CONVERGENT ESTIMATORS FOR THE L_1-MEDIAN OF A BANACH VALUED RANDOM VARIABLE

Benoît Cadre
Laboratoire de Probabilités et Statistiques
CC 051, Place Eugène Bataillon
34095 MONTPELLIER Cdex 05
FRANCE

Abstract Let E be a separable Banach space, which is the dual of a Banach space F. If X is an E-valued random variable, the set of L_1-medians of X is

$$\mathrm{Argmin}_{\alpha \in E} E[\|X-\alpha\| - \|X\|].$$

Assume that this set contains only one element. From any sequence of probability measures $\{\mu_n, n \geq 1\}$ on E, which converges in law to X, we give two approximating sequences of the L_1-median, for the weak* topology induced by F.

AMS 1991 subject classification : 62G05; 60F17; 60G60.
Key-words and phrases : L_1-median, nonparametric estimation, auto-reproducing space, weak* topology .

1. Introduction

Let E be a Banach space over $I\!R$ with norm $\|.\|$ and X be an E-valued Borel random variable defined on the probability space $(\Omega, \mathcal{F}, P)$. By a L_1-median of X (or, of the law of X), we mean a point of E which minimizes φ, where φ is defined for all $\alpha \in E$ by

$$\varphi(\alpha) = E[\|X-\alpha\| - \|X\|].$$

That is, $\alpha_0 \in E$ is a L_1-median of X if

$$\varphi(\alpha_0) = \inf_{\alpha \in E} \varphi(\alpha).$$

In various situations, such L_1-medians were already considered by Bedall and Zimmerman (1978), Bru and Heinich (1985), Gower (1974), Kemperman (1987), Milasevick and Ducharme (1987), Valadier (1984), ... (see also

Averous and Meste (1997) for an extended definition of the L_1-median). Note that this definition does not need the assumption $E[\|X\|] < \infty$. Moreover, in the particular case where $E = I\!R$, we recover the standard definition of the median.

Of course, existence and uniqueness of L_1-medians are not always guaranteed. But Valadier (1984) observed that L_1-medians exist provided E is reflexive (i.e. the closed unit ball is compact for the weak topology), such as $E = L_p$ with $1 < p < \infty$ or $E = I\!R^d$ with $d \geq 1$. Kemperman (1987) proved that it also suffices that E is the dual of a separable Banach space, which includes at least the cases above. It also includes the case where E is the set of all signed and bounded measures on a locally compact metric space. Uniqueness of the L_1-median was already obtained by Gower (1974) when $E = I\!R^d$ with Euclidean norm and the law of X is supported by finitely many points not all on a straight line. More generally, Kemperman (1987) proved that if E is a strictly convex space (i.e. $\|x + y\| < \|x\| + \|y\|$ whenever x and y are not proportional), and if the law of X is not carried by any straight line in E, then X possesses at most one L_1-median. Note that important examples of strictly convex spaces is that of Hilbert spaces and L_p-spaces with $1 < p < \infty$.

In the present paper, one considers the case where E is a separable Banach space which is the dual of a Banach space F i.e. $E = F^*$. Then also F is separable (see Dunford and Schwartz (1958), p.65). In the remainder of the paper, we assume that X possesses only one L_1-median $\alpha_0 \in E$. It is important to know how far the L_1-median of X is stable relative to small perturbations of the law of X. From a Statistical viewpoint, one wants to find convergent estimators for the L_1-median α_0 of X, based on a sequence of copies of X. The main advantage of the L_1-median is that it provides a robust estimator (see Kemperman (1987)). When $E = I\!R^d$, many authors proved strong consistency of a natural estimator of α_0. See for instance Haldane (1948), Kemperman (1987), ... (see also Berlinet et al. (1998) and Cadre and Gannoun (2000) in a conditional context, or Cadre (2000)). Gannoun et al. (1999) construct a robust nonparametric predictor via the conditional L_1-median. In an infinite dimensional case, see Bru and Heinich (1985) and Cadre-Menneteau (2000) (the latter used the L_1-median in order to estimate, in a robust way, the autocorrelation operator of the Banach AR(1) model). In this paper, we consider the general case, and we give two approximating sequences of α_0, for the weak* topology induced by F.

