Adaptive Robust Efficient Methods for Periodic Signal Processing Observed with Colours Noises

Evgeny PCHELINTSEV [1,2], *Serguei PERGAMENSHCHIKOV* [3,4], *Mariana MARCOKOVA* [5,6]

[1]Department of Mathematical Analysis and Theory of Functions, Faculty of Mechanics and Mathematics, Tomsk State University, Lenin Avenue 36, 634050 Tomsk, Russia
[2]International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, Lenin Avenue 36, 634050 Tomsk, Russia
[3]Laboratory of Mathematics Raphael Salem, University of Rouen Normandie, Avenue de l'Universite 12, 76801 Saint Etienne du Rouvray, France
[4]International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, Lenin Avenue 36, 634050 Tomsk, Russia
[5]Department of Structural Mechanics and Applied Mathematics, Faculty of Civil Engineering, University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovak Republic
[6]Department of Electronics and Nanoelectronics, National Research University "Moscow Power Engineering Institute", Krasnokazarmennaya 14, 111250 Moscow, Russia

evgen-pch@yandex.ru, serge.pergamenchtchikov@univ-rouen.fr, mariana.marcokova@fstav.uniza.sk

DOI: 10.15598/aeee.v17i3.3132

Abstract. *In this paper, we consider the problem of robust adaptive efficient estimating a periodic signal observed in the transmission channel with the dependent noise defined by non-Gaussian Ornstein-Uhlenbeck processes with unknown correlation properties. Adaptive model selection procedures, based on the shrinkage weighted least squares estimates, are proposed. The comparison between shrinkage and least squares methods is studied and the advantages of the shrinkage methods are analyzed. Estimation properties for proposed statistical algorithms are studied on the basis of the robust mean square accuracy defined as the maximum mean square estimation error over all possible values of unknown noise parameters. Sharp oracle inequalities for the robust risks have been obtained. The robust efficiency of the model selection procedure has been established.*

Keywords

Asymptotic efficiency, model selection, nonparametric regression, Ornstein-Uhlenbeck process, periodic signals, robust quadratic risk, sharp oracle inequality, shrinkage estimation, weighted least squares estimates.

1. Introduction

In this paper, we consider the estimation problem for the 1-periodic signal $S(t)$ on the basis of observations $(y_t)_{0\le t\le n}$ given by the stochastic differential equation:

$$dy_t = S(t)dt + d\xi_t, \quad 0 \le t \le n, \tag{1}$$

where n is the duration of observation and $(\xi_t)_{0\le t\le n}$ is unobserved colour noise. Note that if $(\xi_t)_{0\le t\le n}$ is Brownian motion, then we obtain the well-known "signal + white noise" model which is very popular in statistical radio-physics (see, for example, [1], [2] and [3]). In this paper, we assume that the useful signal S is distorted by the impulse flow described by the non-Gaussian Ornstein-Uhlenbeck processes, which allows studying the signal estimation problems with dependent pulse noises, i.e. we assume that the noise process $(\xi_t)_{0\le t\le n}$ obeys the equation:

$$d\xi_t = a\xi_t dt + du_t, \tag{2}$$
$$u_t = \acute{n}_1 w_t + \acute{n}_2 z_t \quad \text{and} \quad z_t = x * (\mu - \tilde{\mu})_t\,,$$

where a, $\acute{n}_1$ and $\acute{n}_2$ are some unknown constants, $(w_t)_{t\ge0}$ is the standard Brownian motion, $\mu(ds\,dx)$ is the jump measure with deterministic compensator $\tilde{\mu}(ds\,dx) = ds\Pi(dx)$, $\Pi(\cdot)$ is the Levy measure, i.e. some positive measure on $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$, such that $\Pi(x^2) = 1$ and $\Pi(x^6) < \infty$. Here we use the notation $\Pi(|x|^m) = \int_{\mathbb{R}^*} |y|^m \Pi(dy)$. Note that the Levy measure $\Pi(\mathbb{R}^*)$ could be equal to $+\infty$. We use $*$ for the stochastic integrals with respect to random measures (see [4], Chapters 2 and 3), i.e.:

$$x * (\mu - \tilde{\mu})_t = \int_0^t \int_{\mathbb{R}^*} y (\mu - \tilde{\mu}) (ds, dy) . \tag{3}$$

