CDC
PAPER No. TP3-5

OPTIMAL FILTERING FOR LINEAR STOCHASTIC HEREDITARY DIFFERENTIAL SYSTEMS*

A. Bensoussan
Université de Paris IX, Paris, and I.R.I.A., 78-Rocquencourt, France.

M.C. Delfour
Centre de recherches mathématiques
Université de Montréal, Montréal 101, Canada.

S.K. Mitter
Electronic Systems Laboratory, M.I.T., Cambridge, Mass. 02139, U.S.A.

Abstract

This paper outlines the development of Filtering Theory for stochastic hereditary differential systems. It uses the model of hereditary differential systems developed by Delfour and Mitter and the results of Bensoussan on Estimation theory in Hilbert spaces.

1. Introduction

In this paper we outline the development of Filtering Theory for stochastic hereditary differential systems. To the authors' knowledge, the first paper in this field is the one of H. Kwakernaak[8], where simultaneously the smoothing and filtering problems for linear differential systems with multiple constant time delays are studied. In his paper, Kwakernaak[8] considers a class of linear filters, restricted to the extent that they must be equal to 0 at time t_0, and using a direct argument analogous to the one of Kalman and Bucy[7], he derives the equations of the filter and the covariance operator.

Here we use a different approach which allows us to justify the formal calculations. First, we consider the model of hereditary differential systems developed by Delfour and Mitter[4,5]. They prove that these systems can be modelled by an operational differential equation in a Hilbert space without delays, but with an unbounded operator. This model is very analogous to the one of J.L. Lions[9], already used in the study of partial differential equations. This analogy has already been used by Delfour and Mitter[6] to solve the problem of optimal control for hereditary differential systems. It is therefore very natural to use it for solving the Filtering problem. Of course, the question of modelling the noises in this framework raises new questions, since we are dealing with infinite dimensional spaces. The model we adopt is the one already used by A. Bensoussan[1,2] to solve the Filtering problem for linear distributed parameter systems. The noises are modelled as linear random functionals in a Hilbert space. By virtue of results on Estimation theory in Hilbert spaces (see Bensoussan[2]) it is possible to reduce our problem to a least square functional minimization problem. The problem is then a quadratic control problem for the operational differential equation of Delfour and Mitter, and thus can be solved using the method of these authors. The key point concerns the study of a Riccati equation in a Hilbert space, for which a direct approach is possible but new results of L. Tartar[12] more easily lead to an existence theorem.

Since our objective is mainly to show how the above approach is a useful tool to give a rigorous treatment of the problem, avoiding a lot of technical difficulties inherent in the direct approach, we did not attempt to present the most general results which can be expected. In particular, we did not consider delays in the observation process (which have been considered by Kwakernaak[8]) or distributed parameter systems with delays. For simplicity, we also did not consider the smoothing problem and we concentrated on the Filtering problem. Those problems are considered in Bensoussan[3], using an approach based upon the decoupling method of J.L. Lions[10,11].

* The work of the first and third authors was supported by NSF Grant GK 25 781 as part of a joint research programme between Decision and Control Sciences Group, Electronic Systems Laboratory, M.I.T., and I.R.I.A. The work of the second author was supported by the "Sous-Commission Franco-Québécoise à la recherche scientifique et technologique".

2. System description and formulation of the problem

2.1. Deterministic features

Let H, E and F be three Hilbert spaces. Let $a > 0$, $0 = \theta_0 > \theta_1 > \ldots > \theta_N = -a$ be real numbers. Let $B \in L^\infty(0,T;\mathcal{L}(E,H))$ and $C \in L^\infty(0,T;\mathcal{L}(H,F))$. Furthermore let A_{00} and A_i $(i=1,2,\ldots,N)$ belong to $L^\infty(0,T;\mathcal{L}(H))$, and A_{01} to $L^\infty(0,T;-a,0;\mathcal{L}(H))$. As in Delfour and Mitter[6], we introduce the space $\mathcal{L}^2(-a,0;H)$ (not to be confused with $L^2(-a,0;H)$) of all measurable maps from $[-a,0]$ into H which are square integrable, endowed with the seminorm

$$\|y\| = [|y(0)|^2 + \int_{-a}^{0} |y(\theta)|^2 d\theta]^{1/2}.$$

The quotient space of $\mathcal{L}^2(-a,0;H)$ by its linear subspace of all y such that $\|y\| = 0$ is a Hilbert space which is isometrically isomorphic to the product space $H \times L^2(-a,0;H)$. It will be denoted by $M^2(-a,0;H)$. For $f \in L^2(0,T;H)$, $\xi \in L^2(0,T;E)$ and $h \in M^2(-a,0;H)$, we consider the following linear hereditary differential system

$$\left.\begin{aligned} \frac{dx}{dt}(t) &= A_{00}(t)x(t) + \sum_{i=1}^{N} A_i(t)x(t+\theta_i) \\ &\quad + \int_{-a}^{0} A_{01}(t,\theta)x(t+\theta)d\theta \\ &\quad + B(t)\xi(t) + f(t) \\ x(\theta) &= h(\theta), \quad -a \le \theta \le 0 . \end{aligned}\right\} \tag{1}$$

