A Friendly Introduction To Real Analysis II

Anqiao Ouyang

March 25, 2024

1 Introduction

In the second part of my series on Real Analysis, I will shift the focus from measure theory to the study of integration. This continuation delves into real variable function theory, emphasizing concepts and results in abstract spaces. Through this series, I aim to explore the foundational aspects of integration theory and its applications in the broader context of real analysis.

2 Prerequisite Knowledge

2.1 Riemann Intergral

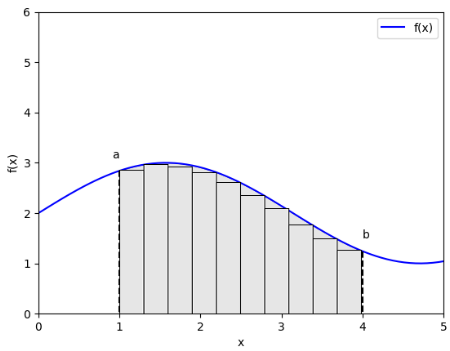
(As far as I know) readers at the high school level or above should have already encountered the concept of definite integrals. The definite integrals mentioned in general textbooks refer to the Riemann integral. For functions defined on a finite closed interval, the Riemann integral provides a method for calculating the area under a curve.

Considering readers with no prior knowledge, I'll first introduce the most basic Riemann integral. The basic idea of the Riemann integral is to divide the interval into many small subintervals, then take samples on each subinterval, and finally find the sum of the products of the function values on these small subintervals and the corresponding subinterval lengths. This sum, in the limit as the subinterval lengths approach zero, gives the value of the definite integral.

2.2 Sampling, Partitioning, and Riemann Sums

Consider real numbers $a < b$, and take a finite number of points in the closed interval $I = [a, b]$ such that $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$. Thus, I is divided into a finite number of subintervals

$$I = [x_0, x_1] \cup [x_1, x_2] \cup \ldots \cup [x_{n-1}, x_n]$$



For each subinterval $I_i = [x_{i-1}, x_i]$, we take a point $x_{i-1} \leq \xi_i \leq x_i$. This operation is called **sampling and partitioning**.

Then, we find the function value at each ξ_i, $f(\xi_i)$. Defining the width of each subinterval as $\Delta x_i = x_i -$

x_{i-1}, we can calculate the area of each subinterval's rectangle. The sum defined by the **Riemann sum** is

$$R(f, x_0, x_1, ..., x_n) = \sum_{i=0}^{n} f(\xi_i) \cdot \Delta x_i$$

It is not difficult to see that as we take more sample partitions, i.e., finer partitions, the Riemann sum gets closer to the actual area under the curve. Therefore, by taking the limit, we can rigorously define the Riemann integral.

2.3 Riemann Integrable and Definition

A function is called Riemann integrable (Integrable) if its Riemann integral exists and is finite.

Consider a function $f(x)$ defined on $[a, b]$. The corresponding Riemann sum is denoted as $R(f, x_0, x_1, ..., x_n)$. I is a real number such that for any positive real number ε, there exists a $\delta > 0$ such that when each Δx is less than δ, there exists a positive integer N such that for all $n > N$, we have $|R(f, x_0, x_1, ..., x_n) - I| < \varepsilon$. As in the intuitive understanding, there will be some error in the actual integral, and this error should be controlled within a certain range. This ε is the requirement for the precision of the Riemann integral.

Thus, we say $f(x)$ is Riemann integrable on $[a, b]$ at this point

$$\lim_{n\to\infty} R(f, x_0, x_1, ..., x_n) = I$$

Considering that the Riemann integral of the function $f(x)$ exists, its Riemann integral is defined as

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{i=0}^{n} f(\xi_i) \cdot \Delta x_i$$

Alternatively, some documents/textbooks may state it as the width of each sample partition approaching zero. Let $\sigma = \max_{0\leq i\leq n}(\Delta x_i)$. If the Riemann integral of the function $f(x)$ exists, we can also denote it as

$$\int_a^b f(x)\mathrm{d}x = \lim_{\sigma\to 0} \sum_{i=0}^{n} f(\xi_i)\Delta x_i$$

2.4 Measure and Integral

We should first understand why we need the Lebesgue integral by starting with the limitations of the Riemann integral.

