SCIREA Journal of Mathematics

http://www.scirea.org/journal/Mathematics

November 13, 2016

Volume 1, Issue1, October 2016

ON PARAMETER ESTIMATION

FOR ORNSTEIN-UHLENBECK PROCESS

LabadzeLevan[1], SokhadzeGrigol[2], KvatadzeZurab[1]

[1]Georgian Technical University

[2]I.Javakhishvili Tbilisi State University

Abstract

An estimation procedure for Ornstein–Uhlenbeck process drift and volatility coefficients is given. The procedure is based on the maximum likelihood principle andplug-in-estimator.

KeyWords： Estimation,MLE,Ornstein-Uhlenbeck processes, plug-in-estimator.

Ornstein-Uhlenbeck processes are known to play a significant part in such fields as quant mechanics, neuron theory, stochastic control, chemical reactions, financial mathematics, etc. ([1-4]). This process and its extensions have been studied in detail (see, e.g., [5–7] and references given in these publications).

It should be noted, however, that the problems of statistical parameter estimation have not been thoroughly investigated yet. Steps made in this direction can be found in [8–12]. A general analysis of the obtained results is given in monograph [13].

The aim of the present work is to develop methods for statistical estimation of drift and volatility coefficients of Ornstein–Uhlenbeck type statistical differential equations. Drift coefficients will be estimated using the maximum likelihood method while for volatility coefficients we use the least squares method. Moreover the equation itself (and Ornstein–Uhlenbeck process, respectively) will be considered in infinite space. In this case an estimation technique from [14-15] will be used. Such estimators, as it results from the above-mentioned works, are consistent and asymptotically normal. Before we state the given problem consider the necessary concepts and results of the above-mentioned works.

Let $\{\Omega, \mathcal{F}, P\}$ denote a fixed complete probability space. Let $\mathbb{E}$ be a separable reflexive Banach space, $\mathbb{E}^n = \mathbb{E} \times \mathbb{E} \times \dots \times \mathbb{E}$, $\mathcal{B}^n$beBorelσ–algebra in$\mathbb{E}^n$–ðo. $\mathbb{B}$is also a separable Banach space, $\Theta \subset \mathbb{B}$is its subset, which plays the part of a parametric set. Consider a random element $X = X(\omega, \theta), \omega \in \Omega, \theta \in \Theta$ taking values in the set $\mathbb{E}$. Suppose we have identically, same as Xdistributed independent observations $X_1, X_2, \dots, X_n$. Let θ be an unknown parameter that we have to estimate on the basis of the given observation by means of a statistics $\hat{\theta}_n(X_1, X_2, \dots, X_n)$. If we consider the sample $X_1, X_2, \dots, X_n$ as the vector $Y = (X_1, X_2, \dots, X_n)$, we obtain a random vector taking values in $\mathbb{E}^n$. Thus we have, respectively, a statistical structure of independent random elements. Naturally, on a $\{\mathbb{E}^n, \mathcal{B}^n\}$-measurable space the distribution of Y will be formed by the equality

$$P_\theta(A) = P(Y^{-1}(A)), \qquad A \in \mathcal{B}^n.$$

As we can see, $P_\theta(\cdot)$is a function of two variables: one variable takes set values, while the other is a space parameter. In a number of cases the above function is smooth with respect to both variables. We will show what is ment. Let $P_{\theta,h}(A) = P_\theta(A + h), h \in \mathbb{E}^n$. We say that $P_\theta(\cdot)$ is differentiable with respect to the set variable in the direction of the vector $h$if there exists an alternating measured$d_h P_\theta$, such that for every$A \in \mathcal{B}^n$the following equality

$$P_{\theta,th}(A) - P_\theta(A) = t d_h P_\theta(A) h + o(t), t \in R.$$

is true. It can be easily shown that $d_h P_\theta \ll P_\theta$and thus there exists a Radon-Nikodym derivative

$$\frac{d_h P_\theta(dx)}{P_\theta(dx)} = \beta_\theta(x, h).$$

Function $\beta_\theta(x, h)$ is called logarithmic derivative.

