On variants of stable quasi-likelihood for Lévy driven SDE

Hiroki Masuda

Kyushu University, JST CREST

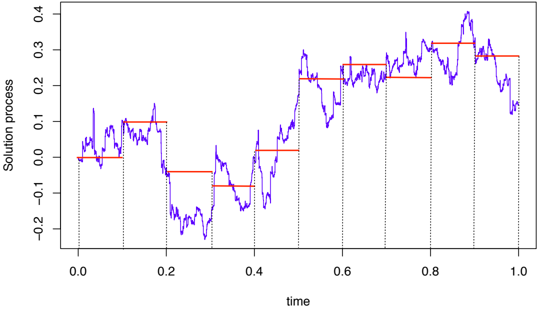
S.A.P.S. $\mathbb{X}$
Le Mans
March 19, 2015

1. Objective
2. SQMLE asymptotics
3. Concluding remarks

Objective: Estimation of $\theta = (\alpha, \gamma)$

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

- **Locally β-stable Lévy process J ($\beta < 2$)**
- **High-frequency data $(X_{t_j})_{j=0}^{n}$, $t_j = t_j^n = jh_n$ for $h_n := T/n$**



1 Objective

2 SQMLE asymptotics

3 Concluding remarks

Goal: Estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$ better than the Gaussian QMLE

$$\left\{\sqrt{n}h_n^{1-1/\beta}(\hat{\alpha}_n - \alpha_0),\ \sqrt{n}(\hat{\gamma}_n - \gamma_0)\right\} \xrightarrow{\mathcal{L}} \text{Mixed Normal.}$$

- **Gaussian QMLE (and/or LSE) does NOT work here (M, 2013).**

Assumptions (1/2)

A1. Regularity of the coefficients (can be weakened)

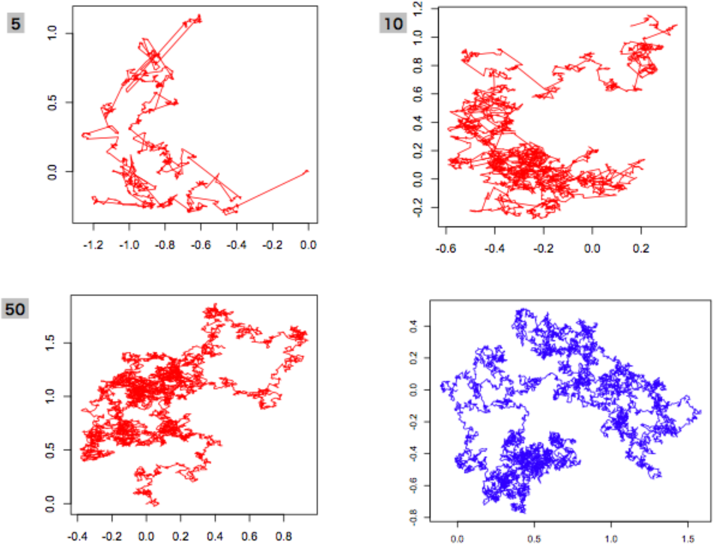
1. a and c are smooth, $a(\cdot, \alpha_0)$ and $c(\cdot, \gamma_0)$ are globally Lipschitz.
2. $\exists c \in (1, \infty)\ \forall (x, \gamma)$: $c^{-1} \le c(x, \gamma) \le c$.

A2. Driving-noise structure

1. $\mathcal{L}(h^{-1/\beta} J_h) \underset{h \to 0}{\Rightarrow}$ β-stable, the C.F. $u \mapsto e^{-|u|^\beta}$ (the pdf $\phi_\beta(\cdot)$)

   e.g. Stable, Tempered stable, Generalized hyperbolic, Meixner, Generalized-z, etc.
2. Admissible range of β:
   - $\beta \in [1, 2)$ if X is a Lévy process;
   - $\beta \in (1, 2)$ if $c(x, \gamma) = \gamma$;
   - $\beta \in (4/3, 2)$ otherwise.
3. The Leb. density f_h of $\mathcal{L}(h^{-1/\beta} J_h)$ fulfils $\sqrt{n} \int |f_h(y) - \phi_\beta(y)| dy \to 0$

