The Local Asymptotic Mixed Normality for Nonsynchronously Observed Diffusion Processes

Teppei Ogihara

JST CREST, CSFI Osaka University

2013.3.13

Introduction

- Model
  $\{Y_t\}_{0\le t\le T} = \{(Y_t^1, Y_t^2)\}_{0\le t\le T}$: a two-dimensional stochastic process satisfying the following SDE :
  $$dY_t = \mu(t, Y_t, \sigma)ds + b(t, Y_t, \sigma)dW_t, \quad t \in [0, T]. \tag{1}$$
  $\{W_t\}_{0\le t\le T}$: a two-dimensional standard Wiener process
  μ, b : Borel functions $\quad \sigma \in \Lambda \subset \mathbb{R}^d$ a bounded open set
  $Y_0 \in L^2$, the distribution of Y_0 does not depend on σ

- Observation
  $\mathcal{S} = \{S^{n,i}\}_{i=0}^{l_{1,n}}$: Observation times of $\{Y_t^1\}_t$ (random)
  $\mathcal{T} = \{T^{n,j}\}_{j=0}^{l_{2,n}}$: Observation times of $\{Y_t^2\}_t$ (random)
  $P_{\sigma,n}$: the distribution of observations $\mathcal{S}, \mathcal{T}, \{Y^1_{S^{n,i}}\}, \{Y^2_{T^{n,j}}\}$
  $(\mathcal{S}, \mathcal{T}) \perp (Y_t, W_t)$, $S^{n,0} = T^{n,0} = 0$, $S^{n,l_{1,n}} = T^{n,l_{2,n}} = T$
  The end time $T > 0$ is fixed. $\max_{i,j} |S^i - S^{i-1}| \vee |T^j - T^{j-1}| \to^p 0$.
  The distributions of $\mathcal{S}$ and $\mathcal{T}$ does not depend on σ.

- We will consider the estimation problem of the true value σ_* of parameter σ with nonsynchronous observations $\mathcal{S}$, $\mathcal{T}$, $\{Y^1_{S^{n,i}}\}$, $\{Y^2_{T^{n,j}}\}$.

Introduction

- A family $\{P_{\sigma,n}\}_{\sigma,n}$ of probability measures on measurable spaces $(\mathcal{X}_n, \mathcal{A}_n)$ are said to satisfy the local asymptotic mixed normality (LAMN) at $\sigma = \sigma_*$ if there exist a sequence $\{b_n\}_{n\in\mathbb{N}}$ of positive numbers, $d \times d$ symmetric positive definite matrices Γ_n, Γ and d-dimensional vectors $\mathcal{N}_n, \mathcal{N}$ such that $b_n \to \infty$,

$$\log \frac{dP_{\sigma_* + {b_n}^{-1/2} u, n}}{dP_{\sigma_*, n}} - \left(u^\star \sqrt{\Gamma_n} \mathcal{N}_n - \frac{1}{2} u^\star \Gamma_n u \right) \to 0$$

in $P_{\sigma_*,n}$-probability, $\mathcal{N} \sim N(0, I_d) \perp \Gamma$ and

$$(\mathcal{N}_n, \Gamma_n) \to^d (\mathcal{N}, \Gamma)$$

as $n \to \infty$ for any $u \in \mathbb{R}^d$.

Introduction

- Jeganathan (1983) proved the minimax inequality :

$$\lim_{\alpha\to\infty}\liminf_{n\to\infty}\sup_{|u|\le\alpha} E_{\sigma_*+b_n^{-1/2}u}[l(|b_n^{1/2}(V_n-\sigma_*-b_n^{-1/2}u)|)] \geq E[l(|\Gamma^{-1/2}\mathcal{N}|)]$$

  for any estimators $\{V_n\}$ and any function $l:[0,\infty)\to[0,\infty)$ which is nondecreasing and $l(0)=0$, when the family $\{P_{\sigma,n}\}$ has the LAMN property.
- This inequality gives lower bound of estimation error for estimators. In particular, this inequality gives lower bound of asymptotic variance of estimators if $l(x)=x^2$.
- On the other hand, Gobet (2001) proved the LAMN property for diffusion processes when the observations are synchronous, equi-spaced : $\{Y_{kT/n}\}_{k=1}^n$ by using a Malliavin calculus approach.
- In the context of high frequency financial data analysis, the observations of two different securities are necessarily nonsynchronous.
- In this work, we prove the LAMN property for nonsynchronously observed diffusion processes.

