Measure Theory & Optimal Transport

Shuailong Zhu

Contents

1 Measure

1.1 Set Family

Proposition 1.1.1 (The relationship among different set families)**.**

$$\pi\text{-system} \to \text{semiring} \to \text{ring} \to \text{field} \to \sigma{-}\text{field} \tag{1.1.1}$$

$$\text{monotone class} \to d\text{-system} \to \sigma\text{-field} \tag{1.1.2}$$

$$\begin{cases} \mathscr{A} \text{ is a monotone class} \\ \mathscr{A} \text{ is a field} \end{cases} \Rightarrow \mathscr{A} \text{ is a } \sigma\text{-field.} \tag{1.1.3}$$

Theorem 1.1.2 (Monotone class theorem)**.** *If $\mathscr{A}$ is a algebra (field) of sets, then $\sigma(\mathscr{A}) = m(\mathscr{A})$ where $m(\mathscr{A})$ denotes the smallest monotone class containing $\mathscr{A}$.*

We begin with the following lemma.

Lemma 1.1.3. *If $\mathscr{A}$ is a ring and $X \in \mathscr{A}$, then $\mathscr{A}$ is a field.*

Proof. It is easy to show with the properties of ring and field.
Property of ring: $\mathscr{R}$ is a π-system; $A, B \in \mathscr{R} \Rightarrow A \cup B, A \backslash B \in \mathscr{R}$.
Property of field: $\mathscr{F}$ is a π-system; $X \in \mathscr{F}; A \in \mathscr{F} \Rightarrow A^c \in \mathscr{F}; A, B \in \mathscr{F} \Rightarrow A \cap B \in \mathscr{F}$.
(The difference between semiring and semialgebra: semiring + $X \in \mathscr{R} \Rightarrow$ semialgebra) □

Proof of Theorem 1.1.2. As $m(\mathscr{A})$ is the smallest σ-field containing $\mathscr{A}$, $m(\mathscr{A})$ is also a monotone class. Then $m(\mathscr{A}) \subseteq \sigma(\mathscr{A})$ since $m(\mathscr{A})$ is the smallest monotone class containing $\mathscr{A}$. Thus, it is enough to prove $\sigma(\mathscr{A}) \subseteq m(\mathscr{A})$. Furthermore, it suffices to show $m(\mathscr{A})$ is a filed by means of relationship 1.1.3. Then, according to Lemma 1.1.3, we only need to verify $m(\mathscr{A})$ is a ring.

$\forall A \in \mathscr{A}$, let

$$\mathscr{G}_A = \{B : B, A \cup B, A \backslash B \in m(\mathscr{A})\}$$

We have

- $\mathscr{G}_A$ is a monotone class. $\forall B_i \uparrow \Rightarrow A \cup B_i \uparrow$, $A \backslash B_i \downarrow \Rightarrow \cup B_i, \cup(A \cup B_i), \cup(A \backslash B_i) \in m(\mathscr{A})$ (according to the definition of monotone class)$\Rightarrow \cup B_i$, $A \cup (\cup B_i)$, $A \backslash (\cup B_i) \in m(\mathscr{A}) \Rightarrow \cup B_i \in \mathscr{G}_A$.
- $\mathscr{A} \subseteq \mathscr{G}_A$. Fix $A \in \mathscr{A}$, since $\mathscr{A}$ is a field (it is also a ring and is closed under the formation of finite unions), $\forall A' \in \mathscr{A}$, $A', A \cup A', A \backslash A' \in \mathscr{A}$. Thus, $\mathscr{A} \subset \mathscr{G}_A$ which indicates $m(\mathscr{A}) \subset \mathscr{G}_A$. Furthermore,

$$A \in \mathscr{A}, B \in m(\mathscr{A}) \Rightarrow A \cup B, A \backslash B \in m(\mathscr{A}) \tag{1.1.4}$$

$\forall B \in m(\mathscr{A})$, let

$$\mathscr{H}_B = \{A : A, A \cup B, A \backslash B \in m(\mathscr{A})\}$$

- Similarly, $\mathscr{H}_B$ is a monotone class.
- According to formula 1.1.4, $\mathscr{A} \subseteq \mathscr{H}_B$. It follows that $m(\mathscr{A}) \subseteq \mathscr{H}_B$,

$$A, B \in m(\mathscr{A}) \Rightarrow A \cup B, A \backslash B \in m(\mathscr{A}) \tag{1.1.5}$$

