J. Japan Statist. Soc.
Vol. 38 No. 1 2008 173–185

ASYMPTOTIC EXPANSION FOR STOCHASTIC PROCESSES: AN OVERVIEW AND EXAMPLES

Yuji Sakamoto* and Nakahiro Yoshida**

The asymptotic expansion method for ϵ-Markov processes with a mixing property is briefly reviewed. It is illustrated by a point process marked by a diffusion process. As a typical application, the expansion formula for the M-estimator based on ϵ-Markov data is exhibited.

Key words and phrases: Asymptotic expansions, ϵ-Markov process, Malliavin calculus.

1. Introduction

The aim of this article is to give an overview of the developments in the theory of the asymptotic expansion for stochastic processes of continuous time. Today we know two typical methods of asymptotic expansion: *the martingale approach* and *the mixing approach*. These methods are complementary to each other. The martingale approach was found first and applied to derive an asymptotic expansion for ergodic diffusion processes. However, if the diffusion process satisfies a sufficiently nice mixing condition, then the mixing approach is more effective. On the other hand, the martingale approach is still useful when the higher-order terms do not obey an asymptotic normal law, which makes it impossible to apply the mixing approach. Such examples are seen in a stochastic regression model with a long memory explanatory variable, and in estimation of a volatility parameter over a finite time interval. In the latter example, the data is strongly time dependent, so that it requires a global estimate of the smoothness of random variables. In this sense, the martingale approach is also called the global approach. Contrarily, the mixing approach is called the local approach since the regularity often comes from a local (in time) estimate of the characteristic function.

We will focus our attention on the mixing approach in this article. In Section 2, we recall a stochastic process having the "ϵ-Markovian" structure as an underlying stochastic process. The ϵ-Markov model written in continuous time may seem to be complicated, however it has an advantage because nonlinear (Markovian) time series models are included in the present model by natural embedding. Section 3 gives an illustrative application. We demonstrate an ap-

Accepted April 14, 2008.

*Graduate School of Human Development and Environment, Kobe University, 3-11 Tsurukabuto, Nada, Kobe 657-8501, Japan. Email: sakamoto@main.h.kobe-u.ac.jp

**University of Tokyo and Japan Science and Technology Agency, Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan. Email: nakahiro@ms.u-tokyo.ac.jp

plication of the Malliavin calculus to verify the condition for the regularity of the distribution. We hope our discussion here explains intuitively how the Malliavin calculus works in the theory of the asymptotic expansion. The use of a support theorem is effective in the treatment of the nondegeneracy of the Malliavin covariance, see Yoshida (2004). Sections 4 and 5 give other applications. We refer the interested reader to: Kusuoka and Yoshida (2000) for the mixing property and asymptotic expansion of diffusion functionals, Sakamoto and Yoshida (2004a) for expansion of a functional admitting a stochastic expansion for the ϵ-Markov process, Sakamoto (2000) for testing problems, Sakamoto and Yoshida (2003) for the degenerate case, Uchida and Yoshida (2001) for Akaike's information criterion and others, Yoshida (1997, 2001) for martingales, Masuda and Yoshida (2005) for application to a stochastic volatility model, and Kutoyants (2004) for the inference for diffusion processes.

2. ϵ-Markov process and asymptotic expansion

We consider a d-dimensional càdlàg process Y and an r-dimensional càdlàg process X on a probability space $(\Omega, \mathcal{F}, P)$. For an interval $I \subset \mathbb{R}_+$, let $\mathcal{B}_I^{X,Y}$ be the σ-field generated by X_t, Y_t $(t \in I)$, and $\mathcal{B}_I^{dX}$ the σ-field generated by $X_t - X_s$ $(s, t \in I)$. Suppose that $\mathcal{B}_{[0,t]}^{X,Y}$ is independent of $\mathcal{B}_{[t,\infty)}^{dX}$ for all $t \in \mathbb{R}_+$, i.e., X is a process with independent increments. Denote by $\mathcal{B}_I^Y$ the σ-field generated by Y_t $(t \in I)$, and we assume that there exists a nonnegative constant ϵ such that Y_t is $\mathcal{B}_{[s-\epsilon,s]}^Y \vee \mathcal{B}_{[s,t]}^{dX}$-measurable for all $s, t > 0$ with $\epsilon \le s \le t$. The process Y is called an *ϵ-Markov process* driven by X. We consider an n-dimensional stochastic process Z such that Z_0 is $\mathcal{B}_{[0]}$-measurable and that the increment $Z_t - Z_s$ is $\mathcal{B}_{[s,t]}$-measurable for every $s, t \in \mathbb{R}_+$, $0 \le s \le t$, where $\mathcal{B}_I = \mathcal{B}_I^Y \vee \mathcal{B}_I^{dX}$. Examples of Y are diffusion processes, nonlinear AR models, cluster processes and diffusion processes with jumps.

We are interested in a higher-order approximation of the distribution of Z_T. For this purpose, we consider a model with a mixing property as well as the existence of moments. Let

$$\alpha(s,t) = \sup_{B_1 \in \mathcal{B}_{[0,s]}, B_2 \in \mathcal{B}_{[t,\infty)}} |P[B_1 \cap B_2] - P[B_1]P[B_2]|$$

and let $\alpha(h) = \sup_{h' \ge h, s \in \mathbb{R}_+} \alpha(s, s+h')$. We will assume

[A1] There exists a constant $a > 0$ such that $\alpha(h) \le a^{-1} e^{-ah}$ for all $h > 0$.

Let $p \in \mathbb{N}$ with $p \ge 3$.

[A2] There exists a positive number h_0 such that

$$E[|Z_0|^{p+1}] + \sup_{t,h: t \in \mathbb{R}_+, h \in [0,h_0]} E[|Z_{t+h} - Z_t|^{p+1}] < \infty$$

and $E[Z_t] = 0$ for all $t \in \mathbb{R}_+$.

