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 – 2018»

(09–11 июля 2018 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
2018

Sharp oracle inequalities for the nonparametric signal estimation in the Lévy regression model *

Beltaief S., Chernoyarov O., Pergamenshchikov S.

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

Abstract

We develop a new model selection method for the adaptive robust efficient nonparametric signal estimation observed with impulse noise defined by a general non Gaussian Lévy processes. On the basis of the developed method we construct the estimation procedures which are analyzed in two settings: in non asymptotic and asymptotic ones. For the first time for such models we show non asymptotic sharp oracle inequalities for the quadratic and for the robust risks, i.e. we show that the constructed procedures are optimal in the sharp oracle inequalities sense. To this end we develop a new mathematical tool for the Lévy regression models in continuous time.

Keywords: nonparametric estimation, model selection, non-asymptotic estimation, robust estimation, oracle inequalities, statistical signal processing techniques and analysis.

In this paper we consider the signal estimation problem on the basis of the observations defined by the nonparametric regression model in continuous time with noises of small intensity, i.e.

$$\mathrm{d}\, y_t = S(t)\mathrm{d}\, t + \varepsilon\, \mathrm{d}\xi_t\,, \quad 0 \le t \le 1\,, \tag{1}$$

where $S(\cdot)$ is an unknown deterministic signal (i.e. $[0,1] \to \mathbb{R}$ nonrandom function), $(\xi_t)_{0\le t\le 1}$ is an unobserved noise and $\varepsilon > 0$ is the noise intensity. We assume that the noise process $(\xi_t)_{0\le t\le 1}$ is defined as

$$\xi_t = \varrho_1 w_t + \varrho_2 z_t \quad \text{and} \quad z_t = x * (\mu - \widetilde{\mu})_t\,, \tag{2}$$

where, ϱ_1 and ϱ_2 are some unknown constants, $(w_t)_{0\le t\le 1}$ is a standard brownian motion, "*" denotes the stochastic integral with respect to the compensated jump measure (see, for example in Jacod and Shiryaev [5] or Cont and Tankov [1] for details), $\mu(\mathrm{d}s\,\mathrm{d}x)$ is a jump measure with

*This work was supported by the RSF grant 17-11-01049 (National Research Tomsk State University).

deterministic compensator $\widetilde{\mu}(\mathrm{d}s\,\mathrm{d}x) = \mathrm{d}s\Pi(\mathrm{d}x)$, $\Pi(\cdot)$ is the unknown Lévy measure, i.e. some positive measure on $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$, such that

$$\Pi(x^2) = 1 \quad \text{and} \quad \Pi(x^4) < \infty\,, \tag{3}$$

where $\Pi(|x|^m) = \int_{\mathbb{R}_*} |z|^m \,\Pi(\mathrm{d}z)$. Note that the measure $\Pi(\mathbb{R}_*)$ could be equal to $+\infty$. In the sequel we will denote by Q the distribution of the process $(\xi_t)_{0\le t\le 1}$. We assume that the parameters ϱ_1 and ϱ_2 satisfy the conditions

$$0 < \check{\varrho}_\varepsilon \le \varrho_1^2 \quad \text{and} \quad \varkappa_Q = \varrho_1^2 + \varrho_2^2 \le \varsigma_\varepsilon^*\,, \tag{4}$$

where the bounds $\check{\varrho}_\varepsilon$ and ς_ε^* are such that for any $\mathbf{b} > 0$

$$\liminf_{\varepsilon\to 0} \varepsilon^{-\mathbf{b}}\, \check{\varrho}_\varepsilon > 0 \quad \text{and} \quad \lim_{\varepsilon\to 0} \varepsilon^{\mathbf{b}}\, \varsigma_\varepsilon^* = 0\,. \tag{5}$$

We denote by $\mathcal{Q}_\varepsilon^*$ the family of all distributions Q of the process (2) in the Skorokhod space $\mathbf{D}[0,1]$ for which the conditions (4) and (5) hold.