□

Theorem 1.1.4 (π-λ theorem)**.** *If $\mathscr{A}$ is a π-system, then $\sigma(\mathscr{A}) = l(\mathscr{A})$ where $l(\mathscr{A})$ denotes the smallest Dynkin system (λ-system) containing $\mathscr{A}$. It is tantamount to: Let $\mathscr{P}$ be a π-system and $\mathscr{L}$ be a Dynkin system with $\mathscr{P} \subseteq \mathscr{L}$, then $\sigma(\mathscr{P}) \subseteq \mathscr{L}$.*

Example 1.1.1 (Corollary 1.3.6 in [2])**.** *$\mathscr{P}_R = \{(-\infty, a] : a \in \mathbb{R}\}$ is a π-system; $\mathscr{R}_R = \{(a, b] : a, b \in \mathbb{R}\}$ is a semiring. The definition of Borel system of sets on $\mathbb{R}$ is*

$$\mathscr{B}_R = \sigma(\mathscr{P}_R) = \sigma(\mathscr{R}_R)$$

1.2 Measurable Mappings

Definition 1.2.1 (Topological measurable space)**.** *For topological space X, we denote the collection of open sets $\mathcal{O}$, we call $\mathscr{B} = \sigma(\mathcal{O})$ Borel algebra (system) of sets on X. And $(X, \mathscr{B})$ is called topological measurable space. But we need to always notice that topology (X, τ) and measurable space $(X, \mathscr{A})$ are different. One defines the open set structure, while another is defined on σ-field.*

Proposition 1.2.1 (Properties of preimage of sets)**.** *......*
As a simple corollary of preimage properties, for any set system $\mathscr{E}$ on Y,

$$\sigma(f^{-1}\mathscr{E}) = f^{-1}\sigma(\mathscr{E})$$

Definition 1.2.2 (Measurable mappings). *$(X, \mathscr{A}) \xrightarrow{f} (Y, \mathscr{F})$ is a measurable mapping if*

$$f^{-1}\mathscr{F} \subseteq \mathscr{A}$$

Theorem 1.2.2. *Let $\mathscr{E}$ be any set system on Y, then*

$$(X, \mathscr{A}) \xrightarrow{f} (Y, \sigma(\mathscr{E})) \text{ is a measurable mapping} \iff f^{-1}\mathscr{E} \subseteq \mathscr{A}$$

which can be proved by means of the corollary in Proposition 1.2.1.

Definition 1.2.3 (Measurable function). *The measurable mapping $f : (X, \mathscr{A}) \to (\overline{\mathbb{R}}, \mathscr{B}_{\bar{\mathbb{R}}}))$ is called measurable function on $(X, \mathscr{A})$. The measurable mapping $f : (X, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}_R))$ is called random variable (or finite measurable function) on $(X, \mathscr{A})$. Note that for image in $\bar{\mathbb{R}}$ of measurable function, we only care about Borel set. With the σ-algebra property of $(X, \mathscr{A})$ and Theorem 1.2.1, we only need to care about $f^{-1}((a, +\infty])$.*

There is a very nice example to understand the difference between topology and measurable space.

Example 1.2.1. *Assume f is continuous, g is measurable. Then $f \circ g$ is measurable, but $g \circ f$ is not necessarily measurable.*

Theorem 1.2.3 (Approximation by simple functions, Proposition 2.1.8 in [1]). *(1) Let $(X, \mathscr{A})$ be a measurable space, let A be a subset and let f be a $[0, +\infty]$-valued measurable function on A. Then there is a sequence $\{f_n\}$ of $[0, \infty)$-valued simple functions that satisfy $f_1(x) \leq f_2(x) \leq \ldots$ and $f(x) = \lim_n f_n(x), x \in A$.*
(2) For any measurable function f, there is a sequence of simple functions such that $|f_n(x)| \leq f(x)$ and $f(x) = \lim_n f_n(x), x \in A$.

If f is bounded, then the convergence is uniform.

