On Adaptive Optimal Prediction of Ornstein-Uhlenbeck Process

Tatiana Dogadova and Vyacheslav Vasiliev

Department of Applied Mathematics and Cybernatics
Tomsk State University, Tomsk, Russia 634050

Abstract

This paper suggests predictors for the Ornstein-Uhlenbeck process based on the truncated estimators of parameters. For these estimators there are established the asymptotic and non-asymptotic properties with guaranteed accuracy. Strong consistency of the obtained estimators is proved. There are investigated asymptotic properties of the predictors and shown their optimality.

Mathematics Subject Classification: 60G25; 60J60; 62F10; 62M20

Keywords: truncated parameter estimation, Ornstein-Uhlenbeck process, adaptive optimal prediction

1 Introduction

Prediction is one of the most challenging and popular problems for modern researchers all over the world. A lot of practically important problems, for instance, predicting economic or technological processes, portfolio building etc. require the development of methods that allow to construct adequate mathematical models and perform statistical processing of such models. In addition, the problem of developing prediction procedures that guarantee specific properties while using small samples of fixed size is extremely topical.

Stochastic differential equations are widely used in modern financial mathematics, for example, in the problem of optimal consumption and investment...
for the financial markets. However, construction of optimal strategies requires knowledge of parameters and functions of the market model, determining the dynamics of the risky assets movement. Therefore, to be used in practical calculations, the optimal financial strategies must be able to assess the unknown parameters and functions in models of financial markets. To sum it up, the development of effective robust statistical methods of estimation of unknown parameters is a momentous problem to solve. The quality of adaptive prediction significantly depends on a choice of estimation method for model parameters. Adaptive prediction problem for discrete-time systems was solved in [1, 2, 7] on the basis of truncated estimators proposed in [5]. In this paper we solve the optimization problem of the predictors built upon truncated estimators in the sense of special risk function similar to discrete-time case considered in [7].

2 Prediction of the Ornstein-Uhlenbeck process

Assume the model

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

with an unknown parameter \( a \), where \( x_0 \) is zero mean random variable with variance \( \sigma_0^2 \) having all the moments, \( w_t \) is a 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 \( m \geq 1 \)

\[ \sup_{t \geq 0} E x_{2m}^t < \infty. \] (2)

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

Using the solution of (1) we obtain the following representation

\[ x_t = \lambda x_{t-u} + \xi_{t,t-u}, \quad t \geq u, \] (3)

where \( \xi_{t,t-u} = \int_{t-u}^{t} e^{a(t-s)} dw_s, \lambda = e^{au} \). Applying properties of the Ito integral it is easy to calculate

\[ E \xi_{t,t-u} = 0, \quad \sigma^2 := E \xi_{t,t-u}^2 = \frac{1}{2a} [\lambda^2 - 1]. \]

Optimal in the mean square sense predictor \( x_t^0 \) for \( x_t \) is the conditional mathematical expectation of \( x_t \) under the condition of \( x_t^{t-u} \) which can be found by (3)

\[ x_t^0 = \lambda x_{t-u}, \quad t \geq u. \] (4)
Since the parameters $a$ and $\lambda$ are unknown, we define the adaptive predictor
\[ \hat{x}_t = \lambda_{t-u} x_{t-u}, \quad t \geq u, \]
where
\[ \lambda_s = e^{\hat{a}_s} u, \quad s \geq 0. \]
Here
\[ \hat{a}_s = \text{proj}_{(-\infty,0]} a_s, \]
a$_s$ is the truncated estimator of the parameter $a$ constructed similar to discrete-time case [5] on the basis of the maximum likelihood estimator
\[ a_s = \frac{\int_0^s x_v dx_v}{\int_0^s x_v^2 dv} \chi\left( \int_0^s x_v^2 dv \geq s \log^{-1} s \right), \quad s > 0. \]

Denote the prediction errors of $x^0_t$ and $\hat{x}_t$ as
\[ e^0_t = x_t - x^0_t = \xi_{t-t-u}, \quad e_t = x_t - \hat{x}_t = (\lambda - \lambda_{t-u}) x_{t-u} + \xi_{t-t-u}, \quad t \geq u. \]

Now we define the loss function
\[ L_t = \frac{A}{t} e^2(t) + t, \quad t \geq u, \]
where
\[ e^2(t) = \frac{1}{t} \int_u^t e_s^2 ds \]
and the parameter $A > 0$ that is the cost of prediction error.

