International Journal of Innovative Computing, Information and Control
Volume **18**, Number **1**, February **2022**
ICIC International ©2022 ISSN 1349-4198
pp. **305–314**

ON THE RATE OF CONVERGENCE OF THE OPTIMAL FILTER WITH DISTURBANCE DECOUPLING PROPERTY FOR LINEAR STOCHASTIC SYSTEMS

Akio Tanikawa

Faculty of Science and Technology
Yamato University
2-5-1, Katayama-cho, Suita-shi, Osaka 564-0082, Japan
tanikawa.a@yamato-u.ac.jp

Received August 2021; revised November 2021

Abstract. *For discrete-time linear stochastic systems with unknown disturbances, we consider the Riccati equation of the error covariance matrices of the optimal filter with disturbance decoupling property. In this paper, when the stochastic processes have constant coefficients, we prove the geometric rate of convergence of the error covariance matrices and the gain matrices of the optimal filter under some simple conditions similar to those for the Kalman filter. We also prove that the difference between the optimal state estimate and the approximated estimate which utilizes the constant gain matrix instead of the optimal gain matrix converges to zero at a geometric rate.*

Keywords: Stochastic systems, Optimal filter, Unknown inputs, Riccati equation

1. **Introduction.**

For discrete-time linear stochastic systems, we consider the Riccati equation of the error covariance matrices of the optimal filter. Since most of modeled systems are not very accurate representations of real systems and are not free from modeling errors, we often need to consider systems with unknown inputs. Thus, we consider the optimal filter with disturbance decoupling property for stochastic systems with unknown disturbances in this paper.

The problem of investigating optimal (or sub-optimal) filters for systems with noise and modeling uncertainty (including unknown disturbances and modelling errors) became to attract research attention in 1990's and then remarkable progress has been made in designing optimal filters (or observers) for stochastic systems with unknown disturbances. Darouach et al. [6, 7] 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. Optimal observers (or filters) with unknown input decoupling property were also studied by Chang and Hsu [2] and by Hou and Müller [9]. In 1996, Chen and Patton [3] proposed ODDO (Optimal Disturbance Decoupling 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 [11] also 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. This paper is closely related to the paper by Chen and Patton [3]. In our papers [24, 25], the recursive procedure ODDO was modified by the author and his colleagues as the optimal filter with disturbance decoupling property. That optimal filter was utilized to derive the optimal smoothers with disturbance decoupling property in [18, 19]. These unknown input

---

DOI: 10.24507/ijicic.18.01.305

decoupled filtering techniques were applied to some specific problems (see e.g., [1, 13]). In [20, 21], the optimal filter with disturbance decoupling property was modified so as to be applicable to nonlinear stochastic systems with unknown disturbances recently. In [22, 23], when the stochastic processes have constant coefficients, we proved convergence of the Riccati equation of the optimal filter with disturbance decoupling property and derived the algebraic Riccati equation (ARE) under some simple conditions similar to those for the Kalman filter.

In this paper, we consider the rate of convergence of the error covariance matrices of the optimal filter. The main motivation of discussing the geometric rate of convergence of these matrices is to show the reason of effectiveness of the optimal filter proposed in [24, 25]. Actually, we observed very quick adjustment of the optimal filter to the change of the system dynamics in the simulations in [24, 25]. The second motivation comes from our expectation of geometric convergence by observing that the limit P of the error covariance matrix is a stabilizing solution of ARE (see Theorem 3.2 in [22]). In Sections 2 and 3, we review preliminary results and give new formulas which play important roles in Section 4. Assuming that the stochastic processes have constant coefficients, we prove the geometric rate of convergence of the error covarience matrices and the gain matrices of the optimal filter in Section 4. We also prove that the difference between the optimal state estimate and the approximated estimate which utilizes the constant gain matrix instead of the optimal gain matrix converges to zero at a geometric rate. Finally, we conclude the paper and present future problem in Section 5.

