Real Analysis

Hai Le **Professor**
LaTeX by Rui Zhou
University of Michigan

Winter 2024

Contents

Chapter 3

Differentiation

3.1 Hardy-Littlewood Maximal Inequality

Definition 3.1.1: L^1-Norm, Lebesgue Space

Suppose $(X, \mathcal{S}\, \mu)$ is a measure space. If $f : X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable, then the L^1-*norm* of f is denoted by $||f||_1$ and is defined by

$$\|f\|_1 := \int |f| \ \mathrm{d}\mu.$$

The Lebesgue space (L^1-space) is defined by

$$L^1(\mu) := \{f : \|f\|_1 < \infty\}.$$

Remark.

- The Lebesgue space is a Banach space, i.e., a complete normed vector space.
- If $f = g$ almost everywhere, we would consider f and g to be the same functions. In fact, they would at most differ in a set of measure 0.
- (Definition of Norm) $\|\cdot\|_1$ is a norm such that
  - $\|f\|_1 \geq 0$, and $\|f\|_1 = 0$ if and only if $f = 0$ almost everywhere.
  - $\|cf\|_1 = |c| \cdot \|f\|_1$ for $c \in \mathbb{R}$.
  - (Triangle Inequality) $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.

Next, we will first detour to two lemmas that are crucial to the induction of Hardy-Littlewood Maximal Inequality: Markov's Inequality and Vitali Covering Lemma.

Lemma 3.1.2: Markov's Inequality

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. If $h \in L^1(\mu)$, then

$$\mu(\{x \in X : |h(x)| > c\}) \leq \frac{1}{c} \cdot \|h\|_1$$

for any $c > 0$.

Proof for Lemma

$$\begin{aligned}
\mu(\{x \in X : |h(x)| > c\}) &= \int_{\{x \in X : |h(x)| > c\}} 1 \, \mathrm{d}\mu \\
&= \frac{1}{c} \int_{\{x \in X : |h(x)| > c\}} c \, \mathrm{d}\mu \\
&\leq \frac{1}{c} \int_{\{x \in X : |h(x)| > c\}} |h| \, \mathrm{d}\mu \\
&\leq \frac{1}{c} \int_X |h| \, \mathrm{d}\mu \\
&= \frac{1}{c} \cdot \|h\|_1
\end{aligned}$$

■

Before introducing Vitali Covering Lemma, we shall formally define the notion of $3I$.

Definition 3.1.3: $3I$

Suppose I is an bounded nonempty open interval of $\mathbb{R}$. $3I$ is defined as an interval with the same center as I and three times the length of I.

Lemma 3.1.4: Vitali Covering Lemma

Suppose $I_1, \cdots, I_n$ is a list of bounded nonempty intervals in $\mathbb{R}$. Then there exists a disjoint sublist $I_{k_1}, \cdots, I_{k_m}$ such that

$$I_1 \cup \cdots \cup I_n \subseteq 3I_{k_1} \cup \cdots \cup 3I_{k_m}.$$

Proof for Lemma

We use the idea of greedy algorithm to take the largest remaining interval that is disjoint from the previously selected intervals.

Let k_1 be such that

$$|I_{k_1}| = \max\{|I_1|, \cdots, |I_n|\}$$

Suppose $k_1, \cdots, k_j$ have been chosen. Let k_{j+1} be such that $\left|I_{k_{j+1}}\right|$ is as large as possible subject to the collection that $I_{k_1}, \cdots, I_{k_j}, I_{k_{j+1}}$ are disjoint. If there is no choice of k_{j+1} such that $I_{k_1}, \cdots, I_{k_{j+1}}$ are disjoint, then the procedure terminates. Because we start with a finite list, the procedure must eventually terminates after some number m of choices.

We want to show that for any $j \in \{1, \cdots, n\}$:

$$I_j \subseteq 3I_{k_1} \cup \cdots \cup 3I_{k_m}$$

If $j \in \{k_1, \cdots, k_m\}$, then we are done.

If $j \notin \{k_1, \cdots, k_m\}$, let I_{k_l} be the first that is not disjoint from I_j. Thus I_j is disjoint from $I_{k_1}, \cdots, I_{k_{l-1}}$. Because I_j is not chosen at the lth-step, which implies

$$|I_j| \leq |I_{k_l}|$$

Also, I_j and I_{k_l} are not disjoint.

$$I_j \subseteq 3I_{k_l} \subseteq 3I_{k_1} \cup \cdots \cup 3I_{k_l}$$

■

3 is the best constant here.

