ОБРАБОТКА ИНФОРМАЦИИ

УДК 519.2
DOI: 10.17223/19988605/43/3

T.V. Dogadova, V.A. Vasiliev

ADAPTIVE PREDICTION OF NON-GAUSSIAN ORNSTEIN–UHLENBECK PROCESS

This study was supported by the Russian Science Foundation under grant No. 17-11-01049 and performed in National Research Tomsk State University.

This paper proposes adaptive predictors of non-Gaussian Ornstein–Uhlenbeck process with unknown parameters. Predictors are based on the truncated parameter estimators. Asymptotic and non-asymptotic properties of the predictors are investigated. In particular, there is found the rate of convergence of the second moment of a prediction error to its minimum value. In addition, there is established an asymptotic optimality of the adaptive predictors in the sense of a special risk function. The structure of the risk function assumes the optimization of both the duration of observations and the prediction quality.
Keywords: truncated parameter estimation; adaptive optimal prediction; non-Gaussian Ornstein–Uhlenbeck process; risk function.

Nowadays mathematical statistics along with economics, financial mathematics, engineering, biology and other fields of science that use mathematical tools for their benefits, are turned to development of predictive methods. Models allowing making predictions of high statistical quality are highly appreciated. Currently, one of the most popular continuous-time models that is extensively used in financial mathematics is a non-Gaussian Ornstein-Uhlenbeck process driven by the Lévy process. Usually in practice the applied models depend on unknown parameters. Estimation problem of the unknown parameters of dynamic systems is a relevant one since the estimators of the dynamic parameters are to be used in various adaptive problems including the problem of adaptive prediction. The quality of adaptive predictors significantly depends on a choice of estimators of the model parameters. One of the most proper methods to solve this problem is the truncated estimation method proposed in [1]. It gives an opportunity to obtain the estimators of guaranteed quality by samples of fixed size under low level of a’priory information on system parameters.

Adaptive prediction problem for discrete-time systems was solved in [2] on the basis of truncated estimators proposed in [1]. Later, the same problem for Gaussian Ornstein-Uhlenbeck process was solved in [3,4] on the basis of truncated estimators. In this paper we propose adaptive predictors of non-Gaussian Ornstein-Uhlenbeck process constructed on the basis of truncated estimators of dynamic parameters which are optimal in the sense of a special risk function. The risk function aims to optimize both the duration of observations and the predictive quality. The risk function of similar structure first appeared in [5] for the problems of parameter estimator’s optimization. Later on, this idea was developed in [6, 7] etc. for optimization of predictors of dynamic discrete-time systems.

1. Problem statement. Optimal prediction

Consider the following regression model:

$$dx_t = ax_t dt + d\xi_t,\ t \geq 0 \tag{1}$$

with zero mean initial condition x_0, having all the moments. Here $\xi_t = \rho_1 W_t + \rho_2 Z_t, \rho_1 \neq 0$ and ρ_2 are some constants, $(W_t, t \geq 0)$ is a standard Wiener process, given on a filtered probability space $(\Omega, F, \{F_t\}_{t\geq 0}, P)$, adapted to a filtration $\{F_t\}_{t\geq 0}$, $Z_t = \sum_{k=1}^{N_t} Y_k$ is a compound Poisson process, where $Y_k, k \geq 0$ are i.i.d. random zero mean variables having all the moments and (N_t) is a Poisson process with the intensity $\lambda > 0$, i.e.

$$N_t = \sum_{j\geq 1} \chi\{T_j \leq t\} \quad \text{and} \quad T_j = \sum_{l=1}^{j} \tau_j.$$

Here $(T_j)_{j\geq 1}$ are jumps of the Poisson process $(N_t)_{t\geq 0}$ and $(\tau_l)_{l\geq 1}$ are i.i.d. random variables that are exponentially distributed with the parameter λ.

It should be noted that for $\rho_2 = 0$ the process (1) is a standard Ornstein–Uhlenbeck process.

