高维微分学——向量值映照的可微性

复旦力学　谢锡麟

2016 年 3 月 15 日

1　知识要素

1.1　向量值映照的可微性定义

可微性作为映照的一种 “局部行为”, 指自变量变化所引起的因变量的变化可由自变量空间至因变量空间之间的线性映照逼近, 误差为一阶无穷小量，如图1所示.

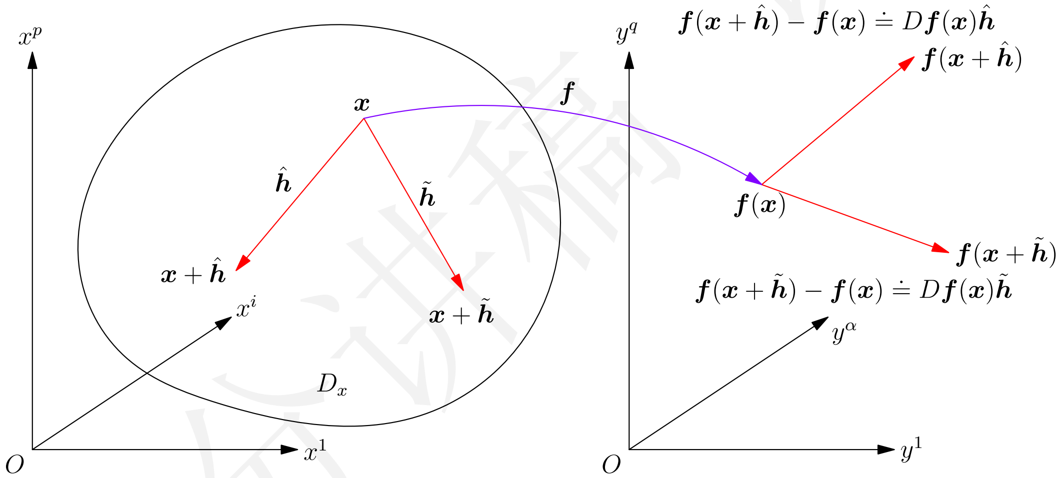

图 1: 向量值映照可微性示意

定义 1.1 (向量值映照可微性)**.** 对一般向量值映照

$$\boldsymbol{f}(\boldsymbol{x}) : \mathbb{R}^m \supset \mathscr{D}_x \ni \boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x}) \in \mathbb{R}^n,$$

$\boldsymbol{x}_0 \in int\mathscr{D}_x$①, 有

$$\boldsymbol{f}(\boldsymbol{x}_0+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{h})+\boldsymbol{o}(|\boldsymbol{h}|_{\mathbb{R}^m}) \in \mathbb{R}^n, \quad \boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}_0) \in \mathscr{L}(\mathbb{R}^m;\mathbb{R}^n)$$

则称 $\boldsymbol{f}(x)$ 在 $\boldsymbol{x}_0$ 点可微, 可称 $\boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{h}) \in \mathbb{R}^n$ 为 $\boldsymbol{f}(\boldsymbol{x})$ 在 $\boldsymbol{x}_0 \in \mathbb{R}^m$ 点的**微分**.

① 表示 $\boldsymbol{x}_0$ 为定义域 $\mathscr{D}_x$ 的内点

微积分讲稿 谢锡麟

以下研究微分 $D\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{h})$ 的表达式. 由 $D\boldsymbol{f}(\boldsymbol{x}_0) \in \mathscr{L}(\mathbb{R}^m;\mathbb{R}^n)$ 可有

$$\begin{aligned} D\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{h}) &= \boldsymbol{Df}(\boldsymbol{x}_0)(h^1\boldsymbol{i}_1 + \cdots + h^m\boldsymbol{i}_m) = \sum_{i=1}^{m} \boldsymbol{Df}(\boldsymbol{x}_0)(\boldsymbol{i}_i)h^i \\ &= [\boldsymbol{Df}(\boldsymbol{x}_0)(\boldsymbol{i}_1), \cdots, \boldsymbol{Df}(\boldsymbol{x}_0)(\boldsymbol{i}_m)] \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix}. \end{aligned}$$

按可微性定义, 特取 $\boldsymbol{h} = \lambda \boldsymbol{i}_i \in \mathbb{R}^m$, 有

$$\boldsymbol{f}(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - \boldsymbol{f}(\boldsymbol{x}_0) = \lambda\boldsymbol{Df}(\boldsymbol{x}_0)(\boldsymbol{i}_i) + \boldsymbol{o}(\lambda) \in \mathbb{R}^n.$$

亦即, 有

$$\exists \lim_{\lambda\to 0\in\mathbb{R}} \frac{\boldsymbol{f}(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - \boldsymbol{f}(\boldsymbol{x}_0)}{\lambda} = \boldsymbol{Df}(\boldsymbol{x}_0)(\boldsymbol{i}_i) \in \mathbb{R}^n.$$

再由存在向量值映照极限等价于存在各分量的极限, 则有

$$\lim_{\lambda\to 0\in\mathbb{R}} \frac{\boldsymbol{f}(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - \boldsymbol{f}(\boldsymbol{x}_0)}{\lambda} = \begin{bmatrix} \lim\limits_{\lambda\to 0\in\mathbb{R}} \dfrac{f^1(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - f^1(\boldsymbol{x}_0)}{\lambda} \\ \vdots \\ \lim\limits_{\lambda\to 0\in\mathbb{R}} \dfrac{f^\alpha(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - f^\alpha(\boldsymbol{x}_0)}{\lambda} \\ \vdots \\ \lim\limits_{\lambda\to 0\in\mathbb{R}} \dfrac{f^n(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - f^n(\boldsymbol{x}_0)}{\lambda} \end{bmatrix}$$

定义向量值映照 $\boldsymbol{f}(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 点相对于自变量第 i 个分量 x^i 的变化率

$$\frac{\partial \boldsymbol{f}}{\partial x^i}(\boldsymbol{x}_0) \triangleq \lim_{\lambda\to 0\in\mathbb{R}} \frac{\boldsymbol{f}(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - \boldsymbol{f}(\boldsymbol{x}_0)}{\lambda} \in \mathbb{R}^n,$$

以及 $\boldsymbol{f}(\boldsymbol{x})$ 的第 α 个分量 $f^\alpha(\boldsymbol{x})$ 相对于 x^i 的变化率

$$\frac{\partial f^\alpha}{\partial x^i}(\boldsymbol{x}_0) \triangleq \lim_{\lambda\to 0\in\mathbb{R}} \frac{f^\alpha(\boldsymbol{x}_0 + \lambda\boldsymbol{i}_i) - f^\alpha(\boldsymbol{x}_0)}{\lambda} \in \mathbb{R},$$

即有

$$\frac{\partial \boldsymbol{f}}{\partial x^i}(\boldsymbol{x}_0) = \begin{bmatrix} \dfrac{\partial f^1}{\partial x^i}(\boldsymbol{x}_0) \\ \vdots \\ \dfrac{\partial f^\alpha}{\partial x^i}(\boldsymbol{x}_0) \\ \vdots \\ \dfrac{\partial f^n}{\partial x^i}(\boldsymbol{x}_0) \end{bmatrix} \in \mathbb{R}^n.$$

综上所述，有

$$\boldsymbol{Df}(\boldsymbol{x}_0) = \left[\frac{\partial \boldsymbol{f}}{\partial x^1}, \cdots, \frac{\partial \boldsymbol{f}}{\partial x^m}\right](\boldsymbol{x}_0) = \begin{bmatrix} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f^n}{\partial x^1} & \cdots & \dfrac{\partial f^n}{\partial x^m} \end{bmatrix}(\boldsymbol{x}_0)$$

微积分讲稿 谢锡麟

矩阵 $\boldsymbol{Df}(\boldsymbol{x}_0) \in \mathbb{R}^{n\times m}$ 称为 **Jacobi 矩阵**。

可定义在向量值映照在某点沿某方向的变化率

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{e}}(\boldsymbol{x}_0) \triangleq \lim_{\lambda\to 0\in\mathbb{R}} \frac{\boldsymbol{f}(\boldsymbol{x}_0+\lambda\boldsymbol{e})-\boldsymbol{f}(\boldsymbol{x}_0)}{\lambda} \in \mathbb{R}^n, \quad \forall\, |\boldsymbol{e}|_{\mathbb{R}^m}=1.$$

按可微性的定义，易见当 $\boldsymbol{f}(\boldsymbol{x})$ 在 $\boldsymbol{x}_0 \in \mathbb{R}^m$ 点可微, 有

$$\exists\, \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{e}}(\boldsymbol{x}_0) = \boldsymbol{Df}(\boldsymbol{x}_0)\,\boldsymbol{e}, \quad \forall\, |\boldsymbol{e}|_{\mathbb{R}^m}=1$$

1.2 复合向量值映照的可微性定理

定理 1.1 (复合向量值映照的可微性定理)**.** 向量值映射

$$\boldsymbol{\theta}(\boldsymbol{x}):\, \mathbb{R}^m \supset \mathscr{D}_\theta \ni \boldsymbol{x} \,\mapsto\, \boldsymbol{\theta}(\boldsymbol{x}) \in \mathbb{R}^n$$

在 $\boldsymbol{x}_0 \in int\mathscr{D}_\theta \subset \mathbb{R}^m$ 点可微; 向量值映射

$$\boldsymbol{\Theta}(\boldsymbol{y}):\, \mathbb{R}^n \supset \mathscr{D}_\Theta \ni \boldsymbol{y} \,\mapsto\, \boldsymbol{\Theta}(\boldsymbol{y}) \in \mathbb{R}^l$$

在 $\boldsymbol{y}_0 = \boldsymbol{\theta}(\boldsymbol{x}_0) \in\ int\mathscr{D}_\Theta \subset \mathbb{R}^n$ 点可微. 则有

1. 在 $\boldsymbol{x}_0$ 点局部存在向量值的复合 $\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x})$;
2. $\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 点可微，且有

$$\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}_0+\triangle\boldsymbol{x}) = \boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}_0) + \boldsymbol{D\Theta}(\boldsymbol{\theta}(\boldsymbol{x}_0))\cdot\boldsymbol{D\theta}(\boldsymbol{x}_0)\triangle\boldsymbol{x} + o(|\triangle\boldsymbol{x}|_{\mathbb{R}^m})$$

证明 证明复合向量值映照的局部存在性。由于，考虑到 $\boldsymbol{x_0} \in int\mathscr{D}_\theta$ 和 $\boldsymbol{y_0} \in int\mathscr{D}_\Theta$，以及可微性保证连续性，则有：

$$\exists\lambda,\mu \in \mathbb{R}^+,\, s.t.\, \boldsymbol{\theta}(B_\lambda(\boldsymbol{x_0})) \subset B_\mu(\boldsymbol{y_0})$$