Definition 3.1.5: Hardy-Littlewood Maximal Function

Suppose $h : \mathbb{R} \to \mathbb{R}$ is a Lebesgue measurable function. Then the *(Hardy-Littlewood) maximal function* of h is the function $h^* : \mathbb{R} \to [0, \infty]$ defined by

$$h^*(b) := \sup_{t>0} \frac{1}{2t} \int_{b-t}^{b+t} |h| \ \mathrm{d}\lambda$$

$\frac{1}{2t} \int_{b-t}^{b+t} |h| \ d\lambda$ represents the average over $[b-t, b+t]$. Hence h^* aims to pick the largest average around b.

Proposition 3.1.6: Hardy-Littlewood Maximal Inequality

Suppose $h \in L^1(\mathbb{R})$. Then

$$\lambda(\{b \in \mathbb{R} : h^*(b) > c\}) \leq \frac{3}{c} \cdot \|h\|_1$$

for any $c > 0$.

Proof for Proposition.

We will use the following fact for a measurable set A:

$$\lambda(A) = \sup\{\lambda(F) : F \subseteq A \text{ and } F \text{ is closed and bounded}\}$$

Let F be a closed and bounded subset of $A = \{b \in \mathbb{R} : h^*(b) > c\}$. We hope to show $\lambda(F) \leq \frac{3}{c} \cdot \|h\|_1$.

For each $b \in F$, there exists $t_b > 0$ such that

$$\frac{1}{2t_b} \int_{b-t_b}^{b+t_b} |h| \ \mathrm{d}\lambda > c$$

Therefore

$$F \subseteq \bigcup_{b \in F} (b - t_b, b + t_b)$$

By Heine-Borel theorem, there exists a finite subcover of F. Hence, there exist $b_1, \cdots, b_n \in F$ such that

$$F \subseteq \bigcup_{k=1}^{n} (b_k - t_{b_k}, b_k + t_{b_k})$$

Denote $I_k = (b_k - t_{b_k}, b_k + t_{b_k})$. By Vitali covering lemma, we can extract a disjoint sublist of $\{I_k\}$ such that

$$\bigcup_{k=1}^{n} I_k \subseteq 3I_{k_1} \cup \cdots \cup 3I_{k_m}$$

Therefore,

$$\begin{aligned}
\lambda(F) &\leq 3\left(\lambda\left(I_{k_1}\right) + \cdots + \lambda\left(I_{k_m}\right)\right) \\
&< \frac{3}{c}\left(\int_{I_{k_1}} |h| \, \mathrm{d}\lambda + \cdots + \int_{I_{k_m}} |h| \, \mathrm{d}\lambda\right) \\
&\leq \frac{3}{c}\int_{-\infty}^{\infty} |h| \, \mathrm{d}\lambda \\
&= \frac{3}{c}\|h\|_1
\end{aligned}$$

There is a "best" constant instead of 3.

3.2 Lebesgue Differentiation Theorem

3.2.1 First Version

Theorem 3.2.1: Lebesgue Differentiation Theorem, First Version

Suppose $F \in L^1(\mathbb{R})$. Then

$$\lim_{t \to 0} \frac{1}{2t} \int_{b-t}^{b+t} |f(x) - f(b)| \, \mathrm{d}\lambda = 0$$

for almost every $b \in \mathbb{R}$.

Proof for Theorem

If f is a continuous every where, then

$$\frac{1}{2t}\int_{b-t}^{b+t} |f(x) - f(b)| \, \mathrm{d}\lambda \leq \sup_{x \in (b-t, b+t)} |f(x) - f(b)|$$

If f is continuous, then $f(x) \to f(b)$ when $t \to 0$.
Unfortunately, $f \in L^1(\mathbb{R})$, we have no concrete information on the continuity of f.

There exists a simple function h such that

$$\|f-h\|_1<\varepsilon$$

There exists a continuous function $g:\mathbb{R}\to\mathbb{R}$ such that

$$\|f-g\|_1<\varepsilon$$

Let $\delta>0$. For each $k>0$, there exists a continuous function $h_k:\mathbb{R}\to\mathbb{R}$ such that

$$\|f-h_k\|_1<\frac{\delta}{k2^k}$$

Let

$$B_k=\left\{b\in\mathbb{R}:\left|f\left(b\right)-f_k\left(b\right)\right|\leq\frac{1}{k}\text{ and }\left(f-f_k\right)^*\left(b\right)\leq\frac{1}{k}\right\}$$

