A CONVEXITY CONDITION IN BANACH SPACES AND THE STRONG LAW OF LARGE NUMBERS[1]

ANATOLE BECK

Introduction. The strong law of large numbers can be shown under certain hypotheses for random variables which take values in Banach spaces. The general statement reads as follows:

THEOREM. *Let $\mathfrak{X}$ be a Banach space and let $\{X_i\}$ be a sequence of independent random $\mathfrak{X}$-variables (see definition below) with $E(X_i)=0$, all $i>0$. Under appropriate conditions on $\mathfrak{X}$ and on $\{X_i\}$, we can then assert that $(1/n)\sum_{i=1}^{n} X_i$ converges to 0 in the strong topology of $\mathfrak{X}$ almost surely.*[2]

In a recent paper [1], this author showed this theorem under the hypotheses that $\mathfrak{X}$ is uniformly convex and that the variances of X_i are uniformly bounded ($\mathrm{Var}(X_i)=E(\|X_i\|^2)$). At the same time, an example was given of a space in which the theorem fails. It is now possible, using the methods of [1], to show a necessary and sufficient condition on the Banach space $\mathfrak{X}$ to yield this particular strong law of large numbers.

A Banach space $\mathfrak{X}$ is said to have property (A) if, for every sequence $\{X_i\}$ of independent random $\mathfrak{X}$-variables with $E(X_i)=0$, all i, and $\mathrm{Var}(X_i)<M$, all i, we have

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \qquad \text{strongly almost surely.}$$

A Banach space $\mathfrak{X}$ is said to have property (B) if there exists an integer $k>0$ and an $\epsilon>0$ such that any choice $a_1, a_2, \cdots, a_k$ of elements from $\mathfrak{X}$ with $\|a_i\|\leqq 1$ gives us

$$\|\pm a_1 \pm a_2 \pm \cdots \pm a_k\| < k(1-\epsilon)$$

for some combination of the $+$ and $-$ signs.

We shall show that these two conditions are equivalent.

1. **Definitions.** Let $\mathfrak{X}$ be a separable Banach space and let

Presented to the Society, January 23, 1961; received by the editors February 14, 1961.

[1] This research was supported by Cornell University under contract with the Office of Naval Research and by the University of Wisconsin under contract number AF 49(638)-868 with the Air Force Office of Scientific Research.

[2] Here and hereafter in this paper unless otherwise specified, all limits are taken as $n\to\infty$.

(S, Σ, m) be a measure space. Then A mapping X from S into $\mathfrak{X}$ is called *strongly measurable* if $X^{-1}(B)=\{s \mid X(s)\in B\}$ is measurable for every Borel set $B\subset\mathfrak{X}$. If X is strongly measurable and if $\int_S \|X(s)\| m ds < \infty$ then it can be shown that there is a $y\in\mathfrak{X}$ such that $x^*(y) = \int_S x^*(X(s)) m ds$ for every $x^*\in\mathfrak{X}^*$. y is defined as the *integral of* X.

A *probability space* (customarily denoted $(\Omega, B, \Pr)$) is a measure space of total measure 1. ($\Pr(\Omega)=1$.) A strongly measurable function X from Ω into $\mathfrak{X}$ is called a *random $\mathfrak{X}$-variable*, and its integral, if it has one, is called its *expectation*, $E(X)$. If $X_1, \cdots, X_m$ are random $\mathfrak{X}$-variables and if for every choice $B_1, \cdots, B_m$ of Borel sets from $\mathfrak{X}$, we have

$$\Pr\{X_1 \in B_1, \cdots, X_m \in B_m\} = \prod_{i=1}^{m} \Pr(X_i \in B_i),$$

then $X_1, \cdots, X_m$ are an *independent* collection of random $\mathfrak{X}$-variables. If, in an infinite collection $\{X_\alpha, \alpha\in A\}$ of random $\mathfrak{X}$-variables, every finite sub-collection is independent, then the infinite collection is said to be independent.