It should be noted that if $a = 0$, then we obtain the Levy regression model considered in [5]. In the case when $\Pi(\cdot) = 0$ we obtain the well-known Gaussian Ornstein-Uhlenbeck regression model introduced in [6] and [7]. The model in the Eq. (1) and Eq. (2) in which the jump process $(z_t)_{t\geq 0}$ is defined by the compound Poisson process was studied in [8] and [9]. However, the compound Poisson processes can describe only the large noise impulses of small fixed frequency, but the telecommunication and location systems may have the impulse noises with any frequency without any condition. We note that in the papers [8] and [9] the proposed statistical procedures are based on the classical weighted least squares estimators.

The main goal of this paper is to develop a new improved adaptive robust efficient signal estimation methods for the non-Gaussian Ornstein-Uhlenbeck noise $(\xi_t)_{0\leq t\leq n}$ based on the general Levy processes with unknown distribution Q. We assume that this distribution belongs to the class Q_n^* defined as a family of all these distributions for which the parameters $-a_* \leq a < 0$, $\acute{n}_1 \geq \xi_*$ and $\acute{n}_1^2 + \acute{n}_2^2 \leq \xi^*$, where a_*, ξ_* and ξ^* are some fixed positive bounds. The quality of an estimate $\hat{S}_n$ of the unknown signal S, i.e. some function of $(y_t)_{0\leq t\leq n}$, will be measured with the robust quadratic risk:

$$R^*\left(\hat{S}_n, S\right) = \sup_{Q\in Q_n^*} R_Q\left(\hat{S}_n, S\right), \tag{4}$$

where

$$R_Q\left(\hat{S}_n, S\right) := E_{Q,S}\left\|\hat{S}_n - S\right\|^2$$
$$\text{and} \tag{5}$$
$$\|S\|^2 = \int_0^1 S^2(t)dt.$$

Here $E_{Q,S}$ is the expectation with respect to the distribution $P_{Q,S}$ of the process in the Eq. (1) with a fixed distribution Q of the noise $(\xi_t)_{0\leq t\leq n}$ and a given function S.

2. Shrinkage Estimation Methods

Let $(\phi_j)_{j\geq 1}$ be a trigonometric basis in $L_2[0, 1]$. We extend these functions by the periodic way on $\mathbb{R}$ i.e. $\phi_j(t) = \phi_j(t + 1)$ for any $t \in \mathbb{R}$. For estimating the unknown function S in the Eq. (1), we consider its Fourier expansion:

$$S(t) = \sum_{j=1}^{\infty} \theta_j \phi_j(t)$$
$$\text{and} \tag{6}$$
$$\theta_j = (S, \phi_j) = \int_0^1 S(t)\phi_j(t)dt.$$

The Fourier coefficients θ_j can be estimated as:

$$\hat{\theta}_{j,n} = \frac{1}{n}\int_0^n \phi_j(t)dy_t. \tag{7}$$

We define a class of weighted least squares estimates for $S(t)$ as:

$$\hat{S}_\lambda = \sum_{j=1}^{n} \lambda(j)\hat{\theta}_{j,n}\phi_j, \tag{8}$$

where the weights $\lambda \in \mathbb{R}^n$ belong to some finite set Λ from $[0, 1]^n$.

Now, for the first $d \leq n$ Fourier coefficients in Eq. (6), we use the improved estimation method proposed for parametric models in [10] and [11]. To this end we set $\tilde{\theta}_n = \left(\hat{\theta}_{j,n}\right)_{1\leq j\leq d}$. In the sequel, we will use the norm $|x|_d^2 = \sum_{j=1}^{d} x_j^2$ for any vector $x = (x_j)_{1\leq j\leq d}$ from $\mathbb{R}^n$. Now we define the shrinkage estimators as:

