

第一章 微分同胚（曲线坐标系）

1.1 微分同胚

1.1.1 双射

设 f 是集合 A 到 B 的映照. 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](#).

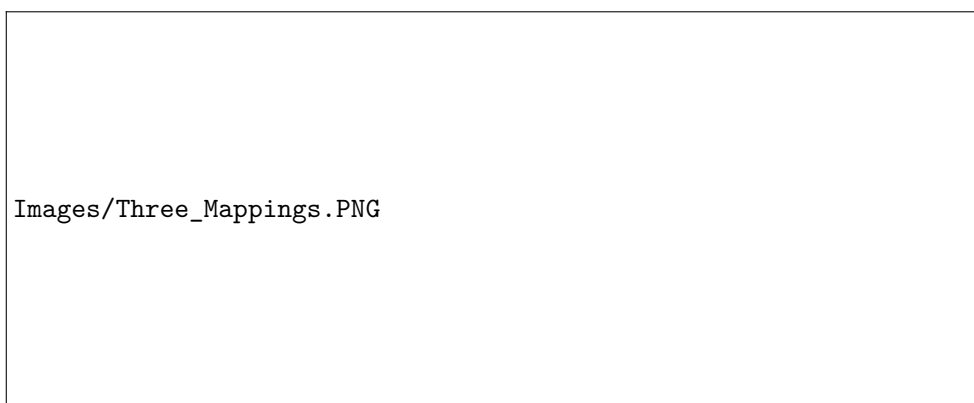


图 1.1: 单射、满射与双射

设开集 $\mathfrak{D}_X, \mathfrak{D}_x \subset \mathbb{R}^m$, 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 $\mathfrak{D}_X \subset \mathbb{R}^m$ 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 的映照

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

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**：

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

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

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

其中的 \mathbf{I}_m 是单位阵。因此

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

用代数的方法总可以求出

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

它是通过求逆运算确定的函数，即位于矩阵 $\mathbf{D}\mathbf{x}$ 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

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

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

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. $\mathfrak{D}_X, \mathfrak{D}_x \subset \mathbb{R}^m$ 均为开集^②；
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 $\mathbf{x}(X)$ 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）。

有了微分同胚，物理域中的位置就可用参数域中的位置等价地进行刻画。因此我们也把微分同胚称为曲线坐标系。

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。在映照 $\mathbf{X}(\mathbf{x})$ 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 $\exists \mathbf{DX}(\mathbf{x}_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ ^③，满足

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

则称向量值映照 $\mathbf{X}(\mathbf{x})$ 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 当然它仍然是 \mathbf{X} 的隐函数： $\mathbf{x} = \mathbf{x}(\mathbf{X})$ 。

② 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。

③ 正如之前已经定义的， \mathbf{DX} 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。

1.2.2 Jacobi 矩阵

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

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

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

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

这样就有

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

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

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

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 $\begin{bmatrix} \mathbf{DX}(\mathbf{x}_0)(\mathbf{e}_1) & \cdots & \mathbf{DX}(\mathbf{x}_0)(\mathbf{e}_m) \end{bmatrix}$ 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

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

即除了 \mathbf{h} 的第 i 各元素为 λ 外, 其余元素均为 0 ($\lambda \neq 0$). 因而有 $\|\mathbf{h}\|_{\mathbb{R}^m} = \lambda$. 代入可微性的定义 (1.19) 式, 可得

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

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:

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

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:

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

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

$$\begin{aligned} & \left[\mathbf{D}\mathbf{X}(\mathbf{x}_0)(\mathbf{e}_1), \dots, \mathbf{D}\mathbf{X}(\mathbf{x}_0)(\mathbf{e}_i), \dots, \mathbf{D}\mathbf{X}(\mathbf{x}_0)(\mathbf{e}_m) \right] \\ &= \left[\frac{\partial \mathbf{X}}{\partial x^1}(\mathbf{x}_0), \dots, \frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}_0), \dots, \frac{\partial \mathbf{X}}{\partial x^m}(\mathbf{x}_0) \right] \in \mathbb{R}^{m \times m}. \end{aligned} \quad (1.28)$$