2. **Preliminary Results (1).**

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

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

Let A_t, B_t, C_t and E_t be known matrices with appropriate dimensions. We suppose that ζ_t and η_t are independent zero mean white noise sequences with covariance matrices I (the identity matrix) and R_t respectively. 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., [12], Chapter 4)).

In [25], we considered the optimal estimate $\hat{x}_{t+1}$ of the state x_{t+1} with the following structure which was originally proposed by Chen and Patton ([3, 4]):

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

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

for $t = 0, 1, 2, \ldots$. Here, $\hat{x}_0$ is chosen to be z_0 for a fixed z_0. We use the notations $e_t := x_t - \hat{x}_t$ and $P_t := \mathbf{E}\left\{e_t {e_t}^T\right\}$ respectively for the state estimation error and its covariance matrix. Here, $\mathbf{E}$ denotes expectation and T denotes transposition of a matrix. We assume that random variables e_0, $\{\eta_t\}$, $\{\zeta_t\}$ are independent. As in [3, 4, 25], we consider state estimate of the form (3) and (4) where the matrices F_{t+1}, T_{t+1}, H_{t+1} and K_{t+1} have the forms respectively

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

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

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

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

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

The next lemma on Equation (6) was obtained and used by Chen and Patton [3, 4]. As in [3, 4], we adopt the standard condition that E_k is a full column rank matrix.

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

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

When this condition holds, the matrix H_{t+1} satisfying (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^1_{t+1} has the form

$$K^1_{t+1} = A^1_{t+1}\left(P_tC_t{}^T - H_tR_t\right)\left(C_tP_tC_t{}^T + R_t\right)^{-1}, \tag{13}$$

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

which is slightly different from the form in [3, 4], we obtained the following proposition in [24] on the optimal filtering algorithm under the next condition which is also assumed in this paper.

Condition 1. *The matrices C_tH_t and R_t are commutative, i.e.,*

$$C_tH_tR_t = R_tC_tH_t. \tag{15}$$

Proposition 2.2. *(Theorem 2.7 in* [24]*) The optimal gain matrix K^1_{t+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:*

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

$$P_{t+1} = A^1_{t+1}M_t{A^1_{t+1}}^T + T_{t+1}S_tS_t{}^TT_{t+1}{}^T + H_{t+1}R_{t+1}H_{t+1}{}^T, \tag{17}$$

where

$$G_t = \left(P_tC_t{}^T - H_tR_t\right)\left(C_tP_tC_t{}^T + R_t\right)^{-1}, \tag{18}$$

and

$$M_t = P_t - G_t\left(C_tP_t - R_tH_t{}^T\right). \tag{19}$$

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

Remark 2.3. *If the matrix R_t has the form*

$$R_t = r_tI$$

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

In [22], we obtained the following simplified form of the Riccati equation (17).

Proposition 2.4. *(Proposition 2.3 in* [22]*) For $t = 0, 1, 2, \ldots$, we have*

$$M_t = (I - G_t C_t) P_t (I - G_t C_t)^T + G_t R_t {G_t}^T, \tag{20}$$

and the Riccati equation

$$\begin{aligned} P_{t+1} &= A^1_{t+1} M_t {A^1_{t+1}}^T + T_{t+1} S_t {S_t}^T {T_{t+1}}^T + H_{t+1} R_{t+1} {H_{t+1}}^T \\ &= A^1_{t+1} (I - G_t C_t) P_t (I - G_t C_t)^T {A^1_{t+1}}^T + A^1_{t+1} G_t R_t {G_t}^T {A^1_{t+1}}^T \\ &\quad + T_{t+1} S_t {S_t}^T {T_{t+1}}^T + H_{t+1} R_{t+1} {H_{t+1}}^T. \end{aligned} \tag{21}$$

3. **Preliminary Results (2).** From now on, we consider the case where all coefficient matrices of the system (1) and (2) are independent of time. Thus, we consider the following discrete-time linear stochastic system with constant coefficients for $t = 0, 1, 2, \ldots$

$$x_{t+1} = Ax_t + Bu + Ed + S\zeta_t, \tag{22}$$

$$y_t = Cx_t + \eta_t. \tag{23}$$

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 (an invertible matrix) respectively. However, P_t, K_t, F_t, G_t and M_t still depend on t.