First, the space of Riemann integrable functions is incomplete. Consider the distance between two functions $f(x)$ and $g(x)$ in the space of Riemann integrable functions on the interval $[a, b]$

$$d(f, g) = \int_a^b |f(x) - g(x)|\,\mathrm{d}x$$

At this point, $(R[a,b], d)$ forms a metric space. We introduce a sequence of continuous functions $\{f_n\}$ on $[a,b]$. If they converge to a function f_m in this space of Riemann integrable functions, then for any positive real number ε, there exists a positive integer N such that when $n > N$, we have $d(f_n, f_m) < \varepsilon$. Recall from Real Analysis I: "If every Cauchy sequence in a metric space converges within that space, then it is a complete metric space." It is not necessarily the case that there exists a Riemann integrable function f_m such that

$$\lim_{n\to\infty} d(f_n, f_m) = 0$$

Thus, the space of Riemann integrable functions is incomplete.

Secondly, the conditions for Riemann integrability are not very friendly. The function must be bounded, and according to Lebesgue's Integrability Criterion, a function is Riemann integrable on a closed interval if and only if the set of its discontinuities has measure zero. Intuitively, this means that a Riemann integrable function must be "essentially" continuous. These conditions are quite stringent, and many "strange" functions in mathematics are restricted by them.

¿ Here's a document with a simple explanation of the properties. For a detailed understanding of Lebesgue's Integrability Criterion, please see [Lebesgue's Integrability Criterion](https://www.math.mcgill.ca/labute

Furthermore, even some common and important properties of Riemann integrals have stringent conditions. For example, for a sequence of functions $\{f_n\}$ on $[a,b]$ to have the property of interchanging limits and integrals

$$\int_a^b f \,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n \,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}x$$

the sequence $\{f_n\}$ must converge uniformly;

To improve the integral and extend it to any measure space, we need a more suitable definition of the integral. We know that a measure is a generalized concept of size for sets, and the Lebesgue integral is an integration method defined on measure spaces. It is a more general definition of the integral of a function over a set, compared to the Riemann integral, and is more flexible, capable of handling more "strange" functions, such as discontinuous or divergent functions.

A few days ago, a freshman friend from the math department asked me, "Since the Lebesgue integral is better than the Riemann integral, why do we still learn the Riemann integral first?" I want to start by saying that, strictly speaking, the Lebesgue integral and the Riemann integral cannot be simply compared in terms of superiority; you can at most say that the Lebesgue integral's definition is more general. Moreover, as we will understand from the definitions later, putting aside the emphasis on measure, the Lebesgue integral cannot be said to be a mere extension of the Riemann integral. They are just two different definitions.

Additionally, mathematics often follows the principle of "specialization in specific fields." Most people cannot avoid the so-called 'calculus,' which generally doesn't delve into the complex theories involving constructing very strange and unusual functions as in 'analysis.' From a learning content perspective, the gap can be quite significant. When we first learn "integrals" in higher mathematics, we might study various formulas and theorems, typically dealing with continuous and differentiable, ideally smooth, functions, and use techniques to calculate the integrals of these functions. However, readers who have studied analysis will notice that we rarely need to calculate a specific integral value here.

3 Introduction to Banach Spaces

To better introduce the Lebesgue integral later, let's first introduce the basic concept of Banach spaces. A more systematic introduction to "Banach spaces" will be covered in subsequent articles on functional analysis. At this stage of real analysis, we mainly focus on what are called L^p spaces.

¿ Note: Starting from this chapter, it is generally assumed that readers have already read the previous article **Friendly Real Analysis Guide I** or have a basic understanding of measure theory, so the basic concepts will not be fully repeated.

