YUIMA SUMMER SCHOOL Brixen (June 26)

Lecture 07
QLA – Quasi-likelihood analysis
+
Lecture 08
Bayesian analysis

Nakahiro Yoshida

Graduate School of Mathematical Sciences, University of Tokyo
Japan Science and Technology Agency CREST
Institute of Statistical Mathematics

Tokyo April 2019

MENU

- **QMLE and QBE**
- **Quick introduction to the QLA theory**
- **Applications of the Quasi-Likelihood Analysis**
- **Slightly deeper discussion (omitted)**

MENU

1. **QMLE and QBE**
2. **Quick introduction to the QLA theory**
3. **Applications of the Quasi-Likelihood Analysis**
4. **Slightly deeper discussion (omitted)**

Quasi likelihood analysis (QLA)

- Θ: a (bounded) open set in $\mathbb{R}^{\mathsf{p}}$, the parameter space
- $T \in \mathbb{T}$ ($\mathbb{T} = \mathbb{Z}_+ = \{0, 1, ...\}, \mathbb{R}_+ = [0, \infty), ...$)
- $\mathbb{H}_T : \Omega \times \overline{\Theta} \to \mathbb{R}$: a random field (quasi-log likelihood function)
- Example. $\{P_{n,\theta} = N(\theta, 1)^n\}_{\theta \in \Theta}$, $T = n \in \mathbb{T} = \mathbb{N}$,

$$\mathbb{H}_n(\theta) = \sum_{j=1}^{n} \log \phi(x_j; \theta, 1) \quad \text{(log likelihood function)}$$

QLA estimators

- $\hat{\theta}_T^M$: the quasi-maximum likelihood estimator (QMLE) defined as

$$\mathbb{H}_T(\hat{\theta}_T^M) = \max_{\theta \in \overline{\Theta}} \mathbb{H}_T(\theta). \tag{1}$$

- $\hat{\theta}_T^B$: the quasi-Bayes estimator (QBE) for a prior density $\pi : \Theta \to \mathbb{R}_+$ defined by

$$\hat{\theta}_T^B = \left(\int_\Theta \exp(\mathbb{H}_T(\theta))\pi(\theta)d\theta \right)^{-1} \int_\Theta \theta \exp(\mathbb{H}_T(\theta))\pi(\theta)d\theta. \tag{2}$$

Assume that π is continuous and $0 < \inf_{\theta \in \Theta} \pi(\theta) \leq \sup_{\theta \in \Theta} \pi(\theta) < \infty$.
Here $\pi(d\theta)$ is a strategy or tuning parameter to estimate the fixed true value θ^*.

Summary of Lectures 07 and 08

- **By the QLA theory, for the QLA estimators (QMLE, QBE), we can prove**
  - **consistency and asymptotic (mixed) normality**
  - **asymptotic optimality**
  - **convergence of the moments of the error**
- **QLA can apply to stochastic processes such as diffusion processes and point processes.**
- **Without deep knowledge of the QLA theory, through YUIMA, we can use many cutting-edge results in statistics of stochastic processes.**

MENU

1. QMLE and QBE
2. **Quick introduction to the QLA theory**
3. Applications of the Quasi-Likelihood Analysis
4. Slightly deeper discussion (omitted)

Quasi likelihood analysis (QLA)

- Example. $N(\theta, 1)$. The log likelihood ratio

$$\mathbb{H}_n(\theta) - \mathbb{H}_n(\theta^*) = \sum_{j=1}^{n} \log \phi(x_j; \theta, 1)/\phi(x_j; \theta^*, 1)$$
$$= (\theta - \theta^*) \sum_{j=1}^{n} x_j - \frac{n}{2}(\theta^2 - \theta^{*\,2})$$

$$\mathbb{H}_n(\theta^* + n^{-1/2}u) - \mathbb{H}_n(\theta^*) = n^{-1/2} \sum_{j=1}^{n} \epsilon_j u - \frac{1}{2}u^2,$$
$$\epsilon_j = x_j - \theta^* \sim N(0, 1)$$

- a quadratic form of the normalized parameter u
- This phenomenon occurs asymptotically in quite many cases if the model is differentiable.

Quasi likelihood analysis (QLA)

- Θ: a bounded open set in $\mathbb{R}^{\mathsf{p}}$, the parameter space
- $T \in \mathbb{T}$ ($\mathbb{T} = \mathbb{Z}_+, \mathbb{R}_+, \ldots$)
- $\mathbb{H}_T : \Omega \times \Theta \to \mathbb{R}$: a random field
- $a_T \in GL(\mathbb{R}^{\mathsf{p}})$, $a_T \to 0$ ($T \to \infty$)
- $\mathbb{U}_T = \{u;\ \theta^* + a_T u \in \Theta\}$
- Quasi-likelihood ratio process

$$\mathbb{Z}_T(u) = \exp\left\{\mathbb{H}_T(\theta^* + a_T u) - \mathbb{H}_T(\theta^*)\right\}$$

Locally asymptotically quadratic $\mathbb{Z}_T$

- $\mathbb{Z}_T$ is Locally Asymptotically Quadratic (LAQ)

$$\mathbb{Z}_T(u) = \exp\left(\Delta_T[u] - \frac{1}{2}\Gamma[u^{\otimes 2}] + r_T(u)\right)$$

- Δ_T: a random vector (linear form)
- Γ: a deterministic or random bilinear form
- $r_T(u) \to^p 0$ as $T \to \infty$
- Notation.
  $v[u] = \sum_i v_i u^i$,
  $M[u^{\otimes 2}] = M[u, u] = \sum_{i,j} M_{i,j} u^i u^j$
  for $v = (v_i)$, $M = (M_{i,j})$ and $u = (u^i)$.

