RESEARCH

Open Access

CrossMark

An axiomatic integral and a multivariate mean value theorem

Milan Merkle[*]

[*]Correspondence: emerkle@etf.rs
Department of Applied
Mathematics, Faculty of Electrical
Engineering, University of Belgrade,
Bulevar Kralja Aleksandra 73,
Belgrade, 11120, Serbia

Abstract

In order to investigate minimal sufficient conditions for an abstract integral to belong to the convex hull of the integrand, we propose a system of axioms under which it happens. If the integrand is a continuous $\mathbf{R}^n$-valued function over a path-connected topological space, we prove that any such integral can be represented as a convex combination of values of the integrand in at most n points, which yields an ultimate multivariate mean value theorem.

MSC: Daniell integral; convex hull; finitely additive measure; integral means

Keywords: 28A30; 26E60; 26B25

1 Introduction and motivation

The basic integral mean value theorem states that for a function X which is continuous on the interval $[a, b]$, there exists a point $t \in (a, b)$ such that

$$\frac{1}{b-a} \int_a^b X(s)\, \mathrm{d}s = X(t). \tag{1}$$

To show that the assumption of continuity is crucial for validity of this theorem, we can take the interval $[-1, 1]$ and define $X(s) = -1$ for $s \in [-1, 0)$ and $X(s) = 1$ for $s \in [0, 1]$. Hence here we do not have a single point $t \in (-1, 1)$ for (1) to be satisfied. However, we can achieve a similar result with a convex combination of values of f in two points:

$$\frac{1}{b-a} \int_a^b X(s)\, \mathrm{d}s = \frac{1}{2} X(t_1) + \frac{1}{2} X(t_2), \quad t_1 \in (-1, 0), t_2 \in (0, 1). \tag{2}$$

It turns out that the difference of one extra point for non-continuous functions remains in a much more general case and for a very broad class of integrals in higher dimensions. This is the topic of this article.

In multivariate case, with $X \in \mathbf{R}^n$, $n \geq 1$, there is an old and seemingly forgotten mean value theorem by Kowalewski [1, 2] as follows.

Theorem A [1] *Let* $x_1, \ldots, x_n$ *be continuous functions in a variable* $t \in [a, b]$. *There exist real numbers* $t_1, \ldots, t_n$ *in* $[a, b]$ *and non-negative numbers* $\lambda_1, \ldots, \lambda_n$, *with* $\sum_{i=1}^n \lambda_i = b - a$,

such that

$$\int_a^b X_k(t)\,\mathrm{d}t = \lambda_1 X_k(t_1) + \cdots + \lambda_n X_k(t_n) \quad \text{for each } k = 1, 2, \ldots, n,$$

or in more compact notation,

$$\int_a^b X(s)\,\mathrm{d}s = \lambda_1 X(t_1) + \cdots + \lambda_n X(t_n), \tag{3}$$

where $X = (X_1, \ldots, X_n)$.

A recent generalization of Theorem A is proved in [3] using the following modification of classical Carathéodory's convex hull theorem.

Theorem B [3] *Let* $C : s \mapsto X(s)$, $s \in I$, *be a continuous curve in* $\mathbf{R}^n$, *where* $I \subset \mathbf{R}$ *is an interval, and let* K *be the convex hull of the curve* C. *Then each* $v \in K$ *can be represented as a convex combination of* n *or fewer points of the curve* C.

In this way, a more general mean value theorem for n-dimensional functions is obtained directly from the fact that the normalized integral should belong to the convex hull of the set of values of the integrand. This is proved in the following theorem for Lebesgue integral with a probability measure.

Theorem C ([3], Lemma 1) *Let* $(S, \mathcal{F}, \mu)$ *be a probability space, and let* $X_i : S \to \mathbf{R}$, $i = 1, \ldots, n$, *be* μ-*integrable functions. Let* $X(t) = (X_1(t), \ldots, X_n(t))$ *for every* $t \in S$. *Then* $\int_S X(t)\,\mathrm{d}\mu(t) \in \mathbf{R}^n$ *is in the convex hull of the set* $X(S) = \{X(t) \mid t \in S\} \subset \mathbf{R}^n$.

Finally, the main result of [3] reads as follows.

