第 2 章 光滑映射的微分及其应用

上一章构建了“光滑范畴”，在该范畴中对象是光滑流形，而态射则是光滑映射。抽象范畴论的基本思想之一是用对象之间的态射来研究对象本身。本章的目的就是深入研究光滑映射，并利用光滑映射研究光滑流形本身。

2.1 光滑映射的微分

本节旨在定义光滑映射的微分。对一个光滑映射在一个给定点取微分，本质上就是在该点附近用线性映射逼近原映射，即“以直代曲”的线性化过程。为此，需要先定义光滑流形在每点处的切空间，作为该线性化映射的承载空间。

2.1.1 切空间

¶ 欧氏空间中光滑映射的微分

首先回顾一下欧氏空间开集间光滑映射的微分。设 U, V 为欧氏空间中的开集，且 $f : U \to V$ 是一个光滑映射. f 在点 $a \in U$ 处的**微分**（或称**切映射**）是一个线性映射 $df_a : \mathbb{R}^n \to \mathbb{R}^m$. 它 (在典范基下的) 矩阵是 f 在 a 处的 **Jacobi 矩阵**, 即

$$df_a = \begin{pmatrix} \frac{\partial f_1}{\partial x^1}(a) & \cdots & \frac{\partial f_1}{\partial x^n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x^1}(a) & \cdots & \frac{\partial f_m}{\partial x^n}(a) \end{pmatrix}.$$

在多变量微积分中已经所看到，线性映射 df_a 在研究光滑映射 f 时扮演了关键角色，因为它本质上是 f 在 a 附近的“线性化”：

$$\lim_{x \to a} \frac{\|f(x) - f(a) - df_a(x-a)\|}{\|x - a\|} = 0.$$

关于微分的一个非常有用的事实是: [1]

命题 2.1.1.（链式法则）

如果映射 $f : U \to V$ 在 $x = a$ 处可微，而映射 $g : V \to W$ 在 $x = f(a)$ 处可微，那么复合映射 $g \circ f : U \to W$ 在 $x = a$ 处也可微, 并且

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

♠

¶ 切向量定义背后的想法

设 M, N 是光滑流形，且 $f : M \to N$ 是光滑映射. 我们希望跟欧氏空间情形一样，将 f 在点 p 处的微分 df_p 定义为对应切空间之间某个线性映射, 作为映射 f 在点 p 附近的线性化. 为此，首先需要解决的问题是: 什么是光滑流形在一点处的切空间?

---

[1] 考虑两个范畴，第一个范畴是以“欧氏空间中‘带点开集’(U, a)”为对象，以“光滑映射”为态射，第二个范畴是以“线性空间”为对象，以“线性映射”为态射. 那么 d 可被视作是从第一个范畴到第二个范畴的一个“函子”，而链式法则只不过是函子性质的一部分.

从熟悉的例子开始. 在数学分析以及古典微分几何中已经学过空间中"曲线的切线"以及"曲面的切平面"的概念. 例如, 若

$$f=(f_1,f_2,f_3):D\to\mathbb{R}^3$$

是空间里的一个曲面的 (局部) 参数方程, 则该曲面在点 $f(u,v)$ 处的切平面是由向量 $(\frac{\partial f_1}{\partial u},\frac{\partial f_2}{\partial u},\frac{\partial f_3}{\partial u})^T$ 和 $(\frac{\partial f_1}{\partial v},\frac{\partial f_2}{\partial v},\frac{\partial f_3}{\partial v})^T$ 所张成的平面, 而该平面恰好是 f 在点 (u,v) 处的微分 $df_{(u,v)}$ 的像集.