The problem is to estimate the function S on the observations $(y_t)_{0\le t\le 1}$ when $\varepsilon \to 0$. Note that if $(\xi_t)_{0\le t\le 1}$ is the brownian motion, then we obtain the "signal+white noise" model which is very popular in statistical radio-physics and is well studied by many authors: Ibragimov and Khasminskii in [4], Pinsker in [11], Kutoyants in [10], Trifonov, Kharin, Chernoyarov and Kalashnikov in [12] and etc. The condition $\varepsilon \to 0$ means that the signal/noise ration goes to the infinity. In this paper we assume that in addition to the intrinsic noise in the radio-electronic system, approximated usually by the gaussian white noise, the useful signal S is distorted by the impulse noise flow defined by the Lévy process with the jumps introduced in the next section. The cause of the appearance of a pulse stream in the radio-electronic systems can be, for example, either external unintended (atmospheric) noises or intentional impulse noises or the errors in the demodulation and in the channel decoding for the binary information symbols. Note that, for the first time the impulse noises for the signal detection problems have been introduced by Kassam in [6] on the basis of the compound Poisson processes. Later, Konev and Pergamenshchikov used the compound Poisson processes in [8,9] for the nonparametric signal estimation problems. However, the compound Poisson process can describe only the large impulses influence of small frequencies. It should be noted that in the telecommunication systems the noise impulses are without limitations on frequencies and, therefore, the compound Poisson models are too restricted for the practical applications. To include the all possible impulse noises we propose to use a general non-gaussian Lévy processes in the observation model (1). In this paper we consider the nonparametric estimation problem in the adaptive setting, i.e. when the regularity of the signal S is unknown. Moreover, we also assume that the distribution Q of the noise process $(\xi_t)_{0\le t\le 1}$ is unknown. It is known only that this distribution belongs to the distribution family $\mathcal{Q}^*_\varepsilon$ defined in the next section. By these reasons we use the robust estimation approach proposed for nonparametric problems by Galtchouk, Konev and Pergamenshchikov in [2, 8, 9]. We set the robust risks as

$$\mathcal{R}^*_\varepsilon(\widehat{S}_\varepsilon, S) = \sup_{Q\in\mathcal{Q}^*_\varepsilon} \mathcal{R}_Q(\widehat{S}_\varepsilon, S) \tag{6}$$

where $\widehat{S}_\varepsilon$ is an estimator (i.e. any measurable function of $(y_t)_{0\le t\le 1}$) and

$$\mathcal{R}_Q(\widehat{S}_\varepsilon, S) := \mathbf{E}_{Q,S}\,\|\widehat{S}_\varepsilon - S\|^2 \quad \text{and} \quad \|S\|^2 = \int_0^1 S^2(t)\mathrm{d}t\,. \tag{7}$$

In this paper we develop a sharp model selection method for the estimating the unknown signal S. The interest to such statistical procedures can be explained by the fact that they provide adaptive solutions for the nonparametric estimation through the non-asymptotic oracle inequalities which give the non-asymptotic upper bound for the quadratic risk including the minimal risk over chosen family of estimators with some coefficient closed to one. Such inequalities were obtained, for example, by Galtchouk and Pergamenshchikov [3] for non Gaussian regression models in discrete time and by Konev and Pergamenshchikov [7] for general regression semimartingale models in continuous time, i.e. when the observation process is given by the following stochastic differential equation

$$\mathrm{d}x_t = S(t)\mathrm{d}t + \mathrm{d}\eta_t\,, \quad 0\le t\le n\,, \quad (n\to\infty)\,, \tag{8}$$

where S is an unknown 1 - periodic signal and the unobserved noise $(\eta_t)_{t\ge 0}$ is square integrated semi-martingale. Note, that for any $0 < t < 1$, setting $\check{x}_t = n^{-1}\sum_{j=1}^n (x_{t+j} - x_j)$, we can represent this model as a model with a small parameter of form (1)

$$\mathrm{d}\check{x}_t = S(t)\mathrm{d}t + \varepsilon\,\mathrm{d}\check{\eta}_t\,, \tag{9}$$

where $\varepsilon = n^{-1/2}$ and $\check{\eta}_t = n^{-1/2}\sum_{j=0}^{n-1}(\eta_{t+j} - \eta_j)$. If $(\eta_t)_{t\ge 0}$ is the Lévy process, then the $\check{\eta}_t$ is the Lévy process as well. But the main difference between the models (1) and (9) is that the jumps in the last one are small, i.e.

$$\Delta\check{\eta}_t = \check{\eta}_t - \check{\eta}_{t-} = \mathrm{O}(n^{-1/2}) = \mathrm{O}(\varepsilon) \quad \text{as} \quad \varepsilon\to 0\,. \tag{10}$$

But there is not such property in the model (1). It should be noted that the property (10) is crucial in the non asymptotic analysis for the observations on the large time intervals, i.e. the methods developed for the model (8) can not be used for the problem (1).

