

УДК 519.2

DOI: 10.17223/19988605/38/3

T.V. Dogadova, V.A. Vasiliev

ADAPTIVE PREDICTION OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH UNKNOWN PARAMETERS

This study (research grant N8.1.55.2015) was supported by The Tomsk State University Academic D.I. Mendeleev Fund Program in 2015-2016 and by RFBR Grant 16-01-00121 A.

This paper proposes adaptive predictors of continuous-time dynamic systems with unknown parameters. Predictors are based on the truncated parameter estimators. In particular, there are considered the Ornstein-Uhlenbeck process and one-parameter stochastic delay differential equation. In this paper the truncated estimation method is first applied to continuous-time systems. Asymptotic and non-asymptotic properties of the predictors are investigated. There is also found the rate of convergence of the second moment of a prediction error to its minimum value.

Keywords: truncated estimation; continuous-time dynamic systems; prediction; delay differential equations; Ornstein-Uhlenbeck process.

Prediction is a momentous problem in modern world. Distinctive fields of science, for instance, economics, financial mathematics, engineering, biology etc. appeal to mathematical tools to get predictors of the real dynamic processes they explore. If a model allows making predictions of high statistical quality it is considered to be beneficial. Since models of dynamic systems often have unknown parameters, we have to deal with estimation problem in order to construct adaptive predictors. The quality of adaptive prediction significantly depends on a choice of estimators of the model parameters. Adaptive prediction problem for discrete-time systems was solved in [1] on the basis of truncated estimators proposed in [2]. In this paper we present analytically investigated predictors of continuous-time systems constructed on the basis of truncated estimators of dynamic parameters.

1. Prediction of Ornstein-Uhlenbeck process

Assume the model

$$dx_t = ax_t dt + dw_t, \quad t \geq 0, \quad (1)$$

with an unknown parameter a , where x_0 is zero mean random variable having finite moments of all order, w_t is the standard Wiener process, x_0 and w_t are mutually independent. Suppose that the process (1) is stable, i.e. the parameter $a < 0$. Note that in this case for every $m \geq 1$

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

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.

Using the solution of the process x_t for given $u > 0$ we have the representation

$$x_t = \lambda x_{t-u} + \xi_{t,t-u}, \quad t \geq u,$$

where $\lambda = e^{au}$, $\xi_{t,t-u} = \int_{t-u}^t e^{a(t-s)} dw_s$, $E \xi_{t,t-u} = 0$ and $\sigma^2 := E \xi_{t,t-u}^2 = \frac{1}{2a} [e^{2au} - 1]$.

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

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

As in practice the parameter a and, as follows, λ 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 estimator a_t of unknown parameter a .

Define adaptive predictor as

$$\hat{x}_t(t-u) = \hat{\lambda}_{t-u} x_{t-u}, \quad t > u, \quad (3)$$

where $\hat{\lambda}_{t-u} = e^{\hat{a}_{t-u}u}$, $t > u$; $\hat{a}_t = \text{proj}_{(-\infty, 0]} a_t$, 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

$$a_t = \frac{\int_0^t x_v dx_v}{\int_0^t x_v^2 dv} \chi \left(\int_0^t x_v^2 dv \geq t \log^{-1} t \right). \quad (4)$$

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

$$E(\hat{a}_t - a)^{2p} \leq \frac{C_p}{t^p} \quad (5)$$

and

$$E(\hat{\lambda}_t - \lambda)^{2p} \leq \frac{C_p}{t^p}, \quad p \geq 1. \quad (6)$$

Proof of Lemma. We prove the property (5) similar to [2]. By the definition (4) of the estimator a_t and using (1) let us find representation for the deviation of the estimator

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

Define

$$g_t = \frac{1}{t} \int_0^t x_v^2 dv, \quad g = -\frac{1}{2a}, \quad f_t = \frac{1}{t} \int_0^t x_v dw_v.$$

Then

$$E(a_t - a)^{2p} = E \left[\frac{f_t}{g_t} \right]^{2p} \cdot \chi[g_t \geq \log^{-1} t] + a^{2p} \cdot P[g_t < \log^{-1} t] =: I_1 + I_2. \quad (7)$$