The remainder of the paper is organized as follows. The second part is dedicated to fixing notations, hypotheses and the presentation of the main results (Theorem 1). Then, we will set out the proofs in the third part. In the last part, we will consider the case of some Hilbert spaces. More precisely, the case $E = L_2(I\!R)$ is considered. The case E an auto-reproducing space is also considered, and it will be derived from it the existence of L_1-medians (in an enlarged definition) on some prehilbertian space of signed and bounded measures.

2. Notations, hypotheses and main results

We will assign to $E = F^*$ the so-called weak* topology induced by F, i.e. the coarsest topology on E such that

$$\begin{array}{ll} E & \to I\!R \\ e & \mapsto < e, f > . \end{array}$$

is continuous, for each $f \in F$ (see Dunford and Schwartz (1958)). If $e \in E$ and $r > 0$, $B(e, r)$ will denote the closed ball

$$B(e, r) = \{x \in E \,:\, \|x - e\| \leq r\}.$$

Recall that the ball $B(e, r)$ is compact for the weak* topology and, because F is separable, the weak* topology on $B(e, r)$ is also metrizable. Finally, if $\{e_n, n \geq 1\}$ is a sequence of elements of E, we will write $e_n \to^* e$ if the sequence $\{e_n, n \geq 1\}$ weak* converges to e i.e. if $< e_n, f > \to < e, f >$ for all $f \in F$.
The task is now to define the approximating sequences of α_0. Let μ be the law of X, and $\{\mu_n, n \geq 1\}$ be a sequence of probability measures on E, which weakly converges to the law of X (for the strong topology on E). For all $n \geq 1$, let φ_n be defined by

$$\varphi_n(\alpha) = \int (\|x - \alpha\| - \|x\|)\mu_n(dx), \ \alpha \in E.$$

Definition of the first approximating sequence
The space E is the dual of a separable Banach space. According to Kemperman ((1987), Theorem 3.6), $\text{Argmin}_E \varphi_n \neq \emptyset$ for all $n \geq 1$. For all $n \geq 1$, we choose $\beta_n \in \text{Argmin}_E \varphi_n$. The sequence $\{\beta_n, n \geq 1\}$ is a natural approximating sequence of α_0.

Definition of the second approximating sequence

Let $\{X_i, i \geq 1\}$ be a sequence of independent copies of X, defined on $(\Omega, \mathcal{F}, P)$. For all $n \geq 1$ and $\omega \in \Omega$, let

$$T_n(\omega) \in \{k \in \{1, \cdots, n\} \, : \, \varphi_n(X_k(\omega)) = \min_{i=1,\cdots,n} \varphi_n(X_i(\omega))\}.$$

The sequence $\{X_{T_n}, n \geq 1\}$ will be proved to be another approximating sequence of α_0.

Theorem 1 *We have:*
i) $\beta_n \to^* \alpha_0$ as $n \to \infty$;
ii) with probability 1 (w.p. 1), $X_{T_n} \to^* \alpha_0$ as $n \to \infty$ if and only if for all $\varepsilon > 0$, $P(X \in B(\alpha_0, \varepsilon)) > 0$.

Remark 1 From a Statistical viewpoint, one considers the case where μ_n is the empirical measure induced by the sample $X_1, \cdots, X_n$, i.e.

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i},$$

where δ_e denotes the Dirac measure on $e \in E$. For, then $\{\mu_n, n \geq 1\}$ converges in law to X w.p. 1 (Dudley (1989), Theorem 11.4.1).

Remark 2 Assume that E is a separable reflexive Banach space. For, then $F = E^*$ and F is separable. Hence the weak* topology induced by F is the weak topology on E. This will be used in Section 4, when $E = L_2(\mathbb{R})$ or E an auto-reproducing space.

Remark 3 Let $E = \mathbb{R}^d$ so that the weak* topology is the strong topology on $\mathbb{R}^d$. Assume that $\{\mu_n, n \geq 1\}$ is the sequence defined in Remark 1. Then i) is the result of Kemperman ((1987), Theorem 2.29). Moreover, ii) gives a new estimator for the L_1-median of X.

Remark 4 The advantage of using the approximating sequence $\{X_{T_n}, n \geq 1\}$ lies in the fact that its values are taken among the n values of the n-sample $X_1, \cdots, X_n$. Moreover, computing $\{X_{T_n}, n \geq 1\}$ is easier than computing $\{\beta_n, n \geq 1\}$.

