On Nonparametric Estimation for SDE with Delay

Yury Kutoyants

Pub. Inst. Stat. Univ. Paris
63, fasc. 2-3, 2019, 11-20
Numéro spécial en l'honneur des 80 ans de Denis Bosq /
Special issue in honour of Denis Bosq's 80th birthday

On Nonparametric Estimation for SDE with Delay

Yury A. Kutoyants,

Le Mans University and Tomsk State University

Abstract: We consider the problem of drift function estimation of inhomogeneous stochastic differential equation with delay. It is shown that kernel-type estimator is consistent and asymptotically efficient.

1. Introduction

This work is devoted to the problems of nonparametric estimation for the non linear stochastic differential equation (SDE) with delay

$$\mathrm{d}X_t = S\left(t, X_{t-\tau}\right)\mathrm{d}t + \varepsilon \mathrm{d}W_t, \quad X_s = x_0,\ s \le 0, \quad 0 \le t \le T, \tag{1.1}$$

where $\varepsilon \to 0$. Such perturbed dynamical systems were studied in [7]. Statistical problems (parametric and nonparametric) related to such models can be found in [18]. We suppose that the function $S\left(t, x\right)$ is unknown. Recall that the limit ($\varepsilon \to 0$) equation is

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = S\left(t, x_{t-\tau}\right), \quad x_s = x_0,\ s \le 0, \quad 0 \le t \le T \tag{1.2}$$

and the trajectories X_t converge uniformly in $t \in [0, T]$ to x_t.

The problem of drift estimation for diffusion processes in *small noise* ($\varepsilon \to 0$) and large samples ($T \to \infty$) are well studied and there are many results in different statements for continuous and discrete time observations.

In this work we consider the problem of drift estimation by the observations $X^T = (X_t, 0 \le t \le T)$. We suppose that the delay $\tau \ge 0$ is known and we estimate the function $S\left(t, x\right)$, where for a given x we take $t = t_x$ defined by the equality $x_{t_x - \tau} = x$. Therefore we estimate a function $S\left(t_x, x\right)$ of one variable x.

We propose kernel-type estimator and study its asymptotic ($\varepsilon \to 0$) behavior. We show its consistency and asymptotic efficiency in order of convergence.

The particularity of this model can be illustrated as follows. If we write (1.1) in integral form and change the variables $t \to t - \tau$, then for $t \ge \tau$ we obtain the representation

$$X_{t-\tau} = x_0 + \int_0^{t-\tau} S\left(s, X_{s-\tau}\right)\mathrm{d}s + \varepsilon W_{t-\tau}$$

AMS 2000 subject classifications: Primary 62M09; secondary 62G20
Keywords and phrases: delayed SDE, nonparametric estimation, drift estimation

Therefore the drift $S(X_{t-\tau})$ in the equation (1.1) has no even the first derivative w.r.t. t.

Statistical problems of parameter estimation related to different generalizations of the model (1.2) with the asymptotic $\varepsilon \to 0$ were treated in the works [1], [11], [12],[13], [17], [18] [22], [23], (see review in [20]). Some statistical problems for the related SDE with delay in the case of asymptotic $T \to \infty$ were studied in the works [8], [9], [14], [20].

Nonparametric estimation problems for perturbed ($\varepsilon \to 0$) SDE ($\varepsilon \to 0$) can be found in [2], [15], [16], [18]. The large samples cases of drift estimation were considered in the works [3], [4],[5], [25], [6], [19].

2. Main Results

For a given $x > x_0$ we define $t = t_x$ such that $x = x_{t_x-\tau}$ and consider the problem of estimation of the value $S(t,x) = S(t_x,x)$.

Introduce the conditions $\mathcal{R}$.

$\mathcal{R}_1$. *The function* $S(t,x)$ *is positive* $\kappa = \inf_{t\in[0,T],x} S(t,x) > 0$.

$\mathcal{R}_2$. *The function* $S(t,x)$ *has two continuous bounded derivatives w.r.t.* t *and* x.

