Non-Gaussian quasi-likelihood estimation of jump processes

Hiroki Masuda

Graduate School of Mathematics, Kyushu University, Japan

CREST and Sakigake International Symposium,
Tokyo Institute of Technology University,
December 15, 2010

Brief summary

"Inference for a class of Stochastic Differential Equations (SDE)"

- When observing a discrete-time but high-frequency sample

$$X_0, X_{h_n}, X_{2h_n}, \dots, X_{nh_n} \quad (h_n \to 0)$$

from the semi-parametric Lévy driven SDE

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t,$$

how can we estimate $\theta_0 = (\alpha_0, \gamma_0)$, the true value of $\theta := (\alpha, \gamma)$?

- We will provide an estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ s.t.

$$\left\{\left(\sqrt{n}h_n^{1-1/\beta}(\hat{\alpha}_n - \alpha_0), \sqrt{n}(\hat{\gamma}_n - \gamma_0)\right)\right\}_{n\in\mathbb{N}} \quad \text{is asymp. normal,}$$

with β denoting the Blumental-Getoor index of the Lévy process Z.

Brief summary

"Inference for a class of Stochastic Differential Equations (SDE)"

- When observing a discrete-time but high-frequency sample

$$X_0, X_{h_n}, X_{2h_n}, \dots, X_{nh_n} \quad (h_n \to 0)$$

from the semi-parametric Lévy driven SDE

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t,$$

how can we estimate $\theta_0 = (\alpha_0, \gamma_0)$, the true value of $\theta := (\alpha, \gamma)$?

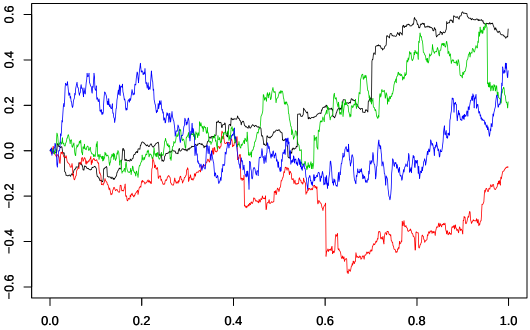
- We will provide an estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ s.t.

$$\left\{ \left(\sqrt{n} h_n^{1-1/\beta} (\hat{\alpha}_n - \alpha_0),\ \sqrt{n}(\hat{\gamma}_n - \gamma_0) \right) \right\}_{n \in \mathbb{N}} \quad \text{is asymp. normal,}$$

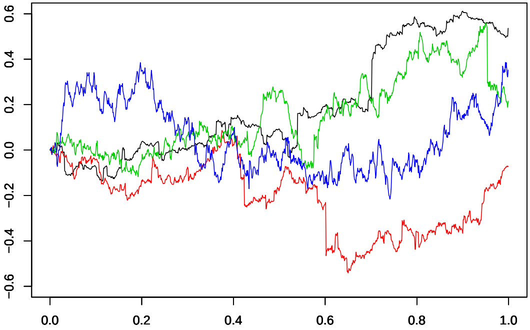
with β denoting the Blumental-Getoor index of the Lévy process Z.

Contents

Statistics for SDE models

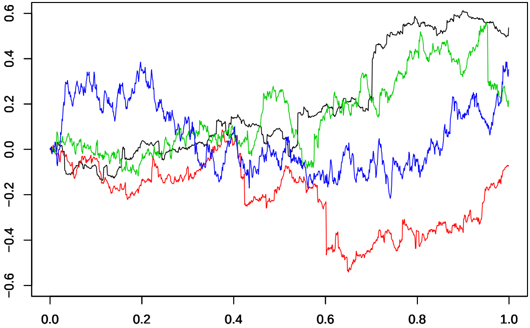


- **Time-varying phenomena $\longleftarrow$ "Stochastic process (SDE) models"**
  - **Mostly, data series exhibits dependence.**
  - **In real world, data is observed at discrete time instants.**



- **“Parameter estimation” is a standing problem in statistics.**
  - **We want a good estimation procedure for a model in question.**

⇒ “Estimation of continuous-time structure from discrete-time sample”.



- **A central issues in stochastic process modelling:**
  - **Continuous?**
  - **Including jumps?**
  - **...or, continuous with jumps?**

Why including jumps?

- Lévy process in finance (Cont and Tankov (2004)): e.g.,
  - Non-Gaussian stable... Mandelbrot (1963)
  - Normal inverse Gaussian... Barndorff-Nielsen (1995)
  - Hyperbolic... Eberlein and Keller (1995)
  - Generalized hyperbolic... Prause (1999), Raible (2000)
  - CGMY (tempered stable)... Carr et al. (2002)
  - Bilateral gamma... Küchler and Tappe (2008)
- Also, signal processing, turbulence, physical science, etc.
  - Non-Gaussian stable... e.g., Nikias and Shao (1995)
  - Semi-heavy tail distributions... Barndoprff-Nielsen (1995)
  - Tempered stable (truncated Lévy flight)... Baeumer and Meerschaert (2010)

$\Rightarrow$ Needs for statistics for jump processes

Why including jumps?

