

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

УДК 519
ББК 22.17
Меж43

Международная научная конференция
«Робастная статистика и финансовая математика – 2019»:
Меж43 сборник статей / под ред. С.М. Пергаменщикова,
 Е.А. Пчелинцева. – Томск : Издательский Дом Томского
 государственного университета, 2019. – 90 с.

ISBN 978-5-94621-849-8

В сборнике представлены статьи, посвященные актуальным проблемам математической статистики и финансовой математики, а также современным подходам и методам решения фундаментальных и прикладных задач.

Для студентов, аспирантов, специалистов в области математической статистики и ее приложений.

УДК 519
ББК 22.17

*Сборник подготовлен к печати при поддержке Российского
научного фонда, проект № 17-11-01049.*

Локальный организационный комитет конференции:

к.ф.-м.н., доцент Пчелинцев Е. А. (председатель оргкомитета);
к.ф.-м.н., доцент Пчелинцев В. А. (ответственный секретарь);
Повзун М. А.;
Шишкова А. А.

ISBN 978-5-94621-849-8

© Томский государственный университет, 2019
© Авторы статей, 2019

Содержание

<i>Burkatovskaya Yu.B., Vasiliev V.A. Truncated Parameter Estimation of AR(1) Process with Additive Noise.....</i>	<i>5</i>
<i>Dachian S., Kutoyants Yu., Yang L. On Hypotheses Testing for Poisson Processes: Regular and Singular Cases.....</i>	<i>11</i>
<i>Galtchouk L., Pergamenshchikov S. Sharp Oracle Inequalities for Drift Estimation Problems from Discrete Data.....</i>	<i>20</i>
<i>Kutoyants Yu.A. On Identification of Continuous Time Hidden Markov Processes: a Survey</i>	<i>26</i>
<i>Pchelintsev E.A. Improved Model Selection for Estimation in Semimartingale Rregression from Discrete Data</i>	<i>36</i>
<i>Povzun M.A., Pchelintsev E.A., Pergamenshchikov S.M. Efficient Nonparametric Estimation of Square-integrable Functions in Continuous Time Regression Models.....</i>	<i>43</i>
<i>Shishkova A.A. The Hedging Problem for Asian Options in Financial Markets with Transaction Costs.....</i>	<i>49</i>
<i>Болдин М.В. Асимптотическая мощность и качественная робастность критерия Пирсона для проверки нормальности авторегрессии.....</i>	<i>54</i>
<i>Воробейчиков С.Э., Удод В.А. Сегментация изображений на основе алгоритма обнаружения разладки</i>	<i>61</i>
<i>Жуман А.Б., Пчелинцев Е.А. Оценивание непараметрической регрессии с полумарковскими шумами.....</i>	<i>67</i>

<i>Перелевский С.С., Пчелинцев Е.А.</i> Асимптотически эффективное оценивание коэффициента сноса в диффузионных процессах	75
<i>Степаненко А.С., Пчелинцев Е.А., Бордзиловский А.С.</i> Построение и анализ регрессионной модели, описывающей зависимость объемного коэффициента нефти от некоторых ее физических характеристик.....	83

Truncated Parameter Estimation of AR(1) Process with Additive Noise ¹

Burkatovskaya Yu. B., Vasiliev V. A.

Tomsk Polytechnical University, Tomsk
Tomsk State University, Tomsk
e-mail: vas@mail.tsu.ru

Abstract

This paper presents a truncated estimator of the dynamic parameter of a stable AR(1) process by observations with additive noise. The estimator is constructed by sample of a fixed size and it has a known upper bound of the mean square deviation. Cases of known and unknown variance of observation noise are considered.

Keywords: autoregressive process, observations with noise, fixed sample size, guaranteed accuracy.

Introduction and problem statement. Development of parameter estimation methods of dynamic systems by samples of finite or fixed size is very important in statistical problems such that model construction and various adaptive problems (prediction, control, filtration etc.).

One of the possibilities for finding estimators with the guaranteed quality of inference using a sample of fixed size is provided by the approach of truncated estimation. Truncated estimators were constructed in [10] for ratio type multivariate functionals by a fixed-size sample. They have guaranteed accuracy in the sense of the L_{2m} -norm, $m \geq 1$. This fact allows one to obtain desired non-asymptotic and asymptotic properties of the estimators. The truncated estimation method was developed in [2] and others for parameter estimation problems in discrete-time dynamic models. Solutions of some non-asymptotic parametric and non-parametric problems can be found also in [1], [5], [9], [6], [7], among others. In particular, [9] established the minimax optimality of the least-squares estimator of the dynamic parameter in AR(1) model.

In this paper, the truncated estimation method introduced in [10] is applied for the parameter estimation of AR(1) by additively-noised observations with unknown noise variance (another applications of this method can be found, e.g., in [3], [4]).

Consider the estimation problem of the parameter λ of the scalar first-order autoregressive process $(x_n)_{n \geq 0}$ satisfying the equation

$$x_n = \lambda x_{n-1} + \xi_n, \quad n \geq 1 \quad (1)$$

¹The work was supported by the RSF, Project No. 17-11-01049.

by observations

$$y_n = x_n + \eta_n, \quad n \geq 0. \quad (2)$$

Process (1) is supposed to be stable, i.e. $|\lambda| < 1$. Introduce the notation $\zeta = (x_0, \xi_1, \eta_0, \dots)$. The processes (ξ_n) , (η_n) and x_0 are supposed to be mutually independent; noises ξ_n and η_n form sequences of i.i.d. random variables such that $E\zeta = 0, E\|\zeta\|^4 < \infty$. Denote $\sigma^2 = E\eta_0^2$. We assume that the variance of ξ_1 is known. Then without loss of generality we put $E\xi_1^2 = 1$.

The main aim of the paper is to construct truncated estimators of $\lambda \in (-1, 1)$ with guaranteed accuracy in the mean square sense by sample of fixed size. Cases of both known and unknown values of σ^2 will be considered.

A similar problem has been solved in, e.g., [11] on the basis of the sequential approach (when the sample size is a random value determined by a special stopping rule) for $\lambda \in (-1, 0) \cap (0, 1)$

Parameter estimation of AR(1) with known noise variance.

To estimate the parameter λ , we use the correlation method. To this end, we obtain from the system (1), (2) the recurrent equation for the observed process $y = (y_n)_{n \geq 0}$:

$$\begin{aligned} y_n &= \lambda y_{n-1} + \delta_n, \quad n \geq 1, \\ \delta_n &= \xi_n + \eta_n - \lambda \eta_{n-1}. \end{aligned} \quad (3)$$

Due to the dependence of noises δ_n , the least squares estimator (LSE) of λ obtained from equation (3) is asymptotically biased, see, e.g., [8], [11]. Equation (3) implies the following formula for correlations of the process (y_n) :

$$E_\lambda y_n y_{n-1} = \lambda E_\lambda (y_{n-1}^2 - \sigma^2), \quad n \geq 1.$$

Hence, the consistent correlation estimator $\hat{\lambda}_n$ of λ has the following form (see [8])

$$\hat{\lambda}_{n,\sigma} = \frac{\sum_{k=1}^n y_k y_{k-1}}{\sum_{k=1}^n (y_{k-1}^2 - \sigma^2)}, \quad n \geq 1. \quad (4)$$

It is easy to verify that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (y_{k-1}^2 - \sigma^2) = \frac{1}{1 - \lambda^2} > 1 \quad P_\lambda - a.s. \quad (5)$$

Thus, according to the general procedure described in [10], it is reasonable to construct the truncated estimator $\tilde{\lambda}_n$ of λ as follows:

$$\tilde{\lambda}_n = \hat{\lambda}_n \cdot \chi\left(\sum_{k=1}^n (y_{k-1}^2 - \sigma^2) \geq hn\right), \quad n \geq 1, \quad (6)$$

where $h \in (0, 1)$ and $\chi(A)$ is the indicator of the set A .

The following Theorem 1 gives the first main result of this paper.

Theorem 1. *Assume the model (1), (2). Then for every $|\lambda| < 1$ and $n \geq 1$, estimator (6) has the property*

$$E_\lambda(\tilde{\lambda}_n - \lambda)^2 \leq \frac{C}{n}. \quad (7)$$

Parameter estimation of AR(1) with unknown noise variance.

To estimate $\lambda \in (-1, 1)$, we use an adaptive modification of estimator (4):

$$\lambda_n^* = \frac{\frac{1}{n} \sum_{k=1}^n y_k y_{k-1}}{\frac{1}{n} \sum_{k=1}^n y_{k-1}^2 - \sigma_n^2}, \quad n > 1. \quad (8)$$

Taking into account (5), we construct the estimator σ_n^2 of σ^2 as follows

$$\sigma_n^2 = \frac{1}{n} \sum_{k=1}^n y_{k-1}^2 - \frac{1}{1 - \lambda_n^2}, \quad n > 1 \quad (9)$$

where λ_n is the pilot estimator of λ

$$\lambda_n = \text{proj}_{[-1,1]} \tilde{\lambda}_n, \quad n > 1, \quad (10)$$

$$\tilde{\lambda}_n = \frac{\sum_{k=2}^n y_k y_{k-2}}{\sum_{k=2}^n y_{k-1} y_{k-2}} \cdot \chi(|\sum_{k=2}^n y_{k-1} y_{k-2}| \geq H_n), \quad n > 1. \quad (11)$$

Here we put $H_n = n(\log n)^{-1}$. According to the general truncated estimation method [10], the multiplier $(\log n)^{-1}$ in the definition of H_n can be any other slowly-decreasing function.

It should be noted that the estimator (10) is constructed on the bases of the correlation (Yule-Walker type) estimator which can not be used if $\lambda = 0$ (see Lemma 1 below). Our main aim is to construct an estimator of λ without this restriction.

Taking into account (9), estimator (8) can be written in the form

$$\lambda_n^* = (1 - \lambda_n^2) \frac{1}{n} \sum_{k=1}^n y_k y_{k-1}, \quad n > 1. \quad (12)$$

Lemma 1. *Assume that in model (1), (2), $E\|\zeta\|^8 < \infty$. Then estimator (10) for every $\lambda \in (-1, 0) \cup (0, 1)$ and $n > 1$ has the following property*

$$E_\lambda(\lambda_n - \lambda)^2 \leq C \left(\frac{1}{n} + \frac{\log^4 n}{n^2} \right).$$

This lm makes possible to obtain the main result of the section.

Theorem 2. Assume that in model (1), (2), $E\|\zeta\|^8 < \infty$. Then for every $|\lambda| < 1$ and $n > 1$, estimator (12) satisfies the following condition

$$E_{\lambda}(\lambda_n^* - \lambda)^2 \leq C \left(\frac{1}{n} + \frac{\log^4 n}{n^2} \right).$$

Simulation Results and Discussion. We conducted numerical simulation of the proposed estimation algorithm.

For every set of the parameters, the experiment was performed 100 times, the number of observations is equal to 100, the parameter of the procedure $h = 0,5$. Table 1 presents the results of simulation.

Here λ and σ are the parameters of model (1), $\tilde{\lambda}_n$ and λ_n^* are the mean estimators of the parameter λ when the noise variance σ^2 is supposed to be known and unknown, correspondingly; $\tilde{d}_n$ and d_n^* are sample standard errors of the corresponding estimators.

Table 1. Simulation results

λ	σ^2	$\tilde{\lambda}_n$	$\tilde{d}_n$	λ_n^*	d_n^*
0,5	0,09	0,477	0,0092	0,452	0,0294
0,5	0,25	0,492	0,0111	0,490	0,0314
0,5	0,49	0,487	0,0150	0,470	0,0488
0,5	0,81	0,475	0,0229	0,418	0,0795
0,5	1	0,473	0,0419	0,424	0,0953
0,8	0,09	0,786	0,0046	0,796	0,0465
0,8	0,25	0,794	0,0054	0,854	0,0793
0,8	0,49	0,786	0,0054	0,789	0,0737
0,8	0,81	0,772	0,0120	0,765	0,1435
0,8	1	0,788	0,0122	0,797	0,1590
0,9	0,09	0,876	0,0038	0,865	0,0772
0,9	0,25	0,889	0,0018	0,913	0,0596
0,9	0,49	0,888	0,0030	0,910	0,1044
0,9	0,81	0,874	0,0044	0,886	0,1822
0,9	1	0,891	0,0028	0,891	0,1780

One can see that $\tilde{d}_n < d_n^*$ in all experiments; thus, if the noise variance is unknown then the standard error increases at least twice (if $\lambda = 0,5$); but d_n^* can be fully ten times larger than $\tilde{d}_n$, if $\lambda = 0,9$. Both deviations increase with the grow of σ^2 , as one should expect; besides, $\tilde{d}_n$ decreases and d_n^* increases with the increase of λ .

References

- [1] Dobrovidov A.V., Koshkin G.M., Vasiliev V.A. Non-parametric state space models. Heber City, Utah: Kendrick Press. 2012. <http://www.kendrickpress.com/nonpara.html>
- [2] Dogadova T., Vasiliev V. Guaranteed parameter estimation of stochastic linear regression by sample of fixed size // Tomsk State University Journal of Control and Computer Science. 2014. Vol. 26, No. 1. P. 39–52.
- [3] Dogadova T., Kusainov M., Vasiliev V. Truncated estimation method and applications // Serdica. Mathematical Journal Bulgarian Academy of Sciences Institute of Mathematics and Informatics. 2017. Vol. 43. P. 221–266.
- [4] Kusainov M.I., Vasiliev V.A. On optimal adaptive prediction of multivariate autoregression // Sequential Analysis. 2015. Vol. 35, No. 2. P. 211–234.
- [5] Mikulski P.W., Monsour M.J. Optimality of the maximum likelihood estimator in first-order autoregressive processes // Journal of Time Series Analysis. 1991. Vol. 12, No. 3. P. 237–253.
- [6] Roll J., Nazin A., Ljung L. A Non-Asymptotic Approach to Local Modelling // The 41st IEEE CDC, Las Vegas, Nevada, USA, 10–13 Dec. 2002 (Regular paper). Predocumentation is available at <http://www.control.isy.liu.se/research/reports/2003/2482.pdf>
- [7] Roll J., Nazin A., Ljung L. Non-linear System Identification Via Direct Weight Optimization // Automatica. 2005. Vol. 41, No. 3. P. 475–490.
- [8] Shneeweiss H. Consistent estimation of a regression with error in the variables // Metrika. 1976. Vol. 23. P. 101–115.
- [9] Shiryaev A.N., Spokoiny V.G. Statistical Experiments and Decisions. Asymptotic Theory // Singapore, New Jersey, London, Hong Kong: World Scientific, 2000.
- [10] Vasiliev V. A Truncated Estimation Method with Guaranteed Accuracy // Annals of Institute of Statistical Mathematics. 2014. Vol. 66. P. 141–163.
- [11] Vasiliev V.A., Konev V.V. The Sequential Parameters Estimation of Dynamic Systems with Noisy Observations // Izvestia Acad. USSR, Technical Cybernetics. 1982. Vol. 6. P. 145–154.

Буркатовская Ю. Б., Васильев В. А. (Томский политехнический университет, Томский государственный университет, Томск, 2019) **Усеченное оценивание параметра процесса $AR(1)$ с аддитивным шумом.**

Аннотация. В данной работе представлена усеченная оценка параметра динамики устойчивого процесса $AR(1)$ по наблюдениям с аддитивным шумом. Оценка построена по выборке фиксированного объема и имеет известную верхнюю границу среднеквадратического отклонения. Рассмотрены случаи известной и неизвестной дисперсии шумов в наблюдениях.

Ключевые слова: процесс авторегрессии, зашумленные наблюдения, фиксированный размер выборки, гарантированная точность.

On Hypotheses Testing for Poisson Processes: Regular and Singular Cases

Dachian S., Kutoyants Yu., Yang L.

University of Lille, Lille, France

Le Mans University, Le Mans, France

Tomsk State University, Tomsk, Russia

Fuzhou University, Fuzhou, China

e-mail: Serguei.Dachian@univ-lille.fr

Abstract

We consider the problem of choosing between two hypotheses concerning an unknown parameter of an observed Poissonian signal. The first hypothesis is simple and the second one is local one-sided composite. The realization and the properties of the proposed tests turn out to be very different depending on the (non-)regularity of the intensity function of the signal. We study three different cases: regular intensity, intensity having a cusp type singularity and discontinuous intensity (change-point type singularity). In each case, we describe the choice of the thresholds and the power functions of the proposed tests both theoretically and by numerical simulations.

Keywords: hypothesis testing, inhomogeneous Poisson processes, asymptotic theory, composite alternatives, regular models, cusp type singularity, change-point type singularity.

We consider the situation when the observations are n independent realizations $X_j = \{X_j(t), t \in [0, \tau]\}$, $j = 1, \dots, n$, of an inhomogeneous Poisson process on the interval $[0, \tau]$ (the constant $\tau > 0$ is supposed to be known) with intensity function $\lambda(\vartheta, t)$, where $\vartheta \in \Theta = [\vartheta_0, b)$ is some unknown parameter. We denote the observation $X^n = \{X_1, \dots, X_n\}$.

Let us note that this model of observation is equivalent to observation on $[0, n\tau]$ of a single realization of a Poisson process with periodic intensity function (given on the first period by $\lambda(\vartheta, t)$), as well as to observation on $[0, \tau]$ of a single realization of a Poisson process with intensity function $n\lambda(\vartheta, t)$, $0 \leq t \leq \tau$.

We suppose that the intensity function $\lambda(\vartheta, t)$ is separated from zero. Then the measures corresponding to different values of ϑ are equivalent (and also equivalent to the case $\lambda \equiv 1$). The likelihood of our model (with respect to n independent standard Poisson processes of intensity 1) is given by (see, for example, [4, 5])

$$L(\vartheta, X^n) = \exp \left\{ \sum_{j=1}^n \int_0^\tau \ln \lambda(\vartheta, t) dX_j(t) - n \int_0^\tau [\lambda(\vartheta, t) - 1] dt \right\}.$$

We also introduce the likelihood ratio

$$L(\vartheta, \vartheta_0, X^n) = \frac{L(\vartheta, X^n)}{L(\vartheta_0, X^n)}.$$

We have to choose between two hypotheses

$$\begin{aligned} \mathcal{H}_0 &: \vartheta = \vartheta_0, \\ \mathcal{H}_1 &: \vartheta > \vartheta_0. \end{aligned}$$

A (randomized) test $\bar{\psi}_n = \bar{\psi}_n(X^n)$ is defined as the probability to accept the hypothesis $\mathcal{H}_1$.

We denote $\mathcal{K}_\varepsilon$ the class of tests $\bar{\psi}_n$ of asymptotic size $\varepsilon \in [0, 1]$:

$$\mathcal{K}_\varepsilon = \left\{ \bar{\psi}_n : \lim_{n \rightarrow \infty} \mathbf{E}_{\vartheta_0} \bar{\psi}_n(X^n) = \varepsilon \right\}.$$

The power function of $\bar{\psi}_n$ is given $\beta(\bar{\psi}_n, \vartheta) = \mathbf{E}_\vartheta \bar{\psi}_n(X^n)$, $\vartheta > \vartheta_0$.

Since for any reasonable (consistent) test, the power function converges to 1, in order to compare the tests asymptotically we place ourselves in the framework of *Pitman's approach* and consider *close* (or *contiguous*) alternatives. More precisely, we put $\vartheta = \vartheta_0 + \varphi_n u$, where $\varphi_n = \varphi_n(\vartheta_0) \searrow 0$ and $u \in \mathbb{U}_n^+ = [0, \varphi_n^{-1}(b - \vartheta_0))$. The rate φ_n must be chosen so that the normalized likelihood ratio process

$$Z_n(u) = \frac{L(\vartheta_0 + \varphi_n u, X^n)}{L(\vartheta_0, X^n)}, \quad u \in \mathbb{U}_n^+,$$

has a non degenerate limit (on $\mathbb{R}_+$).

Now, we have to choose between

$$\begin{aligned} \mathcal{H}_0 &: u = 0, \\ \mathcal{H}_1 &: u > 0, \end{aligned}$$

and the power function of $\bar{\psi}_n$ is $\beta(\bar{\psi}_n, u) = \mathbf{E}_{\vartheta_0 + \varphi_n u} \bar{\psi}_n(X^n)$, $u \in \mathbb{U}_n^+$.

Definition 1. A test $\psi_n^*(X^n) \in \mathcal{K}_\varepsilon$ is called locally asymptotically uniformly most powerful (LAUMP) in the class $\mathcal{K}_\varepsilon$ if its power function $\beta(\psi_n^*, u)$ satisfies

$$\lim_{n \rightarrow \infty} \inf_{0 < u \leq K} [\beta(\psi_n^*, u) - \beta(\bar{\psi}_n, u)] \geq 0$$

for any test $\bar{\psi}_n(X^n) \in \mathcal{K}_\varepsilon$ and any $K > 0$.

In the following section we present the results obtained in [2].

Regular case. We suppose that the following two conditions hold:

Smoothness: the function $\lambda(\vartheta, t)$ is two times continuously differentiable w.r.t. ϑ in $\vartheta = \vartheta_0$, and the Fisher information in ϑ_0 is positive:

$$I(\vartheta_0) = \int_0^\tau \frac{\dot{\lambda}(\vartheta_0, t)^2}{\lambda(\vartheta_0, t)} dt > 0.$$

Here $\dot{\lambda}$ denotes the derivative of λ w.r.t. ϑ .

Distinguishability: for any $\nu > 0$, we have

$$\inf_{\vartheta > \vartheta_0 + \nu} \int_0^\tau \left(\sqrt{\lambda(\vartheta, t)} - \sqrt{\lambda(\vartheta_0, t)} \right)^2 dt > 0.$$

We denote $\varphi_n = \varphi_n(\vartheta_0) = \frac{1}{\sqrt{nI(\vartheta_0)}}$. As shows the following theorem, under the above conditions, the model is *locally asymptotically normal (LAN)*.

Theorem 1. *Under the above conditions, uniformly on $u \in [0, K)$ (for any $K > 0$), we have*

$$Z_n(u) = \frac{L(\vartheta_0 + \varphi_n u, X^n)}{L(\vartheta_0, X^n)} = \exp \left\{ u \Delta_n(\vartheta_0, X^n) - \frac{u^2}{2} + r_n \right\},$$

where

$$\Delta_n(\vartheta_0, X^n) = \frac{1}{\sqrt{nI(\vartheta_0)}} \sum_{j=1}^n \int_0^\tau \frac{\dot{\lambda}(\vartheta_0, t)}{\lambda(\vartheta_0, t)} [dX_j(t) - \lambda(\vartheta_0, t) dt]$$

$$\implies \Delta \sim \mathcal{N}(0, 1) \quad (\text{convergence in law under } \vartheta_0)$$

and $r_n = r_n(\vartheta_0, u, X^n) \xrightarrow{\mathbf{P}} 0$ (convergence in probability under ϑ_0).

This theorem implies the asymptotic behavior of the likelihood ratio under the null hypothesis and under the alternative.

Corollary 1. *Under $\vartheta = \vartheta_0$, the limit of the normalized likelihood ratio $Z_n(u)$, $u \in \mathbb{U}_n^+$, is*

$$Z(u) = \exp \left\{ u \Delta - \frac{u^2}{2} \right\}, \quad u \geq 0.$$

Under $\vartheta = \vartheta_0 + \varphi_n u_$ (with fixed $u_* > 0$), the limit of the normalized likelihood ratio $Z_n(u)$, $u \in \mathbb{U}_n^+$, is*

$$Z(u, u_*) = \exp \left\{ u (\Delta + u_*) - \frac{u^2}{2} \right\}, \quad u \geq 0.$$

Now, we can deduce the thresholds and the properties of the usual tests.

Score function test (SFT). The SFT is defined by

$$\psi_n^*(X^n) = \mathbf{1}_{\{\Delta_n(\vartheta_0, X^n) > z_\varepsilon\}},$$

where z_ε is the $(1 - \varepsilon)$ -quantile of the standard normal distribution.

The test ψ_n^* belongs to $\mathcal{K}_\varepsilon$, for its power function we have

$$\beta(\psi_n^*, u_*) \longrightarrow \beta^*(u_*) = \mathbf{P}\{\Delta > z_\varepsilon - u_*\},$$

and it is LAUMP.

Let us briefly explain the idea behind LAUMP property of the test. It follows from the fact that its limit power $\beta^*(u_*)$ coincides with the limit power of the (most powerful) Neyman-Pearson test in the problem of testing $u = 0$ against $u = u_*$ (for any fixed u_*).

Wald's test (WT). First, let us recall that the *maximum likelihood estimator (MLE)* $\hat{\vartheta}_n$ is given by

$$\hat{\vartheta}_n = \operatorname{argsup}_{\vartheta \in [\vartheta_0, b)} L(\vartheta, \vartheta_0, X^n). \quad (1)$$

The WT is based on the MLE $\hat{\vartheta}_n$ and is defined by

$$\psi_n^\circ(X^n) = \mathbf{1}_{\{\varphi_n^{-1}(\hat{\vartheta}_n - \vartheta_0) > g_\varepsilon\}}. \quad (2)$$

Taking $g_\varepsilon = z_\varepsilon$, the WT belongs to $\mathcal{K}_\varepsilon$, its power function converges to β^* , and it is LAUMP.

General likelihood ratio test (GLRT). The GLRT is defined by

$$\hat{\psi}_n(X^n) = \mathbf{1}_{\{Q(X^n) > h_\varepsilon\}} \quad (3)$$

with

$$Q(X^n) = \sup_{\vartheta \in [\vartheta_0, b)} L(\vartheta, \vartheta_0, X^n). \quad (4)$$

Taking $h_\varepsilon = \exp\{z_\varepsilon^2/2\}$, the GLRT belongs to $\mathcal{K}_\varepsilon$, its power function converges to β^* , and it is LAUMP.

