DEMONSTRATIO MATHEMATICA
Vol. XLII No 4 2009

Artur Bator, Wiesław Zięba

EXPECTATION IN METRIC SPACES AND CHARACTERIZATIONS OF BANACH SPACES

Abstract. We consider different definitions of expectation of random elements taking values in metric spaces. All such definitions are valid also in Banach spaces and in this case the results coincide with the Bochner integral. There may exist an isometry between considered metric space and some Banach space and in this case one can use the Bochner integral instead of expectation in metric space. We give some conditions which ensure existence of such isometry, for two different definitions of expectation in metric space.

1. Introduction

We are dealing with two definitions of expected value of random elements taking values in metric space. Both this definitions have the same property (which is not so obvious in metric spaces): the expectation of any integrable random element is a singleton. This is necessary property for the existence of isometry because Bochner integral in any Banach space is always a singleton.

Let $(\Omega, \mathcal{A}, P)$ be a non-atomic probability space. Assume that $(\mathbb{E}, d)$ is a complete metric space. By $\mathcal{F}$ we denote the Borel σ-field on $\mathbb{E}$.

We will use the following terminology concerning random elements:

- a map $X\colon \Omega \to \mathbb{E}$ will be called random element if for any $A \in \mathcal{F}$
$$X^{-1}(A) \in \mathcal{A}$$
- a random element X will be called simple random element if it takes only finite number of values
- a random element X will be called integrable if a real random variable $d(u, X)$ is integrable for some $u \in \mathbb{E}$
- the space of all integrable random elements will be denoted by $L^1_{\mathbb{E}}$
- the space of all square integrable random elements will be denoted by $L^2_{\mathbb{E}}$.

2000 *Mathematics Subject Classification*: 60B05.
Key words and phrases: Banach space, metric space, expectation.

First we introduce the following set called "metric combination"

$$\sum_{i=1}^{n} \lambda_i u_i = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n$$

which is given recursively in the following way:
For any finite system of nonnegative constants $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and for any subset $\{u_1, u_2, \ldots, u_n\}$ of $\mathbb{E}$:

- if $n = 1$ then $1u_1 = \{u_1\}$.
- Let $n > 1$. Suppose the sets $\sum_{i=1}^{k} \mu_i v_i$ are defined for $k < n$. We define:

$$u \in \sum_{i=1}^{n} \lambda_i u_i$$

if and only if there exist non-empty, disjoint and complementary subsets I_1, I_2 of the set of indices $\{1, 2, \ldots, n\}$ and elements

$$u^1 \in \sum_{i \in I_1} \lambda_i^1 u_i, \quad u^2 \in \sum_{i \in I_2} \lambda_i^2 u_i,$$

where $\lambda_i^1 = \lambda_i / \sum_{j \in I_1} \lambda_j$ i $\lambda_i^2 = \lambda_i / \sum_{j \in I_2} \lambda_j$ such that

$$d(u, u^1) = \Big(\sum_{i \in I_1} \lambda_i\Big) d(u^1, u^2), \quad d(u, u^2) = \Big(\sum_{i \in I_2} \lambda_i\Big) d(u^1, u^2).$$

Remark 1. "Metric combination" may be found in [4].

Next we give different, known convexity properties of metric spaces:

- $(\mathbb{E}, d)$ is called convex if for any $u, v \in \mathbb{E}$ and any $\lambda \in (0, 1)$ there is $w \in \mathbb{E}$ such that

$$w \in \lambda u + (1 - \lambda) v,$$

- $(\mathbb{E}, d)$ is called strictly convex if for any $u, v \in \mathbb{E}$ and any $\lambda \in (0, 1)$ there is exactly one element $w \in \mathbb{E}$ such that

$$\{w\} = \lambda u + (1 - \lambda) v,$$

- $(\mathbb{E}, d)$ is called externally convex if for any $u, v \in \mathbb{E}$ and any $\lambda \in (0, 1)$ there is $w \in \mathbb{E}$ such that

$$u \in \lambda w + (1 - \lambda) v.$$

Using this basic definitions we will be dealing with two different kinds of expectation:

- convex combination expectation
- Fréchet expectation.

First definition was given by Teran and Molchanov in [6] the second one is due to Fréchet [2].

