Statistical inference for volatility and related limit theorems

Nakahiro YOSHIDA

Graduate School of Mathematical Sciences, University of Tokyo,

December 15, 2010

Asymptotic Statistics, Risk and Computation in Finance and Insurance

- **The topics and references I can show here are limited. For more information, please see**
  - **DYNSTOCH web page**

    http://www.math.ku.dk/~michael/dynstoch/

    DYNSTOCH 2010 Angers

    http://dynstoch.math.univ-angers.fr/spip.php?article9&lang=fr
  - **Statistique Asymptotique des Processus Stochastiques**

    http://subaru.univ-lemans.fr/sciences/statist/index.php?page=liste
  - **the references in my papers**

    http://www2.ms.u-tokyo.ac.jp/probstat/?page_id=23

    nakahiro@ms.u-tokyo.ac.jp

Menu

- **Quasi-likelihood analysis (QLA)**
  - **QLA for ergodic diffusion processes**
  - **QLA for ergodic jump-diffusions**
  - **QLA for volatility in finite time-horizon**
- **Nonsynchronous covariance estimation**
  - **Limit theorem**
  - **Lead-lag estimate**
- **Higher-order asymptotics for the realized volatility**

Quasi-likelihood analysis (QLA): QLA for ergodic diffusion processes

QLA for ergodic diffusion processes

A d-dimensional stationary diffusion process satisfying the stochastic differential equation

$$dX_t = a(X_t, \theta_2)dt + b(X_t, \theta_1)dw_t, \quad X_0 = x_0.$$

- **w_t is an r-dimensional standard Wiener process independent of the initial value x_0.**
- **θ_1 and θ_2 are unknown parameters with $\theta_i \in \Theta_i \subset \mathbb{R}^{m_i}$**
- **The distribution of x_0 possibly depends on the parameters.**
- **The true value of the unknown parameter is denoted by $\theta^* = (\theta_1^*, \theta_2^*)$.**

QLA for ergodic diffusion processes: Assumptions*

[D1] (i) **The mappings** $a : \mathbb{R}^d \times \Theta_2 \to \mathbb{R}^d$ **and** $b : \mathbb{R}^d \times \Theta_1 \to \mathbb{R}^d$ **have continuous derivatives satisfying**

$$\sup_{\theta_2 \in \Theta_2} \left| \partial_{\theta_2}^i a(x, \theta_2) \right| \leq C(1 + |x|)^C \quad (0 \leq i \leq 4)$$

and

$$\sup_{\theta_1 \in \Theta_1} \left| \partial_x^j \partial_{\theta_1}^i b(x, \theta_1) \right| \leq C(1 + |x|)^C \quad (0 \leq i \leq 4 \,; 0 \leq j \leq 2)$$

for some constant C**.**

(ii) $B(x, \theta_1) = bb'(x, \theta_1)$ **is elliptic uniformly in** (x, θ_1)**.**

(iii) **For some constant** C**,**

$$\sup_{\theta_2 \in \Theta_2} |a(x_1, \theta_2) - a(x_2, \theta_2)| + \sup_{\theta_1 \in \Theta_1} |b(x_1, \theta_1) - b(x_2, \theta_1)| \\ \leq C|x_1 - x_2| \quad (x_1, x_2 \in \mathbb{R}^d).$$

(iv) $X_0 \in \bigcap_{p>0} L^p(P_{\theta^*})$**.**

QLA for ergodic diffusion processes: Quasi-likelihood function

Now we want to estimate the unknown parameters with the discrete-time observations

$$\mathbf{x}_n = (X_{t_i})_{i=0}^{n},$$

where $t_i = ih$ **with** $h = h_n$ **depending on** $n \in \mathbb{N}$**.**

For this purpose, we consider a quasi-likelihood function

$$\begin{aligned} & p_n(\mathbf{x}_n, \theta) \\ &= \prod_{i=1}^{n} \frac{1}{(2\pi h)^{d/2} |B(X_{t_{i-1}}, \theta_1)|^{1/2}} \\ & \quad \times \exp\left(-\frac{1}{2h} B(X_{t_{i-1}}, \theta_1)^{-1} \left[(\Delta_i X - h a(X_{t_{i-1}}, \theta_2))^{\otimes 2} \right] \right) \end{aligned}$$

with $\Delta_i X = X_{t_i} - X_{t_{i-1}}$**.**

QLA for ergodic diffusion processes: Quasi MLE

The maximum likelihood type estimator

$$\hat{\theta}_n = (\hat{\theta}_{1,n}, \hat{\theta}_{2,n})$$

is an estimator that maximizes $p_n(\mathbf{x}_n, \theta)$ **in** $\theta = (\theta_1, \theta_2) \in \Theta = \Theta_1 \times \Theta_2$**.**

QLA for ergodic diffusion processes: Mixing condition*

We assume a mixing property for X:

[D2] **There exists a positive constant a such that**

$$\alpha_X(h) \leq a^{-1} e^{-ah} \quad (h > 0),$$

where

$$\alpha_X(h) = \sup_{t \in \mathbb{R}_+} \sup_{\substack{A \in \sigma[X_r; r \leq t], \\ B \in \sigma[X_r; r \geq t+h]}} \left| P_{\theta^*}[A \cap B] - P_{\theta^*}[A] P_{\theta^*}[B] \right|.$$

We will assume that $h \to 0$ and $nh^2 \to 0$ as $n \to \infty$. Moreover, we assume that for some positive constant ϵ_0, $nh \geq n^{\epsilon_0}$ for large n.

