Sequential nonparametric pointwise estimation of the drift in ergodic diffusion processes

Jean-Yves Brua

Institut de Recherche Mathématique Avancée,
Université de Strasbourg,
7 rue René-Descartes, 67084 Strasbourg Cedex, France
brua@math.u-strasbg.fr

Session: ”Sequential estimation for dependent observations”, organizer : L. Galtchouk.

Abstract. This study deals with the pointwise estimation of the drift function of an ergodic diffusion. The drift function is assumed to belong to a neighborhood centred on a Lipschitz function and involving a weak Hölder class. The maximal risk of an estimator of the drift is defined over this neighborhood using the absolute error loss and the ergodic density as normalizing factor. Then the optimal convergence rate and the sharp asymptotic lower bound are found for the minimax risk. Eventually an asymptotically efficient kernel estimator, that is an estimator for which the maximal risk attains this bound, is constructed.

Keywords. Asymptotic efficiency, drift, ergodic diffusion, minimax, nonparametric estimation, sequential estimation.

1 Introduction

This paper is devoted to the problem of estimating the drift coefficient of an ergodic diffusion, solution of the stochastic differential equation

$$dX_t = S(X_t)dt + dB_t, \quad 0 \le t \le T \tag{1}$$

where $(B_t)_{t\geq 0}$ is a scalar standard Brownian motion. Assuming that we observe the continuous data $(X_t, 0 \le t \le T)$ and knowing the smoothness of S, we want to estimate the function S at a fixed point $x_0 \in \mathbb{R}$.

Model (1) finds applications in numerous fields, namely financial mathematics, econometrics, stochastic control, filtering and others, see for example Ait-Sahalia (2002), Jiang and Knight (1997) and Liptser and Shiryaev (1978). There is a lot of papers studying the problem of nonparametric estimation of the drift of a diffusion. For instance, Banon (1978) proved the consistency whereas Pham (1981) considered the convergence rate of kernel estimators for our model (1).

Our current purpose is to estimate the drift function at a single point using the absolute error loss to quantify the performance of an estimator through its corresponding risk. We are first interested in obtaining the exact asymptotic behavior of the minimax risk and particularly in finding an asymptotic lower bound of it. Secondly we aim at constructing an estimator of the drift coefficient for which the risk is asymptotically bounded from above by the same constant as for the minimax risk. Such an estimator will be called asymptotically efficient, see Ibragimov and Has’minskii (1981).

The problem of sharp estimation of the drift function in diffusion models has already been treated for some Sobolev classes: Dalalyan and Kutoyants (2002) with known smoothness then Dalalyan (2005) with an unknown one proposed asymptotically efficient estimators of the drift coefficient in model (1) for an L^2-type risk.

The drift estimation in some Hölder classes was handled as well. Galtchouk and Pergamenshchikov (2004) achieved the optimal convergence rate of the minimax risk for the $L^2([a;b], dx)$-loss function as the regularity of the drift function is known but as it remains unknown too. The optimal convergence rate of the minimax risk is also obtained for the pointwise estimation of the drift function with unknown smoothness and for the absolute error loss in Galtchouk and Pergamenshchikov (2001). For any positive power of the absolute error loss and the same Hölder class, the sharp constant for the local minimax risk is given in Galtchouk and Pergamenshchikov (2005) as well as an asymptotically efficient estimator of the drift coefficient with known regularity.

We consider here the pointwise estimation of the drift function belonging to a Hölder class with known smoothness using the absolute error loss. For this problem Galtchouk and Pergamenshchikov (2006) gave the sharp asymptotic lower bound for the local minimax risk and an asymptotically efficient kernel estimator. More precisely they assumed that the drift function belonged to a neighborhood centred on a Lipschitz function. The neighborhood consists in the centre plus a function satisfying a weak Hölder condition (involving a weak Hölder constant) and having a small norm. The asymptotic results were given with the time of observations tending to infinity, the weak Hölder constant and the diameter of the neighborhood to zero. We propose to find the sharp asymptotic lower bound for the minimax risk and an asymptotically efficient estimator without making the neighborhood tend to its centre but by keeping its diameter constant.

