The plan

What is integration

Our goal is to define integation for functions $f: E \subseteq \mathbb{R}^N \mapsto \mathbb{R}$, that is to gives precise meaning of the intuitive ideas of volume/area. Like many concepts in mathematics, this can be done easily and intuitively for functions that are simple, but quickly become subtle and technical once we start to consider more complicated functions. Therefore we try to make the meaning of integration precise from the very start.

An integral is a function defined on (as many as possible) pairs (f, E) where $E \subseteq \mathbb{R}^N$ and $f: \mathbb{R}^N \mapsto \mathbb{R}$:

$$(f, E) \mapsto I(f, E) \in \mathbb{R} \tag{1}$$

satisfying certain rules of operation. The following should be satisfied by any definition of integral.

1. I is linear in f:

$$I(a f + b g) = a I(f) + b I(g); \tag{2}$$

2. I is linear in E: If $E_1 \cap E_2 = \varnothing$, then

$$I(f, E_1 \cup E_2) = I(f, E_1) + I(f, E_2). \tag{3}$$

3. I is monotone in E and f: If $E_1 \subseteq E_2$, and $0 \leqslant f_1 \leqslant f_2$, then

$$I(f_1, E_1) \leqslant I(f_2, E_2). \tag{4}$$

4. I is translation invariant for $f = 1$. For any $\boldsymbol{x}_0 \in \mathbb{R}^N$, define

$$E + \{\boldsymbol{x}_0\} = \{\boldsymbol{x} + \boldsymbol{x}_0 | \, \boldsymbol{x} \in E\}. \tag{5}$$

   Then

$$I(1, E + \{\boldsymbol{x}_0\}) = I(1, E). \tag{6}$$

5. I is invariant under rotation of E for $f = 1$. For any $O \in \mathbb{R}^{N \times N}$ orthogonal, define

$$O E := \{O \boldsymbol{x} | \, \boldsymbol{x} \in E\}. \tag{7}$$

   Then

$$I(1, O E) = I(1, E). \tag{8}$$

6. I is homogeneous in E: Let $a \in \mathbb{R}$. Then

$$I(1, a E) = |a|^N I(1, E). \tag{9}$$

7. I is normalized:

$$I(1, [0, 1] \times \cdots \times [0, 1]) = 1. \tag{10}$$

Problem 1. Are any of the above redundant?

Remark 1. Note that simply listing requirements 1 – 7 does not mean there is indeed such a theory. To show the existence to such a theory we need to give definition of $I(f, E)$. We will see that this may not be possible for all pairs of functions and sets. Furthermore, different ways of defining integration will lead to different "admissible" pairs of functions and sets. Examples are Danielle integration (continuous functions), Riemann integration (functions not too discontinuous), Lebesgue integration (functions that can be defined without Axiom of Choice), and many more.

Remark 2. From this point on there are two ways to proceed. 1. Define integration and then define the measure of a set E to be $I(1, E)$; 2. First develop a measure theory $E \mapsto \mu(E)$, and then identify $\mu(E)$ with the integral $I(1, E)$ and build a integration theory based on properties 1 – 7. We will take the second approach.

Measure theory

A "measure": $E \mapsto \mu(E)$ is a mathematical abstraction of the everyday concepts of length/area/volume. We expect the following to be true for any reasonable measure theory. In the following E, E_1, E_2 are assumed to be measurable.

i. Lineariy: If $E_1 \cap E_2 = \varnothing$, then $E_1 \cup E_2$ is measurable, and

$$\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2). \tag{11}$$

ii. Monotonicity: If $E_1 \subseteq E_2$, then

$$\mu(E_1) \leqslant \mu(E_2). \tag{12}$$

iii. Translation and rotation invariance: $E + \{\boldsymbol{x}_0\}$ and $O\,E$ are measurable, with

$$\mu(E + \{\boldsymbol{x}_0\}) = \mu(E). \tag{13}$$

$$\mu(O\,E) = \mu(E). \tag{14}$$

iv. Homogeneity. Let $a \in \mathbb{R}$. Then $a\,E$ is measurable, and

$$\mu(a\,E) = |a|^N \mu(E). \tag{15}$$

v. Normalized: The unit interval $[0, 1] \times \cdots \times [0, 1]$ is measurable, and

$$\mu([0, 1] \times \cdots \times [0, 1]) = 1. \tag{16}$$

Exercise 1. Check that the above is consistent with the requirement 1 - 7 for integrations if we identify $I(1, E) := \mu(E)$.

