2.4.6 向量值函数的微分

多元向量值函数如何求导/求微分？记住：求导是求变化率，微分是线性化。

¶ 一元向量值函数的微分

先考虑一元向量值函数$\overrightarrow{r}=\overrightarrow{r}(t)=\langle x(t),y(t)\rangle$。

- 怎么对$\overrightarrow{r}$求导？变化率就是差（因变量变化量）商（除以自变量变化量）的极限，而我们既可以对两个向量作差，也可以求向量与数的商，所以自然定义

$$\begin{aligned}\overrightarrow{r}'(t_0)&=\lim_{t\to t_0}\frac{\overrightarrow{r}(t)-\overrightarrow{r}(t_0)}{t-t_0}\\&=\lim_{t\to t_0}\frac{\langle x(t)-x(t_0),y(t)-y(t_0)\rangle}{t-t_0}\\&=\lim_{t\to t_0}\left\langle\frac{x(t)-x(t_0)}{t-t_0},\frac{y(t)-y(t_0)}{t-t_0}\right\rangle\\&=\left\langle\lim_{t\to t_0}\frac{x(t)-x(t_0)}{t-t_0},\lim_{t\to t_0}\frac{y(t)-y(t_0)}{t-t_0}\right\rangle\\&=\langle x'(t_0),y'(t_0)\rangle,\end{aligned}$$

  其中倒数第二个等数用到了命题2.2.10即“点列（向量）极限就是分量极限所对应的点”这个。

  注 2.4.22

  - 从几何上看，$\overrightarrow{r}'(t_0)$ 就是平面参数曲线 $\begin{cases}x=x(t)\\y=y(t)\end{cases}$ 在 $\overrightarrow{r}(t_0)$ 处的一个切向量。
  - 从物理上看，若t代表时间，$\overrightarrow{r}=\overrightarrow{r}(t)$代表质点位置，则$\overrightarrow{r}'(t_0)$恰好就是在时刻$t_0$位移的变化率，即瞬时速度向量。

- 函数$\overrightarrow{r}$的微分是什么？它是一个能“最佳逼近$\overrightarrow{r}=\overrightarrow{r}(t)$的线性映射”。由于其定义域为实数，而值为平面向量，所以它是形如

$$d\overrightarrow{r}:\mathbb{R}\to\mathbb{R}^2,\quad h\mapsto\langle ah,bh\rangle$$

  且“能最佳逼近$\overrightarrow{r}=\overrightarrow{r}(t)=\langle x(t),y(t)\rangle$的线性映射”。由此可见其“分量映射”$h\mapsto ah$ 是“能最佳逼近$x=x(t)$的线性映射”，即$ah=x'(t)h=dx(h)$。同理可知$b=y'(t)$。于是

$$d\overrightarrow{r}:\mathbb{R}\to\mathbb{R}^2,\quad h\mapsto\langle dx(h),dy(h)\rangle$$

也可以用(与t无关的)向量$\overrightarrow{i},\overrightarrow{j}$来表示和理解：

$$\overrightarrow{r}(t)=x(t)\overrightarrow{i}+y(t)\overrightarrow{j},\quad \overrightarrow{r}'(t)=x'(t)\overrightarrow{i}+y'(t)\overrightarrow{j},\quad d\overrightarrow{r}=dx\,\overrightarrow{i}+dy\,\overrightarrow{j}$$

为方便，我们也把向量写成列的形式：

$$\overrightarrow{r}(t)=\begin{pmatrix}x(t)\\y(t)\end{pmatrix},\overrightarrow{r}'(t)=\begin{pmatrix}x'(t)\\y'(t)\end{pmatrix},d\overrightarrow{r}=\begin{pmatrix}dx\\dy\end{pmatrix}$$

¶ 更一般映射的微分

考虑映射 $\overrightarrow{F}:\mathbb{R}^2\to\mathbb{R}^2,\quad \overrightarrow{F}(x,y)=\begin{pmatrix}f_1(x,y)\\f_2(x,y)\end{pmatrix}\Big(=f_1(x,y)\overrightarrow{i}+f_2(x,y)\overrightarrow{j}\Big)$。

- 怎么求导？因为有两个分量，所以可以固定一个变量，视$\overrightarrow{F}$为关于另一个变量的一元向量值函数，从而可以对另一个变量求偏导（沿着给定坐标轴方向的变化率），即

  $$\frac{\partial\overrightarrow{F}}{\partial x}=\begin{pmatrix}\frac{\partial f_1}{\partial x}\\\frac{\partial f_2}{\partial x}\end{pmatrix},\quad \frac{\partial\overrightarrow{F}}{\partial y}=\begin{pmatrix}\frac{\partial f_1}{\partial y}\\\frac{\partial f_2}{\partial y}\end{pmatrix}$$

