JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 56, 195–210 (1976)

A Martingale Approach to State Estimation in Delay-Differential Systems

ARUNABHA BAGCHI

Department of Applied Mathematics, Twente University of Technology, P.O. Box 217, Enschede, The Netherlands

Submitted by Harold J. Kushner

A rigorous derivation of filtering and smoothing equations for linear stochastic systems with time delay is presented. The estimation equations are obtained in term of the innovation process of the problem under consideration. The method used is based on a representation theorem on Gaussian martingales.

1. INTRODUCTION

Many simple rigorous derivations of continuous time Kalman filter equations have been recently given [1–3]. The approach in [3] can be naturally extended to the smoothing problem [4]. In this paper, we follow the same idea to solve the estimation problem in delay-differential systems. The estimation problem in delay-differential systems has been solved by Kwakernaak [5] using the largely heuristic approach of Kalman–Bucy [6]. Many other derivations have since been given, notably by Koivo [7] and Kolmanovskii All these analyses were, however, not wholly convincing. In this paper, we present a rigorous derivation of the state estimation equations in delay-differential systems by exploiting the idea of Balakrishnan [9] of estimating one martingale from another.

2. PROBLEM FORMULATION

Consider the linear stochastic delayed system

$$x(t; \omega) = \sum_{i=0}^{k} \int_0^t A_i(\sigma)\, x(\sigma - h_i; \omega)\, d\sigma + \int_0^t B(\sigma)\, dW(\sigma; \omega) \qquad (2.1)$$

with

$$x(t; \omega) = 0 \qquad \text{for } t \leqslant 0$$

$$Y(t; \omega) = \sum_{i=0}^{k} \int_0^t C_i(\sigma)\, x(\sigma - h_i; \omega)\, d\sigma + \int_0^t D(\sigma)\, dW(\sigma; \omega) \tag{2.2}$$

where $x(t; \omega)$ and $Y(t; \omega)$ are n- and m-dimensional "state" and "output" functions, respectively; $W(t; \omega)$ is a p-dimensional Wiener process and $A_i(t)$, $B(t)$, $C_i(t)$, $D(t)$, $i = 0,..., k$ are appropriate dimensional matrix valued functions. Assume that these coefficient functions are all continuous and $D(t)\, D(t)^* > 0$ on the interval $[0, T]$ of interest where * denotes the transpose. The scalar quantities h_i with $0 = h_0 < h_1 < \cdots < h_k$ are the time delays which occur in the system. The existence of a solution to (1) has been proved in [10].

We shall study here the problem of estimating the state $x(t - \theta)$ with $\theta \geqslant 0$ from the observation of the signal $Y(\sigma; \omega)$, $0 \leqslant \sigma \leqslant t$. Let $\beta(s)$ be the smallest σ-algebra generated by the process $Y(\sigma; \omega)$, $0 \leqslant \sigma \leqslant s$ completed with respect to sets of measure 0 and $\beta(s-)$ the smallest σ-algebra generated by the process $Y(\sigma; \omega)$, $0 \leqslant \sigma < s$ completed with respect to sets of measure 0. Then since $Y(t; \omega)$ is continuous in t with probability one, $\beta(s) = \beta(s-)$.

Let $\hat{x}(t, \theta \mid \tau) = E[x(t - \theta) \mid \beta(\tau)]$ and denote $\hat{x}(t, \theta \mid t)$ by $\hat{x}(t, \theta)$. We know that $\hat{x}(t, \theta)$ is the best mean square estimate of $x(t - \theta)$ given the observation $Y(\sigma)$ for $\sigma \leqslant t$. The problem is called filtering if $\theta = 0$ and smoothing if $\theta > 0$. Contrary to the case of no time delay, the equation for the filtered state $\hat{x}(t, 0)$ involves some smoothed estimates and therefore, it is convenient to consider at the outset the general smoothing problem.

3. Innovation Process

Let us introduce the innovation process

$$Z_0(t; \omega) = Y(t; \omega) - \sum_{i=0}^{k} \int_0^t C_i(\sigma)\, \hat{x}(\sigma, h_i; \omega)\, d\sigma. \tag{3.1}$$

Then we have the following

Lemma 1. *$Z_0(t; \omega)$ is a Gaussian martingale. Moreover,*

$$\lim_{\Delta \to 0} (1/\Delta)\, E\left(\left(\int_t^{t+\Delta} dZ_0(s; \omega)\right)\left(\int_t^{t+\Delta} dZ_0(s; \omega)\right)^* \mid \beta(t)\right) = D(t)\, D(t)^* \quad \textit{in} \quad L_1 .$$

Proof.

