Convergence of Multivariate Quantile Surfaces

Adil Ahidar

Institut de Mathématiques de Toulouse - CERFACS

August 30, 2013

Outline

Outline

Outline

Outline

Outline

Outline

Introduction

Let (X_n) be i.id. on $\mathbb{R}^d$, $P = \mathbb{P}^X$.

Desirable properties for multidimensional quantile:

- Generalization of unidimensionnal quantile
- Estimability by the sample (X_n)
- non parametric
- Interpretability (Mods detection, Mass localisation, $\cdots$)
- Admit ULLN, UCLT, SAP, $\cdots$
- Caracterize the law
- Calculability and Fast simulation

- We will define "quantile surfaces" via a class of shapes
- General classes can be used (curves, manifolds, ...)
- Here we focus on the hyperplans class

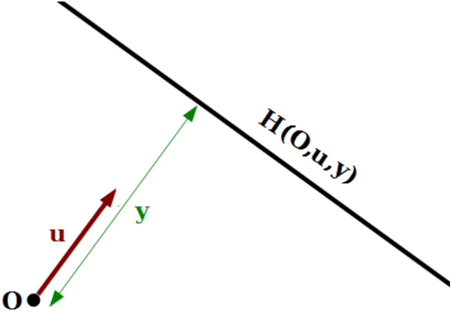
Definitions

Let define the α-th quantile surfaces associated to $\mathbb{P}$ and seen from $O \in \mathbb{R}^d$.

$\mathcal{H}$ the collection of all half-spaces and $\langle \cdot, \cdot \rangle$ the inner product in $\mathbb{R}^d$ and

$$H(O, u, y) = \left\{ x \in \mathbb{R}^d : \langle x - O, u \rangle \leq y \right\} \in \mathcal{H}$$

the halfspace standing at distance $y \in \mathbb{R}$ from O in direction $u \in \mathbb{S}_{d-1}$.



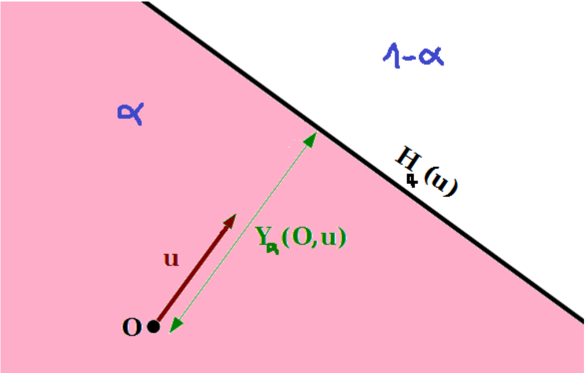
Given $\alpha \in (1/2, 1)$ and a direction $u \in \mathbb{S}_{d-1}$ let

$$Y_\alpha(O, u) = \inf\{y : P(H(O, u, y)) \geq \alpha\}$$

be the u-directional α-th quantile range and

$$H_\alpha(u) = H(O, u, Y_\alpha(O, u))$$

be the u-directional α-th quantile halfspace.



Remark that $Y_\alpha(O,u) = F^{-1}_{\langle X-O,u\rangle}(y)$ and thus $Y_\alpha(O,u)$ is the α-th real quantile of the real random variable $\langle X - O, u\rangle$

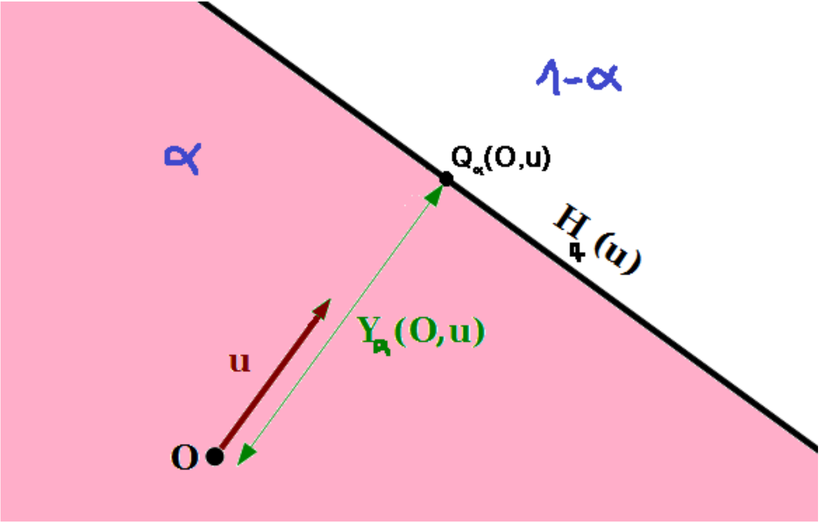
Definition (Multivariate quantile set)

