ASYMPTOTIC BEHAVIOR OF $p$-PREDICTIONS FOR VECTOR VALUED RANDOM VARIABLES

JUAN A. CUESTA AND CARLOS MATRÁN

ABSTRACT. Let $(\Omega, \sigma, \mu)$ be a probability space and let $X$ be a $B$-valued \( \mu \)-essentially bounded random variable, where $(B, \| \|)$ is a uniformly convex Banach space. Given $\alpha$, a sub-$\sigma$-algebra of $\sigma$, the $p$-prediction (1 < $p$ < $\infty$) of $X$ is defined as the best $L_p$-approximation to $X$ by $\alpha$-measurable random variables.

The paper proves that the Pólya algorithm is successful, i.e. the $p$-prediction converges to an "$\infty$-prediction" as $p \to \infty$. First the proof is given for $p$-means ($p$-predictions given the trivial $\sigma$-algebra), and the general case follows from the characterization of the $p$-prediction in terms of the $p$-mean of the identity in $B$ with respect to a regular conditional probability. Notice that the problem was treated in [7], but the proof is not satisfactory (as pointed out in [4]).

1. Introduction. Throughout this paper $(\Omega, \sigma, \mu)$ denotes a probability space, $(B, \| \|)$ is a uniformly convex Banach space, and $L_p(\sigma) = L_p(\Omega, \sigma, \mu, B)$, 1 < $p$ < $\infty$, represents the abstract Lebesgue-Bochner $L_p$-space. If $\alpha$ is a sub-$\sigma$-algebra of $\sigma$, $L_p(\alpha)$ denotes the (closed) subspace of $L_p(\sigma)$ consisting in all the equivalence classes in $L_p(\sigma)$ containing an $\alpha$-measurable function. In this notation we will not make any distinction between a random variable and the equivalence class it represents. Recall that random variables in $L_p(\alpha)$ are strongly $\alpha$-measurable, i.e. they are a.s. limits of finite valued $\alpha$-measurable random variables.

In [1] Ando and Amemiya have introduced the $p$-predictions given a $\sigma$-algebra for real valued random variables. In an analogous way, taking into account that $L_p(\sigma)$, 1 < $p$ < $\infty$, is uniformly convex, we may consider the $p$-prediction of a variable $X \in L_p(\sigma)$ given the sub-$\sigma$-algebra $\alpha$ as the (unique) best $L_p$-approximation to $X$ by elements of $L_p(\alpha)$. Therefore the $p$-prediction will be continuous in $L_p(\sigma)$. However there exist important differences between the real and the abstract cases. For instance, it is well known that if $B = R$ and $p = 2$, the 2-prediction given $\alpha$ coincides with the conditional mean given $\alpha$, while this is not true for general uniformly convex spaces. In fact the conditional mean is always linear (see Diestel and Uhl [8, p. 122]) but the 2-prediction is not linear unless $B$ is a Hilbert space or $\Omega$ is the union of two $\mu$-atoms (Herrndorf [10]).

This paper deals with the study of the limit of $p$-predictions as $p \to \infty$. In the remainder of this work we consider a fixed $\mu$-essentially bounded random variable $X$ (i.e. $X \in L_{\infty}(\sigma)$) and we will prove that the $p$-prediction of $X$ given the (fixed) sub-$\sigma$-algebra $\alpha$ converges to a best $L_{\infty}$-approximation of $X$ by elements of $L_{\infty}(\alpha)$ (or $\infty$-prediction of $X$ given $\alpha$).

For real valued random variables this result was proved in [3]. Notice that the study of the convergence as $p \to \infty$ of $p$-predictions on uniformly convex spaces

Received by the editors June 1, 1985.

