Part II – Probability and Measure

Based on lectures by E. Brieuillard
Notes taken by Bhavik Mehta

Michaelmas 2017

0 Introduction

0.1 Course structure

- Week 1: Lebesgue measure
- Week 2: Abstract measure theory
- Week 3: Integration
- Week 4: Foundations of probability theory
- Week 5: L^p spaces
- Week 6: Modes of convergence
- Week 7: Fourier transform and gaussians
- Week 8: Ergodic Theory

0.2 Historical motivation

Suppose we have a subset $E \subset \mathbb{R}^d$.

1. What does it mean to measure this subset?

   In one dimension, we have some intuition of length, and in two and three dimensions we are familiar with the notions of surface area and volume.

2. Does it make sense to measure every subset?

   This seems reasonable, but it turns out that assigning a measure to every subset can lead to logical contradictions.

3. What is a measure? It should be a function defined on subsets, in particular some assignment $E \to m(E) \in \mathbb{R}$. It should satisfy some properties:

   - Non-negativity: $m(E) \geq 0$ for all E
   - Empty set: $m(\emptyset) = 0$
   - Additivity: $m(E \sqcup F) = m(E) + m(F)$ for any two disjoint sets E and F

- Normalisation: $m([0,1]^d) = 1$
- Translation invariant: $m(E + x) = m(E)$ for all E and all $x \in \mathbb{R}^d$

It's possible to construct pathological 'measures' satisfying all these axioms and defined *on all* subsets of $\mathbb{R}^d$, but they won't be 'nice'. When mathematicians construct such measures, they usually do so on a restricted class of sets, otherwise this leads to contradictions.

If $d \geq 3$, it can be shown that there is no $m : P(\mathbb{R}^d) \to [0, \infty)$ that is also rotation invariant. This is referred to as the Hausdorff–Banach–Tarski paradox. Namely, if we take $B = B(0,1) = \left\{ \mathbf{x} \in \mathbb{R}^d : x_1^2 + \dots x_d^2 \leq 1 \right\}$, then there is a partition

$$B = X_1 \sqcup \cdots \sqcup X_k \sqcup Y_1 \sqcup \cdots \sqcup Y_k$$

and isometries $g_1, \dots, g_k$, $h_1, \dots, h_k$ such that

$$\bigcup g_i X_i = B = \bigcup h_i Y_i$$

1 Lebesgue measure

Definition (Jordan measure of a box)**.** We define the **Jordan measure of the box** by $m(B) \coloneqq \prod_{i=1}^{d} |b_i - a_i|$.

Consider a box $B \subset \mathbb{R}^d$, given by $B = \prod_{i=1}^{d} I_i$, where $I_i = [a_i, b_i]$ are intervals in $\mathbb{R}$ From here, we can define the Jordan measure of the box by $m(B) \coloneqq \prod_{i=1}^{d} |b_i - a_i|$.

Definition. An **elementary subset** of $\mathbb{R}^d$ is a finite union of boxes.

Remark. Every elementary set can be written as a finite union of disjoint boxes. The family of elementary sets is stable under finite unions, finite intersections and set difference. The concern may arise that if the disjoint union can be taken in two different ways, then perhaps we could get different measures, but

$$\begin{aligned} E = \bigsqcup_{i=1}^{N} B_i &= \bigsqcup_{j=1}^{M} B'_j \\ \implies \sum_{i=1}^{N} m(B_i) &= \sum_{j=1}^{M} m(B'_j) \end{aligned}$$

This means that *makes sense* to define $m(E)$ as $\sum_1^N m(B_i)$, for elementary subsets.

Definition. A subset $E \subset \mathbb{R}^d$ is **Jordan measurable** if $\forall \epsilon > 0\ \exists$ elementary sets A, B such that $A \subset E \subset B$ and $m(B \setminus A) < \epsilon$.

Remark. If E is Jordan measurable, then

$$\inf \{\, m(B) \mid E \subset B, B \text{ elementary} \,\} = \sup \{\, m(A) \mid A \subset E, A \text{ elementary} \,\}$$

Proof is left as an exercise for the reader.

Definition. We define the **Jordan measure** of E as this supremum or infimum, and denote it by $m(E)$.

Exercise. m so defined satisfies all the axioms defined earlier.