We need a condition that insures the regularity of the distribution of the increments of Z. Suppose that there exist intervals $I(j) = [u(j), v(j)]$ and sub σ-fields $\mathscr{B}'_{[v(j)-\epsilon,v(j)]}$ of $\mathscr{B}_{[v(j)-\epsilon,v(j)]}$ $(j = 1, \dots, n'(T))$ for which the following conditions hold:

(i) $\epsilon < u(j) \le u(j) + \epsilon < v(j) \le v(j) + \epsilon < u(j+1)$.

(ii) Conditional expectation operator $E[\cdot \mid \mathscr{B}_{[v(j)-\epsilon,v(j)]}] = E[\cdot \mid \mathscr{B}'_{[v(j)-\epsilon,v(j)]}]$ on $b\mathscr{B}_{[v(j),\infty)}$.[1]

(iii) $\liminf_{T\to\infty} n'(T)/T > 0$ and $0 < \inf_{j,T}(v(j) - u(j)) \le \sup_{j,T}(v(j) - u(j)) < \infty$.

Then $(I(j))$ form *dense reduction intervals* with $\hat{C}(j) = \mathscr{B}_{[u(j)-\epsilon,u(j)]} \vee \mathscr{B}'_{[v(j)-\epsilon,v(j)]}$.

[A3] There exist truncation functionals $\psi_j : (\Omega, \mathcal{F}) \to ([0,1], \boldsymbol{B}([0,1]))$ and constants $a, a' \in (0,1)$ and $B > 0$ such that $4a' < (1-a)^2$ and

(i) $\frac{1}{n'(T)} \sum_j E[\sup_{u:|u|\ge B} |E[e^{iu\cdot Z_{I(j)}}\psi_j \mid \hat{C}(j)]|] < a'$ for large T, where $Z_{I(j)} = Z_{v(j)} - Z_{u(j)}$.

(ii) $\frac{1}{n'(T)} \sum_j E[\psi_j] > 1 - a$ for large T.

Remark 1. In order to validate an asymptotic expansion formula, it suffices to assume that $\alpha(h)$ decays at a sufficiently large polynomial order. It is possible to relax [A3], though it is sufficient for our use here; see Yoshida (2004).

As usual, we prepare the notations for the representation of the expansion. Let $\chi_{T,r}$ be the r-th cumulant function of $T^{-1/2}Z_T$ defined by

$$\chi_{T,r}(u) = \frac{d^r}{d\epsilon^r}(\log E[\exp(i\epsilon u \cdot T^{-1/2}Z_T)])\mid_{\epsilon=0} .$$

Obviously, $\chi_{T,r}$'s are polynomials in $u = (u_1, \dots, u_n)$ satisfying

$$\chi_{T,r}(u) = i^r \sum_{\alpha_1,\dots,\alpha_r=1}^{n} \lambda^{\alpha_1\cdots\alpha_r} u_{\alpha_1}\cdots u_{\alpha_r}$$

where $\lambda^{\alpha_1\cdots\alpha_r}$ is the r-th cumulant of $T^{-1/2}Z_T$. Particularly, denote the covariance matrix $(\lambda^{\alpha_1\alpha_2})$ by $\bar{g}$. For any positive definite matrix $\sigma = (\sigma^{\alpha\beta})$, the Hermite polynomials $h_{\alpha_1\cdots\alpha_j}$ are defined by

$$h_{\alpha_1\cdots\alpha_j}(z;\sigma) = \frac{(-1)^j}{\phi(z;\sigma)}\partial_{\alpha_1}\cdots\partial_{\alpha_j}\phi(z;\sigma), \qquad \partial_\alpha = \frac{\partial}{\partial z^\alpha},$$

where $\phi(z;\sigma)$ is the density function of the normal distribution with mean 0 and covariance matrix σ. Define the functions $\tilde{P}_{T,r}(u)$ by the formal Taylor expansion:

$$\exp\left(\sum_{r=2}^{\infty} \frac{\epsilon^{r-2}}{r!}\chi_{T,r}(u)\right) = \exp\left(\frac{1}{2}\chi_{T,2}(u)\right) + \sum_{r=1}^{\infty} \frac{\epsilon^r}{T^{r/2}}\tilde{P}_{T,r}(u).$$

1 For any σ-field $\mathscr{B}$, $b\mathscr{B}$ denotes the set of bounded $\mathscr{B}$-measurable functions.

Denote by $\hat{\Psi}_{T,k}(u)$ the k-th partial sum of the RHS of the above equation with $\epsilon = 1$, and define a signed measure $\Psi_{T,k}$ as the Fourier inversion of $\hat{\Psi}_{T,k}$. It is called the (formal) Edgeworth expansion, and can be represented by the density function

$$\Psi_{T,k}(z) = \sum_{j=0}^{k} T^{-j/2} \Xi_{T,j}(z)\phi(z;\bar{g}),$$

where $\Xi_{T,j}(z)$'s are polynomials in z defined by $\Xi_{T,0}(x) = 1$ and

$$\Xi_{T,j}(z) = \sum_{m=1}^{j} \frac{1}{m!} \sum_{\substack{k_1+\cdots+k_m=j \\ k_1\geq 1,\ldots,k_m\geq 1}} \frac{\bar{\lambda}^{\alpha_1^1\cdots\alpha_{k_1+2}^1}\ldots\bar{\lambda}^{\alpha_1^m\cdots\alpha_{k_m+2}^m}}{(k_1+2)!\cdots(k_m+2)!} h_{\alpha_1^1\cdots\alpha_{k_1+2}^1\cdots\alpha_1^m\cdots\alpha_{k_m+2}^m}(z;\bar{g})$$

for $j \geq 1$ with $\bar{\lambda}^{\alpha_1^i\cdots\alpha_{k_i+2}^i} = T^{k_i/2}\lambda^{\alpha_1^i\cdots\alpha_{k_i+2}^i}$. Note that α_s^i's are indices running from 1 to n and that we here adopt the Einstein summation convention for them. For $M > 0$ and $\gamma > 0$, the set $\mathcal{E}(M,\gamma)$ of measurable functions on $\mathbb{R}^n$ is defined by

$$\mathcal{E}(M,\gamma) = \{f : \mathbb{R}^n \to \mathbb{R},\ \text{measurable},\ |f(x)| \leq M(1+|x|)^\gamma\ (x \in \mathbb{R}^n)\}.$$

