

Notes on Measure Theory

Tianyu Du

June 9, 2020

Contents

1	Sigma Algebras	1
2	Measurable Spaces and Measurable Maps	3
3	Lebesgue Measures and Lebesgue Integrals	4

1 Sigma Algebras

Definition 1.1. For a set X , a set $\mathcal{A} \subseteq \mathcal{P}(X)$ is a σ -**algebra** if it satisfies the following properties:

1. $\emptyset, X \in \mathcal{A}$;
2. for all $A \in \mathcal{A}$, $A^c \in \mathcal{A}$ as well;
3. for a sequence in \mathcal{A} , $\{A_i\}_{i \in \mathbb{N}}$, the union $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$ as well.

An element $A \in \mathcal{A}$ is called a \mathcal{A} -**measurable set**.

Remark 1.1. It's easy to show that the largest σ -algebra of set X is the power set $\mathcal{P}(X)$, and the smallest σ -algebra is $\{\emptyset, X\}$.

Theorem 1.1. Let $\{\mathcal{A}_i\}_{i \in I}$ be the collection of all σ -algebra on X . Then, $\bigcap_{i \in I} \mathcal{A}_i$ is also a σ -algebra on X .

Proof. Clearly, $\emptyset, X \in \bigcap_{i \in I} \mathcal{A}_i$ given that every \mathcal{A}_i is a σ -algebra.

For $A \in \bigcap_{i \in I} \mathcal{A}_i$, $A \in \mathcal{A}_i$ for all $i \in I$. Hence $A^c \in \mathcal{A}_i$ for all $i \in I$. Therefore, $A^c \in \bigcap_{i \in I} \mathcal{A}_i$.

Let $\{F_j\}_{j \in \mathbb{N}}$ be a sequence such that $F_j \in \bigcap_{i \in I} \mathcal{A}_i$ for every j . Then $F_j \in \mathcal{A}_i$ for all i, j since \mathcal{A}_i 's are σ -algebras. Hence, $\bigcup_{j \in \mathbb{N}} F_j \in \mathcal{A}_i$ for all $i \in I$, and $\bigcup_{j \in \mathbb{N}} F_j \in \bigcap_{i \in I} \mathcal{A}_i$. ■

Remark 1.2. The union of σ -algebras are not necessarily a σ -algebra. For example, consider

$$X = \{a, b, c\} \tag{1}$$

$$\mathcal{A}_1 = \{\emptyset, \{a\}, \{b, c\}, X\} \tag{2}$$

$$\mathcal{A}_2 = \{\emptyset, \{b\}, \{a, c\}, X\} \tag{3}$$

$$\mathcal{A}_1 \cup \mathcal{A}_2 = \{\emptyset, \{a\}, \{b\}, \{b, c\}, \{a, c\}, X\} \tag{4}$$

Both \mathcal{A}_1 and \mathcal{A}_2 are σ -algebras, but $\mathcal{A}_1 \cup \mathcal{A}_2$ is not a σ -algebra because $\{a\} \cup \{b\} \notin \mathcal{A}_1 \cup \mathcal{A}_2$.

Definition 1.2. For $\mathcal{M} \subseteq \mathcal{P}(X)$ (\mathcal{A} is not necessarily a σ -algebra), the smallest σ -algebra (by taking intersections) containing \mathcal{M} is defined as the **σ -algebra generated by \mathcal{M}** . The generated σ -algebra is simply the intersection of all σ -algebras that are supersets of \mathcal{M} .

$$\sigma(\mathcal{M}) = \bigcap_{\mathcal{A} \supseteq \mathcal{M} \text{ s.t. } \mathcal{A} \text{ is } \sigma\text{-algebra}} \mathcal{A} \quad (5)$$

Definition 1.3. Let (X, τ) be a topological space, then the **Borel algebra** is σ -algebra generated by the collection of open sets τ .

$$\mathcal{B}(X) := \sigma(\tau) \quad (6)$$

Remark 1.3. We do not use the entire power set for analysis because it's too large to construct a sensible measure on (see Theorem 1.2).

Theorem 1.2. There is no measure μ on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ satisfying the following two conditions: (i) $\mu((a, b]) = b - a$ for every $a < b$ and (ii) $\mu(x + A) = \mu(A)$ for all $a \in \mathbb{R}$ and $A \in \mathcal{P}(\mathbb{R})$.

Proof. Suppose, for contradiction, there exists such a measure μ , then $\mu((0, 1]) = 1 < \infty$.

Claim: the only measure on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ satisfying $\mu((0, 1]) < \infty$ and $\mu(x + A) = \mu(A)$ is the zero measure.