For $\alpha \in (1/2, 1]$, $O \in \mathbb{R}^d$ and $u \in S_{d-1}$ define the u-directional α-th quantile point seen from O to be

$$Q_\alpha(O,u) = O + Y_\alpha(O,u)u \tag{1}$$

and the α-th quantile set seen from O to be the star-shaped collection of points

$$Q_\alpha(O) = \{Q_\alpha(O,u) : u \in \mathbb{S}_{d-1}\}.$$

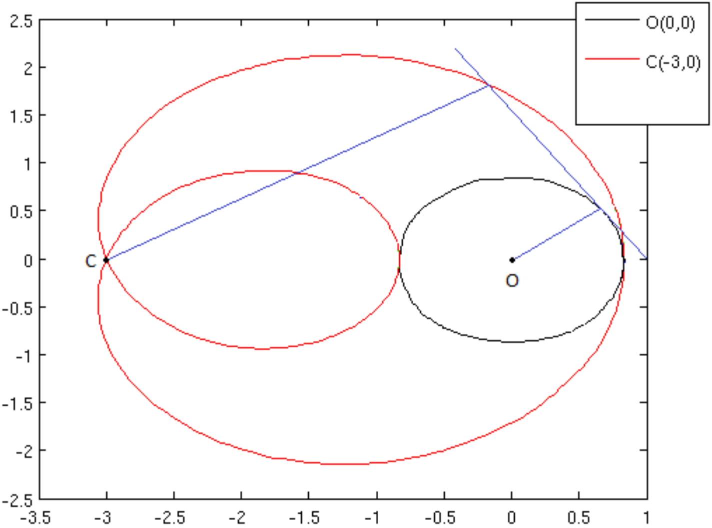

1-α
α
Q_α(O,u)
H(u)
Y_α(O,u)
u
O

Definition

Let

$$\mathcal{H}_\alpha = \{H_\alpha(u) : u \in \mathbb{S}_{d-1}\}$$
$$= \{H : H \text{ is a half-space, } P(H) = \alpha\}$$

$$\mathcal{C}_\alpha = \bigcap_{H \in \mathcal{H}_\alpha} H.$$


2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
O(0,0)
C(-3,0)
C
O

Empirical quantile surfaces

Let $\alpha \in \Delta = [\alpha^-, \alpha^+] \subset (1/2, 1)$ and $O \in \mathbb{R}^d$. Define P_n and $P_{n,O,u}$ as follows,

$$P_n = \frac{1}{n} \sum_{i \leq n} \delta_{X_i}, \quad P_{n,O,u} = \frac{1}{n} \sum_{i \leq n} \delta_{\langle X_i - O, u \rangle},$$

where δ_x is the Dirac mass at $x \in \mathbb{R}^d$ or $x \in \mathbb{R}$.

Definition

For $u \in \mathbb{S}_{d-1}$ *let*

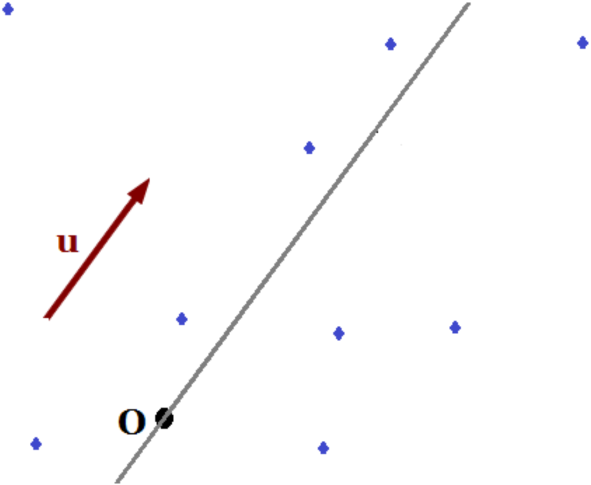
$$\begin{aligned} Y_{n,\alpha}(O,u) &= \inf\left\{y : P_n(H(O,u,y)) \geq \alpha\right\} \\ &= \inf\left\{y : P_{n,O,u}((-\infty,y)) \geq \alpha\right\}. \end{aligned}$$