For measurable function $f : \mathbb{R}^n \to \mathbb{R}$, positive definite matrix σ and $r > 0$, let

$$\omega(f,r,\sigma) = \int_{\mathbb{R}^n} \sup\{|f(x+y) - f(x)| : |y| \leq r\}\phi(x;\sigma)dx.$$

We then have the following asymptotic expansion formula.

THEOREM 2.1. *Assume that $\chi_{T,2}(u) \to -u'\Sigma u$ as $T \to \infty$ for a positive definite matrix Σ. Let Σ^o be a symmetric matrix satisfying $\Sigma < \Sigma^o$. Suppose that the Assumptions* **[A1]**, **[A2]**, **[A3]** *are fulfilled. Then for any $k \in \mathbb{N}$, M, γ, $K > 0$, there exist constants $\delta > 0$ and $c > 0$ such that for $f \in \mathscr{E}(M,\gamma)$,*

$$|E[f(T^{-1/2}Z_T)] - \Psi_{T,k}[f]| \leq c\omega(f, T^{-K}, \Sigma^o) + \epsilon_T^{(k)},$$

where $\epsilon_T^{(k)} = o(T^{-(k+\delta)/2})$ uniformly in $\mathscr{E}(M,\gamma)$, and

$$\Psi_{T,k}[f] = \int_{\mathbb{R}^n} f(z)\Psi_{T,k}(z)dz.$$

Each discrete time model, e.g., m-Markov chains, can be embedded into a continuous time ϵ-Markov process considered here, and therefore the theorem above can be applied to such models.

For any matrix-valued deterministic process (C_t) converging to a non-singular matrix, the distributional expansions of $T^{-1/2}C_T Z_T$ and a transform of it can be derived from this result. Since most of the statistical estimators are asymptotically equivalent to a transform of $T^{-1/2}C_T Z_T$, their distributional asymptotic expansions are also obtained as applications. See Sakamoto and Yoshida (2004a).

3. Point process marked by a diffusion process

We will discuss an illustrative example. For any $T > 0$, let $X = (X_t)_{t\in[0,T]}$ be a d-dimensional stationary diffusion process satisfying a stochastic differential equation

$$dX_t = V_0(X_t)dt + V(X_t)dw_t, \tag{3.1}$$

where $V_0 = (V_0^i)_{i=1,\dots,d} \in C_b^\infty(\mathbb{R}^d;\mathbb{R}^d)$ [2], $V = (V_\alpha^i)_{i=1,\dots,d,\alpha=1,\dots,r} \in C_b^\infty(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^r)$, and w is an r-dimensional standard Wiener process. Let N be a Poisson process with intensity λ and independent of X. We are interested in the asymptotic expansion of the normalized additive functional $\bar{Z}_T$ defined by $\bar{Z}_T = T^{-1/2}Z_T$ and

$$Z_T = \int_0^T g(X_t)dN_t,$$

where $g \in C_\uparrow^\infty(\mathbb{R}^d)$ [3] and $E[g(X_0)] = 0$. In terms of the previous section, X is a 0-Markov process driven by (w, N), and Z is an additive functional of w, N, and X.

Suppose that (i) $E[|X_0|^p] < \infty$ for any $p > 0$ and that (ii)[4] there exists a positive constant a such that

$$\|E[h \mid \mathcal{B}_{[s]}^X] - E[h]\|_1 \le a^{-1}e^{-a(t-s)}\|h\|_\infty$$

for any $s, t \in \mathbb{R}_+$, $s \le t$ and for any bounded $\mathcal{B}_{[t,\infty)}^X$-measurable function h, where $\mathcal{B}_I^X = \sigma[X_t : t \in I]$. Note that (i) and (ii) ensure [A1] and [A2] in the previous section if we regard the Wiener and Poisson processes (w, N) and the diffusion process X in this section as driving process X and ϵ-Markov process Y respectively in the previous section; see Lemma 1 of Yoshida (2004).

Since the characteristic function $\varphi(u)$ of $\bar{Z}_T$ is given by

$$\varphi(u) = E\left[\exp\left\{\int_0^T \lambda(e^{iuT^{-1/2}g(X_t)} - 1)dt\right\}\right],$$

any cumulant of $\bar{Z}_T$ can be easily obtained. For example, the second and third cumulants, $\kappa_T^{(2)}$ and $\kappa_T^{(3)}$, are given by

$$\begin{aligned}
\kappa_T^{(2)} &= E\left[\left(\int_0^T \lambda T^{-1/2}g(X_t)dt\right)^2\right] + E\left[\int_0^T \lambda T^{-1}g(X_t)^2dt\right],\\
\kappa_T^{(3)} &= E\left[\left(\int_0^T \lambda T^{-1/2}g(X_t)dt\right)^3\right]\\
&\quad + 3E\left[\int_0^T \lambda T^{-1}g(X_t)^2dt\int_0^T \lambda T^{-1/2}g(X_t)dt\right]\\
&\quad + E\left[\int_0^T \lambda T^{-3/2}g(X_t)^3dt\right].
\end{aligned}$$

2 C_b^∞ is the space of smooth functions with bounded derivatives of positive order.

3 $C_\uparrow^\infty$ denotes the space of smooth functions, all derivatives of which are of at most polynomial growth.