We should note that we have defined the logarithmic derivative in the direction of constant vectors. This definition can be extended in the direction of the vector field as well. It is also interesting to see that the notion of logarithmic derivative is closely connected to the partial integration formula in infinite-dimensional spaces (see [16]).

Let $\mathbb{H} \subset \mathbb{E}$ be a separable Hilbert space and the plug-in-operator be of Hilbert-Schmidt type. Then we can consider a Hilbert-Schmidt structure (or, in other words, the Gelfand triple)$\mathbb{E}^* \subset \mathbb{H} \subset \mathbb{E}$.

If we view the function $P_\theta(A)$ asparameter function for a fixed A, then using a standard procedure we can define a derivative of this function by means of a parameter in a certain specific direction $\vartheta \in \Theta$. This derivative will be also denoted as $d_\theta P_\theta(A)\vartheta$. This is an alternating measure absolutely continuous with respect to P_θ. The logarithmic derivative as the corresponding Radon-Nikodym derivative in the given class of measures $P_\theta(A)$ is called according to the parameter. It is denoted by the symbol $\rho_\theta(x, \vartheta)$. Hence

$$\rho_\theta(x, \vartheta) = \frac{d_\theta P_\theta(dx)\vartheta}{P_\theta(dx)}.$$

Suppose we have a random element X with the distribution $P_\theta(A)$. Obviouslythe value$X$besides the random parameter also depends onθ–: $X = X(\omega, \theta)$.It is very important that there is a connection between the two above-defined types of logarithmic derivatives.In particular, it is true that. Theorems 1-3 are proved in reference ([15]), but I have changed their forms to use them more effectively to prove theorems 4-5.

Theorem 1 ([15]).Let the following conditions be satisfied (the so called regularity conditions):

1) For elements $X = X(\omega, \theta)$ there exists a derivative $\frac{dX}{d\theta} = X'$in the direction of the vector $\vartheta \in \Theta$. This is a linear map $\Theta \to \mathbb{E}$. Besides, suppose $\|X'\vartheta\|_{\mathbb{E}} \in \mathbb{L}_2(\Omega, P)$;

2) The function$f(x) = E\{(X')\vartheta | X = x\}$ is strictly continuous for each pair(θ, ϑ);

3) The measure family $\{P_\theta, \theta \in \Theta\}$ has a logarithmic derivative - $\rho_\theta(x, \vartheta)$ for a unity of parameter directions which constitute a dense linear subclass in the space$\mathbb{B}_0 \subset \mathbb{B}$. Besides, $\rho_\theta(x, \vartheta) \in \mathbb{L}_2(\mathbb{E}, P_\theta), \vartheta \in \mathbb{B}_0, \theta \in \Theta$;

4) The above family $\{P_\theta, \theta \in \Theta\}$ has a logarithmic derivative - $\beta_\theta(x, h)$, for a class of directions the unity$\mathbb{E}_0$of which constitutes a dense subspace of the space $\mathbb{E}$.Besides,$\beta_\theta(x, h) \in \mathbb{L}_2(\mathbb{E}, P_\theta)$, $h \in \mathbb{E}_0, \theta \in \Theta$.

Then the following equality

$$\rho_\theta(x, \vartheta) =- \beta_\theta \left(x, K_{\theta,\vartheta}(x)\right),$$

is true, where $K_{\theta,\vartheta}(x) = E\{(X')\vartheta | X = x\}$.

It follows from the results of [15] that the parameter θ can be estimated using the infinite-dimentional version of the maximum likelihood principle.

