1999 European Control Conference (ECC)
31 August – 3 September 1999, Karlsruhe, Germany

RECURSIVE-TYPE MODEL-FREE LQ CONTROLLER DESIGN

Y. Kawamura

Faculty of Engineering,
Osaka Electro-Communication University
Fax : +81-720-24-0014 and e-mail : kawamura@isc.osakac.ac.jp

Keywords : LQ control, Riccati Equation, Model-free design, Self-tuning regulator.

Abstract

A new type model-free data-based iterative controller design of the LQ regulator is introduced in this paper. This iterative approach is just corresponding to an extension of the model-based Riccati difference equation, and it is directly constructs the regulator from input-output data of an unknown system without requiring the system model. The result is extended to the model-free design of dynamical output feedback regulators as well.

1 Introduction

The modern control theory has developed on the basis of the system model. For example, the LQ regulator has been designed by making use of the Riccati equation or the transfer function given by system models even for adaptive design. However it often happens that a designed control system has poor performance because of unavoidable modeling error caused by nonlinearity and others of actual systems. Recently, some ideas have been proposed on model-free design of the LQ or LQG regulator [1]-[10]. These approaches may have possibilities of reducing difficulties related to modeling error because they directly optimize the feedback controller from input-output data without resorting to the process of system identification. However, since they obtain the control on the basis of long drawn-out responses, they seems extremely sensitive with respect to irregularity included in actual responses.

A new type (recursive-type) model-free controller design of the LQ regulator is introduced in this paper, which is closely related to the non-recursive model-free iterative (self-tuning) design [1]-[3]. The non-recursive algorithm is a kind of Newton method where the state feedback gain G is iteratively updated by

$$\Delta G = -(\Gamma_k^{bb})^{-1}\,\Gamma_k^{ba}$$

where Γ_k^{ba} and Γ_k^{bb} denote the first and the second sensitivity matrices of the LQ performance in the kth $(0,1,2,\cdots)$ iteration with respect to the gain . It is also shown that each element of the sensitivity matrices is given by an inner product of responses. For example, $\Gamma_k^{ba} =< \boldsymbol{z}^a,\ \boldsymbol{z}^b >$ and $\Gamma_k^{bb} =< \boldsymbol{z}^b,\ \boldsymbol{z}^b >$ for the SISO system where $\boldsymbol{z}^a$ and $\boldsymbol{z}^b$ denote the closed-loop responses of the unit initial state and the unit impulse respectively.

It is well known that the optimal control problem (multi-stage problem) is resolved into a sequence of one-stage optimization problems. The new model-free algorithm is based on this idea so that we may overcome the defect of the model-free design. The recursive form is just corresponding to the model-based Riccati difference equation. In addition, a statistical data processing is introduced which is corresponding to an extension of the Riccati equation.

The algorithm is first explained for the basic state feedback, and it is extended to the dynamical output feedback problem. The latter is based on the extended non-minimal state space model representation [3], so that it is easily applicable to unknown systems without using the conventional state observer.

2 Recursive algorithm

First we study the case that the state is directly measurable. Let $x(t) \in R^n$ be the state and $u(t) \in R^m$ be the control defined on $t = 0, 1, 2, \cdots$. Let

$$z(t) \quad = \quad \begin{pmatrix} Cx(t) \\ Du(t) \end{pmatrix} \tag{1}$$

be the controlled output. Note that $z(t)^T z(t) = x(t)^T C^T Cx(t) + u(t)^T D^T Du(t)$. We study the infinite-horizon LQ control which minimizes the standard LQ performance index

$$J = \sum_{t=0}^{\infty} z(t)^T z(t). \tag{2}$$

It is well known that the infinite-horizon LQ optimal control is given by a suitable time-invariant state feedback $u(t) = Gx(t)$. Therefore, the problem is to optimize the gain G.