QLA for ergodic diffusion processes: Information matrices*

Set

$$
\begin{aligned}
\Gamma_1(\theta^*)[u_1,u_1] &:= \Gamma(\theta_2;\theta^*)[u_1,u_1] \\
&:= \frac{1}{2}\int\Big\{\partial^2_{\theta_1}B^{-1}(x,\theta_1)[u_1^{\otimes 2}, B(x,\theta_1^*)] \\
&\qquad +\partial^2_{\theta_1}\log\frac{|B(x,\theta_1)|}{|B(x,\theta_1^*)|}[u_1^{\otimes 2}]\Big\}\Big|_{\theta_1=\theta_1^*}\,\nu(dx) \\
&= \frac{1}{2}\int \mathrm{tr}\left\{B^{-1}(\partial_{\theta_1}B)B^{-1}(\partial_{\theta_1}B)(x,\theta_1^*)[u_1^{\otimes 2}]\right\}\,\nu(dx)
\end{aligned}
$$

for $u_1\in\mathbb{R}^{m_1}$. **Let**

$$
\Gamma_2(\theta^*)[u_2^{\otimes 2}] = \int_{\mathbb{R}^d} B(x,\theta_1^*)^{-1}[\partial_{\theta_2}a(x,\theta_2^*)[u_2],\partial_{\theta_2}a(x,\theta_2^*)[u_2]]\,\nu(dx)
$$

for $u_1\in\mathbb{R}^{m_2}$.

QLA for ergodic diffusion processes

Theorem 1. Under [D1], [D2] **and certain identifiability conditions, for any sequence of the maximum likelihood type estimators for** $\theta = (\theta_1, \theta_2)$**, it holds that**

$$\left(\sqrt{n}(\hat{\theta}_1 - \theta_1^*), \sqrt{nh}(\hat{\theta}_2 - \theta_2^*)\right)$$
$$\to^d (\zeta_1, \zeta_2) \sim N_{m_1+m_2}\left(0, \mathrm{diag}\left[\Gamma_1(\theta^*)^{-1}, \Gamma_2(\theta^*)^{-1}\right]\right)$$

as $n \to \infty$**. Moreover,**

$$E_{\theta^*}\left[f(\sqrt{n}(\hat{\theta}_1 - \theta_1^*), \sqrt{nh}(\hat{\theta}_2 - \theta_2^*))\right] \to \mathbb{E}\left[f(\zeta_1, \zeta_2)\right]$$

as $n \to \infty$ **for all continuous functions** f **of at most polynomial growth.**

QLA for ergodic diffusion processes: Comments*

1. Prakasa Rao (1983,1988) presented asymptotic results for an ergodic diffusion process under a sampling scheme.
2. The joint weak convergence was given in Yoshida (1992).
3. Kessler (1997) treated a local Gaussian approximation with higher order correction terms to relax the rate of convergence of h to zero.

QLA for ergodic diffusion processes: Bayesian type estimator

The Bayesian method can apply to the stochastic differential equations.

Asymptotic results can be obtained.

QLA for ergodic diffusion processes: Bayesian type estimator

We consider the same estimation problem as before:

- **a d-dimensional stationary diffusion process**

$$dX_t = a(X_t, \theta_2)dt + b(X_t, \theta_1)dw_t, \quad X_0 = x_0.$$

- **the discrete-time observations**

$$\mathbf{x}_n = (X_{t_i})_{i=0}^n,$$

where $t_i = ih$ with $h = h_n$ depending on $n \in \mathbb{N}$.

QLA for ergodic diffusion processes: Bayesian type estimator

- **a quasi-likelihood function**

$$
\begin{aligned}
& p_n(\mathbf{x}_n, \theta) \\
&= \prod_{i=1}^{n} \frac{1}{(2\pi h)^{d/2} |B(X_{t_{i-1}}, \theta_1)|^{1/2}} \\
&\quad \times \exp\left(-\frac{1}{2h} B(X_{t_{i-1}}, \theta_1)^{-1} \left[(\Delta_i X - h a(X_{t_{i-1}}, \theta_2))^{\otimes 2} \right] \right)
\end{aligned}
$$

with $\Delta_i X = X_{t_i} - X_{t_{i-1}}$**.**

QLA for ergodic diffusion processes: Bayesian type estimator

Adaptive Bayesian type estimator (Y 2005)

The adaptive Bayesian type estimator is defined as follows.

(i) First fix a value of θ_2 and compute the Bayesian estimator with some prior distribution of θ_1 by the quasi-likelihood.

(ii) Next by using the first step estimator, we compute the Bayesian estimator for θ_1.

QLA for ergodic diffusion processes: Adaptive Bayesian type estimator

Theorem 2. **For adaptive Bayes type estimator $(\tilde{\theta}_{1,n}, \tilde{\theta}_{2,n})$ for $\theta = (\theta_1, \theta_2)$, it holds that**

$$\left(\sqrt{n}(\tilde{\theta}_{1,n} - \theta_1^*), \sqrt{nh}(\tilde{\theta}_{2,n} - \theta_2^*)\right)$$
$$\to^d (\zeta_1, \zeta_2) \sim N_{m_1+m_2}\left(0, \mathrm{diag}\left[\Gamma_1(\theta^*)^{-1}, \Gamma_2(\theta^*)^{-1}\right]\right)$$

as $n \to \infty$. Moreover,

$$E_{\theta^*}\left[f(\sqrt{n}(\tilde{\theta}_1 - \theta_1^*), \sqrt{nh}(\tilde{\theta}_2 - \theta_2^*))\right] \to \mathbb{E}\left[f(\zeta_1, \zeta_2)\right]$$

as $n \to \infty$ for all continuous functions f of at most polynomial growth.

QLA for ergodic diffusion processes: Comments*

1. The program by Ibragimov-Has'minskii and Kutoyants shows the way of constructing the QLA for stochastic processes.
2. Polynomial type large deviation inequality plays a role. (2005, now on-line of AISM)
3. Recently Uchida and Yoshida proved the same properties under the condition that $nh^p \to 0$ for any $p \geq 2$ by a more precise quasi-likelihood function, using an adaptive method.

QLA for ergodic diffusion processes: Comments

4. Shimizu and Yoshida (2006) obtained asymptotic normality for a maximul likelihood type estimator for a diffusion process with jumps.
5. Ogihara and Yoshida (2009) gave quasi-likelihood analysis and proved asymptotic normality and moment convergence for a diffusion process with jumps.