Example: Likelihood Analysis

- $P_\theta << \nu$, ν: a reference measure (e.g. dx on $\mathbb{R}$)
- $p_\theta(x) = \frac{dP_\theta}{d\nu}(x)$
- likelihood function $\Theta \ni \theta \mapsto L_n(\theta) = \prod_{j=1}^{n} p_\theta(x_j)$
- maximum likelihood estimator $\hat{\theta}_n^M : \mathcal{X}^n \to \Theta$

$$L_n(\hat{\theta}_n^M) = \max_{\theta \in \Theta} L_n(\theta)$$

- Let $\mathbb{H}_n(\theta) = \log L_n(\theta)$. $T = n$.
- Then, for $a_T = n^{-1/2}$,

$$\begin{aligned}\log \mathbb{Z}_n(u) &= \mathbb{H}_n(\theta^* + a_n u) - \mathbb{H}_n(\theta^*) \\ &= \sum_{j=1}^{n} \log \frac{p(x_j, \theta^* + n^{-1/2} u)}{p(x_j, \theta^*)}\end{aligned}$$

Example: Likelihood Analysis

Roughly

$$
\begin{aligned}
&\log \mathbb{Z}_n(u) \\
&= \sum_{j=1}^{n} \log\left(1 + n^{-1/2}\frac{\partial_\theta p_\theta(x_j)}{p_\theta(x_j)}[u] + \frac{1}{2}n^{-1}\frac{\partial_\theta^2 p_\theta(x_j)}{p_\theta(x_j)}[u,u] + \cdots\right) \\
&= \sum_{j=1}^{n}\left\{n^{-1/2}\frac{\partial_\theta p_\theta(x_j)}{p_\theta(x_j)}[u] + \frac{1}{2}n^{-1}\frac{\partial_\theta^2 p_\theta(x_j)}{p_\theta(x_j)}[u,u]\right. \\
&\qquad \left. -\frac{1}{2}n^{-1}\left(\frac{\partial_\theta p_\theta(x_j)}{p_\theta(x_j)}[u]\right)^2 + \cdots\right\} \\
&= n^{-1/2}\sum_{j=1}^{n}\frac{\partial_\theta p_\theta(x_j)}{p_\theta(x_j)}[u] - \frac{1}{2}I(\theta^*)[u,u] + o_p(1),
\end{aligned}
$$

where

$$
I(\theta^*) = E_{\theta^*}\left[\left(\frac{\partial_\theta p_\theta}{p_\theta}(\theta^*)\right)^{\otimes 2}\right] \quad \text{(Fisher information matrix)}
$$

Example: local asymptotic normality (LAN, Le Cam)

$$\begin{aligned} \mathbb{Z}_n(u) &= \exp\left(\Delta_n[u] - \frac{1}{2} I(\theta^*)[u,u] + o_p(1)\right), \\ \Delta_n &\to^d \Delta \sim N_p(0, I(\theta^*)) \end{aligned} \tag{3}$$

as $n = T \to \infty$.

Convergence of the random field and QLA estimators

- LAQ

$$\mathbb{Z}_T(u) = \exp\left(\Delta_T[u] - \frac{1}{2}\Gamma[u^{\otimes 2}] + r_T(u)\right)$$

- Assume $(\Delta_T, \Gamma) \to^d (\Delta, \Gamma)$.
- $\mathbb{Z}(u) = \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u^{\otimes 2}]\right)$
- Then

$$\mathbb{Z}_T \to^{d_f} \mathbb{Z} \quad \text{(finite dimensional convergence)}$$

  More strongly

- Convergence of the random field

$$\mathbb{Z}_T \to^d \mathbb{Z} \quad \text{in } \hat{C} = \{f : \mathbb{R}^{\mathsf{p}} \to \mathbb{R}, \lim_{|u|\to\infty} |f(u)| = 0\}$$

Quasi-maximum likelihood estimator (QMLE)

- For arbitrary $U \subset \mathbb{R}^p$,

$$\sup_{u \in U} \mathbb{Z}_T(u) \to^d \sup_{u \in U} \mathbb{Z}(u)$$

- Therefore, for <u>any</u> sequence of QMLE's,

$$\hat{u}_T^M = \operatorname{argmax} \mathbb{Z}_T \to^d \hat{u} = \operatorname{argmax} \mathbb{Z},$$

that is,

$$\hat{u}_T^M = a_T^{-1}(\hat{\theta}_T^M - \theta^*) \to^d \hat{u} = \Gamma^{-1}\Delta$$

Quasi-Bayesian estimator (QBE)

- By definition,

$$a_T^{-1}(\hat{\theta}_T^B - \theta^*) = \int u\, \mathbb{Z}_T(u)\pi(\theta^* + a_T u)du \Big/ \int \mathbb{Z}_T(u)\pi(\theta^* + a_T u)du$$

- $\left(\int \mathbb{Z}_T(u)du, \int u\, \mathbb{Z}_T(u)du \right) \to^d \left(\int \mathbb{Z}(u)du, \int u\, \mathbb{Z}(u)du \right)$

- Convergence

$$\hat{u}_T^B := a_T^{-1}(\hat{\theta}_T^B - \theta^*) \to^d \int u\, \mathbb{Z}(u)du \Big/ \int \mathbb{Z}(u)du$$
$$= \int u \; \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u,u]\right)du \Big/ \int \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u,u]\right)du$$
$$= \Gamma^{-1}\Delta$$

