2.3 Sard 定理

本节将要证明的主要定理是：对于任意光滑映射 $f: M \to N$, N 中“几乎所有点”(即某个在测度意义或者 Baire 纲意义下可忽略集合的补集中的所有点)都是 f 的正则值. 为此，我们将首先给出临界点/值与正则点/值的概念，然后在上述两种意义下分别证明主要定理。

2.3.1 临界点和临界值

¶ 临界点和临界值：定义

根据隐函数定理，若 $f: U \subset \mathbb{R}^m \to \mathbb{R}^n$ 是一个光滑映射，且 f 在点 a 处是淹没，即 $df_a: \mathbb{R}^m \to \mathbb{R}^n$ 是满射，则 f 在 a 点附近具有比较好的性态（例如 f 的水平集可以被表示成某个映射的图像）. 换而言之，f 的性态变化比较大的点是使得 df_a 不是满射的点. 在分析中这些点被称为 f 的临界点:

定义 2.3.1. （临界点/临界值与正则点/正则值）

设 M, N 为光滑流形，$p \in M, q \in N$，而 $f: M \to N$ 是光滑映射.

(1) 若 $df_p: T_pM \to T_{f(p)}N$ 不是满射[a]，则称 p 是 f 的**临界点**.

(2) 若点 p 是 f 的临界点，则称它的像 $f(p) \in N$ 为 f 的**临界值**.

(3) 若点 p 不是 f 的临界点，则称它是 f 的**正则点**.

(4) 若点 q 不是 f 的临界值，则称它是 f 的**正则值**.

记 f 的所有临界点所组成的集合为 $\mathrm{Crit}(f)$.

[a]本书将临界点定义为那些使得 $\mathrm{rank}(df_p) < \dim N$ 的点，而在一些书中，临界点被定义为那些使得 $\mathrm{rank}(df_p) < \min(\dim M, \dim N)$ 的点.

♣

注 2.3.2. 令 $f: M \to N$ 为一个光滑映射.

- 由定义，任意 $q \in N \setminus \mathrm{Im}(f)$ 都自动是正则值，故映射的正则值未必是映射的值!
- 临界点的像都是临界值，但正则点的像未必是正则值，也有可能是临界值.
- 由定义，正则点构成的集合在 M 中是开集:

$$\begin{aligned} p \text{ 是 } f \text{ 的正则点} &\iff f \text{ 是 } p \text{ 处的淹没} \\ &\iff f \text{ 是 } p \text{ 附近的淹没, 即 } p \text{ “附近” 的点都是正则点.} \end{aligned}$$

  于是临界点的集合 $\mathrm{Crit}(f)$ 在 M 中是闭集.

¶ 例子

对于临界点的概念，我们更熟悉的是它的两种特殊情形:

- 光滑函数$f: \mathbb{R} \to \mathbb{R}$ 的临界点是那些使得 $f'(a) = 0$ 的点 a.
- 光滑函数$f: \mathbb{R}^m \to \mathbb{R}$ 的临界点是那些使得对所有的 i 都有 $\frac{\partial f}{\partial x^i}(a) = 0$ 的点 a.

一般而言，若 $f \in C^\infty(M)$，则 $p \in M$ 是 f 的临界点当且仅当 $df_p = 0$. 特别地，“任何光滑函数的极大值/极小值点是临界点”这个结果在光滑流形上仍然成立:

命题 2.3.3.（极值点是临界点）

设 $f\in C^\infty(M)$ 为光滑函数，且 $p\in M$ 是 f 取到（局部）极大值或极小值的点. 那么 p 是 f 的临界点.

证明 根据注记2.1.9，对于任意 $X_p=\sum a_i\partial_i|_p\in T_pM$，有

$$df_p(X_p)=X_p(f)=\sum a_i\partial_i|_p(f)=\sum a_i\frac{\partial(f\circ\varphi^{-1})}{\partial x^i}(\varphi(p))=0.$$

由 X_p 的任意性可知 p 是 f 的一个临界点. □

特别地，定义在紧流形上的光滑函数都存在至少两个临界点.

