Proceedings of I. Vekua Institute
of Applied Mathematics
Vol. 63, 2013

ON THE MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS OF STOCHASTIC DIFFERENTIAL EQUATIONS

Gubeladze A., Sokhadze G.

Abstract. An evolution model with a random right-hand part and an unknown parameter of the nonlinear member is considered. This parameter is estimated by means of the maximum likelihood method, having first calculated the Radon–Nikodym density.

Keywords and phrases: Maximum likelihood estimation, random perturbation, consistency.

AMS subject classification (2010): 60G30, 62H12 .

Introduction

In the present paper we consider the Cauchy problem for evolution-type quasilinear differential equations with random additive perturbation. In the general case, perturbation can be a random process whose distribution is smooth in the sense that it possesses the logarithmic derivative in the direction of a sufficiently rich set of vectors. Such processes include, in particular, Gaussian random processes. In the Gaussian case, when we have the so-called white noise, such equations can be interpreted as stochastic differential equations. When the linear term contains an unknown parameter, it is of practical interest to estimate this parameter by observations of the equation solution trajectory and thereby to estimate the nonlinearity degree of the considered model. Such an idea was for the first time realized in [1]. Confidence intervals indicating the reliability of estimates were obtained also in that paper.

To obtain the estimate of an unknown parameter in the nonlinear member, we apply a maximum likelihood estimate. For this we preliminarily establish the conditions under which distributions for solutions of such equations are absolutely continuous. In doing so, we use the result of [2].

Preliminary result

Let $\{\Omega, \mathfrak{G}, P\}$ be a probabilistic space. Let us consider the Hilbert space $H_+ \subset H \subset H_-$ with Hilbert–Schmidt embeddings. For scalar products and norms we will use the usual notation with indexes of the corresponding spaces. We denote by $\mathcal{L}_2^- = \mathcal{L}_2([0,a], H_-)$ the space of functions defined on the closed interval $[0,a]$ with values in H_-. In an analogous manner we define the spaces $\mathcal{L}_2^+$ and $\mathcal{L}_2$; these are spaces of functions defined on $[0,a]$, having values in H_+ and H, respectively, square integrable with respect to the respective norms. Thus we obtain a different triple of Hilbert spaces

$$\mathcal{L}_2^+ \subset \mathcal{L}_2 \subset \mathcal{L}_2^- \tag{1}$$

with Hilbert–Schmidt embeddings.

In the space H_-, consider the evolution-type differential equation

$$\frac{dy}{dt} - A(t)y(t) + \theta f(t, y(t)) = \xi(t), \tag{2}$$

with the initial conditions

$$0 \le t \le T, \;\; y(0) = \xi(0) = 0 \pmod{P}. \tag{3}$$

(i) $A(t)$ is a linear, maybe unbounded operator with a dense definition domain $\mathcal{D}(A) \subset H$ not depending on t. Besides, let $A(t)$ be the generating operator of the evolution family $u(t,s)$, which is strongly continuous with respect to the set of variables (see [3]);

(ii) The function $f(t,y)$ is defined on $[0,T] \times H_-$, is bounded and takes its values in H_+. Let there exist the partial derivative $f'_y(t,y)$, which is the Hilbert–Schmidt operator and satisfies the inequality

$$\sup_{y\in\mathcal{D}(A)} \int_0^T [\![f'_y(\tau, y)]\!]^2_{H_-} \, d\tau \le \left\{ \int_0^T \int_0^T [\![u(t,\tau)]\!]^2_{H_-} \, d\tau \, dt \right\}^{-1}, \tag{4}$$

where $[\![K]\!]_M$ denotes, here and in the sequel, the operator norm for the linear operator acting in the Hilbert space M.

(iii) $\xi = \xi(t)$ is a random process with finite second moment on $[0,T]$ and with values in H_-, whose nearly all trajectories are continuous. Besides, let there exist a bounded and continuous function $\lambda(t,x)$: $[0,T] \times H_- \to H_-$, such that for any functional $\varphi \in C^1(\mathcal{L}_2^-)$ and any function $h(t) : [0,T] \to H_+$ from $\mathcal{L}_2^+$ we have

$$E \int_0^T \big([\varphi'(\xi)](t), h(t) \big)_H \, dt = E\varphi(\xi) \int_0^T \big(\lambda(t, \xi(t)), h(t) \big)_H \, dt. \tag{5}$$

Under the above conditions there exists a solution $y(t)$ of the evolutionary problem (2), (3). The distribution of this process is concentrated in $\mathcal{L}_2^-$. We denote it by μ_y.