2 Statement of the problem

In model (1) we are interested in the estimation of the unknown function S at a fixed point $x_0 \in \mathbb{R}$ assuming that S lies in a neighborhood of a function S_0 belonging to

$$\Sigma_{L,M} := \left\{ f : |f(0)| \le L, -L \le \frac{f(x)-f(y)}{x-y} \le -M, \forall x, y \in \mathbb{R} \right\},$$

with $0 < M < L$.

As mentioned in the introduction the neighborhood of the function $S_0 \in \Sigma_{L,M}$ consists in the centre plus an other function lying in

$$\mathcal{U}_{\delta,\beta}(S_0) = \left\{ S : S = S_0 + D, D \in \mathcal{H}^w_{x_0}(\delta,\beta) \right\}$$

where the weak Hölder class is defined by:

$$\mathcal{H}^w_{x_0}(\delta,\beta) = \Big\{ \; D \text{ differentiable } : \sup_{x\in\mathbb{R}} \Big(|D(x)| + |\dot{D}(x)| \Big) \le B; \\ \forall h > 0, \; \left| \int_{-1}^{1} (D(x_0 + zh) - D(x_0)) dz \right| \le \delta h^\beta \Big\},$$

with $\beta \in]1;2[$, $0 < \delta < 1$ and $0 < B < M$.

If $S_0 \in \Sigma_{L,M}$ one has $\mathcal{U}_{\delta,\beta}(S_0) \subset \Sigma_{L+B,M-B}$, so that for all $S \in \mathcal{U}_{\delta,\beta}(S_0)$ the process $(X_t)_{t\ge0}$ is ergodic and there exists an ergodic density q_S. The risk of an estimator $\tilde{S}_T(x_0)$ of $S(x_0)$ is the following

$$\mathcal{R}_{\delta,\beta}(\tilde{S}_T, S_0) = \sup_{S\in\mathcal{U}_{\delta,\beta}(S_0)} \sqrt{2q_S(x_0)}\varphi_T \mathbb{E}_S |\tilde{S}_T(x_0) - S(x_0)|, \quad \varphi_T = T^{\beta/(2\beta+1)}.$$

3 Main results

The asymptotic lower bound of the minimax risk is given in the following theorem.

Theorem 1. *If $S_0 \in \Sigma_{L,M}$ we have for any $\delta \in]0;1[$*

$$\liminf_{T\to\infty} \inf_{\tilde{S}} \mathcal{R}_{\delta,\beta}(\tilde{S}_T, S_0) \ge \mathbb{E}|\xi|, \quad \xi \sim \mathcal{N}(0,1),$$

where the infimum is taken over all estimators of $S(x_0)$.

Proof: For $u > 0$, denote $S_u(x) = S_0(x) + uD_\nu(x)$, where

$$D_\nu(x) = \varphi_T^{-1} V_\nu\left(\frac{x-x_0}{h}\right), \; V_\nu(x) = \nu^{-1} \int_{-\infty}^{\infty} \tilde{Q}_\nu(u) g\left(\frac{u-x}{\nu}\right) du,$$

$$\tilde{Q}_\nu(u) = \mathbb{I}_{\{|u|\le 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\le|u|\le1-\nu\}}, \; g(z) = \begin{cases} c\exp(-(1-z^2)^{-1}), & |z| \le 1; \\ 0, & |z| > 1, \end{cases}$$

with $0 < \nu < 1/4$; the normalizing constant $c > 0$ is such that $\int_{-1}^{1} g(z)dz = 1$.
Now fix $b > 0$ and $\delta > 0$. We can easily see that there exists $T_{b,\nu} > 0$ such that for any $|u| \le b$ and $T \ge T_{b,\nu}$ one has $S_u \in \mathcal{U}_{\delta,\beta}(S_0)$.

Furthermore we may write for all $T \ge T_{b,\nu}$

$$\mathcal{R}_{\delta,\beta}(\tilde{S}_T, S_0) \ge \frac{1}{2b}\int_{-b}^{b} \sqrt{2q_{S_u}(x_0)}\mathbb{E}_{S_u}\Psi_{a,T}(\tilde{S}_T, S_u)du =: I_T(a,b),$$

with $\Psi_{a,T}(\tilde{S}_T, S) = v_a\left(\varphi_T(\tilde{S}_T(x_0) - S(x_0))\right)$ and $v_a(x) = a \wedge |x|$, $a > 0$.

