Efficient estimation for SDEs with jumps from discrete observations

Hilmar Mai
Humboldt-Universität zu Berlin

SAPS VIII
Le Mans

21 March 2011

Ornstein-Uhlenbeck (OU) Processes

Let $(L_t, t \geq 0)$ be a Lévy process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. For every $a \in \mathbb{R}$

$$dX_t = -aX_t\,dt + dL_t, \quad t \in \mathbb{R}_+, \quad X_0 = x, \tag{1}$$

defines an Ornstein-Uhlenbeck process driven by the Lévy process L.
Equivalently,

$$X_t = e^{-at}X_0 + \int_0^t e^{-a(t-s)}dL_s. \tag{2}$$

We develop a maximum likelihood approach to estimate a.

Absolute Continuity

Theorem

- *Let $P^a, P^{a'}$ be two solution measures of the OU equation for the driving Lévy process L with characteristic triplet (b, σ^2, ρ) and initial distributions π and π'. Suppose that $\sigma^2 > 0$ and $\pi' \ll \pi$, then we have*

$$P^{a'} \overset{loc}{\ll} P^a.$$

- *If $\sigma^2 = 0$, then $P^{a'} \perp P^a$.*

Radon-Nikodym Derivative

Proposition

For two solution measures $P^{a'} \overset{loc}{\ll} P^a$ of the OU equation the Radon-Nikodym derivative is

$$\frac{dP_t^{a'}}{dP_t^a} = \frac{dP_0^{a'}}{dP_0^a} \exp\left(\int_0^t \frac{(a' - a)}{\sigma^2} X_{s-} \, dX_s^c - \frac{(a' - a)^2}{2\sigma^2} \int_0^t X_s^2 ds \right),$$

where X^c denotes the continuous martingale part of X under P^a.

Maximum-Likelihood-Estimator (MLE)

For continuous observations of the Ornstein-Uhlenbeck process X the likelihood function $\mathcal{L}$ for the statistical experiment $(\Omega, \mathcal{F}, (\mathcal{F}_t), (P^a)_{a\in\mathbb{R}})$ takes the form

$$\mathcal{L}(a, X^T) = \frac{dP_t^a}{dP_t^0} = \exp\left(-\frac{a}{\sigma^2}\int_0^T X_{s-}\, dX_s^c - \frac{a^2}{2\sigma^2}\int_0^T X_s^2\, ds\right).$$

Hence, the MLE for a is explicitly given by

$$\hat{a}_T = -\frac{\int_0^T X_{s-} dX_s^c}{\int_0^T X_s^2 ds}.$$

Continuous martingale part X^c

By the Lévy-Itô decomposition of L we can write X as

$$X_t = X_0 - a\int_0^t X_s\, ds + \sigma W_t + J_t \quad , t \geq 0,$$

where W is a standard Wiener process and J a quadratic pure jump process in the sense of Protter given by

$$J_t = \int_{\{|x|<1\}} x(N_t(dx) - t\mu(dx)) + bt + \sum_{0\leq s\leq t} \Delta X_s 1_{\{|\Delta X_s|\geq 1\}}.$$

Hence, the continuous martingale part of X under P^a is

$$X^c = W_t - a\int_0^t X_s\, ds.$$

Asymptotic properties of MLE

The form of the likelihood function means that we are in curved exponential family setting (cf. Küchler and Sørensen (1997)).

Theorem

- *If* $\sigma^2 > 0$ *the MLE*

$$\hat{a}_T = -\frac{\int_0^T X_{s-} dX_s^c}{\int_0^T X_s^2 ds}.$$

  exists and is strongly consistent.

- *If furthermore* X *is stationary and* $E_a[X_0^2] < \infty$ *then under* P^a

$$\sqrt{T}(\hat{a}_T - a) \to N\left(0, \frac{\sigma^2}{E_a[X_0^2]}\right) \quad \textit{weakly}$$

  as $T \to \infty$.

