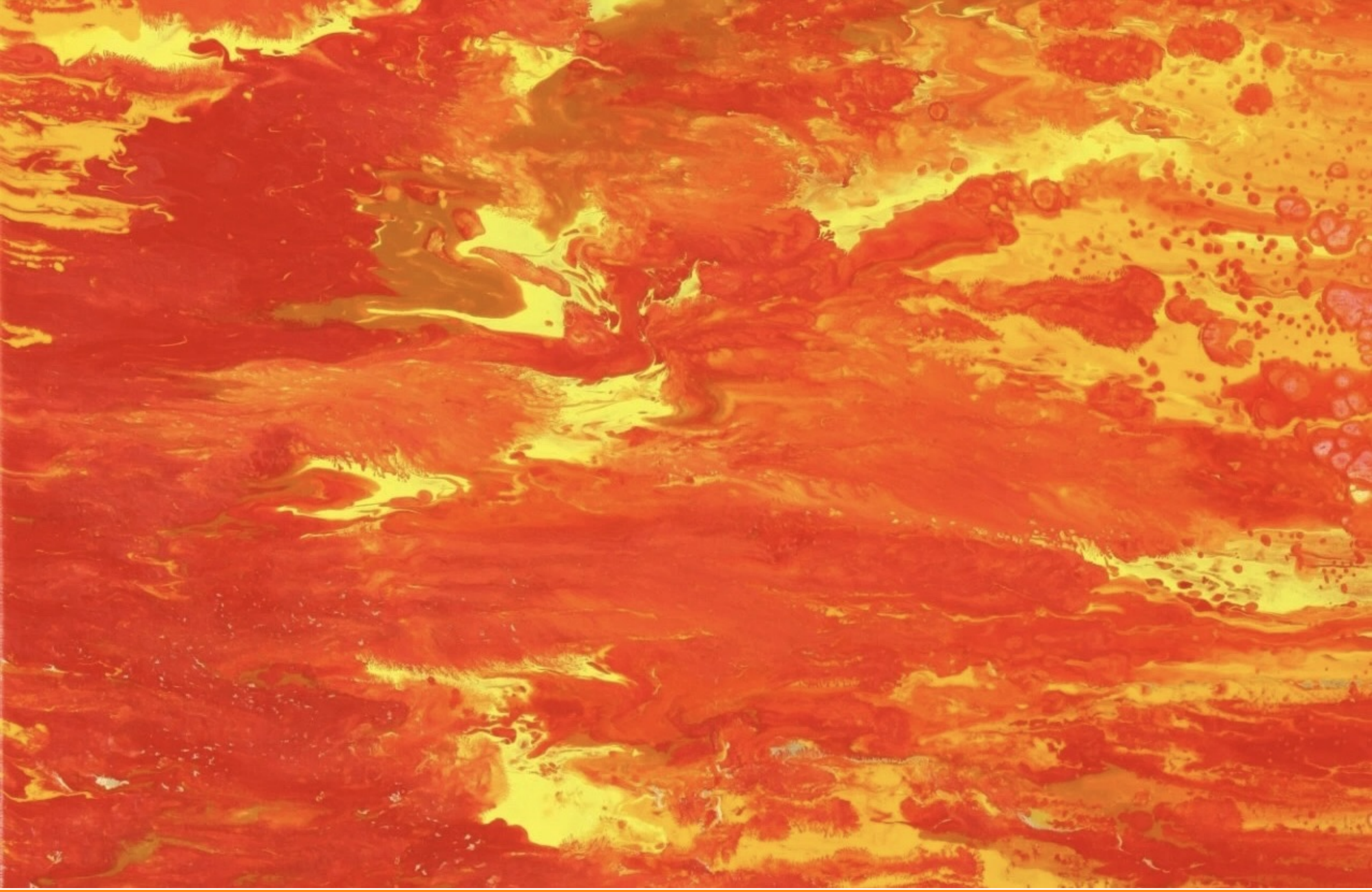

Measure Theory and Real Analysis

Foundation of Advanced Analysis and Probability Theory

Author: Wang Zhi

Institute: University of Science and Technology of China

Date: April 26, 2024

Contents

0.1 Introduction

Measure theory is a quintessential example of theory-building in mathematics instead of problem-solving (though a considerable amount of tricks is used throughout the process), what we usually do first is either to prove a weaker result and then extend it, such as assume $\mu(\Omega) < +\infty$, a function to be positive, or to prove a weaker result, such as proving μ to be finite-additive rather than σ-additive. We would also come to realize that a "countable" statement is not that different from a "finite" statement, though necessary, it's usually easy to extend finite cases to countable cases. In measure theory, we deal mostly with countable sets and many results we will see in this note fail without countability.

