

YSW notes term

FE enthusiast enthusiast

2025.02.17

不会呜呜呜呜呜呜呜呜呜呜呜呜呜呜呜呜

目录

1	Lebesgue 积分理论, 续	3
1.1	Lebesgue积分复习	3
1.2	可积函数空间 L^1	10
1.3	Fubini 引理	12
1.4	Fubini 定理的应用	19
2	Lebesgue 微分定理	22
2.1	微分定理的第一个形式: Lebesgue 基本定理	22
2.2	Lebesgue 微分定理的应用	26
2.3	微分定理的第二个形式与单调函数	31
2.4	绝对连续函数与 Newton-Leibniz 公式	41
3	多元微分论	51
3.1	关于微分	51
3.2	偏导数的交换性与 Taylor 展开	58
3.3	反函数定理, 隐函数定理	61
4	(子)流形论	67
4.1	子流形的定义	67
4.2	子流形之间的光滑映射	72
4.3	Lagrange 乘子法	75
4.4	Hesse 矩阵与凸函数	78
5	多元积分与 Stokes 公式	83
5.1	第一型曲线和曲面积分	83
5.1.1	Stokes 公式	85

目录	2
5.2 曲线和曲面积分	89
5.3 第二型曲线和曲面积分	93
5.4 Brouwer 不动点定理	96
5.5 区域不变性定理	98
6 第一次作业	99
7 第二次作业	112
8 第三次作业	121
9 第四次作业	128
10 第五次作业	135
11 第六次作业	148
12 第七次作业	158
13 第八次作业	169
14 第九次作业	177
15 第十次作业	182
16 第十一次作业	187
17 第十二次作业	196

1 Lebesgue 积分理论, 续

1.1 Lebesgue 积分复习

YSW 讲义中 Lebesgue 积分定义的进路是采用使用**简单函数逼近**的方法, 得到一般**可测函数**的积分. 简单函数

$$f = \sum_{k=1}^N a_k \cdot \chi_{E_k}$$

可以定义自然的 Lebesgue 积分: f 把 E_k 区域内的每个点赋予权重 a_i , 那么每项的贡献应该是 $a_k \cdot m(E_k)$. 因此可定义

$$\int f := \sum_{k=1}^N a_k \cdot m(E_k). \quad (1)$$

可以直观地看出这样的定义和 f 的分拆的选取无关. 这样定义的积分具有

- 线性: $\int (af + bg) = a \int f + b \int g$.
- (逐点) 单调性: 如果 $f \leq g$ ($f(x) \leq g(x), \forall x$), 那么 $\int f \leq \int g$.

我们也可以类似定积分一样定义函数 f 在可测集 E 上的积分 $\int_E f$, 满足

$$\int_E f := \sum_{k=1}^N a_k \cdot m(E \cap E_k).$$

那么对两个不交区域 E, F , 积分有可加性

$$\bullet \int_E f + \int_F f = \int_{E \cup F} f.$$

对可测函数来说有下面的命题, 这是定义 Lebesgue 积分的基础.

Proposition 1.1.1. 任意非负可测函数都可以被一系列单调简单函数列从下逼近.

根据上面的命题, 我们立即想要用**简单函数的极限**来定义一般可测函数的 Lebesgue 积分. 但是这立即会遇到问题: 比如说给定两个简单函数列 $f_k \rightarrow f, g_k \rightarrow f$, 如果要良定义 $\int_E f$, 需要

$$\lim_{k \rightarrow \infty} \int_E f_k = \lim_{k \rightarrow \infty} \int_E g_k. \quad (2)$$

但是当 E 为无穷测度集, 比如说 $E = \mathbb{R}$ 时, 考虑

$$f_k(x) = \begin{cases} 1/k, & x \in [-k, k]; \\ 0, & \text{else.} \end{cases}$$

那么显然 $f_k \rightarrow 0$, 令 $g_k \equiv 0$. 但是 $\int f_k$ 始终为非零常数, 所以上面的事情不一定成立! 所以看来 E 在有限测度时才能按上述的方法良定义其 Lebesgue 积分. 另一个重

要的问题是当 $f_k \rightarrow f$ 时, $\lim_{k \rightarrow \infty} \int f_k$ 可能并不是收敛的. 接下来将定义**有限测度集上有界函数的积分**, 并证明以上两点均不会出现问题.

$$\int_E f := \lim_{k \rightarrow \infty} \int_E \varphi_k, \quad \varphi_k \rightarrow f. \text{ 其中 } m(E) < +\infty, |f| \leq M.$$

下证明右侧极限一定存在, 即 $\int_E \varphi_k$ 构成 Cauchy 列. 想法是利用 Egorov 定理, 对任意 $\varepsilon > 0$ 都可以找到测度小于 ε 的集合 A_ε 使得 φ_k 在 $E \setminus A_\varepsilon$ 上一致收敛至 f . 这蕴含在 $E \setminus A_\varepsilon$ 上存在 N 使得 $n, m > N$ 时 $|\varphi_n(x) - \varphi_m(x)| < \varepsilon$. (因为二者会一致地靠近 $f(x)$) 所以

$$\begin{aligned} \left| \int_E \varphi_n - \int_E \varphi_m \right| &= \left| \int_{E \setminus A_\varepsilon} (\varphi_n - \varphi_m) + \int_{A_\varepsilon} (\varphi_n - \varphi_m) \right| \\ &\leq \varepsilon \cdot m(E) + 2M \cdot \varepsilon. \end{aligned}$$

由 ε 的任意性可知 $\int_E \varphi_n$ 为 Cauchy 列, 最后一步同时用到了 E 的有限测度性与 f 的有界性. 特别地, 对于 (2) 的问题, 只需注意到当 $\varphi_k, \phi_k \rightarrow f$ 时, 序列

$$\varphi_1, \phi_1, \varphi_2, \phi_2, \dots$$

也收敛到 f , 所以根据上面的结论可知这个序列逐项取积分后极限也存在, 这蕴含了

$$\lim_{k \rightarrow \infty} \int_E \varphi_k = \lim_{k \rightarrow \infty} \int_E \phi_k.$$

一般地, 对定义在 \mathbb{R}^k 上的函数, 只要 f 的支集 $\text{supp } f = \{x : f(x) \neq 0\}$ 被包含在一个有限测度集 E 中时, 上面的定义可以完全照搬到 $\int f$ 上. 所以

$$L_0 = \{f : \text{supp } f \text{ 在有限测度集上, } f \text{ 有界}\}$$

上函数的积分 $\int f$ 可以被良定义.

还有另一个需要避免的事情: 在定义 $\int f$ 是我们没有特意排除掉 $\int f = \infty$ 的情形. 这是必要的, 因为“积分为无穷”和“不能定义积分”有很大区别. 那么在考虑线性性时可能对两个积分为 ∞ 的函数做差的情况, 此时值是没有定义的. 所以我们也也许可以先对**非负函数**定义其 Lebesgue 积分, 这样所有“无穷”都是正无穷, 那么所有的加法都是良定义的. 现在对非负可测函数 f , 定义

$$\int_E f = \sup \left\{ \int_E g : 0 \leq g \leq f, g \in L_0 \right\}.$$

我们来看看这个定义为什么是好的: 比如来证明一下积分的线性.

$$\int_E f + \int_E g = \int_E (f + g).$$

一方面, 设

$$\int_E f = \lim_{k \rightarrow \infty} \int_E f_k, \quad \int_E g = \lim_{k \rightarrow \infty} \int_E g_k.$$

则 $f_k \leq f, g_k \leq g$, 故 $f_k + g_k \leq f + g \Rightarrow \int_E (f_k + g_k) \leq \int_E (f + g)$. 所以

$$\int_E (f + g) \geq \lim_{k \rightarrow \infty} \int_E (f_k + g_k) = \int_E f + \int_E g.$$

另一方面, 设

$$\int_E (f + g) = \lim_{k \rightarrow \infty} \int_E \varphi_k, \quad \varphi_k \leq f + g.$$

那么应当把 φ_k 分拆为两项, 两项分别比 f 和 g 小. 取

$$f_k = \min\{f, \varphi_k\}, g_k = \varphi_k - f_k.$$

则如果 $f_k(x) = \varphi_k(x) \leq f(x)$, 则 $g_k(x) = 0 \leq g(x)$. 如果 $f_k(x) = f(x) \leq \varphi_k(x)$, 则 $g_k(x) = \varphi_k(x) - f(x) \leq g(x)$. 无论如何都有

$$f_k + g_k = \varphi_k, \quad 0 \leq f_k \leq f, \quad 0 \leq g_k \leq g.$$

而由 L_0 的定义, $0 \leq f_k \leq f \in L_0$ 迫使 $\text{supp } f_k \subset \text{supp } f$ 且 $|f_k(x)| \leq |f(x)|$. 因此也有 $f_k, g_k \in L_0$. 所以

$$\int_E \varphi_k = \int_E f_k + \int_E g_k \leq \int_E f + \int_E g.$$

所以另一边也成立, 这就证明了积分满足线性. 注意到如果我们只允许 g 为简单函数, 就没有那么好的性质, 但这一进路也可以直接定义 Lebesgue 积分 (见小黄书).

最后对一般的可测函数定义 Lebesgue 积分.

Definition 1.1.2. 定义可测函数 f 为 E 上的**可积函数**, 如果 $\int_E |f| < +\infty$.

那么对可积函数 f , 考虑 f 的正负分离

$$f = f^+ - f^-, \quad f^+(x) = \max(f(x), 0), \quad f^-(x) = -\min(f(x), 0).$$

那么 f^+ 与 f^- 都是非负可测函数并且由非负函数积分的线性可知

$$\int_E |f| = \int_E (f^+ + f^-) = \int_E f^+ + \int_E f^-.$$

所以 f^+, f^- 都是可积的. 那么自然可以定义

$$\int_E f = \int_E f^+ - \int_E f^-.$$

那么对一般可测函数, 可以证明仍然有积分的线性, 单调性, 三角不等式成立.

Remark. 这样的积分定义方法一同定义了高维情形的积分 (只需要 f 取值在 \mathbb{R} 上, 但定义域可以在 \mathbb{R}^n 上), 相比一维 Riemann 积分, Lebesgue 积分舍去了绝对积分和条件积分, 根据定义来看可以认为可积就默认是绝对可积.

至此定义好了 Lebesgue 积分. 然后来考察 Lebesgue 积分的一些性质, 从它和 Riemann 积分的联系开始.

Proposition 1.1.3. 如果 f 在区间 $[a, b]$ 上 Riemann 可积, 那么它也 Lebesgue 可积, 并且两积分相等.

Proof. 这是因为 Riemann 积分本质采取的是用阶梯函数逼近的方式. 我们暂时用 $\int^{\mathcal{R}}$ 来表示 Riemann 积分的值, 用 \int^L 来表示 Lebesgue 积分的值. 由 Riemann 积分的定义, 对可积函数 f 存在两列阶梯函数 ϕ_k 和 φ_k 分别从上和从下逼近 f . 由于区间加细只会导致积分值更加接近 f , 所以可不妨假设 $\phi_k \leq f$ 且 ϕ_k 为递增列, $\varphi_k \geq f$ 且 φ_k 为递减列, 并满足

$$\lim_{k \rightarrow \infty} \int_{[a,b]}^{\mathcal{R}} \phi_k = \int_{[a,b]}^{\mathcal{R}} f = \lim_{k \rightarrow \infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k.$$

那么可以逐点定义 ϕ_k 与 φ_k 收敛到的函数, 分别记为 \underline{f} 与 \bar{f} . 由于 \bar{f}, \underline{f} 均是定义在有限测度集上的有限测度函数, 因此其 Lebesgue 积分存在且有就是阶梯函数积分的极限. 另一方面, 对这些阶梯函数 Riemann 积分和 Lebesgue 积分的定义是一样的, 综上所述可以得到

$$\int_{[a,b]}^L \underline{f} = \lim_{k \rightarrow \infty} \int_{[a,b]}^L \phi_k = \lim_{k \rightarrow \infty} \int_{[a,b]}^{\mathcal{R}} \phi_k = \int_{[a,b]}^{\mathcal{R}} f.$$

同理

$$\int_{[a,b]}^L \bar{f} = \int_{[a,b]}^{\mathcal{R}} f.$$

最后, 由于 $\underline{f} \leq f \leq \bar{f}$, 所以首先 f 是 Lebesgue 可积的, 其次根据单调性直接得到应有

$$\int_{[a,b]}^L f = \int_{[a,b]}^{\mathcal{R}} f.$$



Remark. 所以在做题中有这么一种方法: 把 Riemann 积分替换成 Lebesgue 积分, 并利用如控制收敛定理等较强的命题求出积分值 (而不是不断去验证一致收敛), 并证明最终表达式仍然是 Riemann 可积的, 也可以得到正确的积分值. 其本质是因为 Lebesgue 是更加“完备化”的, 就像函数方程中 $\mathbb{R}^+ \rightarrow \mathbb{R}^+$ 的函数被允许代入了零值, 借助这些本来无定义的值能得到更好的结论是理所应当的事情.

接下来就围绕控制收敛定理一一证明这些有用的结论.

Theorem 1.1.4 (有限测度一致有界收敛定理). 对有限测度集 E 及 E 上的可测函数列 $f_n \rightarrow f$, a.e. 如果

- f_n 几乎处处一致有界 ($\exists M$ s.t. $|f_n| < M$, a.e.),

则

$$\lim_{n \rightarrow \infty} \int_E |f_n - f| = 0, \quad \lim_{n \rightarrow \infty} \int_E f_n = \int_E f.$$

Proof. 和前面证明有限测度集上有界函数积分定义的合理性一样, 还是利用 **Egorov** 定理. 对任意 $\epsilon > 0$, 存在闭集 $A_\epsilon \subset E$ 使得 f_n 在 A_ϵ 上一致收敛到 f , 并且 $m(E \setminus A_\epsilon) < \epsilon$. 因而对于足够大的 n , 我们有

$$|f_n - f| < \epsilon, \quad \forall x \in A_\epsilon.$$

这样我们可以估计

$$\begin{aligned} \int_E |f_n - f| &\leq \int_{A_\epsilon} |f_n - f| + \int_{E \setminus A_\epsilon} |f_n - f| \\ &\leq \epsilon m(A_\epsilon) + 2M m(E \setminus A_\epsilon) \\ &\leq m(E)\epsilon + 2M\epsilon. \end{aligned}$$

由 ϵ 的任意性立得结论. 这个命题完全就是 L_0 中积分定义的推广. $\left\langle \frac{1}{\epsilon} \right\rangle$

对非负函数有如下的引理, 适用范围很广 (看名字就可知这是可测集上的 Fatou 引理的精神延续):

Lemma 1.1.5 (Fatou 引理). 对非负可测函数列 f_n , 总有

$$\int_E \liminf_{n \rightarrow \infty} f_n \leq \liminf_{n \rightarrow \infty} \int_E f_n.$$

Proof. 根据非负可测函数积分的定义,

$$\text{LHS} = \sup \left\{ \int_E g : 0 \leq g \leq \liminf_{n \rightarrow \infty} f_n, g \in L_0 \right\}.$$

所以只需对每个 $g \leq \liminf_{n \rightarrow \infty} f_n$ 证明 $\int_E g \leq \text{RHS}$. 假设结论不成立, 那么一定存在子列 f_{n_k} 及 $\varepsilon > 0$ 使得

$$\int_E g - \varepsilon \geq \int_E f_{n_k}, \quad \forall k \in \mathbb{N}.$$

现在可以利用上 g 的条件: 取 $g_k = \min\{g, f_{n_k}\}$. 则由 $g \in L_0$ 可知 $g_k \in L_0$. 此外

$$\liminf_{k \rightarrow \infty} g_k(x) = \liminf_{k \rightarrow \infty} \min\{f_k(x), g(x)\} \geq g(x).$$

再结合 $g_k(x) \leq g(x)$ 可知 $\lim_{k \rightarrow \infty} g_k(x) = g(x)$. 由于 $g_k(x)$ 有一致上界 (因为 $|g|$ 有上界) 且 $\text{supp } g_k \subset \text{supp } g$ 有限测度, 因此利用上题结论可知

$$\int_E g = \lim_{k \rightarrow \infty} \int_E g_k.$$

所以对足够大的 k 总有 $\int_E g_k > \int_E g - \varepsilon/2$, 因此存在 n_k 使得 $\int_E g_{n_k} - \varepsilon/2 \geq \int_E f_{n_k}$. 这与 $g_{n_k} \leq f_{n_k}$ 矛盾! 因此原命题成立. $\left\langle \frac{1}{\epsilon} \right\rangle$

接下来增强条件, 让 Fatou 引理可以推出积分与极限的交换性.

Theorem 1.1.6 (Beppo-Levi). 对单调上升的非负可测函数列 f_n , 若有 $f_n \rightarrow f$, 则

$$\lim_{n \rightarrow \infty} \int_E f_n = \int_E f.$$

Proof. 这是因为 f_n 单调增与 $f_n \rightarrow f$ 足以推出 $f_n \leq f$, 所以 $\int_E f_n \leq \int_E f$. 取极限后显然可以得到

$$\lim_{n \rightarrow \infty} \int_E f_n \leq \int_E f.$$

另一方面, f_n 和 $\int_E f_n$ 都是单调增列, 此时下极限和极限等同, 所以由 Fatou 引理,

$$\int_E f = \int_E \lim_{n \rightarrow \infty} f_n \leq \lim_{n \rightarrow \infty} \int_E f_n.$$

至此结论成立. \square

这个定理有很好的推论: 因为对一系列非负函数 f_n , 总有 $\sum_{n=1}^m f_n$ 是单调增列.

Corollary 1.1.7 (非负级数的逐项积分). 对一系列非负函数 $f_k(x)$, 若 $\sum_{k=1}^{\infty} f_k(x)$ 无穷级数收敛, 则

$$\int_E \sum_{k=1}^{\infty} f_k(x) = \sum_{k=1}^{\infty} \int_E f_k(x).$$

Proof. 证明是上面问题的直接推论: 取 $S_n(x) = \sum_{k=1}^n f_k(x)$ 并利用上题结论即可. \square

接下来的命题固定 f , 考察集合的测度足够小的情况.

Proposition 1.1.8 (积分的绝对连续性). 假设 f 可积. 若 $\lim_{k \rightarrow \infty} m(E_k) = 0$, 则

$$\lim_{k \rightarrow \infty} \int_{E_k} |f| = 0.$$

Proof. 不妨设 f 非负, 否则用 $|f|$ 来代替 f . 设 $f^N = f \cdot \chi_{\{|x| \leq N, |f(x)| \leq N\}}$. 则 f^N 单调增且收敛到 f , 因此对任意 $\varepsilon > 0$, 对足够大的 N 有

$$\int (f - f^N) < \varepsilon \Rightarrow \int_{E_k} (f - f^N) < \varepsilon, \forall k \in \mathbb{N}.$$

再取足够大的 k 使得 $m(E_k) < \varepsilon/N$, 则

$$\int_{E_k} f^N \leq m(E_k) \cdot N \leq \varepsilon.$$

综上得到

$$\int_{E_k} f \leq 2\varepsilon.$$

对一切足够大的 k 成立, 故原命题成立. \square

Proposition 1.1.9. 对任意可测函数 f 与零测集 E , 均有

$$\int_E f = 0.$$

Proof. 这时因为零测集 E 上所有简单函数的积分均为零. 那么 L_0 中函数的积分是一些简单函数积分的极限, 所以也是零, 进而所有非负函数的积分都是零, 故所有可测函数积分均为零. \square

Corollary 1.1.10. 若可测函数 f, g 几乎处处相等, 则

$$\int_E f = \int_E g.$$

Proof. 设 $E_1 = \{x : f(x) \neq g(x)\}$, 则 $m(E_1) = 0$. 因此

$$\int_E f(x) = \int_{E \setminus E_1} f(x) = \int_{E \setminus E_1} g(x) = \int_E g(x).$$

\square

最后来看 Lebesgue 控制收敛定理.

Theorem 1.1.11 (控制收敛定理). 假设可测函数列 f_n 几乎处处收敛到 f , 并且存在可积函数 g 使得

$$|f_n| \leq g, a.e.$$

那么就有

$$\lim_{n \rightarrow \infty} \int_E |f_n - f| = 0, \quad \lim_{n \rightarrow \infty} \int_E f_n = \int_E f.$$

Proof. 由题意, $\lim_{n \rightarrow \infty} |f_n(x)| \leq g(x)$, 所以 $|f| \leq g, a.e.$ 因此有

$$|f - f_n| \leq 2g, \forall n \in N.$$

故可以对非负函数列 $2g - |f - f_n|$ 利用 Fatou 引理 1.1.5,

$$\int_E \lim_{n \rightarrow \infty} (2g - |f - f_n|) \leq \liminf_{n \rightarrow \infty} \int_E (2g - |f - f_n|).$$

LHS 是几乎处处收敛的, 又因为几乎处处相等的函数积分值一致, 所有可以把下极限替换为极限. RHS 利用线性展开, 得到

$$\int_E 2g - \int_E \lim_{n \rightarrow \infty} |f - f_n| \leq \int_E 2g - \overline{\lim}_{n \rightarrow \infty} \int_E |f - f_n|.$$

又因为 $\lim_{n \rightarrow \infty} |f - f_n| = 0, a.e.$ 所以上式可化简为

$$\overline{\lim}_{n \rightarrow \infty} \int_E |f - f_n| \leq 0.$$

但显然 LHS 每项均非负, 所以

$$\lim_{n \rightarrow \infty} \int_E |f - f_n| = 0.$$

因此

$$\left| \int_E f_n - \int_E f \right| \leq \int_E |f - f_n|.$$

这导致前者极限为零, 证毕. \square

1.2 可积函数空间 L^1

Definition 1.2.1. 对于在可测集 E 上定义的可测函数可积当且仅当

$$\int_E |f| < +\infty.$$

定义

$$L^1(E) = \{f : f \text{ 可积}\}.$$

我们来看 L^1 的性质:

- 线性空间: 若 $f, g \in L^1(E)$, 则对任意 $a, b \in \mathbb{R}$, $af + bg \in L^1(E)$.
- 赋范空间: $\|f\| = \int_E |f|$. 范数定义在函数的 a.e. 等价类上.
范数满足其应有的线性与三角不等式.
- 范数可以诱导距离 $d(f, g) = \|f - g\|$, 同时也诱导新的收敛规则:

$$f_k \xrightarrow{L^1} f \Leftrightarrow \lim_{k \rightarrow \infty} \|f_k - f\| = 0.$$

- 对赋范线性空间下一步是讨论完备性 (即可以包含所有序列极限). 非常好的一点是, $L^1(E)$ 本身就是完备的!

Theorem 1.2.2 (Riesz-Fisher). $L^1(E)$ 是完备的赋范线性空间.

Proof. 考虑 Cauchy 列 $f_k \in L^1$, 满足

$$\forall \varepsilon > 0, \exists N \text{ s.t. } \|f_k - f_l\| < \varepsilon, \forall k, l > N.$$

问题是能否找到 $f \in L^1$ 使得 $\lim_{k \rightarrow \infty} \|f_k - f\| = 0$.


想法是很经典的: 找到 f_k 的一个指数级别收敛的子列, 证明这个子列存在极限, 然后再证明它也是原序列的极限.

首先取子列 f_{k_l} 使得 $\|f_{k_l} - f_{k_{l+1}}\| < 2^{-l}$. 考虑函数

$$g = |f_{k_1}| + |f_{k_2} - f_{k_1}| + \cdots + |f_{k_{l+1}} - f_{k_l}| + \cdots$$

那么 $\|g\|$ 有界, 所以 g 可积, 因此 $|g| < +\infty$, a.e. 所以

$$f = f_{k_1} + (f_{k_2} - f_{k_1}) + \cdots + (f_{k_{l+1}} - f_{k_l}) + \cdots$$

也几乎处处有限, 这样就可以计算 $f - f_{k_i}$, 并得到 $\|f - f_{k_i}\| \rightarrow 0$. 这已经足以推出 $\|f - f_k\| \rightarrow 0$. 

Proposition 1.2.3. 以下函数空间在 $L^1(E)$ 上稠密:

1. 简单函数;

2. 阶梯函数;
3. 具有紧支集的连续函数空间 C_c ;
4. 光滑的有紧支集函数 C_c^∞ .

Proof. 先证明简单函数的情形. 首先将 $f \in L^1$ 正负分离为 $f = f_+ - f_-$. 然后可以取出一列单调升简单函数来从下逼近 f_+ 与 f_- . 然而, 得到的收敛模式为逐点收敛, 逐点收敛不能直接推出 L^1 收敛. 但是我们有直接与这里结论适配的 Beppo-Levi 定理 1.1.6, 所以这个结论等价于 L^1 收敛. 最后再把得到的正负函数列相减就能推出结论.

对于阶梯函数, 不能直接使用“阶梯函数可 a.e. 逼近可测函数”的结论, 因为 a.e. 收敛推不出 L^1 收敛. 但是可以利用上一问的结论, 即用阶梯函数逼近简单函数即可. 只需要利用可测集的定义, 使用方块来逼近可测集就能做到.

最后用光滑函数来逼近阶梯函数. 一维情况下, 可以用光滑函数 (可使用 $e^{-1/x}$ 的进路或者多项式的进路) 逼近分段函数. 想要得到高维的阶梯函数, 只需按分量指定光滑函数并把它们作积. 然后再使用 (2) 的结论即可. \square

由于可积函数可以用上述很好的函数逼近, 所以在证明之后的命题时, 经常只对好的这些函数进行证明, 然后利用稠密性推至一般的函数成立. 下面讨论可积函数在变换中的不变性.

- 平移不变性: 考虑 $\tau_h \in \text{Alt}(\mathbb{R}^d)$, $(\tau_h f)(x) = f(x+h)$. 由于 $m(I) = m(I+h)$, 因此 $\|\tau_h f\| = \|f\|$.
- 伸缩不变性: 考虑 $D_\delta \in \text{Hom}(\mathbb{R}^d)$, $(D_\delta f)(x) = f(\delta x)$. 则 $\|D_\delta f\| = \delta^{-k} \|f\|$.
- 反射不变性. 目前我们不讨论积分的定向, 因此对反射变换 $\bar{\tau}$, 有 $\|\bar{\tau} f\| = \|f\|$.

Proposition 1.2.4 (平移伸缩变换的连续性). 对 $f \in L^1(\mathbb{R}^d)$, 有

$$\lim_{h \rightarrow 0} \|\tau_h f - f\| = 0; \quad \lim_{\delta \rightarrow 1} \|D_\delta f - f\| = 0.$$

Proof. 固定 $\forall \varepsilon > 0$, 存在阶梯函数 g 使得 $\|g - f\| < \varepsilon/3$.

设 $g = \sum_{k=1}^N a_k \chi_{I_k}$, 且 I_k 互不相交. 则 $\|\tau_h f - \tau_h g\| < \varepsilon/3$. 另一方面, 还有

$$\begin{aligned} \|\tau_h g - g\| &= \left\| \sum_{k=1}^N (\chi_{I_k+h} - \chi_{I_k}) a_k \right\| \\ &\leq \left\| \sum_{k=1}^n h \cdot c(I_k) a_k \right\| \leq Ch. \end{aligned}$$

其中 C 是仅和 g 有关的常数. 因此当 h 充分小时, 总有 $\|\tau_h g - g\| \leq \varepsilon/3$. 此时

$$\begin{aligned} \|\tau_h f - f\| &\leq \|f - g\| + \|\tau_h f - \tau_h g\| + \|\tau_h g - g\| \\ &\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

故 $\lim_{h \rightarrow 0} \|\tau_h f - f\| = 0$. \square

1.3 Fubini 引理

对高维的积分, 不存在所谓条件收敛的概念, 这是因为比如在平面上, 不同区间内本身就可以进行“重排”后再相加, 而为了让这些值良定义, 本身就需要绝对收敛.

Theorem 1.3.1 (Fubini 定理). 设 $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, $x \in \mathbb{R}^{d_1}, y \in \mathbb{R}^{d_2}, d = d_1 + d_2$. 现在有函数 $f(x, y)$ 定义在 \mathbb{R}^d 上满足 $f \in L^1(\mathbb{R}^d)$, 是否有

$$\int_{\mathbb{R}^d} f(x, y) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y) \right) = \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x, y) \right)?$$

结论如下:

1. 对几乎所有的 x , 作为 y 的函数 $f(x, y)$ 在 \mathbb{R}^{d_2} 上可积;
2. 对 y 的积分 $\int_{\mathbb{R}^{d_2}} f(x, y) dy$ 定义了 \mathbb{R}^{d_1} 上关于 x 的可积函数;
3. 上述的积分等式关系成立.

定义 \mathcal{F} 是所有满足 Fubini 定理的可积函数组成的空间.

Proposition 1.3.2. \mathcal{F} 是线性空间, 并且非负单调序列的极限如果为可积函数, 那么也属于 \mathcal{F} .

Proof. • 对于线性部分, 只需证 $f + g \in \mathcal{F}$, 数乘部分是显然的. 由条件 1 可知存在 \mathbb{R}^{d_1} 上零测集 X_f, X_g 使得对 $x \notin X_f \cup X_g$, 总有 $f(x, y)$ 与 $g(x, y)$ 作为 y 的函数是 \mathbb{R}^{d_2} 上可积的. 所以 $(f + g)(x, y)$ 也是几乎处处在 \mathbb{R}^{d_2} 上可积的 (这是因为 $L^1(\mathbb{R}^{d_2})$ 是线性空间, 并且 $X_f \cup X_g$ 仍是零测集. 所以条件 1 成立.

又因为 $\int_{\mathbb{R}^{d_1}} f(x, y) dy$ 和 $\int_{\mathbb{R}^{d_1}} g(x, y) dy$ 作为 x 的函数在 \mathbb{R}^{d_1} 上可积. 所以由积分的线性性,

$$\begin{aligned} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} (f(x, y) + g(x, y)) dy &= \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y) dy + \int_{\mathbb{R}^{d_2}} g(x, y) dy \right) dx \\ &= \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} f(x, y) + \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} g(x, y) \\ &= \int_{\mathbb{R}^d} f(x, y) + \int_{\mathbb{R}^d} g(x, y) = \int_{\mathbb{R}^d} (f + g)(x, y). \end{aligned}$$

这就证明了条件 2 和条件 3.

- 再证明极限封闭. 方法与之前类似, 对单调序列 f_n , 由条件 1 取出 \mathbb{R}^{d_1} 中的一列零测集 X_n 使得当 $x \notin X_n$ 时 $f_n(x, y)$ 在 \mathbb{R}^{d_2} 上可积. 由于零测集的可数并仍是零测集, 所以

$$f_n(x, y) \text{ 关于 } y \text{ 在 } \mathbb{R}^{d_2} \text{ 上可积, } \forall y \in \mathbb{N}, \quad x \notin \bigcup_{n=1}^{\infty} X_n.$$

其中 $\bigcup_{n=1}^{\infty} X_n$ 也是零测集. 但这还不能完全推出 $\lim f_n$ 关于 y 可积, 这是因为取极限可能让积分趋向于无穷. 事实上, 当 f_n 单调增时可以使用 Beppo-Levi 定理 1.1.6, f_n 单调减时可以使用控制收敛定理 1.1.11 得到

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d_2}} f_n(x, y) dy = \int_{\mathbb{R}^{d_2}} \lim_{n \rightarrow \infty} f_n(x, y) dy, \quad x \notin \bigcup_{n=1}^{\infty} X_n.$$

但是这并不能保证两式取值非 $+\infty$. 但我们还有没用上的条件: $\lim_{n \rightarrow \infty} f_n$ 在 \mathbb{R}^d 上可积. 首先对 f_n 利用条件 3 可知

$$\int_{\mathbb{R}^d} f_n = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f_n(x, y) dy \right).$$

又因为 $g_n(x) := \int_{\mathbb{R}^{d_2}} f_n(x, y) dy$ 也是非负单调函数列, 所以

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n = \int_{\mathbb{R}^{d_1}} \left(\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d_2}} f_n(x, y) dy \right) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \lim_{n \rightarrow \infty} f_n(x, y) dy \right). \quad (1)$$

最后, 由于 f_n 在 \mathbb{R}^d 上也是非负单调列, 所以还有

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n = \int_{\mathbb{R}^d} \lim_{n \rightarrow \infty} f_n < +\infty.$$

后者有限是条件. 这导致 (1) 也为有限值, 即

$$g(x) := \int_{\mathbb{R}^{d_2}} \lim_{n \rightarrow \infty} f_n(x, y) dy$$

为可积函数. 实际上这已经证明完了条件 2 和条件 3. 最后回过头去, 上式表明 $g(x)$ 几乎处处有限, 设 $g(x)$ 在 C 上取值为 $+\infty$, 则 $m(C) = 0$. 这意味着

$$g(x) \text{ 有定义且有限, } x \notin \bigcup_{n=1}^{\infty} X_n \cup C.$$

后者为零测集, 所以 $\lim_{n \rightarrow \infty} f_n(x, y)$ 几乎处处可积, 这就证完了条件 1. \square

Proposition 1.3.3. \mathcal{F} 包含以下集合的示性函数:

- 方体;
- 有限测度开集;
- 有限测度 G_δ 集;
- 零测集;
- 有限测度可测集.

Proof. (1) 对方体 $I_x \times I_y = I$, 有 $\chi_I = \chi_{I_x} \cdot \chi_{I_y}$. 对任意 x , 均有 $\chi_I(x, y)$ 在 \mathbb{R}^{d_2} 上可积, 并且可以算出

$$\int_{\mathbb{R}^{d_2}} \chi_I(x, y) dy = |I_y| \cdot \chi_{I_x}$$

在 \mathbb{R}^{d_1} 上可积, 所以

$$\int_{\mathbb{R}^d} \chi_I = |I| = |I_y| \cdot |I_x| = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \chi_I \, dy \, dx.$$

(2) 再看开集 O , 开集可以写成方体的并 $O = \bigcup_{k=1}^{\infty} I_k$. 设 $E_k = \bigcup_{l=1}^k I_l$, 则

$$\chi_O = \lim_{k \rightarrow \infty} \chi_{E_k}.$$

因为 χ_{E_k} 为单调增函数列, 且 χ_O 为可积函数 (这需要 $m(O) < +\infty$), 可以利用命题 1.3.2. 所以 $\chi_O \in \mathcal{F}$.

(3) 对 G_δ 集 O , 同样存在有限测度开集列 I_k 满足 $O = \bigcap_{k=1}^{\infty} I_k$. 设 $E_k = \bigcap_{l=1}^k I_l$. 则

$$\chi_O = \lim_{k \rightarrow \infty} \chi_{E_k}.$$

因为 χ_{E_k} 为单调减函数列, 且 χ_O 为可积函数 (这需要 $m(O) < +\infty$), 可以利用命题 1.3.2. 所以 $\chi_O \in \mathcal{F}$.

(4) 零测集才是真正有困难的地方: 因为它不能利用“逼近”的手段得到, 并且其形态可能非常奇怪.

首先我们需要找一个“梯子”: 由于每个集合都能写成 G_δ 集去掉一个零测集, 那么零测集就会成为一个零测 G_δ 集的子集. 因此问题转化为证明如果已知一个零测集 $E \subset \mathcal{F}$. 去证明 E 的子集 $A \in \mathcal{F}$. 首先, 有

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_E(x, y) \, dy \right) \, dx = \int_{\mathbb{R}^d} \chi_E = 0; \quad \int_{\mathbb{R}^d} \chi_A = 0.$$

由于积分为零的非负函数几乎处处为零 (见习题), 因此可推出

$$g(x) = \int_{\mathbb{R}^{d_2}} \chi_E(x, y) \, dy = 0, \text{ a.e.}$$

又因为对几乎所有的 x , $E(x, y)$ 关于 y 可积, 故

$$\chi_E(x, y) = 0, \text{ 对几乎处处的 } x \text{ 和 } y.$$

利用 $A \subset E$ 可知上面命题推出

$$\chi_A(x, y) = 0, \text{ 对几乎处处的 } x \text{ 和 } y.$$

因此对几乎处处的 x 有 $\chi_A(x, y)$ 可积且积分为零, 这导致 $\int_{\mathbb{R}^{d_2}} \chi_A(x, y) \, dy$ 关于 x 几乎处处为零, 所以它也是可积的, 并且积分为零, 这就证明了零测集满足要求.

Remark. 这里的一个主要问题是高维空间的零测集在低维空间上的投影不一定是零测集. 比如说, 一条斜直线在 \mathbb{R}^2 中零测, 但是在两个方向上投影均是 \mathbb{R} . 所以低维空间的积分可能非零. 然而, 条件告诉我们积分是几乎处处为零的, 这是积分为零的充要条件, 所以这已经足够了.

(5) 因为有限测度可测集都可以写成 G_δ 集去掉一个零测集, 由于 \mathcal{F} 为线性空间, 所以其示性函数为二者之差, 因此在 \mathcal{F} 中. $\left\{ \begin{smallmatrix} u \\ i \\ x \end{smallmatrix} \right\}$

现在说明上面的条件已经足以推出 $\mathcal{F} = L^1(\mathbb{R}^d)$.

Proof for 1.3.1. • 由线性可知所有积分有限的简单函数都落在 \mathcal{F} 中.

- 由 Beppo-Levi 定理 1.1.6 和非负可测函数的简单函数收敛定理 1.1.1, 对任何非负可测的函数, 可以找到一列单调简单函数从下逼近它. 若其可积, 则这些简单函数也都可积. 因此由单调列极限封闭性可知所有非负可积函数均在 \mathcal{F} 中.

- 对任何可积函数 f , 将其正负分离为 $f = f^+ - f^-$. 根据线性立即有 $f \in \mathcal{F}$. $\left\{ \begin{smallmatrix} u \\ i \\ x \end{smallmatrix} \right\}$

Theorem 1.3.4 (Tonelli 定理). 对 \mathbb{R}^d 上的非负可测函数 $f(x, y)$,

1. 对几乎所有 x , $f(x, y)$ 作为 y 的函数在 \mathbb{R}^{d_2} 上可测;
2. 作为 x 的函数 $\int_{\mathbb{R}^{d_2}} f(x, y) dy$ 是 \mathbb{R}^{d_1} 上的可测函数;
3. 积分等式

$$\int_{\mathbb{R}^d} f(x, y) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y) dy \right) dx$$

成立.

Proof. 这个命题可以看作是 Fubini 定理的推论. 对于非负可测函数, 唯一的问题在于处理积分为无穷的情况, 这可以依靠上升函数列逼近来解决.

记 $f_n = f \cdot \chi_{|x|+|y| \leq n, f(x, y) \leq n}$. 则 f_n 一定可积. 由 Fubini 定理 1.3.1, 对几乎所有的 x 有 $f_n(x, y)$ 关于 y 可积. 但另一方面又有 f_n 逐点收敛到 f , 所以对这样的 x 有 $f(x, y)$ 可测. 同样地, 对所有 n 均有 $\int_{\mathbb{R}^{d_2}} f_n(x, y) dy$ 为可积函数, 因此也是可测函数. 又由 Beppo-Levi 定理 1.1.6 有

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d_2}} f_n(x, y) dy = \int_{\mathbb{R}^{d_2}} f(x, y) dy.$$

所以 $g_n(x)$ 逐点收敛到 $g(x)$ 且 g_n 可测, 因此 g 也可测. 最后一个积分等式直接对 Fubini 定理的条件 3 两边施加极限即可. $\left\{ \begin{smallmatrix} u \\ i \\ x \end{smallmatrix} \right\}$

Proposition 1.3.5. 设 E 是 \mathbb{R}^d 上的可测集. 对几乎所有 x , 集合 $E^x = \{y : (x, y) \in E\}$ 是 \mathbb{R}^{d_2} 上的可测集. 并且作为 x 的函数 $m(E^x)$ 是 \mathbb{R}^{d_1} 上的可测函数, 满足

$$m(E) = \int_{\mathbb{R}^{d_1}} m(E^x) dx.$$

“高维信息可以得到所有低维信息”，根据 Tonelli 定理 1.3.4 这个命题可以直接被推出。但是比较有趣的是它的逆命题是否成立，即如果所有 E^x, E^y 都是可测集，那么 E 是否为可测集？这是一个很难的问题，但是如果我们承认连续统假设（ \aleph_0 和 c 之间没有其它基数），就可以构造出这样一个不可测集 $E \subset \mathbb{R}^2$ 。

- 首先构造 \mathbb{R} 上的全序 \prec 使得对任意固定的 y , $\{x : x \prec y\}$ 是可数集。
- 由良序定理 (in ZFC), 存在 \mathbb{R} 上的良序关系 \prec , 考虑所有不满足要求的集合 $X = \{y : \text{card}\{x : x \prec y\} > \aleph_0\}$. 如果 X 非空, X 中存在最小元 $\bar{y} \in X$. 由于 $\{x : x \prec \bar{y}\} \subset \mathbb{R}$ 不可数, 因此由连续统假设它的基只能是 c . 这样就可以构造一个从 $\{x : x \prec \bar{y}\}$ 到 \mathbb{R} 的双射, 因此可以让该子集继承 X 的全序并赋到 \mathbb{R} 上, 记这个全序为 \prec^1 .
- 现在如果存在 $x_1 \in \mathbb{R}$ 使得 $\{x : x \prec^1 x_1\}$ 是不可数集, 那么存在 $y_1 \in \prec \bar{y}$ 使得 $\{x : x \prec y_1\}$ 是不可数集, 因此 $y_1 \in X$, 但这与 \bar{y} 是 X 的极小元矛盾! 因此上面的命题成立.
- 接下来定义 $E = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, x \prec y\}$, 那么固定 y 可得 E^y 是可数集, 固定 x 可得 E^x 是全集去掉一个可数集, 因此它们都是可测的, 且 $m(E^y) = 0, m(E^x) = 1$.
- 假设 E 可测, 那么它满足 Fubini 定理, 那么应该有

$$\int_0^1 m(E^x) dx = \int_0^1 m(E^y) dy.$$

但左边结果为 1, 右边结果为 0, 矛盾! 所以 E 为不可测集.

抽象测度论上的乘积测度

考虑两个测度空间 $(X_1, \mathcal{A}_1, \mu_1)$ 和 $(X_2, \mathcal{A}_2, \mu_2)$, 实际上我们可以对 $X_1 \times X_2$ 定义乘积测度. 首先的问题是 $\mathcal{A}_1 \times \mathcal{A}_2$ 不一定是 σ -代数, 我们可以模仿定义 Lebesgue 测度的方式, 使用“方块——外测度——完备化”的方法来构造. 首先方块的定义很明显, 就是 $\mathcal{A}_1 \times \mathcal{A}_2$ 中的集合, 它们的测度也可以被显然地定义: 对 $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$, 可有 $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$. 然后可以用这些“方块”定义 $X_1 \times X_2$ 中所有集合的外测度, 最后再用 Caratheodry 条件进行完备化, 得到所有可测集, 这便得到了乘积测度 μ .

上面的操作告诉我们 Lebesgue 测度本身就可以看作一维测度的积（一维方块是区间），而乘积的对称性带来了 Fubini 定理. 并且，低维的可测集（而不只是方块）可以用来给出它们的积的测度，这引出了下面的结论。（需要指出的是，能用低维推出的高维命题是很稀有的）

Proposition 1.3.6. 1. 如果 E_1, E_2 分别是 $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}$ 中的可测集, $E = E_1 \times E_2$ 是 \mathbb{R}^d 中的可测集, 且有 $m(E) = m(E_1)m(E_2)$.

2. 反过来, 若 $E = E_1 \times E_2$ 是 \mathbb{R}^d 中的可测集, 那么 E_1, E_2 是 $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}$ 中的可测集, 并且 $m(E) = m(E_1)m(E_2)$.
3. 如果 E_1, E_2 中存在一个零测集 (另一个集合不一定可测), 那么此时 E 为零测集.
4. 若 $f(x)$ 是 \mathbb{R}^{d_1} 上可测函数, 那么 $g(x, y) = f(x)$ 是 \mathbb{R}^d 上的可测函数.

Proof. 首先证明 (1), 不妨设 E_1 和 E_2 都是测度有限的可测集, 那么可以考虑用等测包来证明结论: 取出 G_δ 集 F_1, F_2 及零测集 Z_1, Z_2 使得

$$E_1 = F_1 \setminus Z_1, \quad E_2 = F_2 \setminus Z_2.$$

而 $F_1 \times F_2$ 也是 G_δ 集, 所以只用证明 $F_1 \times F_2 \setminus E_1 \times E_2$ 是零测集. 可以通过证明一不等式来得到结论.

Lemma 1.3.7. 对任意集合 E_1, E_2 , 有

$$m^*(E) \leq m^*(E_1) \cdot m^*(E_2), \quad E = E_1 \times E_2.$$

Proof. 考虑覆盖 E_1 的一列矩体 I_1, I_2, \dots 及覆盖 E_2 的一列矩体 J_1, J_2, \dots , 那么可知

$$E_1 \times E_2 \subset \bigcup_{k=1}^{\infty} I_k \times \bigcup_{l=1}^{\infty} J_l = \bigcup_{k,l} I_k \times J_l.$$

最右边仍然是可数个方体的并, 而对方体有 $m(I_k \times J_l) = m(I_k) \cdot m(J_l)$. 那么如果 E_1, E_2 中存在一个外测度为无穷, 那么 $\sum_{k,l} m(I_k)m(J_l) = \infty$, 否则这是一个绝对收敛列, 交换求和顺序即得 $\sum_{k,l} m(I_k)m(J_l) = (\sum_k m(I_k)) \cdot (\sum_l m(J_l))$. 而又有 $\sum_{k,l} m(I_k)m(J_l) \geq m^*(E_1 \times E_2)$, 因此取两者足够接近 $m(E_i)$ 可知 $m^*(E) \leq m^*(E_1) \cdot m^*(E_2)$. □

特别地, 当 E_1, E_2 中有零测集时, 它与任何集合的积都为零测集. 又因为 $(F_1 \times F_2) \setminus (E_1 \times E_2) \subset (Z_1 \times F_2) \cup (Z_2 \times F_1)$, 所以显然这是零测集.

剩下的命题是容易的, 对函数 $\chi_{E_1 \times E_2}$ 利用 Fubini 定理在可测集上的应用 1.3.5, 可以直接证出结论. 最后, (4) 中 $\{(x, y) : g(x, y) > t\} = \{(x, y) : f(x) > t\} = \{x : f(x) > t\} \times \mathbb{R}^{d_2}$ 也是可测集. □

Proposition 1.3.8 (集合测度在线性变换下的转换). 给定线性变换 $L : \mathbb{R}^d \rightarrow \mathbb{R}^d$ 以及 \mathbb{R}^d 上的可测集 E , 则 $L(E)$ 可测, 且有 $m(L(E)) = |\det(L)| \cdot m(E)$.

特别地, 可测集在旋转变换下测度不变. 若 $L \in O(n)$, 则 $m(L(E)) = m(E)$.

Proof. 只需对方体证明结论. 这是因为所有开集的都可以写成一些方体的无交并, 所以对开集命题均成立, 所以对闭集也都成立. 对一般的可测集, 它可以

被开集和闭集逼近, 即对任意 $\varepsilon > 0$, 存在闭集 F 和开集 G 使得 $F \subset E \subset G$ 且 $m(G \setminus F) < \varepsilon$. 因此

$$|\det(L)|(m(E) - \varepsilon) \leq |\det(L)|m(F), \quad |\det(L)|(m(E) + \varepsilon) \geq |\det(L)|m(G).$$

又因为 $m(L(F)) \leq m(L(E)) \leq m(L(G))$, 所以 $m(L(E))$ 与 $m(E)$ 之间差距很小, 令 $\varepsilon \rightarrow 0$ 即可知 $m(L(E)) = |\det L| m(E)$.

不妨假设 I 为单位方体, $0 \leq x_i \leq 1$, 则可设 L 对应的矩阵为 $L = (L_{ij})$.

- 如果 (L_{ij}) 是对角矩阵, 设 $(L_{ij}) = \text{diag}(\lambda_1, \dots, \lambda_n)$, 那么方体恰好扩大为边长分别为 $\lambda_1, \dots, \lambda_n$ 的矩体, 其测度显然为 $\prod_{i=1}^n \lambda_i$.
- 如果 (L_{ij}) 是对角线全为 1 的上三角矩阵, 则

$$L(E) = \{(a_1 + L_{12}a_2 + \dots + L_{1n}a_n, a_2 + \dots + L_{2n}a_n, \dots, L_{nn}a_n) : 0 \leq a_i \leq 1\}.$$

因此可考虑其特征函数 $\chi_{L(E)}$, 利用 Fubini 定理 1.3.1, 固定 (x_2, \dots, x_n) , 考虑

$$\int_{\mathbb{R}} \chi_{L(E)} dx.$$

那么当且仅当 (x_2, \dots, x_n) 落在 $L_1(E) = \{(a_2 + \dots + L_{2n}a_n, \dots, L_{nn}a_n) : 0 \leq a_i \leq 1\}$ 时, 上面的积分才非零, 并且当 x_1 跑遍 \mathbb{R} 时恰好有长度为 1 的部分落在 $L(E)$ 中, 因此此时积分结果总是 1. 这也就是说,

$$\int_{\mathbb{R}} \chi_{L(E)} dx = \chi_{L_1(E)}.$$

所以

$$\int_{\mathbb{R}^n} \chi_{\chi_{L(E)}} = \int_{\mathbb{R}^{n-1}} \chi_{L_1(E)}.$$

因此可以一直递推下去, 最后会得到 $\int_{\mathbb{R}^n} \chi_{\chi_{L(E)}} = 1$.

- 最后证明任意矩阵都满足要求. 这是因为任意矩阵都可以分解为初等矩阵的乘积. 初等矩阵分为三种, 其中扩倍矩阵和置换矩阵下, 结论的成立性是显然的. 而最后剩下的形如 $I + E_{ij}$ 的矩阵已经被包含在上述上三角矩阵的讨论中. (事实上显然上面的命题对下三角矩阵也正确), 由此命题成立. \square

Proposition 1.3.9 (高维积分线性函数换元公式). 若有非退化的 $L: \mathbb{R}^d \rightarrow \mathbb{R}^d$ 及在 \mathbb{R}^d 上的可积函数 $f(x)$, 那么 $f(L(x))$ 也是 \mathbb{R}^d 上的可积函数, 并且

$$\int f(L(x)) dx = \frac{1}{\det L} \int f(x) dx.$$

这可以“理解为”

$$\int f(L(x)) dx = \int f(y) dL^{-1}(y) = \det L^{-1} \int f(y) dy = \frac{1}{\det L} \int f(x) dx.$$

即这个式子从 $dL(x) = \det L dx$ 推出. 这和一维情形的换元公式是类似的, 之后学习多元微积分时会有更广泛的形式.

1.4 Fubini 定理的应用

若 f, g 是 \mathbb{R}^d 上的光滑有紧支集函数, 那么可以定义卷积

$$f * g(x) = \int f(x-y)g(y) \, dy.$$

此时卷积 $f * g$ 是光滑有紧支集函数. 对一般的可积函数 f 和 g , 也能定义卷积, 并且证明方法也很奇特: 先把其当作二元函数延拓到高维再回到低维展开. 记

$$L(x, y) = f(x-y) \cdot g(y).$$

那么 $f(x-y)$ 和 $g(y)$ 均是可测函数 (因为对任意线性映射, f 可测能推出 $f \circ L$ 可测, 所以 $L(x, y)$ 可测, 再利用上它们可积的条件, 得到

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |f(x-y)g(y)| \, dy &= \int_{\mathbb{R}^d} |g(y)| \left(\int_{\mathbb{R}^d} |f(x-y)| \, dx \right) dy \\ &\leq \int_{\mathbb{R}^d} |g(y)| \cdot \|f\|_{\mathcal{L}} \, dy = \|f\|_{L^1} \cdot \|g\|_{L^1}. \end{aligned}$$

因此 $L(x, y)$ 是 \mathbb{R}^{2d} 上的可积函数, 那么固定 x , 它就是对几乎所有的 x 是关于 y 的 \mathbb{R}^d 上的可积函数. 再对此根据 x 积分, 其积分值就可被上面的积分控制, 因此 $f * g$ 可积.

卷积满足

1. 交换律: $f * g = g * f$;
2. 结合律: $(f * g) * h = f * (g * h)$; (可通过变元替换得到, $\int (\int f(x-z-y)g(y) \, dy)h(z) \, dz = \int f(x-z)(\int g(z-y)h(y) \, dy) \, dz$, 将其用 Fubini 定理化为高维积分后, 利用一个线性映射 $(y, z) \mapsto (z-y, y)$ 进行变元替换即可).
3. 分配律: $f * (g + h) = f * g + f * h$.

需要注意的是, 卷积不存在幺元, 所以无法构成一个群.

回忆

$$\begin{aligned} \Gamma(t) &= \int_0^{+\infty} e^{-x} x^{t-1} \, dx; \quad t > 0 \\ B(p, q) &= \int_0^1 x^{p-1} (1-x)^{q-1} \, dx; \quad p, q > 0. \end{aligned}$$

有如下定理:

Theorem 1.4.1 (B 函数和 Γ 函数的关系).

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p, q > 0.$$

Proof. 考虑用 Fubini 定理 1.3.1 把两个积分揉在一起:

$$\begin{aligned}
 & \int_0^{+\infty} e^{-x} x^{p-1} dx \cdot \int_0^{+\infty} e^{-y} y^{q-1} dy \\
 &= \int_{x,y>0} e^{-(x+y)} x^{p-1} y^{q-1} dx dy \\
 &= \int_{x,y>0} e^{-(x+y)} (x+y)^{p+q-2} \left(\frac{x}{x+y}\right)^{p-1} \left(\frac{y}{x+y}\right)^{q-1} dx dy \\
 &= \int_{t,y} e^{-t} t^{p+q-2} \left(1 - \frac{y}{t}\right)^{p-1} \left(\frac{y}{t}\right)^{q-1} dt dy \\
 & \text{(Fubini)} = \int_{t>0} e^{-t} t^{p+q-2} \left(\int_0^t \left(1 - \frac{y}{t}\right)^{p-1} \left(\frac{y}{t}\right)^{q-1} dy\right) dt \\
 &= \int_{t>0} e^{-t} t^{p+q-2} \left(\int_0^1 (1-s)^{p-1} (s)^{q-1} d(ts)\right) dt \\
 &= B(p, q) \cdot \int_{t>0} e^{-t} t^{p+q-1} dt \\
 &= B(p, q) \cdot \Gamma(p+q).
 \end{aligned}$$

证毕.



下面再来考虑 \mathbb{R}^n 中球的体积: 设单位球体积为 C_n , 则半径为 r 的球体积为 $C_n r^n$. 由于对于 n 维的球它的每个纬圆都是 $n-1$ 维的, 并且截距为 x 时其半径是 $\sqrt{1-x^2}$, 因此具有如下递推关系:

$$\begin{aligned}
 C_n &= 2 \int_0^1 C_{n-1} (\sqrt{1-x^2})^{n-1} dx \\
 &= C_{n-1} \int_0^1 (1-y)^{\frac{n-1}{2}} y^{-1/2} dy \quad (y=x^2) \\
 &= C_{n-1} \int_0^1 (1-y)^{\frac{n+1}{2}-1} y^{\frac{1}{2}-1} dy \\
 &= C_{n-1} B\left(\frac{n+1}{2}, \frac{1}{2}\right) \\
 &= C_{n-1} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n+2}{2})}.
 \end{aligned}$$

展开上面的递推式, 得到

$$\begin{aligned}
 C_n &= \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n+2}{2})} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n+1}{2})} \cdots \frac{\Gamma(\frac{4}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{5}{2})} C_2 \\
 &= \frac{\Gamma(\frac{1}{2})^{n-2} \cdot \Gamma(2) \cdot \pi}{\Gamma(\frac{n+2}{2})} = \frac{\pi^{n/2}}{\Gamma(\frac{n+2}{2})}.
 \end{aligned}$$

最后来利用多重积分推导 $\Gamma(1/2) = \sqrt{\pi}$.

$$\begin{aligned}
 \Gamma\left(\frac{1}{2}\right) &= \int_0^{+\infty} e^{-x} x^{-\frac{1}{2}} dx \\
 &= 2 \int_0^{+\infty} e^{-x^2} dx = \int_{-\infty}^{+\infty} e^{-x^2} dx.
 \end{aligned}$$

所以

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) \cdot \Gamma\left(\frac{1}{2}\right) &= \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy \\ &= \int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy.\end{aligned}$$

接下来要进行变量替换, 用到极坐标变元替换公式, 下面给出推导.

Lemma 1.4.2 (极坐标变元替换). 推导

$$dx \wedge dy = r \cdot dr \wedge d\theta.$$

Proof. 这里 x, y, r, θ 都是二元函数 $\mathbb{R}^2 \rightarrow \mathbb{R}$, 我们把右边进行全微分展开. 固定一点 $p(x_0, y_0)$, 则

$$dr|_p = \frac{\partial r}{\partial x} \Big|_p dx|_p + \frac{\partial r}{\partial y} \Big|_p dy|_p.$$

由于 $r(x_0, y_0) = \sqrt{x_0^2 + y_0^2}$, 因此

$$dr|_p = \frac{x_0}{\sqrt{x_0^2 + y_0^2}} dx|_p + \frac{y_0}{\sqrt{x_0^2 + y_0^2}} dy|_p.$$

同理, 由于 $\theta(x_0, y_0) = \tan^{-1} \frac{y_0}{x_0}$, 所以

$$d\theta|_p = \frac{-y_0}{x_0^2 + y_0^2} dx|_p + \frac{x_0}{x_0^2 + y_0^2} dy|_p.$$

所以 $dr, d\theta$ 作为余切向量场 $p \mapsto T_p(\mathbb{R}^2)^*$, 满足

$$dr = \frac{x}{\sqrt{x^2 + y^2}} dx + \frac{y}{\sqrt{x^2 + y^2}} dy, \quad d\theta = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy.$$

从而

$$\begin{aligned}dr \wedge d\theta &= \frac{x}{\sqrt{x^2 + y^2}} \frac{x}{x^2 + y^2} dx \wedge dy + \frac{y}{\sqrt{x^2 + y^2}} \frac{-y}{x^2 + y^2} dy \wedge dx \\ &= \frac{x^2 + y^2}{(x^2 + y^2)^{3/2}} dx \wedge dy = \frac{1}{\sqrt{x^2 + y^2}} dx \wedge dy.\end{aligned}$$

上式等价于 $dx \wedge dy = r \cdot dr \wedge d\theta$.



回到上面的命题, 我们有

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy &= \int_{r>0, \theta \in (0, 2\pi)} e^{-r^2} r dr d\theta \\ &= 2\pi \cdot \int_0^{+\infty} e^{-t} \sqrt{t} d\sqrt{t} \\ &= \pi \cdot \int_0^{+\infty} e^{-t} dt = \pi.\end{aligned}$$

这足以得到 $\Gamma(1/2) = \sqrt{\pi}$.

2 Lebesgue 微分定理

回忆微积分基本定理:

- 函数 f 在 $[a, b]$ 上 Riemann 可积, 在 $x_0 \in (a, b)$ 处连续, 则 $F(x) = \int_a^x f(t) dt$ 在 x_0 处可导, 并且 $F'(x_0) = f(x_0)$.
- 若 F , 导函数 F' Riemann 可积, 则 $F(b) - F(a) = \int_a^b F'(t) dt$.

接下来做的事情基本上是把上面的命题改写为 Lebesgue 可积情形. 上面的第一个命题可以写成

$$f(x) = \lim_{\delta \rightarrow 0} \frac{F(x+\delta) - F(x)}{\delta} = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_x^{x+\delta} f(x) dx.$$

在高维情形下, 把区间 $[x, x+\delta]$ 改成包含 x 的一个区域, 这就可以引出如下一般的情形:

2.1 微分定理的第一个形式: Lebesgue 基本定理

Theorem 2.1.1 (Lebesgue 微分 (基本) 定理). 如果 $f \in L^1(\mathbb{R}^d)$, 则

$$\lim_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B f = f(x), \text{ a.e. } x \in \mathbb{R}^d.$$

问题的入手点是对任意 $f \in L^1(\mathbb{R}^d)$ 定义极大函数

$$Mf(x) = \sup_{B: x \in B} \frac{1}{m(B)} \int_B |f|.$$

这样的考虑在调和和分析中经常出现. 注意到 Mf 的定义在选定某个 B 时与 $f(x)$ 本身无关, 这可以引出一些一致估计.

Proposition 2.1.2 (Hardy-Littlewood). 对 $f \in L^1$, Mf 满足:

1. Mf 可测;
2. 对几乎所有的 x , $|f(x)| \leq Mf(x) < +\infty$, a.e.
3. 存在常数 A 使得对任意 $\alpha > 0$, 有

$$m(\{x : Mf(x) > \alpha\}) \leq \frac{A}{\alpha} \|f\|_{L^1}.$$

其中第三条类似于 Chebyshev 不等式, 回忆在作业中有

$$m(\{x : |f(x)| > \alpha\}) \leq \frac{1}{\alpha} \|f\|_{L^1}.$$

说明 Mf 虽然比 $|f|$ 大, 但是 Mf 和 f 在“分布”的意义下等价.

Proof. (1) 对任意 $\alpha > 0$, 记 $E_\alpha = \{x : Mf(x) > \alpha\}$. 则对任意 $x \in E_\alpha$, 存在开球 B 使得 $x \in B$ 且 $\frac{1}{m(B)} \int_B |f| > \alpha$. 而 B 本身定义不依赖于 x , 所以对 B 中任一点均落在 E_α 中, 所以 $B \subset E_\alpha$, 因此 E_α 为开集. 故 E_α 可测, 所以 Mf 可测.

(2) 只需对任意 N 和 ε_0 证明

$$(P_{\varepsilon_0, N} :=) P = \{x : N > |f(x)| > Mf(x) + \varepsilon_0\}$$

为零测集. 否则假设其正测度, 那么对任意 $0 < \lambda < 1$, 存在集合 $E \subset P$ 及开球 B 满足

$$E \subset B, \quad |E| > \lambda|B|.$$

并且对任意 $x \in E$, 有

$$\begin{aligned} Mf(x) &\geq \frac{1}{|B|} \int_B |f(y)| dy \geq \frac{1}{|B|} \int_E |f(y)| dy \geq \frac{1}{|B|} \int_E (Mf(y) + \varepsilon_0) dy \\ &= \frac{1}{|B|} \left(\int_E Mf(y) dy + \varepsilon_0 |E| \right), \end{aligned}$$

注意到 RHS 与 x 无关, 但出现了相同的形式, 所以两边对 x 进行积分可以得到

$$\int_E Mf(x) dx \geq \frac{|E|}{|B|} \left(\int_E Mf(y) dy + \varepsilon_0 |E| \right).$$

现在关于 x 和 y 的东西是一样的, 所以把它们等同后得到

$$\left(1 - \frac{|E|}{|B|}\right) \int_E Mf(y) dy \geq \frac{|E|^2}{|B|} \varepsilon_0.$$

根据 $|E| > \lambda|B|$, 上式可进一步化简为

$$(1 - \lambda) \int_E Mf(y) dy \geq \lambda|E|\varepsilon_0.$$

因此整个积分存在下界. 另一方面, 根据 P 的上界限制得到 $Mf(y) \leq N$, 所以

$$(1 - \lambda) \cdot N|E| \geq \lambda|E|\varepsilon_0.$$

现在令 $\lambda \rightarrow 1$ 就得到矛盾!

那么最后考虑

$$P_{\varepsilon_0} = \{x : |f(x)| > Mf(x) + \varepsilon_0\} = \bigcup_{N=1}^{\infty} P_{\varepsilon_0, N}, \text{ a.e.}$$

他也是零测集. 因此

$$\{x : |f(x)| > Mf(x)\} = \bigcup_{m=1}^{\infty} P_{\frac{1}{m}}$$

也是零测集. 所以对几乎处处的 x 结论成立.

剩下的证明 Mf 几乎处处有限将由第三问推出.

(3) 首先根据 (1) 中的讨论, 对任意 $x \in E_\alpha$, 可以取一个包含 x 的开球 B 满足

$$x \in B, \quad B \subset E_\alpha, \quad \int_B |f| > \alpha \cdot |B|.$$


这来源于 Mf 的定义, 由此得到 $|B|$ 的一个上界. 想要得到 $m(E_\alpha)$ 的一个上界, 应该取一个足够靠近整个集合的“开覆盖”——类似于方块对可测集的足够接近的覆盖, 从而对上面所有的不等式求和就能“几乎”得到 $m(E_\alpha)$ 的一个上界, 由此推得结论. 这里所谓的“几乎”, 根据结论要求, 只需要量级相同, 但需要对任意集合此上界是一致的. 所以, 问题的关键在于下面的覆盖引理.

Lemma 2.1.3. 对于 \mathbb{R}^d 中有限个开球 $\{B_1, \dots, B_N\}$, 存在 l 个互不相交的开球 $\{B_{n_1}, \dots, B_{n_l}\}$ 使得

$$m\left(\bigcup_{j=1}^N B_j\right) \leq 3^d m\left(\bigcup_{k=1}^l B_{n_k}\right).$$

Proof. 设 B_j 的半径为 r_j , 不妨假设 $r_1 \geq r_2 \geq \dots \geq r_N$. 我们贪心地取出一些球:

- 按照半径从大到小的顺序取一些球, 首先取出半径最大的球.
- 第 k 步时, 如果半径为 r_k 的球与前面已经选取的球有交, 那么就不把它选出来, 否则就把它选出来, 直至最后就可以得到一个互不相交的开球子列.
- 我们证明选出来的那些球球心不变, 半径扩大 3 倍时就可以覆盖所有球构成的集合. 这是因为如果某个球未被选出 (设为 r_k), 它一定与某个半径比它大的球相交 (设为 r_s), 因此两个圆心之间距离小于 $r_k + r_s \leq 2r_s$. 所以小球内的所有点到小球球心距离至多为 $2r_k + r_s \leq 3r_s$, 是故能被半径扩大三倍后的球覆盖.

- 因此 $m\left(\bigcup_{j=1}^N B_j\right) \leq m\left(\bigcup_{k=1}^l (3B_{n_k})\right)$. 这直接导出了结论. 

Remark. 覆盖引理还有各种形式的推广. 比如如果有可数个球, 那么不能把它们按半径排序, 但是可以取出半径大于上确界的一半的球, 因此用同样的方法可知也许把半径扩大五倍就可以覆盖所有球. 此外, 还有对半径上界进行控制的 **Vitali 覆盖引理**, 它告诉我们对任意可测集 E , 可以选取至多可数的且半径不超过 d 的不交开球列 B_k 使得 $m(E) \leq 5^d \cdot \sum |B_k|$.

最后还需指出, Lebesgue 微分定理里面要求 B 是球, 就是因为球满足形如这样的 Covering Lemma, 在其他地方都没有用到球这种模型的特殊性. 所以说, 如果一类图形也能满足这样的覆盖引理 (比如方体), 那么结论也成立. 这样的图形需要具有某种“结构性”.

回到 2.1.2 (3) 的证明: 可以考虑用有界紧集 $K \subset E_\alpha$ 来逼近 E_α . 由上述分析, 对每个 $x \in K$ 选定 $B_x \subset E_\alpha$ 使得

$$|B_x| < \frac{1}{\alpha} \int_B |f|.$$

则 $\bigcup_{x \in K} B_x$ 构成 K 的一个开覆盖, 因此可选择有限个开球 B_1, \dots, B_N 使得它们的并覆盖 K . 由引理 2.1.3, 可再选出子开球列 B_{n_k} 使得

$$m(K) \leq m\left(\bigcup_{j=1}^N B_j\right) \leq 3^d \sum_{k=1}^l m(B_{n_k}) \leq \frac{3^d}{\alpha} \int_{\bigcup B_{n_k}} |f| \leq \frac{3^d}{\alpha} \|f\|_{L^1}.$$

右边与 K 无关, 因此可取 K 的测度充分接近 E_α , 由此得到

$$m(E_\alpha) \leq \frac{3^d}{\alpha} \|f\|_{L^1}.$$

取 $A = 3^d$, 证毕!



现在回到微分定理 2.1.1 的证明, 实际上可以证明更强的命题: 对 $f \in L^1(\mathbb{R}^d)$, 有

$$\lim_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy = 0, \text{ a.e. } x \in \mathbb{R}^d.$$

Proof for 2.1.1. 我们相对一个 L^1 范数相对小的函数应用上述引理, 为此采用连续函数逼近. 对任意 $\varepsilon > 0$, 取连续函数 g 使得 $\|f - g\|_{L^1} < \varepsilon$. 那么对 x 附近利用连续性条件, 那么对任意 $\varepsilon_0 > 0$, 均存在 δ 使得对测度小于 δ 且包含 x 的开球 B , 总有

$$|g(x) - g(y)| < \varepsilon_0.$$

固定 $x \in B$ 满足 $m(B) < \delta$, 对积分进行分拆:

$$\frac{1}{m(B)} \int_B |f(y) - f(x)| dy \leq \frac{1}{m(B)} \int_B (|g(y) - g(x)| + |f(y) - g(y)| + |f(x) - g(x)|) dy.$$

那么第一项可以放为 ε_0 , 第三项与 y 无关, 就是 $|f(x) - g(x)|$, 第二项可以放成 $M(f - g)(x)$ (为了利用上引理 2.1.2, 直接放成 $\|f - g\|_{L^1}$ 显然太差了). 总结一下就是

$$\frac{1}{m(B)} \int_B |f(y) - f(x)| dy \leq \varepsilon_0 + |f(x) - g(x)| + M(f - g)(x). \quad (*)$$

由引理 2.1.2 和 Chebyshev 不等式可得,

$$\begin{aligned} m(\{x : M(f - g)(x) > \alpha\}) &\leq \frac{A}{\alpha} \|f - g\|_{L^1} \leq \frac{A\varepsilon}{\alpha}, \\ m(\{x : |f(x) - g(x)| > \alpha\}) &\leq \frac{1}{\alpha} \|f - g\|_{L^1} \leq \frac{\varepsilon}{\alpha}. \end{aligned}$$

所以这两个式子在“依测度”的意义下都能被 ε 这个参数所控制. 现在取 $\alpha = \varepsilon_0$, 并且注意到 (*) RHS 和 B 的选取无关, 这也就是说

$$m\left(\left\{x : \sup_{x \in B, m(B) < \delta} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy > 3\varepsilon_0\right\}\right) \leq \frac{(A+1)\varepsilon}{\varepsilon_0}. \quad (**)$$

为了消去 (**) 中的 δ , 使用上极限:

$$m\left(\left\{x : \limsup_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy > 3\varepsilon_0\right\}\right) \leq \frac{(A+1)\varepsilon}{\varepsilon_0}. \quad (***)$$

左边式子内的 x 都落在 $(**)$ LHS 集合中. 现在式子左右均和 g 的选取无关, 所以令 $\varepsilon \rightarrow 0$ 可知

$$m\left(\left\{x: \limsup_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy > 3\varepsilon_0\right\}\right) = 0.$$

再令 $\varepsilon_0 \rightarrow 0$ 可知

$$m\left(\left\{x: \limsup_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy > 0\right\}\right) = 0.$$

这和结论等价!



2.2 Lebesgue 微分定理的应用

Definition 2.2.1 (Lebesgue 点). 我们称满足 $|f(x)| < +\infty$, 且

$$\lim_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy = 0$$

的点为 **Lebesgue 点**. 所有这样的点构成 **Lebesgue 集**.

根据 Lebesgue 微分定理, 对 $f \in L^1(\mathbb{R}^d)$, 几乎所有点都是 Lebesgue 点. 注意到 Lebesgue 点的定义具有局部性, 因此可以推广得到如下结论:

Proposition 2.2.2. 局部可积函数空间 $L^1_{loc}(\mathbb{R}^d)$ 定义为所有的 f , 满足对任何有限测度集 E , f 在 E 上的限制 $f\chi_E$ 是 Lebesgue 可积的. 那么 Lebesgue 微分定理对 $L^1_{loc}(\mathbb{R}^d)$ 内的函数也成立.

Proof. 对 $f \in L^1_{loc}(\mathbb{R}^d)$ 与任意 $N > 0$, $f\chi_{B_N(0)} \in L(\mathbb{R}^d)$. 所以 $f\chi_{B_N(0)}$ 内几乎所有点都是 Lebesgue 点, 而在 $B_N(0)$ 中 $f\chi_{B_N(0)}$ 的 Lebesgue 点就是 f 的 Lebesgue 点, 所以 $B_N(0)$ 内几乎所有点都是 f 上的 Lebesgue 点. 由于可数个零测集的并还是零测集, 所以 \mathbb{R}^d 中几乎所有点都是 f 上的 Lebesgue 点.



特别地, 对任意可测集 E , $\chi_E \in L^1_{loc}(\mathbb{R}^d)$ (但当 E 无限测度时它不落在 L^1 中), 有

$$\frac{1}{m(B)} \int_B \chi_E = \frac{m(E \cap B)}{m(B)}.$$

称为 E 在 B 中的**密度**, 那么对它利用 Lebesgue 微分定理, 可以得到:

Theorem 2.2.3 (Lebesgue 密度定理). 称 x 为 E 的**密度点**如果

$$\lim_{x \in B, |B| \rightarrow 0} \frac{m(B \cap E)}{m(B)} = 1.$$

那么对于可测集 E , 几乎所有 E 中点均为 E 中的密度点, 几乎所有 E 之外的点都不是 E 中的密度点.

再来看它在卷积上的应用. 下设 c_d 是 \mathbb{R}^d 中单位球的体积.

$$\begin{aligned} \frac{1}{|B(x, \varepsilon)|} \int_{B(x, \varepsilon)} f &= c_d^{-1} \varepsilon^{-d} \int_{B(0, \varepsilon)} f(x-y) dy \\ &= \int f(x-y) \cdot (c_d^{-1} \varepsilon^{-d} \chi_{B(0, \varepsilon)}) dy \\ &= f * (c_d^{-1} \varepsilon^{-d} \chi_{B(0, \varepsilon)})(x). \end{aligned}$$

那么设

$$K_\varepsilon = c_d^{-1} \varepsilon^{-d} \chi_{B(0, \varepsilon)},$$

根据 Lebesgue 微分定理, LHS 在 $\varepsilon \rightarrow 0$ 时几乎处处趋于 $f(x)$. 所以

$$\lim_{\varepsilon \rightarrow 0} f * K_\varepsilon(x) = f(x), \text{ a.e. } x \in \mathbb{R}^d.$$

我们希望寻找满足这个要求的一族函数 K_ε (这里的定义显然不会是一个很好的函数), 使得在某种意义上对满足某种条件的函数 f , 有

$$f * K_\varepsilon \xrightarrow{?} f.$$

关于这个话题, 我们讨论的这族 K_ε 被称为**卷积核**. 如果卷积核还满足下面的更强的条件, 那么这个卷积核就会在某种意义上成为“恒同量”.

Definition 2.2.4 (恒同逼近). 如果 \mathbb{R}^d 中的卷积核 K_ε 满足

$$\int_{\mathbb{R}^d} K_\varepsilon = 1; \tag{1}$$

$$|K_\varepsilon(x)| \leq A\varepsilon^{-d}; \tag{2}$$

$$|K_\varepsilon(x)| \leq A\varepsilon \cdot |x|^{-d-1}. \tag{3}$$

那么 K_δ 被称为**恒等逼近函数**.