First Bayes test (BT1). First, let us recall that the *Bayes estimator (BE)* $\tilde{\vartheta}_n$ for a *priori* density $p(\theta)$ and quadratic loss function is given by

$$\tilde{\vartheta}_n = \frac{\int_{\vartheta_0}^b \theta p(\theta) L(\theta, \vartheta_0, X^n) d\theta}{\int_{\vartheta_0}^b p(\theta) L(\theta, \vartheta_0, X^n) d\theta}. \quad (5)$$

The (BT1) is a Wald-type test based on the BE $\tilde{\vartheta}_n$ and is defined by

$$\tilde{\psi}_n(X^n) = \mathbf{1}_{\{\varphi_n^{-1}(\tilde{\vartheta}_n - \vartheta_0) > k_\varepsilon\}}. \quad (6)$$

Taking

$$k_\varepsilon = \frac{f(z_\varepsilon)}{F(z_\varepsilon)} + z_\varepsilon,$$

where f and F are the density and the distribution function of Δ , the BT1 belongs to $\mathcal{K}_\varepsilon$, its power function converges to β^* , and it is LAUMP.

Second Bayes test (BT2). For a test $\bar{\psi}_n$, the *mean (or averaged) power* is

$$\beta(\bar{\psi}_n) = \int_{\vartheta_0}^b \beta(\bar{\psi}_n, \theta) p(\theta) d\theta.$$

The BT2 is the test which maximizes the mean power and is defined by

$$\tilde{\psi}_n^*(X^n) = \mathbf{1}_{\{R(X^n) > m_\varepsilon\}} \quad (7)$$

with

$$R(X^n) = \frac{1}{p(\vartheta_0)\varphi_n} \int_{\vartheta_0}^b L(\theta, \vartheta_0, X^n) p(\theta) d\theta. \quad (8)$$

Taking $m_\varepsilon = \frac{F(z_\varepsilon)}{f(z_\varepsilon)}$, the test BT2 belongs to $\mathcal{K}_\varepsilon$, its power function converges to β^* , and it is LAUMP.

In the following two sections we present the results obtained in [3].

First singular case: cusp. Now, we suppose that the intensity function has the form

$\lambda(\vartheta, t) = a|t - \vartheta|^\kappa + h(t)$, $0 \leq t \leq \tau$, $\vartheta \in \Theta = [\vartheta_0, b) \subset (0, \tau)$, where $\kappa \in (0, \frac{1}{2})$, $a \neq 0$ and the function h is continuously differentiable.

Denoting

$$\varphi_n = n^{-\frac{1}{2\kappa+1}} \Gamma_{\vartheta_0}^{-\frac{2}{2\kappa+1}},$$

where

$$\Gamma_{\vartheta_0}^2 = \frac{2a^2 B(\kappa+1, \kappa+1)}{h(\vartheta_0)} \left[\frac{1}{\cos(\pi\kappa)} - 1 \right],$$

we have the following asymptotic behavior of the likelihood ratio under the null hypothesis and under the alternative.

Theorem 2. *Under $\vartheta = \vartheta_0$, the limit of the normalized likelihood ratio $Z_n(u)$, $u \in \mathbb{U}_n^+$, is*

$$Z(u) = \exp \left\{ W^{\kappa+\frac{1}{2}}(u) - \frac{|u|^{2\kappa+1}}{2} \right\}, \quad u \geq 0,$$

where W^H denotes a fractional Brownian motion of Hurst parameter H . Let us note that if ϑ_0 was not a border of Θ , the process Z would be defined for all $u \in \mathbb{R}$.

Under $\vartheta = \vartheta_0 + \varphi_n u_*$ (with fixed $u_* > 0$), the limit of the normalized likelihood ratio $Z_n(u)$, $u \in \mathbb{U}_n^+$, is

$$Z(u, u_*) = \exp \left\{ W^{\kappa+\frac{1}{2}}(u) - \frac{|u - u_*|^{2\kappa+1}}{2} + \frac{|u_*|^{2\kappa+1}}{2} \right\}, \quad u \geq 0.$$

The SFT no longer exist in this singular situation, but the thresholds and the properties of the other four tests can be deduced from the above theorem.

Wald's test (WT). As before, the WT ψ_n° is defined by (1)–(2), but now, in order for ψ_n° belong to $\mathcal{K}_\varepsilon$, one has to chose g_ε as solution of

$$\mathbf{P}\{\hat{u} > g_\varepsilon\} = \varepsilon, \quad \hat{u} = \underset{u \geq 0}{\operatorname{argsup}} Z(u),$$

while its power function converges to β° given by

$$\beta^\circ(u_*) = \mathbf{P}\{\hat{u}_* > g_\varepsilon - u_*\}, \quad \hat{u}_* = \underset{u \geq -u_*}{\operatorname{argsup}} Z(u),$$

and/or

$$\beta^\circ(u_*) = \mathbf{P}\{\hat{u}^* > g_\varepsilon\}, \quad \hat{u}^* = \underset{u \geq 0}{\operatorname{argsup}} Z(u, u_*).$$

General Likelihood Ratio test (GLRT). As before, the GLRT $\hat{\psi}_n$ is defined by (3)–(4), but now, in order for $\hat{\psi}_n$ belong to $\mathcal{K}_\varepsilon$, one has to

chose h_ε as solution of

$$\mathbf{P}\{\hat{Z} > h_\varepsilon\} = \varepsilon, \quad \hat{Z} = \sup_{u \geq 0} Z(u),$$

while its power function converges to $\hat{\beta}$ given by

$$\hat{\beta}(u_*) = \mathbf{P}\{\hat{Z}_* > h_\varepsilon\}, \quad \hat{Z}_* = (Z(-u_*))^{-1} \sup_{u \geq -u_*} Z(u),$$

and/or

$$\hat{\beta}(u_*) = \mathbf{P}\{\hat{Z}^* > h_\varepsilon\}, \quad \hat{Z}^* = \sup_{u \geq 0} Z(u, u_*).$$

First Bayes test (BT1). As before, the BT1 $\tilde{\psi}_n$ is defined by (5)–(6), but now, in order for $\tilde{\psi}_n$ belong to $\mathcal{K}_\varepsilon$, one has to chose k_ε as solution of

$$\mathbf{P}\{\tilde{u} > k_\varepsilon\} = \varepsilon, \quad \tilde{u} = \frac{\int_0^\infty v Z(v) dv}{\int_0^\infty Z(v) dv},$$

while its power function converges to $\tilde{\beta}$ given by

$$\tilde{\beta}(u_*) = \mathbf{P}\{\tilde{u}_* > k_\varepsilon - u_*\}, \quad \tilde{u}_* = \frac{\int_{-u_*}^\infty v Z(v) dv}{\int_{-u_*}^\infty Z(v) dv},$$

and/or

$$\tilde{\beta}(u_*) = \mathbf{P}\{\tilde{u}^* > k_\varepsilon\}, \quad \tilde{u}^* = \frac{\int_0^\infty v Z(v, u_*) dv}{\int_0^\infty Z(v, u_*) dv}.$$

Second Bayes test (BT2). As before, the BT2 $\tilde{\psi}_n^*$ is defined by (7)–(8), but now, in order for $\tilde{\psi}_n^*$ belong to $\mathcal{K}_\varepsilon$, one has to chose m_ε as solution of

$$\mathbf{P}\{\tilde{Z} > m_\varepsilon\} = \varepsilon, \quad \tilde{Z} = \int_0^\infty Z(v) dv,$$

while its power function converges to $\tilde{\beta}^*$ given by

$$\tilde{\beta}^*(u_*) = \mathbf{P}\{\tilde{Z}_* > m_\varepsilon\}, \quad \tilde{Z}_* = Z(-u_*)^{-1} \int_{-u_*}^\infty Z(v) dv,$$

and/or

$$\tilde{\beta}^*(u_*) = \mathbf{P}\{\tilde{Z}^* > m_\varepsilon\}, \quad \tilde{Z}^* = \int_0^\infty Z(v, u_*) dv.$$

For $\kappa = 0.4$, some thresholds of GLRT, WT and BT1 obtained by numerical simulations are given in Table 1. The corresponding limit power functions, as well as the limit Neyman-Pearson envelop are presented in Figure 1.

We can observe that the limit power function of the GLRT is the closest to the limit Neyman-Pearson envelope for small values of u , while the limit power function of the BT1 is the one that tends to 1 (as u becomes large) the most quickly. Finally, we need to say that all

ε	0.01	0.05	0.10	0.2	0.4	0.5
$\ln h_\varepsilon$	2.959	1.641	1.081	0.559	0.159	0.068
g_ε	3.041	1.996	1.521	0.950	0.333	0.166
k_ε	2.864	2.0776	1.720	1.365	1.005	0.885

Table 1. Thresholds in the cusp case

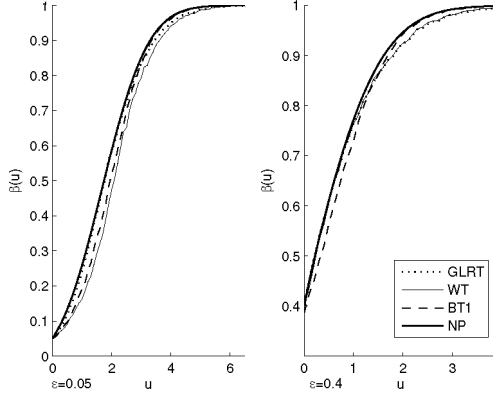


Figure 1. Limit power functions in the cusp case

these limit power functions are perceptibly below the limit Neyman-Pearson envelope, and that the existence of an asymptotically optimal test remains an open question.

Second singular case: change-point. Now, we suppose that the intensity function has the form

$$\lambda(\vartheta, t) = \lambda(t - \vartheta), \quad 0 \leq t \leq \tau, \quad \vartheta \in \Theta = [\vartheta_0, b) \subset (0, \tau),$$

where the function $\lambda(s)$, $s \in [-b, \tau - \vartheta_0]$, has a jump in a given point $t_* \in (-\vartheta_0, \tau - b)$ and is continuously differentiable elsewhere. We denote $\lambda_+ = \lambda(t_* +)$, $\lambda_- = \lambda(t_* -)$ and $\rho = \frac{\lambda_-}{\lambda_+} \neq 1$.

The results are similar to those of the cusp case, with the difference that the rate is now

$$\varphi_n = \frac{1}{n\lambda_+},$$

while the limit likelihood ratio processes are now defined by

$$Z(u) = \begin{cases} \exp\{\ln \rho x_*(u) - (\rho - 1)u\}, & u \geq 0, \\ \exp\{-\ln \rho x_\rho(-u) - (\rho - 1)u\}, & u \leq 0, \end{cases}$$

where $(x_*(u), u \geq 0)$ and $(x_\rho(u), u \geq 0)$ are independent Poisson

process of intensities 1 and ρ respectively (for a more detailed study of this process we refer to [1]), and by

$$Z(u, u_*) = \exp\{\ln \rho x_*(u, u_*) - (\rho - 1)u\}, \quad u \geq 0,$$

where $(x_*(u, u_*), u \geq 0)$ is a Poisson process of intensity function $\mu(u, u_*) = \rho \mathbf{1}_{\{u < u_*\}} + \mathbf{1}_{\{u \geq u_*\}}, u \geq 0$.

For $\rho = 3$, some thresholds of GLRT, WT and BT1 obtained by numerical simulations are given in Table 2. The corresponding limit power functions, as well as the limit Neyman-Pearson envelop are presented in Figure 2. The observations made in the cusp case apply here equally.

ε	0.01	0.05	0.10	0.20	0.40	0.50
$\ln h_\varepsilon$	4.242	2.607	1.922	1.120	0.573	0.191
g_ε	5.990	3.556	2.078	1.045	0.329	0.099
k_ε	6.669	3.937	2.983	2.132	1.402	1.196

Table 2. Thresholds in the change-point case

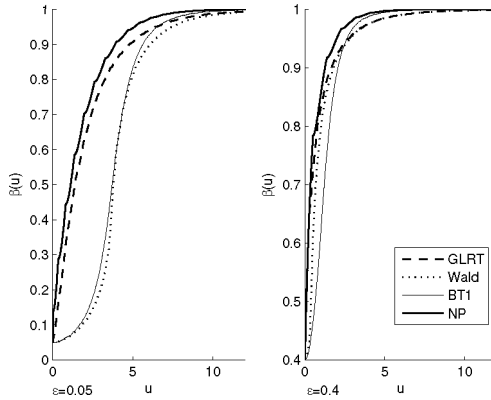


Figure 2. Limit power functions in the change-point case

References

- [1] Dachian S. On limiting likelihood ratio processes of some change-point type statistical models // Journal of Statistical Planning and Inference. 2010. Vol. 140, No. 9. P. 2682–269,.

- [2] Dachian S., Kutoyants Yu. A., Yang L. On hypothesis testing for Poisson processes. Regular case // Communications in Statistics — Theory and Methods. 2016. Vol. 45, No. 23. P. 6816–6832.
- [3] Dachian S., Kutoyants Yu. A., Yang L. On hypothesis testing for Poisson processes. Singular cases // Communications in Statistics — Theory and Methods. 2016. Vol. 45, No. 23. P. 6833–6859.
- [4] Kutoyants Yu. A. Parameter Estimation for Stochastic Processes. Armenian Academy of Sciences, Yerevan, 1980 (in Russian), translation of revised version, Heldermann-Verlag, Berlin, 1984.
- [5] Kutoyants Yu. A. Statistical Inference for Spatial Poisson Processes. Lect. Notes Statist. 134, Springer-Verlag, New York, 1998.

Дашян С., Кутоянц Ю., Янг Л. (Университет Лилля, Лилль, Университет Ле-Мана, Ле-Ман, Франция, Томский государственный университет, Томск, Россия, Университет Фучжоу, Фучжоу, Китай, 2019) **О проверке гипотез для пуассоновских процессов: регулярный и сингулярный случаи.**

Аннотация. Рассматривается проблема выбора между двумя гипотезами относительно неизвестного параметра наблюдаемого пуассоновского сигнала. Первая гипотеза проста, а вторая является сложной локальной односторонней. Реализация и свойства предложенных тестов оказываются очень разными в зависимости от (не)регулярности функции интенсивности сигнала. Изучаем три различных случая: регулярную интенсивность, интенсивность, имеющую особенность типа каспа и разрывную интенсивность (особенность типа точки изменения). В каждом случае описываем выбор порогов и степенных функций предлагаемых испытаний как теоретически, так и путем численного моделирования.

Ключевые слова: проверка гипотез, неоднородные пуассоновские процессы, асимптотическая теория, сложные альтернативы, регулярные модели, особенность типа касп, особенность типа точки переключения.

Sharp Oracle Inequalities for Drift Estimation Problems from Discrete Data ¹

Galtchouk L., Pergamenschchikov S.

Université de Rouen Normandie, Rouen, France

Tomsk State University, Tomsk

e-mail: serge.pergamenchchikov@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

$$dy_t = S(y_t) dt + b(y_t) dW_t, \quad 0 \leq t \leq T, \quad (1)$$

where $(W_t)_{t \geq 0}$ 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

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

where the frequency $\delta = \delta_T \in (0, 1)$ is some function of T which will be specified later and the sample size $N = N(T) = \lfloor T/\delta \rfloor$. 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 $(y_t)_{0 \leq t \leq T}$ is observed.

¹This 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

$$\Sigma_{\mathbf{L}, \mathbf{M}} = \left\{ S \in \mathbf{C}^1(\mathbb{R}) : \sup_{|x| \leq \mathbf{x}_*} (|S(x)| + |\dot{S}(x)|) \leq M, \right. \\ \left. -L \leq \inf_{|x| \geq \mathbf{x}_*} \dot{S}(x) \leq \sup_{|x| \geq \mathbf{x}_*} \dot{S}(x) \leq -1/L \right\}. \quad (3)$$

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

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

Now we set

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

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.

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

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

$$\mathbf{q}_{\vartheta}(x) = \left(\int_{\mathbb{R}} b^{-2}(z) e^{\tilde{S}(z)} dz \right)^{-1} b^{-2}(x) e^{\tilde{S}(x)}, \quad (7)$$

where $\tilde{S}(x) = 2 \int_0^x b^{-2}(v) S(v) dv$ (see, e.g., [3], Ch.4, 18, Th2).

We need the following condition for the observation frequency.

A₁) Assume that the frequency δ in the observations (2) is of the following form

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

where the function l_T is such that,

$$\lim_{T \rightarrow \infty} \frac{l_T}{\ln T} = +\infty. \quad (9)$$

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

$$\mathcal{R}(\hat{S}, S) = \mathbf{E}_{\vartheta} \int_{\mathbf{x}_0}^{\mathbf{x}_1} |\hat{S}(x) - S(x)|^2 dx, \quad (10)$$

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 $(z_k)_{0 \leq k \leq n}$ as

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

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

$$\lim_{T \rightarrow \infty} \frac{n(T)}{\sqrt{T}} = 1. \quad (12)$$

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

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

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

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

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.,

$$S(z_k) = \sum_{j=1}^n \theta_{j,n} \phi_j(z_k), \quad 1 \leq k \leq n, \quad (15)$$

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 $(\hat{S}_k, \tau_k)$ for $S(z_k)$ proposed in [2]. To this end we set

$$Y_k = \hat{S}_k \chi_{\{\tau_k \leq N\}}.$$

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

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

we can write the regression model, i.e.

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

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

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

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

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

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

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

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.

A₂) For any $a > 0$

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

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

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

$$\text{Err}_n(\lambda) = \|\widehat{S}_\lambda - S\|_n^2 \rightarrow \min.$$

In our case, the empirical squared error is equal to

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

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

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

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

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

Finally, we use the following cost function

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

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

$$\widehat{\lambda} = \arg\min_{\lambda \in \Lambda} J_n(\lambda) \quad \text{and} \quad \widehat{S}_* = \widehat{S}_{\widehat{\lambda}}. \quad (25)$$

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 $\hat{S}_*$ defined in (25) satisfies the following inequality

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

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

$$\lim_{T \rightarrow \infty} \frac{\Psi_T}{T^a} = 0. \quad (27)$$

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.

On Identification of Continuous Time Hidden Markov Processes: a Survey ¹

Kutoyants Yu. A.

Le Mans University, Le Mans, France,
Toms State University, Tomsk, Russia
e-mail: YuryKutoyants@univ-lemans.fr

Abstract

We consider several models of partially observed stochastic linear systems and discuss the problems of parameter estimation. We describe the asymptotic properties of the MLE, Bayes estimators and One-step MLE-processes in two different asymptotics: small noise and large samples. In all problems we show the consistency of these estimators, their limit distributions and define asymptotically efficient estimators.

Keywords: partially observed systems, parameter estimation, asymptotic properties, change-point problems.

Introduction. We consider the following model of observations

$$dX_t = f_t Y_t dt + \sigma_t dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T$$

where Y_t is hidden Gaussian process satisfying the SDE

$$dY_t = -a_t Y_t dt + b_t dV_t, \quad Y_0 = y_0, \quad 0 \leq t \leq T$$

The Wiener processes W_t and V_t are supposed always independent.

We suppose that in the case of diffusion processes the functions $f_t = f(\vartheta, t)$, $a_t = a(\vartheta, t)$, $b_t = b(\vartheta, t)$ depend on some finite-dimensional parameter ϑ and we have to estimate it by the observations $X^T = (X_t, 0 \leq t \leq T)$. We describe the properties of the MLE and BE in the asymptotic of *small noise* and *large samples*. Moreover we propose One-step MLE-processes of this parameter, which have several advantages w.r.t. MLE and BE.

In the case of diffusion processes the LR function is

$$L(\vartheta, X^T) = \exp \left\{ \int_0^T \frac{f(\vartheta, t) m(\vartheta, t)}{\sigma_t^2} dX_t - \int_0^T \frac{f(\vartheta, t)^2 m(\vartheta, t)^2}{2\sigma_t^2} dt \right\},$$

where $m(\vartheta, t) = m(t)$ is solution of the K-B filtration equations

$$dm(t) = -a(\vartheta, t) m(t) dt + \frac{\gamma(\vartheta, t) f(\vartheta, t)}{\sigma_t^2} [dX_t - f(\vartheta, t) m(t) dt],$$

$$\frac{\partial \gamma(\vartheta, t)}{\partial t} = -2a(\vartheta, t) \gamma(\vartheta, t) - \frac{\gamma(\vartheta, t)^2 f(\vartheta, t)^2}{\sigma_t^2} + b(\vartheta, t)^2,$$

¹This research was supported by The Tomsk State University competitiveness improvement programme grant No 8.1.18.2018.

with initial values $m(\vartheta, 0) = y_0$ and $\gamma(\vartheta, 0) = 0$. Recall that $\gamma(\vartheta, t) = \mathbf{E}_\vartheta (Y_t - m(\vartheta, t))^2$.

The MLE $\hat{\vartheta}$ is defined as solution of the equation

$$L(\hat{\vartheta}, X^T) = \sup_{\vartheta \in \Theta} L(\vartheta, X^T).$$

Suppose that ϑ is random vector with density $p(\vartheta)$, $\vartheta \in \Theta$. The BE $\hat{\vartheta}$ for quadratic loss function is

$$\tilde{\vartheta} = \frac{\int_{\Theta} \vartheta p(\vartheta) L(\vartheta, X^T) d\vartheta}{\int_{\Theta} p(\vartheta) L(\vartheta, X^T) d\vartheta}.$$

The One-step MLE-process ϑ_t^* , $\tau < t \leq T$ is given by

$$\vartheta_t^* = \bar{\vartheta}_\tau + \mathbb{I}_t(\bar{\vartheta}_\tau)^{-1} \int_\tau^t \frac{\dot{\mathcal{M}}(\bar{\vartheta}_\tau, s)}{\sigma_s^2} [dX_s - \mathcal{M}(\bar{\vartheta}_\tau, s) ds], \quad \tau < t \leq T,$$

where $\bar{\vartheta}_\tau = \bar{\vartheta}_\tau(X_t, 0 \leq t \leq \tau)$ preliminary estimator,

$\mathcal{M}(\vartheta, s) = f(\vartheta, s)m(\vartheta, s)$ and $\mathbb{I}_t(\bar{\vartheta}_\tau)$ is Fisher information.

We are given the system of linear SDEs

$$dX_t = f(\vartheta, t) Y_t dt + \varepsilon_1 \sigma(t) dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T,$$

$$dY_t = -a(\vartheta, t) Y_t dt + \varepsilon_2 b(\vartheta, t) dV_t, \quad Y_0 = y_0.$$

Here $f(\vartheta, t)$, $\sigma(t)$, $a(\vartheta, t)$, $b(\vartheta, t)$ are known functions, $\vartheta \in \Theta \subset \mathcal{R}^d$ is unknown parameter.

The equations are called “state equation” for Y^T (hidden Markov) and “observations equation” for X^T .

The asymptotics are described in terms of parameters $\varepsilon_1 \in (0, 1]$, $\varepsilon_2 \in (0, 1]$, $T > 0$.

To construct estimators of ϑ and to describe the properties of these estimators in different asymptotics.

There are several possibilities for diffusion processes

- Small noises, $\varepsilon_1 = \varepsilon_2 = \varepsilon \rightarrow 0$.
- Small noise, $\varepsilon_1 = \varepsilon \rightarrow 0, \varepsilon_2 = 1$.
- Small noise different rates, $\varepsilon_1 \rightarrow 0, \varepsilon_2 \rightarrow 0$ (not presented).
- Large samples, $T \rightarrow \infty$.

In all these situations we describe the MLE, BE and One-step MLE-processes.

Small noise $\varepsilon_1 = \varepsilon_2 = \varepsilon \rightarrow 0$. Model

$$dX_t = f(\vartheta, t) Y_t dt + \varepsilon \sigma(t) dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T,$$

$$dY_t = -a(\vartheta, t) Y_t dt + \varepsilon b(t) dV_t, \quad Y_0 = y_0 > 0.$$

Equations of K-B for $m(\vartheta, t)|_{\varepsilon=0} = y(\vartheta, t)$ and $\gamma_*(\vartheta, t) = \frac{\gamma(\vartheta, t)}{\varepsilon^2}$ are

$$\begin{aligned} \frac{\partial y(\vartheta, t)}{\partial t} &= - \left[a(\vartheta, t) \gamma_*(\vartheta, t) h_t(\vartheta)^2 \right] y(\vartheta, t) \\ &\quad + \gamma_*(\vartheta, t) h_t(\vartheta) h_t(\vartheta_0) y_t(\vartheta_0), \\ \frac{\partial \gamma_*(\vartheta, t)}{\partial t} &= -2a(\vartheta, t) \gamma_*(\vartheta, t) - \gamma_*(\vartheta, t)^2 h_t(\vartheta)^2 + b(t)^2, \end{aligned}$$

with initial values $y(\vartheta, t) = y_0$, $\gamma_*(\vartheta, 0) = 0$ and $h_t(\vartheta) = \frac{f(\vartheta, t)}{\sigma(t)}$

Introduce the function $M(\vartheta, t) = f(\vartheta, t) y(\vartheta, t)$, vector-function $\dot{M}(\vartheta, t)$, Fisher information matrix

$$\mathbb{I}(\vartheta_0) = \int_0^T \dot{M}(\vartheta_0, t) \dot{M}(\vartheta_0, t)^\top \sigma(t)^{-2} dt.$$

and identifiability condition: for any $\nu > 0$

$$\inf_{\|\vartheta - \vartheta_0\| > \nu} \int_0^T [f(\vartheta, t) y(\vartheta, t) - f(\vartheta_0, t) y_t(\vartheta_0)]^2 dt > 0.$$

The set $\vartheta \subset \mathcal{R}^2$ is open, convex and bounded. The function $p(\vartheta)$, $\vartheta \in \Theta$ is supposed to be continuous and positive. The functions $f(\vartheta, t)$ and $a(\vartheta, t)$ have two continuous bounded derivatives on ϑ . We have the Hajek-Le Cam lower bound

$$\lim_{\nu \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \sup_{|\vartheta - \vartheta_0| < \nu} \varepsilon^{-2} \mathbf{E}_\vartheta \|\bar{\vartheta}_\varepsilon - \vartheta\|^2 \geq \mathbf{E}_{\vartheta_0} \|\zeta\|^2, \quad \zeta \sim \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1}).$$

Theorem 1. [K. 1994] *Suppose that the conditions of regularity are fulfilled. Then the MLE and BE are consistent, asymptotically normal*

$$\frac{\hat{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \Rightarrow \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1}), \quad \frac{\tilde{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \Rightarrow \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1}),$$

for any $p > 0$

$$\mathbf{E}_{\vartheta_0} \left\| \frac{\hat{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \right\|^p \rightarrow \mathbf{E}_{\vartheta_0} \|\zeta\|^p, \quad \mathbf{E}_{\vartheta_0} \left\| \frac{\tilde{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \right\|^p \rightarrow \mathbf{E}_{\vartheta_0} \|\zeta\|^p$$

and the both estimators are asymptotically efficient.