For each random $\mathfrak{X}$-variable X, we define $\mathrm{Var}(X)=E(\|X-E(X)\|^2) = \int_\Omega \|X(\omega)-E(X)\|^2 \Pr d\omega$. For each sequence $\{X_i\}$ of random $\mathfrak{X}$-variables, we define

$$c\{X_i\} = \operatorname*{ess\,sup}_{\Omega} \ \limsup_{n} \left\| \frac{1}{n}\sum_{i=1}^{n} X_i \right\|.$$

A random $\mathfrak{X}$-variable X is called *symmetric* if there is a measure-preserving mapping ϕ of Ω into Ω such that $X(\phi(\omega)) = -X(\omega)$ for (almost) all $\omega\in\Omega$.

2. Condition (B) implies condition (A).

THEOREM 1. *Let $\mathfrak{X}$ be a Banach space satisfying condition* (B) *and let $\{X_i\}$ be a sequence of independent random $\mathfrak{X}$-variables with $E(X_i)=0$ and* $\mathrm{Var}(X_i)<M$, $i=1, 2, \cdots$. *Then*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \qquad \textit{strongly in } \mathfrak{X} \textit{ almost surely.}^{3}$$

Instead of proving Theorem 1, we prove Lemma 2, below, which is the same result under additional hypotheses. The derivation of Theorem 1 from Lemma 2, *i.e.*, the exorcism of the extraneous hypotheses, can be taken verbatim from [1], since the convexity condition is used only in Lemma 2.

[3] *I.e.*, for almost all $\omega\in\Omega$.

LEMMA 2. *If* $\mathfrak{X}$ *satisfies condition* (B) *and* $\{X_i\}$ *is a sequence of random* $\mathfrak{X}$*-variables and*

(1) *the* X_i *are independent,*

(2) $E(X_i)=0$, $i=1, 2, \cdots$,

(3) $\|X_i(\omega)\| \leqq 1$, *all* $i=1, 2, \cdots$; $\omega \in \Omega$,

(4) X_i *is symmetric,* $i=1, 2, \cdots$,

then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \text{ strongly in } \mathfrak{X} \text{ almost surely.}$$

PROOF. We will designate the sequences satisfying these hypotheses as being of type 2.

Then we can read the lemma as saying that if $\mathfrak{X}$ satisfies condition (B) and $\{X_i\}$ is of type 2, then $c\{X_i\}=0$. If we set $C=C(\mathfrak{X})$ $=\sup(c\{X_i\} \mid \{X_i\}$ of type 2), then we are to prove that $C=0$.

We assume, contrarily, that $C\neq 0$. (Note that $c\{X_i\}$ always exists under these hypotheses, and does not exceed 1, so that C exists and is no greater than 1.) We shall derive a contradiction. Choose any $\eta>0$, and let $\{U_i\}$ be chosen so that $\{U_i\}$ is of type 2, and $c\{U_i\}$ $>C-\eta$. Let k and ϵ be the numbers given to us in the definition of condition (B), and consider the random $\mathfrak{X}$-variables

$$V_i = \frac{U_{ki} + U_{ki-1} + \cdots + U_{ki-k+1}}{k}.$$

It is easily seen that $\{V_i\}$ is of type 2, and $c\{V_i\}=c\{U_i\}$. Furthermore, we can show that $E(\|V_i\|)<1-\epsilon/2^k$. Let ϕ_i be chosen so that

$$U_i(\phi_i(\omega)) = -U_i(\omega), \qquad U_j(\phi_i(\omega)) = U_j(\omega),$$

all $\omega\in\Omega$, $i=1, 2, \cdots, j\neq i$.[4] Then if we look at the 2^k mappings

$$\phi_{ki}^{\alpha_k}\phi_{ki-1}^{\alpha_{k-1}} \cdots \phi_{ki-k+1}^{\alpha_1}$$

given by the possible choices of $\alpha_j=0, 1, j=1, \cdots, k$, we see that all these are measure-preserving transformations on Ω, and that for every $\omega\in\Omega$, some one of these, Φ_ω, has the property that

$$\left\|\sum_{j=ki-k+1}^{ki} U_j(\Phi_\omega(\omega))\right\| = \|\pm U_{ki-k+1}(\omega) \pm \cdots \pm U_{ki}(\omega)\| < k(1-\epsilon).$$

Therefore, if we number these 2^k mappings as $\Phi_1, \Phi_2, \cdots, \Phi_{2^k}$, we have

