Quasi-likelihood analysis for the stochastic differential equation with jumps

Teppei Ogihara

(Mitsubishi UFJ Trust Investment Technology Institute)

Joint work with Nakahiro Yoshida

Introduction

We consider $\mathbb{R}^d$-valued stationary process $\{X_t\}$ which satisfies SDE with jumps:

$$dX_t \;=\; a(X_{t-},\theta)dt + b(X_{t-},\sigma)dW_t + \int c(X_{t-},z,\theta)p(dz,dt) \qquad (1)$$

with unknown multidimensional parameters θ and σ.

$\{W_t\}$: r-dimensional standard Brownian motion, $\quad X_t \sim \pi(t \geq 0)$.
p : poisson random measure with mean measure $f_\theta(z)dzdt$,
i.e., for disjoint Borel sets $A_1,\ldots,A_r$,
$p(A_1),\ldots,p(A_r)$ are independent and the law of $p(A_i)$ is Poisson distribution with mean $\int_{A_i} f_\theta dzdt$.
Moreover, $\{W_t\}$ and p are independent.

Introduction

Let $\alpha = (\sigma, \theta)$, then we are interested in the inference for the true value $\alpha^* = (\sigma^*, \theta^*)$ of parameters from discrete observation $\{X_{t_i^n}\}_{1 \le i \le n}$ of $\{X_t\}$

h_n: distance between data, $t_i^n = ih_n$ $(1 \le i \le n)$
We assume there exists positive number c and C, such that
$cn^{-3/5} \le h_n \le Cn^{-4/7}$.
In particular, $nh_n \to \infty$ and $nh_n^2 \to 0$.

We want to argue about maximum-likelihood estimator and Bayes estimator. However, it is difficult in general to obtain strict likelihood function for diffusion process with jumps.

So we will define quasi-likelihood function and check asymptotic behavior of quasi-maximum-likelihood estimator and Bayes type estimator.

Previous Work

Shimizu and Yoshida(2006) constracts a quasi-likelihood function by using threshold which detects jumps, and prove the consistency and the asymptotic normality of the quasi-maximal likelihood estimator, assuming $|f_\theta(z)| \leq C|z|^\gamma$ for some constant $C > 0$ and $\gamma > 3$ near the origin.

We use almost the same quasi-likelihood function. But, our result shows the consistency, the asymptotic normality and higher order moment convergence for the quasi-maximal likelihood estimator and the Bayes type estimator. Moreover, the condition of jump distribution f_θ is weakened and contains the case f_θ is not 0 at origin, like normal distributions.

SDE with Jumps

In the equation (1), if $c(x, z, \theta) = z$ and $\lambda(\theta) := \int f_\theta dz < \infty$, then for $F_\theta(z) = f_\theta(z)/\lambda(\theta)$, there exists
$\{N_t\}$: Poisson process with mean $\lambda(\theta)t$
$\{Y_i\}$: i.i.d.$\sim F_\theta$,
such that

$$\int_0^t \int c(X_{t-}, z, \theta) p(dz, dt) = \sum_{i=1}^{N_t} Y_i$$

In what follows, we will assume $c(x, z, \theta) = z$. However, the following discussion is generalized for general c which satisfies certain conditions.

SDE with Jumps

Example: Lévy OU process

Let $d = 1$,$a(x,\theta) = -\alpha x$, $b(x,\sigma) = \sigma$,
$F_\theta(z)$: probability density of $N(0,v)$, then

$$X_t = X_0 - \alpha \int_0^t X_{t-} dt + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$

and θ can be written as $\theta = (\alpha, \lambda, v)$.

Quasi-likelihood Analysis

Next, we define a quasi-likelihood function. We prepare the notation.

$\Delta X_i^n := X_{t_i^n} - X_{t_{i-1}^n}$
$\bar{X}_{i,n}(\theta) := \Delta X_i^n - h_n a(X_{t_{i-1}^n}, \theta)$
$\bar{X}_{i,n} := \bar{X}_{i,n}(\theta^*)$
$7/16 \leq \rho < 1/2$
$\beta(x, \sigma) := b(x, \sigma) b^T(x, \sigma)$ where T denotes transpose.

