MAXIMUM LIKELIHOOD ESTIMATION IN SKOROHOD STOCHASTIC DIFFERENTIAL EQUATIONS

JAYA P. N. BISHWAL

(Communicated by Edward C. Waymire)

ABSTRACT. Consistency and limit distribution of the maximum likelihood estimator of a parameter in the drift coefficient of an anticipative Skorohod stochastic differential equation satisfying a boundary condition are obtained based on n independent trajectories of the corresponding Skorohod diffusion inside a time interval $[0, T]$ as $n \rightarrow \infty$. The results are illustrated for the anticipative Ornstein-Uhlenbeck process.

1. INTRODUCTION

Parameter estimation in Itô stochastic differential equations has received a lot of attention during the last three decades in view of their wide applications in finance, engineering, physics and biology; see Kutoyants ([9], [10], [11]) and Bishwal ([1]). Here the solution, the diffusion process, is an adapted process, and one of the basic assumptions is that the initial value of the random variable is independent of the whole path of the driving Brownian motion. Whereas it is a natural condition in many situations, where the filtration represents the evolution of the available information, in many cases it is a limitation. In many applications, one needs a model where the initial value of the process depends on the entire path of the driving Brownian motion. For example, in insider trading in finance, an agent, in addition to knowing prices up to the present, has nonpublic advance information of the price of the stock at the terminal time T . The insider takes this advantage in optimizing his portfolio and/or consumption in the market. See Pikovsky and Karatzas ([18]) and Platen and Rebolledo ([16]). Such problems are overcome by modeling by Skorohod stochastic differential equations. The adaptedness (nonanticipativity) of the Itô integral is weakened in the Skorohod integral (see Skorohod ([20]), which allows anticipativity of the integrands. The two main aspects of the Skorohod integral are its total symmetry with respect to time reversal: it generalises the Itô forward and Itô backward integral, and no restriction whatsoever is put on the possible dependence of the integrand on the future increments of the Brownian integrator. The price that has to be paid for this generality consists of some smooth requirements upon the integrand. Nualart and Pardoux ([13]) developed an extended stochastic

Received by the editors August 15, 2008, and, in revised form, May 15, 2009.

2010 *Mathematics Subject Classification.* Primary 62F12, 62M05; Secondary 60F05, 60H05, 60H10.

Key words and phrases. Skorohod stochastic differential equations, Skorohod integral, anticipative Girsanov transformation, maximum likelihood estimator, consistency, limit distribution, Carleman-Fredholm determinant.

calculus for the Skorohod and generalised Stratonovich integrals. This theory allows the study of SDEs where the solution is a nonadapted process. For a review of different types of anticipative SDEs, see Pardoux ([17]). For a detailed treatment of anticipative SDEs, see Nualart ([12]).

2. MODEL AND PRELIMINARIES

Let $(\Omega, \mathcal{F}, P)$ denote the Wiener space; i.e., $\Omega = C_0[0, T]$ is the Wiener space of continuous functions on $[0, T]$ with initial value 0 which is equipped with the supremum norm $\|\cdot\|$, $\mathcal{F}$ is the Borel σ -algebra of subsets of Ω and P is the standard Wiener measure on $(\Omega, \mathcal{F})$. We will represent by $W(\omega, t) = \omega(t)$, $0 \leq t \leq T$, $\omega \in \Omega$ the Wiener process on the canonical probability space $(\Omega, \mathcal{F}, P)$.

On $(\Omega, \mathcal{F}, P)$ consider the homogeneous nonlinear Skorohod stochastic differential equation

$$(2.1) \quad \begin{aligned} dX(t) &= \theta f(X(t))dt + \sigma X(t)dW(t), \quad 0 \leq t \leq T, \\ X(0) &= \psi(X(T)), \end{aligned}$$

where ψ is a known function, f is a known real-valued function defined on $\mathbb{R}$ satisfying the existence and uniqueness of solution to (2.1) (see Nualart ([12])), and $\theta \in \Theta \subset \mathbb{R}$ is the unknown parameter which is to be estimated on the basis of n independent copies $X_1(t), X_2(t), \dots, X_n(t)$ of $\{X(t), 0 \leq t \leq T\}$. Let θ_0 be the true value of the unknown parameter.