$\mathcal{R}_3$. *The delay* $\tau \in (0,T)$ *is known.*

By condition $\mathcal{R}_2$ there exist two constants $M > 0$ and $L > 0$ such that for all $t \in [0,T]$ and all $x \in \mathcal{R}$.

$$\left|S'_t(t,x)\right| \le M, \qquad \text{and} \qquad \left|S'_x(t,x)\right| \le L.$$

Moreover, for $t \in [0,\tau]$ we have $S(t,x_t) = S(t,x_0)$ and

$$x_t = x_0 + \int_0^t S(s,x_0)\,\mathrm{d}s, \qquad 0 \le t \le \tau.$$

Therefore we have two different problems of estimation. For $t \in [0,\tau]$ we estimate the function $S(t,x_0)$ of one variable $t \in (0,\tau)$ and for $t \in (\tau,T]$ we estimate once more the function $S(t,x) = S(t,x)$ of one variable $x \in (x_0, x_{T-\tau})$ because here $t = t_x$ such that $x_{t_x-\tau} = x > x_0$. Note that estimation of the function $f(t) = S(t,x_0), 0 < t < \tau$ is particular case of the similar problem considered in [10], Theorem 7.4.1. That is why we consider estimation of the function $S(t,x)$ for $t > \tau$.

Define the random variable (stopping time)

$$t_{x,\varepsilon} = \inf\{t > \tau : \ X_{t-\tau} \ge x\}$$

and put $t_{x,\varepsilon} = T - \tau$ if $\sup_{0\le s\le T-\tau} X_s < x$. Introduce the kernel-type estimator

$$S_\varepsilon(t,x) = \frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right)\mathrm{d}X_s. \tag{2.1}$$

Here $\varphi_\varepsilon \to 0$ and the bounded kernel $K(u), u \in \mathcal{R}$ satisfies the "usual properties":

$$\int_{-1}^{1} K(u)\,\mathrm{d}u = 1, \qquad \int_{-1}^{1} uK(u)\,\mathrm{d}u = 0, \qquad K(u) = 0, \text{ for } u \notin [-1,1]. \tag{2.2}$$

We have to define the set $\mathcal{X}$ of x admitting consistent estimation of $S(t,x), x \in \mathcal{X}$. Note that

$$x_{t-\tau} = x_0 + \int_0^{t-\tau} S(s, x_{s-\tau})\,\mathrm{d}s, \qquad \tau < t < T - \tau$$

where $S(t,x) > 0$. Therefore, we can consider $\mathcal{X} = (x_0, x_{T-\tau})$, but the value $x_{T-\tau} = x_{T-\tau}(S)$ depends on the unknown function $S(\cdot,\cdot)$ and can not be known. There are at least two possibilities. The first one is to use the condition $\mathcal{R}_1$, which allows us to write $x_{T-\tau} > x_0 + (T-\tau)\kappa \equiv \hat{x}$ and to put $\mathcal{X} = (x_0, \hat{x})$. The second one is to put $\mathcal{X} = (x_0, X_{T-\tau})$ where $X_{T-\tau}$ is estimator of $x_{T-\tau}$. Then to use the estimate $|X_{T-\tau} - x_{T-\tau}| \le C\varepsilon \sup_{0\le s\le T-\tau} |W_t|$ (see (A.1)) and study the introduced estimator (2.1). Here and below the generic constants are denoted $C.$.

Similar problems and estimators of nonparametric estimation for perturbed dynamical system (without delays) were studied in [15], [17] and [18]. Of course, if $S(t,x) = S(t)$ we obtain classical problem of nonparametric estimation of the deterministic signal in white Gaussian noise (see, e.g., [10]).

Theorem 2.1. *Suppose that conditions $\mathcal{R}$ hold then the estimator $S_\varepsilon(t,x), x \in (x_0, \hat{x})$ with $\varphi_\varepsilon = \varepsilon^{2/5}$ is mean-squared consistent*

$$\mathbf{E}_S |S_\varepsilon(t,x) - S(t,x)|^2 \le C\,\varepsilon^{8/5}. \tag{2.3}$$

Proof. Let us introduce the set

$$\mathbb{B}_\varepsilon = \left\{\omega :\ \varphi_\varepsilon^{-1} t_{x,\varepsilon} > 2,\ \ \varphi_\varepsilon^{-1}(T - t_{x,\varepsilon}) > 2\right\}$$

and for $\omega \in \mathbb{B}_\varepsilon$ change the variables $s = t_{x,\varepsilon} + \varphi_\varepsilon u$ in the integral