Suppose that the process (1) is stable, i.e. the parameter $a < 0$. Note that in this case for every $m \geq 1$

$$\sup_{t\geq 0} E x_t^{2m} < \infty.$$

The purpose is to construct a predictor for x_t by observations $x^{t-u} = (x_s)_{0\leq s\leq t-u}$ which is optimal in a sense of the risk function introduced below. Here $u > 0$ is a fixed time delay.

The solution of the process x_t, obtained by the Ito formula, has the form

$$x_t = e^{at} x_0 + \int_0^t e^{a(t-z)} d\xi_z, t \geq 0$$

and for given $u > 0$ we have the representation

$$x_t = b x_{t-u} + \eta_{t,t-u}, \quad t \geq u,$$

where

$$b = e^{au}, \ \eta_{t,t-u} = \int_{t-u}^{t} e^{a(t-s)} d\xi_s, \ E\eta_{t,t-u} = 0 \ \text{and} \ \sigma^2 := D\eta_{t,t-u} = \frac{1}{2a}\left(\rho_1^2 + \lambda \rho_2^2 E Y_1^2\right)[b^2 - 1].$$

Optimal in the mean square sense predictor is the conditional mathematical expectation

$$x_t^0 = b x_{t-u}, \ t > u.$$

2. Adaptive prediction. Model parameter estimators

As in practice the parameter a and, as follows, b are unknown, it is impossible to construct the optimal predictor for real processes. In order to solve the problem of prediction we define an adaptive predictor that is constructed by an estimator a_t of unknown parameter a.

Define adaptive predictor as

$$\hat{x}_t(t-u) = \hat{b}_{t-u} x_{t-u}, \ t > u, \tag{2}$$

where $\hat{b}_{t-u} = e^{\hat{a}_{t-u} u}, \ t > u; \ \hat{a}_t = proj_{(-\infty,0]} a_t, \ a_t$ is the truncated estimator of the parameter a constructed similar to discrete-time case [2] on the basis of the least squares estimator

$$a_t = \frac{\int_0^t x_v dx_v}{\int_0^t x_v^2 dv} \chi\left(\int_0^t x_v^2 dv \geq t \log^{-1} t\right). \tag{3}$$

3. Risk functions and prediction criteria

Denote the prediction errors of x_t^0 and $\hat{x}_t(t-u)$ as

$$e_t^0 := x_t - x_t^0 = \eta_{t,t-u}, \ e_t(t-u) := x_t - \hat{x}_t(t-u) = \left(b - \hat{b}_{t-u}\right) x_{t-u} + \eta_{t,t-u}, t \geq u.$$

Now we define the loss function

$$L_t = \frac{A}{t} e^2(t) + t,\ t \ge u,$$

where

$$e^2(t) = \frac{1}{t}\int_u^t e_s^2(s-u)ds$$

and the parameter A > 0 stands for the cost of prediction error. We also define the risk function $R_t = EL_t$ which has the following form

$$R_t = \frac{A}{t} Ee^2(t) + t$$

and consider optimization problem

$$R_t \to \min_t .$$

For the optimal predictors x_t^0 it is possible to optimize the corresponding risk function directly

$$R_t^0 = E\left(\frac{A}{t}(e^0(t))^2 + t\right) = \frac{A\sigma^2}{t} + t \to \min_t, \tag{4}$$

where

$$(e^0(t))^2 = \frac{1}{t}\int_u^t (e_s^0)^2 ds.$$

In this case the optimal duration of observations T_A^0 and the corresponding value of $R_{T_A^0}^0$ are respectively

$$T_A^0 = A^{1/2}\sigma,\ R_{T_A^0}^0 = 2A^{1/2}\sigma, \tag{5}$$

where $\sigma := \sqrt{\sigma^2}$. However, since *a* and as follows, σ are unknown and both T_A^0 and $R_{T_A^0}^0$ depend on *a*, the optimal predictor can not be used. Then we define the estimator T_A of the optimal time T_A^0 as

