高等微积分讲义 (下)

目录

第一章　连续映射

1.1　多元函数

多元微积分研究多元函数. 我们先定义多元函数. 例如, $f(x,y)=x^2+y$ 是二元函数, 对每个有序对 (x,y), f 赋予函数值 x^2+y.

定义 1.1.1. 定义 n 维欧氏空间 $\mathbf{R}^n$ 为

$$\mathbf{R}^n=\{(x_1,...,x_n)|x_1,...,x_n\in\mathbf{R}\}.$$

把 $\mathbf{R}^n$ 中的元素 $(x_1,...,x_n)$ 称为 n 维欧氏空间中的点, 简记为 $\mathbf{x}=(x_1,...,x_n)$, 称 x_i 为 $\mathbf{x}$ 的第 i 个坐标分量. 经常将 $\mathbf{x}$ 视为一个列向量

$$\mathbf{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

一般的, 对于集合 X,Y, 定义它们的笛卡尔积为所有有序对构成的集合 $X\times Y=\{(x,y)|x\in X,y\in Y\}$, 这样

$$\mathbf{R}^n=\underbrace{\mathbf{R}\times\cdots\times\mathbf{R}}_{n\text{个}}.$$

定义 1.1.2. 设 $D\subset\mathbf{R}^n$. 称映射

$$\begin{aligned}&f:D\to\mathbf{R};\\&(x_1,...,x_n)\to f(x_1,...,x_n)\end{aligned}$$

为 D 上的 n 元函数, 称 D 为函数 f 的定义域.

例 1.1.3. (1) 给定 $1\le i\le n$, 定义第 i 个投影映射 $p_i:\mathbf{R}^n\to\mathbf{R}$ 为

$$p_i(x_1,...,x_n)=x_i.$$

(2) 线性函数 $f(x_1,...,x_n)=\sum_{i=1}^{n} a_i x_i$.

(3) 二次函数 $Q(x_1,...,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$.

定义 1.1.4. 设 $D\subset \mathbf{R}^n$, 称映射 $f:D\to\mathbf{R}^m$ 为 D 上的多元映射. 对于 $1\le i\le m$, 令

$$f_i=p_i\circ f:D\to\mathbf{R},$$

称之为 f 的第 i 个分量函数. f 由它的 m 个分量 $f_1,...,f_m:D\to\mathbf{R}$ 唯一确定

$$f(\mathbf{x})=(f_1(\mathbf{x}),...,f_m(\mathbf{x})),\quad \forall \mathbf{x}\in D.$$

例 1.1.5. 给定 $m\times n$ 的矩阵 $A=(A_{ij})_{1\le i\le m,1\le j\le n}$, 它确定线性映射 $f:\mathbf{R}^n\to\mathbf{R}^m$

$$f(\mathbf{x})=A\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix},$$

用分量函数表示为

$$f(x_1,...,x_n)=(\sum_{j=1}^{n}A_{1j}x_j,...,\sum_{j=1}^{n}A_{mj}x_j).$$

定义 1.1.6. 设 $f:D\to\mathbf{R}^m$ 是映射. 对于 $U\subset D$, 令

$$f(U)=\{\mathbf{y}\in\mathbf{R}^m|\text{存在}\mathbf{x}\in U,\text{使得}\mathbf{y}=f(\mathbf{x})\},$$

称之为 U 在函数 f 下的像集 (image). 对于 $V\subset\mathbf{R}^m$, 令

$$f^{-1}(V)=\{\mathbf{x}\in D:f(\mathbf{x})\in V\}$$

称之为 V 在 f 下的原像集 (pre-image).

1.2 欧氏空间

定义 1.2.1. 集合 X 上的一个度量是指一个映射 $d:X\times X\to\mathbf{R}$, 满足:

(1) 正定性. 对任何 $x,y\in X$ 有 $d(x,y)\ge 0$, 且 $d(x,y)=0$ 当且仅当 $x=y$.

(2) 对称性. 对任何 $x,y\in X$ 有 $d(x,y)=d(y,x)$.

(3) 三角不等式. 对任何 $x,y,z\in X$ 有 $d(x,z)\le d(x,y)+d(y,z)$.

例 1.2.2 (欧氏度量)**.** 设 $\mathbf{x} = (x_1, ..., x_n), \mathbf{y} = (y_1, ..., y_n) \in \mathbf{R}^n$, 定义它们之间的距离为

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

欧氏度量可以用内积叙述. 定义 $\mathbf{x}$ 与 $\mathbf{y}$ 的内积为

$$\mathbf{x} \cdot \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i,$$

则有

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}.$$

特别的, $\mathbf{x}$ 到原点 $\mathbf{0}$ 的距离为 $||\mathbf{x}|| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, 称之为 $\mathbf{x}$ 的模长.

我们来验证前述定义的欧氏度量满足三角不等式. 三个点 $\mathbf{0}, \mathbf{x}, -\mathbf{y}$ 之间的三角不等式为

$$\begin{aligned}
& d(\mathbf{x}, -\mathbf{y}) \leq d(\mathbf{x}, \mathbf{0}) + d(\mathbf{0}, -\mathbf{y}) \\
\iff & \sqrt{\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle} \leq ||\mathbf{x}|| + ||\mathbf{y}|| \\
\iff & \langle \mathbf{x}, \mathbf{y} \rangle \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||
\end{aligned}$$

这是所谓的 Cauchy-Schwartz 不等式.

定理 1.2.3. *设* $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$*, 则有* $\langle \mathbf{x}, \mathbf{y} \rangle \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||$*, 亦即*

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 \leq \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{i=1}^{n} y_i^2 \right).$$

证明. 考虑关于 t 的二次函数

$$\begin{aligned}
f(t) &= \langle t\mathbf{x} + \mathbf{y}, t\mathbf{x} + \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle t^2 + 2 \langle \mathbf{x}, \mathbf{y} \rangle t + \langle \mathbf{y}, \mathbf{y} \rangle,
\end{aligned}$$

由于对任何 $t \in \mathbf{R}$ 都有 $f(t) \geq 0$, 则 f 的判别式小于等于 0, 即有

$$\Delta = (2 \langle \mathbf{x}, \mathbf{y} \rangle)^2 - 4 \langle \mathbf{x}, \mathbf{x} \rangle \cdot \langle \mathbf{y}, \mathbf{y} \rangle \leq 0,$$

由此可得

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||.$$

□

完全类似的, 可得积分版本的 Cauchy-Schwartz 不等式.

命题 1.2.4. 设 f,g 是 $[a,b]$ 上的可积函数, 则有

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \left(\int_a^b f(x)^2dx\right)\cdot\left(\int_a^b g(x)^2dx\right).$$

练习 1.2.5. 设 A,B 是 $n\times n$ 的对称方阵, $\mathbf{v}\in\mathbf{R}^n$ 是给定的向量, 考虑二次函数

$$g(t)=\langle tA\mathbf{v}+B\mathbf{v},tA\mathbf{v}+B\mathbf{v}\rangle\geq 0,\quad \forall t\in\mathbf{R},$$

可得如下不等式

$$\frac{1}{2}\langle(AB+BA)\mathbf{v},\mathbf{v}\rangle\leq\sqrt{\langle A^2\mathbf{v},\mathbf{v}\rangle\cdot\langle B^2\mathbf{v},\mathbf{v}\rangle}.$$

此不等式在复数域上的版本就是所谓的 Uncertainty Principle.

1.2.1 $\mathbf{R}^n$ 的拓扑 (topology)

定义 1.2.6. 设 $\mathbf{a}\in\mathbf{R}^n$. 对于正数 r, 令

$$B_r(\mathbf{a})=B(\mathbf{a};r)=\{\mathbf{x}|d(\mathbf{a},\mathbf{x})<r\},$$

称之为 $\mathbf{a}$ 的开球邻域; 称 $B_r(\mathbf{a})\backslash\{\mathbf{a}\}$ 为 $\mathbf{a}$ 的去心开球邻域.

定义 1.2.7. 设 $X\subset\mathbf{R}^n$, $\mathbf{a}\in X$. 称 $\mathbf{a}$ 是 X 的一个内点, X 是 $\mathbf{a}$ 的一个邻域 (*neighborhood*), 如果存在 $r>0$ 使得 $B_r(\mathbf{a})\subset X$.

定义 1.2.8. 称 U 是 $\mathbf{R}^n$ 的一个开集 (*open set*), 如果 U 中每个点都是 U 的内点. 换句话说, 即对任何 $\mathbf{x}\in U$, 存在 $r>0$, 使得 $B_r(\mathbf{x})\subset U$.

例 1.2.9. (1) $\emptyset$ 与 $\mathbf{R}^n$ 是开集.

(2) 开球邻域 $B_r(\mathbf{x}_0)$ 是开集.

(3) U 是 $\mathbf{R}^n$ 的开集当且仅当 U 可以表示成一族开球邻域的并

$$U=\bigcup_{\lambda\in\Lambda}B(\mathbf{x}_\lambda;r_\lambda).$$

证明*:* 只证明 (3). 一方面, 如果 U 是开集, 则对每个 $\mathbf{x}\in U$, 存在 $B_{r(\mathbf{x})}(\mathbf{x})\subset U$, 由此可得

$$U=\bigcup_{\mathbf{x}\in U}\{\mathbf{x}\}\subset\bigcup_{\mathbf{x}\in U}B_{r(\mathbf{x})}(\mathbf{x})\subset U,$$

即有 $U=\bigcup_{\mathbf{x}\in U}B_{r(\mathbf{x})}(\mathbf{x})$ 是一族开球邻域的并. 另一方面, 如果 $U=\bigcup_{\lambda\in\Lambda}B(\mathbf{x}_\lambda;r_\lambda)$, 对每个点 $\mathbf{x}\in U$, 它属于某个 $B(\mathbf{x}_\lambda;r_\lambda)$, 取正数 r 满足 $r<r_\lambda-d(\mathbf{x},\mathbf{x}_\lambda)$, 则有

$$B_r(\mathbf{x})\subset B(\mathbf{x}_\lambda;r_\lambda)\subset U,$$

这就验证了 U 的每个点都是它的内点, 因而 U 是开集.

$\square$

命题 1.2.10. $\mathbf{R}^n$ *的开集满足如下的拓扑公理:*

(1) *空集与全集是开集.*

(2) *任何一族开集的并集是开集.*

(3) *任何有限个开集的交集是开集.*

定义 1.2.11. 称 $B \subset \mathbf{R}^n$ 是 ($\mathbf{R}^n$ 的) 闭集 (*closed set*), 如果 $\mathbf{R}^n \backslash B$ 是开集.

例 1.2.12. n 维闭球 (*disk*)

$$\overline{B^n} = \{\mathbf{x} = (x_1, ..., x_n) | \sum_{i=1}^n x_i^2 \leq 1\}$$

与 $(n-1)$ 维球面 (*sphere*)

$$S^{n-1} = \{\mathbf{x} = (x_1, ..., x_n) | \sum_{i=1}^n x_i^2 = 1\}$$

都是 $\mathbf{R}^n$ 的闭集.

我们把一般的闭球记作

$$\overline{B}_r(\mathbf{a}) = \{\mathbf{x} \in \mathbf{R}^n | d(\mathbf{a}, \mathbf{x}) \leq r\}.$$

为什么要引入开集的概念呢? 人们希望对一般的几何空间 X, Y, 定义映射 $f: X \to Y$ 的连续性. 粗略的说, 所谓 f 在 $x_0 \in X$ 处连续, 是指 f 把 x_0 某个附近中的所有点都映到 $f(x_0)$ 的附近. 这样, 为了定义连续性, 人们需要先定义好 "附近", 如果 X 上有度量, 可以定义 x_0 的附近为距 x_0 很近的点所构成的集合, 比如前述所说的开球邻域, 由此就可讨论度量空间之间映射的连续性. 更一般的, 如果 X 上没有给定度量呢? 人们想到的办法是: 把 X 中每个点的各种程度的附近都指定出来, 这么一个指定称之为 X 的一个拓扑结构. 这就启发了如下定义.

定义 1.2.13. 集合 X 上的一个拓扑结构是指一个满足如下条件的子集族 $\mathcal{T} \subset 2^X$:

(1) $\emptyset, X$ 属于 $\mathcal{T}$.

(2) $\mathcal{T}$ 中有限个成员的交仍然属于 $\mathcal{T}$, 即对任何 $U_1, ..., U_n \in \mathcal{T}$, 有 $\cap_{i=1}^n U_i \in \mathcal{T}$.

(3) $\mathcal{T}$ 中任意多个成员的并仍然属于 $\mathcal{T}$, 即对任何 $U_i \in \mathcal{T}, i \in I$, 有 $\cup_{i \in I} U_i \in \mathcal{T}$.

称 $\mathcal{T}$ 中成员为这个拓扑结构下的开集 (*open set*), 称有序对 $(X, \mathcal{T})$ 为一个拓扑空间, 在不引起混淆的情况下, 也把拓扑空间简记为 X.

定义 1.2.14. 设 $D \subset \mathbf{R}^n$. 称 D 是有界的, 如果存在某个 $R > 0$, 使得 $D \subset B_R(\mathbf{0})$.

1.3　极限

回忆一元函数的极限. 设 $f : D \to \mathbf{R}$, 则

$$\lim_{x \to x_0} f(x) = A \iff \forall \epsilon > 0, \exists \delta > 0, \forall 0 < |x - x_0| < \delta, 有 |f(x) - A| < \epsilon.$$

容易把这个定义推广到多元函数.

定义 1.3.1. 设 n 元函数 $f : D \to \mathbf{R}$ 在 $\mathbf{x}_0$ 的某个去心邻域 $B_r(\mathbf{x}_0)\backslash\{\mathbf{x}_0\}$ 上有定义. 我们称当 $\mathbf{x}$ 趋近于 $\mathbf{x}_0$ 时 f 的极限为 A, 并记作

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = A,$$

如果对任何 $\epsilon > 0$, 存在某个 $\delta > 0$, 使得只要 $0 < d(\mathbf{x}, \mathbf{x}_0) < \delta$, 则有

$$|f(\mathbf{x}) - A| < \epsilon.$$

可以把极限的定义写成便于推广的形式. $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = A$ 当且仅当对 A 的任何邻域 V, 存在 $\mathbf{x}_0$ 的某个去心邻域 U, 使得 $f(U) \subset V$.

注 1.3.2. (1) $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x})$ 只与 f 在 $\mathbf{x}_0$ 的某个去心邻域上的行为有关, 与 $f(\mathbf{x}_0)$ 无关, f 甚至可在 $\mathbf{x}_0$ 处无定义.

(2) 用坐标分量, 可以把极限表示成

$$\lim_{(x_1, ..., x_n) \to (a_1, ..., a_n)} f(x_1, ..., x_n) = A.$$

定理 1.3.3 (极限的四则运算)**.** 设 $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = A$, $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x}) = B$, 则

(1) $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} (f(\mathbf{x}) \pm g(\mathbf{x})) = A \pm B$.

(2) $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} (f(\mathbf{x}) g(\mathbf{x})) = AB$.

(3) 如果 $B \neq 0$, 则 $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} \dfrac{f(\mathbf{x})}{g(\mathbf{x})} = \dfrac{A}{B}$.

命题 1.3.4 (极限不等式)**.** 设 $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = A$, $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x}) = B$, 且对 $\mathbf{x}_0$ 的某个去心邻域 W 中的每个点 $\mathbf{x}$ 都有 $f(\mathbf{x}) \leq g(\mathbf{x})$, 则 $A \leq B$.

定理 1.3.5 (夹逼定理)**.** 设对 $\mathbf{x}_0$ 的某个去心邻域 W 中的每个点 $\mathbf{x}$ 都有

$$f(\mathbf{x}) \leq g(\mathbf{x}) \leq h(\mathbf{x}).$$

如果 $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x})$, $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} h(\mathbf{x})$ 存在且都等于 L, 则 $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x})$ 也存在且等于 L.

例 1.3.6. 设 $P(x,y)=\sum\limits_{i=0}^{n}a_ix^iy^{n-i}$ 是 n 次齐次的二元多项式, 则对任何 $\alpha<n$, 有

$$\lim_{(x,y)\to(0,0)}\frac{P(x,y)}{(x^2+y^2)^{\alpha/2}}=0.$$

证明: 利用不等式 $|x|,|y|\leq(x^2+y^2)^{1/2}$, 可知对每个 $0\leq i\leq n$ 都有

$$\frac{|x^iy^{n-i}|}{(x^2+y^2)^{\alpha/2}}\leq(x^2+y^2)^{(n-\alpha)/2},\quad\forall(x,y)\neq(0,0),$$

由此可得

$$-(x^2+y^2)^{(n-\alpha)/2}\leq\frac{x^iy^{n-i}}{(x^2+y^2)^{\alpha/2}}\leq(x^2+y^2)^{(n-\alpha)/2},\quad\forall(x,y)\neq(0,0).$$

注意到当 $\alpha<n$ 时, 有 $\lim\limits_{(x,y)\to(0,0)}(x^2+y^2)^{(n-\alpha)/2}=0$, 利用夹逼定理1.3.5可得

$$\lim_{(x,y)\to(0,0)}\frac{x^iy^{n-i}}{(x^2+y^2)^{\alpha/2}}=0,$$

再由定理1.3.3可得所要证明的结论.

□

可以把极限的定义推广到一般的映射.

定义 1.3.7. 设 $D\subset\mathbf{R}^n$, 多元映射 $f:D\to\mathbf{R}^m$ 在 $\mathbf{x}_0$ 的某个去心邻域 $B_r(\mathbf{x}_0)\backslash\{\mathbf{x}_0\}$ 上有定义. 我们称当 $\mathbf{x}$ 趋近于 $\mathbf{x}_0$ 时 f 的极限为 $\mathbf{L}\in\mathbf{R}^m$, 并记作

$$\lim_{\mathbf{x}\to\mathbf{x}_0}f(\mathbf{x})=\mathbf{L},$$

如果对任何 $\epsilon>0$, 存在某个 $\delta>0$, 使得只要 $0<d(\mathbf{x},\mathbf{x}_0)<\delta$, 则有

$$d(f(\mathbf{x}),\mathbf{L})<\epsilon.$$

命题 1.3.8 (写成分量的形式)**.** 设 $f:D\to\mathbf{R}^m$ 的各个分量函数为 $f_1,...,f_m$, $\mathbf{L}=(L_1,...,L_m)$, 则

$$\lim_{\mathbf{x}\to\mathbf{x}_0}f(\mathbf{x})=\mathbf{L}\iff\lim_{\mathbf{x}\to\mathbf{x}_0}f_i(\mathbf{x})=L_i,\forall 1\leq i\leq m.$$

证明: 只需注意到对每个 $1\leq i\leq m$ 有

$$|f_i(\mathbf{x})-L_i|\leq d(f(\mathbf{x}),\mathbf{L})\leq\sum_{j=1}^{m}|f_j(\mathbf{x})-L_j|.$$

□

定理 1.3.9 (复合函数的极限). 设 $f:\mathbf{R}^l\to\mathbf{R}^m, g:\mathbf{R}^m\to\mathbf{R}^n$ 满足 $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})=\mathbf{y}_0$, $\lim\limits_{\mathbf{y}\to\mathbf{y}_0} g(\mathbf{y})=\mathbf{z}_0$.

(1) 如果对 $\mathbf{x}_0$ 的某个去心邻域 W 中的每个点 $\mathbf{x}$ 都有 $f(\mathbf{x})\neq\mathbf{y}_0$, 则 $\lim\limits_{\mathbf{x}\to\mathbf{x}_0}(g\circ f)(\mathbf{x})=\mathbf{z}_0$.

(2) 如果 $g(\mathbf{y}_0)=\mathbf{z}_0$, 则 $\lim\limits_{\mathbf{x}\to\mathbf{x}_0}(g\circ f)(\mathbf{x})=\mathbf{z}_0$.

证明: 与一元函数中相应版本定理的证明类似. □

例 1.3.10. 计算极限 $\lim\limits_{(x,y)\to(0,0)}(1+x^2+y^2)^{1/(x^2+y^2)}$.

解. 定义 $f:\mathbf{R}^2\to\mathbf{R}$ 与 $g:\mathbf{R}\backslash\{0\}\to\mathbf{R}$ 为

$$f(x,y)=x^2+y^2,\quad g(t)=(1+t)^{1/t}.$$

利用定理1.3.9中第 (1) 种情形的结论, 可得

$$\lim_{(x,y)\to(0,0)} g\circ f(x,y)=\lim_{t\to 0} g(t)=e.$$

□

例 1.3.11. 极限 $\lim\limits_{(x,y)\to(0,0)}\dfrac{xy}{x^2+y^2}$ 是否存在?

解. 我们来证明此极限不存在. 用反证法, 令 $g:\mathbf{R}^2\backslash\{\mathbf{0}\}\to\mathbf{R}$ 为

$$g(x,y)=\frac{xy}{x^2+y^2},$$

假设 $\lim\limits_{(x,y)\to(0,0)}\dfrac{xy}{x^2+y^2}=L$. 对于非零实数 k, 定义映射 $\Delta_k:\mathbf{R}\backslash\{0\}\to\mathbf{R}^2\backslash\{\mathbf{0}\}$ 为

$$\Delta_k(t)=(t,kt),\quad \forall t\in\mathbf{R}\backslash\{0\}.$$

利用定理1.3.9中第 (1) 种情形的结论, 可得

$$\lim_{t\to 0} g\circ\Delta_k(t)=\lim_{(x,y)\to(0,0)} g(x,y)=L,$$

但直接计算可得

$$\lim_{t\to 0} g\circ\Delta_k(t)=\lim_{t\to 0}\frac{t\cdot kt}{t^2+(kt)^2}=\frac{k}{1+k^2},$$

矛盾! 这就证明了本例中的极限不存在. □

1.4 连续映射

回忆一元函数连续性的定义. 一元函数 $f:\mathbf{R}\to\mathbf{R}$ 在 x_0 处连续当且仅当 $\lim\limits_{x\to x_0} f(x)=f(x_0)$. 或者等价的, 对任何 $\epsilon>0$, 存在 $\delta>0$, 使得对任何 $|x-x_0|<\delta$, 都有 $|f(x)-f(x_0)|<\epsilon$.

容易把这个定义推广到多元映射.

定义 1.4.1. 设 D 是 $\mathbf{R}^n$ 的子集, 映射 $f:D\to\mathbf{R}^m$ 在 $\mathbf{x}_0$ 的某个开球邻域上有定义. 称 f 在 $\mathbf{x}_0$ 处连续, 如果 f 满足如下彼此等价的条件:

(1) $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})=f(\mathbf{x}_0)$.

(2) 对任何 $\epsilon>0$, 存在 $\delta>0$, 使得对任何 $d(\mathbf{x},\mathbf{x}_0)<\delta$, 都有 $d(f(\mathbf{x}),f(\mathbf{x}_0))<\epsilon$.

(3) 对任何 $\epsilon>0$, 存在 $\delta>0$, 使得 $f(B_\delta(\mathbf{x}_0))\subset B_\epsilon(f(\mathbf{x}_0))$.

(4) 对 $f(\mathbf{x}_0)$ 的任何开球邻域 V, 存在 $\mathbf{x}_0$ 的某个开球邻域 U, 使得 $f(U)\subset V$.

注意到, 上述定义中要求 $\mathbf{x}_0$ 是定义域 D 的内点, 可适当修改推广至更一般的情形.

定义 1.4.2. 设 D 是 $\mathbf{R}^n$ 的子集, $\mathbf{x}_0\in D$. 称映射 $f:D\to\mathbf{R}$ 在 $\mathbf{x}_0$ 处连续, 如果对任何 $\epsilon>0$, 存在 $\delta>0$, 使得 $f(B_\delta(\mathbf{x}_0)\cap D)\subset B_\epsilon(f(\mathbf{x}_0))$.

由此可以定义整体的连续映射.

定义 1.4.3 (连续映射)**.** 设 D 是 $\mathbf{R}^n$ 的子集. 称映射 $f:D\to\mathbf{R}^m$ 是连续的, 并记作 $f\in C(D,\mathbf{R}^m)$, 如果 f 在 D 中每一点处都连续.

粗略的说, 所谓 $f:D\to\mathbf{R}^m$ 是连续的, 是指在映射过程中没有撕裂 D.

例 1.4.4. 投影映射 $p_i:\mathbf{R}^n\to\mathbf{R}$ 是连续的.

命题 1.4.5. 映射 $f:D\to\mathbf{R}^m$ 连续当且仅当它的各个分量函数 $f_i(1\le i\le m)$ 都连续.

证明*:* 与命题1.3.8的证明一样, 只需用到如下不等式

$$|f_i(\mathbf{x})-f_i(\mathbf{x}_0)|\le d(f(\mathbf{x}),f(\mathbf{x}_0))\le\sum_{j=1}^m|f_j(\mathbf{x})-f_j(\mathbf{x}_0)|.$$

□

定理 1.4.6. 设 D 是 $\mathbf{R}^n$ 的子集, 则映射 $f:D\to\mathbf{R}^m$ 是连续映射的充分必要条件是*:* 对 $\mathbf{R}^m$ 中的任何开集 V, 存在 $\mathbf{R}^n$ 的开集 U 使得 $f^{-1}(V)=U\cap D$.

证明: "$\Rightarrow$" 假设 f 是连续映射. 设 V 是 $\mathbf{R}^m$ 的开集, 对每个点 $\mathbf{x}\in f^{-1}(V)$, $f(\mathbf{x})$ 属于开集 V, 由开集的定义可知存在开球邻域 $B_\delta(f(\mathbf{x}))\subset V$. 由于 f 在 $\mathbf{x}$ 处连续, 存在 $\mathbf{x}$ 的开球邻域 U_x 满足 $f(U_x\cap D)\subset B_\delta(f(\mathbf{x}))$. 特别的, $U_x\cap D\subset f^{-1}(V)$, 从而有

$$f^{-1}(V)=\bigcup_{x\in f^{-1}(V)}\{\mathbf{x}\}\subset\bigcup_{x\in f^{-1}(V)}(U_x\cap D)\subset f^{-1}(V).$$

令 $U=\bigcup_{\mathbf{x}\in f^{-1}(V)}U_x$, 它是 $\mathbf{R}^n$ 的开集, 则上式表明 $f^{-1}(V)=U\cap D$.

"$\Leftarrow$" 假设对 $\mathbf{R}^m$ 中的任何开集 V, 都存在 $\mathbf{R}^n$ 的开集 U 使得 $f^{-1}(V)=U\cap D$, 我们来证明 f 在每个点 $\mathbf{x}\in D$ 处连续. 为此, 对 $f(\mathbf{x})$ 的任何开球邻域 W, 由前述假设知存在 $\mathbf{R}^n$ 的开集 U 使得 $f^{-1}(W)=U\cap D$. 特别的, $\mathbf{x}\in f^{-1}(W)\subset U$, 由于 U 是开集, 存在 $\mathbf{x}$ 的开球邻域 U_0, 满足 $U_0\subset U$, 这样就有

$$f(U_0\cap D)\subset f(U\cap D)\subset W,$$

这就证明了 f 在 $\mathbf{x}$ 处连续, 从而完成了整个证明. □

我们先考察定理1.4.6的特例, 取 $D=\mathbf{R}^n$, 得到如下的推论.

推论 1.4.7. *映射 $f:\mathbf{R}^n\to\mathbf{R}^m$ 是连续映射的充分必要条件是对 $\mathbf{R}^m$ 的任何开集 V, 其原像集 $f^{-1}(V)$ 是 $\mathbf{R}^n$ 的开集.*

以后我们会经常用到如下结论.

推论 1.4.8. *设 $f:\mathbf{R}^n\to\mathbf{R}^m$ 是连续映射, 则对任何 $\mathbf{c}\in\mathbf{R}^m$, 集合 $f^{-1}(\{\mathbf{c}\})$ 是 $\mathbf{R}^n$ 的闭集.*

证明: 注意到 $\mathbf{R}^m\backslash\{\mathbf{c}\}$ 是 $\mathbf{R}^m$ 的开集, 利用推论1.4.8即可. □

为了方便记忆, 人们经常说: f 是连续映射的充分必要条件是 f 下开集的原像是开集. 为了让这句话适用于更一般的定义域 D, 可引入如下定义.

定义 1.4.9. 设 D 是 $\mathbf{R}^n$ 的子集, 可按如下方式赋予 D 拓扑结构, 称之为 D 上的子空间拓扑: 定义 $W\subset D$ 是子空间拓扑下的开集, 当且仅当存在 $\mathbf{R}^n$ 的的开集 U 使得 $W=U\cap D$.

在此定义下, 可把定理1.4.6统一叙述成: $f:D\to\mathbf{R}^m$ 是连续映射的充分必要条件是 f 下开集的原像是开集. 人们经常把 "开集的原像是开集" 作为连续性的定义.

1.4.1 连续函数的局部性质

定理 1.4.10 (四则运算保持连续性)**.** *设 $f,g:D\to\mathbf{R}$ 在点 $\mathbf{x}_0\in D$ 处都连续, 则*

(1) *$f\pm g$ 在点 $\mathbf{x}_0\in D$ 处连续.*

(2) *$f\cdot g$ 在点 $\mathbf{x}_0\in D$ 处连续.*

(3) *如果 $g(\mathbf{x}_0)\neq 0$, 则 $\frac{f}{g}$ 在点 $\mathbf{x}_0\in D$ 处连续.*

推论 1.4.11. 设 $f, g \in C(D, \mathbf{R})$, 则 $f \pm g, f \cdot g \in C(D, \mathbf{R})$. 如果 g 在 D 上处处非零, 则 $\frac{f}{g} \in C(D, \mathbf{R})$.

定理 1.4.12 (连续函数的复合是连续的)**.** 设 $f : D \to E$ 在 $\mathbf{x}_0$ 处连续, $g : E \to F$ 在 $f(\mathbf{x}_0)$ 处连续, 则 $g \circ f : D \to F$ 在 $\mathbf{x}_0$ 处连续.

证明: 对于 $g(f(\mathbf{x}_0))$ 的任何邻域 W, 由 g 在 $f(\mathbf{x}_0)$ 处连续的定义, 存在 $f(\mathbf{x}_0)$ 的邻域 V, 使得 $g(V \cap F) \subset W$. 再由 f 在 $\mathbf{x}_0$ 处连续的定义, 存在 $\mathbf{x}_0$ 的邻域 U, 使得 $f(U \cap D) \subset V$. 结合起来可得

$$(g \circ f)(U \cap D) = g(f(U \cap D)) \subset g(V \cap E) \subset W,$$

这就证明了 $g \circ f$ 在 $\mathbf{x}_0$ 处连续. □

推论 1.4.13. 设 D, E, F 都是欧氏空间的子集. 设 $f \in C(D, E), g \in C(E, F)$, 则 $g \circ f \in C(D, F)$.

例 1.4.14. 定义 $f : \mathbf{R}^2 \to \mathbf{R}$ 为

$$f(x, y) = \begin{cases} \frac{x^2+y^2}{|x|+|y|} & \text{如果}(x, y) \neq (0, 0) \\ 0 & \text{如果}(x, y) = (0, 0). \end{cases}$$

f 是否连续?

证明: (1) 投影函数 x, y 是连续函数, 它们复合绝对值函数后所得的函数 $|x|, |y|$ 也连续. 利用推论1.4.11可知 f 在原点之外都连续.

(2) 注意到

$$0 \leq \frac{x^2 + y^2}{|x| + |y|} \leq |x| + |y|, \quad \forall (x, y) \neq (0, 0),$$

利用夹逼定理1.3.5可得

$$\lim_{(x,y)\to(0,0)} \frac{x^2 + y^2}{|x| + |y|} \leq |x| + |y| = 0 = f(0, 0),$$

表明 f 在原点处连续. □

1.5 介值定理与最值定理

1.5.1 介值定理

定义 1.5.1. 称 $\mathbf{R}^n$ 的子集 D 是道路连通的 (*path connected*), 如果 D 中任何两点都可以用 D 中的道路连接. 具体的说, 即对任何 $\mathbf{x}, \mathbf{y} \in D$, 存在连续映射 $p : [0, 1] \to D$ 使得 $p(0) = \mathbf{x}, p(1) = \mathbf{y}$.

例 1.5.2. 当 $n \geq 2$ 时, $(n-1)$ 维球面 S^{n-1} 是道路连通的.

证明: 设 $\mathbf{x}, \mathbf{y}$ 是 S^{n-1} 上任何两点.

(1) 当 $\mathbf{x} \neq -\mathbf{y}$ 时, 道路

$$p(t) = \frac{\mathbf{x} + t(\mathbf{y} - \mathbf{x})}{||\mathbf{x} + t(\mathbf{y} - \mathbf{x})||}, \quad \forall t \in [0, 1],$$

是 S^{n-1} 上连接 $\mathbf{x}$ 到 $\mathbf{y}$ 的道路.

(2) 当 $\mathbf{x} = -\mathbf{y}$ 时, 取点 $\mathbf{z} \in S^{n-1}$ 满足 $\mathbf{z} \neq \mathbf{x}, \mathbf{y}$, 则道路

$$q(t) = \begin{cases} \frac{\mathbf{x}+2t(\mathbf{z}-\mathbf{x})}{||\mathbf{x}+2t(\mathbf{z}-\mathbf{x})||}, & \forall t \in [0, \frac{1}{2}], \\ \frac{\mathbf{z}+(2t-1)(\mathbf{y}-\mathbf{z})}{||\mathbf{z}+(2t-1)(\mathbf{y}-\mathbf{z})||}, & \forall t \in [\frac{1}{2}, 1], \end{cases}$$

是 S^{n-1} 上连接 $\mathbf{x}$ 到 $\mathbf{y}$ 的道路. □

定理 1.5.3 (介值定理)**.** *设 D 是 $\mathbf{R}^n$ 的道路连通的子集, $f : D \to \mathbf{R}$ 是连续函数, 则对任何 $\mathbf{x}_0, \mathbf{x}_1 \in D$, 对任何介于 $f(\mathbf{x}_0), f(\mathbf{x}_1)$ 之间的数 y, 存在 $\mathbf{x} \in D$ 使得 $f(\mathbf{x}) = y$.*

证明: 取一条道路 $p : [0, 1] \to D$ 连接 $\mathbf{x}_0, \mathbf{x}_1$

$$p(0) = \mathbf{x}_0, \quad p(1) = \mathbf{x}_1,$$

则复合映射 $f \circ p : [0, 1] \to \mathbf{R}$ 是连续的一元函数. 由一元函数的介值定理知存在 $t \in [0, 1]$, 使得 $(f \circ p)(t) = y$, 此即 $f(p(t)) = y$. □

例 1.5.4. 当 $n = 1$ 时, 闭区间 $[a, b] \subset \mathbf{R}$ 是道路连通的, 则定理1.5.3断言连续函数 $f : [a, b] \to \mathbf{R}$ 能取到一切中间值, 这就是一元函数的介值定理.

1.5.2 最值定理

定义 1.5.5. 设 X 是 $\mathbf{R}^n$ 的子集, $\mathcal{U} = \{U_\alpha : \alpha \in I\}$($I$ 是指标集) 是由 $\mathbf{R}^n$ 的若干子集构成的子集族. 称 $\mathcal{U}$ 是 X 的一个覆盖, 如果 $X \subset \cup_{\alpha \in I} U_\alpha$. 如果 $\mathcal{U}$ 是 X 的一个覆盖且 $\mathcal{U}$ 中每个成员都是 $\mathbf{R}^n$ 的开集, 则称 $\mathcal{U}$ 是 X 的一个开覆盖.

如果 $\mathcal{U}$ 的子集 $\mathcal{V}$ 满足 $X \subset \cup_{U \in \mathcal{V}} U$, 则称 $\mathcal{V}$ 是一个子覆盖, 进一步如果 $\mathcal{V}$ 还是有限集, 则称 $\mathcal{V}$ 是一个有限子覆盖.

定义 1.5.6. 称 $\mathbf{R}^n$ 的子集 X 是紧致的 (*compact*), 如果 X 的任何开覆盖都有有限子覆盖. 具体的说, 如果对于 X 的任何开覆盖 $\mathcal{U}$, 都可以从 $\mathcal{U}$ 中选出有限多个成员覆盖 X.