Local asymptotic normality

Theorem

Assume that X is stationary and $E_a[X_0^2] < \infty$, then the following holds:

1. *The statistical experiment $\{P^a, a \in \mathbb{R}\}$ is locally asymptotically normal.*
2. *The maximum likelihood estimator $\hat{a}_T$ is asymptotically efficient in the sense of Hájek-Le Cam.*

Influence of jump noise

A very interesting property is that the MLE is in fact robust to small jumps. Define

$$X_t^j(\epsilon) = \int_{|x| \leq \epsilon} x(N_t(dx) - t\mu(dx)),$$

then the resulting estimate remains strongly consistent.

Theorem

Let us assume that X is stationary with $E[X_0^2] < \infty$, $\sigma^2 > 0$ and set $X^{cj}(\epsilon) = X^c + X^j(\epsilon)$. If we define

$$\tilde{a}_T^\epsilon = -\frac{\int_0^T X_{s-} dX_s^{cj}(\epsilon)}{\int_0^T X_s^2 ds},$$

then $\tilde{a}_T^\epsilon \to a$ with probability 1 as $T \to \infty$.

Influence of jumps on estimation error

Theorem

Let X be a stationary Ornstein-Uhlenbeck process with $E_a[X_0^4] < \infty$, then

$$\sqrt{T}(\tilde{a}_T^\epsilon - a) \to N(0, \Sigma(\epsilon)) \text{ as } T \to \infty$$

where

$$\Sigma(\epsilon) = E_a[X_0^2]^{-1}\sigma^2 + E_a[X_0^2]^{-1} \int_{|x|<\epsilon} x^2 \, \mu(dx).$$

Maximum likelihood versus least squares

The LSE for the parameter a is

$$a_T^{LS} = -\frac{\int_0^T X_{s-} dX_s}{\int_0^T X_s^2 ds}$$

From the theorem it follows for the asymptotic variances

$$AVAR_{LSE} - AVAR_{MLE} = E_a[X_0^2] \int_{\mathbb{R}} x^2 \, \mu(dx) > 0.$$

This motivates the jump filtering approach that will be discussed in the next section.

Filtering jumps from discrete observations

- **Observation scheme:**
  Observation points $0 = t_1 < \ldots < t_n = T_n$ such that $T_n \stackrel{n\to\infty}{\longrightarrow} \infty$ and

$$\Delta_n = \max\{|t_{i+1} - t_i|, 1 \leq i \leq n-1\} \stackrel{n\to\infty}{\longrightarrow} 0.$$

- **Discretized MLE with jump filter:**

$$\bar{a}_n := -\frac{\sum_{i=1}^{n} X_{t_i^n} \Delta_i X \mathbf{1}_{\{|\Delta_i X| \leq v_n\}}}{\sum_{i=1}^{n} X_{t_i^n}^2 \Delta_i^n}$$

for $\Delta_i X = X_{t_{i+1}} - X_{t_i}$ and some cut-off sequence $v_n > 0$.

Discretized MLE with jump filter

Theorem

Assume

- *X stationary and*
- *$v_n = \Delta_n^\alpha$ for $\alpha \in (0, 1/2)$.*

Under suitable conditions on the small jumps (next slide)

$$T_n^{1/2}(\bar{a}_n - a) \xrightarrow{\mathcal{D}} N\left(0, \frac{\sigma^2}{E_a[X_0^2]}\right) \text{ as } n \to \infty.$$

Hence, the truncated MLE is **asymptotically efficient** in the sense of Hájek-Le Cam.