上面积分为 1 的条件原因在于 $\|f * K_\varepsilon\|_{L^1} = \|f\|_{L^1} \|K_\varepsilon\|_{L^1}$, 这保证了卷积后积分值不变. 逐点的条件主要是要求无穷远卷积核衰减足够快, 而在 0 附近一致有界. 这在后面两个命题中的证明里体现地很明显. 首先是 L^1 意义下的收敛性:

Proposition 2.2.5. 如果 K_ε 是恒同逼近, 那么对任何 $f \in L^1(\mathbb{R}^d)$, 均有 $K_\varepsilon * f$ 可积, 并且

$$\lim_{\varepsilon \rightarrow 0} \|K_\varepsilon * f - f\|_{L^1} = 0.$$

Proof. 首先先把 $K_\varepsilon * f - f$ 写为

$$\begin{aligned} K_\varepsilon * f - f &= \int K_\varepsilon(x-y) f(y) dy - f(x) \\ &= \int K_\varepsilon(y) (f(x-y) - f(x)) dy. \end{aligned}$$

两边取 L^1 范数, RHS 化为关于 x, y 的重积分:

$$\|K_\varepsilon * f - f\|_{L^1} \leq \iint |K_\varepsilon(y)| \cdot |f(x-y) - f(x)| \, dy \, dx.$$

我们注意到 K_ε 有 (2) 和 (3) 两个上界, 其中 (2) 是对 $|y|$ 较小时的控制, (3) 是对 $|y|$ 较大时候的控制. 所以交换积分后进行放缩

$$\begin{aligned} \|K_\varepsilon * f - f\|_{L^1} &\leq \int_{|y| \leq \delta} A \cdot \varepsilon^{-d} \cdot \left(\int |f(x-y) - f(x)| \, dx \right) dy \\ &\quad + \int_{|y| > \delta} A \cdot \varepsilon |y|^{-1-d} \cdot \left(\int |f(x-y) - f(x)| \, dx \right) dy. \end{aligned}$$

里面的积分也在 y 较小和较大的时候有不同的放缩. y 较小时, $f(x-y) - f(x)$ 可看作 f 在 x 附近的微小平移, 它的积分会有一个上界. 而 y 较大时, 直接放成 $2|f(x)|$ 即可.

$$\|K_\varepsilon * f - f\|_{L^1} \leq \int_{|y| \leq \delta} A \cdot \varepsilon^{-d} \cdot \|\tau_{-y}f - f\|_{L^1} \, dy + \int_{|y| > \delta} A \cdot \varepsilon |y|^{-1-d} \cdot 2\|f\|_{L^1} \, dy.$$

由平移算子的性质, 对任意 $\varepsilon_0 > 0$, 存在 $\delta = \delta(\varepsilon_0)$ 使得

$$\|\tau_h f - f\|_{L^1} \leq \varepsilon_0, \quad \forall |h| \leq \delta.$$

故左边的积分号可以消去, 留下 $\int_{|y| \leq \delta} 1 \, dy = |B_\delta(0)| = \delta^d C_d$. 而右边的积分中只剩下 $\int_{|y| > \delta} |y|^{-1-d} \, dy$, 对其积分可得

$$\int_{|y| > \delta} |y|^{-1-d} \, dy = S_{d-1} \cdot \int_\delta^\infty |y|^{-1-d} \cdot |y|^{d-1} \, dy = S_{d-1} \cdot \int_\delta^\infty y^{-2} \, dy = \frac{S_{d-1}}{\delta}.$$

这样上面的式子可以把积分全部去除, 进一步放缩为

$$\begin{aligned} \|K_\varepsilon * f - f\|_{L^1} &\leq A \cdot \varepsilon^{-d} \cdot \varepsilon_0 \cdot \delta^d C_d + 2\|f\|_{L^1} A \varepsilon \cdot S_{d-1} \delta^{-1} \\ &= (AC_d) \cdot \varepsilon_0 \cdot \delta^d \varepsilon^{-d} + (2\|f\|_{L^1} AC) \cdot \varepsilon \delta^{-1}. \end{aligned}$$

现在我们要证明 $\varepsilon \rightarrow 0$ 时上式 $\rightarrow 0$. 因此应当固定 ε_0, δ , 寻找 ε 使得上式取到可趋于 0 的值, 这需要等式两边均含有 ε_0 的正次幂即可. 比如说取

$$\varepsilon = \delta \cdot \varepsilon_0^{\frac{1}{d+1}},$$

那么就有

$$\|K_\varepsilon * f - f\|_{L^1} \leq (AC_d) \cdot \varepsilon_0^{\frac{1}{d+1}} + (2\|f\|_{L^1} AC) \cdot \varepsilon_0^{\frac{1}{d+1}}.$$

上式在 $\varepsilon \rightarrow 0$ 时趋于 0, 因此

$$\lim_{\varepsilon \rightarrow 0} \|K_\varepsilon * f - f\|_{L^1} = 0.$$

证毕. \square

接下来还有逐点收敛意义下的恒同性, 某种意义上, 这是最直观的一种收敛.

Proposition 2.2.6. 如果 f 可积, K_ε 为恒同逼近, 则 f 在所有 Lebesgue 点处有

$$\lim_{\varepsilon \rightarrow 0} (K_\varepsilon * f)(x) = f(x).$$

特别地, 局部 Lebesgue 可积函数中几乎所有点均是 Lebesgue 点, 所以 $\lim_{\varepsilon \rightarrow 0} (K_\varepsilon * f) \rightarrow f$, a.e.

Proof. 首先还是一样的放缩:

$$|K_\varepsilon * f - f| \leq \int |K_\varepsilon(y)| \cdot |f(x-y) - f(x)| dy.$$

利用 Lebesgue 点的条件, 在

$$\lim_{x \in B, m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy = 0$$

中令 B 为以 x 为球心的标准球, 再利用平移性质得到

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{1}{m(B(0, r))} \int_{B(0, r)} |f(x-y) - f(x)| dy &= 0. \\ \Rightarrow \lim_{r \rightarrow 0} \omega(r) &:= \lim_{r \rightarrow 0} r^{-d} \cdot \int_{B(0, r)} |f(x-y) - f(x)| dy = 0. \end{aligned}$$

这恰好可以与 y 较小的情况的放缩对应. 所以把上面的积分拆分为 y 较小和较大的情形, 依然是与上题类似的拆分:

$$\begin{aligned} |K_\varepsilon * f - f| &\leq \int_{|y| \leq r} A\varepsilon^{-d} |f(x-y) - f(x)| dy \\ &\quad + \int_{|y| > r} A\varepsilon |y|^{-1-d} |f(x-y) - f(x)| dy \\ &\leq A\varepsilon^{-d} \cdot \omega(r) \cdot r^d + A\varepsilon \cdot \int_{|y| > r} |y|^{-1-d} |f(x-y) - f(x)| dy. \end{aligned}$$

显然我们并不能对后面的式子求积分, 所以还得把 $|y|^{-1-d}$ 放掉, 为此采用翻倍截断的方法 (正如证明调和级数发散时的截断), 得到

$$\begin{aligned} \int_{|y| > r} |y|^{-1-d} |f(x-y) - f(x)| dy &\leq \sum_{k=0}^{\infty} \int_{2^k r < |y| \leq 2^{k+1} r} (2^k r)^{-1-d} |f(x-y) - f(x)| dy \\ &= \sum_{k=0}^{\infty} (2^k r)^{-1-d} \cdot \int_{|y| \leq 2^{k+1} r} |f(x-y) - f(x)| dy \\ &= \sum_{k=0}^{\infty} (2^k r)^{-1-d} \cdot \omega(2^{k+1} r) \cdot (2^{k+1} r)^d. \end{aligned}$$

恰好后面的项又能写为与 ω 相关的项, 只不过不一定取值趋向于 0. 然而, 经过简单的计算可发现 ω 确实是有界函数:

$$\begin{aligned} \omega(r) &\leq r^{-d} \cdot \left(\int_{B(0, r)} |f(x-y)| dy + \int_{B(0, r)} |f(x)| dy \right) \\ &\leq r^{-d} \cdot \|f\|_{L^1} + C_d \cdot |f(x)| \\ &\leq \|f\|_{L^1} + C_d \cdot |f(x)|. \quad (r > 1) \end{aligned}$$

在 $r \leq 1$ 时利用连续性可得有界性. 那么可以设 $|\omega(r)| \leq M$, 这就得到

$$\begin{aligned} |K_\varepsilon * f - f| &\leq A \cdot \omega(r) \cdot \varepsilon^{-d} r^d + A\varepsilon \cdot \sum_{k=0}^{\infty} (2^{d-k} \cdot r^{-1} \cdot \omega(2^{k+1}r)) \\ &\leq A \cdot \omega(r) \cdot \varepsilon^{-d} r^d + 2^d \cdot A \cdot r^{-1} \varepsilon \cdot \sum_{k=0}^{\infty} 2^{-k} \cdot \omega(2^{k+1}r). \end{aligned}$$

这很显然诱导我们令 $r = \varepsilon$, 那么整个式子中 ε 就只会出现在 ω 内部出现:

$$|K_\varepsilon * f - f| \leq A \cdot \omega(\varepsilon) + 2^d \cdot A \cdot \sum_{k=0}^{\infty} 2^{-k} \cdot \omega(2^{k+1}\varepsilon).$$

注意不能把后面的 $\omega(2^{k+1}\varepsilon)$ 直接放为 M , 而可以待定 N , 把后面的级数再分拆为 $k \leq N$ 和 $k > N$ 两部分:

$$|K_\varepsilon * f - f| \leq A \cdot \omega(\varepsilon) + 2^d \cdot A \cdot \left(\sum_{k=0}^N 2^{-k} \cdot \omega(2^{k+1}\varepsilon) + 2^{-N} \cdot M \right).$$

那么对任意 δ , 可以先取足够大的 N 使得 $2^d \cdot A \cdot M \cdot 2^{-N} < \delta$. 然后对充分小的 ε 均有

$$\omega(2^k \varepsilon) < \frac{\delta}{2^d A}, k = 0, 1, \dots, N+1.$$

那么就有 $|K_\varepsilon * f - f| < \delta + 2\delta + \delta = 4\delta$. 令 $\delta \rightarrow 0$ 即得结论. \square

最后来看一些例子.

Example 2.2.7 (Friedrichs 光滑化核). 对光滑有紧支集函数 $\varphi \in C_c^\infty(\mathbb{R}^d)$ 满足 $\int_{\mathbb{R}^d} \varphi = 1$, 对 $\varepsilon > 0$, 考虑卷积核 φ_ε , 定义为

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n} \varphi\left(\frac{x}{\varepsilon}\right).$$

所以所有微分运算均可以看作仅在 g 上作微分, 而 f 不发生变化. 后面几个例子也用到这点) 可以证明 φ_ε 是恒同逼近. 那么对任意局部可积函数 $f \in L_{loc}^1(\mathbb{R}^d)$, 就有 $f * \varphi_\varepsilon \in C^\infty(\mathbb{R}^d)$ (因为当 g 光滑时, 可以证明

$$\frac{\partial}{\partial x_i}(f * g)(x) = \left(f * \frac{\partial g}{\partial x_i}\right)(x).$$

, 并且在某种意义上有 $f * \varphi_\varepsilon \rightarrow f$, 这就是“光滑化”一词的来源. 在实际应用中, 可以利用 L^1 收敛或逐点收敛取出 f 的光滑逼近来完成一些命题的证明. 这个逼近的关键在于我们“可以写出表达式”, 一般来说这是做不到的.

Example 2.2.8 (热核). 热方程初值问题

$$(\partial_t - \Delta)u(t, x) = 0, \quad u(0, x) = u_0(x). \quad \Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}.$$

有解. 证法是利用 Fourier 变换, 设

$$\begin{aligned} u(t, x) &= \frac{1}{\sqrt{2\pi}} \int g(t, k) e^{ikx} dk, \\ \Rightarrow g(t, x) &= \frac{1}{\sqrt{2\pi}} \int u(t, k) e^{-ikx} dx. \end{aligned}$$

所以可得到

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{1}{\sqrt{2\pi}} \int -k^2 \cdot g(t, k) e^{ikx} dk, \\ \frac{\partial u}{\partial t} &= \frac{1}{\sqrt{2\pi}} \int \frac{\partial}{\partial t} g(t, k) e^{ikx} dk. \end{aligned}$$

这可得到

$$\frac{\partial}{\partial t} g(t, k) = -k^2 g(t, k). \Rightarrow g(t, k) = g(0, k) e^{-k^2 t}.$$

这就推出

$$u(t, x) = \mathcal{F}(g(0, k) e^{-k^2 t}) = \mathcal{F}(e^{-k^2 t}) * \mathcal{F}(g(0, k)).$$

这来源于作业 34. 由此可见, 卷积将会在 Fourier 变换中频繁出现, 因为它就代表着乘法. 另一方面, 也能同构上例所说的, 先验证 $H_t := \mathcal{F}(e^{-k^2 t})$ 满足热方程, 然后验证所有 $H_t * u_0$ 也满足热方程, 并且也有逐点收敛性成立.

关于热核与 Poisson 核的恒同逼近性, 作业 42 和 43 中有详细解释.

2.3 微分定理的第二个形式与单调函数

Proposition 2.3.1. 对满足什么条件的 F , Newton-Leibniz 公式

$$F(b) - F(a) = \int_a^b F'(x) dx$$

成立?

Example 2.3.2. • Heaviside 函数 $H(x) = \chi_{(0, +\infty)}$, 导数在 0 处不存在, 在其他地方 $H' = 0$. 此时

$$H(1) - H(0) = 1 \neq \int_0^1 H'(x) dx.$$

• Cantor-Lebesgue 函数 $F(x)$ 在 Cantor 集 (零测集) 之外 $F' = 0$. 但此时

$$F(1) - F(0) = 1 \neq \int_0^1 F'(x) dx.$$

• $F(x) = x^2 \sin \frac{1}{x^2} (x \neq 0), F(0) = 0$, 那么 F' 处处存在, 但是 F' 不 Lebesgue 可积, 它也不满足 Newton-Leibniz 公式.

接下来定义的 Dini 导数的重要意义是它对任意的函数都一定存在, 所以可以通过讨论它的性质来得到可微性.

Definition 2.3.3 (Dini 分解).

$$\Delta_h(F)(x) = \frac{F(x+h) - F(x)}{h}, \quad h \neq 0.$$

定义

$$D^+(F)(x) = \limsup_{h \rightarrow 0, h > 0} \Delta_h(F)(x);$$

$$D_+(F)(x) = \liminf_{h \rightarrow 0, h > 0} \Delta_h(F)(x);$$

$$D^-(F)(x) = \limsup_{h \rightarrow 0, h < 0} \Delta_h(F)(x);$$

$$D_-(F)(x) = \liminf_{h \rightarrow 0, h < 0} \Delta_h(F)(x).$$

还可以定义上下导数

$$\overline{D}(F)(x) = \max\{D^+(F)(x), D^-(F)(x)\}; \quad \underline{D}(F)(x) = \min\{D_+(F)(x), D_-(F)(x)\}$$

为了回答问题 2.3.1, 本节的关键在于下面的单调函数 Lebesgue 定理. 接下来的一系列推理会告诉我们单调函数是足够好的函数.

Theorem 2.3.4 (单调函数 Lebesgue 定理). $F(x)$ 在 $[a, b]$ 上单调增, $F'(x)$ 几乎处处存在, 并且非负可积. 那么就有

$$\int_{[a,b]} F'(x) dx \leq F(b) - F(a).$$

上面的例子告诉我们等号肯定不成立.

Proof. 思路可以表述如下:

1. 单调增函数的跳跃点可数, 因此可以分解成跳跃函数和连续单增函数的和.
2. 跳跃函数导数几乎处处为 0.
3. 连续单调增函数几乎处处可导, 其导数满足该不等式.

Remark. 最后不等式的证明不是重点, 前面证明 a.e. 可微是更重要的. 从逻辑上来讲, 这个问题的证明可以只从 2.3.8 开始, 前面做的事情只是为了让单调函数的结构更加清楚.

其中第一点在半年前就讲过, 是因为可以定义左右极限 $F(x+0)$ 和 $F(x-0)$, 其中 x 为不连续点当且仅当 $F(x+0) - F(x-0) > 0$. 那么不连续点 x 就对应两两不交的区间, 因此是可数集, 记为 $\{x_n\}$. 记

$$\alpha_n = F(x_n + 0) - F(x_n - 0) > 0, \beta_n = F(x_n) - F(x_n - 0).$$

那么可以定义跳跃函数

$$j_n(x) = \begin{cases} 0, & x < x_n; \\ \beta_n, & x = x_n; \\ \alpha_n, & x > x_n. \end{cases}$$

它在 x 附近的“跳跃性”与 F 相同. 那么

$$J(x) = J_F(x) := \sum j_n(x),$$

自然地成为 F 的跳跃函数. 直观上来说, $F - J_F$ 会把所有跳跃的地方“抹平”, 从而成为连续函数. $\left\langle \frac{u}{x} \right\rangle$

Lemma 2.3.5 (单调函数分解). $F - J_F$ 是连续单调增函数.

Proof. 只需证明左右极限均等于原函数取值:

$$\lim_{h \rightarrow 0, h > 0} (F(x \pm h) - J_F(x \pm h)) = F(x) - J_F(x).$$

由于 F 单调增, 因此 $F(x \pm 0)$ 总存在. 所以上式可化为证明

$$F(x \pm 0) - F(x) = \lim_{h \rightarrow 0, h > 0} J_F(x \pm h) - J_F(x).$$

由于 $\sum_{n > N} j_n(x)$ 可以被 $\sum_{n > N} \alpha_n$ 一致地控制, 且 $\sum \alpha_n < +\infty$, 所以 J_F 作为级数关于 x 一致收敛, 这导出

$$\begin{aligned} \lim_{h \rightarrow 0, h > 0} J_F(x + h) &= \sum_n \lim_{h \rightarrow 0, h > 0} j_n(x + h), \\ \lim_{h \rightarrow 0, h > 0} j_n(x + h) &= \begin{cases} j_n(x), & x \neq x_n; \\ \alpha_n, & x = x_n. \end{cases} \quad \lim_{h \rightarrow 0, h > 0} j_n(x - h) = \begin{cases} j_n(x), & x \neq x_n; \\ 0, & x = x_n. \end{cases} \end{aligned}$$

因此

$$\lim_{h \rightarrow 0, h > 0} J_F(x \pm h) - J_F(x) = \mathbf{1}_{x=x_n} \cdot \begin{cases} \alpha_n - \beta_n \\ -\beta_n \end{cases}.$$

而根据 α_n, β_n 的定义, 容易知道当 $x = x_n$ 时上面式子实际上就是 $F(x \pm 0) - F(x)$.

而 $x \neq x_n$ 时也有 $F(x \pm 0) - F(x) = 0$, 这就证明了连续性.

再证单调性, 设 $a \leq x < y \leq b$. 如果 x 为 F 的连续点, 那么

$$\begin{aligned} J_F(y) - J_F(x) &= \sum_{x_k \leq y} j_k(y) - \sum_{x_k \leq x} j_k(x) \\ &= \sum_{x < x_k \leq y} j_k(y) \leq F(y) - F(x). \end{aligned}$$

如果 x 为不连续点, 那么

$$\begin{aligned} J_F(y) - J_F(x) &= \sum_{x_k \leq y} j_k(y) - \sum_{x_k \leq x} j_k(x) \\ &= \sum_{x_k < x} \alpha_k + \sum_{x \leq x_k \leq y} j_k(y) - \sum_{x_k < x} \alpha_k - \beta_{x_n=x} \\ &= (\alpha_n - \beta_n) + \sum_{x < x_k \leq y} j_k(y) \leq F(y) - F(x). \end{aligned}$$

这就证明了单调性. $\left\langle \frac{u}{x} \right\rangle$

Proposition 2.3.6. 若 $J(x)$ 为跳跃函数, 则 $J(x)$ 几乎处处可导且导数为零. (回忆 Cantor-Lebesgue 函数满足该性质)

Proof. 回忆 Dini 分解 2.3.3, 我们证明

$$E_\varepsilon = \{x : \overline{D}(J)(x) = \limsup_{h>0, h \rightarrow 0} \frac{J(x+h) - J(x)}{h} > \varepsilon\}$$

为零测集. 现在要对这个函数进行分段, 对任意 $\delta > 0$, 首先取足够大的 N 使得

$$\sum_{k>N} \alpha_k < \delta.$$

这也就是说较大的那些 $J_k(x)$ 的和贡献非常小, 记 $J_N = \sum_{k>N} j_k$, 那么对任意 $x \in [a, b]$, 均有

$$J_N(x) \leq \sum_{k>N} |j_k(x)| \leq \sum_{k>N} \alpha_k < \delta \Rightarrow |J_N(b) - J_N(a)| \leq \delta.$$

这就得到了整体“跃度”的上界. 现在把 $\overline{D}(J)(x)$ 作分段, 得到

$$\overline{D}(J)(x) = \limsup_{h>0, h \rightarrow 0} \frac{J_N(x+h) - J_N(x)}{h} + \sum_{k=1}^N \limsup_{h>0, h \rightarrow 0} \frac{j_k(x+h) - j_k(x)}{h}.$$

显然当 $x \notin \{x_1, \dots, x_N\}$ 时, 右边极限内取值总为零, 那么我们就只需要关心前一项的贡献. 具体而言, 设

$$E_{N,\varepsilon} = \{x : \overline{D}(J_N)(x) > \varepsilon\},$$

那么 E_ε 应满足

$$E_\varepsilon \subset E_{N,\varepsilon} \cup \{x_1, \dots, x_N\} \Rightarrow m(E_\varepsilon) \leq m^*(E_{N,\varepsilon}).$$

考察任意的 $x \in E_{N,\varepsilon}$, 由上极限定义, 存在一列趋于零的 h 使得

$$J_N(x+h) - J_N(x) \geq \varepsilon \cdot h \Rightarrow J_N(x+h) - J_N(x-h) \geq \varepsilon \cdot h. \quad (1)$$

这表明有一列开区间 $(x-h, x+h)$ (记为 $B_{x,i}$) 覆盖 x , 且它们对应的函数值跃度存在一个较好的下界. 如果我们选择一族互不相交的 $B_x \ni x$ 覆盖大部分 $E_{N,\varepsilon}$ 中的点, 那么 $E_{N,\varepsilon}$ 的测度就可以被这族区间的函数值跃度之和控制, 从而被 $J_N(b) - J_N(a) \leq \delta$ 控制, 而 δ 是我们任选的值. 至此, 只需要如同引理 2.1.3 一样运用无穷版本的覆盖引理, 就能类似地完成证明. 我们严格地把它写出来:

Lemma 2.3.7 (无穷覆盖引理). 设 $\{B_\alpha\}_{\alpha \in \mathcal{A}}$ 是 \mathbb{R}^d 中一族半径有界的开球, 那么存在其可数子族 $\{B_j\}_{j \in J \subset \mathcal{A}}$ 使得 B_j 两两不交, 且

$$m^*\left(\bigcup_{\alpha} B_\alpha\right) \leq 5^d \cdot \sum_j m(B_j).$$

假设引理成立, 对 $\bigcup_{x,i} B_{x,i}$ 利用引理, 由于 $E_{N,\varepsilon} \subset \bigcup_{x,i} B_{x,i}$, 因此

$$m^*(E_{N,\varepsilon}) \leq 5 \cdot \sum_j m(B_j).$$

特别地, J 应为可数集. 另一方面, 根据 (1) 可得对每个 $B_j = (p_j, q_j)$ 均有

$$J_N(q_j) - J_N(p_j) \geq \frac{\varepsilon}{2} \cdot m(B_j).$$

由于 J_N 为递增函数且 B_j 两两不交, 所以上式可以两边对 j 求和:

$$\delta \geq J_N(b) - J_N(a) \geq \sum_j J_N(q_j) - J_N(p_j) \geq \frac{\varepsilon}{2} \cdot \sum_j m(B_j) \geq \frac{\varepsilon}{10} m^*(E_{N,\varepsilon}).$$

令 $\delta \rightarrow 0$ 立即可得 $\lim_{N \rightarrow \infty} m^*(E_{N,\varepsilon}) = 0$, 从而根据 $m(E_\varepsilon) \leq m^*(E_{N,\varepsilon})$ 可得

$$m(E_\varepsilon) = 0, \quad \forall \varepsilon > 0.$$

所以对几乎处处的 x 有 $\overline{D}(J)(x) = 0$. 又因为函数单调增所以每个 Dini 函数均非负, 这推出每个 Dini 函数均为零. 所以 J 几乎处处可导且导数为零. \square

Proof for 2.3.7. 我们将完全照抄 2.1.3 的证明. 事实上, 证明思路在那题的注记里写的很清楚了.

设 B_α 的半径为 r_α , 我们 **贪心地** 取出一些球:

- 第一步取出半径至少为 $\frac{1}{2} \sup_{\alpha \in \mathcal{A}} r_\alpha$ 的球记为 B_1 , 把剩下的 **和 B_1 无交** 的球的集合记为 \mathcal{A}_1 .
- 第 $k-1$ 步之后得到的开球族 \mathcal{A}_{k-1} 是所有与 B_1, \dots, B_{k-1} 无交的开球构成的集合. 考虑选出一个半径至少为 $\frac{1}{2} \sup_{\alpha \in \mathcal{A}_{k-1}} r_\alpha$ 的开球记为 B_k , 并且把所有在 \mathcal{A}_{k-1} 中与 B_k 无交的球构成的集合记为 \mathcal{A}_k . 如果 $\mathcal{A}_k = \emptyset$ 则流程结束得到一个有限开球列, 否则得到一个可数开球列.
- 这一步和之前有所不同: 我们要单列出

$$\sum_j m(B_j) = +\infty$$

的情形, 此时原命题显然成立. 因此下面将考虑 $\sum_j m(B_j) < +\infty$ 的情形, 那么就有 $\lim_{j \rightarrow \infty} m(B_j) = 0$. 特别地, 根据选取规则可知也有

$$\lim_{j \rightarrow \infty} \sup_{\alpha \in \mathcal{A}_j} r_\alpha = 0.$$

因此对任意 δ , 存在 N 使得第 N 步后所有半径大于 δ 的球要么已经被选出, 要么和前面选出的球有交.

- 我们证明在 $\sum_j m(B_j) < +\infty$ 的条件下, 选出来的那些球球心不变, 半径扩大 5 倍时就可以覆盖所有球构成的集合. 即

$$\bigcup_{\alpha} B_\alpha \subset \bigcup_j (5B_j).$$

对任意 B_α , 根据上一步的讨论, 它要么被选出要么和已经被选出的某个球 B_j 有交. 如果是前者则已经成立, 如果是后者则由 B_j 的定义可知 $2r_j \geq r_\alpha$, 故 B_α 中的点到 B_j 球心的距离小于 $(r_j + r_\alpha) + r_\alpha \leq 5r_j$. 这意味着 $B_\alpha \subset 5B_j$, 证毕.

- 因此有 $m^*(\bigcup_\alpha B_\alpha) \leq \sum_j m(5B_j)$. 这直接导出了结论. \square

Remark. 应当仔细体会这个证明和有关极大函数的 Hardy-Littlewood 定理 2.1.2 的证明的关系. 虽然一个是积分一个是求导, 但是通过覆盖引理来完成控制, 通过上极限来刻画测度的方式是如出一辙的. 后面证明单调函数几乎处处可微也是用的一样的想法, 并且会有各种不同的控制的方法, 但本质就是对跃度进行控制. 它可以用来证明更有用的 Vitali 覆盖定理, 见 2.4.4.

Proposition 2.3.8. 设连续函数 F 在 $[a, b]$ 上单调增, 则

$$D^+(F)(x) < +\infty, \text{ a.e. } D^+(F)(x) \leq D_-(F)(x), \text{ a.e.}$$

这个命题初一看很怪, 我们先解释它有什么用: 首先我们可以通过这个命题得免费再得到另一个关系式:

$$D^-(F)(x) \leq D_+(F)(x). \quad (1)$$

这是因为我们可以把 $\bar{F}(x) = -F(-x)$ 代入结论, 得到

$$\begin{aligned} D^+(\bar{F})(x) &= \limsup_{h>0, h \rightarrow 0} \frac{\bar{F}(x+h) - \bar{F}(x)}{h} = \limsup_{h>0, h \rightarrow 0} \frac{F(-x-h) - F(-x)}{-h} \\ &= \limsup_{h<0, h \rightarrow 0} \frac{F(-x+h) - F(-x)}{h} = D^-(F)(-x). \\ D_-(\bar{F})(x) &= \liminf_{h<0, h \rightarrow 0} \frac{\bar{F}(x+h) - \bar{F}(x)}{h} = \liminf_{h<0, h \rightarrow 0} \frac{F(-x-h) - F(-x)}{-h} \\ &= \liminf_{h>0, h \rightarrow 0} \frac{F(-x+h) - F(-x)}{h} = D_+(F)(-x). \end{aligned}$$

再用 x 代替 $-x$ 即可得 (1). 再结合上显然的关系式

$$D_+(F)(x) \leq D^+(F)(x), \quad D_-(F)(x) \leq D^-(F)(x).$$

这可得到一个循环

$$D^+(F)(x) \leq D_-(F)(x) \leq D^-(F)(x) \leq D_+(F)(x) \leq D^+(F)(x).$$

因此四个导数相等, 它的推论就是

Corollary 2.3.9. 设 F 在 $[a, b]$ 上单调增, 则 F 几乎处处可导.

Proof for 2.3.8, part 1. 先证明 $D^+(F)(x) < +\infty$, a.e.. 记

$$E_\gamma = \{x \in [a, b] : D^+(F)(x) > \gamma\}.$$

我们先证明 $D^+(F)$ 是可测函数, 这样 E_γ 就是可测集, 就可以取极限. 首先把一族函数的上确界化归成一系列函数的上确界, 注意到对 $\frac{1}{n+1} \leq h \leq \frac{1}{n}$,

$$\frac{n}{n+1} \cdot \frac{F(x + \frac{1}{n+1}) - F(x)}{\frac{1}{n+1}} \leq \frac{F(x+h) - F(x)}{h} \leq \frac{n+1}{n} \cdot \frac{F(x + \frac{1}{n}) - F(x)}{\frac{1}{n}}.$$

所以实际上 $1/n$ 这一列数的极限就已经能够逼近 $D^+(F)(x)$:

$$D^+(F)(x) = \limsup_{h>0, h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \limsup_{n \rightarrow \infty} \frac{F(x + \frac{1}{n}) - F(x)}{\frac{1}{n}}.$$

化成一列函数的上极限后, 由 F 可测可知里面每个函数均可测, 那么 $D^+(F)$ 作为可测函数的上极限也是可测函数. 那么接下来就只需证明可测集 E_γ 的测度趋向于 0. 利用的方法与之前类似, 对 $x \in E_\gamma$, 存在足够小的 h 使得

$$F(x+h) - F(x) > \gamma h.$$

所以可以用一列 $I_{x,i} := (x - h_i, x + h_i)$ 来覆盖 x . 由无穷覆盖引理 2.3.7 可得存在互不相交的的开集列 $(x_j - h_j, x_j + h_j)$ 满足

$$m(E_\gamma) \leq 5 \cdot \sum_j m(x_j - h_j, x_j + h_j) = 10 \sum_j h_j.$$

另一方面, 又有

$$F(x_j + h_j) - F(x_j - h_j) > \gamma \cdot h_j;$$

所以

$$F(b) - F(a) \geq \sum_j (F(x_j + h_j) - F(x_j - h_j)) > \gamma \cdot \sum_j h_j \geq \frac{\gamma}{10} m(E_\gamma).$$

令 $\gamma \rightarrow \infty$ 可得

$$\lim_{\gamma \rightarrow \infty} m(E_\gamma) = 0.$$

这就证明了 $D^+(F)(x) < +\infty$, a.e..



Remark. 这个证明的一部分用到了“单调增函数的 Dini 导数可测”, 实际上在作业 48 中证明了任意可测函数的 Dini 导数也是可测函数.

最后介绍一个小引理 (数分 I 小练习), 可以用来完全解决这个问题.

Lemma 2.3.10 (Riesz 日出引理). 设 G 是 \mathbb{R} 上的实值连续函数, 定义

$$E = \{x : \exists h > 0 \text{ s.t. } G(x+h) > G(x)\}.$$

证明 E 是开集, 因此 E 可以写为可数个不交开区间的并

$$E = \bigcup_{k=1}^{\infty} (a_k, b_k).$$


证明如果 (a_k, b_k) 是有限长度区间, 则 $G(a_k) = G(b_k)$. 如果 G 定义在闭区间 $[a, b]$ 上, 则结论对 $[a, b]$ 作为 \mathbb{R} 上的子拓扑成立, 不过对于包含端点的区间, 仅有 “ \leq ” 关系成立 (这也导致后面应用 Sunrise Lemma 时得到的是不等号而不是等号, 并且还需要注意方向.)

Proof. 首先证明 E 为开集. 这是因为若 $x \in E$, 那么取 $y > x$ 使得 $G(y) > G(x)$, 这意味着存在 ε 使得 $G(y) > G(x) + \varepsilon$. 由于 G 为连续函数, 因此存在 δ 使得对任意 $z \in B_\delta(x)$ 均有 $G(z) \in B_\varepsilon(G(x))$, 所以 $G(z) < G(y)$, 因此 $B_\delta(x) \cap (-\infty, y) \subset E$, 而这是包含 x 的开集.

再证明后半部分: 由于 (a_k, b_k) 为一个满足要求的极大开区间, 因此 $a_k, b_k \notin E$, 导致 $G(b_k) \leq G(a_k)$. 假设 $G(b_k) < G(a_k)$, 则我们称 a_k 的右邻域不能全落在 E 中. 否则设 $G(b_k) < G(a_k) - \varepsilon$, 取 $\delta > 0$ 使得 $G(x + \delta) > G(b_k)$. 接下来考虑闭区间 $[x + \delta, b_k]$, 由于 G 连续故 G 在 $[x + \delta, b_k]$ 中存在极大值, 记为 $G(x_0)$, 由于 $x_0 \neq b_k$ 因此 $x_0 \in E$. 但是一方面对 $x \in (x_0, b]$ 总有 $G(x) \leq G(x_0)$, 另一方面对 $x > b$ 也有 $G(x) \leq G(b) \leq G(x_0)$, 这导致 $x_0 \notin E$, 矛盾. 综上引理成立.

如果 G 仅在闭区间 $[a, b]$ 上有定义, 将其延拓到 \mathbb{R} 上, 延拓方法为

$$G^*(x) = \begin{cases} -(x - a) + G(a), & x < a; \\ G(x), & a \leq x \leq b; \\ (x - b) + G(b), & x > b. \end{cases}$$

那么对 G^* 使用日出引理, 连续性仍然成立, 并且左右原本包含端点的区间一定延拓为了有限区间, 比如说 (c, d) 满足 $c < a < b$, 那么就有 $G(a) \leq G^*(c) = G^*(d)$, 右端点是同理的. 

Back to Proof for 2.3.8. 对连续函数 F , 我们考虑函数 $F(x) - \gamma x$, 则 $E_\gamma \subset E$. 这是英文由 $D^+(F)(x)$ 的定义, 取 $x \in E_\gamma$ 可知

$$\begin{aligned} \exists h > 0 \text{ s.t. } \frac{F(x+h) - F(x)}{h} &> \gamma. \\ \Rightarrow F(x+h) - \gamma(x+h) &> F(x) - \gamma x. \end{aligned}$$

设 $E = \bigcup_{k=1}^{\infty} (a_k, b_k)$, 那么就可以直接把 E_γ 放缩为 E , 得到

$$m(E_\gamma) \leq m(E) = \sum_{k=1}^{\infty} (b_k - a_k).$$

现在对 $F(x) - \gamma x$ 使用日出引理, 则有 $F(b_k) - \gamma b_k \geq F(a_k) - \gamma a_k$, 所以 $F(b_k) - F(a_k) \geq \gamma(b_k - a_k)$. 那么上面的式子可以进一步放缩:

$$m(E_\gamma) \leq m(E) = \sum_{k=1}^{\infty} (b_k - a_k) \leq \gamma^{-1} \sum_{k=1}^{\infty} (F(b_k) - F(a_k)) \leq \gamma^{-1} (F(b) - F(a)).$$

可以看到日出引理相比无穷覆盖引理给出了更好的一个界（之前是 $10\gamma^{-1}$ ）.

再证明

$$D^+(F)(x) \leq D_-(F)(x), \text{ a.e.}$$

对任意非负有理数 $r < R$, 考虑集合（这无疑是整个证明最有技巧性的一步, ysw 课上评价“这是整个实变中我认为最有技巧性的一个证明”）

$$E_{r,R} = \{x \in [a, b] : D^+(F)(x) > R, D_-(F)(x) < r\}.$$

如果 $E_{r,R}$ 测度非零, 那么可以考虑用一个区间来逼近它:（用区间逼近便于使用 Sunrise Lemma, 并且能够用好集合逼近坏集合是测度理论给我们最好的东西. 之后会发现, 本题证明中唯一的自由度来源于这里闭集逼近的选取）接下来将用两次日出引理来给出放缩, 两次日出引理将分别利用 $D^+(F)(x)$ 的下界和 $D_-(F)(x)$ 的上界. 把它们写开得到

$$\limsup_{h>0, h \rightarrow 0} \frac{F(x+h) - F(x)}{h} > R, \quad \liminf_{h>0, h \rightarrow 0} \frac{F(x-h) - F(x)}{-h} < r.$$

所以

$$\begin{aligned} \exists h_1, h_2 > 0 \text{ s.t. } \frac{F(x+h_1) - F(x)}{h_1} > R, \quad \frac{F(x) - F(x-h_2)}{h_2} < r. \\ \Rightarrow F(x+h_1) - R(x+h_1) > F(x) - Rx, \end{aligned} \quad (1)$$

$$F(x) - rx < F(x-h_2) - r(x-h_2). \quad (2)$$

这自然引到日出引理 2.3.10 的形式. 我们发现把日出引理中 E 的定义改成

$$E = \{x : \exists h > 0 \text{ s.t. } G(x-h) > G(x)\}$$

之后, 证明是完全一致的（即左侧的日出引理, 或者说考虑 $H(x) = G(-x)$ 然后在 $[-b, -a]$ 上对 H 使用在右侧的日出引理也可以得到相同的结论）. 因此我们在 $[c, d]$ 上对函数 $F(x) - rx$ 使用日出引理（这里用到了 F 连续. 如果 F 只是一般的单调函数这里就没有机会使用 Sunrise Lemma）得到开集 A 满足

$$A = \bigcup_{k=1}^{\infty} (a_k, b_k), \quad F(b_k) - F(a_k) \geq r(b_k - a_k).$$

再在每个 $[a_k, b_k]$ 上对函数 $F(x) - Rx$ 利用右日出引理, 得到开集 B_k 满足

$$B = \bigcup_{l=1}^{\infty} (a_{kl}, b_{kl}), \quad F(b_{kl}) - F(a_{kl}) \leq R(b_{kl} - a_{kl}).$$

那么就有 $E_{r,R} \subset A \cap B$, 所以可以据此给 $m(E_{r,R})$ 一个放缩:

$$\begin{aligned} m(E_{r,R}) &\leq \sum_{k,l=1}^{\infty} (b_{kl} - a_{kl}) \stackrel{\text{Sunrise B}}{\leq} \frac{1}{R} \sum_{k,l=1}^{\infty} (F(b_{kl}) - F(a_{kl})) \\ &\leq \frac{1}{R} \sum_{k=1}^{\infty} (F(b_k) - F(a_k)) \stackrel{\text{Sunrise A}}{\leq} \frac{r}{R} \sum_{k=1}^{\infty} (b_k - a_k) \leq \frac{r}{R} (d - c). \end{aligned}$$

然而我们 $d - c$ 的选取可以任意接近 $m(E_{r,R})$. 事实上, 由于 $r < R$ 所以可以取 $[c, d]$ 满足

$$E_{r,R} \subset [c, d] \subset [a, b], \quad d - c < \frac{R}{r} m(E_{r,R}).$$

这就和上式矛盾, 因此 $m(E_{r,R}) = 0$. 那么由于 $\{x : D^+(F)(x) > D_-(F)(x)\}$ 就可以写成可数个这样的 $E_{r,R}$ 之并, 所以也是零测集, 证毕. \square

Remark. 这个证明可以直接对一般的递增函数证明, 这是因为证明里利用到连续性的地方都可以被绕过: 我们可以用 Vitali 覆盖引理 2.4.4 来代替日出引理的使用, 而前者不需要连续性条件. 周民强书上有这个结论的证明方式.

最后我们回到单调函数 Lebesgue 定理的证明. 需要指出, 上面做的所有工作加起来证明了单调函数几乎处处可微性, 而下面只是最后的收尾工作.

Proof for 2.3.4. 首先把 F 延拓到 \mathbb{R} 上, 取 $F(x) = \begin{cases} F(a), & x < a; \\ F(x), & a \leq x \leq b; \\ F(b), & x > b. \end{cases}$ 记 $f =$

F' , 考虑逼近 f 的函数列

$$f_n(x) = n(F(x + \frac{1}{n}) - F(x)).$$

由于 F 几乎处处可微, 因此 $f_n \rightarrow f$, a.e. 而 $f_n(x)$ 为非负函数列, 利用 Fatou 引理 1.1.5 可得

$$\begin{aligned} \int_a^b \liminf_{n \rightarrow \infty} f_n \, dx &\leq \liminf_{n \rightarrow \infty} \int_a^b f_n \, dx \\ &= \liminf_{n \rightarrow \infty} \int_a^b n(F(x + \frac{1}{n}) - F(x)) \, dx \\ &= \liminf_{n \rightarrow \infty} \left(n \int_b^{b+\frac{1}{n}} F(x) \, dx - n \int_a^{a+\frac{1}{n}} F(x) \, dx \right) \\ &\quad (\text{根据延拓}) = F(b) - n \limsup_{n \rightarrow \infty} \int_a^{a+\frac{1}{n}} F(x) \, dx \\ &\leq F(b) - F(a). \end{aligned}$$

显然 $\liminf f_n = f$, 所以命题成立. \square

单调函数的 Lebesgue 定理有很多利用, 比如 Riemann 积分下的逐项微分定理就可以推广到单调函数的情形, 这相比一致收敛还是实用不少. (就正如 1.1 章中列举的 Lebesgue 积分带来的好处一样, 现在给出的就是 Lebesgue 微分理论建立之后带来的好处)

Proposition 2.3.11 (逐项微分). 设 $\{f_n\}$ 是 $[a, b]$ 上的递增函数列, 且 $\sum_{n=1}^{\infty} f_n(x)$ 在 $[a, b]$ 上收敛, 则

$$\left(\sum_{n=1}^{\infty} f_n(x) \right)' = \sum_{n=1}^{\infty} f_n'(x), \quad \text{a.e.}$$

Proof. 首先作有限拆分可得

$$\begin{aligned} \left(\sum_{n=1}^{\infty} f_n(x) \right)' &= \left(\sum_{n=1}^N f_n(x) + \sum_{n>N} f_n(x) \right)' \\ &= \sum_{n=1}^N f'_n(x) + \left(\sum_{n>N} f_n(x) \right)'. \end{aligned}$$

而

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N f'_n(x) = \sum_{n=1}^{\infty} f'_n(x),$$

因此只需证

$$\lim_{N \rightarrow \infty} \left(\sum_{n>N} f_n(x) \right)' = 0, \quad \text{a.e.}$$

设 $F_N := \sum_{n>N} f_n(x)$. 由于 f_n 是单调增函数, 故 f'_N 几乎处处取值非负. 又因为

$$F'_N(x) = f'_N(x) + F'_{N+1}(x) \geq F'_{N+1}(x), \quad \text{a.e.}$$

所以 F'_N 是单调降的非负函数列, 因此由 1.1.7 可知其极限和积分交换. 又联想到作业 1 证明非负函数几乎处处为零只需证明其积分为零, 所以只需证明

$$\lim_{N \rightarrow \infty} \int_a^b F'_N(x) = 0. \quad (1)$$

而对 F_N 使用刚刚证明的单调函数 Lebesgue 定理 2.3.4, 就可以直接得到

$$\int_a^b F'_N(x) \leq F_N(b) - F_N(a) \leq \sum_{n>N} f_n(b) - \sum_{n>N} f_n(a).$$

显然级数的收敛性可以保证 $N \rightarrow \infty$ 时 a, b 这两点处从 N 项起取值之和任意小, 所以 (1) 成立, 因此原命题成立. \square

2.4 绝对连续函数与 Newton-Leibniz 公式

现在可以着手回答问题 2.3.1.

Definition 2.4.1 (绝对连续函数). 称定义在 $[a, b]$ 上的函数 F 是绝对连续的, 如果 $\forall \varepsilon > 0, \exists \delta > 0$ 使得对任意两两不交的区间 $(a_k, b_k), k = 1, \dots, n$ 满足

$$\sum_{k=1}^n (b_k - a_k) < \delta,$$

都有

$$\sum_{k=1}^n |F(b_k) - F(a_k)| < \varepsilon.$$

(这可以用极限的方式推广到可数情形, 从而对所有开集都有类似式子成立.)

所有绝对连续函数组成的空间记为 $AC([a, b])$.

绝对连续函数和积分的绝对连续性 1.1.8 有关系. 如果 Newton-Leibniz 公式成立:

$$F(x) - F(a) = \int_a^x F' dy,$$

那么取一个测度足够小的集合 E , 那么积分 $\int_E F' dy$ 就对应了函数 F 在 E 上的变差, 这就引出了对函数的绝对连续性的定义, 同时也证明了如果 F 要满足 Newton-Leibniz 公式, 那么 F 一定是绝对连续函数.


另外, 绝对连续函数也可以看作是 Cantor-Lebesgue 函数的“对立面”: 后者是不满足 Newton-Leibniz 公式的, 所以好的函数需要规避掉这点. 然而, Cantor-Lebesgue 函数的可微性已经达到最好 (几乎处处为零), 所以只能是变差太大出了问题. 因此, 我们期待满足 Newton-Leibniz 公式的函数如果导数几乎处处为零, 那么它应该是常值函数. 实际上, 就可以证明如果一个绝对连续函数的导数几乎处处为零, 那么这个函数一定是常值函数.

Proposition 2.4.2. 如果 $f \in AC([a, b])$ 满足 $f' = 0$, a.e., 那么 f 为常值函数.

Thoughts. 由于 f' 几乎处处为零, 因此对几乎所有的 x , 存在 h_x 使得

$$\left| \frac{f(x+h) - f(x)}{h} \right| \leq \varepsilon, \quad h \in (-h_x, 0) \cup (0, h_x).$$

$$\Leftrightarrow |f(x+h) - f(x)| \leq \varepsilon h, \quad |h| < h_x.$$

接下来的想法依然是利用覆盖引理: 由于可以找到一些区间覆盖几乎所有的 x , 如果可以证明能够取出一个测度足够接近 $b-a$ 的好区间族 E , 那么这些区间族上 f 的变差可以随 ε 的变化而足够小. 假设结论不成立, 不妨设 $f(a) \neq f(b)$, 那么 f 在 $[a, b] \setminus E$ 上的变差可以足够接近 $|f(a) - f(b)|$, 这和绝对连续性矛盾. 然而, 这里需要的覆盖引理需要取出的 E 测度接近 $[a, b]$ 而不是仅仅能够控制它 (之前只需要控制它是因为我们基本上是在证明某个集合测度为零), 因此我们需要更强的覆盖引理, 同时对覆盖本身自然也有更高的条件要求. 下面的 Vitali 定理就给出了这样一种方法, 事实上之前的绝大部分问题 (使用无限覆盖引理或 Sunrise Lemma, 但不包括 Lebesgue 基本定理的证明) 的命题都可以替换成使用 Vitali 引理. 

Definition 2.4.3 (Vitali 覆盖). 我们称闭球族 $\mathcal{B} = \{B_\alpha\}$ 是集合 $E \subset \mathbb{R}^d$ 的 **Vitali 覆盖**, 如果对任意 $x \in E$ 及 $\delta > 0$, 存在闭球 $B_\alpha \in \mathcal{B}$ 使得

$$x \in B_\alpha, \quad m(B_\alpha) < \delta.$$

也就是说, 每个点均可被一个任意小的球盖住.

Lemma 2.4.4 (Vitali 覆盖定理). 如果 $E \subset \mathbb{R}^d$ 且 $m^*(E) < +\infty$, 那么对任意 E 的 Vitali 覆盖 \mathcal{B} 及 $\varepsilon > 0$, 存在有限个两两不交的球 $B_1, \dots, B_n \in \mathcal{B}$ 满足

$$m^*(E \setminus \bigcup_{i=1}^N B_i) < \varepsilon.$$

特别地, 存在可数列 $\{B_i\}$ 使得

$$m^*(E \setminus \bigcup_{i=1}^{\infty} B_i) = 0.$$

Proof. 证明将利用到无穷覆盖引理 2.3.7 的证法. 首先找一个开集 $A \supset E$ 使得

$$m(A) < m^*(E) + \delta.$$

(这一步, 连同题干内闭球的定义都是为了确保在后面取出有限个球后, 剩下的点构成开集, 所以存在每个点存在一个邻域和这些球都不交) 此外, 我们把半径大于 1 的球都去掉 (为了满足 2.3.7 中球的半径存在上界的条件), 再把不完全包含于 A 的球去掉 (为了规避掉引理中得到 $\sum m(B_k) = +\infty$ 的情况, 它的推论是 $m(B_k) \rightarrow 0$), 那么剩下的球仍然是一个 E 的 (而不是 A 的) Vitali 覆盖. 利用引理 2.3.7 中的算法可得

- 存在至多可数个互不相交的球 B_k 使得

$$r(B_k) > \frac{1}{2} \sup\{r(B) : B \in \mathcal{B}, B \cap B_j = \emptyset, j < k\},$$

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_{k=1}^{+\infty} 5B_k.$$

如果只取出有限个球 B_1, \dots, B_N , 我们证明

$$E \subset \bigcup_{k=1}^N B_k. \quad (2)$$

由选取规则可知 \mathcal{B} 中的每个球都和 B_1, \dots, B_N 有交. 假设存在 $x \in E$ 满足 $x \notin \bigcup_{k=1}^N B_k$, 那么由于 $A \setminus \bigcup_{k=1}^N B_k$ 为开集, 因此存在 x 的邻域 N 满足 $N \cap \bigcup_{k=1}^N B_k = \emptyset$. 另一方面, 存在足够小的 δ 使得任意包含 x 且半径小于 δ 的球都落在 N 中, 而存在 $B \in \mathcal{B}$ 满足 $x \in B$, $m(B) < \delta$, 这个球就和前面的球都不交, 矛盾! 因此 (2) 成立, 原命题成立.

下考虑最终取出了可数个球的情形, 它们满足

$$\sum_{k=1}^{\infty} m(B_k) \leq m(A) < +\infty.$$

我们证明只需要取适当的 N , B_1, \dots, B_N 就可以满足题目条件. 关键在于发现 B_{N+1}, \dots 可以控制第 N 步时剩下的那些点:

$$E \setminus \bigcup_{k=1}^N B_k \subset \bigcup_{k>N} 5B_k. \quad (1)$$

这是因为对任意 $x \in E \setminus \bigcup_{k=1}^N B_k$, 存在包含 x 的开球 $C \in \mathcal{B}$ 使得

$$C \subset A \setminus \bigcup_{k=1}^N B_k.$$

由于 $m(B_k) \rightarrow 0$, 因此

$$\sup\{r(B) : B \in \mathcal{B}, B \cap B_j = \emptyset, j < k\} \rightarrow 0,$$

所以 C 要么在第 N 步后被取出, 要么和第 N 步后的某一个取出的球有交. 无论如何, 都能用同样的方法证明 (1) 成立. 由 $\sum m(B_k) < +\infty$, 可以取 N 使得

$$\sum_{k>N} B_k < \frac{\varepsilon}{5^d} \Rightarrow m(E \setminus \bigcup_{k=1}^N B_k) \leq \varepsilon.$$



Remark. 特别地, 对任何包含 E 的开集 A , 我们还可以保证取出的 B_k 满足

$$\sum_{k=1}^N m(B_k) \leq m(A) < m^*(E) + \varepsilon.$$

其次, 证明过程自然保证取出上面的一整列 B_k 后 $E \setminus \bigcup_{i=1}^{\infty} B_i$ 测度为零.

Back to Proof for 2.4.2. 全体

$$[x - h, x + h], \quad h < h_x, x \in E$$

构成了 $E = \{x \in [a, b] : f'(x) = 0\}$ 的一个 Vitali 覆盖. 由 Vitali 覆盖引理 2.4.4, 取出有限个互不相交的区间 $[a_k, a_k]$ 使得

$$m([a, b] \setminus \bigcup_{k=1}^N [a_k, b_k]) = m^*(E \setminus \bigcup_{k=1}^N [a_k, b_k]) < \delta.$$

由于这是有限个区间, 因此可以不妨考虑一个序关系 (这也是 \mathbb{R} 的特性)

$$b_0 := a \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_N < b_N \leq b =: a_{N+1}.$$

那么剩下的区间长度就满足

$$\sum_{k=0}^N (a_{k+1} - b_k) < \delta.$$

一方面, 根据覆盖的性质可知

$$\sum_{k=1}^N |f(b_k) - f(a_k)| \leq \sum_{k=1}^N \varepsilon \cdot (b_k - a_k) \leq \varepsilon \cdot (b - a).$$

另一方面, 由绝对连续性可知对任意 $\lambda > 0$, 当 δ 足够小时总有

$$\sum_{k=0}^N |f(a_{k+1}) - f(b_k)| < \lambda.$$

$$|f(b) - f(a)| \leq \sum_{k=1}^N |f(b_k) - f(a_k)| + \sum_{k=1}^N |f(b_k) - f(a_k)| \leq \varepsilon(b - a) + \lambda.$$

令 $\delta \rightarrow 0$, $\varepsilon \rightarrow 0$, 这就得到 $f(b) = f(a)$. 对 f 中的每个闭子区间利用结论, 就可以证明 f 为常值函数.



接下来考虑绝对连续函数的一些充分条件和必要条件.

Example 2.4.5. 我们考虑绝对连续函数的一些充分条件和必要条件.

- C^1 连续函数是绝对连续的. 这就是在 Riemann 积分下证明 Newton-Leibniz 公式所需的条件.
- Lipschitz 函数是绝对连续的, 它满足

$$|f(y) - f(x)| \leq M \cdot |x - y|.$$

因此它也被称为强绝对连续函数.

- 绝对连续函数是一致连续函数, 令 $n = 1$, 只分割为一段即可.
- 绝对连续函数是有界变差函数. 比如说取 $\varepsilon = 1$ 得到 δ , 那么我们可以把 $[a, b]$ 分割为长度小于 δ 的一些区间, 设该分割为 τ . 对于任何的分割 σ , 考虑 $\sigma \cup \tau$ 分割后变差不减. 然而对于 τ 中的每个小区间, 可以一同取出所有它细分得到的区间, 这些区间长度之和小于 δ , 所以它们整体的变差永远小于 1, 在 $[a, b]$ 上的全变差就不会超过分割 τ 产生的变差, 这一定是一个有限值.

Remark. $BV([a, b]) \subset AC([a, b])$ 是极其有用的信息. 我们刚刚研究完单调函数是几乎处处可微的, 而有界变差函数能写成两个单调函数的差, 所以也是几乎处处可微的! 这就是说我们几乎可以对绝对连续函数用上之前讨论的所有性质 (就像讨论可测函数积分时进行正负分离一样).

- 绝对连续函数把零测集映到零测集. 这是因为可以取测度任意小的开集逼近一个零测集, 然后把开集写成可数个区间的并, 每个区间上存在最小值和最大值 (可能是端点), 把最大值和最小值所对应的子区间取出来代替原区间, 那么得到的 $\sum |F(b_i) - F(a_i)|$ 就可以逼近值域 $F(E)$, 这个值可以任意小, 因此为零. 由 F 的连续性可知 F 将可测集映为可测集 (之前的某次作业).
- $AC([a, b])$ 是一个代数.
- $AC([a, b])$ 在范数

$$\|F\|_{AC} = \sup |F| + \int |F'|$$

下构成完备的 Banach 空间.

利用命题 2.4.2, 实际上也已经足够证明微积分基本定理就和绝对连续性是等价的, 至此完全回答了 2.3.1.

Theorem 2.4.6 (Lebesgue 微积分基本定理). 如果 $F \in AC([a, b])$, 那么 F 几乎处处可微, 且 $F' \in L^1([a, b])$ 满足

$$F(x) - F(a) = \int_a^x F'(y) dy.$$

Proof. 首先 F 几乎处处可微来源于 F 可以写成两个单调函数的差. 其次, 单调函数 Lebesgue 微分定理 2.3.4 告诉我们 F' 是可积的 (如果设 $F = G - H$, 那么 $\int F'$ 可以被 $|G(b) - G(a)| + |H(b) - H(a)|$ 控制). 最后根据 Lebesgue 基本定理 2.1.1, 函数

$$G(x) = \int_a^x F'(y) dy$$

几乎处处满足 $G'(x) = F'(x)$. 所以 $(G - F)'(x) = 0$, a.e. 而 F, G 都绝对连续, 所以 $G - F$ 也绝对连续, 由 2.4.2 可知 $G - F$ 为常值函数. 代入 $x = a$ 即得结论. \square

如果微积分基本定理成立, 那么也有分部积分成立:

Theorem 2.4.7 (分部积分). 对 $F, G \in AC([a, b])$, 有分部积分成立

$$\int_a^b FG' dx + \int_a^b F'G dx = G(b)F(b) - G(a)F(a).$$

最后来回忆 Riemann 积分下的微积分基本定理: 条件是 F 在 $[a, b]$ 上处处可微, 且导函数 F' Riemann 可积, 则 Newton-Leibniz 公式成立. 现在可以证明, F' 只需要是 Lebesgue 可积的, 这大大推广了该命题.

Theorem 2.4.8. 若 F 在 $[a, b]$ 上可微, 且 F' Lebesgue 可积, 则

$$F(b) - F(a) = \int_a^b F'(x) dx.$$

Thoughts. 只需证明 $F \in AC([a, b])$. 这需要我们证明 F 的可微性可以建立对 F 的变差的控制, 而 F' 的可积性告诉我们这个控制的来源应该是 $\int_E |F'|$. 实际上确实有这样的性质成立. \square

Proposition 2.4.9. 设 F 为 $[a, b]$ 上的实值函数. 则对任意 $E \subset [a, b]$ (不一定可测), 如果 F 在 E 上可微, 且 $\sup_{x \in E} |F'(x)| \leq M$; 那么

$$m^*(F(E)) \leq M \cdot m^*(E).$$

Proof. 考虑 E 的 Vitali 覆盖从而控制 E 的外测度. 想法还是和证明 2.4.2 时一样, 只不过把 0 换成了 M . 对任意 $x \in E$, $\varepsilon > 0$, 由 $|F'|$ 的上界可知存在 $\delta_x > 0$ 使得

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq M + \varepsilon, \quad \forall 0 < |h| < \delta_x.$$

那么所有

$$[x-h, x+h], \quad 0 < h < \delta_x, x \in E$$

就构成了 E 的 Vitali 覆盖. 由覆盖定理 2.4.4 的可数列版本可知存在可数个互不相交的区间 $I_i = [x_i - h_i, x_i + h_i]$ 满足

$$m^*(E \setminus \bigcup_{i=1}^{\infty} I_i) = 0, \quad \sum_{i=1}^{+\infty} |I_i| \leq m^*(E) + \varepsilon,$$

只需取一个包含 E 的开集 B 满足 $m^*(E) > m(B) - \varepsilon$, 然后只看那些包含于 B 的集合, 它们也构成一个 Vitali 覆盖. 那么在每个 I_i 上都有

$$\begin{aligned} m^*(F(I_i)) &\leq \sup_{y_i, z_i \in I_i} |F(y_i) - F(z_i)| \leq \sup_{y_i, z_i \in I_i} (|F(y_i) - F(x_i)| + |F(z_i) - F(x_i)|) \\ &\leq \sup_{y_i, z_i \in I_i} (h_i(M + \varepsilon) + h_i(M + \varepsilon)) \\ &= (M + \varepsilon) |I_i|. \end{aligned}$$

所以可以把这些 I_i 进行累和:

$$\sum_{i=1}^{\infty} m^*(F(I_i)) \leq (M + \varepsilon) \sum_{i=1}^{\infty} |I_i| \leq (M + \varepsilon)(m^*(E) + \varepsilon).$$

另一方面, $F(E)$ 也可以被分解为

$$\begin{aligned} F(E) &\subset \bigcup_{i=1}^{\infty} F(I_i) \cup F(E \setminus \bigcup_{i=1}^{\infty} I_i) \\ \Rightarrow m^*(F(E)) &\leq \sum_{i=1}^{\infty} m^*(F(I_i)) + m^*(F(E \setminus \bigcup_{i=1}^{\infty} I_i)) \\ &\leq (M + \varepsilon)(m^*(E) + \varepsilon) + m^*(F(E \setminus \bigcup_{i=1}^{\infty} I_i)). \end{aligned}$$

因此问题转化为零测集的情形. 如果能证明对任意零测集 E_0 , 均有 $m(F(E_0)) = 0$, 那么由于 $E \setminus \bigcup_{i=1}^{\infty} I_i$ 为零测集, 由上式可得到

$$m^*(F(E)) \leq (M + \varepsilon)(m^* + \varepsilon), \quad \forall \varepsilon > 0.$$

令 $\varepsilon \rightarrow 0$ 就得到结论. 当 E_0 为零测集时之前的方法不再奏效, 因为用 $\bigcup I_i$ 来在几乎处处的意义下逼近 E 没有意义.

新的想法是把 E 写成一列集合的极限: 仍然固定 $\varepsilon > 0$, 考虑

$$E_n = \left\{ x \in E : |F(x+h) - F(x)| \leq (M + \varepsilon)h, \forall h \in [0, \frac{1}{n}] \right\}.$$

(我们讨论一下这里 $x \in E$ 的限制, 可能的一个想法是我们想把 E 写成单调列的极限, 而且其实 E_n 和 ε 有关, 最后还要令 $\varepsilon \rightarrow 0$, 所以只有当 $x \in E$ 时才对所有 ε 满足要求. 当然, 如果去掉这一条后把它定义为 $E_{n,\varepsilon}$ 最后令 $n \rightarrow \infty, \varepsilon \rightarrow 0$ 应该还是能得到 E .) 则

$$\lim_{n \rightarrow \infty} E_n = \{x \in E : \exists \delta_x > 0 \text{ s.t. } |F(x+h) - F(x)| \leq (M + \varepsilon)h, \forall h \in [0, \delta_x]\} = E.$$

因此 $\lim_{n \rightarrow \infty} F(E_n) = F(E)$. 取开集 $I_n \supset E_n$ 使得 $m(I_n) \leq m^*(E_n) + \varepsilon$, 可以把 I_n 分解为开区间列 $I_n = \bigcup_{k=1}^{\infty} I_{n,k}$. 考虑每个 $I_{n,k}$, 我们只需要其中在 E_n 中的点的贡献, 即考虑 $I_{n,k} \cap E_n$. 我们发现只要 $|I_{n,k}| < \frac{1}{n}$, 那么就有

$$|F(y) - F(z)| \leq (M + \varepsilon) |I_{n,k}|, \quad \forall y, z \in I_{n,k} \cap E_n.$$

(这一步是整个技巧展现出来的地方: 如果只对 $E \cap I_{n,k}$ 考虑, 那么每个 y, z 对应的 δ_y, δ_z 都可以任意小, 所以 $F(y)$ 和 $F(z)$ 可以没有关系. 而一维情形下开集可以规整地分解使得所有 $|I_{n,k}|$ 累和后恰好就是 $|I_n|$ 没有任何损失, 可以说是把技巧用到极致了.) 因此可以适当细分 I_n 使得每个子区间的长度都小于 $\frac{1}{n}$, 那么就可以得到对任意 k 均有

$$m^*(F(E_n \cap I_{n,k})) \leq \sup_{y, z \in E_n \cap I_{n,k}} |F(y) - F(z)| \leq (M + \varepsilon) |I_{n,k}|.$$

(注意这里 $E_n \cap I_{n,k}$ 可以长的很奇怪, 但是它们的值域确实可以给到一个限制, 直接看 $F(I_{n,k})$ 是不行的.) 把它们累和

$$\begin{aligned} m^*(F(E_n)) &= m^*(F(E_n \cap I_n)) \leq \sum_{k=1}^{+\infty} m^*(F(E_n \cap I_{n,k})) \\ &\leq (M + \varepsilon) \sum_{k=1}^{\infty} |I_{n,k}| \leq (M + \varepsilon) \cdot (m^*(E_n) + \varepsilon). \end{aligned}$$

两边取极限 (可以通过取等测包的技巧证明对单调升集合列外测度和极限也可交换) 可得

$$\begin{aligned} m^*(F(E)) &= \lim_{n \rightarrow \infty} m^*(F(E_n)) \leq (M + \varepsilon) \cdot (\lim_{n \rightarrow \infty} m^*(E_n) + \varepsilon) \\ &= (M + \varepsilon) \cdot (m^*(E) + \varepsilon). \end{aligned}$$

上面的表达式已经和 E_n 无关, 此时令 $\varepsilon \rightarrow 0$ 可知

$$m^*(F(E)) \leq M \cdot m^*(E).$$

因此原命题成立. \square

Remark. YSW 课上讲的版本实际上是对零测集证明了后半命题, 但是实际上稍加修改就可以得到对一般可测集可能这么做. 坦白来说整个证明是我见过最奇怪的证明之一了, 好像什么都没做, 只是用开集逼近任意集合再把开集拆开就起到了覆盖引理的效果并得到了一个很强的结论. 整个结论实际上也可以用在之前的一些命题上, 比如可以证明 $f \in AC, f' = 0, \text{a.e.} \Rightarrow f = c$ 时, 把导数为零的集合记为 E 就可测 $f(E)$ 零测, 而绝对连续函数又把零测集映到零测集, 所以像集零测, 连续函数又说明像集零测就一定是单点集. 我们发现实际上我们只用到了“ f 连续, f 把零测集映到零测集”, 如果还有“ f 几乎处处可微”(或者它的一个充分条件, $f \in BV$), 那么下面的推论 2.4.10 就立即告诉我们 f 绝对连续. 注意 Cantor-Lebesgue 函数就不满足“把零测集映到零测集”, 所以某种意义上这条性质才是绝对连续函数的关键. 完整的证明叙述见习题 65.

Corollary 2.4.10 (Growth Lemma). 若 f 为 $[a, b]$ 上的可测函数, 对任意 $E \subset [a, b]$, 如果 f 在 E 上可微, 则

$$m^*(f(E)) \leq \int_E |f'(x)| dx.$$

Proof. 任给 $\varepsilon > 0$, 设

$$E_n = \{x \in E : (n-1)\varepsilon \leq |f'(x)| < n\varepsilon\}.$$

则根据 2.4.9 可知

$$m^*(f(E_n)) \leq n\varepsilon \cdot m^*(E_n) \leq \varepsilon \cdot m^*(E_n) + \int_{E_n} |f'(x)| dx.$$

对 n 累和可知

$$m^*(f(E)) \leq \sum_{n=1}^{\infty} m^*(f(E_n)) \leq \varepsilon \cdot m(E) + \int_E |f'(x)| dx.$$

令 $\varepsilon \rightarrow 0$ 即得结论. Ⓜ

Back to Proof for 2.4.8. 我们证明 $F \in AC([a, b])$. 由推论 2.4.10 可知对任意 $E \subset [a, b]$ 均有

$$m^*(f(E)) \leq \int_E |f'(x)| dx.$$

由于 $f' \in L^1$, 因此利用积分的绝对连续性就可得 f 的绝对连续性. Ⓜ

Theorem 2.4.11 (Lebesgue 分解). 设 F 在 $[a, b]$ 上单调增, 则 F 可以进行分解:

$$F = F_{AC} + F_C + J.$$

其中 F_{AC} 为绝对连续函数, F_C 为单调增的连续函数且导数几乎处处为零, J 为跳跃函数.

Proof. 见习题 66. 实际上有用的信息已经在 2.3.5 中分析的差不多了, 后面就是人为构造一个绝对连续函数. Ⓜ

我们知道对有界变差的可微函数, 在 Riemann 积分下有下面的命题成立:

$$T_F([a, b]) = \int_a^b |F'(x)| dx.$$

我们可以把命题推广到 Lebesgue 积分下绝对连续函数的情形, 进一步地我们可以把命题推广到向量值函数 $\gamma(t) : [0, 1] \rightarrow \mathbb{R}^d$, 结论就转化为可求长性质.

Proposition 2.4.12. 如果 $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ 的每个分量均是绝对连续函数, 则 $\gamma(t)$ 是可求长曲线, 且

$$\text{Length}(\gamma) = \int_0^1 |\gamma'(t)| dt.$$

Corollary 2.4.13. 如果 F 是绝对连续函数, 则全变差满足

$$T_F([a, b]) = \int_a^b |F'(x)| dx.$$

Proof. 证明一边是简单的:

$$|\gamma(t_i) - \gamma(t_{i-1})| = \left| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right| \leq \int_{t_{i-1}}^{t_i} |\gamma'(t)| dt,$$

因此

$$\begin{aligned} \sup_{a=t_0 < t_1 < \dots < t_n = b} \sum_i |\gamma(t_i) - \gamma(t_{i-1})| &\leq \sup_{a=t_0 < t_1 < \dots < t_n = b} \sum_i \int_{t_{i-1}}^{t_i} |\gamma'(t)| dt \\ &= \int_a^b |\gamma'(t)| dt. \end{aligned}$$

要证明反向的不等号也成立, 需要利用逼近的想法. 由于阶梯函数在 L^1 上稠密 1.2.3, 取向量值阶梯函数 g 使得

$$\int_a^b |\gamma' - g| dt < \varepsilon.$$

记 $\gamma' = g + h$, 定义

$$G(t) = \int_a^t g(s) ds, \quad H(t) = \int_a^t h(s) ds.$$

则由 γ, G, H 均为绝对连续函数, 因此 $(G+H-\gamma)' = 0$, a.e. 可推出 $\gamma = G+H+\gamma(a)$. 于是接下来仔细分析 G : 由于 g 为阶梯函数, 因此 G 应当是分段线性函数, 自然有

$$\text{Length}(G) = \int_a^b |g(t)| dt.$$

那么我们只需证明 $\text{Length}(\gamma)$ 和 $\text{Length}(G)$ 足够接近, $\int_a^b |g(t)| dt$ 和 $\int_a^b |\gamma'(t)| dt$ 足够接近. 事实上根据 $\gamma = G + H + \gamma(a)$, 由定义展开可以证明

$$\text{Length}(\gamma) \geq \text{Length}(G) - \text{Length}(H) \geq \text{Length}(G) - \|h\|_{L^1}.$$

再根据 L^1 范数的性质, 可得

$$\|g\|_{L^1} \geq \|\gamma'\|_{L^1} - \|h\|_{L^1}.$$

两者结合可知

$$\text{Length}(\gamma) \geq \text{Length}(G) - \varepsilon = \|g\|_{L^1} - \varepsilon \geq \|\gamma'\|_{L^1} - 2\varepsilon.$$

证毕. 

Remark. 一些比较细致的讨论:

1. 如果 F 不是绝对连续而只是有界变差的, 那么进行正负变差分解后可以得到

$$T_F([a, b]) \geq \int_a^b |F(t)| dt.$$

并且如果取到等号, 那么 F 就是绝对连续的 (见习题 60. 多元的情形问题比较复杂, 因为不能逐分量地考虑正负性).

2. 还有一个很强的结论：虽然对一般的有界变差函数上面的等号无法取到，但它们的导数却几乎处处相等（想象一下就是 Lebesgue 分解中“Cantor-Lebesgue 项”出现的问题，但它只能影响一个零测集），即

$$T'_F([a, x]) = |F'(t)|, \quad \text{a.e.}$$

证明见习题。习题也告诉我们由有界变差函数得到绝对连续函数的过程实际上是不停把函数“掰成”单调函数的过程。

最后我们考虑换元公式：

Proposition 2.4.14 (Lebesgue 换元公式). 如果 $\varphi : [a, b] \rightarrow [c, d]$ 严格增且绝对连续，故对 $[c, d]$ 上的可积函数 f ，总有

$$\int_c^d f(y) \, dy = \int_a^b f(\varphi(x)) \varphi'(x) \, dx.$$

Proof.



3 多元微分论

我们主要的目标是两个方向：

- 高维的 Newton-Leibniz 公式，即 Stokes 公式。根据一维时的内容，我们得到的结果应当是把区域上一个函数 F 的积分去对应到这个区域的“边界”上去。如何去定义边界上的积分的很重要的问题。
- 隐函数定理。

3.1 关于微分

Definition 3.1.1 (方向导数). 给定连通的开区域 $\Omega \subset \mathbb{R}^n$ 。考虑函数 $f : \Omega \rightarrow \mathbb{R}$ ，对 $x_0 \in \Omega$ ， $\boldsymbol{v} \in \mathbb{R}^n$ ，如果极限

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h\boldsymbol{v}) - f(x_0)}{h}$$

存在，则称 f 在 x_0 处沿 \boldsymbol{v} 的方向导数存在，记作

$$\frac{\partial f}{\partial \boldsymbol{v}} = (\nabla_{\boldsymbol{v}} f)(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h\boldsymbol{v}) - f(x_0)}{h}.$$

对于标准坐标系， \boldsymbol{e}_i 方向的方向导数记为

$$\nabla_{\boldsymbol{e}_i} f(x_0) = \frac{\partial f}{\partial x_i}(x_0)$$

这也被称为 f 在 x_i 方向上的偏导数。

我们发现偏导数的定义是和坐标系选取有关的, 但是从几何上的观点来看, x 的附近的所有方向导数都是地位相同的. 应当有一个量记录函数 f 在 x 附近的所有变化.

Definition 3.1.2 (微分). $f: \Omega \rightarrow \mathbb{R}, x_0 \in \Omega$. 如果存在线性映射 $A: \mathbb{R}^n \rightarrow \mathbb{R}$ 使得对任意 $\mathbf{v} \rightarrow 0, \mathbf{v} \in \mathbb{R}^n$ 均有

$$\begin{aligned} f(x_0 + \mathbf{v}) &= f(x_0) + A\mathbf{v} + o(|\mathbf{v}|). \\ \Leftrightarrow \lim_{|\mathbf{v}| \rightarrow 0} \frac{|f(x_0 + \mathbf{v}) - f(x_0) - A\mathbf{v}|}{|\mathbf{v}|} &= 0. \end{aligned}$$

其中 $|\mathbf{v}|$ 是标准欧氏范数. 称 f 在 x_0 处可微, 并且称线性映射 A 是 f 在 x_0 处的微分

$$A = df|_{x=x_0} = df(x_0).$$

如果 f 在 Ω 的每个点处均可微, 则称 f 在 Ω 上可微.

Proposition 3.1.3. 若 f 在 x_0 处可微, 则 f 在 x_0 处的任意方向导数均存在. 在局部坐标系下,

$$df(x_0)(\mathbf{v}) = (\nabla_{\mathbf{v}} f)(x_0) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x_0) \mathbf{v}_j = (\nabla f)(x_0) \cdot \mathbf{v}.$$

其中 ∇f 是一个向量场, 在选定一组坐标系后, 它可表示为

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right).$$

所以最后一项实际上是 \mathbb{R}^n 中两个向量的内积.

Remark. 严格地说, df 是微分形式 (余切向量场), ∇f 是向量场, 所以 df 作用在一个向量场上, ∇f 和一个向量场作内积. 所以它们从某种意义上是对偶算子. 某种意义上, ∇f 更像一维情形下导数的直接推广. 在上述的各种记号中, 个人还是最喜欢使用 $df|_{x=x_0}$. 它把所有参数都写的很清楚, 如果写它作用在 x_0 再作用在 \mathbf{v} 上显得有些复杂.

Proof. 若 f 可微, 则存在 A 使得

$$f(x_0 + h\mathbf{v}) = f(x_0) + Ah\mathbf{v} + o(|h\mathbf{v}|), \quad \forall h \rightarrow 0.$$

那么

$$(\nabla_{\mathbf{v}} f)(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h\mathbf{v}) - f(x_0)}{h} = A\mathbf{v} + \lim_{h \rightarrow 0} \frac{o(|h\mathbf{v}|)}{h} = A\mathbf{v}.$$

证毕.