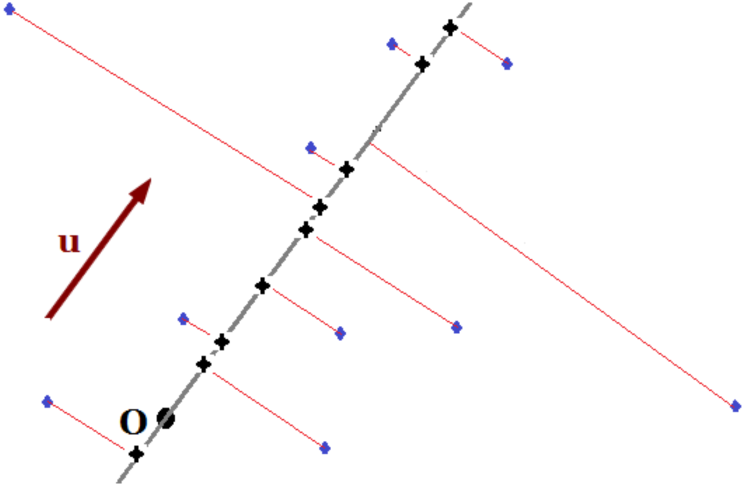
Define the u-directional α*-th empirical quantile point seen from O to be*

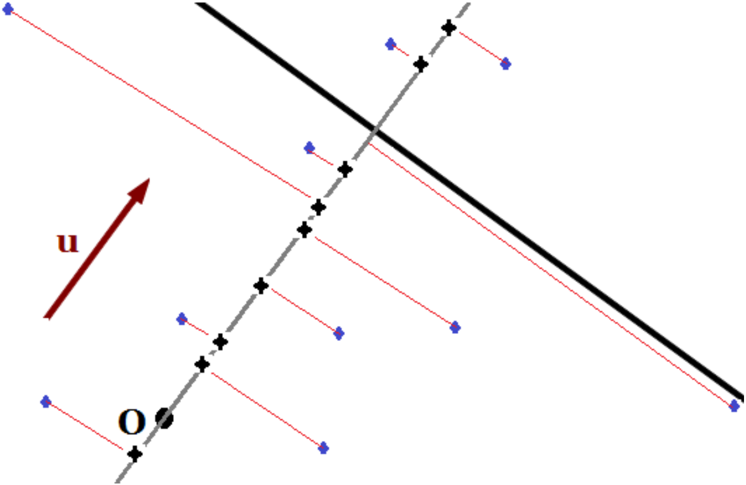
$$Q_{n,\alpha}(O,u) = O + Y_{n,\alpha}(O,u)u$$

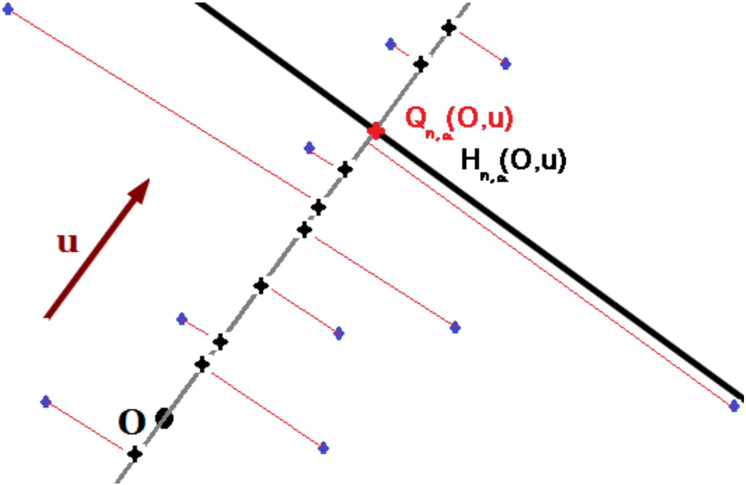
and associate to this point the subjective α*-th quantile half-space*

$$H_{n,\alpha}(u) = H(O,u,Y_{n,\alpha}(O,u)).$$


u
O


u
O


u
O


Q
(O,u)
H
(O,u)
u
O

...