The main goal of this paper is to develop a new model selection method for the adaptive signal estimation problem in the nonparametric regression (1) as $\varepsilon \to 0$.

We estimate the function $S(x)$ for $x \in [0,1]$ by the weighted least squares estimator

$$\widehat{S}_\lambda(x) = \sum_{j=1}^{n} \lambda(j) \widehat{\theta}_{j,\varepsilon} \phi_j(x) \,, \qquad (11)$$

where $n = [1/\varepsilon^2]$, the weights $\lambda = (\lambda(j))_{1\le j\le n}$ belong to some finite set Λ from $[0,1]^n$, $\widehat{\theta}_{j,\varepsilon}$ are estimators for the Fourier coefficients and ϕ_j are basis functions. For example, we can take the trigonometric basis defined as $\mathrm{Tr}_1 \equiv 1$ and for $j \ge 2$

$$\mathrm{Tr}_j(x) = \sqrt{2} \begin{cases} \cos(2\pi[j/2]x) & \text{for even} \quad j\,; \\ \\ \sin(2\pi[j/2]x) & \text{for odd} \quad j\,. \end{cases} \qquad (12)$$

Now we set

$$\iota = \mathrm{card}(\Lambda) \quad \text{and} \quad |\Lambda|_* = \max_{\lambda\in\Lambda} \sum_{j=1}^{n} \mathbf{1}_{\{\lambda_j > 0\}} \,, \qquad (13)$$

where $\mathrm{card}(\Lambda)$ is the number of the vectors in Λ. In the sequel we assume that ι is a function of $\varepsilon > 0$, i.e. $\iota = \iota(\varepsilon)$, such that for any $\mathbf{b} > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{\mathbf{b}} \iota(\varepsilon) = 0 \,. \qquad (14)$$

By the same way as in [8] we introduce the cost function

$$J_\varepsilon(\lambda) = \sum_{j=1}^{n} \lambda^2(j) \widehat{\theta}^2_{j,\varepsilon} - 2\sum_{j=1}^{n} \lambda(j) \widetilde{\theta}_{j,\varepsilon} + \delta\, \widehat{P}_\varepsilon(\lambda) \,, \qquad (15)$$

where $\widetilde{\theta}_{j,\varepsilon} = \widehat{\theta}^2_{j,\varepsilon} - \varepsilon^2 \widehat{\varkappa}_\varepsilon$ and $\widehat{\varkappa}_\varepsilon$ is an estimator for the Fourier coefficient variance and the penalty term

$$\widehat{P}_\varepsilon(\lambda) = \varepsilon^2 \widehat{\varkappa}_\varepsilon |\lambda|^2 \quad \text{and} \quad |\lambda|^2 = \sum_{j=1}^{n} \lambda_j^2 \,.$$

We define the model selection procedure as

$$\widehat{S}_* = \widehat{S}_{\hat{\lambda}} \quad \text{and} \quad \widehat{\lambda} = \mathrm{argmin}_{\lambda\in\Lambda} J_\varepsilon(\lambda) \,. \qquad (16)$$

We recall that the set Λ is finite so $\hat{\lambda}$ exists. In the case when $\hat{\lambda}$ is not unique we take one of them. To estimate the variance of the Fourier coefficient we use the approach proposed by Konev and Pergamenshchikov in [8,9] by making use of the properties of the trigonometric basis (12) which provide a sharp upper bound for the convergence rate for the series of the squares of its Fourier coefficients. We set

$$\widehat{\varkappa}_\varepsilon = \sum_{j=[1/\varepsilon]+1}^{n} \widehat{\tau}^2_{j,\varepsilon}\,, \quad n = [1/\varepsilon^2]\,, \tag{17}$$

where $\widehat{\tau}_{j,\varepsilon}$ are the estimators for the Fourrier coefficients τ_j with respect to the trigonometric basis (12), i.e.

$$\widehat{\tau}_{j,\varepsilon} = \int_0^1 \mathrm{Tr}_j(t)\mathrm{d}\check{y}_t \quad \text{and} \quad \tau_j = \int_0^1 S(t)\mathrm{Tr}_j(t)\mathrm{d}t\,. \tag{18}$$