$$Z_0(t;\omega) = Y(t;\omega) - \sum_{i=0}^{k} \int_0^t C_i(\sigma)\, x(\sigma - h_i;\omega)\, d\sigma$$

$$+ \sum_{i=0}^{k} \int_0^t C_i(\sigma)\, [x(\sigma - h_i;\omega) - \hat{x}(\sigma, h_i;\omega)]\, d\sigma$$

$$= \int_0^t D(\sigma)\, dW(\sigma;\omega) + \sum_{i=0}^{k} \int_0^t C_i(\sigma)\, [x(\sigma - h_i;\omega) - \hat{x}(\sigma, h_i;\omega)]\, d\sigma$$

so that

$$Z_0(t;\omega) - Z_0(s;\omega)$$
$$= \int_s^t D(\sigma)\, dW(\sigma;\omega) + \sum_{i=0}^{k} \int_s^t C_i(\sigma)\, [x(\sigma - h_i;\omega) - \hat{x}(\sigma, h_i;\omega)]\, d\sigma$$

implying that

$$E[((Z_0(t;\omega) - Z_0(s;\omega) \mid \beta(s))] = 0.$$

Let us use the notation

$$e(s, h_i;\omega) = x(s - h_i;\omega) - \hat{x}(s, h_i;\omega)$$

and observe that

$$E\left(\left\| \int_t^{t+\Delta} C_i(s)\, e(s, h_i;\omega)\, ds \right\|^2 \Big| \beta(t)\right) = O(\Delta^2) \quad \text{in} \quad L_1$$

since

$$E(\| e(s, h_i;\omega)\|^2) \leqslant E(\| x(s - h_i;\omega)\|^2)$$

and is bounded in $0 \leqslant s \leqslant T$ where $\|\cdot\|$ denotes the Euclidean norm in appropriate dimensional Euclidean space. Again

$$E\left(\left\| \int_t^{t+\Delta} D(s)\, dW(s;\omega) \right\|^2 \Big| \beta(t)\right) = O(\Delta) \quad \text{in} \quad L_1 .$$

Therefore,

$$\frac{1}{\Delta} E\left(\left(\int_t^{t+\Delta} dZ_0(s;\omega)\right)\left(\int_t^{t+\Delta} dZ_0(s;\omega)\right)^* \Big| \beta(t)\right)$$
$$= \frac{1}{\Delta}\int_t^{t+\Delta} D(s)\, D(s)^*\, ds + O(\Delta^{1/2}) \quad \text{in} \quad L_1$$

from which the result follows.

We next consider the key result we shall use in solving the estimation problem.

THEOREM 1. *Under the assumption that* $D(s)\,D(s)^* > 0$ *for every* s, $0 \leqslant s \leqslant T$, *we have for every* t, $0 \leqslant t \leqslant T$, $\beta(t)$ = *smallest* σ-*algebra generated by* $\{Z_0(s;\omega), s \leqslant t\}$.

To prove the theorem, we need the following:

LEMMA 2. *Under the assumption that* $D(s)\,D(s)^* > 0$ *for* $0 \leqslant s \leqslant T$, *we can write*

$$\sum_{i=0}^{k} C_i(t)\,\hat{x}(t, h_i;\omega) = \int_0^t h(t,s)\,dY(s;\omega) \qquad 0 \leqslant t \leqslant T \tag{*}$$

where

$$\int_0^T\int_0^t \| h(t,s)\|^2\,ds\,dt < \infty. \tag{**}$$

Proof. First notice that the solution of (2.1) can be written as

$$x(t;\omega) = \int_0^t \Phi(t,\tau)\,B(\tau)\,dW(\tau;\omega) \tag{3.2}$$

where the transition matrix $\Phi(t,\tau)$ satisfies the equation

$$\begin{aligned} (d/dt)\,\Phi(t,\tau) &= \sum_i A_i(t)\,\Phi(t-h_i,\tau) \qquad &&\text{for } t \geqslant \tau \\ \Phi(\tau,\tau) &= I && \\ \Phi(t,\tau) &= 0 &&\text{for } t < \tau. \end{aligned} \tag{3.3}$$

With the last condition, it is clear that we can write

$$x(t-h_i;\omega) = \int_0^t \Phi(t-h_i,\tau)\,B(\tau)\,dW(\tau;\omega).$$

Then for m-by-m function $f(\cdot)$

$$\int_0^t f(s)\,dY(s;\omega) = \int_0^t (f(s)\,D(s) + 1(s))\,dW(s;\omega)$$

where

$$1(s) = \sum_{i=0}^{k}\int_s^t f(\sigma)\,C_i(\sigma)\,\Phi(\sigma-h_i,s)\,d\sigma B(s).$$

Define the operator L by

$$Lf = g; \qquad g(s) = f(s)\, D(s) + \sum_{i=0}^{k} \int_s^t f(\sigma)\, C_i(\sigma)\, \Phi(\sigma - h_i\, , s)\, d\sigma B(s).$$

It then follows that

$$E\left(\left(\int_0^t f(s)\, dY(s; \omega)\right)\left(\int_0^t q(s)\, dY(s; \omega)^*\right)\right) = \int_0^t p(s)\, q(s)^*\, ds$$

where $L^*Lf = p$.