下面给出 $\partial \mathbf{X} / \partial x^i (\mathbf{x}_0)$ 的计算式. 根据定义, 有

$$\begin{aligned} \frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}_0) &:= \lim_{\lambda \rightarrow 0} \frac{\mathbf{X}(\mathbf{x}_0 + \lambda \mathbf{e}_i) - \mathbf{X}(\mathbf{x}_0)}{\lambda} \in \mathbb{R}^m \\ &= \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \cdot \left(\begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix}(\mathbf{x}_0 + \lambda \mathbf{e}_i) - \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix}(\mathbf{x}_0) \right) \\ &= \lim_{\lambda \rightarrow 0} \begin{bmatrix} \frac{X^1(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X^1(\mathbf{x}_0)}{\lambda} \\ \vdots \\ \frac{X^m(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X^m(\mathbf{x}_0)}{\lambda} \end{bmatrix}. \end{aligned} \quad (1.29)$$

向量极限存在的充要条件是各分量极限均存在, 即存在

$$\frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}_0) := \lim_{\lambda \rightarrow 0} \frac{X^\alpha(\mathbf{x}_0 + \lambda \mathbf{e}_i) - X^\alpha(\mathbf{x}_0)}{\lambda} \in \mathbb{R}, \quad (1.30)$$

其中的 $\alpha = 1, \dots, m$. 这其实就是我们熟知的多元函数偏导数的定义. 用它来表示向量值映照的偏导数, 可有

$$\frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial X^1}{\partial x^i}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial X^m}{\partial x^i}(\mathbf{x}_0) \end{bmatrix} = \sum_{\alpha=1}^m \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}_0) \mathbf{e}_\alpha. \quad (1.31)$$

向量值映照 \mathbf{X} 关于 x^i 的偏导数, 从代数的角度来看, 是 Jacobi 矩阵的第 i 列; 从几何的角度来看, 则是物理域中 x^i 线的切向量; 从计算的角度来看, 又是 (该映照) 每个分量偏导数的组合.

现在我们重新回到 Jacobi 矩阵. 情况已经十分明了: 只需把之前获得的各列并起来, 就可以得到完整的 Jacobi 矩阵. 于是

$$\begin{aligned} \mathbf{D}\mathbf{X}(\mathbf{x}_0)(\mathbf{h}) &= \left[\frac{\partial \mathbf{X}}{\partial x^1}, \dots, \frac{\partial \mathbf{X}}{\partial x^m} \right](\mathbf{x}_0)(\mathbf{h}) \\ &= \begin{bmatrix} \frac{\partial X^1}{\partial x^1} & \dots & \frac{\partial X^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial X^m}{\partial x^1} & \dots & \frac{\partial X^m}{\partial x^m} \end{bmatrix}(\mathbf{x}_0) \cdot \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix}. \end{aligned} \quad (1.32)$$

这与 1.1.2 小节中 (1.11) 式给出的定义是完全一致的.

1.2.3 偏导数的几何意义

这一小节中, 我们要回过头来, 澄清向量值映照偏导数的几何意义.

如图 1.2, $\mathbf{X}(\mathbf{x})$ 是定义域空间 $\mathfrak{D}_{\mathbf{x}} \subset \mathbb{R}^m$ 到值域空间 $\mathfrak{D}_{\mathbf{X}} \subset \mathbb{R}^m$ 的向量值映照. 在定义域空间 $\mathfrak{D}_{\mathbf{x}}$ 中, 过点 \mathbf{x}_0 作一条平行于 x^i 轴的直线, 称为 x^i -线. x^i 轴定义了向量 \mathbf{e}_i , 因而 x^i -线上的任意一点均可表示为 $\mathbf{x}_0 + \lambda \mathbf{e}_i$, 其中 $\lambda \in \mathbb{R}$.



图 1.2: 向量值映照偏导数的几何意义

在 $\mathbf{X}(\mathbf{x})$ 的作用下, 点 \mathbf{x}_0 被映照到 $\mathbf{X}(\mathbf{x}_0)$, 而 $\mathbf{x}_0 + \lambda \mathbf{e}_i$ 则被映照到了 $\mathbf{X}(\mathbf{x}_0 + \lambda \mathbf{e}_i)$. 这样一来, x^i -线也就被映照到了值域空间 $\mathfrak{D}_{\mathbf{X}}$ 中, 成为一条曲线.