Here $\zeta \sim \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1})$.

One-step MLE-process. Suppose that the functions $f(\vartheta, 0)$, $a(\vartheta, 0)$ have two continuous bounded derivatives on $\vartheta \in \Theta = (\alpha, \beta)$ and $t \in [0, T]$ and the following condition is fulfilled

$$\inf_{\vartheta \in \Theta} |f(\vartheta, 0)| > 0, \quad \inf_{\vartheta \in \Theta} \left| \dot{f}(\vartheta, 0) \right| > 0.$$

Introduce functions

$$x_t(\vartheta) = \int_0^t f(\vartheta, s) y_s(\vartheta) ds, \quad y_t(\vartheta) = y_0 \exp \left\{ - \int_0^t a(\vartheta, s) ds \right\}.$$

and

$$x_m(t) = \inf_{\vartheta \in \Theta} x_t(\vartheta) < \sup_{\vartheta \in \Theta} x_t(\vartheta) = x_M(t).$$

First we propose a consistent preliminary estimator by observations $X^{\tau_\varepsilon} = (X_t, 0 \leq t \leq \tau_\varepsilon)$, where $\tau_\varepsilon = \varepsilon^\delta, 0 < \delta < 2$.

Preliminary estimator $\bar{\vartheta}_{\tau_\varepsilon}$ we construct as follows

$$\bar{\vartheta}_{\tau_\varepsilon} = \alpha \mathbb{I}_{\{\mathbb{B}_\varepsilon^-\}} + \mu_\varepsilon \mathbb{I}_{\{\mathbb{B}_\varepsilon\}} + \beta \mathbb{I}_{\{\mathbb{B}_\varepsilon^+\}},$$

where the sets $\mathbb{B}_\varepsilon = \{x_m(\tau_\varepsilon) < X_{\tau_\varepsilon} < x_M(\tau_\varepsilon)\}$ $\mathbb{B}_\varepsilon^- = \{X_{\tau_\varepsilon} \leq x_m(\tau_\varepsilon)\}$, $\mathbb{B}_\varepsilon^+ = \{X_{\tau_\varepsilon} \geq x_M(\tau_\varepsilon)\}$ and μ_ε is solution of the equation $x_{\tau_\varepsilon}(\mu_\varepsilon) = X_{\tau_\varepsilon}$.

Theorem 2. [K., Zhou, 2019] *The estimator $\bar{\vartheta}_{\tau_\varepsilon}$ is consistent and for any $p > 0$*

$$\sup_{\vartheta_0 \in \Theta} \mathbf{E}_{\vartheta_0} |\bar{\vartheta}_{\tau_\varepsilon} - \vartheta_0|^p \leq C \left(\frac{\varepsilon}{\sqrt{\tau_\varepsilon}} \right)^p \rightarrow 0.$$

Define Fisher information in the case of observations $X_s, \tau \leq s \leq t$

$$\mathbb{I}_\tau(\vartheta, t) = \int_\tau^t \left[\frac{\dot{f}(\vartheta, s) y_s(\vartheta) + f(\vartheta, s) \dot{y}(\vartheta, s)}{\sigma(s)} \right]^2 ds,$$

and put $\mathbb{I}(\vartheta, t) = \mathbb{I}_0(\vartheta, t)$. The Hajek-Le Cam lower bound is

$$\lim_{\nu \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \sup_{|\vartheta - \vartheta_0| < \nu} \varepsilon^{-2} \mathbf{E}_{\vartheta} (\vartheta_{t,\varepsilon} - \vartheta)^2 \geq \mathbb{I}(\vartheta_0, t)^{-1}.$$

We call the estimator $\vartheta_{t,\varepsilon}^*$ asymptotically efficient if

$$\lim_{\nu \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \sup_{|\vartheta - \vartheta_0| < \nu} \varepsilon^{-2} \mathbf{E}_{\vartheta} (\vartheta_{t,\varepsilon}^* - \vartheta)^2 = \mathbb{I}(\vartheta_0, t)^{-1}.$$

One-step MLE-process $\vartheta_{t,\varepsilon}^*, \tau_\varepsilon < t \leq T$ is

$$\vartheta_{t,\varepsilon}^* = \bar{\vartheta}_{\tau_\varepsilon} + \frac{1}{\mathbb{I}_{\tau_\varepsilon}(\bar{\vartheta}_{\tau_\varepsilon}, t)} \int_{\tau_\varepsilon}^t \frac{\dot{\mathcal{M}}(\bar{\vartheta}_{\tau_\varepsilon}, s)}{\sigma(s)^2} [dX_s - M(\bar{\vartheta}_{\tau_\varepsilon}, s) ds],$$

where $\dot{\mathcal{M}}(\vartheta, s) = \dot{f}(\vartheta, s) y_s(\vartheta) + f(\vartheta, s) \dot{y}(\vartheta, s)$ and

$$M(\vartheta, s) = f(\vartheta, s) m(\vartheta, s).$$

Theorem 3. [K., Zhou, 2019] *Let $\delta \in (0, 1)$, then the estimator $\vartheta_{t,\varepsilon}^*, \tau_\varepsilon < t \leq T$ is consistent and asymptotically normal*

$$\varepsilon^{-1} (\vartheta_{t,\varepsilon}^* - \vartheta_0) \Rightarrow \zeta_t \sim \mathcal{N}(0, \mathbb{I}(\vartheta_0, t)^{-1}).$$

The moments converge: for any $p > 0$

$$\varepsilon^{-p} \mathbf{E}_{\vartheta_0} |\vartheta_{t,\varepsilon}^* - \vartheta_0|^p \rightarrow \mathbf{E}_{\vartheta_0} |\zeta_t|^p$$

and it is asymptotically efficient for all $t \in (0, T]$.

Small noise $\varepsilon_1 = \varepsilon \rightarrow 0, \varepsilon_2 = 1$. The corresponding system of equations is

$$dX_t = f(\vartheta, t) Y_t dt + \varepsilon \sigma(t) dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T,$$

$$dY_t = -a(\vartheta, t) Y_t dt + b(\vartheta, t) dV_t, \quad Y_0 = y_0 > 0,$$

We have to estimate ϑ and describe the properties of estimators as $\varepsilon \rightarrow 0$. The limit system is

$$x_t(\vartheta) = \int_0^t f(\vartheta, s) Y_s ds,$$

$$dY_t = -a(\vartheta, t) Y_t dt + b(\vartheta, t) dV_t, \quad Y_0 = y_0 > 0,$$

and we are in the situation, where it is possible to estimate ϑ by $x_t, 0 \leq t \leq T$ without error.

Introduce notation $S(\vartheta, t) = f(\vartheta, t) b(\vartheta, t)$

$$G(\vartheta, \vartheta_0) = \int_0^T \frac{[S(\vartheta, t) - S(\vartheta_0, t)]^2}{2S(\vartheta, t) \sigma(t)} dt,$$

$$\mathbb{I}_0(\vartheta) = \int_0^T \frac{S(\vartheta, t)}{2\sigma(t)} \left(\frac{\partial}{\partial \vartheta} \ln[S(\vartheta, t)] \right)^2 dt,$$

and condition: for any $\nu > 0$

$$\kappa = \inf_{\vartheta \in \Theta} \inf_{0 \leq t \leq T} f(\vartheta, t) > 0, \quad \inf_{0 \leq t \leq T} \sigma(t) > 0, \quad \inf_{\vartheta \in \Theta} \inf_{0 \leq t \leq T} b(\vartheta, t) > 0.$$

$$\inf_{\vartheta_0 \in \Theta} \inf_{|\vartheta - \vartheta_0| > \nu} G(\vartheta, \vartheta_0) > 0, \quad \inf_{\vartheta \in \Theta} \mathbb{I}_0(\vartheta) > 0.$$

The Hajek-Le Cam's bound is

$$\lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \sup_{|\vartheta - \vartheta_0| < \delta} \varepsilon^{-1} \mathbf{E}_{\vartheta} |\bar{\vartheta}_{\varepsilon} - \vartheta|^2 \geq \mathbb{I}_0(\vartheta_0)^{-1}.$$

Theorem 4. [K.2019] *The MLE $\hat{\vartheta}_{\varepsilon}$ and BE $\tilde{\vartheta}_{\varepsilon}$ are consistent, asymptotically normal*

$$\frac{\hat{\vartheta}_{\varepsilon} - \vartheta_0}{\sqrt{\varepsilon}} \Rightarrow \zeta \sim \mathcal{N}(0, \mathbb{I}_0(\vartheta_0)^{-1}), \quad \frac{\tilde{\vartheta}_{\varepsilon} - \vartheta_0}{\sqrt{\varepsilon}} \Rightarrow \zeta,$$

the moments converge: for any $p > 0$

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}_{\vartheta_0} \left| \frac{\hat{\vartheta}_{\varepsilon} - \vartheta_0}{\sqrt{\varepsilon}} \right|^p = \mathbf{E}_{\vartheta_0} |\zeta|^p, \quad \lim_{\varepsilon \rightarrow 0} \mathbf{E}_{\vartheta_0} \left| \frac{\tilde{\vartheta}_{\varepsilon} - \vartheta_0}{\sqrt{\varepsilon}} \right|^p = \mathbf{E}_{\vartheta_0} |\zeta|^p$$

and the both estimators are asymptotically efficient.

Hidden Ergodic O-U Process, $T \rightarrow \infty$. Consider the model of observations

$$dX_t = c(\vartheta) Y_t dt + \sigma dW_t, \quad X_0, \quad 0 \leq t \leq T,$$

$$dY_t = -a(\vartheta) Y_t dt + b(\vartheta) dV_t, \quad Y_0,$$

Here W_t, V_t independent Wiener processes, $a(\cdot), b(\cdot), c(\cdot)$ are known smooth functions and constant $\sigma > 0$. $X^T = (X_t, 0 \leq t \leq T)$ are observations and Y^T is hidden Markov process. Denote $m_t(\vartheta) = \mathbf{E}_{\vartheta}(Y_t | X_s, 0 \leq s \leq t)$. Then Kalman-Bucy filtration is

$$dm_t(\vartheta) = -a(\vartheta) m_t(\vartheta) dt + \frac{c(\vartheta) \gamma_t(\vartheta)}{\sigma^2} [dX_t - c(\vartheta) m_t(\vartheta) dt],$$

$$\frac{\partial \gamma_t(\vartheta)}{\partial t} = -2a(\vartheta) \gamma_t(\vartheta) - \frac{c(\vartheta)^2 \gamma_t(\vartheta)^2}{\sigma^2} + b(\vartheta)^2.$$

Introduce Fisher information

$$\mathbb{I}(\vartheta) = \frac{\dot{a}(\vartheta)^2}{2a(\vartheta)} - \frac{2\dot{a}(\vartheta)\dot{r}(\vartheta)}{a(\vartheta) + r(\vartheta)} + \frac{\dot{r}(\vartheta)^2}{2r(\vartheta)},$$

where

$$r(\vartheta) = \sqrt{a(\vartheta)^2 + \frac{b(\vartheta)^2 c(\vartheta)^2}{\sigma^2}}.$$

Conditions

$$\inf_{|\vartheta - \vartheta_0| > \nu} (|a(\vartheta) - a(\vartheta_0)| + |r(\vartheta) - r(\vartheta_0)|) > 0,$$

$$\inf_{\vartheta} (|\dot{a}(\vartheta)| + |\dot{r}(\vartheta)|) > 0.$$

Properties of MLE and BE

Theorem 5. [K. 1984, 2004] *The MLE $\hat{\vartheta}_T$ and the BE $\tilde{\vartheta}_T$ are consistent, and asymptotically normal, i.e.,*

$$\sqrt{T}(\hat{\vartheta}_T - \vartheta_0) \Rightarrow \zeta \sim \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1}), \quad \sqrt{T}(\tilde{\vartheta}_T - \vartheta_0) \Rightarrow \zeta.$$

Moreover the moments converge, i.e., for any $p > 0$

$$T^{\frac{p}{2}} \mathbf{E}_{\vartheta_0} \left| \hat{\vartheta}_T - \vartheta_0 \right|^p \rightarrow \mathbf{E}_{\vartheta_0} |\zeta|^p, \quad T^{\frac{p}{2}} \mathbf{E}_{\vartheta_0} \left| \tilde{\vartheta}_T - \vartheta_0 \right|^p \rightarrow \mathbf{E}_{\vartheta_0} |\zeta|^p$$

and both estimators are asymptotically efficient.

One-step MLE-process. We are given a partially observed linear system, defined by the equations

$$dX_t = fY_t dt + \sigma dW_t, \quad X_0 = 0,$$

$$dY_t = -aY_t dt + b dV_t, \quad Y_0 = \xi,$$

The model is defined by the parameters $f \neq 0, \sigma^2 \neq 0, a > 0, b \neq 0$. Note that σ^2 can be estimated without error and the process X_t can be written as

$$X_t = fb \int_0^t \int_0^s e^{-a(s-r)} dV_r ds + \sigma W_t + o(1).$$

Hence the parameters f and b can not be estimated separately. We consider one-dimensional parameters $\vartheta = f, \vartheta = a$ and $\vartheta = b$ and then we discuss how to estimate $\vartheta = (f, a)$ and $\vartheta = (a, b)$. where $a \neq 0, \sigma \neq 0, b \neq 0$ and $f > 0$ are constants.

Suppose that the unknown parameter is $\vartheta = a$. The random process $m(\vartheta, \cdot)$ is solution of the Kalman-Bucy filtering equations

$$dm(\vartheta, t) = - \left[\vartheta + \frac{\gamma(\vartheta, t) f^2}{\sigma^2} \right] m(\vartheta, t) dt + \frac{\gamma(\vartheta, t) f}{\sigma^2} dX_t,$$

$$\frac{\partial \gamma(\vartheta, t)}{\partial t} = -2\vartheta \gamma(\vartheta, t) - \frac{\gamma(\vartheta, t)^2 f^2}{\sigma^2} + b^2, \quad \gamma(\vartheta, 0) = d^2.$$

We introduce One-step MLE-process $\vartheta_t^*, \tau \leq t \leq T$, which has two advantages. First, its numerical calculation is much more simple than that of the MLE and, second, this estimator has a recurrent structure and can be used for the joint estimation of the hidden process Y_t and

the parameter ϑ .

Preliminary estimator is constructed using an asymptotically negligible amount of the observations $X^K = (X_t, 0 \leq t \leq K)$, where $K = T^\delta$, $\delta \in (1/2, 1)$.

Introduce the statistic $\mathbb{S}_K$ and the function $\Phi(\vartheta)$, $\vartheta \in \Theta$:

$$\mathbb{S}_K = \frac{1}{K} \sum_{k=1}^K [X_k - X_{k-1}]^2 \longrightarrow \Phi(\vartheta) = \frac{a^2 b^2}{\vartheta^3} [e^{-\vartheta} - 1 + \vartheta] + \sigma^2.$$

In the cases $\vartheta = b$ and $\vartheta = a$ the counterparts of the latter function are

$$\Phi_*(\vartheta) = \frac{a^2 \vartheta^2}{f^3} [e^{-f} - 1 + f] + \sigma^2, \quad \hat{\Phi}(\vartheta) = \frac{\vartheta^2 b^2}{f^3} [e^{-f} - 1 + f] + \sigma^2$$

respectively.

Note that the functions $\Phi(\vartheta)$, $\Phi_*(\vartheta)$, $\hat{\Phi}(\vartheta)$, $\alpha < \vartheta < \beta$ are strictly monotone. Define the preliminary estimator of a as $\bar{\vartheta}_K$, base the observations X^K :

$$\bar{\vartheta}_K = \vartheta_K^* \mathbb{I}_{\{\mathcal{A}_K\}} + \alpha \mathbb{I}_{\{\mathcal{A}_K^-\}} + \beta \mathbb{I}_{\{\mathcal{A}_K^+\}}.$$

Here ϑ_K^* is the root of equation $\hat{\Phi}(\vartheta_K^*) = \mathbb{S}_K$ and $\mathcal{A}_K, \mathcal{A}_K^-, \mathcal{A}_K^+$ are the sets

$$\begin{aligned} \mathcal{A}_K &= \left\{ \omega : \hat{\Phi}(\alpha) < \mathbb{S}_K < \hat{\Phi}(\beta) \right\}, \quad \mathcal{A}_K^- = \left\{ \omega : \mathbb{S}_K \leq \hat{\Phi}(\alpha) \right\}, \\ \mathcal{A}_K^+ &= \left\{ \omega : \mathbb{S}_K \geq \hat{\Phi}(\beta) \right\}. \end{aligned}$$

Proposition 1. *The estimator $\bar{\vartheta}_K$ is consistent, uniformly on compacts $[\bar{\alpha}, \bar{\beta}] \subset \Theta$, and*

$$\sup_{\vartheta_0 \in \Theta} \mathbf{E}_{\vartheta_0} |\bar{\vartheta}_K - \vartheta_0|^2 \leq \frac{C}{K}$$

with some constant $C > 0$.

The Fisher information for this model of observations is

$$\mathbb{I}(\vartheta) = \frac{1}{2\vartheta} - \frac{2\dot{r}(\vartheta)}{r(\vartheta) + \vartheta} + \frac{\dot{r}(\vartheta)^2}{2r(\vartheta)}.$$

and derivative processes $\dot{m}(\vartheta, t)$ satisfies the equation

$$\begin{aligned} d\dot{m}(\vartheta, t) &= - \left[\vartheta + \frac{\gamma(\vartheta) f^2}{\sigma^2} \right] \dot{m}(\vartheta, t) dt + \frac{\dot{\gamma}(\vartheta) f}{\sigma^2} dX_t \\ &\quad - \left[1 + \frac{\dot{\gamma}(\vartheta) f^2}{\sigma^2} \right] m(\vartheta, t) dt. \end{aligned}$$

The One-step MLE-process $\vartheta_{t,T}^*, T^\delta < t \leq T$ is introduced as follows

$$\vartheta_{t,T}^* = \bar{\vartheta}_{T^\delta} + \frac{f}{\sigma^2 t \mathbb{I}(\bar{\vartheta}_{T^\delta})} \int_{T^\delta}^t \dot{m}(\bar{\vartheta}_{T^\delta}, s) [dX_s - f m(\bar{\vartheta}_{T^\delta}, s) ds].$$

Let us put $t = \tau T$ and denote $\vartheta_{\tau T, T}^* = \vartheta_T^*(\tau)$, $T^{\delta-1} < \tau \leq 1$.

Theorem 6. [K. 2019] *One-step MLE-process $\vartheta_T^*(\tau)$, $T^{\delta-1} < \tau \leq 1$*

with $\delta \in (1/2, 1)$ is consistent: for any $\nu > 0$ and any $\tau \in (0, 1]$

$$\lim_{T \rightarrow \infty} -\mathbf{P}_{\vartheta_0} (|\vartheta_T^*(\tau) - \vartheta_0| > \nu) = 0,$$

and asymptotically normal

$$\sqrt{\tau T} (\vartheta_T^*(\tau) - \vartheta_0) \Rightarrow \mathcal{N} \left(0, \mathbf{I}(\vartheta_0)^{-1} \right).$$

Moreover, the random process $\zeta_T(\tau) = \sqrt{T\mathbf{I}(\vartheta_0)} (\vartheta_T^*(\tau) - \vartheta_0)$, $\kappa \leq \tau \leq 1$ for any $\kappa \in (0, 1)$ converges in $(\mathcal{C}[\kappa, 1], \mathcal{B})$ to the Wiener process.

Note that the process $\vartheta_t^*, T^\delta < t \leq T$ can be written in recurrent form

$$d\vartheta_t^* = -\frac{\vartheta_t^*}{t - T^\delta} dt + \frac{f\dot{m}(\bar{\vartheta}_{T^\delta}, t)}{\sigma^2 t \mathbf{I}(\bar{\vartheta}_{T^\delta})} [dX_t - f m(\bar{\vartheta}_{T^\delta}, t) dt]$$

and we can introduce the adaptive filtering equations as follows

$$dm_t = - \left[\vartheta_t^* + \frac{\gamma(\vartheta_t^*) f^2}{\sigma^2} \right] m_t dt + \frac{\gamma(\vartheta_t^*) f}{\sigma^2} dX_t, \quad T^\delta < t \leq T,$$

$$\gamma(\vartheta_t^*) = \frac{\vartheta_t^* \sigma^2}{f^2} \left(\sqrt{1 + \frac{b^2 f^2}{(\vartheta_t^*)^2 \sigma^2}} - 1 \right)$$

with the initial value $m_{T^\delta} = m(\bar{\vartheta}_{T^\delta}, T^\delta)$.

Hidden telegraph process. Consider the following model of observations

$$dX_t = Y_t dt + dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T,$$

where Y_t is a hidden telegraph process with two states: y_1, y_2 and rate matrix

$$\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

The unknown parameter is $\vartheta = (\lambda, \mu)$ and we have to estimate it by the observations $X^T = (X_t, 0 \leq t \leq T)$. The One-step MLE-process we construct in two steps. First we define the method of moments estimator $\bar{\vartheta}_{T^\delta}$ obtained by the observations $X^{T^\delta} = (X_t, 0 \leq t \leq T^\delta)$, where $\delta \in (1/2, 1)$, verify its consistency and then we introduce the One-step estimator-process by the relation

$$\vartheta_{t,T}^* = \bar{\vartheta}_{T^\delta} + t^{-1} \mathbb{I}_t(\bar{\vartheta}_{T^\delta})^{-1} \int_{T^\delta}^t \dot{m}(\bar{\vartheta}_{T^\delta}, s) [dX_s - m(\bar{\vartheta}_{T^\delta}, s) ds].$$

Here the vector

$$\dot{m}(\vartheta, s) = (y_1 - y_2) \frac{\partial \pi_\lambda(s, \vartheta)}{\partial \vartheta} = (y_1 - y_2) \left(\frac{\partial \pi(s, \vartheta)}{\partial \lambda}, \frac{\partial \pi(st, \vartheta)}{\partial \mu} \right)^*$$

and the empirical Fisher information matrix $\mathbb{I}_t(\vartheta)$ is

$$\mathbb{I}_t(\vartheta) = \frac{1}{t} \int_{T^\delta}^t \dot{m}(\vartheta, s) \dot{m}(\vartheta, s)^* ds \longrightarrow \mathbb{I}(\vartheta)$$

as $t \rightarrow \infty$ by the law of large numbers. We denoted

$m(\vartheta, t) = \mathbf{E}_\vartheta(Y_t | X_s, 0 \leq s \leq t)$, $\pi(t, \vartheta) = \mathbf{P}_\vartheta(Y_t = y_1 | X_s, 0 \leq s \leq t)$.

We show that [6] the one-step MLE-process is consistent and it is asymptotically normal

$$\sqrt{T}(\vartheta_T^* - \vartheta_0) \implies \mathcal{N}(0, \mathbb{I}(\vartheta_0)^{-1}).$$

Remark. Note that extended review including the mentioned in this work results and several other results for the same models of observations is given in the work [13].

References

- [1] Chigansky P. Maximum likelihood estimation for hidden Markov models in continuous time // Statist. Inference Stoch. Processes. 2009. Vol. 12, No. 2. P. 139–163.
- [2] Elliott R.J., Aggoun L., Moor J.B. Hidden Markov Models. - N.Y.: Springer, 1995.
- [3] Iacus S., Yoshida N. Estimation for discretely observed telegraph process // Theory Probab. Math. Statist. 2009. Vol. 78. P. 37–47.
- [4] Ibragimov I.A., Khasminskii R. Statistical Estimation - Asymptotic Theory. - New York: Springer-Verlag, 1981.
- [5] Kamatani K., Uchida M. Hybrid multi-step estimators for stochastic differential equations based on sampled data // Statist. Inf. Stoch. Proc. 2015. Vol. 18, No. 2. P. 177–204.
- [6] Khasminskii R.Z., Kutoyants Yu. A. On parameter estimation of hidden telegraph process // Bernoulli. 2018. Vol. 24, No. 3. P. 2064–2090.
- [7] Kutoyants Yu.A. Parameter Estimation for Stochastic Processes. Yerevan (in Russian) Translated in 1984, Heldermann, Berlin, 1980.
- [8] Kutoyants Yu.A. Identification of Dynamical Systems with Small Noise. - Dordrecht: Kluwer, 1994.
- [9] Kutoyants Yu. A. On the multi-step MLE-process for ergodic diffusion // Stoch. Proces. & Appl. 2017. Vol. 127. P. 2243–2261.
- [10] Kutoyants Yu. A. On parameter estimation of hidden Ornstein-Uhlenbeck process // Journal of Multivariate Analysis. 2019. Vol. 169, No 1. P. 248–263.