Example 0.1 For a sequence of measurable functions, the function defined by pointwise sup is also a measurable function, however, if we take an uncountable set $\mathcal{N}$ and notice that $\mathbb{1}_{\mathcal{N}} = \underset{t\in\mathcal{N}}{sup}\, \mathbb{1}_t$, then this function is measurable because of uncountability.

One technique that's been used many times in this note is the following: given a class of objects(sets or functions), called $\mathcal{B}$, we want to show that it has property F. When it is hard (or impossible) to verify explicitly, consider a subclass $\mathcal{C} \subset \mathcal{B}$ of all objects with property F, and prove that $\mathcal{B} = \mathcal{C}$ by using other overall characters of $\mathcal{B}$.

Example 0.2 proving that the preimage of a Borel set is in the σ-algebra for measurable functions, since we cannot write out a Borel set explicitly, we make use of the minimal property of Borel algebra as it is generated by all open sets; Stein's approach to proving Fubini's theorem.

One can easily notice that the Lebesgue measure is used much more often than the Borel measure, the reason is not simply that the former is an extension of the latter, most importantly, the Lebesgue measure is **a complete measure**, which enables many desired properties to hold.

Example 0.3 for two functions that agree almost surely, if one is Lebesgue measurable, so does the other, this provides us with a fundamental view that functions that agree almost surely are "the same", but this **does not** hold for Borel measurable functions.

Chapter 1 Theory of Sets and Measure

1.1 Algebra

Hi! Let's start with two most important definitions in this note and start to build the foundation of

Definition 1.1

A family of sets $\mathcal{P} \subset 2^{\Omega}$ is called a semi-algebra if the following three conditions are met

(1) $\emptyset \in \mathcal{P}$

(2) $\forall A \in \mathcal{P}$, A^c *can be written as a finite union of elements in* $\mathcal{P}$.

(3) $\forall A, B \in \mathcal{P}$, $A \cap B \in \mathcal{P}$

is called an algebra if

(1) $\emptyset \in \mathcal{P}$

(2) closed under complement.

(3) closed under finite intersection(union).

and called a σ-algebra if

(1) $\emptyset \in \mathcal{P}$

(2) closed under complement.

(3) closed under countable union.

♣

Definition 1.2

Let $\mathcal{F}$ be a σ-algebra, a set function $\mu : \mathcal{F} \to [0, +\infty] = \overline{\mathbf{R}}_+$ is said to be a measure if

(a) $\mu(\emptyset) = 0$

(b) *σ-additivity,* $\forall E_i \in \mathcal{F}, E_i \cap E_j = \emptyset$, $\mu(\bigcup_{i \geq 1} E_i) = \sum_{i \geq 1} \mu(E_i)$

♣

Remark the triplet $(\Omega, \mathcal{F}, \mu)$ is called a measure space and we assume it to be the background of our discussion without referring to it explicitly later.

You can verify two facts now: any intersection (needless to be finite or countable) of algebras remains an algebra, and any intersection of σ-algebras remains a σ-algebra.

Definition 1.3

the algebra generated by a set E is the intersection of all algebras that contains E, denoted $\mathcal{Q}(E)$, i.e the smallest algebra containing E, similarly define σ-algebra generated by E, denoted $\mathcal{F}(E)$.

♣

then start form a semi-algebra, we can write $\mathcal{Q}(E)$ explicitly.

Proposition 1.1

$\mathcal{Q}(E)$ is the collection of all finite union of sets in E.

♠

This enables us to extend a σ-additive/finite additive set function in E to $\mathcal{Q}(E)$ while preserving σ-additivity/finite additivity, this function takes $\emptyset$ to 0, of course. Unfortunately, σ-algebra generated by an algebra has a much finer structure and does not have a simple expression, which we will discuss in the next section. The continuity of measure is a frequently used property enabled by the σ-additivity.

