A GEOMETRIC PROOF FOR SOME SIMPLE CASES OF MARTINGALE CONVERGENCE

BRAILEY SIMS

§0 *INTRODUCTION.*

The use of real valued martingales in probability theory has grown steadily since their detailed introduction by Doob in 1950. Vector valued martingales were first studied explicitly in 1960 by Chatterji and independently Scalora, and were rapidly recognised as a powerful tool for the study of certain Banach space structures (see, for example, Diestel and Uhl, 1977).

Besides introducing the idea of a vector valued martingale this article gives a "geometric" proof of the following convergence theorem which makes no use of probability theory and involves a minimum of measure theory.

THEOREM 1. *Let* (Σ_n, f_n) *be a martingale in* $L_p(\Sigma,\mu;X)$ *where* $1 < p < \infty$ *and* X *is a uniformly convex Banach space. Then there exists* $f_\infty \in L_p(\Sigma,\mu;X)$ *with* $\|f_n - f_\infty\|_p \to 0$ *if (and only if)* $\|f_n\|_p \le K$ *for all* n *and some* $K > 0$.

Our proof is particularly simple when $p = 2$ and X is an Hilbert space. Since this case would usually be introduced first, the proof for other cases has been relegated to footnotes and remarks.

The article owes much to the suggestions of my colleagues, in particular, Mr. L. Kavalieris, made during a series of seminars at the University of New England.

§1 *PRELIMINARIES.*

Notations.

Throughout the article, Σ denotes a σ-algebra of subsets of Ω and μ a countably additive probability measure on Σ (that is $\mu(E) \ge 0$ for all $E \in \Sigma$ and $\mu(\Omega) = 1$).

H is a Hilbert space over the real field, with inner product $(\cdot,\cdot)$, and X is a real Banach space. We say X *is uniformly convex* if: every sequence (x_n) for which $\|x_n\|$ and $\left\|\frac{x_n + x_m}{2}\right\|$ have a common limit as $n,m \to \infty$, is a Cauchy sequence (and hence convergent).[1] It is an easy consequence of the parallelogram rule that H is uniformly convex $\left(\left\|\frac{x_n - x_m}{2}\right\|^2 = \tfrac{1}{2}\|x_n\|^2 + \tfrac{1}{2}\|x_m\|^2 - \left\|\frac{x_n + x_m}{2}\right\|^2\right)$.

Bochner integration for vector valued functions.[2]

The reader interested only in real valued martingales could pass over this material provided $\mathbb{R}$ (the real numbers) is everywhere substituted for H or X and integrals

---

(1) This form of the definition is most suited for our purpose. The equivalence with Clarkson's original definition being a simple matter.

(2) For details see either Dunford and Schwartz, 1958 or Diestel and Uhl, 1977.

are interpreted in the sense of Lebesgue.

A function $f\colon \Omega \to X$ is μ-*measurable* if there exists a sequence of simple functions (S_n) with $\|S_n(\omega) - f(\omega)\| \to 0$ for μ-almost all $\omega \in \Omega$.

A *simple function* is of the form

$$S = \sum_{i=1}^{N} x_i \chi_{E_i}$$

for some $x_1, x_2, \ldots, x_N \in X$ and $E_1, \ldots, E_N \in \Sigma$, which, without loss of generality, may be assumed to be pairwise disjoint. Here and elsewhere, χ_E is the characteristic function of the set E. The integral of such a simple function f is defined in the obvious way;

$$\int_E S \, d\mu = \sum_{i=1}^{N} \mu(E \cap E_i) x_i, \quad E \in \Sigma.$$

A μ-measurable function is *integrable* if there exists a sequence (S_n) of simple functions with

$$\int_\Omega \|S_n(\omega) - f(\omega)\| \, d\mu \to 0 . \quad \text{(Lebesgue integral)}$$

In this case the norm limit, as $n \to \infty$, of $\int_E S_n \, d\mu$ exists for each $E \in \Sigma$ and is independent of any particular choice for (S_n). By definition, $\int_E f \, d\mu$ is this common limit.

For $1 \le p < \infty$ we denote by $L_p(\Sigma, \mu; X)$ the space of $f\colon \Omega \to X$ for which $\|f\|_p = (\int_\Omega \|f(\omega)\|^p \, d\mu)^{\frac{1}{p}} < \infty$.

It is a significant result that the elements of $L_1(\Sigma, \mu; X)$ are precisely those functions which are integrable in the above sense. It follows from Hölder's inequality and the finiteness of μ that the elements of $L_p(\Sigma, \mu, X)$ are integrable. Further, as in the scalar case, $\|\cdot\|_p$ is a Banach space norm for this space.(3)

LEMMA 2. *$L_2(\Sigma, \mu; H)$ is a Hilbert space.*

PROOF. $(f,g) = \int_\Omega (f(\omega), g(\omega)) \, d\mu$ defines an inner-product on $L_2(\Sigma, \mu; H)$.