Theorem D ([3], Theorem 2) *For an interval* $I \subseteq \mathbf{R}$, *let* μ *be a finite positive measure on the Borel sigma-field of* I. *Let* X_k, $k = 1, \ldots, n$, $n \geq 1$, *be continuous functions on* I *integrable on* I *with respect to the measure* μ. *Then there exist points* $t_1, \ldots, t_n$ *in* I *and non-negative numbers* $\lambda_1, \ldots, \lambda_n$, *with* $\sum_{i=1}^n \lambda_i = \mu(I)$, *such that*

$$\int_I X_k(s)\,\mathrm{d}\mu(s) = \sum_{i=1}^n \lambda_i X_k(t_i), \quad k = 1, \ldots, n. \tag{4}$$

Let us note that without the continuity assumption we still may use Carathéodory's convex hull theorem which would yield (4) with $n + 1$ points t_i and the same number of λ_i's, which shows that the example at the beginning of the text well describes the general situation in $\mathbf{R}^n$.

In this paper we give a more general theorem of the type (4), tracing the steps of the proof in [3] in a much more general context. In Section 3 we show that a result like (4) holds if the integral over I is replaced with a general linear functional on some function space, under a system of axioms, whereas the interval I can be replaced by a topological space which is path connected.

To reach this goal, we need to extend Theorem C. In Section 2 we show that Theorem C holds for any linear functional which satisfies a condition slightly stronger than positivity, and where functions X_k are defined over an arbitrary non-empty set.

Applications of such a very general mean value theorem are numerous, and we are not discussing particular applications in this paper. Let us just mention that as shown in [3], the theorems of this type can be considered as an aid to construct quadrature rules or their approximative versions. See also a recent paper [4] for another application related to integrals. Another advantage of the approach presented in this paper is that the results are widely applicable to different kinds of integrals treated as linear functionals over some space of functions.

2 Axioms and their consequences

We start with an arbitrary non-empty set S with an algebra $\mathcal{F}$ (may be a sigma algebra as well) of its subsets. Therefore, $\mathcal{F}$ contains S and if a set A is a result of finitely many set operations over sets in $\mathcal{F}$, then $A \in \mathcal{F}$.

Let $\mathcal{S}$ be a family of functions $X : S \to \mathbf{R}$ which satisfies the following conditions:

(C1) If $X_1, X_2 \in \mathcal{S}$, then $aX_1 + bX_2 \in \mathcal{S}$ for all $a, b \in \mathbf{R}$.
(C2) For $B \in \mathcal{F}$, the indicator function $I_B(\cdot)$ belongs to $\mathcal{S}$.
(C3) For $X \in \mathcal{S}$ and any interval J, the set $\{s \in S : X(s) \in J\}$ is in $\mathcal{F}$.
(C4) For $X \in \mathcal{S}$ such that $X(s) \geq 0$ for all $s \in S$, it holds that $X \cdot I_{X \in J} \in \mathcal{S}$ for any interval J.

Note that from (C1) and (C2) it follows that all constants are in $\mathcal{S}$. Let us also note that functions in $\mathcal{S}$ are not assumed to be bounded.

Let E be a functional defined on $\mathcal{S}$ and taking values in $\mathbf{R}$ such that the following axioms hold.

(A1) $\mathrm{E}c = c$ for any constant c;
(A2) $\mathrm{E}(\sum_{i=1}^{m} \alpha_i X_i) = \sum_{i=1}^{m} \alpha_i \mathrm{E}(X_i)$ for $a_j \in \mathbf{R}$ and for $X_i \in \mathcal{S}$, $i = 1, \ldots, m$.

Now we may define a set function P as

$$P(B) = \mathrm{E}\big(I_B(\cdot)\big), \quad B \in \mathcal{F}, \tag{5}$$

and consider yet another condition related to P:

(C5) If $X \in \mathcal{S}$ and $P(N) = 0$, then $X \cdot I_N \in \mathcal{S}$, and $\mathrm{E}(X \cdot I_N) = 0$.

The last axiom that we propose is as follows.

(A3) For $X \in \mathcal{S}$, if $\mathrm{E}X = 0$, then either $P(X = 0) = 1$ or there exist $s_1, s_2 \in S$ such that $X(s_1)X(s_2) < 0$,

or equivalently (see Lemma 2.3 below)

(A3′) For $X \in \mathcal{S}$, if $X(s) \geq 0$ for all $s \in S$ and $\mathrm{E}X = 0$, then $P(X = 0) = 1$.

Finally we extend the functional E to act on functions with values in $\mathbf{R}^n$. Let $X = (X_1, \ldots X_n)$ be a function from S to $\mathbf{R}^n$, where we assume that X_i, $i = 1, \ldots, n$, satisfy axioms and conditions above, then we define

$$\mathrm{E}(X) := (\mathrm{E}X_1, \ldots, \mathrm{E}X_n).$$

The central result of this section is Theorem 2.10, where we show that under (A1)-(A2)-(A3), E(X) belongs to the convex hull of $X(S)$.

