Least Squares Estimator for Ornstein-Uhlenbeck Processes Driven by Stable Lévy Motions

Hongwei Long

Joint work with Yaozhong Hu (University of Kansas)

Florida Atlantic University

2009 Barrett Lectures at the University of Tennessee

April 17-18, 2009

Outline

1. Formulation of LSE for Generalized O-U Processes

2. Strong Consistency of LSE

3. Asymptotic Distribution of LSE

4. Simulations

1. Formulation of LSE for Generalized O-U Processes

- Generalized Ornstein-Uhlenbeck processes $\{X_t\}$:

$$dX_t = -\theta_0 X_t dt + dZ_t, \quad X_0 = x. \tag{1}$$

on some probability space $(\Omega, \mathcal{F}, P)$, where $\theta_0 > 0$ is an unknown parameter.

- $\{Z_t, t \geq 0\}$ is a standard symmetric α-stable Lévy motion ($\alpha \in (1,2)$) so that Z_1 has a symmetric α-stable distribution $S_\alpha(1,0,0)$ with c.f.

$$u \mapsto \exp(-|u|^\alpha).$$

- **Remark**: (1) In general, a random variable η is said to have a stable distribution with index of stability $\alpha \in (0,2]$, scale parameter $\sigma \in (0,\infty)$, skewness parameter $\beta \in [-1,1]$, and location parameter $\mu \in (-\infty,\infty)$ if it has characteristic function of the following form:

$$\phi_\eta(u) = E\exp\{iu\eta\} = \begin{cases} \exp\left\{-\sigma^\alpha |u|^\alpha \left(1 - i\beta \text{sgn}(u) \tan\frac{\alpha\pi}{2}\right) + i\mu u\right\}, & \text{if } \alpha \neq 1, \\ \exp\left\{-\sigma|u| \left(1 + i\beta\frac{2}{\pi}\text{sgn}(u)\log|u|\right) + i\mu u\right\}, & \text{if } \alpha = 1. \end{cases}$$

We denote $\eta \sim S_\alpha(\sigma,\beta,\mu)$.
(2) Stable distributions are introduced by Paul Lévy (1925). Applications in modeling financial data: Mandelbrot (1963), Fama (1965), Officer (1972), Mittnik and Rachev (2001), Rachev (2003), Nolan (2005).
(3) Stable distributions are heavy-tailed and for $\eta \sim S_\alpha(\sigma,\beta,\mu)$, $E(|\eta|^q) = \infty$ ($q \geq \alpha$).

(4) Brownian motion is almost surely continuous, while stable Lévy motion is a pure jump process (more appropriate to model extreme events).

- Discrete observations: $(X_{t_i})_{i=0}^n$ with $t_i = ih$.

- Our goal: (i) Construct LSE of θ_0 based on discrete observations; (ii) Study the high frequency ($h \to 0$) asymptotics of the LSE.

- Historical Comments:
  Least squares method is a classical method: parameter estimation in regression models, time series models, diffusion models, and jump-diffusion models. Several related references: Dorogovcev (1976), Le Breton (1976), Prakasa Rao (1983), Kasonga (1988), Masuda (2005), Shimizu (2006), Shimizu and Yoshida (2006).

- Our model is simple but can not be covered by those models mentioned above. This is due to the infinite variance property of α-stable Levy motions.

- Our model has been used in physics, e.g., in geophysical turbulence and climate dynamical changes [see Ditlevsen (1999), Schertzer et al (2001)].

- The SDE (1) can be approximated by Euler scheme [see Jacod (2004), Jacod et al (2005)]

$$X_{t_i} = X_{t_{i-1}} - \theta X_{t_{i-1}} \Delta t_i + \Delta Z_{t_i},$$

where $\Delta t_i = t_i - t_{i-1} = h$ and $\Delta Z_{t_i} = Z_{t_i} - Z_{t_{i-1}}$.

- We may regard $X_{t_{i-1}} - \theta X_{t_{i-1}} \Delta t_i$ as a prediction of X_{t_i}.