Now, we specify the weight coefficients $(\lambda(j))_{1\le j\le n}$. Consider a numerical grid of the form

$$\mathcal{A} = \{1,\ldots,k^*\} \times \{r_1,\ldots,r_{\mathbf{m}}\}\,, \tag{19}$$

where $r_i = i\,\varpi$ and $\mathbf{m} = [1/\varpi^2]$. We assume that both the parameters $k^* \ge 1$ and $0 < \varpi < 1$ are functions of ε, i.e. $k^* = k^*_\varepsilon$ and $\varpi = \varpi_\varepsilon$ such that

$$\lim_{\varepsilon\to 0}\left(\frac{1}{k^*_\varepsilon} + \frac{k^*_\varepsilon}{|\ln\varepsilon|}\right) = 0 \quad \text{and} \quad \lim_{\varepsilon\to 0}\left(\varpi_\varepsilon + \frac{\varepsilon^{\mathbf{b}}}{\varpi_\varepsilon}\right) = 0 \tag{20}$$

for any $\mathbf{b} > 0$. One can take, for example, for $0 < \varepsilon < 1$

$$\varpi_\varepsilon = |\ln\varepsilon|^{-1} \quad \text{and} \quad k^*_\varepsilon = k^*_0 + \sqrt{|\ln\varepsilon|}\,, \tag{21}$$

where $k^*_0 \ge 0$ is some fixed constant. For each $\alpha = (\beta, r) \in \mathcal{A}$, we introduce the weights $\lambda_\alpha = (\lambda_\alpha(j))_{1\le j\le n}$ from $\mathbb{R}^n$ as

$$\lambda_\alpha(j) = \mathbf{1}_{\{1\le j<j_*\}} + \left(1 - (j/\omega_\alpha)^\beta\right)\,\mathbf{1}_{\{j_*\le j\le\omega_\alpha\}}\,, \tag{22}$$

where $j_* = j_*(\alpha) = [\omega_\alpha/|\ln\varepsilon|]$, $\omega_\alpha = \mathrm{d}_\beta\,(r\,\upsilon_\varepsilon)^{1/(2\beta+1)}$,

$$\mathrm{d}_\beta = \left(\frac{(\beta+1)(2\beta+1)}{\pi^{2\beta}\beta}\right)^{1/(2\beta+1)}\,, \qquad \upsilon_\varepsilon = \frac{1}{\varepsilon^2\,\varsigma^*_\varepsilon} \tag{23}$$

and the threshold ς^*_ε is introduced in (4). Now we define the set Λ as

$$\Lambda = \{\lambda_\alpha\,,\,\alpha\in\mathcal{A}\}\,. \tag{24}$$

Note that in this case $\iota = k^*\mathbf{m}$ and the conditions (20) imply directly the property (14). Moreover, from (22) we find that for any $\alpha \in \mathcal{A}$

$$\sum_{j=1}^{n} \lambda_\alpha(j) \le \omega_\alpha \le d_*\,r_{\mathbf{m}}^{1/3}\,\upsilon_\varepsilon^{1/3} \quad \text{and} \quad d_* = \sup_{\beta\ge 1}\mathrm{d}_\beta\,.$$

Therefore, the conditions (20) imply that for any $\mathbf{b} > 0$

$$\lim_{\varepsilon\to 0}\varepsilon^{2/3+\mathbf{b}}|\Lambda|_* = 0\,. \tag{25}$$

Remark 1. *It should be emphasized that the weight coefficients defined by the set* (24) *are used by Konev and Pergamenshchikov in [8, 9] for continuous time regression models to show the asymptotic efficiency.*

Theorem 1. *Assume that for the model* (1) *the condition* (3) *holds and the unknown function* $S(\cdot)$ *is continuously differentiable. Then for* *any* $0 < \delta < 1/6$ *and for any* $\varepsilon > 0$ *for which* $|\Lambda|_* \le 1/\varepsilon$*, the robust risks for the procedure* (16) *satisfies the following oracle inequality*

$$\mathcal{R}^*_\varepsilon(\widehat{S}_*, S) \le \frac{1+3\delta}{1-3\delta} \min_{\lambda\in\Lambda} \mathcal{R}^*_\varepsilon(\widehat{S}_\lambda, S) + \varepsilon^2 \frac{\mathbf{U}^*_\varepsilon(S)}{\delta}\,, \qquad (26)$$

where the term $\mathbf{U}^*_\varepsilon(S) > 0$ *is such that for any* $r > 0$ *and* $\mathbf{b} > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{\mathbf{b}} \sup_{\|\dot{S}\|\le r} \mathbf{U}^*_\varepsilon(S) = 0\,. \qquad (27)$$

Now taking into account the property (25) we can deduce the following theorem for the procedure (16) with the weight coefficients (24).