定理 1.5.7 (Heine-Borel)**.** *$\mathbf{R}^n$ 的子集 X 是紧致的当且仅当 X 是 $\mathbf{R}^n$ 的有界闭子集.*

证明: "$\Rightarrow$" 设 X 是紧致的. 考虑 X 的开覆盖 $\mathcal{U}=\{B_n(\mathbf{0}):n\in\mathbf{Z}_+\}$, 它有有限子覆盖, 可知存在 N 使得 $X\subset B_n(\mathbf{0})$, 即 X 是有界的. 另一方面, 对任何 $\mathbf{y}\in\mathbf{R}^n\backslash X$, 考虑 X 的开覆盖 $\mathcal{V}=\{\mathbf{R}^n\backslash\overline{B}_{1/n}(y):n\in\mathbf{Z}_+\}$, 它有有限子覆盖, 可知存在 N 使得 $X\subset\mathbf{R}^n\backslash\overline{B}_{1/N}(y)$, 从而有

$$B_{1/N}(y)\subset\overline{B}_{1/N}(y)\subset\mathbf{R}^n\backslash X,$$

这表明 $R^n\backslash X$ 的每个点 $\mathbf{y}$ 都是它的内点, 证明了 $R^n\backslash X$ 是开集, 即 X 是闭集.

"$\Leftarrow$" 设 X 是有界闭子集, $\mathcal{U}$ 是 X 的任何开覆盖, 我们来证明它有有限子覆盖. 由于 X 有界, 取一个闭矩形体 $C=[a_1,b_1]\times\cdots\times[a_n,b_n]$ 包含 X, 考虑子集族

$$\mathcal{U}'=\mathcal{U}\cup\{\mathbf{R}^n\backslash X\},$$

由于 X 是闭集, 则 $\mathcal{U}'$ 是 C 的开覆盖. 我们只需证明 $\mathcal{U}'$ 有 C 的有限子覆盖, 因为在此有限子覆盖中忽略成员 $\mathbf{R}^n\backslash X$, 即得 $\mathcal{U}$ 有 X 的有限子覆盖. 用反证法, 假设 $\mathcal{U}'$ 不含 C 的有限子覆盖, 令 $C_1=C$, 把 C_1 的每条棱都从中点剖分, 将 C_1 剖分成 2^n 个小矩形体, 则其中必有一个小矩形体 (记为 C_2) 在 $\mathcal{U}'$ 中不存在有限子覆盖, 依此进行下去, 可得到一族矩形体

$$C_1\supset C_2\supset\cdots\supset C_m\supset\cdots$$

其中每个 C_m 在 $\mathcal{U}'$ 中不存在有限子覆盖, 且它们的边长趋于零. 利用区间套定理可知 $\bigcap_{m=1}^{\infty}C_m$ 是单点集, 设为 $\{\mathbf{x}\}$, 则 $\mathbf{x}$ 属于 $\mathcal{U}'$ 中某个成员 W. 注意到 W 是开集, 存在 $B_\delta(\mathbf{x})\subset W$, 这样当 m 充分大时有 $C_m\subset W$, 与前述 C_m 在 $\mathcal{U}'$ 中不存在有限子覆盖矛盾! 这就完成了整个证明. □

定理 1.5.8. *设 X 是 $\mathbf{R}^n$ 的紧致子集, $f:X\to\mathbf{R}^m$ 是连续映射, 则 $f(X)$ 是 $\mathbf{R}^m$ 的紧致子集.*

证明: 设 $\mathcal{V}=\{V_\alpha:\alpha\in I\}$ 是 $f(X)$ 在 $\mathbf{R}^m$ 中的任何开覆盖. 对每个 V_α, 由定理1.4.6 知存在 $\mathbf{R}^n$ 的开集 U_α 使得 $f^{-1}(V_\alpha)=U_\alpha\cap X$. 考虑子集族 $\mathcal{U}=\{U_\alpha:\alpha\in I\}$, 它是 X 的开覆盖. 由于 X 紧致, $\mathcal{U}$ 中存在有限子覆盖

$$X\subset U_{\alpha_1}\cup\cdots\cup U_{\alpha_k},$$

相应的得到 $f(X)$ 在 $\mathcal{V}$ 中存在有限子覆盖

$$f(X)\subset V_{\alpha_1}\cup\cdots\cap V_{\alpha_k},$$

这就证明了 $f(X)$ 紧致. □

推论 1.5.9 (最值定理: 有界闭集上的连续函数能取到最大与最小值)**.** *设 D 是 $\mathbf{R}^n$ 的有界的闭集, $f:D\to\mathbf{R}$ 是连续函数, 则 f 在 D 上能取到最大值与最小值. 具体的说, 存在 $\mathbf{x}_{min},\mathbf{x}_{max}\in D$ 使得:*

$$f(\mathbf{x}_{min})\leq f(x)\leq f(\mathbf{x}_{max}),\quad\forall\mathbf{x}\in D.$$

证明: 由定理1.5.7知 D 紧致, 利用定理1.5.8可知 $f(D)$ 也紧致, 再一次利用定理1.5.7的结论可得 $f(D)$ 是有界闭集. 设 $M=\sup f(D)$, 来证明 $M\in f(D)$, 从而说明了 $f(D)$ 有最大值. 用反证法, 假设 $M\notin f(D)$, 则 M 是开集 $\mathbf{R}\backslash f(D)$ 的内点, 因而存在 $B_\delta(M)\subset\mathbf{R}\backslash f(D)$, 这与 M 是 $f(D)$ 的上确界矛盾! □

推论 1.5.10 (有界性定理: 有界闭集上的连续函数一定有界)**.** *设 D 是 $\mathbf{R}^n$ 的有界闭子集, $f:D\to\mathbf{R}$ 是连续函数, 则存在常数 M 使得:*

$$|f(\mathbf{x})|\le M,\quad \forall\mathbf{x}\in D.$$

例 1.5.11. 设 $A=(a_{ij})_{1\le i,j\le n}$ 是对称的实矩阵, 它给出二次函数

$$Q(x_1,...,x_n)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j,$$

由最值定理可知 Q 在 S^{n-1} 上能取到最大值与最小值. 以后, 利用 Lagrange 乘子法, 我们可由上述结果证明 A 一定有实的特征根.

第二章　多元函数微分学

2.1　导数

一元函数 $f:\mathbf{R}\to\mathbf{R}$ 的导数

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$$

反映了 f 在 x_0 处的变化率.

问题 2.1.1. *设 $D\subset\mathbf{R}^n$. 如何描述 $f:D\to\mathbf{R}$ 在点 $\mathbf{x}$ 处的变化率?*

为了研究多元函数 $f:D\to\mathbf{R}$ 在 $\mathbf{x}$ 附近的行为, 可以选取一条过 $\mathbf{x}$ 的道路

$$p:\mathbf{R}\to D,\quad p(0)=\mathbf{x},$$

通过研究一元函数 $f\circ p$, 获得 f 在 $\mathbf{x}$ 附近的部分信息.

定义 2.1.2. 设 $\mathbf{x}\in\mathbf{R}^n$. 过 $\mathbf{x}$ 的一条直线是指一个映射 $p:\mathbf{R}\to\mathbf{R}^n$,

$$p(t)=\mathbf{x}+t\mathbf{v},\quad\forall t\in\mathbf{R},$$

其中 $\mathbf{v}\in\mathbf{R}^n$, 称之为直线 p 的 "方向".

可用坐标分量更加具体的表述. 设

$$\mathbf{x}=(x_1,...,x_n),\quad\mathbf{v}=(v_1,...,v_n),$$

则有

$$p(t)=(x_1+tv_1,...,x_n+tv_n),\quad\forall t\in\mathbf{R}.$$

例 2.1.3. 对 $1\le k\le n$, 令 $\mathbf{e}_k=(0,...0,1,0...,0)$, 它的第 k 个分量为 1, 其他分量都等于 0, 称之为 $\mathbf{R}^n$ 的第 k 个坐标的单位正方向, 简称为第 k 个坐标方向.

定义 2.1.4 (方向导数)**.** 设 D 是 $\mathbf{R}^n$ 的子集, 函数 $f: D \to \mathbf{R}$ 在点 $\mathbf{x}$ 的某个开球邻域中有定义. 定义 f 在 $\mathbf{x}$ 处沿 $\mathbf{v}$ 方向的方向导数为

$$\nabla_{\mathbf{v}} f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}} = \frac{d}{dt}|_{t=0} f(\mathbf{x} + t\mathbf{v}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t},$$

它反映了 f 在 $\mathbf{x}$ 处沿 $\mathbf{v}$ 方向的变化率. 设 $\mathbf{x} = (x_1, ..., x_n), \mathbf{v} = (v_1, ..., v_n)$, 则

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}} = \lim_{t \to 0} \frac{f(x_1 + tv_1, ..., x_n + tv_n) - f(x_1, ..., x_n)}{t},$$

我们也把上述方向导数记作

$$\frac{\partial f}{\partial \mathbf{v}}|_{(x_1,...,x_n)} \text{ 或} \frac{\partial f(x_1, ..., x_n)}{\partial \mathbf{v}}.$$

定义 2.1.5 (偏导数)**.** 把 f 在 $\mathbf{x} = (x_1, ..., x_n)$ 处沿第 i 个坐标方向 $\mathbf{e}_i$ 的方向导数称为 f 在 $\mathbf{x} = (x_1, ..., x_n)$ 处的第 i 个偏导数, 记作

$$\frac{\partial f}{\partial x_i}|_{(x_1,...,x_n)}, \quad \text{或} \frac{\partial f(x_1, ..., x_n)}{\partial x_i},$$

它反映了 f 在 $(x_1, ..., x_n)$ 处沿第 i 个坐标方向的变化率. 具体的说,

$$\frac{\partial f(x_1, ..., x_n)}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, ..., x_i + \Delta x_i, ..., x_n) - f(x_1, ..., x_i, ..., x_n)}{\Delta x_i}.$$

计算偏导数 $\frac{\partial f}{\partial x_i}$ 时, 把其他 $x_j (j \neq i)$ 都视为常数, 把 $f(x_1, ..., x_n)$ 看成 x_i 的一元函数, 计算它关于 x_i 的导数即可.

注 2.1.6. 一般也把二元函数写作 $f(x, y)$, 这样, f 的各个偏导数也分别记作

$$\frac{\partial f}{\partial x} = f_x = f_1', \quad \frac{\partial f}{\partial y} = f_y = f_2'.$$

类似的, 如果把三元函数写作 $f(x, y, z)$, 则也把 f 的第 3 个偏导数记作

$$\frac{\partial f}{\partial z} = f_z = f_3'$$

例 2.1.7. 求 $f(x, y) = \sqrt{x^2 + y^2}$ 的方向导数.

解. 对于方向 $\mathbf{v} = (a, b)$,

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{d}{dt}|_{t=0} \sqrt{(x + at)^2 + (y + bt)^2} = \frac{ax + by}{\sqrt{x^2 + y^2}}, \quad \forall (x, y) \neq (0, 0).$$

特别的, f 的两个偏导数分别为

$$\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}.$$

注意到, f 在 $(0,0)$ 处连续, 但 f 在该点处对任何非零方向的方向导数都不存在, 因为对 $\mathbf{v}=(a,b)\neq(0,0)$, 按照定义式有

$$\frac{\partial f}{\partial \mathbf{v}}|_{(0,0)}=\frac{d}{dt}|_{t=0}f(ta,tb)=\frac{d}{dt}|_{t=0}\sqrt{(ta)^2+(tb)^2}=\lim_{t\to 0}\frac{|t|\sqrt{a^2+b^2}}{t},$$

此极限不存在. 这个例子表明, 连续函数不一定有偏导数或方向导数. □

例 2.1.8. 定义函数 $f:\mathbf{R}^2\to\mathbf{R}$ 为

$$f(x,y)=\begin{cases}\frac{|y|\sqrt{x^2+y^2}}{x} & \text{如果}x\neq 0,\\ 0 & \text{如果}x=0,\end{cases}$$

则 f 在 $(0,0)$ 处有各个方向导数, 但在 $(0,0)$ 处不连续. 由此可知, 对于多元函数, 可导并不一定连续.

证明: 对任何单位方向 $\mathbf{v}=(\cos\theta,\sin\theta)$, 由定义直接计算, 当 $\cos\theta\neq 0$ 时,

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}|_{(0,0)}&=\lim_{t\to 0}\frac{f(t\cos\theta,t\sin\theta)-f(0,0)}{t}\\&=\lim_{t\to 0}\frac{\frac{|t\sin\theta|\cdot|t|}{t\cos\theta}}{t}\\&=\frac{|\sin\theta|}{\cos\theta}.\end{aligned}$$

另外, 当 $\cos\theta=0$ 时显然有 $\frac{\partial f}{\partial \mathbf{v}}|_{(0,0)}=0$. 这样, f 在 $(0,0)$ 处有各个方向的方向导数.

我们来证明 f 在 $(0,0)$ 处不连续. 否则的话, 对任何 $\epsilon>0$, 存在 $\delta>0$ 使得对任何 $0<\sqrt{x^2+y^2}<\delta$, 有 $|f(x,y)|<\epsilon$. 取 $(x,y)=\frac{\delta}{2}(\cos\theta,\sin\theta)$, 则 $f(x,y)=\frac{\delta}{2}\cdot\frac{|\sin\theta|}{\cos\theta}$, 由于当 $\theta\in(0,\frac{\pi}{2})$ 时, $\frac{|\sin\theta|}{\cos\theta}=\tan\theta$ 是无界的, 可以取 θ 使 $f(x,y)=\frac{\delta}{2}\cdot\frac{|\sin\theta|}{\cos\theta}>\epsilon$, 矛盾! □

例 2.1.9. 令 $f(x,y,z)=x^2+y^2-z^2$, 则

$$f_x=2x,\quad f_y=2y,\quad f_z=-2z.$$

对任何方向 $\mathbf{v}=(a,b,c)$, 有

$$\frac{\partial f}{\partial \mathbf{v}}=2ax+2by-2cz.$$

2.2 微分

2.2.1 函数的微分

粗略的说, 函数的微分就是它的线性近似.

定义 2.2.1 (微分)**.** 设函数 f 在 $\mathbf{x}_0$ 的某个开球邻域中有定义. 如果存在线性函数

$$\begin{aligned} L: \quad & \mathbf{R}^n \to \mathbf{R}, \\ & \mathbf{h} \to L(\mathbf{h}) \end{aligned}$$

使得在 $\mathbf{x}_0$ 附近, 有

$$f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + L(\mathbf{h}) + \alpha(\mathbf{h}),$$

其中 α 满足 $\lim\limits_{\mathbf{h}\to 0}\dfrac{\alpha(\mathbf{h})}{||\mathbf{h}||}=0$(即当 $\mathbf{h}\to 0$ 时, α 是比 $||\mathbf{h}||$ 更高阶的无穷小), 则称 f 在 $\mathbf{x}_0$ 处可微, 并把上述线性函数 L 称为 f 在 $\mathbf{x}_0$ 处的微分, 记作 $df_{\mathbf{x}_0}: \mathbf{R}^n \to \mathbf{R}$.

引理 2.2.2. *满足定义2.2.1中条件的线性映射是唯一的.*

证明: 用反证法, 假设两个不同的线性映射 L, L' 满足条件, 则有

$$\lim_{\mathbf{h}\to 0}\frac{(L-L')(\mathbf{h})}{||\mathbf{h}||}=0.$$

设 $(L-L')(\mathbf{h})=\sum\limits_{i=1}^{n} a_i h_i$, 且 $a_1 \neq 0$. 由复合函数极限定理可得

$$\lim_{t\to 0}\frac{(L-L')(a_1t,0,...,0)}{||(a_1t,0,...,0)||}=0,$$

但上式左边等于

$$\lim_{t\to 0}\frac{a_1^2t}{|a_1t|}=|a_1|\cdot\lim_{t\to 0}\frac{t}{|t|},$$

此极限不存在, 矛盾! □

例 2.2.3. 线性函数的微分是它自己. 设 $f(\mathbf{x})=\sum\limits_{i=1}^{n} a_i x_i$ 是线性函数, 则取 $L=f$, $\alpha=0$ 可得

$$f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + L(\mathbf{h}) + \alpha(\mathbf{h}),$$

这表明 f 在每点 $\mathbf{x}_0$ 的微分都等于 f. 特别的, 第 i 个投影映射 $p_i:\mathbf{R}^n\to\mathbf{R}$ 的微分为

$$dp_i(\mathbf{h})=h_i, \quad \forall \mathbf{h}=(h_1,...,h_n)\in\mathbf{R}^n.$$

例 2.2.4. 投影函数 $p_1:\mathbf{R}^2\to\mathbf{R}, \quad p_1(x,y)=x$ 的微分为

$$(dp_1)_{(x_0,y_0)}(s,t)=s.$$

类似的, 投影函数 $p_2:\mathbf{R}^2\to\mathbf{R}, \quad p_2(x,y)=y$ 的微分为

$$(dp_2)_{(x_0,y_0)}(s,t)=t.$$

把 dp_1, dp_2 分别记作 dx, dy, 则 $dx(s,t)=s, dy(s,t)=t$.

定理 2.2.5 (可微 $\Rightarrow$ 连续)**.** 设 $f:\mathbf{R}^n\to\mathbf{R}$ 在 $\mathbf{x}_0$ 处可微, 则 f 在 $\mathbf{x}_0$ 处连续.

定理 2.2.6 (可微 $\Rightarrow$ 可导)**.** 设 $f:\mathbf{R}^n\to\mathbf{R}$ 在 $\mathbf{x}_0$ 处可微, 则 f 在 $\mathbf{x}_0$ 处有各个方向导数 *(*特别的, 有各个偏导数*)*, 且对任何方向 $\mathbf{v}=(\mathbf{v}_1,...,\mathbf{v}_n)$, 有

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}_0}=df_{\mathbf{x}_0}(\mathbf{v})=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}|_{\mathbf{x}_0}\cdot\mathbf{v}_i.$$

定理2.2.5,2.2.6的证明. 设 f 在 $\mathbf{x}_0$ 处可微, 则存在线性函数

$$\begin{aligned}L:&\mathbf{R}^n\to\mathbf{R},\\ &\mathbf{h}\to L(\mathbf{h})\end{aligned}$$

使得在 $\mathbf{x}_0$ 附近, 有

$$f(\mathbf{x}_0+\mathbf{h})=f(\mathbf{x}_0)+L(\mathbf{h})+\alpha(\mathbf{h}),$$

其中 α 满足 $\lim\limits_{\mathbf{h}\to 0}\dfrac{\alpha(\mathbf{h})}{|\mathbf{h}|}=0$.

(1) 计算极限

$$\lim_{\mathbf{x}\to\mathbf{x}_0}f(\mathbf{x})=\lim_{\mathbf{h}\to 0}f(\mathbf{x}_0+\mathbf{h})=\lim_{\mathbf{h}\to 0}(f(\mathbf{x}_0)+L(\mathbf{h})+\alpha(\mathbf{h}))=f(\mathbf{x}_0),$$

所以 f 在 $\mathbf{x}_0$ 处连续.

(2) 计算方向导数

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}_0}=\lim_{t\to 0}\frac{f(\mathbf{x}_0+t\mathbf{v})-f(\mathbf{x}_0)}{t}=\lim_{t\to 0}\frac{L(t\mathbf{v})+\alpha(t\mathbf{v})}{t}=L(\mathbf{v}),$$

此即

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}_0}=df_{\mathbf{x}_0}(\mathbf{v}).$$

$\square$

这个定理表明, 如果 f 在某个点处可微, 则 f 在该点的微分是唯一的.

注 2.2.7. 利用例2.2.4中的记号, 可把定理2.2.6的结论叙述成: 如果 f 在 (x_0,y_0) 处可微, 则有

$$df_{(x_0,y_0)}(s,t)=\frac{\partial f(x_0,y_0)}{\partial x}dx_{(x_0,y_0)}(s,t)+\frac{\partial f(x_0,y_0)}{\partial y}dy_{(x_0,y_0)}(s,t),\quad \forall(s,t)\in\mathbf{R}^2.$$

所以, 作为线性映射, 有

$$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy.$$

定理 2.2.8 (可微的充分条件)**.** 设 $D\subset\mathbf{R}^n$, $f:D\to\mathbf{R}$ 在 $\mathbf{x}_0$ 的某个邻域中处处有各个偏导数, 且偏导函数 $\frac{\partial f}{\partial x_1},...,\frac{\partial f}{\partial x_n}$ 都在 $\mathbf{x}_0$ 处连续, 则 f 在 $\mathbf{x}_0$ 处可微.

证明*:* 设 $\mathbf{x}_0=(x_1,...,x_n)$, 利用一元函数的 Lagrange 中值定理, 存在 $\theta_1,...,\theta_n\in(0,1)$, 使得

$$
\begin{aligned}
&f(x_1+h_1,x_2+h_2,...,x_n+h_n)-f(x_1,...,x_n)\\
=&\sum_{i=1}^n\left(f(x_1+h_1,...,x_i+h_i,x_{i+1},...,x_n)-f(x_1+h_1,...,x_{i-1}+h_{i-1},x_i,...,x_n)\right)\\
=&\sum_{i=1}^n h_i\cdot f_i(x_1+h_1,...,x_{i-1}+h_{i-1},x_i+\theta_ih_i,x_{i+1},...,x_n)\\
=&\left(\sum_{i=1}^n h_i\cdot f_i(x_1,...,x_n)\right)+\alpha(h_1,...,h_n),
\end{aligned}
$$

其中函数 α 为

$$\alpha(h_1,...,h_n)=\sum_{i=1}^n h_i\cdot\left(f_i(x_1+h_1,...,x_{i-1}+h_{i-1},x_i+\theta_ih_i,x_{i+1},...,x_n)-f_i(x_1,...,x_n)\right).$$

注意到, 对 $\mathbf{h}\neq\mathbf{0}$, 对每个 $1\le i\le n$, 有 $|\frac{h_i}{|\mathbf{h}|}|\le 1$, 由此可得

$$0\le|\frac{\alpha(\mathbf{h})}{|\mathbf{h}|}|\le\sum_{i=1}^n|f_i(x_1+h_1,...,x_{i-1}+h_{i-1},x_i+\theta_ih_i,x_{i+1},...,x_n)-f_i(x_1,...,x_n)|,\tag{2.1}$$

由于 $f_i=\frac{\partial f}{\partial x_i}$ 在 $(x_1,...,x_n)$ 处连续, 且 $\theta_i\in(0,1)$, 则

$$
\begin{aligned}
&\lim_{\mathbf{h}\to\mathbf{0}}\sum_{i=1}^n|f_i(x_1+h_1,...,x_{i-1}+h_{i-1},x_i+\theta_ih_i,x_{i+1},...,x_n)-f_i(x_1,...,x_n)|\\
=&\sum_{i=1}^n\lim_{\mathbf{h}\to\mathbf{0}}|f_i(x_1+h_1,...,x_{i-1}+h_{i-1},x_i+\theta_ih_i,x_{i+1},...,x_n)-f_i(x_1,...,x_n)|\\
=&0,
\end{aligned}
$$

这样, 由 (1) 式及夹逼定理可得

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{\alpha(\mathbf{h})}{|\mathbf{h}|}=0,$$

这就证明了 f 在 $\mathbf{x}_0$ 处可微.

□

定义 2.2.9. 设 D 是 $\mathbf{R}^n$ 的开集, 称 $f:D\to\mathbf{R}$ 是 C^1 光滑函数, 记作 $f\in C^1(D,\mathbf{R})$, 如果 f 在 D 内各个偏导数 $\frac{\partial f}{\partial x_1},...,\frac{\partial f}{\partial x_n}$ 都存在且连续.

称映射 $g:D\to\mathbf{R}^m$ 是 C^1 光滑映射, 记作 $g\in C^1(D,\mathbf{R}^m)$, 如果 g 的每个分量 $g_1,...,g_m$ 都是 C^1 光滑函数.

推论 2.2.10. 设 D 是 $\mathbf{R}^n$ 的开集, $f \in C^1(D, \mathbf{R})$, 则 f 在 D 内处处可微, 且有

$$df_{\mathbf{x}}(\mathbf{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}|_{\mathbf{x}} \cdot h_i, \quad \forall \mathbf{h} = (h_1, ..., h_n) \in \mathbf{R}^n$$

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}|_{\mathbf{x}} \cdot v_i, \quad \forall \mathbf{v} = (v_1, ..., v_n) \in \mathbf{R}^n.$$

2.2.2 映射的微分

容易把多元函数的微分推广至映射的情形.

定义 2.2.11. 设 $f : \mathbf{R}^n \to \mathbf{R}^m$. 如果存在线性映射

$$\begin{aligned} L: \quad & \mathbf{R}^n \to \mathbf{R}^m, \\ & \mathbf{h} \to L(\mathbf{h}) \end{aligned}$$

使得在 $\mathbf{x}_0$ 附近, 有

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + L(\mathbf{h}) + \alpha(\mathbf{h}),$$

其中 α 满足 $\lim\limits_{\mathbf{h}\to 0} \dfrac{\alpha(\mathbf{h})}{||\mathbf{h}||} = 0$(即当 $\mathbf{h} \to 0$ 时, α 是比 $||\mathbf{h}||$ 更高阶的无穷小), 则称 f 在 $\mathbf{x}_0$ 处可微, 并把上述线性映射 L 称为 f 在 $\mathbf{x}_0$ 处的微分, 记作 $df_{\mathbf{x}_0} : \mathbf{R}^n \to \mathbf{R}^m$.

怎么用偏导数表示微分 $df_{\mathbf{x}_0}$?

命题 2.2.12. 设 $f = (f_1, ..., f_m) : \mathbf{R}^n \to \mathbf{R}^m$, 则

$$f \text{ 在} \mathbf{x}_0 \text{ 处可微} \iff f_1, ..., f_m \text{ 在} \mathbf{x}_0 \text{ 处可微}.$$

在上述条件下, f 的微分 $df_{\mathbf{x}_0} : \mathbf{R}^n \to \mathbf{R}^m$ 由如下公式给出

$$df_{\mathbf{x}_0}(\begin{pmatrix} h_1 \\ ... \\ h_n \end{pmatrix}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_1}{\partial x_n}|_{\mathbf{x}_0} \\ ... & ... & ... \\ \frac{\partial f_m}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_m}{\partial x_n}|_{\mathbf{x}_0} \end{pmatrix} \begin{pmatrix} h_1 \\ ... \\ h_n \end{pmatrix}.$$

证明: f 在 $\mathbf{x}_0$ 处可微的定义是, 如果存在线性映射 $L : \mathbf{R}^n \to \mathbf{R}^m$ 使得在 $\mathbf{x}_0$ 附近有 $f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + L(\mathbf{h}) + \alpha(\mathbf{h})$, 且 $\lim\limits_{\mathbf{h}\to 0} \dfrac{\alpha(\mathbf{h})}{||\mathbf{h}||} = 0$. 设 $L = (L_1, ..., L_m)$, $\alpha = (\alpha_1, ..., \alpha_m)$, 则由上述定义等价于: 对每个 $1 \le i \le m$ 有 $f_i(\mathbf{x}_0 + \mathbf{h}) = f_i(\mathbf{x}_0) + L_i(\mathbf{h}) + \alpha_i(\mathbf{h})$, 且 $\lim\limits_{\mathbf{h}\to 0} \dfrac{\alpha_i(\mathbf{h})}{||\mathbf{h}||} = 0$, 即对每个 i, f_i 在 $\mathbf{x}_0$ 处都可微.

进一步, 从上述证明过程中可知, $df_{\mathbf{x}_0}$ 的第 i 个分量函数等于 $(df_i)_{\mathbf{x}_0}$, 在利用定理2.2.6的结论可得到 $df_{\mathbf{x}_0}$ 的矩阵表示. □

推论 2.2.13 (C^1 光滑 $\Rightarrow$ 可微)**.** 设 D 是 $\mathbf{R}^n$ 的开集, $f : D \to \mathbf{R}^m$ 是 C^1 光滑映射, 则 f 在 D 上处处可微, 且它的微分为

$$df_{\mathbf{x}_0}(\begin{pmatrix} h_1 \\ ... \\ h_n \end{pmatrix}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_1}{\partial x_n}|_{\mathbf{x}_0} \\ ... & ... & ... \\ \frac{\partial f_m}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_m}{\partial x_n}|_{\mathbf{x}_0} \end{pmatrix} \begin{pmatrix} h_1 \\ ... \\ h_n \end{pmatrix}.$$

定义 $m \times n$ 的矩阵

$$J_f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & ... & \frac{\partial f_1}{\partial x_n} \\ ... & ... & ... \\ \frac{\partial f_m}{\partial x_1} & ... & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

称为 f 的雅可比矩阵 (Jacobian), 它在 $\mathbf{x}_0$ 处的值

$$J_f(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_1}{\partial x_n}|_{\mathbf{x}_0} \\ ... & ... & ... \\ \frac{\partial f_m}{\partial x_1}|_{\mathbf{x}_0} & ... & \frac{\partial f_m}{\partial x_n}|_{\mathbf{x}_0} \end{pmatrix}$$

称为 f 在 $\mathbf{x}_0$ 处的雅可比矩阵, 它是微分 $df_{\mathbf{x}_0}$ 在标准基底下的矩阵表示.

注 2.2.14. 也可以把雅可比矩阵记作 $\frac{\partial(f_1,...,f_m)}{\partial(x_1,...,x_n)}|_{\mathbf{x}_0}$.

例 2.2.15. (1) $f : \mathbf{R} \to \mathbf{R}$ 的雅可比雅可比矩阵为 $J_f(x_0) = (f'(x_0))$.

(2) $f : \mathbf{R}^2 \to \mathbf{R}$ 的雅可比矩阵为 $J_f(x_0, y_0) = (\frac{\partial f}{\partial x}|_{(x_0,y_0)}, \frac{\partial f}{\partial y}|_{(x_0,y_0)})$.

(3) 设 $\gamma : \mathbf{R} \to \mathbf{R}^2$ 为

$$\gamma(t) = (x(t), y(t)),$$

则 γ 的雅可比矩阵为

$$J_\gamma(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

(4) 设 $f : \mathbf{R}^2 \to \mathbf{R}^2$ 为

$$f(x, y) = (u(x, y), v(x, y)),$$

则 f 的雅可比矩阵为

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

2.3 复合函数的导函数: 链式法则

回忆一元复合函数的求导法则.

定理 2.3.1. 设 $f:\mathbf{R}\to\mathbf{R}$ 在 x 处可微, $g:\mathbf{R}\to\mathbf{R}$ 在 $f(x)$ 处可微, 则 $g\circ f:\mathbf{R}\to\mathbf{R}$ 在 x 处可微, 且有

$$d(g\circ f)_x=(dg)_{f(x)}\circ df_x,$$

由此可得

$$(g\circ f)'(x)=g'(f(x))\cdot f'(x).$$

可以把上述定理粗略的叙述成:"线性近似" 保持映射的复合. 不难把其推广到高维的情形.

定理 2.3.2 (复合的微分等于微分的复合)**.** 设 $f:\mathbf{R}^n\to\mathbf{R}^m$ 在 $\mathbf{x}_0$ 处可微, $g:\mathbf{R}^m\to\mathbf{R}^l$ 在 $f(\mathbf{x}_0)$ 处可微, 则 $g\circ f:\mathbf{R}^n\to\mathbf{R}^l$ 在 $\mathbf{x}_0$ 处可微, 且有

$$d(g\circ f)_{\mathbf{x}_0}=dg_{f(\mathbf{x}_0)}\circ df_{\mathbf{x}_0}.$$

证明: 由可微的定义, 存在线性映射 $A:\mathbf{R}^n\to\mathbf{R}^m$, $B:\mathbf{R}^m\to\mathbf{R}^l$, 使得

$$\begin{aligned}f(\mathbf{x}_0+\mathbf{h})&=f(\mathbf{x}_0)+A(\mathbf{h})+\alpha(\mathbf{h}),\\ g(f(\mathbf{x}_0)+\mathbf{v})&=g(f(\mathbf{x}_0))+B(\mathbf{v})+\beta(\mathbf{v}),\end{aligned}$$

其中, 映射 α,β 满足

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{\alpha(\mathbf{h})}{|\mathbf{h}|}=\mathbf{0},\quad \lim_{\mathbf{v}\to\mathbf{0}}\frac{\beta(\mathbf{v})}{|\mathbf{v}|}=\mathbf{0}.$$

由此可得

$$\begin{aligned}&(g\circ f)(\mathbf{x}_0+\mathbf{h})=g\left(f(\mathbf{x}_0)+A(\mathbf{h})+\alpha(\mathbf{h})\right)\\ =&g(f(\mathbf{x}_0))+B\left(A(\mathbf{h})\right)+B(\alpha(\mathbf{h}))+\beta\left(A(\mathbf{h})+\alpha(\mathbf{h})\right).\end{aligned}$$

(1) 我们来证明 $\lim\limits_{\mathbf{h}\to\mathbf{0}}\dfrac{B(\alpha(\mathbf{h}))}{|\mathbf{h}|}=\mathbf{0}$.

为此, 注意到 B 是线性映射, 因而是连续的, 从而有

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{B(\alpha(\mathbf{h}))}{|\mathbf{h}|}=\lim_{\mathbf{h}\to\mathbf{0}}B\left(\frac{\alpha(\mathbf{h})}{|\mathbf{h}|}\right)=B\left(\lim_{\mathbf{h}\to\mathbf{0}}\frac{\alpha(\mathbf{h})}{|\mathbf{h}|}\right)=B(\mathbf{0})=\mathbf{0}.$$

(2) 我们来证明 $\lim\limits_{\mathbf{h}\to\mathbf{0}}\dfrac{\beta\left(A(\mathbf{h})+\alpha(\mathbf{h})\right)}{|\mathbf{h}|}=\mathbf{0}$.

定义 $p:\mathbf{R}^n\to\mathbf{R}^m$ 为

$$p(\mathbf{h})=A(\mathbf{h})+\alpha(\mathbf{h}),$$

定义 $q:\mathbf{R}^m\to\mathbf{R}^l$ 为

$$q(\mathbf{v})=\begin{cases}\frac{\beta(\mathbf{v})}{|\mathbf{v}|}, & \text{如果}\mathbf{v}\neq\mathbf{0},\\ \mathbf{0}, & \text{如果}\mathbf{v}=\mathbf{0},\end{cases}$$

则有

$$\lim_{\mathbf{h}\to\mathbf{0}}p(\mathbf{h})=\mathbf{0},\quad \lim_{\mathbf{v}\to\mathbf{0}}q(\mathbf{v})=q(\mathbf{0})=\mathbf{0},$$

由复合映射的极限定理, 有

$$\lim_{\mathbf{h}\to\mathbf{0}}q\circ p(\mathbf{h})=\mathbf{0}.$$

在 $\mathbf{0}\in\mathbf{R}^n$ 的某个去心邻域中, $\frac{|A(\mathbf{h})+\alpha(\mathbf{h})|}{|\mathbf{h}|}$ 是有界的, 因而有

$$\lim_{\mathbf{h}\to 0}(q\circ p(\mathbf{h}))\cdot\frac{|A(\mathbf{h})+\alpha(\mathbf{h})|}{|\mathbf{h}|}=\mathbf{0}.$$

注意到

$$(q\circ p(\mathbf{h}))\cdot\frac{|A(\mathbf{h})+\alpha(\mathbf{h})|}{|\mathbf{h}|}=\frac{\beta\left(A(\mathbf{h})+\alpha(\mathbf{h})\right)}{|\mathbf{h}|},\quad \forall\mathbf{h}\in\mathbf{R}^n,$$

这就证明了 $\lim\limits_{\mathbf{h}\to\mathbf{0}}\dfrac{\beta\left(A(\mathbf{h})+\alpha(\mathbf{h})\right)}{|\mathbf{h}|}=\mathbf{0}$.

