

第一章 微分同胚

1.1 微分同胚

1.1.1 双射

设 f 是集合 A 到 B 的映射. 如果 A 中不同的元素有不同的像, 则称 f 为**单射** (也叫 “一对一”); 如果 B 中每个元素都是 A 中元素的像, 则称 f 为**满射**; 如果 f 既是单射又是满射, 则称 f 为**双射** (也叫 “一一对应”). 三种情况的示意图见 [图 1.1](#).

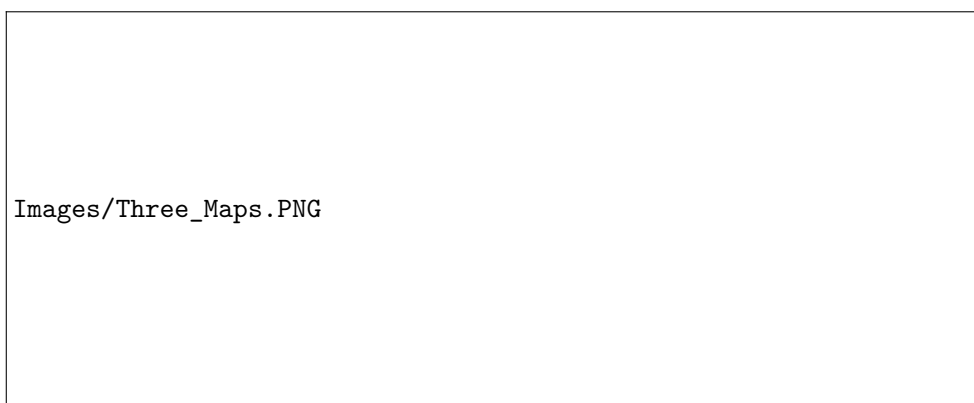


图 1.1: 单射、满射与双射

设开集 $\mathfrak{D}_X, \mathfrak{D}_x \in \mathbb{R}^m$, 它们之间存在双射, 即一一对应关系:

$$X(x) : \mathfrak{D}_x \ni x = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix} \mapsto X(x) = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix} (x) \in \mathfrak{D}_X. \quad (1.1)$$

由于该映照实现了 \mathfrak{D}_x 到 \mathfrak{D}_X 之间的双射, 因此它存在逆映照:

$$x(X) : \mathfrak{D}_X \ni X = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix} \mapsto x(X) = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix} (X) \in \mathfrak{D}_x. \quad (1.2)$$

我们把 \mathfrak{D}_X 称为**物理域**, 它是实际物理事件发生的区域; \mathfrak{D}_x 则称为**参数域**. 由于物理域通常较为复杂, 因此我们常把参数域取为规整的形状, 以便之后的处理.

设物理量 $f(\mathbf{X})$ 定义在物理域 $\mathfrak{D}_X \in \mathbb{R}^m$ 上^①，则 f 就定义了一个场：

$$f : \mathfrak{D}_X \ni \mathbf{X} \mapsto f(\mathbf{X}). \quad (1.3)$$

所谓的“场”，就是自变量用位置刻画的映照。它可以是**标量场**，如温度、压强、密度等，此时 $f(\mathbf{X}) \in \mathbb{R}$ ；也可以是**向量场**，如速度、加速度、力等，此时 $f(\mathbf{X}) \in \mathbb{R}^m$ ；对于更深入的物理、力学研究，往往还需引入**张量场**，此时 $f(\mathbf{X}) \in \mathcal{T}^r(\mathbb{R}^m)$ 。

\mathbf{X} 存在于物理域 \mathfrak{D}_X 中，我们称它为**物理坐标**。由于上文已经定义了 \mathfrak{D}_x 到 \mathfrak{D}_X 之间的双射（不是 f ！），因此 \mathfrak{D}_x 中就有唯一的 \mathbf{x} 与 \mathbf{X} 相对应，它称为**参数坐标**（也叫**曲线坐标**）。又因为物理域 \mathfrak{D}_X 上已经定义了场 $f(\mathbf{X})$ ，参数域中必然唯一存在场 $\tilde{f}(\mathbf{x})$ 与之对应：

$$\tilde{f} : \mathfrak{D}_x \ni \mathbf{x} \mapsto \tilde{f}(\mathbf{x}) = f \circ \mathbf{X}(\mathbf{x}) = f(\mathbf{X}(\mathbf{x})). \quad (1.4)$$