Main steps of the proof

1. Choose the threshold v_n such that the continuous part is approximated in the limit.
2. Show that $\bar{a}_n$ has the same asymptotic behavior as the following benchmark estimator

$$\hat{a}_n = -\frac{\sum_{i=1}^{n} X_{t_i^n} \Delta_i X(u_n)}{\sum_{i=1}^{n} X_{t_i^n}^2 \Delta_i^n}.$$

3. Prove a CLT for the benchmark $\hat{a}_n$.
4. Finally, show that the drift is negligible and

$$T_n^{1/2}(\bar{a}_n - \hat{a}_n) \xrightarrow{p} 0.$$

Identifying the jumps

Define for $n \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$

$$A_n^i = \left\{\omega \in \Omega : \mathbf{1}_{\{|\Delta_i X| \leq v_n\}} = \mathbf{1}_{\{\Delta_i N(\overline{B}_{u_n})=0\}}\right\}$$

where $B_{u_n} = [-u_n, u_n]$ and $N(B_{u_n})$ counts the jumps of L in B_{u_n}.

Lemma

Let $v_n, u_n \downarrow 0$ such that for the Lévy measure μ of L

- $\frac{\mu(B_{2v_n} \setminus B_{u_n})}{\mu(\overline{B}_{u_n})} = o(T_n^{-1})$ *and*
- $u_n^2 v_n^{-2} = o(T_n^{-1})$.

Then, it follows that for $A_n = \bigcap_{i=1}^n A_n^i$ we have

$$P(A_n) \to 1 \quad \text{as } n \to \infty.$$

Sketch of the proof

Two types of errors can appear:

(i) Jumps were not detected:

$$P(|\Delta_i X| \leq v_n, \Delta_i N(\overline{B}_{u_n}) > 0)$$
$$= \Delta_n P(|\Delta_i X| \leq v_n | \Delta_i N(\overline{B}_{u_n}) = 1) + O(\Delta_n^2)$$
$$\sim \Delta_n P(|\Delta_i X| \leq v_n, \Delta_i N(\bar{B}_{2v_n}) = 0, \Delta_i N(B_{2v_n} \backslash B_{u_n}) = 1$$
$$|\Delta_i N(\overline{B}_{u_n}) = 1) + O(\Delta_n^2)$$
$$= O\left(\Delta_n \frac{\mu(B_{2v_n} \backslash B_{u_n})}{\mu(\overline{B}_{u_n})}\right)$$

(ii) Very large increment of the continuous or small jumps part:

$$
\begin{aligned}
&P(|\Delta_i X| > v_n, \Delta_i N(\overline{B}_{u_n}) = 0 \\
&= P(|\Delta_i X(u_n)| > v_n) P(\Delta_i N(\overline{B}_{u_n}) = 0) \\
&\leq C \Delta_n u_n^2 v_n^{-2}
\end{aligned}
$$

where we used the independent increments property of L and independence of W and J.

The benchmark estimator

To proof the CLT for $\bar{a}_n$ we introduce a benchmark estimator

$$\hat{a}_n = -\frac{\sum_{i=1}^n X_{t_i} \Delta_i X(u_n)}{\sum_{i=1}^n X_{t_i}^2 \Delta_i^n}.$$

Lemma

Under the Assumptions of the previous lemma and if $\Delta_n^{1/2} \mu(\bar{B}_{u_n}) = o(T_n^{-1})$ it follows that

$$\left| \sum_{i=1}^n X_{t_i^n} (\Delta_i X \mathbf{1}_{\{|\Delta_i X| \leq v_n\}} - \Delta_i X(u_n)) \right| = o_p(1)$$

as $n \to \infty$.

Sketch of the proof

On A_n we have

$$\sum_{i=1}^{n} \left(\Delta_i X \mathbf{1}_{\{|\Delta_i X| \le v_n\}} - \Delta_i X(u_n)\right) = \sum_{i=1}^{n} \left(\Delta_i X \mathbf{1}_{\{\Delta_i N(\bar{B}_{u_n})=0\}} - \Delta_i X(u_n)\right)$$

Then we obtain the following bound

$$E\left|\mathbf{1}_{A_n} \sum_{i=1}^{n} X_{t_i^n}(\Delta_i X \mathbf{1}_{\{\Delta_i N(\bar{B}_{u_n})=0\}} - \Delta_i X(u_n))\right| = O(\Delta_n^{1/2} \mu(\bar{B}_{u_n})).$$

CLT for benchmark estimator

The following lemma leads to a CLT for the benchmark estimator.

