INFORMATION AND CONTROL 19, 232–245 (1971)

Optimal Estimator for Linear Stochastic Systems Described by Functional Differential Equations*

A. J. KOIVO

School of Electrical Engineering, Purdue University, Lafayette, Indiana 47907

The optimal state estimator for linear systems described by functional differential equations is constructed. The problem is solved by deriving equations for the unbiased estimation error as well as for its covariance. A functional of the estimation error at the terminal time is minimized by an optimal choice of the gain matrices of the estimator. It is accomplished by the application of the matrix maximum principle. The applicability of the optimal estimator is illustrated by an example.

INTRODUCTION

The optimal estimation of linear dynamical systems, the behavior of which is described by ordinary differential equations, is well known. The problem has also been investigated recently in systems which are governed by partial differential equations (Balakrishnan and Lions (1967)) and by differential equations with delayed arguments (Kwakernaak (1967)). The solution to the estimation problem in the former case is obtained by means of a least-squares fit procedure in an infinite-dimensional space. The solution to the filtering problem in the framework of differential-difference equations presented by Kwakernaak (1967) represents the best linear estimator.

The purpose here is to present the optimal solution to the estimation problem for dynamical systems described by functional differential equations. The problem is solved by assuming the mathematical description for the estimator. Equations for the unbiased estimation error are obtained. Then the equations for the covariance of the estimation error are derived. The use of the first- and second-order moments results in a deterministic optimization problem whereby a functional of the estimation error at the terminal time is minimized. The equations for determining the optimal

* This work was supported in part by NSF under Grant No. 18225 and AFOSR under Grant No. 69-1776.

gains of the estimator are derived. The optimal solution furnishes an estimator which is best for the estimators of the assumed structure. An example is presented to illustrate the operation of the optimal estimator.

Problem Statement

The state transition in stochastic linear systems with time delay is governed by

$$dx(t) = \left[A_0(t)x(t) + A_1(t)x(t-\tau) + \int_{-\tau}^{0} A_2(t,\sigma)x(t+\sigma)\,d\sigma\right] dt + D(t)\,d\xi_1(t) \tag{1}$$

where the n-vector $x_t = x(t+\sigma)$, $\sigma \in [-\tau, 0]$, signifies the state of the system in the Banach space $\mathscr{B}$ of continuous functions over the interval $[t_0 - \tau, T]$; $A_0(t)$, $A_1(t)$, and $D(t)$ are matrices of appropriate dimensions with continuous bounded elements; $A_2(t, \sigma)$ is a Kernel which is continuous in t and σ. Each component in the m-dimensional noise term $d\xi_1$ signifies an independent Brownian motion process whose statistical behavior is specified by

$$E\{(\xi_1^i(t_2) - \xi_1^i(t_1))(\xi_1^i(t_2) - \xi_1^i(t_1))\} = |\,t_2 - t_1\,| Q_1^i, \quad i = 1,..., m. \tag{2}$$

The observations are performed according to the following equation:

$$dz_1(t) = \left[H_0(t)x(t) + H_1(t)x(t-\tau) + \int_{-\tau}^{0} H_2(t,\sigma)x(t+\sigma)\,d\sigma\right] dt + d\eta_1(t), \tag{3}$$

where $H_0(t)$ and $H_1(t)$ are matrices of appropriate dimensions with bounded continuous elements; the Kernel $H_2(t, \sigma)$ is bounded and continuous with respect to both arguments. Each component in the r-dimensional term $d\eta_1(t)$ represents a Brownian motion process which is statistically independent upon the other components of $d\eta_1$ as well as upon the $d\xi_1$-process for all time; its statistical behavior is given by

$$E\{(\eta_1^i(t_2) - \eta_1^i(t_1))(\eta_1^i(t_2) - \eta_1^i(t_1))\} = |\,t_2 - t_1\,| R_1^i, \quad i = 1, 2,..., r. \tag{4}$$

It may be mentioned that the integral terms in Eqs. (1) and (3) may be encountered, for example, in equations describing the behavior of some electrical or hydraulic networks.

The problem is to determine the optimal estimate $\hat{x}_t = \hat{x}(t + \sigma)$ of the state $x_t = x(t + \sigma)$, $\sigma \in [-\tau, 0]$, given the measurements $z(t)$ up to time t so as to minimize a quadratic functional of the estimation error at the terminal time T, i.e.,

$$J = \left|\left|\left| [x(T) - \hat{x}(T)] + \int_{-\tau}^{0} [x(T + \sigma) - \hat{x}(T + \sigma)]\, d\sigma \right|\right|\right|^2, \tag{5}$$

where $T \geqslant t_0$ and $|||\, x\, |||^2 = E\{x'x\}$.