ISBN 978-3-9524173-5-5

The recursive model-free algorithm is given as follows: Let $T_k = \{t_k, t_k{+}1, t_k{+}2, \cdots, t_k{+}N_k{-}1\}$, $(k = 0, 1, 2, \cdots)$ denote a sequence of periods such that $t_{k+1} \geq t_k + N_k$. Suppose that a matrix P_k, H_k and a gain G_k are given at the beginning of the T_k. Observe response signals of the closed-loop system by applying the control input

$$u(t) = G_k x(t) + v(t) \tag{3}$$

on the period T_k where $v(t)$ is an additional signal (white noise is typical). The recursive algorithm is given as follows:

$$\begin{aligned} P_{k+1} &= \Gamma_k^{aa} - \Gamma_k^{ab} \, (\Gamma_k^{bb} + H_k)^{-1} \\ &\quad * (\Gamma_k^{bb} + 2H_k) \, (\Gamma_k^{bb} + H_k)^{-1} \, \Gamma_k^{ba}, \qquad (4) \\ G_{k+1} &= G_k - (\Gamma_k^{bb} + H_k)^{-1} \, \Gamma_k^{ba}. \qquad (5) \end{aligned}$$

The first equation is related to the Riccati equation and the second is an extension of the above Newton method. The matrices Γ_k^{aa}, $\cdots$, Γ_k^{aa} are sub-matrices of the $(n+m, n+m)$ non-negative matrix

$$\Gamma_k = \begin{pmatrix} \Gamma_k^{aa} & \Gamma_k^{ab} \\ \Gamma_k^{ba} & \Gamma_k^{bb} \end{pmatrix}, \tag{6}$$

which is called the fundamental sensitivity matrix [11].

In this recursive algorithm, the sensitivity matrix is given by the above response data as

$$\begin{aligned} \Gamma_k &= \Xi_k^{-1} \Bigg[\sum_{t \in T_k} \sum_{s \in T_k} \xi(t) \{ z(t)^T z(s) \\ &\quad + x(t+1)^T P_k x(s+1) \} \xi(s)^T \Bigg] \, \Xi_k^{-1}, \qquad (7) \end{aligned}$$

$$\Xi_k = \sum_{t \in T_k} \xi(t) \, \xi(t)^T + \delta I,$$

$$\xi(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}$$

where δ is a positive number and I is the unit matrix. The supplementary non-negative (m, m) matrix H_k is given recursively as

$$\begin{aligned} H_{k+1} &= \lambda_{1k} \, H_k + \lambda_{2k} \, \Gamma_k^{bb}, \qquad (8) \\ 0 &\leq \lambda_{1k} \leq 1, \\ 0 &\leq \lambda_{2k} \leq 1. \end{aligned}$$

Note that (7) contains no system parameters. The initial conditions are

$$P_0 \geq 0, \qquad H_0 \geq 0. \tag{9}$$

for this recursive relation. The initial gain G_0 need not be a stabilizing gain. Suppose that $P_0 \geq 0$, G_0 and $H_0 \geq 0$ are given for an unknown system. Then (3) - (8) give Γ_k, P_{k+1}, G_{k+1} and H_{k+1} ($k = 0, 1, 2, \cdots$ in order) at the end of T_k. We refer to this model-free algorithm as the recursive type. The order of the calculation of Γ_k becomes $O(n+m)$ in each t by using the matrix inversion lemma [11].

3 Relation to the model-based design

Although this algorithm is developed for us to overcome certain range of nonlinearity and others, first we study the case of the ideal linear system:

$$x(t+1) = Ax(t) + Bu(t). \tag{10}$$

It is well known that the LQ optimal gain is given by the limit $(k \to \infty)$ of the matrix Riccati difference equation and the related gain equation:

$$\begin{aligned} P_{k+1} &= A^T P_k A + C^T C \\ &\quad - A^T P_k B \, (B^T P_k B + D^T D)^{-1} B^T P_k A, \qquad (11) \\ G_{k+1} &= -(B^T P_k B + D^T D)^{-1} B^T P_k A \qquad (12) \end{aligned}$$

with the terminal condition $P_0 \geq 0$ under the quite loose condition of stabilizability and detectability. The model-free algorithm is related to these equations as follows;

Theorem 1. Suppose that $N_k \geq m + n$ and δ is a sufficiently small positive number. Then (3) - (9) are equivalent to the matrix equations:

$$\begin{aligned} P_{k+1} &= (A + BG_{k+1})^T P_k (A + BG_{k+1}) \\ &\quad + C^T C + G_{k+1}^T D^T D G_{k+1}, \qquad (13) \\ G_{k+1} &= G_k - \{ B^T P_k B + D^T D + H_k \}^{-1} \\ &\quad * \{ B^T P_k (A + BG_k) + D^T D G_k \}. \qquad (14) \end{aligned}$$

(Proof) Define the sub-matrices of Γ_k by

$$\begin{aligned} \Gamma_k^{aa} &= (A + BG_k)^T P_k (A + BG_k) + C^T C \\ &\quad + G_k^T D^T D G_k, \qquad (15) \\ \Gamma_k^{ab} &= (A + BG_k)^T P_k B + G_k^T D^T D, \\ \Gamma_k^{bb} &= B^T P_k B + D^T D. \end{aligned}$$

