Optimization of Observations for LQG Control Systems by an Information Theoretic Approach

Yoshiki TAKEUCHI and Masatoshi INOUE
Dept. of Information Science, Osaka Univ. of Education
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 an optimal selection of the gain matrix of the noisy observation for stationary LQG stochastic control systems. By introducing an information theoretic criterion based on a generalized Water Filling Theorem, we obtain a set of the gains which maximize the mutual information between the system state and the observations. Then, equations are obtained to find the optimal gain matrix from this set of gains which produces the best performance of the optimal LQG regulator.

1 Introduction

In this paper, we consider an optimization problem of observations for the LQG optimal control systems. Since, as it is well-known, the separation theorem holds for this class of systems, the optimal control input is a linear function of the filtered state estimate which is generated by the Kalman filter.

For systems without the control input, we have already considered the optimization of the observation which is used for the Kalman filter by the following two steps [6]:

(i) Information theoretic optimization to maximize the mutual information between the signal and the observation subject to a power constraint concerned with the innovations process;

(ii) Optimization of the Kalman filter's performance, i.e., minimization of the estimation error variance.

In the situation of this paper, however, since we are concerned with the LQG control system, our attention should be paid first to the value of the control performance criterion. Therefore, in this paper, step (ii)' is replaced by

(ii) Optimization of the performance of the LQG Regulator, i.e., minimization of performance criterion of the control.

There does not exist much difference between the mathematical features of problems (i)-(ii)' and (i)-(ii). So, we can obtain the set of gains which is the solution of (i) based on a generalized Water Filling Theorem [6]. We will, however, introduce a new method of solution in (ii). As we will see, (ii) is an optimization problem with variables which are components of an orthogonal matrix. This problem is converted to one without constraint by introducing an n-dimensional polar coordinate system. Also, for this unconstrained problem, we will show the connection rule of the angular parameters which makes us possible to find an interior point with the same value as any exterior point near the boundary of the domain. A numerical example shows this connection rule really works greatly and we can easily get the solution of (i)-(ii).

Mathematical symbols, in this paper, are used in the following way. \( R \) is the space of all real numbers, i.e., \( R = (-\infty, \infty) \). 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 \( x \in R^n \), \( |x| = \sqrt{x^T x} \). 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 \([A]_k\) simply by \( a_k \). 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 non-negative 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 \( E\{\cdot|\mathcal{B}\}, \mathcal{B} \subset \mathcal{F} \) the conditional expectation, given \( \mathcal{B} \), 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 and Preliminaries

2.1 Stationary Optimal LQG Regulator System

Let \( x = \{x_i(\omega); t = 0, 1, \cdots\} \) denote an \( n \)-dimensional Gaussian stochastic process described by

\[
\begin{align*}
x_t(\omega) &= A x_{t-1}(\omega) + C u(t) + G w_t(\omega), \\
x_0(\omega) &= x^0(\omega),
\end{align*}
\]

where \( A \in R^{nxn}, C \in R^{nxr}, G \in R^{nxd}, x^0(\omega) \) is a Gaussian random vector with mean \( \hat{x}^0 \) and covariance \( Q^0 \), \( u = \{u(t); t = 0, 1, \cdots\} \) is a \( r \)-dimensional control
where \( y = \{ y_t(\omega) ; t = 1, 2, \cdots \} \) is an \( m \)-dimensional observation process, \( H \in R^{mxn} \), \( R \in R^{nxn} \), and 
\( v = \{ v_t(\omega) ; t = 1, 2, \cdots \} \) is a \( d_v \)-dimensional standard white Gaussian noise sequence. We will assume the following two conditions throughout this paper.

\[
y_t(\omega) = H x_t(\omega) + R v_t(\omega),
\]
(2)

where \( y = \{ y_t(\omega) ; t = 1, 2, \cdots \} \) is an \( m \)-dimensional observation process, \( H \in R^{mxn} \), \( R \in R^{nxn} \), and 
\( v = \{ v_t(\omega) ; t = 1, 2, \cdots \} \) is a \( d_v \)-dimensional standard white Gaussian noise sequence. We will assume the following two conditions throughout this paper.

\[
(C-1) \quad R_0 \triangleq R R^T > 0, \\
(C-2) \quad x^0(\omega), \ w \text{ and } v \text{ are mutually independent.}
\]

