МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Международная лаборатория статистики случайных
процессов и количественного финансового анализа

Международная научная конференция «Робастная статистика и финансовая математика – 2019»

(04–06 июля 2019 г.)

Сборник статей

Под редакцией
д-ра физ.-мат. наук, профессора С.М. Пергаменщикова,
канд. физ.-мат. наук, доцента Е.А. Пчелинцева

Томск
Издательский Дом Томского государственного университета
2019

Sharp Oracle Inequalities for Drift Estimation Problems from Discrete Data [1]

Galtchouk L., Pergamenshchikov S.

Université de Rouen Normandie, Rouen, France
Tomsk State University, Tomsk
e-mail: serge.pergamenchtchikov@univ-rouen.fr

Abstract

In this paper we consider the non parametric drift estimation problem for the ergodic diffusion processes on the basis of the observations in the fixed discrete time moments in the case when the diffusion coefficients are unknown. Using the truncated sequential point wise kernel estimators we develop a new model selection method for arbitrary orthonormal basis on the grid. For the constructed model selection procedure we show sharp oracle inequalities, i.e. we show, that it is optimal in the sharp oracle inequalities sense.

Keywords: adaptive non parametric drift estimation; discrete data; non asymptotic estimation; model selection; quadratic risk; sequential estimation; sharp oracle inequality.

In this paper we consider the diffusion process defined by the following stochastic process

$$\mathrm{d}y_t = S(y_t)\,\mathrm{d}t + b(y_t)\,\mathrm{d}W_t\,, \quad 0 \le t \le T\,, \tag{1}$$

where $(W_t)_{t\ge 0}$ is a standard Wiener process, the initial value y_0 is a given constant. The problem is to estimate the nonparametric drift $S(\cdot)$ on the discrete observations

$$(y_{t_j})_{1\le j\le N}\,, \quad t_j = j\delta\,, \tag{2}$$

where the frequency $\delta = \delta_T \in (0,1)$ is some function of T which will be specified later and the sample size $N = N(T) = [T/\delta]$. We denote here by $[a]$ the whole part of a. We consider this problem in the case when the diffusion coefficient $b(\cdot)$ is unknown, i.e. this is a nuisance parameter.

These models are very important for various fields in the theory of stochastic processes such that optimal control, identification problem, filtration, financial markets, insurance etc (see, for example, [9, 6, 5, 7, 8, 10]). Nonparametric estimation problems of the drift S were studied in a number of papers in the case of complete observations, i.e. when the whole trajectory $(y_t)_{0\le t\le T}$ is observed.

[1] This work is supported by the RSF Grant no. 17-11-01049.

To obtain the uniform ergodicity property for the process (1) we use the functional class introduced in [2], i.e. for any fixed $\mathbf{L} \geq 1$, $\mathbf{M} > 0$ and $\mathbf{x}_* > |\mathbf{x}_0| + |\mathbf{x}_1|$ we set

$$\Sigma_{\mathbf{L},\mathbf{M}} = \Bigg\{ S \in \mathbf{C}^1(\mathbb{R}) \; : \; \sup_{|x| \leq \mathbf{x}_*} \Big(|S(x)| + |\dot{S}(x)| \Big) \leq M \,,$$

$$-L \leq \inf_{|x| \geq \mathbf{x}_*} \dot{S}(x) \leq \sup_{|x| \geq \mathbf{x}_*} \dot{S}(x) \leq -1/L \Bigg\} \,. \tag{3}$$

Moreover, for some fixed parameters $0 < \mathbf{b}_{min} \leq \mathbf{b}_{max} < \infty$ we denote by $\mathcal{B}$ the class of the functions b from $\mathbf{C}^2(\mathbb{R})$ such that

$$\mathbf{b}_{min} \leq \inf_{x \in \mathbb{R}} |b(x)| \leq \sup_{x \in \mathbb{R}} \max \Big(|b(x)| \,, |\dot{b}(x)| \,, |\ddot{b}(x)| \Big) \leq \mathbf{b}_{max} \,. \tag{4}$$

Now we set

$$\Theta = \Sigma_{\mathbf{L},\mathbf{M}} \times \mathcal{B} = \{ (S, b) \; : \; S \in \Sigma \quad \text{and} \quad b \in \mathcal{B} \} \,. \tag{5}$$

It is easy to see that the functions from $\Sigma_{\mathbf{L},\mathbf{M}}$ are uniformly bounded on $[\mathbf{x}_0, \mathbf{x}_1]$, i.e.

