Chapter 7

An Approach to Hybrid Smoothing for Linear Discrete-Time Systems with Non-Gaussian Noises

Gou Nakura

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51385

1. Introduction

It is very important to consider simultaneous estimation of both system states and inaccessible modes for hybrid systems with unknown modes [4,7,24,25]. This estimation is called hybrid estimation. By the hybrid estimation we often want to know both a current mode and system state at each time through information of observation. However there exist cases that we want to know distributions of modes on long run time interval rather than each estimate of the modes themselves at each time to grasp global performance over long time intervals, for example, distributions of active modes in solar systems [5,21], distributions of active agents on formation or consensus via hybrid systems representation and so on.

Much work has been done for smoothing theory for both of continuous- and discrete-time systems ([1,2,3,6,8,9,10,12,13,14,15,18,19,20,22,23] and so on). Various researchers have studied the smoothing problems by various approach, for example, maximum likelihood approach [9,13,19], projection approach [14] and so on. It is well known that smoothers (noncausal estimators) more effectively estimates the states than filters (causal estimators) because of more information of observation.

It is well known that utilization of accumulated information of observation improves estimation performance. Nevertheless, on research of estimation for hybrid systems, little work has been done from the point of view of the noncausal information of observation, i.e., smoothing. In [9] Helmick et al. have presented a fixed-interval smoothing algorithm for discrete-time Markovian jump systems by maximum likelihood (ML) approach. However they have considered only the case with fully accessible modes and their approach is based on approximate approach to probability density functions (PDFs). Therefore they have presented only a nearly optimal smoothing algorithm. While it is significant that optimality is guaranteed for estimation algorithms, in [4] and [7] Costa et al. have presented LMMSE (linear minimum mean square estimate) filters to estimate both system states and inaccessible modes for continuous- and discrete-time Markovian jump systems affected by wide sense white noises, but in these LMMSE filters theory the optimality of estimation isn''t always guaranteed in the meaning that these filters aren''t always MMSE (minimum mean square estimate). To the best of the author''s knowledge the optimal smoothing problems in the cases with inaccessible modes have not yet fully investigated.

In this chapter we study hybrid estimation for linear discrete-time systems with non-Gaussian noises. The concerned systems are general hybrid systems given below which aren''t restricted to Markovian jump systems [4,5,7] and where added noises aren''t restricted to be Gaussian. It is assumed that modes of the systems are not directly accessible throughout this paper. We consider optimal estimation problems to find both estimated states of the systems and an optimal candidate of the distributions of the modes over the finite time interval. We adopt most probable trajectory (MPT) approach to guarantee the optimality of estimation methods. On this approach, given information of observation, we consider optimal control problems where we seek optimal control by which averaged noises energies are minimized for averaged systems throughout the mode distributions. In [24,25] Zhang has presented hybrid filtering algorithm by MPT approach for the continuous- and discrete-time hybrid systems. We consider both filtering and smoothing problems for discrete-time hybrid systems in this chapter. We can expect better estimation performance by taking into consideration noncausal information of observations. The hybrid smoother is realized by two filters approach [2,8,12,20,22,23]. Finally we give numerical examples and verify that we can obtain better estimation performance by smoothing than filtering.

The organization of this chapter is as follows. In section 2 we describe the systems and problem formulation. In section 3 we present the hybrid estimation algorithms by the MPT approach over the finite time interval. In subsection 3.1 we review the hybrid filtering algorithm and in subsection 3.2 we design the backward filters and present the hybrid smoothing algorithm by the two filters approach. In section 4 we consider numerical examples and verify the effectiveness of the estimation algorithms presented in this chapter. In the Appendix we present the principles of hybrid optimality, which give the basis of validity for the hybrid estimation algorithms presented in this chapter

2. Systems and Problems Formulation

Let (Ω, F, P) be a probability space and, on this space, we consider the following system with mode transitions and noises which aren''t restricted to be Gaussian.

$$\begin{aligned} x(k+1) &= A_d(k,\ \theta(k))x(k) + w_d(k,\ \theta(k)), \\ x(0) &= x_0,\ \theta(0) = i_0 \\ y(k) &= H_d(k,\ \theta(k))x(k) + v_d(k,\ \theta(k)) \end{aligned} \tag{1}$$

where $x \in \mathbf{R^n}$ is the state, $w_d \in \mathbf{R^n}$ is the exogenous random noise, $v_d \in \mathbf{R^k}$ is the measurement noise, and $y \in \mathbf{R^k}$ is the measured output. x_0 is an unknown initial state and it is assumed that a distribution of initial modes i_0 is given. The noises $w_d(\cdot,\ \cdot)$ and $v_d(\cdot,\ \cdot)$ aren''t restricted to be Gaussian.

We assume that all these matrices are of compatible dimensions.

Let $\mathsf{M} = \{1,\ 2,\ \cdots,\ m\}$ denote the state space of $\theta(k)$. In this chapter it is assumed that the probability distribution of $\theta(\cdot)$ is unknown or inaccessible. But it is also assumed that a finite number of candidate distributions and the true probability distribution is among the candidate distributions. Let $r \in \mathsf{N}_0 = \{1,\ 2,\ \cdots,\ n_0\}$, and let $\mathsf{P} = \{\phi^{(1)}(\cdot),\ \cdots,\ \phi^{(n_0)}(\cdot)\}$ denote the set of such candidate distributions on M, i.e., for $r \in \mathsf{N}_0$ and $k \in [0,\ N]$, $\phi^{(r)}(k) = (\phi_1^{(r)}(k),\ \cdots,\ \phi_m^{(r)}(k))$ with $\phi_i^{(r)}(k) \geq 0$ and $\sum_{i=1}^{m} \phi_i^{(r)}(k) = 1$.

The fixed-interval optimal hybrid estimation problems we address in this chapter for the system (1) are to find the MPT (most probable trajectory) estimate of $x(k)$, $k \in [0,\ N]$, over the finite horizon $[0,N]$, using the information available on the known part of the observation $y(\cdot)$ for the given distributions of initial mode i_0 and initial state x_0. We define the following performance indices for $r \in \mathsf{N}_0$ and $k \in [0,\ N]$:

$$\begin{aligned} J_{0k}^{(r)}(x_0,\ w_d,\ v_d) := \sum_{l=0}^{k-1} \sum_{i=1}^{m} \phi^{(r)}(l)\big(w_d^{'}(l,\ i)M_d(l,\ i)w_d(l,\ i) + v_d^{'}(l,\ i)N_d(l,\ i)v_d(l,\ i)\big) \\ + (x_0 - \hat{x}_0)' D_0 (x_0 - \hat{x}(0)) \end{aligned} \tag{2}$$

$$\begin{aligned} J_{0N}^{(r)}(x_0,\ w_d,\ v_d) := \sum_{l=0}^{N-1} \sum_{i=1}^{m} \phi^{(r)}(l)\big(w_d^{'}(l,\ i)M_d(l,\ i)w_d(l,\ i) + v_d^{'}(l,\ i)N_d(l,\ i)v_d(l,\ i)\big) \\ + (x_0 - \hat{x}_0)' D_0 (x_0 - \hat{x}(0)) \\ + (x(N) - \hat{x}_N)' D_N (x(N) - \hat{x}_N) \end{aligned} \tag{3}$$