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 \( \mathcal{Y}_t = \{ y_s(\omega) ; s = 1, 2, \cdots, t \} \) is given by the Kalman filter:

\[
\begin{align*}
\hat{x}_{t+1|t}(\omega) &= A \hat{x}_{t|t-1}(\omega) + C u(t - 1) \\
\hat{y}_t(\omega) &= \hat{x}_{t|t}(\omega) + Q^T H^T (H Q H^T + R_0)^{-1} \hat{y}_t(\omega),
\end{align*}
\]
(3)

and

\[
\begin{align*}
Q' &= A Q A' + G G' \\
Q &= Q' - Q^T H^T (H Q H^T + R_0)^{-1} H Q',
\end{align*}
\]
(4)

where

\[
\begin{align*}
\hat{x}_{t+1|t} &\triangleq E[x_t(\omega)|\mathcal{Y}_t], \\
Q' &\triangleq E[(x_t(\omega) - \hat{x}_{t|t}(\omega))(x_t(\omega) - \hat{x}_{t|t}(\omega))^T], \\
Q &\triangleq E[(x_t(\omega) - \hat{x}_{t|t}(\omega))(\hat{x}_t(\omega) - \hat{x}_{t|t}(\omega))^T].
\end{align*}
\]
(5)

Also, \( \bar{y} = \{ \bar{y}_t(\omega) ; t = 1, 2, \cdots \} \) in (3) is the innovations process:

\[
\bar{y}_t(\omega) \triangleq y_t(\omega) - H \hat{x}_{t|t-1}(\omega) = H (x_t(\omega) - \hat{x}_{t|t-1}(\omega)) + R v_t(\omega).
\]
(8)

We are concerned with the stationary optimal LQG control system with (1) and (2), i.e., we determine the control input \( u = \{ u(t) ; t = 0, 1, \cdots \} \) in the way to minimize

\[
\bar{J} \triangleq \lim_{T \to \infty} E \left\{ \frac{1}{T} \sum_{t=1}^{T} E \left[ x_t(\omega)^T M x_t(\omega) + u(t - 1) N u(t - 1) \right] \right\},
\]
(9)

where \( M \in R^{mn} \) and \( N \in R^{nxn} \) 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

\[
u^*(t) \triangleq -(C W C + N)^{-1} C W A \hat{x}_{t|t}(\omega),
\]
(10)

where \( W \in R^{mxn} \) is given by the positive definite solution of the matrix Riccati equation:

\[
\begin{align*}
W &= A' Y A + M \\
Y &= W - W C (C W C + N)^{-1} C W.
\end{align*}
\]
(11)

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

\[
\bar{J}^* = \text{tr}[A'(W - Y)A Q] + \text{tr}[W G G'].
\]
(12)

### 2.2 Optimization of the Observations

In this paper, the optimization of \( H \in R^{mxn} \) is done by the following two steps.

(i) Optimization by an information theoretic criterion:

Find a set, say \( \mathcal{H} \), of the values of \( H \) which maximize the mutual information between \( x \) and \( y \) subject to a power condition concerned with the innovations process.

(ii) Find \( H \in \mathcal{H} \) which minimizes the performance index \( \bar{J}^* \) of the LQG optimal control system (1) and (2) with (10) and (11).

As for (i), the problem in the optimal transmission framework is described as follows [5, 6].

**Problem 1 (Optimal Selection (i) of \( H \))** Find \( H \in R^{mxn} \) such that

\[
I(x, y) = \frac{1}{2} \log \left( \frac{\det[Q H' H + R]}{\det R} \right) \to \text{max.},
\]
(13)

subject to (4) and the power condition:

\[
P \triangleq \text{tr}[H Q H'] \leq p,
\]
(14)

where \( p \) is a maximum admissible value of \( P \).

In the next subsection, we will give some explanations about the optimal transmission problem and how we get the innovations process and Problem 1 in it.

On the other hand, step (ii) is rather simple. Since only \( Q \) is dependent on \( H \) in (12), this part of the problem is that of finding \( H \in \mathcal{H} \) which minimizes

\[
\bar{J}^* \triangleq \text{tr}[A'(W - Y) A Q].
\]
(15)

### 2.3 Optimization of \( H \) from an Information Theoretic Point of View

Fig. 1 Transmission channels with feedback.