The base for our results is the following known theorem:

THEOREM 1. (Andalafte et al. [1]) *Let* $(\mathbb{E}, d)$ *be complete, convex and externally convex metric space. The space is isometric with a strictly convex real Banach space if and only if for any triplet* a_1, a_2, a_3 *the set* $\sum_{i=1}^{3} \frac{1}{3} a_i$ *is a singleton in* $\mathbb{E}$.

2. Convex combination expectation

We will start with some facts concerning convex combination operation and definition of expectation given in [6].

2.1. Definition of convex combination expectation

On $(\mathbb{E}, d)$ introduce a *convex combination operation* which for all $n \geq 2$, numbers $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ satisfying $\sum_{i=1}^{n} \lambda_1 = 1$, and all $u_1, u_2, \ldots, u_n \in \mathbb{E}$ produces an element of $\mathbb{E}$ denoted by $[\lambda_i u_i]_{i=1}^{n}$ or $[\lambda_1 u_1; \lambda_2 u_2; \ldots; \lambda_n u_n]$. Assume that $[1u] = u$ for every $u \in \mathbb{E}$.

Assume that convex combination operation satisfy the following conditions:

(i) $[\lambda_i u_i]_{i=1}^{n} = [\lambda_{\sigma(i)} u_{\sigma(i)}]_{i=1}^{n}$ for any permutation σ of $1, 2, \ldots, n$;

(ii) $[\lambda_i u_i]_{i=1}^{n+2} = [\lambda_1 u_1; \lambda_2 u_2; \ldots; (\lambda_{n+1} + \lambda_{n+2})[\frac{\lambda_{n+j}}{\lambda_{n+1}+\lambda_{n+2}} u_{n+j}]_{j=1}^{2}]$;

(iii) for any sequence of numbers $\lambda^{(k)} \to \lambda \in (0,1)$; $k \to \infty$

$$[\lambda^{(k)} u; \left(1 - \lambda^{(k)}\right) v] \to [\lambda u; (1-\lambda) v]; \; k \to \infty;$$

(iv) $\forall(\lambda \in (0,1)) \; \forall(u_1, u_2, v_1, v_2 \in \mathbb{E};)$

$$d([\lambda u_1; (1-\lambda) u_2], [\lambda v_1; (1-\lambda) v_2]) \leq \lambda d(u_1, v_1) + (1-\lambda) d(u_2, v_2).$$

The following properties follow from (i)-(iv)

1) for every $u_1, u_2, \ldots, u_{nm} \in \mathbb{E}$ and $\lambda_1, \ldots, \lambda_n, \alpha_1, \ldots, \alpha_m > 0$ with $\sum_{i=1}^{n} \lambda_i = \sum_{j=1}^{m} \alpha_j = 1$,

$$[\lambda_i [\alpha_j u_{ij}]_{j=1}^{m}]_{i=1}^{n} = [\lambda_i \alpha_j u_{ij}]_{j=1, i=1}^{j=m, i=n},$$

2) the convex combination operation is jointly continuous in its $2n$ arguments.

The following assumption concerns the limiting behaviour of the convex combination operation:

(v) for each $u \in \mathbb{E}$, there exists $\lim_{n \to \infty} [n^{-1} u]_{i=1}^{n}$, which will be denoted by Ku.

For the ease of reference, a separable, complete metric space $\mathbb{E}$ with the convex combination operation satisfying conditions (i)-(v) will be called convex combination space.

Introduce some more terminology which will be useful later on.

- A point $u \in \mathbb{E}$ is called convexely decomposable if

$$u = [\lambda_i u]_{i=1}^n$$

  for all $n \geq 2$ and $\lambda_1, \ldots, \lambda_n > 0$ with $\sum \lambda_i = 1$.
- Convex combination is called unbiased if $Kx = x$ for all $x \in \mathbb{E}$.
- The metric space $(\mathbb{E}, d)$ is called convexifiable if it admits an unbiased convex combination.