- Contrast function:

$$\rho_n(\theta) = \sum_{i=1}^{n} \left| X_{t_i} - X_{t_{i-1}} + \theta X_{t_{i-1}} \cdot \Delta t_{i-1} \right|^2. \qquad (2)$$

- LSE of θ_0:

$$\hat{\theta}_n = -\frac{\sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}}) X_{t_{i-1}}}{h \sum_{i=1}^{n} X_{t_{i-1}}^2}. \qquad (3)$$

- An equivalent expression:

$$\hat{\theta}_n = \frac{1 - e^{-\theta_0 h}}{h} - \frac{\sum_{i=1}^{n} X_{t_{i-1}} \cdot \int_{t_{i-1}}^{t_i} e^{-\theta_0 (t_i - s)} dZ_s}{h \sum_{i=1}^{n} X_{t_{i-1}}^2}. \qquad (4)$$

- Strong consistency and asymptotic distribution of $\hat{\theta}_n$.

2. Strong Consistency of the LSE

2.1. Stable stochastic integrals

- $L^\alpha_{a.s}$ is the class of real-valued predictable processes F such that for every $T > 0$, $\int_0^T |F(t,\omega)|^\alpha dt < \infty$ a.s.

- First define stochastic integral w.r.t Z_t for simple predictable process and then for general predictable process in $L^\alpha_{a.s.}$: $\int_0^t F(s)dZ_s$. [see Rosinski and Woyczynski (1986), Kallenberg (1992)]

- **Random time change**:
Let $F \in L^\alpha_{a.s}$ such that $\tau(u) = \int_0^u |F|^\alpha dt \to \infty$ as $u \to \infty$. Then, there exists an independent process Z' (with same distribution as Z) such that
$$\int_0^t F dZ = Z'(\tau(t)).$$

- **Limit Theorem**: Let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be increasing. If $\tau(u) \to \infty$ a.s. as $u \to \infty$, then
$$\limsup_{t\to\infty} \left| \int_0^t F dZ \right| / \varphi(\tau(t)) = 0 \text{ or } = \infty \quad a.s$$
according to
$$\int_1^\infty \varphi(t)^{-\alpha} dt < \infty \text{ or } = \infty.$$

2.2. Strong Consistency

- X_t is ergodic ($\theta_0 > 0$) and $X_t \Rightarrow X_\infty$ as $t \to \infty$, where $X_\infty = \int_0^\infty e^{-\theta_0 t} dZ_t$ [see Sato (1999)].
- **Theorem 1**: Assume that $h \to 0$ and $t_n = nh \to \infty$. Then,

$$\hat{\theta}_n \to \theta_0 \text{ almost surely as } n \to \infty. \quad (5)$$

- **Proof ideas**:

  – Let $\phi_n(t) = \sum_{i=1}^n X_{t_{i-1}} e^{-\theta(t_i - t)} 1_{(t_{i-1}, t_i]}(t)$ and $\tau_n(t_n) = \int_0^{t_n} |\phi_n(t)|^\alpha dt$.

  – Rewrite:

$$\hat{\theta}_n = \frac{1 - e^{-\theta_0 h}}{h} - \frac{\int_0^{t_n} \phi_n(t) dZ_t}{\tau_n(t_n)} \cdot \frac{\tau_n(t_n)}{h \sum_{i=1}^n X_{t_{i-1}}^2}. \quad (6)$$

  – Ergodic theorem:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n |X_{t_{i-1}}|^2 = E[X_\infty^2] = \infty \quad \text{a.s.}$$

  – Ergodic theorem plus limit theorem imply the almost sure convergence.

3. Asymptotic Distribution of LSE

3.1. Main Result

- **(A1)**: As $n \to \infty$, $h \to 0$, $nh^{1+\alpha}/\log n \to 0$, $nh^{2\alpha-1}\log n \to \infty$, and $nh^{2-\alpha/2+\rho} \to \infty$ for some $\rho > 0$.