In this case, we proved convergence of the sequence of matrices $\{P_t\}$ to a matrix P as $t \to \infty$ in the previous paper [22] (see Proposition 3.1 and Proposition 3.2 in this section). The main purpose of this paper is to show the geometric rate of convergence of $\{P_t\}$ and $\{M_t\}$. The following proofs in this section and the next section were inspired by those for the Kalman filter by Chui and Chen [5].

In order to discuss convergence of $\{P_t\}$, we give two definitions. (A^1, S) is said to be *stabilizable* iff there is a matrix L such that $A^1 + SL$ is asymptotically stable (i.e., the absolute value of the eigenvalues of $A^1 + SL < 1$). (C, A^1) is said to be *detectable* iff there is a matrix L such that $A^1 + LC$ is asymptotically stable.

We now cite two propositions from [22].

Proposition 3.1. *(Theorem 3.1 in* [22]*) 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 P as $t \to \infty$ and P satisfies the equation*

$$\begin{aligned} P = {} & A^1 \left\{ P - \left(PC^T - HR\right) \left(CPC^T + R\right)^{-1} \left(CP - RH^T\right) \right\} \left(A^1\right)^T \\ & + TSS^T T^T + HRH^T \end{aligned} \tag{24}$$

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

$$G := \left(PC^T - HR\right) \left(CPC^T + R\right)^{-1}, \tag{25}$$

$$M := P - G\left(CP - RH^T\right), \tag{26}$$

we also have

$$G_t \longrightarrow G,\ M_t \longrightarrow M \ (as\ t \longrightarrow \infty),$$

where G_t and M_t are defined in Proposition 2.2.

We can show the following equation similarly to Proof of Lemma 6.4 in Chui and Chen [5], which plays an important role for proving the geomertic convergence of the sequence $\{P_t\}$.

Lemma 3.2. *Suppose that (C, A^1) is detectable and that $P_0 = O$. Then, P can be defined by $P = \lim_{t\to\infty} P_t$ as in Proposition 3.1. Moreover, we have*

$$P - P_{t+1} = A^1 \left(I - GC\right) \left(P - P_t\right) \left(I - G_t C\right)^T \left(A^1\right)^T . \tag{27}$$

Remark 3.3. *When we suppose that* (A^1, S) *is stabilizable and* (C, A^1) *is detectable, we obtain the same conclusions as in Lemma 3.2 for any initial matrix* $P_0 \geq 0$.

Lemma 3.4. *Let the matrix* Q *defined by*

$$Q = \begin{bmatrix} S & A^1 S & \left(A^1\right)^2 S & \ldots & \left(A^1\right)^{n-1} S \end{bmatrix} \tag{28}$$

have full rank (i.e., $\mathrm{rank}[Q] = n$*). Then, for any initial matrix* $P_0 \geq 0$*, we have* $P_t > 0$ *for* $t \geq n+1$*. Consequently, we have* $P > 0$*.*

Proof: Using (21) t times, we have

$$\begin{aligned} P_t = {} & \left[A^1 \left(I - G_{t-1}C\right) \;\; A^1 \left(I - G_{t-2}C\right) \;\; \ldots \;\; A^1 \left(I - G_0 C\right)\right] P_0 \left[A^1 \left(I - G_{t-1}C\right) \;\; A^1(I \right. \\ & \left. - \; G_{t-2}C) \;\; \ldots \;\; A^1 \left(I - G_0 C\right)\right]^T \\ & + \left[A^1 \left(I - G_{t-1}C\right) \;\; A^1 \left(I - G_{t-2}C\right) \;\; \ldots \;\; A^1 \left(I - G_1 C\right)\right] \left\{A^1 G_0 R {G_0}^T \left(A^1\right)^T \right. \\ & \left. + T S S^T T^T + H R H^T\right\} \left[A^1 \left(I - G_{t-1}C\right) \;\; A^1 \left(I - G_{t-2}C\right) \;\; \ldots \;\; A^1 \left(I - G_1 C\right)\right]^T \\ & + \cdots \\ & + \left[A^1 \left(I - G_{t-1}C\right)\right] \left\{A^1 G_{t-2} R {G_{t-2}}^T \left(A^1\right)^T + T S S^T T^T + H R H^T\right\} \left[A^1(I \right. \\ & \left. - \; G_{t-1}C)\right]^T + A^1 G_{t-1} R {G_{t-1}}^T \left(A^1\right)^T + T S S^T T^T + H R H^T. \end{aligned} \tag{29}$$