另一方面, 所有方向导数存在不一定能推出 f 可微. 我们有如下的例子:

Example 3.1.4.

$$f(x, y) = \begin{cases} \frac{y^2}{x}, & x \neq 0, \\ 0, & x = 0 \end{cases}.$$

f 在 $(0, 0)$ 处的任意方向导数存在, 但 f 在 $(0, 0)$ 处不可微, 因为它的取值不能构成一个线性映射.

然而我们还是想通过方向导数 (或者更好地, 偏导数) 存在来推出微分存在. 有一个充分条件可以来干成这件事情.

Proposition 3.1.5. 若 f 的所有偏导数 $\frac{\partial f}{\partial x_j}$ 均在其邻域 Ω 上存在且在 x_0 附近连续, 那么 f 在 x_0 处可微.

Proof. 在选定坐标轴后, n 个方向上的信息 $df|_{x=x_0}(\mathbf{v}_i) = \frac{\partial f}{\partial \mathbf{x}_i}$ 已经足以通过线性确定整个线性映射 $df|_{x=x_0}$ (或者梯度 $(\nabla f)(x_0)$). 此时我们需要考虑

$$f(x_0 + \mathbf{v}) - f(x_0) - (\nabla f)(x_0) \cdot \mathbf{v} \quad (df|_{x=x_0}(\mathbf{v})).$$

而我们只知道沿坐标轴方向的偏导数的信息, 因此我们要从 x_0 到达 $x_0 + \mathbf{v}$, 只需要通过平行于坐标轴的折线到达它即可. 由此, 设 $\mathbf{v} = (v_1, v_2, \dots, v_n)$, 则可以设

$$\mathbf{u}_j = (v_1, \dots, v_j, 0, \dots, 0).$$

把上面的式子作拆分可得:

$$\begin{aligned} & f(x_0 + \mathbf{v}) - f(x_0) - (\nabla f)(x_0) \cdot \mathbf{v} \\ &= \sum_{j=1}^n \left(f(x_0 + \mathbf{u}_j) - f(x_0 + \mathbf{u}_{j-1}) - \frac{\partial f}{\partial \mathbf{x}_j}(x_0) \mathbf{v}_j \right) \\ &= \sum_{j=1}^n \left(\frac{\partial f}{\partial \mathbf{x}_j}(x_0 + \tilde{\mathbf{u}}_j) \mathbf{v}_j - \frac{\partial f}{\partial \mathbf{x}_j}(x_0) \mathbf{v}_j \right). \end{aligned}$$

其中最后一步事实上是对一元函数 $g(\lambda) = f(x_0 + \mathbf{u}_{j-1} + \lambda \mathbf{v}_j)$ 使用 Lagrange 中值定理得到满足要求的 $\tilde{\mathbf{u}}_j$ 在 \mathbf{u}_j 和 \mathbf{u}_{j-1} 之间的线段上. 然后再利用偏导数的连续性可知当 $|\mathbf{v}| \rightarrow 0$ 时的确有每个分量的系数都 $\rightarrow 0$, 因此

$$f(x_0 + \mathbf{v}) - f(x_0) - (\nabla f)(x_0) \cdot \mathbf{v} \rightarrow 0.$$

证毕.



Remark. 我们还有一种想法是利用积分. 虽然它没法用于完全解决这个问题, 但是仍然具有参考价值. 简单来说, 如果 f 的正则性足够好, 那么 $f(x_0 + \mathbf{v}) - f(x_0)$ 可以被 Newton-Leibniz 公式展开为

$$\begin{aligned} f(x_0 + \mathbf{v}) - f(x_0) &= \int_0^1 \frac{d}{dt} f(x_0 + t\mathbf{v}) dt \\ &= \int_0^1 \frac{\partial f}{\partial \mathbf{v}}(x_0 + t\mathbf{v}) \mathbf{v} dt \\ &= \int_0^1 \frac{\partial f}{\partial \mathbf{v}}(x_0 + t\mathbf{v}) dt \cdot \mathbf{v}. \end{aligned}$$

这就把问题化归到了一维的情形. 这样的想法在后面 3.2.4 中有利用.

一维的导数还有一件事可以推广:

Proposition 3.1.6. 若 $f: \Omega \rightarrow \mathbb{R}$ 在 x_0 处可微, 且在 x_0 在其邻域内取到最值, 则 $df(x_0) = 0$.

接下来对任意的映射定义微分. 事实上, 在一元微分学的时候就已经可以定义向量值函数, 只需要逐分量定义即可. 所以这一步扩展实际上并不费劲, 并且我们会看到它们在矩阵的表达中的确有一定的对偶性.

Definition 3.1.7 (映射的微分). 设 $\Omega \subset \mathbb{R}^n$, $\Omega' \subset \mathbb{R}^m$, 对映射 $f: \Omega \rightarrow \Omega'$ 如果存在线性映射

$$df|_{x=x_0} = df(x_0): \mathbb{R}^n \rightarrow \mathbb{R}^m$$

使得当 $|v| \rightarrow 0$ 时有

$$\begin{aligned} f(x_0 + v) &= f(x_0) + df|_{x=x_0}v + o(|v|). \\ \Leftrightarrow \lim_{|v| \rightarrow 0} \frac{|f(x_0 + v) - f(x_0) - df|_{x=x_0}v|}{|v|} &= 0. \end{aligned}$$

则称 f 在 x_0 处可微, 并且称线性映射 $df|_{x=x_0}$ 是 f 在 x_0 处的微分. 如果 f 在 Ω 上每点处均可微, 则称 f 为 Ω 上的可微映射.

根据其定义方式可知, 一般的 Banach 空间 (有范数, 有线性) 也可定义微分的概念 (实际上只需要让一些定义有意义即可).

Proposition 3.1.8. 对映射 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 设 $f(x) = (f_1(x), \dots, f_m(x))$. 那么 f 在 x_0 处可微当且仅当每个 $f_j(x)$ 均在 x_0 处可微.

特别地, 在选定一个坐标系后, df 唯一确定一个 $m \times n$ 矩阵

$$\left(\frac{\partial f_j}{\partial x_i} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

其中每个分量均是一个 $\mathbb{R} \rightarrow \mathbb{R}$ 的函数. 在选定取值 x_0 后, 这个矩阵就是线性映射 $df|_{x=x_0}$ 在这组基下对应的矩阵. 我们把这个矩阵成为 f 在 x_0 处的 **Jacobi 矩阵**, 记为 $\text{Jac}(f)$ 或 $J(f)$.

Definition 3.1.9 (k 阶可导). 设 f 是一个 $\mathbb{R}^n \rightarrow \mathbb{R}^m$ 的映射, k 次可导指的是所有 $\leq k$ 次的高阶导数分量

$$\frac{\partial^l f}{(\partial x_1)^{l_1} \cdots (\partial x_n)^{l_n}}, \quad l \leq k, l_1 + \cdots + l_n = l.$$

均存在. 我们记 $C^k(\Omega, \mathbb{R}^m)$ 是从 Ω 到 \mathbb{R}^m 的 k 次可导且导数连续的映射组成的空间. 这个记号实际上用到了把偏导数看作偏导算子 $\frac{\partial}{\partial x_i}$ 的想法, 它把函数空间打到函数空间.

Proposition 3.1.10 (链式法则). 设 $\Omega_j \subset \mathbb{R}^{m_j}$ 为开集, $j = 1, 2, 3$. 设 $f: \Omega_1 \rightarrow \Omega_2$, $g: \Omega_2 \rightarrow \Omega_3$. 设 f 在 $\mathbf{x}_1 \in \Omega_1$ 处可微, g 在 $\mathbf{x}_2 = f(\mathbf{x}_1) \in \Omega_2$ 处可微, 则复合映射 $g \circ f$ 在 \mathbf{x}_1 处可微且

$$d(g \circ f)|_{\mathbf{x}=\mathbf{x}_1} = dg|_{y=f(\mathbf{x}_1)} \circ df|_{\mathbf{x}=\mathbf{x}_1}.$$

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{f} & \Omega_2 \\ & \searrow g \circ f & \downarrow g \\ & & \Omega_3 \end{array} \quad \hookrightarrow \quad \begin{array}{ccc} \mathbb{R}^{m_1} & \xrightarrow{df|_{\mathbf{x}=\mathbf{x}_1}} & \mathbb{R}^{m_2} \\ & \searrow d(g \circ f)|_{\mathbf{x}=\mathbf{x}_1} & \downarrow dg|_{y=f(\mathbf{x}_1)} \\ & & \mathbb{R}^{m_3} \end{array}$$

Remark. 这一点实际上就是在说微分运算具有类似于“函子性”的性质成立, 它在对象上把带基点的欧氏空间映到线性空间, 在态射上把 C^1 映射映到线性映射.

Proof. 对任意 $\mathbf{h} \in \mathbb{R}^{m_1}, l \in \mathbb{R}^{m_2}$, 均有对任意 $\mathbf{h} \in \mathbb{R}^{m_1}, l \in \mathbb{R}^{m_2}$, 均有

$$\begin{aligned} f(\mathbf{x}_1 + \mathbf{h}) &= f(\mathbf{x}_1) + df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h}); \\ g(f(\mathbf{x}_1) + l) &= g(f(\mathbf{x}_1)) + dg|_{f(\mathbf{x}_1)} l + \Delta(l). \end{aligned}$$

那么把 $g \circ f(\mathbf{x}_1 + \mathbf{h})$ 写开可得

$$\begin{aligned} g \circ f(\mathbf{x}_1 + \mathbf{h}) - g \circ f(\mathbf{x}_1) &= g(f(\mathbf{x}_1) + df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})) - g(f(\mathbf{x}_1)) \\ &= g(f(\mathbf{x}_1)) + dg|_{f(\mathbf{x}_1)} (df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})) + \Delta(df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})) - g(f(\mathbf{x}_1)) \\ &= dg|_{f(\mathbf{x}_1)} \circ df|_{\mathbf{x}_1} (\mathbf{h}) + dg|_{f(\mathbf{x}_1)} \delta(\mathbf{h}) + \Delta(df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})). \end{aligned}$$

所以当 $\|\mathbf{h}\| \rightarrow 0$ 时, 只需证明

$$\lim_{\mathbf{h} \rightarrow 0} \left\| \frac{dg|_{f(\mathbf{x}_1)} \delta(\mathbf{h}) + \Delta(df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h}))}{\mathbf{h}} \right\| = 0.$$

前者是因为可找到常数 C 使得 $\|dg|_{f(\mathbf{x}_1)} \delta(\mathbf{h})\|$ 被 $C \|\delta(\mathbf{h})\|$ 控制, 然后由 $\delta(\mathbf{h})$ 的定义即得. 后者是因为

$$\lim_{\mathbf{h} \rightarrow 0} \left\| \frac{\Delta(df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h}))}{\mathbf{h}} \right\| = \lim_{\mathbf{h} \rightarrow 0} \left\| \frac{\Delta(df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h}))}{df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})} \right\| \cdot \left\| \frac{df|_{\mathbf{x}_1} \mathbf{h} + \delta(\mathbf{h})}{\mathbf{h}} \right\|.$$

前者为 0, 后者是有限值, 因为 $\|df|_{\mathbf{x}_1} \mathbf{h}\|$ 可以被 $C' \|\mathbf{h}\|$ 所控制, 因此整体为 0. \square

Remark. 实际运用中, 多会采用偏导数的版本. 比如说, 我们设 Ω_1 上有参数系 (x_1, \dots, x_n) , Ω_2 上有参数系 (y_1, \dots, y_m) . 将上面链式法则中左右在某个 \mathbf{x}_i 处取值, 得到

$$\frac{\partial(g \circ f)}{\partial x_i} \Big|_{\mathbf{x}} = dg|_{f(\mathbf{x})} \left(\frac{\partial f}{\partial x_i} \Big|_{\mathbf{x}} \right).$$

这时候需要把 f 写成 $f = (f_1, \dots, f_m)$, 那么上式就可继续化简为

$$\frac{\partial(g \circ f)}{\partial x_i}|_x = \sum_{j=1}^m \frac{\partial f_j}{\partial x_i}|_x \cdot dg|_{f(x)}(\mathbf{y}_j) = \sum_{j=1}^m \frac{\partial f_j}{\partial x_i}|_x \cdot \frac{\partial g}{\partial \mathbf{y}_j}|_{f(x)}.$$

当然, 如果从 Jacobi 矩阵的角度来理解它就是矩阵乘法在把 $\text{Jac}(g)$ 看作 $1 \times m$ 矩阵后单格的展开式, 其中 $\frac{\partial f_j}{\partial x_i}|_x$ 是一个数, $\frac{\partial g}{\partial \mathbf{y}_j}|_{f(x)}$ 是一个向量.

对于变量替换, 问题还更复杂些: 在一组参数 (y_1, \dots, y_n) 下给定映射 $g: \Omega_1 \rightarrow \Omega_2$, 那么在参数 (x_1, \dots, x_n) 下 f 的偏导数表达式是什么? 利用上面的式子, 我们可以发现只要设 $y_i = f_i(x_1, \dots, x_n)$ 得到一个映射 $f: \Omega_1 \rightarrow \Omega_1$, 那么就有

$$\frac{\partial(g \circ f)}{\partial x_i}|_x = \sum_{j=1}^n \frac{\partial f_j}{\partial x_i}|_x \cdot \frac{\partial g}{\partial \mathbf{y}_j}|_{f(x)}.$$

到目前为止都没什么问题. 问题在于在变量替换中我们把 y_i 和 f_i 等同起来了, 还把 g 和 $g \circ f$ 等同起来了 (这从应用来看是自然的: 比如你本来有函数 $g(y)$, 换元 $y = \cos x$, 那么你会直接写 $g(\cos x)$, 这时候 y 充当了函数的角色, g 作为一个映射在几何上也没有发生变化, 只是原本标记为 $y = \cos x$ 的点变成了标记为 x 的点.) 无论如何, 我们最后得到的式子长成这样:

$$\frac{\partial g}{\partial x_i}|_x = \sum_{j=1}^n \frac{\partial y_j}{\partial x_i}|_x \cdot \frac{\partial g}{\partial \mathbf{y}_j}|_y.$$

其中 x 和 y 在几何上对应了相同的点. 这个式子很便于记忆 (虽然我并不喜欢), 总之在利用时我们还是需要注意每个东西真正对应的是什麼. 特别地, 可以发现所有 $\frac{\partial y_j}{\partial x_i}$ 和 g 的选取无关, 只和参数替换本身的性质相关.

(我希望这里是我最后一次讨论这件事情喵……)

Corollary 3.1.11. 对 $\Omega_1 \subset \mathbb{R}^{n_1}$, $\Omega_2 \subset \mathbb{R}^{n_2}$, $f: \Omega_1 \rightarrow \Omega_2$. 若 f 为双射且 f 和 f^{-1} 分别在 x 和 $f(x)$ 处可微, 则 $n_1 = n_2$ 且 $df(x)$ 是可逆的.

Proof. 我们可以对 $f \circ f^{-1}$ 与 $f^{-1} \circ f$ 利用微分的函子性, 得到 $df|_{x_0}$ 和 $df^{-1}|_{f(x_0)}$ 正复合后都是 id , 因此它们均可逆, 特别地 $n_1 = n_2$. $\left(\begin{smallmatrix} \circ \\ \vdots \\ \circ \end{smallmatrix}\right)$

Example 3.1.12 (指数映射的微分). 考虑矩阵可积上的映射

$$\exp: M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}),$$

$$A \mapsto e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

那么根据定义可知

$$\begin{aligned} \exp(A + \mathbf{v}) - \exp(A) &= \sum_{k=0}^{+\infty} \frac{(A + \mathbf{v})^k}{k!} - \sum_{k=0}^{+\infty} \frac{A^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \left((A + \mathbf{v})^k - A^k \right) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \left(\sum_{j=0}^{k-1} A^j \mathbf{v} A^{k-1-j} \right) + O(\|\mathbf{v}\|^2) \cdot I_n. \end{aligned}$$

其中上式是因为对 $(A + \mathbf{v})^k$ 作二项式展开并且按照含有 \mathbf{v} 的次数分类, 并且把 \mathbf{v} 的高次项放为其范数乘以单位矩阵就可得到结论. 那么考虑微分 $d\exp|_A$ 在方向 \mathbf{v} 上的作用, 得到

$$d\exp|_A(\mathbf{v}) = \sum_{k=1}^{\infty} \frac{1}{k!} \left(\sum_{j=0}^{k-1} A^j \mathbf{v} A^{k-1-j} \right).$$

如果 \mathbf{v} 和 A 交换, 那么就可得到

$$d\exp|_A(\mathbf{v}) = \sum_{k=1}^{\infty} \frac{1}{k!} \cdot k A^{k-1} \mathbf{v} = e^A \cdot \mathbf{v}.$$

在不交换时肯并没有好的性质.

Theorem 3.1.13 (积分换元公式 (万恶之源)). 给定映射 $\varphi: V \rightarrow V$ 满足 φ 和 φ^{-1} 均连续可微, 则

$$\int_{\varphi(U)} f(\mathbf{y}) d\mathbf{y} = \int_U f(\varphi(\mathbf{x})) \cdot \det(J_\varphi(\mathbf{x})) d\mathbf{x}.$$

其中 $J_\varphi(\mathbf{x})$ 为 φ 在 x 处的 Jacobi 矩阵.

这个命题的正确性来源于线性映射版本的变元替换公式, 条件里的可微保证了 φ 在局部上形如线性映射. 事实上, 更加本质的是微分形式本身就具有

$$d(\varphi(\mathbf{x})) = \det(J_\varphi(\mathbf{x})) d\mathbf{x}$$

的表达式.

Proof. 我们依次证明如下结论:

- 对方体 I ,

$$m(\varphi(I)) = \int_I \det(J_\varphi(\mathbf{x})) d\mathbf{x}.$$

(你是否回忆起 Growth Lemma 2.4.10?) 对任意 $x \in I$, 线性映射 $d\varphi|_x$ 和 $d\varphi^{-1}|_{\varphi(x)}$ 的

对任意 $\varepsilon > 0$, 将 I 再进行细分为一些以 x_i 为中心的小方块 I_i , 使得

$$|\varphi(x) - \varphi(x_i) - d\varphi|_{x_i}(x - x_i)| \leq M\varepsilon \cdot |x - x_i|.$$

这代表对 I_i 中的任何点 x , $\varphi(x)$ 到 $\varphi(x_i)$ 的距离和这个线性映射所给出的距离差的不会太多, 因此总体来说 $\varphi(I_i)$ 这个像和 $d\varphi|_{x_i}$ 的像 (这是一个平行多面体, 可以直接计算体积) 不会差太多. 我们可以得到存在常数 k 不依赖于 ε 使得

$$m(\varphi(I_i)) \leq (\det(J_\varphi(x_i)) + Mk\varepsilon) \cdot m(I_i).$$

另一边的证明有两种进路: 一种是反过来再对 $\varphi(I_i)$ 构建一个内核, 它能给出 $\varphi(I_i)$ 的测度的一个下界, 另一种是对 φ^{-1} 代入这个单边的不等式得到反向的不等号.

- 对零测集 E 总有 $m(\varphi(E)) = 0$, 即零测集的像为零测集. 我们可以用方体的并来逼近这个零测集



Example 3.1.14. 在 \mathbb{R}^3 上考虑坐标变换

$$(0, +\infty) \times (0, \pi) \times (0, 2\pi) \rightarrow \mathbb{R} \times \mathbb{R} \times \mathbb{R},$$

$$(r, \theta, \varphi) \mapsto (r \sin \theta \sin \varphi, r \sin \theta \cos \varphi, r \cos \theta) =: (x, y, z).$$

考虑这个坐标变换的 Jacobi 矩阵:

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}.$$

其行列式 $|\det(J(r, \theta, \varphi))| = r^2 \sin \theta$. 所以

$$\int_{\mathbb{R}^3} f(x, y, z) = \int_0^{+\infty} \int_0^\pi \int_0^{2\pi} f(r, \theta, \varphi) r^2 \sin \theta \, dr \, d\theta \, d\varphi.$$

3.2 偏导数的交换性与 Taylor 展开

先引入一些记号:

- $C^k(\Omega)$ 为 k 此连续可微函数组成的空间;
- $C^\infty(\Omega)$ 为光滑函数组成的空间, 它是任意次可微的.

Definition 3.2.1 (微分同胚). 对 \mathbb{R}^n 中的两个区域 U, V , 如果存在双射 $f: U \rightarrow V$ 使得 f, f^{-1} 都是光滑的, 则称 U 和 V 是微分同胚 (光滑同胚). 用 $C^\infty(U, V)$ 表示所有从 U 到 V 的微分同胚.

在微分结构中同胚的空间应当被认为是等同的 (在这个范畴下的同构), 即所有微分性质应当在同胚的空间中对应等同.

Example 3.2.2. $f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$ 是光滑的, 然而 f^{-1} 不是光滑的, 所以这不是微分同胚.

Definition 3.2.3 (拉回). 假设 Φ 为 $U \rightarrow V$ 的微分同胚, 则对任意 $f \in C^k(V)$, 有 $f \circ \Phi \in C^k(U)$. 我们就称 $\Phi^*: C^k(V) \rightarrow C^k(U), \Phi^*(f) = f \circ \Phi$ 为 f 在 Φ 下的拉回.

我们证明 Φ^* 的定义是良好的: 对 k 归纳, 当 $k=0$ 时 f 连续可推出 $f \circ \Phi$ 连续. 假设命题对 $k=n$ 成立, 考虑 $f \in C^{n+1}(V)$, 记 $\Phi = (\Phi_1, \dots, \Phi_d)$, 则

$$\frac{\partial(f \circ \Phi)}{\partial x_i} \Big|_x = \sum_{j=1}^d \frac{\partial f}{\partial y_j} \Big|_{\Phi(x)} \cdot \frac{\partial \Phi_j}{\partial x_i} \Big|_x.$$

由 $f \in C^{n+1}(V)$ 可知 $\frac{\partial f}{\partial \mathbf{y}_j} \in C^n(V)$. 根据归纳假设, 应当有

$$\frac{\partial f}{\partial \mathbf{y}_j} \circ \Phi \in C^n(U).$$

根据 Φ 的光滑性可知 $\frac{\partial \Phi_j}{\partial \mathbf{x}_i} \in C^\infty(U)$. 两者结合起来, 就能得到所有 $\frac{\partial(f \circ \Phi)}{\partial \mathbf{x}_i}$ 都是 C^n 的. 根据定义就可知 $f \circ \Phi$ 是 C^{n+1} 的, 因此拉回的定义是合理的.

Remark. 喜欢对偶映射吗.

下面的问题关于偏导数交换, 在理论上具有一定的作用, 他证明了只要函数部分满足 C^1 就能推出偏导数交换.

Theorem 3.2.4 (Clairant-Schwarz, 偏导数交换). 给定 \mathbb{R}^n 上的开集 Ω ($n \geq 2$) 与函数 $f: \Omega \rightarrow \mathbb{R}$. 假设在 Ω 上

$$\frac{\partial f}{\partial \mathbf{x}_i}, \frac{\partial f}{\partial \mathbf{x}_j}, \frac{\partial}{\partial \mathbf{x}_i} \left(\frac{\partial f}{\partial \mathbf{x}_j} \right)$$

存在且连续, 则

$$\frac{\partial}{\partial \mathbf{x}_j} \left(\frac{\partial f}{\partial \mathbf{x}_i} \right)$$

存在且

$$\frac{\partial}{\partial \mathbf{x}_i} \left(\frac{\partial f}{\partial \mathbf{x}_j} \right) = \frac{\partial}{\partial \mathbf{x}_j} \left(\frac{\partial f}{\partial \mathbf{x}_i} \right).$$

Proof. 问题本质就是极限的交换问题, 我们将利用积分中的 Fubini 定理来完成这个“交换”. 由于只涉及到两个方向上的偏导, 因此只需考虑 $n = 2$ 的情形, 我们用 x, y 来代替 x_i, x_j . 首先按定义展开:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{y}} \left(\frac{\partial f}{\partial \mathbf{x}} \right) (x_0, y_0) &= \lim_{s \rightarrow 0} \frac{\frac{\partial f}{\partial \mathbf{x}}|_{(x_0, y_0+s)} - \frac{\partial f}{\partial \mathbf{x}}|_{(x_0, y_0)}}{s} \\ &= \lim_{s \rightarrow 0} s^{-1} \lim_{t \rightarrow 0} \frac{f(x_0+t, y_0+s) - f(x_0, y_0+s) - f(x_0+t, y_0) + f(x_0, y_0)}{t}. \end{aligned}$$

根据条件的 C^1 性, 可以把最后的四项重新组合使用 Newton-Leibniz 公式写成积分形式:

$$\begin{aligned} & f(x_0+t, y_0+s) - f(x_0, y_0+s) - f(x_0+t, y_0) + f(x_0, y_0) \\ &= (f(x_0+t, y_0+s) - f(x_0+t, y_0)) - (f(x_0, y_0+s) - f(x_0, y_0)) \\ & \left(\frac{\partial f}{\partial \mathbf{y}} C^0 \right) = \int_0^s \left(\frac{\partial f}{\partial \mathbf{y}}(x_0+t, y_0+\tilde{s}) - \frac{\partial f}{\partial \mathbf{y}}(x_0, y_0+\tilde{s}) \right) d\tilde{s} \\ & \left(\frac{\partial}{\partial \mathbf{x}_i} \left(\frac{\partial f}{\partial \mathbf{x}_j} \right) C^0 \right) = \int_0^s \left(\int_0^t \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f}{\partial \mathbf{y}}(x_0+\tilde{t}, y_0+\tilde{s}) \right) \right) d\tilde{t} \right) d\tilde{s} \\ & (\text{Fubini}) = \int_0^t \left(\int_0^s \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f}{\partial \mathbf{y}}(x_0+\tilde{t}, y_0+\tilde{s}) \right) \right) d\tilde{s} \right) d\tilde{t}. \end{aligned}$$

然后再对它求极限:

$$\begin{aligned} & \lim_{s \rightarrow 0} \frac{1}{s} \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \left(\int_0^s \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f}{\partial \mathbf{y}}(x_0 + \tilde{t}, y_0 + \tilde{s}) \right) \right) d\tilde{s} \right) d\tilde{t} \\ &= \lim_{s \rightarrow 0} \frac{1}{s} \int_0^s \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f}{\partial \mathbf{y}}(x_0, y_0 + \tilde{s}) \right) \right) d\tilde{s} \\ &= \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f}{\partial \mathbf{y}}(x_0, y_0) \right). \end{aligned}$$

这就证明了偏导数存在且可交换.



Definition 3.2.5 (多重偏导数). 记 $\alpha = (\alpha_1, \dots, \alpha_n)$ 为多重指标, $\alpha_i \in \mathbb{Z}^{\geq 0}$. 我们用 $\partial^\alpha f$ 表示 α_1 个 $\frac{\partial}{\partial x_1}$, \dots , α_n 个 $\frac{\partial}{\partial x_n}$. 那么记

$$\partial^\alpha f = \left(\frac{\partial}{\partial x_1} \right)^{\alpha_1} \left(\frac{\partial}{\partial x_2} \right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_n} \right)^{\alpha_n} f.$$

我们也把它记为

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} f.$$

Example 3.2.6. 考虑函数

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & (x, y) \neq (0, 0); \\ 0, & (x, y) = (0, 0). \end{cases}$$

则

$$\frac{\partial f}{\partial x}(x, y) = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}.$$

从定义上可以算出 $\frac{\partial f}{\partial x}(x, y)$ 连续, 同理也有 $\frac{\partial f}{\partial y}$ 连续. 然而

$$\frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial x} \right) (0, 0) = \lim_{y \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0)}{y} = \lim_{y \rightarrow 0} \frac{-y - 0}{y} = -1.$$

由对称性可知 $\frac{\partial f}{\partial x} \left(\frac{\partial f}{\partial y} \right) (0, 0) = 1 \neq \frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial x} \right) (0, 0)$. 其原因在于虽然一阶导数存在且连续, 然而二阶导数不连续, 导致偏导数无法交换. 可以直观地看到这件事情: 初始分母是 4 次的分子是 2 次的, 求一次导后分子分母的次数差变为 1, 求两次导时变为 0. 但是只有分子次数大于分母时才可以控制 $(0, 0)$ 附近的收敛性, 正如 $x^{1/2}$ 在求导一次后也无法保证连续性一样.

接下来讨论 Taylor 公式的推广. 由于其证明完全就是化归为一维情况后再进行讨论, 所以实用价值并不高.

Theorem 3.2.7. 设 $\Omega \subset \mathbb{R}^n$ 为凸集, f 在 Ω 上 $k+1$ 次连续可微, 对任意 $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in \Omega$, 存在 $\theta \in [0, 1]$ 使得

$$f(\mathbf{y}) = \sum_{|\alpha| \leq k} \frac{(\partial^\alpha f)(\mathbf{x})}{\alpha!} (\mathbf{y} - \mathbf{x})^\alpha + \sum_{|\alpha| = k+1} \frac{\partial^\alpha f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))}{\alpha!} (\mathbf{y} - \mathbf{x})^\alpha.$$

其中有多重指标记号

$$\alpha! = \alpha_1! \dots \alpha_n!, \quad (\mathbf{y} - \mathbf{x})^\alpha = (y_1 - x_1)^{\alpha_1} \dots (y_n - x_n)^{\alpha_n}.$$

Proof. 条件里的凸集保证了可以化归到一元情况利用一维 Taylor 展开. 定义

$$g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})).$$

我们归纳证明

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (\partial^\alpha f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\alpha.$$

当 $k = 0$ 时命题成立, 假设 $k - 1$ 时有

$$g^{(k-1)}(t) = \sum_{|\alpha|=k-1} \frac{(k-1)!}{\alpha!} (\partial^\alpha f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\alpha.$$

接下来对它求一次导数得到

$$g^{(k)}(t) = \sum_{|\alpha|=k-1} \frac{(k-1)!}{\alpha!} \sum_{j=1}^n \left(\frac{\partial}{\partial x_j} (\partial^\alpha f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (y_j - x_j) \right) (\mathbf{y} - \mathbf{x})^\alpha.$$

我们将进行交换求和, 固定 β 使得 $|\beta| = k$, 则存在一些 $|\alpha| = k - 1$ 使得 α 在某个分量上加一就得到 β , 即它们在网格上相邻. 我们称这样的 α 构成的集合为 $N(\beta)$, 则

$$\begin{aligned} g^{(k)}(t) &= \sum_{|\beta|=k} \sum_{\alpha \in N(\beta)} \frac{(k-1)!}{\alpha!} (\partial^\beta f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\beta \\ &= \sum_{|\beta|=k} (\partial^\beta f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\beta \sum_{\alpha \in N(\beta)} \frac{(k-1)!}{\alpha!} \\ &= \sum_{|\beta|=k} (\partial^\beta f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\beta \sum_{\alpha \in N(\beta)} \frac{(k-1)! \sum_{i=1}^n \beta_i}{\beta!} \\ &= \sum_{|\beta|=k} \frac{k!}{\beta!} (\partial^\beta f)(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) (\mathbf{y} - \mathbf{x})^\beta \sum_{\alpha \in N(\beta)} 1. \end{aligned}$$

证毕.



3.3 反函数定理, 隐函数定理

一维情况下我们知道如果 f 是区间 (a, b) 上的连续单射, 那么 f 一定单调, 那么 $f((a, b))$ 一定是一个区间 (即: f 为开映射). 这样的话反函数 f^{-1} 存在且连续.

如果 f 还具有一定的正则性, 比如如果 f 在 (a, b) 上 C^1 并且 $f' \neq 0$, 那么反函数存在且反函数也是 C^1 的. 这个命题将在高维中得到推广:

Theorem 3.3.1 (反函数定理). 设 f 是开区域 U 上的连续可微函数, 假设 $df|_{\mathbf{x}=\mathbf{x}_0}$ 为双射, 那么存在 \mathbf{x}_0 的开邻域 $U_{\mathbf{x}_0}$ 及 $f(\mathbf{x}_0)$ 的邻域 $V_{f(\mathbf{x}_0)}$ 使得 f 是从 $U_{\mathbf{x}_0}$ 到 $V_{f(\mathbf{x}_0)}$ 的连续双射, 且逆映射连续可微 (f 是 C^1 同胚).

我们自然想问在 C^0 时有没有像一维情形一样类似的结论, 实际上有下面在代数拓扑中非常金典的区域不变性定理:

Theorem 3.3.2 (区域不变性定理). 若 $f: B \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 为连续单射, 则 f 为开映射, 且逆映射 $f^{-1}: f(B) \rightarrow B$ 连续.

这个定理的证明和 Brouwer 不动点定理有关, 证明非常拓扑而非分析 (在wzq讲义上有).

我们回到反函数定理, 问题的关键在于如何用正则性条件证明 f 是局部双射, 我们先来看单射部分. 反函数定理的条件告诉我们 f 在局部的动向很像线性函数: 根据微分的定义,

$$f(\mathbf{x}) = f(\mathbf{x}_0) + df|_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0) + o(\mathbf{x} - \mathbf{x}_0).$$

我们想要的是在 $|\mathbf{x}_1 - \mathbf{x}_2|$ 较小时证明 $|f(\mathbf{x}_1) - f(\mathbf{x}_2)|$ 一定不是零. 然而这时最后一项 $o(\mathbf{x} - \mathbf{x}_0)$ 难以控制, 只能把握大致的动向, 无法保证“每个点非零”类似的性质. 因此如果要考察 $|f(\mathbf{x}_1) - f(\mathbf{x}_2)|$, 我们将采用积分的表达式. 记

$$g(t) = f(t \cdot \mathbf{x}_1 + (1-t) \cdot \mathbf{x}_2).$$

则 $g(1) - g(0) = f(\mathbf{x}_1) - f(\mathbf{x}_2)$, 所以

$$\begin{aligned} f(\mathbf{x}_1) - f(\mathbf{x}_2) &= \int_0^1 g'(t) dt = \int_0^1 df|_{t \cdot \mathbf{x}_1 + (1-t) \cdot \mathbf{x}_2}(\mathbf{x}_1 - \mathbf{x}_2) dt \\ &= \int_0^1 df|_{\mathbf{x}_0}(\mathbf{x}_1 - \mathbf{x}_2) dt + \int_0^1 (df|_{t \cdot \mathbf{x}_1 + (1-t) \cdot \mathbf{x}_2} - df|_{\mathbf{x}_0})(\mathbf{x}_1 - \mathbf{x}_2) dt. \end{aligned}$$

由于 $df|_{\mathbf{x}_0}$ 可逆, 故存在仅和 $df|_{\mathbf{x}_0}$ 有关的常数 C 使得

$$|df|_{\mathbf{x}_0}(\mathbf{x}_1 - \mathbf{x}_2)| \geq C |\mathbf{x}_1 - \mathbf{x}_2|.$$

由 $f \in C^1$ 可知对任意 $\varepsilon > 0$, 存在 r 使得对任意 $|\mathbf{z} - \mathbf{x}| < r$ 均有

$$\|df|_{\mathbf{z}} - df|_{\mathbf{x}}\| < \varepsilon.$$

那么对 $\mathbf{x}_1, \mathbf{x}_2 \in B_r(\mathbf{x})$, 总有

$$|(df|_{t \cdot \mathbf{x}_1 + (1-t) \cdot \mathbf{x}_2} - df|_{\mathbf{x}_0})(\mathbf{x}_1 - \mathbf{x}_2)| \leq \varepsilon |\mathbf{x}_1 - \mathbf{x}_2|.$$

这就导致

$$\int_0^1 (df|_{t \cdot \mathbf{x}_1 + (1-t) \cdot \mathbf{x}_2} - df|_{\mathbf{x}_0})(\mathbf{x}_1 - \mathbf{x}_2) dt \leq \varepsilon |\mathbf{x}_1 - \mathbf{x}_2|.$$

取 $\varepsilon < C$ 就可知对任意 $\mathbf{x}_1, \mathbf{x}_2 \in B_r(\mathbf{x})$,

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \geq (C - \varepsilon) |\mathbf{x}_1 - \mathbf{x}_2|.$$

故 f 在 $B_r(\mathbf{x}_0)$ 上为单射, 再证明 f 在某个小邻域内是满射. 如果固定 $\mathbf{y}_0 = f(\mathbf{x}_0)$ 邻域内的 \mathbf{y} , 则问题转化为一个存在性命题. 我们利用 Lusion Lemma 来解决这个存在性问题, 在证明常微分方程解的存在唯一性时已经和它打过交道.

Lemma 3.3.3 (Lusion Lemma (压缩映像定理)). 设 (X, d) 是一个完备度量空间, $T: X \rightarrow X$ 是一个压缩映像, 即存在正常数 $\gamma < 1$, 使得对任意 $x, y \in X$, 均有

$$d(Tx, Ty) \leq \gamma d(x, y).$$

那么 T 存在唯一的不动点 x_0 .

可是我们立即遇到问题: 如何构造出一个压缩映像 T ? 为了寻找 y 的原像, 我们可以构建一个把 $f(x_0)$ 的邻域拉回 x_0 的邻域的映射, 比如说

$$T(x) = f(x) - y + x.$$

那么 T 的不动点就是 y 的原像, 但是

$$T(x_1) - T(x_2) = f(x_1) - f(x_2) + x_1 - x_2$$

没办法和 $x_1 - x_2$ 的模长建立联系. 解决方法是通过复合上一个可逆函数来把 $df|_x$ 打到 id : 待定 $A: V \rightarrow V$, 考虑

$$g = A \circ (f(x) - y), \quad dg|_{x_0} = dA|_{f(x_0)-y} \circ df|_{x_0}.$$

取 $A(z) = f(x_0) - y + (df|_{x_0})^{-1}(z - f(x_0) + y)$, 那么就有

$$dA|_{f(x_0)-y} = (df|_{x_0})^{-1} \Rightarrow dg|_{x_0} = \text{id}.$$

注意这里的 A 依赖于 x_0 的选取, 然而我们也只需要局部的性质就足够了. 因此考虑

$$T(x) = A \circ (f(x) - y) + x,$$

那么仍然有 T 的不动点为 y 的原像, 而另一方面, 两边求微分可得

$$dT|_{x_0} = \text{id} - \text{id} = 0.$$

接下来的操作就和第一部分证明单射时比较像了, 由 $T \in C^1$ 可知对任意 $\varepsilon > 0$ 均存在 r_1 使得对任意 $|z - x_0| \leq r_1$ 均有

$$\|df|_z - df|_x\| < \varepsilon.$$

而对任意 $x_1, x_2 \in \overline{B_{r_1}(x_0)}$, 均有

$$\begin{aligned} |f(x_1) - f(x_2)| &= \left| \int_0^1 (df|_{t \cdot x_1 + (1-t) \cdot x_2} - df|_{x_0})(x_1 - x_2) dt \right| \\ &\leq \varepsilon \cdot |x_1 - x_2|. \end{aligned}$$

所以取 $\varepsilon = \frac{1}{42}$ 并取出满足要求的 r_1 , 那么 T 在 $\overline{B_{r_1}(x_0)}$ 上是压缩映像. 最后考虑 $T(x)$ 的值域, 我们知道存在常数 C 使得

$$|T(x) - x_0| \leq f(x_0) - y + C \cdot |f(x) - f(x_0)| + x - x_0.$$

: 问题的关键在于利用好 df 在 x_0 处非退化的性质. 从点集拓扑中我们知道从紧集到 Hausdorff 空间的连续单射一定是开映射, 在欧氏空间中表达为下面的引理:

Lemma 3.3.4. 设 $V \subset \mathbb{R}^n$ 为有界闭集, 则 $f: V \rightarrow f(V) \subset \mathbb{R}^m$ 为闭映射.

Proof. V 为紧集, 紧集的闭子集也是紧集, 连续映射把紧集映到紧集, 而在欧氏空间中紧集也一定是闭集. 所以 f 是闭映射. \square

Lemma 3.3.5. 设 f 为 U 到 V 的连续双射, 且在 \overline{U} 上也连续. 若 $f \in C^1$ 且 df 在 U 上处处非退化, 则 f^{-1} 也是 C^1 的.

Proof. 首先证明 $f^{-1}: V \rightarrow U$ 是连续函数. 这是因为 $f: \overline{U} \rightarrow f(\overline{U})$ 为闭映射, 所以限制在 U 上也是闭映射, 故 f^{-1} 连续. 因此只需证明 f^{-1} 是 C^1 的.

固定 $\mathbf{x}_0 \in U$, 考虑 $f(\mathbf{x}_0) \in V$. 由可微性存在 A 使得

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + A\mathbf{v} + o(\mathbf{v}). \quad (1)$$

设函数 $E(\delta)$ 满足

$$E(\delta) = f^{-1}(f(\mathbf{x}_0) + \delta) - \mathbf{x}_0 - A^{-1}\delta,$$

则只需证明 $E(\delta) = o(\delta)$. 把上式重新代入 (1) 中得到

$$\begin{aligned} f(\mathbf{x}_0) + \delta &= f(\mathbf{x}_0 + A^{-1}\delta + E(\delta)) \\ &= f(\mathbf{x}_0) + A(A^{-1}\delta + E(\delta)) + o(A^{-1}\delta + E(\delta)) \\ &\Rightarrow A \cdot E(\delta) = o(A^{-1}\delta + E(\delta)). \end{aligned}$$

根据 f^{-1} 的连续性可知

$$\lim_{\delta \rightarrow 0} E(\delta) = 0.$$

所以 $o(A^{-1}\delta + E(\delta)) = o(\delta)$, 从而 $E(\delta) = A^{-1}o(\delta) = o(\delta)$. \square

接下来回到反函数定理的证明.

Back to 3.3.1. 通过对 f 的平移和复合上一个可逆线性变换可以不妨设

$$\mathbf{x}_0 = \mathbf{0}, \quad f(\mathbf{x}_0) = \mathbf{0}, \quad df(\mathbf{x}_0) = \text{id}.$$

这会对后面的证明有简化作用. 首先证明存在 $U \ni \mathbf{0}$ 使得 f 是单射. 尝试之后会发现如果我们只利用无穷小表达式会得到类似 $o(x) - o(y)$ 的项, 不一定能得到它非零. 因此将采用积分表达式来精确地写出来:

$$f(\mathbf{x}) - f(\mathbf{y}) = \int_0^1 df|_{t\mathbf{x}+(1-t)\mathbf{y}}(\mathbf{x} - \mathbf{y}) dt.$$

根据 f 的 C^1 性质及 df 在 0 处非退化, 对任意 $\varepsilon > 0$, 存在 δ 使得对任意 $|\mathbf{z}| < \delta$ 均有


$$\|df|_{\mathbf{z}} - \text{id}\| < \varepsilon.$$

那么对任意 $|\mathbf{x}|, |\mathbf{y}| < \delta$ 均有

$$|(\mathrm{d}f_{t\mathbf{x}+(1-t)\mathbf{y}} - \mathrm{id})(\mathbf{x} - \mathbf{y})| \leq \varepsilon \cdot |\mathbf{x} - \mathbf{y}|.$$

这就导致

$$\begin{aligned} \left| \int_0^1 \mathrm{d}f_{t\mathbf{x}+(1-t)\mathbf{y}}(\mathbf{x} - \mathbf{y}) \, dt \right| &\geq \int_0^1 (1 - \varepsilon) |\mathbf{x} - \mathbf{y}| \, dt \\ &= (1 - \varepsilon) |\mathbf{x} - \mathbf{y}|. \end{aligned}$$

这样令 $\varepsilon = 1/2$, $U = B(\mathbf{0}, \delta)$ 即可知 $f|_U$ 为单射, 我们令 $V = f(U) \ni \mathbf{0}$, 则 $U \rightarrow V$ 为满射. 

再证明 f^{-1} 的存在连续性, 即证明 f^{-1} 为开映射. 如果利用区域不变性定理问题已经解决, 但事实上我们现在需要通过 $\mathrm{d}f$ 的更特殊的性质来证明这个命题. 开映射的证明可以化归为一个局部性质: 只需证明在 $\mathbf{0}$ 附近满足要求. 根据刚刚的推理 $\mathbf{0}$ 附近有一个闭邻域 $\overline{B}(\mathbf{0}, \delta)$ 上是连续单射, 下证明 $f(\mathbf{0})$ 是 $f(\overline{B}(\mathbf{0}, \delta))$ 的内点, 即

$$\exists \delta_1 > 0 \text{ s.t. } \forall \mathbf{y} \in \mathbb{R}^n, |\mathbf{y}| < \delta_1, \exists \mathbf{x} \in \overline{B}(\mathbf{0}, \delta) \text{ s.t. } f(\mathbf{x}) = \mathbf{y}.$$

这又是某种意义上的存在性的不动点定理. 我们利用和 Lyapunov Lemma 类似的想法, 找到一列 x_n 不断修正与 $f(x_n)$ 与 \mathbf{y} 之间的误差, 从而最后由连续性得到 $f(\lim_{n \rightarrow \infty} x_n) = \mathbf{y}$.

归纳定义

$$\mathbf{x}_0 = \mathbf{y}, \quad \mathbf{x}_{k+1} = \mathbf{y} - (f(\mathbf{x}_k) - \mathbf{x}_k).$$

可以归纳证明

$$|\mathbf{x}_{k+1}| = |\mathbf{y} - o(\mathbf{x}_k)| \leq |\mathbf{y}| + o(|\mathbf{x}_k|) \leq \delta_1 + \varepsilon_0 |\delta| \leq \delta.$$

再证明 \mathbf{x}_k 收敛:

$$\begin{aligned} |\mathbf{x}_{k+1} - \mathbf{x}_k| &= |o(\mathbf{x}_k) - o(\mathbf{x}_{k-1})| = |f(\mathbf{x}_k) - \mathbf{x}_k - f(\mathbf{x}_{k-1}) + \mathbf{x}_{k-1}| \\ &\leq \varepsilon_0 |\mathbf{x}_k - \mathbf{x}_{k-1}|. \end{aligned}$$

则 $\{x_k\}$ 为 $\overline{B}(\mathbf{0}, \delta)$ 中的 Cauchy 列.


Corollary 3.3.6. 设 $k \geq 2$, $f \in C^k(\Omega)$. 则在局部上存在逆映射 $f^{-1} \in C^k(V)$.

Proof. 链式法则

$$\mathrm{d}f^{-1}(\mathbf{y}) = (\mathrm{d}f)^{-1}(f^{-1}(\mathbf{y})).$$

如果把 $\mathrm{d}f$ 和 $\mathrm{d}f^{-1}$ 都表示为矩阵, 分别记为 A 和 B , 则条件告诉我们 $A_{ij} \in C^k(\Omega)$, 并且链式法则给出

$$B(\mathbf{y}) = A^{-1}(f^{-1}(\mathbf{y})).$$

所以可以逐项写出 B 的每个矩阵元, 它是 A 中矩阵元的一些初等复合, 所以也是 C^k 的. 

接下来将引出隐函数定理. 对一个开区域 $\Omega \subset \mathbb{R}^{n+p} = \mathbb{R}^n \times \mathbb{R}^p$, 可以把它看作 $f(\mathbf{x}, \mathbf{y})$. 那么固定某个 \mathbf{x}_0 就可得到一个关于 \mathbf{y} 的 $\mathbb{R}^p \rightarrow \mathbb{R}^p$ 的映射, 这个映射的微分就记为 $d_{\mathbf{y}}f(\mathbf{x}_0, \mathbf{y})$.

Theorem 3.3.7 (隐函数定理). 设 $f: \Omega \rightarrow \mathbb{R}^p$ 连续可微, 如果存在 $(\mathbf{x}^*, \mathbf{y}^*) \in \Omega$ 使得 $f(\mathbf{x}^*, \mathbf{y}^*) = 0$ 且 $d_{\mathbf{y}}f(\mathbf{x}^*, \mathbf{y}^*)$ 可逆, 则存在 \mathbf{x}^* 的开邻域 $U \subset \mathbb{R}^n$ 和 \mathbf{y}^* 的邻域 $V \subset \mathbb{R}^p$ 及连续可微函数 $\phi: U \rightarrow V$ 满足对任意 $\mathbf{x} \in U$, 有

$$f(\mathbf{x}, \phi(\mathbf{x})) = 0, \quad d\phi|_{\mathbf{x}} = -(d_{\mathbf{y}}f|_{(\mathbf{x}, \phi(\mathbf{x}))})^{-1} \cdot d_{\mathbf{x}}f|_{(\mathbf{x}, \phi(\mathbf{x}))}.$$

Remark. 隐函数定理具有很强的几何意义. 反函数定理告诉我们如果有一个高维空间到自己的映射那么可以找到一个反函数, 而如果一个从高维空间打到低维空间的映射, 那么每个点的原像都是一个空间, 从线性映射的观点来看, 如果从 $n+d$ 维空间到 d 维空间的映射秩为 d , 那么它的核维数就是 n . 那么隐函数定理就算把这个命题推广到了局部线性函数, 即 C^1 函数上.

Proof. 证明思路是利用反函数定理来构造一个 $\mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+p}$ 的 C^1 函数. 构造

$$F: \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n \times \mathbb{R}^p, \quad (x, y) \mapsto (x, f(x, y)).$$

那么对 $\mathbf{p} \in \mathbb{R}^n \times \mathbb{R}^p$, $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, $\mathbf{v}_1 \in \mathbb{R}^n$, $\mathbf{v}_2 \in \mathbb{R}^p$, 有

$$dF|_{\mathbf{p}}(\mathbf{v}) = \mathbf{v}_1 + d_{\mathbf{x}}f|_{\mathbf{p}}(\mathbf{v}_1) + d_{\mathbf{y}}f|_{\mathbf{p}}(\mathbf{v}_2).$$

因此 $dF|_{\mathbf{p}}(\mathbf{v}) = 0 \Rightarrow \mathbf{v}_1 = 0$, 从而 $d_{\mathbf{y}}f|_{\mathbf{p}}(\mathbf{v}_2) = 0$. 根据 $d_{\mathbf{y}}f$ 的可逆性, $\mathbf{v}_2 = 0$, 从而 $\mathbf{v} = 0$.

所以 $dF|_{\mathbf{p}}$ 可逆, 并且根据 f 的 C^1 性可推出 F 也是 C^1 的. 利用反函数定理, 存在 $\mathbf{p} = (\mathbf{x}^*, \mathbf{y}^*)$ 的邻域 U 和 $F(\mathbf{x}^*, \mathbf{y}^*) = (\mathbf{x}^*, 0)$ 的邻域 V 使得 $F: U \rightarrow V$ 是 C^1 同胚. 取 $V_x \subset \mathbb{R}^n$ 使得 $V_x \times \{0\}$ 的某个邻域落在 V 中. 那么就可以使用

$$\phi: V_x \rightarrow \mathbb{R}^p, \quad F^{-1}(x, 0) = (x, \phi(x))$$

来定义 ϕ . 设投影映射 $p: \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ 把 (x, y) 映到 y , 嵌入映射 $\iota: \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^p$ 把 x 映到 $(x, 0)$, 那么实际上就有 $\phi = p \circ F^{-1} \circ \iota$. 从而 ϕ 是一些 C^1 映射的复合, 故也是 C^1 的. 此外, $d\phi$ 也可以写出具体表达式:

$$d\phi|_x = dp|_{(x, \phi(x))} \circ dF^{-1}|_{(x, 0)} \circ d\iota|_x.$$

这个表达式比较抽象, 如果为 \mathbb{R}^n 和 \mathbb{R}^p 都选择一组基, 那么线性映射可表达为矩阵形式, 而 p 和 ι 对应的矩阵在每点处相同:

$$d\phi|_x = \begin{bmatrix} 0 & I_p \end{bmatrix} \begin{bmatrix} I_n & 0 \\ A & B \end{bmatrix}^{-1} \begin{bmatrix} I_n \\ 0 \end{bmatrix}.$$

其中 $A = d_{\mathbf{x}}\phi|_{(x, \phi(x))}$, $B = d_{\mathbf{y}}\phi|_{(x, \phi(x))}$. 显然根据分块矩阵的运算规则我们只关心矩阵左下角的部分, 同样地设 $dF^{-1}|_{(x, 0)}$ 写成如上的分块矩阵形式, 即

$$\begin{bmatrix} D' & C' \\ A' & B' \end{bmatrix} \begin{bmatrix} I_n & 0 \\ A & B \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_p \end{bmatrix}.$$

考虑第二行得到 $A' + B'A = 0$ 以及 $B'B = I_p$, 因此 $A' = -B^{-1}A$. 用线性映射形式写下来即是

$$d\phi|_x = -(d_{\mathbf{y}}f|_{(x, \phi(x))})^{-1} \cdot d_{\mathbf{x}}f|_{(x, \phi(x))}.$$

这就得到了结论. \square

Remark. 我们可以把问题的描述变得再一般些, 考虑 $\mathbf{p} \in \Omega \subset \mathbb{R}^{n+p}$, 如果 $f(\mathbf{p}) = 0$ 并且 $\text{rk } df|_{\mathbf{p}} = p$ (注意: 这个描述本身是只和整体坐标有关的) 那么可以在 df 对应的矩阵中选取线性无关的 p 列, 它们张成的空间记为 \mathbb{R}^p , 然后再把 \mathbf{p} 的一个邻域上的点写成 $\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ 的形式, 并选取一个邻域使得 $f(x, y) = 0 \Leftrightarrow y = \Phi(x)$. 这样做的好处是让坐标的选取包含在局部性质之内, 从而更加吻合后续子流形的定义.

4 (子) 流形论

4.1 子流形的定义

反函数定理和隐函数定理可以和线性方程组的理论相对照: 考虑线性映射 $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$, 对应矩阵 $A \in \mathbb{R}^{n \times m}$.

则若 $n = m$ 且 $\text{rk } T = n$, 则对任意的 $y \in \mathbb{R}^n$ 存在唯一的 $x = T^{-1}y$ 使得上述方程成立. 若 $n < m$ 且 $\text{rk } T = n$, 则对任意的 $y \in \mathbb{R}^n$, $T^{-1}(y)$ 是一个 $n - m$ 维超平面. 而反函数定理和隐函数定理就对应了它们的推广情形, 即如果有一个映射局部上很接近 T , 满足 $f(x) = Tx + o(x)$, 那么反函数定理和隐函数定理在局部上仍成立.

Definition 4.1.1 (d -维子流形). 设 $M \subset \mathbb{R}^n$ 非空. 若存在 $d \geq 0$ 使得对任意 $x \in M$, 存在开集 $U \subset \mathbb{R}^n, U \ni x$ 以及 \mathbb{R}^n 中的开集 V , 还有微分同胚 $\Phi: U \rightarrow V$ 使得

$$\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\}).$$

那么称 M 是 \mathbb{R}^n 中的一个 d -维微分子流形, 称 d 为 M 的维数, 记作 $\dim M = d$; $n - d$ 为 M 的余维数, 记作 $\text{codim } M = n - d$.

Remark. 子流形的定义和一般的流形有一定区别, 一般的流形给定开覆盖 \mathcal{U} , 对每个开集 $B \in \mathcal{U}$ 存在一个 φ_B 把 B 打到欧氏空间 \mathbb{R}^n 中的开集 \tilde{B} , 且 φ_B 是开集. 并

且对 $B_1, B_2 \in \mathcal{U}$, 如果 $B_1 \cap B_2 \neq \emptyset$, 则 φ_1, φ_2 分别给出了两个从 $B_1 \cap B_2$ 打到 \mathbb{R}^n 的映射, 记它们的像集分别为 A_1, A_2 . 那么 $\varphi_2^{-1} \circ \varphi_1$ 给出了 $A_1 \rightarrow A_2$ 的同胚, 并且它是一个 $\mathbb{R}^n \rightarrow \mathbb{R}^n$ 的映射, 我们就让这个映射被赋予 C^k 的微分结构. 子流形是流形的一个特殊情况, 但是由于任何流形都可以嵌入到某个足够高维的 \mathbb{R}^n , 所以也可以被实现为上述的一个子流形. 此外, 子流形的正则性完全取决于 Φ 的正则性, 如果说 Φ 是 C^k 的那么这个流形就是 C^k 流形, 当 $k=0$ 时没有微分结构, 称为拓扑流形.

如果我们在 $V \cap (\mathbb{R}^d \times \{0\})$ 中选取一些坐标函数 y_1, \dots, y_d , 那么可以把这些函数拉回到 M 上, 构成集合 $\{\Phi^*(y_i)\}_{1 \leq i \leq d}$. 它们构成了 M 上的坐标曲线, 它们给出了 M 上的局部坐标系, 即每个 $z \in U \cap M$ 拥有一个局部坐标 (把 $\Phi(z)$ 对应的坐标作为其局部坐标?) 局部坐标的选取不唯一, 所以一个点在几何上不变, 但是在代数上可以有不同坐标的表达形式.

隐函数定理就给我们一族子流形, 这是因为如果有 $f: \mathbb{R}^n \rightarrow \mathbb{R}^d$ 满足处处有 $\text{rk } df|_{(x,y)} = d$, 那么对任意 $x \in \mathbb{R}^n$ 满足 $f(x) = 0 \in \mathbb{R}^d$, 就可选取一组坐标

$$\mathbb{R}^n = \mathbb{R}^{n-d} \times \mathbb{R}^d, x \mapsto (x_1, x_2)$$

使得 $d_{x_2} f|_{(x_1, x_2)}$ 可逆, 考虑隐函数定理中得到的函数 $F = p \times f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, 它是某个 $x = (x_1, x_2)$ 的邻域 U 和 $F(x) = (x_1, 0)$ 的邻域 V 之间的同胚, 另一方面,

$$F(f^{-1}(0) \cap U) = V \cap (\mathbb{R}^{n-d} \times \{0\}).$$

从而 $f^{-1}(0)$ 是子流形, 对每个 $x \in F^{-1}(0)$, F 给出了子流形上的微分同胚, 满足 $\dim F^{-1}(0) = n - d$. 我们还可以对 $f - c$ 利用这个命题, 得到所有的 $f^{-1}(c)$, $c \in \mathbb{R}^d$ 都是 \mathbb{R}^n 中维数为 $n - d$ 的子流形. 我们用纤维的定义来总结一下上面做的事情.

Definition 4.1.2 (纤维). 设 $f: \Omega \rightarrow \mathbb{R}^p$ 为光滑映射, $\Omega \subset \mathbb{R}^n$ 为开集. 对 $c \in \mathbb{R}^p$, c 上的纤维定义为 $f^{-1}(c) \subset \Omega$.

如果对任意 $x \in f^{-1}(c)$, $\text{rk } df|_x = d$, 那么 $f^{-1}(c)$ 是一个余维数为 d 的子流形.

我们经常提到光滑曲线和光滑曲面的概念, 子流形的定义可以让我们对它们下一个精准的定义.

Definition 4.1.3. 若 $M \subset \mathbb{R}^n$ 是 1 维子流形, 则称 M 为 \mathbb{R}^n 中的光滑曲线, 若 $M \subset \mathbb{R}^n$ 是 2 维子流形, 则称为 \mathbb{R}^n 中的光滑曲面. 若 M 是余 1 维子流形, 则称为 \mathbb{R}^n 中的光滑超曲面.

比如说 \mathbb{R}^3 中的柱面是光滑曲面, 因为它是 $f(x, y, z) = x^2 + y^2 - 1$ 的零点集.

Lemma 4.1.4. 设 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 为光滑函数. 若 f 满足对任意 $x_0 \in f^{-1}(0)$ 有 $df|_{x_0} \neq 0$, 则 $f^{-1}(0)$ 是光滑超曲面. 这种曲面称为由一个方程整体定义的光滑超曲面.

Proof. 这是因为 $df|_{x_0} \neq 0 \Leftrightarrow \operatorname{rk} df|_{x_0} = 1$, 所以 $f^{-1}(0)$ 自然是余 1 维子流形. \square

我们还有隐函数定理的反向结论, 隐函数定理说的是函数在一个点处原像是子流形, 我们接下来要证明一个函数的像在局部上也是子流形. 从而在光滑映射下, 流形的像是流形.

Theorem 4.1.5. 设 $\Omega \subset \mathbb{R}^n$ 为开集, $f: \Omega \rightarrow \mathbb{R}^{n+p}$ 光滑. 如果 $x^* \in \Omega$ 满足 $\operatorname{rk} df|_{x^*} = n$, 则存在开集 $U \subset \mathbb{R}^n$ 且 $x^* \in U$, 使得 $f(U) \subset \mathbb{R}^{n+p}$ 是 n 维子流形.

Proof. 由于 $\operatorname{rk} df|_{x^*} = n$, 因此可以选取 \mathbb{R}^{n+p} 的一组坐标

$$\mathbb{R}^{n+p} \rightarrow \mathbb{R}^n \times \mathbb{R}^d, \quad x \mapsto (x_1, x_2).$$

满足 $d(p \circ f)|_{x^*}$ 可逆, 其中 p_1 为投影映射 $x \mapsto x_1$. 则由反函数定理, 存在 x^* 的开邻域 U 以及 $(p \circ f)(x^*)$ 的开邻域 V 使得 U 和 V 微分同胚. 考虑映射

$$\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^d, \quad x_1 \mapsto (p_2 \circ f) \circ (p_1 \circ f)^{-1}(x_1)$$

它是光滑的, 并且对任意 $x_1 \in V$,

$$(x_1, x_2) \in f(U) \Leftrightarrow x_2 = \Phi(x_1).$$

从而 $f(U) = \{(x_1, \Phi(x_1)) : x_1 \in V\}$. 我们证明 $f(U)$ 是 n 维子流形. 这是因为只需考虑映射 $\Psi: (x_1, x_2) \mapsto (x_1, x_2 - \Phi(x_1))$, 满足 $d\Psi|_{(x_1, x_2)} = \operatorname{id} - d\Phi|_{x_1}$ 一定可逆. (证明它可逆只需要 Φ 可微, 不需要其它性质了) 从而 Ψ 是 $V \times \mathbb{R}^d \rightarrow V \times \mathbb{R}^d$ 的光滑同胚, 并且满足 $\Psi(f(U) \cap (V \times \mathbb{R}^d)) = V \times \{0\}$, 所以 $f(U)$ 为子流形. \square

Example 4.1.6. Mobius 带是集合

$$M = \left\{ (x, y, z) : \begin{cases} x = (1 + r \cos \frac{\theta}{2}) \cos \theta; \\ x = (1 + r \cos \frac{\theta}{2}) \sin \theta; \\ z = r \sin \frac{\theta}{2}. \end{cases}, \theta \in [0, 2\pi], r \in (-1, 1) \right\}.$$

记 $e(\theta) = (\cos \theta, \sin \theta, 0)$, $e_0 = (0, 0, 1)$. 则 Mobius 带使用参数 (θ, r) 表示为

$$\phi: (\theta, r) \mapsto e(\theta) + r \cos \frac{\theta}{2} e'(\theta) + r \sin \frac{\theta}{2} e_0.$$

因此

$$\begin{aligned} \phi_\theta &= e'(\theta) - \frac{r}{2} \sin \frac{\theta}{2} e(\theta) + r \cos \frac{\theta}{2} e'(\theta) + \frac{r}{2} \cos \frac{\theta}{2} e_0. \\ \phi_r &= \cos \frac{\theta}{2} e(\theta) + \sin \frac{\theta}{2} e_0. \end{aligned}$$

我们发现 $e_0, e(\theta), e'(\theta)$ 构成了一个单位正交标架, 因此可以很容易计算它们的内积得到 $\langle \phi_\theta, \phi_r \rangle = 0$. 因此最后只需证明它们都不是零, 这是显见的. 从而, M 的确是一个光滑曲面, 因此可以被实现为一个光滑函数 $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ 的零点.

然而, Mobius 带是熟知的不可定向曲面. 不可定向曲面的显著特征是:

Proposition 4.1.7. 不存在光滑函数 $f \in C^\infty(\mathbb{R}^3)$ 使得

$$M = \{(x, y, z) : f(x, y, z) = 0\}$$

且 df 在 M 上每点不为零, 即 M 不能用一个方程的零点来整体定义.

关键在于不能定义一个处处非退化的法向量场使得在每一点处的所谓“定向”都被唯一确定, 一个点处的法向量转动一周后变为其反向, 从而整体性不能满足.

Proof. 反设存在这样的 f , 则 $f(\mathbf{e}(\theta) + r \cos \frac{\theta}{2} \mathbf{e}_0 + r \sin \frac{\theta}{2} \mathbf{e}_0) = 0$.

对 θ 求导, 在 $r = 0$ 处取值:

$$\frac{\partial f}{\partial \theta}|_{(\theta, 0)} = df|_{\mathbf{e}(\theta)}(\mathbf{e}'(\theta)) = 0.$$

对 r 求导, 在 $r = 0$ 处取值:

$$\frac{\partial f}{\partial r}|_{(\theta, 0)} = df|_{\mathbf{e}(\theta)}(\cos \frac{\theta}{2} \mathbf{e}_0 + \sin \frac{\theta}{2} \mathbf{e}_0) = 0.$$

由于 $df|_{\mathbf{e}(\theta)} \neq 0$, 而它已经在两个方向上取值为 0, 所以一定在第三个方向

$$\mathbf{v}(\theta) = \cos \frac{\theta}{2} \mathbf{e}_0 - \sin \frac{\theta}{2} \mathbf{e}(\theta)$$

上有 $df|_{\mathbf{e}(\theta)}(\mathbf{v}(\theta)) \neq 0$ 对任意 θ 均成立. 然而, 取 $\theta = 0, 2\pi$ 分别代入得

$$\mathbf{e}(0) = \mathbf{e}(2\pi), \quad \mathbf{v}(0) = -\mathbf{v}(2\pi).$$

所以 $df|_{\mathbf{e}(0)}(\mathbf{v}(0))$ 与 $df|_{\mathbf{e}(2\pi)}(\mathbf{v}(2\pi))$ 反号, 因此由介值原理可知存在一个 θ 使得 $df|_{\mathbf{e}(\theta)} = 0$, 矛盾. □

接下来将证明子流形的原像也是子流形.

Proposition 4.1.8. 设 $\Omega \subset \mathbb{R}^n$, $f: \Omega \rightarrow \mathbb{R}^m$ 满足 $\text{rk } df = m$. 则如果 $S \subset \mathbb{R}^m$ 是子流形, 那么 $f^{-1}(S)$ 是余维数为 $\text{codim } S$ 的子流形.

特别地, $0 \in \mathbb{R}^m$ 是 0 维子流形, $\text{codim}\{0\} = m$, 所以 $f^{-1}(0)$ 余维数为 m .

Proof. 设 $\dim S = s$, 考虑下面的交换图表

$$\begin{array}{ccccc} \mathbb{R}^n \supset \Omega & \xrightarrow{f} & \mathbb{R}^m & \xrightarrow{\Phi} & \mathbb{R}^m \\ \uparrow & & \uparrow & & \uparrow \\ f^{-1}(S) & \xrightarrow{f} & S & \xrightarrow{\Phi} & \mathbb{R}^s \times \{0\} \end{array}$$

这可以引出从 \mathbb{R}^n 到 \mathbb{R}^{m-s} 的映射 F , 把 x 打到 $\Phi \circ f$ 的后 $m-s$ 个分量, 那么 $f^{-1}(S) = F^{-1}(0)$, 并且 $\text{rk } dF = m-s$. 所以 $f^{-1}(S)$ 是维数为 $n - (m-s)$ 的子流形. □

接下来将定义切空间, 定义的标准是, 我们不借助任何选取坐标系相关的信息来刻画切空间.

Definition 4.1.9 (切空间). 给定 $p \in M$, 考虑所有参数化曲线 $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \subset \mathbb{R}^n$, 使得 $\gamma(0) = p$, 那么这给出一个

$$\gamma'(0) = \frac{d\gamma}{dt}|_{t=0}.$$

定义 M 的切空间为 M 上所有这样的向量 $\gamma'(0)$ 构成的空间 (集合), 即

$$T_p M = \{\gamma'(0) \in \mathbb{R}^n \mid \gamma: (-\varepsilon, \varepsilon) \rightarrow M \subset \mathbb{R}^n, \gamma \in C^\infty, \gamma(0) = p\}.$$

直观感觉是切空间为线性空间, 证明方法是先证明切空间在集合上与平坦空间之间的对应, 那么平坦空间在线性映射下的像集就一定是线性空间.

Proposition 4.1.10 (切空间的推出). 设 $\Phi: U \rightarrow V$ 为微分同胚, $M \subset U$ 为子流形, $p \in U$. 那么 $\Phi(M)$ 是子流形, 且 $\Phi(p)$ 的切空间为

$$T_{\Phi(p)} \Phi(M) = d\Phi|_{x=p}(T_p M)$$

Remark. 虽然 pullback 和 pushforward 的记号不太好记, 但是对于一个给定的 object 很容易判断它应该被推出还是拉回. 比如说, 对一个函数 $f: \Omega \rightarrow \mathbb{R}$ 只能定义 $f \circ \Phi$, 对一条曲线 $\gamma: \mathbb{R} \rightarrow \Omega$ 只能定义 $\Phi \circ \gamma$. 因此 Φ 在 f 和 γ 上的作用就清楚了, 它们只能被这么定义. 事实上这就是 0 维切向量场和余切向量场之间的区别.

Proof. 对 M 上经过 p 点的参数曲线 $\gamma(t): (-\varepsilon, \varepsilon) \rightarrow M$, $\gamma(0) = p$,

$$\Phi \circ \gamma: (-\varepsilon, \varepsilon) \rightarrow \Phi(M), \quad \Phi \circ \gamma(0) = \Phi(p).$$

那么 $\Phi \circ \gamma$ 的切向量

$$\frac{d}{dt} \Phi \circ \gamma(t)|_0 = d\Phi|_{\gamma(0)}(\gamma'(0)), \quad \gamma'(0) \in T_{\Phi(p)} \Phi(M).$$

从而

$$d\Phi|_p T_p M \subset T_{\Phi(p)} \Phi(M).$$

反过来也可知

$$d\Phi^{-1}|_{\Phi(p)} T_{\Phi(p)} \Phi(M) \subset T_p M.$$

根据 $d\Phi|_p \circ d\Phi^{-1}|_{\Phi(p)} = \text{id}$ 就可得两者是相等的.



Theorem 4.1.11. $T_p M$ 是 \mathbb{R}^n 的 d -维线性子空间.

Proof. 根据子流形的定义, 可以把 $T_p M$ 和平坦空间中的切空间建立对应.

$$T_p M = d\Phi^{-1}|_{\Phi(p)}(T_{\Phi(p)} \Phi(M)).$$

而

$$(T_{\Phi(p)} \Phi(M)) = \mathbb{R}^d,$$

所以 $T_p M$ 是 \mathbb{R}^d 在一个非退化线性映射下的像, 所以是 d 维子空间.



Proposition 4.1.12. 假设 f 是从 $\mathbb{R}^{d+n} \rightarrow \mathbb{R}^n$ 的光滑映射, $df|_p$ 的秩为 n . 设 $M = f^{-1}(f(p))$ 为 d 维子流形, 则 $T_p M = \ker df|_p$.

Proof. 一方面, $f: M \rightarrow \{f(p)\}$ 是光滑映射, 从而显然 $df|_p(T_p M) = \{0\}$, 所以 $T_p M \subset \ker df|_p$.

另一方面, 由于 $\dim \ker df|_p = (d+n) - n = d$ 并且 $\dim T_p M = d$, 所以二者维数相同, 从而 $\dim T_p M = \dim \ker df|_p$. \square

Remark. 如果把第一步仔细写下来, 考虑 M 上的曲线 γ 使得 $\gamma(0) = p$ 以及 $\gamma'(0) = v \in T_p M$, 那么 $f \circ \gamma \equiv f(p)$, 从而两边在 p 附近求导可知

$$df|_p \circ d\gamma|_0 = 0 \Rightarrow df|_p(\gamma'(0)) = df|_p(v) = 0.$$

这就是把上个命题重新证了一遍, 和把 $\{f(p)\}$ 看作 0 维流形没什么区别.

Example 4.1.13 (超曲面切空间和法向量). 假设 f 是 \mathbb{R}^n 中由一个函数整体定义的超曲面 ($n-1$ 维) $f \in C^\infty(\mathbb{R}^n)$. 在 \mathbb{R}^n 中选取一组坐标系, 则

$$T_p M = \ker df|_p = \{v = (v_1, \dots, v_n) : \sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i = 0.\}$$

那么把

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) \in \mathbb{R}^n$$

看作一个向量, 记为梯度 $\nabla f|_p$, 它就和切空间里的所有向量垂直. 由于 df 非退化, 那么可以定义其单位化

$$\frac{\nabla f|_p}{|\nabla f|_p|}$$

为超曲面在 p 处的单位法向量.

Example 4.1.14 (柱面).

$$f(x, y, z) = x^2 - y^2 - 1.$$

给定 (x, y, z) , 计算梯度向量:

$$\nabla f|_{(x,y)} = (2x, 2y, 0), \quad x^2 + y^2 = 1.$$

所以 $(-y, x, 0)$ 和 $(0, 0, 1)$ 张成 (x, y, z) 处的切空间.

4.2 子流形之间的光滑映射

Definition 4.2.1. 设 M, N 是光滑流形, f 是从 M 到 N 的映射, 如果对任意 $p \in M$ 及微分同胚 $\Phi_1: p \in M \cap U \rightarrow U' \subset \mathbb{R}^{\dim M}$, $\Phi_2: f(p) \in N \cap V \rightarrow V' \subset \mathbb{R}^{\dim N}$ 使得

$$\begin{array}{ccc} M \cap U & \xrightarrow{f} & N \cap V \\ \downarrow \Phi_1 & & \downarrow \Phi_2 \\ U' & \xrightarrow{\Phi_2 f \Phi_1^{-1}} & V' \end{array}$$

构造出的 $\Phi_2 f \Phi_1^{-1}$ 是 $U' \rightarrow V'$ 光滑的, 那么称 f 流形之间的光滑映射. 特别地, 如果 $M \subset \mathbb{R}^m, N \subset \mathbb{R}^n$ 是子流形, 那么只需

$$\begin{array}{ccc} V \cap \{\mathbb{R}^d \times \{0\}\} & \xrightarrow{\Phi^{-1}} & M \cap U \\ & \searrow f \cdot \Phi^{-1} & \downarrow f \\ & & N \end{array}$$

给出的 $f \circ \Phi^{-1}$ 光滑即可. 显然根据交换图表可知光滑映射的定义和 U, V 的选取无关.

Lemma 4.2.2. f 是子流形 M 到 N 的光滑映射当且仅当对任意 $p \in M$, 存在包含 p 的开集 $U \subset \mathbb{R}^m$ 和 C^∞ 映射 $F: U \rightarrow \mathbb{R}^n$ 使得

$$f|_{U \cap M} = F|_{U \cap M}.$$

Proof. 如果 F 存在, 那么

$$\begin{array}{ccccc} V \cap \{\mathbb{R}^d \times \{0\}\} & \xrightarrow{\Phi^{-1}} & M \cap U & \xrightarrow{\iota} & U \subset \mathbb{R}^m \\ & \searrow ? & \downarrow f & \swarrow F & \\ & & N \subset \mathbb{R}^n & & \end{array}$$

交换, 保证了一个 $V \cap \{\mathbb{R}^d \times \{0\}\} \rightarrow N$ 的光滑映射. 另一方面, 如果 f 是子流形之间的光滑映射, 则

$$\begin{array}{ccc} V & \xrightarrow{\Phi^{-1}} & U \\ \downarrow p & & \\ V \cap \{\mathbb{R}^d \times \{0\}\} & \xrightarrow{\Phi^{-1}} & M \cap U \\ & \searrow C^\infty & \downarrow f \\ & & N \subset \mathbb{R}^n \end{array}$$

那么 $F = (f \circ \Phi^{-1}) \circ p \circ \Phi$ 就给出了 $U \rightarrow N$ 的光滑函数, 并且满足 $F \circ \tau = f$, 所以 $f|_{U \cap M} = F|_{U \cap M}$. □

Remark. 如果 F 是大的空间 $\mathbb{R}^m \rightarrow \mathbb{R}^n$ 上的一个光滑映射, 子流形 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 满足 $F(M) \subset N$, 那么 $F|_M$ 构成了 $M \rightarrow N$ 的光滑映射. 也就是说, 实际运用时不一定是把 f 延拓到 F , 也可以直接把 F 限制到 f 来自动给出一个子流形之间的光滑映射.