- Exercise. In the LAN case (3), we obtain

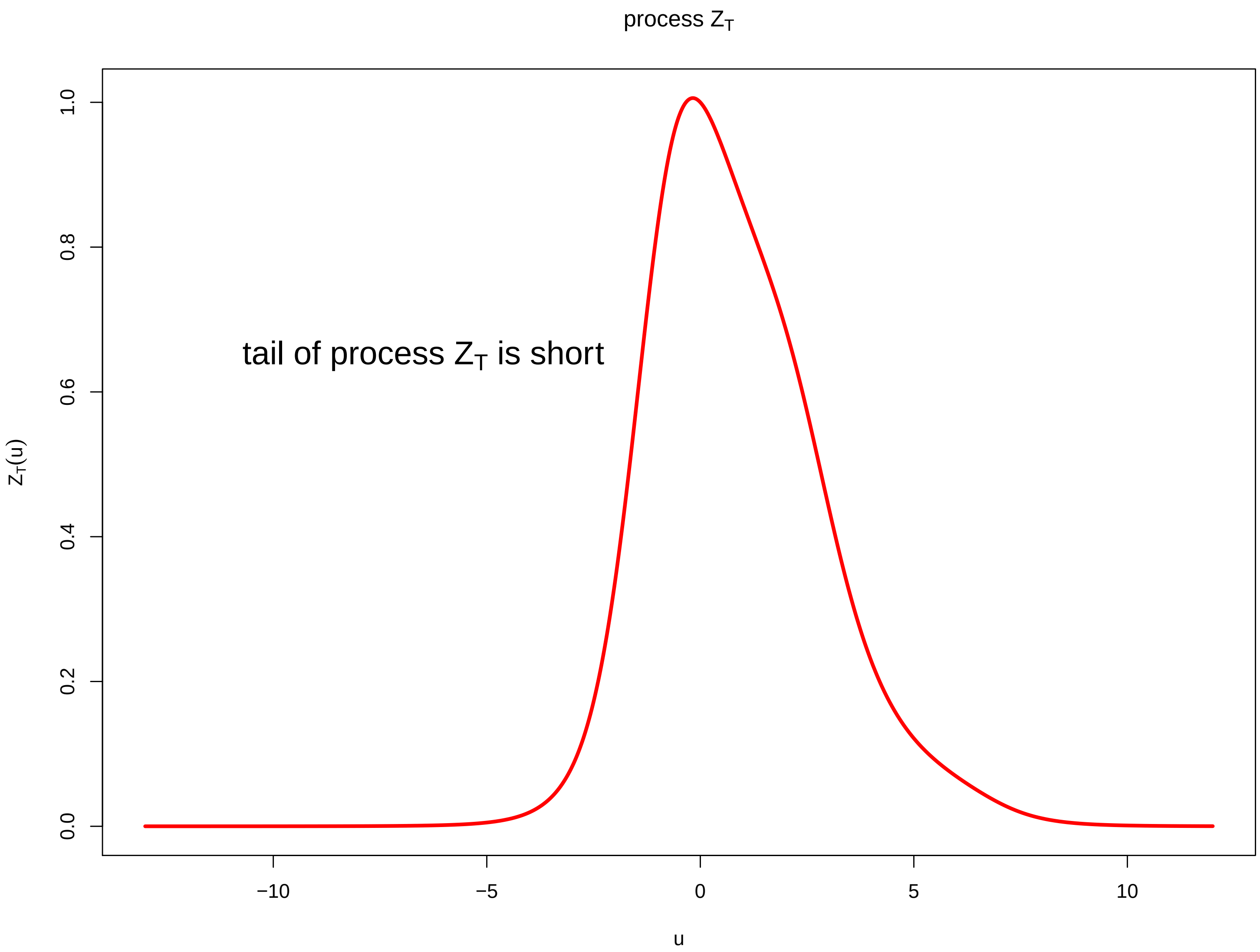
$$\hat{u}_T^{\mathsf{A}} \to^d N_{\mathsf{p}}\left(0, I(\theta^*)^{-1}\right) \quad (\mathsf{A} = M, B)$$

- The random field approach works quite well.

However, a basic question is there

- Is it possible to control $\int u\, \mathbb{Z}_T(u)du$?
  - The region of the integral
$$\mathbb{U}_T = \{u;\ \theta^* + a_T u \in \Theta\} \to \mathbb{R}^{\mathsf{p}}$$
    as $T \to \infty$ even when Θ is bounded.
- Estimate of $\mathbb{Z}_T$ at the tail is essential. (See the plot on the next page.)
  $\Rightarrow$ Large deviation for the random field $\mathbb{Z}_T$
- See the last section.

Random field $\mathbb{Z}_T$



MENU

1. QMLE and QBE
2. Quick introduction to the QLA theory
3. **Applications of the Quasi-Likelihood Analysis**
4. Slightly deeper discussion (omitted)

Let's apply the QLA to stochastic processes

We shall discuss some applications of the QLA theory:

- **ergodic diffusion**
- **non-ergodic diffusion**

Quasi likelihood analysis for ergodic diffusion processes

An ergodic diffusion process

- We consider a stationary diffusion process satisfying the stochastic differential equation

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

- $w = (w_t)_{t \in \mathbb{R}_+}$: an r-dimensional standard Wiener process
- $\theta_i \in \Theta_i \subset \mathbb{R}^{\mathsf{p}_i}$: unknown parameters $(i = 1, 2)$
- $a : \mathbb{R}^{\mathsf{d}} \times \overline{\Theta}_2 \to \mathbb{R}^{\mathsf{d}}$
- $b : \mathbb{R}^{\mathsf{d}} \times \overline{\Theta}_1 \to \mathbb{R}^{\mathsf{d}} \otimes \mathbb{R}^{\mathsf{r}}$
- The true value of $\theta \in \Theta_1 \times \Theta_2$ will be denoted by $\theta^* = (\theta_1^*, \theta_2^*)$.