Let the α-th empirical quantile set seen from O be

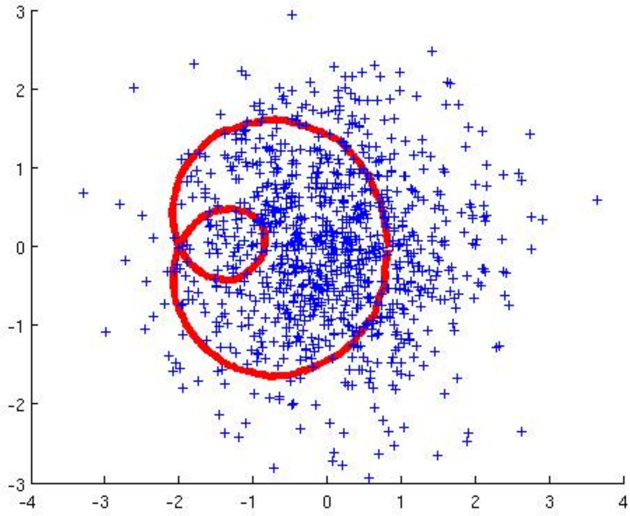
$$Q_{n,\alpha}(O) = \{Q_{n,\alpha}(O,u) : u \in \mathbb{S}_{d-1}\}.$$

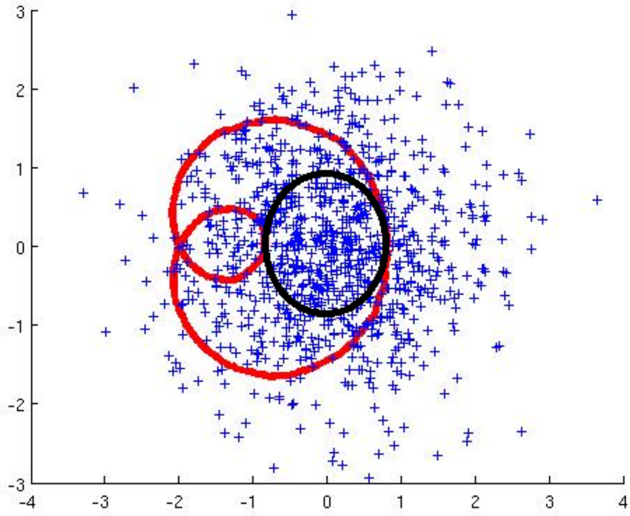
The half-spaces indexed by points $Q_{n,\alpha}(O,u)$ are collected into

$$\mathcal{H}_{n,\alpha} = \{H_{n,\alpha}(u) : u \in \mathbb{S}_{d-1}\}$$

from which we further define the random convex set

$$\mathcal{C}_{n,\alpha} = \bigcap_{H \in \mathcal{H}_{n,\alpha}} H.$$

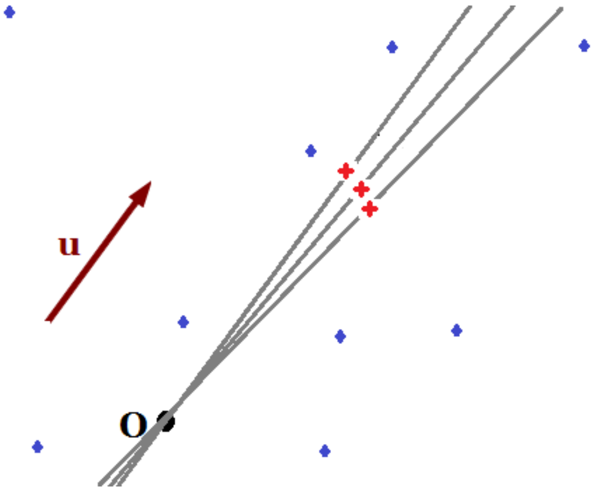

3
2
1
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
3
4


3
2
1
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
3
4

Proposition

The sets $\partial\mathcal{C}_{n,\alpha}$ and $Q_{n,\alpha}(O)$ are closed surfaces and piecewise spherical. We also have

$$\mathcal{H}_{n,\alpha} = \left\{ H : H \text{ is a half-space, } P_n(H) \in \left[\alpha, \alpha + \frac{d}{n}\right] \right\}.$$


u
O

Proposition

The processes $\mathbb{Q}_n - \mathbb{Q}$ *and* $\mathbb{Y}_n - \mathbb{Y}$ **do not depend on** O *and*

$$\sqrt{n}(\mathbb{Q}_n - \mathbb{Q})(u, \alpha) = \sqrt{n}(\mathbb{Y}_n - \mathbb{Y})(u, \alpha)u \quad (2)$$