The problem is first reformulated as a distributed parameter system.

Reformulation of the Estimation Problem

It can formally be shown by the method of characteristics for partial differential equations that an alternative (formal) description for system (1) is

$$\frac{dx(t, 0)}{dt} = \sum_{i=0}^{2} A_i x(t, -\tau_i) + D(t)\xi(t), \tag{6}$$

$$\frac{\partial x(t, \theta)}{\partial t} = \frac{\partial x(t, \theta)}{\partial \theta}; \quad -\tau \leqslant \theta < 0, \tag{7}$$

where $x(t, \theta)$ is written for $x(t + \theta)$; $\tau_0 = 0$, $\tau_1 = \tau$, $\tau_2 = (\cdot)$; $A_0 = A_0(t)$, $A_1 = A_1(t)$, $A_2 x(t, \cdot) = \int_{-\tau}^{0} A_2(t, \sigma) x(t, \sigma)\, d\sigma$; and $x(t, \theta)$ is defined over $[t_0, T]x[-\tau, 0]$. The m-dimensional term $\xi(t)$ represents a white Gaussian zero-mean noise process with

$$E\{\xi(t_1)\xi'(t_2)\} = Q\,\delta(t_1 - t_2), \qquad Q > 0 \qquad \text{for all } t. \tag{7}$$

The observation Eq. (3) is rewritten formally as

$$z(t) = \sum_{i=0}^{2} H_i x(t, -\tau_i) + \eta(t), \tag{8}$$

where $H_0 = H_0(t)$, $H_1 = H_1(t)$ and $H_2 x(t, \cdot) = \int_{-\tau}^{0} H_2(t, \sigma) x(t, \sigma)\, d\sigma$. The measurement $z(t)$ is specified by $z(t)\, dt = dz_1(t)$; and the r-dimensional term $\eta(t)$ signifies a white Gaussian zero-mean process with

$$E[\eta(t_1)\eta'(t_2)] = R(t)\,\delta(t_1 - t_2), \qquad R > 0 \qquad \text{for all } t. \tag{9}$$

The problem now is to determine the optimal estimate $\hat{x}(t, \sigma)$ of the state vector $x(t, \sigma)$ for $t \in [t_0, T]$ and $\sigma \in [-\tau, 0]$ so as to minimize

$$J = \mathrm{Tr}\left[P(T, 0, 0) + \int_{-\tau}^{0} [P(T, 0, \alpha) + P(T, \alpha, 0)]\, d\alpha + \int_{-\tau}^{0} d\sigma \int_{-\tau}^{0} P(T, \sigma, \alpha)\, d\alpha\right], \tag{10}$$

where $P(T, \sigma, \alpha) = E\{[x(T, \sigma) - \hat{x}(T, \sigma)][x(T, \alpha) - \hat{x}(T, \alpha)]'\}$ for $-\tau \leqslant \sigma$, $\alpha \leqslant 0$; $\mathrm{Tr}[\cdot]$ signifies the trace of the matrix $[\cdot]$; and the measurements $z(t)$ are given over the range $t_0 \leqslant t \leqslant T$.

The problem is solved by assuming the mathematical form of the estimator. It will be specified so that the estimates are unbiased, and that the minimum of the functional of the estimation error is attained.

Form of the Unbiased Estimator

Since the systems (1) and (3) are linear, it is assumed that (i) the state transitions in the estimator are linear; (ii) the measurement $z(t)$ provides a linear action in the estimation procedure. Now let $\hat{x}(t, \theta) = \hat{x}(t + \theta)$ represent an estimate of $x(t, \theta) = x(t + \theta)$ based on the measurement $z(t)$ up to time t. Conditions (i) and (ii) are then satisfied by an estimator of the form

$$\frac{d\hat{x}(t, 0)}{dt} = \sum_{i=0}^{2} F_i \hat{x}(t, -\tau_i) + G(t, 0) z(t), \tag{11}$$

$$\frac{\partial \hat{x}(t, \theta)}{\partial t} = F_3 \frac{\partial \hat{x}(t, \theta)}{\partial \theta} + \sum_{k=4}^{6} F_k \hat{x}(t, -\tau_{k-4}) + G(t, \theta) z(t), \tag{12}$$

where the n-vector $\hat{x}(t, \theta)$ is defined over $[t_0, T]x[-\tau, 0]$; $F_2 \hat{x}(t, \cdot) = \int_{-\tau}^{0} F_2(t, \sigma)\hat{x}(t, \sigma)\, d\sigma$; $F_6 \hat{x}(t, \cdot) = \int_{-\tau}^{0} F_6(t, \sigma)\hat{x}(t, \sigma)\, d\sigma$; F_i, $i = 0, 1$ and F_j, $j = 2, 3, 4, 5, 6$, are bounded continuous matrices of appropriate dimensions on the domain of definition, the gain matrix $G(t, \theta)$ defined over $[t_0, T]x[-\tau, 0]$ operates on the observations. It is noted that Eq. (12) provides smoothed estimates.

