Maximum likelihood estimation for jump diffusions

Hilmar Mai
Figure 1: Jumps in the trajectory of DM/USD exchange rate, sampled at 5-minute intervals.
Jump diffusions - an overview

\[ dX_t = \gamma(X_t, t) \, dt + \sigma(X_{t-}, t) \, dL_t, \quad t \in [0, T] \]
\[ X_0 = x \]

**Mathematical finance**

- Option pricing: Merton [1976]
- Stochastic volatility models: Heston [1993], Barndorff-Nielsen and Shephard [2001]
- Interest rate models: Cox, Ingersoll and Ross [1985]

**Condensed matter physics**
Atomic diffusion modelling in Hall and Ross [1981]

**Neuroscience**
Neuronal membrane potential in Jahn, Berg, Hounsgaard and Ditlevsen [2011]
Consider an SDE with driver \((L_t, t \geq 0)\) on a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)\)

\[
dX_t = \gamma(X_t, t, \theta) \, dt + \sigma(X_{t-}, t) \, dL_t, \quad t \in [0, T]
\]

\[X_0 = x\]

\(L\) a Lévy process and parameter \(\theta \in A \subset \mathbb{R}\). We call the solution \(X\) a jump diffusion.

Recall that \(L\) is a Lévy process if

- \(L\) has independent and stationary increments,
- \(L\) is continuous in probability, i.e. \(p - \lim_{t \to s} L_t = L_s\)

How to estimate \(\theta\) efficiently when

- \((X_t)_{t \in [0, T]}\) observed continuously?
- given discrete observations \(X_{t_1}, \ldots, X_{t_n}\)?
Define

\[ K_t(x, A) = \int_{\mathbb{R}^d} 1_{A \setminus \{0\}}(\gamma(t, x)) \mu(dy), \]
\[ \beta(\theta, s, x) = \delta(\theta, s, x) + b + \int_{|x|>1} y K_s t(x, dy), \]
\[ c(s, x) = \gamma(t, x)\gamma(t, x)', \]
\[ a(\theta, \theta', s, x) = \beta(\theta, s, X) - \beta(\theta', s, x). \]

Under suitable conditions on the coefficients of \( X \) the Hellinger process \( h \) for the absolute continuity problem induced by \( X \) is given by

\[ h_t(\theta, \theta') = \int_0^t (c(s, X_t)^{-1} a(\theta, \theta', s, X_t))' c(s, X_t) (c(s, X_t)^{-1} a(\theta, \theta', s, X_t)). \]
**Theorem**

Suppose that the coefficients are locally Lipschitz and $c$ is strictly positive definite. Then

(i) $P^\theta_{t \loc} \ll P^\theta'_{t}$ if and only if

$$P^\theta (h_t(\theta, \theta')) < \infty = P^\theta' (h_t(\theta, \theta') < \infty) = 1.$$ 

(ii) If (ii) holds the likelihood function is given by

$$\frac{dP^\theta_t}{dP^{\theta'}_t} = \exp \left[ \int_0^t c(s, X_{s-})^{-1} a(\theta, \theta', s, X_{s-}) \, dX^c_s ight. 
- \left. \frac{1}{2} \int_0^t a(\theta, \theta', s, X_s)' c(s, X_{s-})^{-1} a(\theta, \theta', s, X_s) \, ds \right]$$

where $X^c$ denotes the continuous martingale part under $P^\theta'$. 

---

Likelihood theory for jump diffusions

Drift estimation for jump diffusions · Haindorf Seminar 2012 · Page 6 (24)
Lévy-driven Ornstein-Uhlenbeck processes

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

\[
dX_t = -aX_t \, dt + dL_t, \quad t \in \mathbb{R}_+, \quad X_0 = x,
\]

defines an Ornstein-Uhlenbeck process driven by the Lévy process \(L\).

