9.2.4 复合函数的微分和一阶微分形式不变形

以二元函数为例, 设

$$u = f(\xi, \zeta)$$

定义在区域 D 上的一个二元函数,

$$\xi = \phi(x, y),\ \zeta = \psi(x, y)$$

定义在区域 Δ 上. 为了使复合函数

$$u = f(\phi(x, y), \psi(x, y))$$

有意义, 则要假设

$$(\phi(x, y), \psi(x, y)) \in D$$

对任何 $(x, y) \in \Delta$ 成立. 以下总是作这样的假设.

为了书写的简明和可推广性, 我们有时也采用映射的写法. 记

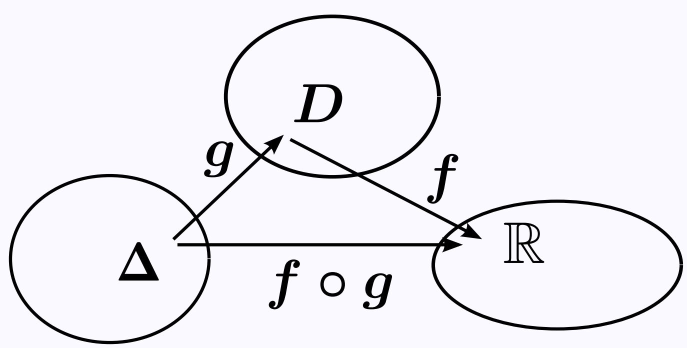
$$\boldsymbol{g} = (\xi, \zeta),$$

这里 ξ, ζ 都是定义在 $\mathbb{R}^2$ 中区域 $\boldsymbol{\Delta}$ 上的二元函数. 这样

$$\boldsymbol{g} : \boldsymbol{\Delta} \to \boldsymbol{D} \subset \mathbb{R}^2$$

是一个映射.

设 D 也是 $\mathbb{R}^2$ 中区域, 且 $g(\Delta) \subset D$. 若二元函数 f 在 D 上有定义, 则 $f \circ g$ 就是定义在 Δ 上的二元函数.



在讨论可微性之前, 首先要解决复合函数的连续性问题.

定理 1 设 Δ 和 D 都是 $\mathbb{R}^2$ 中区域. 若 $g:\Delta\to D$ 在 $x_0\in\Delta$ 连续, $f:D\to\mathbb{R}$ 在 $y_0=g(x_0)$ 连续, 则复合函数 $f\circ g:\Delta\to\mathbb{R}$ 在 x_0 连续.

证明 因为 f 在 y_0 连续, 所以对任意 $\varepsilon>0$ 存在 $\delta_1>0$, 使得当 $y\in B(y_0,\delta_1)\cap D$ 时, 有

$$|f(y)-f(y_0)|<\varepsilon.$$

因为 g 在 x_0 连续, 所以对 δ_1 存在 $\delta>0$ 使得当 $x\in B(x_0,\delta)\cap\Delta$ 时, 有

$$|g(x)-g(x_0)|<\delta_1,$$

即, $y=g(x)\in B(y_0,\delta_1)\cap D$. 由此知当 $x\in B(x_0,\delta)\cap\Delta$ 时, 有

$$|f\circ g(x)-f\circ g(x_0)|<\varepsilon.$$

这就说明 $f\circ g$ 在 x_0 连续.

定理 2 设 $\boldsymbol{\Delta}$ 和 $\boldsymbol{D}$ 都是 $\mathbb{R}^2$ 中区域. 若 $\boldsymbol{g}:\boldsymbol{\Delta}\to\boldsymbol{D}$ 在 $\boldsymbol{x}\in\boldsymbol{\Delta}$ 可微, $\boldsymbol{f}:\boldsymbol{D}\to\mathbb{R}$ 在 $\boldsymbol{y}=\boldsymbol{g}(\boldsymbol{x})$ 可微, 则复合函数 $\boldsymbol{f}\circ\boldsymbol{g}:\boldsymbol{\Delta}\to\mathbb{R}$ 在 $\boldsymbol{x}$ 可微, 且

$$J(f\circ g)(x)=Jf(y)Jg(x). \tag{9.1}$$

.

