1

Banach space valued martingales

We start by recalling the definition and basic properties of conditional expectations. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{B} \subset \mathcal{A}$ be a σ-subalgebra. On $L_2(\Omega, \mathcal{A}, \mathbb{P})$, the conditional expectation $\mathbb{E}^{\mathcal{B}}$ can be defined simply as the orthogonal projection onto the subspace $L_2(\Omega, \mathcal{B}, \mathbb{P})$. One then shows that it extends to a positive contraction on $L_p(\Omega, \mathcal{A}, \mathbb{P})$ for all $1 \le p \le \infty$, taking values in $L_p(\Omega, \mathcal{B}, \mathbb{P})$. The resulting operator $\mathbb{E}^{\mathcal{B}} : L_p(\Omega, \mathcal{A}, \mathbb{P}) \to L_p(\Omega, \mathcal{B}, \mathbb{P})$ is a linear projection and is characterized by the property

$$\forall f \in L_p(\Omega, \mathcal{A}, \mathbb{P})\ \forall h \in L_\infty(\Omega, \mathcal{B}, \mathbb{P}) \qquad \mathbb{E}^{\mathcal{B}}(hf) = h\mathbb{E}^{\mathcal{B}}(f). \tag{1.1}$$

Here ‘positive’ really means positivity preserving, i.e. for any $f \in L_p$,

$$f \ge 0 \Rightarrow T(f) \ge 0.$$

As usual, we often abbreviate ‘almost everywhere’ by ‘a.e.’ and ‘almost surely’ by ‘a.s.’

We will now consider conditional expectation operators on Banach space valued L_p-spaces.

1.1 Banach space valued L_p-spaces

Let $(\Omega, \mathcal{A}, m)$ be a measure space. Let B be a Banach space. We will denote by $F(B)$ the space of all measurable simple functions, i.e. the functions $f\colon\ \Omega \to B$ for which there is a partition of Ω, say, $\Omega = A_1 \cup \cdots \cup A_N$ with $A_k \in \mathcal{A}$, and elements $b_k \in B$ such that

$$\forall \omega \in \Omega \qquad f(\omega) = \sum\nolimits_1^N 1_{A_k}(\omega) b_k. \tag{1.2}$$

Equivalently, $F(B)$ is the space of all measurable functions $f\colon\ \Omega \to B$ taking only finitely many values.

Definition 1.1. We will say that a function $f\colon\ \Omega \to B$ is Bochner measurable if there is a sequence (f_n) in $F(B)$ tending to f pointwise.

Let $1 \le p \le \infty$. We denote by $L_p(\Omega, \mathcal{A}, m; B)$ the space of (equivalence classes of) Bochner measurable functions $f\colon\ \Omega \to B$ such that $\int \|f\|_B^p\, dm < \infty$ for $p < \infty$, and $\text{ess sup}\|f(\cdot)\|_B < \infty$ for $p = \infty$. As usual, two functions that are equal a.e. are identified. We equip this space with the norm

$$\|f\|_{L_p(B)} = \left(\int \|f\|_B^p\, dm\right)^{1/p} \quad \text{for} \quad p < \infty$$

$$\|f\|_{L_\infty(B)} = \text{ess sup}\|f(\cdot)\|_B \quad \text{for} \quad p = \infty,$$

with which it becomes a Banach space.

We will often use the following elementary consequence of Fubini's theorem:

$$\int \|f\|_B^p\, dm = \int_0^\infty p t^{p-1} m(\{\|f\|_B > t\})dt. \tag{1.3}$$

Of course, this definition of $L_p(B)$ coincides with the usual one in the scalar valued case, i.e. if $B = \mathbb{R}$ (or $\mathbb{C}$). In that case, we often denote simply by $L_p(\Omega, \mathcal{A}, m)$ (or sometimes $L_p(m)$, or even L_p) the resulting space of scalar valued functions.

For brevity, we will often write simply $L_p(m; B)$, or, if there is no risk of confusion, simply $L_p(B)$, instead of $L_p(\Omega, \mathcal{A}, m; B)$.

