Least Squares Estimators for Stochastic Differential Equations Driven by Small Lévy Noises

Hongwei Long*
Department of Mathematical Sciences, Florida Atlantic University, Boca Raton
Florida 33431-0991, USA-hlong@fau.edu

Chunhua Ma
School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071
PR China-mach@nankai.edu.cn

Yasutaka Shimizu
Department of Applied Mathematics, Waseda University, Tokyo, Japan-shimizu@waseda.jp

Abstract

We study parameter estimation for discretely observed stochastic differential equations driven by small Levy noises. There have been many applications of small noise asymptotics to mathematical finance and insurance. Using small noise models we can deal with both applications and statistical inference. We do not impose Lipschitz condition on the dispersion coefficient function and any moment condition on the driving Levy process, which greatly enhances the applicability of our results to many practical models. Under certain regularity conditions on the drift and dispersion functions, we obtain consistency and rate of convergence of the least squares estimator (LSE) of the drift parameter based on discrete observations. The asymptotic distribution of the LSE in our general framework is shown to be the convolution of a normal distribution and a distribution related to the jump part of the driving Levy process. We provide an example on a two-factor financial model driven by stable noises.

Keywords: parameter estimation; discrete observations; consistency of LSE; asymptotic distribution of LSE.

1. Introduction

Let $\mathbb{R}_0^r = \mathbb{R}^r \setminus 0$ and let $\mu(du)$ be a σ-finite measure on $\mathbb{R}_0^r$ satisfying $\int_{\mathbb{R}_0^r}(|u|^2 \wedge 1)\mu(du) < \infty$ with $|u| = (\sum_{i=1}^r u_i^2)^{1/2}$. Let $\{B_t = (B_t^1, \cdots, B_t^r) : t \geq 0\}$ be an r-dimensional standard Brownian motion and let $N(dt, du)$ be a Poisson random measure on $(0,\infty) \times \mathbb{R}_0^r$ with intensity measure $dt\mu(du)$. Suppose that $\{B_t\}$ and $\{N(dt, du)\}$ are independent of each other. Then an r-dimensional Lévy process $\{L_t\}$ can be given as

$$L_t = B_t + \int_0^t \int_{|u|>1} uN(ds, du) + \int_0^t \int_{|u|\leq 1} u\tilde{N}(ds, du), \tag{1}$$

where $\tilde{N}(ds, du) = N(ds, du) - ds\mu(du)$. Let us consider a family of d-dimensional stochastic processes defined as the solution of

$$dX_t^\varepsilon = b(X_t^\varepsilon, \theta)dt + \varepsilon\sigma(X_{t-}^\varepsilon)dL_t, \quad t \in [0, 1]; \quad X_0^\varepsilon = x, \tag{2}$$

where $\theta \in \Theta = \bar{\Theta}_0$, the closure of an open convex bounded subsets Θ_0 of $\mathbb{R}^p$. The function $b(x, \theta) = (b_k(x, \theta))_{k=1}^d$ is $\mathbb{R}^d$-valued and defined on $\mathbb{R}^d \times \bar{\Theta}$; the function $\sigma(x) = (\sigma_{kl}(x))_{d\times r}$ is defined on $\mathbb{R}^d$ and takes values on the space of matrices $\mathbb{R}^d \otimes \mathbb{R}^r$; the initial value $x \in \mathbb{R}^d$, and $\varepsilon > 0$ are known constant. A stochastic process of form (2) has long been used in the financial world and has been the fundamental tool in financial modeling. Examples of (2) include (i) the multivariate diffusion process defined by

$$dX_t^\varepsilon = b(X_t^\varepsilon, \theta)dt + \varepsilon\sigma(X_t^\varepsilon)dB_t,$$

