



ELSEVIER

Stochastic Processes and their Applications 51 (1994) 259–267

stochastic
processes
and their
applications

On minimum uniform metric estimate of parameters of diffusion-type processes

Yury Kutoyants^{a,*}, Philippe Pilibossian^b

^aUniversité du Maine, Département Mathématiques, B.P. 535–72017 Le Mans, France and Université Paris VI, CNRS URA 1321, 4, pl. Jussieu, 75252 Paris Cedex 05, France

^bUniversité Paris VI, UFR 920 Mathématiques, CNRS URA 1321, 4, pl. Jussieu, Tour 45, 46, 75252 Paris Cedex 05, France

Received 2 May 1991; revised 12 August 1993

Abstract

The problem of finite-dimensional parameter estimation for a diffusion-type process is considered. The proposed minimum distance estimate is introduced as a point where the supremum norm of the difference between the observations and the corresponding deterministic (limit) solution attains its minimum. Under some regularity conditions the consistency of this estimate is established as the diffusion coefficient tends to zero and the limit distribution is described.

Key words: Parameter estimation; Diffusion-type process; Diffusion coefficient; Limit distribution

AMS Subject Classification: 62M09, 62F12, 62C20.

1. Introduction

Let us consider the problem of estimating the unknown parameter θ of a diffusion-type process:

$$dX_t = S_t(\theta, X) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T, \quad (1)$$

where $S_t(\cdot, X)$ is a known measurable, nonanticipative functional, $\theta \in \Theta$, an open bounded set in \mathbb{R}^d , $d \geq 1$, W_t is a Wiener process and $\varepsilon \in (0, 1]$ (see, for example, Liptser and Shiriyayev, 1977).

*Corresponding author.

We are interested in the limiting properties of estimates of θ as $\varepsilon \rightarrow 0$. This corresponds to the presence of small noise in the dynamical system:

$$\frac{dx_t}{dt} = S_t(\theta, x), \quad x_0, \quad 0 \leq t \leq T. \quad (2)$$

The properties of maximum likelihood and Bayes estimates in this situation are known (see, for example, Kutoyants, 1984). Another approach is to consider the minimum norm estimates of θ , which could be described, for example, as follows. Set

$$\hat{\theta}_\varepsilon = \arg \min_{\theta \in \bar{\Theta}} \|X - x(\theta)\|, \quad (3)$$

where $\bar{\Theta}$ is the closure of Θ and

$$\|f\|^2 = \int_0^T f(t)^2 d\mu_t,$$

for some finite measure μ and we denote by $x_t(\theta) = x_t$ the solution of Eq. (2). The consistency and asymptotic normality of these minimum norm estimates are established in Kutoyants (1991).

It is interesting also to consider the properties of estimate (3) with respect to the uniform metric:

$$\theta_\varepsilon^* = \arg \min_{\theta \in \bar{\Theta}} \sup_{0 \leq t \leq T} |X_t - x_t(\theta)|. \quad (4)$$

This paper is mainly concerned with the description of the asymptotics of the estimate given by (3). If the equation

$$\sup_{0 \leq t \leq T} |X_t - x_t(\theta_\varepsilon^*)| = \inf_{\theta \in \bar{\Theta}} \sup_{0 \leq t \leq T} |X_t - x_t(\theta)|$$

has several solutions, then we call minimum distance estimate (MDE) any one of them, for instance, that having a minimal norm.

We will prove the consistency of this estimate and describe its limit distribution.

There exists a direct analogy between this problem and the classical minimum distance estimation problem, where starting with n i.i.d. observations with distribution function $F(\theta_0, x)$, the empirical distribution function $\hat{F}_n(x)$ is constructed, and then the minimum distance estimate θ_n^* is defined by the equation

$$\theta_n^* = \arg \min_{\theta \in \bar{\Theta}} \|\hat{F}_n - F(\theta)\|,$$

where $\|\cdot\|$ is a norm in some Banach space B such that $\hat{F}_n, F(\theta) \in B$ and $\|\hat{F}_n - F(\theta_0)\| \rightarrow 0$ as $n \rightarrow \infty$ (Pollard, 1980; Millar, 1984).