Then the solution x in $[0,T]$ belongs to $AC^2(0,T;H)$, the space of absolutely continuous functions from $[0,T] \to H$, with a derivative in $L^2(0,T;H)$. The main result that we need is that the map

$$(h,\xi) \mapsto x : M^2 \times L^2(0,T;E) \to AC^2(0,T;H) \tag{2}$$

is affine and continuous.

For t in $[0,T]$, we can define the state $\tilde{x}(t)$ in $M^2(-a,0;H)$ by

$$\tilde{x}(t)(\theta) = \begin{Bmatrix} x(t+\theta), & t+\theta \ge 0 \\ h(t+\theta), & t+\theta < 0 \end{Bmatrix} . \tag{3}$$

For all h in $AC^2(-a,0;H)$, $\tilde{x}(t)$ is the unique solution in $AC^2(0,T;H)$ of

$$\left.\begin{aligned} \frac{d\tilde{x}}{dt}(t) &= \tilde{A}(t)\tilde{x}(t) + \tilde{B}(t)\xi(t) + \tilde{f}(t), \text{ a.e. in } [0,T] \\ \tilde{x}(0) &= h , \end{aligned}\right\} \tag{4}$$

where $\tilde{B}(t) \in \mathcal{L}(E,M^2)$ and $\tilde{f}(t) \in M^2$ are defined from B and f and $\tilde{A}(t)$ is a family of unbounded operators in M^2 defined from A_{00}, A_i and A_{01}.

2.2. Stochastic features

We now consider a noisy initial condition, that is

$$\left.\begin{aligned} x(0) &= h(0) + \zeta^0 \\ x(\theta) &= h(\theta) + \zeta^1(\theta), \; -a \le \theta < 0, \end{aligned}\right\} \tag{5}$$

where $\zeta = (\zeta^0,\zeta^1)$ belong to $M^2(-a,0;H)$. From now on ξ and ζ will be the noise at the input and the noise in the initial datum, respectively. We shall also assume an observation of the form

$$z(t) = C(t)x(t) + \eta(t), \tag{6}$$

where η represents the error in measurement. As in Bensoussan[2], $\{\zeta^0,\zeta^1,\xi,\eta\}$ will be modelled as a Gaussian linear random functional on the Hilbert space

$$\Phi = H \times L^2(-a,0;H) \times L^2(0,T;E) \times L^2(0,T;F) \tag{7}$$

with zero mean and covariance operator

$$\Lambda = \begin{bmatrix} P_0 & 0 & 0 & 0 \\ 0 & P_1(\theta) & 0 & 0 \\ 0 & 0 & Q(t) & 0 \\ 0 & 0 & 0 & R(t) \end{bmatrix}. \tag{8}$$

It will also be convenient to consider the covariance operator P in $\mathcal{L}(M^2)$ defined as follows

$$(Ph,\bar{h})_{M^2} = (P_0h^0,\bar{h}^0) + \int_{-a}^{0}(P_1(\theta)h^1(\theta),\bar{h}^1(\theta))d\theta. \tag{9}$$

In view of (2) and the properties of the image of a linear random functional under an affine continuous map, we can look at $x(t)$ as a Gaussian linear random functional on H (for any t), where the mean of $x(t)$, $\bar{x}(t)$, is a solution of equation (1) with $\xi=0$ and $\zeta=0$. But it is easy to check that the mean of $\tilde{x}(t)$, $\bar{\tilde{x}}(t)$, is obtained from the mean of $x(t)$ and the mean $\bar{h}(\theta)$ of $h(\theta)$ as follows

$$\bar{\tilde{x}}(t)(\theta) = \begin{cases} \bar{x}(t+\theta), & t+\theta\geq 0 \\ \bar{h}(t+\theta), & t+\theta<0 \end{cases}. \tag{10}$$