$$\begin{aligned}
&\frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) S(s, x_{s-\tau})\,\mathrm{d}s \\
&= \int_{-\frac{t_{x,\varepsilon}}{\varphi_\varepsilon}}^{\frac{T-t_{x,\varepsilon}}{\varphi_\varepsilon}} K(u)\, S\left(t_{x,\varepsilon} + \varphi_\varepsilon u, x_{t_{x,\varepsilon}+\varphi_\varepsilon u-\tau}\right)\mathrm{d}u \\
&= \int_{-1}^{1} K(u)\, S\left(t_{x,\varepsilon} + \varphi_\varepsilon u, x_{t_{x,\varepsilon}+\varphi_\varepsilon u-\tau}\right)\mathrm{d}u \\
&= S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) + \varphi_\varepsilon^2 R_\varepsilon^{(1)}\left(t_{x,\varepsilon}\right) \\
&+ \varphi_\varepsilon \int_{-1}^{1} uK(u)\,\mathrm{d}u \left[S'_t\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) + S'_x\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-2\tau}\right)\right] \\
&= S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) + \varphi_\varepsilon^2 R_\varepsilon^{(1)}\left(t_{x,\varepsilon}\right). \qquad (2.4)
\end{aligned}$$

Here we used Taylor expansion of $S\left(t_{x,\varepsilon}+\varphi_\varepsilon u, x_{t_{x,\varepsilon}+\varphi_\varepsilon u-\tau}\right)$ by the powers of φ_ε and the properties (2.2). Of course, $\varphi_\varepsilon^2 R_\varepsilon^{(1)}\left(t_{x,\varepsilon}\right)$ is the last term of Taylor expansion.

The equalities

$$x = x_{t_x-\tau} = x_0 + \int_0^{t_x-\tau} S\left(s, x_{s-\tau}\right) \mathrm{d}s,$$

$$x = X_{t_{x,\varepsilon}-\tau} = x_0 + \int_0^{t_{x,\varepsilon}-\tau} S\left(s, X_{s-\tau}\right) \mathrm{d}s + \varepsilon W_{t_{x,\varepsilon}-\tau}$$

allow us to write

$$0 = X_{t_{x,\varepsilon}-\tau} - x_{t_x-\tau} = \int_0^{t_{x,\varepsilon}-\tau} S\left(s, X_{s-\tau}\right) \mathrm{d}s - \int_0^{t_x-\tau} S\left(s, x_{s-\tau}\right) \mathrm{d}s + \varepsilon W_{t_{x,\varepsilon}-\tau}$$

and

$$\int_{t_x-\tau}^{t_{x,\varepsilon}-\tau} S\left(s, x_{s-\tau}\right) \mathrm{d}s = \int_0^{t_{x,\varepsilon}-\tau} \left[S\left(s, X_{s-\tau}\right) - S\left(s, x_{s-\tau}\right)\right] \mathrm{d}s + \varepsilon W_{t_{x,\varepsilon}-\tau}.$$

Hence

$$\begin{aligned}\left|\int_{t_x-\tau}^{t_{x,\varepsilon}-\tau} S\left(s, x_{s-\tau}\right) \mathrm{d}s\right| &\le \int_0^{t_{x,\varepsilon}-\tau} \left|S\left(s, X_{s-\tau}\right) - S\left(s, x_{s-\tau}\right)\right| \mathrm{d}s + \varepsilon \left|W_{t_{x,\varepsilon}-\tau}\right| \\ &\le L \int_0^{T-\tau} \left|X_{s-\tau} - x_{s-\tau}\right| \mathrm{d}s + \varepsilon \left|W_{t_{x,\varepsilon}-\tau}\right| \le \varepsilon H\left(t_{x,\varepsilon}\right),\end{aligned}$$

where the random function $H\left(t_{x,\varepsilon}\right)$ has bounded moments.

We have as well

$$\left|\int_{t_x-\tau}^{t_{x,\varepsilon}-\tau} S\left(s, x_{s-\tau}\right) \mathrm{d}s\right| \ge \kappa \left|t_{x,\varepsilon} - t_x\right|.$$

From the last two estimates we have

$$\left|t_{x,\varepsilon} - t_x\right| \le \varepsilon \kappa^{-1} H\left(t_{x,\varepsilon}\right) \to 0. \tag{2.5}$$

We can write

$$S_\varepsilon(t,x) = \frac{1}{\varphi_\varepsilon} \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) S\left(s, X_{s-\tau}\right) \mathrm{d}s + \frac{\varepsilon}{\varphi_\varepsilon} \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) \mathrm{d}W_s.$$