Example 4.2.3. 设 $M = N = \mathbb{S}^2$ 为 \mathbb{R}^3 中的单位球面. 给定正交变换 $A \in O_3$, 则 $A|_{\mathbb{S}^2}$ 是 $\mathbb{S}^2 \rightarrow \mathbb{S}^2$ 的映射, 由于 A 在 \mathbb{R}^3 上光滑, 那么 A 在 \mathbb{S}^2 上也光滑.

Definition 4.2.4 (流形上的微分). 假设 $M \subset \mathbb{R}^m, N \subset \mathbb{R}^n$ 为子流形, 设 $f: M \rightarrow N$ 光滑. 则对任意 $v \in T_p M$, 存在曲线 $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \subset \mathbb{R}^m$ 使得 $\gamma(0) = p$ 且 $\gamma'(0) = v$.

那么一定有 $f \circ \gamma: (-\varepsilon, \varepsilon) \rightarrow N \subset \mathbb{R}^n$ 满足 $f \circ p(0) = f(p)$, 从而 $(f \circ p)'(0) \in T_{f(p)} N$, 这就可以定义

$$df|_p: T_p M \rightarrow T_{f(p)} N, \quad v \mapsto (f \circ p)'(0).$$

称 df 为映射 f 的微分, $df|_p$ 为 f 在 p 处的微分.

我们还待以证明 $df|_p$ 是线性映射. 根据子流形光滑映射的性质, f 可看成是 \mathbb{R}^m 到 \mathbb{R}^n 的光滑映射 F 在 M 上的限制

$$\frac{df \circ \gamma(t)}{dt} \Big|_{t=0} = \frac{dF(\gamma(t))}{dt} \Big|_{t=0} = dF|_p(\gamma'(0)).$$

这就仅和 $\gamma'(0) = v$ 的选取有关. 所以

$$df|_p(v) = dF|_p(v)$$

是线性映射, 因为 $dF|_p$ 是线性映射.



接下来给出向量场的定义, 它和 df 所代表的余切向量场是对偶的.

Definition 4.2.5 (向量场和切丛). 设 $M \subset \mathbb{R}^n$, 对 $p \in M$, $T_p M \subset \mathbb{R}^n$ 是线性子空间. 定义切丛为

$$TM = \bigsqcup_{p \in M} T_p M = \{(x, y) : x \in M, y \in T_x M\}.$$

对切丛上的一个元素 $(x, y) \in TM$, 存在投影映射 $\pi : TM \rightarrow M$, $(x, y) \mapsto x$ 把 (x, y) 打回它在 M 上对应的点.

在切丛上的一个截面就是切向量场, 它为每个 $p \in M$ 指定一个向量 $X(p) \in T_p M$. 这等价于说有

$$X : M \rightarrow TM, \quad p \mapsto (p, X(p)).$$

我们可以讨论向量场的光滑性质, 这是因为在每个局部上存在 $\Phi : U \cap M \rightarrow V \subset \mathbb{R}^d$, 它诱导的 $d\Phi$ 把每个 $X(p)$ 送到 $d\Phi|_p X(p)$, $p \in U \cap M$. 这就得到了 V 上的一个切向量场 $d\Phi \circ X \circ \Phi^{-1}$, 它也可以被看作 $\mathbb{R}^n \rightarrow \mathbb{R}^n$ 的函数. 用欧氏空间中的光滑性可以用来定义 X 的光滑性.

当然, 在 M 为子流形时, 可直接有 $T_p M \subset \mathbb{R}^n$, 从而 X 可直接看作 $M \rightarrow \mathbb{R}^n$ 的函数, 而又有 $M \subset \mathbb{R}^n$, 所以可以直接讨论 X 在欧氏空间上的光滑性. 记 M 上所有光滑切向量场构成集合为 $\Gamma(M, TM) = \mathfrak{X}(M)$.

切向量场可以参与运算. 对 M 上所有光滑函数 f , 可以定义

$$f : \mathfrak{X}(M) \rightarrow \mathfrak{X}, \quad (fX)(p) = f(p)X(p).$$

同时, 全体光滑切向量场构成加法群, 从而 $\mathfrak{X}(M)$ 具有 $C^\infty(M)$ -模结构.

Proposition 4.2.6. 设 $M \subset \mathbb{R}^n$ 为子流形, 那么

$$TM = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n : x \in M, v \in T_x M\}$$


是 \mathbb{R}^{2n} 中的子流形, 并且 $\dim TM = 2 \dim M$.

Proof. 设 φ 是通常所说的光滑同胚的逆, 它把 $\mathbb{R}^d \times \{0\}$ 上的点映到 M 上. 定义映射

$$T\varphi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n, \quad (x, v) \mapsto (\varphi(x), d\varphi|_x(v)).$$

首先 φ 构成从 $\mathbb{R}^n \rightarrow \mathbb{R}^n$ 的微分同胚, 而 $d\varphi|_p$ 构成 $\mathbb{R}^n \cong T_p\mathbb{R}^n \rightarrow T_p\mathbb{R}^n \supset T_pM$ 的同胚, 所以 $T\varphi$ 定义了从 \mathbb{R}^{2n} 到 \mathbb{R}^{2n} 的微分同胚. 因此只需证明 $T\varphi$ 在局部上构成到切丛的微分同胚. 这是因为

$$\varphi(\mathbb{R}^d \times \{0\}) = M, \quad d\varphi|_p(\mathbb{R}^d \times \{0\}) = T_pM.$$

从而它的像在局部上就是 TM . 这就同时证明了 TM 是子流形并且 $\dim TM = 2\dim M$. 

回忆一个光滑函数 $f: M \rightarrow N$ 的微分 df 满足 $df|_p: T_pM \rightarrow T_{f(p)}N \hookrightarrow TN$. 那么当 p 变化时可以取遍切丛中的每个向量, 从而可以通过切丛来记录微分得到的所有信息, 即定义

$$df: TM \rightarrow TN, \quad df(p, v) := d|_p(v) = \frac{d}{dt}f \circ \gamma(t)|_{t=0}.$$

刚刚说到 f 可以作用在 X 上得到一个切向量场, 然而 X 也可以作用在 f 上的得到一个光滑函数: 定义

$$(Xf)(p) = \nabla_{X(p)}f(p) = \frac{\partial f}{\partial \mathbf{X}(p)}|_p.$$

如果选取一组局部坐标, 那么我们可以把 X 写为

$$X(p) = \sum_{i=1}^d X^i(p) \frac{\partial}{\partial \mathbf{x}^i}.$$

其中 $X^i \in C^\infty(M)$, $\frac{\partial}{\partial \mathbf{x}^i}$ 是切空间的一组基. 那么 X 显然可以被“作用”在 f 上, 这可以被写为

$$Xf = \sum_{i=1}^d X^i \frac{\partial f}{\partial \mathbf{x}^i}.$$

两边作用在 p 上得到

$$Xf(p) = \sum_{i=1}^d X^i(p) \frac{\partial f}{\partial \mathbf{x}^i}|_p.$$

Proposition 4.2.7. $X: C^\infty(M) \rightarrow C^\infty$, $X \mapsto Xf$ 是良定的.

4.3 Lagrange 乘子法

Proposition 4.3.1. 设 $M \subset \mathbb{R}^m$ 为子流形, $f \in C^\infty(M)$. 如果 p 是 f 的局部极值点, 那么 $df|_p = 0$.

Proof. 只需把问题简化到一维情形. 我们对任意 $v \in T_p M$ 证明 $df|_p(v) = 0$. 这样可以找到曲线

$$\gamma : (-\varepsilon, \varepsilon \rightarrow M), \gamma(0) = p, \gamma'(0) = v.$$

从而 $f \circ \gamma$ 是 $(-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ 的一维函数, 并且 0 是其极值点. 由一维微积分的理论就可知

$$df|_p(v) = (f \circ \gamma)'(0) = 0.$$



我们介绍 Lagrange 乘子法: 背景是想对 \mathbb{R}^n 上的光滑函数 f 求极值. 一般来说 $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ 还满足一些约束条件

$$\begin{cases} g_1(x_1, \dots, x_n) = 0, \\ g_2(x_1, \dots, x_n) = 0 \\ \dots \\ g_m(x_1, \dots, x_n) = 0 \end{cases}$$

上面的约束方程可以看作给定一个光滑函数 $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, 要在子流形 $M := g^{-1}(0)$ 上求函数 f 的极值 (这里我们先不妨让 g 满足隐函数定理的条件, 即这些 dg_i 线性无关, 否则总可以删去一些 relation 使得这个条件成立).

那么根据上面的命题 4.3.1, $x \in M$ 取极值的必要条件是 $df|_x = 0$, 即 $T_x M \subset \ker df|_x$. 然而我们知道 $T_x M = \ker dg|_x$, 所以条件化为

$$\ker dg|_x \subset \ker df|_x.$$

这里 $df : \mathbb{R}^n \rightarrow \mathbb{R}$, $dg : \mathbb{R}^n \rightarrow \mathbb{R}^m$, 为了把它完全 reduce 到线性函数的情形, 只需注意到再把 g 拆开就可以了:

$$\bigcap_{i=1}^m \ker dg_i|_x = \ker dg|_x \subset \ker df|_x.$$

根据线性代数的知识可知: 上面的事情等价于 $df|_x \in \text{span}(dg_1|_x, \dots, dg_m|_x)$. 从而存在 $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ 使得

$$df|_x = \sum_{i=1}^m \lambda_i dg_i|_x. \quad (*)$$

因此标准的 Lagrange 乘子法考虑新的函数 $\mathcal{L} : \mathbb{R}^m \times M \rightarrow \mathbb{R}$,

$$\mathcal{L}(\lambda, x) = f(x) - \sum_{j=1}^m \lambda_j g_j(x)$$

那么根据 (*) 可知极值点满足存在一个 $\lambda \in \mathbb{R}^m$ 使得 $d\mathcal{L} = 0$. 反过来, 如果仅考虑 \mathcal{L} 这个函数, 那么它的极值点将成为 f 在子流形上的极值点的候选人. (*) 式更便于计算的形式是利用 \mathbb{R}^n 上的内积把 dg_i 和 ∇g_i 等同起来, 那么就有

$$\nabla f|_x \in \text{span}(\nabla g_i|_x).$$

Example 4.3.2. 如果约束方程仅有一个, 即 $g: \mathbb{R}^n \rightarrow \mathbb{R}, g(x) = 0$. 那么极值点满足

$$df|_x = \lambda \cdot dg|_x.$$

根据线性函数与法向量的等同,

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right) = \lambda \cdot \left(\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n} \right).$$

即 ∇f 和 ∇g 平行. (这时微分的限制达到最高, 但是 g 本身的限制只有一个.)

如果固定 z , 考虑在 $f(x) = d(x, z)^2$ 在曲面上的极值. 那么

$$\nabla f|_x = 2(x_1 - z_1, \dots, x_n - z_n)$$

与 x 和 z 的连线平行, 而 $\nabla g|_x$ 就是曲面在 x 处的法向量. 从而 f 的极值点有一个很直观的几何刻画: 极值点 x 应满足曲面 $g^{-1}(0)$ 在 x 处的法向量就在 x 和 z 的连线上. 我们可以很容易从平面, 球面上证明这里的几何直观是成立的.

Example 4.3.3. 固定 $\triangle ABC$, P 在三角形内部. 从 P 到三边的垂足为 A', B', C' , 找到点 P 使得

$$|PA'| \cdot |PB'| \cdot |PC'|$$

最大. 这个命题首先可以进行转化:

$$|PA'| \cdot |PB'| \cdot |PC'| = \frac{8|S_{\triangle PBC}| \cdot |S_{\triangle PAB}| \cdot |S_{\triangle PCA}|}{|AB| \cdot |BC| \cdot |CA|}.$$

那么问题转化为已知 $a + b + c = t$, 求 abc 的最大值. 这是 $AM - GM$ 不等式, 但是也可以用 Lagrange 乘子法解决. 记

$$f(a, b, c) = abc, \quad g(a, b, c) = a + b + c.$$

那么要在子流形 $g^{-1}(t)$ 上求极值. 极值点应满足 $\nabla f|_x$ 和 $\nabla g|_x$ 共线, 即

$$(bc, ac, ab) = \lambda \cdot (1, 1, 1).$$

从而 $bc = ac = ab$, 即 $a = b = c$. 这就证明了均值不等式等号成立在 $a = b = c$ 时取到. 对一般的 n 元均值不等式, 也能用相同的方法得到结论.

Theorem 4.3.4 (反函数定理的推广). 设 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是两个 d 维子流形, $f: M \rightarrow N$ 是光滑映射. 设 M 上有一点 p , $f(p) = q$. 假设

$$df|_p: T_p M \rightarrow T_q N \text{ 可逆};$$

那么存在 \mathbb{R}^m 中包含 p 的开邻域 U , \mathbb{R}^n 中包含 q 的开邻域 V 使得

$$f|_{U \cap M}: U \cap M \rightarrow V \cap N$$

是微分同胚, 且 $f|_{U \cap M}$ 的逆存在光滑.

Remark. 这个命题和原本的反函数定理是类似的, 我们可以看到 M 在 p 附近以及 N 在 q 附近都是 \mathbb{R}^d , 那么 f 在局部上就退化为 $\mathbb{R}^d \rightarrow \mathbb{R}^d$ 的反函数定理. 换句话说, 反函数定理是局部性质, 这就已经足以让我们把它推广到 d 维流形上.

类似地, 还有隐函数定理的推广.

Proposition 4.3.5 (隐函数定理的推广). 设 $M \subset \mathbb{R}^m$, $N \subset \mathbb{R}^n$, $f: M \rightarrow N$ 光滑. 设 $S \subset N$ 是 N 的子流形, 并且对任意 $x \in f^{-1}(S)$ 均有 $\text{rk } df|_x = \dim N$; 那么 $f^{-1}(S)$ 是 M 的子流形, 并且

$$\text{codim}_M f^{-1}(S) = \text{codim}_N S \Rightarrow \dim M - \dim f^{-1}(S) = \dim N - \dim S.$$

证明也是和上面一样的, 在局部上 M 和 N 都是欧氏空间, 然后 S 和 $f^{-1}(S)$ 就可以看作相对于这个空间的子流形, 然后利用欧氏空间中的隐函数定理.

4.4 Hesse 矩阵与凸函数

设 $f \in C^2(\Omega, \mathbb{R})$, 其中 $\Omega \subset \mathbb{R}^n$, 接下来要考虑的是 f 的二阶导. 首先 df 是余切向量场, 也就是说 df 作用在一个 (光滑) 切向量场上, 给出一个 (光滑) 函数. 那么 d^2f 是什么? 把 d 作用在 df 上, 现在选定一个切向量场, $d(df)(X)$ 应当是一个余切向量场, 把它作用在向量场 Y 上得到一个函数. 也就是说, 输入两个光滑切向量场, 吐出一个光滑函数. 在每个点附近看, $d^2f|_p$ 就是 $T_p M$ 上的双线性函数.

选定局部坐标 (x^1, \dots, x^n) , 那么 df 可以写为

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n.$$

非常自然地定义 d^2f 的方式是:

$$d^2f = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) dx_j \otimes dx_i.$$

那么它的确和我们刚刚说的是同一个东西. 由于 $f \in C^2$, 所以根据偏导数可交换性, 上式的 $dx_j \otimes dx_i$ 和 $dx_i \otimes dx_j$ 系数相等, 从而也可以利用对称张量

$$dx_i dx_j = \frac{1}{2}(dx_i \otimes dx_j + dx_j \otimes dx_i)$$

写作

$$d^2f = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) dx_i dx_j.$$

从而 $d^2f|_p$ 是对称双线性函数 (但不一定是内积, 因为可能不满足正定性). 无论如何, 双线性函数具有矩阵形式, 称为 **Hesse 矩阵**. 如果考虑 $d(\nabla f)$, 由于 ∇f 可以看作一个 $\Omega \rightarrow \mathbb{R}^n$ 的光滑映射, 那么 $d(\nabla f)$ 对应的 Jacobi 矩阵就是:

$$H_f := \text{Jac}_{x \mapsto \nabla f|_x} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

当然它也可以被看作双线性函数对应的矩阵, 给定切向量 $\mathbf{v}, \mathbf{w} \in T_p M \cong \mathbb{R}^{n \times 1}$,

$$\mathbf{v}^t H_f(p) \mathbf{w} = d^2 f|_p(\mathbf{v}, \mathbf{w}) = d^2 f|_p(\mathbf{w}, \mathbf{v}).$$

回忆多元 Taylor 公式 3.2.7, 我们截断到二阶.

$$\begin{aligned} f(x) = & f(x_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_0} (x_i - (x_0)_i) + \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_i} \Big|_{x_0} (x_i - (x_0)_i)^2 \\ & + \sum_{i < j} \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_0} (x_i - (x_0)_i)(x_j - (x_0)_j) + o(|x - x_0|^2). \end{aligned}$$

利用微分, 上面的式子可以被大大简化, 而且不依赖于具体坐标了:

$$\begin{aligned} df|_{x_0}(x - x_0) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x_0} (x_i - (x_0)_i); \\ d^2 f|_{x_0}(x - x_0, x - x_0) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_0} (x_i - (x_0)_i)(x_j - (x_0)_j). \end{aligned}$$

所以

$$f(x) = f(x_0) + df|_{x_0}(x - x_0) + \frac{1}{2} d^2 f|_{x_0}(x - x_0, x - x_0) + o(|x - x_0|^2).$$

显然这也和一维的情形是完全相容的. 接下来我们由此研究 f 的极值点.

Proposition 4.4.1. 设 $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^d$ 为开集. 若 $x_0 \in \Omega$ 是 f 的极小值点, 那么 $df|_{x_0} = 0$ 且 $d^2 f|_{x_0}$ 半正定. (由于 $d^2 f|_{x_0}$ 对称, 所以这等价于特征值非负)

Proof. $df|_{x_0} = 0$ 在之前 4.3.1 已经证明过, 关键是附带的二阶导信息. 根据 Taylor 公式,

$$0 \leq f(x) - f(x_0) = \frac{1}{2} d^2 f|_{x_0}(x - x_0, x - x_0) + o(|x - x_0|^2).$$

对任意方向 \mathbf{v} , 根据双线性性可知

$$d^2 f|_{x_0}(\lambda \mathbf{v}, \lambda \mathbf{v}) = \lambda^2 d^2 f|_{x_0}(\mathbf{v}, \mathbf{v}).$$

由此我们可以取 $x = x_0 + \lambda \mathbf{v}$, 则 Taylor 公式化归为

$$\lambda^2 \cdot \frac{1}{2} d^2 f|_{x_0}(\mathbf{v}, \mathbf{v}) + o(\lambda^2 |\mathbf{v}|^2) \geq 0.$$

两边除以 λ^2 后再令 $\lambda \rightarrow 0$ 可得 $d^2 f|_{x_0}(\mathbf{v}, \mathbf{v}) \geq 0$, 即 $d^2 f|_{x_0}$ 半正定. $\hat{\mathbb{V}}$

和一维情形一样地: $df|_{x_0} = 0$ 是极值点的必要非充分条件, 但是如果再加上 $d^2 f|_{x_0}$ 严格正定, 就可以反推出极值点.

Proposition 4.4.2. 若 $f \in C^2(\Omega)$, $x_0 \in \Omega$, $df|_{x_0} = 0$, 那么如果 $d^2 f|_{x_0}$ 正定, 则 x_0 是 f 的局部最小值点.

Proof. 只需要把 Taylor 公式反过来用, 证明存在常数 C 使得对任意 $\mathbf{v} \in \mathbb{R}^n$ (现在讨论的 Ω 是一个 \mathbb{R}^n 中的区域, 不是一个流形了, 所以切空间可以被自然等同起来.) 均有

$$d^2 f|_{x_0}(\mathbf{v}, \mathbf{v}) \geq C \cdot |\mathbf{v}|^2,$$

这样就存在 ε 使得当 $|\mathbf{v}| < \varepsilon$ 时

$$\frac{o(|\mathbf{v}|^2)}{|\mathbf{v}|^2} < \frac{C}{4} \Rightarrow f(x) - f(x_0) > \frac{C}{2} |\mathbf{v}|^2 - \frac{C}{4} |\mathbf{v}|^2 > 0.$$

从而 x_0 是局部最小值点 (甚至可以证明是局部唯一最小值点, 和一维一样). 接下来的问题完全转化为了高代问题: 题目中有两个内积, 一个是 \mathbb{R}^n 上的标准内积, 另一个是 $d^2 f|_{x_0}$ (正定性让它确实成为了一个内积). 它们诱导的范数分别记为 $\|\cdot\|_1$ 和 $\|\cdot\|_2$, 那么只需证明存在常数 $C > 0$ 使得

$$\|\mathbf{v}\|_2 \geq C \cdot \|\mathbf{v}\|_1, \quad \forall \mathbf{v} \in \mathbb{R}^n.$$

取在第二个范数下的标准正交基 $\{\mathbf{e}_i\}_{1 \leq i \leq n}$, 则对任意 $\mathbf{v} \in \mathbb{R}^n$,

$$\begin{aligned} \|\mathbf{v}\|_1 &= \left\| \sum_{i=1}^n \langle \mathbf{e}_i, \mathbf{v} \rangle_2 \mathbf{e}_i \right\|_1 \leq \sum_{i=1}^n |\langle \mathbf{e}_i, \mathbf{v} \rangle_2| \|\mathbf{e}_i\|_1 \\ &\leq \sqrt{\left(\sum_{i=1}^n |\langle \mathbf{e}_i, \mathbf{v} \rangle_2|^2 \right) \left(\sum_{i=1}^n \|\mathbf{e}_i\|_1^2 \right)} \\ &= \|\mathbf{v}\|_2 \cdot \sqrt{\sum_{i=1}^n \|\mathbf{e}_i\|_1^2}. \end{aligned}$$

后者是与 \mathbf{v} 无关的常数, 证毕. \square

Remark. 后半部分证明实际上说的是“欧氏空间中的任意两个范数等价”.

下面尝试将凸函数的定义推广到高维:

Definition 4.4.3 (凸集). 凸集 $\Omega \subset \mathbb{R}^n$ 满足 $\forall x, y \in \Omega, 0 \leq t \leq 1, tx + (1-t)y \in \Omega$.

Definition 4.4.4 (凸函数). 给定凸集 Ω , 如果 $f: \Omega \rightarrow \mathbb{R}$ 满足对任意 $x, y \in \Omega$, $t \in [0, 1]$ 有

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y),$$

则称 f 为凸函数. 如果 $\forall t \in (0, 1)$, 有

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y),$$

那么称 f 为严格凸的.

高维情形仍然有 Jensen 不等式成立:

Proposition 4.4.5. 设 Ω 为 \mathbb{R}^d 中的凸集, 则对任意 $x_1, \dots, x_n \in \Omega$, $t_1, \dots, t_n \in [0, 1]$, $t_1 + \dots + t_n = 1$, 则

$$f(t_1 x_1 + \dots + t_n x_n) \leq t_1 f(x_1) + \dots + t_n f(x_n).$$

我们知道一维情形中凸函数是连续函数很好证明, 在高维情形这个命题也成立.

Proposition 4.4.6. 若 $\Omega \subset \mathbb{R}^n$ 为凸开集, $f: \Omega \rightarrow \mathbb{R}$ 为凸函数, 则 f 是连续函数.

Proof. 利用一维情形可知 f 在每个方向上都是连续函数, 关键在于在还不足以推出 f 连续. 首先对 Ω 作平移后可以不妨设 $0 \in \Omega$, 然后不妨缩放整个区域使得所有单位向量 $e_i \in \Omega$. 待定 $R > 0$, 考虑 0 周围的一个小邻域 $B_R(0)$. 只需证明 $|f(x) - f(0)|$ 可被 $|x|$ 控制. 做法的灵感来源于对 $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$, 当 x, y 模长较小时, 可以用 x, e_1, \dots, e_n 的重心坐标表示 y , 用 y, e_1, \dots, e_n 的重心坐标表示 x , 从而可以利用凸函数的条件. 其表达式为:

$$y = \frac{1 - \sum_i y_i}{1 - \sum_i x_i} x + (y_1 - \frac{1 - \sum_i y_i}{1 - \sum_i x_i} x_1) e_1 + \dots + (y_n - \frac{1 - \sum_i y_i}{1 - \sum_i x_i} x_n) e_n.$$

所以我们分别代入 $x = 0$ 和 $y = 0$ 可以得到:

$$\begin{aligned} 0 &= \frac{1}{1 - \sum_i x_i} x - \frac{x_1}{1 - \sum_i x_i} e_1 - \dots - \frac{x_n}{1 - \sum_i x_i} e_n, \\ x &= (1 - \sum_i x_i) \cdot 0 + x_1 e_1 + \dots + x_n e_n. \end{aligned}$$

由于只有在系数均为正的时候才可以利用凸函数性质, 所以可以取 $e_i^x = \pm e_i$ 使得 $x = \sum_i |x_i| e_i^x$, 从而

$$\begin{aligned} 0 &= \frac{1}{1 + \sum_i |x_i|} x + \frac{|x_1|}{1 + \sum_i |x_i|} (-e_1^x) + \dots + \frac{|x_n|}{1 + \sum_i |x_i|} (-e_n^x), \\ x &= (1 - \sum_i |x_i|) \cdot 0 + |x_1| e_1^x + \dots + |x_n| e_n^x. \end{aligned}$$

现在可以利用凸函数性质, 得到

$$\begin{aligned} f(0) &\leq \frac{1}{1 + \sum_i |x_i|} f(x) + \frac{|x_1|}{1 + \sum_i |x_i|} f(-e_1^x) + \dots + \frac{|x_n|}{1 + \sum_i |x_i|} f(-e_n^x), \\ f(x) &\leq (1 - \sum_i |x_i|) \cdot f(0) + |x_1| f(e_1^x) + \dots + |x_n| f(e_n^x). \end{aligned}$$

从而

$$\begin{aligned} f(0) - f(x) &\leq - \sum_i |x_i| f(0) + |x_1| f(-e_1^x) + \dots + |x_n| f(-e_n^x); \\ f(x) - f(0) &\leq - \sum_i |x_i| f(0) + |x_1| f(e_1^x) + \dots + |x_n| f(e_n^x). \end{aligned}$$

由于 $f(0), f(\pm e_i)$ 都是和 x 无关的常数, 因此存在常数 C 使得

$$|f(0) - f(x)| \leq C \cdot \sum_i |x_i| \leq C \sqrt{n} \cdot |x|.$$

从而 f 在 0 附近连续, 证毕.



Proposition 4.4.7. 若 Ω 是 \mathbb{R}^n 中的凸开集, $f: \Omega \rightarrow \mathbb{R}$ 可微, 则 f 为凸函数当且仅当

$$f(x) \geq f(x_0) + \mathrm{d}f|_{x_0}(x - x_0), \quad \forall x, x_0 \in \Omega.$$

这个的证明就和一维差不多了: 根据 f 为凸函数以及微分的定义,

$$\begin{aligned} f(tx + (1-t)x_0) &= f(x_0 + t(x - x_0)) = f(x_0) + \mathrm{d}f|_{x_0}(t(x - x_0)) + o(t(x - x_0)) \\ &\leq tf(x) + (1-t)f(x_0), \quad \forall 0 \leq t \leq 1, \quad x, x_0 \in \Omega. \end{aligned}$$

将上面的式子适当改造为

$$\begin{aligned} tf(x) - tf(x_0) - t\mathrm{d}f|_{x_0}(x - x_0) - o(t) &\geq 0. \\ \Rightarrow f(x) - f(x_0) - \mathrm{d}f|_{x_0}(x - x_0) - o(1) &\geq 0. \\ (t \rightarrow 0) \Rightarrow f(x) - f(x_0) - \mathrm{d}f|_{x_0}(x - x_0) &\geq 0. \end{aligned}$$

反过来, 注意到条件中 x 和 x_0 都有任意性, 为了把微分消掉可配凑出下式

$$\begin{aligned} f(x) &\geq f(tx + (1-t)x_0) + (1-t)\mathrm{d}f|_{tx+(1-t)x_0}(x - x_0); \\ f(x_0) &\geq f(tx + (1-t)x_0) + t\mathrm{d}f|_{tx+(1-t)x_0}(x_0 - x). \end{aligned}$$

从而两边乘系数相加就得到

$$\begin{aligned} tf(x) + (1-t)f(x_0) &\geq tf(tx + (1-t)x_0) + (1-t)f(tx + (1-t)x_0) \\ &= f(tx + (1-t)x_0). \end{aligned}$$

Proposition 4.4.8. 设 $f \in C^2(\Omega)$, 则 f 为凸函数当且仅当 d^2f 在每个点均半正定. 进一步, 如果 d^2f 在每点处均正定, 那么 f 是严格凸的.

Proof. 将上一题的结论和 Taylor 公式结合起来可得 f 是凸函数当且仅当对任意 $x, x_0 \in \Omega$ 均有

$$\begin{aligned} f(x) &= f(x_0) + \mathrm{d}f|_{x_0}(x - x_0) + \frac{1}{2}\mathrm{d}^2f|_{x_0}(x - x_0, x - x_0) + o(|x - x_0|^2) \\ &\geq f(x_0) + \mathrm{d}f|_{x_0}(x - x_0) \end{aligned} \quad (*)$$

如果 f 是凸函数, 用 $tx + (1-t)x_0$ 代替 x 可知

$$\begin{aligned} f(tx + (1-t)x_0) &= f(x_0) + t\mathrm{d}f|_{x_0}(x - x_0) + \frac{t^2}{2}\mathrm{d}^2f|_{x_0}(x - x_0, x - x_0) + o(t^2) \\ &\geq f(x_0) + t\mathrm{d}f|_{x_0}(x - x_0). \end{aligned}$$

此时不等号成立, 上式等价于 $\mathrm{d}^2f|_{x_0}(x - x_0, x - x_0) + o(1) \geq 0 \Rightarrow \mathrm{d}^2f|_{x_0}(x - x_0, x - x_0) \geq 0$. 根据 x 选取的任意性, d^2f 是半正定的.

反过来, 如果 d^2f 在每点处均正定, 那么在命题 4.4.2 中我们证明了存在常数 C 使得

$$\mathrm{d}^2f|_{x_0}(x - x_0, x - x_0) \geq C \cdot |x - x_0|^2.$$

从而每点 x_0 附近均有 (*) 成立且不取等, 对一般情形只需要把局部的拼起来即可?

考虑一元函数

$$H(t) = f(x_0 + t(x - x_0)) - f(x_0) - t df|_{x_0}(x - x_0)$$

满足 $H''(t) \geq 0$, 从而是凸函数.



5 多元积分与 Stokes 公式

这一章的主要目标来源于定义边界上的 Riemann 积分.

5.1 第一型曲线和曲面积分

Proposition 5.1.1. 给定 $f: M(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$, 给出 $\int_M f$ 的定义.

我们有两个问题需要指出: 其一, 一般来说我们希望在流形上对一个最高次微分形式 ω 来定义积分, 记为 $\int_M \omega$. 定义方式如下: 如果我们把 ω 拉回到某个坐标卡 (φ, U, V) 上

$$\varphi^*(\omega) = f \cdot dx^1 \wedge dx^2 \wedge \cdots \wedge dx^m \quad (1)$$

并在坐标卡上以标准的 Lebesgue 积分来定义微分形式在 U 上的积分:

$$\int_U \omega := \int_V \varphi^* \omega = \int_V f(x) dx^1 \cdots dx^m.$$

由于我们本身是对一个“形式”作积分, 所以在不同的坐标卡选取下, $dx^1 \wedge \cdots \wedge dx^m$ 到 $dy^1 \wedge \cdots \wedge dy^m$ 的变换会自动给出一个过渡矩阵, 事实上可以证明 (在不关心定向的时候) 上述定义配合上单位分解就可以良定义 ω 在流形的积分.

反之, 如果我们直接把 f 拉回到某个坐标卡上化归为 \mathbb{R}^m 上的积分, 那么比如说一条曲线 $\gamma: [0, 1] \rightarrow \mathbb{R}^n$ 显然它可以有不同的参数化, 不同的参数化应当对应的长度是一样的, 而它的表达式 $\int |\gamma'(t)| dt$ 也说明了它是由 M 所处的空间 \mathbb{R}^n 典范地诱导的, 我们需要找到这个诱导出的形式.

其二, 传统定义出的积分并不借助外微分, 从而对定向并不敏感. 比如说用 $ds = \sqrt{(dx)^2 + (dy)^2}$ 定义出的东西并不是 2-form, 它只是在每个点处喂一个切向量进去吐出一个非负实数. 同样地还有 $dS = \sqrt{(dx \wedge dy)^2 + (dy \wedge dz)^2 + (dz \wedge dx)^2}$, 它给两个切向量吐出来一个非负实数. 但这才是我们这里想要的, 我们就把它记作 $d\sigma$. φ^* 仍然可以把 $d\sigma$ 拉回到 V 上, 但是能不能写成 (1) 的形式? 答案是可以的, 如果定义的好. 比如上面的 dS , $\varphi^*(dS)$ 实际上被 $\varphi^*(dS)(x^1, x^2)$ 所唯一决定, 因为每个分量上都具有双重交错线性. 我们最后定义出的 $d\sigma$ 就会具有这样的形式.

回到一维情形, 进一步地考虑一维情形的启示: $\gamma'(t) \in T_p M$, 而 $|\gamma'(t)|$ 指 $\gamma'(t)$ 在嵌入 $T_p M \hookrightarrow \mathbb{R}^n$ 下被赋予的范数. 因此我们可以抽象出一个 $d\ell \in \Omega^1(\mathbb{R}^n)$, 它在每个点 p 上的作用是把 $v \in T_p M$ 打到 $|v|$. 这样的话, 将曲线的像实现为 \mathbb{R}^n 中的

一维子流形 M , 那么 $f: M \rightarrow \mathbb{R}, x \mapsto 1$ 的积分就应当是所谓曲线的长度, 它就可以被定义为 $\int_M d\ell$. 这是因为如果把 γ 实现为 M 的局部坐标, 那么就有

$$\gamma^*(d\ell) = |\gamma'(t)| \cdot dt.$$

只需要展开验证: $\gamma^*(d\ell)$ 把向量 $\mathbf{v} \in T_t([0, 1])$ 打到 $d\gamma|_t(\mathbf{v})$, 然后再被 $d\ell$ 打到 $|d\gamma|_t(\mathbf{v})|$. 由于 $T_t([0, 1])$ 是一维空间, 设其单位向量为 \mathbf{e} , 我们常说的 $\gamma'(t) = d\gamma|_t(\mathbf{e})$, 因而 $|d\gamma|_t(\mathbf{v})| = |\gamma'(t) \cdot dt(\mathbf{v})|$, 因此上式成立. 从而我们有理由相信, 把这样的定义推广到高维定义出一个 $d\sigma$, 然后给定坐标卡 $\varphi: V(\subset \mathbb{R}^m) \rightarrow U(\subset M)$ 并用

$$\int_U f := \int_U f d\sigma = \int_V (f \circ \varphi) \cdot \varphi^*(d\sigma)$$

来定义积分是合理的. 我们想寻找的 $d\sigma$ 应当继承 \mathbb{R}^n 的结构: 它在每个点 p 处把切空间上的 m 个向量 $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m)$ 嵌入到 \mathbb{R}^n 中, 然后打到它们在 \mathbb{R}^n 中张成的体积. 暂时把这个值记为

$$\text{vol}(\mathbf{v}_1, \dots, \mathbf{v}_m).$$

事实上, 如果设

$$\begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_m \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix} A$$

其中 $A \in \mathbb{R}^{n \times m}$, 那么

$$\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_m = \sum \det A_{i_1, \dots, i_m} \cdot \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_m}.$$

其中 A_{i_1, \dots, i_m} 代表 A 的 i_1, \dots, i_m 行围成的 $m \times m$ 矩阵. 通过定义 $\bigwedge^m(V)$ 上的内积, 可以定义 $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ 张成的体积就是

$$\|\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_m\| = \sqrt{\sum (\det A_{i_1, \dots, i_m})^2}.$$

这个表达式难以计算, 实际上利用 Cauchy-Binet 公式就可以给出它的一个好看的形式

$$\sqrt{\sum (\det A_{i_1, \dots, i_m})^2} = \sqrt{\det(A^t A)} = \sqrt{\det(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{1 \leq i, j \leq m}}.$$

其中最后一个等式有更强的推广:

$$\langle \mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_m, \mathbf{w}_1 \wedge \dots \wedge \mathbf{w}_m \rangle = \det(\langle \mathbf{v}_i, \mathbf{w}_j \rangle)_{1 \leq i, j \leq m}.$$

这就是 **Gram 矩阵** 的来源. 从而我们的确定义了一个好的 $d\sigma$:

$$d\sigma = \sqrt{\sum_{i_1 < \dots < i_m} (d\mathbf{x}^{i_1} \wedge \dots \wedge d\mathbf{x}^{i_m})^2}.$$

(注意这不是一个 form.) 现在只需要在给定 (φ, U, V) 的情况下把这个 σ 拉回到 V 上去. 给定局部坐标 (x^1, \dots, x^m) , 那么

$$\varphi^*(d\sigma)(\mathbf{x}^1, \dots, \mathbf{x}^m) = d\sigma(d\varphi(\mathbf{x}^1), \dots, d\varphi(\mathbf{x}^m)).$$

再把 $T_p M$ 嵌入到 \mathbb{R}^n 中 (表示为列向量), 那么 $(d\varphi(\mathbf{x}^1), \dots, d\varphi(\mathbf{x}^m))$ 立刻化为 Jacobi 矩阵 $\text{Jac}_\varphi \in \mathbb{R}^{n \times m}$. 根据上面得到的 m 个向量张成的体积公式,

$$\varphi^*(d\sigma)(\mathbf{x}^1, \dots, \mathbf{x}^m) = \sqrt{\det(\text{Jac}_\varphi^t \cdot \text{Jac}_\varphi)}.$$

从而

$$\varphi^*(d\sigma) = \sqrt{\det(\text{Jac}_\varphi^t \cdot \text{Jac}_\varphi)} dx^1 \dots dx^m.$$

至此, 我们得到了在参数化后积分的表达式:

$$\int_U f := \int_V f(\varphi(x)) \sqrt{\det(\text{Jac}_\varphi^t(x) \cdot \text{Jac}_\varphi(x))} dx^1 \dots dx^m.$$

至于各种良定义性, 它的证明仅依赖于 $\bigwedge^m(V)$ 上内积定义的良好性和微分形式拉回在不同参数选取下的良定性. 这些都是不需要繁杂计算容易验证的.

Remark. 如果我们想要把定向考虑进去, 只需要对这里的 σ 进行一些改装并加上正负性. 回顾定向的定义: 称 M 局部可定向如果可以在每点 p 处选定一个切空间上的定向 o_p , 使得 $d\varphi: \mathbb{R}^m \xrightarrow{\sim} T_p M$ 把 \mathbb{R}^m 上的标准定向送到 o_p . 直观地说, 我们像选定一个“全局标架场”, 使得每个点的局部标架场都是光滑的.

由此我们可以在每个切空间上确实定义一个微分形式 ω , 它符合定向 o_p 并且取值的绝对值是 σ . 比如一维情形下, 如果在 \mathbb{R}^1 的正负两个定向中选择一个 (记为 v), 然后把 $\varphi_*(v)$ 定向为正, 那么每个时刻都 correspond to 这个定向, 所以从 0 积到 1 值不会改变, 但是如果反过来积分每刻就会反定向, 符号就会相反.

特别地, 对余 1 维流形选定底空间的一组定向相当于选定了一个法向量, 因为 $\bigwedge^{n-1}(V)$ 可以通过 star map 对应到 V 上, 从而可以被一个向量直观体现出来. 所以余 1 维曲面可定向当且仅当可以选定一个光滑法向量场, 可以看出 Mobius 带不满足这个要求.

5.1.1 Stokes 公式

Definition 5.1.2 (有界带边区域). $\Omega \subset \mathbb{R}^n$ 是有界带边区域, 如果对任意边界点 $x \in \partial\Omega$, 存在开集 $U, V \subset \mathbb{R}^n$ 以及连续微分同胚 $\Phi: U \rightarrow V$ 使得

$$\Phi(U \cap \Omega) = V \cap \{x_n \geq 0\}, \quad \Phi(U \cap \partial\Omega) = V \cap \{x_n = 0\}.$$

其中 x_n 为 \mathbb{R}^n 中坐标的第 n 个分量.

如果 Φ 是连续可微的, 称 x 为正则点, 否则称为奇异点.

Remark. 这里的定义似乎默认了 $\partial\Omega$ 是余 1 维流形. 并且由于这个区域已经给定了, 所以自然地存在“内部”和“外部”的观念, 这样就可以选定连续的外法向量族, 从而也选定了一个合法的定向.

此外, 如果使用 Hausdorff 测度, 那么我们希望 Ω 是 n 维的, $\partial\Omega$ 是 $n-1$ 维的, 而奇异点相对于 $\partial\Omega$ 来说是零测的, 所以它应当是类似于 $n-2$ 维的东西, 这就和 Hausdorff 维数连接起来了.

Theorem 5.1.3 (The REAL Stokes Formula). 设 M 是一个 m 维带边光滑定向流形, 其边界上赋予 M 的诱导定向. 则对任意紧支光滑形式 $\omega \in \Omega^{m-1}(M)$, 有

$$\int_{\partial M} \iota_{\partial M}^* \omega = \int_M d\omega.$$

其中 $\iota_{\partial M}: \partial M \hookrightarrow M$ 是包含映射.

这里讨论的命题不涉及微分形式, 并且只考察 M 是某个 \mathbb{R}^n 的子空间的情形. 也就是说, M 是嵌入 \mathbb{R}^n 的 n 维流形, 而 ∂M 是 $n-1$ 维流形. 此时, 在空间 \mathbb{R}^n 中存在单位外法向量的概念: 这是因为每个 $n-1$ 维切空间在嵌入 \mathbb{R}^n 时都典范对应到一个 1 维子空间, 而“区域”又严格定义了内部和外部的概念.


Lemma 5.1.4 (单位外法向量). 设 $\Omega \subset \mathbb{R}^n$ 是有界带边区域, 那么对任意的正则点 $p \in \partial\Omega$, 存在 \mathbb{R}^n 中唯一的单位向量 $v(p)$ 使得

- $v(p) \perp T_p \partial\Omega$;
- $|v(p)| = 1$;
- $\exists \varepsilon > 0$ s.t. $p - tv(p) \in \Omega, \forall 0 \leq t < \varepsilon$.

它就是 p 处的单位外法向量.

Proof. 按定义拉回到平坦空间中即可. 记 f 在 V 中把 (x_1, \dots, x_n) 打到 x_n , 那么 ∇f 是指向内部 (上半超平面) 的法向量, 那么把它拉回到 U 上得到 $\Phi_*(f)$, 然后考虑

$$v(p) = -\frac{\nabla_p(\Phi_* f)}{|\nabla_p(\Phi_* f)|}.$$

注意这里的定向和 Φ 的定义有关, 如果 Φ 把区域内部打到了负半超平面, 那么定义出 f 拉回后就不需要带负号. 

Example 5.1.5 (方块的 Stokes 公式). \mathbb{R}^n 中的方体 $I = [a_1, b_1] \times \dots \times [a_n, b_n]$. 若 Φ 为定义在 \mathbb{R}^n 上的连续可微函数, 则

$$\begin{aligned} \int_I \frac{\partial \Phi}{\partial x_n} dx &= \int_{I_{n-1}} \left(\int_{a_n}^{b_n} \frac{\partial \Phi}{\partial x_n} dx_n \right) dx_1 \dots dx_{n-1} \\ &= \int_{I_{n-1}} (\Phi(x_1, \dots, x_{n-1}, b_n) - \Phi(x_1, \dots, x_{n-1}, a_n)) dx_1 \dots dx_{n-1} \\ &= \int_{\partial I} \Phi \cdot \nu_n dx_1 \dots dx_{n-1} \end{aligned}$$

其中 ν_n 为点 $x \in \partial I$ 处的单位外法向量的第 n 个分量. 这个 ν_n 只留下了 $[a_1, b_1] \times \dots \times [a_{n-1}, b_{n-1}] \times \{a\}$ or $\{b\}$ 对应的像.

Theorem 5.1.6 (Stokes). 对任意有界带边区域 D ,

$$\int_D \frac{\partial \Phi}{\partial x_i} dx = \int_{\partial D} \Phi \cdot \nu_i d\sigma.$$

其中 $d\sigma$ 为边界测度, ν_i 是 x 处的外指法向量的第 i 个分量.

我们可以用 Real Stokes 公式来推出这个结论：对函数 $\Phi : D \rightarrow \mathbb{R}$ ，考虑在 5.1.3 中代入

$$\omega = \Phi \cdot dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$$

那么 Real Stokes 公式得到

$$\int_{\partial D} \Phi dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n = \int_D (-1)^{i-1} \frac{\partial \Phi}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n.$$

因为其它的项都已经被消去. 那么只需要证明

$$dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n = (-1)^{i-1} \nu_i \cdot d\sigma.$$

设切空间的一组（保定向）基为 e_1, \dots, e_{n-1} ，那么可设

$$e_1 \wedge \cdots \wedge e_{n-1} = \sum_{i=1}^n a_i \cdot x^1 \wedge \cdots \wedge \widehat{x^i} \wedge \cdots \wedge x^n.$$

那么 $d\sigma$ 转化系数为 $\sqrt{\sum_{i=1}^n a_i^2}$ ，而 $dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$ 转化系数为 a_i . 另一方面，两边取 Hodge star map 可以把 $n-1$ 维外积和向量进行对应，这就得到

$$n = \sum_{i=1}^n a_i \cdot x^i.$$

从而单位外法向量的第 i 分量系数为 $\frac{a_i}{\sqrt{\sum a_i^2}}$ ，所以

$$\nu_i \cdot \sqrt{\sum a_i^2} = a_i.$$

这就证明了上面的定理.（至少说明它是对的）.

Remark. 证明的想法就是把 D 分成一些小方块，对小方块上的积分转化为其边界上的积分，而 D 内部的小方块公共部分的积分由于法向量反向所以互相抵消，而在边界上不能抵消，所以留下了边界上的值. 拆分的过程利用单位分解把积分拆为小区域上的积分，然后对于那些比较糟糕的方块，可以把它再细分之和证明可以和标准小方体之间建立同胚，因而也是满足要求的. 与 $\int_D d\Phi = \int_{\partial D} \Phi$

大部分曲面没有办法用一个坐标系写出来，我们要使用单位分解把它 reduce 到一个小坐标卡内的积分.

Lemma 5.1.7. 设 K 为 \mathbb{R}^n 中紧集， $K \subset U$ ， U 为开集. 则存在光滑函数 f 使得其支集落在 U 中，并且在 K 上有 $f > 0$.（也可以要求 f 在 K 上取值为 1）

Proof. 一方面，根据 K 紧且 U^c 闭，则 $\text{dist}(K, U^c) > 0$ ，即这两个区域可以完全被一些“等高线”完全分离开. 记 $\delta = \frac{1}{4}\text{dist}(K, U^c)$ ，取 $K \subset U_1 = \{x \in \mathbb{R}^n, \text{dist}(x, K) < \delta\}$ 为开集. 首先存在一个满足要求的连续函数 f （在一般的拓扑空间中，是 Urysohn 引理）：

$$f(x) = \frac{\text{dist}(x, U_1^c)}{\text{dist}(x, K) + \text{dist}(x, U_1^c)}$$

这个 f 在 K 上为 1 且在 U_1^c 上为零. 接下来要把它光滑化, 只需要找一个光滑函数并且把它卷积上这个 f 即可. 如果更一般地希望在 K 上取值为 1, 那么可以通过适当的调整把它变成 K 上取值为 1 的函数. 为此我们再细分 U_1 为 $K \subset U_2 \subset K_1 \subset U_1$, 然后取光滑的 f 在 K_1 上恒大于 0, 在 U_1^c 上恒为零. 因此我们可取光滑函数 g 在 K 上恒为 0, 在 U_2^c 上恒为 1. 从而函数

$$h := \frac{f}{f+g}$$

是良定义的, 并且在 K 上取值恒为 1.



单位分解希望对于给定的一族流形的图册 $\{U_i\}$, 将函数 1 分解为一些函数的和, 使得每个函数的支集都完全包含在某个 U_i 中, 然后把每个 U_i 拉回到欧氏空间上即可. 严格来说就是下面的引理:

Lemma 5.1.8 (紧集上的单位分解). 设 $K \subset \mathbb{R}^n$ 为紧集, $\{U_1, \dots, U_k\}$ 是 K 的覆盖, 那么存在光滑函数 $f_1(x), \dots, f_k(x)$ 使得

$$1 = f_1(x) + \dots + f_k(x), \quad \forall x \in K$$

并且 $\text{supp } f_i \subset U_i$, $1 \leq i \leq k$.

Proof. 根据上面的引理, 可以选择一些 g_i 在 $K_i \subset U_i$ 上为正并且 $\text{supp } g_i \subset U_i$.

我们希望取一个函数 $g(x)$ 在所以我们对 U_i 进行细分定义开集

$$K_i^\delta = \{x \in U_i : \text{dist}(x, U_i^c) > \delta\}$$



再证明 Stokes 性质是保微分同胚的, 即如果 Φ 是从方体 I 到 Ω 的 C^1 同胚, 那么 Ω 满足 Stokes 公式.

Theorem 5.1.9. I 是 \mathbb{R}^n 中的方体, Φ 是从 I 到 Ω 的 C^1 的同胚, 那么 Ω 满足 Stokes 公式.

Proof. 根据积分换元公式, 由于 $x = \Phi(y)$, $y \in I$, 因此对任意 $\varphi : I \rightarrow \mathbb{R}$,

$$\int_{\Phi(I)} d\varphi|_x dx = \int_I d\varphi|_{\Phi(y)} d\Phi(y) = \int_I d(\varphi \circ \Phi)|_y \circ d\Phi^{-1}|_x |\text{Jac}_\Phi(y)| dy.$$

不妨假设 $\text{Jac}_\Phi(y) > 0$, 固定 y , 考虑矩阵 $A := \text{Jac}_\Phi(y) = \left(\frac{\partial \Phi_i}{\partial y_j} \right)_{1 \leq i \leq n, 1 \leq j \leq n}$. 设 A 的逆矩阵元素为 $(A^{-1})_{ij} = (\det A)^{-1} A_{ij}^*$, 根据偏导数交换性 $\partial_{y_k} A_{ij} = \partial_{y_j} A_{ik}$, 可以证明技术上的命题

$$\partial_{y_j} (A^{ji} \det A) = 0.$$

然后代入原式可以消去其中一些项,

$$d(\varphi \circ \Phi)|_y \circ d\Phi^{-1}|_x |\text{Jac}_\Phi(y)| dy = \int_I \varphi(\Phi(y)) A^{ji} \det A \nu_j d\sigma$$

要证明

$$\int_{\Phi(I)} \frac{\partial \varphi}{\partial x_i} dx = \int_{\partial \Phi(I)} \varphi \cdot \nu_i d\sigma$$

只需证明

$$\int_{(\partial I)_i} \varphi(\Phi(y)) A^{ji} \det A \cdot \nu_j d\sigma = \int_{\Phi(\partial I)_i} \varphi \cdot \tilde{\nu}_i d\sigma$$



5.2 曲线和曲面积分

按传统说法积分分为**第一型**和**第二型**两种. 然而根据其具体运算形式, 我们可以这么说:

- 第一型积分是指按“子流形测度”来积分, 常用记号为 $\int f d\sigma$, 这里的 $d\sigma$ 也被称作体元素, 它在每个切空间上 input k 个向量 (流形维数) output 一个非负实数 (它们张成的体积), 特别地在作为子流形, 其诱导的题源素把切空间中的 k 个向量打到它们嵌入 \mathbb{R}^n 中张成的 k 维图形的体积, 可以写成下述投影的形式

$$d\sigma = \sqrt{\sum_{i_1 < \dots < i_k} (dx_{i_1} \wedge \dots \wedge dx_{i_k})^2}.$$

这有多种理解方式, 比如说可以看成“投影在 $\binom{n}{k}$ 个分量上所得的面积构成的向量模长”, 或者“在 $\bigwedge^k(\mathbb{R}^n)$ 上赋予以 $dx_{i_1} \wedge \dots \wedge dx_{i_k}$ 维标准正交基后诱导出的范数的取值”, 或者“这些向量构成的 Gram 矩阵 $(\langle v_i, v_j \rangle)$ 的行列式. 总之, 它和子流形的定向无关, 并且在不同的参数替换下所得结果差一个 $|\det \text{Jac} \Phi_2^{-1} \circ \Phi_1|$, 这可以追溯到 Lebesgue 积分中的参数替换所差的系数就是 Jacobi 矩阵行列式的绝对值.

- 第二型积分是指“对形式积分”, 常用记号为 $\int \omega$. 其中, 对于嵌入在 \mathbb{R}^n 内的 k 维流形, 我们将对 k 维形式作积分, 此时 ω 有具体表达形式

$$\omega = \sum_{i_1 < \dots < i_k} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}. \quad f_{i_1 \dots i_k} : M \rightarrow \mathbb{R}.$$

对形式积分最大的特点在于与定向相关. 即如果选择两个不同的坐标卡, 那么它们之间的系数会差一个 $|\det \text{Jac} \Phi_2^{-1} \circ \Phi_1|$, 没有绝对值. 这就导致选择不同的到平坦空间的拉回可能给出不同的 $\int \omega$. 为了让 ω 良定义, 我们发现只能对可定向曲面选取一个定向后进行积分. 接下来将具体探讨“定向”选取的定义以及形式的积分和一般而言所说的第二型曲面积分之间的关系.

定向选取：有很多种定义定向的方法，个人认为标架法最容易利用. 每个点处的切空间 $T_p M$ 是一个 k 维线性空间，一个好的定向是对每个切空间选取一个定向，使得每个点处对应的定向满足相容性条件. 所以规定定向的方式分两步.


首先是考虑 \mathbb{R}^k 的两种标架定向. 对两个标架 $\{e^1, \dots, e^k\}$ 和 $\{\tilde{e}^1, \dots, \tilde{e}^k\}$ ，很自然地如果这两个标架之间的过度矩阵行列式为正则二者同定向，若行列式为负则二者反定向. 且在给定坐标系时， $\{x^1, \dots, x^k\}$ 成为典范的正标架类. 还有另外一种理解方式：选取非零的 $\omega \in \bigwedge^k \mathbb{R}^k$ ，考虑 $e^1 \wedge \dots \wedge e^k = a\omega$. 如果 $a > 0$ 则称为正定向， $a < 0$ 则称为负定向，它们构成了 $\bigwedge^k \mathbb{R}^k$ 的两个连通分支. 我们把 ω 所在定向类记作 $[\omega]$.

在嵌入 \mathbb{R}^n 后这两种定义仍然可用，比如 1 维时曲线的两个方向是非常清楚的，且和这里的定义相契合. 与之类似地，在余 1 维时我们也可以直接利用法向来指定定向. 每一点处恰好有两个方向的法向量，通过选取 \mathbb{R}^n 的典范定向 $\epsilon = x^1 \wedge \dots \wedge x^n$.

Definition 5.2.1 (用法向量定向). 我们称 $\omega \in T_p M$ 和 n 定向相同如果 $[n \wedge \omega] = [\epsilon]$. 一般来说， k 重外积和 $n - k$ 重外积之间可以使用 Hodge star map 来转化定向.

接下来从局部定向过渡到整体定向. 首先在平坦空间中，每个点处的切空间可以通过平移典范地对应在一起，那么只需要它们定向相同就可以了，这就完成了平坦空间中定向相容的定义. 对于一般的流形，在每点 p 附近存在一个同胚于欧氏的邻域，我们要求这个邻域上每点的定向在拉回到平坦空间后是相容的即可. 换言之， p 这一点处的定向可以决定它的邻域上的定向.

Proposition 5.2.2. k 维光滑流形 M 是可定向的当且仅当 M 上存在处处非零的光滑 k 形式 μ .

Proof. 右推左是容易的（因为变化连续），左推右利用单位分解把局部上的好形式拼起来即可. 

Corollary 5.2.3. 余 1 维光滑流形 M 是可定向的当且仅当 M 上存在处处非零的光滑法向量场.

接下来研究这样的定向定义是否和传统的一些定义是相容的. 我们先来看 Stokes 公式，一般来说 Stokes 公式可以表达为

$$\int_M d\omega = \int_{\partial M} \omega.$$

这涉及到在边界上的诱导定向的选取. 具体选取可以通过“ ω 支在微分同胚于 \mathbb{R}_+^n 的坐标卡 U 中”的 Stokes 公式的证明看出来. 设

$$M = \{(x^1, \dots, x^n) \in \mathbb{R}^n : x^1 \geq 0\}, \quad \omega = \sum_{i=1}^n (-1)^{i-1} f_i dx^1 \wedge \dots \wedge \widehat{dx^i} \wedge \dots \wedge dx^n.$$

那么可以求出 ∂M 和 $d\omega$ 的表达式：

$$\partial M = \{(x^1, \dots, x^n) \in \mathbb{R}^n : x^1 = 0\}, \quad d\omega = \sum_{i=1}^n \frac{\partial f_i}{\partial x^i} dx^1 \wedge \dots \wedge dx^n.$$

以定向为 $\mathbf{x}^1 \wedge \cdots \wedge \mathbf{x}^n$ 的坐标卡计算 $\int d\omega$:

$$\int_M d\omega = \int_{\mathbb{R}^{n-1}} \int_0^{+\infty} \frac{\partial f_1}{\partial \mathbf{x}^1} dx^1 \cdots dx^n + \sum_{i=2}^n \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{+\infty} \frac{\partial f_i}{\partial \mathbf{x}^i} dx^i \right) dx^1 \cdots \widehat{dx^i} \cdots dx^n.$$

根据 ω 的紧支性可知 f_i 都是紧支的, 因此后一个求和内部是零, 前一个求和内部会得到 $-f_1(0, x^2, \dots, x^n)$. 所以

$$\int_M d\omega = \int_{x_2, \dots, x_n \in \mathbb{R}} -f_1(0, x_2, \dots, x_n) dx^2 \cdots dx^n.$$

另一方面, 现在 f_1 的取值 $(0, x^2, \dots, x^n)$ 就是 ∂M , 事实上如果以 $\mathbf{x}^2 \wedge \cdots \wedge \mathbf{x}^n$ 维定向计算 $\int_{\partial M} \omega$, 可以得到

$$\int_{\partial M} \omega = \int_{x_2, \dots, x_n \in \mathbb{R}} f_1(0, x_2, \dots, x_n) dx^2 \cdots dx^n = - \int_M d\omega.$$

所以诱导定向应该是反过来的: 即以 $-\mathbf{x}^2 \wedge \cdots \wedge \mathbf{x}^n$ 作为 ∂M 的诱导定向. 由于实际上可以认为 \mathbf{x}^1 指向了 M 的内部, 而 $-\mathbf{x}^1$ 指向了 M 的外部, 因此可以这样形容诱导定向:

“给定 M 上的定向 $[\omega]$, 它在 $p \in \partial M$ 上的诱导定向是取 p 上符合 M 定向的标架并让第一个分量外指, 剩下的部分就构成了 ∂M 的切空间上的标架并确定出一个定向.”

特别地, 当 M 是在 \mathbb{R}^n 中的 n 维流形时, 自动选取典范的定向 $[\epsilon]$; ∂M 是余 1 维的, 上面的第一个分量外指就实现为该点处的一个“外法向量”, 根据法向量定向的定义, 选取外法向量直接给出了 ∂M 上的定向. 这也就给出了之前 Stokes 公式中外法向量出现的合理性.

下一个问题所谓“和法向作内积”来定义的第二型曲面积分是否和上面所说的定向是相容的.

$$\int_L P dx + Q dy \stackrel{?}{=} \int_L \langle (Q, -P), \mathbf{n} \rangle d\sigma = \int_L \langle (Q, -P), d\sigma \rangle. \quad (1)$$

在一维情形下, 单位切向量也可以用来定义方向, 所以人们也会给出下面的定义式

$$\int_L P dx + Q dy \stackrel{?}{=} \int_L \langle (P, Q), \boldsymbol{\tau} \rangle d\sigma. \quad (2)$$

(1) 可以推广到所有余 1 维子流形, (2) 可以推广到所有 1 维子流形. 我们先来看 (2) 的推广: 选定一条曲线和一个定向, $\boldsymbol{\tau}$ 代表选定方向上曲线的单位切向量. 考虑这条曲线的一个符合定向的参数化 $\gamma: [0, 1] \rightarrow L$, 记 $\gamma(t) = (x^1(t), \dots, x^n(t))$. 分别按定义写下两个表达式:

$$\begin{aligned} \int_L \sum_{i=1}^n f_i dx^i &= \int_0^1 \sum_{i=1}^n f_i(x^i)'(t) dt. \\ \int_L \langle (f_1, \dots, f_n), \boldsymbol{\tau} \rangle d\sigma &= \int_0^1 \left\langle (f_1, \dots, f_n), \frac{((x^1)'(t), \dots, (x^n)'(t))}{|\gamma'(t)|} \right\rangle \cdot |\gamma'(t)| dt \\ &= \int_0^1 \sum_{i=1}^n f_i(x^i)'(t) dt. \end{aligned}$$

从而二者相等，即把 (2) 作为定义式是合理的。

再来讨论 (1) 的推广：先写出推广形式：我们证明如果 \mathbf{n} 确定的定向和 $[\omega]$ 确定的定向一样，那么

$$\int_M \sum_{i=1}^n (-1)^{i-1} f_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n = \int_M \langle (f_1, \dots, f_n), \mathbf{n} \rangle d\sigma.$$

前一个式子中的定向隐藏在定向当中（不是显性的），后一个式子在 $d\sigma$ 上没有定向，定向完全包含在 \mathbf{n} 的信息当中。还是只需要看支在一个坐标卡的像中的情形。选择一个和定向相容的坐标卡 $\Phi: U \rightarrow M$ ，其中 U 上的坐标用 y^1, \dots, y^{n-1} 来表示。首先初步化简两边的式子：设

$$\frac{\partial \Phi}{\partial y^1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial y^{n-1}} = \sum_{i=1}^n a_i \mathbf{x}^1 \wedge \cdots \wedge \widehat{\mathbf{x}^i} \wedge \cdots \wedge \mathbf{x}^n. \quad (3)$$

那么

$$\begin{aligned} \int_M \sum_{i=1}^n (-1)^{i-1} f_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n &= \int_U \sum_{i=1}^n (-1)^{i-1} f_i(\Phi(y)) \cdot a_i dy^1 \cdots dy^{n-1}. \\ \int_M \langle (f_1, \dots, f_n), \mathbf{n} \rangle d\sigma &= \int_U \sum_{i=1}^n f_i(\Phi(y)) \mathbf{n}_i(\Phi(y)) \cdot \sqrt{a_1^2 + \cdots + a_n^2} dy^{n-1} \cdots dy^1. \end{aligned}$$

问题归结于 \mathbf{n} 的具体表达式。实际上，对 (3) 的两边作在左侧的 Hodge 对偶，即把 $\omega \in \bigwedge^k(V)$ 对应到唯一的 $\star \omega \in \bigwedge^{n-k}(V)$ 使得

$$(\star \omega) \wedge \eta = \langle \omega, \eta \rangle \epsilon, \quad \forall \eta \in \bigwedge^k(V).$$

其中 ϵ 是事先选定好的 $\bigwedge^n(V)$ 中的体积形式，这里就取为 $\mathbf{x}^1 \wedge \cdots \wedge \mathbf{x}^n$ 。这样，把 (3) 两边作用 \star 后的结果记为 \mathbf{t} ，那么根据 RHS：

$$\mathbf{t} = \sum_{i=1}^n (-1)^{i-1} a_i \mathbf{x}^i.$$

再根据 LHS 可知 \mathbf{t} 和 $\frac{\partial \Phi}{\partial y^j}$ 张成的超平面（也就是切空间）垂直，并且

$$\mathbf{t} \wedge \frac{\partial \Phi}{\partial y^1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial y^{n-1}} = \mathbf{x}^1 \wedge \cdots \wedge \mathbf{x}^n.$$

而 \mathbf{n} 也和切平面垂直，并且如果 \mathbf{n} 和标架指示的定向相同，那么

$$\mathbf{n} \wedge \frac{\partial \Phi}{\partial y^1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial y^{n-1}} = c \cdot \mathbf{x}^1 \wedge \cdots \wedge \mathbf{x}^n$$

且 $c > 0$ 。这意味着 \mathbf{n} 是 \mathbf{t} 的正常数倍，又因为 \mathbf{n} 是单位向量，所以

$$\mathbf{n} = \frac{1}{\sqrt{a_1^2 + \cdots + a_n^2}} \sum_{i=1}^n (-1)^{i-1} a_i \mathbf{x}^i.$$

这直接得到

$$\mathbf{n}_i(\Phi(y)) \cdot \sqrt{a_1^2 + \cdots + a_n^2} = (-1)^{i-1} a_i.$$

从而两个积分的确相等, (1) 成立. 进一步地, 通过 Stokes 公式可知, 如果上面的 M 是某个 ∂D , 则

$$\int_{\partial D} \langle (f_1, \dots, f_n), \mathbf{n} \rangle d\sigma = \int_D \sum_{i=1}^n \frac{\partial f_i}{\partial x^i} dx^1 \wedge \dots \wedge dx^n.$$

通过选取典范的 D 的定向 $[\epsilon]$, \mathbf{n} 自动成为外指法向量. 不得不说这个形式是有用的: 因为 LHS 的表达形式有个更简洁的表达方式, 即在每个点 $p \in D \subset \mathbb{R}^n$ 处把 (f_1, \dots, f_n) 看作一个切空间中的向量, 那么就得到了一个 D 上的光滑向量场, 把它记为 X . X 在 ∂D 上未必是切向量场 (否则上面积分就是 0 了), LHS 就成为了 $\int_{\partial D} \langle X, \mathbf{n} \rangle d\sigma$, 类似于某种 “流量”. 这时 RHS 也可以将其赋予意义, 即所谓向量场的散度. 在直角坐标系下, 它具有直接的表达形式

$$\operatorname{div} X = \sum_{i=1}^n \frac{\partial}{\partial x^i} X_i.$$

一般地, 要从一个 vector field 得到一个 0-form, 可以把它先对偶到 $n-1$ -form, 作外微分运算得到 n -form, 最后再对偶回 0-form. 通过这样的想法, 可以证明 $\operatorname{div} X$ 是一个几何量, 在标准坐标系下具有这样简洁的表达形式而在其它坐标系下不一定永远形如此 (从量纲上也能看出来).

5.3 第二型曲线和曲面积分

定义向量场 X , 在一组局部坐标系下它可以写成

$$X = \sum_{i=1}^n X^i \cdot \frac{\partial}{\partial x_i}.$$

其中 X^i 可以视作 $\Omega \rightarrow \mathbb{R}$ 的函数. 从而, X 可以和 $\Omega \cong U \rightarrow \mathbb{R}^n$ 的函数 $f = (X^1, \dots, X^n)$ 进行对应.

Definition 5.3.1 (散度). 向量场 X 的散度记为

$$\operatorname{div}(X) = \sum_{i=1}^n \frac{\partial}{\partial x_i} X^i.$$

散度在某种意义上是 df 的 trace, 所以它在坐标变换下不变, 是个几何量.

Definition 5.3.2 (Laplace 算子). 二阶微分算子 (Laplace 算子) 为 $\Delta = \operatorname{div} \circ \nabla$. 它作用在函数 f 上得到 $\operatorname{div}(\nabla f)$.

Laplace 算子也是几何量, 因为 f 对应的梯度向量场是固定的. (但是梯度向量在流形上不是几何量?)

Theorem 5.3.3 (散度定理). 设 $\Omega \subset \mathbb{R}^n$ 为有界带边区域, ν 为 $\partial\Omega$ 的单位外法向量, 取 $d\sigma$ 为 $\partial\Omega$ 上的测度, 则对 \mathbb{R}^n 中任意的向量场 X , 均有

$$\int_{\Omega} \operatorname{div}(X) dx = \int_{\partial\Omega} \langle X, \nu \rangle d\sigma.$$

证明即是每个分量上的 Stokes 公式相加, 或者说考虑

$$\omega = dx^1 \wedge \cdots \wedge dx^n(X, \cdot, \dots, \cdot) = \sum_{i=1}^n (-1)^{i-1} \frac{\partial X}{\partial x^i} dx^1 \wedge \cdots \widehat{dx^i} \cdots \wedge dx^n.$$

代入 Real Stokes Formula.

Example 5.3.4 (求体积). 求体积可以转化为某种对边界的求积分. 当 $n = 2$ 时, 边界 $\partial\Omega$ 是曲线 (分段 C^1), 将曲线局部参数化为 $\gamma'(t) = (x'(t), y'(t))$. 求法向方向:

$$\nu = \frac{(y'(t), -x'(t))}{|\gamma'(t)|} \Rightarrow \nu \cdot d\sigma = \frac{(y'(t), -x'(t))}{|\gamma'(t)|} \cdot |\gamma'(t)| dt = (dy, -dx).$$

从而

$$\int_{\partial\Omega} \langle (Q, -P), \nu \rangle d\sigma = \int_{\partial\Omega} Q dy + P dx.$$

Definition 5.3.5 (第二型曲线积分). 设有曲线 $\gamma(t)$ 与向量场 X , 曲线积分

$$\int_{\gamma} \langle X, \nu \rangle d\sigma = \int_{\gamma} \sum_{i=1}^n X_i dx_i.$$

Definition 5.3.6 (微分形式). 设 $\Omega \subset \mathbb{R}^n$ 为开区域, $x = (x_1, \dots, x_n)$. 每个点 x 处的切空间 $T_x\Omega \cong \mathbb{R}^n$, 设它的一组 (全局的) 基为 $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$. 那么向量场可写成

$$X = \sum_{i=1}^n X^i(x) \frac{\partial}{\partial x_i}$$

的形式, 其中 $X^i(x)$ 为 Ω 上的函数. 设 $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$ 的一组对偶基为 $\{dx_1, \dots, dx_n\}$. 它们张成对偶空间 $T_x\Omega^*$. 余切丛的一个截面就是 Ω 上的 1 阶微分形式, 它可以写为

$$\omega = \sum_{i=1}^n \omega^i(x) dx_i, \quad \omega^i(x) \in C(\Omega, \mathbb{R}).$$

Definition 5.3.7 (高阶微分形式). k 阶微分形式在每点 x 处是一个 n 阶交错线性函数 $\text{Alt}^k(T_x\Omega, \mathbb{R})$, 从而是 $\bigwedge^k T_x^*\Omega$ 中的元素, 它的一组基是

$$\{dx_{i_1} \wedge \cdots \wedge dx_{i_k}\}, \quad i_1 < i_2 < \cdots < i_k.$$

记 k 阶微分形式构成全空间为 $\bigwedge^k(\Omega)$.

Example 5.3.8 (链复形). 因为 $d^2=0$, 所以外微分运算能够形成一个链复形.

$$\bigwedge^0(\Omega) \xrightarrow{d} \bigwedge^1(\Omega) \xrightarrow{d} \cdots \xrightarrow{d} \bigwedge^n(\Omega) \xrightarrow{d} 0.$$

在平坦空间中, 上同调均为 0, 或者说闭的微分形式 ($\ker d$) 都是恰当的 ($\text{im } d$), 这是 Poincare 的定理.

Proof. 先看简单的情况, 对 n 阶微分形式 $\omega = f dx_1 \wedge \cdots \wedge dx_n$, 找

$$\omega_1 = \sum_{i=1}^n g_i dx_1 \wedge \cdots \widehat{dx_i} \cdots \wedge dx_n$$

使得 $\sum (-1)^{i-1} \frac{\partial g_i}{\partial x_i} = f$. 这样我们只需要取后面的 $g_i (i \geq 2)$ 均为零, 选择 g_1 使得 $\frac{\partial g}{\partial x_1} = f$. 这只需要取积分

$$g(x_1, \dots, x_n) := \int_0^{x_1} f(t, x_2, \dots, x_n) dt + f(0, x_2, \dots, x_n).$$

现在从上往下归纳: 对 $\omega \in \bigwedge^k(\mathbb{R}^n)$, 设

$$\omega = \omega_1 + dx_1 \wedge \omega_2, \quad \omega_2 \in \bigwedge^{k-1}(\mathbb{R}^n).$$



Proposition 5.3.9. 设有单连通区域 $\Omega \subset \mathbb{R}^n$ 及 1 阶微分形式

$$\omega = \sum_{i=1}^n \omega^i(x) dx_i.$$

那么 ω 有原函数 f 当且仅当 $d\omega = 0$. 但是如果在区域并不足够好的时候, 已知 $d\omega = 0$ 是否能找到 f 使得 $\omega = df$?

- f 存在 $\Leftrightarrow \omega$ 的积分与路径无关.

1 阶微分形式可沿曲线做积分

$$\int_{\gamma} \omega = \int_0^1 \sum_{i=1}^n \omega^i(\gamma(t)) x'_i(t) dt.$$

取 $\gamma(0) = P$, $\gamma(1) = Q$, γ 是连接 P, Q 的曲线. 如果 ω 的积分和路径无关, 那么可定义函数

$$f(x) = \int_{\gamma} \omega, \quad \gamma(0) = P, \gamma(1) = x.$$

那么 $df = \omega$.

Proposition 5.3.10. 已知 $d\omega = 0$, 下述命题等价:

- 存在 f 使得 $\omega = df$.
- ω 的积分仅依赖于端点的选取.
- ω 沿环路积分总是零.

Theorem 5.3.11 (Stokes 公式). 设 D 是 \mathbb{R}^n 中的 $k+1$ 维光滑带边可定向子流形, 则对任意 $\omega \in \bigwedge^k(\mathbb{R}^n)$ 有

$$\int_D d\omega = \int_{\partial D} \omega.$$

这里的曲面不一定是余 1 维的, 所以需要其它的定义定向的方式. 大概的想法是: 在任意点处选定某种“定向”, 使得它的一个邻域内定向是一致的. 如果这样的定向可以全局定义, 那么就称曲面是可定向的. 比如在 v6 中, 我们希望任意两个坐

标卡的定向在交集处是相容的, 相容性来源于交集处有两种参数选择, 从而存在一个 $\mathbb{R}^k \rightarrow \mathbb{R}^k$ 的参数变换, 我们希望其行列式为正. 曲面可定向当且仅当存在处处非零的体积形式. 比如在代数拓扑中一般以 $H_k(M, M-x) \cong \mathbb{Z}$ 的某个生成元定向, 曲面可定向当且仅当

考虑 2-形式

$$\omega = P dx \wedge dy + Q dy \wedge dz + R dz \wedge dx.$$

代入 Stokes 公式:

$$\int_D d\omega = \int_{\partial D} \omega = \int_D \left(\frac{\partial P}{\partial z} + \frac{\partial Q}{\partial x} + \frac{\partial R}{\partial y} \right) dx \wedge dy \wedge dz.$$

D 上的微分形式 $dx \wedge dy \wedge dz$ 实际上已经选定好了 D 的定向, 并由此可诱导出 ∂D 上的定向, 具体如下: 如果在 $x \in \partial D$ 附近的坐标卡 U_α 内 M 由 $x_k \geq 0$ 给出, 并且

5.4 Brouwer 不动点定理

考虑 $D^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n, x_1^2 + \dots + x_n^2 \leq 1\}$, 其边界 $\partial D^n = S^{n-1}$ 为 $n-1$ 维单位球面.