且 $B_\lambda(\boldsymbol{x_0}) \subset \mathscr{D}_\theta$ 和 $B_\mu(\boldsymbol{y_0}) \subset \mathscr{D}_\Theta$。由此可构造：

$$\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}):\, B_\lambda(\boldsymbol{x_0}) \ni \boldsymbol{x} \,\mapsto\, \boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}) \equiv \boldsymbol{\Theta}(\boldsymbol{\theta}(\boldsymbol{x})) \in \mathbb{R}^l$$

证明可微性。基于 $\boldsymbol{\Theta}(\boldsymbol{y}) \in \mathbb{R}^l$ 在 $\boldsymbol{y_0} = \boldsymbol{\theta}(\boldsymbol{x_0}) \in \mathbb{R}^n$ 点的可微性，即有：

$$\boldsymbol{\Theta}(\boldsymbol{y_0}+\triangle\boldsymbol{y}) = \boldsymbol{\Theta}(\boldsymbol{y_0}) + \boldsymbol{D\Theta}(\boldsymbol{y_0})\triangle\boldsymbol{y} + o(|\triangle\boldsymbol{y}|_{\mathbb{R}^n})$$

引入

$$\Phi(\triangle\boldsymbol{y}):\, \mathbb{R}^n \supset B_\mu(0) \ni \triangle\boldsymbol{y} \,\mapsto\, \Phi(\triangle\boldsymbol{y}) \triangleq \begin{cases} \frac{o(|\triangle\boldsymbol{y}|_{\mathbb{R}^n})}{|\triangle\boldsymbol{y}|_{\mathbb{R}^n}} & \triangle\boldsymbol{y} \neq \boldsymbol{0} \in \mathbb{R}^n \\ 0 & \triangle\boldsymbol{y} = \boldsymbol{0} \in \mathbb{R}^n \end{cases} \in \mathbb{R}^l$$

则有

$$\boldsymbol{\Theta}(y_0+\triangle\boldsymbol{y}) = \boldsymbol{\Theta}(\boldsymbol{y_0}) + \boldsymbol{D\Theta}(\boldsymbol{y_0})\triangle\boldsymbol{y} + \boldsymbol{\Phi}(\triangle\boldsymbol{y})\cdot|\triangle\boldsymbol{y}|_{\mathbb{R}^n}, \quad ,\forall\triangle\boldsymbol{y} \in B_\mu(0) \subset \mathbb{R}^n$$

微积分讲稿 谢锡麟

基于 $\boldsymbol{\theta}(\boldsymbol{x}) \in \mathbb{R}^n$ 在 $\boldsymbol{x_0} \in \mathbb{R}^m$ 点的可微性，取

$$\triangle \boldsymbol{y} := \boldsymbol{\theta}(\boldsymbol{x_0} + \triangle \boldsymbol{x}) - \boldsymbol{\theta}(x_0) = \boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m}), \quad \forall \triangle \boldsymbol{x} \in \mathring{B}_\lambda(0) \subset \mathbb{R}^m$$

带入上式，则有

$$\begin{aligned}\boldsymbol{\Theta} \circ \boldsymbol{\theta}(\boldsymbol{x_0} + \triangle \boldsymbol{x}) &= \boldsymbol{\Theta} \circ \boldsymbol{\theta}(\boldsymbol{x_0}) + \boldsymbol{D\Theta}(\boldsymbol{y_0}) \cdot [\boldsymbol{D\theta}(\boldsymbol{x_0})\triangle \boldsymbol{x} + o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})] \\ &\quad + \boldsymbol{\Phi}(\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})) \cdot |\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n}, \quad \forall \triangle \boldsymbol{x} \in \mathring{B}_\lambda(0) \subset \mathbb{R}^m\end{aligned}$$

考虑到：

$$\lim_{\triangle \boldsymbol{x} \to 0 \in \mathbb{R}^m} \frac{\boldsymbol{D\Theta}(y_0) \cdot o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})}{|\triangle x|_{\mathbb{R}^m}} \in \mathbb{R}^l$$

$$\Leftrightarrow$$

$$\lim_{\triangle \boldsymbol{x} \to 0 \in \mathbb{R}^m} \frac{1}{|\triangle x|_{\mathbb{R}^m}} \cdot \begin{bmatrix} \frac{\partial \Theta^\alpha}{\partial y^1}, \cdots, \frac{\partial \Theta^\alpha}{\partial y^n} \end{bmatrix} (\boldsymbol{y_0}) \cdot \begin{bmatrix} o^1(|\triangle x|_{\mathbb{R}^m}) \\ \vdots \\ o^n(|\triangle x|_{\mathbb{R}^m}) \end{bmatrix} = 0 \in \mathbb{R}, \quad \alpha = 1, \cdots, n$$

即有：

$$\boldsymbol{D\Theta}(\boldsymbol{y_0}) \cdot o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m}) = \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m}) \in \mathbb{R}^l$$

考虑到：

$$\lim_{\triangle \boldsymbol{x} \to \boldsymbol{0} \in \mathbb{R}^m} [\boldsymbol{D\theta}(\boldsymbol{x_0})\triangle \boldsymbol{x} + o(|\triangle x|_{\mathbb{R}^m})] = \boldsymbol{0} \in \mathbb{R}^n, \quad \lim_{\triangle \boldsymbol{y} \to \boldsymbol{0} \in \mathbb{R}^n} \boldsymbol{\Phi}(\triangle \boldsymbol{y}) = 0 = \boldsymbol{\Phi}(\boldsymbol{0}) \in \mathbb{R}^l$$

以及包含性条件：

$$\boldsymbol{D\theta}(\boldsymbol{x_0})\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m}) \in B_\mu(0) \subset \mathbb{R}^n, \quad \forall \boldsymbol{x} \in \mathring{B}_\lambda(0) \subset \mathbb{R}^m$$

按复合向量值映照的极限定理有

$$\lim_{\triangle \boldsymbol{x} \to 0 \in \mathbb{R}^m} \Phi(\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle x|_{\mathbb{R}^m})) = \boldsymbol{0} \in \mathbb{R}^l$$

考虑到：

$$\begin{aligned}&\frac{|\boldsymbol{D\theta}(\boldsymbol{x}_0)\triangle \boldsymbol{x} + o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} \\ &\leq \frac{|\boldsymbol{D\theta}(\boldsymbol{x}_0)\triangle \boldsymbol{x}|_{\mathbb{R}^n}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} + \frac{|o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} \qquad \leq \frac{|D\boldsymbol{\theta}(\boldsymbol{x}_0)|_{\mathbb{R}^{n\times m}} \cdot |\triangle \boldsymbol{x}|_{\mathbb{R}^m}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} + \frac{|o(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} \\ &\leq |D\boldsymbol{\theta}(\boldsymbol{x}_0)|_{\mathbb{R}^{n\times m}} + 1 \quad \text{当：} |\triangle \boldsymbol{x}|_{\mathbb{R}^m} \ll 1\end{aligned}$$

此处 $|\boldsymbol{D\theta}(\boldsymbol{x}_0)|_{\mathbb{R}^{n\times m}} \triangleq \sqrt{\sum_{\alpha=1}^n \sum_{j=1}^m |\partial \boldsymbol{\theta}^\alpha / \partial x^j|^2}$。故有：

$$\lim_{\triangle \boldsymbol{x} \to \boldsymbol{0} \in \mathbb{R}^m} \frac{\Phi(\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})) \cdot |\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n}}{|\triangle \boldsymbol{x}|_{\mathbb{R}^m}} = \boldsymbol{0} \in \mathbb{R}^l$$

即有：

$$\boldsymbol{\Phi}(\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})) \cdot |\boldsymbol{D\theta}(x_0)\triangle \boldsymbol{x} + \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m})|_{\mathbb{R}^n} = \boldsymbol{o}(|\triangle \boldsymbol{x}|_{\mathbb{R}^m}) \in \mathbb{R}^l$$

微积分讲稿 谢锡麟

综上所述，则有：

$$\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}_0+\triangle\boldsymbol{x})=\boldsymbol{\Theta}\circ\boldsymbol{\theta}(\boldsymbol{x}_0)+\boldsymbol{D\Theta}(\boldsymbol{\theta}(\boldsymbol{x}_0))\cdot\boldsymbol{D\theta}(\boldsymbol{x}_0)\triangle\boldsymbol{x}+o(|\triangle\boldsymbol{x}|_{\mathbb{R}^m})\in\mathbb{R}^l,\quad\forall\triangle\boldsymbol{x}\in\mathring{B}_\lambda(0)\subset\mathbb{R}^m$$

□

如果 $\boldsymbol{g}$ 为向量值映射

$$\boldsymbol{g}:\mathbb{R}^n\longrightarrow\mathbb{R}^l;\quad\boldsymbol{y}=\begin{bmatrix}y^1\\\vdots\\y^n\end{bmatrix}\longmapsto\boldsymbol{g}(\boldsymbol{y})=\begin{bmatrix}g^1(y^1,\cdots,y^n)\\\vdots\\g^l(y^1,\cdots,y^n)\end{bmatrix}$$

则 $\boldsymbol{g}$ 和 $\boldsymbol{f}$ 复合成向量值映射

$$\boldsymbol{g}\circ\boldsymbol{f}:\mathbb{R}^m\longrightarrow\mathbb{R}^l;\quad\boldsymbol{x}=\begin{bmatrix}x^1\\\vdots\\x^m\end{bmatrix}\longmapsto(\boldsymbol{g}\circ\boldsymbol{f})(\boldsymbol{x})=\begin{bmatrix}(g\circ f)^1(x^1,\cdots,x^m)\\\vdots\\(g\circ f)^l(x^1,\cdots,x^m)\end{bmatrix}.$$

按复合向量值映照的可微性定量, 复合映照 $(\boldsymbol{g}\circ\boldsymbol{f})(\boldsymbol{x})$ 的 *Jacobi* 矩阵为

$$\boldsymbol{D}(\boldsymbol{g}\circ\boldsymbol{f})(\boldsymbol{x})=\boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{x}))\cdot\boldsymbol{Df}(\boldsymbol{x}),$$