公式 (9.1) 表示了复合函数 $f\circ g$ 的偏导数与 f 和 g 的偏导数之间的关系, 称为复合函数偏导数的**链式法则**. 设 $x=(x_1,x_2)$, $y=g(x)$, 即, $y_1=g_1(x_1,x_2)$, $y_2=g_2(x_1,x_2)$, $z=f(y_1,y_2)=f\circ g$, 则 (9.1) 就是

$$\left(\frac{\partial z}{\partial x_1},\frac{\partial z}{\partial x_2}\right)=\left(\frac{\partial z}{\partial y_1},\frac{\partial z}{\partial y_2}\right)\begin{pmatrix}\dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2}\\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2}\end{pmatrix} \tag{9.2}$$

展开就是

$$\frac{\partial z}{\partial x_1}=\frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial x_1}+\frac{\partial z}{\partial y_2}\frac{\partial y_2}{\partial x_1} \tag{9.3}$$

$$\frac{\partial z}{\partial x_2}=\frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial x_2}+\frac{\partial z}{\partial y_2}\frac{\partial y_2}{\partial x_2} \tag{9.4}$$

证明 对于 $x \in \Delta$ 取 $h \in \mathbb{R}^2$ 的模充分小, 使 $x + h \in \Delta$. 因为 g 在 x 可微, 所以

$$g_1(x+h) - g_1(x) = Jg_1(x) \cdot h + o(h) \tag{9.5}$$

$$g_2(x+h) - g_2(x) = Jg_2(x) \cdot h + o(h). \tag{9.6}$$

g 的可微性保证了它的连续性, 因此当 h 趋于零时, $k = g(x+h) - g(x)$ 也趋于零, 这里把 k 看出列向量. 因为 f 在 $y = g(x)$ 可微, 所以

$$f(y+k) - f(y) = Jf(y) \cdot k + o(k). \tag{9.7}$$

(9.5) 和 (9.6) 合并起来可以写成 $k = Jg(x)h + o(h)$, 将之代入 (9.7) 并注意到 $o(k) = o(h)$, 就得到

$$f(g(x+h)) - f(g(x)) = Jf(y) \cdot \big(Jg(x) \cdot h + o(h)\big) + o(k) = Jf(y) \cdot Jg(x) \cdot h + o(h).$$

这就表明复合函数 $f \circ g$ 在 x 可微, 且 (9.1) 成立.

一阶微分形式的不变性

回到二元函数的情形. 函数 $u = f(\xi, \zeta) = f(\phi(x, y), \psi(x, y))$ 既是变量 ξ, ζ 的函数, 也是变量 x, y 的函数, 因此关于它的微分形式有两种表示方式, 其一是 $d\xi$, $d\zeta$ 的线性组合, 其二是 dx, dy 的线性组合

$$du = \frac{\partial u}{\partial \xi} d\xi + \frac{\partial u}{\partial \zeta} d\zeta, \quad du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$$

注意到, 当 ξ, ζ 分别是x, y 的函数时

$$d\xi = \frac{\partial \xi}{\partial x} dx + \frac{\partial \xi}{\partial y} dy \quad d\zeta = \frac{\partial \zeta}{\partial x} dx + \frac{\partial \zeta}{\partial y} dy$$

上述关于 du 的两种微分形式的表达方式是相等的

$$\begin{aligned} du &= \frac{\partial u}{\partial \xi} d\xi + \frac{\partial u}{\partial \zeta} d\zeta \\ &= \frac{\partial u}{\partial \xi} \left(\frac{\partial \xi}{\partial x} dx + \frac{\partial \xi}{\partial y} dy \right) + \frac{\partial u}{\partial \zeta} \left(\frac{\partial \zeta}{\partial x} dx + \frac{\partial \zeta}{\partial y} dy \right) \\ &= \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \end{aligned}$$

这就是所谓的一阶微分形式的不变性.

例 1 设 $u = f(x, \xi, \eta)$, $\xi = x^2 + y^2$, $\eta = x^2 + y^2 + z^2$. 求复合函数的全微分 du.

解 根据一阶微分形式的不变性, 有

$$\begin{aligned} du &= f_1' dx + f_2' d\xi + f_3' d\eta \\ &= f_1' dx + f_2'(2x dx + 2y dy) + f_3'(2x dx + 2y dy + 2z dz) \\ &= (f_1' + 2x f_2' + 2x f_3') dx + 2y(f_2' + f_3') dy + 2z f_3' dz. \end{aligned}$$

例 2 解微分方程 $xydz + yzdx + zxdy = 0$.

解 令 $u = xyz$. 则

$$du = xydz + yzdx + zxdy = 0.$$

根据推论 ??, 知 u 为常数, 即, $xyz = c$. 这就是所给微分方程的通积分.

9.2.5 向量值函数的微分与微商

设 $\vec{r}(t)$ ($t \in [\alpha, \beta]$) 是空间中一元向量值函数. 如果极限

$$\vec{r}'(t) = \frac{d\vec{r}}{dt}(t) := \lim_{\Delta t \to 0} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{\Delta t}$$

存在, 那么就称之为 $\vec{r}(t)$ 关于 t 的导数, 它还是一个向量.