From (8), we notice that the innovations process \( \bar{y} \) is described by the same equation which was used in the discussion of the optimal transmission problem [5]. The optimal transmission problem for a set of parallel transmission channels with the construction shown in Fig. 1 is
described as follows. Here, we will also assume that \( x = \{ x(t); t = 1, 2, \ldots \} \) which denotes the signal is an \( n \)-dimensional Gaussian stochastic process given by (1).

In what follows, on the other hand, we temporarily reset the definition of \( \tilde{y} \) to denote the output of the channels. When we send \( x \) through the set of \( m \)-parallel channels, we employ an encoding scheme. Namely, we send an encoded signal \( \tilde{\beta} = \{ \tilde{\beta}(x, \tilde{y}); t = 1, 2, \ldots \} \) which is a function of both signal \( x \) and output \( y \) given by (1).

\[ y(t, \tilde{y}) = \tilde{\beta}(x, \tilde{y}) + Rv(t), \]

where \( \tilde{\beta}(x, \tilde{y}) = Hx - \phi(t, \tilde{y}) \).

We assume \( \phi(t, \tilde{y}) \) is a non-anticipative function of \( \tilde{y} \):

\[ \phi(t, \tilde{y}) = \phi(t, \tilde{y}_1, \ldots, \tilde{y}_m), \]

namely, the feedback component \( \phi(t, \tilde{y}) \) must be a measurable random vector, where

\[ \tilde{y}_i = y(t, \tilde{y}), \quad t = 1, 2, \ldots. \]

Now, the optimal transmission problem is described by

\[ I(x, \tilde{y}) = \frac{1}{2} \log \left\{ \frac{\det |HQH' + R_0|}{\det |R_0|} \right\} \rightarrow \max, \]

subject to the power condition:

\[ P = E \left\{ |\beta(x, \tilde{y})|^2 \right\} \leq p, \]

where \( p \) is the maximum admissible transmission power at each time.

Now, by (17), the optimal transmission problem is that of finding \( \phi(t, \tilde{y}) \) and \( H \) for which we have (21) and (22). By Theorem 2, and by the duality of (21) and (22), the problem is divided into the following two subproblems:

**Problem S1** For a fixed \( H \), find \( \phi(t, \tilde{y}) \) such that

\[ P = E \left\{ |\beta(x, \tilde{y})|^2 \right\} \rightarrow \min. \]  

**Problem S2** For the solution \( \phi(t, \tilde{y}) = \phi' (t, \tilde{y}) \) of Problem S1, find \( H \) such that (21) holds under (22).

As for Problem S1, it can be seen from (22) and (17) that \( \phi(t, \tilde{y}) \) is optimal when

\[ P = E \left\{ |H(t)x(t) - \phi(t, \tilde{y})|^2 \right\}, \]

is minimized. According to the least-squares estimation theory, the optimal selection of \( \phi(t, \tilde{y}) \) is given by the following theorem.

**Theorem 3** The solution of Problem S1 is given by

\[ \phi(t, \tilde{y}) = H'x(t), \]

where \( x(t) \) is the prediction of \( x(t) \) defined by (5). When (25) is used, we have

\[ P = tr(HQH'). \]

If we take \( \phi(t, \tilde{y}) \) by (25), we see from (16) and (17) that the output of the transmission channels is described by (8), i.e., the channel output is now the innovations process. Also, since \( I(x, \tilde{y}) = I(x, y) \) by Theorem 2, it is seen that Problem S2 is reduced to Problem 1 in the previous subsection.

### 3 Optimal Selection of \( H \) by the Solution of Optimal Transmission Problem

As we explained in Section 2.2, the optimization of \( H \) has two steps. For the first step, i.e., for Problem 1, the condition of the optimality is given by the following two theorems [6].

**Theorem 4 (A Generalized Water Filling Theorem)** In addition to (C-1) and (C-2), assume

\[ GG' > 0. \]

Let \( \{ \psi_1, \psi_2, \ldots, \psi_m \} \) and \( \{ y_1, y_2, \ldots, y_m \} \) denote the set of eigen-values of \( R_0 \) and the corresponding set of eigenvectors, respectively. Also, let \( \Psi \triangleq \text{diag}(\psi_1, \psi_2, \ldots, \psi_m) \) and \( \Gamma \triangleq \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_m) \), i.e., \( R_0 = \Gamma \Psi \Gamma' \). Without loss of generality, we can assume that

\[ 0 < \psi_1 \leq \psi_2 \leq \cdots \leq \psi_m. \]