- **Lévy process in finance (Cont and Tankov (2004)): e.g.,**
  - **Non-Gaussian stable... Mandelbrot (1963)**
  - **Normal inverse Gaussian... Barndorff-Nielsen (1995)**
  - **Hyperbolic... Eberlein and Keller (1995)**
  - **Generalized hyperbolic... Prause (1999), Raible (2000)**
  - **CGMY (tempered stable)... Carr et al. (2002)**
  - **Bilateral gamma... Küchler and Tappe (2008)**
- **Also, signal processing, turbulence, physical science, etc.**
  - **Non-Gaussian stable... e.g., Nikias and Shao (1995)**
  - **Semi-heavy tail distributions... Barndoprff-Nielsen (1995)**
  - **Tempered stable (truncated Lévy flight)... Baeumer and Meerschaert (2010)**

$\Rightarrow$ **Needs for statistics for jump processes**

Why including jumps?

- **Lévy process in finance (Cont and Tankov (2004)): e.g.,**
  - **Non-Gaussian stable... Mandelbrot (1963)**
  - **Normal inverse Gaussian... Barndorff-Nielsen (1995)**
  - **Hyperbolic... Eberlein and Keller (1995)**
  - **Generalized hyperbolic... Prause (1999), Raible (2000)**
  - **CGMY (tempered stable)... Carr et al. (2002)**
  - **Bilateral gamma... Küchler and Tappe (2008)**
- **Also, signal processing, turbulence, physical science, etc.**
  - **Non-Gaussian stable... e.g., Nikias and Shao (1995)**
  - **Semi-heavy tail distributions... Barndoprff-Nielsen (1995)**
  - **Tempered stable (truncated Lévy flight)... Baeumer and Meerschaert (2010)**

$\Rightarrow$ **Needs for statistics for jump processes**

Non-Gaussian behavior in small time

- **In high-frequency data framework, jumps may be more conspicuous.**
- **Empirical evidence in financial returns, Grabchak and Samorodnitsky (2010):**
  - **Distribution tails appear to become:**
    - ⋆ **less heavy for less frequent (e.g. monthly) returns,**
    - ⋆ **than for more frequent (e.g. daily) returns.**
  - **Tempered heavy-tail models are reasonable.**

Maximum-Likelihood Estimation for Markov models

- **Maximum-Likelihood Estimator (MLE) is theoretically preferred.**
- **Data $Y_{t_1}, \dots, Y_{t_n}$ from a Markov process (Y_t)**
- **The MLE is defined to be the "argmax" of the log-likelihood function**

$$\theta \mapsto \log p_\theta(Y_{t_1}, \dots, Y_{t_n}) = \sum_{j=1}^{n} \log p_\theta(Y_{t_j} | Y_{t_{j-1}}).$$

- **For Y SDE, the transition density $p_\theta(y|x)$ is mostly unknown. What proxy can we make use of? How can we proceed in practice?**

Maximum-Likelihood Estimation for Markov models

- Maximum-Likelihood Estimator (MLE) is theoretically preferred.
- Data $Y_{t_1}, \dots, Y_{t_n}$ from a Markov process (Y_t)
- The MLE is defined to be the "argmax" of the log-likelihood function

$$\theta \mapsto \log p_\theta(Y_{t_1}, \dots, Y_{t_n}) = \sum_{j=1}^{n} \log p_\theta(Y_{t_j} | Y_{t_{j-1}}).$$

- For Y SDE, the transition density $p_\theta(y|x)$ is mostly unknown. What proxy can we make use of? How can we proceed in practice?

Maximum-Likelihood Estimation for Markov models

- **Maximum-Likelihood Estimator (MLE) is theoretically preferred.**
- **Data $Y_{t_1}, \dots, Y_{t_n}$ from a Markov process (Y_t)**
- **The MLE is defined to be the "argmax" of the log-likelihood function**

$$\theta \mapsto \log p_\theta(Y_{t_1}, \dots, Y_{t_n}) = \sum_{j=1}^{n} \log p_\theta(Y_{t_j} | Y_{t_{j-1}}).$$

- **For Y SDE, the transition density $p_\theta(y|x)$ is mostly unknown. What proxy can we make use of? How can we proceed in practice?**

Maximum-Likelihood Estimation for Markov models

- **Maximum-Likelihood Estimator (MLE) is theoretically preferred.**
- **Data $Y_{t_1}, \dots, Y_{t_n}$ from a Markov process (Y_t)**
- **The MLE is defined to be the "argmax" of the log-likelihood function**

$$\theta \mapsto \log p_\theta(Y_{t_1}, \dots, Y_{t_n}) = \sum_{j=1}^{n} \log p_\theta(Y_{t_j} | Y_{t_{j-1}}).$$

- **For Y SDE, the transition density $p_\theta(y|x)$ is mostly unknown. What proxy can we make use of? How can we proceed in practice?**

Gaussian Quasi-Likelihood Estimator (GQLE)

- Consists of fitting one-step conditional mean and variances:
  - Originally due to Wedderburn (1974);
  - A kind of generalized method of moments.

  To formulate the estimation procedure, it is enough to have

$$E[Y_{t_j}|Y_{t_{j-1}}] = m_{j-1}(\theta) \text{ and } Var[Y_{t_j}|Y_{t_{j-1}}] = v_{j-1}(\theta).$$

  explicitly.