Given $\varphi_1, \dots, \varphi_N \in L_p$ and $b_1, \dots, b_N \in B$, we can define a function $f\colon\ \Omega \to B$ in $L_p(B)$ by setting $f(\omega) = \sum_1^N \varphi_k(\omega) b_k$. We will denote this function by $\sum_1^N \varphi_k \otimes b_k$ and by $L_p \otimes B$ the subspace of $L_p(B)$ formed of all such functions.

Proposition 1.2. *Let $1 \le p < \infty$. Each of the subspaces $F(B) \cap L_p(B)$ and $L_p \otimes B \subset L_p(B)$ is dense in $L_p(B)$. More generally, for any subspace $V \subset L_p$ dense in L_p, $V \otimes B$ is dense in $L_p(B)$.*

Proof. Consider $f \in L_p(B)$. Let $f_n \in F(B)$ be such that $f_n \to f$ pointwise. Then $\|f_n(\cdot)\|_B \to \|f(\cdot)\|_B$ pointwise, so that if we set $g_n(\omega) = f_n(\omega) 1_{\{\|f_n\| < 2\|f\|\}}$, we still have $g_n \to f$ pointwise and, in addition, $\sup_n \|g_n - f\| \le \sup_n \|g_n\| + \|f\| \le 3\|f\|$. Therefore, by dominated convergence, we must have $\int \|g_n - f\|_B^p\, dm \to 0$ and, of course, $g_n \in F(B) \cap L_p(B)$. This proves the first assertion. The second and third ones are then obvious since $F(B) \cap L_p(B) \subset L_p \otimes B$ (indeed, we can take $\varphi_k = 1_{A_k}$ with $m(A_k) < \infty$, as in (1.2)). $\square$

Remark 1.3. If B is finite-dimensional, then $F(B)$ is dense in $L_\infty(B)$, but this is no longer true in the infinite-dimensional case, because the unit ball of B is not compact.

Remark 1.4. Let $(\Omega, \mathcal{A})$ be a compact space equipped with its Borel σ-algebra (e.g. $\mathbb{T}$) and a Radon measure m; then any continuous B-valued function on Ω is Bochner measurable and the space $C(\Omega; B)$ of all such functions is included in $L_\infty(\Omega, \mathcal{A}, m; B)$. Moreover, $C(\Omega) = C(\Omega; \mathbb{R})$ is dense in $L_p(m; \mathbb{R})$ for any $1 \le p < \infty$. This remains true on a locally compact and σ-compact space (e.g. if $\Omega = \mathbb{R}$) provided $C(\Omega; B)$ is replaced by the space of compactly supported continuous functions.

We now turn to the definition of the integral of a function in $L_1(B)$. Consider a function f of the form (1.2) in $L_1(B) \cap F(B)$. We define

$$\int f\, dm = \sum\nolimits_1^N m(A_k) b_k.$$

This defines a continuous linear map from $L_1(B) \cap F(B)$ to B, since we have obviously, by the triangle inequality,

$$\left\| \int f\, dm \right\| \le \sum m(A_k)\|b_k\| = \|f\|_{L_1(B)}.$$

By density, this linear map admits an extension defined on the whole of $L_1(B)$, which we still denote by $\int f\, dm$ when $f \in L_1(B)$. The extension clearly satisfies the following fundamental inequality, called Jensen's inequality:

$$\forall f \in L_1(B) \qquad \left\| \int f\, dm \right\|_B \le \int \|f\|_B\, dm = \|f\|_{L_1(B)}. \tag{1.4}$$

This extends the linear map $f \to \int f\, dm$ from the scalar valued case to the B-valued one. More generally, let $(\Omega', \mathcal{A}', m')$ be another measure space and let $T\colon\ L_1(\Omega, \mathcal{A}, m) \to L_1(\Omega', \mathcal{A}', m')$ be a bounded operator. We may clearly define unambiguously a linear operator $T_0\colon\ F(B) \cap L_1(m; B) \to L_1(m', B)$ by setting, for any f of the form (1.2),

$$T_0(f) = \sum\nolimits_1^N T(1_{A_k}) b_k.$$

We have clearly, by the triangle inequality,

$$\|T_0(f)\|_{L_1(m';B)} \le \sum\nolimits_1^N \|T(1_{A_k})\|\,\|b_k\| \le \|T\| \sum m(A_k)\|b_k\| = \|T\|\,\|f\|_{L_1(B)}.$$