Suppose we have $T: \mathbb{E} \to R$ -measurable map (statistics) and $g(\theta) = E_\theta T(X)$. The infinite-dimentional version of the Cramer-Rao inequality is true. In particular,

Theorem 2 ([15]).Let Conditions 1-4 of Theorem 1be satisfied. Suppose the following equality

$$d_\theta \int_{\mathbb{E}} T(x) P_\theta(dx) = \int_{\mathbb{E}} T(x)\, d_\vartheta P_\theta(dx).$$

is true. Then

$$VarT(x) \geq \frac{\left(g'_\vartheta(\theta)\right)^2}{E_\theta \rho_\theta^2(X; \vartheta)} = \frac{\left(g'_\vartheta(\theta)\right)^2}{E_\theta \beta_\theta^2 \left(X; E\left(X'_\theta \vartheta | X\right)\right)}.$$

Note that the expression $\mathcal{I}(\theta)\vartheta = E_\theta \rho_\theta^2(X; \vartheta) = E_\theta \beta_\theta^2\left(X; E(X'_\theta \vartheta | X)\right)$ is called the Fisher information.

Now consider the repeated statistical structure

$$\left\{\mathbb{E}^n, \mathfrak{B}^n, \{P_\theta, \theta \in \Theta\}\right\} = \left\{\mathbb{E}, \mathfrak{B}, \{p_\theta, \theta \in \Theta\}\right\}^n.$$

Theorem 3 [15]. Suppose in the statistical structure $\left\{\mathbb{E}, \mathfrak{B}, \{p_\theta, \theta \in \Theta\}\right\}$ the measure family $\{p_\theta, \theta \in \Theta\}$ has a logarithmic derivative dense everywhere in the direction of the parameter ϑ.

Then in the repeated statistical structure the class of measures$\{P_\theta, \theta \in \Theta\}$also has a logarithmic derivativewith respect to the parameter $(\vartheta, \vartheta, \dots, \vartheta)$ in the direction $L_\theta\big((x_1, \dots, x_n); (\vartheta_1, \dots, \vartheta_n)\big)$and

$$L_\theta\big((x_1, \dots, x_n); (\vartheta_1, \dots, \vartheta_n)\big) = \sum_{k=1}^{n} \rho_\theta(x_k; \vartheta) = -\sum_{k=1}^{n} \beta_\theta(x_k; E\{X_k'\vartheta | X_k = x_k\}).$$

The given results allow us to establish the maximum likelihood principle for the general case.Since in the future we will only have to estimate a finite number of parameters we will show the principle in the case when the parametric space$\mathbb{B}$is finite-dimensional.

Consider the equation

$$\sum_{k=1}^{n} \rho_\theta(x_k; \vartheta) = 0,$$

For every$\vartheta \in \mathbb{B}$. If this equation has a solution with respect to , such that

$$\frac{d}{d\theta} \rho_\theta(x; \vartheta)$$

is negatively defined, then the solution iscalled maximum likelihood estimation ofthe unknown parameter θ.

It follows from the results of [15] that maximum likelihood estimation is consistent and asymptotically normalparameter estimation.

The basic point for the application of this technique is to find a logarithmic derivative for the given class of measures. We will present a simple tool helping to calculate the logarithmic derivative.

Theorem 4. Let the class of measures$\{P_\theta, \theta \in \Theta\}$ be dominated by aσ–finiteμmeasure and the Radon-Nikodym derivative $\frac{dP_\theta}{d\mu}(x) = \pi(\theta, x)$be continuously differentiable with respect toθ. Then the class of measures $\{P_\theta, \theta \in \Theta\}$has a logarithmic derivative with respect to the parameter $\rho_\theta(x; \vartheta)$ and the following equality is true:

$$\rho_\theta(x; \vartheta) = \frac{grad\pi(\theta, x)\vartheta}{\pi(\theta, x)}.$$

The proof is simple. It follows from an obvious implication

$$\frac{dP_\theta}{d\mu}(x) = \pi(\theta, x) \Longrightarrow \frac{dd_\vartheta P_\theta}{d\mu}(x) = grad\pi(\theta, x)\vartheta.$$

Let $\{\Omega, \mathcal{F}, P\}$ be a fixed probability space. $H_+ \subset H \subset H_-$ is a triple of Hilbert spaces with quasi-kernel imbedding. In these spaces we write the index of the space itself as the index of the scalar product and norms. The plug-in-operator $i: H \to H\_$ is the Hilbert-Schmidt operator. Coupling of elements of spaces H_+ and H_- can be expressed by a scalar product of the space H. Every space to be examined is assumed to be separable, so that $i^*: H_+ \to H$.

