PREDICTION AND FILTERING OF FUNCTIONALS OF SOLUTIONS OF NONLINEAR DIFFERENTIAL EQUATIONS WITH RANDOM FUNCTIONS

MATHEMATICS

1970

SovietRxiv

Abstract

Full Text

UDC 519.217

MATHEMATICS

A. D. SHATASHVILI

PREDICTION AND FILTERING OF FUNCTIONALS OF SOLUTIONS OF NONLINEAR DIFFERENTIAL EQUATIONS WITH RANDOM FUNCTIONS

(Presented by Academician Yu. V. Linnik on February 27, 1970)

In the present note we consider random processes with values in a finite-dimensional Euclidean space E, which are solutions of differential equations of the form

$$dx(t)/dt + af(t, x(t)) = \xi'(t) \quad (0 \le t \le a), \qquad x(0) = \xi(0) = 0, \qquad (1)$$

where $\xi'(t)$ is a Gaussian random process, a is a certain parameter, and $f(t, x)$ is a continuous function of both variables in the domain $[0, a] \times E$, taking its values in E. In E we denote by $(\cdot, \cdot)$ and $\|\cdot\|$, respectively, the scalar product and the norm.

The aim of the work is to establish formulas for optimal prediction and filtering for solutions of differential equations of the form (1).

We note that at present the theory of linear prediction and filtering has been completely developed for arbitrary random processes. At the same time, the determination of optimal estimates in prediction and filtering problems encounters very great difficulties connected with the need to involve all finite-dimensional distributions of the processes.

At present, results on optimal prediction and filtering are known only for certain classes of Markov processes ([1–4]). The solution of equation (1), generally speaking, is not Markovian. However, in the case when $\xi'(t)$ is a Gaussian process of "white-noise" type, $x(t)$ will be a diffusion Markov process, so that the class of processes under study may be regarded as an extension of the class of Markov diffusion processes for which the prediction and filtering problems were solved in papers ([1–4]).

Let $x(t)$ be a random process defined for $t \in [0, a]$ and taking values in the m-dimensional Euclidean space E. Denote by $\mathfrak{F}_t$ the σ-algebra of events generated by the quantities $x(s)$ for $s \leq t$, $t \in [0, a]$. If η is some random variable measurable with respect to $\mathfrak{F}_a$ (such variables are naturally called functionals of the process $x(t)$), then the optimal mean-square estimate of the quantity η with respect to the σ-algebra $\mathfrak{B} \subset \mathfrak{F}_a$ is the $\mathfrak{B}$-measurable random variable $\hat{\eta}$ for which $\mathbf{M}(\eta - \hat{\eta})^2$ assumes its minimal value. If, moreover, $\mathbf{M}\eta^2 < \infty$, then there exists a best estimate $\hat{\eta}$ of the form

$$\eta = \mathbf{M}(\eta/\mathfrak{B}). \tag{2}$$

In considering extrapolation and filtering problems, as η one takes quantities $\eta = h(x(t))$, and as the σ-algebra $\mathfrak{B}$, the σ-algebra $\mathfrak{F}_T^g$, generated by the values $g(x(s))$ for $s \leq T$, where $h(\cdot)$ and $g(\cdot)$ are certain measurable functions defined on E.

In deriving formulas for the best prediction and filter, the following lemma is proved.

Lemma. *Suppose that $x(t)$ and $\xi(t)$ are certain random processes with values in E, defined on the interval $[0, a]$. Let*

the measure μ_1, generated by the process $x(t)$, is absolutely continuous with respect to the measure μ_2, generated by the process $\xi(t)$, and let $\rho_\alpha(\cdot)$ denote the corresponding density. Let, further, $h(\cdot)$ and $g(\cdot)$ be some measurable functions defined on E. If we denote

$$\alpha(\xi, t) = \mathbf{M}\{h(\xi(t))\rho_t(\xi(\cdot))/\mathfrak{F}_T^{g^*}\}/\rho_T(\xi(\cdot)), \tag{3}$$

where $\mathfrak{F}_T^{g^*}$ denotes the σ-algebra generated by the values $g(\xi(s))$ for $s \leq T$, then the optimal estimate $\hat{h}(x(t))$ of the functional $h(x(t))$ has the form

$$\hat{h}(x(t)) = \alpha(x, t). \tag{4}$$

Suppose now that $\xi(t)$ is a Gaussian process with values in E, and that $g(\cdot)$ and $h(\cdot)$ are linear functionals on E (it is precisely such functionals that are considered in problems of prediction and filtering). Then the computation of the conditional expectation (3) reduces to the computation of an unconditional one in the following way.