Hence on the set $\mathbb{B}_\varepsilon$ we have

$$\begin{aligned}&\left|S_\varepsilon(t,x) - S(t,x)\right| = \left|S_\varepsilon\left(t_x, x_{t_x-\tau}\right) - S\left(t_x, x_{t_x-\tau}\right)\right| \\ &\le \left|\frac{1}{\varphi_\varepsilon} \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) \left[S\left(s, x_{s-\tau}\right) - S\left(t_x, x_{t_x-\tau}\right)\right] \mathrm{d}s\right| \\ &\quad + \left|\frac{1}{\varphi_\varepsilon} \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) \left[S\left(s, X_{s-\tau}\right) - S\left(s, x_{s-\tau}\right)\right] \mathrm{d}s\right| \\ &\quad + \left|\frac{\varepsilon}{\varphi_\varepsilon} \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) \mathrm{d}W_s\right| \\ &\le \left|S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) - S\left(t_x, x_{t_x-\tau}\right)\right| + \varphi_\varepsilon^2 \left|R_\varepsilon^{(1)}\left(t_{x,\varepsilon}\right)\right| + \left|R_\varepsilon^{(2)}\left(t_{x,\varepsilon}\right)\right| + \left|R_\varepsilon^{(3)}\left(t_{x,\varepsilon}\right)\right|\end{aligned}$$

with obvious notation.

The first term in the last line can be estimated as follows

$$\begin{aligned}\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} &\left|S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) - S\left(t_x, x_{t_x-\tau}\right)\right|^2 \\ &\le 2M\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} \left|t_{x,\varepsilon} - t_x\right|^2 + 2\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} \left|x_{t_{x,\varepsilon}-\tau} - x_{t_x-\tau}\right|^2 \\ &\le 2M\mathbf{E}_S \left|t_{x,\varepsilon} - t_x\right|^2 + 2\mathbf{E}_S \left|x_{t_{x,\varepsilon}-\tau} - x_{t_x-\tau}\right|^2 \le C\,\varepsilon^2.\end{aligned}$$

For the last estimate see (2.5).

Further,

$$\begin{aligned}\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} R_\varepsilon^{(3)}\left(t_{x,\varepsilon}\right)^2 &\le \mathbf{E}_S R_\varepsilon^{(3)}\left(t_{x,\varepsilon}\right)^2 = \frac{\varepsilon^2}{\varphi_\varepsilon^2}\mathbf{E}_S \int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right)^2 \mathrm{d}s \\ &= \frac{\varepsilon^2}{\varphi_\varepsilon}\mathbf{E}_S \int_{-\frac{t_{x,\varepsilon}}{\varphi_\varepsilon}}^{\frac{T-t_{x,\varepsilon}}{\varphi_\varepsilon}} K\left(u\right)^2 \mathrm{d}u \le \frac{\varepsilon^2}{\varphi_\varepsilon}\int_{-1}^1 K\left(u\right)^2 \mathrm{d}u = C\,\frac{\varepsilon^2}{\varphi_\varepsilon}\end{aligned}$$

and (below $t_{x,\varepsilon,u} = t_{x,\varepsilon} + \varphi_\varepsilon u$)

$$\begin{aligned}\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} R_\varepsilon^{(2)}\left(t_{x,\varepsilon}\right)^2 &\le \mathbf{E}_S R_\varepsilon^{(2)}\left(t_{x,\varepsilon}\right)^2 \\ &= \mathbf{E}_S \left|\frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right)\left[S\left(s, X_{s-\tau}\right) - S\left(s, x_{s-\tau}\right)\right]\mathrm{d}s\right|^2 \\ &\le 2\mathbf{E}_S \int_{-1}^1 K\left(u\right)^2 \left|S\left(t_{x,\varepsilon,u}, X_{t_{x,\varepsilon,u}-\tau}\right) - S\left(t_{x,\varepsilon,u}, x_{t_{x,\varepsilon,u}-\tau}\right)\right|^2 \mathrm{d}u \\ &\le 2L^2 \int_{-1}^1 K\left(u\right)^2 \mathbf{E}_S \left|X_{t_{x,\varepsilon,u}-\tau} - x_{t_{x,\varepsilon,u}-\tau}\right|^2 \mathrm{d}u \le C\,\varepsilon^2.\end{aligned}$$

Here we used the estimate (A.1).