$$T_A = \inf\{\mathrm{t} \ge \mathrm{t_A} : t \ge A^{1/2}\sigma_{t_A}\}, \tag{6}$$

where $t_A := A^{1/2}\log^{-1} A = o(A^{1/2}).$ Here $\sigma_t := \sqrt{\sigma_t^2}$ is the estimator of unknown σ,

$$\sigma_t^2 = \frac{1}{t-u}\int_u^t \left(x_s - \hat{b}_t x_{s-u}\right)^2 ds.$$

The estimator is defined like that since $\sigma^2 = E\eta_{t,t-u}^2 = E(x_t - bx_{t-u})^2.$

4. Properties of parameter estimators and adaptive predictors

Estimators a_t, b_t and σ_t that are used in construction of adaptive predictors have the properties given in Lemma below which can be proven similar to [3]. Compare to [3] this way of estimation of the variance σ^2 does not require the knowledge of parameters the model parameters. In this particular case it is not dependant on the true values of parameters $\rho_1, \rho_2, EY_1^2\lambda$ and their estimators. Moreover, the upper boundary for the moments of deviation of σ_t^2 is more accurate than in [3].

In what follows, *C* will denote a generic non-negative constant whose value is not critical (and not always the same).

Lemma 1. Assume the model (1). Then for $t-u > s_0 := \exp(2|a|)$ and some numbers *C* estimators $\hat{a}_t$ and $\hat{b}_t$ have the properties:

$$E\left(\hat{a}_t-a\right)^{2p}\le\frac{C}{t^p} \tag{7}$$

and

$$E\left(\hat{b}_t-b\right)^{2p}\le\frac{C}{t^p},\ \ p\ge 1. \tag{8}$$

$$E\left(\sigma_t^2-\sigma^2\right)^{2p}\le\frac{C}{t^p},\ p\ge 1.$$

Proof. We prove the property (7) similar to [3]. By the definition (3) of the estimator a_t and using (1) let us find representation for the deviation of the estimator

$$a_t-a=\frac{\int_0^t x_v d\xi_v}{\int_0^t x_v^2 dv}\chi\left(\int_0^t x_v^2 dv\ge t\log^{-1}t\right)-a\cdot\chi\left(\int_0^t x_v^2 dv< t\log^{-1}t\right).$$

Define

$$g_t=\frac{1}{t}\int_0^t x_v^2 dv,\ \ g=-\frac{1}{2a}\left(\rho_1^2+\rho_2^2 EY_1^2\lambda\right)>0,\ \ f_t=\frac{1}{t}\int_0^t x_v d\xi_v.$$

Then

$$E(a_t-a)^{2p}=E\left[\frac{f_t}{g_t}\right]^{2p}\cdot\chi\left[g_t\ge\log^{-1}t\right]+a^{2p}\cdot P\left[g_t<\log^{-1}t\right]=:I_1+I_2. \tag{9}$$

Using the Cauchy-Schwarz-Bunyakovsky inequality for the first summand we get:

$$I_1=E\left[\frac{f_t}{g}+f_t\frac{g-g_t}{gg_t}\right]^{2p}\cdot\chi\left[g_t\ge\log^{-1}t\right]=\frac{Ef_t^{2p}}{g^{2p}}\cdot\left(1+\sum_{k=1}^{2p}C_{2p}^k\cdot\frac{(g-g_t)^k}{g_t^k}\right)\cdot\chi\left[g_t\ge\log^{-1}t\right]\le$$

$$\le\frac{C}{t^p}+C\cdot\log t\cdot Ef_t^{2p}\cdot\left|g-g_t\right|\le\frac{C}{t^p}+C\cdot\log t\cdot(Ef_t^{4p}\cdot E(g-g_t)^2)^{\frac{1}{2}}\le$$

$$\le\frac{C}{t^p}+C\cdot\log t\cdot\frac{1}{t^p}\cdot(E(g-g_t)^2)^{\frac{1}{2}}.$$

Now we estimate the moments $E\left(g-g_t\right)^{2m}$.