Theorem 5.4.1. 设 $f: D^n \rightarrow D^n$ 为连续映射, 则存在 $x \in D^n$ 使得 $f(x) = x$.

Proof. 第一步是把问题转化到光滑的情形, 即只需证明光滑映射 $g: D^n \rightarrow D^n$ 存在不动点. 这是因为假设存在连续的 f 使得 f 无不动点, 根据 D^n 紧可知 $\|f(x) - x\|$ 作为关于 x 的函数存在非零的最小值, 记为 C . 这样就可以取一个光滑函数 g 来逼近 f . 其中一个方法是利用光滑化核 ρ_ε , 并且考虑 $f * \rho_\varepsilon$, 待定 δ 并选取 ε 使得 $\|f - f * \rho_\varepsilon\| < \delta$. 由于 $f * \rho_\varepsilon$ 有可能映出界外但半径不会超过 $1 + \delta$, 我们考虑 $f_\delta := (f * \rho_\varepsilon)/(1 + \delta)$. 现在 $\|f_\delta - f\| \leq 2\delta$, 那么选取 δ 使得 $\delta < C/2$ 即可知 f_δ 不存在不动点. 从而如果对光滑的 g 命题成立那么对连续的 f 命题也成立.

接下来是金典的映到边界的操作. 如果光滑的 f 无不动点, 那么考虑射线 $\overrightarrow{f(x)}$ 与边界 ∂D^n 的交, 记为 $g(x)$. 这样 $g(x)$ 就成为 $D^n \rightarrow \partial D^n$ 的光滑映射, 并且满足 $f|_{\partial D^n} = \text{id}$.

先来看 $n = 2$ 的情形. 将 F 表示为

$$F(x, y) = (u(x, y), v(x, y)) \in \mathbb{R}^2.$$

考虑积分

$$I = \int_{\partial D^2} u dx - v du.$$

由于在边界上 $F = \text{id}$, 所以上面的积分可以直接被解出来:

$$I = \int_0^{2\pi} \cos \theta (d \sin \theta) - \sin \theta (d \cos \theta) = 2\pi.$$

另一方面, 可以利用 Stokes 公式把 I 转化为内部的积分: 我们用 f_x 来表示 f 在 x 上的偏导数.

$$\begin{aligned} I &= \int_{\partial D^2} u(v_x dy + v_y dx) - v(u_x dx - u_y dy) \\ &= \int_{\partial D^2} (uv_x - vu_x) dx + (uv_y - vu_y) dy \\ (\text{Stokes}) &= \int_{D^2} (uv_x - vu_x)_y dy \wedge dx + (uv_y - vu_y)_x dx \wedge dy \\ (\text{偏导数交换}) &= 2 \int_{D^2} (u_x v_y - v_x u_y) dx \wedge dy. \end{aligned}$$

然而, 对 $u^2 + v^2 = 1$ 分别对 x 和 y 方向作偏导可得 $uu_x + vv_x = 0$ 以及 $uu_y + vv_y = 0$. 因此对



Theorem 5.4.2 (毛球定理). 在 S^{2k} ($k \geq 1$) 上不存处处非零的光滑 (连续) 向量场.

Proof. 反证存在 $X(x)$ 处处非零, 不妨设 $|X(x)| = 1$. 考虑

$$\Omega = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \frac{1}{2} < x_1^2 + \dots + x_n^2 < 2\}, \quad n = 2k + 1.$$

定义 Ω 上的光滑向量场

$$v(x) = rX(r^{-1}x), \quad r = |x|.$$

即延拓到 S^{2k} 的一个邻域. 考虑对这个向量场进行微扰. 对 $t \in \mathbb{R}$, 考虑映射 $\Phi_t(x) = x + tv(x)$, $x \in \Omega$. 根据 $x \perp v(x)$ 可知 $|\Phi_t(x)| = \sqrt{1+t^2}|x|$, 当 t 足够小时可以使

$$|\Phi_t(x) - \Phi_t(y)| = |x - y + t(v(x) - v(y))| \geq |x - y| - t|v(x) - v(y)| > \frac{1}{2}|x - y|.$$

这是因为根据垂直关系, $|v(x) - v(y)|$ 的下界能被 $|x - y|$ 控制吗? 从而 Φ_t 是 $\Omega \rightarrow \Phi_t(\Omega)$ 的微分同胚 (习题 87). 我们证明

$$\Phi_t(\Omega) = \{(x_1, \dots, x_n) \mid \frac{1}{2}\sqrt{1+t^2} < x_1^2 + \dots + x_n^2 < 2\sqrt{1+t^2}\}.$$

只需证明 S^{n-1} 的像集为 $\sqrt{1+t^2}S^{n-1}$. 证明我认为是和上述习题中一样的. 一方面,

$$m(\Phi_t(\Omega)) = (1+t^2)^{\frac{2k+1}{2}} m(\Omega).$$

另一方面, 换元公式

$$m(\Phi_t(\Omega)) = \int_{\Omega} |\det(I + t dv)| dx$$

是关于 t 的 n 次多项式, 矛盾.



回到 brouwer 不动点定理的证明.

Proof. 证明在偶数维时 Brouwer 不动点定理成立. 反证 $f(x) \neq x$, 构造非零向量

$$w(x) = x - \frac{1 - |x|^2}{1 - \langle x, f(x) \rangle} f(x), \quad \forall x \in D^n.$$



下面考虑一般情形. 如果 $n - 1 = 2k - 1$ 是奇数, 在点 $(x_1, x_2, \dots, x_{2k-1}, x_{2k}) \in S^{n-1}$ 处指定切向量

$$v(x_1, \dots, x_{2k}) = (-x_2, x_1, -x_4, x_3, \dots, -x_{2k}, x_{2k-1}) \neq 0$$

构造

$$X(x, \lambda \sqrt{1 - |x|^2}) = (-\lambda \sqrt{1 - |x|^2} w(x), \frac{\langle w(x), x \rangle}{\sqrt{1 - |x|^2}}), \quad \lambda = \pm 1.$$

下面就是机械地证明 X 非零.

5.5 区域不变性定理

若 $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是连续单射, 则 f 是开映射.

Proof. 设 B 为包含原点的闭球, $f: B \rightarrow \mathbb{R}^n$ 是连续单射, 我们要证明 $f(0)$ 落在 $f(B)$ 的内部. 记集合 $f(B) = Y$ 是闭集, 由于 f 是 $B \rightarrow Y$ 的双射, 因此可定义逆映射 $f^{-1}: Y \rightarrow B$. 延拓 f^{-1} 到整个 \mathbb{R}^n (Urysohn 引理) 得到连续函数 $G: \mathbb{R}^n \rightarrow B$, 那么条件就翻译为 G 在 Y 上有唯一的零点 $f(0)$.

Lemma 5.5.1. G 的定义同上. 如果存在 $\hat{G}: Y \rightarrow \mathbb{R}^n$ 满足 $|\hat{G}(y) - G(y)| \leq 1, \forall y \in Y$. 那么 \hat{G} 在 Y 上存在零点.

Proof. 构造 $F: B \rightarrow \mathbb{R}^n, F(x) = x - \hat{G}(f(x))$. 则 $|F(x)| \leq |G(f(x)) - \hat{G}(f(x))| \leq 1$. 从而 F 把闭圆盘映到闭圆盘内, 所以由 Brouwer 不动点定理可知存在 x 使得 $F(x) = x$, 即 \hat{G} 有零点.

回到原题, 如果 $f(0)$ 不是 Y 的内点, 那么存在 $c \notin Y$ 和 $f(0)$ 距离足够近. 取 r 使得 $|c - f(0)| < 2r \Rightarrow |G(c) - 0| < \frac{1}{42}$. 不妨设 $c = 0$, 即 0 不在 Y 内但距离 $f(0)$ 很近.

考虑紧集 $K_0 = \{y: |y| = r\}$ 以及 $K_1 = \{y \in Y: |y| \geq r\}$. 因此可在 $K_0 \cup K_1$ 上利用 Weierstrass 定理, 取多项式 P 使得

$$|P(y) - G(y)| < \delta, \quad \forall y \in K_1 \cup K_0.$$

由于 G 在 K_1 上无零点, 因此存在非零下界, 可以取 P 在 K_1 上无零点, 然后根据 $P(K_0)$ 零测知再平移一个小的 $c_0 \in \mathbb{R}^n$ 就可以使得 $c_0 + P$ 在 $K_0 \cup K_1$ 上无零点.

定义 $\hat{G}(y) = P(\Phi(y)), \forall y \in Y$, 其中 $\Phi(y) = \max\{\frac{r}{|y|}, 1\}y$.

- 当 $|y| \geq r$ 时 $\Phi(y) = y \Rightarrow |\hat{G}(y) - G(y)| = |P(y) - G(y)| < \delta$.
- 当 $|y| < r$ 时 $|G(y)| < \frac{1}{5}$, 故 $|y - f(0)| \leq 2r$, 因此 $|\hat{G}(y) - G(y)| \leq \delta + \frac{1}{42} + \frac{1}{42} < 1$.

这与之前的引理矛盾.



6 第一次作业

Problem 1

对非负函数 f , 如果 $\int f = 0$, 那么必定有 $f = 0$, a.e.

Proof. 记

$$E_n = \left\{ x : f(x) \geq \frac{1}{n} \right\}.$$

则我们称 $m(E_n) = 0$, 这是因为 $0 \leq 1/n \cdot \chi_{E_n} \leq f$, 故

$$0 = \int f \geq \int \frac{1}{n} \cdot \chi_{E_n} = \frac{m(E_n)}{n} \Rightarrow m(E_n) = 0.$$

因此,

$$m\{x : f(x) \neq 0\} = m\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \rightarrow \infty} m(E_n) = 0.$$

这推出 $f = 0$, a.e.



Problem 2

对于可积函数 f , 如果对于任何开集 O 均有 $\int_O f = 0$, 那么 $f = 0$, a.e.

Proof. 仍然记

$$E_n = \left\{ x : f(x) \geq \frac{1}{n} \right\}.$$

取 E_n 的一个等测包 G_n , 它是一个 G_δ 集, 故可设

$$G_n = \bigcap_{i=1}^{+\infty} O_{n,i}, \quad O_{n,i} \text{ 是下降开集列.}$$

由于

$$\int_O f = \int f \cdot \chi_O;$$

且所有 $f \cdot \chi_{O_{n,i}}$ 均可被可积函数 $|f|$ 控制, 因此由控制收敛定理可知

$$\int_{G_n} f = \int f \cdot \chi_{G_n} = \int f \cdot \lim_{i \rightarrow \infty} \chi_{O_{n,i}} = \lim_{i \rightarrow \infty} \int_{O_{n,i}} f = 0.$$

又因为 $m(G_n \setminus E_n) = 0$, 所以 $\int_{G_n \setminus E_n} f = 0$, 故 $\int_{E_n} f = 0$. 同上题可知 $m(E_n) = 0$, 所以 $m\{x : f(x) > 0\} = 0$. 同理可知 $m\{x : f(x) < 0\} = 0$, 所以 $f = 0$, a.e.



Problem 3

(Chebyshev 不等式) 如果 f 可积, 那么对任意 $\alpha > 0$ 均有


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

Proof. 记

$$E_\alpha = \{x : |f(x)| > \alpha\}.$$

则 $0 \leq \alpha \cdot \chi_{E_\alpha} \leq |f|$. 由单调性可知

$$\alpha \cdot m(E_\alpha) = \int \alpha \cdot \chi_{E_\alpha} \leq \int |f|.$$

这与需证式等价. 

Problem 4

对于非负可测函数 f , 定义

$$E_k = \{x : f(x) > 2^k\}, \quad F_k = \{x : 2^k < f(x) \leq 2^{k+1}\}.$$

证明 f 可积等价于 $\sum_{k=-\infty}^{+\infty} 2^k m(E_k)$ 收敛, 也等价于 $\sum_{k=-\infty}^{+\infty} 2^k m(F_k)$ 收敛. 这里划分值域也可用其它趋于无穷的序列, 比如 $\{n\}$.

本题证明后一个结论需要 f 几乎处处有限.

Proof. 先证明 f 可积与 $\sum_{k=-\infty}^{+\infty} 2^k m(F_k)$ 收敛等价. 由 F_k 的定义可知

$$\sum_{k=-\infty}^{+\infty} 2^k \cdot \chi_{F_k} \leq f \leq \sum_{k=-\infty}^{+\infty} 2^{k+1} \cdot \chi_{F_k}.$$

它们都是非负可测函数. 所以取积分可得

$$\sum_{k=-\infty}^{+\infty} 2^k \cdot m(F_k) \leq \int f \leq \sum_{k=-\infty}^{+\infty} 2^{k+1} \cdot m(F_k) = 2 \cdot \sum_{k=-\infty}^{+\infty} 2^k \cdot m(F_k).$$

因此 $\sum_{k=-\infty}^{+\infty} 2^k m(F_k)$ 收敛可推出 f 可积; f 可积可推出 $\sum_{k=-\infty}^{+\infty} 2^k m(F_k) < +\infty$, 即级数收敛.

最后证明 $\sum_{k=-\infty}^{+\infty} 2^k m(F_k)$ 收敛可推出 $\sum_{k=-\infty}^{+\infty} 2^k m(E_k)$ 收敛 (另一侧显然). 注意到 $E_k = \bigcup_{l=k}^{\infty} F_l$, 且 F_l 两两不交. 由于所有 E_k, F_l 均是可测集, 所以

$$m(E_k) = \sum_{l=k}^{\infty} m(F_l) \leq 2^{-k} \cdot \sum_{l=-\infty}^{+\infty} 2^l m(F_l) < +\infty.$$

由于 E_k 测度有限, 所以可作级数相减得 $m(E_k) - m(E_{k+1}) = m(F_k)$. 因此

$$\sum_{k=-\infty}^{+\infty} 2^k m(F_k) = \sum_{k=-\infty}^{+\infty} 2^k (m(E_k) - m(E_{k+1})) = \sum_{k=-\infty}^{+\infty} (2^k - 2^{k-1}) \cdot m(E_k) < +\infty.$$

因此 $\sum_{k=-\infty}^{+\infty} 2^k m(E_k)$ 收敛, 证毕.



Problem 5

考虑可测集 E 上的非负可测函数 f , 如果 $m(E) < \infty$, 那么 f 可积等价于

$$\sum_{n=0}^{\infty} m(\{x \in E : f(x) \geq n\}) < +\infty. \quad (1)$$

注意此题与上题的区别, 上题中不要求 E 为有限测度, 原因在于分割的足够细致 (靠近 0 的地方), 而此题的分割在 0 处跨度为 1, 因而要求全集 E 有限测度.

Proof. 一方面, 已知 (1) 成立, 下证 f 可积. 记

$$E_n = \{x \in E : f(x) \geq n\}, \quad F_n = \{x \in E : n \leq f(x) < n+1\}.$$

首先由级数收敛可知

$$m(\{x : f(x) = +\infty\}) = m\left(\lim_{n \rightarrow \infty} E_n\right) = \lim_{n \rightarrow \infty} m(E_n) = 0.$$

因此不妨设 f 取值有限. 其次, 由 E 有限测度可知 E_n, F_n 均有限测度, 所以

$$m(E_n) = \sum_{k=n}^{\infty} m(F_k).$$

代入得

$$\sum_{n=0}^{\infty} m(E_n) = \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} m(F_k) = \sum_{k=0}^{\infty} (k+1) m(F_k) < +\infty.$$

而 f 满足

$$f \leq \sum_{n=0}^{\infty} (n+1) \cdot \chi_{F_n}.$$

因此

$$\int f \leq \sum_{n=0}^{\infty} (n+1) \cdot m(F_n) = \sum_{n=0}^{\infty} m(E_n) < +\infty.$$

所以 f 可积. 另一方面, 假设 f 可积, 则也有 f 几乎处处有限, 因此不妨设 f 取值有限. 其次, 有

$$\sum_{n=0}^{\infty} n \cdot \chi_{F_n} \leq f \Rightarrow \sum_{n=0}^{\infty} n \cdot m(F_n) \leq \int f < +\infty.$$

所以

$$\sum_{n=0}^{\infty} (n+1) \cdot m(F_n) \leq \sum_{n=0}^{\infty} 2n \cdot m(F_n) + m(F_0) < +\infty.$$

因此 $\sum_{n=0}^{\infty} m(E_n) < +\infty$, 证毕.



Problem 6

此题是关于 Lebesgue 积分与广义积分的关系. 对于高维的定义在无穷测度集 E 上的广义无穷积分, 如果对任意可测连通有界闭区域 $F \subset E$, f 在 F 上的积分都能定义; 且对任意 $\varepsilon > 0$, 均能找到 N 使得对任意可测连通有界闭区域 F 满足 $B_N(0) \cap E \subset F$, 均有

$$\left| \int_E f - I \right| < \varepsilon,$$

则称广义积分 $\int_E f$ 收敛, 其值为 I . 对于一般的函数, 一维与高维有本质上的区别. 证明如下结论:

- (a) 在 \mathbb{R}^d , $d \geq 2$ 上, 如果 f 是广义可积的 (可只证明无穷积分或者奇异积分), 那么 f 也必定是 Lebesgue 可积的, 并且积分相同.
- (b) 在 \mathbb{R} 上, 如果 f 的广义积分存在并且是 Lebesgue 可积的, 那么 Lebesgue 积分等于广义积分.
- (c) 在 $(0, 1)$ 上, 证明函数 $f(x) = \frac{1}{x} \sin(\frac{1}{x^3})$ 广义积分存在, 但不是 Lebesgue 可积的.

Proof. (a) 首先证明在有界闭区域上, 若高维 Riemann 积分存在, 那么 Lebesgue 积分也存在且两个值相等. 设这个区域为 M , 暂分别用 $\int^{\mathcal{R}}$ 和 $\int^{\mathcal{L}}$ 来表示 Riemann 积分和 Lebesgue 积分. 高维 Riemann 积分把区域 M 划分为一些小区域 M_i , 并要求当所有小区域 M_i 的直径都趋于 0 时, 其 Dardoux 上下和收敛到一个值 I . 因此函数 f Riemann 可积当且仅当存在两列阶梯函数

$$\underline{f}_k = \sum_{i=1}^n m_i \cdot \chi_{D_{k,i}}, \quad \overline{f}_k = \sum_{i=1}^n M_i \cdot \chi_{D_{k,i}}$$

满足 $\underline{f}_k \leq f, \overline{f}_k \geq f$, 且它们的积分的极限相等. 由于只需要分划的足够细, 因此可不妨取 \underline{f}_k 和 \overline{f}_k 都是单调列, 这样可以设它们分别逐点收敛到 \underline{f} 和 \overline{f} . 由于 $\underline{f}_k, \overline{f}_k$ 也是 Lebesgue 意义下的简单函数, 并且它们在 Riemann 积分和 Lebesgue 积分下的取值相同, 因此

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_D^{\mathcal{R}} \underline{f}_k &= \lim_{k \rightarrow \infty} \int_D^{\mathcal{L}} \underline{f}_k = \int_D^{\mathcal{L}} \underline{f}; \\ \lim_{k \rightarrow \infty} \int_D^{\mathcal{R}} \overline{f}_k &= \lim_{k \rightarrow \infty} \int_D^{\mathcal{L}} \overline{f}_k = \int_D^{\mathcal{L}} \overline{f}. \end{aligned}$$

而 Riemann 可积意味着

$$\lim_{k \rightarrow \infty} \int_D^{\mathcal{R}} \underline{f}_k = \lim_{k \rightarrow \infty} \int_D^{\mathcal{R}} \overline{f}_k;$$

所以 $\underline{f} \leq f \leq \overline{f}$ 且前后两者的 Lebesgue 积分相等. 这推出 f 可积, 且积分就和上面那些值均相等. 故

$$\int_D^{\mathcal{R}} f = \int_D^{\mathcal{L}} f.$$

下面考虑一般的定义在无界闭区域 D 上的函数 f . 通过把无定义的点都取值为 0 可以把 f 延拓到 \mathbb{R}^d 上, 下在 \mathbb{R}^d 上证明结论. 首先证明当 $d \geq 2$ 时, 任意闭区域内积分存在一致上界. 这是因为由广义积分收敛的定义, 存在 I 与正整数 N 使得

$$B_N(0) \subset E \Rightarrow \left| \int_E^{\mathcal{L}} f - I \right| = \left| \int_E^{\mathcal{R}} f - I \right| \leq 1, \quad E \text{ 为有界闭区域}.$$

所以包含 $B_N(0)$ 的有界闭区域的积分存在一致上界. 对一般的区域 E , 存在一条道路连接 E 和 $B_N(0)$ 并且当 $d \geq 2$ 时可保证其测度为 0. 考虑 $E, \overline{B_N(0)}$ 还有这条道路的并集记为 E^* , 这是一个新的有界闭区域, 因此

$$\left| \int_{E^*}^{\mathcal{L}} f \right| \leq I + 1.$$

再把 E^* 拆开, 得到

$$\left| \int_E^{\mathcal{L}} f + \int_{B_N(0)}^{\mathcal{L}} f \right| \leq I + 1.$$

由于 $\int_{B_N(0)}^{\mathcal{L}}$ 也为固定的有限值, 因此对任意 E , $|\int_E^{\mathcal{L}} f|$ 存在一致上界.

再证明对有界的可测集 E , 其 Lebesgue 积分总存在且能被一致上界控制. 这是因为有限测度集可以被一系列不交的方体任意逼近, 这列方体之间可以使用零测的道路连接, 然后再通过与上面同样的方法, 把它们与 $B_N(0)$ 连接起来组成一个大的闭区域, 再把多余的道路和 $B_N(0)$ 扔掉即可. 整个区域 E 的积分是用上述方法构造出的逐渐逼近 E 的区域的积分的极限, 那么这个极限值存在并且能被一致上界控制.

最后考虑 $P^+ = \{x : f(x) \geq 0\}$, 则

$$\lim_{n \rightarrow \infty} \int_{P^+ \cap B_n(0)}^{\mathcal{L}} f = \int_{P^+}^{\mathcal{L}} f = \int f^+.$$

由于左边为单调列且存在一致上界, 所以 $\int^{\mathcal{L}} f^+$ 有限. 同理可知 $\int^{\mathcal{L}} f^-$ 有限, 这就得到了 f 为 Lebesgue 可积函数!

(b) 我们对任何的 d 一同证明 (b) 的结论. 由于 f 为 Lebesgue 可积的, 因此利用控制收敛定理可知

$$\lim_{n \rightarrow \infty} \int_{B_n(0)}^{\mathcal{L}} f = \int^{\mathcal{L}} f.$$

但左侧所有 Lebesgue 积分和 Riemann 积分相同, 并且替换为 Riemann 积分后构成了一个穷尽闭区域列, 它的极限是 $\int^{\mathcal{R}} f$. 因此

$$\int^{\mathcal{R}} f = \int^{\mathcal{L}} f.$$

(c)

$$\int_0^1 \frac{1}{x} \sin\left(\frac{1}{x^3}\right) dx = \int_{\infty}^1 t^{1/3} \cdot \sin(t) dt^{-1/3} = \frac{1}{3} \int_1^{\infty} t^{-1} \cdot \sin(t) dt.$$

课上证明过后者是收敛的, 所以 Riemann 积分存在. 但是

$$\int_1^{\infty} t^{-1} \cdot \sin(t) dt$$

不是绝对收敛的, 所以 $\int^{\mathcal{L}} |f| = +\infty$, 所以不是 Lebesgue 可积的.



Problem 7

对于一般的非负函数 f (不一定可测), 定义 f 的 Lebesgue 上、下积分

$$\begin{aligned} \overline{\int} f &= \inf \left\{ \int g : f \leq g, g \text{ 是简单函数} \right\}, \\ \underline{\int} f &= \sup \left\{ \int g : 0 \leq g \leq f, g \text{ 是简单函数} \right\}. \end{aligned}$$

证明如果 f 可测, 那么 f 的 Lebesgue 积分 (课上定义) 等于 $\underline{\int} f$. 因而可用下积分来定义 Lebesgue 积分.

Proof. 首先需要对 $f \in \mathcal{L}_0$ 证明两种定义相同. 即如果有一列简单函数满足 $f_n \rightarrow f$ (我们已经证明了不同的收敛列极限相同), 则

$$\sup \left\{ \int g : 0 \leq g \leq f, g \text{ 是简单函数} \right\} = \lim_{n \rightarrow \infty} \int f_n.$$

首先存在一列非负简单函数 f_n 从下逼近 f , 因此 $\text{LHS} \geq \text{RHS}$. 另一方面, 对任意 $0 \leq g \leq f$, 取函数列 $g_n = \max\{g, f_n\}$, 那么 g_n 也是简单函数, 并且有 $\int g_n \geq \int g$, 所以 $\text{RHS} \geq \int g$. 由 g 的任意性可知 $\text{RHS} \geq \text{LHS}$.

再证明对非负可测函数两种定义相同. 只需注意到

$$\begin{aligned} & \sup \left\{ \int g : 0 \leq g \leq f, g \in \mathcal{L}_0 \right\} \\ &= \sup \left\{ \sup \left\{ \int h : 0 \leq h \leq g, h \text{ 是简单函数} \right\} : 0 \leq g \leq f, g \in \mathcal{L}_0 \right\} \\ &= \sup \left\{ \int h : 0 \leq h \leq f, h \text{ 是简单函数} \right\}. \end{aligned}$$

所以两种定义相同.



Problem 8

如果 f 非负有界且有紧支集, 证明 f 可测等价于其上下积分相等.

Proof. 设 f 的支集落在有限测度集 E 上, f 的一个上界为 M . 先证明如果 f 可测, 那么其上下积分相等. 由于 $0 \leq f \leq M$, 因此两个确界的值均有限. 由于总有

$$\int f \leq \overline{\int} f,$$

所以只需证明存在两简单函数列 $\underline{f}_k \leq f$ 与 $\overline{f}_k \geq f$ 使得 $\int(\overline{f}_k - \underline{f}_k) \rightarrow 0$. 对每个 k , 设

$$E_{k,i} = \{x : \frac{i}{k} \leq f(x) < \frac{i+1}{k}\}, \quad 0 \leq i \leq [kM].$$

那么所有 $E_{k,i}$ 均是可测集, 因此可定义简单函数

$$\underline{f}_k = \sum_i \frac{i}{k} \cdot \chi_{E_{k,i}} \leq f; \quad \overline{f}_k = \sum_k \frac{i+1}{k} \cdot \chi_{E_{k,i}} \geq f. \quad (1)$$

那么就有

$$\lim_{k \rightarrow \infty} \left(\int g_k - \int f_k \right) \leq \lim_{k \rightarrow \infty} \int \left(\frac{1}{k} \cdot m(E) \right) = 0.$$

这导致下积分大于等于上积分, 所以上下积分相等.

另一方面, 证明如果上下积分相等, 下证可以找到一列简单函数 a.e. 收敛到 f , 这足以证明 f 可测. 首先取一列 $f_k \leq f$ 使得

$$\lim_{k \rightarrow \infty} \int f_k = \int f.$$

通过取 $f'_k = \max\{f_1, \dots, f_k\}$, 可以不妨设 f_k 上升. 由于 f_k 存在上界, 所以它逐点收敛到一个函数 g . 记 $F_n = \{x : f(x) - g(x) \geq 1/n\}$, 则对任意简单函数 $h \geq f$, 均有

$$\int h - \int f_k \geq \int \frac{1}{n} \cdot \chi_{F_n} = \frac{m(F_n)}{n}.$$

两边取极限可知 $\overline{\int} f - \underline{\int} f \geq \frac{m(F_n)}{n}$, 由于左边为零, 因此 $m(F_n) = 0$. 那么

$$m(\{x : f(x) \neq g(x)\}) = m\left(\bigcup_{n=1}^{\infty} F_n\right) = 0.$$

所以 $f_k \rightarrow f$, a.e. 因此 f 可测.



Remark. 后半段证明可以分别取两列 f_k, g_k 从下和从上逼近 f . 它们的极限有大小关系而且积分相同所以只差一个零测集, 并且和 f 也只差一个零测集, 因此 f 是一列可测函数的极限.

Problem 9

考虑 \mathbb{R} 上的有界函数 f , 其不连续点集记为 D . 如果 D 只有可数个极限点, 证明 f 在区间 $[a, b]$ 上是 Riemann 可积的.

Proof. 由于 f 有界, 因此只需证明 f 在 $[a, b]$ 上不连续点零测. 事实上, 可以证明 D 零测.

考虑 $E = \{x \in D : x \text{ 不是极限点}\}$, 则根据极限点的定义存在 x 为中心的小开球 $B_{\delta_x}(x)$ 使得 $B_{\delta_x}(x) \cap D = \{x\}$. 我们证明

$$B_{\delta_x/2}(x) \cap B_{\delta_y/2}(y) = \emptyset, \quad x, y \in E.$$

这是因为 $y \notin B_{\delta_x}(x), x \notin B_{\delta_y}(y)$, 所以 $\delta_x, \delta_y < |y - x|$, 它们取半后的和也小于 $|y - x|$.

那么此时可以从每个 $B_{\delta_x/2}(x)$ 中选取一个有理数 $q(x)$, 且 $x \mapsto q(x)$ 是单射, 所以 E 为可数集. 而又有 $D \setminus E$ 为 D 中的极限点集可数, 所以 D 可数, 因此 D 零测, 证毕. \square

Problem 10

如果可测函数列 f_k 依测度收敛到 f , 并且有可积函数 F 使得

$$|f_k(x)| \leq F(x).$$

那么有 f 可积, 并且

$$\lim_{k \rightarrow \infty} \int f_k = \int f.$$

Proof. 对于 $f_k \xrightarrow{m} f$, 存在其子列 f_{k_i} 满足 $f_{k_i} \rightarrow f$, a.e. 因此可对 f_{k_i} 使用控制收敛定理得到 $\lim_{k \rightarrow \infty} \int f_k = \int f$. \square

Problem 11

此题为 Fatou 引理更一般的形式, 称为 Fatou-Lebesgue 定理. 给定可积函数列 f_n , 如果存在可积函数 $g(x)$ 使得 $|f_n(x)| \leq g(x)$, 那么 $\liminf f_n$ 与 $\limsup f_n$ 均可积, 并且

$$\int \liminf_{n \rightarrow \infty} f_n \leq \liminf_{n \rightarrow \infty} \int f_n \leq \limsup_{n \rightarrow \infty} \int f_n \leq \int \limsup_{n \rightarrow \infty} f_n.$$

Proof. 由条件知 $g + f_n, g - f_n$ 均是非负可积函数, 因此可以利用 Fatou 引理得到

$$\int \liminf_{n \rightarrow \infty} (g \pm f_n) \leq \liminf_{n \rightarrow \infty} \int (g \pm f_n).$$

函数列的下极限是逐点定义的, 此时每个 $g(x)$ 可看作一个常数, 可以从下极限中拆出来, 再结合积分的线性可知

$$\int g + \int \liminf_{n \rightarrow \infty} (\pm f_n) \leq \int g + \liminf_{n \rightarrow \infty} \int (\pm f_n).$$

由 g 可积, 取正号时即得第一个不等号, 取负号时由于

$$\int \liminf_{n \rightarrow \infty} (-f_n) = - \int \limsup_{n \rightarrow \infty} f_n; \quad \liminf_{n \rightarrow \infty} \int (-f_n) = - \limsup_{n \rightarrow \infty} \int f_n,$$

所以第三个不等号成立, 中间的不等号由定义就可以推出.



Problem 12

记所有可测函数等价类 (在几乎处处相等的意义下) 组成的空间为 $L^0(E)$. 如果 $0 < m(E) < \infty$, 定义 $L^0(E)$ 上的距离函数

$$d(f, g) = \int_E \frac{|f - g|}{1 + |f - g|}.$$

证明在此距离下收敛等价于依测度收敛. 本题疑似需要 f, g 几乎处处有限.

Proof. 几乎处处有限函数列 f_k 依测度收敛到 f 当且仅当对任意 $\varepsilon > 0$, 均有

$$\lim_{k \rightarrow \infty} m(\{x : |f_k(x) - f(x)| > \varepsilon\}) = 0.$$

因此记 $E_{k,\varepsilon} = \{x : |f_k(x) - f(x)| > \varepsilon\}$, 则

$$\begin{aligned} \lim_{k \rightarrow \infty} d(f_k, f) &= \lim_{k \rightarrow \infty} \int_E \frac{|f_k - f|}{1 + |f_k - f|} \\ &= \lim_{k \rightarrow \infty} \left(\int_{E_{k,\varepsilon}} \frac{|f_k - f|}{1 + |f_k - f|} + \int_{E \setminus E_{k,\varepsilon}} \frac{|f_k - f|}{1 + |f_k - f|} \right) \\ &\leq \lim_{k \rightarrow \infty} (m(E_{k,\varepsilon}) \cdot 1 + m(E) \cdot \varepsilon) \\ &= m(E) \cdot \varepsilon. \end{aligned}$$

由 ε 的任意性可知 $\lim_{k \rightarrow \infty} d(f_k, f) = 0$. 另一方面, 又有

$$\begin{aligned} \lim_{k \rightarrow \infty} d(f_k, f) &\geq \lim_{k \rightarrow \infty} \int_{E_{k,\varepsilon}} \frac{|f_k - f|}{1 + |f_k - f|} \\ &\geq \lim_{k \rightarrow \infty} m(E_{k,\varepsilon}) \cdot \frac{\varepsilon}{1 + \varepsilon}. \end{aligned}$$

所以 $\lim_{k \rightarrow \infty} d(f_k, f) = 0$ 可推出 $\lim_{k \rightarrow \infty} m(E_{k,\varepsilon}) = 0$. 因此两种收敛条件等价.



Problem 13

如果 f 在实轴 \mathbb{R} 上可积, 定义

$$F(x) = \int_{-\infty}^x f(t) dt.$$

证明 $F(x)$ 是 \mathbb{R} 上的连续函数.

Proof. 只需证明对任意 x_0 ,

$$\lim_{x \rightarrow x_0} \int_{x_0}^x f(t) dt = 0.$$

首先证明 f 有界时命题成立, 设 $|f| \leq M$, 则

$$\lim_{x \rightarrow x_0} \int_{x_0}^x |f(t)| dt \leq \lim_{x \rightarrow x_0} |x - x_0| \cdot M = 0.$$

对一般的 f , 考虑 $f_k = f \cdot \chi_{\{x: |f(x)| \leq k\}}$, 由 f 可积, 利用控制收敛定理可知

$$\lim_{k \rightarrow \infty} \int f_k = \int f.$$

所以对任意 ε , 取 k 使得 $|\int f_k - \int f| < \varepsilon$, 那么就有

$$\varepsilon > \lim_{x \rightarrow x_0} \left| \int_{x_0}^x f_k(t) - f(t) dt \right| = \lim_{x \rightarrow x_0} \int_{x_0}^x |f(t)| dt.$$

令 $\varepsilon \rightarrow 0$ 即可得 $\text{RHS} = 0$. 因此 $F(x)$ 在 \mathbb{R} 上连续.



Problem 14

如果 f 在 \mathbb{R}^n 上可积, E 是 \mathbb{R}^n 上的紧集, 证明

$$\lim_{|y| \rightarrow \infty} \int_{E+\{y\}} f = 0.$$

Proof. 首先不妨设 f 非负, 否则用 $|f|$ 代替 f , 则条件与需证式都不发生变化.

考虑函数列 $f_N = f \cdot \chi_{B_N(0)}$, 则 $f_N \rightarrow f$. 由控制收敛定理,

$$\lim_{N \rightarrow \infty} \int f \cdot \chi_{B_N(0)} = \int f.$$

又因为 E 为紧集, 故 E 为有界可测集, 故存在 M 使得 $E \subset B_M(0)$. 因此对足够大的 N , 当 $|y| > N$, 时, 有

$$\int_{E+\{y\}} f = \int f \cdot \chi_{E+\{y\}} \leq \int f \cdot \chi_{B_M(0)+\{y\}} \leq \int f - \int f \cdot \chi_{B_{N-M}(0)}.$$

两边取极限可知

$$\lim_{|y| \rightarrow \infty} \int_{E+\{y\}} f = 0.$$



Problem 15

证明如果 f_k 在 L^1 中收敛到 f , 那么 f_k 依测度收敛到 f . 即是

$$f_k \xrightarrow{L^1} f \Rightarrow f_k \xrightarrow{m} f.$$

Proof. 只需证明对任意 $\varepsilon > 0$, 均有

$$\lim_{k \rightarrow \infty} m(\{x : |f_k(x) - f(x)| > \varepsilon\}) = 0.$$

记 $E_k = \{x : |f_k(x) - f(x)| > \varepsilon\}$, 则

$$0 = \lim_{k \rightarrow \infty} \int |f_k - f| \geq \lim_{k \rightarrow \infty} \int \varepsilon \cdot \chi_{E_k} = \varepsilon \cdot \lim_{k \rightarrow \infty} m(E_k).$$

这足以证明 $\lim_{k \rightarrow \infty} m(E_k) = 0$.



Problem 16

如果 f_k 几乎处处或者依测度收敛到 f , 并且 $\|f_k\|_{L_1} \rightarrow \|f\|_{L_1}$, 那么有 $f_k \xrightarrow{L_1} f$.

Proof. 固定 $\varepsilon > 0$, 我们考虑 $f_k \xrightarrow{m} f$ 的情形. 首先根据积分的绝对连续性, 取出 $\delta > 0$ 使得对任意 E 满足 $m(E) \leq \delta$, 均有

$$\int_E |f| \leq \varepsilon. \quad (1)$$

再取有限测度集 B 使得

$$\int_{B^c} |f| \leq \varepsilon. \quad (2)$$

考虑 f_k 和 f 在 B 上的限制. 由依测度收敛的性质, 可找到 N 使得对任意 $k > N$, 有

$$m(\{x \in B : |f_k(x) - f(x)| > \frac{\varepsilon}{m(B)}\}) < \delta \quad (3)$$

$$|\|f\| - \|f_k\|| < \varepsilon. \quad (4)$$

记 $B' = \{x \in B : |f_k(x) - f(x)| \leq \frac{\varepsilon}{m(B)}\}$, 那么就有

$$\left| \int_{B'} |f_k| - |f| \right| \leq \int_{B'} |f_k - f| \leq \frac{\varepsilon}{m(B)} \cdot m(B) = \varepsilon.$$

此外, 还有

$$\int_{B'^c} |f| = \int_{B \setminus B'} |f| + \int_{B^c} |f| \leq \varepsilon + \varepsilon = 2\varepsilon.$$

其中前一项是因为条件 (3) 告诉我们 $m(B \setminus B') < \delta$, 再用 (1) 得到上界, 第二项是因为 (2). 此外, 由 (4) 还可知

$$\left| \int_{B'^c} |f_k| - |f| + \int_{B'} |f_k| - |f| \right| < \varepsilon.$$

而第二项根据前面放缩可知已经可被 ε 控制, 所以

$$\left| \int_{B'^c} |f_k| - |f| \right| < 2\varepsilon.$$

结合上面所有命题,

$$\begin{aligned} \int |f_k - f| &\leq \int_{B'} |f_k - f| + \int_{B'^c} |f_k - f| \\ &\leq \int_{B'} |f_k - f| + \int_{B'^c} 2|f| + \int_{B'^c} |f_k| - |f| \\ &\leq \varepsilon + 4\varepsilon + 2\varepsilon = 6\varepsilon. \end{aligned}$$

由 ε 的任意性,

$$\lim_{k \rightarrow \infty} \int |f_k - f| = 0.$$

这和结论是等价的.



Remark. Fatou 估计:

$$\int \liminf_{k \rightarrow \infty} (2|f| - (|f| + |f_k|) - |f_k - f|) \leq \liminf_{k \rightarrow \infty} \int (2|f| - (|f| + |f_k|) - |f_k - f|)$$

展开即得结论.

Problem 17

如果 $f_k \xrightarrow{L^1} f$, 并且 $\sum_{k=1}^{\infty} \|f_k - f\| < \infty$, 证明 f_k 近一致收敛到 f .

Proof. f_k 近一致收敛到 f 等价于对任意 $\varepsilon > 0$, 均有

$$\lim_{n \rightarrow \infty} m \left(\bigcup_{k=n}^{\infty} \{x : |f_k(x) - f(x)| \geq \varepsilon\} \right) = 0.$$

设 $E_k = \{x : |f_k(x) - f(x)| \geq \varepsilon\}$, 则

$$\infty > \sum_{k=1}^{\infty} \|f_k - f\| = \sum_{k=1}^{\infty} \int |f_k - f| \geq \sum_{k=1}^{\infty} \int \varepsilon \cdot \chi_{E_k} = \varepsilon \sum_{k=1}^{\infty} m(E_k).$$

因此 $\sum_{k=1}^{\infty} m(E_k) = 0$, 所以由 Borel-Cantelli 定理,

$$m \left(\overline{\lim}_{k \rightarrow \infty} E_k \right) = 0.$$

这和近一致收敛的命题相同.



Problem 18

证明伸缩变换是连续的. 即

$$\lim_{\delta \rightarrow 1} \|D_{\delta} f - f\| = 0.$$

Proof. 固定任意 $\varepsilon > 0$. 由于阶梯函数在 L^1 空间中是稠密的, 所以存在阶梯函数 g 使得

$$\|f - g\| \leq \varepsilon, \quad \|D_{\delta} f - D_{\delta} g\| \leq \delta^{-1} \varepsilon.$$

设 $g = \sum_{k=1}^N a_k \cdot \chi_{I_k}$. 则对一个矩体 I , 有 $D_{\delta} \chi_I = \chi_{\delta^{-1}I}$. 所以

$$\begin{aligned} \|D_{\delta} g - g\| &= \left\| \sum_{k=1}^N a_k \cdot \chi_{\delta^{-1}I_k} - \sum_{k=1}^N a_k \cdot \chi_{I_k} \right\| \\ &\leq \sum_{k=1}^N |a_k| \cdot \|\chi_{\delta^{-1}I_k} - \chi_{I_k}\|. \end{aligned}$$

由于 $m((\delta^{-1}I_k) \triangle I_k)$ 可被关于 δ 的多项式控制, 且 $\delta = 1$ 时取值为零, 因此存在多项式 f_k 使得

$$\|D_{\delta} g - g\| \leq \sum_{k=1}^N |a_k| \cdot |f_k(\delta)|.$$

且 $f_k(1) = 0$. 因此

$$\lim_{\delta \rightarrow 1} \|D_\delta g - g\| = 0.$$

证毕.



Problem 19

证明 $L^1(\mathbb{R}^d)$ 是可分的距离空间, 即是证明存在可数稠密子集.

Proof. 我们知道全体阶梯函数构成 $L^1(\mathbb{R}^d)$ 的稠密子集, 那么可以证明

$$X = \{f : f = \sum_{k=1}^N a_k \cdot \chi_{I_k}, a_k \text{ 为有理数}, I_k \text{ 为边长有理的矩体}\}$$

也是稠密子集. 这是因为每个矩体都能写成一些下降有理边长矩体的极限, 每个 a_k 都能写成一列递减有理数的极限. 它们可组装成 X 中的单调下降函数列, 由控制收敛定理可知这列函数列可 L_1 收敛到任意阶梯函数. 那么只需证明 X 是可数集. 首先

$$a_k \cdot \chi_{I_k}, \quad a_k \in \mathbb{Q}, I_k \text{ 边长均} \in \mathbb{Q}.$$

共有可数种. 则

$$\text{card } X \leq \sum_{k=1}^{\infty} \aleph_0^k = \aleph_0.$$

所以 X 为可数集.



Problem 20

假设 $f \in L^1(\mathbb{R}), a > 0$. 证明级数

$$S(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{x}{a} + n\right)$$

在 \mathbb{R} 上几乎处处绝对收敛, 并且函数 $S(x)$ 是以 a 为周期的周期函数且 $S(x) \in L^1((0, a))$.

Proof. 由于 $f \in L^1(\mathbb{R})$, 故 f 几乎处处有限. 其次, 由于 $|f|$ 可积, 因此

$$\int |f| = \sum_{n=-\infty}^{\infty} \int_n^{n+1} |f(x)| = \sum_{n=-\infty}^{\infty} \int_0^1 |f(x+n)|.$$

由非负级数逐项积分定理可知

$$+\infty > \int |f| = \int_0^1 \sum_{n=-\infty}^{\infty} |f(x+n)|.$$

所以 $\sum_{n=-\infty}^{\infty} |f(x+n)|$ 几乎处处有限, 即 $S(x)$ 几乎处处收敛. 而对于绝对收敛的那些点, 可以讲级数重排, 即

$$\begin{aligned} \cdots + f\left(\frac{x}{a} - 1\right) + f\left(\frac{x}{a}\right) + f\left(\frac{x}{a} + 1\right) + \cdots &= \cdots + f\left(\frac{x}{a}\right) + f\left(\frac{x}{a} + 1\right) + f\left(\frac{x}{a} + 2\right) + \cdots \\ &= \cdots + f\left(\frac{x+a}{a} - 1\right) + f\left(\frac{x+a}{a}\right) + f\left(\frac{x+a}{a} + 1\right) + \cdots \end{aligned}$$

其中 LHS 为 $S(x)$, RHS 为 $S(x+a)$, 所以 $S(x+a) = S(x)$. 另一方面, 又有

$$\begin{aligned} \int_0^a |S(x)| &\leq \int_0^a \sum_{n=-\infty}^{\infty} \left| f\left(\frac{x}{a} + n\right) \right| \leq \sum_{n=-\infty}^{\infty} \int_0^a \left| f\left(\frac{x}{a} + n\right) \right| \\ &\leq \sum_{n=-\infty}^{\infty} \int_0^1 |f(x+n)| \leq \int |f| < +\infty. \end{aligned}$$

所以 $S(x)$ 为可积函数.



7 第二次作业

Problem 21

此题是 Riemann-Lebesgue 引理的推广. 假设 f 在 $[-\pi, \pi]$ 上 (Lebesgue?) 可积, 这样可定义 Fourier 级数

$$f \sim \sum_{n \in \mathbb{Z}} \hat{f}_n e^{inx}, \quad \hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

证明 \hat{f}_n 有界, 并且 $\lim_{n \rightarrow \infty} \hat{f}_n = 0$.

Proof. 由 f 可积可知

$$\int_{-\pi}^{\pi} |f| dx < +\infty.$$

而由于恒有 $|e^{-inx}| = 1$, 所以

$$\int_{-\pi}^{\pi} |f(x) e^{-inx}| dx \leq \int_{-\pi}^{\pi} |f| dx < +\infty.$$

且此上界是一致的, 因此 \hat{f}_n 有界. 若 g 是可导的且 $g' \in \mathcal{L}^1$, 则积分 Riemann 可积, 因此可使用 Riemann 分部积分的结论得到

$$\begin{aligned} |2\pi \cdot \hat{g}_n| &= \left| \int_{-\pi}^{\pi} g(x) \cdot \left(\frac{e^{-inx}}{-in} \right)' dx \right| \\ &= \left| g(x) \cdot \frac{e^{-inx}}{-in} \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} g'(x) \cdot \left(\frac{e^{-inx}}{-in} \right) dx \right| \\ &\leq \left| \frac{g(\pi) + g(-\pi)}{n} \right| + \frac{\|g'\|_{\mathcal{L}^1}}{n}. \end{aligned}$$

因此

$$\lim_{n \rightarrow \infty} \hat{g}_n = 0.$$

另一方面, 由 \mathcal{L}^1 空间的性质可知 光滑有紧支集函数在 \mathcal{L}^1 空间内稠密, 所以对任意 $\varepsilon > 0$, 可选择光滑有紧支集函数 g 使得

$$\|f - g\| \leq \varepsilon.$$

此时就有

$$\begin{aligned} |\hat{f}_n| &\leq \left| \int_{-\pi}^{\pi} (f(x) - g(x)) e^{-inx} dx \right| + \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-inx} dx \right| \\ &\leq \|f - g\| + |\hat{g}_n|. \end{aligned}$$

令 $n \rightarrow \infty$ 可知 $\lim_{n \rightarrow \infty} |\hat{f}_n| \leq \varepsilon$, 因此 $\lim_{n \rightarrow \infty} \hat{f}_n = 0$.



Remark. 更加有道理的是用阶梯函数逼近, 这是因为振荡函数的本质是对任何一段区间 $[a, b]$, 总有

$$\lim_{n \rightarrow \infty} \int_a^b F_n = 0.$$

Problem 22

上题对应的连续情形是如下结论: 假设 $f \in \mathcal{L}^1(\mathbb{R}^d)$, 定义 f 的 **Fourier 变换**

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, \xi \rangle} dx.$$

其中 $\langle \cdot, \cdot \rangle$ 代表内积. 证明 \hat{f} 一致连续, 有界, 且 $\lim_{|\xi| \rightarrow \infty} \hat{f}(\xi) = 0$.

Proof. 先证有界性:

$$|\hat{f}(\xi)| \leq \int_{\mathbb{R}^d} |f(x)| dx = \|f\| < +\infty.$$

一致连续性: 对任意 $\varepsilon > 0$, 存在 N 使得

$$\left| \int_{|x| \leq N} |f| dx - \int_{\mathbb{R}^d} |f| dx \right| < \varepsilon.$$

待定 δ 并取任意的 ξ, ζ 满足 $|\xi - \zeta| < \delta$, 则

$$\begin{aligned} |\hat{f}(\xi) - \hat{f}(\zeta)| &\leq \int_{\mathbb{R}^d} |f| \cdot |e^{-2\pi i \langle x, \xi \rangle} - e^{-2\pi i \langle x, \zeta \rangle}| dx \\ &\leq \int_{\mathbb{R}^d} |f| \cdot |e^{-2\pi i \langle x, \xi - \zeta \rangle} - 1| dx \\ &\leq \int_{|x| \leq N} |f| \cdot |e^{-2\pi i N \cdot \delta} - 1| dx + \int_{|x| > N} 2 \cdot |f| dx \\ &\leq |2 - 2 \cos(-2\pi N \cdot \delta)| \|f\| + 2\varepsilon. \end{aligned}$$

上面的最后几步均在 δ 较小时成立, 取 $\delta > 0$ 使得

$$|2 - 2 \cos(-2\pi N \cdot \delta)| \leq \frac{\varepsilon}{\|f\|},$$

则就有 $|\hat{f}(\xi) - \hat{f}(\zeta)| \leq 3\varepsilon$, 这就证明了一致收敛性.

最后来证明 $\lim_{|\xi| \rightarrow \infty} \hat{f}(\xi) = 0$. 设 ξ 方向的单位向量为 v , 那么可以把 \mathbb{R}^n 拆分为 v 所在一维子空间与和 v 垂直的超平面, 并把每个点表示为 (x, y) . 由 Fubini 定理,

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x, y) e^{-2\pi i \langle x, \xi \rangle} dx dy = \int_{\mathbb{R}} e^{-2\pi i x |\xi|} \left(\int_{\mathbb{R}^{d-1}} f(x, y) dy \right) dx.$$

再对 $f(x, y)$ 利用 Fubini 定理, 可知如果设

$$F(x) := \int_{\mathbb{R}^{d-1}} f(x, y) dy, \text{ 则 } \int_{\mathbb{R}} F(x) dx = \int_{\mathbb{R}^d} f.$$

因此 F 可积, 问题化归为一维情形. 由上题结论易知对任意闭区间 I , 均有

$$\lim_{|\xi| \rightarrow \infty} \int_I e^{-2\pi i \cdot x |\xi|} F(x) dx = 0.$$

对任意 ε , 取 I 使得

$$\left| \int_{\mathbb{R}} |F| dx - \int_I |F| dx \right| < \varepsilon.$$

则

$$\begin{aligned} \lim_{|\xi| \rightarrow \infty} \left| \int_{\mathbb{R}} e^{-2\pi i \cdot x |\xi|} F(x) dx \right| &\leq \int_{\mathbb{R} \setminus I} |F| dx + \lim_{|\xi| \rightarrow \infty} \left| \int_I e^{-2\pi i \cdot x |\xi|} F(x) dx \right| \\ &\leq \varepsilon + 0 = \varepsilon. \end{aligned}$$

由 ε 的任意性可知结论成立.



Problem 23

如果 $f \in \mathcal{L}^1(\mathbb{R})$, 对于正整数 n 与任意常数 $p > 0$, 证明

$$\lim_{n \rightarrow \infty} n^{-p} f(nx) = 0, \quad a.e. x \in \mathbb{R}.$$

Proof. 由于 $n^{-p} \leq 1$, 所以 $n^{-p} f(nx)$ 对所有 n 均是可积函数, 且

$$\int_{\mathbb{R}} |n^{-p} f(nx)| dx = n^{-p-1} \int_{\mathbb{R}} |f| dx \leq \|f\| \cdot n^{-p-1}.$$

因此

$$\sum_{n=1}^{\infty} \int_{\mathbb{R}} |n^{-p} f(nx)| dx \leq \|f\| \cdot \sum_{n=1}^{\infty} \frac{1}{n^{p+1}} < +\infty.$$

由正项级数收敛定理可知

$$\int_{\mathbb{R}} \sum_{n=1}^{\infty} |n^{-p} f(nx)| dx < +\infty.$$

因此

$$\sum_{n=1}^{\infty} |n^{-p} f(nx)| \text{ 可积 } \Rightarrow \sum_{n=1}^{\infty} |n^{-p} f(nx)| < +\infty, \quad a.e.$$

对那些级数收敛的 x ,

$$\lim_{n \rightarrow \infty} n^{-p} f(nx) = 0.$$

所以几乎处处的 x 满足要求.



Remark. 当时没做出来的一题, 证明进路是: 设 F_n 是非负函数, 要证

$$\lim_{n \rightarrow \infty} F_n(x) = 0, \text{ a.e.}$$

一个充分条件是

$$\sum_{n=1}^{\infty} F_n(x) < +\infty, \text{ a.e.}$$

为了消去 a.e., 这等价于

$$\int_{\mathbb{R}} \sum_{n=1}^{\infty} F_n(x) < +\infty.$$

如果 F_n 是非负级数, 那么就可以交换级数与积分, 从而实际上是从 $\int F_n$ 的性质推得了 F_n 的性质. 但是, 非负函数列未必有

$$\int_{\mathbb{R}} \lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} F_n(x)$$

所以不能通过直接求极限来解决问题.

Problem 24

假设 f 是 \mathbb{R}^d 上的可积函数. 对 $\alpha > 0$, 定义 $E_\alpha = \{x : |f(x)| > \alpha\}$. 证明:

$$\int_{\mathbb{R}^d} |f(x)| dx = \int_0^\infty m(E_\alpha) d\alpha.$$

Proof. 考虑 $\Phi : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$, 满足

$$\Phi(x, y) = \begin{cases} 1, & \text{如果 } 0 \leq y \leq |f(x)|, \\ 0, & \text{else.} \end{cases}$$

这是一个非负可测函数, 因此利用 Tonelli 定理可知

$$\begin{aligned} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \Phi(x, y) dy \right) dx &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}^d} \Phi(x, y) dx \right) dy. \\ \text{LHS} &= \int_{\mathbb{R}^d} |f(x)| dx, \quad \text{RHS} = \int_{\mathbb{R}} m(E_y) dy. \end{aligned}$$

所以两者相等.



Remark. 整个解释见题 29.

Problem 25

在 \mathbb{R}^2 中考虑集合 $A = \{0\} \times [0, 1]$, $B = \mathbb{N} \times \{0\}$, 这里 \mathbb{N} 是 $[0, 1]$ 上的不可测集. 证明 $A + B = \{x + y : x \in A, y \in B\}$ 是不可测的.

此题表明对于可测集 A, B , 和集 $A + B$ 与差集 $A - B$ 不一定是可测集.

Proof. 假设 $A + B$ 为可测集, 则对几乎所有的 y , $E_y = \{x : (x, y) \in A + B\}$ 为可测集. 但是 $E_y = \mathbb{N}$, $y \in [0, 1]$. 故对一个正测度的 y 有 E_y 不可测, 矛盾.



Problem 26

如果 $E_1 \subset \mathbb{R}^{d_1}, E_2 \subset \mathbb{R}^{d_2}$, 那么乘积集 $E_1 \times E_2$ 在 $\mathbb{R}^{d_1+d_2}$ 中的外测度满足关系

$$m^*(E) \leq m^*(E_1) \cdot m^*(E_2).$$

Proof. 考虑覆盖 E_1 的一列矩体 I_1, I_2, \dots 及覆盖 E_2 的一列矩体 J_1, J_2, \dots , 那么可知

$$E_1 \times E_2 \subset \bigcup_{k=1}^{\infty} I_k \times \bigcup_{l=1}^{\infty} J_l = \bigcup_{k,l} I_k \times J_l.$$

最右边仍然是可数个方体的并, 而对方体有 $m(I_k \times J_l) = m(I_k) \cdot m(J_l)$. 那么如果 E_1, E_2 中存在一个外测度为无穷, 那么 $\sum_{k,l} m(I_k)m(J_l) = \infty$, 否则这是一个绝对收敛列, 交换求和顺序即得 $\sum_{k,l} m(I_k)m(J_l) = (\sum_k m(I_k)) \cdot (\sum_l m(J_l))$. 而又有 $\sum_{k,l} m(I_k)m(J_l) \geq m^*(E_1 \times E_2)$, 因此取两者足够接近 $m(E_i)$ 可知 $m^*(E) \leq m^*(E_1) \cdot m^*(E_2)$. \square

Problem 27

定义 $\delta = (\delta_1, \dots, \delta_d)$, $\delta_j \neq 0, 1 \leq \delta_j \leq d$. 对于可测集 E , 定义

$$E^\delta = \{(\delta_1 x_1, \dots, \delta_d x_d) : (x_1, \dots, x_d) \in E\}.$$

证明 E^δ 可测, 并且 $m(E^\delta) = |\delta_1 \cdots \delta_d| m(E)$.

Proof. 对矩体 I , 显然有 $m(I^\delta) = |\delta_1 \cdots \delta_d| m(I)$. 那么对任意集合 X , 考虑覆盖它的一列矩体 I_1, I_2, \dots . 那么 $I_1^\delta, I_2^\delta, \dots$ 覆盖 X^δ , 所以

$$\sum_{n=1}^{\infty} m(I_n^\delta) \geq m^*(X^\delta) \Rightarrow |\delta_1 \cdots \delta_d| \cdot \sum_{n=1}^{\infty} m(I_n) \geq m^*(X^\delta).$$

则取 $\sum_{n=1}^{\infty} m(I_n)$ 足够接近 $m^*(X)$ 可得

$$|\delta_1 \cdots \delta_d| \cdot m^*(X) \geq m^*(X^\delta).$$

记 $\delta^{-1} = (\delta_1^{-1}, \dots, \delta_d^{-1})$, 则同理可知

$$|\delta_1^{-1} \cdots \delta_d^{-1}| \cdot m^*(X^\delta) \geq m^*(X).$$

因此 $|\delta_1 \cdots \delta_d| \cdot m^*(X) = m^*(X^\delta)$. 另一方面, 对开集 G 与闭集 F , 也有 G^δ 和 F^δ 是开集和闭集. 因此对任意 $\varepsilon > 0$, 取 $F \subset E \subset G$ 使得 $m(G \setminus F) < \varepsilon$, 所以

$$F^\delta \subset E^\delta \subset G^\delta, \quad m(G^\delta \setminus F^\delta) = |\delta_1 \cdots \delta_d| \cdot \varepsilon.$$

取 $\varepsilon \rightarrow 0$ 就可知 E^δ 可测, 并且

$$|\delta_1 \cdots \delta_d| \cdot (m(E) - \varepsilon) \leq m(E^\delta) \leq |\delta_1 \cdots \delta_d| \cdot (m(E) + \varepsilon).$$

所以 $m(E^\delta) = |\delta_1 \cdots \delta_d| \cdot m(E)$. \square

Problem 28

假设 f 是 \mathbb{R}^d 上的可测函数, 考虑图像 $\Gamma = \{(x, y) \in \mathbb{R}^d \times \mathbb{R}, y = f(x)\}$. 证明 Γ 是 \mathbb{R}^{d+1} 上的零测集.

之前的习题中表明如果 $f(x)$ 是连续的, 那么图像是零测集. 此题表明可以把连续函数推广到一般可测函数.

Proof. 首先证明 Γ 为可测集. 对任意 n ,

$$E_{k,n} = \{x : \frac{k}{n} \leq f(x) < \frac{k+1}{n}\}$$

是可测集, 并且


$$\Gamma \subset \bigcup_k E_{k,n} \times [\frac{k}{n}, \frac{k+1}{n}) =: F_n.$$

后者为可测集. 另一方面, 对任意 $(x, y) \notin \Gamma$, 存在 n 使得 $|y - f(x)| > \frac{1}{n}$, 因此 $y \notin F_n$. 所以 (x, y) 和 $(x, f(x))$ 在第 n 次分划中不会落在同一个 $E_{k,n}$ 中, 因此

$$\Gamma = \bigcap_n F_n \text{ 为可测集.}$$

现在设 χ 为 Γ 的示性函数, 它作用在 $\mathbb{R}^d \times \mathbb{R}$ 上. 则

$$m(\Gamma) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \chi(x, y) dy \right) dx.$$

固定 x , 则恰好有一个 y 使得 $\chi(x, y) = 1$, 它在 \mathbb{R} 上测度也是零, 因此积分后得到 0, 即 Γ 零测. 

Remark. 只是证明 Γ 零测的话, 可以直接由

$$\Gamma \subset \bigcup_k E_{k,n} \times [\frac{k}{n}, \frac{k+1}{n})$$

得到

$$m^*(\Gamma \cap [-N, N]) \leq \sum_k m(E_{k,n} \cap [-N, N]) \cdot \frac{1}{n} = 2N/n.$$

由 n 的任意性可知 $m^*(\Gamma \cap [-N, N]) = 0$, 从而它是零测集. 由于一系列零测集的极限也是零测集, 所以 Γ 是零测集.

Problem 29

给定 \mathbb{R}^d 上的非负实值函数 $f(x)$, 考虑 \mathbb{R}^{d+1} 中图像下面的区域

$$\Omega_f = \{(x, y) : 0 \leq y \leq f(x)\}.$$

证明 f 可测当且仅当 Ω_f 是 \mathbb{R}^{d+1} 中的可测集, 并且 $m(\Omega_f) = \int_{\mathbb{R}^d} f$.

Proof. 若已知 f 可测, 则可以用非负简单函数上升列 h_n 逼近 f . 一方面, 对简单函数

$$h = \sum_{i=1}^k a_i \cdot \chi_{E_i},$$

其中 a_i 两两不等, E_i 两两不交, 则有

$$\Omega_h = \bigsqcup_{i=1}^k E_i \times [0, a_i] \Rightarrow m(\Omega_h) = \sum_{i=1}^k a_i \cdot m(E_i) = \int_{\mathbb{R}^d} h.$$

又由 Beppo-Levi 定理, $\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} h_n = \int_{\mathbb{R}^d} f$.

另一方面, 考虑 $\lim_{n \rightarrow \infty} \Omega_{h_n}$. 对任意 $(x, y) \in \Omega_f$, $y < f(x)$, 由于 $h_n(x)$ 为上升列且收敛到 f , 故存在 n 使得 $0 \leq y \leq h_n(x)$, 此时有 $(x, y) \in \Omega_{h_n}$. 而对 $(x, y) \notin \Omega_f$, 如果 $y < 0$ 则显然 $(x, y) \notin \lim_{n \rightarrow \infty} \Omega_{h_n}$, 若 $y > f(x)$ 则 $y > f(x) \geq h_n(x)$, 同样有 $(x, y) \notin \lim_{n \rightarrow \infty} \Omega_{h_n}$. 因此我们得到了


$$\lim_{n \rightarrow \infty} \Omega_{h_n} \cup \Gamma = \Omega_f.$$

由上题结论可知 Γ 为零测集, 所以 Ω_f 为可测集. 并且由于 Ω_{h_n} 为上升列, 因此

$$m(\Omega_f) = \lim_{n \rightarrow \infty} m(\Omega_{h_n}) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} h_n = \int_{\mathbb{R}^d} f.$$

最后如果 Ω_f 为可测集, 则

$$\{y = a\} \cap \Omega_f = \{(x, a) : f(x) \geq a\},$$

所以 $\{x : f(x) \geq a\}$ 为可测集, 因此 f 可测, 这就回到上面的推理过程. 

Remark. 由于 Ω_f 为可测集, 考虑其特征函数 χ_{Ω_f} , 因此利用 Tonelli 定理可推出

$$\begin{aligned} m(\Omega_f) &= \int_{\mathbb{R}^n} \{y : 0 \leq y \leq f(x)\} = \int_{\mathbb{R}^n} f(x) \, dx \\ &= \int_0^\infty \{x : f(x) \geq y\}. \end{aligned}$$

后一行是第 24 题的结论, 那里的证明并不严谨.

Problem 30

考虑非退化线性变换 $L : \mathbb{R}^d \rightarrow \mathbb{R}^d$ 以及 \mathbb{R}^d 上的可积函数 f , 那么 $f \circ L$ 也是 \mathbb{R}^d 上的可积函数, 并且

$$\int f(L(x)) = \frac{1}{|\det L|} \int f(x).$$

Proof. 将 f 正负分离为 $f = f^+ - f^-$. 定义

$$E_\alpha = \begin{cases} \{x : f(x) > \alpha\}, & \alpha > 0, \\ \emptyset, & \alpha = 0, \\ \{x : f(x) < \alpha\}, & \alpha < 0. \end{cases}$$

则根据第 4 题的结论,

$$\int_{\mathbb{R}^d} f = \int_0^\infty m(E_\alpha) d\alpha - \int_0^\infty m(E_{-\alpha}) d\alpha. \quad (1)$$

另一方面, 把 $f \circ L$ 对应的集合记为 E'_α , 则对 $\alpha > 0$,

$$E'_\alpha = \{x : f(L(x)) > \alpha\} = \{x : L(x) \in E_\alpha\} = L^{-1}(E_\alpha).$$

同理可知小于零的情形, 因此 $E'_\alpha = L^{-1}(E_\alpha)$, 这推出

$$m(E'_\alpha) = \frac{1}{|\det L|} m(E_\alpha).$$

而根据 (1), 积分中的每项均被替换成原来的 $\frac{1}{|\det L|}$ 倍, 所以

$$\int f(L(x)) = \frac{1}{|\det L|} \int f(x).$$



Remark. 也可以用和上题差不多的证明方法, 先证明对简单函数都成立, 然后逼近 f .

Problem 31

如果线性变换 L 退化, f 是 \mathbb{R}^d 上的可测函数, 举例说明 $f(L(x))$ 不一定可测.

Proof. 取 \mathbb{R} 中的不可测集 \mathcal{N} , 并取 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 满足

$$f(x, y) = \begin{cases} 1, & x \in \mathcal{N}, y = 0; \\ 0, & \text{else.} \end{cases}$$

则 $\mathcal{N} \times \{0\}$ 为 \mathbb{R}^2 中的零测集, 故 f 为可测函数.

但取 $L(x, y) = (x, 0)$ 为退化线性映射, 则 $(f \circ L)^{-1}(1) = \mathcal{N} \times \mathbb{R}$. 而 $\mathcal{N} \times \mathbb{R}$ 为不可测集 (否则它在几乎所有方向上的切片均可测, 与 \mathcal{N} 的不可测性矛盾), 故 $f \circ L$ 不是可测函数.



Problem 32

如果 f 是 \mathbb{R}^d 上的可测函数, 证明 $g(x, y) = f(x - y)$ 作为 \mathbb{R}^{2d} 上的函数可测.

Proof. 首先证明 $h: \mathbb{R}^{2d} \rightarrow \mathbb{R}$, $h(x, y) = f(x)$ 可测, 这是因为

$$\{(x, y) : h(x, y) \geq a\} = \{x : f(x) \geq a\} \times \mathbb{R}^d.$$

后者为两个可测集的积, 因此也是可测集, 故 h 为可测函数. 另一方面, $L: (x, y) \mapsto (x - y, y)$ 是非退化线性映射, 因此也是连续映射, 所以 $h \circ L = g$ 是可测函数.




Problem 33

证明在 \mathcal{L}^1 中不存在一个函数 I 使得 $I * f = f, \forall f \in \mathcal{L}^1$.

Proof. 考虑单点集 $\{0\}$ 的特征函数 f , 则积分

$$1 = I * f(0) = \int I(-y)f(y) dy$$

几乎处处为零, 所以 RHS=0, 矛盾. 故这样的 I 不存在. 

Remark. 比较没那么搞笑的一个版本是证明不存在 I 使得 $I * f = f$, a.e. 但其实也没有很难, 只需利用积分的绝对连续性, 取足够小的某个 χ_A 即可.

Problem 34

如果 f 可积, 定义 f 的 Fourier 变换

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, \xi \rangle} dx.$$

证明: 如果 f, g 均可积, 则

$$\widehat{f * g}(\xi) = \hat{f}(\xi) \hat{g}(\xi).$$

Proof. 直接展开表达式:

$$\begin{aligned} \hat{f}(\xi) \hat{g}(\xi) &= \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, \xi \rangle} dx \cdot \int_{\mathbb{R}^d} g(y) e^{-2\pi i \langle y, \xi \rangle} dy \\ &= \int_{\mathbb{R}^{2d}} f(x) g(y) e^{-2\pi i \langle x+y, \xi \rangle} dx dy. \end{aligned}$$

进行变元替换 $(x, y) \mapsto (x+y, y)$, 定义 $t := x+y$, 则上式可化为

$$\begin{aligned} \int_{\mathbb{R}^{2d}} f(t-y) g(y) e^{-2\pi i \langle t, \xi \rangle} dt dy &= \int_{\mathbb{R}^d} (f(t-y) g(y) dy) e^{-2\pi i \langle t, \xi \rangle} dt \\ &= \int_{\mathbb{R}^d} (f * g)(t) e^{-2\pi i \langle t, \xi \rangle} dt = \widehat{f * g}(\xi). \end{aligned}$$

证毕. 

8 第三次作业

Problem 35

此题的目的是定义弱 L^1 空间 L^1_ω : 如果 f 可测, 定义

$$\lambda_f(t) = m(\{x : |f(x)| > t\}).$$

再定义

$$\|f\|_{L^1_\omega} = \sup_{t>0} (t\lambda_f(t)).$$

定义 L^1_ω 是所有满足 $\|f\|_{L^1_\omega} < \infty$ 的函数组成的空间. 证明 $L^1 \subset L^1_\omega$, 即是证明如果 f 可积, 那么 $f \in L^1_\omega$.

Proof. 只需证明 $\|f\|_{L^1} \geq \|f\|_{L^1_\omega}$. 这是因为对可积函数 f ,

$$\int |f| \geq \int |f| \cdot \chi_{\{x: |f(x)| > t\}} \geq \int t \cdot \chi_{\{x: |f(x)| > t\}} = t \cdot \lambda_f(t)$$

对任意 $t > 0$ 成立. 所以

$$\|f\|_{L^1} \geq \sup_{t>0} (t\lambda_f(t)) = \|f\|_{L^1_\omega}.$$

故 $L^1 \subset L^1_\omega$.



Problem 36

证明 L^1_ω 是线性空间, 但上题定义的 $\|\cdot\|_{L^1_\omega}$ 不是范数, 而是拟范数, 满足

$$\|f + g\|_{L^1_\omega} \leq 2(\|f\|_{L^1_\omega} + \|g\|_{L^1_\omega}), \quad \forall f, g \in L^1_\omega. \quad (1)$$

并且 2 是最优系数.

Proof. 先证明不等式 (1) 成立, 这是因为

$$\{x : |(f + g)(x)| > t\} \subset \{x : |f(x)| > \frac{t}{2}\} \cup \{x : |g(x)| > \frac{t}{2}\}.$$

所以

$$\begin{aligned} \lambda_{f+g}(t) &\leq \lambda_f\left(\frac{t}{2}\right) + \lambda_g\left(\frac{t}{2}\right), \\ t \cdot \lambda_{f+g}(t) &\leq 2 \cdot \left(\frac{t}{2} \lambda_f\left(\frac{t}{2}\right) + \frac{t}{2} \lambda_g\left(\frac{t}{2}\right)\right) \leq 2(\|f\|_{L^1_\omega} + \|g\|_{L^1_\omega}). \end{aligned}$$

由 t 的任意性可知

$$\|f + g\|_{L^1_\omega} \leq 2(\|f\|_{L^1_\omega} + \|g\|_{L^1_\omega}).$$

构造一组取等的例子：取

$$f = \begin{cases} x, & x \in [0, 2]; \\ 0, & \text{else.} \end{cases} \quad g = \begin{cases} 2 - x, & x \in [0, 2]; \\ 0, & \text{else.} \end{cases}$$


$$t \cdot \lambda_f(t) = t(2 - t) \leq 1, \quad t \cdot \lambda_g(t) \leq 1.$$

因此 $\|f\|_{L^1_\omega}, \|g\|_{L^1_\omega} \leq 1$. 然而在 $[0, 2]$ 上 $f + g = 2$, 故 $\|f + g\|_{L^1_\omega} \geq 4$. 此时

$$\|f + g\|_{L^1_\omega} \geq 2(\|f\|_{L^1_\omega} + \|g\|_{L^1_\omega}).$$

因此等号能被取到. 再证明 L^1_ω 是线性空间, 由 (1) 可得 $\|f + g\|_{L^1_\omega}$ 有限, 故 $f + g \in L^1_\omega$. 另一方面,

$$\{x : |(\mu f)(x)| > t\} = \{x : |f(x)| > t/|\mu|\} \Rightarrow \lambda_{\mu f}(t) = \lambda_f(t/|\mu|).$$

故 $\|\mu f\|_{L^1_\omega} = |\mu| \cdot \|f\|_{L^1_\omega} < +\infty$. 因此 $\mu f \in L^1_\omega$, 证毕. 

Problem 37

更进一步我们不能在 L^1_ω 中找到一个与 $\|\cdot\|_{L^1_\omega}$ 等价的范数 (即是满足三角不等式). 即是证明不存在 L^1_ω 上的范数 $\|\cdot\|$ 及常数 C 使得

$$C^{-1} \|f\| \leq \|f\|_{L^1_\omega} \leq C \|f\|, \quad \forall f \in L^1_\omega. \quad (1)$$

Proof. 首先证明一维情形: 考虑阶梯函数

$$f = \sum_{k=1}^n \frac{1}{k+1} \cdot \chi_{[k, k+1]}.$$

那么 $f \leq 1/x$, 又因为 $\|1/x\|_{L^1_\omega} = 1$, 因此 $\|f\|_{L^1_\omega} \leq 1$. 设置换 $\sigma = (1 \ 2 \ \dots \ n)$, 记

$$f_i = \sum_{k=1}^n \frac{1}{k+1} \cdot \chi_{[\sigma^i(k), \sigma^i(k)+1]}.$$

则有 $\lambda_f(t) = \lambda_{f_i}(t)$, 因此对每个 f_i 均有 $\|f_i\|_{L^1_\omega} \leq 1$. 现在考虑

$$\begin{aligned} \sum_{i=0}^{n-1} f_i &= \sum_{i=0}^{n-1} \sum_{k=1}^n \frac{1}{k+1} \cdot \chi_{[\sigma^i(k), \sigma^i(k)+1]} \\ &= \sum_{k=1}^n \chi_{[k, k+1]} \cdot \sum_{i=0}^{n-1} \frac{1}{\sigma^{-i}(k) + 1} \\ &= \sum_{k=1}^n \chi_{[k, k+1]} \cdot \sum_{t=2}^{n+1} \frac{1}{t} = \left(\sum_{t=2}^{n+1} \frac{1}{t} \right) \cdot \chi_{[1, n+1]}. \end{aligned}$$

因此

$$\left\| \sum_{i=0}^{n-1} f_i \right\|_{L^1_\omega} = n \cdot \left(\sum_{t=2}^{n+1} \frac{1}{t} \right).$$

假设条件 (1) 成立, 则

$$Cn \geq \sum_{i=0}^{n-1} C \|f_i\|_{L^1_\omega} \geq \sum_{i=0}^{n-1} \|f_i\| \geq \left\| \sum_{i=0}^{n-1} f_i \right\| \geq C^{-1} \cdot \left\| \sum_{i=0}^{n-1} f_i \right\|_{L^1_\omega} \geq C^{-1} n \cdot \left(\sum_{t=2}^{n+1} \frac{1}{t} \right).$$

这意味着

$$\sum_{t=2}^{n+1} \frac{1}{t} \leq C^2, \quad \forall n \in \mathbb{N}.$$

这和调和级数发散矛盾, 因此原命题成立. \square

Remark. 非常有意思的题, 关键在于发现如何把一些好的函数叠在一起使得它们的贡献被平均化.

