Hybrid Optimization of the Observations for Stationary LQG Control Systems*

Yoshiki TAKEUCHI[†] and Ryo MATSUMI[†]

In this paper, we consider an optimization problem for observations of stationary LQG stochastic control systems which employ the stationary Kalman filter. The performance of the Kalman filter and that of the LQG stochastic optimal control are both dependent on the gain matrix in the linear observation. One of the authors has already developed methods of optimizing this gain matrix based on the estimation or control individual performance under a quadratic performance criterion. This paper discusses a hybrid problem by taking into account of both estimator and regulator performances. By introducing the eigenvalues-eigenvectors representation of a nonnegative definite symmetric matrix, the condition of optimality is derived. Also, numerical calculations are easily carried out by introducing multi-dimensional polar coordinates systems.

1. Introduction

In this paper, we consider an optimization problem for observations of the LQG optimal control systems. This type of problem is rather classical, and there were a number of studies.^{[1]–[16]} A brief review of the researches in this topic is given in [12]. The performance of the estimator by the stationary Kalman filter and that of the stationary LQG regulator both depend on the value of the gain matrix in the linear observations. One of the authors already considered optimization problems concerned with the observation gain matrix based on each of

- (1) performance of the stationary Kalman filter^[15];
- (2) performance of the stationary LQG regulator^[16].

In [15] and [16], the condition of optimality, for each case based on (1) and (2), is shown, and are obtained easily calculable numerical algorithms by introducing multi-dimensional polar coordinates systems representation^{[15],[16]} of an orthogonal matrix and by applying the connection rules^{[15],[16]} of the angular parameters.

From mathematical point of view, the two problems by (1) and (2) are nearly equivalent, and both of the two performance criterions have the same form which is an “weighted estimation error variance” plus a quadratic function of the gain matrix that accounts for the energy or cost of the observations. Namely, if we take the same weight matrix, then they are equivalent.

In real applications, however, the characters of the two weight matrices, i.e., the properties of the problems based on (1) and (2), are considerably different in the following points.

- (i) The weight matrix in (1) is often taken to be a diagonal matrix or an identity matrix multiplied by a scalar value; the variation of the eigenvalues are not so large; and any of them are rarely taken to be zero ;
- (ii) The weight matrix in (2) is dependent on the system matrix and the solution of the matrix Riccati equation by which we obtain the optimal feedback gain in the control input; there often arises the case in which the variation of the eigenvalues is considerably large and/or some of them are zero.

Namely, in order to optimize the observation gain matrix by taking into account of both estimator and control performances, it is better for us to use a hybrid type weight matrix of (1) and (2).

In this paper, we discuss the optimization of the observation with this hybrid type weight matrix for the estimation error variance. We will show the condition of optimality. Then, by introducing multi-dimensional polar coordinates systems representation^{[15],[16]} of an orthogonal matrix and by applying the connection rules^{[15],[16]} of the angular parameters, the problem is converted to a form which is easily solved by a numerical optimization method. Also, a set of formulas to compute partial derivatives with respect to the angular parameters and the eigenvalues are shown, which make us possible to use the gradient methods in the numerical optimization. A new result of numerical computation, using the steepest descent method, is provided to show that we can easily get a solution and that how the hybrid problem is related

* Manuscript Received Date: May 21, 2012

[†] Dept. of Information Science, Osaka University of Education; 4-698-1 Asahigaoka, Kashiwara city, Osaka 582-8582, JAPAN

Key Words: optimization of observations, LQG stochastic control, state estimation, Gaussian processes.

with (1) and (2). A preliminary version of this paper has been presented in [17].

Mathematical symbols, in this paper, are used in the following way. $\mathbf{R}$ is the space of all real numbers, i.e., $\mathbf{R} \triangleq (-\infty, \infty)$. For positive integers m and n , $\mathbf{R}^n$ and $\mathbf{R}^{m \times n}$ denote the spaces of n -dimensional vectors and $m \times n$ -dimensional matrices whose components take values in $\mathbf{R}$. The prime denotes the transpose of a vector or a matrix and the Euclidean norm is $|\cdot|$. Thus, for $x \in \mathbf{R}^n$, $|x| = \sqrt{x'x}$. The identity matrix of any dimension is denoted by I . The components of a matrix are denoted by using subscripts. Thus, $[A]_{ij}$ is the (i, j) -component of A . In the case where no confusion may arise, we denote $[A]_{ij}$ simply by a_{ij} . 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. For any pair of matrices A and B , $A \otimes B$ denotes the Kronecker product of A and B , and $\text{vec}(A)$ is the vector formed by stacking the columns of A into a single column vector. The triplet $(\Omega, \mathcal{F}, P)$ is a complete probability space where Ω is a sample space with elementary events ω , $\mathcal{F}$ is a σ -field of subsets of Ω , and P is a probability measure. $E\{\cdot\}$ denotes the expectation and $E\{\cdot|\mathcal{G}\}$, $\mathcal{G} \subset \mathcal{F}$ the conditional expectation, given $\mathcal{G}$, with respect to P . $\sigma\{\cdot\}$ is the minimal sub- σ -field of $\mathcal{F}$ with respect to which the family of $\mathcal{F}$ -measurable sets or random variables $\{\cdot\}$ is measurable.