Then we can rewrite (4), which contains G_{k+1}, to

$$\begin{aligned} P_{k+1} &= \Gamma_k^{aa} + \Gamma_k^{ab} \, \Delta G_k + \Delta G_k^T \, \Gamma_k^{ba} \\ &\quad + \Delta G_k^T \, \Gamma_k^{bb} \, \Delta G_k, \qquad (16) \\ \Delta G_k &= G_{k+1} - G_k. \end{aligned}$$

Substitute (15) into (6) and substitute the resultant ΔG_k into (16). Then we obtain (4) and (5).

On the other hand, (15) is equivalent to

$$\begin{aligned} \Gamma_k &= \begin{pmatrix} (A + BG_k)^T \\ B^T \end{pmatrix} P_k \, (\, A + BG_k \quad B \,) \\ &\quad + \begin{pmatrix} C^T & G_k^T D^T \\ 0 & D^T \end{pmatrix} \begin{pmatrix} C & 0 \\ DG_k & D \end{pmatrix}. \qquad (17) \end{aligned}$$

Multiply (17) by $\xi(t)\xi(t)^T$ from the left side and $\xi(s)\xi(s)^T$ from the right side respectively. Note that (10), (1) (3) and the definition of $\xi(t)$ imply

$$x(t+1) = (A + BG_k \;\; B)\,\xi(t), \tag{18}$$

$$z(t) = \begin{pmatrix} C & 0 \\ DG_k & D \end{pmatrix} \xi(t). \tag{19}$$

Substitute these relations into $\xi(t)\xi(t)^T \Gamma_k \, \xi(s)\xi(s)^T$. It follows that

$$\xi(t)\xi(t)^T \Gamma_k \, \xi(s)\xi(s)^T = \xi(t)\{x(t+1)^T P_k \, x(s+1) + z(t)^T z(s)\}\xi(s)^T. \tag{20}$$

Sum up this equation with respect to $t \in T_k$ and $s \in T_k$, and multiply it by Ξ_k^{-1} from both sides. Then we obtain (7).

4 Convergence property of the gain

Suppose the special case that $H_0 = 0$ and $\lambda_{1k} \equiv \lambda_{2k} \equiv 0$. Then we can easily show that (13) and (14) are equivalent to the matrix Riccati difference equation and the related LQ gain equation. This relation guarantees, under the above special case, that P_k and G_k converge to the non-negative solution P^* of the algebraic Riccati equation and the related time-invariant LQ optimal gain G^* respectively. Then we can perform the LQ control by removing $v(t)$ after we optimize the feedback gain. We can extend this result to the general case.

Theorem 2.[12] Suppose the standard condition: (A, B) is stabilizable, (C, A) is detectable and $D^T D > 0$ in addition the assumption of Theorem 1. Then G_k given by the iterative algorithm (3) - (9) converges to G^* as $k \to \infty$

According to this result, the convergent gain is identical regardless of λ_{1k}, λ_{2k} and H_0, In addition, the gain is identical with what is given by the standard model-based LQ design. However, these results become different if the system is not ideal because of nonlinearity and others. As a countermeasure, the standard model-based design contains a statistical data processing in the system identification. On the other hand, the basic model-free algorithm with $H_k \equiv 0$ had no statistical data processing [11]. As a result, its convergence ability was poor for practical use. The matrix H_k is introduced in order to deal with non-ideal systems.

Note that (4) are equivalent to

$$P_{k+1} = \Gamma_k^{aa} + \Gamma_k^{ab}\,\Delta G_k + \Delta G_k^T\,\Gamma_k^{ba} + \Delta G_k^T\,\Gamma_k^{bb}\,\Delta G_k \tag{21}$$

where $\Delta G_k = G_{k+1} - G_k$. We know that this relation is the Taylor expansion of P_{k+1} at the gain G_k with respect to gain change ΔG_k. Therefore, Γ_k^{aa}, Γ_k^{ab} and Γ_k^{bb} denote the 0th, 1st and 2nd sensitivities of P_{k+1} with respect to the gain change, respectively. Suppose $H_k = 0$. Then (5) is a kind of Newton method to optimize the gain, and P_{k+1} is optimized provided that the system is exactly linear. If the system has nonlinearity, the Taylor expansion has higher-order terms. Then P_{k+1} is not optimized and the sequence P_k often diverges.