Next we introduce some basic tools from Malliavin calculus that will be used throughout the text. Let $C_b^\infty(\mathbb{R}^n)$ be the set of C^∞ functions $g : \mathbb{R}^n \rightarrow \mathbb{R}$ which are bounded and have bounded derivatives of all orders. The class of all real random variables of the form $F(\omega) = g(W(t_1), W(t_2), \dots, W(t_n))$, $g \in C_b^\infty(\mathbb{R}^n)$, called Wiener functionals, is denoted by $\mathcal{S}$. $\mathbb{D}^{1,p}$ designates the Banach space which is the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = \{E|F|^p\}^{1/p} + (E[\int_0^T |D_s F|^2 ds]^{p/2})^{1/p},$$

where

$$D_s F := \sum_{i=1}^n \frac{\partial g}{\partial x_i}(W(t_1), W(t_2), \dots, W(t_n)) I_{[0, t_i]}(t).$$

More generally, the k -th order derivative of F is defined as the k -parameter process given by $D_{s_1 \dots s_k}^{(k)}(F) = D_{s_1} D_{s_2} \dots D_{s_k} F$.

$\mathbb{D}^{\alpha,p}$ is defined analogously, and its associated norm is denoted by $\|\cdot\|_{\alpha,p}$. That is,

$$\|F\|_{\alpha,p} = \{E|F|^p\}^{1/p} + (E[\int_0^T \dots \int_0^T |D_{s_1 \dots s_\alpha}^\alpha F|^2 ds_1 \dots ds_\alpha]^{p/2})^{1/p}.$$

$\mathcal{S}$ is dense in $(\mathbb{D}^{1,p}, \|\cdot\|_{1,p})$, $p \geq 1$. Set $\mathbb{D}^{1,\infty} = \bigcap_{p \geq 1} \mathbb{D}^{1,p}$. The adjoint of the closed unbounded operator $D : L^2(\Omega) \rightarrow L^2([0, T] \times \Omega)$ is usually denoted by δ and is called the Skorohod integral. The domain of δ is the class of processes $u \in L^2([0, T] \times \Omega)$ such that

$$|E(\int_0^T D_t F u_t dt)| \leq C \|F\|_{L^2(\Omega)},$$

for all $F \in \mathcal{S}$. Here C is a constant which may depend on u . In the case that u is in the domain of δ , then $\delta(u)$ is the square integrable random variable defined by the duality relation

$$E(\delta(u)F) = E\left(\int_0^1 D_t F u_t dt\right)$$

for all $F \in \mathbb{D}^{1,2}$. The Skorohod integral $\delta(u)$ turns out to be an extension of the classical Itô integral and it allows one to integrate processes u not necessarily adapted. For this reason one also writes $\delta(u) = \int_0^T u(s) dW(s)$.

For any real number $p \geq 1$ and any integer $n \geq 1$ we set $\mathbb{L}^{n,p} = \mathbb{L}^p([0, T]; \mathbb{D}^{n,p})$ and $\mathbb{L}^{1,\infty} = \bigcap_{p \geq 1} \mathbb{L}^{1,p}$. The processes u of the space $\mathbb{L}^{1,2}$ satisfy $uI_{[0,t]} \in \text{Dom } \delta$ for each $t \in [0, T]$, and for those processes one can define the indefinite Skorohod integral $Y_t = \delta(uI_{[0,T]})$ denoted also by $\int_0^t u(s) dW(s)$. This process has continuous modification provided $u \in \mathbb{L}^{1,2}$ and $E(\int_0^T (\int_0^T |D_s u_r|^2 ds)^p dr) < \infty$ for some $p > 2$. The trajectory of this process is continuous but highly irregular. Imkeller ([6]) showed that one can represent the Skorohod integral as the composition of a Gaussian semimartingale depending on an infinite dimensional parameter with a Gaussian vector.

We need the following nonadapted extension of the Girsanov theorem proved by Kusuoka ([8], Theorem 6.4). See also Theorem 4.1.2 in Nualart ([12]).

