bar{P}$ without solving (17) one by one. The following proofs in this section were inspired by those for the Kalman filter by Hewer [9] and Katayama [17].

From now on, we consider the following discrete-time linear stochastic system with constant coefficients for $t = 0, 1, 2, \ldots$:

$$x_{t+1} = Ax_t + Bu + Ed + S\zeta_t, \tag{30}$$

$$y_t = Cx_t + \eta_t. \tag{31}$$

Namely, the matrices A_t, B_t, C_t, E_t, S_t, R_t and the vectors u_t and 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. ζ_t and η_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 a real number α (with $0 \le \alpha \le 1$) and a symmetric matrix U, we set

$$\Phi(U) = A^1 \left\{ U - \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1}\left(CU - \alpha RH^T\right) \right\} \left(A^1\right)^T + TSS^TT^T + HRH^T. \tag{32}$$

Then, it is easy to observe that the sequence $\{P_t\}$ defined by (17) satisfies $P_1 = \Phi(P_0)$ with $\alpha = 0$ and $P_{t+1} = \Phi(P_t)$ with $\alpha = 1$ for $t \ge 1$. 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 $Q_2 \ge Q_1$, then $\Phi(Q_2) \ge \Phi(Q_1)$.* □

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

$$\begin{aligned}
&\Phi(Q_2) - \Phi(Q_1) = \int_0^1 \frac{d}{ds}\Phi(U(s))ds \\
&= A^1 \left[\int_0^1 \left\{ \frac{d}{ds}U(s) - \left(\frac{d}{ds}U(s)\right) C^T \left(CUC^T + R\right)^{-1}\left(CU - \alpha RH^T\right) \right.\right. \\
&\quad + \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1} C \left(\frac{d}{ds}U(s)\right) C^T \left(CUC^T + R\right)^{-1} \\
&\quad \left.\left. \times \left(CU - \alpha RH^T\right) - \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1} C \left(\frac{d}{ds}U(s)\right)\right\} ds\right] \left(A^1\right)^T \\
&= A^1 \left[\int_0^1 \left\{ (Q_2 - Q_1) - (Q_2 - Q_1)\, C^T \left(CUC^T + R\right)^{-1}\left(CU - \alpha RH^T\right) \right.\right. \\
&\quad + \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1} C\,(Q_2 - Q_1)\, C^T \left(CUC^T + R\right)^{-1} \\
&\quad \left.\left. \times \left(CU - \alpha RH^T\right) - \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1} C\,(Q_2 - Q_1) \right\} ds\right] \left(A^1\right)^T \\
&= A^1 \left[\int_0^1 \left\{ \left(I - \left(UC^T - \alpha HR\right)\left(CUC^T + R\right)^{-1} C\right)(Q_2 - Q_1) \right.\right. \\
&\quad \left.\left. \times \left(I - C^T\left(CUC^T + R\right)^{-1}\left(CU - \alpha RH^T\right)\right)\right\} ds\right] \left(A^1\right)^T \\
&= A^1 \left[\int_0^1 W(s)\,(Q_2 - Q_1)\,W(s)^T ds\right] \left(A^1\right)^T \ge 0,
\end{aligned}$$

where

$$W(s) = I - \left(U(s)C^T - \alpha HR\right)\left(CU(s)C^T + R\right)^{-1} C.$$

□

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

$$P_0 (= O) \le P_1 \le P_2 \le P_3 \le \cdots \tag{33}$$

We now give two definitions to discuss convergence of the sequence of matrices $\{P_t\}$. (A^1, S) is said to be *stabilizable* iff there is a matrix L such that $A^1 + SL$ is asymptotically stable. (C, A^1) is said to be *detectable* iff 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 $L \in \mathbf{R}^n \times \mathbf{R}^m$ such that $\breve{A} := A^1 - LC$ is asymptotically stable. Instead of the dynamical system (16), we consider the following filter by substituting L into $A^1 G_t$