从几何上看一般来说 $\vec{r}(t)$ 的轨迹是空间中一条曲线, 而 $\vec{r}'(t)$ 就是在点 $\vec{r}(t)$ 处的切线方向, 并指向参数增加的方向. 可以这样看, 设曲线上两点 M_0 和 M 对应的位置向量分别是 $\vec{r}(t_0)$ 和 $\vec{r}(t)$. 于是割线向量 $\overrightarrow{M_0M} = \vec{r}(t) - \vec{r}(t_0)$, 故 $\frac{\vec{r}(t)-\vec{r}(t_0)}{t-t_0}$ 就与 $\overrightarrow{M_0M}$ 共线, 并指向参数的增加方向. 当 $t \to t_0$ 时, $\vec{r}(t)$ 在 t_0 的导数 $\vec{r}'(t_0)$ 就是曲线在 M_0 的切向量, 切向量的方向指向参数增加的方向.

在直角坐标系中, 如果 $\vec{r}(t)$ 的坐标是 $(x(t), y(t), z(t))$, 则

$$\vec{r}'(t) = (x'(t), y'(t), z'(t)).$$

若 $x(t), y(t), z(t)$ 有二阶导数, 则 $\vec{r}(t)$ 也有二阶导数, 且

$$\vec{r}''(t) = (x''(t), y''(t), z''(t)).$$

设 $\vec{a}(t), \vec{b}(t)$ 是空间中可导的向量值函数, $f(t)$ 是可导函数, 则有

$$\frac{d}{dt}(f\vec{a}) = f\frac{d\vec{a}}{dt} + \frac{df}{dt}\vec{a}, \tag{9.8}$$

$$\frac{d}{dt}(\vec{a}\cdot\vec{b}) = \frac{d\vec{a}}{dt}\cdot\vec{b} + \vec{a}\cdot\frac{d\vec{b}}{dt}, \tag{9.9}$$

$$\frac{d}{dt}(\vec{a}\times\vec{b}) = \frac{d\vec{a}}{dt}\times\vec{b} + \vec{a}\times\frac{d\vec{b}}{dt}. \tag{9.10}$$

设 $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 是一个 n 元向量值函数, 即

$$f(x) = (f_1(x), \cdots, f_m(x))^T,\ x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$$

其中的每一个分量 $y_j = f_j(x)$ 都是一个 n 元函数, T 表示转置. 我们定义 f 的微分就是对每一个分量的微分, 因此有

$$df(x) = (dy_1(x), \cdots, dy_m(x))^T$$

返回 全屏 关闭 退出

利用 n 元函数微分的规则

$$dy_j = \sum_{i=1}^{n} \frac{\partial y_j}{\partial x_i} dx_i$$

并利用线性代数中关于矩阵的运算, 我们有

$$df = (dy_1, \cdots, dy_m)^T = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = (Jf)dx,$$

其中

$$Jf = J_x(y) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

称之为向量值函数的 Jacobi 矩阵.

当 $n = m$ 时, Jacobi 矩阵的行列式简记为

$$\det J_x(y) = \frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}$$

返回 全屏 关闭 退出

并称之为函数的 Jacobi 行列式. 这种 $\mathbb{R}^n \to \mathbb{R}^n$ 向量值函数又称为坐标变换.

特别, 对于 $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ 可微的向量值函数

$$f = (x(u,v), y(u,v)), \quad \text{或} \quad x = x(u,v),\ y = y(u,v)$$

$$Jf = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}, \quad \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

定理 3 设有区域 $D \subset \mathbb{R}^n$, 映射 $f, g: D \to \mathbb{R}^m$ 在 D 中可微, 则有

1° $J(cf) = cJf$, 其中 c 是常数;

2° $J(f+g) = Jf + Jg$;

3° $J(f \cdot g) = g^T(Jf) + f^T(Jg)$, 其中 $f \cdot g$ 表示 $\mathbb{R}^m$ 中的内积, 右端的 f, g 理解为列向量.

定理 4 设 $\Delta \subset \mathbb{R}^n$ 和 $D \subset \mathbb{R}^m$ 都是区域. 若 $g : \Delta \to D$ 在 $x \in \Delta$ 可微, $f : D \to \mathbb{R}^l$ 在 $y = g(x)$ 可微, 则复合映射 $f \circ g : \Delta \to \mathbb{R}^l$ 在 x 可微, 且

$$J(f \circ g)(x) = Jf(y)Jg(x). \tag{9.11}$$

.