Denoting $\mathbb{P}_S$ the distribution of the process (X_t) in $C([0;T])$ as the drift function is S, we have thanks to Lemma 4.2 in Galtchouk and Pergamenshchikov (2006)

$$\rho_T(u) := \frac{d\mathbb{P}_{S_u}}{d\mathbb{P}_{S_0}} = \exp\{u\Delta_T - \frac{1}{2}u^2\sigma_\nu^2 + r_T(u)\}, \ \ \forall u > 0,$$

where $\Delta_T = \int_0^T D_\nu(X_t)dB_t$ and $\sigma_\nu^2 = q_{S_0}(x_0)\int_{-1}^{1} V_\nu^2(z)dz$.

Moreover we can assume that ξ_ν and Δ_T are independent and we have

$$\Delta_T \xrightarrow[T\to\infty]{\mathcal{L}_{\mathbb{P}_{S_0}}} \xi_\nu, \ \ \xi_\nu \sim \mathcal{N}(0, \sigma_\nu^2) \quad \text{and} \quad r_T(u) \xrightarrow[T\to\infty]{\mathbb{P}_{S_0}} 0, \tag{2}$$

Then write

$$I_T(a,b) \ge \frac{1}{2b}\int_{-b}^{b} \sqrt{2q_{S_u}(x_0)}\mathbb{E}_{S_0}\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_u)\rho_T^0(u)du + \delta_T^0(a,b) =: J_T(a,b) + \delta_T^0(a,b),$$

where $B_d = \{|\Delta_T| \le d\}$, $d = \sigma_\nu^2(b - \sqrt{b})$, $\rho_T^0(u) = \exp(u\Delta_T - u^2\sigma_\nu^2/2)$.

Using the fact that the family $\{\rho_T(u), T > 0\}$ is uniformly integrable and the convergence (2) of $r_T(u)$, one can show that

$$\sup_{\tilde{S}_T} \delta_T^0(a,b) \xrightarrow[T\to\infty]{} 0. \tag{3}$$

Now we consider the quantity

$$\frac{1}{2b}\int_{-b}^{b} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_u)\left(\sqrt{2q_{S_u}(x_0)} - \sqrt{2q_{S_0}(x_0)}\right)\rho_T^0(u)du =: J_T(a,b) - K_T(a,b).$$

Bounded convergence yields $q_{S_{\nu,u}}(x_0) \xrightarrow[T\to\infty]{} q_{S_0}(x_0)$ and then

$$\sup_{\tilde{S}_T} |J_T(a,b) - K_T(a,b)| \xrightarrow[T\to\infty]{} 0. \tag{4}$$

As a consequence only the study of the quantity $K_T(a,b)$ remains. So rewrite

$$\rho_T^0(u) = \zeta_T \exp(-\sigma_\nu^2(u - \tilde{\Delta}_T)^2/2), \quad \zeta_T = \exp(\Delta_T^2/2\sigma_\nu^2), \quad \tilde{\Delta}_T = \Delta_T/\sigma_\nu^2$$

and put $g_T = \varphi_T(\tilde{S}_T(x_0) - S_0(x_0))$, $\tilde{g}_T = g_T - \tilde{\Delta}_T$. Then one successively has

$$\begin{aligned}
K_T(a,b) &= \frac{1}{2b}\int_{-b}^{b} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_u)\sqrt{2q_{S_0}(x_0)}\rho_T^0(u)du \\
&= \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-b}^{b} v_a(u - g_T)\exp\left(-\sigma_\nu^2(u - \tilde{\Delta}_T)^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&\ge \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} v_a(u - \tilde{g}_T)\exp\left(-\sigma_\nu^2 u^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&\ge \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} v_a(u)\exp\left(-\sigma_\nu^2 u^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&=: \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} |u|\exp\left(-\sigma_\nu^2 u^2/2\right)\sqrt{2q_{S_0}(x_0)}du + \delta_T^1(a,b),
\end{aligned}$$

the second inequality arising from Anderson's lemma (see Ibragimov and Has'minskii, 1981, Lemma 10.2, p.157).
Noticing that

$$\lim_{T\to\infty} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T = 2\sigma_\nu(b-\sqrt{b})/\sqrt{2\pi},$$

one obtains for all $b > 0$,

$$\lim_{a\to\infty} \liminf_{T\to\infty} \delta_T^1(a,b) = 0, \tag{5}$$

and

$$\liminf_{a\to\infty} \liminf_{T\to\infty} K_T(a,b) \geq \frac{b-\sqrt{b}}{b}\frac{\sqrt{2q_{S_0}(x_0)}\sigma_\nu}{\sqrt{2\pi}} \int_{-\sqrt{b}}^{\sqrt{b}} |u| \exp(-u^2\sigma_\nu^2/2)du. \tag{6}$$