We now suppose $y^T P_t y = 0$. Then, from the last row of (29), we have

$$y^T T S S^T T^T y = 0 \text{ and so } y^T T S = 0, \tag{30}$$

$$y^T H R H^T y = 0 \text{ and so } y^T H = 0, \tag{31}$$

$$y^T A^1 G_{t-1} R {G_{t-1}}^T \left(A^1\right)^T y = 0 \text{ and so } y^T A^1 G_{t-1} = 0. \tag{32}$$

It then follows from (30) and (31) that

$$0 = y^T T S = y^T (I - HC) S = y^T S - y^T H C S = y^T S. \tag{33}$$

Namely, the equation

$$y^T S = 0 \tag{34}$$

holds. Similarly, from the second last row of (29), we have

$$y^T A^1 \left(I - G_{t-1}C\right) T S = 0, \tag{35}$$

$$y^T A^1 \left(I - G_{t-1}C\right) H = 0, \tag{36}$$

$$y^T A^1 \left(I - G_{t-1}C\right) A^1 G_{t-2} = 0. \tag{37}$$

It then follows from Equations (35) and (36) that

$$0 = y^T A^1 \left(I - G_{t-1}C\right) (I - HC) S = y^T A^1 \left(I - G_{t-1}C\right) S.$$

Substituting (32) into this, we have

$$y^T A^1 S = 0. \tag{38}$$

Repeating this procedure as in Proof of Lemma 6.6 in Chui and Chen [5], we have

$$y^T \left(A^1\right)^j S = 0 \quad (j = 0, 1, \ldots, t-1), \tag{39}$$

and so

$$y^T \left(A^1\right)^j S y = 0 \quad (j = 0, 1, \ldots, t-1). \tag{40}$$

Letting $t \geq n$, we obtain $y = 0$ since the matrix Q has full rank. Thus, we have $P_t > 0$ for $t \geq n+1$ and so $P > 0$. □

Remark 3.5. *When the matrix Q has full rank, the system (22) and (23) is said to satisfy reachability. If the system satisfies reachability, then it is stabilizable (see page 237 of* [12]*).*

Let the matrix B_t be defined by

$$B_t = \left[A^1 \left(I - G_{t-1}C\right)\right] \left[A^1 \left(I - G_{t-2}C\right)\right] \dots \left[A^1 \left(I - G_{n+1}C\right)\right]$$

for $t = n+2, n+3, \dots$ with $B_{n+1} := I$. Then, using (27) repeatedly, we have

$$P - P_t = \left[A^1 \left(I - GC\right)\right]^{t-n-1} \left(P - P_{n+1}\right) {B_t}^T \tag{41}$$

for $t = n+1, n+2, \dots$. We can easily prove boundedness of $\{B_t\}$ as follows.

Lemma 3.6. *Suppose that the system (22) and (23) satisfies reachability (i.e., rank$[Q] = n$) and detectability. Then, for any initial matrix $P_0 \geq O$, we have*

$$B_t {B_t}^T \leq U, \text{ for } t \geq n+1$$

for some constant matrix U. Consequently, if $B_t = \left({b_{ij}}^{(t)}\right)$, then

$$\left|{b_{ij}}^{(t)}\right| \leq \tilde{u}$$

for some constant $\tilde{u}$ (> 0) and for all i, j and t $(\geq n+1)$.