where $\hat{x}_0$ is an initial estimate of x_0 and $\hat{x}_{\mathrm{N}}$ is a terminal estimate of $x(N)$. $M_d(l,\ i) > O$, $N_d(l,\ i) \geq O$, $D_0 > O$ and $D_N > O$ are symmetric matrices which reflect the uncertainties on the noises $w_d(\cdot,\ \cdot)$ and $v_d(\cdot,\ \cdot)$ with the estimates $\hat{x}_0$ and $\hat{x}_{\mathrm{N}}$. Thus these performance indices mean the energies of noises, initial and terminal estimates under some uncertainties averaged by the mode distributions for each $r \in \mathsf{N}_0$. We consider the optimization problems to decide $w_d(\cdot, i)$, $v_d(\cdot, i)$ and $r \in \mathrm{N}_0$ minimizing $J_{0k}^{(r)}$ and $J_{0N}^{(r)}$ utilizing the known parts of the observed information $\mathrm{Y}_N = \{y(l) \mid 0 \le l \le N\}$.

Since the mode $\theta(k)$ at each time is inaccessible, we cannot directly design estimators for the system (1) including the unknown modes. Also, even if the modes are accessible, the computational complexity can exponentially increase with k if we directly design the estimators for the system (1) including $\theta(k)$ explicitly. Hence we introduce the system averaged through the mode distributions for each $r \in \mathrm{N}_0$.

For notational simplicity, we adopt the following notation.

$$\overline{F}^{(r)}(k) = \sum_{i=1}^{m} \phi_i^{(r)}(k) F(k, i)$$

for a matrix function $F(k,i)$ and $r \in \mathrm{N}_0$. Similarly $\overline{F_1 F_2}^{(r)}(k) = \sum_{i=1}^{m} \phi_i^{(r)}(k) F_1(k, i) F_2(k, i)$ for matrix functions $F_1(k,i)$ and $F_2(k,i)$ and so on. Using these notations, we can shift the drift term in the system (1) to $\overline{A_d}^{(r)}(k)$ as follows:

$$x(k+1) = \overline{A_d}^{(r)}(k) x(k) + w_d(k)$$

where

$$w_d(k) = w_d^{(r)}(k) = (A_d(k, \theta(k)) - \overline{A_d}^{(r)}(k)) x(k) + w_d(k, \theta(k)).$$

By replacing the system noise $w_d(k, i)$ by $(\overline{A_d}^{(r)}(k) - A_d(k, i)) x(k) + w_d(k)$ and the observation noise $v_d(k, i)$ by $y(k) - H_d(k, i) x(k)$ in the performance indices (2) and (3), we define

$$\begin{aligned} & L^{(r)}(k, x, w_d, y) \\ & := \sum_{i=1}^{m} \phi_i^{(r)}(k) ([(\overline{A_d}^{(r)}(k) - A_d(k, i)) x + w_d]' \\ & \qquad \times M_d(k, i) [(\overline{A_d}^{(r)}(k) - A_d(k, i)) x + w_d] \\ & \qquad + (y - H_d(k, i) x)' N_d(k, i) (y - H_d(k, i) x)). \end{aligned}$$

Then we can define the following performance indices:

$$J_f^{(r)}(k, x, w_d(\cdot)) := \sum_{l=0}^{k-1} L^{(r)}(l, x(l), w_d(l), y(l)) + \Phi_0(x(0)) \tag{4}$$

$$J_b^{(r)}(k, x, w_d(\cdot)) := \sum_{l=k}^{N-1} L^{(r)}(l, x(l), w_d(l), y(l)) + \Phi_N(x(N)) \tag{5}$$

$$J_s^{(r)}(k, x, w_d(\cdot)) := J_f^{(r)}(k, x, w_d(\cdot)) + J_b^{(r)}(k, x, w_d(\cdot))$$

where $\Phi_0(x(\cdot)) = (x(\cdot) - \hat{x}_0)' D_0 (x(\cdot) - \hat{x}_0)$ and $\Phi_N(x(\cdot)) = (x(\cdot) - \hat{x}_N)' D_N (x(\cdot) - \hat{x}_N)$.

http://dx.doi.org/10.5772/51385

We consider the optimal control problems to minimize $J_f^{(r)}$ and $J_s^{(r)} = J_f^{(r)} + J_b^{(r)}$ for the given parts of Y_N. Let $V_f^{(r)}(k, x)$ and $V_b^{(r)}(k, x)$ be the value functions of these control problems as follows:

$$V_f^{(r)}(k, x) := \inf_{w_d(\cdot)} J_f^{(r)}(k, x, w_d)$$
$$V_b^{(r)}(k, x) := \inf_{w_d(\cdot)} J_b^{(r)}(k, x, w_d)$$
$$V_s^{(r)}(k, x) := V_f^{(r)}(k, x) + V_b^{(r)}(k, x)$$
$$w_{d,f}^{(r)*}(k) := \arg\min\{J_f^{(r)}(k, x, w_d(k)) : w \in R^n\}$$
$$w_{d,b}^{(r)*}(k) := \arg\min\{J_b^{(r)}(k, x, w_d(k)) : w \in R^n\}$$
$$w_{d,s}^{(r)*}(k) := \arg\min\{J_s^{(r)}(k, x, w_d(k)) : w \in R^n\}$$

Then define

$$\hat{x}_f^{(r)}(k) := \arg\min\{V_f^{(r)}(k, x) : x \in R^n\},$$
$$V_f^{(r)}(k) := V_f^{(r)}\left(k, \hat{x}_f^{(r)}(k)\right)$$

and

$$\hat{r}_f(k) := arg\min\{V_f^{(r)}(k) : r \in \mathrm{N}_0\}.$$

Then the most probable distribution is $\phi^{(\hat{r}_f(k))}(\cdot)$. Let $\hat{x}_f(k) = \hat{x}_f^{(\hat{r}_f(k))}(k)$ and we have

$$V_f^{(\hat{r}_f(k))}(k, \hat{x}_f(k)) \le V_f^{(r)}\left(k, \hat{x}_f^{(r)}(k)\right)$$
$$\le V_f^{(r)}(k, x) = J_f^{(r)}\left(k, x, w_{d,f}^{(r)*}(k)\right) \le J_f^{(r)}(k, x, w_d(k)).$$

Also define

$$\hat{x}_s^{(r)}(k) := arg\min\{V_s^{(r)}(k, x) : x \in \boldsymbol{R}^n\},$$
$$V_s^{(r)}(k) := V_s^{(r)}(k, \hat{x}_s^{(r)}(k))$$

and

$$\hat{r}_s(k) := arg\min\{V_s^{(r)}(k) : r \in \mathrm{N}_0\}.$$

Then the most probable distribution is $\phi^{(\hat{r}_s(k))}(\cdot)$. Let $\hat{x}_s(k) = \hat{x}_s^{(\hat{r}_s(k))}(k)$ and we have

$$V_f^{(\hat{r}_f(k))}(k, \hat{x}_f(k)) \le V_f^{(r)}\left(k, \hat{x}_f^{(r)}(k)\right)$$
$$\le V_f^{(r)}(k, x) = J_f^{(r)}\left(k, x, w_{d,f}^{(r)*}(k)\right) \le J_f^{(r)}(k, x, w_d(k)).$$

Now we define the following optimal estimators in the sense of most probable trajectory (MPT).