1.3 Measure

Definition 1.3.1. *A measure (or a countably additive measure) on $\mathscr{A}$ (could be semiring, ring, algebra, σ-algebra, $\cdots$) is a function $\mu : \mathscr{A} \to [0, +\infty]$ that satisfies $\mu(\varnothing) = 0$ and is countably additive.*

Theorem 1.3.1 (Measure on semiring). *The measure on a semiring exhibits monotonicity, subadditivity, semiadditivity, upper semicontinuity, and lower semicontinuity.*

Lemma 1.3.2. *If $\mathscr{R}$ is a semiring,*

$$r(\mathscr{R}) = \bigcup_{n=1}^{\infty} \left\{ \bigcup_{k=1}^{n} A_k : \{A_k \in \mathscr{R}, k = 1, \ldots, n\} \text{are disjoint} \right\} \tag{1.3.1}$$

where $r(\cdot)$ is the ring generation function on the set system. See Theorem 1.3.2 in [2].

Proof of Theorem 1.3.1. Semiadditivity. Assume $A_1, A_2, \ldots \in \mathscr{R}$ and $\bigcup_{n=1}^{\infty} A_n \in \mathscr{R}$, then $A_1, A_2, \ldots \in r(\mathscr{R}) \Rightarrow \bigcup_{i=1}^{n-1} A_i \in r(\mathscr{R}) \Rightarrow A_n \backslash \bigcup_{i=1}^{n-1} A_i \in r(\mathscr{R})$ (definition of ring). Thus, according to Lemma 1.3.2, $\exists$ disjoint sets $\{C_{n,k} \in \mathscr{R}, k = 1, \ldots, k_n\}$ such that,

$$A_n \backslash \bigcup_{i=1}^{n-1} A_i = \bigcup_{k=1}^{k_n} C_{n,k}$$

Similarly,

$$A_n \backslash \bigcup_{k=1}^{k_n} C_{n,k} = \bigcup_{l=1}^{l_n} D_{n,l}$$

We have

$$A_n = (\bigcup_{k=1}^{k_n} C_{n,k}) \cup (\bigcup_{l=1}^{l_n} D_{n,l}) \tag{1.3.2}$$

Then we can easily derive $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \bigcup_{n=1}^{\infty} \mu(A_n)$ by means of additivity of μ. For other properties see Proposition 2.1.4 in [2]. ∎

1.4 Outer Measure

Definition 1.4.1. *Let X be a set, and let $\mathcal{P}(X)$ be the collection of all subsets of X. An outer measure on X is a function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ such that it is a monotone and countably subadditive function with $\mu^*(\varnothing) = 0$.*
(a) $\mu^*(\varnothing) = 0$,
(b) if $A \subseteq B \subseteq X$, then $\mu^*(A) \leq \mu^*(B)$, and
(c) if $\{A_n\}$ is an infinite sequence of subsets of X, then $\mu^*(\cup A_n) \leq \sum \mu^*(A_n)$.
Remark: Measure is defined on σ-algebra, while outer measure is defined on a set X.

Theorem 1.4.1 (Construction of outer measure, Theorem 2.2.1 in [2]). *Let $\mathscr{E}$ be a set system and $\varnothing \in \mathscr{E}$, $\mu : \mathscr{E} \to [0, +\infty]$ be a non-negative set function satisfying $\mu(\varnothing) = 0$. Then the function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$*

$$\mu^*(A) = inf \left\{ \sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathscr{E}, A \subseteq \bigcup_{i=1}^{\infty} B_n \right\} \tag{1.4.1}$$

is an outer measure.

We can prove Lebesgue Outer Measure is an outer measure.

Definition 1.4.2 (μ^*-measurable). *Let μ^* be an outer measure on X, we say A is μ^*-measurable if*

$$\mu^*(D) = \mu^*(D \cap A) + \mu^*(D \cap A^c)$$

for any $D \in \mathscr{A}$ We denote all μ^*-measurable sets $\mathscr{F}_{\mu^*}$.

Definition 1.4.3 (σ-finiteness). *Let μ be a measure on a measurable space $(X, \mathscr{A})$. Then μ is a finite measure if $\mu(X) < +\infty$ and is a σ-finite measure if X is the union of a sequence $A_1, A_2, ...$ of sets that belong to A and satisfy $\mu(A_i) < +\infty$ for each i.*

Let X be any set with at least two points, take the trivial σ -algebra $\mathscr{F} = \{X, \varnothing\}$, and define μ on F by $\mu(X) = \mu(\varnothing) = 0$.