- Let $C_\alpha = \left(\int_0^\infty x^{-\alpha}\sin(x)dx\right)^{-1} = [\Gamma(1-\alpha)\cos(\pi\alpha/2)]^{-1}$, $\sigma_1 = C_{\alpha/2}^{-2/\alpha}$, and $\sigma_2 = C_\alpha^{-1/\alpha}$.

- **Theorem 2**: Under the condition (A1), we have

$$\left(\frac{n}{\log n}\right)^{1/\alpha} h^{1/\alpha}(\hat{\theta}_n - \theta_0) \Rightarrow \frac{2\theta_0(\alpha\theta_0)^{-1/\alpha}\tilde{Y}}{Y_0}, \qquad (7)$$

  where Y_0 and $\tilde{Y}$ are independent stable random variables, Y_0 is positive $\alpha/2$-stable with distribution $S_{\alpha/2}(\sigma_1, 1, 0)$, $\tilde{Y}$ is symmetric α-stable with distribution $S_\alpha(\sigma_2, 0, 0)$.

- **Remark**: The rate of convergence $\left(\frac{\log n}{nh}\right)^{1/\alpha}$ is considerably faster than the rate $(nh)^{-1/2}$ in the classical Brownian motion case.

3.2. Some Preliminaries

- We have (denoting X_{t_i} by X_i)

$$\begin{aligned}
&\left(\frac{n}{\log n}\right)^{1/\alpha} h^{1/\alpha}(\hat{\theta}_n - \theta_0) \\
= &\left(\frac{n}{\log n}\right)^{1/\alpha} h^{1/\alpha}[h^{-1}(1 - e^{-\theta_0 h}) - \theta_0] \\
&- \frac{(n \log n)^{-1/\alpha} h^{-1/\alpha} \sum_{i=1}^{n} X_{i-1} \int_{t_{i-1}}^{t_i} e^{-\theta_0(t_i - s)} dZ_s}{n^{-2/\alpha} h^{1-\frac{2}{\alpha}} \sum_{i=1}^{n} X_{i-1}^2} \\
:= &\ \Lambda_n - \frac{\Phi_1(n)}{\Phi_2(n)}. \qquad (8)
\end{aligned}$$

- Note that

$$X_i = e^{-\theta_0 i h} X_0 + \sum_{k=1}^{i} e^{-\theta_0 i h} \int_{t_{k-1}}^{t_k} e^{\theta_0 s} dZ_s. \qquad (9)$$

- Let $V_{k-1} = \displaystyle\int_{t_{k-1}}^{t_k} e^{\theta_0 s} dZ_s$. and $\tau_{k-1} = \int_{t_{k-1}}^{t_k} |e^{\theta_0 s}|^\alpha ds$, $U_{k-1} = V_{k-1}/\tau_{k-1}^{\frac{1}{\alpha}}$. By scaling property of stable distribution, we know that $U_0, U_1, U_2, \ldots$ are i.i.d. with the same distribution $S_\alpha(1, 0, 0)$.

- X_i can be represented as

$$
\begin{aligned}
X_i &= e^{-\theta_0 ih} X_0 + \left(\frac{e^{\alpha\theta_0 h}-1}{\alpha\theta_0}\right)^{1/\alpha} \sum_{k=1}^{i} e^{-\theta_0(i-k+1)h} U_{k-1} \\
&= c_{i,h} X_0 + \gamma_h \sum_{j=1}^{i} c_{j,h} U_{i-j}, \qquad (10)
\end{aligned}
$$

where $c_{i,h} = e^{-\theta_0 ih}$ and $\gamma_h = \left(\frac{e^{\alpha\theta_0 h}-1}{\alpha\theta_0}\right)^{1/\alpha}$.