The terms chosen for Eqs. (11) and (12) are sufficient for the construction of an estimator which generates unbiased estimates. Indeed, the plant matrices F_i, $i = 0, \ldots, 6$ will now be determined so that the estimator of the assumed structure generates unbiased estimates. The gain matrix $G(t, \theta)$ will be specified such that the minimum value of the functional given by Eq. (10) is attained.

The use of the estimator specified by Eqs. (10) and (11) to estimate the states of system (6) and (7) results in an estimation error

$$e(t, \theta) = x(t, \theta) - \hat{x}(t, \theta); \qquad -\tau \leqslant \theta \leqslant 0. \tag{13}$$

Equations which determine the evolution of the estimation error $e(t, \theta)$ in time t and θ can be obtained by subtracting Eqs. (11) and (12) from Eqs. (6) and (7), respectively, and substituting Eq. (8) for $z(t)$. It can then be observed by performing the expectation operation that unbiased estimates are achieved if and only if F_i and F_k satisfy the following relations almost everywhere on $t \in [t_0 - \tau, T]$ and $\theta \in [-\tau, 0]$:

$$F_i(t) = A_i(t) - G(t, 0)H_i(t); \qquad i = 0, 1, \tag{14}$$

$$F_2(t, \sigma) = A_2(t, \sigma) - G(t, 0)H_2(t, \sigma), \tag{15}$$

$$F_3(t, \theta) = I \text{ (identity matrix)}, \tag{16}$$

$$F_j(t, \theta) = -G(t, \theta)H_{j-4}(t); \qquad j = 4, 5, \tag{17}$$

$$F_6(t, \sigma) = -G(t, \sigma)H_2(t, \sigma). \tag{18}$$

Equations (11) and (12) for the estimator can now be written as follows:

$$\frac{d\hat{x}(t, 0)}{dt} = \sum_{i=0}^{2} A_i\hat{x}(t, -\tau_i) + G(t, 0)\left[z(t) - \sum_{k=0}^{2} H_k\hat{x}(t, -\tau_k)\right], \tag{19}$$

$$\frac{\partial\hat{x}(t, \theta)}{\partial t} = \frac{\partial\hat{x}(t, \theta)}{\partial\theta} + G(t, \theta)\left[z(t) - \sum_{k=0}^{2} H_k\hat{x}(t, -\tau_k)\right] \quad \text{for} \quad -\tau \leqslant \theta < 0. \tag{20}$$

Equations (19) and (20) establish the realization of unbiased filtering and smoothing estimates, respectively, based on the measurements $z(t)$ up to time t. The gain matrices $G(t, 0)$ and $G(t, \theta)$ for $\tau \leqslant \theta < 0$ will now be determined so that the functional of the estimation error given by expression (10) is minimized.

Minimum Variance Estimator

To determine the evolution of the variance of the estimation error, the dynamical equations for the estimation error $e(t, \theta)$ defined over $[t_0, T]x[-\tau, 0]$ are written by means of Eqs. (11)–(20):

$$\frac{de(t, 0)}{dt} = \sum_{i=0}^{2} A_i e(t, -\tau_i) + D(t)\xi(t) - G(t, 0) \sum_{k=0}^{2} H_k e(t, -\tau_k) - G(t, 0)\eta(t), \tag{21}$$

$$\frac{\partial e(t, \theta)}{\partial t} = \frac{\partial e(t, \theta)}{\partial \theta} - G(t, \theta) \sum_{k=0}^{2} H_k e(t, -\tau_k) - G(t, \theta)\eta(t), \tag{22}$$

where Eq. (22) specifies the smoothing error when the measurements are available up to time t. Equations (21) and (22) are then used in deriving the evolution of the covariance $P(t, \sigma, \alpha)$ of the estimation error over $[t_0 - \tau, T]x[-\tau, 0]^2$.