2. Problem Formulation

2.1 Stationary Optimal LQG Regulator System

Let $\mathbf{x} \equiv \{x_t(\omega); t=0, 1, \dots\}$ denote the state process of an n -dimensional Gaussian stochastic system described by

$$\begin{cases} x_{t+1}(\omega) = Ax_t(\omega) + Cu(t) + Gw_t(\omega), \\ x_0(\omega) = x^0(\omega), \end{cases} \quad (1)$$

where $A \in \mathbf{R}^{n \times n}$, $C \in \mathbf{R}^{n \times r}$, $G \in \mathbf{R}^{n \times d_1}$, $x^0(\omega)$ is a Gaussian random vector with mean $\hat{x}^0$ and covariance Q^0 , $\mathbf{u} \equiv \{u(t); t=0, 1, \dots\}$ is a r -dimensional control input, and $\mathbf{w} \equiv \{w_t(\omega); t=0, 1, \dots\}$ is a d_1 -dimensional standard white Gaussian noise sequence. Suppose that the value of $\mathbf{x}$ is not available but we have m -dimensional observations described by

$$y_t(\omega) = Hx_t(\omega) + Rv_t(\omega), \quad (2)$$

where $\mathbf{y} \equiv \{y_t(\omega); t=1, 2, \dots\}$ is an m -dimensional observation process, $H \in \mathbf{R}^{m \times n}$, $R \in \mathbf{R}^{m \times d_2}$, and $\mathbf{v} \equiv \{v_t(\omega); t=1, 2, \dots\}$ is a d_2 -dimensional standard white Gaussian noise sequence. Throughout this paper, we will assume that the following two conditions are satisfied.

(C-1) $R_0 \triangleq RR' > 0$,

(C-2) $x^0(\omega)$, $\mathbf{w}$ and $\mathbf{v}$ are mutually independent.

It is well-known that the least-squares estimate $\hat{x}_{t|t}(\omega) \triangleq E\{x_t(\omega)|\mathcal{Y}_t\}$ of $x_t(\omega)$ based on $\mathcal{Y}_t \triangleq \sigma\{y_s(\omega); s=$

$1, 2, \dots, t\}$ is given by the Kalman filter:

$$\begin{cases} \hat{x}_{t|t-1}(\omega) = A\hat{x}_{t-1|t-1}(\omega) + Cu(t-1) \\ \hat{x}_{t|t}(\omega) = \hat{x}_{t|t-1}(\omega) \\ \quad + Q^-H'\{HQ^-H' + R_0\}^{-1}\tilde{y}_t(\omega), \end{cases} \quad (3)$$

and

$$\begin{cases} Q^- = AQ'A' + GG', \\ Q = Q^- - Q^-H'\{HQ^-H' + R_0\}^{-1}HQ^-, \end{cases} \quad (4)$$

where

$$\hat{x}_{t|t-1}(\omega) \triangleq E\{x_t(\omega)|\mathcal{Y}_{t-1}\}, \quad (5)$$

$$Q^- \triangleq E\{[x_t(\omega) - \hat{x}_{t|t-1}(\omega)] \times [x_t(\omega) - \hat{x}_{t|t-1}(\omega)]'\}, \quad (6)$$

$$Q \triangleq E\{[x_t(\omega) - \hat{x}_{t|t}(\omega)][x_t(\omega) - \hat{x}_{t|t}(\omega)]'\}. \quad (7)$$

Also, $\tilde{\mathbf{y}} \equiv \{\tilde{y}_t(\omega); t=1, 2, \dots\}$ in (3) is the innovations process:

$$\begin{aligned} \tilde{y}_t(\omega) &\triangleq y_t(\omega) - H\hat{x}_{t|t-1}(\omega) \\ &= H\{x_t(\omega) - \hat{x}_{t|t-1}(\omega)\} + Rv_t(\omega). \end{aligned} \quad (8)$$

The stationary optimal input $\mathbf{u} \equiv \{u(t); t=0, 1, \dots\}$ is determined based on the well-known solution of the LQG regulator problem with the quadratic performance criterion:

$$\bar{J}_C \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \sum_{t=1}^T [x_t'(\omega) M x_t(\omega) + u'(N) u] \right\}, \quad (9)$$

where $u \triangleq u(t-1)$, and $M \in \mathbf{R}^{n \times n}$ and $N \in \mathbf{R}^{r \times r}$ are, respectively, non-negative definite and positive definite symmetric matrices. As it is well-known, for (1), (2) and (9), the optimal control is given by

$$u^*(t) \triangleq -\{C'WC + N\}^{-1}C'WA\hat{x}_{t|t}(\omega), \quad (10)$$

where $W \in \mathbf{R}^{n \times n}$ is given by the positive definite solution of the matrix Riccati equation:

$$\begin{cases} W = A'YA + M, \\ Y = W - WC\{C'WC + N\}^{-1}C'W. \end{cases} \quad (11)$$

From (9)-(11), the minimal value of $\bar{J}_C$ is given by

$$\bar{J}_C^* = \text{tr}[A'(W - Y)AQ] + \text{tr}[WGG']. \quad (12)$$

2.2 Estimation-Control Hybrid Optimization of the Observations

One of the authors has already considered two types of problems for the observation given by (2). For the optimization of the performance of the least-squares state estimator, we employed the following criterion:^[15]

$$\hat{J}_E \triangleq \text{tr}[\tilde{M}Q] + \text{tr}[H\tilde{N}H'], \quad (13)$$

where $\text{tr}[\tilde{M}Q]$ is the weighted estimation error variance, i.e., for $x_t \equiv x_t(\omega)$ and $\hat{x}_{t|t} \equiv \hat{x}_{t|t}(\omega)$,

$$\text{tr}[\tilde{M}Q] = E\{[x_t - \hat{x}_{t|t}]'\tilde{M}[x_t - \hat{x}_{t|t}]\}, \quad (14)$$

and $\tilde{M} \in \mathbf{R}^{n \times n}$ and $\tilde{N} \in \mathbf{R}^{r \times r}$ are, respectively, non-negative definite and positive definite symmetric matrices. The second term in the right-hand side of (13) denotes the cost or the energy consumed by the observation.