Quasi-likelihood analysis (QLA): QLA for volatility in the finite time-horizon

QLA for volatility in the finite time-horizon

- **An m-dimensional Itô process satisfying the stochastic differential equation**

$$dY_t = b_t dt + \sigma(X_t, \theta) dw_t, \quad t \in [0, T], \tag{1}$$

- w: **an r-dimensional standard Wiener process on some stochastic basis** $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, P)$
- b **and** X: **progressively measurable processes with values in** $\mathbf{R}^m$ **and** $\mathbf{R}^{\mathsf{d}}$**, respectively.** b **is unobservable, completely unknown.**
- σ: **an** $\mathbf{R}^m \otimes \mathbf{R}^r$**-valued function defined on** $\mathbf{R}^{\mathsf{d}} \times \Theta$,
- Θ: **is a bounded domain in** $\mathbf{R}^p$
- θ^* **denotes the true value of** θ.

QLA for volatility in the finite time-horizon

- **Data:** $\mathbf{Z}_n = (X_{t_k}, Y_{t_k})_{0 \le k \le n}$ **with** $t_k = kh$ **for** $h = h_n = T/n$.
- **For example, when** $b_t = b(Y_t, t)$ **and** $X_t = (Y_t, t)$**,** Y **is the time-inhomogeneous diffusion process.**

QLA for volatility in the finite time-horizon

- **Asymptotic theory of estimation of the volatility parameter with high frequency data observed on a fixed interval has been developed.**
  - **Dohnal (1987): the local asymptotic mixed normality (LAMN) property for the likelihood**
  - **Genon-Catalot and Jacod (1993, 1994): the asymptotic mixed normality of the minimum contrast estimator**

QLA for volatility*

- We will see the asymptotic mixed normality and convergence of moments of both the maximum likelihood type estimator and the Bayesian type estimator for a quasi-likelihood function.
- The Ibragimov-Has'minskii-Kutoyants scheme is applied.
- A key point is to obtain the polynomial type large deviation inequality for the statistical random field.

QLA for volatility: Recall the model

- **An m-dimensional Itô process satisfying the stochastic differential equation**

$$dY_t = b_t dt + \sigma(X_t, \theta) dw_t, \quad t \in [0, T], \tag{2}$$

- **Data:** $\mathbf{Z}_n = (X_{t_k}, Y_{t_k})_{0 \le k \le n}$ **with** $t_k = kh$ **for** $h = h_n = T/n$**.**

QLA for volatility

- **Quasi log likelihood function:**

$$\mathbb{H}_n(\theta) = -\frac{nm}{2}\log(2\pi h) - \frac{1}{2}\sum_{k=1}^{n}\Bigg\{\log\det S(X_{t_{k-1}},\theta) + h^{-1}S^{-1}(X_{t_{k-1}},\theta)[(\Delta_k Y)^{\otimes 2}]\Bigg\},$$

$$S = \sigma^{\otimes 2}.$$

QLA for volatility

- $\hat{\theta}_n$: **the maximum likelihood type estimator defined as**

$$\mathbb{H}_n(\hat{\theta}_n) = \sup_{\theta\in\Theta} \mathbb{H}_n(\theta). \tag{3}$$

- **Let** $\tilde{\theta}_n$ **be the Bayes type estimator for a prior density** $\pi : \Theta \to \mathbf{R}_+$ **defined as**

$$\tilde{\theta}_n = \left(\int_\Theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta\right)^{-1} \int_\Theta \theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta. \tag{4}$$

We assume that π **is continuous and** $0 < \inf_{\theta\in\Theta} \pi(\theta) \leq \sup_{\theta\in\Theta} \pi(\theta) < \infty$**.**

QLA for volatility*

- **Define the random field $\mathbb{Z}_n(u)$ for $u \in \mathbb{U}_n$ by**

$$\mathbb{Z}_n(u) = \exp\left\{\mathbb{H}_n\left(\theta^* + \frac{1}{\sqrt{n}}u\right) - \mathbb{H}_n(\theta^*)\right\},$$

- **Let**

$$\mathbb{Z}(u) = \exp\left(\Gamma(\theta^*)^{1/2}\zeta[u] - \frac{1}{2}\Gamma(\theta^*)[u,u]\right),$$

where $\Gamma(\theta^*) = (\Gamma^{ij}(\theta^*))_{i,j=1,\ldots,p}$ with

$$\Gamma^{ij}(\theta^*) = \frac{1}{2T}\int_0^T tr\left((\partial_{\theta_i}S)S^{-1}(\partial_{\theta_j}S)S^{-1}(X_t,\theta^*)\right)dt$$

and ζ is a p-dimensional standard normal random variable independent of $\Gamma(\theta^*)$.

QLA for volatility*

- **By setting** $\tilde{u}_n = \sqrt{n}(\tilde{\theta}_n - \theta^*)$**,**

$$\begin{aligned} \tilde{u}_n &= \left(\int_{\mathbb{U}_n} \mathbb{Z}_n(u)\pi(\theta^* + (1/\sqrt{n})u)du\right)^{-1} \\ &\quad \times \int_{\mathbb{U}_n} u\mathbb{Z}_n(u)\pi(\theta^* + (1/\sqrt{n})u)du. \end{aligned} \tag{5}$$

- **Let**

$$\tilde{u} = \left(\int_{\mathbf{R}^p} \mathbb{Z}(u)du\right)^{-1} \int_{\mathbf{R}^p} u\mathbb{Z}(u)du \ \left(= \Gamma(\theta^*)^{-1/2}\zeta\right) \tag{6}$$

QLA for volatility*

- **Then the convergences**

  $\tilde{u}_n = \sqrt{n}(\tilde{\theta}_n - \theta^*) \to^d \tilde{u}$ and $E[f(\tilde{u}_n)] \to E[f(\tilde{u})]$ $(f \in C_\uparrow(\mathbb{R}^p)$

  as well as the quasi MLE $\hat{\theta}_n$ follow from the convergence

  $\mathbb{Z}_n \to^d \mathbb{Z}$ plus

  Polynomial type large deviation inequality

$$P\left[\sup_{u \in \mathbb{R}^p : |u| \geq r} \mathbb{Z}_n(u) \geq e^{-r}\right] \leq \frac{C_L}{r^L}$$

QLA for volatility: Quasi-MLE

Theorem 3.