$$s^* = \sup_{a \leq x \leq b} \sup_{S \in \Sigma_{\mathbf{L},\mathbf{M}}} S^2(x) \; < \infty \,. \tag{6}$$

It should be noted that, for any $\vartheta \in \Theta$, there exists an invariant density for the process (1) which is defined as

$$\mathbf{q}_\vartheta(x) = \left(\int_{\mathbb{R}} b^{-2}(z) \, e^{\widetilde{S}(z)} \mathrm{d}z \right)^{-1} b^{-2}(x) \, e^{\widetilde{S}(x)} \,, \tag{7}$$

where $\widetilde{S}(x) = 2 \int_0^x b^{-2}(v) S(v) \mathrm{d}v$ (see,e.g., [3], Ch.4, 18, Th2).
We need the following condition for the observation frequency.

$\mathbf{A}_1$) *Assume that that the frequency δ in the observations* (2) *is of the following form*

$$\delta = \delta_T = \frac{1}{(T+1) l_T} \,, \tag{8}$$

where the function l_T is such that,

$$\lim_{T \to \infty} \frac{l_T}{\ln T} = +\infty \,. \tag{9}$$

For example, we can take $l_T = (\ln T)^{1+\iota}$ for some $\iota > 0$.
In this paper for estimation problem we consider the quadratic risk defined as

$$\mathcal{R}(\widehat{S}, S) = \mathbf{E}_\vartheta \int_{\mathbf{x}_0}^{\mathbf{x}_1} |\widehat{S}(x) - S(x)|^2 \mathrm{d}x \,, \tag{10}$$

where $\mathbf{E}_\vartheta$ is the expectation with respect to the distribution of the process (1) with the parameters $\vartheta \in \Theta$.

Remark 1. *Note that we consider the problem estimation only for the drift function S, i.e. in this case the diffusion function b is considered as a nuisance parameter.*

To construct the model selection procedures first we define the sieve $(z_k)_{0\le k\le n}$ as

$$z_k = \mathbf{x}_0 + \frac{k(\mathbf{x}_1 - \mathbf{x}_0)}{n}\,, \quad 0 \le k \le n\,, \tag{11}$$

where $n = n(T)$ is an integer-valued function of T such that

$$\lim_{T\to\infty} \frac{n(T)}{\sqrt{T}} = 1\,. \tag{12}$$

Moreover, we fixe some basis $(\phi_j)_{j\ge1}$ in $\mathcal{L}_2([\mathbf{x}_0, \mathbf{x}_1])$ such that for any $1 \le i, j \le n$,

$$(\phi_i\,,\,\phi_j)_n = \frac{\mathbf{x}_1 - \mathbf{x}_0}{n} \sum_{l=1}^{n} \phi_i(z_l)\phi_j(z_l) = \mathbf{1}_{\{i=j\}}\,. \tag{13}$$

For example, we can take the trigonometric basis defined as $\mathrm{Tr}_1(x) \equiv 1/\sqrt{\mathbf{x}_1 - \mathbf{x}_0}$ and for $j \ge 2$

$$\mathrm{Tr}_j(x) = \sqrt{\frac{2}{\mathbf{x}_1 - \mathbf{x}_0}} \begin{cases} \cos(2\pi[j/2]\, l_0(x)) & \text{for even} \quad j\,; \\ \sin(2\pi[j/2] l_0(x)) & \text{for odd} \quad j\,, \end{cases} \tag{14}$$

where $[a]$ denotes integer part of a and $l_0(x) = (x - \mathbf{x}_0)/(\mathbf{x}_1 - \mathbf{x}_0)$. Note that if n is odd, then this basis is orthonormal for the empirical inner production, i.e. satisfies the property (13). By making use of this property we define the discrete Fourier representation for S on the sieve (11), i.e.,

$$S(z_k) = \sum_{j=1}^{n} \theta_{j,n}\phi_j(z_k)\,, \qquad 1 \le k \le n\,, \tag{15}$$

where

$$\theta_{j,n} = (S, \phi_j)_n = \frac{\mathbf{x}_1 - \mathbf{x}_0}{n} \sum_{l=1}^{n} S(z_l)\phi_j(z_l)\,.$$