On the other hand, from the view point of the LQG regulator performance, the observation is better when we have a smaller value of $\tilde{J}_C^*$ given by (12). Since the second term in $\tilde{J}_C^*$ is independent of H , the performance criterion was taken to be [16]

$$\hat{J}_C \triangleq \text{tr}[A'(W - Y)AQ] + \text{tr}[H\tilde{N}H']. \quad (15)$$

As we notice, the two problems given respectively by (13) and (15) are mathematically nearly equivalent, i.e., $\hat{J}_E = \hat{J}_C$ if we take $\tilde{M} = A'(W - Y)A$. However, we should note that the characters of the two matrices: $\tilde{M}$ and $A'(W - Y)A$ are considerably different in real applications.

- (i) $\tilde{M}$ is often taken to be a diagonal matrix or an identity matrix multiplied by a scalar value, the variation of the eigenvalues are not so large, and any of them are rarely taken to be zero ;
- (ii) Since $A'(W - Y)A$ is dependent on A , and on W and Y which are the solution of (11), there often arises the case in which the variation of the eigenvalues is large and/or some of them are zero.

Thus, if we want to optimize the observation by taking into account of both the estimation and control performances, we must introduce a hybrid performance criterion given by

$$\hat{J}^\alpha \triangleq \text{tr}[\{\alpha\tilde{M} + (1 - \alpha)A'(W - Y)A\}Q] + \text{tr}[H\tilde{N}H'], \quad 0 \leq \alpha \leq 1, \quad (16)$$

where α is a scalar parameter which denotes the weights for the estimation and control performances. For example, if $\alpha = 1$ then we have $\hat{J}^1 = \hat{J}_E$, and if $\alpha = 0$ then $\hat{J}^0 = \hat{J}_C$.

Thus, we are now concerned with the following problem.

【Problem 1】 Find $H \in \mathbf{R}^{m \times n}$ such that (16) is mini-mized subject to (4).

(Remark 1) Since Problem 1 is concerned with the design of the observation mechanism, all computations are carried out by off-line. Hence, if we obtain the solution of Problem 1 for a set of values of α , we can select a desirable value of α by examining the changes of the minimal values of $\hat{J}_E$ and $\hat{J}_C$ with respect to α .

The mathematical treatment and result of Problem 1 are almost equivalent to those of the problems discussed in Takeuchi [16]. We will show the outline of them below for convenience.

First, let us denote

$$\tilde{N}^{1/2}H'R_0^{-1}H\tilde{N}^{1/2} = \tilde{U}\tilde{\Xi}\tilde{U}', \quad (17)$$

where

$$\tilde{\Xi} \triangleq \text{diag}(\xi_1, \xi_2, \dots, \xi_{\tilde{m}}), \quad \xi_i > 0, \quad i = 1, 2, \dots, \tilde{m}, \quad (18)$$

$$\tilde{m} \triangleq \text{rank}[H] \ (\leq m), \quad (19)$$

and $\tilde{U} = [u_1 \ u_2 \ \dots \ u_{\tilde{m}}]$ is the set of eigenvectors of the matrix given by the left-hand side of (17) corresponding to the positive eigenvalues $\xi_i, i = 1, 2, \dots, \tilde{m}$. Note that we have $\tilde{U}'\tilde{U} = I$. Also, we can always find a pair of orthogonal and diagonal matrices for which we have (17) because $\tilde{N}^{1/2}H'R_0^{-1}H\tilde{N}^{1/2}$ is a nonnegative definite symmetric matrix [18; Sec. 8.2 Corollary 2]. Without loss of generality, we can assume that

$$\xi_1 \geq \xi_2 \geq \dots \geq \xi_{\tilde{m}} > 0. \quad (20)$$

Since (17) implies that

$$H'R_0^{-1}H = \tilde{N}^{-1/2}\tilde{U}\tilde{\Xi}\tilde{U}'\tilde{N}^{-1/2}, \quad (21)$$

we have the following theorem which guarantees that we can take H in the form:

$$H = R_0^{1/2}\tilde{F}\tilde{\Xi}^{1/2}\tilde{U}'\tilde{N}^{-1/2}, \quad (22)$$

where $\tilde{F} \in \mathbf{R}^{m \times \tilde{m}}$ denotes the first $\tilde{m}$ columns of an orthogonal matrix $\Gamma \in \mathbf{R}^{m \times m}$ such that $\Gamma\Gamma' = \Gamma'\Gamma = I$.

【Theorem 1】 [15] Assume (C-1) and (C-2). Then, any $H \in \mathbf{R}^{m \times n}$ which satisfies (17) for a fixed set of values $\tilde{U}$, $\tilde{\Xi}$ and $\tilde{N}$ yields the same value of Q .