The equations for the covariance are obtained by forming $dP(t, 0, 0)/dt$, $\partial P(t, \theta_1, 0)/\partial t$, and $\partial P(t, \theta_1, \theta_2)/\partial t$. For example,

$$dP(t, 0, 0)/dt = \lim_{\Delta \to 0} E[e(t + \Delta, 0)\, e'(t + \Delta, 0) - e(t, 0) - e(t, 0)]/\Delta$$

where

$$e(t + \Delta, 0) = e(t, 0) + \frac{de(t, 0)}{dt} \Delta + 0(\Delta^2);$$

$$\begin{aligned}\frac{dP(t, 0, 0)}{dt} = \sum_{i=0}^{1} \{&[A_i(t) - G(t, 0)H_i(t)]P(t, -\tau_i, 0) \\ &+ P(t, 0, -\tau_i)[A_i'(t) - H_i'(t)G'(t, 0)]\} \\ &+ \int_{-\tau}^{0} \{[A_2(t, \sigma) - G(t, 0)H_2(t, \sigma)]P(t, \sigma, 0) \\ &+ P(t, 0, \sigma)[A_2'(t, \sigma) - H_2'(t, \sigma)G'(t, 0)]\}\, d\sigma \\ &+ D(t)Q(t)D'(t) + G(t, 0)RG'(t, 0).\end{aligned} \tag{23}$$

In a similar manner, the following equations are obtained:

$$\frac{\partial P(t, \theta_1, 0)}{\partial t} = \frac{\partial P(t, \theta_1, 0)}{\partial \theta_1} + \sum_{i=0}^{1} P(t, \theta_1, -\tau_i)[A_i'(t) - H_i'(t)G'(t, 0)]$$
$$+ \int_{-\tau}^{0} P(t, \theta_1, \alpha)[A_2'(t, \alpha) - H_2'(t, \alpha)G'(t, 0)]\, d\alpha$$
$$- \sum_{i=0}^{1} G(t, \theta_1)H_i(t)P(t, -\tau_i, 0)$$
$$- \int_{-\tau}^{0} G(t, \theta_1)H_2(t, \sigma)P(t, \sigma, 0)\, d\sigma + G(t, \theta_1)R(t)G'(t, 0); \tag{24}$$

$$\frac{\partial P(t, \theta_1, \theta_2)}{\partial t}$$
$$= \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_1} + \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_2} - G(t, \theta_1) \sum_{i=0}^{1} H_i(t)P(t, -\tau_i, \theta_2)$$
$$- \sum_{i=0}^{1} P(t, \theta_1, -\tau_i)H_i'(t)G'(t, \theta_2) - G(t, \theta_1) \int_{-\tau}^{0} H_2(t, \sigma)P(t, \sigma, \theta_2)\, d\sigma$$
$$- \int_{-\tau}^{0} P(t, \theta_1, \alpha)H_2'(t, \alpha)\, d\alpha G'(t, \theta_2) + G(t, \theta_1)R(t)G'(t, \theta_2). \tag{25}$$

The task is now to determine the optimal gains $G(t, 0)$ and $G(t, \theta_1)$ such that the minimum value of expression (10) is attained. Thus, the problem has been transformed to a deterministic optimization problem. It will be solved here by applying the matrix maximum principle (Athans, 1967).

The covariance matrices of the estimation error play now the role of the state. The Hamiltonian function for systems (10), (23), (24), and (25) is written by introducing an adjoint matrix $\Lambda(t, \theta_1, \theta_2)$ of dimension $n \times n$ defined on $[t_0, T] \times [-\tau, 0]^2$:

$$H(P, \Lambda, G, t) = \mathrm{Tr} \Big\{ \Lambda(t, 0, 0) \frac{dP(t, 0, 0)}{dt}$$
$$+ \int_{-\tau}^{0} \Lambda(t, \theta_1, 0) \frac{\partial P(t, \theta_1, 0)}{\partial t}\, d\theta_1$$
$$+ \int_{-\tau}^{0} \Lambda(t, 0, \theta_2) \frac{\partial P(t, 0, \theta_2)}{\partial t}\, d\theta_2$$
$$+ \int_{-\tau}^{0} d\theta_1 \int_{-\tau}^{0} d\theta_2 \Lambda(t, \theta_1, \theta_2) \frac{\partial P(t, \theta_1, \theta_2)}{\partial t} \Big\}. \tag{26}$$