结合 (1), (2) 可知, 复合映射 $g\circ f$ 在 $\mathbf{x}_0$ 处可微, 且其微分为

$$d(g\circ f)_{\mathbf{x}_0}(\mathbf{h})=B\left(A(\mathbf{h})\right)=dg_{f(\mathbf{x}_0)}(df_{\mathbf{x}_0}(\mathbf{h})),$$

这就完成了定理的证明. □

推论 2.3.3. 设 $f:\mathbf{R}^n\to\mathbf{R}^m$ 在 $\mathbf{x}_0$ 处可微, $g:\mathbf{R}^m\to\mathbf{R}^l$ 在 $f(\mathbf{x}_0)$ 处可微, 则

$$J_{g\circ f}(\mathbf{x}_0)=J_g(f(\mathbf{x}_0))J_f(\mathbf{x}_0).$$

定理 2.3.4 (链式法则, *chain rule*)**.** 设 $f:\mathbf{R}^n\to\mathbf{R}^m$ 与 $g:\mathbf{R}^m\to\mathbf{R}^l$ 都是 C^1 光滑的, 则对任何 $\mathbf{x}\in\mathbf{R}^n$, 有

$$J_{g\circ f}(\mathbf{x})=J_g(f(\mathbf{x}))J_f(\mathbf{x}).$$

设 $\mathbf{R}^n$ 的坐标为 $x_1,...,x_n$, $\mathbf{R}^m$ 的坐标为 $y_1,...,y_m$, 则可把上述链式法则具体写出来:

$$\frac{\partial(g\circ f)_i}{\partial x_j}|_{\mathbf{x}}=\sum_{k=1}^m\frac{\partial g_i}{\partial y_k}|_{f(\mathbf{x})}\cdot\frac{\partial f_k}{\partial x_j}|_{\mathbf{x}},\quad \forall 1\le i\le l,\quad 1\le j\le n.$$

例 2.3.5. 考虑 $l=1$ 的特殊情形. 若记 $y_k=f_k(x_1,...,x_n), z=g(y_1,...,y_m)$, 则得到链式法则的经典表述

$$\frac{\partial z}{\partial x_j}|_{(x_1,...,x_n)}=\sum_{k=1}^m\frac{\partial z}{\partial y_k}|_{(y_1,...,y_m)}\cdot\frac{\partial y_k}{\partial x_j}|_{(x_1,...,x_n)},$$

或者进一步记作

$$\frac{\partial z}{\partial x_j}=\sum_{k=1}^m\frac{\partial z}{\partial y_k}\frac{\partial y_k}{\partial x_j}.$$

假设下面例子中的映射都是 C^1 光滑的.

例 2.3.6. 设 f 是 C^1 光滑的, $\mathbf{x}=(x_1,...,x_n),\mathbf{v}=(q_1,...,q_n)$, 则 f 沿 $\mathbf{v}$ 的方向导数为:

$$\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}}=\frac{df(x_1+tq_1,...,x_n+tq_n)}{dt}|_{t=0}=\sum_{i=1}^n\frac{\partial f(x_1,...,x_n)}{\partial x_i}\cdot q_i.$$

定义

$$\nabla f|_{(x_1,...,x_n)}=\mathrm{grad}f|_{(x_1,...,x_n)}:=(\frac{\partial f(x_1,...,x_n)}{\partial x_1},...,\frac{\partial f(x_1,...,x_n)}{\partial x_n}),$$

称之为 f 在 $(x_1,...,x_n)$ 处的梯度方向 (*gradiant*), 这是 f 在 $(x_1,...,x_n)$ 点处增长最快的方向. 实际上, 对任何 $\mathbf{v}\neq 0$, 有:

$$\frac{\frac{\partial f}{\partial \mathbf{v}}|_{\mathbf{x}}}{|\mathbf{v}|}=\frac{\mathrm{grad}f|_{\mathbf{x}}\cdot\mathbf{v}}{|\mathbf{v}|}\leq|\mathrm{grad}f|_{\mathbf{x}}|,$$

当 $\mathbf{v}=\lambda\mathrm{grad}f|_{\mathbf{x}}(\lambda\in\mathbf{R}^+)$ 时上述等号成立.

命题 2.3.7. *如果 f 可微, 且 $gradf|_{\mathbf{p}_0}\neq 0$, 则 f 在 $\mathbf{p}_0$ 处沿梯度方向 $gradf|_{\mathbf{p}_0}$ 增加最快, 沿负梯度方向 $-gradf|_{\mathbf{p}_0}$ 增加最慢.*

例 2.3.8. 设 $f:\mathbf{R}^n\to\mathbf{R}$ 为二次函数

$$f(x_1,...,x_n)=x_1^2+...+x_k^2-x_{k+1}^2-...-x_{k+l}^2,$$

其中 $k+l\leq n$. 则 f 的梯度方向为

$$\mathrm{grad}f|_{(x_1,...,x_n)}=(2x_1,...,2x_k,-2x_{k+1},...,-2x_{k+l},0,...,0).$$

例 2.3.9. 设 $f:\mathbf{R}^3\to\mathbf{R}$ 是 C^1 光滑的, 则:

$$\frac{d}{dt}f(tx,ty,tz)=f_x(tx,ty,tz)x+f_y(tx,ty,tz)y+f_z(tx,ty,tz)z.$$

结合 Newton-Leibniz 公式可得

$$\begin{aligned}&f(x,y,z)-f(0,0,0)\\=&\int_0^1\frac{d}{dt}f(tx,ty,tz)dt\\=&x\int_0^1 f_x(tx,ty,tz)dt+y\int_0^1 f_y(tx,ty,tz)dt+z\int_0^1 f_z(tx,ty,tz)dt.\end{aligned}$$

例 2.3.10 (Euler 等式)**.** 称 $f:\mathbf{R}^3\to\mathbf{R}$ 是 n 次齐次函数, 如果对任何 $x,y,z,t\in\mathbf{R}$, 有

$$f(tx,ty,tz)=t^nf(x,y,z).$$

如果 f 是 C^1 光滑的 n 次齐次函数, 则有

$$xf_x(x,y,z)+yf_y(x,y,z)+zf_z(x,y,z)=nf(x,y,z).$$

例 2.3.11. 设 $z=f(x,y)$. 利用坐标变换公式 $x=r\cos\theta, y=r\sin\theta$ 可以把 z 看成 r,θ 的函数. 计算 $\frac{\partial z}{\partial r},\frac{\partial z}{\partial\theta}$.

$$\frac{\partial z}{\partial r}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial r}=\cos\theta\frac{\partial z}{\partial x}+\sin\theta\frac{\partial z}{\partial y},$$

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

例 2.3.12. 设 $f:\mathbf{R}^2\to\mathbf{R}$ 是 C^1 光滑函数. 定义 $g(x,y)=f(e^y,xy)$, 则 g 的偏导数为:

$$\frac{\partial g}{\partial x}=f_x(e^y,xy)\cdot 0+f_y(e^y,xy)\cdot y,$$

$$\frac{\partial g}{\partial y}=f_x(e^y,xy)\cdot e^y+f_y(e^y,xy)\cdot x,$$

例 2.3.13. 设 $f:\mathbf{R}^3\to\mathbf{R}$, $g:\mathbf{R}^2\to\mathbf{R}$ 都是 C^1 光滑函数. 定义

$$\phi(x,y)=f(x,y,g(x,y)),$$

则 ϕ 的偏导数为:

$$\frac{\partial\phi}{\partial x}=f_1'(x,y,g(x,y))+f_3'(x,y,g(x,y))g_x(x,y),$$

$$\frac{\partial\phi}{\partial y}=f_2'(x,y,g(x,y))+f_3'(x,y,g(x,y))g_y(x,y),$$

其中 f_1',f_2',f_3' 分别表示 f 对三个坐标方向的偏导数.

命题 2.3.14 (逆映射的微分). 设 C^1 光滑函数 $f:\mathbf{R}^n\to\mathbf{R}^n$ 与 $g:\mathbf{R}^n\to\mathbf{R}^n$ 互为逆映射, 即

$$g\circ f(\mathbf{x})=\mathbf{x},\quad f\circ g(\mathbf{x})=\mathbf{x},\quad \forall \mathbf{x}\in\mathbf{R}^n.$$

则对任何 $\mathbf{x}\in\mathbf{R}^n$ 有 $dg_{f(\mathbf{x})}=(df_{\mathbf{x}})^{-1}$. 写成矩阵的形式, 有

$$J_{f^{-1}}(f(\mathbf{x}))=(J_f(\mathbf{x}))^{-\mathbf{1}}.$$

特别的, f 的雅可比矩阵 $J(f)_{\mathbf{x}}$ 处处都是可逆矩阵.

例 2.3.15. (1) $n=1$ 时, 得 $(f^{-1})'(f(x))=\frac{1}{f'(x)}$.

(2) $n=2$ 时, 设 $f:\mathbf{R}^2\to\mathbf{R}^2$ 为:

$$(x,y)\to(u(x,y),v(x,y)).$$

设 f 与 f^{-1} 都是 C^1 光滑的, 则 x,y 可以表示成 u,v 的 C^1 光滑函数 $x=x(u,v),y=y(u,v)$, 且可以计算它们的偏导数:

$$\begin{pmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{pmatrix}|_{(u(x,y),v(x,y))}=\begin{pmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{pmatrix}^{-1}|_{(x,y)}.$$

例 2.3.16. 3 维欧氏空间中的坐标变换公式

$$\begin{cases}x=r\sin\phi\cos\theta\\ y=r\sin\phi\sin\theta\\ z=r\cos\phi\end{cases}$$

把 r,ϕ,θ 表示成 x,y,z 的函数, 计算它们的偏导数.

2.4 高阶导数

定义 2.4.1. 设 f 在 D 中处处有偏导数 $\frac{\partial f}{\partial x_i}$, 则可以考虑偏导函数 $\frac{\partial f}{\partial x_i}$ 在某点 $\mathbf{x}_0$ 处对 x_j 的偏导, 称之为 f 在 $\mathbf{x}_0$ 处的二阶偏导数, 记作

$$\frac{\partial(\frac{\partial f}{\partial x_i})}{\partial x_j}|_{\mathbf{x}_0}=\frac{\partial^2 f}{\partial x_j\partial x_i}|_{\mathbf{x}_0}=\partial^2_{x_jx_i}f(\mathbf{x}_0)=f_{x_ix_j}(\mathbf{x}_0)=f''_{x_ix_j}(\mathbf{x}_0).$$

类似的, 可以归纳的定义 f 的 k 阶偏导数

$$\frac{\partial^k f}{\partial x_{i_k}...\partial x_{i_1}}|_{\mathbf{x}_0}=\partial^k_{x_{i_k}\cdots x_{i_1}}f(\mathbf{x}_0)=\frac{\partial}{\partial x_{i_k}}\left(\frac{\partial^{k-1}f}{\partial x_{i_{k-1}}...\partial x_{i_1}}\right)|_{\mathbf{x}_0}.$$

例 2.4.2. 设 f 在某个开集内有偏导数 $\frac{\partial f}{\partial x}=f_x$, $\frac{\partial f}{\partial y}=f_y$, 则可以考虑偏导函数的偏导数, 即 f 的二阶偏导数:

$$\frac{\partial^2 f}{\partial x^2}=f_{xx}=\frac{\partial}{\partial x}(\frac{\partial f}{\partial x}),$$

$$\frac{\partial^2 f}{\partial y\partial x}=f_{xy}=\frac{\partial}{\partial y}(\frac{\partial f}{\partial x}),$$

$$\frac{\partial^2 f}{\partial x\partial y}=f_{yx}=\frac{\partial}{\partial x}(\frac{\partial f}{\partial y}),$$

$$\frac{\partial^2 f}{\partial y^2}=f_{yy}=\frac{\partial}{\partial y}(\frac{\partial f}{\partial y}).$$

例 2.4.3. (1) 定义 $h:\mathbf{R}^2\backslash\{(0,0)\}\to\mathbf{R}$ 为 $h(x,y)=\log\sqrt{x^2+y^2}$. 计算 $\Delta h=\frac{\partial^2 h}{\partial x^2}+\frac{\partial^2 h}{\partial y^2}$.

(2) 设正整数 $n\geq 3$. 定义函数 $f:\mathbf{R}^n\to\mathbf{R}$ 为

$$f(\mathbf{x})=\frac{1}{|\mathbf{x}|^{n-2}}=\frac{1}{(\sum_{i=1}^n x_i^2)^{\frac{n-2}{2}}}.$$

计算

$$\Delta f=\left(\sum_{i=1}^n\frac{\partial^2}{\partial x_i^2}\right)f.$$

例 2.4.4. 令 $f(x,y)=e^{x^2+xy+y^2}$, 则有

$$f_{xy}=f_{yx}=e^{x^2+xy+y^2}(2x+y)(x+2y)+e^{x^2+xy+y^2},$$

似乎总有 $f_{xy}=f_{yx}$?

例 2.4.5. 定义函数

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

则 $f_{xy}(0,0)=-1$, $f_{yx}(0,0)=1$.

定理 2.4.6. 设二元函数 f 在点 (x_0,y_0) 的某邻域中处处有二阶偏导数 f_{xy},f_{yx}. 如果 f_{xy} 与 f_{yx} 在 (x_0,y_0) 处都连续, 则有 $f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0)$.

证明*:* 考虑极限

$$F=\lim_{(s,t)\to(0,0),st\neq 0}\frac{f(x_0+s,y_0+t)-f(x_0+s,y_0)-f(x_0,y_0+t)+f(x_0,y_0)}{st}.$$

注意到

$$\begin{aligned}&f(x_0+s,y_0+t)-f(x_0+s,y_0)-f(x_0,y_0+t)+f(x_0,y_0)\\=&(f(x_0+s,y_0+t)-f(x_0+s,y_0))-(f(x_0,y_0+t)-f(x_0,y_0)),\end{aligned}$$

对固定的 t, 令 $g(x)=f(x,y_0+t)-f(x,y_0)$, 它是 x 的函数. 对 g 使用 Lagrange 中值定理,

$$g(x_0+s)-g(x_0)=sg'(x_0+\theta_1 s)=s(f_x(x_0+\theta_1 s,y_0+t)-f_x(x_0+\theta_1 s,y_0)),$$

再对 $f_x(x_0+\theta_1 s,\cdot)$ 使用 Lagrange 中值定理, 得

$$g(x_0+s)-g(x_0)=stf_{xy}(x_0+\theta_1 s,y_0+\theta_2 t),$$

其中 $0<\theta_1,\theta_2<1$. 由于 f_{xy} 在 (x_0,y_0) 处连续, 有

$$F=\lim_{(s,t)\to(0,0),st\neq 0}f_{xy}(x_0+\theta_1 s,y_0+\theta_2 t)=f_{xy}(x_0,y_0).$$

对称的, 如果 f_{yx} 在 (x_0,y_0) 处连续, 则 $F=f_{yx}(x_0,y_0)$. 这就完成了证明. □

定理 2.4.7 (二阶偏导的对称性, Clairaut-Schwarz)**.** 设 $D\subset\mathbf{R}^n$, 函数 $f:D\to\mathbf{R}$ 在 $\mathbf{x}_0$ 附近有二阶偏导数 $\frac{\partial^2 f}{\partial x_i\partial x_j},\frac{\partial^2 f}{\partial x_j\partial x_i}$. 如果 $\frac{\partial^2 f}{\partial x_i\partial x_j}$ 与 $\frac{\partial^2 f}{\partial x_j\partial x_i}$ 在 $\mathbf{x}_0$ 处连续, 则有 $\frac{\partial^2 f}{\partial x_j\partial x_i}|_{\mathbf{x}_0}=\frac{\partial^2 f}{\partial x_i\partial x_j}|_{\mathbf{x}_0}$.

证明*:* 不妨设 $i<j$, $\mathbf{x}_0=(a_1,\dots a_n)$. 定义二元函数 $g(x,y)$ 为

$$g(x,y)=f(a_1,...,a_{i-1},x,a_{i+1},...,a_{j-1},y,a_{j+1},...,a_n),$$

则有

$$g_{xy}(a_i,a_j)=f_{x_ix_j}(a_1,...,a_n),\quad g_{yx}(a_i,a_j)=f_{x_jx_i}(a_1,...,a_n),$$

由定理2.4.6的结论可得 $f_{x_ix_j}(\mathbf{x}_0)=f_{x_jx_i}(\mathbf{x}_0)$. □

定义 2.4.8. 设 $D\subset\mathbf{R}^n$ 是开集, 称 $f:D\to\mathbf{R}$ 是 C^2 光滑的, 如果 f 在 D 上所有偏导数与二阶偏导数都存在且连续, 记作 $f\in C^2(D;\mathbf{R})$.

推论 2.4.9. 设 D 是 $\mathbf{R}^n$ 的开集. 如果 $f\in C^2(D;\mathbf{R})$, 则在 D 上有 $\frac{\partial^2 f}{\partial x_j\partial x_i}=\frac{\partial^2 f}{\partial x_i\partial x_j}$.

2.5 多元函数的 Taylor 公式

设 $f:\mathbf{R}^n\to\mathbf{R}$ 是光滑函数. 我们用多项式函数来逼近 f, 这就是 Taylor 公式. 对于 $\mathbf{x}=(x_1,...,x_n),\mathbf{y}=(y_1,...,y_n)$, 考虑一元函数

$$g(t)=f(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))=f(x_1+t(y_1-x_1),...,x_n+t(y_n-x_n)).$$

由一元函数的 Taylor 公式可知, 存在 $\theta \in (0,1)$ 使得

$$g(1) = g(0) + \frac{g^{(1)}(0)}{1!} + ... + \frac{g^{(m-1)}(0)}{(m-1)!} + \frac{g^{(m)}(\theta)}{m!}. \tag{2.2}$$

下面我们来计算 g 的高阶导数. 利用链式法则, 有

$$\begin{aligned}
g^{(1)}(t) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}|_{\mathbf{x}+t(\mathbf{y}-\mathbf{x})} \cdot (y_i - x_i), \\
g^{(2)}(t) &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}|_{\mathbf{x}+t(\mathbf{y}-\mathbf{x})} \cdot (y_j - x_j) \right) \cdot (y_i - x_i) \\
&= \sum_{1 \le i,j \le n} \frac{\partial^2 f}{\partial x_j \partial x_i}|_{\mathbf{x}+t(\mathbf{y}-\mathbf{x})} \cdot (y_j - x_j)(y_i - x_i), \\
&\ldots\ldots \\
g^{(k)}(t) &= \sum_{1 \le i_1, ..., i_k \le n} \frac{\partial^k f}{\partial x_{i_k} ... \partial x_{i_1}}|_{\mathbf{x}+t(\mathbf{y}-\mathbf{x})} \cdot (y_{i_k} - x_{i_k}) .... (y_{i_1} - x_{i_1}), \\
&\ldots\ldots
\end{aligned}$$

注意到, 在计算 $g^{(m)}(t)$ 时, 需要使用 m 次链式法则, 因此需要假设 f 是 C^m 光滑的. 把这些计算结果代入 (2.2) 式, 即得到 f 的 Taylor 公式.

定理 2.5.1. *设 D 是 $\mathbf{R}^n$ 的开集, $\mathbf{x} = (x_1, ..., x_n), \mathbf{y} = (y_1, ..., y_n) \in D$ 且线段 $\overline{\mathbf{xy}} \subset D$. 设 $f \in C^m(D, \mathbf{R})$, 则存在 $\theta \in (0,1)$, 使得*

$$\begin{aligned}
f(\mathbf{y}) =& f(\mathbf{x}) + \frac{1}{1!} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}|_{\mathbf{x}} \cdot (y_i - x_i) + \cdots \\
&+ \frac{1}{(m-1)!} \sum_{1 \le i_1, ..., i_{m-1} \le n} \frac{\partial^{m-1} f}{\partial x_{i_{m-1}} ... \partial x_{i_1}}|_{\mathbf{x}} \cdot (y_{i_{m-1}} - x_{i_{m-1}}) .... (y_{i_1} - x_{i_1}) \\
&+ \frac{1}{m!} \sum_{1 \le i_1, ..., i_m \le n} \frac{\partial^m f}{\partial x_{i_m} ... \partial x_{i_1}}|_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})} \cdot (y_{i_m} - x_{i_m}) .... (y_{i_1} - x_{i_1}).
\end{aligned}$$

这里, 我们特别关心 $m = 1, 2$ 的情形.

定理 2.5.2 (微分中值定理)**.** *设 D 是 $\mathbf{R}^n$ 的开集, $\mathbf{x} = (x_1, ..., x_n), \mathbf{y} = (y_1, ..., y_n) \in D$ 且线段 $\overline{\mathbf{xy}} \subset D$. 设 $f : D \to \mathbf{R}$ 是 C^1 光滑函数, 则存在 $0 < \theta < 1$, 使得*

$$f(y_1, ..., y_n) - f(x_1, ..., x_n) = \sum_{i=1}^{n} (y_i - x_i) \frac{\partial f}{\partial x_i}|_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})}.$$

也可以把微分中值定理写成

$$\begin{aligned} f(\mathbf{y}) &= f(\mathbf{x}) + J_f(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x}) \\ &= f(\mathbf{x}) + \nabla f(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))\cdot(\mathbf{y}-\mathbf{x}), \end{aligned}$$

将一次项分别用 f 的雅可比矩阵与 f 的梯度向量表示.

例 2.5.3. 设 D 是 $\mathbf{R}^2$ 的开集, $P_0=(x_0,y_0), P_1=(x_0+\Delta x, y_0+\Delta y)\in D$ 且线段 $\overline{P_0P_1}\subset D$. 设 $f: D\to\mathbf{R}$ 是 C^1 光滑函数, 则存在 $0<\theta<1$, 使得

$$\begin{aligned} & f(x_0+\Delta x, y_0+\Delta y) - f(x_0,y_0) \\ =& f_x(x_0+\theta\Delta x, y_0+\theta\Delta y)\cdot\Delta x + f_y(x_0+\theta\Delta x, y_0+\theta\Delta y)\cdot\Delta y. \end{aligned}$$

推论 2.5.4. *设 D 是 $\mathbf{R}^n$ 的道路连通的开集. 如果 $f: D\to\mathbf{R}$ 在 D 上的各个偏导函数恒为 0, 则 f 是常值函数.*

证明: 对任何开球 $B\subset D$, 对 B 任何两点 $\mathbf{x},\mathbf{y}$, 由微分中值定理, 存在 $\theta\in(0,1)$, 使得

$$f(\mathbf{y})-f(\mathbf{x}) = \sum_{i=1}^n \partial_i f(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))\cdot(y_i-x_i) = 0,$$

可知 f 在 D 中的任何开球上都是常值.

对 D 中任何两点 $\mathbf{x}_0,\mathbf{x}_1$, 由于 D 道路连通, 存在连续道路 $p:[0,1]\to D$ 使得 $p(0)=\mathbf{x}_0, p(1)=\mathbf{x}_1$. 对每个 $t\in[0,1]$, 由于 D 是开集, 存在以 $p(t)$ 为球心的开球 $B_t\subset D$, 由前述 f 在每个 B_t 上是常值. 注意到, 道路 p 的像集 $Y=p([0,1])$ 是紧致集合 $[0,1]$ 在连续映射 p 下的像集, 因而 Y 紧致. 这样, Y 的开覆盖 $\{B_t\}_{t\in[0,1]}$ 有有限的子覆盖, 设为

$$Y\subset B_{t_1}\cup\cdots\cup B_{t_k},$$

由此可知 f 在 Y 上只取有限个值, 即连续函数 $f\circ p$ 只有有限个像点, 结合介值定理可得 $f\circ p$ 只能有一个像点, 特别的 $f(p(0))=f(p(1))$. 这就证明了 f 在 D 中任何两个点上的值都相等, 即 f 在 D 上是常值. □

定理 2.5.5 (展开至二阶的带 Lagrange 余项的 Taylor 公式)**.** *设 D 是 $\mathbf{R}^n$ 的开集, $\mathbf{x}=(x_1,...,x_n), \mathbf{y}=(y_1,...,y_n)\in D$ 且线段 $\overline{\mathbf{xy}}\subset D$. 设 $f: D\to\mathbf{R}$ 是 C^2 光滑函数, 则存在 $0<\theta<1$, 使得*

$$\begin{aligned} f(y_1,...,y_n) = f(x_1,...,x_n) &+ \sum_{i=1}^n \frac{\partial f}{\partial x_i}|_{\mathbf{x}}\cdot(y_i-x_i) \\ &+ \frac{1}{2}\sum_{1\le i,j\le n}\frac{\partial^2 f}{\partial x_j\partial x_i}|_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})}\cdot(y_j-x_j)(y_i-x_i). \end{aligned}$$

定义 2.5.6. 定义 f 在 $\mathbf{x}$ 处的海瑟矩阵 (*Hessian*) 为

$$H_f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}|_{\mathbf{x}} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}|_{\mathbf{x}} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}|_{\mathbf{x}} & \cdots & \frac{\partial^2 f}{\partial x_n^2}|_{\mathbf{x}} \end{pmatrix},$$

如果 f 是 C^2 光滑的, 则 $H_f(\mathbf{x})$ 是对称矩阵.

这样, 可以把展开至 2 阶的 Taylor 公式写成

$$f(\mathbf{y}) = f(\mathbf{x}) + J_f(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \frac{1}{2}(\mathbf{y}-\mathbf{x})^T H_f(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x}).$$

例 2.5.7. 设 $f \in C^2(D)$. 对任何两点 $P_0 = (x_0, y_0), P_1 = (x_0+\Delta x, y_0+\Delta y) \in D$, 如果线段 $\overline{P_0P_1} \subset D$, 则存在 $0 < \theta < 1$, 使得

$$\begin{aligned} f(x_0+\Delta x, y_0+\Delta y) = {} & f(x_0,y_0) + f_x(x_0,y_0)\Delta x + f_y(x_0,y_0)\Delta y \\ & + \frac{1}{2}\left((\Delta x)^2 f_{xx} + 2\Delta x \Delta y f_{xy} + (\Delta y)^2 f_{yy}\right)|_{(x_0+\theta\Delta x, y_0+\theta\Delta y)}. \end{aligned}$$

2.5.1 带 Peano 余项的 Taylor 公式

利用带 Lagrange 余项的 Taylor 公式, 可以得到带 Peano 余项的 Taylor 公式.

定理 2.5.8. 设 $f: D \to \mathbf{R}$ 在点 $\mathbf{x}_0$ 的某个开球邻域中是 C^m 光滑的, 则当 $\mathbf{h} = (h_1, ..., h_n) \to 0$ 时, 有

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h}) = {} & f(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}|_{\mathbf{x}_0} \cdot h_i + ... \\ & + \frac{1}{m!} \sum_{1 \le i_1, \ldots, i_m \le n} \frac{\partial^m f}{\partial x_{i_m} ... \partial x_{i_1}}|_{\mathbf{x}_0} \cdot h_{i_m} ... h_{i_1} + o(|\mathbf{h}|^m). \end{aligned}$$

证明*:* 利用定理2.5.1可得

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h}) = {} & f(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}|_{\mathbf{x}_0} \cdot h_i + \cdots + \frac{1}{(m-1)!} \sum_{1 \le i_1, \ldots, i_{m-1} \le n} \frac{\partial^{m-1} f}{\partial x_{i_{m-1}} ... \partial x_{i_1}}|_{\mathbf{x}_0} \cdot h_{i_{m-1}} ... h_{i_1} \\ & + \frac{1}{m!} \sum_{1 \le i_1, \ldots, i_m \le n} \frac{\partial^m f}{\partial x_{i_m} ... \partial x_{i_1}}|_{\mathbf{x}_0+\theta\mathbf{h}} \cdot h_{i_m} ... h_{i_1}, \end{aligned}$$

由此可得

$$f(\mathbf{x}_0+\mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}|_{\mathbf{x}_0} \cdot h_i + \cdots + \frac{1}{m!} \sum_{1 \le i_1, \ldots, i_m \le n} \frac{\partial^m f}{\partial x_{i_m} ... \partial x_{i_1}}|_{\mathbf{x}_0} \cdot h_{i_m} ... h_{i_1} + R(\mathbf{h}),$$

其中余项 $R(\mathbf{h})$ 为

$$R(\mathbf{h})=\frac{1}{m!}\sum_{1\le i_1,\dots,i_m\le n}\left(\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0+\theta\mathbf{h}}-\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0}\right)\cdot h_{i_m}....h_{i_1}.$$

利用绝对值不等式以及 $\frac{|h_i|}{|\mathbf{h}|}\le 1$ 可得

$$\frac{|R(\mathbf{h})|}{|\mathbf{h}|^m}\le\frac{1}{m!}\sum_{1\le i_1,\dots,i_m\le n}\left|\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0+\theta\mathbf{h}}-\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0}\right|.$$

由 f 在 $\mathbf{x}_0$ 附近 C^m 光滑, 可知 $\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}$ 在 $\mathbf{x}_0$ 处连续, 结合 $\theta\in(0,1)$ 可知

$$\lim_{\mathbf{h}\to 0}\left|\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0+\theta\mathbf{h}}-\frac{\partial^m f}{\partial x_{i_m}\dots\partial x_{i_1}}|_{\mathbf{x}_0}\right|=0.$$

这样, 利用夹逼定理可得

$$\lim_{\mathbf{h}\to 0}\frac{R(\mathbf{h})}{|\mathbf{h}|^m}=0,$$

这就完成了定理的证明. □

比较常用的是展开至 2 阶的带 Peano 余项的 Taylor 公式.

推论 2.5.9. 设 $f:D\to\mathbf{R}$ 在点 $\mathbf{x}_0$ 的某个开邻域中是 C^2 光滑的, 则当 $\mathbf{h}=(h_1,...,h_n)\to 0$ 时, 有

$$f(\mathbf{x}_0+\mathbf{h})=f(\mathbf{x}_0)+J_f(\mathbf{x}_0)\mathbf{h}+\frac{1}{2}\mathbf{h}^T H_f(\mathbf{x}_0)\mathbf{h}+o(|\mathbf{h}|^2).$$

例 2.5.10. 设 $f\in C^1(D)$, 则在 (x_0,y_0) 附近有

$$f(x_0+\Delta x,y_0+\Delta y)=f(x_0,y_0)+f_x(x_0,y_0)\Delta x+f_y(x_0,y_0)\Delta y+o\left(\sqrt{(\Delta x)^2+(\Delta y)^2}\right).$$

类似的, 如果 $f\in C^2(D)$, 则在 (x_0,y_0) 附近有

$$\begin{aligned}f(x_0+\Delta x,y_0+\Delta y)&=f(x_0,y_0)+f_x(x_0,y_0)\Delta x+f_y(x_0,y_0)\Delta y\\&\quad+\frac{1}{2}\left((\Delta x)^2f_{xx}+2\Delta x\Delta yf_{xy}+(\Delta y)^2f_{yy}\right)|_{(x_0,y_0)}+o\left((\Delta x)^2+(\Delta y)^2\right).\end{aligned}$$

例 2.5.11. 求 $f(x,y)=\ln(2+x+y+xy)$ 在 $(0,0)$ 附近的 Taylor 公式.

2.6　反函数定理

定义 2.6.1. 设 $D \subset \mathbf{R}^n, E \subset \mathbf{R}^m$ 是欧氏空间的开集, $f: D \to E$ 是 C^1 光滑映射. 称 f 有 C^1 光滑的逆, 如果存在 C^1 光滑映射 $g: E \to D$ 满足:

$$g \circ f = Id_D, \quad f \circ g = Id_E.$$

称这样的 g 为 f 的 (C^1 光滑的) 逆映射, 记作 f^{-1}.

命题 2.6.2. *设 $f: D \to E$ 有 C^1 光滑的逆 $f^{-1}: E \to D$, 则对任何 $\mathbf{x} \in D$, 有*

$$(df^{-1})_{f(\mathbf{x})} \circ (df)_{\mathbf{x}} = Id, \quad (df)_{\mathbf{x}} \circ (df^{-1})_{f(\mathbf{x})} = Id.$$

写成矩阵形式即

$$J(f^{-1})_{f(\mathbf{x})} \cdot J(f)_{\mathbf{x}} = I_{n\times n}, \quad J(f)_{\mathbf{x}} \cdot J(f^{-1})_{f(\mathbf{x})} = I_{m\times m}.$$

特别的, $m = n$(D 与 E 的维数 dimension 相同), 且 $df_{\mathbf{x}}$ 是可逆的线性映射, $J(f)_{\mathbf{x}}$ 是可逆矩阵.

注 2.6.3. 如果 $f: D \to E$ 有 C^1 光滑的逆, 则 D 与 E 的形状 "本质上一样", 可认为 f 把 D 与 E"等同" 起来.

问题 2.6.4. *(1) 对于给定的映射 $f: D \to E$, 它是否有 C^1 光滑的逆映射?*

(2) 对于给定的 D, E, 是否存在 C^1 光滑映射 $f: D \to E$ 使得 f 有 C^1 光滑的逆?

注 2.6.5. 这些都是非常困难的问题, 数学中回答这些问题的分支叫 "拓扑学"(*Topology*).

找一个整体上可逆的映射很困难, 但很容易保证在一个点附近可逆.

定义 2.6.6. 称 U 是 $\mathbf{x}$ 的开邻域, 如果 $\mathbf{x} \in U$ 且 U 是开集.

定理 2.6.7 (反函数定理)**.** *设 D 是 $\mathbf{R}^n$ 的开集, $f: D \to \mathbf{R}^n$ 是 C^1 光滑映射. 如果 f 在 $\mathbf{x}_0 \in D$ 处的微分 $df_{\mathbf{x_0}}$ 是可逆的 (或等价的, 雅可比矩阵 $J_f(\mathbf{x}_0)$ 是可逆矩阵), 则存在 $\mathbf{x}_0$ 的开邻域 U 与 $f(\mathbf{x_0})$ 的开邻域 V, 使得 $f: U \to V$ 是双射, 且它的逆映射 $f^{-1}: V \to U$ 是 C^1 光滑的. 更进一步, 对 $\forall \mathbf{x} \in U$, 有*

$$(df^{-1})_{f(\mathbf{x})} = (df_{\mathbf{x}})^{-1}, \quad J_{f^{-1}}(f(\mathbf{x})) = (J_f(\mathbf{x}))^{-1}.$$

反函数定理的解释. 适当的坐标平移以后, 可假设 $\mathbf{x}_0 = \mathbf{0}, f(\mathbf{x}_0) = \mathbf{0}$. 令 $A = J_f(\mathbf{0})$, 它是可逆矩阵.

(1) 非常粗略的看法. 在 $\mathbf{x}_0 = \mathbf{0}$ 附近

$$f(\mathbf{x}) = A\mathbf{x} + (\text{比一阶更高阶的误差项}),$$

因此在 $\mathbf{0}$ 附近 f 可近似成

$$f(\mathbf{x}) \sim A\mathbf{x},$$

由于 A 是可逆矩阵, 线性映射 $\mathbf{x} \to A\mathbf{x}$ 有逆 $\mathbf{y} \to \mathbf{A^{-1}y}$, 所以有理由期待 f 在 $\mathbf{0}$ 附近也可逆.

(2) 稍微精确一点的看法. 在 $\mathbf{x}_0 = \mathbf{0}$ 附近有

$$f(\mathbf{x}) = A\mathbf{x} + g(\mathbf{x}),$$

其中 g 是比一阶更高阶的误差项 (换句话说, $g(\mathbf{x})$ 与 $A\mathbf{x}$ 相比很小). 下面我们来证明 f 在 $\mathbf{0}$ 附近是满的, 即对给定的 $\mathbf{y}$, 求解方程:

$$A\mathbf{x} + g(\mathbf{x}) = \mathbf{y},$$

物理学家们喜欢用如下方法:

$$\begin{aligned}
\mathbf{x} &= A^{-1}(\mathbf{y} - g(\mathbf{x})) \quad \rightsquigarrow \text{得到一级近似解} \mathbf{x}_1 = A^{-1}\mathbf{y}, \\
&= A^{-1}(\mathbf{y} - g(A^{-1}(\mathbf{y} - g(\mathbf{x})))) \quad \rightsquigarrow \text{得到二级近似解} \mathbf{x}_2 = A^{-1}(\mathbf{y} - g(A^{-1}\mathbf{y})), \\
&= A^{-1}(\mathbf{y} - g(A^{-1}(\mathbf{y} - g(A^{-1}(\mathbf{y} - g(\mathbf{x})))))) \\
&\quad \rightsquigarrow \text{得到三级近似解} \mathbf{x}_3 = A^{-1}(\mathbf{y} - g(A^{-1}(\mathbf{y} - g(A^{-1}\mathbf{y})))), \\
&= \cdots\cdots\cdots\cdots
\end{aligned}$$

因为 g 很小, 可相信上述过程给出越来越好的近似解 $\mathbf{x}_n$.