By the Ito formula for x_t^2 from [8] it is true that

$$x_t^2=x_0^2+2\int_0^t x_{s-}dx_s+\rho_1^2 t+\sum_{0\le s\le t}\left(\Delta x_s\right)^2=$$

$$=x_0^2+2a\int_0^t x_s^2 ds+\rho_1\int_0^t x_s dW_s+\rho_2\sum_{0\le s\le t}x_{s-}\Delta Z_s+\rho_1^2 t+\rho_2^2\sum_{0\le s\le t}\left(\Delta Z_s\right)^2.$$

Note that

$$\sum_{0\le s\le t}\left(\Delta Z_s\right)^2=\sum_{k=1}^{N_t}Y_k^2=\sum_{k=1}^{+\infty}Y_k^2\chi\left\{T_j\le t\right\},$$

where $\triangle x_s=x_s-x_{s-}$. Then by making use of the representation

$$g_t-g=\frac{1}{2a}\frac{x_t^2-x_0^2}{t}-\frac{\rho_1}{2at}\int_0^t x_s dW_s-\frac{\rho_2}{2at}\sum_{0\le s\le t}x_{s-}\Delta Z_s-\frac{\rho_2^2}{2a}\left[\frac{1}{t}\sum_{0\le s\le t}\left(\Delta Z_s\right)^2-\lambda EY_1^2\right] \tag{10}$$

and the strong law of large numbers one can show that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t x_s^2 ds=-\frac{1}{2a}\left(\rho_1^2+\rho_2^2 EY_1^2\lambda\right)\quad\text{a.s.}$$

By (10) for every $m \ge 1$ it holds

$$E\left(g_t - g\right)^{2m} \le \frac{C}{t^m}. \qquad (11)$$

Then

$$I_1 \le \frac{C}{t^p} + C\frac{\log t}{t^{p+1/2}} \le \frac{C}{t^p}, \;\; t > u \qquad (12)$$

and applying (11) , as well as the Chebyshev inequality for $t > s_0$ we have

$$I_2 \le a^{2p} \cdot P\left(\left|g_t - g\right| \ge g - \log^{-1} t\right) \le \frac{a^{2p}}{\left(g - \log^{-1} t\right)^{4p}} \cdot E\left(g_t - g\right)^{4p} \le \frac{C}{t^{2p}}. \qquad (13)$$

From (9), (12), (13) and definition of $\hat{a}_t$ the property (7) follows.

The assertion (8) follows from the obvious inequality $\left|\hat{b}_t - b\right| \le u \cdot e^{u|a|} \cdot \left|\hat{a}_t - a\right|$, which can be obtained by the Taylor expansion for the exponent $\exp\left(\left(\hat{a}_t - a\right)u\right)$.

Using the following representation of the estimator σ_t^2

$$\sigma_t^2 = \frac{1}{t-u}\int_u^t \left(x_s - \hat{b}_t x_{s-u}\right)^2 ds = \frac{1}{t-u}\left\{\int_u^t x_{s-u}^2 ds \cdot \left(b - \hat{b}_t\right)^2 + \int_u^t \eta_{s,s-u}^2 ds + 2\int_u^t x_{s-u} \cdot \eta_{s,s-u} ds \cdot \left(b - \hat{b}_t\right)\right\}$$

we get its deviation in a form

$$\sigma_t^2 - \sigma^2 = \frac{1}{t-u}\int_u^t \left(\eta_{s,s-u}^2 - \sigma^2\right) ds + \frac{1}{t-u}\int_u^t x_{s-u}^2 ds \cdot \left(\hat{b}_t - b\right)^2 + \frac{2}{t-u}\int_u^t x_{s-u} \cdot \eta_{s,s-u} ds \cdot \left(b - \hat{b}_t\right) := I_1 + I_2 + I_3.$$