2.1 Definition

Given the matrices M_d, N_d, D_0 and D_N, $(\hat{r}_f(k), \hat{x}_f(k))$, $k \geq 0$, is called an optimal filter (in the MPT sense) if it minimizes $V_f^{(r)}(k, x)$. $(\hat{r}_s(k), \hat{x}_s(k))$, $0 \leq k \leq N$ is called an optimal smoother (in the MPT sense) if it minimizes $V_s^{(r)}(k, x)$.

Then we formulate the following optimal hybrid estimation problems for the performance indices (4) and (5).

The Optimal Hybrid Filtering Problem for Linear Discrete-Time Systems:

Find the pair $(\hat{r}_f(l), \hat{x}_f^{(\hat{r}_f(l))}(l))$, $l \in [0, k]$ minimizing the performance index (4) based on the causal part $\mathrm{Y}_k = \{y(l) \mid 0 \leq l \leq k\}$ of the observed information Y_N.

The Optimal Hybrid Smoothing Problem for Linear Discrete-Time Systems:

Find the pair $(\hat{r}_s(k), \hat{x}_s^{(\hat{r}_s(k))}(k))$, $k \in [0, N]$ minimizing the performance index (4)+(5) based on the whole observed information Y_N.

Remark 2.1. *In general, if we directly adopt dynamic programming (DP) method for mode-dependent systems, it can arise that computational complexity increases exponentially with time k ([5,11]). On the other hand in this chapter we consider the averaged systems and averaged performance indices for them with regard to the candidates of the mode distributions. Note that this introduction of the averaged systems and performance indices prevents the computational complexity from increasing exponentially by applying the dynamical programming (DP) method as seen in the next section.*

3. Hybrid Estimation Algorithms

We assume the following condition:

A1: The matrices $\overline{A_d}^{(r)}(k)$, $k = 0, 1, \cdots$ are invertible.

Remark 3.1. *As described in [24], note that* ***A1*** *is the reasonable assumption in the discrete-time models. First we consider the following continuous-time model:*

$$\dot{x}_c(t) = A_c(t, \theta_c(t))x_c(t) + w_c(t, \theta_c(t))$$

where $\theta_c(t) \in M$, $t \geq 0$ is the switching mode process. If we discretize this model with stepsize $h \geq 0$, let $x(k) = x_c(kh)$ and the following discretized equation holds:

$$x(k+1) = [I + h\,A_c(kh, \theta_c(kh))]x(k) + w_d(k, \theta(k))$$

where $w_d(k, \theta(k)) = h\,w_c(kh, \theta_c(kh))$. Let

$$A_d(k, i) = I + A_c(kh, i)$$

and then we obtain

$$\begin{aligned}\overline{A_d}^{(r)}(k) &= \sum_{i=1}^{m} \phi_i^{(r)}(k) A_d(k, i) \\ &= \sum_{i=1}^{m} \phi_i^{(r)}(k)[I + h\, A_c(kh, i)] \\ &= I + h \sum_{i=1}^{m} \phi_i^{(r)}(k) A_c(kh, i).\end{aligned}$$

If we assume that $A_c(t, i)$ is uniformly bounded, then $\overline{A_d}^{(r)}(k)$, $k = 0, 1, \cdots$ is invertible for h small enough.

3.1 Optimal Hybrid Filtering

The dynamic programming (DP) equations associated with the forward control problem to minimize $J_f^{(r)}$ with regard to $w_d(\cdot)$ are given as follows:

$$V_f^{(r)}(k+1, x) = \min_{w_d} \{ L^{(r)}(k, \overline{A_d}^{(r),-1}(k)(x - w_d), w_d, y(k)) + V_f^{(r)}(k, x) \}$$
$$V_f^{(r)}(0, x) = \Phi_0(x), \quad r \in \mathrm{N}_0$$

Let

$$V_f^{(r)}(k, x) = x' K_f^{(r)}(k) x + 2(p_f^{(r)}(k))' x + q_f^{(r)}(k) \tag{6}$$

for some functions $K_f^{(r)}$ and $q_f^{(r)}$ with appropriate dimensions. Then we obtain the following minimizing $w_d(\cdot)$.

$$w_{d,f}^{(r)*}(k, x) = x - \overline{A_d}^{(r)}(k) S_d^{(r)}(k) \left(\overline{A_d' M_d}^{(r)}(k) x + \overline{H_d' N_d}^{(r)}(k) y(k) - p_f^{(r)}(k) \right)$$

where

$$S_d^{(r)}(k) = [\overline{A_d' M_d A_d}^{(r)}(k) + \overline{H_d' N_d H_d} + K_f^{(r)}(k)]^{-1}.$$

Then we obtain the following matrix difference equations, forward vector equations and scalar equations with initial conditions:

$$K_f^{(r)}(k+1) = \overline{M_d}^{(r)}(k) - \overline{M_d A_d}^{(r)}(k) S_d^{(r)}(k) \overline{A_d' M_d}^{(r)}(k), \quad K_f^{(r)}(0) = D_0 \tag{7}$$

$$p_f^{(r)}(k+1) = -\overline{M_d A_d}^{(r)}(k) S_d^{(r)} \left[\overline{H_d' N_d}^{(r)}(k) y(k) - p_f^{(r)}(k) \right], \quad p_f^{(r)}(0) = -D_0 \hat{x}_0 \tag{8}$$

$$\begin{aligned} q_f^{(r)}(k+1) = &-\left[\overline{H_d' N_d}^{(r)}(k) y(k) - p_f^{(r)}(k) \right]' S_d^{(r)}(k) \left[\overline{H_d' N_d}^{(r)}(k) y(k) - p_f^{(r)}(k) \right] \\ &+ y'(k) \overline{N_d}^{(r)}(k) y(k) + q_f^{(r)}(k), \quad q_f^{(r)}(0) = \hat{x}_0 D_0 \hat{x}_0 \end{aligned} \tag{9}$$