(ii) the Vasicek model with jumps or the Lévy driven Ornstein-Uhlenbeck process defined by

$$dX_t^\varepsilon = \kappa(\beta - X_t^\varepsilon)dt + \varepsilon dL_t,$$

where κ and β are positive constant and L_t can be choosed to be a standard symmetric α-stable Lévy process (see Hu and Long (2009), Masuda (2010), and Fasen (2013)) or a (positive) Lévy subordinator (see Barndorff-Nielsen and Shephard (2001)); (iii) the Cox-Ingersoll-Ross (CIR) model driven by α-stable Lévy processes defined by

$$dX_t^\varepsilon = \kappa(\beta - X_t^\varepsilon)dt + \varepsilon \sqrt[\alpha]{X_{t-}^\varepsilon}\, dL_t,$$

where $\{L_t\}$ is a spectrally positive α-stable process with $1 < \alpha < 2$; see Fu and Li (2010) and Li and Ma (2015).

Assume that the only unknown quantity in (2) is the parameter θ. We denote the true value of the parameter by θ_0 and assume that $\theta \in \Theta$. Suppose that this process is observed at regularly spaced time points $\{t_k = k/n,\ k = 1, 2, \ldots, n\}$. The purpose of this paper is to study the least squares estimator for the true value θ_0 based on the sampling data $(X_{t_k})_{k=1}^n$ with small dispersion ε and large sample size n.

For statistical inference of jump-diffusions, we refer to Sørensen (2004), Masuda (2005), Shimizu and Yoshida (2006), Shimizu (2006), and Ogihara and Yoshida (2011). The small diffusion asymptotic $\varepsilon \to 0$ has been extensively studied with wide applications to real world problems; see Takahashi and Yoshida (2004) and Uchida and Yoshida (2004) for the applications to contingent claim pricing. For small diffusion asymptotics for parameter estimators in diffusion models, we refer to Long et al. (2013) for literature review.

Long (2009) studied the parameter estimation problem for discretely observed one-dimensional Ornstein-Uhlenbeck processes with small Lévy noises. In that paper, the drift function is linear in both x and θ ($(b(x, \theta) = -\theta x)$, the driving Lévy process is $L_t = aB_t + bZ_t$, where a and b are known constants, $(B_t, t \geq 0)$ is the standard Brownian motion and Z_t is a α-stable Lévy motion independent of $(B_t, t \geq 0)$. The consistency and rate of convergence of the least squares estimator are established. The asymptotic distribution of the LSE is shown to be the convolution of a normal distribution and a stable distribution. Ma (2010) extended the results of Long (2009) to the case when the driving noise is a general Lévy process. Ma and Yang (2014) also studied the parameter estimation problem for discretely observed CIR model with small Lévy noises. Recently Long et al. (2013) discussed the statistical estimation of the drift parameter for a class of Lévy driven SDEs with drift function $b(x, \theta)$ to be nonlinear in both x and θ. However they only consider the simple version of (2) with constant coefficient in the Lévy jump term, i.e., $\sigma(\cdot)$ is a constant matrix.

In this paper, we consider a more general class of stochastic processes with small Lévy noises defined by (2). We are interested in estimating the drift parameter θ based on discrete observations $\{X_{t_i}\}_{i=1}^n$ when $\varepsilon \to 0$ and $n \to \infty$. We shall use the least squares method to obtain an asymptotically consistent estimator and discuss the asymptotic behavior of the LSE. We show that the asymptotic distribution of the LSE is the convolution of a normal distribution and a distribution related to the jump part of the driving Lévy process.

The paper is organized as follows. In Section 2, we propose our least squares estimator in our general framework and state the main results, which provide consistency and asymptotic behavior of the LSE. In Section 3, we present a two-factor financial model as an example.