- The GQLE is formally given by the argmax of

$$\theta \mapsto \sum_{j=1}^{n} \log \left\{ \frac{1}{\sqrt{v_{j-1}(\theta)}} \phi\left(Y_{t_j} - m_{j-1}(\theta) \right) \right\},$$

  ϕ the $\mathcal{N}(0,1)$-density.

Gaussian Quasi-Likelihood Estimator (GQLE)

- **Consists of fitting one-step conditional mean and variances:**
  - **Originally due to Wedderburn (1974);**
  - **A kind of generalized method of moments.**

  To formulate the estimation procedure, it is enough to have

$$E[Y_{t_j}|Y_{t_{j-1}}] = m_{j-1}(\theta) \text{ and } Var[Y_{t_j}|Y_{t_{j-1}}] = v_{j-1}(\theta).$$

  explicitly.
- **The GQLE is formally given by the argmax of**

$$\theta \mapsto \sum_{j=1}^{n} \log\left\{ \frac{1}{\sqrt{v_{j-1}(\theta)}} \phi\left(Y_{t_j} - m_{j-1}(\theta) \right) \right\},$$

  ϕ **the** $\mathcal{N}(0,1)$**-density.**

- **The GQLE is popular in "non-Gaussian" noise time-series analysis: e.g., with i.i.d. ϵ_n s.t. $E[\epsilon_n] = 0$ and $E[\epsilon_n^2] = 1$,**
  - **GARCH type model, Straumann and Mikosch (2006),**

$$Y_n = \sigma_n \epsilon_n, \quad n \in \mathbb{N},$$

$$\sigma_n = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{n-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{n-j}^2.$$

  - **Multivariate causal time series, Bardet and Wintenburger (2009),**

$$Y_n = M_\theta(Y_{n-1}, Y_{n-2}, \dots)\epsilon_n + f_\theta(Y_{n-1}, Y_{n-2}, \dots).$$

- **Question. How about using GQLE methodology for the SDE model...?**

- **The GQLE is popular in "non-Gaussian" noise time-series analysis: e.g., with i.i.d. ϵ_n s.t. $E[\epsilon_n] = 0$ and $E[\epsilon_n^2] = 1$,**
  - **GARCH type model, Straumann and Mikosch (2006),**

$$Y_n = \sigma_n \epsilon_n, \quad n \in \mathbb{N},$$

$$\sigma_n = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{n-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{n-j}^2.$$

  - **Multivariate causal time series, Bardet and Wintenburger (2009),**

$$Y_n = M_\theta(Y_{n-1}, Y_{n-2}, \dots)\epsilon_n + f_\theta(Y_{n-1}, Y_{n-2}, \dots).$$

- **Question. How about using GQLE methodology for the SDE model...?**

GQLE for discretely observed Lévy driven SDE *

- Based on $X_{h_n}, X_{2h_n}, \dots, X_{nh_n}$ stemming from the ergodic

$$dX_t = a(X_t, \alpha)dt + c(X_t, \gamma)dZ_t,$$

we want to estimate $\theta = (\alpha, \gamma)$, where Z is a Lévy process s.t. $E[Z_t] = 0$ and $E[Z_t^2] = t$.

- "Aggressive" approximation $\mathcal{L}(Z_{h_n}) \approx \mathcal{N}(0, h_n)$ for small h_n:

$$\begin{aligned} X_{jh_n} &\approx X_{(j-1)h_n} + a(X_{(j-1)h_n}, \alpha_0)h_n \\ &\qquad + c(X_{(j-1)h_n}, \gamma_0)(Z_{jh_n} - Z_{(j-1)h_n}) \\ &\sim \mathcal{N}\left(X_{(j-1)h_n} + a(X_{(j-1)h_n}, \alpha_0)h_n, c(X_{(j-1)h_n}, \gamma_0)^2 h_n\right), \end{aligned}$$

making the GQLE procedure explicit.

*M (2010, preprint) and the references therein.

GQLE for discretely observed Lévy driven SDE *

- **Based on $X_{h_n}, X_{2h_n}, \dots, X_{nh_n}$ stemming from the ergodic**

$$dX_t = a(X_t, \alpha)dt + c(X_t, \gamma)dZ_t,$$

  we want to estimate $\theta = (\alpha, \gamma)$, where Z is a Lévy process s.t. $E[Z_t] = 0$ and $E[Z_t^2] = t$.
- **"Aggressive" approximation $\mathcal{L}(Z_{h_n}) \approx \mathcal{N}(0, h_n)$ for small h_n:**

$$\begin{aligned} X_{jh_n} &\approx X_{(j-1)h_n} + a(X_{(j-1)h_n}, \alpha_0)h_n \\ &\qquad + c(X_{(j-1)h_n}, \gamma_0)(Z_{jh_n} - Z_{(j-1)h_n}) \\ &\sim \mathcal{N}\left(X_{(j-1)h_n} + a(X_{(j-1)h_n}, \alpha_0)h_n, c(X_{(j-1)h_n}, \gamma_0)^2 h_n\right), \end{aligned}$$

  making the GQLE procedure explicit.

*M (2010, preprint) and the references therein.