Under certain regularity conditions this estimate is known to be consistent (Parr and Schucany, 1982), and for a Hilbert space B , it is also asymptotically normal (Millar, 1984); see also the references in Parr, 1981). The limit distribution of $\sqrt{n}(\theta_n^* - \theta_0)$ for Banach spaces is non-Gaussian (Pollard, 1980; Litell and Rao, 1982). Millar (1984) has proved the asymptotic optimality in some sense of this estimate for a Hilbertian B .

2. Consistency

We investigate the properties of θ_ε^* under the following conditions.

(CI) The functional $S_t(\theta, \cdot)$ is measurable, nonanticipative and satisfies the following inequalities: for all $t \in [0, T]$ and $X, Y \in \mathcal{C}[0, T]$,

$$|S_t(\theta, X) - S_t(\theta, Y)| \leq L_1 \int_0^t |X_s - Y_s| ds + L_2 |X_t - Y_t|,$$

$$|S_t(\theta, X)| \leq L_1 \int_0^t (1 + |X_s|) ds + L_2 (1 + |X_t|), \quad (5)$$

where L_1, L_2 are positive constants.

(CII) The function $S_t(\theta, x)$ is measurable with respect to (t, θ) and for any $v > 0$, we have

$$g(v) = \inf_{|\theta - \theta_0| \geq v} \sup_{0 \leq t \leq T} |x_t(\theta) - x_t(\theta_0)| > 0.$$

Let us denote by $P_{\theta}^{(e)}$ the measure induced by process (1) in the measurable space $(C[0, T], B[0, T])$ of continuous functions on $[0, T]$ ($B[0, T]$ is a sigma algebra of borelian subsets).

Theorem 1. *Let conditions (CI) and (CII) be fulfilled. Then*

$$P_{\theta_0}^{(e)} \{ |\theta_\varepsilon^* - \theta_0| \geq v \} \leq 2 \exp \left\{ -\gamma \frac{g(v)^2}{\varepsilon^2} \right\}, \quad (6)$$

where γ is some positive constant.

Proof. Condition (CI) ensures the existence and uniqueness of a strong solution of Eq. (1) (Kutoyants, 1984, Theorem 4.6.). Moreover, under condition (5), with $P_{\theta_0}^{(e)}$ probability one, we also have the inequality

$$\sup_{0 \leq t \leq T} |X_t - x_t(\theta_0)| \leq C\varepsilon \sup_{0 \leq t \leq T} |W_t|, \quad (7)$$

where $C > 0$ (Kutoyants, 1991, Lemma 3.4.4).

In the following we denote the sup-norm by $\|\cdot\|$. Let us introduce the set

$$H_0 = H_0(v) = \left\{ \omega : \inf_{|\theta - \theta_0| < v} \|X - x(\theta)\| < \inf_{|\theta - \theta_0| \geq v} \|X - x(\theta)\| \right\}. \quad (8)$$

Note that, for $\omega \in H_0$, the MDE θ_ε^* satisfies $\theta_\varepsilon^* \in \{\theta : |\theta - \theta_0| < v\}$.

Thus, we have

$$P_{\theta_0}^{(e)} \{ |\theta_\varepsilon^* - \theta_0| \geq v \} = P_{\theta_0}^{(e)} \{ H_0^c \}$$

so that we need to estimate the probability of H_0^c (complement of H_0):

$$\begin{aligned}
 P_{\theta_0}^{(\varepsilon)} \{H_0^c\} &\leq P_{\theta_0}^{(\varepsilon)} \left\{ \inf_{|\theta - \theta_0| < \nu} (\|X - x(\theta_0)\| + \|x(\theta) - x(\theta_0)\|) \right. \\
 &\quad \left. \geq \inf_{|\theta - \theta_0| \geq \nu} (\|x(\theta) - x(\theta_0)\| - \|X - x(\theta_0)\|) \right\} \\
 &\leq P_{\theta_0}^{(\varepsilon)} \left\{ \|X - x(\theta_0)\| \geq g(\nu) - C\varepsilon \sup_{0 \leq t \leq T} |W_t| \right\} \\
 &\leq P \left\{ 2C\varepsilon \sup_{0 \leq t \leq T} |W_t| \geq g(\nu) \right\} = P \left\{ \sup_{0 \leq t \leq T} |W_t| \geq \frac{g(\nu)}{2C\varepsilon} \right\} \\
 &\leq 4P \left\{ W_T > \frac{g(\nu)}{2C\varepsilon} \right\} \leq 2 \exp \left\{ -\frac{g(\nu)^2}{8TC^2\varepsilon^2} \right\}.
 \end{aligned}$$

