Asymptotics of the maximal radius of an L^r-optimal sequence of quantizers.

Abass SAGNA (join work with Gilles Pagès)
abass.sagna@upmc.fr

Laboratoire de Probabilitès et Modèles Aléatoires,
Université Pierre et Marie Curie

Conference on Numerical Methods in Finance
April 15-17 2009

Plan

Plan

Quantization problem

Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ be a $r.v.$ with distribution $\mathbb{P}_X = P$ and $X \in L^r(\mathbb{P})$. Quantizing X (w.r.t the L^r norm, at level n) consists on finding the best approximation of X (w.r.t the L^r norm) by an $\mathbb{R}^d$ $r.v.$ $q(X)$ of cardinal at most n, with q a Borel function on $\mathbb{R}^d$. The associated L^r-mean quantization error is given by

$$e_{n,r}(X) = \inf\{\|X - q(X)\|_r, q : \mathbb{R}^d \stackrel{\text{Borel}}{\longrightarrow} \mathbb{R}^d,\ \mathrm{card}(q(\mathbb{R}^d)) \le n\}.$$

Let $q : \mathbb{R}^d \mapsto \alpha$ and set

$$\widehat{X}^\alpha = \sum_{a \in \alpha} a \mathbf{1}_{\{X \in C_a(\alpha)\}} \tag{1}$$

with

$$C_a(\alpha) \subset \{x \in \mathbb{R}^d : |x - a| = \min_{b \in \alpha} |x - b|\}. \tag{2}$$

Then the L^r mean quantization problem reads

$$e_{n,r}(X) \;=\; \inf\{\|X - \widehat{X}^\alpha\|_r, \alpha \subset \mathbb{R}^d,\ \mathrm{card}(\alpha) \le n\} \tag{3}$$

Zador Theorem

Definition

A sequence of quantizers $(\alpha_n)_{n\geq 1}$ *is* $L^r(\mathbb{P}_X)$*-optimal if for every* $n \geq 1$, α_n *realize the infimum (with is in fact a minimum) in* (3).

1. The L^r-mean quantization error $e_{n,r}(X) \searrow 0$ as $n \to +\infty$.
2. Let $P = P_a + P_s$. Let $X \in L^{r+\eta}(\mathbb{P})$ for some $\eta > 0$. Then

$$\lim_{n\to+\infty} n^{1/d} e_{n,r}(P) = Q_r(P)^{1/r} \tag{4}$$

with $Q_r(P) = J_{r,d} \left\| \frac{dP_a}{d\lambda_d} \right\|_{\frac{d}{d+r}}$.

Applications to Numerical probabilities

▷ We estimate $\mathbb{E}f(X)$ by $\mathbb{E}f(\widehat{X}) = \sum_{i=1}^{n} f(x_i^\star)\mathbb{P}(X \in C_i(x_i^\star))$ owing to the weighted empirical measure theorem: $\sum_{i=1}^{n} p_i^\star \delta_{x_i^\star} \Longrightarrow P.$

▷ In nonlinear filtering problem: if X is a signal process and Y the observed process we estimation $\mathbb{E}(f(X_n)|(Y_1, \cdots, Y_n) = y)$ by $\sum_{i=1}^{N_n} \widehat{\Pi}_y^i f(x_i)$.

• Bally, Pagès, Printems: A quantization tree method for pricing and hedging multidimensional American options.

• Pagès, Pham: Optimal quantization methods for nonlinear filtering with discrete-time observations.

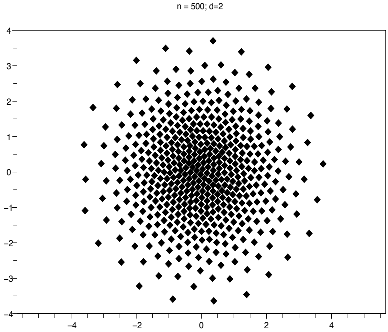
• Pagès, Pham, Printems: An Optimal markovian quantization algorithm for multidimensional stochastic control problems.