一般地, 若 M 是 $\mathbb{R}^N$ 中的一个具体流形(即将学到的 Whitney 嵌入定理说明这总是正确的), 那么总可以选择 p 附近的一个坐标卡 (φ,U,V), 使得 $\varphi^{-1}:V\to U$ 是一个微分同胚 (这个映射可被视为是该流形的局部参数方程). 将从 M 到 $\mathbb{R}^N$ 的嵌入记为 $\iota:M\hookrightarrow\mathbb{R}^N$, 就得到欧氏空间中开集之间的一个光滑映射

$$\iota\circ\varphi^{-1}:V\to\mathbb{R}^N,$$

从而可以跟空间中曲面情形类似, 把切空间 T_pM 定义为以下线性映射

$$d(\iota\circ\varphi^{-1})_{\varphi(p)}:\mathbb{R}^n\to\mathbb{R}^N$$

的像集. 当然, 需要验证用这种方式定义的空间 T_pM 不依赖于坐标卡的选取, 并且, 因为从 M 到欧氏空间的嵌入方式不唯一, 还需要研究不同嵌入所得的切空间 T_pM 之间的关系.

因为目前我们并不先验地知道光滑流形是否可被嵌入欧氏空间, 而且又没有一个简洁优美的"几何图像"去实现一个抽象的流形, 下面将仅使用 M 本身的信息去内蕴地定义切空间 T_pM. 为了理解下文中将要给出的"光滑流形在每一点处切空间"的抽象定义, 我们先仔细考察欧氏空间的情况. 基本思路是:

- 在给定点 a 处的任意向量 $\vec{v}\in\mathbb{R}^n$ 都可被视作在 a 处的方向导数,
- 方向导数有一个纯代数的刻画, 且该刻画可以被推广到光滑流形上.

于是, 可以将切空间定义为由这些"用代数方法定义的方向导数"所构成的线性空间!

¶ 欧氏空间中方向导数的代数刻画

先回顾一下: 对于任意 $\vec{v}\in\mathbb{R}^n$, n 元函数 f 在 x 处沿着方向 v 的**方向导数**是

$$D_{\vec{v}}^a f:=\lim_{h\to 0}\frac{f(a+h\vec{v})-f(x)}{h}=\frac{d}{dt}\Big|_{t=0}f(a+t\vec{v})=df_a(\vec{v}).$$

因此对于每个给定的点 a 以及向量 $\vec{v}$, 都有一个算子

$$D_{\vec{v}}^a:C^\infty(\mathbb{R}^n)\to\mathbb{R}.$$

在坐标表示下, 如果 $\vec{v}=\langle v^1,\cdots,v^n\rangle^T$, 那么由链式法则可得 $D_{\vec{v}}^a f=\sum v^i\frac{\partial f}{\partial x^i}(a)$. 换句话说, 作为 $C^\infty(\mathbb{R}^n)$ 上的一个算子, 有

$$D_{\vec{v}}^a=\sum_i v^i\frac{\partial}{\partial x^i}\Big|_{x=a}.$$

当然, $D_{\vec{v}}^a:C^\infty(\mathbb{R}^n)\to\mathbb{R}$ 是一个非常特殊的算子: 它是一个线性算子

$$D_{\vec{v}}^a(\alpha f+\beta g)=\alpha D_{\vec{v}}^a f+\beta D_{\vec{v}}^a g,\quad\forall\alpha,\beta\in\mathbb{R}$$

并且它满足（在 a 处的）**Leibnitz 法则**:

$$D_{\vec{v}}^{a}(fg)=f(a)D_{\vec{v}}^{a}g+g(a)D_{\vec{v}}^{a}f.$$

反之，这两个性质刻画了方向导数：

命题 2.1.2. （方向导数的代数刻画）

如果 $D: C^\infty(\mathbb{R}^n)\to\mathbb{R}$ 是线性的并且满足在 a 处的 Leibnitz 法则，即

$$D(fg)=f(a)D(g)+g(a)D(f),$$

那么存在 a 处的某个向量 $\vec{v}$ 使得 $D=D_{\vec{v}}^{a}$.