(Proof) We have the assertion immediately by noting the fact that the second relation of (4) can be rewritten as

$$Q^{-1} = (Q^-)^{-1} + H'R_0^{-1}H. \quad (23)$$

Thus, without loss of generality, we can take H as the form given by (22) which is an expression of H with property (17) and/or (21). Thus, the problem has been converted to the optimization with respect to $\tilde{F} \in \mathbf{R}^{m \times \tilde{m}}$, $\tilde{U} \in \mathbf{R}^{n \times \tilde{m}}$ and $\tilde{\Xi} \triangleq \text{diag}(\xi_1, \xi_2, \dots, \xi_{\tilde{m}})$.

(Remark 2) Theorem 1 and (21) imply that there are a number of minimum points in the space of H , and that by the change of variables $H \rightarrow (\tilde{U}, \tilde{\Xi})$, these minimum points are mapped into a single point in the space of $(\tilde{U}, \tilde{\Xi})$. This is a main reason that the method of optimization of H in this paper is efficient.

3. The Condition of Optimality

As we see from (21), (22) and (23), the value of Q is independent of that of $\tilde{F} \in \mathbf{R}^{m \times \tilde{m}}$ and is determined by (21), (23) and the first part of (4). Hence, the optimal value of $\tilde{F}$ should be determined in such a way that $\text{tr}[H\tilde{N}H']$, the second term in (16), is minimized.

【Theorem 2】 Assume (C-1)-(C-2). Then the optimal value of $\tilde{F} \in \mathbf{R}^{m \times \tilde{m}}$ is given by the set of eigenvectors of R_0 corresponding to the first $\tilde{m}$ eigenvalues in ascending order, i.e.,

$$R_0 = [\tilde{F} \ \bar{F}] \Psi \begin{bmatrix} \tilde{F}' \\ \bar{F}' \end{bmatrix}, \quad (24)$$

$$\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_{\tilde{m}}, \dots, \psi_m),$$

$$\psi_1 \leq \psi_2 \leq \dots \leq \psi_{\tilde{m}} \leq \dots \leq \psi_m.$$

(Remark 3) If we take $\tilde{F} \in \mathbf{R}^{m \times \tilde{m}}$ according to Theorem 2, we have

$$H = \tilde{F} \tilde{\Psi}^{1/2} \tilde{\Xi}^{1/2} \tilde{U}' \tilde{N}^{-1/2}, \quad (25)$$

where

$$\tilde{\Psi} \triangleq \text{diag}(\psi_1, \psi_2, \dots, \psi_{\tilde{m}}). \quad (26)$$

(Remark 4) For the existence and the uniqueness of nonnegative definite and positive definite solutions of (4), explicit conditions are obtained by the transformation of (4) into the continuous-time-type Riccati equation:

$$\hat{A}\hat{Q} + \hat{Q}\hat{A}' + \hat{G}\hat{G}' - \hat{Q}\hat{H}'\hat{H}\hat{Q} = 0,$$

where $\hat{A}$, $\hat{G}$ and $\hat{H}$ are given by [16; Sec. 13.5],

$$\begin{aligned} \hat{A} &\triangleq (A - I)(A + I)^{-1} - W_1' W_2^{-1} \hat{D}, \\ \hat{H} &\triangleq W_2^{-1/2} \hat{D}, \quad \hat{G}\hat{G}' \triangleq G G' - W_1 W_2^{-1} W_1, \\ \hat{D} &\triangleq 2\hat{B}'(A' + I)^{-2}, \quad \hat{B} \triangleq H' R_0^{-1/2}, \\ W_1 &\triangleq -G G'(A' + I)\hat{B}, \\ W_2 &\triangleq I + \hat{B}'(A + I)^{-1} G G'(A' + I)^{-1} \hat{B}. \end{aligned}$$

By applying the result of [19], a positive-definite solution $\hat{Q}$ is obtained if $(\hat{A}, \hat{G})$ is controllable and $(\hat{H}, \hat{A})$ is observable. Furthermore, from (21) and (23), we have

$$Q = Q^- - Q^- \tilde{N}^{-1/2} \tilde{U} \times \{ \tilde{U}' \tilde{N}^{-1/2} Q^- \tilde{N}^{-1/2} \tilde{U} + \tilde{\Xi}^{-1} \}^{-1} \tilde{U}' \tilde{N}^{-1/2} Q^-.$$

Thus, the condition is obtained in terms of A , G , $\tilde{N}$, $\tilde{U}$ and $\tilde{\Xi}$ by the following substitutions:

$$H \rightarrow \tilde{U}' \tilde{N}^{-1/2}, \quad R_0 \rightarrow \tilde{\Xi}^{-1}.$$

As we will see, this condition is satisfied for the example given in Section 7, and there, we have the positive-definite solution of (4).

Now, we have converted Problem 1 to the following form.