Equations (23), (24), and (25) are then substituted for the derivatives. The application of the matrix maximum principle yields necessary conditions for the optimal solution: If $G^*(t, 0)$ and $G^*(t, \theta)$, $t \geqslant t_0$, $-\tau < \theta < 0$ are the optimal gain matrices that result in the minimum of expression (10) and in the trajectory $P(t, \sigma, \alpha)$, $t \geqslant t_0$, $-\tau \leqslant \sigma, \alpha \leqslant 0$, then there exists an adjoint matrix $\Lambda(t, \sigma, \alpha)$, $t \geqslant t_0$, $-\tau \leqslant \sigma, \alpha \leqslant 0$ such that conditions (23)–(25) and (27)–(35) hold:

$$\frac{d\Lambda(t, 0, 0)}{dt} = -[A_0' + H_0'G'(t, 0)]\Lambda(t, 0, 0) - \Lambda(t, 0, 0)[A_0 + G(t, 0)H_0] - 2\Lambda(t, 0, 0) - \int_{-\tau}^{0} \Lambda(t, \theta_1, 0)G(t, \theta_1)\, d\theta_1 H_0 - H_0' \int_{-\tau}^{0} G'(t, \theta_2)\Lambda(t, 0, \theta_2)\, d\theta_2 . \tag{27}$$

$$\frac{\partial\Lambda(t, \theta_1, 0)}{\partial t} = \frac{\partial\Lambda(t, \theta_1, 0)}{\partial\theta_1} - [A_0' - H_0'G'(t, 0)]\Lambda(t, \theta_1, 0) - \Lambda(t, \theta_1, 0) - \Lambda(t, 0, 0)[A_2(t, \theta_1) - G(t, 0)H_2(t, \theta_1)] - H_0' \int_{-\tau}^{0} G'(t, \theta_2)\Lambda(t, \theta_1, \theta_2)\, d\theta_2 + \int_{-\tau}^{0} \Lambda(t, \sigma, 0)G(t, \sigma)\, d\sigma\, H_2(t, \theta_1). \tag{28}$$

Equation (28) transposed is satisfied by $\Lambda(t, 0, \theta_1)$:

$$\frac{\partial\Lambda(t, \theta_1, \theta_2)}{\partial t} = \frac{\partial\Lambda(t, \theta_1, \theta_2)}{\partial\theta_1} + \frac{\partial\Lambda(t, \theta_1, \theta_2)}{\partial\theta_2} \tag{29}$$

$$- [A_2'(t, \theta_2) - H_2'(t, \theta_2)G'(t, 0)]\Lambda(t, \theta_1, 0) - \Lambda(t, 0, \theta_2)[A_2(t, \theta_1) - G(t, 0)H_2(t, \theta_1)] + \int_{-\tau}^{0} d\alpha[\Lambda(t, \alpha, \theta_2)G(t, \alpha)H_2(t, \theta_1) + H_2'(t, \theta_2)G'(t, \alpha)\Lambda(t, \theta_1, \alpha)]\, d\alpha; \tag{30}$$

$$\Lambda(T, 0, 0) = I; \quad \Lambda(T, \theta_1, 0) = \Lambda(T, 0, \theta_2) = I; \quad \Lambda(T, \theta_1, \theta_2) = I; \tag{31}$$

$$\Lambda(t, -\tau, \theta_2) = = \Lambda(t, 0, \theta_2)[A_1 - G(t, 0)H_1] - \int_{-\tau}^{0} \Lambda(t, \alpha, \theta_2)G(t, \alpha)\, d\alpha\, H_1 \qquad -\tau < \theta_2 \leqslant 0, \tag{32}$$

where $\theta_1, \theta_2 \in [-\tau, 0)$. Equations (27)–(32) must be satisfied by $\Lambda(t, \theta_1, \theta_2)$ almost everywhere on $[t_0, T] \times [-\tau, 0] \times [-\tau, 0]$.

The optimal gain matrix is determined by differentiating Eq. (26) with respect to $G(t, 0)$:

$$\Lambda(t, 0, 0)\left[G(t, 0)R - \sum_{l=0}^{1} P(t, 0, -\tau_l)H_l' - \int_{-\tau}^{0} P(t, 0, \sigma)H_2'(t, \sigma)\, d\sigma\right]$$

$$+\left[RG(t, 0) - \sum_{l=0}^{1} H_l P(t, -\tau_l, 0) - \int_{\tau}^{0} H_2(t, \sigma)P(t, \sigma, 0)\, d\sigma\right]\Lambda'(t, 0, 0)$$

$$+\int_{-\tau}^{0} d\theta_1 \Big\{\Lambda(t, \theta_1, 0)\Big[G(t, \theta_1)R - \sum_{l=0}^{1} P(t, \theta_1, -\tau_l)H_l'$$

$$- \int_{-\tau}^{0} P(t, \theta_1, \sigma)H_2'(t, \sigma)\, d\sigma\Big]$$

$$+\Big[RG(t, \theta_1) - \sum_{l=0}^{1} H_l P(t, -\tau_l, \theta_1)$$

$$- \int_{-\tau}^{0} H_2(t, \sigma)P(t, \sigma, \theta_1)\, d\sigma\Big]\, \Lambda(t, 0, \theta_1)\Big\} = 0. \tag{33}$$

