Multi-step estimation procedure for stable Ornstein-Uhlenbeck processes

Hiroki Masuda

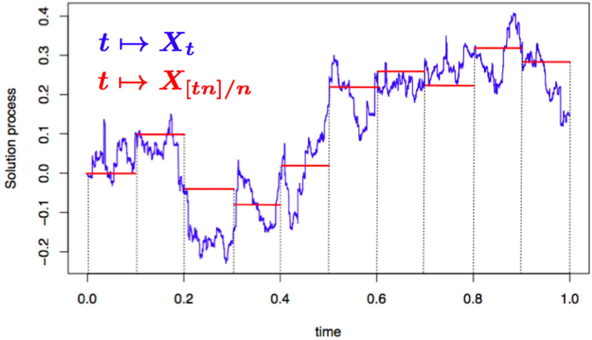
Institute of Mathematics for Industry
Kyushu University

“Stochastic processes and their statistics in Finance”
Okinawa, October 27, 2013

Put simply

- **Parametric estimation**
- **of the stable Ornstein-Uhlenbeck processes**
- **based on discrete-time but high-frequency infill sampling.**

$$\{P_\theta := \mathcal{L}(X) : \theta \in \Theta\}$$



Symmetric β-stable Lévy process J

$$E[\exp(iuJ_t)] = \exp\{-t|u|^\beta\} \sim S_\beta(1), \quad \beta \in (0,2).$$

- **Characterized by the Lévy density:**
  - $g(z) = \frac{1}{2}\sigma^\beta \left\{ \frac{1}{\beta}\Gamma(1-\beta)\cos\frac{\beta\pi}{2} \right\}^{-1} |z|^{-\beta-1}, \ z \neq 0$
- $P(J_1 \in dy) = \phi_\beta(y)dy$:
  - **Positive density:** $\forall y \in \mathbb{R}, \ \phi_\beta(y) > 0$
  - ϕ_β **is smooth in** $(y,\beta) \in \mathbb{R} \times (0,2)$
- **Selfsimilarity:** $J_t \overset{d}{=} t^{1/\beta} J_1$
- **Lack of finite variance:** $E(|J_t|^q) < \infty$ **iff** $q \in (-1, \beta)$.

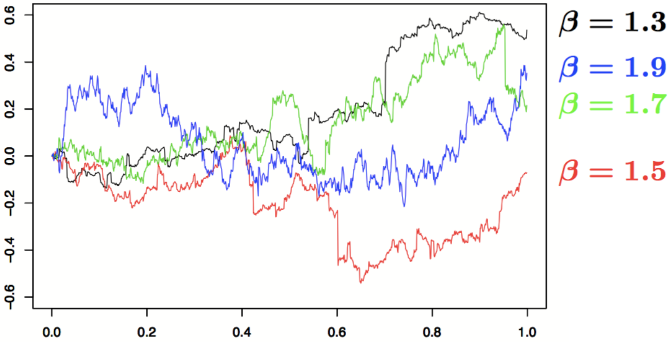
Model setup

Stable-Lévy driven Ornstein-Uhlenbeck process

$$dX_t = -\lambda X_t dt + \sigma dJ_t, \quad X_0 = x_0.$$

- Equidistant sample over a fixed period $[0, T]$:

$$(X_{t_1}, \dots, X_{t_n}), \quad t_j = jh, \quad h = h_n := T/n$$



Goal

$$dX_t = -\lambda X_t dt + \sigma dJ_t, \qquad X_0 = x_0, \ (X_{t_j})_{j=0}^n$$

Local Asymptotic Mixed Normality (LAMN) for $\lambda \in \mathbb{R}$

Local quadratic approximation of the likelihood ratio ($r_n \to \infty$):

$$\log \frac{P_{\lambda+u/r_n}}{P_\lambda}(X_{t_1}, \dots, X_{t_n}) = \Delta_n(T)u - \frac{1}{2}\Gamma_0(T)u^2 + o_p(1),$$

clarifying the ideal asymptotic phenomenon in estimation of λ:

- Optimal rate of convergence;
- Minimal asymptotic variance;
- Efficiency (in test) of the score statistics $\Delta_n(T)$.