Here we have used the well-known properties of norms, the equality

$$\inf_{|\theta - \theta_0| < \nu} \|x(\theta) - x(\theta_0)\| = 0,$$

inequality (7) and the following property of Wiener processes:

$$P \left\{ \sup_{0 \leq t \leq T} W_t > a \right\} = 2P \{ W_T > a \} \leq \exp \left\{ -\frac{a^2}{2T} \right\}. \quad \square$$

3. Limit distribution

To describe the limit behavior of the normed difference $u_\varepsilon = \varepsilon^{-1}(\theta_\varepsilon^* - \theta_0)$ we specialize the model and we suppose that the trend functional of the process (2) is of the form

$$S_t(\theta, X) = V(\theta, t, X_t) + \int_0^t K(\theta, t, s, X_s) ds. \quad (9)$$

We also need the following additional conditions:

(CIII) The measurable functions $V(\theta, t, x)$ and $K(\theta, t, s, x)$ have two continuous bounded derivatives with respect to θ and x .

Let us denote by $\dot{x}_t(\theta)$ the vector of derivatives of $x_t(\theta)$ with respect to θ (CIII allows one to prove its existence) and consider the matrix

$$J_t(\theta) = \dot{x}_t(\theta) \dot{x}_t(\theta)^T$$

(where T denotes transposition).

(CIV)

$$\inf_{\theta \in \Theta} \inf_{|e|=1} \sup_{0 \leq t \leq T} (e, J_t(\theta) e) > 0,$$

where e is a unit vector in \mathbb{R}^d and (\cdot, \cdot) is an inner product.

We introduce also a Gaussian process $x_t^{(1)} = x_t^{(1)}(\theta)$, which satisfies the equation

$$dx_t^{(1)} = [V'_x(\theta, t, x_t) x_t^{(1)} + \int_0^t K'_x(\theta, t, s, x_s) x_s^{(1)} ds] dt + dW_t, \quad x_0^{(1)} = 0, \quad 0 \leq t \leq T, \quad (10)$$

and is a derivative with probability one of X_t with respect to ε at $\varepsilon = 0$ (see Kutoyants (submitted), Chapter 7). Here V'_x and K'_x are the derivatives of $V(\vartheta, t, x)$ and $K(\vartheta, t, s, x)$ with respect to x .

The random variable $\xi = \xi(\theta_0)$ is defined by the equality

$$\|x^{(1)} - (\xi, \dot{x}(\theta_0))\| = \inf_{u \in \mathbb{R}^d} \|x^{(1)} - (u, \dot{x}(\theta_0))\|.$$

(CV) This equation has a unique solution ξ with probability 1.

Theorem 2. Under conditions (CII)–(CV) we have

$$P_{\theta_0} - \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (\theta_\varepsilon^* - \theta_0) = \xi.$$

Proof. First we localize the problem. Let $v = v_\varepsilon = \varepsilon \lambda_\varepsilon \rightarrow 0$ and $\lambda_\varepsilon \rightarrow +\infty$ as $\varepsilon \rightarrow 0$ and H_0 be defined as in (8). Then for $\omega \in H_0$ we have $|\theta_\varepsilon^* - \theta_0| < v_\varepsilon$.

Let us denote

$$F(u) = \sup_{0 \leq t \leq T} |x_t(\theta_0 + u) - x_t(\theta_0)|^2.$$

As it follows from condition (CIII),

$$\sup_{0 \leq t \leq T} |x_t(\theta_0 + u) - x_t(\theta_0) - (u, \dot{x}_t(\theta_0))| = \mathcal{O}(|u|^2).$$

So if we introduce

$$\kappa_0 = \kappa(\theta_0) = \inf_{|e|=1} \sup_{0 \leq t \leq T} (e, J_t(\theta_0) e),$$

which is by condition (CIV) positive, then we can find a neighborhood V of zero such that

$$\inf_{u \in V} \frac{F(u)}{|u|^2} \geq \frac{1}{2} \kappa_0,$$

and we have for $u \in V$: $F(u) \geq \frac{1}{2} \kappa_0 |u|^2$.