  除了这两个“偏导数”向量，还可以求该映射沿着特定方向的变化率：对于单位向量$\overrightarrow{e}=\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$，有

  $$\frac{\partial\overrightarrow{F}}{\partial\overrightarrow{e}}=\begin{pmatrix}\frac{\partial f_1}{\partial\overrightarrow{e}}\\\frac{\partial f_2}{\partial\overrightarrow{e}}\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x}\cos\theta+\frac{\partial f_1}{\partial y}\sin\theta\\\frac{\partial f_2}{\partial x}\cos\theta+\frac{\partial f_2}{\partial y}\sin\theta\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x}&\frac{\partial f_1}{\partial y}\\\frac{\partial f_2}{\partial x}&\frac{\partial f_2}{\partial y}\end{pmatrix}\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}.$$

  这里我们看到“两个偏导数向量并列放在一起所得的矩阵”起了关键作用，这个矩阵称为映射$\overrightarrow{F}$的 **Jacobi 矩阵**：

  $$J(\overrightarrow{F})=\frac{\partial(f_1,f_2)}{\partial(x,y)}=\begin{pmatrix}\frac{\partial f_1}{\partial x}&\frac{\partial f_1}{\partial y}\\\frac{\partial f_2}{\partial x}&\frac{\partial f_2}{\partial y}\end{pmatrix}.$$

- 怎么求$\overrightarrow{F}$的微分？它是 $\overrightarrow{F}$ 的线性化，因而是一个线性映射

  $$d\overrightarrow{F}_{(x_0,y_0)}:\mathbb{R}^2\to\mathbb{R}^2,\quad \begin{pmatrix}h\\k\end{pmatrix}\longmapsto\Big(\quad ?\quad\Big)\begin{pmatrix}h\\k\end{pmatrix}$$

  线性映射是由矩阵给出的，这里(?)只能是 Jacobi 矩阵。【这一点也可由上述计算$\frac{\partial\overrightarrow{F}}{\partial\overrightarrow{e}}$的结果中看出来：$\overrightarrow{F}$沿着方向$\overrightarrow{e}$的线性化（最优逼近）是由 Jacobi 矩阵乘以向量$\overrightarrow{e}$ 给出，所以Jacobi 矩阵所代表的线性映射是$\overrightarrow{F}$整体的线性化（最优逼近）】

一般地，若 $F:\mathbb{R}^m\to\mathbb{R}^n$，$(x_1,\dots,x_m)\mapsto(f_1,\dots,f_n)$。则 Jacobi 矩阵为 $n\times m$ 矩阵：

$$JF=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}&\cdots&\frac{\partial f_1}{\partial x_m}\\\vdots&\ddots&\vdots\\\frac{\partial f_n}{\partial x_1}&\cdots&\frac{\partial f_n}{\partial x_m}\end{pmatrix}_{n\times m}.$$

而F的微分就是由矩阵JF给出的线性映射 $df:\mathbb{R}^m\to\mathbb{R}^n$，$\overrightarrow{h}\mapsto(JF)\overrightarrow{h}$。

¶ 复合映射的微分

设有映射 $F:\mathbb{R}^m\to\mathbb{R}^n$ 及 $G:\mathbb{R}^n\to\mathbb{R}^k$，且二者可复合（即 F 的像在 G 的定义域内）。则有复合映射 $G\circ F$。那么，F 在 P 处的微分 dF_P，G 在 $F(P)$ 处的微分 $dG_{F(P)}$，以及 $G\circ F$ 在 P 处的微分 $d(G\circ F)_P$，三者之间有什么关系？答案是自然的："$G\circ F$ 的线性化当然就是'G 在 $F(P)$ 处的线性化'与'F 在 P 处线性化'这两个线性映射的复合！"

定理 2.4.23: 链式法则 (Chain Rule)

设映射 $F:\mathbb{R}^m\to\mathbb{R}^n$ 与$G:\mathbb{R}^n\to\mathbb{R}^k$可复合（即 F 的像在 G 的定义域内），则

$$d(G\circ F)_P=(dG)_{F(P)}\circ(dF)_P.$$

对应到矩阵（线性映射的复合对应矩阵乘法）：

$$J(G\circ F)_P=J(G)_{F(P)}\cdot J(F)_P.$$

注意在上式中，$J(G\circ F)$是映射$G\circ F$在点P处的Jacobi矩阵，是一个$k\times m$ 矩阵；$J(G)_{F(P)}$是映射G在点$F(P)$处的Jacobi矩阵，是一个$k\times n$ 矩阵；$J(F)_P$是映射F在点P处的Jacobi矩阵，是一个$n\times m$ 矩阵。