Assumptions

We first consider assumptions to prove the LAMN property. The first one is the assumptions for coefficient μ, b in the SDE (1).

A1.

- For $i, j \geq 0, 0 \leq k \leq 4, 0 \leq 2i + j \leq 3, 0 \leq l \leq 1$, the derivatives $\partial_t^i \partial_x^j \partial_\sigma^k b, \partial_t^i \partial_x^j \partial_\sigma^l \mu$ exist and continuous with respect to (t, x, σ). Moreover, $\partial_x \mu, \partial_x b$ are bounded uniformly in $[0, T] \times \mathbb{R}^2 \times \Lambda$ and $\partial_\sigma^k b$ can be extended to a continuous function on $[0, T] \times \mathbb{R}^2 \times \bar{\Lambda}$.
- $b(t, x, \sigma)$ is a symmetric positive definite matrix for any (t, x, σ).

Condition $[A1]$ is almost the same as the assumptions in Gobet (2001).

Assumptions

Our approach to prove the LAMN property is approximating the log-likelihood ratio $\log(dP_{\sigma_*+b_n^{-1/2}u,n}/dP_{\sigma_*,n})$ by the difference of quasi-log-likelihood function $H_n(\sigma)$ proposed in O. and Yoshida (2012). In that paper, a LAMN type property of $H_n(\sigma_*+b_n^{-1/2}u)-H_n(\sigma_*)$ are proved under the following condition about observation times.

A2. There exists an exponential α-mixing simple point process $\{N_t\}_{t\geq 0}=\{(N_t^1,N_t^2)\}_{t\geq 0}$ with stationary increment such that $N_0=0$ and

$$S^i=\inf\{t\geq 0; N^1_{b_nt}\geq i\},\quad T^j=\inf\{t\geq 0; N^2_{b_nt}\geq j\}.$$

Moreover,

$$E[|N_1|^q]<\infty,\quad \limsup_{u\to\infty}\max_{i=1,2}u^qP[N_u^i=0]<\infty$$

for any $q>0$.

Assumptions

- $[A2]$ can be easily checked when N is a Poisson process :

$$\lim_{u \to \infty} u^q P[N_u^i = 0] = \lim_{u \to \infty} u^q e^{-\lambda_i u} = 0.$$

- The LAMN type property of the quasi-likelihood function can be proved under weaker conditions than $[A2]$. See O. and Yoshida (2012) for more details.

The main theorem

Theorem 1

Assume $[A1], [A2]$*. Then the family* $\{P_{\sigma,n}\}_{\sigma,n}$ *generated by nonsynchronous observations* $\mathcal{S}, \mathcal{T}, \{Y^1_{S^i}\}, \{Y^2_{T^j}\}_j$ *has the LAMN property at* $\sigma = \sigma_*$*.*

- Theorem 1 leads the minimax inequality :

$$\lim_{\alpha\to\infty} \liminf_{n\to\infty} \sup_{|u|\leq\alpha} E_{\sigma_*+b_n^{-1/2}u}[l(|b_n^{1/2}(V_n - \sigma_* - b_n^{-1/2}u)|)] \geq E[l(|\Gamma^{-1/2}\mathcal{N}|)]$$

for any estimators $\{V_n\}$ and any function $l : [0,\infty) \to [0,\infty)$ which is nondecreasing and $l(0) = 0$.

The main theorem

We shortly look back the results in O. and Yoshida (2012).