An ergodic diffusion process

- Assume a mixing property: there exists $a > 0$ such that

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

where

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

- Consequently, we have

$$\frac{1}{T}\int_0^T g(X_t)dt \to^p \int_{\mathbb{R}^{\mathsf{d}}} g(x)\nu(dx) \quad (T\to\infty)$$

for every bounded measurable function g, where $\nu = \nu_{\theta^*}$ is the invariant measure of X.

An ergodic diffusion process

- Data $\mathrm{x}_n = (X_{t_j})_{j=0,1,\ldots,n}$, $t_j = t_j^n = jh$, $h = h_n$
- Estimate $\theta = (\theta_1, \theta_2)$ based on the data $(X_{t_j})_{j=0,1,\ldots,n}$.
- Assume $h \to 0$, $nh \to \infty$ and $nh^2 \to 0$ as $n \to \infty$. That is, long-term high frequency data.
- $B(x, \theta_1) = (bb^\star)(x, \theta_1)$, assumed uniformly non-degenerate
- quasi-likelihood function

$$p_n(\mathrm{x}_n, \theta) = \prod_{j=1}^{n} \frac{1}{(2\pi h)^{\mathrm{d}/2} \det B(X_{t_{j-1}}, \theta_1)^{1/2}} \times \exp\left(-\frac{1}{2h} B(X_{t_{j-1}}, \theta_1)^{-1} \left[\left(\Delta_j X - h a(X_{t_{j-1}}, \theta_2 \right)^{\otimes 2} \right] \right)$$

where $\Delta_j X = X_{t_j} - X_{t_{j-1}}$.

QMLE

- Equivalently, we consider the quasi-log likelihood function

$$\begin{aligned}\mathbb{H}_n(\theta) &= \log\left\{(2\pi h)^{n\mathrm{d}/2} p_n(\mathrm{x}_n, \theta)\right\} \\ &= -\frac{1}{2}\sum_{j=1}^{n}\Big\{h^{-1}B(X_{t_{j-1}}, \theta_1)^{-1}\big[(\Delta_j X - ha(X_{t_{j-1}}, \theta_2)^{\otimes 2}\big] \\ &\qquad + \log\det B(X_{t_{j-1}}, \theta_1)\Big\}.\end{aligned}$$

- The QMLE $\hat{\theta}_n^M = (\hat{\theta}_{1,n}^M, \hat{\theta}_{2,n}^M)$ is any measurable mapping of the data such that

$$\mathbb{H}_n(\hat{\theta}_n^M) = \max_{\theta\in\overline{\Theta}_1\times\overline{\Theta}_2} \mathbb{H}_n(\theta).$$

Information matrices

- Let

$$\boldsymbol{\Gamma_1(\theta^*)[u_1^{\otimes 2}] = \frac{1}{2}\int_{\mathbb{R}^d} \mathrm{Tr}\{B^{-1}(\partial_{\theta_1}B[u_1])B^{-1}(\partial_{\theta_1}B[u_1])(x,\theta_1^*)\}\nu(dx)}$$

for $u_1 \in \mathbb{R}^{p_1}$, and

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

for $u_2 \in \mathbb{R}^{p_2}$.

QMLE

- Applying the QLA theory twice, to θ_1 first and to θ_2 next, we obtain

  Theorem 1. For any sequence of M-estimators for $\theta = (\theta_1, \theta_2)$,

$$E\left[f\left(\sqrt{n}(\hat{\theta}_{1,n}^M - \theta_1^*), \sqrt{nh}(\hat{\theta}_{2,n}^M - \theta_2^*)\right)\right] \to E\left[f(\zeta_1, \zeta_2)\right]$$

  as $n \to \infty$ for $f \in C_p(\mathbb{R}^{\mathsf{p}_1+\mathsf{p}_2})$, where

$$(\zeta_1, \zeta_2) \sim N_{\mathsf{p}_1+\mathsf{p}_2}\left(0, \mathrm{diag}\left[\Gamma_1(\theta^*)^{-1}, \Gamma_2(\theta^*)^{-1}\right]\right).$$

QMLE by YUIMA
(cf. Iacus and Yoshida Springer p.85)

- a SDE model

$$dX_t = (2 - \theta_2 X_t)dt + (1 + X_t^2)^{\theta_1} dw_t, \quad X_0 = 1$$

- Let $\theta_1^* = 0.2$, $\theta_2^* = 0.3$.
- By YUIMA simulate, generate sampled data X_{t_j} with $t_j = jn^{-2/3}$, $n = 750$.
- For the simulated data, apply YUIMA qmle to estimate θ.

```
        Estimate   Std. Error
theta1 0.1969182 0.008095453
theta2 0.2998350 0.126410524
```

Adaptive quasi-Bayesian estimator (adaBayes)

- **The quasi-Bayesian estimator can be defined for $\mathbb{H}_n(\boldsymbol{\theta})$ to estimate parameters simultaneouly.**
- **However, an adaptive method is superior to it from computational point of view. Numerical integration (even with MCMC) becomes easier if the dimension is reduced.**
- **The scheme of the adaptive quasi-Bayesian estimator ("adaBayes") is as follows.**

Adaptive quasi-Bayesian estimator (adaBayes)

- **Step 1.**

$$\hat{\theta}_{1,n}^{aB} = \left[\int_{\Theta_1} \exp\left(\mathbb{H}_n(\theta_1, \theta_2^0)\right) \pi_1(\theta_1) d\theta_1 \right]^{-1} \times \int_{\Theta_1} \theta_1 \exp\left(\mathbb{H}_n(\theta_1, \theta_2^0)\right) \pi_1(\theta_1) d\theta_1$$

where θ_2^0 is any value of θ_2.