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

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

$$\begin{aligned} I_1 &= E \left[\frac{f_t}{g} + f_t \frac{g - g_t}{gg_t} \right]^{2p} \cdot \chi[g_t \geq \log^{-1} t] = \frac{E f_t^{2p}}{g^{2p}} \cdot \left(1 + \sum_{k=1}^{2p} C_{2p}^k \cdot \frac{(g - g_t)^k}{g_t^k} \right) \cdot \chi[g_t \geq \log^{-1} t] \leq \\ &\leq \frac{C}{t^p} + C \cdot \log t \cdot E f_t^{2p} \cdot |g - g_t| \leq \frac{C}{t^p} + C \cdot \log t \cdot (E f_t^{4p} \cdot E(g - g_t)^2)^{\frac{1}{2}} \leq \\ &\leq \frac{C}{t^p} + C \cdot \log t \cdot \frac{1}{t^p} \cdot (E(g - g_t)^2)^{\frac{1}{2}}. \end{aligned}$$

By the Ito formula for x_t^2 it is true that

$$dx_t^2 = 2ax_t^2 dt + 2x_t dw_t + dt$$

and then

$$g_t - g = \frac{1}{2a} \frac{x_t^2}{t} - \frac{1}{at} \int_0^t x_t dw_t.$$

Thus, by making use of the properties of the Ito integral and (2) for every $m \geq 1$ it holds

$$E(g_t - g)^{2m} \leq \frac{C}{t^m}. \quad (8)$$

Then

$$I_1 \leq \frac{C}{t^p} + C \frac{\log t}{t^{p+1/2}}, \quad t > u, \quad (9)$$

and applying (8), as well as the Chebyshev inequality for $t > s_0$ we have

$$I_2 \leq a^{2p} \cdot P(|g_t - g| \geq g - \log^{-1} t) \leq \frac{a^{2p}}{(g - \log^{-1} t)^{4p}} \cdot E(g_t - g)^{4p} \leq \frac{C}{t^{2p}}. \quad (10)$$

From (7), (9), (10) and definition of \hat{a}_t the property (5) follows.

The assertion (6) follows from the obvious inequality $|\hat{\lambda}_t - \lambda| \leq u|\hat{a}_t - a|$, which can be obtained by the Taylor expansion for the exponent $\exp((\hat{a}_t - a)u)$. Lemma is proven.

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

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

Thus, for some C

$$\lim_{t \rightarrow \infty} t \cdot (Ee_t^2(t-u) - \sigma^2) \leq C$$

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

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

2. Delay differential equation

Assume $w = (w_t, t \geq 0)$ is a real-valued standard Wiener process, a and b are real numbers and $x = (x_t, t \geq -r)$ is a solution of the stochastic delay differential equation

$$dx_t = ax_t dt + bx_{t-r} dt + dw_t, \quad t \geq 0, \quad (11)$$

with some fixed initial condition $x_t = X_0(t), t \in [-r, 0]$, where $X_0(\cdot)$ is a continuous stochastic process independent of $w(\cdot)$.

The solution x of (11) exists, it is pathwise uniquely determined and can be represented as (see, e.g., [3, 4]):

$$x_t = x_0(t)X_0(0) + b \int_{-r}^0 x_0(t-s-r)X_0(s)ds + \int_0^t x_0(t-s)dw_s, \quad t \geq 0.$$

Obviously, it has continuous paths for $t \geq 0$ with probability one and, conditionally on $X_0(\cdot)$, x is a Gaussian zero mean process. Here $(x_0(t), t \geq -r)$ denotes the so-called fundamental solution of the deterministic equation

$$\dot{x}(t) = ax_0(t) + bx_0(t-r), \quad x_0(0) = 1 \text{ and } x_0(t) = 0, t \in [-r, 0].$$

Assume in the next section that the solution x of (11) is being observed, the parameter $a = 0$ and the parameter b is unknown and has to be estimated.

Estimation problems for stochastic differential equations with time delay have been considered in a few papers up to now – see, e.g., [3; 5–7] and the references therein.

2.1. One-parameter delay differential equation

In this section we consider the differential equation with time-delay of a special structure. Assume the model

$$dx_t = bx_{t-r}dt + dw_t, \quad t \geq 0, \quad (12)$$

with an unknown parameter b such that the process (12) is stable. Note it is stable when the parameter $b \in (-\pi/2, 0)$, see [3]. Now we construct optimal and adaptive predictors for the process (12). Optimal in the mean square sense predictor is the conditional mathematical expectation

$$z_t^{(k)}(t-u) = E(x_t | x_{t-u})$$

that satisfies the following equation

$$\begin{aligned} z_t^{(k)}(t-u) &= x_{t-u} + b \int_{t-u}^{[t-(u-r)] \wedge t} x_{v-r} dv + b \int_{[t-(u-kr)] \wedge t}^t z_{v-kr}^{(0)}(t-u) dv + \\ &+ b \sum_{i=1}^{k-1} \int_{t-r}^t z_{v-(k-i)r}^{(i)}(t-u) dv, \quad kr < u \leq (k+1)r, \quad t > u. \end{aligned} \quad (13)$$

Here $\alpha \wedge \beta$ means the minimum between α and β .