Thus we can state the following:

Proposition 1.5. *Given a bounded operator $T\colon L_1(\Omega, \mathcal{A}, m) \to L_1(\Omega', \mathcal{A}', m')$, there is a unique bounded linear map $\widetilde{T}\colon L_1(\Omega, \mathcal{A}, m; B) \to L_1(\Omega', \mathcal{A}', m'; B)$ such that*

$$\forall \varphi \in L_1(\Omega, \mathcal{A}, m)\ \forall b \in B \qquad \widetilde{T}(\varphi \otimes b) = T(\varphi)b. \tag{1.5}$$

Moreover, we have $\|\widetilde{T}\| = \|T\|$.

Proof. By the density of $F(B) \cap L_1(B)$ in $L_1(B)$, the (continuous) map T_0 admits a unique continuous linear extension $\widetilde{T}$ from $L_1(m; B)$ to $L_1(m'; B)$, with $\|\widetilde{T}\| \le \|T_0\| \le \|T\|$. If φ is a simple function in L_1, then (1.5) is clear by definition of T_0. Approximating φ in L_1 by a simple function, we see that (1.5) is true in general. The unicity of $\widetilde{T}$ is clear since (1.5) implies that $\widetilde{T}$ coincides with T_0 on $F(B) \cap L_1(B)$. Finally, considering a fixed b with $\|b\| = 1$, we easily derive from (1.5) that $\|T\| \le \|\widetilde{T}\|$, so we obtain $\|T\| = \|\widetilde{T}\|$. □

It is not true in general that a bounded operator on L_p (or from L_p to L_q) extends boundedly to $L_p(B)$ as in the preceding proposition for $p = 1$. Nevertheless, it is true for *positive* operators, as follows:

Proposition 1.6. *Let $1 \le p, q \le \infty$. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an arbitrary measure space and let $T\colon\ L_p(\Omega) \to L_q(\Omega)$ be a bounded linear operator. Let*

$$T \otimes I_B\colon\ L_p(\Omega, \mathbb{P}) \otimes B \to L_q(\Omega, \mathbb{P}) \otimes B$$

be the unique linear operator such that

$$\forall \varphi \in L_p(\Omega, \mathbb{P}) \quad \forall x \in B \qquad (T \otimes I_B)(\varphi \otimes x) = T(\varphi) \otimes x.$$

If T is positive (i.e. if T preserves non-negative functions), then $T \otimes I_B$ extends to a bounded operator $\widetilde{T \otimes I_B}$ from $L_p(\Omega, \mathbb{P}; B)$ to $L_q(\Omega, \mathbb{P}; B)$, which has the same norm as T, i.e.

$$\|\widetilde{T \otimes I_B}\|_{L_p(B) \to L_q(B)} = \|T\|_{L_p \to L_q}.$$

Proof. It clearly suffices to show that

$$\forall f \in L_p(\Omega, \mathbb{P}) \otimes B \qquad \|(T \otimes I_B) f(\cdot)\|_B \overset{\text{a.s.}}{\le} T(\|f(\cdot)\|_B). \tag{1.6}$$

For that purpose, we can assume B separable (or even finite-dimensional) so that there is a countable subset $D \subset B^*$ verifying

$$\forall x \in B \qquad \|x\| = \sup_{\xi \in D} |\xi(x)|.$$

Clearly, for any ξ in B^*, we have

$$\langle \xi, (T \otimes I_B) f(\cdot) \rangle = T(\langle \xi, f(\cdot) \rangle),$$

and hence by the positivity of T for any finite subset $D' \subset D$,

$$\sup_{\xi \in D'} |\langle \xi, (T \otimes I_B) f(\cdot) \rangle| \overset{\text{a.s.}}{\leq} T(\sup_{\xi \in D} |\langle \xi, f(\cdot) \rangle|);$$

therefore we obtain (1.6), and the proposition follows. □

Remark. Let B_1 be another Banach space and let $u\colon\ B \to B_1$ be a bounded operator. Then, for any f in $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$, we have

$$\widetilde{T \otimes I_{B_1}}(u(f)) = u[\widetilde{T \otimes I_B}(f)].$$

In particular, if we set $g = \widetilde{T \otimes I_B}(f)$, then for any ξ in B^*, we have

$$T(\xi(f)) = \xi(g). \tag{1.7}$$

Indeed, this is immediately checked for f in $L_p(\Omega, \mathbb{P}) \otimes B$, and the general case is obtained after completion.