例 2.3.4. 北极点 $(0,\cdots,0,1)$ 和南极点 $(0,\cdots,0,-1)$ 是"高度函数"

$$f:S^n\to\mathbb{R},\qquad f(x^1,\cdots,x^{n+1})=x^{n+1}$$

的临界点. 不难验证 S^n 上其他点都是 f 的正则点，因此 f 的临界值只有 1 和 -1.

例 2.3.5. 对于以下两种极端情况，

- $f:M\to N$ 是常值映射，即 $f(p)\equiv q_0\in N$，
- $f:M\to N$ 是光滑映射，但是 $\dim M<\dim N$，

M 中的任意一点都是临界点，因此像集 $f(M)$ 中的任意一点都是 f 的临界值.

例 2.3.6. 存在一个光滑函数 $f\in C^\infty(\mathbb{R})$，其临界值在 $\mathbb{R}$ 中稠密：记 $\mathbb{Q}=\{r_1,r_2,\cdots\}$. 取一个定义在 $\mathbb{R}$ 上的光滑鼓包函数 f_0，使得

$$\mathrm{supp}(f_0)\subset(-1/3,1/3)\quad 且\quad f_0\equiv 1\text{ on }(-1/4,1/4).$$

令

$$f(x)=\sum_{k=1}^{\infty}r_kf_0(x-k).$$

则每个 $k\in\mathbb{N}$ 都是 f 的一个临界点，从而 f 的临界值包含 $f(\mathbb{N})=\mathbb{Q}$.（思考题：找到临界值集合不可数的光滑函数.）

2.3.2 Sard 定理

¶ 光滑流形上的零测集

首先解释一下在光滑流形上哪些集合"在测度意义下可忽略"，即"零测集". 注意我们还没有在 M 或 N 上引入任何测度结构. 事实上，如果仅仅需要讨论光滑流形 M 上的"零测集"，那么并不需要在 M 上引入测度：只需把欧氏空间中的"Lebesgue 零测"这个概念移植过去即可.

在将欧氏空间里的概念移植到光滑流形上时，需要非常小心：虽然看起来利用局部坐标卡，就可以将欧氏空间上的 Lebesgue 测度"移植"到流形上，但是这样得到的"测度"并不是在流形上良好定义的的测度，因为同一个区域在该"测度"下的值取决于局部坐标卡的选择. 事实上，测度结构是流形上的一种额外的结构. 仅有光滑结构时，是无法典范地定义测度结构的.（对于一个可定向流形，每个体积形式都给出了一个测度，见本书后文）.

然而，即使不引入测度结构，依然可以在光滑流形上谈论"一个集合是否是零测集"：

借助于欧氏空间的 Lebesgue 测度，虽然对于给定的集合，它在坐标映射下像的 Lebesgue 测度值一般而言依赖于局部坐标卡的选取，但是“该像集是否是零测集”这一点是不依赖于坐标卡的选取的. 回忆一下，集合 $A \subset \mathbb{R}^n$ 是**零测集**，是指对于任意 $\varepsilon > 0$，存在可数个 n 维开箱体 $U_i \in \mathbb{R}^n$，使得

$$A \subset \bigcup_i U_i \qquad 且 \qquad \sum_i \mathrm{volume}(U_i) < \varepsilon.$$

对于 $\mathbb{R}^n$ 中的 Lebesgue 零测集，有以下性质：

(i) 可数个零测集的并仍然是零测集.

(ii) 若 $A \subset \mathbb{R}^n$ 是零测集，且 $f : \mathbb{R}^n \to \mathbb{R}^n$ 是光滑映射，则 $f(A)$ 是零测集. [4]

(iii) (Fubini 定理的特殊形式) 设 A 为 $\mathbb{R}^n$ 的可测子集，且对于任意 $c \in \mathbb{R}^r$，“切片” $A \cap (\{c\} \times \mathbb{R}^{n-r})$ 在 $\mathbb{R}^{n-r}$ 中都是零测集，那么 A 是 $\mathbb{R}^n$ 中的零测集.