$$\begin{aligned}\breve{x}_{t+1} &= \left(A^1 - LC\right)\breve{x}_t + Ly_t + Hy_{t+1} + TBu \\ &= \left(A^1 - LC\right)\breve{x}_t + L\left(Cx_t + \eta_t\right) + H\left(Cx_{t+1} + \eta_{t+1}\right) + TBu \\ &= \left(A - HCA - LC\right)\breve{x}_t + LCx_t + L\eta_t + HC\left(Ax_t + Ed + S\zeta_t\right) + H\eta_{t+1} + TBu\end{aligned}$$

with $\breve{x}_0 = \bar{x}_0$. Hence, we have

$$\begin{aligned}x_{t+1} - \breve{x}_{t+1} &= \left(A - HCA - LC\right)\left(x_t - \breve{x}_t\right) + \left(I - HC\right)Ed + \left(I - HC\right)S\zeta_t \\ &\quad + \left(I - HC - T\right)Bu - L\eta_t - H\eta_{t+1} \\ &= \left(A - HCA - LC\right)\left(x_t - \breve{x}_t\right) + TS\zeta_t - L\eta_t - H\eta_{t+1}.\end{aligned}$$

Using the notations $\breve{e}_t := x_t - \breve{x}_t$, $\breve{P}_t := \mathbf{E}\left\{\breve{e}_t\breve{e}_t^T\right\}$, we have

$$\breve{P}_{t+1} = \breve{A}\breve{P}_t\breve{A}^T + LRL^T + HRH^T + TSS^TT^T + \breve{A}HRL^T + LRH^T\breve{A}^T \tag{34}$$

with $\breve{P}_0 = \Sigma_0$ due to Lemma 1.1 in [26]. It then follows from $\breve{P}_0 = \Sigma_0 = P_0$ that

$$\begin{aligned}\breve{P}_t &= \breve{A}^t\Sigma_0\left(\breve{A}^T\right)^t + \sum_{k=0}^{t-1}\breve{A}^k\left(LRL^T + HRH^T + TSS^TT^T + \breve{A}HRL^T\right. \\ &\quad \left. + LRH^T\breve{A}^T\right)\left(\breve{A}^T\right)^k.\end{aligned} \tag{35}$$

Since $\breve{P}_t$ is identical to P_t (the optimal covariance matrix) when we choose $L = A^1G_t$, we note that $P_t \leq \breve{P}_t$. Due to asymptotic stability of $\breve{A}$, the right hand side of (35) converges as $t \to \infty$. Hence, we have

$$\breve{P}_t \leq \Sigma_0 + \sum_{k=0}^{\infty}\breve{A}^k\left(LRL^T + HRH^T + TSS^TT^T + \breve{A}HRL^T + LRH^T\breve{A}^T\right)\left(\breve{A}^T\right)^k < \infty$$

and boundedness of $\{P_t\}$. □

In view of (33) and Lemma 3.2, we can obtain the following convergence results of the sequence $\{P_t\}$.

Theorem 3.1. *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 $\bar{P}$ as $t \to \infty$ and $\bar{P}$ satisfies the equation*

$$\begin{aligned}\bar{P} &= A^1\left\{\bar{P} - \left(\bar{P}C^T - HR\right)\left(C\bar{P}C^T + R\right)^{-1}\left(C\bar{P} - RH^T\right)\right\}\left(A^1\right)^T \\ &\quad + TSS^TT^T + HRH^T\end{aligned} \tag{36}$$

which is called algebraic Riccati equation (ARE). Moreover, using the definitions

$$\bar{G} := \left(\bar{P}C^T - HR\right)\left(C\bar{P}C^T + R\right)^{-1}, \tag{37}$$

$$\bar{M} := \bar{P} - \bar{G}\left(C\bar{P} - RH^T\right), \tag{38}$$

we also have

$$G_t \longrightarrow \bar{G}, \qquad M_t \longrightarrow \bar{M} \qquad (\text{as } t \longrightarrow \infty)$$

where G_t and M_t are defined in Proposition 2.1. □

Proof: Since $\{P_t\}$ is bounded above, there exists a positive definite matrix W such that the inequalities