3.1 Normed Spaces

In normed spaces, the concept of length is introduced for vectors, called the norm. Consider a normed space K and its vector space X. The norm is a mapping

$$\|\cdot\| : X \to \mathbb{R}_{\geq 0}$$

and satisfies the following properties

+ Non-negativity: The value of the norm is a non-negative real number. + Homogeneity: For a vector $\mathbf{v}$ and any real number α, it satisfies $\|\alpha x\| = |\alpha| \cdot \|x\|$. This means that scaling the vector by a real number scales its norm accordingly. + Triangle Inequality: The norm satisfies $\|x + y\| \leq \|x\| + \|y\|$, meaning the norm of the sum of two vectors does not exceed the sum of their norms.

If there exists a constant n such that every point in the space can be represented by at most n ordered parameters (using a set of ordered parameters to uniquely represent a point in the space), then this space is finite-dimensional, and n is the dimension. A finite-dimensional normed space of dimension n is isomorphic to $\mathbb{R}^n$.

3.2 Classical Banach Spaces

The definition of a Banach space is straightforward: it is a complete normed space, meaning a vector space with a complete norm, where every Cauchy sequence converges.

For example, Euclidean space is a typical Banach space, with its norm being the Euclidean norm (L2) $\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$. Another example is the Hilbert space in functional analysis, which is a complete inner product space where the norm is induced by an inner product that satisfies the parallelogram law.

In functional analysis, we often study infinite-dimensional function spaces. However, in the context of real analysis, we only need to focus on the function spaces of Lebesgue integrable functions, known as L^p spaces.

3.3 p-Norm

In normed vector spaces, the distance metric is generally defined as the p-th root of the sum of the p-th powers of its components, hence the term p-norm. For example, when $p = 1$, it is known as the Manhattan norm (L1), given by $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$; when $p = 2$, it is the Euclidean norm (L2), given by $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$.

When the space is finite-dimensional or countably infinite, for a vector $\mathbf{x}$ in $\mathbb{R}^n$ and a non-negative real number p, the p-norm is defined as:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

However, when the space is infinite-dimensional and uncountable, we cannot use the above method for finite or countably infinite dimensions to define the norm.

Here is the definition. Given a measure space $(X, \mathcal{M}, \mu)$ and a generalized real-valued function $f : X \to \overline{\mathbb{R}}$, where f is a measurable function on X and $p \in [0, \infty]$, the p-norm of f, denoted $\|f\|_p \in [0, \infty]$, is defined as follows. When p is finite:

$$\|f\|_p = \left(\int_X |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}}$$

When $p = \infty$, it is called the essential supremum norm and is defined as:

$$\|f\|_\infty = \inf\{M \in \mathbb{R}_{\geq 0} : |f| \leq M\}$$

Generally, when $\mu(X) > 0$, $\|f\|_\infty$ is referred to as the essential supremum of f, which represents the supremum in the measure-theoretic sense: ignoring sets of measure zero, it is the supremum of the function over the remaining set.

It is important to note that, although we often refer to p-norms, when $p < 1$, the p-norm does not satisfy the definition of a norm. When $p = 0$, homogeneity is not satisfied because $\|kx\|_0 \neq |k| \cdot \|x\|_0$ may not hold. When $0 < p < 1$, the triangle inequality is not satisfied. For example, consider two non-zero vectors $x = 1$ and $y = -1$:

$$\|x + y\|_p = (|1|^p + |(-1)|^p)^{\frac{1}{p}} = 2^{\frac{1}{p}} \|x\|_p + \|y\|_p = 1 + 1 = 2$$

Clearly, in this case, $2^{1/p} > 2$, thus the triangle inequality is not satisfied.