Proof: By Lemma 3.2 in [22], there exists a matrix $W > O$ such that $P_t \leq W$. Using (21) in Proposition 2.4 repeatedly, we have

$$\begin{aligned} W \geq P_t &\geq \left[A^1 \left(I - G_{t-1}C\right)\right] P_{t-1} \left[A^1 \left(I - G_{t-1}C\right)\right]^T \\ &\geq \left[A^1 \left(I - G_{t-1}C\right)\right] \left[A^1 \left(I - G_{t-2}C\right)\right] P_{t-2} \left[A^1 \left(I - G_{t-2}C\right)\right]^T \left[A^1 \left(I - G_{t-1}C\right)\right]^T \\ &\geq \cdots \\ &\geq B_t P_{n+1} {B_t}^T. \end{aligned}$$

Let $\lambda_{\min}$ be the smallest eigenvalue of P_{n+1}. Noting that $P_{n+1} \geq \lambda_{\min} I$, we have

$$W \geq B_t \lambda_{\min} I {B_t}^T = \lambda_{\min} B_t {B_t}^T.$$

Thus, we have proved the lemma. □

4. **Geometric Convergence.** In this section, we will prove the geometric rate of convergence of the error covariance matrices $\{P_t\}$ and the gain matrices $\{G_t\}$ of the optimal filter with disturbance decoupling property under some simple conditions similar to those for the Kalman filter. We will also prove that the difference between the optimal state estimate and the approximated estimate which utilizes the constant gain matrix instead of the optimal gain matrix converges to zero at a geometric rate. These results explain the reason why we observe excellent simulation results of the optimal filter (16) [22].

Theorem 4.1. *Suppose that (A^1, S) is reachable 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 (24))) as $t \to \infty$, where $P > 0$. Moreover, the rate of convergence is geometric, i.e.,*

$$\operatorname{tr}\left[(P_t - P)(P_t - P)^T\right] \leq C r^t, \tag{42}$$

for some r with $0 < r < 1$ and $C > 0$. Consequently,

$$\operatorname{tr}\left[(G_t - G)(G_t - G)^T\right] \leq C' r^t, \tag{43}$$

for some $C' > 0$.

Proof: Let F be defined by $A^1(I-GC)$. Using (41), Lemma 3.6 and boundedness of $\{P_t\}$, we have

$$\begin{aligned}(P_t-P)(P_t-P)^T &= F^{t-n-1}(P_{n+1}-P)B_t{}^T B_t(P_{n+1}-P)\left(F^{t-n-1}\right)^T\\ &\le F^{t-n-1}\Omega\left(F^{t-n-1}\right)^T\end{aligned}$$

for some non-negative definite symmetric constant matrix Ω. Hence, by applying the simple inequality

$$\operatorname{tr}\left[(MN)(MN)^T\right] \le \operatorname{tr}\left[MM^T\right]\operatorname{tr}\left[NN^T\right] \tag{44}$$

for matrices M and N (see Lemma 1.7 in [5]), we obtain

$$\begin{aligned}\operatorname{tr}\left[(P_t-P)(P_t-P)^T\right] &\le \operatorname{tr}\left[F^{t-n-1}\Omega^{1/2}\left(F^{t-n-1}\Omega^{1/2}\right)^T\right]\\ &\le \operatorname{tr}\left[F^{t-n-1}\left(F^{t-n-1}\right)^T\right]\operatorname{tr}\left[\Omega^{1/2}\left(\Omega^{1/2}\right)^T\right]\\ &\le Cr^t \end{aligned}\tag{45}$$

for some r with $0<r<1$ and $C>0$, where we applied Lemma 1.10 in [5] to the last inequality above since the absolute value of eigenvalues of F is smaller than 1 (see Theorem 3.2 in [22]).