4 Veretennikov (1987, 1997) and Kusuoka and Yoshida (2000) provided sufficient conditions for (ii).

The principal parts of the above expectations can be represented by the Green function for the generator

$$\mathcal{A} = \sum_{i=1}^{d} V_0^i(x)\frac{\partial}{\partial x^i} + \frac{1}{2}\sum_{i,j}^{d}\sum_{k=1}^{r} V_k^i(x)V_k^j(x)\frac{\partial^2}{\partial x^i \partial x^j}$$

of the diffusion process X. For a measurable function $f : \mathbb{R}^d \to \mathbb{R}$, let $G\langle f\rangle$ be a function such that $\mathcal{A}G\langle f\rangle = f - \nu(f)$, and $[f] = -V'\nabla G\langle f\rangle$, where $\nu(f)$ is the expectation of f w.r.t. the stationary distribution ν of X: $\nu(f) = \int_{\mathbb{R}^d} f(x)\nu(dx)$. We then have

$$\kappa_T^{(2)} = \lambda^2\nu([g]\cdot[g]) + \lambda\nu(g^2) + O(T^{-1})$$

and

$$\kappa_T^{(3)} = \frac{1}{\sqrt{T}}\{3\lambda^3\nu([[g]\cdot[g]]\cdot[g]) + 3\lambda^2\nu([g]\cdot[g^2]) + \lambda\nu(g^3)\} + O(T^{-3/2}).$$

For simplicity, we will assume uniform ellipticity: there exists a positive constant c_0 such that

$$\xi' V(x)V(x)'\xi \geq c_0|\xi|^2 \qquad (\xi \in \mathbb{R}^d).$$

Let $0 < t_1 < t_0$, and let $\bar{Y}$ be the solution to the variational equation:

$$\begin{aligned} d\bar{Y}_t &= \begin{bmatrix} \partial_x V_0(X_t) & 0 \\ \partial_x \mathcal{A}g(X_t)1_{\{t\leq t_1\}} & 0 \end{bmatrix} \bar{Y}_t dt + \sum_{\alpha=1}^{r} \begin{bmatrix} \partial_x V_\alpha(X_t) & 0 \\ \partial_x((\partial_x g)V_\alpha)(X_t)1_{\{t\leq t_1\}} & 0 \end{bmatrix} \bar{Y}_t dw^\alpha \\ \bar{Y}_0 &= I_{d+1}. \end{aligned}$$

We apply the partial Malliavin calculus that shifts only the Wiener part and keeps the initial value X_0 and the Poisson part unchanged. Define a process g_t by the stochastic differential equation

$$\begin{aligned} dg_t &= 1_{\{t\leq t_1\}}\mathcal{A}g(X_t)dt + \sum_{\alpha=1}^{r} 1_{\{t\leq t_1\}}\partial_x g(X_t)V_\alpha(X_t)dw_t^\alpha, \\ g_0 &= g(X_0). \end{aligned}$$

For an element $h = (h_t^\alpha)_{\alpha=1,\dots,r;t\in[0,t_0]}$ of the Cameron-Martin space, the H-derivative $(D_h X_t, D_h g_t)$ satisfies the stochastic differential equations

$$\begin{aligned} dD_hX_t &= \partial_x V_0(X_t)D_hX_t dt + \partial_x V_\alpha(X_t)D_hX_t dw_t^\alpha + V_\alpha(X_t)\dot{h}_t^\alpha dt, \\ dD_h g_t &= 1_{\{t\leq t_1\}}\partial_x\mathcal{A}g(X_t)D_hX_t dt + 1_{\{t\leq t_1\}}\partial_x((\partial_x g)V_\alpha)(X_t)D_hX_t dw_t^\alpha \\ &\quad + 1_{\{t\leq t_1\}}\partial_x g(X_t)V_\alpha(X_t)\dot{h}_t^\alpha dt, \\ D_hX_0 &= 0, \qquad D_h g_0 = 0 \end{aligned}$$

by using Einstein's convention. This implies that

$$\begin{bmatrix} D_hX_t \\ D_h g_t \end{bmatrix} = \sum_{\alpha=1}^{r} \bar{Y}_t \int_0^t \bar{Y}_s^{-1} \begin{bmatrix} V_\alpha(X_s) \\ 1_{\{s\leq t_1\}}\partial_x g(X_s)V_\alpha(X_s) \end{bmatrix} \dot{h}_s^\alpha ds.$$

Then, on the event $\{\Delta N_{t_1} = 1, N_{t_0} = 1\}$, the Malliavin covariance matrix σ of

$$(X_{t_0}, Z_{t_0}) = (X_{t_0}, g(X_{t_1})) = (X_{t_0}, g_{t_0})$$

corresponding to the shift of $w\,|_{[0,t_0]}$ is given by

$$\sigma = \int_0^{t_0} \bar{Y}_{t_0}\bar{Y}_t^{-1} \begin{bmatrix} v(X_t) & 1_{[0,t_1]}v(X_t)(\partial_x g(X_t))' \\ \text{sym.} & 1_{[0,t_1]}\partial_x g(X_t)v(X_t)(\partial_x g(X_t))' \end{bmatrix} (\bar{Y}_t^{-1})'\bar{Y}_{t_0}'\,dt,$$

where

$$v(x) = \sum_{\alpha=1}^{r} V_\alpha(x)V_\alpha(x)'.$$

Take a point $x_* \in \mathbb{R}$ such that $\partial_x g(x_*) \neq 0$. We can choose sufficiently small $t_0 > 0$ so that

$$\sigma_* := \int_0^{t_0} \begin{bmatrix} v(x_*) & 1_{[0,t_1]}v(x_*)(\partial_x g(x_*))' \\ \text{sym.} & 1_{[0,t_1]}\partial_x g(x_*)v(x_*)(\partial_x g(x_*))' \end{bmatrix} dt$$

is positive definite uniformly in t_1 on some nonempty interval $[t_-, t_+] \subset (0, t_0)$; it is possible if one takes small t_+ for t_0. Fix $\varphi \in C^\infty(\mathbb{R}; [0,1])$ such that $\varphi(v) = 1$ if $|v| \leq \frac{1}{2}$ and $\varphi(v) = 0$ if $|v| \geq 1$. Define a functional ψ for $c > 0$ by

$$\psi = \varphi(c^{-1}|X_0 - x_*|^2)\varphi(c^{-1}|\sigma(\sigma_*)^{-1} - I_{d+1}|^2)1_{\{N_{t_-}=0, N_{t_+}=N_{t_0}=1\}}.$$

We note that σ and σ_* in this representation depends on the single jump time t_1.