【Problem 2】 For convenience, let

$$K(\alpha) \triangleq \alpha \tilde{M} + (1 - \alpha) A'(W - Y)A, \quad 0 \leq \alpha \leq 1, \quad (27)$$

Then, find $\tilde{U} \in \mathbf{R}^{n \times \tilde{m}}$ and $\tilde{\Xi} = \text{diag}(\xi_1, \xi_2, \dots, \xi_{\tilde{m}})$ such that

$$\hat{J}^\alpha = \text{tr}[K(\alpha)Q] + \text{tr}[\tilde{\Psi} \tilde{\Xi}] \rightarrow \min., \quad (28)$$

subject to (4), (20), (25) and

$$\tilde{U}' \tilde{U} = I. \quad (29)$$

For Problem 2, let us define the Lagrangean by

$$L(\tilde{\Xi}, \tilde{U}, \tilde{\Lambda}) \triangleq \text{tr}[K(\alpha)Q] + \text{tr}[\tilde{\Psi} \tilde{\Xi}] + \text{tr}[\tilde{\Lambda}(\tilde{U}' \tilde{U} - I)], \quad (30)$$

where $\tilde{\Lambda} \in \mathbf{R}^{\tilde{m} \times \tilde{m}}$ is a symmetric matrix whose (i, j) -component is a Lagrange multiplier for the same component of (29), i.e.,

$$\begin{aligned} \text{tr}[\tilde{\Lambda}(\tilde{U}' \tilde{U} - I)] &= \sum_{i=1}^{\tilde{m}} \sum_{j=1}^{\tilde{m}} \lambda_{ij} [(\tilde{U}' \tilde{U} - I)]_{ji} \\ &= \sum_{i=1}^{\tilde{m}} \sum_{j=1}^{\tilde{m}} \lambda_{ij} [(\tilde{U}' \tilde{U} - I)]_{ij}. \end{aligned} \quad (31)$$

For Problem 2 and (30), we have the following result.

【Theorem 3】 (Condition of Optimality) Assume (C-1), (C-2) and

(C-3) For $F \triangleq Q(Q^-)^{-1}A$, the set:

$$\mathcal{H} \triangleq \{(\tilde{U}, \tilde{\Xi}); \det|F \otimes F - I| \neq 0\},$$

is not empty.

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

$$\tilde{N}^{-1/2} Q X Q \tilde{N}^{-1/2} \tilde{U} = \tilde{U} \tilde{\Psi}, \quad (32)$$

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

$$X = F' X F + K(\alpha). \quad (33)$$

Since $\tilde{\Psi}$ is a diagonal matrix given by (26), (32) implies that

【Corollary 1】 Assume (C-1)-(C-3). The optimal $(\tilde{U}, \tilde{\Xi})$ is such that

- (i) Each column vector of $\tilde{U}$ is an eigenvector of $\tilde{N}^{-1/2} Q X Q \tilde{N}^{-1/2}$ 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 $\tilde{N}^{-1/2} Q X Q \tilde{N}^{-1/2}$ corresponding to $\tilde{U}$ coincide with the first $\tilde{m}$ eigenvalues of R_0 .

(Remark 5) Clearly, the condition: $\det|F \otimes F - I| \neq 0$ is equivalent to the one that no eigenvalue of $F \otimes F$ is equal to 1. Let $\mu_i, i = 1, 2, \dots, n$ denote the eigenvalues of $F \triangleq Q(Q^-)^{-1}A$. Then, the condition holds if and only if the following two conditions are fully satisfied.

- (i) $\mu_i \neq 1, i = 1, 2, \dots, n$.
- (ii) $\mu_i \mu_j \neq 1, i < j, i, j = 1, 2, \dots, n$.

(Remark 6) As we see, the set of optimal values $(Q, Q^-, \tilde{U}, \tilde{\Xi})$ is given by a solution of the set of equations (4), (22), (29), (32) and (33). Although it is not easy to get an analytical solution of these equations, these relations together with the properties given in Corollary 1 are applicable in constructing a recursive numerical algorithms.

4. Representation of $\tilde{U}$ by Polar Coordinates Systems

In this section, we convert the constrained problem given by (28) and (29) to an unconstrained one by introducing multi-dimensional polar coordinates systems in $\mathbf{R}^n$. Let us denote

$$\tilde{U} = T(n) \left\{ \prod_{\ell=1}^{\bar{k}} \begin{bmatrix} I_{\ell \times \ell} & 0 \\ 0 & T(n-\ell) \end{bmatrix} \right\} \begin{bmatrix} I_{\bar{m} \times \bar{m}} \\ 0 \end{bmatrix}, \quad (34)$$

where

$$\bar{k} \triangleq \min(\bar{m} - 1, n - 2), \quad (35)$$

and, for $k = n, n-1, \dots, n-\bar{k}$,

$$T(k) \triangleq T(k; \theta_{k1}, \theta_{k2}, \dots, \theta_{kk-1}) \\ = [z_{k1} \ z_{k2} \ \dots \ z_{kk}], \quad (36)$$

$$z_{k1} = \begin{bmatrix} \prod_{i=1}^{k-1} \cos \theta_{ki} \\ \sin \theta_{k1} \prod_{i=2}^{k-1} \cos \theta_{ki} \\ \vdots \\ \sin \theta_{kj-1} \prod_{i=j}^{k-1} \cos \theta_{ki} \\ \vdots \\ \sin \theta_{kk-2} \cos \theta_{kk-1} \\ \sin \theta_{kk-1} \end{bmatrix}, \quad (37a)$$

$$z_{k\ell} = \begin{bmatrix} -\sin \theta_{kk-\ell+1} \prod_{i=1}^{k-\ell} \cos \theta_{ki} \\ -\sin \theta_{kk-\ell+1} \sin \theta_{k1} \prod_{i=2}^{k-\ell} \cos \theta_{ki} \\ \vdots \\ -\sin \theta_{kk-\ell+1} \sin \theta_{kj-1} \prod_{i=j}^{k-\ell} \cos \theta_{ki} \\ \vdots \\ -\sin \theta_{kk-\ell+1} \sin \theta_{kk-\ell} \\ \cos \theta_{kk-\ell+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (37b)$$

$$\ell = 2, 3, \dots, k-1,$$

$$z_{kk} = \begin{bmatrix} -\sin \theta_{k1} \\ \cos \theta_{k1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (37c)$$

and

$$0 \leq \theta_{k1} \leq 2\pi, \quad -\frac{\pi}{2} \leq \theta_{ki} \leq \frac{\pi}{2}, \quad i = 2, 3, \dots, k-1. \quad (38)$$

Then, it can be easily seen that $T'(k)T(k) = T(k)T'(k) = I_{k \times k}$ and, hence, we have (29) for $\tilde{U}$ given by (34). The fact that all values in $\mathbf{R}^{n \times \bar{m}}$ of $\tilde{U}$ with property (29) is given by (34)-(38) is shown in [14], [15].

