PMM U.S.S.R.,Vol.51,No.5,pp.607-612,1987
Printed in Great Britain

0021-8928/87 $10.00+0.00

ESTIMATION OF THE SOLUTIONS OF LINEAR STOCHASTIC INTEGRAL EQUATIONS*

V.B. KOLMANOVSKII and L.E. SHAIKHET

The problem of the optimal estimate (filtering) (optimal in the mean-square sense) of a partially observed process specified by a linear stochastic Volterra equation is solved. Numerical examples are given.

The problem of filtering for systems with discrete delay has been the subject of many investigations. Nevertheless, in a number of problems in control, aeroautoelasticity, the mechanics of a continuous medium etc., when constructing a model of different technical devices, equations arise in which Stieltjes integrals are used to describe the aftereffect /1-4/. These equations also contain systems with discrete delay as special cases. The approach used below to solve the filtering problem is based on an investigation of the dual problem of optimal control, which is well-known in the case of systems with discrete delay. An important difference in the case considered here is the fact that the dual problem is described by integral and not differential equations. This leads to the need to develop and use different conditions of the optimality of the dual problem.

1. Formulation of the problem. We will consider the problem of the optimal estimate (optimal in the mean-square sense) $m(T)$ of the vector $x(T)$, described by the equation

$$x(t)=x_0+\int_0^t [d_sK(t,s)]\,x(s)+\int_0^t \sigma_0(s)\,d\xi_0(s) \tag{1.1}$$
$$x(t)\in R_n,\ t\in[0,T]$$

The observations $y(t)\in R_m,\ 0\leqslant t\leqslant T$ satisfy the relation

$$dy(t)=A(t)x(t-h)dt+\sigma_1(t)d\xi_1(t),\ x(s)=0,\ -h\leqslant s<0 \tag{1.2}$$

The random vector x_0 has a non-degenerate Gaussian probability distribution with zero expectation, specified by the covariance matrix $D_0=\mathrm{M}\,x_0x_0'$ (the prime denotes transposition), and is independent of the mutually independent standard Wiener processes $\xi_0(t)$ and $\xi_1(t)$, whose dimensions are arbitrary. The presence of delay $h\geqslant 0$ in the measurement channel (1.2) is due to the finiteness of the time required to make the observations and to process the results of these observations. The matrices $\sigma_0,\ \sigma_1,\ A,\ K$ are piecewise continuous, and the elements of the matrix $K(t,s)$ have a bounded variation with respect to $s\in[0,T]$ uniformly with respect to $t\in[0,T]$. The first integral on the right-hand side of (1.1) is understood in the Stieltjes sense, while the stochastic integral is understood in the Ito sense. We will put $N_i=\sigma_i\sigma_i'$. The noise in the measurements (1.2) is non-degenerate, i.e. the matrix $N_1(t)$ is uniformly positive definite. For the further development it is convenient to assume that all the coefficients in (1.1) and (1.2) outside the interval $[0,T]$ are zero.

It is well-known that $m(T)$ is equal to the conditional expectation: $m(T)=\mathrm{M}(x(T)/G_T)$, where G_T is the minimum σ-algebra, generated by the process $y(t),\ 0\leqslant t\leqslant T$. In this case, in view of the fact that the combined probability density of the process $(x(t),y(t))$ is Gaussian, the following representation holds /5, 6/:

$$m(T)=\int_0^T u(t)\,dy(t) \tag{1.3}$$

(the deterministic matrix $u(t)$ remains to be defined).

2. The dual problem of optimal control. We will construct a problem of optimal control, the solution of which defines the kernel $u(t)$ of estimate (1.3). Since estimate (1.3) is optimal in the mean-square, the matrix $u(t)$ should minimize the functional

$$J(u)=\mathrm{M}\left|x(T)-\int_0^T u(t)\,dy(t)\right|^2 \tag{2.1}$$

*Prikl.Matem.Mekhan.,51,5,775-781,1987

which is the error of the estimation. We will transform this functional. We will represent the solution of Eq.(1.1) in the form

$$x(t)=\psi(t,0)\,x_0+\int_0^t \psi(t,s)\,\sigma_0(s)\,d\xi_0(s) \tag{2.2}$$

The deterministic matrix $\psi(t,s)$ is such that $\psi(t,t)=I$, where I is the unit matrix, $\psi(t,s)=0$ when $s>t$ and

$$\psi(t,s)=I+\int_{s+0}^{t}[d_\tau K(t,\tau)]\,\psi(\tau,s). \quad s\leqslant t$$