We also define the risk function $R_t = E L_t$ which has the following form
\[ R_t = A E e^2(t) + t \]
and consider optimization problem
\[ R_t \to \min_t. \]

For the optimal predictors $x^0_t$ it is possible to optimize the corresponding risk function
\[ R^0_t = E \left( \frac{A}{t} (e^0(t))^2 + t \right) = \frac{A \sigma^2}{t} + t \to \min_t, \]
where $(e^0(t))^2 = \frac{1}{t} \int_u^t (e^0_s)^2 ds.$
In this case the optimal duration of observations $T_A^0$ and the corresponding value of $R_A^0$ are respectively

$$T_A^0 = A^{1/2} \sigma, \quad R_A^0 = 2A^{1/2} \sigma, \quad (11)$$

where \( \sigma := \sqrt{\sigma^2} \).

However, since \( a \) and as follows, \( \sigma \) are unknown and both $T_A^0$ and $R_A^0$ 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

$$T_A = \inf\{t \geq t_A : t \geq A^{1/2} \sigma_t \}, \quad (12)$$

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

$$\sigma_t^2 = \frac{1}{2} \theta_t \cdot [\lambda_t^2 - 1]$$

and \( \theta_t \) is the truncated estimator of \( \theta = a^{-1} \) defined as follows

$$\theta_t = a_t^{-1} \cdot \chi[a_t \leq -\log^{-1} t], \quad t > 0.$$  

Estimators \( a_t, \lambda_t \) and \( \sigma_t \) that are used in construction of adaptive predictors have the properties given in Lemma 2.1 below which can be proved similar to discrete-time case [5].

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

**Lemma 2.1.** Assume the model (1). Then the estimators \( a_t, \lambda_t \) and \( \sigma_t \) are strongly consistent. Moreover, for \( t - u > s_0 := \exp(2|a|) \) the following properties hold:

$$E(a_t - a)^{2p} \leq \frac{C}{t^p}, \quad (13)$$

and

$$E(\lambda_t - \lambda)^{2p} \leq \frac{C}{t^p}, \quad p \geq 1, \quad (14)$$

$$E(\sigma_t^2 - \sigma^2)^{2p} \leq \frac{C \log^{2p} t}{t^p}, \quad p \geq 1. \quad (15)$$

Analogously to [3], [4] and [7], 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 $R_A^0$ and the obviously modified risk

$$\bar{R}_A = A \cdot E\frac{1}{T_A} e^2(T_A) + ET_A. \quad (16)$$
Theorem 2.2. Assume the model (1) and \( t_A \) that is defined in (12). Let the predictors \( \hat{x}_i \) be defined by (5), the times \( T_A^0, T_A \) and the risk functions \( R_i^0, \bar{R}_A \) defined by (11), (12) and (10), (16) respectively. Then for every \( a < 0 \)

\[
i) \quad \frac{T_A}{T_A^0} \xrightarrow{a \to \infty} 1 \quad \text{a.s.}; \tag{17}
\]

\[
ii) \quad \frac{E T_A}{T_A^0} \xrightarrow{a \to \infty} 1; \tag{18}
\]

\[
iii) \quad \frac{\bar{R}_A}{R_A^0} \xrightarrow{a \to \infty} 1. \tag{19}
\]