Resulting phenomenon and a practical caution

$$dX_t = a(X_t, \alpha)dt + c(X_t, \gamma)dZ_t$$

- **The GQLE $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ are asymptotically normal:**

$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V') \quad \text{if } \nu(\mathbb{R}) = 0;$$

$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{nh_n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V'') \quad \text{if } \nu(\mathbb{R}) > 0,$$

where ν is the Lévy measure of Z.

- Existence of "any" jump part in Z slows down the convergence rate.
- An obvious practical problem: Jumps are present or not?

Resulting phenomenon and a practical caution

$$dX_t = a(X_t, \alpha)dt + c(X_t, \gamma)dZ_t$$

- The GQLE $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ are asymptotically normal:

$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V') \quad \text{if } \nu(\mathbb{R}) = 0;$$
$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{nh_n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V'') \quad \text{if } \nu(\mathbb{R}) > 0,$$

  where ν is the Lévy measure of Z.
- Existence of "any" jump part in Z slows down the convergence rate.
- An obvious practical problem: Jumps are present or not?

Resulting phenomenon and a practical caution

$$dX_t = a(X_t, \alpha)dt + c(X_t, \gamma)dZ_t$$

- The GQLE $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ are asymptotically normal:

$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V') \quad \text{if } \nu(\mathbb{R}) = 0;$$

$$\left(\sqrt{nh_n}(\hat{\alpha}_n - \alpha_0), \sqrt{nh_n}(\hat{\gamma}_n - \gamma_0)\right) \to^d \mathcal{N}(0, V'') \quad \text{if } \nu(\mathbb{R}) > 0,$$

  where ν is the Lévy measure of Z.
- Existence of "any" jump part in Z slows down the convergence rate.
- An obvious practical problem: Jumps are present or not?

Test statistics for the noise normality

$$\mathcal{T}_n := \frac{n}{6}\left\{\hat{\Phi}_n^{(3)} - \frac{3\sqrt{h_n}}{n}\sum_{j=1}^{n}\partial_x c(X_{(j-1)h_n}, \hat{\gamma}_n)\right\}^2 + \frac{n}{24}(\hat{\Phi}_n^{(4)} - 3)^2$$

$$\hat{\epsilon}_{nj} := \frac{X_{jh_n} - X_{(j-1)h_n} - a(X_{(j-1)h_n}, \hat{\alpha}_n)h_n}{c(X_{(j-1)h_n}, \hat{\gamma}_n)\sqrt{h_n}}, \quad \bar{\hat{\epsilon}}_n := \frac{1}{n}\sum_{j=1}^{n}\hat{\epsilon}_{nj},$$

$$\hat{\Psi}_n^{(k)} := \frac{1}{n}\sum_{j=1}^{n}(\hat{\epsilon}_{nj} - \bar{\hat{\epsilon}}_n)^k, \quad \hat{\Phi}_n^{(k)} := \frac{\hat{\Psi}_n^{(k)}}{(\hat{\Psi}_n^{(2)})^{k/2}}.$$

- Consistent and asymptotically distribution-free test:
  - $\mathcal{T}_n \to^d \chi^2(2)$ under $\mathcal{H}_0 : \nu(\mathbb{R}) = 0$;
  - $\mathcal{T}_n \to^p \infty$ under $\mathcal{H}_1 : \nu(\mathbb{R}) \in (0, \infty]$.
- We may proceed as follows: Using $\mathcal{T}_n$ with the GQLE,
  - $\mathcal{H}_0$ not rejected $\Rightarrow$ follow diffusion estimation procedures,
  - $\mathcal{H}_0$ rejected $\Rightarrow$ incorporate a jump part, or adopt a pure-jump noise.

Test statistics for the noise normality

$$\mathcal{T}_n := \frac{n}{6}\left\{\hat{\Phi}_n^{(3)} - \frac{3\sqrt{h_n}}{n}\sum_{j=1}^n \partial_x c(X_{(j-1)h_n}, \hat{\gamma}_n)\right\}^2 + \frac{n}{24}(\hat{\Phi}_n^{(4)} - 3)^2$$

$$\hat{\epsilon}_{nj} := \frac{X_{jh_n} - X_{(j-1)h_n} - a(X_{(j-1)h_n}, \hat{\alpha}_n)h_n}{c(X_{(j-1)h_n}, \hat{\gamma}_n)\sqrt{h_n}}, \quad \bar{\hat{\epsilon}}_n := \frac{1}{n}\sum_{j=1}^n \hat{\epsilon}_{nj},$$

$$\hat{\Psi}_n^{(k)} := \frac{1}{n}\sum_{j=1}^n (\hat{\epsilon}_{nj} - \bar{\hat{\epsilon}}_n)^k, \quad \hat{\Phi}_n^{(k)} := \frac{\hat{\Psi}_n^{(k)}}{(\hat{\Psi}_n^{(2)})^{k/2}}.$$

- **Consistent and asymptotically distribution-free test:**
  - $\mathcal{T}_n \to^d \chi^2(2)$ under $\mathcal{H}_0 : \nu(\mathbb{R}) = 0$;
  - $\mathcal{T}_n \to^p \infty$ under $\mathcal{H}_1 : \nu(\mathbb{R}) \in (0, \infty]$.
- We may proceed as follows: Using $\mathcal{T}_n$ with the GQLE,
  - $\mathcal{H}_0$ not rejected $\Rightarrow$ follow diffusion estimation procedures,
  - $\mathcal{H}_0$ rejected $\Rightarrow$ incorporate a jump part, or adopt a pure-jump noise.

