Improved signal processing observed with semi-Markov noises

E A Pchelintsev, S S Perelevskiy, M A Povzun
Tomsk State University, 36 Lenin avenue, Tomsk, Russia
E-mail: evgen-pch@yandex.ru

Abstract. This paper considers the problem of improved adaptive estimating a periodic signal observed in the transmission channel with the dependent noise defined by semi-Markov processes with unknown probabilities and spectral characteristics. Improved adaptive model selection procedure is proposed. The comparison between improved and least squares methods is studied. Sharp oracle inequality for the procedure risk is obtained. The efficiency of the proposed procedure is established. The Monte Carlo simulation results are given.

1. Introduction. Problem statement
We consider the estimation problem for the 1-periodic signal $S(t)$ from the observations $(y_t)_{0 \leq t \leq n}$ given by the stochastic differential equation

$$\text{d} y_t = S(t) \text{d} t + \text{d} \xi_t, \quad 0 \leq t \leq n,$$

where $n$ is the duration of observation and the noise process $(\xi_t)_{0 \leq t \leq n}$ is defined as

$$\xi_t = \sigma_1 w_t + \sigma_2 L_t + \sigma_3 z_t.$$

Here $\sigma_1$, $\sigma_2$ and $\sigma_3$ are unknown coefficients, $(w_t)_{t \geq 0}$ is the Brownian motion process, $(L_t)_{t \geq 0}$ is the jump Levy process, i.e.

$$L_t = \int_0^t \int \{0\} x \cdot (\mu(ds,dx) - \tilde{\mu}(ds,dx)),$$

where $\mu$ is some jump measure with deterministic compensator $\tilde{\mu}$ (see for details [1]). The process $(z_t)_{t \geq 0}$ is a pure jump semi-Markov process (see, for example, [2]) with the following form

$$z_t = \sum_{i=0}^{N_t} Y_i,$$

where $(Y_i)_{i \geq 0}$ is a sequence of independent and identically distributed random variables with zero mean, unit variance and bounded fourth moment; $(N_t)_{t \geq 0}$ is a counting process (see, for example, [3]) defined as

$$N_t = \sum_{l=1}^{\infty} 1_{\{T_l \leq t\}}, \quad T_l = \sum_{k=1}^{l} \tau_k,$$
where $1_A$ is an indicator of the set $A$, $(\tau_k)_{k \geq 1}$ is a sequence of independent and identically distributed positive integrated random variables with some known distribution.

Note that, if in the equation (2) $\sigma_2 = \sigma_3 = 0$, then we have the "signal+white noise" model which widely uses in statistical radiophysics ([4] and [5]). In this paper we suppose that in addition to intrinsic noises in radio-electronic systems, approximated usually by the gaussian white or colour noise, the useful signal $S$ is distorted by the impulse flow modeled by semi-Markov–Levy processes defined in the equation (2). The examples of a pulse stream are either external unintended (atmospheric) or intentional impulse noises or errors. Recall that, the impulse noises in the detection signal problems through compound Poisson processes have been studied by Kassam [6], Flaksman [7], Konev and Pergamenshchikov [8], Konev, Pergamenshchikov and Pchelintsev [9] for parametric and nonparametric signal estimation problems. Since the real technical systems (for example, telecommunication or navigation systems) are functioning under noise impulses having different sizes and different frequencies, then the models with compound Poisson processes are not used for these ones. In these cases we can use the models with the general Levy noises [10, 11] or non-Gaussian Ornstein–Uhlenbeck noises driven by Levy processes [12, 13]. Such process allow us to consider continuous time models with dependent observations. Unfortunately, the dependence for the Ornstein–Uhlenbeck processes decreases with a geometric rate. Thus, when the observations duration tends to infinity, we obtain quickly the "signal+white noise" model. For this reason we need to study the dependent observations models with the noise processes of the type (2) for which the dependence persists for a sufficient large observations duration. We can study the problem of the signals processing observed under long impulse noise impact with a memory or in the presence of "against signals".

The main goal is to develop a new improved adaptive robust efficient signal estimation method for the model defined by equations (1)–(2). We assume that the noise process distribution $Q$ is unknown and belongs to the class $Q_n$ defined as a family of all distributions with the parameters $\sigma_1 \geq \sigma$ and $\sigma_1^2 + \sigma_2^2 + \sigma_3^2 / E_{Q_T} \leq \sigma$, where $\sigma$ and $\sigma$ are some positive bounds. In general case, the parameters $\sigma_1, \sigma_2, \sigma_3$ and their bounds can be dependent on $n$.