亦即为

$$\begin{bmatrix}\frac{\partial(g\circ f)^1}{\partial x^1}&\cdots&\frac{\partial(g\circ f)^1}{\partial x^m}\\\vdots&\cdots&\vdots\\\frac{\partial(g\circ f)^l}{\partial x^1}&\cdots&\frac{\partial(g\circ f)^l}{\partial x^m}\end{bmatrix}^{l\times m}=\begin{bmatrix}\frac{\partial g^1}{\partial f^1}&\cdots&\frac{\partial g^1}{\partial f^n}\\\vdots&\cdots&\vdots\\\frac{\partial g^l}{\partial f^l}&\cdots&\frac{\partial g^l}{\partial f^n}\end{bmatrix}^{l\times n}\cdot\begin{bmatrix}\frac{\partial f^1}{\partial x^1}&\cdots&\frac{\partial f^1}{\partial x^m}\\\vdots&\cdots&\vdots\\\frac{\partial f^n}{\partial x^1}&\cdots&\frac{\partial f^n}{\partial x^m}\end{bmatrix}^{n\times m}.$$

对于上述向量值映射的每个分量，即为多元函数

$$(g\circ f)^p(x)=g^p(\boldsymbol{f}(\boldsymbol{x}))=g^p(f^1(x^1,\cdots,x^m),\cdots,f^n(x^1,\cdots,x^m))\quad\forall\,p=1,\cdots l,$$

向量值映射的每个分量遵循以下链式求导法则：

$$\frac{\partial(g\circ f)^p}{\partial x^k}(x)=\sum_{j=1}^n\frac{\partial g^p}{\partial f^j}(\boldsymbol{f}(\boldsymbol{x}))\cdot\frac{\partial f^j}{\partial x^k}\quad\forall\,k=1,\cdots m;\forall\,p=1,\cdots l.\tag{1}$$

如有 $\hat{\boldsymbol{f}}(\boldsymbol{x})=\boldsymbol{f}(\boldsymbol{\eta}_1(\boldsymbol{x}),\cdots,\boldsymbol{\eta}_q(\boldsymbol{x}))\in\mathbb{R}^l$ 式中 $\boldsymbol{\eta}_1(\boldsymbol{x})\in\mathbb{R}^{n_1},\cdots,\boldsymbol{\eta}_q(\boldsymbol{x})\in\mathbb{R}^{n_q};\boldsymbol{x}\in\mathbb{R}^m$，则有

$$\boldsymbol{D}\hat{\boldsymbol{f}}(\boldsymbol{x})=\boldsymbol{D}_{\boldsymbol{\eta}_1}\boldsymbol{f}(\boldsymbol{\eta}_1(\boldsymbol{x}),\cdots,\boldsymbol{\eta}_q(\boldsymbol{x}))\boldsymbol{D\eta}_1(\boldsymbol{x})+\cdots+\boldsymbol{D}_{\boldsymbol{\eta}_q}\boldsymbol{f}(\boldsymbol{\eta}_1(\boldsymbol{x}),\cdots,\boldsymbol{\eta}_q(\boldsymbol{x}))\boldsymbol{D\eta}_q(\boldsymbol{x})\in\mathbb{R}^{l\times(n_1+\cdots+n_q)}$$

分析：

$$\hat{\boldsymbol{f}}(\boldsymbol{x}):\mathbb{R}^m\ni\boldsymbol{x}\mapsto\hat{\boldsymbol{f}}(\boldsymbol{x})\triangleq\boldsymbol{f}(\boldsymbol{\eta}_1(\boldsymbol{x}),\cdots,\boldsymbol{\eta}_q(\boldsymbol{x}))\equiv\boldsymbol{f}\circ\boldsymbol{\eta}(\boldsymbol{x})\in\mathbb{R}^l$$

此处

$$\mathbb{R}^m\ni\boldsymbol{x}\mapsto\boldsymbol{\eta}(\boldsymbol{x})=\begin{bmatrix}\boldsymbol{\eta}_1(\boldsymbol{x})\\\vdots\\\boldsymbol{\eta}_q(\boldsymbol{x})\end{bmatrix}\in\mathbb{R}^{n_1+\cdots+n_q}$$

微积分讲稿 谢锡麟

按复合映照可微性定理，则有

$$\begin{aligned}\boldsymbol{D}\hat{\boldsymbol{f}}(\boldsymbol{x}) &= \left[\boldsymbol{D}_{\eta_1}\boldsymbol{f},\cdots,\boldsymbol{D}_{\eta_q}\boldsymbol{f}\right](\boldsymbol{\eta}(\boldsymbol{x}))\begin{bmatrix}\boldsymbol{D}\boldsymbol{\eta}_1(\boldsymbol{x})\\ \vdots\\ \boldsymbol{D}\boldsymbol{\eta}_q(\boldsymbol{x})\end{bmatrix}\\ &= \boldsymbol{D}_{\eta_1}\boldsymbol{f}(\boldsymbol{\eta}(\boldsymbol{x}))\boldsymbol{D}\boldsymbol{\eta}_1(\boldsymbol{x})+\cdots+\boldsymbol{D}_{\eta_q}\boldsymbol{f}(\boldsymbol{\eta}(\boldsymbol{x}))\boldsymbol{D}\boldsymbol{\eta}_q(\boldsymbol{x})\end{aligned}$$

1.3 高阶偏导数

无论对多元函数还是向量值映照, 如有存在对自变量某个分量的一阶偏导函数, 则可进一步考虑其对自变量某个分量的偏导数, 含有二次或二次以上的偏导数可统称为**高阶偏导数**.

例如, 对多元函数

$$f(\boldsymbol{x}):\mathbb{R}^m\supset\mathscr{D}_x\ni\boldsymbol{x}\mapsto f(\boldsymbol{x})\in\mathbb{R}$$

存在关于 x^i 的偏导函数

$$\frac{\partial f}{\partial x^i}(\boldsymbol{x}):\mathbb{R}^m\supset\mathscr{D}_x\ni\boldsymbol{x}\mapsto\frac{\partial f}{\partial x^i}(\boldsymbol{x})\in\mathbb{R}.$$

则可进一步考虑 $\dfrac{\partial f}{\partial x^i}(\boldsymbol{x})$ 在 $\boldsymbol{x}$ 点关于 x^j 的变化率

$$\frac{\partial^2 f}{\partial x^j\partial x^i}(\boldsymbol{x})\triangleq\frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^i}\right)(\boldsymbol{x})\triangleq\lim_{\lambda\to 0\in\mathbb{R}}\frac{\frac{\partial f}{\partial x^i}(\boldsymbol{x}+\lambda\boldsymbol{i}_j)-\frac{\partial f}{\partial x^i}(\boldsymbol{x})}{\lambda}.$$

可称 $\dfrac{\partial^2 f}{\partial x^j\partial x^i}(\boldsymbol{x})$ 为 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}$ 点的二阶偏导数, 如果 $i\neq j$ 又可称为混合偏导数.

值得指出, 一般情况混合偏导数并不一定相等, 亦即

$$\frac{\partial^2 f}{\partial x^i\partial x^j}\triangleq\lim_{\lambda\to 0\in\mathbb{R}}\frac{\frac{\partial f}{\partial x^j}(\boldsymbol{x}+\lambda\boldsymbol{i}_i)-\frac{\partial f}{\partial x^j}(\boldsymbol{x})}{\lambda}\neq\lim_{\lambda\to 0\in\mathbb{R}}\frac{\frac{\partial f}{\partial x^i}(\boldsymbol{x}+\lambda\boldsymbol{i}_j)-\frac{\partial f}{\partial x^i}(\boldsymbol{x})}{\lambda}\triangleq\frac{\partial^2 f}{\partial x^j\partial x^i}(\boldsymbol{x}).$$

2 应用事例

2.1 分片函数的偏导数

事例 1 (一阶偏导数都不连续，不可微)**.**

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

连续性：全空间连续

一阶偏导数：

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases}\dfrac{2xy^3}{(x^2+y^2)^2} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

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

微积分讲稿 谢锡麟

事例 2 (一阶偏导数都不连续，可微)**.**

$$f(x,y)=\begin{cases} xy\,sin\dfrac{1}{\sqrt{x^2+y^2}} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

连续性：全空间连续

一阶偏导数：

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases} y\,sin\dfrac{1}{\sqrt{x^2+y^2}}-\dfrac{x^2y}{(x^2+y^2)^{3/2}}\,cos\dfrac{1}{\sqrt{x^2+y^2}} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases} x\,sin\dfrac{1}{\sqrt{x^2+y^2}}-\dfrac{xy^2}{(x^2+y^2)^{3/2}}\,cos\dfrac{1}{\sqrt{x^2+y^2}} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

事例 3 (一阶偏导数都连续，可微)**.**

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

连续性：全空间连续

一阶偏导数：

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases} y\,\dfrac{x^2-y^2}{x^2+y^2}+xy\,\dfrac{4xy^2}{(x^2+y^2)^2} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

$$\frac{\partial f}{\partial y}(x,y)=\begin{cases} x\,\dfrac{x^2-y^2}{x^2+y^2}-xy\,\dfrac{4xy^2}{(x^2+y^2)^2} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

3 拓广深化

3.1 单参数向量值映照的变化率

性质 3.1. 设 $\boldsymbol{A}(t),\boldsymbol{B}(t)\in\mathbb{R}^3$ 为两个单参数向量值映照, 如果 $\exists\,\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t),\ \dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^3$, 则有

1. $\exists\,\dfrac{\mathrm{d}}{\mathrm{d}t}(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)=\alpha\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)+\beta\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^3,\quad\forall\,\alpha,\beta\in\mathbb{R};$
2. $\exists\,\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{A},\,\boldsymbol{B}\right)_{\mathbb{R}^3}(t)=\left(\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t),\,\boldsymbol{B}(t)\right)_{\mathbb{R}^3}+\left(\boldsymbol{A}(t),\,\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right)_{\mathbb{R}^3}\in\mathbb{R};$
3. $\exists\,\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\times\boldsymbol{B})(t)=\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\times\boldsymbol{B}(t)+\boldsymbol{A}(t)\times\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^3.$

微积分讲稿 谢锡麟

证明 方法一 按映照观点. $(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)\in\mathbb{R}^3$ 可理解为

$$\mathbb{R}\ni t\mapsto(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)=\alpha\begin{bmatrix}A^1\\A^2\\A^3\end{bmatrix}(t)+\beta\begin{bmatrix}B^1\\B^2\\B^3\end{bmatrix}(t)=\begin{bmatrix}(\alpha A^1+\beta B^1)(t)\\(\alpha A^2+\beta B^2)(t)\\(\alpha A^3+\beta B^3)(t)\end{bmatrix}\in\mathbb{R}^3.$$