The functional ψ is smooth in Malliavin's sense, moreover $P[\psi > 0] > 0$ since the path X is drifting near x_* with positive probability as a consequence of the support theorem. Choose sufficiently small c, then σ is uniformly nondegenerate on the event $\{\psi > 0\}$. By using the integration-by-parts formula in the Malliavin calculus, we can apply Theorem 2 of Yoshida (2004), p. 571 to validate the formal Edgeworth expansion of $\mathcal{L}\{\bar{Z}_T\}$ as Theorem 2.1.

Here we will briefly sketch the role of σ. Applying the integration-by-parts formula in the Malliavin calculus, one obtains

$$iuE[e^{iuZ_{t_0}}\psi \mid X_0, X_{t_0}] = E[e^{iuZ_{t_0}}\sigma^{-p}\Psi \mid X_0, X_{t_0}]$$

for all $u \in \mathbb{R}$, where p is some positive number and Ψ is a functional that is smooth in the Malliavin's sense and vanishes on $\{\psi = 0\}$; see Yoshida (2004) for details. Thus

$$\begin{aligned} E\left[\sup_{|u|\geq B} |E[e^{iuZ_{t_0}}\psi \mid X_0, X_{t_0}]|\right] &= E\left[\sup_{|u|\geq B} \frac{|E[e^{iuZ_{t_0}}\sigma^{-p}\Psi \mid X_0, X_{t_0}]|}{|u|}\right] \\ &\leq \frac{1}{B}E[\sigma^{-p}|\Psi|], \end{aligned}$$

and the right-hand side tends to zero as $B \to \infty$, which, together with the stationarity, implies Condition [A3].

4. Nonparametric estimator for diffusion process

Suppose that X is the stationary diffusion process satisfying the stochastic differential equation (3.1), and assume the exponential mixing condition and $X_0 \in \cap_{p>1} L^p$ as well as the smoothness of the coefficients of the stochastic differential equation. However, we only consider the case where the dimension of X is one ($d = 1$). Let $F \in C_{\uparrow}^{\infty}(\mathbb{R})$, and $\vartheta = E[F(X_0)]$. In this section, we will present an asymptotic expansion of the nonparametric estimator

$$\vartheta_T^* = \frac{1}{T}\int_0^T F(X_t)dt$$

for ϑ, which is asymptotically normal and asymptotically efficient in a nonparametric sense (see Kutoyants and Yoshida (2007)). We consider the functional

$$Z_T = \int_0^T q(X_t)dt,$$

where $q(x) = F(x) - \vartheta$.

We assume that the invariant probability measure ν has the support $\mathbb{R}$, and it is given by the density

$$\frac{d\nu}{dx} = \frac{n(x)}{\int_{-\infty}^{\infty} n(y)dy}, \tag{4.1}$$

where

$$n(x) = \frac{1}{V^2(x)p(x)}, \qquad p(x) = \exp\left\{-2\int_0^x \frac{V_0(u)}{V(u)^2}du\right\},$$
$$\int_{-\infty}^{\infty} n(x)dx < \infty.$$

In the non-parametric estimation considered here, the role of Fisher information does play the quantity

$$\mathrm{I}_* = \left\{4E\left[\left(\frac{\mathcal{M}(X_0)}{V(X_0)\frac{d\nu}{dx}(X_0)}\right)^2\right]\right\}^{-1}, \tag{4.2}$$

where $\mathcal{M}(y) = E[(F(X_0) - \vartheta)1_{\{X_0<y\}}]$. For continuous function $f : \mathbb{R} \to \mathbb{R}$, the Green function G_f can be expressed explicitly as

$$G_f(x) = \int_{-\infty}^x p(y)\left(\int_{-\infty}^y 2f(v)n(v)dv\right)dy$$

whenever it exists. We write

$$[f] = -V\nabla G_{f-E[f(X_0)]}.$$

Define the set of functions $\mathcal{C}$ by

$$\mathcal{C} = \left\{ f \in C_{\uparrow}(\mathbb{R}) \,\middle|\, \int_{-\infty}^{\infty} f(x)n(x)dx = 0; \right.$$
$$\left. p(\cdot)\int_{-\infty}^{\cdot} f(x)n(x)dx \in L^1((-\infty,0]); [f], G_f \in C_{\uparrow}(\mathbb{R}) \right\}.$$

Fix $I_\star \in (0, I_*)$ arbitrarily.

THEOREM 4.2 (Kutoyants and Yoshida (2007)). *Let $k \in \mathbb{N}$, and let M, γ, $K > 0$. Suppose that F is not a constant. Then*

(1) *There exist constants $\delta > 0$ and $c > 0$ such that for $h \in \mathcal{E}(M,\gamma)$,*

$$|E[h(\sqrt{T}(\vartheta_T^* - \vartheta))] - \Psi_{T,k}[h]| \le c\omega(h, T^{-K}, I_\star^{-1}) + \epsilon_T^{(k)},$$

where $\epsilon_T^{(k)} = o(T^{-(k+\delta)/2})$ uniformly.

