LSE-type estimation for stochastic processes with small Lévy noise

Yasutaka Shimizu

Dept. Applied Mathematics, Waseda University

Asymptotical Statistics of Stochastic Processes X
March 17–20th, 2015@ Le Mans

Introduction: SDE with small noise

X^ϵ: d-dim stocahstaic process,

$$\mathrm{d}X_t^\epsilon = b(X_t^\epsilon, \theta_0)\,\mathrm{d}t + \epsilon \cdot \mathrm{d}Q_t^\epsilon, \quad X_0^\epsilon = x.$$

- $x \in \mathbb{R}^d$; $\theta_0 \in \Theta_0(\subset \mathbb{R}^p)$ (open): a parameter.
- $b : \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$.
- Q^ϵ: a d-dim stochastic process, $\to Q$ ($\epsilon \downarrow 0$), (*small noise asymptotics*);

$$\mathrm{d}X_t^0 = b(X_t^0, \theta_0)\,\mathrm{d}t, \quad X_0^0 = x.$$

Applications: finance and insurance, e.g., **Yoshida (1992a)**, **Kunitomo and Takahashi (2001)**, **Takahashi and Yoshida (2004)**, **Uchida and Yoshida (2004)**, **Pavlyukevich (2008)**, **etc.**

Statistical infernece: Estimating θ_0 from $\{X_{t_k^n}\}_{k=0,1,2,\dots,n}$, $t_k^n = k/n \in [0,1]$
Asymptotic theory: under $n \to \infty$, $\epsilon \to 0$.

Eariler works

- Small diffusions: $\mathrm{d}Q_t^\epsilon \equiv \sigma(X_t^\epsilon, \sigma_0)\mathrm{d}W_t$ (W:Wiener process)
  - Continuous obs.: **Kutoyants (1982,1984)**, **Laredo (1990)**, **Yoshida (1992)**.
  - Discrete obs.: **Sørensen and Uchida (2003)**, **Gloter and Sørensen (2009)**

Eariler works

- Small diffusions: $\mathrm{d}Q_t^\epsilon \equiv \sigma(X_t^\epsilon, \sigma_0)\mathrm{d}W_t$ (W:Wiener process)
  - Continuous obs.: **Kutoyants (1982,1984)**, **Laredo (1990)**, **Yoshida (1992)**.
  - Discrete obs.: **Sørensen and Uchida (2003)**, **Gloter and Sørensen (2009)**
- Small Lévy noise: $Q^\epsilon \equiv Q$ (Lévy process)
  - **Long (2009)**, **Ma (2010)** (Lévy O-U: $b(x, \theta) = -\theta x$)
  - **Long, Shimizu and Sun (2013)**: (b: non-linear)

Eariler works

- Small diffusions: $\mathrm{d}Q_t^\epsilon \equiv \sigma(X_t^\epsilon, \sigma_0)\mathrm{d}W_t$ (W:Wiener process)
  - Continuous obs.: **Kutoyants (1982,1984)**, **Laredo (1990)**, **Yoshida (1992)**.
  - Discrete obs.: **Sørensen and Uchida (2003)**, **Gloter and Sørensen (2009)**
- Small Lévy noise: $Q^\epsilon \equiv Q$ (Lévy process)
  - **Long (2009)**, **Ma (2010)** (Lévy O-U: $b(x,\theta) = -\theta x$)
  - **Long, Shimizu and Sun (2013)**: (b: non-linear)

$$\widehat{\theta}_{n,\epsilon}^{LSE} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2$$

where $\Delta_n := t_k^n - t_{k-1}^n (= 1/n), \Delta_k^n X = X_{t_k^n} - X_{t_{k-1}^n}$.