NIG processes & Wiener process



Construction of estimator: two approximations

- **Locally stable (non-Gaussian!) approximation:**

$$\epsilon_{nj}(\theta;\beta) := \frac{\Delta_j X - a_{j-1}(\alpha)h_n}{h_n^{1/\beta} c_{j-1}(\gamma)} \approx \frac{\Delta_j J}{h_n^{1/\beta}} \approx \text{(i.i.d. } \beta\text{-stable)}$$

- **Transition density approximation: in the Fourier inversion,**

$$p_n(X_{jh_n} | X_{(j-1)h_n};\theta) \approx \frac{1}{c_{j-1}(\gamma)h_n^{1/\beta}} f_{h_n}(\epsilon_j(\theta;\beta)) \quad \text{(1. Euler approx.)}$$

$$\approx \frac{1}{c_{j-1}(\gamma)h_n^{1/\beta}} \phi_\beta(\epsilon_j(\theta;\beta)) \quad \text{(2. Stable approx.)}$$

Stable Quasi-Maximum Likelihood Estimator $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\gamma}_n)$; SQMLE

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

Assumptions (2/2)

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

A3. Identifiability

$(\alpha,\gamma) = (\alpha_0,\gamma_0)$ iff we have a.s. both:

1. $\displaystyle\int_0^1 \frac{\partial_\alpha a(X_t,\alpha)}{c(X_t,\gamma)^2}\{a(X_t,\alpha_0) - a(X_t,\alpha)\}\int \partial g\left(\frac{c(X_t,\gamma_0)}{c(X_t,\gamma)}y\right)\phi_\beta(y)dydt = 0$
2. $\displaystyle\int_0^1 \frac{\partial_\gamma c(X_t,\gamma)}{c(X_t,\gamma)}\int\left\{1 + \frac{c(X_t,\gamma_0)}{c(X_t,\gamma)}yg\left(\frac{c(X_t,\gamma_0)}{c(X_t,\gamma)}y\right)\right\}\phi_\beta(y)dydt = 0$

Asymptotic behavior of SQMLE

$$dX_t = a(X_t,\alpha)dt + c(X_{t-},\gamma)dJ_t, \quad (X_{jT/n})_{j=0}^n$$

Theorem 1 (Asymptotic Mixed Normality of the SQMLE)

$$\begin{pmatrix} n^{1/\beta-1/2}(\hat{\alpha}_n-\alpha_0) \\ \sqrt{n}(\hat{\gamma}_n-\gamma_0) \end{pmatrix} \xrightarrow{\mathcal{L}} MN\left(0, \mathrm{diag}[\Sigma_{T,\alpha}(\theta_0)^{-1}, \Sigma_{T,\gamma}(\theta_0)^{-1}]\right)$$

$$\Sigma_{T,\alpha}(\theta_0) := T^{2(1-1/\beta)}\frac{1}{T}\int_0^T \frac{\{\partial_\alpha a(X_t,\alpha_0)\}^{\otimes 2}}{c(X_t,\gamma_0)^2}dt \cdot \int \frac{\{\partial\phi_\beta(y)\}^2}{\phi_\beta(y)}dy,$$

$$\Sigma_{T,\gamma}(\theta_0) := \frac{1}{T}\int_0^T \frac{\{\partial_\gamma c(X_t,\gamma_0)\}^{\otimes 2}}{c(X_t,\gamma_0)^2}dt \cdot \int \frac{\{\phi_\beta(y)+y\partial\phi_\beta(y)\}^2}{\phi_\beta(y)}dy$$

- Under $t_n \equiv T$, no ergodicity condition and no unit-root problem.
- Efficiency? Clément and Gloter (2015, SPA; LAMN), Ivanenko et al. (2014, arxiv)
- β is assumed to be known...