Lemma

Let X be stationary with finite second moments. Set $\tilde{X}(u_n) = \sigma W + J(u_n)$ then

$$T_n^{-1/2} \sum_{i=1}^{n} X_{t_i^n} \Delta_i^n \tilde{X}(u_n) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2 E_a[X_0^2]\right) \text{ as } n \to \infty.$$

Sketch of the proof

The Lemma follows from a CLT for martingale triangular arrays. Therefore, we check

- Lindeberg condition

$$L_n = T_n^{-1} E_a \left[\sum_{i=1}^{n} X_{t_i^n}^2 \Delta_i^n \tilde{X}(u_n)^2 \mathbf{1}_{\{T_n^{-1/2} |X_{t_i^n} \Delta_i^n \tilde{X}(u_n)| > \epsilon\}} | \mathcal{F}_{t_i^n} \right] \xrightarrow{p} 0$$

- and Convergence of the quadratic variation

$$T^{-1} \sum_{i=1}^{n} X_{t_i^n}^2 \Delta_i^n \tilde{X}(u_m)^2 \xrightarrow{p} \sigma^2 E_a[X_0^2] \text{ as } \Delta_n \to 0.$$

For the empirical quadratic variation

$$T^{-1} \sum_{i=1}^{n} X_{t_i^n}^2 \Delta_i^n \tilde{X}(u_m)^2 \xrightarrow{ucp} T^{-1} \int_0^T X_t^2 d[\tilde{X}(u_m)]_t \text{ as } \Delta_n \to 0.$$

and the ergodic theorem implies as $T \to \infty$

$$T^{-1} \int_0^T X_t^2 d[\tilde{X}(u_m)]_t \xrightarrow{p} \sigma^2 E_a[X_0^2] + E_a[X_0^2] \int_{-u_m}^{u_m} x^2 \; \mu(dx)$$

such that finally

$$\sigma^2 E_a[X_0^2] + E_a[X_0^2] \int_{-u_m}^{u_m} x^2 \; \mu(dx) \to \sigma^2 E_a[X_0^2] \text{ as } u_m \to 0.$$

End of the Proof

By the previous lemma a CLT for the benchmark estimator follows

$$T_n^{1/2}(\hat{a}_n - a) \xrightarrow{\mathcal{D}} N(0, \frac{\sigma^2}{E_a[X_0^2]}).$$

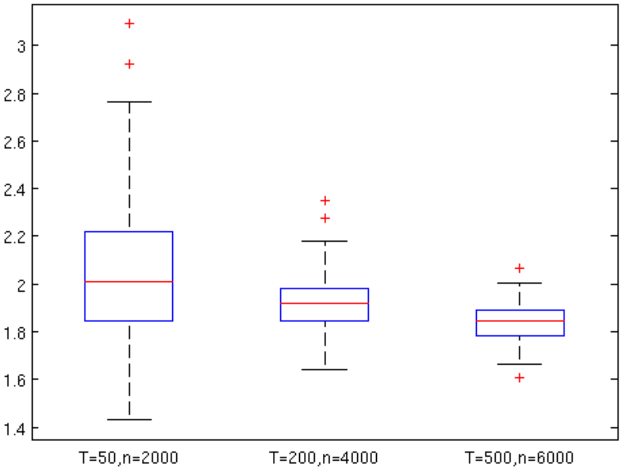
and

$$T_n^{1/2}(\bar{a}_n - \hat{a}_n) \xrightarrow{p} 0$$

as we have shown before.

Simulations

Box plot for $\hat{a}_n$ from a Wiener process plus compound Poisson (intensity $\lambda = 4$, N(0,1)-jumps) driver and true parameter $a = 2$.



Summary

- The MLE takes an explicit form, is asymptotically normal and efficient.
- The jumps lead to an inefficient LSE in this model.
- The discretized MLE attains the efficiency bound from the continuous case when a jump filter is employed.
- The estimator can be directly computed and performs well also for finite sample size.