Remark 5 It is easy to understand the condition of ii). Indeed, if there exists $\varepsilon > 0$ such that w.p. 1 $X \notin B_{\mathcal{B}}(\mu, \varepsilon)$, then by construction, $\{X_{T_n}, n \geq 1\}$ can not converge to α_0 !

Remark 6 Assume that the law of X is

$$\sum_{k=1}^{p} a_k \, \delta_{b_k},$$

where, for all $k = 1, \cdots, p$, $a_k > 0$, $b_k \in E$ and $\sum_{k=1}^{p} a_k = 1$. Then w.p. 1, $\{X_{T_n}, n \geq 1\}$ weak* converges to α_0 if and only if there exists k such that $b_k = \alpha_0$.

3. Proofs

Lemma 1 below was proved by Berlinet et al. ((1998), Lemma 1) when $E = I\!R^d$ and μ_n is a Nadaraya-Watson type estimator of a conditionnal measure.

Lemma 1 *We have:*

$$\sup_n |\frac{\varphi_n(\alpha)}{\|\alpha\|} - 1| \to 0, \text{ as } \|\alpha\| \to \infty.$$

Proof For all $p \geq 1$ and for all $\|\alpha\|$ large enough,

$$\sup_n |\frac{\varphi_n(\alpha)}{\|\alpha\|} - 1| \leq 2\frac{p}{\|\alpha\|} + 2\int_{\{\|x\|>p\}} \mu_n(dx).$$

Hence, for all $p \geq 1$,

$$\limsup_{\|\alpha\|\to\infty} \sup_n |\frac{\varphi_n(\alpha)}{\|\alpha\|} - 1| \leq 2\sup_n \int_{\{\|x\|>p\}} \mu_n(dx). \quad (1)$$

Let $\mathcal{C} = \{r > 0 : B(0, r)$ is a μ-continuity set$\}$ and let $\{p_k, k \geq 1\}$ be an increasing sequence of elements in $\mathcal{C}$. If $n, k \geq 1$, let

$$I_n(k) = \int_{\{\|x\|>p_k\}} \mu_n(dx) \text{ and } I(k) = \int_{\{\|x\|>p_k\}} \mu(dx).$$

Then, $I_n(k) \to I(k)$ if $n \to \infty$ and for all $n \geq 1$, the sequences $I_n(.)$ and $I(.)$ are non-increasing. It is then a classical exercise to prove that

$$\sup_k |I_n(k) - I(k)| \to_{n\to\infty} 0.$$

Let $\varepsilon > 0$. There exists $N \geq 1$ such that for all $n \geq N$, $k \geq 1$: $I_n(k) \leq I(k) + \varepsilon$. Hence, for all $k \geq 1$,

$$\sup_n I_n(k) \leq \sup_{n<N} I_n(k) + \sup_{n \geq N} I_n(k) \leq \sup_{n<N} I_n(k) + I(k) + \varepsilon.$$

By monotone convergence, we deduce that for all $\varepsilon > 0$,

$$\limsup_k \sup_n I_n(k) \leq \varepsilon,$$

so that the right-hand side in (1) vanishes as $p \to \infty$. □

Lemma 2 *We have:*
i) $\sup_n \|\beta_n\| < \infty$ *and* $\lim \varphi(\beta_n) = \varphi(\alpha_0)$*;*
ii) for P*-a.e.* $\omega \in \Omega$, $\sup_n \|X_{T_n}(\omega)\| < \infty$ *and* $\limsup_n \varphi(X_{T_n}(\omega)) = \varphi(\alpha_0)$ *if for all* $\varepsilon > 0$, $P(X \in B(\alpha_0, \varepsilon)) > 0$.

Proof i) Let $r_1 = \sup_n \|\beta_n\|$. Since $\beta_n \in \text{Argmin}_E \varphi_n$ for all $n \geq 1$, Lemma 1 shows that $r_1 < \infty$. For all $n \geq 1$:

$$\begin{aligned} |\varphi_n(\beta_n) - \varphi(\beta_n)| &= |\int (\|x - \beta_n\| - \|x\|)(\mu_n - \mu)(dx)| \\ &= 3\max(1, \|\beta_n\|) |\int \frac{\|x - \beta_n\| - \|x\|}{3\max(1, \|\beta_n\|)} (\mu_n - \mu)(dx)| \\ &\leq 3\max(1, r_1)\pi(\mu_n, \mu), \end{aligned}$$

if $\pi(.,.)$ denotes the BL-metric (see Dudley (1989), Chapter 11). Hence, by assumption:

$$\varphi_n(\beta_n) - \varphi(\beta_n) \to 0 \text{ as } n \to \infty. \quad (2)$$