[4] Possibly this may require a change to an equivalent sequence of random variables in an isomorphic measure space. The truth of this lemma survives such a transplanting.

$$\sum_{r=1}^{2^k}\left\|\sum_{j=ki-k+1}^{ki} U_j(\Phi_r(\omega))\right\| < k(2^k-1)+k(1-\epsilon) = k(2^k-\epsilon)$$

for each $\omega\in\Omega$, and therefore

$$2^k E\left(\left\|\sum_{j=ki-k+1}^{ki} U_j\right\|\right) = E\left(\sum_{r=1}^{2^k}\left\|\sum_{j=ki-k+1}^{ki} U_j(\Phi_r(\cdot))\right\|\right) < k(2^k-\epsilon),$$

so that

$$E(\|V_i\|) = \frac{1}{k}E\left(\left\|\sum_{j=ki-k+1}^{ki} U_j\right\|\right) < \frac{1}{k}\cdot\frac{k(2^k-\epsilon)}{2^k} = 1-\frac{\epsilon}{2^k}\cdot$$

Remember that ϵ and k are constants depending only on the space $\mathfrak{X}$.

Let $t>1/\eta^3$, and for each $i>0$, define

$$W_i = \frac{V_{ti}+V_{ti-1}+\cdots+V_{ti-t+1}}{t}.$$

We easily see that $\{W_i\}$ is of type 2, and that $c\{W_i\}=c\{V_i\}$. Since $\operatorname{Var}(\|V_i\|)<1$, all $i=1, 2, \cdots$, and the V_i are independent, we have

$$\operatorname{Var}\left(\sum_{j=ti-t+1}^{ti}\left\|\frac{V_j}{t}\right\|\right) < \frac{1}{t}\cdot$$

Thus, for each $i>0$,

$$\begin{aligned}\Pr\left\{\|W_i\| > 1-\frac{\epsilon}{2^k}+\eta\right\} &= \Pr\left\{\left\|\sum_{j=ti-t+1}^{ti}\frac{V_j}{t}\right\| > 1-\frac{\epsilon}{2^k}+\eta\right\}\\ &\leqq \Pr\left\{\sum_{j=ti-t+1}^{ti}\left\|\frac{V_j}{t}\right\| > 1-\frac{\epsilon}{2^k}+\eta\right\}\\ &< \frac{1/t}{\eta^2} < \eta,\end{aligned}$$

by Chebyshev's inequality. Using the independe: ce of the W_i, we define

$$\left.\begin{aligned} Y_i = W_i,\quad Z_i = 0 \quad &\text{if } \|W_i\| \leqq 1-\frac{\epsilon}{2^k}+\eta,\\ Y_i = 0,\quad Z_i = W_i \quad &\text{if } \|W_i\| > 1-\frac{\epsilon}{2^k}+\eta,\end{aligned}\right\} \quad i = 1, 2, \cdots.$$

Then $\{Y_i\}$ is of type 2, and in fact $\|Y_i(\omega)\| \leqq 1-\epsilon/2^k+\eta$, all $\omega\in\Omega$, $i=1, 2, \cdots$, so that $c\{Y_i\} \leqq C(1-\epsilon/2^k+\eta)$. The sequence $\{\|Z_i\|\}$ is independent, and since $\|Z_i(\omega)\| \leqq 1$, all $\omega\in\Omega$, $i=1, 2, \cdots$ and $\Pr\{Z_i=0\} > 1-\eta$, we have $E(\|Z_i\|) < \eta$. Thus $c\{Z_i\} \leqq c\{\|Z_i\|\} \leqq \eta$, by the (real-valued) strong law of large numbers. It is easily seen that $c\{W_i\} \leqq c\{Y_i\} + c\{Z_i\}$, and thus

$$\begin{aligned} C - \eta < c\{U_i\} &= c\{V_i\} = c\{W_i\} \\ &\leqq c\{Y_i\} + c\{Z_i\} \\ &\leqq C\left(1 - \frac{\epsilon}{2^k} + \eta\right) + \eta. \end{aligned}$$

Since $C \leqq 1$, we have $C(\epsilon/2^k) < 3\eta$ for every $\eta > 0$ which is impossible if $C > 0$. This proves the lemma.