Theorem 2.1. *Let $V : \Omega \rightarrow \Omega$ be a mapping of the form*

$$V(t, \omega) = \omega(t) + \int_0^t U(s, \omega) ds,$$

where U is a measurable mapping from Ω into $H = L^2(0, T)$ and suppose that the following conditions are satisfied:

- (i) V is bijective.
- (ii) For all $\omega \in \Omega$, there exists a Hilbert-Schmidt operator $DU(\omega)$ from H into itself such that

$$\|U(\omega + \int_0^\bullet h_s ds) - U(\omega) - DU(\omega)h\|_H = o(\|h\|_H)$$

for all $\omega \in \Omega$ as $\|h\|_H \rightarrow 0$,

- (b) $h \rightarrow DU(\omega + \int_0^\bullet h_s ds)$ is continuous from H into $L^2([0, T]^2)$, the space of Hilbert-Schmidt operators 0 for all ω ,
- (c) $I + DU(\omega) : H \rightarrow H$ is invertible.

Then if Q is the measure on $(\Omega, \mathcal{F})$ such that $F = QV^{-1}$, then Q is absolutely continuous with respect to P and

$$\frac{dQ}{dP} = |d_c(-DU)| \exp\left(-\int_0^T U(t) dW_t - \frac{1}{2} \int_0^T U_t^2 dt\right),$$

where $d_c(-DU)$ denotes the Carleman-Fredholm determinant of the Hilbert-Schmidt operator $-DU$ and $\int_0^T U(t) dW(t)$ is the Skorohod integral.

We recall that the Carleman-Fredholm determinant of a Hilbert-Schmidt operator B from $L^2[0, T]$ into itself is defined by the product expansion

$$d_c(B) = \prod_j (1 - \lambda_j) \exp(\lambda_j),$$

where the λ_j 's are the nonzero eigenvalues of B counted as many times as their multiplicities; see Simon ([19]). In particular, if the operator B is nuclear, then

$$d_c(B) = \det(I - B) \exp\{\text{trace } B\}.$$

Thus if DU is nuclear, then

$$d_c(-DU) = \det(I + DU) \exp\{\text{trace } (-DU)\}.$$

3. MAIN RESULTS

Let P_θ^T be the measure generated by the process $\{X(t), 0 \leq t \leq T\}$ on the space $C([0, T], \mathbb{R}), \mathcal{B}([0, T], \mathbb{R})$ equipped with the supremum norm.

By Theorem 2.1, we have

$$\begin{aligned} \frac{dP_{\theta_0}^T}{dP_{\theta^T}} &= d_c(I + (\theta - \theta_0)DX) \exp\{\text{trace } (-DX)\} \\ &\quad \times \exp\left(-(\theta - \theta_0) \int_0^T f(X(t))dW(t) - \frac{(\theta - \theta_0)^2}{2} \int_0^T f^2(X(t))dt\right) \\ &= d_c(I + (\theta - \theta_0)DX) \exp\{\text{trace } (-DX)\} \\ &\quad \times \exp\left(-(\theta - \theta_0) \int_0^T f(X(t))dX(t) - \frac{\theta^2 - \theta_0^2}{2} \int_0^T f^2(X(t))dt\right). \end{aligned}$$

Thus the likelihood based on n independent trajectories $X_1(t), X_2(t), \dots, X_n(t)$ of $\{X(t), 0 \leq t \leq T\}$ is given by

$$\begin{aligned} L_{n,T}(\theta) &= d_c(I + (\theta - \theta_0)DX) \exp\{\text{trace } (-DX)\} \\ &\quad \times \exp\left((\theta - \theta_0) \sum_{i=1}^n \int_0^T f(X_i(t))dX_i(t) - \frac{\theta^2 - \theta_0^2}{2} \sum_{i=1}^n \int_0^T f^2(X_i(t))dt\right), \end{aligned}$$

and the MLE is given by

$$\hat{\theta}_n = \frac{\sum_{i=1}^n \int_0^T f(X_i(t))dX_i(t)}{\sum_{i=1}^n \int_0^T f^2(X_i(t))dt}.$$

