Approximate martingale estimating functions under small perturbations of dynamical systems

Uchida, Masayuki
Faculty of Mathematics, Kyushu University

https://hdl.handle.net/2324/3354

出版情報：MHF Preprint Series. MHF2005-9, 2005-03-03. 九州大学大学院数理学研究院
バージョン：
権利関係：

MHF Preprint Series

Kyushu University
21st Century COE Program
Development of Dynamic Mathematics with
High Functionality

Approximate martingale estimating functions under small perturbations of dynamical systems

Masayuki Uchida

MHF 2005-9

(Received March 3, 2005)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

Approximate martingale estimating functions under small perturbations of dynamical systems

Masayuki Uchida
Faculty of Mathematics, Kyushu University
Ropponmatsu, Fukuoka 810-8560, Japan
E-Mail: uchida@math.kyushu-u.ac.jp

Abstract. An approximate martingale estimating function based on eigenfunctions is proposed for an estimation problem about an unknown drift parameter for a one-dimensional diffusion process with small perturbed parameter ε from discrete time observations at n regularly spaced time points k/n, $k = 0, 1, \ldots, n$. We study asymptotic properties of an M-estimator derived from the approximate martingale estimating function as $\varepsilon \to 0$ and $n \to \infty$ simultaneously.

AMS 2000 subject classifications: Primary 62F12, 62M05; Secondary 60J60.

Key words and phrases: diffusion process with small dispersion parameter, discrete time observation, eigenfunction, asymptotic efficiency, parametric inference.

Abbreviated Title: Estimation for dynamical systems.

1 Introduction

Consider a family of one-dimensional diffusion processes defined by the following stochastic differential equations

$$\begin{aligned} dX_t &= b(X_t, \theta)dt + \varepsilon\sigma(X_t)dw_t, \ t \in [0, 1], \ \varepsilon \in (0, 1], \qquad (1) \\ X_0 &= x_0, \end{aligned}$$

where w is a one-dimensional standard Wiener process, the diffusion coefficient $\sigma : \mathbf{R} \to \mathbf{R}$ and ε are known and the drift $b : \mathbf{R} \times \bar{\Theta} \to \mathbf{R}$ is known apart from the parameter θ. Here $\bar{\Theta}$ denotes the closure of Θ and Θ is an open bounded convex subset of $\mathbf{R}^p$. We treat discrete time observations of X at equidistant time points $t_k = k/n$, that is, $X_{t_0}, X_{t_1}, \ldots, X_{t_n}$. The asymptotics is when ε tends to 0 and n tends to ∞ simultaneously.

A family of diffusion processes defined by (1) is an important class of dynamical systems with small perturbations. For dynamical systems with small perturbations, see Azencott (1982) and Freidlin and Wentzell (1984). For applications of dynamical systems with small perturbations to mathematical finance, see Yoshida (1992b), Kunitomo and Takahashi (2001), Takahashi and Yoshida (2004), Uchida and Yoshida (2004b) and references therein. Asymptotically statistical inference for dynamical systems with small perturbations from continuous time observations is well developed, see Kutoyants (1984, 1994), Yoshida (1992a, 2003) and Uchida and Yoshida (2004a). Moreover, there are a number of works on asymptotically parametric estimation based on discrete observations, see Genon-Catalot (1990), Laredo (1990), Sørensen (2000), Sørensen and Uchida (2003), Uchida (2004a, 2004b) and Matsuzaki and Uchida (2005).

As for estimation of a drift parameter for a diffusion process with a small perturbation, the following three papers treated asymptotically efficient estimators in a general setting. Genon-Catalot (1990) proposed two kinds of estimators for an unknown drift parameter. One is derived from a contrast function based on a discretization of the log likelihood function of continuously observed sample path. The estimator has asymptotic efficiency under $(\varepsilon n)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. The other is obtained from the contrast function based on the first order stochastic expansion of X with respect to ε. The second estimator is asymptotically efficient under $\varepsilon\sqrt{n} = O(1)$ as $\varepsilon \to 0$ and $n \to \infty$. For multi-dimensional diffusion processes with small perturbations, Laredo (1990) studied asymptotically efficient estimators of drift parameters under $(\varepsilon n^2)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. Uchida (2004a) applied a martingale estimating function presented in Bibby and Sørensen (1995) to a multi-dimensional diffusion process with a small noise. For the diffusion process (1), the martingale estimating function $\mathcal{M}_{\varepsilon,n}(\theta) = (\mathcal{M}^{(i)}_{\varepsilon,n}(\theta))_{i=1,2,\dots,p}$ is as follows:

$$\mathcal{M}^{(i)}_{\varepsilon,n}(\theta) = \sum_{k=1}^{n} \left(\frac{\partial b}{\partial \theta_i}\right)(X_{t_{k-1}}, \theta)\sigma^{-2}(X_{t_{k-1}})(X_{t_k} - E_\theta[X_{t_{k-1}}|X_{t_k}]) \tag{2}$$

for $i = 1, \dots, p$. An estimator obtained from the martingale estimating function $\mathcal{M}_{\varepsilon,n}(\theta)$ has asymptotic efficiency as $\varepsilon \to 0$ and $n \to \infty$. From the viewpoint of the first order asymptotic theory, all the above four estimators are asymptotically efficient. Therefore, it seems that there is no difficulty for efficient estimation of a drift parameter. However, there is an example for which the four estimators do not work.

In order to explain the motivation of this paper, we consider the following diffusion process with a small noise defined by

$$\begin{aligned} dX_t &= \theta_1 X_t(\theta_2 - X_t)dt + \varepsilon X_t dw_t, \ t \in [0, T], \ \varepsilon \in (0, 1], \\ X_0 &= x_0, \end{aligned} \tag{3}$$

where θ_1 and θ_2 are unknown parameters. The diffusion process (3) is often used as a model for the growth of a population of size X_t in a stochastic, crowded environment. For details of this model, see (5.3.9) on page 78 in Øksendal (2003). If a monthly data is observable, one has that $t_k = k/12$ for $k = 0, 1, \dots, 12T$, that is, $n = 12$. For simplicity, we set that $\varepsilon = 0.01$ and $T = 1$. Under this situation, in order to estimate unknown parameters θ_1 and θ_2, we first try to compute the martingale estimating function $\mathcal{M}_{\varepsilon,n}(\theta)$. However, for the diffusion process (3), the conditional expectation $E_\theta[X_{t_{k-1}}|X_{t_k}]$ does not have an explicit form. Next, we consider the second contrast function presented in Genon-Catalot (1990) since the estimator has asymptotic efficiency under $\epsilon\sqrt{n} = O(1)$. However, we see that the contrast function can not be explicitly obtained from the diffusion process (3). Moreover, the estimator proposed by Laredo (1990) can not be used in this situation since $(\varepsilon n^2)^{-1}$ is not so small. Neither can the estimator derived from the first contrast function of Genon-Catalot (1990).

To overcome the difficulties, Uchida (2004a) investigated approximate martingale estimating functions for multi-dimensional diffusion processes with small perturbations. For the diffusion process (1), the approximate martingale estimating function $\mathcal{G}_{\varepsilon,n,\ell}(\theta) = (\mathcal{G}^{(i)}_{\varepsilon,n,\ell}(\theta))_{i=1,2,\dots,p}$ is as follows: for $\ell \geq 1$ and $i = 1, \dots, p$,

$$\mathcal{G}^{(i)}_{\varepsilon,n,\ell}(\theta) = \sum_{k=1}^{n} \left(\frac{\partial b}{\partial \theta_i}\right)(X_{t_{k-1}}, \theta)\sigma^{-2}(X_{t_{k-1}})P_{k,\ell}(\theta), \tag{4}$$

where $P_{k,\ell}(\theta) = X_{t_k} - \sum_{j=0}^{\ell} \frac{1}{j!n^j} \tilde{L}^j_\theta g(X_{t_{k-1}})$, $g(x) = x$ and $\tilde{L}_\theta g(x) = b(x, \theta)\frac{\partial}{\partial x} g(x)$. He also showed that an estimator obtained from the approximate martingale estimating function $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$

is asymptotically efficient under $(\varepsilon n^\ell)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. Since the situation we consider is when $\varepsilon = 0.01$ and $n = 12$, it suffices to set that $\ell = 3$. It is certain that the estimating function $\mathcal{G}_{\varepsilon,n,3}(\theta)$ is obtained explicitly. Therefore, it seems that we attain our goal. However, there is a serious problem that an explicit estimator can not be derived since the estimating function has a complicated expression. For this reason, Matsuzaki and Uchida (2005) studied a one-step estimator based on the Newton-Raphson method as follows. For an initial estimator $\hat{\theta}^{(0)}_{\varepsilon,n,\ell}$, a one-step estimator $\hat{\theta}^{(1)}_{\varepsilon,n,\ell}$ is defined by

$$\hat{\theta}^{(1)}_{\varepsilon,n,\ell} = \hat{\theta}^{(0)}_{\varepsilon,n,\ell} - \left[\frac{\partial \mathcal{G}_{\varepsilon,n,\ell}}{\partial \theta}(\hat{\theta}^{(0)}_{\varepsilon,n,\ell})\right]^{-1} \mathcal{G}_{\varepsilon,n,\ell}(\hat{\theta}^{(0)}_{\varepsilon,n,\ell}). \tag{5}$$

Under the assumption that an initial estimator $\hat{\theta}^{(0)}_{\varepsilon,n,\ell}$ satisfies that $\varepsilon^{-1}(\hat{\theta}^{(0)}_{\varepsilon,n,\ell} - \theta_0) = O_p(1)$ under $(\varepsilon n^\ell)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$, the one-step estimator $\hat{\theta}^{(1)}_{\varepsilon,n,\ell}$ has asymptotic efficiency. For more details of one-step estimators, see Matsuzaki and Uchida (2005). Therefore, for the population model (3), in order to derive the one-step estimator with asymptotic efficiency under $(\varepsilon n^3)^{-1} \to 0$, it suffices to use an initial estimator with asymptotic normality under $(\varepsilon n^3)^{-1} \to 0$.

For the reason stated as above, in this paper, we are interested in estimating functions generating estimators with asymptotic normality under a general condition that $\varepsilon \to 0$ and $n \to \infty$. However, we note that our final goal is to obtain an asymptotically efficient estimator by using the one-step estimator based on the approximate martingale estimating function $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$. Consequently, it is needless to discuss optimality of the proposed estimating function. Since Bibby and Sørensen (1995) presented martingale estimating functions for discretely observed ergodic diffusion processes, various types of estimating functions have been studied for discretely observed diffusion processes; see Bibby and Sørensen (1996), Kessler (1997), Sørensen (1997), Kessler and Sorensen (1999), Bibby et al. (2002), Aït-Sahalia and Mykland (2003, 2004) and references therein. Kessler and Sørensen (1999) investigated a new type of martingale estimating function based on an eigenfunction for the infinitesimal generator of a diffusion process. They also showed asymptotic normality of an estimator obtained from the martingale estimating function for a discretely observed ergodic diffusion process. Uchida (2004b) applied a martingale estimating function proposed by Kessler and Sørensen (1999) to a diffusion process with a small perturbation given by (1). However, there is a disadvantage that it is not generally easy to find out an eigenfunction $\phi(x,\theta,\varepsilon)$ with an eigenvalue $\Lambda(\theta,\varepsilon)$ such that $L_\theta\phi(x,\theta,\varepsilon) = -\Lambda(\theta,\varepsilon)\phi(x,\theta,\varepsilon)$, where L_θ is the infinitesimal generator of the diffusion (1):

$$L_\theta = b(x,\theta)\frac{\partial}{\partial x} + \frac{1}{2}\varepsilon^2\sigma^2(x)\frac{\partial^2}{\partial x^2}.$$

In order to conquer the difficulty, in this paper, we consider an estimating function based on an eigenfunction $\varphi(x,\theta)$ with an eigenvalue $\lambda(\theta)$ such that $\tilde{L}_\theta\varphi(x,\theta) = -\lambda(\theta)\varphi(x,\theta)$, where $\tilde{L}_\theta$ is the differential operator:

$$\tilde{L}_\theta = b(x,\theta)\frac{\partial}{\partial x}.$$

We notice that L_θ is the second order differential operator while $\tilde{L}_\theta$ is the first order differential operator. Compared with the infinitesimal generator L_θ, it is easy to obtain an eigenfunction with an eigenvalue for the differential operator $\tilde{L}_\theta$. An estimating function constructed by using an eigenfunction with an eigenvalue for $\tilde{L}_\theta$ does not have a martingale property, but it can be asymptotically equivalent to a martingale estimating function. For this reason, an estimating function based on an eigenfunction with an eigenvalue for $\tilde{L}_\theta$ is called an approximate martingale estimating function. The details are given in section 2. Under the condition that

$\varepsilon \to 0$ and $n \to \infty$, consistency and asymptotic normality of an M-estimator obtained from an approximate martingale estimating function are proved. We also mention that for a special model, an estimator can be asymptotically efficient as $\varepsilon \to 0$ and $n = 1$. The situation where $n = 1$ means that the initial value $X_0 = x_0$ and the terminal value $X_1 = x_1$ are only observed.