We substitute (2.2) and (1.2) into (2.1), giving

$$J(u)=\mathrm{M}\Big|\psi(t,0)\,x_0+\int_0^T\psi(T,s)\,\sigma_0(s)\,d\xi_0(s)-$$
$$\int_0^T u(t)\Big[A(t)\Big(\psi(t-h,0)\,x_0+\int_0^{t-h}\psi(t-h,s)\,\sigma_0(s)\,d\xi_0(s)\Big)\,dt+$$
$$\sigma_1(t)\,d\xi_1(t)\Big]\Big|^2$$

and we introduce the matrix

$$\alpha(s)=\psi(T,s)-\int_{s+h}^{T}u(t)\,A(t)\,\psi(t-h,s)\,dt,\quad \alpha(T)=I \tag{2.3}$$

In view of (2.3) and the properties of the stochastic Ito integrals, we obtain (tr stands for the trace of the matrix)

$$J(u)=\mathrm{M}\Big|\alpha(0)\,x_0+\int_0^T\alpha(s)\,\sigma_0(s)\,d\xi_0(s)-\int_0^T u(t)\,\sigma_1(t)\,d\xi_1(t)\Big|^2= \tag{2.4}$$
$$\operatorname{tr}\Big[\alpha(0)\,D_0\alpha'(0)+\int_0^T(\alpha(t)\,N_0(t)\,\alpha'(t)+u(t)\,N_1(t)\,u'(t))\,dt\Big]$$

Hence, the problem of determining the best estimate, in the mean-square sense, of the quantity $x(T)$ reduces to problem (2.3), (2.4) of the optimal control of the deterministic matrix process $\alpha(t)$ with the quadratic criterion of quality (2.4). After the optimal control $u_0(t)$ of problem (2.3), (2.4) is obtained, the best estimate $m(T)$ of the quantity $x(T)$ is defined by Eq.(1.3) for $u(t)=u_0(t)$, and the error of the estimation is equal to $J(u_0)$.

Notes. 1°. A similar method was employed in /7/ to solve the problem of optimal filtering for stochastic differential equations with delay. Another approach to the filtering of linear integral equations was proposed in /8/. Note that as in /9/ we can also write the dual control problem of the form (2.3), (2.4) for the case of several delays in the measurement channel, and also for filtered (coloured) or correlated excitations ξ_0 and ξ_1.

2°. If in Eqs.(1.1), (1.2) we put

$$h=0,\quad K(t,s)=\int_0^s B(\tau)\,d\tau \tag{2.5}$$

we obtain the formulation of the Kalman-Bucy filtering problem /5/. We will show how the equation of the Kalman-Bucy filter is obtained from the relations derived above. When conditions (2.5) are satisfied, the function ψ in representation (2.2) reduces to the fundamental matrix of the ordinary differential equation $x^{\cdot}=Bx$. Eq.(2.3) for the process $\alpha(t)$ takes the form

$$\alpha^{\cdot}(t)=-\alpha(t)\,B(t)+u(t)\,A(t),\quad 0\leqslant t\leqslant T,\quad \alpha(T)=I \tag{2.6}$$

The solution of the linear-quadratic problem (2.6), (2.4) shows that expression (1.3) is the Cauchy formula for the solution $m(t)$ of the equation

$$dm(t)=(B-DA_1)\,mdt+DA'N_1^{-1}dy(t),\quad m(0)=0$$

The quantity $J(u_0)=\operatorname{tr}D(T)$, where $D(t)=\mathrm{M}(x(t)-m(t))(x(t)-m(t))'$, and the Riccati matrix equation holds for $D(t)$.