还可以把这些矩阵全部写出来：若

$$\begin{aligned}G(y_1,\cdots,y_n)&=\langle g_1(y_1,\cdots,y_n),\cdots,g_k(y_1,\cdots,y_n)\rangle,\\F(x_1,\cdots,x_m)&=\langle f_1(x_1,\cdots,x_m),\cdots,f_n(x_1,\cdots,x_m)\rangle\end{aligned}$$

并记复合函数为

$$G\circ F(x_1,\cdots,x_m)=\langle h_1(x_1,\cdots,x_m),\cdots,h_k(x_1,\cdots,x_m)\rangle$$

则

$$\begin{pmatrix}\frac{\partial h_1}{\partial x_1}&\cdots&\frac{\partial h_1}{\partial x_m}\\\vdots&\ddots&\vdots\\\frac{\partial h_k}{\partial x_1}&\cdots&\frac{\partial h_k}{\partial x_m}\end{pmatrix}_{k\times m}=\begin{pmatrix}\frac{\partial g_1}{\partial x_1}&\cdots&\frac{\partial g_1}{\partial x_n}\\\vdots&\ddots&\vdots\\\frac{\partial g_k}{\partial x_1}&\cdots&\frac{\partial g_k}{\partial x_n}\end{pmatrix}_{k\times n}\begin{pmatrix}\frac{\partial f_1}{\partial x_1}&\cdots&\frac{\partial f_1}{\partial x_m}\\\vdots&\ddots&\vdots\\\frac{\partial f_n}{\partial x_1}&\cdots&\frac{\partial f_n}{\partial x_m}\end{pmatrix}_{n\times m}$$

注意这个矩阵方程事实上包含了 km个方程。

¶ 各种情形下的链式法则

下面列出多种不同情形复合的具体链式法则：

1. **全导数**：若 $\vec{r}(t)=\langle x(t),y(t)\rangle$ 是可微的，$f:\mathbb{R}^2\to\mathbb{R}$ 是可微二元函数，则复合函数$h(t)=f(x(t),y(t))$是可微的一元函数，且

$$h'(t)=\frac{\partial f}{\partial x}(x(t),y(t))x'(t)+\frac{\partial f}{\partial y}(x(t),y(t))y'(t)=\nabla f\cdot\vec{r}\,'.$$

例如，若 $f(x,y) = xy,\quad x = f_1(t), y = f_2(t)$，则$h(t) = f_1(t)f_2(t)$。由 $\frac{\partial f}{\partial x} = y, \frac{\partial f}{\partial y} = x$，得

$$(f_1(t)f_2(t))' = h'(t) = y(t)x'(t) + x(t)y'(t).$$

（注：这就像“高射炮打蚊子”，用二元微分推导一元乘积法则）。

2. 若 $f : \mathbb{R}^2 \to \mathbb{R}, g : \mathbb{R} \to \mathbb{R}$ 可微，则 $h = g \circ f : \mathbb{R}^2 \to \mathbb{R}$ 可微，且

$$\left(\frac{\partial(g \circ f)}{\partial x}, \frac{\partial(g \circ f)}{\partial y}\right) = \left(g'(f(x,y))\frac{\partial f}{\partial x}, \quad g'(f(x,y))\frac{\partial f}{\partial y}\right).$$

3. **变量代换**：若 $z = f(u,v)$，而 $u = u(x,y), v = v(x,y)$。则复合函数 $z = f(u(x,y), v(x,y))$（看作 x,y 的函数）的偏导数为

$$\begin{aligned}\frac{\partial z}{\partial x}(x,y) &= \frac{\partial z}{\partial u}(u(x,y),v(x,y))\frac{\partial u}{\partial x}(x,y) + \frac{\partial z}{\partial v}(u(x,y),v(x,y))\frac{\partial v}{\partial x}(x,y),\\ \frac{\partial z}{\partial y}(x,y) &= \frac{\partial z}{\partial u}(u(x,y),v(x,y))\frac{\partial u}{\partial y}(x,y) + \frac{\partial z}{\partial v}(u(x,y),v(x,y))\frac{\partial v}{\partial y}(x,y).\end{aligned}$$

上式用矩阵可简记为

$$\begin{pmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial u} \\ \frac{\partial f}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