2. Consistency and asymptotic distribution of least squares estimators

We start this section by presenting the necessary conditions and the construction of the estimator. Let $X^0 = (X_t^0, t \geq 0)$ be the solution to the underlying ordinary differential equation (ODE) under the true value of the drift parameter:

$$dX_t^0 = b(X_t^0, \theta_0)dt, \quad X_0^0 = x_0.$$

For a multi-index $m = (m_1, \ldots, m_k)$, we define a derivative operator in $z \in \mathbb{R}^k$ as $\partial_z^m := \partial_{z_1}^{m_1} \cdots \partial_{z_k}^{m_k}$, where $\partial_{z_i}^{m_i} := \partial^{m_i}/\partial z_i^{m_i}$. Let $C^{k,l}(\mathbb{R}^d \times \Theta; \mathbb{R}^q)$ be the space of all functions $f : \mathbb{R}^d \times \Theta \to \mathbb{R}^q$ which is k and l times continuously differentiable with respect to x and θ, respectively. Moreover $C_\uparrow^{k,l}(\mathbb{R}^d \times \Theta; \mathbb{R}^q)$ is a class of $f \in C^{k,l}(\mathbb{R}^d \times \Theta; \mathbb{R}^q)$ satisfying that $\sup_{\theta\in\Theta} |\partial_\theta^\alpha \partial_x^\beta f(x, \theta)| \leq C(1+|x|)^\lambda$ for universal positive constants C and

λ, where $\alpha = (\alpha_1, \ldots, \alpha_p)$ and $\beta = (\beta_1, \ldots, \beta_d)$ are multi-indices with $0 \le \sum_{i=1}^p \alpha_i \le l$ and $0 \le \sum_{i=1}^d \beta_i \le k$, respectively.

Now let us introduce the following set of assumptions.

(A1) For all $\varepsilon > 0$, the SDE (2) admits a unique strong solution X^ε on some probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$.

(A2) There exists a constant $K > 0$ such that

$$|b(x,\theta) - b(y,\theta)| \le K|x-y|; \quad |b(x,\theta)| + |\sigma(x)| \le K(1+|x|)$$

for each $x, y \in \mathbb{R}$ and $\theta \in \bar{\Theta}$.

(A3) $b(\cdot,\cdot) \in C_{\uparrow}^{2,3}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$.

(A4) σ is continuous and there exists some open convex subset $\mathcal{U}$ of $\mathbb{R}^d$ such that $X_t^0 \in \mathcal{U}$ for all $t \in [0,1]$, and σ is smooth on $\mathcal{U}$. Moreover $\sigma\sigma^T(x)$ is invertible on $\mathcal{U}$.

(A5) $\theta \neq \theta_0 \;\Leftrightarrow\; b(X_t^0, \theta) \neq b(X_t^0, \theta_0)$ for at least one value of $t \in [0,1]$.

(B) $\varepsilon = \varepsilon_n \to 0$ and $n\varepsilon \to \infty$ as $n \to \infty$.

Consider the following contrast function

$$\Psi_{n,\varepsilon}(\theta) = \Big(\sum_{k=1}^n \varepsilon^{-2} n P_k^T(\theta)\Lambda_{k-1}^{-1} P_k(\theta)\Big) 1_{\{Z>0\}}, \tag{3}$$

where

$$\begin{aligned} P_k(\theta) &= X_{t_k} - X_{t_{k-1}} - \frac{1}{n} b(X_{t_{k-1}}, \theta), \\ \Lambda_k &= [\sigma\sigma^T](X_{t_k}) \end{aligned}$$

and the random variable $Z = \inf_{k=0,\cdots,n-1} \det \Lambda$ is introduced to ensure that $\Psi_{n,\varepsilon}$ is well defined. Let $\hat{\theta}_{n,\varepsilon}$ be a minimum contrast estimator, i.e. a family of random variables satisfying

$$\hat{\theta}_{n,\varepsilon} := \arg\min_{\theta\in\Theta} \Psi_{n,\varepsilon}(\theta).$$

Since minimizing $\Psi_{n,\varepsilon}(\theta)$ is equivalent to minimizing

$$\Phi_{n,\varepsilon}(\theta) := \varepsilon^2(\Psi_{n,\varepsilon}(\theta) - \Psi_{n,\varepsilon}(\theta_0)),$$

we may write the LSE as

$$\hat{\theta}_{n,\varepsilon} = \arg\min_{\theta\in\Theta} \Phi_{n,\varepsilon}(\theta). \tag{4}$$