因为任意流形 M 可以被可数多坐标卡覆盖，并且每个坐标卡将 M 中的一个开集等同于 $\mathbb{R}^n$ 中的一个开集，以下定义跟坐标卡的选取无关 (从而是良定的)：

定义 2.3.7. (光滑流形上的零测集)

设 M 是光滑流形，$A \subset M$ 是一个子集. 若对于任意 $p \in A$，可以找到 M 的在 p 附近的一个坐标卡 (φ, U, V)，使得 $\varphi(A \cap U)$ 是 V 中的零测集，则称 A 是 M 中的一个**零测集**. ♣

显然，光滑流形上可数多个零测集的并仍然是零测集.

¶ Sard 定理

在做完这些准备工作后，现在可以陈述并证明 Sard 定理，它是微分拓扑中非常重要也非常有用的一个定理：[5]

定理 2.3.8. (Sard 定理)

对于任意光滑映射 $f : M \to N$，f 的临界值集合在 N 中是零测集. ♡

注意

- Sard 定理并没有断言 f 的临界点集合是 M 中的零测集. 事实上，我们已经看到，甚至可能 M 上的所有点都是临界点.
- 如果 $n = \dim N = 0$，那么 f 没有临界点 (因此也没有临界值). 下面总假设 $n > 0$.

因为总是可以用至多可数个坐标卡 U_i 覆盖 M，使得每个 $f(U_i)$ 都包含于 N 的某个坐标卡 X_i，而且零测集的可数并仍然是零测集，所以只需要证明 Sard 定理对于欧氏空间中开集之间的光滑映射成立即可，即只需证明

[4] 注意连续函数可以将 $\mathbb{R}^n$ 中的零测集映射到 $\mathbb{R}^n$ 中的正测度集，于是在拓扑流形上用这种方式定义的“零测集”不是良定的.

[5] 该定理在 $m = 1$ 的情形首先由 Morse 于 1939 年证明，后来 1942 年由 Sard 将它推广到了一般情况. 无穷维的 Banach 流形版本的 Sard 定理则是 1965 年由 S. Smale 证明的.

定理 2.3.9.（欧氏空间光滑映射的 Sard 定理） 如果 $U \subset \mathbb{R}^m$ 是欧氏空间中的开集，并且 $f : U \to \mathbb{R}^n$ 是光滑映射，那么 f 的临界值集合是 $\mathbb{R}^n$ 中的零测集. ♡

证明 首先观察到如果 $m < n$, 那么整个像集 $f(U)$ 都是临界点。为了说明 $f(U)$ 是 $\mathbb{R}^n$ 中的零测集，将 U 等同于 $U \times \mathbb{R}^{n-m} \subset \mathbb{R}^n$ 中的子集 $U \times \{0\}$. 定义映射

$$\tilde{f} : U \times \mathbb{R}^{n-m} \to \mathbb{R}^n, \quad \tilde{f}(x, y) = f(x),$$

则 $\tilde{f}$ 是同维数的欧氏空间之间的光滑映射，并且 $f(U) = \tilde{f}(U \times \{0\})$。由于 $U \times \{0\}$ 是 $\mathbb{R}^n$ 中的零测集, 所以像集 $\tilde{f}(U \times \{0\})$ 是 $\mathbb{R}^n$ 中的零测集.

