Conditions of Existence of the Solution of the Observation Optimization for the Kalman-Bucy Filter

Yoshiki TAKEUCHI
Dept. of Information Science, Osaka Kyoiku University
4-698-1 Asahigaoka, Kashiwara City, Osaka 582-8582, Japan
Tel. +81-729-78-3669, Fax. +81-729-78-3554
e-mail: takeuti@cc.osaka-kyoiku.ac.jp

Abstract

In this paper, we are concerned with a problem of optimization of the linear observations which are used in the stationary Kalman-Bucy filter. Especially, we consider the optimization of the gain matrix in the observation. In the previous works the author, by introducing a change of variables using eigenvalues and eigenvectors of a symmetric matrix, the condition of optimality of these variables was obtained. In this paper, it is shown that the expression of the optimal gain obtained in the previous works is unique, and there does not exist any other gain matrix which produces the same value of the cost function. It will be also shown in this paper that under what condition the optimal gain fails to exist and that how one can avoid the situation by modifying the weight matrix in the cost function.

1 Introduction

This paper is concerned with optimization of the gain matrix in the linear observation for the stationary Kalman-Bucy filter, under a cost function which is quadratic in both the estimation error and the observation gain. This kind of problem was extensively discussed in [1]-[5], [8], [10]-[18]. In the previous works of the author, [7]-[17] are concerned with the discrete-time systems, and the corresponding results for the continuous-time systems are reported in [18].

The summary of the result obtained in [18] is as follows. The coefficient of the quadratic term in the matrix Riccati equation, as you know, is a symmetric matrix which is a quadratic function of the gain matrix weighted by the inverse of the noise covariance matrix. In what follows, this matrix will be referred to as the symmetric signal-to-noise ratio matrix (SSNRM).

(i) By introducing a change of variables by the eigenvalues and eigenvectors of SSNRM, the problem becomes the optimization with respect to these variables and an additional matrix with orthonormal column vectors.

(ii) The optimal value of the additional matrix is simply given by a set of eigenvectors of the noise covariance matrix corresponding to the ascending order of the eigenvalues.

(iii) For the pair of the eigenvalues and eigenvectors of SSNRM, the condition of optimality was obtained. It is such that the eigenvectors of SSNRM are that of a symmetric matrix which is given by a solution of Lyapunov equation, and that the eigenvalues of this matrix coincide with that of the noise covariance matrix.

In this paper, it will be shown first that there is no other optimal gain matrix with a different expression than the one obtained in [18]. At the time when the author reported [18], this optimal gain was thought to be one of the expressions which produce the same value of the cost function. But as will be seen, the set of observation gains with the same value of SSNRM is completely described by the additional matrix in the above (ii).

In numerical studies of the computation of the optimal solution, the author sometimes encountered the cases where the optimal observation gain with the assigned rank does not exist. It will be also shown in this paper that under what condition this situation arises and that how one can avoid the situation by modifying the weight matrix in the cost function.

Mathematical symbols, in this paper, are used in the following way. \( R \) is the space of all real numbers, i.e., \( R \equiv (-\infty, \infty) \). For positive integers \( m \) and \( n \), \( R^m \) and \( R^{nxn} \) denote the spaces of \( n \)-dimensional vectors and \( m \times n \)-dimensional matrices whose components take values in \( R \). The prime denotes the transpose of a vector or a matrix and the Euclidean norm is \( | \cdot | \). Thus, for \( x \in R^m \), \( |x| = \sqrt{x'x} \). The identity matrix of any dimension is denoted by \( I \). The components of a matrix are denoted by using subscripts. Thus, \([A]_y\) is the \((i, j)\)-component of \( A \). In the case where no confusion may arise, we denote \([A]_y\) simply by \( a_y \). If \( A \) is a square matrix, \( \det[A] \) and \( \text{tr}[A] \) respectively denote the determinant and the trace of \( A \). We use \( A > 0 \) and \( A \geq 0 \) to denote that \( A \) is positive definite and nonnegative definite, respectively. The triplet \((\Omega, \mathcal{F}, P)\) is a complete probability space where \( \Omega \) is a sample space with elementary events \( \omega \), \( \mathcal{F} \) is a \( \sigma \)-field of subsets of \( \Omega \), and \( P \) is a probability measure. \( E[\cdot] \) denotes the expectation and \( E[\cdot|\cdot] \), \( \cdot \subset \mathcal{F} \) the conditional expectation, given \( \cdot \), with respect to \( P \). \( \sigma[\cdot] \) is the minimal sub-\( \sigma \)-field of \( \mathcal{F} \) with respect to which the family of \( \mathcal{F} \)-measurable sets or random variables \{\cdot\} is measurable.
2 Problem Formulation