The operator L^* is defined by

$$L^*f = g; \qquad g(s) = f(s)\, D(s)^* + \sum_{i=0}^{k} \int_0^s f(\sigma)\, B(\sigma)^*\, \Phi(s - h_i\, , \sigma)^*\, d\sigma C_i(s)^*.$$

Next

$$E\left(\left(\sum_{i=0}^{k} C_i(t)\, x(t - h_i; \omega)\right)\left(\int_0^t q(s)\, dY(s; \omega)\right)^*\right)$$

$$= E\left(\left(\sum_{i=0}^{k} C_i(t) \int_0^t \Phi(t - h_i; s)\, B(s)\, dW(s; \omega)\right)\left(\int_0^t q(s)\, dY(s; \omega)\right)^*\right)$$

$$= \int_0^t u(s)\, q(s)^*\, ds$$

where

$$u = Lv \qquad \text{and} \qquad v(s) = \sum_{i=0}^{k} C_i(t)\, \Phi(t - h_i\, , s)\, B(s), \qquad 0 \leqslant s \leqslant t.$$

Therefore,

$$E\left(\left(\sum_{i=0}^{k} C_i(t)\, x(t - h_i; \omega) - \int_0^t h(t, s)\, dY(s, \omega)\right)\left(\int_0^t q(s)\, dY(s; \omega)\right)^*\right)$$

$$= \int_0^t c(s)\, q(s)^*\, ds$$

where $L^*Lh - L^*v = -c$. Hence (*) holds if and only if $c(s) = 0$ or, $L^*Lh = L^*v$.

But L^*L has a bounded inverse because $D(s)\, D(s)^* > 0$ and therefore, there exists a function $h(t, s)$ satisfying (*). A wellknown sufficient condition for the square integrability of $h(t, s)$ is [11, p. 165]

$$E \int_0^T \left\| \sum_{i=0}^{k} C_i(t)\, x(t - h_i; \omega) \right\|^2 dt < \infty.$$

This condition can be verified in a straightforward manner and (**) follows.

Proof of Theorem 1. For m-by-m square integrable matrix function $f(s)$ in $[0, t]$, we have

$$\int_0^t f(s)\, dZ_0(s; \omega) = \int_0^t f(s)\, dY(s; \omega) - \sum_{i=0}^{k} \int_0^t f(s)\, C_i \hat{x}(s, h_i; \omega)\, ds.$$

But by the above lemma,

$$\int_0^t f(s) \sum_{i=0}^{k} C_i(s)\, \hat{x}(s, h_i; \omega)\, ds = \int_0^t f(s) \int_0^s h(s, \sigma)\, dY(\sigma; \omega)\, ds$$
$$= \int_0^t \left(\int_\sigma^t f(s)\, h(s, \sigma)\, ds \right) dY(\sigma; \omega).$$

Define a new operator K by

$$Kf = g; \qquad g(\sigma) = f(\sigma) - \int_\sigma^t f(s)\, h(s, \sigma)\, ds, \qquad 0 \leqslant \sigma \leqslant t.$$

K differs from identity by a Volterra operator with square integrable kernel and therefore, has a bounded inverse. For m-by-m function $g(s)$ square integrable in $[0, t]$ we then have

$$\int_0^t g(s)\, dY(s; \omega) = \int_0^t f(s)\, dZ_0(s; \omega), \qquad f = K^{-1}g.$$

Hence the random variables

$$\int_0^t g(s)\, dY(s; \omega)$$

are measurable with respect to the smallest σ-algebra generated by $\{Z_0(s; \omega), s \leqslant t\}$ and so $\beta(t)$ is contained in that algebra.

The reverse inclusion is immediate from definition and the theorem is established.

4. Estimation Equations

We first state the basic representation theorem for Gaussian martingales on which our method of estimation is based. The detailed proof can be found in Balakrishnan [9, pp. 118–123].

THEOREM 2. *Let $Z_i(t; \omega)$ $i = 1, 2$ denote two martingales with respect to the same growing σ-algebra $\beta(t)$ and let*

$$E(\| Z_i(T; \omega) - Z_i(0; \omega)\|^2) < \infty, \qquad i = 1, 2. \tag{4.1}$$

Suppose that for i, j fixed, $0 \leqslant t \leqslant T$,

$$\lim_{\Delta \to 0}(1/\Delta)\, E((Z_i(t+\Delta) - Z_i(t))\,(Z_j(t+\Delta) - Z_j(t))^* \mid \beta(t)) = P_{ij}(t) \qquad (4.2)$$

where the convergence of the random variable on the left is in L_1 *and* $P_{ij}(t)$ *is a nonrandom function that is piecewise continuous,* $0 \leqslant t \leqslant T$.

Then for $0 \leqslant s \leqslant t \leqslant T$,

$$E\left(\left(\int_s^t dZ_i(\sigma;\omega)\right)\left(\int_s^t dZ_j(\sigma;\omega)\right)^* \Big| \beta(s)\right) = \int_s^t P_{ij}(\sigma)\, d\sigma.$$

COROLLARY *Assume that the martingales* $Z_i(s;\omega)$ *are Gaussian,* $0 \leqslant s \leqslant T$ *and the* $Z_i(0;\omega) = 0$, $i = 1, 2$. *Assume further that* (4.1) *holds and* (4.2) *holds for* $i = j = 2$ *and for* $i = 1, j = 2$. *Let* $\beta_2(t)$ *be the smallest* σ*–algebra generated by* $Z_2(s;\omega)$, $s \leqslant t$. *Then*

$$E[(Z_1(t;\omega) \mid \beta_2(t))] = \int_0^t \nu_{12}(s)\, dZ_2(s;\omega), \qquad 0 \leqslant t \leqslant T$$

where $\nu_{12}(s)$ *defined as the limit*

$$\nu_{12}(s) = \lim_{\epsilon \to 0} P_{12}(s)\,(P_{22}(s) + \epsilon I)^{-1} \qquad \text{a.e.} \qquad 0 \leqslant s \leqslant T$$

where I is the identity matrix of appropriate dimension.