3.4 L^p Spaces

L^p spaces are special normed linear spaces consisting of all functions defined on a domain with finite p-norm. These spaces are defined over measurable functions with respect to Lebesgue measure, where the p-norm is given by an integral with respect to the Lebesgue measure. Thus, L^p spaces are sometimes referred to as Lebesgue spaces. L^p spaces are also typical Banach spaces, and in particular, when $p = 2$, they are known as Hilbert spaces, such as $L^2(S^1)$ used in Fourier analysis. ¡sub¿This will be discussed further at a later time.¡/sub¿

Formally, L^p spaces can be real or complex. For simplicity, the field $\mathbb{K}$ is taken to be $\mathbb{R}$ or $\mathbb{C}$. Given a measure space $(X, \mathcal{M}, \mu)$, for $1 \leq p < \infty$, the L^p space is defined as:

$$L^p(X, \mu) = \left\{ f : X \to \mathbb{K} \;\middle|\; f \text{ is measurable}, \ \|f\|_p < \infty \right\}$$

¿ For simplicity, $L^p(X, \mu)$ will be abbreviated as $L^p(X)$ when there is no need for further clarification, and L^p will be used to refer to this class of spaces in general.

L^p spaces are vector spaces, typically defined through addition of functions and scalar multiplication. For all $f, g \in L^p(X)$ and $\lambda \in \mathbb{K}$, we have: + $(f+g)(x) = f(x)+g(x)$ + $\lambda f(x) = (\lambda f)(x)$

In practice with Lebesgue measure, we often use $L^p(\mathbb{R}^n, m)$, which is one of the most important and typical examples in real analysis. Sometimes, we also restrict the space to a specific interval, such as $L^p([a, b])$.

3.4.1 Hölder's Inequality

To start with, the Cauchy-Schwarz inequality states that for two vectors $\mathbf{v}$ and $\mathbf{u}$ in an inner product space V over real or complex numbers (i.e., $\mathbb{K}$), we have:

$$\|\langle \mathbf{v}, \mathbf{u} \rangle\| \leq \|\mathbf{v}\| \cdot \|\mathbf{u}\|$$

Hölder's inequality is considered one of the most important inequalities in the context of L^p spaces because it generalizes the commonly used Cauchy-Schwarz inequality. For real numbers $p, q \geq 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, and functions $f \in L^p(X)$ and $g \in L^q(X)$, Hölder's inequality is given by:

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$$

It is easy to see that when we apply the Cauchy-Schwarz inequality to functions in L^2 space, i.e., taking $p = q = 2$, Hölder's inequality reduces to the Cauchy-Schwarz inequality.

3.4.2 Minkowski Inequality

Consider finite positive real numbers $1 \leq p \leq \infty$ and functions $f \in L^p(X)$, $g \in L^q(X)$, then:

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

For $0 \leq p \leq 1$, the inequality becomes:

$$\|f + g\|_p \geq \|f\|_p + \|g\|_p$$

Clearly, when $p = 2$, the Minkowski inequality reduces to the triangle inequality in Euclidean space under the inner product setting. For other values of p, this inequality provides a useful estimate of norms in function spaces. Additionally, the Minkowski inequality has an integral form. For real numbers $p > 1$, if f and g are integrable on $[a, b]$, then:

$$\left(\int_a^b (|f + g|)^p \, \mathrm{d}x \right)^{\frac{1}{p}} \leq \left(\int_a^b |f|^p \, \mathrm{d}x \right)^{\frac{1}{p}} + \left(\int_a^b |g|^p \, \mathrm{d}x \right)^{\frac{1}{p}}$$

Similarly, for $0 \leq p \leq 1$ and non-negative f and g, the inequality is reversed.

Alright, that's enough for now. The concepts we've covered are sufficient for our upcoming discussions in "Real Analysis" within this article. More general topics will be discussed later in functional analysis.

4 Lebesgue Integration

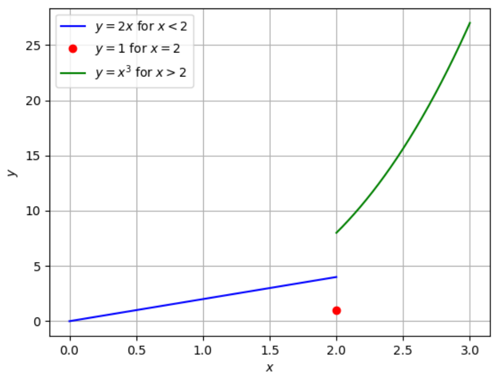
4.1 ”Almost Everywhere”

The term ”almost everywhere” refers to the property where, for a set considering a property P, if the set of elements that do not satisfy this property has measure zero, then we say that the property P holds almost everywhere. (Unless otherwise specified, we assume Lebesgue measure.)