Eariler works

- Small diffusions: $\mathrm{d}Q_t^\epsilon \equiv \sigma(X_t^\epsilon, \sigma_0)\mathrm{d}W_t$ (W:Wiener process)
  - Continuous obs.: **Kutoyants (1982,1984)**, **Laredo (1990)**, **Yoshida (1992)**.
  - Discrete obs.: **Sørensen and Uchida (2003)**, **Gloter and Sørensen (2009)**
- Small Lévy noise: $Q^\epsilon \equiv Q$ (Lévy process)
  - **Long (2009)**, **Ma (2010)** (Lévy O-U: $b(x, \theta) = -\theta x$)
  - **Long, Shimizu and Sun (2013)**: (b: non-linear)

$$\widehat{\theta}_{n,\epsilon}^{LSE} := \arg\min_{\theta\in\Theta} \sum_{k=1}^n \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2$$

where $\Delta_n := t_k^n - t_{k-1}^n (= 1/n), \Delta_k^n X = X_{t_k^n} - X_{t_{k-1}^n}$.

$$\Longrightarrow \quad \epsilon^{-1}\left(\widehat{\theta}_{n,\epsilon}^{LSE} - \theta_0\right) \xrightarrow{\mathbb{P}} \zeta := \int_0^1 \xi(t, \theta_0)\, \mathrm{d}Q_t,$$

as $n \to \infty, \epsilon \to 0$ and $n\epsilon \to \infty$.

Eariler works

- Small diffusions: $\mathrm{d}Q_t^\epsilon \equiv \sigma(X_t^\epsilon, \sigma_0)\mathrm{d}W_t$ (W:Wiener process)
  - Continuous obs.: **Kutoyants (1982,1984)**, **Laredo (1990)**, **Yoshida (1992)**.
  - Discrete obs.: **Sørensen and Uchida (2003)**, **Gloter and Sørensen (2009)**
- Small Lévy noise: $Q^\epsilon \equiv Q$ (Lévy process)
  - **Long (2009)**, **Ma (2010)** (Lévy O-U: $b(x,\theta) = -\theta x$)
  - **Long, Shimizu and Sun (2013)**: (b: non-linear)

$$\widehat{\theta}_{n,\epsilon}^{LSE} := \arg\min_{\theta\in\Theta} \sum_{k=1}^n \left|\Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n\right|^2$$

where $\Delta_n := t_k^n - t_{k-1}^n (= 1/n), \Delta_k^n X = X_{t_k^n} - X_{t_{k-1}^n}$.

$$\Longrightarrow \quad \epsilon^{-1}\left(\widehat{\theta}_{n,\epsilon}^{LSE} - \theta_0\right) \xrightarrow{\mathbb{P}} \zeta := \int_0^1 \xi(t,\theta_0)\,\mathrm{d}Q_t,$$

as $n\to\infty, \epsilon\to 0$ and $n\epsilon\to\infty$.

- But, this LSE is unstable in finite sample performance.

A problem of LSE: simulation example

Consider a 1-dim Lévy driven SDE $X^\epsilon = (X_t^\epsilon)_{t\geq 0}$:

$$b(x,\theta) = -\frac{\theta x}{\sqrt{1+x^2}}, \quad Q_t = S_t^\alpha,$$

- S^α: a standard symmetric α-stable process with

$$\mathbb{E}[e^{itS_1^\alpha}] = e^{-|z|^\alpha}, \quad \alpha \in (0,2)$$

- Set:

$$X_0 = 1, \quad (\theta,\alpha) = (1.0, 1.5).$$

- Generate 10,000 sample paths, and calculate mean and s.d. of $\widehat{\theta}_{n,\epsilon}^{LSE}$.

Results

Theoretically: "$\widehat{\theta}_{n,\epsilon}^{LSE} \xrightarrow{\mathbb{P}} \theta_0, \quad \epsilon \to 0, \ n \to \infty$".

$\epsilon = 0.4$	$n = 1000$	$n = 3000$	$n = 5000$	True
$\widehat{\theta}_{n,\epsilon}^{LSE}$	1.72415	1.79755	1.78645	1.0
(s.d.)	(3.5426)	(4.2814)	(2.9579)	

$\epsilon = 0.05$	$n = 1000$	$n = 3000$	$n = 5000$	True
$\widehat{\theta}_{n,\epsilon}^{LSE}$	1.05963	1.04438	1.06135	1.0
(s.d.)	(1.3026)	(0.6773)	(0.7344)	

Table: LSE is biased and unstable (large s.d.) even if ϵ is small.