- [11] Kutoyants Yu. A. On parameter estimation of hidden ergodic Ornstein-Uhlenbeck process // 2019 (arXiv:1902.08500) submitted.
- [12] Kutoyants Yu. A. On localization of source by hidden Gaussian processes // 2019, submitted.
- [13] Kutoyants Yu. A. Parameter estimation for continuous time hidden Markov processes: a survey // 2019, To appear in Automation and Remote Control, 2019.
- [14] Kutoyants Yu. A., Zhou L. On parameter estimation of the hidden Gaussian process in perturbed SDE // 2019 (arXiv:1904.09750) submitted.
- [15] Liptser R.S., Shiryaev A.N. Statistics of Random Processes, 2-nd ed., vol. 1. - N.Y.: Springer, 2001.
- [16] Wonham W. M. Some applications of stochastic differential equations to optimal non-linear filtering // SIAM J. Contr. 1965. Ser. A, No 2. P. 347–369.

Кутоянц Ю.А. (Университет Ле-Мана, Ле-Ман, Франция, Томский государственный университет, Томск, Россия, 2019) **Об идентификации скрытых марковских процессов с непрерывным временем: обзор.**

Аннотация. Мы рассматриваем несколько моделей частично наблюдаемых стохастических линейных систем и обсуждаем задачи оценки параметров этих моделей. Мы описываем асимптотические свойства оценок максимального правдоподобия, байесовских оценок Одношаговых ОМП в случае двух асимптотик: малого шума и больших выборок. Во всех задачах мы устанавливаем состоятельность этих оценок, их предельные распределения и определяем асимптотически эффективные оценки.

Ключевые слова: частично наблюдаемые системы, оценка параметров, асимптотические свойства, проблемы разладки.

Improved model selection for estimation in semimartingale regression from discrete data ¹

Pchelintsev E. A.

Tomsk State University, Tomsk
e-mail: evgen-pch@yandex.ru

Abstract

We consider the adaptive nonparametric estimation problem for a function of a continuous time regression model with semimartingale noises by incomplete observations. We consider the regression model defined by non-Gaussian Ornstein–Uhlenbeck processes. A model selection procedure, based on the shrinkage weighted least squares estimates, is proposed. The sharp oracle inequalities for the robust risks are obtained. the robust efficiency property has been established in adaptive setting.

Keywords: improved estimation, least squares estimates, robust quadratic risk, Ornstein–Uhlenbeck process, semimartingale regression, model selection, sharp oracle inequality, asymptotic efficiency.

Introduction. Statement of a problem. In this paper we consider the following continuous time regression model

$$dy_t = S(t)dt + d\xi_t, \quad 0 \leq t \leq n, \quad (1)$$

where S is an unknown 1-periodic $\mathbb{R} \rightarrow \mathbb{R}$ function from $\mathbf{L}_2[0, 1]$, $(\xi_t)_{t \geq 0}$ is an unobservable noise which is defined by a non-Gaussian Ornstein–Uhlenbeck process with the Lévy subordinator

$$d\xi_t = a\xi_t dt + du_t, \quad \xi_0 = 0, \quad (2)$$

where

$$u_t = \varrho_1 w_t + \varrho_2 z_t \quad \text{and} \quad z_t = x * (\mu - \tilde{\mu})_t. \quad (3)$$

Here $(w_t)_{t \geq 0}$ is a standard Brownian motion, "*" denotes the stochastic integral with respect to the compensated jump measure $\mu(ds, dx)$ with deterministic compensator $\tilde{\mu}(ds dx) = ds\Pi(dx)$, i.e.

$$z_t = \int_0^t \int_{\mathbb{R}_*} v(\mu - \tilde{\mu})(ds dv) \quad \text{and} \quad \mathbb{R}_* = \mathbb{R} \setminus \{0\},$$

$\Pi(\cdot)$ is the Lévy measure on $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$, (see, for example in [1]), such that

$$\Pi(x^2) = 1 \quad \text{and} \quad \Pi(x^8) < \infty. \quad (4)$$

¹The work was supported by the RSF, the project No 17-11-01049.

We use the notation $\Pi(|x|^m) = \int_{\mathbb{R}^*} |z|^m \Pi(dz)$. Moreover, we assume that the nuisance parameters $a \leq 0$, ϱ_1 and ϱ_2 satisfy the conditions

$$-a_{max} \leq a \leq 0, \quad 0 < \underline{\varrho} \leq \varrho_1^2 \quad \text{and} \quad \sigma_Q = \varrho_1^2 + \varrho_2^2 \leq \varsigma^*, \quad (5)$$

where the bounds a_{max} , $\underline{\varrho}$ and ς^* are functions of n , i.e. $a_{max} = a_{max}(n)$, $\underline{\varrho} = \varrho_n$ and $\varsigma^* = \varsigma_n^*$, such that for any $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \frac{a_{max}(n) + \varsigma_n^*}{n^\epsilon} = 0 \quad \text{and} \quad \liminf_{n \rightarrow \infty} n^\epsilon \underline{\varrho}_n > 0. \quad (6)$$

We denote by $\mathcal{Q}_n$ the family of all distributions of process (1) – (2) on the Skorokhod space $\mathbf{D}[0, n]$ satisfying the conditions (5) – (6). It should be noted that the process (2) is conditionally-Gaussian square integrated semimartingale with respect to σ -algebra $\mathcal{G} = \sigma\{z_t, t \geq 0\}$ which is generated by jump process $(z_t)_{t \geq 0}$.

The problem is to estimate the unknown function S in the model (1) on the basis of observations

$$(y_{t_j})_{0 \leq j \leq np}, \quad t_j = \frac{j}{p}, \quad (7)$$

where the observations frequency p is some fixed integer number. For this problem we use the quadratic risk, which for any estimate $\hat{S}$, is defined as

$$\mathcal{R}_Q(\hat{S}, S) := \mathbf{E}_{Q,S} \|\hat{S} - S\|^2 \quad \text{and} \quad \|f\|^2 := \int_0^1 f^2(t) dt, \quad (8)$$

where $\mathbf{E}_{Q,S}$ stands for the expectation with respect to the distribution $\mathbf{P}_{Q,S}$ of the process (1) with a fixed distribution Q of the noise $(\xi_t)_{0 \leq t \leq n}$ and a given function S . Moreover, in the case when the distribution Q is unknown we use also the robust risk

$$\mathcal{R}^*(\hat{S}, S) = \sup_{Q \in \mathcal{Q}_n} \mathcal{R}_Q(\hat{S}, S). \quad (9)$$

To study the estimation problem for non-Gaussian observations (1) in the papers [2, 8, 7, 5, 6] it was introduced impulse noises defined through the semi-Markov or compound Poisson processes with unknown impulse distributions. However, the semi-Markov or compound Poisson processes can describe the impulse influence of only one fixed frequency. It should be noted that in the telecommunication systems, the noise impulses are without limitations on frequencies and, therefore, such models are too restricted for practical applications. To include all possible impulse noises, in [3, 4] it was proposed to use general non-Gaussian semimartingale processes. Later, for semimartingale models in the papers [9, 10, 11, 12] the authors developed the improved (shrinkage) nonparametric estimation methods. It should be emphasized, that in all these papers the improved estimation problems are studied only for the complete observations cases, i.e. when the all trajectory $(y_t)_{0 \leq t \leq n}$ is accessed to be observed.

Our main goal in this paper is to develop improved estimation meth-

ods for the incomplete observations, i.e. when the process (1) can be observed only in the fixed time moments (7). As an example, we consider the regression model (1) with the noise defined by non-Gaussian Ornstein–Uhlenbeck process with unknown distribution.

Improved estimation. For estimating the unknown function S in (1) we will use it's Fourier expansion with respect to an orthonormal basis $(\phi_j)_{j \geq 1}$ in $\mathbf{L}_2[0, 1]$. We extend these functions by the periodic way on $\mathbb{R}$, i.e. $\phi_j(t) = \phi_j(t + 1)$ for any $t \in \mathbb{R}$. Assume that the basis functions are uniformly bounded, i.e. for some constant $\phi_* \geq 1$, which may be depend on n ,

$$\sup_{0 \leq j \leq n} \sup_{0 \leq t \leq 1} |\phi_j(t)| \leq \phi_* < \infty. \quad (10)$$

Moreover we will use such basis that the restrictions of the functions $(\phi_j)_{1 \leq j \leq p}$, on the sampling lattice

$$\mathcal{T}_p = \{t_1, \dots, t_p\}, \quad t_j = j/p,$$

form an orthonormal basis in the Hilbert space $\mathbb{R}^{\mathcal{T}_p}$ with the inner product

$$(x, y)_p = \frac{1}{p} \sum_{j=1}^p x(t_j) y(t_j) \quad \text{for } x, y \in \mathbb{R}^{\mathcal{T}_p}, \quad (11)$$

i.e. $(\phi_i, \phi_j)_p = \mathbf{1}_{\{i=j\}}$. We put the norm $\|x\|_p = \sqrt{(x, x)_p}$.

We write the discrete Fourier expansion of the unknown function S on the lattice $\mathcal{T}_p$ in the form

$$S(t) = \sum_{j=1}^p \theta_{j,p} \phi_j(t),$$

where the corresponding Fourier coefficients

$$\theta_{j,p} = (S, \phi_j)_p \quad (12)$$

can be estimated from the discrete data by the formulae

$$\hat{\theta}_{j,p} = \frac{1}{n} \int_0^n \psi_{j,p}(t) dy_t, \quad \psi_{j,p}(t) = \sum_{k=1}^{np} \phi_j(t_k) \mathbf{1}_{(t_{k-1}, t_k]}(t). \quad (13)$$

As in [6] we define a class of weighted least squares estimates for $S(t)$ as

$$\hat{S}_\gamma(t) = \sum_{j=1}^p \gamma(j) \hat{\theta}_{j,n} \psi_{j,p}(t), \quad (14)$$

where the weights $\gamma = (\gamma(j))_{1 \leq j \leq p} \in \mathbb{R}^p$ belong to some finite set Γ from $[0, 1]^p$. We will use here the set Γ from [10].

For the first $d > 6$ Fourier coefficients in (14) we will use the improved estimation method proposed for parametric models in [8]. To this end we set $\tilde{\theta}_p = (\hat{\theta}_{j,p})_{1 \leq j \leq d}$. 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 $\mathbb{R}^d$. Now we define the

shrinkage estimators as

$$\theta_{j,p}^* = (1 - g(j)) \hat{\theta}_{j,p}, \quad (15)$$

where $g(j) = (\mathbf{c}_n / |\tilde{\theta}_p|_d) \mathbf{1}_{\{1 \leq j \leq d\}}$,

$$\mathbf{c}_n = \frac{\underline{g}_n (d - 6)}{2 \left(r + \sqrt{d\kappa/n} \right) n} \quad \text{and} \quad \kappa = \sup_{Q \in \mathcal{Q}_n} \kappa_Q.$$

The positive parameter r may be dependent of n , i.e. $r = r_n$, and such that

$$\lim_{n \rightarrow \infty} n^{-\epsilon} r_n = 0 \quad \text{for any} \quad \epsilon > 0. \quad (16)$$

Now we set shrinkage estimates for S

$$S_\gamma^*(t) = \sum_{j=1}^p \gamma(j) \theta_{j,p}^* \psi_{j,p}(t). \quad (17)$$

We compare the estimators (14) and (17) through the difference

$$\Delta_Q(S) := \mathcal{R}_Q(S_\gamma^*, S) - \mathcal{R}_Q(\hat{S}_\gamma, S).$$

Now we obtain the non asymptotic bound for this comparative risk.

Let now we set

$$p_0 = \frac{\sqrt{d}\phi_* L}{\mathbf{c}_n} + 1, \quad L = \sup_{0 \leq s, t \leq 1} \frac{|S(t) - S(s)|}{|t - s|}. \quad (18)$$

Theorem 1. *Assume that the conditions $\mathbf{D}_1) - \mathbf{D}_2)$ hold. Moreover, assume that the function S is Lipschitzian. Then for any $p \geq p_0$*

$$\sup_{Q \in \mathcal{Q}_n} \sup_{\|S\| \leq r} \Delta_Q(S) < 0. \quad (19)$$

Model selection. Main results. In order to obtain a good estimate, we will use a rule to choose a weight vector $\gamma \in \Gamma$ in (17) as in [10]. We denote the improved model selection procedure

$$S^* = S_{\gamma^*}^*, \quad (20)$$

where γ^* is defined in [10].

Theorem 2. *Under some conditions for any $n \geq 2$ and $0 < \rho < 1/2$*

$$\mathcal{R}^*(S^*, S) \leq \frac{1 + 5\rho}{1 - \rho} \min_{\gamma \in \Gamma} \mathcal{R}_n^*(S_\gamma^*, S) + \frac{1}{\rho n} \mathbf{U}_n,$$

where the coefficient $\mathbf{U}_n$ is such that for any $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \frac{\mathbf{U}_n}{n^\epsilon} = 0. \quad (21)$$

This sharp oracle inequality allows us to prove the asymptotic efficiency property for proposed model selection procedure.

In order to study the asymptotic efficiency we need the functional Sobolev ball $W_{k,\mathbf{r}}$ that is defined in [10]. We denote by Σ_n all estimators $\hat{S}_n$ i.e. any $\sigma\{y_t, 0 \leq t \leq n\}$ measurable functions. In the sequel we

denote by Q^* the distribution of the process $(y_t)_{0 \leq t \leq n}$ with $\xi_t = \varsigma^* w_t$, i.e. white noise model with the intensity ς^* .

Theorem 3. Assume that $Q^* \in \mathcal{Q}_n$. The robust risk (9) admits the following lower bound

$$\liminf_{n \rightarrow \infty} \inf_{\hat{S}_n \in \Sigma_n} v_n^{2k/(2k+1)} \sup_{S \in W_{k,\mathbf{r}}} \mathcal{R}^*(\hat{S}_n, S) \geq l_k(\mathbf{r}),$$

where $l_k(\mathbf{r}) = ((2k+1)\mathbf{r})^{1/(2k+1)} (k/(\pi(k+1)))^{2k/(2k+1)}$ and $v_n = n/\varsigma^*$.

We show that this lower bound is sharp in the following sense.

Theorem 4. Assume that $Q^* \in \mathcal{Q}_n$ and there exists $\epsilon > 0$ such that $\lim_{n \rightarrow \infty} n^{5/6+\epsilon}/p = 0$. Then the robust risk of the model selection procedure (20) satisfies the following upper bound

$$\limsup_{n \rightarrow \infty} v_n^{2k/(2k+1)} \sup_{S \in W_{k,\mathbf{r}}} \mathcal{R}^*(S^*, S) \leq l_k(\mathbf{r}).$$

It is clear that these theorems imply the following efficient property.

Corollary 1. Assume that $Q^* \in \mathcal{Q}_n$. Then the model selection procedure (20) is asymptotically efficient, i.e.

$$\lim_{n \rightarrow \infty} v_n^{2k/(2k+1)} \sup_{S \in W_{k,\mathbf{r}}} \mathcal{R}^*(S^*, S) = l_k(\mathbf{r}).$$

References

- [1] Cont R., Tankov P. Financial Modelling with Jump Processes. London: Chapman & Hall, 2004.
- [2] Barbu V., Beltaief S., Pergamenshchikov S. M. Robust adaptive efficient estimation for semi - Markov nonparametric regression models // Statistical inference for stochastic processes. 2019. Vol. 22, No 2. P. 187–231.
- [3] Konev V.V., Pergamenshchikov S.M. Nonparametric estimation in a semimartingale regression model. Part 1. Oracle Inequalities // Tomsk State University Journal of Mathematics and Mechanics. 2009. No 7. P. 23–41.
- [4] Konev V.V., Pergamenshchikov S.M. Nonparametric estimation in a semimartingale regression model. Part 2. Robust asymptotic efficiency // Tomsk State University Journal of Mathematics and Mechanics. 2009. No 8. P. 31–45.

- [5] Konev V.V., Pergamenshchikov S.M. Efficient robust nonparametric in a semimartingale regression model // Annals of the Institute of Henri Poincaré. Probab. and Stat. 2012. Vol. 48, No 4. P. 1217–1244.
- [6] Konev V.V., Pergamenshchikov S.M. Robust model selection for a semimartingale continuous time regression from discrete data // Stochastic processes and their applications. 2015. Vol. 125. P. 294–326.
- [7] Konev V., Pergamenshchikov S. and Pchelintsev E. Estimation of a regression with the pulse type noise from discrete data // Theory Probab. Appl. 2014. Vol. 58, No 3. P. 442–457.
- [8] Pchelintsev E. Improved estimation in a non-Gaussian parametric regression // Stat. Inference Stoch. Process. 2013. Vol. 16, No 1. P. 15–28.
- [9] Pchelintsev E., Pergamenshchikov S. Oracle inequalities for the stochastic differential equations // Stat. Inference Stoch. Process. 2018. Vol. 21 No 2. P. 469–483.
- [10] Pchelintsev E., Pergamenshchikov S. Adaptive model selection method for a conditionally Gaussian semimartingale regression in continuous time // Tomsk State University Journal of Mathematics and Mechanics. 2019. No 58. P. 14 – 31.
- [11] Pchelintsev E.A., Pchelintsev V.A., Pergamenshchikov S. M. Non asymptotic sharp oracle inequality for the improved model selection procedures for the adaptive nonparametric signal estimation problem // Komunikacie. 2018. Vol. 20, No 1. P. 72–76.
- [12] Pchelintsev E.A., Pchelintsev V.A., Pergamenshchikov S. M. Improved robust model selection methods for a Lévy nonparametric regression in continuous time // Journal of Nonparametric Statistics. 2019. Vol. 31, No 3. P. 612–628.
- [13] Pinsker, M.S. Optimal filtration of square integrable signals in gaussian white noise // Problems of Transimission information. 1981. Vol. 17. P. 120–133.

Пчелинцев Е. А. (Томский государственный университет, Томск, 2019) **Улучшенный метод выбора модели для оценивания в семимартингальной регрессии по дискретным данным.**

Аннотация. Рассматривается задача адаптивного непараметрического оценивания функции в модели непрерывной семимартингальной регрессии по неполным наблюдениям. Рассматривается модель регрессии с шумами, определяемыми негауссовскими процессами Орнштейна - Уленбека. Предложена процедура выбора модели, основанная на улучшенных взвешенных оценках наименьших квадратов. Получены точные оракульные неравенства для робастных рисков. Установлено свойство робастной эффективности в адаптивной обстановке.

Ключевые слова: улучшенное оценивание, оценка наименьших квадратов, робастный квадратичный риск, процесс Орнштейна - Уленбека, семимартингальная регрессия, выбор модели, точное оракульное неравенство, асимптотическая эффективность.

Efficient Nonparametric Estimation of Square-integrable Functions in Continuous Time Regression Models ¹

Povzun M. A., Pchelintsev E. A., Pergamenschikov S. M.

Tomsk State University, Tomsk,
Rouen University, Rouen, France
e-mail: povzunyasha@gmail.com

Abstract

In this paper we study an asymptotic efficiency property of the weighted least squares estimates for unknown square integrable functions in Gaussian regression models. We use the Pinsker approach. It is established that it is impossible to directly use the Pinsker method. Some conditions and an a priori distribution for the Fourier coefficients are obtained under which the asymptotic lower bound for the mean square risk was proved.

Keywords: regression model, weighted least squares estimates, asymptotic efficiency, Pinsker method.

We will observe the following process

$$dy_t = \theta(t)dt + \varepsilon dW_t, \quad 0 \leq t \leq 1. \quad (1)$$

This model is a non-parametric regression in continuous time with noises of small intensity, where $\theta(t) \in \mathcal{L}_2[0, 1]$ – unknown function, $(W_t)_{0 \leq t \leq 1}$ is Brownian motions and $\varepsilon > 0$ is the noise intensity. Another words we have got a "signal+white noise" model. Such models are widely used in statistical radio-physics and financial mathematics.

Our goal is to study an asymptotic efficiency property of estimate the unknown function $\theta(t)$ as $\varepsilon \rightarrow 0$, i.e. in the case when the ratio "signal/noise" tends to infinity.

The asymptotically efficient estimates for Gaussian models have been constructed in [3, 6, 9]. In [1, 2] was proposed a new approach to study an asymptotic efficiency for estimating problem in the non-parametric models via sharp oracle inequalities method. This method has been developed for semimartingale regression models in continuous time by [4, 5]. For improved estimates of unknown regression functions was expanded such approach in [7, 8].

Estimation of the function will be considered in the sense of mean-square accuracy:

$$\mathcal{R}(\hat{\theta}, \theta) := \mathbf{E}\|\hat{\theta} - \theta\|^2 \quad \text{and} \quad \|\theta\|^2 = \int_0^1 \theta^2(t)dt. \quad (2)$$

¹The work was supported by the RSF, the project No 17-11-01049.

Let (φ_j) be an orthonormal basis, $(\varphi_j) \in \mathcal{L}_2[0, 1]$ and we represent $\theta(t)$ as

$$\theta(t) = \sum_{j=1}^{\infty} \theta_j \varphi_j(t), \quad (3)$$

where the Fourier coefficients can be represented as

$$\theta_j = \int_0^1 \theta(t) \varphi_j(t) dt.$$

Further we need the following set

$$\Theta = \left\{ \theta(t) \in \mathcal{L}_2[0, 1] : \sum_{j=1}^{\infty} a_j \theta_j^2 \leq \mathbf{r} \right\}, \quad (4)$$

where $\mathbf{r} > 0$.

In order to construct the estimate, one needs to know how many parameters should be choosing. In the sequel we will use the following threshold

$$n_* = \max_{l \geq 1} \left\{ \sqrt{a_l} \sum_{j=1}^l \sqrt{a_j} - \sum_{j=1}^l a_j \leq \frac{\mathbf{r}}{\varepsilon^2} \right\}. \quad (5)$$

Note that n_* is finite, since the sums, in view of (4), are bounded by the constant $\mathbf{r}$.

So, we turn on from the non-parametric estimation problem to the parametric one. The Fourier coefficients in (3) can be estimated as

$$\hat{\theta}_j = \int_0^1 \varphi_j(t) dy_t = \theta_j + \varepsilon \xi_j, \quad (6)$$

where

$$\xi_j = \int_0^1 \varphi_j(t) dW_t.$$

Now we consider projection estimates in the following form:

$$\tilde{\theta}_j = \alpha_j \hat{\theta}_j,$$

where $(\alpha_j)_{j \geq 1}$ is a non-random sequence.

We find such a case that the difference between the estimate and the true value of the function will be minimal.

We set

$$\nu^2 = \inf_v \sup_{\theta \in \Theta} \sum_{j=1}^{\infty} \mathbf{E}_{\theta} \left(\alpha_j \hat{\theta}_j - \theta_j \right)^2, \quad (7)$$

where $\inf$ is taken over all coefficients $v = (v_i)_{i \geq 1}$. Using Lemma of Pinsker [?] we can find that

$$\nu^2 = \varepsilon^2 \sum_{j=1}^{n_*} \frac{\theta_j^{*2}}{\theta_j^{*2} + \varepsilon^2} = n_* \varepsilon^2 - \frac{\varepsilon^2 (\sum_{j=1}^{n_*} \sqrt{a_j})^2}{\mathbf{r}/\varepsilon^2 + \sum_{j=1}^{n_*} a_j}. \quad (8)$$

where optimal points $(\theta_j^*)_{1 \leq j \leq n_*}$ are defined as:

$$\theta_j^{*2} = \varepsilon^2 \left(\frac{\mu}{\sqrt{a_j}} - 1 \right) \quad \text{and} \quad \mu = \frac{\mathbf{r}/\varepsilon^2 + \sum_{j=1}^{n_*} a_j}{\sum_{j=1}^{n_*} \sqrt{a_j}}. \quad (9)$$

The parameter μ is taken from the next condition

$$\sum_{j=1}^{n_*} a_j \theta_j^* = \mathbf{r}. \quad (10)$$

In the future, we will need important conditions for the proof of the theorems:

A₁) A sequence $(a_j)_{j \geq 1}$ increasing numbers such that $a_j > a_{j-1}$ and

$$\sum_{j=1}^{\infty} \sqrt{a_j} (\sqrt{a_{j+1}} - \sqrt{a_j}) = +\infty;$$

For a fixed $h > 0$

$$N_h = \max_{j \geq 1} \{a_j \leq h\}. \quad (11)$$

A₂) For any $0 < \gamma < 1$,

$$N_h - N_{\gamma h} - \frac{1}{\sqrt{a_{N_h}}} \sum_{j=N_{\gamma h}+1}^{N_h} \sqrt{a_j} \rightarrow \infty.$$

Without these conditions, Pinsker theorem cannot be proved.

First we obtain the lower bound.

Theorem 1. *For the model (1) the following lower bound holds*

$$\lim_{\varepsilon \rightarrow 0} \sup_{\theta \in \Theta} \inf_{\hat{\theta}} \frac{\mathcal{R}(\hat{\theta}, \theta)}{\nu^2} \geq 1. \quad (12)$$

where $\inf$ is taken over all possible estimates $\hat{\theta}$.