- **Step 2.**

$$\hat{\theta}_{2,n}^{aB} = \left[\int_{\Theta_2} \exp\left(\mathbb{H}_n(\hat{\theta}_{1,n}^{aB}, \theta_2)\right) \pi_2(\theta_2) d\theta_2 \right]^{-1} \times \int_{\Theta_2} \theta_2 \exp\left(\mathbb{H}_n(\hat{\theta}_{1,n}^{aB}, \theta_2)\right) \pi_2(\theta_2) d\theta_2.$$

An ergodic diffusion process

- Apply the QLA theory, we obtain

  Theorem 2. For the adaptive Bayesian estimator for $\theta = (\theta_1, \theta_2)$,

  $$E\left[f\left(\sqrt{n}(\hat{\theta}^{aB}_{1,n} - \theta^*_1), \sqrt{nh}(\hat{\theta}^{aB}_{2,n} - \theta^*_2)\right)\right] \to E\big[f(\zeta_1, \zeta_2)\big]$$

  as $n \to \infty$ for $f \in C_p(\mathbb{R}^{\mathsf{p}_1+\mathsf{p}_2})$, where

  $$(\zeta_1, \zeta_2) \sim N_{\mathsf{p}_1+\mathsf{p}_2}\big(0, \mathrm{diag}\big[\Gamma_1(\theta^*)^{-1}, \Gamma_2(\theta^*)^{-1}\big]\big).$$

- Asymptotic properties are the same as QMLE.
- It is a commonly observed fact that the MLE and the BE perform in the same fashion at the first-order asymptotics.
- Reference. Yoshida AISM2011

QBE by YUIMA (cf. Iacus and Yoshida Springer p.87)

- a SDE model

$$dX_t = (2 - \theta_2 X_t)dt + (1 + X_t^2)^{\theta_1} dw_t, \quad X_0 = 1$$

- Let $\theta_1^* = 0.2$, $\theta_2^* = 0.3$.
- By YUIMA simulate, generate sampled data X_{t_j} with $t_j = jn^{-2/3}$, $n = 750$.
- For the simulated data, apply YUIMA adaBayes to estimate θ with a MCMC method.

QBE by YUIMA (cf. Iacus and Yoshida Springer p.87)

```
prior <- list(theta2=list(measure.type="code",
              df="dunif(theta2,0,1)"),
              theta1=list(measure.type="code",
              df="dunif(theta1,0,1)"))
bayes1 <- adaBayes(yuima, start=param.init, prior=prior,
                   lower=lower,upper=upper, method="mcmc")
```

```
         Estimate   Std. Error
theta1 0.1974995 0.008112845
theta2 0.3487866 0.126663874
```

- **The convergence of $\hat{\theta}_2^B$ is slow.**

$$\left(\sqrt{n}(\hat{\theta}_{1,n}^{aB} - \theta_1^*), \sqrt{nh}(\hat{\theta}_{2,n}^{aB} - \theta_2^*)\right) \to^d (\zeta_1, \zeta_2)$$

QBE by YUIMA (cf. Iacus and Yoshida Springer p.87)

- **Try qmle and adaBayes for $n = 2750$.**
- **The estimation of $\hat{\theta}_n^B$ is improved:**

```
> coef(summary(bayes1))
        Estimate  Std. Error
theta1 0.1978142 0.003730354
theta2 0.2925331 0.088708241
> coef(summary(mle1))
        Estimate  Std. Error
theta1 0.1979697 0.003732584
theta2 0.2914936 0.088761680
```

Quasi likelihood analysis for volatility

Stochastic regression model

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

$$dY_t = b_t dt + \sigma(X_t, \theta) dw_t, \quad t \in [0, T], \tag{4}$$

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

- **Data:**

$\mathrm{Z}_n = (X_{t_j}, Y_{t_j})_{0\leq j\leq n}$ **with** $t_j = jh$
for $h = h_n = T/n$. T **is fixed.**

- **For example, when** $b_t = b(Y_t, t)$ **and** $X_t = (Y_t, t)$, Y **is the time-inhomogeneous diffusion process.**
- **Ergodicity is not assumed.**
- **Remark. Even if the drift coefficient** b_t **is parametrically modeled, it is known that, under a finite time horizon, consistent estimatimation of the drift parameter is impossible. So, we are interested in the parameter** θ **in the diffusion coefficient only.**

Quasi likelihood

- **Quasi log likelihood function:**

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

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

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

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

QLA estimators

- $\hat{\theta}_n^M$: the quasi-maximum likelihood estimator (QMLE) defined as

$$\mathbb{H}_n(\hat{\theta}_n^M) = \sup_{\theta \in \overline{\Theta}} \mathbb{H}_n(\theta). \tag{5}$$

- $\hat{\theta}_n^B$: the quasi-Bayesian estimator (QBE) for a prior density $\pi : \Theta \to \mathbb{R}_+$ is defined by

$$\hat{\theta}_n^B = \left(\int_\Theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta \right)^{-1} \int_\Theta \theta \exp(\mathbb{H}_n(\theta))\pi(\theta)d\theta. \tag{6}$$

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

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: $\mathrm{Z}_n = (X_{t_j}, Y_{t_j})_{0 \le j \le n}$ with $t_j = jh$ for $h = h_n = T/n$.

Asymptotic properties of the QLA estimators

- By the QLA theory, we obtain

  Theorem 3. For the $\mathsf{A} \in \{M, B\}$,

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

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

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

  Here ζ is a standard Gaussian random vector $\perp\!\!\!\perp$ $\Gamma(\theta^*)$.