根据前面的定义, 当 $\lambda \rightarrow 0$ 时,

$$\frac{\mathbf{X}(\mathbf{x}_0 + \lambda \mathbf{e}_i) - \mathbf{X}(\mathbf{x}_0)}{\lambda} \rightarrow \frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}_0). \quad (1.33)$$

对应到图 1.2 中, 就是 x^i -线 (值域空间中) 在 $\mathbf{X}(\mathbf{x}_0)$ 处的切向量.

完全类似, 在定义域空间 $\mathfrak{D}_{\mathbf{x}}$ 中, 过点 \mathbf{x}_0 作出 x^j -线 (自然是平行于 x^j 轴), 其上的点可以表示为 $\mathbf{x}_0 + \lambda \mathbf{e}_j$. 映射到值域空间 $\mathfrak{D}_{\mathbf{X}}$ 上, 则成为 $\mathbf{X}(\mathbf{x}_0 + \lambda \mathbf{e}_j)$. 很显然,

$$\frac{\partial \mathbf{X}}{\partial x^j}(\mathbf{x}_0) = \frac{\mathbf{X}(\mathbf{x}_0 + \lambda \mathbf{e}_j) - \mathbf{X}(\mathbf{x}_0)}{\lambda} \quad (1.34)$$

就是 x^j -线在 $\mathbf{X}(\mathbf{x}_0)$ 处的切向量. 在定义域空间中, x^i -线作为直线共有 m 条, 它们之间互相垂直. 作用到值域空间后, 这样的 x^i -线尽管变为了曲线, 但仍为 m 条. 相应的切向量, 自然也有 m 个.

1.3 局部基

这里的讨论基于曲线坐标系 (即微分同胚) $\mathbf{X}(\mathbf{x}) \in \mathcal{C}^p(\mathfrak{D}_{\mathbf{x}}; \mathfrak{D}_{\mathbf{X}})$.

1.3.1 局部协变基

我们已经知道， $\mathbf{X}(\mathbf{x})$ 的 Jacobi 矩阵可以表示为

$$\mathbf{D}\mathbf{X}(\mathbf{x}) = \left[\frac{\partial \mathbf{X}}{\partial x^1}, \dots, \frac{\partial \mathbf{X}}{\partial x^i}, \dots, \frac{\partial \mathbf{X}}{\partial x^m} \right](\mathbf{x}) \in \mathbb{R}^{m \times m}, \quad (1.35)$$

式中的

$$\frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}) = \lim_{\lambda \rightarrow 0} \frac{\mathbf{X}(\mathbf{x} + \lambda \mathbf{e}_i) - \mathbf{X}(\mathbf{x})}{\lambda}. \quad (1.36)$$

在参数域 \mathfrak{D}_x 中作出 x^i -线. 映照到物理域后, 它变成一条曲线, 我们仍称之为 x^i -线. 1.2.3 小节已经说明, (1.36) 式表示物理域中 x^i -线的切向量. 在张量分析中, 我们通常把它记作 $\mathbf{g}_i(\mathbf{x})$.

由于微分同胚要求是双射, 因而 Jacobi 矩阵

$$\mathbf{D}\mathbf{X}(\mathbf{x}) = \left[\mathbf{g}_1, \dots, \mathbf{g}_i, \dots, \mathbf{g}_m \right](\mathbf{x}) \in \mathbb{R}^{m \times m} \quad (1.37)$$

必须是非奇异的. 这等价于

$$\left\{ \mathbf{g}_i(\mathbf{x}) = \frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}) \right\}_{i=1}^m \subset \mathbb{R}^m \quad (1.38)$$

线性无关. 由此, 它们可以构成 \mathbb{R}^m 上的一组基.

用任意的 $\mathbf{x} \in \mathfrak{D}_x$ 均可构建一组基. 但选取不同的 \mathbf{x} , 将会使所得基的取向有所不同. 因而这种基称为局部协变基. 和之前一样, 我们用“协变”表示指标在下方.