Moreover, for all $n \geq 1$, $\varphi_n(\beta_n) = \inf_E \varphi_n \leq \varphi_n(\alpha_0)$, so that:

$$\limsup_n \varphi_n(\beta_n) \leq \varphi(\alpha_0).$$

By the above and (2):

$$\limsup_n \varphi(\beta_n) \leq \varphi(\alpha_0),$$

hence i), since $\varphi(\beta_n) \geq \varphi(\alpha_0)$ for all $n \geq 1$.
ii) According to Lemma 1, there exists $A > 0$ such that for all $\|\alpha\| \geq A$ and for all $n \geq 1$, $\varphi_n(\alpha) \geq \|\alpha\|/2$. Let $r_2 = \max(A, 2\|X_1\|)$ and assume that for some $p \geq 1$ and $\omega \in \Omega$, $\|X_{T_p}(\omega)\| > r_2(\omega)$. On one hand, we then we have

$$\varphi_p(X_{T_p}(\omega)) \geq \frac{1}{2}\|X_{T_p}(\omega)\| > \frac{1}{2} r_2(\omega),$$

and on the other hand, since $\varphi_p(\alpha) \leq \|\alpha\|$ for all $\alpha \in E$:

$$\varphi_p(X_{T_p}(\omega)) = \min_{i=1,\cdots,n} \varphi_p(X_i(\omega)) \leq \varphi_p(X_1(\omega)) \leq \|X_1(\omega)\|.$$

Consequently $\|X_1(\omega)\| > r_2(\omega)$, which is impossible. It follows that w.p. 1

$$\sup_n \|X_{T_n}\| \leq r_2 < \infty.$$

As in the proof of (2), one can now prove that w.p. 1:

$$\varphi_n(X_{T_n}) - \varphi(X_{T_n}) \to 0 \text{ as } n \to \infty. \quad (3)$$

Moreover, for all $n \geq 1$ and $k = 1, \cdots, n$: $\varphi_n(X_{T_n}) \leq \varphi_n(X_k)$. By assumption, we then have w.p. 1 : $\limsup_n \varphi_n(X_{T_n}) \leq \varphi(X_k)$, for all $k \geq 1$. From the above and (3) it follows that w.p. 1:

$$\limsup_n \varphi(X_{T_n}) \leq \min_{k=1,\cdots,p} \varphi(X_k), \ \forall p \geq 1. \quad (4)$$

The task is now to find the limit of the rightmost term in (4). For $\varepsilon > 0$ and $p \geq 1$, let:

$$N(p,\varepsilon) = [\min_{k=1,\cdots,p} \varphi(X_k) \geq \varphi(\alpha_0) + 2\varepsilon, \ \exists j \in \{1,\cdots,p\} \, : \, X_j \in B(\alpha_0,\varepsilon)].$$

Then, since for all $\alpha_1, \alpha_2 \in E$: $|\varphi(\alpha_1) - \varphi(\alpha_2)| \leq \|\alpha_1 - \alpha_2\|$, we have:

$$\begin{aligned} N(p,\varepsilon) &\subset [\varphi(X_j) \geq \varphi(\alpha_0) + 2\varepsilon, \ \exists j \in \{1,\cdots,p\} \, : \, X_j \in B(\alpha_0,\varepsilon)] \\ &\subset [\varepsilon + \varphi(\alpha_0) \geq \varphi(\alpha_0) + 2\varepsilon] = \emptyset. \end{aligned}$$

By independence of the X_i's we deduce that

$$\begin{aligned} P(|\min_{k=1,\cdots,p} \varphi(X_k) - \varphi(\alpha_0)| \geq 2\varepsilon) &\leq P(E(p,\varepsilon)) + P(X \notin B(\alpha_0,\varepsilon))^p \\ &\leq (1 - P(X \in B(\alpha_0,\varepsilon))^p. \end{aligned}$$