We are now in the position to solve the estimation problem. For fixed t and θ, $\hat{x}(t, \theta \mid s)$ is a martingale in s with respect to the growing σ-algebra $\beta(s)$ and therefore, by corollary to Theorem 2,

$$\hat{x}(t, \theta \mid s) = \int_0^s K(t, \theta, \tau)\, dZ_0(\tau;\omega)$$

where $K(t, \theta, \tau) = P_{12}(t, \theta, \tau)\, P_{22}(\tau)^{-1}$ with P_{12} and P_{22} given by

$$P_{12}(t, \theta, \tau) = \lim_{\Delta \to 0}(1/\Delta)\, E((\hat{x}(t, \theta \mid \tau + \Delta) - \hat{x}(t, \theta \mid \tau))\,(Z_0(\tau+\Delta) - Z_0(\tau))^* \mid \beta(\tau))$$

and

$$P_{22}(\tau) = \lim_{\Delta \to 0}(1/\Delta)\, E((Z_0(\tau+\Delta) - Z_0(\tau))\,(Z_0(\tau+\Delta) - Z_0(\tau))^* \mid \beta(\tau)).$$

From Lemma 1, $P_{22}(\tau) = D(\tau)\, D(\tau)^*$.

Now $\hat{x}(t, \theta \mid s)$ being a martingale in s for fixed t and θ, we have from Doob [12, Theorem 4.3, p. 355]

$$\lim_{s \to t-} \hat{x}(t, \theta \mid s) = E[x(t-\theta) \mid \beta(t-)] = E[x(t-\theta) \mid \beta(t)] = \hat{x}(t, \theta)$$

and therefore, taking limit as $s \to t-$, we get

$$\hat{x}(t, \theta) = \int_0^t K(t, \theta, \tau)\, dZ_0(\tau; \omega) \tag{4.3}$$

where $K(t, \theta, \tau) = P_{12}(t, \theta, \tau)\,(D(\tau)\, D(\tau)^*)^{-1}$.

Now

$$\begin{aligned} &P_{12}(t, \theta, \tau) \\ &\quad = \lim_{\Delta \to 0}(1/\Delta)\, E((\hat{x}(t, \theta \mid \tau + \Delta) - \hat{x}(t, \theta \mid \tau))\,(Z_0(\tau + \Delta) - Z_0(\tau))^* \mid \beta(\tau)) \\ &\quad = \lim_{\Delta \to 0}(1/\Delta)\, E((x(t - \theta) - E(x(t - \theta) \mid \beta(\tau)) - (x(t - \theta) \\ &\qquad - E(x(t - \theta) \mid \beta(\tau + \Delta)))\,(Z_0(\tau + \Delta) - Z_0(\tau))^* \mid \beta(\tau)). \end{aligned}$$

$x(t - \theta) - E(x(t - \theta) \mid \beta(\tau + \Delta))$ is uncorrelated with $Y(\sigma; \omega)$, $\sigma \leqslant \tau + \Delta$ and hence with $Z_0(\sigma; \omega)$, $\sigma \leqslant \tau + \Delta$. It is also uncorrelated with (and hence independent of) the random variables generating $\beta(\tau)$. Therefore

$$E\left[(x(t - \theta) - E(x(t - \theta) \mid \beta(\tau + \Delta)))\left(\int_\tau^{\tau+\Delta} dZ_0(\sigma; \omega)\right)^* \mid \beta(\tau)\right] = 0.$$

Furthermore, $x(t - \theta) - E[x(t - \theta) \mid \beta(\tau)]$ being independent of the random variables generating $\beta(\tau)$,

$$P_{12}(t, \theta, \tau) = \lim_{\Delta \to 0}(1/\Delta)\, E\left((x(t - \theta) - E(x(t - \theta) \mid \beta(\tau)))\left(\int_\tau^{\tau+\Delta} dZ_0(\sigma; \omega)\right)^*\right).$$

Let us now assume that the state and observation noises are independent of each other. Mathematically, this means that $BD^* = 0$. Notice that

$$dZ_0(\sigma; \omega) = \sum_{i=0}^{k} C_i(\sigma)\,[x(\sigma - h_i; \omega) - \hat{x}(\sigma, h_i; \omega)]\, d\sigma + D(\sigma)\, dW(\sigma; \omega).$$

We have, therefore,

$$\begin{aligned} &P_{12}(t, \theta, \tau) \\ &\quad = \sum_{i=0}^{k} \lim_{\Delta \to 0}(1/\Delta)\, E\Bigl((x(t - \theta) - E(x(t - \theta) \mid \beta(\tau))) \\ &\qquad \times \left(\int_\tau^{\tau+\Delta} C_i(\sigma)\,[x(\sigma - h_i) - E(x(\sigma - h_i) \mid \beta(\sigma))]\right)^*\Bigr)\, d\sigma \\ &\qquad + \lim_{\Delta \to 0}(1/\Delta)\, E\left((x(t - \theta) - E(x(t - \theta) \mid \beta(\tau)))\left(\int_\tau^{\tau+\Delta} D(\sigma)\, dW(\sigma)\right)^*\right). \end{aligned}$$

If $\tau \geqslant t - \theta$, $x(t-\theta) - E(x(t-\theta) \mid \beta(\tau))$ is independent of $W(\sigma) - W(\sigma')$, $\tau \leqslant \sigma$, $\sigma' \leqslant \tau + \Delta$ and hence the second term vanishes. If $\tau < t - \theta$, writing

$$x(t-\theta) = \int_0^\tau \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma) + \int_\tau^{t-\theta} \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma)$$

and using the fact that $BD^* = 0$, the second term vanishes again.