证明 对于任意 $f\in C^\infty(\mathbb{R}^n)$, 有

$$f(x)=f(a)+\int_0^1\frac{d}{dt}f(a+t(x-a))dt=f(a)+\sum_{i=1}^{n}(x^i-a^i)h_i(x),$$

其中

$$h_i(x)=\int_0^1\frac{\partial f}{\partial x^i}(a+t(x-a))dt.$$

另一方面，可以用 Leibnitz 法则计算 $D(1)$，其中 1 表示恒取常值 1 的函数：

$$D(1)=D(1\cdot 1)=2D(1)\Longrightarrow D(1)=0.$$

再结合线性性，对于任意常数 c 都有 $D(c)=0$. 因此

$$D(f)=0+\sum_{i=1}^{n}D(x^i)h_i(a)+\sum_{i=1}^{n}(a^i-a^i)D(h_i)=\sum_{i=1}^{n}D(x^i)\frac{\partial f}{\partial x^i}(a).$$

由此可得，作为 $C^\infty(\mathbb{R}^n)$ 上的算子，

$$D=\sum_{i=1}^{n}D(x^i)\left.\frac{\partial}{\partial x^i}\right|_{x=a}.$$

于是只要令 $\vec{v}=\langle D(x^1),\cdots,D(x^n)\rangle$, 就有 $D=D_{\vec{v}}^{a}$. □

注 2.1.3. 一般地，人们把满足 Leibnitz 性质的线性映射叫做导子：

定义 2.1.4. （导子）

若 A 是域 k(例如 $\mathbb{R}$) 一个代数（例如 $C^\infty(U)$、$C^\infty(M)$），B 是 A 上的一个双模（例如 $\mathbb{R}$、$C^\infty(M)$ 等），且线性算子 $d: A\to B$ 满足 Leibnitz 法则

$$d(uv)=(du)v+u(dv),$$

则称 d 为 A 的一个（取值于 B 的）**导子**。

不难验证对于给定的 A 和 B，所有导子组成一个线性空间。本书后续章节中还将出现很多对应于不同代数的导子，例如向量场、李导数等。

下面考虑 “(几何) 向量-(代数) 导子” 对应关系

$$\vec{v}\rightsquigarrow D_{\vec{v}}^{a}.$$

我们有

- 该对应是从 (由在 a 处的全部切向量构成的) 线性空间 $\mathbb{R}^n$ 到 “由在 a 点处的全部

导子构成的线性空间 $\mathcal{D}$" 的线性映射:

$$D^a_{\alpha\vec{v}+\beta\vec{w}} = \alpha D^a_{\vec{v}} + \beta D^a_{\vec{w}}.$$

- 它是单射: 如果 $\vec{v}_1 \neq \vec{v}_2$, 那么 $D^a_{\vec{v}_1} \neq D^a_{\vec{v}_2}$ (请读者尝试去证明它).
- 它是满射: 这正是命题2.1.2的结论.

因此 "点 a 处的切向量 $\vec{v}$ 构成的线性空间" 与 "点 a 处的导子构成的线性空间" 是线性同构的, 即可以将点 a 处所有切向量的向量空间等同于点 a 处所有导数的向量空间!

¶ 光滑流形在一点处的切空间

现在回到光滑流形的情形. 虽然在抽象框架里并没有 "几何向量", 但仍然有全体光滑函数构成的代数 $C^\infty(M)$. 跟欧氏空间情形一样, 可以代数地定义在一点处的导子, 并称之为该点处的切向量:

定义 2.1.5. (切向量)

令 M 为一个 n-维光滑流形.

(1) 若 $\mathbb{R}$-线性映射 $X_p : C^\infty(M) \to \mathbb{R}$ 在点 $p \in M$ 处满足 Leibnitz 法则

$$X_p(fg) = f(p)X_p(g) + X_p(f)g(p), \qquad \forall f, g \in C^\infty(M),$$

则称 X_p 为 M 在 p 点处的一个**切向量**.