$$\theta_{j,n}^* = (1 - g(j))\,\hat{\theta}_{j,n}, \tag{9}$$

where $g(j) = \frac{c_n}{|\tilde{\theta}_n|_d} 1_{\{1\leq j\leq d\}}$, 1_A is the indicator of the set A and c_n is some known parameter such that $c_n \approx \frac{d}{n}$ as $n \to \infty$. Now we introduce a class of shrinkage weighted least squares estimates for S as:

$$S_\lambda^* = \sum_{j=1}^{n} \lambda(j)\theta_{j,n}^*\phi_j. \tag{10}$$

We denote the difference of quadratic risks of the estimates in Eq. (10) and Eq. (8) as $\Delta_Q(S) := R_Q(S_\lambda^*, S) - R_Q\left(\hat{S}_\lambda, S\right)$. Now for this deviation, we obtain the following result.

Theorem 1 *Assume that for any vector $\lambda \in \Lambda$ there exists some fixed integer $d = d(\lambda)$ such that their first d components equal to one. Then for any $n \geq 1$ and $r > 0$:*

$$\sup_{Q\in Q_n} \sup_{\|S\|\leq r} \Delta_Q(S) < -c_n^2. \tag{11}$$

The inequality in Eq. (11) means that non-asymptotically, i.e. for any $n \geq 1$ the estimate in the

Eq. (10) outperforms in mean square accuracy the estimate in the Eq. (8). Moreover, as we will see below, $nc_n \to \infty$ as $n \to \infty$. This means that the improvement effect in the nonparametric case is more significant than for parametric regression [11].

3. Model Selection Procedure

This Section gives the construction of a model selection procedure for estimating a function S in the Eq. (1) on the basis of improved weighted least squares estimates and states the sharp oracle inequality for the robust risk of the proposed procedure.

The model selection procedure for the unknown function S in the Eq. (1) will be constructed on the basis of a family of estimates $(S^*_\lambda)_{\lambda\in\Lambda}$. The performance of any estimate S^*_λ will be measured by the empirical squared error:

$$\mathrm{Err}_n(\lambda) = \|S^*_\lambda - S\|^2 . \tag{12}$$

In order to obtain a good estimate, we have to write a rule to choose a weight vector $\lambda \in \Lambda$ in the Eq. (6). It is obvious that the best approach is to minimize the empirical squared error with respect to λ. Making use the estimate definition in the Eq. (6) and the Fourier transformation of S implies:

$$\mathrm{Err}_n(\lambda) = \sum_{j=1}^{n} \lambda^2(j)\left(\theta^*_{j,n}\right)^2 - 2\sum_{j=1}^{n} \lambda(j)\theta^*_{j,n}\theta_j + \sum_{j=1}^{n} \theta_j^2 . \tag{13}$$

Since the Fourier coefficients $(\theta_j)_{j\geq1}$ are unknown, the weight coefficients $(\lambda_j)_{j\geq1}$ cannot be found by minimizing this quantity. To circumvent this difficulty one needs to replace the terms $\theta^*_{j,n}\theta_j$ by their estimators $\tilde{\theta}_{j,n}$. We set:

$$\tilde{\theta}_{j,n} = \theta^*_{j,n}\hat{\theta}_{j,n} - \frac{\hat{\sigma}_n}{n}, \tag{14}$$

where $\hat{\sigma}_n$ is the estimate for the noise variance of $\sigma_Q = E_Q\xi^2_{j,n}$ which we choose in the following form:

$$\hat{\sigma}_n = \sum_{j=[\sqrt{n}]+1}^{n} \hat{t}^2_{j,n} \quad \text{and} \quad \hat{t}_{j,n} = \frac{1}{n}\int_0^n \phi_j(t)\mathrm{d}y_t . \tag{15}$$

