

Module 8 Hardy-Littlewood maximal function and Lebesgue differentiation theorem

8.1 Hardy-Littlewood max function and max theorem [Fol 3.4]

目前我们 finish 了 Folland 的 Ch1, Ch2.

我们将先跳过 Radon-Nikodym differentiation theory, 在讲完 L^p space theory 后再回到 Radon-Nikodym differentiation theory. 但是我们将会先将 differentiation theory 中的一个特殊部分: HL max theorem 和 Lebesgue differentiation theorem, 因为它们在 L^p space theory 中需要被用到.

此 lec 对应: Folland 3.4(1)

Differentiation theory 的 overview:

Radon-Nikodym derivative 并不是 classical calculus 的扩展 (classical calculus 表示变量之间的相对变化), 而是专门针对: 同一个 measure space 上, 一个测度对于另一个测度 (要求它们之间绝对连续) 的变化率.

$$d\nu = f d\mu$$

从而

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

这使得我们可以更改一个积分 with respect to 的测度.

其核心定理: Radon-Nikodym Theorem, 表示了在一般的 measure space 上, 这两个测度满足一定条件下, 这个 Radon-Nikodym derivative 的存在性; 而 LDT 提出一种在 **Euclidean space** 上, 求 **Radon Nikodym derivative** 的方法.

LDT 本身是一种将积分信息转换为点态信息的手段. 它表示对于 **locally integrable** 的函数, 局部积分平均值可以收敛到函数本身, **a.e.**

因而在知道 ν 和 μ 的情况下, LDT 提供了类比经典微积分中 "导数是局部变化率的极限" 的观点: 在 Euclidean space 上, Radon Nikodym derivative 等于局部均值:

$$f(x) = \lim_{r \rightarrow 0} \frac{\nu(B(r, x))}{\mu(B(r, x))} \text{ for a.e. } x$$

Radon Nikodym derivative 的应用: 比如在概率论中, **pdf/pmf** 都是 **cdf** 对于 **Lebesgue measure** 的 **Radon Nikodym derivative**; 在贝叶斯推理中, 给定先验 prior 和观测数据的分布, 后验分布 posterior 的密度可以通过 Radon-Nikodym 导数计算. 下面介绍概念:

8.1.1 L^1_{loc} and local average**Def 8.1 (locally integrable)**

如果 measurable $f : \mathbb{R}^n \rightarrow \mathbb{C}$ 在任意 bounded subset of \mathbb{R}^n 上的 integral 都 $< \infty$, 则称 function $f : \mathbb{R}^n \rightarrow \mathbb{C}$ 是 locally integrable 的, 写作 $f \in L^1_{loc}(m)$.



Remark 等价于 f 在任意 compact set 上的 integral 都 $< \infty$:

$$\int_K f \, dm < \infty$$

for all K ; 或者 f 在任意 open ball 上的 integral 都 $< \infty$:

$$\int_{B(r,x)} f \, dm < \infty$$

for all x, r .

这是比 $f \in L^1(m)$ 更弱的条件. For example: 连续函数一定是 L^1_{loc} 的.

Example 8.1 考虑

$$f(x) := |x|^p, \quad x \in \mathbb{R}^n$$

(使用 polar coord) 可验证:

$$f \in L^1_{loc}(m) \iff p > -n$$

Def 8.2 (average)

对于 $f \in L^1_{loc}(m)$, 以及 bounded and Lebesgue measurable $E \subset \mathbb{R}^n$ with $m(E) > 0$, 我们定义:

$$\bar{f}_E := \frac{1}{m(E)} \int_E f \, dm$$

为 f 在 E 上的 **average value**.

特别地, 当 E 为一个 ball $B(r, x)$ 时, 我们可以写作:

$$A_r f(x) := \bar{f}_{B(r,x)}$$

表示它在 x 为中心的 r 为半径的 ball 上的 average.

**Lemma 8.1**

对于任意 $f \in L^1_{loc}$, $A_r f(x)$ 都是 jointly continuous in r and x 的. ($r > 0, x \in \mathbb{R}^n$)



Proof Suppose $(x_j, r_j) \rightarrow (x, r)$ in $\mathbb{R}^n \times \mathbb{R}_{>0}$.