Goal

$$dX_t = -\lambda X_t dt + \sigma dJ_t, \qquad X_0 = x_0,\ (X_{t_j})_{j=0}^{n}$$

Local Asymptotic Mixed Normality (LAMN) for $\lambda \in \mathbb{R}$

Local quadratic approximation of the likelihood ratio ($r_n \to \infty$):

$$\log \frac{P_{\lambda+u/r_n}}{P_\lambda}(X_{t_1}, \dots, X_{t_n}) = \Delta_n(T)u - \frac{1}{2}\Gamma_0(T)u^2 + o_p(1),$$

clarifying the ideal asymptotic phenomenon in estimation of λ:

- Optimal rate of convergence;
- Minimal asymptotic variance;
- Efficiency (in test) of the score statistics $\Delta_n(T)$.

Practical estimators for the unknown parameters

MLE is not quite convenient...so resort to something else.

Remark: Case of Gaussian OU process

$$dX_t = -\lambda X_t dt + \sigma dw_t$$

- **LAN or LAMN available for λ only under**

$$T_n := nh_n \to \infty$$

  - **Ergodic ($\lambda > 0$) $\Rightarrow$ LAN (Local Asymptotic Normality).**
  - **Non-ergodic ($\lambda < 0$) $\Rightarrow$ LAMN (Local Asymptotic Mixed Normality).**
  - **Unit-root ($\lambda = 0$): Locally Asymptotically Brownian Functional.**

Our question

Case of the β-stable driven case for $\beta < 2$...?

Remark: Case of Gaussian OU process

$$dX_t = -\lambda X_t dt + \sigma dw_t$$

- **LAN or LAMN available for λ only under**

$$T_n := nh_n \to \infty$$

  - **Ergodic ($\lambda > 0$) $\Rightarrow$ LAN (Local Asymptotic Normality).**
  - **Non-ergodic ($\lambda < 0$) $\Rightarrow$ LAMN (Local Asymptotic Mixed Normality).**
  - **Unit-root ($\lambda = 0$): Locally Asymptotically Brownian Functional.**

Our question

Case of the β-stable driven case for $\beta < 2$...?

Answer

Entirely different phenomena

Remark: How about the Least-Squares Estimator...?

- **The LSE** $\tilde{\lambda}_n := \operatorname*{arginf}_{\lambda>0} \sum_{j=1}^{n} (X_{t_j} - X_{t_{j-1}} + \lambda X_{t_{j-1}} h)^2$.

Remark: How about the Least-Squares Estimator...?

- **The LSE** $\tilde{\lambda}_n := \operatorname*{arginf}_{\lambda>0} \sum_{j=1}^{n} (X_{t_j} - X_{t_{j-1}} + \lambda X_{t_{j-1}} h)^2$.
- **Explicit asymptotic hebavior, cf. Hu and Long (2009):**

$$\left(\frac{T_n}{\log n}\right)^{1/\beta} (\tilde{\lambda}_n - \lambda_0) \to^d \frac{S'_\beta}{S''^{+}_{\beta/2}}.$$

  - **Ergodicity is essential.**
  - $T = T_n \to \infty$ **inevitable.**
  - **Construction of a confidence interval may not be straightforward.**

Transition probability

- For any $0 \leq s \leq t$,

$$X_t = e^{-\lambda(t-s)} X_s + \sigma \int_s^t e^{-\lambda(t-u)} dJ_u.$$

- Stable-integral property:

$$\mathcal{L}\left(\int_{t_{j-1}}^{t_j} e^{-\lambda(t_j-u)} dJ_u \right) = S_\beta(\kappa_h(\lambda)),$$

where

$$\kappa_h(\lambda) := \left\{ \frac{1-\exp(-\lambda h)}{\lambda\beta} \right\}^{1/\beta} \sim h^{1/\beta}$$