We, therefore, get

$$\begin{aligned} P_{12}(t, \theta, \tau) &= \sum_{i=0}^{k} \lim_{\Delta \to 0} (1/\Delta)\, E\Big((x(t-\theta) - E(x(t-\theta) \mid \beta(\tau))) \\ &\qquad \times \Big(\int_\tau^{\tau+\Delta} C_i(\sigma)\, [x(\sigma - h_i) - E(x(\sigma - h_i) \mid \beta(\sigma))]\Big)^*\Big)\, d\sigma \\ &= \sum_{i=0}^{k} E[(x(t-\theta) - E(x(t-\theta) \mid \beta(\tau)))\, (x(\tau - h_i) \\ &\qquad - E(x(\tau - h_i) \mid \beta(\tau))^*]\, C_i(\tau)^*. \end{aligned}$$

It then follows that

$$\begin{aligned} K(t, \theta, \tau) &= \sum_{i=0}^{k} E[(x(t-\theta) - E(x(t-\theta) \mid \beta(\tau)))\, (x(\tau - h_i) \\ &\qquad - E(x(\tau - h_i) \mid \beta(\tau)))^*]\, C_i(\tau)^*\, (D(\tau) D(\tau)^*)^{-1}. \end{aligned} \tag{4.4}$$

Differentiability of $K(t, \theta, \tau)$ in t and θ follows from the smoothness property of the covariance of the process $x(t-\theta)$. Furthermore $K(t, \theta, \tau)$ is clearly seen to be a function of $t - \theta$ and τ. It then follows that

$$\frac{\partial K(t, \theta, \tau)}{\partial t} + \frac{\partial K(t, \theta, \tau)}{\partial \theta} = 0, \qquad 0 \leqslant \tau \leqslant t. \tag{4.5}$$

Now

$$\begin{aligned} x(t-\theta) &= \int_0^\tau \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma) + \int_\tau^{t-\theta} \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } \tau < t - \theta \\ &= \int_0^\tau \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma) \qquad \text{for } \tau \geqslant t - \theta \end{aligned}$$

so that

$$E(x(t-\theta) \mid \beta(\tau)) = E\Big(\int_0^\tau \Phi(t-\theta, \sigma) B(\sigma)\, dW(\sigma) \mid \beta(\tau)\Big).$$

Furthermore, for $\tau < t - \theta$

$$E\left[\left(\int_{\tau}^{t-\theta} \Phi(t-\theta, \sigma)\, B(\sigma)\, dW(\sigma)\right) (x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))^*\right] = 0.$$

Hence, we finally have

$$\begin{aligned} &E[(x(t-\theta) - E(x(t-\theta) \mid \beta(\tau)))\,(x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))^*] \\ &\quad = E\left[\left(\int_0^{\tau} \Phi(t-\theta, \sigma)\, B(\sigma)\, dW(\sigma) - E\left(\int_0^{\tau} \Phi(t-\theta, \sigma)\, B(\sigma)\, dW(\sigma) \mid \beta(\tau)\right)\right)\right. \\ &\qquad \left. \times\, (x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))^*\right) . \end{aligned} \tag{4.6}$$

We shall use (4.6) to derive the differential equation for $K(t, 0, \tau)$. Thus, using (4.5) and noting that $\tau \leqslant t$,

$$\begin{aligned} K(t, \theta, \tau) &= \sum_{i=0}^{k} E[(x(t) - E(x(t) \mid \beta(\tau)))\,(x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))^*] \\ &\qquad \times C_i(\tau)^*\,(D(\tau)\, D(\tau)^*)^{-1} \\ &= \sum_{i=0}^{k} E\left[\left(\int_0^{\tau} \Phi(t, \sigma)\, B(\sigma)\, dW(\sigma) - E\left(\int_0^{\tau} \Phi(t, \sigma)\, B(\sigma)\, dW(\sigma) \mid \beta(\tau)\right)\right)\right. \\ &\qquad \left. \times\, (x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))\ \right] C_i(\tau)^*\,(D(\tau)\, D(\tau)^*)^{-1} \end{aligned}$$

so that, using (3.3)