By condition (CII) the function $F(u)$ is positive outside of V . Hence there exists a positive constant κ such that the inequality

$$F(u) \geq \kappa |u|^2$$

holds for all $u \in \Theta - \theta_0$. Thus we obtain

$$\inf_{|u| > v_\varepsilon} \sup_{0 \leq t \leq T} |x_t(\theta_0 + u) - x_t(\theta_0)|^2 \geq \kappa v_\varepsilon^2.$$

Hence $g(v) \geq \sqrt{\kappa} v_\varepsilon$, and from (6) we obtain

$$P_{\theta_0}^{(\varepsilon)} \{ |\theta_\varepsilon^* - \theta_0| \geq v_\varepsilon \} \leq 2 \exp \left\{ -\gamma \kappa \frac{v_\varepsilon^2}{\varepsilon^2} \right\} \leq 2 \exp \{ -\gamma \kappa \lambda_\varepsilon^2 \} \rightarrow 0.$$

Let us now consider the behavior of the norm $\|X - x(\theta)\|$, for $\theta \in \{\theta: |\theta - \theta_0| < v_\varepsilon\}$. We have $\theta = \theta_0 + \varepsilon u$ and

$$\varepsilon^{-1} \|X - x(\theta)\| = \left\| \frac{X - x(\theta_0)}{\varepsilon} - \frac{x(\theta) - x(\theta_0)}{\varepsilon} \right\| = \|x^{(1)} - (u, \dot{x}(\theta_0)) - r + q\|,$$

where

$$q_t = \frac{X_t - x_t(\theta_0)}{\varepsilon} - x_t^{(1)}$$

and

$$r_t = \frac{X_t(\theta_0 + \varepsilon u) - x_t(\theta_0)}{\varepsilon} - (u, \dot{x}_t(\theta_0)).$$

Condition (CIII) allows us to use the Taylor formula and to write

$$\begin{aligned} & \sup_{0 \leq t \leq T} \left| \frac{x_t(\theta_0 + \varepsilon u) - x_t(\theta_0)}{\varepsilon} - (u, \dot{x}_t(\theta_0)) \right| \\ &= \sup_{0 \leq t \leq T} |(u, (\dot{x}_t(\hat{\theta}_0) - \dot{x}_t(\theta_0)))| \leq |u| \sup_{0 \leq t \leq T} |\dot{x}_t(\hat{\theta}_0) - \dot{x}_t(\theta_0)| \leq C\varepsilon |u|^2 \leq C\varepsilon \lambda_\varepsilon^2. \end{aligned}$$

We so obtain

$$\sup_{|u| < \lambda_\varepsilon} \sup_{0 \leq t \leq T} |r_t| \leq C\varepsilon \lambda_\varepsilon^2.$$

Using (1), (7), (9) and (10) we can write

$$\begin{aligned} |q_t| &= \left| \frac{X_t - x_t}{\varepsilon} - x_t^{(1)} \right| \\ &= \left| \int_0^t \left[\frac{S_v(\theta_0, X) - S_v(\theta_0, x)}{\varepsilon} - V'_x(\theta_0, v, x_v) x_v^{(1)} \right] \right| \end{aligned}$$

