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

Международная научная конференция «Робастная статистика и финансовая математика – 2018»

(09–11 июля 2018 г.)

Сборник статей

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

Томск
Издательский Дом Томского государственного университета
2018

Sequential model selection method for a nonparametric autoregression *

Arkoun O., Brua J.-Y., Pergamenshchikov S.

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

Abstract

In this paper for the first time the nonparametric autoregression estimation problem for the quadratic risks is considered. To this end we develop a new adaptive sequential model selection method based on the efficient sequential kernel estimators proposed by Arkoun and Pergamenshchikov (2016). Moreover, we develop a new analytical tool for general regression models to obtain the non asymptotic sharp oracle inequalities for both usual quadratic and robust quadratic risks. Then, we show that the constructed sequential model selection procedure is optimal in the sense of oracle inequalities.

Keywords: nonparametric estimation, nonparametric autoregression, non-asymptotic estimation, robust risk, model selection, sharp oracle inequalities.

One of the standard linear models in general theory of time series is the autoregressive model (see, for example, [1] and the references therein). Natural extensions for such models are nonparametric autoregressive models which are defined by

$$y_k = S(x_k)y_{k-1} + \xi_k \quad \text{and} \quad x_k = a + \frac{k(b-a)}{n}\,, \qquad (1)$$

where $S(\cdot) \in \mathbf{L}_2[a,b]$ is unknown function, $a < b$ are fixed known constants, $1 \le k \le n$, the initial value y_0 is a constant and the noise $(\xi_k)_{k\ge 1}$ is i.i.d. sequence of unobservable random variables with $\mathbf{E}\xi_1 = 0$ and $\mathbf{E}\xi_1^2 = 1$.

The problem is to estimate the function S on the basis of the observations $(y_k)_{1\le k\le n}$ under the condition that the noise distribution is unknown. The minimax estimation problem for the model (1) has been treated for the first time in [3] and [8] in the nonadaptive case, i.e. for the known regularity of the function S. Then, in [2] it is proposed to use

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

the sequential analysis method for the adaptive pointwise estimation problem in the case where the unknown Hölder regularity is less than one and it is also shown that for the model (1), the adaptive pointwise estimation is possible only in the sequential analysis framework.

Here we study sequential estimation methods for a smooth function S that belongs to a certain Lipschitzian ε-stability set $\Theta_{\epsilon,L}$ (see [4]) for the quadratic risk defined as

$$\mathcal{R}_p(\widehat{S}_n, S) \,=\, \mathbf{E}_{p,S}\|\widehat{S}_n - S\|^2\,, \quad \|S\|^2 = \int_a^b S^2(x)\mathrm{d}x\,, \tag{2}$$

where $\widehat{S}_n$ is an estimator of S based on observations $(y_k)_{1\le k\le n}$ and $\mathbf{E}_{p,S}$ is the expectation with respect to the distribution law $\mathbf{P}_{p,S}$ of the process $(y_k)_{1\le k\le n}$ given the coefficient S and the distribution density p of the random variables $(\xi_k)_{1\le k\le n}$. Moreover, taking into account that the distribution p is unknown, we use the robust nonparametric estimation approach proposed in [5]. To this end we set the robust risk as

$$\mathcal{R}^*(\widehat{S}_n, S) \,=\, \sup_{p\in\mathcal{P}} \mathcal{R}_p(\widehat{S}_n, S)\,, \tag{3}$$

where $\mathcal{P}$ is a family of distributions depending on a fixed parameter $\varsigma \ge 1$.

To give a pointwise estimation of S, we follow the approach proposed in [7], i.e. we pass to a discrete time regression model by making use of the truncated sequential procedure introduced in [4]. More precisely, at any point $(z_l)_{1\le l\le d}$ of a partition of the interval $[a,b]$, we define a sequential procedure (τ_l, S_l^*) with a stopping rule τ_l and an estimator S_l^* of $S(z_l)$.