In order to extract the expressions for $G(t, 0)$ and $G(t, \theta_1)$, one first observes that the terms in Eq. (30) for $d\Lambda(t, 0, 0)/dt$ appear symmetrical. Also, since $\Lambda(t, 0, \theta_1)$ satisfies an equation obtained by transposing Eq. (28), the terms in the equation $\partial[\Lambda(t, \theta_1, 0) + \Lambda(t, 0, \theta_1)]/\partial t$ appear as symmetric; and the same holds for Eq. (29). Because the terminal conditions (30) consist of symmetric matrices, it follows that $\Lambda(t, 0, 0)$ is symmetric, $\Lambda(t, \theta_1, 0) = \Lambda'(t, 0, \theta_1)$, and $\Lambda(t, \theta_1, \theta_2) = \Lambda'(t, \theta_2, \theta_1)$ for all t and $\theta_1, \theta_2 \in [-\tau, 0]$. Consequently, the terms multiplying $\Lambda(t, 0, 0)$ and $\Lambda'(t, 0, 0)$ as well as $\Lambda(t, 0, \theta_1)$ and $\Lambda(t, \theta_1, 0)$ in Eq. (33) can be combined. The resulting expression implies that

$$G^*(t, 0) = \left[\sum_{k=0}^{1} P(t, 0, -\tau_k)H_k' + \int_{-\tau}^{0} P(t, 0, \sigma)H_2'(t, \sigma)\, d\sigma\right] R^{-1}, \tag{34}$$

$$G^*(t, \theta) = \left[\sum_{k=0}^{1} P(t, \theta, -\tau_k)H_k' + \int_{-\tau}^{0} P(t, \theta, \sigma)H_2'(t, \sigma)\, d\sigma\right] R^{-1}, \tag{35}$$

on $t \geqslant t_0$ and $-\tau \leqslant \theta < 0$ are the candidates for the optimal gain matrices.

The covariance Eqs. (23)–(25) are rewritten for the optimal estimation by substituting expressions (34) and (35) for the optimal gain matrices. The results are:

$$\frac{dP(t,0,0)}{dt} = \sum_{i=0}^{2} [A_i P(t, -\tau_i, 0) + P(t, 0, -\tau_i) A_i']$$

$$- \sum_{i,k=0}^{2} P(t, 0, -\tau_i) H_i' R^{-1} H_k P(t, -\tau_k, 0) + D(t) Q D'(t), \quad (36)$$

where

$$A_2 P(t, \cdot, 0) = \int_{-\tau}^{0} A_2(t, \sigma) P(t, \sigma, 0)\, d\sigma$$

and

$$H_2 P(t, \cdot, 0) = \int_{-\tau}^{0} H_2(t, \sigma) P(t, \sigma, 0)\, d\sigma.$$

Moreover, $P(t, -\tau_i, 0) = P'(t, 0, -\tau_i)$. Similarly, one can write Eqs. (24) and (25) for the case that the optimal gains are employed:

$$\frac{\partial P(t, \theta_1, 0)}{\partial t} = \frac{\partial P(t, \theta_1, 0)}{\partial \theta_1} + \sum_{i=0}^{2} P(t, \theta_1, -\tau_i) A_i'$$

$$- \sum_{i,k=0}^{2} P(t, \theta_1, -\tau_k) H_k' R^{-1} H_i P(t, -\tau_i, 0), \quad (37)$$

$$\frac{\partial P(t, \theta_1, \theta_2)}{\partial t} = \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_1} + \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_2}$$

$$- \sum_{i,k=0}^{2} P(t, \theta_1, -\tau_k) H_k' R^{-1} H_i P(t, -\tau_i, \theta_2). \quad (38)$$

Equations (36)–(38) describe the evolution of the covariance of the estimation error. They correspond to the usual Riccati-equation encountered in conjunction with the estimation in linear systems described by ordinary differential equations. Equations (36)–(38) are nonlinear partial differential equations.

The optimal estimator of the structure given by Eqs. (11) and (12) is now completely specified by Eqs. (19) and (20), where the optimal gain matrices are determined by Eqs. (34) and (35). If $A_2(t, \sigma) \equiv 0$ and $H_2(t, \sigma) \equiv 0$, the estimator equations are the same as those obtained by Kwakernaak (1967) who derived the best linear estimator for linear systems exhibiting time delays.