1.5 The Extension of Measure

Definition 1.5.1 (Completeness and Regularity). *Let $(X, \mathscr{F}, \mu)$ be a measure space. The measure μ (or the measure space $(X, \mathscr{F}, \mu)$ is complete if the relations $A \in \mathscr{F}$, $\mu(A) = 0$, and $B \subseteq A$ together imply that $B \in A$. A positive μ is regular, if any compact set has finite measure; any measurable set can be approximated from above by open sets; any open set can be approximated from below by compact sets*

Theorem 1.5.1 (General extension theorem, Theorem 2.2.2 in [2]). *If μ^* is an outer measure, them $\mathscr{F}_{\mu^*}$ is a σ-field, and $(X, \mathscr{F}_{\mu^*}, \mu^*)$ is a complete measurable space.*

We want to extend the measure μ on $\mathscr{E}$ to a larger set system $\mathscr{F}$. Can we construct the outer measure μ^* using Formula 1.4.1 and extend it to $\mathscr{F}$ by means of Theorem 1.5.1?
The answer is no. There is no restriction to ensure $\mathscr{E} \subseteq \mathscr{F}_{\mu^*}$

Theorem 1.5.2 (Carathéodory extension theorem I). *Let μ be a σ-finite measure on an semialgebra $\mathscr{S}$. Then μ has a unique extension to $\sigma(\mathscr{S})$. See [2].*

Theorem 1.5.3 (Carathéodory extension theorem II). *Let μ be a σ-finite measure on an algebra $\mathscr{A}$. Then μ has a unique extension to $\sigma(\mathscr{A})$. See [3].*

In the following of this chapter, we focus on extending the measure on semiring $\mathscr{R}_R$ in Example 1.1.1.

Proposition 1.5.4. *Let F be a nondecreasing right-continuous function on $\mathbb{R}$ and for any $a < b \in \mathbb{R}$,*

$$\mu((a, b]) = F(b) - F(a)$$

otherwise $\mu((a, b]) = 0$. Then μ is a measure on $\mathscr{R}_R$.

Definition 1.5.2 (Definition I of Lebesgue measure). *(1) According to Theorem 1.5.2, μ has a unique extension on $\sigma(\mathscr{R}_R) = \mathscr{B}_R$.*

(2) According to Theorem 1.5.1, the outer measure μ_F^* is also a measure on $\mathscr{F}_{\mu_F^*}$. We call sets in $\mathscr{F}_{\mu_F^*}$ Lebesgue-Stieljes measurable sets and μ_F^* on $\mathscr{F}_{\mu_F^*}$ Lebesgue measure. If $F(x) = x$, then μ_F^* is called Lesbegue measure.

(3) Note we have $\sigma(\mathscr{R}) \subseteq \mathscr{F}_{\mu_F^*}$, which is mentioned in the proof of Theorem 2.3.2 in [2]. Theorem 2.3.4 in [2] and Appendix A.2 in [3] contain more details about the difference between $\sigma(\mathscr{R})$ and $\mathscr{F}_{\mu_F^*}$. Proposition 2.1.11 in [1] presents a Lebesgue measurable set which is not a Borel set.

(4) Theorem 2.4.2 in [2] gives a more accurate description: Let μ^* be the outer measure generated by measure μ defined on semiring $\mathscr{R}$, then $(X, \mathscr{F}_{\mu^*}, \mu^*)$ is the completion of $(X, \sigma(\mathscr{R}), \mu^*)$.

Definition 1.5.3 (Definition II of Lebesgue measure). *We can also start with Lebesgue outer measure denoted as λ^*.*

$$\lambda^*(A) = inf\left\{\sum_i (a_i - b_i) : (a_i, b_i) \text{ is a open interval.}\right\} \tag{1.5.1}$$

Lebesgue measurable subset is the set that is λ^*-measurable in accord with Definition 1.4.2, the set of which is denoted as M_{λ^*}. Then Lebesgue measure is λ^* restricted on M_{λ^*} and is denoted by λ.