Since the parameter b in the definition of the optimal predictors $z_t^{(k)}(t-u)$ is unknown, we define the adaptive predictor by formula (13) replacing the unknown b with \hat{b}_{t-u} , where \hat{b}_{t-u} is the projection

$$\hat{b}_{t-u} = \text{proj}_{[-\pi/2, 0]} b_{t-u} \quad (14)$$

of the truncated estimator of the parameter b which is defined as follows

$$b_t = \frac{\int_r^t x_{v-r} dx_v}{\int_r^t x_{v-r}^2 dv} \chi \left(\int_r^t x_{v-r}^2 dv \geq t \log^{-1} t \right), \quad t > \max(u, r).$$

Define the numbers $\sigma_0^2 = \int_0^\infty x_0^2(v) dv$ and $s_0 = \max \{ r, \exp(\sigma_0^{-2}) \}$, where $x_0(\cdot)$ is the solution of the characteristic equation $\dot{x}_0(v) = bx_0(v-r)$, $v \geq 0$ of the process (12) with $x_0(0) = 1$ and $x_0(v) = 0$, when $v \in [-r, 0)$.

For $t > u + s_0$ estimators (14) have the properties

$$E(\hat{b}_t - b)^{2p} \leq \frac{C_p}{t^p}, \quad p \geq 1,$$

which can be proven similarly to Lemma.

Denote the adaptive prediction error and rewrite it in the form

$$e_t^{(k)}(t-u) := x_t - \hat{z}_t^{(k)}(t-u) = e_t^0(t-u) + \hat{e}_t^{(k)}(t-u),$$

where $e_t^0(t-u) = x_t - E(x_t | x_{t-u})$ and $\hat{e}_t^{(k)}(t-u) = z_t^{(k)}(t-u) - \hat{z}_t^{(k)}(t-u)$. Then for every fixed $k \geq 0$ the following limit inequalities hold

$$\overline{\lim}_{t \rightarrow \infty} t \cdot \left(E(e_t^{(k)}(t-u))^2 - \sigma_0^2 \right) \leq C.$$

If it is known that $b \in [b_0, b_1]$, $-\pi/2 < b_0 < b_1 < 0$, then for $t-u > s_1 = \max \{ r, \exp(\sigma_1^{-2}) \}$, where $\sigma_1^2 = \inf_{b \in [b_0, b_1]} \sigma_0^2$ the non-asymptotic property is fulfilled

$$E(e_t^{(k)}(t-u))^2 - \sigma_0^2 \leq \frac{C}{t}.$$

3. Simulations

To confirm the convergence of the truncated estimators (4) and the properties of predictors (3) constructed on the basis of these estimators we made the simulations.

For this purpose we used the software package MATLAB. In Tables I, II the average

$$a_T = \frac{1}{100} \cdot \sum_{k=1}^{100} a_T(k)$$

of estimators $a_T(k)$ constructed according to (4) by the k -th realization $x^{(k)} = (x_t^{(k)})$, $k = 1, \dots, 100$ of the process (1) and its quality characteristics (empirical mean square error):

$$S_a^2(T) = \frac{1}{100} \sum_{k=1}^{100} (a_T(k) - a)^2$$

for different durations of observations T are given. Define the discretization step h and the number of discrete observations $N = h^{-1}T$.

Table 1

Estimators a_T and their quality characteristics $S_a^2(T)$ for $h = 0,1$

A	$N = 1000, T = 100$		$N = 2000, T = 200$		$N = 5000, T = 500$	
	a_T	$S_a^2(T)$	a_T	$S_a^2(T)$	a_T	$S_a^2(T)$
-0,3	-0,3135	0,0065	-0,3051	0,0030	-0,3070	0,0011
-0,5	-0,5113	0,0082	-0,5180	0,0059	-0,4969	0,0017
-0,8	-0,8283	0,0154	-0,8034	0,0074	-0,7949	0,0033
-1	-1,0126	0,0186	-0,9979	0,0111	-1,0017	0,0051