由此, 可有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)&=\begin{bmatrix}\left(\alpha\dfrac{\mathrm{d}A^1}{\mathrm{d}t}+\beta\dfrac{\mathrm{d}B^1}{\mathrm{d}t}\right)(t)\\\left(\alpha\dfrac{\mathrm{d}A^2}{\mathrm{d}t}+\beta\dfrac{\mathrm{d}B^2}{\mathrm{d}t}\right)(t)\\\left(\alpha\dfrac{\mathrm{d}A^3}{\mathrm{d}t}+\beta\dfrac{\mathrm{d}B^3}{\mathrm{d}t}\right)(t)\end{bmatrix}=\alpha\begin{bmatrix}\dfrac{\mathrm{d}A^1}{\mathrm{d}t}\\\dfrac{\mathrm{d}A^2}{\mathrm{d}t}\\\dfrac{\mathrm{d}A^3}{\mathrm{d}t}\end{bmatrix}(t)+\beta\begin{bmatrix}\dfrac{\mathrm{d}B^1}{\mathrm{d}t}\\\dfrac{\mathrm{d}B^2}{\mathrm{d}t}\\\dfrac{\mathrm{d}B^3}{\mathrm{d}t}\end{bmatrix}(t)\\&=\alpha\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)+\beta\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^3.\end{aligned}$$

类似地, 有

$$\mathbb{R}\ni t\mapsto(\boldsymbol{A},\boldsymbol{B})_{\mathbb{R}^3}(t)=(A^1B^1)(t)+(A^2B^2)(t)+(A^3B^3)(t)\in\mathbb{R}.$$

由此, 可有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A},\boldsymbol{B})_{\mathbb{R}^3}(t)&=\left(\frac{\mathrm{d}A^1}{\mathrm{d}t}B^1\right)(t)+\left(\frac{\mathrm{d}A^2}{\mathrm{d}t}B^2\right)(t)+\left(\frac{\mathrm{d}A^3}{\mathrm{d}t}B^3\right)(t)+\left(A^1\frac{\mathrm{d}B^1}{\mathrm{d}t}\right)(t)\\&\quad+\left(A^2\frac{\mathrm{d}B^2}{\mathrm{d}t}\right)(t)+\left(A^3\frac{\mathrm{d}B^3}{\mathrm{d}t}\right)(t)\\&=\left(\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t),\,\boldsymbol{B}(t)\right)_{\mathbb{R}^3}+\left(\boldsymbol{A}(t),\,\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right)_{\mathbb{R}^3}\in\mathbb{R}.\end{aligned}$$

另有

$$\mathbb{R}\ni t\mapsto(\boldsymbol{A}\times\boldsymbol{B})(t)\triangleq\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\A^1&A^2&A^3\\B^1&B^2&B^3\end{vmatrix}(t)=\begin{bmatrix}A^2B^3-A^3B^2\\A^3B^1-A^1B^3\\A^1B^2-A^2B^1\end{bmatrix}(t)\in\mathbb{R}^3,$$

可有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\times\boldsymbol{B})(t)&=\begin{bmatrix}\dfrac{\mathrm{d}A^2}{\mathrm{d}t}B^3-\dfrac{\mathrm{d}A^3}{\mathrm{d}t}B^2\\\dfrac{\mathrm{d}A^3}{\mathrm{d}t}B^1-\dfrac{\mathrm{d}A^1}{\mathrm{d}t}B^3\\\dfrac{\mathrm{d}A^1}{\mathrm{d}t}B^2-\dfrac{\mathrm{d}A^2}{\mathrm{d}t}B^1\end{bmatrix}(t)+\begin{bmatrix}A^2\dfrac{\mathrm{d}B^3}{\mathrm{d}t}-A^3\dfrac{\mathrm{d}B^2}{\mathrm{d}t}\\A^3\dfrac{\mathrm{d}B^1}{\mathrm{d}t}-A^1\dfrac{\mathrm{d}B^3}{\mathrm{d}t}\\A^1\dfrac{\mathrm{d}B^2}{\mathrm{d}t}-A^2\dfrac{\mathrm{d}B^1}{\mathrm{d}t}\end{bmatrix}(t)\\&=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\times\boldsymbol{B}(t)+\boldsymbol{A}(t)\times\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^3.\end{aligned}$$

方法二 按极限分析. 考虑到

$$\exists\,\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\triangleq\lim_{\Delta t\to0}\frac{\boldsymbol{A}(t+\Delta t)-\boldsymbol{A}(t)}{\Delta t}\in\mathbb{R}^3$$

微积分讲稿 谢锡麟

等价于

$$\boldsymbol{A}(t+\Delta t)=\boldsymbol{A}(t)+\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\Delta t+o(\Delta t),$$

故有

$$\begin{aligned}(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t+\Delta t)&\equiv\alpha\boldsymbol{A}(t+\Delta t)+\beta\boldsymbol{B}(t+\Delta t)\\&=\alpha\left[\boldsymbol{A}(t)+\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\Delta t+o(\Delta t)\right]+\beta\left[\boldsymbol{B}(t)+\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\Delta t+o(\Delta t)\right]\\&=(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)+\left[\alpha\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)+\beta\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right]\Delta t+o(\Delta t),\end{aligned}$$

亦即 $\exists\dfrac{\mathrm{d}}{\mathrm{d}t}(\alpha\boldsymbol{A}+\beta\boldsymbol{B})(t)=\alpha\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)+\beta\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)$.

类似地, 有

$$\begin{aligned}(\boldsymbol{A},\boldsymbol{B})_{\mathbb{R}^3}(t+\Delta t)&\equiv(\boldsymbol{A}(t+\Delta t),\boldsymbol{B}(t+\Delta t))_{\mathbb{R}^3}\\&=\left(\boldsymbol{A}(t)+\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\Delta t+o(\Delta t),\,\boldsymbol{B}(t)+\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\Delta t+o(\Delta t)\right)_{\mathbb{R}^3}\\&=(\boldsymbol{A},\,\boldsymbol{B})_{\mathbb{R}^3}(t)+\left[\left(\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t),\,\boldsymbol{B}(t)\right)_{\mathbb{R}^3}+\left(\boldsymbol{A}(t),\,\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right)_{\mathbb{R}^3}\right]\Delta+o(\Delta t)\end{aligned}$$

$$\begin{aligned}(\boldsymbol{A}\times\boldsymbol{B})(t+\Delta t)&\equiv\boldsymbol{A}(t+\Delta t)\times\boldsymbol{B}(t+\Delta t)\\&=\left[\boldsymbol{A}(t)+\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\Delta t+o(\Delta t)\right]\times\left[\boldsymbol{B}(t)+\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\Delta t+o(\Delta t)\right]\\&=(\boldsymbol{A}\times\boldsymbol{B})(t)+\left[\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\times\boldsymbol{B}(t)+\boldsymbol{A}(t)\times\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right]\Delta t+o(\Delta t),\end{aligned}$$

由此即有 $\exists\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A},\,\boldsymbol{B})_{\mathbb{R}^3}(t)=\left(\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t),\,\boldsymbol{B}(t)\right)_{\mathbb{R}^3}+\left(\boldsymbol{A}(t),\,\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\right)_{\mathbb{R}^3}$.

另外, 可有 $\exists\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\times\boldsymbol{B})(t)=\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\times\boldsymbol{B}(t)+\boldsymbol{A}(t)\times\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)$.

□

性质 3.2. 1. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 为常数矩阵, $\boldsymbol{b}(t)\in\mathbb{R}^n$, 有

$$\mathbb{R}\ni t\mapsto\boldsymbol{A}\boldsymbol{b}(t)\in\mathbb{R}^m$$

则有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\boldsymbol{b})(t)=\boldsymbol{A}\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t)\in\mathbb{R}^m.$$

2. 设 $\boldsymbol{A}(t)\in\mathbb{R}^{m\times n}$，$\boldsymbol{b}(t)\in\mathbb{R}^n$, 有

$$\mathbb{R}\ni t\mapsto(\boldsymbol{A}\boldsymbol{b})(t)=\boldsymbol{A}(t)\boldsymbol{b}(t)\in\mathbb{R}^m$$

则有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\boldsymbol{b})(t)=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{b}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t)\in\mathbb{R}^m,$$

此处

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t):=\begin{bmatrix}\dfrac{\mathrm{d}A_{11}}{\mathrm{d}t}&\cdots&\dfrac{\mathrm{d}A_{1n}}{\mathrm{d}t}\\\vdots&&\vdots\\\dfrac{\mathrm{d}A_{m1}}{\mathrm{d}t}&\cdots&\dfrac{\mathrm{d}A_{mn}}{\mathrm{d}t}\end{bmatrix}(t).$$

3. 设 $\boldsymbol{A}(t)\in\mathbb{R}^{m\times n}$，$\boldsymbol{B}(t)\in\mathbb{R}^{n\times l}$, 有

$$\mathbb{R}\ni t\mapsto(\boldsymbol{AB})(t)=\boldsymbol{A}(t)\boldsymbol{B}(t)\in\mathbb{R}^{m\times l},$$

则有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{AB})(t)=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t).$$

4. 设 $\boldsymbol{A}(t)\in\mathbb{R}^{m\times n},\boldsymbol{B}(t)\in\mathbb{R}^{n\times l},\boldsymbol{C}(t)\in\mathbb{R}^{l\times p}$, 有

$$\mathbb{R}\ni t\mapsto(\boldsymbol{ABC})(t)=\boldsymbol{A}(t)\boldsymbol{B}(t)\boldsymbol{C}(t)\in\mathbb{R}^{m\times p},$$

则有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{ABC})(t)=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\boldsymbol{B}(t)\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}(t)\in\mathbb{R}^{m\times p}.$$

证明 1. 考虑

$$\boldsymbol{Ab}(t)=\begin{bmatrix}A_{11}&\cdots&A_{1n}\\ \vdots&&\vdots\\ A_{i1}&\cdots&A_{in}\\ \vdots&&\vdots\\ A_{m1}&\cdots&A_{mn}\end{bmatrix}\begin{bmatrix}b^1\\ \vdots\\ b^n\end{bmatrix}(t)\in\mathbb{R}^m$$

则有

$$(\boldsymbol{Ab})_i(t)=\sum_{s=1}^n A_{is}b^s(t)$$

所以

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{Ab})_i(t)=\sum_{s=1}^n A_{is}\frac{\mathrm{d}b^s}{\mathrm{d}t}(t).$$

因此，有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{Ab})(t)=\boldsymbol{A}\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t)\in\mathbb{R}^n.$$

2. 由 $(\boldsymbol{Ab})_i(t)=\sum\limits_{s=1}^n A_{is}(t)b^s(t)$, 有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{Ab})_i(t)=\sum_{s=1}^n\frac{\mathrm{d}A_{is}}{\mathrm{d}t}(t)b^s(t)+\sum_{s=1}^n A_{is}(t)\frac{\mathrm{d},b^s}{\mathrm{d}t}(t)$$

即有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{Ab})(t)=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{b}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t)\in\mathbb{R}^n.$$

3. 由

$$(\boldsymbol{AB})_{ij}(t)=\sum_{k=1}^n A_{ik}(t)B_{kj}(t),\quad 1\leqslant i\leqslant m,1\leqslant j\leqslant l$$