Equivalently,

\[
X_t = e^{-at}X_0 + \int_0^t e^{-a(t-s)}dL_s.
\]  \hspace{1cm} (3)

The likelihood function is

\[
\frac{dP^a_T}{dP^a_0} = \frac{dP^a_0}{dP^a_T} \exp \left( \int_0^t \frac{(a' - a)}{\sigma^2} X_s - dX^c_s - \frac{(a' - a)^2}{2\sigma^2} \int_0^t X^2_s \, ds \right),
\]

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

The MLE vor \(a\) is

\[
\hat{a}_T = -\frac{\int_0^T X_s - dX^c_s}{\int_0^T X^2_s \, ds}.
\]
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
  \[
  \hat{a}_T = -\frac{\int_0^T X_s - dX^c_s}{\int_0^T X^2_s ds}.
  \]
  exists and is strongly consistent.

- If furthermore $X$ is stationary and $E_a[X_0^2] < \infty$ then under $P^a$
  \[
  \sqrt{T}(\hat{a}_T - a) \rightarrow N\left(0, \frac{\sigma^2}{E_a[X_0^2]}\right) \quad \text{weakly}
  \]
  as $T \rightarrow \infty$. 
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 $\hat{\alpha}_T$ is asymptotically efficient in the sense of Hájek-Le Cam.
Lévy-driven CIR processes

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

\[ dX_t = -aX_t dt + \sqrt{X_t} dL_t; \quad 0 \leq t \leq T \]

defines **Cox-Ingersoll-Ross process** \(X\) driven by the Lévy process \(L\). No explicit solution is known for this SDE.

The likelihood function for continuous-time observations is

\[
\frac{dP_T^a}{dP_T^0} = \exp \left[ -\frac{1}{\sigma^2} (X_T^c - X_0^c) - \frac{a}{\sigma^2} \int_0^T X_t \, dt \right].
\]

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

The MLE vor \(a\) is

\[
\hat{a}_T = -\frac{X_T^c - X_0^c}{\int_0^T X_t \, dt}.
\] (4)
Theorem

Let $X$ be stationary, $\sigma^2 > 0$ and $E[X_0^2] < \infty$. Then the MLE $\hat{a}_T$ is consistent and

$$\sqrt{T}(\hat{a}_T - a) \xrightarrow{D} N(0, \sigma^2 E[|X_0|^{-1}])$$

as $T \to \infty$.

Moreover, the model is locally asymptotically normal such that the MLE is efficient in the sense of Hájek-Le Cam.
Linear stochastic delay equations

Let $L$ be a Lévy process with Lévy–Khintchine triplet $(\gamma, \sigma^2, \mu)$ and consider the following linear stochastic delay differential equation (SDDE) driven by $L$.

$$dX_t = aX_t\ dt + bX_{t-1}\ dt + dL_t, \quad t > 0$$

$$(6)$$

$$X_t = X_t^0, \quad t \in [-1, 0]$$

$$(7)$$

where $a, b \in \mathbb{R}$ and $X^0 : [-1, 0] \times \Omega \to \mathbb{R}$ is the initial process with càdlàg trajectories that is assumed to be independent of $L$.

The likelihood function reads

$$L(\theta, X, T) = \frac{dP_T^\theta}{dP_T^{(0,0)}} = \exp \left( \theta' V_T - \frac{1}{2} \theta' I_T \theta \right),$$

where

$$V_T = \left( \int_0^T X_t\ dX^c_t \int_0^T X_{t-1}\ dX^c_t \right)$$

and $I_T$ is the observed Fisher information

$$I_T = \begin{pmatrix} \int_0^T X_t^2\ dt & \int_0^T X_t X_{t-1}\ dt \\ \int_0^T X_t X_{t-1}\ dt & \int_0^T X_{t-1}^2\ dt \end{pmatrix}. $$
Consistency and asymptotic normality

**Theorem**

Assume that $X$ is a stationary solution, then the MLE $\hat{\theta}_T$ is strongly consistent, i.e. under $P^a$

$$\hat{\theta}_T \overset{a.s.}{\longrightarrow} \theta \text{ as } t \to \infty.$$  