$$\begin{aligned} \frac{\partial K(t, 0, \tau)}{\partial t} &= \sum_{i=0}^{k} \sum_{j=0}^{k} A_j(t)\, E\left[\left(\int_0^{\tau} \Phi(t - h_j, \sigma)\, B(\sigma)\, dW(\sigma)\right.\right. \\ &\qquad \left. - E\left(\int_0^{\tau} \Phi(t - h_j, \sigma)\, B(\sigma)\, dW(\sigma) \mid \beta(\tau)\right)\right) \\ &\qquad \left. \times\, (x(\tau - h_i) - E(x(\tau - h_i) \mid \beta(\tau)))^*\right] C_i(\tau)^*\,(D(\tau)\, D(\tau)^*)^{-1} \\ &= \sum_{j=0}^{k} A_j(t) \sum_{i=0}^{k} E\left[\left(\int_0^{\tau} \Phi(t - h_j, \sigma)\, B(\sigma)\, dW(\sigma)\right.\right. \\ &\qquad \left. - E\left(\int_0^{\tau} \Phi(t - h_j, \sigma)\, B(\sigma)\, dW(\sigma) \mid \beta(\tau)\right)\right) (x(\tau - h_i) \\ &\qquad \left. - E(x(\tau - h_i) \mid \beta(\tau)))^*\right] C_i(\tau)^*\,(D(\tau)\, D(\tau)^*)^{-1} \\ &= \sum_{j=0}^{k} A_j(t)\, K(t, h_j, \tau) \end{aligned}$$

yielding the equation

$$\frac{\partial K(t, 0, \tau)}{\partial t} - \sum_{j=0}^{k} A_j(t)\, K(t, h_j, \tau) = 0 \qquad 0 \leqslant \tau \leqslant t. \tag{4.7}$$

From (4.5) and (4.7) the desired estimation equations can be readily obtained. Thus, from (4.3),

$$\hat{x}(t, \theta) = \int_0^t K(t, \theta, \tau)\, dZ_0(\tau; \omega).$$

Therefore,

$$d_t\hat{x}(t, \theta) = K(t, \theta, t)\, dZ_0(t; \omega) + \int_0^t \frac{\partial K(t, \theta, \tau)}{\partial t}\, dZ_0(\tau; \omega)\, dt$$

$$\frac{\partial \hat{x}(t, \theta)}{\partial \theta}\, dt = \int_0^t \frac{\partial K(t, \theta, \tau)}{\partial \theta}\, dZ_0(\tau; \omega)\, dt.$$

The last equation holds on verifying from (4.4) that

$$\int_0^t \left\| \frac{\partial}{\partial \theta} K(t, \theta, \tau)\, D(\tau) \right\|^2 d\tau < \infty.$$

Adding these two, we get using (4.5),

$$d_t\hat{x}(t, \theta) + \frac{\partial \hat{x}(t, \theta)}{\partial \theta}\, dt = K(t, \theta, t)\, dZ_0(t; \omega). \tag{4.8}$$

Similarly, we have

$$d_t\hat{x}(t, 0) = K(t, 0, t)\, dZ_0(t; \omega) + \int_0^t \frac{\partial K(t, 0, \tau)}{\partial t}\, dZ_0(\tau; \omega)\, dt$$

$$= K(t, 0, t)\, dZ_0(t; \omega) + \sum_{i=0}^{k} A_i(t) \int_0^t K(t, h_i, \tau)\, dZ_0(\tau; \omega)\, dt$$

by using (4.7)

$$d_t\hat{x}(t, 0) = K(t, 0, t)\, dZ_0(t; \omega) + \sum_{i=0}^{k} A_i(t)\, \hat{x}(t, h_i)\, dt$$

giving the following equation

$$d_t\hat{x}(t, 0) - \sum_{i=0}^{k} A_i(t)\, \hat{x}(t, h_i)\, dt = K(t, 0, t)\, dZ_0(t; \omega). \tag{4.9}$$

Relations (4.8) and (4.9) take the place of the single optimal filtering equation of the Kalman–Bucy problem. Equation (4.9) constitutes a boundary condition to the partial-differential equation (4.8). $\hat{x}(0, \theta) = 0$ for $\theta \geqslant 0$ gives the initial condition.

5. Covariance Equations

Define the covariance matrix of the estimate

$$P(t, \theta_1, \theta_2) = E[(x(t-\theta_1) - \hat{x}(t, \theta_1))(x(t-\theta_2) - \hat{x}(t, \theta_2))^*].$$

Then from (4.4)

$$K(t, \theta, \tau) = \sum_{i=0}^{k} P(\tau, \tau-(t-\theta), h_i)\, C_i(\tau)^* (D(\tau) D(\tau)^*)^{-1}. \qquad (5.1)$$

Simple calculation yields

$$\begin{aligned} P(t, \theta_1, \theta_2) & \\ &= E[x(t-\theta_1)\, x(t-\theta_2)^*] - E[\hat{x}(t, \theta_1)\, \hat{x}(t, \theta_2)^*] \\ &= E\left[\left(\int_0^{t-\theta_1} \Phi(t-\theta_1, \sigma) B(\sigma) dW(\sigma)\right)\left(\int_0^{t-\theta_2} \Phi(t-\theta_2, \sigma) B(\sigma)\, dW(\sigma)\right)^*\right] \\ &\quad - E\left[\left(\int_0^t K(t, \theta_1, \tau)\, dZ_0(\tau)\right)\left(\int_0^t K(t, \theta_2, \tau)\, dZ_0(\tau)\right)^*\right] \\ &= \int_0^{t-\theta_1} \Phi(t-\theta_1, \sigma) B(\sigma) B(\sigma)^* \Phi(t-\theta_2, \sigma)^*\, d\sigma \\ &\quad - \int_0^t K(t, \theta_1, \tau)(D(\tau) D(\tau)^*) K(t, \theta_2, \tau)^*\, d\tau. \end{aligned}$$