Theorem 3.1. Under the assumptions, $\hat{\theta}_n$ is a strongly consistent estimator of θ_0 , i.e.,

$$\hat{\theta}_n \rightarrow \theta_0 \text{ a.s. as } n \rightarrow \infty.$$

Proof. Note that

$$\hat{\theta}_n - \theta_0 = \frac{\sum_{i=1}^n \int_0^T f(X_i(t))dW_i(t)}{\sum_{i=1}^n \int_0^T f^2(X_i(t))dt} = \frac{\frac{1}{n} \sum_{i=1}^n \int_0^T f(X_i(t))dW_i(t)}{\frac{1}{n} \sum_{i=1}^n \int_0^T f^2(X_i(t))dt}.$$

By the strong law of large numbers, the numerator above converges to zero and the denominator converges to $I(\theta_0)$. Hence the result follows. $\square$

Theorem 3.2. *Under the assumptions, $\hat{\theta}_n$ is asymptotically normally distributed i.e.,*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow^D \mathcal{N}(0, I^{-1}(\theta_0)) \text{ as } n \rightarrow \infty.$$

Proof. Note that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{\sqrt{n} \sum_{i=1}^n \int_0^T f(X_i(t)) dW_i(t)}{\sum_{i=1}^n \int_0^T f^2(X_i(t)) dt} = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^T f(X_i(t)) dW_i(t)}{\frac{1}{n} \sum_{i=1}^n \int_0^T f^2(X_i(t)) dt}.$$

By the central limit theorem, the numerator above converges in distribution to $\mathcal{N}(0, I(\theta_0))$. On the other hand, by the strong law of large numbers, the denominator converges a.s. to $I(\theta_0)$, where

$$I(\theta_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^T f^2(X_i(t)) dt = E(f^2(X_1(0))).$$

Hence

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow^D \mathcal{N}(0, I^{-1}(\theta_0)).$$

This completes the proof. $\square$

4. ANTICIPATIVE ORNSTEIN-UHLENBECK PROCESS

Consider the anticipative Ornstein-Uhlenbeck process

$$dX(t) = -\theta X(t)dt + dW(t), \quad 0 \leq t \leq T$$

with the initial condition

$$X(0) = X(T),$$

where $\theta > 0$. Note that

$$X(T) = e^{-\theta T} X(0) + e^{-\theta T} \int_0^T e^{\theta s} dW(s).$$

Since $X(0) = X(T)$,

$$X(0) = X(T) = \frac{1}{e^{\theta T} - 1} \int_0^T e^{\theta s} dW(s).$$

Recall that a process is said to be a Markov process if the past and the future are conditionally independent given the present. A process is said to be a Markov random field if the process outside an interval and inside an interval are conditionally independent given the process at the boundary. More precisely, $\{Y_t, t \in [0, T]\}$ on $\mathbb{R}$ is called a Markov random field if for all a, b with $0 \leq a \leq b \leq T$, the conditional distribution of $\{Y(t), t \in [0, T]\}$ given $\{Y(t), t \in [0, T] \setminus (a, b)\}$ depends only on $(Y(a), Y(b))$. Jamison ([7]) shows that any Markov process is a Markov field, but the converse is not necessarily true. It can be shown that the process $\{X(t), t \in \mathbb{R}\}$ is stationary square integrable.

We show that the process $\{X(t), t \in [0, T]\}$ is a Markov random field. By stationarity and periodicity, it suffices to show that, for $0 < b < T$, the conditional distribution of $\{X(t), t \in [b, T]\}$ given $\{X(t), t \in [0, b]\}$ depends on $(X(b), X(T))$ only. Let us introduce the process $\{X^0(t), t \in [0, T]\}$ which solves

$$dX^0(t) = -\theta X^0(t)dt + dW(t), \quad 0 \leq t \leq T$$

with the initial condition

$$X^0(0) = 0,$$

that is,

$$X^0(t) = e^{-\theta t} \int_0^t e^{\theta s} dW(s) = X(t) - e^{-\theta t} X(0), \quad t \in [0, T].$$