Table 2

Estimators a_T and their quality characteristics $S_a^2(T)$ for $h = 0,15$

A	$N = 1000, T = 150$		$N = 2000, T = 300$		$N = 5000, T = 750$	
	a_T	$S_a^2(T)$	a_T	$S_a^2(T)$	a_T	$S_a^2(T)$
-0,3	-0,3054	0,0050	-0,3098	0,0019	-0,3042	0,0001
-0,5	-0,5107	0,0053	-0,5081	0,0036	-0,5050	0,0014
-0,8	-0,8263	0,0124	-0,8114	0,0042	-0,8041	0,0020
-1	-1,0146	0,0111	-1,0165	0,0061	-1,0122	0,0028

According to the simulation results given in tables 1, 2 we can say that the empirical mean square error (MSE) $S_a^2(T)$ of the estimators a_T becomes less with growth of the sample size (see fig. 1 as well). It means that the estimator's value becomes closer to the true value of the parameter. This fact proves the effectiveness of the presented estimation procedure.

There were also constructed predictors of the observed process by the obtained estimators of $a = -0,5$ with $u = 1, h = 0,1$:

$$\hat{x}_t(t-1) = e^{\hat{a}_{t-1}} x_{t-1}, \quad t > 1.$$

The results are presented on the graphs below. The solid one identifies the real values of the observed process, the stipple one shows the predictor's value (see fig. 2).

For the fig. 3 there was computed the empirical mean square prediction error that is equal to

$$S_x^2(l) = \frac{1}{100} \sum_{k=1}^{100} \left(x_{lh}^{(k)} - \hat{x}_{lh}^{(k)}(lh-1) \right)^2, \quad l = \overline{10, N}$$

with $T = 100, N = 1000, u = 1, h = 0,1$. It shows that the prediction error converges to the value $\sigma^2 = 1 - e^{-1} \approx 0,6321$ which corresponds to the obtained theoretical results.

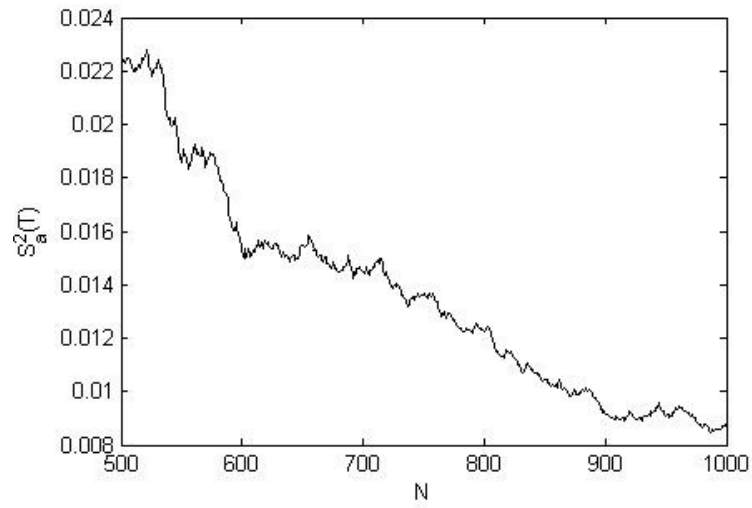


Fig. 1. Empirical MSE $S_a^2(T)$ for $a = -0,5; T = 100; h = 0,1; N = 1000$

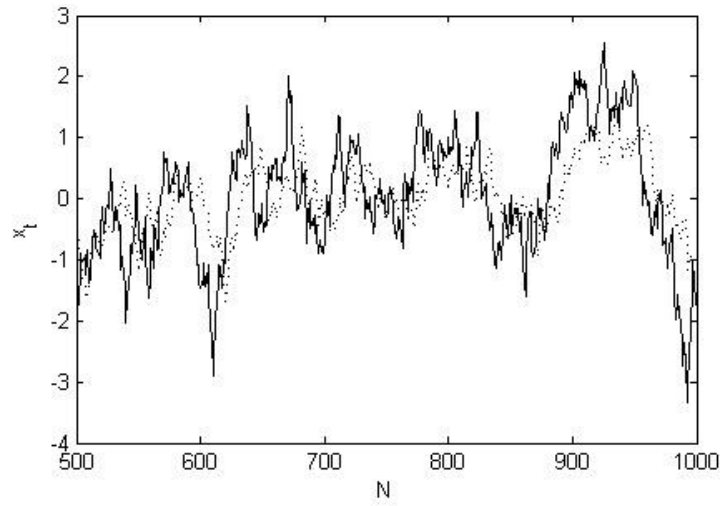


Fig. 2. Process x_t (—) and adaptive predictor $\hat{x}_t(t-1)$ (· · ·)