Let \( x = \{x_t(\omega) ; 0 \leq t < \infty\} \) be an \( n \)-dimensional Gaussian stochastic process generated by the linear system:

\[
\begin{align*}
    dx_t(\omega) &= Ax_t(\omega)dt + Gdw_t(\omega) \\
    x_0(\omega) &= x^0(\omega),
\end{align*}
\]

where \( A \in \mathbb{R}^{n \times n} \), \( G \in \mathbb{R}^{n \times d} \), \( x^0(\omega) \) is a Gaussian random vector with mean \( \bar{x} \) and covariance \( \Omega^0 \), and \( w = \{w_t(\omega) ; 0 \leq t < \infty\} \) is a \( d \)-dimensional standard Brownian motion process. Suppose that the value of \( x \) is not directly available but we have linear observations described by

\[
dy_t(\omega) = H x_t(\omega)dt + R dv_t(\omega),
\]

where \( y = \{y_t(\omega) ; 0 \leq t < \infty\} \) is an \( m \)-dimensional observation process, \( H \in \mathbb{R}^{m \times n} \), \( R \in \mathbb{R}^{m \times d} \), and \( v = \{v_t(\omega) ; 0 \leq t < \infty\} \) is a \( d \)-dimensional standard Brownian motion process. Throughout this paper, we will assume that the following two conditions are satisfied.

\[
\begin{align*}
    (C-1) & \quad G^2 > 0 \quad \text{and} \quad R \triangleq RR' > 0. \quad (C-2) & \quad x^0(\omega), w \text{ and } v \text{ are mutually independent.}
\end{align*}
\]

In this paper, we are concerned with the stationary Kalman-Bucy filter. Namely, under the assumption that \( x \) is a wide-sense stationary process, the least-squares estimate is given by the stationary Kalman-Bucy filter:

\[
\begin{align*}
    \hat{x}_t(\omega) &= A\hat{x}_t(\omega)dt + QH'R_0^{-1}\{dy_t(\omega) - H\hat{x}_t(\omega)dt\} \\
    \hat{\sigma}_t(\omega) &= E[x^0(\omega)],
\end{align*}
\]

and

\[
AQ + QA' + GG' - QH'R_0^{-1}HQ = 0,
\]

due to the stationarity of \( x \).

For the optimization of performance of the least-squares state estimator (3) and (4), we introduce the following quadratic performance criterion: \([19]\)

\[
J \triangleq \text{tr}[MQ] + \text{tr}[HNH'],
\]

where \( \text{tr}[MQ] \) is an weighted estimation error variance, i.e.,

\[
\text{tr}[MQ] = E\{[x_1(\omega) - \hat{x}_1(\omega)][x_1(\omega) - \hat{x}_1(\omega)]'\},
\]

and \( M \in \mathbb{R}^{m \times m} \) and \( N \in \mathbb{R}^{m \times m} \) are positive-definite symmetric matrices. In (6), the second term in the right-hand side denotes the cost or the energy consumed by the observation.

Thus, we are now concerned with the following problem.

[Problem 1] Find \( H \in \mathbb{R}^{m \times n} \) such that (6) is minimized subject to (4).

For Problem 1, it was shown in Takeuchi [18; Theorem 1] that this problem is reduced to the one of the simple case where \( N = I \). Also, this is done by a simple transformation of \( x, A, G, M \) and \( H \) using the square root matrix of \( N \). For this reason, we are, from now on, only concerned with the case where \( N = I \) and \( J \) is given by

\[
J = \text{tr}[MQ] + \text{tr}[HH'].
\]

Namely, we are concerned with

\[
[\text{Problem 1}'] \text{ Find } H \in \mathbb{R}^{m \times n} \text{ such that (8) is minimized subject to (4).}
\]

As was shown in Takeuchi [18], it is possible to obtain the condition of optimality by introducing a parameter transformation using the symmetric matrix \( H'R_0^{-1}H \).

Note that this symmetric matrix is the one which may be called a symmetric signal-to-noise ratio matrix (SSNRM). Also, it was shown in [18] that among the values of \( H \in \mathbb{R}^{m \times n} \) for which we have the same value of \( Q \), there exists a set of values that can be simply represented by using the positive eigenvalues and the corresponding eigenvectors of SSNRM. Theorem 2 below is a new result of this paper that all \( H \in \mathbb{R}^{m \times n} \) can be represented in the above mentioned form by using SSNRM. Then, we can easily obtain the optimal value of \( H \) which is unique in the space \( \mathbb{R}^{m \times n} \).