Definition (Riemann integrable function)**.** A function $f : [a, b] \to \mathbb{R}$ is **Riemann integrable** if all its Riemann sums converge. Formally, the integral $I(f) \in \mathbb{R}$ exists if $\forall \epsilon > 0$, we can find $\delta > 0$ such that for every partition P of $[a, b]$ of width $\tau(P) < \delta$, we have $|S(f, P) - I(f)| < \epsilon$, where we recall the following

A partition P given by $a = t_0 < t_1 < \cdots < t_n = b$ has width

$$\tau(P) = \max_{0 \leq i < N} |t_{i+1} - t_i|$$

and

$$S(f, P) = \sum_{i=1}^{N-1} f(x_i)(t_{i+1} - t_i)$$

where $x_i \in [t_i, t_{i+1}]$

Proposition. f is Riemann integrable if and only if

$$E^+ = \{ (x,t) \in \mathbb{R}^2 \mid 0 \le t \le f(x) \}$$
$$E^- = \{ (x,t) \in \mathbb{R}^2 \mid f(x) \le t \le 0 \}$$

are both Jordan measurable.

However the Jordan measure is not perfect. For instance, the complement of a Jordan measurable set is not Jordan measurable. Also, we can take an (infinite) union of Jordan measurable sets and produce a set which is not Jordan measurable. In addition, there are simple sets that are not Jordan measurable, and simple functions that are not Riemann integrable.

Example.

(*i*)

$$\mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$$

This is not Riemann integrable, as seen in earlier Analysis courses.

(*ii*) $\mathbb{Q}$ or even $\mathbb{Q} \cap [0,1]$ is not a Jordan measurable subset of $\mathbb{R}$. In fact no dense countable subset of an interval in $\mathbb{R}$ can be Jordan measurable.

Another problem is with limits of functions: a pointwise limit of Riemann integrable functions is not always Riemann integrable.

Example. $f_n(x) = \mathbb{1}_{\frac{1}{n!}\mathbb{Z} \cap [0,1]}$ is Riemann integrable, but $f(x) = \lim_{n \to 0} f_n(x) = \mathbb{1}_{\mathbb{Q} \cap [0,1]}$ is not.

In comparison to the Jordan measure, the ideas of Lebesgue were to remove the containment $A \subset E$, and to allow a countable union of boxes instead of just a finite union.

We begin by defining m^*, the *Lebesgue outer measure*:

Definition (Lebesgue outer measure)**.** For a subset $F \subseteq \mathbb{R}^d$, the **Lebesgue outer measure** m^* is given by

$$m^*(F) = \inf \left\{ \sum_{n \ge 1} m(B_n) \;\middle|\; F \subset \bigcup_{n \ge 1} B_n,\ B_n \text{ are boxes} \right\}$$

This is defined for all subsets of $\mathbb{R}^d$, but it is not additive on all subsets.

Remark.

- If you let $m^{*,J}(F)$ (the Jordan outer measure) be defined similarly, but for finitely many boxes B_n instead of countably many, then $m^*(F) \le m^{*,J}(F)$.
- But the inequality can be strict, for instance for $F = \mathbb{Q} \cap [0,1]$ we have $m^*(F) = 0$ but $m^{*,J}(F) = 1$.

The Lebesgue outer measure satisfies

1. $m^*(\emptyset) = 0$

2. $m^*(E) \leq m^*(F)$ if $E \subseteq F$ (monotone)

3. $m^*\left(\bigcup_{i\geq 1} E_i\right) \leq \sum_{i\geq 1} m^*(E_i)$ (countable subadditivity)

Example (Vitali's example)**.** Let E be a set of representations of the cosets of the subgroup $(\mathbb{Q}, +) \subset (\mathbb{R}, +)$. We can choose $E \subset [0, 1]$, such that

$$\forall x \in \mathbb{R}\ \exists! e \in E \text{ such that } x - e \in \mathbb{Q}$$

So the family $\{E + r\}_{r\in\mathbb{Q}}$ is a disjoint family of subsets of $\mathbb{R}$, a partition. By translational invariance,

$$m^*(E + r) = m^*(E) \quad \forall r \in \mathbb{Q}$$

If m^* were additive, we can consider distinct $r_1, \ldots, r_N \in \mathbb{Q} \cap [0, 1]$ so

$$m^*\left(\bigcup_{i=1}^{N} E + r_i\right) = Nm^*(E)$$

but

$$\bigcup_{i=1}^{N} E + r_i \subseteq [0, 2]$$

$$\implies m^*\left(\bigcup E + r_i\right) \leq m^*([0, 2]) \leq 2$$

So for any $N \in \mathbb{N}$, we have $Nm^*(E) \leq 2$, hence $m^*(E) = 0$. On the other hand, $[0, 1] \subseteq \bigcup_{r\in\mathbb{Q}} E + r$, so by countable subadditivity $1 = m^*([0, 1]) \leq \sum_{r\in\mathbb{Q}} m^*(E + r) = 0$, so we have a contradiction. This shows that m^* is not additive on all subsets.