1.3.2 局部逆变基; 对偶关系

有了局部协变基 $\{\mathbf{g}_i(\mathbf{x})\}_{i=1}^m$, 根据 交叉引用第一章 (已注释掉) 小节中的讨论, 必然唯一存在与之对应的局部逆变基 $\{\mathbf{g}^i(\mathbf{x})\}_{i=1}^m$, 满足

$$\left[\mathbf{g}^1(\mathbf{x}), \dots, \mathbf{g}^m(\mathbf{x}) \right]^\top \left[\mathbf{g}_1(\mathbf{x}), \dots, \mathbf{g}_m(\mathbf{x}) \right] = \begin{bmatrix} (\mathbf{g}^1)^\top \\ \vdots \\ (\mathbf{g}^m)^\top \end{bmatrix}(\mathbf{x}) \cdot \mathbf{D}\mathbf{X}(\mathbf{x}) = \mathbf{I}_m. \quad (1.39)$$

下面我们来寻找逆变基 $\{\mathbf{g}^i(\mathbf{x})\}_{i=1}^m$ 的具体表示. 考虑到^①

$$\mathbf{X}(\mathbf{x}(\mathbf{X})) = \mathbf{X} \in \mathbb{R}^m, \quad (1.40)$$

并利用复合映照可微性定理, 可知

$$\mathbf{D}\mathbf{X}(\mathbf{x}(\mathbf{X})) \cdot \mathbf{D}\mathbf{x}(\mathbf{X}) = \mathbf{I}_m, \quad (1.41)$$

即有

$$\mathbf{D}\mathbf{x}(\mathbf{X}) = (\mathbf{D}\mathbf{X})^{-1}(\mathbf{x}(\mathbf{X})). \quad (1.42)$$

于是

$$\begin{bmatrix} (\mathbf{g}^1)^\top \\ \vdots \\ (\mathbf{g}^m)^\top \end{bmatrix}(\mathbf{x}) = (\mathbf{D}\mathbf{X})^{-1}(\mathbf{x}) = \mathbf{D}\mathbf{x}(\mathbf{X}) = \begin{bmatrix} \frac{\partial x^1}{\partial X^1} & \dots & \frac{\partial x^1}{\partial X^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^m}{\partial X^1} & \dots & \frac{\partial x^m}{\partial X^m} \end{bmatrix}(\mathbf{X}). \quad (1.43)$$

^① 这里的几步推导在 1.1.2 小节中也有所涉及.

这样我们就得到了局部逆变基的具体表示（注意转置）：

$$\mathbf{g}^i(\mathbf{x}) = \begin{bmatrix} \frac{\partial x^i}{\partial X^1} \\ \vdots \\ \frac{\partial x^i}{\partial X^m} \end{bmatrix}(\mathbf{X}) = \sum_{\alpha=1}^m \frac{\partial x^i}{\partial X^\alpha}(\mathbf{X}) \mathbf{e}_\alpha. \quad (1.44)$$

定义标量场 $f(\mathbf{x})$ 的梯度为

$$\nabla f(\mathbf{x}) \triangleq \sum_{\alpha=1}^m \frac{\partial f}{\partial x^\alpha}(\mathbf{x}) \mathbf{e}_\alpha, \quad (1.45)$$

则局部逆变基又可以表示成

$$\mathbf{g}^i(\mathbf{x}) = \nabla x^i(\mathbf{X}). \quad (1.46)$$

此处的梯度实际上就是我们熟知的三维情况在 m 维下的推广。

在 1.2.3 小节中已经指出，局部协变基的几何意义是 x^i -线的切向量。现在，我们来讨论局部逆变基的几何意义。



图 1.3: 局部逆变基的几何意义

如图 1.3 所示，在参数空间中，过点 \mathbf{x} 作垂直于 x^i 轴的平面，记为 x^i -面。在 x^i -面上，自然有 $x^i = \text{const}$ 。映照到物理空间后， x^i -面变为一个曲面，其上仍有 $x^i(\mathbf{X}) = \text{const}$ ，即它是一个等值面。等值面的梯度方向显然与该曲面的法向相同。因此，局部逆变基 $\mathbf{g}^i(\mathbf{x})$ 的几何意义就是 x^i -面的法向量。