Proof of Theorem 2.2: First we prove the assertion (i). By definitions of \( T_A \) and \( T_A^0 \) we have

\[
\frac{T_A}{T_A^0} = \frac{t_A}{A^{1/2} \sigma} \cdot \chi[T_A = t_A] + \frac{\sigma t_A}{\sigma} \cdot \chi[T_A > t_A]
\]

\[
= \frac{t_A}{A^{1/2} \sigma} \cdot \chi[t_A \geq A^{1/2} \sigma t_A] + \frac{\sigma t_A}{\sigma} \cdot \chi[A^{1/2} \sigma t_A > t_A] \xrightarrow{a \to \infty} 1 \quad \text{a.s.}
\]

The property (17) is proven.

Now we prove the assertion (ii). The following representation will be used

\[
\frac{T_A}{T_A^0} = 1 + \frac{T_A - A^{1/2} \cdot \sigma}{A^{1/2} \cdot \sigma} = 1 + S_A.
\]

Rewrite \( S_A \) in a form

\[
S_A = \frac{t_A - A^{1/2} \cdot \sigma}{A^{1/2} \cdot \sigma} \cdot \chi[T_A = t_A] + \frac{T_A - A^{1/2} \cdot \sigma}{A^{1/2} \cdot \sigma} \cdot \chi[T_A > t_A]
\]

\[
= \frac{t_A - A^{1/2} \cdot \sigma}{A^{1/2} \cdot \sigma} \cdot \chi[t_A \geq A^{1/2} \sigma t_A] + \frac{\sigma t_A - \sigma}{\sigma} \cdot \chi[A^{1/2} \sigma t_A > t_A].
\]

Then, by the third assertion of Lemma

\[
|ES_A| \leq P[t_A \geq A^{1/2} \sigma t_A] + \sigma^{-1} E|\sigma t_A - \sigma|
\]

\[
\leq P[\sigma t_A \leq \log^{-1} A] + \sigma^{-2} \sqrt{E(\sigma t_A^2 - \sigma^2)^2} \xrightarrow{a \to \infty} 0.
\]

The property (18) is proven.

Prove the assertion (iii).

The left-hand side in this assertion can be rewritten as

\[
\frac{R_A}{R_A^0} = \frac{1}{2} \left( A^{1/2} E \frac{1}{T_A \sigma} e^2(T_A) + \frac{E T_A}{A^{1/2} \sigma} \right).
\]
Then from (17) it follows that to prove (18) it is enough to show the convergency
\[ A^{1/2}E \frac{1}{T_A \sigma} e^2(T_A) \rightarrow 1. \] (20)

For some \( \varepsilon \in (0, \sigma) \) we denote
\[ T' = (\sigma - \varepsilon) \cdot A^{1/2}, \quad T'' = (\sigma + \varepsilon) \cdot A^{1/2}. \]

Further we use the following properties
\[ P(T_A < T') \leq C \frac{\log^3 m A}{A^{m/2}}, \quad P(T_A > T'') \leq C \frac{\log^3 m A}{A^{m/2}}, \] (21)
which are fulfilled for every \( m \geq 1 \). Indeed,
\[ P(T_A < T') = P(A^{1/2} \cdot \sigma_t_A < (\sigma - \varepsilon)A^{1/2}) \leq P(\sigma_t_A - \sigma > \varepsilon) \]
\[ \leq \frac{1}{(\varepsilon \sigma)^{2m}} E(\sigma_t_A^2 - \sigma^2)^{2m} \leq C \frac{\log^2 m (A^{1/2} \log^{-1} A) \cdot \log m A}{A^{m/2}} \leq C \frac{\log^3 m A}{A^{m/2}}. \]

Analogously, we have
\[ P(T_A > T'') = P(T'' < A^{1/2} \sigma_t_A) = P(\sigma + \varepsilon < \sigma_t_A) \]
\[ \leq P(\sigma_t_A - \sigma > \varepsilon) \leq (\varepsilon \sigma)^{-2m} \cdot E(\sigma_t_A^2 - \sigma^2)^{2m} \leq C \frac{\log^3 m A}{A^{m/2}}. \]