A similar axiomatic approach is applied in Daniell's integral, and there are other axiomatic systems in the literature for different purposes like in [5] for Riemann integrals in connection to evaluating the length of a curve, general means in [6], finitely additive probabilities (FAPs) in [7] and applications in a recent article [8]. The system of axioms applied in this article differs from others in the literature in the conditions that allow non-absolute integrals, as well as in axiom (A3), which is a slightly stronger condition than the usual positivity. The reason for introducing this system of axioms is that it provides conditions under which EX belongs to the convex hull of $X(S)$ (Theorem 2.10) independently of the kind of integrals that are considered.

Now we are going to derive some additional properties as consequences from the axioms.

Lemma 2.1 *Under the system of axioms* (A1)-(A2)-(A3) *or* (A1)-(A2)-(A3′), *assuming also conditions* (C1)-(C2), *the set function P defined on $\mathcal{F}$ with* (5) *is a finitely additive probability on* $(S, \mathcal{F})$.

Proof Since $I_S(s) = 1$ for all $s \in S$, we have that $P(S) = 1$. For disjoint sets $B_1, \ldots, B_m$, using (A2) we get additivity

$$P\left(\bigcup_{i=1}^m B_i\right) = E\left(\sum_{i=1}^m I_{B_i}\right) = \sum_{i=1}^m P(B_i).$$

Let us now show that $P(B) \geq 0$ for all $B \in \mathcal{F}$. Indeed, suppose that $P(B) = -\varepsilon$ for some $\varepsilon > 0$. This implies (by (A1) and (A2)) that $E(I_B + \varepsilon) = 0$; on the other hand, $I_B(s) + \varepsilon > 0$ for all $s \in S$, which contradicts both (A3) and (A3′), and this ends the proof. $\square$

From now on, the quintuplet $(S, \mathcal{F}, \mathcal{S}, E, P)$ will be assumed to be as defined above in the framework of axioms (A1)-(A2)-(A3) and conditions (C1)-(C5) (if not specified differently). The letter X will be reserved for elements of $\mathcal{S}$, and P is a set function derived from E as in (5).

Lemma 2.2 (Positivity) *If for all $s \in S$, $X(s) \geq 0$, then $EX \geq 0$.*

Proof Suppose $X(s) \geq 0$ for all $s \in S$, and $EX = -\varepsilon$ for some $\varepsilon > 0$. Then, by (A1) and (A2), $E(X + \varepsilon) = 0$, whereas $X + \varepsilon > 0$. This contradicts (A3). Therefore, if $X \geq 0$, then $EX \geq 0$. $\square$

Lemma 2.3 *Assuming that* (A1) *and* (A2) *hold, axioms* (A3) *and* (A3′) *are equivalent.*

Proof In Lemma 2.1 we already proved the property that P is a FAP follows with either (A3) or (A3′), so we may use that property in both parts of the present proof.

Assume that (A1)-(A2)-(A3) hold and suppose that for all $s \in S$, $X(s) \geq 0$ and $EX = 0$. Then by (A3) it follows that $P(X = 0) = 1$. Therefore, (A3′) holds.

Now assume that (A1)-(A2)-(A3′) hold, but not (A3). Then there exists $X \in \mathcal{S}$ such that: (a) $EX = 0$, $P(X = 0) < 1$, $X \geq 0$, or (b) $EX = 0$, $P(X = 0) < 1$, $X \leq 0$. In the case (a), using

(A3′) we find that $P(X = 0) = 1$, which is a contradiction to $P(X = 0) < 1$. The case (b) can be reduced to (a) with the function $Y = -X$. □

Due to the equivalence established in Lemma 2.3, in the rest of the article we refer to (A3) as being either (A3) or (A3′).