Note that now that $\widetilde{T \otimes I_B}$ makes sense, the preceding argument can be repeated to show that

$$\forall f \in L_p(\Omega, \mathbb{P}; B) \qquad \|(\widetilde{T \otimes I_B}) f\|_B \overset{\text{a.e.}}{\leq} T(\|f(\cdot)\|_B). \tag{1.8}$$

A priori, in the preceding (1.8), we implicitly assume that B is a real Banach space, but actually, if B is a complex space (and T is $\mathbb{C}$-linear on complex valued L_p), we may consider B a fortiori as a real space and then (1.8) remains valid.

Remark 1.7. Let $1 \leq p \leq \infty$. Let B_1, B_2 be Banach spaces. We will denote by

$$B_1 \oplus_p B_2$$

the direct sum equipped with the norm defined by

$$\forall (x_1, x_2) \in B_1 \oplus B_2 \quad \|(x_1, x_2)\| = (\|x_1\|^p + \|x_2\|^p)^{1/p}. \tag{1.9}$$

More generally, for any family $(B_i)_{i \in I}$ of Banach spaces, we denote by

$$\left(\oplus \sum\nolimits_{i \in I} B_i\right)_p$$

the direct sum, 'in the sense of ℓ_p', which means the subspace formed of those $x = (x(i)) \in \prod_{i \in I} B_i$ such that $\sum_{i \in I} \|x(i)\|^p < \infty$ equipped with the norm $x \mapsto (\sum_{i \in I} \|x(i)\|^p)^{1/p}$. When the spaces B_i are all identical to a space B, the latter space will be denoted by $\ell_p(I; B)$. When $I = \mathbb{N}$, we denote this simply by $\ell_p(B)$. Note that $\ell_p(B) = L_p(\mathbb{N}, \nu; B)$, where ν is the counting measure on $\mathbb{N}$.

Remark 1.8. Occasionally, mainly in Chapters 4 and 8, we will consider quasi-Banach spaces. Here is a brief discussion of this notion. When $0 < p < 1$, the space $E = L_p(m)$ over any measure m equipped with its usual distance

$d(f, g) = \|f - g\| = (\int |f - g|^p dm)^{1/p}$ is complete but fails the usual triangle inequality and is not even locally convex (except when finite-dimensional). However, for any $f, g \in E$, it is well known that

$$\|f + g\| \leq (\|f\|^p + \|g\|^p)^{1/p}. \tag{1.10}$$

We call this the p-triangle inequality, and any space satisfying this is called p-normed. A space that is p-normed for some $0 < p < 1$ is often called quasi-normed. It suffices for this that there be a constant c such that $\|f + g\| \leq c(\|f\| + \|g\|)$ for any $f, g \in E$. When it is complete, we refer to it as a quasi-Banach space. Clearly, for any normed (or merely p-normed) space B, the space $L_p(m; B)$, which can be defined as earlier, is again p-normed. Thus, for any quasi-Banach space B, the space $L_p(m; B)$ is a quasi-Banach space. We refer to [45] for more information on quasi-Banach spaces.