Note that the performance of any estimator $\widehat{S}$ will be measured by the empirical squared error

$$\|\widehat{S} - S\|_d^2 = \frac{b-a}{d}\sum_{l=1}^{d}(\widehat{S}(z_l) - S(z_l))^2\,.$$

For $Y_l = S_l^*$ with $1 \le l \le d$, we come to the regression equation on some set $\Gamma \subseteq \Omega$:

$$Y_l = S(z_l) + \zeta_l\,, \quad 1 \le l \le d\,, \tag{4}$$

the noise sequence $(\zeta_l)_{1\le l\le d}$ having a complex structure, namely,

$$\zeta_l \,=\, \eta_l + \varpi_l\,,$$

where $(\eta_l)_{1\le l\le d}$ is a "main noise" sequence of uncorrelated random variables and $(\varpi_l)_{1\le l\le n}$ is a sequence of random variables for which

$$\mathbf{u}_d^* = \mathbf{E}_{p,S}\mathbf{1}_\Gamma\|\varpi\|_d^2 < \infty\,.$$

Moreover, for any $1 \le l \le d$ and $p \in \mathcal{P}$, the random variable η_l will satisfy some properties, in particular

$$\mathbf{E}_{p,S}\left(\eta_l\,|\mathcal{G}_l\right)=0\,\text{and}\quad \mathbf{E}_{p,S}\left(\eta_l^2\,|\mathcal{G}_l\right)=\sigma_l^2\,,$$

and there exist some constants $0<\sigma_{0,*}<\sigma_{1,*}$ such that

$$\sigma_{0,*}\le\min_{1\le l\le d}\sigma_l^2\le\max_{1\le l\le d}\sigma_l^2\le\sigma_{1,*}\,.$$

In addition, we show that the probability of Γ^c goes to zero faster than any power function of the number of observations n.

In order to estimate the function S in model (12) we make use of the estimator family $(\widehat{S}_\lambda\,,\lambda\in\Lambda)$, where $\widehat{S}_\lambda$ is a weighted least square estimator with the Pinsker weights.

We construct the set Λ as

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

where $\mathcal{A}$ is a numerical grid with cardinal ν. For each $\alpha=(\beta,\mathbf{l})\in\mathcal{A}$, we define the weight sequence

$$\lambda_\alpha=(\lambda_\alpha(j))_{1\le j\le n}$$

with the elements

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

where $j_*=1+[\ln n]$, $\omega_\alpha=(\mathrm{d}_\beta\,\mathbf{l}\,n)^{1/(2\beta+1)}$ and

$$\mathrm{d}_\beta=\frac{(\beta+1)(2\beta+1)}{\pi^{2\beta}\beta}\,.$$

For this family, similarly to [3], we construct a special selection rule, i.e. a random variable $\widehat{\lambda}$ with values in some set Λ, for which we define the selection estimator as $\widehat{S}_*=\widehat{S}_{\widehat{\lambda}}$.

First we obtain the sharp oracle inequality for this selection model procedure for the general regression model (12).

Theorem 1. *There exists some constant $\mathbf{l}^*>0$ such that for any weight vectors set Λ, any $p\in\mathcal{P}$, any $n\ge1$ and $0<\delta\le1/12$, the selection estimator $\widehat{S}_*$ satisfies the following oracle inequality*

$$\mathbf{E}_{p,S}\|\widehat{S}_*-S\|_d^2\le\frac{1+4\delta}{1-6\delta}\min_{\lambda\in\Lambda}\mathbf{E}_{p,S}\|\widehat{S}_\lambda-S\|_d^2$$

$$+\,\mathbf{l}^*\frac{\nu\varsigma^2}{\delta}\left(\frac{\sigma_{1,*}^2}{\sigma_{0,*}d}+\mathbf{u}_d^*+\delta^2\sqrt{\mathbf{P}_S(\Gamma^c)}\right)\,.$$

Then we obtain the oracle inequality for the quadratic risks (2).