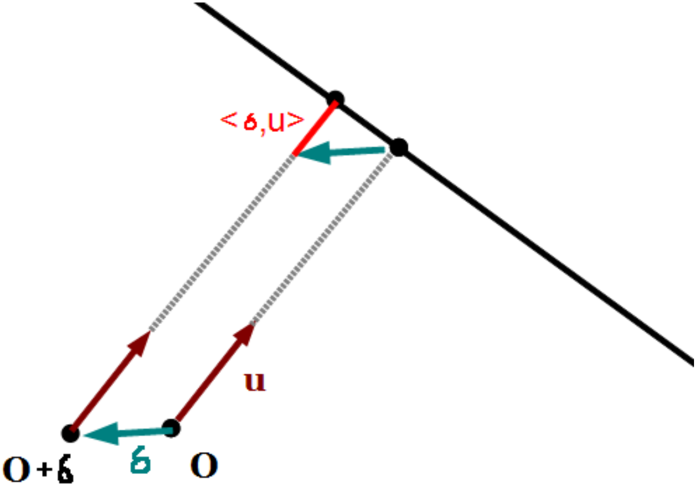
is a $\mathbb{R}^d$*-valued empirical process indexed by* $S_{d-1} \times \Delta$.

when O moves we have

$$\mathbb{Q}_\alpha(O + \delta, u) = \mathbb{Q}_\alpha(O, u) + \delta - \langle \delta, u \rangle \, u$$

and

$$\mathbb{Q}_{n,\alpha}(O + \delta, u) = \mathbb{Q}_{n,\alpha}(O, u) + \delta - \langle \delta, u \rangle \, u$$


<σ,u>
u
O+σ
σ
O

Directional regularity assumptions

Let $\Delta = [\alpha^-, \alpha^+] \subset \Delta_0 = \left(\alpha_0^-, \alpha_0^+\right)$, where $1/2 < \alpha_0^- < \alpha_0^+ < 1$.

Assumption A0

(A_0): for all $u \in \mathbb{S}_{d-1}$, for all open interval $I \subset F^{-1}_{\langle X-O,u\rangle}(\Delta)$

$$\mathbb{P}_{\langle X-O,u\rangle}(I) > 0$$

and for all $x \in \mathbb{R}$

$$\mathbb{P}_{\langle X-O,u\rangle}(\{x\}) = 0$$

Proposition

Assume that $d > 1$ and (A_0) holds. Then the sets $Q_\alpha(O)$, $\alpha \in \Delta$, are closed surfaces with at most two connex components and are unions of convex surfaces. Moreover $(u, \alpha) \mapsto \mathbb{Q}(u, \alpha)$ is continuous on $\mathbb{S}_{d-1} \times \Delta$.

From now we suppose for all $u \in \mathbb{S}_{d-1}$, the random variable $\langle X, u\rangle$ has a density $f_{\langle X,u\rangle}$ on $F^{-1}_{\langle X,u\rangle}(\Delta_0)$. This implies A_0

Notation

$$h_u = h_{\langle X,u\rangle} = f_{\langle X,u\rangle} \circ F^{-1}_{\langle X,u\rangle}. \tag{3}$$

Assumption A1

The function $(u,\alpha) \mapsto h(u,\alpha) = h_u(\alpha)$ is continuous on $\mathbb{S}_{d-1} \times \Delta$ and

$$0 < m \leq \inf_{\alpha\in\Delta_0} \inf_{u\in\mathbb{S}_{d-1}} h_u(\alpha) \leq \sup_{\alpha\in\Delta_0} \sup_{u\in\mathbb{S}_{d-1}} h_u(\alpha) \leq M < \infty.$$

Remark:

- Assumption (A_1) **does not imply** that P has a density on $\mathbb{R}^d$. However it implies that $P(H) = 0$ for all $H \in \mathcal{H}$.
- Under (A_1) $h_u = f_{\langle X-O,u\rangle} \circ F^{-1}_{\langle X-O,u\rangle}$ **do not depend** on O.
- Allows *supp* $\mathbb{P}$ with **low dimension**.