It is emphasized that the derivation presented here is different. Moreover, the optimal estimator is derived for functional differential equations.

Example

Consider a scalar system defined by

$$\dot{x}(t) = -0.2x(t) - 0.4x(t-.3) - 0.2\int_{-.3}^{0} x(t+\sigma)\,d\sigma + \xi(t), \tag{44}$$

$$z(t) = x(t) + x(t-.3) + \int_{-.3}^{0} x(t+\sigma)\,d\sigma + \eta(t), \tag{45}$$

where $\xi(t)$ and $\eta(t)$ represent statistically independent white Gaussian zero-mean processes with $E[\xi^2(t)] = 0.2$ and $E[\eta^2(t)] = 0.2$; $z(t)$ represents the available data contaminated by noise. The problem is to determine the optimal estimate $\hat{x}(t)$ that establishes the minimum of the functional (10) from the available data $z(t)$; that is, the minimum of expression (5). The unbiased estimator is specified by Eqs. (19) and (20):

$$\frac{d\hat{x}(t,0)}{dt} = \sum_{i=0}^{2} a_i\hat{x}(t,-\tau_i) + g(t,0)\left[z(t) - \sum_{i=0}^{2}\hat{x}(t,-\tau_i)\right], \tag{46}$$

$$\frac{\partial\hat{x}(t,\theta)}{\partial t} = \frac{\partial\hat{x}(t,\theta)}{\partial\theta} + g(t,\theta)\left[z(t) - \sum_{i=0}^{2}\hat{x}(t,-\tau_i)\right], \tag{47}$$

where $a_0 = -0.2$, $a_1 = -0.4$, $a_2 = -0.2$, $\tau_0 = 0$, $\tau_1 = -0.3$, and $\tau_2 = (\cdot)$; i.e., $\hat{x}(t,\cdot) = \int_{-.3}^{0}\hat{x}(t+\sigma)\,d\sigma$. The optimal gain values are specified by Eqs. (34) and (35):

$$g(t,0) = 5\left[\sum_{i=0}^{2} P(t,0,-\tau_i)\right]; \qquad g(t,\theta) = 5\left[\sum_{i=0}^{2} P(t,\theta,-\tau_i)\right]. \tag{48}$$

The initial condition for Eqs. (46) and (47) is $\hat{x}(t_0,\sigma) = 0$ for $-\tau \leqslant \sigma \leqslant 0$. Equations (36)–(38), in this case, become:

$$\frac{dP(t,0,0)}{dt} = -2\left[0.2P(t,0,0) + 0.4P(t,-0.3,0) + 0.2\int_{-.3}^{0} P(t,\sigma,0)\,d\sigma\right]$$
$$-5\left[P(t,0,0) + P(t,-0.3,0) + \int_{-.3}^{0} P(t,\sigma,0)\,d\sigma\right]^2 + 0.2, \tag{49}$$

$$\frac{\partial P(t, \theta, 0)}{\partial t} - \frac{\partial P(t, \theta, 0)}{\partial \theta}$$
$$= -0.2P(t, \theta, 0) - 0.4P(t, \theta, -0.3) - 0.2 \int_{-.3}^{0} P(t, \theta_1, \sigma)\, d\sigma$$
$$- 5 \left[P(t, \theta, 0) + P(t, \theta, -0.3) + \int_{-.3}^{0} P(t, \theta, \alpha)\, d\alpha\right]$$
$$\times \left[P(t, 0, 0) + P(t, -0.3, 0) + \int_{-.3}^{0} P(t, \sigma, 0)\, d\sigma\right], \tag{50}$$

$$\frac{\partial P(t, \theta_1, \theta_2)}{\partial t} - \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_1} - \frac{\partial P(t, \theta_1, \theta_2)}{\partial \theta_2}$$
$$= -5 \left[P(t, \theta_1, 0) + P(t, \theta_1, -0.3) + \int_{-.3}^{0} P(t, \theta_1, \sigma)\, d\sigma\right]$$
$$\times \left[P(t, 0, \theta_2) + P(t, -0.3, \theta_2) + \int_{-.3}^{0} P(t, \alpha, \theta_2)\, d\alpha\right]. \tag{51}$$

Equations (46), (47), and (49)–(51) are solved numerically by using the method of characteristics, along which a discretization is performed relative to θ_1 and θ_2 in steps of $\Delta\theta_1 = \Delta\theta_2 = 0.1$. The operation of the optimal estimator is illustrated in Figs. 1(a) and 1(b), which display the evolution of