Convexification operator K defined by (v) has the following properties:

(a) the operator K is linear, that is

$$K[\lambda_j u_j]_{j=1}^m = [\lambda_j K u_j]_{j=1}^m,$$

(b) the image $K(\mathbb{E})$ of $\mathbb{E}$ under K coincides with the family of convexely decomposable elements of $\mathbb{E}$,

(c) the operator K is idempotent on $\mathbb{E}$, that is $K(Ku) = Ku$,

(d) for any numbers $\lambda_1 + \lambda_2 + \lambda_3 = 1;\ \ \lambda_i \geq 0;\ \ i = 1, 2, 3;$

$$[\lambda_1 u; \lambda_2 Kv; \lambda_3 Kv] = [\lambda_1 u; (1 - \lambda_1) Kv],$$

(e) K is non-expansive with respect to the metric d, that is

$$d(Ku; Kv) \leq d(u, v).$$

Now we are ready to give the definition of expectation.

Definition 1. For simple, integrable random element X taking values $u_1, u_2, \ldots, u_n$ with probabilities $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively, define expectation of X by

$$EX = [\lambda_i, Ku_i]_{i=1}^n.$$

Remark 2. Any integrable random element may be approximated by simple random element and this definition may be extended on $L^1_{\mathbb{E}}$. (For details see [6]).

2.2. Characterization

Now we are in position to state the following result:

Theorem 2. *Let $(\mathbb{E}, d)$ be complete, strictly convex and externally convex metric space. If $\mathbb{E}$ is convexifiable then it is isometric with a strictly convex real Banach space.*

Proof. We know that $Ku = u$ for any $u \in \mathbb{E}$.

Furthermore by property (iv) for any $u, v \in \mathbb{E}$ and any $\lambda \in (0,1)$ we have:

$$\begin{aligned} d\left([\lambda u; (1-\lambda)v]\,; u\right) &= d\left([\lambda u; (1-\lambda)v]\,; Ku\right) \\ &= d\left([\lambda Ku; (1-\lambda)Kv]\,; [\lambda Ku; (1-\lambda)Ku]\right) \\ &\leq (1-\lambda)d(u,v). \end{aligned}$$

Similarly one can show that $d\left([\lambda u; (1-\lambda)v]\,; v\right) \leq \lambda d(u,v)$.

Note that by triangle inequality:

$$\begin{aligned} d(u;v) &\leq d\left(u; [\lambda u; (1-\lambda)v]\right) + d\left([\lambda u; (1-\lambda)v]\,; v\right) \\ &\leq (1-\lambda)d(u;v) + \lambda d(u,v) = d(u,v). \end{aligned}$$

So it is clear that

$$d\left([\lambda u; (1-\lambda)v]\,; v\right) = \lambda d(u,v) \quad \text{and} \quad d\left([\lambda u; (1-\lambda)v]\,; u\right) = (1-\lambda)d(u,v).$$

Now by strict convexity metric combination of two points is also a singleton. Thus we have just proved that for any $u, v \in \mathbb{E}$

$$[\lambda u; (1-\lambda)v] = \lambda u + (1-\lambda)v,$$

where the right hand side is metric combination operation introduced above.

Now note that the metric combination of three points $\frac{1}{3}u_1 + \frac{1}{3}u_2 + \frac{1}{3}u_3$ may consist three different points of the form $\frac{1}{3}u_i + \frac{2}{3}(\frac{1}{2}u_j + \frac{1}{2}u_k)$ for different choices of $i, j, k \in \{1,2,3\}$. But on the other hand by ii):

$$\frac{1}{3}u_i + \frac{2}{3}\left(\frac{1}{2}u_j + \frac{1}{2}u_k\right) = \left[\frac{1}{3}u_i; \frac{2}{3}\left[\frac{1}{2}u_j; \frac{1}{2}u_k\right]\right] = \left[\frac{1}{3}u_i; \frac{1}{3}u_j; \frac{1}{3}u_k\right]$$

for any choice of $i, j, k \in \{1,2,3\}$.