Since $\{X_t^0, t \in [0, T]\}$ is a Markov process and hence also a Markov random field, the conditional distribution of $\{X^0(t), t \in [b, T]\}$ given $\{X^0(t), t \in [0, b], X^0(T)\}$ depends on $(X^0(b), X^0(T))$ only. By observing that $\{X^0(t), t \in [b, T]\}$ is in one-to-one correspondence with $\{X(t), t \in [b, T]\}$, that $(X^0(b), X^0(T))$ is in one-to-one correspondence with $(X(b), X(T))$ and that $\{X^0(t), t \in [0, b], X^0(T)\}$ is in one-to-one correspondence with $\{X(t), t \in [0, b]\}$, it follows that the conditional distribution of $\{X(t), t \in [b, T]\}$ given $\{X(t), t \in [0, b]\}$ depends on $(X(b), X(T))$ only. Hence $\{X(t), t \in [0, T]\}$ is a Markov random field. However, $\{X(t), t \in [0, T]\}$ is not a Markov process, since due to the boundary conditions, $X(0)$ and $X(T)$ are not conditionally independent of $X(t)$.

Note that the initial condition depends on the entire path of the driving Brownian motion W . As is well known, the initial condition in an Ornstein-Uhlenbeck process is usually taken to be independent of W . Here X is not adapted to the filtration generated by W . This type of O-U process is useful in statistical shape analysis; see Grenander ([5]). By Theorem 2.1, we have

$$\begin{aligned} \frac{dP_{\theta_0}^T}{dP_{\theta^T}} &= d_c(I + (\theta_0 - \theta)DX) \exp\{\text{trace}(-DX)\} \\ &\quad \times \exp\left(-(\theta_0 - \theta) \int_0^T X(t) dW(t) - \frac{(\theta_0 - \theta)^2}{2} \int_0^T X^2(t) dt\right) \\ &= d_c(I + (\theta_0 - \theta)DX) \exp\{\text{trace}(-DX)\} \\ &\quad \times \exp\left((\theta_0 - \theta) \int_0^T X(t) dX(t) - \frac{\theta_0^2 - \theta^2}{2} \int_0^T X^2(t) dt\right). \end{aligned}$$

Thus the likelihood ratio based on n independent trajectories $X_1(t), X_2(t), \dots, X_n(t)$ of $\{X(t), 0 \leq t \leq T\}$ is given by

$$\begin{aligned} L_{n,T}(\theta) &:= d_c(I + (\theta - \theta_0)DX) \exp\{\text{trace}(-DX)\} \\ &\quad \times \sum_{i=1}^n \exp\left((\theta - \theta_0) \int_0^T X_i(t) dX_i(t) - \frac{\theta^2 - \theta_0^2}{2} \int_0^T X_i^2(t) dt\right), \end{aligned}$$

and the MLE is given by

$$\hat{\theta}_n = \frac{-\sum_{i=1}^n \int_0^T X_i(t) dX_i(t)}{\sum_{i=1}^n \int_0^T X_i^2(t) dt}.$$

Note that the Carleman-Fredholm determinant is deterministic in this case. Furthermore

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{-\sqrt{n} \sum_{i=1}^n \int_0^T X_i(t) dW_i(t)}{\sum_{i=1}^n \int_0^T X_i^2(t) dt} = \frac{-\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^T X_i(t) dW_i(t)}{\frac{1}{n} \sum_{i=1}^n \int_0^T X_i^2(t) dt}.$$

By the central limit theorem, the numerator above converges in distribution to $\mathcal{N}(0, V^{-1}(\theta_0))$ as $n \rightarrow \infty$, where

$$V(\theta) = \frac{1}{2\theta} \frac{e^{\theta/2} + e^{-\theta/2}}{e^{\theta/2} - e^{-\theta/2}}.$$

On the other hand, by the strong law of large numbers, the denominator converges a.s. to $V(\theta_0)$ as $n \rightarrow \infty$.

Hence

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow^D \mathcal{N}(0, V^{-1}(\theta_0)) \text{ as } n \rightarrow \infty.$$

Note that in the nonanticipative case, $V(\theta) = \frac{1}{2\theta}$.

