SAPS IX, Le Mans, March 10, 2013

Nondegeneracy of statistical random field and statistics for stochastic processes

Nakahiro Yoshida (University of Tokyo)

In this talk

- **Nondegeneracy of statistical random field**
- **Applications to statistical analysis for volatility under high frequent sampling and finite time-horizon.**
  - **construction of quasi likelihood analysis**
  - **a criterion for nondegeneracy of statistical random field**
  - **second order statistical analysis for volatility**

Volatility estimation: QLA and nondegeneracy of statistical random field,

Stochastic regression model

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

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

- 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** $\mathbb{R}^m$ **and** $\mathbb{R}^{\mathsf{d}}$**, respectively.** b **is unobservable, completely unknown.**
- σ: **an** $\mathbb{R}^m \otimes \mathbb{R}^r$**-valued function defined on** $\mathbb{R}^{\mathsf{d}} \times \Theta$,
- Θ: **a bounded domain in** $\mathbb{R}^p$
- θ^* **denotes the true value of** θ.

- **Data:** $\mathbf{Z}_n = (X_{t_j}, Y_{t_j})_{0 \le j \le n}$ **with** $t_j = jh$ **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.**

Quasi likelihood

- **Quasi log likelihood function:**

$$\begin{aligned} \mathbb{H}_n(\theta) &= -\frac{nm}{2}\log(2\pi h) - \frac{1}{2}\sum_{j=1}^{n}\left\{ \log\det S(X_{t_{j-1}},\theta) \right. \\ &\qquad \left. + h^{-1}S^{-1}(X_{t_{j-1}},\theta)[(\Delta_j Y)^{\otimes 2}] \right\}, \end{aligned}$$

$$\begin{aligned} S &= \sigma^{\otimes 2} = \sigma\sigma', \\ \Delta_j Y &= Y_{t_j} - Y_{t_{j-1}}. \end{aligned}$$

QLA estimators

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

$$\mathbb{H}_n(\hat{\theta}_n) = \sup_{\theta\in\Theta} \mathbb{H}_n(\theta). \tag{2}$$

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

$$\tilde{\theta}_n = \left(\int_\Theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta\right)^{-1} \int_\Theta \theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta. \tag{3}$$

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

Quasi likelihood analysis (QLA)

Definition 1. QLA $\ni$

- (quasi) likelihood random field
- quasi MLE
- quasi Bayesian estimator
- (polynomial type) large deviation estimates for the quasi likelihood random field
- tail probability estimate for the QLA estimators, convergence of moments.

Convergence of moments of estimator is very important

- $P[|\sqrt{n}(\hat{\theta}-\theta)|>r] \lesssim \frac{1}{r^L} \Leftrightarrow E[|\sqrt{n}(\hat{\theta}-\theta)|^p]<\infty$
- **Decision theory: risk function** $E_\theta[W(\sqrt{n}(\hat{\theta}-\theta))]$.
  - **Attainability of the risk bound by the estimator of interest.**
- **Prediction and Information criteria:**
  - **Expanding E[log-likelihood**$(\hat{\theta}_n)$**].**
- **Asymptotic expansion of** $E[f(\sqrt{n}(\hat{\theta}_n-\theta))]$ **for measurable functions** f.
  - **Rigorous estimate of the higher-order terms in the stochastic expansion.**
- **Machinery is necessary:**
  "(at least polynomial type) large deviation inequality"

Large deviation of likelihood ratio random fields and likelihood analysis

- **The framework of the likelihood analysis was established by Ibragimov and Hasminskii (TPA1972, TPA1973, Springer1981)**
- **Likelihood analysis of semimartingales by Kutoyants (Heldermann1984, Springer1998, Springer2004).**

Recall the model

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

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

- **Data:** $\mathbf{Z}_n = (X_{t_j}, Y_{t_j})_{0 \le j \le n}$ **with** $t_j = jh$ **for** $h = h_n = T/n$**.**
- **This is non-ergodic statistics.**
- **Polynomial type large deviation holds in a general setting (e.g. LAQ structure), and consequently, the QLA is obtained (Y ISMRM2006, AISM2011). We apply this QLA to the volatility estimation problem.**

A Key index for separation of distributions (1)