Implementation and computation

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

1. **Unknown β contained:**
   - **Joint optimization w.r.t. (θ, β) leads to the singularity problem: The asymptotic Fisher information $|\mathcal{I}(\beta,\gamma)| \equiv 0$ for β-stable Lévy process (Aït-Sahalia and Jacod, 2008; HM, 2009).**

$\Rightarrow$ **Want to make a $\hat{\beta}_n$ separately and plug-in it.**

2. **Time-consuming optimization through ϕ_β plugged-in by $\epsilon_{nj}(\theta;\beta)$:**
   - **Libraries available (R, MATLAB, ...), but rather heavy.**

$\Rightarrow$ **Want to replace ϕ_β with a more handy one.**

A simplified model setting

- Suppose A1 and A2 for the model

$$dX_t = a(X_t, \alpha)dt + \gamma dJ_t, \qquad (X_{jT/n})_{j=0}^{n}.$$

A3′. Fake SQMLE (but it works!)

$$\hat{\alpha}_n^* \in \operatorname{argmax}_{\alpha} \sum_{j=1}^{n} \log \psi \left(h_n^{-1/\hat{\beta}_n} (\Delta_j X - h_n a_{j-1}(\alpha)) \right)$$

- ψ even, $\log \psi \in \mathcal{C}_b^4(\mathbb{R})$, $\sup_y |y|^k |\partial_y^k \log \psi(y)| < \infty$ $(k \geq 0)$.
- Identifiability conditions (ψ analogue to A3).
- $\sqrt{n}(\hat{\beta}_n - \beta_0) = O_p(1)$ and $\beta \in (1, 4/3)$.

Bypassing heavy computation load (very special case)

Theorem 2 (Asymptotic mixed normality of the modified estimator)

A1, A2, A3′ $\Rightarrow n^{1/\beta-1/2}(\hat{\alpha}_n^* - \alpha_0)$ is asymptotically mixed-normal.

- Todorov-Tauchen and Todorov $\sqrt{n}$-consistent $\hat{\beta}_n$ (2011, 2013):
  - Partly used the 2nd-order power-variation statistics.
- Cauchy (possibly fake!) quasi-likelihood is the case:

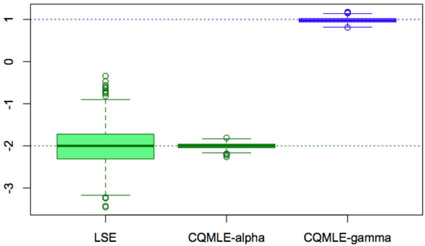
$$\log\psi(x) \leftarrow -\log(1+x^2).$$

- Unfortunately, estimation of γ in such a way would lead to a biased estimate; try another martingale estimating function (Z-estimation)?

Example 1: Normal inverse Gaussian (NIG) Lévy process

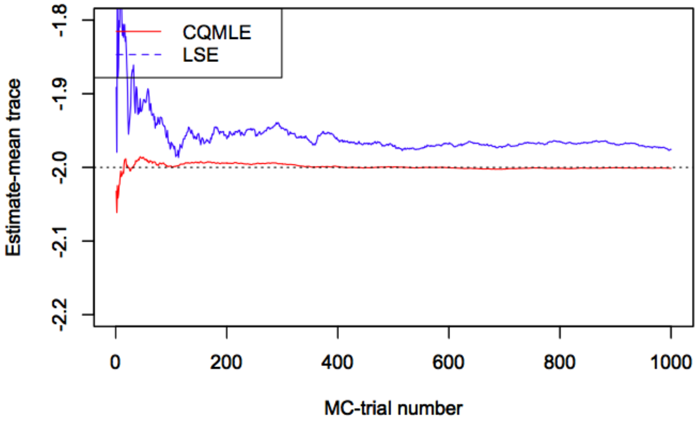
- $X_t = \alpha t + \gamma J_t$ with $\mathcal{L}(J_t) = NIG(a, 0, t, 0)$ for (unknown) $a > 0$: then, $(\gamma t)^{-1}(X_t - \alpha t) \sim NIG(at, 0, 1, 0) \xrightarrow{\mathcal{L}}$ Cauchy.
- $T = 1$, $\theta_0 = (\alpha_0, \gamma_0) \leftarrow (-2, 1)$, $\beta = 1$, and $a = 5$.
- 1000 Monte Carlo iterations with $n = 500$.