Main Results

Theorem (Uniform consistency)

Under (A_1) we have

$$\lim_{n\to\infty} \sup_{u\in\mathbb{S}_{d-1}} \sup_{\alpha\in\Delta} |Y_{n,\alpha}(O,u) - Y_\alpha(O,u)| = \lim_{n\to\infty} \|\mathbb{Y}_n - \mathbb{Y}\|_{S_{d-1}\times\Delta} = 0 \ a.s.$$

Corollary

Under (A_1) we have

$$\lim_{n\to\infty} \sup_{O\in\mathcal{D}} \sup_{u\in\mathbb{S}_{d-1}} \sup_{\alpha\in\Delta} \|\mathbb{Q}_{\alpha,n}(O,u) - \mathbb{Q}_\alpha(O,u)\| = 0 \quad a.s.$$

$$\lim_{n\to\infty} \sup_{O\in\mathcal{D}} \sup_{\alpha\in\Delta} d_H(\mathbb{Q}_{\alpha,n}(O), \mathbb{Q}_\alpha(O)) = 0 \quad a.s.$$

$$\lim_{n\to\infty} \sup_{\alpha\in\Delta} d_H(\mathcal{C}_{\alpha,n}, \mathcal{C}_\alpha) = 0 \quad a.s.$$

To provide the UCLT we define the limiting gaussian process $\mathbb{G}_P$

Definition

Let B_P be the P-Brownian bridge indexed by half-spaces, that is the zero mean Gaussian process on $\mathcal{H}$ with
$cov(B_P(H), B_P(H')) = P(H \cap H') - P(H)P(H')$, *for* $(H, H') \in \mathcal{H} \times \mathcal{H}$

For O,u fixed we have

$$cov(B_{O,u}(y), B_{O,u}(y')) = \min(F_{\langle X-O,u \rangle}(y), F_{\langle X-O,u \rangle}(y')) - F_{\langle X-O,u \rangle}(y)F_{\langle X-O,u \rangle}(y').$$

with notation $B_{O,u}(y) := B_P(H(O, u, y))$

Definition

For $(u, \alpha) \in \mathbb{S}_{d-1} \times \Delta$ *we define*

$$\mathbb{G}_P\,(u,\alpha) = \frac{B_P(H(O,u,Y(O,u,\alpha)))}{h(u,\alpha)} = \frac{B_P(H_\alpha(u))}{h(u,\alpha)}. \qquad (4)$$

For $(u_1, \alpha_1), (u_2, \alpha_2) \in \mathbb{S}_{d-1} \times \Delta$ we have

$$cov(\mathbb{G}_P\,(u_1,\alpha_1)\,, \mathbb{G}_P\,(u_2,\alpha_2)) = \frac{P(H_{\alpha_1}(u_1) \cap H_{\alpha_2}(u_2)) - \alpha_1\alpha_2}{h(u_1,\alpha_1)h(u_2,\alpha_2)}.$$

and the correlation

$$\begin{aligned} c(u_1,\alpha_1,u_2,\alpha_2) &= corr(\mathbb{G}_P(u_1,\alpha_1), \mathbb{G}_{P,\alpha}(u_2,\alpha_2)) \\ &= \frac{P(H_{\alpha_1}(u_1) \cap H_{\alpha_2}(u_2)) - \alpha_1\alpha_2}{\sqrt{\alpha_1(1-\alpha_1)}\sqrt{\alpha_2(1-\alpha_2)}} \\ &\in \left[-\sqrt{\frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1\alpha_2}}, \sqrt{\frac{(\alpha_1 \vee \alpha_2)(1-\alpha_1 \wedge \alpha_2)}{(\alpha_1 \wedge \alpha_2)(1-\alpha_1 \vee \alpha_2)}} \;\right] \end{aligned}$$

Let $\mathcal{B}_\infty$ be the set of all bounded real functions on $\mathbb{S}_{d-1} \times \Delta$, endowed with the supremum norm.

Theorem (Uniform central limit theorem)

If P satisfies (A_1) then the sequence $\sqrt{n}(\mathbb{Y}_n - \mathbb{Y})$ weakly converges to $\mathbb{G}_P$ on $\mathcal{B}_\infty$.