(a) $\sqrt{n}(\hat{\theta}_n - \theta^*) \to^{d_s(\mathcal{F}_T)} \Gamma(\theta^*)^{-1/2}\zeta$

(b) For all continuous functions f of at most polynomial growth,

$$E\left[f(\sqrt{n}(\hat{\theta}_n - \theta^*))\right] \to \mathbb{E}\left[f(\Gamma(\theta^*)^{-1/2}\zeta)\right]$$

as $n \to \infty$.

QLA for volatility: Quasi-Bayesian Estimator

Theorem 4. (Uchida and Y 2008)

(a) $\sqrt{n}(\tilde{\theta}_n - \theta^*) \to^{d_s(\mathcal{F}_T)} \Gamma(\theta^*)^{-1/2}\zeta$

(b) For all continuous functions f **of at most polynomial growth,**

$$E\left[f(\sqrt{n}(\tilde{\theta}_n - \theta^*))\right] \to \mathbb{E}\left[f(\Gamma(\theta^*)^{-1/2}\zeta)\right]$$

as $n \to \infty$**.**

QLA for volatility: Examples and simulation results

- **Consider the one-dimensional diffusion process**

  $$dX_t = X_t dt + \exp\{\theta \sin^2 X_t\} dw_t, \quad t \in [0,1], \quad X_0 = 0,$$

  where $\theta \in [-\pi, \pi]$**.**

- **the uniform prior** $\pi(\theta)$

- **The simulations were done for each** $h_n = 1/50$**,** $1/250$**,** $1/500$**.**

- **For the true model with** $\theta^* = 1$**,** 10000 **independent sample paths are generated by the Milstein scheme, and the means and the standard deviations of the estimators are computed and shown in Table 1 below.**

Table 1:

h_n	$\hat{\theta}_n$		$\tilde{\theta}_n$	
	mean	**s.d.**	**mean**	**s.d.**
1/50	**0.90938**	**0.55704**	**0.97465**	**0.47647**
1/250	**0.98181**	**0.23022**	**0.99714**	**0.22370**
1/500	**0.99354**	**0.16436**	**1.00164**	**0.16236**

- **The statistical model is completely degenerate at $t = 0$. Naïve nondegeneracy conditions cannot apply. However, it is solved by another machinery (Uchida and Y).**

Nonsynchronous covariance estimation

Nonsynchronous covariance estimation

- $(X_t, Y_t)_{t\in[0,T]}$ **: two-dimensional Iô process**
- **A semiparametric problem arises if we want to estimate the (possibly random) "parameter"**

$$\theta = [X, Y].$$

Nonsynchronous covariance estimation

- **If the two sequences of data are synchronously observed, the sum of cross products**

$$\sum_{i=1}^{N_1} \Delta_i X \Delta_i Y$$

is a natural estimator of θ because it may converge in probability to θ:

$$\sum_{i=1}^{N_1} \Delta_i X \Delta_i Y \to^p \theta$$

if the maximum lag of the time points tends to 0 in probability, as it is well known in the stochastic analysis.

Indeed, we can regard this as a definition of $[X, Y]$.

Nonsynchronous covariance estimation

Non-synchronous sampling.

- **The families**
  $\Pi_1 = \{I^i, i = 1, \ldots, N_1\}$ **and** $\Pi_2 = \{J^j, j = 1, \ldots, N_2\}$ **are partitions of the interval** $[0, T]$ **corresponding to the observing times of** X_1 **and** X_2 **respectively.**

- **Notation**
  $\Delta_i X = \int_{I^i} dX_t$ **and** $\Delta_j Y = \int_{J^j} dY_t$**.**

Nonsynchronous covariance estimation

- **Naïve synchronization**
  - **If one applies the "realized volatility" estimator to the real tick data, a certain interpolation such as previous-tick interpolation and linear interpolation will be necessary.**
  - **However, it is known that such a naive synchronization causes estimation bias.**
  - **Nonsynchronicity can cause "Epps effect".**

Nonsynchronous covariance estimation

- **Malliavin and Mancino (2002) have proposed a Fourier transform based estimator.**
- **Reno (2003) utilizes it to investigate numerically biases of Epps-type in case of a bivariate continuous-time version of GARCH(1,1) process.**

Nonsynchronous covariance estimation

- **For estimation of θ, Hayashi and Yoshida proposed**

$$\hat{\theta} = \sum_i \sum_j \Delta_i X \Delta_j Y \, 1_{\{I^i \cap J^j \neq \emptyset\}}$$

- **This estimator satisfies:**
  - **No interpolation is used so that it does not depend on any tuning parameter such as the grid size.**
  - **It is a finite sum. No cut-off number is involved.**
  - **It attains asymptotically minimum variance.**
  - **The summation is essentially one-dimensional.**

Nonsynchronous covariance estimation*

Theoretical statistical requires the basic asymptotic properties:

- **consistency of the estimator**
- **asymptotic distribution of the error**
- **efficiency and optimality**
- **precise approximation to the error distribution**