(2) *The signed-measure $d\Psi_{T,1}$ has a density $d\Psi_{T,1}(z)/dz = p_{T,1}(z)$ with*

$$p_{T,1}(z) = \phi(z;\kappa_T^{(2)})\left(1 + \frac{1}{6}\kappa_T^{(3)} h_3(z;\kappa_T^{(2)})\right),$$

where $\kappa_T^{(r)}$ is the r-th cumulant of $\sqrt{T}(\vartheta_T^* - \vartheta)$. Moreover, if q and $[q]^2 - \nu([q]^2)$ belong to $\mathcal{C}$, then

$$p_{T,1}(z) = p_{T,1}^*(z) + R_T(z),$$

where

$$p_{T,1}^*(z) = \phi(z;0,\mathrm{I}_*^{-1})\left(1 + \frac{1}{2\sqrt{T}}E[[[q]^2][q](X_0)]h_3(z;\mathrm{I}_*^{-1})\right)$$

and

$$\lim_{T\to\infty} \sqrt{T} \sup_{z\in\mathbb{R}} \{|R_T(z)|\exp(bz^2)\} = 0$$

for some positive constant b. In particular,

$$\left| E[h(\sqrt{T}(\vartheta_T^* - \vartheta))] - \int_{\mathbb{R}} h(z)p_{T,1}^*(z)dz \right| \le c\omega(h, T^{-K}, I_\star^{-1}) + \tilde{\epsilon}_T$$

with $\tilde{\epsilon}_T = o(1/\sqrt{T})$ uniformly in $h \in \mathcal{E}(M,\gamma)$.

See Kutoyants and Yoshida (2007) for details of this theorem.

5. M-estimator

Let Θ be an open bounded convex set included in $\mathbb{R}^p$. Let $(X_t)_{t\in[0,\infty)}$ be an ϵ-Markov process defined on some probability space, and for each $T > 0$, denote by $\mathfrak{X}_T$ the path space of $X^T = (X_t)_{t\in[0,T]}$. Let ψ_T be an $\mathbb{R}^p$-valued function defined on $\mathfrak{X}_T \times \Theta$. Fix $\theta_0 \in \Theta$ arbitrarily, and let $\tilde{\Theta} \subset \Theta$ be a subset including

θ_0. For the existence of the M-estimator corresponding to ψ_T, it is possible to show that for given $m > 0$ and $\gamma \in (3/4, 1)$, under some regularity conditions, there exists a subset $\tilde{\mathfrak{X}}_T \subset \mathfrak{X}_T$ such that $P[X^T \in \tilde{\mathfrak{X}}_T] = 1 - o(T^{-m/2})$ as $T \to \infty$ and that

$$X^T \in \tilde{\mathfrak{X}}_T \Rightarrow \exists_1 \hat{\theta}_T \in \tilde{\Theta} \text{ s.t. } \psi_T(X^T, \hat{\theta}_T) = 0, \quad \text{and} \quad |\hat{\theta}_T - \theta_0| < T^{-\gamma/2}.$$

For the details, see Theorem 6.1 of Sakamoto and Yoshida (2004a). We refer to any extension of $\hat{\theta}_T$ defined on the whole sample space $\mathfrak{X}_T$ or the pull-back of that map to the probability space as the M-estimator for θ_0, and also denote it by $\hat{\theta}_T$.

Denote the a-th elements of θ and $\psi(X^T, \theta)$ by θ^a and $\psi_{a;}(X^T, \theta)$ respectively, and denote $\delta_{a_1} \cdots \delta_{a_k} \psi_{a;}(X^T, \theta)$ by $\psi_{a;a_1 \cdots a_k}(X^T, \theta)$ where $\delta_a = \partial/\partial\theta^a$. Let $\Delta_{a;}(\theta) = E[\psi_{a;}(X^T, \theta)]$, $\bar{\nu}_{a;} = T^{-1}E[\psi_{a;}(X^T, \theta_0)]$, and $\bar{\nu}_{a;a_1 \cdots a_k} = T^{-1}E[\psi_{a;a_1 \cdots a_k}(X^T, \theta_0)]$. Put

$$Z_{a;} = T^{-1/2}(\psi_{a;}(X^T, \theta_0) - E[\psi_{a;}(X^T, \theta_0)])$$

and

$$Z_{a;A} = T^{-1/2}(\psi_{a;A}(X^T, \theta_0) - E[\psi_{a;A}(X^T, \theta_0)])$$

for any index sequence $A = a_1 \cdots a_k$ whose elements $a_j (j = 1, \ldots, k)$ run from 1 to p.

Suppose that $\sup_{T,\theta} \Delta_{a;}(\theta) < \infty$ and $(\bar{\nu}_{a;b})_{a,b=1}^p$ is nonsingular. Denote the inverse matrix of $(\bar{\nu}_{a;b})_{a,b=1}^p$ by $(\bar{\nu}^{a;b})_{a,b=1}^p$, and let $\Delta^{a;} = -\bar{\nu}^{a;a'}\Delta_{a';}(\theta_0)$, $\bar{\nu}^{a;}{}_A = -\bar{\nu}^{a;a'}\bar{\nu}_{a';A}$ $Z^{a;} = -\bar{\nu}^{a;a'}Z_{a';}$, and $Z^{a;}{}_A = -\bar{\nu}^{a;a'}Z_{a';A}$ for any index sequence A.