\mathbf{x} 与 \mathbf{X} 是完全等价的，因而 \tilde{f} 与 f 也是完全等价的，所以同样有

$$f(\mathbf{X}) = \tilde{f}(\mathbf{x}(\mathbf{X})). \quad (1.5)$$

物理域中的场要满足守恒定律，如质量守恒、动量守恒、能量守恒等。从数学上看，这些守恒定律就是 $f(\mathbf{X})$ 需要满足的一系列偏微分方程。将场变换到参数域后，它仍要满足这些方程。但我们已经设法将参数域取得较为规整，故在其上进行数值求解就会相当方便。

1.1.2 参数域方程

上文已经提到，物理域中的场 $f(\mathbf{X})$ 需满足守恒定律，这等价于一系列偏微分方程（PDE）。在物理学和力学中，用到的 PDE 通常是二阶的，它们可以写成

$$\forall \mathbf{X} \in \mathfrak{D}_X, \quad \sum_{\alpha=1}^m A_{\alpha}(\mathbf{X}) \frac{\partial f}{\partial X^{\alpha}}(\mathbf{X}) + \sum_{\alpha=1}^m \sum_{\beta=1}^m B_{\alpha\beta}(\mathbf{X}) \frac{\partial^2 f}{\partial X^{\beta} \partial X^{\alpha}}(\mathbf{X}) = 0 \quad (1.6)$$

的形式。我们的目标是把该物理域方程转化为参数域方程，即关于 $\tilde{f}(\mathbf{x})$ 的 PDE。多元微积分中已经提供了解决方案：**链式求导法则**。

考虑到

$$f(\mathbf{X}) = \tilde{f}(\mathbf{x}(\mathbf{X})) = \tilde{f}(x^1(\mathbf{X}), \dots, x^m(\mathbf{X})), \quad (1.7)$$

于是有

$$\frac{\partial f}{\partial X^{\alpha}}(\mathbf{X}) = \sum_{s=1}^m \frac{\partial \tilde{f}}{\partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \frac{\partial x^s}{\partial X^{\alpha}}(\mathbf{X}). \quad (1.8)$$

这里用到的链式法则，由复合映照可微性定理驱动，它要求 \tilde{f} 关于 \mathbf{x} 可微，同时 \mathbf{x} 关于 \mathbf{X} 可微。

对于更高阶的项，往往需要更强的条件。一般地，我们要求

$$\begin{cases} \mathbf{X}(\mathbf{x}) \in \mathcal{C}^p(\mathfrak{D}_x; \mathbb{R}^m); \\ \mathbf{x}(\mathbf{X}) \in \mathcal{C}^p(\mathfrak{D}_X; \mathbb{R}^m). \end{cases} \quad (1.9-a)$$

$$(1.9-b)$$

这里的 \mathcal{C}^p 指直至 p 阶偏导数（存在且）连续的映照全体； $p=1$ 时，它就等价于可微。至于 p 的具体取值，则由 PDE 的阶数所决定。

^① 实际的物理事件当然只会发生在三维 Euclid 空间中（只就“空间”而言），但在数学上也可以推广到 m 维。

通常情况下，已知条件所给定的往往都是 \mathfrak{D}_x 到 \mathfrak{D}_X 的映射

$$\mathbf{X}(\mathbf{x}) : \mathfrak{D}_x \ni \mathbf{x} = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix} \mapsto \mathbf{X}(\mathbf{x}) = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix} (\mathbf{x}) \in \mathfrak{D}_X, \quad (1.10)$$

用它不好直接得到式 (1.8) 中的 $\partial x^s / \partial X^\alpha$ 项，但获得它的“倒数” $\partial X^\alpha / \partial x^s$ 却很容易，只需利用 **Jacobi** 矩阵：

$$D\mathbf{X}(\mathbf{x}) \triangleq \begin{bmatrix} \frac{\partial X^1}{\partial x^1} & \cdots & \frac{\partial X^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial X^m}{\partial x^1} & \cdots & \frac{\partial X^m}{\partial x^m} \end{bmatrix} (\mathbf{x}) \in \mathbb{R}^{m \times m}, \quad (1.11)$$