Let

\[
n \triangleq \text{rank}[H] \leq m.
\]

Then, since \( H'R_0^{-1}H \) is a nonnegative-definite symmetric matrix, we can represent this matrix as

\[
H'R_0^{-1}H = \hat{U}\hat{\Xi}\hat{U}^{'},
\]

where \( \hat{U} = [u_1, u_2, \ldots, u_m] \in \mathbb{R}^{m \times m} \) is the set of eigenvectors of the SSNRM in the left-hand side of (10) corresponding to the positive eigenvalues \( \xi_i, i = 1, 2, \ldots, m \), and

\[
\hat{\Xi} = \text{diag}(\xi_1, \xi_2, \ldots, \xi_m), \quad \xi_i > 0, i = 1, 2, \ldots, m.
\]

Without loss of generality, we can assume that

\[
\xi_1 \geq \xi_2 \geq \cdots \geq \xi_m > 0.
\]

For any \( \hat{U} \in \mathbb{R}^{m \times n} \) with property \( \hat{U}'\hat{U} = I \) and for any \( \hat{\Xi} \) with properties (11) and (12), let

\[
\mathcal{H}(\hat{U}, \hat{\Xi}) \triangleq \{H \in \mathbb{R}^{m \times n} ; H'R_0^{-1}H = \hat{U}\hat{\Xi}\hat{U}'\}.
\]

Thus, \( \mathcal{H}(\hat{U}, \hat{\Xi}) \) is the set of values of \( H \) for which \( H'R_0^{-1}H \) has fixed positive eigenvalues \( \xi_i \) and the corresponding eigenvectors \( u_i, i = 1, 2, \ldots, m \).

We have the following theorems which guarantee that we can always take \( H \) in the form (14) given below.

\[\text{Theorem 1}[13]-[18] \text{ Assume (C-1) and (C-2). Then, we have the following properties.}
\]

(i) \( Q \), the solution of (4), is completely determined by \( (A, GG', \hat{U}, \hat{\Xi}) \).

(ii) For a fixed value of \( (\hat{U}, \hat{\Xi}) \) and for any \( H \in \mathcal{H}(\hat{U}, \hat{\Xi}) \), we have the same value of \( Q \).
\textbf{Theorem 2 (A Representation of $H$)} Assume (C-1) and (C-2). Then, for any fixed set of $(\tilde{U}, \tilde{E})$ with the property $\tilde{U}^T\tilde{U} = I$, any $H \in \mathcal{H}(\tilde{U}, \tilde{E})$ is represented by

$$H = R_0^{1/2} \tilde{E}^{1/2} \tilde{U},$$

where $\tilde{U} \in \mathbb{R}^{m \times m}$ is a matrix whose all column vectors form an orthonormal system in $\mathbb{R}^m$, i.e., $\tilde{U}^T = I$.

Namely, (14) is a necessary and sufficient condition for $H \in \mathcal{H}(\tilde{U}, \tilde{E})$. In other words, the degree of freedom in the space $\mathcal{H}(\tilde{U}, \tilde{E})$ is exactly the same as the one given by $\tilde{U} \in \mathbb{R}^{m \times m}$ with $\tilde{U}^T = I$.

(Proof) It is clear that (14) is sufficient for $H \in \mathcal{H}(\tilde{U}, \tilde{E})$. Hence, we will show that (14) is necessary for $H \in \mathcal{H}(\tilde{U}, \tilde{E})$. First, let $H \in \mathcal{H}(\tilde{U}, \tilde{E})$ and define $\tilde{H} = H\tilde{U}^{1/2}$. Hence, it follows that $\tilde{H} = H\tilde{U}^{1/2}$ with $\tilde{H}^T = I$.

From (16), it is seen that by defining $\tilde{H} = H\tilde{U}^{1/2}$, we have $\tilde{H} = H\tilde{U}^{1/2}$. Thus, from the second equality of (18), we have $\tilde{H}^T = H\tilde{U}^{1/2}$.

For $\tilde{H} \in \mathbb{R}^{m \times m}$, let $\tilde{H} \in \mathbb{R}^{m \times m}$ be a matrix such that $U \triangleq \{U, \tilde{U}\} \in \mathbb{R}^{m \times m}$ becomes an orthogonal matrix, i.e., $U^T \tilde{U}^T = I$. Since $\tilde{U}^T \tilde{U}^T = 0$, it follows that $U^T \tilde{U}^T = 0$. Thus, all rows of $H$ are orthogonal to all the column vectors of $\tilde{U}$. Hence, any row of $H$ is a linear combination of the column vectors of $\tilde{U}$, namely, we can write

$$H = K \tilde{U}, \quad K \in \mathbb{R}^{m \times m}.$$ (21)

From (21) and (19), we have

$$H = K \tilde{U} = R_0^{1/2} \tilde{E}^{1/2} \tilde{U}.$$ (22)

Then, substituting (22) into (21), we have (14). This completes the proof.

Thus, any $H$ in $\mathcal{H}(\tilde{U}, \tilde{E})$ can be represented in the form (14). So, the problem has been converted to the one of the optimization with respect to $\tilde{U} \in \mathbb{R}^{m \times m}$, $\tilde{E} \in \mathbb{R}^{m \times m}$ and $\tilde{E} \triangleq \text{diag}(\xi_1, \xi_2, \ldots, \xi_m)$.