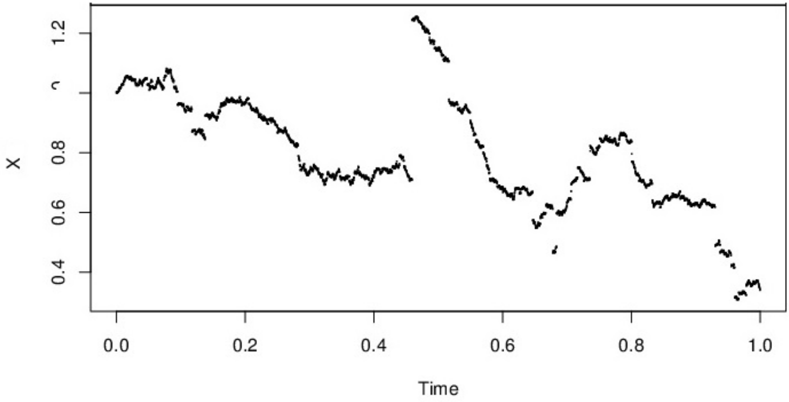

Figure: A path of X

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{"Filtered LSE"}: \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{"Filtered LSE"}: \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \leq \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

Contributions:

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{"Filtered LSE"} : \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

Contributions:

1. It can be asymptotically equivalent to the usual LSE for suitable choice of $\delta_{n,\epsilon} \downarrow 0$.

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{"Filtered LSE"} : \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \leq \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

Contributions:

1. It can be asymptotically equivalent to the usual LSE for suitable choice of $\delta_{n,\epsilon} \downarrow 0$.
2. Numerically: bias and variance can be reduced even if ϵ is not so "small".

“Filtered” LSEs

Proposal: Cut “large” shocks from the contrast function:

$$\text{“Filtered LSE”} : \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate “large” shocks causing a bias to drift estimation.

Contributions:

1. It can be asymptotically equivalent to the usual LSE for suitable choice of $\delta_{n,\epsilon} \downarrow 0$.
2. Numerically: bias and variance can be reduced even if ϵ is not so “small”.
3. *Mighty convergence*: Not only $\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \xrightarrow{\mathbb{P}} \zeta$ (same as the LSE),

$$\mathbb{E}\left[f\left(\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \right) \right] \to \mathbb{E}[f(\zeta)]$$

for every continuous f, of polynomial growth.

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{"Filtered LSE"} : \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

Contributions:

1. It can be asymptotically equivalent to the usual LSE for suitable choice of $\delta_{n,\epsilon} \downarrow 0$.
2. Numerically: bias and variance can be reduced even if ϵ is not so "small".
3. *Mighty convergence*: Not only $\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \xrightarrow{\mathbb{P}} \zeta$ (same as the LSE),

   $$\mathbb{E}\left[f\left(\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \right) \right] \to \mathbb{E}[f(\zeta)]$$

   for every continuous f, of polynomial growth.
4. (Extension to the case where $Q^\epsilon \to Q$: a semimartingale: **S. (2014)**)

"Filtered" LSEs

Proposal: Cut "large" shocks from the contrast function:

$$\text{“Filtered LSE”}: \ \widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^{n} \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

$\delta_{n,\epsilon}$: *Threshold* to eliminate "large" shocks causing a bias to drift estimation.

Contributions:

1. It can be asymptotically equivalent to the usual LSE for suitable choice of $\delta_{n,\epsilon} \downarrow 0$.
2. Numerically: bias and variance can be reduced even if ϵ is not so "small".
3. *Mighty convergence*: Not only $\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \xrightarrow{\mathbb{P}} \zeta$ (same as the LSE),

$$\mathbb{E}\left[f\left(\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \right) \right] \to \mathbb{E}[f(\zeta)]$$

   for every continuous f, of polynomial growth.
4. (Extension to the case where $Q^\epsilon \to Q$: a semimartingale: **S. (2014)**)
5. A possibility of asymptotic noramlity for appropriate $\delta_{n,\epsilon}$ when $Q^\epsilon \equiv$ Lévy.