- The quasi-log-likelihood function $H_n(\sigma)$ is defined as follows :

$$H_n(\sigma) = -\frac{1}{2}Z^\star S^{-1}Z - \frac{1}{2}\log\det S \qquad (2)$$

where $\star$ represents transpose,

$$Z = (((Y^1_{S^i} - Y^1_{S^{i-1}})/\sqrt{|I^i|})^\star_i, ((Y^2_{T^j} - Y^2_{T^{j-1}})/\sqrt{|J^j|})^\star_j)^\star,$$

$$S = \begin{pmatrix} \mathrm{diag}(\{|b^1_i|^2\}_{i=1}^l) & \left\{b^1_i \cdot b^2_j \frac{|I^i \cap J^j|}{\sqrt{|I^i|}\sqrt{|J^j|}}\right\}_{i,j} \\ \left\{b^1_i \cdot b^2_j \frac{|I^i \cap J^j|}{\sqrt{|I^i|}\sqrt{|J^j|}}\right\}_{j,i} & \mathrm{diag}(\{|b^2_j|^2\}_{j=1}^m) \end{pmatrix},$$

$I^i = [S^{i-1}, S^i), J^j = [T^{j-1}, T^j)$ and b^1_i and b^2_j are values obtained by substituting the latest observed X before the time points S^{i-1} and T^{j-1}, respectively.

- If μ equals 0 and b and $\mathcal{S}, \mathcal{T}$ are deterministic, Z follows a normal distribution and S is an approximate covariance matrix of Z.

The main theorem

- Let $\Gamma = -\mathrm{P}\text{-}\lim_{n\to\infty}\partial_\sigma^2 H_n(\sigma_*)$ and $\hat{\sigma}_n$ be the quasi-maximum likelihood estimator defined by $H_n(\sigma)$, i.e. $\hat{\sigma}_n = \mathrm{argmax} H_n(\sigma)$.
- O. and Yoshida (2012) proved that there exists a random variable $\mathcal{N} \sim N(0, I_d) \perp \Gamma$ such that

$$\hat{\sigma}_n \to^p \sigma_*, \quad b_n^{1/2}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma^{-1/2}\mathcal{N},$$
$$E[f(b_n^{1/2}(\hat{\sigma}_n - \sigma_*))] \to E[f(\Gamma^{-1/2}\mathcal{N})]$$

  for any continuous function f of at most polynomial growth, where $\to^{s\text{-}\mathcal{L}}$ represents stable convergence.
- Similar results for the Bayes type estimator $\tilde{\sigma}_n$ are also proved.

The main theorem

- Γ in O. and Yoshida (2012) coincides with Γ in Theorem 1.
- By Theorem 1 and the results in O. and Yoshida (2012), we can prove that the qMLE $\hat{\sigma}_n$ and the Bayes type estimator $\tilde{\sigma}_n$ attain the lower bound in the minimax inequality for any continuous loss function l of at most polynomial growth.
- Then $\hat{\sigma}_n$ and $\tilde{\sigma}_n$ are asymptotically efficient in this sense.
  (We need to change the assumptions about drift coefficient μ and conditions for separation to uniform estimate in σ_* in O. and Yoshida (2012)).

Example

We calculate the qMLE $\hat{\sigma}_n$ for the simple model:

$$\left\{ \begin{array}{lcl} dY_t^1 & = & \sigma_1 dW_t^1 \\ dY_t^2 & = & \sigma_3 dW_t^1 + \sigma_2 dW_t^2 \end{array} \right.$$

where the parameter $\sigma = (\sigma_1, \sigma_2, \sigma_3)$. Let $\{N_t^1\}, \{N_t^2\}$ be independent Poisson processes with parameters λ_1, λ_2, and $S^i = \inf\{t \geq 0; N_{nt}^1 \geq i\}$, $T^j = \inf\{t \geq 0; N_{nt}^2 \geq j\}$.