Then, for Problem 1, the condition of optimality of \( H \in R^{m\times n} \) is given by

\[ HQH' = \Gamma \Xi \Gamma', \]

where

\[ \Xi \triangleq \text{diag}(\xi_1, \xi_2, \ldots, \xi_m), \]

\[ \xi_i \triangleq \max \{ 0, \alpha - \psi_i \}, \quad i \in \{1, 2, \ldots, m\} \]
\[ \alpha = \frac{p + \sum_{i=1}^{n} \psi_i}{m}, \]  
(31)

and

\[ m = \text{rank}(\Xi) = \{ \text{Number of positive } \zeta_i, 1 \leq i \leq m \}. \]  
(32)

Furthermore, for any \( H \) given by (28), we have

\[ HQH' = \Xi(\Xi + \Psi)^{-1}\Psi \Gamma'. \]  
(33)

and

\[ I(x, y) = \frac{1}{2} \log \{ \det(\Xi\Psi^{-1} + I) \}. \]  
(34)

Please see the proof and Remarks 1-5 of the above theorem given in [6] since we do not repeat them here again.

«Theorem 5 (Reduction of Riccati Equation)» Assume (C-1)-(C-3). For simplicity, let us denote

\[ \Xi = \begin{bmatrix} \xi & 0 \\ 0 & 0 \end{bmatrix}, \]  
(35)

\[ \Xi = \text{diag}(\xi_1, \xi_2, \cdots, \xi_m), \]  
(36)

and

\[ \Gamma = \begin{bmatrix} \Gamma_1 & \Gamma_2 \\
\end{bmatrix}, \]  
(37)

Then, the solution of Problem 1 is given by the following set of equations.

\[ Q^+ = QA' + GG', \]  
(38)

\[ Q = (Q')^{1/2} K(Q')^{1/2}, \]  
(39)

\[ K = I - \frac{1}{\alpha} U'U', \]  
(40)

\[ H = \Gamma \Xi^{-1/2} U(Q'^{-1/2}, \]  
(41)

where \( U \in \mathbb{R}^{n \times m} \) is any matrix with the property

\[ UU' = I. \]  
(42)

As we can see from (41) and (42) that \( H \) is not unique but we have the set of values:

\[ \mathcal{H} = \{ H \in \mathbb{R}^{n \times m} ; H = \Gamma \Xi^{-1/2} U(Q'^{-1/2}, \]  
(43)

Hence, in order to determine \( H \) uniquely by (ii) in Section 2.2, we have to find a solution of the following problem:

[Problem 2 (Optimal Selection (ii) of \( H \))] Find \( \hat{U} \in \mathbb{R}^{n \times m} \) such that

\[ J^* = \text{tr}[A'(W - Y)A] \rightarrow \min., \]  
(44)

subject to (38)-(40) and (42).

4 Representation of \( \hat{U} \) by an \( n \)-dimensional Polar Coordinate System

In this section, we convert the constrained problem given by (42) and (44) to an unconstrained one by introducing an \( n \)-dimensional polar coordinates system in \( \mathbb{R}^n \). Let us denote

\[ \hat{U} = T(n) \begin{bmatrix} \hat{r} & 0 \\ 0 & T(n - \ell) \end{bmatrix} \begin{bmatrix} I_{\ell \times \ell} & 0 \\ 0 & 0 \end{bmatrix}, \]  
(45)

where

\[ \hat{k} = \min(m - 1, n - 2), \]  
(46)

and, for \( k = n, n - 1, \cdots, n - \hat{k} \),

\[ T(k) = T(k; \theta_{k_1}, \theta_{k_2}, \cdots, \theta_{k_{\ell - 1}}) = [z_{k_1} z_{k_2} \cdots z_{\ell}], \]  
(47)

and

\[ 0 \leq \theta_{ki} \leq 2\pi, \quad -\frac{\pi}{2} \leq \theta_{ki} \leq \frac{\pi}{2}, \quad i = 2, 3, \cdots, k - 1. \]  
(49)

Then, it can be easily seen that \( T(k)T(k) = T(k)T'(k) = I_{k \times k} \) and, hence, we have (42) for \( \hat{U} \) given by (45). The fact that all values in \( \mathbb{R}^{n \times m} \) of \( \hat{U} \) with property (42) is given by (45)-(49) is easily seen as follows. Let \( U(k) \) denote any \( k \times k \) - dimensional orthogonal matrix. Hence, \( U = U(n) \). Note that the first column vector of \( T(k) \), i.e., \( z_{k_1} \), can take any value in \( \mathbb{R}^k \) with \( |z_{k_1}| = 1 \).