3 The Condition of Optimality

3.1 The Optimization of $\hat{\Gamma}$ over $\mathcal{H}(\tilde{U}, \tilde{E})$

By Theorem 2, if we find the optimal value of $\hat{\Gamma}$, then we can determine the optimal observation gain matrix $H$ for a fixed value of $(\tilde{U}, \tilde{E})$. As we see from (4) and (10), the value of $Q$ is independent of that of $\hat{\Gamma} \in \mathbb{R}^{m \times m}$. Hence, the optimal value of $\hat{\Gamma}$ should be determined in such a way that $\text{tr}(HH^T)$, the second term in (8), is minimized. This minimization is done by the exactly same way as the discrete-time case.\(^{[13]}\)

\textbf{Theorem 3\(^{[18], [13]}\)} Assume (C-1)-(C-2). Then the optimal value of $\hat{\Gamma} \in \mathbb{R}^{m \times m}$ is given by the set of eigenvectors of $R_0$ corresponding to the first $\tilde{m}$ eigenvalues in ascending order, i.e., we have the relation

$$\hat{\Gamma} = \tilde{K},$$ (23)

where $\tilde{K} \in \mathbb{R}^{m \times m}$ is given by

$$R_0 = \left[\tilde{K} \tilde{K}^T\right] \Psi \left[\begin{array}{c} \tilde{K} \tilde{K}^T \end{array}\right], \quad \Psi = \text{diag}(\psi_1, \psi_2, \ldots, \psi_m, \ldots, \psi_m),$$

$$\psi_1 \leq \psi_2 \leq \cdots \leq \psi_m \leq \cdots \leq \psi_m.$$ (24)

\textbf{Remark 1} It has been shown by Theorem 3, that the optimal value of $\hat{\Gamma}$ does not depend on $(\tilde{U}, \tilde{E})$ but it is always given by the first $\tilde{m}$ eigenvectors of $R_0$.

\textbf{Remark 2} The following (i)-(iv) should be noted.

(i) If we take $\hat{\Gamma} \in \mathbb{R}^{m \times m}$ as (23), then we have $H = \hat{\Gamma} \tilde{E}^{1/2} \tilde{U}^T$,

$$H = \hat{\Gamma} \tilde{E}^{1/2} \tilde{U}^T,$$ (25)

where $\tilde{E} \triangleq \text{diag}(\psi_1, \psi_2, \ldots, \psi_m)$.

(ii) For any value of $(\tilde{U}, \tilde{E})$ which satisfies (11), (12) and $U^T \tilde{U}^T = I$,

$$\tilde{U}^T \hat{\Gamma} \tilde{U} = 1,$$ (27)

$H$ is optimal when $H_0(\omega) \in \mathcal{H}(\tilde{U}, \tilde{E})$. As we see from (4) and (10), the value of $Q$ is independent of that of $\hat{\Gamma} \in \mathbb{R}^{m \times m}$. Hence, $H_0(\omega)$ should be determined in such a way that $\text{tr}(HH^T)$, the second term in (8), is minimized. This minimization is done by the exactly same way as the discrete-time case.\(^{[13]}\)

(iii) Let

$$H_{\tilde{E}} = \tilde{K} \tilde{E}^{1/2} \tilde{U}^T,$$ (28)

and

$$H_{\text{SNR}} = \tilde{K} \tilde{E}^{1/2} \tilde{U}^T,$$ (29)

Then, we have the quadratic relations: $\tilde{R}_0 \triangleq \tilde{K} \tilde{E}^{1/2} \tilde{H}_0$ and $H \tilde{R}_0^{-1} H = H_{\text{SNR}} H_{\text{SNR}}^T$, and (25) implies

$$H = H_{\tilde{E}} H_{\text{SNR}}^T.$$ (30)

Namely, $H$ is given by a product of the square root components of the noise covariance and SSNRM.
(iv) Both (30) and (14) can be regarded as matrix generalizations of the trivial relation for the scalar case:

$$H = \sqrt{R_0} \times \sqrt{R_0^*}.$$  \hspace{1cm} (31)

It should be noted, however, that (30) is true only when $\Gamma$ in (14) is selected as $\Gamma = \Gamma$ and that $H_{b_0}$ is a square root matrix of not $R_0$ but $\tilde{R}_0$, the part of $R_0$ composed of the $\tilde{m}$-set of the eigenvalues and eigenvectors.  

3.2 The Condition Optimality of $U$, $\tilde{Z}$

From (25), it is seen that the rest part of Problem 1' can be written in the following form.