We have just proved that two mentioned operations: convex combination and metric combination are actually the same. We can use Theorem 1 to finish the proof. ∎

3. Fréchet expectation

3.1. Definition of Fréchet expectation

Let us consider expectation defined as an element of a metric space $(\mathbb{E}, d)$ minimizing "variance" i.e.

$$E_F X = \{a \in \mathbb{E} : \; Ed^2(a,X) = \min_{u \in \mathbb{E}} Ed^2(u,X)\}.$$

This expectation does not need to be a singleton, it may be a set. There are however spaces in which this expectation is always a singleton for any square integrable random element. The spaces are known as global NPC spaces. Here we recall the definition:

Definition 2. $(\mathbb{E}, d)$ is called a global NPC space if and only if it is a complete metric space with nonpositive curvature in the following sense:

$$\inf_{z\in\mathbb{E}} \int_{\mathbb{E}} d^2(x,z)q(dx) \le \frac{1}{2}\int_{\mathbb{E}}\int_{\mathbb{E}} d^2(x,y)q(dx)q(dy)$$

for all discrete probability measures q on $\mathbb{E}$.

The following characterization of global NPC spaces is also known:

Proposition 1. (Sturm [5]) *A complete metric space $(\mathbb{E}, d)$ is a global NPC space if and only if:*

i) *it is a geodesic space, that is, any two points $\gamma_0, \gamma_1 \in \mathbb{E}$ can be joined by a (continuous) curve $\gamma : [0,1] \to \mathbb{E}$ such that $d(\gamma_0,\gamma_1) = l_d(\gamma)$ where the length of γ is defined as*

$$l_d(\gamma) := \sup\Big\{\sum_{k=1}^{n} d(\gamma_{t_k}, \gamma_{t_{k+1}}) : 0 \le t_0 \le t_1 \le \ldots \le t_n \le 1, n \in \mathbb{N}\Big\},$$

ii) *and it has nonpositive curvature in the sense of A. D. Alexandrov, which means that for every point z, every geodesic $t \mapsto \gamma_t$ (parametrized proportionally to arclength, as always) and every $t \in [0,1]$,*

$$d^2(z,\gamma_t) \le (1-t)d^2(z,\gamma_0) + td^2(z,\gamma_1) - t(1-t)d^2(\gamma_0,\gamma_1).$$

Proposition 2. *Note that global NPC space is always strictly convex.*

Proof. The convexity of the space is clear from i) because geodesic joining two points $u, v \in \mathbb{E}$ consists of the points $\lambda u + (1-\lambda)v$. Assume that for some points $u, v \in \mathbb{E}$ there are two geodesic γ, γ' joining this two points but by negative curvature property ii) taking $z = \gamma'_t$ we have:

$$\begin{aligned} d^2(\gamma'_t,\gamma_t) &\le (1-t)d^2(\gamma'_t,\gamma_0) + td^2(\gamma'_t,\gamma_1) - t(1-t)d^2(\gamma_0,\gamma_1) \\ &= (1-t)t^2d^2(\gamma_0,\gamma_1) + t(1-t)^2d^2(\gamma_0,\gamma_1) - t(1-t)d^2(\gamma_0,\gamma_1) = 0. \end{aligned}$$

This contradiction ends the proof. ∎

Remark 3. Note that for the Fréchet expectation the following property:

$$(*) \qquad E_F(X) = E_F(E_F(X|Y))$$

does not hold in general even for random elements in global NPC space. However we know, that if we use this definition for any given Hilbert space it coincides with Bochner integral. So condition $(*)$ is true in this case.

For detailed discussion of Fréchet expectation in the global NPC spaces see [5].

3.2. Characterization

We will start with the following easy result:

Lemma 1. *Let $(\mathbb{E}, d)$ be a global NPC metric space. Consider random element X taking only two values u and v with probabilities λ and $1-\lambda$ respectively. Let γ_t be a geodesic joining u and v. Then $E_F X = \gamma_\lambda$.*

Proof. Suppose that $E_F X = \gamma_t$ for some $t \in (0,1)$. Note first that

$$d(\gamma_t, \gamma_0) = (1-t)d(\gamma_0, \gamma_1) = (1-t)D,$$
$$d(\gamma_t, \gamma_1) = td(\gamma_0, \gamma_1) = tD,$$

where $D = d(u,v)$. And we have

$$\begin{aligned} Ed^2(X, E_F X) &= P(X=u)d^2(E_F X, u) + P(X=v)d^2(E_F X, v) \\ &= \lambda d^2(\gamma_t, \gamma_0) + (1-\lambda)d^2(\gamma_t, \gamma_1) \\ &= \lambda(1-t)^2 D^2 + (1-\lambda)t^2 D^2 = D^2\left(t^2 - 2\lambda t + \lambda\right) = f(t). \end{aligned}$$

The function $f(t)$ reaches its minimum for $t = \lambda$. So it is true that if we assume that $E_F X \in \gamma$ then $E_F X = \gamma_\lambda$.