For any given k, by letting $\partial V_f^{(r)} / \partial x = 0$, we obtain

$$K_f^{(r)}(k)x + p_f^{(r)}(k) = 0.$$

Since it can be shown that the matrix $K_f^{(r)}(k)$ is positive-definite, we obtain

$$\hat{x}_f^{(r)}(k) = -(K_f^{(r)}(k))^{-1} p_f^{(r)}(k)$$

as the minimizer of $V_f^{(r)}(k, x)$. Then we obtain

$$\begin{aligned} \hat{x}_f^{(r)}(k+1) &= -({}^{K_f^{(r)}k1}_{\quad(+)} p_f^{(r)}(k+1) \\ &= -[\overline{M_d}^{(r)}(k) - \overline{M_d A_d}^{(r)}(k) S_d^{(r)}(k) \overline{A_d' M_d}^{(r)}(k)]^{-1} \overline{M_d A_d}^{(r)}(k) S_d^{(r)} \\ &\quad \times [\overline{H_d' N_d}^{(r)}(k) y(k) + K_f^{(r)}(k) \hat{x}_f^{(r)}(k)],\ \hat{x}_f^{(r)}(0) = \hat{x}_0 \end{aligned} \tag{10}$$

and

$$\begin{aligned} q_f^{(r)}(k+1) = &-[\overline{H_d' N_d}^{(r)}(k) y(k) + K_f^{(r)}(k) \hat{x}_f^{(r)}(k)]' S_d^{(r)}(k) \\ &\times [\overline{H_d' N_d}^{(r)}(k) y(k) + K_f^{(r)}(k) \hat{x}_f^{(r)}(k)] \\ &+ y'(k) \overline{N_d}^{(r)}(k) y(k) + q_f^{(r)}(k),\ q_f^{(r)}(0) = \hat{x}_0 D_0 \hat{x}_0 \end{aligned} \tag{11}$$

We also obtain

$$V_f^{(r)}(k) = -(\hat{x}_f^{(r)}(k))' K_f^{(r)}(k) \hat{x}_f^{(r)}(k) + q_f^{(r)}(k).$$

Now we have the following filtering algorithm, which gives the solution of **the Optimal Hybrid Filtering Problem for Linear Continuous-Time Systems**.

****Optimal hybrid filtering algorithm****

Step 1) Obtain $K_f^{(r)}(k)$, $\hat{x}_f^{(r)}(k)$ and $q_f^{(r)}(k)$ for $r \in \mathrm{N}_0$ and $k \in [0, N]$ by solving (7), (10) and (11) with initial conditions.

Step 2) Choose $\hat{r}_f(k)$ that minimizes

$$V_f^{(r)}(k) = -(\hat{x}_f^{(r)}(k))' K_f^{(r)}(k) \hat{x}_f^{(r)}(k) + q_f^{(r)}(k).$$

Then the most probable distribution is $\phi^{(\hat{r}_f(k))}(k)$ and the optimal filter is given by

$$(\hat{r}_f(k), \hat{x}_f(k)) = (\hat{r}_f(k), \hat{x}_f^{(r(k))}(k)).$$

3.2 Optimal Hybrid Smoothing

The dynamic programming (DP) equations associated with the backward control problem to minimize $J_b^{(r)}$ with regard to $w_d(\cdot)$ are given as follows:

$$V_b^{(r)}(k,\ x)=\min_{w_d}\{L^{(r)}(k,\ x,\ w_d,\ y(k))+V_b^{(r)}(k+1,\ \overline{A}_d^{(r)}(k)x+w_d)\}$$

$$V_b^{(r)}(N,\ x)=\Phi_N(x),\ \ r\in \mathrm{N}_0$$

Let

$$V_b^{(r)}(k,\ x)=x'K_b^{(r)}(k)x+2(p_b^{(r)}(k))'x+q_b^{(r)}(k) \tag{12}$$

for some functions $K_b^{(r)}$, $p_b^{(r)}$ and $q_b^{(r)}$ with appropriate dimensions. Then we obtain the following minimizing $w_d(\cdot)$.

$$w_{d,b}^{(r)*}(k,\ x)=\{-\overline{A_d}^{(r)}(k)+T_d^{(r)}(k)\overline{M_dA_d}(k)\}x(k)-T_d^{(r)}(k)p_f^{(r)}(k+1)$$

where

$$T_d^{(r)}(k)=[\overline{M_d}^{(r)}(k)+K_b^{(r)}(k+1)]^{-1}.$$

Let

$$V_b^{(r)}(k,\ x)=x'K_b^{(r)}(k)x+2(p_b^{(r)}(k))'x+q_b^{(r)}(k) \tag{13}$$

for some functions $K_b^{(r)}$, $p_b^{(r)}$ and $q_b^{(r)}$ with appropriate dimensions. Then we obtain the following matrix difference equations, backward vector equations and scalar equations with terminal conditions:

$$\begin{gathered}K_b^{(r)}(k)=\overline{A_d'M_dA_d}^{(r)}(k)-\overline{A_d'M_d}^{(r)}(k)T_d^{(r)}(k)\overline{M_dA_d}^{(r)}(k)+\overline{H_d'N_dH_d}^{(r)}(k),\\ K_b^{(r)}(N)=D_N\end{gathered} \tag{14}$$

$$p_b^{(r)}(k)=\overline{A_d'M_d}(k)T_d^{(r)}(k)p_b^{(r)}(k+1)\text{ - }\overline{H_dN_d}^{(r)}(k)y(k)],\ \ p_b^{(r)}(N)=-D_N\hat{x}_N \tag{15}$$

$$\begin{aligned}q_b^{(r)}(k)=&-p_b^{(r)'}(k+1)T_d^{(r)}(k)p_b^{(r)}(k+1)+q_b^{(r)}(k+1)\\ &+y'(k)\overline{N}_d(k+1)y(k),\ \ q_b^{(r)}(N)=\hat{x}'_ND_N\hat{x}_N\end{aligned} \tag{16}$$

For any given k, by letting $\partial V_b^{(r)}/\partial x=0$, we obtain

$$K_b^{(r)}(k)x+p_b^{(r)}(k)=0.$$

Since it can be also shown that the matrix $K_b^{(r)}(k)$ is positive-definite, we obtain

$$\hat{x}_b^{(r)}(k)=-(K_b^{(r)}(k))^{-1}p_b^{(r)}(k)$$

as the minimizer of $V_b^{(r)}(k,\ x)$. Then we obtain

$$\begin{aligned}\hat{x}_b^{(r)}(k) &= -(K_b^{(r)}(k))^{-1} p_b^{(r)}(k) \\ &= [\overline{A_d{}' M_d A_d}^{(r)}(k) - \overline{A_d{}' M_d}^{(r)}(k) T_d^{(r)}(k) \overline{M_d A_d}^{(r)}(k) + \overline{H_d{}' N_d H_d}^{(r)}(k)]^{-1} \\ &\quad \times [\overline{A_d{}' M_d}(k) T_d^{(r)}(k) K_b^{(r)}(k+1) \hat{x}_b^{(r)}(k+1) + \overline{H_d{}' N_d}^{(r)}(k) y(k)], \quad \hat{x}_b^{(r)}(N) = \hat{x}_N\end{aligned} \tag{17}$$

and

$$\begin{aligned}q_b^{(r)}(k) &= -\hat{x}_b^{(r)\prime}(k+1) K_b^{(r)}(k+1) T_d^{(r)}(k) K_b^{(r)}(k+1) \hat{x}_b^{(r)}(k+1) \\ &\quad + q_b^{(r)}(k+1) + y'(k) \overline{N}_d(k+1) y(k), \quad q_b^{(r)}(N) = \hat{x}_N{}' D_N \hat{x}_N.\end{aligned} \tag{18}$$