1980 Mathematics Subject Classification (1985 Revision). Primary 60G30, 28B05, 60B12.

©1987 American Mathematical Society

0002-9939/87 $1.00 + $.25 per page

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
was carried out in [7], but the proof of the existence of the limit is not satisfactory as pointed out in [4]. (Also in connection with the Pólya algorithm see [5].) Our proof consists of two stages. First we prove that the p-prediction, $1 < p < \infty$, given $\alpha$ can be obtained as the p-prediction (given the trivial $\sigma$-algebra) with respect to a regular conditional probability. For notational convenience, the p-prediction given the trivial $\sigma$-algebra or best $L_p$-approximation by constants will be called p-mean, in connection with the real case. At a second stage we will prove that the p-mean converges, as $p \to \infty$, to the Chebysev center of the probabilistic support of $X$ (the Midrange in the real case).

2. $p$-predictions and regular conditional probabilities. This section is devoted to establish the relation between the p-prediction of $X$ and the conditional probability distribution of $X$ given $\alpha$.

Given any metric space $(E, d)$, $\beta_E$ denotes the Borel $\sigma$-algebra on $E$. The weak convergence of measures will be denoted by $\overset{w}{\to}$.

Recall that an $\alpha$-measurable function is strongly $\alpha$-measurable iff its image is a.s. contained in a separable subset of $B$.

Let $P_X$ denote the probability measure induced by $X$ on $(B, \beta_B)$. There exists a unique smaller closed set in $B$ of $P_X$-probability one. Moreover this set, which we denote by $S(X)$, is separable (it is called the support of the probability $P_X$). Therefore the set $\Gamma_{S(X)}$ of all probability measures defined on $\beta_{S(X)}$ is separable and metrizable with the topology associated to the weak convergence of measures (see Parthasarathy [11, p. 43]).

Now, considering the sub-$\sigma$-algebra $\alpha$ of $\sigma$, since $S(X)$ is separable, there exists (see Ash [2, p. 265]) a function $Q_\alpha : \Omega \times \beta_{S(X)} \to R$ such that:

(a) For every $w \in \Omega$, $Q_\alpha(w, \cdot)$ is a probability measure on $\beta_{S(X)}$ (hence we can also consider it as a probability on $\beta_B$).

(b) For every $A \in \beta_{S(X)}$, $Q_\alpha(\cdot, A)$ is a version of $\mu[X \in A/\alpha]$, the conditional distribution of $X$ given $\alpha$.

Also, since $X$ is $\mu$-essentially bounded, we can choose $Q_\alpha$ verifying

(c) For every $w \in \Omega$ the identity map on $B$ is $Q_\alpha(w, \cdot)$-essentially bounded.

$Q_\alpha$ will be called a regular conditional probability (R.C.P.) of $X$ given $\alpha$, and we have

**Proposition 2.1.** The mapping $T : \Omega \to \Gamma_{S(X)}$ defined by $T(w) = Q_\alpha(w, \cdot)$ is strongly $\alpha$-measurable (we consider on $\Gamma_{S(X)}$ the Borel $\sigma$-algebra associated with the weak convergence of measures).

**Proof.** Let $F$ be the set of all continuous bounded real valued functions defined on $B$. Then the family of subsets of $\Gamma_{S(X)}$,

$$\left\{ Q \in \Gamma_{S(X)}, \left| \int f \, dQ - \int f \, dP \right| < \delta \right\}, \quad \delta > 0, \quad f \in F, \quad P \in \Gamma_{S(X)},$$

is a subbase of the topology on $\Gamma_{S(X)}$, and, since $\Gamma_{S(X)}$ is separable, it suffices to show that for $\delta > 0$, $f \in F$, and $P \in \Gamma_{S(X)}$:

$$\left\{ w/ \left| \int f(t)Q_\alpha(w, dt) - \int f(t)P(dt) \right| < \delta \right\} \in \alpha.$$

But this is obvious because the map $w \to \int f(t)Q_\alpha(w, dt)$ is $\alpha$-measurable. $\square$
In the next proposition we use the Skorohod representation theorem for weak convergence of probability measures (see Skorohod [12]), stated for convenience in the following way.

"Let $P_n, \ n = 0, 1, 2, \ldots$, belong to $\Gamma_S(X)$ and $P_n \xrightarrow{w} P_0$. Then there exist $S(X)$-valued random variables $Y_0, Y_1, Y_2, \ldots$ defined on an appropriate probability space $(W, \mathcal{F}, L)$ such that:

(i) $Y_n, \ n = 0, 1, 2, \ldots$, induces on $(S(X), \mathcal{B}_S(X))$ the probability $P_n$.