$$0 \leq v^T P_t v \leq v^T W v$$

hold for any n-vector v. Let us choose $v = [0 \ldots 0\,1\,0 \ldots 0]^T$ with 1 being placed at the ith component. Setting $P_t = \left(p_{ij}{}^{(t)}\right)$, we have

$$v^T P_t v = p_{ii}{}^{(t)} \longrightarrow p_{ii} \qquad (\text{as } t \longrightarrow \infty)$$

for some non-negative number p_{ii} since the sequence $p_{ii}{}^{(t)}$ is a non-negative monotonic non-decreasing and bounded sequence. Next, choosing $v = [0 \cdots 0\,1\,0 \cdots 0\,1\,0 \cdots 0]^T$ with two 1's being placed at the ith and jth components, we have

$$v^T P_t v = p_{ii}{}^{(t)} + p_{ij}{}^{(t)} + p_{ji}{}^{(t)} + p_{jj}{}^{(t)} = p_{ii}{}^{(t)} + 2p_{ij}{}^{(t)} + p_{jj}{}^{(t)} \longrightarrow q \qquad (\text{as } t \longrightarrow \infty)$$

for some non-negative number q. Noticing that $p_{ii}{}^{(t)} \to p_{ii}$, we have

$$p_{ij}{}^{(t)} \longrightarrow \frac{1}{2}\left(q - p_{ii} - p_{jj}\right) \qquad (\text{as } t \longrightarrow \infty)$$

Thus, we have $P_t \to \bar{P}$ for some non-negative and symmetric matrix $\bar{P}$ since P_t is non-negative and symmetric. The rest of the conclusions immediately follow from this. □

Remark 3.1. In view of Lemma 2.3 and Proposition 2.3, we have the following forms of ARE:

$$\begin{aligned} \bar{P} &= A^1 \left\{ \left(I - \bar{G}C\right) \bar{P} \left(I - \bar{G}C\right)^T + \bar{G}R\bar{G}^T + \bar{G}RH^T + HR\bar{G}^T \right\} \left(A^1\right)^T \\ &\quad + TSS^T T^T + HRH^T, \qquad (39) \\ \bar{P} &= A^1 \left\{ \left(I - \bar{G}C\right) \bar{P} \left(I - \bar{G}C\right)^T + \bar{G}R\bar{G}^T \right\} \left(A^1\right)^T + TSS^T T^T + HRH^T. \qquad (40) \end{aligned}$$

□

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 3.3. *Using the notation ψ defined by*

$$\psi(P, G) = (I - GC)\,P\,(I - GC)^T + GRG^T + GRH^T + HRG^T,$$

we have

$$\begin{aligned} \psi\left(P^{(1)}, G^{(1)}\right) - \psi\left(P^{(2)}, G^{(2)}\right) &= \left(I - G^{(1)}C\right)\left(P^{(1)} - P^{(2)}\right)\left(I - G^{(1)}C\right)^T \\ &\quad + \left(G^{(1)} - G^{(2)}\right)\left(CP^{(2)}C^T + R\right)\left(G^{(1)} - G^{(2)}\right)^T, \qquad (41) \end{aligned}$$

where $G^{(i)} = \left(P^{(i)}C^T - HR\right)\left(CP^{(i)}C^T + R\right)^{-1}$. □

For a solution P of ARE, we put $G = \left(PC^T - HR\right)\left(CPC^T + R\right)^{-1}$ and call P a stabilizing solution of ARE if $\widetilde{A} := A^1(I - GC)$ is asymptotically stable.

Theorem 3.2. *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., Equation (36)). Moreover, P is a stabilizing solution of ARE.* □

Proof: Existence of a non-negative definite solution of ARE has been shown in Theorem 3.1 (i.e., $\bar{P}$ in that theorem). From now on, we use the notation P instead of $\bar{P}$. We similarly use the notations G and M respectively instead of $\bar{G}$ and $\bar{M}$.