These estimates allow us to write

$$\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} \left|S_\varepsilon\left(t,x\right) - S\left(t,x\right)\right|^2 \le C\,\varphi_\varepsilon^4 + C\,\varepsilon^2 + C\,\frac{\varepsilon^2}{\varphi_\varepsilon}. \tag{2.6}$$

The probability of the compliment of the set $\mathbb{B}_\varepsilon$ we estimate as follows

$$\begin{aligned}\mathbf{P}_S\left(\mathbb{B}_\varepsilon^c\right) &\le \mathbf{P}_S\left(\varphi_\varepsilon^{-1} t_{x,\varepsilon} \le 2\right) + \mathbf{P}_S\left(\varphi_\varepsilon^{-1}\left(T - t_{x,\varepsilon}\right) \le 2\right) \\ &= \mathbf{P}_S\left(t_x - t_{x,\varepsilon} \ge t_x - 2\varphi_\varepsilon\right) + \mathbf{P}_S\left(T - t_{x,\varepsilon} \le 2\varphi_\varepsilon\right).\end{aligned}$$

We have for $\tau > 4\varphi_\varepsilon$

$$\begin{aligned}\mathbf{P}_S\left(t_x - t_{x,\varepsilon} \ge t_x - 2\varphi_\varepsilon\right) &\le \mathbf{P}_S\left(\left|t_x - t_{x,\varepsilon}\right| \ge t_x - \frac{\tau}{2}\right) \le \mathbf{P}_S\left(\left|t_x - t_{x,\varepsilon}\right| \ge \frac{\tau}{2}\right) \\ &\le \mathbf{P}_S\left(C\varepsilon \sup_{0\le t\le T}\left|W_t\right| \ge \frac{\tau}{2}\right) \le Ce^{-c\varepsilon^{-2}}\end{aligned}$$

with positive constants $C > 0, c > 0$. Here we used (2.5). The similar estimate we have for the probability $\mathbf{P}_S\left(T - t_{x,\varepsilon} \le 2\varphi_\varepsilon\right)$.

This allow us to estimate the mean squared error on the set $\mathbb{B}^c$

$$\left(\mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon^c\}} \left|S_\varepsilon(t,x) - S(t,x)\right|^2\right)^2 \le \mathbf{E}_S \mathbb{I}_{\{\mathbb{B}_\varepsilon^c\}} \mathbf{E}_S \left|S_\varepsilon(t,x) - S(t,x)\right|^4 \le C e^{-c\varepsilon^{-2}}.$$

Hence

$$\mathbf{E}_S \left|S_\varepsilon(t,x) - S(t,x)\right|^2 \le C\varphi_\varepsilon^4 + C\varepsilon^2 + C\frac{\varepsilon^2}{\varphi_\varepsilon} + Ce^{-\frac{c}{2}\varepsilon^{-2}}$$

As usual in such situations the optimal rate $\varphi_\varepsilon = \varepsilon^{2/5}$ we obtain from the equation

$$\varphi_\varepsilon^4 = \frac{\varepsilon^2}{\varphi_\varepsilon}.$$

Therefore the estimate (2.3) is proved. □

Remark. Suppose that the function $S(t,x)$ is Lipshits and the function $\Phi(t) = S\left(t, x_{t-\tau}(\vartheta_0)\right), t \in [0,T]$ has M continuous derivatives. Moreover, the M-th derivative is Holder of order $\mu \in (0,1)$:

$$\left|\frac{\mathrm{d}^M \Phi(t)}{\mathrm{d}t^M} - \frac{\mathrm{d}^M \Phi(t_1)}{\mathrm{d}t_1^M}\right| \le H\left|t - t_1\right|^\mu.$$

The set of functions $S(t, x_{t-\tau})$ with the same constants $C > 0$ and μ we denote as $\mathcal{F}_{M+\mu}$.