Split the proof of (20) into 3 parts:

1) \( A^{1/2} \cdot E \frac{1}{T_A \sigma} e^2(T_A) \cdot \chi[T_A < T'] \rightarrow 0 \) \hspace{1cm} (22)

2) \( A^{1/2} \cdot E \frac{1}{T_A \sigma} e^2(T_A) \cdot \chi[T_A > T''] \rightarrow 0 \) \hspace{1cm} (23)

3) \( A^{1/2} \cdot E \frac{1}{T_A \sigma} e^2(T_A) \cdot \chi[T' \leq T_A \leq T''] \rightarrow 1 \) \hspace{1cm} (24)

as \( A \rightarrow \infty \).

By definition of \( e^2(t) \) we have
\[ A^{1/2} \cdot E \frac{1}{T_A \sigma} e^2(T_A) = A^{1/2} \cdot \frac{1}{T_A^2} \int_u^T e^2_t dt = A^{1/2} \cdot \frac{1}{T_A^2} \int_u^T (\lambda_{t-u} - \lambda)^2 x^2_{t-u} dt \]
\[ + 2A^{1/2} \cdot \frac{1}{T_A^2} \int_u^T (\lambda_{t-u} - \lambda) x_{t-u} \cdot \xi_{t,t-u} dt + A^{1/2} \cdot \frac{1}{T_A^2} \int_u^T \xi^2_{t,t-u} dt \]
\[ =: I_1 + I_2 + I_3. \] (25)
Prove 1). By the Cauchy-Schwarz-Bunyakovsky inequality, (2) and Lemma 2.1
\[
EI_1 \cdot \chi[T_A < T'] \leq \frac{A^{1/2}}{t_A^2} \int_u^{T'} E(\lambda_{t-u} - \lambda)^2 x_t^2 dt
\]
\[
\leq \frac{A^{1/2}}{t_A^2} \int_u^{T'} \sqrt{E(\lambda_{t-u} - \lambda)^4 \cdot E x_t^4 dt} \leq C \frac{A^{1/2}}{t_A} \int_u^{T'} \frac{dt}{t} \leq C \frac{A^{1/2} \log A}{t_A^2} \to 0.
\]
By Cauchy-Schwarz-Bunyakovsky inequality and Lemma 2.1 we obtain
\[
E|I_2| \cdot \chi[T_A < T'] \leq \frac{2A^{1/2}}{t_A^2} E \int_u^{T_A} (\lambda_{t-u} - \lambda)x_{t-u} \xi_{t-u} dt \cdot \chi[T_A < T']
\]
\[
\leq \frac{2A^{1/2}}{t_A^2} \int_u^{T'} \sqrt{E(\lambda_{t-u} - \lambda)^2 x_t^2 \xi_{t-u}^2 dt}
\]
\[
\leq \frac{2A^{1/2}}{t_A^2} \int_u^{T'} (E(\lambda_{t-u} - \lambda)^4 \cdot E x_t^4) ^{1/4} dt
\]
\[
\leq C \frac{A^{1/2}}{t_A} \int_u^{T'} \frac{dt}{t^{1/2}} \leq C \frac{A^{1/2} \cdot (T')^{1/2}}{t_A^2} \leq C \frac{\log^2 A}{A^{1/4}} \to 0.
\]
Using (21) with \( m = 2 \) as \( A \to \infty \) we get
\[
EI_3 \cdot \chi[T_A < T'] \leq \frac{A^{1/2}}{t_A^2} \int_u^{T'} E \xi_{t-u}^2 \cdot \chi[T_A < T'] dt
\]
\[
\leq \frac{A^{1/2}}{t_A^2} \int_u^{T'} \sqrt{E \xi_{t-u}^2 dt} \cdot P_2(T_A < T') \leq C \frac{A^{1/2}}{t_A^2} \cdot A^{1/2} \cdot \frac{\log^6 A}{A} = C \frac{\log^6 A}{A} \to 0.
\]
Prove (23). Using (21) with \( m = 2 \) and (25) we get
\[
EI_1 \cdot \chi[T_A > T''] = A^{1/2} E \frac{1}{T_A^2} \int_u^{T_A} (\lambda_{t-u} - \lambda)^2 x_{t-u}^2 dt \cdot \chi[T_A > T'']
\]
\[
\leq \frac{A^{1/2}}{\sigma} \cdot P^{1/2}(T_A > T'') \sqrt{E \sup_{s \geq T''} \frac{1}{s^4} \left( \int_u^{s} (\lambda_{t-u} - \lambda)^2 x_{t-u}^2 dt \right)^2}
\]
\[
\leq C \log^3 A \sqrt{E \sup_{s \geq T''} \frac{1}{s^3} \int_u^{s} (\lambda_{t-u} - \lambda)^4 x_{t-u}^4 dt}.
\]
For simplification we assume that the time \( T'' \) is integer. Then
\[
C \log^3 A \sqrt{\sum_{n \geq T''} E \sup_{n \leq s \leq n + 1} \frac{1}{s^3} \int_u^{s} (\lambda_{t-u} - \lambda)^4 x_{t-u}^4 dt}
\]
\[ \leq C \log^3 A \sqrt{\sum_{n \geq T''} \frac{1}{n^3} \int_u^{n+1} dt} \leq C \log^3 A \cdot \sqrt{\sum_{n \geq T''} \frac{1}{n^3}} \]