This paper is organized as follows. In section 2, we propose an approximate martingale estimating function based on an eigenfunction and state asymptotic results on consistency and asymptotic normality of an M-estimator obtained from the estimating function. Section 3 gives two examples of non-linear diffusion models and simulation studies on the estimators for each model. A discussion on the relation between our estimator and other estimators and a conclusion of this paper are given in section 4. In section 5, the asymptotic results presented in section 2 are proved.

2 Approximate martingale estimating functions

Let θ_0 be the true value of θ and $\theta_0 \in \Theta$. Suppose that the equation (1) has a unique strong solution on $[0, 1]$. Let X_t^0 be the solution of the differential equation corresponding to $\varepsilon = 0$, i.e., $dX_t^0 = b(X_t^0, \theta_0)dt$, $X_0^0 = x_0$. A^* denotes the transpose of a matrix A. Let P_θ be the law of the solution of (1). For the first order differential operator $\tilde{L}_\theta = b(x, \theta)(\partial/\partial x)$, we define the domain $\mathcal{D}(\tilde{L}_\theta)$ of $\tilde{L}_\theta$ as a family of functions $\varphi \in L^2(P_\theta)$ for which $\tilde{L}_\theta\varphi$ is well defined. For example, $\mathcal{D}(\tilde{L}_\theta) = \{\varphi \in L^2(P_\theta) \mid \varphi$ is absolutely continuous and $b(\cdot, \theta)\partial_\cdot\varphi \in L^2(P_\theta)\}$. For $\varphi(x, \theta) \in \mathcal{D}(\tilde{L}_\theta)$, we call $\varphi(x, \theta)$ an eigenfunction for $\tilde{L}_\theta$ associated with eigenvalue $\lambda(\theta)$ if $\tilde{L}_\theta\varphi(x, \theta) = -\lambda(\theta)\varphi(x, \theta)$ for all x in the state space of X under P_θ. Moreover, we introduce several notation as follows.

1. $C_\uparrow^{\infty,k}(\mathbf{R} \times \Theta; \mathbf{R})$ is the space of all functions f satisfying the following two conditions: (i) $f(x, \theta)$ is an $\mathbf{R}$-valued function on $\mathbf{R} \times \Theta$ that is infinitely differentiable with respect to x and continuously differentiable with respect to θ up to order k, (ii) for $n \geq 0$ and $0 \leq |\nu| \leq k$, there exists $C > 0$ such that $\sup_{\theta\in\Theta} |\delta^\nu \partial_x^n f| \leq C(1 + |x|)^C$ for all x. Here $\partial_x = \partial/\partial x$ and $\nu = (\nu_1, \cdots, \nu_p)$ is a multi-index, $|\nu| = \nu_1 + \cdots + \nu_p$, $\delta^\nu = \delta_1^{\nu_1} \cdots \delta_p^{\nu_p}$, $\delta_j = \partial/\partial\theta_j$, $j = 1, \cdots, p$.
2. $C_\uparrow^\infty(\mathbf{R}; \mathbf{R})$ is the set of all functions f of class $C^\infty(\mathbf{R}; \mathbf{R})$ such that f and all of its derivatives have polynomial growth.
3. $C_b^k(\Theta; \mathbf{R})$ is the space of all functions f satisfying the following two conditions: (i) $f(\theta)$ is an $\mathbf{R}$-valued function on Θ that is continuously differentiable with respect to θ up to order k, (ii) for $0 \leq |\nu| \leq k$, there exists $C > 0$ such that $|\delta^\nu f| \leq C$ for all θ.
4. R is a function $\bar{\Theta} \times (0, 1] \times \mathbf{R} \to \mathbf{R}$ for which there exists a constant $C > 0$ such that $|R(\theta, a, x)| \leq aC(1 + |x|)^C$ for all θ, a, x.

We make the assumptions as follows.

- A1. For all $m > 0$, $\sup_t E[|X_t|^m] < \infty$.
- A2. $b(x, \theta) \in C_\uparrow^{\infty,3}(\mathbf{R} \times \bar{\Theta}; \mathbf{R})$, $\sigma(x) \in C_\uparrow^\infty(\mathbf{R}; \mathbf{R}^r)$.
- A3. There exists an eigenfunction $\varphi(x, \theta) \in C_\uparrow^{\infty,2}(\mathbf{R} \times \bar{\Theta}; \mathbf{R})$ with eigenvalue $\lambda(\theta) \in C_b^2(\bar{\Theta}; \mathbf{R})$ such that $\tilde{L}_\theta\varphi(x, \theta) = -\lambda(\theta)\varphi(x, \theta)$ for all x in the state space of X under P_θ.
- A4. $\inf_x \sigma^2(x) > 0$, $\sigma^{-2}(x) \in C_\uparrow^\infty(\mathbf{R}; \mathbf{R})$.

In the same way as in Bibby and Sørensen (1995), we obtain a martingale estimating function $M_{\varepsilon,n}(\theta) = (M_{\varepsilon,n}^{(i)}(\theta))_{1\le i\le p}$, where for $i = 1, 2, \ldots, p$,

$$M_{\varepsilon,n}^{(i)}(\theta) = \sum_{k=1}^{n} (\delta_i b)(X_{t_{k-1}}, \theta)\sigma^{-2}(X_{t_{k-1}})(\varphi(X_{t_k}, \theta) - E_\theta[\varphi(X_{t_k}, \theta)|X_{t_{k-1}}]). \tag{6}$$

However, the conditional expectation $E_\theta[\varphi(X_{t_k}, \theta)|X_{t_{k-1}}]$ does not generally have an explicit form except for several special cases. For special cases, see Kessler and Sørensen (1999) and Uchida (2004b). We therefore consider an approximation of the conditional expectation. It follows from Ito's formula that

$$\begin{aligned}
& e^{\lambda(\theta)t_k}\varphi(X_{t_k}, \theta) - e^{\lambda(\theta)t_{k-1}}\varphi(X_{t_{k-1}}, \theta) \\
= & \int_{t_{k-1}}^{t_k} e^{\lambda(\theta)s}(\lambda(\theta)\varphi(X_s, \theta) + \tilde{L}_\theta\varphi(X_s, \theta))ds + \varepsilon\int_{t_{k-1}}^{t_k} e^{\lambda(\theta)s}(\partial_x\varphi)(X_s, \theta)\sigma(X_s)dw_s \\
& + \frac{1}{2}\varepsilon^2\int_{t_{k-1}}^{t_k} e^{\lambda(\theta)s}(\partial_x^2\varphi)(X_s, \theta)\sigma^2(X_s)ds \\
= & \varepsilon\int_{t_{k-1}}^{t_k} e^{\lambda(\theta)s}(\partial_x\varphi)(X_s, \theta)\sigma(X_s)dw_s + \frac{1}{2}\varepsilon^2\int_{t_{k-1}}^{t_k} e^{\lambda(\theta)s}(\partial_x^2\varphi)(X_s, \theta)\sigma^2(X_s)ds.
\end{aligned}$$

By lemma 6 in Kessler (1997), we can show that

$$\begin{aligned}
E_\theta[\varphi(X_{t_k}, \theta)|X_{t_{k-1}}] &= e^{-\lambda(\theta)/n}\varphi(X_{t_{k-1}}, \theta) \\
&+ \frac{1}{2}\varepsilon^2\int_0^{1/n} E_\theta[e^{-\lambda(\theta)(1/n-u)}(\partial_x^2\varphi)(X_{t_{k-1}+u}, \theta)\sigma^2(X_{t_{k-1}+u})|X_{t_{k-1}}]du \\
&= e^{-\lambda(\theta)/n}\varphi(X_{t_{k-1}}, \theta) + R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right).
\end{aligned} \tag{7}$$

It follows from (7) that $E_\theta[\varphi(X_{t_k}, \theta)|X_{t_{k-1}}]$ can be approximated to $e^{-\lambda(\theta)/n}\varphi(X_{t_{k-1}}, \theta)$ when ε is small. Thus, we consider an estimating function $G_{\varepsilon,n}(\theta) = (G_{\varepsilon,n}^{(i)}(\theta))_{1\le i\le p}$: for $i = 1, 2, \ldots, p$,

$$G_{\varepsilon,n}^{(i)}(\theta) = \sum_{k=1}^{n} (\delta_i b)(X_{t_{k-1}}, \theta)\sigma^{-2}(X_{t_{k-1}})\left[\varphi(X_{t_k}, \theta) - e^{-\lambda(\theta)/n}\varphi(X_{t_{k-1}}, \theta)\right]. \tag{8}$$

Note that the estimating function $G_{\varepsilon,n}(\theta)$ defined by (8) is asymptotically equivalent to the martingale estimating function $M_{\varepsilon,n}(\theta)$ defined by (6), see the proof of lemma 2 put later on. In this sense, the estimating function $G_{\varepsilon,n}(\theta)$ is called an approximate martingale estimating function in this paper.

Let $K_{\varepsilon,n}(\theta) = (K_{\varepsilon,n}^{(ij)}(\theta))_{1\le i,j\le p}$ and $K_{\varepsilon,n}^{(ij)}(\theta) = \delta_j G_{\varepsilon,n}^{(i)}(\theta)$. In order to prove that an M-estimator, which is a solution of the estimating equation $G_{\varepsilon,n}(\theta) = 0$, has consistency and asymptotic normality, it is essential to show both the convergence of $K_{\varepsilon,n}(\theta)$ in probability uniformly in θ and the convergence of $G_{\varepsilon,n}(\theta_0)$ in distribution, see lemmas 1 and 2 put later on. Let $\xrightarrow{p}$ and $\xrightarrow{d}$ denote the convergence in probability and the convergence in distribution, respectively. In lemma 1, we define that $K(\theta) = (K^{(ij)}(\theta))_{1\le i,j\le p}$ and

$$\begin{aligned}
& K^{(ij)}(\theta) \\
= & \int_0^1 (\delta_j\delta_i b)(X_s^0, \theta)\sigma^{-2}(X_s^0)\{b(X_s^0, \theta_0)\partial_x\varphi(X_s^0, \theta) + \lambda(\theta)\varphi(X_s^0, \theta)\}ds \\
+ & \int_0^1 (\delta_i b)(X_s^0, \theta)\sigma^{-2}(X_s^0)\{b(X_s^0, \theta_0)\partial_x\delta_j\varphi(X_s^0, \theta) + \lambda(\theta)\delta_j\varphi(X_s^0, \theta) + (\delta_j\lambda)(\theta)\varphi(X_s^0, \theta)\}ds.
\end{aligned}$$

In lemma 2, we set that $A(\theta_0) = (A^{(ij)}(\theta_0))_{1\le i,j\le p}$, where

$$A^{(ij)}(\theta_0) = \int_0^1 (\delta_i b)(X_s^0, \theta_0)\sigma^{-2}(X_s^0)(\partial_x\varphi)^2(X_s^0, \theta_0)(\delta_j b)(X_s^0, \theta_0)ds.$$

Lemma 1 *Assume A1–A4. Then, as $\varepsilon \to 0$ and $n \to \infty$,*

$$\sup_{\theta\in\bar\Theta} |K_{\varepsilon,n}(\theta) - K(\theta)| \stackrel{p}{\longrightarrow} 0.$$

Lemma 2 *Assume A1–A4. Then, as $\varepsilon \to 0$ and $n \to \infty$,*

$$\varepsilon^{-1}G_{\varepsilon,n}(\theta_0) \stackrel{d}{\longrightarrow} N(0, A(\theta_0)).$$

Let $\hat\theta_{\varepsilon,n}$ be an M-estimator defined as a solution of the estimating equation that $G_{\varepsilon,n}(\theta) = 0$. We then have the following asymptotic result of the M-estimator $\hat\theta_{\varepsilon,n}$.