$\varphi_n(z)$ is a smooth function such that
$\varphi_n(z) = 0$ near the points where $f_\theta(z)$ equals to 0, and
$\varphi_n(z) = 1$, if z is far from those points.
The exact definition of φ_n is in an assumption [H10].

Quasi-likelihood Analysis

We define log quasi-likelihood function $H_n(\alpha)$ of $\{\Delta X_1^n, \Delta X_2^n, \ldots, \Delta X_n^n\}$ by

$$
\begin{aligned}
H_n(\alpha) \;=\; & -\frac{1}{2h_n}\sum_{i=1}^{n} \bar{X}_{i,n}^T(\theta)\beta^{-1}(X_{t_{i-1}^n},\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le h^\rho\}} \\
& -\frac{1}{2}\sum_{i=1}^{n} \log\beta(X_{t_{i-1}^n},\sigma)1_{\{|\Delta X_i^n|\le h^\rho\}} \\
& +\sum_{i=1}^{n}\{\log f_\theta(\Delta X_i^n)\}\varphi_n(\Delta X_i^n)1_{\{|\Delta X_i^n|>h^\rho\}} \\
& -h_n\sum_{i=1}^{n}\int f_\theta(z)\varphi_n(z)dz. \qquad (2)
\end{aligned}
$$

Quasi-likelihood Analysis

Intuitive mean of $H_n(\alpha)$ is as following:
If jumps do not occur in the time interval $(t_{i-1}^n, t_i^n]$, then

$$\begin{aligned}\Delta X_i^n &= X_{t_i^n} - X_{t_{i-1}^n} \\ &= \int_{t_{i-1}^n}^{t_i^n} a(X_{t-}, \theta^*)dt + \int_{t_{i-1}^n}^{t_i^n} b(X_{t-}, \sigma^*)dW_t \\ &\approx h_n a(X_{t_{i-1}^n}, \theta^*) + b(X_{t_{i-1}^n}, \sigma^*)(W_{t_i^n} - W_{t_{i-1}^n}) \\ &\approx N(h_n a(X_{t_{i-1}^n}, \theta^*), \beta(X_{t_{i-1}^n}, \sigma^*)h_n)\end{aligned}$$

The log density of the last distribution appeared in the first and the second term of H_n.

Quasi-likelihood Analysis

If a jump occurs in the time interval $(t_{i-1}^n, t_i^n]$,
we neglect except jump term, then

$$\Delta X_i^n \;=\; X_{t_i^n} - X_{t_{i-1}^n} = Y_j \sim F_\theta,$$

so log likelihood becomes

$$\log F_\theta(\Delta X_i^n) = \log f_\theta(\Delta X_i^n) - \log \lambda(\theta).$$

It is the third term of H_n.

Therefore, H_n is a quasi-likelihood function such that
if $|\Delta X_i^n| \leq h_n^\rho$, then H_n judge that jumps do not occur, and
if $|\Delta X_i^n| > h_n^\rho$, then H_n judge that a jumps occur.

Main Theorems

Now, we denote the assumptions of main theorems.

[A1] a and b satisfy Lipschitz condition,
i.e., there exists L such that

$$|a(x,\theta^*) - a(y,\theta^*)| + |b(x,\sigma^*) - b(y,\sigma^*)| \le L|x-y|$$

[A2] a and b are differentiable and there exists a positive constant C such that

$$|\partial_x^k b(x,\sigma)|, |\partial_\theta^l a(x,\theta)|, |\partial_\sigma^l b(x,\sigma)| \le C(1+|x|)^C$$

for $l = 0,1,2,3,4; k = 1,2$.

[A3] $\inf_{x,\sigma} \det \beta(x,\sigma) > 0$.