(ii) $Y_n \rightarrow Y_0 \text{ L-a.s.}"

Now, let $Q \in \Gamma_S(X)$. Define $H_p(Q), 1 < p < \infty$, to be the $p$-mean, computed in $L_p(B, \mathcal{B}_B, Q, B)$, of the identity on $B$. Then

**PROPOSITION 2.2.** $H_p: \Gamma_S(X) \rightarrow B$ is a continuous map.

**PROOF.** Let $Q_n \in \Gamma_S(X), \ n = 0, 1, 2, \ldots$, such that $Q_n \xrightarrow{w} Q_0$ (recall that $\Gamma_S(X)$ is metrizable) and let $Y_n, \ n = 0, 1, 2, \ldots$, be the $S(X)$-valued random variables obtained from Skorohod's theorem. Also observe that the $p$-mean of a variable depends only on the probability induced by the variable (the law of the random variable).

But, as $S(X)$ is bounded, $Y_n \rightarrow Y_0$ also in $L_p$ (a distinct $L_p$), so the result is obtained from the continuity of the $p$-mean (recall that $L_p(W, \mathcal{F}, L, B)$ is also uniformly convex). \[\square\]

**COROLLARY 2.3.** The map $M_p: \Omega \rightarrow B$ defined by $M_p(w) = H_p(Q_\alpha(w, \cdot))$ is strongly $\alpha$-measurable.

**PROOF.** By Propositions 2.1 and 2.2 $M_p$ is $\alpha$-measurable, but moreover $M_p(\Omega) \subset H_p(\Gamma_S(X))$ which is separable as the image of a separable set by a continuous function. \[\square\]

The fundamental theorem in this section is based in the following result.

**LEMMA 2.4.** Let $X$ be a $B$-valued random variable on $(\Omega, \sigma, \mu)$ and let $H: B \times \Omega \rightarrow R$ be $\mathcal{B}_B \times \alpha$-measurable such that $H(X, \text{Id}) \in L_1(\Omega, \sigma, \mu, R)$. Then $\int H(t, \cdot) Q_\alpha(\cdot, dt)$ is a version of the conditional mean of $H(X, \text{Id})$ given the sub-$\sigma$-algebra $\alpha$ of $\sigma$.

**PROOF.** A standard reasoning proves the statement (begin with $H = I_{D \times A}$, where $D \in \mathcal{B}_B, A \in \alpha$). \[\square\]

**THEOREM 2.5.** The random variable $M_p$ defined in Corollary 2.3 is a version of the $p$-prediction of $X$ given $\alpha$ (i.e. the $p$-prediction of $X$ given $\alpha$ coincides with the $p$-mean of the identity with respect to $Q_\alpha$).

**PROOF.** Let $g \in L_p(\alpha)$. By Lemma 2.4,

$$
\|X - g\|_p^p = \int \left[ \int \|t - g(w)\|^p Q_\alpha(w, dt) \right] \mu(dw)
\geq \int \left[ \int \|t - M_p(w)\|^p Q_\alpha(w, dt) \right] \mu(dw) = \|X - M_p\|_p^p;
$$

hence Corollary 2.3 and the uniqueness of $p$-predictions prove the result. \[\square\]
3. **The Pólya algorithm for p-predictions (vector valued case).** The existence and uniqueness of the best $L_\infty$-approximation by constants in uniformly convex spaces is well known (see Garkavi [9] or Singer [13]). In fact, the Chebyshev center of $S(X)$ (which we denote by $\pi_\infty$) is the best $L_\infty$-approximation to $X$. In this section we will prove that the $p_n$-predictions given $\alpha$ converge on an $\infty$-prediction given $\alpha^n$ as $p_n \to \infty$. Hence the Pólya algorithm holds for these $L_p$-approximations.