Assume that $E_F X \notin \gamma$ then we can write

$$d(E_F X, u) = (1-t+a)d(u,v) = (1-t+a)D,$$
$$d(E_F X, v) = td(u,v) = tD,$$

where $a > 0$. So we obtain:

$$\begin{aligned} Ed^2(X, E_F X) &= P(X=u)d^2(E_F X, u) + P(X=v)d^2(E_F X, v) \\ &= \lambda(1-t+a)^2 D^2 + (1-\lambda)t^2 D^2 \\ &= D^2\left(t^2 - 2\lambda t + \lambda\right) + D^2\left(a^2\lambda + 2a\lambda(1-t)\right) \\ &= f(t) + C(t) = F(t) \end{aligned}$$

and it is enough to note that $C(t) > 0$ so minimum can not be obtained anywhere besides γ. ■

Theorem 3. *Let E_F be a Fréchet expectation operator defined on externally convex, global NPC space $(\mathbb{E}, d)$. If the condition $E_F X = E_F(E_F(X|Y))$ is satisfied for any square integrable random element X and any random element Y taking two values then $(\mathbb{E}, d)$ is isometric with some strictly convex real Banach space.*

Proof. It is enough to show that for any random variable X taking values u_1, u_2, u_3 with probabilities $\lambda_1, \lambda_2, \lambda_3$ respectively the following equation holds:

$$E_F X = \sum_{i=1}^{3} \lambda_i u_i.$$

Again the right hand side is understood as metric combination.

Put $v_i = \frac{\lambda_j}{1-\lambda_i} u_j + \frac{\lambda_k}{1-\lambda_i} u_k$ for i, j, k chosen from the set $\{1, 2, 3\}$ and define for $i = 1, 2, 3$ the random variables

$$Y_i(\omega) = \begin{cases} u_i, & X = u_i; \\ v_i, & X \neq u_i. \end{cases}$$

Note that by the last lemma we have

$$\begin{aligned} E_F X &= E_F(E_F(X|Y_i)) \\ &= \lambda_i E_F(X|X = u_i) + (1 - \lambda_i) E_F(X| \neq u_i) \\ &= \lambda_i u_i + (1 - \lambda_i) v_i, \end{aligned}$$

for $i = 1, 2, 3$. But on the other hand by definition of metric combination operation we get

$$\sum_{i=1}^{3} \lambda_i u_i = \bigcup_{i=1}^{3} \{\lambda_i u_i + (1 - \lambda_i) v_i\} = E_F X$$

and $E_F X$ is always a singleton in global NPC spaces.

Thus conditions of Theorem 1 are satisfied and $(\mathbb{E}, d)$ is isometric with some strictly convex real Banach space. ∎

References

[1] E. Z. Andalafte, J. E. Valentine, S. G. Wayment, *Triangle median properties which characterize Banach spaces*, Houston J. Math. 5, 3 (1979), 307–312.

[2] M. Fréchet, *On definition of expectation for any random element*, Giorn. Inst. Ital. Att. 19 (1956), 1–15.

[3] W. Herer, *Mathematical expectation and martingales of random subsets of a metric space*, Prob. Math. Statist. 11, 2 (1991), 291–304.

[4] W. Herer, *Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature*, Prob. Math. Statist. 13, 2 (1992), 59–70.

[5] K. T. Sturm, *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, Ann. Probab. 30 (2002), 1195–1222.

[6] P. Teran, I. Molchanov, *The law of large numbers in a mertic space with a convex combination operation*, J. Theoret. Probab. 19, 4 (2006), 875–897.

INSTITUTE OF MATHEMATICS
MARIE CURIE–SKŁODOWSKA UNIVERSITY
pl. Marii Curie–Skłodowskiej 1
20–031 LUBLIN, POLAND
E-mails: Artur.Bator@umcs.lublin.pl;
Wieslaw.Zieba@umcs.lublin.pl

Received October 22, 2008; revised version January 29, 2009.