Nonsynchronous covariance estimation: Consistency

The estimator $\hat{\theta}$ is consistent as the maximum lag of the observation times tends to 0 in probability

$$\hat{\theta} \longrightarrow^{p} \theta$$

whenever

$$\max_{i,j}\{|I^i|, |J^j|\} \longrightarrow^{p} 0.$$

Nonsynchronous covariance estimation: Nonsynchronous covariance process*

- $\mathcal{B} = (\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$: **a stochastic basis**
- $X = (X_t)_{t\in\mathbb{R}_+}$ **and** $Y = (Y_t)_{t\in\mathbb{R}_+}$: **Itô processes**
- $\left(S^i\right)_{i\in\mathbb{Z}_+}$ **and** $\left(T^j\right)_{j\in\mathbb{Z}_+}$: **two sequences of stopping times that are increasing a.s.,** $S^i \uparrow \infty$ **and** $T^j \uparrow \infty$, **and** $S^0 = 0$, $T^0 = 0$
- **Random intervals and indicator functions:**

$$I^i = \left[S^{i-1}, S^i\right), \quad J^j = \left[T^{j-1}, T^j\right),$$

$$I^i_t = 1_{\left[S^{i-1}, S^i\right)}(t), \quad J^j_t = 1_{\left[T^{j-1}, T^j\right)}(t),$$

$$I^i(t) = \left[S^{i-1} \wedge t, S^i \wedge t\right), \quad J^j(t) = \left[T^{j-1} \wedge t, T^j \wedge t\right),$$

$$r_n(t) = \sup_{i\in\mathbb{N}} |I^i(t)| \vee \sup_{j\in\mathbb{N}} |J^j(t)|.$$

Nonsynchronous covariance estimation: Nonsynchronous covariance process

For a stochastic process V and an interval I, let $V(I)_t = \int^t 1_I(s-)dV_s$.

Definition 1. **The nonsynchronous covariation process of X and Y associated with sampling designs $\mathcal{I} := \left(I^i\right)_{i\in\mathbb{N}}$ and $\mathcal{J} := \left(J^j\right)_{j\in\mathbb{N}}$ is the process**

$$\{X, Y\}_t = \sum_{i,j=1}^{\infty} X(I^i)_t Y(J^j)_t 1_{\left\{I^i(t)\cap J^j(t)\neq\varnothing\right\}}$$

Nonsynchronous covariance estimation: Stable convergence of the estimation error

- **The estimation error of $\{X, Y\}$ is given by**

$$M_t^n := \{X, Y\}_t - [X, Y]_t = \textstyle\sum_{i,j} L_t^{ij} K_t^{ij}, \qquad (7)$$

where $K_t^{ij} = 1_{\{I^i(t) \cap J^j(t) \neq \varnothing\}}$ **and**

$$L_t^{ij} = \left(I_-^i \cdot X\right)_- \cdot \left(J_-^j \cdot Y\right)_t + \left(J_-^j \cdot Y\right)_- \cdot \left(I_-^i \cdot X\right)_t .$$

- **When X and Y are local martingales, M_t^n is a local martingale with**

$$[M^n, M^n]_t = \sum_{i,j,i',j'} \left(K_-^{ij} K_-^{i'j'}\right) \cdot \left[L^{ij}, L^{i'j'}\right]_t \qquad (8)$$

Nonsynchronous covariance estimation: Stable convergence of the estimation error*

A sequence of random elements X^n defined on a probability space $(\Omega, \mathcal{F}, P)$ is said to converge stably in law to a random element X defined on an appropriate extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ of $(\Omega, \mathcal{F}, P)$ if $E\left[Y g\left(X^n\right)\right] \to E\left[Y g\left(X\right)\right]$ for any $\mathcal{F}$-measurable and bounded random variable Y and any bounded and continuous function g.

Nonsynchronous covariance estimation: Stable convergence of the estimation error

Denote $[X]=[X,X]$ **and** $[Y]=[Y,Y]$ **as usual. Let**

$$\bar{V}_t^n = \sum_{i,j}[X]\left(I^i(t)\right)[Y]\left(J^j(t)\right)K_t^{ij} + \sum_i [X,Y]\left(I^i(t)\right)^2 + \sum_j [X,Y]\left(J^j(t)\right)^2 - \sum_{i,j}[X,Y]\left(\left(I^i\cap J^j\right)(t)\right)^2.$$

[A1] There exists an F-adapted, nondecreasing, continuous process $(V_t)_{t\in\mathbb{R}_+}$ **such that** $b_n^{-1}\bar{V}_t^n \to^P V_t$ **as** $n\to\infty$ **for every** t**.**

Nonsynchronous covariance estimation: Stable convergence of the estimation error

Theorem 5. (HY 2006, 2008, 2010 SPA on-line) Suppose that $[A1]$ and a regularity condition are fulfilled and that there exists an F-predictable process w such that $V_\cdot = \int_0^\cdot w_s^2 ds$. Then

$$b_n^{-\frac{1}{2}} M^n \quad \overset{\text{Stably}}{\longrightarrow} \quad M$$

in $C(\mathbb{R}_+)$ as $n \to \infty$, where $M = \int_0^\cdot w_s d\widetilde{W}_s$ and $\widetilde{W}$ is a one-dimensional Wiener process (defined on an extension of $\mathcal{B}$) independent of $\mathcal{F}$.

Nonsynchronous covariance estimation: Convergence of the sampling measures*

The empirical distribution functions of the sampling times are defined by

$$H_n^1(t) := \sum_i |I^i(t)|^2, \; H_n^2(t) := \sum_j |J^j(t)|^2,$$

$$H_n^{1\cap 2}(t) := \sum_{i,j} |(I^i \cap J^j)(t)|^2, \; H_n^{1*2}(t) := \sum_{i,j} |I^i(t)||J^j(t)|K_t^{ij},$$

where $|\cdot|$ is the Lebesgue measure.

Nonsynchronous covariance estimation: Convergence of the sampling measures*

[A1′] There exists a possibly random, nondecreasing, functions $H^1, H^2, H^{1\cap 2}$ and H^{1*2} on $[0,T]$, such that each $H^k = \int_0^t h_s^k ds$ for some density h^k, and that $b_n^{-1} H_n^k(t) \xrightarrow{P} H^k(t)$ as $n \to \infty$ for every $t \in \mathbb{R}_+$ and $k = 1, 2, 1\cap 2, 1*2$.