In equation (2), θ is a real but unknown parameter. Using the observation of the solution $\tilde{y}(t)$ of problem (2), (3), we are to construct the estimate of the maximal likelihood $\widehat{y}(t)$.

Let us rewrite problem (2), (3) in terms of the triple of spaces (1). In the space $\mathcal{L}_2^-$, the expression $\frac{d}{dt} - A(t)$ generates, generally speaking, the unbounded linear operator $\mathcal{A}$:

$$(\mathcal{A}\varphi)(t) = \frac{d\varphi}{dt} - A(t)\varphi(t), \;\; \varphi(0) = 0,$$

with definition domain $\mathcal{D}(\mathcal{A})$, which by the condition (i) is densely embedded into $\mathcal{L}_2^-$. The inverse operator to $\mathcal{A}$ exists and has the form

$$(\mathcal{A}^{-1}\varphi)(t) \overset{def}{=} (U\varphi)(t) = \int_0^t u(t,s)\varphi(s)\,ds. \tag{6}$$

The function $f(t,y) : [0,T] \times H_- \to H_+$ generated, in $\mathcal{L}_2^-$, the nonlinear operator F :

$$[F(\varphi)](t) = f(t, \varphi(t))$$

and by the condition (ii) we have $F : \mathcal{L}_2^- \to \mathcal{L}_2^+$. Moreover, $F(\varphi)$ is differentiable and

$$[\![F'(\varphi)]\!]_{\mathcal{L}} < 1.$$

The random process $\xi(t)$ in $\mathcal{L}_2^-$ can be treated as a random element that has the logarithmic derivative along constant directions $\mathcal{L}_2^+$ of the form $\lambda(\xi)$. As seen from (5), if we write it in terms of the space $\mathcal{L}_2$, then

$$E(\varphi'(\xi), h)_{\mathcal{L}_2} = E\varphi(\xi)(\lambda(\xi), h)_{\mathcal{L}_2}, \quad h \in \mathcal{L}_2^+.$$

Along with problem (2), (3), let us consider the linearized problem

$$\frac{dx(t)}{dt} - A(t)x(t) = \xi(t), \tag{7}$$

with the initial conditions

$$0 \le t \le T, \quad y(0) = \xi(0) = 0 \pmod{P}. \tag{8}$$

Under the conditions (i) and (iii) the Cauchy problem (7), (8) has a unique solution $x(t)$. It generates the probability measure or distribution in the space $\mathcal{L}_2^-$. We denote this measure by μ_x.

The measures μ_y and μ_x are related through a smooth nonlinear transform in the space $\mathcal{L}_2^-$. This transform can be written as

$$\psi = \varphi + UF(\varphi). \tag{9}$$

To transform (9) we can apply the main theorem of [2] and establish the equivalence of the measures μ_y and μ_x. In this case we can write explicitly the corresponding Radon–Nikodym density. Moreover, we can calculate the Fredholm determinant that figures in the density formula. Let us do this.

To begin with, we find the spectral radius of the compact operator $UF' \overset{def}{=} V$. This is an integral operator. Let $v(t,s)$ be its kernel. Then

$$v(t,s) = 0 \text{ for } s > t.$$

Let

$$M = \max_{t,s} \|v(t,s)\|_H$$

and

$$N_x = \int_0^T \|x(t)\|_H^2 \, dt.$$

We have the estimates

$$\big\|(Vx)(\,\cdot\,)\big\|_{\mathcal{L}_2}^2 = \bigg\| \int_0^{\cdot} v(\,\cdot\,,s)x(s)\,ds \bigg\|_{\mathcal{L}_2}^2 \le$$

$$\le \int_0^{\cdot} \|v(\,\cdot\,,s)\|_H^2 \|x(s)\|_H^2 \, ds \le N_x M^2,$$

$$\big\|(V^2x)(\,\cdot\,)\big\|_{\mathcal{L}_2}^2 = \bigg\| \int_0^{\cdot} v(\,\cdot\,,s)\bigg[\int_0^s v(s,\tau)x(\tau)\,d\tau\bigg]\,ds \bigg\|_{\mathcal{L}_2}^2 \le N_x M^4 T$$

and, in general,

$$\big\|(V^nx)(\,\cdot\,)\big\|_{\mathcal{L}_2}^2 \le N_x M^{2n} \frac{t^{n-1}}{(n-1)!} \le N_x M^{2n} \frac{T^{n-1}}{(n-1)!}\,.$$