Let us estimate the mathematical expectation of each summand separately. For the first one consider the case of $p = 1$

$$EI_1^2 = \frac{1}{(t-u)^2} E\left(\int_u^t \left(\eta_{s,s-u}^2 - \sigma^2\right) ds\right)^2 = \frac{1}{(t-u)^2}\int_u^t\int_u^t E\left(\eta_{s,s-u}^2 - \sigma^2\right)\left(\eta_{v,v-u}^2 - \sigma^2\right) \cdot \chi\left[|v-s| \le u\right] dv ds \le$$

$$\le \frac{1}{(t-u)^2}\int_u^t\int_u^t \chi\left[|v-s| \le u\right] dv ds = \frac{1}{(t-u)^2}\int_u^t\int_{s-u}^{s+u} dv ds = \frac{2u}{(t-u)^2}\int_u^t dv = \frac{2u}{(t-u)}.$$

Similarly, for an arbitrary number *p* we get the inequality

$$EI_1^{2p} \le \frac{C}{(\mathrm{t-u})^p}.$$

We apply the Cauchy–Schwarz–Bunyakovsky inequality and (6) to the third summand and since the process x_t is stable it holds

$$EI_2^{2p} \le \frac{1}{(t-u)^{2p}}\sqrt{E\left(\int_u^t x_{s-u}^2 ds\right)^{4p}} \cdot \sqrt{E \cdot \left(\hat{b}_t - b\right)^{4p}} \le \frac{C}{t^p}.$$

Using the independency of $\eta_{t,t-u}$ and x_{t-u}, the Cauchy–Schwarz–Bunyakovsky inequality and (8), we obtain

$$EI_3^{2p} \le \frac{4}{(t-u)^{2p}}\sqrt{E\left(\int_u^t x_{s-u} \cdot \eta_{s,s-u} ds\right)^{4p}} \cdot \sqrt{E\left(b - \hat{b}_t\right)^{2p}} \le \frac{C}{t^p}.$$

Lemma is proven.

Now we are ready to investigate the statistical properties of the adaptive predictor (3). The prediction error has the form

$$e_t(t-u) := x_t - \hat{x}_t(t-u) = \left(b - \hat{b}_{t-u}\right)x_{t-u} + \xi_{t,t-u}.$$

Thus, for some C

$$\overline{\lim_{t\to\infty}}\, t\cdot\left(Ee_t^2(t-u)-\sigma^2\right)\le C$$

and if there is a'priori information that $|a|\le L$ then

$$Ee_t^2(t-u)-\sigma^2\le\frac{C}{t},\quad t>u+\exp(2L).$$

Analogously to [4], our purpose is to prove the asymptotic equivalence of T_A and T_A^0 in the almost surely and mean senses and the optimality of the presented adaptive prediction procedure in the sense of equivalence of $R^0_{T_A^0}$ and the obviously modified risk

$$\bar{R}_A=A\cdot E\frac{1}{T_A}e^2(T_A)+ET_A. \qquad (14)$$

Theorem 1. Assume the model (1). Let the predictors $\hat{x}_t(t-u)$ be defined by (2), the times T_A^0, T_A and the risk functions $R^0_{T_A^0}$, $\bar{R}_A$ defined by (5), (6) and (14). Then for every $a < 0$

i) $\dfrac{T_A}{T_A^0}\xrightarrow[A\to\infty]{}1$ a.s.;

ii) $\dfrac{ET_A}{T_A^0}\xrightarrow[A\to\infty]{}1$;

iii) $\dfrac{\bar{R}_A}{R^0_{T_A^0}}\xrightarrow[A\to\infty]{}1$.

Proof of Theorem 1 in general is similar to one from [4].