Corollary

Finite dimensional marginal laws convergence Fix $(O_1, u_1, \alpha_1), \ldots, (O_k, u_k, \alpha_k)$ in $\mathbb{R}^d \times \mathbb{S}_{d-1} \times \Delta$. Under (A_1) we have

$$\sqrt{n} \begin{pmatrix} Y_n(O_1, u_1, \alpha_1) - Y(O_1, u_1, \alpha_1) \\ \ldots \\ Y_n(O_k, u_k, \alpha_k) - Y(O_k, u_k, \alpha_k) \end{pmatrix} \to_{law} \mathcal{N}(0_k, \Sigma)$$

where the limiting covariance matrix Σ *has coordinates*

$$\Sigma_{i,j} = \frac{P(H_{\alpha_i}(u_i) \cap H_{\alpha_j}(u_j)) - \alpha_i \alpha_j}{h_{\alpha_i}(u_i) h_{\alpha_j}(u_j)}.$$

Assumption A2

(A_2) :(A_1) holds and $h(u,\alpha) = h_u(\alpha)$ is differentiable on $\mathbb{S}_{d-1} \times \Delta_0$ in variables (u,α) with uniformly bounded derivatives.

Theorem (Uniform strong approximation with rate)

if (A_1) *then one can construct on the same probability space* $(\Omega, \mathcal{T}, \mathbb{P})$ *an i.i.d. sequence* X_n *with law* P *and* $\mathbb{G}_n$ *versions of* $\mathbb{G}_P$ *such that for* $O \in \mathbb{R}^d, \alpha \in \Delta, u \in \mathbb{S}_{d-1}$

$$\mathbb{Y}_n(O, u, \alpha) = \mathbb{Y}(O, u, \alpha) + \frac{\mathbb{G}_n(u, \alpha)}{\sqrt{n}} + \frac{\mathbb{Z}_n(u, \alpha)}{\sqrt{n}} \quad (5)$$

where $\mathbb{Z}_n = \sqrt{n}(\mathbb{Y}_n - \mathbb{Y}) - \mathbb{G}_n$ *is such that*

$$\lim_{n \to \infty} \|\mathbb{Z}_n\|_{S_{d-1} \times \Delta} = \lim_{n \to \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathbb{Z}_n(u, \alpha)| = 0 \quad a.s.$$

...

If P moreover satisfies (A_2) *then* $\mathbb{G}_n$ *can be constructed such that for* $v_1 = v_2 = 1/4$, $w_1 = 1/2$, $w_2 > 1$ *and, if* $d \geq 3$, $v_d = 1/(3+4d)$, $w_d = ...$, *there exists* $c_\theta(m, M, d) > 0$ *and* $n_\theta(m, M, d) > 0$ *such that we have, for all* $n > n_\theta$,

$$\mathbb{P}\left(\|\mathbb{Z}_n\|_{S_{d-1}\times\Delta} \geq c_\theta \frac{(\log n)^{w_d}}{n^{v_d}} \right) \leq \frac{1}{n^\theta}. \tag{6}$$

Let $\Lambda_n = \sqrt{n}(P_n - P)$ be the empirical process indexed by $\mathcal{H}$ and define

$$\mathbb{E}_n(u, \alpha) = \Lambda_n(H_\alpha(u)) = \sqrt{n}(P_n(H_\alpha(u)) - \alpha)$$

its restriction to

$$\mathcal{H}_\Delta = \bigcup_{\alpha \in \Delta} \mathcal{H}_\alpha = \{H : H \text{ is a half-space, } P(H) \in \Delta\}$$

Theorem (Bahadur-Kiefer type representation of multivariate quantiles)

If P satisfies (A_1)then we have

$$\lim_{n\to\infty} \left\| \sqrt{n}(\mathbb{Y}_n - \mathbb{Y})h + \mathbb{E}_n \right\|_{S_{d-1}\times\Delta} = 0 \quad a.s. \tag{7}$$

and if moreover satisfies (A_2) then it holds

$$\left\| \sqrt{n}(\mathbb{Y}_n - \mathbb{Y})h + \mathbb{E}_n \right\|_{S_{d-1}\times\Delta} = O_{a.s.}\left(\frac{(\log n)^{w_d}}{n^{v_d}} \right). \tag{8}$$

steps of the proof

Step I: Approximate the empirical process

- $\mathcal{F}$ is pointwise measurable, uniformly bounded
- for $\nu_0 > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{\nu_0}}, \quad 0 < \varepsilon < 1$$

- if $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{C}\}$, $\mathcal{C}$ *VC-class* then $\nu_0 = VC - 1$