Hence we obtain

$$[\![V^n]\!]_{\mathcal{L}_2} = \sup_{\|x\|=1} \|V^n x\|_{\mathcal{L}_2} = \sup_{\|x\|=1} \int_0^T \big\|(V^nx)(t)\big\|_H^2 \, dt \le M^{2n} \frac{T^n}{(n-1)!}\,.$$

Therefore, for the spectral radius r of the operator $V = UF'$ we get that

$$r = \lim_{n\to\infty} \sqrt[n]{[\![V^n]\!]_{\mathcal{L}_2}} = 0$$

and the Neumann series $\sum_{n=1}^{\infty} V^n x$ converges uniformly. Hence we conclude that

$$\widetilde{\det}(I + UF') = \det(I + UF') = 1.$$

Applying the above-mentioned theorem to transform (9) we conclude that under the conditions (i), (ii) and (iii) we have $\mu_y \sim \mu_x$ and the Radon–Nikodym density is written as

$$\frac{d\mu_y}{d\mu_x}(z) =$$

$$= \exp\bigg\{\theta \int_0^T \int_0^1 \Big(\lambda\Big(t, \frac{dz}{dt} - A(t)z(t) + \tau\theta f(t, z(t))\Big), f(t, z(t))\Big)_H \, d\tau \, dt\bigg\}. \quad (10)$$

Main result

Let us apply the obtained result and formula (10) to the Gaussian case. Assume that the conditions (i) and (ii) are fulfilled. Instead of (iii), we require the fulfillment of a weaker condition:

(iv) $\xi(t)$ is a random Gaussian process on $[0,T]$ with values in H, whose nearly all trajectories are continuous; $E\xi(t)=0$, while its correlation operator kernel $R(t,s)$ satisfies the inequality

$$\int_0^T \|R(t,t)\|_2 \, dt < \infty. \tag{11}$$

We define the operator $K(t,s)$ from the operator relation

$$R(t,s) = \int_0^T K(t,\tau)K^*(s,\tau)\, d\tau. \tag{12}$$

Let $\mathfrak{R}$ and $\mathfrak{K}$ be integral operators in $\mathcal{L}_2([0,T],H)$ with kernels $R(t,s)$ and $K(t,s)$, respectively:

$$(\mathfrak{R}\varphi)(\mathfrak{t}) = \int_0^T R(t,s)\varphi(s)\, ds, \quad \varphi \in \mathcal{L}_2([0,T],H),$$

$$(\mathfrak{K}\varphi)(\mathfrak{t}) = \int_0^T K(t,s)\varphi(s)\, ds, \quad \varphi \in \mathcal{L}_2([0,T],H)$$

and introduce the spaces $X_+ = \mathfrak{R}H$, $X = \mathfrak{K}H$ and $X_- = H$ with the scalar products

$$\langle x,y\rangle_+ = (\mathfrak{R}^{-1}x, \mathfrak{R}^{-1}y)_H, \quad \langle x,y\rangle = (\mathfrak{K}^{-1}x, \mathfrak{K}^{-1}y)_H.$$

In this notation, problems (2), (3) and (7), (8) can be considered in the equipped space

$$X_+ \subset X \subset X_-,$$

$\xi(t)$ taking its values in X_- and the space X having the unit correlation operator $\delta(t-s)$. This makes it possible to use in the equipped space

$$\mathcal{L}_2^+ = \mathcal{L}_2([0,T],X_+) \subset \mathcal{L}_2 = \mathcal{L}_2([0,T],X) \subset \mathcal{L}_2^-([0,T],X_-)$$

the appropriate result from [2]. For this, we define the Wiener process $w(t)$ from the equality

$$\xi(t) = \int_0^T K(t,s)\, dw(s). \tag{13}$$

Then we make sure that the following statement is valid.