2 Integrals and Convergence

2.1 Integral of Non-Negative Functions

Theorem 2.1.1 (Monotone convergence theorem). *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and $f_1, f_2, ...$ be $[0, +\infty]$-valued $\mathscr{A}$-measurable functions on X. Suppose that $f_1(x) \leq f_2(x) \leq ...$ and $f(x) = \lim_{n\to\infty} f_n(x)$ hold at μ-almost every x in X. Then*

$$\int f d\mu = lim_n \int f_n d\mu \tag{2.1.1}$$

In this theorem the functions f and $f_1, f_2, ...$ are only assumed to be nonnegative and measurable; there are no assumptions about whether they are integrable.

Proof. To prove monotone convergence theorem, we need to use "Continuity of measure on sets". □

There are two important corollaries of monotone convergence theorem, specially the interchangibility between addition and integration for non-negative functions is a result for the corresponding property of simple functions and monotone convergence theorem.

Corollary 2.1.2 (Beppo Levi's theorem). *If X is a measurable space, and $f_1, f_2, ...$ are a sequence of $[0, \infty]$-valued measurable functions, then*

$$\int \sum f_i d\mu = \sum \int f_i d\mu$$

Corollary 2.1.3 (Interchange between addition and integration). *If X is a measurable space, and f, g are a two $[0, \infty]$-valued measurable functions, then*

$$\int (f+g)d\mu = \int f d\mu + \int g d\mu$$

The following lemma is equivalent to monotone convergence theorem. However, we only show how to prove Fatou's lemma using monotone convergence theorem in this note. Also, the intuition is, can we have the exchange result as Formula 2.1.1 for any sequence of $[0, \infty]$-valued functions.

Theorem 2.1.4 (Fatou's lemma, Theorem 2.4.4 in [1]). *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let $\{f_n\}$ be a sequence of $[0, +\infty]$-valued $\mathscr{A}$-measurable functions on X. Then*

$$\int \varliminf_n f_n d\mu \leq \varliminf_n \int f_n d\mu \tag{2.1.2}$$

Proof. We define a monotone sequence $g_k = \inf_{m \geq k}\{g_m\}$, and we have $\varliminf_{n\to\infty} f_n = \sup_{n\to\infty} g_n = \lim_{n\to\infty} g_n$ by definition. Then we can use the monotone convergence theorem. □

Example 2.1.1. *Assume $f_n(x) = n\mathbb{1}_{[0,\frac{1}{n}]}$. The relation in Formula 2.1.2 is stricly "<".*

Lemma 2.1.5. *Let f be a non-negative $[0, \infty]$-valued $\mathscr{A}$-measurable function such that $\int f d\mu$ is finite. Then f is finite almost everywhere. We can use this corollary to prove Borel-Cantelli lemma.*

2.2 Integral of Signed Functions

Proposition 2.2.1 (Triangular inequality for absolute integral). *If f is an absolutely-integral function,*

$$\begin{aligned}
|\int f d\mu| &= |\int f^+ d\mu - \int f^- d\mu| \\
&\leq \int f^+ d\mu + \int f^- d\mu \\
&= \int f^+ d\mu + \int f^- d\mu && \text{(using Corollary 2.1.3)} \\
&= \int |f| d\mu
\end{aligned}$$

Theorem 2.2.2 (Dominated convergence theorem.). *Let $(X, \mathscr{A}, \mu)$ be a measure space, let g be $[0, +\infty]$-valued integrable function on X, and let f and $f_1, f_2, ...$ be $[0, +\infty]$-valued $\mathscr{A}$-measurable functions on X and such that $f(x) = \lim_n f_n(x)$ and $|f_n(x)| \leq g(x)$ hold at μ-almost every x in X, where g is absolutely integrable. Then f and $f_1, f_2, ...$ are integrable, and*

$$\int f d\mu = lim_n \int f_n d\mu$$

Corollary 2.2.3. *We can prove the interchangeability between limit and integral for uniform convergence using dominated convergence theorem.*

2.3 Riemann Integral and Lebesgue Integral

About the consistency between Riemann integral and Lebesgue integral, [1] gives a classical description while [6] gives a very sharp argument from a clever angle.