Let A be a linear, possibly unbounded operator in H the domain $\mathcal{D}(A)$ of which is dense in H. Furthermore we assume that A represents a strictly continuous semigroup generator denoted in the following way: $S(t) = e^{tA}$. Besides, let $B: H_- \to H$ be a Hilbert-Schmidt type operator and $w_t, t \geq 0$ be a Wiener process in H_-.

In the Hilbert space triple $H_+ \subset H \subset H_-$ consider an Ornstein–Uhlenbeck type stochastic differential equation

$$dX_t = -AX_t dt + \sigma B dw_t, \quad X_0 = x_{0,} \ \ t \geq 0, \tag{1}$$

where $\sigma > 0$ is an unknown parameter. It is called the volatility parameter. Suppose we have $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ observations of the process X_t. The observation vector is

$$X^{(n)} = (X_1, X_2, \dots, X_n), \text{ where: } X_m = X_{t_m}, m = 1, \dots, n.$$

Equation (1) should be understood as a symbolic representation of the following integral equality:

$$X_t = e^{-At} - \sigma \int_0^t e^{-A(t-s)} B dw_s. \tag{2}$$

Correspondingly, X_t is a Gaussian process in H.

Let the operator A (we call it the drift operator) have the form :

$$A = \sum_{k=1}^{m} a_k A_k, \tag{3}$$

$A_k, k = 1,2, \dots, m$ are known linear operatorsand either all or some of the operators can be unbounded, but the domain of definition of each of them includes (or coincides with)$\mathcal{D}(A)$.In expression (3) $a_k, k = 1,2, \dots, m$are unknown parameters to be estimated based onobservations $X^{(n)}$.

First we deal with the estimation of the drift parameters.We estimate them using the maximum likelihood method described above.It is well-known (see, e.g. [17])that this method is ineffective when volatility coefficientsare estimated, but we use it for the estimation of the parameters $a_k, k = 1,2, \dots, m$. Then by means of the “plug-in-operator”we estimate volatility using the least squares method.

Let$\theta = (a_1, a_2, \dots, a_m)^T$ denote the vector of unknown parameters.Here and in the forthcoming the upper index$T$denotesvector (matrix) transpose.The solutionX_t of Equation (1) represents a continuous process and it belongs to the space of continuous functions $C[0;T]$. In this space we takeσ–algebra$\mathfrak{B}[0;T]$ of Borel sets. The random process (Ornstein–Uhlenbeck process) $X_t$on a measurable space$\{C[0;T], \mathfrak{B}[0;T]\}$is distributed by the following equality

$$P_\theta(A) = P\{X^{-1}(A)\}, A \in \mathfrak{B}[0;T].$$

In order to apply the maximum likelihood principleit is necessary to calculate the logarithmic derivative of this class of measures.The logarithmic derivative can be calculated both with respect to the parameterand with respect to the set variable. It follows from the connection established by Theorem 1. In our case it is easier to show it using Theorem 4.So we calculate the logarithmic derivative with respect to the parameter.Since the parameter vector is an element of a finite-dimensionalspace it makes no difference in which direction we performderivation.We are goingt to use the Radon-Nikodym derivative.