Theorem 1 *Let $\gamma \in (0, 1)$. Suppose that A1–A4 hold true. Moreover, suppose that there exists an open set $\tilde\Theta$ including θ_0 such that*

$$\inf_{\theta_1,\theta_2\in\tilde\Theta,|x|=1} \left| \left(\int_0^1 K(\theta_1 + s(\theta_2 - \theta_1))ds \right)^* x \right| > 0.$$

Then, as $\varepsilon \to 0$ and $n \to \infty$,

$$P_{\theta_0}[(\exists_1\hat\theta_{\varepsilon,n} \in \tilde\Theta \text{ such that } G_{\varepsilon,n}(\hat\theta_{\varepsilon,n}) = 0) \text{ and } (|\hat\theta_{\varepsilon,n} - \theta_0| \le \varepsilon^\gamma)] \to 1$$

and

$$\varepsilon^{-1}(\hat\theta_{\varepsilon,n} - \theta_0) \stackrel{d}{\longrightarrow} N\left(0, K(\theta_0)^{-1}A(\theta_0)(K^*(\theta_0))^{-1}\right).$$

Note that $K(\theta)$ used in the additional assumption of theorem 1 is the same as in lemma 1. Moreover, it follows from A3 that $K(\theta_0)$ appearing in the asymptotic variance of $\varepsilon^{-1}(\hat\theta_{\varepsilon,n} - \theta_0)$ has the following simple form.

$$K^{(ij)}(\theta_0) = \int_0^1 \frac{(\delta_i b)(X_s^0, \theta_0)}{\sigma^2(X_s^0)} \{b(X_s^0, \theta_0)\partial_x\delta_j\varphi(X_s^0, \theta_0) + \lambda(\theta_0)\delta_j\varphi(X_s^0, \theta_0) + (\delta_j\lambda)(\theta_0)\varphi(X_s^0, \theta_0)\}ds.$$

In order to obtain theorem 1, we can relax the assumptions A2–A4. By using a well-known localization argument, A2–A4 can be replaced with milder regularity conditions about b, σ and φ in the neighborhood of the path of X_t^0.

The asymptotic variance of $\varepsilon^{-1}(\hat\theta_{\varepsilon,n} - \theta_0)$ depends on the eigenfunction $\varphi(x, \theta)$ and the eigenvalue $\lambda(\theta)$. For this reason, it seems that it is very much important to propose an **optimal** estimating function and to choose a **nice** eigenfunction $\varphi(x, \theta)$ with an eigenvalue $\lambda(\theta)$ for which the asymptotic variance of $\varepsilon^{-1}(\hat\theta_{\varepsilon,n} - \theta_0)$ becomes as small as possible. However, we do not have to care the above optimality. As stated in section 1, we have already taken the efficient estimating function $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$ in (4) and the efficient one-step estimator in (5). Therefore, using the M-estimator presented in theorem 1 as an initial estimator, we can necessarily get an asymptotically efficient estimator.

Although the estimating function we treat is released from a discussion on optimality, one still has the question why only the estimating function $G_{\varepsilon,n}(\theta)$ defined by (8) is considered. We never say that the estimating function is best. Since we are going to use the estimator obtained from the estimating function as an initial estimator, we must get a **unique** root of the estimating equation and it is better to obtain the root **explicitly**. For this reason, we proposed the estimating function with a simple expression. Thus, we completely understand that the important thing in this paper is to obtain an estimating function which derives an **unique** estimator with an **explicit** form, but not to consider an optimal estimating function. One of the answers is our estimating function $G_{\varepsilon,n}(\theta)$ in (8).

The next question is how to choose an eigenfunction $\varphi(x,\theta)$ with an eigenvalue $\lambda(\theta)$. It suffices to find out a non-trivial eigenfunction $\varphi(x,\theta)$ associated with an eigenvalue $\lambda(\theta)$ for which A3 holds. It follows from A3 that

$$\frac{\varphi(x,\theta)}{\varphi(x_0,\theta)} = \exp\left\{-\lambda(\theta)\int_{x_0}^{x}\frac{1}{b(y,\theta)}dy\right\}.$$

In the case that $b(x,\theta) = \theta g(x)$ for a function g, in order to get a simple estimating function, it is natural to set that $\lambda(\theta) = -\theta$. We then have that

$$\frac{\varphi(x)}{\varphi(x_0)} = \exp\left\{\int_{x_0}^{x}\frac{1}{g(y)}dy\right\}.$$

For example, if $g(x) = \cos x$, one has that

$$\varphi(x) = \frac{\cos(x/2) + \sin(x/2)}{\cos(x/2) - \sin(x/2)}.$$

As another illustration with one parameter, we will consider the hyperbolic diffusion process in section 3.1 put later on. An example of a model with two unknown parameters is the case that $b(x,\theta) = \theta_1/x - \theta_2 x$. Since

$$\frac{\varphi(x,\theta)}{\varphi(x_0,\theta)} = \exp\left\{-\lambda(\theta)\int_{x_0}^{x}\frac{y}{\theta_1 - \theta_2 y^2}dy\right\},$$

one has that $\lambda(\theta) = 2\theta_2$ and $\varphi(x,\theta) = \theta_1 - \theta_2 x^2$. As another example with two parameters, the population growth model is treated in section 3.2 put later on.

Theorem 1 ensures the asymptotic normality of the M-estimator $\hat{\theta}_{\varepsilon,n}$ as $\varepsilon \to 0$ and $n \to \infty$, while as for somewhat special cases, it is possible to obtain an asymptotically efficient estimator as $\varepsilon \to 0$. It means that we can obtain an asymptotically efficient estimator when ε is small but n is not large. We will spend the rest of this section on a contribution of this direction.

Consider the model (1) where $\theta \in \Theta \subset \mathbf{R}$ and $b(x,\theta) = \theta g(x)$. In this case, set that

$$\lambda(\theta) = -\theta, \quad \varphi(x) = \exp\left\{\int_{x_0}^{x}\frac{1}{g(y)}dy\right\}.$$

If $n = 1$, that is, we get only the initial value $X_0 = x_0$ and the terminal value $X_1 = x_1$, it follows from (8) that the estimator is

$$\hat{\theta}_{\varepsilon,1} \quad = \quad \log\varphi(X_1) - \log\varphi(X_0).$$

Let $I(\theta_0) = \int_0^1 g^2(X_s^0)/\sigma^2(X_s^0)ds$ and $J(\theta_0) = \int_0^1 \sigma^2(X_s^0)/g^2(X_s^0)ds$, where $I(\theta_0)$ is the asymptotic Fisher information of the diffusion process from the continuous time observations. Suppose that $I(\theta_0) \neq 0$ and that $J(\theta_0) < \infty$. Note that $J(\theta_0) \geq I(\theta_0)^{-1}$.

Under the assumption that $\int_0^1 (\partial_x g)(X_s^0)\sigma^2(X_s^0)/g^2(X_s^0)ds < \infty$, Ito's formula yields that

$$\begin{aligned}\log\varphi(X_1) - \log\varphi(X_0) &= \int_0^1 \frac{\tilde{L}_{\theta_0}\varphi(X_s)}{\varphi(X_s)}ds + \varepsilon\int_0^1 (\partial_x \log\varphi)(X_s)\sigma(X_s)dw_s \\ &\quad +\frac{1}{2}\varepsilon^2\int_0^1 (\partial_x^2 \log\varphi)(X_s)\sigma^2(X_s)ds \\ &= \theta_0 + \varepsilon\int_0^1 (\partial_x \log\varphi)(X_s)\sigma(X_s)dw_s + O_p(\varepsilon^2).\end{aligned}$$

It is easy to show that in distribution under P_{θ_0}, as $\varepsilon \to 0$,

$$\varepsilon^{-1}(\hat{\theta}_{\varepsilon,1} - \theta_0) \to N(0, J(\theta_0)).$$

Moreover, if $g^2(x) = t\sigma^2(x)$ for some $t > 0$, then

$$\varepsilon^{-1}(\hat{\theta}_{\varepsilon,1} - \theta_0) \to N\left(0, I(\theta_0)^{-1}\right)$$

in distribution under P_{θ_0} as $\varepsilon \to 0$.

For example, if we treat a small random perturbation of the geometric Brownian motion, which is the case that $b(x,\theta) = \theta x$ and $\sigma(x) = x$, we can derive the asymptotically efficient estimator by only two observations, $X_0 = x_0$ and $X_1 = x_1$. Even though it is certain that this result can be applied to the restricted models, it seems attractive to be able to construct an asymptotically efficient estimator even if $n = 1$. For the estimator proposed in Laredo (1990), we can observe a similar phenomenon.

3 Examples and simulation studies

Through simulations, we examine the asymptotic behaviour of both the estimator $\hat{\theta}_{\varepsilon,n}$ stated in section 2 and the one-step estimator $\hat{\theta}^{(1)}_{\varepsilon,n,\ell}$ with $\ell = 3$, which is constructed by using the estimator $\hat{\theta}_{\varepsilon,n}$ as an initial estimator. For details, see (5). For two non-linear models, the simulations are done for each $\varepsilon = 0.1, 0.05, 0.01$ and $n = 5, 10, 50$. For a true parameter value θ_0 and an initial value x_0, 10000 independent sample paths are generated by the Euler-Maruyama scheme. For the Euler-Maruyama scheme, see Kloeden and Platen (1992). In order to evaluate the two estimators $\hat{\theta}_{\varepsilon,n}$ and $\hat{\theta}^{(1)}_{\varepsilon,n,3}$, we also examine the estimator in corollary 1 of Genon-Catalot (1990) and the estimator in Laredo (1990), which are called the simple estimator $\hat{\theta}^{(S)}_{\varepsilon,n}$ and Laredo's estimator $\hat{\theta}^{(L)}_{\varepsilon,n}$, respectively. We note that the second contrast function of Genon-Catalot (1990) and the martingale estimating function of Uchida (2004a) are not obtained explicitly for two non-linear models we treat. For each of the four estimators, that is, $\hat{\theta}^{(S)}_{\varepsilon,n}$, $\hat{\theta}^{(L)}_{\varepsilon,n}$, $\hat{\theta}_{\varepsilon,n}$ and $\hat{\theta}^{(1)}_{\varepsilon,n,3}$, its mean and its standard deviation are computed.

3.1 The hyperbolic diffusion model

Consider a one-dimensional diffusion process defined by the stochastic differential equation

$$dX_t = \theta\frac{X_t}{\sqrt{1+X_t^2}}dt + \varepsilon dw_t,\ t \in [0,1],\ \varepsilon \in (0,1],\ X_0 = x_0,$$

where x_0 and ε are known constants and θ is an unknown parameter. This model is called the hyperbolic diffusion process. For details, see Bibby and Sørensen (1995).

Solving the differential equation $\tilde{L}_\theta \varphi(x,\theta) = -\lambda(\theta)\varphi(x,\theta)$, where $\tilde{L}_\theta = \theta \frac{x}{\sqrt{1+x^2}}\partial_x$, one has

$$\varphi(x,\theta) = \exp\left\{-\frac{\lambda(\theta)}{\theta}\left(\sqrt{1+x^2} + \log\frac{x}{1+\sqrt{1+x^2}}\right)\right\}.$$

Therefore, by setting that $\lambda(\theta) = -\theta$, a non-trivial eigenfunction is

$$\varphi(x) = \frac{x}{1+\sqrt{1+x^2}}\exp\{\sqrt{1+x^2}\}.$$

Note that we can obtain the eigenfunction $\varphi(x)$ independent of a parameter θ by putting that $\lambda(\theta) = -\theta$. The approximate martingale estimating function is given by

$$G_{\varepsilon,n}(\theta) = \sum_{k=1}^{n} \frac{X_{t_{k-1}}}{\sqrt{1+X_{t_{k-1}}^2}}\left[\varphi(X_{t_k}) - e^{\theta/n}\varphi(X_{t_{k-1}})\right].$$

By the estimating equation $G_{\varepsilon,n}(\theta) = 0$, one has the estimator

$$\hat{\theta}_{\varepsilon,n} = n\left\{\log\left(\sum_{k=1}^{n}\frac{X_{t_{k-1}}}{\sqrt{1+X_{t_{k-1}}^2}}\varphi(X_{t_k})\right) - \log\left(\sum_{k=1}^{n}\frac{X_{t_{k-1}}}{\sqrt{1+X_{t_{k-1}}^2}}\varphi(X_{t_{k-1}})\right)\right\}.$$

It follows from Theorem 1 that the asymptotic variance of $\varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0)$ is

$$\frac{A(\theta_0)}{K(\theta_0)^2} = \frac{\int_0^1 (\varphi(X_s^0))^2 ds}{\left(\int_0^1 (X_s^0/\sqrt{1+(X_s^0)^2})\varphi(X_s^0)ds\right)^2},$$

where X_t^0 is a solution of the differential equation $dX_t^0 = \theta_0 X_t^0/\sqrt{1+(X_t^0)^2}dt$, $X_0^0 = x_0$.