数学家们证明了, 对于模长很小的 $\mathbf{y}$, 点列

$$\mathbf{x}_0 = \mathbf{0}, \quad \mathbf{x}_n = A^{-1}(\mathbf{y} - g(\mathbf{x}_{n-1})),$$

有极限, 设为 $\mathbf{x} = \lim\limits_{n\to\infty} \mathbf{x}_n$. 对递推式 $\mathbf{x}_n = A^{-1}(\mathbf{y} - g(\mathbf{x}_{n-1}))$ 两边取极限, 得

$$\mathbf{x} = A^{-1}(\mathbf{y} - g(\mathbf{x})),$$

因此 $\mathbf{x}$ 是精确解.

□

定理 2.6.8 (压缩映像)**.** 设 (X,d) 是完备的度量空间, $T:X\to X$ 是压缩映射, 即存在 $0<c<1$ 使得

$$d(T(x),T(y))\le c\cdot d(x,y),\quad \forall x,y\in X,$$

则 T 有唯一的不动点.

证明: 任取 $x_0\in X$, 考虑点列

$$x_k=\underbrace{T\circ\cdots\circ T}_{k\text{个}}(x_0),\quad \forall k\in\mathbf{Z}_+.$$

注意到, 对任何正整数 m,k, 有

$$\begin{aligned}
d(x_m,x_{m+k})&=d(T(x_{m-1}),T(x_{m+k-1}))\le c\cdot d(x_{m-1},x_{m-1+k})\\
&\le\cdots\le c^m\cdot d(x_0,x_k)\le c^m\cdot(d(x_0,x_1)+\cdots+d(x_{k-1},x_k))\\
&\le c^m\cdot\big(d(x_0,x_1)+\cdots+c^{k-1}d(x_0,x_1)\big)=c^m(1+c+\cdots+c^{k-1})d(x_0,x_1)\\
&<\frac{c^m}{1-c}d(x_0,x_1),
\end{aligned}$$

可知 $\{x_k\}$ 是 Cauchy 列. 由于 X 完备, 故 $\{x_k\}$ 有极限, 设为 $\lim\limits_{k\to\infty}x_k=x^*$.

映射 T 是连续的, 因为在任何点 a 处, 对任何 $\epsilon>0$, 取 $\delta=\frac{\epsilon}{c}$, 则只要 x 满足 $d(a,x)<\delta$, 就有 $d(T(a),T(x))\le c\cdot d(a,x)<c\delta=\epsilon$. 这样, 利用 T 的连续性可得

$$x^*=\lim_{k\to\infty}x_{k+1}=\lim_{k\to\infty}T(x_k)=T(\lim_{k\to\infty}x_k)=T(x^*),$$

即 x^* 是 T 的不动点. 若 y 也是 T 的不动点, 则有

$$d(x^*,y)=d(T(x^*),T(y))\le c\cdot d(x^*,y),$$

结合 $c\in(0,1)$ 可知 $d(x^*,y)=0$, 这表明 x^* 是 T 唯一的不动点. □

反函数定理的证明. 设 $J_f(\mathbf{x}_0)=A$, 不妨设 $\mathbf{x}_0=\mathbf{0},f(\mathbf{x}_0)=\mathbf{0},A=I$, 因为可以转化成考虑映射

$$\widetilde{f}(\mathbf{x})=A^{-1}\cdot(f(\mathbf{x}+\mathbf{x}_0)-f(\mathbf{x}_0))$$

在 $\mathbf{x}=\mathbf{0}$ 附近的可逆性. 在此假设下, f 在 $\mathbf{0}$ 附近可写成 $f(\mathbf{x})=\mathbf{x}+\alpha(\mathbf{x})$, 其中 α 是 C^1 光滑的, 且 $J_\alpha(\mathbf{0})=0$. 由 J_α 的矩阵元 $\frac{\partial\alpha_i}{\partial x_j}$ 的连续性, 存在正数 R, 使得

$$|\frac{\partial\alpha_i}{\partial x_j}(\mathbf{x})|\le\frac{1}{2n^2},\quad \forall|\mathbf{x}|<2R,\forall i,j.$$

对任何 $\mathbf{x},\mathbf{y}\in B_{2R}(\mathbf{0})$, 对每个 $1\le i\le n$, 由微分中值定理知存在 θ 使得

$$\alpha_i(\mathbf{y})-\alpha_i(\mathbf{x})=\sum_{j=1}^{n}\frac{\partial\alpha_i}{\partial x_j}(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))(y_j-x_j),$$

由此可得,

$$|\alpha_i(\mathbf{y})-\alpha_i(\mathbf{x})|\le\sum_{j=1}^{n}\frac{1}{2n^2}\cdot|y_j-x_j|\le\frac{1}{2n}\cdot|\mathbf{y}-\mathbf{x}|,$$

从而有

$$|\alpha(\mathbf{y})-\alpha(\mathbf{x})|\le\sum_{i=1}^{n}|\alpha_i(\mathbf{x})-\alpha_i(\mathbf{y})|\le\frac{1}{2}\cdot|\mathbf{y}-\mathbf{x}|. \tag{2.3}$$

我们证明如下几个结论.

(1) f 在 $B_{2R}(\mathbf{0})$ 上是单射. 这是由于, 若 $\mathbf{x},\mathbf{y}\in B_{2R}(\mathbf{0})$ 满足 $f(\mathbf{x})=f(\mathbf{y})$, 则有 $\mathbf{x}-\mathbf{y}=\alpha(\mathbf{y})-\alpha(\mathbf{x})$, 结合不等式 (2.3) 可知 $\mathbf{x}=\mathbf{y}$.

(2) f 在 $\mathbf{0}$ 附近是满射. 我们证明对任何 $|\mathbf{w}|<\frac{R}{2}$, 存在 $\mathbf{x}\in\overline{B}_R(\mathbf{0})$ 使得 $f(\mathbf{x})=\mathbf{w}$. 考虑映射 $T(\mathbf{x})=\mathbf{w}-\alpha(\mathbf{x})$. 对每个 $\mathbf{x}\in\overline{B}_R(\mathbf{0})$, 由不等式 (2.3) 可知

$$|T(\mathbf{x})|\le|\mathbf{w}|+|\alpha(\mathbf{x})|<\frac{R}{2}+\frac{1}{2}|\mathbf{x}|\le R,$$

因此 T 是完备度量空间 $\overline{B}_R(\mathbf{0})$ 到自身的映射. 注意到, 由不等式 (2.3) 有

$$|T(\mathbf{x})-T(\mathbf{y})|=|\alpha(\mathbf{y})-\alpha(\mathbf{x})|\le\frac{1}{2}\cdot|\mathbf{y}-\mathbf{x}|,$$

可知 T 是压缩映射. 由定理2.6.8知 T 有不动点, 即存在 $\mathbf{x}\in\overline{B}_R(\mathbf{0})$ 使得 $\mathbf{w}-\alpha(\mathbf{x})=\mathbf{x}$, 亦即 $f(\mathbf{x})=\mathbf{w}$.

结合这两个结论, 令 $V=B_{\frac{R}{2}}(\mathbf{0})$, $U=f^{-1}(V)\cap B_{2R}(\mathbf{0})$, 则 $f:U\to V$ 是双射.

□

引理 2.6.9. *设 U,V 是 $\mathbf{R}^n$ 的开集, $f:U\to V$ 是 C^1 光滑的双射, 且对任何 $x\in U$ 都有 $J_f(x)$ 可逆, 则 $f^{-1}:V\to U$ 也是 C^1 光滑映射.*

证明: (1) f^{-1} 连续. 来证明 f^{-1} 在任何点 $f(x_0)$ 处连续, 为此对任何 $\epsilon>0$, 由上述定理中结论 (2) 的证明可知, 存在 $R<\frac{\epsilon}{2}$, 使得对任何 $w\in B_{\frac{R}{2}}(f(x_0))$ 都存在 $x\in\overline{B}_R(x_0)$ 使得 $f(x)=w$. 令 $\delta=\frac{R}{2}$, 则对每个 $w\in B_\delta(f(x_0))$ 有 $f^{-1}(w)\in\overline{B}_R(x_0)\subset B_\epsilon(x_0)$, 即验证了 $f^{-1}(B_\delta(f(x_0)))\subset B_\epsilon(x_0)$, 从而验证了 f^{-1} 在 $f(x_0)$ 处连续.

(2) f^{-1} 可微. 来证明 f^{-1} 在点 $f(x)$ 处可微. 不妨假设 $J_f(x)=I$, 由 f 在 x 处可微的定义有 $f(x+h)=f(x)+h+\alpha(h)$ 且 $\lim_{h\to0}\frac{\alpha(h)}{|h|}=0$. 令

$$f^{-1}(f(x)+v)=x+v+\beta(v),$$

只需证明 $\lim_{v\to 0}\frac{\beta(v)}{|v|}=0$, 就表明 f^{-1} 在 $f(x)$ 处可微. 令 $h=f^{-1}(f(x)+v)-x$ 换元, 利用复合极限定理可得

$$\lim_{v\to 0}\frac{\beta(v)}{|v|}=\lim_{v\to 0}\frac{f^{-1}(f(x)+v)-x-v}{|v|}=\lim_{h\to 0}\frac{-\alpha(h)}{|h+\alpha(h)|}=0.$$

(3) f^{-1} 是 C^1 光滑的. 由 (2) 的结论, f^{-1} 处处可微, 则由链式法则可得

$$J_{f^{-1}}(y)=\left(J_f(f^{-1}(y))\right)^{-1}.$$

注意到, 由 $f\in C^1$ 可知 J_f 的矩阵元都是连续函数, 而 $(J_f)^{-1}$ 的矩阵元都可表示成 J_f 矩阵元的有理式, 因而 $(J_f)^{-1}$ 的矩阵元都是连续函数, 再复合连续函数 $f^{-1}(y)$, 所得的仍然是连续函数. 这就证明了 $J_{f^{-1}}(y)$ 的矩阵元都是连续的, 即 f^{-1} 是 C^1 光滑映射.

□

例 2.6.10. $n=1$. 设 $f:D\to\mathbf{R}$ 有连续的导函数 f', 且 $f'(x_0)\neq 0$. 反函数定理告诉我们 f 在 x_0 附近可逆, 且逆映射也是 C^1 光滑的. 可以把这个结论用更朴素的语言叙述成: 设 $y=y(x)$ 是 x 的 C^1 光滑函数, 如果 $y'(x_0)\neq 0$, 则在 x_0 附近, x 可以表示成 y 的函数 $x=x(y)$, 且这个函数也是 C^1 光滑的.

例 2.6.11. $f:\mathbf{R}\to\mathbf{R}, f(x)=x^3$ 在 $x=0$ 附近没有 C^1 光滑的逆, 但在其他点附近有 C^1 光滑的逆 $f^{-1}(y)=y^{\frac{1}{3}}$.

例 2.6.12. $n=2$. 设 $D\subset\mathbf{R}^2$ 是开集, $f:D\to\mathbf{R}^2$ 是 C^1 光滑的, 它的分量为

$$f(x,y)=(u(x,y),v(x,y)).$$

反函数定理断言, 如果

$$J(f)_{(x_0,y_0)}=\begin{pmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{pmatrix}\Big|_{(x_0,y_0)}$$

是可逆矩阵, 则 f 在 (x_0,y_0) 的某个附近中是双射, 且它的逆映射是 C^1 光滑的. 更加具体的说, 在 (x_0,y_0) 附近可以把 x,y 表示成 u,v 的 C^1 光滑函数

$$x=x(u,v),\quad y=y(u,v).$$

例 2.6.13. 直角坐标与极坐标的变换公式为

$$\begin{cases}x=r\cos\theta,\\ y=r\sin\theta.\end{cases}$$

当 $r \neq 0$ 时, 上述坐标变换在 (r, θ) 附近有 C^1 光滑的逆变换. 但在 $(r = 0, \theta)$ 附近没有 C^1 光滑的逆变换.

例 2.6.14. 直角坐标与球坐标的变换法则为

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases}$$

当 $r\sin\theta \neq 0$ 时, 上述坐标变换在 (r, θ, ϕ) 附近有 C^1 光滑的逆变换; 当 $r\sin\theta = 0$ 时, 在 (r, θ, ϕ) 附近没有 C^1 光滑的逆变换.

2.7 隐函数定理

2.7.1 Motivation

问题 2.7.1. *给定 C^1 光滑函数 $F : \mathbf{R}^2 \to \mathbf{R}$. 求解方程 $F(x, y) = 0$.*

假设已经找到某个解 (x_0, y_0) 满足上述方程, 能否在它附近找到更多的解? 这等价于找 $\Delta x, \Delta y$ 使得:

$$F(x_0 + \Delta, y_0 + \Delta y) = 0.$$

注意到当 $\Delta x, \Delta y$ 很小时, 有如下近似

$$F(x_0 + \Delta, y_0 + \Delta y) \sim F(x_0, y_0) + F_x(x_0, y_0)\Delta x + F_y(x_0, y_0)\Delta y,$$

因此当 $F_y(x_0, y_0) \neq 0$ 时, 对于任何很小的 Δx, 都可以找到方程 $F(x, y) = 0$ 的一个近似解

$$(x_0 + \Delta x, y_0 - \frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}\Delta x).$$

问题 2.7.2. *对 x_0 附近的 x, 能否找到精确解 $F(x, y) = 0$?*

2.7.2 隐函数定理

定理 2.7.3 (隐函数定理)**.** *设 $F : \mathbf{R}^2 \to \mathbf{R}$ 是 C^1 光滑函数, $F(x_0, y_0) = 0$.*

(1) *如果 $F_y(x_0, y_0) \neq 0$, 则存在 x_0 的某个开邻域 U, 以及 C^1 光滑函数 $g : U \to \mathbf{R}^1$, 使得*

$$F(x, g(x)) = 0, \quad \forall x \in U.$$

进一步, 有

$$g'(x) = -\frac{F_x(x, g(x))}{F_y(x, g(x))}, \quad \forall x \in U.$$

(2) 如果 $F_x(x_0, y_0) \neq 0$, 则存在 y_0 的某个开邻域 V, 以及 C^1 光滑函数 $h : V \to \mathbf{R}^1$, 使得

$$F(h(y), y) = 0, \quad \forall y \in V.$$

进一步, 有

$$h'(y) = -\frac{F_y(h(y), y)}{F_x(h(y), y)}, \quad \forall y \in V.$$

证明*:* 设 $F_y(x_0, y_0) \neq 0$. 定义 $f : \mathbf{R}^2 \to \mathbf{R}^2$ 为

$$f(x, y) = (x, F(x, y)), \quad \forall (x, y) \in \mathbf{R}^2.$$

f 在点 (x_0, y_0) 处的雅可比矩阵为

$$J(f)_{(x_0,y_0)} = \begin{pmatrix} 1 & 0 \\ F_x(x_0, y_0) & F_y(x_0, y_0) \end{pmatrix},$$

它的行列式为

$$detJ(f)_{(x_0,y_0)} = F_y(x_0, y_0) \neq 0,$$

特别的, $J(f)_{(x_0,y_0)}$ 是可逆矩阵. 由反函数定理, 存在 (x_0, y_0) 的某个开邻域 N 与 $(x_0, 0)$ 的某个开邻域 W, 使得 $f : N \to W$ 是双射, 且 $f^{-1} : W \to N$ 是 C^1 光滑的. 这样就有

$$\begin{aligned} (x, y) \in N \text{ 满足方程} F(x, y) = 0 \iff & \quad (x, y) \in N \text{ 且} f(x, y) = (x, 0) \\ \iff & \quad (x, 0) \in W \text{ 且} (x, y) = f^{-1}(x, 0) \\ \iff & \quad (x, 0) \in W \text{ 且} y = p_2(f^{-1}(x, 0)) \end{aligned}$$

其中 $p_2 : \mathbf{R}^2 \to \mathbf{R}$ 是向第二个坐标轴的投影. 取 x_0 在 $\mathbf{R}$ 中的开集 U 满足 $U \times \{0\} \subset W$. 定义 $g : U \to \mathbf{R}$ 为

$$g(x) = p_2(f^{-1}(x, 0)),$$

则 g 是 C^1 光滑函数, 且满足: 对于 (x_0, y_0) 附近的点 (x, y),

$$F(x, y) = 0 \iff y = g(x).$$

最后我们来计算 $g'(x)$. 对恒等式

$$F(x, g(x)) = 0 \quad \forall x \in U$$

两边求导得

$$F_x(x, g(x)) + g'(x)F_y(x, g(x)) = 0,$$

如果 $F_y(x,g(x))\neq 0$, 则

$$g'(x)=-\frac{F_x(x,g(x))}{F_y(x,g(x))}.$$

□

注 2.7.4. (1) 我们实际上证明了更强的结论. 如果 $F_y(x_0,y_0)\neq 0$, 则在 (x_0,y_0) 附近, 方程 $F(x,y)=0$ 的解全部由某个 C^1 光滑函数 g 给出:

$$F(x,y)=0 \iff y=g(x).$$

特别的, 在 (x_0,y_0) 附近, 解集 $Z(F)=\{(x,y)|F(x,y)=0\}$ 的形状与 $\mathbf{R}$ 一样.

(2) 类似的, 如果 $F_x(x_0,y_0)\neq 0$, 则 $Z(F)=\{(x,y)|F(x,y)=0\}$ 的形状与 $\mathbf{R}$ 一样.

(3) 如果 $F_x(x_0,y_0)=F_y(x_0,y_0)=0$, 则解集 $Z(F)=\{(x,y)|F(x,y)=0\}$ 在 (x_0,y_0) 附近的的形状可能比较奇怪. 例如 $F(x,y)=xy=0$ 的解集在 $(0,0)$ 附近就形如 "+".

定义 2.7.5. 称上述定理中的函数 g(或 h) 为由方程 $F(x,y)=0$ 决定的隐函数.

定理 2.7.6 (隐函数定理)**.** *设 $F:\mathbf{R}^3\to\mathbf{R}$ 是 C^1 光滑映射, $F(x_0,y_0,z_0)=0$. 如果 $F_z(x_0,y_0,z_0)\neq 0$, 则存在 (x_0,y_0) 的某个开邻域 U 以及 C^1 光滑函数 $g:U\to\mathbf{R}^1$, 使得*

$$F(x,y,g(x,y))=0,\quad \forall (x,y)\in U.$$

进一步, 有

$$\begin{cases}\frac{\partial g(x,y)}{\partial x}=-\frac{F_x(x,y,g(x,y))}{F_z(x,y,g(x,y))},\\ \frac{\partial g(x,y)}{\partial y}=-\frac{F_y(x,y,g(x,y))}{F_z(x,y,g(x,y))}.\end{cases}$$

注 2.7.7. 如果 $F_x(x_0,y_0,z_0),F_y(x_0,y_0,z_0),F_z(x_0,y_0,z_0)$ 不全为 0, 则方程 $F(x,y,z)=0$ 的解集

$$Z(F)=\{(x,y,z)|F(x,y,z)=0\}$$

在 (x_0,y_0,z_0) 附近的形状与 $\mathbf{R}^2$ 一样.

定理 2.7.8 (隐函数定理)**.** *设 $F,G:\mathbf{R}^3\to\mathbf{R}$ 是 C^1 光滑映射, $F(x_0,y_0,z_0)=G(x_0,y_0,z_0)=0$. 如果矩阵*

$$\begin{pmatrix} F_y(x_0,y_0,z_0) & F_z(x_0,y_0,z_0)\\ G_y(x_0,y_0,z_0) & G_z(x_0,y_0,z_0)\end{pmatrix}$$

是可逆矩阵, 则存在 x 的开邻域 U 以及 C^1 光滑函数 $g,h:U\to\mathbf{R}$, 使得对任何 $x\in U$, 有

$$\begin{cases}F(x,g(x),h(x))=0,\\ G(x,g(x),h(x))=0.\end{cases}$$

进一步, 有

$$\begin{pmatrix} g'(x)\\ h'(x)\end{pmatrix}=-\begin{pmatrix} F_y(x,g(x),h(x)) & F_z(x,g(x),h(x))\\ G_y(x,g(x),h(x)) & G_z(x,g(x),h(x))\end{pmatrix}^{-1}\cdot\begin{pmatrix} F_x(x,g(x),h(x))\\ G_x(x,g(x),h(x))\end{pmatrix}.$$

注 2.7.9. 如果矩阵

$$\begin{pmatrix} F_x(x_0,y_0,z_0) & F_y(x_0,y_0,z_0) & F_z(x_0,y_0,z_0)\\ G_x(x_0,y_0,z_0) & G_y(x_0,y_0,z_0) & G_z(x_0,y_0,z_0)\end{pmatrix}$$

的秩为 2, 则方程组

$$\begin{cases}F(x,y,z)=0,\\ G(x,y,z)=0.\end{cases}$$

的解集在 (x_0,y_0,z_0) 附近的形状与 $\mathbf{R}$ 一样.

例 2.7.10. 方程 $x^3+y^3-3axy=0, a>0$ 确定的隐函数.

例 2.7.11. 方程 $x^2+y^2+z^2-1=0$ 确定的隐函数.

例 2.7.12. 方程组

$$\begin{cases}x^2+y^2+z^2-1=0,\\ x^2+2y^2-z^2-1=0.\end{cases}$$

确定的隐函数.

2.8 一般维数的隐函数定理

问题 2.8.1. *给定 $k\le n$ 个 C^1 光滑的函数 $F_1,...,F_k:\mathbf{R}^n\to\mathbf{R}$, 考虑它们的公共零点集合*

$$S=\{(x_1,...,x_n)\in\mathbf{R}^n|F_i(x_1,...,x_n)=0,\forall 1\le i\le k\}.$$

S 的形状是怎么样的? 设 $\mathbf{x}_0=(a_1,...,a_n)\in S$, S 在 $\mathbf{x}_0$ 附近的形状是怎样的?

定理 2.8.2 (隐函数定理)**.** 给定 $k \le n$ 个 C^1 光滑的函数 $F_1, ..., F_k : \mathbf{R}^n \to \mathbf{R}$, 设 $\mathbf{x}_0 = (a_1, ..., a_n)$ 是它们的公共零点. 如果 $k \times k$ 矩阵

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_{n-k+1}}\Big|_{\mathbf{x}_0} & \cdots & \frac{\partial F_1}{\partial x_n}\Big|_{\mathbf{x}_0} \\ \cdots & \cdots & \cdots \\ \frac{\partial F_k}{\partial x_{n-k+1}}\Big|_{\mathbf{x}_0} & \cdots & \frac{\partial F_k}{\partial x_n}\Big|_{\mathbf{x}_0} \end{pmatrix}$$

是可逆矩阵, 则存在 $(a_1, ..., a_{n-k})$ 在 $\mathbf{R}^{n-k}$ 中的开邻域 U, 以及 k 个 C^1 光滑的函数 $g_{n-k+1}, ..., g_n : U \to \mathbf{R}$, 使得对每个 $(n-k+1) \le j \le n$ 有 $g_j(a_1, ..., a_{n-k}) = a_j$, 且对每个 $(x_1, ..., x_{n-k}) \in U$, 有

$$\begin{cases} F_1(x_1, ..., x_{n-k}, g_{n-k+1}(x_1, ..., x_{n-k}), ..., g_n(x_1, ..., x_{n-k})) = 0, \\ ...... \\ F_k(x_1, ..., x_{n-k}, g_{n-k+1}(x_1, ..., x_{n-k}), ..., g_n(x_1, ..., x_{n-k})) = 0. \end{cases}$$

例 2.8.3. 方程组

$$\begin{cases} x^2 + y^2 + z^2 - 1 = 0, \\ x^2 + 2y^2 - z^2 - 1 = 0. \end{cases}$$

确定的隐函数.

2.9 多元函数的极值

2.9.1 临界点

问题 2.9.1. 求多元函数 $f : K \to \mathbf{R}$ 的最大值与最小值.

先引入局部最大值与局部最小值的概念.

定义 2.9.2. 设 $\mathbf{x}_0$ 是 K 的内点, 称 $\mathbf{x}_0$ 是 f 的极大值点, 如果存在 $\mathbf{x}_0$ 的开邻域 $U \subset K$, 使得

$$f(\mathbf{x}_0) \ge f(\mathbf{x}), \quad \forall \mathbf{x} \in U.$$

类似的, 可定义 f 的极小值点. 将 f 的极大值点与极小值点统称为 f 的极值点.

命题 2.9.3 (极值点的必要条件, Fermart 引理)**.** 设 $\mathbf{x}_0$ 是 f 的极值点, 且 f 在 $\mathbf{x}_0$ 处沿方向 $\mathbf{v}$ 的方向导数 $\nabla_{\mathbf{v}} f(\mathbf{x}_0)$ 存在, 则有 $\nabla_{\mathbf{v}} f(\mathbf{x}_0) = 0$. 特别的, 如果 f 在 $\mathbf{x}_0$ 处各个偏导数存在, 则 f 在 $\mathbf{x}_0$ 处的各个偏导数都等于 0.

证明*:* 设 $f(\mathbf{x}_0)$ 是 f 在 $\mathbf{x}_0$ 的开邻域 U 中的最值点. 考虑直线道路 $p(t) = \mathbf{x}_0 + t\mathbf{v}$, 由映射 p 在 $t = 0$ 处连续, 存在 0 在 $\mathbf{R}$ 中的开邻域 W, 使得 $p(W) \subset U$. 这样, $t = 0$ 是一元函数 $f(p(t))$ 的极值点, 从而有 $\frac{d}{dt}|_{t=0} f(p(t)) = 0$, 此即 $\nabla_{\mathbf{v}} f(\mathbf{x}_0) = 0$.

□

定义 2.9.4. 设 $D \subset \mathbf{R}^n$, $\mathbf{x}$ 是 D 的内点. 称 $\mathbf{x}$ 是 $f : D \to \mathbf{R}$ 的临界点 (*critical point*), 如果

$$\frac{\partial f}{\partial x_i}|_{\mathbf{x}} = 0, \quad \forall 1 \le i \le n.$$

这样, 对于光滑函数, {极值点} ⊂ {临界点}, 由此得到求解最值问题的方法.

(1) 存在性. 如果 K 是有界闭集, 则连续函数 f 在 K 上能取到最大值. 设 $\mathbf{x}_0$ 是最大值点.

(2) 讨论最大值点 $\mathbf{x}_0$ 的候选点. 如果 $\mathbf{x}_0$ 是 K 的内点, 则它是 f 的极大值点, 由 Fermart 引理可得它是 f 的临界点, 将 f 的所有临界点所构成的集合记为 $\mathrm{Crit}(f)$; 如果 $\mathbf{x}_0$ 不是 K 的内点, 称之为 K 的边界点, 将 K 的所有边界点所构成的集合记为 ∂K.

结合上述讨论, 可得

$$\max_{\mathbf{x} \in K} f(\mathbf{x}) = \max\{f(\mathbf{x}) : \mathbf{x} \in \mathrm{Crit}(f) \cup \partial K\}.$$

例 2.9.5. 设 $x, y \ge 0$ 且 $x + y \le \pi$. 求 $\sin x \cdot \sin y \cdot \sin(\pi - x - y)$ 的最大值.

例 2.9.6 (最小二乘法 *the method of least squares*, Gauss)**.** 已知 (理论预言) y 按如下方式依赖于自变量 x,

$$y = f(x; \alpha_1, ..., \alpha_m),$$

其中 $\alpha_1, ..., \alpha_m$ 是未知的参数. 设有大量的实验数据 $(x_1, y_1), ..., (x_n, y_n)$. 如何选取参数 $\alpha_1, ..., \alpha_m$ 使得理论预言与实验结果拟合的最好?

最小二乘法考虑所有误差的平方和

$$E(\alpha_1, ..., \alpha_m) = \sum_{i=1}^{n} (f(x_i; \alpha_1, ..., \alpha_m) - y_i)^2.$$

计算 $E(\alpha_1, ..., \alpha_m)$ 的最小值, 可以认为最小值点给出最佳拟合方案.

我们按照如下步骤求最小值.

(1) 存在性. 利用最值定理或者不等式估计证明 E 能取到最小值, 设最小值点为 $(\alpha_1, ..., \alpha_m)$.

(2) 最小值点是临界点, 它满足临界点的方程

$$\begin{cases} \frac{\partial E}{\partial \alpha_1} = \sum_{i=1}^n 2(f(x_i;\alpha_1,...,\alpha_m) - y_i)\frac{\partial f}{\partial \alpha_1} = 0, \\ ... \\ \frac{\partial E}{\partial \alpha_m} = \sum_{i=1}^n 2(f(x_i;\alpha_1,...,\alpha_m) - y_i)\frac{\partial f}{\partial \alpha_m} = 0, \end{cases}$$

求解这个方程组, 找出所有临界点. 比较临界点处 $E(\alpha_1,...,\alpha_m)$ 的值, 它们的最小值就是 E 在 $\mathbf{R}^m$ 上的最小值.

例 2.9.7 (线性最小二乘法)**.** 理论预言 y 线性的依赖于自变量 x,

$$y = \alpha x + \beta.$$

已有实验数据 $(x_1,y_1),...,(x_n,y_n)$, 其中 $n > 1$ 且 $x_1,...,x_n$ 互不相同. 求最佳拟合.

解答. 定义函数 $E : \mathbf{R}^2 \to \mathbf{R}$ 为

$$E(\alpha,\beta) = \sum_{i=1}^n (y_i - \alpha x_i - \beta)^2,$$

我们来求 $\min E$. 它的临界点方程为

$$\begin{cases} \frac{\partial E}{\partial \alpha} = \sum_{i=1}^n 2(y_i - \alpha x_i - \beta)\cdot(-x_i) = 0, \\ \frac{\partial E}{\partial \beta} = \sum_{i=1}^n 2(y_i - \alpha x_i - \beta)\cdot(-1) = 0, \end{cases}$$

此即为

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix},$$

上述方程有唯一解 (α_0,β_0). 我们来证明 $\min E = E(\alpha_0,\beta_0)$. 为此, 利用带 Lagrange 余项的 Taylor 公式, 对任何 $(\alpha,\beta) \in \mathbf{R}^2$, 有

$$E(\alpha,\beta) = E(\alpha_0,\beta_0) + (\alpha-\alpha_0, \beta-\beta_0) \cdot \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \cdot \begin{pmatrix} \alpha-\alpha_0 \\ \beta-\beta_0 \end{pmatrix},$$

利用二次型的正定性, 可得 $E(\alpha,\beta) \geq E(\alpha_0,\beta_0)$. □

2.9.2 判断临界点是否为极值点

问题 2.9.8. 设 f 是 C^1 光滑函数, Fermart 引理说

$$\{f \text{ 的极值点}\} \subset \{f \text{ 的临界点}\}.$$

如何判断某个临界点是否为极值点呢?

利用展开至 2 阶的 Taylor 公式可得如下结果.

定理 2.9.9. 设 D 是 $\mathbf{R}^n$ 的子集, $f: D \to \mathbf{R}$ 是 C^2 光滑函数, $\mathbf{x}_0$ 是 D 的内点, 且是 f 的临界点. 令

$$H_f(\mathbf{x}_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}|_{\mathbf{x}_0} \right)_{1 \le i,j \le n},$$

称为 f 在 $\mathbf{x}_0$ 处的 Hessian*(海瑟矩阵)*.

(1) 如果 $H_f(\mathbf{x}_0)$ 正定, 则 $\mathbf{x}_0$ 是极小值点.

(2) 如果 $H_f(\mathbf{x}_0)$ 负定, 则 $\mathbf{x}_0$ 是极大值点.

(3) 如果 $H_f(\mathbf{x}_0)$ 不定, 则 $\mathbf{x}_0$ 不是极值点.

证明*:* (1) 考虑 Hessian 的顺序主子式

$$D_k(\mathbf{x}) = \det \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{1 \le i,j \le k}, \quad \forall 1 \le k \le n,$$

它们都是连续函数. 由假设 $H_f(\mathbf{x}_0)$ 正定, 可知对每个 $1 \le k \le n$ 有 $D(\mathbf{x}_0) > 0$. 利用函数 $D_k(1 \le k \le n)$ 的连续性, 存在 $\mathbf{x}_0$ 的开球邻域 U, 使得在 U 中每个 $D_k(1 \le k \le n)$ 都是恒正的, 从而对每个 $\mathbf{y} \in U$, Hessian $H_f(\mathbf{y})$ 都是正定矩阵.

由 $\mathbf{x}_0$ 是 f 的临界点, 利用 f 在 $\mathbf{x}_0$ 处的 Taylor 公式可知对每个 $\mathbf{x} \in U$, 存在 $\mathbf{y} \in U$ 使得

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \cdot H_f(\mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}_0).$$

由前述 $H_f(\mathbf{y})$ 是正定矩阵, 可得 $\mathbf{x}_0$ 是 f 的极小值点.

(3) 设 $H_f(\mathbf{x}_0)$ 不定, 则存在 $\mathbf{v}_-$ 与 $\mathbf{v}_+$, 使得

$$\mathbf{v}_-^T H_f(\mathbf{x}_0) \mathbf{v}_- < 0, \quad \mathbf{v}_+^T H_f(\mathbf{x}_0) \mathbf{v}_+ > 0.$$

注意到, 利用带 Peano 余项的 Taylor 公式, 有如下极限式

$$\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}_-) - f(\mathbf{x}_0)}{t^2} = \mathbf{v}_-^T H_f(\mathbf{x}_0) \mathbf{v}_- < 0,$$
$$\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}_+) - f(\mathbf{x}_0)}{t^2} = \mathbf{v}_+^T H_f(\mathbf{x}_0) \mathbf{v}_+ > 0,$$

由此可知 $\mathbf{x}_0$ 不是极值点. □

我们把上述结果用于 2 维的情形.

引理 2.9.10. 二次型 $Ax^2 + 2Bxy + Cy^2$ 的正定性如下表所示*:*

$AC-B^2>0, A>0$	正定
$AC-B^2>0, A<0$	负定
$AC-B^2<0$	不定
$AC-B^2=0$	半正定或半负定

推论 2.9.11. 设 f 是 C^2 光滑函数, (x_0, y_0) 是 f 的临界点. 记

$$A=f_{xx}(x_0,y_0),\quad B=f_{xy}(x_0,y_0),\quad C=f_{yy}(x_0,y_0).$$

(1) 如果 $A>0, AC-B^2>0$, 则 (x_0,y_0) 是极小值点.

(2) 如果 $A<0, AC-B^2>0$, 则 (x_0,y_0) 是极大值点.

(3) 如果 $AC-B^2<0$, 则 (x_0,y_0) 不是极值点.

2.10 条件最值问题

2.10.1 单个约束的最值问题

经常要在某种约束条件下求函数的最值.

问题 2.10.1. 设 (x,y) 满足约束条件 $g(x,y)=0$, 求 $\max f(x,y)$. 这里我们假设 f,g 都是 C^1 光滑的.

令 $S=\{(x,y)|g(x,y)=0\}$ 为方程 $g(x,y)=0$ 的解集, 要计算 f 在 S 上的最大值. 可按如下方法求解.

(1) 存在性. 如果能证明 S 是有界闭集, 则连续函数 f 在 S 上能取到最大值. 设 $P_0=(x_0,y_0)$ 是最大值点.

(2) 分析 P_0 可能的候选点, 分两种情况.

(2.1) 如果 $(g_x(x_0,y_0), g_y(x_0,y_0))\neq(0,0)$, 则由隐函数定理知在 P_0 附近 S 的形状与 $\mathbf{R}^1$ 一样. 人们称这种点 P_0 为 S 的光滑点.

定义 2.10.2. 设 $P_0=(x_0,y_0)$ 是 S 的光滑点. 如果存在 P_0 的邻域 U, 使得对任何 $P\in U\cap S$ 都有 $f(P)\leq f(P_0)$, 则称 P_0 为 f 的条件极大值点.