• Callegaro, Sagna: Estimation of the default probability in partial information model by an hybrid MC-Quantization

• Pagès, Printems: Functional quantization for numerics with an application to option pricing.

Plan

Definition of the maximal radius



In this talk we are interested to the asymptotics of the maximal radius sequence of an $L^r(P)$-optimal sequence of quantizers (α_n) defined for every $n \geq 1$ by

$$\rho(\alpha_n) := \rho_n := \max\{|a|,\ a \in \alpha_n\}. \tag{5}$$

limit of the maximal radius

Proposition

If $\mathrm{card}(\mathrm{supp}(P)) = +\infty$ *then for every* $L^r(P)$*-optimal sequence of quantizers* (α_n),

$$\lim_{n\to+\infty} \rho_n = \sup_{n\geq 1} \rho_n = \sup\{|x|,\ x \in \mathrm{supp}(P)\}. \tag{6}$$

The norm $|\cdot|$ *is Euclidean.*

▷ A Idea of the proof: One shows that

- $$\liminf_{n\to+\infty} \rho_n \geq \sup\{|x|,\ x \in \mathrm{supp}(P)\}. \tag{7}$$
- For every n such that $\mathrm{card}(\mathrm{supp}(P)) \geq n$,

$$\alpha \subset \overline{\mathrm{Conv}(\mathrm{supp}(P))} \quad \text{and} \quad \rho_n \leq \sup\{|x|,\ x \in \mathrm{supp}(P)\}. \tag{8}$$

Plan

Plan

Main result

We will make the following hypothesis on P in the general settings

$$\textbf{(H)} \equiv \qquad P(dx) \geq \varepsilon_0 \, \mathbf{1}_{\{x \in \bar{B}(x_0, r_0)\}} \lambda_d(dx), \quad \varepsilon_0,\ r_0 > 0,\ x_0 \in \mathbb{R}^d.$$

Theorem

Suppose X has an unbounded support and **(H)** *holds. Let $(\alpha_n)_{n \geq 1}$ be $L^r(P)$-optimal. Then, for every $r > 0$, for every $d \geq 1$,*

$$\lim_{\varepsilon \searrow 0} \, \underline{\lim}_{n} \left(n^{1+\frac{r}{d}} \mathbb{E}\left(|X|^r \mathbf{1}_{\{|X| > \frac{\rho_n}{2+\varepsilon}\}} \right) \right) \geq C_{r,d}. \qquad (9)$$

Rem. If $d = 1, r \geq 1$, under Assumption $(\mathbf{G_r})$: $P = f \cdot \lambda_d$ where f is non-increasing toward 0 on $[A, +\infty)$ for some $A > 0$ and

$$\lim_{y \to +\infty} \int_1^{+\infty} (u-1)^{r-1} \frac{f(uy)}{f(y)} du = 0$$

then (9) holds with "$1 + \varepsilon$" instead of "$2 + \varepsilon$".

toward the proof of the theorem

- ▷ Let $M(\alpha_n) = \{a \in \alpha_n,\ |a| = \rho_n\}$. Then, $\forall\, \varepsilon > 0,\ \exists\, n_\varepsilon$ t.q. $\forall n \geq n_\varepsilon$,

$$\forall\ a \in M(\alpha_n),\ \forall\ \xi \in C_a(\alpha_n), \qquad |\xi| \geq \frac{\rho_n}{2+\varepsilon}. \tag{10}$$

  with "$1+\varepsilon$" in place of "$2+\varepsilon$" in (10) when $d = 1,\ r \geq 1$ and if $(\mathbf{G_r})$ holds.