Table 1 shows means and standard deviations of the four estimators for $\theta_0 = 10$ and $x_0 = 0.1$. For the case that $n \le 10$, $\hat{\theta}_{\varepsilon,n}^{(S)}$ has a considerable bias and $\hat{\theta}_{\varepsilon,n}^{(L)}$ has a small bias. Even if n is small, both $\hat{\theta}_{\varepsilon,n}$ and $\hat{\theta}_{\varepsilon,n,3}^{(1)}$ are unbiased and $\hat{\theta}_{\varepsilon,n,3}^{(1)}$ works well with a small variance.

Table 1: (The hyperbolic diffusion model) The mean and standard deviation (s.d.) of the four estimators for 10000 independent simulated sample paths with $\theta_0 = 10$ and $x_0 = 0.1$.

		$\hat{\theta}_{\varepsilon,n}^{(S)}$		$\hat{\theta}_{\varepsilon,n}^{(L)}$		$\hat{\theta}_{\varepsilon,n}$		$\hat{\theta}_{\varepsilon,n,3}^{(1)}$	
ε	n	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
0.10	5	10.61274	0.13204	9.82586	0.11806	10.00009	0.19902	9.96889	0.12561
0.10	10	10.31695	0.12320	9.95102	0.11901	10.00290	0.22007	9.99505	0.11989
0.10	50	10.06267	0.11989	9.99447	0.11929	10.00237	0.22974	9.99461	0.11932
0.05	5	10.61570	0.06611	9.82735	0.05875	9.99996	0.09947	9.96288	0.06379
0.05	10	10.31708	0.06145	9.95090	0.05937	10.00136	0.10995	9.99622	0.05981
0.05	50	10.06357	0.05980	9.99433	0.05950	10.00109	0.11475	9.99574	0.05952
0.01	5	10.61671	0.01323	9.82809	0.01173	9.99983	0.01989	9.96164	0.01285
0.01	10	10.31735	0.01228	9.95119	0.01186	10.00012	0.02198	9.99689	0.01195
0.01	50	10.06417	0.01195	9.99462	0.01189	10.00007	0.02293	9.99643	0.01189

3.2 The population growth model

We consider another non-linear model defined by the stochastic differential equation

$$dX_t = \alpha X_t(\beta - X_t)dt + \varepsilon X_t dw_t,\ t \in [0,1],\ \varepsilon \in (0,1],\ X_0 = x_0,$$

where x_0 and ε are known constants and α and β are unknown parameters. Here we assume that the state space is the positive real line. The diffusion process is one of population models, and often used as a population growth model in a stochastic, crowded environment. The parameter $\beta > 0$ is called the carrying capacity of the environment. The parameter $\alpha \in \mathbf{R}$ is a measure of the quality of the environment. For more details, see (5.3.9) on page 78 in Øksendal (2003).

Let $\theta = (\alpha, \beta)$. In this case, $L_\theta = \alpha x(\beta - x)\partial_x$. By solving the differential equation $L_\theta \varphi(x,\theta) = -\lambda(\theta)\varphi(x,\theta)$, an eigenfunction is given by

$$\varphi(x,\theta) = \left(\frac{\beta - x}{x}\right)^{\frac{\lambda(\theta)}{\alpha\beta}}.$$

Therefore, it is natural to set that $\lambda(\theta) = \alpha\beta$ because a non-trivial eigenfunction is given by

$$\varphi(x,\beta) = (\beta - x)/x$$

and the eigenfunction is independent of α. The approximate martingale estimating functions with respect to α and β are

$$G^{(1)}_{\varepsilon,n}(\theta) = \sum_{k=1}^{n} X^{-1}_{t_{k-1}}(\beta - X_{t_{k-1}})(\varphi(X_{t_k},\beta) - e^{-\alpha\beta/n}\varphi(X_{t_{k-1}},\beta)),$$

$$G^{(2)}_{\varepsilon,n}(\theta) = \sum_{k=1}^{n} \alpha X^{-1}_{t_{k-1}}(\varphi(X_{t_k},\beta) - e^{-\alpha\beta/n}\varphi(X_{t_{k-1}},\beta)),$$

respectively. It then follows that

$$\hat{\alpha}_{\varepsilon,n} = -\frac{n}{\hat{\beta}_{\varepsilon,n}} \log \frac{\sum_{k=1}^{n}\varphi(X_{t_k},\hat{\beta}_{\varepsilon,n})}{\sum_{k=1}^{n}\varphi(X_{t_{k-1}},\hat{\beta}_{\varepsilon,n})},$$

$$\hat{\beta}_{\varepsilon,n} = \frac{\sum_{k=1}^{n}\frac{1}{X_{t_{k-1}}}\left(\sum_{k=1}^{n}\frac{1}{X_{t_{k-1}}} - \sum_{k=1}^{n}\frac{1}{X_{t_k}}\right) + n\left(\sum_{k=1}^{n}\frac{1}{X_{t_{k-1}}X_{t_k}} - \sum_{k=1}^{n}\frac{1}{X^2_{t_{k-1}}}\right)}{\sum_{k=1}^{n}\frac{1}{X_{t_{k-1}}X_{t_k}}\sum_{k=1}^{n}\frac{1}{X_{t_{k-1}}} - \sum_{k=1}^{n}\frac{1}{X_{t_k}}\sum_{k=1}^{n}\frac{1}{X^2_{t_{k-1}}}}.$$

An easy computation implies that $K(\theta_0)$ and $A(\theta_0)$, which are concerned with the asymptotic variance of $\varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0)$, are given by

$$K^{(11)}(\theta_0) = \beta_0 \int_0^1 \left(\frac{\beta_0}{X^0_s} - 1\right)^2 ds, \quad K^{(22)}(\theta_0) = \alpha_0^2\beta_0 \int_0^1 \left(\frac{1}{X^0_s}\right)^2 ds,$$

$$K^{(12)}(\theta_0) = K^{(21)}(\theta_0) = \alpha_0\beta_0 \int_0^1 \left(\frac{\beta_0}{X^0_s} - 1\right)\frac{1}{X^0_s} ds,$$

$$A^{(11)}(\theta_0) = \beta_0^2 \int_0^1 \left(\frac{\beta_0}{X^0_s} - 1\right)^2 \left(\frac{1}{X^0_s}\right)^2 ds, \quad A^{(22)}(\theta_0) = \alpha_0^2\beta_0^2 \int_0^1 \left(\frac{1}{X^0_s}\right)^4 ds,$$

$$A^{(12)}(\theta_0) = K^{(21)}(\theta_0) = \alpha_0\beta_0^2 \int_0^1 \left(\frac{\beta_0}{X^0_s} - 1\right)\left(\frac{1}{X^0_s}\right)^3 ds,$$

where X_t^0 is a solution of the differential equation: $dX_t^0 = \alpha_0 X_t^0(\beta_0 - X_t^0)dt$, $X_0^0 = x_0$, that is,

$$X_t^0 = \frac{\exp\{\alpha_0\beta_0 t\}}{1/x_0 + (\exp\{\alpha_0\beta_0 t\} - 1)/\beta_0}.$$

Under the situation where $\alpha_0 = 0.5$, $\beta_0 = 20$ and $x_0 = 1$, simulation results of the four estimators for α and β are given in tables 2 and 3, respectively. Even if $n = 5$, both $\hat{\alpha}_{\varepsilon,n}$ and $\hat{\alpha}^{(1)}_{\varepsilon,n,3}$ are unbiased, while $\hat{\alpha}^{(S)}_{\varepsilon,n}$ has a considerable bias and $\hat{\alpha}^{(L)}_{\varepsilon,n}$ has a small bias. Although $\hat{\alpha}_{\varepsilon,n}$ has a greater standard deviation than $\hat{\alpha}^{(L)}_{\varepsilon,n}$, we see that $\hat{\alpha}^{(1)}_{\varepsilon,n,3}$ recovers a reasonable standard deviation in all cases. We can say that $\hat{\beta}^{(1)}_{\varepsilon,n,3}$ performs quite well in all cases. For estimation of β, both $\hat{\beta}_{\varepsilon,n}$ and $\hat{\beta}^{(1)}_{\varepsilon,n,3}$ are unbiased in all cases. In particular, $\hat{\beta}^{(1)}_{\varepsilon,n,3}$ works well for the case that $n \geq 10$. Furthermore, it is mentioning that $\hat{\beta}^{(1)}_{\varepsilon,n,3}$ is a good estimator even if $\varepsilon = 0.1$ or $n = 5$.

Table 2: (The population growth model) The mean and standard deviation (s.d.) of the four estimators for 10000 independent simulated sample paths with $\alpha_0 = 0.5$, $\beta_0 = 20$ and $x_0 = 1$.

		$\hat{\alpha}^{(S)}_{\varepsilon,n}$		$\hat{\alpha}^{(L)}_{\varepsilon,n}$		$\hat{\alpha}_{\varepsilon,n}$		$\hat{\alpha}^{(1)}_{\varepsilon,n,3}$	
ε	n	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
0.10	5	1.06660	0.04736	0.53440	0.01782	0.49784	0.01846	0.49702	0.01682
0.10	10	0.72208	0.02541	0.50706	0.01543	0.49770	0.02613	0.49716	0.01500
0.10	50	0.53587	0.01617	0.49796	0.01463	0.49816	0.03321	0.50139	0.01540
0.05	5	1.06625	0.02359	0.53455	0.00889	0.49768	0.00923	0.49717	0.00841
0.05	10	0.72170	0.01267	0.50736	0.00771	0.49736	0.01306	0.49676	0.00749
0.05	50	0.53536	0.00807	0.49845	0.00731	0.49739	0.01657	0.49895	0.00743
0.01	5	1.06616	0.00471	0.53461	0.00178	0.49766	0.00185	0.49723	0.00168
0.01	10	0.72156	0.00253	0.50744	0.00154	0.49723	0.00261	0.49663	0.00149
0.01	50	0.53518	0.00161	0.49860	0.00146	0.49709	0.00331	0.49820	0.00146

Table 3: (The population growth model) The mean and standard deviation (s.d.) of the four estimators for 10000 independent simulated sample paths with $\alpha_0 = 0.5$, $\beta_0 = 20$ and $x_0 = 1$.

		$\hat{\beta}^{(S)}_{\varepsilon,n}$		$\hat{\beta}^{(L)}_{\varepsilon,n}$		$\hat{\beta}_{\varepsilon,n}$		$\hat{\beta}^{(1)}_{\varepsilon,n,3}$	
ε	n	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
0.10	5	18.23044	0.28863	19.57295	0.29307	20.01498	0.33629	19.93971	0.31131
0.10	10	19.04767	0.26286	19.89671	0.27549	20.03362	0.50570	20.00064	0.27844
0.10	50	19.80643	0.26379	20.01152	0.26848	20.04057	0.69266	19.92236	0.30189
0.05	5	18.23015	0.14397	19.57255	0.14619	20.01901	0.16806	19.93278	0.15524
0.05	10	19.04734	0.13112	19.89487	0.13739	20.03012	0.25282	20.01452	0.13891
0.05	50	19.80733	0.13161	20.00758	0.13384	20.03408	0.34647	19.99032	0.13718
0.01	5	18.22932	0.02874	19.57185	0.02919	20.01929	0.03364	19.92993	0.03098
0.01	10	19.04663	0.02618	19.89377	0.02743	20.02902	0.05071	20.01837	0.02773
0.01	50	19.80708	0.02628	20.00580	0.02671	20.03249	0.06956	20.01018	0.02678

4 Discussion

In order to estimate an unknown parameter θ in the model (1), several kinds of estimators can be considered. It is natural to ask the following question. "Which of the estimators is best in the sense of the first order asymptotics?" As far as the author knows, the answer is as follows.