It seems important to apply the idea of reducing gain change in stochastic gradient methods. The non-negative matrix H_k plays this role. Therefore, the convergence property of P_k is different from the standard Riccati difference equation. The introduction of H_k and the condition (8) are given on the analogy of Landau's adaptive scheme [13]. Note also that (13) is different from the usual system error equation. .

5 Dynamical output feedback

The state feedback is assumed in the above basic model-free LQ control design. We extend the result to the output feedback problem. The standard approach to the problem is based on the separation into state observation and state feedback. However the separation makes the problem rather difficult for unknown systems from the viewpoint of the model-free design. Here we apply the idea of the non-minimal state-space model [3].

Let S be the system given by (10)

$$y(t) = C' x(t) \tag{22}$$

where $y(t) \in R^r$ is a measurable output. Suppose that $z(t)$ is composed of measurable signals as

$$z(t) = \begin{pmatrix} C" y(t) \\ Du(t) \end{pmatrix}. \tag{23}$$

It follows that $C = C"C'$. The detectability of (C, A) is satisfied provided that (C', A) is detectable and $C"$ is column full-rank.

The input-output relation of S is

$$\begin{aligned} y(t) &= a_1 y(t-1) + a_2 y(t-2) + \cdots + a_{\bar{n}} y(t-\bar{n}) \\ &+ b_1 u(t-1) + b_2 u(t-2) + \cdots + b_{\bar{n}} u(t-\bar{n}) \end{aligned} \tag{24}$$

where $\bar{n}$ is a suitable integer. In fact this relation is familiar for a SISO system as $\bar{n} = n$.

Let $\bar{S}$ be the $m(\bar{n}-1) + r\bar{n}$th order extended state space model:

$$\begin{aligned} X(t+1) &= \bar{A}X(t) + \bar{B}u(t)\;, \\ y(t) &= \bar{C}'X(t)\;, \end{aligned} \tag{25}$$

$$\bar{z}(t) = \begin{pmatrix} \bar{C}X(t) \\ \bar{D}u(t) \end{pmatrix},$$

where

$$\bar{A} = \begin{pmatrix} a_1 & \dots & a_{\bar{n}} & b_2 & \dots & b_{\bar{n}} \\ & \boldsymbol{I}_{\bar{n}-1} & 0 & & \mathbf{0} & \\ 0 & \dots & 0 & 0 & \dots & 0 \\ & \mathbf{0} & & & \boldsymbol{I}_{\bar{n}-2} & 0 \end{pmatrix}, \quad (26)$$

$$\begin{aligned} \bar{B} &= (b_1 \ 0 \ \dots \ 0 \ I \ 0 \ \dots \ 0)^T, \\ \bar{C}' &= (I \ 0 \ \dots \ 0 \ 0 \ 0 \ \dots \ 0), \\ \bar{C} &= C"\bar{C}', \\ \bar{D} &= D \end{aligned}$$

and the state variable is

$$\begin{aligned} X(t) = \big(& y(t)^T \ \dots \ y(t-\bar{n}+1)^T \\ & u(t-1)^T \ \dots \ u(t-\bar{n}+1)^T \big)^T. \quad (27) \end{aligned}$$

In the equations, $\boldsymbol{I}_{\bar{n}-1}$ is the unit matrix with $(\bar{n}-1)$ diagonal blocks and $\bar{B}$ has the unit matrix as the $(\bar{n}+1)$th block. This is formally a state equation, and the following relations hold though it is not a minimal realizing system.

Theorem 3 Suppose that (A, B) is stabilizable, (C, A) is detectable and $D^T D > 0$ for the original system. Suppose $\bar{n}$ is the observability index of the observable sub-system of S or over. Then $\bar{S}$ satisfies the following relations [3]:

(1) S and $\bar{S}$ have the same input-output $(\boldsymbol{u}, \boldsymbol{y})$ relation provided that $X(0)$ is given at $t = 0$.

(2) $X(t) \to 0$ $(t \to \infty)$ implies $x(t) \to 0$.

(3) $(\bar{A}, \bar{B})$ is stabilizable, $(\bar{C}, \bar{A})$ is detectable and $\bar{D}^T \bar{D} > 0$.

The relation (3) guarantees the existence of the LQ control of the extended state space model $\bar{S}$. In addition, (2) guarantees that this LQ control stabilizes the original system S as well. This state feedback of $\bar{S}$ means a dynamical output feedback of S. We refer to this control as the LQ output-feedback control.