Nonsynchronous covariance estimation: Convergence of the sampling measures*

Theorem 6. (HY 2006, 2008)
Suppose that $[A1']$ and certain regularity conditions are fulfilled, and that each $[X]$, $[Y]$ and $[X,Y]$ is absolutely continuous with a bounded derivative a.s. Then

$$b_n^{-1/2}\left(\{X,Y\} - [X,Y]\right) \xrightarrow{\mathcal{L}} M$$

in $C(\mathbb{R}_+)$ as $n \to \infty$, where M is the process given in Theorem 5 with w_s given by

$$w_s = \sqrt{[X]'_s[Y]'_s h_s^{1*2} + \left([X,Y]'\right)^2\left(h_s^1 + h_s^2 - h_s^{1\cap 2}\right)}. \quad (9)$$

Nonsynchronous covariance estimation: Example: Poisson sampling

The partitions Π_i is given by a Poison random measure on $[0,T]$ with intensity np_i for each $i = 1,2$. Suppose that $\Pi = (\Pi_n^1, \Pi_n^2)$ is independent of (X,Y).

If the functions σ_1, σ_2 and ρ are continuous, then the sequence $\sqrt{n}(\hat{\theta}_n - \theta)$ converges in distribution to a centered Gaussian random variable with variance

$$\mathfrak{c} = \left(\frac{2}{p_1} + \frac{2}{p_2}\right) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 (1+\rho_t^2) dt - \frac{2}{p_1+p_2} \int_0^T (\sigma_{1,t}\sigma_{2,t}\rho_t)^2 dt.$$

Nonsynchronous covariance estimation:* Comments

- Related works are Barndorff-Nielsen and Shephard (2004), and Mykland and Zhang (2006), and Hoshikawa, Kanatani, Nagai and Nishiyama (2008).
- There is vast literature on nonsynchronisity with microstructure noise. Robert and Rosenbaum (2008) gave a new insight into the nonsynchronous covariance estimator under microstructure noise. See also Ubukata and Oya (2008).
- Recently, Markus Bibinger proposed a rate-optimal estimator of a new version of the nonsynchronous covariance estimator to overcome the microstructure noise.

- **It is possible to derive asymptotic expansion of M_T^n in the case without feedback to the diffusion coefficient, where the first order limit is central (Dalayan and Y, to appear in AIHP).**

Nonsynchronous covariance estimation: Data analysis with YUIMA Package

```
> load(file=”ba.data”)
> load(file=”ge.data”)
> load(file=”gm.data”)
> load(file=”cc.data”)
> all.yuima<-cbind.yuima(ba.data,ge.data,gm.data,cc.data)
>cce(all.yuima)
```

	[,1]	[,2]	[,3]	[,4]
[1,]	**9.138171e-04**	**7.284301e-05**	**1.139381e-04**	**1.220833e-04**
[2,]	**7.284301e-05**	**8.312598e-04**	**5.703226e-05**	**8.153857e-05**
[3,]	**1.139381e-04**	**5.703226e-05**	**3.617391e-04**	**5.319538e-05**
[4,]	**1.220833e-04**	**8.153857e-05**	**5.319538e-05**	**3.014167e-04**

Nonsynchronous covariance estimation: Lead-lag estimation

- **Let** $X = (X_t)_{t \in \mathbb{R}_+}$ **and** $\overset{\circ}{Y} = (\overset{\circ}{Y}_t)_{t \in [-\theta^*, \infty)}$ **be Itô processes for a suitable filtration, and assume that** $Y = (Y_t)_{t \in \mathbb{R}_+}$ **is given by** $Y_t = \overset{\circ}{Y}_{t-\theta^*}$**.**
- **Estimation of** θ^* **deserve investigation because when** $\theta^* > 0$**,** X **is regarded as the leader and** Y **as the follower.**
- **We propose a lead-lag estimator and provide the convergence rate. This is a joint work with M. Hoffmann and M. Rosenbaum.**

Nonsynchronous covariance estimation: Lead-lag estimation

- **In this situation, we proposed the estimator**

$$\hat{\theta}_n = \mathbf{argmax}\ |U^n(\theta)|,$$

where

$$U^n(\theta) = \sum_{I,J:\overline{I}\le T} X(I)Y(J)1_{\{I\cap J_{-\theta}\neq\emptyset\}}.$$

Nonsynchronous covariance estimation: Lead-lag estimation

- **We can prove the consistency of $\hat{\theta}_n$.**

Theorem 7. **Under certain regularity conditions,**

$$\bar{r}_n^{-1}(\hat{\theta}_n - \theta^*) \to^p 0 \quad n \to \infty$$

on the event $\{[X, \overset{\circ}{Y}]_T \neq 0\}$ for a sequence of positive constants $\bar{r}_n$ tending to 0 as $n \to \infty$ such that $r_n/\bar{r}_n \to^p 0$ as $n \to \infty$. Here r_n is the maximum length of the inter-arrival times of observations in $[0, T]$.

Nonsynchronous covariance estimation: Lead-lag estimation

sec. cor.

BA-CCE

1.4751087 0.2348809

BA-GE

-18.1460249 0.1311659

BA-GM

-4.1453611 0.1692068

CCE-GE

-27.4679106 0.1760048

CCE-GM

120.3912058 0.2170557

GE-GM

1.7497747 0.1282431

Higher-order asymptotics for the realized volatility

Asymptotic expansion

- **Small σ expansion**
  - **Watanabe (AP1987), Kusuoka and Stroock (JFA1991)**
  - **Applications to statistics:**
    Y (PTRF1992,1993),
    Dermoune and Kutoyants (Stochastics1995),
    Sakamoto and Y (JMA1996, SISP1998),
    Uchida and Y (SISP2004),
    Masuda and Y (StatProbLet2004), ...

- **Application to option pricing:**
  Y (JJSS1992*),
  Kunitomo and Takahashi (MathFinance2001),
  Uchida and Y (SISP2004),
  Takahashi and Y (SISP2004, JJSS2005),
  Osajima (SSRN2007),
  Takahashi and Takehara (2009,2010),
  Andersen and Hutchings (SSRN2009),
  Antonov and Misirpashaev (SSRN2009),
  Chenxu Li (ColumbiaUniv2010),
  ....