M.M. Day [1941] proves the following.

LEMMA 3. *$L_p(\Sigma, \mu; X)$ is uniformly convex if $1 < p < \infty$ and X is uniformly convex.*

---

(3) Strictly, the space of equivalence classes modulo almost everywhere zero functions.

Conditional Expectations.

Let Σ_0 be a proper sub σ-algebra of Σ, then $L_p(\Sigma_0,\mu;X)$ is a proper subspace of $L_p(\Sigma,\mu;X)$.

Clearly for any given $f \in L_p(\Sigma,\mu;X)$ there can be at most one $f_0 \in L_p(\Sigma_0,\mu;X)$ with $\int_E f\,d\mu = \int_E f_0\,d\mu$ for all $E \in \Sigma_0$.

When such a f_0 exists it will be referred to as the *conditional expectation* of f with respect to Σ_0 and denoted by $E(f|\Sigma_0)$.

LEMMA 4. *$f_0 \in M_0 = L_2(\Sigma_0,\mu;H)$ is the conditional expectation of f with respect to Σ_0 if and only if $(f-f_0,g) = 0$ for all $g \in M_0$.*

PROOF. It is sufficient to consider $g = x\chi_E$ where $x \in H$ and $E \in \Sigma_0$ (M_0 is the closed linear span of such functions). Now,

$$(f - f_0, x\,\chi_E) = \int_\Omega ((f - f_0)(\omega), x\,\chi_E(\omega))\,d\mu$$

$$= \int_E ((f - f_0)(\omega), x)\,d\mu = \left(\int_E (f - f_0)(\omega)\,d\mu, x\right).$$

Thus $(f - f_0, x\,\chi_E) = 0$ if and only if

$$\left(\int_E (f - f_0)(\omega)\,d\mu, x\right) = 0 \quad \text{for all } x \in H.$$

That is, if and only if $\int_E (f - f_0)(\omega)\,d\mu = 0$.

While it is not necessary for our subsequent work, one important consequence of this lemma is the existence of $E(f|\Sigma_0)$ for every $f \in L_2(\Sigma,\mu;H)$ and sub σ-algebra Σ_0. This follows since f_0 is the foot of the perpendicular from f to M_0 if and only if f_0 is the closest point from M_0 to f, the existence of which is ensured by the uniform convexity of $L_2(\Sigma,\mu;H)$.

Another consequence is the following.

COROLLARY 5. *For $f \in L_2(\Sigma,\mu;H)$, $\|E(f|\Sigma_0)\|_2 \le \|f\|_2$.[(4)] Indeed the mapping $f \to E(f|\Sigma_0)$ is a norm one linear projection onto $L_2(\Sigma_0,\mu;H)$.*

PROOF.

$$\begin{aligned}\|f_0\|^2 &= (f_0, f_0)\\ &= (f - (f - f_0), f_0)\\ &= (f, f_0), \text{ by lemma 4.}\\ &\le \|f\|\|f_0\|, \text{ by the Cauchy-Schwartz inequality.}\end{aligned}$$

§2 *MARTINGALES.*

Definition.

By a *martingale* (Σ_n, f_n) *in* $L_p(\Sigma,\mu;X)$ we mean a nested sequence $\Sigma_0 \le \Sigma_1 \le \ldots \le \Sigma_n \le \ldots \le \Sigma$ of sub σ-algebras and a sequence of functions $f_0, f_1, \ldots, f_n, \ldots$ with $f_n \in L_p(\Sigma_n,\mu;X)$, which satisfy the *martingale condition* $f_m = E(f_n|\Sigma_m)$ for all $n \ge m$.(5)

As an example note that for any nested sequences of sub σ-algebras $\Sigma_0 \le \Sigma_1 \le \ldots \le \Sigma_n \le \ldots \le \Sigma$ and any $f \in L_2(\Sigma,\mu;H)$, $(\Sigma_n, E(f|\Sigma_n))$ is a martingale. To check the martingale condition observe that for $n \ge m$ and $E \in \Sigma_m$ we have

$$\int_E E(f|\Sigma_m)\, d\mu = \int_E f\, d\mu, \text{ by the definition of } E(\cdot|\cdot),$$

$$= \int_E E(f|\Sigma_n)\, d\mu, \text{ again by the definition of } E(\cdot|\cdot) \text{ and the fact that } E \in \Sigma_n.$$

Proof of the Main Theorem.