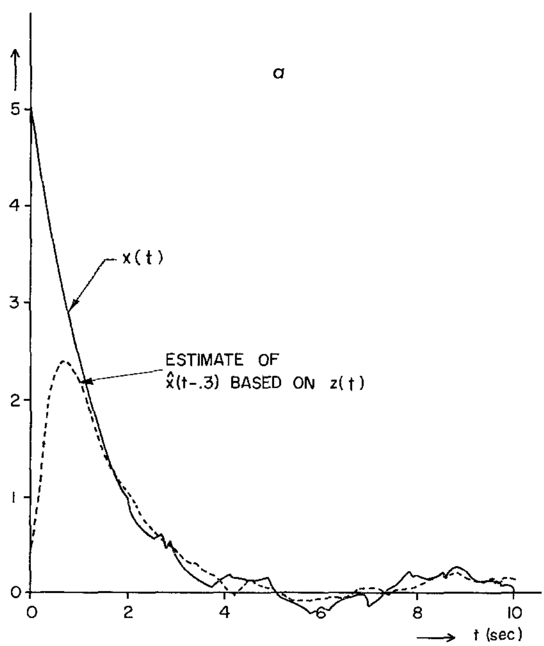

FIGURE 1(a)

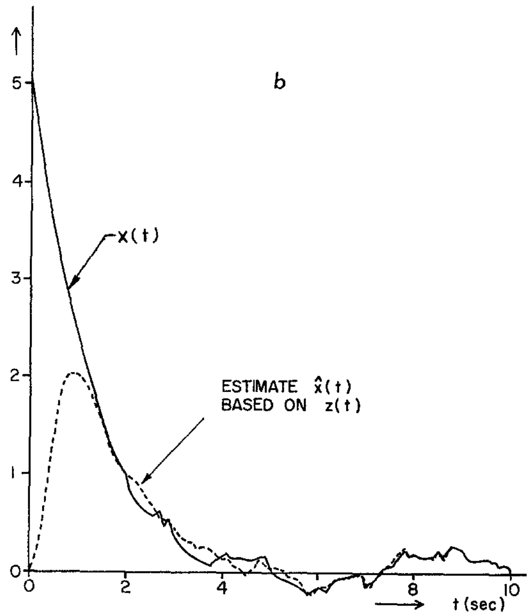

FIGURE 1(b)

the actual state $x(t)$ and the estimated states $\hat{x}(t, 0)$ (filtering solution) and $\hat{x}(t, -0.3)$ (smoothing solution) based on the measurement $z(t)$ available up to time t.

The computations of the solution to the optimal estimator problem in this scalar case demonstrates the complexity involved, particularly in a high-dimensional case.

CONCLUSIONS

The solution to the optimal estimation problem in systems described by functional differential equations is presented. The mathematical description of the estimator is assumed. The estimator is so specified that unbiased estimates are obtained. Then the gain matrix of the estimator is determined so that a functional of the estimation error is minimized. Partial differential equations which specify the evolution of the error covariance associated with the optimal gain are derived. The applicability of the solution is illustrated by an example.

RECEIVED: September 30, 1970

REFERENCES

ATHANS, M. (1967), The matrix maximum principle, *Information and Control* **11**, 592–606.

ATHANS, M., AND E. TSE (1967), A direct derivation of the Kalman-Bucy filter using the maximum principle, *IEEE Trans. Automatic Control* **AC-12**, 690–697.

BALAKRISHNAN, A. V., AND J. L. LIONS (1967), State estimation for infinite-dimensional systems, *J. Comp. Syst. Sci.* **1**, No. 4, 391–403.

DRIVER, R. D. (1960), "Delay-Differential Equations and Applications to a Two-Body Problem of Classical Electro-Dynamics," Report, University of Minnesota, Minneapolis, Minn.

HIRATSUKA, S., AND A. ICHIKAWA (1969), Optimal control of systems with transportation lags, *IEEE Trans. Automatic Control* **AC-14**, 237–247.

KUSHNER, H. J., AND D. I. BARNEA (1969), "On the Control of a Linear Functional-Differential Equation with Quadratic Cost," Report, Division of Applied Mathematics and Engineering, Brown University, Providence, R. I.

KWAKERNAAK, H. (1967), "Optimal Filtering in Linear Systems with Time Delays," *IEEE Trans. Automatic Control* **AC-12**, 169–173.

LAKSHMIKANTHAM, V., AND S. LEELA (1969), "Differential and Integral Inequalities, Theory and Applications," Academic Press, New York.