Proposition 1.2

(a) for any increasing measurable set sequence $\{E_i | i = 1, 2, \cdots\}, E_i \subset E_{i+1}$*,* μ *is continuous from below in that* $\mu(\lim_{n\to\infty} E_i) = \lim_{n\to\infty} \mu(E_i)$*,* μ *is also continous from above if* $\exists E_{i_0}$ *s.t* $\mu(E_{i_0}) < +\infty$*.*

(b)

1.2 Monotone Class

Definition 1.4

$\mathcal{M} \subset 2^\Omega$, $A_k \in \mathcal{M}, A_k \uparrow A \to A \in \mathcal{M}$*, similarly,* $A_k \in \mathcal{M}, A_k \downarrow A \to A \in \mathcal{M}$*, i.e class is closed under increasing and decreasing sequences, then* $\mathcal{M}$ *is called a monotone class.*

we prove one result that will be used later for a more important theorem.

Theorem 1.1

$\mathcal{Q}$ *is an algebra, then* $\forall E, \mathcal{Q} \subset E$*, we have* E *is a* σ*-algebra* $\iff$ E *is a monotone class.*

1.3 Caratheodory Theorem

Definition 1.5 (σ-finite)

you are an idiot

Theorem 1.2

Ω *is* σ*-finite,* $\mathcal{Q}$ *an algebra with a* σ*-additive function* $\nu : \mathcal{Q} \to \overline{\mathbf{R}}_+$*, it can be extended uniquely to a measure in* $\mathcal{Q} = \mathcal{F}$*.*

Remark the statement also holds if we replace "algebra" by "semi-algebra", or "σ-additive" by "finite-additive". But once the σ-finite condition is removed then it may fail.

1.4 Hahn-Jordan Decomposition

First, we introduce the concept of signed measure and state some properties of it.

Definition 1.6 (signed measure)

a set function that takes value in extended reals and with σ*-additivity.*

Remark $\mu(\cap_{i=0}^{\infty} E_i) = \sum_{i=0}^{\infty} \mu(E_i)$, since the left-hand side does not depend on the order, the series on the right-hand side is absolutely convergent.

Lemma 1.1

If a signed measure does not obtains $+\infty$ *then it is bounded.*

Proof (This is a false proof) Suppose contradiction, $\mu(A_n) \geq 1 + \mu(\cup_{i=1}^{n-1} A_i), \mu(A_1) \geq 1$, use the technique to turn them disjoint $\{B_n\}$ and compute $\mu(B_n) \geq 1$, this fails because signed measure does not have monoticity.

Proposition 1.3

(1) $E \subset F$, $f(E) = \infty$, *then* $f(F) = \infty$, *same holds for* $-\infty$.

(2) f *cannot take value* ∞ *and* $-\infty$ *at the same time.*

(3) f *is continuous from above and from below.*

Next, we assume f to have domain $(-\infty, \infty]$.

Theorem 1.3 (Hahn-Jordan Decomposition)

$\exists P \in \mathcal{F}, N = P^c$, *such that* $\forall E \in P, F \in N \Rightarrow f(E) \geq 0, f(F) \leq 0$.

Proof Assume $+\infty$ is not attained, and define $\mu^+(A) = \underset{E \subset A, E \in \mathcal{F}}{sup} \mu(E)$, then prove this is σ-additive by

Remark the decomposition is not unique. But is unique up to null set.

Corollary 1.1

any signed measure can be decomposed into the difference between a positive part and a negative part, just like measurable functions.

1.5 Radon-Nikodym Decomposition

The Hahn-Jordan decomposition can be used to derive another decomposition of the signed measure, this time the decomposition will be unique, but first, let's get familiar with the concept of singularity.

Definition 1.7 (singularity)

two measures μ, ν *are singular of each other if* $\exists E \in \mathcal{F}$, $\mu(E) = 0, \nu(E^c) = 0$, *we use the notation* $\mu\nu$.

Theorem 1.4 (Radon-Nikodym)

for σ*-finite signed measure* ν, *given a* σ*-finite measure* μ

(1) $\nu = \nu_1 + \nu_2$ *s.t.* $\nu_1 \ll \mu$ *and* $\nu_2\mu$.

(2) uniqueness hold.