The quality of any estimate $\hat{S}_n$ (i.e. any function of observations $(y_t)_{0 \leq t \leq n}$) of the unknown signal $S$ will be measured with the robust quadratic risk

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

where $R_Q(\cdot, \cdot)$ is the usual quadratic risk defined as

$$R_Q(\hat{S}_n, S) := E_{Q,S} \| \hat{S}_n - S \|^2$$

and $\| S \|^2 = \int_0^1 S^2(t) dt$.

Here $E_{Q,S}$ is the expectation with respect to the distribution $P_{Q,S}$ of the process in the equation (1) with a fixed distribution $Q$ of the noise (2) and a given function $S$.

2. Improved estimation procedure

In order to estimate the unknown signal $S$ we use here the improved model selection procedure for regression models in continuous time from [14] defined via the shrinkage weighted least squares estimates. We consider the Fourier expansion of the function $S$ from $L_2[0, 1]$

$$S(t) = \sum_{j=1}^{\infty} a_j f_j(t) \quad \text{with} \quad a_j = \int_0^1 S(t) f_j(t) dt,$$
where \((f_j)_{j \geq 1}\) be an orthonormal trigonometric basis in \(L^2[0,1]\), i.e.

\[
f_j(x) = \sqrt{2} \begin{cases} 
\cos(2\pi [j/2] x) & \text{for even } j, \\
\sin(2\pi [j/2] x) & \text{for odd } j,
\end{cases}
\]

\([x]\) denotes integer part of \(x\). So, for estimating the signal \(S\) one needs to estimate the Fourier coefficients \(a_j\) and to replace them in the Fourier representation by their estimators. Given the fact that the functions \(S\) and \(f_j\) are 1-periodic and replacing the differential \(S(t) dt\) by the stochastic differential \(dy_t\) (which is observed), we can write the least squares estimate for \(a_j\) on the time interval \([0,n]\)

\[
\hat{a}_j = \frac{1}{n} \int_0^n f_j(t) dy_t.
\]

In order to obtain efficient estimator for \(S\) we consider weighted least squares estimators (instead of ordinary projection estimators [5]) of the form

\[
\hat{S}_{\lambda}(t) = \sum_{j=1}^n \lambda_j \hat{a}_j f_j(t),
\]

(4)

where the weights \(\lambda = (\lambda_j)_{1 \leq j \leq n} \in \mathbb{R}^n\) belong to some finite set \(\Lambda\) which is defined in the equations (3.12)–(3.13) from [14]. From here we have that for any vector \(\lambda \in \Lambda\) there exists some fixed integer \(d = d(\lambda)\) such that their first \(d\) components are equal to one, i.e. \(\lambda_j = 1\) for \(1 \leq j \leq d\). Such weight coefficients were defined by Nussbaum [15] to construct the efficient estimation for the nonparametric discrete time regression.

The main goal of our paper is to improve the quality of nonparametric signals estimation. For this we change the first \(d\) estimates \(\hat{a}_j\) in the equation (4) by the shrinkage estimates

\[
a_j^* = (1 - g_\lambda(j)) \hat{a}_j \quad \text{and} \quad g_\lambda(j) = \frac{c_n}{|\hat{a}_\lambda|_d} 1_{\{1 \leq j \leq d\}},
\]

where

\[
c_n = c_n(\lambda) = \frac{(d-1)\sigma}{(r + \sqrt{d\sigma/n})n},
\]

and \(|\hat{a}_\lambda|_d^2 = \sum_{j=1}^d \hat{a}_j^2\) (we recall that the parameter \(d\) depends on \(\lambda\)). Here \(r\) is a function of \(n\) with

\[
\lim_{n \to \infty} r(n) = \infty \quad \text{and} \quad \lim_{n \to \infty} n^{-\epsilon} r(n) = 0
\]

for any \(\epsilon > 0\).

Now we define a family of shrinkage weighted least squares estimates for \(S\) as

\[
S^*_\lambda(t) = \sum_{j=1}^n \lambda_j a_j^* f_j,
\]

(5)

We denote the difference of the quadratic risks of the estimates (5) and (4) as

\[
\Delta_Q(S) := R_Q(S^*_\lambda, S) - R_Q(\hat{S}_\lambda, S).
\]

For this quantity uniformly over the distribution family \(Q_n\) the following inequality holds

\[
\sup_{Q \in Q_n} \sup_{\|S\| \leq r} \Delta_Q(S) \leq -c_n^2.
\]
It means that non-asymptotically, i.e. for fixed $n \geq 1$, the estimate (5) exceeds in terms of the mean square robust accuracy the estimate (4) and the minimal gain is equal to $c^2_n$. In view of the definition of the $c_n$, one has that $nc_n \to \infty$ as $n \to \infty$. This property implies that the proposed improved method is sufficiently efficient for nonparametric models.