Integrate Equation (1) and rewrite it in the following way

$$Y_t - x_0 = X_t + \int_0^t AX_s ds, \tag{4}$$

where

$$Y_t = \sigma B \gamma_t.$$

$\gamma_t$is the so called "white noise" in the spaceH_- (i.e. this is the corresponding measure of the Gaussian process inH_-, the correlation operatorof which is unit operator in H). Obviously, the average of Y_t is 0and the correlation operator equals $\sigma^2 BB^*$. The latter is a kernel $-$ typeoperator. (4) is a transformation into the Hilbert space and it is possible to use the theorem of measure transformation theorem (see [18])for it.According to this theorem, if the operator$AB$is bounded, the measure P_θ is equivalent to the measure μ, where μ is the distribution of the processY_t.We can write the Radon-Nikodymderivative :

$$\pi(x,\theta) = \frac{dP_\theta}{d\mu}(x) =$$

$$= exp\left\{-\frac{1}{\sigma^2}\int_0^T (B^*AX_s, dX_s)_H - \frac{1}{2\sigma^2}\int_0^T \|B^*AX_s\|_H^2 ds\right\}. \tag{5}$$

In order to apply Theorem4 introduce the following notation.

$$b = (b_1, b_2, \dots, b_m)^T, \qquad a = (a_1, a_2, \dots, a_m)^T, \qquad \left(C_{ij}\right)_{i,j=1}^m,$$

where

$$b_k = \int_0^T (B^*A_kX_s, dX_s)_H, \qquad k = 1,2,\dots,m,$$

$$C_{ij} = \int_0^T \langle B^*A_iX_s, B^*A_jX_s\rangle_H \, ds, i,j = 1,2,\dots,m.$$

Consequently,

$$\rho(x,\theta) = -\frac{1}{\sigma^2}b - \frac{1}{\sigma 2}Ca.$$

It follows that

$$\frac{d\rho(x,\theta)}{d\theta} = -\frac{1}{\sigma^2}\begin{pmatrix} C_{11} & 0 & \dots & 0 \\ 0 & C_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & C_{mm} \end{pmatrix}.$$

As one can see, the latter is a negatively defined matrix.

Since all conditions of Theorem 4 are satisfiedit is necessary to find the solution of the maximum likelihood equation:

$$E\left\{\left(-\frac{1}{\sigma^2}b-\frac{1}{\sigma^2}C\hat{a}_n\right)|X^{(n)}\right\}=0.$$

The solution of this equation is

$$\hat{a}_n=\left(E\{C|X^{(n)}\}\right)^{-1}\left(E\{b|X^{(n)}\}\right). \tag{6}$$

As it follows from (6),the estimator of the operator$A$can be written in the following way

$$\hat{A}=\sum_{k=1}^{m}\hat{a}_kA_k. \tag{7}$$

The results of [15] imply that the estimator (7) is consistent. Moreover this work implies that the asymptotic normality theorem is also true and

$$\sqrt{n}(\hat{a}_n-a)\to N(0,\mathcal{I}^{-1}(a)),$$

where$\mathcal{I}^{-1}(a)$ is Fisher information.

So that the following theorem is true.

Theorem 5. Let for Ornstein–Uhlenbeck Equation (1) the following conditions be satisfied:

1. $A$is a linear, possiblyunbounded, operator in$H$, with the domain of definition$\mathcal{D}(A)$dense in H. It has the form of (3). Besides, A is a strictly continuous semi group generator$S(t)=e^{tA}$;
2. $B: H_{-}\to H$is a Hilbert-Schmidt type operator linear operator whileAB is bounded;

Then estimator (7) is consistent for the drift parameter and asymptotically normal.

Example. Let $H=R, m=1, \sigma B=1, A_1=1, x_0=0, a_1=\theta$. Then Formula (6) leads to a classical formula (see [19])

$$\hat{\theta}_n=\frac{E\left(\int_0^T X_s dX_s\,|X^{(n)}\right)}{E\left(\int_0^T X_s^2 ds|X^{(n)}\right)}.$$

When the consistent estimator $\hat{A}_n$, of the operator$A$ has been constructedwe can construct the estimator of the parameterσ. We will need the so called“plug-in-estimator”.