Remark.

- This contruction uses the axiom of choice to define E.
- We will soon define the Lebesgue measure of a Lebesgue measurable set as the outer measure of that set, and this will be additive on Lebesgue measurable sets. This means Vitali's set E was *not* Lebesgue measurable.

Definition (Lebesgue measurable subset)**.** $E \subseteq \mathbb{R}^d$ is **Lebesgue measurable** if $\forall \epsilon > 0$, $\exists C \coloneqq \bigcup_n B_n$, a countable union of boxes such that $m^*(C \setminus E) < \epsilon$ and $E \subseteq C$.

Example (Middle-thirds Cantor set)**.** This is a compact subset C of $[0, 1]$, defined as follows. Start with the interval $I_0 = [0, 1]$ and remove the middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$, leaving $I_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Then remove the middle third of each interval here, giving $I_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$. Repeat, then define $C = \bigcap_{n\geq 0} I_n$. Equivalently, take a ternary expansion of $x \in [0, 1]$, and define C as the set of x which has *none* of its digits equal to 2.

Remark. I_n is a finite union of intervals, so it is an elementary set (link). In particular, $m(I_n) = 2^n/3^n \to 0$ as $n \to \infty$, so C is Jordan measurable with measure 0. Every Jordan measurable set is Lebesgue measurable (clear from the definition).

We can define a 'fat' Cantor set which is Jordan measurable but not Lebesgue measurable.

Remark. What is the outer measure of a Vitali set? $m^*(E) > 0$ by the argument given above, but we can create $E \subseteq [0, \frac{1}{n}]$, or inside any interval. So, $m^*(E)$ depends on the choice of E and could be arbitrarily small, but must be positive.

Remark. In the definition of the Lebesgue outer measure, we used closed boxes, but open or half-open boxes would not change the definition (same for Jordan measure).

Remark. Every null set is Lebesgue measurable.

Definition. A **null set** $E \subseteq \mathbb{R}^d$ is a subset E such that $m^*(E) = 0$.

We state two propositions which will be proved together.

Proposition 1. The family $\mathcal{L}$ of Lebesgue measurable subsets of $\mathbb{R}^d$ is stable under

a) countable unions: if $E_n \in \mathcal{L}$ for all $n \in \mathbb{N}$, then $\bigcup_{n\geq 1} E_n \in \mathcal{L}$.

b) complementation: if $E \in \mathcal{L}$, then $E^c \in \mathcal{L}$.

c) countable intersections: if $E_n \in \mathcal{L}$ for all $n \in \mathbb{N}$, then $\bigcap_{n\geq 1} E_n \in \mathcal{L}$.

Proposition 2. Every closed (open) subset of $\mathbb{R}^d$ is Lebesgue measurable.

First note that part c) follows from parts a) and b), since $(\bigcap_n E_n)^c = \bigcup_n E_n^c$. We'll first show a).

Proof of proposition 1a). Let $(E_n)_{n\geq 1}$ be a countable family in $\mathcal{L}$, and pick $\epsilon > 0$. By definition, $\exists C_n \coloneqq \bigcup_{i\geq 1} B_i^{(n)}$ such that $m^*(C_n \setminus E_n) < \frac{\epsilon}{2^n}$. Then $\bigcup E_n \subseteq \bigcup C_n = \bigcup_{n,i} B_i^{(n)}$ is still a countable union of boxes.

Finally,

$$m^*\left(\bigcup B_i^{(n)} \setminus \bigcup E_n\right) \leq \sum_n m^*(C_n \setminus E_n) \leq \sum_{n\geq 1} \frac{\epsilon}{2^n} \leq \epsilon$$

by countable subadditivity of m^*, proving a). □

Next, prove a lemma which will help to prove proposition 2.

Lemma. Every open subset of $\mathbb{R}^d$ is a countable union of open boxes.

Proof. For $r \in \mathbb{Q}_{\geq 0}$ and $s \in \mathbb{Q}^d$, let $B_{s,r} = \left\{ x \in \mathbb{R}^d \;\middle|\; |x_i - s_i| < \frac{1}{r} \text{ for } i = 1, \dots, d \right\}$ $(B_{s,r})_{s,r}$ form a countable family and any open set $U \subseteq \mathbb{R}^d$ is the union of those $B_{s,r}$'s it contains. □

Now we prove open and closed sets are Lebesgue measurable.

Proof of proposition 2. By part a) of proposition 1 and the lemma, we see that every open set in $\mathbb{R}^d$ is Lebesgue measurable. Now, let's show that closed sets are Lebesgue measurable.