- $\mathbb{Y}_n(\theta) = \frac{1}{n}\big(\mathbb{H}_n(\theta) - \mathbb{H}_n(\theta^*)\big)$ **: quasi log likelihood ratio**
- $\mathbb{Y}$ **: the limit random field of** $\mathbb{Y}_n$**, more precisely,**

$$\sup_{n\in\mathbb{N}} \left\| \sup_{\theta\in\Theta} \sqrt{n}|\mathbb{Y}_n(\theta) - \mathbb{Y}(\theta)| \right\|_q < \infty$$

for every $q > 1$**. Next slide for** $\mathbb{Y}(\theta)$**.**

[NB Quasi log likelihood function:

$$\mathbb{H}_n(\theta) = -\frac{nm}{2}\log(2\pi h) - \frac{1}{2}\sum_{j=1}^{n}\Bigg\{\log\det S(X_{t_{j-1}}, \theta)$$

$$+h^{-1}S^{-1}(X_{t_{j-1}}, \theta)[(\Delta_j Y)^{\otimes 2}]\Bigg\},$$

$$S = \sigma^{\otimes 2} = \sigma\sigma', \quad \Delta_j Y = Y_{t_j} - Y_{t_{j-1}}.]$$

A Key index for separation of distributions (2)

Key index:

$$\chi_0 = \inf_{\theta \neq \theta^*} \frac{-\mathbb{Y}(\theta)}{|\theta - \theta^*|^2}.$$

For the stochastic regression model,

$$\mathbb{Y}(\theta) = -\frac{1}{2T}\int_0^T \left\{ \log\left(\frac{\det S(X_t,\theta)}{\det S(X_t,\theta^*)}\right) + \mathbf{tr}\left(S^{-1}(X_t,\theta)S(X_t,\theta^*) - I_d\right)\right\} dt.$$

[H2] For every $L > 0$, there exists $c_L > 0$ such that

$$P\left[\chi_0 \leq r^{-1}\right] \leq \frac{c_L}{r^L}$$

for all $r > 0$.

- **Remark. If is possible to reformulate the problem by weakening** $[H2]$ **for a specific** L **(cf. Y AISM2011). Naturally, the resulting order of moments will be limited, not an arbitrary positive number, under such a condition.**

QLA for volatility under $[H2]$:

Under $[H2]$, we can obtain the asymptotic mixed normality and moment convergence of the quasi ML and quasi Bayesian estimators, e.g.,

Theorem 1. (Uchida and Y) Under $[H2]$,

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

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

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

as $n \to \infty$.

Here $\Gamma(\theta^*)$ is a random matrix and ζ is a standard Gaussian random vector independent of $\Gamma(\theta^*)$.

- **Non-ergodic statistics!**
  The identifiability of the statistical model is random!

Proof (1) Limit theorem

- **Mixed normal limit theorem (stable convergence):**

$$\sum_{j=1}^{n} A(t_{j-1})[(\Delta_j Y)^{\otimes 2} - B(t_{j-1})] \to^{d_s(\mathcal{F}_T)} \textbf{mixed normal}$$

after scaling.

Proof (2) Quasi likelihood ratio random field

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

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

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

Proof (3) [H2] implies PLD

- **[H2]+LAQ** $\Rightarrow^{Y}$ **Polynomial type large deviation (PLD) inequality**

$$P\left[\sup_{u\in\mathbb{R}^p:|u|\geq r} \mathbb{Z}_n(u) \geq e^{-r}\right] \leq \frac{C_L}{r^L} \quad (r>0)$$

- $\mathbb{H}_n \Rightarrow^{J} [\mathbb{Z}_n \to^{d_f} \mathbb{Z}] \Rightarrow [\mathbb{Z}_n \to^{d} \mathbb{Z}]_{compact}$
- **PLD** $+ [\mathbb{Z}_n \to^{d} \mathbb{Z}]_{compact} \Rightarrow^{IHK}$ **QLA** $\square$

- **References:**
  - **Y (ISM ResearchMemorandum2006, AISM2011)**
  - **Uchida and Y (ISM ResearchMemorandum2011, arXiv2012)**

Then, what ensures [H2]?