(2.2) 如果 $(g_x(x_0,y_0), g_y(x_0,y_0))=(0,0)$, 则称 P_0 为 S 的奇异点. 对于奇异点, 没有系统的办法, 只能具体分析. 好在奇异点很少 (甚至不存在).

总结一下, f 在 S 上的最大值点或者是奇异点, 或者是条件极大值点. 奇异点一般来说很少, 可以通过解方程组

$$g_x(x,y)=g_y(x,y)=0$$

一一枚举. 下面我们来确定条件极值点.

设 P_0 是 f 的条件极大值点, 不妨设 $g_y(x_0,y_0)\neq 0$. 由隐函数定理知在 x_0 附近 y 可以表示成 x 的隐函数 $y=y(x)$. 由假设知 x_0 是函数

$$x\to f(x,y(x))$$

的极大值点, 从而有

$$\frac{df(x,y(x))}{dx}|_{x_0}=0,$$

此即

$$f_x(x_0,y_0)+f_y(x_0,y_0)\cdot(-\frac{g_x(x_0,y_0)}{g_y(x_0,y_0)})=0.$$

令 $\lambda=\frac{f_y(x_0,y_0)}{g_y(x_0,y_0)}$, 则

$$\begin{cases} f_x(x_0,y_0)-\lambda g_x(x_0,y_0)=0,\\ f_y(x_0,y_0)-\lambda g_y(x_0,y_0)=0.\end{cases} \tag{2.4}$$

类似的, 如果 $g_x(x_0,y_0)\neq 0$, 同样可知存在 λ 使得方程组 (1) 成立. 这就得到条件极值点的一个必要条件.

命题 2.10.3. *如果 (x_0,y_0) 是函数 f 在约束 $g(x,y)=0$ 下的条件极大 (或极小) 值点, 则存在实数 λ 使得*

$$\begin{cases} f_x(x_0,y_0)-\lambda g_x(x_0,y_0)=0,\\ f_y(x_0,y_0)-\lambda g_y(x_0,y_0)=0,\\ g(x_0,y_0)=0.\end{cases} \tag{2.5}$$

通过求解方程组 (2), 可以找出所有潜在的条件极值点, 把它们处 f 的值与奇异点处 f 的值相比较, 所得的最大值就是 f 在约束条件 $g(x,y)=0$ 下的最大值.

如果引入函数 $F(x,y,\lambda)=f(x,y)-\lambda g(x,y)$, 则上述方程组 (2) 可以写成:

$$\begin{cases} \frac{\partial F}{\partial x}|_{(x_0,y_0)}=0,\\ \frac{\partial F}{\partial y}|_{(x_0,y_0)}=0,\\ \frac{\partial F}{\partial \lambda}|_{(x_0,y_0)}=0,\end{cases} \tag{2.6}$$

这恰好是 F 的临界点方程!

命题 2.10.4 (Lagrange 乘子法)**.** *如果 (x_0, y_0) 是函数 f 在约束 $g(x,y)=0$ 下的条件极大 (或极小) 值点, 则存在实数 λ 使得 (x_0, y_0, λ) 满足函数 $F(x,y,\lambda)=f(x,y)-\lambda g(x,y)$ 的临界点方程 (3).*

类似的, 考虑约束 $g(x,y,z)=0$ 下 $f(x,y,z)$ 的最值问题. 令 $Z=\{(x,y,z)|g(x,y,z)=0\}$ 为 g 的零点集. 如果 $(x_0,y_0,z_0)\in Z$ 处

$$(g_x(x_0,y_0,z_0), g_y(x_0,y_0,z_0), g_z(x_0,y_0,z_0)) \neq (0,0,0),$$

则称 (x_0,y_0,z_0) 是 Z 的光滑点, 否则称之为 Z 的奇异点.

如果 (x_0,y_0,z_0) 是 Z 的光滑点, 且存在开邻域 U 使得

$$f(x,y,z) \leq f(x_0,y_0,z_0), \quad \forall (x,y,z) \in U \cap Z,$$

则称 (x_0,y_0,z_0) 是 f 在约束 $g=0$ 下的条件极大值点. 下面我们来确定所有条件极大值点.

由对称性, 不妨假设 $g_z(x_0,y_0,z_0)\neq 0$. 由隐函数定理可从 $g(x,y,z)=0$ 解出隐函数 $z=z(x,y)$, 如果 (x_0,y_0,z_0) 是 f 的条件极值点, 则 (x_0,y_0) 是函数 $f(x,y,z(x,y))$ 的 (无条件) 极值点, 从而满足方程组

$$\begin{cases} f_x(x_0,y_0,z_0)+f_z(x_0,y_0,z_0)\cdot z_x(x_0,y_0)=0, \\ f_y(x_0,y_0,z_0)+f_z(x_0,y_0,z_0)\cdot z_y(x_0,y_0)=0, \end{cases}$$

将 $z_x=-\frac{g_x}{g_z}, z_y=-\frac{g_y}{g_z}$ 代入上述方程, 令 $\lambda=\frac{f_z(x_0,y_0,z_0)}{g_z(x_0,y_0,z_0)}$, 可得

$$\begin{cases} f_x(x_0,y_0,z_0)-\lambda g_x(x_0,y_0,z_0)=0 \\ f_y(x_0,y_0,z_0)-\lambda g_y(x_0,y_0,z_0)=0 \\ f_z(x_0,y_0,z_0)-\lambda g_z(x_0,y_0,z_0)=0 \\ g(x_0,y_0,z_0)=0 \end{cases} \tag{2.7}$$

这恰是函数 $F(x,y,z,\lambda)=f(x,y,z)-\lambda g(x,y,z)$ 的临界点方程.

命题 2.10.5 (Lagrange 乘子法)**.** *如果 (x_0,y_0,z_0) 是函数 f 在约束 $g(x,y,z)=0$ 下的条件极大 (或极小) 值点, 则存在 λ 使得 (x_0,y_0,z_0,λ) 满足函数 $F(x,y,z,\lambda)=f(x,y,z)-\lambda g(x,y,z)$ 的临界点方程 (4).*

例 2.10.6. 给定正整数 n 与正数 a. 设非负实数 $x_1,x_2,...,x_n$ 满足 $x_1+...+x_n=a$. 求 $x_1x_2...x_n$ 的最大值.

证明: 令 $f(x_1,x_2,...,x_n)=x_1x_2...x_n$, $g(x_1,x_2,...,x_n)=x_1+...+x_n-a$,

$$S=\{(x_1,...,x_n)|g(x_1,...,x_n)=0,x_i\geq 0,\forall i\}.$$

显然 S 是有界闭集, 由最值定理, 连续函数 f 在 S 上能取到最大值, 设 $(\overline{x}_1,...,\overline{x}_n)$ 是最大值点.

S 的所有内点构成的集合为 (称为 S 的内部)

$$\overset{\circ}{S}=\{(x_1,...,x_n)|\sum_{i=1}^{n}x_i=a,x_1>0,x_2>0,...,x_n>0\}.$$

注意到 f 在 $\partial S=S-\overset{\circ}{S}$ 上恒等于 0, 最大值点一定是 S 的内点, 即有 $(\overline{x}_1,...,\overline{x}_n)\in\overset{\circ}{S}$. 其次, $\frac{\partial g}{\partial x_n}=1\neq 0$, 由隐函数定理知 $(\overline{x}_1,...,\overline{x}_n)$ 是 S 的光滑点, 因此它是 f 的条件极值点. 利用 Lagrange 乘子法, 存在实数 $\overline{\lambda}$, 使得 $(\overline{x}_1,...,\overline{x}_n,\overline{\lambda})$ 满足函数 $F(x_1,...,x_n,\lambda)=f(x_1,...,x_n)-\lambda g(x_1,...,x_n)$ 的临界点方程:

$$\begin{cases}\frac{\partial F}{\partial x_1}=x_2...x_n-\lambda=0,\\ ...\\ \frac{\partial F}{\partial x_n}=x_1...x_{n-1}-\lambda=0,\\ \frac{\partial F}{\partial \lambda}=-(x_1+...+x_n-a)=0,\end{cases}$$

解得 $\overline{x}_1=...=\overline{x}_n=\frac{a}{n}$, 由此得 f 的最大值为 $(\frac{a}{n})^n$, 在点 $(\frac{a}{n},...,\frac{a}{n})$ 处取得. □

定义 2.10.7. 设 $f,g:\mathbf{R}^n\to\mathbf{R}$ 是 C^1 光滑函数, 令

$$Z=\{(x_1,...,x_n)|g(x_1,...,x_n)=0\}$$

为 g 的零点集合. 设 $\mathbf{a}=(a_1,...,a_n)\in Z$, 称 $\mathbf{a}$ 是 f 在约束 $g(x_1,...,x_n)=0$ 下的条件极大值点, 如果它满足如下两个条件:

(1) 存在 $\mathbf{a}$ 的开球邻域 U, 使得对任何 $\mathbf{x}\in U\cap Z$, 有 $f(\mathbf{a})\geq f(\mathbf{x})$.

(2) $\mathbf{a}$ 是 Z 的光滑点, 即 $(\frac{\partial g}{\partial x_1}|_{\mathbf{a}},...,\frac{\partial g}{\partial x_n}|_{\mathbf{a}})\neq(0,...,0)$.

定理 2.10.8. 设 $\mathbf{a}=(a_1,...,a_n)$ 是 f 在约束 $g(x_1,...,x_n)=0$ 下的条件极值点, 则存在 $\lambda\in\mathbf{R}$, 使得 $(a_1,...,a_n,\lambda)$ 是函数

$$F(x_1,...,x_n,\lambda)=f(x_1,...,x_n)-\lambda g(x_1,...,x_n)$$

的临界点.

例 2.10.9. 给定对称矩阵 $A=(A_{ij})_{1\le i,j\le n}$. 求函数

$$f(x_1,...,x_n)=\sum_{i,j}A_{ij}x_ix_j$$

在球面

$$S^{n-1}:=\{(x_1,...,x_n)|x_1^2+...+x_n^2=1\}$$

上的最大值.

这样, 我们就证明了如下结果:

定理 2.10.10. *实对称方阵至少有一个实特征根.*

2.10.2 多个约束的最值问题

更加一般的, 给定正整数 $m<n$, 设有 m 个约束

$$\begin{cases} g_1(x_1,...,x_n)=0, \\ g_2(x_1,...,x_n)=0, \\ ... \\ g_m(x_1,...,x_n)=0, \end{cases}$$

求 $f(x_1,...,x_n)$ 的最值.

令 $Z=\{(x_1,...,x_n)|g_1(x_1,...,x_n)=...=g_m(x_1,...,x_n)=0\}$ 为 $g_1,...,g_m$ 的公共零点集, 设 $(\overline{x}_1,...,\overline{x}_n)\in Z$. 如果雅可比矩阵

$$(\frac{\partial g_i}{\partial x_j}|_{(\overline{x}_1,...,\overline{x}_n)})_{1\le i\le m,1\le j\le n}$$

是满秩矩阵, 则称 $(\overline{x}_1,...,\overline{x}_n)$ 是 Z 的光滑点, 否则称之为 Z 的奇异点.

如果 $(\overline{x}_1,...,\overline{x}_n)$ 是 Z 的光滑点, 且存在开邻域 U 使得

$$f(x_1,...,x_n)\le f(\overline{x}_1,...,\overline{x}_n),\quad \forall(x_1,...,x_n)\in U\cap Z,$$

则称 $(\overline{x}_1,...,\overline{x}_n)$ 是 f 在约束 $g_1=...=g_m=0$ 下的条件极大值点. 下面我们来确定所有条件极大值点.

定理 2.10.11 (Lagrange 乘子法)**.** *设 $(\overline{x}_1,...,\overline{x}_n)$ 是 f 在约束 $g_1=...=g_m=0$ 下的条件极大值点, 则存在 $\overline{\lambda}_1,...,\overline{\lambda}_m$ 使得 $(\overline{x}_1,...,\overline{x}_n,\overline{\lambda}_1,...,\overline{\lambda}_m)$ 满足函数*

$$F(x_1,...,x_n,\lambda_1,...,\lambda_m):=f(x_1,...,x_n)-\lambda_1g_1(x_1,...,x_n)-...-\lambda_mg_m(x_1,...,x_n)$$

的临界点方程

$$\begin{cases} \frac{\partial F}{\partial x_1} = 0, \\ \dots \\ \frac{\partial F}{\partial x_n} = 0, \\ \frac{\partial F}{\partial \lambda_1} = 0, \\ \dots \\ \frac{\partial F}{\partial \lambda_m} = 0. \end{cases}$$

例 2.10.12. 在约束条件

$$\begin{cases} x + y + z = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

下, 求函数 $f(x, y, z) = x^3 + y^3 + z^3$ 的最大值与最小值.

2.11 微分学在几何中的应用

2.11.1 向量

定义 2.11.1. 设 $\mathbf{a} = (a_1, ..., a_n), \mathbf{b} = (b_1, ..., b_n) \in \mathbf{R}^n$, 定义它们的内积 (*inner product*) 为:

$$\mathbf{a} \cdot \mathbf{b} := \sum_{i=1}^{n} a_i b_i.$$

显然有 $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$.

命题 2.11.2. 设 $\mathbf{a} = (a_1, ..., a_n), \mathbf{b} = (b_1, ..., b_n)$, 点 $\mathbf{a}$ 到直线 $\mathbf{ob}$ 的距离为

$$d = \frac{\sqrt{|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2}}{|\mathbf{b}|}.$$

注 2.11.3. 由此可得 Cauchy-Schwartz 不等式 $|\mathbf{a} \cdot \mathbf{b}| \le |\mathbf{a}||\mathbf{b}|$.

命题 2.11.4. 由向量 $\mathbf{a} = (a_1, ..., a_n)$ 与 $\mathbf{b} = (b_1, ..., b_n)$ 张成的平行四边形的面积为

$$S = \sqrt{|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2}.$$

推论 2.11.5. 由平面向量 $\mathbf{v} = (v_1, v_2)$ 与 $\mathbf{w} = (w_1, w_2)$ 张成的平行四边形的面积为

$$S = |v_1 w_2 - v_2 w_1| = |det \begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix}|.$$

推论 2.11.6. 由空间向量 $\mathbf{a}=(a_1,a_2,a_3)$ 与 $\mathbf{b}=(b_1,b_2,b_3)$ 张成的平行四边形的面积为

$$S=\sqrt{(a_2b_3-a_3b_2)^2+(a_3b_1-a_1b_3)^2+(a_1b_2-a_2b_1)^2}.$$

定义 2.11.7. 定义向量 $\mathbf{a}=(a_1,a_2,a_3)$ 与 $\mathbf{b}=(b_1,b_2,b_3)$ 的叉积 (*cross product*) 为

$$\mathbf{a}\times\mathbf{b}=(a_2b_3-a_3b_2,a_3b_1-a_1b_3,a_1b_2-a_2b_1).$$

如果用 $\mathbf{i}=(1,0,0),\mathbf{j}=(0,1,0),\mathbf{k}=(0,0,1)$ 表示 $\mathbf{R}^3$ 的标准基底, 则叉乘可以表示成

$$\mathbf{a}\times\mathbf{b}=\begin{pmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{pmatrix}$$

由推论2.11.6可知, $\mathbf{a}\times\mathbf{b}$ 的长度等于 $\mathbf{a},\mathbf{b}$ 张成的平行四边形的面积

$$S=|\mathbf{a}\times\mathbf{b}|.$$

命题 2.11.8. 由空间向量 $\mathbf{a}=(a_1,a_2,a_3)$, $\mathbf{b}=(b_1,b_2,b_3)$ 与 $\mathbf{c}=(c_1,c_2,c_3)$ 张成的平行六面体的体积为

$$V=|(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|=|det\begin{pmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{pmatrix}|.$$

2.11.2 右手法则

把人们生活在其中的三维空间记作 V. 如果选定基底

$$\mathbf{e}_x,\quad \mathbf{e}_y,\quad \mathbf{e}_z\in V,$$

则可用如下方式将 V 等同成 $\mathbf{R}^3$:

$$\Phi:\mathbf{R}^3\to V,$$

$$\Phi(x,y,z)=x\mathbf{e}_x+y\mathbf{e}_y+z\mathbf{e}_z.$$

定义 2.11.9. 设 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是 V 中的向量, 称 $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成右手系, 如果当右手的拇指指向 $\mathbf{a}$, 食指指向 $\mathbf{b}$ 时, 中指恰好指向 $\mathbf{c}$.

类似的, 对于平面 P, 如果选定基底 $\mathbf{e}'_x,\mathbf{e}'_y$, 则可将 P 等同成 $\mathbf{R}^2$:

$$\Psi(x,y)=x\mathbf{e}'_x+y\mathbf{e}'_y.$$

以后我们都不加说明的用以上方式把 V 等同成 $\mathbf{R}^3$, 将 P 等同成 $\mathbf{R}^2$.

命题 2.11.10. 设 $\mathbf{e}_y'$ 是由 $\mathbf{e}_x'$ 逆时针旋转 $\frac{\pi}{2}$ 角得到, 则对 P 中的两个向量 $\mathbf{a}=(a_1,a_2),\mathbf{b}=(b_1,b_2)$, $\mathbf{b}$ 由 $\mathbf{a}$ 逆时针旋转某个 $\theta\in(0,\pi)$ 角得到的充分必要条件是

$$\det\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}>0.$$

证明: 设

$$\frac{\mathbf{a}}{|\mathbf{a}|}=(\cos\alpha,\sin\alpha),\quad \alpha\in[0,2\pi),$$

则有

$$\begin{aligned} &\mathbf{b}\text{由}\mathbf{a}\text{逆时针旋转某个}\theta\in(0,\pi)\text{角得到}\\ \Longleftrightarrow\ &(b_1+b_2i)e^{-\alpha i}\text{由}(a_1+a_2i)e^{-\alpha i}\text{逆时针旋转某个}\theta\in(0,\pi)\text{角得到}\\ \Longleftrightarrow\ &\mathrm{Im}((b_1+b_2i)e^{-\alpha i})>0\\ \Longleftrightarrow\ &\det\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}>0. \end{aligned}$$

□

注 2.11.11. 在上述证明过程中, 我们把 $\mathbf{R}^2$ 等同成复平面 $\mathbf{C}$,

$$\mathbf{R}^2\ni(x,y)\to x+iy\in\mathbf{C},$$

在 $\mathbf{C}$ 中, 把 $x+yi$ 逆时针旋转 θ 角所得的复数是 $(x+yi)e^{i\theta}$.

命题 2.11.12. 设 $\mathbf{e}_x,\mathbf{e}_y,\mathbf{e}_z$ 构成右手系, 则 V 中的三个向量

$$\mathbf{a}=(a_1,a_2,a_3),\quad \mathbf{b}=(b_1,b_2,b_3),\quad \mathbf{c}=(c_1,c_2,c_3),$$

构成右手系的充分必要条件是

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}>0.$$

也可以把上述充分必要条件叙述成

$$\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}>0.$$

证明: 因为要讨论 $\mathbf{a},\mathbf{b},\mathbf{c}$ 是否构成右手系, 不妨设它们线性无关 (即不共面).

(1) 如果 $\mathbf{a}, \mathbf{b}$ 都属于 Oxy 平面. 设观察者站在点 $(0,0,1)$ 处, 在他看来 $\mathbf{a}, \mathbf{b}$ 的位置关系有两种可能:

(1.1) $\mathbf{b}$ 由 $\mathbf{a}$ 逆时针旋转某个 $\theta \in (0, \pi)$ 角得到.

(1.2) $\mathbf{b}$ 由 $\mathbf{a}$ 顺时针旋转某个 $\theta \in (0, \pi)$ 角得到.

由命题2.11.2, 情形 (1.1) 的条件等价于

$$\det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} > 0.$$

在此条件下, 有:

$$\begin{aligned} \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ 构成右手系 } &\iff c_3 > 0 \\ &\iff \det \begin{pmatrix} a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \\ c_1 & c_2 & c_3 \end{pmatrix} > 0. \end{aligned}$$

类似的, 情形 (1.2) 的条件等价于

$$\det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} < 0.$$

在此条件下, 有:

$$\begin{aligned} \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ 构成右手系 } &\iff c_3 < 0 \\ &\iff \det \begin{pmatrix} a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \\ c_1 & c_2 & c_3 \end{pmatrix} > 0. \end{aligned}$$

(2) 如果 $\mathbf{a}, \mathbf{b}$ 处于一般位置, 设

$$\mathbf{n} = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \quad \theta \in [0, \pi), \phi \in [0, 2\pi).$$

考虑 $\mathbf{R}^3$ 的旋转:

$$R_3(\phi) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

$$R_2(\theta)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

令 $T^{-1} = R_3(\phi)R_2(-\theta)R_3(-\phi)$, 则

$$T^{-1}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{pmatrix},$$

此即 $T(\mathbf{n}) = (0,0,1)^t$. 注意到 T 保持垂直关系, 由 $\mathbf{a}, \mathbf{b} \perp \mathbf{n}$ 可得

$$T(\mathbf{a}), T(\mathbf{b}) \perp T(\mathbf{n}),$$

特别的,

$$T(\mathbf{a}), T(\mathbf{b}) \in 平面 Oxy.$$

这样, 利用 (1) 的结论, 有:

$$\begin{aligned} &\mathbf{a}, \mathbf{b}, \mathbf{c} \text{ 构成右手系} \\ \Longleftrightarrow\ & T\mathbf{a}, T\mathbf{b}, T\mathbf{c} \text{ 构成右手系} \\ \Longleftrightarrow\ & (T\mathbf{a} \times T\mathbf{b}) \cdot T\mathbf{c} > 0 \\ \Longleftrightarrow\ & (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} > 0 \end{aligned}$$

□

推论 2.11.13. 设向量 $\mathbf{a}, \mathbf{b} \in \mathbf{R}^3$ 线性无关, 则 $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ 构成右手系.

证明: 当 $\mathbf{a}, \mathbf{b}$ 线性无关时, $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$, 则

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) = |\mathbf{a} \times \mathbf{b}|^2 > 0.$$

□

2.11.3 曲线

定义 2.11.14. 称 $L \subset \mathbf{R}^3$ 为 $\mathbf{R}^3$ 中的 (C^1 光滑的) 曲线, 如果对每个点 $\mathbf{x}_0 \in L$, 都存在开邻域 U 与 C^1 光滑函数 $f, g : U \to \mathbf{R}$, 使得

(1) $L \cap U = \{(x,y,z) \in U | f(x,y,z) = g(x,y,z) = 0\}$.

(2) 矩阵

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{pmatrix}|_{\mathbf{x}_0}$$

是满秩的.

例 2.11.15. 设 $f, g : \mathbf{R}^3 \to \mathbf{R}$ 是 C^1 光滑函数, 考虑它们的公共零点集

$$L = \{(x,y,z) | f(x,y,z) = g(x,y,z) = 0\}.$$

如果对任何 $(x,y,z) \in L$, 矩阵

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{pmatrix}|_{(x,y,z)}$$

都是满秩的, 则 L 满足定义2.11.14中的条件, 是 $\mathbf{R}^3$ 中的曲线.

例 2.11.16. 由上例可知, 单位圆周

$$S^1 := \{(x,y,z) | x^2 + y^2 + z^2 - 1 = 0, z = 0\}$$

是 $\mathbf{R}^3$ 中的曲线.

定义 2.11.17. 设 L 是 $\mathbf{R}^3$ 中的曲线, $\mathbf{x}_0 \in L$. 称 $\mathbf{v} \in \mathbf{R}^3$ 是 L 在 $\mathbf{x}_0$ 处的一个切向量, 如果存在正数 ϵ 及 C^1 光滑映射 $p : (-\epsilon, \epsilon) \to L$, 使得

$$p(0) = \mathbf{x}_0, \quad p'(0) = \mathbf{v}.$$

设 L 在 $\mathbf{x}_0$ 处的所有切向量构成的集合为 $T_{\mathbf{x}_0}L$, 称之为 L 在 $\mathbf{x}_0$ 处的切空间.

例 2.11.18. 设 $L, \mathbf{x}_0, f, g$ 满足定义2.11.14中的条件, 我们来证明

$$T_{\mathbf{x}_0}L = \{\lambda(f_x, f_y, f_z)|_{\mathbf{x}_0} \times (g_x, g_y, g_z)|_{\mathbf{x}_0} : \lambda \in \mathbf{R}\}.$$

证明: 一方面, 如果 $\mathbf{v} \in T_{\mathbf{x}_0}L$, 则存在 C^1 光滑映射 $p : (-\epsilon, \epsilon) \to L$, 使得 $p(0) = \mathbf{x}_0$, $p'(0) = \mathbf{v}$. 设 $p(t) = (x(t), y(t), z(t))$, 则对任何 $t \in (-\epsilon, \epsilon)$, 有

$$\begin{cases} f(x(t), y(t), z(t)) = 0, \\ g(x(t), y(t), z(t)) = 0. \end{cases}$$

在 $t = 0$ 处求导, 可得

$$(f_x, f_y, f_z)|_{\mathbf{x}_0} \cdot \mathbf{v} = (g_x, g_y, g_z)|_{\mathbf{x}_0} \cdot \mathbf{v} = 0.$$

由于 $(f_x, f_y, f_z)|_{\mathbf{x}_0} \times (g_x, g_y, g_z)|_{\mathbf{x}_0} \neq (0,0,0)$, 则存在实数 λ 使得

$$\mathbf{v} = \lambda(f_x, f_y, f_z)|_{\mathbf{x}_0} \times (g_x, g_y, g_z)|_{\mathbf{x}_0}.$$

另一方面, 利用隐函数定理可以证明: 对任何实数 λ, 都有

$$\lambda(f_x, f_y, f_z)|_{\mathbf{x}_0} \times (g_x, g_y, g_z)|_{\mathbf{x}_0} \in T_{\mathbf{x}_0}L.$$

□

例 2.11.19. 由上例可知,

$$T_{(x,y,0)}S^1 = \{\lambda(2x, 2y, 2z) \times (0,0,1) | \lambda \in \mathbf{R}\} = \{\lambda(y, -x, 0) | \lambda \in \mathbf{R}\}.$$

例 2.11.20. 设曲线 L 由参数方程给出

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t), \end{cases}$$

则 L 在点 $(x(t_0), y(t_0), z(t_0))$ 处的切空间为

$$T_{(x(t_0),y(t_0),z(t_0))}L = \{\lambda(x'(t_0), y'(t_0), z'(t_0)) | \lambda \in \mathbf{R}\}.$$

定义 2.11.21. 设 L 是曲线, $\mathbf{x}_0 \in L$. 过点 $\mathbf{x}_0$ 做与 $T_{\mathbf{x}_0}L$ 平行的直线, 把所得的直线称之为 L 在 $\mathbf{x}_0$ 处的切线.

例 2.11.22. 设 $L, \mathbf{x}_0, f, g$ 满足定义2.11.14中的条件, 则 L 在 $\mathbf{x}_0$ 处的切线为

$$\mathbf{x}_0 + \lambda(f_x, f_y, f_z)|_{\mathbf{x}_0} \times (g_x, g_y, g_z)|_{\mathbf{x}_0}, \quad \lambda \in \mathbf{R}.$$

2.11.4 曲面

定义 2.11.23. 称 $S \subset \mathbf{R}^3$ 为 (C^1 光滑的) 曲面, 如果对每个点 $\mathbf{x}_0 \in S$, 存在开邻域 U 与 C^1 光滑函数 $f : U \to \mathbf{R}$, 使得

(1) $S \cap U = \{(x, y, z) \in U | f(x, y, z) = 0\}$.

(2) $(f_x, f_y, f_z)|_{\mathbf{x}_0} \neq (0,0,0)$.

例 2.11.24. 设 $f : \mathbf{R}^3 \to \mathbf{R}$ 是 C^1 光滑函数, 考虑它的零点集

$$S = \{(x, y, z) | f(x, y, z) = 0\}.$$

如果对任何 $(x, y, z) \in S$, 都有 $(f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)) \neq (0,0,0)$, 则 S 满足定义2.11.23中的条件, 是 C^1 光滑的曲面.

例 2.11.25. 由上例可知, 单位球面

$$S^2 := \{(x,y,z)|x^2+y^2+z^2-1=0\}$$

是 C^1 光滑的曲面.

例 2.11.26. 定义曲面

$$T^2 = \{(x,y,z)|(\sqrt{x^2+y^2}-1)^2+z^2=\frac{1}{4}\}.$$

容易验证它是 C^1 光滑的曲面. T^2 的形状与轮胎表面一样, 人们称之为二维环面 (torus).

定义 2.11.27. 设 S 是 C^1 光滑的曲面, $\mathbf{x}_0 \in S$. 称 $\mathbf{v} \in \mathbf{R}^3$ 是 S 在 $\mathbf{x}_0$ 处的一个切向量, 如果存在正数 ϵ 及 C^1 光滑映射 $p:(-\epsilon,\epsilon)\to S$, 使得

$$p(0)=\mathbf{x}_0, \quad p'(0)=\mathbf{v}.$$

设 S 在 $\mathbf{x}_0$ 处的所有切向量构成的集合为 $T_{\mathbf{x}_0}S$, 称之为 S 在 $\mathbf{x}_0$ 处的切空间.

定义 2.11.28. 称 $\mathbf{w}$ 是 S 在 $\mathbf{x}_0$ 处的一个法向量, 如果对任何 $\mathbf{v}\in T_{\mathbf{x}_0}S$, 有 $\mathbf{w}\cdot\mathbf{v}=0$.

例 2.11.29. 设 $S, \mathbf{x}_0, f$ 满足定义2.11.23中的条件, 我们来证明

$$T_{\mathbf{x}_0}S = \{\mathbf{v}|\mathbf{v}\cdot(f_x,f_y,f_z)|_{\mathbf{x}_0}=0\},$$

由此可知 S 在 $\mathbf{x}_0$ 处的法向量都形如 $\lambda(f_x,f_y,f_z)|_{\mathbf{x}_0}$, 其中 $\lambda\in\mathbf{R}$.

证明: 一方面, 如果 $\mathbf{v}\in T_{\mathbf{x}_0}S$, 则存在 C^1 光滑映射 $p:(-\epsilon,\epsilon)\to S$, 使得 $p(0)=\mathbf{x}_0$, $p'(0)=\mathbf{v}$. 设 $p(t)=(x(t),y(t),z(t))$, 则对任何 t, 有

$$f(x(t),y(t),z(t))=0,$$

在 $t=0$ 处求导, 可得

$$(f_x,f_y,f_z)|_{\mathbf{x}_0}\cdot\mathbf{v}=0. \tag{2.8}$$

另一方面, 利用隐函数定理可以证明: 只要 $\mathbf{v}$ 满足 (1) 式, 则 $\mathbf{v}\in T_{\mathbf{x}_0}S$. □

例 2.11.30. 由上例可知,

$$T_{(x,y,z)}S^2 = \{\mathbf{v}|\mathbf{v}\cdot(x,y,z)=0\}.$$

S^2 在 (x,y,z) 处的所有法向量为 $\{\lambda\cdot(x,y,z)|\lambda\in\mathbf{R}\}$.

例 2.11.31. 设曲面 S 由参数方程给出

$$\begin{cases} x = x(u,v) \\ y = y(u,v) \\ z = z(u,v), \end{cases}$$

则 S 在点 $\mathbf{p} = (x(u_0,v_0), y(u_0,v_0), z(u_0,v_0))$ 处的切空间为

$$T_{\mathbf{p}}S = \{\lambda(x_u, y_u, z_u)|_{(u_0,v_0)} + \mu(x_v, y_v, z_v)|_{(u_0,v_0)} : \lambda, \mu \in \mathbf{R}\},$$

S 在 $\mathbf{p}$ 处的所有法向量为

$$\nu(x_u, y_u, z_u)|_{(u_0,v_0)} \times (x_v, y_v, z_v)|_{(u_0,v_0)}, \quad \nu \in \mathbf{R}.$$

定义 2.11.32. 设 S 是曲面, $\mathbf{x}_0 \in S$. 过点 $\mathbf{x}_0$ 做与 $T_{\mathbf{x}_0}S$ 平行的平面, 把所得的平面称之为 S 在 $\mathbf{x}_0$ 处的切平面.

例 2.11.33. 设 $S, \mathbf{x}_0, f$ 满足定义2.11.23中的条件, 则 S 在 $\mathbf{x}_0$ 处的切平面为

$$(\mathbf{x} - \mathbf{x}_0) \cdot (f_x, f_y, f_z)|_{\mathbf{x}_0} = 0.$$

第三章　多元函数的积分

3.1　二重积分

3.1.1　定义

定义 3.1.1. 二重积分是 Riemann 和的极限

$$\iint_D f(x,y)dxdy := \lim \sum_i f(x_i,y_i)\cdot area(D_i),$$

也记作 $\iint_D f(x,y)d\sigma$, 称 $dxdy$ 与 $d\sigma$ 为面积微元.

几何含义. 可以把 $\iint_D f(x,y)dxdy$ 解读成:

(1) 函数图像下方柱体体积的代数和.

(2) D 的带权重 (权重为 f) 的面积.

例 3.1.2. $\iint_D dxdy$ 等于 D 的面积.

命题 3.1.3. *如果 f 在 D 上可积, 则 f 是 D 上的有界函数.*

命题 3.1.4. *对于“较好”的 D, 如果 $f\in C(D)$, 则 f 在 D 上可积.*

3.1.2　基本性质

(1) 关于被积函数是线性的.

$$\iint_D (af(x,y)+bg(x,y))dxdy = a\iint_D f(x,y)dxdy + b\iint_D g(x,y)dxdy.$$

(2) 关于积分区域是可加的. 如果 D_1, D_2 至多在边界上相交, 则

$$\iint_{D_1\cup D_2} f(x,y)dxdy = \iint_{D_1} f(x,y)dxdy + \iint_{D_2} f(x,y)dxdy.$$

(3) 绝对值不等式:

$$|\iint_D f(x,y)dxdy| \le \iint_D |f(x,y)|dxdy.$$

(4) 如果 $f(x,y) \ge 0, \forall (x,y) \in D$, 则 $\iint_D f(x,y)dxdy \ge 0$. 一般的, 如果 $f(x,y) \ge g(x,y), \forall (x,y) \in D$, 则

$$\iint_D f(x,y)dxdy \ge \iint_D g(x,y)dxdy.$$

(5) 积分中值定理. 设 $f \in C(D)$, 则存在 $(x_0,y_0) \in D$ 使得

$$\iint_D f(x,y)dxdy = f(x_0,y_0)\cdot \mathrm{area}(D).$$

(1) 设 $I = [a_1,a_2]\times[b_1,b_2] = \{(x,y)|a_1 \le x \le a_2, b_1 \le y \le b_2\}$ 是矩形, 定义 $f: I \to \mathbf{R}$ 在 I 上的二重积分为

$$\iint_I f(x,y)d\sigma = \lim_{\max\{\Delta x_i,\Delta y_j\}\to 0} \sum_{i,j} f(\xi_{ij})\Delta x_i \Delta y_j,$$

其中 $\{\xi_{ij} \in [x_{i-1},x_i]\times[y_{j-1},y_j]\}_{i,j}$ 是任意的选点方案.

(2) 设 $D \subset \mathbf{R}^2$ 是有界集合, $f: D \to \mathbf{R}$. 取矩形 I 包含 D, 令

$$f\cdot\chi_D(x,y) = f_D(x,y) = \begin{cases} f(x,y), & \text{如果}(x,y)\in D \\ 0, & \text{如果}(x,y)\notin D \end{cases}$$

定义 f 在 D 上的积分为

$$\iint_D f(x,y)d\sigma = \iint_I f_D(x,y)d\sigma.$$

定理 3.1.5. *(Riemann-Lebesgue) f 在 D 上可积的充分必要条件是: f 在 D 上有界, 且 f_D 的不连续点构成的集合是 $\mathbf{R}^2$ 的零测集.*

例 3.1.6. 设 $D \subset \mathbf{R}^2$ 是有界集合, 定义 D 的面积为

$$area(D) = \iint_D 1\cdot d\sigma.$$

注 3.1.7. 这样, 可以把 $\iint_D f(x,y)d\sigma$ 解释为 D 的权重为 f 的 “面积”.