Take the commonly mentioned example of a ”function that is continuous almost everywhere.” Consider the piecewise function:

$$f(x) = \begin{cases} y = 2x & \text{if } x < 2 \\ y = 1 & \text{if } x = 2 \\ y = x^3 & \text{if } x > 2 \end{cases}$$

It is clear that the function is discontinuous at $x = 2$, but this discontinuity occurs at just one point, which has measure zero. Therefore, we say that the function f is continuous almost everywhere.



4.2 Measurable and Integrable

Clearly, the Lebesgue measurability of a function refers to the property of the function being measurable in the measure space. Given a measure space $(X, \mathcal{M}, \mu)$, a function $f : X \to \mathbb{R}$ is Lebesgue measurable if for every real number a, the set $f^{-1}((-\infty, a]) = \{x \in X : f(x) \leq a\}$ is measurable in that measure space. This definition involves the preimages of intervals of the form $(-\infty, a]$.

4.2.1 Lusin’s Theorem

In real analysis, Lusin’s theorem states that every measurable function is almost continuous.

Given a Lebesgue measure space $(X, \mathcal{M}, m)$, for a measurable function f that is almost everywhere finite on X, then for every $\varepsilon > 0$, there exists a closed set $C_\varepsilon \subset X$ such that $m(X \backslash C_\varepsilon) \leq \varepsilon$, and f is continuous on C_ε.

This can also be extended to higher dimensions. For a regular Borel measure μ and its measurable function $f : \mathbb{R}^n \to \mathbb{R}^m$, if μ is a measurable set in $\mathbb{R}^n$ with finite measure, then for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset X$ such that $\mu(X \backslash K_\varepsilon) < \varepsilon$, and f is continuous on K_ε.

The significance of this theorem is that any measurable function on the real line can be approximated by continuous functions on its subset, with arbitrarily small error.

¡hr¿

Regarding integrability, we previously discussed the Riemann integral, which is defined on bounded closed intervals and only applies to bounded functions that meet the stringent Riemann integrability conditions. When we attempt to extend the Riemann integral to the entire real line, the Riemann integral may be undefined or infinite in some cases, such as when the function has unbounded points on the integration interval or when the limits of integration approach infinity. This makes the Riemann integral improper, which can lead to a loss of some desirable properties.

In contrast, the Lebesgue integral can be defined not only on bounded closed intervals but also on more general measure spaces, such as subsets of the real line or more general measure spaces.

To continue with the previous definitions, we decompose the function f into two parts

$$f^- = \max\{-f, 0\}$$
$$f^+ = \max\{f, 0\}$$
$$f = f^+ - f^-$$

Here, f^- and f^+ are referred to as the negative and positive parts of f, respectively. Both f^- and f^+ are non-negative measurable functions. If

$$\int_X f^- \mathrm{d}\mu < \infty \text{ and } \int_X f^+ \mathrm{d}\mu < \infty$$

that is, if both are finite, then the function f is said to be Lebesgue integrable. When considering $|f| = f^+ + f^-$, if

$$\int_X |f| \, \mathrm{d}\mu < \infty$$

then f is said to be absolutely integrable.

4.3 Basic Properties

Since proving certain special properties of the Lebesgue integral requires using some basic properties of the Lebesgue integral, let's first list a few of these properties. Given a measure space $(X, \mathcal{M}, \mu)$, and $A, B \in \mathcal{M}$, as well as $f, g, h \in L(X)$

$$g \le h, \quad \int_X g \, \mathrm{d}\mu \le \int_X h \, \mathrm{d}\mu \left| \int_A f \, \mathrm{d}\mu \right| \le \int_A |f| \, \mathrm{d}\mu \mu(A) = 0, \ \int_A f \, \mathrm{d}\mu = 0$$

Since L^p spaces are linear spaces, the Lebesgue integral has the **linearity property** as follows

$$c \in \mathbb{C}, \quad \int_X c \, f \, \mathrm{d}\mu = c \int_X f \, \mathrm{d}\mu \int_A (f + g) \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu + \int_X g \, \mathrm{d}\mu$$