Note that in the case when $b(x,\theta)$ is nonlinear in both x and θ, it is generally very difficult or impossible to obtain an explicit formula for the least squares estimator $\hat{\theta}_{n,\varepsilon}$. However, we can use some nice criteria in statistical inference (see Chapter 5 of Van der Vaart (1998)) to establish the consistency of the LSE as well as its asymptotic behaviors (asymptotic distribution and rate of convergence). The main results of this paper are the following asymptotics of the LSE $\hat{\theta}_{n,\varepsilon}$ with high frequency ($n \to \infty$) and small dispersion ($\varepsilon \to 0$).

Theorem 1. *Under conditions (A1)–(A5), we have*

$$\hat{\theta}_{n,\varepsilon} \xrightarrow{P_{\theta_0}} \theta_0,$$

as $\varepsilon \to 0$ *and* $n \to \infty$.

Introduce the matrix $I(\theta_0) = \left(I^{ij}(\theta_0)\right)_{1\le i,j\le p}$, where

$$I^{ij}(\theta) = \int_0^1 (\partial_{\theta_i} b)^T (X_s^0, \theta)[\sigma\sigma^T]^{-1}(X_s^0)\partial_{\theta_j} b(X_s^0, \theta)ds. \tag{5}$$

Theorem 2. *Assume (A1)-(A5), (B) and that $\theta_0 \in \Theta$ with the matrix $I(\theta_0)$ given in (5) being positive definite. Then*

$$\varepsilon^{-1}(\hat{\theta}_{n,\varepsilon} - \theta_0) \xrightarrow{P_{\theta_0}} (I(\theta_0))^{-1} \left(\int_0^1 (\partial_{\theta_i} b)^T (X_t^0, \theta_0)[\sigma\sigma^T]^{-1}(X_t^0)\sigma(X_t^0)dL_t\right)^T_{1\le i\le p}, \tag{6}$$

as $\varepsilon \to 0$ and $n \to \infty$.

Remark. When $\sigma(x)$ is a constant matrix $\sigma = (\sigma_{ij})_{d\times r}$, our main results Theorems 1 and 2 reduce to Theorems 2.1 and 2.2 of Long et al. (2013). The limiting distribution in (6) is the convolution of a normal distribution and a distribution related to the jump part of the driving Lévy process L_t. In particular, if the Lévy process L_t is a linear combination of Brownian motion and α-stable motion, then the limiting distribution is the convolution of a normal distribution and a stable distribution.

3. An example: A two-factor model driven by stable noises

An affine two-factor model $X = (Y, R)$ is defined by

$$dY_t = (R_t + \mu_1)dt + \varepsilon dB_t^1, \quad Y_0 = y_0 \in \mathbb{R},$$

$$dR_t = \mu_2(m - R_t)dt + \varepsilon\sqrt{R_t}\left(\rho dB_t^1 + (1-\rho^2)^{1/2}dB_t^2\right), \quad R_0 = r_0 > 0,$$

where $B = (B^1, B^2)$ is two-dimensional Brownian, the unknown parameter $\theta = (\mu_1, \mu_2, m) \in \mathbb{R}\times(0,\infty)^2$ and the known constant $\rho \in (0,1)$. Intuitively, Y is the log price of some asset and R represents the short term interest rate, which follows the classical Cox-Ingersoll-Ross model; see Longstaff and Schwartz (1995). The parameter ρ allows correlation between innovation terms of the two factors. The estimator $\hat{\theta}_{n,\varepsilon}$ is given by (4). Gloter and Sørensen (2009) explored the behavior of $\hat{\theta}_{n,\varepsilon}$ for finite samples using Monte Carlo simulations.