- For each $j \leq n$:

$$\mathcal{L}(X_{t_j} | X_{t_{j-1}} = x) = \delta_{x\exp(-\lambda h)} * S_\beta(\sigma\kappa_h(\theta))$$

The likelihood function in question

Loglikelihood (w.r.t. λ)

$$\ell_n(\lambda) = \sum_{j=1}^{n} \log\left\{\frac{1}{\sigma\kappa_h(\lambda)}\phi_\beta(\epsilon_j(\lambda))\right\}$$

where

$$\epsilon_j(\lambda) := \frac{1}{\sigma\kappa_h(\lambda)}\left(X_{t_j} - e^{-\lambda h}X_{t_{j-1}}\right) \overset{P_\lambda}{\sim} S_\beta(1)$$

- **For $k \in \mathbb{N}$,**

$$\begin{aligned}\partial_\lambda^k \ell_n(\lambda) &= \sum_{j=1}^{n}\left\{-\partial_\lambda^k \log\kappa_h(\lambda) + \partial_\lambda^k \log\phi_\beta(\epsilon_j(\lambda))\right\} \\ &= o(1) + \sum_{j=1}^{n} \partial_\lambda^k \log\phi_\beta(\epsilon_j(\lambda)), \quad n \to \infty.\end{aligned}$$

1. Objective and Background

2. Claims

3. Simulation

4. Summary and Remarks

Stochastic expansion: $r_n := n^{1/\beta-1/2}$-rate localization

$$\ell_n\left(\lambda+\frac{u}{r_n}\right)-\ell_n(\lambda)=\frac{u}{r_n}\partial_\lambda\ell_n(\lambda)-\frac{1}{2}\left(-\frac{u^2}{r_n^2}\partial_\lambda^2\ell_n(\lambda)\right)+R_n(u)$$

1. **Stable convergence in law of the martingale term**

$$\exists\text{r.v. } \Delta_0,\quad (r_n^{-1}\partial_\lambda\ell_n(\lambda),F_n)\xrightarrow{\mathcal{L}}(\Delta_0,F_0)\quad \text{for any } F_n\xrightarrow{p}F_0.$$

2. **Law of large numbers for the quadratic term**

$$\exists\text{positive r.v. } \Gamma_0,\quad -r_n^{-2}\partial_\lambda^2\ell_n(\lambda)\xrightarrow{p}\Gamma_0.$$

3. **Negligibility of the remainder term**

$$\forall u,\quad R_n(u)=O_p(n^{-1/2})\xrightarrow{p}0$$

Main claim: LAMN for $\lambda \in \mathbb{R}$ when T is fixed

- $dX_t = -\lambda X_t dt + \sigma dJ_t, \quad \lambda \in \mathbb{R}$, $(X_{jT/n})_{j=0}^n$, $r_n = n^{1/\beta - 1/2}$

Theorem (LAMN)

$$\forall u \in \mathbb{R}, \quad \ell_n\left(\lambda + \frac{u}{r_n}\right) - \ell_n(\lambda) = \Delta_n(T)u - \frac{1}{2}\Gamma_0(T)u^2 + o_p(1)$$

- $\Delta_n(T) \xrightarrow{\mathcal{L}_s} \Delta_0(T) \sim \Gamma_0(T)^{-1/2}\eta$ with $\eta \sim N(0,1) \perp\!\!\!\perp J$.
- $\Gamma_0(T) := \left\{\sigma^{-2}\int\left(\frac{\partial\phi_\beta(y)}{\phi_\beta(y)}\right)^2 \phi_\beta(y)dy\right\} T^{1-2/\beta}\int_0^T X_t^2 dt$.