The Lebesgue integral has the **countable additivity over regions of integration**, i.e.,

$$A \cap B = \emptyset, \quad \int_{A \cup B} f \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu$$

Furthermore, if f is integrable over $X = \bigcup_{i=1}^{\infty} E_i$, where each E_i is measurable and mutually disjoint, i.e., $\cap_{i=1}^{\infty} E_i = \emptyset$, then

$$f \ \mathrm{d}\mu = \sum_{i=1}^{\infty} \int_{E_i} f \ \mathrm{d}\mu$$

4.3.1 Convergence Theorems

Here, we present some theorems describing the convergence properties of function sequences, which make it easier to interchange limits and integrals. Given a measure space $(X, \mathcal{M}, \mu)$.

First, consider the Monotone Convergence Theorem. Suppose we have a sequence of measurable functions $\{f_n\}$ defined on X that is pointwise monotonically increasing or decreasing, i.e., for almost every $x \in X$, $f_1 \leq f_2 \leq \ldots \leq f_n \leq \ldots$ and they converge pointwise to a function f, $\lim_{n\to\infty} f_n = f$. Then we have

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

Next, let's discuss the Lebesgue Dominated Convergence Theorem.

Consider a sequence of measurable real functions $\{f_n\}$ defined on X such that $\forall x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$. If there exists a non-negative integrable function $g : X \to [0, \infty)$ such that for every $n \geq 0$ and for all $x \in X$, $|f_n(x)| \leq g(x)$, then both f and each f_n are integrable, and we can interchange the order of integration and limits:

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

¡br¿

The key difference is that the Monotone Convergence Theorem focuses on non-negative measurable functions that are pointwise increasing or decreasing, while the Lebesgue Dominated Convergence Theorem requires a more general sequence of pointwise convergent functions. The Monotone Convergence Theorem is simpler, requiring only the monotonicity and existence of the limit of the function sequence; the Lebesgue Dominated Convergence Theorem is more general, needing an integrable function to dominate the absolute values of the sequence, and thus can handle a broader range of function sequences.

4.4 Linear Combinations of Indicator Functions

An indicator function is a function that reflects whether an element belongs to a given set.

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Many texts categorize finite linear combinations of indicator functions as a type of Lebesgue integral definition, called ”simple functions.” Consider non-negative real numbers $a_1, a_2, \ldots, a_n$, measurable sets $E_1, E_2, \ldots, E_n$ in X, and the indicator functions χ_{E_i} for each E_i. A non-negative simple function is defined as:

$$\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$$

The Lebesgue integral of this function is defined as:

$$\int_X \phi \, \mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

It is agreed that $0 \cdot \infty = 0$, meaning that if the measure of a set is infinite, the integral of any function over that set is zero.

It is important to note that the value of the integral does not depend on the specific way the simple function is represented as a linear combination of indicator functions. This means that the integral value is uniquely determined and does not depend on our choice of coefficients and measurable sets.

The simple Lebesgue integral can be seen as a generalization of the finite additivity concept of measures. Let's revisit the finite additivity of measures from this perspective. For a collection of disjoint measurable sets $E_1, E_2, \ldots, E_n$, finite additivity is:

$$\mu\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mu(E_i)$$

A simple function is defined as $\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$, and according to the properties of the Lebesgue integral, we have:

$$\int_X \phi \, \mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

The simple Lebesgue integral essentially involves a weighted sum of the product of the measure of sets and their associated weights, generalizing the concept of measure additivity to a broader context.

4.5 Non-negative Measurable Functions

To handle more general functions, we further extend the definition to non-negative measurable functions on $\overline{\mathbb{R}}$. With the foundation laid earlier, this is quite straightforward.