**Problem 2** Find $\tilde{U} \in R^{m \times n}$ and $\tilde{Z} = \text{diag}(\xi_1, \xi_2, \ldots, \xi_m)$ such that

$$J = \text{tr}[MQ] + \text{tr}[\Psi\tilde{Z}] \rightarrow \text{min.},$$  \hspace{1cm} (32)

subject to (4), (10), (11), (12) and

$$\tilde{U}'\tilde{U} = I.$$  \hspace{1cm} (33)

For Problem 2, let us define the Lagrangean by

$$L(\tilde{Z}, \tilde{U}, \lambda) \triangleq \text{tr}[MQ] + \text{tr}[\Psi\tilde{Z}] + \text{tr}[\Lambda(\tilde{U}'\tilde{U} - I)],$$  \hspace{1cm} (34)

where $\Lambda \in R^{m \times n}$ is a symmetric matrix whose $(i, j)$-component is a Lagrange multiplier for the same component of (33), i.e.,

$$\text{tr}[\Lambda(\tilde{U}'\tilde{U} - I)] = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} [(\tilde{U}'\tilde{U} - I)]_{ij},$$  \hspace{1cm} (35)

where

For Problem 2, we have the following result.

**Theorem 4** \[21\] Assume (C-1) and (C-2). If $(\tilde{U}', A)$ is detectable and $\tilde{Z}$ satisfies (12), then we have a unique positive-definite solution $Q$ of (4) with (10), and $A = Q\tilde{U}'\tilde{U}'$ is an asymptotically stable matrix.  

**Theorem 5 (Condition of Optimality)** \[18\] Let us define $F \in R^{m \times n}$ by

$$F \triangleq A - Q\tilde{U}\tilde{Z}\tilde{U}'. $$  \hspace{1cm} (36)

Assume (C-1), (C-2) and (C-3).

Then, the condition of optimality of $\tilde{U}$ and $\tilde{Z}$ is given by

$$QXQ\tilde{U} = \tilde{U}\tilde{Z},$$  \hspace{1cm} (37)

where $X \in R^{m \times n}$ is a solution of

$$F'X + XF + M = 0.$$  \hspace{1cm} (38)

Since $\tilde{Z}$ is a diagonal matrix given by (26), (37) implies that

«**Corollary 1**» Assume (C-1)-(C-3). The optimal $(\tilde{U}, \tilde{Z})$ is such that

(i) Each column vector of $\tilde{U}$ is an eigenvector of $QXQ$ which is a symmetric matrix.

(ii) The order of column vectors in $\tilde{U}$ is the one that the corresponding eigenvalues are in ascending order.

(iii) The $\tilde{m}$ eigenvalues of $QXQ$ that corresponds to $\tilde{U}$ coincide with the first $\tilde{m}$ eigenvalues of $R_0$.  

The proof of Theorem 5 is already given in Takeuchi \[18\] and we will not repeat it here. However, it should be noted that in the process of the proof of Theorem 5, we have shown

$$\frac{\partial}{\partial U} L(\tilde{Z}, \tilde{U}, \lambda) = -2QXQ\tilde{U} + 2\tilde{U}\lambda,$$  \hspace{1cm} (39)

and

$$\frac{\partial}{\partial \xi_i} L(\tilde{Z}, \tilde{U}, \lambda) = -u_{i}QXQ_{u_{i}} + \psi_{i}.$$  \hspace{1cm} (40)

Thus, we have

«**Corollary 2**» Assume (C-1)-(C-3). For simplicity, let $W \triangleq QXQ$ and let $u_{i}$ denotes the $i$th column vector of $\tilde{U}$.

(i) For any $\tilde{Z} > 0$ which satisfies (12), the necessary condition of optimality of $\tilde{U}$ is given by

$$\frac{\partial}{\partial U} L(\tilde{Z}, \tilde{U}, \lambda) = -2W\tilde{U} + 2\tilde{U}\lambda = 0,$$  \hspace{1cm} (41)

which implies that $u_{i}$ is optimal when it is an eigenvector of the symmetric matrix $W$.

(ii) For any $\tilde{U} \in \mathcal{G}(A)$, the necessary condition of optimality of $\tilde{Z}$ is given by

$$\frac{\partial}{\partial \xi_i} L(\tilde{Z}, \tilde{U}, \lambda) = -u_{i}W_{u_{i}} + \psi_{i} = 0.$$  \hspace{1cm} (42)

for $i = 1, 2, \ldots, \tilde{m}$, under the assumption

$$\xi_1 > \xi_2 > \cdots > \xi_m > 0.$$  \hspace{1cm} (43)

(iii) When $u_{i}W_{u_{i}} > \psi_{i}$, we have $\frac{\partial}{\partial \xi_i} L(\tilde{Z}, \tilde{U}, \lambda) < 0$.

(iv) When $u_{i}W_{u_{i}} < \psi_{i}$, we have $\frac{\partial}{\partial \xi_i} L(\tilde{Z}, \tilde{U}, \lambda) > 0$.