(3) $\exists f$ *which is measurable under current context,* $\nu_1(A) = \int_A f d\mu, \forall A \in \mathcal{F}$.

Proof Sketch of the proof: assume ν to be a measure instead of a signed measure, and further assume both ν, μ to be finite instead of σ-finite. Consider $\mathcal{H} = \{f \in \mathcal{F} | f \geq 0, \int_A f d\mu \leq \nu(A), \forall A \in \mathcal{F}\}$, and let $a = \underset{f \in \mathcal{H}}{sup} \int_\Omega f d\mu$, we want to show that there exists a $g \in \mathcal{H}$ s.t. $a = \int_\Omega g d\mu$, note here that g is not unique but only up to a zero measure set, then define $\nu_1(A) = \int_A g d\mu$. In order to show singularity, consider $\sigma_n = \nu_2 - \frac{1}{n}\mu$ and by Hahn decomposition it has a positive part $P_n \in \mathcal{F}$, $\int_A g + \frac{1}{n}\mathbb{1}_{P_n} d\mu \leq \nu_1(A) + \frac{1}{n}\mu(A \cap P_n) \leq \nu_1(A) + \nu_2(A) = \nu(A)$, this show that $g + \frac{1}{n}\mathbb{1}_{P_n} \in \mathcal{H}$, which in turns shows by sup that $\mu(P_n) = 0$, finally let $P = \bigcup_{n=1}^{\infty} P_n$.

1.6 Vitali's Lemma

You're probably familiar with the definition of compactness, which states one can choose a countable subcover from any open cover of a compact set. Vitali's lemma is also a lemma concerning covering, but first, let's discuss what is a Vitali (an Italian mathematician) covering.

Definition 1.8

consider $\mathbb{R}$ with Lebesgue measure, $\forall E$ (not necessarily a measurable set), $\mathcal{V} = \{I_\alpha | \alpha \in \mathcal{I}\}$, here $\mathcal{I}$ is the index set and I_α to be an interval, if $\forall \epsilon > 0, \forall x \in E, \exists I_\alpha \in \mathcal{V}, 0 < |I_\alpha| < \epsilon$ s.t $x \in I_\alpha$, then $\mathcal{L}$ is said to be a Vitali cover of E. ♣

We claim that for every Vitali cover, we can find a subcover (almost) consisting of countable disjoint intervals.

Theorem 1.5 (Vitali's Lemma)

$\forall$Vitali cover $\mathcal{V}$ of E, $\exists \mathcal{L} \subset \mathcal{V}, \mathcal{L} = \{I_j | j = 1, 2, \cdots\}$, $I_i \cap I_j = \emptyset$, and most importantly $\mu^*(E \setminus \bigcup_{i \geq 1} I_i) = 0$, here μ^* is the Lebesgue exterior measure. ♡

Proof

1.7 Non Measurable Sets

Even though we've discussed a lot about the theory of sets, one foundational problem remains which is that all sets seem to be measurable, is there even a non-measurable at all? sure, there are actually many non-measurable sets which shows that it is necessary to restrict ourselves to measurable sets, we will present one example first and then we show that $\forall E, \mu^*(E) > 0, \exists F \subset E \subset \mathbb{R}$, F is a non-measurable set.

Example 1.1 consider an equivalence relation in $[0, 1]$: $x \sim y \iff x - y \in \mathbb{Q}$, then choose one and only one element from each equivalent class $x_\alpha \in \epsilon_\alpha, \alpha \in A$, where A is the index set and $\mathcal{N} = \{x_\alpha | \alpha \in A\}$, obviously this construction relies on the choice of axioms. Denote all rationals in $[-1, 1]$ as $\{r_k\}$, then $\mathcal{N}_k = \mathcal{N} + r_k, [0, 1] \subset \bigcup_{k \geq 1} \mathcal{N}_k \subset [-1, 2]$. If $\mathcal{N}$ is a measurable set then so is $\bigcup_{k \geq 1} \mathcal{N}_k$, then by the σ-additivity, $1 \leq \sum_{k \geq 1} \mu(\mathcal{N}_k) \leq 3$, which is impossible.

Read the next theorem as entertainment rather than something you must remember.