We now prove asymptotic stability of the matrix $\widetilde{A} := A^1(I-GC)$ in a similar way to the case of the Kalman filter. Suppose that $A^1(I-GC)$ is not asymptotically stable. Then, there exist $v \in \mathbf{C}^n$ and $\lambda \in \mathbf{C}(|\lambda| \geq 1)$ such that

$$(I-GC)^T\left(A^1\right)^T v = \lambda v \tag{42}$$

Since P is a solution of ARE, we have

$$P = A^1\left\{(I-GC)P(I-GC)^T + GRG^T\right\}\left(A^1\right)^T + TSS^TT^T + HRH^T. \tag{43}$$

By virtue of (42) and (43), the equality

$$\left(1-|\lambda|^2\right)v^*Pv = v^*A^1GRG^T\left(A^1\right)^T v + v^*TSS^TT^Tv + v^*HRH^Tv \tag{44}$$

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

$$v^*A^1G = 0, \quad v^*TS = 0 \quad \text{and} \quad v^*H = 0.$$

Notice that the last two equalities imply

$$S^T\bar{v} = S^T\left(T^T\bar{v} + C^TH^T\bar{v}\right) = 0,$$

where $\bar{v}$ is the complex conjugate of v. From these equalities and (42), we have

$$v^*A^1 = \lambda v^*, \quad v^*S = 0, \quad |\lambda| \geq 1.$$

This means that (A^1, S) is not stabilizable. This contradicts our assumption. Thus, $\widetilde{A} := A^1(I-GC)$ 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

$$P^{(i)} = A^1\psi\left(P^{(i)}, G^{(i)}\right)\left(A^1\right)^T + TSS^TT^T + HRH^T$$

holds, where $G^{(i)} = \left(P^{(i)}C^T - HR\right)\left(CP^{(i)}C^T + R\right)^{-1}$, $i = 1, 2$. It then follows from Lemma 3.3 that the equality

$$\begin{aligned} P^{(1)} - P^{(2)} &= A^1\left\{\psi\left(P^{(1)}, G^{(1)}\right) - \psi\left(P^{(2)}, G^{(2)}\right)\right\}\left(A^1\right)^T \\ &= A^1\left(I - G^{(1)}C\right)\left(P^{(1)} - P^{(2)}\right)\left(I - G^{(1)}C\right)^T\left(A^1\right)^T \\ &\quad + A^1\left(G^{(1)} - G^{(2)}\right)\left(CP^{(2)}C^T + R\right)\left(G^{(1)} - G^{(2)}\right)^T\left(A^1\right)^T \end{aligned}$$

holds. Since asymptotic stability of $\widetilde{A} = A^1\left(I - G^{(1)}C\right)$ follows from the first part of Proof, we have the equality

$$P^{(1)} - P^{(2)} = \sum_{k=0}^{\infty} \widetilde{A}^k A^1\left(G^{(1)} - G^{(2)}\right)\left(CP^{(2)}C^T + R\right)\left(G^{(1)} - G^{(2)}\right)^T\left(A^1\right)^T\left(\widetilde{A}^T\right)^k$$

Thus, we have $P^{(1)} - P^{(2)} \geq O$. Since we can exchange the indexes 1 and 2 and $\check{A} = A^1\left(I - G^{(2)}C\right)$ is asymptotically stable, we have $P^{(2)} - P^{(1)} \geq O$ 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 3.1.

Theorem 3.3. *Suppose that (A^1, S) is stabilizable and (C, A^1) is detectable. Then, for any $P_0 \geq 0$, the sequence $\{P_t\}$ given by (17) converges to P (the solution of ARE (i.e., Equation (36)). 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 3.1 and Theorem 3.2, we have that $\lim_{t\to\infty} P_t = P$. We denote by $\widetilde{P}_t$ the solution of the Riccati equation (17) with any P_0 satisfying $P_0 \geq O$. Namely, we suppose that $\widetilde{P}_0$ is an arbitrary non-negative definite matrix (see Remark 3.2). We will prove $\lim_{t\to\infty} \widetilde{P}_t = P$.