则有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{AB})_{ij}(t)=\sum_{k=1}^n\frac{\mathrm{d}A_{ik}}{\mathrm{d}t}(t)B_{kj}(t)+\sum_{k=1}^n A_{ik}(t)\frac{\mathrm{d}B_{kj}}{\mathrm{d}t}(t),$$

微积分讲稿 谢锡麟

即有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{AB})(t)=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\in\mathbb{R}^{n\times l}.$$

4. 由 (3) 即有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{ABC})(t)&=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{BC}}{\mathrm{d}t}(t)\\&=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\boldsymbol{C}(t)+\boldsymbol{B}(t)\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}(t)\right)\\&=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}(t)\boldsymbol{B}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}(t)\boldsymbol{C}(t)+\boldsymbol{A}(t)\boldsymbol{B}(t)\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}(t).\end{aligned}$$

□

3.2 旋转的数学机制及其应用

3.2.1 角速度向量

设 $\{\boldsymbol{i}_i\}_{i=1}^m\subset\mathbb{R}^m$ 为典则基, $\{\boldsymbol{e}_i(t)\}_{i=1}^m\subset\mathbb{R}^m$ 为随时间 t 变化的单位正交基, 则可建立关系式

$$\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_m\end{bmatrix}(t)=\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}\begin{bmatrix}P_{11} & \cdots & P_{1m}\\ \vdots & & \vdots\\ P_{m1} & \cdots & P_{mm}\end{bmatrix}(t).$$

记作 $\boldsymbol{e}(t)=\boldsymbol{iP}(t)$。由于 $\{\boldsymbol{e}_i(t)\}_{i=1}^m$ 为单位正交基, 即有

$$\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_m^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{e}_1 & \cdots\boldsymbol{e}_m\end{bmatrix}=\boldsymbol{e}^{\mathrm{T}}(t)\boldsymbol{e}(t)=\boldsymbol{I}_m,$$

则有 $\boldsymbol{P}^{\mathrm{T}}(t)\boldsymbol{P}(t)=\boldsymbol{I}_m$, 亦即 $\boldsymbol{P}(t)\in\mathbf{Orth}$.

考虑 $\forall\,\boldsymbol{b}(t)\in\mathbb{R}^m$, 则有

$$\boldsymbol{b}=\begin{cases}\begin{bmatrix}\boldsymbol{i}_1 & \cdots & \boldsymbol{i}_m\end{bmatrix}\begin{bmatrix}\overset{i}{b}{}^1(t)\\ \vdots\\ \overset{i}{b}{}^m(t)\end{bmatrix}=:\boldsymbol{i}\overset{i}{\boldsymbol{b}}(t)\\ \begin{bmatrix}\boldsymbol{e}_1(t) & \cdots & \boldsymbol{e}_m(t)\end{bmatrix}\begin{bmatrix}\overset{e}{b}{}^1(t)\\ \vdots\\ \overset{e}{b}{}^m(t)\end{bmatrix}=:\boldsymbol{e}(t)\overset{e}{\boldsymbol{b}}(t).\end{cases}$$

由于 $\boldsymbol{e}(t)=\boldsymbol{iP}(t)$, 则有

$$\begin{cases}\overset{i}{\boldsymbol{b}}(t)=\boldsymbol{P}(t)\overset{e}{\boldsymbol{b}}(t)\\ \overset{e}{\boldsymbol{b}}(t)=\boldsymbol{P}^{\mathrm{T}}(t)\overset{i}{\boldsymbol{b}}(t)\end{cases}.$$

微积分讲稿 谢锡麟

计算 $\boldsymbol{b}(t)$ 在典则基 $\{\boldsymbol{i}_i\}_{i=1}^m$ 下的变化率

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{e}\overset{e}{\boldsymbol{b}})(t) = \dot{\boldsymbol{e}}(t)\overset{e}{\boldsymbol{b}}(t) + \boldsymbol{e}(t)\frac{\mathrm{d}\overset{e}{\boldsymbol{b}}}{\mathrm{d}t}(t) = \boldsymbol{e}(t)\frac{\mathrm{d}\overset{e}{\boldsymbol{b}}}{\mathrm{d}t}(t) + \boldsymbol{i}\dot{\boldsymbol{P}}(t)\overset{e}{\boldsymbol{b}}(t)\\ &= \boldsymbol{e}(t)\frac{\mathrm{d}\overset{e}{\boldsymbol{b}}}{\mathrm{d}t}(t) + \boldsymbol{e}(t)(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t)\overset{e}{\boldsymbol{b}}(t).\end{aligned}$$

考虑到 $\boldsymbol{P}^{\mathrm{T}}(t)\boldsymbol{P}(t) = \boldsymbol{I}_m$, 可有

$$\boldsymbol{0} = (\dot{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{P})(t) + (\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t) = (\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})^{\mathrm{T}}(t) + (\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t),$$

故有 $(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t) \in \mathbf{Skw}$, 即为反对称矩阵. 记

$$(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t) = \begin{bmatrix} 0 & -\omega_3 & \omega_2\\ \omega_3 & 0 & -\omega_1\\ -\omega_2 & \omega_1 & 0\end{bmatrix}(t)$$

则有

$$\begin{aligned}\boldsymbol{e}(t)(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t)\overset{e}{\boldsymbol{b}}(t) &= \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix}(t)\begin{bmatrix} 0 & -\omega_3 & \omega_2\\ \omega_3 & 0 & -\omega_1\\ -\omega_2 & \omega_1 & 0\end{bmatrix}(t)\begin{bmatrix}\overset{e}{b}{}^1\\ \overset{e}{b}{}^2\\ \overset{e}{b}{}^3\end{bmatrix}(t)\\ &= \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix}(t)\begin{bmatrix}\omega_2\overset{e}{b}{}^3 - \omega_3\overset{e}{b}{}^2\\ \omega_3\overset{e}{b}{}^1 - \omega_1\overset{e}{b}{}^3\\ \omega_1\overset{e}{b}{}^2 - \omega_2\overset{e}{b}{}^1\end{bmatrix}(t) = \begin{vmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\\ \omega_1 & \omega_2 & \omega_3\\ \overset{e}{b}{}^1 & \overset{e}{b}{}^2 & \overset{e}{b}{}^3\end{vmatrix}\\ &= \boldsymbol{\omega}(t)\times\boldsymbol{b}(t),\end{aligned}$$

式中 $\boldsymbol{\omega}(t) = \omega_1\boldsymbol{e}_1 + \omega_2\boldsymbol{e}_2 + \omega_3\boldsymbol{e}_3$ 可称为反对称矩阵 $(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t)$ 的伴随向量, 亦可称为**角速度向量**.

引入 $\boldsymbol{b}(t)$ 在正交基 $\{\boldsymbol{e}_i(t)\}_{i=1}^m$ 下的变化率为

$$\frac{\overset{e}{\mathrm{d}}\boldsymbol{b}}{\mathrm{d}t}(t) = \frac{\overset{e}{\mathrm{d}}}{\mathrm{d}t}(\boldsymbol{e}\overset{e}{\boldsymbol{b}})(t) \triangleq \boldsymbol{e}(t)\frac{\mathrm{d}\overset{e}{\boldsymbol{b}}}{\mathrm{d}t}(t).$$

由此, 可有 $\boldsymbol{b}(t) \in \mathbb{R}^3$ 在典则基下变化率与在运动正交基 $\{\boldsymbol{e}_i(t)\}_{i=1}^m$ 下的变化率之间的关系为

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t) = \frac{\overset{e}{\mathrm{d}}\boldsymbol{b}}{\mathrm{d}t}(t) + \boldsymbol{\omega}(t)\times\boldsymbol{b}(t).$$

3.2.2 研究简单的旋转

如图2的定轴旋转, 可有

$$\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix} = \begin{bmatrix}\boldsymbol{i}_1 & \boldsymbol{i}_2 & \boldsymbol{i}_3\end{bmatrix}\begin{bmatrix}\cos\theta(t) & -\sin\theta(t) & 0\\ \sin\theta(t) & \cos\theta(t) & 0\\ 0 & 0 & 1\end{bmatrix} = \boldsymbol{i}\boldsymbol{P}(t),$$

微积分讲稿 谢锡麟

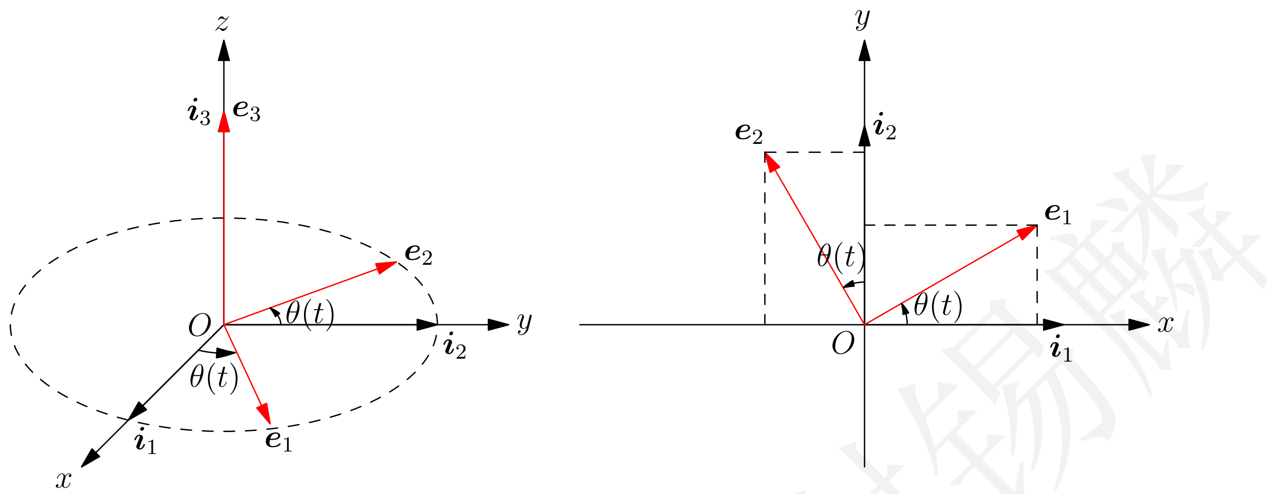

图 2: 空间定轴旋转

则有

$$
\begin{aligned}
(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t) &= \begin{bmatrix} \cos\theta(t) & \sin\theta(t) & 0 \\ -\sin\theta(t) & \cos\theta(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\sin\theta(t) & -\cos\theta(t) & 0 \\ \cos\theta(t) & -\sin\theta(t) & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{\theta}(t) \\
&= \begin{bmatrix} 0 & -\dot{\theta}(t) & 0 \\ \dot{\theta}(t) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} =: \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}.
\end{aligned}
$$

所以

$$
\boldsymbol{\omega}(t) = \dot{\theta}(t)\boldsymbol{e}_3 = \dot{\theta}(t)\boldsymbol{i}_3 \in \mathbb{R}^3
$$

即为定轴转动的角速度.