To estimate $S(\cdot)$ we will use the point-wise sequential procedure $(\widehat{S}_k, \tau_k)$ for $S(z_k)$ proposed in [2]. To this end we set

$$Y_k = \widehat{S}_k \chi_{\{\tau_k \le N\}}\,.$$

We remind that τ_k is the random number (stoping time) of observations used for the estimation of $S(z_k)$. Note that on the set

$$\Gamma = \cap_{k=1}^{n} \{\tau_k \le N\}$$

we can write the regression model, i.e.

$$Y_k = S(z_k) + \sigma(z_k)\xi_k + \mathbf{g}_k\,, \tag{16}$$

where $(\xi_k)_{1\le k\le n}$ is the i.i.d. Gaussian $(0, 1)$ sequence and the sequence $\mathbf{g}_k$ is the sufficiently quick vanish sequence when $T \to \infty$. Moreover, using the regression model (16) we estimate the Fourrier coefficients for S as

$$\widehat{\theta}_{j,n} = (Y, \phi_j)_n = \frac{\mathbf{x}_1 - \mathbf{x}_0}{n} \sum_{l=1}^{n} Y_l \phi_j(z_l)\,. \tag{17}$$

We estimate the values $(S(z_k))_{1\le k\le n}$ by the weighted least squares estimators

$$\widehat{S}_\lambda(z_k) = \sum_{j=1}^{n} \lambda(j)\widehat{\theta}_{j,n}\phi_j(z_k)\,, \qquad 1 \le k \le n\,, \tag{18}$$

where the weight vector $\lambda = (\lambda(1), \ldots, \lambda(n))'$ belongs to some finite set Λ from $[0,1]^n$. In the sequel we denote by ν the cardinal number of the set Λ, $\nu = \mathrm{card}(\Lambda)$, which is a function of T, i.e. $\nu = \nu_T$. Moreover, we set the norm for Λ as

$$\Lambda_* = \max_{\lambda\in\Lambda} \sum_{j=1}^{n} \lambda(j) \tag{19}$$

which can be a function of T, i.e. $\Lambda_* = \Lambda_*(T)$. We assume that the basis functions and the weights set Λ satisfy the following condition.

$\mathbf{A}_2$)*For any* $a > 0$

$$\lim_{T\to\infty} \frac{\phi_T^* + \nu_T}{T^a} = 0 \quad \textit{and} \quad \lim_{T\to\infty} \frac{\Lambda_*(T)}{T^{1/6+a}} = 0\,, \tag{20}$$

where $\phi_T^* = \max_{1\le j\le n} \max_{\mathbf{x}_0\le x\le \mathbf{x}_1} |\phi_j(x)|$.

Now one needs to choose a cost function in order to define an optimal weight $\lambda \in \Lambda$. A best candidate for the cost function should be the empirical squared error given by the relation

$$\mathrm{Err}_n(\lambda) = \|\widehat{S}_\lambda - S\|_n^2 \to \min\,.$$

In our case, the empirical squared error is equal to

$$\mathrm{Err}_n(\lambda) \;=\; \sum_{j=1}^{n} \lambda^2(j)\widehat{\theta}_{j,n}^2 \;-2\sum_{j=1}^{n} \lambda(j)\widehat{\theta}_{j,n}\,\theta_{j,n} \;+\; \sum_{j=1}^{n} \theta_{j,n}^2\,. \tag{21}$$

Since coefficients $\theta_{j,n}$ are unknown, we need to replace the term $\widehat{\theta}_{j,n}\,\theta_{j,n}$ by some estimator which we choose as

$$\widetilde{\theta}_{j,n} = \widehat{\theta}_{j,n}^2 - \frac{\mathbf{x}_1 - \mathbf{x}_0}{n}\widehat{\sigma}_{j,n} \quad \text{and} \quad \widehat{\sigma}_{j,n} = \frac{\mathbf{x}_1 - \mathbf{x}_0}{n}\sum_{l=1}^{n} \widehat{\sigma}_l \phi_j^2(z_l)\,, \tag{22}$$

where $\widehat{\sigma}_l$ is the estimator for the coefficient σ_l^2 defined through a special sequential modification for the quadratic characteristic proposed in [4]. Now, to penalize the replacement the term $\theta_{j,n}\widehat{\theta}_{j,n}$ with the estimator (22) we are needed the additinal positive term defined as

$$\widehat{P}(\lambda) = \frac{\mathbf{x}_1 - \mathbf{x}_0}{n}\sum_{j=1}^{n} \lambda^2(j)\widehat{\sigma}_{j,n}\,. \tag{23}$$