于是 for sure:

$$m(B(x_j, r_j)) \xrightarrow{j \rightarrow \infty} m(B(x, r))$$

并且 by DCT (取 $\chi_{B(r_0+1, x_0)}$ 作为 bound) 可以得到:

$$\int f \chi_{B(x_j, r_j)} \rightarrow \int f \chi_{B(x, r)}$$

8.1.2 Hardy-Littlewood maximal function

Def 8.3 (Hardy-Littlewood maximal function)

对于 $f \in L^1_{loc}$, 我们定义它的 HL maximal function 为:

$$Hf(x) := \sup_{r>0} A_r|f|(x)$$



HF maximal 函数 $Hf(x)$ 表示 f 的绝对值函数在 x 处能取到最大的 local average.

Theorem 8.1 (HL maximal function 是 measurable 的)

对于任意 $f \in L^1_{loc}$, Hf 都是 measurable 的.



Proof Follows from lemma.

$$(Hf)^{-1}((a, \infty)) = \bigcup_{r>0} (A_r|f|)^{-1}((a, \infty))$$

是 open 的, 因为 $A_r|f|$ ctn, ctn function 下 open set 的 preimage 也 open.

Corollary 8.1

如果 $f : \mathbb{R}^n \rightarrow [0, \infty)$ 是 lower semictn 的, 那么它一定 Borel (thus Lebesgue) measurable.



8.1.3 Vitali-type covering lemma

对于一个 ball $B = B(x, r)$ 以及一个 constant c , 我们定义:

$$cB := B(x, cr)$$

Lemma 8.2 (Vitali-type covering lemma)

For given collection of balls $\{B_j \subset \mathbb{R}^n\}_{j=1}^k$, 存在 **disjoint** subcollection $\{B_{j_1}, \dots, B_{j_m}\}$ 使得

$$\bigcup_{j=1}^k B_j \subset \bigcup_{i=1}^m (3B_{j_i})$$

(于是,

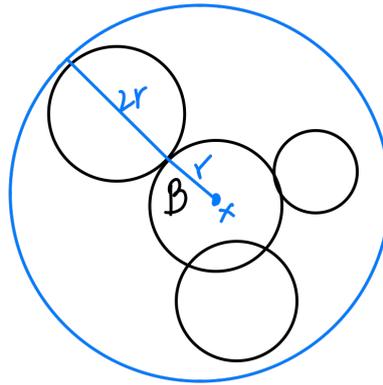
$$m \left(\bigcup_{j=1}^k B_j \right) \leq 3^n m \left(\bigcup_{i=1}^m (3B_{j_i}) \right)$$

)



Proof Greedy Algorithm: 直接按照半径大小排序, 取出最大的 disjoint subcollection.

Prove without words:



(每次都选择下一个和前面所有更大的球不 intersect 的最大球; 在这个过程中, 所有和前面更大的球有 intersection 的球都被包括在该球的三倍球里.)

8.1.4 Hardy-Littlewood maximal theorem

Theorem 8.2 (HL maximal theorem)

For $L^1(m^n)$, take constant $C := 3^n$, 则对于任意 $f \in L^1(m^n)$, 都有:

$$m(\{x : Hf(x) > \alpha\}) \leq \frac{C}{\alpha} \int |f|$$



Proof Set

$$E_\alpha := \{x : Hf(x) > \alpha\} \subset \mathbb{R}^n$$

因而 by def of Hf , 对于任意的 $x \in E_\alpha$ 都存在 r_x 使得

$$m(B(x, r_x)) < \frac{1}{\alpha} \int_{B(x, r_x)} |f|$$

对于 compact $K \in E_\alpha$, 一定存在 finite subcovering $\{B(x_i, r_i)\}$ covers K .

于是 Apply Vitali-type covering Lemma:

$$m(K) \leq \sum_i m(3B(x_i, r_i)) = 3^n \sum_i m(B(x_i, r_i)) \leq \frac{3^n}{\alpha} \int |f|$$

于是 by inner regularity, taking sup over all compact subsets 得证.

Remark 这是 HL max Thm 的一部分, 另一部分是 L^p space 下的. 这一部分表示了 HL maximal operator 的 **weakly boundedness**.

Notice: H 是一个从 L^p space (此处为 L^1) 到它自身的 (nonlinear) operator.

recall, 一个 operator 的 **(strongly) boundedness** 表示:

$$\|Tf\| \leq C\|f\|, \quad \text{for all } f \text{ (in the function space)}$$

而 weakly boundedness 表示一定的可控制性: 越大的函数值, 能取到这个函数值的点的占比 (with respect to measure) 越小.

(will prove: 对于 $p > 1$, H 具有 **strongly boundedness**.)