Main Theorems

[A4] There exists potitive constants r and K such that $f_{\theta^*}(z)1_{\{|z|\leq r\}} \leq K|z|^{1-d}$ and

$$\sup_\theta \int |z|^p f_\theta(z)dz < \infty (p > 0).$$

[A5] f_θ is diffenretiable in θ and

$$\begin{aligned} \sup_\theta \int |\partial_\theta^k f_\theta(z)|dz &< \infty \\ \sup_\theta \int |\partial_\theta^k \log f_\theta(z)|^l \times f_{\theta^*}(z)dz &< \infty \end{aligned}$$

for $0 \leq k, l \leq 4$.

Main Theorems

[A6](exponential mixing property)

$\{X_t\}$ is stable and there exists a positive number c such that

$$\sup_{t\in\mathbb{R}_+} \sup_{A\in\sigma[X_r:r\le t],B\in\sigma[X_r:r\ge t+h]} |P[A\cap B]-P[A]P[B]| \le e^{-ch}/c \quad (h>0).$$

[A7] $\sup_{t\ge 0} E[|X_t|^p]<\infty$ for any $p\ge 1$.

Main Theorems

The condition [A6] implies that there exists a distribution π, such that for any π-integrable function f,
$\frac{1}{T}\int_0^T f(X_t)dt \to^P \int f(x)\pi(dx)$ as $n \to \infty$

Masuda(2008) prove conditions $[A6]$ and $[A7]$, under easily checkable conditions for coefficients a and b.

Masuda, H.:On Stability of Diffusions with Compound-Poisson Jumps. Bulletin of Informatics and Cybernetics, Vol. 40, 61–74 (2008)

Main Theorems

Next two conditions are about the function φ_n.

[A8] $\varphi_n(z)$ satisfies $0 \leq \varphi_n(z) \leq 1$ and there exists a potive constant M such that $\varphi_n = 0$ on D_n where

$$D_n = \cup_{k=0}^{4} \left\{ z; \sup_\theta |\partial_\theta^k \log f_\theta(z)| \geq \frac{M}{\epsilon_n^{k\vee 1}} \right\} \cup \cup_{k=0}^{4} \left\{ z; \sup_\theta |\partial_z \partial_\theta^k \log f_\theta(z)| \geq \frac{M}{\epsilon_n^{k+1}} \right\}$$

Moreover, $\partial_z \varphi_n = 0$ on D_n and $\sup_z |\partial_z \varphi_n| = O(\epsilon_n^{-1})$ where $\epsilon_n = h_n^{1/16}$.

Main Theorems

[A9] There exists positive constants a and C such that

$$\sup_{\theta} \int |\partial_{\theta}^{k} \log f_{\theta}(z)| \times f_{\theta^*}(z)(1-\varphi_n(z))dz \leq Ch_n^a$$

$$\sup_{\theta} \int |\partial_{\theta}^{k} f_{\theta}(z)|(1-\varphi_n(z))dz \leq Ch_n^a.$$

for $k = 0, 2$.

Main Theorems

[A10] Let

$$
\begin{aligned}
Y^1(\sigma : \sigma^*) &= \frac{1}{2}\int tr(I_d - \beta^{-1}(x,\sigma)\beta(x,\sigma^*))\pi(dx) \\
&\quad -\frac{1}{2}\int \log \frac{\det \beta(x,\sigma)}{\det \beta(x,\sigma^*)}\pi(dx) \\
Y^2(\theta;\alpha^*) &= -\frac{1}{2}\int (a(x,\theta)-a(x,\theta^*))^T\beta^{-1}(x,\sigma^*) \\
&\quad \times(a(x,\theta)-a(x,\theta^*))\pi(dx) \\
&\quad +\int (\log f_\theta(z) - \log f_{\theta^*}(z))f_{\theta^*}(z)dz \\
&\quad -(\lambda(\theta)-\lambda(\theta^*)),
\end{aligned}
$$

then there exists positive constants χ and χ' such that

$$
Y^1(\sigma;\sigma^*) \le -\chi|\sigma-\sigma^*|^2, \quad Y^2(\theta;\alpha^*) \le -\chi'|\theta-\theta^*|^2.
$$