Remark 3. Similar to the situation of integration, it may not be possible to assigne a measure to every set and stay satisfied with i – v.

One illustration of this point is the Banach-Tarski paradox, where the unit ball B can be writte nas $B = S_1 \cup \cdots \cup S_n$ where $S_i \cap S_j = \varnothing$, and then S_n can be translated and rotated to form two disjoint balls each differs from B only by a translation.

Thus the sets $S_1, \ldots, S_n$ cannot all be measurable for any measure satisfying i – iii.

Exercise 2. Explain why $S_1, \ldots, S_n$ cannot all be measurable for any measure satisfying i – iii.

Remark 4. Again, simply listing i – v does not tell us that there exists such a measure. On the other hand, from i – v we can derive many properties that a measure must satisfy, and consequently have a clear idea how a measure should be defined.

Lemma 5. *$\mu(\varnothing) = 0$.*

Proof. For any E measurable, by linearity we have

$$\mu(E) = \mu(E \cup \varnothing) = \mu(E) + \mu(\varnothing) \Longrightarrow \mu(\varnothing) = 0. \tag{17}$$

The proof ends. $\square$

Exercise 3. Let $E_1, \ldots, E_n \subseteq \mathbb{R}^N$ be measurable and $E_i \cap E_j = \varnothing$ for each pair i, j. Then if $E_1 \cup \cdots \cup E_n$ is measurable, there must hold $\mu(E_1 \cup \cdots \cup E_n) = \mu(E_1) + \cdots + \mu(E_n)$.

Lemma 6. *For any measurable $E \subseteq \mathbb{R}^N$, $\mu(E) \geqslant 0$.*

Proof. By monotonicity $\varnothing \subseteq E$ therefore $\mu(E) \geqslant \mu(\varnothing) = 0$. $\square$

Lemma 7. *If $(0,1) \times \cdots \times (0,1)$ is measurable, then $\mu((0,1) \times \cdots \times (0,1)) = 1$.*

Proof. Define $C_\varepsilon = \begin{pmatrix} \varepsilon \\ \vdots \\ \varepsilon \end{pmatrix} + (1 - 2\varepsilon)([0,1] \times \cdots [0,1])$. Then we have $I(1, C_\varepsilon) = (1 - 2\varepsilon)^N$. Since

$$C_\varepsilon \subseteq (0,1) \times \cdots \times (0,1) \subseteq [0,1] \times \cdots \times [0,1] \tag{18}$$

we have

$$(1 - 2\varepsilon)^N \leqslant I(1, (0,1) \times \cdots \times (0,1)) \leqslant 1. \tag{19}$$

Now as $\varepsilon > 0$ is arbitrary, we must have $I(1, (0,1) \times \cdots \times (0,1)) = 1$. $\square$

Exercise 4. From i - v, what can you say about the measure of an arbitrary compact interval in $\mathbb{R}^N$?

Integration theory

Some properties of integrals will also follow directly from the natural requirements. For example,

Lemma 8. *Let $f \geqslant 0$ be integrable. Then $I(f, E_1 \cup E_2) \leqslant I(f, E_1) + I(f, E_2)$.*

Proof. We have

$$I(f, E_1 \cup E_2) = I(f, E_1) + I(f, E_2 - E_1) \leqslant I(f, E_1) + I(f, E_2). \tag{20}$$

The proof ends. $\square$

On the other hand, some functions are naturally integrable:

Now different approaches can be taken to extend the definition to more complicated functions.

1. First define a measure μ. The domain of μ must include all the smooth domains, for example those whose boundaries are smooth curves.
2. Once this is done, the integration of piecewise constant functions can be easily defined.
3. From here it is possible to define integrals for continuous functions and maybe some other functions by approximating f from above and below using piecewise constant functions.

Remark 9. Step 3 in the definition of Lebesgue integral is slightly different since Lebesgue's theory allows f to be unbounded.

Exercise 5. If you have Lebesgue measure at hand, how would you define Lebesgue integral? Compare with what Lebesgue actually did.