现在来验证一下对偶关系。

$$(\mathbf{g}_i(\mathbf{x}), \mathbf{g}^j(\mathbf{x}))_{\mathbb{R}^m} = \left(\frac{\partial X}{\partial x^i}(\mathbf{x}), \nabla x^j(\mathbf{X}) \right)_{\mathbb{R}^m}$$

$$= \left(\sum_{\alpha=1}^m \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}) \mathbf{e}_\alpha, \sum_{\beta=1}^m \frac{\partial x^j}{\partial X^\beta}(\mathbf{X}) \mathbf{e}_\beta \right)_{\mathbb{R}^m}$$

利用内积的线性性，有

$$\begin{aligned} &= \sum_{\alpha=1}^m \sum_{\beta=1}^m \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}) \frac{\partial x^j}{\partial X^\beta}(\mathbf{X}) \cdot (\mathbf{e}_\alpha, \mathbf{e}_\beta)_{\mathbb{R}^m} \\ &= \sum_{\alpha=1}^m \sum_{\beta=1}^m \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}) \frac{\partial x^j}{\partial X^\beta}(\mathbf{X}) \cdot \delta_{\alpha\beta} \end{aligned}$$

合并掉指标 β ，可得

$$\begin{aligned} &= \sum_{\alpha=1}^m \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}) \frac{\partial x^j}{\partial X^\alpha}(\mathbf{X}) \\ &= \sum_{\alpha=1}^m \frac{\partial x^j}{\partial X^\alpha}(\mathbf{X}) \frac{\partial X^\alpha}{\partial x^i}(\mathbf{x}). \end{aligned} \quad (1.47)$$

最后一步求和号中的第一项位于 Jacobi 矩阵 $\mathbf{D}\mathbf{x}(\mathbf{X})$ 的第 j 行第 α 列，而第二项位于 $\mathbf{D}\mathbf{X}(\mathbf{x})$ 的第 α 行第 i 列，因此关于 α 的求和结果便是乘积矩阵的第 j 行第 i 列。根据式 (1.41)，这两个 Jacobi 矩阵的乘积为单位阵，所以有

$$(\mathbf{g}_i(\mathbf{x}), \mathbf{g}^j(\mathbf{x}))_{\mathbb{R}^m} = \delta_i^j. \quad (1.48)$$

总结一下我们得到的结果。对于体积形态的连续介质，存在着

$$\begin{cases} \text{局部协变基:} & \left\{ \mathbf{g}_i(\mathbf{x}) \triangleq \frac{\partial \mathbf{X}}{\partial x^i}(\mathbf{x}) \right\}_{i=1}^m, \\ \text{局部逆变基:} & \left\{ \mathbf{g}^i(\mathbf{x}) \triangleq \nabla x^i(\mathbf{X}) \right\}_{i=1}^m, \end{cases}$$

它们满足对偶关系

$$(\mathbf{g}_i(\mathbf{x}), \mathbf{g}^j(\mathbf{x}))_{\mathbb{R}^m} = \delta_i^j. \quad (1.50)$$

这样，在研究连续介质中的一个点时，我们就有三种基可以使用：局部协变基、局部逆变基，当然还有典则基 $\{\mathbf{e}_i\}_{i=1}^m$ 。

1.4 标架运动方程

1.4.1 向量在局部基下的表示

对于 \mathbb{R}^m 空间中的任意一个向量 \mathbf{b} ，它可以用典则基表示：

$$\mathbf{b} = \sum_{\alpha=1}^m b_\alpha \mathbf{e}_\alpha = b_\alpha \mathbf{e}_\alpha. \quad (1.51)$$

第二步省略掉了求和号，这是根据 **Einstein 求和约定**：指标出现两次，则表示对它求和。^① 根据之前一小节的结论， \mathbf{b} 还可以用局部协变基和局部逆变基来表示：

$$\mathbf{b} = b^i \mathbf{g}_i = b_j \mathbf{g}^j, \quad (1.52)$$

^① 在交叉引用第一章（已注释掉）小节中，还要求重复指标一上一下。典则基不分协变、逆变，标号均在下方，可以视为一个特例。

式中,

$$b^i = (\mathbf{b}, \mathbf{g}^i)_{\mathbb{R}^m} \quad (1.53-a)$$

和

$$b_j = (\mathbf{b}, \mathbf{g}_j)_{\mathbb{R}^m} \quad (1.53-b)$$

分别称为向量 \mathbf{b} 的**逆变分量**和**协变分量**. 注意, 这里同样用到了 Einstein 求和约定.