Main Theorems

Now, we state main theorems. Let

$$\begin{aligned}\Gamma_1 &= \frac{1}{2}\int tr(\beta^{-1}\partial_\sigma\beta\beta^{-1}\partial_\sigma\beta(x,\sigma^*)\pi(dx) \\ \Gamma_2 &= \int \partial_\theta a^T(x,\theta^*)\beta^{-1}(x,\sigma^*)\partial_\theta a(x,\theta^*)\pi(dx) \\ &\quad + \int\int \frac{\partial_\theta f_{\theta^*}(z)\times\partial_\theta f_{\theta^*}(z)}{f_{\theta^*}(z)}dz,\end{aligned}$$

where T denotes transpose. Γ_1 and Γ_2 are the limit of normalized second order differential $-\frac{1}{n}\partial_\sigma^2 H_n(\alpha^*)$ and $-\frac{1}{nh_n}\partial_\theta^2 H_n(\alpha^*)$ of $H_n(\alpha)$ as $n\to\infty$.

Main Theorems

Y^1 and Y^2 are the limits of

$$Y_n^1(\alpha) = \frac{1}{nh}(H_n(\sigma, \theta) - H_n(\sigma^*, \theta)),$$

and

$$Y_n^2(\theta) = \frac{1}{n}(H_n(\widehat{\sigma}_n, \theta) - H_n(\widehat{\sigma}_n, \theta^*)),$$

respectively, where $\widehat{\sigma}_n$ is the quasi-maximal likelihood estimator for σ^* and the condition $[A10]$ is a separability condition for parameters.

Main Theorems

Let $\widehat{\alpha}_n = (\widehat{\sigma}_n, \widehat{\theta}_n)$ be a random variable such that $H_n(\widehat{\alpha}_n) = \max_{\sigma,\theta} H_n(\alpha)$.

Theorem 1. Assume $[A1] - [A10]$.
Then the quasi-maximum-likelihood estimator $\widehat{\alpha}_n$ satisfies

$$\widehat{\alpha}_n \to^P \alpha^*$$

and

$$\widehat{u}_n := (\sqrt{n}(\widehat{\sigma}_n - \sigma^*), \sqrt{nh_n}(\widehat{\theta}_n - \theta^*)) \to^d \widehat{u}$$

as $n \to \infty$. Moreover,

$$E[f(\widehat{u}_n)] \to E[f(\widehat{u})] \quad (n \to \infty)$$

where $\widehat{u} \sim N(0, diag(\Gamma_1^{-1}, \Gamma_2^{-1}))$ and f is at most polynomial growth, i.e., there exists C such that $|f(x)| \leq C(1 + |x|)^C$.

Main Theorems

Next, we see Bayes type estimators. We define Bayes type estimators for σ and θ separately because the rate of convergences are different.

We define Bayes type estimators $\tilde{\sigma}_n, \tilde{\theta}_n$ for σ and θ with prior distribution $\pi_1(\sigma)$ and $\pi_2(\theta)$ by

$$\begin{aligned} \tilde{\sigma}_n &= \left\{ \int \exp(H_n(\sigma, \theta^\star))\pi_1(d\sigma) \right\}^{-1} \times \int \sigma \exp(H_n(\sigma, \theta^\star))\pi_1(d\sigma) \\ \tilde{\theta}_n &= \left\{ \int \exp(H_n(\tilde{\sigma}_n, \theta))\pi_2(d\theta) \right\}^{-1} \times \int \theta \exp(H_n(\tilde{\sigma}_n, \theta))\pi_2(d\theta) \end{aligned}$$

where $\theta^\star$ is an arbitrary constant.