Introduce additional condition on the kernel $K(\cdot)$

$$\int_{-1}^{1} u^m K(u)\,\mathrm{d}u = 0, \qquad m = 1, \dots, M. \tag{2.7}$$

Using the equality (2.4) and Taylor expansion of the function $\Phi\left(t_{x,\varepsilon} + \varphi_\varepsilon u\right) = S\left(t_{x,\varepsilon} + \varphi_\varepsilon u, x_{t_{x,\varepsilon}} - \varphi_\varepsilon u - \tau\right)$ by the powers of φ_ε we obtain the expression

$$\begin{aligned}
\frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s - t_{x,\varepsilon}}{\varphi_\varepsilon}\right) S(s, x_{s-\tau})\,\mathrm{d}s &= \int_{-1}^{1} K(u)\,\Phi\left(t_{x,\varepsilon} + \varphi_\varepsilon u\right)\mathrm{d}u \\
&= S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) + \sum_{m=1}^{M} \int_{-1}^{1} u^m K(u)\,\mathrm{d}u\,\frac{\Phi^{(m)}\left(t_{x,\varepsilon}\right)}{m!}\varphi_\varepsilon^m + Q_\varepsilon \\
&= S\left(t_{x,\varepsilon}, x_{t_{x,\varepsilon}-\tau}\right) + Q_\varepsilon.
\end{aligned}$$

The last term is estimated as follows

$$|Q_\varepsilon| \le \frac{\varphi_\varepsilon^M}{M!}\int_{-1}^{1} \left|u^M K(u)\right| \left|\Phi^{(M)}\left(t_{x,\varepsilon} + \tilde{\varphi}_\varepsilon u\right) - \Phi^{(M)}\left(t_{x,\varepsilon}\right)\right| \mathrm{d}u \le C\varphi_\varepsilon^{M+\mu}. \tag{2.8}$$

The equation (2.6) became

$$\mathbf{E}_S \left| S_\varepsilon (t,x) - S(t,x) \right|^2 \le C \varphi_\varepsilon^{2(M+\mu)} + C\varepsilon^2 + C\frac{\varepsilon^2}{\varphi_\varepsilon}.$$

Now the equation $\varphi_\varepsilon^{2(M+\mu)} = \varepsilon^2/\varphi_\varepsilon$ provides the rate $\varphi_\varepsilon = \varepsilon^{\frac{1}{M+\mu+\frac{1}{2}}}$ and

$$\mathbf{E}_S \left| S_\varepsilon (t,x) - S(t,x) \right|^2 \le C\varepsilon^{-\frac{2(M+\mu)}{M+\mu+1/2}}. \tag{2.9}$$

Hence slightly modifying the proof we obtain for any $p > 0$ the estimate

$$\sup_{S(\cdot)\in\mathcal{F}_{M+\mu}} \varepsilon^{-\frac{p(M+\mu)}{M+\mu+1/2}} \mathbf{E}_S \left| S_\varepsilon (t,x) - S(t,x) \right|^p \le C.$$

Recall that for the model of observations (1.1) we have the lower bound ($t = t_x$ such that $x_{t_x-\tau} = x$)

$$\varliminf_{\varepsilon\to 0} \inf_{\bar S_\varepsilon (t,x)} \sup_{S(\cdot)\in\mathcal{F}_{M+\mu}} \varepsilon^{-\frac{2(M+\mu)}{M+\mu+1/2}} \mathbf{E}_S \left| \bar S_\varepsilon (t,x) - S(t,x) \right|^2 > 0, \tag{2.10}$$

where inf is taken over all possible estimators $\bar S_\varepsilon (t,x)$ of the function $S(t,x)$ (Theorem 4.3, [18]). The inequality (2.10) shows that the estimator converging to $S(t,x)$ faster than $S_\varepsilon (t,x)$ does not exist and therefore $S_\varepsilon (t,x)$ is asymptotically efficient in order of convergence.

Remark. Suppose that the delay τ is unknown and we have to estimate the function $f(t) = S(t, x_{t-\tau}), 0 < t < T$. The function $f(t)$ is M times continuously differentiable and the M-th derivative is Hölder of order $\mu \in (0,1)$. It is possible to estimate the function $f(t)$ using even more simple kernel-type estimator than above:

$$\begin{aligned}\hat f_\varepsilon (t) &= \frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s-t}{\varphi_\varepsilon}\right) \mathrm{d}X_s \\ &= \frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s-t}{\varphi_\varepsilon}\right) S(s, X_{s-\tau})\,\mathrm{d}s + \frac{\varepsilon}{\varphi_\varepsilon}\int_0^T K\left(\frac{s-t}{\varphi_\varepsilon}\right) \mathrm{d}W_s.\end{aligned}$$