Remarking here that $\sigma_\nu \to 2q_{S_0}(x_0)$ as $\nu \to 0$ and limiting $b \to \infty$ before $\nu \to 0$ in (6) yield with (3), (4) and (5) :

$$\liminf_{T\to\infty} \mathcal{R}_{\delta,\beta}(\tilde{S}_T, S_0) \geq \mathbb{E}|\xi|, \quad \xi \sim \mathcal{N}(0,1).$$

□

In order to exhibit an asymptotically efficient estimator of $S(x_0)$ we begin with estimating the ergodic density at the point x_0 through the observations $\{X_t, t \leq t_0\}$, where $t_0 = T^{2\gamma}$, $\gamma_* < \gamma < 1/2$ and $\gamma_* = \frac{\beta-1}{2\beta+1}$. Let

$$\hat{q}_T(x_0) = \frac{1}{2t_0 l_T}\int_0^{t_0} Q\left(\frac{X_t - x_0}{l_T}\right)dt,$$

with $Q = \mathbb{I}_{[-1;1]}$ and $l_T = o(1/\sqrt{T})$ as $T \to \infty$.
Then for $H > 0$ we define the sequential procedure $(\tau_H, S_T^*(x_0))$ as

$$\tau_H = \inf\{t \geq t_0 : \int_{t_0}^t Q\left(\frac{X_t - x_0}{h}\right)dt \geq H\},$$

$$S_T^*(x_0) = \frac{1}{H}\int_{t_0}^{\tau_H} Q\left(\frac{X_t - x_0}{h}\right)dX_t\mathbb{I}_{\{\tau_H \leq T\}}. \tag{7}$$

We choose the bandwidth $h = h_T = T^{-1/(2\beta+1)}$ and the level $H = H_T = (T - t_0)(2\tilde{q}_T(x_0) - \varepsilon_T)h_T$, where $\tilde{q}_T(x_0) = \max(\hat{q}_T(x_0), \nu_T^{-1/2})$, $\varepsilon_T = 1/(\nu_T T^{\gamma_*})$ and $\nu_T = \ln T$.
The centre S_0 of the considered neighborhood needs to verify an additional condition described by the following property:

$$\lim_{y\to x_0} \frac{\dot{S}_0(y) - \dot{S}_0(x_0)}{|y - x_0|^{\beta-1}} = 0. \tag{8}$$

We are now able to give the upper bound of the risk for the estimator (7).

Theorem 2. *Let $S_0 \in \Sigma_{L,M}$, $\beta \in]1;2[$ and assume that the condition (8) is satisfied. Then one has*

$$\limsup_{\delta\to 0}\limsup_{T\to\infty} \mathcal{R}_{\delta,\beta}(S_T^*(x_0), S_0) \leq \mathbb{E}|\xi|, \quad \xi \sim \mathcal{N}(0,1).$$

As a consequence of Theorems 1 and 2, the estimator (7) of the drift coefficient is asymptotically efficient.
PROOF: Let $S \in \mathcal{U}_{\delta,\beta}(S_0)$ and parse the error of estimation as

$$S_T^*(x_0) - S(x_0) = \left(B_T - G_T + \frac{\xi_T}{\sqrt{H_T}}\right)\mathbb{I}_{\{\tau_H \leq T\}} - S(x_0)\mathbb{I}_{\{\tau_H > T\}},$$

where

$$B_T = \frac{1}{H_T}\int_{t_0}^{T} Q\left(\frac{X_t - x_0}{h}\right)(S(X_t) - S(x_0))dt,$$
$$G_T = \frac{1}{H_T}\int_{\tau_H}^{T} Q\left(\frac{X_t - x_0}{h}\right)(S(X_t) - S(x_0))dt,$$
$$\xi_T = \frac{1}{\sqrt{H_T}}\int_{t_0}^{T} Q\left(\frac{X_t - x_0}{h}\right)dB_t.$$