\[ \leq C \log^3 A \frac{1}{T''} \leq C \frac{\log^3 A}{A^{1/2}} \to 0. \]

Similarly we obtain:

\[ E|I_2| \cdot \chi[T_A > T''] \leq \frac{2A^{1/2}}{\sigma} \sqrt{E \sup_{s \geq T''} \frac{1}{s^4} \left( \int_u^s (\lambda_{t-u} - \lambda)x_{t-u} \xi_{t,u} dt \right)^2} \]

\[ \leq C \log^3 A \cdot \sqrt{E \sup_{s \geq T''} \frac{1}{s^3} \int_u^s (\lambda_{t-u} - \lambda)^2 x_{t-u}^2 \xi_{t,u}^2 dt} \leq C \cdot \frac{\log^{7/2} A}{A^{1/2}} \to 0, \]

as well as

\[ EI_3 \cdot \chi[T_A \geq T''] \leq \frac{A^{1/2}}{\sigma} P^{1/2}(T_A > T'') \cdot \sqrt{E \sup_{s \geq T''} \frac{1}{s^4} \left( \int_u^s \xi_{t,u}^2 dt \right)^2} \]

\[ \leq C \frac{\log^3 A}{A^{1/2}} \to 0. \]

Finally, for the (iii) assertion of the Theorem 2.2 let us decompose (24) in a following way

\[ EI_3 \cdot \chi[T' \leq T_A \leq T''] = J_1 + J_2 + J_3 + J_4, \]

where

\[ J_1 = A^{1/2} \cdot E \frac{1}{T_A^2 \sigma} \int_u^{T_A} (\lambda_{t-u} - \lambda)^2 x_{t-u}^2 dt \cdot \chi[T' \leq T_A \leq T''], \]

\[ J_2 = 2A^{1/2} \cdot E \frac{1}{T_A^2 \sigma} \int_u^{T_A} (\lambda_{t-u} - \lambda)x_{t-u} \xi_t dt \cdot \chi[T' \leq T_A \leq T''], \]

\[ J_3 = A^{1/2} \cdot E \frac{1}{T_A^2 \sigma} \int_u^{T_A} (\xi_{t,u}^2 - \sigma^2) dt \cdot \chi[T' \leq T_A \leq T''], \]

\[ J_4 = E \frac{A^{1/2} \sigma}{T_A} \cdot \chi[T' \leq T_A \leq T'']. \]