To prove the claim, let $I := (0, 1]$ and define the following equivalence relation on I :

$$x \sim y \iff x - y \in \mathbb{Q} \quad (7)$$

the corresponding equivalence class of x on I can be written as

$$[x] = \{x + r : r \in \mathbb{Q} \wedge x + r \in I\} \quad (8)$$

The collection of all such equivalence classes, \mathcal{A} , is a disjoint decomposition of I . (for every $x \in I$, $[x]$ must in \mathcal{A} and $x \in [x]$ trivially. If there exists different $[x] \neq [y]$ but $[x] \cap [y] \neq \emptyset$, take $z \in [x] \cap [y]$, by the transitivity of equivalence relation, $x \sim z \sim y$. Therefore, $[x] = [y]$, contradiction.)

For each $[x] \in \mathcal{A}$, take exactly one $a_x \in [x]$ and define set $A := \{a_x : [x] \in \mathcal{A}\}$. As a result, set A satisfies the following two properties:

1. $\forall x \in I, \exists a_x \in A \text{ s.t. } a_x \in [x]$.
2. $\forall x, y \in A, x \sim y \implies x = y$.

Since $\mathbb{Q} \cup (-1, 1]$ is countable, let $(r_n)_{n \in \mathbb{N}}$ be an enumeration of all elements in it.

For each $n \in \mathbb{N}$, define $A_n := r_n + A$.

Note that for any m, n such that $A_m \cap A_n \neq \emptyset$, take $x \in A_m \cap A_n$. By definition,

$$x = r_n + a_n \quad (9)$$

$$x = r_m + a_m \quad (10)$$

where $a_n, a_m \in A$ and $r_n, r_m \in \mathbb{Q}$. Consequently,

$$a_n - a_m = r_m - r_n \in \mathbb{Q} \quad (11)$$

Therefore, $a_n \sim a_m$. By the second property of A , $a_n = a_m$. Thus, $r_m = r_n$ and $m = n$.

Take the counterposition of what we just proved, $m \neq n \implies A_m \cap A_n = \emptyset$.

Let $z \in (0, 1]$, there exists some $a \in A$ such that $z \in [a]$. That is, $z = x + r$ for some $r \in \mathbb{Q} \cap (-1, 1]$. There must exist some $m \in \mathbb{N}$ such that $r_m = r$, and consequently, $z \in A_m$.

Therefore, $(0, 1] \subseteq \bigcup_{n \in \mathbb{N}} A_n \subseteq (-1, 2]$ (the second relation is obvious). Moreover,

$$\mu((0, 1]) \leq \mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \mu((-1, 2]) = \mu((-1, 0]) + \mu((0, 1]) + \mu((1, 2]) = 3\mu((0, 1]) \quad (12)$$

Note that we just proved $\bigcup_{n \in \mathbb{N}} A_n$ is a disjoint union, hence,

$$\mu((0, 1]) \leq \sum_{n=1}^{\infty} \mu(A_n) \leq 3\mu((0, 1]) \quad (13)$$

$$\implies ((0, 1])\mu \leq \sum_{n=1}^{\infty} \mu(A + r_n) \leq 3\mu((0, 1]) \quad (14)$$

$$\implies \mu((0, 1]) \leq \sum_{n=1}^{\infty} \mu(A) \leq 3\mu((0, 1]) \quad (15)$$

Since $\mu((0, 1])$ is finite, the only value $\mu(A)$ can take is zero, and $\mu(I) = 0$ as well. Consequently, for any set $S \in \mathcal{P}(\mathbb{R})$, if $S \subseteq I$, then $\mu(S) \leq \mu(I)$ and $\mu(S) = 0$. Otherwise, let $l = \lfloor \inf(S) \rfloor$ and $u = \lceil \sup(S) \rceil$.

$$I \subseteq S \subseteq \bigcup_{n=l}^u (n, n+1] \quad (16)$$

Therefore,

$$0 \leq \mu(S) \leq \sum_{n=l}^u \mu(n + (0, 1]) = \sum_{n=l}^u \mu((0, 1]) = 0 \quad (17)$$

It's shown that $\mu(S) = 0$ for every $S \subseteq \mathcal{P}(\mathbb{R})$.

This leads to a contradiction to the first property required ($\mu((a, b]) = b - a$). ■

2 Measurable Spaces and Measurable Maps

Definition 2.1. Let (X_1, \mathcal{A}_1) and (X_2, \mathcal{A}_2) be two measurable spaces. A function $f : X_1 \rightarrow X_2$ is a **measurable map** with respect to \mathcal{A}_1 and \mathcal{A}_2 (sometimes written as $f : (X_1, \mathcal{A}_1) \rightarrow (X_2, \mathcal{A}_2)$) if

$$f^{-1}(A_2) \in \mathcal{A}_1 \quad \forall A_2 \in \mathcal{A}_2 \quad (18)$$

That is, the pre-image of every set in \mathcal{A}_2 is an element in \mathcal{A}_1 as well.