它是一个方阵。

有了 **Jacobi** 矩阵，施加一些手法就可以得到所需要的 $\partial x^s / \partial X^\alpha$ 项。考虑到

$$\forall \mathbf{X} \in \mathfrak{D}_X, \quad \mathbf{X}(\mathbf{x}(\mathbf{X})) = \mathbf{X}, \quad (1.12)$$

并且其中的 $\mathbf{X}(\mathbf{x})$ 和 $\mathbf{x}(\mathbf{X})$ 均可微，可以得到

$$D\mathbf{X}(\mathbf{x}(\mathbf{X})) \cdot D\mathbf{x}(\mathbf{X}) = \mathbf{I}_{m \times m}, \quad (1.13)$$

其中的 $\mathbf{I}_{m \times m}$ 是单位阵。因此

$$D\mathbf{x}(\mathbf{X}) \triangleq \begin{bmatrix} \frac{\partial x^1}{\partial X^1} & \cdots & \frac{\partial x^1}{\partial X^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^m}{\partial X^1} & \cdots & \frac{\partial x^m}{\partial X^m} \end{bmatrix} (\mathbf{X}) = (D\mathbf{X})^{-1}(\mathbf{x}) = \begin{bmatrix} \frac{\partial X^1}{\partial x^1} & \cdots & \frac{\partial X^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial X^m}{\partial x^1} & \cdots & \frac{\partial X^m}{\partial x^m} \end{bmatrix}^{-1}(\mathbf{x}). \quad (1.14)$$

用代数的方法总可以求出

$$\frac{\partial x^s}{\partial X^\alpha} =: \varphi_\alpha^s, \quad (1.15)$$

它是通过求逆运算确定的函数，即位于矩阵 $D\mathbf{x}$ 第 s 行第 α 列的元素。这样就有

$$\frac{\partial f}{\partial X^\alpha}(\mathbf{X}) = \sum_{s=1}^m \frac{\partial \tilde{f}}{\partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \varphi_\alpha^s(\mathbf{x}(\mathbf{X})). \quad (1.16)$$

接下来处理二阶偏导数。由上式，

$$\begin{aligned} \frac{\partial^2 f}{\partial X^\beta \partial X^\alpha}(\mathbf{X}) &= \sum_{s=1}^m \left[\left(\sum_{k=1}^m \frac{\partial^2 \tilde{f}}{\partial x^k \partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \frac{\partial x^s}{\partial X^\beta}(\mathbf{X}) \right) \cdot \varphi_\alpha^s(\mathbf{x}(\mathbf{X})) \right. \\ &\quad \left. + \frac{\partial \tilde{f}}{\partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \left(\sum_{k=1}^m \frac{\partial \varphi_\alpha^s}{\partial x^k}(\mathbf{x}(\mathbf{X})) \cdot \frac{\partial x^k}{\partial X^\beta}(\mathbf{X}) \right) \right] \end{aligned}$$

继续利用式 (1.15)，有

$$\begin{aligned} &= \sum_{s=1}^m \left[\left(\sum_{k=1}^m \frac{\partial^2 \tilde{f}}{\partial x^k \partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \varphi_\beta^s(\mathbf{x}(\mathbf{X})) \right) \cdot \varphi_\alpha^s(\mathbf{x}(\mathbf{X})) \right. \\ &\quad \left. + \frac{\partial \tilde{f}}{\partial x^s}(\mathbf{x}(\mathbf{X})) \cdot \left(\sum_{k=1}^m \frac{\partial \varphi_\alpha^s}{\partial x^k}(\mathbf{x}(\mathbf{X})) \cdot \varphi_\beta^k(\mathbf{x}(\mathbf{X})) \right) \right]. \quad (1.17) \end{aligned}$$

这样，就把一阶和二阶偏导数项全部用关于 \mathbf{x} 的函数^①表达了出来。换句话说，我们已经把物理域

^① 当然它仍然是 \mathbf{X} 的隐函数： $\mathbf{x} = \mathbf{x}(\mathbf{X})$ 。

中 f 关于 \mathbf{X} 的 PDE，转化成了参数域中 \tilde{f} 关于 \mathbf{x} 的 PDE。这就是上文要实现的目标。

1.1.3 微分同胚的定义

上文已经指出了 \mathfrak{D}_x 到 \mathfrak{D}_X 的映照 $\mathbf{X}(\mathbf{x})$ 所需满足的一些条件。这里再次罗列如下：