- ▷ Let $(\alpha_n)_{n\geq 1}$ be $L^r(P)$-optimal. Suppose that **(H)** holds. Then for large enough n,

$$e^r_{n,r}(X) - e^r_{n+1,r}(X) \geq C_{r,d}\ n^{-\frac{r+d}{d}}. \tag{11}$$

From now on, set for every $r > 0$,

$$\bar{F}_r(x) = \mathbb{E}\left(|X|^r \mathbf{1}_{\{|X|>x\}}\right) \text{ and } f^{\leftarrow}(x) = \inf\{t > 0,\ f(t) \geq x\}.$$

Asympt. from the upper bound of the asympt. inverse of $-log\,\bar{F}_r(\cdot)$

Suppose that X has an unbounded support and that (**H**) is fulfilled.
h is an asymptotic inverse function of g if $h(g(x)) = x + o(x)$

Proposition

(a) *If $\psi_r \in R_\delta$ is measurable nondecreasing, and $\psi_r(-\log \bar{F}_r(x)) \geq x + o(x)$ as $x \to +\infty$, then*

$$\overline{\lim_n} \frac{\rho_n}{\psi_r(\log(n))} \leq c_{r,d} \left(1 + \frac{r}{d}\right)^{\delta}. \tag{12}$$

with $c_{r,d} = 1$ if $d = 1$, $r \geq 1$ and $(\mathbf{G_r})$ holds; $c_{r,d} = 2$ otherwise.
(b) *If $\psi_r \in R_\delta$ is measurable nondecreasing, and $\psi_r(-\log \bar{F}_r(e^x)) \geq x + o(x)$ when $x \to +\infty$, then*

$$\overline{\lim_n} \frac{\log \rho_n}{\psi_r(\log(n))} \leq \left(1 + \frac{r}{d}\right)^{\delta}. \tag{13}$$

Choice of ψ_r for distributions with expon. and polyn. tails:

Criterion

(a) *Let* $\theta^\star = \sup\left\{\theta > 0,\ \limsup_{x\to+\infty}\ e^{\theta x^\kappa}\bar{F}_r(x) < +\infty\right\}$,

for some κ*, then* $(\psi_r(x) = (\frac{x}{\theta})^{1/\kappa},\ \forall\theta \in (0,\theta^\star)\ )$

$$\overline{\lim_n}\ \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \le \frac{c_{r,d}}{(\theta^\star)^{1/\kappa}}\left(1+\frac{r}{d}\right)^{1/\kappa}. \tag{14}$$

(b) *Let* $\zeta^\star = \sup\left\{\zeta > 0,\ \limsup_{x\to+\infty}\ x^{\zeta-r}\bar{F}_r(x) < +\infty\right\}$.

Then $(\psi_r(x) = \frac{x}{\zeta-r},\ \forall\zeta \in (0,\zeta^\star)\ )$

$$\overline{\lim_n}\ \frac{\log\rho_n}{\log(n)} \le \frac{1}{\zeta^\star - r}\frac{r+d}{d}. \tag{15}$$

Asymp. for a specified density

Corollary

(a) *If* $f(x) \sim |x|^c \, e^{-\vartheta |x|^\kappa}$ *as* $|x| \to +\infty$; $\vartheta, \kappa > 0$; $c > -d$

then $\theta^\star = \vartheta$ *and*

$$\overline{\lim_n} \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \leq \frac{c_{r,d}}{\vartheta^{1/\kappa}} \left(1 + \frac{r}{d}\right)^{1/\kappa}. \tag{16}$$

(b) *If* $f(x) \sim \dfrac{(\log |x|)^\beta}{|x|^c}$ *as* $|x| \to +\infty$; $\beta \in \mathbb{R}$, $c > r + d$

then $\zeta^\star = c - d$ *and*

$$\overline{\lim_n} \, \frac{\log \rho_n}{\log(n)} \leq \frac{1}{c - d - r} \frac{r + d}{d}. \tag{17}$$

Plan

Preliminaries

Definition and limit of the maximal radius

convergence rate

- Limit superior
- Limit inferior

Bibliography

Distribution of type (r, s)

Any L^r-optimal sequence $(\alpha_n)_{n\geq 1}$, $s \leq r$, is L^s-**rate-optimal** *i.e.*

$$\overline{\lim_n}\ n^{1/d}\|X - \widehat{X}^{\alpha_n}\|_s < +\infty. \tag{18}$$

If $s > r$ (and $X \in L^s(\mathbb{P})$) this rate optimality usually fails. So is always the case then $s > r + d$ and $\lambda_d(f > 0) = \infty$.