Remark 1.9. Let $(\Omega, \mathcal{A}, m)$ be a σ-finite measure space. We will denote by $L_0(\Omega, \mathcal{A}, m; B)$ (or simply $L_0(m; B)$) the set of equivalence classes (modulo equality almost everywhere) of Bochner measurable B-valued functions f. When m is finite, $L_0(\Omega, \mathcal{A}, m; B)$ is classically equipped with a metrizable topological vector space structure. One possible distance defining this topology is

$$\forall f, g \in L_0(m; B) \quad d(f, g) = \int \frac{\|f - g\|}{1 + \|f - g\|} dm.$$

Consider a sequence (f_n) in $L_0(m; B)$. Then we have

$$d(f_n, f) \to 0 \Leftrightarrow m(\{\|f_n - f\| > \varepsilon\}) \to 0 \quad \forall \varepsilon > 0.$$

When $m(\Omega) = \infty$, assuming m σ-finite, let $w \in L_1(m)$ be such that $w > 0$ a.e. so that the measure $m' = w \cdot m$ is finite and equivalent to m. We may obviously identify $L_0(m; B)$ with $L_0(m'; B)$. We then define the topological vector space structure of $L_0(m; B)$ by transplanting the one just defined on $L_0(m'; B)$. Then f_n tends to f in this topology iff, for any measurable subset $A \subset \Omega$ with $m(A) < \infty$, the restriction of f_n to A converges in measure to $f_{|A}$ in the preceding sense (this shows in particular that this is independent of the choice of w). We then say that f_n tends to f 'in measure' (or 'in probability' if $m(\Omega) = 1$). It is sufficient (resp. necessary) for this that f_n tends to f almost everywhere (resp. that (f_n) admits a subsequence converging a.e. to f).

In conclusion, for any σ-finite measure space, and any $0 < p \leq \infty$, we have a continuous injective linear embedding $L_p(m; B) \subset L_0(m; B)$. All these facts are easily checked.

1.2 Banach space valued conditional expectation

In particular, Proposition 1.6 applied with $p = q$, is valid for $T = \mathbb{E}^{\mathcal{B}}$. For any f in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$, we denote again simply by $\mathbb{E}^{\mathcal{B}}(f)$ the function $\widetilde{T \otimes I_B}(f)$ for $T = \mathbb{E}^{\mathcal{B}}$.

Proposition 1.10. *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space. The conditional expectation* $f \mapsto \mathbb{E}^{\mathcal{B}}(f)$ *is an operator of norm* 1 *on* $L_p(B)$ *for all* $1 \le p \le \infty$*, satisfying*

$$\forall f \in L_p(\Omega, \mathcal{A}, \mathbb{P}; B)\ \forall h \in L_\infty(\Omega, \mathcal{B}, \mathbb{P}) \quad \mathbb{E}^{\mathcal{B}}(hf) = h\mathbb{E}^{\mathcal{B}}(f). \quad (1.11)$$

The operator $\mathbb{E}^{\mathcal{B}}$ *is a norm* 1 *projection from* $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ *to* $L_p(\Omega, \mathcal{B}, \mathbb{P}; B)$*, viewed as a subspace of* $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$*.*

Proof. By (1.1), Proposition 1.6 applied to $T = \mathbb{E}^{\mathcal{B}}$ with $p = q$ produces an operator P of norm 1 on $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$, such that $P(hx) = hP(x)$ for any $x \in L_p(\mathcal{A}, \mathbb{P}) \otimes B$ and $h \in L_\infty(\Omega, \mathcal{B}, \mathbb{P})$ and also such that $P(x) = x$ for any $x \in L_p(\mathcal{B}, \mathbb{P}) \otimes B$. Therefore, by density, $P(hx) = hP(x)$ holds for any $x \in L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$. Also $P(x) = x$ holds for any $x \in \overline{L_p(\mathcal{B}, \mathbb{P}) \otimes B} \subset L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$. Since we may clearly identify the latter subspace with $L_p(\Omega, \mathcal{B}, \mathbb{P}; B)$, the statement follows. □

Remark 1.11. Let $1 \le p, p' \le \infty$ such that $p^{-1} + p'^{-1} = 1$. Let $f \in L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$. Note that $g = \mathbb{E}^{\mathcal{B}}(f) \in L_p(\Omega, \mathcal{B}, \mathbb{P}; B)$ satisfies the following:

$$\forall E \in \mathcal{B} \quad \int_E g\,d\mathbb{P} = \int_E f\,d\mathbb{P}. \quad (1.12)$$