For this change in the empirical squared error, one has to pay some penalty. Thus, one comes to the cost function of the form:

$$J_n(\lambda) = \sum_{j=1}^{n} \lambda^2(j)\left(\theta^*_{j,n}\right)^2 - 2\sum_{j=1}^{n} \lambda(j)\tilde{\theta}_{j,n} + \delta\hat{P}_n(\lambda), \tag{16}$$

where δ is some positive constant and $\hat{P}_n(\lambda)$ is the penalty term defined as:

$$\hat{P}_n(\lambda) = \frac{\hat{\sigma}_n |\lambda|_n^2}{n}. \tag{17}$$

Substituting the weight coefficients, minimizing the cost function:

$$\lambda^* = \underset{\lambda\in\Lambda}{\mathrm{argmin}}\, J_n(\lambda) \tag{18}$$

in the Eq. (10) leads to the improved model selection procedure:

$$S^* = S^*_{\lambda^*}. \tag{19}$$

It will be noted that λ^* exists because Λ is a finite set. If the minimizing sequence in the Eq. (18) λ^* is not unique, one can take any minimizer. In the case, when the value of σ_Q is known, one can take $\hat{\sigma}_n = \sigma_Q$ and $P_n(\lambda) = \sigma_Q |\lambda|_n^2 n^{-1}$.

Theorem 2 *For any $n \geq 2$ and $0 < \delta < \frac{1}{2}$, the robust risks defined in the Eq. (4) of estimate in the Eq. (19) for continuously differentiable function S satisfies the oracle inequality:*

$$R^*\left(S^*_{\lambda^*}, S\right) \leq \frac{1+5\delta}{1-\delta} \min_{\lambda\in\Lambda} R^*\left(S^*_\lambda, S\right) + \frac{B^*_n}{n\delta}, \tag{20}$$

where the term B^*_n is independent of S and such that $B^*_n n^{-\epsilon} \to 0$ as $n \to \infty$ for any $\epsilon > 0$.

The inequality in Eq. (20) allows us to establish that the procedure in the Eq. (19) is optimal in the oracle inequalities sense. This property enables to provide asymptotic efficiency in the adaptive setting, i.e. when information about the signal regularity is unknown.

4. Asymptotic Efficiency

In order to study the asymptotic efficiency, we define the following functional Sobolev ball:

$$W_{k,r} = \left\{ f \in C_p^k[0,1] : \sum_{i=0}^{k} \left\| f^{(i)} \right\|^2 \leq r \right\}, \tag{21}$$

where $r > 0$ and $k \geq 1$ are some unknown parameters, $C_p^k[0,1]$ is the space of k times differentiable 1-periodic functions such that for any $0 \leq i \leq k-1 : f^{(i)}(0) = f^{(i)}(1)$. In order to formulate our asymptotic results we set:

$$v_n = \frac{n}{\xi^*}, l_k(r) = \left((2k+1)\, r\right)^{\frac{1}{(2k+1)}} \left(\frac{k}{\pi\,(k+1)}\right)^{\frac{2k}{(2k+1)}} \tag{22}$$

and we denote by Σ_n of all estimates $\hat{S}_n$ of S measurable with respect to the σ-algebra generated by the process in the Eq. (1).

Theorem 3 *The robust risk defined in the Eq. (4) admits the following asymptotic lower bound:*

$$\liminf_{n\to\infty} \inf_{\hat{S}_n\in\Sigma_n} v_n^{2k/(2k+1)} \sup_{S\in W_{k,r}} R^*\left(\hat{S}_n, S\right) \geq l_k(r). \tag{23}$$

This lower bound is sharp in the following sense.

Theorem 4 *The robust risk defined in the Eq. (4) for the estimating procedure in the Eq. (19) has the following asymptotic upper bound:*

$$\limsup_{n\to\infty} v_n^{2k/(2k+1)} \sup_{S\in W_{k,r}} R^*(S^*,S) \le l_k(r). \tag{24}$$

Theorem 3 and Thm. 4 imply that the model selection procedure S^* is efficient and the parameter $l_k(r)$ defined in the Eq. (22) is the Pinsker constant in this case [3].