- **? $\Rightarrow$ [H2]**
- **Analytic criterion: Nondegeneracy of a certain tensor field over supp $\mathcal{L}\{X_0\} \times \Theta$ ensures [H2]. (U-Y, SAPS 2009)**
- **Geometric criterion (U-Y SAPS 2013, arXiv2012)**

NB. $\chi_0 = \inf_{\theta \neq \theta^*} \dfrac{-\mathbb{Y}(\theta)}{|\theta - \theta^*|^2}$.

[H2] $P\left[\chi_0 \leq r^{-1}\right] \leq \dfrac{c_L}{r^L}$ **for all** $r > 0$.

Exercise.

- **A sufficient condition for [H2] is that**

$$\inf_{\substack{\omega\in\Omega,\theta\in\Theta\setminus\{\theta_*\}\\ t\in[0,T]}} \left\{ \log\left(\frac{\det S(X_t,\theta)}{\det S(X_t,\theta^*)}\right) \right.$$
$$\left. +\mathbf{tr}\left(S^{-1}(X_t,\theta)S(X_t,\theta^*)-I_d\right)\right\} / |\theta-\theta^*|^2 > 0 \ \ a.s.$$

- **It is too naïve as it breaks, for example, in a simple model**

$$dX_t = (1+X_t^2)^\theta dw_t, \ \ X_0 = 0.$$

Support function

Let

$$Q(x,\theta,\theta^*) = \mathbf{tr}\left(S(x,\theta)^{-1}S(x,\theta^*) - I_d\right) - \log\det\left(S(x,\theta)^{-1}S(x,\theta^*)\right)$$

then

$$-2\mathbb{Y}(\theta) = \frac{1}{T}\int_0^T Q(X_t,\theta,\theta^*)dt.$$

A support function f is a function such that

$$Q(x,\theta,\theta^*)|\theta-\theta^*|^{-2} \geq |f(x,\theta)|^{\varrho},$$

for a constant $\varrho \in (0,\infty)$. Recall

$$\chi_0 = \inf_{\theta\neq\theta^*}\frac{-\mathbb{Y}(\theta)}{|\theta-\theta^*|^2} \geq \inf_{\theta\neq\theta^*}\frac{1}{2T}\int_0^T |f(X_t,\theta)|^{\varrho}dt.$$

Analytic criterion: nondegeneracy of a tensor field

- **For simplicity, let** $\mathsf{d} = 1$ **and suppose that** X **is a nondegenerate Itô process.**
- **Suppose that** $\mathcal{X}_0$ **is a neighborhood of compact supp**$\mathcal{L}\{X_0\}$**, and that** Θ **is compact.**
- **For each** $(x_0, \theta) \in \mathcal{X}_0 \times \Theta$**,** $\max_{j=0,...,J-1} \left|\partial_x^j f(x_0, \theta)\right| > 0$**.**

Then [H2] holds.

- **Remarks.**
  - **Similar condition in the multi-dimensional case.**
  - **It is possible to give a condition for a degenerate diffusion on manifold. However the condition becomes much more complicated. (Uchida and Y LeMans2009, ISM RM2011, Paris2012)**

Geometric criterion

- **Example.**
  - $f(x,\theta) = x_1 x_2 (x_1 - \theta_1 x_2^2)(\theta_2 x_1 + x_2^2)$
  - $X = (X_t) = (X_{1,t}, X_{2,t})$: **a nondegenerate diffusion with uniform initial distribution on** $\mathrm{supp}\mathcal{L}\{X_0\} = \{0\} \times [0,1]$.
  - **Show**

$$P\left[\inf_{\theta} \int_0^1 |f(X_t,\theta)|^2 dt < \frac{1}{r}\right] \leq \frac{C_L}{r^L}.$$

  - **The null set** $\{f = 0\}$ **is not a regular submanifold.**

Geometric criterion

[A3′] supp$\mathcal{L}\{X_0\}$ **is compact, there exists a function** $f : U \times \Theta \to \mathbb{R}$ **for some open neighborhood** U **of supp**$\mathcal{L}\{X_0\}$ **and the following conditions are satisfied.**

(i) For some $\varrho \in (0, \infty)$ **,** $Q(x, \theta, \theta^*)|\theta - \theta^*|^{-2} \geq |f(x, \theta)|^\varrho$ **for all** $(x, \theta) \in U \times (\Theta \setminus \{\theta^*\})$**.**