Since P_t and $\widetilde{P}_t$ are solutions of the Riccati equation (17), we have by Lemma 3.3

$$\begin{aligned}\widetilde{P}_{t+1} - P_{t+1} &= A^1 \left\{\psi\left(\widetilde{P}_t, \widetilde{G}_t\right) - \psi(P_t, G_t)\right\} \left(A^1\right)^T \\ &= A^1\left(I - \widetilde{G}_t C\right)\left(\widetilde{P}_t - P_t\right)\left(I - \widetilde{G}_t C\right)^T \left(A^1\right)^T \\ &\quad + A^1\left(\widetilde{G}_t - G_t\right)\left(CP_tC^T + R\right)\left(\widetilde{G}_t - G_t\right)^T\left(A^1\right)^T, \end{aligned} \tag{45}$$

where $\widetilde{G}_t = \left(\widetilde{P}_t C^T - HR\right)\left(C\widetilde{P}_t C^T + R\right)^{-1}$.

Notice that $\widetilde{P}_0 \geq O$ (= the initial matrix of $\{P_t\}$). Suppose that $P_1 \leq \widetilde{P}_1, \cdots, P_t \leq \widetilde{P}_t$. Then, (45) implies that $P_{t+1} \leq \widetilde{P}_{t+1}$. Thus, we obtain

$$P_t \leq \widetilde{P}_t, \quad t = 0, 1, \ldots. \tag{46}$$

Since $\left\{\widetilde{P}_t\right\}$ is monotone non-decreasing (Lemma 3.1) and bounded (Lemma 3.2), we have $\widetilde{P}_t \to \widetilde{P}$ (as $t \to \infty$), where $\widetilde{P}$ is a solution of the ARE (36).

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

$$\begin{aligned}\widehat{P}_t &= \widetilde{A}^t \widetilde{P}_0 \left(\widetilde{A}^T\right)^t + \sum_{k=0}^{t-1} \widetilde{A}^k \Big\{ A^1\widetilde{G}R\widetilde{G}^T\left(A^1\right)^T + \widetilde{A}HR\widetilde{G}^T\left(A^1\right)^T + A^1\widetilde{G}RH^T\widetilde{A}^T \\ &\quad + TSS^TT^T + HRH^T \Big\}\left(\widetilde{A}^T\right)^k. \\ &= \widetilde{A}^t \widetilde{P}_0 \left(\widetilde{A}^T\right)^t + \sum_{k=0}^{t-1} \widetilde{A}^k \left\{ A^1\widetilde{G}R\widetilde{G}^T\left(A^1\right)^T + TSS^TT^T + HRH^T \right\}\left(\widetilde{A}^T\right)^k. \end{aligned} \tag{47}$$

Since this gain $A^1\widetilde{G}$ does not minimize the error covariance matrix, we have $P_t \leq \widehat{P}_t$. Notice that $\widetilde{A}$ is asymptotically stable due to Theorem 3.2. By letting $t \to \infty$, we have

$$\lim_{t\to\infty} \widehat{P}_t = \sum_{k=0}^{\infty} \widetilde{A}^k \left\{A^1\widetilde{G}R\widetilde{G}^T\left(A^1\right)^T + TSS^TT^T + HRH^T\right\}\left(\widetilde{A}^T\right)^k,$$

where the right hand side is a solution of the ARE (39). Thus, from (46), we obtain

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

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

Remark 3.2. In view of Lemma 3.1, we suppose that P_0 in Theorem 3.3 (and $\widetilde{P}_0$ in its proof) need to satisfy $0 \leq P_0 \leq P$ (and $0 \leq \widetilde{P}_0 \leq P$). □