For any extended M-estimator $\hat{\theta}_T$ and any bounded function $\beta \in C^2(\Theta; \mathbb{R}^p)$ with bounded first and second derivatives, define a modified M-estimator $\hat{\theta}_T^*$ by

$$\hat{\theta}_T^* = \hat{\theta}_T - \frac{1}{T}\beta(\hat{\theta}_T)$$

and let R_3^a be a random variable defined by

$$\begin{aligned}(5.1)\quad \sqrt{T}(\hat{\theta}_T^* - \theta_0)^a \;=\; & Z^{a;} + \frac{1}{\sqrt{T}}\left(Z^{a;}{}_b Z^{b;} + \frac{1}{2}\bar{\nu}^{a;}{}_{bc}Z^{b;}Z^{c;} + \Delta^{a;} - \beta^a\right) \\ & + \frac{1}{T}\biggl(\frac{1}{6}(\bar{\nu}^{a;}{}_{bcd} + 3\bar{\nu}^{a;}{}_{be}\bar{\nu}^{e;}{}_{cd})Z^{b;}Z^{c;}Z^{d;} \\ & \qquad + \bar{\nu}^{a;}{}_{bc}Z^{b;}Z^{c;}{}_d Z^{d;} + \frac{1}{2}\bar{\nu}^{b;}{}_{cd}Z^{a;}{}_b Z^{c;}Z^{d;} \\ & \qquad + \frac{1}{2}Z^{a;}{}_{bc}Z^{b;}Z^{c;} + Z^{a;}{}_b Z^{b;}{}_c Z^{c;} \\ & \qquad - Z^{b;}\delta_b\beta^a + \Delta^{b;}(Z^{a;}{}_b + \bar{\nu}^{a;}{}_{bc}Z^{c;})\biggr) + \frac{1}{T\sqrt{T}}R_3^a,\end{aligned}$$

where $a = 1, \ldots, p$, and β^a is the a-th element of β. For the remainder term R_3^a above, it can be proved that for given $m > 0$, under some conditions, there exist $C > 0$ and $\epsilon > 0$ such that

$$(5.2)\qquad P[T^{-1/2}|R_3^a| \leq CT^{-\epsilon/2}, a = 1, \ldots, p] = 1 - o(T^{-m/2}).$$

See Theorem 6.2 of Sakamoto and Yoshida (2004a).

Let $Z_T^{(0)} = T^{1/2}(Z_{1;}, \dots, Z_{p;})$ and $Z_T^{(1)} = T^{1/2}(\underbrace{Z_{1;1}, \dots, Z_{p;p}}_{p^2}, \underbrace{Z_{1;11}, \dots, Z_{p;pp}}_{p^3})$.

We require the following conditions:

(i) $(\mathrm{Cov}[Z^{a;}, Z^{b;}])_{a,b=1}^p$ is non-singular;

(ii) there exists a random variable $\dot{Z}_T$ consisting of the elements of $Z_T^{(1)}$, and
 a) $\mathrm{Cov}(T^{-1/2}Z_T^*)$ converges to a positive matrix, where $Z_T^* = (Z_T^{(0)}, \dot{Z}_T)$;
 b) $\ddot{Z}_T = L\dot{Z}_T$ for some matrix L, where $\ddot{Z}_T$ is the random variable consisting of the elements of $Z_T^{(1)}$ except those of $\dot{Z}_T$;
 c) $(Z_t^*)_{t\in[0,\infty)}$ is an additive functional of the ϵ-Markov process $(X_t)_{t\in[0,\infty)}$ driven by some process.

Let $(g^{ab}) = (\mathrm{Cov}[Z^{a;}, Z^{b;}])$, $(g_{ab}) = (g^{ab})^{-1}$, $V^{a;}{}_{a_1\cdots a_k,b} = \mathrm{Cov}[Z^{a;}{}_{a_1\cdots a_k}, Z^{b';}]g_{b'b}$, $\tilde{\mu}^{a;}{}_{bc} = (V^{a;}{}_{b,c} + V^{a;}{}_{c,b} + \bar{\nu}^{a;}{}_{bc})/2$, $\tilde{\eta}^{a;}{}_{b,c} = V^{a;}{}_{b,c} + \bar{\nu}^{a;}{}_{bc}$, and

$$U^{a;}{}_{bcd} = \frac{1}{6}\left(\bar{\nu}^{a;}{}_{bcd} + \sum_{(bc,d)}^{[3]} V^{a;}{}_{bc,d}\right) + \frac{1}{3}\sum_{(bc,d)}^{[3]} \tilde{\mu}^{d';}{}_{bc}\tilde{\eta}^{a;}{}_{d',d}.$$

Put $\tilde{M}^{a;}{}_{b,}{}^{c;}{}_{d} = E[Z^{a;}{}_b Z^{c;}{}_d] - V^{a;}{}_{b,b'}V^{c;}{}_{d,d'}E[Z^{b'}Z^{d'}]$, $\tilde{N}^{a;}{}_{,}{}^{b;}{}_{,}{}^{c;}{}_{d} = T^{1/2}E[Z^a Z^{b;} \cdot (Z^{c;}{}_d - V^{c;}{}_{d,d'}Z^{d'})]$, $\bar{\lambda}^{abc} = T^{1/2}\,\mathrm{Cum}[Z^{a;}, Z^{b;}, Z^{c;}]$, $H^{abcd} = T\,\mathrm{Cum}[Z^{a;}, Z^{b;}, Z^{c;}, Z^{d;}]$. We then obtain a third order asymptotic expansion of the distribution of the modified M-estimator $\hat{\theta}_T^*$ under certain regularity conditions [C0]–[C4] in Sakamoto and Yoshida (2004a).