Since the set of column vectors of \( \hat{U} \) is an orthogonal system, we have, for \( k = n, \) \( U(n) = [z_{k_1} z_{k_2} \cdots z_m] B \) where \( B \in \mathbb{R}^{n \times (n-1)} \), i.e., the second to \( n \)th column vectors of \( U(n) \) are linear combinations of \( z_{k_1}, \cdots, z_m \). In order to satisfy the property \( U(n)U'(n) = I, \) \( B \) must satisfy \( BB' = I. \) Hence, we have \( B = U(n - 1) \).

Thus, we have converted Problem 2 with constraint (42) to the one with the unconstrained angular variables given by (49) for \( k = n, n - 1, \cdots, n - k \).

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

For simplicity, let

\[ \Theta = \{ \theta_{ki}, \ 2 \leq i \leq n, \ 1 \leq k \leq n - 1 \}. \]  
(50)

Clearly, \( \hat{U} = \hat{U}(\Theta) \) is a periodic function of \( \Theta \). Hence, for an exterior point of the domain given by (49), there
always exists an interior point (of the domain) for which \( J^* \)
has the same value as the exterior point. If we can find
such a pair of values of \( \Theta \), the implementation of
the searching algorithm of finding the minimal point of \( J^* \)
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:

\[
0 \leq \theta_{kl} \leq \pi, \quad -\frac{\pi}{2} \leq \theta_{kl} \leq \frac{\pi}{2}, \quad i = 2, 3, \ldots, k - 1, \quad k = n, n - 1, \ldots, n - k, \tag{51}
\]

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

**Theorem 6 (Connections of Angular Parameters)**

For \( \hat{U} \) given by (45), we have the following rules for the
case when only one parameter violates (51). For

\[
k = n, n - 1, \ldots, k, \tag{52}
\]

and assume that all \( \theta_{kl}, k \neq q \) satisfy (51). Then,

(i) When \( \theta_{kl} \in \left[ -\pi, 0 \right) \) or \( \theta_{kl} \in (\pi, 2\pi] \) and \( \theta_{ql} \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right], i = 2, 3, \ldots, q - 1 \), the value of \( J^* \) is equal to the one at the point by the following substitutions.

(a) \( q : \) odd,

\[
\begin{align*}
\theta_{kl} & \to \theta_{kl} \pm \pi, \quad \theta_{ql} \to -\theta_{ql}, i = 2, 3, \ldots, q - 1, \\
\theta_{(2j-1,2j)} & \to -\theta_{(2j-1,2j)}, \\
\theta_{(2j,2j+1)} & \to \pi - \theta_{(2j-1,2j)}, \quad j = 2, 3, \ldots, (q - 1)/2, \\
\theta_{21} & \to \pi - \theta_{21}.
\end{align*}
\]

(b) \( q : \) even,

\[
\begin{align*}
\theta_{kl} & \to \theta_{kl} \pm \pi, \quad \theta_{ql} \to -\theta_{ql}, i = 2, 3, \ldots, q - 1, \\
\theta_{(2j-1,2j)} & \to -\theta_{(2j-1,2j)}, \\
\theta_{(2j,2j+1)} & \to \pi - \theta_{(2j-1,2j)}, \quad j = 1, 2, \ldots, (q - 2)/2.
\end{align*}
\]

(ii) When \( \theta_{ql,1} \in \left[ -\frac{\pi}{2}, -\frac{\pi}{4} \right] \) or \( \theta_{ql,1} \in \left( \frac{\pi}{4}, \frac{\pi}{2} \right], \) and \( \theta_{ql,1} \in [0, \pi] \) and \( \theta_{ql,1} \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right], i = 2, 3, \ldots, q - 2, \) the value of \( J^* \) is equal to the one at the point by the following substitutions.