For proving (43), we write

$$\begin{aligned}(G_t-G) &= \left(P_tC^T-HR\right)\left(CP_tC^T+R\right)^{-1}-\left(PC^T-HR\right)\left(CPC^T+R\right)^{-1}\\ &= (P_t-P)C^T\left(CP_tC^T+R\right)^{-1}+\left(PC^T-HR\right)\left(CP_tC^T+R\right)^{-1}C(P\\ &\quad -P_t)C^T\left(CPC^T+R\right)^{-1}.\end{aligned}$$

Using the inequality

$$(U+V)(U+V)^T \le 2\left(UU^T+VV^T\right)$$

for matrices U and V (see Exercise 6.5 in [5]), we obtain

$$\begin{aligned}(G_t-G)(G_t-G)^T &\le 2(P_t-P)C^T\left(CP_tC^T+R\right)^{-1}\left(CP_tC^T+R\right)^{-1}C(P_t-P)\\ &\quad +2\left(PC^T-HR\right)\left(CP_tC^T+R\right)^{-1}C(P-P_t)C^T\left(CPC^T\right.\\ &\quad \left.+R\right)^{-1}\left(CPC^T+R\right)^{-1}C(P-P_t)C^T\left(CP_tC^T\right.\\ &\quad \left.+R\right)^{-1}\left(CP-RH^T\right).\end{aligned}$$

Moreover, due to $P_0 \le P_t$ (see Lemma 3.1 in [22]), we have $\left(CP_0C^T+R\right) \le \left(CP_tC^T+R\right)$ and so

$$\left(CP_tC^T+R\right)^{-1} \le \left(CP_0C^T+R\right)^{-1}.$$

Hence, we obtain

$$\operatorname{tr}\left[\left(CP_tC^T+R\right)^{-1}\left(CP_tC^T+R\right)^{-1}\right] \le \left(\operatorname{tr}\left[\left(CP_0C^T+R\right)^{-1}\right]\right)^2$$

(see Lemma 1.9 in [5]). Then, by (44), we have

$$\begin{aligned}\operatorname{tr}\left[(G_t-G)(G_t-G)^T\right] &\le 2\operatorname{tr}\left[(P_t-P)(P_t-P)^T\right]\operatorname{tr}\left[CC^T\right]\left(\operatorname{tr}\left[\left(CP_0C^T+R\right)^{-1}\right]\right)^2\\ &\quad +2\operatorname{tr}\left[\left(PC^T-HR\right)\left(PC^T-HR\right)^T\right]\left(\operatorname{tr}\left[\left(CP_0C^T\right.\right.\right.\\ &\quad \left.\left.\left.+R\right)^{-1}\right]\right)^2\operatorname{tr}\left[CC^T\right]\operatorname{tr}\left[(P_t-P)(P_t\right.\\ &\quad \left.-P)^T\right]\operatorname{tr}\left[CC^T\right]\operatorname{tr}\left[\left(CPC^T+R\right)^{-1}\left(CPC^T+R\right)^{-1}\right]\end{aligned}$$

$$\begin{aligned}&\le C_1 \mathrm{tr}\left[(P_t - P)(P_t - P)^T\right] \\ &\le C_2 r^t,\end{aligned}$$

where C_1 (> 0) and C_2 (> 0) are constants independent of t and we used (45) at the last inequality. □

We now discuss the last problem: How much will be the error of the state estimate from the optimal estimate $\hat{x}_t$ when we use the gain matrix G instead of G_t? Thus, noting that $\{\hat{x}_t\}$ is defined iteratively by (16) with the optimal gain matrix, we define the modified one $\{\vec{x}_t\}$ by

$$\vec{x}_{t+1} = A^1 \left\{\vec{x}_t + G\left(y_t - C\vec{x}_t\right)\right\} + Hy_{t+1} + TBu, \tag{46}$$

where G_t in (16) was replaced by G. Then, in view of (46), we can easily obtain the following simple iterative formula on $\vec{x}_t - x_t$.