Moreover, consider ν^2 in the following form

$$\nu^2 = \sum_{j=n_0+1}^{n_1} \frac{\theta_j^{*2} \varepsilon^2}{\theta_j^{*2} + \varepsilon^2}. \quad (13)$$

This is a key property, which is necessary for the proof of the lower bound.

Proposition 1. *The condition A_2 imply*

$$\lim_{\varepsilon \rightarrow 0} \frac{\nu^2}{\varepsilon^2} = +\infty, \quad (14)$$

where

$$n_0 = \max_{j \geq 1} \left\{ a_j \leq \frac{\mu^2}{(1+q)^2} \right\} \quad \text{and} \quad n_1 = \max_{j \geq 1} \left\{ a_j \leq \frac{\mu^2 q^2}{(1+q)^2} \right\}. \quad (15)$$

If $a_j = e^{2j}$, then conditions $A_1) - A_2)$ are not satisfied, so

$$\lim_{\varepsilon \rightarrow 0} \frac{\nu^2}{\varepsilon^2} < +\infty. \quad (16)$$

It means that the theorem of Pinsker without conditions is not true. To confirm the obtained results, we consider the following example. Let $a_j = j^{2k}$ and k is fixed. Also fix some $q > 0$. Check the condition A_1):

$$\sum_{j=1}^{\infty} \sqrt{j^{2k}} (\sqrt{(j+1)^{2k}} - \sqrt{j^{2k}}) = +\infty.$$

As you can see, this condition is satisfied. Next, we check the condition A_2):

$$N_h - N_{\gamma h} - \sum_{j=1}^{N_h} \left(\frac{j}{N_h} \right)^k + \sum_{j=1}^{N_{\gamma h}} \left(\frac{j}{N_h} \right)^k = +\infty,$$

we obtain the condition A_2 satisfy. Thus, this coefficients can be used for estimation procedure. Next, we calculate some more characteristics. Note that, in equation (5) for n_*

$$n_* \approx \left(\frac{\mathbf{r}}{\varepsilon^2} \right)^{\frac{1}{2k+1}} \quad \text{and} \quad \mu \approx \left(\frac{1}{\varepsilon^2} \right)^{\frac{k}{2k+1}}.$$

Note that if we take $a_j = e^{2j}$, we get a counter-example. So, conditions A_1 and A_2 are not satisfied for this.

We estimate the function $\theta(t)$ as

$$\tilde{\theta}(t) = \sum_{j=1}^{n_*} \lambda_j \hat{\theta}_j \varphi_j(t), \quad \lambda_j = \left(1 - \frac{\sqrt{a_j}}{\mu} \right). \quad (17)$$

Let us proceed to the formulation of the theorem on the upper boundary.

Theorem 2. *The estimator (17) admits the following upper bound*

$$\limsup_{\varepsilon \rightarrow 0} \sup_{\theta \in \Theta} \frac{\mathcal{R}(\tilde{\theta}, \theta)}{\nu^2} \leq 1. \quad (18)$$

Using these theorems, we can formulate the efficiency property of estimates.

Corollary 1. *The theorems 2 and 1 imply*

$$\limsup_{\varepsilon \rightarrow 0} \sup_{\theta \in \Theta} \frac{\mathcal{R}(\tilde{\theta}, \theta)}{\nu^2} = 1. \quad (19)$$

In this paper we constructed an estimator for which we showed efficiency property. To this end we use the approach proposed by Pinsker (1980) for this problem. Unfortunately, we cannot use directly the Pinsker method for since his main theorem about lower bound is not true without conditions on the coefficient of the Sobolev ball. We found this conditions for which we provide the efficiency property. The main difficulty is to find the prior distribution for Fourier coefficient. Moreover, we gave the constrictive sufficient conditions for which this prop-

erty holds. And as example we checked the obtained conditions for Sobolev coefficients which goes to infinity power function.

References

- [1] Galtchouk L.I., Pergamenshchikov S.M. Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes // Stat. Inference Stoch. Processes. 2006. Vol.9, No1. P. 1–16.
- [2] Galtchouk L., Pergamenshchikov S. Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression // Journal of Korean Statistical Society. 2009. Vol. 38, No 4. P. 305–322.
- [3] Ibragimov I. A., Khasminskii R. Z. Statistical Estimation: Asymptotic Theory. – New York: Springer, 1981.
- [4] Konev V. V., Pergamenshchikov S. M. Nonparametric estimation in a semimartingale regression model. Part 2. Robust asymptotic efficiency // Journal of Mathematics and Mechanics of Tomsk State University. 2009. No 4. P. 31–45.
- [5] Konev V.V., Pergamenshchikov S.M. Efficient robust nonparametric estimation in a semimartingale regression model // Ann. Inst. Henri Poincaré Probab. Stat. 2012. Vol.48, No 4. P. 1217–1244.
- [6] Kutoyants Yu. A. Identification of dynamical systems with small noise. – Dordrecht: Kluwer Academic Publishers Group, 1994.
- [7] Pchelintsev E., Pergamenshchikov S. Oracle inequalities for the stochastic differential equations // Stat. Inference Stoch. Process. 2018. Vol. 21, No 2. P. 469–483.
- [8] Pchelintsev E.A., Pchelintsev V.A., Pergamenshchikov S.M. Improved robust model selection methods for a Lévy nonparametric regression in continuous time // Journal of Nonparametric Statistics. 2019. Vol. 31, No 3. P. 612–628.
- [9] Pinsker M.S. Optimal filtration of square integrable signals in Gaussian white noise // Problems of Transimission information. 1981. Vol. 17. P. 120–133.

Повзун М. А., Пчелинцев Е. А., Пергаменщиков С. М. (Томский государственный университет, Томск, Руанский Университет, Руан, 2019) **Эффективное непараметрическое оценивание квадратично интегрируемых функций в моделях непрерывной регрессии.**

Аннотация. В данной работе изучено свойство асимптотической эффективности взвешенных оценок наименьших квадратов для неизвестных квадратично интегрируемых функций в регрессионных моделях с гауссовскими шумами. Применяется подход Пинскера. Установлено, что напрямую использовать данный метод невозможно. Получены некоторые условия и априорное распределение для коэффициентов Фурье, при которых доказана асимптотическая нижняя оценка для среднеквадратичного риска.

Ключевые слова: регрессионная модель, взвешенные оценки наименьших квадратов, асимптотическая эффективность, метод Пинскера.

The Hedging Problem for Asian Options in Financial Markets with Transaction Costs ¹

Shishkova A. A.

Tomsk State University, Tomsk
e-mail: alshishkovatomsk@gmail.com

Abstract

In this paper we study asymptotic property for the portfolio value with transaction cost in the Black-Scholes model with risky asset without drift and risk-free asset with interest rate $r = 0$. We use the modification of Leland's strategy. Main result of study is obtained sufficient conditions, which provide asymptotic hedging.

Keywords: Asian option, hedging strategy, Brownian motion, stochastic differential equations, Black and Scholes model.

Introduction. Models with proportional transaction costs were considered as early as the 1970s. Magill and Constantinides [1] suggested in 1976 the consumption-investment model which is generalization of the Merton model of 1973 [2]. However, the article written by H. Leland [3] in 1985 became more important for practical application. Leland's strategy provides an easy way to effectively eliminate the risks associated with transaction costs. This method is based on the idea that transaction costs can be offset by increasing the volatility parameter in the Black-Scholes strategy, that is the delta strategy obtained from a changed Black-Scholes equation with an appropriate modified volatility ensures an approximately complete replication as expected. The major goal in Leland's algorithm is to explore the asymptotic behavior of the hedging error (difference between the terminal value of portfolio and the payoff function) as the number of transaction goes to infinity.

For a trader or an investor the main task is not only the saving but also the multiplication of its capital. Many risks can be avoided with the help of one popular and very effective technique hedging. The option is hedged to protect its value from the risk of price movement of the underlying asset in an unfavorable direction. Our goal is to extend hedging methods for the hedging problem of Asian option for the financial markets with transaction costs. To this end we use the approximative hedging approach proposed Leland, Kabanov, Safarian,

¹The work was supported by the Ministry of Science and Higher Education of the Russian Federation, Goszadanie No 2.3208.2017/4.6.

Pergamenshchikov, Lepinette [3, 4, 5, 6]. Note that in all this paper the hedging strategy is based on the delta-strategy. But for Asian option one needs to change basic strategy, i.e. to pass from delta-strategy to Asian hedging strategy constructed in [7].

Market model and settings. We consider the continuous time classical Black-Scholes model on financial market with risk-free asset (bond) and risky asset (stock). For simplicity we suppose that the risk-free rate $r = 0$, i.e. the bond price is constant over time $B_t = 1$ throughout this article. Let $(\Omega, \mathcal{F}_1, (\mathcal{F}_t)_{0 \leq t \leq 1}, \mathbf{P})$ be the standard filtered probability space with $\mathcal{F}_t = \sigma(W_s, 0 \leq s \leq t)$ and W is a Wiener process. The asset price process S_t given by

$$dS_t = \sigma S_t dW_t, \quad 0 \leq t \leq 1 \quad (1)$$

and admit the following explicit form

$$S_t = S_0 e^{\sigma W_t - \sigma^2 t/2}.$$

Remark that S_t is a martingale under measure $\mathbf{P}$. The model is considered on the interval $[0, 1]$ where 1 is a maturity of the Asian option with payoff function

$$f_1 = \left(\int_0^1 S_u du - K \right)_+.$$

The hedging problem for the Asian call option with the terminal payoff f_1 is to choose the admissible self-financing strategy (β_t, γ_t) such that

$$V_1 = V_0 + \int_0^1 \gamma_u dS_u \geq f_1, \quad \text{a.s.}$$

Here V_0 is an initial capital, β_t and γ_t are quantity of the risk-free asset and risk asset respectively.

To construct a hedging strategy in the case of model (1) without transaction costs apply the representation theorem for quadratic integrated martingale to the following martingale

$$M_t = \mathbf{E}(f_1 | \mathcal{F}_t). \quad (2)$$

We will find the square integrable process $(\alpha_t)_{0 \leq t \leq 1}$ adapted w.r.t. $\mathcal{F}_t$ such that for all $t \in [0, 1]$

$$M_t = M_0 + \int_0^t \alpha_s dW_s. \quad (3)$$

In [7] we obtained the formulas for strategy $\Pi = (\beta_t, \gamma_t)_{0 \leq t \leq 1}$

$$\gamma_t = \alpha_t / \sigma S_t,$$

$$\beta_t = \mathbf{E}f_1 + \int_0^t \alpha_s dW_s - \gamma_t S_t.$$

and we found the formula for calculating martingale coefficients in (3)

$$\alpha_t = \sigma G'_y(t, \xi_t, S_t) S_t,$$

where

$$G(t, x, y) = \mathbf{E}(x + y\tilde{\eta}_v - K)_+$$

and

$$\xi_t = \int_0^t S_v dv, \quad \tilde{\eta}_v = \int_0^v \exp\{\sigma W_u - \sigma^2 u/2\} du.$$

Moreover $G(t, x, y)$ is the unique solution of the following equation

$$\begin{cases} G'_t(t, x, y) + yG'_x(t, x, y) + \frac{\sigma^2}{2}y^2G''_{yy}(t, x, y) = 0 \\ G(1, x, y) = (x - K)_+. \end{cases} \quad (4)$$

Definition of strategies for the Asian options with transaction costs. Suppose that traders have to pay for a successful transaction some fee which is proportional to the trading volume. We assume that the cost proportion $\kappa_n = \kappa_0 n^{-\alpha}$. To compensate the transaction cost Leland [3] suggested to correct the volatility. The new parameter $\hat{\sigma}$ we have to put in the PDE (4) and calculate the strategy again with a new volatility. Applying the Leland approach we modify the strategy as follows

$$\gamma_t^n = \sum_{i=1}^n \hat{G}'_y(t_{j-1}, \xi_{t_{j-1}}, S_{t_{j-1}}) \chi_{(t_{j-1}, t_j]}(t),$$

where $\hat{G}'_y(t, x, y)$ is the solution of the equation (4) with parameter $\hat{\sigma}$.

Moreover $\hat{G}'_y(t, x, y)$ has the following form

$$\hat{G}'_y(t, x, y) = \int_b^\infty z \hat{q}(v, z) dz, \quad b = \frac{(K - x)_+}{y}$$

here $\hat{q}(v, z)$ is a density of random variable $\tilde{\eta}_v$ with new parameter $\hat{\sigma}$ and given by

$$\hat{q}(v, z) = \mathbf{E} \left(\frac{\varphi_{0,1}(\hat{a}(t, z))}{\hat{K}(v, \hat{a}(t, z))} \right),$$

$$\hat{K}(v, \hat{a}(t, z)) = \hat{\sigma} \int_0^v u \exp\{\hat{\sigma} \tilde{W}_u - \frac{\hat{\sigma}^2 u}{2} + \hat{\sigma} u \hat{a}(t, z)\}, \quad \tilde{W}_u = W_u - u W_1.$$

This form of density has been received in the article [7]. The portfolio value at t with the initial capital $V_0 = \hat{G}(0, \xi_0, S_0)$ has the form

$$V_t^n = \hat{G}(0, \xi_0, S_0) + \int_0^t \gamma_u^n dS_u - \kappa_n J_n, \quad (5)$$

where the total trading volume is given by

$$J_n = \sum_{j=1}^n S_{t_j} |\gamma_{t_j}^n - \gamma_{t_{j-1}}^n|.$$

In order to keep the hedging strategy it is necessary to satisfy the

following condition

$$V_1^n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} f_1.$$

For this we need to consider a hedging error

$$V_1^n - f_1 = \int_0^1 (\gamma_t^n - \hat{\gamma}_t) dS_t + \frac{\hat{\sigma}^2 - \sigma^2}{2} \int_0^1 \hat{G}_{yy}''(t, \xi_t, S_t) S_t^2 dt - \kappa_n J_n.$$

Main result. Recall that we need to choose new parameter volatility as

$$\hat{\sigma}^2 = \sigma^2 + \sigma \sqrt{n} \kappa_n \sqrt{\frac{8}{\pi}}, \quad \kappa_n = \kappa_0 n^{-\alpha}. \quad (6)$$

Then the following theorem is hold.

Theorem 1. For $\alpha = 1/2$ in (6) the portfolio value V_1^n converges in probability to the payoff f_1 as $n \rightarrow \infty$.

The proof of this theorem is based on the following lemmas.

Lemma 1. Let

$$B_n = \kappa_n \sum_{j=1}^n S_{t_j} |H(t_j, \xi_{t_j}, S_{t_j}) - H(t_{j-1}, \xi_{t_{j-1}}, S_{t_{j-1}})|, \quad (7)$$

$$\kappa_n \xrightarrow[n \rightarrow \infty]{} 0$$

Then

$$\mathbf{P} - \lim_{n \rightarrow \infty} B_n = 0 \quad (8)$$

Lemma 2. Let

$$D_n^{(y)} = \kappa_n \sum_{j=1}^n d_j^{(y)}$$

$$d_j^{(y)} = S_{t_j} |H(t_{j-1}, \xi_{t_{j-1}}, S_{t_j}) - H(t_{j-1}, \xi_{t_{j-1}}, S_{t_{j-1}})|$$

$$- H_y(t_{j-1}, \xi_{t_{j-1}}, S_{t_{j-1}})(S_{t_j} - S_{t_{j-1}})|.$$

Then

$$\mathbf{P} - \lim_{n \rightarrow \infty} D_n^{(y)} = 0. \quad (9)$$

Lemma 3. Let $\beta(t)$ is continuous consistent function almost sure. Then

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \beta(t_{j-1}) |S_{t_j} - S_{t_{j-1}}| \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \sqrt{\frac{2}{\pi}} \sigma \int_0^1 S_t \beta(t) dt.$$

References

- [1] Magill M.J.P., Constantinides G.M. Portfolio selection with transaction costs // J.Econ.Theory. 1976. Vol. 13. P. 245–263.

- [2] Merton R.C. Optimum consumption and portfolio rules in a continuous time model // J.Econ.Theory. 1971 Vol. 3. P. 373–413.
- [3] Leland H. Option pricing and replication with transaction costs // Journal of Finance. 1985. Vol. 40. P. 1283–1301.
- [4] Kabanov Yu.M., Safarian M. On Leland's strategy of option pricing with transaction costs // Finance and Stochastics. 1997. Vol. 1. P. 239–250.
- [5] Pergamenshchikov S. Limit theorem for Leland's strategy // Annals of Applied Probability. 2003. Vol. 13. P. 1099–1118.
- [6] Lepinette E. Marche avec couts de transaction: approximation de Leland et arbitrage.- PhD thesis, Universite de Franche-Comte Besanson, 2008.
- [7] Shishkova A.A. Calculation of Asian options for the Black Scholes model // Journal of Tomsk State University. Mathematics and Mechanics. 2018. Vol. 51. P. 48–63.

Шишкова А.А. (Томский государственный университет, Томск, 2019) **Задача хеджирования для азиатских опционов на финансовых рынках с транзакционными издержками.**

Аннотация. В данной статье изучены асимптотические свойства портфеля с транзакционными издержками в модели Блэка-Шоулса с рисковым активом без дрефта и безрисковым активом с процентной ставкой $r = 0$. Мы используем модификацию стратегии Леланда. Основным результатом являются полученные достаточные условия, которые обеспечивают асимптотическое хеджирование.

Ключевые слова: азиатский опцион, хеджирующая стратегия, Броуновское движение, модель Блэка-Шоулса, стохастические дифференциальные уравнения.

Асимптотическая мощность и качественная робастность критерия Пирсона для проверки нормальности авторегрессии

Болдин М. В.

Московский государственный университет им. М.В. Ломоносова,
Москва
e-mail: boldin_m@hotmail.com

Аннотация

Мы рассматриваем стационарную линейную $AR(p)$ модель в ситуации, когда наблюдения содержат грубые ошибки (выбросы или засорения). Распределение выбросов неизвестно и произвольно, их интенсивность $\gamma n^{-1/2}$, $\gamma \geq 0$ неизвестно, n — объем данных. Параметры авторегрессии неизвестны, распределение инноваций тоже. По остаткам от $n^{1/2}$ -состоятельных оценок параметров строится подобие эмпирической функции распределения и на ней основывается тест типа хи-квадрат Пирсона для проверки нормальности распределения инноваций. Найдена асимптотическая мощность теста при локальных альтернативах. Установлена качественная робастность теста при гипотезе и альтернативах.

Ключевые слова: авторегрессия, выбросы, остатки, эмпирическая функция распределения, тест хи-квадрат Пирсона, робастность, локальные альтернативы.

Введение и постановка задачи. Рассмотрим линейную $AR(p)$ модель

$$u_t = \beta_1 u_{t-1} + \dots + \beta_p u_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (1)$$

где $\{\varepsilon_t\}$ — независимые одинаково распределенные случайные величины (н.о.р.с.в.) с неизвестной функцией распределения (ф.р.) $G(x)$; $E\varepsilon_1 = 0$, $0 < E\varepsilon_1^2 < \infty$; $\beta = (\beta_1, \dots, \beta_p)^T$ — вектор неизвестных параметров, таких что корни соответствующего характеристического уравнения по модулю меньше единицы.

В этой заметке мы рассматриваем модель (1) с выбросами в наблюдениях. А именно, предполагается, что наблюдаются переменные

$$y_t = u_t + z_t^{\gamma n} \xi_t, \quad t = 1 - p, \dots, n, \quad (2)$$

где $u_{1-p}, \dots, u_n$ — выборка из стационарного решения $\{u_t\}$ уравнения (1), $\{z_t^{\gamma n}\}$ — н.о.р.с.в., распределенные по закону Бернулли с параметром γ_n ,

$$\gamma_n = \min(1, \frac{\gamma}{\sqrt{n}}), \quad \gamma \geq 0 \text{ неизвестно.}$$

Кроме того, $\{\xi_t\}$ – н.о.р.с.в. с неизвестным и произвольным распределением Π . Последовательности $\{u_t\}$, $\{z_t^{\gamma_n}\}$, $\{\xi_t\}$ предполагаются независимыми.

Случайные величины $\{\xi_t\}$ интерпретируются как выбросы (засорения), γ_n – уровень засорения. При $\gamma = 0$ получаем модель (1) без засорений.

Схема типа (2) с постоянным не зависящим от n уровнем засорения введена в [4], где используется для характеристики робастности оценок функционалами влияния. Схема (2) с зависящим от n уровнем засорения – локальный вариант схемы засорения из [4].

Мы интересуемся в этой работе проверкой по наблюдениям (2) гипотезы

$$H_\Phi: G(x) \in \{\Phi(x/\theta), \theta > 0\},$$

где $\Phi(x)$ стандартная нормальная ф.р. Гипотеза H_Φ , напомним, эквивалентна нормальности самой авторегрессионной последовательности $\{u_t\}$.

Предположение о нормальности авторегрессии обеспечивает оптимальность общепотребительных процедур наименьших квадратов оценивания и проверки гипотез в авторегрессии. Это объясняет актуальность для статистической теории и приложений рассматриваемой далее задачи.

Эта задача (помимо прочих) уже решалась в [1]. В этой работе в схеме (1) – (2) изучалась остаточная эмпирическая функция распределения (о.э.ф.р.) и основанные на ней тесты для проверки гипотез относительно $G(x)$.

Вот как строится о.э.ф.р. Рассмотрим схему (1) – (2). Пусть $\hat{\beta}_n = (\hat{\beta}_{1n}, \dots, \hat{\beta}_{pn})^T$ будет любая равномерно по $\gamma \leq \Gamma < \infty$ $n^{1/2}$ -состоятельная оценка β . Величины

$$\hat{\varepsilon}_t = y_t - \hat{\beta}_{1n}y_{t-1} - \dots - \hat{\beta}_{pn}y_{t-p}, \quad t = 1, \dots, n, \quad (3)$$

называются остатками, а функция

$$\hat{G}_n(x) = n^{-1} \sum_{t=1}^n I(\hat{\varepsilon}_t \leq x), \quad x \in \mathbb{R}^1, \quad (4)$$

– остаточная эмпирическая функция распределения. Здесь и далее $I(\cdot)$ обозначает индикатор события.

Функция $\hat{G}_n(x)$ – аналог гипотетической э.ф.р.

$$G_n(x) = n^{-1} \sum_{t=1}^n I(\varepsilon_t \leq x)$$

ненаблюдаемых величин $\varepsilon_1, \dots, \varepsilon_n$.

В [1] показано: если $G(x)$ дважды дифференцируема с производной $g(x) = G'(x)$ и $\sup_x |g'(x)| < \infty$, то

$$\sup_{\gamma \leq \Gamma} P(|n^{1/2}[\hat{G}_n(x) - G_n(x)] - \gamma \Delta(x, \Pi)| > \delta) \rightarrow 0, \quad n \rightarrow \infty, \quad (5)$$

при любых $0 \leq \Gamma < \infty$, $\delta > 0$. Здесь сдвиг, возникающий из-за засорений, имеет вид:

$$\Delta(x, \Pi) = \sum_{j=0}^p [EG(x + \beta_j \xi_1) - G(x)], \quad \beta_0 = -1.$$

Поскольку статистики типа хи-квадрат Пирсона являются простыми функционалами соответствующей э.ф.р., утверждения (5) достаточно, чтобы построить по $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n$ тесты типа хи-квадрат Пирсона для проверки гипотезы H_Φ .

Соответствующая статистика $\hat{\chi}_n^2$ будет определена в следующем Разделе 2. Ее асимптотическое распределение при гипотезе было найдено в [1]. Разумеется, предельное распределение $\hat{\chi}_n^2$ при гипотезе зависит от γ и Π . Важно, что предельный уровень значимости $\alpha(\gamma, \Pi)$ соответствующего теста таков, что если для $0 < \alpha < 1$ $\alpha(0, \Pi) = \alpha$, то

$$\sup_{\Pi} |\alpha(\gamma, \Pi) - \alpha| \rightarrow 0, \quad \gamma \rightarrow 0. \quad (6)$$

Соотношение (6) означает равностепенную непрерывность семейства $\{\alpha(\gamma, \Pi)\}$ по γ в точке $\gamma = 0$. Мы интерпретируем его как асимптотическую качественную устойчивость предельного уровня значимости нашего теста хи-квадрат относительно засорений. Такая устойчивость означает, что при малых γ гипотезу H_Φ можно проверять примерно на асимптотическом уровне α независимо от Π . Этим свойством, кстати, не обладают тесты типа Колмогорова и омега-квадрат для H_Φ , построенные в схеме (1) – (2). См. Замечание 2.3 в [1].

Еще раз подчеркнем: упомянутые результаты были получены только при гипотезе H_Φ , мощность теста при локальных (близких) альтернативах не исследовалась. А значит, не исследовалась и качественная робастность теста при таких альтернативах. В этой работе мы устраним отмеченный пробел и исследуем мощность теста при альтернативах, сближающихся с гипотезой при $n \rightarrow \infty$ со скоростью $O(n^{1/2})$.

Неизвестную при H_Φ дисперсию инноваций будем обозначать θ_0^2 . Тогда при H_Φ ф.р. $G(x) = \Phi(x/\theta_0)$. Далее будем изучать альтернативную к H_Φ ситуацию. А именно, будем предполагать, что ф.р. инноваций $G(x)$ является смесью:

$$G(x) \in \{(1 - \rho_n)\Phi(x/\theta_0) + \rho_n H(x)\}, \quad (7)$$

$$H(x) - \text{ф.р.}, \quad \rho_n = \min(1, \frac{\rho}{\sqrt{n}}), \quad \rho \geq 0.$$

Предположение (7) будем понимать как альтернативную к H_Φ гипотезу и обозначать A_Φ . Разумеется, при $\rho = 0$ H_Φ и A_Φ совпадают. Гипотезу A_Φ будем называть локальной альтернативой (хотя никакой последовательности регулярных статистических экспериментов

мы не определяем).