命题 3.1.8. *设 D_1, D_2 都是有界集合, f 在 $D_1\cup D_2$ 上可积分, 则*

$$\iint_{D_1\cup D_2} f(x,y)d\sigma = \iint_{D_1} f(x,y)d\sigma + \iint_{D_2} f(x,y)d\sigma - \iint_{D_1\cap D_2} f(x,y)d\sigma.$$

特别的, 如果 $D_1\cap D_2$ 是有长度的曲线, 则有

$$\iint_{D_1\cup D_2} f(x,y)d\sigma = \iint_{D_1} f(x,y)d\sigma + \iint_{D_2} f(x,y)d\sigma.$$

3.1.3 计算方法: 化成累次积分

定理 3.1.9 (Fubini 定理)**.** 设 f 在 $I=[a_1,a_2]\times[b_1,b_2]$ 上可积, 则

$$\iint_I f(x,y)d\sigma=\int_{a_1}^{a_2}dx\int_{b_1}^{b_2}f(x,y)dy=\int_{b_1}^{b_2}dy\int_{a_1}^{a_2}f(x,y)dx.$$

推论 3.1.10. 设 $D=\{(x,y)|a\le x\le b,\phi_1(x)\le y\le\phi_2(x)\}$, 其中 ϕ_1,ϕ_2 是连续函数, 则对任何连续函数 $f:D\to\mathbf{R}$, 有

$$\iint_D f(x,y)d\sigma=\int_a^b dx\int_{\phi_1(x)}^{\phi_2(x)}f(x,y)dy.$$

类似的, 如果 $D'=\{(x,y)|c\le y\le d,\psi_1(y)\le x\le\psi_2(y)\}$, 其中 ψ_1,ψ_2 是连续函数, 则对任何连续函数 $f:D\to\mathbf{R}$, 有

$$\iint_{D'} f(x,y)d\sigma=\int_c^d dy\int_{\psi_1(y)}^{\psi_2(y)}f(x,y)dx.$$

推论 3.1.11. 如果 D 能表示成上述两种形式, 则对任何连续函数 f, 有

$$\int_a^b dx\int_{\phi_1(x)}^{\phi_2(x)}f(x,y)dy=\int_c^d dy\int_{\psi_1(y)}^{\psi_2(y)}f(x,y)dx,$$

即可以交换累次积分的次序

$$\int dx\int f(x,y)dy=\int dy\int f(x,y)dx.$$

例 3.1.12. 设 $D=\{(x,y)|a\le x\le b,\phi_1(x)\le y\le\phi_2(x)\}$, 则有

$$area(D)=\int_a^b(\phi_2(x)-\phi_1(x))dx.$$

例 3.1.13. 设积分区域是闭圆盘

$$\begin{aligned}D=&\{(x,y)|x^2+y^2\le r^2\}\\=&\{(x,y)|-r\le x\le r,-\sqrt{r^2-x^2}\le y\le\sqrt{r^2-x^2}\}\\=&\{(x,y)|-r\le y\le r,-\sqrt{r^2-y^2}\le x\le\sqrt{r^2-y^2}\},\end{aligned}$$

则对任何连续函数 $f:D\to\mathbf{R}$, 有

$$\iint_D f(x,y)dxdy=\int_{-r}^r dy\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}}f(x,y)dx=\int_{-r}^r dx\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}f(x,y)dy.$$

例 3.1.14. 设 $f:\mathbf{R}\to\mathbf{R}$ 是恒正的连续函数. 证明:

$$\iint_{x^2+y^2\leq R^2}\frac{f(x)}{f(x)+f(y)}dxdy=\iint_{x^2+y^2\leq R^2}\frac{f(y)}{f(x)+f(y)}dxdy=\frac{\pi R^2}{2}.$$

例 3.1.15. 定义 $\mathbf{R}^2$ 的子集 D 为

$$D=\{(x,y)|x,y\geq 0,x+y\leq 1\}.$$

对正数 a,b,c, 有

$$\iint_D\frac{1}{(ax+by+c(1-x-y))^3}dxdy=\frac{1}{2abc}.$$

例 3.1.16 (Schwinger trick and Feynman Parameters)**.** 设 V 是 $\mathbf{R}^n$ 的子集

$$V=\{(x_1,...,x_n)|x_1,...,x_n\geq 0,x_1+...+x_n\leq 1,\}$$

则对正数 $A_1,...,A_{n+1}$, 有

$$\int...\int_V\frac{1}{(A_1x_1+...+A_nx_n+A_{n+1}(1-x_1-...-x_n))^{n+1}}dx_1...dx_n=\frac{1}{n!\cdot A_1...A_{n+1}}.$$

"This is the way Schwinger presented the method of combining propagators. An interesting anecdote of physics history is that Schwinger remained bitter that a virtually identical mathematical trick became commonly known as Feynman parameters. Why two brilliant physicists, each of whom had an appropriately won a Nobel prize, should fight over what is essentially a trivial mathematical trick, is an interesting question in the sociology of physicists."

具体可参考`https://en.wikipedia.org/wiki/Feynman_parametrization`

3.1.4 换元公式

定理 3.1.17. 设 D' 到 D 的坐标变换公式

$$\phi:D'\to D,$$

$$(u,v)\to(x(u,v),y(u,v)),$$

是 C^1 光滑的, 且有 C^1 光滑的逆, 则对任何连续函数 $f:D\to\mathbf{R}$, 有如下换元公式:

$$\iint_D f(x,y)dxdy=\iint_{D'}f(x(u,v),y(u,v))\cdot|detJ(\phi)|\cdot dudv.$$

注 3.1.18. (1) 我们称 $\phi:D'\to D$ 为 C^1 微分同胚 (*diffeomorphism*), 如果它满足上述定理中的条件: ϕ 是 C^1 光滑的, 且有 C^1 光滑的逆 ϕ^{-1}.

(2) 换元公式中的因子 $|detJ(\phi)|$ 反映了坐标变换 ϕ 在每点附近面积扩大或缩小的倍数.

例 3.1.19 (椭圆的面积)**.**

$$\iint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}\leq 1} dxdy$$

证明*:* 有可逆的坐标变换:

$$\{(u,v): u^2+v^2\leq 1\} \xrightarrow{\phi} \{(x,y): \frac{x^2}{a^2}+\frac{y^2}{b^2}\leq 1\},$$

$$\phi(u,v)=(au,bv).$$

注意 $detJ(\phi)=ab$, 由换元法则

$$\iint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}\leq 1} dxdy = \iint_{u^2+v^2\leq 1} |detJ(\phi)|dudv = \pi ab.$$

□

例 3.1.20 (极坐标下的重积分)**.** 平面上的点可以用极坐标 (r,θ) 表示, 即存在光滑的坐标变换

$$\mathbf{R}^+\times[0,2\pi] \xrightarrow{\phi} \mathbf{R}^2-\{(0,0)\}$$

$$(r,\theta)\longrightarrow (x=r\cos\theta, y=r\sin\theta).$$

如果限制在 $\mathbf{R}^+\times[0,\pi]$ 上, 它有光滑的逆映射 $(x,y)\xrightarrow{\phi^{-1}}(r=\sqrt{x^2+y^2}, \theta=\arccos\frac{x}{\sqrt{x^2+y^2}})$

ϕ 的雅可比矩阵为

$$\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

雅可比行列式为 r. 由换元公式得

$$\iint_{D\cap\mathbf{R}^2_+} f(x,y)dxdy = \iint_{\phi^{-1}(D\cap\mathbf{R}^2_+)} f(r\cos\theta, r\sin\theta)rdrd\theta,$$

类似的, 下半平面中积分可化成

$$\iint_{D\cap\mathbf{R}^2_-} f(x,y)dxdy = \iint_{\phi^{-1}(D\cap\mathbf{R}^2_-)} f(r\cos\theta, r\sin\theta)rdrd\theta.$$

由于积分对区域的可加性, 得

$$\iint_D f(x,y)dxdy = \iint_{\phi^{-1}(D)} f(r\cos\theta, r\sin\theta)rdrd\theta,$$

例 3.1.21. 设 D 为圆盘 $x^2+y^2\leq a^2$. 计算 $\iint_D e^{-x^2-y^2}dxdy$.

证明:
$$\iint_D e^{-x^2-y^2}dxdy=\iint_{0\le r\le a,0\le\theta\le 2\pi}e^{-r^2}rdrd\theta=\int_0^a 2\pi e^{-r^2}rdr=\pi(1-e^{-a^2}).$$
□

上述结果对 $a\to\infty$ 取极限, 得
$$\lim_{a\to\infty}\iint_{x^2+y^2\le a^2}e^{-x^2-y^2}dxdy=\pi,$$
把左边的极限记作 $\iint_{\mathbf{R}^2}e^{-x^2-y^2}dxdy$, 则
$$\pi=\iint_{\mathbf{R}^2}e^{-x^2-y^2}dxdy=\int_{-\infty}^{+\infty}e^{-x^2}dx\int_{-\infty}^{+\infty}e^{-y^2}dy=(\int_{-\infty}^{+\infty}e^{-x^2}dx)^2,$$
从而
$$\int_{-\infty}^{+\infty}e^{-x^2}dx=\sqrt{\pi}.$$
这是著名的 Gauss 积分, 也写成
$$\int_0^{+\infty}e^{-x^2}dx=\frac{\sqrt{\pi}}{2},$$
它在概率论和物理中经常出现.

Gauss 积分的严格推导. 注意
$$\int_{-a}^{a}e^{-x^2}dx\int_{-a}^{a}e^{-y^2}dy=\iint_{[-a,a]\times[-a,a]}e^{-x^2-y^2}dxdy,$$
而
$$\iint_{x^2+y^2\le a^2}e^{-x^2-y^2}dxdy\le\iint_{[-a,a]\times[-a,a]}e^{-x^2-y^2}dxdy\le\iint_{x^2+y^2\le 2a^2}e^{-x^2-y^2}dxdy,$$
所以
$$\sqrt{\pi(1-e^{-a^2})}\le\int_{-a}^{a}e^{-x^2}dx\le\sqrt{\pi(1-e^{-2a^2})}.$$
利用夹逼定理知
$$\lim_{a\to\infty}\int_{-a}^{+a}e^{-x^2}dx=\sqrt{\pi}.$$

例 3.1.22. 设常数 a,b 不全为 0, $f:\mathbf{R}\to\mathbf{R}$ 是连续映射. 计算二重积分
$$\iint_{x^2+y^2\le R^2}f(ax+by)dxdy.$$
证明: 考虑如下坐标变换
$$\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}\frac{a}{\sqrt{a^2+b^2}}&\frac{b}{\sqrt{a^2+b^2}}\\\frac{-b}{\sqrt{a^2+b^2}}&\frac{a}{\sqrt{a^2+b^2}}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$
□

3.2 三重积分

完全类似的, 可以定义三重积分:

$$\iiint_{\Omega} f(x,y,z)dxdydz = \lim \sum f(x_i,y_i,z_i)\ \text{Volume}(\Omega_i).$$

例 3.2.1. $\iiint_{\Omega} dxdydz$ 等于 Ω 的体积, 由此可以把 $\iiint_{\Omega} f(x,y,z)dxdydz$ 看成 Ω 的带权重的体积.

3.2.1 化成累次积分

对于比较简单的积分区域, 重积分可以化成累次积分.

定理 3.2.2. *设积分区域为*

$$\Omega = \{(x,y,z)|(x,y)\in D, z_1(x,y)\le z\le z_2(x,y)\},$$

其中 z_1, z_2 是 D 上的连续函数, 则对任何连续函数 $f:\Omega\to\mathbf{R}$, 有:

$$\iiint_{\Omega} f(x,y,z)dxdydz = \iint_{D} dxdy\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)dz.$$

定理 3.2.3. *设积分区域 Ω 介于平面 $z=a$ 和平面 $z=b$ 之间, 并且对任何 $z_0\in[a,b]$, Ω 与平面 $z=z_0$ 的交集 D_{z_0} 是平面上的有界闭集, 则对任何连续函数 $f:\Omega\to\mathbf{R}$, 有*

$$\iiint_{\Omega} f(x,y,z)dxdydz = \int_a^b dz\iint_{D_z} f(x,y,z)dxdy.$$

例 3.2.4. n 维球体

$$B^n(r) := \{(x_1,...,x_n)|\sum_{i=1}^n x_i^2\le r^2\}$$

的体积.

换元公式可以推广到一般维数, 这里我们只叙述三维的情形.

定理 3.2.5. *设 $\phi:\Omega'\to\Omega$ 是 C^1 微分同胚*

$$\phi(u,v,w) = (x(u,v,w), y(u,v,w), z(u,v,w)).$$

对任何连续函数 $f:\Omega\to\mathbf{R}$, 有如下换元公式

$$\iiint_{\Omega} f(x,y,z)dxdydz = \iiint_{\Omega'} f(x(u,v,w),y(u,v,w),z(u,v,w))\cdot|detJ(\phi)|\cdot dudvdw.$$

其中坐标变换的雅可比矩阵是

$$J(\phi)=\begin{pmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{pmatrix},$$

换元公式中的因子 $|detJ(\phi)|$ 反映了坐标变换在每点附近体积扩大或缩小的倍数.

定理 3.2.6. 设 $V', V\subseteq \mathbf{R}^n$, $\phi: V'\to V$ 是 C^1 微分同胚, 则对任何连续函数 $f\in C(V)$, 有

$$\int\dots\int_V f dvol=\int\dots\int_{V'}(f\circ\phi)\cdot|\det J(\phi)|dvol.$$

例 3.2.7 (椭球的体积)**.**

$$\iiint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1} dxdydz$$

证明*:* 考虑如下微分同胚:

$$\{(u,v): u^2+v^2+w^2\le 1\}\xrightarrow{\phi}\{(x,y,z):\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1\},$$

$$\phi(u,v,w)=(au,bv,cw).$$

注意到 $detJ(\phi)=abc$, 由换元法则可得:

$$\iiint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1} dxdy=\iiint_{u^2+v^2+w^2\le 1}|detJ(\phi)|dudvdw=\frac{4}{3}\pi abc.$$

□

例 3.2.8 (柱坐标下的三重积分)**.** 柱坐标 (r,θ,z) 与直角坐标的变换关系为

$$\begin{cases} x=r\cos\theta \\ y=r\sin\theta \\ z=z \end{cases}$$

坐标变换的雅可比矩阵为

$$\begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

对应的雅可比行列式的值为 $detJ(\phi)=r$. 因此, 三重积分可用柱坐标表示成:

$$\iiint_\Omega f(x,y,z)dxdydz=\iiint f(r\cos\theta,r\sin\theta,z)rdrd\theta dz.$$

例 3.2.9 (球坐标下的三重积分)**.** 球坐标 $(r,\varphi,\theta)\in\mathbf{R}_{\geq 0}\times[0,2\pi]\times[0,\pi]$ 与直角坐标的变换关系为

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases}$$

坐标变换的雅可比矩阵为

$$\begin{pmatrix} \sin\theta\cos\varphi & -r\sin\theta\sin\varphi & r\cos\theta\cos\varphi \\ \sin\theta\sin\varphi & r\sin\theta\cos\varphi & r\cos\theta\sin\varphi \\ \cos\theta & 0 & -r\sin\theta \end{pmatrix}.$$

对应的雅可比行列式的值为 $detJ(\phi)=-r^2\sin\theta$. 因此, 三重积分用球坐标表示成:

$$\iiint_\Omega f(x,y,z)dxdydz=\iiint f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)r^2\sin\theta drd\varphi d\theta.$$

例 3.2.10. 设 $\Omega=\{(x,y,z)|\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\leq 1\}$. 计算

$$\iiint_\Omega(1-\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2})^{\frac{3}{2}}dxdydz.$$

完全类似的, 可以定义三重积分:

$$\iiint_\Omega f(x,y,z)dxdydz=\lim\sum f(x_i,y_i,z_i)\cdot\mathrm{volume}(\Omega_i).$$

例 3.2.11. $\iiint_\Omega dxdydz$ 等于 Ω 的体积, 由此可以把 $\iiint_\Omega f(x,y,z)dxdydz$ 看成 Ω 的带权重的体积.

对于比较简单的积分区域, 重积分可以化成累次积分.

定理 3.2.12. *设积分区域为*

$$\Omega=\{(x,y,z)|(x,y)\in D, z_1(x,y)\leq z\leq z_2(x,y)\},$$

其中 z_1,z_2 是 D 上的连续函数, 则对任何连续函数 $f:\Omega\to\mathbf{R}$, 有:

$$\iiint_\Omega f(x,y,z)dxdydz=\iint_D dxdy\int_{z_1(x,y)}^{z_2(x,y)}f(x,y,z)dz.$$

定理 3.2.13. *设积分区域 Ω 介于平面 $z=a$ 和平面 $z=b$ 之间, 并且对任何 $z_0\in[a,b]$, Ω 与平面 $z=z_0$ 的交集 D_{z_0} 是平面上的有界闭集, 则对任何连续函数 $f:\Omega\to\mathbf{R}$, 有*

$$\iiint_\Omega f(x,y,z)dxdydz=\int_a^b dz\iint_{D_z}f(x,y,z)dxdy.$$

例 3.2.14. 设 $\Omega = \{(x,y,z)|x^2+y^2 \le 1, x^2+z^2 \le 1, y^2+z^2 \le 1\}$. 求 Ω 的体积.

例 3.2.15. n 维球体

$$B^n(r) := \{(x_1,...,x_n)|\sum_{i=1}^n x_i^2 \le r^2\}$$

的体积.

例 3.2.16. 设 $f \in C^{(n)}(\mathbf{R},\mathbf{R})$, 计算

$$\int ... \int_{x_1+...+x_n \le 1, x_i \ge 0} f^{(n)}\left(A_1x_1 + ... + A_nx_n + A_0(1-x_1-...-x_n)\right) dx_1...dx_n.$$

例 3.2.17. 给定正数 $A_0,...,A_n$, 计算

$$\int ... \int_{x_1+...+x_n \le 1, x_i \ge 0} \left(A_1x_1 + ... + A_nx_n + A_0(1-x_1-...-x_n)\right)^{-n-1} dx_1...dx_n.$$

换元公式可以推广到一般维数.

定理 3.2.18. 设 $\phi : \Omega' \to \Omega$ 是 C^1 微分同胚

$$\phi(u,v,w) = (x(u,v,w), y(u,v,w), z(u,v,w)).$$

对任何连续函数 $f : \Omega \to \mathbf{R}$, 有如下换元公式

$$\iiint_\Omega f(x,y,z)dxdydz = \iiint_{\Omega'} f(x(u,v,w), y(u,v,w), z(u,v,w)) \cdot |detJ(\phi)| \cdot dudvdw.$$

其中坐标变换的雅可比矩阵是

$$J(\phi) = \begin{pmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{pmatrix},$$

换元公式中的因子 $|detJ(\phi)|$ 反映了坐标变换在每点附近引起体积扩大或缩小的倍数.

定理 3.2.19. 设 $V', V \subseteq \mathbf{R}^n$, $\phi : V' \to V$ 是 C^1 微分同胚, 则对任何连续函数 $f \in C(V)$, 有

$$\int ... \int_V f dvol = \int ... \int_{V'} (f \circ \phi) \cdot |\det J(\phi)| \cdot dvol.$$

例 3.2.20 (对称性)**.** 设 $\Omega \subset \mathbf{R}^3$, 令

$$\Omega' = \{(x,y,z)|(x,y,-z) \in \Omega\},$$

则对 Ω 上的可积函数 f, 有

$$\iiint_{\Omega} f(x,y,z)dxdydz = \iiint_{\Omega'} f(x,y,-z)dxdydz.$$

特别的, 如果 Ω 关于 z 方向的反射是对称的 (即 $\Omega = \Omega'$), f 关于 z 方向的反射是奇函数 (即 $f(x,y,-z) = -f(x,y,z)$), 则有 $\iiint_{\Omega} f(x,y,z)dxdydz = 0$.

例 3.2.21. 设 $V = \{(x_1,...,x_n)|0 \le x_1 \le x_2 \le ... \le x_n \le 1\}$, 求 V 的 n 维体积, 即计算积分

$$\int ... \int_V dx_1...dx_n.$$

例 3.2.22. 计算椭球的体积

$$\iiint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1} dxdydz.$$

证明*:* 考虑如下微分同胚:

$$\{(u,v): u^2+v^2+w^2 \le 1\} \xrightarrow{\phi} \{(x,y,z): \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \le 1\},$$

$$\phi(u,v,w) = (au,bv,cw).$$

注意到 $detJ(\phi) = abc$, 由换元法则可得:

$$\iiint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1} dxdy = \iiint_{u^2+v^2+w^2\le 1} |detJ(\phi)|dudvdw = \frac{4}{3}\pi abc.$$

□

例 3.2.23 (柱坐标下的三重积分)**.** 柱坐标 (r,θ,z) 与直角坐标的变换关系为

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{cases}$$

坐标变换的雅可比矩阵为

$$\begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

对应的雅可比行列式的值为 $detJ(\phi) = r$. 因此, 三重积分可用柱坐标表示成

$$\iiint_{\Omega} f(x,y,z)dxdydz = \iiint f(r\cos\theta, r\sin\theta, z)rdrd\theta dz.$$

例 3.2.24 (球坐标下的三重积分)**.** 球坐标 $(r,\varphi,\theta)\in\mathbf{R}_{\geq 0}\times[0,2\pi]\times[0,\pi]$ 与直角坐标的变换关系为

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases}$$

坐标变换的雅可比矩阵为

$$\begin{pmatrix} \sin\theta\cos\varphi & -r\sin\theta\sin\varphi & r\cos\theta\cos\varphi \\ \sin\theta\sin\varphi & r\sin\theta\cos\varphi & r\cos\theta\sin\varphi \\ \cos\theta & 0 & -r\sin\theta \end{pmatrix}.$$

对应的雅可比行列式的值为 $detJ(\phi)=-r^2\sin\theta$. 因此, 三重积分用球坐标表示成

$$\iiint_\Omega f(x,y,z)dxdydz = \iiint f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)r^2\sin\theta drd\varphi d\theta.$$

例 3.2.25. 设 $\Omega=\{(x,y,z)|\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\leq 1\}$. 计算

$$\iiint_\Omega (1-\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2})^{\frac{3}{2}}dxdydz.$$

3.3 第一型曲线积分

问题 3.3.1. *设曲线 $L\subset\mathbf{R}^3$ 上每点 (x,y,z) 处的 (线) 密度为 $\rho(x,y,z)$. 如何计算 L 的质量?*

先把 L 剖分 $L=\bigcup_{i=1}^n L_i$, 由此得到质量的近似表达式:

$$\sum_{i=1}^n \rho(x_i,y_i,z_i)\cdot length(L_i),$$

当剖分得越来越细时, 考虑上述 Riemann 和的极限, 可以认为该极限就是 L 的质量.

定义 3.3.2 (第一型的曲线积分)**.** 设 L 是三维空间中的分段光滑曲线, $f: L\to\mathbf{R}$ 是 L 上的函数. 如果不论怎样剖分以及选点 $(x_i,y_i,z_i)\in L_i$, 只要剖分得越来越细, Riemann 和

$$\sum_{i=1}^n f(x_i,y_i,z_i)\cdot length(L_i),$$

总趋近于同一个极限, 则称这个极限为 f 在 L 上的第一型曲线积分, 记做

$$\int_L f(x,y,z)dl,$$

称 f 为被积函数, L 为积分曲线, dl 为弧长微元.

第一型曲线积分有如下性质.

(1) 关于被积函数是线性的,

$$\int_L (\alpha f(x,y,z) + \beta g(x,y,z))dl = \alpha \int_L f(x,y,z)dl + \beta \int_L g(x,y,z)dl.$$

(2) 关于积分曲线是可加的. 如果 L_1, L_2 至多在端点相交, 则有

$$\int_{L_1 \cup L_2} f(x,y,z)dl = \int_{L_1} f(x,y,z)dl + \int_{L_2} f(x,y,z)dl.$$

3.3.1 计算方法

命题 3.3.3. *设 L 由参数方程*

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$$

给出, 其中 $t \in [a,b]$(这里要求 $a < b$) 且 $x(t), y(t), z(t)$ 都是 C^1 光滑映射, 则对任何连续函数 f, 有

$$\int_L f(x,y,z)dl = \int_a^b f(x(t),y(t),z(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}dt.$$

例 3.3.4. 设 L 是平面曲线, 由参数方程

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

给出, 其中 $t \in [a,b]$, $x(t), y(t)$ 是 C^1 光滑映射.

可以把 L 视为空间曲线, 由参数方程

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) = 0 \end{cases}$$

给出, 利用上述命题可得

$$\int_L f(x,y)dl = \int_a^b f(x(t),y(t))\sqrt{x'(t)^2 + y'(t)^2}dt.$$

例 3.3.5. 设平面曲线 L 是函数 g 的图像

$$L = \{(x, y) | y = g(x), a \le x \le b\}.$$

则 L 可由参数方程

$$\begin{cases} x = x \\ y = g(x) \end{cases}$$

给出, 其中参数 $x \in [a, b]$. 利用上例的结果, 可得

$$\int_L f(x, y) dl = \int_a^b f(x, g(x)) \sqrt{1 + g'(x)^2} dx.$$

例 3.3.6. 设 L 由参数方程

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$$

给出, 其中 $t \in [a, b]$ 且 $x(t), y(t), z(t)$ 都是 C^1 光滑映射, 则 L 的长度为

$$\text{length}(L) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt.$$

例 3.3.7. 设

$$L = \{(x, y, z) \in \mathbf{R}^3 | x^2 + y^2 + z^2 = 1, x + y + z = 1\}.$$

(1) 求 L 的一个参数方程表示.

(2) 计算 $\int_L f(x, y, z) dl$.

解. 引入坐标变化

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

□

例 3.3.8 (两点之间直线段最短)**.** 设 $x(t), y(t), z(t)$ 都是 C^1 光滑映射, 则有

$$\int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt \ge \sqrt{(x(b) - x(a))^2 + (y(b) - y(a))^2 + (z(b) - z(a))^2}.$$

证明: 由 Cauchy-Schwartz 不等式, 有

$$\begin{aligned}&x'(t)\cdot(x(b)-x(a))+y'(t)\cdot(y(b)-y(a))+z'(t)\cdot(z(b)-z(a))\\ \leq&\sqrt{x'(t)^2+y'(t)^2+z'(t)^2}\cdot\sqrt{(x(b)-x(a))^2+(y(b)-y(a))^2+(z(b)-z(a))^2},\end{aligned}$$

对 t 积分, 即得要证的结论. □

例 3.3.9. 考虑单位开圆盘 $D=\{(x,y)|x^2+y^2<1\}$, 定义 D 中 C^1 曲线 $p:[0,1]\to D$ 的长度为

$$L(p)=\int_0^1\frac{2}{1-(x(t)^2+y(t)^2)}\sqrt{x'(t)^2+y'(t)^2}dt,$$

其中 $p(t)=(x(t),y(t))$. 用这种方式定义了曲线长度后, 人们把 D 称之为 "双曲圆盘".

假设 $p(0)=(0,0)$, $|p(1)|=r$. 证明: $L(p)\geq\ln\frac{1+r}{1-r}$.

证明: 由 Cauchy-Schwartz 不等式, 有

$$\sqrt{x'(t)^2+y'(t)^2}\sqrt{x(t)^2+y(t)^2}\geq x(t)x'(t)+y(t)y'(t),$$

由此可得

$$\begin{aligned}L(p)&=\int_0^1\frac{2}{1-(x(t)^2+y(t)^2)}\sqrt{x'(t)^2+y'(t)^2}dt\\&\geq\int_0^1\frac{2}{1-(x(t)^2+y(t)^2)}\cdot\frac{x(t)x'(t)+y(t)y'(t)}{\sqrt{x(t)^2+y(t)^2}}dt\end{aligned}$$

□

3.4 第一型曲面积分

问题 3.4.1. *设曲面 S 上每点 (x,y,z) 处的 (面积) 密度为 $\rho(x,y,z)$. 如何计算 S 的质量?*

先把 S 剖分 $S=\cup_{i=1}^n S_i$, 由此得到质量的近似表达式:

$$\sum_{i=1}^n\rho(x_i,y_i,z_i)\cdot area(S_i),$$

当剖分得越来越细时, 考虑上述 Riemann 和的极限, 可以认为其极限就是 S 的质量.

定义 3.4.2 (第一型的曲面积分)**.** 设 S 是三维空间中的分块光滑曲面, $f:S\to\mathbf{R}$ 是 S 上的函数. 如果不论怎样剖分以及选点 $(x_i,y_i,z_i)\in S_i$, 只要剖分得越来越细, Riemann 和

$$\sum_{i=1}^n f(x_i,y_i,z_i)\cdot area(S_i),$$

总趋近于同一个极限, 则称这个极限为 f 在 S 上的第一型曲面积分, 记做

$$\iint_S f(x,y,z)dS.$$

其中 f 称为被积函数, S 称为积分曲面, dS 称为面积微元.

第一型曲面积分有如下性质.

(1) 关于被积函数是线性的,

$$\iint_S (\alpha f(x,y,z)+\beta g(x,y,z))dS=\alpha\iint_S f(x,y,z)dS+\beta\iint_S g(x,y,z)dS.$$

(2) 关于积分曲面是可加的. 如果 $S_1\cap S_2$ 是有限长的曲线, 则有

$$\iint_{S_1\cup S_2} f(x,y,z)dS=\iint_{S_1} f(x,y,z)dS+\iint_{S_2} f(x,y,z)dS.$$

3.4.1 计算方法

命题 3.4.3. 设曲面 S 由参数方程

$$\begin{cases} x=x(u,v)\\ y=y(u,v)\\ z=z(u,v)\end{cases}$$

给出, 其中 $(u,v)\in D\subseteq \mathbf{R}^2$, $x(u,v),y(u,v),z(u,v)$ 都是 C^1 光滑映射, 则对任何连续函数 f, 有

$$\iint_S f(x,y,z)dS=\iint_D f(x(u,v),y(u,v),z(u,v))\cdot|(x_u,y_u,z_u)\times(x_v,y_v,z_v)|\cdot dudv.$$

推论 3.4.4. 设 $g:D\to\mathbf{R}$ 是 C^1 光滑函数, S 是 g 的图像

$$S=\{(x,y,z)|(x,y)\in D, z=g(x,y)\},$$

则对任何连续函数 f, 有

$$\iint_S f(x,y,z)dS=\iint_D f(x,y,g(x,y))\sqrt{g_x^2+g_y^2+1}dxdy.$$

推论 3.4.5. 设曲面 S 由参数方程

$$\begin{cases} x=x(u,v)\\ y=y(u,v)\\ z=z(u,v)\end{cases}$$

给出, 其中 $(u,v)\in D\subseteq\mathbf{R}^2$, $x(u,v),y(u,v),z(u,v)$ 都是 C^1 光滑映射, 则 S 的面积为

$$area(S)=\iint_D |(x_u,y_u,z_u)\times(x_v,y_v,z_v)|\cdot dudv.$$

推论 3.4.6. 设 $g:D\to\mathbf{R}$ 是 C^1 光滑函数, S 是 g 的图像

$$S=\{(x,y,z)|(x,y)\in D,z=g(x,y)\},$$

则 S 的面积为

$$area(S)=\iint_D \sqrt{g_x^2+g_y^2+1}dxdy.$$

例 3.4.7. 求曲面积分

$$\iint_S x^2y^2dS,$$

其中 S 为上半球面: $z=\sqrt{R^2-x^2-y^2}$, $x^2+y^2\le R^2$.

例 3.4.8. 求曲面积分

$$\iint_S (x^2+y^2+z^2)dS,$$

其中 S 为圆柱面: $x^2+y^2=R^2$, $0\le z\le H$.

例 3.4.9. 考虑 (质量) 面密度为 ρ 的均匀球面 $x^2+y^2+z^2=R^2$. 计算它产生的引力场在点 $(0,0,l)$ 处的场强, 其中 $l\neq R$.

解. 引力场在点 $(0,0,l)$ 处的场强为

$$\mathbf{F}=\iint_S \frac{(x,y,z-l)}{(x^2+y^2+(z-l)^2)^{3/2}}\rho dS,$$

由对称性, 有

$$\iint_S \frac{x}{(x^2+y^2+(z-l)^2)^{3/2}}dS=\iint_S \frac{y}{(x^2+y^2+(z-l)^2)^{3/2}}dS=0.$$

下面来计算 $\iint_S \frac{z-l}{(x^2+y^2+(z-l)^2)^{3/2}}dS$. 利用球坐标下球面的参数表示, 可得

$$\begin{aligned}&\iint_S \frac{z-l}{(x^2+y^2+(z-l)^2)^{3/2}}dS\\ =&R^2\int_0^{2\pi}d\phi\int_0^{\pi}\frac{R\cos\theta-l}{(R^2-2Rl\cos\theta+l^2)^{3/2}}\sin\theta d\theta\\ =&2\pi R^2\int_0^{\pi}\frac{R\cos\theta-l}{(R^2-2Rl\cos\theta+l^2)^{3/2}}\sin\theta d\theta,\end{aligned}$$

这样, 计算归结于如下例子中的 Riemann 积分. □

例 3.4.10. 给定正数 $a \neq 1$, 计算定积分

$$\int_0^\pi \frac{(\cos x - a)\sin x}{(1+a^2-2a\cos x)^{3/2}}dx.$$

解答. 令 $u=\cos x$ 换元, 可得:

$$\begin{aligned}
&\int_0^\pi \frac{(\cos x - a)\sin x}{(1+a^2-2a\cos x)^{3/2}}dx\\
=&\int_0^\pi \frac{(\cos x - a)\cdot(-1)\cdot d(\cos x)}{(1+a^2-2a\cos x)^{3/2}}\\
=&\int_1^{-1} \frac{(u-a)\cdot(-1)\cdot du}{(1+a^2-2au)^{3/2}}\\
=&\int_{-1}^{1} \frac{(u-a)du}{(1+a^2-2au)^{3/2}}\\
=&\int_{-1}^{1} \frac{u-a}{a}\cdot((1+a^2-2au)^{-1/2})'du\\
=&\frac{u-a}{a}(1+a^2-2au)^{-1/2}|_{-1}^{1} - \int_{-1}^{1}\frac{1}{a}(1+a^2-2au)^{-1/2}du\\
=&\frac{u-a}{a}(1+a^2-2au)^{-1/2}|_{-1}^{1} + \frac{1}{a^2}(1+a^2-2au)^{1/2}|_{-1}^{1}\\
=&\frac{1-a}{a}\frac{1}{|1-a|}+\frac{1}{a^2}|1-a| - \frac{-1-a}{a}\frac{1}{1+a} - \frac{1}{a^2}(1+a)\\
=&\frac{1-a}{a^2|1-a|}-\frac{1}{a^2}.
\end{aligned}$$

所以有

$$\int_0^\pi \frac{(\cos x - a)\sin x}{(1+a^2-2a\cos x)^{3/2}}dx = \begin{cases} 0 & \text{如果} 0<a<1,\\ -\frac{2}{a^2} & \text{如果} a>1.\end{cases}$$

□

3.5 第二型曲线积分

Newton 力学方程为

$$\begin{cases} mx''(t)=F_1,\\ my''(t)=F_2,\\ mz''(t)=F_3,\end{cases}$$

由此可得

$$m(x''(t),y''(t),z''(t))\cdot(x'(t),y'(t),z'(t))=(F_1,F_2,F_3)\cdot(x'(t),y'(t),z'(t)),$$

其中 "·" 表示内积. 对上式积分得

$$\frac{1}{2}m[x'(t)^2 + y'(t)^2 + z'(t)^2]|_{t_0}^{t_1} = \int_{t_0}^{t_1}(F_1, F_2, F_3)\cdot(x'(t), y'(t), z'(t))dt.$$

在物理学中我们把上式解读成:

$$\text{动能的改变} = \text{外力做的功}.$$

外力 $\mathbf{F}$ 是矢量值的函数 $\mathbf{F}(x,y,z) = (F_1(x,y,z), F_2(x,y,z), F_3(x,y,z))$. 定义它对运动质点从 t_0 到 t_1 时间段中做的功为

$$W = \int_{t_0}^{t_1}(F_1, F_2, F_3)\cdot(x'(t), y'(t), z'(t))dt,$$

其中 $(x'(t), y'(t), z'(t))dt$ 是质点的 "位移微元"$d\mathbf{r} = (x'(t), y'(t), z'(t))dt$. 因此

$$\text{功} = \int \text{力}\cdot\text{位移微元}.$$

一般的, 我们引入以下定义.