Main Theorems

Theorem 2. Assume $[A1] - [A10]$.
Then the Bayes type estimator $(\tilde{\sigma}_n, \tilde{\theta}_n)$ satisfies

$$(\tilde{\sigma}_n, \tilde{\theta}_n) \to^P \alpha^*$$

$$\tilde{u}_n := (\sqrt{n}(\tilde{\sigma}_n - \sigma^*), \sqrt{nh_n}(\tilde{\theta}_n - \theta^*)) \to^d \hat{u}$$

as $n \to \infty$. Moreover,

$$E[f(\tilde{u}_n)] \to E[f(\hat{u})] \quad (n \to \infty),$$

where $\hat{u}$ and f are the same as those in Theorem 1.

Example

For the Lev́y OU process

$$X_t = X_0 - \alpha \int_0^t X_{t-} dt + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \quad Y_i \sim N(0, v), \quad \theta = (\alpha, \lambda, v),$$

we can check conditions $[A1] - [A10]$ with $\varphi_n \equiv 1$, $\Gamma^1 = 2/(\sigma^*)^2$ and $\Gamma^2 = diag(1/(\sigma^*)^2, 1/\lambda^*, \lambda^*/(2v))$.

So for the quasi-maximal likelihood estimator $(\widehat{\sigma}_n, \widehat{\theta}_n)$ and the Bayes type estimator $(\tilde{\sigma}_n, \tilde{\theta}_n)$ for (σ^*, θ^*), we have

$$\begin{aligned} (\sqrt{n}(\widehat{\sigma}_n - \sigma^*), \sqrt{nh_n}(\widehat{\theta}_n - \theta^*)) &\to N(0, \Gamma^{-1}), \\ (\sqrt{n}(\tilde{\sigma}_n - \sigma^*), \sqrt{nh_n}(\tilde{\theta}_n - \theta^*)) &\to N(0, \Gamma^{-1}), \end{aligned}$$

as $n \to \infty$, where $\Gamma = diag(2/(\sigma^*)^2, 1/(\sigma^*)^2, 1/\lambda^*, \lambda^*/(2v))$.

Polynomial Large Deviation Theory

For the proof of Theorem 1 and 2, we use "polynomial type large deviation inequality."

In general, let
$H_n(\upsilon,\tau)$: function, $\xi = (\upsilon,\tau)$,
a_n: sequence of positive numbers, $a_n \to 0$ $(n \to \infty)$, $b_n = a_n^{-2}$,

$$\begin{aligned} Z_n(u,\tau,\upsilon^*) &:= \exp\{H_n(\upsilon^* + a_n u,\tau) - H_n(\upsilon^*,\tau)\} \\ Y_n(\upsilon,\tau;\upsilon^*) &:= \frac{1}{b_n}(H_n(\upsilon,\tau) - H_n(\upsilon^*,\tau)) \\ \Gamma_n(\upsilon,\tau) &:= -\frac{1}{b_n}\partial_\upsilon^2 H_n(\upsilon,\tau), \end{aligned}$$

Let $\Gamma(\tau,\xi^*)$ and $Y(\upsilon,\tau;\upsilon^*)$ be the limit of Γ_n and Y_n, respectively.

Polynomial Large Deviation Theory

Let $L > 0, \alpha > 0, \beta = \alpha/(1-\alpha), 0 < \beta_1 < 1/2, \beta_2 \geq 0,$
$0 < \rho_1 < 1 \rho_2 > 0.$

[C1] For $M_1 = L(\beta - \rho_1)^{-1}$ and $M_2 = L(\frac{2\beta_1}{1-\alpha} - \rho_1)^{-1}$,

$$\sup_{n \in \mathbb{N}} E\left[\left(b_n^{-1} \sup_{(\upsilon,\tau) \in \Theta \times \mathcal{T}} |\partial_\upsilon^3 H_n(\upsilon, \tau)|\right)^{M_1}\right] < \infty.$$

$$\sup_{n \in \mathbb{N}} E\left[\sup_{\tau \in \mathcal{T}} (b_n^{\beta_1} |\Gamma_n(\upsilon^*, \tau) - \Gamma(\tau; \xi^*)|)^{M_2}\right] < \infty.$$

[C2] $\rho_1 < \beta \wedge \frac{2\beta_1}{1-\alpha}$, $\alpha < \rho_2/2$, $1 - 2\beta_2 - \rho_2 > 0$.