Theorem. *Let for problem* (2), (3) *the conditions* (i), (ii) *and* (iv) *be fulfilled. Then if the integral Fredholm equation of first kind*

$$f(t,y(t)) = \int_0^T K(t,s)g(s,y)\, ds$$

is solvable with respect to $g(t,y)$, then the measures μ_y and μ_x are equivalent and

$$\frac{d\mu_y}{d\mu_x}(u) = \exp\Big\{ -\theta\int_0^T \big(g(s,u(\cdot)), dw(s)\big)_H - \frac{\theta^2}{2}\int_0^T \big\|g(s,u(\cdot))\big\|_H^2\, ds\Big\}, \tag{14}$$

where the Wiener process $w(t)$ is defined from (13)*.*

Note that in (14) the first summand in the exponent is understood as an expanded Skorokhod stochastic integral.

Stochastic equations

Let us consider the case where the right-hand part of equation (2) is "generalized white" noise. Then an equation is called a stochastic differential equation in the Hilbert space H and it is usually written in the form

$$dy(t) - A(t)y(t)\,dt + \theta f(t,y(t))\,dt = dw(t) \tag{15}$$

with the initial condition

$$y(0) = 0 \pmod{P}. \tag{16}$$

Note that problem (15), (16) is understood as an equivalent form of the following stochastic integral equation

$$y(t) + \theta\int_0^t u(t,s) f(s,y(s))\,ds = \int_0^t u(t,s)\,dw(s).$$

Corollary. Under the conditions (i) and (ii), for the stochastic problem (15), (16) we have the equivalence of the measures μ_y and μ_x. In that case, the Radon–Nikodym density has the form

$$\frac{d\mu_y}{d\mu_x}(u) = \exp\Big\{ -\theta\int_0^T \big(f(s,u(\cdot)), dw(s)\big)_H - \frac{\theta^2}{2}\int_0^T \big\|f(s,u(\cdot))\big\|_H^2\, ds\Big\}. \tag{17}$$

Estimation of the maximal likelihood of the parameter θ

Under appropriate conditions, the likelihood functions (14) and (17) enable us to find a maximal likelihood estimate (MLE) for the unknown parameter in the nonlinear member if we use the observed solution of the equation.

Under the conditions of the theorem, for equation (2) with initial condition (3) a MLE has the form

$$\widehat{\theta}(u_0) = -\frac{\int_0^T (g(s,u_0(s)), dw(s))_H}{\int_0^T \|g(s,u_0(s))\|_H^2\, ds},$$

while, under the conditions of the corollary, for problem (15), (16) a MLE has the form

$$\widehat{\theta}(u_0) = -\frac{\int_0^T (f(s, u_0(s)), dw(s))_H}{\int_0^T \|f(s, u_0(s))\|_H^2 \, ds}.$$

In these formulas, $u_0(t)$ is the observed solution of problem (2), (3) (or of (15), (16)), respectively.

Using Bernstein inequalities for sums of independent random variables [1], we can construct confident domains for the above estimates. For this we can use the method of [1].

R E F E R E N C E S

1. Bondarev B.V. An estimate for the nonlinearity parameter in initial-boundary value problems with Gaussian perturbations. (Russian) *Teor. Veroyatnost. i Primenen.* **36**, 3 (1991), 553–560; translation in *Theory Probab. Appl.* **36**, 3 (1991), 570–578 (1992).
2. Sokhadze G.A. On the absolute continuity of smooth measures. (Ukrainian) translated from *Teor. Ĭmovīr. Mat. Stat.*, **49** (1993), 225-230; *Theory Probab. Math. Statist.*, **49** (1994), 159–163 (1995).
3. Showalter R. E. Hilbert Space Methods for Partial Differential Equations. *Monographs and Studies in Mathematics*, Vol. 1. *Pitman, London-San Francisco, Calif.-Melbourne*, 1977.
4. Bernšteĭn S. N. Collected works. Vol. IV: Theory of probability. Mathematical statistics. 1911-1946. (Russian) *Izdat. "Nauka", Moscow*, 1964.

Received 15.07.2013; revised 01.10.2013; accepted 26.11.2013.

Authors' addresses:

A. Gubeladze
Kiev T. Shevchenko National University
1, Glushkova St., Kiev
Ukraine

G. Sokhadze
I. Vekua Institute of Applied Mathematics of
Iv. Javakhishvili Tbilisi State University
2, University St., Tbilisi 0186
Georgia
E-mail: grigol.sokhadze@tsu.ge