Definition

*Let $s, r > 0, s > r$. A r.v. $X \in L^s(\mathbb{P})$ has an (r, s)-**distribution** if any L^r-optimal sequence of quantizers $(\alpha_n)_{n\geq 1}$ is L^s-rate-optimal.*

Suppose that X has an (r, s)-distribution for some $\nu > 0$ and set

$$\nu_X^\star := \sup\{\nu > 0 \text{ s.t. } X \text{ has an } (r, r+\nu)\text{-distribution}\}$$

main result

Theorem

Suppose X has an $(r, r+\nu)$-distribution for some $\nu > 0$. Let $(\alpha_n)_{n\geq 1}$ be an $L^r(P)$-optimal sequence of quantizers. Then,

$$\overline{\lim_{n}} \sup_{u>1} \left(\left(1 - 1/u\right)^{r+\nu} \; n^{\frac{r+\nu}{d}} \bar{F}_{r+\nu}\left(u\rho_n\right) \right) < +\infty. \qquad (19)$$

How to choose $\psi_{r,\nu}$?

Criterion

i) *Let* $\kappa > 0$ *and set*

$$\theta_\star = \inf\left\{\theta > 0,\ \liminf_{x\to+\infty} e^{\theta x^\kappa}\mathbb{P}(|X| > x) > 0\right\}. \qquad (20)$$

Then

$$\lim_n \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \geq \frac{1}{(\theta_\star)^{1/\kappa}}\left(\frac{r+\nu^\star}{d}\right)^{1/\kappa}. \qquad (21)$$

ii) *Let*

$$\zeta_\star = \inf\left\{\zeta > 0,\ \forall \nu \in (0,\nu^\star),\ \liminf_{x\to+\infty} x^{\zeta-r-\nu}\bar{F}_{r+\nu}(x) > 0\right\}, \qquad (22)$$

then

$$\lim_n \frac{\log \rho_n}{\log(n)} \geq \frac{1}{\zeta_\star - r - \nu^\star}\,\frac{r+\nu^\star}{d}. \qquad (23)$$

Corollary

If

$$f(x) \sim |x|^c \, e^{-\vartheta |x|^\kappa}, \; as \; |x| \to +\infty; \; \vartheta, \kappa > 0; \; c > -d \tag{24}$$

then

$$\nu^\star = d \quad and \quad \theta^\star = \theta_\star = \vartheta,$$

and for every $r > 0$, for every $d \geq 1$,

$$\frac{1}{\vartheta^{1/\kappa}} \left(1 + \frac{r}{d}\right)^{1/\kappa} \leq \varliminf_n \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \leq \varlimsup_n \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \leq \frac{2}{\vartheta^{1/\kappa}} \left(1 + \frac{r}{d}\right)^{1/\kappa} \tag{25}$$

If $d = 1$, $r \geq 1$ we have

$$\lim_{n \to +\infty} \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} = \left(\frac{r+1}{\vartheta}\right)^{1/\kappa}. \tag{26}$$

Corollary

(b) *If*

$$f(x) \sim \frac{(\log|x|)^{\beta}}{|x|^{c}} \quad as \quad |x| \to +\infty,\ \beta \in \mathbb{R},\ c > r + d \tag{27}$$

then

$$\nu^{\star} = d\left(1 - \frac{r+d}{c}\right) \in (0, d) \quad and \quad \zeta_{\star} = \zeta^{\star} = c - d$$

and for every $r > 0$, *for every* $d \geq 1$,

$$\lim_{n \to +\infty} \frac{\log \rho_n}{\log(n)} = \frac{1}{c - r - d} \frac{r + d}{d}. \tag{28}$$

Examples

(1) • If $X \sim \mathcal{N}(0, I_d)$, $\forall r > 0$, $\forall d \geq 1$,

$$\sqrt{\frac{2(r+d)}{d}} \leq \underline{\lim}_n \frac{\rho_n}{\sqrt{\log(n)}} \leq \overline{\lim}_n \frac{\rho_n}{\sqrt{\log(n)}} \leq 2\sqrt{\frac{2(r+d)}{d}}. \quad (29)$$