As a result $\bar{\tilde{x}}(t)$ is a solution of the state equation

$$\left.\begin{aligned} &\frac{d}{dt}\bar{\tilde{x}}(t) = \tilde{A}(t)\bar{\tilde{x}}(t) + \tilde{f}(t), \text{ a.e. in } [0,T], \\ &\bar{\tilde{x}}(0) = h \end{aligned}\right\} \tag{11}$$

and the covariance operator of $\tilde{x}(t)$, $\Gamma(t)$, is a "weak solution" of the equation

$$\left.\begin{aligned} &\frac{d\Gamma}{dt}(t) = \tilde{A}(t)\Gamma(t) + \Gamma(t)\tilde{A}(t)^* + \tilde{B}(t)Q(t)\tilde{B}(t)^*, \\ &\Gamma(0) = P. \end{aligned}\right\} \tag{12}$$

2.3. Formulation of the problem

For each T we want to determine the best estimator of the linear random functional $x(T)$ with respect to the linear random functional $z(s)$, $0 \leq s \leq T$. It is a linear random functional $\hat{x}(T)$ which can be obtained (see Bensoussan[2]) through the following control problem. We start with the deterministic system (1), the initial data (5) at time 0, where ξ and $\zeta = (\zeta^0,\zeta^1)$ are considered as control variables, and the following cost function

$$J_T(\xi,\zeta) = (P^{-1}\zeta,\zeta) + \int_0^T (Q(t)^{-1}\xi(t),\xi(t))dt + \int_0^T (R(t)^{-1}(z(t)-C(t)x(t)),z(t)-C(t)x(t))dt. \tag{13}$$

3. Solution of the control problem

At this point it is technically advantageous to work in the state space. For this purpose we redefine the cost function in terms of the state $\tilde{x}(t)$

$$\left.\begin{aligned} &\frac{d\tilde{x}}{dt}(t) = \tilde{A}(t)\tilde{x}(t) + \tilde{B}(t)\xi(t) + \tilde{f}(t) \\ &\tilde{x}(0) = h + \zeta \end{aligned}\right\} \tag{14}$$

$$J_T(\xi,\zeta) = (P^{-1}\zeta,\zeta) + \int_0^T (Q(t)^{-1}\xi(t),\xi(t))dt + \int_0^T (R(t)^{-1}(z(t)-\tilde{C}(t)\tilde{x}(t)),z(t)-\tilde{C}(t)\tilde{x}(t))dt, \tag{15}$$

where $\tilde{B}(t)\xi = (B(t)\xi,0)$, $\tilde{f}(t) = (f(t),0)$ and $\tilde{C}(t)h = C(t)h(0)$.

The pair $(\hat{\xi},\hat{\zeta})$ which minimizes the cost function over all (ξ,ζ) in $L^2(0,T;E) \times M^2$ is characterized by Euler's equation

$$\begin{aligned} &(P^{-1}\hat{\zeta},\zeta) + \int_0^T (Q(t)^{-1}\hat{\xi}(t),\xi(t))dt \\ &\quad + \int_0^T (R(t)^{-1}\tilde{C}(t)\hat{y}(t),\tilde{C}(t)y(t))dt \\ &\quad = \int_0^T (R(t)^{-1}\tilde{C}(t)y(t),z(t)-\tilde{C}(t)\bar{y}(t))dt, \ \forall\, \xi,\zeta, \end{aligned} \tag{16}$$

where

$$\left.\begin{aligned} &\frac{dy}{dt}(t) = \tilde{A}(t)y(t) + \tilde{B}(t)\xi(t) \\ &y(0) = \zeta, \end{aligned}\right\} \tag{17}$$

$$\left.\begin{aligned} &\frac{d\bar{y}}{dt}(t) = \tilde{A}(t)\bar{y}(t) + f(t) \\ &\bar{y}(0) = h. \end{aligned}\right\} \tag{18}$$

By introducing the adjoint system

$$\left.\begin{aligned} &\frac{d\hat{p}}{dt}(t) + \tilde{A}(t)^*\hat{p}(t) \\ &\qquad + \tilde{C}(t)^*R(t)^{-1}\tilde{C}(t)\hat{\tilde{x}}(t)-\tilde{C}(t)^*R(t)^{-1}z(t) = 0 \\ &\hat{p}(T) = 0, \end{aligned}\right\} \tag{19}$$

where $\hat{x}$ is the solution of (1) corresponding to $(\hat{\xi},\hat{\zeta})$ and $\hat{\tilde{x}}$ is the state constructed from $\hat{x}$ and $h+\hat{\zeta}$. Finally $(\hat{\xi},\hat{\zeta})$ is characterized as follows

$$\left.\begin{aligned} &\hat{\zeta} = -P\,\hat{p}(0) \\ &\hat{\xi}(t) = -Q(t)\tilde{B}(t)^*\hat{p}(t), \end{aligned}\right\} \tag{20}$$

and

$$\left.\begin{aligned} &\frac{d\hat{\tilde{x}}}{dt}(t) = \tilde{A}(t)\hat{\tilde{x}}(t) - \tilde{B}(t)Q(t)\tilde{B}(t)^*\hat{p}(t) + \tilde{f}(t) \\ &\hat{\tilde{x}}(0) = h - P\,\hat{p}(0). \end{aligned}\right\} \tag{21}$$