- Non-ergodic statistics
- Reference. Uchida and Yoshida SPA2013

Remark

- **YUIMA qmle returns some estimated value of θ_2 even when the time horizon nh is not sufficiently large.**
- **The user should be careful to use such estimated value.**
- **There is no theoretical backing unless all condition are satisfied.**
- **If the process is not ergodic, then any long-term observation cannot ensure the correctness of the estimation of the drift parameter.**

MENU

1. QMLE and QBE
2. Quick introduction to the QLA theory
3. Applications of the Quasi-Likelihood Analysis
4. **Slightly deeper discussion (omitted)**

三春滝桜

2011年 春

http://www.takizakura.com/

Quasi likelihood analysis (QLA)

- $\boldsymbol{\Theta}$: a bounded open set in $\mathbb{R}^{\mathsf{p}}$, the parameter space
- $\boldsymbol{T} \in \mathbb{T}$ $(\mathbb{T} = \mathbb{Z}_+, \mathbb{R}_+, \ldots)$
- $\mathbb{H}_{\boldsymbol{T}} : \boldsymbol{\Omega} \times \boldsymbol{\Theta} \to \mathbb{R}$: a random field
- $\boldsymbol{\theta}^*$: "true" value of $\boldsymbol{\theta}$
- $\boldsymbol{a}_{\boldsymbol{T}} \in \mathrm{GL}(\mathbb{R}^{\mathsf{p}})$, $\boldsymbol{a}_{\boldsymbol{T}} \to \mathbf{0}$ $(\boldsymbol{T} \to \infty)$
- $\mathbb{U}_{\boldsymbol{T}} = \{\boldsymbol{u} \in \mathbb{R}^{\mathsf{p}};\ \boldsymbol{\theta}^* + \boldsymbol{a}_{\boldsymbol{T}}\boldsymbol{u} \in \boldsymbol{\Theta}\}$

Quasi likelihood ratio process

$$\mathbb{Z}_{\boldsymbol{T}}(\boldsymbol{u}) = \exp\left\{\mathbb{H}_{\boldsymbol{T}}(\boldsymbol{\theta}^* + \boldsymbol{a}_{\boldsymbol{T}}\boldsymbol{u}) - \mathbb{H}_{\boldsymbol{T}}(\boldsymbol{\theta}^*)\right\}$$

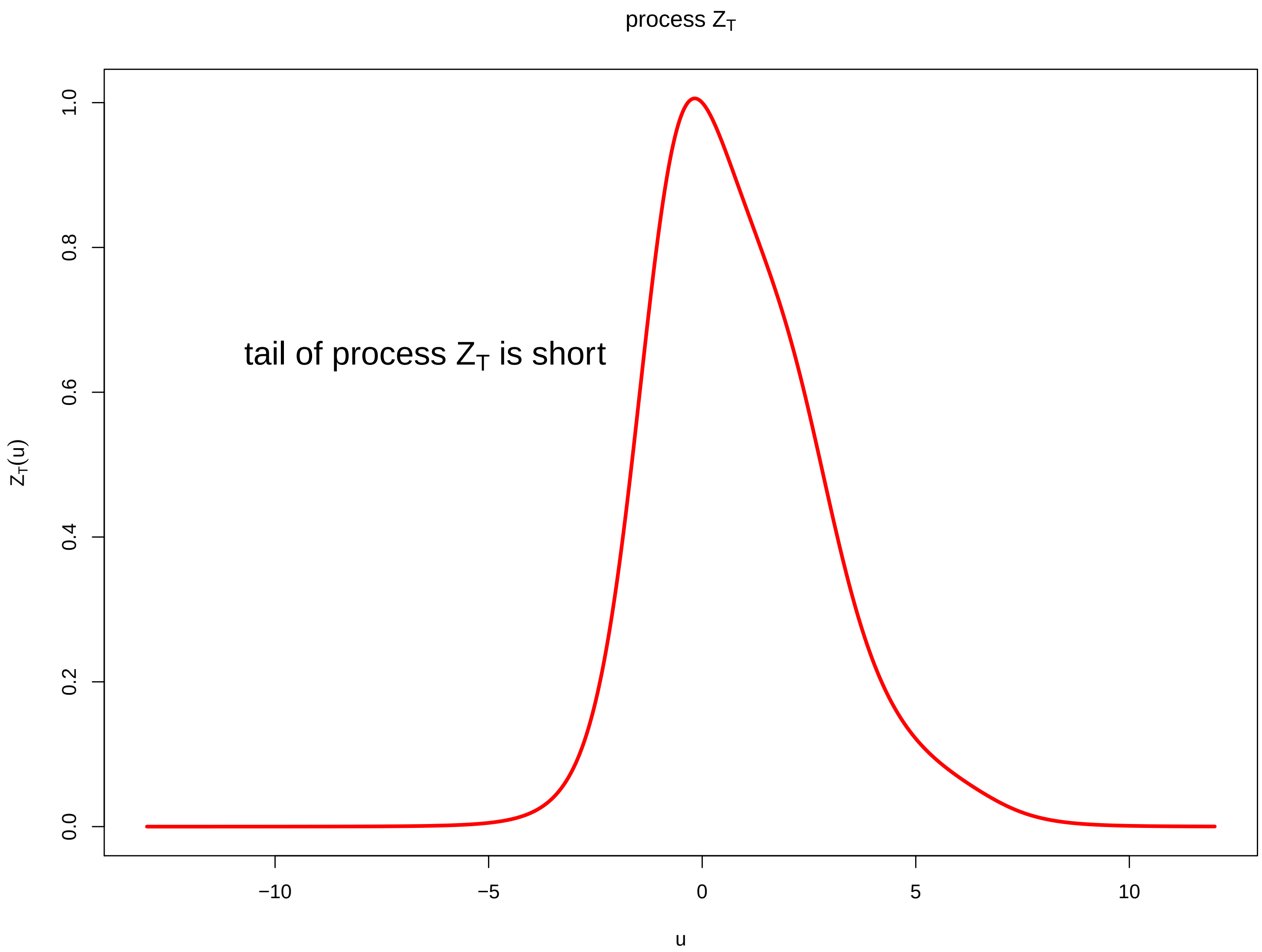
Locally asymptotically quadratic random field