Integrate Equation (1)

$$X_T + x_0 + \int_0^T \hat{A}_n X_s \, ds = \sigma B w_T.$$

Square both sides of itand apply conditional expectation operator.We have

$$\hat{\sigma}_n^2 = \frac{E \left\| X_T + x_0 + \int_0^T \hat{A}_n X_s \, ds | X^{(n)} \right\|_H^2}{E \| B w_T | X^{(n)} \|_H^2}.$$

Since$\hat{A}_n$is a consistent estimator of the operator A, $\hat{\sigma}_n^2$ is a consistent estimator of the true value of σ^2

References

[1] Athreya S. R., Bass R. F., Gordina M., Perkins E. A. Infinite Dimensional Stochastic Differential Equations of Ornstein-Uhlenbeck Type. Stochastic Processes and their Applications, 116, 381–406. 2006.

[2] Babilua P., Nadaraya E., Sokhadze G. On the limit properties of maximal likelihood estimators in a Hilbert space. Georgian Mathematical Journal. Vol. 22, Issue 2. 171-178, 2015

[3] Benguria R., Kac M. Quantum Langevin Equation. Physical Review Letters, 46, p. 1-4. 1981.

[4] Bishwal J. P. Parameter Estimation in Stochastic Differential Equations. Springer-Verlag Berlin Heidelberg. 2008.

[5] Brouste A., Iacus S. M. Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima R package. arXiv:1112.3777v1.

[6] Curtain R. Markov Processes Generated by Linear Stochastic Evolution Equations. Stochastics, 5, p. 135-165. 1981.

[7] Daletskiy Y. L., Belopol’skaya Y. I. Stochastic Equations and Differential Geometry (in Russian), VyshchaShkola, Kiev, 1989.

[8] Gao M. Free Ornstein-Uhlenbeck Processes. Journal of Mathematical Analysis and Applications, 322, p. 177-192. 2006.

[9] Gubeladze A., Sokhadze G. On the Maximum Likelihood Estimation of Stochastic Differential Equations. Proceedings of I. Vekua Institute of Applied Mathematics. Vol. 63. 1-7 2013.

[10] Kott T. Statistical Inference for Generalized Mean Reversion Processes.Diss. of Ruhr-niversitat Bochum. September. 2010.

[11] Liptser R.S., Shiryaev A.N. Statistics of Random Processes. Springer-Verlag. 1978.

[12] Maslowski B., Pospisil J. Ergodicity and Parameter Estimates for Infinite-Dimensional Fractional Ornstein-Uhlenbeck Process. Applied Mathematics and Optimization. 57(3), 401-429. 2008.

[13] McKeague I. W. Estimation for Infinite Dimensional Ornstein-Uhlenbeck Processes. Florida State University Statistics Report No. M674. 9 p. 1983.

[14] Statistical estimation of multivariate Ornstein–Uhlenbeck processes and applications to co-integration. Journal of Econometrics, Volume 172, Issue 2, p. 325–337. 2013.

[15] Stein E., Stein J. Stock Price Distribution with Stochastic Volatility: An Analytic Approach. Review of Financial Studies, 4(4), p. 727-752. 1991.

[16] Sokhadze G. On the Absolute Continuity of Smooth Measures. Theoryof Probability and Mathematical Statistics, 491996

[17] Valdivieso L., Schoutens W., Tuerlinckx F. Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type. Stat. Infer Stoch Process, 12:1, p. 1-19. 2009.

[18] Walsh J. B. A Stochastic Model for Neural Response. Adv. Appl. Probability, 13, p. 231-281. 1981.

[19] Wittig T. A Dynamical Theory of Generalized Ornstein-Uhlenbeck Processes. Ph. D. dissertation, Michigan State University. 1981.