First we want to show that

$$\lim_{\delta\to 0}\limsup_{T\to\infty}\sup_{S\in\mathcal{U}_{\delta,\beta}(S_0)} \mathbb{E}_S\sqrt{2q_S(x_0)}\varphi_T|B_T|\mathbb{I}_{\{\tau_H\le T\}} = 0. \tag{9}$$

We begin with writing

$$B_T = \frac{T}{(2\tilde{q}_T(x_0) - \varepsilon_T)(T - t_0)}\left(\frac{T - t_0}{T}m(f_h) + \frac{1}{\sqrt{T}}\Delta_{t_0,T}(f_h)\right), \tag{10}$$

where $f_h(y) = \phi_h(y)(S(y) - S(x_0))$, $\phi_h(y) = \frac{1}{h}Q\left(\frac{y-x_0}{h}\right)$, $m(f) = \int f(y)q_S(y)dy$ and $\Delta_{t_0,T}(f) = \frac{1}{\sqrt{T}}\int_{t_0}^{T}(f(X_t) - m(f))dt$.

We can rewrite the term $m(f_h)$ as

$$\begin{aligned} m(f_h) &= \int_{-1}^{1}(S(x_0 + hz) - S(x_0))(q_S(x_0 + hz) - q_S(x_0))dz + q_S(x_0)\int_{-1}^{1}(S(x_0 + hz) - S(x_0))dz \\ &=: m_1(h) + q_S(x_0)m_0(h). \end{aligned}$$

Putting $r(y) := S_0(y) - S_0(x_0) - \dot{S}_0(x_0)(y - x_0)$, we easily get

$$\begin{aligned} |m_0(h)| &= \left|\int_{-1}^{1} r(x_0 + zh)dz + \int_{-1}^{1}(D(x_0 + zh) - D(x_0))dz\right| \\ &\le 2\sup_{|u|\le h}\frac{|r(x_0 + u)|}{|u|^\beta}h^\beta + \delta h^\beta. \end{aligned}$$

Using the "zero-constant" Hölder condition (8) one can show that

$$\lim_{\delta\to 0}\lim_{T\to\infty}\varphi_T|m_0(h)| = 0. \tag{11}$$

Moreover it is not difficult to see that there exists a constant $C = C(L, M, B)$ such that

$$\varphi_T|m_1(h)| \le C\varphi_T h^2 = Ch^{2-\beta}. \tag{12}$$

In addition, thanks to inequality (A.1) in Galtchouk and Pergamenshchikov (2006), there exists a constant $\kappa > 0$ such that for all $\lambda > 0$:

$$\sup_{T\ge 1}\sup_{0\le t_0\le T}\sup_{S\in\Sigma_{L+B,M-B}} \mathbb{P}_S\left(|\Delta_{t_0,T}(f_h)| > \lambda\right) \le 2e^{-\kappa\lambda^2}. \tag{13}$$

Hence it is easy to prove that $\varphi_T T^{-1/2}\mathbb{E}_S|\Delta_{t_0,T}(f_h)|$ tends to zero as $T \to \infty$ uniformly in $S \in \mathcal{U}_{\delta,\beta}(S_0)$.

Now remark that $q_*(x_0) := \inf_{S\in\Sigma_{L+B,M-B}} q_S(x_0) > 0$ and write for sufficiently large T

$$\frac{1}{2\tilde{q}_T(x_0) - \varepsilon_T} \le \frac{1}{\tilde{q}_T(x_0)} \le \left|\frac{1}{\tilde{q}_T(x_0)} - \frac{1}{q_S(x_0)}\right| + \frac{1}{q_*(x_0)}.$$

Assertion (9) follows then from (10), (11), (12) and Lemma A.2 in Galtchouk and Pergamenshchikov (2006).