Lemma 4.2. *For $t = 0, 1, 2, \ldots$, we have*

$$\vec{x}_{t+1} - x_{t+1} = A^1(I - GC)\left(\vec{x}_t - x_t\right) + A^1 G\eta_t - TS\zeta_t + H\eta_{t+1}. \tag{47}$$

Now, we can prove that the difference between the optimal state estimate $\hat{x}_t$ and the approximated estimate $\vec{x}_t$ tends to zero at a geometric rate.

Theorem 4.3. *Suppose that (A^1, S) is reachable and (C, A^1) is detectable. Then, there exist a real number r $(0 < r < 1)$ and a positive constant C such that*

$$\mathrm{tr}\left[\mathbf{E}\left\{\epsilon_t {\epsilon_t}^T\right\}\right] \le Cr^t, \tag{48}$$

where $\epsilon_t := \hat{x}_t - \vec{x}_t$.

Proof: Denote $\epsilon_t := \hat{x}_t - \vec{x}_t$ and $\delta_t := x_t - \hat{x}_t$. Noting that

$$\begin{aligned}\hat{x}_t &= A^1\hat{x}_{t-1} + A^1 G_{t-1}\left(y_{t-1} - C\hat{x}_{t-1}\right) + Hy_t + TBu \\ &= A^1\hat{x}_{t-1} + A^1 G\left(y_{t-1} - C\hat{x}_{t-1}\right) + Hy_t + TBu + A^1\left(G_{t-1} - G\right)\left(y_{t-1} - C\hat{x}_{t-1}\right)\end{aligned}$$

and

$$\vec{x}_t = A^1\vec{x}_{t-1} + A^1 G\left(y_{t-1} - C\vec{x}_{t-1}\right) + Hy_t + TBu,$$

we have

$$\begin{aligned}\epsilon_t &= A^1\left(\hat{x}_{t-1} - \vec{x}_{t-1}\right) - A^1 GC\left(\hat{x}_{t-1} - \vec{x}_{t-1}\right) + A^1\left(G_{t-1} - G\right)\left(Cx_{t-1} + \eta_{t-1} - C\hat{x}_{t-1}\right) \\ &= A^1\left(I - GC\right)\epsilon_{t-1} + A^1\left(G_{t-1} - G\right)\left(C\delta_{t-1} + \eta_{t-1}\right) \\ &= A^1\left(I - GC\right)\epsilon_{t-1} + A^1\left(G_{t-1} - G\right)C\delta_{t-1} + A^1\left(G_{t-1} - G\right)\eta_{t-1}.\end{aligned}$$

Then, similarly to Proof of Theorem 6.3 in Chui and Chen [5], we have

$$\mathbf{E}\left\{\epsilon_t {\epsilon_t}^T\right\} \le C_4 {r_1}^t,$$

for some positive constant C_4. □

Remark 4.4. *Reachability of (A^1, S) is equivalent to the condition that*

$$\mathrm{rank}\begin{bmatrix} A^1 - \lambda I & S \end{bmatrix} = n \tag{49}$$

holds for any $\lambda \in \mathbf{C}$. Detectability of (C, A^1) is equivalent to the condition that

$$\mathrm{rank}\begin{bmatrix} A^1 - \lambda I \\ C \end{bmatrix} = n \tag{50}$$

holds for any $\lambda \in \mathbf{C}$ with $|\lambda| \ge 1$ (see [12] *for the details of those equivalences). We can easily see that the system (22) and (23) with*

$$\mathrm{rank}[S] = \mathrm{rank}[C] = n \tag{51}$$

satisfies reachability and detectability.

As a practical example, Chen and Patton considered the linearized discrete-time model of a simplified longitudinal flight control system (22) and (23) where the state variables $x = [\eta_y \ \omega_z \ \delta_z]^T$ are pitch angle δ_z, pitch rate ω_z and normal velocity η_y, the control input is elevator control signal, and the actual system parameter matrices are given in [3, 4]. This flight control system satisfies the condition (51). They showed the simulation results which indicate the effectiveness of their method ODDO over the standard Kalman filter. We also had the simulation results which compared our method in Proposition 2.2 and the Kalman filter and showed the effectiveness of our method over the Kalman filter in [22]. This is one of the motivations for considering the rate of convergence of the error covariance matrices in this paper.