Problem 38

如果 f 可积, 并且不是几乎处处为 0, 那么存在常数 $c > 0$ 使得

$$Mf(x) > c|x|^{-d}, \quad \forall |x| > 1.$$

这表明 Mf 不是可积的.

Proof. 由于 $\int |f| > 0$ 当且仅当 f 几乎处处不为零, 因此由题意 $\int |f| > 0$. 因此可找到 N 使得

$$C := \int_{B_N(0)} |f| > 0.$$

因此对任意 $|x| > 1$ 均有

$$\begin{aligned} C &\leq \int_{B_{|x|+N}(x)} |f| \leq m(B_{|x|+N}(x)) \cdot Mf(x) \leq C_d(|x|+N)^d \cdot Mf(x) \\ &\leq C_d(N+1)^d \cdot |x|^d \cdot Mf(x). \end{aligned}$$

最后一步用到了 x 的模长下界. 因此

$$Mf(x) \geq \frac{C}{C_d(N+1)^d} \cdot |x|^{-d}.$$

前者即可设为常数 c , 证毕. \square

Problem 39

上题表明极大函数不可积, 主要原因在于无穷远处衰减太慢. 此题表明存在局部不可积的极大函数:

$$f = x^{-1} |\log x|^{-2} \chi_{\{0 < x < \frac{1}{2}\}}.$$

证明 f 可积但存在常数 $c > 0$ 使得

$$Mf(x) > cx^{-1} |\log x|^{-1}, \quad 0 < x < \frac{1}{2}.$$

特别地, 对应的极大函数在 0 处不可积.

Proof. 注意到

$$((\log x)^{-1})' = -(\log x)^{-2} \cdot x^{-1}.$$

因此

$$\int^{\mathbb{R}} f = -(\log x)^{-1} \Big|_0^{\frac{1}{2}} = (\log 2)^{-1} < +\infty.$$

因此 f 也是 Lebesgue 可积的. 另一方面, 又有


$$\int_0^t f = -(\log x)^{-1} \Big|_0^t = |\log t|^{-1}.$$

故取 $(0, t + \varepsilon)$ 是包含 t 的开球, 得到

$$Mf(x) \geq (t + \varepsilon)^{-1} |\log(t + \varepsilon)|^{-1}.$$

令 $\varepsilon \rightarrow 0$ 即可知

$$Mf(x) > \frac{1}{2} \cdot x^{-1} |\log x|^{-1}, \quad 0 < x < \frac{1}{2}.$$

而 $(\log(-\log x))' = x^{-1}(-\log x)^{-1}$, 因此它在 0 处的反常积分不可积, 因此 0 附近的 Lebesgue 积分也不可积, 故 Mf 在 0 附近不可积. 

Problem 40

如果 E 是 $[0, 1]$ 上的可测集满足对任意 $[0, 1]$ 上的区间 I 以及某个 $\alpha > 0$ 均有 $m(E \cap I) \geq \alpha m(I)$, 证明 $m(E) = 1$.

Proof. 由于 E 为可测集, 考虑对 χ_E 利用 Lebesgue 微分定理. 那么对几乎处处的 x 均有


$$\mathbf{1}_{x \in E} = \lim_{x \in B, m(B) \rightarrow 0} \frac{m(E \cap B)}{m(B)}.$$

假设 $m(E) < 1$, 则 $m([0, 1] \setminus E)$ 不是零测集, 所以存在 $x \notin E$ 满足

$$0 = \lim_{x \in B, m(B) \rightarrow 0} \frac{m(E \cap B)}{m(B)}.$$

取区间列 $x \in I_n \subset [0, 1]$ 使得其长度趋近于 0, 则存在 n 使得

$$\frac{m(E \cap I_n)}{m(I_n)} < \alpha.$$

这与条件矛盾. 所以 $m(E) = 1$. 

Problem 41

假设 E 为 \mathbb{R}^d 上的可测集, 对任意 $x \in \mathbb{R}^d$, 回忆距离函数

$$d(x, E) = \inf_{y \in E} |x - y|.$$

证明对几乎所有的 $x \in E$ 均有

$$\lim_{y \rightarrow 0} \frac{d(x + y, E)}{|y|} = 0.$$

Proof. 同上题对 χ_E 使用 Lebesgue 微分定理可知对几乎所有的 $x \in E$, 均有

$$\lim_{x \in B, m(B) \rightarrow 0} \frac{m(E \cap B)}{m(B)} = 1.$$

则对任意 $1/2 > \varepsilon > 0$, 存在 δ_0 使得对任意 $\delta < \delta_0$, 均有

$$\frac{m(E \cap B_\delta(x))}{m(B_\delta(x))} > 1 - \varepsilon^d.$$

那么对任意 $|y| < \delta_0/2$, 取 $\delta = 2|y|$. 那么开球 $B_{\varepsilon\delta}(x+y) \subset B_\delta(x)$, 并且

$$\frac{m(B_{\varepsilon\delta}(x+y))}{m(B_\delta(x))} = \varepsilon^d.$$

因此 $B_{\varepsilon\delta}(x+y) \cap E \neq \emptyset$, 这就意味着 $d(x+y, E) < \varepsilon\delta$. 此时

$$\frac{d(x+y, E)}{|y|} < \frac{\varepsilon\delta}{\delta/2} = 2\varepsilon, \quad \forall |y| < \frac{\delta_0}{2}.$$

令 $\varepsilon \rightarrow 0$, 这就证明了

$$\lim_{y \rightarrow 0} \frac{d(x+y, E)}{|y|} = 0.$$



Remark. 这里的 $|y|$ 就可以看作 x, y 张成的区间, 这个区间里至少有长度为 $d(x+y, E)$ 的部分不在 E 中, 所以在区间长度 $\rightarrow 0$ 时直接得到结论.

2022 年期中第十题: 给定可测集 $E \subset \mathbb{R}$, 证明

$$\int_{|y| \leq 1} \frac{d(x+y, E)}{|y|^2} dy < +\infty, \quad \text{a.e. } x \in E.$$

Proof.

$$\int_{|y| \leq 1} \frac{d(x+y, E)}{|y|^2} dy = \int_{|z-x| \leq 1} \frac{d(z, E)}{|x-z|^2} dz.$$

这是一个关于 x 的函数, 只需证明这个函数关于 x 的积分有限, 那么这个函数就几乎处处有限. 当 E 为有界集时, 设 $E \subset [-M, M]$, 则

$$\begin{aligned} \int_{x \in E} \int_{|z-x| \leq 1} \frac{d(z, E)}{|x-z|^2} dz dx &= \int_{x \in E} \int_{z \notin E} \frac{d(z, E)}{|x-z|^2} \cdot \chi_{|x-z| \leq 1} dz dx \\ &= \int_{z \notin E} d(z, E) \int_{x \in E} \frac{1}{|x-z|^2} \cdot \chi_{|x-z| \leq 1} dx dz \\ &\leq \int_{z \notin E} d(z, E) \int_{d(z, E) \leq |x-z| \leq 1} \frac{1}{|x-z|^2} dx dz \\ &= \int_{z \notin E} d(z, E) \cdot 2(d(z, E)^{-1} - 1) \cdot \chi_{d(z, E) \leq 1} dz \\ &\leq 2(2M + 2). \end{aligned}$$

那么只需把 E 划分为一些有界区域 $E \cap [k, k+1]$ 之并, 在每个区域内这个函数都几乎处处有限, 那么可数个区域的并中仍然几乎处处有限.



Problem 42

我们称满足条件

$$\int K_\varepsilon = 1, \quad |K_\varepsilon| \leq A \min\{\varepsilon^{-d}, \varepsilon |x|^{-d-1}\}$$

的卷积核 $\{K_\varepsilon\}$ 为**恒同逼近**. 证明热核

$$H_t = (4\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t}}$$

则在取 $K_\varepsilon = H_{\varepsilon^2}$ 时, K_ε 是恒同逼近.

Proof. 熟知

$$\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}.$$

因此对 d 维的 x 满足

$$\int e^{-|x|^2} dx = \int e^{-(x_1^2 + \dots + x_d^2)} dx_1 \dots dx_d = (\sqrt{\pi})^d.$$

换元之后得到

$$\begin{aligned} \int H_t dx &= (4\pi t)^{-\frac{d}{2}} \cdot \int e^{-\frac{|x|^2}{4t}} dx \\ &= (4\pi t)^{-\frac{d}{2}} \cdot (2\sqrt{t})^d \cdot \int e^{-|y|^2} dy \\ &= \pi^{-\frac{d}{2}} \cdot (\sqrt{\pi})^d = 1. \end{aligned}$$

再来看上界: 一方面,

$$K_\varepsilon = (4\pi\varepsilon^2)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4\varepsilon^2}} \leq (4\pi)^{-\frac{d}{2}} \cdot \varepsilon^{-d}.$$

另一方面, 指数函数满足存在常数 $B > 0$, 使得

$$B \cdot e^x \geq x^{\frac{d+1}{2}}, \quad \forall x \geq 0.$$

所以对任意 $x \leq 0$, 均有

$$B \cdot |x|^{-\frac{d+1}{2}} \geq e^{-x}.$$

因此

$$K_\varepsilon \leq (4\pi\varepsilon^2)^{-\frac{d}{2}} \cdot B \cdot \left(\frac{|x|^2}{4\varepsilon^2}\right)^{-\frac{d+1}{2}} = 2\pi^{-\frac{d}{2}} B \cdot \varepsilon |x|^{-d-1}.$$

取两个常数的较大者

$$A = \max\{(4\pi)^{-\frac{d}{2}}, 2\pi^{-\frac{d}{2}} B\}$$

即可得 K_ε 是恒同逼近.



Problem 43

证明 Poisson 核

$$P_y(x) = \frac{c_d y}{(y^2 + |x|^2)^{\frac{d+1}{2}}}, \quad c_d = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}}$$

在取 $K_\varepsilon = P_\varepsilon$ 时为恒同逼近.

Proof. 先计算积分:

$$\begin{aligned} \int P_y dx &= c_d y \int (y^2 + |x|^2)^{-\frac{d+1}{2}} dx \\ (B_d \text{ 为单位球体积}) &= c_d y \cdot dB_d \cdot \int_0^{+\infty} r^{d-1} (y^2 + r^2)^{-\frac{d+1}{2}} dr \\ (r = yt) &= c_d y \cdot dB_d \cdot \int_0^{+\infty} y^{d-1} t^{d-1} y^{-d+1} (1+t^2)^{-\frac{d+1}{2}} d(yt) \\ &= c_d \cdot dB_d \cdot \int_0^{+\infty} t^{d-1} (1+t^2)^{-\frac{d+1}{2}} dt \\ (s = t^2) &= c_d \cdot dB_d \cdot \int_0^{+\infty} s^{\frac{d-1}{2}} (1+s)^{-\frac{d+1}{2}} d(\sqrt{s}) \\ &= \frac{1}{2} c_d \cdot dB_d \cdot \int_0^{+\infty} s^{\frac{d-2}{2}} (1+s)^{-\frac{d+1}{2}} ds \\ &= \frac{1}{2} c_d \cdot dB_d \cdot \int_0^{+\infty} \left(\frac{s}{1+s} \right)^{\frac{d-2}{2}} \left(\frac{1}{1+s} \right)^{\frac{3}{2}} ds \\ (g = \frac{s}{s+1}) &= \frac{1}{2} c_d \cdot dB_d \cdot \int_0^1 g^{\frac{d-2}{2}} (1-g)^{\frac{3}{2}} \cdot (1-g)^{-2} dg \\ &= \frac{1}{2} c_d \cdot dB_d \cdot \int_0^1 g^{\frac{d-2}{2}} (1-g)^{-\frac{1}{2}} dg \\ &= \frac{d}{2} \cdot \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \cdot \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d+2}{2})} \cdot B\left(\frac{d}{2}, \frac{1}{2}\right) \\ &= \frac{d}{2\sqrt{\pi}} \cdot \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d+2}{2})} \cdot \frac{\Gamma(\frac{d}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{d+1}{2})} \\ &= \frac{d}{2} \cdot \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d+2}{2})} = 1. \end{aligned}$$

验证上界是容易的, 这是因为

$$P_y(x) \leq \min\left\{\frac{c_d y}{y^{d+1}}, \frac{c_d y}{|x|^{d+1}}\right\} = c_d \min\{y^{-d}, y|x|^{-d-1}\}.$$

取 $A = c_d$ 就得到收敛性的结论.

Problem 44

如果 K_ε 为恒同逼近, 证明存在常数 c 使得对任何可积函数 f 均有

$$\sup_{\varepsilon > 0} |K_\varepsilon * f(x)| \leq c M f(x), \quad \forall x \in \mathbb{R}^d.$$

Proof.

$$\begin{aligned}
 |K_\varepsilon * f(x)| &\leq \int |K_\varepsilon(x-y) \cdot f(y)| dy \\
 &\leq \int_{|x-y| \leq \varepsilon} |K_\varepsilon(x-y)f(y)| dy + \sum_{k=0}^{+\infty} \int_{2^k \varepsilon < |x-y| \leq 2^{k+1} \varepsilon} |K_\varepsilon(x-y)f(y)| dy \\
 &\leq A\varepsilon^{-d} \int_{|x-y| \leq \varepsilon} |f(y)| dy + \sum_{k=0}^{+\infty} A\varepsilon(2^k \varepsilon)^{-d-1} \int_{|x-y| \leq 2^{k+1} \varepsilon} |f(y)| dy \\
 &\leq A\varepsilon^{-d} Mf(x) \cdot C_d \varepsilon^d + \sum_{k=0}^{\infty} A\varepsilon(2^k \varepsilon)^{-d-1} \cdot Mf(x) \cdot C_d(2^{k+1} \varepsilon)^d \\
 &\leq Mf(x) \cdot (AC_d + \sum_{k=0}^{\infty} 2^{d-k} AC_d) \\
 &= Mf(x) \cdot (AC_d + 2^{d+1} AC_d).
 \end{aligned}$$

记 $c = AC_d + 2^{d+1} AC_d$, 这是与 x, ε, f 均无关的常数, 所以可以得到

$$\sup_{\varepsilon > 0} |K_\varepsilon * f(x)| \leq cMf(x), \quad \forall f, \forall x \in \mathbb{R}^d.$$



Remark. 放缩是这样的, 这题所做的放缩的理解是必须和 ε 及 $\|f\|_{L^1}$ 均无关, 只能用极大函数的定义进行放缩, 而且为了保证 ε 不出现截断选择很唯一. 并且, 逐倍放缩还真的可以用 $(2^k)^{-d-1}$ 来控制住半径为 $(2^{k+1})^{-d}$ 的球, 因为多了一个 2^k 出来.

9 第四次作业

Problem 45

证明函数 $F(x) = \begin{cases} x^2 \sin(x^{-2}), & x \neq 0 \\ 0, & x = 0 \end{cases}$ 导数存在, 但 F' 在 $[0, 1]$ 上不可积.

Proof. 由于 $\mathbb{R} \setminus \{0\}$ 是开集, 所以当 $x \neq 0$ 时

$$F'(x) = 2x \sin(x^{-2}) - 2x^{-1} \cos(x^{-2}).$$

当 $x = 0$ 时,

$$\lim_{x \rightarrow 0} \left| \frac{F(x) - F(0)}{x} \right| = \lim_{x \rightarrow 0} |x \sin(x^{-2})| \leq \lim_{x \rightarrow 0} |x| = 0.$$

所以 $f'(0) = 0$ 存在. 然而对 $a_k = (2k\pi + \frac{\pi}{2})^{-1/2}$, $b_k = (2k\pi + \pi)^{-1/2}$, 有

$$\int_{b_k}^{a_k} F' = F(b_k) - F(a_k) = 0 - \frac{1}{2k\pi + \frac{\pi}{2}}.$$

所以

$$\int_0^1 |F'| \geq \sum_k \int_{a_k}^{b_k} |F'| = \sum_k \frac{1}{2k\pi + \frac{\pi}{2}} = +\infty.$$

所以 F' 不是 Lebesgue 可积的.



Problem 46

如果 F 在 $[a, b]$ 上连续, 且对任意 $a \leq x \leq b$ 均有 $D^+(F)(x) \geq 0$. 证明 F 是单调增的.

Proof. 假设存在 $m < n$ 使得 $F(m) > F(n)$, 考虑函数

$$G(x) = F(x) - \frac{F(n) - F(m)}{n - m}(x - m).$$

则 $G(m) = G(n) = F(m)$, 并且条件化为

$$\limsup_{h>0, h \rightarrow 0} \frac{G(x+h) - G(x)}{h} + \frac{F(n) - F(m)}{n - m} \geq 0. \quad (1)$$

记

$$\varepsilon = -\frac{F(n) - F(m)}{2(n - m)} > 0.$$

那么 (1) 可推出

$$\limsup_{h>0, h \rightarrow 0} \frac{G(x+h) - G(x)}{h} > \varepsilon. \quad (2)$$

记 $M = \{x \in [m, n] : G(x) - \varepsilon x \geq G(m) - \varepsilon m\}$, 则 $m \in M$. 考虑 $\sup M$:

- 一方面 $\sup M \in M$, 这是因为 G 具有连续性, 所以取一列 $x_n \rightarrow \sup M$ 就能得到

$$\lim_{n \rightarrow \infty} (G(x_n) - \varepsilon x_n) = G(\sup M) - \varepsilon \sup M \geq G(m) - \varepsilon m.$$

- 另一方面假设 $\sup M < n$, 则由 (2) 存在足够小的 h 使得 $\sup M + h < n$ 且

$$\frac{G(\sup M + h) - G(\sup M)}{h} > \varepsilon \Rightarrow \sup M + h \in M.$$

这与 $\sup M$ 的定义矛盾!

所以 $\sup M = n$, 这与 $G(m) = G(n)$ 矛盾. 所以假设错误, F 为单调增函数.



Remark. 更方便的另解是直接利用 Sunrise Lemma. 假设 $F(a) > F(b)$, 那么适当加上一个线性函数平移到 G , 使得

$$G(a) > G(b), \quad D^+(G)(x) > 0.$$

那么对任意 $x \in [a, b)$, 存在 $h > 0$ 使得 $G(x+h) > G(x)$. 并且, G 还是连续函数, 所以 G 满足 Sunrise Lemma 的利用条件, 直接得到 $G(a) \leq G(b)$, 矛盾.

Problem 47

如果 F 是定义在 $[a, b]$ 上的连续函数, 证明 Dini 导数 $D^+(F)(x)$ 可测.

Proof. 由于

$$\limsup_{h \rightarrow 0, h > 0} g(h) = \lim_{h \rightarrow 0, h > 0} \sup_{t \in (0, h]} g(t),$$

且 $\sup g(t)$ 关于 h 单调减, 所以有

$$\lim_{h \rightarrow 0, h > 0} \sup_{t \in (0, h]} g(t) = \lim_{N \rightarrow \infty} \sup_{t \in (0, 1/N]} g(t).$$

回到原题, 记

$$g(x, h) = \frac{f(x+h) - f(x)}{h}.$$

则固定 x , $g(x, h)$ 在 $(0, 1/N]$ 上是关于 h 的连续函数, 所以

$$\sup_{h \in (0, 1/N]} g(x, h) = \sup_{h \in (0, 1/N] \cap \mathbb{Q}} g(x, h).$$

现在固定 h , 把后者视为关于 x 的函数可知这是可数个连续函数的上确界, 所以是可测函数. 最后 $D^+(F)(x)$ 是一些可测函数的极限, 所以也是可测函数. \square

Problem 48

考虑跳跃函数

$$J(x) = \sum_{n=1}^{\infty} j_n(x).$$

证明

$$\limsup_{h \rightarrow 0} \frac{J(x+h) - J(x)}{h}$$

是可测的.

Proof. 我们对一般的可测函数 J 完成证明.

$$\limsup_{h \rightarrow 0} \frac{J(x+h) - J(x)}{h} = \lim_{N \rightarrow \infty} \sup_{h \in (0, 1/N)} \frac{J(x+h) - J(x)}{h}.$$

而

$$\begin{aligned} & \left\{ x : \sup_{h \in (0, 1/N)} \frac{J(x+h) - J(x)}{h} > \alpha \right\} \\ &= \{ x : \exists h \in (0, 1/N) \text{ s.t. } J(x+h) - \alpha(x+h) > J(x) - \alpha x \} \\ &= \left\{ x : \sup_{h \in (0, 1/N)} (J(x+h) - \alpha(x+h)) > J(x) - \alpha x \right\}. \end{aligned}$$

然而考虑函数

$$G(x) = \sup_{h \in (0, 1/N)} (J(x+h) - \alpha(x+h)),$$

可以证明 G 是可测函数: 若 $G(x) > \beta$, 则 $\exists y \in (x, x+1/N)$ s.t. $J(y) - \alpha y > \beta$, 所以 $\forall z \in (y-1/N, y)$ 均有 $G(z) > \beta$. 而 $(y-1/N, y)$ 是 x 的一个邻域, 这导致


$$\{x : G(x) > \beta\}$$

为开集, 所以 G 为可测函数. 而又因为 $J(x) - \alpha x$ 为可测函数, 因此

$$\{x : G(x) - J(x) + \alpha x > 0\}$$

为可测集, 这就证明了

$$\sup_{h \in (0, 1/N)} \frac{J(x+h) - J(x)}{h}$$

是可测函数. 而 $D^+(F)(x)$ 作为一系列可测函数的极限也是可测函数. 

Remark. 神证明, 值得的多次品味: 灵活利用了证明 f 可测的两大手段: 直接利用定义以及可测函数对各种运算的封闭性. 特别地,

$$G(x) = \sup_{t \in (x, x+\varepsilon)} g(t)$$

当 g 可测时, G 也可测.

Problem 49

如果 F 是 $[a, b]$ 上的单调增函数. J_F 是相对应的跳跃函数, 证明


$$\sum_{a < x_k < b} j_k(x) \leq F(b-0) - F(a+0), \quad \forall a \leq x \leq b.$$

Proof. 由于 LHS 为单调增函数, 因此只需证明 $x = b$ 的情形. 此时对任意 $x_k \in (a, b)$, 有 $j_k(x) = F(x_k+0) - F(x_k-0)$. 由于 F 为单调增函数, 故区间

$$(F(x_k-0), F(x_k+0))$$

两两不交, 但是又有

$$\bigcup_{a < x_k < b} (F(x_k-0), F(x_k+0)) \subset (F(a+0), F(b-0)).$$

计算两边测度即得结论. 

Problem 50

如果 $J(x) = \sum_{k=1}^{\infty} j_k(x)$ 为跳跃函数.

$$E_\varepsilon = \{x : \overline{D}(J)(x) > \varepsilon\}, \quad E_{\varepsilon, N} = \{x : \overline{D}(J_N)(x) > \varepsilon\}, \quad J_N(x) = \sum_{k \leq N} j_k(x).$$

证明

$$E_\varepsilon \subset E_{\varepsilon, N} \cup \{x_1, \dots, x_N\}.$$

Proof. 由于 J 作为级数和是一致收敛的, 所以

$$\begin{aligned} \overline{D}(J)(x) &= \limsup_{h \rightarrow 0} \frac{J(x+h) - J(x)}{h} \\ &= \limsup_{h \rightarrow 0} \sum_{k=1}^{\infty} \frac{j_k(x+h) - j_k(x)}{h}. \end{aligned}$$

下证明

$$\limsup_{h \rightarrow 0} (F + G) \leq \limsup_{h \rightarrow 0} F + \limsup_{h \rightarrow 0} G.$$

这是因为

$$\sup_{x \in [-\delta, \delta]} (F + G) \leq \sup_{x \in [-\delta, \delta]} F + \sup_{x \in [-\delta, \delta]} G.$$

两边取极限即可. 将无穷求和作有限次拆分, 就可以得到

$$\begin{aligned} \overline{D}(J)(x) &\leq \sum_{k=1}^N \limsup_{h \rightarrow 0} \frac{j_k(x+h) - j_k(x)}{h} + \limsup_{h \rightarrow 0} \sum_{k > N} \frac{j_k(x+h) - j_k(x)}{h} \\ &= \sum_{k=1}^N \limsup_{h \rightarrow 0} \frac{j_k(x+h) - j_k(x)}{h} + \overline{D}(J_N)(x). \end{aligned}$$

当 $x \notin \{x_1, \dots, x_N\}$ 时, 存在 x 的一个邻域内 j_k 为常值, 此时前一项为 0. 因此当 $\overline{D}(J)(x) > \varepsilon$ 时一定有 $\overline{D}(J_N)(x) > \varepsilon$. 这已经足以证明

$$E_\varepsilon \subset E_{\varepsilon, N} \cup \{x_1, \dots, x_N\}.$$



Remark. 主要的注意点是 \limsup 的一些性质, 对利用 Fatou 引理也有用. 比如说最基础的有如果 $f \leq g$, 那么

$$\limsup_{x \rightarrow x_0} f(x) \leq \limsup_{x \rightarrow x_0} g(x).$$

以及常数也可以提出来:

$$\limsup_{x \rightarrow x_0} (f(x) + C) = \limsup_{x \rightarrow x_0} f(x) + C.$$

还有本题得到的:

$$\limsup_{x \rightarrow x_0} (f(x) + g(x)) \leq \limsup_{x \rightarrow x_0} f(x) + \limsup_{x \rightarrow x_0} g(x).$$

它们的证明都可以通过 $\limsup_{x \rightarrow x_0} = \lim_{t \rightarrow 0} \sup_{x_0 - t < x < x_0 + t}$ 来证出. 比较有趣的是下面的

$$\lim_{x \rightarrow x_0} g(x) \text{ 存在} \Rightarrow \limsup_{x \rightarrow x_0} (f(x) + g(x)) = \limsup_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x).$$

这可以通过上面几条组合起来证出.

Problem 51

构造一个单调增函数使得其不连续点集恰好是有理数点集.

Proof. 设 $\mathbb{Q} = \{x_1, x_2, \dots\}$, 则取

$$f_n = \begin{cases} \frac{1}{3^n}, & x > x_n; \\ 0, & x \leq x_n. \end{cases}$$

那么 $F = \sum_{n=1}^{\infty} f_n$ 是单调的收敛级数. 接下来证明 F 在 \mathbb{Q} 上不连续: 首先对任意 x_n 取其邻域 $(x_n - \delta, x_n + \delta)$ 使得

$$x_1, \dots, x_{n-1} \notin (x_n - \delta, x_n + \delta).$$

那么对任意 $y \in (x_n, x_n + \delta)$, 有

$$F(y) - F(x_n) = \sum_{k=n}^{\infty} (f_k(y) - f_k(x_n)) \geq \frac{1}{3^n} - \sum_{k=n+1}^{\infty} \frac{1}{3^k} = \frac{1}{2 \cdot 3^n}.$$

而对 $x \notin \mathbb{Q}$, 则对任意 N , 存在 δ 使得

$$x_1, \dots, x_n \notin (x - \delta, x + \delta).$$

因此对任意 $y \in (x - \delta, x + \delta)$, 有

$$|F(y) - F(x)| \leq \sum_{k=n+1}^{\infty} |f_k(y) - f_k(x)| \leq \sum_{k=n+1}^{\infty} \frac{1}{3^k} < \frac{1}{3^n}.$$

这就证明了

$$\lim_{y \rightarrow x} f(y) = f(x).$$

所以 F 单调增且在 \mathbb{Q} 上不连续.



Remark. 另外有一题要求 f 恰好有理点处不可微而在无理点处可微, 构造方法是把全体有理点 $\{q_n\}$ 取出来, 然后构造

$$f(x) = \sum_{n=1}^{\infty} \frac{|x - q_n|}{3^n}.$$

构造还是很精妙的, 它使得在每个有理点附近导数有一项正负号出现问题. 另一种更直观的理解是对这题中的函数 F 求积分, 即对 f_n 积分后再相加:

$$\int_{q_n}^x f_n(t) dt = \begin{cases} \frac{x-q_n}{3^n}, & x > x_n; \\ 0, & x \leq x_n. \end{cases}$$

也可以证明这样定义出来的 F 是满足要求的.

Problem 52

在 \mathbb{R} 上, 对可测函数 f 定义单边极大函数

$$\mathcal{M}_+ f(x) = \sup_{h>0} \frac{1}{h} \int_x^{x+h} |f(y)| dy$$

记 $E_\alpha^+ = \{x \in \mathbb{R} : \mathcal{M}_+ f(x) > \alpha\}$, 证明

$$m(E_\alpha^+) = \frac{1}{\alpha} \int_{E_\alpha^+} |f(y)| dy.$$

f 是否要求可积?

Proof. 如果 f 在任意有限长度区间上可积, 我们考虑函数

$$F(s) = \int_0^s (|f(y)| - \alpha) dy,$$

那么 $F(s) < +\infty$, 所以可以作如下拆分:

$$\begin{aligned} s \in E_\alpha^+ &\Leftrightarrow \exists t > s \text{ s.t. } \frac{1}{t-s} \int_s^t |f(y)| dy > \alpha \\ &\Leftrightarrow \exists t > s \text{ s.t. } \int_0^t (|f(y)| - \alpha) dy > \int_0^s (|f(y)| - \alpha) dy \\ &\Leftrightarrow \exists t > s \text{ s.t. } F(t) > F(s). \end{aligned}$$

这恰好是 Sunrise Lemma 的利用条件. 由于 f 在任意有限长度区间上可积, 故 F 在 \mathbb{R} 上连续, 所以可设

$$E_\alpha^+ = \bigcup_k (a_k, b_k).$$

- 如果不存在无穷区间, 则

$$\begin{aligned} F(a_k) = F(b_k) &\Rightarrow \int_{a_k}^{b_k} (|f(y)| - \alpha) dy = 0 \\ &\Rightarrow \int_{E_\alpha} (|f(y)| - \alpha) dy = 0 \\ &\Rightarrow \int_{E_\alpha} |f(y)| dy = \alpha \cdot m(E_\alpha). \end{aligned}$$

- 如果存在右无穷区间 $(a, +\infty)$, 对 $x_0 \in (a, +\infty)$ 考虑

$$D = \{x \in [x_0, +\infty) : F(x) \geq F(x_0)\}.$$

我们称 $\sup D = +\infty$, 否则设 $\sup D = x_1$, 则由 F 的连续性可知 $x_1 \in D$. 又因为 $x_1 \in (a, +\infty)$, 故存在 $x_2 > x_1$ 使得 $F(x_2) > F(x_1) \geq F(x_0)$, 所以 $x_2 \in D$, 与 x_1 是上确界矛盾! 因此对任意 N , 可取 $x \in D$ 使得 $x > x_0 + N$. 由于

$$F(x) \geq F(x_0) \Rightarrow \int_{x_0}^x |f(y)| dy \geq \alpha \cdot (x - x_0) \geq \alpha \cdot N,$$

因此

$$\int_{x_0}^{+\infty} |f(y)| dy = \infty = m(E_\alpha^+).$$

- 如果存在左无穷区间 $(-\infty, b)$ 而不存在右无穷区间, 则对任意 $x > b$ 有 $F(b) \geq F(x)$. 同样地, 对任意 $x_0 \in (-\infty, b)$ 考虑

$$D = \{x \in [x_0, +\infty) : F(x) \geq F(x_0)\}.$$

同样的论证可以得到 $\sup D \geq b$. 因此存在 $x \geq b$ 使得 $F(b) \geq F(x) \geq F(x_0)$. 因此

$$\int_{x_0}^b |f(y)| dy \geq \alpha \cdot (b - x_0).$$

令 $x_0 \rightarrow -\infty$ 即可得

$$\int_{-\infty}^b |f(y)| dy = \infty = m(E_\alpha^+).$$

最后, 如果 f 满足存在有限长度区间 $I = [a, b]$ 使得

$$\int_a^b |f(y)| dy = +\infty,$$

则对任意 $x < a$ 均有 $\mathcal{M}_+ f(x) = +\infty$, 所以每个 E_α^+ 测度均为无穷.

如果 $x \notin E_\alpha^+$, 则对任意 $h > 0$ 均有

$$\int_x^{x+h} |f(y)| dy \leq h\alpha.$$

取 $s = \inf\{x : x \notin E_\alpha^+\}$, 则 $I \setminus [s, +\infty)$ 上函数的积分也为 $+\infty$ (因为只去掉了积分值有限的一段), 而我们有 $I \setminus [s, +\infty) \subset E_\alpha^+$, 所以

$$\int_{E_\alpha^+} |f(y)| dy = +\infty.$$

所以也是满足要求的, 综上命题成立.



10 第五次作业

Problem 53

如果 F 是定义在 $[a, b]$ 上的可微函数, 并且 F' 是 Riemman 可积的, 证明 F 是有界变差函数, 并且

$$T_F([a, b]) = \int_a^b |F'(x)| dx.$$

Proof. 由 $T_F([a, b])$ 的定义可知

$$\begin{aligned} T_F([a, b]) &= \sup_{\sigma} |F(x_{k+1}) - F(x_k)| = \sup_{\sigma} (x_{k+1} - x_k) |F'(\xi_k)| \\ &= \int_a^b |F'(x)| dx. \end{aligned}$$

最后一步是 F' 的 Riemann 积分的定义. 由于 Riemann 可积能推出 Lebesgue 可积因此两者相同. \square

Problem 54

如果 $F \in BV([a, b])$, 定义 $G(x) = T_F([a, x])$, 那么 G 几乎处处可微并且 $G'(x) = |F'(x)|$, a.e.

Proof. 考虑正变差和负变差 $P_F(a, x)$ 和 $N_F(a, x)$, 它们满足

$$\begin{aligned} G(x) &= T_F(a, x) = P_F(a, x) + N_F(a, x); \\ F(x) - F(a) &= P_F(a, x) - N_F(a, x). \end{aligned}$$

而 P_F 和 N_F 均是非负单调增函数, 所以

$$G(x) - F(x) + F(a) = 2N_F(a, x)$$

是非负单调增函数. 由单调增函数 Lebesgue 定理,

$$\int_a^b (G'(x) - F'(x)) dx \leq G(b) - F(b) - G(a) + F(a) = G(b) - (F(b) - F(a)).$$

由此可得

$$\int_a^b (G'(x) - |F'(x)|) dx \leq T_F(a, b) - (F(b) - F(a)). \quad (1)$$

接下来我们构造一系列函数 F_n 满足

- $|F'_n(x)| = |F'(x)|$, a.e., $T_{F_n}(a, b) = T_F(a, b)$.
- $T_{F_n}(a, b) - (F_n(b) - F_n(a)) < \varepsilon$.

取分割 $\sigma = \{a = x_0, x_1, \dots, x_m = b\}$ 使得

$$T_F(a, b) - \sum_{i=0}^{m-1} |F(x_{i+1}) - F(x_i)| < \varepsilon.$$

取

$$F_n(x) = F(a) + \sum_{i=0}^{s-1} |F(x_{i+1}) - F(x_i)| + \operatorname{sgn}(F(x_{s+1}) - F(x_s)) \cdot (F(x) - F(x_s)),$$

$$x_s \leq x \leq x_{s+1}.$$

那么 F_n 满足

$$F_n(x_s) = F(a) + \sum_{i=0}^{s-1} |F(x_{i+1}) - F(x_i)|;$$

$$F_n(x) - F_n(x_s) = F(x) - F(x_s), \quad x_s \leq x \leq x_{s+1}.$$

$$\text{或 } F_n(x) - F_n(x_s) = -(F(x) - F(x_s)), \quad x_s \leq x \leq x_{s+1}.$$

因此 $|F'_n(x)| = |F'(x)|$, $x \in (x_s, x_{s+1})$ 且 $T_{F_n}(x_s, x_{s+1}) = T_F(x_s, x_{s+1})$, 这推出

$$T_F(a, b) = \sum_{i=0}^{m-1} T_F(x_i, x_{i+1}) = \sum_{i=0}^{m-1} T_{F_n}(x_i, x_{i+1}) = T_{F_n}(a, b); \quad (2)$$

$$F_n(b) - F_n(a) = \sum_{i=0}^{m-1} |F(x_{i+1}) - F(x_i)| > T_F(a, b) - \varepsilon. \quad (3)$$

因此将 F_n 代入 (1) 可知

$$\begin{aligned} \int_a^b (G'(x) - |F'_n(x)|) dx &\leq T_{F_n}(a, b) - (F_n(b) - F_n(a)). \\ \int_a^b (G'(x) - |F'(x)|) dx &\leq \varepsilon. \end{aligned}$$

由 ε 的任意性可得

$$\int_a^b (G'(x) - |F'(x)|) dx = 0.$$

由于 $G'(x) - |F'(x)|$ 为非负函数, 故 $G'(x) - |F'(x)| = 0$, a.e. 证毕. \square

Problem 55

根据定义直接证明 Cantor-Lebesgue 函数不是绝对连续的.

Proof. 由 Cantor 集的构造过程可知 $[0, 1] \setminus \mathcal{C}$ 可以写成一些区间的并

$$[0, 1] \setminus \mathcal{C} = \bigcup_{k=1}^{\infty} I_k.$$

并且 Cantor-Lebesgue 函数在每段 I_k 上均为常值. 由于这些区间的测度之和为 1, 因此对任意 $\varepsilon > 0$ 存在 N 使得

$$\sum_{k=1}^N m(I_k) > 1 - \varepsilon.$$

我们把这些开区间按照左右关系重新排列为 I_{i_1}, \dots, I_{i_k} , 并且设

$$I_{i_k} = (a_k, b_k), \quad b_0 := 0 \leq a_1 < b_1 < a_2 < b_2 < \dots < a_N < b_N \leq 1 := a_{N+1}.$$

考虑它的变差

$$\begin{aligned} 1 = F(1) - F(0) &= \sum_{k=0}^N (F(a_{k+1}) - F(b_k)) + \sum_{k=1}^N (F(b_k) - F(a_k)) \\ &= \sum_{k=0}^N (F(a_{k+1}) - F(b_k)), \\ \sum_{k=0}^N (a_{k+1} - b_k) &\leq \varepsilon. \end{aligned}$$

那么 Cantor-Lebesgue 函数就不是绝对连续函数.



Problem 56

绝对连续函数的复合不一定是绝对连续函数: 证明 (在 $[0, 1]$ 上) $f = x^3 \sin(x^{-1})$, $g = y^{\frac{1}{3}}$ 都是绝对连续函数, 但是 $g \circ f$ 不是绝对连续函数.

Proof. 先证明 f 在 $[0, 1]$ 上 C^1 , 这是因为

$$f' = 3x^2 \sin(x^{-1}) + x^3 \cos(x^{-1}) \cdot (-x^{-2}) = 3x^2 \sin(x^{-1}) - x \cos(x^{-1}).$$

在 $x > 0$ 时有 f' 恒为连续函数, 而当 $x \rightarrow 0$ 时 $f' \rightarrow 0$. 而由定义可知确实有 $f'(0) = 0$, 所以 f 是 C^1 函数从而也是绝对连续函数.

再证明 $g(y) = y^{\frac{1}{3}}$ 是绝对连续函数. 这是因为 $y^{-\frac{1}{3}}$ 是上凸递增函数, 所以对任意长度为 ε 的一些区间其变差最大值在 $[0, \varepsilon]$ 处取到. 所以对任意 $\varepsilon < \delta^3$ 均有长度为 ε 的区间变差不超过 δ , 根据定义结论成立.

现在考虑 $g \circ f = x \sin^{\frac{1}{3}}(x^{-1})$. 则考虑 $x_k^{-1} = 2k\pi + \frac{\pi}{2}$. $y_k^{-1} = 2k\pi + \frac{3\pi}{2}$ 可知 $g \circ f(x_k) = x_k$, $g \circ f(y_k) = -y_k$. 因此区间列 $\bigcup_k [y_k, x_k]$ 上的变差是

$$\sum_k (y_k + x_k) = \sum_k \left((2k\pi + \frac{\pi}{2})^{-1} + (2k\pi + \frac{3\pi}{2})^{-1} \right) = +\infty.$$

因此对任意小的 ε , 总可以找到 $x_k < \varepsilon$, 并且 $[y_i, x_i] (i \geq k)$ 上的变差之和仍然为 $+\infty$, 可取出其中有限多个区间使得其变差大于某个固定的 N , 并且这些区间长度之和小于 ε . 所以 $g \circ f$ 不是绝对连续函数.



Remark. 直观上来说确实不会有保复合性成立, 因为无法控制像集.

Problem 57

证明 $AC([a, b])$ 是 $BV([a, b])$ 的线性子空间.

Proof. 我们先证明 $AC([a, b])$ 是线性空间. 设 $f, g \in AC([a, b])$, 考虑 $f + \lambda g$. 对任意 $\delta > 0$ 取 ε 使得对长度和小于 ε 的有限区间并 I , 均有 f, g 在 I 上变差小于 δ . 因此 $f + \lambda g$ 在 I 上变差小于 $(1 + |\lambda|)\delta$. 当 $\delta \rightarrow 0$ 时命题对任意小的数均成立, 所以 $f + \lambda g$ 是绝对连续函数.

另一方面, 绝对连续函数一定是有界变差函数. 在条件中取 $\varepsilon = 1$ 得到 δ , 然后把 $[a, b]$ 分割为长度小于 δ 的一些区间, 设该分割为 τ . 对于任何的分割 σ , 考虑 $\sigma \cup \tau$ 分割后变差不减. 然而对于 τ 中的每个小区间, 可以一同取出所有它细分得到的区间, 这些区间长度之和小于 δ , 所以它们整体的变差永远小于 1, 在 $[a, b]$ 上的全变差就不会超过分割 τ 产生的变差, 这一定是一个有限值.

综上, $AC([a, b])$ 是 $BV([a, b])$ 的一个线性子空间.



Problem 58

证明 $AC([a, b])$ 在范数 $\|f\| = \sup |f| + \|f'\|_{L^1}$ 下构成 Banach 空间.

Proof. 由于 $\sup |f|$ 是范数且 L^1 是 a.e. 意义下的范数 (仅在非退化性时函数 f' 几乎处处为零等价于 f' 的 L^1 范数为零), 所以 $\|f\|$ 是一个范数. 下证明该范数完备, 即证任意 Cauchy 列收敛.

任取函数列 $f_n \in AC([a, b])$ 使得对任意 $\varepsilon > 0$, 存在 N 使得对任意 $n, m > N$ 均有 $\|f_n - f_m\| < \varepsilon$. 那么对任意 $x \in [a, b]$ 均有

$$|f_n(x) - f_m(x)| \leq \sup_{x \in [a, b]} |f_n - f_m| < \varepsilon.$$

因此 f_n 逐点收敛到一个函数 f , 只需证明这个函数是 AC 的. 由于 f_n 都是绝对连续函数, 因此它们满足

$$\int_a^x f'_n = f_n(x) - f_n(a).$$

两边取极限可得

$$\lim_{n \rightarrow \infty} \int_a^x f'_n = \lim_{n \rightarrow \infty} (f_n(x) - f_n(a)) = f(x) - f(a).$$

由于 L^1 是完备赋范空间且 f'_n 是 L^1 空间里的 Cauchy 列, 所以存在 $g \in L^1$ 使得 $f'_n \xrightarrow{L^1} g$. 这将导致对任意可测集 E ,

$$\lim_{n \rightarrow \infty} \int_E |f'_n - g| = 0 \Rightarrow \lim_{n \rightarrow \infty} \int_E f'_n = \int_E g.$$

特别地对每个 x 取 $E = [a, x]$, 这就得到

$$\int_a^x g = f(x) - f(a), \quad \forall x \in [a, b].$$

由积分的连续性就推出 f 为绝对连续函数, 并且 $f' = g$, a.e. 那么分别根据逐点收敛和 L^1 收敛可得

$$\lim_{n \rightarrow \infty} \sup |f - f_n| = 0, \quad \lim_{n \rightarrow \infty} \|f' - f'_n\| = 0.$$

所以 $\lim_{n \rightarrow \infty} \|f - f_n\| = 0$, 因此 f_n 确实在该范数下收敛到 f 且 $f \in AC$, 证毕.



Remark. 这题有更强的结论: 如果 f'_n 是 $[a, b]$ 上的 L^1 Cauchy 列, 并且只要它们的初值 $f_n(a)$ 是 Cauchy 列那么 f_n 就收敛到一个 AC 函数.

这是因为我们不必选择上面的 a.e. 收敛的那个 f , 可以直接选定 $f'_n \xrightarrow{L^1} g$, 那么直接令

$$f(x) = \lim_{n \rightarrow \infty} f_n(a) + \int_a^x g(t) dt$$

同样也可以得到 $f_n \rightarrow f$ (不是 a.e.), 这是因为

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} (f_n(a) + \int_a^x f_n(t) dt) = f(a) + \int_a^x g(t) dt = f(x).$$

由积分的绝对连续性显然推出 f AC.

Problem 59

证明 $AC([a, b])$ 是可分的, 即是存在可数稠密子集.

Proof. 首先证明对所有 $[a, b]$ 上的阶梯函数 g ,

$$G(x) = G(a) + \int_a^x g, \quad G(a) \in \mathbb{R}, g \text{ 阶梯}$$

构成了 $AC([a, b])$ 的稠密子集.

对 $F \in AC([a, b])$, $F' \in L^1$, 由此可取出一列阶梯函数 g_n 使得 $g_n \xrightarrow{L^1} F'$. 具体而言,

$$\lim_{n \rightarrow \infty} \int_a^b |g_n - F'| = 0.$$

那么考虑 $G_n = F(a) + \int_a^x g_n$, 则 G_n 满足 $G'_n = g_n$, a.e. 因此我们有

$$|G_n(x) - F(x)| = \left| \int_a^x (g_n - F') \right| \leq \int_a^b |g_n - F'|; \quad (1)$$

$$\lim_{n \rightarrow \infty} \|G_n - F\| = \sup |G_n - F| + \|G'_n - F'\|_{L^1} \leq \lim_{n \rightarrow \infty} 2 \|g_n - F'\|_{L^1} = 0.$$

再证明可以用在有理点分割且取值为有理数的阶梯函数与起始值落在 \mathbb{Q} 中的 G 来逼近所有 G . 前者是因为根据稠密性可以取这样的一些阶梯函数列 g_n 逐点逼近 g 且它们均有界, 所以根据控制收敛定理 $g_n \xrightarrow{L^1} g$. 此外, 取 $G_n = q_n + \int_a^x g_n$ 且 $\lim_{n \rightarrow \infty} |q_n - G(a)| = 0$, 则 (1) 变为

$$|G_n(x) - G(x)| = \left| \int_a^x (g_n - G') + (q_n - G(a)) \right| \leq \int_a^b |g_n - G'| + |q_n - G(a)|;$$

$$\lim_{n \rightarrow \infty} \sup |G_n - G| \leq \lim_{n \rightarrow \infty} \|G'_n - G'\|_{L^1} + \lim_{n \rightarrow \infty} |q_n - G(a)| = 0.$$

因此这列 G_n 可以逼近 G , 而它们构成可数集, 所以 $AC([a, b])$ 可分.



Problem 60

证明如果有界变差函数 F 满足

$$\int_a^b |F'| dx = T_F(a, b)$$

那么 F 是绝对连续函数.

Proof. 将 T_F 写为正变差 P_F 与负变差 N_F 之和, 则可以得到


$$\begin{aligned} T_F(a, x) &= P_F(a, x) + N_F(a, x); \\ F(x) - F(a) &= P_F(a, x) - N_F(a, x). \end{aligned}$$

条件转化为

$$\int_a^b |P'_F - N'_F| = P_F(b) + N_F(b).$$

由于 P_F, N_F 都是单调增函数, 因此

$$\begin{aligned} P_F(x) + N_F(x) &\geq \int_a^x (P'_F + N'_F) \geq \int_a^x |P'_F - N'_F|. \\ P_F(b) + N_F(b) - P_F(x) - N_F(x) &\geq \int_x^b (P'_F + N'_F) \geq \int_x^b |P'_F - N'_F|. \end{aligned}$$

求和可知上述不等号均取等, 因此 $P_F + N_F$ 为绝对连续函数. 由于它们均单调增, 所以在一段区间内 P_F 和 N_F 分别的变差一定比 $P_F + N_F$ 小. 这推出 P_F 和 N_F 均是绝对连续函数, 所以 $F(x) - F(a) = P_F - N_F$ 也是绝对连续函数, 故 F 是绝对连续函数. 

Remark. 一般来讲, 根据正负变差分解容易看出恒有

$$T'_F = P'_F + N'_F \geq |P'_F - N'_F| = |F'|$$

成立, 所以总有

$$T_F([a, b]) = T_F([a, b]) - T_F([a, a]) \geq \int_a^b T'_F \geq \int_a^b |F'|$$

成立, 这对任何 $F \in BV$ 成立.

Problem 61

如果 F 在 $[a, b]$ 上是连续的凸函数, 证明 $F \in AC([a, b])$.

Proof. 由于 F 为凸函数, 因此对任意的 $x < y$ 和 $h > 0$, 均有

$$F(x+h) - F(x) \leq F(y+h) - F(y).$$

由于 F 连续, 故可以取出 x_0 使得

$$F(x_0) = \min_{x \in [a, b]} F(x).$$

那么 F 在 $[x_0, b]$ 上单调增, 在 $[a, x_0]$ 上单调减. 现在对任意的两两不交的区间 $I_k = (a_k, b_k)$, $k = 1, \dots, n$, 不妨设 $a \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_n \leq b_n \leq b$, 设 $b_k - a_k = h_k$, 则一定有

$$a_k \leq b - h_n - \dots - h_k.$$

因此

$$\sum_{k=1}^n |F(a_k + h_k) - F(a_k)| \leq \sum_{k=1}^n \left| F(b - \sum_{i=k+1}^n h_i) - F(b - \sum_{i=k}^n h_i) \right|.$$

当 $\sum h_k \leq b - x_0$ 时, 所有 $F(b - \sum_{i=k}^n h_i)$ 均处在一个单调区间内, 上述绝对值均可去除:

$$\sum_{k=1}^n |F(a_k + h_k) - F(a_k)| \leq F(b) - F(b - \sum_{i=1}^n h_i).$$

而对任意 $\delta > 0$, 由 F 的连续性存在 $\varepsilon < b - x_0$ 使得 $F(b) - F(b - \varepsilon) < \delta$. 此时对总长度小于 ε 的有限区间族 I , F 在 I 上的总变差至多为 $F(b) - F(b - \varepsilon) < \delta$. \square

Problem 62

如果 f 在 $[0, 1]$ 上绝对连续, 定义

$$f_h(x) = \frac{f(x+h) - f(x)}{h}, \quad h > 0.$$

对于上述定义中取值大于 1 的点, 定义 $f(x) = f(1)$. 证明 $\lim_{h \rightarrow 0} \|f_h - f'\|_{L^1} = 0$.

Proof. 固定 $h > 0$, 则由 $f \in AC([0, 1])$ 可知

$$f_h(x) = \frac{1}{h} \int_x^{x+h} f'(t) dt.$$

因此

$$\begin{aligned} \|f_h - f'\|_{L^1} &= \int_a^b \left| \int_x^{x+h} \frac{1}{h} f'(t) dt - f'(x) \right| dx \\ &\leq \frac{1}{h} \int_a^b \int_x^{x+h} |f'(t) - f'(x)| dt dx \\ &= \frac{1}{h} \int_0^h \int_a^b |f'(x+s) - f'(x)| dx ds \\ &= \frac{1}{h} \int_0^h \|\tau_s f' - f'\|_{L^1} ds. \end{aligned}$$

积分号可交换来源于 $f \in AC([a, b]) \Rightarrow f' \in L^1$. 现在由于

$$\lim_{s \rightarrow 0} \|\tau_s f' - f'\|_{L^1} = 0,$$

因此对任意 $\varepsilon > 0$, 存在 h_0 使得对任意 $|s| \leq h_0$ 均有 $\|\tau_s f' - f'\|_{L^1} \leq \varepsilon$. 这可以推出

$$\frac{1}{h} \int_0^h \|\tau_s f' - f'\|_{L^1} ds \leq \frac{1}{h} \int_0^h \varepsilon ds = \varepsilon, \quad \forall h < h_0.$$

所以

$$\lim_{h \rightarrow \infty} \|f_h - f'\|_{L^1} = 0.$$

证毕.



Problem 63

证明 F 满足 Lipschitz 条件

$$|F(x) - F(y)| \leq M |x - y| \quad (1)$$

等价于 F 绝对连续且 $|F'| \leq M$, a.e.

Proof. 若 F 满足 (1), 则显然 $F \in AC$. 后者来源于对任意 $x \in [a, b]$,

$$\left| \frac{F(x+h) - F(x)}{h} \right| \leq M, \quad \forall h > 0.$$

取 $h \rightarrow 0$ 即得对任意可微点均有 $|F'| \leq M$. 而 $F \in AC$ 故 F 几乎处处可导, 所以 $|F'| \leq M$, a.e. 另一方面, 如果 F 绝对连续且 $|F'| \leq M$, a.e. 假设存在 $c < d$ 使得

$$|F(c) - F(d)| > M |c - d|,$$

则设 $E = \{x \in [c, d] : F \text{ 在 } x \text{ 处可微}\}$, 则由 Growth Lemma 可知

$$m^*(F(E)) \leq M m^*(E) = M(d - c).$$

由于 F 绝对连续, 则 F 把零测集映到零测集. 所以 $m^*(F([c, d])) \leq M(d - c)$. 但是又因为 F 连续, 故事实上有

$$m^*(F([c, d])) > |F(c) - F(d)| > M |c - d|.$$

矛盾. 所以两者等价.



Remark. 应该不需要用 Growth Lemma, 如果 F 绝对连续就直接有

$$|F(x) - F(y)| \leq \int_x^y |F'(t)| dt \leq \int_x^y M dt = M |x - y|.$$

从而 F 是 Lipschitz 的.

Problem 64

证明绝对连续函数满足 Lusin(N) 性质: 如果 $N \subset [a, b]$ 为零测集, F 为 $[a, b]$ 上的绝对连续函数, 那么 $F(N)$ 为零测集. 进一步证明 F 把可测集映到可测集.

Proof. 对任意 $\varepsilon > 0$, 取 δ 使得只要有限区间族的总测度和小于 δ 就有其变差小于 ε . 取开集 $O \supset N$ 满足 $m(O) < \delta$. 由于 $O \subset \mathbb{R}$, 设

$$O = \bigcup_{k=1}^{\infty} (a_k, b_k).$$

则对每个小区间 $[a_k, b_k]$, 由 F 的连续性可取出 c_k, d_k 使得

$$F(c_k) = \max_{x \in [a_k, b_k]} F(x), \quad F(d_k) = \min_{x \in [a_k, b_k]} F(x).$$

记 $I_k = (\min\{c_k, d_k\}, \max\{c_k, d_k\})$. 则对任意 N ,

$$m\left(\bigcup_{k=1}^N I_k\right) < \delta \Rightarrow \sum_{k=1}^N |F(c_k) - F(d_k)| < \varepsilon.$$

所以


$$\sum_{k=1}^{\infty} |F(c_k) - F(d_k)| \leq \varepsilon.$$

由连续性可知:

$$F([a_k, b_k]) \subset [F(d_k), F(c_k)];$$

$$m^*(F(N)) \leq m^*(F(O)) \leq \sum_{k=1}^{\infty} m^*(F([a_k, b_k])) \leq \sum_{k=1}^{\infty} |F(c_k) - F(d_k)| \leq \varepsilon.$$

令 $\varepsilon \rightarrow 0$ 即可得 $F(N)$ 为零测集.

由于 F 为连续函数, 因此 F 把闭区间映到闭区间, 所以 F 把开区间映到可测集, 故 F 把 G_δ 集映到可测集. 又因为 F 把零测集映到零测集, 而任意可测集都可以写为 G_δ 集去掉一个零测集, 所以 F 把任何可测集映到可测集. 

Problem 65

如果 F 是 $[a, b]$ 上的连续有界变差函数且把零测集映到零测集, 那么 F 为绝对连续函数.

Proof. 由于 F 是 $[a, b]$ 上的有界变差函数, 故 F 几乎处处可微. 设 E 为 F 的所有可微点构成的集合. 则对任意 $A \subset [a, b]$, 由 Growth Lemma 可知

$$m^*(F(E \cap A)) \leq \int_{E \cap A} |F'|.$$

由于 F 把零测集映到零测集, 故


$$m(F(E \setminus A)) = 0.$$

而根据 $F \in BV([a, b])$ 可知 $F' \in L^1$, 所以根据积分的连绝对续性可知对任意 $\varepsilon > 0$, 存在 δ 使得只要 $m^*(E) \leq \delta$ 就有

$$m^*(F(E)) \leq \int_E |F'| \leq \varepsilon.$$

特别地, 当 $E = \bigcup_{k=1}^n [a_k, b_k]$ 时, 由 F 的连续性可知

$$|F(b_k) - F(a_k)| \leq m^*(F([a_k, b_k])) \Rightarrow \sum_{k=1}^n |F(b_k) - F(a_k)| \leq m^*(F(E)) \leq \varepsilon.$$

因此 $F \in AC([a, b])$. 另一边的三条性质中零测集映到零测集在上题中已证, 剩下两者显然. 

Problem 66

完成 Lebesgue 分解定理的证明：对于定义在 $[a, b]$ 上的单调增函数 F ，我们有分解

$$F = F_{AC} + F_C + J$$

其中每个函数均单调增，且满足

- F_{AC} 为绝对连续函数；
- F_C 为连续函数且 $F'_C = 0$, a.e.
- J 为跳跃函数，即存在一些 j_n 满足 $J(x) = \sum_{n=1}^{\infty} j_n(x)$ ，其中

$$j_n(x) = \begin{cases} 0, & a \leq x < x_n; \\ \beta_n, & x = x_n; \\ \alpha_n, & x_n < x \leq b \end{cases}, 0 \leq \beta_n \leq \alpha_n.$$

并且在差一个常数的意义下上述分解唯一。

Proof. 首先把单调增函数分解为连续函数和跳跃函数的和：所有不连续点 x 构成一个可数集 $\{x_n\}$. 记

$$\alpha_n = F(x_n + 0) - F(x_n - 0), \quad \beta_n = F(x_n) - F(x_n - 0).$$

那么可以定义跳跃函数

$$j_n(x) = \begin{cases} 0, & x < x_n; \\ \beta_n, & x = x_n; \\ \alpha_n, & x > x_n. \end{cases}$$

记 $J_n = \sum_{n=1}^{\infty} j_n(x)$ ，我们来证明 $F - J_F$ 是单调增的连续函数。只需证明左右极限均等于原函数取值：

$$\lim_{h \rightarrow 0, h > 0} (F(x \pm h) - J_F(x \pm h)) = F(x) - J_F(x).$$

由于 F 单调增，因此 $F(x \pm 0)$ 总存在。所以上式可化为证明

$$F(x \pm 0) - F(x) = \lim_{h \rightarrow 0, h > 0} J_F(x \pm h) - J_F(x).$$

由于 $\sum_{n > N} j_n(x)$ 可以被 $\sum_{n > N} \alpha_n$ 一致地控制，且 $\sum \alpha_n < +\infty$ ，所以 J_F 作为级

数关于 x 一致收敛, 这导出

$$\lim_{h \rightarrow 0, h > 0} J_F(x+h) = \sum_n \lim_{h \rightarrow 0, h > 0} j_n(x+h),$$

$$\lim_{h \rightarrow 0, h > 0} j_n(x+h) = \begin{cases} j_n(x), & x \neq x_n; \\ \alpha_n, & x = x_n. \end{cases} \quad \lim_{h \rightarrow 0, h > 0} j_n(x-h) = \begin{cases} j_n(x), & x \neq x_n; \\ 0, & x = x_n. \end{cases}$$

因此

$$\lim_{h \rightarrow 0, h > 0} J_F(x \pm h) - J_F(x) = \mathbf{1}_{x=x_n} \cdot \begin{cases} \alpha_n - \beta_n \\ -\beta_n \end{cases}.$$

而根据 α_n, β_n 的定义, 容易知道当 $x = x_n$ 时上面式子实际上就是 $F(x \pm 0) - F(x)$.

而 $x \neq x_n$ 时也有 $F(x \pm 0) - F(x) = 0$, 这就证明了连续性.

再证单调性, 设 $a \leq x < y \leq b$. 如果 x 为 F 的连续点, 那么

$$\begin{aligned} J_F(y) - J_F(x) &= \sum_{x_k \leq y} j_k(y) - \sum_{x_k \leq x} j_k(x) \\ &= \sum_{x < x_k \leq y} j_k(y) \leq F(y) - F(x). \end{aligned}$$

如果 x 为不连续点, 那么

$$\begin{aligned} J_F(y) - J_F(x) &= \sum_{x_k \leq y} j_k(y) - \sum_{x_k \leq x} j_k(x) \\ &= \sum_{x_k < x} \alpha_k + \sum_{x \leq x_k \leq y} j_k(y) - \sum_{x_k < x} \alpha_k - \beta_{x_n=x} \\ &= (\alpha_n - \beta_n) + \sum_{x < x_k \leq y} j_k(y) \leq F(y) - F(x). \end{aligned}$$

这就证明了单调性.

记 $F - J_F = G$, 下面把 G 分解为导函数几乎处处为零的函数和绝对连续函数之和. 考虑

$$F_{AC}(x) = G(a) + \int_a^x G'.$$

根据单调增函数的 Lebesgue 定理可知 $G' \in L^1$ 且恒有 $F_{AC}(x) \leq G(x)$, 因此令 $F_C = G - F_{AC}$ 可知 F_C 为单调增函数. 由 $G' \in L^1$ 及积分的绝对连续性可知 F_{AC} 为绝对连续函数. 再根据 Lebesgue 基本定理, 对所有 Lebesgue 点 x 有

$$F'_C(x) = G'(x) - F'_{AC}(x) = G'(x) - \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} G' = G'(x) - G'(x) = 0.$$

因此 $F'_C = 0$, a.e. 综上, F 可被分为满足要求的三个函数之和.



Problem 67

假设 F 是 \mathbb{R} 上处处可微的单调增函数, 满足

$$\lim_{x \rightarrow -\infty} F(x) = A, \quad \lim_{x \rightarrow \infty} F(x) = B.$$

证明

$$\int_{\mathbb{R}} F'(x) dx = B - A.$$

Proof. 考虑 F 在区间 $[-k, k]$ 上的限制, 记为 F_k . 由于 F_k 是单调增函数, 故 F'_k 几乎处处存在且 $F'_k \in L^1$. 因此利用 Riemann 情形的 Lebesgue 微积分基本定理可得

$$\int_{-k}^k F'(x) dx = F(k) - F(-k).$$

由于 $\lim_{k \rightarrow \infty} (F(k) - F(-k)) = B - A < +\infty$, 且 F'_k 为逐点逼近 F' 的非负函数列, 所以

$$\int_{\mathbb{R}} F' = \lim_{k \rightarrow \infty} \int_{-k}^k F' = \lim_{k \rightarrow \infty} (F(k) - F(-k)) = B - A.$$

证毕.



Remark. 单调+零测集映到零测集, 或者单调+处处可微都可推出 AC.

Problem 68

在 $[0, 1]$ 上考虑函数

$$F(x) = \begin{cases} x^a \sin(x^{-b}), & 0 < x \leq 1 \\ 0, & x = 0 \end{cases}.$$

对怎样的 α , F 在 $[0, 1]$ 上的 C^α 连续函数?

Proof. 由于 F 在任意的 $[\varepsilon, 1]$ 上是 C^1 的, 从而也是 Lipschitz 的. 所以只需看 $[0, \varepsilon]$ 上函数的动态. 首先讨论掉一些平凡情况: C^α 函数一定是连续函数, 所以

$$\lim_{x \rightarrow 0} F(x) = 0 \Rightarrow \lim_{x \rightarrow 0} |F(x)| = 0.$$

因此当 $a > 0$ 时 $\lim_{x \rightarrow 0} |F(x)| \leq \lim_{x \rightarrow 0} |x^a| = 0$, 当 $a \leq 0$ 时 F 在零处不连续, 所以下只讨论 $a > 0$ 的情形.

若 $b < 0$, 则 F 在某个 $[0, \varepsilon]$ 上是单调的并且

$$F(x) = x^a \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(x^{-b})^{2n-1}}{(2n-1)!} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)!} x^{a-b(2n-1)}.$$

对函数 $f(x) = x^t$, 当 $t \geq 1$ 时 f 在 $[0, 1]$ 上 Lipschitz 连续, 当 $t < 0$ 时 f 在 $[0, 1]$ 均不 Holder 连续, 当 $0 < t < 1$ 时可证明 f 恰好是 t -Holder 连续的. 这是因为 f 在

$[0, 1]$ 上上凸, 所以 $|x^t - y^t| \leq |0 - (x - y)^t|$. 而取 $x = 0$ 也可知对 $\alpha > t$ 函数不是 α -连续的. 所以此时 $F \in C^{a-b}$.

接下来考虑 $a > 0, b \geq 0$ 的情形. 取

$$x_k = (2k\pi + \frac{\pi}{2})^{-1/b}, \quad y_k = (2k\pi - \frac{\pi}{2})^{-1/b}$$

那么如果 F 是 α -Holder 连续的, 则


$$|x_k^a + y_k^a| \leq M \cdot |x_k - y_k|^\alpha.$$

只需考虑是否存在 M 使得

$$\left| (2k + \frac{1}{2})^{-a/b} + (2k - \frac{1}{2})^{-a/b} \right| \leq M \cdot \left| (2k + \frac{1}{2})^{-1/b} + (2k - \frac{1}{2})^{-1/b} \right|^\alpha.$$

而 $\text{LHS} = \Theta(k^{-a/b}), \text{RHS} = \Theta(k^{-\alpha/b-\alpha})$, 所以我们需要

$$\alpha \leq \frac{a}{b+1}.$$

所以 F 是 $\min(\frac{a}{b+1}, 1)$ -Holder 连续的. 

11 第六次作业

Problem 69: 微分的唯一性

考虑映射 $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, 假设存在两个线性映射 $A_i: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($i = 1, 2$), 使得对任意 $\mathbf{v} \rightarrow 0$ 及 $i = 1, 2$ 均有

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + A_i(\mathbf{v}) + o(\mathbf{v}).$$

证明 $A_1 = A_2$.

Proof. 假设 $A_1 \neq A_2$, 取 \mathbf{v} 使得 $\|\mathbf{v}\| = 1$ 且 $(A_1 - A_2)\mathbf{v} = \mathbf{t} \neq 0$. 那么对任意 $\lambda > 0$ 总有

$$\begin{aligned} f(\mathbf{x}_0 + \lambda\mathbf{v}) &= f(\mathbf{x}_0) + A_i(\lambda\mathbf{v}) + o(\lambda). \\ \Rightarrow A_1(\lambda\mathbf{v}) - A_2(\lambda\mathbf{v}) &= o(\lambda). \\ \Rightarrow (A_1 - A_2)(\mathbf{v}) &= o(1) = 0. \end{aligned}$$

矛盾. 所以 $A_1 = A_2$. 

Problem 70

$$\text{考察函数 } f(x, y) = \begin{cases} (x^2 + y^2) \sin \frac{1}{\sqrt{x^2 + y^2}}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

计算 f 的偏导数并证明 f 的两个偏导数在 $(0, 0)$ 处均不连续但 f 在 $(0, 0)$ 处可微.

Proof. 首先在 $(0,0)$ 处,

$$\lim_{(x,y) \rightarrow (0,0)} \left| \frac{(x^2 + y^2) \sin \frac{1}{\sqrt{x^2 + y^2}}}{\sqrt{x^2 + y^2}} \right| = \lim_{r \rightarrow 0} |r \sin r^{-1}| \leq \lim_{r \rightarrow 0} |r| = 0.$$

然而在 $y = 0$ 上

$$\frac{\partial f}{\partial x} = \begin{cases} 2x \sin x^{-1} - \cos x^{-1}, & x > 0; \\ 0, & x = 0. \end{cases}$$

这个函数在 $x = 0$ 处不连续, 所以偏导数在 $(0,0)$ 附近不连续.



Problem 71

假设函数 f 在 Ω 上的所有偏导数 $\frac{\partial f}{\partial x_i}$ 都存在并且一致有界, 证明 f 在 Ω 上连续. f 是否一定在 Ω 上可微? 如果是, 请给出证明, 否则举出反例.

Proof. 设对任意 $i = 1, \dots, n$ 和 $\mathbf{x} \in \Omega$ 均有

$$\left| \frac{\partial f}{\partial x_i}(\mathbf{x}) \right| < M.$$

固定 i, \mathbf{x} , 考虑一元函数 $g(t) = f(\mathbf{x} + t\mathbf{e}_i)$. 根据上述条件 $g'(t)$ 存在且 $|g'(t)| < M$. 因此 $(Mt - g)' > 0$ 及 $(g + Mt)' > 0$ 恒成立, 故 $|g(t) - g(0)| \leq Mt$. 这意味着

$$|f(\mathbf{x} + t\mathbf{e}_i) - f(\mathbf{x})| \leq Mt. \quad (1)$$

现在对任意 $\mathbf{x}, \mathbf{y} \in \Omega$ 满足 $|\mathbf{x} - \mathbf{y}| < \varepsilon$, 设 $\mathbf{x} - \mathbf{y} = \sum_{i=1}^n a_i \mathbf{e}_i$, 则由 (1) 可知

$$\left| f(\mathbf{y} + \sum_{i=1}^m a_i \mathbf{e}_i) - f(\mathbf{y} + \sum_{i=1}^{m+1} a_i \mathbf{e}_i) \right| \leq M a_{m+1}, \quad m = 0, \dots, n-1.$$

所以

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq M \cdot \sum_{m=1}^n a_m \leq Mn |\mathbf{x} - \mathbf{y}|.$$

所以 f 在 Ω 上连续. 另一方面, f 不一定可微: 构造

$$f(x, y) = \frac{x^3 + y^3}{x^2 + y^2}.$$

在 $\Omega = (-1, 1) \times (-1, 1)$ 上计算偏导:

$$\frac{\partial f}{\partial x} = \frac{3x^2(x^2 + y^2) - 2x(x^3 + y^3)}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y} = \frac{3y^2(x^2 + y^2) - 2y(x^3 + y^3)}{(x^2 + y^2)^2}.$$

那么偏导数有界, 因为

$$5(x^2 + y^2)^2 \geq 3x^2(x^2 + y^2) + 2x^4, (x^2 + y^2)^2 \geq |xy^3|.$$

所以

$$\left| \frac{\partial f}{\partial x} \right| \leq 6, \quad \frac{\partial f}{\partial x}|_{(0,0)} = \lim_{x \rightarrow 0} x = 0.$$

同理 f 关于 y 的偏导数也有界. 但是

$$f(x, kx) = \frac{k^3 + 1}{k^2 + 1}x \Rightarrow df|_{(0,0)}(1, k) = \frac{k^3 + 1}{k^2 + 1}.$$

显然这不是一个线性函数, 所以 f 在 $(0, 0)$ 处不可微.



Problem 72

设 $\Omega \subset \mathbb{R}^2$ 为开集, 我们用 (x, y) 来表示 \mathbb{R}^2 上的坐标. 函数 $f: \Omega \rightarrow \mathbb{R}$ 的偏导数 $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 处处存在. 如果偏导数 $\frac{\partial f}{\partial y}$ 在 Ω 上连续, 证明 f 在 Ω 上可微.

Proof. 对任意 $(x_0, y_0) \in \Omega$ 及 $(u, v) \in \mathbb{R}^2$, 我们证明

$$\lim_{h \rightarrow 0} \frac{f(x_0 + hu, y_0 + hv) - f(x_0, y_0)}{h} = u \frac{\partial f}{\partial x}|_{(x_0, y_0)} + v \frac{\partial f}{\partial y}|_{(x_0, y_0)}.$$

根据定义:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + hu, y_0) - f(x_0, y_0)}{h} = u \frac{\partial f}{\partial x}|_{(x_0, y_0)}.$$

因此只需证明

$$\lim_{h \rightarrow 0} \frac{f(x_0 + hu, y_0 + hv) - f(x_0 + hu, y_0)}{h} = v \frac{\partial f}{\partial y}|_{(x_0, y_0)}.$$

这是因为由于 $g(t) = f(x_0 + hu, y_0 + tv)$ 是 C^1 的, 所以存在 $\xi \in [0, h]$ 使得

$$g(h) - g(0) = g'(\xi) \cdot h = \frac{\partial f}{\partial (vy)}|_{(x_0 + hu, y_0 + \xi v)} \cdot h = v \frac{\partial f}{\partial y}|_{(x_0 + hu, y_0 + \xi v)} \cdot h.$$

所以

$$\lim_{h \rightarrow 0} \frac{f(x_0 + hu, y_0 + hv) - f(x_0 + hu, y_0)}{h} = \lim_{h \rightarrow 0} v \frac{\partial f}{\partial y}|_{(x_0 + hu, y_0 + \xi v)} = v \frac{\partial f}{\partial y}|_{(x_0, y_0)}.$$

最后一步根据 $\frac{\partial f}{\partial y}$ 连续及 $\lim_{h \rightarrow 0}(x_0 + hu, y_0 + \xi v) = (x_0, y_0)$ 得到.



Problem 73

考虑在 $\mathbb{R}^{n \times n}$ 上定义的行列式函数

$$\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \quad A \mapsto \det A.$$

任意给定 $A \in \mathbb{R}^{n \times n}$, 试计算 $d \det|_{x=A}$.

Proof. 设 $X = (x_{ij})$, 如果把每个输入 x_{ij} 看作一个变元, 那么

$$\det X = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \cdot x_{1\sigma(1)} \cdots x_{n\sigma(n)}$$

是关于 x_{ij} 的线性多项式. 因此对 x_{ij} 求导后可得:

$$\frac{\partial \det}{\partial x_{ij}} = \sum_{\sigma \in S_n, \sigma(i)=j} \operatorname{sgn} \sigma \cdot \prod_{k \neq i} x_{k\sigma(k)} = C_{ij}.$$

其中 C_{ij} 为 X 中去掉 i 行 j 列后所得的矩阵的行列式. 因此

$$\left. \frac{\partial \det}{\partial x_{ij}} \right|_{\mathbf{x}=A} = A_{ij}.$$

其中 A_{ij} 为 A 关于 i 行 j 列的代数余子式.



Problem 74

计算下列函数的偏导数:

$$(1) f(x, y, z) = x^{y^z}, \quad (2) f(x, y, z) = \tan \frac{xy}{z^2}.$$

Proof. (1) 化简表达式:

$$f(x, y, z) = x^{y^z} = e^{y^z \log x} = e^{e^{z \log y} \log x}.$$

计算偏导数

$$\begin{aligned} \frac{\partial f}{\partial x} &= y^z x^{y^z-1}; \quad \frac{\partial f}{\partial y} = e^{y^z \log x} z y^{z-1} = x^{y^z} z y^{z-1}. \\ \frac{\partial f}{\partial z} &= e^{e^{z \log y} \log x} \cdot e^{z \log y} \log x \cdot \log y = x^{y^z} y^z \log x \cdot \log y. \end{aligned}$$

(2)

$$\begin{aligned} \frac{\partial f}{\partial x} &= \sec^2 \frac{xy}{z^2} \cdot \frac{y}{z^2}, \quad \frac{\partial f}{\partial y} = \sec^2 \frac{xy}{z^2} \cdot \frac{x}{z^2}; \\ \frac{\partial f}{\partial z} &= \sec^2 \frac{xy}{z^2} \cdot -2xyz^{-3}. \end{aligned}$$



Problem 75

假设 f 为可微函数, 用 f 的偏导数表示下列函数的偏导数:

$$(1) u(x, y, z) = f(x, xy, xyz); \quad (2) u(x, y) = f(\log x + \frac{1}{y}).$$

Proof. 设 $f = f(u, v, w)$, 它对应的偏导数为 $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w}$.