Locally Asymptotically Quadratic (LAQ)

$$\mathbb{Z}_T(u) = \exp\left(\Delta_T[u] - \frac{1}{2}\Gamma[u^{\otimes 2}] + r_T(u)\right)$$

- Δ_T: a random vector (linear form)
- Γ: a random bilinear form
- $r_T(u) \to^p 0$ as $T \to \infty$
- Notation.
  $v[u] = \sum_i v_i u^i$,
  $M[u^{\otimes 2}] = \sum_{i,j} M_{i,j} u^i u^j$
  for $v = (v_i)$, $M = (M_{i,j})$ and $u = (u^i)$.
- Since $\mathbb{Z}_T$ is exponential of a nearly quadratic function, we expect fast decay of $\mathbb{Z}_T$ .....

Polynomial type large deviation (PLD) inequality: "tail of $\mathbb{Z}_T$ is short"



Polynomial type large deviation (PLD) inequality

- $V_T(r) = \{u \in \mathbb{U}_T;\ |u| \geq r\}$
- For $\exists\alpha \in (1,2)$, $\forall L > 0$, $\exists C_L$ such that

$$P\left[\sup_{u \in V_T(r)} \mathbb{Z}_T(u) \geq e^{-r^\alpha}\right] \leq \frac{C_L}{r^L} \quad (r > 0,\ T \in \mathbb{T})$$

NB:

$$\mathbb{Z}_T(u) = \exp\left\{\mathbb{H}_T(\theta^* + a_T u) - \mathbb{H}_T(\theta^*)\right\}$$

- LAQ + nondegeneracy of χ_0 implies PLD (Y 2011AISM), where χ_0 is ....

Key index χ_0

- $b_T = \{\lambda_{\min}(a_T' a_T)\}^{-1}$ (ex. $b_T = T$)
- $\mathbb{Y}_T(\theta) = \frac{1}{b_T}\big(\mathbb{H}_T(\theta) - \mathbb{H}_T(\theta^*)\big)$
- $\mathbb{Y}_T(\theta) \to^{L^p} \mathbb{Y}(\theta)$ $(T \to \infty)$ (and slightly more)
- **Assumption.**
  $\exists$ a positive r.v. χ_0 and a positive constant ρ s.t.

$$\mathbb{Y}(\theta) = \mathbb{Y}(\theta) - \mathbb{Y}(\theta^*) \leq -\chi_0|\theta - \theta^*|^\rho$$

- Nondegeneracy of the key index: for $\forall L > 0$, $\exists C_L$,

$$P[\chi_0 \leq r^{-1}] \leq \frac{C_L}{r^L} \quad (r > 0)$$

- Nondegeneracy of the key index $\Rightarrow$ PLD inequality

L^p-boundedness of the quasi likelihood estimators

- $\hat{u}_T^M := a_T^{-1}(\hat{\theta}_T^M - \theta^*)$
- Tail probability

$$P[|\hat{u}_T^M| \geq r] \leq P\left[\sup_{u \in V_T(r)} \mathbb{Z}_T(u) \geq 1\right] \leq \frac{C_L}{r^L}$$

In particular, $\sup_T \|\hat{u}_T\|_p < \infty$.

- Similarly, for $\hat{u}_T^B = a_T^{-1}(\hat{\theta}_T^B - \theta^*)$,

$$P[|\hat{u}_T^B| \geq r] \leq \frac{C_L}{r^L}$$

In particular, $\sup_T \|\hat{u}_T^B\|_p < \infty$.