Further to construct a good estimate, we need some rule to choose a weight vector $\lambda \in \Lambda$ in (5). To obtain efficient estimate, the coefficients $\lambda_j$ in (5) need to be chosen depending on the regularity of the unknown function $S$ (see, for example, [16]). But we consider the adaptive case, and we choose the weight coefficients on the basis of the model selection method proposed in [8]. We introduce the cost function of the form

$$J_n(\lambda) = \sum_{j=1}^{n} \lambda_j^2 (a_j^*)^2 - 2 \sum_{j=1}^{n} \lambda_j \bar{\theta}_j + \delta \hat{P}_n(\lambda),$$

where $\delta$ is a positive parameter, $\hat{P}_n(\lambda)$ is the "penalty" defined as

$$\hat{P}_n(\lambda) = \frac{\hat{\sigma}_n |\lambda|_2^2}{n}$$

and

$$\bar{\theta}_j = a_j^* \hat{a}_j - \frac{\hat{\sigma}_n}{n}, \quad \hat{\sigma}_n = \sum_{j=[\sqrt{n}]+1}^{n} \hat{\tau}_j^2, \quad \text{with} \quad \hat{\tau}_j = \frac{1}{n} \int_{0}^{n} f_j(t)dy.$$

Now we define the improved model selection procedure as

$$S^* = S^*_\lambda^* \quad \text{and} \quad \lambda^* = \text{argmin}_{\lambda \in \Lambda} J_n(\lambda).$$

Since $\Lambda$ is a finite set, then $\lambda^*$ exists. If the such $\lambda^*$ is not unique, one can choose any of them.

Further we will study the properties for the robust quadratic risk (3) of the constructed model selection procedure (6). Firstly we obtain a non-asymptotic sharp oracle inequality. Under some assumptions on the noise process distribution (see [2] for details) one has that for any $n \geq 1$ and $0 < \delta < 1/3$, the robust risk (3) of estimate (6) for continuously differentiable function $S$ satisfies the oracle inequality

$$R^*_n(S^*, S) \leq (1 + \phi(\delta)) \min_{\lambda \in \Lambda} R^*_n(S^*_\lambda, S) + B_n \frac{(1 + ||\dot{S}||_2^2)}{n^\delta},$$

where terms $\phi(\delta)$ and $B_n$ are independent of $S$ and for any $\varepsilon > 0$

$$\lim_{n \to \infty} \frac{B_n}{n^\varepsilon} = 0, \quad \lim_{n \to \infty} \frac{\phi(\delta)}{n^\varepsilon} = 0.$$

Note that for the signal processing in the adaptive setting we need in the model selection procedure (6) use the parameter $\delta$ defined as a function of $n$ and for any $\varepsilon > 0$

$$\lim_{n \to \infty} \delta_n = 0 \quad \text{and} \quad \lim_{n \to \infty} n^\varepsilon \delta_n = \infty.$$

For example, we can put $\delta_n = (3 + \log n)^{-1}$.

Then, using the oracle inequality (7), we can establish the asymptotic efficient property of the constructed model selection procedure for the estimating of unknown signal $S$ in the model (1)–(2) which belongs to some Sobolev ball $W$ from $L_2[0, n]$. Namely, for the robust quadratic risk (3) of estimate (6) we have

$$\lim_{n \to \infty} \frac{\inf_{\hat{S}} \sup_{S \in W} R^*_n(\hat{S}, S)}{\sup_{S \in W} R^*_n(S^*, S)} = 1.$$
3. Monte Carlo simulation

In this section, we present the results of the Monte Carlo experiments. We put in the model (1) the function

\[ S(t) = t \sin(2\pi t) + t^2(1 - t) \cos(4\pi t) \]  (8)

on \([0, 1]\) and the noise process \(\xi_t\) is defined as

\[ \xi_t = 0.5 w_t + 0.5 z_t. \]

Here \(z_t\) is a semi-Markov process with \(Y_j \sim N(0, 1)\) and \(\tau_k \sim \chi^2_3\) (see [2]). We simulate the estimate (6) with the weight coefficients in which

\[ k_n = 100 + \sqrt{\ln(n + 1)}, \quad r_i = i / \ln(n + 1), \quad m = \ln^2(n + 1), \quad \sigma = 0.5 \quad \text{and} \quad \delta = (3 + \log n)^{-1}. \]

We define the empirical risk as

\[ R(S^*, S) = \frac{1}{p} \sum_{j=1}^{p} \hat{E}(S_n^*(t_j) - S(t_j))^2, \]

\[ \hat{E}(S_n^*(\cdot) - S(\cdot))^2 = \frac{1}{N} \sum_{l=1}^{N} (S_{n,l}^*(\cdot) - S(\cdot))^2, \]

where \(p = 100001\) and \(N = 1000\) are the observation frequency and replications respectively.