3.2.3 研究一般的旋转

如图3的定点旋转, 可有

$$
\begin{bmatrix} \tilde{e}_1 & \tilde{e}_2 & \tilde{e}_3 \end{bmatrix}(t) = \begin{bmatrix} \boldsymbol{i}_1 & \boldsymbol{i}_2 & \boldsymbol{i}_3 \end{bmatrix} \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) & 0 \\ \sin\alpha(t) & \cos\alpha(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} =: \boldsymbol{i}\boldsymbol{A}(t),
$$

$$
\begin{bmatrix} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \end{bmatrix}(t) = \begin{bmatrix} \tilde{e}_1 & \tilde{e}_2 & \tilde{e}_3 \end{bmatrix}(t) \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta(t) & -\sin\beta(t) \\ 0 & \sin\beta(t) & \cos\beta(t) \end{bmatrix} =: \tilde{e}(t)\boldsymbol{B}(t),
$$

$$
\begin{bmatrix} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \end{bmatrix}(t) = \begin{bmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \end{bmatrix}(t) \begin{bmatrix} \cos\gamma(t) & -\sin\gamma(t) & 0 \\ \sin\gamma(t) & \cos\gamma(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} =: \hat{e}(t)\boldsymbol{C}(t).
$$

微积分讲稿 谢锡麟

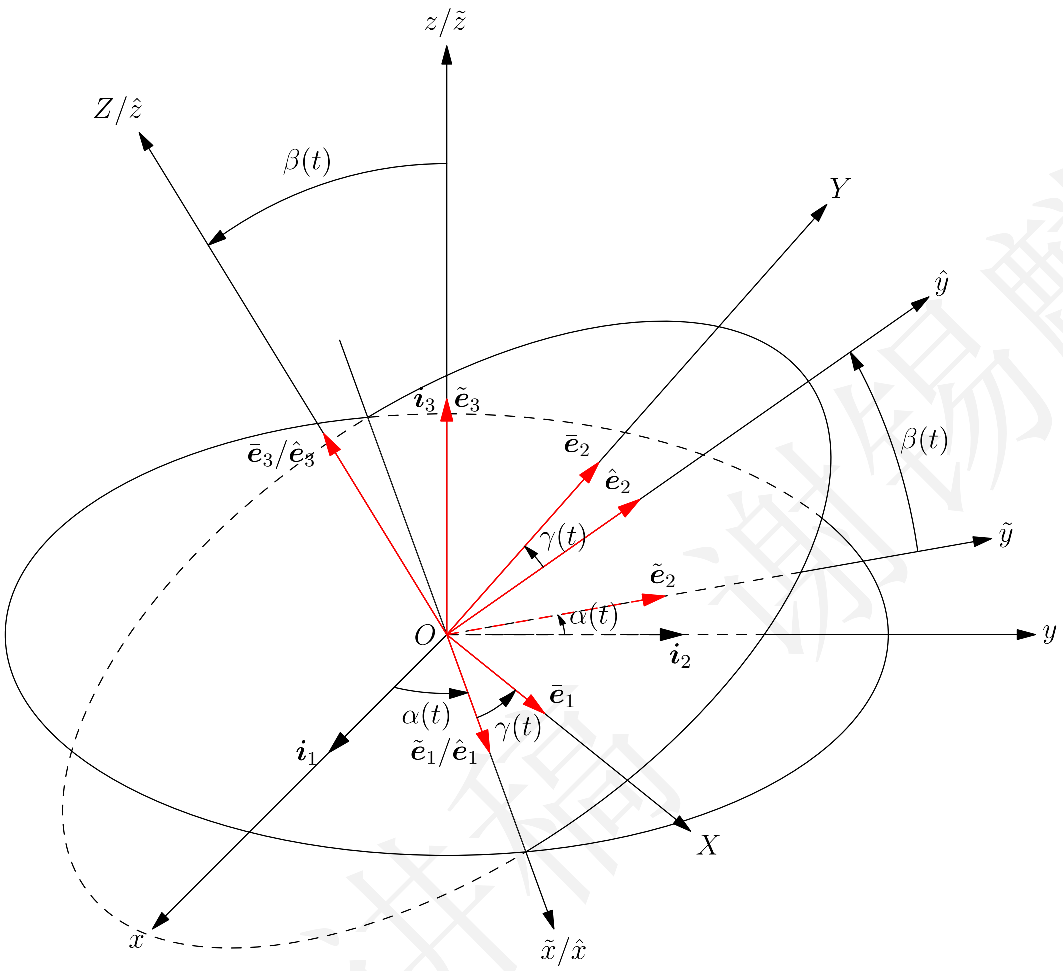

图 3: 空间定点旋转

计算可得 $(\boldsymbol{A}^{\mathrm{T}}\dot{\boldsymbol{A}})(t), (\boldsymbol{B}^{\mathrm{T}}\dot{\boldsymbol{B}})(t), (\boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{C}})(t)$ 对应的伴随向量分别为

$$\boldsymbol{\omega}_A(t) := \dot{\alpha}(t)\tilde{\boldsymbol{e}}_3(t), \quad \boldsymbol{\omega}_B(t) := \dot{\beta}(t)\hat{\boldsymbol{e}}_1(t), \quad \boldsymbol{\omega}_C(t) := \dot{\gamma}(t)\bar{\boldsymbol{e}}_3(t).$$

综上，可有

$$\bar{\boldsymbol{e}}(t) = \hat{\boldsymbol{e}}(t)\boldsymbol{C}(t) = \tilde{\boldsymbol{e}}\boldsymbol{B}(t)\boldsymbol{C}(t) = \boldsymbol{i}\boldsymbol{A}(t)\boldsymbol{B}(t)\boldsymbol{C}(t) =: \boldsymbol{i}\boldsymbol{P}(t)$$

以及

$$\dot{\boldsymbol{P}}(t) = \dot{\boldsymbol{A}}(t)\boldsymbol{B}(t)\boldsymbol{C}(t) + \boldsymbol{A}(t)\dot{\boldsymbol{B}}(t)\boldsymbol{C}(t) + \boldsymbol{A}(t)\boldsymbol{B}(t)\dot{\boldsymbol{C}}(t) \in \mathbb{R}^{3\times 3}$$

则有

$$\begin{aligned}(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t) &= \boldsymbol{C}^{\mathrm{T}}(t)\boldsymbol{B}^{\mathrm{T}}(t)(\boldsymbol{A}^{\mathrm{T}}\dot{\boldsymbol{A}})(t)\boldsymbol{B}(t)\boldsymbol{C}(t) + \boldsymbol{C}^{\mathrm{T}}(t)(\boldsymbol{B}^{\mathrm{T}}\dot{\boldsymbol{B}})(t)\boldsymbol{C}(t) + (\boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{C}})(t) \\ &= (\boldsymbol{B}\boldsymbol{C})^{\mathrm{T}}(t)(\boldsymbol{A}^{\mathrm{T}}\dot{\boldsymbol{A}})(t)(\boldsymbol{B}\boldsymbol{C})(t) + \boldsymbol{C}^{\mathrm{T}}(t)(\boldsymbol{B}^{\mathrm{T}}\dot{\boldsymbol{B}})(t)\boldsymbol{C}(t) + (\boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{C}})(t).\end{aligned}$$

已有

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t) = \bar{\boldsymbol{e}}(t)\frac{\mathrm{d}\bar{\boldsymbol{b}}}{\mathrm{d}t}(t) + \bar{\boldsymbol{e}}(t)(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t)\bar{\boldsymbol{b}}(t),$$

微积分讲稿 谢锡麟

此处 $\bar{\boldsymbol{b}}(t)$ 为 $\boldsymbol{b}(t)$ 在单位正交基 $\{\bar{\boldsymbol{e}}_i(t)\}_{i=1}^3$ 下的坐标. 现有

$$\bar{\boldsymbol{e}}(t)(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})(t)\bar{\boldsymbol{b}}(t) = \bar{\boldsymbol{e}}(\boldsymbol{BC})^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\dot{\boldsymbol{A}})(\boldsymbol{BC})\bar{\boldsymbol{b}} + \bar{\boldsymbol{e}}\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{B}^{\mathrm{T}}\dot{\boldsymbol{B}})\boldsymbol{C}\bar{\boldsymbol{C}} + \bar{\boldsymbol{e}}(\boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{C}})\bar{\boldsymbol{b}}.$$

再考虑到 $\bar{\boldsymbol{e}} = \tilde{\boldsymbol{e}}(\boldsymbol{BC})$, 有

$$\boldsymbol{b} = \bar{\boldsymbol{e}}\bar{\boldsymbol{b}} = \tilde{\boldsymbol{e}}(\boldsymbol{BC})\bar{\boldsymbol{b}} = \tilde{\boldsymbol{e}}\tilde{\boldsymbol{b}},$$

因此有 $\tilde{\boldsymbol{b}} = (\boldsymbol{BC})\bar{\boldsymbol{b}}$. 同理由 $\bar{\boldsymbol{e}} = \hat{\boldsymbol{e}}\boldsymbol{C}$, 可有

$$\boldsymbol{b} = \bar{\boldsymbol{e}}\bar{\boldsymbol{b}} = \hat{\boldsymbol{e}}\boldsymbol{C}\bar{\boldsymbol{b}},$$

因此有 $\hat{\boldsymbol{b}} = \boldsymbol{C}\bar{\boldsymbol{b}}$, 此处 $\tilde{\boldsymbol{b}}$ 和 $\hat{\boldsymbol{b}}$ 即为 $\boldsymbol{b}(t)$ 在单位正交基 $\{\tilde{\boldsymbol{e}}_i(t)\}_{i=1}^3$ 和 $\{\hat{\boldsymbol{e}}_i(t)\}_{i=1}^3$ 下的坐标.