We can now introduce the Riccati equation of Delfour and Mitter[6]

$$\left.\begin{aligned} &\frac{d\Pi}{dt}(t) = \tilde{A}(t)\Pi(t) + \Pi(t)\tilde{A}(t)^* \\ &\qquad - \Pi(t)\tilde{C}(t)^*R(t)^{-1}\tilde{C}(t)\Pi(t) \\ &\qquad + \tilde{B}(t)Q(t)\tilde{B}(t)^* \\ &\Pi(0) = P \end{aligned}\right\} \tag{22}$$

and prove that $\hat{\tilde{x}}(T)$ is the solution of the following equation

$$\left.\begin{aligned}\frac{dy}{dt}(t) &= \tilde{A}(t)y(t) + \tilde{f}(t) \\ &\quad + \Pi(t)\tilde{C}^*(t)R(t)^{-1}(z(t)-\tilde{C}(t)y(t)) \\ y(0) &= h.\end{aligned}\right\} \quad (23)$$

The variable z makes equation (23) a stochastic equation, where the solution y(t) is to be interpreted as a linear random functional. Furthermore, if we introduce the estimation error

$$\varepsilon(t) = \hat{\tilde{x}}(t) - \tilde{x}(t),$$

then $\varepsilon(t)$ is a linear random functional with mean 0 and covariance operator $\Pi(t)$.

Remark 1. By virtue of the equivalence between (1) and (4) the equations of the filter and the covariance operator can be obtained in a form very analogous to the Kalman-Bucy filter.

Remark 2. The equivalence between the Filtering problem and the least square control problem is true only when P,Q(t) and R(t) are invertible. In fact, we can also obtain (22) and (23) when P and Q(t) are not invertible by using duality arguments as in Bensoussan[2].

Remark 3. As was pointed out in the introduction our main objective was to show how our approach can be used to obtain a rigorous treatment of the Filtering problem. The case where we have delays in the observation, namely

$$z(t) = C_0(t)x(t) + C_1(t)x(t-a),$$

is more interesting from the practical standpoint. This necessitates the use of the space $AC^2(-a,0;H)$ as state space rather than the space $M^2(-a,0;H)$. With that state space the map

$$\tilde{x}(t) \mapsto \tilde{C}(t)\tilde{x}(t) = C_0(t)x(t) + C_1(t)x(t-a) : AC^2(-a,0;H) \to E$$

is linear and continuous. To deal with this problem in the state space $M^2(-a,0;H)$ would require an unbounded operator $\tilde{C}(t)$. Henceforth it is clear that the method will be applicable and that the equations for the mean and the covariance will have the same form.

References

1. A. Bensoussan, Identification et filtrage, Cahiers de l'I.R.I.A., No. 1, Paris, Feb. 1969.
2. A. Bensoussan, Filtrage optimal des systèmes linéaires, Dunod, Paris, 1971.
3. A. Bensoussan, Filtrage des systèmes linéaires avec retard, I.R.I.A. report INF 7118/71027; Oct. 1971.
4. M.C. Delfour and S.K. Mitter, Hereditary differential systems with constant delays. I. General case, J. Differential Equations, 12 (1972).
5. M.C. Delfour and S.K. Mitter, Hereditary differential systems with constant delays. II. A class of affine systems and the adjoint problem. To appear in J. Differential Equations.
6. M.C. Delfour and S.K. Mitter, Controllability, observability and optimal feedback control of affine hereditary differential systems, SIAM J. on Control, 10 (1972), pp. 298-328.
7. R.E. Kalman and R.S. Bucy, New results in linear filtering and prediction theory, J. of Basic Engineering, Trans. ASME, Series D, Vol. 83 (1961), p. 95.
8. H. Kwakernaak, Optimal filtering in linear systems with time delays, IEEE Trans. on Automatic Control, AC-12 (1967), pp. 169-173.
9. J.L. Lions, Equations différentielles opérationnelles, Springer Verlag, Berlin, 1961.
10. J.L. Lions, Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles, Dunod, Paris, 1968; English translation, by S.K. Mitter, Springer Verlag, Berlin, 1971.
11. J.L. Lions, Contrôle optimal de systèmes avec retard, unpublished.
12. L. Tartar, to appear.