Remark 2.4 Suppose that we have a quintuplet $(S, \mathcal{F}, \mathcal{S}, \mathrm{E}, P)$ that satisfies assumptions (C1)-(C5) and axioms (A1)-(A2), where P is defined with (5). Let $S^* \in \mathcal{F}$ be a subset of S with $P(S^*) > 0$ such that $X \cdot I_S^* \in \mathcal{S}$. Define $\mathcal{F}^* = \{B \cap S^* \mid B \in \mathcal{F}\}$. Next, for each $X \in \mathcal{S}$, we define the function $X^* := X|_{S^*}$ - that is, X restricted to S^*, and let $\mathcal{S}^*$ be the space of all those mappings. We define a linear functional E^* on $\mathcal{S}$ as

$$\mathrm{E}^*\left(X^*\right) = \frac{1}{P(S^*)} \mathrm{E}(X \cdot I_{S^*}), \quad X^* = X|_{S^*},$$

and the set function P^* as

$$P^*\left(B^*\right) = \mathrm{E}^*(I_{B^*}) = \frac{P(B^*)}{P(S^*)} = \frac{P(B \cap S^*)}{P(S^*)}, \quad B^* = B \cap S^* \in \mathcal{F}^*.$$

In this way we get a new quintuplet with $(S^*, \mathcal{F}^*, \mathcal{S}^*, \mathrm{E}^*, P^*)$, and it is not difficult to see that the new quintuplet inherits conditions (C1)-(C5) and axioms (A1)-(A2) as well as (A3) if it is satisfied with the original quintuplet.

In the next lemma we prove Markov's inequality from the axioms.

Lemma 2.5 *Let $X \in \mathcal{S}$ and $X(s) \geq 0$ for all $s \in S$. Then*

$$P(X > \varepsilon) \leq \frac{\mathrm{E}X}{\varepsilon}, \quad \varepsilon > 0. \tag{6}$$

Proof Since $X \geq 0$, we use (C4) to conclude that

$$\mathrm{E}X = \mathrm{E}(X \cdot I_{0 \leq X \leq \varepsilon}) + \mathrm{E}(X \cdot I_{X > \varepsilon}).$$

Further, we have

$$X(s) \cdot I_{0 \leq X \leq \varepsilon}(s) \geq 0, \qquad X(s) \cdot I_{X > \varepsilon}(s) \geq \varepsilon I_{X > \varepsilon}(s),$$

and then use positivity (Lemma 2.2) and (A1)-(A2) to conclude that $\mathrm{E}X \geq \varepsilon P(X > \varepsilon)$. □

Lemma 2.6 *Assuming* (A1)-(A2), (C1)-(C5) *and Markov's inequality,* (A3) *holds if P is countably additive probability.*

Proof Let $X \in \mathcal{S}$ such that $X \geq 0$ and $\mathrm{E}X = 0$. We need to show that $P(X > 0) = 0$. By Markov's inequality we have that $P(X > \varepsilon) = 0$ for any $\varepsilon > 0$, and so, using the countable additivity, we get

$$P(X > 0) = P\left(\bigcup_{n=1}^{\infty} \left\{X > \frac{1}{n}\right\}\right) = \lim_{n \to +\infty} P\left(X > \frac{1}{n}\right) = 0$$

as desired. □

Remark 2.7 Consider the case where P is a countably additive probability on $(S, \mathcal{F})$, $\mathcal{F}$ is a sigma algebra, and $EX = \int_S X(s)\, dP(s)$. Axioms (A1)-(A2) and conditions (C1)-(C5) are clearly satisfied, and Markov's inequality can be proved from properties of the integral, so by Lemma 2.6, axiom (A3) also holds.

Let us now recall some facts about FAPs. A probability P which is defined on an algebra $\mathcal{F}$ of subsets of the set S is *purely finitely additive* if $\nu \equiv 0$ is the only countably additive measure with the property that $\nu(B) \le P(B)$ for all $B \in \mathcal{F}$. A purely finitely additive probability P is *strongly finitely additive-SFAP* if there exist countably many disjoint sets $H_1, H_2, \ldots \in \mathcal{F}$ such that

$$\bigcup_{i=1}^{+\infty} H_i = S \quad \text{and} \quad P(H_i) = 0 \quad \text{for all } i. \tag{7}$$

For every probability P on $\mathcal{F}$ there exists a countably additive probability P_c and a purely finitely additive probability P_d such that $P = \lambda P_c + (1 - \lambda) P_d$ for some $\lambda \in [0, 1]$. This decomposition is unique (except for $\lambda = 0$ or $\lambda = 1$, when it is trivially non-unique). For more details see [9].