THEOREM 5.3. *Let $\hat{g} > \lim_{T\to\infty} g$, M and γ' be positive constants. Suppose that the conditions* [A1], [A2], [A3] *for Z_T^* in place of Z_T hold true. Then there exist constants $c > 0$, $\tilde{C} > 0$, $\tilde{\epsilon} > 0$ such that for any function $f \in \mathcal{E}(M, \gamma')$,*

$$\left|E[f(\sqrt{T}(\hat{\theta}_T^* - \theta_0))] - \int dy^{(0)} f(y^{(0)}) q_{T,2}(y^{(0)})\right| \le c\omega(f, \tilde{C}T^{-(\tilde{\epsilon}+2)/2}, \hat{g}) + o(T^{-1}), \tag{5.3}$$

where

$$\begin{aligned} q_{T,2}(y^{(0)}) = \phi(y^{(0)}; g^{ab})\Bigg(1 &+ \frac{1}{6\sqrt{T}}c^{abc}h_{abc}(y^{(0)}; g^{ab}) \\ &+ \frac{1}{\sqrt{T}}(\tilde{\mu}^{a;}{}_{cd}g^{cd} - \tilde{\beta}^a)h_a(y^{(0)}; g^{ab}) \\ &+ \frac{1}{2T}A^{ab}h_{ab}(y^{(0)}; g^{ab}) \\ &+ \frac{1}{24T}c^{abcd}h_{abcd}(y^{(0)}; g^{ab}) \end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad + \frac{1}{72T} c^{abc} c^{def} h_{abcdef}(y^{(0)}; g^{ab})\Bigg),\\
c^{abc} &= \bar{\lambda}^{abc} + 6\tilde{\mu}^{c;}{}_{a'b'} g^{a'a} g^{b'b},\\
A^{ab} &= 2(\bar{\lambda}^{acd} + \tilde{\mu}^{a;}{}_{c'd'} g^{c'c} g^{d'd})\tilde{\mu}^{b;}{}_{cd} + 2\delta_c^{c'} \tilde{N}^{a;}{}_{,}{}^{c;}{}_{,}{}^{b;}{}_{c'} + g^{cc'} \tilde{M}^{b;}{}_{c,}{}^{a;}{}_{c'}\\
&\quad + 2((\Delta^{c;} \tilde{\eta}^{a;}{}_{c,b'} - \delta_{b'}\beta^{a}) + \delta_{b_1}^{a_1} \tilde{M}^{a;}{}_{a_1}{}^{b_1;}{}_{b'} + 3U^{a;}{}_{cdb'} g^{cd}) g^{b'b}\\
&\quad + (\tilde{\mu}^{a;}{}_{cd} g^{cd} - \tilde{\beta}^{a})(\tilde{\mu}^{b;}{}_{ef} g^{ef} - \tilde{\beta}^{b}),\\
c^{abcd} &= H^{abcd} + 4c^{abc}(\tilde{\mu}^{d;}{}_{ef} g^{ef} - \tilde{\beta}^{d}) + 24(\bar{\lambda}^{abe} + 2\tilde{\mu}^{a;}{}_{b'e'} g^{b'b} g^{e'e})\tilde{\mu}^{c;}{}_{d'e} g^{d'd}\\
&\quad + 12(g^{bb'} g^{dd'} \tilde{M}^{c;}{}_{d',}{}^{a;}{}_{b'} + \tilde{N}^{a;}{}_{,}{}^{b;}{}_{,}{}^{c;}{}_{d'} g^{dd'}) + 24U^{a;}{}_{b'c'd'} g^{b'b} g^{c'c} g^{d'd}.
\end{aligned}
$$

As for this result, we refer the reader to Sakamoto and Yoshida (1998, 2004a, 2004b).

Acknowledgements

This work was in part supported by the Grants-in-Aid for Scientific Research of JSPS (Japan Society for the Promotion of Science), and by the Cooperative Research Program of the Institute of Statistical Mathematics.

References

Kusuoka, S. and Yoshida, N. (2000). Malliavin calculus, geometric mixing, and expansion of diffusion functionals, *Probab. Theory Related Fields*, **116**(4), 457–484.

Kutoyants, Y. A. (2004). *Statistical Inference for Ergodic Diffusion Processes*, Springer Series in Statistics, Springer-Verlag London Ltd., London.

Kutoyants, Y. and Yoshida, N. (2007). Moment estimation for ergodic diffusion processes, *Bernoulli*, **13**(4), 933–951.

Masuda, H. and Yoshida, N. (2005). Asymptotic expansion for Barndorff-Nielsen and Shephard's stochastic volatility model, *Stochastic Process. Appl.*, **115**(7), 1167–1186.

Sakamoto, Y. (2000). Asymptotic expansions of test statistics for stochastic processes, *Theory of Statistical Analysis and Its Applications (3)*, volume 130 of *Cooperative Research Report*, pp. 19–32, The Institute of Statistical Mathematics.

Sakamoto, Y. and Yoshida, N. (1998). Third order asymptotic expansion for diffusion process, *Theory of Statistical Analysis and Its Applications*, volume 107 of *Cooperative Research Report*, pp. 53–60, The Institute of Statistical Mathematics.

Sakamoto, Y. and Yoshida, N. (2003). Asymptotic expansion under degeneracy, *J. Japan Statist. Soc.*, **33**(2), 145–156.

Sakamoto, Y. and Yoshida, N. (2004a). Asymptotic expansion formulas for functionals of ϵ-Markov processes with a mixing property, *Ann. Inst. Statist. Math.*, **56**(3), 545–597.

Sakamoto, Y. and Yoshida, N. (2004b). Third order asymptotic expansion of M-estimators for diffusion processes, *Ann. Inst. Statist. Math.* (to appear).

Uchida, M. and Yoshida, N. (2001). Information criteria in model selection for mixing processes, *Stat. Inference Stoch. Process.*, **4**(1), 73–98.

Veretennikov, A. Y. (1987). Bounds for the mixing rate in the theory of stochastic equations, *Theory of Probability and its Applications*, **32**(2), 273–281.

Veretennikov, A. Y. (1997). On polynomial mixing bounds for stochastic differential equations, *Stochastic Process. Appl.*, **70**(1), 115–127.

Yoshida, N. (1997). Malliavin calculus and asymptotic expansion for martingales, *Probab. Theory Related Fields*, **109**(3), 301–342.

Yoshida, N. (2001). Malliavin calculus and martingale expansion, *Bull. Sci. Math.*, **125**(6-7), 431–456. Rencontre Franco-Japonaise de Probabilités (Paris, 2000).

Yoshida, N. (2004). Partial mixing and Edgeworth expansion, *Probab. Theory Related Fields*, **129**(4), 559–624.