Notes on Measure Theory

Tianyu Du

June 13, 2020

Contents

1	Sigma Algebra	1
2	Measurable Spaces and Measurable Maps	4
3	Lebesgue Measures and Lebesgue Integrals	5

1 Sigma Algebra

Definition 1.1. For a set X, a set $A \subseteq \mathcal{P}(X)$ 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 $\bigcup_{i\in\mathbb{N}}\in\mathcal{A}$ 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 $\{A_i\}_{i\in I}$ be the collection of all σ -algebra on X. Then, $\bigcap_{i\in I} A_i$ 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 $F_j \in A_i$ for all i, j since A_i 's are σ -algebra. Hence, $\bigcup_{j \in \mathbb{N}} F_j \in A_i$ 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 -algebra are not necessarily a σ -algebra. For example, consider

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

$$A_1 = \{\emptyset, \{a\}, \{b, c\}, X\}$$
 (2)

$$A_2 = \{\emptyset, \{b\}, \{a, c\}, X\}$$
(3)

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

$$\tag{4}$$

Both A_1 and A_2 are σ -algebra, but $A_1 \cup A_2$ is not a σ -algebra because $\{a\} \cup \{b\} \notin A_1 \cup A_2$.

Definition 1.2. For $\mathcal{M} \subseteq \mathcal{P}(X)$ (\mathcal{M} 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 -algebra that are supersets of \mathcal{M} .

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

The σ -algebra generated by \mathcal{M} is therefore the smallest σ -algebra containing \mathcal{M} .

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 .

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

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).

Definition 1.4. For a measurable space (X, \mathcal{A}) , a map $\mu : \mathcal{A} \to [0, \infty]$ is a **measure** if μ satisfies

- 1. $\mu(\emptyset) = 0$.
- 2. $(\sigma$ -addivitity) $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$, where $A_i \in \mathcal{A}$ for all i and $A_i \cap A_j = \emptyset$ for all $i \neq j$.

Example 1.1. For an element $x \in X$, the **Dirac measure**, δ_x , on a measurable space (X, A) is defined as

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \tag{7}$$

Definition 1.5. For a measurable space (X, \mathcal{A}) and a measure μ defined on it, the triple (X, \mathcal{A}, μ) is a **measure space**.

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 defien the following equivalence relation on I:

$$x \sim y \iff x - y \in \mathbb{Q} \tag{8}$$

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

$$[x] = \{x + r : r \in \mathbb{Q} \land x + r \in I\} \tag{9}$$

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,

$$x = r_n + a_n \tag{10}$$

$$x = r_m + a_m \tag{11}$$

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

$$a_n - a_m = r_m - r_n \in \mathbb{Q} \tag{12}$$

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 $z \in [x]$. 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,

$$\mu((0,1]) \le \mu(\bigcup_{n \in \mathbb{N}} A_n) \le \mu((-1,2]) = \mu((-1,0]) + \mu((0,1]) + \mu((1,2]) = 3\mu((0,1])$$
 (13)

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

$$\mu((0,1]) \le \sum_{n=1}^{\infty} \mu(A_n) \le 3\mu((0,1]) \tag{14}$$

$$\implies ((0,1])\mu \le \sum_{n=1}^{\infty} \mu(A+r_n) \le 3\mu((0,1]) \tag{15}$$

$$\implies \mu((0,1]) \le \sum_{n=1}^{\infty} \mu(A) \le 3\mu((0,1]) \tag{16}$$

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

 $u = \lceil \sup(S) \rceil$.

$$I \subseteq S \subseteq \bigcup_{n=l}^{u} (n, n+1] \tag{17}$$

Therefore,

$$0 \le \mu(S) \le \sum_{n=l}^{u} \mu(n + (0,1]) = \sum_{n=l}^{u} \mu((0,1]) = 0$$
(18)

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, A_1) and (X_2, A_2) be two measurable spaces. A function $f: X_1 \to X_2$ is a **measurable map** with respect to A_1 and A_2 (sometimes written as $f: (X_1, A_1) \to (X_2, A_2)$) if

$$f^{-1}(A_2) \in \mathcal{A}_1 \quad \forall A_2 \in \mathcal{A}_2 \tag{19}$$

That is, the pre-image of every set in A_2 is an element in A_1 as well.

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

$$\mathcal{X}_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \tag{20}$$

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 $\{0,1\}$:

$$\mathcal{X}_A^{-1}(\varnothing) = \varnothing \tag{21}$$

$$\mathcal{X}_A^{-1}(\{0,1\}) = X \tag{22}$$

$$\mathcal{X}_A^{-1}(\{0\}) = A^c \tag{23}$$

$$\mathcal{X}_{A}^{-1}(\{1\}) = A \tag{24}$$

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 $f: (X_1, \mathcal{A}_1) \to (X_2, \mathcal{A}_2)$ and $g: (X_2, \mathcal{A}_2) \to (X_3, \mathcal{A}_3)$ 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 $f, g : \Omega \to \mathbb{R}$, f + g, f - g and |f| are measurable.