For a non-negative measurable function $f : X \to [0, \infty]$, the Lebesgue integral is sometimes referred to as the lower integral, denoted with a underline:

$$\underline{\int_X} f \, \mathrm{d}\mu = \sup\left\{\int_X \phi \, \mathrm{d}\mu : \phi \text{ is a simple function}, 0 \leq \phi \leq f\right\}$$

Alternatively, it can be defined as the upper integral:

$$\overline{\int_X} f \, \mathrm{d}\mu = \inf\left\{\int_X \phi \, \mathrm{d}\mu : \phi \text{ is a simple function}, 0 \leq f \leq \phi\right\}$$

Despite the different notations, these two definitions are equivalent. Furthermore, it is evident that for any simple function ϕ, its integral is:

$$\int_X \phi \, \mathrm{d}\mu = \underline{\int_X} \phi \, \mathrm{d}\mu = \overline{\int_X} \phi \, \mathrm{d}\mu$$

4.5.1 Term-by-Term Integration Theorem

Integration and summation operations are interchangeable. Consider a sequence of non-negative measurable functions $\{f_n\}$, then:

$$\int_X \sum_{k=1}^{\infty} f_k \, \mathrm{d}\mu = \sum_{k=1}^{\infty} \int_X f_k \, \mathrm{d}\mu$$

Its proof is straightforward.

Given a measure space $(X, \mathcal{M}, \mu)$ and a sequence of non-negative measurable functions $\{f_n\}$, define $S_n = \sum_{k=1}^{n} f_k$ which is a non-negative measurable function on X, increasing and converging to f. Consider the limit as n approaches ∞:

$$\lim_{n\to\infty} S_n = \sum_{k=1}^{\infty} f_k$$

Applying the Monotone Convergence Theorem and the linearity property to S_n

$$\int_X \lim_{n\to\infty} S_n \ = \int_X \sum_{k=1}^{\infty} f_k \, \mathrm{d}\mu$$

$$\lim_{n\to\infty} \int_X S_n = \lim_{n\to\infty} \int_X \sum_{k=1}^{n} f_k \, \mathrm{d}\mu = \lim_{n\to\infty} \sum_{k=1}^{n} \int_X f_k \, \mathrm{d}\mu = \sum_{k=1}^{\infty} \int_X f_k \, \mathrm{d}\mu$$

$$\int_X \sum_{k=1}^{\infty} f_k \, \mathrm{d}\mu = \sum_{k=1}^{\infty} \int_X f_k \, \mathrm{d}\mu$$

4.5.2 Chebyshev's Inequality

Chebyshev's Inequality, which can be described using measure theory, states that for a measurable set $E \subset \mathbb{R}^n$ and a measurable non-negative function $f : E \to [0, \infty]$, the following holds for any positive real number λ:

$$\mu(\{f > \lambda\}) \leq \frac{1}{\lambda} \int_E f \, \mathrm{d}x$$

In my view, Chebyshev's Inequality is primarily used in probability theory (though it is often expressed using measure theory). In probability theory, it describes the deviation of a random variable from its expected value. Specifically, it states that the deviation of a random variable from its mean is controlled by its variance.

4.6 Arbitrary Measurable Functions

Furthermore, for any measurable function f on X, the Lebesgue integral is defined as:

$$\int_X f \ \mathrm{d}\mu = \int_X f^- \ \mathrm{d}\mu - \int_X f^+ \ \mathrm{d}\mu$$

4.7 Fatou's Lemma

In an exam, if you consider the scores of the worst performers in each subject to calculate the total score, it will not be higher than the score of the worst performer overall. Similarly, for the integral of a sequence of non-negative measurable functions, the integral of the infimum of the limit will not be smaller than the limit inferior of the integrals. Fatou's Lemma is crucial because it is used to prove the Lebesgue Dominated Convergence Theorem.

For a sequence of non-negative measurable functions $\{f_n\}$, the following inequality holds:

$$\int_X \liminf_{n\to\infty} f_n \ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \ \mathrm{d}\mu$$

The proof of Fatou's Lemma is quite simple and can be done using the ***Monotone Convergence Theorem***.

Proof. Given a measure space $(X, \mathcal{M}, \mu)$ and a sequence of measurable functions $\{f_n\}$ defined on X, let $g_k = \inf_{n\geq k} f_n$. The functions g_k are the infima of a finite number of measurable functions and $\liminf_{n\to\infty} f_n = \lim_{n\to\infty} g_n$.