2.4 L^p Space

Definition 2.4.1 (L^∞). *Assume (X, μ) is a measurable space, $f : X \to \bar{\mathbb{R}}$ is measurable.*

$$\|f\|_{L^\infty} = esssup_{x\in X}|f|(x) = \inf\{\alpha : |f| < \alpha \ for \ a.e.x\}$$

Theorem 2.4.1 (Holder inequality). *Let (X, μ) be a measure space, and let p and q satisfy $1 \leq p \leq +\infty$, $1 \leq q \leq +\infty$, and $1/p + 1/q = 1$. If $f \in L^p(X)$ and $g \in L^p(X)$, then fg belongs to $L^1(X)$ and satisfies*

$$|fg|d\mu \leq \|f\|_{L^p}\|g\|_{L^q}$$

Proposition 2.4.2 (Properties of L^p)**.**
(1) L^p is a vector space.
(2) Minkowski inequality. Let $f \in L^p, g \in L^p$ then $f+g \in L^p$ and $\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$

Proof. We can use holder inequality to prove (2) in Proposition 2.4.2. □

Theorem 2.4.3 (Completeness of $L^p, 1 \leq p \leq \infty$)**.** *Let (X, μ) be a measure space, and let p satisfy $1 \leq p \leq +\infty$. Then $L^p(X, \mu)$ is complete under the norm $\|\cdot\|_{L^p}$. That is for any Cauchy sequence $\{f_m\}$ in L^p, there exsits $f \in L^p(X)$, and $\{m_k\}$ such that*
(1) $f_{m_k} \to f$ almost everywhere;
(2) $|f_{m_k}| \leq g$, where $g \in L^p$.
and also

$$\lim_{m\to\infty} \|f - f_m\|_{L^p} = 0$$

Before the proof, we want to show, neither "convergence in $L^p \Rightarrow$ almost everywhere convergence" or " almost everywhere convergence $\Rightarrow$ convergence in L^p" are correct.
(1) Counterexample of "convergence in $L^p \Rightarrow$ almost everywhere convergence". The indicator functions of interval $[0, 1/2], [1/2, 1], [0, 1/4], [1/4, 1/2], ...$ is a good exmaple. Actually convergence in $L^p \Rightarrow$ convergence in measure, and $\Rightarrow$ there is a subsequence that is almost everywhere convergent.
(2) Counter example of "almost everywhere convergence $\Rightarrow$ convergence in L^p". Consider $f_m(x) = m \cdot \mathbb{1}[0, \frac{1}{m}]$

Proof. Strategy: find a sparse sequence first. Find a subsequence $\{f_{m_k}\} \in L^p(X)$ such that

$$\|f_{m_k} - f_{m_{k+1}}\|_{L^p} \leq 2^{-k}$$

Define

$$f(x) = f_{m_1}(x) + \sum_{k=1}^{\infty} f_{m_{k+1}}(x) - f_{m_k}(x)$$

$$g(x) = |f_{m_1}| + \sum_{k=1}^{\infty} |f_{m_{k+1}}(x) - f_{m_k}(x)|$$

where $\|g\|_{L^p} \leq \|f_{m_1}\|_{L^p} + \sum_{k=1}^{\infty} \|f_{m_{k+1}} - f_{m_k}\|_{L^p} \leq \|f_{m_1}\|_{L^p} + 1$. Thus, g is finite almost everywhere. To continue the proof, we need to firstly show that f is well defined. By the finiteness of g, we have f_{m_k} is absolutely convergence almost everywhere, and therefore f_{m_k} is convergent to f almost everywhere.

To show $\lim_{m\to\infty} \|f - f_m\|_{L^p} = 0$, we can first show that $\lim_{k\to\infty} \|f - f_{m_k}\|_{L^p} = 0$. With $\|f - f_{m_k}\|_{L^p} = \sum_{j=k}^{\infty} \|f_{m_{j+1}} - f_{m_j}\|_{L^p}$, $f - f_{m_k} \in L^p$ and thus $f \in L^p$. Also,

$$\|f - f_{m_k}\|_{L^p} \leq 2^{-(k-1)} \to 0$$

□

Theorem 2.4.4 (Separability of $L^p, 1 \leq p < \infty$)**.** *L^p is separable if $1 \leq p < \infty$.*

2.5 Between Measurable Function and Continuous Functions

Lusin's theorem, converse theorem of Lusin's theorem and Frechet's theorem.