Models

X^ϵ: 1-dim (for simplicity of notation)

$$\mathrm{d}X_t^\epsilon = b(X_t^\epsilon, \theta_0)\,\mathrm{d}t + \epsilon \cdot \mathrm{d}Q_t, \quad X_0^\epsilon = x.$$

- $x \in \mathbb{R}$ (possibly random); $\theta_0 \in \Theta_0$ (open) $\subset \mathbb{R}^p$.
- $b : \mathbb{R} \times \overline{\Theta}_0 \to \mathbb{R}$.
- Q: a Lévy process s.t. $\psi(u) := \log \mathbb{E}\left[\exp(iuQ_1)\right]$,

$$\psi(u) = i\mu u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}^d} \left(e^{iuz} - 1 - \frac{iuz}{1+z^2} \right) \nu(\mathrm{d}z), \quad u \in \mathbb{R},$$

  where $\mu, \sigma \in \mathbb{R}$, ν is the Lévy measure.
- Estimation of θ_0 from

$$\{X_{t_k^n}\}_{k=0,1,2,\dots,n}, \quad t_k^n = k/n \in [0,1], \quad \Delta_n = 1/n.$$

  by a Filtered LSE:

$$\widehat{\theta}_{n,\epsilon} := \arg\min_{\theta\in\Theta} \sum_{k=1}^n \left| \Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\Delta_n \right|^2 \mathbf{1}_{\{|\Delta_k^n X| \le \delta_{n,\epsilon}\}},$$

Assumptions

A1 $|b(x,\theta) - b(y,\theta)| \le C|x-y|$ for $x, y \in \mathbb{R}$, $\theta \in \Theta$, under which the ODE

$$\mathrm{d}X_t^0 = b(X_t^0, \theta_0)\,\mathrm{d}t, \quad X_0^0 = x,$$

has the unique solution.

A2 $b \in C^{2,3}(\mathbb{R} \times \Theta; \mathbb{R})$ and

$$\sup_{\theta\in\Theta} |\nabla_x^k \nabla_\theta^l b(x,\theta)| \le C(1+|x|)^C \quad (k = 0,1,2,\ l = 0,1,2,3)$$

A3 $\theta \ne \theta' \Leftrightarrow b(X_t^0, \theta) \ne b(X_t^0, \theta')$ for at least one value of $t \in [0,1]$.

A4 $I(\theta_0) := \int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \nabla_\theta b(X_t^0, \theta_0)\,\mathrm{d}t$ is positive definite.

Assumptions on Q

Q1[γ] For $\gamma > 0$ such that,

$$\mathbb{P}\left\{\sup_{t\in(0,\Delta_n]} |Q_t| > \Delta_n^\gamma\right\} = o_p(1).$$

Q2[q] For given $q > 0$,

$$\int_{|z|>1} |z|^q \, \nu(\mathrm{d}z) < \infty.$$

On the condition Q1[γ] I

Assume: Q is a Lévy process with chractersitic exponent

$$\psi(u) = i\mu u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}^d} \left(e^{iuz} - 1 - \frac{iuz}{1+|z|^2} \right) \nu(\mathrm{d}z), \quad u \in \mathbb{R}.$$

On the condition Q1[γ] I

Assume: Q is a Lévy process with chractersitic exponent

$$\psi(u) = i\mu u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}^d} \left(e^{iuz} - 1 - \frac{iuz}{1+|z|^2} \right) \nu(\mathrm{d}z), \quad u \in \mathbb{R}.$$

- Q: a Wiener process, Q1[γ] holds for $\gamma \in (0, 1/2)$ since

$$\mathbb{P}\left\{ \sup_{t\in(0,\Delta_n]} |Q_t| > \Delta_n^\gamma \right\} = 2\left(1 - \Phi\left(\frac{\Delta_n^\gamma}{\sqrt{\Delta_n}}\right)\right) \to 0.$$

See e.g., **Doob (1949)**.