Lemma 2.8 *Assuming axioms* (A1)-(A2), *conditions* (C1)-(C5) *and positivity, if P is a SFAP, the condition of axiom* (A3) *is not satisfied.*

Proof Let P be a SFAP, and let H_i be a partition of S as in (7). Define $X(s) = 1/i$ if $s \in H_i$. Further,

$$X(s) \le 1 \cdot I_{H_1}(s) + \frac{1}{2} \cdot I_{H_2}(s) + \cdots + \frac{1}{k} \cdot I_{\{H_k \cup H_{k+1} \cup \cdots\}}$$

and so (by positivity) $0 \le EX \le 1/k$ for every $k > 0$, hence $EX = 0$. This contradicts (A3).
□

Example 2.9 Let $S = [0, +\infty)$ and let P be the probability defined by the non-principal ultrafilter of Banach limit as $s \to +\infty$. Let $X(s) = e^{-s}$. Then $X \ge 0$ and $EX = 0$, but $P(X = 0) = 0$. In this case the convex hull $K(X) = (0, 1]$ and $EX \notin K(X)$.

Theorem 2.10 *Let $(S, \mathcal{F}, \mathcal{S}, E, P)$ be a quintuplet as defined above, and let $X = (X_1, \ldots, X_n)$, where $X_i \in \mathcal{S}$ for all i. Assuming that axioms* (A1)-(A2)-(A3) *and conditions* (C1)-(C5) *hold, EX belongs to the convex hull of the set $X(S) = \{X(t) \mid t \in S\} \subset \mathbf{R}^n$.*

Proof Without loss of generality we may assume that for all i, $EX_i = 0$ (otherwise, if $EX_i = c_i$, we can observe $E(X_i - c_i) = 0$). Let K denote the convex hull of the set $X(S) \in \mathbf{R}^n$. We now prove that $0 \in K$ by induction on n. Let $n = 1$. By (A3), $EX = 0$ implies that either $X(s) = 0$ for some $s \in S$ or there are $s_1, s_2 \in S$ such that $X(s_1) > 0$ and $X(s_2) < 0$. In both cases it follows that $0 \in K$.

Now assume that the statement of the theorem is valid for all dimensions from 1 to $n - 1$ for all quintuplets $(S, \mathcal{F}, \mathcal{S}, E, P)$ that satisfy the conditions mentioned in the statement of the theorem. Let now X be a vector function with values in $\mathbf{R}^n$.

If every hyperplane π that contains 0 has the property that the set $X(S)$ has a non-empty intersection with both of two open half-spaces with π as boundary, then $0 \in K$ (see [10] for details). Otherwise, suppose that

$$L(s) := \sum_{k=1}^{n} a_k X_k(s) \geq 0 \quad \text{for every } s \in S, \tag{8}$$

with some real numbers $a_1, \ldots, a_n$ such that $\sum a_k^2 > 0$. By linearity (A2) we have that $\mathrm{E}L(t) = 0$, which is (by (A3')) possible together with (8) only if $L(t) = 0$ for all $t \in S \setminus N$, where $\mu(N) = 0$. Assuming that $a_n \neq 0$, we find that

$$X_n(s) = -\sum_{k=1}^{n-1} \frac{a_k}{a_n} X_k(s) \quad \text{for every } s \in S \setminus N. \tag{9}$$

In other words, a separating hyperplane exists only if there exists a linear relation among n given functions with probability one. In order to eliminate X_n and to reduce the system to $n - 1$ functions, we consider functions $X_i^*(s) = X_i(s)$ on the restricted domain $S^* = S \setminus N$ ($i = 1, \ldots, n - 1$) and the corresponding functional E^*. Let K^* be convex hull of $X^*(S^*) \in \mathbf{R}^{n-1}$.

By hereditary property (Remark 2.4), we have that $\mathrm{E}^*(X_i^*) = \mathrm{E}(X_i \cdot I_{S \setminus N}) = 0$ by (C4). Note that $K^* \subset K$. By induction assumption, the statement of the theorem holds for dimension $n - 1$, and so

$$\sum_{i=1}^{m} \lambda_i X_k(t_i) = 0, \quad k = 1, \ldots, n - 1, \tag{10}$$

with some $t_1, \ldots, t_m \in S^* \subset S$ (here we use the fact that $X_i^*(s) = X_i(s)$ for $s \in S^*$). Finally, using (9) and (10) we find that also

$$\sum_{i=1}^{m} \lambda_i X_n(t_i) = 0, \tag{11}$$

and so, the statement of the theorem holds for dimension n. □

Remark 2.11 Theorem 2.10 provides sufficient conditions for $\mathrm{E}X$ to belong to the convex hull of $X(S)$. However, by inspection of the proof, we can see that axiom (A3) is also necessary, assuming (A1)-(A2) and conditions (C1)-(C5).

Now, as a corollary to Theorem 2.10, using Carathéodory's theorem on representation of convex hull in finite dimension, we get the following result.