Finally, we use the following cost function

$$J_n(\lambda) \;=\; \sum_{j=1}^{n} \lambda^2(j)\widehat{\theta}_{j,n}^2 \;-2\sum_{j=1}^{n} \lambda(j)\widetilde{\theta}_{j,n} \;+\; \rho\,\widehat{P}_n(\lambda)\,, \tag{24}$$

where the positive coefficient $0 < \rho < 1$ will be specified later. We define the model selection procedure as

$$\widehat{\lambda} = \mathrm{argmin}_{\lambda\in\Lambda} J_n(\lambda) \quad \text{and} \quad \widehat{S}_* = \widehat{S}_{\widehat{\lambda}}\,. \tag{25}$$

Theorem 1. *Assume that the conditions* $\mathbf{A}_1) - \mathbf{A}_2)$ *hold. Then, there exists* $\mathbf{c}^* > 0$ *such that for any* $0 < \rho < 1/6$*, the estimation procedure* $\widehat{S}_*$ *defined in* (25) *satisfies the following inequality*

$$\mathcal{R}(\widehat{S}_*, S) \leq \frac{(1+\rho)^2(1+4\rho)}{1-3\rho} \min_{\lambda \in \Lambda} R_n(\widehat{S}_\lambda, S) + \frac{\Psi_T}{\rho} \tag{26}$$

where the term Ψ_T *is such that for any* $a > 0$

$$\lim_{T \to \infty} \frac{\Psi_T}{T^a} = 0\,. \tag{27}$$

Remark 2. *It should be noted that similarly to [1] we will use the inequality* (26) *to provide the efficiency property in adaptive setting, i.e. through the sharp oracle inequality without using the regularity properties of unknown function we can estimate from above the risk for the model selection procedure by the risk effective value corresponding to the Sobolev ball.*

References

[1] Galtchouk L., Pergamenshchikov S. Adaptive sequential estimation for ergodic diffusion processes in quadratic metric // Journal of Nonparametric Statistics. 2011. V. 23, No 2. P. 255–285.

[2] Galtchouk L., Pergamenshchikov S. Efficient pointwise estimation based on discrete data in ergodic nonparametric diffusions // Bernoulli. 2015. V. 21, No 4. P. 2569–2594.

[3] Gihkman I.I., Skorohod A.V. Stochastic differential equations. Kiev: Naukova Dumka; P.I. - 1968.

[4] Jacod J. Non-parametric kernel estimation of the coefficient of a diffusion // Scand. J. Statist. 2000. V. 27, No 1. P. 83–96.

[5] Kabanov Yu.M., Pergamenshchikov S.M. Two Scale Stochastic Systems: Asymptotic Analysis and Control. - Berlin, New York: Springer-Verlag Applications of Mathematics, Stochastic Modeling and Applied Probability, 49. P.I. - 2003.

[6] Karatzas I., Shreve S.E. Methods of Mathematical Finance. - New York: Springer; P.I. - 1998.

[7] Kutoyants Yu. A. Identification of Dynamical Systems with Small Noise. - Dordrecht: Kluwer; P.I. - 1994, 298p.

[8] Kutoyants Yu.A. Statistical Inferences for Ergodic Diffusion Processes. - London: Springer Series in Statistics; P.I. - 2004, 496 p.

[9] Lamberton D., Lapeyre B. Introduction to stochastic calculus applied to finance. - London: Chapman & Hall; P.I. - 1996.

[10] Liptser R.Sh., Shiryaev A.N. Statistics of a random process. - Berlin, Heidelberg, New York: Springer-Verlag, v.I, II; P.I. - 1978.

Гальчук Л., Пергаменщиков С. (Руанский университет, Руан, Франция, Томский государственный университет, Томск, 2019)
Точные оракульные неравенства для задач оценивания сноса по дискретным данным.

Аннотация. В данной работе рассматривается задача оценивания непараметрического сноса для эргодических диффузионных процессов на основе наблюдений в фиксированные дискретные моменты времени в случае, когда коэффициенты диффузии неизвестны. Используя усеченные последовательные точечные ядерные оценки, разрабатывается новый метод выбора модели для произвольного ортонормированного базиса на сетке. Для построенной процедуры выбора модели получены точные оракульные неравенства, т.е. показывается, что она оптимальна в смысле точных оракульных неравенств.

Ключевые слова: адаптивное непараметрическое оценивание; дискретные данные; неасимптотическое оценивание; выбор модели; квадратичный риск; последовательное оценивание; точные оракульные неравенства.