5. Monte Carlo Simulations

In this section, we report the results of a Monte Carlo experiment to assess the performance of the proposed model selection procedure in the Eq. (19). In the Eq. (1) we choose 1-periodic function S which is defined as $S(t) = t\sin(2\pi t) + t^2(1-t)\cos(2\pi t)$, for $0 \le t \le 1$. We simulate the Eq. (1) with the noise process defined as:

$$d\xi_t = -\xi_t dt + 0.5dw_t + 0.5dz_t, \tag{25}$$

where $z_t = \sum_{j=1}^{N_t} Y_j$, N_t is a Poisson process with the intensity $\lambda = 1$ and $(Y_j)_{j\ge1}$ is i.i.d. Gaussian $(0,1)$. We use the model selection procedure defined in the Eq. (19) with the weights proposed in [8]: $k^* = 100 + \sqrt{\ln n}$, $\epsilon = \frac{1}{\ln n}$ and $m = \left[\frac{1}{\epsilon^2}\right]$. We used the cost function with $\delta = (3+\ln n)^{-2}$. We define the empirical risk as $\bar{R}\left(\tilde{S},S\right) = \frac{1}{p}\sum_{j=1}^{p}\hat{E}\left(\tilde{S}_n(t_j) - S(t_j)\right)^2$ and $\hat{E}\left(\tilde{S}_n(\cdot) - S(\cdot)\right)^2 = \frac{1}{N}\sum_{l=1}^{N}\left(\tilde{S}_n^l(\cdot) - S(\cdot)\right)^2$ with the frequency of observations $p = 100001$ and numbers of replications $N = 10000$.

Table 1 gives the values for the sample risks for different numbers of observation period n.

Tab. 1: Empirical risks.

n	$\bar{R}\left(\tilde{S},S\right)$	$\bar{R}(S^*,S)$	$\bar{R}\left(\tilde{S},S\right)/\bar{R}(S^*,S)$
100	0.0457	0.0289	1.6
200	0.0216	0.0089	2.4
500	0.0133	0.0021	6.3
1000	0.098	0.0011	8.9

6. Conclusion

In this paper, we considered the problem of nonparametric signal processing on the basis of the observations with the dependent non-Gaussian impulse noises. We developed adaptive efficient statistical model selection procedures based on the shrinkage methods and we have shown that the shrinkage estimation methods considerably improve the non-asymptotic estimation accuracy. The obtained theoretical results are confirmed by the numerical simulation. It turns out that numerically the improvement effect may increase 10 times. Next, for the developed statistical methods we obtained the adaptive efficiency property, which means that we provide the best mean squares accuracy without using the smoothness information about the form of unknown signal. Moreover, in this paper, we studied the accuracy properties for the proposed methods on the basis of the robust approach, i.e. uniformly over all possible unknown noise distributions. This allows us to synthesize the statistical algorithms possessing the high noise immunity properties. The results (their satisfactory concordance with the corresponding experimental data) can be used for the estimation of the signals. Such problems are of a great importance in the fields of radio-and-hydroacoustic communications and positioning, radio-and-hydrolocation, etc. (see [12] and references therein).

Acknowledgment

The results of this work are supported by the Ministry of Science and Higher Education of the Russian Federation in the framework of the research project no. 2.3208.2017/4.6. The second author is partially supported by the Russian Federal Professor Program, project no. 1.472.2016/1.4 (Ministry of Science and Higher Education of the Russian Federation) and by the project XterM-Feder, University of Rouen. The results of Sec. 4. and Sec. 5. are supported by the RSF grant number 17-11-01049.

References

[1] IBRAGIMOV, I. A. and R. Z. KHASMINSKII. *Statistical Estimation: Asymptotic Theory.* 1st ed. New York: Springer, 1981. ISBN 978-1-4899-0027-2.

[2] KASSAM, S. A. *Signal Detection in Non-Gaussian Noise.* 1st ed. New York: Springer, 1988. ISBN 978-1-4612-3834-8.

[3] CHERNOYAROV, O. V., M. VACULIK, A. SHIRIKYAN and A. V. SALNIKOVA. Statistical Analysis of Fast Fluctuating Random Signals with Arbitrary - Function Envelope and Unknown Parameters. *Communications - Scientific Letters of the University of Zilina.* 2015, vol. 17, no. 1a, pp. 35–43. ISSN 1335-4205.

[4] CONT, R. and P. TANKOV. *Financial Modelling with Jump Processes*. 1st ed. Boca Raton: Chapman & Hall, 2003. ISBN 1-58488-413-4.