If $d = 1$ and $r \geq 1$ we have

$$\lim_{n \to +\infty} \frac{\rho_n}{\sqrt{\log(n)}} = \sqrt{2(r+1)}. \quad (30)$$

• If $X \sim \Gamma(a, \lambda)$, $a > 0, \lambda > 0$ or if X has a double gamma distribution then for every $\forall r \geq 1$,

$$\lim_{n \to +\infty} \frac{\rho_n}{\log(n)} = \frac{r+1}{\lambda}; \quad (31)$$

and $\forall r \in (0, 1)$,

$$\frac{r+1}{\lambda} \leq \underline{\lim}_n \frac{\rho_n}{\log(n)} \leq \overline{\lim}_n \frac{\rho_n}{\log(n)} \leq \frac{2(r+1)}{\lambda}. \quad (32)$$

Examples

• The logistic distribution has the same asymptotics as the exponential distribution with parameter $\lambda = 1$ since $f(x) \sim e^{-x}$.
• For the Weibull distribution with parameter $\kappa > 0$: $\forall r \geq 1$,

$$\lim_{n\to+\infty} \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} = (r+1)^{1/\kappa}.$$

When $r \in (0,1)$,

$$(r+1)^{1/\kappa} \leq \underline{\lim}_n \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \leq \overline{\lim}_n \frac{\rho_n}{\big(\log(n)\big)^{1/\kappa}} \leq 2(r+1)^{1/\kappa}.$$

(2) Suppose X has a Pareto distribution with $\gamma > r$:
$f(x) = \gamma x^{-(\gamma+1)}\mathbf{1}_{\{x>1\}}$.

$$\lim_{n\to+\infty} \frac{\log \rho_n}{\log(n)} = \frac{r+1}{\gamma - r}. \tag{33}$$

An alternative approach

Theorem

Let $X \sim P$ with $P_a \neq 0$. Assume $(\alpha_n)_{n\geq 1}$ is an $L^r(P)$-optimal sequence of n-quantizers. Let $(X_k)_{k\geq 1}$ be an i.i.d sequence of $\mathbb{R}^d$-valued random variables with probability distribution P. Then for every $\nu \in (0, \nu_X^\star)$,

$$\liminf_{n\to+\infty} \left(\rho_n - \mathbb{E}\big(\max_{k\leq [n^{(r+\nu)/d}]} |X_k|\big)\right) \geq -\ C_\nu \tag{34}$$

where C_ν is a positive real constant.

▷ For the exponential distribution

$$\liminf_{n\to+\infty} \frac{\rho_n}{\log(n)} \geq \frac{r+1}{\lambda}$$

▷ For the Pareto distribution we have

$$\liminf_{n\to+\infty} \frac{\log \rho_n}{\log(n)} \geq \frac{r+1}{\gamma+1}.$$

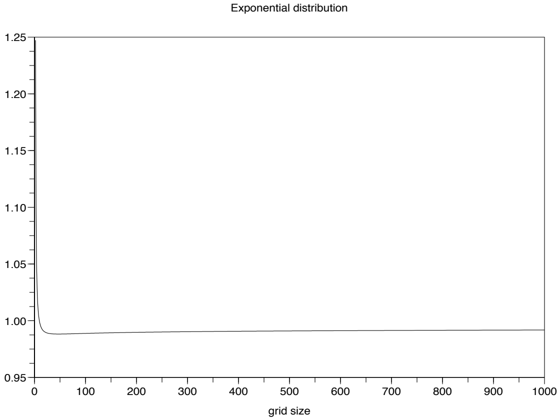


Figure: $\frac{\rho_n}{3\log(n)}$, $1 \leq n \leq 1\,000$ for the exponential distribution.