(a) \( q : \) odd,

\[
\begin{align*}
\theta_{ql,1} & \to \theta_{ql,1} \pm \pi, \\
\theta_{(2j,2j-1)} & \to \pi - \theta_{(2j,2j-1)}, \quad j = 1, 2, \ldots, (q - 1)/2, \\
\theta_{(2j-1,2j-2,2j-1)} & \to \theta_{(2j-1,2j-2,2j)}, \quad j = 2, 3, \ldots, (q - 1)/2.
\end{align*}
\]

(b) \( q : \) even,

\[
\begin{align*}
\theta_{ql,1} & \to \theta_{ql,1} \pm \pi, \\
\theta_{(2j-1,2j)} & \to \pi - \theta_{(2j-1,2j)}, \quad j = 1, 2, \ldots, (q - 2)/2, \\
\theta_{(2j,2j+1)} & \to \pi - \theta_{(2j,2j+1)}, \quad j = 1, 2, \ldots, (q - 2)/2, \\
\theta_{1} & \to \pi - \theta_{1}.
\end{align*}
\]

(iii) When \( \theta_{ql} \in [0, \pi] \) or \( \theta_{ql} \in \left( \frac{\pi}{2}, \frac{3\pi}{2} \right], \) \( \tau = 2, 3, \ldots, q - 1, \) and \( \theta_{ql} \in [0, \pi] \) and \( \theta_{ql} \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right], \)
\( i \neq \tau, i = 2, 3, \ldots, q - 1, \) the value of \( J^* \) is equal to
the one at the point by the following substitutions.

(a) \( 2\tau < q, \)

\[
\begin{align*}
\theta_{ql} & \to \theta_{ql} \pm \pi, \quad \theta_{ql} \to -\theta_{ql}, i = \tau + 1, \tau + 2, \ldots, q - 1, \\
\theta_{(2\tau-1,2\tau)} & \to -\theta_{(2\tau-1,2\tau)}, \\
\theta_{(2\tau,2\tau+1)} & \to \pi - \theta_{(2\tau-1,2\tau)}, \quad \theta_{(2\tau+1,2\tau+2)} = -\theta_{(2\tau+1,2\tau+2)}.
\end{align*}
\]

(b) \( 2\tau \geq q, \)

\[
\begin{align*}
\theta_{ql} & \to \theta_{ql} \pm \pi, \quad \theta_{ql} \to -\theta_{ql}, i = \tau + 1, \tau + 2, \ldots, q - 1, \\
\theta_{(2\tau-1,2\tau)} & \to -\theta_{(2\tau-1,2\tau)}, \\
\theta_{(2\tau,2\tau+1)} & \to \pi - \theta_{(2\tau-1,2\tau)}, \quad \theta_{(2\tau+1,2\tau+2)} = -\theta_{(2\tau+1,2\tau+2)}.
\end{align*}
\]

(Proof) It is easily seen that the replacements (i)-(iii)
produce the same value of \( \hat{U} \) except for two column vectors
with the opposite sign. Since \( K \) in (42) is a quadratic
function of \( \hat{U} \), we have the same value of \( K \) by the
replacements (i)-(iii). This completes the proof.

6 A Numerical Example

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

**Example 1.** Let us consider a 3-dimensional system with

\[
A = \begin{bmatrix}
0.5 & 0.3 & 0.1 \\
0.2 & 0.4 & 0.2 \\
0.1 & 0.5 & 0.6
\end{bmatrix}, \quad G = \begin{bmatrix}
0.5 & 0 & 0 \\
0 & 0.5 & 0 \\
0 & 0 & 0.5
\end{bmatrix},
\]

\[
C = \begin{bmatrix}
1.2 & 0.1 & 0.2 \\
0.2 & 1.4 & 0.2 \\
0.1 & 0.5 & 1.1
\end{bmatrix}, \quad M = I, \quad N = I.
\]

\[
R = \begin{bmatrix}
\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\
-\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\
\frac{1}{4} & -\frac{1}{4} & \frac{1}{4}
\end{bmatrix}, \quad \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = 0.049, \quad 0.100, \quad 0.444.
\]

\[
= \begin{bmatrix}
0.1015 & -0.2128 \\
0.8673 & -0.3437 \\
-0.3437 & 0.8231
\end{bmatrix}
\]

i.e.,
\[ \Psi = \text{diag}(0.49, 1.0, 1.44), \quad \Gamma = \begin{bmatrix} \frac{1}{1.4} & -\frac{1}{1.4} & \frac{1}{1.4} \\ -\frac{1}{1.4} & \frac{1}{1.4} & -\frac{1}{1.4} \\ \frac{1}{1.4} & -\frac{1}{1.4} & \frac{1}{1.4} \end{bmatrix}. \]

For the above system and observations, we made numerical computations for three cases of the value of \( p \), i.e., \( p = 2.0, \quad p = 1.2 \) and \( p = 0.4 \). For these cases, the values of \( \tilde{m}, \alpha, \tilde{\Psi}, \tilde{\Xi} \) and \( I(x, y) \) are summarized in Table 1.