4. **Numerical Simulation.** We applied the optimal filtering algorithm given in Section 2 and the (standard) Kalman filter to a simple example and compared the results. We consider the discrete-time model of a simplified longitudinal flight control system given in Chen and Patton [5]:

$$x_{t+1} = A_t x_t + B_t u_t + E_t d_t + S_t \zeta_t, \tag{48}$$

$$y_t = C_t x_t + \eta_t, \tag{49}$$

where the state variables x are normal velocity η_y, pitch angle δ_z and pitch rate ω_z, i.e., $x = [\eta_y, \omega_z, \delta_z]^T$. And the control input u_t is elevator control signal. Here, the system matrices are

$$A_t = \begin{bmatrix} 0.9944 & -0.1203 & -0.4302 \\ 0.0017 & 0.9902 & -0.0747 \\ 0 & 0.8187 & 0 \end{bmatrix},$$

$$B_t = \begin{bmatrix} 0.4252 \\ -0.0082 \\ 0.1813 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and $C_t = I_{3\times3}$. The matrices related to the noise sequences are: $S_t = \mathrm{diag}\{0.1, 0.1, 0.01\}$ and $R_t = 0.1^2 I_{3\times3}$. In this case, we can easily observe that (A^1, S) is stabilizable and (C, A^1) is detectable. The term $E_t d_t$ is used to represent the parameter perturbation in matrices A_t and B_t:

$$E_t = \Delta A_t x_t + \Delta B_t u_t = E \left\{ \begin{bmatrix} \Delta a_{11} & \Delta a_{12} & \Delta a_{13} \\ \Delta a_{21} & \Delta a_{22} & \Delta a_{23} \end{bmatrix} x_t + \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \end{bmatrix} u_t \right\},$$

where Δa_{ij} and Δb_i ($i = 1, 2$; $j = 1, 2, 3$) are perturbations in aerodynamic and control coefficients. We consider these terms as unknown disturbances which should be decoupled from the state estimation using the method given in Section 2. In the simulation, the aerodynamic coefficients are perturbed by $\pm 50\%$, i.e., $\Delta a_{ij} = -0.5 a_{ij}$ and $\Delta b_i = 0.5 b_i$.

The stochastic system (48)-(49) is obviously detectable and stabilizable. The initial state of system (48) is given by $x(0) = [0, 0, 0]^T$, and the known input is given by $u(t) \equiv 10$. We set $\hat{x}_0 = [0, 0, 0]^T$ and $P_0 = 0.1^2 I_{3\times3}$.

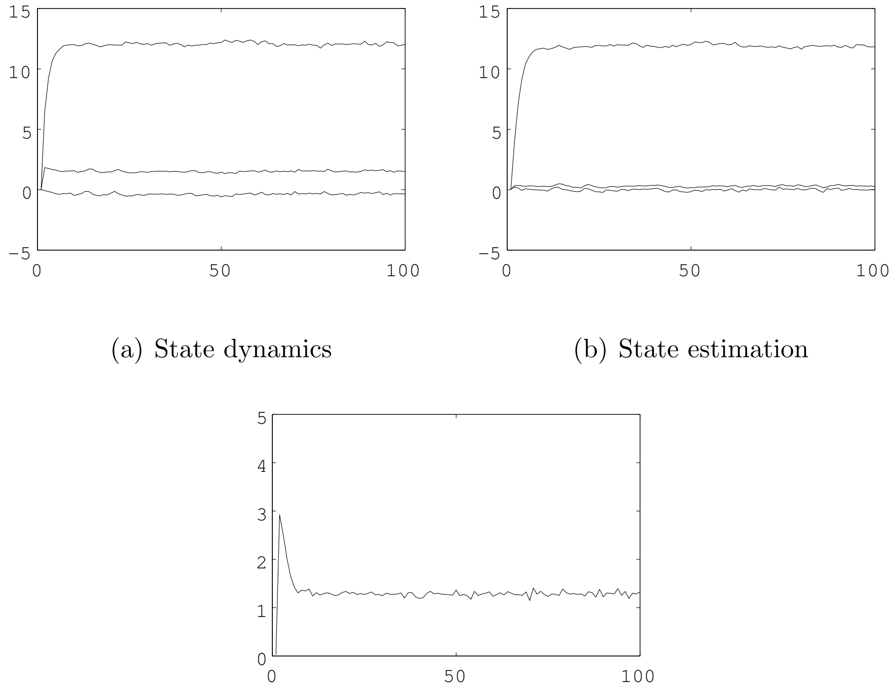
The simulation results by the standard Kalman filter (not disturbance decoupled) for $1 \le t \le 100$ are shown in Figure 1. While 1(a) depicts the trajectory of stochastic system (48), the trajectory of the state estimate $\hat{x}_t$ by the Kalman filter is shown in 1(b) and the absolute value of the estimation error $\hat{x}_t - x_t$ is shown in 1(c).