(I) If the conditional expectation $E_\theta[X_{t_{k-1}}|X_{t_k}]$ can be obtained explicitly, we should use the martingale estimating function $\mathcal{M}_{\varepsilon,n}(\theta)$ in (2). The estimator derived from $\mathcal{M}_{\varepsilon,n}(\theta)$ has asymptotic efficiency as $\varepsilon \to 0$ and $n \to \infty$. For example, when $b(x,\theta) = \theta_1 + \theta_2 x$ for $\theta_2 \neq 0$, one has that $E_\theta[X_{t_k}|X_{t_{k-1}}] = \exp\{\theta_2/n\}X_{t_{k-1}} + (\exp\{\theta_2/n\} - 1)\theta_1/\theta_2$. However, there is a disadvantage that $E_\theta[X_{t_{k-1}}|X_{t_k}]$ does not generally have an explicit form for diffusion processes.

(II) If we can not obtain the estimator of (I), the second best is the estimator obtained from the second contrast function presented in Genon-Catalot (1990). The estimator is asymptotically efficient under $\varepsilon\sqrt{n} = O(1)$ as $\varepsilon \to 0$ and $n \to \infty$. Note that in order to obtain the contrast function explicitly, we need to compute the solution X_t^0 satisfying that $dX_t^0 = b(X_t^0, \theta)dt$, $X_0^0 = x_0$. Moreover, we have to get $H_t = \exp \int_0^t \frac{\partial b}{\partial x}(X_s^0, \theta)ds$ and $K_t = \int_{t_{k-1}}^{t_k} H_s^{-2}\sigma^2(X_s)ds$.

(III) If the estimator of (II) can not be obtained explicitly, we try to get the estimator proposed by Laredo (1990), which has asymptotic efficiency when $(\varepsilon n^2)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. The contrast function is constructed by using the function $V(x,\theta)$ satisfying that $\partial V(x,\theta)/\partial x = \sigma^{-2}(x)b(x,\theta)$, $V(x_0,\theta) = 0$. However, it is not certain that $V(x,\theta)$ can be obtained explicitly.

The above three estimating functions do not generally have explicit expressions, while the following two estimating functions can be necessarily obtained.

(IV) If the estimating functions in (I), (II), (III) can not be obtained explicitly and if the asymptotics we treat is when $(\varepsilon n)^{-1} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$, it is enough to get the asymptotically efficient estimator obtained from the first contrast function presented in Genon-Catalot (1990). However, as seen in two examples of section 3, the estimator has a considerable bias when ε is not so small and n is not so large.

(V) If we face to the situation for which (I)-(IV) do not work, we consider the approximate martingale estimating function $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$ in (4). In order to obtain an asymptotically efficient estimator derived from $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$, we use a one-step estimator with an initial estimator obtained from the approximate martingale estimating function $G_{\varepsilon,n}(\theta)$ in (8). For this reason, we should turn our attention to simplicity of $G_{\varepsilon,n}(\theta)$ rather than optimality of $G_{\varepsilon,n}(\theta)$.

In conclusion, we can say that in order to obtain an asymptotically efficient estimator, we should seek a suitable estimator among (I)-(V) based on the asymptotics we treat. However, we must note that even if the estimating functions of (I)-(IV) are explicitly obtained, it is not always to get an explicit estimator from the estimating equation. For such ill cases, we need to consider a one-step estimator and it is important to use the estimator proposed in this paper as an initial estimator. If these procedures are niggling, we will recommend to use the estimator in (V). As seen in this paper, $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$ in (4) is always obtained explicitly for diffusion processes, and in most cases an initial estimator can be derived from $G_{\varepsilon,n}(\theta)$ in (8).

Finally, we discuss an extension of this paper to a multivariate diffusion process with a small perturbation. Both $\mathcal{G}_{\varepsilon,n,\ell}(\theta)$ in (4) and the one-step estimator in (5) work for a multivariate diffusion process with small noise. Therefore, it is important to get an initial estimator for the multivariate diffusion process. From the theoretical point of view, it is possible to extend the result of this paper to a d-dimensional diffusion process. However, in practice, we need to find an explicit eigenfunction for $\tilde{L}_\theta = \sum_{j=1}^d b^j(x,\theta)\frac{\partial}{\partial x_j}$. Compared with a one-dimensional diffusion process, this is an arduous undertaking. Therefore, we need to consider the more convenient procedure as a future work.

5 Proofs

For proofs of the results, we introduce notation and functions. Let $U(x,\theta) = \varphi(x,\theta) - \varphi(X_{t_{k-1}},\theta)$ and $\mathcal{G}_k^n = \sigma(w_s; s \le t_k)$. Define that $C_\uparrow^{1,1}(\mathbf{R}\times\Theta;\mathbf{R})$ is the space of all functions $f : \mathbf{R}\times\Theta \to \mathbf{R}$ such that (i) f is continuously differentiable with respect to x and θ, (ii) f and its derivatives are of polynomial growth in x uniformly in θ.

In order to show lemmas 1 and 2, we need the following lemma.

Lemma 3 *Let $f \in C_\uparrow^{1,1}(\mathbf{R}\times\bar{\Theta};\mathbf{R})$. Assume A1–A3. Then, as $\varepsilon \to 0$ and $n \to \infty$,*
(i)

$$\sup_{\theta\in\bar{\Theta}} \left| \frac{1}{n}\sum_{k=1}^n f(X_{t_{k-1}},\theta) - \int_0^1 f(X_s^0,\theta)ds \right| \xrightarrow{p} 0,$$

(ii)

$$\sup_{\theta\in\bar{\Theta}} \left| \sum_{k=1}^n f(X_{t_{k-1}},\theta)U(X_{t_k},\theta) - \int_0^1 f(X_s^0,\theta)b(X_s^0,\theta_0)\partial_x\varphi(X_s^0,\theta)ds \right| \xrightarrow{p} 0,$$

(iii) for $i = 1,\dots,p$,

$$\sup_{\theta\in\bar{\Theta}} \left| \sum_{k=1}^n f(X_{t_{k-1}},\theta)(\delta_i U)(X_{t_k},\theta) - \int_0^1 f(X_s^0,\theta)b(X_s^0,\theta_0)\partial_x\delta_i\varphi(X_s^0,\theta)ds \right| \xrightarrow{p} 0.$$

Proof of Lemma 3. (i) We can prove (i) along the same lines as the proof of lemma 8 in Kessler (1997). See also lemma 4-(i) in Uchida (2004b). ∎

(ii) By noting that

$$L_{\theta_0}U(x,\theta) \;=\; b(x,\theta_0)\partial_x\varphi(x,\theta) + \frac{1}{2}\varepsilon^2\sigma^2(x)\partial_x^2\varphi(x,\theta),$$

Lemma 1 in Florens-Zmirou (1989) implies that

$$\begin{aligned} E_{\theta_0}[U(X_{t_k},\theta)|\mathcal{G}_{k-1}^n] &= U(X_{t_{k-1}},\theta) + \frac{1}{n}L_{\theta_0}U(X_{t_{k-1}},\theta) \\ &\quad + \int_0^{\frac{1}{n}}\int_0^{u_1} E_{\theta_0}[L_{\theta_0}^2 U(X_{t_{k-1}+u_2},\theta)|\mathcal{G}_{k-1}^n]du_2du_1 \\ &= \frac{1}{n}b(X_{t_{k-1}},\theta_0)\partial_x\varphi(X_{t_{k-1}},\theta) \\ &\quad + R\left(\theta,\frac{\varepsilon^2}{n},X_{t_{k-1}}\right) + R\left(\theta,\frac{1}{n^2},X_{t_{k-1}}\right). \end{aligned} \tag{9}$$

Similarly, we see that

$$\begin{aligned} & E_{\theta_0}[(U(X_{t_k},\theta))^2|\mathcal{G}_{k-1}^n] \\ =\ & (U(X_{t_{k-1}},\theta))^2 + \frac{1}{n}L_{\theta_0}(U(X_{t_{k-1}},\theta))^2 + \int_0^{\frac{1}{n}}\int_0^{u_1} E_{\theta_0}[L_{\theta_0}^2(U(X_{t_{k-1}+u_2},\theta))^2|\mathcal{G}_{k-1}^n]du_2du_1 \\ =\ & \frac{\varepsilon^2}{n}\sigma^2(X_{t_{k-1}})(\partial_x\varphi(X_{t_{k-1}},\theta))^2 + R\left(\theta,\frac{1}{n^2},X_{t_{k-1}}\right). \end{aligned} \tag{10}$$

Let $\xi_k(\theta) = f(X_{t_{k-1}}, \theta)\{U(X_{t_k}, \theta) - b(X_{t_{k-1}}, \theta_0)\partial_x\varphi(X_{t_{k-1}}, \theta)/n\}$. By (9), (10) and lemma 3–(i), we see that as $\varepsilon \to 0$ and $n \to \infty$,

$$\sum_{k=1}^n E[\xi_k(\theta)|\mathcal{G}_{k-1}^n] = \sum_{k=1}^n \left\{ R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right)\right\} \xrightarrow{p} 0,$$

$$\sum_{k=1}^n E[(\xi_k(\theta))^2|\mathcal{G}_{k-1}^n] = \sum_{k=1}^n \left\{ R\left(\theta, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\theta, \frac{1}{n^2}, X_{t_{k-1}}\right)\right\} \xrightarrow{p} 0.$$

Lemma 9 in Genon-Catalot and Jacod (1993) yields that $\sum_{k=1}^n \xi_k(\theta) \xrightarrow{p} 0$ as $\varepsilon \to 0$ and $n \to \infty$. For the tightness of $\sum_{k=1}^n \xi_k(\cdot)$, it suffices to show the following two inequalities, see for instance theorem 20 in appendix I in Ibragimov and Has'minskii (1981):

$$E_{\theta_0}\left[\left(\sum_{k=1}^n \xi_k(\theta)\right)^{2m}\right] \leq C, \tag{11}$$

$$E_{\theta_0}\left[\left(\sum_{k=1}^n \xi_k(\theta_2) - \sum_{k=1}^n \xi_k(\theta_1)\right)^{2m}\right] \leq C|\theta_2 - \theta_1|^{2m}, \tag{12}$$

for $\theta, \theta_1, \theta_2 \in \bar{\Theta}$, where $m > p/2$.