Some consequences and messages

$$\forall u \in \mathbb{R}, \quad \ell_n\left(\lambda + \frac{u}{r_n}\right) - \ell_n(\lambda) = \Delta_n(T)u - \frac{1}{2}\Gamma_0(T)u^2 + o_p(1)$$

- An efficient $\hat{\lambda}_n$ should fulfil that

$$n^{1/\beta-1/2}(\hat{\lambda}_n - \lambda) \xrightarrow{\mathcal{L}} MN\left(0, \sigma^2 C(\beta) T^{2/\beta-1}\left(\int_0^T X_t^2 dt\right)^{-1}\right)$$

  the asymptotic Mixed Normality;
  - Getting useless as $\beta \to 2$ (should be!),
  - Asymptotic random variance is λ-free, but λ affects $\mathcal{L}(\Gamma_0(T)^{-1})$.
- No unit-root type problem (cf. Szimayer and Maller (2004));
  - Unified asymptotics, whatever λ_0 is, unlike AR time series models.
- Quantitative distributional theory for each fixed T.

Remark: Rate-optimal estimator of $\lambda \in \mathbb{R}$ [1]

- $dX_t = -\lambda X_t dt + \sigma dJ_t, \quad (X_{t_j})_{j=0}^n,\ t_j = jT/n$

LAD (Least Absolute Deviation) estimator

$$\hat{\lambda}_n \leftarrow \operatorname*{argmin}_{\lambda} \sum_{j=1}^{n} \left| X_{t_j} - e^{-\lambda T/n} X_{t_{j-1}} \right|$$

$$n^{1/\beta - 1/2}(\hat{\lambda}_n - \lambda) \xrightarrow{\mathcal{L}} v_0^{-1/2} \eta$$

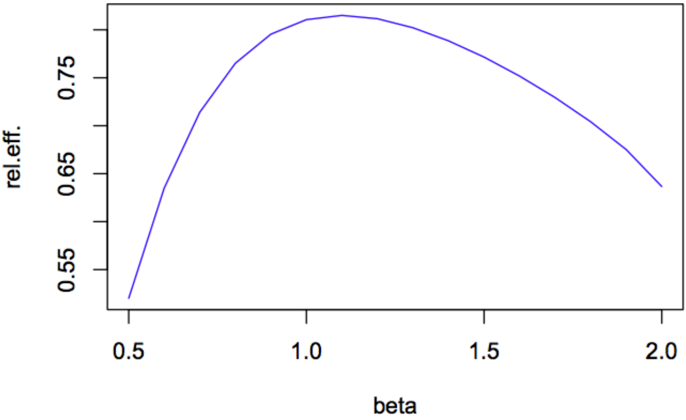
where $\eta \sim \mathcal{N}(0, I_2) \perp\!\!\!\perp (X_0, Z)$ and

$$v_0 := 4\sigma^{-2} \phi_\beta(0)^2 T^{1-2/\beta} \int_0^T X_t^2 dt.$$

[1] Applicable to more general locally stable OU processes, M (2013).

- **Relative efficiency LAD/MLE** [2]**:**

$$4\phi_\beta(0)^2 \left\{ \int \left(\frac{\partial \phi_\beta(y)}{\phi_\beta(y)} \right)^2 \phi_\beta(y) dy \right\}^{-1}$$



[2] Matsui and Takemura (2006) for the plot

Remark: Simple estimators of β and σ [3]

- $\Delta_j X := X_{jT/n} - X_{(j-1)T/n}$, $p \in (0, \beta/2)$, $\beta \in (2/3, 2)$.
- $V_n'(p) := \sum_{j=2}^{n} |\Delta_j X - \Delta_{j-1} X|^p$
- $V_n''(p) := \sum_{j=4}^{n} |\Delta_j X - \Delta_{j-1} X + \Delta_{j-2} X - \Delta_{j-3} X|^p$

[3] Application of Todorov (2013).