3. Condition (A) implies condition (B).

THEOREM 3. *Let $\mathfrak{X}$ be a Banach space in which condition* (B) *fails. Then there is in $\mathfrak{X}$ a sequence $\{X_i\}$ of type* 2 *such that*

$$c\{X_i\} = 1.$$ [5]

PROOF. If $\mathfrak{X}$ fails to meet condition (B), then for every k and ϵ, there are k vectors $a_1, \cdots, a_k$ in the unit ball of $\mathfrak{X}$ such that $\|\pm a_1 \pm \cdots \pm a_k\| \geqq k(1-\epsilon)$ for every choice of signs. We now choose any sequences $\{\epsilon_i\}$ and $\{\delta_i\}$ of positive real numbers with $\epsilon_n \to 0$ and $\delta_n \to 0$. We now proceed as follows:

We choose an integer k_1 with $k_1 > (1-\delta_1)/\delta_1$ and a set of elements $a_1^{(1)}, \cdots, a_{k_1}^{(1)}$ such that

$$\|\pm a_1^{(1)} \pm a_2^{(1)} \pm \cdots \pm a_{k_1}^{(1)}\| \geqq k_1(1-\epsilon_1)$$

for all choices of signs. Then, for each $n > 1$, we set

$$m_n = \sum_{i=1}^{n-1} k_i$$

and choose

$$k_n > \frac{1-\delta_n}{\delta_n} m_n$$

and k_n elements $a_1^{(n)}, \cdots, a_{k_n}^{(n)}$ such that

$$\|\pm a_1^{(n)} \pm \cdots \pm a_{k_n}^{(n)}\| \geqq k_n(1-\epsilon_n)$$

[5] For definition of *type* 2, see proof of Lemma 2.

for all choices of signs. This gives us

$$\frac{k_n}{m_{n+1}} > 1 - \delta_n$$

and

$$\frac{m_n}{m_{n+1}} < \delta_n.$$

For any integer i, we have $m_j < i \leqq m_{j+1}$ for some value of j, *i.e.*, $i = m_j + r$, where $1 \leqq r \leqq k_j$. Define $b_i = a_r^{(j)}$. This gives us a sequence $\{b_i\}$ of elements of $\mathfrak{X}$. We define a sequence $\{X_i\}$ of random $\mathfrak{X}$-variables by requiring that the X_i be independent and that $\Pr\{X_i = b_i\} = \Pr\{X_i = -b_i\} = 1/2$. Then $\{X_i\}$ is a sequence of type 2, and for each j and each $\omega \in \Omega$, we have

$$\begin{aligned}\left\| \frac{1}{m_{j+1}} \sum_{i=1}^{m_{j+1}} X_i(\omega) \right\| &\geqq \left\| \frac{1}{m_{j+1}} \sum_{i=m_j+1}^{m_{j+1}} X_i(\omega) \right\| - \left\| \frac{1}{m_{j+1}} \sum_{i=1}^{m_j} X_i(\omega) \right\| \\ &\geqq \frac{1}{m_{j+1}} \left\| \pm a_1^{(j)} \pm a_2^{(j)} \pm \cdots \pm a_{k_j}^{(j)} \right\| - \frac{1}{m_{j+1}} m_j \\ &\geqq \frac{1}{m_{j+1}} k_j(1 - \epsilon_j) - \frac{m}{m_{j+1}} \\ &> (1 - \delta_j)(1 - \epsilon_j) - \delta_j.\end{aligned}$$

Thus, for every $\omega \in \Omega$,

$$\limsup_n \left\| \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \right\| = 1,$$

so that $c\{X_i\} = 1$, as required.

COROLLARY 4. *If $\mathfrak{X}$ is a Banach space, then the constant $C(\mathfrak{X})$, defined in the proof of Lemma* 2, *can take only the values* 0 *or* 1.

BIBLIOGRAPHY

1. Anatole Beck, *Une loi forte des grandes nombres dans des espaces de Banach uniformément convexes*, Ann. Inst. H. Poincaré **16** (1958), 35–45.

CORNELL UNIVERSITY AND
UNIVERSITY OF WISCONSIN