3. Solution of the dual problem. To construct the matrix $u_0(t)$ we obtain the optimal-control problem (2.3), (2.4). Suppose first of all that $h \geqslant T$. Then, as follows from (2.4), the process $\alpha(s) = \psi(T, s)$ is also independent of the control $u(t)$. Then the optimal control $u_0(t) = 0$. Hence, in view of (1.3), (2.4) the optimal estimate $m(T)$ and the estimation error $J(u_0)$ satisfy the relations

$$m(T) = 0, \quad T \leqslant h \tag{3.1}$$
$$J(u_0) = \operatorname{tr}\left[\psi(T, 0) D_0 \psi'(T, 0) + \int_0^T \psi(T, t) N_0(t) \psi'(T, t)\, dt\right]$$

Suppose now that $h < T$. It can be seen from (2.3) that in this case when $s \in [T-h, T]$ the process $\alpha(s)$ is independent of the control and is determined by the relation $\alpha(s) = \psi(T, s)$, and when $s \in [0, T-h]$ it depends on the control $u(t)$ only when $t \in [h, T]$. Consequently, $u_0(t) = 0$ when $t \in [0, h)$, and the functional (2.4) can be represented in the form

$$J(u) = \operatorname{tr}\Big[\int_{T-h}^{T} \psi(T, s) N_0(s) \psi'(T, s)\, ds + \alpha(0) D_0 \alpha'(0) + \tag{3.2}$$
$$\int_0^{T-h} \alpha(s) N_0(s) \alpha'(s)\, ds + \int_h^T u(s) N_1(s) u'(s)\, ds\Big]$$

The solution of the problem of minimizing the functional (3.2) on trajectories of system (2.3) (see below) shows that $u_0(t)$ is the solution of the Fredholm integral equation

$$u_0(\tau + h) = \left[R(T, \tau) - \int_0^{T-h} u_0(t+h) A(t+h) R(t, \tau)\, dt\right] \times \tag{3.3}$$
$$A'(\tau + h) N_1^{-1}(\tau + h), \quad 0 \leqslant \tau \leqslant T - h$$
$$R(t, \tau) = Mx(t) x'(\tau) = \psi(t, 0) D_0 \psi'(\tau, 0) + \tag{3.4}$$
$$\int_0^{\min(t, \tau)} \psi(t, s) N_0(s) \psi'(\tau, s)\, ds$$

where ($R(t, \tau)$ is the correlation matrix of the process $x(t)$, defined by virtue of (2.2)). The solution of Eq.(3.3), which defines the optimal estimate, exists and is unique.

To derive Eq.(3.3) we will use the necessary condition for optimality /10/, which is as follows. Suppose $u_0(t)$ is the optimal control of problem (2.3), (3.2). We will introduce the control

$$u_\varepsilon(t) = \begin{cases} v, & t \in (\tau, \tau + \varepsilon], \quad h < \tau < \tau + \varepsilon < T \\ u_0(t), & t \in [h, T] \setminus (\tau, \tau + \varepsilon] \end{cases}$$

Here v is an arbitrary constant matrix of the same dimensions as u_0. Then, the necessary condition for optimality is the non-negativity of the functional

$$J_1(u_0) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [J(u_\varepsilon) - J(u_0)]$$

We will evaluate $J_1(u_0)$. Suppose $\alpha_0(s)$ and $\alpha_\varepsilon(s)$ are processes defined by relation (2.3) with the control $u_0(t)$ and $u_\varepsilon(t)$ respectively. On the basis of (2.3) we have

$$\frac{1}{\varepsilon}[J(u_\varepsilon) - J(u_0)] = \operatorname{tr}\Big[(\alpha_\varepsilon(0) + \alpha_0(0)) D_0 q_\varepsilon'(0) +$$
$$\int_0^{\tau+\varepsilon} (\alpha_\varepsilon(s) + \alpha_0(s)) N_0(s) q_\varepsilon'(s)\, ds + \int_\tau^{\tau+\varepsilon} (v N_1(s) v' - u_0(s) N_1(s) u_0'(s))\, ds\Big]$$
$$q_\varepsilon(s) = \frac{1}{\varepsilon}(\alpha_\varepsilon(s) - \alpha_0(s)) = -\frac{1}{\varepsilon} \int_\tau^{\tau+\varepsilon} (v - u_0(t)) A(t) \psi(t - h, s)\, dt$$