Remark: Simple estimators of β and σ [3]

- $\Delta_j X := X_{jT/n} - X_{(j-1)T/n}$, $p \in (0, \beta/2)$, $\beta \in (2/3, 2)$.
- $V'_n(p) := \sum_{j=2}^{n} |\Delta_j X - \Delta_{j-1} X|^p$
- $V''_n(p) := \sum_{j=4}^{n} |\Delta_j X - \Delta_{j-1} X + \Delta_{j-2} X - \Delta_{j-3} X|^p$

- $\sqrt{n}$-asymptotically normal **estimators:**

$$\hat{\beta}_n := p \log(2) / \log\{V''_n(p)/V'_n(p)\},$$

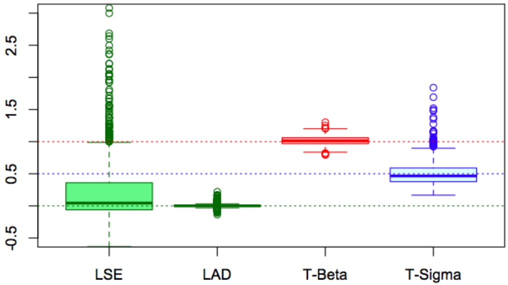
$$\hat{\sigma}_n := T^{-1/\hat{\beta}_n} \left\{ C(p, \hat{\beta}_n) n^{p/\hat{\beta}_n - 1} V'_n(p) \right\}^{1/p}.$$

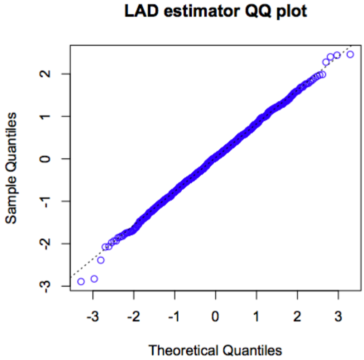
[3] Application of Todorov (2013).

Setup:

- $dX_t = -\lambda X_t dt + \sigma dJ_t$ **with** $\mathcal{L}(J_1) = \mathcal{S}_\beta(1)$ **and** $X_0 = 0$.
- $T \leftarrow 5$ **and** $n = 1000$, **with** 1000 **MC-iterations.**
- $(\lambda, \beta, \sigma) \leftarrow (0, 1, 0.5),\ (-1, 1, 0.5),\ (1, 1.5, 0.5)$.

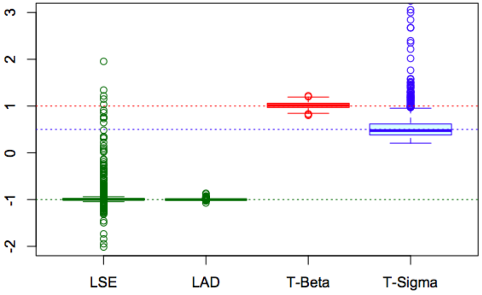
$(\lambda, \beta, \sigma) \leftarrow (0, 1, 0.5)$: Stable Lévy process

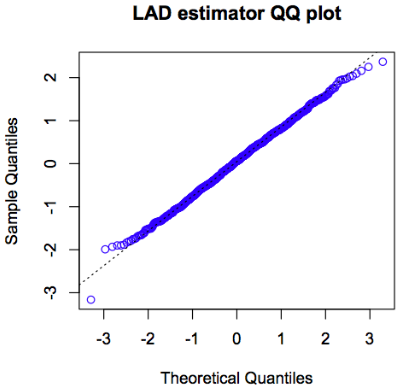




- LSE also shown for comparison.
- QQ plot of randomly-normed LAD estimator $n^{1/\beta-1/2}\sqrt{\int_0^T X_t^2 dt}(\hat{\lambda}_n - \lambda)$

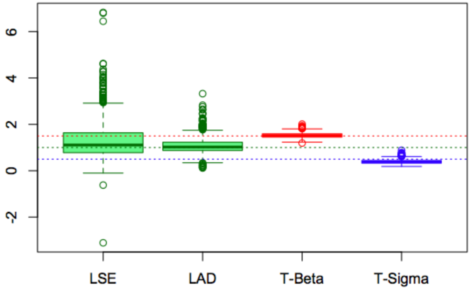
$(\lambda, \beta, \sigma) \leftarrow (-1, 1, 0.5)$: Non-ergodic case