The simulation results by the optimal filter with disturbance decoupling property given in Section 2 are shown in Figure 2. While 2(a) depicts the trajectory of stochastic system (48), the trajectory of the state estimate $\hat{x}_t$ by the optimal filter in Section 2 is shown in 2(b) and the absolute value of the estimation error $\hat{x}_t - x_t$ is shown in 2(c).

It can be seen from Figures 1 and 2 that the optimal filter proposed in Section 2 works better than the standard Kalman filter for the simple stochastic system with unknown inputs given above.

5. **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. Moreover, we also proved asymptotic stability of the

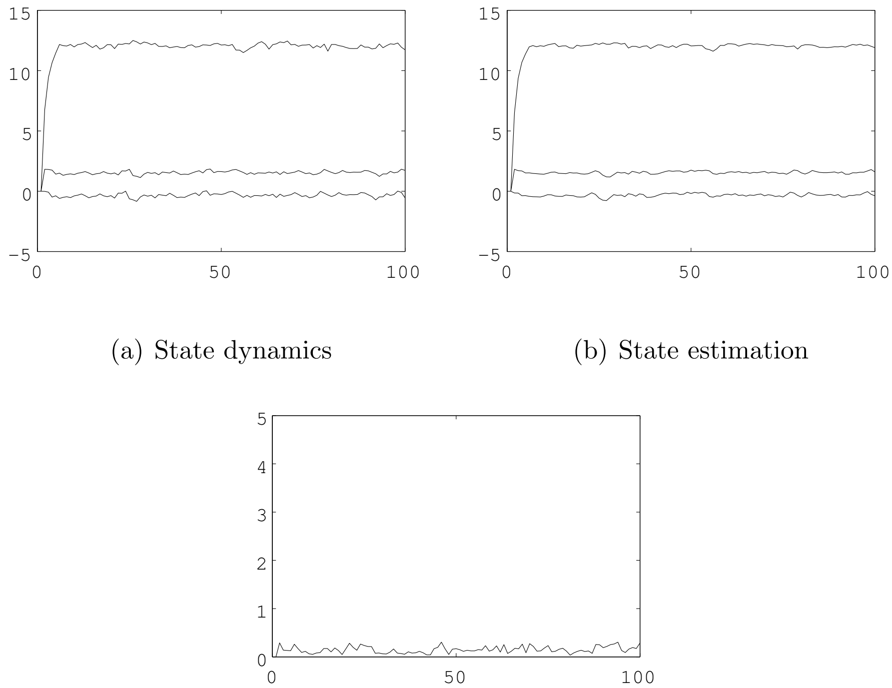
systems whose optimal gains are determined by the ARE. Finally, one of the interesting future research directions will be to find out the rate of the convergence $P_t \to \bar{P}$ as $t \to \infty$.