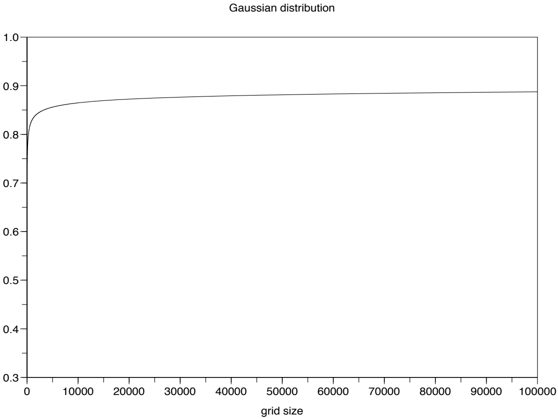


Figure: $\frac{\rho_n}{3\log(n)}$, $1 \leq n \leq 10\,000$ for the normal distribution.

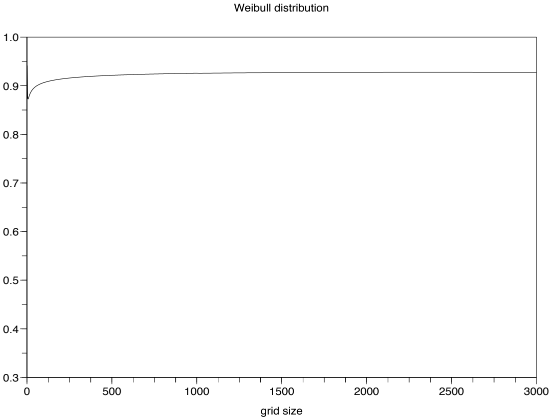


Figure: $\frac{\rho_n}{3\log(n)}$, $1 \leq n \leq 3\,000$ for the Weibull distribution.

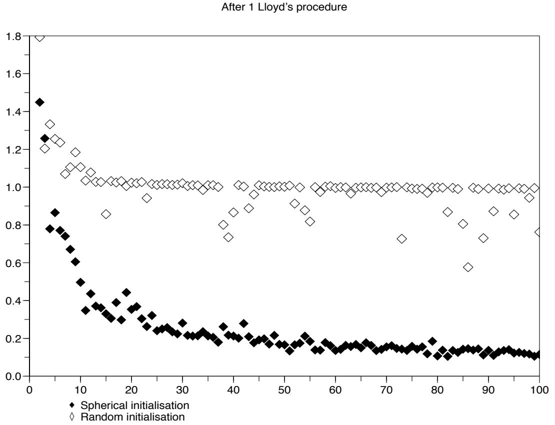


Figure: Comparaison of the distortions for different types of initializations after 1 Lloyd.

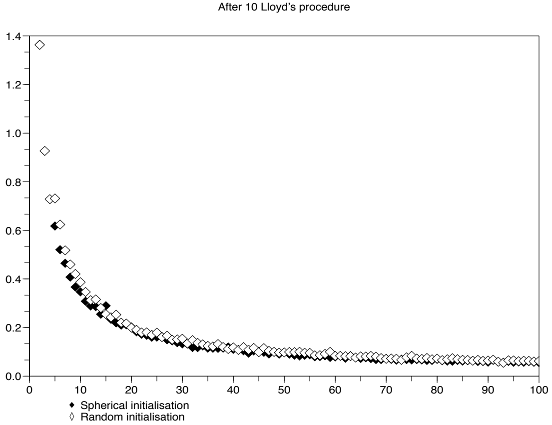


Figure: Comparaison of the distortions for different types of initializations after 10 Lloyd.

Plan

The bibliography.

- Graf, S., Luschgy, H., *Foundations of Quantization for Probability Distributions*, Lecture Notes in Mathematics, Vol 1730, Berlin, 2000.
- Graf, S., Luschgy, H., Pagès G., Distorsion mismatch in the quantization of probability measures, *ESAIM: Probability and Statistics*, 12, 127-153 (2008)..
- Delattre, S., Graf, S, Luschgy, H., Pagès G., Quantization of probability distributions under norm-based distortion measures, Statistics & Decisions 22, 261-282 (2004).
- Fort, J.-C. , Pagès G., Asymptotics of optimal quantizers for some scalar distributions, *Journal of Computational and Applied Mathematics* 146, 253-275 (2002).
- Pagès, G., A space quantization method for numerical integration, *Journal of Computational and Applied Mathematics* 89, 1-38 (1997).