Theorem 2.1. Let (X, \mathcal{A}) be a measurable space, then the indicator (characteristic) function for any $A \in \mathcal{A}$, $\mathcal{X}_A : X \rightarrow \mathbb{R}$, is measurable with respect to \mathcal{A} and $\mathcal{B}(\mathbb{R})$.

$$\mathcal{X}_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad (19)$$

Proof. Since \mathcal{X}_A can only take values from $\{0, 1\}$, the pre-image of any set $\not\subseteq \{0, 1\}$ is undefined. We only need to consider pre-images of subsets of $\{0, 1\}$:

$$\mathcal{X}_A^{-1}(\emptyset) = \emptyset \quad (20)$$

$$\mathcal{X}_A^{-1}(\{0, 1\}) = X \quad (21)$$

$$\mathcal{X}_A^{-1}(\{0\}) = A^c \quad (22)$$

$$\mathcal{X}_A^{-1}(\{1\}) = A \quad (23)$$

Therefore, \mathcal{X}_A is measurable. ■

Theorem 2.2. The composition of measurable maps is measurable.

Proof. For measurable spaces (X_1, \mathcal{A}_1) , (X_2, \mathcal{A}_2) , and (X_3, \mathcal{A}_3) , let $f : (X_1, \mathcal{A}_1) \rightarrow (X_2, \mathcal{A}_2)$ and $g : (X_2, \mathcal{A}_2) \rightarrow (X_3, \mathcal{A}_3)$ be two measurable functions.

Let $A_3 \in \mathcal{A}_3$, $A_2 := g^{-1}(A_3) \in \mathcal{A}_2$. Similarly, $A_1 := f^{-1}(A_2) \in \mathcal{A}_1$ as well. Note that $A_1 = (g \circ f)^{-1}(A_3)$, therefore, $g \circ f$ is measurable. ■

Theorem 2.3. For measurable spaces (X, \mathcal{A}) and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and measurable maps $f, g : \Omega \rightarrow \mathbb{R}$, $f + g$, $f - g$ and $|f|$ are measurable.

Proof. ■

3 Lebesgue Measures and Lebesgue Integrals

Definition 3.1. Let (X, \mathcal{A}, μ) be a measure space and for any $A \in \mathcal{A}$, the **Lebesgue integral** of indicator function \mathcal{X}_A on X is defined to be $\mu(A) \in [0, \infty]$.

$$\int_X \mathcal{X}_A d\mu := \mu(A) \quad (24)$$

Definition 3.2. A function $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a **simple function** (also termed step function and stair-case function) if there exists finitely many $A_1, \dots, A_n \in \mathcal{A}$ and $c_1, \dots, c_n \in \mathbb{R}$ such that

$$f = \sum_{i=1}^n c_i \mathcal{X}_{A_i} \quad (25)$$

That is, a function f is simple if it can be expressed as a linear combination of *finitely* many indicators.

Let \mathbb{S}^+ denote the set of non-negative simple functions.

$$\mathbb{S}^+ := \{f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) \mid f \text{ is simple and } f \geq 0\} \quad (26)$$

Since simple functions only take finitely many values, every $f \in \mathbb{S}^+$ can be written as

$$f = \sum_{t \in f(X)} t \mathcal{X}_{\{x \in X \mid f(x)=t\}} = \sum_{i=1}^n c_i \mathcal{X}_{A_i}, \quad c_i \geq 0 \quad (27)$$

Theorem 3.1. Simple functions are measurable.

Definition 3.3 (Lebesgue integral for \mathbb{S}^+). For $f \in \mathbb{S}^+$ such that $f = \sum_{i=1}^n c_i \mathcal{X}_{A_i}$ with $c_i \geq 0$, the **Lebesgue integral** of f with respect to μ is

$$I(f) = \int_X f \, d\mu := \sum_{i=1}^n c_i \mu(A_i) \in [0, \infty] \quad (28)$$

Theorem 3.2. The Lebesgue integral of $f, g \in \mathbb{S}^+$ satisfies

1. $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$ for $\alpha, \beta \geq 0$,
2. $f \leq g \implies I(f) \leq I(g)$.

Proof. ■

Definition 3.4 (Lebesgue integral for non-negative functions). For $f \geq 0$, the **Lebesgue integral** of f with respect to measure μ is

$$I(f) = \int_X f \, d\mu := \sup \left\{ \int_X s \, d\mu : s \in \mathbb{S}^+ \text{ and } s \leq f \right\} \quad (29)$$

Definition 3.5. A function f is μ -**integrable** if $\int_X f \, d\mu < \infty$.

Theorem 3.3. Let $f, g : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable functions, if $f = g$ except a μ -measure-zero set, that is,

$$\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0 \quad (30)$$

then $\int_X f \, d\mu = \int_X g \, d\mu$.