- **Remark**: For symmmetric α-stable random variable $U_1 \sim S_\alpha(1,0,0)$, we have

$$
\lim_{x\to\infty} x^\alpha P(U_1 > x) = C_\alpha/2 \quad \text{and} \quad \lim_{x\to\infty} x^\alpha P(U_1 < -x) = C_\alpha/2.
$$

- We introduce some scaling quantities:

$$
\begin{aligned}
a_n &= \inf\{x : P\left(|U_1| > x\right) \le n^{-1}\} \quad \text{and} \\
\tilde{a}_n &= \inf\{x : P\left(|U_0 U_1| > x\right) \le n^{-1}\}.
\end{aligned}
$$

- Thanks to the tail behavior, we may take

$$
a_n = (C_\alpha n)^{\frac{1}{\alpha}} \quad \text{and} \quad \tilde{a}_n = C_\alpha^{\frac{2}{\alpha}} (n \log n)^{\frac{1}{\alpha}}.
$$

- **Lemma 1** [Davis and Resnick (1986)]: Let $\{U_i\}_{i=0}^{\infty}$ be i.i.d. with the same stable distribution $S_\alpha(1,0,0)$. Then, for a_n and $\tilde{a}_n$ defined as above, we have for $m \in \mathbb{N}$

$$\left(a_n^{-2} \sum_{i=1}^{n} U_i^2, \tilde{a}_n^{-1} \sum_{i=1}^{n} U_i U_{i+1}, \cdots, \tilde{a}_n^{-1} \sum_{i=1}^{n} U_i U_{i+m} \right) \Rightarrow (Y_0, Y_1, \cdots, Y_m), \tag{11}$$

where $Y_0, Y_1, \cdots, Y_m$ are independent stable random variables, Y_0 is positive $\alpha/2$-stable with distribution $S_{\alpha/2}(\sigma_1, 1, 0)$, $Y_1, \cdots, Y_m$ are i.i.d. symmetric α-stable with distribution $S_\alpha(\sigma_2, 0, 0)$.

3.3. Proof Ideas for Theorem 2

- The asymptotic behavior of $\left(\frac{n}{\log n}\right)^{1/\alpha} h^{1/\alpha}(\hat{\theta}_n - \theta_0)$ will be determined by the asymptotic behavior of Λ_n, $\Phi_1(n)$ and $\Phi_2(n)$.
- It is obvious that $\Lambda_n \to 0$ as $n \to \infty$ under condition (A1).
- **Proposition 1**: Under condition (A1), we have

$$\Phi_2(n) \Rightarrow \frac{C_\alpha^{2/\alpha} Y_0}{2\theta_0}, \tag{12}$$

  where Y_0 is a random variable with positively skewed stable distribution $S_{\alpha/2}(\sigma_1, 1, 0)$ as specified in Theorem 2.
- Proof idea of Proposition 1:

  – Decomposition of $\Phi_2(n)$:

$$\begin{aligned} \Phi_2(n) &= n^{-2/\alpha} h^{1-2/\alpha} \sum_{i=1}^{n} X_{i-1}^2 \\ &= n^{-2/\alpha} h^{1-2/\alpha} X_0^2 + n^{-2/\alpha} h^{1-2/\alpha} \sum_{i=1}^{n-1} X_i^2 \\ &:= \Phi_{2,1}(n) + \Phi_{2,2}(n). \end{aligned} \tag{13}$$

and

$$
\begin{aligned}
&\Phi_{2,2}(n) \\
= \;& n^{-2/\alpha}h^{1-2/\alpha}\sum_{i=1}^{n-1}\left[c_{i,h}X_0+\gamma_h\sum_{j=1}^{i}c_{j,h}U_{i-j}\right]^2 \\
= \;& n^{-2/\alpha}h^{1-2/\alpha}\sum_{i=1}^{n-1}c_{i,h}^2X_0^2 \\
& +2n^{-2/\alpha}h^{1-2/\alpha}\gamma_h\sum_{i=1}^{n-1}c_{i,h}X_0\sum_{j=1}^{i}c_{j,h}U_{i-j} \\
& +n^{-2/\alpha}h^{1-2/\alpha}\gamma_h^2\sum_{i=1}^{n-1}\sum_{j=1}^{i}c_{j,h}^2U_{i-j}^2 \qquad (14) \\
& +n^{-2/\alpha}h^{1-2/\alpha}\gamma_h^2\sum_{i=1}^{n-1}\sum_{j=1}^{i}\sum_{k=1,k\neq j}^{i}c_{j,h}c_{k,h}U_{i-j}U_{i-k}.
\end{aligned}
$$