4. **一般情形**：若 $y = y(x_1, \ldots, x_m)$，且 $x_i = x_i(t_1, \ldots, t_n)$。则

$$\frac{\partial y}{\partial t_k} = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i}\frac{\partial x_i}{\partial t_k} = \frac{\partial y}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \cdots + \frac{\partial y}{\partial x_m}\frac{\partial x_m}{\partial t_k}.$$

5. 再来一个稍微不一样的例子：$h(x,y) = f(u(x,y), v(x,y), w(x,y), x)$。注意这里 f 有四个槽位，第四个槽位直接是 x。求 $\frac{\partial h}{\partial x}$ 时，既要考虑 u,v,w 带来的变化，也要考虑显式 x 的变化。

$$\begin{aligned}\frac{\partial h}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} + \frac{\partial f}{\partial x}\\ \frac{\partial h}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial y}.\end{aligned}$$

（注：第一个公式中最后一个 $\frac{\partial f}{\partial x}$ 指的是 f 对其第四个变量求偏导）。

¶ 例子：极坐标下的拉普拉斯算子

例 2.4.24: 极坐标变换

设 $z = f(x,y)$，其中f可微。在极坐标下 $\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \rightsquigarrow z = f(r\cos\theta, r\sin\theta)$。下面用直角坐标下的一阶偏导和二阶偏导表示极坐标下的一阶偏导和二阶偏导。

解： **1. 一阶偏导数：**

$$\frac{\partial z}{\partial r}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial r}=\frac{\partial z}{\partial x}\cos\theta+\frac{\partial z}{\partial y}\sin\theta$$
$$\frac{\partial z}{\partial \theta}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta}=\frac{\partial z}{\partial x}(-r\sin\theta)+\frac{\partial z}{\partial y}(r\cos\theta)$$

2. 二阶偏导数：

$$\begin{aligned}\frac{\partial^2 z}{\partial r^2}&=\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial x}\cos\theta+\frac{\partial z}{\partial y}\sin\theta\right)\\&=\left(\frac{\partial^2 z}{\partial x^2}\cos\theta+\frac{\partial^2 z}{\partial y\partial x}\sin\theta\right)\cos\theta+\left(\frac{\partial^2 z}{\partial x\partial y}\cos\theta+\frac{\partial^2 z}{\partial y^2}\sin\theta\right)\sin\theta\\&=\frac{\partial^2 z}{\partial x^2}\cos^2\theta+2\frac{\partial^2 z}{\partial x\partial y}\cos\theta\sin\theta+\frac{\partial^2 z}{\partial y^2}\sin^2\theta\end{aligned}$$

对于 $\frac{\partial^2 z}{\partial\theta^2}$，注意求导时不仅要对 z_x, z_y 求导（链式法则），还要对系数 $-r\sin\theta$ 等求导（乘积法则）：

$$\begin{aligned}\frac{\partial^2 z}{\partial \theta^2}&=\frac{\partial}{\partial \theta}\left(z_x(-r\sin\theta)+z_y(r\cos\theta)\right)\\&=\frac{\partial(z_x)}{\partial\theta}(-r\sin\theta)+z_x\frac{\partial}{\partial\theta}(-r\sin\theta)+\frac{\partial(z_y)}{\partial\theta}(r\cos\theta)+z_y\frac{\partial}{\partial\theta}(r\cos\theta)\\&=\left(z_{xx}(-r\sin\theta)+z_{xy}(r\cos\theta)\right)(-r\sin\theta)-z_x(r\cos\theta)\\&\quad+\left(z_{yx}(-r\sin\theta)+z_{yy}(r\cos\theta)\right)(r\cos\theta)-z_y(r\sin\theta)\\&=r^2\left(z_{xx}\sin^2\theta-2z_{xy}\sin\theta\cos\theta+z_{yy}\cos^2\theta\right)-r\underbrace{\left(z_x\cos\theta+z_y\sin\theta\right)}_{=\partial z/\partial r}\end{aligned}$$

3. 比较与合并： 将 $\frac{\partial^2 z}{\partial r^2}$ 与 $\frac{1}{r^2}\frac{\partial^2 z}{\partial\theta^2}$ 相加，交叉项抵消，$\sin^2+\cos^2=1$，整理得：

$$\frac{\partial^2 z}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 z}{\partial\theta^2}=\left(\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right)-\frac{1}{r}\frac{\partial z}{\partial r}$$

移项即得如下重要公式

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=\frac{\partial^2 z}{\partial r^2}+\frac{1}{r}\frac{\partial z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 z}{\partial\theta^2}$$

□

定义 2.4.25: Laplace算子

称

$$\Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\ \text{或}\ \Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$$

为Laplace算子。

它是数学和物理中最最最重要的微分算子，没有之一。