[C3] $\Gamma(\tau; \xi^*)$ is positive definite, uniformly in τ.

Polynomial Large Deviation Theory

[C4] For $M_3 = L(1-\rho_1)^{-1}$ and $M_4 = L(1-2\beta_2-\rho_2)^{-1}$,

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\tau\in\mathcal{T}} |a_n \partial_\upsilon H_n(\upsilon^*, \tau)|\right)^{M_3}\right] < \infty.$$

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\upsilon,\tau} b_n^{\frac{1}{2}-\beta_2} |Y_n(\upsilon,\tau;\upsilon^*) - Y(\upsilon,\tau;\upsilon^*)|\right)^{M_4}\right] < \infty.$$

[C5] There exists a positive constant χ such that $Y(\upsilon,\tau;\upsilon^*) \leq -\chi|\upsilon-\upsilon^*|^2$.

Polynomial Large Deviation Theory

Theorem 3. (N.Yoshida) Assume [C1]-[C5]. Then there exists a constant C_L such that

$$E\left[\sup_{(u,\tau)\in V_n(r)\times\mathcal{T}} Z_n(u,\tau;\upsilon^*) \geq e^{-r/2}\right] \leq \frac{C_L}{r^L} \tag{3}$$

for any $r > 0$, where $V_n(r) := \{u \in \mathbb{R}^m; \upsilon^* + a_n u \in \Theta, |u| \geq r\}$.

Polynomial Large Deviation Theory

We use Theorem 3 to prove Theorem 1 for functions

$$
\begin{aligned}
\sigma &\to H_n(\sigma,\theta) \quad (\upsilon \to \sigma, \tau \to \theta) \\
\theta &\to H_n(\widehat{\sigma}_n,\theta) \quad (\upsilon \to \theta).
\end{aligned}
$$

Then under certain moment condition of H_n, $\partial_\sigma H_n$ and $\partial_\theta H_n$, we have

$$
\begin{aligned}
P\Big[\Big|\frac{1}{n}(\widehat{\sigma}_n - \sigma^*)\Big| > r\Big] &\le E\Big[\sup_{(u,\theta)\in V_n(r)\times\Theta} Z_n(u,\theta;\sigma^*) \ge 1\Big] \le \frac{C_L}{r^L} \\
P\Big[\Big|\frac{1}{nh}(\widehat{\theta}_n - \theta^*)\Big| > r\Big] &\le E\Big[\sup_{u\in V_n(r)} Z_n(u;\theta^*) \ge 1\Big] \le \frac{C_L}{r^L}
\end{aligned}
$$

These inequalities make possible to treat high order moments. These inequalities are key of the proof of Theorem 1 and 2.

References

[1] Cont, R. and Tankov, P.: *Financial modeling with Jump Processes*, Chapman & Hall/CRC, Financial Mathematics Series (2004)

[2] Ibragimov, I. A. and Has'minskii, R. Z.: *Statistical Estimation*, Springer-Verlag, Berlin (1981)

[3] Masuda, H.:Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* 117, no. 1, 35–56 (2007)

[4] Masuda, H.:On Stability of Diffusions with Compound-Poisson Jumps. Bulletin of Informatics and Cybernetics, Vol. 40, 61–74 (2008)

References

[5] Shimizu, Y.:M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Stat. Inference Stoch. Process.* 9, no. 2, 179–225 (2006)

[6] Shimizu, Y.:Density estimation of Levy measures for discretely observed diffusion processes with jumps. *J. Japan Statist. Soc.* 36, 37-62 (2007)

[7] Shimizu, Y. and Yoshida, N.:Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations. *Stat. Inference Stoch. Process.* 9, 227-277 (2006)

[8] Yoshida, N.: Polynomial Type Large Deviation Inequalities and Quasi-Likelihood Analysis for Stochastic Differential Equations. preprint, to appear in *Ann. Inst. Statist. Math* (2005)