Цель настоящей работы – найти при A_Φ асимптотическую мощность теста со статистикой $\hat{\chi}_n^2$ и установить асимптотическую качественную робастность теста при локальных альтернативах (7).

Разумеется, для решения этой задачи нам понадобится аналог разложения (5) при альтернативе A_Φ . Необходимый результат был получен в [2], и мы приведем его далее.

В заключение раздела заметим, что в линейных и нелинейных авторегрессионных схемах без засорений о.э.ф.р. исследовалась во многих работах, см. библиографию в [1]. В схемах без засорений удается получить даже равномерные по x разложения о.э.ф.р. и построить тесты типа Колмогорова и омега-квадрат для H_Φ , см. [3]. Но мощность этих тестов при локальных альтернативах не исследовалась.

Основные результаты работы представлены в разделе 2.

Основные результаты. Обозначим ф.р. инноваций при A_Φ из (7) через $A_n(x)$, тогда

$$A_n(x) = (1 - \rho_n)\Phi(x/\theta_0) + \rho_n H(x) \quad (8)$$

при некоторых неизвестных $\theta_0 > 0$, $\rho \geq 0$ и ф.р. $H(x)$.

Нам потребуются два условия.

Условие (i). Случайная величина с ф.р. $H(x)$ имеет нулевое среднее и конечную дисперсию.

Условие (ii). Ф.р. $H(x)$ дважды дифференцируема с ограниченной второй производной.

Опишем статистику типа хи-квадрат Пирсона для проверки H_Φ

Пусть $\hat{\beta}_n$ будет любая равномерно по $\rho \leq R$ и $\gamma \leq \Gamma$ $n^{1/2}$ -состоятельная при A_Φ оценка β . (R, Γ здесь и далее – любые конечные неотрицательные числа.) Например, годится GM-оценка из Раздела 2.3 в [1].

Пусть остатки $\{\hat{\varepsilon}_t, t = 1, \dots, n\}$ определены соотношением (3). Разобьем действительную прямую на m непересекающихся полуинтервалов $B_1, \dots, B_m$,

$B_j = (x_{j-1}, x_j]$, $j = 1, \dots, m$, $m > 2$, $-\infty = x_0 < x_1 < \dots < x_m = \infty$, и положим

$$p_j(\theta) := \Phi(x_j/\theta) - \Phi(x_{j-1}/\theta).$$

Обозначим через $\hat{\nu}_j$ число остатков среди $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n$, попавших в B_j .

Оценку θ_0 (обозначим ее $\hat{\theta}_n$) будем основывать на решении уравнения

$$\sum_{j=1}^n \frac{\hat{\nu}_j}{p_j(\hat{\theta}_n)} p_j'(\hat{\theta}_n) = 0, \quad (9)$$

в котором штрих означает производную. Точнее, оценку полагаем равной равномерно по $\rho \leq R$, $\gamma \leq \Gamma$ $n^{1/2}$ -состоятельному решению (9) (оно существует с вероятностью, стремящейся к единице при $n \rightarrow \infty$), и произвольному положительному числу, если упомянутого решения нет. Такая оценка $\hat{\theta}_n$ асимптотически нормальна и слабая сходимость к предельному гауссовскому закону равномерна по $\rho \leq R$, $\gamma \leq \Gamma$.

Уравнение (9) – аналог обычного уравнения видоизмененного метода минимума хи-квадрат, в котором неизвестные частоты $\nu_1, \dots, \nu_n$ заменены на оценки $\hat{\nu}_1, \dots, \hat{\nu}_n$. (Здесь ν_j – число с.в. среди $\varepsilon_1, \dots, \varepsilon_n$, попавших в B_j .)

Очевидно,

$$\hat{\nu}_j = n[\hat{G}_n(x_j) - \hat{G}_n(x_{j-1})], \quad j = 1, \dots, m. \quad (10)$$

о.э.ф.р. $\hat{G}_n(x)$ определена в (4).

Интересующая нас статистика типа хи-квадрат Пирсона для проверки H_Φ имеет вид

$$\hat{\chi}_n^2 = \sum_{j=1}^m \frac{(\hat{\nu}_j - np_j(\hat{\theta}_n))^2}{np_j(\hat{\theta}_n)}. \quad (11)$$

Это аналог стандартной статистики Пирсона с оцененным методом минимума хи-квадрат параметром θ_0 .

В силу (9) – (11) статистика $\hat{\chi}_n^2$ является функционалом от $\hat{G}_n(x)$. Поэтому ее асимптотическое исследование основывается на следующем стохастическом разложении.

Теорема 1. Пусть верна альтернатива A_Φ , т.е. функция распределения инноваций $A_n(x)$ задается соотношением (8). Предположим, что для $H(x)$ выполнены Условия (i) –(ii). Пусть $G_0(x) := \Phi(x/\theta_0)$,

$$\Delta_0(x, \Pi) := \sum_{j=0}^p [EG_0(x + \beta_j \xi_1) - G_0(x)], \quad \beta_0 = -1.$$

Тогда при любом $\delta > 0$

$$\sup_{\rho \leq R, \gamma \leq \Gamma} P(|n^{1/2}[\hat{G}_n(x) - G_n(x)] - \gamma \Delta_0(x, \Pi)| > \delta) \rightarrow 0, \quad n \rightarrow \infty.$$

Напомним, $G_n(x)$ – э.ф.р. величин $\varepsilon_1, \dots, \varepsilon_n$.

Сформулированная здесь теорема 1 следует из теоремы 2.1 и следствия 2.1, доказанных в [2].

Чтобы сформулировать основной результат настоящей работы – теорему 2, нам потребуются некоторые обозначения.

Определим диагональную матрицу

$$\mathbf{P}_0 = \text{diag}\{p_1(\theta_0), \dots, p_m(\theta_0)\},$$

и векторы $\delta_0(\Pi) = (\delta_1^0(\Pi), \dots, \delta_m^0(\Pi))^T$ с компонентами $\delta_j^0(\Pi) :=$

$$\Delta_0(x_j, \Pi) - \Delta_0(x_{j-1}, \Pi),$$

$$\mathbf{p}_0 = (p_1(\theta_0), \dots, p_m(\theta_0))^T,$$

$$\mathbf{p}'_0 = (p'_1(\theta_0), \dots, p'_m(\theta_0))^T, \quad \mathbf{b}_0 = \mathbf{P}_0^{-1/2} \mathbf{p}'_0, \quad \mathbf{e}_0 = \mathbf{b}_0/|\mathbf{b}_0|,$$

$$\mathbf{p}_H = (p_1^H, \dots, p_m^H)^T \text{ с компонентами } p_j^H := H(x_j) - H(x_{j-1}).$$

Отметим, что вектор $\mathbf{p}_0$ имеет ненулевые компоненты, а вектор $\mathbf{b}_0$ ненулевой, так что матрица $\mathbf{P}_0^{-1/2}$ и вектор $\mathbf{e}_0$ определены.

Далее $F_{m-2}(x, \lambda^2(\rho, \gamma, \Pi))$ обозначает ф.р. нецентрального хи-квадрата с $m - 2$ степенями свободы и параметром нецентральности $\lambda^2(\rho, \gamma, \Pi)$. Через $|\cdot|$ будем обозначать Евклидову норму вектора или матрицы, а через $\mathbf{E}_m$ — единичную матрицу порядка m .

Теорема 2. *Предположим, что верна альтернатива A_Φ . Тогда при Условиях (i) – (ii) для любых конечных $R \geq 0$, $\Gamma \geq 0$ и для любого $\delta > 0$*

$$\sup_{x \in \mathbb{R}^1, \rho \leq R, \gamma \leq \Gamma} |\mathbb{P}(\hat{\chi}_n^2 \leq x) - F_{m-2}(x, \lambda^2(\rho, \gamma, \Pi))| \rightarrow 0, \quad n \rightarrow \infty.$$

Параметр нецентральности

$$\lambda^2(\rho, \gamma, \Pi) = |(\mathbf{E}_m - \mathbf{e}_0 \mathbf{e}_0^T) \mathbf{P}_0^{-1/2} [\rho(\mathbf{p}_H - \mathbf{p}_0) + \gamma \delta_0(\Pi)]|^2.$$

Теорема 2 означает равномерную по $\rho \leq R$, $\gamma \leq \Gamma$ слабую сходимость $\hat{\chi}_n^2$ к случайной величине с ф.р. $F_{m-2}(x, \lambda^2(\rho, \gamma, \Pi))$.

Доказательство теоремы 2 довольно кропотливо и здесь не приводится. Оно основано на Теореме 1 и схема рассуждений аналогична доказательству теоремы 2.3 в [1].

Обозначим ф.р. хи-квадрата с $m - 2$ степенями свободы через $F_{m-2}(x)$, и пусть $(1 - \alpha)$ -квантиль этой ф.р. с $0 < \alpha < 1$ будет $\chi_{m-2}(1 - \alpha)$.

Будем отвергать H_Φ , если

$$\hat{\chi}_n^2 > \chi_{m-2}(1 - \alpha). \quad (12)$$

Мощность такого теста равна

$$W_n(\rho, \gamma, \Pi) = P(\hat{\chi}_n^2 > \chi_{m-2}(1 - \alpha)).$$

В силу теоремы 2 эта мощность сходится при $n \rightarrow \infty$ равномерно по $\rho \leq R$, $\gamma \leq \Gamma$ к асимптотической мощности

$$W(\rho, \gamma, \Pi) = 1 - F_{m-2}(\chi_{m-2}(1 - \alpha), \lambda^2(\rho, \gamma, \Pi)), \quad W(0, 0, \Pi) = \alpha.$$

Используя неравенство

$$|F_k(x, \lambda_1^2) - F_k(x, \lambda_2^2)| \leq 2 \sup_{x \in \mathbb{R}^1} \phi(x) |\lambda_1 - \lambda_2|,$$

где $\phi(x)$ стандартная гауссовская плотность, и определение $\lambda^2(\rho, \gamma, \Pi)$ в теореме 2, мы получаем:

$$\begin{aligned} & \sup_{\rho \geq 0, \Pi} |W(\rho, \gamma, \Pi) - W(\rho, 0, \Pi)| \\ & \leq \gamma \sqrt{2/\pi} |(\mathbf{E}_m - \mathbf{e}_0 \mathbf{e}_0^T) \mathbf{P}_0^{-1/2}| \sup_{\Pi} |\delta^0(\Pi)| \rightarrow 0, \quad \gamma \rightarrow 0. \end{aligned} \quad (13)$$

Соотношение (13) означает равностепенную непрерывность семейства $\{W(\rho, \gamma, \Pi)\}$ по γ в точке $\gamma = 0$. Мы интерпретируем это свой-

ство как асимптотическую качественную робастность теста (12). Такая робастность означает, что при малых γ независимо от распределения засорений Π гипотезу H_F все еще можно проверять примерно с асимптотическим уровнем α , и тест будет иметь примерно такую же асимптотическую мощность при A_F с $\rho > 0$, как в схеме без засорений. Весьма привлекательное свойство.

Список литературы

- [1] Boldin M.V. and Petriev M.N. On the Empirical Distribution Function of Residuals in Autoregression with Outliers and Pearson's Chi-Square Type Tests // Math. Methods Statist. 2018. Vol. 27, No. 4. P. 1–17.
- [2] Boldin M.V. On the Power of Pearson's Test under Local Alternatives in Autoregression with Outliers // Math. Methods Statist. 2019. Vol. 28, No. 1. P. 144–154.
- [3] Болдин М.В., Симонова Г.И., Тюрин Ю.Н. Знаковый статистический анализ линейных моделей. М.: Наука; P.I. 1997, 285 с.
- [4] Martin B.D. and Yohai V.J. Influence functionals for Time Series // The Ann. of Statist. 1986. Vol. 14. P. 781–818.

Boldin M. V. (Lomonosov Moscow State University, Moscow, 2019)
Asymptotic Power and Qualitative Robustness of Pearson's Test for Testing the Normality Hypothesis in Autoregression.

Abstract. We consider a stationary linear $AR(p)$ model with observations subject to gross errors (outliers). The distribution of outliers is unknown and arbitrary, their intensity is $\gamma n^{-1/2}$ with an unknown γ , n is the sample size. The autoregression parameters are unknown, they are estimated by any estimator which is $n^{1/2}$ -consistent. Using the residuals from the estimated autoregression, we construct a kind of empirical distribution function, which enables us to construct a test of Pearson's chi-square type for testing of normality of autoregression. We establish qualitative robustness of this test and their asymptotic power under local alternatives.

Key words: autoregression, residuals, empirical distribution function, outliers, Pearson's chi-square test, robustness, close alternatives.

Сегментация изображений на основе алгоритма обнаружения разладки¹

Воробейчиков С. Э., Удод В. А.

Национальный исследовательский Томский государственный
университет, Томск
e-mail: sev@mail.tsu.ru

Аннотация

Рассматривается задача сегментации цифрового изображения на однородные в смысле постоянного значения яркости фрагменты. Для определения границ фрагментов используется алгоритм обнаружения разладки [2], который модифицирован для обнаружения многократных разладок с различными уровнями среднего значения яркости на отдельных участках. Построенный многоуровневый алгоритм применяется последовательно к каждой строке и столбцу изображения. В каждом из полученных фрагментов вычисляется среднее значение яркости и все элементы фрагмента получают это значение. Приводятся результаты численного моделирования.

Ключевые слова: алгоритм сегментации, цифровое изображение, обнаружение момента разладки.

Постановка задачи. Одной из задач, возникающих при обработке изображений, является их сегментация на непересекающиеся фрагменты, однородные по некоторому признаку (цвету, яркости и т.д.). Этап сегментации является предварительным и может использоваться в дальнейшем в задачах дефектоскопии [1, 3].

Предполагается, что цифровое изображение размера $n \times m$ описывается уравнениями

$$B(i, j) = B_f + F(i, j) + B_N(i, j) \quad i = \overline{1, n}; \quad j = \overline{1, m}. \quad (1)$$

Здесь B_f - фон; $F(i, j)$ - функция "яркости" изображения (сигнальная составляющая), обусловленная наличием объектов на изображении; $B_N(i, j)$ - шум. Задача – произвести сегментацию изображения на однородные по яркости фрагменты. Для этого предлагается использовать модифицированный алгоритм обнаружения разладки (кумулятивных сумм) [2].

¹Работа выполнена при финансовой поддержке Министерства образования и науки РФ, Госзадание 2.3208.2017/4.6

Алгоритм обнаружения разладки. Рассмотрим ситуацию, когда последовательность наблюдений x_n является последовательностью независимых гауссовских случайных величин, у которых дисперсия является постоянной величиной, а среднее значение в момент разладки меняется с a_0 на a_1 . Тогда согласно [2] алгоритм кумулятивных сумм (АКС) определяется следующим образом: вычисляются последовательно суммы

$$S_{n+1} = \max(S_n + x_{n+1} - r, 0), \quad S_0 = 0, \quad r = (a_1 + a_0)/2, \quad a_1 > a_0, \quad (2)$$

и решение о наличии разладки принимается, когда значение S_n превысит некоторый порог h , являющийся параметром алгоритма. В случае $a_1 < a_0$ используется другая сумма

$$C_{n+1} = \max(C_n - x_{n+1} + r, 0), \quad C_0 = 0. \quad (3)$$

В качестве оценки момента разладки принимается значение $k + 1$, где k – то значение n , при котором сумма S_n (или C_n) последний раз равнялась нулю перед принятием решения о наличии разладки.

Отметим, что АКС ориентирован на обнаружение однократной разладки, связанной с увеличением (или уменьшением) среднего значения в наблюдаемой последовательности. При сегментации отдельной строки изображения может происходить многократное изменение среднего значения. Рассмотрим следующий пример. Пусть процесс $x(n)$ описывается уравнениями

$$x_n = s_n + \xi_n, \quad \xi_n \sim N(0, 0.5), \quad n = 1, \dots, 500,$$

$$s_n = \begin{cases} 1, & n \leq 300, \quad n > 400 \\ 2, & 301 \leq n \leq 400. \end{cases}$$

Таким образом, среднее значение процесса x_n изменяется дважды: увеличивается на единицу при $n = 300$ и уменьшается на единицу в точке $n = 401$ и на интервале наблюдений имеется два момента разладки. График процесса x_n представлен на рисунке 1. Далее

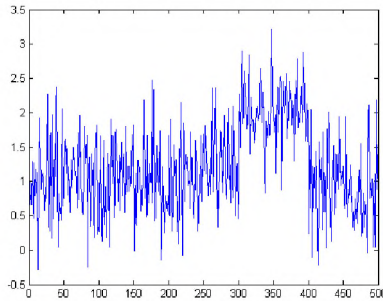


Рис. 1. Наблюдаемый процесс x_n

построим графики кумулятивных сумм

$$\begin{aligned} S_{n+1} &= \max(S_n + x_{n+1} - 1.5, 0), \\ C_{n+1} &= \max(C_n - x_{n+1} + 1.5, 0), \quad h = 10.0. \end{aligned}$$

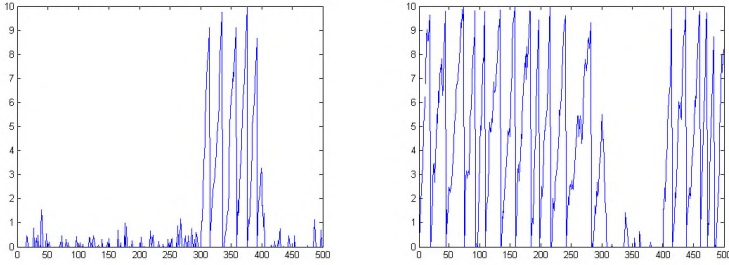


Рис. 2. Кумулятивные суммы S_n и C_n

Отметим, что при превышении любой суммой порога h ее значение обнуляется и суммирование продолжается со следующего значения x_n . Из рисунков 2 видно, что сумма S_n , ориентированная на обнаружение увеличения среднего, принимает значения, близкие к нулю на интервалах $[1, 300]$ и $(401, 500]$ и несколько раз достигает порога h на отрезке $[300, 400]$, где среднее значение наблюдаемого процесса имеет большее значение. Соответственно, сумма C_n многократно превышает порог h на интервалах $[1, 300]$ и $(401, 500]$ и принимает значения, близкие к нулю на отрезке $[300, 400]$.

Для обнаружения многократной разладки на интервале наблюдений будем использовать одновременно две суммы S_n и C_n и фиксировать все моменты t_i превышений суммами порога h . Если в последовательности t_i несколько значений подряд получены за счет превышения порога одной и той же суммой, то будем оставлять только первое из этих значений. В результате в последовательности t_i будут оставлены только те значения, в которых соседние элементы получены при превышении порога различными типами сумм. При этом будут получены оценки моментов разладки. Отметим, что алгоритм обнаружения многократной разладки с использованием только одной суммы S_n рассматривался в [4]. При этом требовалось иметь всю реализацию процесса x_n . Достоинством предложенного в данной работе алгоритма обнаружения многократных разладок является возможность использования его при последовательном поступлении данных.

Сегментация изображений. Описанный алгоритм обнаружения многократной разладки используем для выделения однородных по яркости участков цифрового изображения. Зададим возрастающие граничные значения яркостей B_k , $k = 0, \dots, N$. Например, эти значения могут выбираться равномерно в диапазоне от минимального до максимального значения яркостей точек изображения. Далее выделяются те участки изображения, которые можно считать однородными для диапазона яркостей $(B_{k-1}, B_k]$. Сначала значение k полагается равным N . Изображение "просматривается" построчно и на каждой строке выделяются участки однородности "k-го уровня" (строчные сегменты "k-го уровня") с помощью описанного выше алгоритма обнаружения многократной разладки $S_j = \max(S_{j-1} + x_j - r_k, 0)$, $S_0 = 0$, $r_k = (B_k + B_{k-1})/2$, $j = 1, \dots, m$, $C_j = \max(C_{j-1} - x_j + r_k, 0)$, $C_0 = 0$, $j = 1, \dots, m$, $h = 2(B_k - B_{k-1})$. Здесь x_j представляет собой j -й элемент текущей строки изображения. Далее участки однородности "k-го уровня" выделяются аналогичным образом на каждом столбце изображения (столбцовые сегменты "k-го уровня"). В результате объединения полученных сегментов получаем множество Q_k .

Затем значение k уменьшается на 1 и этап выделения однородных сегментов следующего уровня повторяется. При этом во множество Q_k включаем только те точки изображения, которые не вошли во множества Q_p , $p = k + 1, \dots, N$, которые были получены на предыдущих шагах. В результате каждая точка изображения будет включена только в одно из множеств Q_k .

На следующем шаге алгоритма каждое из множеств Q_k разбивается на непересекающиеся подмножества. На каждом из полученных подмножеств вычисляется среднее значение яркости и каждая точка сегментированного подмножества получает значение, равное вычисленному среднему. В результате будет получено сегментированное изображение.

Результаты моделирования. Для иллюстрации работы алгоритма сегментации рассмотрим следующий пример. Смоделированное изображение имеет размер 100×100 . Значение фона B_f в (1) равно 100. Объектом на изображении является квадрат с тремя уровнями яркости внутри вложенных квадратов, отличающимися по яркости друг от друга и от фона на 10 единиц (от 110 до 130). Изображение искажено шумом $B_N(i, j)$, представляющим собой независимые гауссовские случайные величины с нулевым средним и среднеквадратическим отклонением $\sigma = 5.56$. Исходное и зашумленное изображения представлены на рисунке 3.

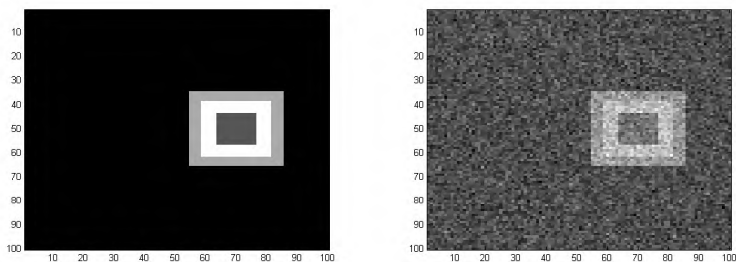


Рис. 3. Исходное и зашумленное изображение

Результаты применения алгоритма сегментации приведены на рисунке 4.

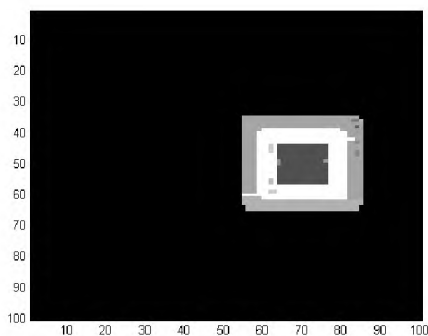


Рис. 4. Сегментированное изображение

По результатам моделирования можно сделать вывод, что предложенный алгоритм сегментации обеспечивает достаточно высокое качество при высоком уровне шумов.

Список литературы

- [1] Baniukiewicz P. Automated Defect Recognition and Identification in Digital Radiography// J. Nondestruct. Eval. 2014. V. 33. No. 3. P. 327-334. doi: 10.1007/s10921-013-0216-6.
- [2] Page E.S. Continuous Inspection Schemes// Biometrika. 1954. V. 4. No. 1. P. 141-154.

- [3] Park J.S., Kim J.K. Calculation of effective atomic number and normal density using a source weighting method in a dual energy X-ray inspection system// Journal of the Korean physical society. 2011. V. 59. No. 4. P. 2709-2713.
- [4] Vorobeychikov S.E., Chakhlov S.V., Udod V.A. A Cumulative sums algorithm for segmentation of digital X-ray images//J. Nondestruct. Eval. 2019. V. 38. doi: 10.1007/s10921-019-0616-3.

Vorobeychikov S. E., Udod V. A. (Tomsk State University, Tomsk, 2019) **Image segmentation on the base of change point algorithm.**

Abstract. The problem of image segmentation is considered. A change-point detection algorithm ([2]) is applied to detect the edges of segments homogeneous in brightness. This algorithm has been modified to detect the multiple change-points with different levels of brightness. The constructed algorithm is repeatedly applying to every row and every column of the image. The average value of the brightness over each fragment is calculated and each point of the fragment gets this value. The results of simulation are presented.

Key words: segmentation algorithm, digital image, change-point detection.

Оценивание непараметрической регрессии с полумарковскими шумами¹

Жуман А. Б., Пчелинцев Е. А.

Томский государственный университет, Томск

e-mail: aigerima_z@list.ru

Аннотация

В данной работе рассматривается задача оценивания неизвестной функции в непрерывной регрессии с полумарковскими шумами. Приводятся результаты численного моделирования эмпирических рисков взвешенных оценок наименьших квадратов и их улучшенных версий, а также исследуется влияние мешающих параметров шума на качество процедур оценивания.

Ключевые слова: непараметрическая регрессия, взвешенные оценки наименьших квадратов, полумарковский процесс, среднеквадратическая точность.