Thus, we have converted Problem 2 with constraint (29) to the one with the unconstrained angular variables given by (38) for $k = n, n-1, \dots, n-\bar{k}$.

5. A Method of Connections of Angular Parameters at the Boundary

For simplicity, let

$$\Theta \triangleq \{\theta_{ki}, i = 1, 2, \dots, k-1, k = n, \dots, n-\bar{k}\}. \quad (39)$$

Clearly, $\tilde{U} = \tilde{U}(\Theta)$ is a periodic function of Θ . Hence, for an exterior point of the domain given by (38), there always exists an interior point (of the domain) for which $\hat{J}$ has the same value as the exterior point. If we can find such a pair of values of Θ , the implementation of the searching algorithm of finding the minimal point of $\hat{J}$ becomes much simpler by replacing the exterior point generated by the algorithm with the corresponding interior point. This is because, in usual cases of optimization, we must stop searching when the algorithm generates an exterior point. In such a case, we usually take a nearest boundary point and memorize the value of the objective function at that point as a local minimal point. However, by using the replacement of the exterior point with the interior one, we can continue searching until we really reach a minimal point. By the following theorem, we show that it suffices to search over the set with $1/2^{k+1}$ size:

$$0 \leq \theta_{k1} \leq \pi, \quad -\frac{\pi}{2} \leq \theta_{ki} \leq \frac{\pi}{2}, \quad (40) \\ i = 2, 3, \dots, k-1, \quad k = n, n-1, \dots, n-\bar{k},$$

and obtain the rule of replacement of exterior points of (40) with the corresponding interior points.

【Theorem 4】 (Connections of Angular Parameters)^[14] For $\tilde{U}$ given by (34), we have the following rules for the case when only one parameter violates (40). For $k = n, n-1, \dots, \bar{k}$, let

$$\theta_k \triangleq [\theta_{k1} \ \theta_{k2} \ \dots \ \theta_{kk-1}]', \quad (41)$$

$$q \in \{n, n-1, \dots, n-\bar{k}\}, \quad (42)$$

and assume that all $\theta_k, k \neq q$ satisfy (40). Then,

(i) When $\theta_{q1} \in [-\pi, 0)$ or $\theta_{q1} \in (\pi, 2\pi]$ and $\theta_{qi} \in [-\pi/2, \pi/2], i = 2, 3, \dots, q-1$, $\hat{J}^*$ keeps the same value after the following replacements.

(a) q : odd,

(44) and (49) are trivial by using (28), (34) and (37). Let us show (44). Substituting (21) into (23), we have

$$Q^{-1} = (Q^-)^{-1} + \tilde{N}^{-1/2} \tilde{U} \tilde{\Xi} \tilde{U}' \tilde{N}^{-1/2}. \quad (50)$$

Then, noting that

$$Q_{\theta_{ki}} = \frac{\partial Q}{\partial \theta_{ki}} = \frac{\partial (Q^-)^{-1}}{\partial \theta_{ki}} = -Q \frac{\partial Q^-}{\partial \theta_{ki}} Q, \quad (51)$$

and that

$$\frac{\partial (Q^-)^{-1}}{\partial \theta_{ki}} = -(Q^-)^{-1} \frac{\partial Q^-}{\partial \theta_{ki}} (Q^-)^{-1}, \quad (52)$$

together with the fact that the first relation of (4) implies

$$\frac{\partial Q^-}{\partial \theta_{ki}} = A \frac{\partial Q}{\partial \theta_{ki}} A' = A Q_{\theta_{ki}} A', \quad (53)$$

we have (44) because $F \triangleq Q(Q^-)^{-1}A$. Also, (49) can be obtained by the same procedure.

(Remark 7) As we see from (44) and (49), each of $Q_{\theta_{ki}}$ and Q_{ξ_i} is determined as a solution of a Lyapunov type matrix equation for any value of $(\tilde{U}, \tilde{\Xi}, Q, Q^-)$.

7. A Numerical Example

In this section, we will give an illustrative example for a 3-dimensional LQG system for signal $\mathbf{x}$ and observations $\mathbf{y}$ ($n=m=3$).