Then for $G = \{|I^i \cap J^j|/\sqrt{|I^i||J^j|}\}_{i,j}$, we obtain

$$\begin{array}{rcl} H_n(\sigma) & = & -\frac{1}{2} Z^\star \begin{pmatrix} \sigma_1^2 I_{l_{1,n}} & \sigma_1\sigma_3 G \\ \sigma_1\sigma_3 G^* & (\sigma_2^2+\sigma_3^2) I_{l_{2,n}} \end{pmatrix}^{-1} Z \\ & & -\frac{1}{2} \log\det \begin{pmatrix} \sigma_1^2 I_{l_{1,n}} & \sigma_1\sigma_3 G \\ \sigma_1\sigma_3 G^* & (\sigma_2^2+\sigma_3^2) I_{l_{2,n}} \end{pmatrix}. \end{array}$$

The qMLE $\hat{\sigma}_n = (\hat{\sigma}_{1,n}, \hat{\sigma}_{2,n}, \hat{\sigma}_{3,n})$ is obtained as σ which maximizes H_n.

Example

By O. and Yoshida (2012), we can see that

$$\hat{\sigma}_n \to^p \sigma_*, \quad \sqrt{n}(\hat{\sigma}_n - \sigma_*) \to^d N(0, \Gamma^{-1}).$$

Then the estimator $\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$ of the quadratic covariation $< Y^1, Y^2 >_T$ satisfies

$$\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T \to^p < Y^1, Y^2 >_T,$$

$$\sqrt{n}(\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T - < Y^1, Y^2 >_T) \to^d N(0, v)$$

for some $v > 0$ by Delta method.

Let $\rho = \sigma_3/\sqrt{\sigma_2^2 + \sigma_3^2}$ and

$$\begin{aligned} a(\rho') &= T^{-1}\text{P-lim}_{n\to\infty}\mathrm{tr}((I_{l_{1,n}} - (\rho')^2 GG^\star)^{-1}), \\ c(\rho') &= T^{-1}\text{P-lim}_{n\to\infty}\mathrm{tr}((I_{l_{2,n}} - (\rho')^2 G^\star G)^{-1}), \end{aligned}$$

and $\mathcal{A}(\rho') = a(\rho') - a(0) = c(\rho') - c(0)$ for $\rho' \in (-1, 1)$. Then we obtain

$$v = T\sigma_{1,*}^2\sigma_{3,*}^2 \frac{2a(\rho)c(\rho) + \partial_\rho\mathcal{A}(\rho)\rho(a(\rho) + c(\rho))}{-2a(\rho)c(\rho)\mathcal{A}(\rho) + \partial_\rho\mathcal{A}(\rho)\rho(a_0 c(\rho) + c_0 a(\rho))}.$$

Example

We change the parameters $\lambda_1, \lambda_2, \sigma$, see behavior of $\hat{\sigma}_n$ and compare the estimator $\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$ of covariation with the Hayashi-Yoshida estimator

$$HY = \sum_{i,j}(Y^1_{S^i} - Y^1_{S^{i-1}})(Y^2_{T^j} - Y^2_{T^{j-1}})1_{I^i \cap J^j \neq \emptyset}.$$

According to Hayashi and Yoshida (2008),
$\sqrt{n}(HY - \sigma_{1,*}\sigma_{3,*}T) \to^d N(0, v_0)$, where

$$v_0 = T\sigma^2_{1,*}\sigma^2_{3,*}\left\{(1+\rho^{-2})\left(\frac{2}{\lambda_1} + \frac{2}{\lambda_2}\right) - \frac{2}{\lambda_1 + \lambda_2}\right\}.$$

Since v can be calculated numerically, we calculate and compare v_0 and v.