Theorem 2. *Assume that for the model* (1) *the condition* (3) *holds. Then the model selection procedure* (16) *constructed through the weight coefficients* (24) *with the conditions* (20) *satisfies the oracle inequality* (26) *with the property* (27).

Remark 2. *Note that the similar sharp oracle inequalities were obtained in the papers [3] and [8] for the model selection procedures based on the trigonometric basis functions. In this paper we obtain these inequalities for the model selection procedures based on any arbitrary orthogonal basic function in* $\mathbf{L}_2[0,1]$. *We use the trigonometric functions only for the estimator* (17).

References

1. Cont R., Tankov P. Financial Modelling with Jump Processes. Chapman & Hall - 2004.

2. Galtchouk, L.I. and Pergamenshchikov, S.M. Asymptotically efficient estimates for non parametric regression models. // Statistics and Probability Letters. 2006. V. 76. No. 8. P. 852–860.

3. Galtchouk, L.I. and Pergamenshchikov, S. M. Sharp non-asymptotic oracle inequalities for nonparametric heteroscedastic regression models.// Journal of Nonparametric Statistics. 2009. V. 21. No. 1. P. 1 - 16.

4. Ibragimov I. A. and Khasminskii R. Z. Statistical Estimation: Asymptotic Theory, Springer, New York -1981.

5. Jacod J., Shiryaev A.N. Limit theorems for stochastic processes, 2nd edition, Springer, Berlin -2002.

6. Kassam, S.A. Signal detection in non - Gaussian noise, New York: Springer-Verlag Inc., IX, -1988.

7. Konev V. V. and Pergamenshchikov S. M. Nonparametric estimation in a semimartingale regression model. Part 1. Oracle Inequal-

ities. // Journal of Mathematics and Mechanics of Tomsk State University. 2009. V. 3. P. 23–41.

8. Konev V. V. and Pergamenshchikov S. M. Efficient robust nonparametric estimation in a semimartingale regression model.//Ann. Inst. Henri Poincaré Probab. Stat. 2012. V. 48. No. 4. P. 1217–1244.

9. Konev V. V. and Pergamenshchikov S. M. Robust model selection for a semimartingale continuous time regression from discrete data.// Stochastic processes and their applications. 2015. V. 125. P. 294 – 326.

10. Kutoyants, Yu. A. Identification of dynamical systems with small noise. Mathematics and its Applications, 300, Kluwer Academic Publishers Group, Dordrecht - 1994.

11. Pinsker, M.S. Optimal filtration of square integrable signals in gaussian white noise.// Problems of Transimission information. 1981. V. 17. P. 120–133.

12. Trifonov A.P., Kharin A.V., Chernoyarov O.V., Kalashnikov K. S. Determining the number of radio signals with unknown phases. // International Journal on Communications Antenna and Propagation. 2015. V. 5. No. 6. P. 367–374.

Белтайеф С., Чернояров О. В., Пергаменщиков С. М. (Руанский университет, Руан, Национальный исследовательский университет "МЭИ" , Москва, Томский государственный университет, Томск, 2018) **Точные оракульные неравенства для оценивания непараметрического сигнала в модели регрессии Леви.**

Аннотация. Разрабатывается новый метод выбора модели для адаптивного робастного эффективного оценивания непараметрического сигнала, наблюдаемого с импульсным шумом, определяемым общими негауссовскими процессами Леви. На основе разработанного метода строятся процедуры оценивания, которые анализируются в неасимптотическом и асимптотическом случаях. Впервые для таких моделей получены неасимптотические точные оракульные неравенства для квадратичного и для нробастного рисков, т.е. устанавливается, что построенные процедуры оптимальны в смысле оракульных неравенств. Разрабатывается новый математический инструментарий для анализа регрессионных моделей Леви в непрерывном времени.

Ключевые слова: непараметрическое оценивание, выбор модели, неасимптотическое оценивание, оракульные неравенства, статистические методы обработки сигналов.