综上，可有

$$\begin{aligned}
\bar{\boldsymbol{e}}(\boldsymbol{P}^{\mathrm{T}}\dot{\boldsymbol{P}})\bar{\boldsymbol{b}} &= \tilde{\boldsymbol{e}}(\boldsymbol{A}^{\mathrm{T}}\dot{\boldsymbol{A}})\tilde{\boldsymbol{b}} + \hat{\boldsymbol{e}}(\boldsymbol{B}^{\mathrm{T}}\dot{\boldsymbol{B}})\hat{\boldsymbol{b}} + \bar{\boldsymbol{e}}(\boldsymbol{C}^{\mathrm{T}}\dot{\boldsymbol{C}})\bar{\boldsymbol{b}} \\
&= \boldsymbol{\omega}_A \times \boldsymbol{b} + \boldsymbol{\omega}_B \times \boldsymbol{b} + \boldsymbol{\omega}_C \times \boldsymbol{b} \\
&= (\boldsymbol{\omega}_A + \boldsymbol{\omega}_B + \boldsymbol{\omega}_C) \times \boldsymbol{b} = \bar{\boldsymbol{\omega}} \times \boldsymbol{b}.
\end{aligned}$$

由此，有

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}(t) &= \frac{\bar{\mathrm{d}}\boldsymbol{b}}{\mathrm{d}t}(t) + \bar{\boldsymbol{\omega}}(t) \times \boldsymbol{b}(t) \\
&= \frac{\bar{\mathrm{d}}\boldsymbol{b}}{\mathrm{d}t}(t) + (\boldsymbol{\omega}_A + \boldsymbol{\omega}_B + \boldsymbol{\omega}_C) \times \boldsymbol{b}(t) \\
&= \frac{\bar{\mathrm{d}}\boldsymbol{b}}{\mathrm{d}t}(t) + (\dot{\alpha}\tilde{e}_3 + \dot{\beta}\hat{e}_1 + \dot{\gamma}\bar{e})(t) \times \boldsymbol{b}(t).
\end{aligned}$$

上述结论称为**角速度合成定理**.

3.3 曲线坐标系

有一类特殊的向量值映照称为$\mathscr{C}^p$ **微分同胚**, 在高维微积分、流形上的微积分、力学等方面都有着极为重要的应用.

定义 3.1 ($\mathscr{C}^p$ 微分同胚)**.** 对映照

$$\boldsymbol{X}(\boldsymbol{x}) : \mathbb{R}^m \supset D_x \ni \boldsymbol{x} \mapsto \boldsymbol{X}(\boldsymbol{x}) \in \mathbb{R}^m,$$

满足:

1. 定义域 D_x 和值域 $D_X := \boldsymbol{X}(D_x)$ 均为开集;
2. $\boldsymbol{X}(\boldsymbol{x})$ 实现 D_x 和 D_X 之间的双射;
3. $\boldsymbol{X}(\boldsymbol{x})$ 及其逆映照 $\boldsymbol{x}(\boldsymbol{X})$ 均为 $\mathscr{C}^p$ 映照,

则称映照 $\boldsymbol{X}(\boldsymbol{x})$ 实现集合 D_x 与 D_X 之间的**微分同胚**, 记作 $\boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p(D_x; D_X)$.

微积分讲稿 谢锡麟

本书称 D_X 为**物理区域**, 亦即物理事件实际发生的区域, 如连续介质实际所在的区域; 称 D_x 为**参数区域**. 由于存在 $\boldsymbol{x} \in D_x$ 同 $\boldsymbol{X} \in D_X$ 之间的一一对应关系, 故物理区域中的位置刻画可等价地对应至参数区域中的位置刻画, 籍此也称 $\boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p(D_x; D_X)$ 为**曲线坐标系**.

考虑 $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{x})$ 的 Jacobi 阵

$$\begin{aligned} D\boldsymbol{X}(\boldsymbol{x}) &= \begin{pmatrix} \dfrac{\partial X^1}{\partial x^1} & \cdots & \dfrac{\partial X^1}{\partial x^i} & \cdots & \dfrac{\partial X^1}{\partial x^m} \\ \vdots & & \vdots & & \vdots \\ \dfrac{\partial X^m}{\partial x^1} & \cdots & \dfrac{\partial X^m}{\partial x^i} & \cdots & \dfrac{\partial X^m}{\partial x^m} \end{pmatrix} (\boldsymbol{x}) \\ &=: \begin{pmatrix} \boldsymbol{g}_1 & \cdots & \boldsymbol{g}_i & \cdots & \boldsymbol{g}_m \end{pmatrix} (\boldsymbol{x}) \in \mathbb{R}^{m\times m}, \end{aligned}$$

式中

$$\boldsymbol{g}_i(\boldsymbol{x}) \triangleq \lim_{\lambda\to 0} \frac{\boldsymbol{X}(\boldsymbol{x}+\lambda\boldsymbol{i}_i) - \boldsymbol{X}(\boldsymbol{x})}{\lambda} \in \mathbb{R}^m,$$

其几何意义为物理空间中 x^i 曲线在 $\boldsymbol{X}(\boldsymbol{x})$ 点的切向量. 由于 $D\boldsymbol{X}(\boldsymbol{x})$ 非奇异, 因此 $\{\boldsymbol{g}_i(\boldsymbol{x})\}_{i=1}^m$ 为线性无关向量组, 亦即成为 $\mathbb{R}^m$ 中的一个基, 且这种基随空间位置变化, 称为**曲线坐标系的局部协变基**.

考虑 $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X})$ 的 Jacobi 阵

$$\begin{aligned} D\boldsymbol{x}(\boldsymbol{X}) &= \begin{pmatrix} \dfrac{\partial x^1}{\partial X^1} & \cdots & \dfrac{\partial x^1}{\partial X^i} & \cdots & \dfrac{\partial x^1}{\partial X^m} \\ \vdots & & \vdots & & \vdots \\ \dfrac{\partial x^m}{\partial X^1} & \cdots & \dfrac{\partial x^m}{\partial X^i} & \cdots & \dfrac{\partial x^m}{\partial X^m} \end{pmatrix} \\ &=: \begin{pmatrix} \boldsymbol{g}^1 & \cdots & \boldsymbol{g}^i & \cdots & \boldsymbol{g}^m \end{pmatrix}^{\mathrm{T}} (\boldsymbol{X}) \in \mathbb{R}^{m\times m}, \end{aligned}$$

式中

$$\boldsymbol{g}^i(\boldsymbol{X}) \triangleq \begin{pmatrix} \dfrac{\partial x^i}{\partial X^1} & \cdots & \dfrac{\partial x^i}{\partial X^m} \end{pmatrix}^{\mathrm{T}} (\boldsymbol{X}) = \frac{\partial x^i}{\partial X^\alpha}(\boldsymbol{X})\boldsymbol{i}_\alpha \triangleq \operatorname{grad} x^i(\boldsymbol{X}),$$

其几何意义为物理空间中 $x^i(\boldsymbol{X})$ 在点 $\boldsymbol{X}$ 的梯度或者曲面 $x^i(\boldsymbol{X}) =$ 常数 的法向量方向.

考虑到 $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X})$ 为 $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{x})$ 的逆映照, 即有 $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{x}(\boldsymbol{X}))$ 或 $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X}(\boldsymbol{x}))$. 按复合向量值映照的可微性定理, 有

$$\boldsymbol{D}\boldsymbol{x}(X)\boldsymbol{D}\boldsymbol{X}(\boldsymbol{x}) = \begin{pmatrix} (\boldsymbol{g}_1, \boldsymbol{g}^1)_{\mathbb{R}^m} & \cdots & (\boldsymbol{g}_1, \boldsymbol{g}^i)_{\mathbb{R}^m} & \cdots & (\boldsymbol{g}_1, \boldsymbol{g}^m)_{\mathbb{R}^m} \\ \vdots & & \vdots & & \vdots \\ (\boldsymbol{g}_m, \boldsymbol{g}^1)_{\mathbb{R}^m} & \cdots & (\boldsymbol{g}_m, \boldsymbol{g}^i)_{\mathbb{R}^m} & \cdots & (\boldsymbol{g}_m, \boldsymbol{g}^m)_{\mathbb{R}^m} \end{pmatrix} = \boldsymbol{I} \in \mathbb{R}^{m\times m}.$$

由此, 可有 $\{\boldsymbol{g}^i(\boldsymbol{x})\}_{i=1}^m$ 亦为 $\mathbb{R}^m$ 中的一个基, 且满足对偶关系 $(\boldsymbol{g}_i, \boldsymbol{g}^j)_{\mathbb{R}^m} = \delta_i^j$, $i, j = 1, \cdots, m$. 可称 $\{\boldsymbol{g}^i(\boldsymbol{x})\}_{i=1}^m$ 为**曲线坐标系的局部逆变基**.

基于曲线坐标系的局部协变基及逆变基, 可引入

$$g_{ij}(\boldsymbol{x}) = \big(\boldsymbol{g}_i(\boldsymbol{x}), \boldsymbol{g}_j(\boldsymbol{x})\big)_{\mathbb{R}^m}, \quad g^{ij}(\boldsymbol{x}) = \big(\boldsymbol{g}^i(\boldsymbol{x}), \boldsymbol{g}^j(\boldsymbol{x})\big)_{\mathbb{R}^m},$$

可称 $\{g_{ij}(\boldsymbol{x})\}_{i,j=1}^m$ 或 $\{g^{ij}(\boldsymbol{x})\}_{i,j=1}^m$ 为曲线坐标系的度量.

三维 Euclid 空间中曲线坐标系如图4所示, 高维情况类似.

引入曲线坐标系 (微分同胚) 有两方面意义:

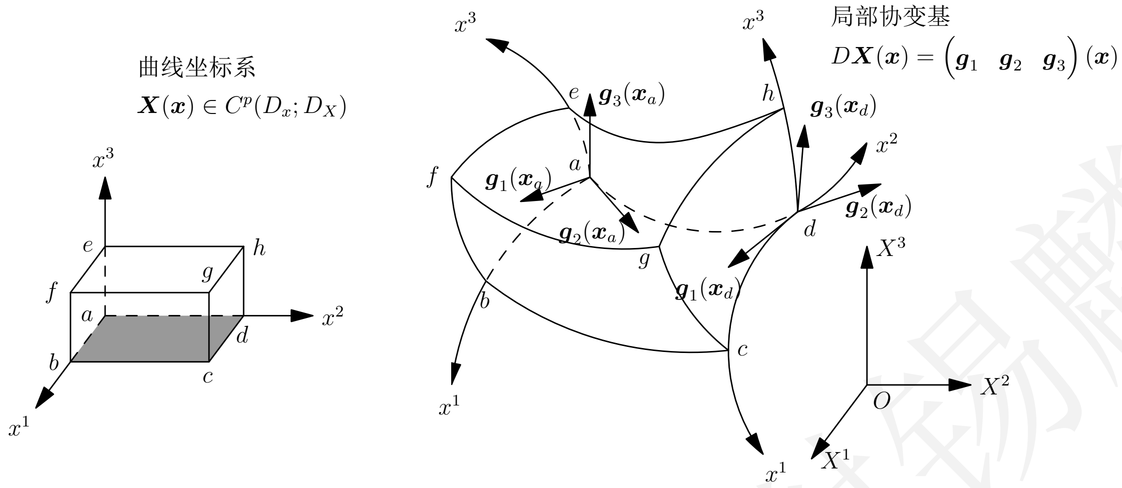

图 4: 三维 Euclid 空间中曲线坐标系示意

1. 将几何形态不规则的物理区域变换为几何形态规则的参数区域;
2. 可利用曲线坐标系自身诱导的局部基 (协变基及逆变基) 展开物理量满足的自然界守恒律方程, 以获得相对局部基的分量方程. 此种类型的方程有益于建立力学等物理过程同几何之间的关系.