In particular, $\mathbb{E}(g) = \mathbb{E}(f)$. More generally,

$$\forall h \in L_{p'}(\Omega, \mathcal{B}, \mathbb{P}) \qquad \int hg\,d\mathbb{P} = \int hf\,d\mathbb{P}, \quad (1.13)$$

and also

$$\forall h \in L_{p'}(\Omega, \mathcal{A}, \mathbb{P}) \quad \int hg\,d\mathbb{P} = \int (\mathbb{E}^{\mathcal{B}}h)g\,d\mathbb{P}. \quad (1.14)$$

Indeed (among many other ways to verify these equalities), they are easy to check by 'scalarization', by a simple deduction from the scalar case. More precisely, for each of these identities, we have to check the equality of two vectors in B, say, $x = y$, and this is the same as $\xi(x) = \xi(y)$ for all ξ in B^*. But we know by (1.7) that $\langle \xi, g(\cdot)\rangle = \mathbb{E}^{\mathcal{B}}\langle \xi, f\rangle$, so to check (1.12)–(1.14), it suffices to check them for $B = \mathbb{R}$ (or $B = \mathbb{C}$ in the case of complex scalars) and for all scalar valued f. But in that case, they are easy to deduce from the basic (1.1). Indeed, by the density of L_∞ in $L_{p'}$, (1.1) remains valid when $h \in L_{p'}$, so we derive (1.12) and (1.13) from it. To check (1.14), inverting the roles of g and h, we find by (1.12) with $E = \Omega$ (and with f replaced by hg) that

$$\int hg d\mathbb{P} = \int \mathbb{E}^{\mathcal{B}}(hg) d\mathbb{P} = \int (\mathbb{E}^{\mathcal{B}} h) g d\mathbb{P}.$$

In the sequel, we will refer to the preceding way to check a vector valued functional identity as the *scalarization principle*.

We will sometimes use the following generalization of (1.8).

Proposition 1.12. *Let $f \in L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$. For any continuous convex function $\varphi : B \to \mathbb{R}$ such that $\varphi(f) \in L_1$, we have*

$$\varphi(\mathbb{E}^{\mathcal{B}}(f)) \le \mathbb{E}^{\mathcal{B}} \varphi(f). \tag{1.15}$$

In particular,

$$\|\mathbb{E}^{\mathcal{B}}(f)\| \le \mathbb{E}^{\mathcal{B}} \|f\|. \tag{1.16}$$

Proof. We may clearly assume B separable. Recall that a real valued convex function on B is continuous iff it is locally bounded above ([82, p. 93]). By Hahn-Banach, we have $\varphi(x) = \sup_{a \in \mathcal{C}} a(x)$, where $\mathcal{C}$ is the collection of all continuous real valued affine functions a on B such that $a \le \varphi$. By the separability of B, we can find a countable subcollection $\mathcal{C}'$ such that $\varphi(x) = \sup_{a \in \mathcal{C}'} f(x)$.

By (1.7), we have $a(\mathbb{E}^{\mathcal{B}}(f)) = \mathbb{E}^{\mathcal{B}}(a(f))$ for any $a \in B^*$, and hence, also for any affine continuous function a on B (since this is obvious for constant functions). Thus we have $a(\mathbb{E}^{\mathcal{B}} f) = \mathbb{E}^{\mathcal{B}} a(f) \le \mathbb{E}^{\mathcal{B}} \varphi(f)$ for any $a \in \mathcal{C}'$, and hence, taking the sup over $a \in \mathcal{C}'$, we find (1.15). □

Remark 1.13. To conform with the tradition, we assumed that $\mathbb{P}$ is a probability in Corollary 1.10. However, essentially the same results remain valid if we merely assume that the restriction of $\mathbb{P}$ to $\mathcal{B}$ is σ-finite (or that $L_\infty \cap L_1$ is weak*-dense in L_∞ both for $\mathcal{B}$ and $\mathcal{A}$). Indeed, we can still define $\mathbb{E}^{\mathcal{B}}$ on $L_2(\Omega, \mathcal{A}, \mathbb{P})$ as the orthogonal projection onto $L_2(\Omega, \mathcal{B}, \mathbb{P})$, and although it is rarely used, it is still true that the resulting operator on $L_p(\Omega, \mathcal{A}, \mathbb{P})$ preserves positivity and extends to an operator of norm 1 on $L_p(\Omega, \mathcal{A}, \mathbb{P}; B)$ for all $1 \le p \le \infty$, satisfying (1.11), (1.13) and (1.14). See Remark 6.20 for an illustration.