We are now in a position to prove Theorem 1.

---

(4) A proof of this result for general $p \in [1,\infty)$ is the following. By their density we may assume that f is a simple function, then

$$\|E(f,\Sigma_0)\|_p^p = \int_\Omega \|E(\sum_{i=1}^n x_i \chi_{E_i}|\Sigma_0)\|^p\, d\mu$$
$$= \int_\Omega \|\sum_{i=1}^n x_i\, E(\chi_{E_i}|\Sigma_0)\|^p\, d\mu$$
$$\le \int_\Omega (\sum_{i=1}^n \|x_i\|\, E(\chi_{E_i}|\Sigma_0))^p\, d\mu$$
$$= \int_\Omega (E(\sum_{i=1}^n \|x_i\|\, \chi_{E_i}|\Sigma_0))^p\, d\mu$$
$$\le \int_\Omega E((\sum_{i=1}^n \|x_i\|\, \chi_{E_i})^p|\Sigma_0)\, d\mu, \text{ by Jensen's inequality.}$$
$$= \int_\Omega (\sum_{i=1}^n \|x_i\|\, \chi_{E_i})^p\, d\mu, \text{ by definition of } E(\cdot|\Sigma_0) \text{ and the fact } \Omega \in \Sigma_0.$$
$$= \int_\Omega \|f\|^p\, d\mu = \|f\|_p^p\, .$$

(5) Setting $\mu_n(E) = \int_E f_n\, d\mu$, for all $E \in \Sigma_n$, the martingale condition may be re-expressed as: For $n \ge m$ the restriction of μ_n to Σ_m is μ_m.

Let (Σ_n, f_n) be a martingale in $L_2(\Sigma,\mu;H)$ with $\|f_n\|_2 \le K$ for all n. Let $M_n = L_2(\Sigma_n,\mu;H)$, then we have the following structure.

A nested sequence of closed subspaces $M_0 \le M_1 \le \ldots \le M_n \le \ldots$ and a uniformly bounded sequence $f_0, f_1, \ldots, f_n, \ldots$ with $f_n \in M_n$.

By the martingale condition and corollary 5 we have $\|f_n\|_2 \le \|f_m\|_2$ for $n \le m$. Thus $\|f_0\|_2, \|f_1\|_2, \ldots, \|f_n\|_2, \ldots$ is an increasing sequence of real numbers bounded above by K, and so convergent from below to some real number k.

Further, for $n < m$

$$2\|f_m\|_2 \ge \|f_n\|_2 + \|f_m\|_2 \ge \|f_n + f_m\|_2$$

$$\ge \|E(f_n + f_m|\Sigma_n)\|_2 = 2\|f_n\|_2$$

(again, by corollary 5 and the martingale condition), so

$$\left\|\frac{f_n + f_m}{2}\right\|_2 \to k \quad \text{as} \quad m,n \to \infty.$$

Hence, by the uniform convexity of $L_2(\Sigma,\mu;H)$, $f_0, f_1, \ldots, f_n, \ldots$ is a Cauchy sequence and so by the completeness of the space, converges to some $f_\infty \in L_2(\Sigma,\mu;H)$.

If Σ_∞ denotes the σ-algebra generated by $\bigcup_n \Sigma_n$, it is trivial to check that $f_\infty \in \overline{\bigcup_n M_n} = L_2(\Sigma_\infty,\mu;H)$ – the proper home for our martingale limit.

On the basis of our earlier observations, the proof for arbitrary $p \in (1,\infty)$ and uniformly convex Banach space X is merely a paraphrase of the above.

REFERENCES

[1] M.M. Day, "Some more uniformly convex spaces", *Bull. Amer. Math. Soc.* 47 (1941) 504-507.

[2] J. Diestel and J.J. Uhl Jr., "Vector measures", *Amer. Math. Soc.*, Math. Survey No. 15 (1977).

[3] N. Dunford and J.T. Schwartz, *Linear Operators*, Part I. Interscience, N.Y. (1958).

University of New England

THE LARGEST PRIME

Readers of the Los Angeles Times (November 16, 1978) will no doubt know of the exploits of two freshmen students of USC Hayward, Laura Nichel and Curt Noll, who, for a high school computing project, showed that $2^{21701} - 1$ (*6553* digits) is prime. Noll continued this work and in February 1979 found that $2^{23209} - 1$ (*6987* digits) is prime. A more recent announcement of the L.A. Times (May 31, 1979) reveals that the the next Mersenne prime is $2^{44497} - 1$ (*13395* digits) a result due to Harry Nelson and David Slowinski.