Ito's formula implies that $\xi_k(\theta) = A_{k,1}(\theta) + A_{k,2}(\theta) - A_{k,3}(\theta)$, where

$$\begin{aligned} A_{k,1}(\theta) &= f(X_{t_{k-1}}, \theta)\int_{t_{k-1}}^{t_k} L_{\theta_0}\varphi(X_{t_k}, \theta)ds, \\ A_{k,2}(\theta) &= \varepsilon f(X_{t_{k-1}}, \theta)\int_{t_{k-1}}^{t_k} (\partial_x\varphi)(X_s, \theta)\sigma(X_s)dw_s, \\ A_{k,3}(\theta) &= \frac{1}{n} f(X_{t_{k-1}}, \theta) b(X_{t_{k-1}}, \theta_0)\partial_x\varphi(X_{t_{k-1}}, \theta). \end{aligned}$$

It follows from the Burkholder-Davis-Gundy inequality that

$$\begin{aligned} E_{\theta_0}\left[\left|\sum_{k=1}^n A_{k,2}(\theta)\right|^{2m}\right] &\leq C_{2m}\varepsilon^{2m} E_{\theta_0}\left[\left(\sum_{k=1}^n \int_{t_{k-1}}^{t_k} f(X_{t_{k-1}}, \theta)^2(\partial_x\varphi)(X_s, \theta)\sigma^2(X_s)ds\right)^m\right] \\ &\leq C_{2m}\varepsilon^{2m} C. \end{aligned}$$

By lemma 6 in Kessler (1997), we have

$$\begin{aligned} E_{\theta_0}\left[\left|\sum_{k=1}^n A_{k,1}(\theta)\right|^{2m}\right] &\leq n^{2m-1}\sum_{k=1}^n E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k} |f(X_{t_{k-1}}, \theta) L_{\theta_0}\varphi(X_{t_k}, \theta)|ds\right)^{2m}\right] \\ &\leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} E_{\theta_0}\left[|f(X_{t_{k-1}}, \theta)|^{2m} E_{\theta_0}[|L_{\theta_0}\varphi(X_{t_k}, \theta)|^{2m}|\mathcal{G}_{k-1}^n]\right] ds \\ &\leq n \cdot \frac{1}{n} \cdot C, \end{aligned}$$

and

$$\begin{aligned} E_{\theta_0}\left[\left|\sum_{k=1}^n A_{k,3}(\theta)\right|^{2m}\right] &\leq \frac{1}{n}\sum_{k=1}^n E_{\theta_0}\left[|f(X_{t_{k-1}}, \theta) b(X_{t_{k-1}}, \theta_0)\partial_x\varphi(X_{t_{k-1}}, \theta)|^{2m}\right] \\ &\leq C. \end{aligned}$$

From these inequalities, we obtain the inequality (11) and similarly it is possible to prove the inequality (12). This completes the proof of (ii). By a similar way, (iii) can be shown. ∎

Proof of Lemma 1. By an easy computation, one has

$$
\begin{aligned}
K^{(ij)}_{\varepsilon,n}(\theta) &= \sum_{k=1}^{n}(\delta_j\delta_i b)(X_{t_{k-1}},\theta)\sigma^{-2}(X_{t_{k-1}})\left[\varphi(X_{t_k},\theta)-\varphi(X_{t_{k-1}},\theta)\right] \\
&\quad+\sum_{k=1}^{n}(\delta_j\delta_i b)(X_{t_{k-1}},\theta)\sigma^{-2}(X_{t_{k-1}})(1-e^{-\lambda(\theta)/n})\varphi(X_{t_{k-1}},\theta) \\
&\quad+\sum_{k=1}^{n}(\delta_i b)(X_{t_{k-1}},\theta)\sigma^{-2}(X_{t_{k-1}})\left[\delta_j\varphi(X_{t_k},\theta)-\delta_j\varphi(X_{t_{k-1}},\theta)\right] \\
&\quad+\sum_{k=1}^{n}(\delta_i b)(X_{t_{k-1}},\theta)\sigma^{-2}(X_{t_{k-1}})(1-e^{-\lambda(\theta)/n})\delta_j\varphi(X_{t_{k-1}},\theta) \\
&\quad+\sum_{k=1}^{n}(\delta_i b)(X_{t_{k-1}},\theta)\sigma^{-2}(X_{t_{k-1}})\frac{(\delta_j\lambda)(\theta)}{n}e^{-\lambda(\theta)/n}\varphi(X_{t_{k-1}},\theta).
\end{aligned}
$$

It follows from lemma 3 that we complete the proof. ∎

Proof of Lemma 2. By lemma 6 in Kessler (1997),

$$
E_{\theta_0}[\varphi(X_{t_k},\theta_0)|X_{t_{k-1}}]-e^{-\lambda(\theta_0)/n}\varphi(X_{t_{k-1}},\theta_0) \;=\; R\left(\theta_0,\frac{\varepsilon^2}{n},X_{t_{k-1}}\right). \tag{13}
$$

Lemma 3–(i) implies that as $\varepsilon\to 0$ and $n\to\infty$,

$$
\varepsilon^{-1}M^{(i)}_{\varepsilon,n}(\theta_0)-\varepsilon^{-1}G^{(i)}_{\varepsilon,n}(\theta_0)=\sum_{k=1}^{n}R\left(\theta_0,\frac{\varepsilon}{n},X_{t_{k-1}}\right)\xrightarrow{p}0.
$$

As $\varepsilon^{-1}G_{\varepsilon,n}(\theta_0)$ is asymptotic equivalent to $\varepsilon^{-1}M_{\varepsilon,n}(\theta_0)$ in the sense as above, it suffices to show asymptotic normality of $\varepsilon^{-1}M_{\varepsilon,n}(\theta_0)$. The predictable quadratic variation of the martingale $\varepsilon^{-1}M_{\varepsilon,n}(\theta_0)$ is as follows.

$$
\varepsilon^{-2}<M^{(i)}(\theta_0),M^{(j)}(\theta_0)>_n \;=\; \varepsilon^{-2}\sum_{k=1}^{n}(\delta_i b)(X_{t_{k-1}},\theta_0)\sigma^{-4}(X_{t_{k-1}})v(X_{t_{k-1}},\theta_0)(\delta_j b)(X_{t_{k-1}},\theta_0),
$$

where $v(X_{t_{k-1}},\theta_0)=E_{\theta_0}[(\varphi(X_{t_k},\theta_0)-E_{\theta_0}[\varphi(X_{t_k},\theta_0)|X_{t_{k-1}}])^2|X_{t_{k-1}}]$.

In order to estimate $v(X_{t_{k-1}},\theta_0)$, we will use the results based on the infinitesimal generator L_{θ_0} as follows.

$$
\begin{aligned}
L_{\theta_0}(U(x,\theta_0))^2 &= 2(-\lambda(\theta_0))\varphi(x,\theta_0)U(x,\theta_0)+\varepsilon^2\sigma^2(x)((\partial_x^2\varphi)(x,\theta_0)U(x,\theta_0)+(\partial_x\varphi)^2(x,\theta_0)), \\
L_{\theta_0}(\varphi(x,\theta_0))^2 &= 2(-\lambda(\theta_0))(\varphi(x,\theta_0))^2+\varepsilon^2\sigma^2(x)((\partial_x^2\varphi)(x,\theta_0)\varphi(x,\theta_0)+(\partial_x\varphi)^2(x,\theta_0)),
\end{aligned}
$$

and that for $l\geq 2$,

$$
L^l_{\theta_0}(U(x,\theta_0))^2 \;=\; 2(-\lambda(\theta_0))^l\varphi(x,\theta_0)U(x,\theta_0)+(2^l-2)(-\lambda(\theta_0))^l(\varphi(x,\theta_0))^2+R(\theta_0,\varepsilon^2,x).
$$

Note that

$$
v(X_{t_{k-1}},\theta_0) \;=\; E_{\theta_0}[(\varphi(X_{t_k},\theta_0)-\varphi(X_{t_{k-1}},\theta_0))^2|X_{t_{k-1}}]-(\varphi(X_{t_{k-1}},\theta_0)-E_{\theta_0}[\varphi(X_{t_k},\theta_0)|X_{t_{k-1}}])^2
$$

$$
\begin{aligned}
&= E_{\theta_0}[(U(X_{t_k},\theta_0))^2|X_{t_{k-1}}] - \left((1-e^{-\lambda(\theta_0)/n})\varphi(X_{t_{k-1}},\theta_0) + R\left(\theta_0, \frac{\varepsilon^2}{n}, X_{t_{k-1}}\right)\right)^2 \\
&= E_{\theta_0}[(U(X_{t_k},\theta_0))^2|X_{t_{k-1}}] - ((1-e^{-\lambda(\theta_0)/n})\varphi(X_{t_{k-1}},\theta_0))^2 + R\left(\theta_0, \frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right).
\end{aligned}
$$

Using lemma 1 in Florens-Zmirou (1989), lemma 6 in Kessler (1997) and the results on L_{θ_0} as above, we see that for any $m \geq 1$,

$$
\begin{aligned}
v(X_{t_{k-1}},\theta_0) &= \sum_{l=1}^{2m} \frac{1}{l!n^l} L_{\theta_0}^l (U(X_{t_{k-1}},\theta_0))^2 - (1-2e^{-\lambda(\theta_0)/n} + e^{-2\lambda(\theta_0)/n})(\varphi(X_{t_{k-1}},\theta_0))^2 \\
&\quad + R\left(\theta_0, \frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right) \\
&= \frac{\varepsilon^2}{n}\sigma^2(X_{t_{k-1}})(\partial_x\varphi)^2(X_{t_{k-1}},\theta_0) \\
&\quad + \sum_{l=2}^{2m} \frac{1}{l!n^l} L_{\theta_0}^l (U(X_{t_{k-1}},\theta_0))^2 - \sum_{l=2}^{2m} \frac{(2^l-2)(-\lambda(\theta_0))^l}{l!n^l}(\varphi(X_{t_{k-1}},\theta_0))^2 \\
&\quad + R\left(\theta_0, \frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right) \\
&= \frac{\varepsilon^2}{n}\sigma^2(X_{t_{k-1}})(\partial_x\varphi)^2(X_{t_{k-1}},\theta_0) + R\left(\theta_0, \frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\theta_0, \frac{1}{n^{2m+1}}, X_{t_{k-1}}\right).
\end{aligned}
$$

It then follows that $\varepsilon^{-2} < M^{(i)}, M^{(j)} >_n \xrightarrow{p} A^{(ij)}(\theta_0)$ as $\varepsilon \to 0$ and $n \to \infty$. The central limit theorem for martingales implies that $\varepsilon^{-1}G_{\varepsilon,n}(\theta_0) \xrightarrow{d} N(0, A(\theta_0))$ as $\varepsilon \to 0$ and $n \to \infty$. This completes the proof. ∎

Proof of Theorem 1. In the same way as the proof of theorem 6.1 in Sakamoto and Yoshida (2004), we prove the existence, the uniqueness and the consistency of $\hat{\theta}_{\varepsilon,n}$. By A5, there exist a constant $C > 0$ and an open set $\tilde{\Theta}$ including θ_0 such that

$$
\inf_{\theta_1,\theta_2\in\tilde{\Theta},|x|=1} \left| \left(\int_0^1 K(\theta_1 + s(\theta_2-\theta_1))ds\right)^* x \right| > 2C.
$$

For such a $C > 0$, let $\mathcal{X}_{\varepsilon,n,0}$ denote the subset of the sample space $\mathcal{X}_{\varepsilon,n}$ defined by

$$
\begin{aligned}
\mathcal{X}_{\varepsilon,n,0} &= \left\{ X \in \mathcal{X}_{\varepsilon,n} \,\middle|\, \sup_{\theta\in\tilde{\Theta}} |K_{\varepsilon,n}(\theta) - K(\theta)| < \frac{C}{2}, \quad |\varepsilon^{-\gamma}G_{\varepsilon,n}(\theta_0)| < C, \right. \\
&\qquad \left. \inf_{\theta_1,\theta_2\in\tilde{\Theta},|x|=1} \left| -\left(\int_0^1 K_{\varepsilon,n}(\theta_1 + s(\theta_2-\theta_1))ds\right)^* x \right| > C \right\}.
\end{aligned}
$$

By lemma 2, we see that as $\varepsilon \to 0$ and $n \to \infty$,

$$
P_{\theta_0}\left[|\varepsilon^{-\gamma}G_{\varepsilon,n}(\theta_0)| \geq C\right] \to 0. \tag{14}
$$

An easy estimates implies that

$$
P_{\theta_0}\left[\inf_{\theta_1,\theta_2\in\tilde{\Theta},|x|=1} \left| -\left(\int_0^1 K_{\varepsilon,n}(\theta_1 + s(\theta_2-\theta_1))ds\right)^* x \right| \leq C\right]
$$

$$\leq P_{\theta_0}\left[\inf_{\theta_1,\theta_2\in\tilde\Theta,|x|=1}\left\{-\left|\left(\int_0^1(K_{\varepsilon,n}(\theta_1+s(\theta_2-\theta_1))-K(\theta_1+s(\theta_2-\theta_1))ds\right)^* x\right|\right.\right.$$
$$\left.\left.+\left|\left(\int_0^1 K(\theta_1+s(\theta_2-\theta_1))ds\right)^* x\right|\right\}\leq C\right]$$
$$\leq P_{\theta_0}\left[\sup_{\theta\in\tilde\Theta}|K_{\varepsilon,n}(\theta)-K(\theta)|\geq C\right].$$