As $S(s, X_{s-\tau}) = S(s, x_{s-\tau}) + O(\varepsilon)$ and

$$\mathbf{E}_S \left(\frac{\varepsilon}{\varphi_\varepsilon}\int_0^T K\left(\frac{s-t}{\varphi_\varepsilon}\right) \mathrm{d}W_s\right)^2 = \frac{\varepsilon^2}{\varphi_\varepsilon}\int_{-\frac{t}{\varphi_\varepsilon}}^{\frac{T-t}{\varphi_\varepsilon}} K(u)^2\,\mathrm{d}u \le C\frac{\varepsilon^2}{\varphi_\varepsilon}$$

it is sufficient to study the integral

$$J_\varepsilon = \frac{1}{\varphi_\varepsilon}\int_0^T K\left(\frac{s-t}{\varphi_\varepsilon}\right) S(s, x_{s-\tau})\,\mathrm{d}s.$$

Suppose that the kernel $K(\cdot)$ satisfies the conditions (2.2) and (2.7). Then using Taylor formula once more we can write

$$J_\varepsilon = \int_{-1}^{1} K(u) \left[f(t) + \sum_{m=1}^{M} \frac{\varphi_\varepsilon^m u^m}{m!} f^{(m)}(t) \right] \mathrm{d}u + \varphi_\varepsilon^M R_\varepsilon = f(t) + \varphi_\varepsilon^M R_\varepsilon.$$

For the last term of Taylor expansion we have estimate similar to (2.8)

$$\varphi_\varepsilon^M |R_\varepsilon| \le C \varphi_\varepsilon^{M+\mu}.$$

Therefore we have the same normalizing function $\varphi_\varepsilon = \varepsilon^{\frac{1}{M+\mu+\frac{1}{2}}}$ and the same estimate for the mean squared error as in (2.9).

A. Appendix section

Suppose that the function $S(t,x), t \in [0,T], x \in \mathcal{R}$ has continuous bounded derivative $|S'_x(t,x)| \le L$ w.r.t. x, then the equation (1.1) has a unique strong solution (see Theorem in [24]). Moreover, we have the following estimates (with probability 1)

$$\begin{aligned} |X_t - x_t| &\le \int_0^t |S(s, X_{s-\tau}) - S(s, x_{s-\tau})| \, \mathrm{d}s + \varepsilon |W_t| \\ &\le L \int_0^t |X_{s-\tau} - x_{s-\tau}| \, \mathrm{d}s + \varepsilon |W_t| \le L \int_0^t |X_s - x_s| \, \mathrm{d}s + \varepsilon \sup_{0 \le s \le T} |W_s|. \end{aligned}$$

Hence by Gronwall-Bellman lemma we can write

$$|X_t - x_t| \le \varepsilon C \widehat{W}_T, \tag{A.1}$$

where we denoted $\widehat{W}_T = \sup_{0 \le s \le T} |W_s|$.

For the second moment we have

$$\begin{aligned} \mathbf{E}_S |X_t - x_t|^2 &\le 2L^2 T \int_0^t \mathbf{E}_S |X_{s-\tau} - x_{s-\tau}|^2 \, \mathrm{d}s + 2\varepsilon^2 T \\ &\le 2L^2 T \int_0^t \mathbf{E}_S |X_s - x_s|^2 \, \mathrm{d}s + 2\varepsilon^2 T \end{aligned}$$

and by the same lemma

$$\mathbf{E}_S |X_t - x_t|^2 \le C \varepsilon^2 \tag{A.2}$$

Let us write the formal derivative $X_t^{(1)}$ of X_t w.r.t. ε, then we obtain the equation

$$\mathrm{d}X_t^{(1)} = S'_x(t, X_{t-\tau}) X_{t-\tau}^{(1)} \mathrm{d}t + \mathrm{d}W_t, \qquad X_s^{(1)} = 0, \text{ for } s \le 0, \quad 0 \le t \le T.$$

The Gaussian process $X_t^{(1)}|_{\varepsilon=0} = x_t^{(1)}$ satisfies the linear equation

$$\mathrm{d}x_t^{(1)} = S'_x(t, x_{t-\tau}) x_{t-\tau}^{(1)} \mathrm{d}t + \mathrm{d}W_t, \qquad x_s^{(1)} = 0, \text{ for } s \le 0, \quad 0 \le t \le T.$$

The proof of (A.1) can be done using the standard technique based on Gronwall-Bellman lemma.

References

[1] Apoyan G.T. (1986). Parameter estimation of non differentiable trend coefficient. (in Russian) *Uchen. Zap. Erevan Univ.* **1**, 33-62.