REFERENCES

- [1] Bishwal, J.P.N. (2008) : *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, 1923, Springer-Verlag, Berlin. MR2360279 (2009c:62007)
- [2] Buckdahn, R. (1992) : Skorohod stochastic differential equations of diffusion type, *Prob. Theor. Rel. Fields* **93**, 297–323. MR1180703 (93g:60128)
- [3] Buckdahn, R. (1994) : *Anticipative Girsanov transformations and Skorohod stochastic differential equations*, *Memoirs Amer. Math. Soc.* 111, no. 533. MR1219706 (95c:60053)
- [4] Buckdahn, R. and Nualart, D. (1993) : Skorohod stochastic differential equations with boundary conditions, *Stoch. Stoch. Reports* **45**, 211–235. MR1306932 (95j:60085)
- [5] Grenander, U. (1993) : *General Pattern Theory. A Mathematical Study of Regular Structures*, The Clarendon Press, Oxford University Press, New York. MR1270904 (96e:68118)
- [6] Imkeller, P. (1993) : Existence and continuity of occupation densities of stochastic integral processes, *Ann. Probab.* **21**, 1050–1072. MR1217580 (94c:60089)
- [7] Jamison, B. (1974) : Reciprocal processes, *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **30**, 65–86. MR0359016 (50:11471)
- [8] Kusuoka, S. (1982) : The nonlinear transformation of Gaussian measure on Banach space and its absolute continuity. I, *J. Fac. Sci. Univ. Tokyo, Sect. IA Math.* **29**, 567–597. MR687592 (84j:60015a)
- [9] Kutoyants, Yu. A. (1984) : *Parameter Estimation for Stochastic Processes* (translated and edited by B.L.S. Prakasa Rao), Heldermann-Verlag, Berlin. MR777685 (86b:62145)
- [10] Kutoyants, Yu. A. (1994) : *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht. MR1332492 (97b:93093)
- [11] Kutoyants, Yu. A. (2004) : *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag, London-New York-Berlin. MR2144185 (2006b:62005)
- [12] Nualart, D. (1995) : *Malliavin Calculus and Related Topics*, Springer-Verlag, New York-Berlin. MR1344217 (96k:60130)
- [13] Nualart, D. and Pardoux, É. (1988) : Stochastic calculus with anticipating integrands, *Prob. Theor. Relat. Fields* **78**, 535–581. MR950346 (89h:60089)
- [14] Nualart, D. and Pardoux, É. (1991) : Boundary value problems for stochastic differential equations, *Ann. Probab.* **19**, 1118–1144. MR1112409 (92j:60072)
- [15] Ocone, D. and Pardoux, É. (1989) : Linear stochastic differential equations with boundary conditions, *Prob. Theory Relat. Fields* **82**, 489–526. MR1002898 (91a:60154)
- [16] Platen, E. and Rebolledo, R. (1994) : Pricing via anticipative stochastic calculus, *Adv. Appl. Probab.* **26**, 1006–1021. MR1303874 (96e:90005)
- [17] Pardoux, É. (1990) : Applications of anticipating stochastic calculus to stochastic differential equations, *Lect. Notes in Math.*, 1444, Springer, Berlin, 63–105. MR1078843 (92b:60049)
- [18] Pikovsky, I. and Karatzas, I. (1996) : Anticipative portfolio optimization, *Advances in Applied Probability* **28**, 1095–1122. MR1418248 (98b:90020)

- [19] Simon, B. (1979) : *Trace Ideals and Their Applications*, London Math. Society, Lecture Notes Series, 35, Cambridge University Press, Cambridge-New York. MR541149 (80k:47048)
- [20] Skorohod, A.V. (1975) : On a generalization of the stochastic integral, *Theory Probab. Appl.* **20**, 219-223. MR0391258 (52:12079)

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE, 376 FRETWELL BUILDING, 9201 UNIVERSITY CITY BOULEVARD, CHARLOTTE, NORTH CAROLINA 28223-0001

E-mail address: J.Bishwal@unccl.edu