Lemma 1 yields that as $\varepsilon\to 0$ and $n\to\infty$,

$$P_{\theta_0}\left[\inf_{\theta_1,\theta_2\in\tilde\Theta,|x|=1}\left|-\left(\int_0^1 K_{\varepsilon,n}(\theta_1+s(\theta_2-\theta_1))ds\right)^* x\right|\leq C\right]\to 0. \qquad (15)$$

For $C>0$, there exists $N_0(C)\in(0,1]$ such that for any $\varepsilon\in(0,N_0(C))$ and $|\delta|\leq 1$, $\{\theta : |\theta-\theta_0|\leq\varepsilon^\gamma\}\subset\tilde\Theta$ and $|K(\theta_0+\delta\varepsilon^\gamma)-K(\theta_0)|<C/2$. For $X\in\mathcal{X}_{\varepsilon,n,0}$, let $\hat I(u)$ be a function $\hat I:\{u\in\mathbf{R}^p:|u|\leq 1\}\to\mathbf{R}^p\otimes\mathbf{R}^p$ defined by

$$\hat I(u)=-\int_0^1 K_{\varepsilon,n}(\theta_0+\varepsilon^\gamma u\xi)d\xi.$$

It then follows that for $x\in\mathbf{R}^p$ satisfying $|x|=1$,

$$|(\hat I(u)+K(\theta_0))^* x| \leq \int_0^1|K_{\varepsilon,n}(\theta_0+\varepsilon^\gamma u\xi)-K(\theta_0+\varepsilon^\gamma u\xi)|d\xi+\int_0^1|K(\theta_0+\varepsilon^\gamma u\xi)-K(\theta_0)|d\xi$$
$$\leq \sup_{\theta\in\tilde\Theta}|K_{\varepsilon,n}(\theta)-K(\theta)|+\frac{C}{2}<C.$$

As we see that

$$\inf_{|x|=1}|(\hat I(u))^* x|\geq\inf_{|x|=1}\Big(|-(K(\theta_0))^* x|-|(\hat I(u))^* x+(K(\theta_0))^* x|\Big)>C,$$

$\hat I(u)$ is invertible on $\mathcal{X}_{\varepsilon,n,0}$. For $X\in\mathcal{X}_{\varepsilon,n,0}$, let H be a function on $\{u\in\mathbf{R}^p:|u|\leq 1\}$ defined by

$$H(u)=\varepsilon^{-\gamma}\check I(u)G_{\varepsilon,n}(\theta_0),$$

where $\check I(u)=\hat I^{-1}(u)$. Let $\tilde\rho(A)$ and $\rho(A)$ are the minimum and the maximum eigenvalues of a matrix A, respectively. It then follows that for $X\in\mathcal{X}_{\varepsilon,n,0}$,

$$|H(u)| \leq C\sup_{|x|\leq 1}|(\check I(u))^* x|=C\sqrt{\rho(\check I^*(u)\check I(u))}=\frac{C}{\sqrt{\tilde\rho(\hat I(u)\hat I^*(u))}}=\frac{C}{\inf_{|x|=1}|(\hat I(u))^* x|}\leq 1.$$

Thus, Brouwer's fixed point theorem implies that for $X\in\mathcal{X}_{\varepsilon,n,0}$, there exists a $\hat u\in\{u:|u|\leq 1\}$ such that $H(\hat u)=\hat u$. Moreover, setting $\hat\theta_{\varepsilon,n}=\theta_0+\varepsilon^\gamma\hat u$ and using Taylor's theorem, one has

$$G_{\varepsilon,n}(\hat\theta_{\varepsilon,n}) = G_{\varepsilon,n}(\theta_0)+\int_0^1 K_{\varepsilon,n}(\theta_0+\varepsilon^\gamma\hat u\xi)d\xi\varepsilon^\gamma\hat u=\varepsilon^\gamma\hat I(\hat u)(H(\hat u)-\hat u)=0.$$

As we see that $\int_0^1 K_{\varepsilon,n}(\theta_1+s(\theta_2-\theta_1))ds$ is non-singular uniformly in $\theta_1,\theta_2\in\tilde\Theta$ on $\mathcal{X}_{\varepsilon,n,0}$, for $X\in\mathcal{X}_{\varepsilon,n,0}$, there exists a unique $\hat\theta_{\varepsilon,n}\in\tilde\Theta$ such that $G_{\varepsilon,n}(\hat\theta_{\varepsilon,n})=0$ and $\hat\theta_{\varepsilon,n}$ lies in the ε^γ-neighborhood of θ_0. Moreover, lemma 1, (14) and (15) yields that $P_{\theta_0}[\mathcal{X}^c_{\varepsilon,n,0}]\to 0$ as $\varepsilon\to 0$ and $n\to\infty$. Therefore, we have that as $\varepsilon\to 0$ and $n\to\infty$,

$$P_{\theta_0}[(\exists_1\hat\theta_{\varepsilon,n}\in\tilde\Theta \text{ such that } G_{\varepsilon,n}(\hat\theta_{\varepsilon,n})=0) \text{ and } (|\hat\theta_{\varepsilon,n}-\theta_0|\leq\varepsilon^\gamma)]\geq P_{\theta_0}[\mathcal{X}_{\varepsilon,n,0}]\to 1.$$

This completes the proof of the existence, the uniqueness and the consistency.

Next, we prove the asymptotic normality of $\hat{\theta}_{\varepsilon,n}$. As θ_0 is in Θ, $B(\theta_0;\rho) \subset \Theta$ for sufficiently small $\rho > 0$, where $B(\theta_0;\rho) = \{\theta : |\theta - \theta_0| \le \rho\}$. Taylor's formula yields that if $\hat{\theta}_{\varepsilon,n} \in B(\theta_0;\rho)$,

$$D_{\varepsilon,n}S_{\varepsilon,n} = \varepsilon^{-1}G_{\varepsilon,n}(\hat{\theta}_{\varepsilon,n}) - \varepsilon^{-1}G_{\varepsilon,n}(\theta_0),$$

where $D_{\varepsilon,n} = \int_0^1 K_{\varepsilon,n}\left(\theta_0 + u(\hat{\theta}_{\varepsilon,n} - \theta_0)\right)du$ and $S_{\varepsilon,n} = \varepsilon^{-1}(\hat{\theta}_{\varepsilon,n} - \theta_0)$. By A5, there exists a constant $C > 0$ such that $\inf_{|x|=1}|(K(\theta_0))^* x| > 2C$. For such a $C > 0$, there exist $N_1(C) > 0$ and $N_2(C) > 0$ such that for any $(\varepsilon, n) \in (0, N_1(C)) \times (N_2(C), \infty)$ and $\delta \in [0,1]$, $B(\theta_0;\eta_{\varepsilon,n}) \subset \Theta$ and $|K(\theta_0 + \delta\eta_{\varepsilon,n}) - K(\theta_0)| < C/2$, where $\eta_{\varepsilon,n} \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. We set $\mathcal{N}(C) = (0, N_1(C)) \times (N_2(C), \infty)$. For $C > 0$, let $\mathcal{C}_{\varepsilon,n}$ denote the subset of the sample space $\mathcal{X}_{\varepsilon,n}$ defined by

$$\mathcal{C}_{\varepsilon,n} = \left\{ X \in \mathcal{X}_{\varepsilon,n} \;\middle|\; \hat{\theta}_{\varepsilon,n} \in B(\theta_0;\eta_{\varepsilon,n}), \quad \sup_{\theta\in\bar{\Theta}} |K_{\varepsilon,n}(\theta) - K(\theta)| < \frac{C}{2} \right\}.$$

By the consistency of $\hat{\theta}_{\varepsilon,n}$, there exists a sequence $\{B(\theta_0;\eta_{\varepsilon,n})\}$ such that $\eta_{\varepsilon,n} \to 0$ and $P_{\theta_0}[\hat{\theta}_{\varepsilon,n} \in B(\theta_0;\eta_{\varepsilon,n})] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$. It then follows that $P_{\theta_0}[\mathcal{C}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$. For any $(\varepsilon, n) \in \mathcal{N}(C)$ and $|u| < 1$, we see that on $\mathcal{C}_{\varepsilon,n}$,

$$\begin{aligned}
& \sup_{|x|=1} |(-D_{\varepsilon,n} + K(\theta_0))^* x| \\
\le\ & \sup_{|x|=1} \left\{ |(-D_{\varepsilon,n} + \int_0^1 K(\theta_0 + u(\hat{\theta}_{\varepsilon,n} - \theta_0))du)^* x| + |(K(\theta_0) - \int_0^1 K(\theta_0 + u(\hat{\theta}_{\varepsilon,n} - \theta_0))du)^* x| \right\} \\
\le\ & \sup_{|\theta-\theta_0|<\eta_{\varepsilon,n}} |\varepsilon^2 K_{\varepsilon,n}(\theta) - K(\theta)| + \frac{C}{2} < C.
\end{aligned}$$

It then follows that for any $(\varepsilon, n) \in \mathcal{N}(C)$, on $\mathcal{C}_{\varepsilon,n}$,

$$\inf_{|x|=1} |(D_{\varepsilon,n})^* x| \ \ge\ \inf_{|x|=1} |(K(\theta_0))^* x| - \sup_{|x|=1} |(-D_{\varepsilon,n} + K(\theta_0))^* x| > C.$$

By the above estimate, we see that for any $(\varepsilon, n) \in \mathcal{N}(C)$, $P_{\theta_0}[\mathcal{D}_{\varepsilon,n}] \ge P_{\theta_0}[\mathcal{C}_{\varepsilon,n}]$, where $\mathcal{D}_{\varepsilon,n} = \{D_{\varepsilon,n} \text{ is invertible}\}$. We obtain that $P_{\theta_0}[\mathcal{D}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$. We define $E_{\varepsilon,n}$ as follows: $E_{\varepsilon,n} = D_{\varepsilon,n}$ on $\mathcal{E}_{\varepsilon,n}$ and $E_{\varepsilon,n} = J_p$ on $\mathcal{E}^c_{\varepsilon,n}$, where $\mathcal{E}_{\varepsilon,n} = \{\hat{\theta}_{\varepsilon,n} \in \Theta\} \cap \mathcal{D}_{\varepsilon,n}$ and J_p is the $p \times p$ identity matrix. As it follows that $P_{\theta_0}[\mathcal{E}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$, we see that $|E_{\varepsilon,n} - K(\theta_0)|1_{\mathcal{E}_{\varepsilon,n}} \le |D_{\varepsilon,n} - K(\theta_0)|$ and $1_{\mathcal{E}_{\varepsilon,n}} \xrightarrow{p} 1$ as $\varepsilon \to 0$ and $n \to \infty$. By using the estimate that $P_{\theta_0}[\hat{\theta}_{\varepsilon,n} \in \Theta^c \cup B(\theta_0;\eta_{\varepsilon,n})^c] \le P_{\theta_0}[\hat{\theta}_{\varepsilon,n} \in \Theta^c] + P_{\theta_0}[\hat{\theta}_{\varepsilon,n} \in B(\theta_0;\eta_{\varepsilon,n})^c] \to 0$ as $\varepsilon \to 0$ and $n \to \infty$, it follows that $1_{\{\hat{\theta}_{\varepsilon,n} \in \Theta^c \cup B(\theta_0;\eta_{\varepsilon,n})^c\}} \xrightarrow{p} 0$ as $\varepsilon \to 0$ and $n \to \infty$. Define $R_{\varepsilon,n} = D_{\varepsilon,n} - K_{\varepsilon,n}(\theta_0)$. Lemma 1 yields that $|R_{\varepsilon,n}| \cdot 1_{\{\hat{\theta}_{\varepsilon,n} \in \Theta \cap B(\theta_0;\eta_{\varepsilon,n})\}} \le \sup_{\theta \in B(\theta_0;\eta_{\varepsilon,n})} |K_{\varepsilon,n}(\theta) - K_{\varepsilon,n}(\theta_0)| \xrightarrow{p} 0$ as $\varepsilon \to 0$ and $n \to \infty$. Thus, $R_{\varepsilon,n} \xrightarrow{p} 0$ as $\varepsilon \to 0$ and $n \to \infty$. By using lemma 1, $D_{\varepsilon,n} \xrightarrow{p} -K(\theta_0)$ as $\varepsilon \to 0$ and $n \to \infty$. Therefore, $E_{\varepsilon,n} \xrightarrow{p} K(\theta_0)$ as $\varepsilon \to 0$ and $n \to \infty$. It is easy to obtain that $S_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}} = E^{-1}_{\varepsilon,n}D_{\varepsilon,n}S_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}} = E^{-1}_{\varepsilon,n}(\varepsilon^{-1}G_{\varepsilon,n}(\hat{\theta}_{\varepsilon,n}) - \varepsilon^{-1}G_{\varepsilon,n}(\theta_0))1_{\mathcal{E}_{\varepsilon,n}}$. Lemma 2 implies that $S_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}} \xrightarrow{d} N(0, K(\theta_0)^{-1}A(\theta_0)(K^*)^{-1}(\theta_0))$ as $\varepsilon \to 0$ and $n \to \infty$. This completes the proof. ∎

Acknowledgements

This work is partially supported by the Research Fund for Scientists of the Ministry of Education, Culture, Sports, Science and Technology, and by Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality, of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Aït-Sahalia, Y. and Mykland, P. A. (2003). The effects of random and discrete sampling when estimating continuous-time diffusions. *Econometrica* **71**, 483-549.