Theorem 2. *There exists some constant $\mathbf{l}^*>0$ such that for any weight vectors set Λ, any continuously differentiable function S, any $p\in\mathcal{P}$, any $n\ge1$ and $0<\delta\le1/12$, the selection estimator $\widehat{S}_*$*

satisfies the following oracle inequality

$$\mathcal{R}_p(\widehat{S}_*, S) \leq \frac{(1+4\delta)(1+\delta)^2}{1-6\delta} \min_{\lambda\in\Lambda} \mathcal{R}_p(\widehat{S}_\lambda, S)$$

$$+ \mathbf{l}^* \frac{\varsigma^2 \nu}{\delta} \left(\frac{\|\dot{S}\|^2}{d^2} + \frac{\sigma_{1,*}^2}{\sigma_{0,*} d} + \mathbf{u}_d^* + \delta^2 \sqrt{\mathbf{P}_S(\Gamma^c)} \right) .$$

Furthermore, assuming that the cardinal $\nu = \nu(n)$ of Λ and the parameter $\varsigma = \varsigma(n)$ in the density family $\mathcal{P}$ are functions of the number observations n of the form $o(n^{\check{\delta}})$ for any $\check{\delta} > 0$, we obtain the oracle inequality for the estimation problem for the model (1).

Theorem 3. *For any $p \in \mathcal{P}$, $S \in \Theta_{\epsilon,L}$, $n \geq 3$ and $0 < \delta \leq 1/12$, the selection estimator $\widehat{S}_*$ satisfies the following oracle inequality*

$$\mathcal{R}_p(\widehat{S}_*, S) \leq \frac{(1+4\delta)(1+\delta)^2}{1-6\delta} \min_{\lambda\in\Lambda} \mathcal{R}_p(\widehat{S}_\lambda, S) + \frac{\check{\mathbf{B}}_n(p)}{\delta n} ,$$

where the term $\check{\mathbf{B}}_n(p)$ is such that for any $\check{\delta} > 0$

$$\lim_{n\to\infty} \frac{\check{\mathbf{B}}_n(p)}{n^{\check{\delta}}} = 0 .$$

Eventually, we obtain the same inequality for the robust risk (3).

References

1. Anderson, T.W. The Statistical Analysis of Time Series, New York, Wiley - 1994.

2. Arkoun, O. Sequential Adaptive Estimators in Nonparametric Autoregressive Models// Sequential Anal. 2011. V. 30. No. 2. P. 228-246.

3. Arkoun, O., Pergamenchtchikov, S. Nonparametric Estimation for an Autoregressive Model// Journal of Mathematics and Mechanics of Tomsk State University. 2008. V. 2. P. 20-30.

4. Arkoun, O., Pergamenchtchikov, S. Sequential Robust Estimation for Nonparametric Autoregressive Models.// Sequential Anal. 2016. V. 35. No. 4. P. 489-515.

5. Galtchouk, L., Pergamenshchikov, S. Asymptotically Efficient Estimates for Nonparametric Regression Models// Statist. Prob. Lett. 2006. V. 76. No. 8. P. 852-860.

6. Galtchouk, L.I., Pergamenshchikov, S.M. Sharp non-asymptotic oracle inequalities for nonparametric heteroscedastic regression models // J. Nonparametr. Statist. 2009. V. 21. No. 1. P. 1-18.

7. Galtchouk, L. and Pergamenshchikov, S. Adaptive Sequential Estimation for Ergodic Diffusion Processes in Quadratic Metric // J. Nonparametr. Statist. 2011. V. 23. No. 2. P. 255-285.

8. Moulines, E., Priouret, P., Roueff, F. On Recursive Estimation for Time Varying Autoregressive Processes// Ann. Statist. 2005. V. 33. No. 6. P. 2610-2654.

Аркун У., Бруа Ж.-И., Пергаменщиков С. М. (Руанский университет, Руан, Томский государственный университет, Томск, 2018) **Последовательный метод выбора модели для непараметрической авторегрессии**

Аннотация. В работе рассматривается задача непараметрического оценивания авторегрессии для квадратичных рисков. Разрабатывается новый адаптивный последовательный метод выбора модели, основанный на эффективных последовательных ядерных оценках, предложенных Аркун и Пергаменщиковым (2016). Кроме того, разрабатывается новый аналитический инструмент для общих моделей регрессии для получения несимптотических точных оракульных неравенств как для обычных квадратичных, так и для робастных квадратичных рисков. Устанавливается, что построенная процедура последовательного выбора модели оптимальна в смысле оракульных неравенств.

Ключевые слова: непараметрическое оценивание, непараметрическая авторегрессия, неасимптотическое оценивание, робастный риск, выбор модели, точные оракульные неравенства.