Conclusion

Adaptive prediction problem of the non-Gaussian Ornstein-Uhlenbeck process is solved. Non-asymptotic upper bound of the second moment of the prediction errors is found. It is shown that this bound is inverse proportional to the duration of observations. Non-asymptotic properties of adaptive predictors are obtained due to the usage of the truncated estimators [1] of the unknown parameters constructed by samples of fixed size. This method can be applied to various problems of parametric and non-parametric statistics.

In this paper we propose adaptive optimal predictors of non-Gaussian Ornstein-Uhlenbeck process. Their optimality in the sense of a special risk function is shown. The used risk function makes it possible to optimize the duration of observations along with the prediction quality.

The authors are very thankful to S. Pergamenshchikov for the helpful comments and remarks.

REFERENCES

1. Vasiliev, V.A. (2014) A Truncated Estimation Method with Guaranteed Accuracy. *Annals of Institute of Statistical Mathematics.* 66. pp. 141–163. DOI: 10.1007/s10463-013-0409-x
2. Kusainov, M.I. & Vasiliev, V.A. (2015) On optimal adaptive prediction of multivariate autoregression. *Sequential Analysis: Design Methods and Applications.* 34. pp. 211–234. DOI: 10.1080/07474946.2015.1030977
3. Dogadova, T.V. & Vasiliev, V.A. (2017) Adaptive prediction of stochastic differential equations with unknown parameters. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 38. pp. 17–23. DOI: 10.17223/19988605/38/3
4. Dogadova, T.V. & Vasiliev, V.A. (2017) On adaptive optimal prediction of Ornstein-Uhlenbeck process. *Applied Mathematical Science*. 11. pp. 591–600. DOI: 10.12988/ams.2017.7247
5. Chernoff, H. (1972) *Sequential Analysis and Optimal Design*. SIAM, Philadelphia.
6. Sriram, T.N. (1988) Sequential estimation of the autoregressive parameter in a first order autoregressive process. *Sequential Analysis.* 7(1). pp. 53–74. DOI: 10.1080/07474948808836142
7. Sriram, T.N. & Ross, I. (2014) Sequential estimation for time series models. *Sequential Analysis*. 33(2). pp. 136–157.
8. Liptzer, R.S. & Shiryaev, A.N. (1989) *Theory of Martingales*. Dordrecht: Kluwer.

Received: January 19, 2018

Догадова Т.В., Васильев В.А. АДАПТИВНОЕ ПРОГНОЗИРОВАНИЕ НЕГАУССОВСКОГО ПРОЦЕССА ОРНШТЕЙНА–УЛЕНБЕКА. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика.* 2018. № 43. С. 26–32

DOI: 10.17223/19988605/43/3

В работе предлагаются адаптивные прогнозы негауссовского процесса Орнштейна–Уленбека с неизвестными параметрами. Прогнозы основаны на усеченных оценках параметров. Исследуются асимптотические и неасимптотические свойства прогнозов. В частности, найдена скорость сходимости второго момента ошибки прогнозирования к ее минимальному значению. Кроме того, установлена асимптотическая оптимальность адаптивных прогнозов в смысле особой функции риска. Структура функции риска предполагает оптимизацию как длительности наблюдений, так и качества прогнозирования.

Ключевые слова: усеченное оценивание параметров; адаптивное оптимальное прогнозирование; негауссовский процесс Орнштейна–Уленбека; функция риска.

VASILIEV Vyacheslav Arthurovich (Doktor of Physics and Mathematics, Professor of the Department of High Mathematics and Mathematical Modelling of National Research Tomsk State University, Russian Federation).
E-mail: vas@mail.tsu.ru

DOGADOVA Tatiana Valeryevna (PhD Student of the Department of High Mathematics and Mathematical Modelling of National Research Tomsk State University, Russian Federation).
E-mail: aurora1900@mail.ru

Dogadova T.V., Vasiliev V.A. (2018) ADAPTIVE PREDICTION OF NON-GAUSSIAN ORNSTEIN–UHLENBECK PROCESS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitelnaja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 43. pp. 26–32