Hence it follows that

$$\lim_{\varepsilon \to 0} q_\varepsilon(s) = -(v - u_0(\tau)) A(\tau) \psi(\tau - h, s)$$

Consequently

$$J_1(u_0) = \mathrm{tr}\,[(v - u_0(\tau))\, N_1(\tau)\, (v - u_0(\tau))' + 2\,(u_0(\tau)\, N_1(\tau) - S(\tau)\, A'(\tau))\,(v - u_0(\tau))'] \geqslant 0$$

$$S(\tau) = \alpha_0(0)\, D_0 \psi'(\tau - h, 0) + \int_0^{\tau-h} \alpha_0(s)\, N_0(s)\, \psi'(\tau - h, s)\, ds$$

For this inequality to be correct for any $\tau \in [h, T]$ and an arbitrary constant matrix v it is necessary and sufficient that the optimal control have the form

$$u_0(\tau) = S(\tau)\, A'(\tau)\, N_1^{-1}(\tau) \tag{3.5}$$

Substituting (2.3) with $\alpha(s) = \alpha_0(s)$, $u(s) = u_0(s)$ into (3.5), and taking (3.4) into account, we obtain (3.3).

Hence, the correctness of Eq.(3.3) for the function $u_0(t)$ is established. The existence of the solution of Eq.(3.3) follows from the existence of the function $u_0(t)$ as the kernel in estimate (1.3).

We will obtain a certain representation for $u_0(t)$. Consider the matrix

$$B(T, \tau) = \mathrm{M}\,(x(T) - m(T))\,(x(\tau) - m(\tau))' = R(T, \tau) - B_1(T, \tau) - B_2(T, \tau)$$

$$B_1(T, \tau) = \mathrm{M} m(T)\, x'(\tau), \quad B_2(T, \tau) = \mathrm{M}\,(x(T) - m(T))\, m'(\tau)$$

Since

$$B_1(T, \tau) = \int_0^{T-h} u_0(t + h)\, A(t + h)\, R(t, \tau)\, dt$$

$$B_2(T, \tau) = \mathrm{M}\,[\mathrm{M}\,(x(T) - m(T)/G_T]\, m'(\tau) = 0$$

from (3.3) it follows that the representation

$$u_0(\tau) = B(T, \tau - h)\, A(\tau)\, N_1^{-1}(\tau), \quad h \leqslant \tau \leqslant T$$

of the solution of Eq.(3.3) is correct provided that the solution of this equation is unique.

We will prove the uniqueness of the solution of Eq.(3.3). Suppose $\Delta u(t)$ is the difference between two different solutions of Eq.(3.3). Then

$$\Delta u(\tau + h) + \int_0^{T-h} \Delta u(t + h)\, A(t + h)\, R(t, \tau)\, dt A'(\tau + h)\, N_1^{-1}(\tau + h) = 0$$

Multiplying this relation on the right by $N_1(\tau + h)\, \Delta u'(\tau + h)$ and then integrating with respect to $\tau \in [0, T - h]$ we obtain that the sum of the two integrals equals zero. Since they are both non-negative, each of them must be equal to zero. In particular, for almost all $\tau \in [h, T]$ we have $\Delta u(\tau)\, N_1(\tau)\, \Delta u'(\tau) = 0$. Hence, $\Delta u(\tau) = 0$ almost everywhere is $[h, T]$. Consequently, the solution of Eq.(3.3) is unique.

Hence, the optimal control of problem (2.3), (3.2) is zero when $s \in [0, h]$, and when $s \in [h, T]$ it is a unique solution of Eq.(3.3).