将 $\mathbf{b} = b^i \mathbf{g}_i$ 的两边分别与 \mathbf{g}^j 作内积, 可有

$$(\mathbf{b}, \mathbf{g}^j)_{\mathbb{R}^m} = (b^i \mathbf{g}_i, \mathbf{g}^j)_{\mathbb{R}^m}$$

利用内积的线性性, 提出系数:

$$= b^i (\mathbf{g}_i, \mathbf{g}^j)_{\mathbb{R}^m}$$

利用对偶关系 (1.50) 式, 可有

$$= b^i \delta_i^j = b^j, \quad (1.54)$$

这就得到了逆变分量的表示式 (1.53-a). 同理, 将 $\mathbf{b} = b_j \mathbf{g}^j$ 的两边分别与 \mathbf{g}_i 作内积, 就有

$$(\mathbf{b}, \mathbf{g}_i)_{\mathbb{R}^m} = (b_j \mathbf{g}^j, \mathbf{g}_i)_{\mathbb{R}^m} = b_j (\mathbf{g}^j, \mathbf{g}_i)_{\mathbb{R}^m} = b_j \delta_i^j = b_i, \quad (1.55)$$

这便是协变分量的表示 (1.53-b) 式.

1.4.2 局部基的偏导数

所谓局部基 (或曰“活动标架”), 顾名思义, 它在不同的点上往往是不同的. 根据之前的定义, 我们有

$$\mathbf{g}_i(\mathbf{x}) : \mathfrak{D}_x \ni \mathbf{x} \mapsto \mathbf{g}_i(\mathbf{x}) = \begin{bmatrix} \frac{\partial X^1}{\partial x^i} \\ \vdots \\ \frac{\partial X^m}{\partial x^i} \end{bmatrix} (\mathbf{x}) \in \mathbb{R}^m, \quad (1.56-a)$$

$$\mathbf{g}^i(\mathbf{x}) : \mathfrak{D}_x \ni \mathbf{x} \mapsto \mathbf{g}^i(\mathbf{x}) = \begin{bmatrix} \frac{\partial x^i}{\partial X^1} \\ \vdots \\ \frac{\partial x^i}{\partial X^m} \end{bmatrix} (\mathbf{X}(\mathbf{x})) \in \mathbb{R}^m. \quad (1.56-b)$$

从映照的角度来看, 局部基定义了新的向量值映照, 其定义域仍为参数域, 而值域则为 \mathbb{R}^m 空间. 这样一来, 我们在 1.2 节中所引入的操作均可完全类似地应用在局部基上. 例如, 我们可以来求局部基的 Jacobi 矩阵:

$$\begin{cases} D\mathbf{g}_i(\mathbf{x}) = \left[\frac{\partial \mathbf{g}_i}{\partial x^1}, \dots, \frac{\partial \mathbf{g}_i}{\partial x^m} \right] (\mathbf{x}) \in \mathbb{R}^{m \times m}, \end{cases} \quad (1.57-a)$$

$$\begin{cases} D\mathbf{g}^i(\mathbf{x}) = \left[\frac{\partial \mathbf{g}^i}{\partial x^1}, \dots, \frac{\partial \mathbf{g}^i}{\partial x^m} \right] (\mathbf{x}) \in \mathbb{R}^{m \times m}. \end{cases} \quad (1.57-b)$$

Jacobi 矩阵中的每一列都是局部基作为整体相对自变量第 j 个分量的变化率，即偏导数：

$$\left\{ \frac{\partial \mathbf{g}_i}{\partial x^j}(\mathbf{x}) \triangleq \lim_{\lambda \rightarrow 0} \frac{\mathbf{g}_i(\mathbf{x} + \lambda \mathbf{e}_j) - \mathbf{g}_i(\mathbf{x})}{\lambda} \in \mathbb{R}^m, \right. \quad (1.58-a)$$

$$\left. \frac{\partial \mathbf{g}^i}{\partial x^j}(\mathbf{x}) \triangleq \lim_{\lambda \rightarrow 0} \frac{\mathbf{g}^i(\mathbf{x} + \lambda \mathbf{e}_j) - \mathbf{g}^i(\mathbf{x})}{\lambda} \in \mathbb{R}^m. \right. \quad (1.58-b)$$