Test statistics for the noise normality

$$\mathcal{T}_n := \frac{n}{6}\left\{\hat{\Phi}_n^{(3)} - \frac{3\sqrt{h_n}}{n}\sum_{j=1}^{n}\partial_x c(X_{(j-1)h_n},\hat{\gamma}_n)\right\}^2 + \frac{n}{24}(\hat{\Phi}_n^{(4)} - 3)^2$$

$$\hat{\epsilon}_{nj} := \frac{X_{jh_n} - X_{(j-1)h_n} - a(X_{(j-1)h_n},\hat{\alpha}_n)h_n}{c(X_{(j-1)h_n},\hat{\gamma}_n)\sqrt{h_n}}, \quad \bar{\hat{\epsilon}}_n := \frac{1}{n}\sum_{j=1}^{n}\hat{\epsilon}_{nj},$$

$$\hat{\Psi}_n^{(k)} := \frac{1}{n}\sum_{j=1}^{n}(\hat{\epsilon}_{nj} - \bar{\hat{\epsilon}}_n)^k, \quad \hat{\Phi}_n^{(k)} := \frac{\hat{\Psi}_n^{(k)}}{(\hat{\Psi}_n^{(2)})^{k/2}}.$$

- Consistent and asymptotically distribution-free test:
  - $\mathcal{T}_n \to^d \chi^2(2)$ under $\mathcal{H}_0 : \nu(\mathbb{R}) = 0$;
  - $\mathcal{T}_n \to^p \infty$ under $\mathcal{H}_1 : \nu(\mathbb{R}) \in (0,\infty]$.
- We may proceed as follows: Using $\mathcal{T}_n$ with the GQLE,
  - $\mathcal{H}_0$ not rejected $\Rightarrow$ follow diffusion estimation procedures,
  - $\mathcal{H}_0$ rejected $\Rightarrow$ incorporate a jump part, or adopt a pure-jump noise.

Some important previous studies, some remarks

- **Jump detection filter may work well.**
  (Mancini, Shimizu and Yoshida, Shimizu, Ogihara and Yoshida.)
  - **Asymptotically efficient, may work well for compound Poisson jumps.**
  - **In principle, the coexistence of Wiener and Poisson parts makes estimation problem difficult when pursuing estimation efficiency.**
- **What will theoretically occur in general?**
  - **We do not known any general optimal behavior of estimators.**
  - **LAN results known only for very particular cases.**

Some important previous studies, some remarks

- **Jump detection filter may work well.**
  (Mancini, Shimizu and Yoshida, Shimizu, Ogihara and Yoshida.)
  - **Asymptotically efficient, may work well for compound Poisson jumps.**
  - **In principle, the coexistence of Wiener and Poisson parts makes estimation problem difficult when pursuing estimation efficiency.**
- **What will theoretically occur in general?**
  - **We do not known any general optimal behavior of estimators.**
  - **LAN results known only for very particular cases.**

Our goal of this talk is to

- **Provide an estimator of the true value of $\theta = (\alpha, \gamma)$ in**

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

  based on $X_0, X_{h_n}, X_{2h_n}, \dots, X_{nh_n}$ ($h_n \to 0$).
- **We want to deal with pure-jump Z with higher degree of activity; e.g. Generalized hyperbolic, Meixner, CGMY, etc.**
- **We here do not adopt:**
  - **the GQLE, unsatisfactory while usable, in the presence of any jump;**
  - **the jump detection filter approach, a nice device with a good choice of fine-tuning parameter**
    - ★ **under the presence of a Wiener part,**
    - ★ **when jump activity is finite (or moderate).**

Contents

Non-Gaussian Quasi-Likelihood Estimation (NGQLE)

Target:

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t, \quad \eta := \mathcal{L}(X_0)$$

- Z is a **pure-jump Lévy process of infinite activity.**
- **The parameter** $\theta := (\alpha, \gamma) \in \Theta_\alpha \times \Theta_\gamma = \Theta \subset \mathbb{R}^p$, **a bounded convex domain, the true value** $\theta_0 := (\alpha_0, \gamma_0) \in \Theta$.

Notation:

- $\Delta_j Y := Y_{jh_n} - Y_{(j-1)h_n}$ **for a process** Y;
- $f_{j-1}(\theta) := f(X_{(j-1)h_n}, \theta)$ **for any function of the form** $f(x, \theta)$.

A1. Regularity of the coefficients

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

1. a and c are smooth in $\mathbb{R} \times \Theta$.
2. $a(\cdot, \alpha_0)$ and $c(\cdot, \gamma_0)$ are globally Lipschitz.
3. $\exists c \in (1, \infty)$ s.t. $\forall (x, \gamma)$: $0 < c^{-1} \leq c(x, \gamma) \leq c$.
4. If X is not a Lévy process, then $\exists c', M > 0$ s.t. $\forall |x| \geq M$: $xa(x, \alpha_0) \leq -c'|x|^2$.