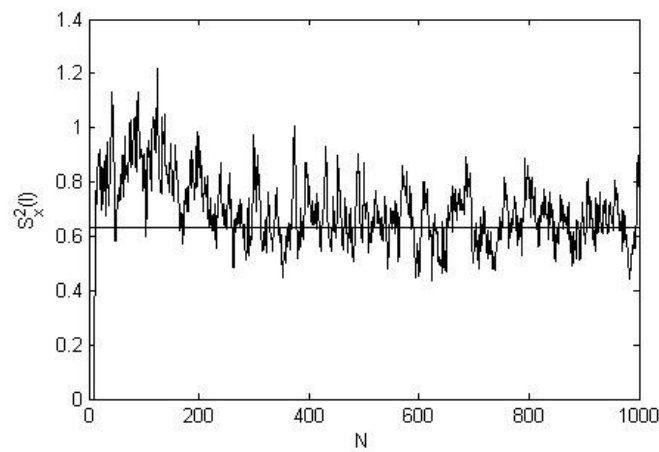


Fig. 3. Empirical MSE $S_x^2(t)$

Conclusion

Adaptive prediction problem of the Ornstein-Uhlenbeck process and one-parameter linear stochastic delay differential equation is solved. Non-asymptotic upper boundary of the prediction error is found. It is shown

that this boundary is inverse proportional to the duration of observations. Non-asymptotic properties of adaptive predictors are obtained due to the usage of the truncated estimators [2] 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 it is first employed to continuous-time systems. The obtained results can be used for the optimization problem of proposed predictors in the sense of special risk functions similar to discrete-time models [1].

REFERENCES

1. Kusainov, M.I. & Vasiliev, V.A. (2015) On optimal adaptive prediction of multivariate autoregression. *Sequential Analysis: Design Methods and Applications*. 34(2). pp. 211–234. DOI: 10.1080/07474946.2015.1030977
2. 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
3. Guschin, A.A. & Kuchler, U. (1999) Asymptotic Inference for a Linear Stochastic Differential Equation with Time Delay. *Bernoulli*. 5(6). pp. 1059–1098.
4. Myschkis, A.D. (1972) Linear Differential Equations with Delayed Argument. Moscow: Nauka. (In Russian).
5. Kuchler, U. & Vasiliev, V.A. (2001) On Sequential Parameter Estimation for Some Linear Stochastic Differential Equations with Time Delay. *Sequential Analysis*. 20. pp. 117–146. DOI: 10.1081/SQA-100106052
6. Kuchler, U. & Vasiliev, V.A. (2010) On guaranteed parameter estimation of a multiparameter linear regression process. *Automatica, Journal of IFAC, Elsevier*. 46(4). pp. 637–646. DOI: 10.1016/j.automatica.2010.01.003
7. Kuchler, U. & Kutoyants, Yu. A. (2000) Delay estimation for some stationary diffusion-type processes. *Scandinavian Journal of Statistics*. 27(3). pp. 405–414. DOI: 10.1111/1467-9469.00197

Dogadova Tatiana Valerievna. E-mail: aurora1900@mail.ru

Vasiliev Vyacheslav Arturovich, Dr. Science, prof. E-mail: vas@mail.tsu.ru

National Research Tomsk State University, Russian Federation

Поступила в редакцию 3 августа 2016 г.

Догадова Татьяна Валерьевна, Васильев Вячеслав Артурович (Национальный исследовательский Томский государственный университет, Российская Федерация).

Адаптивное прогнозирование стохастических систем с непрерывным временем.

Ключевые слова: адаптивные прогнозы; усеченное оценивание; системы с непрерывным временем, дифференциальные уравнения с запаздыванием; процесс Орнштейна–Уленбека.

DOI: 10.17223/19988605/38/3

В данной работе представлены адаптивные прогнозы для динамических систем с непрерывным временем и неизвестными параметрами динамики. Прогнозы строятся на основе усеченных оценок параметров. В частности, рассматриваются процесс Орнштейна–Уленбека и однопараметрическое стохастическое дифференциальное уравнение с запаздыванием. В статье усеченный метод оценивания впервые применяется в системах с непрерывным временем. Исследованы асимптотические и неасимптотические свойства прогнозов. Найдена скорость сходимости второго момента ошибки прогноза к ее минимальному значению.