下面澄清局部基偏导数的几何意义。如图 1.4 所示，在参数空间中，过点 \mathbf{x} 作出 x^j -线，并在其上取点 $\mathbf{x} + \lambda \mathbf{e}_j$ 。分别过点 \mathbf{x} 和 $\mathbf{x} + \lambda \mathbf{e}_j$ 作出 x^i -线，于是 $\partial \mathbf{g}_i / \partial x^j(\mathbf{x})$ 就表示 $\mathbf{g}_i(\mathbf{x})$ （即 x^i -线的切向量）沿 x^j -线的变化率。同理，过点 \mathbf{x} 和 $\mathbf{x} + \lambda \mathbf{e}_j$ 作出 x^i -面，则 $\partial \mathbf{g}^i / \partial x^j(\mathbf{x})$ 就表示 $\mathbf{g}^i(\mathbf{x})$ （即 x^i -面的法向量）沿 x^j -线的变化率。

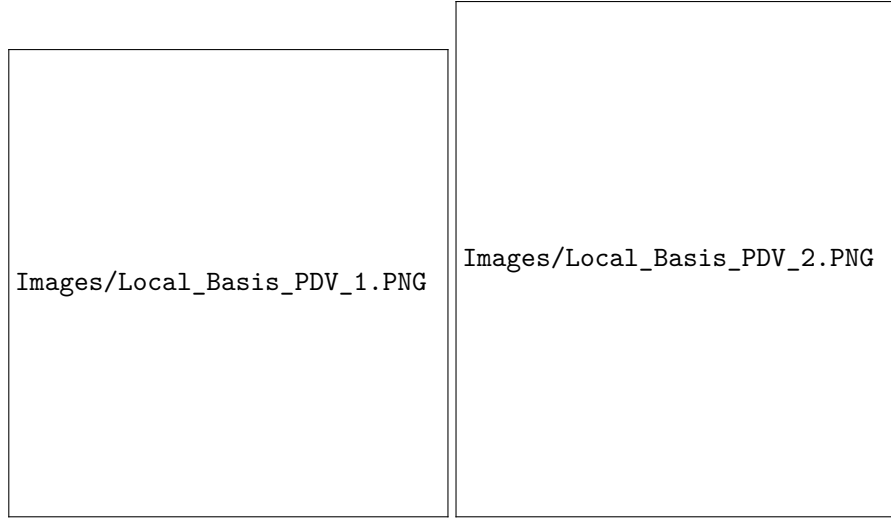


图 1.4: 局部基偏导数的几何意义

1.4.3 Christoffel 符号

考察 $\partial \mathbf{g}_i / \partial x^j(\mathbf{x})$ ，它是 \mathbb{R}^m 空间中的一个向量。于是，它可以用（局部）协变基或逆变基来表示：

$$\frac{\partial \mathbf{g}_i}{\partial x^j}(\mathbf{x}) = \begin{cases} \left(\frac{\partial \mathbf{g}_i}{\partial x^j}, \mathbf{g}^k \right)_{\mathbb{R}^m} \mathbf{g}_k, \\ \left(\frac{\partial \mathbf{g}_i}{\partial x^j}, \mathbf{g}_k \right)_{\mathbb{R}^m} \mathbf{g}^k. \end{cases} \quad (1.59)$$

引入第一类 Christoffel 符号

$$\Gamma_{ji,k} := \left(\frac{\partial \mathbf{g}_i}{\partial x^j}, \mathbf{g}_k \right)_{\mathbb{R}^m} \quad (1.60)$$

和第二类 Christoffel 符号

$$\Gamma^k_{ji} := \left(\frac{\partial \mathbf{g}_i}{\partial x^j}, \mathbf{g}^k \right)_{\mathbb{R}^m} \quad (1.61)$$

则式 (1.59) 可以写成

$$\frac{\partial \mathbf{g}_i}{\partial x^j}(\mathbf{x}) = \begin{cases} \Gamma^k_{ji} \mathbf{g}_k, \\ \Gamma_{ji,k} \mathbf{g}^k. \end{cases} \quad (1.62-a)$$

$$(1.62-b)$$

引用 1: 式 (1.62) 引用 2: 式 (1.62-a) 引用 3: 式 (1.62-b)