(ii) For each $x_0 \in U$**, there exist a neighborhood** V **in** U **of** x_0 **and a covering** $\{\Theta_k\}_{k=1,...,\bar{k}}$ **of** Θ **such that for each** $k = 1, ..., \bar{k}$**, there exist** $\xi_0 \in \mathbb{S}$**,** $J \in \mathbb{N}$**, some positive numbers** M, c, ϵ_0**,** K_j **(**$j = 1, ..., J$**) and some functions** $\Psi_j : P_{\xi_0}^{\perp} V \times \Theta_k \to \mathbb{R}$ **such that**

(a) each function $P_{\xi_0}^{\perp} V \ni y \mapsto \Psi_j(y, \theta) \in \mathbb{R}$ **is** M**-Lipschitz continuous for all** $\theta \in \Theta_k$**,**

(b) for $(x,\theta) \in V \times \Theta_k$**,**

$$|f(x,\theta)| \geq c \prod_{j=1}^{J} \left(|\xi_0 \cdot x - \Psi_j(P_{\xi_0}^{\perp} x, \theta)| \wedge \epsilon_0 \right)^{K_j}.$$

Remarks

- **In** $[A3']$**,** $\bar{k}$ **may depend on** x_0**.**
- **Note that** $\{x \in V;\ f(x,\theta) = 0\} \subset \bigcup_{j=1}^{J} \{x \in V;\ \xi_0 \cdot x = \Psi_j(P_{\xi_0}^{\perp} x, \theta)\}$ **under** $[A3']$**(ii), that is, the graph of the functions** Ψ_j **covers locally the null set of** f**.**

Theorem 2. (Uchida-Y 2012 arXiv) $[A3']$ + **nondegenerate Itô process** $X \Rightarrow [H2]$**, and hence QLA.**

Summary

- **We constructed QLA for the stochastic regression model**

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

- **Tail probability estimates for QLA estimators:**

$$\textbf{(QMLE)} \quad \sup_n P[\sqrt{n}|\hat{\theta}_n - \theta| > r] \lesssim \frac{1}{r^L},$$

$$\textbf{(QBE)} \quad \sup_n P[\sqrt{n}|\tilde{\theta}_n - \theta| > r] \lesssim \frac{1}{r^L}.$$

- **The geometric criterion gives the nondegeneracy of the key index χ_0.**

Some remarks on QLA for ergodic diffusion

- **QLA for ergodic diffusion: Y (ISM RM2006, AISM2011)**
  - **QMLE, adaptive Bayes estimator, PLD, moment convergence**
- **QLA and adaptive procedures: Uchida and Y (LeMans2011, SPA2012)**
  - **reduced sampling rate from** $nh^2 \to 0$ **to** $nh^p \to 0$
  - **PLD, moment convergence**
- **Information criterion CIC: Uchida (AISM2010)**
  - **PLD, Malliavin calculus**
- **Misspecification affects convergence rate of the volatility parameter: Uchida and Y (ESAIM2011)**
  - **The convergence rate of the volatility estimator becomes** $\sqrt{T}$**, not** $\sqrt{n}$**.**

Some remarks on QLA for ergodic jump-diffusion

- **Ergodic jump-diffusion: Shimizu and Y (SISP2006)**
- **QLA for ergodic jump-diffusion: Ogihara and Y (Tokyo2009, SISP2011)**
- **QLA for ergodic Lévy driven SDE: Masuda (2012MIPreprintSeries)**

More on QL methods

- **QMLE of volatility in nonsynchronous sampling scheme: Ogihara and Y (Paris2012)**
- **Change point problem for volatility: Iacus and Y (SPA2012) Convergence of QL random field to a mixture of double sided Wiener functionals.**
- **Distcriminant analysis: Uchida and Y (Hokkaido2012, SAPS2013)**

Asymptotic expansion of the QMLE as an application of QLA

Asymptotic expansion of the QMLE

- **SDE**

$$dX_t = b(X_t)dt + \sigma(X_t, \theta)dw_t, \quad t \in [0, 1]$$

- **Aim: asymptotic expansion of the distribution of the QMLE** $\hat{\theta}_n$
- **We need**
  - **the polynomial type large deviation inequality for** $\sqrt{n}(\hat{\theta}_n - \theta^*)$
  - **the martingale expansion under mixed normal limit.**