We also obtain

$$V_b^{(r)}(k) = -(\hat{x}_b^{(r)}(k))' K_b^{(r)}(k) \hat{x}_b^{(r)}(k) + q_b^{(r)}(k).$$

Using (6) and (12), we can express $V_s^{(r)}(k,\ x)$ as

$$V_s^{(r)}(k,\ x) = x'[K_f^{(r)}(k) + K_b^{(r)}(k)]x + 2[p_f^{(r)}(k) + p_b^{(r)}(k)]' x + q_f^{(r)}(k) + q_b^{(r)}(k)$$

Let

$$\partial V_s^{(r)} / \partial x = 0$$

and we obtain the following form.

$$\hat{x}_s^{(r)}(k) = -[K_f^{(r)}(k) + K_b^{(r)}(k)]^{-1}(p_f^{(r)}(k) + p_b^{(r)}(k))$$

Since $p_f^{(r)}(k) = -K_f^{(r)}(k)\hat{x}_f^{(r)}(k)$ and $p_b^{(r)}(k) = -K_b^{(r)}(k)\hat{x}_b^{(r)}(k)$, for each candidate r of given distributions, we can obtain the following form of smoothed estimate at time k by the forward and backward filtered estimates.

$$\hat{x}_s^{(r)}(k) = K_s^{(r)}(k)[K_f^{(r)}(k)\hat{x}_f^{(r)}(k) + K_b^{(r)}(k)\hat{x}_b^{(r)}(k)]$$

where $K_s^{(r)}(k) = [K_f^{(r)}(k) + K_b^{(r)}(k)]^{-1}$.

Now we have the following smoothing algorithm, which gives the solution of **the Optimal Hybrid Smoothing Problem for Linear Continuous-Time Systems**.

****Optimal hybrid smoothing algorithm****

Step 1) Obtain $K_b^{(r)}(k)$, $\hat{x}_b^{(r)}(k)$ and $q_b^{(r)}(k)$ for $r \in \mathrm{N}_0$ and $k \in [0,\ N]$ by solving (14), (17) and (18) with terminal conditions.

Step 2) Choose $\hat{r}_s(k)$ that minimizes

$$V_s^{(r)}(k) = V_f^{(r)}(k) + V_b^{(r)}(k)$$

where

http://dx.doi.org/10.5772/51385

$$V_b^{(r)}(k) = -\hat{x}_b^{(r)'}(k) K_b^{(r)}(k) \hat{x}_b^{(r)}(k) + q_b^{(r)}(k).$$

Then the most probable distribution is $\phi^{(\hat{r}_s(k))}(k)$ and the optimal smoother is given by

$$\begin{aligned}(\hat{r}_s(k), \hat{x}_s(k)) &= (\hat{r}_s(k), \hat{x}_s^{(\hat{r}_s(k))}(k)) \\ &= (\hat{r}_s(k), K_s^{(\hat{r}_s(k))}(k)[K_f^{(\hat{r}_s(k))}(k)\hat{x}_f^{(\hat{r}_s(k))}(k) + K_b^{(\hat{r}_s(k))}(k)\hat{x}_b^{(\hat{r}_s(k))}(k)])\end{aligned}$$

where $K_s^{(\hat{r}_s(k))}(k) = [K_f^{(\hat{r}_s(k))}(k) + K_b^{(\hat{r}_s(k))}(k)]^{-1}$.

Remark 3.2. *Note that, if the system (1) is a single mode system, i.e., the system (1) is independent of θ(k), the forms of the filter and smoother presented in this section are reduced to the well-known ones of the Kalman filter and smoother.*

4. Numerical Examples

In this section, we study numerical examples to demonstrate the effectiveness of the presented design algorithms.

We consider the following two mode systems and assume that the system parameters are as follows:

$$\begin{aligned} x(k+1) &= A_d(k, \theta(k))x(k) + w_d(k, \theta(k)), \\ x(0) &= x_0, \ \theta(0) = i_0 \\ y(k) &= H_d(k, \theta(k))x(k) + v_d(k, \theta(k)) \end{aligned} \tag{19}$$

where

· Mode 1: $A_1 = \begin{bmatrix} 0 & 1 \\ -0.8 & 0.6 \end{bmatrix}$, · Mode 2: $A_2 = \begin{bmatrix} 0.5 & 1 \\ -0.4 & 0.6 \end{bmatrix}$,

$$H = [1, 0]$$

and

$$M(t, i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ N(t, i) = 1, \ D_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

for i=*1,2*. We set $\hat{x}_0 = col(-0.1, 0)$ and the distribution of the initial mode i_0 as (1/ 2, 1/ 2). $w_d(\cdot, \cdot)$ and $v_d(\cdot, \cdot)$ are stochastic noises which aren''t restricted to be Gaussian white. The candidates of mode distributions are given as follows:

$$\phi^{(1)} = \left(\frac{2}{5}, \frac{3}{5}\right), \ \phi^{(2)} = \left(\frac{1}{2}, \frac{1}{2}\right), \ \phi^{(3)} = \left(\frac{3}{5}, \frac{2}{5}\right)$$

The paths of $\theta(k)$ are generated randomly, and the performances are compared under the same circumstance, that is, the same set of the paths so that the performances can be easily compared.

We consider the whole system (19) with the true mode distribution $\phi^{(3)}$ over the time interval $k \in [0, 100]$. We verify the effectiveness of the presented hybrid estimation algorithms and compare the estimation performances for the optimal filtering and smoothing algorithms. In order to carry out these algorithms we solve the forward or backward triplet of the difference equations (7)(10)(11) or (14)(17)(18) with the initial or terminal conditions for given observation $y(\cdot)$ and each candidate $r=1,2,3$ of given distributions, and obtain the pair $(\hat{r}_f(k), \hat{x}_f(k))$ minimizing $V_f^{(r)}(k, x)$ in the filtering case or the pair $(\hat{r}_s(k), \hat{x}_s(k))$ minimizing $V_s^{(r)}(k, x)$ in the smoothing case for $k \in [0, 100]$.

Filtered and smoothed values of the first components of the whole system states are given by Fig. 1 and Fig. 2 respectively. Fig. 3 and Fig. 4 show the square errors between the states and filtered values, and the states and smoothed values respectively. The mean square errors over the time interval *[0, 100]* are 0.0276 in the filtering case, and 0.0151 in the smoothing case respectively. From these figures and calculation results it is shown that the smoother gives better estimation than the filter. Filtered and smoothed values of the second components of the whole system states are given by Fig. 5 and Fig. 6 respectively. Fig. 7 and Fig. 8 show the square errors between the states and filtered values, and the states and smoothed values respectively. The mean square errors over the time interval *[0, 100]* are 0.0151 in the filtering case, and 0.0118 in the smoothing case respectively. From these figures and calculation results it is shown that the smoother gives better estimation than the filter. Filtered and smoothed mode distributions are given by Fig. 9 and Fig. 10. Notice that the vertical axes show the candidates of the mode distributions not the modes themselves. In Fig. 9 the filtered values of the mode distributions rapidly change to be left undecided. To the contrary in Fig. 10 the smoothed values of the mode distributions are firmly decided. Through these ten figures it is shown that the optimal smoother presented in this chapter gives better estimate performance than the optimal filter presented in the previous work [24] from the point of view of both state and modes estimation.