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 $\mu(A) \in [0, \infty]$.

$$\int_{X} \mathcal{X}_{A} d\mu := \mu(A) \tag{25}$$

Definition 3.2. A function $f:(X,\mathcal{A})\to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ 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

$$f = \sum_{i=1}^{n} c_i \mathcal{X}_{A_i} \tag{26}$$

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

Let S^+ denote the set of non-negative simple functions.

$$\mathbb{S}^+ := \{ f : (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \mid f \text{ is simple and } f \ge 0 \}$$
 (27)

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

$$f = \sum_{t \in f(X)} t \mathcal{X}_{\{x \in X | f(x) = t\}} = \sum_{i=1}^{n} c_i \mathcal{X}_{A_i}, \quad c_i \ge 0$$
 (28)

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

$$I(f) = \int_{X} f \ d\mu := \sum_{i=1}^{n} c_{i} \mu(A_{i}) \in [0, \infty]$$
 (29)

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

- 1. $I(\alpha f + \beta q) = \alpha I(f) + \beta I(q)$ for $\alpha, \beta > 0$,
- $2. f < q \implies I(f) < I(q).$

Proof.

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

$$I(f) = \int_X f \ d\mu := \sup \left\{ \int_X s \ d\mu : s \in \mathbb{S}^+ \text{ and } s \le f \right\}$$
 (30)

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

Theorem 3.3. Let $f, g: (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable functions, if $0 \leq f = g$ except a μ -measure-zero set, that is,

$$\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0 \tag{31}$$

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

Lemma 3.1. Let $h: X \to [0, \infty)$ 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

$$h = \sum_{t \in h(X)} t \mathcal{X}_{\{x \in X | h(x) = t\}} = \sum_{t \in h(X) \setminus \{0\}} t \mathcal{X}_{\{x \in X | h(x) = t\}}$$
(32)

define

$$\tilde{h} = \begin{cases} h(x) & \text{if } x \in \tilde{X} \\ a & \text{if } x \in \tilde{X}^c \end{cases}$$
(33)

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

$$\int_{X} \tilde{h} \ d\mu = \sum_{t \in \tilde{h}(X)} t\mu(\{x \in X | \tilde{h}(x) = t\})$$
(34)

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

$$\tag{35}$$

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}}^{t \in h(X) \setminus \{0\}} t \mu(\{x \in \tilde{X} | h(x) = t\})$$
(36)

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} | h(x) = t\}) + \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \underbrace{\mu(\{x \in \tilde{X}^c | h(x) = t\})}_{=0}$$
(37)

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in \tilde{X} | h(x) = t\}) \cup \{x \in \tilde{X}^c | h(x) = t\})$$
(38)

$$= \sum_{t \in h(\tilde{X}) \setminus \{0\}} t \mu(\{x \in X | h(x) = t\}) + \sum_{t' \in h(X) \setminus (h(\tilde{X}) \cup \{0\})} t' \mu(\{x \in X | h(x) = t'\})$$
(39)

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

measure zero.

$$(44) = \sum_{t \in h(X) \setminus \{0\}} t\mu(\{x \in X | h(x) = t\}) = \int_X h \ d\mu \tag{40}$$

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 $\tilde{X} := \{x \in X : f(x) \neq g(x)\}$, for each simple function h in above set, define

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

$$\tag{41}$$

Then,

$$\int_{X} f \ d\mu = \sup \left\{ I(h) \mid h \in \mathbb{S}^{+}, h \le f \text{ on } X \right\}$$
 (42)

$$= \sup \left\{ I(\tilde{h}) \mid h \in \mathbb{S}^+, h \le f \text{ on } X \right\}$$
 (43)

$$= \sup \left\{ I(\tilde{h}) \mid h \in \mathbb{S}^+, h \le f \text{ on } \tilde{X} \right\}$$
 (44)

$$= \sup \left\{ I(\tilde{h}) \mid h \in \mathbb{S}^+, h \le g \text{ on } \tilde{X} \right\}$$
 (45)

$$= \int_{X} g \ d\mu \tag{46}$$

Where equation (44) holds because the value of h on \tilde{X}^c does not affect $I(\tilde{h})$.

Theorem 3.4. Let $f,g:(X,\mathcal{A})\to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ be measurable functions, if $0\leq f\leq g$ except a μ -measure-zero set, then $\int_X f\ d\mu\leq \int_X g\ d\mu$.