Table 1 gives the empirical risks of the improved estimate (6) and the estimate (3.15) from [2] for different \(n\). Table 2 gives the empirical risks of the the estimate (3.15) from [2] and it’s improved version for different \(n\).

**Table 1.** Risks comparison of estimates (4) and (5) for different optimal \(\lambda\).

| \(n\) | \(\mathcal{R}(\hat{S}_{\lambda}^*, S)\) | \(\mathcal{R}(S_{\lambda}^*, S)\) | \(\mathcal{R}(\hat{S}_{\lambda}^*, S) / \mathcal{R}(S_{\lambda}^*, S)\) |
|-----|-------------------|-------------------|-------------------|
| 20  | 3.1605            | 0.9447            | 3.35              |
| 100 | 0.8815            | 0.2604            | 3.39              |
| 1000| 0.1198            | 0.0277            | 4.32              |
| 10000| 0.0307           | 0.0056            | 5.48              |

**Table 2.** Risks comparison of estimates (4) and (5) for the same optimal \(\hat{\lambda}\).

| \(n\) | \(\mathcal{R}(\hat{S}_{\hat{\lambda}}^*, S)\) | \(\mathcal{R}(S_{\hat{\lambda}}^*, S)\) | \(\mathcal{R}(\hat{S}_{\hat{\lambda}}^*, S) / \mathcal{R}(S_{\hat{\lambda}}^*, S)\) |
|-----|-------------------|-------------------|-------------------|
| 20  | 3.1605            | 1.1569            | 2.73              |
| 100 | 0.8815            | 0.7962            | 1.11              |
| 1000| 0.1198            | 0.0963            | 1.24              |
| 10000| 0.0307           | 0.0062            | 4.95              |

Figures 1–4 show the graphs of the real signal (8) (bold line) and its estimates (4) (continuous line) and (6) (dashed line) for different \(n\).

From the Table 2 we can see that theoretical result on the improvement effect is confirmed by the Monte Carlo simulations. Besides, for the proposed improved procedure, Table 1 and Figures 1–4 show that the gain is considerable for non large \(n\).
4. Conclusion
In this paper, we considered the problem of nonparametric signal processing on the basis of the observations with the dependent non-Gaussian semi-Markov impulse noises. We developed improved adaptive efficient statistical model selection method and we have shown that such approach improve the non-asymptotic accuracy of estimation. The obtained theoretical results are confirmed by the numerical simulation. For the developed statistical methods we obtained the adaptive efficiency property, i.e. we provide the best mean squares accuracy without using the smoothness information about the form of unknown signal. As well, 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 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 [17] and references therein).

Acknowledgments
This work was supported by the Grant of President of Russian Federation (research project no. MK-834.2020.9).

References
[1] Jacod J and Shiryaev A N 2002 Limit theorems for stochastic processes (Berlin: Springer).
[2] Barbu V S, Beltaief S and Pergamenshchikov S M 2019 Stat. Inference Stoch. Process 22 pp 187-231.
[3] Mikosch T 2004 Non-life insurance mathematics. An introduction with stochastic processes (Berlin: Springer).
[4] Ibragimov I A and Khasminskii R Z 1981 Statistical Estimation: Asymptotic Theory (New York: Springer).
[5] Kassam S A 1988 Signal detection in non-Gaussian noise (New York: Springer).
[6] Flaksman A G 2002 Radiophysics and Quantum Electronics 45 pp 977-988.
[7] Konev V V and Pergamenshchikov S M 2012 Ann. Inst. Henri Poincaré Probab. Stat. 48 pp 1217-1244.
[8] Pechlīntsev E A, Pechlīntsev V A and Pergamenshchikov S M 2018 Communications – Scientific Letters of the University of Žilina 20 (1) pp 72-76.
[9] Pechlīntsev E A, Pechlīntsev V A and Pergamenshchikov S M 2019 Journal of Nonparametric Statistics 31 (3) pp 612-628.
[10] Pechlīntsev E A, Pergamenshchikov S M and Marcokova M 2019 Advances in Electrical and Electronic Engineering 17 pp 270-274.
[11] Pechlīntsev E A and Pergamenshchikov S M 2019 Tomsk State University Journal of Mathematics and Mechanics 58 pp 14-31.
[12] Pechlīntsev E A and Pergamenshchikov S M 2018 Stat. Inference Stoch. Process 21 pp 469-483.
[13] Nussbaum M 1985 Ann. Statist. 13 pp 984-997.
[14] Pinsker M S 1981 Problems of Transmission information 17 pp 120-133.
[15] Chernoyarov O V, Kutoyants Yu A and Marcokova M 2018 Communications - Scientific Letters of the University of Žilina 20 (1) pp 66-71.