**Theorem**

Assume that $X$ is a stationary solution, then

$$T^{1/2} \left( \hat{\theta}_T - \theta \right) \overset{D}{\longrightarrow} N \left( 0, \Sigma^{-1} \right)$$

where

$$\Sigma = \sigma^2 \begin{pmatrix}
\int_0^\infty x_0(s)^2 \, ds & \int_0^\infty x_0(s) x_0(s+1) \, ds \\
\int_0^\infty x_0(s) x_0(s+1) \, ds & \int_0^\infty x_0(s)^2 \, ds
\end{pmatrix};$$

and $x_0$ denotes the fundamental solution.
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

$$X^c = W_t - a \int_0^t X_s \, ds.$$
Influence of jump noise

An interesting question is the influence of jumps on the MLE. Define

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

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 $X^{c,j}(\epsilon) = X^c + X^j(\epsilon)$. If we define

$$\tilde{\alpha}_T^\epsilon = -\frac{\int_0^T X_s - dX^{c,j}_s(\epsilon)}{\int_0^T X_s^2 ds},$$

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

Let $X$ be a stationary Ornstein-Uhlenbeck process with $E_a[X_0^2] < \infty$, then

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

where

$$\Sigma(\epsilon) = E_a[X_0^2]^{-1} \sigma^2 + E_a[X_0^2]^{-1} \int_{|x| < \epsilon} x^2 \mu(dx).$$
The LSE for the parameter $a$ is

$$ a_{LS}^T = -\frac{\int_0^T X_s - dX_s}{\int_0^T X_s^2 ds} $$

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 $T_n \xrightarrow{n \to \infty} \infty$ and
  \[
  \Delta_n = \max\{|t_{i+1} - t_i|, 1 \leq i \leq n - 1\} \xrightarrow{n \to \infty} 0.
  \]

- **Discretized MLE with jump filter:**
  \[
  \bar{a}_n := -\frac{\sum_{i=1}^{n} X_{t_i}^n \Delta_i X 1_{\{\Delta_i X \leq v_n\}}}{\sum_{i=1}^{n} X_{t_i}^2 \Delta_i^n}
  \]
  for $\Delta_i X = X_{t_{i+1}} - X_{t_i}$ and some cut-off sequence $v_n > 0$. 
Assumption

1. There exists $\alpha \in (0, 2)$ such that as $v \downarrow 0$

$$\int_{-v}^{v} x^2 \mu(dx) = O(v^{2-\alpha}). \quad (8)$$

2. There exists $\delta > 0$ such that for all $\epsilon \leq \delta$

$$E[\Delta_i M 1_{\{|\Delta_i M| \leq \epsilon\}}] = 0$$

Theorem

If Assumption 1. and 2. hold and $v_n = \Delta_n^\beta$ for $\beta \in (0, 1/2)$ such that $T_n \Delta_n^{1/2-\beta} = o(1)$ as $n \to \infty$ then

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

Furthermore, the estimator is asymptotically efficient.

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

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

3. Prove a CLT for the benchmark $\hat{a}_n$.

4. Finally, show that the drift is negligible and

$$T_n^{1/2} (\bar{a}_n - \hat{a}_n) \overset{p}{\rightarrow} 0.$$
Identifying the jumps

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

$$A^i_n = \left\{ \omega \in \Omega : 1\{\Delta_i X \leq v_n\} = 1\{\Delta_i N(B^c_{u_n}) = 0\} \right\}$$

where $B_{u_n} = [-u_n, u_n]$, its complement $B^c_{u_n}$ 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 $\mu$ of $L$

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

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

$$P(A_n) \to 1 \text{ as } n \to \infty.$$
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 $\Delta_n^{1/2} \mu(B_{u_n}^c) = o(T_n^{-1})$ it follows that

$$\left| \frac{1}{n-1} \sum_{i=0}^{n-1} X_{t_i} \left( \Delta_i X 1_{\{|\Delta_i X| \leq v_n\}} - \Delta_i X(u_n) \right) \right| = o_p(1)$$

as $n \to \infty$. 
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

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

- We have developed a maximum likelihood approach for drift estimation in jump diffusions.
- The jumps lead to an inefficient LSE in this class of processes.
- Several examples like Ornstein-Uhlenbeck, Cox-Ingersoll-Ross and linear stochastic delay equations lead to explicit MLEs.
- Strong optimality properties have been demonstrated for the MLE is these models.
- For the OU process the discretized MLE attains the efficiency bound from the continuous case when a jump filter is employed.
- The estimator can be computed directly and performs well for finite sample size.