接下来采用归纳法. 定理对于 $m = 0$ 显然是成立的，因为任意可数集都是零测集. 下面证明：若定理对 $m-1$ 成立的，则它对 m 也成立. 记 $C = \mathrm{Crit}(f)$ 是 f 的所有临界点的集合。为了证明 $f(C)$ 在 N 中为零测集，记

$$C_j = \{x \in U \mid \partial^\alpha f_i(x) = 0, \forall |\alpha| \leq j, \forall i\}.$$

显然，对于任意正整数 k,

$$f(C) = f(C \backslash C_1) \cup f(C_1 \backslash C_2) \cup \cdots \cup f(C_{k-1} \backslash C_k) \cup f(C_k).$$

下面将按照 J. Milnor 的证明方法, 分三步证明 $f(C)$ 是零测集:

步骤 1: $f(C \backslash C_1)$ 是零测集.

步骤 2: 对于任意 i, $f(C_i \backslash C_{i+1})$ 是零测集.

步骤 3: 当 k 足够大时，$f(C_k)$ 是零测集,

步骤 1 的证明. 对于任意 $x \in C \backslash C_1$，下面构造开集 $U_x \ni x$ 使得 $f(U_x \cap C)$ 是零测集. 因为 $C \backslash C_1$ 能被可数多这样的开集覆盖 (由第二可数性), 所以 $f(C \backslash C_1)$ 是零测集.

注意到如果 $n = 1$, 那么点 x 是 f 的临界点当且仅当对于所有的 i, $\frac{\partial f}{\partial x^i}(x) = 0$. 于是 $C \backslash C_1 = \emptyset$, 从而 $f(C \backslash C_1)$ 必然是零测集. 下面假设 $m \geq n > 1$. 因为 $x \notin C_1$, 在 x 处存在某些偏导数非零. 不妨设 $\frac{\partial f_1}{\partial x^1} \neq 0$. 考虑

$$h : U \to \mathbb{R}^m, \quad h(x) = (f_1(x), x^2, \cdots, x^m).$$

则 dh_x 是非奇异的，从而根据反函数定理, h 将 x 的某个邻域 U_x 微分同胚地映射到 $\mathbb{R}^m$ 中的某开集 V. 于是复合映射 $g = f \circ h^{-1}$ 将 V 映到 $\mathbb{R}^n$ 中. 此外，因为 h^{-1} 是 V 上的一个微分同胚, dh^{-1} 在 V 上处处是线性同构. 因此 g 的临界值集合恰好是 $f(U_x \cap C)$.

根据定义，映射 g 具有以下形式

$$g(t, x^2, \cdots, x^m) = (t, g_2, \cdots, g_n).$$

因此对于每个 t, g 诱导了一个光滑映射 $g^t : (\{t\} \times \mathbb{R}^{m-1}) \cap V \to \{t\} \times \mathbb{R}^{n-1}$. 此外,

$$dg = \begin{pmatrix} 1 & 0 \\ * & \left(\frac{\partial (g^t)_i}{\partial x^j}\right)_{i,j \geq 2} \end{pmatrix}.$$

于是 $(\{t\} \times \mathbb{R}^{m-1}) \cap V$ 中的一个点是 g^t 的临界点当且仅当它是 g 的临界点. 但是由归纳假设，Sard 定理对于 $m-1$ 成立, 从而对于每个 g^t 成立. 因此 g^t 的临界值集合在 $\{t\} \times \mathbb{R}^{n-1}$ 中是零测集. 于是由 Fubini 定理的特殊情形可知 g 的临界值集合是零测集.

步骤 2 的证明. 对于每个 $x \in C_i \backslash C_{i+1}$, 取满足 $|\alpha| = i$ 的某个多重指标 α, 使得

- 偏导数 $w := \partial^\alpha f$ 在 C_i 上为零,
- w 的一阶偏导数至少有一个在 x 处不为零. 不妨设 $\frac{\partial w}{\partial x^1}(x) \neq 0$.