【Example 1】 Let us consider a 3-dimensional system with

$$\begin{aligned} A &= \begin{bmatrix} 0.5 & 0.3 & 0.1 \\ 0.2 & 0.4 & 0.2 \\ 0.1 & 0.5 & 0.6 \end{bmatrix}, \quad G = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}, \\ C &= \begin{bmatrix} 1.2 & 0.1 & 0.2 \\ 0.2 & 1.4 & 0.2 \\ 0.1 & 0.5 & 1.1 \end{bmatrix}, \quad M = 50.0I, \quad N = I, \\ R_0 &= \begin{bmatrix} \frac{1}{8} & -\frac{3\sqrt{3}}{8} & \frac{3}{4} \\ -\frac{3\sqrt{3}}{8} & \frac{5}{8} & \frac{\sqrt{3}}{4} \\ \frac{3}{8} & -\frac{\sqrt{3}}{8} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0.49 & 0 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0 & 1.44 \end{bmatrix} \\ &\times \begin{bmatrix} \frac{1}{8} & -\frac{3\sqrt{3}}{8} & -\frac{3}{4} \\ -\frac{3\sqrt{3}}{8} & \frac{5}{8} & -\frac{\sqrt{3}}{4} \\ \frac{3}{8} & \frac{\sqrt{3}}{8} & -\frac{1}{4} \end{bmatrix} \\ &= \begin{bmatrix} 1.2395 & 0.1015 & -0.2128 \\ 0.1015 & 0.8673 & -0.3437 \\ -0.2128 & -0.3437 & 0.8231 \end{bmatrix}, \end{aligned}$$

i.e.,

$$\begin{aligned} \Psi &= \text{diag}(0.49, 1.0, 1.44), \\ \Gamma &= \begin{bmatrix} \frac{1}{8} & -\frac{3\sqrt{3}}{8} & \frac{3}{4} \\ -\frac{3\sqrt{3}}{8} & \frac{5}{8} & \frac{\sqrt{3}}{4} \\ \frac{3}{8} & -\frac{\sqrt{3}}{8} & \frac{1}{4} \end{bmatrix}. \end{aligned}$$

For the above system and observations, we have

$$W = \begin{bmatrix} 50.2999 & 0.2837 & 0.1517 \\ 0.2837 & 50.5010 & 0.4172 \\ 0.1517 & 0.4172 & 50.4148 \end{bmatrix},$$

and

$$Y = \begin{bmatrix} 0.7209 & -0.0854 & -0.1348 \\ -0.0854 & 0.6918 & -0.4141 \\ -0.1348 & -0.4141 & 0.9677 \end{bmatrix}.$$

By taking $\tilde{M} = 50.0 \cdot I$ and $\tilde{N} = (1/70.0) \cdot I$, we have made numerical computations for three values of $\tilde{m}$: $\tilde{m} = 3$, $\tilde{m} = 2$ and $\tilde{m} = 1$, and six values of α : $\alpha = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. We have carried out the optimization by using the steepest descent method which applies the derivatives given in Section 6 with respect to the angular parameters: θ_{31}, θ_{32} and θ_{21} , and $\tilde{\Xi} \triangleq \text{diag}(\xi_1, \dots, \xi_{\tilde{m}})$, and by making use of the connection rule shown in Section 5. The result are summarized in Table 1, Table 2 and Table 3. These results are similar as those in [17] except for small errors which is inherent in the numerical computations with the alternate search algorithm. In Fig. 1, the results of the optimization of θ_{31}, θ_{32} and θ_{21} by the steepest descent method for $\tilde{m} = 3$ is shown. The corresponding runs of ξ_1, ξ_2 and ξ_3 with the initial values $\xi_1 = 13.0, \xi_2 = 11.0$ and $\xi_3 = 10.0$ are shown in Fig. 2. Also, the corresponding values of $\hat{J}^\alpha$, the hybrid performance criterion, are shown in Fig. 3. As we see from Figs. 1-3, we have better convergences compared with the alternate algorithm [17]. Thus, we got good convergences to the optimal value shown in Table 1, Table 2 and Table 3.

Table 1 The optimal values of $\Theta = \{\theta_{31}, \theta_{32}, \theta_{21}\}$, $\tilde{\Xi}$, and $\hat{J}^\alpha$ for $\tilde{m} = 3$

		$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
Θ	θ_{31}	1.0354	1.0252	1.0082	0.9957	0.9736	0.9248
	θ_{32}	0.6367	0.6483	0.6569	0.6711	0.6842	0.7249
	θ_{21}	0.7936	0.7813	0.7629	0.7506	0.7302	0.6977
$\tilde{\Xi}$	ξ_1	7.8311	7.7324	7.5951	7.5507	7.5847	7.5718
	ξ_2	0.1473	1.0506	1.8097	2.4754	3.0769	3.6306
	ξ_3	0.0000	0.0000	0.0000	0.5931	1.2878	1.8849
$\hat{J}^\alpha$		11.6058	16.1248	20.2454	24.0061	27.2659	30.1738

Table 2 The optimal values of $\Theta = \{\theta_{31}, \theta_{32}, \theta_{21}\}$, $\tilde{\Xi}$, and $\hat{J}^\alpha$ for $\tilde{m} = 2$

		$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
Θ	θ_{31}	1.0257	1.0153	1.0098	0.9842	0.9552	0.9287
	θ_{32}	0.6381	0.6445	0.6621	0.6733	0.7265	0.7266
	θ_{21}	0.7844	0.7706	0.7629	0.7445	0.7138	0.6946
$\tilde{\Xi}$	ξ_1	7.5633	7.5782	7.5906	7.5928	7.5057	7.5763
	ξ_2	0.1470	1.0491	1.8071	2.4738	3.5617	3.6260
$\hat{J}^\alpha$		11.6016	16.1227	20.2453	24.1329	27.8996	31.4592