On the condition Q1[γ] I

Assume: Q is a Lévy process with chractersitic exponent

$$\psi(u) = i\mu u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}^d} \left(e^{iuz} - 1 - \frac{iuz}{1+|z|^2} \right) \nu(\mathrm{d}z), \quad u \in \mathbb{R}.$$

- Q: a Wiener process, Q1[γ] holds for $\gamma \in (0, 1/2)$ since

$$\mathbb{P}\left\{ \sup_{t\in(0,\Delta_n]} |Q_t| > \Delta_n^\gamma \right\} = 2\left(1 - \Phi\left(\frac{\Delta_n^\gamma}{\sqrt{\Delta_n}}\right)\right) \to 0.$$

  See e.g., **Doob (1949)**.

- Q: a symmetric α-stable process with $\alpha \in (1, 2)$, $\gamma \in (0, \alpha^{-1})$, since

$$\mathbb{P}\left\{ \sup_{t\in(0,\Delta_n]} |Q_t| > \Delta_n^\gamma \right\} = O\left(\Delta_n^{1-\gamma\alpha}\right) \to 0$$

  See **Julian (2007)**.

On the condition Q1[γ] II

For Q with

$$\psi(u) = i\mu u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}^d} \left(e^{iuz} - 1 - \frac{iuz}{1+|z|^2} \right) \nu(\mathrm{d}z), \quad u \in \mathbb{R}.$$

Let

$$\begin{aligned} h(x) := & \int_{|z|>x} \nu(\mathrm{d}z) + x^{-2} \int_{|z|\le x} |z|^2 \nu(\mathrm{d}z) \\ & + x^{-1} \left| \mu + \int_{|z|\le x} \frac{z|z|^2}{1+|z|^2} \nu(\mathrm{d}z) - \int_{|z|>x} \frac{z}{1+|z|^2} \nu(\mathrm{d}z) \right|. \end{aligned}$$

On the condition Q1[γ] III

Theorem

Suppose that there exists a constant

$$\beta := \inf \left\{ \eta > 0 : \limsup_{x \to 0} x^{\eta} h(x) = 0 \right\}.$$

Then the condition Q1[γ] holds true for any $\gamma \in (0, \gamma_0)$, where

$$\gamma_0 = \begin{cases} \beta^{-1} \wedge 1 & (\sigma = 0) \\ (\beta \vee 2)^{-1} & (\sigma \neq 0) \end{cases}.$$

In particular, if $\int_{|z| \le 1} |z| \nu(\mathrm{d}z) < \infty$ then β is given by the Blumenthal-Getoor index:

$$\beta = \inf \left\{ \eta > 0 : \int_{|z| \le 1} |z|^{\eta} \, \nu(\mathrm{d}z) < \infty \right\} \le 1.$$

Main results

Theorem (Consistency & Asymptotic dist.)

Suppose A1–A3, Q1[γ], and that a sequence $\{\delta_{n,\epsilon}\}$ satisfies that

$$n\delta_{n,\epsilon} \to \infty, \quad \delta_{n,\epsilon} n^{\gamma} \epsilon^{-1} \to \infty.$$

Then $\widehat{\theta}_{n,\epsilon}$ is asymptotically equivalent to $\widehat{\theta}^{LSE}_{n,\epsilon}$, and

$$\widehat{\theta}_{n,\epsilon} \xrightarrow{\mathbb{P}} \theta_0.$$

In addition, suppose that $n\epsilon \to \infty$. Then

$$\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \xrightarrow{\mathbb{P}} \zeta := I^{-1}(\theta_0) \int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \, \mathrm{d}Q_t.$$

Main results

Theorem (Consistency & Asymptotic dist.)