http://dx.doi.org/10.5772/51385

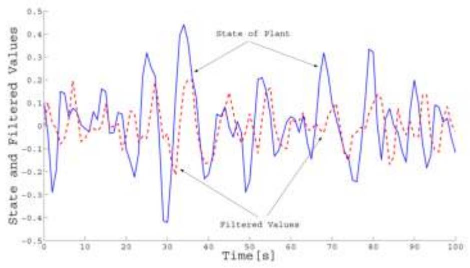

Figure 1. The state of the system and filtered values: 1st components

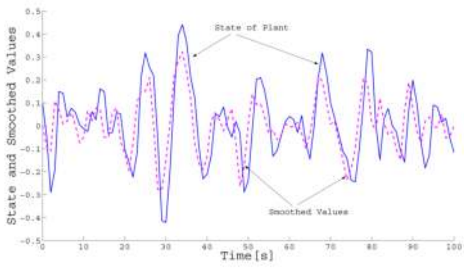

Figure 2. The state of the system and smoothed values: 1st components

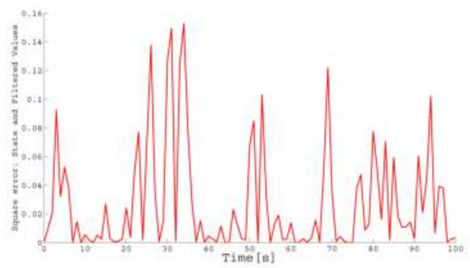

Figure 3. The square errors between the state and filtered values: 1st components

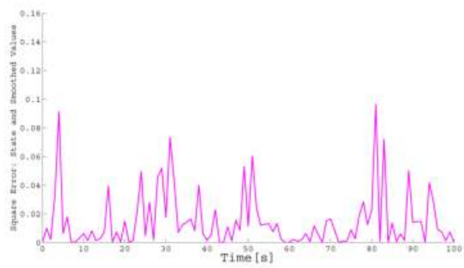

Figure 4. The square errors between the state and smoothed values: 1st components

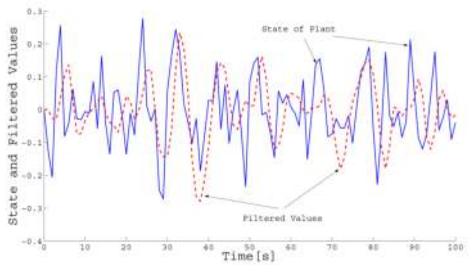

Figure 5. The state of the system and filtered values: 2nd components

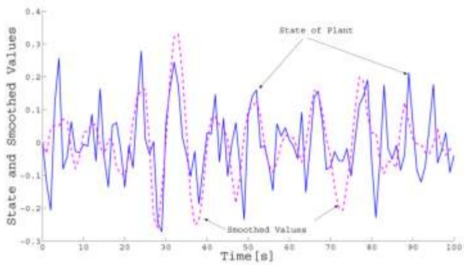

Figure 6. The state of the system and smoothed values: 2nd components

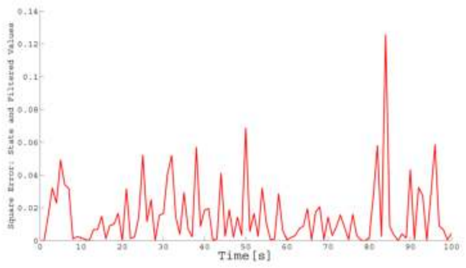

Figure 7. The square errors between the state and filtered values: 2nd components

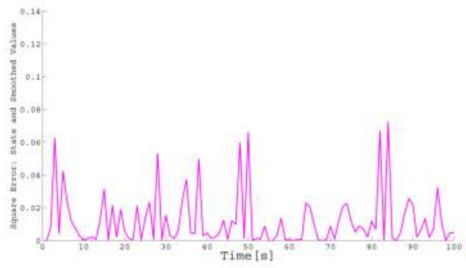

Figure 8. The square errors between the state and smoothed values: 2nd components

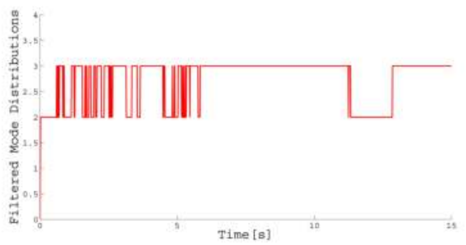

Figure 9. The filtered mode distributions

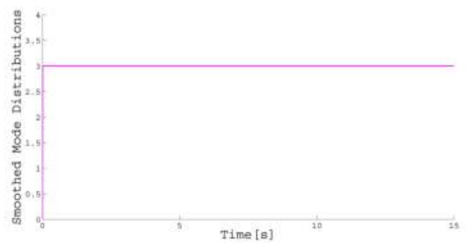

Figure 10. The smoothed mode distributions

5. Concluding Remarks

In this chapter we have studied the state and mode estimation problems for linear discrete-time hybrid systems over the fixed time interval. The systems aren''t restricted to the Markovian jump systems and the added noises aren''t restricted to be Gaussian. With regard to concrete examples of the systems considered in this chapter, refer to [24,25]. Those examples show that the systems and estimation algorithms presented in this chapter cover extreme broad classes of dynamical systems affected by the noises not to be restricted to be Gaussian. We have adopted the MPT approach. The state and mode estimation approach adopted in this paper guarantees the optimality of estimation performance in the meaning of MPT different from the previous work ([4,7]).

In this chapter we have considered the problems that both system state and modes are estimated. However we have considered the problems that the distributions of the modes over the fixed time interval not the modes themselves are estimated to grasp the global behavior of the hybrid systems over the long time intervals. In order to estimate both the system state and distributions of the modes we have introduced the averaged performance indices with respect to the candidates of the mode distributions for the averaged systems. This introduc-

http://dx.doi.org/10.5772/51385

tion of the averaged systems and performance indices prevents the computational complexity from increasing exponentially with time passage. For these performance indices we have formulated the optimal filtering and smoothing problems based on the available observed information. The estimation problems have been reduced to the optimal control problems to find the noises minimizing the introduced performance indices. We have derived the forward and backward matrix difference equations and the forward and backward filter equations with the initial and terminal conditions respectively, which give the necessary conditions for the solvability of the optimal estimation problems. Then we have presented the optimal hybrid smoothing algorithm by the two filters approach. Finally we have studied the numerical examples to compare the estimation performances by filtering and smoothing. We have obtained the better estimation performance by the smoothing algorithm than the filtering algorithm from the point of view of both state and modes estimation.

With regard to continuous-time cases, refer to [16,17,25]. In particular, in [17,25], the cases that concerned systems are assumed to be Markovian jump systems is also considered. In these papers the concept of quasi-stationary distributions is introduced for the Markovian mode processes and near optimality of limiting estimators with the quasi-stationary distributions is shown. It is well known that the concept of quasi-stationary distribution is very important and highly practical to grasp behavior of stochastic processes over long run time. As a further research issue it is very significant that the quasi-stationary distributions of stochastic mode processes and estimator with these distributions are investigated for the discrete-time hybrid systems.