Remark 我们可以 compare the HL ineq 和 Markov ineq. Markov ineq 表示, 对于任意 $f \in L^1(\mu)$ 都有

$$m(\{f > \alpha\}) \leq \frac{1}{\alpha} \int |f|$$

8.2 Lebesgue differentiation Theorem [Fol 3.4]

对应: Folland 3.4(2)

Def 8.4 (Lebesgue set)

我们定义一个函数 $f \in L^1_{loc}(\mathbb{R}^n)$ 的 Lebesgue set 为:

$$L_f := \{x \in \mathbb{R}^n \mid \lim_{r \rightarrow 0^+} \int_{B(x,r)} |f(y) - f(x)| dy = 0\}$$

其中每个 point 被称为一个 **Lebesgue point**.



Remark 一个 Lebesgue point x 即满足:

$$\lim_{r \rightarrow 0^+} A_r |f - f(x)|(x) = 0$$

同时可证明, 这个条件也等价于:

$$\lim_{r \rightarrow 0^+} A_r f(x) = f(x)$$

即: 该点附近的 **average value of f** 等于 f 在该点的值. 因为

$$A_r f(x) - f(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} f dy - f(x) \quad (8.1)$$

$$= \frac{1}{m(B(x,r))} \int_{B(x,r)} f dy - \frac{f(x)}{m(B(x,r))} \int_{B(x,r)} 1 dy \quad (8.2)$$

$$= \frac{1}{m(B(x,r))} \int_{B(x,r)} [f - f(x)] dy \quad (8.3)$$

$$= A_r [f - f(x)](x) \quad (8.4)$$

它是 $A_r |f - f(x)|(x)$ 或其负数. 因而这两个函数趋向于 0 的趋势是相同的.

8.2.1 original LDT: locally L^1 函数几乎每一点附近的函数均值都等于这一点的值

Theorem 8.3 (Lebesgue differentiation theorem)

对于任意的 $f \in L^1_{loc}(\mathbb{R}^n)$, L_f is Leb mble and $m(L_f^c) = 0$.



Proof 由于对于任意 x 都有:

$$f(x) = \lim_{N \rightarrow \infty} f \chi_{B(0,N)}(x)$$

所以 it suffices to prove the statement for $f \chi_{B(0,N)}$ for 任意 N .

注意, $f \chi_{B(0,N)}$ 是一个 L^1 function. 因而只需要 prove the statement for $f \in L^1(\mathbb{R}^n)$ 就可以 generalize it to $f \in L^1_{loc}(\mathbb{R}^n)$. 因而 **WLOG suppose** $f \in L^1(\mathbb{R}^n)$.

首先, it is true for $f \in C_c^0(\mathbb{R}^n)$. (容易 check: 在某点连续性的定义能够 imply 在这一点均值等于这一点的值).

In other cases, 我们需要利用 $C_c^0(\mathbb{R}^n)$ 在 $L^1(\mathbb{R}^n)$ 中的 density.

Let

$$Q(x, r) := A_r |f - f(x)|(x)$$

这是一个 nonnegative function. 注意, 上个 lec 中我们证明了 $A_r f(x)$ 是 jointly continuous in r and x 的. 因而 $Q(x, r)$ 也是 jointly continuous in r and x 的.

我们随后定义:

$$Q(x) := \limsup_{r \rightarrow 0^+} Q(x, r)$$

于是 $x \mapsto Q$ 相当于 maximal function 的一个变体. 容易验证它也是 measurable 的. 我们 WTS:

$$m(\{x : Q(x) > 0\}) = 0$$

等价于 show:

$$m(\{Q > \alpha\}) = 0 \quad \text{for all } \alpha = \frac{1}{n}, n \in \mathbb{N}$$

Let $\epsilon > 0$. By density of $C_c^0(\mathbb{R}^n)$ in $L^1(\mathbb{R}^n)$, 我们可以 pick $g \in C_c^0(\mathbb{R}^n)$ s.t.

$$\int |f - g| < \epsilon$$

By triangular ineq in L^1 , for a.e. $x \in \mathbb{R}^n$ we have:

$$A_r |f - f(x)|(x) \leq A_r |f - g|(x) + A_r |g - g(x)|(x) + A_r |f(x) - g(x)|(x)$$

where

$$\text{(constant)} A_r |f(x) - g(x)|(x) = |f(x) - g(x)|$$

因而 putting $r \rightarrow 0^+$, we have $A_r |f - f(x)|(x) \rightarrow Q(x)$, 以及 $A_r |f - g|(x) \rightarrow H(f - g)(x)$; 从而上述不等式变为:

$$Q(x) \leq H(f - g)(x) + 0 + |f(x) - g(x)|$$

从而一定有:

$$m(\{Q > \alpha\}) \leq m(\{H(f - g) > \frac{\alpha}{2}\}) + m(\{|f - g| > \frac{\alpha}{2}\})$$