由反函数定理知

$$h : U \to \mathbb{R}^m, \quad h(x) = \big(w(x), x^2, \cdots, x^m\big)$$

将 x 的邻域 U_x 微分同胚地映为 $\mathbb{R}^m$ 中的开集 V. 因为 $h(C_i \cap U_x) \subset \{0\} \times \mathbb{R}^{m-1}$, 所以映射 $g = f \circ h^{-1}$ 的 "类型为 C_i 的临界点" 都在超平面 $\{0\} \times \mathbb{R}^{m-1}$ 上. 令

$$\bar{g} : (\{0\} \times \mathbb{R}^{m-1}) \cap V \to \mathbb{R}^n$$

为 g 的限制映射, 则 g 的 "类型为 C_i 的临界点" 恰好是 $\bar{g}$ 的临界点. 由归纳假设, $\bar{g}$ 的临界值集是 $\mathbb{R}^n$ 中的零测集, 故 g 的 "类型为 C_i 的临界点" 的像为零测集, 即 $f(C_i \cap U_x)$ 为零测集. 由于 $C_i \setminus C_{i+1}$ 能被可数个这样的集合 U_x 覆盖, $f(C_i \setminus C_{i+1})$ 是零测集.

步骤 3 的证明. 令 $Q \subset U$ 为一个边长为 δ 的 m 维闭方体. 下证: 对于 $k > \frac{m}{n} - 1$, $f(C_k \cap Q)$ 是零测集. 因为 C_k 能被可数多个这样的方体覆盖, 所以 $f(C_k)$ 是零测集.

由 Q 的紧性以及 C_k 的定义, 对于满足 $x \in C_k \cap Q$ 且 $x + h \in Q$ 的 (x, h), f 的 Taylor 展开形如

$$f(x + h) = f(x) + R(x, h),$$

其中余项 $|R(x, h)| < a|h|^{k+1}$, 且常数 a 仅依赖于 f 和 Q. 接着将 Q 细分为 l^m 个边长是 $\frac{\delta}{l}$ 的小方体, 并令 Q_1 为细分中包含点 $x \in C_k$ 的小方体. 那么 Q_1 中的任意点可以写作 $x + h$, 且 $|h| \leq \sqrt{m}\frac{\delta}{l}$. 于是由上述 Taylor 展开的余项估计, $f(Q_1)$ 落在一个边长为 $\frac{b}{l^{k+1}}$ 且中心为 $f(x)$ 的方体中, 其中常数 $b = 2a(\sqrt{m}\delta)^{k+1}$. 因此 $f(C_k \cap Q)$ 包含于至多 l^m 个小方体的并集中, 且这些小方体的总体积为

$$\text{体积} \leq l^m (\frac{b}{l^{k+1}})^n = b^n l^{m-(k+1)n}.$$

因为 $k > \frac{m}{n} - 1$, 所以当 $l \to \infty$ 时 体积 $\to 0$. 故 $f(C_k \cap Q)$ 是零测集. $\square$

注 2.3.10. 设 $r \geq 1$ 且 M, N 都是 C^r 流形。

(1) 根据上述证明, 若 $m \leq n$, 则对于任意 C^1 映射 $f : M \to N$, Sard 定理依然成立。

(2) 若 $m \leq n$, 且 $f : M \to N$ 是一个 C^r 映射, 则当 $r \geq 1 + m - n$ 时 Sard 定理依然成立, 但其证明更加复杂, 参见 [6], [7].

- 上述证明对于这种更一般的情形并不适用, 因为 "步骤 2" 中对 $\bar{g}$ 用了归纳假设, 而 $\bar{g}$ 的光滑性不够。
- 这里关于 r 的条件不能去掉, 反例见习题。

2.3.3 Sard 定理的一个变体

在一些应用中, 需要处理的未必是有限维流形之间的光滑映射, 而是某些 "无穷维流形之间的光滑映射", 此时因为模型空间 (例如无穷维 Banach 空间或者 Hilbert 空间) 中不再有 "平移不变 Lebesgue 测度", 所以用 "零测" 这个概念表述的 Sard 定理无法直接推广到该情形。好在天无绝人之路, 还有用其它方式来表述 "拓扑空间里某个集合