(2) 称在 p 处的全体切向量所构成的线性空间 T_pM 为 M 在 p 点处的**切空间**. ♣

应用 Leibnitz 法则和线性性易得: 如果 $f \equiv c$ 是一个常值函数, 那么 $X_p(f) = 0$. 更一般地,

引理 2.1.6. (局部常值函数的导数)

如果在 p 点的某个邻域里有 $f = c$, 那么 $X_p(f) = 0$. ♢

证明 取 M 上的鼓包函数 φ, 使得 φ 在 p 的附近恒为 1, 且在集合 $f \neq c$ 上恒为 0. 则

$$(f-c)\varphi \equiv 0.$$

因此

$$0 = X_p((f-c)\varphi) = (f(p)-c)X_p(\varphi) + X_p(f)\varphi(p) = X_p(f).$$ □

特别地, 如果在 p 点的某邻域里有 $f = g$ [2], 那么 $X_p(f) = X_p(g)$. 换而言之, $X_p(f)$ 这个数由 f 在 p 的邻域里的值决定. 因此可以将定义2.1.5中的 $C^\infty(M)$ 替换为 $C^\infty(U)$, 其中 U 是任意包含 p 的开集:

命题 2.1.7. (切空间是局部的)

设 M 是一个光滑流形, U 是 $p \subset M$ 任一开邻域, 那么作为线性空间,

$$T_pM \simeq T_pU.$$ ♠

[2] 注意对于不同的函数对, 这里取的邻域可以不同。人们用**芽**的语言来描述这种局部性: 如果在 p 的某邻域内有 $f = g$, 则我们称 f 和 g 在 p 点处定义了**相同的芽**. 不难验证 "在 p 点处定义了相同的芽" 是了 $C^\infty(M)$ 上(或更一般地, 在 $C^\infty(M,N)$ 上)的一个等价关系. 当研究局部性质的时候, 在芽上处理起来一般而言更加便利.

2.1.2 光滑映射的微分

¶ 光滑流形之间的光滑映射的微分

现在定义光滑流形之间光滑映射的微分. 我们知道, 欧氏空间中开集之间的光滑映射 $f: U \to V$ 在点 a 处的微分是一个线性映射

$$df_a : T_aU = \mathbb{R}^n_x \to T_{f(a)}V = \mathbb{R}^m_y,$$

其矩阵是 f 在 a 处的 Jacobi 矩阵 $(\frac{\partial f_i}{\partial x^j}(a))$. 为了将这个概念推广到流形之间的光滑映射, 需要仔细考察 T_aU 的两种解释: 我们已经看到了可以将 a 点处 (几何的) 向量 $\vec{v} \in \mathbb{R}^n$ 等同于 a 处 (代数的) 导子 $D^a_{\vec{v}} = \sum v^i \frac{\partial}{\partial x^i}|_{x=a}$. 注意到从几何上看,

$$df_a(\vec{v}) = \left(\frac{\partial f_i}{\partial x^j}(a)\right)\vec{v} = \left\langle \sum_j \frac{\partial f_1}{\partial x^j}(a)v^j, \cdots, \sum_j \frac{\partial f_m}{\partial x^j}(a)v^j \right\rangle^T.$$

上式右侧的向量是 $\mathbb{R}^m_y$ 中的一个几何向量, 它可以被代数地解释为在 V 里点 $f(a)$ 处的导子, 即把 $g \in C^\infty(\mathbb{R}^m_y)$ 映为

$$\sum_i \sum_j v^j \frac{\partial f_i}{\partial x^j}(a) \frac{\partial g}{\partial y^i} = \sum_j v^j \left.\frac{\partial}{\partial x^j}\right|_{x=a} (g \circ f) = D^a_{\vec{v}}(g \circ f)$$

的映射.

上面的计算表明向量 $df_a(\vec{v})$ 对应的导子恰好是在点 $f(a)$ 处 "将 $g \in C^\infty(\mathbb{R}^m)$ 映射到 $D^a_{\vec{v}}(g \circ f)$" 的那个导子. 由此启发我们定义

定义 2.1.8. (光滑映射的微分)

对于任意光滑映射 $f: M \to N$. 以及任意点 $p \in M$, f 在 p 处的 **微分**是一个线性映射 $df_p : T_pM \to T_{f(p)}N$, 其定义由下式给出:

$$df_p(X_p)(g) = X_p(g \circ f), \qquad \forall X_p \in T_pM, g \in C^\infty(N)$$