– By using some basic inequalities, truncation techniques, and Karamata's theorem, we can prove that

$$
\Phi_2(n)-n^{-2/\alpha}h^{1-2/\alpha}\gamma_h^2\sum_{i=1}^{n-1}\sum_{j=1}^{i}c_{j,h}^2U_{i-j}^2 \to_P 0. \qquad (15)
$$

– By using some basic inequalities and Lemma 1, we can show that

$$n^{-2/\alpha}h^{(\alpha-2)/\alpha}\gamma_h^2\sum_{i=1}^{n-1}\sum_{j=1}^{i}c_{j,h}^2U_{i-j}^2 \Rightarrow \frac{C_\alpha^{2/\alpha}Y_0}{2\theta_0}, \quad (16)$$

where Y_0 is a random variable with positively skewed stable distribution $S_{\alpha/2}(\sigma_1, 1, 0)$ as specified in Theorem 2.

- **Proposition 2**: Under the condition (A1), we have

$$\Phi_1(n) \Rightarrow C_\alpha^{2/\alpha}Y \quad (17)$$

as $n \to \infty$, where Y is a random variable with stable distribution $S_\alpha((\alpha\theta_0)^{-1/\alpha}\sigma_2, 0, 0)$.

- Proof idea is similar to that of Proposition 1, i.e., using some decomposition technique, basic inequalities, Lemma 1 and Skorohod's representation theorem.
- Finally, the proof of **Theorem 2** is completed by combining Propositions 1 and 2.

4. Simulations

- We apply our estimator to the following stochastic differential equation:

$$dX_t = -\theta_0 X_t dt + dZ_t\,, \quad X_0 = 1\,,$$

where $\theta_0 = 2$ and Z_t is a stable process with index $\alpha = 1.8$. We simulate the process on the interval $[0, T]$ with $T = 200$.

- We plot $\hat{\theta}_T = \hat{\theta}_n$ (where $T = nh$) as a function of T for $h = 0.05$ (Figure 1) and $h = 0.01$ (Figure 2).

- The following table describes $\theta(25), \cdots, \theta(200)$ for different choice of h.

$T =$	25	50	75	100	125	150	175	200
$h = 0.1$	1.3304	1.4764	1.4409	1.4593	1.5903	1.6104	1.4234	1.3817
$h = 0.05$	1.7580	1.7326	1.7333	1.7051	1.4785	1.5509	1.5431	1.4053
h=0.033	2.3924	2.4681	2.4417	2.2373	2.2200	2.2094	2.1755	2.2002
h=0.025	1.8936	1.8320	1.8429	1.8323	1.8505	1.7615	1.7981	1.8109
h=0.02	2.0268	2.1199	2.1277	2.1162	2.1311	2.1529	2.1794	2.1734
h=0.0167	2.4107	2.5096	2.5188	2.4964	2.5093	2.5160	2.5260	2.5398
h=0.0143	2.2751	2.2514	2.1568	2.2191	2.0080	1.9516	1.6916	1.6516
h=0.0125	2.1116	1.9001	1.9310	1.9313	1.9313	1.9454	1.9353	1.9282

- We need to let both T goes to infinity and h goes to 0 to have the convergence of $\hat{\theta}_T$ to θ_0.