If $X(0)$ is given at $t = 0$, the recursive model-free LQ control design algorithm is applicable to $\bar{S}$ by replacing $x(t)$, G_k, n and $\boldsymbol{z}$ with $X(t)$, $\bar{G}_k$, $2\bar{n}-1$ and $\bar{\boldsymbol{z}}$ respectively. Data on A, B, C' (or a_i, b_i) and $x(t)$ are unnecessary for this calculation. If $X(0)$ is unknown, we can replace it with an estimate $\hat{X}(t)$. This control is related to the standard approach with a state observer as follows [3].

Theorem 4. Suppose the same assumption as Lemma 1. Then the LQ output-feedback control is equivalent to the combination of the standard LQ control with dead-beat state observation, namely

$$\bar{G}^* X(t) = G^* x(t)$$

holds after $X(t)$ is known.

Note that $X(t)$ becomes known at $t \geq \bar{n}-1$ even if $X(0)$ is unknown. The LQ output-feedback control agrees with the standard LQ control after this.

5.1 Simulation results

In order to test the convergence property of the recursive model-free algorithm, we applied it to the following system with irregular responses. Consider the system

$$\begin{aligned} \begin{pmatrix} x_1(t+1) \\ x_2(t+1) \end{pmatrix} &= \begin{pmatrix} a_{11} & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0 \end{pmatrix} u(t), \\ z(t) &= \begin{pmatrix} \frac{\sqrt{2}}{2} x(t) \\ u(t) \end{pmatrix} \end{aligned}$$

We express the irregularity of the system by the irregular time-varying parameter

$$a_{11} = 1.2 + r \sin(t/10) w(t)$$

where r is a non-negative number and $w(t)$ is the uniformly distributed random noise on the range $[0, 1]$ with the expectation. The system is ideal only if $r = 0$.

The algorithm is applied by assuming that the system equation is unknown though $x(t)$, $u(t)$ $v(t)$ and $z(t)$ are available. A zero-mean white noise with the unit variance is used as $v(t)$. The initial parameters are chosen as $P_0 = 0$, $H_0 = 0$, and $G_0 = 0$ for all cases. Therefor the initial system is unstable. The positive number δ is 10^{-6}. The period T_k is $\{3k,\ 3k+1\ 3k+2\}$ or $\{4k,\ 4k+1,\ 4k+2,\ 4k+3\}$. The initial state is $x(0) = 0$. The state is reset to $x(t+1) = 0$ provided $||x(t)|| > 10$ in order to avoid divergence of the state. In addition, P_k is restricted so that $\{rmtrace\}(P_k) \leq 10000$.

Fig. 1 and Fig. 2 show the change of the gain elements g_1 and g_2. where $\lambda_{1k} \equiv \lambda_{2k} \equiv 0$ for the former and $\lambda_{1k} \equiv 0.995$, $\lambda_{2k} \equiv 1$ for the latter. The solid lines express the change of the first element g_1 of the gain and and the broken lines express that of the second element g_2. Both of them converge to the LQ optimal values $g_1^* = -2$ and $g_2^* = 0$, respectively though their convergence rate are different.

Fig.3 shows the change of gain for $r = 0.5$ where the other parameters are the same as Fig.1. The gain does not converge and the basic Newton method with $H_k \equiv 0$ can not give the stabilizing gain. On the other hand, The gain converges in Fig.4 - Fig.6. where $r = 0.1,\ 0.5,\ 1.5$ respectively. The other parameters are the same as Fig.2. The convergent gain of Fig.4 is almost the same as the optimal gain. Fig.7 and Fig.8 show the change of $q(t) = z(t)^T z(t)$ for the above cases r=0.5 and r=1.5 respectively. Therefore can see that stabilizing gains are obtained in these cases.

Fig. 9 shows the result of the output feedback problem:

$$y(t) = 1.5 x_2(t)$$

is measurable in stead of the state of the above example (r=0). This is the same example as was discussed in [3]. The non-minimal model $\bar{S}$ has the state

$$X(t) = \big(y(t) \ \ y(t-1) \ \ u(t-1) \big) \quad (28)$$

by letting $\bar{n} = 2$. The solid line, the broken line and the dotted line express the gain elements of $y(t)$, $y(t-1)$ and $u(t-1)$ respectively. Other parameters of the algorithm is the same as Fig.2. We see that the gain converges to the optimal gain $\bar{G}^* = (-1.6 \;\; 0 \;\; -0.4)$.

k=4t