$$\begin{aligned}
& - \int_0^v K'_x(\theta_0, v, h, x_h) x_h^{(1)} dh \Big] dv \Big| \\
& \leq \int_0^t \left| \frac{V(\theta_0, s, X_s) - V(\theta_0, s, x_s)}{\varepsilon} - V'_x(\theta_0, s, x_s) x_s^{(1)} \right| ds \\
& \quad + \int_0^t \int_0^s \left| \frac{K(\theta_0, s, v, X_v) - K(\theta_0, s, v, x_v)}{\varepsilon} - K'_x(\theta_0, s, v, x_v) x_v^{(1)} \right| dv ds \\
& \leq \int_0^t \left| V'_x(\theta_0, s, \tilde{X}_s) \frac{X_s - x_s}{\varepsilon} - V'_x(\theta_0, s, x_s) x_s^{(1)} \right| ds \\
& \quad + \int_0^t \int_0^s \left| K'_x(\theta_0, s, v, \tilde{X}_v) \frac{X_v - x_v}{\varepsilon} - K'_x(\theta_0, s, v, x_v) x_v^{(1)} \right| dv ds \\
& \leq \int_0^t |V'_x(\theta_0, s, \tilde{X}_s)| \left| \frac{X_s - x_s}{\varepsilon} - x_s^{(1)} \right| ds + \int_0^t |V'_x(\theta_0, s, \tilde{X}_s)| \\
& \quad - V'_x(\theta_0, s, x_s)|x_s^{(1)}| ds + \int_0^t \int_0^s |K'_x(\theta_0, s, v, \tilde{X}_v)| \left| \frac{X_v - x_v}{\varepsilon} - x_v^{(1)} \right| dv ds \\
& \quad + \int_0^t \int_0^s |K'_x(\theta_0, s, v, \tilde{X}_v) - K'_x(\theta_0, s, v, x_v)| |x_v^{(1)}| dv ds \\
& \leq C_1 \int_0^t |q_s| ds + C_2 \int_0^t \int_0^s |q_v| dv ds + C_3 \varepsilon \sup_{0 \leq t \leq T} |W_t| \sup_{0 \leq t \leq T} |x_t^{(1)}|,
\end{aligned}$$

with some constants $C_1 > 0$, $i = 1, 2, 3$.

From (10), condition (CH1) and Lemma 4.13 of Kutoyants (1984) we obtain

$$\sup_{0 \leq t \leq T} |x_t^{(1)}| \leq C \sup_{0 \leq t \leq T} |W_t|.$$

This allows us to write the inequality

$$|q_t| \leq C_1 \int_0^t |q_s| ds + C_2 \int_0^t \int_0^s |q_v| dv ds + \tilde{C}_3 \varepsilon \sup_{0 \leq t \leq T} |W_t|^2$$

and using Lemma 4.13 of Kutoyants (1984) once more, we finally obtain the desired expression

$$\sup_{0 \leq t \leq T} |\varepsilon^{-1}(X_t - x_t(\theta_0)) - x_t^{(1)}| \leq C \varepsilon \sup_{0 \leq t \leq T} |W_t|^2. \quad (12)$$

Now we return to the original problem and consider the difference

$$\begin{aligned}
& \sup_{|u| < \lambda_t} \sup_{0 \leq t \leq T} \left| \frac{X_t - x_t(\theta_0 + \varepsilon u)}{\varepsilon} - (x_t^{(1)} - (u, \dot{x}_t(\theta_0))) \right| \\
& \leq \sup_{|u| < \lambda_t} \sup_{0 \leq t \leq T} \left\{ \left| \frac{X_t - x_t(\theta_0)}{\varepsilon} - x_t^{(1)} \right| + \left| \frac{x_t(\theta_0 + \varepsilon u) - x_t(\theta_0)}{\varepsilon} - (u, \dot{x}_t(\theta_0)) \right| \right\}
\end{aligned}$$

$$\leq \sup_{0 \leq t \leq T} |q_t| + \sup_{|u| < \lambda_\varepsilon} \sup_{0 \leq t \leq T} |r_t| \leq C\varepsilon \sup_{0 \leq t \leq T} |W_t|^2 + C\varepsilon \lambda_\varepsilon^2,$$

where we have used (11) and (12).

Hence, if we choose λ_ε such that $\varepsilon \lambda_\varepsilon^2 \rightarrow 0$ when $\varepsilon \rightarrow 0$, then with probability one we obtain

$$\sup_{|u| \leq \lambda_\varepsilon} \left| \frac{\|X - x(\theta_0 + \varepsilon u)\|}{\varepsilon} - \|x^{(1)} - (u, \hat{x}(\theta_0))\| \right| \rightarrow 0.$$

Therefore, for $t \in [0, T]$, we have uniform convergence of continuous functions of u towards a continuous function and the minimizer (as a continuous functional of the trajectory) converges to the minimizer of the limit process.

$$\arg \inf_{|u| < \lambda_\varepsilon} \|X - x(\theta_0 + \varepsilon u)\| \sim \arg \inf_{|u| < \lambda_\varepsilon} \|x^{(1)} - (u, \hat{x}(\theta_0))\| \rightarrow \xi. \quad \square$$