Applying the Borel-Cantelli Lemma (see also the assumption in ii)), we conclude that w.p. 1:

$$\min_{k=1,\cdots,p} \varphi(X_k) \to_{p\to\infty} \varphi(\alpha_0).$$

According to (4) and the above result, we have Assertion ii). □

Proof of Theorem 1 We only prove ii). Assume that there exists $\varepsilon > 0$ such that $P(X \in B(\alpha_0, \varepsilon)) = 0$. Since $X_{T_n} \in \{X_1, \cdots, X_n\}$ for all $n \geq 1$, $\{X_{T_n}, n \geq 1\}$ can not converge to α_0.
Assume now that for all $\varepsilon > 0$, $P(X \in B(\alpha_0, \varepsilon)) > 0$. Let

$$\Omega_1 = [\sup_n \|X_{T_n}\| < \infty \text{ and } \lim_n \varphi(X_{T_n}) = \varphi(\alpha_0)].$$

Lemma 2 ii) shows that $P(\Omega_1) = 1$. Fix $\omega \in \Omega_1$ and let us denote by r the real number $r = \sup_n \|X_{T_n}(\omega)\|$. The sequence $\{X_{T_n}(\omega), n \geq 1\}$ is in the ball $B(0, r)$, which is compact and metrizable for the weak* topology (see the beginning of Section 2). Hence from every subsequence, one can extract a subsequence $\{n_k, k \geq 1\}$ and $\gamma \in B(0, r)$ such that

$$X_{T_{n_k}}(\omega) \to^* \gamma, \text{ as } k \to \infty.$$

Since for all $e \in E$, $\|e\| = \sup_{f \in F, \|f\| \leq 1} | < e, f > |$, the function $e \mapsto \|e\|$ defined on E is weak* lower semi continuous. Therefore, for all $e \in E$:

$$\|e - \gamma\| - \|e\| \leq \liminf_k (\|e - X_{T_{n_k}}(\omega)\| - \|e\|).$$

Using Fatou, one can conclude that

$$\varphi(\gamma) \leq \liminf_k \varphi(X_{T_{n_k}}(\omega)).$$

By the very definition of Ω_1, $\varphi(\gamma) \leq \varphi(\alpha_0)$. Since α_0 is the unique L_1-median of X, this clearly forces $\gamma = \alpha_0$. We conclude that from every subsequence of $\{X_{T_n}(\omega), n \geq 1\}$, one can extract a subsequence which weakly* converges to α_0, hence that $X_{T_n}(\omega) \to^* \alpha_0$ as $n \to \infty$, and finally that w.p. 1 $X_{T_n} \to^* \alpha_0$ as $n \to \infty$ because $P(\Omega_1) = 1$. □

4. Applications

4.1 The case $E = L_2(I\!R)$: strong convergence to the L_1-median

In this Section we assume that $E = L_2(I\!R)$ and for all $\varepsilon > 0$, $P(X \in B(\alpha_0, \varepsilon)) > 0$. Theorem 1 ii) shows that w.p. 1, the sequence $\{X_{T_n}, n \geq 1\}$ converges to α_0 for the weak topology. One wants to improve this result : we look for conditions so that this convergence holds for the strong topology.

Corollary 1 *Let $A \subset I\!R$ be a bounded set. Assume that for P-a.e. $\omega \in \Omega$, the support of $X(\omega)$ is contained in A and $X(\omega)$ is absolutely continuous,* *with derivative $X'(\omega)$. Furthermore, assume that there exists $c > 0$ such that w.p. 1 $\|X'\| \leq c$. Then w.p. 1 $\|X_{T_n} - \alpha_0\| \to 0$ as $n \to \infty$.*

Proof Let $\widehat{f}$ be the Fourier transform of $f \in L_2(I\!\!R)$. By assumption, for all $n \geq 1$, $X_{T_n}(t) = 0$ if $t \notin A$. Note that for all $\gamma_1, \gamma_2 \in I\!\!R$ and $t \in A$, we have

$$|\exp(i\gamma_1 t) - \exp(i\gamma_2 t)| \leq |\gamma_1 - \gamma_2| \sup_{t \in A} |t|.$$