Theorem 2.12 *Assume that axioms* (A1)-(A2)-(A3) *and conditions* (C1)-(C5) *hold on* $(S, \mathcal{F}, \mathcal{S}, \mathrm{E}, P)$. *Let* $X = (X_1, \ldots, X_n)$, *where* $X_i \in \mathcal{S}$ *for all i. Then there are points* $t_0, \ldots, t_n$ *and a discrete probability law given by probabilities* $\lambda_0, \ldots, \lambda_n$ *so that*

$$\mathrm{E}X_i = \sum_{j=0}^{n} \lambda_j X_i(t_j), \quad i = 1, \ldots n. \tag{12}$$

Theorem 2.12 is the most general mean value theorem for axiomatic integral. Due to Remark 2.7, the statement of this theorem applies with $EX_i = \int_S X_i(s)\,d\mu(s)$, μ is a countably additive probability measure on $(S, \mathcal{F})$ where $\mathcal{F}$ is a sigma algebra, and X_i are $(S, \mathcal{F}) - (\mathbf{R}, \mathcal{B})$ measurable and integrable functions ($\mathcal{B}$ is the Borel sigma-field on $\mathbf{R}$).

In the next section we consider the case of continuous functions X_i.

3 Mean value theorem for continuous multivariate mappings

Definition 3.1 A path from a point a to a point b in a topological space S is a continuous mapping $f : [0,1] \to S$ such that $f(0) = a$ and $f(1) = b$. A space S is path connected if for any two points $a, b \in S$ there exists a path that connects them.

Let us remark that any topological vector space is path connected. A path that connects points a and b is given by $f(\lambda) = \lambda a + (1 - \lambda)b$, $\lambda \in [0,1]$. The same is true for a convex subset S of any topological space.

The following result is a generalization of Theorem B.

Theorem 3.2 *Let $X : t \mapsto X(t)$, $t \in S$, be a continuous function defined on a path-connected topological space S with values in $\mathbf{R}^n$, and let K be the convex hull of the set $X(S)$. Then each $v \in K$ can be represented as a convex combination of n or fewer points of the set $X(S)$.*

Proof By Carathéodory's theorem, any $v \in K$ can be represented as a convex combination of at most $n + 1$ points of the set $X(S)$. Without loss of generality, assume that $v = 0$. Therefore, there exist $t_j \in S$ and $v_j \geq 0$, $0 \leq j \leq n$, such that $t_i \neq t_j$ for $i \neq j$, $v_0 + \cdots + v_n = 1$, and

$$v_0 x(t_0) + v_1 x(t_1) + \cdots + v_n x(t_n) = 0. \tag{13}$$

We may also assume that all $n + 1$ points $x(t_j)$ do not belong to one hyperplane in $\mathbf{R}^n$ (in particular, $x(t_i) \neq x(t_j)$ for $i \neq j$) and that the numbers v_j are all positive; otherwise, at least one term from (13) can be eliminated. Now we apply the following reasoning.

Denote by $p_j(x)$, $0 \leq j \leq n$, the coordinates of the vector $x \in \mathbf{R}^n$ with respect to the coordinate system with the origin at 0, and with the vector base consisting of vectors $x(t_j)$, $j = 1, \ldots, n$ (that is, $x = \sum_{j=1}^{n} p_j(x) x(t_j)$). Then from (13) we find that $p_j(x(t_0)) = -v_j/v_0 < 0$, $j = 1, \ldots, n$, *i.e.*, the coordinates of the vector $x(t_0)$ are negative. The coordinates of vectors $x(t_j)$, $j = 1, 2, \ldots, n$, are non-negative: $p_j(x(t_j)) = 1$ and $p_k(x(t_j)) = 0$ for $k \neq j$. Now consider a path $t = t(\lambda)$, $\lambda \in [0,1]$ which connects points t_0 and t_1, so that $t(0) = t_0$ and $t(1) = t_1$. The functions $\lambda \mapsto p_j(x(t(\lambda))) := f_j(\lambda)$ are continuous as mappings from $[0,1]$ to $\mathbf{R}^n$ and $f_j(0) < 0$ for all $j = 1, \ldots, n$, whereas $f_1(1) = 1$ and $f_j(1) = 0$ for $j > 1$. Therefore, for each of functions f_j there exists one or more points $\lambda \in (0,1]$ such that $f_j(\lambda) = 0$. Since the set $N = \bigcup_{j=1}^{n} f_j^{-1}(\{0\})$ is a closed and non-empty subset of $(0,1]$, there exists $\lambda_0 = \min N$, $\lambda_0 > 0$. Let $\bar{t} := t(\lambda_0)$. From this construction it follows that there exists (at least one) k such that $p_k(x(\bar{t})) = 0$ and $p_j(x(\bar{t})) < 0$ for $j \neq k$. Hence,

$$x(\bar{t}) - \sum_{1 \leq j \leq n, j \neq k} p_j(x(\bar{t})) x(t_j) = 0,$$

and it follows that $v = 0$ is a convex combination of points $x(\bar{t})$ and $x(t_j)$, $j = 1, \ldots, n$, $j \neq k$.