我们分别以 HL max Thm 和 Chebyshev's Thm bound 住右边这两个式子, 得到:

$$m(\{Q > \alpha\}) \leq \frac{2 \cdot 3^n}{\alpha} \int |f - g| + \frac{2}{\alpha} \int |f - g| \leq \frac{2}{\alpha} (3^n + 1) \epsilon \xrightarrow{\epsilon \rightarrow 0} 0$$

从而得证.

Corollary 8.2

$$x \in L_{loc}^1(\mathbb{R}^n) \implies \lim_{r \rightarrow 0^+} A_r f(x) = f(x) \text{ a.e.}$$



Proof If $x \in L_f$ then

$$\left| \frac{1}{|B(x,r)|} \int_{B(x,r)} (f(y) - f(x)) dy \right| \leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)| dy \xrightarrow{r \rightarrow 0} 0$$

Remark 我们先前已经 verify 了: Lebesgue point 的定义 $\lim_{r \rightarrow 0^+} A_r |f - f(x)|(x) = 0$ 和 $\lim_{r \rightarrow 0^+} A_r f(x) = f(x)$ 是等价的.

因而 Lebesgue Differentiation Thm 即表示: **locally** L^1 函数几乎每一点附近的函数均值都等于这一点的值.

实则, locally L^1 是一个很宽松的条件. 基本上, 只要是 measurable function 且不要有过多的 unbounded points, 这个函数就是 locally L^1 的. 所以 LDT 表示的是大多数可测函数几乎每一点附近的函数均值都等于这一点上的值.

Remark 在 Hw 8 中将 prove LDT 的 L^p 版本: 对于任意 $1 \leq p < \infty$, 如果 $f \in L^p(\mathbb{R})$, 都有

$$\lim_{r \rightarrow 0} \frac{1}{2r} \int_{x-r}^{x+r} |f(y) - f(x)|^p dy = 0 \quad \text{for a.e. } x$$

8.2.2 density of a set at a point

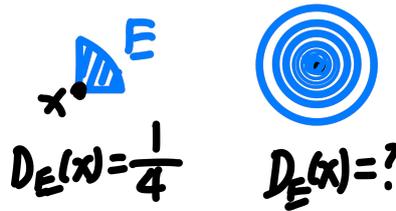
Def 8.5 (density of a set at a point)

对于 $E \subset \mathbb{R}^n$ Lebesgue measurable (which implies: $\chi_E \in L^1_{loc}$), 我们定义:

$$D_E(x) := \lim_{r \rightarrow 0^+} \frac{m(E \cap B(x, r))}{m(B(x, r))}$$



Remark density, 一个点附近一个集合占的密度, 理解很直观.



这是一个极限行为, 表示这个集合在这个点附近 locally 占的比例趋向于的极限.

如果是一个分布连续的集合, 比如一个扇形等, 那么就很好求. 但是如果是右图这样分布比较断断续续的集合, 会需要仔细考虑极限.

不过, LDT 告诉我们:

Corollary 8.3

对于 $E \subset \mathbb{R}^n$ Lebesgue measurable (which implies: $\chi_E \in L^1_{loc}$), 一定有:

$$D_E(x) = \begin{cases} 1, & \text{for a.e. } x \in E \\ 0, & \text{for a.e. } x \in E^c \end{cases}$$

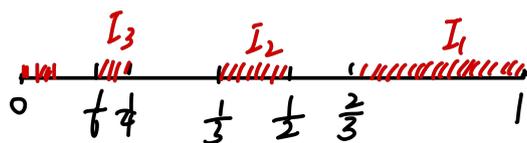


Proof 因为这个 indicator function 是 measurable 的, 以及 locally L^1 的. 所以它在 x 处的 density 就变成了它在 x 处的均值, 从而在 E 上 a.e. 为 1, 在 E^c 上 a.e. 为 0 (函数值).

Example 8.2 我们这里介绍一些 behavior 比较特殊的集合, 空间每点上这个集合的 density. 考虑

$$E := \{0\} \cup \bigcup_{j=0}^{\infty} \left[\frac{2}{3} \cdot \frac{1}{2^j}, \frac{1}{2^j} \right]$$

这是一个 closed set.



对于 $x \in E^\circ$, $D_E(x) = 1$.