* X is then ergodic under the true image measure P_0, the invariant measure denoted by $\pi_0(dx)$.

A2. Driving noise

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

1. $\mathcal{L}(Z_1)$ is symmetric around 0, and the Lévy measure ν of Z fulfils

$$\nu(dz) = {}^{\exists}g_0(z)dz \quad \text{s.t.} \quad g_0(z) = \frac{c_0}{|z|^{1+\beta}}\{1 + O(|z|)\}, \quad |z| \to 0.$$

* $\mathcal{L}(h^{-1/\beta}Z_h) \underset{h\to 0}{\Rightarrow}$ β-stable law with the C.F. $u \mapsto \exp(-|u|^\beta)$ for some $\beta \in (0,2)$: ϕ_β denotes the density.

2. $\mathcal{L}(h^{-1/\beta}Z_h)$ admits a positive density $f_h(y)$ s.t.:
There exist constant $\epsilon_n \to 0$ and Lebesgue-integrable λ s.t.

$$\sqrt{n}\int |f_h(y) - \phi_\beta(y)|dy \to 0.$$

* This holds for, e.g., the NIG Z if $nh_n^{2-\kappa} \to 0$ for some $\kappa > 0$.

A2. Driving noise

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

1. $\mathcal{L}(Z_1)$ is symmetric around 0, and the Lévy measure ν of Z fulfils

$$\nu(dz) = {}^{\exists} g_0(z)dz \quad \text{s.t.} \quad g_0(z) = \frac{c_0}{|z|^{1+\beta}}\{1 + O(|z|)\}, \quad |z| \to 0.$$

* $\mathcal{L}(h^{-1/\beta}Z_h) \underset{h\to 0}{\Rightarrow}$ β-stable law with the C.F. $u \mapsto \exp(-|u|^{\beta})$ for some $\beta \in (0,2)$: ϕ_β denotes the density.

2. $\mathcal{L}(h^{-1/\beta}Z_h)$ admits a positive density $f_h(y)$ s.t.:
There exist constant $\epsilon_n \to 0$ and Lebesgue-integrable λ s.t.

$$\sqrt{n}\int |f_h(y) - \phi_\beta(y)|dy \to 0.$$

* This holds for, e.g., the NIG Z if $nh_n^{2-\kappa} \to 0$ for some $\kappa > 0$.

A3. Sampling rate

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

1. $\beta \geq 1$ if X is a Lévy process (we do not need $nh_n \to \infty$).
2. Otherwise, $\beta > 1$, $nh_n \to \infty$, and
   $\exists \epsilon_0 > 0$ s.t. $\limsup_{n\to\infty} nh_n^{3-2/\beta-\epsilon_0} < \infty$.

A4. Weight function; for heavy-tailed cases

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

1. $W : \mathbb{R} \to \mathbb{R}_+$ **is bounded.**
2. **There exists a function** $K : \mathbb{R} \to \mathbb{R}_+$ **s.t.**
   1. $\sup_{\theta\in\Theta} W(x)\{|\partial_\alpha a(x,\alpha)| + |\partial_\alpha a(x,\alpha)|^2 + |\partial_\alpha^2 a(x,\alpha)| + |\partial_\gamma c(x,\gamma)| + |\partial_\gamma c(x,\gamma)|^2 + |\partial_\gamma^2 c(x,\gamma)|\} \le K(x)$,
   2. $\sup_{t\in\mathbb{R}_+} E_0[K(X_t)] < \infty$.

A5. Nonsingularity and identifiability

For $g(y) := \frac{\partial\phi_\beta}{\phi_\beta}(y)$,

1. $\det\left\{ \int W(x)\frac{[\partial_\alpha a(x,\alpha_0)]^{\otimes 2}}{c(x,\gamma_0)^2}\pi_0(dx)\right\}\cdot\det\left\{ \int W(x)\frac{[\partial_\gamma c(x,\gamma_0)]^{\otimes 2}}{c(x,\gamma_0)^2}\pi_0(dx)\right\} \neq 0.$
2. $\iint W(x)\frac{\partial_\alpha a(x,\alpha)}{c(x,\gamma)^2}\{a(x,\alpha_0)-a(x,\alpha)\}\partial g(\frac{c(x,\gamma_0)}{c(x,\gamma)}y)\phi_\beta(y)dy\pi_0(dx)=0$ **iff** $\theta=\theta_0$.
3. $\iint W(x)\frac{\partial_\gamma c(x,\gamma)}{c(x,\gamma)}\left\{1+\frac{c(x,\gamma_0)}{c(x,\gamma)}yg(\frac{c(x,\gamma_0)}{c(x,\gamma)}y)\right\}\phi_\beta(y)dy\pi_0(dx)=0$ **iff** $\theta=\theta_0$.