Suppose A1–A3, Q1[γ], and that a sequence $\{\delta_{n,\epsilon}\}$ satisfies that

$$n\delta_{n,\epsilon} \to \infty, \quad \delta_{n,\epsilon} n^{\gamma} \epsilon^{-1} \to \infty.$$

Then $\widehat{\theta}_{n,\epsilon}$ is asymptotically equivalent to $\widehat{\theta}^{LSE}_{n,\epsilon}$, and

$$\widehat{\theta}_{n,\epsilon} \xrightarrow{\mathbb{P}} \theta_0.$$

In addition, suppose that $n\epsilon \to \infty$. Then

$$\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \xrightarrow{\mathbb{P}} \zeta := I^{-1}(\theta_0) \int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \mathrm{d}Q_t.$$

Theorem (Mighty convergence)

Suppose the same assumptions as above, and that Q2[q] holds true for any $q > 0$. Then

$$\mathbb{E}\left[f\left(\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta_0) \right) \right] \to \mathbb{E}[f(\zeta)],$$

for every continuous function $f : \mathbb{R}^p \to \mathbb{R}$, of polynomial growth.

Numerical results

Consider a 1-dim Lévy driven SDE $X^\epsilon = (X^\epsilon_t)_{t\geq 0}$:

$$b(x,\theta) = -\frac{\theta x}{\sqrt{1+x^2}}, \quad Q_t = S^\alpha_t,$$

- S^α: a standard symmetric α-stable process with

$$\mathbb{E}[e^{itS_1^\alpha}] = e^{-|z|^\alpha}, \quad \alpha \in (0,2)$$

- Set:

$$X_0 = 1, \quad (\theta, \alpha) = (1, 1.5).$$

- Compare $\widehat{\theta}^{LSE}_{n,\epsilon}$ and our proposal (filtered LSE) $\widehat{\theta}_{n,\epsilon}$.

LSE vs. **Filtered** LSE: $\delta_{n,\epsilon} = \epsilon/5$

$\epsilon = 0.4$	$n = 1000$	$n = 3000$	$n = 5000$	True
$\widehat{\theta}_{n,\epsilon}^{LSE}$	1.72415	1.79755	1.78645	1.0
(s.d.)	(3.5426)	(4.2814)	(2.9579)	
$\widehat{\theta}_{n,\epsilon}$	1.14289	1.15283	1.16222	1.0
(s.d.)	(0.8721)	(0.88426)	(0.8856)	

$\epsilon = 0.05$	$n = 1000$	$n = 3000$	$n = 5000$	True
$\widehat{\theta}_{n,\epsilon}^{LSE}$	1.05963	1.04438	1.06135	1.0
(s.d.)	(1.3026)	(0.6773)	(0.7344)	
$\widehat{\theta}_{n,\epsilon}$	0.98972	0.99936	1.00039	1.0
(s.d.)	(0.10098)	(0.1031)	(0.1037)	

Table: The perfomance for $\widehat{\theta}_{n,\epsilon}$ is drastically improved (unbiased and stable!).

Normal QQ-plot for LSE and Filtered: $\epsilon = 0.05, n = 5000$

$$\epsilon^{-1}\left(\widehat{\theta}_{n,\epsilon} - \theta_0\right) \xrightarrow{\mathbb{P}} I^{-1}(\theta_0)\int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \, \mathrm{d}Q_t$$

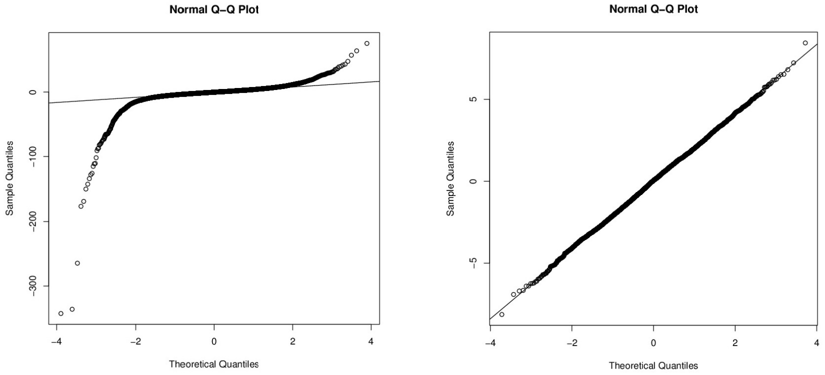

Figure: Normal QQ-plots for $\epsilon^{-1}(\widehat{\theta}_{n,\epsilon}^{LSE} - \theta)$ (left) and $\epsilon^{-1}(\widehat{\theta}_{n,\epsilon} - \theta)$ (right).