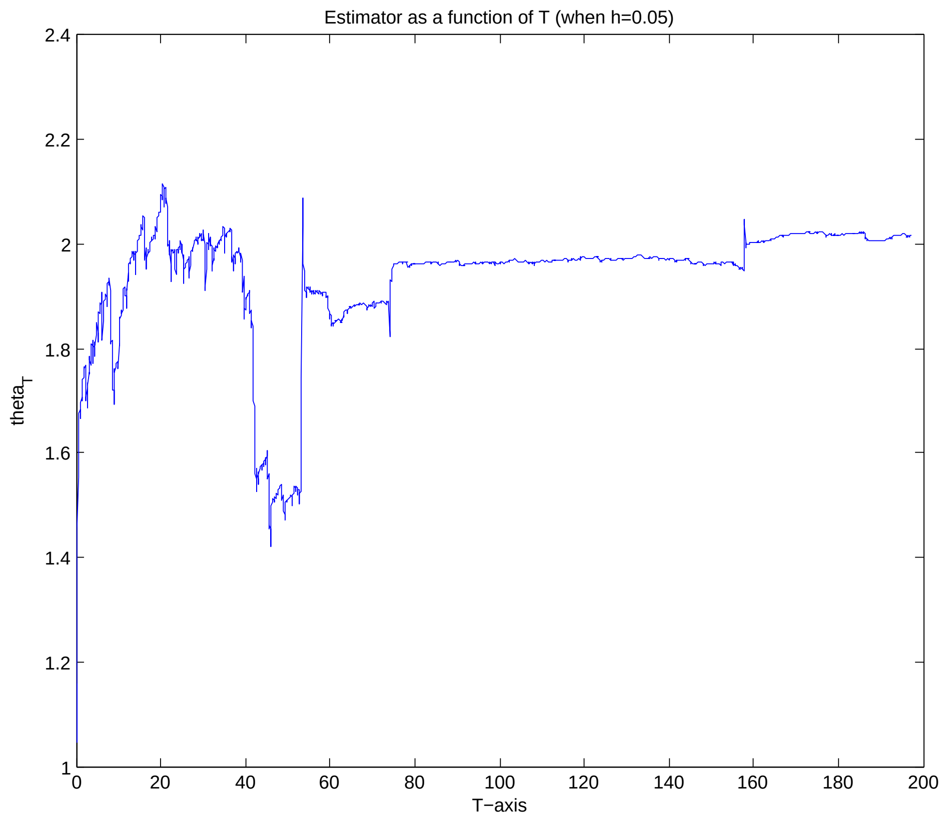

Figure 1: LSE for $\theta_0 = 2$ when $h = 0.05$

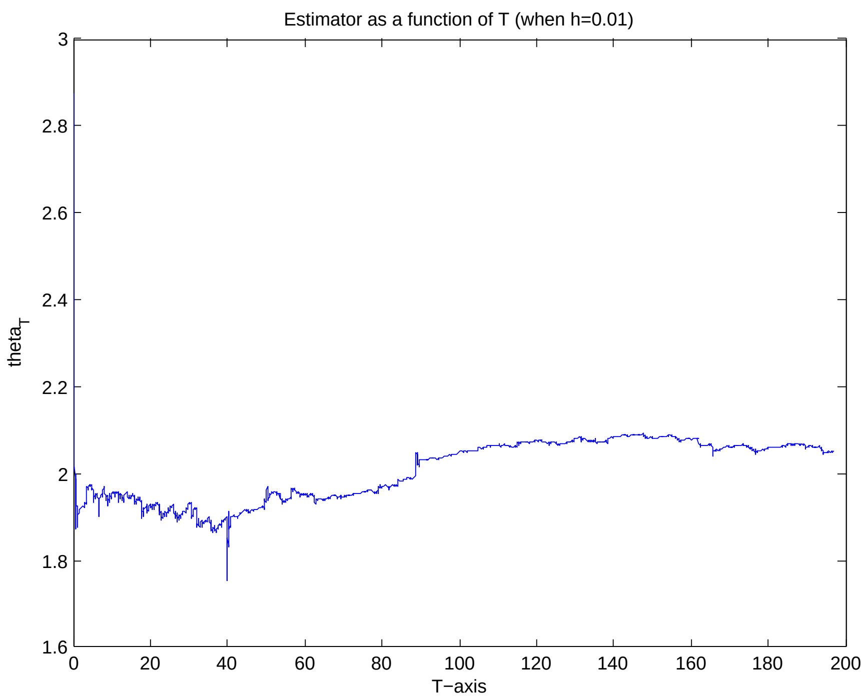

Figure 2: LSE for $\theta_0 = 2$ when $h = 0.01$