(1) 设 $p(x, y, z) = (x, xy, xyz)$, 则

$$\begin{aligned}\frac{\partial u}{\partial \mathbf{x}}|_{(x,y,z)} &= \mathrm{d}(f \circ p)|_{(x,y,z)}(\mathbf{x}) = \mathrm{d}f|_{(x,xy,xyz)}(1, y, yz) \\ &= \frac{\partial f}{\partial \mathbf{u}}|_{(x,xy,xyz)} + y \frac{\partial f}{\partial \mathbf{v}}|_{(x,xy,xyz)} + yz \frac{\partial f}{\partial \mathbf{w}}|_{(x,xy,xyz)} \cdot \\ \frac{\partial u}{\partial \mathbf{y}}|_{(x,y,z)} &= \mathrm{d}(f \circ p)|_{(x,y,z)}(\mathbf{y}) = \mathrm{d}f|_{(x,xy,xyz)}(0, x, xz) \\ &= x \frac{\partial f}{\partial \mathbf{v}}|_{(x,xy,xyz)} + xz \frac{\partial f}{\partial \mathbf{w}}|_{(x,xy,xyz)} \cdot \\ \frac{\partial u}{\partial \mathbf{z}}|_{(x,y,z)} &= \mathrm{d}(f \circ p)|_{(x,y,z)}(\mathbf{z}) = \mathrm{d}f|_{(x,xy,xyz)}(0, 0, xy) \\ &= xy \frac{\partial f}{\partial \mathbf{w}}|_{(x,xy,xyz)} \cdot\end{aligned}$$

(2) 设 $p(x, y) = \log x + \frac{1}{y}$, 则

$$\begin{aligned}\frac{\partial u}{\partial \mathbf{x}}|_{(x,y)} &= \mathrm{d}(f \circ p)|_{(x,y)}(\mathbf{x}) = \mathrm{d}f|_{\log x + y^{-1}}(x^{-1}) = x^{-1} f'(\log x + y^{-1}). \\ \frac{\partial u}{\partial \mathbf{y}}|_{(x,y)} &= \mathrm{d}(f \circ p)|_{(x,y)}(\mathbf{y}) = \mathrm{d}f|_{\log x + y^{-1}}(-y^{-2}) = -y^{-2} f'(\log x + y^{-1}).\end{aligned}$$



Problem 76

求如下坐标变换的 Jacobi 矩阵 $J(f)$ 并计算 $\det J(f)$:

- (1) $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, (u, v) \mapsto (e^u \cos v, e^u \sin v);$
- (2) $f: \mathbb{R}^2 \setminus (0, 0) \rightarrow \mathbb{R}^2 \setminus (0, 0), (u, v) \mapsto (\frac{u}{u^2 + v^2}, \frac{v}{u^2 + v^2}).$

Proof. (1)

$$\frac{\partial f}{\partial \mathbf{u}} = (e^u \cos v, e^u \sin v), \quad \frac{\partial f}{\partial \mathbf{v}} = (-e^u \sin v, e^u \cos v).$$

所以它确定的 Jacobi 矩阵就是

$$J(f) = \begin{pmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos v \end{pmatrix}, \quad \det J(f) = e^{2u}(\cos^2 v + \sin^2 v) = e^{2u}.$$

(2)

$$\frac{\partial f}{\partial \mathbf{u}} = \left(\frac{v^2 - u^2}{(u^2 + v^2)^2}, \frac{-2uv}{(u^2 + v^2)^2} \right), \quad \frac{\partial f}{\partial \mathbf{v}} = \left(\frac{-2uv}{(u^2 + v^2)^2}, \frac{u^2 - v^2}{(u^2 + v^2)^2} \right).$$

所以它确定的 Jacobi 矩阵就是

$$\begin{aligned}J(f) &= \begin{pmatrix} \frac{v^2 - u^2}{(u^2 + v^2)^2} & \frac{-2uv}{(u^2 + v^2)^2} \\ \frac{-2uv}{(u^2 + v^2)^2} & \frac{u^2 - v^2}{(u^2 + v^2)^2} \end{pmatrix}, \quad \det J(f) = -\frac{(u^2 - v^2)^2}{(u^2 + v^2)^4} - \frac{4u^2 v^2}{(u^2 + v^2)^4} \\ &= -\frac{1}{(u^2 + v^2)^2}.\end{aligned}$$



Problem 77

我们考虑 \mathbb{R}^3 上的柱面坐标系: $x = r \cos \theta, y = r \sin \theta, z = z$. 这个坐标变换用映射来写就是

$$\Phi: \mathbb{R}_{>0} \times (0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3, \quad (r, \theta, z) \mapsto (r \cos \theta, r \sin \theta, z).$$

设 f 是 \mathbb{R}^3 上的二次可微函数. 当 $(x, y) \neq (0, 0)$ 时, 试通过计算来证明:

$$\begin{aligned} \frac{\partial f}{\partial r} &= \cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial \theta} = -r \sin \theta \frac{\partial f}{\partial x} + r \cos \theta \frac{\partial f}{\partial y}, \\ \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2 &= \left(\frac{\partial f}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial f}{\partial \theta} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2, \\ \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} &= \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}. \end{aligned}$$

Proof. 计算一阶偏导:

$$\begin{aligned} \frac{\partial f}{\partial r}|_{(r, \theta, z)} &= \mathrm{d}f|_{(x, y, z)}(\cos \theta, \sin \theta, 0) = \cos \theta \frac{\partial f}{\partial x}|_{(x, y, z)} + \sin \theta \frac{\partial f}{\partial y}|_{(x, y, z)}. \\ \frac{\partial f}{\partial \theta}|_{(r, \theta, z)} &= \mathrm{d}f|_{(x, y, z)}(-r \sin \theta, r \cos \theta, 0) = -r \sin \theta \frac{\partial f}{\partial x}|_{(x, y, z)} + r \cos \theta \frac{\partial f}{\partial y}|_{(x, y, z)}. \\ \frac{\partial f}{\partial z}|_{(r, \theta, z)} &= \mathrm{d}f|_{(x, y, z)}(0, 0, 1) = \frac{\partial f}{\partial z}|_{(x, y, z)}. \end{aligned}$$

$$\begin{aligned} \left(\frac{\partial f}{\partial x} \right)^2 + \frac{1}{r^2} \left(\frac{\partial f}{\partial \theta} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2 &= (\cos^2 \theta + \sin^2 \theta) \left(\frac{\partial f}{\partial x} \right)^2 \\ &+ (\sin^2 \theta + \cos^2 \theta) \left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2 = \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2. \end{aligned}$$

计算二阶偏导:

$$\begin{aligned} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) &= \frac{1}{r} \left(\cos \theta \frac{\partial f}{\partial x} + r \cos \theta (\cos \theta \frac{\partial^2 f}{\partial x^2} + \sin \theta \frac{\partial^2 f}{\partial x \partial y}) \right. \\ &\quad \left. + \sin \theta \frac{\partial f}{\partial y} + r \sin \theta (\cos \theta \frac{\partial^2 f}{\partial x \partial y} + \sin \theta \frac{\partial^2 f}{\partial y^2}) \right). \\ \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} &= \frac{1}{r^2} \left(-r \cos \theta \frac{\partial f}{\partial x} - r \sin \theta (-r \sin \theta \frac{\partial^2 f}{\partial x^2} + r \cos \theta \frac{\partial^2 f}{\partial x \partial y}) \right. \\ &\quad \left. + \frac{1}{r^2} \left(-r \sin \theta \frac{\partial f}{\partial y} + r \cos \theta (-r \sin \theta \frac{\partial^2 f}{\partial x \partial y} + r \cos \theta \frac{\partial^2 f}{\partial y^2}) \right) \right) \end{aligned}$$

上述两者相加即得

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

再加上 z 的贡献即得结论.



Problem 78

我们考虑 \mathbb{R}^3 上的球坐标系

$$(x, y, z) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta).$$

证明:

$$\begin{aligned} \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 &= \left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2} \left(\frac{\partial f}{\partial \theta}\right)^2 + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial f}{\partial \varphi}\right)^2, \\ \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2}. \end{aligned}$$

Proof. 算:

$$\begin{aligned} \frac{\partial f}{\partial r} &= \sin \theta \cos \varphi \frac{\partial f}{\partial x} + \sin \theta \sin \varphi \frac{\partial f}{\partial y} + \cos \theta \frac{\partial f}{\partial z}; \\ \frac{\partial f}{\partial \theta} &= r \cos \theta \cos \varphi \frac{\partial f}{\partial x} + r \cos \theta \sin \varphi \frac{\partial f}{\partial y} - r \sin \theta \frac{\partial f}{\partial z}; \\ \frac{\partial f}{\partial \varphi} &= -r \sin \theta \sin \varphi \frac{\partial f}{\partial x} + r \sin \theta \cos \varphi \frac{\partial f}{\partial y}. \\ \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) &= 2r \sin \theta \cos \varphi \frac{\partial f}{\partial x} + 2r \sin \theta \sin \varphi \frac{\partial f}{\partial y} + 2r \cos \theta \frac{\partial f}{\partial z} \\ &\quad + r^2 \sin \theta \cos \varphi \left(\sin \theta \cos \varphi \frac{\partial^2 f}{\partial x^2} + \sin \theta \sin \varphi \frac{\partial^2 f}{\partial x \partial y} + \cos \theta \frac{\partial^2 f}{\partial x \partial z} \right) \\ &\quad + r^2 \sin \theta \sin \varphi \left(\sin \theta \cos \varphi \frac{\partial^2 f}{\partial x \partial y} + \sin \theta \sin \varphi \frac{\partial^2 f}{\partial y^2} + \cos \theta \frac{\partial^2 f}{\partial y \partial z} \right) \\ &\quad + r^2 \cos \theta \left(\sin \theta \cos \varphi \frac{\partial^2 f}{\partial x \partial z} + \sin \theta \sin \varphi \frac{\partial^2 f}{\partial y \partial z} + \cos \theta \frac{\partial^2 f}{\partial z^2} \right). \\ \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) &= r(\cos^2 \theta - \sin^2 \theta) \cos \varphi \frac{\partial f}{\partial x} + r(\cos^2 \theta - \sin^2 \theta) \sin \varphi \frac{\partial f}{\partial y} - 2r \sin \theta \cos \theta \frac{\partial f}{\partial z} \\ &\quad + r \cos \theta \sin \theta \cos \varphi \left(r \cos \theta \cos \varphi \frac{\partial^2 f}{\partial x^2} + r \cos \theta \sin \varphi \frac{\partial^2 f}{\partial x \partial y} - r \sin \theta \frac{\partial^2 f}{\partial x \partial z} \right) \\ &\quad + r \cos \theta \sin \theta \sin \varphi \left(r \cos \theta \cos \varphi \frac{\partial^2 f}{\partial x \partial y} + r \cos \theta \sin \varphi \frac{\partial^2 f}{\partial y^2} - r \sin \theta \frac{\partial^2 f}{\partial y \partial z} \right) \\ &\quad - r \sin^2 \theta \left(r \cos \theta \cos \varphi \frac{\partial^2 f}{\partial x \partial z} + r \cos \theta \sin \varphi \frac{\partial^2 f}{\partial y \partial z} - r \sin \theta \frac{\partial^2 f}{\partial z^2} \right). \\ \frac{\partial^2 f}{\partial \varphi^2} &= -r \sin \theta \cos \varphi \frac{\partial f}{\partial x} - r \sin \theta \sin \varphi \frac{\partial f}{\partial y} \\ &\quad - r \sin \theta \sin \varphi \left(-r \sin \theta \sin \varphi \frac{\partial^2 f}{\partial x^2} + r \sin \theta \cos \varphi \frac{\partial^2 f}{\partial x \partial y} \right) \\ &\quad + r \sin \theta \cos \varphi \left(-r \sin \theta \sin \varphi \frac{\partial^2 f}{\partial x \partial y} + r \sin \theta \cos \varphi \frac{\partial^2 f}{\partial y^2} \right). \end{aligned}$$

所以

$$\begin{aligned}
 & \left(\frac{\partial f}{\partial \mathbf{r}}\right)^2 + \frac{1}{r^2} \left(\frac{\partial f}{\partial \theta}\right)^2 + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial f}{\partial \varphi}\right)^2 \\
 &= (\sin \theta \cos \varphi \frac{\partial f}{\partial x} + \sin \theta \sin \varphi \frac{\partial f}{\partial y} + \cos \theta \frac{\partial f}{\partial z})^2 + (\cos \theta \cos \varphi \frac{\partial f}{\partial x} + \cos \theta \sin \varphi \frac{\partial f}{\partial y} \\
 &\quad - \sin \theta \frac{\partial f}{\partial z})^2 + (\sin \varphi \frac{\partial f}{\partial x} - \cos \varphi \frac{\partial f}{\partial y})^2 \\
 &= \sin^2 \theta (\cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y})^2 + \cos^2 \theta (\cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y})^2 \\
 &\quad + (\sin \varphi \frac{\partial f}{\partial x} - \cos \varphi \frac{\partial f}{\partial y})^2 + \left(\frac{\partial f}{\partial z}\right)^2 \\
 &= \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2.
 \end{aligned}$$

$$\begin{aligned}
 & \frac{1}{r^2} \frac{\partial}{\partial \mathbf{r}} \left(r^2 \frac{\partial f}{\partial \mathbf{r}} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2} \\
 &= \left(\frac{2 \sin \theta \cos \varphi}{r} + \frac{(1 - 2 \sin^2 \theta) \cos \varphi}{r \sin \theta} - \frac{\cos \varphi}{r \sin \theta} \right) \frac{\partial f}{\partial x} \\
 &\quad + \left(\frac{2 \sin \theta \sin \varphi}{r} + \frac{(1 - 2 \sin^2 \theta) \sin \varphi}{r \sin \theta} - \frac{\sin \varphi}{r \sin \theta} \right) \frac{\partial f}{\partial y} \\
 &\quad + \left(\frac{2 \cos \theta}{r} - \frac{2 \cos \theta}{r} \right) \frac{\partial f}{\partial z} \\
 &\quad + (\sin^2 \theta \cos^2 \varphi + \cos^2 \theta \cos^2 \varphi + \sin^2 \varphi) \frac{\partial^2 f}{\partial x^2} \\
 &\quad + (\sin^2 \theta \sin \varphi \cos \varphi + \cos^2 \theta \sin \varphi \cos \varphi - \sin \varphi \cos \varphi) \frac{\partial^2 f}{\partial x \partial y} \\
 &\quad + (\sin \theta \cos \theta \cos \varphi - \cos \theta \sin \theta \cos \varphi) \frac{\partial^2 f}{\partial x \partial z} \\
 &\quad + (\sin^2 \theta \sin \varphi \cos \varphi + \cos^2 \theta \sin \varphi \cos \varphi + \sin \varphi \cos \varphi) \frac{\partial^2 f}{\partial x \partial y} \\
 &\quad + (\sin^2 \theta \sin^2 \varphi + \cos^2 \theta \sin^2 \varphi + \cos^2 \varphi) \frac{\partial^2 f}{\partial y^2} \\
 &\quad + (\sin \theta \cos \theta \sin \varphi - \cos \theta \sin \theta \sin \varphi) \frac{\partial^2 f}{\partial y \partial z} \\
 &\quad + (\cos \theta \sin \theta \cos \varphi - \sin \theta \cos \theta \cos \varphi) \frac{\partial^2 f}{\partial x \partial z} \\
 &\quad + (\cos \theta \sin \theta \sin \varphi - \sin \theta \cos \theta \sin \varphi) \frac{\partial^2 f}{\partial y \partial z} \\
 &\quad + (\cos^2 \theta + \sin^2 \theta) \frac{\partial^2 f}{\partial z^2} \\
 &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.
 \end{aligned}$$



Problem 79

对于常数 $k \in \mathbb{R}$, 我们称 \mathbb{R}^n 上的函数是 k 次齐次的, 如果对任意的 $\lambda > 0, x \neq 0$, 有 $f(\lambda x) = \lambda^k f(x)$. 证明线性函数是齐次函数.

Proof. 线性性蕴含 $f(\lambda x) = \lambda f(x)$.



Problem 80

假设 f 是可微的, 证明: f 是 k 次齐次函数当且仅当它满足 Euler 等式

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = kf.$$

Proof. 如果 f 是 k 次齐次函数, 对 $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, 我们有

$$f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x}).$$

对 λ 求导并在 \mathbf{x} 处取值即得

$$\left. \frac{df}{d\lambda} \right|_{\lambda=1} = kf(\mathbf{x}).$$

而当 $\mathbf{x} \neq 0$ 时,

$$\left. \frac{df}{d\lambda} \right|_{\lambda=1} = \lim_{t \rightarrow 0} \frac{|f(\mathbf{x} + t\mathbf{x}) - f(\mathbf{x})|}{t} = df|_{\mathbf{x}}(\mathbf{x}) = \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}.$$

当 $\mathbf{x} = 0$ 时, 需证式退化为 $kf(\mathbf{0}) = \mathbf{0}$, 显然成立.

另一方面, 若已知 Euler 等式, 则对任意 $\mathbf{x} \neq 0$ 均有

$$df|_{\mathbf{x}}(\mathbf{x}) = kf(\mathbf{x}).$$

则固定任何一个方向, 考虑该方向上的单位向量记为 \mathbf{e} , 那么将 $\lambda \mathbf{e}$ 代入上式可得

$$kf(\lambda \mathbf{e}) = df|_{\lambda \mathbf{e}}(\lambda \mathbf{e}) = \lambda df|_{\lambda \mathbf{e}}(\mathbf{e}) = \lambda \frac{\partial f}{\partial \mathbf{e}}(\lambda \mathbf{e}).$$

考虑一元函数 $g(\lambda) = f(\lambda \mathbf{e})$, 则 $kg(\lambda) = \lambda g'(\lambda)$, 即 $(\lambda^{-k}g)' = 0$, 所以 $g = C \cdot \lambda^k = g(1) \cdot \lambda^k$. 这就证明了 $f(\lambda \mathbf{e}) = \lambda^k f(\mathbf{e})$, 从而对任意向量结论均成立.



Problem 81

假设可微函数 f 是 k 阶线性函数. 证明对任意 $\mathbf{v} \in \mathbb{R}^n$, $\nabla_{\mathbf{v}} f$ 是 $k-1$ 次齐次函数.

Proof. 对表达式关于方向 \mathbf{v} 求偏导:

$$\begin{aligned} \nabla_{\mathbf{v}} f(\lambda \mathbf{x}) &= \lim_{t \rightarrow 0} \frac{f(\lambda \mathbf{x} + t\mathbf{v}) - f(\lambda \mathbf{x})}{t} = \lim_{t \rightarrow 0} \frac{\lambda^k (f(\mathbf{x} + t\lambda^{-1}\mathbf{v}) - f(\mathbf{x}))}{t} \\ &= \lambda^{k-1} \nabla_{\mathbf{v}} f(\mathbf{x}). \end{aligned}$$

即得结论.



Problem 82

$$\text{令 } f(x_1, \dots, x_n) = \det \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \\ \vdots & \ddots & \vdots \\ x_1^{n-1} & \dots & x_n^{n-1} \end{pmatrix}. \text{ 证明:}$$

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = \frac{n(n-1)}{2} f, \quad \sum_{i=1}^n \frac{\partial f}{\partial x_i} = 0.$$

Proof. 根据行列式的性质:

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda x_1 & \dots & \lambda x_n \\ \vdots & \ddots & \vdots \\ \lambda^{n-1} x_1^{n-1} & \dots & \lambda^{n-1} x_n^{n-1} \end{pmatrix} = \lambda^{\frac{n(n-1)}{2}} f(x_1, \dots, x_n).$$

根据 Euler 公式就可得到前一个等式. 后一个等式来源于 Vandemonde 行列式

$$\det \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \\ \vdots & \ddots & \vdots \\ x_1^{n-1} & \dots & x_n^{n-1} \end{pmatrix} = \prod_{i < j} (x_j - x_i).$$

所以实际上有

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i} = \sum_{i=1}^n \left(\sum_{p < q} \chi_{i,p,q} \cdot \prod_{\substack{i < j \\ (i,j) \neq (p,q)}} (x_j - x_i) \right).$$

其中

$$\chi_{i,p,q} = \begin{cases} 1, & i = q; \\ -1, & i = p; \\ 0, & \text{else.} \end{cases}$$

因此交换求和后可得到

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i} = \sum_{p < q} \prod_{\substack{i < j \\ (i,j) \neq (p,q)}} (x_j - x_i) \cdot \left(\sum_{i=1}^n \chi_{i,p,q} \right) = 0.$$



12 第七次作业

Problem 83

考虑映射函数 $f: \mathbb{R} \rightarrow \mathbb{R}$, 其中

$$f(x) = \begin{cases} x + x^2 \sin(\frac{\pi}{x}), & x \neq 0; \\ 0, & x = 0. \end{cases}$$


证明 f 在 \mathbb{R} 上可微分并且 $df(0) \neq 0$; 但是在 0 的任意小邻域内 f 均不是单射. 请对比反函数定理的条件, 说明有哪个条件没有被满足.

Proof. 根据定义可知

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = \lim_{x \rightarrow 0} (1 + x \sin(\frac{\pi}{x})) = 1.$$

所以 f 在 0 处可微且 $df(0) \neq 0$. 此外 f 在非零处是 C^1 的, 满足

$$\begin{aligned} f'(x) &= 1 + 2x \sin(\frac{\pi}{x}) + x^2 \cos(\frac{\pi}{x}) \cdot -\frac{\pi}{x^2} \\ &= 1 + 2x \sin(\frac{\pi}{x}) - \pi \cos(\frac{\pi}{x}). \end{aligned}$$

所以 $\lim_{x \rightarrow 0} f'(x)$ 不存在, 因此 f 在 0 附近不是 C^1 的, 所以不满足反函数定理: 这是因为在 0 附近导函数总是有正有负 (取 $\cos \frac{\pi}{x} = \pm 1$), 所以 f 不单调, 从而也不是单射. 

Problem 84

如果 f, g 是区域 Ω 上 k 次连续可微的函数, 证明 fg 也是 Ω 上的 k 次可微函数.

Proof. 当 $k = 0$ 时, f, g 连续显然可推出 fg 是连续的. 对 k 归纳, 假设 $k - 1$ 时成立, 考察 k 时情形.

由于 $f, g \in C^k$, 所以对任意 i, j 均有

$$\frac{\partial f}{\partial x_i}, \frac{\partial g}{\partial x_i} \in C^{k-1} \Rightarrow f \cdot \frac{\partial g}{\partial x_i}, g \cdot \frac{\partial f}{\partial x_i} \in C^{k-1}.$$

根据求导法则可知

$$\frac{\partial(fg)}{\partial x_i} = f \cdot \frac{\partial g}{\partial x_i} + g \cdot \frac{\partial f}{\partial x_i} \in C^{k-1}.$$

所以 fg 每个分量上均 C^{k-1} , 故 fg 是 C^k 的. 

Problem 85


考虑映射 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (e^x \cos y, e^x \sin y)$. 证明 f 在 $(x, y) \neq (0, 0)$ 处都满足反函数定理的要求但 f 不是单射也不是满射.

Proof. 计算偏导数:

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}}|_{(x,y)} &= (e^x \cos y, e^x \sin y) = f(x, y). \\ \frac{\partial f}{\partial \mathbf{y}}|_{(x,y)} &= (-e^x \sin y, e^x \cos y) = f(x, y + \frac{\pi}{2}).\end{aligned}$$

f 的偏导数均连续, 所以 f 是 C^1 的. 而在 (x, y) 处 Jacobi 行列式为

$$e^{2x} \cos^2 y + e^{2x} \sin^2 y = e^{2x} \neq 0.$$

所以 f 满足反函数定理的要求. 但 f 不满, 因为 $f^{-1}(0, 0) = \emptyset$, f 不单因为 $f(x, y) = f(x, y + 2\pi)$. 

Problem 86


我们考虑映射

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (2x - y + x^2y - 2y^5, x + 3y - 4x^2y^2)$$

证明: 存在 0 的开邻域 U 和 V 使得 $f: U \rightarrow V$ 是微分同胚.

Proof. 由于 f 的每个分量为一些多项式, 所以 f 的任意阶偏导均连续, 故 f 光滑. 下面验证反函数定理条件: 首先 $f(0, 0) = (0, 0)$, 所以两个开邻域均含 0. 其次求偏导可知

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}} &= (2 + 2xy, 1 - 8xy^2), \quad \frac{\partial f}{\partial \mathbf{y}} = (-1 + x^2 - 10y^4, 3 - 8x^2y). \\ \frac{\partial f}{\partial \mathbf{x}}|_{(0,0)} &= (2, 1), \quad \frac{\partial f}{\partial \mathbf{y}}|_{(0,0)} = (-1, 3).\end{aligned}$$

因此 $\det J_f|_{(0,0)} \neq 0$, 所以 $df|_{(0,0)}$ 可逆, 因此由反函数定理存在 U, V 使得结论成立. 

Problem 87

假设 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是连续可微的. 如果存在 $a > 0$ 使得对任意的 $x, y \in \mathbb{R}^n$, 我们都有

$$|f(x) - f(y)| \geq a |x - y|,$$

那么 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是 C^1 的微分同胚.

Proof. 由于 f 是 C^1 的, 那么只需证 f 是双射且 f^{-1} 也是 C^1 的. 首先证明 df 处处非退化. 根据 df 的定义,

$$f(x + \mathbf{v}) = f(x) + df|_x(\mathbf{v}) + o(\mathbf{v}).$$

代入条件可知对任意 \mathbf{v} 均有


$$|df|_x(\mathbf{v}) + o(\mathbf{v})| \geq a \cdot |\mathbf{v}|.$$

假设存在 \mathbf{v} 使得 $df|_x(\mathbf{v}) = 0$, 那么在上式中用 $\lambda\mathbf{v}$ 代替 \mathbf{v} 可得

$$|o(\lambda\mathbf{v})| \geq a \cdot |\lambda\mathbf{v}|.$$

所以令 $\lambda \rightarrow 0$ 就可知 $0 \geq a$, 矛盾.

所以 df 处处非退化, 那么由反函数定理, 对任意 $f(x) \in \mathbb{R}^n$, 存在开集 $V \ni x$ 及开集 $f(V) \ni f(x)$ 使得 f 在 V 上是双射, 而 $f(V) \subset \text{im } f$, 所以 $\text{im } f$ 是开集. 设有一列 $f(x_n) \rightarrow y$, 根据条件可知 x_n 也是 \mathbb{R}^n 上的 Cauchy 列, 可设 $x_n \rightarrow x$. 由 f 连续可直接得到 $f(x) = y$, 因此 $\text{im } f$ 为闭集.

综上 $\text{im } f$ 既开又闭, 所以 $\text{im } f$ 要么是全集要么是空集, 显然它不是空集, 故 f 为满射. 又由反函数定理可知 f^{-1} 在整个 \mathbb{R}^2 上 C^1 (因为 C^1 是局部性质), 证毕. 

Problem 88

假设 $f: \mathbb{R} \rightarrow \mathbb{R}$ 是 C^∞ 的并且存在 $0 \leq \alpha < 1$, 使得 $|f'(x)| \leq \alpha$. 定义

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x + f(y), y + f(x)).$$

Proof. • 由 $f \in C^\infty$ 可知 F 的任意阶偏导均存在连续, 所以 F 是光滑微分映射.

• 不妨设 $|y_2 - y_1| \geq |x_1 - x_2|$, 计算得

$$\begin{aligned} |F(x_1, y_1) - F(x_2, y_2)| &= |(x_1 - x_2 + f(y_1) - f(y_2), y_1 - y_2 + f(x_1) - f(x_2))| \\ &\geq |y_1 - y_2 + f(x_1) - f(x_2)|. \end{aligned}$$


根据 $|f'(x)| \leq \alpha$ 和微分中值定理可知

$$|f(x_1) - f(x_2)| = |f'(\xi)| \cdot |x_1 - x_2| \leq \alpha |x_1 - x_2| \leq \alpha |y_1 - y_2|.$$

因此

$$\begin{aligned} |F(x_1, y_1) - F(x_2, y_2)| &\geq |y_1 - y_2| - |f(x_1) - f(x_2)| \geq (1 - \alpha) |y_1 - y_2| \\ &\geq \frac{1 - \alpha}{\sqrt{2}} \sqrt{(y_1 - y_2)^2 + (x_1 - x_2)^2}. \end{aligned}$$

因此 F 满足上一题的条件, 因此 F 是 C^1 同胚, 从而是双射且是开映射.

• 最后证明 F^{-1} 存在且 F^{-1} 是 C^∞ 的, 因为 dF 作为矩阵函数每个分量均是 C^∞ 的, 而 $dF^{-1}|_p = (d|_{F^{-1}(p)})^{-1}$ 作为逆矩阵也是 C^∞ 的, 所以 F^{-1} 是 C^∞ 的, 从而 F 是微分同胚. 

Problem 89

$f, g \in C^1(\mathbb{R})$ 是实值函数, 假设对任意的 $x, y \in \mathbb{R}$, 都有 $f'(x) \neq g'(y)$. 我们定义

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x + y, f(x) + g(y)).$$

证明 $F(\mathbb{R}^2)$ 是开集且 $F: \mathbb{R}^2 \rightarrow F(\mathbb{R}^2)$ 是 C^1 微分同胚.

Proof. 计算偏导数得

$$\frac{\partial F}{\partial x}|_{(x,y)} = (1, f'(x)), \quad \frac{\partial F}{\partial y}|_{(x,y)} = (1, g'(y)).$$

所以 $\det dF|_{(x,y)} = g'(y) - f'(x)$, 所以 dF 处处可逆, 因此对任意 $p \in \mathbb{R}^2$, 存在 p 的开邻域 U_p 和 $F(p)$ 的开邻域 $V_{F(p)}$ 使得 $F: U_p \rightarrow V_{F(p)}$ 是双射. 从而

$$F(p) \in V_{F(p)} \subset F(\mathbb{R}^2) \Rightarrow F(\mathbb{R}^2) = \bigcup_p V_{F(p)}$$

是开集. 接下来证明 F 在全局上是单射, 这是因为如果有 $(x_1, y_1) \neq (x_2, y_2)$ 使得 $F(x_1, y_1) = F(x_2, y_2)$, 即

$$x_1 + y_1 = x_2 + y_2, \quad f(x_1) + g(y_1) = f(x_2) + g(y_2).$$

则 $x_1 \neq x_2, y_1 \neq y_2$, 由 $f, g \in C^1(\mathbb{R})$, 利用微分中值定理得

$$(x_1 - x_2)f'(\xi) = (y_2 - y_1)g'(\zeta) \rightarrow f'(\xi) = g'(\zeta).$$

这与条件矛盾, 所以 F 是单射, 因此可在 $F(\mathbb{R}^2)$ 上讨论 F^{-1} , 由于 F^{-1} 在每个局部上是 C^1 的, 因此整个 F^{-1} 是 C^1 的, 所以 F 是 \mathbb{R}^2 到其像集的微分同胚. \square

Problem 90

对 $A \in \mathbb{R}^{n \times n}$, 定义

$$\|A\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|Ax|}{|x|}.$$

其中 $|\cdot|$ 是欧氏空间上的标准距离, 证明这是一个范数并且和范数

$$\|A\|_1 = \sum_{i,j=1}^n |a_{ij}|$$

等价.

Proof. 首先证明 $\|\cdot\|$ 是范数: 显然 $\|A\| \geq 0$, 当 $A \neq 0$ 时存在 x 使得 $Ax \neq 0$, 从而 $\|A\| > 0$. 对 $A, B \in \mathbb{R}^{n \times n}$, 有

$$\|\lambda A\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|\lambda Ax|}{|x|} = |\lambda| \cdot \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|Ax|}{|x|} = |\lambda| \|A\|.$$

$$\|A + B\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|Ax + Bx|}{|x|} \leq \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|Ax|}{|x|} + \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|Bx|}{|x|} = \|A\| + \|B\|.$$

再证明范数等价, 设 $x = \sum_i x_i e_i$, 那么就有

$$Ax = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right) e_i.$$

所以

$$\begin{aligned} |Ax| &\leq \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right)^2} \leq \sqrt{\sum_{i=1}^n \left(\|A\|_1 \cdot \sum_{j=1}^n |x_j| \right)^2} \\ &\leq \sqrt{n} \|A\|_1 \cdot \sum_{j=1}^n |x_j| \leq n\sqrt{n} \|A\|_1 \cdot |x|. \end{aligned}$$

所以 $\|A\| \leq n\sqrt{n} \|A\|_1$. 另一方面, 分别取 $x = e_i$ 可得

$$\|A\| \geq \max_{1 \leq i \leq n} \frac{|Ae_i|}{|e_i|} = \max_{1 \leq i \leq n} \sqrt{\sum_{j=1}^n a_{ij}^2} \geq \max_{1 \leq i \leq n} \max_{1 \leq j \leq n} |a_{ij}| \geq \frac{1}{n^2} \|A\|_1.$$

所以两个范数等价.



Problem 91

证明 $GL_n(\mathbb{R})$ 是 $\mathbb{R}^{n \times n}$ 中的开集, 并且映射

$$\text{inv} : GL_n(\mathbb{R}) \rightarrow GL_n(\mathbb{R}), \quad A \mapsto A^{-1}.$$

是 C^1 的映射. 进一步地, 对于 $X \in \mathbb{R}^{n \times n}$,

$$d \text{inv}|_A(X) = -A^{-1}XA^{-1}.$$

Proof. 由于 $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$, $A \mapsto \det A$ 是连续函数, 因此 $\det^{-1}(\mathbb{R} \setminus \{0\}) = GL_n(\mathbb{R})$ 是开集.

计算 $d \text{inv}|_A$ 可知

$$\begin{aligned} d \text{inv}|_A(X) &= \lim_{\lambda \rightarrow 0} \frac{(A + \lambda X)^{-1} - A^{-1}}{\lambda} = \lim_{\lambda \rightarrow 0} A^{-1} \frac{(I + \lambda A^{-1}X)^{-1} - I}{\lambda} \\ &= \lim_{\lambda \rightarrow 0} A^{-1} \frac{\sum_{i=0}^{\infty} (-1)^i \lambda^i (A^{-1}X)^i - I}{\lambda} \\ &= \lim_{\lambda \rightarrow 0} A^{-1} (-XA^{-1} + \sum_{i=2}^{\infty} \lambda^{i-1} (-1)^i (A^{-1}X)^i) = -A^{-1}XA^{-1}. \end{aligned}$$

因此 $d \text{inv}|_A$ 的确是关于 X 线性映射, 并且关于 A 连续. 所以 inv 是 C^1 的.



Problem 92


定义平方运算 $\Theta : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$, $A \mapsto A^2$. 证明 Θ 连续可微并计算其微分 $d\Theta|_A$. 据此证明存在 $\varepsilon > 0$, 使得当 $\|B - I\| < \varepsilon$ 时存在 $A \in \mathbb{R}^{n \times n}$ 使得 $A^2 = B$.

其中 $\|\cdot\|$ 为第 8 题中的范数.

Proof. 计算 $d\Theta|_A$ 可知

$$d\Theta|_A(X) = \lim_{\lambda \rightarrow 0} \frac{(A + \lambda X)^2 - A^2}{\lambda} = \lim_{\lambda \rightarrow 0} (AX + XA + \lambda X^2) = AX + XA.$$

所以 $d\Theta|_A$ 是关于 X 的线性映射并且关于 A 是连续的, 所以 Θ 是 C^1 的.

令 $A = I$, 则 $d\Theta|_I = 2\text{id}$ 是可逆的. 由反函数定理, 存在 I 的开邻域 U, V 使得 Θ 是 $U \rightarrow V$ 的双射. 特别地, V 中的每个点均存在原像. 这也就是说, 存在 $\varepsilon > 0$ 使得当 $\|B - I\| < \varepsilon$ 时, 存在 A 使得 $A^2 = B$. 

Problem 93

假设 $n = 2$, $U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. 是否存在 I 的邻域 $\mathcal{J} \subset \mathbb{R}^{2 \times 2}$, U 的邻域 $\mathcal{U} \subset \mathbb{R}^{2 \times 2}$ 以及连续可微的映射 $\Psi: \mathcal{J} \rightarrow \mathcal{U}$ 使得 $\Psi(I) = U$ 且对任意的 $A \in \mathcal{J}$ 均有 $\Psi(A)^2 = A$?

Proof. 假设满足要求的 Ψ 存在, 对条件式两边微分得


$$d\Psi^2|_A = d\text{id}|_A \Rightarrow d\Theta|_{\Psi(A)} \circ d\Psi|_A = \text{id}.$$

令 $A = I$, 两边代入 X :

$$d\Theta|_U(d\Psi_I(X)) = X \Rightarrow U(d\Psi_I(X)) + (d\Psi_I(X))U = X.$$

设 $d\Psi_I(X) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 那么上式可化为

$$\begin{pmatrix} -a & -b \\ c & d \end{pmatrix} + \begin{pmatrix} -a & b \\ -c & d \end{pmatrix} = \begin{pmatrix} -2a & 0 \\ 0 & 2d \end{pmatrix} = X.$$

但是如果取 X 为非对角阵, 这就不可能成立, 因此满足条件的 Ψ 不存在. 

Problem 94

我们定义立方运算

$$\mathcal{C}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, \quad A \mapsto A^3.$$

证明 \mathcal{C} 是连续可微的, 并计算它的微分 $d\mathcal{C}|_A$. 进一步证明

$$\|d\mathcal{C}|_A - 3\text{id}_{\mathbb{R}^{n \times n}}\| \leq 6\|A - \text{id}_{\mathbb{R}^n}\| + 3\|A - \text{id}_{\mathbb{R}^n}\|^2.$$

其中 $\|\cdot\|$ 是第 8 题中内积, 它可以看作是对线性映射所定义的.

Proof. 计算微分可得

$$\begin{aligned} d\mathcal{C}|_A(X) &= \lim_{\lambda \rightarrow 0} \frac{(A + \lambda X)^3 - A^3}{\lambda} \\ &= \lim_{\lambda \rightarrow 0} ((A^2 X + A X A + X A^2) + \lambda \cdot *) \\ &= A^2 X + A X A + X A^2. \end{aligned}$$

因此

$$\text{LHS} = \sup_{X \in \mathbb{R}^{n \times n}, X \neq 0} \frac{|A^2 X + A X A + X A^2 - 3X|}{|X|}.$$

这里的 $|X| = \sqrt{\sum_{i,j} x_{ij}^2}$. 令 $B = A - I$, 则

$$\begin{aligned} &A^2 X + A X A + X A^2 - 3X \\ &= (B^2 + 2B + I)X + (B + I)X(B + I) + X(B^2 + 2B + I) - 3X \\ &= B^2 X + B X B + X B^2 + 3B X + 3X B. \end{aligned}$$

先考虑最简单的情形, 对任意 $C \in \mathbb{R}^{n \times n}$, 如果设 $X = \begin{pmatrix} X_1 & X_2 & \dots & X_n \end{pmatrix}$, 那么 $CX = \begin{pmatrix} CX_1 & CX_2 & \dots & CX_n \end{pmatrix}$. 所以

$$\frac{|CX|}{|X|} = \frac{\sqrt{\sum_{i=1}^n |CX_i|^2}}{\sqrt{\sum_{i=1}^n |X_i|^2}} \leq \max_{1 \leq i \leq n, X_i \neq 0} \frac{|CX_i|}{|X_i|} \leq \sup_{V \in \mathbb{R}^{n \times 1}, V \neq 0} \frac{|CV|}{|V|} = \|C\|.$$

同理有 $\frac{|XC|}{|X|} \leq \|C\|$, 所以

$$\begin{aligned} \frac{|B^2 X|}{|X|} &= \frac{|B(BX)|}{|BX|} \cdot \frac{|BX|}{|X|} \leq \|B\|^2; \\ \frac{|B X B|}{|X|} &= \frac{|B(XB)|}{|XB|} \cdot \frac{|XB|}{|X|} \leq \|B\|^2; \\ \frac{|X B^2|}{|X|} &= \frac{|(XB)B|}{|XB|} \cdot \frac{|XB|}{|X|} \leq \|B\|^2. \end{aligned}$$

综上,

$$\begin{aligned} \text{LHS} &\leq \sup_{X \in \mathbb{R}^{n \times n}, X \neq 0} \frac{|B^2 X| + |B X B| + |X B^2| + 3|BX| + 3|XB|}{|X|} \\ &\leq 3\|B\|^2 + 6\|B\|. \end{aligned}$$

证毕.



Problem 95

令

$$B_{1/3}(I_n) = \{X \in \mathbb{R}^{n \times n} : \|X - I_n\| < \frac{1}{3}\}.$$

证明: $\mathcal{C}(B_{1/3}(I_n))$ 是 $\mathbb{R}^{n \times n}$ 中的开集并且

$$\mathcal{C} : B_{1/3}(I_n) \rightarrow \mathcal{C}(B_{1/3}(I_n))$$

是微分同胚.

Proof. 对 $X_1, X_2 \in B_{1/3}(I_n)$, 设 $X_i = I_n + Y_i$, 则 $\|Y_i\| < 1/3$. 计算范数得

$$\|X_1^3 - X_2^3\| = \|(I_n + Y_1)^3 - (I_n + Y_2)^3\| = \|3(Y_1 - Y_2) + 3(Y_1^2 - Y_2^2) + (Y_1^3 - Y_2^3)\|.$$

由于

$$\|AB\| = \sup_{x \in \mathbb{R}^{n \times 1}, x \neq 0} \frac{|ABx|}{|x|} \leq \sup_{Bx \in \mathbb{R}^{n \times 1}, Bx \neq 0} \frac{|A(Bx)|}{|Bx|} \cdot \sup_{x \in \mathbb{R}^{n \times 1}, x \neq 0} \frac{|Bx|}{|x|} \leq \|A\| \|B\|,$$


所以

$$\|Y_1^2 - Y_2^2\| \leq \|Y_1 - Y_2\| \|Y_1 + Y_2\| \leq \frac{2}{3} \|Y_1 - Y_2\|;$$

$$\|Y_1^3 - Y_2^3\| \leq \|Y_1 - Y_2\| \|Y_1^2 + Y_1 Y_2 + Y_2^2\| \leq \frac{1}{3} \|Y_1 - Y_2\|.$$

这推出

$$\begin{aligned} \|X_1^3 - X_2^3\| &\geq 3 \|Y_1 - Y_2\| - 3 \|Y_1^2 - Y_2^2\| - \|Y_1^3 - Y_2^3\| \\ &\geq (3 - 2 - \frac{1}{3}) \|Y_1 - Y_2\| = \frac{2}{3} \|X_1 - X_2\|. \end{aligned}$$

从而 \mathcal{C} 在 $B_{1/3}(I_n)$ 上为单射, 因此利用第 5 题的结论 (范数使用 $\|\cdot\|$) 可知 $d\mathcal{C}$ 处处非退化, 从而根据反函数定理 \mathcal{C} 在 $B_{1/3}(I_n)$ 上是开映射, 所以 \mathcal{C} 是 $B_{1/3}(I_n) \rightarrow \mathcal{C}(B_{1/3}(I_n))$ 的微分同胚, 证毕. 

Problem 96

设 $a, b \in \mathbb{R}$. 定义映射

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x + a \sin y, y + b \sin x).$$

证明 f 是 $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ 的微分同胚当且仅当 $ab \in (-1, 1)$.

Proof. 当 $|ab| \geq 1$ 时, 我们证明 f 不是微分同胚. 计算 df :

$$\frac{\partial f}{\partial x}|_{(x,y)} = (1, b \cos x), \quad \frac{\partial f}{\partial y}|_{(x,y)} = (a \cos y, 1).$$

所以

$$\det df|_{(x,y)} = 1 - ab \cos x \cos y.$$

如果 $|ab| \geq 1$, 那么一定存在一组 (x, y) 使得 $\det df|_{(x,y)} = 0$. 但如果 f 为微分同胚, 那么

$$df^{-1}|_{f(x,y)} \circ df|_{(x,y)} = \text{id}$$

所以 df 应当处处非退化, 矛盾.

下证 $|ab| < 1$ 时 f 是微分同胚, 显然 f 是 C^∞ 的, 并且 df 处处非退化.

首先证明 f 是单射, 假设存在 $(x_1, y_1) \neq (x_2, y_2)$ 使得

$$\begin{cases} x_1 + a \sin y_1 = x_2 + a \sin y_2; \\ y_1 + b \sin x_1 = y_2 + b \sin x_2. \end{cases}$$

对 $\sin x$ 使用微分中值定理可得

$$\begin{aligned} |x_1 - x_2| &= |a| |\sin y_1 - \sin y_2| = |a| |\cos \xi| |y_1 - y_2| \leq |a| |y_1 - y_2|; \\ |y_1 - y_2| &\leq |b| |x_1 - x_2|. \end{aligned}$$

两式结合可知 $|ab| \geq 1$, 矛盾.

再证明 f 为满射. 首先注意到 $f(x + 2m\pi, y + 2n\pi) = f(x, y) + 2\pi(m, n)$, 因此首先考虑 f 在闭集 $A = [0, 2\pi] \times [0, 2\pi]$ 上的像.

- $\text{im } f$ 为开集, 这是因为 f 在 \mathbb{R}^2 上为开映射, 这由 df 处处非退化得到.
- 下证 $\text{im } f$ 为闭集, 考虑一列 \mathbb{R}^2 中的收敛子列 $\mathbf{p}_i \rightarrow \mathbf{p}$. 如果 $\mathbf{p}_i \in \text{im } f$, 我们证明 $\mathbf{p} \in \text{im } f$. 设 $f(\mathbf{q}_i) = \mathbf{p}_i$, 则存在 N 使得对任意 $i > N$ 均有 $|\mathbf{p}_i - \mathbf{p}| < 1$, 因此不妨设所有 \mathbf{p}_i 均在 $B_{\mathbf{p}}(1)$ 内, 则 $|f(\mathbf{q}_1) - \mathbf{p}| < 1$.
- 注意到

$$f(x + 2m\pi, y + 2n\pi) = f(x, y) + 2\pi(m, n).$$

而 $A = [0, 2\pi] \times [0, 2\pi]$ 是紧集, 所以 $f(A)$ 也是紧集, 即有界闭集, 因此可找到 $R > 0$ 使得 $f(A) \subset B_R(0)$.

现在将 \mathbb{R}^2 按以 2π 为大小的网格划分, 每一块为 $[2m\pi, (2m+2)\pi) \times [2n\pi, (2n+2)\pi)$. 则对任意 $\mathbf{q} \in \mathbb{R}^2$, 存在唯一的 (m, n) 使得 $\mathbf{q} - 2\pi(m, n)$ 和 \mathbf{q}_1 落在同一块内, 满足

$$\begin{aligned} |\mathbf{q} - 2\pi(m, n) - \mathbf{q}_1| &< 4\pi \\ \Rightarrow 2\pi\sqrt{m^2 + n^2} &= |2\pi(m, n)| > |\mathbf{q} - \mathbf{q}_1| - 4\pi. \end{aligned}$$

则


$$\begin{aligned} |f(\mathbf{q} - 2\pi(m, n)) - f(\mathbf{q}_1)| &< R \\ \Rightarrow |f(\mathbf{q}) - f(\mathbf{q}_1)| &> |2\pi(m, n)| - R = 2\pi\sqrt{m^2 + n^2} - R. \\ \Rightarrow |f(\mathbf{q}) - \mathbf{p}| &> 2\pi\sqrt{m^2 + n^2} - R - 1 > |\mathbf{q} - \mathbf{q}_1| - 4\pi - R - 1. \end{aligned}$$

因此只要 $|\mathbf{q} - \mathbf{q}_1| > 4\pi + R + 2$, 那么就有 $|f(\mathbf{q}) - \mathbf{p}| > 1$, 从而 $\mathbf{q} \notin \{\mathbf{q}_i\}$.

- 因此 $\{q_i\} \subset \overline{B_{4\pi+R+2}(q_1)}$, 而这是一个紧集, 所以一定存在 $\{q_i\}$ 的收敛子列 $q_{i_k} \rightarrow r$. 由连续性可知

$$f(r) = \lim_{k \rightarrow \infty} f(q_{i_k}) = \lim_{k \rightarrow \infty} p_{i_k} = p.$$

从而 $p \in \text{im } f$, 证毕.

所以 $\text{im } f$ 既开又闭并且非空, 故 $\text{im } f = \mathbb{R}^2$, 即 f 为满射. 所以 f^{-1} 是良定义的, 根据反函数定理它在每个局部上都是 C^∞ 的, 从而 f^{-1} 是 C^∞ 的, f 为 $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ 的微分同胚. 

Problem 97

假设 $f(\lambda, x)$ 是 $\mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ 的连续可微映射, 其中 $\lambda \in \mathbb{R}^m$, $x \in \mathbb{R}^n$. 假设存在常数 $k < 1$ 使得对任意 $\lambda \in \mathbb{R}^m$, 映射 $g_\lambda(x) = f(\lambda, x)$ 满足

$$\|dg|_x\| < k.$$


证明对任意 $\lambda \in \mathbb{R}^m$, 存在唯一的 $x_\lambda \in \mathbb{R}^n$ 使得 $f(\lambda, x_\lambda) = x_\lambda$. 特别地, 这给出映射

$$\varphi: \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad \lambda \mapsto x_\lambda.$$

Proof. 如果函数 $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是 C^1 的且处处有 $\|dF|_x\| < k < 1$, 则对任意 $x, y \in \mathbb{R}^n$,

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_0^1 dF|_{tx+(1-t)y}(x-y) dt \right| \\ &\leq \int_0^1 k |x-y| dt = k |x-y|. \end{aligned}$$

由于 \mathbb{R}^n 是完备度量空间, 由压缩映像定理, F 存在唯一的不动点.

那么对任意 $\lambda \in \mathbb{R}^m$, 由于 $f(\lambda, x)$ 是 C^1 的, 因此它的所有偏导数连续, 而 $g_\lambda(x)$ 的偏导数都是 f 的偏导数, 故 g_λ 也是 C^1 的, 通过上面的论证就可知 g_λ 存在唯一的不动点, 就是 x_λ . 

Problem 98

证明上题中的 φ 是 C^1 的, 并计算其微分.

Proof. 考虑映射

$$F: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n, \quad (\lambda, x) \mapsto x - f(\lambda, x).$$

那么 F 是 C^1 的, 并且

$$dF_\lambda|_x = \text{id} - df_\lambda|_x.$$

对任意 $|\mathbf{v}| \neq 0$ 有

$$|\mathrm{d}F_\lambda|_x(\mathbf{v})| = |\mathbf{v} - \mathrm{d}f_\lambda|_x(\mathbf{v})| \geq |\mathbf{v}| - k|\mathbf{v}| > 0.$$

所以 $\mathrm{d}F_\lambda|_x$ 处处非退化. 现在给定 $\lambda_0 \in \mathbb{R}^m$, 由上一题结论, 存在唯一的 $\varphi(\lambda_0)$ 使得

$$F(\lambda_0, \varphi(\lambda_0)) = \varphi(\lambda_0) - \varphi(\lambda_0) = 0.$$

由隐函数定理, 存在 $U \ni \lambda_0$ 和 $V \ni \varphi(\lambda_0)$ 和 C^1 的 $\phi: U \rightarrow V$ 使得

$$f(\lambda, x) = 0 \Leftrightarrow x = \phi(\lambda), \quad \lambda \in U, \phi(\lambda) \in V.$$

由唯一性可知在 U 上有 $\phi = \varphi$, 而 C^1 是局部性质, 从而在 \mathbb{R}^m 上有 φ 是 C^1 的并且

$$\begin{aligned} \mathrm{d}\varphi|_\lambda &= -(\mathrm{d}_x F|_{(\lambda, x)})^{-1} \circ (\mathrm{d}_\lambda F|_{(\lambda, x)}) \\ &= -(\mathrm{id} - \mathrm{d}_x f|_{(\lambda, x)})^{-1} \circ (-\mathrm{d}_\lambda f|_{(\lambda, x)}) \\ &= (\mathrm{id} - \mathrm{d}_x f|_{(\lambda, x)})^{-1} \circ \mathrm{d}_\lambda f|_{(\lambda, x)}. \end{aligned}$$

证毕.



Problem 99

证明对任意 $t \in \mathbb{R}$, 存在唯一的 C^1 映射 $t \mapsto (x(t), y(t)) \in \mathbb{R}^2$, 使得

$$\begin{cases} x = \frac{1}{2} \sin(x+y) + t - 1; \\ y = \frac{1}{2} \cos(x-y) - t + \frac{1}{2}. \end{cases}$$

进一步证明 $(x(t), y(t)) = (0, 0)$ 当且仅当 $t = 1$ 并计算 $x'(1)$ 和 $y'(1)$.

Proof. 考虑

$$F_t: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto \left(\frac{1}{2} \sin(x+y) + t - 1, \frac{1}{2} \cos(x-y) - t + \frac{1}{2} \right).$$

计算其微分:

$$\begin{aligned} \frac{\partial F_t}{\partial \mathbf{x}} &= \left(\frac{1}{2} \cos(x+y), -\frac{1}{2} \sin(x-y) \right); \\ \frac{\partial F_t}{\partial \mathbf{y}} &= \left(\frac{1}{2} \cos(x+y), \frac{1}{2} \sin(x-y) \right). \end{aligned}$$

因此

$$|\det F_t|_{(x, y)}| = \left| \frac{1}{2} \cos(x+y) \sin(x-y) \right| \leq \frac{1}{2} < 1.$$

由第 15 题结论, 存在唯一的一组 (x_t, y_t) 使得 (x_t, y_t) 为 F_t 的不动点. 由于

$$F: \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad (t, x, y) \mapsto F_t(x, y)$$

显然是 C^1 的, 所以 $t \mapsto (x(t), y(t))$ 也是 C^1 的.

下证明 $(x(t), y(t)) = (0, 0)$ 当且仅当 $t = 1$. 把 $(0, 0)$ 代入上式后得到

$$0 = 0 + t - 1, \quad 0 = \frac{1}{2} - t + \frac{1}{2} \Rightarrow t = 1.$$

根据 16 题结论, 可以计算 $\varphi: t \mapsto (x(t), y(t))$ 的微分:

$$\begin{aligned} dF|_{(t,x,y)} &= \begin{pmatrix} 1 & \frac{1}{2} \cos(x+y) & \frac{1}{2} \cos(x+y) \\ -1 & -\frac{1}{2} \sin(x-y) & \frac{1}{2} \sin(x-y) \end{pmatrix}. \\ dF|_{(1,0,0)} &= \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 0 \end{pmatrix}. \\ d\varphi|_1 &= (\text{id} - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{pmatrix})^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

因此 $x'(1) = 1, y'(1) = -1$.



13 第八次作业

Problem 100

设 $M \subset \mathbb{R}^n$ 和 $N \subset \mathbb{R}^m$ 是光滑子流形. 证明 $M \times N \subset \mathbb{R}^{n+m}$ 也是光滑子流形.

Proof. 对任意 $(x, y) \in M \times N$, 取 $x \in U_1 \subset \mathbb{R}^n, y \in U_2 \subset \mathbb{R}^m$ 以及 $V_1 \in \mathbb{R}^n, V_2 \subset \mathbb{R}^m$ 和光滑同胚 $\Phi_i: U_i \rightarrow V_i$ 使得

$$\Phi_1(M \cap U_1) = V_1 \cap (\mathbb{R}^n \times \{0\}), \quad \Phi_2(N \cap U_2) = V_2 \cap (\mathbb{R}^m \times \{0\}).$$



则 $(x, y) \in U_1 \times U_2$, 并且 $(M \times N) \cap (U_1 \times U_2) = (M \cap U_1) \times (N \cap U_2)$, 并且

$$\begin{aligned} \Phi_1 \times \Phi_2: U_1 \times U_2 &\xrightarrow{\sim} V_1 \times V_2; \\ (M \times N) \cap (U_1 \times U_2) &\xrightarrow{\sim} (V_1 \cap (\mathbb{R}^n \times \{0\})) \times (V_2 \cap (\mathbb{R}^m \times \{0\})) \\ &= (V_1 \times V_2) \cap (\mathbb{R}^n \times \{0\} \times \mathbb{R}^m \times \{0\}) \\ &= (V_1 \times V_2) \cap (\mathbb{R}^{n+m} \times \{0\}). \end{aligned}$$

所以 $\Phi_1 \times \Phi_2$ 满足 (x, y) 附近的子流形条件, 因此 $M \times N$ 是光滑子流形.

Problem 101

假设 $\Omega \subset \mathbb{R}^n$ 是开集, f_1, \dots, f_k 是给定的 k 个光滑函数, 其中 $k \leq n$. 令

$$Z_i = \{x \in \Omega : f_i(x) = 0\}.$$


假设 $x \in \bigcap_{i \leq k} Z_i$ 且 $df_1|_x, df_2|_x, \dots, df_k|_x$ 线性无关, 证明: 存在 x 附近的开集 U , 使得对任意的 $l \leq k$, $\bigcap_{i \leq l} Z_i \cap U$ 是维数为 $n - l$ 的子流形.

Proof. 对任意 $l \leq k$, 考虑函数

$$F_l : \Omega \rightarrow \mathbb{R}^l, \quad x \mapsto (f_1(x), \dots, f_l(x)).$$

则

$$\bigcap_{i \leq k} Z_i = F_l^{-1}(0).$$

另一方面, $dF_l|_x$ 作为 $k \times n$ 矩阵, 其行向量分别对应线性映射 $df_1|_x, \dots, df_l|_x$. 从而 $\text{rk } dF_l|_x = l$, 因此由隐函数定理, 存在 x 的邻域 U 使得 $U \cap F_l^{-1}(0)$ 是子流形, 这就证明了结论. 

Problem 102


假设 $f(x, y, z)$ 和 $g(x, y, z)$ 是 \mathbb{R}^3 上的光滑函数, 如果对于任意同时满足 $f(x_0, y_0, z_0) = 0$ 和 $g(x_0, y_0, z_0) = 0$ 的点 (x_0, y_0, z_0) , 如果如下两个向量线性无关:

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \Big|_{(x_0, y_0, z_0)}, \quad \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z} \right) \Big|_{(x_0, y_0, z_0)}$$

证明:

$$\gamma = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = 0, g(x, y, z) = 0\}$$

是光滑曲线.

Proof. 构造 $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2, p \mapsto (f(p), g(p))$, 则 $\gamma = F^{-1}(0)$. 由条件知 $\text{rk } F = 2$ 对任意 $x \in F^{-1}(0)$ 均成立, 所以 $F^{-1}(0)$ 是余 2 维子流形, 即光滑曲线. 

Problem 103

设 $\Omega \subset \mathbb{R}^n$ 为开集, $f : \Omega \rightarrow \mathbb{R}^m$ 是光滑映射. 证明 f 的图像

$$\Gamma_f = \{(x, y) \in \mathbb{R}^{n+m} : x \in \mathbb{R}^n, y \in \mathbb{R}^m, y = f(x)\}$$

是 \mathbb{R}^{n+m} 中的 n 维光滑子流形.

Proof. 考虑 $F: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m, (x, y) \mapsto y - f(x)$. 则

$$F^{-1}(0) = \Gamma_f.$$

只需考虑 dF , 注意到

$$dF|_{(x,y)}(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_2 - df|_x(\mathbf{v}_1).$$

对任意 $\mathbf{v}_1 \in \mathbb{R}^n$, 存在唯一的 $\mathbf{v}_2 \in \mathbb{R}^m$ 使得 $\ker dF|_{(x,y)}(\mathbf{v}_1, \mathbf{v}_2) = 0$, 从而可建立一个从 \mathbb{R}^n 到 $\ker dF|_{(x,y)}$ 的双射, $\dim \ker dF|_{(x,y)} = n$, 所以 $\operatorname{rk} dF|_{(x,y)} = m$ 恒成立, 所以 Γ_f 是余 m 维光滑子流形, 故是 n 维光滑子流形. \square

Problem 104

证明如下两个方程定义出 \mathbb{R}^3 中的一条光滑曲线:

$$\gamma: x^2 + y^2 + z^2 = 14, \quad x^3 + y^3 + z^3 = 36.$$

对于曲线上的每一点, 计算它的切空间.

Proof. 记 $f(x, y, z) = x^2 + y^2 + z^2 - 14$, $g(x, y, z) = x^3 + y^3 + z^3 - 36$. 则

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \Big|_{(x,y,z)} = (2x, 2y, 2z), \quad \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z} \right) \Big|_{(x,y,z)} = (3x^2, 3y^2, 3z^2).$$

则所有 (x, y, z) 处的切向量和上述两者垂直. 首先先证明它们线性无关, 否则存在不全为零的 λ, μ 使得 x, y, z 都是方程

$$3\lambda t^2 + 2\mu t = 0$$

的根. 那么 x, y, z 是一些零和一些相等项, 从而 $at^2 = 14, at^3 = 36, a = 1, 2, 3$. 简单验证地它们都不可能成立, 从而在 γ 上这两个向量始终线性无关, 因此 γ 确实是一条光滑曲线. 它的切空间是

$$\operatorname{span}(yz^2 - zy^2, zx^2 - xz^2, xy^2 - yx^2).$$

\square

Problem 105

在 \mathbb{R}^2 中考虑如下的集合

$$M = \{(x, y) : x^3 + y^3 - 3xy = 1\}.$$

证明 M 是光滑曲线并计算在任何给定点 (x_0, y_0) 处的切空间.

Proof. 计算

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \Big|_{(x_0, y_0)} = (3x_0^2 - 3y_0, 3y_0^2 - 3x_0).$$

我们有对 $(x_0, y_0) \in M$ 均有 $(3x_0^2 - 3y_0, 3y_0^2 - 3x_0) \neq 0$, 否则 $x_0^2 = y_0, y_0^2 = x_0$, 这推出 $(x_0, y_0) = (0, 0)$ 或 $(1, 1)$, 它们都不满足方程, 从而 M 是光滑曲线, M 在 (x_0, y_0) 处切空间为

$$\text{span}(y_0^2 - x_0, y_0 - x_0^2).$$



Problem 106

假设 $M \subset \mathbb{R}^n$ 是子流形且 $\dim M \geq 1$, $p \in M$. 那么存在光滑曲线

$$\gamma : (-1, 1) \rightarrow \mathbb{R}^n$$

使得 $\gamma(0) = p$, $\gamma'(0) \neq 0$ 并且 $\gamma((-1, 1)) \subset M$.

Proof. 取 p 的邻域 $U \cap M$, $V \subset \mathbb{R}^n$ 及光滑同胚 $\Phi : U \rightarrow V$ 使得 $\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$. 可以取 $V \cap (\mathbb{R}^d \times \{0\})$ 中的一条包含 $\Phi(p)$ 的小线段 $\gamma_1 : (-1, 1) \rightarrow V \cap (\mathbb{R}^d \times \{0\})$ 使得 $\gamma_1(0) = \Phi(p)$, 显然 $\gamma_1'(0) \neq 0$. 那么 $\Phi^{-1} \circ \gamma_1$ 就是满足要求的曲线, 这是因为 $d\Phi|_p$ 是双射, 所以 $d\Phi^{-1}(\gamma_1'(0)) \neq 0$ 并且

$$\Phi^{-1}(\gamma_1((-1, 1))) \subset \Phi^{-1}(V \cap (\mathbb{R}^d \times \{0\})) \subset M.$$

证毕.



Problem 107

设 $V = \{(x, y) \in \mathbb{R}^2 : y = |x|\}$ 是 \mathbb{R}^2 中的锥, 证明 V 不是 \mathbb{R}^2 中的光滑子流形.

Proof. 假设 V 是光滑子流形, 利用上题结论可得到 V 上过 $(0, 0)$ 的一条光滑曲线 γ , 使得

$$\gamma(0) = (0, 0), \quad \gamma'(0) \neq (0, 0).$$

然而如果设 $\gamma(t) = (x(t), |x(t)|)$, 由于 $x(t)$ 为单射, 所以在 0 附近的正负性一定改变, 不妨设当 $t < 0$ 时 $x(t) < 0$, 当 $t > 0$ 时 $x(t) > 0$. 那么就有

$$\lim_{t \rightarrow 0^-} \frac{|x(t)| - |x(0)|}{t} \leq 0, \quad \lim_{t \rightarrow 0^+} \frac{|x(t)| - |x(0)|}{t} \geq 0.$$

这导致 $\gamma'(0) = (0, 0)$, 矛盾. 从而 V 不是光滑子流形.



Problem 108

对任意 $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, 定义映射

$$\mathbb{R}^n \rightarrow \mathbb{R}[X], \quad c \mapsto X^n + c_1 X^{n-1} + \dots + c_n =: P_c.$$

给定 $b \in \mathbb{R}^n$, 我们假设 P_b 有一个实数根 x_b 并且 x_b 的重数为 1. 证明存在 b 在 \mathbb{R}^n 中的开邻域 U 和 x_b 在 \mathbb{R} 中的开邻域 V , 使得对任意 $c \in U$, P_c 在 V 中恰有一个根 $z(c)$ 并且重数是 1. 进一步证明映射

$$V \rightarrow \mathbb{R}, \quad c \mapsto z(c)$$

是光滑的.

Proof. 定义映射

$$\Phi: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}, \quad (c, z) \mapsto P_c(z).$$

Φ 是光滑映射, 因为 Φ 只是关于其分量的多项式函数. 计算微分可得

$$\left(\frac{\partial \Phi}{\partial c_1}, \dots, \frac{\partial \Phi}{\partial c_n}, \frac{\partial \Phi}{\partial z} \right) \Big|_{(c, z)} = (z^{n-1}, \dots, z^0, P'_c(z)).$$

考虑 (b, x_b) 使得 $P_b(x_b) = 0$ 且 x_b 重数为 1, 即 $P'_b(x_b) \neq 0$. 这就意味着

$$\Phi(b, x_b) = 0, \quad \frac{\partial \Phi}{\partial z} \Big|_{(b, x_b)} \neq 0.$$

后者即 $\text{rk } d_z \Phi|_{(b, x_b)} = 1$ 恒成立. 那么根据隐函数定理, 存在 $U \ni b$, $V \ni x_b$ 以及 $\Psi: U \rightarrow V$ 使得

$$z \in V, \Phi(c, z) = 0 \Leftrightarrow z = \Psi(c).$$

所以对任意 $c \in U$, 存在唯一的 $z \in V$ 使得 $P_c(z) = 0$, 并且仍然有 $d_z \Phi|_{(c, \Psi(c))}$ 可逆, 所以 $P'_c(\Psi(c)) \neq 0$, 即 $\Psi(c)$ 为 P_c 的根且重数为 1.

根据隐函数定理的结论, 由于 Φ 是光滑的, 所以 Ψ 为 $U \rightarrow V \subset \mathbb{R}$ 的光滑函数.



Problem 109

考虑实系数多项式 $P(X) = X^n + c_1 X^{n-1} + \cdots + c_n = 0$. 如果对 $(c_1^*, \dots, c_n^*) \in \mathbb{R}^n$, $P(X)$ 恰好有 n 个两两不同的实根

$$z_1(c_1^*, \dots, c_n^*) < \cdots < z_n(c_1^*, \dots, c_n^*).$$

证明: 存在 $(c_1^*, \dots, c_n^*) \in \mathbb{R}^n$ 的开邻域 Ω 和 Ω 上的光滑函数 z_1, \dots, z_n , 使得对任意的 $(c_1, \dots, c_n) \in \Omega$, 我们有

$$z_1(c_1, \dots, c_n) < z_2(c_1, \dots, c_n) < \cdots < z_n(c_1, \dots, c_n)$$

并且

$$P(z_i(c_1, \dots, c_n)) = 0.$$

Proof. 记 $c^* = (c_1^*, \dots, c_n^*)$. 对每个 i 对 $z_i(c^*)$ 利用上一题结论可知存在 c^* 的邻域 U_i 和 $z_i(c^*)$ 的邻域 V_i 以及函数 $z_i: U_i \rightarrow V_i$ 使得任意的 $c \in U_i$ 都在 V_i 中存在一根 $z_i(c)$. 适当缩小 V_i 使得 V_i 两两不交 (这总是能做到, 因为 $c^* \in z_i^{-1}(V_i) \cap U_i$ 仍然是开集并且 z_i 在小区域上继承了唯一性) 那么对任意 $c \in \bigcap_i U_i$, $z_i(c)$ 两两不同 (且满足序关系), 所以它们决定了 P_c 的所有根

$$z_1(c) < z_2(c) < \cdots < z_n(c).$$

证毕.



Problem 110

考虑 $n \times n$ 的实矩阵 $A = (A_{ij})_{1 \leq i, j \leq n}$, 假设 A 有 n 个不同的实特征值 $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. 证明: 存在 $\varepsilon > 0$, 使得对任意的 $n \times n$ 的实矩阵 $B = (B_{ij})_{1 \leq i, j \leq n}$, 如果对每对指标 (i, j) , $|A_{ij} - B_{ij}| < \varepsilon$, 那么 B 有 n 个不同的实特征值 $\lambda_1(B) < \lambda_2(B) < \cdots < \lambda_n(B)$.

进一步证明, 如果将每个 λ_i 看做是 B 的系数 B_{ij} 的函数, 这些函数在区域 $|B_{ij} - A_{ij}| < \varepsilon$ (其中 $1 \leq i, j \leq n$) 中是光滑函数.

Proof. 记 $\text{Char } A$ 为 A 的特征多项式, 则 $\text{Char } A = \det(X \cdot I_n - A)$. 考虑函数

$$\Phi: \mathbb{R}^{n \times n} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (A, z) \mapsto (\text{Char } A)(z).$$

那么 Φ 是关于这 $n^2 + 1$ 个变元的多项式, 则 Φ 是光滑函数. 根据题意, 存在 $\lambda_1 < \cdots < \lambda_n$ 使得

$$\Phi(A, \lambda_i) = 0.$$

那么 λ_i 都是一重根, 满足 $(\text{Char } A)'(\lambda_i) \neq 0$. 那么就可以在每个 λ_i 处使用隐函数定

理, 得到 $U_i \ni A$ 以及 $V_i \ni \lambda_i$ 使得 V_i 两两交为空且 $\lambda_i : U_i \rightarrow V_i$ 满足

$$\lambda \in V_i, \Phi(B, \lambda) = 0 \Leftrightarrow \lambda = \lambda_i(B).$$

即 $(\text{Char } B)(\lambda) = 0$. 所以对 $B \in \bigcap_i U_i$, $\lambda_i(B)$ 构成了 $\text{Char } B$ 的 n 个不同根, 即 B 有 n 个不同的实特征值.

最后由于 $\bigcap_i U_i$ 为开集, 所以存在开球

$$B_\varepsilon(A) = \{B : |A_{ij} - B_{ij}| < \varepsilon\} \subset \bigcap_i U_i.$$

对任意 $B \in B_\varepsilon(A)$, B 有 n 个不同实特征值且 λ_i 是区域上的光滑函数, 证毕. \square

Problem 111

证明 $SL_n(\mathbb{R})$ 是 $\text{Mat}_n(\mathbb{R}) = \mathbb{R}^{n \times n}$ 中的光滑子流形, 并计算 $SL_n(\mathbb{R})$ 的维数.

Proof. $SL_n(\mathbb{R}) = \det^{-1}(1)$. 则由于 $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$, 我们只需证明 $d \det$ 在 $SL_n(\mathbb{R})$ 上非零. 如果 $d \det|_A = 0$, 那么 A 在任意 (i, j) 处的主子式为零, 这导致 $\text{adj } A = 0$, 从而 $\det A = 0$. 从而 $d \det$ 在 $SL_n(\mathbb{R})$ 上非零, 故 $SL_n(\mathbb{R})$ 是余 1 维光滑子流形, 维数是 $n^2 - 1$. \square

Problem 112

证明 $SL_n(\mathbb{R})$ 在 I_n 处的切空间是

$$\mathfrak{sl}_n(\mathbb{R}) = \{a \in \text{Mat}_n(\mathbb{R}) : \text{trace } a = 0\}.$$

Proof. $SL_n(\mathbb{R})$ 在 I_n 处的切空间为 $\ker d \det|_{I_n}$. 由于

$$\begin{aligned} d \det|_{I_n}(X) &= \lim_{\lambda \rightarrow 0} \frac{\det(I_n + \lambda X) - \det(I_n)}{\lambda} \\ &= \lim_{\lambda \rightarrow 0} \frac{(-\lambda)^n (\text{Char } X)(-\lambda^{-1}) - 1}{\lambda}. \end{aligned}$$

如果设

$$\text{Char } X = X^n + a_{n-1}X^{n-1} + \cdots + a_0,$$

则

$$\begin{aligned} (-\lambda)^n (\text{Char } X)(-\lambda^{-1}) &= 1 + a_{n-1}(-\lambda) + \cdots + a_0(-\lambda)^n, \\ d \det|_{I_n}(X) &= \lim_{\lambda \rightarrow 0} \frac{a_{n-1}(-\lambda) + \cdots + a_0(-\lambda)^n}{\lambda} = -a_{n-1}. \end{aligned}$$

而根据特征多项式的性质可知 $a_{n-1} = -\text{trace } A$, 所以 $d \det|_{I_n}(X) = \text{trace } X$, 所以切空间为

$$\ker d \det|_{I_n} = \{X : \text{trace } X = 0\} = \mathfrak{sl}_n(\mathbb{R}).$$

证毕. \square

Problem 113

证明 O_n 是 $\text{Mat}_n(\mathbb{R}) = \mathbb{R}^{n \times n}$ 的光滑子流形并计算其维数.

Proof. 考虑映射 $f: \text{Mat}_n(\mathbb{R}) \rightarrow \text{Sym}_n(\mathbb{R})$, $A \mapsto AA^t - I_n$. 这是因为 $(AA^t - I_n)^t = AA^t - I_n$, 所以 $f(A)$ 一定是对称矩阵. 则 $O_n = f^{-1}(0)$. 下考虑 df :

$$\begin{aligned} df|_A(X) &= \lim_{\lambda \rightarrow 0} \frac{f(A + \lambda X) - f(A)}{\lambda} \\ &= \lim_{\lambda \rightarrow 0} \frac{(A + \lambda X)(A^t + \lambda X^t) - I_n - (AA^t - I_n)}{\lambda} \\ &= XA^t + AX^t. \end{aligned}$$

我们声称

$$\ker df|_A = \{S : S^t + S = 0\} \cdot A.$$

首先如果 $XA^t + AX^t = 0$, 那么 $XA^t + (XA^t)^t = 0$, 所以 XA^t 是反对称矩阵, 设 $XA^t = S$. 由于 $A^t = A^{-1}$, 故 $X = SA$. 另一方面, 如果 $X = SA$, 则

$$XA^t + AX^t = SAA^t + AA^tS^t = S + S^t = 0.$$

由于反对称矩阵构成的线性空间维数为 $\frac{(n-1)n}{2}$, 所以对任意 $A \in O_n$ 均有

$$\text{rk } df|_A = \frac{n(n+1)}{2} = \dim \text{Sym}_n(\mathbb{R}).$$

由隐函数定理, O_n 是余维数为 $\frac{n(n+1)}{2}$ 的子流形, 其维数为 $\frac{n(n-1)}{2}$.