Stochastic expansion of the QMLE

- **Expand** $\partial_\theta \mathbb{H}_n(\hat{\theta}_n) = 0$ **around** θ^***.**

$$\begin{aligned} &\sqrt{n}(\hat{\theta}_n - \theta^*) \\ = &\sqrt{n}\ddot{\mathbb{H}}_n(\theta^*)^{-1}\Big\{ -\dot{\mathbb{H}}_n(\theta^*) - \frac{1}{2}\dddot{\mathbb{H}}_n(\theta^*)(\hat{\theta}_n - \theta^*)^2 \Big\} \\ &+ O_{L^{\infty-}}(n^{-1/2-\epsilon}) \end{aligned}$$

- **The higher-oder term is estimated by the polynomial type large deviation inequality for the QMLE**

- **Thus, the QMLE admits a (very involved) stochastic expansion.**

$$
\begin{aligned}
Z_n &= M_n + \frac{1}{\sqrt{n}} N_n \\
M_n &= \sqrt{n} \sum_j 2^{-1} S^{-1} \dot{S}(X_{t_{j-1}}, \theta^*) \{(\Delta_j w)^2 - n^{-1}\} \\
N_n &= \cdots
\end{aligned}
$$

- **Randomly scaled perturbed martingale:**

$$
\sqrt{n}(\hat{\theta}_n - \theta^*) = \Gamma(\theta^*)^{-1} Z_n
$$

Formula: Martingale expansion in mixed normal limit

- **Perturbed martingale**

$$Z_n = M_1^n + W_n + r_n N_n, \quad r_n \to 0$$

- F_n**: a reference variable, e.g.,**

$$F_n = F_\infty = \int_0^1 \beta(X_t)dt.$$

Often $W_n = 0$

- $C_n := \langle M^n \rangle_1 \to^p C_\infty$
- $M_1^n \to^{d_s}$ **Mixed Normal**$(0, C_\infty)$
- **Random symbols (symbols with random coefficients)**

$$\varsigma(z, iu, iv) = \sum_j c_j(z, \omega)(iu)^{m_j}(iv)^{n_j}$$

- $\underline{\sigma}$: **Adaptive random symbol**
- $\overline{\sigma}$: **Anticipative random symbol**
- **Full random symbol** $\sigma = \underline{\sigma} + \bar{\sigma}$
- **The random symbol** $\overline{\sigma}$ **is defined by the Malliavin derivatives.**

  More precisely,
- **Adaptive random symbol:**

$$\underline{\sigma}(z, iu, iv) = \frac{1}{2}\tilde{C}_\infty(z)^{j,k}(iu_j)(iu_k) + \tilde{W}_\infty(z)^j(iu_j) + \tilde{N}_\infty(z)^j(iu_j) + \tilde{F}_\infty(z)^l(iv_j) \tag{4}$$

for $u \in \mathbb{R}^d$ **and** $v \in \mathbb{R}^{d_1}$**.**
For details, see Y 2012arXiv.

- **Anticipative random symbol (in the case of a quadratic form of the principal part):**

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

with

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

for some functions a and β, $\alpha = a^2$.

- **Asymptotic expansion formula:**

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

with Watanabe's delta functional.

  - **We can see mixed normality in the formula.**

Theorem 3. **Under suitable non-degeneracy condition, the density** $p_n(z,x)$ **gives a 2nd order approximation to** $\mathcal{L}\{(Z_n,F_n)\}$**, i.e., the approximation error is** $o(r_n)$**.**

(Martingale expansion, ISM RM2010, arXiv2012)

Asymptotic expansion of the QMLE

Martingale expansion can apply to QMLE to obtain asymptotic expansion.

- $p_n(z,x)$: **asy. exp. for** (Z_n, F_n),
  F_n **is a reference variable, for example,** $F_n = \Gamma(\theta^*)$.
- $\mathbb{T}_n(z,x) = x^{-1}z$
- **Asymptotic expansion of the QMLE**

$$q_n = (\mathbb{T}_n)_* p_n$$

- **Summary. PLD plays an essential role in derivation of the asymptotic expansion of the quasi maximum likelihood estimator.**