Can the filtered LSEs be asymptotically noraml?

Can the filtered LSEs be asymptotically noraml?

- Suppose that *all the jumps* $|\Delta Q_t|$ *are observable*.

Can the filtered LSEs be asymptotically noraml?

- Suppose that *all the jumps $|\Delta Q_t|$ are observable*.
- We can judge that, for given $c > 0$,

$$\sup_{t\in(t_k,t_{k+1}]} |\Delta Q_t| \le c \quad \text{or} \quad \sup_{t\in(t_k,t_{k+1}]} |\Delta Q_t| > c$$

Can the filtered LSEs be asymptotically noraml?

- Suppose that *all the jumps $|\Delta Q_t|$ are observable*.
- We can judge that, for given $c > 0$,

$$\sup_{t\in(t_k,t_{k+1}]} |\Delta Q_t| \le c \quad \text{or} \quad \sup_{t\in(t_k,t_{k+1}]} |\Delta Q_t| > c$$

- $\widetilde{\theta}_{n,\epsilon,\delta} = \arg\min_{\theta\in\Theta} \widetilde{\Phi}_{n,\epsilon}(\theta)$:

$$\widetilde{\Phi}_{n,\epsilon}(\theta) = \sum_{k=1}^{n} |\Delta_k^n X - b(X_{t_{k-1}^n}, \theta)\cdot\Delta_n|^2 \mathbf{1}_{\{\|\Delta Q\|_k^* \le \delta/\epsilon\}},$$

where

$$\|\Delta Q_t\|_k^* := \sup_{t\in(t_k,t_{k+1}]} |\Delta Q_t|.$$

Finite activity case

Suppose that $\int_{|z|\leq 1} \nu(\mathrm{d}z) < \infty$, $\sigma^2 > 0$:

$$Q_t = \sigma W_t + \sum_{i=1}^{N_t} U_i, \tag{1}$$

where $N_t \sim Po(\lambda t)$ with $\lambda = \int_{\mathbb{R}} \nu(\mathrm{d}z)$; $U_i \sim \lambda^{-1}\nu$ (IID).

Theorem

Suppose that Q is as above, and that A1–A4 hold true. Moreover, suppose that

$$\delta/\epsilon \to 0.$$

Then

$$\epsilon^{-1}(\widetilde{\theta}_{n,\epsilon,\delta} - \theta_0) \xrightarrow{\mathbb{P}} I^{-1}(\theta_0) \int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \, \mathrm{d}W_t.$$

Hence $\widetilde{\theta}_{n,\epsilon,\delta}$ is asymptotically normal.

Remarks:

- $\widehat{\theta}_{n,\epsilon}$:

$$n\delta_{n,\epsilon} \to \infty, \quad \delta_{n,\epsilon} n^{\gamma} \epsilon^{-1} \to \infty.$$

We can take $\delta_{n,\epsilon} \to 0$, but **not so fast**. Then

$$LSE \ \sim \ Filtered \ LSE$$

Remarks:

- $\widehat{\theta}_{n,\epsilon}$:

$$n\delta_{n,\epsilon} \to \infty, \quad \delta_{n,\epsilon} n^{\gamma} \epsilon^{-1} \to \infty.$$

We can take $\delta_{n,\epsilon} \to 0$, but **not so fast**. Then

$$\textit{LSE} \ \sim \ \textit{Filtered LSE}$$

- $\widetilde{\theta}_{n,\epsilon,\delta}$:

$$\delta/\epsilon \to 0.$$

$\delta \to 0$ **faster!** Then

$$\textit{Filtered LSE} \ \sim \ \textit{Asymptotically Normal}$$

Infinite activity case

Suppose that $\int_{|z|<1} \nu(\mathrm{d}z) = \infty$: infinitely many jumps occur in any $[0, \Delta]$