Corollary

We can construct $\{X_n\}$ and $\{\mathbb{G}_n\}$ on Ω such that

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > c_\theta \frac{(\log n)^{\tau_0}}{n^\tau}\right) \leq \frac{1}{n^\theta}$$

where

$$\tau = \frac{1}{2 + 5\nu_0}, \quad \tau_0 = \frac{4 + 5\nu_0}{4 + 10\nu_0}$$

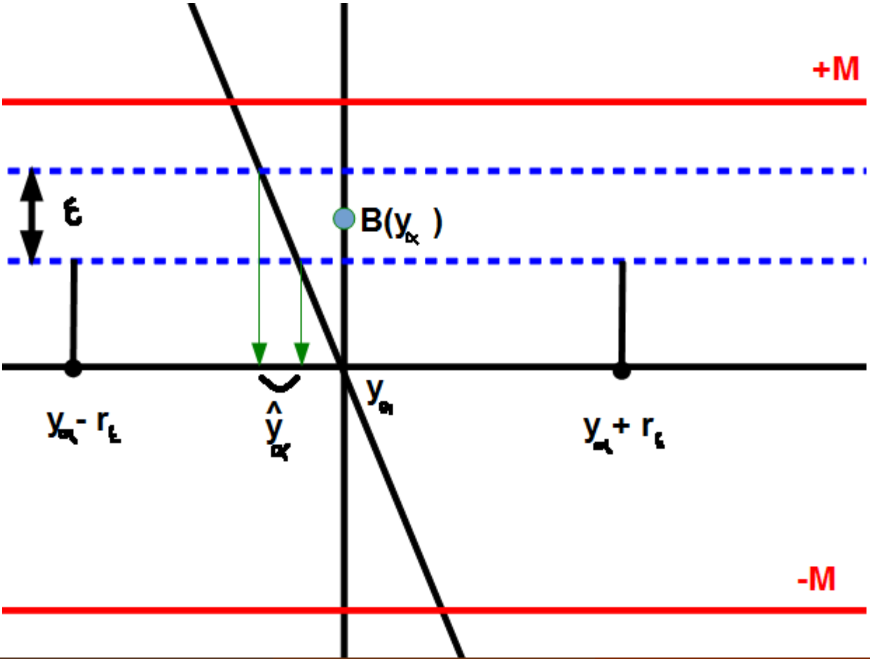
step II

- Discretization of the direction
- Control of the Modulus of continuity and deviation of the supremum of processes
- Given one direction and one quantile level α the process $\mathbb{G}_P$ reduces to a standart Brownian bridge on $\mathbb{R}$

For $\xi_n^U = \sqrt{n}(\alpha_n^U - B_n^U)$ and $\xi_n = \xi_n^U \circ F$, $B_n = B_n^U \circ F$ and $y_\alpha = F^{-1}(\alpha)$, $\hat{y} = F_n^{-1}(\alpha)$ We have

$$\begin{aligned}\hat{y} &= \inf\{y : F_n(y) \geq \alpha\}\\ &= \inf\left\{y : F(y) + \frac{\alpha_n(y)}{\sqrt{n}} \geq F(y_\alpha)\right\}\\ &= \inf\left\{y : B_n(y) \geq \sqrt{n}\left(F(y_\alpha) - F(y)\right) - \frac{\xi_n(y)}{\sqrt{n}}\right\}\\ &\approx \inf\left\{y : B(y) \geq -\sqrt{n} f\left(y_\alpha\right)\left(y - y_\alpha\right)\right\}\end{aligned}$$

we have
$B(y) \in [B(y_\alpha) - \epsilon,\ B(y_\alpha) + \epsilon]$ with large probability on $[y_\alpha - \delta_\epsilon,\ y_\alpha - \delta_\epsilon]$


+M
ε
B(y)
y - r
y
y + r
-M

$$B(y_\alpha) \pm \epsilon = -\sqrt{n}h(y_\alpha)\left(\hat{y} - y_\alpha\right)$$

donc

$$\begin{aligned}\hat{y} &= -\frac{B(y_\alpha) \pm \epsilon}{\sqrt{n}h(y_\alpha)} + y_\alpha \\ &\approx -\frac{B(y_\alpha)}{\sqrt{n}h(y_\alpha)} + y_\alpha\end{aligned}$$

- Approximation (with speed) of the directional quantile via the Gaussian that approximate the empirical process
- Joint region of confidence
- Fast simulation of the quantile surface

- Extension to general classes (geodesics,...)
- Mass localisation via different classes and exploration with O
- New depth notion via $\mathbb{Q}_{\alpha}(O)$
- mods detetction

thank you for your attention