Then

$$\begin{aligned}\mathbb{R}\backslash B_k&=\left\{b\in\mathbb{R}:\left|f\left(b\right)-f_k\left(b\right)\right|>\frac{1}{k}\text{ or }\left(f-f_k\right)^*\left(b\right)>\frac{1}{k}\right\}\\&=\left\{b\in\mathbb{R}:\left|f\left(b\right)-f_k\left(b\right)\right|>\frac{1}{k}\right\}\cup\left\{b\in\mathbb{R}:\left(f-f_k\right)^*\left(b\right)>\frac{1}{k}\right\}\\&:=A\cup B\end{aligned}$$

Apply Markov's and maximal inequality:

$$\begin{aligned}\lambda\left(A\right)\leq k\cdot\|f-h_k\|_1<k\cdot\frac{\delta}{k2^k}&=\frac{\delta}{2^k}\\\lambda\left(A\right)\leq 3k\cdot\|f-h_k\|_1&<\frac{3\delta}{2^k}\end{aligned}$$

Therefore,

$$\lambda\mathbb{R}\backslash B_k\leq\frac{4\delta}{2^k}=\frac{\delta}{2^{k-2}}$$

Let

$$B=\bigcap_{k=1}^{\infty}B_k$$

Then

$$\lambda\left(\mathbb{R}\backslash B\right)\leq\sum_{k=1}^{\infty}\frac{\delta}{2^{k-2}}=4\delta$$

Let $b\in B$ and $t>0$. For each $k>0$, we have

$$\begin{aligned}\frac{1}{2t}\int_{b-t}^{b+t}\left|f-f\left(b\right)\right|\ \mathrm{d}\lambda&\leq\frac{1}{2t}\int_{b-t}^{b+t}\left(\left|f\left(x\right)-f_k\left(x\right)\right|+\left|f_k\left(x\right)-f_k\left(b\right)\right|+\left|f_k\left(b\right)-f\left(b\right)\right|\right)\ \mathrm{d}\lambda\\&\leq\left(f-h_k\right)^*\left(b\right)+\frac{1}{2t}\int_{b-t}^{b+t}\left|h_k\left(x\right)-h_k\left(b\right)\right|\ \mathrm{d}\lambda+\left|f_k\left(b\right)-f\left(b\right)\right|\\&\leq\frac{1}{k}+\frac{1}{2t}\int_{b-t}^{b+t}\left|h_k\left(x\right)-h_k\left(b\right)\right|\ \mathrm{d}\lambda+\frac{1}{k}\end{aligned}$$

The intuition of the first version of Lebesgue Differentiation Theorem is that, when taking the average over a small interval around a point b, the average will converge the functional value $f(b)$ at b.

3.2.2 Second Version

The first version of Lebesgue Differentiation Theorem does not bring in the concept of derivative, which we have learned and got familiar in calculus. Soon we will see the second version of Lebesgue Differentiation Theorem interacts well with the derivative. The following gives the formal definition of derivative:

Definition 3.2.2: Derivative

Suppose $g : I \to \mathbb{R}$ is a function defined on an open interval I, and $b \in I$. The *derivative* of g at $b \in I$, denoted as $g'(b)$, is defined as

$$g'(b) = \lim_{t \to 0} \frac{g(b+t) - g(b)}{t}.$$

Theorem 3.2.3

Suppose $f \in L^1$. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) := \int_{-\infty}^{x} f \, \mathrm{d}\lambda.$$

Then

$$g'(b) = f(b)$$

if f is continuous at b.

Proof for Theorem

We have

$$\begin{aligned}
\left| \frac{g(b+t) - g(b)}{t} - f(b) \right| &= \left| \frac{\int_b^{b+t} f \, \mathrm{d}\lambda}{t} - f(b) \right| \\
&= \frac{1}{t} \left| \int_b^{b+t} (f(x) - f(b)) \, \mathrm{d}\lambda \right| \\
&\leq \sup_{x \in (b, b+t)} |f(x) - f(b)| \xrightarrow{t \to 0} 0
\end{aligned}$$

■

Theorem 3.2.4: Lebesgue Differentiation Theorem, Second Version

$g'(b) = f(b)$ for almost every $b \in \mathbb{R}$.

Proof for Theorem

$$\begin{aligned}\left|\frac{g(b+t)-g(b)}{t}-f(b)\right| &= \left|\frac{\int_b^{b+t} f \, \mathrm{d}\lambda}{t}-f(b)\right| \\ &= \frac{1}{t}\left|\int_b^{b+t} (f(x)-f(b)) \, \mathrm{d}\lambda\right| \xrightarrow{t\to 0} 0\end{aligned}$$

for almost every $b \in \mathbb{R}$, by the first version of Lebesgue Differentiation Theorem. ■