The random variable ξ could be approximated in the following way. Let us consider points

$$0 = t_1^{(n)} < t_2^{(n)} < \dots \leq t_n^{(n)} = T \quad \text{such that} \quad \max_{1 \leq j \leq n-1} |t_{j+1}^{(n)} - t_j^{(n)}| \rightarrow 0$$

when $n \rightarrow +\infty$ and the random variables

$$\xi_n = \arg \inf_{|u| < \lambda_\varepsilon} \max_{1 \leq j \leq n} |x_{t_j}^{(1)} - (u, \hat{x}_{t_j}(\theta_0))|. \quad (13)$$

The process $x_t^{(1)}$ and the vector function \hat{x}_t are continuous with respect to t ; hence $\xi_n \rightarrow \xi$ as $n \rightarrow +\infty$. Then $x_{t_j}^{(1)}$ and \hat{x}_{t_j} appearing in (13) can be calculated. Moreover, the inner product therein is linear with respect to the components of u . This expression is similar to the linear regression model with minimax criteria (Millar, 1984). Theorem 3 in Pilibossian (1977) can then be applied, for $n > d$. The direct resolution method, developed in the same paper, could also be applied to calculate the value of ξ_n .

Remark. The Condition (CV) in the one-dimensional ($d = 1$) case could be easily verified if we suppose that $\hat{x}_t(\theta_0) > 0$ for all $t \in (0, T]$. Denote $\varphi_t(u) = |x_t^{(1)} - u\hat{x}_t(\theta_0)|$ and $\varphi(u) = \sup_{0 \leq t \leq T} \varphi_t(u)$. The function $\varphi(u)$ is convex on u and $\varphi(u) \rightarrow \infty$ as $|u| \rightarrow \infty$, so the minimum of this function exists and is reached on the set $U^* = \{u^*: \varphi(u^*) = \inf \varphi(u)\}$. From the convexity of $\varphi(u)$ it also follows that $U^* = [\alpha, \beta]$ and suppose that $\alpha \neq \beta$. Fix $u_0 \in (\alpha, \beta)$ and denote $t_0(u_0) = \arg \sup_{0 \leq t \leq T} \varphi_t(u_0)$. Then $\varphi_{t_0}(u_0) = \varphi(u_0) = \inf \varphi(u)$. Let $u_1 \neq u_0$, $u_1 \in U^*$. Introduce $t_1 = t_1(u_1)$. By the condition $\varphi_{t_0}(u_0) \neq \varphi_{t_0}(u_1)$. Let $\varphi_{t_0}(u_0) < \varphi_{t_0}(u_1)$, but $\varphi_{t_0}(u_1) < \varphi_{t_1}(u_1) = \varphi_{t_0}(u_0) = \varphi(u^*)$. This contradiction proves the result.

References

- Yu.A. Kutoyants, *Parameter Estimation for Stochastic Processes* (Heldermann, Berlin, 1984).
- Yu.A. Kutoyants, Minimum distance parameter estimation for diffusion type observations, *C.R. Acad. Sci. Paris, Ser. I* 312 (1991) 637–642.
- Yu.A., Kutoyants, Identification of dynamical systems with small noise, submitted for publication.
- R.S. Liptser and A.N. Shirayev, *Statistics of Random Processes, I. General Theory* (Springer, New York, 1977).
- R.C. Littel and P.V. Rao, On minimum distance estimation of location based on the Kolmogorov statistic, *Comm. Statist. Theory Methods* 11 (1982) 1793–1807.
- P. Millar, A general approach to the optimality of minimum distance estimators, *Trans. Amer. Math. Soc.* 286 (1) (1984) 377–418.
- W.C. Parr, Minimum distance estimation: survey and bibliography, *Comm. Statist. A* 10 (1981) 1205–1224.
- W.C. Parr and W.R. Schucarry, Minimum distance estimation and components of goodness-of-fit statistics, *J. Roy. Statist. Soc. Ser. B* 44 (1982) 178–189.
- Ph. Pilibossian, Régression du minimax, *C.R. Acad. Sci. Paris*, 288 (1977) 921–923.
- D. Pollard, The minimum distance method of testing, *Metrika* 27 (1980) 43–70.