Table 3 The optimal values of $\Theta = \{\theta_{31}, \theta_{32}\}$, $\tilde{\Xi}$, and $\hat{J}^\alpha$ for $\tilde{m} = 1$

		$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
Θ	θ_{31}	1.0229	1.0228	1.0228	0.9078	0.9079	0.9154
	θ_{32}	0.6426	0.6774	0.6774	0.7113	0.7114	0.7235
$\tilde{\Xi}$	ξ_1	7.0000	7.0000	7.0000	7.0000	7.0000	7.0000
$\hat{J}^\alpha$		11.6290	16.6033	21.5628	26.5386	31.4614	36.3850

Fig. 1 The results of the optimization of θ_{31} , θ_{32} and θ_{21} for $\tilde{m} = 3$ and $\alpha = 0.6$ Fig. 2 The results of the optimization of ξ_1 , ξ_2 and ξ_3 for $\tilde{m} = 3$ and $\alpha = 0.6$

References

[1] M. Athans: On the determination of optimal costly measurement strategies for linear stochastic systems; *Prep. 5th IFAC World Congress*, pp. 1–11 (1972)

Fig. 3 The result of the optimization by the steepest descent method for $\tilde{m} = 3$ and $\alpha = 0.6$ Fig. 4 Comparison of the performances of estimation and control for $\tilde{m} = 3$

- [2] A. Logothetis, A. Isaksson and E. Skadas: On sensor scheduling via information theoretic criteria; *Prep. ACC, 1999* (1999)
- [3] M. Aoki and M. T. Li: Optimal discrete time control system with cost for observation; *IEEE Trans. Automatic Control*, Vol. AC-14, No. 2, pp. 165–175 (1969)
- [4] A. Sano and M. Terao: Measurement optimization in optimal process control; *Automatica*, Vol. 5, pp. 705–714 (1970)
- [5] L. Meir, J. Peschon and R. M. Dressler: Optimal control of measurement systems; *IEEE Trans. on Automatic Control*, Vol. AC-14, pp. 528–536 (1967)
- [6] V. D. Vanderlinde and A. Lavi: Optimal observation policies in linear stochastic systems; *Prep. Of Joint Automatic Control Conference* (1968)
- [7] S. Ihara: Coding theory in white Gaussian channel with feedback; *J. Multi-variate Analysis*, Vol. 4, pp. 74–87 (1974)
- [8] Y. Takeuchi: The optimal transmission of non-Gaussian signals through a noisy channel with feedback; *IEEE Trans. Inform. Theory*, Vol. IT-40, No. 5, pp. 1624–1629 (1994)
- [9] Y. Takeuchi and A. Ida: Optimal transmission of a set of Gaussian signals through discrete-time chan-

- nels with feedback; *Proc. of 35th ISCIE SSS*, pp. 118–123 (2004)
- [10] Y. Takeuchi: An information theoretic approach to optimization of linear observations for the Kalman-Bucy filter; *International Journal of Innovative Computing, Information & Control*, Vol. 1, No. 3, pp. 401–416 (2005)
- [11] Y. Takeuchi: Optimal transmission of a set of discrete-time Gaussian signals through channels with feedback; *International Journal of Innovative Computing, Information & Control*, Vol. 2, No. 5, pp. 927–942 (2006)
- [12] 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*, Vol. 4, No. 1, pp. 211–230 (2008)
- [13] Y. Takeuchi and A. Hirata: Optimization of linear observations for the stationary Kalman filter based on a generalized water filling theorem; *International Journal of Innovative Computing, Information & Control*, Vol. 5, No. 1, pp. 75–85 (2009)
- [14] Y. Takeuchi and M. Inoue: Optimization of observations for LQG control systems by an information theoretic approach; *International Journal of Innovative Computing, Information & Control*, Vol. 6, No. 1, pp. 75–87 (2010)
- [15] Y. Takeuchi: Optimization of linear observations for the stationary Kalman filter under a quadratic performance criterion; *International Journal of Innovative Computing, Information & Control*, Vol. 7, No. 1, pp. 85–99 (2011)
- [16] Y. Takeuchi: Optimization of the observations for stationary LQG stochastic control systems under a quadratic criterion; *International Journal of Innovative Computing, Information & Control*, Vol. 8, No. 3, pp. 2313–2328 (2012)
- [17] Y. Takeuchi and R. Matsumi: Optimization of the observations for stationary LQG stochastic control systems under a quadratic criterion; *Proceedings of the 43rd ISCIE SSS*, pp. 27–34 (2012)
- [18] S. Kodama and N. Suda: *Matrix Theory for Systems Control* (in Japanese), SICE (1978)
- [19] V. Kucera: A contribution to matrix quadratic equations; *IEEE Trans. Automatic Control*, Vol. AC-17, pp. 344–347 (1972)

Authors

Yoshiki TAKEUCHI (Member)

Yoshiki Takeuchi was born in Kyoto, Japan, on May 10, 1951. He received the Doctor of Engineering degree from Kyoto University in 1980. He is currently a professor in the Department of Information Science, Osaka University of Education (Osaka Kyoiku University), Japan. He has been working in the research field of the state estimation and control of stochastic systems. Especially, his main interest, in recent years, is the optimization of the observation in stochastic systems.

Ryo MATSUMI

Ryo Matsumi was born in Gojo City in Nara Prefecture, Japan, on September 3, 1988. He received the BS degree from Osaka University of Education (Osaka Kyoiku University) in 2011. He is currently a student in the Graduate School, Osaka University of Education, Japan.