1. $\mathfrak{D}_X, \mathfrak{D}_x \in \mathbb{R}^m$ 均为开集^①；
2. 存在 \mathfrak{D}_x 同 \mathfrak{D}_X 之间的双射 $\mathbf{X}(\mathbf{x})$ ，即存在一一对应关系；
3. $\mathbf{X}(\mathbf{x})$ 和它的逆映照 $\mathbf{x}(\mathbf{X})$ 满足一定的正则性要求。

对第 3 点要稍作说明。

如果满足这三点，则称 $\mathbf{X}(\mathbf{x})$ 为 \mathfrak{D}_x 与 \mathfrak{D}_X 之间的 \mathcal{C}^p -微分同胚，记为 $\mathbf{X}(\mathbf{x}) \in \mathcal{C}^p(\mathfrak{D}_x; \mathfrak{D}_X)$ 。把物理域中的一个部分对应到参数域上的一个部分，需要的仅仅是双射这一条件；而要使得物理域中所满足的 PDE 能够转换到参数域上，就需要“过去”和“回来”都满足 p 阶偏导数连续的条件（即正则性要求）。

1.2 向量值映照的可微性

1.2.1 可微性的定义

设 \mathbf{x}_0 是参数域 \mathfrak{D}_x 中的一个内点。在映射 $\mathbf{X}(\mathbf{x})$ 的作用下，它对应到物理域 \mathfrak{D}_X 中的点 $\mathbf{X}(\mathbf{x}_0)$ 。参数域是一个开集。根据开集的定义，必然存在一个实数 $\lambda > 0$ ，使得以 \mathbf{x}_0 为球心、 λ 为半径的球能够完全落在定义域 \mathfrak{D}_x 内，即

$$\mathfrak{B}_\lambda(\mathbf{x}_0) \subset \mathfrak{D}_x, \quad (1.18)$$

其中的 $\mathfrak{B}_\lambda(\mathbf{x}_0)$ 表示 \mathbf{x}_0 的 λ 邻域。

如果 $\exists D\mathbf{X}(\mathbf{x}_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ ^②，满足

$$\forall \mathbf{x}_0 + \mathbf{h} \in \mathfrak{B}_\lambda(\mathbf{x}_0), \quad \mathbf{X}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{X}(\mathbf{x}_0) = D\mathbf{X}(\mathbf{x}_0)(\mathbf{h}) + o(\|\mathbf{h}\|_{\mathbb{R}^m}) \in \mathbb{R}^m, \quad (1.19)$$

则称向量值映射 $\mathbf{X}(\mathbf{x})$ 在 \mathbf{x}_0 点可微。其中， $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ 表示从 \mathbb{R}^m 到 \mathbb{R}^m 的线性变换全体。

根据这个定义，所谓可微性，指由自变量变化所引起的因变量变化，可以用一个线性变换近似，而误差为一阶无穷小量。自变量可见到因变量空间最简单的映照形式就是线性映照（线性变换），因而具有可微性的向量值映照具有至关重要的作用。

1.2.2 Jacobi 矩阵

下面我们研究 $D\mathbf{X}(\mathbf{x}_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ 的表达形式。由于 $\mathbf{h} \in \mathbb{R}^m$ ，所以

$$\mathbf{h} = \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix} = h^1 \mathbf{e}_1 + \cdots + h^i \mathbf{e}_i + \cdots + h^m \mathbf{e}_m. \quad (1.20)$$

① 用形象化的语言来说，如果在区域中的任意一点都可以吹出一个球，并能使球上的每个点都落在区域内，那么这个区域就是开集。这是复合映照可微性定理的一个要求。

② 正如之前已经定义的， $D\mathbf{X}$ 已经用来表示 Jacobi 矩阵。这里还是请先暂时将它视为一种记号，其具体形式将在下一小节给出。

另一方面, $DX(\mathbf{x}_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ 具有线性性:

$$\forall \alpha, \beta \in \mathbb{R} \text{ 和 } \tilde{\mathbf{h}}, \hat{\mathbf{h}} \in \mathbb{R}^m, \quad DX(\mathbf{x}_0)(\alpha\tilde{\mathbf{h}} + \beta\hat{\mathbf{h}}) = \alpha DX(\mathbf{x}_0)(\tilde{\mathbf{h}}) + \beta DX(\mathbf{x}_0)(\hat{\mathbf{h}}). \quad (1.21)$$