Therefore,

$$P(t, \theta_1, \theta_2) = Q(t, \theta_1, \theta_2) - R(t, \theta_1, \theta_2) \qquad (5.2)$$

where

$$\begin{aligned} Q(t, \theta_1, \theta_2) &= \int_0^{t-\theta_1} \Phi(t-\theta_1, \sigma) B(\sigma) B(\sigma)^* \Phi(t-\theta_2, \sigma)^*\, d\sigma \qquad \text{for } \theta_1 \geqslant \theta_2 \\ &= \int_0^{t-\theta_2} \Phi(t-\theta_1, \sigma) B(\sigma) B(\sigma)^* \Phi(t-\theta_2, \sigma)^*\, d\sigma \qquad \text{for } \theta_1 \leqslant \theta_2 \end{aligned}$$

and

$$R(t, \theta_1, \theta_2) = \int_0^t K(t, \theta_1, \tau)(D(\tau) D(\tau)^*) K(t, \theta_2, \tau)^*\, d\tau.$$

Notice that $Q(t, \theta_1, \theta_2)$ is a function of $t - \theta_1$ and $t - \theta_2$ and therefore,

$$\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial \theta_1} + \frac{\partial Q}{\partial \theta_2} = 0.$$

On the other hand,

$$\frac{\partial R}{\partial t} + \frac{\partial R}{\partial \theta_1} + \frac{\partial R}{\partial \theta_2}$$
$$= K(t, \theta_1, t)(D(t) D(t)^*) K(t, \theta_2, t)^*$$
$$+ \int_0^t \left[\frac{\partial K(t, \theta_1, \tau)}{\partial t} + \frac{\partial K(t, \theta_1, \tau)}{\partial \theta_1}\right] (D(\tau) D(\tau)^*) K(t, \theta_2, \tau)^* \, d\tau$$
$$+ \int_0^t K(t, \theta_1, \tau)(D(\tau) D(\tau)^*) \left[\frac{\partial K(t, \theta_2, \tau)}{\partial t} + \frac{\partial K(t, \theta_2, \tau)}{\partial \theta_2}\right]^* d\tau.$$

The second and third term vanishes by virtue of (4.5) and using (5.1), we get

$$\frac{\partial P(t, \theta_1, \theta_2)}{\partial t} + \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_1} + \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_2}$$
$$= - \sum_{i,j=0}^{k} P(t, \theta_1, h_i) C_i(t)^* (D(t) D(t)^*)^{-1} C_j(t) P(t, h_j, \theta_2)$$
$$\theta_1 \geqslant 0, \quad \theta_2 \geqslant 0. \tag{5.3}$$

Putting $\theta_2 = 0$ in (5.2),

$$P(t, \theta_1, 0) = Q(t, \theta_1, 0) - R(t, \theta_1, 0)$$

where

$$Q(t, \theta_1, 0) = \int_0^{t-\theta_1} \Phi(t - \theta_1, \sigma) B(\sigma) B(\sigma)^* \Phi(t, \sigma)^* \, d\sigma$$

and

$$R(t, \theta_1, 0) = \int_0^t K(t, \theta_1, \tau)(D(\tau) D(\tau)^*) K(t, 0, \tau)^* \, d\tau$$

$$\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial \theta_1} = \int_0^{t-\theta_1} \Phi(t - \theta_1, \sigma) B(\sigma) B(\sigma)^* \frac{d}{dt} \Phi(t, \sigma)^* \, d\sigma$$
$$= \sum_{i=0}^{k} \int_0^{t-\theta_1} \Phi(t - \theta_1, \sigma) B(\sigma) B(\sigma)^* \Phi(t - h_i, \sigma)^* \, d\sigma A_i(t)^*$$

while, using (4.5) again, we get

$$\frac{\partial R}{\partial t} + \frac{\partial R}{\partial \theta_1} = K(t, \theta_1, t)(D(t) D(t)^*) K(t, 0, t)^*$$
$$+ \int_0^t K(t, \theta_1, \tau)(D(\tau) D(\tau)^*) \frac{\partial K(t, 0, \tau)^*}{\partial t} \, d\tau.$$

Using (4.7) and with rearrangement of terms and then using (5.2) and (5.1), we get

$$\frac{\partial P(t, \theta_1, 0)}{\partial t} + \frac{\partial P(t, \theta_1, 0)}{\partial \theta_1}$$

$$= \sum_{i=0}^{k} P(t, \theta_1, h_i) A_i(t)^* \tag{5.4}$$

$$- \sum_{i,j=0}^{k} P(t, \theta_1, h_i) C_i(t)^* (D(t) D(t))^* C_j(t) P(t, h_j, 0); \qquad \theta_1 \geqslant 0.$$

Similarly, setting $\theta_1 = 0$ in $P(t, \theta_1, \theta_2)$, we obtain

$$\frac{\partial P(t, 0, \theta_2)}{\partial t} + \frac{\partial P(t, 0, \theta_2)}{\partial \theta_2}$$

$$= \sum_{i=0}^{k} A_i(t) P(t, h_i, \theta_2) \tag{5.5}$$

$$- \sum_{i,j} P(t, 0, h_i) C_i(t)^* (D(t) D(t)^*)^{-1} C_j(t) P(t, h_j, \theta_2); \qquad \theta_2 \geqslant 0.$$