Construction of our estimator

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

- **Again, the naive Euler type approximation:**

$$\begin{aligned} X_{jh_n} &\approx^{P_0} X_{(j-1)h_n} + a_{j-1}(\alpha_0)h_n + c_{j-1}(\gamma_0)\Delta_j Z \\ &= X_{(j-1)h_n} + a_{j-1}(\alpha_0)h_n + c_{j-1}(\gamma_0)h_n^{1/\beta} \cdot \frac{\Delta_j Z}{h_n^{1/\beta}} \end{aligned}$$

$$\therefore \epsilon_{nj}(\theta_0) := \frac{\Delta_j X - a_{j-1}(\alpha_0)h_n}{h_n^{1/\beta} c_{j-1}(\gamma_0)} \approx \beta\text{-stable, in law (density } \phi_\beta\text{).}$$

- **We define our estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ through the quasi-likelihood:**

$$\hat{\theta}_n \in \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{j=1}^{n} W_{j-1} \log \left\{ \frac{1}{h_n^{1/\beta} c_{j-1}(\gamma)} \phi_\beta\left(\epsilon_{nj}(\theta)\right) \right\}.$$

Construction of our estimator

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t$$

- **Again, the naive Euler type approximation:**

$$\begin{aligned} X_{jh_n} &\approx^{P_0} X_{(j-1)h_n} + a_{j-1}(\alpha_0)h_n + c_{j-1}(\gamma_0)\Delta_j Z \\ &= X_{(j-1)h_n} + a_{j-1}(\alpha_0)h_n + c_{j-1}(\gamma_0)h_n^{1/\beta} \cdot \frac{\Delta_j Z}{h_n^{1/\beta}} \end{aligned}$$

$$\therefore \epsilon_{nj}(\theta_0) := \frac{\Delta_j X - a_{j-1}(\alpha_0)h_n}{h_n^{1/\beta} c_{j-1}(\gamma_0)} \approx \beta\text{-stable, in law (density } \phi_\beta).$$

- **We define our estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ through the quasi-likelihood:**

$$\hat{\theta}_n \in \operatorname*{argmax}_{\theta \in \Theta} \sum_{j=1}^{n} W_{j-1} \log \left\{ \frac{1}{h_n^{1/\beta} c_{j-1}(\gamma)} \phi_\beta\left(\epsilon_{nj}(\theta)\right) \right\}.$$

Main claim: Asymptotic Normality

Under the aforementioned assumptions, the estimator is A.N.:

$$\left(\sqrt{n}h_n^{1-1/\beta}(\hat{\alpha}_n-\alpha_0), \sqrt{n}(\hat{\gamma}_n-\gamma_0)\right) \Rightarrow \mathcal{N}\left(0, \mathrm{diag}[U(\theta_0)^{-1}, V(\theta_0)^{-1}]\right),$$

where

$$U(\theta_0)=\int W(x)\frac{\{\partial_\alpha a(x,\alpha_0)\}^{\otimes 2}}{c(x,\gamma_0)^2}\pi_0(dx)\cdot\int\frac{\partial\phi_\beta(y)^2}{\phi_\beta(y)}dy,$$

$$V(\theta_0)=\int W(x)\frac{\{\partial_\gamma c(x,\gamma_0)\}^{\otimes 2}}{c(x,\gamma_0)^2}\pi_0(dx)\cdot\int\frac{\{\phi_\beta(y)+y\partial\phi_\beta(y)\}^2}{\phi_\beta(y)}dy$$

Remarks

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t.$$

Contrast	Rates	
	α	γ
Gaussian QL	$\sqrt{nh_n}$	$\sqrt{nh_n}$
Non-Gaussian (Stable) QL	$\sqrt{n}h_n^{1-1/\beta}$	$\sqrt{n}$

- GQLE is easier to use, but NGQLE has better performance.
- Both are somewhat robust for the specification of the Lévy measure.
- The technical conditions imposed are, unfortunately, not so mild.
- However, we conjecture that the NGQLE is asymptotically optimal.

Remarks

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t.$$

Contrast	Rates α	Rates γ
Gaussian QL	$\sqrt{nh_n}$	$\sqrt{nh_n}$
Non-Gaussian (Stable) QL	$\sqrt{n}h_n^{1-1/\beta}$	$\sqrt{n}$

- GQLE is easier to use, but NGQLE has better performance.
- Both are somewhat robust for the specification of the Lévy measure.
- The technical conditions imposed are, unfortunately, not so mild.
- However, we conjecture that the NGQLE is asymptotically optimal.

Remarks

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t.$$

Contrast	Rates α	Rates γ
Gaussian QL	$\sqrt{nh_n}$	$\sqrt{nh_n}$
Non-Gaussian (Stable) QL	$\sqrt{n}h_n^{1-1/\beta}$	$\sqrt{n}$

- **GQLE is easier to use, but NGQLE has better performance.**
- **Both are somewhat robust for the specification of the Lévy measure.**
- **The technical conditions imposed are, unfortunately, not so mild.**
- **However, we conjecture that the NGQLE is asymptotically optimal.**

Remarks

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t.$$

Contrast	Rates α	Rates γ
Gaussian QL	$\sqrt{nh_n}$	$\sqrt{nh_n}$
Non-Gaussian (Stable) QL	$\sqrt{n}h_n^{1-1/\beta}$	$\sqrt{n}$

- GQLE is easier to use, but NGQLE has better performance.
- Both are somewhat robust for the specification of the Lévy measure.
- The technical conditions imposed are, unfortunately, not so mild.
- However, we conjecture that the NGQLE is asymptotically optimal.