(v) When $u_{i}W_{u_{i}} = \psi_{i}$, we have $\frac{\partial}{\partial \xi_i} L(\tilde{Z}, \tilde{U}, \lambda) = 0$.  

(Proof) The condition (41) is clear from (39). We can easily see that if we have (41), then $u_{i}$ is an eigenvector of $W$ as follows. From (41), we have

$$\tilde{U}W\tilde{U}\tilde{Z} = \tilde{Z}.$$  \hspace{1cm} (44)

Since $\Lambda$ is a symmetric matrix, the product of the symmetric matrix $\tilde{U}W\tilde{U}$ and the diagonal matrix $\tilde{Z}$ must be symmetric. Since we are assuming (43), this implies that $\tilde{U}W\tilde{U}$ is a diagonal matrix. Thus, each $u_{i}$ is an eigenvector of $W$. The rest part (ii)-(v) are clear from (40).
4 The Conditions of Existence and Non-Existence of the Solution

4.1 The Special Case: \( \tilde{m} = m = n \)

For simplicity, let us denote the optimal value of \((\hat{U}^o, \hat{Z})\) of Problem 2 for \( \tilde{m} = m = n \) by \((\hat{U}^o, \hat{Z})\):

\[
\hat{U}^o_{\tilde{m}} = [u^x_1 u^x_2 \ldots u^x_m],
\]

\[
\hat{Z}^o_{\tilde{m}} = \text{diag}(\hat{\xi}_1^o, \hat{\xi}_2^o, \ldots, \hat{\xi}_n^o).
\]

(45)

Also, for \( \tilde{m} = n - 1 \) and \( m = n \) by \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\):

\[
\hat{U}^o_{n-1} = [u^o_1 u^o_2 \ldots u^o_{n-1}],
\]

\[
\hat{Z}^o_{n-1} = \text{diag}(\hat{\xi}_1^o, \hat{\xi}_2^o, \ldots, \hat{\xi}_{n-1}^o).
\]

(46)

For \( \tilde{m} = m = n \), if we accept \( \hat{\xi}_n \) takes zero and consider the optimization of the rest variables:

\[
\hat{U}_{n-1} = [u_1 u_2 \ldots u_{n-1}],
\]

\[
\hat{Z}_{n-1} = \text{diag}(\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_{n-1}).
\]

(47)

Then, the solution is clearly given by \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\) and \(W = W_{n-1}\), the optimal value of \(W = QXQ^T\) for \( \tilde{m} = n - 1 \). Hence, the \( n - 1 \) eigenvectors of \(W = W_{n-1}\) in the eigenvalue-ascending-order coincide with the corresponding column vectors of \(U_{n-1}\). In this situation, the \( n \) th column vector \( u_n \) of \( \hat{U}^o \) is uniquely determined by the relation:

\[
u^o_n u^o_n = 0, \quad u^o_n u^o_n = 1,
\]

(48)

except for the fact that if \( u_n \) satisfies (48) then \( -u_n \) also does. Therefore, we have the following theorem.

**Theorem 6** Assume (C-1)-(C-3). Also, let us assume that \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\) exists and that if \((\hat{U}^o, \hat{Z}^o)\) exists in addition, then there is no change of order of the first \( k - 1 \) column vectors between \(\hat{U}^o\) and \(\hat{U}^o_{n-1}\), i.e., we will assume

\[
[u^o_{n-1} u^o_{n-1}^T] > [u^o_{n-1} u^o_{n-1}^T], \quad \ell \neq k,
\]

\(k = 1, 2, \ldots, n - 1\).

(49)

For \( \tilde{m} = m = n \) and for the problem at the point \((\hat{U}_n, \hat{Z}_n) = ([\hat{U}^o_n u_n], \text{diag}(\hat{\xi}_n^o, 0))\), we have the following properties:

(i) \( u_n \) given by (48) is the \( n \) th eigenvector of \( W \).

(ii) The \( n \) th eigenvalue of \( W \) is given by \( u^TWu_n \).

(iii) If we have

\[
u^TWu_n > \psi_n,
\]

then there exists a solution \((\hat{U}^o, \hat{Z}^o)\) of Problem 2.

(iv) If we have

\[
u^TWu_n \leq \psi_n,
\]

then there does not exist any solution \((\hat{U}^o, \hat{Z}^o)\) of Problem 2.

**Proof** The properties (i) and (ii) are clear. As for (iii), because of (50) and Corollary 2 (iii), there exists a positive value of \( \hat{\xi}_n \) for which \( J \) given by (32) takes a smaller value than the present point \( \hat{\xi}_n = 0 \). Hence, there exists a solution \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\) in the region \( \hat{\xi}_n > 0 \). Finally, (iv) is true by the fact that \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\) is the point in which \( J \) given by (32) takes the minimal value under the condition \( \hat{\xi}_n = 0 \), and for any \( \hat{\xi}_n > 0 \), \( J \) given by (32) takes a larger value than the one at the present point because of (51) and Corollary 2 (iv). This completes the proof.