5. Conclusion.

We consider discrete-time linear stochastic systems with unknown inputs (or disturbances) and investigate the Riccati equation of the error covariance matrices of the optimal filter with disturbance decoupling property. In this paper, when the stochastic processes have constant coefficients, we proved the geometric rate of convergence of the error covariance matrices and the gain matrices of the optimal filter under some simple conditions similar to those for the Kalman filter. We also proved that the difference between the optimal state estimate and the approximated estimate which utilizes the constant gain matrix G instead of the optimal gain matrix G_t converges to zero at a geometric rate. Finally, the assumptions of the theorems in Section 4 are slightly strong, and weakening them will be the future problem.

Acknowledgments. The author wishes to express gratitude to the associate editor and the anonymous referee for valuable comments which have improved the earlier version of the paper.

REFERENCES

[1] 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, CO, pp.5381-5387, 2003.

[2] 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.

[3] J. Chen and R. J. Patton, Optimal filtering and robust fault diagnosis of stochastic systems with unknown disturbances, *IEEE Proc. of Control Theory Applications*, vol.143, no.1, pp.31-36, 1996.

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

[5] C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, 3rd Edition, Springer-Verlag, Berlinl Heidelberg, Germany, 1999.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] T. Katayama, *Applied Kalman Filtering*, 1st Edition, Asakura-Shoten, Tokyo, Japan, 2000 (in Japanese).

[13] 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.
[14] B. P. Molinari, The stabilizing solution of the discrete algebraic Riccati equation, *IEEE Trans. Automatic Control*, vol.20, no.3, pp.396-399, 1975.
[15] R. J. Patton, P. M. Frank and R. N. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Application*, Prentice Hall, 1996.
[16] 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.
[17] 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.
[18] 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.
[19] 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.
[20] A. Tanikawa, Numerical simulation of a new class of filtering algorithms for nonlinear stochastic systems with unknown disturbances, *Memoirs of Osaka Institute of Technology*, vol.62, no.2, pp.1-12, 2017.
[21] 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, 2018.
[22] A. Tanikawa, Analysis of the Riccati equation of the optimal filter with disturbance decoupling property for linear stochastic systems, *International Journal of Innovative Computing, Information and Control*, vol.15, no.3, pp.811-824, 2019.
[23] A. Tanikawa, On the Riccati equation of the optimal filter with disturbance decoupling property for linear stochastic systems, *Proc. of the 50th ISCIE Int. Symp. on Stochastic Systems Theory and Its Appl.*, Kyoto, Japan, pp.7-12, 2019.
[24] A. Tanikawa and H. Mukai, *Minimum Variance State Estimators with Disturbance Decoupling Property for Optimal Filtering Problems with Unknown Inputs*, Unpublished Lecture Note, 2006.
[25] 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.
[26] Y. Xie, J. Liu and L. Wang, H^{∞} filtering design for Takagi-Sugeno fuzzy model with immeasurable premise variable by applying a switching method, *ICIC Express Letters*, vol.14, no.3, pp.257-264, 2020.

Author Biography

Akio Tanikawa received BS and MS in Mathematics, Nagoya University, Nagoya, Japan, in 1977 and 1979 respectively, and D.Sc. in Systems Science and Mathematics, Washington University, St. Louis, USA, in 1982. He was with Toyota College of Technology, Toyota, Japan from 1983 to 1988, with Kanazawa University, Kanazawa, Japan from 1988 to 1996, with Washington University, St. Louis, USA from 1996 to 2001, with Oita University, Oita, Japan from 2001 to 2003, and with Osaka Institute of Technology, Osaka, Japan from 2003 to 2019.

Since 2020 he has been a professor at Yamato University, Osaka, Japan. His current research interests are in stochastic systems, system identification, optimal control, and game theory.