* http://www.journalarchive.jst.go.jp/jnlpdf.php?cdjournal=jjss1970&cdvol=22&noissue=2&startpage=139&lang=ja&from=jnltoc

- **Mixing expansion:**
  - **Kusuoka and Y (PTRF2000), Y (PTRF2004)**
  - **Applications to statistics:**
    Y (PTRF97),
    Sakamoto and Y (JJSS2003, AISM2004, AISM2009,
    JJSS2008, CommStat2010),
    Uchida and Y (SISP2006,SUTJMath2006),
    Kutoyants and Y (SISP2007), ...
  - **Applications to finance: Masuda and Y (SPA2005)**
  - **Regenerative method: Fukasawa (PTRF2008)**

- **Distributional martingale expansion (Central limit)**
  - **Yoshida (PTRF1997)**
  - **Statistics: Y, Sakamoto and Y (SISP1998),**
  - **Finance: Fukasawa (FinanceStoch2009)**
- **Here we discuss the martingale expansion in mixed normal limit and its application.**

Question: Quadratic form for a diffusion process

- **stochastic differential equation**

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dw_s.$$

- **quadratic form of the increments of X:**

$$U_n = \sum_{j=1}^n c(X_{t_{j-1}})(\Delta_j X)^2,$$

where $\Delta_j X = X_{t_j} - X_{t_{j-1}}$ **and** $t_j = j/n$**.**

- **Give the asymptotic expansion for the normalized error**

$$Z_n = \sqrt{n}(U_n - U_\infty),$$

where $U_\infty = \int_0^1 c(X_s)\sigma(X_s)^2 ds$**.**

Stochastic expansion

$$Z_n = M_1^n + \frac{1}{\sqrt{n}} N_n,$$

where

$$M_t^n = \sqrt{n} \sum_{j=1}^{n} 2 c_{t_{j-1}} \sigma_{t_{j-1}}^2 \int_{t_{j-1}}^{t} \int_{t_{j-1}}^{s} dw_r dw_s,$$

and

$$N_n = 6n \sum_{j=1}^{n} c_{t_{j-1}} \sigma_{t_{j-1}} \sigma_{t_{j-1}}^{[1]} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} \int_{t_{j-1}}^{s} dw_u dw_s dw_t$$
$$+2 \sum_{j=1}^{n} c_{t_{j-1}} b_{t_{j-1}} \sigma_{t_{j-1}} \int_{t_{j-1}}^{t_j} dw_t$$

$$
\begin{aligned}
&+2n\sum_{j=1}^{n} c_{t_{j-1}}\sigma_{t_{j-1}}\sigma^{[1]}_{t_{j-1}}\int_{t_{j-1}}^{t_j}(t-t_{j-1})dw_t\\
&+2n\sum_{j=1}^{n} c_{t_{j-1}}\sigma_{t_{j-1}}b_{t_{j-1}}\int_{t_{j-1}}^{t_j}\int_{t_{j-1}}^{t} dw_s dt\\
&+n^{-1}\sum_{j=1}^{n} c_{t_{j-1}}b^2_{t_{j-1}}+n^{-1}\sum_{j=1}^{n} c_{t_{j-1}}\sigma_{t_{j-1}}b^{[1]}_{t_{j-1}}\\
&-n\sum_{j=1}^{n} c^{[1]}_{t_{j-1}}\sigma^2_{t_{j-1}}\int_{t_{j-1}}^{t_j}\int_{t_{j-1}}^{t} dw_s dt\\
&-\frac{1}{2n}\sum_{j=1}^{n} c^{[0]}_{t_{j-1}}\sigma^2_{t_{j-1}}-\frac{1}{n}\sum_{j=1}^{n} c^{[1]}_{t_{j-1}}\sigma_{t_{j-1}}\sigma^{[1]}_{t_{j-1}}+o_M(1).
\end{aligned}
$$

Here $o_M(1)$ denotes a term of $o(1)$ as $n \to \infty$ with respect to $\mathbb{D}_{s,p}$-norms of any order.

We wrote b_t for $b(X_t)$ and σ_t for $\sigma(X_t)$. The Itô decomposition of $\sigma_t = \sigma(X_t)$ is denoted by

$$\sigma_t = \sigma_0 + \int_0^t \sigma_s^{[1]} dw_s + \int_0^t \sigma_s^{[0]} ds.$$

Though $\sigma_s^{[1]}$ and $\sigma_s^{[2]}$ have a simple expression with b, σ and X_s, those symbols are convenient to simplify the notation. This rule will be applied for other functionals.

Reference variable

For a reference variable, we will consider

$$F_n = \frac{1}{n}\sum_{j=1}^{n} \beta(X_{t_{j-1}}) \ \ \textbf{or} \ \ F_n = F_\infty := \int_0^1 \beta(X_t)dt.$$

Nondegeneracy*

Let $a(x) = c(x)\sigma(x)^2$**. Let**

$$V_0(x_1, x_2) = \begin{bmatrix} b(x_1) - \frac{1}{2}\sigma(x_1)\partial_{x_1}\sigma(x_1) \\ \beta(x_1) \end{bmatrix} \quad \text{and} \quad V_1(x_1, x_2) = \begin{bmatrix} \sigma(x_1) \\ 0 \end{bmatrix}$$

for $x_1 \in \mathbb{R}$ **and** $x_2 \in \mathbb{R}^{d_1}$**. The Lie algebra generated by**

$$V_1,\ [V_i, V_j]\ (i, j = 0, 1),\ [V_i, [V_j, V_k]]\ (i, j, k = 0, 1), \ldots.$$