4.2 The General Case: \( \tilde{m} \leq m \leq n \)

In this section, for simplicity, let us denote the optimal value of \((\hat{U}, \hat{Z})\) for \( \tilde{m} = k (\leq m \leq n) \) by \((\hat{U}^o, \hat{Z}^o)\):

\[
\hat{U}^o_k = [u^o_1 u^o_2 \ldots u^o_k],
\]

\[
\hat{Z}^o_k = \text{diag}(\hat{\xi}_{1k}, \hat{\xi}_{2k}, \ldots, \hat{\xi}_{kk}).
\]

(52)

Also, for \( \tilde{m} = k - 1 \) by \((\hat{U}^o_{k-1}, \hat{Z}^o_{k-1})\):

\[
\hat{U}^o_{k-1} = [u^o_1 u^o_2 \ldots u^o_{k-1}],
\]

\[
\hat{Z}^o_{k-1} = \text{diag}(\hat{\xi}_{1:k-1}, \hat{\xi}_{2:k-1}, \ldots, \hat{\xi}_{k:k-1}).
\]

(53)

**Theorem 7** Assume (C-1)-(C-3). For \( n \), let \( m (\leq n) \) be fixed, and let \( 1 \leq k \leq n \). Also, let us assume that \((\hat{U}^o_{n-1}, \hat{Z}^o_{n-1})\) exists and that \((\hat{U}^o, \hat{Z}^o)\) exists in addition, then there is no change of order of the first \( k - 1 \) column vectors between \(\hat{U}^o\) and \(\hat{U}^o_{k-1}\), i.e., we will assume

\[
[u^o_{kq} u^o_{kq}^T] > [u^o_{kq} u^o_{kq}^T], \quad \ell \neq q,
\]

\(q = 1, 2, \ldots, k - 1\).

(54)

For \( \tilde{m} = k \) and for the problem at the point \((\hat{U}_k, \hat{Z}_k) = ([\hat{U}^o_k u_k], \text{diag}(\hat{\xi}_k^o, 0))\), we have the following properties:

(i) If we have at least one unit vector \( \eta \in R^k \), \( \eta^\top \eta = 1 \) which satisfies both \( \eta^\top \hat{U}^o_k = 0 \) and

\[
\eta^\top W \eta > \psi_k,
\]

then there exists a solution \((\hat{U}^o, \hat{Z}^o)\) of Problem 2 for \( \tilde{m} = k (\leq m \leq n) \).

(55)

(ii) For any unit vector \( \eta \in R^k \), \( \eta^\top \eta = 1 \) which satisfies \( \eta^\top \hat{U}^o_{k-1} = 0 \), if we always have

\[
\eta^\top W \eta \leq \psi_k,
\]

then there does not exist any solution \((\hat{U}^o, \hat{Z}^o)\) of Problem 2 for \( \tilde{m} \geq k \).

**Proof** Based on the solution \((\hat{U}^o_{k-1}, \hat{Z}^o_{k-1})\) for \( \tilde{m} = k - 1 \), let us consider finding \((\hat{U}^o, \hat{Z}^o)\) for \( \tilde{m} = k \) starting with \( \hat{\xi}_k = 0 \). First, (i) is easily seen by the fact that because of (55) and Corollary 2 (iii), there exists a positive value of \( \hat{\xi}_k \) for which \( J \) given by (32) takes a smaller value than the present point \( \hat{\xi}_k = 0 \). Hence, there exists a solution \((\hat{U}^o_{k-1}, \hat{Z}^o_{k-1})\) in the region \( \hat{\xi}_k > 0 \). Also, (ii) is true by the fact that \((\hat{U}^o_{k-1}, \hat{Z}^o_{k-1})\) is the point in which \( J \) given by (32) takes the minimal value under the condition \( \hat{\xi}_k = 0 \), and for any \( \hat{\xi}_k > 0 \), \( J \) given by (32) takes a larger value than the one at the present point because of (56) and Corollary 2 (iv). This completes the proof.
4.3 An Efficient Selection of $M$ to Avoid the Non-Existence of the Solution

We see from Theorem 6 (iv) and Theorem 7 (ii) that the solution of Problem 2 does not exist for small values of $W = QQ'$. Since $Q$, the solution of (4), is uniquely determined by assigning $(U, \Xi)$, we have to get a larger value of $X$ in order to let $W$ be larger. Because (38) is a Lyapunov equation, we have a small value of $X$ when $M$ is small and $F$ given by (36) is strongly asymptotically stable. Note that the left-hand side of (38) is linear in both $X$ and $M$. Hence, it is easily seen that if we have a solution $X = X_0$ of (38) for $M = M_0$, then we have $X = \alpha X_0$ for $M = \alpha M_0, \alpha > 0$. Thus, by taking $\alpha$ in $M = \alpha M_0$ sufficiently large enough to satisfy (55), we can rather easily avoid the situation where Problem 2 has no solution.