However, real rates and asset price do not evolve continuously in time. Typical stylized facts of financial time series as asset returns, exchange rates and interest rates are jumps and a heavy tailed distribution in the sense that the second moment is infinite. These characteristics were already noticed in the 60s by the influential works of Fama (1965) and Mandelbrot and Taylor (1967). Thus, α-stable distributions as generalization of a Gaussian distribution have often been discussed as more realistic models for asset returns than the usual normal distribution. Inspired by these ideas, it is natural to suggest the following model defined by

$$dY_t^\varepsilon = (R_t^\varepsilon + \mu_1)dt + \varepsilon dL_t^1, \quad Y_0 = y_0 \in \mathbb{R}, \tag{7}$$

$$dR_t^\varepsilon = \mu_2(m - R_t^\varepsilon)dt + \varepsilon\sqrt[\alpha]{R_t^\varepsilon}\left(\rho dL_t^1 + (1-\rho^\alpha)^{1/\alpha}dL_t^2\right), \quad R_0 = r_0 > 0, \tag{8}$$

where $\{L_t^1\}$ and $\{L_t^2\}$ are independent spectrally positive α-stable process with $1 < \alpha < 2$. More precisely, the corresponding Lévy measure is given by

$$\lambda(dz) = \frac{1_{\{z>0\}}dz}{\Gamma(-\alpha)z^{1+\alpha}}.$$

In the above model, the second component is called the stable CIR model. We refer to Fu and Li (2010) for the pathwise uniqueness of the positive strong solution of SDEs. The asymptotic estimation problem was studied by Li and Ma (2015).

Note that $X_t^\varepsilon = (Y_t^\varepsilon, R_t^\varepsilon)$ is a two-dimensional stochastic process defined in (2) with drift function $b(x,\theta) = (r+\mu_1, \mu_2(m-r))$ and dispersion matrix function

$$\sigma(x) = \begin{pmatrix} 1 & 0 \\ \rho r^{\frac{1}{\alpha}} & (1-\rho^\alpha)^{\frac{1}{\alpha}} r^{\frac{1}{\alpha}} \end{pmatrix},$$

where $x = (y,r) \in \mathbb{R}^2$ and $\theta = (\theta_1,\theta_2,\theta_3) = (\mu_1,\mu_2,m)$. We assume that the true value of θ is $\theta_0 = (\mu_1^0, \mu_2^0, m^0) \in \Theta_0 \subset \mathbb{R} \times (0,\infty)^2$. Then $X_t^0 = (Y_t^0, R_t^0)$ satisfies the following differential equation

$$dX_t^0 = b(X_t^0, \theta_0)dt, \quad X_0^0 = x_0 = (y_0, r_0).$$

The explicit solution is given by

$$Y_t^0 = y_0 + (m^0 - \mu_1^0)t + (r_0 - m^0)(1 - e^{-\mu_2^0 t})/\mu_2^0$$

and

$$R_t^0 = m^0 + (r_0 - m^0)e^{-\mu_2^0 t}.$$

We can use standard calculus method to find the LSE $\hat{\theta}_{n,\varepsilon} = (\hat{\theta}_{n,\varepsilon,1}, \hat{\theta}_{n,\varepsilon,2}, \hat{\theta}_{n,\varepsilon,3})^T$ by solving the system of equations

$$\frac{\partial \Psi_{n,\varepsilon}(\theta)}{\partial \theta_i} = 0, \quad i = 1,2,3.$$

Since $\Psi_{n,\varepsilon}(\theta)$ is a nonlinear function of θ, there is no explicit solution to the aforementioned system of equations. But we can solve them numerically. Furthermore we find that

$$\partial_{\theta_1} b(x,\theta) = (1,0)^T, \quad \partial_{\theta_2} b(x,\theta) = (0, m-r)^T, \quad \partial_{\theta_3} b(x,\theta) = (0,\mu_2)^T,$$

and

$$(\sigma\sigma^T)^{-1}(x) = (1-\rho^\alpha)^{-\frac{2}{\alpha}} r^{-\frac{2}{\alpha}} \begin{pmatrix} \rho^2 r^{\frac{2}{\alpha}} + (1-\rho^\alpha)^{\frac{2}{\alpha}} r^{\frac{2}{\alpha}} & -\rho r^{\frac{1}{\alpha}} \\ -\rho r^{\frac{1}{\alpha}} & 1 \end{pmatrix}.$$