Remark. By classical results due to Ron Douglas [213] and T. Ando [102], conditional expectations can be characterized as the only norm 1 projections on L_p ($1 \le p \neq 2 < \infty$) that preserve the constant function 1. This is also true for $p = 2$ if one restricts to positivity-preserving operators.

1.3 Martingales: basic properties

Let B be a Banach space. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A sequence $(M_n)_{n\geq 0}$ in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$ is called a martingale if there exists an increasing sequence of σ-subalgebras $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_n \subset \cdots \subset \mathcal{A}$ (this is called 'a filtration') such that for each $n \geq 0$, M_n is $\mathcal{A}_n$-measurable and satisfies

$$M_n = \mathbb{E}^{\mathcal{A}_n}(M_{n+1}). \tag{1.17}$$

This implies, of course, that

$$\forall n < m \qquad M_n = \mathbb{E}^{\mathcal{A}_n} M_m.$$

In particular, if (M_n) is a B-valued martingale, (1.12) yields

$$\forall n \leq m \quad \forall A \in \mathcal{A}_n \qquad \int_A M_n d\mathbb{P} = \int_A M_m d\mathbb{P}. \tag{1.18}$$

A sequence of random variables (M_n) is called 'adapted to the filtration' $(\mathcal{A}_n)_{n\geq 0}$ if M_n is $\mathcal{A}_n$-measurable for each $n \geq 0$. Note that the martingale property $M_n = \mathbb{E}^{\mathcal{A}_n}(M_{n+1})$ automatically implies that (M_n) is adapted to $(\mathcal{A}_n)$. Of course, the minimal choice of a filtration to which (M_n) is adapted is simply the filtration $\mathcal{M}_n = \sigma(M_0, M_1, \ldots, M_n)$. Moreover, if (M_n) is a martingale in the preceding sense with respect to some filtration $(\mathcal{A}_n)$, then it is a fortiori a martingale with respect to $(\mathcal{M}_n)$. Indeed, we have obviously $\mathcal{M}_n \subset \mathcal{A}_n$ for all n, and hence applying $\mathbb{E}^{\mathcal{M}_n}$ to both sides of (1.17) implies $M_n = \mathbb{E}^{\mathcal{M}_n}(\mathbb{E}^{\mathcal{A}_n} M_{n+1}) = \mathbb{E}^{\mathcal{M}_n} M_{n+1}$.

An adapted sequence of random variables (M_n) is called 'predictable' if M_n is $\mathcal{A}_{n-1}$-measurable for each $n \geq 1$. Of course, the predictable sequences of interest to us will not be martingales, since predictable martingales must form a constant sequence.

We will also need the definition of a submartingale. A sequence $(M_n)_{n\geq 0}$ of real valued random variables in L_1 is called a submartingale if there are σ-subalgebras $\mathcal{A}_n$, as previously, such that M_n is $\mathcal{A}_n$-measurable and

$$\forall n \geq 0 \qquad M_n \leq \mathbb{E}^{\mathcal{A}_n} M_{n+1}.$$

This implies, of course, that

$$\forall n < m \qquad M_n \leq \mathbb{E}^{\mathcal{A}_n} M_m.$$

For example, if (M_n) is a B-valued martingale in $L_1(B)$, then, for any continuous convex function $\varphi : B \to \mathbb{R}$ such that $\varphi(M_n) \in L_1$ for all n, the sequence $(\varphi(M_n))$ is a submartingale. Indeed, by (1.15), we have

$$\varphi(M_n) \leq \mathbb{E}^{\mathcal{A}_n} \varphi(M_{n+1}). \tag{1.19}$$

A fortiori, taking the expectation of both sides, we have (for future reference)

$$\mathbb{E}\varphi(M_n) \le \mathbb{E}\varphi(M_{n+1}). \tag{1.20}$$

More generally, if I is any partially ordered set, then a collection $(M_i)_{i\in I}$ in $L_1(\Omega, \mathbb{P}; B)$ is called a martingale (indexed by I) if there are σ-subalgebras $\mathcal{A}_i \subset \mathcal{A}$ such that $\mathcal{A}_i \subset \mathcal{A}_j$ whenever $i < j$ and $M_i = \mathbb{E}^{\mathcal{A}_i} M_j$.