Some additional notation will be employed: $\pi_p(X/\alpha)$ is the $p$-prediction of $X$ given $\alpha$, $\pi_p$ is the $p$-mean of $X$ ($1 < p < \infty$), and $V_p = \inf\{\|X - h\|_p, h \in B\}$ ($1 < p \leq \infty$). Note that, obviously, $V_\infty \geq V_p$ for all $p$.

We need a previous theorem.

**Theorem 3.1.** If $0 < m < V_\infty$ then there exists a $\delta > 0$ such that for every $h \in B$, $\mu\{w/\|X(w) - h\| \leq V_\infty - m\} \leq 1 - \delta$.

**Proof.** Suppose not. Then there exist $m < V_\infty$ and a sequence $(h_k)_k$ in $B$ such that for each $k$: $\mu\{\|X - h_k\| \leq V_\infty - m\} > 1 - 1/k$.

Moreover, since $S(X)$ is bounded, we can choose the sequence verifying $h_k \to h_0$ weakly in $B$.

Let $t \in S(X)$ and let $\tau < m/2$; the definition of $S(X)$ entails that

$$\mu\{w/\|X(w) - t\| < \tau\} > 0.$$ 

Therefore there exists $n_0$ such that for $k \geq n_0$,

$$\{\|X - h_k\| \leq V_\infty - m\} \cap \{\|X - t\| < \tau\} \neq \emptyset,$$

so $\|h_k - t\| \leq V_\infty - m/2$. But this implies $\|h_0 - t\| \leq \lim_k \|h_k - t\| \leq V_\infty - m/2$ for all $t \in S(X)$, contradicting the definition of $V_\infty$. \square

**Theorem 3.2.** There exists $\pi_\infty(X/\alpha) \in L_\infty(\alpha)$ such that $\pi_{p_n}(X/\alpha) \to \pi_\infty(X/\alpha)$ a.s. as $p_n \to \infty$, and $\|X - \pi_\infty(X/\alpha)\|_\infty \leq \|X - h\|_\infty$ for every $h \in L_\infty(\alpha)$.

**Proof.** From Theorem 2.5 it is obvious that for proving the convergence it suffices to consider the case in which $\alpha$ is the trivial $\sigma$-algebra. We will prove that $\pi_{p_n} \to \pi_\infty$ (the Chebyshev center of $S(X)$).

If not, then there exist $\tau > 0$ and a subsequence, which we denote as the initial, such that $\|\pi_{p_n} - \pi_\infty\| > \tau$ for all $n$.

Let $m > 0$ and suppose that $t$ is in $B$ and verifies $\|t - \pi_\infty\| \leq V_\infty$ and $\|t - \pi_{p_n}\| \leq V_\infty + m$. Then, if we call $\theta$ to the modulus of convexity of $B$,

$$\|t - \frac{1}{2}(\pi_{p_n} + \pi_\infty)\| \leq [1 - \theta(\tau/(V_\infty + m))](V_\infty + m).$$

Take $m_0 > 0$. As $\theta$ is a nondecreasing function, choosing $0 < m < m_0$ small enough, there exists $\tau > 0$ such that

$$(\ast) \quad \|t - \frac{1}{2}(\pi_\infty + \pi_{p_n})\| \leq V_\infty - \tau.$$

Since $\mu\{w/\|X(w) - \pi_\infty\| \leq V_\infty\} = 1$ by definition of $V_\infty$, $(\ast)$ and Theorem 3.1, with $\frac{1}{2}(\pi_\infty + \pi_{p_n})$ as $h$, imply $\mu\{w/\|X(w) - \pi_{p_n}\| \leq V_\infty + m\} \leq 1 - \delta(\tau)$. Therefore $V_{p_n} \geq (V_\infty + m)(\delta(\tau))^{1/p_n}$ which contradicts that $V_\infty \geq V_p$ for all $p$.

Finally, the fact that $\pi_\infty(X/\alpha)$ is an $\infty$-prediction is obtained from Egoroff's theorem with similar techniques to those employed in [3]. \square
REFERENCES

5. ___, On the Pólya algorithm in isotonic approximation, Preprint.