In certain cases the solution of Eq.(3.3) can be obtained in explicit form. Suppose, for example, that $N_0 = 0$, i.e. Eq.(1.1) does not contain random perturbations. We will seek a solution of Eq. (3.3) in the form

$$u_0(t) = \psi(T, 0) F \psi'(t - h, 0) A'(t) N_1^{-1}(t) \tag{3.6}$$

To determine the matrix F we will substitute (3.6) into (3.3). We obtain, taking (3.4) into account, when $N_0 = 0$, an equation whose solution is

$$F = (D_0^{-1} + Q)^{-1} \tag{3.7}$$

$$Q = \int_0^{T-h} \psi'(t, 0)\, A_1(t + h)\, \psi(t, 0)\, dt, \quad A_1 = A' N_1^{-1} A$$

The optimal estimate for $N_0 = 0$ is determined by expression (1.3) with kernel (3.6), (3.7). The corresponding error of the estimation is

$$J(u_0) = \mathrm{tr}\,[\psi(T, 0) F \psi'(T, 0)] \tag{3.8}$$

The problem of the dependence of the estimation error on the value of the delay h in the observations is of interest. Since the measurments (1.2) in the interval $[0, h]$ do not carry any information on process (1.1), the error of the estimation as a function of h does not decrease as h, $0 \leqslant h \leqslant T$ increases and is constant with respect to h when $h \geqslant T$. Eq.(3.8) confirms this. In fact, it follows from this equation for a constant matrix A, that

$$\partial J(u_0)/dh = \mathrm{tr}\,[\psi(T, 0)F\psi'(T-h, 0)A_1\psi(T-h, 0)F\psi'(T, 0)] \geqslant 0$$

Note that in situations differing from (1.1), (1.2), an increase in the delay h in the observation channel may lead to a reduction in $J(u_0)$, i.e. to an increase in the estimation accuracy /7/.

4. Numerical solution of the filtering problem. Different effective procedures for finding numerical solutions of Fredholm equations of the form (3.3) can be described in numerous handbooks (for example, /11/).

We will consdier the number solution of the problem of constructing an estimate that is optimal in the mean-square sense, for a partially observed process $(x(t), y(t))$, specified by the equations

$$x^{\cdot}(t) = a\int_0^t x(s)\,ds + \sigma_0\xi_0^{\cdot}(t); \quad x(0) = x_0; \quad t \in [0, T] \tag{4.1}$$

$$y^{\cdot}(t) = x(t-h) + \xi_1^{\cdot}(t); \quad x(s) = 0, \quad s < 0; \quad h \in [0, T] \tag{4.2}$$

Here a and σ_0 are aribtrary constants, x_0 is a Gaussian random quantity with zero mean and unit variance, and $\xi_0(t)$ and $\xi_1(t)$ are standard Wiener processes, independent of one another and of x_0.

Integrating (4.1) we obtain

$$x(t) = x_0 + a\int_0^t (t-s)\,x(s)\,ds + \sigma_0\xi_0(t) \tag{4.3}$$

Eq.(3.3) for system (4.3), (4.2) has the form

$$u_0(\tau+h) = R(T, \tau) - \int_0^{T-h} u_0(t+h)\,R(t, \tau)\,dt \tag{4.4}$$

The correlation function $R(t, \tau)$ is defined by relation (3.4) in which $D_0 = 1$, $N_0 = \sigma_0^2$ while the function $\psi(t, s)$ from representations (2.2), depending on the sign of the coefficient a, is equal to

$$\psi(t, s) = \begin{cases} \mathrm{ch}\sqrt{a}\,(t-s), & a \geqslant 0 \\ \cos\sqrt{|a|}\,(t-s), & a < 0 \end{cases}$$

Eq.(4.4) was solved by the method of successive approximations. We used as the initial approximation the analytical solution of Eq.(4.4) with $\sigma_0 = 0$

$$u_0(t) = \psi(T, 0)\,\psi(t-h, 0)\left[1 + \int_0^{T-h} \psi^2(s, 0)\,ds\right]^{-1}, \quad t \in [h, T]$$