	LSE α	Cauchy QMLE α	Cauchy QMLE γ
Mean	-1.968	-1.997	0.980
S.D.	0.443	0.062	0.063
Max	-0.473	-1.758	1.178
Min	-3.524	-2.244	0.796



Estimation-mean traces

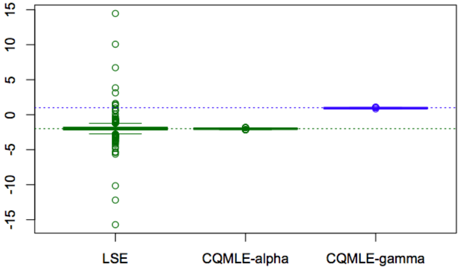
- **Plots of $l \mapsto \frac{1}{l}\sum_{k=1}^{l} \hat{\alpha}_n^{(l)}$ for independent estimates $\{\hat{\alpha}_n^{(l)}\}_{l=1}^{L}$.**



Example 2: β-stable Lévy process

- $X_t = \alpha t + \gamma J_t$ with $\mathcal{L}(J_t) = S_\beta(t^{1/\beta})$.
- $T = 1$, $\theta_0 = (\alpha_0, \gamma_0) \leftarrow (-2, 1)$ and $\beta = 1.3$.
- 1000 Monte Carlo iterations with $n = 1000$.

	LSE α	Cauchy QMLE α	Cauchy QMLE γ
Mean	-1.962	-2.002	0.937
S.D.	1.098	0.050	0.034
Max	14.468	-1.803	1.092
Min	-15.714	-2.170	0.836



1. Objective
2. SQMLE asymptotics
3. Concluding remarks

Summary: Value of β is crucial

Locally stable QMLE under the two approximations

$$dX_t = a(X_t, \alpha)dt + c(X_{t-}, \gamma)dJ_t, \quad (X_{jT/n})_{j=0}^{n},$$

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

- Euler approximation $\Delta_j X \approx a(X_{t_{j-1}}, \alpha)h_n + c(X_{t_{j-1}}, \gamma)\Delta_j J$;
- Small-time stable approximation $\mathcal{L}(h^{-1/\beta} J_h) \Rightarrow S_\beta(1)$.

- $1 \le \beta < 2$ known $\Rightarrow$ need $\beta > 4/3$ for state-dependent $c(x, \gamma)$.
- $1 \le \beta < 2$ unknown $\Rightarrow$ require $\beta < 4/3$ for using $\sqrt{n}$-consistent $\hat{\beta}_n$.

- Estimation of γ would be rather fragile against misspecifications:
  - $c(x, \gamma)$ misspecification;
  - Contrast (ψ) misspecification.
- Cauchy QMLE ($\beta = 1$) $\Rightarrow$ requires separate care (ongoing).

Some references

- Aït-Sahalia, Y. and Jacod, J. (2008), Fisher's information for discretely sampled Lévy processes. Econometrica 76, 727–761.
- Bertoin, J. and Doney, R. A. (1997), Spitzer's condition for random walks and Lévy processes. Ann. Inst. H. Poincaré Probab. Statist. 33, 167–178.
- Clément, E. and Gloter, A. (2015), Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable Lévy processes. Stoch. Proc. Appl.
- Ivanenko, D. O., Kulik, A, M. and Masuda, H. (2014), Uniform LAN property of locally stable Lévy process observed at high frequency. arxiv:1411.1516
- Masuda, H. (2013), Convergence of Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency. Ann. Statist. 41, 1593–1641.
- Masuda, H., Non-Gaussian quasi-likelihood for locally stable processes, in preparation.
- Masuda, H., Optimal estimation of stable Ornstein-Uhlenbeck regression models, in preparation.
- Todorov, V. (2013), Power variation from second order differences for pure jump semimartingales. Stochastic Process. Appl. 123, 2829–2850.
- Todorov, V. and Tauchen, G. (2011), Limit theorems for power variations of pure-jump processes with application to activity estimation. Ann. Appl. Probab. 21, 546–588.