Now let us show that for all $\delta \in (0,1)$

$$\lim_{T\to\infty} \sup_{S\in\mathcal{U}_{\delta,\beta}(S_0)} \mathbb{E}_S \sqrt{2q_S(x_0)}\varphi_T |G_T| \mathbb{I}_{(\tau_H \le T)} = 0. \tag{14}$$

Applying Taylor's formula to q_S at the second order, one obtains

$$\sup_{S\in\mathcal{U}_{\delta,\beta}(S_0)} |m(\phi_h) - 2q_S(x_0)| = \sup_{S\in\mathcal{U}_{\delta,\beta}(S_0)} \frac{1}{2}\left|\int_{-1}^{1} u^2 h^2 \ddot{q}_S(x_0+\theta u h)du\right| \le C(L,M,B)h^2. \tag{15}$$

Then we have successively

$$\begin{aligned} |G_T| &\le \frac{1}{H_T}\int_{\tau_H}^{T} h\phi_h(X_t)|S(X_t)-S(x_0)|dt \le \frac{(L+B)h^2}{H_T}\left(\int_{t_0}^{T}\phi_h(X_t)dt - \frac{H_T}{h}\right) \\ &\le \frac{(L+B)h^2}{H_T}\left(\sqrt{T}\Delta_{t_0,T}(\phi_h) + (T-t_0)m(\phi_h) - (T-t_0)(2\tilde{q}_T(x_0)-\varepsilon_T)\right) \\ &\le \frac{(L+B)h^2}{H_T}\left(\sqrt{T}\Delta_{t_0,T}(\phi_h) + (T-t_0)|2\tilde{q}_T(x_0)-m(\phi_h)| + \varepsilon_T(T-t_0)\right). \end{aligned}$$

We finally get (14) from (15), inequality (A.1) and Lemma 3.2 in Galtchouk and Pergamenshchikov (2006).

Eventually since ξ_T is a Gaussian standard random variable, one has

$$\begin{aligned} &\left|\sqrt{2q_S(x_0)}\varphi_T \mathbb{E}_S \frac{|\xi_T|}{\sqrt{H_T}} - \mathbb{E}_S|\xi|\right| = \left|\frac{\sqrt{2q_S(x_0)}T^{\beta/(2\beta+1)}}{\sqrt{(T-t_0)(2\tilde{q}_T(x_0)-\varepsilon_T)h}} - 1\right|\mathbb{E}|\xi| \\ &\le \left|\frac{\sqrt{2q_S(x_0)}(1-\frac{t_0}{T})^{-1/2}}{\sqrt{2\tilde{q}_T(x_0)-\varepsilon_T}} - \frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}}\right|\mathbb{E}|\xi| + \left|\frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}} - 1\right|\mathbb{E}|\xi|. \end{aligned} \tag{16}$$

It is easy to show that the second part of (16) tends to 0 uniformly on $\mathcal{U}_{\delta,\beta}(S_0)$ as $T\to\infty$. Combining this with (9), (14) and Lemma A.1 in Galtchouk and Pergamenshchikov (2006) finishes the proof of theorem 2. □

References

Ait-Sahalia, Y.: Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach, *Econometrica* **70**(1) (2002), 223–262.

Banon, G.: Nonparametric identification for diffusion processes, *SIAM J. Control Optim.* **16**(3) (1978), 380–395.

Dalalyan, A.: Sharp adaptive estimation of the drift function for ergodic diffusions, *Ann. Statist.* **33**(6) (2005), 2507–2528.

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

Galtchouk, L. and Pergamenshchikov, S.: Sequential nonparametric adaptive estimation of the drift coefficient in diffusion processes, *Math. Methods Statist.* **10**(3) (2001), 316–330.

Galtchouk, L.and Pergamenshchikov, S.: Nonparametric sequential estimation of the drift in diffusion processes via model selection, *Math. Methods Statist.* **13**(1) (2004), 25–49.

Galtchouk, L. and Pergamenshchikov, S.: Nonparametric sequential minimax estimation of the drift coefficient in diffusion processes, *Sequential Anal.* **24**(3) (2005), 303–330.

Galtchouk, L. and Pergamenshchikov, S.: Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes, *Stat. Inference Stoch. Process.* **9**(1) (2006), 1–16.

Ibragimov, I. A. and Has'minskii, R. Z.: *Statistical Estimation: Asymptotic Theory*, Springer, Berlin, New York, 1981.

Jiang, G. J. and Knight, J. L.: A nonparametric approach to the estimation of diffusion processes with an application to a short-term interest rate model, *Econometric Theory* **13**(5) (1997), 615–645.

Liptser, R. S. and Shiryaev, A. N.: *Statistics of random processes.I and II*, Springer-Verlag, New York, 1978.

Pham D. T.: Nonparametric estimation of the drift coefficient in the diffusion equation, *Math. Operationsforsch. Statist. Ser. Statist.* **12**(1) (1981), 61–73.