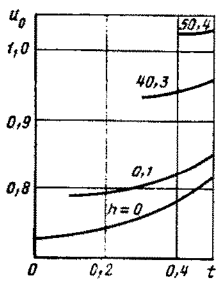

Fig.1

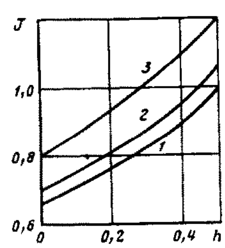

Fig.2

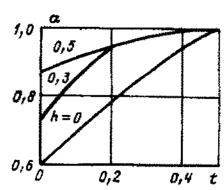

Fig.3

The calculations were carried out for constant $T = 1/2$ and values of the coefficients $a = -1, 0,$ and 1. In Fig.1 we show graphs of the function $u_0(t)$ for $t \in [h, T]$, $\sigma_0 = 0{,}1$ for different values of h. In Fig.2 we show the estimation error J as a function of the delay h for $a = 0$ and three values of σ_0: $\sigma_0 = 0$; $\sigma_0 = 0{,}1$; $\sigma_0 = 0{,}4$ (curves 1, 2 and 3 respectively). In

Fig.3 for $\sigma_0 = 0$ and different values of h we show graphs of the auxiliary controlled process $\alpha(t)$, $t \in [0, T]$, which depends on the control only when $t \in [0, T-h]$. The graphs are similar for the two other values of σ_0^2.

Note of extrapolation. The problem of extrapolation is to obtain the best estimate, in the mean-square sense, of the vector $x(T_1)$ at the instant $T_1 > T$ from measurements $y(t)$ in the interval $0 \leqslant t \leqslant T$. This problem reduces to the filtering problem. For this we put $A_0(t) = A(t)$ when $0 \leqslant t \leqslant T$ and $A_0(t) = 0$ when $t > T$. We will now consider the auxiliary problem of the filtering of the vector $x(T_1)$ (which satisfies Eq.(1.1) when $0 \leqslant t \leqslant T_1$) from the results of measurements (1.2) in the interval $[0, T_1]$, where in Eq.(1.2) instead of $A(t)$ we have $A_0(t)$. Then, in view of the independence of x_0, ξ_0, ξ_1, the solution of the auxiliary filtering problem will simultaneously also be a solution of the initial extrapolation problem. Hence, relations defining the solution of the extrapolation problem are obtained from the corresponding formulas of Sects.2 and 3, everywhere in which T is replaced by T_1 and $A(t)$ by $A_0(t)$.

REFERENCES

1. ASTAPOV I.S., BELOTSERKOVSKII S.M., KACHANOV B.O. and KOCHETKOV A.O., Systems of integro-differential equations describing the unsteady motion of bodies in a continuous medium, Diff. Urav. 18, 9, 1982.
2. DOGANOVSKII S.A. and IVANOV V.A., Delay Devices and their Application in Automatic Systems, Mashinostroyenie, Moscow, 1966.
3. VAN TREES H.L., Detection, estimation, and modulation theory. N.Y., Wiley, 1971.
4. KOLMANOVSKII V.B. and NOSOV V.R., Systems with neutral-type after-effect, Avtomatika i Telemekhanika, 1, 1984.
5. LIPTSER R.SH. and SHIRYAYEV A.N., Statistics of Random Processes, Non-linear Filtering and Allied Problems, Nauka, Moscow, 1974.
6. GIKHMAN I.I. and SKOROKHOD A.V., Stochastic Differential Equations, Nauka, Dumka Kiev, 1968.
7. CHERNOVSK'KO F.L. and KOLMANOVSKII V.B., Optimal control for Random Perturbations, Nauka, Moscow, 1978.
8. KLEPTSINA M.L. and VERETENNIKOV A.YU., On filtering and properties of conditional laws of Ito - Volterra processes, Statistics and control of stochastic processes. Steklow seminar. Moscow, 1984. N.Y., Optimization Software Inc. Publ. Divis. 1985.
9. KILMANOVSKII V.B. and MEIZENBERG T.L., Optimal estimates of the state of a system and some problems of controlling equations with after-effect, PMM, 41, 3, 1977.
10. SHAIKHET L.E., On otpimal control of Volterra's equations, Probl. of Control and Inform. Theory, 13, 3, 1984.
11. VERLAN A.F. and SIZIKOV V.S., Integral Equations, Methods, Algorithms and Programs, Nauk. Dumka, Kiev, 1986.

Translated by R.C.G.