Proof. By definition of Lebesgue integral,

$$\int_{X} f \ d\mu = \sup \left\{ I(h) \mid h \in \mathbb{S}^{+}, h \le f \text{ on } X \right\}$$
(47)

Let $\tilde{X} := \{x \in X : f(x) \leq g(x)\}$, for each simple function h in above set, define

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

$$\tag{48}$$

Then $h \leq f \iff \tilde{h} \leq f$, and $I(h) = I(\tilde{h})$ by Lemma 3.1.

$$\sup \left\{ I(h) \mid h \in \mathbb{S}^+, h \le f \text{ on } X \right\} = \sup \left\{ I(\tilde{h}) \mid h \in \mathbb{S}^+, h \le f \text{ on } \tilde{X} \right\}$$
 (49)

$$\leq \sup \left\{ I(\tilde{h}) \mid h \in \mathbb{S}^+, h \leq g \text{ on } \tilde{X} \right\}$$
 (50)

$$= \int_X g \ d\mu \tag{51}$$

Therefore,

$$\int_{X} f \ d\mu \le \int_{X} g \ d\mu \tag{52}$$

Theorem 3.5. Let $f:(X,\mathcal{A})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 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 $(f_n : X \to [0, \infty))_{n \in \mathbb{N}}$ be a sequence of measurable functions such that

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

Then,

$$\lim_{n \to \infty} \int_X f_n \ d\mu = \int_X \lim_{n \to \infty} f_n \ d\mu = \int_X f \ d\mu \tag{53}$$

Proof. Since $f_n \leq f_{n+1}$ almost everywhere, and $f_n \to f$ point-wisely almost everywhere, therefore,

$$f_n \le f_{n+1} \le f \quad \forall n \in \mathbb{N} \text{ except a set with zero measure}$$
 (54)

Consequently,

$$\int_{X} f_n \ d\mu \le \int_{X} f_{n+1} \ d\mu \le \int_{X} f \ d\mu \quad \forall n \in \mathbb{N}$$
 (55)

As a result,

$$\lim_{n \to \infty} \int_{Y} f_n \ d\mu \le \int_{Y} f \ d\mu \tag{56}$$

Let h be a simple function such that $0 \le h \le f$, let $\varepsilon > 0$, define

$$X_n := \{ x \in X \mid f_n(x) \ge (1 - \varepsilon)h(x) \}$$

$$\tag{57}$$

$$\tilde{X} := \bigcup_{n=1}^{\infty} X_n \tag{58}$$

Note that $f_{n+1} \geq f_n$ implies $X_n \subseteq X_{n+1}$ and $\lim_{n\to\infty} X_n = \tilde{X}$. Moreover, because the monotonicity and point-wise convergence happen almost everywhere in X, almost all $x \in X$ are in some X_n with n sufficiently large, hence $\mu(\tilde{X}^c) = 0$.

Because $X_n \subseteq X$ and $f_n \ge 0$, for any $n \in \mathbb{N}$,

$$\int_{X} f_n \ d\mu \ge \int_{X_n} f_n \ d\mu \ge \int_{X_n} (1 - \varepsilon)h \ d\mu \tag{59}$$

$$\implies \lim_{n \to \infty} \int_{X} f_n \ d\mu \ge \lim_{n \to \infty} \int_{X_n} (1 - \varepsilon) h \ d\mu \tag{60}$$

$$= \int_{\tilde{X}} (1 - \varepsilon) h \ d\mu \tag{61}$$

$$= \int_{X} (1 - \varepsilon)h \ d\mu \tag{62}$$

Where the last equality holds because $\mu(\tilde{X}^c) = 0$.

Since this inequality holds for all $\varepsilon > 0$,

$$\lim_{n \to \infty} \int_X f_n \ d\mu \ge \lim_{\varepsilon \to 0^+} \int_X (1 - \varepsilon) h \ d\mu = \int_X h \ d\mu \tag{63}$$

moreover, this inequality holds for all $0 \le h \le f$,

$$\lim_{n \to \infty} \int_X f_n \ d\mu \ge \int_X f \ d\mu \tag{64}$$

Results (56) and (64) lead to the conclusion.

Corollary 3.1. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions, $g_n: X \to [0, \infty]$, then the series

$$\sum_{n=1}^{\infty} g_n : X \to [0, \infty] \tag{65}$$

is measurable, and

$$\int_{X} \sum_{n=1}^{\infty} g_n \ d\mu = \sum_{n=1}^{\infty} \int_{X} g_n \ d\mu \tag{66}$$

Proposition 3.1. Infimum of measurable functions is measurable.

Proposition 3.2. Limit of measurable functions is measurable.