♣

注 2.1.9. 对于 $\mathbb{R}$, 可以将 $T_t\mathbb{R}$ 等同于 $\mathbb{R}$,

$$c\frac{d}{dt} \in T_t\mathbb{R} \leftrightsquigarrow c \in \mathbb{R}.$$

这个对应还可以用下述方式如下理解:

> 取 $g(t) = t$, 则 $g \in C^\infty(R)$, 且跟向量 $c\frac{d}{dt}$ 所对应的实数 c 恰好就是向量 $c\frac{d}{dt}$ 作用在函数 g 上所得的结果。

特别地, 对于任意光滑函数 $f \in C^\infty(M)$ 以及任意 $X_p \in T_pM$, 在上述等同下, 跟向量 $df_p(X_p)$ 对应的实数就是

$$df_p(X_p)(g) = X_p(g \circ f) = X_p(f).$$

于是我们得到了如下非常有用的公式:

$$df_p(X_p) = X_p(f), \qquad \forall f \in C^\infty(M), \forall X_p \in T_pM,$$

¶ 微分的函子性

微分是研究光滑映射以及微分流形时最重要的工具，因为它把流形之间“非线性”的光滑映射转化为了线性空间之间的线性映射。事实上，微分 d 是从“带点光滑流形范畴”(态射为光滑映射) 到线性空间范畴 (态射是线性映射) 的函子：

定理 2.1.10. (d 的函子性)

设 M, N, P 为光滑流形，$p \in M$。

(1) 对于恒等映射 $f = \mathrm{Id}_M$，有 $df_p = \mathrm{Id}_{T_pM}$.

(2) (**链式法则**) 设 $f \in C^\infty(M, N)$，$g \in C^\infty(N, P)$，那么 $d(g \circ f)_p = dg_{f(p)} \circ df_p$. ♡

证明 对于任意 $X_p \in T_pM$ 和 $h \in C^\infty(P)$,

$$d(g \circ f)_p(X_p)(h) = X_p(h \circ g \circ f) = df_p(X_p)(h \circ g) = dg_{f(p)}(df_p(X_p))(h).$$

□

下面是函子性的标准运用：

命题 2.1.11. (从微分同胚到线性同构)

如果 $f : M \to N$ 是一个微分同胚，那么 $df_p : T_pM \to T_{f(p)}N$ 是一个线性同构. ♠

证明 对 $f^{-1} \circ f = \mathrm{Id}_M$ 和 $f \circ f^{-1} = \mathrm{Id}_N$ 应用链式法则，可以得到

$$(df^{-1})_{f(p)} \circ df_p = \mathrm{Id}_{T_pM} \quad \text{和} \quad df_p \circ (df^{-1})_{f(p)} = \mathrm{Id}_{T_{f(p)}N}.$$

于是 df_p 是一个线性同构. □

特别地，

推论 2.1.12. (切空间的维数)

如果 $\dim M = n$, 那么对于 p 处的任意局部坐标卡 (φ, U, V), 有

$$T_pM = \mathrm{span}\{\partial_1, \cdots, \partial_n\},$$

其中 $\partial_i := d\varphi^{-1}(\frac{\partial}{\partial x^i})$. 特别地，$T_pM$ 是一个 n 维线性空间. ♡

证明 令 (φ, U, V) 为 p 附近的坐标卡，则 $T_{\varphi(p)}V = \mathrm{span}(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^n})$. 另一方面，由例1.2.29，$\varphi : U \to V$ 是一个微分同胚. 由此可得

$$T_pM = T_pU = d\varphi^{-1}(T_{f(p)}V),$$

从而结论得证。 □

在这样的坐标卡里，∂_i 可以用以下具体的公式表示:

$$\partial_i : C^\infty(U) \to \mathbb{R}, \quad \partial_i(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial x^i}(\varphi(p)).$$

注 2.1.13. 结合上述两个推论，可知微分同胚的光滑流形维数一定相同。在 §1.1 曾提到过，拓扑版本“维数不变性”的证明需要用到比较深奥的拓扑工具，这里我们看到光滑版本的证明却是如此简单，由此可见线性化的威力。