Then, setting $c_1 = \sup_{t \in A} |t|$, one has for all $n \geq 1$ and $\gamma_1, \gamma_2 \in I\!\!R$:

$$\begin{aligned}
|\widehat{X_{T_n}}(\gamma_1) - \widehat{X_{T_n}}(\gamma_2)| &= |\int_A (\exp(i\gamma_1 t) - \exp(i\gamma_2 t)) X_{T_n}(t) dt| \\
&\leq c_1 |\gamma_1 - \gamma_2| \int_A |X_{T_n}(t)| dt \\
&\leq c_1 \, \lambda(A)^{1/2} \sup_n \|X_{T_n}\| \, |\gamma_1 - \gamma_2|,
\end{aligned}$$

where λ denotes the Lebesgue measure. Therefore the sequence $\{\widehat{X_{T_n}}, n \geq 1\}$ is equicontinous w.p. 1 by Lemma 2 ii). Let $K \subset I\!\!R$ be a compact set and let $C(K)$ be the set of all real continuous functions defined on K. By Ascoli and the above it follows that w.p. 1, the sequence $\{\widehat{X_{T_n}}, n \geq 1\}$ is relatively compact in $C(K)$. But by Theorem 1 ii), $\{\widehat{X_{T_n}}, n \geq 1\}$ weakly converges to $\widehat{\alpha_0}$ w.p. 1. From what has already been proved, we conclude that w.p. 1:

$$\|(\widehat{X_{T_n}} - \widehat{\alpha_0}) I_K\| \to 0, \text{ as } n \to \infty. \quad (5)$$

If $p \geq 1$, let $K_p = [-p, p]$. For all $n, p \geq 1$:

$$\begin{aligned}
\|\widehat{X_{T_n}} - \widehat{\alpha_0}\| &\leq \|(\widehat{X_{T_n}} - \widehat{\alpha_0}) I_{K_p}\| + \|(\widehat{X_{T_n}} - \widehat{\alpha_0}) I_{K_p^c}\| \\
&\leq \|(\widehat{X_{T_n}} - \widehat{\alpha_0}) I_{K_p}\| + \sup_n \|\widehat{X_{T_n}} I_{K_p^c}\| + \|\widehat{\alpha_0} I_{K_p^c}\|. \quad (6)
\end{aligned}$$

By assumptions, using the well-known properties of the Fourier transform:

$$\begin{aligned}
\sup_n \|\widehat{X_{T_n}} I_{K_p^c}\|^2 &\leq \sup_n \int_{K_p^c} |\widehat{X_n}(\gamma)|^2 d\gamma = \sup_n \int_{K_p^c} \frac{1}{\gamma^2} |\widehat{X'_n}(\gamma)|^2 d\gamma \\
&\leq \frac{1}{p^2} \sup_n \int_{K_p^c} |\widehat{X'_n}(\gamma)|^2 d\gamma \leq \frac{1}{p^2} \sup_n \int |X'_n(\gamma)|^2 d\gamma \leq (\frac{c}{p})^2.
\end{aligned}$$

From the above and (5), (6), we have w.p. 1 and for all $p \geq 1$:

$$\limsup_n \|\widehat{X_{T_n}} - \widehat{\alpha_0}\| \leq \frac{c}{p} + \|\widehat{\alpha_0} I_{K_p^c}\|.$$

Since $\widehat{\alpha_0} \in L_2(I\!R)$, the rightmost term vanishes as $p \to \infty$. Corollary 1 is then a straightforward consequence of the Plancherel Formula. $\square$

4.2 The case of random measures and auto-reproducing spaces

Sometimes, it may happens that E is not a Banach space. Then, the existence result of Kemperman (1987) does not hold. However, in particular situations, one may use Theorem 1 to derive the existence of the L_1-median. The following is from Suquet ((1990) and (1993)). Let χ be a metric space, and let $\mathcal{M}$ be the set of all signed bounded measures on χ. Assume that there exists a sequence of real, measurable and bounded functions $\{f_i, i \geq 1\}$ on χ, which characterizes the signed measures as follows:

$$\forall i \geq 1, \int_\chi f_i d\mu = 0 \Longrightarrow \mu = 0,$$

where $\mu \in \mathcal{M}$. Furthermore, assume that $\sum_{i\geq 1} \sup_\chi |f_i|^2 < \infty$, and let K be defined by

$$K(x,y) = \sum_{i\geq 1} f_i(x) f_i(y), \ x, y \in \chi.$$

Then K is a bounded, measurable reproducing kernel on χ^2. Let $H_K(= E)$ be the auto-reproducing functional space associated with K. The formula

$$< \mu, \nu >_K = \int_{\chi^2} K d\mu \otimes \nu, \ \mu, \nu \in \mathcal{M}$$

defines a scalar product on $\mathcal{M}$. Moreover, the application

$$\psi : \begin{array}{ll} \mathcal{M} & \to H_K \\ \mu & \mapsto \int_\chi K(., y) d\mu(y) \end{array}$$

is an isometric injection from $(\mathcal{M}, \|.\|_K)$ to the Hilbert space H_K. Finally, recall that $\psi(\mathcal{M})$ is dense in H_K.