Theorem 1.6

$\forall E, \mu^*(E) > 0, \exists F \subset E \subset \mathbb{R}$, F is a non-measurable set ♡

Proof firstly, a lemma: **any measurable subset $\mathcal{E} \subset \mathcal{N}$ has zero measure.**

$\mathcal{E}_k = \mathcal{E} + r_k, [0, 1] \subset \bigcup_{k \geq 1} \mathcal{E}_k \subset [-1, 2]$, but $\mu(\bigcup_{k \geq 1} \mathcal{E}_k) = \sum_{k \geq 1} \mu(\mathcal{E}_k)$, this is a series of constant which indicates that $\mu(\mathcal{E}) = 0$.

The next key fact is that $\mathbb{R}$ **is the disjoint union of all $\mathcal{N}_k$, with r_k be all rationals in** $\mathbb{R}$, then $E = \bigcup_{k \geq 1} (E \cap \mathcal{N}_k)$, if all of $E \cap \mathcal{N}_k$ is measurable then by lemma we know $\mu(E) = 0$, which is a contradiction, thus one subset is non-measurable.

1.8 Complete Measure

We've already understand the extension of a σ-additive set function ν from an algebra $\mathcal{Q}$ to the generated σ-algebra $\mathcal{F} = \sigma(\mathcal{Q})$, now we claim that under certain conditions, it is possible to go even further, extending the function μ to a measure on a larger σ-algebra $\hat{\mathcal{F}}$, denoted $\hat{\mu}$, which is unique and a complete measure. So let's see the definition first.

Definition 1.9

A measure $\mu : \mathcal{F} \to R^+ \cup \{+\infty\}$ *is complete if* $\forall F \subset E \in \mathcal{F}, \mu(E) = 0$, *then* $F \in \mathcal{F}$. ♣

1.9 Approximation

1.10 Product Measure

The definition of product measure is similar to that of product topology, in the sense that in both cases, we take the product of sets $\mathcal{F}_1 \times \mathcal{F}_2$ first, and then consider the open sets or σ-algebra generated by them, denoted $\mathcal{F}_1 * \mathcal{F}_2$, but the difference here is that we need to define a set function here in $\mathcal{F}_1 \times \mathcal{F}_2$ which is σ-additive, thus can be extended to $\mathcal{F}_1 * \mathcal{F}_2$.

Next, we extend the above process to define a countable product of measure spaces.

1.11 Convergence of Measure

Chapter 2 Measurable Function and Integration

2.1 Measurable Function

2.2 Littlewood's Three Principles

Although the notions of measurable sets and measurable functions represent new tools, we should not overlook their relation to the older concepts they replaced. Littlewood aptly summarized these connections in the form of three principles that provide a useful intuitive guide in the initial study of the theory.

(i) Every set is nearly a finite union of intervals. (see 2.4)

(ii) Every function is nearly continuous.

(iii) Every convergent sequence is nearly uniformly convergent.

Theorem 2.1 (Egorov)

♡

Proof warning: this proof is classic but it can be a real pain in the ass reading it.

By classic, I mean the idea of using simple sets operations to describe a complex set such as the set of all convergent points

$$\bigcap_{k\geq 1}\bigcup_{n\geq 1}\bigcap_{l\geq n}\{x : |f_l(x) - f(x)| \geq \frac{1}{k}\} = \{x| \lim_{n\to\infty} f_n(x) = f(x)\}$$

Theorem 2.2 (Lusin)

♡

2.3 Integrals and Properties

2.4 Convergence Theorems of Integrals

The first convergence theorem we'll introduce is the **monotone convergence theorem**.

Theorem 2.3 (monotone convergence theorem)

♡

Proof

Remark this theorem shows the existence of the measure determined by a non-negative integrable function as follows. For $f \geq 0$ defined on $(\Omega, \mathcal{F}, \mu)$

$$\mu_f(A) = \int_A f d\mu$$

we shall verify that this is indeed a measure,

We point out that the monotone convergence theorem is equivalent to the following Fatou's lemma, and we give two separate proofs for the two statements respectively.

Lemma 2.1 (Fatou)

$f_n \geq 0$, each a measurable function(needless to have finite integral),then

$$\int \underline{\lim} f_n d\mu \leq \underline{\lim} \int f_n d\mu$$

♡

Proof

Next corollary will be used many times in real analysis.

Corollary 2.1 (convergence of non-negative function series)

♡

we can dig even more treasure out of the monotone convergence theorem, such as the definition of absolute continuity.