对于 $x \notin E$, $D_E(x) = 0$.

对于 $x \in \partial E$, $D_E(x) = \frac{1}{2}$. (区间的一边在 E 里, 一边不在 E 里)

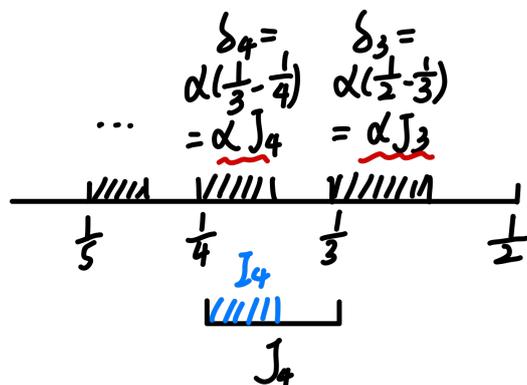
对于 $x \in 0$, $D_E(0)$ **undefined**. 因为每段空心和实心的地方, 这个比例的落差都非常大. (容易证明这个极限不存在.)

而反观, 任取 $\alpha \in (0, 1)$, 那么取

$$E := \bigcup_{n=1}^{\infty} \left(\frac{1}{n}, \frac{1}{n} + \frac{\alpha}{n(n-1)} \right)$$

则有:

$$D_E(0) = \frac{\alpha}{2}$$



这里的关键在于, harmonic seq 随着 n 的增长而缩小的速度非常慢. 在 n 较大的情况下, 背景区间 J_n 几乎与 J_{n+1} 具有相同的长度, 因此正如我们所知, $m(J_n)/m(\cup_{k>n} J_k) = 0$. 所以, 无论 r 位于实心部分 I_n 还是空心部分 $J_n \setminus I_n$ 都不太重要.

另一方面, 我们刚才的例子使用 geometric seq 作为背景区间 J_n 的构建块, 则 fail, 因为 J_n 的长度与 $\cup_{k>n} J_k$ 相比太大了, 有 $m(J_n) = m(\cup_{k>n} J_k)$, 因此无论 r 位于实心区间还是空心区间 这使得 0 处的密度无法定义.

8.2.3 generalized LDT

generalized LDT 表示对形状不规则 (未必是 ball) 的收敛行为, LDT 的 statement 仍然 stay true. 即, 只要 a family of Lebesgue mble sets E_r shrink nicely to x , LDT 就满足.

Theorem 8.4 (generalized LDT)

对于 $f \in L^1_{loc}(\mathbb{R}^n)$, 任意的 $x \in L_f$, 令 $\{E_r(x)\}$ 为 a family of Lebesgue measurable sets, 其中对于每个 $E_r(x)$ 都有:

$$E_r(x) \subset B(x, r)$$

并且

$$m(E_r(x)) > \alpha m(B(x, r))$$

for some $0 < \alpha < 1$.

则有:

$$\lim_{r \rightarrow 0^+} \frac{\int_{E_r(x)} |f(y) - f(x)| dy}{m(E_r(x))} = 0 \quad \text{for a.e. } x$$



证明很简单, 因为

$$\lim_{r \rightarrow 0^+} \frac{\int_{E_r(x)} |f(y) - f(x)| dy}{m(E_r(x))} \leq \lim_{r \rightarrow 0^+} \frac{\int_{B(x, r)} |f(y) - f(x)| dy}{m(E_r(x))} \leq \lim_{r \rightarrow 0^+} \frac{\int_{B(x, r)} |f(y) - f(x)| dy}{\alpha m(B(x, r))} = 0$$

Remark

LDT 最重要的作用是之一定义了 **在 Lebesgue 积分理论下的一种形式的 FTC:**

Theorem 8.5 (FTC in Lebesgue)

对于 $f \in L^1_{loc}(\mathbb{R})$, 任取 $x \in L_f$, 都有

$$\lim_{r \rightarrow 0} \frac{1}{r} \int_x^{x+r} f dm = f(x)$$



更 generally, 它表示由一个函数 f 在 measurable set 上对 measure m 的积分值定义出来的 measure ν , 对于 m 的 Radon-Nikodym 导数.

$$\lim_{r \rightarrow 0^+} \frac{\nu(E_r)}{m(E_r)} = f(x)$$

表示了 Euclidean space 上, 从 measure 比的极限恢复 Radon-Nikodym 导数的方法.

之后系统学习 Radon-Nikodym Theory, 将详细展开, 并引入更多的不等式和 FTC 形式. 而我们接下来将先展开 L^p space theory.