[2] Apoyan G.T. and Kutoyants Yu. A. (1985). On state estimation of dynamical system perturbed by small noise, *Proc. 5-th Pannonian Symp. Math. Statist.*, Visegrád, 247-253.

[3] Banon G. (1980). Nonparametric identification for diffusion processes, *SIAM J.Contr. and Optim.* **16**, 3, 380-395.

[4] Bosq, D. (1998) *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statistics, 110 (2-nd edition), Springer, New York.

[5] Bosq, D. and Blanke, D. (2007) *Inference and Prediction in Large Dimensions*, J. Wiley & Sons,

[6] Dalalyan, A., Kutoyants, Yu. A. (2002) Asymptotically efficient trend coefficient estimation for ergodic diffusion. *Math. Methods Statist.*, **11**, 4, 402–427.

[7] Freidlin M.I. and Wentsell A.D. (1984). *Random Perturbations of Dynamical Systems*, Springer, New York.

[8] Gushchin, A.A. and Küchler, U. (1999). Asymptotic inference for a linear stochastic differential equation with time delay *Bernoulli* **5**, 6, 1059-1098.

[9] Gushchin, A.A. and Küchler, U. (2000). On stationary solutions of delay differential equations driven by a Lévy process. *Stochastic Process. Appl.* **88**, 2, 195-211.

[10] Ibragimov I.A. and Has'minskii R.Z. (1981). *Statistical Estimation: Asymptotic Theory* , Springer, Berlin.

[11] Benyahia, W. (2012). *Estimation paramétrique de la densité des retards dans un processus de diffusion*, These PhD, Université Abou-Bakr Belkaid, Tlemcen.

[12] Korso Feciane, M. (2010) Sur l'estimation des retards multiples dans une diffusion non lineair, *Pub.Inst.Stat.Univ.Paris*, **LIV**, 3, 41-66.

[13] Korso Feciane, M. (2010). *Estimation dans un processus de diffusion non linéaire à retards et tests*, These PhD, Université Abou-Bakr Belkaid, Tlemcen.

[14] Küchler, U. Kutoyants, Yu. A. (2000). Delay estimation for some stationary diffusion-type processes, *Scandinavian Journal of Statistics*, **27**, 3, 405-414.

[15] Kutoyants, Yu.A. (1985). On nonparametric estimation of the trend coefficient of diffusion process. In *Statistics and Control of Stochastic Processes (Steklov Seminar 1984)*, Optimization Software, New York, 230-250.

[16] Kutoyants Yu.A. (1988). On efficient nonparametric estimation of the trend coefficient. In *Statistics and Control of Stochastic Processes (Steklov Seminar 1985)*, Optimization Software, New York, 196-207.

[17] Kutoyants, Yu. A. (1988). An example of estimating a parameter of a non-differentiable drift coefficient. *Theory Probab. Appl.*, **33**, 1, 175-179.

[18] Kutoyants, Yu. A. (1994). *Identification of Dynamical Systems with Small Noise*. Kluwer, Dordrecht.

[19] KUTOYANTS, YU. A. (2004). *Statistical Inference for Ergodic Diffusion Processes.* Springer, London.

[20] KUTOYANTS, YU. A. (2005). On delay estimation for stochastic differential equations. *Stochastics and Dynamics*, **5**, 2, 333-342.

[21] KUTOYANTS, YU. A. AND MOURID, T. (1992). Estimation dans un modèle autorégressif avec retards. *C.R.Acad.Paris*, 315, ser **I**, 455-458.

[22] KUTOYANTS, YU. A. AND MOURID, T. (1994). Estimation par la distance minimale pour un processus de type diffusion avec retards. *Publications de l'ISUP.*, **38**, 2, 3-18

[23] KUTOYANTS, YU. A., MOURID, T. AND BOSQ, D. (1992). Estimation paramétrique d'un processus de diffusion avec retards. *Ann. Inst. Henri Poincaré* , **28**, 1, 95-106.

[24] LIPTSER, R. S. AND SHIRYAEV, A. N. (2001). *Statistics of Random Processes.* 2-nd ed., Vol. 1, Springer, New York.

[25] NGUYEN H.T., PHAM D.T. (1982) Identification of nonstationary diffusion model by the method of sieves, *SIAM J.Contr.and Optim.* **20**, 5, 603-611.

Av. O. Messiaen, 72085, Le Mans, France
e-mail: kutoyants@univ-lemans.fr