We start with the assessment of \( J_1 \)

\[ J_1 \leq \frac{A^{1/2}}{(T'')^2 \sigma} \int_u^{T''} E(\lambda_{t-u} - \lambda)^2 x_{t-u}^2 dt \]
\[ \leq C A^{-1/2} \cdot \int_u^{A^{1/2}(\sigma+\varepsilon)} \sqrt{E(\lambda_{t-u} - \lambda)^4 \cdot E x^4_{t-u}} \, dt \leq C \frac{\log A}{A^{1/2}} \to 0. \]

\[ J_2 \leq \frac{2A^{1/2}}{(T')^2 \sigma} \sqrt{E \left( \int_u^{T_A} (\lambda_{t-u} - \lambda) \xi_{t,t-u} \, dt \right)^2} \cdot \chi[T_A \leq T''] \]

\[ \leq \frac{2A^{1/2}}{(T')^2 \sigma} \sqrt{\int_u^{T''} E (\lambda_{t-u} - \lambda)^2 \xi_{t,t-u}^2 \, dt} \leq C \frac{\log^{1/2} A}{A^{1/4}} \to 0. \]

Denote \( N_A = [u^{-1} \cdot A^{1/2}(\sigma + \varepsilon)] + 1 \), where \([b]_1\) signifies integer part of number \(b\). Using the maximal inequality for martingales we obtain

\[ J_3 \leq C \frac{1}{A^{1/2}} \sqrt{E \sup_{s \leq T''} \left( \int_u^s (\xi_{t,t-u}^2 - \sigma^2) \, dt \right)^2} \]

\[ \leq C \frac{1}{A^{1/2}} \cdot \sqrt{E \sup_{t \leq N_A} \left( \sum_{n=1}^{N_A} \int_{(n-1)u}^{nu} (\xi_{t,t-u}^2 - \sigma^2) \, dt \right)^2} \]

\[ \leq \frac{C}{A^{1/2}} \cdot \sqrt{\sum_{n=1}^{N_A} E \left[ \int_{(n-1)u}^{nu} (\xi_{t,t-u}^2 - \sigma^2) \, dt \right]^2} \leq \frac{C}{A^{1/2}} \cdot \sqrt{N_A} \leq \frac{C}{A^{1/4}} \to 0. \]

Now we show the convergence of \( J_4 \to 1 \) as \( A \to \infty \). \( J_4 \) can be rewritten in a form

\[ J_4 = P[T' \leq T_A \leq T''] + L_A, \]

where

\[ L_A = \frac{1}{T_A} E \left[ A^{1/2} \sigma - T_A \right] \cdot \chi[T' \leq T_A \leq T'']. \]

By the definitions of \( T_A, T', T'' \) and according to Lemma 2.1 the probability

\[ P[T' \leq T_A \leq T''] \to 1 \quad \text{as} \quad A \to \infty. \]

Note that for \( A \gg 1 \) it holds that \( T' > t_A \), then \( T_A = A^{1/2} \sigma t_A \) and

\[ L_A \leq \frac{1}{T'} E[T_A - A^{1/2} \sigma] \cdot \chi[T_A \geq T'] = C \cdot E[\sigma t_A - \sigma] \leq C \cdot \sqrt{E(\sigma_{t_A}^2 - \sigma^2)} \to 0. \]

**Acknowledgements.** Tomsk State University, International Laboratory of Statistics of Stochastic Processes and Quantitative Finance. This study (research) was supported by The Ministry of Education and Science of the Russian Federation, Goszadanie No 2.3208.2017/PCh
References


Received: February 17, 2017; Published: March 2, 2017