Simulation

Table: $T = 1, (\lambda_1, \lambda_2) = (1, 1), (\sigma_1, \sigma_2, \sigma_3) = (1, 1, 0.5)$

n		50	100	500
	true value			
$\hat{\sigma}_{1,n}$	1	0.994 (0.102)	0.998 (0.070)	0.999 (0.031)
$\hat{\sigma}_{2,n}$	1	0.968 (0.129)	0.983 (0.091)	0.996 (0.040)
$\hat{\sigma}_{3,n}$	0.5	0.499 (0.224)	0.502 (0.154)	0.5 (0.067)
$\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$	0.5	0.5 (0.238)	0.503 (0.165)	0.5 (0.071)
HY	0.5	0.501 (0.336)	0.504 (0.236)	0.5 (0.106)
$\sqrt{v/n}$		(0.228)	(0.161)	(0.072)
$\sqrt{v_0/n}$		(0.339)	(0.239)	(0.107)

Simulation

Table: $T = 1, (\lambda_1, \lambda_2) = (1, 1), (\sigma_1, \sigma_2, \sigma_3) = (0.5, 2, 1)$

n		50	100	500
	true value			
$\hat{\sigma}_{1,n}$	0.5	0.497 (0.050)	0.499 (0.035)	0.499 (0.015)
$\hat{\sigma}_{2,n}$	2	1.936 (0.259)	1.968 (0.181)	1.995 (0.079)
$\hat{\sigma}_{3,n}$	1	0.986 (0.449)	0.996 (0.307)	0.997 (0.135)
$\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$	0.5	0.495 (0.239)	0.499 (0.164)	0.498 (0.072)
HY	0.5	0.498 (0.335)	0.499 (0.237)	0.498 (0.108)
$\sqrt{v/n}$		(0.228)	(0.161)	(0.072)
$\sqrt{v_0/n}$		(0.339)	(0.239)	(0.107)

Outline of the proof

We see outline of the proof of main theorem. To approximate $\log(dP_{\sigma_*+b_n^{-1/2}u,n}/dP_{\sigma_*,n})$ by $H_n(\sigma_* + b_n^{-1/2}) - H_n(\sigma_*)$, we consider the following SDE :

$$\begin{cases} dY_s^{r,\sigma} &= \{(1-r)\mu(s, Y_s^{r,\sigma}, \sigma) + r\mu(0, z_0, \sigma)\}ds \\ &\quad +\{(1-r)b(s, Y_s^{r,\sigma}, \sigma) + rb(0, z_0, \sigma)\}dW_s, \quad s \in [0, t] \\ Y_0^{r,\sigma} &= z_0 \end{cases}$$

for $r \in [0, 1]$.

- If $r = 0$, the SDE corresponds to (1).
- If $r = 1$, the SDE is Euler-Maruyama type.

By using these SDE, we consider an approximate density function $\mathbb{P}^r_{\sigma,n}$ for some subdivided grids of original observations.

- $\mathbb{P}^0_{\sigma,n}$: density function of the original probability measure $P_{\sigma,n}$
- $\mathbb{P}^1_{\sigma,n}$: density function of Euler-Maruyama type approximation

Outline of the proof

Let $\check{\mathbb{P}}^r_{\sigma,n}$ be the density function for synchronous observations obtained by unifying $\mathcal{S}$ and $\mathcal{T}$, then $\mathbb{P}^r_{\sigma,n}$ is some integration of $\check{\mathbb{P}}^r_{\sigma,n}$. Hence

$$\log \frac{\mathbb{P}^1_{\sigma,n}}{\mathbb{P}^0_{\sigma,n}} = \int_0^1 \frac{\partial_r \mathbb{P}^r_{\sigma,n}}{\mathbb{P}^r_{\sigma,n}} dr = \int_0^1 \frac{1}{\mathbb{P}^r_{\sigma,n}} \int \frac{\partial_r \check{\mathbb{P}}^r_{\sigma,n}}{\check{\mathbb{P}}^r_{\sigma,n}} \check{\mathbb{P}}^r_{\sigma,n} dr.$$