Let Y be a random variable on $(\Omega, \mathcal{F}, P)$, taking values in $\overline{\mathcal{M}}$ (where the closure holds for the norm $\|.\|_K$) and such that $P(Y \in \mathcal{M}) = 1$ (i.e. Y is a random measure). Let $\overline{\varphi}$ be the function defined on $\mathcal{M}$ by

$$\overline{\varphi}(\mu) = E[\|Y - \mu\|_K - \|Y\|_K], \ \mu \in \mathcal{M}.$$

Though $(\mathcal{M}, \|.\|_K)$ is not a Banach space, an L_1-median of Y is again defined to be any point in $\mathcal{M}$ which minimizes $\overline{\varphi}$. This Section deals with the

existence of L_1-medians of Y, in the prehilbertian space $(\mathcal{M}, \|.\|_K)$. But $\mathcal{M}$ is not complete, hence this can not be a consequence of the general results by Kemperman (1987).
In the following, χ will be assumed to be separable or locally compact. Suquet (1993) proved that $\psi(\mathcal{M})$ is then a borel set in H_K. Let $X = \psi(Y)$. The Hilbertian random variable X satisfies $P(X \in \psi(\mathcal{M})) = 1$. As usual, φ denotes the function defined by

$$\varphi(\alpha) = E[\|X - \alpha\| - \|X\|], \ \alpha \in H_K.$$

(Here, $\|.\|$ is the norm on H_K.)
Assume that the law of Y is not carried by any straight line. But, ψ is linear and $P(X \in \psi(\mathcal{M})) = 1$. Therefore the law of X is not carried by any straight line. Since H_K is a strictly convex Hilbert space (hence reflexive), it follows from Valadier (1984) and Kemperman ((1987), Theorem 2.17) that X possesses only one L_1-median $\alpha_0 \in H_K$, i.e. $\text{Argmin}_{H_K}\varphi = \{\alpha_0\}$. We are looking for conditions which imply that $\alpha_0 \in \psi(\mathcal{M})$, hence the existence of the L_1-median of Y.
Let $\{Y_i, i \geq 1\}$ be independent copies of Y defined on $(\Omega, \mathcal{F}, P)$ and let $X_i = \psi(Y_i)$ for all $i \geq 1$. Following the notations of Section 1, define for all $n \geq 1$:

$$\varphi_n(\alpha) = \frac{1}{n}\sum_{i=1}^{n}(\|X_i - \alpha\| - \|X_i\|), \ \alpha \in H_K;$$
$$T_n \in \{k \in \{1, \cdots, n\} \, : \, \varphi_n(X_k) = \min_{i=1,\cdots,n} \varphi_n(X_i)\}.$$

Recall that $\psi(\mathcal{M})$ is dense in H_K, so that one can find a sequence $\{m_n, n \geq 1\}$ in $\mathcal{M}$ such that $\|\psi(m_n) - \alpha_0\| \to 0$ as $n \to \infty$.

Corollary 2 *Assume that for all $\varepsilon > 0$, $n \geq 1$: $P(\|Y - m_n\|_K \leq \varepsilon) > 0$. Then Y possesses only one L_1-median $\nu_0 \in \mathcal{M}$ and moreover w.p. 1 $\|Y_{T_n} - \nu_0\|_K \to 0$ as $n \to \infty$ if one of the following conditions is satisfied:*
i) χ is separable and there exists $Z \in \mathcal{M}$ such that w.p. 1 : $|Y| \ll Z$;
ii) χ is locally compact and there exists $c > 0$ such that w.p. 1 : $|Y|(\chi) \leq c$.