In particular, when the index set is

$$I = \{0, -1, -2, \dots\},$$

the corresponding sequence is usually called a 'reverse martingale'.

The following convergence theorem is fundamental.

Theorem 1.14. *Let $(\mathcal{A}_n)$ be a fixed increasing sequence of σ-subalgebras of $\mathcal{A}$. Let $\mathcal{A}_\infty$ be the σ-algebra generated by $\bigcup_{n\ge 0} \mathcal{A}_n$. Let $1 \le p < \infty$ and consider M in $L_p(\Omega, \mathbb{P}; B)$. Let us define $M_n = \mathbb{E}^{\mathcal{A}_n}(M)$. Then $(M_n)_{n\ge 0}$ is a martingale such that $M_n \to \mathbb{E}^{\mathcal{A}_\infty}(M)$ in $L_p(\Omega, \mathbb{P}; B)$ when $n \to \infty$.*

Proof. Note that since $\mathcal{A}_n \subset \mathcal{A}_{n+1}$, we have $\mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_{n+1}} = \mathbb{E}^{\mathcal{A}_n}$, and similarly, $\mathbb{E}^{\mathcal{A}_n}\mathbb{E}^{\mathcal{A}_\infty} = \mathbb{E}^{\mathcal{A}_n}$. Replacing M by $\mathbb{E}^{\mathcal{A}_\infty} M$, we can assume without loss of generality that M is $\mathcal{A}_\infty$-measurable. We will use the following fact: the union $\bigcup_n L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)$ is dense in $L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B)$. Indeed, let $\mathcal{C}$ be the class of all sets A such that $1_A \in \overline{\bigcup_n L_\infty(\Omega, \mathcal{A}_n, \mathbb{P})}$, where the closure is meant in $L_p(\Omega, \mathbb{P})$ (recall $p < \infty$). Clearly $\mathcal{C} \supset \bigcup_{n\ge 0} \mathcal{A}_n$ and $\mathcal{C}$ is a σ-algebra, hence $\mathcal{C} \supset \mathcal{A}_\infty$. This gives the scalar case version of the preceding fact. Now, any f in $L_p(\Omega, \mathcal{A}_\infty, \mathbb{P}; B)$ can be approximated (by definition of the spaces $L_p(B)$) by functions of the form $\sum_1^n 1_{A_i} x_i$ with $x_i \in B$ and $A_i \in \mathcal{A}_\infty$. But since $1_{A_i} \in \overline{\bigcup_n L_\infty(\Omega, \mathcal{A}_n, \mathbb{P})}$, we clearly have $f \in \overline{\bigcup_n L_p(\Omega, \mathcal{A}_n, \mathbb{P}; B)}$, as announced.

We can now prove Theorem 1.14. Let $\varepsilon > 0$. By the preceding fact, there is an integer k and g in $L_p(\Omega, \mathcal{A}_k, \mathbb{P}; B)$ such that $\|M - g\|_p < \varepsilon$. We have then $g = \mathbb{E}^{\mathcal{A}_n} g$ for all $n \ge k$, hence

$$\forall n \ge k \qquad M_n - M = \mathbb{E}^{\mathcal{A}_n}(M - g) + g - M$$

and, finally,

$$\begin{aligned}\|M_n - M\|_p &\le \|\mathbb{E}^{\mathcal{A}_n}(M - g)\|_p + \|g - M\|_p \\ &\le 2\varepsilon.\end{aligned}$$

This completes the proof. □

Definition 1.15. Let B be a Banach space and let $(M_n)_{n\ge 0}$ be a sequence in $L_1(\Omega, \mathcal{A}, \mathbb{P}; B)$. We will say that (M_n) is uniformly integrable if the sequence of non-negative random variables $(\|M_n(\cdot)\|)_{n\ge 0}$ is uniformly integrable. More