3.4 曲线坐标系下的速度与加速度表示形式

在引入曲线坐标系 $\boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p(D_x; D_X)$ 的情况下, 首先在参数域中定义曲线

$$\boldsymbol{\Gamma}_x(\lambda) : \mathbb{R} \supset [\alpha, \beta] \ni \lambda \mapsto \boldsymbol{\Gamma}_x(\lambda) \equiv \boldsymbol{x}(\lambda) = \begin{pmatrix} x^1(\lambda) \\ \vdots \\ x^m(\lambda) \end{pmatrix} \in \mathbb{R}^m,$$

其中 $\{x^i\}_{i=1}^m$ 为曲线坐标系中的坐标.

然后, 基于曲线坐标系, 有物理域中的**体积中的曲线**

$$\boldsymbol{\Gamma}(\lambda) : \mathbb{R} \supset [\alpha, \beta] \ni t \mapsto \boldsymbol{\Gamma}(\lambda) \equiv \boldsymbol{X}(\lambda) = \boldsymbol{X} \circ \boldsymbol{x}(\lambda) = \begin{pmatrix} X^1(\lambda) \\ \vdots \\ X^m(\lambda) \end{pmatrix}$$

$$= \begin{pmatrix} X^1(\boldsymbol{x}(\lambda)) \\ \vdots \\ X^m(\boldsymbol{x}(\lambda)) \end{pmatrix} \in \mathbb{R}^m.$$

体积中的曲线, 如图5所示.

当曲线以时间为参数时, 也称曲线为**轨迹**.

速度定义为位置对时间的变化率, 即

$$\boldsymbol{v}(t) \triangleq \lim_{\Delta t \to 0} \frac{\boldsymbol{X}(t + \Delta t) - \boldsymbol{X}(t)}{\Delta t} = \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}(t) \in \mathbb{R}^m.$$

微积分讲稿 谢锡麟

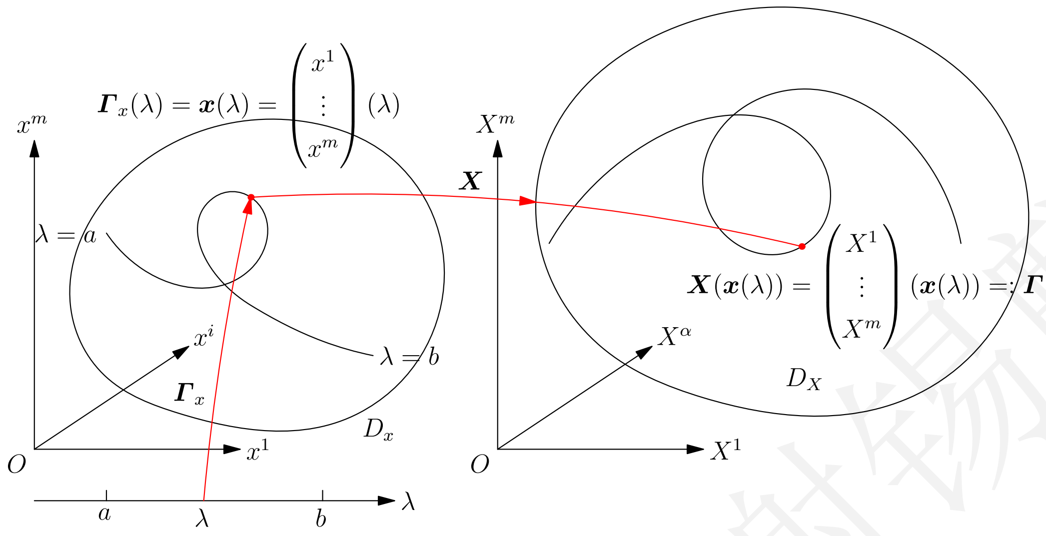

图 5: 体积中的曲线示意

考虑到 $\boldsymbol{X}(t) = \boldsymbol{X} \circ \boldsymbol{x}(t)$, 速度可以表示为

$$\begin{aligned}\boldsymbol{v}(t) = \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}(t) &= D\boldsymbol{X}(\boldsymbol{x})\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) \\ &= \Big(\boldsymbol{g}_1(\boldsymbol{x}), \cdots, \boldsymbol{g}_m(\boldsymbol{x})\Big)\begin{pmatrix} \dot{x}^1(t) \\ \vdots \\ \dot{x}^m(t) \end{pmatrix} = \dot{x}^i(t)\boldsymbol{g}_i(\boldsymbol{x}(t)) \in \mathbb{R}^m,\end{aligned}$$

此处 $\dot{x}^i(t)$ 表示函数 $x^i(t)$ 对时间 t 的导数, 本书采用此记号, 以后不再特别说明.

加速度定义为速度对时间的变化率, 即

$$\boldsymbol{a}(t) \triangleq \lim_{\Delta t \to 0} \frac{\boldsymbol{v}(t+\Delta t) - \boldsymbol{v}(t)}{\Delta t} = \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t),$$

表示为分量形式即为

$$\boldsymbol{a}(t) = a^i(t)\boldsymbol{g}_i(\boldsymbol{x}(t)) = a_i(t)\boldsymbol{g}^i(\boldsymbol{x}(t)),$$

其中

$$a^i(t) = \big(\boldsymbol{a}(t), \boldsymbol{g}^i(\boldsymbol{x}(t))\big)_{\mathbb{R}^3}, \quad a_i(t) = (\boldsymbol{a}(t), \boldsymbol{g}_i(\boldsymbol{x}(t)))_{\mathbb{R}^m}.$$

下面考虑 $a_i(t)$, 有

$$\begin{aligned}a_i(t) = (\boldsymbol{a}(t), \boldsymbol{g}_i(\boldsymbol{x}(t)))_{\mathbb{R}^m} &= \left(\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t), \boldsymbol{g}_i(\boldsymbol{x}(t))\right)_{\mathbb{R}^m} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}(t), \boldsymbol{g}_i(\boldsymbol{x}(t)))_{\mathbb{R}^m} - \left(\boldsymbol{v}(t), \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}_i(\boldsymbol{x}(t))\right)_{\mathbb{R}^m}.\end{aligned}$$

为了计算上式, 引入二元向量值函数 $\hat{\boldsymbol{v}}(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \dot{x}^i\boldsymbol{g}_i(\boldsymbol{x})$, 式中的 $\boldsymbol{x}$ 和 $\dot{\boldsymbol{x}}$ 是相互独立的变量. 易见此函数满足如下性质:

1. $\hat{\boldsymbol{v}}(\boldsymbol{x}(t), \dot{\boldsymbol{x}}(t)) = \dot{x}^i(t)\boldsymbol{g}_i(\boldsymbol{x}(t)) = \boldsymbol{v}(t)$;

微积分讲稿 谢锡麟

2. $\dfrac{\partial \hat{\boldsymbol{v}}}{\partial x^j}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \dot{x}^i \dfrac{\partial \boldsymbol{g}_i}{\partial x^j}(\boldsymbol{x}) = \dot{x}^i \dfrac{\partial \boldsymbol{g}_j}{\partial x^i}(\boldsymbol{x})$, 所以有

$$\frac{\partial \hat{\boldsymbol{v}}}{\partial x^j}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)) = \dot{x}^i(t) \frac{\partial \boldsymbol{g}_j}{\partial x^i}(\boldsymbol{x}(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}_j(\boldsymbol{x}(t)),$$

式中 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 表示全导数;

3. $\dfrac{\partial \hat{\boldsymbol{v}}}{\partial \dot{x}^j}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \dfrac{\partial \dot{x}^i}{\partial \dot{x}^j}\boldsymbol{g}_i(\boldsymbol{x}) = \delta^i_j \boldsymbol{g}_i(\boldsymbol{x}) = \boldsymbol{g}_j(\boldsymbol{x})$, 所以有

$$\frac{\partial \hat{\boldsymbol{v}}}{\partial \dot{x}^j}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)) = \boldsymbol{g}_j(\boldsymbol{x}(t)).$$

再引入二元函数 $\hat{T}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \dfrac{1}{2}\left(\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}}),\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}})\right)_{\mathbb{R}^3} = \dfrac{1}{2}\left|\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}})\right|^2_{\mathbb{R}^3}$, 满足如下性质:

1. $\hat{T}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)) = \dfrac{1}{2}\left|\hat{\boldsymbol{v}}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t))\right|^2_{\mathbb{R}^3} = \dfrac{1}{2}\left|\boldsymbol{v}(t)\right|^2_{\mathbb{R}^3}$ 即为单位质量质点的动能;
2. $\dfrac{\partial \hat{T}}{\partial x^i}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \left(\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}}), \dfrac{\partial \hat{\boldsymbol{v}}}{\partial x^i}(\boldsymbol{x},\dot{\boldsymbol{x}})\right)_{\mathbb{R}^3}$;
3. $\dfrac{\partial \hat{T}}{\partial \dot{x}^i}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \left(\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}}), \dfrac{\partial \hat{\boldsymbol{v}}}{\partial \dot{x}^i}(\boldsymbol{x},\dot{\boldsymbol{x}})\right)_{\mathbb{R}^3}$.

利用函数 $\hat{\boldsymbol{v}}(\boldsymbol{x},\dot{\boldsymbol{x}})$、$\hat{T}(\boldsymbol{x},\dot{\boldsymbol{x}})$ 及其性质, 可得

$$\begin{aligned} a_i(t) =& \frac{\mathrm{d}}{\mathrm{d}t}\left(\hat{\boldsymbol{v}}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)), \frac{\partial \hat{\boldsymbol{v}}}{\partial \dot{x}^i}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t))\right)_{\mathbb{R}^3} - \left(\hat{\boldsymbol{v}}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)), \frac{\partial \hat{\boldsymbol{v}}}{\partial x^i}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t))\right)_{\mathbb{R}^3} \\ =& \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \hat{T}}{\partial \dot{x}^i}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)) - \frac{\partial \hat{T}}{\partial x^i}(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)). \end{aligned}$$

此方程称为**曲线坐标系下加速度表示的 Lagrange 方程**.

3.5　速度场的分解

4　建立路径

微积分讲稿 谢锡麟