It's enough to show that *compact* subsets of $\mathbb{R}^d$ are in $\mathcal{L}$ because every closed subset $F \subseteq \mathbb{R}^d$ is a countable union of compact sets:

$\mathbb{R}^d = \bigcup_{n\geq 1} A_n$, where A_n is an annulus $\{ x \in \mathbb{R}^d \mid n - 1 \leq \|x\| \leq n \}$, so we can write $F = \bigcup_{n\geq 1}(A_n \cap F)$, where $A_n \cap F$ is compact.

So, let $F \subseteq \mathbb{R}^d$ be a compact subset. By definition of $m^*(F)$, $\forall k \geq 1$, $\exists$ a countable union of open boxes $\bigcup_n B_n^{(k)}$ such that $F \subseteq \bigcup_n B_n^{(k)}$ and $m^*(F) + \frac{1}{2^k} \geq \sum_n m(B_n^{(k)})$.

Note:

- Up to subdividing each $B_n^{(k)}$ into a finite number of smaller boxes, we can assume that each $B_n^{(k)}$ has diameter less than $\frac{1}{2^k}$.
- Without loss of generality we can assume that each box meets F.
- Finally, since F is compact there is a finite subcover, so we can assume that there are only finite many boxes at each step, that is, $F \subseteq \bigcup_{n=1}^{N_k} B_n^{(k)}$ for $N_k \in \mathbb{N}$.

Let $U_k = \bigcup_{n=1}^{N_k} B_n^{(k)}$, so $F \subseteq U_k$ for all k and F meets each box $B_n^{(k)}$. Then we must have $F = \bigcap_{k \geq 1} U_k$, because if $x \in \bigcap U_k$, for any k we can find some $x_k \in F$ such that x and x_k lie in the same box $B_n^{(k)}$, and so $\|x - x_k\|_\infty \leq \frac{1}{2^k}$. F is compact, and $x_k \in F$ has a limit point, so $x \in F$.

We need to show that $m^*(U_k \setminus F)$ tends to 0, because this implies that F is Lebesgue measurable.

Claim that if A, B are two disjoint compact subsets of $\mathbb{R}^d$ then $m^*(A \cup B) = m^*(A) + m^*(B)$. This is clear from definition as we can choose disjoint covers by open boxes.

Apply this to $A = F$ and $B = \overline{U_k} \setminus U_{k+1}$ to get

$$\begin{aligned} m^*(\overline{U_k} \setminus U_{k+1}) + m^*(F) &\leq m^*(\overline{U_k} \setminus U_{k+1} \cup F) \\ &\leq m^*(\overline{U_k}) \\ &= m^*(U_k) \\ &\leq m^*(F) + \frac{1}{2^k} \end{aligned}$$

so $m^*(U_k \setminus U_{k+1}) \leq \frac{1}{2^k}$ Since $F = \bigcap_k U_k$, by countable subadditivity of m^*, we get

$$\begin{aligned} m^*(U_k \setminus F) &\leq \sum_{i \geq k} m^*(U_i \setminus U_{i+1}) \\ &\leq \sum_{i \geq k} \frac{1}{2^i} \\ &\leq \frac{1}{2^{k-1}} \to 0 \end{aligned}$$

□

Finally, we show that the complement of a set in $\mathcal{L}$ is in $\mathcal{L}$.

Proof of proposition 1b). We start with $E \in \mathcal{L}$. By definition, $\forall n$, there is a countable family of open boxes C_n with $E \subseteq C_n$ and $m^*(C_n \setminus E) < \frac{1}{n}$.

Taking complements, we see $C_n^c \subseteq E^c$ and $C_n \setminus E = E^c \setminus C_n^c$, and note C_n is open and C_n^c is closed.

By proposition 2, C_n^c is Lebesgue measurable, and by part a) of proposition 1, so is $\bigcup_n C_n^c$. But $\bigcup_n C_n^c \subseteq E^c$ and

$$m^*\left(E^c \setminus \bigcup_n C_n^c\right) \leq m^*(E^c \setminus C_n^c) = m^*(C_n \setminus E) < \frac{1}{n}$$

.

Hence $m^*(E^c \setminus \bigcup_n C_n^c) = 0$, so

$$E^c = \underbrace{\bigcup_n C_n^c}_{\in \mathcal{L}} \cup \underbrace{E^c \setminus \bigcup_n C_n^c}_{\text{is null hence is in } \mathcal{L}}$$

so by proposition 1 a), $E \in \mathcal{L}$. $\square$