Appendix: Principle of Hybrid Optimality

With regard to the optimal control problems considered in this chapter, it is obvious that principle of optimality does not hold for the optimal trajectory $x^*(\cdot)$ with optimal control input $w_d^{(r)*}(\cdot)$ and each performance index for each mode distribution candidate $r \in \mathrm{N}_0$. However, for the pair $(\hat{r}(\cdot), x(\cdot))$ of the optimal mode distribution candidate and optimal trajectory with the optimal control inputs $w_d^{(r)*}(\cdot)$, the following principles of hybrid optimality hold. These principles give a basis of validity for the hybrid estimation algorithms presented in this chapter.

Consider the following system

$$x(k+1) = \bar{A}_d^{(r)}(k)x(k) + w_d(k) \tag{20}$$

$$y(k) = H_d(k,\ \theta(k))x(k) + v_d(k,\ \theta(k))$$

and the performance indices

$$J_f^{(r)}(k,\ x,\ w_d(\cdot)) = \sum_{l=0}^{k-1} L^{\ (r)}(l,\ x(l),\ w_d(l),\ y(l)) + \Phi_0(x(0)) \tag{21}$$

$$J_b^{(r)}(k,\ x,\ w_d(\cdot)) = \sum_{l=k}^{N-1} L^{\ (r)}(l,\ x(l),\ w_d(l),\ y(l)) + \Phi_N(x(N)) \tag{22}$$

$$J_s^{(r)}(k,\ x,\ w_d(\cdot)) = J_f^{(r)}(k,\ x,\ w_d(\cdot)) + J_b^{(r)}(k,\ x,\ w_d(\cdot)) \tag{23}$$

where

$$\begin{aligned} & L^{\ (r)}(k,\ x,\ w_d,\ y) \\ &= \sum_{i=1}^{m} \phi_i^{(r)}(k)([(\overline{A_d}^{(r)}(k) - A_d(k,\ i))x + w_d]' \\ &\qquad \times M_d(k,\ i)[(\overline{A_d}^{(r)}(k) - A_d(k,\ i))x + w_d] \\ &\quad + (y - H_d(k,\ i)x)' N_d(k,\ i)(y - H_d(k,\ i)x)). \end{aligned}$$

We consider the following three optimal control problems for the system (20) and the performance indices (21)-(23):

Problem (A): Forward Optimal Control Problem

Consider the system (20) with the initial state $x(0)$. Find the pair $(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^*}(l))$, $l \in [0,\ k]$ minimizing the value of the performance index (21) based on the causal part $\mathrm{Y}_k = \{y(l) \mid 0 \le l \le k\}$ of the observed information Y_N .

Problem (B): Backward Optimal Control Problem

Consider the system (20) with the initial state $x(k)$. Find the pair $(\hat{r}_b(l),\ w_{d,b}^{(\hat{r}_b(l))^*}(l))$, $l \in [k,\ N]$ minimizing the value of the performance index (22) based on the anti-causal part $\overline{Y}_k = \{y(l) \mid k \le l \le N\}$ of the observed information Y_N .

Problem (C): Fixed-Interval Optimal Control Problem

Consider the system (20) with the initial state $x(0)$. Find the pair $(\hat{r}_s(k),\ w_{d,s}^{(\hat{r}_s(k))^*}(k))$, $k \in [0,\ N]$ minimizing the value of the performance index (23) based on the whole observed information Y_N .

Proposition A (Principle of Hybrid Optimality (A)) *Consider the optimal control problem (A) on the time interval [0,k]. Also consider the optimal control problem minimizing the performance index*

$$J_f^{(r)}(\tau,\ x,\ w_d(\cdot)) = \sum_{l=0}^{\tau-1} L^{\ (r)}(l,\ x(l),\ w_d(l),\ y(l)) + \Phi_0(x(0)),\ 1 < \tau < \mathrm{k} \tag{24}$$

for the system (20) with the initial state x(0) on the partial time interval [0,τ] and then let the pair of optimal control inputs be $\left(\hat{\hat{r}}_f(l)\, w_{d,f}^{(\hat{\hat{r}}_f(l))^{**}}(l).\right)$, $l\in[0,\tau]$. *Then* $\left(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l)\right)=\left(\hat{\hat{r}}_f(l),\ w_{d,f}^{(\hat{\hat{r}})^{**}}(l)\right)$, $l\in[0,\tau]$ *holds.*

Proposition B (Principle of Hybrid Optimality (B)) *Consider the optimal control problem (B) on the time interval [k,N]. Also consider the optimal control problem minimizing the performance index*

$$J_b^{(r)}(\tau,\ x,\ w_d(\cdot))=\sum_{l=\tau}^{N-1} L^{\ (r)}(l,\ x(l),\ w_d(l),\ y(l))+\Phi_N(x(N)),\ k<\tau<N-1 \tag{25}$$

for the system (20) with the initial state $x(\tau)$ *on the partial time interval* $[\tau,\ N]$ *and then let the pair of optimal control inputs be* $\left(\hat{\hat{r}}_b(l) w_{d,b}^{(\hat{\hat{r}}_b(l))^{**}}(l)\right)$, $l\in[\tau,\ N]$. *Then* $\left(\hat{r}_b(l),\ w_{d,b}^{(\hat{r}_b(l))^{*}}(l)\right)=\left(\hat{\hat{r}}_b(l) w_{d,b}^{(\hat{\hat{r}}_b(l))^{**}}(l)\right)$, $l\in[\tau,\ N]$ *holds.*

Theorem C (Principle of Hybrid Optimality (C)) *Consider the optimal control problem (C) on the fixed time interval [0,N]. Split the performance index (23) into the two parts (21) and (22) and also consider the optimal control problems (A) on [0,k] and (B) on [k,N] for the system (20) with initial state x(0) and x(k) respectively. Then at each time k* $\left(\hat{r}_f(l)\, w_{d,f}^{(\hat{r}_f(l))^{*}}(l)\right)$, $l\in[0,\ k]$ *and* $\left(\hat{r}_b(l),\ w_{d,b}^{(\hat{r}_b(l))^{*}}(l)\right)$, $l\in[k,\ N]$ *are optimal input minimizing the values of the (21) and (22) to be used in order to compose the solution of the fixed-interval optimal control problem (C), i.e., any time* $k\in[0,\ N]$ $(\hat{r}_s(l),\ w_{d,s}^{(\hat{r}_f(l))^{*}}(l))=(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l)),\ l\in[0,\ k]$ *and* $(\hat{r}_s(l),\ w_{d,s}^{(\hat{r}_f(l))^{*}}(l))=(\hat{r}_b(l),\ w_{d,f}^{(\hat{r}_b(l))^{*}}(l)),\quad l\in[k+1,\ N]$ *hold.*

In this appendix we give only a proof of Proposition 7.1. The others can be shown by the similar arguments.