(a) State dynamics (b) State estimation

(c) Estimation error

Figure 1. Simulation result via Kalman filter



(a) State dynamics (b) State estimation

(c) Estimation error

Figure 2. Simulation result via the filter in Section 2

Acknowledgments. The author wishes to thank Hisae Tanikawa for her encouragement and support. He also wishes to express gratitude to Professor Takashi Saito and Professor Hideo Araki for kind advice, and to the associate editor and the anonymous referees for valuable comments which have improved the earlier version of the paper.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979.
[2] A. E. Jr. Bryson and Y. C. Ho, *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, Massachusetts, 1969.
[3] F. Caliskan, H. Mukai, N. Katz and A. Tanikawa, Game estimators for air combat games with unknown enemy inputs, *Proc. of American Control Conference*, Denver, Colorado, pp.5381-5387, 2003.
[4] S. Chang and P. Hsu, State estimation using general structured observers for linear systems with unknown input, *Proc. of the 2nd European Control Conference: ECC'93*, Groningen, Holland, pp.1794-1799, 1993.
[5] J. Chen 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] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic Publishers, Norwell, Massachusetts, 1999.
[7] M. Darouach, M. Zasadzinski and J. Y. Keller, State estimation for discrete systems with unknown inputs using state estimation of singular systems, *Proc. of American Control Conference*, pp.3014-3015, 1992.
[8] M. Darouach, 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.
[9] G. A. Hewer, Analysis of a discrete matrix Riccati equation of linear control and Kalman filtering, *J. Math. Anal. Appl.*, vol.42, no.1, pp.226-236, 1973.
[10] M. Hou and P. C. Müller, Unknown input decoupled Kalman filter for time-varying systems, *Proc. of the 2nd European Control Conference: ECC'93*, Groningen, Holland, pp.2266-2270, 1993.
[11] M. Hou and P. C. Müller, Disturbance decoupled observer design: A unified viewpoint, *IEEE Trans. Automatic Control*, vol.39, no.6, pp.1338-1341, 1994.
[12] M. Hou and R. J. Patton, Optimal filtering for systems with unknown inputs, *IEEE Trans. Automatic Control*, vol.43, no.3, pp.445-449, 1998.
[13] T. Kailath, A view of three decades of linear filtering theory, *IEEE Trans. Inform. Theory*, vol.20, no.2, pp.146-181, 1974.
[14] T. Kailath, *Lectures on Linear Least-Squares Estimation*, Springer, 1976.
[15] R. E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng.*, vol.82D, no.1, pp.34-45, 1960.
[16] R. E. Kalman, New methods in Wiener filtering theory, *Proc. of the 1st Symp. Eng. Appl. of Random Function Theory and Probability*, pp.270-388, 1963.
[17] T. Katayama, *Applied Kalman Filtering*, New Edition, Asakura-Shoten, Tokyo, Japan, 2000 (in Japanese).
[18] Y. Li and M. Chen, Composite anti-disturbance control for singular stochastic Markovian jump system, *ICIC Express Letters*, vol.12, no.4, pp.321-330, 2018.
[19] B. P. Molinari, The stabilizing solution of the discrete algebraic Riccati equation, *IEEE Trans. Automatic Control*, vol.20, no.3, pp.396-399, 1975.
[20] R. J. Patton, P. M. Frank and R. N. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Application*, Prentice Hall, 1996.
[21] H. J. Payne and L. M. Silverman, On the discrete time algebraic Riccati equation, *IEEE Trans. Automatic Control*, vol.18, no.3, pp.226-234, 1973.
[22] Y. Sawada and A. Tanikawa, Optimal filtering and robust fault diagnosis of stochastic systems with unknown inputs and colored observation noises, *Proc. of the 5th IASTED Conf. Decision and Control*, Tsukuba, Japan, pp.149-154, 2002.
[23] A. Tanikawa, On a smoother for discrete-time linear stochastic systems with unknown disturbances, *International Journal of Innovative Computing, Information and Control*, vol.2, no.5, pp.907-916, 2006.

[24] A. Tanikawa, On new smoothing algorithms for discrete-time linear stochastic systems with unknown disturbances, *International Journal of Innovative Computing, Information and Control*, vol.4, no.1, pp.15-24, 2008.
[25] A. Tanikawa, Optimal filters with disturbance decoupling property for nonlinear stochastic systems with unknown disturbances, *Proc. of the 49th ISCIE Int. Symp. on Stochastic Systems Theory and Its Appl.*, Hiroshima, Japan, pp.21-26, 2017.
[26] A. Tanikawa and H. Mukai, *Minimum Variance State Estimators with Disturbance Decoupling Property for Optimal Filtering Problems with Unknown Inputs*, unpublished lecture note, 2006.
[27] A. Tanikawa and Y. Sawada, Minimum variance state estimators with disturbance decoupling property for optimal filtering problems with unknown inputs, *Proc. of the 35th ISCIE Int. Symp. on Stochastic Systems Theory and Its Appl.*, Ube, Japan, pp.96-99, 2003.