Aït-Sahalia, Y. and Mykland, P. A. (2004). Estimators of diffusions with discrete observations: a general theory. *Ann. Statist.* **32**, 2186-2222.

Azencott, R. (1982). Formule de Taylor stochastique et développement asymptotique d'intégrales de Feynmann. *Séminaire de Probabilités XVI; Supplément: Géométrie Différentielle Stochastique. Lecture Notes In Math.* **921**, 237-285 Springer Verlag, Berlin.

Bibby, B. M. and Sørensen, M. (1995). Martingale estimating functions for discretely observed diffusion processes. *Bernoulli* **1**, 17–39.

Bibby, B. M. and Sørensen, M. (1996). On estimation for discretely observed diffusions: a review. *Theory Stoch. Process.* **2**, 49–56.

Bibby, B. M., Jacobsen, M. and Sørensen, M. (2002). Estimating functions for discretely sampled diffusion-type models. preprint.

Florens-Zmirou, D. (1989). Approximate discrete time schemes for statistics of diffusion processes. *Statistics* **20**, 547–557.

Freidlin, M. I. and Wentzell, A. D. (1984). *Random perturbations of dynamical systems.* Springer-Verlag, New York.

Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations. *Statistics* **21**, 99–116.

Genon-Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. Henri Poincaré Probab. Statist.* **29**, 119–151.

Ibragimov, I. A. and Has'minskii, R. Z. (1981). *Statistical estimation.* Springer Verlag, New York.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24**, 211–229.

Kessler, M. and Sørensen, M. (1999). Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli* **5**, 299–314.

Kloeden, P. E. and Platen, E. (1992). *Numerical solution of stochastic differential equations.* Springer-Verlag, New York.

Kunitomo, N. and Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance* **11**, 117-151.

Kutoyants, Yu. A. (1984). Parameter estimation for stochastic processes. Prakasa Rao, B.L.S. (ed.) Heldermann, Berlin.

Kutoyants, Yu. A. (1994). *Identification of dynamical systems with small noise.* Kluwer, Dordrecht.

Laredo, C. F. (1990). A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. *Ann. Statist.* **18**, 1158–1171.

Matsuzaki, R. and Uchida, M. (2005). One-step estimator for diffusion processes with small dispersion parameters from discrete observations. preprint.

Øksendal, B. (2003). *Stochastic differential equations. An introduction with applications.* Sixth edition. Universitext. Springer-Verlag, Berlin.

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

Sørensen, M. (1997). Estimating functions for discretely observed diffusions: a review. In Basawa, I. V., Godambe, V. P. and Taylor, R. L., editors, Selected Proceedings of the Symposium on Estimating Functions, Hayward: Institute of Mathematical Statistics. IMS Lecture Notes-Monograph Series, **32**, 305–325.

Sørensen, M. (2000). Small dispersion asymptotics for diffusion martingale estimating functions. Preprint No. 2000-2, Department of Statistics and Operations Research, University of Copenhagen. http://www.stat.ku.dk/research/preprint/

Sørensen, M. and Uchida, M. (2003). Small diffusion asymptotics for discretely sampled stochastic differential equations. *Bernoulli* **9**, 1051–1069.

Takahashi, A. and Yoshida, N. (2004). An asymptotic expansion scheme for optimal investment problems. *Stat. Inference Stoch. Process.* **7**, 153-188.

Uchida, M. (2004a). Estimation for discretely observed small diffusions based on approximate martingale estimating functions. *Scand. J. Statist.* **31**, 553-566.

Uchida, M. (2004b). Martingale estimating functions based on eigenfunctions for discretely observed small diffusions. (submitted.)

Uchida, M. and Yoshida, N. (2004a). Information criteria for small diffusions via the theory of Malliavin-Watanabe. *Statist. Infer. Stochast. Process.* **7**, 35-67.

Uchida, M. and Yoshida, N. (2004b). Asymptotic expansion for small diffusions applied to option pricing. *Statist. Infer. Stochast. Process.* **7**, 189-223.

Yoshida, N. (1992a). Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe. *Probab. Theory Relat. Fields* **92**, 275-311.

Yoshida, N. (1992b). Asymptotic expansion for statistics related to small diffusions. *J. Japan Statist. Soc.* **22**, 139-159.

Yoshida, N. (2003). Conditional expansions and their applications. *Stochastic Process. Appl.* **107**, 53-81.

List of MHF Preprint Series, Kyushu University

21st Century COE Program
Development of Dynamic Mathematics with High Functionality

MHF2003-1 Mitsuhiro T. NAKAO, Kouji HASHIMOTO & Yoshitaka WATANABE
A numerical method to verify the invertibility of linear elliptic operators with applications to nonlinear problems

MHF2003-2 Masahisa TABATA & Daisuke TAGAMI
Error estimates of finite element methods for nonstationary thermal convection problems with temperature-dependent coefficients

MHF2003-3 Tomohiro ANDO, Sadanori KONISHI & Seiya IMOTO
Adaptive learning machines for nonlinear classification and Bayesian information criteria

MHF2003-4 Kazuhiro YOKOYAMA
On systems of algebraic equations with parametric exponents

MHF2003-5 Masao ISHIKAWA & Masato WAKAYAMA
Applications of Minor Summation Formulas III, Plücker relations, Lattice paths and Pfaffian identities

MHF2003-6 Atsushi SUZUKI & Masahisa TABATA
Finite element matrices in congruent subdomains and their effective use for large-scale computations

MHF2003-7 Setsuo TANIGUCHI
Stochastic oscillatory integrals - asymptotic and exact expressions for quadratic phase functions -

MHF2003-8 Shoki MIYAMOTO & Atsushi YOSHIKAWA
Computable sequences in the Sobolev spaces

MHF2003-9 Toru FUJII & Takashi YANAGAWA
Wavelet based estimate for non-linear and non-stationary auto-regressive model

MHF2003-10 Atsushi YOSHIKAWA
Maple and wave-front tracking — an experiment

MHF2003-11 Masanobu KANEKO
On the local factor of the zeta function of quadratic orders

MHF2003-12 Hidefumi KAWASAKI
Conjugate-set game for a nonlinear programming problem

MHF2004-1 Koji YONEMOTO & Takashi YANAGAWA
Estimating the Lyapunov exponent from chaotic time series with dynamic noise

MHF2004-2 Rui YAMAGUCHI, Eiko TSUCHIYA & Tomoyuki HIGUCHI
State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors

MHF2004-3 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Cubic pencils and Painlevé Hamiltonians

MHF2004-4 Atsushi KAWAGUCHI, Koji YONEMOTO & Takashi YANAGAWA
Estimating the correlation dimension from a chaotic system with dynamic noise

MHF2004-5 Atsushi KAWAGUCHI, Kentarou KITAMURA, Koji YONEMOTO, Takashi YANAGAWA & Kiyofumi YUMOTO
Detection of auroral breakups using the correlation dimension

MHF2004-6 Ryo IKOTA, Masayasu MIMURA & Tatsuyuki NAKAKI
A methodology for numerical simulations to a singular limit

MHF2004-7 Ryo IKOTA & Eiji YANAGIDA
Stability of stationary interfaces of binary-tree type

MHF2004-8 Yuko ARAKI, Sadanori KONISHI & Seiya IMOTO
Functional discriminant analysis for gene expression data via radial basis expansion

MHF2004-9 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Hypergeometric solutions to the q - Painlevé equations

MHF2004-10 Raimundas VIDŪNAS
Expressions for values of the gamma function

MHF2004-11 Raimundas VIDŪNAS
Transformations of Gauss hypergeometric functions

MHF2004-12 Koji NAKAGAWA & Masakazu SUZUKI
Mathematical knowledge browser

MHF2004-13 Ken-ichi MARUNO, Wen-Xiu MA & Masayuki OIKAWA
Generalized Casorati determinant and Positon-Negaton-Type solutions of the Toda lattice equation

MHF2004-14 Nalini JOSHI, Kenji KAJIWARA & Marta MAZZOCCO
Generating function associated with the determinant formula for the solutions of the Painlevé II equation

MHF2004-15 Kouji HASHIMOTO, Ryohei ABE, Mitsuhiro T. NAKAO & Yoshitaka WATANABE
Numerical verification methods of solutions for nonlinear singularly perturbed problem

MHF2004-16 Ken-ichi MARUNO & Gino BIONDINI
Resonance and web structure in discrete soliton systems: the two-dimensional Toda lattice and its fully discrete and ultra-discrete versions

MHF2004-17 Ryuei NISHII & Shinto EGUCHI
Supervised image classification in Markov random field models with Jeffreys divergence

MHF2004-18 Kouji HASHIMOTO, Kenta KOBAYASHI & Mitsuhiro T. NAKAO
Numerical verification methods of solutions for the free boundary problem

MHF2004-19 Hiroki MASUDA
Ergodicity and exponential β-mixing bounds for a strong solution of Lévy-driven stochastic differential equations

MHF2004-20 Setsuo TANIGUCHI
The Brownian sheet and the reflectionless potentials

MHF2004-21 Ryuei NISHII & Shinto EGUCHI
Supervised image classification based on AdaBoost with contextual weak classifiers

MHF2004-22 Hideki KOSAKI
On intersections of domains of unbounded positive operators

MHF2004-23 Masahisa TABATA & Shoichi FUJIMA
Robustness of a characteristic finite element scheme of second order in time increment

MHF2004-24 Ken-ichi MARUNO, Adrian ANKIEWICZ & Nail AKHMEDIEV
Dissipative solitons of the discrete complex cubic-quintic Ginzburg-Landau equation

MHF2004-25 Raimundas VIDŪNAS
Degenerate Gauss hypergeometric functions

MHF2004-26 Ryo IKOTA
The boundedness of propagation speeds of disturbances for reaction-diffusion systems

MHF2004-27 Ryusuke KON
Convex dominates concave: an exclusion principle in discrete-time Kolmogorov systems

MHF2004-28 Ryusuke KON
Multiple attractors in host-parasitoid interactions: coexistence and extinction

MHF2004-29 Kentaro IHARA, Masanobu KANEKO & Don ZAGIER
Derivation and double shuffle relations for multiple zeta values

MHF2004-30 Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Generalized partitioned quantum cellular automata and quantization of classical CA

MHF2005-1 Hideki KOSAKI
Matrix trace inequalities related to uncertainty principle

MHF2005-2 Masahisa TABATA
Discrepancy between theory and real computation on the stability of some finite element schemes

MHF2005-3 Yuko ARAKI & Sadanori KONISHI
Functional regression modeling via regularized basis expansions and model selection

MHF2005-4 Yuko ARAKI & Sadanori KONISHI
Functional discriminant analysis via regularized basis expansions

MHF2005-5 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Point configurations, Cremona transformations and the elliptic difference Painlevé equations

MHF2005-6 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Construction of hypergeometric solutions to the q - Painlevé equations

MHF2005-7 Hiroki MASUDA
Simple estimators for non-linear Markovian trend from sampled data:
I. ergodic cases

MHF2005-8 Hiroki MASUDA & Nakahiro YOSHIDA
Edgeworth expansion for a class of Ornstein-Uhlenbeck-based models

MHF2005-9 Masayuki UCHIDA
Approximate martingale estimating functions under small perturbations of dynamical systems