(Proof of Proposition A) We split the pair $\left(\hat{r}_f(l) w_{d,f}^{(\hat{r}_f(l))^{*}}(l)\right)$, $l\in[0,\ k]$ of the optimal control inputs into the following two parts:

$$(\hat{r}_{f,1},\ w_{d,f,1}^{(\hat{r}_{f,1})^{*}})=(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l))\ l\in[0,\ \tau]$$
$$(\hat{r}_{f,2},\ w_{d,f,2}^{(\hat{r}_{f,2})^{*}})=(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l))\ l\in[\tau+1,\ k]$$

Now we assume

$$(\hat{r}_{f,1},\ w_{d,f,1}^{(\hat{r}_{f,1})^{*}})\neq(\hat{\hat{r}}_f,\ w_{d,f}^{(\hat{\hat{r}}_f)^{**}})\ on\,[0,\ \tau]$$

Then there exists the pair of control inputs $\left(\hat{\hat{r}}_f,\ w_{d,f}^{(\hat{r}_f)^{**}}\right)$ giving less value of the performance index (24) than $\left(\hat{r}_{f,1},\ w_{d,f}^{(\hat{r}_{f,1})^{*}}\right)$ and so the pair of control inputs consisting of $\left(\hat{\hat{r}}_f, w_{d,f}^{(\hat{\hat{r}}_f)^{**}}\right)$ and $\left(\hat{r}_{f,2},\ w_{d,f,2}^{(\hat{r}_{f,2})^{*}}\right)$ gives less value of the performance index (21) than $\left(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l)\right)$, $l\in[0,\ k]$. This contradicts with the optimality of $\left(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f(l))^{*}}(l)\right)$, $l\in[0,\ k]$. Therefore $(\hat{r}_f(l),\ w_{d,f}^{(\hat{r}_f)^{*}})=(\hat{\hat{r}}_f(l),\ w_{d,f}^{(\hat{\hat{r}}_f(l))^{**}})$, $l\in[0,\ \tau]$ holds. (Q.E.D.).

Author details

Gou Nakura*

Address all correspondence to: gg9925_fiesta@ybb.ne.jp

References

[1] Basar, T. (1991). Optimal Performance Levels for Minimax Filters, Predictors and Smoothers. *Syst. Contr. Lett.*, 16(5), 309-317.

[2] Blanco, E., Neveux, P., & Thomas, G. (2006). The H_∞ Fixed-Interval Smoothing Problem for Continuous Systems. *IEEE Trans. Signal Process.*, 54(11), 4085-4090.

[3] Blom, H. A. P., & Bar-Shalom, Y. (1990). Time-reversion of a hybrid state stochastic difference system with a jump-linear smoothing application. *IEEE Trans. Inform. Theory*, 36(4), 836-847.

[4] Costa, O. L. V. (1994). Linear Minimum Mean Square Error Estimation for Discrete-Time Markovian Jump Linear Systems. *IEEE Trans. Automat. Contr.*, 39(8), 1685-1689.

[5] Costa, O. L. V., Fragoso, M. D., & Marques, R. P. (2005). Discrete-Time Markov Jump Linear Systems. Springer, London.

[6] Cox, H. (1964). Estimation of State Variables via Dynamic Programming. *Proc. 1964 Joint Automatic Control Conf.*, Stanford, California, 376-381.

[7] Fragoso, M. D., Costa, O. L. V., Baczynski, J., & Rocha, N. (2005). Optimal linear mean square filter for continuous-time jump linear systems. *IEEE Trans. Automat. Contr.*, 50(9), 1364-1369.

[8] Fraser, D. C., & Potter, J. E. (1969). The Optimum Linear Smoother as a Combination of Two Optimum Linear Filters. *IEEE Trans. Automat. Contr.*, AC-14(4), 387-390.

[9] Helmick, R. E., Blair, W. D., & Hoffman, S. A. (1995). Fixed-Interval Smoothing for Markovian Switching Systems. *IEEE Trans. Inform. Theory*, 41(6), 1845-1855.

[10] Jazwinski, A. H. (1966). Filtering for Nonlinear Dynamical Systems. *IEEE Trans. Automat. Contr.*, 11, 765-766.

[11] Lincoln, B., & Rantzer, A. (2006). Relaxing Dynamic Programming. *IEEE Trans. Automat. Contr.*, 51(8), 1249-1260.

[12] Ljung, L., & Kailath, T. (1976). Backwards Markovian Models for Second-Order Stochastic Processes. *IEEE Trans. Inform. Theory*, 22(4), 488-491.

[13] Mayne, D. Q. (1966). A Solution of the Smoothing Problem for Linear Dynamic Systems. *Automatica*, 4(2), 73-92.

[14] Meditch, J. S. (1967). Orthogonal Projection and Discrete Optimal Linear Smoothing. *J. SIAM Control*, 5(1), 74-89.

[15] Nakura, G. (2010). H_∞ Estimation for Linear Discrete-Time Markovian Jump Systems by Game Theoretic Approach. *Proceedings of the 42nd ISCIE International Symposium on Stochastic Systems Theory and Its Applications (SSS10)*, Okayama, Japan, 163-169.

[16] Nakura, G. (2011). An Approach to Noncausal Hybrid Estimation for Linear Continuous-Time Systems with Non-Gaussian Noises. *Proceedings of SICE Annual Conference 2011*, Tokyo, Japan, 979-984.

[17] Nakura, G. (2011). An Approach to Hybrid Smoothing for Linear Continuous-Time Systems with Non-Gaussian Noises. *Proceedings of the 43rd ISCIE International Symposium on Stochastic Systems Theory and Its Applications (SSS11)*, Shiga, Japan, 63-72.

[18] Rauch, H. E. (1963). Solutions to the Linear Smoothing Problem. *IEEE Trans., Automat., Contr.*, AC-8, 371-372.

[19] Rauch, H. E., Tung, F., & Striebel, C. T. (1965). Maximum Likelihood Estimates of Linear Dynamic Systems. *AIAA J.*, 3(8), 1445-1450.

[20] Sidhu, G. S., & Desai, U. B. (1976). New Smoothing Algorithms Based on Reversed-Time Lumped Models. *IEEE Trans. Automat. Contr.*, 21, 538-541.

[21] Sworder, D. D., & Rogers, R. O. (1983). An LQG Solution to a Control Problem with Solar Thermal Receiver. *IEEE Trans. Automat. Contr.*, 28, 971-978.

[22] Verghese, G., & Kailath, T. (1979). A Further Note on Backwards Markovian Models. *IEEE Trans. Inf. Theory*, IT-25(1), 121-124.

[23] Wall, J. E., Jr., Willsky, A. S., & Sandell, N. R., Jr. (1981). On the Fixed-Interval Smoothing Problem. *Stochastics*, 5, 1-41.

[24] Zhang, Q. (1999). Optimal Filtering of Discrete-Time Hybrid Systems. *J. Optim. Theory Appl.*, 100(1), 123-144.

[25] Zhang, Q. (2000). Hybrid Filtering for Linear Systems with Non-Gaussian Disturbances. *IEEE Trans. Automat. Contr.*, 45, 50-61.