定义 3.5.1. 设曲线 γ 为

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$$

其中 $t \in [a,b]$ 且 $x(t), y(t), z(t)$ 都是 (分段)C^1 光滑映射. 设

$$\mathbf{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$$

是连续的矢量值函数, 称积分

$$\int_{t_0}^{t_1}[P(x(t),y(t),z(t))x'(t) + Q(x(t),y(t),z(t))y'(t) + R(x(t),y(y),z(t))z'(t)]dt$$

为 $\mathbf{F}$ 在 γ 上的第二型曲线积分, 记做

$$\int_\gamma Pdx + Qdy + Rdz \text{ 或者} \int_\gamma \mathbf{F}(x,y,z)\cdot d\mathbf{r},$$

其中曲线 γ 称为积分路径.

3.5.1 参数曲线上的第二型曲线积分

定义 3.5.2. 设参数曲线 $\gamma : [a, b] \to \mathbf{R}^3$ 为

$$\begin{cases} x = x(t), \\ y = y(t), \\ z = z(t). \end{cases}$$

对于矢量值函数 $\mathbf{F}(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z))$, 称积分

$$\int_a^b [P(x(t), y(t), z(t))x'(t) + Q(x(t), y(t), z(t))y'(t) + R(x(t), y(y), z(t))z'(t)]dt$$

为 $\mathbf{F}$ 在 γ 上的第二型曲线积分, 记作

$$\int_\gamma Pdx + Qdy + Rdz \text{ 或} \int_\gamma \mathbf{F}(x, y, z) \cdot d\mathbf{r},$$

称 γ 称为积分路径.

例 3.5.3. 如果 γ 是平面曲线

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

$\mathbf{F}(x, y) = (P(x, y), Q(x, y))$ 是平面上的矢量值函数, 则类似的可定义 $\mathbf{F}$ 在 γ 上的第二型曲线积分:

$$\int_\gamma Pdx + Qdy := \int_a^b (P(x(t), y(t))x'(t) + Q(x(t), y(t))y'(t))dt.$$

如果令 $z(t) \equiv 0$, 可把 γ 视空间曲线, 则上述定义式是定义3.5.2的特例.

例 3.5.4 (第二型曲线积分可化成第一型曲线积分)**.** 设 $\gamma : [a, b] \to \mathbf{R}^3$ 是参数曲线, 它的像集 $\mathrm{Im}(\gamma) = \gamma([a, b])$ 是 $\mathbf{R}^3$ 中的 “几何曲线”. 注意到

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t)dt = \int_a^b \mathbf{F}(\gamma(t)) \cdot \frac{\gamma'(t)}{|\gamma'(t)|}|\gamma'(t)|dt,$$

令 $\mathbf{e}(\gamma(t)) = \frac{\gamma'(t)}{|\gamma'(t)|}$, 它是 γ 上的矢量值函数 (单位切向量场). 上式表明

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_\gamma (\mathbf{F} \cdot \mathbf{e})dl,$$

这就把第二型曲线积分化成了第一型曲线积分.

3.5.2 曲线的定向

定义 3.5.5 (曲线的定向). 曲线 L 的一个定向是指如下数据: 对每个点 $(x,y,z) \in L$, 指定该点处 L 的一个单位长度的切向量 $\mathbf{e}(x,y,z)$, 要求这一族切向量 $\{\mathbf{e}(x,y,z)\}_{(x,y,z)\in L}$ 随着 $(x,y,z) \in L$ 连续变化, 即要求映射

$$L \to \mathbf{R}^3,$$
$$(x,y,z) \to \mathbf{e}(x,y,z)$$

是连续的.

给定了一个定向的曲线称为定向曲线, 记作 $\vec{L}$.

例 3.5.6. 曲线 $I=[a,b] \subseteq \mathbf{R} \subseteq \mathbf{R}^3$ 上有两个定向 $\{\mathbf{e}_1(x)\}_{x\in I}$(指向 x 轴正方向) 与 $\{\mathbf{e}_2(x)\}_{x\in I}$(指向 x 轴负方向):

$$\mathbf{e}_1(x) = (1,0,0), \quad \forall x \in I;$$
$$\mathbf{e}_2(x) = -(1,0,0), \quad \forall x \in I.$$

例 3.5.7. 曲线 $S^1 = \{(x,y)|x^2+y^2=1\} \subseteq \mathbf{R}^2$ 上有两个定向 $\{\mathbf{e}_1(x,y)\}_{(x,y)\in S^1}$(逆时针方向) 与 $\{\mathbf{e}_2(x,y)\}_{(x,y)\in S^1}$(顺时针方向):

$$\mathbf{e}_1(x,y) = (-y,x), \quad \forall (x,y) \in S^1;$$
$$\mathbf{e}_2(x,y) = (y,-x), \quad \forall (x,y) \in S^1.$$

例 3.5.8. 设 L 是连通的曲线, 则 L 至多有两个不同的定向. 事实上, 人们可以证明: 连通的曲线都恰好有两个定向.

例 3.5.9. 设 $\gamma:[a,b] \to \mathbf{R}^3$ 是参数曲线, 令 $L = \mathrm{Im}(\gamma)$ 是 γ 的像集, 则 γ 给出 L 的一个定向:

$$\mathbf{e}_{\gamma(t)} = \frac{\gamma'(t)}{|\gamma'(t)|}, \quad \forall t \in [a,b].$$

所以, 曲线的参数化给出曲线的一个定向.

3.5.3 定向曲线上的第二型曲线积分

定义 3.5.10. 设 $\vec{L}$ 是定向曲线, 其定向由 $\{\mathbf{e}(x,y,z)\}_{(x,y,z)\in L}$ 给出. 对于连续的矢量值函数 $\mathbf{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$, 定义其在 $\vec{L}$ 上的第二型曲线积分为

$$\int_{\vec{L}} \mathbf{F}\cdot d\mathbf{r} = \int_{\vec{L}} Pdx + Qdy + Rdz := \int_L (\mathbf{F}(x,y,z)\cdot \mathbf{e}(x,y,z))dl.$$

如果 $\vec{L}$ 是封闭的, 也把上述积分记作 $\oint_{\vec{L}} \mathbf{F}\cdot d\mathbf{r}$ 或 $\oint_{\vec{L}} Pdx+Qdy+Rdz$.

第二型曲线积分的简单性质.

(1) 积分对被积函数是线性的. 设 $\mathbf{F}_1, \mathbf{F}_2$ 都是连续映射, 则对任何常数 c_1, c_2, 有

$$\int_{\vec{L}} (c_1\mathbf{F}_1(x,y,z) + c_2\mathbf{F}_2(x,y,z)) \cdot d\mathbf{r} = c_1 \int_{\vec{L}} \mathbf{F}_1(x,y,z) \cdot d\mathbf{r} + c_2 \int_{\vec{L}} \mathbf{F}_2(x,y,z) \cdot d\mathbf{r};$$

(2) 如果将 L 切成两段之并 $L = L_1 \cup L_2$, 则 L 的定向 $\{\mathbf{e}(x,y,z)\}_{(x,y,z)\in L}$ 给出 L_1 与 L_2 的定向. 对连续的向量值函数 $\mathbf{F}$, 有

$$\int_{\vec{L}} \mathbf{F}(x,y,z) \cdot d\mathbf{r} = \int_{\vec{L}_1} \mathbf{F}(x,y,z) \cdot d\mathbf{r} + \int_{\vec{L}_2} \mathbf{F}(x,y,z) \cdot d\mathbf{r}.$$

(3) 如果对 L 赋予相反的定向 $\{-\mathbf{e}(x,y,z)\}_{(x,y,z)\in L}$, 称所得的定向曲线为 $\vec{L}$ 的反定向曲线, 记作 $-\vec{L}$. 对连续的向量值函数 $\mathbf{F}$, 有

$$\int_{-\vec{L}} \mathbf{F}(x,y,z) \cdot d\mathbf{r} = -\int_{\vec{L}} \mathbf{F}(x,y,z) \cdot d\mathbf{r}.$$

3.5.4 参数曲线与定向曲线上第二型积分的关系

设 $\vec{L}$ 是定向曲线, 其定向由 $\{\mathbf{e}(x,y,z)\}_{(x,y,z)\in L}$ 给出. 如果参数曲线 $\gamma : [a,b] \to \mathbf{R}^3$ 满足如下条件:

(1) γ 是单射, 且 $\mathrm{Im}(\gamma) = L$;

(2) 对任何 $t \in [a,b]$, 有

$$\frac{\gamma'(t)}{|\gamma'(t)|} = \mathbf{e}(x(t), y(t), z(t));$$

则称 γ 是 $\vec{L}$ 的与定向相容的参数化.

命题 3.5.11. 设 γ 是 $\vec{L}$ 的与定向相容的参数化, 则对连续的向量值函数 $\mathbf{F}$ 有

$$\int_{\vec{L}} \mathbf{F} \cdot d\mathbf{r} = \int_{\gamma} \mathbf{F} \cdot d\mathbf{r}.$$

例 3.5.12. 计算

$$\oint_C \frac{-(x+y)dx + (x-y)dy}{x^2+y^2},$$

其中 C 为逆时针方向的圆周 $x^2 + y^2 = a^2 (a > 0)$.

例 3.5.13. 计算

$$\oint_L xydx + yzdy + zxdz,$$

其中 L 为椭圆周

$$\begin{cases} x^2 + y^2 = 1 \\ x + y + z = 1 \end{cases}$$

积分方向从 $(1,0,0)$ 沿着劣弧指向 $(0,1,0)$.

3.6 第二型曲面积分

3.6.1 Motivation

问题 3.6.1. 设三维流体在每个点 (x,y,z) 处的流速为

$$\mathbf{v}(x,y,z) = (v_1(x,y,z), v_2(x,y,z), v_3(x,y,z)).$$

求流体在单位时间内流出曲面 S 的流量.

沿用我们熟悉的 "剖分-近似-求和-取极限" 的办法:

(1) 把 S 剖分成 $S = \bigcup_{i=1}^n S_i$, 在每小块 S_i 上取点 $(x_i, y_i, z_i) \in S_i$.

(2) S_i 附近流体的速度近似为 $\mathbf{v}(x_i, y_i, z_i)$, S_i 上单位外法方向近似为 $\mathbf{n}(x_i, y_i, z_i)$, 则单位时间内流出 S_i 的流体流量近似为 $\mathbf{v}(x_i, y_i, z_i) \cdot \mathbf{n}(x_i, y_i, z_i)\mathrm{area}(S_i)$.

(3) 单位时间内流出 S 的流体流量近似为

$$\sum_{i=1}^n \mathbf{v}(x_i, y_i, z_i) \cdot \mathbf{n}(x_i, y_i, z_i)\mathrm{area}(S_i).$$

(4) 当剖分得越来越细时, 上述和式的极限

$$\lim \sum_{i=1}^n \mathbf{v}(x_i, y_i, z_i) \cdot \mathbf{n}(x_i, y_i, z_i)\mathrm{area}(S_i)$$

就是流体在单位时间内流出 S 的流量, 这恰好是 $\mathbf{v} \cdot \mathbf{n}$ 在 S 上的第一型曲面积分

$$\iint_S \mathbf{v}(x,y,z) \cdot \mathbf{n}(x,y,z) dS.$$

人们把形如

$$\iint_S (\mathbf{F}(x,y,z) \cdot \mathbf{n}(x,y,z)) dS$$

的积分称为第二型曲面积分.

3.6.2 曲面的定向

定义 3.6.2. 曲面 S 的一个定向, 记作 $\mathcal{O} = \{\mathbf{n}(x,y,z)\}_{(x,y,z)\in S}$, 是指如下数据: 对每个点 $(x,y,z)\in S$, 指定该点处 S 的一个法向量 $\mathbf{n}(x,y,z)$, 要求:

(1) $|\mathbf{n}(x,y,z)| = 1, \quad \forall (x,y,z)\in S.$

(2) $\mathbf{n}(x,y,z)$ 随 (x,y,z) 连续变化, 具体的说, 即映射

$$S \to \mathbf{R}^3,$$

$$(x,y,z) \to \mathbf{n}(x,y,z)$$

是连续的.

例 3.6.3. 球面 S^2 有两个不同的定向

$$\mathcal{O}_1 = \{\mathbf{n_1}(x,y,z) = (x,y,z)\},$$

$$\mathcal{O}_2 = \{\mathbf{n}_2(x,y,z) = -(x,y,z)\}.$$

例 3.6.4. 连通曲面上至多有两个不同的定向. 实际上, 如果 $\mathcal{O}_1 = \{\mathbf{n}_1(x,y,z)\}_{(x,y,z)\in S}, \mathcal{O}_2 = \{\mathbf{n}_2(x,y,z)\}_{(x,y,z)\in S}$ 是 S 的两个定向, 则对任何 $(x,y,z)\in S$, $\mathbf{n}_1(x,y,z)$ 与 $\mathbf{n}_2(x,y,z)$ 或者相等或者相反. 特别的, $\frac{\mathbf{n}_1(x,y,z)}{\mathbf{n}_2(x,y,z)}$ 是 S 上的连续函数, 取值 ± 1. 当 S 连通时, $\frac{\mathbf{n}_1}{\mathbf{n}_2}$ 只能恒等于 1 或恒等于 -1.

定义 3.6.5. 如果 S 上存在一个定向, 则称 S 是可定向曲面; 否则, 则称 S 是不可定向曲面.
赋予了一个定向的曲面称为定向的曲面.

例 3.6.6 (不可定向曲面)**.** 莫比乌斯带 (Mobius band), 克莱因瓶 (Klein Bottle).

例 3.6.7 (函数图像是可定向的)**.** 设 S 是 C^1 光滑函数 $g : \mathbf{R}^2 \to \mathbf{R}$ 的图像

$$S = \{(x,y,z)|z = g(x,y), (x,y)\in D\} = \{(x,y,z)|z - g(x,y) = 0, (x,y)\in D\}.$$

S 在 $(x,y,g(x,y))$ 处的法向量都形如 $\lambda(-g_x, -g_y, 1)$, 其中 $\lambda\in\mathbf{R}$. 定义

$$\mathbf{n}(x,y,g(x,y)) = \frac{(-g_x,-g_y,1)}{|(-g_x,-g_y,1)|},$$

它给出 S 的一个定向. 由于 $\mathbf{n}$ 与 z 轴正方向的夹角小于 $\frac{\pi}{2}$, 是指向上方的, 我们称这个定向为 "指向上方的".

例 3.6.8 (三维区域的边界是可定向的)**.** 设 Ω 是三维区域, $S = \partial\Omega$, 则 S 是可定向曲面. 实际上, 可以定义 $\mathbf{n}$ 处处指向 Ω 外面. 以后, 我们称这个定向为 "指向外面的".

3.6.3 第二型曲面积分

为了定义第二型曲面积分, 我们需要事先指定每点处的 $\mathbf{n}(x,y,z)$, 即事先指定 S 的一个定向. 换句话说, 我们只能在定向曲面上定义第二型曲面积分.

定义 3.6.9 (第二型曲面积分)**.** 设 S 是定向曲面, $\{\mathbf{n}(x,y,z)\}_{(x,y,z)\in S}$ 是定向. 对于矢量值函数

$$\mathbf{F}(x,y,z)=(P(x,y,z),Q(x,y,z),R(x,y,z)),$$

定义它在 S 上的第二型的曲面积分为

$$\iint_S \mathbf{F}(x,y,z)\cdot d\mathbf{S}:=\iint_S(\mathbf{F}(x,y,z)\cdot\mathbf{n}(x,y,z))dS.$$

第二型曲面积分的简单性质.

(1) 对被积函数是线性的

$$\iint_S(\alpha\mathbf{F}(x,y,z)+\beta\mathbf{G}(x,y,z))\cdot d\mathbf{S}=\alpha\iint_S\mathbf{F}(x,y,z)\cdot d\mathbf{S}+\beta\iint_S\mathbf{G}(x,y,z)\cdot d\mathbf{S}.$$

(2) 对积分区域是可加的. 把 S 切开成 S_1,S_2 的并 $S=S_1\cup S_2$, 则 S 的定向诱导 S_1 与 S_2 的定向, 对任何连续的 $\mathbf{F}$, 有

$$\iint_S\mathbf{F}(x,y,z)\cdot d\mathbf{S}=\iint_{S_1}\mathbf{F}(x,y,z)\cdot d\mathbf{S}+\iint_{S_2}\mathbf{F}(x,y,z)\cdot d\mathbf{S}.$$

(3) 设 $(S,\mathcal{O}=\{\mathbf{n}(x,y,z)\}_{(x,y,z)\in S})$ 是定向曲面, 则 $\{-\mathbf{n}(x,y,z)\}_{(x,y,z)\in S}$ 给出 S 的另一个定向, 称为 $\mathcal{O}$ 的反定向, 记作 $-\mathcal{O}$. 在 S 上赋予定向 $-\mathcal{O}$, 所得的定向曲面记作 $-S$, 称为 S 的反定向曲面, 对任何连续的 $\mathbf{F}$, 有

$$\iint_{-S}\mathbf{F}(x,y,z)\cdot d\mathbf{S}=-\iint_S\mathbf{F}(x,y,z)\cdot d\mathbf{S}.$$

3.6.4 曲面的参数化

定义 3.6.10. 设 $D\subset\mathbf{R}^2$, 称 C^1 光滑的映射 $\phi:D\to\mathbf{R}^3,\phi(u,v)=(x(u,v),y(u,v),z(u,v))$ 是曲面 S 的一个参数化, 如果它满足如下条件:

(1) $S=\mathrm{Im}(\phi)$ 且 ϕ 至多在边界上不单.

(2) 对任何 $(u,v)\in D$, 有 $(x_u,y_u,z_u)\times(x_v,y_v,z_v)\neq(0,0,0)$.

例 3.6.11. 设 $\phi : D \to \mathbf{R}^3$ 是曲面 S 的一个参数化, $\mathbf{p} = (x(u,v), y(u,v), z(u,v)) \in S$, 则 S 在 $\mathbf{p}$ 处的切空间为

$$T_{\mathbf{p}}S = \{a(x_u, y_u, z_u) + b(x_v, y_v, z_v) | a, b \in \mathbf{R}\}.$$

由此可知 S 在 $\mathbf{p}$ 处的所有法向量为

$$\{\lambda(x_u, y_u, z_u) \times (x_v, y_v, z_v) | \lambda \in \mathbf{R}\}.$$

例 3.6.12 (参数化给出定向)**.** 设 $\phi : D \to \mathbf{R}^3$ 是曲面 S 的一个参数化, 则它给出 S 的一个定向:

$$\mathbf{n}_\phi(x(u,v), y(u,v), z(u,z)) = \frac{(x_u, y_u, z_u) \times (x_v, y_v, z_v)}{|(x_u, y_u, z_u) \times (x_v, y_v, z_v)|}.$$

以后, 我们称上述定向为 "参数化确定的 (给出的) 定向".

定义 3.6.13. 设 S 是定向曲面, $\mathcal{O} = \{\mathbf{n}(x,y,z)\}_{(x,y,z)\in S}$ 是其定向. 称 S 的参数化 $\phi : D \to \mathbf{R}^3$ 是与定向相容的, 如果

$$\mathbf{n}_\phi(x(u,v), y(u,v), z(u,z)) = \mathbf{n}(x(u,v), y(u,v), z(u,v)), \quad \forall (u,v) \in D.$$

3.6.5　第二型曲面积分的计算方法

命题 3.6.14. 设 S 是定向的曲面, $\phi : D \to \mathbf{R}^3$ 是与定向相容的参数化,

$$\phi(u,v) = (x(u,v), y(u,v), z(u,v)).$$

对任何连续映射 $\mathbf{F} = (P, Q, R)$, 有

$$\begin{aligned}
&\iint_S \mathbf{F} \cdot d\mathbf{S} \\
=& \iint_D (P, Q, R) \cdot ((x_u, y_u, z_u) \times (x_v, y_v, z_v))\, dudv \\
=& \iint_D \det \begin{pmatrix} P & Q & R \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{pmatrix} dudv \\
=& \iint_D ((y_u z_v - y_v z_u)P + (z_u x_v - z_v x_u)Q + (x_u y_v - x_v y_u)R)\, dudv.
\end{aligned}$$

为了便于记忆, 我们可以做如下形式化的运算:

$$dydz = (y_u du + y_v dv)(z_u du + z_v dv),$$

如果约定

$$dudu = 0, \quad dvdv = 0, \quad dvdu = -dudv,$$

则有

$$dydz = (y_u z_v - y_v z_u)dudv.$$

类似的, 有

$$dzdx = (z_u x_v - z_v x_u)dudv, \quad dxdy = (x_u y_v - x_v y_u)dudv,$$

这样, 第二型积分可以写成

$$\iint_S (P, Q, R) \cdot d\mathbf{S} = \iint_S Pdydz + Qdzdx + Rdxdy,$$

以后, 我们经常把第二型曲面积分写成上式右边的形式.

例 3.6.15. 设曲面 S 是函数图像

$$S = \{(x, y, z)|z = g(x, y), (x, y) \in D\},$$

取指向上方的定向, 则对任何连续映射 $\mathbf{F}(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z))$, 由上述命题可得:

$$\begin{aligned}
&\iint_S \mathbf{F}(x, y, z) \cdot \mathbf{dS} \\
=&\iint_D (P, Q, R) \cdot ((1, 0, g_x) \times (0, 1, g_y))dxdy \\
=&\iint_D [-g_x P(x, y, g(x, y)) - g_y Q(x, y, g(x, y)) + R(x, y, g(x, y))]dxdy.
\end{aligned}$$

例 3.6.16. 设 S 为球面 $x^2 + y^2 + z^2 = 1$, 取指向外面的定向. 设 $\mathbf{F}(x, y, z) = \frac{(x,y,z)}{(x^2+y^2+z^2)^{3/2}}$, 计算

$$\iint_S \mathbf{F} \cdot d\mathbf{S}.$$

例 3.6.17. 设 S 为上半球面 $z = \sqrt{R^2 - x^2 - y^2}$, 取指向上方的定向. 计算

$$\iint_S yzdzdx + zxdxdy.$$

3.7 Green 公式

例 3.7.1. 考虑平面上的曲边梯形区域 (为了方便起见, 称为第一类曲边梯形)

$$D = \{(x, y)|a \le x \le b, y_1(x) \le y \le y_2(x)\},$$

设 $C = \partial D$, 定向为逆时针方向. 设 $P(x, y)$ 是 D 上的 C^1 光滑函数, 计算

$$\oint_C P dx.$$

解. C 分成四段 $C = C_1C_2C_3C_4$,

$$C_1 = \{(x, y_1(x))|a \le x \le d\},$$

$$C_3 = \{(x, y_2(x))|a \le x \le d\},$$

C_2, C_4 分别是 D 的最右边和最左边的竖直的边界, 这里 C_1 取从左至右的定向, C_3 取从右至左的定向.

由定义可得

$$\int_{C_1} P dx = \int_a^b P(x, y_1(x))dx, \quad \int_{C_2} P dx = 0,$$

$$\int_{C_3} P dx = -\int_a^b P(x, y_2(x))dx, \quad \int_{C_4} P dx = 0.$$

从而有

$$\oint_C P dx = -\int_a^b \left(P(x, y_2(x)) - P(x, y_1(x))\right) dx. \tag{3.1}$$

利用 Newton-Leibniz 公式

$$P(x, y_2(x)) - P(x, y_1(x)) = \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x, y)}{\partial y} dy,$$

代入 (1) 式可得

$$\oint_C P dx = -\int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x, y)}{\partial y} dy = -\iint_D \frac{\partial P(x, y)}{\partial y} dx dy.$$

□

上述例子表明, 对于第一类曲边梯形我们可以把边界上的曲线积分化成区域上的二重积分. 这与 Newton-Leibniz 公式类似, 因为 Newton-Leibniz 公式

$$F|_{\partial[a,b]} = \int_{[a,b]} F'(x)dx$$

把闭区间边界上的求和 (0 维积分) 化成闭区间上的 (1 维) 积分. 因此, 我们希望上述例子中的现象对一般的平面区域也成立.

定义 3.7.2. 称平面曲线 C 为简单闭曲线, 如果它是封闭的 (起点和终点重合) 且没有自交点 (端点除外).

具体的说, 一条简单闭曲线是指连续映射 $\gamma:[0,1]\to\mathbf{R}^2$, 且满足条件:

(1) $\gamma(0)=\gamma(1)$,

(2) 对任何 $t_1\neq t_2$, 如果 $\{t_1,t_2\}\neq\{0,1\}$, 则 $\gamma(t_1)\neq\gamma(t_2)$.

Jordan 发现如下事实

定理 3.7.3 (Jordan 曲线定理)**.** *设 $C\subseteq\mathbf{R}^2$ 是简单闭曲线, 则 $\mathbf{R}^2\backslash C$ 不连通, 它是两个不相交的连通部分的并. 这两个部分其中一个是有界的, 称为 C 的内部, 另一个是无界的, 称为 C 的外部. 更进一步, C 的内部与圆盘 $\mathring{D}=\{(x,y)|x^2+y^2<1\}$ 同胚 (形状"本质上"一样).*

粗略的说, 所谓 X 与 Y 同胚是指 X 与 Y 的形状"本质上"是一样的. 稍微具体的说, 想象 X 与 Y 都是由橡皮材料做成的, 则可以适当的揉捏 X, 要求不能撕裂也不能粘合, 能使它变成 Y.

追踪揉捏过程中每一个点的去向, 得到映射 $f:X\to Y$. 容易看出:

(1) 不撕裂等价于说 f 连续;

(2) 不粘合等价于 f^{-1} 不撕裂, 从而等价于 f^{-1} 连续.

因此, 人们给出如下定义

定义 3.7.4. 称 $f:X\to Y$ 为同胚映射, 如果 f 是一一映射, 且 f 与 f^{-1} 都是连续映射. 称 X 与 Y 同胚, 如果存在某个同胚映射 $f:X\to Y$.

回到 Green 公式. 设 C 是分段光滑的简单闭曲线, 定向为逆时针方向, D 是 C 的内部. 可以把 D 剖分成有限个第一类曲边梯形的并

$$D=\bigcup_{i=1}^{n}D_i,$$

利用前述例子中的计算结果, 有

$$\oint_{\partial D_i}Pdx=-\iint_{D_i}\frac{\partial P(x,y)}{\partial y}dxdy,$$

求和得

$$\sum_{i=1}^{n}\oint_{\partial D_i}Pdx=-\sum_{i=1}^{n}\iint_{D_i}\frac{\partial P(x,y)}{\partial y}dxdy=-\iint_{D}\frac{\partial P(x,y)}{\partial y}dxdy,$$

注意到剖分中每段分割线对上式左边贡献两项, 分别为在它上面沿正反两个方向的第二型曲线积分, 这两项相互抵消, 只剩下在 C 上的曲线积分. 所以

$$\oint_{\partial D} Pdx = -\iint_D \frac{\partial P(x,y)}{\partial y}dxdy.$$

更加一般的, 设 E 是挖了 m 个洞的 (拓扑) 圆盘, 设其最外面的边界为 C_0, 每个洞的边界分别为 $C_1,...,C_m$. 则对每个洞割两刀, 可以把 E 分割成若干个拓扑圆盘的并, 把分割后每个拓扑圆盘上的 Green 公式相加, 并注意到每段分割线上都积分两次且定向相反, 所以有

$$\oint_{C_0} Pdx + \sum_{i=1}^m \oint_{C_i} Pdx = -\iint_E \frac{\partial P(x,y)}{\partial y}dxdy,$$

其中 C_0 取逆时针定向, $C_1,...,C_m$ 取顺时针定向.

完全类似的, 我们可以考虑积分 $\oint_{\partial E} Qdy$, 先在局部模型 (称为第二类曲边梯形)

$$D' = \{(x,y)|c \le y \le d, x_1(y) \le x \le x_2(y)\},$$

上计算

$$\oint_{\partial D'} Qdy = \iint_D \frac{\partial Q(x,y)}{\partial x}dxdy,$$

再利用上述剖分的手法可以证明: 对挖了 m 个洞的 (拓扑) 圆盘 E, 有

$$\oint_{C_0} Qdy + \sum_{i=1}^m \oint_{C_i} Qdy = \iint_E \frac{\partial Q(x,y)}{\partial x}dxdy.$$

总结一下, 写成如下的定理.

定理 3.7.5 (Green 公式)**.** *设 D 是挖了 m 个洞的 (拓扑) 圆盘, 设其最外面的边界为 C_0, 取逆时针定向; 设每个洞的边界分别为 $C_1,...,C_m$, 取顺时针定向, 设 $C_0,...,C_m$ 都是分段光滑的闭曲线. 设 $P(x,y),Q(x,y)$ 是 D 上的 C^1 光滑函数, 则有*

$$\oint_{C_0}(Pdx+Qdy) + \sum_{i=1}^m \oint_{C_i}(Pdx+Qdy) = \iint_D (\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y})dxdy.$$

为了写成统一的形式, 人们规定 ∂D 上的正定向为: 如果依该定向沿着边界走, 则 D 总在左边, 把带有该定向的边界记作 ∂D^+. 容易验证: C_0 上的正定向为逆时针方向, $C_1,...,C_m$ 上的正定向为顺时针方向.

所以 Green 公式可以写成

$$\boxed{\oint_{\partial D^+}(Pdx+Qdy) = \iint_D \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}\right)dxdy.}$$

例 3.7.6 (面积可化成曲线积分)**.** 若 $\frac{\partial Q(x,y)}{\partial x}-\frac{\partial P(x,y)}{\partial y}\equiv 1$, 则

$$\oint_{\partial D^+} Pdx+Qdy=\iint_D\left(\frac{\partial Q(x,y)}{\partial x}-\frac{\partial P(x,y)}{\partial y}\right)dxdy=\iint_D dxdy=\mathrm{area}(D).$$

特别的

$$\oint_{\partial D^+} xdy=\oint_{\partial D^+} -ydx=\frac{1}{2}\oint_{\partial D^+} xdy-ydx=\mathrm{area}(D).$$

例 3.7.7. 计算椭圆盘 $\frac{x^2}{a^2}+\frac{y^2}{b^2}\le 1$ 的面积.

例 3.7.8. 设平面流体的速度是 $\mathbf{v}(x,y)=(v_1(x,y),v_2(x,y))$, 则单位时间内流体通过闭曲线 $C=\partial D$ 的流量为

$$\Phi=\oint_C \mathbf{v}\cdot\mathbf{n}dl,$$

其中 $\mathbf{n}$ 表示 C 的单位外法向量, 指向 D 外部. 证明:

$$\oint_C \mathbf{v}\cdot\mathbf{n}dl=\iint_D\left(\frac{\partial v_1}{\partial x}+\frac{\partial v_2}{\partial y}\right)dxdy.$$

例 3.7.9. $C=\partial D$, 设 u 是 D 上的 C^2 光滑函数, 则

$$\oint_{\partial D}\frac{\partial u}{\partial \mathbf{n}}dl=\iint_D \Delta u dxdy,$$

其中 $\Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$ 是 Laplace 算子.

例 3.7.10 (Gauss 积分)**.** 计算曲线积分

$$\oint_L\frac{-ydx+xdy}{x^2+y^2},$$

其中 L 为逆时针定向的封闭曲线, $(0,0)\notin L$.

例 3.7.11. 计算积分

$$\iint_{\mathbf{R}^2}\ln\sqrt{x^2+y^2}\Delta f(x,y)dxdy.$$

3.8 Green 公式的应用

本小节内容不作要求.

命题 3.8.1. *设 n 是非零整数, R 是正实数,*

$$B(R)=\{(x,y)|x^2+y^2\le R^2\},$$

则不存在 C^2 光滑映射 $f:B(R)\to\mathbf{R}^2\backslash\{(0,0)\}$ 使得

$$f(z)=z^n,\quad \forall z\in\partial B(R)=\{z\in\mathbf{C}:|z|=R\}.\tag{3.2}$$

注 3.8.2. 这里我们将 $\mathbf{R}^2$ 与复平面 $\mathbf{C}$ 等同, 点 $(x,y)\in\mathbf{R}^2$ 对应于复数 $x+yi$. 上述命题中的条件 (1) 可翻译成

$$f(R\cos\theta, R\sin\theta) = R^n(\cos n\theta, \sin n\theta), \quad \forall\theta\in[0,2\pi].$$

证明: 用反证法, 假设存在 C^2 光滑映射 $f: B(R)\to\mathbf{R}^2\backslash\{(0,0)\}$

$$f(x,y) = (u(x,y), v(x,y))$$

使得 (1) 成立, 则 $u^2(x,y)+v^2(x,y)\neq 0, \forall(x,y)\in B(R)$.

考虑如下的第二型曲线积分:

$$I = \oint_{\partial B(R)} \frac{uv_x - u_x v}{u^2(x,y)+v^2(x,y)}dx + \frac{uv_y - u_y v}{u^2(x,y)+v^2(x,y)}dy. \tag{3.3}$$

一方面, 由 Green 公式可得:

$$\begin{aligned}
&\oint_{\partial B(R)} \frac{uv_x - u_x v}{u^2(x,y)+v^2(x,y)}dx + \frac{uv_y - u_y v}{u^2(x,y)+v^2(x,y)}dy \\
&= \iint_{B(R)} \left(\frac{\partial}{\partial x}\frac{uv_y - u_y v}{u^2(x,y)+v^2(x,y)} - \frac{\partial}{\partial y}\frac{uv_x - u_x v}{u^2(x,y)+v^2(x,y)}\right) dxdy \\
&= \iint_{B(R)} 0 dxdy = 0.
\end{aligned}$$

另一方面, 取 $\partial B(R)$ 的参数化实现

$$\gamma(\theta) = R(\cos\theta, \sin\theta), \quad \theta\in[0,2\pi],$$

直接计算, 可得:

$$I = \int_0^{2\pi} \frac{(-v_x R\sin\theta + v_y R\cos\theta)u - (-u_x R\sin\theta + u_y R\cos\theta)v}{u^2+v^2}d\theta. \tag{3.4}$$

注意到

$$u(R\cos\theta, R\sin\theta) = R^n\cos n\theta, \quad v(R\cos\theta, R\sin\theta) = R^n\sin n\theta,$$

求导可得

$$-u_x R\sin\theta + u_y R\cos\theta = -nR^n\sin n\theta, \quad -v_x R\sin\theta + v_y R\cos\theta = nR^n\cos n\theta,$$

代入 (3) 式, 得:

$$\begin{aligned}
I &= \int_0^{2\pi} \frac{nR^n\cos n\theta\cdot R^n\cos n\theta + nR^n\sin n\theta\cdot R^n\sin n\theta}{R^{2n}}d\theta \\
&= \int_0^{2\pi} nd\theta = 2n\pi,
\end{aligned}$$

与之前的计算结果 $I=0$ 矛盾! 这就完成了命题的证明. □

利用上述命题, 我们可以证明两个著名的定理.

定理 3.8.3 (Brouwer 不动点定理)**.** 设 $B=\{(x,y)|x^2+y^2\le 1\}$ 为单位圆盘, $f:B\to B$ 是 C^2 光滑映射, 则存在 $\mathbf{x}\in B$ 使得 $f(\mathbf{x})=\mathbf{x}$.

定理 3.8.4 (代数基本定理, Gauss)**.** 设 $P(z)=z^n+a_1z^{n-1}+...+a_{n-1}z+a_n$ 是多项式, $n\in\mathbf{Z}_+$, 则存在 $z\in\mathbf{C}$ 使得 $P(z)=0$.