用图表示就是, 当

$$(x_1, \cdots x_n) \xrightarrow{g} (y_1, \cdots, y_m) \xrightarrow{f} (z_1, \cdots, z_l)$$

都可微时, $f \circ g$ 也可微. (9.11) 就是

$$\begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial z_l}{\partial x_1} & \cdots & \frac{\partial z_l}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \cdots & \frac{\partial z_1}{\partial y_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial z_l}{\partial y_1} & \cdots & \frac{\partial z_l}{\partial y_m} \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \tag{9.12}$$

例 3 首先回顾一下方向导数的定义. 所谓在一点 (x_0, y_0) 的方向导数其实就是复合函数

$$u = f(x_0 + t\cos\alpha, y_0 + t\cos\beta)$$

对变量 t 求导, 这里 $\vec{e} = (\cos\alpha, \cos\beta)$ 是单位向量. 所以根据复合函数求导规则, 有

$$\frac{\partial u}{\partial \vec{e}} = \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\cos\beta$$

一般来说, 即使方向不是单位向量, 如对一般方向向量 $\vec{l} = (l_1, l_2)$, 仍然有

$$\frac{\partial u}{\partial \vec{l}} = \frac{df(x_0 + tl_1, x_0 + tl_2)}{dt} = \frac{\partial f}{\partial x}l_1 + \frac{\partial f}{\partial y}l_2$$

注意, 我们通常习惯用 “d” 表示对只有一个变量的导数, 而用 “∂” 表示对多个变量的偏导数.

例 4 设 $u = f(x, y, z)$, $y = \varphi(x, r)$, $z = \psi(x, y, r)$. 求 $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial r}$.

解 这里经过复合后, 把 u 当成自变量 x, r 的二元函数

$$\begin{aligned} u &= f(x, y, z) \\ &= f(x,\ \varphi(x, r),\ \psi(x, y, r)) \\ &= f(x,\ \varphi(x, r),\ \psi(x, \varphi(x, r), r)) \\ &= u(x, r). \end{aligned}$$

因此我们有

$$\begin{aligned} \frac{\partial u}{\partial x} &= f'_x + f'_y\varphi'_x + f'_z(\psi'_x + \psi'_y\varphi'_x), \\ \frac{\partial u}{\partial r} &= f'_y\varphi'_r + f'_z(\psi'_y\varphi'_r + \psi'_r). \end{aligned}$$

复合函数的高阶偏导数

在计算复合函数的高阶偏导数时, 只要反复应用复合函数求导的链式法则即可.

例 5 设 $f(x,y)$ 有连续的二阶偏导数. 求复合函数 $z=f\left(xy,\frac{y}{x}\right)$ 的三个二阶偏导数.

解 复合函数 $z=f\left(xy,\frac{y}{x}\right)$ 是 x,y 的二元函数, 一阶偏导数为

$$\frac{\partial z}{\partial x}=yf_1'-\frac{y}{x^2}f_2',\quad \frac{\partial z}{\partial y}=xf_1'+\frac{1}{x}f_2',$$

二阶偏导数为

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2}&=y\left(yf_{11}''-\frac{y}{x^2}f_{12}''\right)-\frac{y}{x^2}\left(yf_{21}''-\frac{y}{x^2}f_{22}''\right)+\frac{2y}{x^3}f_2'\\&=y^2f_{11}''-\frac{2y^2}{x^2}f_{12}''+\frac{y^2}{x^4}f_{22}''+\frac{2y}{x^3}f_2',\end{aligned}$$

同理, 有

$$\frac{\partial^2 z}{\partial x\partial y}=xyf_{11}''-\frac{y}{x^3}f_{22}''+f_1'-\frac{1}{x^2}f_2',\quad \frac{\partial^2 z}{\partial y^2}=x^2f_{11}''+2f_{12}''+\frac{1}{x^2}f_{22}''.$$

例 6 设 $\varphi(x), \psi(x)$ 二阶连续可导. 求证: 函数

$$u = \varphi(x - at) + \psi(x + at)$$

满足波动方程

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}.$$

证明 显然 u 是二阶连续可导的二元函数.

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \varphi'(x - at) + \psi'(x + at), \\
\frac{\partial u}{\partial t} &= -a\varphi'(x - at) + a\psi'(x + at), \\
\frac{\partial^2 u}{\partial x^2} &= \varphi''(x - at) + \psi''(x + at), \\
\frac{\partial^2 u}{\partial t^2} &= a^2\varphi''(x - at) + a^2\psi''(x + at) \\
&= a^2 \frac{\partial^2 u}{\partial x^2}.
\end{aligned}$$