Постановка задачи. Рассмотрим следующую регрессионную модель с непрерывным временем

$$dy_t = S(t)dt + d\xi_t, \quad 0 \leq t \leq n, \quad (1)$$

где $S(\cdot)$ – неизвестная 1-периодическая $\mathbb{R} \rightarrow \mathbb{R}$ функция из пространства $L_2[0, 1]$, процесс $(\xi_t)_{t \geq 0}$ – квадратично интегрируемый семимартингал шум, определяемый равенством

$$\xi_t = \rho_1 w_t + \rho_2 z_t, \quad (2)$$

где ρ_1, ρ_2 – неизвестные коэффициенты, $(w_t)_{t \geq 0}$ – винеровский процесс, $(z_t)_{t \geq 0}$ – полумарковский процесс вида

$$z_t = \sum_{i=1}^{N_t} Y_i, \quad (3)$$

где $(Y_i)_{i \geq 1}$ – последовательность независимых одинаково распределенных случайных величин с $\mathbf{E}Y_i = 0$, $\mathbf{E}Y_i^2 = 1$ и $\mathbf{E}Y_i^4 < \infty$. Здесь процесс N_t – обобщенный считающий процесс, определенный как

$$N_t = \sum_{k=1}^{\infty} 1_{\{T_k \leq t\}} \quad \text{и} \quad T_k = \sum_{l=1}^k \tau_l, \quad (4)$$

где $(\tau_l)_{l \geq 1}$ – последовательность независимых одинаково распределенных случайных величин с математическим ожиданием $\tilde{\tau} = \mathbf{E}\tau_1 > 0$. Процесс (2) подробно изучен в работе [1].

¹The work was supported by the RSF, the project No 17-11-01049.

Задача состоит в том, чтобы оценить неизвестную функцию $S(\cdot)$ в модели (1) на основе наблюдений $(y_t)_{0 \leq t \leq n}$ и исследовать влияние мешающих параметров шумового процесса (2) на качество оценивания.

Отметим, что регрессионные модели типа (1) широко применяются в задачах обработки сигналов и на финансовых рынках [2, 3].

Взвешенные оценки наименьших квадратов. Для оценивания неизвестной функции $S(\cdot)$ в модели (1) используем ее разложение в ряд Фурье по тригонометрическому базису $(\phi_j)_{j \geq 1}$ в $L_2[0, 1]$, который определяется как

$$\phi_1 \equiv 1, \quad \phi_j = \sqrt{2} \text{Tr}_j(2\pi [j/2] x), \quad j \geq 2, \quad (5)$$

где функция $\text{Tr}_j(\cdot) = \cos(\cdot)$ для четных j и $\text{Tr}_j(\cdot) = \sin(\cdot)$ для нечетных j , $[x]$ — целая часть x .

Для эффективного оценивания необходимо использовать взвешенные оценки наименьших квадратов, определяемые формулой

$$\hat{S}_\lambda(t) = \sum_{j=1}^n \lambda(j) \hat{\theta}_{j,n} \phi_j(t), \quad (6)$$

где

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

а коэффициенты $\lambda = (\lambda(j))_{1 \leq j \leq n}$ принадлежат некоторому конечному множеству Λ из $[0, 1]^n$.

Определим весовые коэффициенты $(\lambda(j))_{1 \leq j \leq n}$. Рассмотрим для некоторого фиксированного $0 < \varepsilon < 1$ числовую сетку вида

$$\mathcal{A} = \{1, \dots, k^*\} \times \{\varepsilon, \dots, m\varepsilon\}, \quad (8)$$

где $m = \lceil 1/\varepsilon^2 \rceil$. Предполагаем, что оба параметра $k^* > 0$ и ε — функции от n , т.е. $k^* = k^*(n)$ и $\varepsilon = \varepsilon(n)$ такие, что

$$\begin{cases} \lim_{n \rightarrow \infty} k^*(n) = +\infty, \\ \lim_{n \rightarrow \infty} \frac{k^*(n)}{\ln n} = 0, \\ \lim_{n \rightarrow \infty} \varepsilon(n) = 0, \\ \lim_{n \rightarrow \infty} n^\delta \varepsilon(n) = +\infty \end{cases} \quad (9)$$

для любого $\delta > 0$.

Для каждого $\alpha = (\beta, l) \in \mathcal{A}$ введем последовательность весов $\lambda_\alpha = (\lambda(j))_{1 \leq j \leq n}$ с элементами

$$\lambda_\alpha(j) = 1_{\{1 \leq j \leq d\}} + \left(1 - (j/\omega_\alpha)^\beta\right) 1_{\{d < j \leq \omega_\alpha\}}, \quad (10)$$

где $d = \lfloor \ln v_n \rfloor$, $\omega_\alpha = (d_\beta l v_n)^{1/(2\beta+1)}$, $d_\beta = \frac{(\beta+1)(2\beta+1)}{\pi^{2\beta}\beta}$, $v_n = \frac{n}{\sigma}$, $\sigma = \rho_1^2 + \rho_2^2/\tilde{\tau}$ [1].

Улучшенные взвешенные оценки наименьших квадратов. Теперь рассмотрим метод улучшенного оценивания для модели (1), предложенный в [2].

Предложение 1. *Существует такое $n_0 \geq 1$, что для любого $n \geq n_0$ существует σ -алгебра $\mathcal{G}_n$, для которой случайный вектор $\tilde{\xi}_{d,n} = (\xi_{j,n})_{1 \leq j \leq d}$ является условно гауссовским относительно $\mathcal{G}_n$ на $\mathbb{R}^d$ с матрицей ковариаций*

$$G_n = (\mathbf{E}(\xi_{i,n}\xi_{j,n}|\mathcal{G}_n))_{1 \leq i,j \leq d} \quad (11)$$

и для некоторой неслучайной константы $l_n^ > 0$*

$$\inf_{Q \in \mathcal{Q}_n} \mathbf{tr} G_n - \lambda_{\max}(G_n) \geq l_n^*, \quad (12)$$

где $\lambda_{\max}(G_n)$ – максимальное собственное значение матрицы G_n .

Для модели (1) – (2) имеем следующий результат.

Предложение 2. *Пусть в модели (1) шумовой процесс описывается как в (3), тогда $l_n^* = (d-1)\underline{\varrho}$, где $0 < \underline{\varrho} \leq \rho_1^2$.*

Доказательства данных предложений можно найти в [3].

Теперь для первых d коэффициентов Фурье используем улучшенные оценки вида

$$\theta_{j,n}^* = \left((1 - g(j))\hat{\theta}_{j,n} \right), \quad (13)$$

где $g(j) = \frac{c_n}{|\tilde{\theta}_n|_d} 1_{\{1 \leq j \leq d\}}$, $\tilde{\theta}_n = (\hat{\theta}_{j,n})_{1 \leq j \leq d}$, $|x|_d = \sqrt{\sum_{j=1}^d x_j^2}$ и

$$c_n = \frac{l_n^*}{(r_n^* + \sqrt{d/v_n})n}. \quad (14)$$

Положительный параметр r_n^* удовлетворяет следующим условиям

$$\lim_{n \rightarrow \infty} r_n^* = \infty \quad \text{и} \quad \forall \delta > 0 \quad \lim_{n \rightarrow \infty} \frac{r_n^*}{n^\delta} = 0. \quad (15)$$

Тогда класс улучшенных взвешенных оценок наименьших квадратов для функции $S(\cdot)$ из (1) примет следующий вид

$$S_\lambda^*(t) = \sum_{j=1}^n \lambda(j) \theta_{j,n}^* \phi_j(t). \quad (16)$$

Численное моделирование. В (1) выберем 1-периодическую функцию, определенную при $0 \leq t \leq 1$, как

$$S(t) = t \sin(2\pi t) + t^2(1-t) \cos(4\pi t). \quad (17)$$

Для осуществления симуляции модели (1) необходимо перейти к ее дискретизации (см., например, [4])

$$\Delta w_j = S(t_j)\Delta t + \rho_1 \Delta w_j + \rho_2 \Delta z_j, \quad j = \overline{1, np}, \quad (18)$$

$$\Delta w_j = \eta_j \sqrt{\Delta t} \quad \text{и} \quad \Delta z_j = \begin{cases} 0, & \text{если } \nexists k : t_{j-1} < T_k < t_j, \\ Y_k, & \text{иначе,} \end{cases} \quad (19)$$

где (η_j) , (Y_j) – взаимно независимые последовательности независимых случайных величин, имеющих стандартное нормальное распределение, $T_k = \sum_{l=1}^k \tau_l$, (τ_l) – последовательность независимых одинаково распределенных случайных величин, имеющих распределение хи-квадрат с k степенями свободы.

Для вычисления весовых коэффициентов положим $k^* = 100 + \sqrt{\ln n}$, $\varepsilon = \frac{1}{\ln n}$, $m = \lceil \ln^2 n \rceil$.

Оценку МНК будем строить по формуле (6). Также будем строить две улучшенные оценки МНК. Для первой оценки в формуле (14) в качестве l_n^* будем брать $l_n^* = (d-1)\rho_1^2$ из предложения 2, для второй оценки возьмем $l_n^{**} = \text{tr } G_n - \lambda_{\max}(G_n)$ согласно предложению 1. Тогда в соответствии с (16) получим улучшенные оценки вида

$$S_\lambda^*(t) = \sum_{j=1}^n \lambda(j) \theta_{j,n}^* \phi_j(t) \quad \text{и} \quad S_\lambda^{**}(t) = \sum_{j=1}^n \lambda(j) \theta_{j,n}^{**} \phi_j(t) \quad (20)$$

соответственно.

Для сравнения численных результатов используем эмпирический риск, который определяется по следующей формуле

$$R(\hat{S}, S) = \frac{1}{M} \sum_{k=1}^M \left(\frac{1}{np} \sum_{j=1}^{np} \left(\hat{S}_\lambda(t_j) - S(t_j) \right)^2 \right), \quad (21)$$

где $p = 1000$ и $M = 100$. Аналогичным образом определяются $R(S^*, S)$ и $R(S^{**}, S)$.

Обозначим $\Delta_1 = R(\hat{S}, S) - R(S^*, S)$, $\Delta_2 = R(\hat{S}, S) - R(S^{**}, S)$. Исследуем влияние мешающих параметров на качество оценки. Рассмотрим следующие случаи.

1. Изменение параметра n при $\rho_1 = \rho_2 = 0.5$, $\tilde{\tau} = k = 3$. Значения квадратичных рисков представлены в таблице 1. На рисунке 1 продемонстрированы графики Δ_1 , Δ_2 , где сплошная кривая соответствует функции Δ_1 , а пунктирная кривая – функции Δ_2 .

2. Изменение параметра $\tilde{\tau} = k$ при $\rho_1 = \rho_2 = 0.5$, $n = 100$. Значения квадратичных рисков представлены в таблице 2.

3. Изменение параметра ρ_1 при $\tilde{\tau} = k = 3$, $\rho_2 = 0.5$, $n = 1000$. Значения квадратичных рисков представлены в таблице 3. На рисунке 2 продемонстрированы графики Δ_1 , Δ_2 .

4. Изменение параметра ρ_2 при $\tilde{\tau} = k = 3$, $\rho_1 = 0.5$, $n = 1000$. Значения квадратичных рисков представлены в таблице 4. На рисунке 3 продемонстрированы графики Δ_1 , Δ_2 .

Таблица 1 – Квадратичные риски при изменении параметра n

n	$R(\hat{S}, S)$	$R(S^*, S)$	$R(S^{**}, S)$
20	0.06612133	0.06351045	0.06293257
50	0.03259358	0.03193463	0.03174086
100	0.0172123	0.01706185	0.01702013
200	0.01118756	0.01113515	0.01111913
500	0.005962003	0.005946709	0.005941997

Таблица 2 – Квадратичные риски при изменении параметра $\tilde{\tau} = k$

k	3	4	5	6
$R(\hat{S}, S)$	0.01721230	0.01644610	0.01664828	0.01612601
$R(S^*, S)$	0.01706185	0.01631981	0.01650714	0.01600129
$R(S^{**}, S)$	0.01702013	0.01629527	0.01648733	0.01598728

k	7	8	9	10
$R(\hat{S}, S)$	0.01563492	0.01595340	0.01542107	0.01592017
$R(S^*, S)$	0.01550909	0.01581666	0.01527805	0.01578663
$R(S^{**}, S)$	0.01549431	0.01580341	0.01526778	0.01577674

Таблица 3 – Квадратичные риски при изменении параметра ρ_1

ρ_1	0.5	1	1.5	2	2.5
$R(\hat{S}, S)$	0.003005	0.007384	0.014434	0.020801	0.031100
$R(S^*, S)$	0.003001	0.007360	0.014329	0.020597	0.030640
$R(S^{**}, S)$	0.003000	0.007358	0.014325	0.020594	0.030635

ρ_1	3	3.5	4	4.5	5
$R(\hat{S}, S)$	0.037657	0.050111	0.064565	0.068434	0.082685
$R(S^*, S)$	0.037080	0.049101	0.062929	0.067214	0.080897
$R(S^{**}, S)$	0.037075	0.049094	0.062921	0.067209	0.080892

На основе результатов, представленных в таблицах 1 - 4, а также на рисунках 1 - 3, можно сделать следующие выводы:

- При увеличении параметра n наблюдается монотонное уменьшение эмпирических рисков.
- При увеличении параметра $\tilde{\tau} = k$ наблюдается уменьшение эмпирических рисков, но отсутствует монотонность.

Таблица 4 – Квадратичные риски при изменении параметра ρ_2

ρ_2	0.5	1	1.5	2	2.5
$R(S, S)$	0.003005	0.003738	0.005140	0.007241	0.009970
$R(S^*, S)$	0.003001	0.003735	0.005136	0.007235	0.009962
$R(S^{**}, S)$	0.003000	0.003731	0.005125	0.007208	0.009904

ρ_2	3	3.5	4	4.5	5
$R(\tilde{S}, S)$	0.012472	0.016252	0.020559	0.024985	0.030338
$R(S^*, S)$	0.012464	0.016242	0.020546	0.024974	0.030325
$R(S^{**}, S)$	0.012381	0.016096	0.020307	0.024693	0.029908

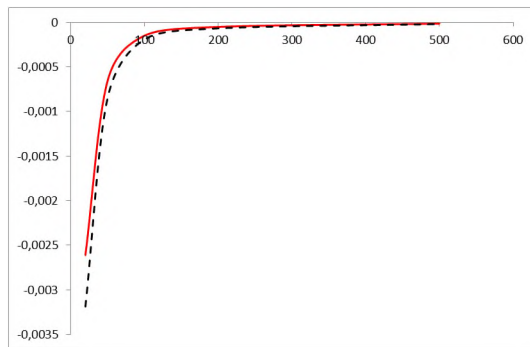


Рис. 1. Графики Δ_1, Δ_2 при изменении параметра n

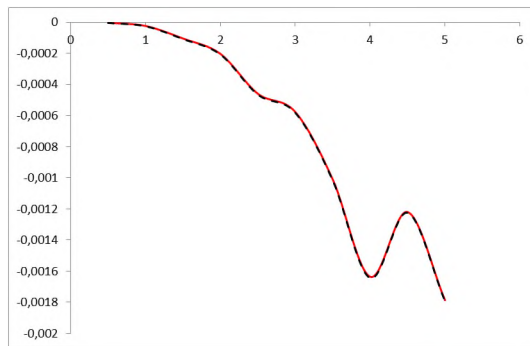


Рис. 2. Графики Δ_1, Δ_2 при изменении параметра ρ_1

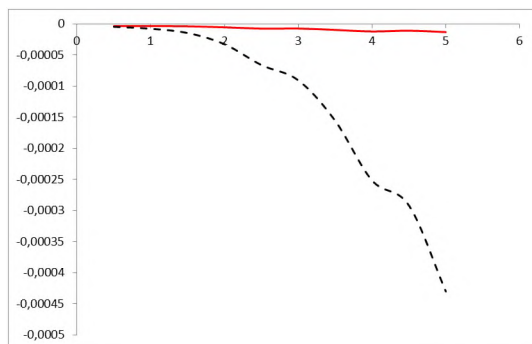


Рис. 3. Графики Δ_1 , Δ_2 при изменении параметра ρ_2

- При увеличении параметров ρ_1 и ρ_2 наблюдается монотонное увеличение эмпирических рисков.
- При увеличении параметра ρ_2 наблюдается увеличение разницы между улучшенными оценками.
- Во всех рассмотренных случаях оценка S^{**} является наиболее эффективной в смысле среднеквадратической точности. Это означает, что при построении улучшенной оценки рекомендуется использовать доступную априорную информацию о мешающих параметрах шумового процесса.

Список литературы

- [1] Barbu V.S., Beltaief S., Pergamenschikov S.M. Robust adaptive efficient estimation for semi-Markov nonparametric regression models // Statistical Inference for Stochastic Processes. 2019. Vol. in press. P. 1 - 45.
- [2] Pchelintsev E.A., Pergamenschikov S.M. Oracle inequalities for the stochastic differential equations // Statistical Inference for Stochastic Processes. 2018. Vol. 21, № 2. P. 469-483.
- [3] Pchelintsev E.A., Pchelintsev V.A., Pergamenschikov S.M. Improved robust model selection methods for a Lévy nonparametric regression in continuous time // Journal of Nonparametric Statistics. 2019. Vol. 31, № 3. P. 612–628.

- [4] Конев В.В., Пергаменщиков С.М., Пчелинцев Е.А. Оценивание регрессии при семимартингальных шумах. Улучшенные проекционные оценки / уч. пособие. - Томск: Томск. гос. ун-т, 2011. - 50 с.

Zhuman A. B., Pchelintsev E. A. (Tomsk State University, Tomsk, 2019) **Estimation of nonparametric regression with semi-Markov noises.**

Abstract. In this paper, we consider the problem of estimating an unknown function in continuous regression with semi-Markov noises. The results of numerical modeling of empirical risks of weighted least squares estimates and their improved modifications are presented, and the effect of interfering noise parameters on the quality of estimation procedures is studied.

Key words: nonparametric regression, weighted least squares estimates, semi-Markov process, mean-square accuracy.

Асимптотически эффективное оценивание коэффициента сноса в диффузионных процессах¹

Перелевский С. С., Пчелинцев Е. А.

Томский государственный университет, Томск
e-mail: slavaperelevskiy@mail.ru

Аннотация

В статье исследуются асимптотические свойства процедуры выбора модели для оценки неизвестного коэффициента сноса в диффузионных процессах. Показано, что процедура асимптотически эффективна, т.е. установлено, что асимптотический квадратичный риск процедуры совпадает с постоянной Пинскера, которая обеспечивает точную нижнюю границу квадратичного риска по всем возможным оценкам.

Ключевые слова: улучшенная оценка, диффузионный процесс, среднеквадратичная точность, оракульное неравенство, константа Пинскера, асимптотическая эффективность.

Введение. Рассматривается задача асимптотически эффективного оценивания неизвестного коэффициента сноса в диффузионном процессе, описываемом следующим стохастическим дифференциальным уравнением:

$$dy_t = S(y_t) dt + dw_t, \quad 0 \leq t \leq T, \quad (1)$$

где $(w_t)_{t \geq 0}$ – скалярный стандартный винеровский процесс, начальное значение y_0 – некоторая заданная константа, и $S(\cdot)$ – неизвестный коэффициент сноса. Такие модели широко используются на финансовых рынках, в радиофизике и т.д. [1]. Задача состоит в том, чтобы оценить функцию $S(x)$, $x \in [a, b]$, по наблюдениям $(y_t)_{0 \leq t \leq T}$. Основная цель этой статьи – доказать свойство асимптотической эффективности улучшенной процедуры выбора модели, предложенной в [2] для оценивания функции S в (1).

Понятие асимптотической эффективности связано с оптимальной скоростью сходимости минимаксного риска, т.е. важной проблемой в результатах оптимальности является исследование точной асимптотики минимаксного риска. Такие задача, например, изучались для модели гетероскедастичной регрессии Ефроймовичем [3] и Пинскером [4]. Чтобы доказать асимптотическую эффективность

¹Работа выполнена при финансовой поддержке РФФ, проект No 17-11-01049.

процедуры, необходимо показать, что ее асимптотический квадратичный риск совпадает с нижней границей, определяемой константой Пинскера [5, 6]. В данной статье задача решается с использованием подхода, основанного на методах выбора моделей и оракульных неравенств. Напомним, что метод выбора модели появился в новаторских работах [7] и [8]. Кроме того, Баррон, Бирже и Массар [9], Массар [10] и Кнайп [11] разработали этот метод для получения неасимптотических оракульных неравенств в моделях непараметрической регрессии с гауссовским шумом в дискретном времени. К сожалению, этот метод не может быть применен в нашем случае для доказательства свойства асимптотической эффективности, поскольку коэффициент в основном слагаемом в полученных неравенствах больше единицы. По этой причине в данной статье будем использовать метод, предложенный в [12].

Статья посвящена оценке неизвестной функции $S(x)$, $a \leq x \leq b$, в смысле среднеквадратического риска

$$\mathcal{R}(\hat{S}_T, S) = \mathbf{E}_S \|\hat{S}_T - S\|^2, \quad \|S\|^2 = \int_a^b S^2(x) dx, \quad (2)$$

где $\hat{S}_T$ – некоторая оценка S по наблюдениям $(y_t)_{0 \leq t \leq T}$, $a < b$ – некоторые действительные числа. Здесь $\mathbf{E}_S$ – математическое ожидание относительно распределения $\mathbf{P}_S$ случайного процесса $(y_t)_{0 \leq t \leq T}$. Чтобы получить надежную оценку функции S , необходимо, чтобы процесс (1) обладал свойством эргодичности. Для этого предположим, что неизвестная функция S принадлежит следующему функциональному классу:

$$\Sigma_{L,N} = \{S \in Lip_L(\mathbb{R}) : |S(N)| \leq L; \forall |x| \geq N, \exists \dot{S}(x) \in \mathbf{C}(\mathbb{R})$$

$$\text{такая, что } -L \leq \inf_{|x| \geq N} \dot{S}(x) \leq \sup_{|x| \geq N} \dot{S}(x) \leq -1/L\}, \quad (3)$$

где $L > 1$, $N > |a| + |b|$, $\dot{S}(x)$ – производная $S(x)$. Для оценки сноса S в (1) Гальчук и Пергаменщиков [13] предложили применить последовательный подход. Первый шаг – переход к модели регрессии с дискретным временем с использованием усеченной последовательной процедуры, представленной в [5]. С этой целью в любой точке x_k равномерного разбиения отрезка $[a, b]$, определим последовательную процедуру (τ_k, S_k^*) с правилом остановки τ_k и оценки S_k^* . Для $Y_k = S_k^* + 1 \leq k \leq n$, приходим к уравнению регрессии на некотором множестве $\Gamma \subseteq \Omega$ ($\sup_{S \in \Sigma_{L,N}} \mathbf{P}_S(\Gamma^c) \leq \Pi_T$, где $\lim_{T \rightarrow \infty} T^m \Pi_T = 0$ для любого $m > 0$):

$$Y_k = S(x_k) + \zeta_k. \quad (4)$$

Здесь, в отличие от классической регрессионной модели, последовательность шумов $(\zeta_k)_{1 \leq k \leq n}$ имеет сложную структуру, а имен-

но,

$$\zeta_k = \sigma_k \xi_k + \delta_k, \quad (5)$$

где $(\sigma_k)_{1 \leq k \leq n}$ – последовательность некоторых наблюдаемых случайных величин, $(\delta_k)_{1 \leq k \leq n}$ – последовательность ограниченных случайных величин и $(\xi_k)_{1 \leq k \leq n}$ – последовательность н.о.р. случайных величин из $\mathcal{N}(0, 1)$, которые не зависят от $(\sigma_k)_{1 \leq k \leq n}$.

Чтобы оценить функцию S в модели (4), мы используем метод выбора модели, основанный на улучшенных взвешенных оценках наименьших квадратов, предложенных в [14]. Усовершенствованный метод оценивания в непараметрических регрессионных моделях был разработан в [15, 16, 17].

Оракульное неравенство. Для того чтобы оценить неизвестную функцию в модели (4), используем улучшенные взвешенные оценки наименьших квадратов, определенные в [2],

$$S_\lambda^*(x_l) = \sum_{j=1}^n \lambda(j) \theta_{j,n}^* \phi_j(x_l) \mathbf{1}_\Gamma, \quad 1 \leq l \leq n, \quad (6)$$

где $(\phi_j)_{j \leq 1}$ является системой ортонормированных функций, вектор весовых коэффициентов $\lambda = (\lambda_1, \dots, \lambda_n)$ принадлежит некоторому конечному множеству Λ из $[0, 1]^n$,

$$\theta_{j,n}^* = \left(1 - \frac{c(d)}{\|\hat{\theta}_n\|} \mathbf{1}_{\{1 \leq j \leq d\}} \right) \hat{\theta}_{j,n}, \quad \|\tilde{\theta}_n\|^2 = \sum_{j=1}^d \hat{\theta}_{j,n}^2,$$

$$\hat{\theta}_{j,n} = \frac{b-a}{n} \sum_{l=1}^n Y_l \phi_j(x_l).$$

Здесь коэффициент $d \approx n^\epsilon / \ln n$, $0 < \epsilon < 1$, $c(d) \approx d/n$. Теперь определим оценку для S in (1). Положим для любого $a \leq x \leq b$

$$S_\lambda^*(x) = S_\lambda^*(x_1) \mathbf{1}_{\{a \leq x \leq x_1\}} + \sum_{l=2}^n S_\lambda^*(x_l) \mathbf{1}_{\{x_{l-1} < x \leq x_l\}}. \quad (7)$$

Чтобы выбрать из этого семейства наилучшую оценку, мы должны написать правило для определения вектора весовых коэффициентов $\lambda \in \Lambda$ в (7). Очевидно, что лучший способ – минимизировать эмпирическую квадратическую ошибку по λ :

$$\text{Err}_n(\lambda) = \|S_\lambda^* - S\|_n^2 \rightarrow \min.$$

Используя (7) и преобразование Фурье для S , получаем

$$\text{Err}_n(\lambda) = \sum_{j=1}^n \lambda^2(j) \theta_{j,n}^{*2} - 2 \sum_{j=1}^n \lambda(j) \theta_{j,n}^* \theta_{j,n} + \sum_{j=1}^n \theta_{j,n}^2.$$

Поскольку коэффициент $\theta_{j,n}$ неизвестен, нам нужно заменить величины $\theta_{j,n}^*$ некоторыми оценками, которые выбираем как

$$\tilde{\vartheta}_{j,n} = \hat{\theta}_{j,n} \theta_{j,n}^* - \frac{b-a}{n} s_{j,n} \quad \text{с} \quad s_{j,n} = \frac{b-a}{n} \sum_{l=1}^n \sigma_l^2 \phi_j^2(x_l).$$

За эту замену в эмпирической квадратичной ошибке нужно заплатить "штраф". Определяем платежную функцию

$$J_n(\lambda) = \sum_{j=1}^n \lambda^2(j) \theta_{j,n}^{*2} - 2 \sum_{j=1}^n \lambda(j) \tilde{\vartheta}_{j,n} + \rho P_n(\lambda),$$

где пенализационное слагаемое определяется как

$$P_n(\lambda) = \frac{b-a}{n} \sum_{j=1}^n \lambda^2(j) s_{j,n}$$

и $0 < \rho < 1$ – некоторая положительная постоянная, которая будет выбрана позже. Положим

$$\hat{\lambda} = \operatorname{argmin}_{\lambda \in \Lambda} J_n(\lambda)$$

и определим оценку S в виде (7):

$$S^*(x) = S_{\hat{\lambda}}^*(x) \quad \text{для} \quad a \leq x \leq b. \quad (8)$$

Теперь получим неасимптотическую верхнюю оценку для квадратического риска (8).