Theorem

Suppose Q is as above, and that

- *A1–A4, Q1[γ], and that Q2[q] hold true for any $q > 0$; $n\epsilon\Delta_n^\gamma \to 0$;*
- *for $\exists\, c \in (0,1)$ s.t. $\dfrac{\lambda(\delta/\epsilon)}{n \log n} \to c$, where $\lambda(\kappa) := \int_{|z|>\kappa} \nu(\mathrm{d}z)$;*
- *$\exists\, \rho \in (0,1)$ s.t. $\sigma^\rho(\delta/\epsilon) \log n \to \infty$;*

  $n\epsilon \cdot \sigma(\delta/\epsilon) \to \infty$, where $\sigma^2(\kappa) = \int_{|z|\le\kappa} |z|^2 \nu(\mathrm{d}z)$;
- *for each $\kappa > 0$, $\sigma\left(\kappa\sigma(\delta/\epsilon) \wedge \delta/\epsilon\right) \sim \sigma(\delta/\epsilon)$.*

Then there exists a d-dimensional B.M. B, independent of $X_0 = x$, such that

$$(\sigma(\delta/\epsilon)\epsilon)^{-1}\left(\widetilde{\theta}_{n,\epsilon,\delta} - \theta_0\right) \xrightarrow{\mathcal{D}} I^{-1}(\theta_0) \int_0^1 \nabla_\theta b(X_t^0, \theta_0)^\top \, \mathrm{d}B_t,$$

Hence $\widetilde{\theta}_{n,\epsilon,\delta}$ is asymptotically normal.

Remarks:

- Q is written as

$$Q_t = \alpha_\delta t + \sigma W_t + \int_0^t \int_{|z|\le\delta} z\,\widetilde{N}(\mathrm{d}t,\mathrm{d}z) + \int_0^t \int_{|z|>\delta} z\,N(\mathrm{d}t,\mathrm{d}z),$$

where:

  - α_δ, σ: constants;
  - N is a jump-counting measure of Q:

$$N(\mathrm{d}t,\mathrm{d}z) = \sum_{s\in \mathrm{d}t} \mathbf{1}_{\{\Delta Q_s \in \mathrm{d}z\}}, \quad \mathbb{E}\left[N(\mathrm{d}t,\mathrm{d}z)\right] = \nu(\mathrm{d}z)\mathrm{d}t;$$

  - $\widetilde{N}(\mathrm{d}t,\mathrm{d}z) = N(\mathrm{d}t,\mathrm{d}z) - \nu(\mathrm{d}z)\mathrm{d}t$.

- **Fact** (**Asmussen and Rosinski (2001)**): under some regularities,

$$\sigma(\delta)^{-1} Q^{(\delta)} \xrightarrow{\mathcal{D}} B \quad \text{in } \mathbb{D}[0,1], \quad \delta \to 0,$$

for $Q_t^{(\delta)} := \int_0^t \int_{|z|\le\delta} z\,\widetilde{N}(\mathrm{d}s,\mathrm{d}z)$.

Concluding remarks

- We proposed filtered LSEs for drift estimation of SDE with small noises.
- Much more "stable" estimation under finite samples than usual LSE.
- Can the filtered LSEs be asymptotically noraml?
  $\Rightarrow$ Justified under an ideal assumption that all the jumps are specified.
  - Take the threshold $\delta \to 0$ faster!.
  - How to justificate in discretely observed case?
  - If it is possible, then *testing hypothesis* for θ can be easy.
- **Remark**: This LSE-type estimation of drift is possible in the case where

$$\sup_{t\in[0,1]} |Q_t^\epsilon - Q_t| \xrightarrow{\mathbb{P}} 0, \quad Q: \textit{ a semimartingale}$$

under some regularities: e.g.,

$$Q_t^\epsilon = \int_0^t c(X_s^\epsilon)\,\mathrm{d}Q_s + \epsilon R_t, \quad R\text{: something.}$$

Thank you for your attention!