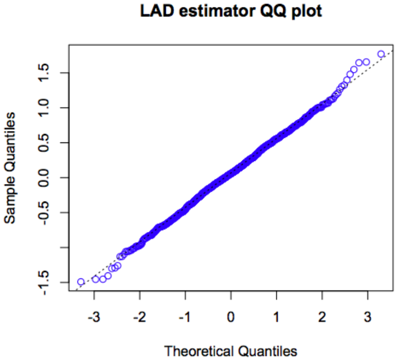




- **LSE also shown for comparison.**
- **QQ plot of randomly-normed LAD estimator** $n^{1/\beta - 1/2}\sqrt{\int_0^T X_t^2 dt}(\hat{\lambda}_n - \lambda)$

$(\lambda, \beta, \sigma) \leftarrow (1, 1.5, 0.5)$: Ergodic case





- **LSE also shown for comparison.**
- **QQ plot of randomly-normed LAD estimator** $n^{1/\beta-1/2}\sqrt{\int_0^T X_t^2 dt}(\hat{\lambda}_n - \lambda)$

Remark: Case of $T = T_n \to \infty$?

- **Faster but intractable convergence expected:**

$$n^{1/\beta} h^{1-1/\beta} (\hat{\lambda}_n - \lambda) \xrightarrow{\mathcal{L}} \text{Non-trivial law}$$

- **Local quadratic approximation no longer true:**

$$\frac{1}{\{n^{1/\beta} h^{1-1/\beta}\}^k} \partial_\lambda^k \ell_n(\lambda) = O_p(1), \qquad k \in \mathbb{Z}_+.$$

  - **Caused by the fact, e.g. Davis et al. (1992):**

$$\left(\frac{1}{n^{1/\beta}}\right)^k \sum_{j=1}^{n} X_{t_{j-1}}^k = O_p(1), \qquad k \in \mathbb{Z}_+.$$

Remark: Optimality for more parameters?

$$dX_t = (\gamma - \lambda X_t)dt + \sigma dJ_t$$

- $\mathcal{L}(J_1) = S_\beta(1)$, $\beta \in (0,2)$.
- $(X_{t_j})_{j=0}^n$ with $t_j = jT/n$ for fixed $T > 0$.

Remark: Optimality for more parameters?

$$dX_t = (\gamma - \lambda X_t)dt + \sigma dJ_t$$

- $\mathcal{L}(J_1) = S_\beta(1)$, $\beta \in (0,2)$.
- $(X_{t_j})_{j=0}^n$ with $t_j = jT/n$ for fixed $T > 0$.

1. $\theta = (\lambda, \gamma, \sigma) \in \mathbb{R} \times \mathbb{R} \times (0, \infty)$
   $\Rightarrow$ LAMN at rate

$$\left(n^{1/\beta - 1/2},\ n^{1/\beta - 1/2},\ \sqrt{n}\right)$$

Remark: Optimality for more parameters?

$$dX_t = (\gamma - \lambda X_t)dt + \sigma dJ_t$$

- $\mathcal{L}(J_1) = S_\beta(1)$, $\beta \in (0, 2)$.
- $(X_{t_j})_{j=0}^n$ with $t_j = jT/n$ for fixed $T > 0$.

1. $\theta = (\lambda, \gamma, \sigma) \in \mathbb{R} \times \mathbb{R} \times (0, \infty)$
   $\Rightarrow$ LAMN at rate
   $$\left(n^{1/\beta - 1/2},\ n^{1/\beta - 1/2},\ \sqrt{n}\right)$$

2. $\theta = (\lambda, \gamma, \sigma, \beta) \in \mathbb{R} \times \mathbb{R} \times (0, \infty) \times (0, 2)$
   $\Rightarrow$ Constantly singular Fisher information...
   $$\left(n^{1/\beta - 1/2},\ n^{1/\beta - 1/2},\ \sqrt{n},\ \sqrt{n}\log n\right)$$

   - Ait-Sahalia and Jacod (2008), M (2009); stable-Lévy process case.