□

As a direct corollary to Theorems 2.10 and 3.2, we have the following mean value theorem for continuous multivariate mappings.

Theorem 3.3 *Let $(S, \mathcal{F}, \mathcal{S}, \mathrm{E}, P)$ be a quintuplet as defined in Section 2, where S is a path-connected topological space. Under conditions* (C1)-(C5) *and axioms* (A1)-(A3), *let X_i, $i = 1, \ldots, n$, be continuous functions from $\mathcal{S}$. Then there exist points $t_1, \ldots, t_n$ in S and non-negative numbers $\lambda_1, \ldots, \lambda_n$, with $\sum_{i=1}^n \lambda_i = 1$, such that*

$$\mathrm{E}X_k = \sum_{i=1}^n \lambda_i X_k(t_i), \quad k = 1, \ldots, n.$$

Remark 3.4 Since Theorem 3.2 is independent of axioms and conditions of Section 2, the statement of Theorem 3.3 holds whenever Theorem 2.12 holds. In particular, it holds whenever $\mathrm{E}X_i = \int_S X_i(s)\, d\mu(s)$, where μ is a countably additive probability measure.

In fact, all what Theorem 3.3 says is that we can save one point in the representation (12) of Theorem 2.12. Although it might look not much significant, in some applications it makes difference. For example, Karamata's representation for covariance in [11] based on Kowalewski's original result for $n = 2$ strongly depends on two points and nothing similar can be derived with three points.

4 Some particular cases and open problems
4.1 Riemann and Lebesgue integral on $\mathbf{R}^d$

For $d = 1$, let $-\infty < a < b < +\infty$ and let X be a Riemann integrable function on $[a, b]$. Define

$$\mathrm{E}X = \frac{1}{b-a} \int_a^b X(s)\, ds. \tag{14}$$

In terms of previous notations, $S = [a, b]$, and $\mathcal{S}$ is the family of Riemann integrable functions on $\mathcal{S}$. A natural choice for algebra $\mathcal{F}$ of sets related to Riemann integral should be the algebra of intervals in $[a, b]$, which can be defined as the collection of all subintervals of $[a, b]$ (including singletons and an empty set) and their finite unions. The corresponding probability is defined then as follows: If $A = \bigcup_{i=1}^k J_i$, where J_i are intervals,

$$P(A) = \frac{1}{b-a} \int_a^b I_A(s)\, ds = \sum_{i=1}^k \lambda(J_i),$$

where $\lambda(J_i)$ is the length of J_i. This is Jordan probability measure on $[a, b]$, and it is well known that, even for continuous bounded functions, the set $\{s \in S : X(s) \le c\}$ where c is a real number does not obligatory belong to $\mathcal{F}$ (this is probably first shown by an example in [12]). This fact makes it impossible to use our system of axioms directly because condition (C3) does not hold. Nevertheless, we can proceed by noticing that the algebra $\mathcal{F}$ as described above is a sub-algebra of the Borel sigma algebra $\mathcal{B}$ generated by open sets in $[a, b]$, and since Riemann and Lebesgue integrals coincide if the integrand is Riemann integrable, we can proceed in this way. In more common notations, let us consider a general case of a functional E based on Lebesgue integral on $\mathbf{R}^d$, $d \ge 1$:

$$\mathrm{E}f = \frac{1}{V(D)} \int_D f(s_1, \ldots, s_d)\, ds_1 \cdots ds_d, \tag{15}$$

or, in shorthand,

$$E f = \frac{1}{V(D)} \int_D f(s) \, d\lambda(s),$$

where λ is the Lebesgue measure on $\mathbf{R}^d$ restricted to D, and $V(D) = \lambda(D) > 0$, where D is a convex (or, in more generality, path-connected) subset of $\mathbf{R}^d$. The underlying probability measure in our construction of Section 2 is $P(\cdot) = \frac{1}{V(D)} \lambda(\cdot)$. Let $f_1, \ldots, f_n$ be continuous functions $D \mapsto \mathbf{R}$ such that $E f_i$ is as defined in (15). Then, by Theorem 3.3, we have that there are points $x_1, \ldots, x_n \in D$ and non-negative numbers $\lambda_1, \ldots, \lambda_n$ with $\sum_{i=1}^n \lambda_i = 1$ such that