Table 1. The values of \( \tilde{m}, \alpha, \tilde{\Psi}, \tilde{\Xi} \) and \( I(x, y) \) for \( p = 2.0, \quad p = 1.2 \) and \( p = 0.4 \).

<table>
<thead>
<tr>
<th>( p )</th>
<th>( \tilde{m} )</th>
<th>( \alpha )</th>
<th>( \tilde{\Psi} )</th>
<th>( \tilde{\Xi} )</th>
<th>( I(x, y) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>1.643</td>
<td>1.44</td>
<td>diag(0.49, 1.0, 1.44)</td>
<td>diag(1.15, 0.64, 0.20)</td>
<td>0.9194</td>
</tr>
<tr>
<td>1.2</td>
<td>1.345</td>
<td>1.0</td>
<td>diag(0.49, 1.0)</td>
<td>diag(0.855, 0.345)</td>
<td>0.6531</td>
</tr>
<tr>
<td>0.4</td>
<td>0.890</td>
<td>0.49</td>
<td>0.40</td>
<td>0.2984</td>
<td></td>
</tr>
</tbody>
</table>

Here we should note that for \( p = 0.4 \), we have \( \tilde{m} = 1 \) and \( \xi_2 = \xi_3 = 0 \) since \( \psi_1 = 1.44 > 0.89 = \alpha \) and \( \psi_2 = 1.0 > 0.89 = \alpha \). This implies that we cannot assign, due to the constraint \( p = 0.4 \), any energy for the second and third eigen-vector components in \( \Gamma \) of \( HQ'H' \) because these components are very much corrupted by the observation noise \( (\psi_1 = 1.44, \psi_2 = 1.0) \). Thus, for \( p = 0.4 \), \( \tilde{U} \in \mathbb{R}^{3\times 1} \), namely, \( \tilde{U} \) is a 3-dimensional vector, and the degree of freedom of \( H \in \mathbb{R}^{3\times 3} \) is restricted to that of a 3-dimensional vector, i.e., \( \text{rank}[H] = 1 \). In the same way, for \( p = 1.2 \), we have \( \tilde{m} = 2 \) and \( \xi_3 = 0 \) since \( \psi_1 = 1.44 > 1.345 = \alpha \).

We have carried out the optimization by a simple alternate search algorithm with respect to the angular parameters: \( \theta_{31}, \theta_{32} \) and \( \theta_{33} \) by making use of the connection rule developed in Section 5. The result are summarized in Table 2.

Table 2. The result of alternate search algorithm for \( p = 2.0, \quad p = 1.2 \) and \( p = 0.4 \).

<table>
<thead>
<tr>
<th>( p )</th>
<th>( \tilde{m} )</th>
<th>( \tilde{J}^* )</th>
<th>( \theta_{31} )</th>
<th>( \theta_{32} )</th>
<th>( \theta_{33} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>3</td>
<td>0.1891</td>
<td>1.0083</td>
<td>0.6257</td>
<td>0.8053</td>
</tr>
<tr>
<td>1.2</td>
<td>3</td>
<td>0.2501</td>
<td>0.9991</td>
<td>0.6340</td>
<td>0.7973</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>0.4882</td>
<td>0.9700</td>
<td>0.6613</td>
<td></td>
</tr>
</tbody>
</table>

In Fig. 2, the results of 16 runs of the alternate search are shown for \( p = 2.0 \) (\( \tilde{m} = 3 \)). We got convergence for all runs to the optimal value shown in Table 2. In order to help the understanding of the results shown in Fig. 2, the plot of \( \min_{\theta_{31}} \tilde{J}^* \) for all \( 0 \leq \theta_{31} \leq 2\pi \) and \( -\pi/2 \leq \theta_{32} \leq \pi/2 \) is shown in Fig. 3. From these results, we see that we can easily compute the optimal gain matrix and also easily improve the performance of the stationary LQG regulator.

References