[5] PCHELINTSEV, E., V. PCHELINTSEV and S. PERGAMENSHCHIKOV. Non asymptotic sharp oracle inequalities for the improved model selection procedures for the adaptive nonparametric signal estimation problem. *Communications - Scientific Letters of the University of Zilina*. 2018, vol. 20, no. 1, pp. 72–76. ISSN 1335-4205.

[6] HOPFNER, R. and Y. A. KUTOYANTS. On LAN for parametrized continuous periodic signals in a time-inhomogeneous diffusion. *Statistical Decisions*. 2009, vol. 27, iss. 4, pp. 309–326. ISSN 0721-2631. DOI: 10.1524/stnd.2009.1064.

[7] HOPFNER, R. and Y. A. KUTOYANTS. Estimating discontinuous periodic signals in a time-inhomogeneous diffusion. *Statistical Inference for Stochastic Processes*. 2010, vol. 13, iss. 3, pp. 193–230. ISSN 1387-0874. DOI: 10.1007/s11203-010-9046-7.

[8] KONEV, V. V. and S. PERGAMENSHCHIKOV. Efficient robust nonparametric estimation in a semimartingale regression model. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*. 2012, vol. 48, no. 4, pp. 1217–1244. ISSN 0246-0203. DOI: 10.1214/12-AIHP488.

[9] KONEV, V. V. and S. PERGAMENSHCHIKOV. Robust model selection for a semimartingale continuous time regression from discrete data. *Stochastic Processes and their Applications*. 2015, vol. 125, iss. 1, pp. 294-326. ISSN 0304-4149. DOI: 10.1016/j.spa.2014.08.003.

[10] KONEV, V. V., S. PERGAMENSHCHIKOV and E. PCHELINTSEV. Estimation of a regression with the pulse type noise from discrete data. *Theory of Probability and Its Applications*. 2014, vol. 58, iss. 3, pp. 442–457. ISSN 0040-585X. DOI: 10.1137/S0040585X9798662X.

[11] PCHELINTSEV, E. Improved estimation in a non-Gaussian parametric regression. *Statistical Inference for Stochastic Processes*. 2013, vol. 16, iss. 1, pp. 15–28. ISSN 1387-0874. DOI: 10.1007/s11203-013-9075-0.

[12] CHERNOYAROV, O. V., Y. A. KUTOYANTS and M. MARCOKOVA. On frequency estimation for partially observed system with small noises in state and observation equations. *Communications - Scientific Letters of the University of Zilina*. 2018, vol. 20, no. 1, pp. 66–71. ISSN 1335-4205.

About Authors

Evgeny PCHELINTSEV was born in Almaty, Kazakhstan. He received his M.Sc. in Mathematics from Tomsk State University in 2009, Ph.D. in Applied Mathematics and Statistics from Tomsk State University and Rouen University in 2012. His research interests include statistical modelling and identification of stochastic dynamic systems, soft computing, and simulation.

Serguei PERGAMENSHCHIKOV was born in Tomsk, Russia. He received his M.Sc. from Tomsk State University in 1980, Ph.D. in Probability Theory and Statistics from Tomsk State University in 1986, Doctor of Sciences in Mathematics from Central Economics and Mathematics Institute of RAS in 1994. His research interests include statistical modelling and identification of stochastic dynamic systems, stochastic differential equations, optimization, and control.

Mariana MARCOKOVA was born in Trencianska Tepla, Slovak Republic. She received the Master degree in teacher training of mathematics and physics at Comenius University in Bratislava, Slovak Republic in 1970. Since 1970 she worked as a teacher of mathematics at several mathematical departments at University of Zilina, Slovak Republic (former the University of Transport in Zilina), in the last 25 years as an associate professor of mathematics and in the last 5 years also as the associate professor at Moscow Power Engineering Institute, Russia. After doctoral studies in mathematical analysis at Palacky University in Olomouc, Czech Republic, she obtained the degrees: RNDr. and CSc. Her research interests include special functions, orthogonal polynomials, and applied mathematics.