3 Some Complementary Parts

3.1 Modes of Convergence

Definition 3.1.1 (Almost everywhere finite and almost everywhere bounded). *Almost everywhere finiteness is a pointwise property while almost everywhere boundedness is a global property.*

Definition 3.1.2 (Modes of convergence).
(1) Uniform convergence;
(2) Pointwise convergence;
(3) Almost everywhere convergence;
(4) Convergence in measure;
(5) Almost uniform convergence (not almost everywhere);
(6) Convergence in mean (generally, convergence in L_p);
(7) Convergence in distribution (in context of probability theory; is in correspondence with weak convergence of cdf [2])

Proposition 3.1.1 (Relations of different convergence modes). *Let $(X, \sigma(\mathscr{A}), \mu^*)$ be a measure space, and let f and $f_1, f_2, \cdots$ be almost everywhere finite $\mathscr{A}$-measurable functions on X to $\bar{\mathbb{R}}$.*
(1)

$$\begin{Bmatrix} \mu \text{ is finite} + \text{almost everywhere convergence} \Rightarrow \text{ convergence in measure} \\ \text{convergence in measure} \Rightarrow \text{a subsequence of } \{f_n\} \text{ converges to } f \text{ almost everywhere} \end{Bmatrix}$$

(2)

$$\begin{Bmatrix} \text{almost uniform convergence} \Rightarrow \text{almost everywhere convergence} \\ \text{(Egoroff's Theorem) } \mu \text{ is finite+ almosteverywhere convergence} \Rightarrow \text{almost uniform convergence} \end{Bmatrix}$$

(3) Convergence in L_p $\Rightarrow$ convergence in measure.

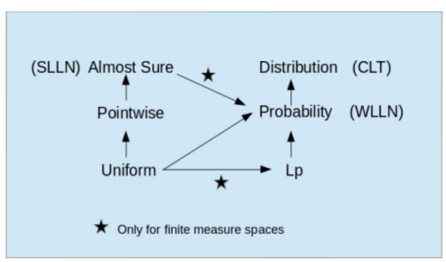

Figure 1: Relations of different convergence modes. [4]

3.2 Some Results about Differentiation

The most classical result is: Vitali covering theorem $\Rightarrow$ Lebesgue Theorem.

Theorem 3.2.1 (Lebesgue Theorem). *Let $F : \mathbb{R} \to \mathbb{R}$ be nondecreasing. Then F is differentiable almost everywhere (in Lebesgue measure).*

Theorem 3.2.2 (Alexandrov Theorem). *Let $F : \mathbb{R} \to \mathbb{R}$ be convex. Then F is secondly differentiable almost everywhere.*

3.3 Riesz Representation Theorem

The regularity of a measure allows many approximations and calculations that would be impossible without it. In particular, various linear functionals can be represented in a useful way with regular measures. It is carefully discussed in Section 7 in [1], and it is widely used in comparison between different topologies in Optimal Transport [7].

$$\begin{aligned}\mathcal{M}(X) &:= \{\text{finite signed measures on } X\} && (3.3.1)\\ &= C_c(X)^* := \{\text{continuous compactly supported functions}\}^* && (3.3.2)\\ &= C_0(X)^* := \{\text{continuous functions vanishing at } \infty\}^* && (3.3.3)\end{aligned}$$

4 Signed Measure

We need to understand the relationship between signed measure, absolutely continuous function, and function of bounded variation.

4.1 Hahn Decomposition

Definition 4.1.1 (Signed measure)**.** *A signed measure on $\mathscr{A}$ is a function $\mu : \mathscr{A} \to \bar{\mathbb{R}}$ such that,*
(1) $\mu(\varnothing) = 0$,
(2) μ is countably additive,
(3) Either $\mu(A) \in [-\infty, +\infty)$ or $\mu(A) \in (-\infty, +\infty]$.

Definition 4.1.2 (Positive set)**.** *A subset A of X is a positive set if $A \in \mathscr{A}$ and each $\mathscr{A}$-measurable subset E of A satisfies $\mu(A) \geq 0$.*

Theorem 4.1.1 (Hahn Decomposition Theorem)**.** *Let $(X, \mathscr{A})$ be a measurable space, and let μ be a signed measure on $(X, \mathscr{A})$. Then there are disjoint subsets P and N of X such that P is a positive set for μ, N is a negative set for μ, and $X = P \cup N$. The decomposition in essentially unique.*