Теорема 1. Пусть $\Lambda \subset [0, 1]^n$ любое конечное множество, такое, что первые $d \leq n$ компонент весового вектора λ равны 1. Тогда для любого $n \geq 3$ и $0 < \rho < 1/6$, оценка (8) удовлетворяет следующему оракульному неравенству

$$\mathbf{E}_S \|S^* - S\|_n^2 \leq \frac{1+6\rho}{1-6\rho} \min_{\lambda \in \Lambda} \mathbf{E}_S \|\hat{S}_\lambda - S\|_n^2 + \frac{\Psi_n(\rho)}{n},$$

где $\lim_{n \rightarrow \infty} \Psi_n(\rho)/n = 0$.

Теперь рассмотрим задачу оценивания в модели (1) через модель (4). Применим процедуру оценивания (8) со специальным множеством весов, введенным в [5] к схеме регрессии (4). Обозначая $S_\alpha^* = S_{\lambda_\alpha}^*$, запишем

$$S^* = S_{\hat{\alpha}}^* \quad \text{с} \quad \hat{\alpha} = \operatorname{argmin}_{\alpha \in A_\varepsilon} J_n(\lambda_\alpha).$$

Теорема 2. Предположим, что $S \in \Sigma_{L,N}$ и количество точек $n = n(T) \rightarrow \infty$ при $T \rightarrow \infty$ в модели (4). Тогда процедура S^* удовлетворяет для любого $T \geq 32$, следующему неравенству

$$\mathcal{R}(S^*, S) \leq \frac{(1+\rho)^2(1+6\rho)}{1-6\rho} \min_{\alpha \in A_\varepsilon} \mathcal{R}(S_\alpha^*, S) + \frac{\mathcal{B}_T(\rho)}{n},$$

где $\lim_{T \rightarrow \infty} \mathcal{B}_T(\rho)/n(T) = 0$.

Асимптотическая эффективность. Для изучения асимптотической эффективности определим следующий функциональный шар

Соболева.

$$W_{k,r}^0 = \{f \in \mathbf{C}_0^k([a, b]) : \sum_{j=0}^k \|f^{(j)}\|^2 \leq r\}, \quad (9)$$

где $r > 0$ и $k \geq 1$ – некоторые неизвестные параметры, $\mathbf{C}_0^k([a, b])$ – пространство k раз дифференцируемых функций $f : \mathbb{R} \rightarrow \mathbb{R}$ таких, что

$$f^{(i)}(x) = 0, \quad \text{для } 0 \leq i \leq k-1 \text{ и } x \notin [a, b].$$

Будем называть такие функции *периодическими на отрезке* $[a, b]$. Пусть S_0 будет фиксированной $k+1$ раз непрерывно-дифференцируемой функцией из $\Sigma_{L,N}$. Положим

$$\Theta_{k,r} = \{S = S_0 + f, f \in W_{k,r}^0\}. \quad (10)$$

Чтобы сформулировать асимптотические результаты, определим следующий нормирующий коэффициент

$$\gamma(S) = ((1+2k)r)^{1/(2k+1)} \left(\frac{J(S)k}{\pi(k+1)} \right)^{2k/(2k+1)} \quad (11)$$

где

$$J(S) = \int_a^b \frac{1}{q_S(x)} dx, \quad q_S(x) = \frac{\exp\{2 \int_0^x S(z) dz\}}{\int_{-\infty}^{+\infty} \exp\{2 \int_0^y S(z) dz\} dy}.$$

Хорошо известно, что для любого $S \in \Theta_{k,r}$ оптимальная скорость сходимости $T^{-2k/(2k+1)}$ (см., например, [18]). Имеем следующую асимптотическую верхнюю оценку для квадратического риска.

Теорема 3. *Квадратический риск (2) для процедуры оценивания S^* имеет следующую асимптотическую верхнюю границу*

$$\limsup_{T \rightarrow \infty} T^{2k/(2k+1)} \sup_{S \in \Theta_{k,r}} \frac{\mathcal{R}(S^*, S)}{\gamma(S)} \leq 1. \quad (12)$$

Эта верхняя граница точна в следующем смысле.

Теорема 4. *Для любой оценки $\hat{S}$ из S измеримой относительно $\mathcal{F}_T^y$,*

$$\liminf_{T \rightarrow \infty} \inf_{\hat{S}} T^{2k/(2k+1)} \sup_{S \in \Theta_{k,r}} \frac{\mathcal{R}(\hat{S}, S)}{\gamma(S)} \geq 1, \quad (13)$$

где $\mathcal{F}_T^y$ – σ -алгебра, порожденная наблюдениями $(y_t)_{0 \leq t \leq T}$.

Замечание 1. Неравенства (14) и (13) означают, что функция (11) – константа Пинскера в данном случае [4].

Следствие 1. *Из теорем 2 и 3 следует, что процедура выбора модели S^* , определенная в (8), является асимптотически эффективной, т.е.*

$$\lim_{T \rightarrow \infty} T^{2k/(2k+1)} \sup_{S \in \Theta_{k,r}} \frac{\mathcal{R}(S^*, S)}{\gamma(S)} = 1. \quad (14)$$

Численное моделирование. Предположим, что в модели (1)

$$S(x) = x^2 \sin(2\pi x) + x^2(1-x) \cos(4\pi x).$$

Для весовых коэффициентов выбираем $n = T$,

$$k^* = 100 + \sqrt{\ln n}, \quad \varepsilon = \frac{1}{\ln n}, \quad m = \ln^2 n, \quad \omega_\alpha = 100 + (A_\beta t n)^{\frac{1}{2\beta+1}}.$$

Эмпирический риск:

$$\mathcal{R}(S^*, S) = \frac{1}{1000} \sum_{m=1}^{1000} \|S_m^* - S\|_n^2.$$

В таблице 1 приведены результаты поведения эмпирических сред-неквадратических рисков для предложенной процедуры оценива-ния (8). Из таблицы 1 видно, что с увеличением количества на-

Таблица 1. Эмпирические квадратические асимптотические риски

n	501	1001	2001	10001
$T^{2k/(2k+1)} \frac{\mathcal{R}(S^*, S)}{\gamma(S)}$	4.7257	2.0856	1.0072	0.9012

блюдений n нормализованные эмпирические среднеквадратические риски стремятся к единице, что численно подтверждает следствие 1.

Список литературы

- [1] Kutoyants Yu. (2004). *Statistical inference for ergodic diffusion processes*. Springer-Verlag, London.
- [2] Pchelintsev E.A., Perelevskiy S.S., Makarova I.A. (2018). Improved nonparametric estimation of the drift in diffusion processes. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*. Vol. **160**, No **2**, pp. 364–372.
- [3] Efroimovich S. (2007). Sequential design and estimation in heteroscedastic nonparametric regression. *Sequential Analysis*. No **26**, pp. 3–25. Springer, New York, 1998.
- [4] Pinsker M.S. (1981) Optimal filtration of square integrable signals in gaussian white noise. *Problems Transimis. information*. No **17**, pp. 120–133.
- [5] Galtchouk L., Pergamenschikov S. (2006) Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes. *Statistical Inference for Stochastic Process*. Vol. **9**, No. **1**, pp. 1–16.

- [6] Ibragimov I.A., Khasminskii R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer, New York.
- [7] Akaike H. (1974). A new look at the statistical model identification. *IEEE Trans. on Automatic Control*. Vol. **19**, No. **7**, pp. 716–723.
- [8] Mallows C. (1973). Some comments on C_p . *Technometrics*. Vol. **15**, pp. 661–675.
- [9] Barron A., Birgé L., Massart P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Relat. Fields*. Vol. **113**, pp. 301–415.
- [10] Massart P. (2005). *A non-asymptotic theory for model selection*. European Congress of Mathematics. Eur. Math. Soc., Zurich.
- [11] Kneip A. (1994). Ordered linear smoothers. *Annals of Statistics*. Vol. **22**, pp. 835–866.
- [12] Galtchouk L.I., Pergamenshchikov S.M. (2009). Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression. *J. Korean Stat. Soc.* Vol. **38**, No **4**, pp. 305–322.
- [13] Galtchouk L.I., Pergamenshchikov S.M. (2011). Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. *Journal of Nonparametric Statistics*. Vol. **23**, No **2**, pp. 255–285.
- [14] Pchelintsev E., Perelevskiy S. (2019). Adaptive estimation in heteroscedastic nonparametric regression. *Tomsk State University Journal of Mathematics and Mechanics*. No. **57**, pp. 38–52.
- [15] Pchelintsev E.A., Pergamenshchikov S.M. (2018). Oracle inequalities for the stochastic differential equations. *Statistical Inference for Stochastic Processes*. Vol. **21**, No **2**, pp. 469–483.
- [16] Pchelintsev E., Pchelintsev V., Pergamenshchikov S. (2018). 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*. Vol. **20**, No **1**, pp. 73–77.
- [17] Pchelintsev E. A., Pergamenshchikov S. M. (2019). Improved model selection Method for an adaptive estimation in semimartingale regression models. *Tomsk State University Journal of Mathematics and Mechanics*. Vol. **58**, pp. 14–31.

- [18] Galtchouk L.I., Pergamenshchikov S.M. (2004). Nonparametric sequential estimation of the drift in diffusion via model selection. *Mathematical Methods of Statistics*. Vol. **13**, pp. 25–49.

Perelevskiy S., Pchelintsev E. (Tomsk State University, Tomsk, 2019) **Asymptotically efficient estimation of a drift coefficient in diffusion processes**

Abstract. The article studies the asymptotic properties of the model selection procedure for estimation an unknown drift coefficient in diffusion processes. It is shown that the procedure is asymptotically efficient, i.e. it is established that the asymptotic quadratic risk of the procedure coincides with the Pinsker constant, which provides an exact lower bound of the quadratic risk for all possible estimates.

Key words: Improved estimation, stochastic diffusion process, mean-square accuracy, oracle inequalities, Pinsker constant, asymptotic efficiency.

Построение и анализ регрессионной модели, описывающей зависимость объемного коэффициента нефти от газосодержания и плотности газа¹

Степаненко А. С., Пчелинцев Е. А., Бордзиловский А. С.

Томский государственный университет, Томск,
ООО "СИАМ Мастер" Инжиниринговый центр, Томск
e-mail: d107252@rambler.ru
ABordzilovskiy@integra.ru

Аннотация

В работе развит подход для прогнозирования объемного коэффициента нефти, который является очень важным параметром нефтегазовой отрасли. Для анализа использованы данные с 5 различных месторождений общим объемом 573 наблюдений. Для описания зависимости объемного коэффициента нефти от газосодержания и плотности газа предложена нелинейная регрессионная модель. Установлено статистически значимое качество построенной модели.

Ключевые слова: нелинейная регрессия, объемный коэффициент нефти, теорема Гаусса-Маркова, оценки МНК, гетероскедастичность.

Введение. Проблема количественного описания физико-химических свойств нефтей и газов является очень актуальной при планировании и анализе разработки разработки нефтегазовых месторождений. От достоверности оценки параметров пластовых флюидов зависит не только корректность оценки начальных запасов нефтяных залежей, но и принятие решения о наиболее эффективном способе их разработки. Количественное описание свойств подземных флюидов является важнейшей составляющей модели трехмерной подземной фильтрации, описывающей процессы в углеводородных залежах. Однако, ввиду трудностей связанных с отбором представительной (качественной) пробы пластового флюида и трудоемкости полномасштабных лабораторных исследований, зачастую получается что на начальном этапе планирования разработки отсутствует детальная лабораторная информация о пластовых флюидах [5]. В данном случае инженерам, планирующим разработку залежей, приходится делать очень серьезные предположения о

¹Статья написана в рамках научного проекта (№ 8.1.18.2018), выполненного при поддержке Программы повышения конкурентоспособности ТГУ.

свойствах флюидов, например предполагать их аналогичными соседним месторождениям или пластам, что негативно сказывается на корректности принятия решений. Одним из перспективных направлений решения этой проблемы является предсказание ключевых свойств пластовых флюидов без отбора и полномасштабного исследования представительной глубинной пробы. Такое прогнозирование может выполняться на основе легко замеряемых на поверхности параметров, таких как: плотности нефтяной и газовой фазы в стандартных условиях, а также газового фактора. В данной статье предложен метод прогнозирования свойств подземных флюидов на основе нелинейной регрессионной модели.

В общем виде множественная линейная регрессия описывается уравнением вида:

$$y_i = f(X) + \varepsilon_i,$$

где ε_i – случайные ошибки модели, а $f(X)$ – некоторая неизвестная функция, аппроксимирующая описываемое нами явление, зависящее от матрицы случайных факторов X .

Важным аспектом регрессионного анализа является спецификация и анализ качества регрессионной модели. Сформулируем данные требования для модели:

1. $y_i = f(X) + \varepsilon_i, i = 1, 2, \dots, n$.
2. $E(\varepsilon_i) = 0, i = 1, 2, \dots, n$.
3. $E(\varepsilon_i \varepsilon_j) = \begin{cases} \sigma^2, & i = j, \\ 0, & i \neq j. \end{cases}$
4. $(x^{(1)}, x^{(2)}, \dots, x^{(p)})$ – неслучайные переменные, а ранг матрицы $X = 1 + p < n$, где X матрица наблюдаемых значений объясняющей переменной:

$$X = \begin{pmatrix} 1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(p)} \\ 1 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(p)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(p)} \end{pmatrix}.$$

Из вышеперечисленных требований вытекает, что функция $f(X)$ имеет вид:

$$f(X) = E(y|X) = \theta_0 + \theta_1 x^{(1)} + \theta_2 x^{(2)} + \dots + \theta_p x^{(p)},$$

где $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ не являются случайными величинами, следовательно вся "случайность" описываемой нами величины y_i зависит только от остатков ε_i . В регрессионной модели требуется некоррелируемость остатков, что и было описано в 3 пункте. Так же в третьем пункте требуется однородность остатков [4].

Нелинейные модели регрессии и линеаризация. Часто на практике встречается ситуация при которой не всегда получается

выявить только линейные зависимости, так же встречаются нелинейные зависимости.

На первом же этапе построения линейной регрессии требуется попытаться подобрать такие преобразования к анализируемым переменным $y, x^{(1)}, \dots, x^{(p)}$, которые позволили бы представить искомую зависимость в виде линейного соотношения между преобразованными переменными, другими словами, если $\varphi_0, \varphi_1, \dots, \varphi_p$ - те самые искомые функции, которые определяют переход к преобразованным переменным, то есть $\tilde{y} = \phi_0(y)$, $\tilde{x}^{(1)} = \phi(x^{(1)}, \dots, \tilde{x}^{(p)}) = \phi(x^{(p)})$, то связь между y и $X = (x^{(1)}, \dots, x^{(p)})$ может быть представлена в виде линейной функции регрессии $\tilde{y}$ по $\tilde{X}$, а именно:

$$\tilde{y} = \theta_0 + \theta_1 \tilde{x}^{(1)} + \dots + \theta_p \tilde{x}^{(p)} + \varepsilon_i, i = 1, 2, \dots, n;$$

эту часть исследования обычно называют процедурой линеаризации модели [1].

В случае невозможности линеаризации модели приходится использовать искомую регрессионную зависимость в терминах исходных переменных, а именно:

$$y_i = f(X_i; \Theta) + \varepsilon_i;$$

если спецификация регрессионных остатков ε_i соответствует условиям классической модели, то для вычисления МНК-оценок $\hat{\Theta}_{\text{МНК}}$ векторного параметра Θ решается оптимизационная задача вида

$$\hat{\Theta}_{\text{МНК}} = \arg \min_{\Theta} \sum_{i=1}^n (y_i - f(X_i; \Theta))^2.$$

Для решения данной задач используются алгоритмы нелинейного программирования [2].

Для проверки адекватности построенной регрессионной модели и дальнейшего имитационного моделирования "белого шума" требуется выполнение условий теоремы Гаусса-Маркова для ошибок модели [1].

Теорема 1. *Оценки по методу наименьших квадратов являются лучшими в классе линейных несмещённых оценок, если:*

1. *Модель правильно специфицирована,*
2. *$\text{rang}(X) = k$,*
3. *$E(\varepsilon_i) = 0$,*
4. *$\text{Cov}(\varepsilon) = \sigma^2 I$.*

Результаты моделирования. При моделировании была проведена первичная статистическая обработка данных, взятых с открытого источника SPE [6]. В связи с тем, что данная выборка собрана из 5 различных месторождений, была проведена проверка на однородность исследуемых данных. Так же, в ходе исследования было

выявлено, что рассматриваемые выборки отличаются лишь сдвигом, поэтому для дальнейшего анализа данные были стандартизованы.

Была построена матрица корреляций между всеми рассматриваемыми параметрами. Из таблицы ниже видно, что интересующий нас объемный коэффициент нефти (y_i) имеет достаточно сильную зависимость от параметров "Газосодержание" ($x_i^{(1)}$) и "Плотность газа" ($x_i^{(2)}$), так же в таблице использованы обозначения для "Плотность нефти" ($x_i^{(3)}$) и "Температура" ($x_i^{(4)}$).

Таблица 1. Матрица корреляций

	$x_i^{(1)}$	$x_i^{(2)}$	$x_i^{(3)}$	$x_i^{(4)}$	y_i
$x_i^{(1)}$	1				
$x_i^{(2)}$	-0,183	1			
$x_i^{(3)}$	-0,253	-0,068	1		
$x_i^{(4)}$	0,115	0,405	-0,287	1	
y_i	0,841	-0,433	-0,103	-0,027	1

Так же наблюдается статистическая зависимость между параметром температура и плотность газа, по модулю совпадающая с зависимостью между объемным коэффициентом нефти и плотностью газа, в связи с этим наблюдается явление мультиколлинеарности. Напомним, что мультиколлинеарность - явление множественной коррелируемой связи между объясняющими переменными.

Оценка параметров проводилась с помощью метода наименьших квадратов. Были найдены оптимальные значения параметров и построена следующая модель:

$$y_i = -0.095 + 0.823x_i^{(1)} + \frac{0.056}{1.052 + x_i^{(2)}} + \varepsilon_i, \quad (1)$$

где $x_i^{(1)}$ - газосодержание, $x_i^{(2)}$ - плотность газа, ε_i - остатки модели.

Значение суммы квадратов остатков составило 0, а коэффициент детерминации R^2 0.739, что свидетельствует о достаточно хорошей аппроксимации. Так же использовался критерий Фишера, по его результатам сделан вывод о значимости коэффициента детерминации(уровень значимости 95

На представленном ниже графике видно, что построенная нами модель достаточно хорошо описывает интересующий нас параметр.

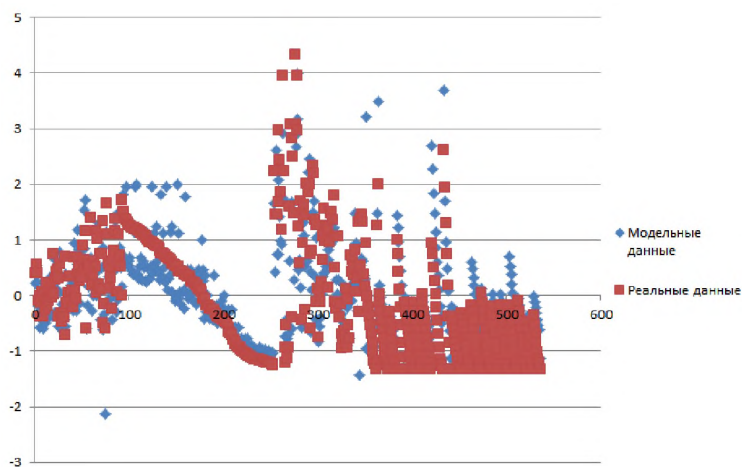


Рис. 1. Реальные и модельные данные

Так же в работе проверялись условия теоремы Гаусса-Маркова. Для проверки нормальности рассмотрим гистограмму ошибок.

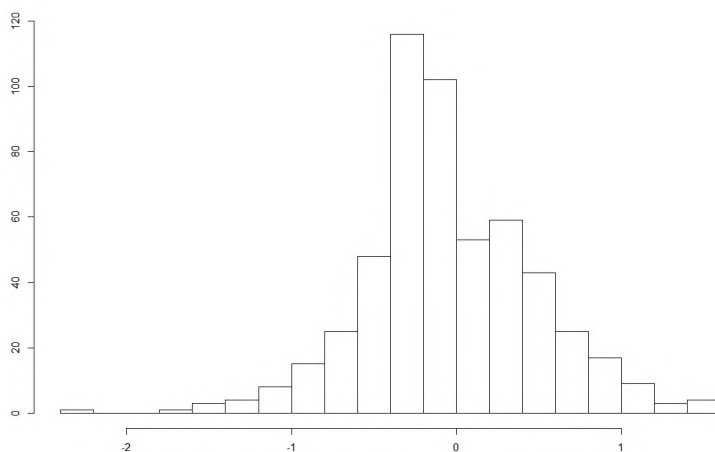


Рис. 2. Гистограмма частот ошибок модели

По рисункам выше можно сделать вывод о том, что ошибки модели имеют нормальное распределение. Значение эмперического среднего составило 0. Методом скользящего среднего рассчитаны выборочные дисперсии, значение не постоянно, на основе этого можно сделать вывод о том, что ошибки не являются однородными. Таким образом, полученная нами модель является двухфакторной нелинейной гетероскедастичной регрессией.

Список литературы

- [1] Айвазян С.А. Эконометрика - 2: продвинутый курс с приложениями в финансах: Учеб. / С.А.Айвазян, Д. Фантаццини; - М.: Магистр: НИЦ ИНФРА-М, 2014. - 944 с.
- [2] Измаилов, А.Ф. Численные методы оптимизации / А.Ф. Измаилов, М.В. Солодов. - М.: Физматлит, 2008. - 320 с.
- [3] Пчелинцев Е.А., Филимонова Ю.О. Исследование качества моделей временных рядов с применением бутстрап-методов // Международная научная конференция "Робастная статистика и финансовая математика - 2017" С.81-86.
- [4] Ширяев А.Н. Основы стохастической финансовой математики. Том 1. Факты. Модели. – М. : ФАЗИС, 1998. – 512 с.
- [5] Pedersen K. S., Christensen P.L. Phase Behavior of Petroleum Reservoir Fluids, CRC Press, 2007. - 407 p.
- [6] Society of Petroleum Engineering, URL: <https://www.spe.org/> (Дата обращения: 17.02.2019).

Stepanenko A. S., Pchelintsev E. A., Bordzilovskiy A. S. (Tomsk State University, ООО "SIAM Master" Engineering Centre, Tomsk, 2019) **Construction and analysis of regression models for description of the dependence of the formation volume factor of solution Gas-Oil ratio and gas density.**

Abstract. This paper presents an approach for prediction oil formation volume factor (FVF) for oil fields. This parameter is very important for field development planning. The approach is developed from a total of 573 experimentally obtained FVF from 5 different oil fields. Non-linear regression model is proposed for FVF prediction and strong dependence of solution Gas-Oil ratio and gas density is identified. Statistical quality of developed non-linear regression model is proved.

Key words: non-linear regression, formation volume factor, Gauss-Markov theorem, least squares estimates, heteroscedasticity.

Научное издание

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

Томск, 04–06 июля 2019 г.

Сборник статей

Издание подготовлено в авторской редакции

Подписано к печати 15.10.2019 г. Формат 60 × 84 1/16.

Бумага для офисной техники. Гарнитура Times.

Усл. печ. л. 5,2.

Тираж 30 экз. Заказ № 4029.

Отпечатано на оборудовании

Издательского Дома

Томского государственного университета

634050, г. Томск, пр. Ленина, 36

Тел. 8+(382-2)–52-98-49

Сайт: <http://publish.tsu.ru>

E-mail: rio.tsu@mail.ru

ISBN 978-5-94621-849-8



9 785946 218498