Represent the Gaussian process $\xi(t)$ in the form

$$\xi(t) = \bar{\eta}_T(t) + \varepsilon_T(t), \tag{5}$$

where $\bar{\eta}_T(t)$ is the projection of the process $\xi(t)$ onto the space of random variables generated by the values $g(\xi(s))$, when $s \leq T$, $T \in [0, a]$, and $\varepsilon_T(t) = \xi(t) - \bar{\eta}_T(t)$ is a Gaussian process independent of the σ-algebra $\mathfrak{F}_T^{g^*}$. Therefore

$$\hat{h}(x(t)) = \frac{\mathbf{M}\{h(u(t)+\varepsilon_T(t))\rho_t^*(u(\cdot)+\varepsilon_T(\cdot))\}}{\rho_T(\xi(\cdot))}\Bigg|_{\substack{u(\cdot)=\bar{\eta}_T(\cdot),\\ \xi(\cdot)=x(\cdot)}}. \tag{6}$$

In [5] a formula was obtained for the density of the measure μ_1, corresponding to the solution $x(t)$ of equation (1), with respect to the measure μ_0, corresponding to the Gaussian process $\xi(t)$. The density has the following form:

$$\frac{d\mu_1}{d\mu_0}(x_0) = \exp\left\{-\alpha\int_0^a (b(t), dw(t)) - \frac{\alpha^2}{2}\int_0^a \|b(t)\|^2\, dt\right\}, \tag{7}$$

where the random function $b(t)$ and the Wiener process $w(t)$ take values in E and are defined in terms of the function $\mathfrak{F}(t,x)$ and the correlation matrix of the Gaussian process $\xi'(t)$ (see [5]).

Taking formula (7) into account, expression (6) can be rewritten as follows:

$$\hat{h}(x(t)) = \frac{\mathbf{M}\left[h(u(t)+\varepsilon_T(t))\exp\left\{-\alpha\int_0^t (b(s), dw(s)) - \frac{\alpha^2}{2}\int_0^t \|b(s)\|^2\, ds\right\}\right]}{\exp\left\{-\alpha\int_0^T (b(s), dw(s)) - \frac{\alpha^2}{2}\int_0^T \|b(s)\|^2\, ds\right\}}\Bigg|_{\substack{u(\cdot)=\bar{\eta}_T(\cdot),\\ \xi(\cdot)=x(\cdot)}}. \tag{8}$$

Formula (8) contains within it the formulas of optimal extrapolation and filtering, depending on the choice of $g(\cdot)$ and $h(\cdot)$ and on the relation between t and T. For example, if $g = h = x$ and $t > T$, then formula (8) gives the formula for the optimal prediction of the process $x(t)$, while if the vector x has the form $x = (g; h)$, and $t \leq T$, then formula (8) gives the formula for the optimal filter of one component x from another observed one.

Let the nonlinearity entering equation (1) be small, i.e., let α be a small quantity. Then the right-hand side of formula (8) can be expanded in powers of α.

Indeed, suppose that $\{\varphi_k(t)\}$ and $\{\lambda_k\}$ are the eigenfunctions and eigenvalues of the correlation matrix of the Gaussian process $\xi'(t)$. Let the representation (5) hold. Then the numerator and denominator of formula (8) are expanded in powers of α, and the optimal estimate of the functional $\hat{h}(x(t))$ has the form

$$\hat{h}(x(t)) = C_1 - \alpha C_2 + \alpha^2 C_3 + o(\alpha^2), \tag{9}$$

where the coefficients C_1, C_2, and C_3 are determined from the linear system of algebraic equations

$$C_1 = A_1, \qquad C_2 = A_2 - A_1 B_2,$$

$$C_3 = A_3 - 2B_2(A_2 - A_1B_2) - A_1B_3, \tag{10}$$

and the quantities A_1, A_2, A_3, B_2, and B_3 are found from the relations

$$A_1 = \mathbf{M}h(u(t) + \varepsilon_T(t)) \mid u(t) = \bar{\eta}_T(t), \tag{11}$$

$$A_2 = \mathbf{M}\left\{h(u(t) + \varepsilon_T(t)) \left[\sum_{k=1}^{\infty} \frac{1}{\lambda_k} \int_0^t \int_0^t (f(z, u(z) + \varepsilon_T(z)), \varphi_k(z))(v(s) + \varepsilon'_T(s), \varphi_k(s))\, dz\, ds\right]\right\} \Bigg| \begin{matrix} u(t) = \bar{\eta}_T(t \\ u(\cdot) = \bar{\eta}_T(\cdot \\ v(\cdot) = \bar{\eta}'_T(\cdot \end{matrix} \tag{12}$$

$$A_3 = \mathbf{M}\left\{h(u(t) + \varepsilon_T(t)) \left[A_2^2 - \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(\int_0^t (f(s, u(s) + \varepsilon_T(s)), \varphi_k(s))\, ds\right)^2\right]\right\} \mid u(\cdot) = \bar{\eta}_T(\cdot), \tag{13}$$

$$B_2 = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \int_0^T \int_0^T (f(t, x(t)), \varphi_k(t))(x'(s), \varphi_k(s))\, dt\, ds, \tag{14}$$

$$B_3 = B_2^2 - \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(\int_0^T (f(t, x(t)), \varphi_k(t))\, dt\right)^2. \tag{15}$$

Formula (9) shows that, for small nonlinearities, the deviation of the optimal prediction (filter) from the linear one has the same order as the order of the nonlinearity. Therefore, the use of optimal estimates instead of linear ones may lead to substantial improvements.

Donetsk Computing Center
of the Academy of Sciences of the Ukrainian SSR

Received
20 II 1970

CITED LITERATURE

1. R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Moscow, 1966.

2. A. N. Shiryaev, "Probability Theory and Its Applications," 11, 200 (1966).

3. A. N. Shiryaev, "Problems of Information Transmission," 3, 1966.

4. R. Sh. Liptser, "Probability Theory and Its Applications," 11, 3, 1528 (1966).

5. A. D. Shatashvili, *Probability Theory and Mathematical Statistics*, issue II, Kazan, 1970.

Note: Figure translations are in progress. See original paper for figures.