at (x_1, x_2) **is denoted by Lie**$[V_0; V_1](x_1, x_2)$**.**

Assume that supp(X_0) **is compact. Moreover, for nondegeneracy, we assume**

[H1] $\inf_{x\in\mathbb{R}} |a(x)| > 0.$

[H2] **Lie**$[V_0; V_1](X_0, 0) = \mathbb{R}^{1+d_1}$ **a.s.**

2nd order specification*

$$
\begin{aligned}
(M_\infty, \mathring{C}_\infty, N_\infty) =^d \Big(& \int_0^1 \sqrt{2}a(X_s)dB_s, \\
& \int_0^1 \frac{4\sqrt{2}}{3}a(X_s)^2 dB_s + \int_0^1 \frac{4}{3}a(X_s)^2 dB'_s, \\
& \int_0^1 q_s dB''_s + \int_0^1 h_s ds \Big),
\end{aligned}
$$

where (B, B', B'') is a three-dimensional standard Wiener process, independent of $\mathcal{F}$, defined on the extension $\bar{\Omega}$, and

$$
h_t = c_t b_t^2 + c_t b_t^{[1]} \sigma_t - \frac{1}{2} c_t^{[0]} \sigma_t^2 - c_t^{[1]} \sigma_t \sigma_t^{[1]}.
$$

Adaptive random symbol

The adaptive random symbol $\underline{\sigma}(z, iu, iv)$ **is given by**

$$\begin{aligned}\underline{\sigma}(z, iu, iv) = & \frac{2z}{3}\int_0^1 a(X_s)^3 ds\Big(\int_0^1 a(X_s)^2 ds\Big)^{-1}(iu)^2 \\ & +iu\int_0^1 h_t dt.\end{aligned}$$

Anticipative random symbol (1)

The random symbol $\sigma_{s,r}(iu,iv)$ admits the expression

$$
\begin{aligned}
&\sigma_{s,r}(iu,iv)\\
= \; & u^2\int_r^s \alpha'(X_t)D_rX_tdt\Big(-u^2\int_s^1 \alpha'(X_t)D_sX_tdt+i\int_s^1 \beta'(X_t)[v]D_sX_tdt\Big)\\
&+\Big(-u^2\int_r^1 \alpha'(X_t)D_rX_tdt+i\int_r^1 \beta'(X_t)[v]D_rX_tdt\Big)\\
&\cdot\Big(-u^2\int_s^1 \alpha'(X_t)D_sX_tdt+i\int_s^1 \beta'(X_t)[v]D_sX_tdt\Big)\\
&+\Big(-u^2\int_s^1\{\alpha''(X_t)D_rX_tD_sX_t+\alpha'(X_t)D_rD_sX_t\}dt\\
&+i\int_s^1\{\beta''(X_t)[v]D_rX_tD_sX_t+\beta'(X_t)[v]D_rD_sX_t\}dt\Big)
\end{aligned}
$$

for $r\le s$, where the prime $'$ stands for the derivative in $x_1\in\mathbb{R}$.

The processes $D_s X_t$ and $D_r D_s X_t$ are determined according to routine; for example, $D_s X_t$ satisfies the equation

$$D_s X_t = \sigma(X_s) + \int_s^t \beta'(X_t) D_s X_t dt + \int_s^t \sigma'(X_t) D_s X_t dw_t$$

for $s \leq t$. $D_r D_s X_t$ admits a similar equation.

Anticipative random symbol (2)

Now we obtain the anticipative random symbol

$$\bar{\sigma}(iu, iv) = \int_0^1 iu\, a(X_s)\sigma_{s,s}(iu, iv)\, ds$$

with

$$\begin{aligned}\sigma_{s,s}(iu, iv) = & \Big(- u^2 \int_s^1 \alpha'(X_t) D_s X_t dt + i \int_s^1 \beta'(X_t)[v] D_s X_t dt \Big)^2 \\ & -u^2 \int_s^1 \{\alpha''(X_t)(D_s X_t)^2 + \alpha'(X_t) D_s D_s X_t\} dt \\ & +i \int_s^1 \{\beta''(X_t)[v](D_s X_t)^2 + \beta'(X_t)[v] D_s D_s X_t\} dt\end{aligned}$$

Formula: Martingale expansion in mixed normal limits (Y 2008) *

Set

$$\sigma = \underline{\sigma} + \bar{\sigma}, \tag{10}$$

$$\begin{aligned} p_n(z,x) &= E\left[\phi(z;W_\infty,C_\infty)\middle|F_\infty = x\right]p^{F_\infty}(x) \\ &\quad + r_n E\left[\sigma(z,\partial_z,\partial_x)^*\left\{\phi(z;W_\infty,C_\infty)\middle|F_\infty = x\right]p^{F_\infty}(x)\right\}. \end{aligned}$$

With Watanabe's delta functional, we can write

$$\begin{aligned} p_n(z,x) &= E\left[\phi(z;W_\infty,C_\infty)\delta_x(F_\infty)\right] \\ &\quad + r_n E\left[\sigma(z,\partial_z,\partial_x)^*\left\{\phi(z;W_\infty,C_\infty)\delta_x(F_\infty)\right\}\right]. \end{aligned}$$

$*$ **Available as a preprint.**

Asymptotic expansion: the quadratic form

Theorem 8. Suppose that [H1] **and** [H2] **are satisfied. Then for any positive numbers** M **and** γ**,**

$$\sup_{f \in \mathcal{E}(M,\gamma)} \left| E\big[f(Z_n, F_n)\big] - \int_{\mathbb{R}^{1+d_1}} f(z,x) p_n(z,x) dz dx \right| = o\left(\frac{1}{\sqrt{n}}\right)$$

as $n \to \infty$**, where** $\mathcal{E}(M,\gamma)$ **is the set of measurable functions** $f : \mathbb{R}^{1+d_1} \to \mathbb{R}$ **satisfying** $|f(z,x)| \leq M(1 + |z| + |x|)^\gamma$ **for all** $(z,x) \in \mathbb{R} \times \mathbb{R}^{d_1}$**.**

Comment: Martingale expansion in mixed normal limits

It is possible to give the asymptotic expansion of the conditional law $\mathcal{L}\{Z_n|F_n\}$ in the same framework we applied for the expansion of the joint law $\mathcal{L}\{(Z_n, F_n)\}$.