Then for any $M > 0$,

$$E\left[\left(\log \frac{\mathbb{P}^1_{\sigma,n}}{\mathbb{P}^0_{\sigma,n}}\right)^2 1_{\{\sup_{0\le r\le 1} |\log(\mathbb{P}^r_{\sigma,n}/\mathbb{P}^0_{\sigma_*,n})|\le M\}} \middle| \mathcal{S}, \mathcal{T}\right]$$

$$\le \quad e^M \int_0^1 \int \left(\frac{\partial_r \check{\mathbb{P}}^r_{\sigma,n}}{\check{\mathbb{P}}^r_{\sigma,n}}\right)^2 \check{\mathbb{P}}^r_{\sigma,n} dr \le e^M \sup_r \int \left(\frac{\partial_r \check{\mathbb{P}}^r_{\sigma,n}}{\check{\mathbb{P}}^r_{\sigma,n}}\right)^2 \check{\mathbb{P}}^r_{\sigma,n}.$$

Therefore

$$\log \frac{dP_{\sigma_*+b_n^{-1/2}u,n}}{dP_{\sigma_*,n}} = \log \frac{\mathbb{P}^0_{\sigma_*+b_n^{-1/2}u,n}}{\mathbb{P}^0_{\sigma_*,n}} \sim \log \frac{\mathbb{P}^1_{\sigma_*+b_n^{-1/2}u,n}}{\mathbb{P}^1_{\sigma_*,n}}$$

if $\{\sup_{0\le r\le 1} |\log(\mathbb{P}^r_{\sigma_*+b_n^{-1/2}u,n}/\mathbb{P}^0_{\sigma_*,n})|\}_n$: tight and

$$\sup_r \int \left(\partial_r \check{\mathbb{P}}^r_{\sigma_*+b_n^{-1/2}u,n}/\check{\mathbb{P}}^r_{\sigma_*+b_n^{-1/2}u,n}\right)^2 \check{\mathbb{P}}^r_{\sigma_*+b_n^{-1/2}u,n} \to 0. \qquad (3)$$

Outline of the proof

- tightness of $\{\sup_{0\leq r\leq 1}|\log(\mathbb{P}^r_{\sigma_*+b_n^{-1/2}u,n}/\mathbb{P}^0_{\sigma_*,n})|\}_n$ is proved by using tightness and a moment condition of corresponding density ratio for synchronous density $\check{\mathbb{P}}^r_{\sigma,n}$.
- The relation (3) is proved by Maliavin calculus approach.
- Asymptotical equivalence of $\log(\mathbb{P}^1_{\sigma_*+b_n^{-1/2}u,n}/\mathbb{P}^1_{\sigma_*,n})$ and $H_n(\sigma_*+b_n^{-1/2}u)-H_n(\sigma_*)$ is proved by a little time shift in $b(t,Y_t^1,Y_t^2)$ in $\mathbb{P}^1_{\sigma,n}$ and calculating integral.

Then the proof is completed since $H_n(\sigma_*+b_n^{-1/2}u)-H_n(\sigma_*)$ has a LAMN type property by O. and Yoshida (2012). $\square$

Concluding Remarks

- We proved the LAMN property for diffusion processes with nonsynchronous observations.
- The quasi-maximum likelihood estimator $\hat{\sigma}_n$ and the Bayes type estimator $\tilde{\sigma}_n$ proposed by O. and Yoshida (2012) are proved to be asymptotically efficient.
- In the example of deterministic diffusion coefficients, $\hat{\sigma}_n$ works better than the Hayashi-Yoshida estimator. (The asymptotic variance of $\hat{\sigma}_n$ is about a half of that of the Hayashi-Yoshida estimator.)

References

1. Gobet, E. (2001) Local Asymptotic mixed normality property for elliptic diffusion : a Malliavin calculus approach. Bernoulli, 7, 899-912.
2. Jeganathan, P. (1983) Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. Sankhya Ser. A, 45, 66-87.
3. Ogihara, T. and N. Yoshida. (2012) Quasi-Likelihood Analysis for Stochastic Regression Models with Nonsynchronous Observations. arXiv:1212.4911