According to the Monotone Convergence Theorem, we have

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

Substituting this in, we get

$$\begin{aligned}
\int_X \liminf_{n\to\infty} f \ \mathrm{d}\mu &= \int_X \lim_{n\to\infty} g_n \ \mathrm{d}\mu \\
&= \lim_{n\to\infty} \int_X g_n \ \mathrm{d}\mu \\
&= \liminf_{n\to\infty} \int_X g_n \ \mathrm{d}\mu \\
&\leq \liminf_{n\to\infty} \int_X f_n \ \mathrm{d}\mu
\end{aligned}$$

■

Fatou's Lemma not only applies to sequences of functions taking non-negative values but can also be extended to functions taking values in $\overline{\mathbb{R}}$ under certain conditions. When the range of functions is $\overline{\mathbb{R}}$, if there exists an integrable non-negative function g such that $g \leq f_n$ for all n, then the inequality still holds.

If there exists a function f such that $\lim_{n\to\infty} f_n = f$ almost everywhere in X, then we have

$$\int_X \liminf_{n\to\infty} f_n \ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \ \mathrm{d}\mu$$

where the inequality holds in the case where f_n can take negative values as well.

$$\int_X f \ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \ \mathrm{d}\mu$$

4.8 Improper Integrals

Improper integrals refer to a classification of integrals that includes cases where the limits of integration are infinite or where the integrand has discontinuities within the interval of integration. Specifically, these are classified as **improper integrals** and **singular integrals**.

Here is a classic example of an improper integral:

$$\int_0^\infty \frac{\sin x}{x} \mathrm{d}x = \frac{\pi}{2}$$

This integral is well-known and frequently encountered in various contexts. However, if you naively try to express this as $\displaystyle\int_{[0,\infty)} \frac{\sin x}{x} \mathrm{d}x$, it does not hold in general. The standard definition of the Lebesgue integral involves the difference between the positive and negative parts of the function. For the integrand $\frac{\sin x}{x}$, its positive and negative parts oscillate as x approaches infinity, leading to convergence issues but not yielding finite values in the usual sense. Therefore, $\frac{\sin x}{x}$ does not have finite integral values for its positive and negative parts over $[0, \infty)$. However, when both parts are infinite, their difference can still be Lebesgue integrable in the limit sense.

The discussion of improper integrals brings us back to a fundamental question: "Is every Riemann integrable function necessarily Lebesgue integrable?" The conclusion is no. This underscores the point made earlier in "Measure and Integration" that the Lebesgue integral should not be considered merely an extension of the Riemann integral.

Next?

The introduction to the most fundamental aspects of real analysis is now complete. The main articles on Mathematical Analysis - Real Analysis will shift towards topics on abstract spaces. This series generally covers the basics and serves as a familiarization with the concepts for beginners. Therefore, for more advanced topics in this area, I may consider writing specialized content (though it will not be part of the "Introduction to Real Analysis" series).

When I translated this article into English, I had already graduated from junior high school. Due to the long time span and the increasing workload of self-study, my writing plan may change in the future.

References

[1] Cornell University. (n.d.). *The Lebesgue integral.* Retrieved from `https://www.math.cornell.edu/`

[2] Nelson, B. (n.d.). *The Lebesgue integral.* Retrieved from `https://www.math.purdue.edu/`

[3] Trinity College Dublin. (n.d.). *Chapter 4. The dominated convergence theorem and applications.* Retrieved from `https://www.tcd.ie/`

[4] Stump, J. (n.d.). *Introduction to the Lebesgue integral.* Retrieved from `https://www.math.washington.edu/`

[5] Stepaniants, G. (n.d.). *The Lebesgue integral, Chebyshev's inequality, and the Weierstrass approximation theorem.* Retrieved from `https://www.math.udel.edu/`

[6] University of Washington Math Department. (n.d.). *Fubini's theorem.* Retrieved from `https://www.math.washington.edu/`