Problem 114

证明 $O_n(\mathbb{R})$ 在 I_n 处的切空间为

$$\mathfrak{o}_n = \{a \in \text{Mat}_n(\mathbb{R}) : a^t + a = 0\}.$$

Proof. 在上一题中令 $A = I_n$ 即得

$$\ker df|_{I_n} = \{a \in \text{Mat}_n(\mathbb{R}) : a^t + a = 0\}.$$

这就是 O_n 在 I_n 处的切空间.



Problem 115

令 $G = SL_n$ 或 O_n . 证明映射

$$\varphi: G \rightarrow G, \quad g \mapsto g^{-1}$$

是光滑的.

Proof. 首先证明 φ 是封闭的. 若 $g \in SL_n$, 则


$$\det(g) = 1 \Rightarrow \det(g^{-1}) = (\det g)^{-1} = 1 \Rightarrow g^{-1} \in SL_n.$$

若 $g \in O_n$, 则

$$(g^{-1})^t(g^{-1}) = [(g^t)g]^{-1} = I_n \Rightarrow g^{-1} \in O_n.$$

另一方面, 注意到对 $A \in G$, φ 不改变行列式的取值, 所以

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A = \frac{1}{c} \operatorname{adj} A$$

其中 $c = \pm 1$. 右边矩阵的每个系数都是关于 A 的矩阵元的多项式, 所以 φ 在 A 附近是光滑的, 从而 φ 在 G 上光滑. 

Problem 116

令 $G = SL_n$ (或 O_n), $\mathfrak{g} = \mathfrak{sl}_n$ (或 \mathfrak{o}_n), 证明指数映射 \exp 把 \mathfrak{g} 打到 G .

Proof. 取函数 φ 使得 $G = \varphi^{-1}(c)$. 对 $A \in \mathfrak{g}_n$, 考察 $f(t) = e^{tA}$, 则

$$f(0) = I_n, f'(0) = A.$$

因此

$$(\varphi \circ f)'(0) = d\varphi|_{f(0)}(f'(0)) = d\varphi|_{I_n}(A) = 0.$$

这是因为 A 落在 I_n 的切空间 \mathfrak{g} 中, 所以也落在 $\ker d\varphi|_{I_n}$ 中. 再考虑一般的 t , 则


$$\begin{aligned} f(t) &= e^{tA}, f'(t) = Ae^{tA}. \\ (\varphi \circ f)'(t) &= d\varphi|_{f(t)}(f'(t)) = d\varphi|_{e^{tA}}(Ae^{tA}) \\ &= \lim_{\lambda \rightarrow 0} \frac{\varphi(e^{tA} + \lambda \cdot Ae^{tA}) - \varphi(e^{tA})}{\lambda}. \end{aligned}$$

当 $\varphi = \det$ 时, $\varphi(AB) = \varphi(A)\varphi(B)$. 从而

$$(\varphi \circ f)'(t) = \varphi(e^{tA}) \cdot d\varphi|_{I_n}(A) = 0.$$

当 $\varphi(A) = AA^t$ 时, $\varphi(AB) = A\varphi(B)A^t$, 从而

$$(\varphi \circ f)'(t) = (e^{tA}) \cdot d\varphi|_{I_n}(A) \cdot (e^{tA})^t = 0.$$

因此 $(\varphi \circ f)' = 0$ 恒成立, 所以 $\varphi(f(t)) = \varphi(f(0)) = c$. 特别地, 令 $t = 1$ 可得 $\varphi(e^A) = c$, 从而 $e^A \in \varphi^{-1}(c) = G$. 

14 第九次作业

Problem 117

假设 $M \subset \mathbb{R}^m, N \subset \mathbb{R}^n$ 是子流形, $f: M \rightarrow N$ 是光滑映射. 证明: $df: TM \rightarrow TN$ 是子流形之间的光滑映射.

Proof. 根据微分的定义,

$$df : TM \rightarrow TN, \quad (p, v) \mapsto (f(p), df|_p(v)).$$

对任何点 $p \in M$, 根据 f 的光滑性, 存在 M 上的开集 $U \ni p$, N 上的开集 $V \ni f(p)$ 以及光滑同胚

$$\Phi_1 : U \rightarrow U' \subset \mathbb{R}^{d_1}, \quad \Phi_2 : V \rightarrow V' \subset \mathbb{R}^{d_2}$$

使得 $\Phi_2 \circ f \circ \Phi_1^{-1}$ 是光滑的. 考虑 Φ_1 和 Φ_2 在 p 与 $f(p)$ 附近的微分:

$$\begin{aligned} d\Phi_1 : \bigsqcup_{x \in U} T_x M &\xrightarrow{\sim} \bigsqcup_{x' \in U'} T_{x'} \mathbb{R}^{d_1} \cong U' \times \mathbb{R}^{d_1} \subset \mathbb{R}^{2d_1}; \\ d\Phi_2 : \bigsqcup_{y \in V} T_y N &\xrightarrow{\sim} \bigsqcup_{y' \in V'} T_{y'} \mathbb{R}^{d_2} \cong V' \times \mathbb{R}^{d_2} \subset \mathbb{R}^{2d_2}. \end{aligned}$$

$d\Phi_1$ 和 $d\Phi_2$ 都是对应区域之间的光滑同胚. 那么 $d\Phi_2 \circ df \circ d\Phi_1^{-1}$ 定义了

$$\begin{aligned} (x', v') &\mapsto (\Phi_1^{-1}(x'), d\Phi_1^{-1}|_{x'}(v')) \mapsto (f \circ \Phi_1^{-1}(x'), df|_{\Phi_1^{-1}(x')} \circ d\Phi_1^{-1}|_{x'}(v')) \\ &\mapsto (\Phi_2 \circ f \circ \Phi_1^{-1}(x'), d\Phi_2|_{f \circ \Phi_1^{-1}(x')} \circ df|_{\Phi_1^{-1}(x')} \circ d\Phi_1^{-1}|_{x'}(v')) \\ &\mapsto (\Phi_2 \circ f \circ \Phi_1^{-1}(x'), d(\Phi_2 \circ f \circ \Phi_1^{-1})|_{x'}(v')). \end{aligned}$$

由于 $\Phi_2 \circ f \circ \Phi_1^{-1}$ 是光滑的, 因此作为 $\mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$ 的映射我们已经证明了它在切丛上的映射也是光滑的, 从而 $d\Phi_2 \circ df \circ d\Phi_1^{-1}$ 是光滑映射, 所以 df 是 $\bigsqcup_{x \in U} T_x M \rightarrow \bigsqcup_{y \in V} T_y N$ 的光滑映射, 根据 p 的任意性, 这就给出了 $TM \rightarrow TN$ 的光滑映射. $\left\langle \begin{smallmatrix} a \\ i \\ x \end{smallmatrix} \right\rangle$

Problem 118

假设实数 $a_{ij} = a_{ji}$, 求二次型 $f(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$ 在单位球面 \mathbb{S}^{n-1} 上的条件极值.

Proof. 单位球面可以看作函数 $g(x) = x_1^2 + \cdots + x_n^2 - 1$ 的零点集, 即 $\mathbb{S}^{n-1} = g^{-1}(0)$. 如果 f 在 $g^{-1}(0)$ 上点 x 处取到极值, 那么应当有

$$\nabla f|_x \in \text{span}(\nabla g|_x).$$

计算得

$$\begin{aligned} \nabla f|_x &= (2 \sum_{j=1}^n a_{1j} x_j, 2 \sum_{j=1}^n a_{2j} x_j, \dots, \sum_{j=1}^n a_{nj} x_j); \\ \nabla g|_x &= (2x_1, 2x_2, \dots, 2x_n). \end{aligned}$$

设

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix};$$

那么条件转化为

$$AX \in \text{span}(X).$$

即 X 是矩阵 A 的特征向量.



Problem 119

求椭球

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

的最大体积的内接长方体.

Proof. 对椭球施加一个仿射变换

$$F : (x, y, z) \mapsto (x/a, y/b, z/c).$$

那么椭球在 F 下的像是球 $C : x^2 + y^2 + z^2 = 1$, 内接长方体在 F 下的像是球的内接平行六面体 (因为仿射变换保平行).

由于任意面截球均是圆, 所以平行六面体的一个面是圆的内接平行四边形. 由于平行四边形的中心也是圆心, 所以两条对角线均是直径, 这导致该平行四边形实际上是矩形, 同理可知六个面都是矩形, 这导致 F 的像一定是长方体, 并且由对面全等可知对面一定落在两个相对于原点对称的平面内, 所以长方体的中心是原点.

现在设 F 的像的三个边长分别为 a, b, c , 那么 $(a/2)^2 + (b/2)^2 + (c/2)^2 = 1$. 记 $f(a, b, c) = a^2 + b^2 + c^2 - 4$ 而 $g(a, b, c) = abc$. 要使 g 取到最值, 则需要 $\nabla_f(a, b, c)$ 和 $\nabla_g(a, b, c)$ 平行. 计算

$$\nabla_f(a, b, c) = (2a, 2b, 2c), \quad g(a, b, c) = (bc, ac, ab).$$

则 $a^2 = b^2 = c^2$, 因此 $|a| = |b| = |c|$, 这推出 $|a| = |b| = |c| = \frac{2}{\sqrt{3}}$. 并且当 $a = b = c = \frac{2}{\sqrt{3}}$ 时取到最大值, 在 $a = b = -c = \frac{2}{\sqrt{3}}$ 时取到最小值. 因此 F 的像的体积最大值为

$$G\left(\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right) = \frac{8}{3\sqrt{3}}.$$

从而原矩体的体积最大值为

$$\frac{8}{3\sqrt{3}}abc.$$



Problem 120

f 是 \mathbb{R} 上的非负连续函数并且如下的反常积分的值为 1:

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

假设 $[a, b]$ 是使得 $\int_a^b f(x) dx = \frac{1}{2}$ 的长度最短的区间, 证明:

$$f(a) = f(b).$$

Proof. 由于 f 非负可积, 因此 $f \cdot \chi_U$ 均可积, 所以可以考虑

$$F(a) = \int_0^a f(x) dx.$$

那么根据 f 连续可知 F 可微并且 $F'(a) = f(a)$. 定义函数

$$H, G: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad H(x, y) = y - x, \quad G(x, y) = F(y) - F(x).$$

则满足要求的 $(a, b) \in G^{-1}(1/2)$, 并且 $H(a, b)$ 在 $G^{-1}(1/2)$ 上取到极值, 从而 $\nabla H|_{(a,b)}$ 和 $\nabla G|_{(a,b)}$ 平行, 这导致

$$(-1, 1) = \lambda \cdot (-F'(a), F'(b)) \Rightarrow f(a) = F'(a) = F'(b) = f(b).$$

这就证明了 $f(a) = f(b)$.



Problem 121

假设 a 是给定的正实数, 考虑 n 元函数

$$f(x_1, \dots, x_n) = \sum_{k=1}^n x_k \log x_k.$$

试求其最小值, 其中我们要求 $x_1 + \dots + x_n = a$, $x_i > 0$.

Proof. 设

$$g(x_1, \dots, x_n) = x_1 + \dots + x_n$$

那么如果 $(x_1, \dots, x_n) \in g^{-1}(a)$ 让 $f(x_1, \dots, x_k)$ 取到极值, 则 $\nabla g|_{(x_1, \dots, x_k)}$ 和 $\nabla f|_{(x_1, \dots, x_k)}$ 平行, 而

$$\nabla f|_{(x_1, \dots, x_k)} = (\log x_1 + 1, \dots, \log x_n + 1)^t;$$

$$\nabla g|_{(x_1, \dots, x_k)} = (1, \dots, 1)^t.$$

从而 $\log x_1 = \dots = \log x_n$, 所以 $x_1 = \dots = x_n = a/n$. 此时 f 取到极值 $a \log(a/n)$, 并且由于这是唯一的极值且 f 不存在最大值 (当 $x_i \rightarrow 0$ 时 $f \rightarrow \infty$), 所以当 f 在此时取到最小值 $a \log(a/n)$.



Problem 122

考虑 \mathbb{R}^3 中的以原点为中心单位球 \mathbb{S}^2 , 其北极为 $(0, 0, 1)$, 南极为 $(0, 0, -1)$. 记 $\Sigma: z = -1$. 对任意的 $p \in \mathbb{S}^2$, $p \neq N$, N 与 p 的连线与 Σ 恰好交于一个点 $\pi(p)$. 这样我们有一一对应

$$\pi: \mathbb{S}^2 - \{N\} \rightarrow \Sigma, \quad p \mapsto \pi(p).$$

这个映射被称作是从北极到南极切平面的**球极投影**. 证明上面的球极投影是光滑的共形映射 (保角映射), 即对于任意 $\mathbf{v}, \mathbf{w} \in T_p \mathbb{S}^2$, 我们有

$$\frac{\langle d\pi|_p(\mathbf{v}), d\pi|_p(\mathbf{w}) \rangle}{|d\pi|_p(\mathbf{v})| |d\pi|_p(\mathbf{w})|} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{v}| |\mathbf{w}|}.$$

这里的内积指 \mathbb{R}^3 中的标准内积.

Proof. 考虑坐标变换 $(x, y, z) \mapsto (x, y, z - 1)$, N 成为新的坐标原点, $\Sigma: z = -2$. 注意到这不改变需证式. 记得到的新的球面为 C .

对 $p = (x, y, z) \in C \setminus \{N\}$, 根据 $\pi(p) = \lambda p$ 并且 $\pi(p) \in \Sigma$ 可知

$$\pi(p) = \frac{-2}{z} \cdot (x, y, z).$$

又因为 $x^2 + y^2 + (z + 1)^2 = 1$, 所以 $-2z = x^2 + y^2 + z^2 = |p|^2$, 从而上式可化简为

$$\pi(p) = \frac{4p}{|p|^2} = \frac{4p}{\langle p, p \rangle}.$$

现在可以计算微分:

$$d\pi|_p(\mathbf{v}) = \frac{4\mathbf{v}}{\langle p, p \rangle} - \frac{8 \langle p, \mathbf{v} \rangle p}{\langle p, p \rangle^2}.$$

因此计算内积可知

$$\begin{aligned} \langle d\pi|_p(\mathbf{v}), d\pi|_p(\mathbf{w}) \rangle &= \frac{16 \langle \mathbf{v}, \mathbf{w} \rangle}{\langle p, p \rangle^2} - \frac{64 \langle p, \mathbf{v} \rangle \langle p, \mathbf{w} \rangle}{\langle p, p \rangle^3} + \frac{64 \langle p, \mathbf{w} \rangle \langle p, \mathbf{v} \rangle \langle p, p \rangle}{\langle p, p \rangle^4} \\ &= \frac{16 \langle \mathbf{v}, \mathbf{w} \rangle}{\langle p, p \rangle^2}. \end{aligned}$$

所以左边式子可写为

$$\frac{\langle d\pi|_p(\mathbf{v}), d\pi|_p(\mathbf{w}) \rangle}{|d\pi|_p(\mathbf{v})| |d\pi|_p(\mathbf{w})|} = \frac{16 \langle \mathbf{v}, \mathbf{w} \rangle}{\langle p, p \rangle^2} \cdot \frac{\langle p, p \rangle \cdot \langle p, p \rangle}{4|\mathbf{v}| \cdot 4|\mathbf{w}|} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{v}| |\mathbf{w}|}.$$

证毕.



15 第十次作业

Problem 123

设 $\Omega \subset \mathbb{R}^n$ 为凸集, f 和 g 为 Ω 上的凸函数, 试证明:

$$\max(f, g) : \Omega \rightarrow \mathbb{R}, \quad x \mapsto \max(f(x), g(x))$$

也是凸函数.


Proof. 对任意 $x, y \in \Omega$, 根据 f, g 的凸性可知

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \quad g(tx + (1-t)y) \leq tg(x) + (1-t)g(y).$$

从而

$$t \max(f, g)(x) + (1-t) \max(f, g)(y) \geq tf(x) + (1-t)f(y) \geq f(tx + (1-t)y);$$

$$t \max(f, g)(x) + (1-t) \max(f, g)(y) \geq tg(x) + (1-t)g(y) \geq g(tx + (1-t)y).$$

所以 $t \max(f, g)(x) + (1-t) \max(f, g)(y) \geq \max(f, g)(tx + (1-t)y)$, 所以 $\max(f, g)$ 是凸函数. 

Problem 124

在 \mathbb{R}^3 中考虑方程

$$z^3 + 2z + e^{z-x-y^2} = \cos(x-y+z) \quad (*)$$

所定义的集合 Σ . 证明: 存在开集 $U \subset \mathbb{R}^2$ 使得 $0 \in U$ 以及函数 $\varphi : U \rightarrow \mathbb{R}$ 使得 $\Sigma \cap \{(x, y, z) : (x, y) \in U\}$ 是函数 $\varphi(x, y)$ 的图像 $\{(x, y, z) : z = \varphi(x, y), (x, y) \in U\}$. 进一步计算 $\Sigma \cap \{(x, y, z) : (x, y) \in U\}$ 在 $(0, 0, \varphi(0, 0))$ 处的切空间并计算 φ 的 Hesse 矩阵.

Proof. 记

$$\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \Phi(x, y, z) = z^3 + 2z + e^{z-x-y^2} - \cos(x-y+z).$$

则

$$\frac{\partial \Phi}{\partial \mathbf{z}} = 3z^2 + 2 + e^{z-x-y^2} + \sin(x-y+z).$$

在 (*) 中令 $x = y = 0$ 可知 $z^3 + 2z + e^z = \cos z$. 对函数 $f(z) = z^3 + 2z + e^z - \cos z$ 求导可知 $f'(z) = 2 + 3z^2 + e^z + \sin z > 0$ 恒成立并且 $f(0) = 0$, 从而其唯一解为 $z = 0$.

因而在上式中令 $x = y = z = 0$ 可知 $\frac{\partial \Phi}{\partial \mathbf{z}}|_{(0,0,0)} = 3 \neq 0$, 所以存在 $(0, 0)$ 的邻域 $U \subset \mathbb{R}^2$ 以及 $\phi : U \rightarrow \mathbb{R}$ 使得 $(x, y, z) \in \Phi^{-1}(0) \Leftrightarrow z = \phi(x, y)$.

接下来将 z 视为 $\phi(x, y)$, 对 (*) 两边分别关于 x, y 求偏导可知

$$\begin{aligned} 3\phi^2 \cdot \frac{\partial \phi}{\partial x} + 2 \frac{\partial \phi}{\partial x} + e^{\phi-x-y^2} \cdot \left(\frac{\partial \phi}{\partial x} - 1\right) &= -\sin(x-y+\phi) \cdot \left(1 + \frac{\partial \phi}{\partial x}\right); \\ 3\phi^2 \cdot \frac{\partial \phi}{\partial y} + 2 \frac{\partial \phi}{\partial y} + e^{\phi-x-y^2} \cdot \left(\frac{\partial \phi}{\partial y} - 2y\right) &= -\sin(x-y+\phi) \cdot \left(-1 + \frac{\partial \phi}{\partial y}\right). \end{aligned}$$

在上面两式中令 $x = y = 0$ 可知 $\phi = 0$, 所以

$$2 \frac{\partial z}{\partial x}|_{(0,0)} + \frac{\partial z}{\partial x}|_{(0,0)} - 1 = 0, \quad 2 \frac{\partial z}{\partial y}|_{(0,0)} + \frac{\partial z}{\partial y}|_{(0,0)} = 0.$$

从而 $\frac{\partial z}{\partial x}|_{(0,0)} = \frac{1}{3}, \frac{\partial z}{\partial y}|_{(0,0)} = 0$. 所以 $(x, y) \mapsto (x, y, \varphi(x, y))$ 在 $(0, 0)$ 处切空间的一组基为 $(1, 0, 1/3)$ 和 $(0, 1, 0)$, 张成切空间为

$$\text{span}\{(3, 0, 1), (0, 1, 0)\}.$$

最后计算 Hesse 矩阵: 对上式再对 x, y 求一次偏导并直接代入 $x = y = \phi = 0$:

$$\begin{aligned} 2 \frac{\partial^2 \phi}{\partial x \partial x} + \frac{\partial^2 \phi}{\partial x \partial x} + \left(\frac{\partial \phi}{\partial x} - 1\right)^2 &= -(1 + \frac{\partial \phi}{\partial x})^2. \\ 2 \frac{\partial^2 \phi}{\partial x \partial y} + \frac{\partial^2 \phi}{\partial x \partial y} + \frac{\partial \phi}{\partial y} \cdot \left(\frac{\partial \phi}{\partial x} - 1\right) &= -(1 + \frac{\partial \phi}{\partial x})(-1 + \frac{\partial \phi}{\partial y}). \\ 2 \frac{\partial^2 \phi}{\partial y \partial y} + \left(\frac{\partial^2 \phi}{\partial y \partial y} - 2\right) + \left(\frac{\partial \phi}{\partial y}\right)^2 &= -(-1 + \frac{\partial \phi}{\partial y})^2. \end{aligned}$$

解得

$$\frac{\partial^2 \phi}{\partial x \partial x} = -\frac{20}{27}, \quad \frac{\partial^2 \phi}{\partial x \partial y} = \frac{4}{9}, \quad \frac{\partial^2 \phi}{\partial y \partial y} = \frac{1}{3}.$$



Problem 125: Hadamard 引理

设 $\Omega \subset \mathbb{R}^n$ 为凸开集, $0 \in \Omega$, 函数 $f \in C^\infty(\Omega)$ 且 $f(0) = 0$, 证明存在光滑向量值函数 $g: \Omega \rightarrow \mathbb{R}^n$, 使得对任意 $x \in \Omega$, 都有

$$f(x) = \langle x, g(x) \rangle, \quad g(0) = \nabla f(0).$$

Proof. 由 Ω 是凸集可知对任意 $x \in \Omega$ 以及 $t \in [0, 1]$ 都有 $tx \in \Omega$. 由 Newton-Leibniz 公式:

$$f(x) = f(x) - f(0) = \int_0^1 df|_{tx}(x) dt = \int_0^1 \langle \nabla f(tx), x \rangle dt = \left\langle \int_0^1 \nabla f(tx) dt, x \right\rangle.$$

从而取

$$g(x) = \int_0^1 \nabla f(tx) dt.$$

此时由控制收敛定理,

$$\lim_{x \rightarrow 0} g(x) = \int_0^1 \nabla f(0) ds = \nabla f(0).$$

计算 g 的微分:

$$dg|_x(v) = \lim_{\lambda \rightarrow 0} \int_0^1 \frac{\nabla f(t(x + \lambda v)) - \nabla f(tx)}{\lambda} dt = \int_0^1 d(\nabla f)|_{tx}(v) dt.$$

由于每次求导后得到的积分都是有界的, 所以不断将其求导后可知正规函数光滑.



Problem 126

仍然假设 f 是区域 Ω 上的光滑函数满足 $f(0) = 0$ 且 $df|_0 = 0$ (称为 f 的临界点). 如果 $\nabla^2 f(0)$ 非退化, 则称 0 为 f 的非退化临界点. 证明: 存在 0 的开邻域 $U \subset \Omega$ 以及 U 上 $n \times n$ 光滑对称可逆矩阵函数 $H(x)$, 使得对任意 $x \in U$ 有 $f(x) = x^t H x$.

Proof. 根据带积分余项的 Taylor 公式:

$$f(x) = f(x_0) + df|_{x_0}(x - x_0) + \frac{1}{2} \int_0^1 (1-t) d^2 f|_{(1-t)x_0+tx}(x - x_0, x - x_0) dt.$$

其证明只需要令 $g(t) = f((1-t)x_0 + tx)$ 然后利用一维情形即可. 代入 $x_0 = 0$ 即可知

$$\begin{aligned} f(x) &= \frac{1}{2} \int_0^1 (1-t) d^2 f|_{tx}(x, x) dt = \frac{1}{2} \int_0^1 (1-t) x^t (\nabla^2 f(tx)) x dt \\ &= x^t \left(\int_0^1 \frac{1-t}{2} (\nabla^2 f(tx)) dt \right) x =: x^t H(x) x. \end{aligned}$$

计算 $H(0)$: 由控制收敛定理,

$$\lim_{x \rightarrow 0} \int_0^1 \frac{1-t}{2} (\nabla^2 f(tx)) dt = \int_0^1 \frac{1-t}{2} dt \cdot \nabla^2 f(0) = \frac{1}{4} \nabla^2 f(0).$$

所以可以良定义 $H(0) = \frac{1}{4} \nabla^2 f(0)$. 由于每个 $\nabla^2 f(tx)$ 都是对称矩阵, 所以最终得到的 $H(x)$ 也是对称矩阵. 此外, 由于 $H(x)$ 的变化是连续的, 并且 $H(0)$ 是可逆的, 所以存在一个 $H(0)$ 的邻域 V 使得任意 V 内矩阵均可逆, 然后再取 0 的开邻域 $U \subset \Omega$ 使得 $H(U) \subset V$, 这样每个 $H(x)$ 就都是可逆的.



Problem 127

假设 U 是 \mathbb{R}^n 中包含原点的开邻域, $H(x)$ 是定义在 U 上的 $n \times n$ 对称光滑矩阵实函数. 如果 $H(0)$ 非退化并且有 n 个不同的特征值, 证明存在 0 的邻域 $V \subset U$ 使得 $H(x)$ 在 V 上每一点 x 处都存在 n 个不同的特征值 $\lambda_i(x)$ 以及对应的单位特征向量 $v_i(x)$ 使得 λ_i, v_i 都是 V 上的光滑向量值函数.

Proof. 根据第 8 次作业第 11 题结论, 如果一个 $n \times n$ 矩阵 A 有 n 个不同的特征值, 那么存在其 (在矩阵范数下的) 邻域 U 使得任意 $B \in U$ 都有 n 个不同特征值并且 $\lambda_i(B)$ 是关于 B 的光滑函数. 类似地, 我们再证明存在邻域 V 使得对任意 $B \in V$, 均有 n 个不同的单位特征向量 $v_i(B)$ 且 v_i 都是光滑的.

考虑函数

$$\Phi: \mathbb{R}^{n \times n} \times \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}, (A, x_1, \dots, x_n, \lambda) \mapsto \left(\sum_{i=1}^n a_{1i}x_i - \lambda \cdot x_1, \sum_{i=1}^n a_{2i}x_i - \lambda \cdot x_2, \dots, \sum_{i=1}^n a_{ni}x_i - \lambda \cdot x_n, \sum_{i=1}^n x_i^2 - 1 \right).$$

由于 A 有 n 个不同的特征值, 所以存在 n 个 \mathbf{v}_i 使得 $\Phi(A, \mathbf{v}_i, \lambda_i) = 0$. 我们在 $(A, \mathbf{v}_i, \lambda_i)$ 处对后 $n+1$ 个分量求偏导.

$$\begin{aligned} \frac{\partial \Phi}{\partial \mathbf{x}_j} &= (a_{1j}, \dots, a_{jj} - \lambda_i, \dots, a_{nj}, 2x_j); \\ \frac{\partial \Phi}{\partial \lambda} &= (-x_1, \dots, -x_n, 0). \end{aligned}$$

我们证明它们线性无关, 只需证明它们拼成的矩阵

$$\begin{pmatrix} A - \lambda_i I & -\mathbf{v}_i \\ 2\mathbf{v}_i^t & 0 \end{pmatrix}$$

非退化. 否则设存在 (\mathbf{x}, q) 使得

$$\begin{pmatrix} A - \lambda_i I & -\mathbf{v}_i \\ 2\mathbf{v}_i^t & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ q \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

则可得 $(A - \lambda_i I)\mathbf{x} - q\mathbf{v}_i = 0$ 且 $2\mathbf{v}_i^t \mathbf{x} = 0$. 由前一个式子知 $(A - \lambda_i I)\mathbf{x} \in \text{span}(\mathbf{v}_i)$. 但是如果取出所有特征向量 \mathbf{v}_i 可知

$$(A - \lambda_i I) \left(\sum_j p_j \mathbf{v}_j \right) = \sum_j (\lambda_j - \lambda_i) p_j \mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \widehat{\mathbf{v}_i}, \dots, \mathbf{v}_n).$$

从而 $\text{im}(A - \lambda_i I) \subset \text{span}(\mathbf{v}_1, \dots, \widehat{\mathbf{v}_i}, \dots, \mathbf{v}_n)$, 从而 $(A - \lambda_i I)\mathbf{x} = 0$, 因此 $\mathbf{x} \in \text{span}(\mathbf{v}_i)$. 但是 $\mathbf{v}_i^t \mathbf{x} = 0$, 所以只能 $\mathbf{x} = 0$, 这导致也有 $q = 0$. 从而, 矩阵是非退化的.

因此可以利用隐函数定理在每个 $(A, \mathbf{v}_i, \lambda_i)$ 周围找一个开邻域 V_i 使得在 V_i 内的所有矩阵存在特征值 $\lambda_i(B)$ 以及单位特征向量 $\mathbf{v}_i(B)$, 且这两个函数都是光滑的. 适当选取这些 V_i 的交就可以得到 n 个关于特征向量的光滑函数 $\mathbf{v}_i(B)$.

回到原题, 由于 $H(x)$ 是关于 x 的光滑矩阵, 所以可取 $H(0)$ 的邻域 N 使得邻域内的所有矩阵都满足题目要求, 然后再取 0 的邻域 V 使得 $H(x) \in N, \forall x \in V$, 这个 V 就满足题目要求.



Problem 128

假设 f 是 \mathbb{R}^n 中包含原点的区域 Ω 上的光滑函数, 满足 $f(0) = 0$ 以及 $df|_0 = 0$, $d^2f|_0$ 非退化, 即是 0 是 f 的非退化临界点. 证明存在非负整数 $m \leq n$ 以及 0 的开邻域 V 和 $U \subset \Omega$ 以及光滑同胚映射 $\varphi: V \rightarrow U$ 使得

$$\varphi(0) = 0, f(\varphi(y)) = -\sum_{i=1}^m y_i^2 + \sum_{i=m+1}^n y_i^2, \quad \forall y \in V.$$

这里的非负整数 m 称为非退化临界点 0 的指标, 只依赖于 $\nabla^2 f|_0$, 与局部坐标系的选取无关. 特别地在 U 上 0 是唯一的非退化临界点, 也即非退化临界点是孤立的.

Proof. 根据第 4 题结论可知存在 0 的邻域 $U \subset \Omega$ 以及光滑可逆对称矩阵函数 $H: U \rightarrow \mathbb{R}^{n \times n}$ 使得 $f(x) = x^t H(x) x$, $f(0) = \frac{1}{4} \nabla^2 f(0)$. 不妨对坐标系施加一个可逆线性变换使得 $\nabla^2 f(0) \circ T$ 有 n 个不同的特征值, 这不会影响需证的结论. 为了记号方便, 我们把它记为双线性函数形式 $\langle x, x \rangle_{H(x)}$.

考虑 $H(0)$ 的特征向量组成的一组标准正交基. 使用第 5 题结论, 存在 $0 \in U_1 \subset U$ 使得存在光滑函数 $\mathbf{v}_i: U_1 \rightarrow \mathbb{S}^{n-1}$ 和 $\lambda_i: U_1 \rightarrow \mathbb{R}$ 使得固定 x , $\lambda_i(x)$ 两两不同并且 $\mathbf{v}_i(x)$ 是 $H(x)$ 关于特征值 $\lambda_i(x)$ 的单位特征向量. 根据 $H(x)$ 的可逆性, $\lambda_i(x) \neq 0$ 对任意 x 成立, 从而 λ_i 的正负性在整个 U_1 上保持一致, 所以正负性仅依赖于 $H(0)$ 的特征值的正负性, 从而仅依赖于 $\nabla^2 f|_0$ 的特征值正负性, 特征值仅依赖于 $\nabla^2 f|_0$ 所对应的双线性映射 $d^2f|_0$, 从而和坐标选取无关. 因此存在常数 m , 使得 λ_i 中有 m 个恒负, 有 $n-m$ 个恒正, 不妨设前 m 个为负.

由于 $H(x)$ 有 n 个不同的特征值, 从而 $\mathbf{v}_i(x)$ 一定是标准正交的, 它们满足

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle_H = \delta_{ij} \cdot \lambda_i.$$

内积记号 $\langle f, g \rangle_H = h$ 表示在任意 $x \in U$ 上取值后均有 $\langle f(x), g(x) \rangle_{H(x)} = h(x)$. 从而我们可以选取 $\mathbf{v}'_i = |\lambda_i|^{-1/2} \mathbf{v}_i$, 现在 $\langle \mathbf{v}'_i, \mathbf{v}'_i \rangle_H = \pm 1$. 这里前 i 个向量对应的为 -1 , 后 $n-i$ 个向量对应的为 1 .

我们现在设

$$x = \sum_{i=1}^n b_i(x) \cdot \mathbf{v}'_i(x).$$

那么 $b_i(x)$ 是 $U_1 \rightarrow \mathbb{R}$ 的光滑函数, 它们可以拼出函数 $b: U_1 \rightarrow \mathbb{R}^n$, $b = (b_1, \dots, b_n)$. $b_i(x)$ 满足

$$f(x) = \langle x, x \rangle_{H(x)} = -\sum_{i=1}^m b_i(x)^2 + \sum_{i=m+1}^n b_i(x)^2.$$

由于 \mathbf{v}'_i 都是非零正交的, x 和 $b(x)$ 之间的关系可以用一个坐标变换矩阵函数表示:

$$B(x) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix}.$$

考虑 $b_i(x)$ 的微分在 $(0, \dots, 0)$ 处的取值:

$$\frac{\partial b_i}{\partial \mathbf{x}_j} \Big|_{(0, \dots, 0)} = \frac{\partial}{\partial \mathbf{x}_j} \Big|_{(0, \dots, 0)} (b_{i1}(x)x_1 + \dots + b_{in}(x)x_n) = b_{ij}(0).$$

从而 $(\frac{\partial b_i}{\partial \mathbf{x}_j} \Big|_{(0, \dots, 0)})$ 恰好就是 $B(0)$, 所以也是可逆矩阵, 因此微分 $db|_{(0, \dots, 0)}$ 可逆. 从而由反函数定理, 存在 $U_2 \subset U_1$ 使得 b 在 $U_2 \rightarrow \text{im } U_2 := V$ 上是同胚, 所以反过来对 $y \in V$,

$$f \circ b^{-1}(y) = - \sum_{i=1}^m y_i^2 + \sum_{i=m+1}^n y_i^2.$$

最后证明 0 是 U 上的唯一非退化临界点: 对结论式两边求微分

$$df|_{b^{-1}(y)} \circ db^{-1}|_y(\mathbf{y}_i) = \epsilon_i \cdot 2y_i.$$

所以如果 $df|_{b^{-1}(y)} = 0$, 那么 $y_i = 0$ 对任意 i 成立, 从而在 U_2 上只有在 $x = b^{-1}(0) = 0$ 时才有 $df|_x = 0$, 这就代表 0 是 U_2 上唯一的非退化临界点.



16 第十一次作业

Problem 129

计算 \mathbb{R}^3 中曲线 C 的长度, 其中参数曲线 C 由 $t \mapsto (a \cos t, b \sin t, bt)$ 给出, 其中 $t \in [0, 2\pi]$; 计算 $\int_C \frac{z^2}{x^2 + y^2} ds$, 其中 ds 为子流形测度.

Proof. 积不了.



Problem 130

假设 C 是 \mathbb{R}^3 中 $x^2 + y^2 + z^2 = a^2$ 和 $x + y + z = 0$ 所截出的曲线, 试计算 $\int_C xy ds$.

Proof. C 是一大圆, 通过某个把 $z = 0$ 上的圆打到 $x + y + z = 0$ 上的圆的正交变换, 可以设 C 的参数方程为

$$\gamma: [0, 2\pi] \rightarrow C, \quad \theta \mapsto a \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix}$$

将这个矩阵记为 P , 列向量记为 $\mathbf{v}(\theta)$. 那么可以计算积分:

$$\begin{aligned}\int_C xy \, ds &= \int_0^{2\pi} x(\theta)y(\theta) |\gamma'(\theta)| \, d\theta \\ &= a^3 \int_0^{2\pi} \left(\frac{1}{\sqrt{6}} \cos \theta + \frac{1}{\sqrt{2}} \sin \theta \right) \left(\frac{1}{\sqrt{6}} \cos \theta - \frac{1}{\sqrt{2}} \sin \theta \right) |P\mathbf{v}'(\theta)| \, d\theta \\ &= a^3 \int_0^{2\pi} \left(\frac{1}{6} \cos^2 \theta - \frac{1}{2} \sin^2 \theta \right) |\mathbf{v}'(\theta)| \, d\theta \\ &= a^3 \int_0^{2\pi} \left(\frac{\cos 2\theta}{3} - \frac{1}{6} \right) d\theta = -\frac{\pi}{3} a^3.\end{aligned}$$



Problem 131

C 是参数曲线 $(a(t - \sin t), a(1 - \cos t))$, 其中 $t \in [0, 2\pi]$. 计算积分 $\int_C y^2 \, ds$.

Proof.

$$\begin{aligned}\int_C y^2 \, ds &= \int_0^{2\pi} a^2(1 - \cos t)^2 \cdot \sqrt{a^2(1 - \cos t)^2 + a^2 \sin^2 t} \, dt \\ &= \sqrt{2} a^3 \int_0^{2\pi} (1 - \cos t)^{2.5} \, dt \\ &= 8a^3 \int_0^{2\pi} \sin^5(t/2) \, dt = 16a^3 \int_0^{\pi} \sin^5 x \, dx \\ &= 16a^3 \cdot \frac{1}{16} \int_0^{\pi} \sin 5x - 5 \sin 3x + 10 \sin x \, dx \\ &= a^3 \cdot \left(\frac{2}{5} - \frac{10}{3} + 20 \right) = \frac{256}{15} a^3.\end{aligned}$$



Problem 132

设平面上的曲线在极坐标下的方程为 $r = f(\theta)$, 其中 f 是连续可微的, $\theta \in [\alpha, \beta]$. 证明 C 的长度为

$$\int_{\alpha}^{\beta} \sqrt{|f(\theta)|^2 + |f'(\theta)|^2} \, d\theta.$$

Proof. 考虑 $\gamma: [\alpha, \beta] \rightarrow \mathbb{R}^2$, $\theta \mapsto (f(\theta) \cos \theta, f(\theta) \sin \theta)$. 则曲线长度为

$$\begin{aligned}\int_C ds &= \int_{\alpha}^{\beta} |\gamma'(\theta)| \, d\theta \\ &= \int_{\alpha}^{\beta} \sqrt{(f'(\theta) \cos \theta - f(\theta) \sin \theta)^2 + (f'(\theta) \sin \theta + f(\theta) \cos \theta)^2} \, d\theta \\ &= \int_{\alpha}^{\beta} \sqrt{f(\theta)^2 + f'(\theta)^2} \, d\theta.\end{aligned}$$



Problem 133

计算 $x^2 + y^2 + z^2 = a^2$ 与 $x^2 + y^2 = ax$ 相交得到的曲线在球面上围成的面积.

Proof. 相交得到的曲线在 $x-y$ 平面上的投影就是 $x^2 + y^2 = ax$ (实际上应该在上方和下方投影了两份), 设它围成的区域为 U . 因此考虑

$$\Phi: U \rightarrow \mathbb{R}^3, \quad (x, y) \mapsto (x, y, \sqrt{a^2 - x^2 - y^2}).$$

设 $\mathbf{v}_1 = (1, 0), \mathbf{v}_2 = (0, 1) \in T_p(U)$, 则

$$\Phi_* \mathbf{v}_1 = (1, 0, -x(a^2 - x^2 - y^2)^{-1/2}), \quad \Phi_* \mathbf{v}_2 = (0, 1, -y(a^2 - x^2 - y^2)^{-1/2}).$$

从而可以求出

$$\|\Phi_* \mathbf{v}_1 \wedge \Phi_* \mathbf{v}_2\| = \sqrt{1 + (x^2 + y^2)(a^2 - x^2 - y^2)^{-1}} = \sqrt{a^2(a^2 - x^2 - y^2)^{-1}}.$$

所以曲线在球面上围成的面积是

$$\begin{aligned} 2 \int_U \sqrt{a^2(a^2 - x^2 - y^2)^{-1}} \, dx \, dy &= 2a \int_0^a \left(\int_{-\sqrt{ax-x^2}}^{\sqrt{ax-x^2}} \sqrt{a^2(a^2 - x^2 - y^2)^{-1}} \, dy \right) dx \\ &= 2a \int_0^a 2 \arcsin \left(\frac{\sqrt{ax-x^2}}{\sqrt{a^2-x^2}} \right) dx \\ &= 4a \int_0^a \arcsin \frac{\sqrt{x}}{\sqrt{a+x}} \, dx \\ (x = a \tan^2 \theta) &= 4a^2 \int_0^{\frac{\pi}{4}} \theta \cdot (\tan^2 \theta)' \, d\theta \\ &= 4a^2 \left(\frac{\pi}{4} - \int_0^{\frac{\pi}{4}} \tan^2 \theta \, d\theta \right) \\ &= 4a^2 \left(\frac{\pi}{2} - 1 \right) = (2\pi - 4)a^2. \end{aligned}$$



Problem 134

试计算参数曲面 Σ 的面积, 其中 Σ 如下定义:

$$[0, 2\pi] \times [0, 2\pi] \rightarrow \mathbb{R}^3, \quad (\theta, \varphi) \mapsto ((b + a \cos \theta) \cos \varphi, (b + a \cos \theta) \sin \varphi, a \sin \theta).$$

其中我们要求 $0 < a < b$.

Proof. 考虑曲面片 C 关于 θ 和 φ 的偏导:

$$\begin{aligned} \frac{\partial C}{\partial \theta}|_{(\theta, \varphi)} &= (-a \sin \theta \cos \varphi, -a \sin \theta \sin \varphi, a \cos \theta); \\ \frac{\partial C}{\partial \varphi}|_{(\theta, \varphi)} &= (-(b + a \cos \theta) \sin \varphi, (b + a \cos \theta) \cos \varphi, 0). \end{aligned}$$

它们张成的图形对应三个分量分别为 $-a \sin \theta (b + a \cos \theta) (\cos^2 \varphi + \sin^2 \varphi) \mathbf{e}_1 \wedge \mathbf{e}_2$, $-a \cos \theta \sin \varphi (b + a \cos \theta) \mathbf{e}_3 \wedge \mathbf{e}_1$, $a \cos \theta \cos \varphi (b + a \cos \theta) \mathbf{e}_3 \wedge \mathbf{e}_2$.

所以面积系数是

$$a(b + a \cos \theta) \sqrt{\sin^2 \theta + \cos^2 \theta \sin^2 \varphi + \cos^2 \theta \cos^2 \varphi} = a(b + a \cos \theta).$$

计算整个 Σ 的面积:

$$\int_{[0, 2\pi] \times [0, 2\pi]} a(b + a \cos \theta) d\theta d\varphi = 4\pi^2 ab + 2\pi a^2 \int_0^{2\pi} \cos \theta d\theta = 4\pi^2 ab.$$



Problem 135

C 是参数曲线 $(a(t - \sin t), a(1 - \cos t))$, 其中 $t \in [0, 2\pi]$. 将 C 在 \mathbb{R}^3 中绕着 x 轴旋转一周得到的曲面记为 Σ , 求 Σ 的面积.

Proof. 对 Σ 进行参数化:

$$f: [0, 2\pi] \times [0, 2\pi] \rightarrow \Sigma, (t, s) \mapsto (a(t - \sin t), a(1 - \cos t) \sin s, a(1 - \cos t) \cos s).$$

计算偏导数:

$$\begin{aligned} \frac{\partial f}{\partial t}|_{(t,s)} &= a \cdot (1 - \cos t, \sin t \sin s, \sin t \cos s); \\ \frac{\partial f}{\partial s}|_{(t,s)} &= a(1 - \cos t) \cdot (0, \cos s, -\sin s). \end{aligned}$$

从而面积系数是

$$\begin{aligned} & a^2(1 - \cos t) \|(1 - \cos t) \cos s \mathbf{e}_1 \wedge \mathbf{e}_2 - (1 - \cos t) \sin s \mathbf{e}_1 \wedge \mathbf{e}_3 - \sin t \mathbf{e}_2 \wedge \mathbf{e}_3\| \\ &= a^2(1 - \cos t) \sqrt{(1 - \cos t)^2 \cos^2 s + (1 - \cos t)^2 \sin^2 s + \sin^2 t} \\ &= a^2(1 - \cos t) \sqrt{2 - 2 \cos t}. \end{aligned}$$

Ω 的面积为

$$\begin{aligned} & \int_{[0, 2\pi] \times [0, 2\pi]} a^2(1 - \cos t) \sqrt{2 - 2 \cos t} dt ds = 2\sqrt{2}\pi a^2 \int_0^{2\pi} (1 - \cos t)^{3/2} dt \\ &= 2\sqrt{2}\pi a^2 \int_0^{2\pi} 2\sqrt{2} \sin^3 \frac{t}{2} dt = 16\pi a^2 \int_0^\pi \sin^3 x dx \\ &= 16\pi a^2 \cdot \frac{1}{4} (-3 \cos x + \frac{1}{3} \cos 3x)|_0^\pi = \frac{64\pi}{3} a^2. \end{aligned}$$



Problem 136

C 是平面上的抛物线段 $x = y^2$, 其中 $x \in [0, a]$. 将 C 在 \mathbb{R}^3 中绕着 x 轴旋转一周得到的曲面记为 Σ , 试计算 Σ 的面积.

Proof. 对 Σ 进行参数化:

$$f: [-\sqrt{a}, \sqrt{a}] \times [0, \pi] \rightarrow \Sigma, (y, \theta) \mapsto (y^2, y \cos \theta, y \sin \theta).$$

计算偏导数:

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{y}}|_{(y,\theta)} &= (2y, \cos \theta, \sin \theta), \\ \frac{\partial f}{\partial \theta}|_{(y,\theta)} &= (0, -y \sin \theta, y \cos \theta) = y(0, -\sin \theta, \cos \theta).\end{aligned}$$

从而面积系数是

$$\begin{aligned}& \|y\| \|-2y \sin \theta \mathbf{e}_1 \wedge \mathbf{e}_2 + 2y \cos \theta \mathbf{e}_1 \wedge \mathbf{e}_3 + \mathbf{e}_2 \wedge \mathbf{e}_3\| \\ &= |y| \sqrt{4y^2 \sin^2 \theta + 4y^2 \cos^2 \theta + 1} = |y| \sqrt{4y^2 + 1}.\end{aligned}$$

计算 Ω 的面积:

$$\begin{aligned}\int_{[-\sqrt{a}, \sqrt{a}] \times [0, \pi]} |y| \sqrt{4y^2 + 1} \, dy \, d\theta &= \pi \int_{-\sqrt{a}}^{\sqrt{a}} \sqrt{y^2(4y^2 + 1)} \, dy \\ &= 2\pi \int_0^{\sqrt{a}} \sqrt{y^2(4y^2 + 1)} \, dy.\end{aligned}$$

换元 $z = 4y^2 + 1$, 则 $dz = 8y \, dy$, 从而上式等于

$$2\pi \int_1^{4a+1} \frac{1}{8} \sqrt{z} \, dz = \frac{\pi}{4} \cdot \frac{2}{3} [(4a+1)^{3/2} - 1] = \frac{\pi}{6} [(4a+1)^{3/2} - 1].$$



Problem 137

Σ 是锥面 $z^2 = x^2 + y^2$ 被柱面 $x^2 + y^2 = 2ax$ 所截下的部分, 计算曲面积分

$$\int_{\Sigma} x^2 y^2 + y^2 z^2 + z^2 x^2 \, d\sigma.$$

Proof. 锥面被柱面截得的部分在 xy -平面上的投影恰好是 $x^2 + y^2 = 2ax$ 的内部, 并且上下各有一份; 将 z 轴上方的部分记为 U . 那么可以考虑

$$f: U \rightarrow \Sigma^+, \quad (x, y) \mapsto (x, y, \sqrt{x^2 + y^2}).$$

计算偏导数:

$$\frac{\partial f}{\partial \mathbf{x}} = (1, 0, x(x^2 + y^2)^{-1/2}), \quad \frac{\partial f}{\partial \mathbf{y}} = (0, 1, y(x^2 + y^2)^{-1/2}).$$

从而面积系数是

$$\begin{aligned}& \|\mathbf{e}_1 \wedge \mathbf{e}_2 + x(x^2 + y^2)^{-1/2} \mathbf{e}_3 \wedge \mathbf{e}_2 + y(x^2 + y^2)^{-1/2} \mathbf{e}_1 \wedge \mathbf{e}_3\| \\ &= \sqrt{1 + x^2(x^2 + y^2)^{-1} + y^2(x^2 + y^2)^{-1}} = \sqrt{2}.\end{aligned}$$

计算面积积分

$$\int_{\Sigma^+} x^2 y^2 + y^2 z^2 + z^2 x^2 \, d\sigma = \int_U \sqrt{2} \cdot (x^2 y^2 + (x^2 + y^2)^2) \, dx \, dy.$$

变元替换 $z = x - a$:

$$\begin{aligned} & \sqrt{2} \int_{z^2+y^2 \leq a^2} (z^4 + 4az^3 + 6a^2z^2 + 4a^3z + a^4 + y^4 + 3z^2y^2 + 6azy^2 + 3a^2y^2) dz dy \\ &= \sqrt{2} \int_0^a \int_0^{2\pi} r \cdot (r^4(1 + \sin^2 \theta \cos^2 \theta) + r^3 \sin \theta(6a - 2a \sin^2 \theta) \\ & \quad + r^2(3a^2 + 3a^2 \sin^2 \theta) + r4a^3 \sin \theta + a^4) d\theta dr \\ &= 2\sqrt{2}\pi \int_0^a \left(\frac{9}{8}r^5 + \frac{9}{2}a^2r^3 + a^4r\right) dr = 2\sqrt{2}\pi \left(\frac{3}{16} + \frac{9}{8} + \frac{1}{2}\right)a^6 = \frac{29\sqrt{2}\pi}{8}a^6. \end{aligned}$$

所以最终结果为

$$\int_{\Sigma} x^2y^2 + y^2z^2 + z^2x^2 d\sigma = 2 \int_{\Sigma^+} x^2y^2 + y^2z^2 + z^2x^2 d\sigma = \frac{29\sqrt{2}\pi}{4}a^6.$$



Problem 138

Σ 是 \mathbb{R}^3 中的柱面 $x^2 + y^2 = a^2$, 其中 $z \in [0, h]$, 计算曲面积分

$$\int_{\Sigma} x^4 + y^4 d\sigma.$$

Proof. 将柱面进行参数化

$$f : [0, 2\pi] \times [0, h] \rightarrow \Sigma, (\theta, z) \mapsto (a \cos \theta, a \sin \theta, z).$$

计算偏导数

$$\begin{aligned} \frac{\partial f}{\partial \theta}|_{(\theta, z)} &= (-a \sin \theta, a \cos \theta, 0), \\ \frac{\partial f}{\partial z}|_{(\theta, z)} &= (0, 0, 1). \end{aligned}$$

计算面积系数

$$\| -a \sin \theta \mathbf{e}_1 \wedge \mathbf{e}_3 + a \cos \theta \mathbf{e}_2 \wedge \mathbf{e}_3 \| = |a|.$$

计算曲面积分:

$$\begin{aligned} \int_{\Sigma} x^4 + y^4 d\sigma &= \int_{[0, 2\pi] \times [0, h]} |a| a^4 \sin^4 \theta + a^4 \cos^4 \theta d\theta dz \\ &= |a|^5 h \int_0^{2\pi} 1 - 2 \sin^2 \theta \cos^2 \theta d\theta \\ &= |a|^5 h \int_0^{2\pi} \frac{3 + \cos 4\theta}{4} d\theta \\ &= \frac{3}{2} \pi |a|^5 h. \end{aligned}$$



Problem 139

Σ 是 \mathbb{R}^3 中的半球面 $x^2 + y^2 + z^2 = 1$, 其中 $z \geq 0$. 计算曲面积分

$$\int_{\Sigma} x + y + z \, d\sigma.$$

Proof. 考虑半球面在 xy -平面上的投影, 记 $U: x^2 + y^2 \leq 1$. 定义参数化

$$f: U \rightarrow \Sigma, \quad (x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2}).$$

计算偏导数:

$$\frac{\partial f}{\partial x}|_{(x,y)} = (1, 0, -x(1 - x^2 - y^2)^{-1/2}), \quad \frac{\partial f}{\partial y}|_{(x,y)} = (0, 1, -y(1 - x^2 - y^2)^{-1/2}).$$

计算面积系数:

$$\begin{aligned} & \| \mathbf{e}_1 \wedge \mathbf{e}_2 - y(1 - x^2 - y^2)^{-1/2} \mathbf{e}_1 \wedge \mathbf{e}_3 - x(1 - x^2 - y^2)^{-1/2} \mathbf{e}_3 \wedge \mathbf{e}_2 \| \\ &= \sqrt{1 + y^2(1 - x^2 - y^2)^{-1} + x^2(1 - x^2 - y^2)^{-1}} = (1 - x^2 - y^2)^{-1/2}. \end{aligned}$$

计算曲面积分

$$\begin{aligned} \int_{\Sigma} x + y + z \, d\sigma &= \int_U (x + y + \sqrt{1 - x^2 - y^2})(1 - x^2 - y^2)^{-1/2} \, dx \, dy \\ &= \pi + \int_U (r \cos \theta + r \sin \theta)(1 - r^2)^{-1/2} r \, dr \, d\theta \\ &= \pi + \int_0^1 r^2(1 - r^2)^{-1/2} \int_0^{2\pi} (\cos \theta + \sin \theta) \, d\theta \, dr = \pi. \end{aligned}$$



Problem 140

Σ 是 \mathbb{R}^3 中的球面 $x^2 + y^2 + z^2 = 1$, 计算曲面积分

$$\int_{\Sigma} x^2 \, d\sigma.$$

Proof. 考虑球面在 yz -平面上的投影, 根据对称性可以只考虑半球面的投影, 最终结果为其两倍. 由上题结论可知面积系数为 $(1 - y^2 - z^2)^{-1/2}$, 从而曲面积分的结果为

$$\begin{aligned} & 2 \int_U (1 - y^2 - z^2) \cdot (1 - y^2 - z^2)^{-1/2} \, dy \, dz = 2 \int_U \sqrt{1 - r^2} r \, dr \, d\theta \\ &= 4\pi \int_0^1 r \sqrt{1 - r^2} \, dr \stackrel{s=1-r^2}{=} 2\pi \int_0^1 \sqrt{s} \, ds = \frac{4\pi}{3}. \end{aligned}$$



Problem 141

Σ 是 \mathbb{R}^3 中的球面上的小帽子 $x^2 + y^2 + z^2 = 1$, $z \geq h > 0$. 计算曲面积分

$$\int_{\Sigma} \frac{1}{z} \, d\sigma.$$

Proof. 取 U 为圆盘 $x^2 + y^2 \leq 1 - h^2$, 则可以同理定义

$$f: U \rightarrow \Sigma, (x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2}).$$

则同上题可知面积系数为 $(1 - x^2 - y^2)^{-1/2}$, 计算积分:

$$\begin{aligned} \int_{\Sigma} (1 - x^2 - y^2)^{-1/2} \cdot (1 - x^2 - y^2)^{-1/2} dx dy &= \int_U (1 - r^2)^{-1} r dr d\theta \\ &= 2\pi \int_0^{\sqrt{1-h^2}} (1 - r^2)^{-1} r dr = \pi \int_0^{\sqrt{1-h^2}} \frac{1}{1-r} - \frac{1}{1+r} dr \\ &= \pi \cdot (-\log(1 - r^2)) \Big|_0^{\sqrt{1-h^2}} = -2\pi \log h. \end{aligned}$$



Problem 142

Σ 是 \mathbb{R}^3 中的参数曲面: $(u, v) \mapsto (u \cos v, u \sin v, v)$, 其中 $(u, v) \in [0, a] \times [0, 2\pi]$, 计算曲面积分

$$\int_{\Sigma} z d\sigma.$$

Proof. 计算偏导数:

$$\frac{\partial f}{\partial \mathbf{u}} = (\cos v, \sin v, 0), \quad \frac{\partial f}{\partial \mathbf{v}} = (-u \sin v, u \cos v, 1).$$

计算面积系数:

$$\begin{aligned} &\|u\mathbf{e}_1 \wedge \mathbf{e}_2 + \cos v \mathbf{e}_1 \wedge \mathbf{e}_3 + \sin v \mathbf{e}_2 \wedge \mathbf{e}_3\| \\ &= \sqrt{u^2 + 1}. \end{aligned}$$

计算曲面积分:

$$\begin{aligned} \int_{\Sigma} z d\sigma &= \int_{[0, a] \times [0, 2\pi]} v \sqrt{u^2 + 1} du dv \\ &= \int_0^a \sqrt{u^2 + 1} \int_0^{2\pi} v dv du = 2\pi^2 \cdot \int_0^a \sqrt{u^2 + 1} du \\ &= \pi^2 \cdot (\log(\sqrt{a^2 + 1} + a) + a \sqrt{a^2 + 1}). \end{aligned}$$



Problem 143

Σ 是 \mathbb{R}^3 中的球面 $x^2 + y^2 + z^2 = 1$, f 是 \mathbb{R} 上的连续函数, 证明:

$$\int_{\Sigma} f(ax + by + cz) d\sigma = 2\pi \int_{-1}^1 f((a^2 + b^2 + c^2)^{\frac{1}{2}} t) dt.$$

Proof. 设 $\mathbf{e}_1, \mathbf{e}_2$ 为平面 $ax + by + cz = 0$ 上的正交单位向量, \mathbf{e}_3 为平面的单位法向量. 从而可以给出 Σ 的一个参数化:

$$\varphi: [-1, 1] \times [0, 2\pi] \rightarrow \Sigma, (t, \theta) \mapsto t\mathbf{e}_3 + \sqrt{1 - t^2}(\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2).$$

计算偏导数:

$$\begin{aligned}\frac{\partial \varphi}{\partial t} &= \mathbf{e}_3 - t(1-t^2)^{-1/2}(\cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2), \\ \frac{\partial \varphi}{\partial \theta} &= \sqrt{1-t^2}(-\sin \theta \mathbf{e}_1 + \cos \theta \mathbf{e}_2).\end{aligned}$$

计算面积系数:

$$\begin{aligned}& \sqrt{1-t^2} \left\| -t(1-t^2)^{-1/2} \mathbf{e}_1 \wedge \mathbf{e}_2 - \sin \theta \mathbf{e}_3 \wedge \mathbf{e}_1 + \cos \theta \mathbf{e}_3 \wedge \mathbf{e}_2 \right\| \\ &= \sqrt{1-t^2} \cdot \sqrt{t^2(1-t^2)^{-1} + 1} = 1.\end{aligned}$$

由于 $\varphi(t, \theta)$ 落在与 $ax + by + cz = 0$ 平行且距离为 t 的平面上, 这个平面为 $ax + by + cz = t\sqrt{a^2 + b^2 + c^2}$. 从而积分可以化简为

$$\begin{aligned}\int_{\Sigma} f(ax + by + cz) d\sigma &= \int_{-1}^1 \int_0^{2\pi} f(t\sqrt{a^2 + b^2 + c^2}) d\theta dt \\ &= 2\pi \int_{-1}^1 f(t\sqrt{a^2 + b^2 + c^2}) dt.\end{aligned}$$

证毕.



Problem 144

设物体的质量由密度函数 ρ 给出, 那么在区域 V 内的总质量为 $\int_V \rho(x) dx$, V 的质心坐标由下述函数确定:

$$\bar{x}_i = \frac{\int_V x_i \rho(x) dx}{\int_V \rho(x) dx}.$$

求抛物面 $x^2 + y^2 = 2az$ 和球面 $x^2 + y^2 + z^2 = 3a^2$ 围成的均质 ($\rho \equiv 1$) 物体的质心坐标.

Proof. 抛物面和球面的交满足方程 $z^2 + 2az = 3a^2$, 解得 $z = a$ 或 $-3a$. 根据 $2az \geq 0$ 可知 $z = a$, 因而两个曲面交于 $C: x^2 + y^2 = 2a^2$.

两个曲面围成的区域被 C 分为两部分, 其下方由 $x^2 + y^2 = 2az$ 和 C 围成 (记为 Σ_1), 上方由 $x^2 + y^2 + z^2 = 3a^2$ 和 C 围成 (记为 Σ_2), 我们分两部分计算. 记平面 $z = h$ 和物体的交为 D_h .

$$\begin{aligned}\int_{\Sigma_1} \mathbf{x} dx &= \int_0^a \left(\int_{x^2+y^2=2az} \mathbf{x} dx dy \right) dz = \int_0^a 2\pi z (0, 0, z) dz \\ &= 2\pi (0, 0, \frac{1}{3}a^3) = (0, 0, \frac{2}{3}\pi a^4).\end{aligned}$$

$$\int_{\Sigma_2} \mathbf{x} dx = \int_a^{\sqrt{3}a} \pi(3a^2 - z^2)(0, 0, z) dz = \frac{3}{2}a^2 z^2 - \frac{1}{4}z^4 \Big|_a^{\sqrt{3}a} = (0, 0, \pi a^4).$$

另一方面直接计算体积:

$$\int_{\Sigma_1} d\mathbf{x} = \int_0^a 2\pi z dz = \pi a^3, \quad \int_{\Sigma_2} d\mathbf{x} = \int_a^{\sqrt{3}a} \pi(3a^2 - z^2) dz = (2\sqrt{3} - \frac{8}{3})\pi a^3.$$

所以质心坐标的 z 分量为

$$\left(\frac{5}{3}\pi a^4\right)/(2\sqrt{3}-\frac{5}{3})\pi a^3 = \frac{5a}{6\sqrt{3}-5}.$$



17 第十二次作业

Problem 145: Co-Area 公式

设 $\Omega \subset \mathbb{R}^n$ 是开集, $F: \Omega \rightarrow \mathbb{R}$ 是光滑映射. 假设对任意的 $x \in \Omega$, $dF|_x \neq 0$, 记 $\Sigma_t = F^{-1}(t)$ 是余 1 维子流形. 那么, 对任意 U 上的可积函数 f , 我们有

$$\int_{\Omega} f(x) dx = \int_{\mathbb{R}} \left(\int_{\Sigma_t} \frac{f}{|\nabla F|} d\sigma_t \right) dt.$$

Proof. 首先证明特殊的情形: 如果 $V, U \subset \mathbb{R}^n$ 之间存在同胚 $\Phi: V \rightarrow U$, 并且 F 将 $x \in U$ 打到 $\Phi^{-1}(x)$ 的第 n 个分量. 固定点 $x = (x_1, \dots, x_n) \in V$,

$$\begin{aligned} dx &= \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \wedge \frac{\partial \Phi}{\partial x_n} \right| dx_1 \cdots dx_n. \\ d\sigma_{x_n} &= \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \right| dx_1 \cdots dx_{n-1}. \end{aligned}$$

其中 $|e_1 \wedge \cdots \wedge e_k|$ 代表这些向量在 \mathbb{R}^n 中张成的多面体体积. 下证明

$$\left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \right| = \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \wedge \frac{\partial \Phi}{\partial x_n} \right| \cdot |\nabla F|_{\Phi(x)}.$$

根据 $F \circ \Phi = x_n$ 可知 $dF|_{\Phi(x)} \circ d\Phi|_x = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$, 从而

$$\left\langle \nabla F|_{\Phi(x)}, \frac{\partial \Phi}{\partial x_i} \right\rangle = \delta_{in}, \quad i = 1, \dots, n.$$

这意味着 $\nabla F|_{\Phi(x)}$ 与前 $n-1$ 个向量垂直, 因此和法向量 \mathbf{n} 垂直. 此外, 第 n 个向量在 \mathbf{n} 上的投影长度为 $1/|\nabla F|_{\Phi(x)}$. 所以

$$\begin{aligned} \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \wedge \frac{\partial \Phi}{\partial x_n} \right| &= \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \wedge \frac{1}{|\nabla F|_{\Phi(x)}} \mathbf{n} \right| \\ &= \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \right| \cdot \frac{1}{|\nabla F|_{\Phi(x)}}. \end{aligned}$$

从而对任意函数 f ,

$$\begin{aligned}\int_U f(y) dy &= \int_V f(\Phi(x)) \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \wedge \frac{\partial \Phi}{\partial x_n} \right| dx_1 \cdots dx_n \\ &= \int_V \frac{f(\Phi(x))}{|\nabla F|_{\Phi(x)}} \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \right| dx_1 \cdots dx_n \\ (\text{Fubini}) &= \int_{\mathbb{R}} \left(\int_{x_n=t} \frac{f(\Phi(x))}{|\nabla F|_{\Phi(x)}} \left| \frac{\partial \Phi}{\partial x_1} \wedge \cdots \wedge \frac{\partial \Phi}{\partial x_{n-1}} \right| dx_1 \cdots dx_{n-1} \right) dx_n \\ &= \int_{\mathbb{R}} \left(\int_{F^{-1}(t)} \frac{f}{\nabla F} d\sigma_t \right) dt.\end{aligned}$$

对一般情形的证明: 一方面对任意点 $x \in \Omega$, 由 $dF|_x \neq 0$ 及隐函数定理, 存在其开邻域 U_x 使得有一个 $V_x \subset \mathbb{R}^n$ 和同胚 $\Phi_x: V_x \rightarrow U_x$, 使得 F 将 x 打到 $\Phi^{-1}(x)$ 的第 n 个分量. 另一方面, 如果命题对可积函数 f, g 成立那么也对 $f + g$ 成立, 而已经证明命题对所有支集在 U_x 中的函数 f 均成立. 首先根据 Ω 可以写成可数个不交方块的并, 因此只需对方块考虑命题. 由于边界测度为零, 所以可以看它的闭包是一个紧集, 紧集上存在一个 U_x 的开覆盖, 因此可被有限个 U_x 进行子覆盖, 然后把 f 分散到支集为 U_x 的函数之和即可. \square

Problem 146

给定 \mathbb{R}^n 上的光滑函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 考虑 \mathbb{R}^{n+1} 中的区域

$$\Omega = \{(x, x_n) : x \in \mathbb{R}^n, x_n \in \mathbb{R}, x_n \leq f(x)\}.$$

证明边界上的单位外法向量为

$$\left(-\frac{\nabla f}{\sqrt{1 + |\nabla f|^2}}, \frac{1}{\sqrt{1 + |\nabla f|^2}} \right).$$

Proof. 根据 f 的连续性, $\{(x, x_n) : x_n < f(x)\}$ 落在 Ω 的内部, 从而

$$\partial\Omega = \{(x, f(x)) : x \in \mathbb{R}^n\}.$$

考虑 $\Phi: \mathbb{R}^n \rightarrow \Omega, x \mapsto (x, f(x))$, 那么

$$\frac{\partial \Phi}{\partial x_i} = (e_i, \frac{\partial f}{\partial x_i}).$$

其中 e_i 为第 i 个分量上的单位向量. 那么单位外法向量和这些向量都是垂直的, 直接验证:

$$\left\langle e_i, -\frac{\nabla f}{\sqrt{1 + |\nabla f|^2}} \right\rangle + \frac{\partial f}{\partial x_i} \cdot \frac{1}{\sqrt{1 + |\nabla f|^2}} = \left(-\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_i} \right) \frac{1}{\sqrt{1 + |\nabla f|^2}} = 0.$$

容易得到它是单位向量, 最后根据它的第 $n+1$ 方向指向正半轴方向可知它是外法向量. \square

Problem 147

假设 $K \subset U_1 \subset U_2$ 为 \mathbb{R}^n 中的有界集合, K 为紧集, U_1, U_2 为开集. 证明存在非负光滑函数 f 使得 f 在 K 上为 0, 在 $U_2 \setminus U_1$ 上为正.

Proof. 首先看一维情形, 考虑定义在 \mathbb{R} 上的光滑函数

$$f_1(x) = \begin{cases} e^{-1/x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

那么可选取

$$f_2(x) = \frac{f_1(x)}{f_1(x) + f_1(1-x)}$$

f_2 是良定义的, 因为不会同时有 $x \leq 0$ 和 $1-x \leq 0$ 成立. 并且 f_2 在 $(-\infty, 0]$ 上取值为 0 而在 $[1, +\infty)$ 上取值为 1.

最后考虑定义在 \mathbb{R}^n 上的函数 f_3 满足 $f_3 = f_2(2 - |x|)$, 那么 f_3 在 $|x| \leq 1$ 时取值为 1, 在 $1 \leq |x| \leq 2$ 时取值在 $[0, 1]$ 内, 在 $|x| \geq 2$ 时取值为 0.

回到原题, 对任意 $x \in K$ 选取一个以 x 为中心的开圆盘 V_x 使得 $2V_x \subset U_1$, 那么这些 V_x 构成了 K 的一个开覆盖, 从而存在有限子覆盖, 记为 V_1, \dots, V_k . 用 Φ_i 将开圆盘 V_i 同胚到 $B_1(0)$ 上, $2V_i$ 同胚到 $B_2(0)$ 上, 取函数 $g_i = \Phi_i^{-1}(f_3)$, 它在 V_i 上取值为 1, 在 $2V_i$ 之外取值为零. 将这些 g_i 求和之后得到的函数 g 满足非负, 在 K 上取值均至少为 1 而在 U_1 之外取值为零.

最后取 $f = 1 - f_2 \circ g$, 那么 f 在 K 上取值为 0, 在 U_1 之外取值为 1, 满足题目条件, 并且每一步得到的函数均光滑推出 f 光滑, 证毕. \square

Problem 148

假设 ρ 是 \mathbb{R}^3 上非负可积的函数 (如果一个物体的质量函数是 ρ 的话), 它所对应的引力势能 $\Phi(x)$ 和引力场 $G(x)$ 由下式给出:

$$\Phi(x) = - \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|} dy, \quad G(x) = -\nabla \Phi(x)$$

我们假设区域 $V \subset \mathbb{R}^3$ 是开区域, $\rho_0 > 0$ 是常数,

$$\rho(x) = \begin{cases} \rho_0, & x \in V \\ 0, & x \in \mathbb{R}^3 \setminus V \end{cases}$$

试计算如下情形下的 $\Phi(x)$ 和 $G(x)$:

- (1) $V = \{x \in \mathbb{R}^3 : |x|^2 < R^2\}$ 是球体.
- (2) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 < R^2, 0 < x_3 < h\}$ 是圆柱体.
- (3) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_3^2 > \frac{h^2}{R^2}(x_1^2 + x_2^2), 0 < x_3 < h\}$ 是圆锥体.

Proof. (1)

$$\Phi(x) = - \int_V \frac{\rho_0}{|x-y|} dy = -\rho_0 \int_V \frac{1}{|x-y|} dy.$$

设 $|x| = a$, 通过施加一个正交变换容易证明 $\Phi(x) = \Phi(0, 0, a)$. 我们进行球坐标参数化

$$f: [0, R] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0, 2\pi] \rightarrow V, \quad (r, \theta, \phi) \mapsto (r \cos \theta \sin \phi, r \cos \theta \cos \phi, r \sin \theta).$$

计算转移系数:

$$dx \wedge dy \wedge dz = r^2 \cos \theta dr \wedge d\theta \wedge d\phi.$$

所以所求积分化简为

$$\begin{aligned} \Phi(x) &= -\rho_0 \int_0^R \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \frac{r^2 \cos \theta}{\sqrt{r^2 + a^2 - 2ar \sin \theta}} dr d\theta d\phi \\ &= -2\pi \rho_0 \int_0^R \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{r^2 \cos \theta}{\sqrt{r^2 + a^2 - 2ar \sin \theta}} dr d\theta \\ (u = \sin \theta) &= -2\pi \rho_0 \int_0^R \left(\int_{-1}^1 \frac{r^2}{\sqrt{r^2 + a^2 - 2aru}} du \right) dr \\ &= -2\pi \rho_0 \int_0^R r^2 \cdot \frac{-2}{2ar} \sqrt{r^2 + a^2 - 2aru} \Big|_{-1}^1 dr \\ &= -2\pi \rho_0 \int_0^R \frac{r}{a} (|r+a| - |r-a|) dr. \end{aligned}$$

当 $a > R$ 时, 上面积分可化简为

$$\int_0^R r(|r+a| - |r-a|) dr = \int_0^R 2r^2 dr = \frac{2}{3}R^3.$$

当 $a \leq R$ 时, 上面积分可化简为

$$\int_0^R r(|r+a| - |r-a|) dr = \int_0^a 2r^2 dr + \int_a^R 2ar dr = aR^2 - \frac{1}{3}a^3.$$

这得到结果:

$$\Phi(x) = -\frac{4\pi\rho_0}{3|x|} R^3 (|x| > R), \quad \Phi(x) = -2\pi\rho_0(R^2 - \frac{1}{3}|x|^2) (|x| \leq R).$$

计算梯度:

$$\nabla \Phi(x) = \frac{4\pi\rho_0 R^3}{3|x|^3} x (|x| > R), \quad \nabla \Phi(x) = \frac{4\pi\rho_0}{3} x.$$



Problem 149

假设

$$\rho(x) = \begin{cases} f(|x|), & r < |x| < R; \\ 0, & |x| > R \text{ 或 } |x| < r. \end{cases}$$

其中 f 是在 $[r, R]$ 上定义的非负可积函数. 证明:

- (1) 对球外一点 ($|x| > R$) 产生的引力与球的所有质量集中于球心的质点产生的引力一样.
- (2) 球壳对球壳内任意一点 ($|x| < r$) 产生的引力 (合力) 为零.

Proof. 我们只需证明对任意球壳 $\Omega: |x| = c$ 作为子流形在球外一点产生的引力相当于质量集中在球心; 在球内一点产生的引力为零. 接上一问的参数化, 并且根据对称性不妨设 $x = (0, 0, a)$, 那么产生的引力为

$$\begin{aligned} \int_{\Omega} \frac{y-x}{|y-x|^3} d\sigma &= \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \frac{(c \cos \theta \sin \phi, c \cos \theta \cos \phi, c \sin \theta - a)}{(c^2 + a^2 - 2ac \sin \theta)^{3/2}} c^2 \cos \theta d\theta d\phi \\ &= 2\pi c^2 \int_{-\pi/2}^{\pi/2} \frac{(0, 0, c \sin \theta - a)}{(c^2 + a^2 - 2ac \sin \theta)^{3/2}} \cos \theta d\theta. \end{aligned}$$

因此今后我只考虑第三个分量上的积分值. 考虑换元 $u = \sin \theta$,

$$\begin{aligned} \text{上式} &= 2\pi c^2 \int_{-1}^1 \frac{cu - a}{(c^2 + a^2 - 2acu)^{3/2}} du \\ &= 2\pi c^2 \int_{-1}^1 \frac{cu - a}{ac} ((c^2 + a^2 - 2acu)^{-0.5})' du \\ &= 2\pi c^2 \left(\frac{cu - a}{ac} (c^2 + a^2 - 2acu)^{-0.5} \Big|_{-1}^1 - \int_{-1}^1 \frac{1}{a} (c^2 + a^2 - 2acu)^{-0.5} du \right) \\ &= 2\pi c^2 \left(\frac{c-a}{ac|c-a|} + \frac{c+a}{ac|c+a|} - \left(-\frac{1}{a^2c} (c^2 + a^2 - 2acu)^{0.5} \right) \Big|_{-1}^1 \right) \\ &= 2\pi c^2 \left(\frac{c-a}{ac|c-a|} + \frac{c+a}{ac|c+a|} + \frac{|c-a|}{a^2c} - \frac{|c+a|}{a^2c} \right). \end{aligned}$$

当 $a < c$ 时,

$$\text{上式} = 2\pi c^2 \left(\frac{2}{ac} - \frac{2a}{a^2c} \right) = 0.$$

当 $a \geq c$ 时,

$$\text{上式} = 2\pi c^2 \left(\frac{-2}{a^2} \right) = \frac{-4\pi c^2}{a^2}.$$

因为球壳的表面积就是 $4\pi c^2$, 所以这个式子和把质量完全放在球心的质点产生的引力一样, 证毕. 