Proof Let $\varepsilon > 0$. For all n large enough, $\|\psi(m_n) - \alpha_0\| \leq \varepsilon/2$. It follows that

$$\|X - \alpha_0\| \leq \|Y - m_n\|_K + \|\psi(m_n) - \alpha_0\| \leq \|Y - m_n\|_K + \frac{\varepsilon}{2},$$

and, in consequence,

$$P(\|Y - m_n\|_K \leq \frac{\varepsilon}{2}) \leq P(\|X - \alpha_0\| \leq \varepsilon).$$

By assumption, this proves that $P(\|X - \alpha_0\| \leq \varepsilon) > 0$. Therefore w.p. 1, $\{X_{T_n}, n \geq 1\}$ weakly converges to α_0, according to Theorem 1 ii).
Assume i) satisfied. Then w.p. 1, for all $n \geq 1$, $|Y_{T_n}| \ll Z$. By Suquet ((1990), Theorem V.2.4), the set $\{X_{T_n}, n \geq 1\}$ is relatively compact w.p. 1, and its closure is contained in $\psi(\mathcal{M})$. Consequently, $\alpha_0 \in \psi(\mathcal{M})$ and if $\nu_0 \in \mathcal{M}$ is such that $\alpha_0 = \psi(\nu_0)$ we have w.p. 1, $\|Y_{T_n} - \nu_0\|_K \to 0$ as $n \to \infty$. Note that ν_0 is unique, since ψ is injective.
Assume ii) satisfied. Then w.p. 1, $\sup_n |Y_{T_n}|(\chi) \leq c$. Following the proof of Suquet ((1993), Theorem V.4), we deduce that the set $\{X_{T_n}, n \geq 1\}$ is relatively compact w.p. 1, and there exists $\nu_0 \in \mathcal{M}$ such that $\alpha_0 = \psi(\nu_0)$. Obviously, we then have w.p. 1, $\|Y_{T_n} - \nu_0\|_K \to 0$ as $n \to \infty$. □

References

A. Averous, M. Meste (1997). Median Balls : An Extension of the Interquantile Intervals to Multivariate Distributions, *J. Multiv. Anal.*, **63**, 222-241.

F.K. Bedall, H. Zimmerman (1979). The mediancentre, *Applied Statistics*, **28**, 325-328.

A. Berlinet, B. Cadre, A. Gannoun (1998). On the conditional L_1-median and its estimation, *Submitted*.

B. Bru, H. Heinich (1985). Meilleures approximations et médianes conditionnelles, *Ann. Inst. Henri Poincaré*, **21**, 197-224.

B. Cadre (2000). Hölder properties of the L_1-median. Rate of convergence in the nonparametric uniform estimation of the conditional L_1-median (mixing case), *Submitted*.

B. Cadre, A. Gannoun (2000). Asymptotic normality of consistent estimate of the conditional L_1-median, *Publ. Inst. Stat. Univ. Paris*, **XXXXIV**, 13-35.

B. Cadre, L. Menneteau (2000). Estimation of the autocorrelation operator for the ARB(1) model, *Submitted*.

R.M. Dudley (1989). Real Analysis and Probability, Wadsworth, Belmont.

N. Dunford, J.T. Schwartz (1958). Linear Operators, part I : General Theory, Interscience Publishers, New-York.

A. Gannoun, B. Cadre, A. Berlinet (1999). L'estimation non paramétrique de la médiane spatiale conditionnelle en vue de la prévision, *Actes du Colloque de la Société Mathématique Tunisienne*, Tabarka 22-24 Mars, 46-55.
J.C. Gower (1974). The mediancentre, *Applied Statistics*, **23**, 466-470.
J.B.S. Haldane (1948). A note on the median of multivariate distribution, *Biometrika*, **73**, 414-415.
J.H.D. Kemperman (1987). The median of finite measure of Banach space, *Statistical data analysis based on the L_1-norm and related methods*, Y. Dodge (Ed.), (North-Holland, Amsterdam), 217-230.
P. Milasevick, G.R. Ducharme (1987). Uniqueness of spatial median, *Ann. Statist.*, **15**, 1332-1333.
C. Suquet (1990). Une topologie préhilbertienne sur l'espace des mesures á signe bornées, *Publ. Inst. Stat. Univ. Paris*, **XXXV**, 51-77.
C. Suquet (1993). Convergences stochastiques des suites de mesures aléatoires signées considérées comme variables aléatoires hilbertiennes, *Publ. Inst. Stat. Univ. Paris*, **XXXVII**, 71-99.
M. Valadier (1984). La multi-application médianes conditionnelles, *Z. W.*, **67**, 279-282.