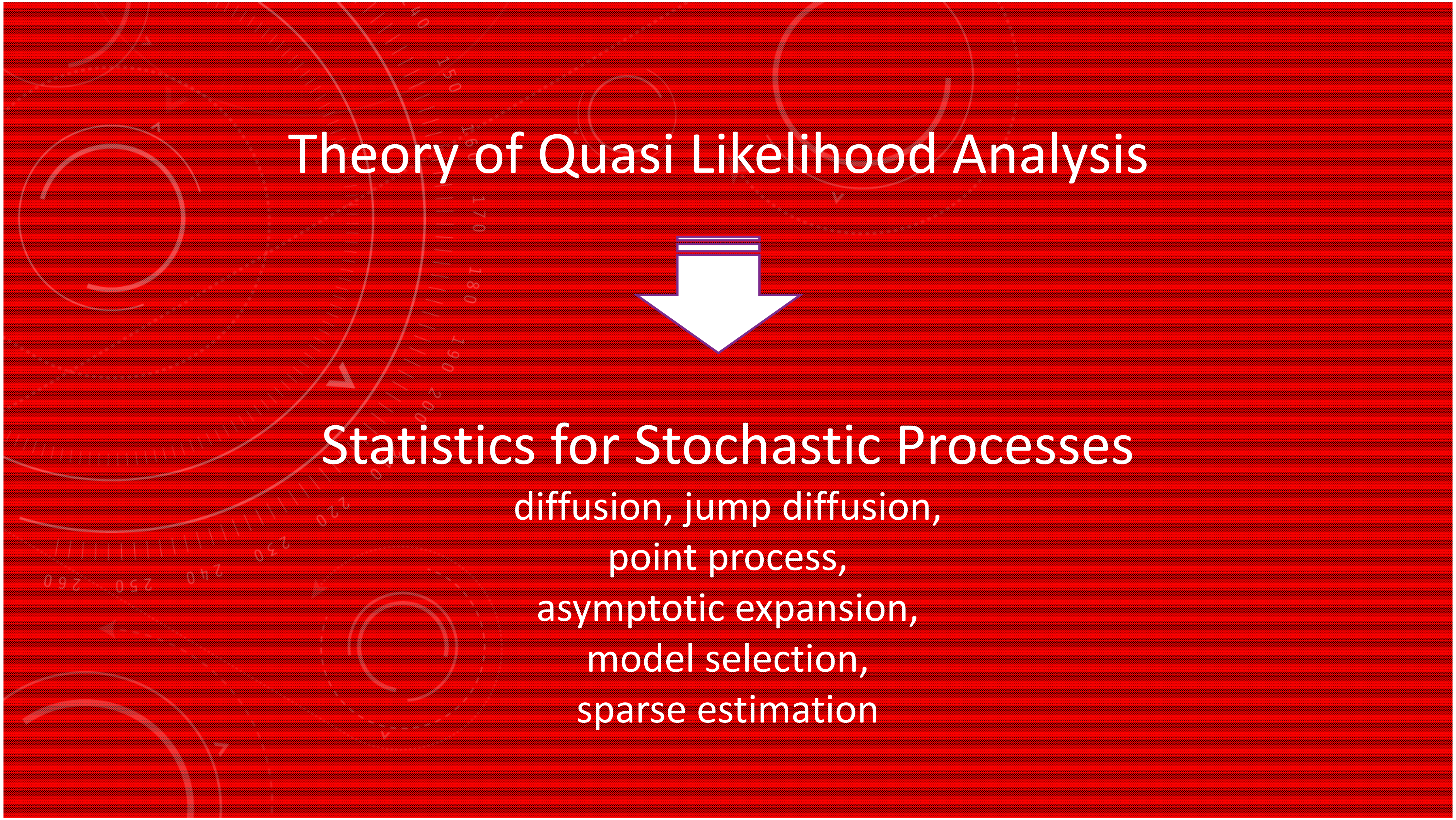
Scheme of the quasi likelihood analysis (QLA) based on Ibragimov-Has'minskii and Kutoyants program

$$\begin{cases} \cdot \text{ LAQ} \quad \mathbb{Z}_T(u) = \exp\left(\Delta_T[u] - \frac{1}{2}\Gamma[u^{\otimes 2}] + r_T(u)\right) \\ \cdot \text{ Limit theorem} \quad (\Delta_T, \Gamma) \to^d (\Delta, \Gamma) \\ \cdot \text{ PLD for } \mathbb{Z}_T(u)[\Leftarrow \text{nondegeneracy of } \chi_0] \end{cases}$$

$\Rightarrow$

$$\begin{cases} \cdot \ \mathbb{Z}_T \to^d \mathbb{Z} = \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u^{\otimes 2}]\right) \quad \text{in } \hat{C} \\ \cdot \ \hat{u}_T^M = a_T^{-1}(\hat{\theta}_T^M - \theta^*), \ \hat{u}_T^B = a_T^{-1}(\hat{\theta}_T^B - \theta^*) \to^d \Gamma^{-1}\Delta \\ \cdot \ L^p\text{-boundedness of } \{\hat{u}_T^M\}_T \text{ and } \{\hat{u}_T^B\}_T \end{cases}$$

where $\hat{C} = \{f \in C(\mathbb{R}^{\mathsf{p}});\ \lim_{|u|\to\infty} |f(u)| = 0\}$


Theory of Quasi Likelihood Analysis
Statistics for Stochastic Processes
diffusion, jump diffusion,
point process,
asymptotic expansion,
model selection,
sparse estimation

in summary: Scheme of the QLA

- **LAQ** $\mathbb{Z}_T(u) = \exp\left(\Delta_T[u] - \frac{1}{2}\Gamma[u^{\otimes 2}] + r_T(u)\right)$
  $+$ Moment conditions $\Rightarrow$ **PLD**
- **LAQ+Limit theorem** $(\Delta_T, \Gamma) \to^d (\Delta, \Gamma)$ + **PLD**
  $\Rightarrow$ Convergence of the random field

  $\mathbb{Z}_T \to^d \mathbb{Z}$ in $\hat{C} = \{f : \mathbb{R}^{\mathsf{p}} \to \mathbb{R}, \lim_{|u|\to\infty} |f(u)| = 0\}$

  where $\mathbb{Z}(u) = \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u^{\otimes 2}]\right)$
- Consequently

$$\hat{u}_T^M,\ \hat{u}_T^B \to^d \Gamma^{-1}\Delta$$

  $+$ L^p-boundedness of $\{\hat{u}_T^M\}_T$ and $\{\hat{u}_T^B\}_T$.

Quasi Likelihood Analysis: Summary

QLA: a systematic inferential framework $\ni$

- **(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.**

NB

- **QLA does not depend on a particular structure of the model.**

Recent applications of the Quasi Likelihood Analysis

- **sampled diffusion processes (Y AISM2011, Uchida and Y SPA2013)**
- **jump-diffusion processes (Ogihara and Y SISP2011)**
- **non-synchronous samapling (Ogihara and Y SPA2014)**
- **model selection (Uchida AISM2010, Uchida and Y 2016)**
- **asymptotic expansion (Y 2016)**
- **point processes (Clinet and Y SPA2016, Muni Toke and Y QF2016, QF2019?, Ogihara and Y arXiv2015)**
- **sparse estimation, penalized methods (Umezu-Shimizu-Masuda-Ninomiya, Kinoshita-Y, Suzuki-Y)**