Remarks

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dZ_t.$$

Contrast	Rates α	Rates γ
Gaussian QL	$\sqrt{nh_n}$	$\sqrt{nh_n}$
Non-Gaussian (Stable) QL	$\sqrt{n}h_n^{1-1/\beta}$	$\sqrt{n}$

- GQLE is easier to use, but NGQLE has better performance.
- Both are somewhat robust for the specification of the Lévy measure.
- The technical conditions imposed are, unfortunately, not so mild.
- However, we conjecture that the NGQLE is asymptotically optimal.

A small numerical example: NIG Lévy process

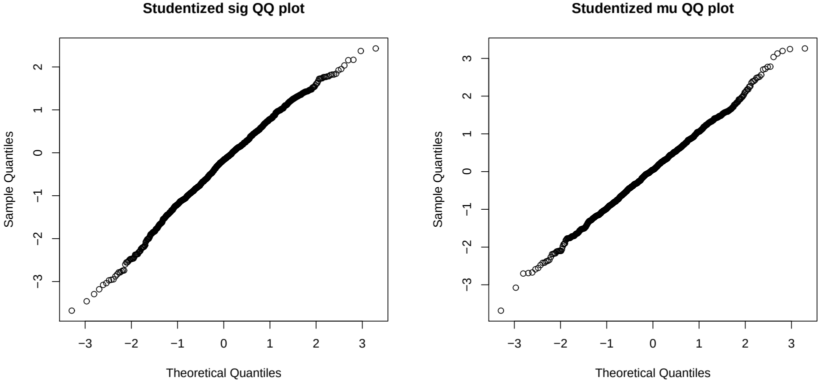
- We set $X_t = \alpha t + \gamma Z_t$ with $\mathcal{L}(Z_t) = NIG(a, 0, t, 0)$ for some (unknown) $a > 0$, hence

$$\frac{X_t - \alpha t}{\gamma t} \sim NIG(at, 0, 1, 0) \to^d \text{standard Cauchy.}$$

- $\theta_0 = (\alpha_0, \gamma_0) \leftarrow (-3, 2)$, $\beta = 1$, and $a = 2$.
- 1000 iterations with $n = 500$ and $h_n = 1/n$.
- Results.

	Sample median	Stable QLE α	Stable QLE γ
Mean	-2.9961	-2.9942	1.9781
S.D.	0.1430	0.1272	0.1237
Max	-2.5186	-2.5852	2.3635
Min	-3.4808	-3.4704	1.6225

Achieving the normality of the NGQLE



Summary and concluding remarks

- **The essential assumption: $\mathcal{L}(h^{-1/\beta}Z_h)$ is approximately β-stable.**
- Without imposing $nh_n \to \infty$ for all cases?: A suitable weak limit theorem is necessary for identifying possible limit distribution.
- Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.
- Estimation of the Blumental-Getoor index β: For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- Large deviation for the random fields, giving convergence of moments?
- Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)

Summary and concluding remarks

- The essential assumption: $\mathcal{L}(h^{-1/\beta}Z_h)$ is approximately β-stable.
- Without imposing $nh_n \to \infty$ for all cases?:
  A suitable weak limit theorem is necessary for identifying possible limit distribution.
- Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.
- Estimation of the Blumental-Getoor index β:
  For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- Large deviation for the random fields, giving convergence of moments?
- Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)

Summary and concluding remarks

- **The essential assumption: $\mathcal{L}(h^{-1/\beta} Z_h)$ is approximately β-stable.**
- **Without imposing $nh_n \to \infty$ for all cases?:**
  A suitable weak limit theorem is necessary for identifying possible limit distribution.
- **Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.**
- **Estimation of the Blumental-Getoor index β:**
  For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- **Large deviation for the random fields, giving convergence of moments?**
- **Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)**

Summary and concluding remarks

- **The essential assumption: $\mathcal{L}(h^{-1/\beta} Z_h)$ is approximately β-stable.**
- **Without imposing $nh_n \to \infty$ for all cases?:**
  A suitable weak limit theorem is necessary for identifying possible limit distribution.
- **Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.**
- **Estimation of the Blumental-Getoor index β:**
  For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- **Large deviation for the random fields, giving convergence of moments?**
- **Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)**

Summary and concluding remarks

- **The essential assumption: $\mathcal{L}(h^{-1/\beta}Z_h)$ is approximately β-stable.**
- **Without imposing $nh_n \to \infty$ for all cases?:**
  A suitable weak limit theorem is necessary for identifying possible limit distribution.
- **Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.**
- **Estimation of the Blumental-Getoor index β:**
  For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- **Large deviation for the random fields, giving convergence of moments?**
- **Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)**

Summary and concluding remarks

- **The essential assumption: $\mathcal{L}(h^{-1/\beta} Z_h)$ is approximately β-stable.**
- **Without imposing $nh_n \to \infty$ for all cases?:**
  A suitable weak limit theorem is necessary for identifying possible limit distribution.
- **Want to utilize the Cauchy quasi-likelihood ($\beta = 1$) for SDE.**
- **Estimation of the Blumental-Getoor index β:**
  For Lévy driven OUP, we can apply LAD type estimate (M, 2010).
- **Large deviation for the random fields, giving convergence of moments?**
- **Adaptive estimation for jump SDEs? (Uchida and Yoshida (2010) for diffusions)**