Lastly, setting $\theta_1 = \theta_2 = 0$,

$$P(t, 0, 0) = \int_0^t \Phi(t, \sigma) B(\sigma) B(\sigma)^* \Phi(t, \sigma)^* \, d\sigma$$

$$- \int_0^t K(t, 0, \tau) (D(\tau) D(\tau)^*) K(t, 0, \tau)^* \, d\tau$$

so that

$$\frac{dP(t, 0, 0)}{dt} = \sum_{i=0}^{k} A_i(t) \Phi(t - h_i, \sigma) B(\sigma) B(\sigma)^* \Phi(t, \sigma)^* \, d\sigma$$

$$+ \sum_{i=0}^{k} \Phi(t, \sigma) B(\sigma) B(\sigma)^* \Phi(t - h_i, \sigma)^* A_i(t)^* \, d\sigma + B(t) B(t)^*$$

$$- \sum_{0}^{t} \frac{\partial K(t, 0, \tau)}{\partial t} (D(\tau) D(\tau)^*) K(t, 0, \tau)^* \, d\tau$$

$$- \int_0^t K(t, 0, \tau) (D(\tau) D(\tau)^*) \frac{\partial K(t, 0, \tau)^*}{\partial t} \, d\tau$$

$$- K(t, 0, t) (D(t) D(t)^*) K(t, 0, t).$$

Using (4.7) and with rearrangement of terms and then using (5.2) and (5.1), we get

$$\frac{dP(t,0,0)}{dt} = \sum_{i=0}^{k} A_i(t)\,P(t,h_i,0) + \sum_{i=0}^{k} P(t,0,h_i)\,A_i(t)^* + B(t)\,B(t)^*$$
$$- \sum_{i,j=0}^{k} P(t,0,h_i)\,C_i(t)^*\,(D(t)\,D(t)^*)^{-1}\,C_j(t)\,P(t,h_j,0). \quad (5.6)$$

Equation (5.3) is the partial differential equation satisfied by the covariance matrix of the estimate $P(t, \theta_1, \theta_2)$ with equation (5.4), (5.5), and (5.6) constituting the boundary conditions. With our choice of $x(t; \omega) = 0$ for $t \leqslant 0$, the initial condition is $P(0, \theta_1, \theta_2) = 0$ for $\theta_1 \geqslant 0$ and $\theta_2 \geqslant 0$.

6. Conclusion

The present paper derives the state estimation equations for linear stochastic systems with time delays both in the system and the observation. The complete solution of the problem involves uniqueness and stability considerations of the filter equations. Mitter and Vinter [13] discusses this question in the more extended framework of hereditary differential systems. Our solution is based on the concept of innovation and representation theorem on Gaussian martingales. When time delay occurs, the single filtering equation of Kalman–Bucy is replaced by a partial-differential equation together with a boundary condition. The covariance matrix of the estimate error satisfies a complicated partial-differential equation with three boundary conditions. The standard filtering and smoothing equations for linear systems with no time delay can be derived from the basic equation (4.4) after setting $A_i(t) = C_i(t) = 0$ for $i \geqslant 1$.

References

1. A. V. Balakrishnan, A martingale approach to linear recursive state estimation, *SIAM J. Control* **10** (1972), 754–766.
2. T. Kailath, A note on least squares estimations by the innovation method, *SIAM J. Control* **10** (1972), 477–486.
3. A. Bagchi, A new martingale approach to Kalman filtering, Memorandum no. 55, Department of Applied Mathematics, Twente University of Technology, Enschede, The Netherlands, October 1974.
4. A. Bagchi, A martingale approach to continuous time linear smoothing, *SIAM J. Appl. Math.* **28** (1975), 276–281.
5. H. Kwakernaak, Optimal filtering in linear systems with time delays. *IEEE Trans. Automatic Control* **AC-12** (1967), 169–173.

6. R. E. KALMAN AND R. S. BUCY, New results in linear filtering and prediction theory *J. Basic Eng. Trans. ASME* **83** (1961), 95–108.
7. A. J. KOIVO,Optimal estimator for linear stochastic systems described by functional differential equations, *Inf. Control* **19** (1971), 232–245.
8. V. B. KOLMANOVSKII, On the filtering of certain stochastic processes with lag, *Automation Remote Control* June (1974), 36–42; (Russian Original in *Automat. Telemek.* **35** (1974), 42–49.
9. A. V. BALAKRISHNAN, Stochastic differential systems 1, *in* "Lecture Notes in Economics and Mathematical Systems," No. 84, Springer–Verlag, Berlin, 1973.
10. K. ITO AND M. NISIO, On stationary solutions of a stochastic differential equation, *J. Math. Kyoto Univ.* **4** (1964), 1–75.
11. T. KAILATH, A view of three decades of linear filtering theory, *IEEE Trans. Inform. Theory* **IT-20** (1974), 146–181.
12. J. L. DOOB, "Stochastic Processes," Wiley, New York, 1953.
13. S. K. MITTER AND R. B. VINTER, Filtering for linear stochastic hereditary differential systems, *in* "Control Theory, Numerical Methods and Computer Systems Modelling" International Symposium, Rocquencourt, June 17–21, 1974); "Lecture Notes in Economics and Mathematical Systems," No. 107, Springer–Verlag, Berlin, 1975.