Theorem 4.1.2 (Jordan Decomposition Theorem)**.** *Every signed measure is the difference of two positive measures, at least one of which is finite.*
(1) Let μ be a signed measure on $(X, \mathscr{A})$. Choose a Hahn decomposition (P, N) for μ and then define functions μ^+ and μ^- on $\mathscr{A}$ by

$$\mu^+(A) = \mu(A \cap P) \tag{4.1.1}$$
$$\mu^-(A) = -\mu(A \cap N) \tag{4.1.2}$$

(2) Actually,

$$\mu^+(A) = sup\{\mu(B) : B \in \mathscr{A} \text{ and } B \subseteq A)\} \tag{4.1.3}$$
$$\mu^-(A) = sup\{-\mu(B) : B \in \mathscr{A} \text{ and } B \subseteq A)\} \tag{4.1.4}$$

Proposition 4.1.3. *Let $(X, \mathscr{A})$ be a measurable space. Then the spaces $M(X, \mathscr{A}, \mathbb{R})$ (finite signed measure) is complete under the total variation norm $||\mu|| = |\mu|(X)$ where the variation is*

$$|\mu| = \mu^+ + \mu^- \tag{4.1.5}$$

4.2 Radon–Nikodym Theorem

Definition 4.2.1 (Absolutely consinuous)**.** *Let $(X, \mathscr{A})$ be a measurable space, and let μ and υ be positive measures on $(X, \mathscr{A})$. Then υ is absolutely continuous with respect to μ ($\upsilon \ll \mu$) if*

$$\mu(A) = 0 \Rightarrow \upsilon(A) = 0, \forall A \in \mathscr{A}$$

Let υ' be a signed measure on $(X, \mathscr{A})$. We say $\upsilon' \ll \mu$ if $|\upsilon| \ll \mu$.

Theorem 4.2.1 (Radon–Nikodym theorem)**.** *Let $(X, \mathscr{A})$ be a measurable space, and let μ and υ be σ-finite positive measures on $(X, \mathscr{A})$. If $\upsilon \ll \mu$, then there is an $\mathscr{A}$-measurable function $g : X \to [0, +\infty)$ such that $\upsilon(A) = \int_A g d\mu$ holds for each $A \in \mathscr{A}$. The function g is unique up to μ-almost everywhere equality.*

Proof. TBD see [1]. □

Theorem 4.2.2 (Radon–Nikodym theorem (signed))**.** *Let $(X, \mathscr{A})$ be a measurable space, let μ be a σ-finite positive measure on $(X, \mathscr{A})$, and let υ be a finite signed measure on $(X, \mathscr{A})$. If $\upsilon \ll \mu$, then there is a function g that belongs to $(X, \mathscr{A}, \mu, \mathbb{R})$ and satisfies $\upsilon(A) = \int_A g d\mu$ for each $A \in \mathscr{A}$. The function g is unique up to μ-almost everywhere equality.*

4.3 Lebesgue Decomposition

Definition 4.3.1 (Sigularity)**.** *Let μ, υ be signed measures. Then μ and υ are mutually singular if $\exists N \in \mathscr{A}$ such that $|\mu|(N) = |\upsilon|(N^c) = 0$.*

Theorem 4.3.1 (Lebesgue decomposition)**.** *Let $(X, \mathscr{A})$ be a σ-finite measurable space, let μ, υ be a σ-finite signed measure on $(X, \mathscr{A})$, then there exists two σ-finite signed measures μ_c, μ_s such that*
(1) $\mu = \mu_c + \mu_s$
(2) $\mu_c \ll \mu$
(3) $\mu_s \perp \upsilon$

Proof. See [2]. □

References

[1] Donald L Cohn. *Measure theory.* Vol. 5. Springer, 2013.

[2] Shihong Cheng. *Foundation of Measure Theory and Probability Theory.* Peking University Press, 2004.

[3] Richard Durret. *Probability: Theory and Examples.* Duxbury Press.

[4] David Mandel. *Notes on Modes of Convergence in Probability Theory.* URL: https://www.math.fsu.edu/~dmandel/Primers/Convergence.pdf.

[5] John K. Hunter. *Notes on Measure Theory.* URL: https://www.math.ucdavis.edu/~hunter/measure_theory/measure_notes.pdf.

[6] Terence Tao. *Analysis II.* Hindustan Book Agency.

[7] Alessio Figalli and Federico Glaudo. *An invitation to optimal transport, Wasserstein distances, and gradient flows.* 2021.