这样就有

$$\begin{aligned} DX(\mathbf{x}_0)(\mathbf{h}) &= DX(\mathbf{x}_0)(h^1 \mathbf{e}_1 + \cdots + h^i \mathbf{e}_i + \cdots + h^m \mathbf{e}_m) \\ &= h^1 DX(\mathbf{x}_0)(\mathbf{e}_1) + \cdots + h^i DX(\mathbf{x}_0)(\mathbf{e}_i) + \cdots + h^m DX(\mathbf{x}_0)(\mathbf{e}_m) \end{aligned} \quad (1.22)$$

注意到 $h^i \in \mathbb{R}$ 以及 $DX(\mathbf{x}_0)(\mathbf{e}_i) \in \mathbb{R}^m$, 因而该式可以用矩阵形式表述:

$$= \begin{bmatrix} DX(\mathbf{x}_0)(\mathbf{e}_1) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_m) \end{bmatrix} \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix}. \quad (1.23)$$

最后一步要用到分块矩阵的思想: 左侧的矩阵为 1 “行” m 列, 每一 “行” 是一个 m 维列向量; 右侧的矩阵 (向量) 则为 m 行 1 列. 两者相乘, 得到 1 “行” 1 列的矩阵 (当然实际为 m 行), 即之前的 (1.22) 式. 在线性代数中, $m \times m$ 的矩阵 $\begin{bmatrix} DX(\mathbf{x}_0)(\mathbf{e}_1) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_m) \end{bmatrix}$ 通常称为变换矩阵 (也叫过渡矩阵).

接下来要搞清楚变换矩阵的具体形式. 取

$$\mathbf{h} = [0, \dots, \lambda, \dots, 0]^T = \lambda \mathbf{e}_i \in \mathbb{R}^m, \quad (1.24)$$

即除了 \mathbf{h} 的第 i 各元素为 λ 外, 其余元素均为 0 ($\lambda \neq 0$). 代入可微性的定义 (1.19) 式, 有

$$\begin{aligned} X(\mathbf{x}_0 + \mathbf{h}) - X(\mathbf{x}_0) &= X(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X(\mathbf{x}_0) \\ &= \begin{bmatrix} DX(\mathbf{x}_0)(\mathbf{e}_1) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_i) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_m) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \lambda \\ \vdots \\ 0 \end{bmatrix} + o(\lambda) \\ &= \lambda \cdot DX(\mathbf{x}_0)(\mathbf{e}_i) + o(\lambda). \end{aligned} \quad (1.25)$$

由于 λ 是非零实数, 故可以在等式两边同时除以 λ 并取极限:

$$\lim_{\lambda \rightarrow 0} \frac{X(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X(\mathbf{x}_0)}{\lambda} = DX(\mathbf{x}_0)(\mathbf{e}_i), \quad (1.26)$$

这里的 $o(\lambda)$ 根据其定义自然趋于 0. 该式左侧极限中的分子部分, 是自变量 \mathbf{x} 第 i 个分量的变化所引起因变量的变化; 而分母, 则是自变量第 i 个分量的变化大小. 我们引入下面的记号:

$$\frac{\partial X}{\partial x^i}(\mathbf{x}_0) := \lim_{\lambda \rightarrow 0} \frac{X(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X(\mathbf{x}_0)}{\lambda} \in \mathbb{R}^m, \quad (1.27)$$

它表示因变量 $\mathbf{X} \in \mathbb{R}^m$ 作为一个整体, 相对于自变量 $\mathbf{x} \in \mathbb{R}^m$ 第 i 个分量 $x^i \in \mathbb{R}$ 的 “变化率”, 即 \mathbf{X} 关于 x^i (在 \mathbf{x}_0 处) 的偏导数. 由于我们没有定义向量的除法, 因此自变量作为整体所引起因变量的变化, 是没有意义的. 利用偏导数的定义, 可有

$$\begin{aligned} &\begin{bmatrix} DX(\mathbf{x}_0)(\mathbf{e}_1) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_i) & \cdots & DX(\mathbf{x}_0)(\mathbf{e}_m) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial X}{\partial x^1}(\mathbf{x}_0) & \cdots & \frac{\partial X}{\partial x^i}(\mathbf{x}_0) & \cdots & \frac{\partial X}{\partial x^m}(\mathbf{x}_0) \end{bmatrix} \in \mathbb{R}^{m \times m}. \end{aligned} \quad (1.28)$$