Definition 2.1 (absolute continuity of measure)

♣

and the absolute continuity of Lebesgue integral.

Proposition 2.1

♠

The last one happens to be the most important one, in the sense that it is sometimes referred to as the soul of the Lebesgue integral theory.

Theorem 2.4 (dominated convergence theorem)

♡

2.5 Convergence Modes

We will first discuss three types of convergence modes in this section, you can find adequate resources for more modes of convergence in my notes for probability theory and from many other standard texts about probability theory.

Definition 2.2 (uniform convergent almost surely)

$\exists E \in \mathcal{F}, \mu(E^c) = 0$, *s.t.* $\forall \epsilon > 0, \exists n_\epsilon, \forall n \geq n_\epsilon \Rightarrow |f_n(x) - f(x)| \leq \epsilon, \forall x \in E$.

♣

Remark here we emphasize that the order of the statement matters a lot, if we exchange "$\exists E \in \mathcal{F}, \mu(E^c) = 0$" with "$\forall \epsilon > 0''$, then what we end up with is called "almost surely uniform convergent", if you're discerning enough you might have found this is exactly the mode of convergence in Egorov's theorem and funny enough the name of the mode is also obtained by exchanging the order, which is almost surely making me laugh.

Recall that in calculus we've shown that uniform convergence is equivalent to convergence in L^∞ norm, inspired by this we define essential sup and the distance induced by it.

Definition 2.3 (ess sup)

♣

and we can define the distance naturally, left to the readers to check that it's indeed a distance.

Theorem 2.5 (unif a.s $\iff$ converge in ess sup)

♡

Proposition 2.2

unif a.s $\Rightarrow$ ***a.s unif*** $\Rightarrow$***a.s***, *and a restatement of Egorov's theorem, given* $\mu(\Omega) < +\infty$, ***a.s*** $\Rightarrow$ ***a.s unif***.

Example 2.1

2.6 L_p Space and Bounded Linear Operators

We first present two well-known inequalities without proof.

Theorem 2.6

(a) $\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p \le +\infty$, $\|fg\|_1 \le \|f\|_p \|g\|_q$
(b) $\|f+g\|_p \le \|f\|_p + \|g\|_p$, $1 \le p \le +\infty$

Corollary 2.2

$\|f\|_p$ *is a norm for* $1 \le p \le +\infty$.

Example 2.2L_p **conv** $\not\Leftrightarrow$ **unif conv**
Example 2.3L_p **conv** $\not\Leftrightarrow$ **a.s conv**

the main result is that $L_p = \{f \in \mathcal{F} : \|f\|_p < +\infty\}$ space is a complete norm space, in other words, a Banach space.

Theorem 2.7

L_p *is complete, in the sense that every Cauchy sequence converges to a function in* L_p.

Proof Sketch of the proof: given a Cauchy sequence in L_p, find a subsequence that converges rapidly enough such that it is a.s convergent on $F \in \mathcal{F}, \mu(F^c) = 0$, thus the limit function is defined by the pointwise convergence of this subsequence and we consider here its equivalent class though still using the symbol f to refer to it, then prove that $f \in L_p$ and the most desired convergence.

Remark for a simple treatment in the case of L_1 see Stein's book for reference, it expresses the limiting function as a series and used monotone convergence to show the series converges a.s.

Next, we discuss the bounded linear operators in L_p which form a subspace of the second dual space of L_p, it is different from what we've seen in linear algebra that the second dual space is isomorphic to the original space, here we note that L_p is infinite-dimensional and thus has more intricate behavior. Such an operator T takes in a linear function(an element in the dual space of L_p) and outputs a real number in a linear fashion, and most importantly its norm is bounded, for some constant $c > 0$

$$|T(f)| \le c\|f\|_p$$

and the smallest such c is defined to be the norm of T. One urgent concern is whether there is an explicit expression for the bounded linear operator T, and we are happy to claim that the answer is positive.

Theorem 2.8

$\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p \le +\infty$, *T is a bounded linear operator* $\iff \exists g \in L_q$ *s.t* $T(f) = \int fgd\mu, \forall f \in L_p$

Proof

Remark the above expression is actually an inner product in the dual space of L_p, and this theorem coincides with the Riesz representation theorem.