Lemma 3.1. Let $h : X \rightarrow [0, \infty)$ be a simple function, for any $\tilde{X} \subseteq X$ such that $\mu(\tilde{X}^c) = 0$, $\int_X h \, d\mu$ is independent from the value of h on \tilde{X}^c .

Proof. of Lemma 3.1. Since h is a simple function, it takes only finitely many values and can be written as

$$h = \sum_{t \in h(X)} t \mathcal{X}_{\{x \in X \mid h(x)=t\}} = \sum_{t \in h(X) \setminus \{0\}} t \mathcal{X}_{\{x \in X \mid h(x)=t\}} \quad (31)$$

define

$$\tilde{h} = \begin{cases} h(x) & \text{if } x \in \tilde{X} \\ a & \text{if } x \in \tilde{X}^c \end{cases} \quad (32)$$

where $a \in [0, \infty)$ takes an arbitrary value, and $\tilde{h} \in \mathbb{S}^+$ as well.

$$\int_X \tilde{h} \, d\mu = \sum_{t \in \tilde{h}(X)} t \mu(\{x \in X \mid \tilde{h}(x) = t\}) \quad (33)$$

$$= \underbrace{a \mu(\tilde{X}^c)}_{=0} + \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} \mid h(x) = t\}) \quad (34)$$

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} \mid h(x) = t\}) \quad (35)$$

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} \mid h(x) = t\}) + \sum_{t \in h(\tilde{X}) \setminus \{0\}} \underbrace{t \mu(\{x \in \tilde{X}^c \mid h(x) = t\})}_{=0} \quad (36)$$

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} \mid h(x) = t\} \cup \{x \in \tilde{X}^c \mid h(x) = t\}) \quad (37)$$

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in X \mid h(x) = t\}) + \sum_{t' \in h(X) \setminus (h(\tilde{X}) \cup \{0\})} t' \mu(\{x \in X \mid h(x) = t'\}) \quad (38)$$

Note that t' 's are values that are attained in \tilde{X}^c only, therefore, $\{x \in X \mid h(x) = t'\} \subseteq \tilde{X}^c$ and have measure zero.

$$(38) = \sum_{t \in h(X) \setminus \{0\}} t \mu(\{x \in X \mid h(x) = t\}) = \int_X h \, d\mu \quad (39)$$

Hence, the value of $\int_X h \, d\mu$ is the same no matter how we change h 's values on \tilde{X}^c . ■

Proof. of Theorem 3.3. Let $\tilde{X} := \{x \in X : f(x) \neq g(x)\}$. Define

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \tilde{X} \\ 0 & \text{if } x \in \tilde{X}^c \end{cases} \quad \tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in \tilde{X} \\ 0 & \text{if } x \in \tilde{X}^c \end{cases} \quad (40)$$

So that $\tilde{f} = \tilde{g}$ for all $x \in X$. By Lemma 3.1,

$$\int_X f \, d\mu = \int_X \tilde{f} \, d\mu = \int_X \tilde{g} \, d\mu = \int_X g \, d\mu \quad (41)$$

■

Theorem 3.4. Let $f, g : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable functions, if $f \leq g$ except a μ -measure-zero set, then $\int_X f \, d\mu \leq \int_X g \, d\mu$.

Proof. Let $\tilde{X} := \{x \in X : f(x) \leq g(x)\}$. Define

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \tilde{X} \\ 0 & \text{if } x \in \tilde{X}^c \end{cases} \quad \tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in \tilde{X} \\ 0 & \text{if } x \in \tilde{X}^c \end{cases} \quad (42)$$

So that $\tilde{f} \leq \tilde{g}$ everywhere. By Lemma 3.1,

$$\int_X f \, d\mu = \int_X \tilde{f} \, d\mu \leq \int_X \tilde{g} \, d\mu = \int_X g \, d\mu \quad (43)$$

■

Theorem 3.5. Let $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable functions, $f = 0$ except a μ -measure-zero set if and only if $\int_X f \, d\mu = 0$.

Proof. Similar to previous proofs. ■

Theorem 3.6 (Monotone Convergence Theorem). For measure space (X, \mathcal{A}, μ) , let $(f_n : X \rightarrow [0, \infty))_{n \in \mathbb{N}}$ be a sequence of measurable functions such that

1. $f_n \leq f_{n+1}$ except for a μ -measure-zero set,
2. $\lim_{n \rightarrow \infty} f_n$ converges point-wisely to f except for a μ -measure-zero set.

Then,

$$\lim_{n \rightarrow \infty} \int_X f_n \, d\mu = \int_X \lim_{n \rightarrow \infty} f_n \, d\mu = \int_X f \, d\mu \quad (44)$$

Proof. ■