在大小上可忽略”这个现象，即采用 Baire 纲集的语言。从某种意义上说，Baire 纲集的语言是无穷维空间中测度的一个替代品，在泛函分析中已经看到过 Baire 纲集起到的基本性作用。下面给出用 Baire 纲集表述的 Sard 定理。当然本书并不研究无穷维流形，所以这里考虑的依然是 (有限维) 光滑流形流形之间的光滑映射。

首先回忆一下拓扑空间中的 Baire 纲集的三个相关概念，即 Baire 第一纲集（也称作贫集）、第二纲集、剩余集，它们分别对应于零测集、正测度集、全测度集：

定义 2.3.11. （Baire 第一纲集、第二纲集、剩余集）

设 A 是拓扑空间 X 的子集。

(1) 若 A 可以表示成可数个稠密开集的交，则称它为**剩余集**(residue set)。

(2) 若 A 是某个剩余集的补集，则称它为 **Baire 第一纲集**。

(3) 若 A 不是第一纲集，则称它 **Baire 第二纲集**。

若使某个性质成立的点集是剩余集，则称该性质为**通有性质**(generic property).

注意根据定义，一个集合是 Baire 第一纲集当且仅当它是可数个无处稠密闭集的并。例如，$\mathbb{R}$ 中的无理数集合是剩余集，而有理数集合则是第一纲集。

注 2.3.12. 即使对于 $\mathbb{R}$ 上的标准拓扑和标准测度而言，Baire 第一纲集、第二纲集、剩余集跟零测集、正测度集、全测度集也并不一致。例如，胖 Cantor 集合就是正测度的第一纲集，而 $\cap_j \cup_k (r_k - \frac{1}{2^{j+k}}, r_k + \frac{1}{2^{j+k}})$（其中 $\{r_k\}$ 是全体有理数）则是零测的剩余集。

注 2.3.13. 因为流形都是局部紧 Hausdorff 空间，所以根据点集拓扑相关内容，流形都是 Baire 空间（即可数个稠密开集的交依然是稠密的）。

下面给出用 Baire 纲集语言表述的 Sard 定理，即“正则值是 N 里面‘通用’的点都是正则点”(注意该定理跟定理2.3.8互不包含)：

定理 2.3.14. (Sard 定理：Baire 纲集的变体)

设 $f: M \to N$ 是光滑映射，则 f 的正则值集合是 N 中的剩余集.

证明 记 f 的临界点集合为 C。下证任意 $x \in C$ 都有一个邻域 U_x，使得限制在 U_x 后 f 的正则值集合 $N \setminus f(C \cap \overline{U_x})$ 是 N 中的稠密开集。因为 C 可以被可数个这样的开集 $\{U_1, U_2, \cdots\}$ 所覆盖，所以正则值的集合

$$N \setminus f(C) = N \setminus \bigcup_i f(C \cap \overline{U_i}) = \bigcap_i N \setminus f(C \cap \overline{U_i})$$

是一个剩余集.

为了证明这样的 U_x 的存在性，只需选取 x 的坐标邻域 U_x 使得 $\overline{U_x}$ 是紧集. 因为 C 是 M 中的闭集，$C \cap \overline{U_x}$ 也是紧集。故 $f(C \cap \overline{U_x})$ 是紧集，从而也是 N 中的闭集。于是 $N \setminus f(C \cap \overline{U_x})$ 是 N 中的开集.

另一方面，根据 Sard 定理，$f(C \cap \overline{U_x})$ 具有零测度。故对于任意 $y \in N$ 以及 y 的任意小邻域 V，必有

$$V \cap (N \setminus f(C \cap \overline{U_x})) \neq \emptyset,$$

否则 $V \subset f(C \cap \overline{U_x})$，跟 $f(C \cap \overline{U_x})$ 是零测集矛盾. 故 $N \setminus f(C \cap \overline{U_x})$ 在 N 中稠密. □