$$\frac{1}{V(D)} \int_D f_1(s_1, \ldots, s_d) \, ds_1 \cdots ds_d = \lambda_1 f_1(x_1) + \cdots + \lambda_n f_1(x_n)$$

$$\vdots$$

$$\frac{1}{V(D)} \int_D f_n(s_1, \ldots, s_d) \, ds_1 \cdots ds_d = \lambda_1 f_n(x_1) + \cdots + \lambda_n f_n(x_n).$$

In words, this result shows that for an arbitrary system of n integrals with continuous integrands, there exists an exact quadrature rule with n points in D with coefficients λ_i which are the same for all integrals. Note that x_i are d-dimensional points, so in fact here we have dn scalar parameters.

4.2 Integrals with respect to countably additive probability measure

As already noted, Theorem 3.3 holds for all integrals based on countably additive probability measure. Suppose that S is a path-connected topological space, and let μ be a countably additive probability measure on $(S, \mathcal{F})$, where $\mathcal{F}$ is a sigma algebra of Borel subsets of S. Let $f_1, \ldots, f_n$ be continuous mappings from S to $\mathbf{R}$, and suppose that

$$E f_i = \int_S f(s) \, d\mu(s), \quad i = 1, \ldots, n \tag{16}$$

is finite. In particular, let $S = C[0, T]$, the space of continuous functions on the interval $[0, T]$ with the supremum norm. Then the measure μ defines a stochastic process on $[0, T]$, $s(t)$, $0 \le t \le T$, and $f_i(s)$ is a continuous functional of trajectories of the process. By Theorem 3.3, we have that the system of expectations (16) can be represented as

$$E f_i = \sum_{j=1}^n \lambda_j f_i(x_j), \quad i = 1, \ldots, n$$

for some $x_j \in C[0, T]$ and $\lambda_j \ge 0$ with $\sum_j \lambda_j = 1$.

4.3 Open question

We showed in Section 2 that axiom (A3) does not hold for (integrals based on) SFAPs, so by Remark 2.11, the mean value theorem does not hold in this case. It would be of interest to describe classes of finitely additive probabilities for which axiom (A3) holds or does not hold in terms of some structural properties of measures.

Competing interests

Acknowledgements

References

1. Kowalewski, G: Ein Mittelwertsatz für ein System von n Integralen. Zeitschrf. für Math. und Phys. (Schlömilch Z.) **42**, 153-157 (1895)
2. Kowalewski, G: Bemerkungen zu dem Mittelwertsatze für ein System von n Integralen. Zeitschrf. für Math. und Phys. (Schlömilch Z.) **43**, 118-120 (1896)
3. Janković, S, Merkle, M: A mean value theorem for systems of integrals. J. Math. Anal. Appl. **342**, 334-339 (2008)
4. Heydari, M, Avazzdeh, Z, Navabpour, H, Loghmani, GB: Numerical solution of Fredholm integral equations of the second kind by using integral mean value theorem II. High dimensional problems. Appl. Math. Model. **37**, 432-442 (2013)
5. Gillman, L: An axiomatic approach to the integral. Am. Math. Mon. **100**(1), 16-25 (1993)
6. Ostrowski, AM: On an integral inequality. Aequ. Math. **4**, 358-373 (1970)
7. Schervish, MJ, Seidenfeld, T, Kadane, JB: On the equivalence of conglomerability and disintegrability for unbounded random variables. Technical Report 864, Carnegie Mellon University, Pittsburgh, PA (2008)
8. Schervish, MJ, Seidenfeld, T, Kadane, JB: Dominating countably many forecasts. Ann. Stat. **42**, 728-756 (2014)
9. Yosida, K, Hewitt, E: Finitely additive measures. Trans. Am. Math. Soc. **72**, 46-66 (1952)
10. Rockafellar, RT: Convex Analysis. Princeton University Press, Princeton (1997)
11. Karamata, J: Sur certain inégalités relatives aux quotients et à la différence de $\int fg$ et $\int f \int g$. Publ. Inst. Math. (Belgr.) **2**, 131-145 (1948)
12. Frink, O Jr: Jordan measure and Riemann integration. Ann. Math. **34**, 518-526 (1933)