Final remark: More general nonlinear SDE?

- **Estimation of $\theta := (\alpha, \gamma)$ in**

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

when observing $(X_{h_n}, X_{2h_n}, \ldots, X_{nh_n})$, $h = 1/n$.

[4] Presented at Dynstoch meeting 2012 Paris.

Final remark: More general nonlinear SDE?

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

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

when observing $(X_{h_n}, X_{2h_n}, \dots, X_{nh_n})$, $h = 1/n$.

$\Rightarrow$ The non-Gaussian stable quasi-likelihood [4]:

$$\hat{\theta}_n \in \underset{\theta \in \Theta}{\mathrm{argmax}} \sum_{j=1}^{n} \log \left\{ \frac{1}{h^{1/\beta} c(X_{t_{j-1}}, \gamma)} \phi_\beta \left(\frac{\Delta_j X - ha(X_{t_{j-1}}, \alpha)}{h^{1/\beta} c(X_{t_{j-1}}, \gamma)} \right) \right\}.$$

- Todorov's index estimator $\hat{\beta}_n$ still usable.

[4] Presented at Dynstoch meeting 2012 Paris.

- **Asymptotic mixed normality valid, M (2013):**

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

where

$$U(\theta_0) = \int_0^1 \frac{\{\partial_\alpha a(X_t, \alpha_0)\}^{\otimes 2}}{c(X_t, \gamma_0)^2} dt \cdot \int \frac{\partial \phi_\beta(y)^2}{\phi_\beta(y)} dy,$$
$$V(\theta_0) = \int_0^1 \frac{\{\partial_\gamma c(X_t, \gamma_0)\}^{\otimes 2}}{c(X_t, \gamma_0)^2} dt \cdot \int \frac{\{\phi_\beta(y) + y\partial \phi_\beta(y)\}^2}{\phi_\beta(y)} dy,$$

- **Asymptotic mixed normality valid, M (2013):**

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

where

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

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

Conjecture (ongoing study; not derived yet)

LAMN holds true: the stable QMLE is asymptotically optimal.

Some references

- Aït-Sahalia, Y. and Jacod, J. (2008), Fisher's information for discretely sampled Lévy processes. Econometrica 76, 727–761.
- Davis, R. A., Knight, K. and Liu, J. (1992), M-estimation for autoregressions with infinite variance. Stochastic Process. Appl. 40, 145–180.
- Hu, Y. and Long, H. (2009), Least squares estimator for Ornstein-Uhlenbeck processes driven by α-stable motions. Stochastic Processes Appl. 119, 2465–2480.
- Masuda, H. (2009), Joint estimation of discretely observed stable Lévy processes with symmetric Lévy density. J. Japan Statist. Soc. 39, 49–75.
- Masuda, H. (2013), On optimal estimation of stable OU processes: infill asymptotics. In preparation.
- Masuda, H. (2013), Laplace quasi-likelihood for locally stable OU processes. In preparation.
- Masuda, H. (2013), Local stable contrast functions for estimating discretely observed jump SDEs. In preparation.
- Matsui, M. and Takemura, A. (2006), Some improvements in numerical evaluation of symmetric stable density and its derivatives. Comm. Statist. Theory Methods 35, 149–172.
- Szimayer, A. and Maller, R. (2004), Testing for Mean Reversion in Processes of Ornstein-Uhlenbeck Type. Statist. Infer. for Stoch. Processes 7, 95–133.
- Todorov, V. (2013), Power variation from second order differences for pure jump semimartingales. Stoch. Process Appl. 123, 2829–2850.