5 Conclusions

In this paper, we considered the optimization problem of the observation gain matrix for the stationary Kalman-Bucy filter. For a fixed value of SSNRM, the rest degree of freedom is always described by a matrix multiplication with the same number of m-dimensional columns as the rank of the observation gain matrix. It is always true that the optimal value of this matrix is such that the column vectors are the required number of eigenvectors of the noise covariance matrix corresponding to the eigenvalues in ascending order.

Roughly speaking, the results in Section 4 are as follows.
(i) There are cases where the optimal observation gain with the assigned rank does not exist. This occurs when the weight matrix $M$ is small and the optimal point is located in the infeasible region with non-positive eigenvalues of SSNRM.
(ii) One can rather easily avoid this situation by increasing $M$, the weight matrix in the cost function, to an appropriate value.

As for the numerical algorithm of computing the optimal observation gain, the fixed point iteration method works well and effective. The discussions and the results on this theme will be reported in a separate paper in near future.

References
[1] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," Prep. 5th IFAC World Congress, Paris, 1/11, 1972.
[2] A. Logothetis, A. Isaksson and E. Skadas, "On sensor scheduling via information theoretic criteria," Prep. ACC, 1999.
[3] M. Aoki, and M. T. Li, "Optimal discrete time control system with cost for observation," IEEE Trans. Automatic Control, AC-14, 2, 165/175, 1969.
[4] A. Sano and M. Terao, "Measurement optimization in optimal process control," Automatica, 5, 705/714, 1970.
[5] S. Ibara, "Coding theory in white Gaussian channel with feedback," J. Multivariate Analysis, 4, 74/87, 1974.
[6] Y. Takeuchi, "The optimal transmission of non-Gaussian signals through a noisy channel with feedback," IEEE Trans. Inform. Theory, IT-40, 5, 1624/1629, 1994.
[7] Y. Takeuchi and A. Ida, "Optimal Transmission of a set of Gaussian Signals through Discrete-Time Channels with Feedback," Proc. of 35th ISCIE SSS, Ube, Oct. 30-31, 2003, 118/123, 2004.
[8] Y. Takeuchi, "An information theoretic approach to optimization of linear observations for the Kalman-Bucy filter," International Journal of Innovative Computing, Information & Control, 1-3, 401/416, 2005.
[9] Y. Takeuchi, "Optimal transmission of a set of discrete-time Gaussian signals through channels with feedback," International Journal of Innovative Computing, Information & Control, 2-5, 927/942, 2006.
[10] Y. Takeuchi, "Optimization of linear observations for the stationary Kalman filter based on a generalized Water Filling Theorem," International Journal of Innovative Computing, Information & Control, 4-1, 211/230, 2008.
[11] Y. Takeuchi and A. Hirata, "A method of optimization of linear observations for the Kalman filter based on a generalized Water Filling Theorem," International Journal of Innovative Computing, Information & Control, 5-1, 75/85, 2009.
[12] Y. Takeuchi and M Inoue, "Optimization of observations for LQG control systems by an information theoretic approach," International Journal of Innovative Computing, Information & Control, 6-1, 75/87, 2010.
[13] Y. Takeuchi, "Optimization of linear observations for the stationary Kalman filter under a quadratic performance criterion," International Journal of Innovative Computing, Information & Control, 7-1, 85/99, 2011.
[14] Y. Takeuchi, "Optimization of the observations for stationary LQG stochastic control systems under a quadratic criterion," International Journal of Innovative Computing, Information & Control, 8-3, 2313/2328, 2012.
[15] Y. Takeuchi and R. Matsumi, "Hybrid optimization of the observations for stationary LQG control systems by an information theoretic approach," Trans. ISCIE, 25-12, 366/374, 2012.
[16] Y. Takeuchi and H. Nakai, "A New Method of Optimization of Linear Observations for the Stationary Kalman Filter," Proc. of 44th ISCIE SSS, Tokyo, Nov. 1-2, 2012, 21/28, 2013.
[17] Y. Takeuchi and H. Nakai, "An Improved Method of Optimization of Linear Observations for the Stationary Kalman Filter," Proc. of 45th ISCIE SSS, Okinawa, Nov. 1-2, 2013, 1/8, 2014.
[18] Y. Takeuchi, "Optimization of Linear Observations for the Stationary Kalman-Bucy Filter," Trans. ISCIE, 28-12, 467/475, 2015.
[19] S. Kodama, and N. Suda, Matrix Theory for Systems Control (in Japanese), SICE, 1978.
[20] V. Kučera, "A contribution to matrix quadratic equations," IEEE Trans. Automatic Control, AC-17, 344/347, 1972.