It is easy and straightforward to compute the information matrix $I(\theta_0) = (I^{ij}(\theta_0))_{1\le i,j\le 3}$ in (5) and the limiting random vector in (6) of Theorem 2. We shall omit the details here.

References

Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein- Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol*, **63**, 167-241.

Fama, E.F. (1965). The behavior of stock market prices. *J. Business*, **38**, 34-105.

Fasen, V. (2013). Statistical estimation of multivariate Ornstein-Uhlenbeck processes and applications to co-integration. *J. Econometrics* **172**, 325-337.

Fu, Z.F. and Li, Z.H. (2010). Stochastic equations of non-negative processes with jumps. *Stochastic Processes and their Applications* **120**, 306-330.

Gloter, A. and Sørensen, M. (2009). Estimation for stochastic differential equations with a small diffusion coefficient. *Stochastic Process. Appl.* **119**, 679-699.

Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein-Uhlenbeck processes driven by α-stable *Stochastic Process. Appl.* **119**, 2465-2480.

Kunitomo, N. and A. Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims. *Math. Finance* **11**, 117-151.

Li, Z.H. and Ma, C.H. (2015). Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. To appear in *Stochastic Process. Appl.* .

Long, H. (2009). Least squares estimator for discretely observed Ornstein-Uhlenbeck processes with small Lévy noises. *Statistics and Probability Letters* **79**, 2076-2085.

Long, H., Shimizu, Y. and Sun, W. (2013). Least squares estimators for discretely observed stochastic processes driven by small Lévy noises. *Journal of Multivariate Analysis* **116**, 422-439.

Longstaff, F. and Schwartz, E. (1995). A simple approach to valuing risky fixed and floating rate debt. *J. Finance* **1**, 789-819.

Ma, C.H. (2010). A note on "Least squares estimator for discretely observed Ornstein-Uhlenbeck processes with small Lévy noises". *Statistics and Probability Letters* **80**, 1528-1531.

Ma, C.H. and Yang, X. (2014). Parameters estimation for the CIR model driven by α-stable noises. *Statistics and Probability Letters* **94**, 1-11.

Mandelbrot, B. and Taylor, H.M. (1967). On the distribution of stock price differences. *Oper. Res.*, **15**, 1057-1062.

Masuda, H. (2005). Simple estimators for parametric Markovian trend of ergodic processes based on sampled data. *J. Japan Statist. Soc.* **35**, 147-170.

Masuda, H. (2010). Approximate self-weighted LAD estimation of discretely observed ergodic Ornstein-Uhlenbeck processes. *Electronic Journal of Statistics* **4**, 525-565.

Ogihara, T. and Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.* **14**, 189-229.

Shimizu, Y. (2006). M-estimation for discretely observed ergodic diffusion processes with infinite jumps. *Stat. Inference Stoch. Process.* **9**, 179-225.

Shimizu, Y. and Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.* **9**, 227-277.

Sørensen, H. (2004). Parameter inference for diffusion processes observed at discrete points in time: a survey. *Internat. Statist. Rev.* **72**, 337-354.

Takahashi, A. and Yoshida, N. (2004). An asymptotic expansion scheme for optimal investment problems. *Stat. Inference Stoch. Process.* **7**, 153-188.

Uchida, M. and Yoshida, N. (2004). Asymptotic expansion for small diffusions applied to option pricing. *Stat. Inference Stoch. Process.* **7**, 189-223.

van der Vaart, A.W. (1998). Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics **3**, Cambridge University Press.