3.9 Gauss 公式

例 3.9.1. 考虑空间中的曲边圆柱体区域 (为了方便起见, 称为第三类曲边圆柱体)

$$V_3=\{(x,y,z)|(x,y)\in D, z_1(x,y)\le z\le z_2(x,y)\},$$

令 $S_3=\partial V_3$, 取指向外面的定向. 设 $R(x,y,z)$ 是 V_3 上的 C^1 光滑函数, 求

$$\iint_{S_3} Rdxdy.$$

证明:

$$\begin{aligned}
&\iint_{S_3} Rdxdy\\
=&\iint_D (R(x,y,z_2(x,y))-R(x,y,z_1(x,y)))dxdy\\
=&\iint_D dxdy\int_{z_1(x,y)}^{z_2(x,y)}\frac{\partial R}{\partial z}dz\\
=&\iiint_{V_3}\frac{\partial R}{\partial z}dxdydz.
\end{aligned}$$

□

仿照 Green 公式的做法, 可以把上述结果推广到一般的三维区域. 只需把一般的三维区域 Ω 分割成有限个第三类曲边圆柱体的并, 对每一小块用例3.9.1的结论, 再把所得到的式子相加, 即得:

$$\iint_{\partial\Omega} Pdydz=\iiint_\Omega \frac{\partial P}{\partial x}dxdydz,$$

其中 $\partial\Omega$ 取指向外面的定向.

完全类似的, 对于另外两个方向的曲边圆柱体

$$V_1=\{(x,y,z)|(y,z)\in D, x_1(y,z)\le x\le x_2(y,z)\},$$

$$V_2=\{(x,y,z)|(x,z)\in D, y_1(x,z)\le y\le y_2(x,z)\},$$

分别有

$$\iint_{\partial V_1} Pdydz=\iiint_{V_1}\frac{\partial P}{\partial x}dxdydz,$$

$$\iint_{\partial V_2} Qdydz=\iiint_{V_2}\frac{\partial Q}{\partial y}dxdydz.$$

把这些局部的计算结果推广到一般的三维区域, 得到 Gauss 公式.

定理 3.9.2 (Gauss 公式)**.** 设三维有界闭区域 Ω 的边界 $\partial\Omega$ 是分块光滑的曲面, 取指向 Ω 外面的定向. 设 P,Q,R 是 Ω 上的 C^1 光滑函数, 则有

$$\iint_{\partial\Omega} Pdydz+Qdzdy+Rdxdy=\iiint_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dxdydz$$

例 3.9.3. 设闭曲面 S 是三维闭区域 Ω 的边界, 取指向外面的定向. 已知 $(0,0,0)\notin S$, 计算

$$\iint_{S^+}\frac{xdydz+ydzdx+zdxdy}{(x^2+y^2+z^2)^{3/2}}.$$

例 3.9.4. 计算积分

$$\iiint_{\mathbf{R}^3}\frac{1}{\sqrt{x^2+y^2+z^2}}\Delta f(x,y,z)dxdydz.$$

例 3.9.5 (Feynman 规则的玩具模型 (Toy model))**.** 设 $M=(M_{ij})_{1\le i,j\le n}$ 是正定矩阵, 令

$$S(x_1,...,x_n)=-\frac{1}{2}\sum_{i,j}M_{ij}x_ix_j.$$

对多项式 $f(x_1,...,x_n)$, 定义

$$\langle f\rangle=\int\cdots\int_{\mathbf{R}^n}f(x_1,...,x_n)e^{S(x_1,...,x_n)}|dx_1...dx_n|,$$

称为 f 的 correlation function, 这是物理学中关心的对象, 我们来计算 f 的 correlation function.

解. 令 $B(r)=\{(x_1,...,x_n)|\sum_{i=1}^n x_i^2\le r\}$ 为 n 维球体, 其边界 $S(r)=\partial B(r)$ 取指向球体外面的定向. Correlation function 的具体定义为

$$\langle f\rangle=\lim_{r\to\infty}\int\cdots\int_{\mathbf{B}(r)}f(x_1,...,x_n)e^{S(x_1,...,x_n)}|dx_1...dx_n|.$$

先以 $n=3$ 为例, 对多项式 $P(x,y,z)$, 由 Gauss 公式, 有

$$\iint_{S(r)}P(x,y,z)e^{S(x,y,z)}dydz=\iiint_{B(r)}\frac{\partial}{\partial x}(Pe^S)|dxdydz|,$$

对 $r\to\infty$ 取极限, 可得

$$0=\lim_{r\to\infty}\iint_{S(r)}P(x,y,z)e^{S(x,y,z)}dydz=\lim_{r\to\infty}\iiint_{B(r)}\frac{\partial}{\partial x}(Pe^S)|dxdydz|=\iiint_{\mathbf{R}^3}\frac{\partial}{\partial x}(Pe^S)|dxdydz|,$$

即有

$$\langle\frac{\partial P}{\partial x}\rangle+\langle\frac{\partial S}{\partial x}P\rangle=0.$$

类似的, 有

$$\langle\frac{\partial P}{\partial y}\rangle+\langle\frac{\partial S}{\partial y}P\rangle=0,\quad \langle\frac{\partial P}{\partial z}\rangle+\langle\frac{\partial S}{\partial z}P\rangle=0.$$

注意到, $\frac{\partial S}{\partial x}P$ 的次数比 $\frac{\partial P}{\partial x}$ 的次数大 2, 则上述结果给出了计算 correlation function 的递推式.

对一般的维数, 由 Stokes 定理, 有

$$\int_{S(r)}Pe^Sdx_1...\widehat{dx_i}...dx_n=(-1)^{i-1}\int_{B(r)}\frac{\partial}{\partial x_i}(Pe^S)|dx_1...dx_n|,$$

对 $r\to\infty$ 取极限, 可得

$$\langle\frac{\partial P}{\partial x_i}\rangle=-\langle\frac{\partial S}{\partial x_i}P\rangle=\sum_{j=1}^{n}M_{ij}\langle x_jP\rangle.$$

取 $P=x_kQ$, 有

$$\sum_j M_{ij}\langle x_jx_kQ\rangle=\langle\frac{\partial}{\partial x_i}(x_kQ)\rangle.$$

把上式视为矩阵等式, 可得

$$\langle x_ix_jQ\rangle=\sum_{k=1}^{n}(M^{-1})_{ik}\langle\frac{\partial}{\partial x_k}(x_jQ)\rangle.$$

设 Q 是单项式 $Q=x_{i_1}...x_{i_t}$, 则

$$\begin{aligned}\langle x_ix_jx_{i_1}...x_{i_t}\rangle&=\sum_{k=1}^{n}(M^{-1})_{ik}\langle\frac{\partial}{\partial x_k}(x_jx_{i_1}...x_{i_t})\rangle\\&=\sum_{k=1}^{n}(M^{-1})_{ik}\langle\delta_{kj}x_{i_1}...x_{i_t}\rangle+\sum_{k=1}^{n}(M^{-1})_{ik}\langle x_j\delta_{ki_1}x_{i_2}...x_{i_t}\rangle+...\\&+\sum_{k=1}^{n}(M^{-1})_{ik}\langle x_jx_{i_1}...x_{i_{t-1}\delta_{ki_t}}\rangle\\&=(M^{-1})_{ij}\langle x_{i_1}...x_{i_t}\rangle+(M^{-1})_{ii_1}\langle x_jx_{i_2}...x_{i_t}\rangle+...+(M^{-1})_{ii_t}\langle x_jx_{i_1}...x_{i_{t-1}}\rangle.\end{aligned}$$

这就得到了计算 $\langle x_ix_jx_{i_1}...x_{i_t}\rangle$ 的方法: 在平面上画 $t+2$ 个点, 分别标记为 $x_i,x_j,x_{i_1},...,x_{i_t}$, 将 x_i 与后面的某个点 (设其标记为 x_k) 用一条边相连, 在该边上放置 $(M^{-1})_{ik}$; 类似的, 把剩下

的 t 个点也两两用边相连, 并在每条边上标记 M^{-1} 的相应的矩阵元. 这样得到一个图, 把该图所有边上所标的数相乘. 对所有可能的图, 每个图都由此规则给出一个乘积, 则 $\langle x_i x_j x_{i_1} \dots x_{i_t} \rangle$ 等于把所有图给出的乘积求和, 最后再乘以常数 $\langle 1 \rangle = \frac{(\sqrt{2\pi})^n}{\sqrt{\det M}}$.

上述计算方法来自量子物理学, 是著名物理学家 Feynman 发现的, 称为 Feynman rule.

□

3.10 Stokes 公式

Green 公式把平面曲线上的线积分转化成曲线所围成区域上的面积分. 一个自然的问题是: 能否推广到空间曲线? 具体的说, 设 S 是曲面, 是否有

$$\oint_{\partial S} Pdx + Qdy + Rdz = \iint_S ?$$

先考虑简单的情形. 设 S 是参数曲面

$$\begin{cases} x = x(u,v), \\ y = y(u,v), \\ z = z(u,v), \end{cases}$$

其中 $(u,v) \in D$, D 是参数平面上的有界闭区域, $x(u,v), y(u,v), z(u,v)$ 是光滑映射.

考虑映射

$$\phi : D \to \mathbf{R}^3,$$

$$\phi(u,v) = (x(u,v), y(u,v), z(u,v)),$$

ϕ 把 D 等同成空间曲面 S. 特别的, ϕ 也把两者的边界等同起来

$$\partial S = \phi(\partial D),$$

∂D 的正定向诱导 ∂S 的一个定向, 我们称之为 ∂S 的正定向, 记作 ∂S^+.

下面我们来计算

$$\oint_{\partial S^+} Pdx + Qdy + Rdz.$$

取 ∂D^+ 的一个参数化实现

$$\begin{cases} u = u(t), \\ v = v(t), \end{cases}$$

它自动诱导 ∂S^+ 的参数化实现

$$\begin{cases} x = x(u(t), v(t)), \\ y = y(u(t), v(t)), \\ z = z(u(t), v(t)), \end{cases}$$

由此可得:

$$\begin{aligned}
&\oint_{\partial S^+} P dx \\
&= \int P(x(u(t), v(t)), y(u(t), v(t)), z(u(t), v(t)))(x_u u'(t) + x_v v'(t)) dt \\
&= \oint_{\partial D^+} P(x(u,v), y(u,v), z(u,v))(x_u du + x_v dv) \\
&= \iint_D \left(\frac{\partial (x_v P(x,y,z))}{\partial u} - \frac{\partial (x_u P(x,y,z))}{\partial v} \right) dudv \\
&= \iint_D \left((x_v y_u - x_u y_v) P_y + (x_v z_u - x_u z_v) P_z \right) dudv \\
&= \iint_S \frac{\partial P}{\partial z} dzdx - \frac{\partial P}{\partial y} dxdy.
\end{aligned}$$

类似的, 有:

$$\oint_{\partial S^+} Q dy = \iint_S -\frac{\partial Q}{\partial z} dydz + \frac{\partial Q}{\partial x} dxdy.$$
$$\oint_{\partial S^+} R dz = \iint_S \frac{\partial R}{\partial y} dydz - \frac{\partial R}{\partial x} dzdx.$$

把这三个式子相加, 得:

$$\oint_{\partial S^+} P dx + Q dy + R dz = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy.$$

这样, 对于参数曲面我们推广了 Green 公式.

对于一般的曲面 S, 可以适当的把 S 分成有限个小块的并, 使得每一小块都有参数化表示. 利用上面的计算结果, 在每一小块上都可以把边界上的曲线积分化成内部的曲面积分, 把所有这些等式相加, 即得到著名的 Stokes 公式.

定理 3.10.1 (Stokes 公式)**.** *设定向曲面 S 的边界 ∂S 是分段光滑的, 按照右手法则对 ∂S 规定一个定向, 把取好这个定向的边界记作 ∂S^+. 设 P, Q, R 是 S 上的 C^1 光滑函数, 则有*

$$\oint_{\partial S^+} P dx + Q dy + R dz = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy.$$

∂S 的定向规则: 当右手的拇指在 S 的切平面中垂直于 ∂S 指向 S 外面, 食指指向 ∂S^+ 的方向时, 中指恰好指向 S 的法方向 $\mathbf{n}$.

3.10.1 Stokes 公式中的定向

在 Stokes 公式的证明中, 我们是按如下方式对 S 与 ∂S 赋予定向的.

- 取 S 的一个参数化实现 $\phi : D \to S$.
- ∂D 是平面区域的边界, 给它赋予正定向, 并任取它的一个与定向相容的参数化 $\gamma : [0,1] \to \partial D$.
- $\phi \circ \gamma : [0,1] \to \partial S$ 给 ∂S 赋予一个定向 $\{\mathbf{e}\}$.
- 对曲面 S 赋予定向 $\{\mathbf{n}\}$:

$$\mathbf{n}(x(u,v), y(u,v), z(u,v)) = \frac{(x_u, y_u, z_u) \times (x_v, y_v, z_v)}{|(x_u, y_u, z_u) \times (x_v, y_v, z_v)|}.$$

我们称 $(\partial S, \{\mathbf{e}\})$ 为定向曲面 $(S, \{\mathbf{n}\})$ 的正定向边界, 可以用如下方法更加直接的描述.

(1) S 的正法向量为

$$\mathbf{n}(x(u,v), y(u,v), z(u,v)) = \frac{(x_u, y_u, z_u) \times (x_v, y_v, z_v)}{|(x_u, y_u, z_u) \times (x_v, y_v, z_v)|}.$$

(2) ∂S 的定向. ∂D 是平面区域 D 的边界, 赋予正定向, 并取一个与定向相容的参数化 $\gamma : [0,1] \to \partial D$, 设 $\gamma'(t) = \mathbf{e}' = (\mathbf{e}_1', \mathbf{e}_2')$. 复合映射 $\phi \circ \gamma : [0,1] \to \partial S$ 是 ∂S 的参数化, 赋予 ∂S 一个定向 (不一定是单位长度的)

$$\mathbf{e} = (\phi \circ \gamma)'(t).$$

直接计算可得:

$$\begin{aligned}\mathbf{e} =& (\phi \circ \gamma)'(t) \\ =& (x_u \mathbf{e}_1' + x_v \mathbf{e}_2', y_u \mathbf{e}_1' + y_v \mathbf{e}_2', z_u \mathbf{e}_1' + z_v \mathbf{e}_2') \\ =& \mathbf{e}_1'(x_u, y_u, z_u) + \mathbf{e}_2'(x_v, y_v, z_v).\end{aligned}$$

(3) 位于 S 的切平面内且指向 S 外面的方向. 在 ∂D 上每一点处选取一个指向 D 外面的向量 $\mathbf{w}' = (\mathbf{w}_1', \mathbf{w}_2')$, 设这个向量由一条道路 $p : (-1,0] \to D$ 代表

$$p(0) \in \partial D, \quad p'(0) = \mathbf{w}',$$

$\phi \circ p$ 是 S 中的一条道路, 终点在 ∂S 上, 指向 S 外面, 设 $(\phi \circ p)'(0) = \mathbf{w}$, 则

$$\begin{aligned}\mathbf{w} =& (x_u \mathbf{w}_1' + x_v \mathbf{w}_2', y_u \mathbf{w}_1' + y_v \mathbf{w}_2', z_u \mathbf{w}_1' + z_v \mathbf{w}_2') \\ =& \mathbf{w}_1'(x_u, y_u, z_u) + \mathbf{w}_2'(x_v, y_v, z_v).\end{aligned}$$

引理 3.10.2. $\mathbf{w}, \mathbf{e}, \mathbf{n}$ 构成右手系.

证明: 在 D 所在的参数平面内, $\mathbf{e}'$ 可由 $\mathbf{w}'$ 逆时针旋转某个 $\theta \in (0, \pi)$ 角得到, 则

$$\det \begin{pmatrix} \mathbf{w}_1' & \mathbf{w}_2' \\ \mathbf{e}_1' & \mathbf{e}_2' \end{pmatrix} > 0.$$

由此可得:

$$(\mathbf{w} \times \mathbf{e}) \cdot \mathbf{n} = (\mathbf{w}_1'\mathbf{e}_2' - \mathbf{w}_2'\mathbf{e}_1')|(x_u, y_u, z_u) \times (x_v, y_v, z_v)| > 0,$$

可知 $\mathbf{w}, \mathbf{e}, \mathbf{n}$ 构成右手系. □

总结一下, 我们给出了曲面边界的定向法则, 这是 Stokes 公式中用到的定向.

定义 3.10.3 (曲面边界定向的右手法则)**.** 当右手的拇指在 S 的切平面中指向 S 外面, 食指指向 ∂S 的正定向时, 中指恰好指向 S 的法方向 $\mathbf{n}$.

3.10.2 Stokes 公式的应用

Stokes 公式把曲面边界上的积分转化成曲面上的积分.

例 3.10.4. Green 公式是 Stokes 公式的特例.

例 3.10.5. 设 $S \subset \mathbf{R}^3$ 是封闭的定向曲面, 则对任何 $P, Q, R \in C^1(S; \mathbf{R}^1)$, 有

$$\iint_S (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})dydz + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})dzdx + (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy = 0.$$

3.11 Stokes 定理

如何记住 Gauss 公式与 Stokes 公式? 如果做如下形式化的约定:

(1) 设 f 是光滑函数, 则 $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$;

(2) 设 ω 是由符号 dx, dy, dz 组成的字符串, 则

$$d(f\omega) = (df)\omega + fd(\omega);$$

(3) $d(dx) = d(dy) = d(dz) = 0$;

(4) $dxdx = dydy = dzdz = 0$, $dydx = -dxdy, dzdy = -dydz, dxdz = -dzdx$,

则可以把 Newton-Leibniz 公式, Green 公式, Gauss 公式, Stokes 公式写成统一的形式.

(1) Newton-Leibniz 公式

$$\int_{\partial[a,b]} F = \int_{[a,b]} dF,$$

其中 $dF = F'(x)dx$.

(2) Green 公式

$$\int_{\partial D^+} (Pdx + Qdy) = \int_D d(Pdx + Qdy),$$

其中

$$d(Pdx + Qdy) = (P_x dx + P_y dy)dx + (Q_x dx + Q_y dy)dy = (Q_x - P_y)dxdy.$$

(3) Gauss 公式

$$\int_{\partial\Omega^+} Pdydz + Qdzdy + Rdxdy = \int_\Omega d(Pdydz + Qdzdy + Rdxdy),$$

其中

$$\begin{aligned} & d(Pdydz + Qdzdy + Rdxdy) \\ =& (P_x dx + P_y dy + P_z dz)dydz + etc. \\ =& (P_x + Q_y + R_z)dxdydz. \end{aligned}$$

(4) Stokes 公式

$$\int_{\partial S^+} Pdx + Qdy + Rdz = \int_S d(Pdx + Qdy + Rdz),$$

其中

$$\begin{aligned} & d(Pdx + Qdy + Rdz) \\ =& (P_x dx + P_y dy + P_z dz)dx + etc. \\ =& (R_y - Q_z)dydz + (P_z - R_x)dzdx + (Q_x - P_y)dxdy. \end{aligned}$$

人们把这个统一的公式写成如下的定理.

定理 3.11.1 (Stokes 定理)**.** *ω 在区域边界上的积分等于 $d\omega$ 在区域上的积分:*

$$\int_{\partial D} \omega = \int_D d\omega.$$

Stokes 定理不但对低维 $(1, 2, 3$ 维) 的区域成立, 对一般有限维的区域 D 也成立, 它是微积分中最核心的结果.

定理 3.11.2 (Stokes 公式)**.** 设定向曲面 S 的边界 ∂S 是分段光滑的, 按照右手法则对 ∂S 规定一个定向, 把取好这个定向的边界记作 ∂S^+. 设 P, Q, R 是 S 上的 C^1 光滑函数, 则有

$$\oint_{\partial S^+} Pdx + Qdy + Rdz = \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy.$$

∂S 的定向规则: 当右手的拇指在 S 的切平面中垂直于 ∂S 指向 S 外面, 食指指向 ∂S^+ 的方向时, 中指恰好指向 S 的法方向 $\mathbf{n}$.

例 3.11.3. 计算

$$\oint_L xydx + yzdy + zxdz,$$

其中 L 为曲线

$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ x + y + z = 1 \end{cases}$$

积分方向从 $(1, 0, 0)$ 指向 $(0, 1, 0)$.

3.12 场论初步

在物理中我们经常碰到各种场, 比如说引力场, 电磁场; 在数学上我们可以给出如下定义.

定义 3.12.1. 区域 $V \subset \mathbf{R}^3$ 上的一个标量场是指一个光滑映射 $f : V \to \mathbf{R}$; 区域 V 上的一个矢量场是指一个光滑映射 $\mathbf{F} : V \to \mathbf{R}^3$.

注 3.12.2. 称 f 是 C^1 光滑的, 如果 f 的偏导数都连续; 称 f 是 C^2 光滑的, 如果 f 的 2 阶偏导数都连续;...... 以此类推, 称 f 是 (C^∞) 光滑的, 如果 f 各阶偏导数都连续.

如果场随时间演化, 可以分别描述成光滑映射

$$V \times \mathbf{R} \to \mathbf{R},$$

$$V \times \mathbf{R} \to \mathbf{R}^3,$$

其中定义域中的因子 $\mathbf{R}$ 表示时间轴.

3.12.1 标量场

空间 X 上的标量场就是光滑函数 $f : X \to \mathbf{R}$.

例 3.12.3. 怎么描绘地球的表面? 测量每点处的海拔, 得到函数 $h : S^2 \to \mathbf{R}$, 可以看成 S^2 上的一个标量场.

对任何 $c \in \mathbf{R}$, 所有场强为 c 的点构成一个等值面/等高面/等势面

$$X_c = f^{-1}(c),$$

由所有的等高面 $\{X_c\}_{c\in\mathbf{R}}$ 可以重构出 X.

标量场 f 的梯度场定义为

$$\mathrm{grad} f = \nabla f := (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}),$$

它是等高面的法方向, 处处 "与等高面垂直". 事实上, 设

$$\gamma : \mathbf{R} \to X_c,$$

$$\gamma(t) = (x(t), y(t), z(t)),$$

是等高面 X_c 中的一条道路, 则 γ 的方向为 $\gamma'(t) = (x'(t), y'(t), z'(t))$. 由于

$$f((x(t), y(t), z(t)) \equiv c,$$

两边对 t 求导得

$$f_x(\gamma(t))x'(t) + f_y(\gamma(t))y'(t) + f_z(\gamma(t))z'(t) = 0,$$

这表明 $(\nabla f)(\gamma(t))$ 垂直于 $\gamma'(t)$. 由 γ 的任意性可知梯度方向处处 "垂直于等高面".

3.12.2 矢量场的散度 (divergence)

矢量场 $\mathbf{F} = (P, Q, R)$ 通过定向曲面 S 的通量为

$$\iint_S \mathbf{F} \cdot \mathbf{n} dS = \iint_S P dydz + Q dzdx + R dxdy.$$

由 Gauss 公式, $\mathbf{F}$ 通过封闭曲面 ∂V^+ 的通量为

$$\iint_{\partial V^+} \mathbf{F} \cdot \mathbf{n} dS = \iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dxdydz.$$

定义 3.12.4. 定义向量场 $\mathbf{F} = (P, Q, R)$ 的散度为

$$\mathrm{div}\mathbf{F} = \nabla \cdot \mathbf{F} := \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

这样, Gauss 公式可以写成

$$\iint_{\partial V^+} \mathbf{F} \cdot \mathbf{n} dS = \iiint_V \nabla \cdot \mathbf{F} dxdydz.$$

例 3.12.5. 设 $\mathbf{R}^3$ 中每点处的电荷密度为 ρ, 则电场强度为

$$\mathbf{E}(x,y,z)=\iiint_{\mathbf{R}^3}\rho(x',y',z')\frac{(x-x',y-y',z-z')}{((x-x')^2+(y-y')^2+(z-z')^2)^{3/2}}dx'dy'dz'.$$

一方面, 如果 ∂V 上 ρ 处处为 0, 则 $\mathbf{E}$ 的通量为

$$\begin{aligned}&\iint_{\partial V}\mathbf{E}\cdot d\mathbf{S}\\ =&\iint_{\partial V}dS\iiint_{\mathbf{R}^3}\rho(x',y',z')\frac{(x-x',y-y',z-z')\cdot\mathbf{n}(x,y,z)}{((x-x')^2+(y-y')^2+(z-z')^2)^{3/2}}dx'dy'dz'\\ =&\iiint_{\mathbf{R}^3}\rho(x',y',z')dx'dy'dz'\iint_{\partial V}\frac{(x-x',y-y',z-z')\cdot\mathbf{n}(x,y,z)}{((x-x')^2+(y-y')^2+(z-z')^2)^{3/2}}dS\\ =&\iiint_V 4\pi\rho(x',y',z')dx'dy'dz',\end{aligned}$$

这里我们用到了

$$\iint_{\partial V}\frac{(x-x',y-y',z-z')\cdot\mathbf{n}(x,y,z)}{((x-x')^2+(y-y')^2+(z-z')^2)^{3/2}}dS=\begin{cases}4\pi, & \text{if } (x',y',z')\in V,\\ 0, & \text{if } (x',y',z')\notin V.\end{cases}$$

另一方面, 由 Gauss 公式, 有

$$\iint_{\partial V^+}\mathbf{E}\cdot\mathbf{n}dS=\iiint_V(\nabla\cdot\mathbf{E})dxdydz.$$

结合这两方面可知, 电场满足方程

$$\nabla\cdot\mathbf{E}=4\pi\rho.$$

例 3.12.6. 磁场的散度为

$$\nabla\cdot\mathbf{B}=0.$$

3.12.3 矢量场的旋度 (curl)

矢量场 $\mathbf{F}$ 沿定向曲线 L 的积分为

$$\int_L\mathbf{F}\cdot d\mathbf{l}=\int_L(Pdx+Qdy+Rdz).$$

由 Stokes 公式, $\mathbf{F}$ 沿封闭曲线 ∂S^+ 的积分为

$$\begin{aligned}&\int_{\partial S^+}\mathbf{F}\cdot d\mathbf{l}\\ =&\iint_S(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z})dydz+(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x})dzdx+(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y})dxdy\\ =&\iint_S(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z},\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x},\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y})\cdot d\mathbf{S}.\end{aligned}$$

定义 3.12.7. 定义向量场 $\mathbf{F}=(P,Q,R)$ 的旋度为

$$\mathrm{rot}\mathbf{F}=\mathrm{curl}(\mathbf{F})=\nabla\times\mathbf{F}:=(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z},\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x},\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}).$$

这样, Stokes 公式可写成

$$\int_{\partial S^+}\mathbf{F}\cdot d\mathbf{l}=\iint_S\nabla\times\mathbf{F}\cdot d\mathbf{S}.$$

例 3.12.8. 由法拉第定律, 感生电流的电势等于磁通量的变化率, 即

$$-\oint_{\partial S}\mathbf{E}\cdot d\mathbf{r}=\frac{\partial}{\partial t}\iint_S\mathbf{B}\cdot d\mathbf{S},$$

利用 Stokes 公式可得

$$-\iint_S\nabla\times\mathbf{E}\cdot d\mathbf{S}=\frac{\partial}{\partial t}\iint_S\mathbf{B}\cdot d\mathbf{S}.$$

这样, 就得到 Maxwell 方程组中的一个

$$\nabla\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}=0.$$

3.13 Poincare 引理

本节中假设所有标量场与矢量场都是 C^∞ 光滑的, 即它们的各个高阶偏导数都存在且连续.

命题 3.13.1. (1) *如果* $\mathbf{F}=\nabla f$, *则* $\nabla\times\mathbf{F}\equiv 0$.

(2) *如果* $\mathbf{B}=\nabla\times\mathbf{A}$, *则* $\nabla\cdot\mathbf{B}\equiv 0$.

问题 3.13.2. *上述命题的逆命题是否成立?*

例 3.13.3 (电磁势)**.** 真空中的 Maxwell 方程为

$$\begin{cases}\nabla\cdot\mathbf{E}=0,\\ \nabla\times\mathbf{E}=-\dfrac{\partial\mathbf{B}}{\partial t},\\ \nabla\cdot\mathbf{B}=0,\\ \nabla\times\mathbf{B}=\dfrac{\partial\mathbf{E}}{\partial t}.\end{cases}$$

如果上述两个命题的逆命题都成立, 则由 $\nabla\cdot\mathbf{B}=0$ 知存在 $\mathbf{A}$ 使得

$$\mathbf{B}=\nabla\times\mathbf{A},$$

代入第二个方程得

$$\nabla \times (\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0,$$

所以存在 ϕ, 使得

$$-\nabla \phi = \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}.$$

电磁学中称 $\phi, \mathbf{A}$ 为电磁场 $\mathbf{E}, \mathbf{B}$ 的电磁势.

例 3.13.4 (用微分形式表示 Maxwell 方程)**.** 设 $\mathbf{E}, \mathbf{B}$ 是电磁场, $\phi, \mathbf{A}$ 为电磁势. 令

$$\alpha = \phi dt + A_x dx + A_y dy + A_z dz$$

为 $\mathbf{R}^4$ 中的 1-形式, 其外微分为

$$\begin{aligned}
d\alpha &= \left(\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t}\right) dxdt + \left(\frac{\partial \phi}{\partial y} - \frac{\partial A_y}{\partial t}\right) dydt + \left(\frac{\partial \phi}{\partial z} - \frac{\partial A_z}{\partial t}\right) dzdt \\
&\quad + \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) dydz + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) dzdx + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dxdy \\
&= E_x dxdt + E_y dydt + E_z dzdt + B_x dydz + B_y dzdx + B_z dxdy,
\end{aligned}$$

它的六个系数恰好是电磁场 $\mathbf{E}, \mathbf{B}$ 的各个分量.

3.13.1

定义 3.13.5. 称矢量场 $\mathbf{F}$ 为保守场, 如果 $\mathbf{F}$ 对质点所做的功只与道路的起点终点有关, 与路径无关.

命题 3.13.6. *如下命题彼此等价.*

(1) $\mathbf{F}$ *是保守场.*

(2) $\mathbf{F}$ *对质点沿任何封闭道路 C 所做的功都等于 0,*

$$\oint_C \mathbf{F} \cdot d\mathbf{l} = 0.$$

(3) $\mathbf{F}$ *是有势场, 即存在光滑函数 ϕ 使得 $\mathbf{F} = \nabla \phi$.*

注 3.13.7. 在物理学中, 称满足条件 $\mathbf{F} = \nabla \phi$ 的 ϕ 为 $\mathbf{F}$ 的势; 在数学中, 称满足条件 $d\phi = Pdx + Qdy + Rdz$ 的 ϕ 为 $Pdx + Qdy + Rdz$ 的原函数. 它可由如下方法求出:

$$\phi(x, y, z) = \phi(x_0, y_0, z_0) + \int_L Pdx + Qdy + Rdz,$$

其中 (x_0, y_0, z_0) 是任意选定的基点, L 是从 (x_0, y_0, z_0) 到 (x, y, z) 的任何光滑曲线.

定理 3.13.8 (Poincare 引理的 1-维部分)**.** 设 X 是 $\mathbf{R}^3$ 的开集且没有 1-维的 "洞", 即对 X 中的任何封闭曲线 C, 存在曲面 $S \subseteq X$ 使得 $\partial S = X$. 设 $\mathbf{F}$ 是 X 上的光滑矢量场, 则

$$\mathbf{F}\text{是保守场} \Leftrightarrow \nabla \times \mathbf{F} \equiv 0.$$

注 3.13.9. 特别的, $\mathbf{R}^3$ 上的矢量场是保守场当且仅当它是无旋的.

例 3.13.10. 考虑 $X = \mathbf{R}^3$ 的情形. 假设光滑矢量场 $\mathbf{F} = (P, Q, R)$ 满足 $\nabla \times \mathbf{F} \equiv 0$. 定义

$$\phi(x, y, z) = x \int_0^1 P(tx, ty, tz)dt + y \int_0^1 Q(tx, ty, tz)dt + z \int_0^1 R(tx, ty, tz)dt.$$

容易验证 $\nabla\phi = \mathbf{F}$.

例 3.13.11. 令 $Y = \{(x, y, z) \in \mathbf{R}^3 | x^2 + y^2 \neq 0\}$, 则 $\mathbf{F} = (\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0)$ 在 Y 上是无旋的, 但不是 Y 上的保守场. 注意到, Y 有 1-维的 "洞", 因为 Y 中的封闭曲线 $C = \{(x, y, 0) | x^2 + y^2 = 1\}$, 它不是 Y 中任何曲面的边界, 所以这个例子并不与 Poincare 引理矛盾.

上述例子来自于电磁学. 事实上, 有如下一般的例子.

例 3.13.12. 设 $C \subseteq \mathbf{R}^3$ 是定向的闭曲线, 里面有单位电流. 由 Biot-Savart 定律, C 产生的磁场为

$$\mathbf{B}(x, y, z) = \oint_C \frac{d\mathbf{l}' \times (x - x', y - y', z - z')}{((x - x')^2 + (y - y')^2 + (z - z')^2)^{3/2}}, \quad \forall (x, y, z) \notin C.$$

直接计算, 可得 $\mathbf{B}$ 的旋度为

$$(\nabla \times \mathbf{B})(x, y, z) = 0, \quad \forall (x, y, z) \notin C.$$

例 3.13.13 (linking number)**.** 在历史上, Gauss, Maxwell 仔细研究了上述例子, 设质点在磁场 $\mathbf{B}$ 的作用下沿定向曲线 C' 运动 (C' 与 C 不相交), 则磁场所做的功为

$$lk(C, C') = \oint_{C'} \oint_C \frac{d\mathbf{l}' \times (x - x', y - y', z - z') \cdot d\mathbf{l}}{((x - x')^2 + (y - y')^2 + (z - z')^2)^{3/2}},$$

拓扑学中称 $lk(C, C')$ 为 C 与 C' 的 *linking number*(环绕数), 是 C 绕 C' 的圈数.

3.13.2

定理 3.13.14 (Poincare 引理的 2-维部分)**.** 设 X 是 $\mathbf{R}^3$ 的开集且没有 2-维的 "洞", 即对 X 中的任何封闭曲面 S, 存在 $V \subseteq X$ 使得 $\partial V = S$. 设 $\mathbf{B}$ 是 X 上的光滑矢量场, 则如下条件彼此等价.

(1) 对任何两个定向曲面 S_1, S_2, 只要 $\partial S_1^+ = \partial S_2^+$, 则有

$$\iint_{S_1} \mathbf{B} \cdot d\mathbf{S} = \iint_{S_2} \mathbf{B} \cdot d\mathbf{S}.$$

(2) 对任何封闭曲面 Σ, 都有

$$\iint_{\Sigma} \mathbf{B} \cdot d\mathbf{S} = 0.$$

(3) $\nabla \cdot \mathbf{B} \equiv 0$.

(4) 存在光滑矢量场 $\mathbf{A}$ 使得 $\nabla \times \mathbf{A} = \mathbf{B}$.

例 3.13.15. 考虑 $X = \mathbf{R}^3$ 的情形. 假设光滑矢量场 $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3)$ 满足 $\nabla \cdot \mathbf{B} \equiv 0$. 定义

$$\mathbf{A}_1(x,y,z) = z\int_0^1 \mathbf{B}_2(tx,ty,tz)tdt - y\int_0^1 \mathbf{B}_3(tx,ty,tz)tdt,$$

$$\mathbf{A}_2(x,y,z) = x\int_0^1 \mathbf{B}_3(tx,ty,tz)tdt - z\int_0^1 \mathbf{B}_1(tx,ty,tz)tdt,$$

$$\mathbf{A}_3(x,y,z) = y\int_0^1 \mathbf{B}_1(tx,ty,tz)tdt - x\int_0^1 \mathbf{B}_2(tx,ty,tz)tdt,$$

容易验证 $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3)$ 满足

$$\nabla \times \mathbf{A} = \mathbf{B}.$$

3.14 de Rham 理论简介

在日常生活中, 人们经常会讨论一个几何物体有没有 "洞", 或者有几个 "洞". 如何用数学语言描述几何物体的这种性质呢?

3.14.1

假设我们生活在空间 X 中. 为了简单起见, 先考虑 1 维的 "洞". 首先, 什么是 1 维的 "洞"? 由于我们生活在 X 中, 只能用 X 中的东西来描述. 人们只能看到或探测到洞的边界, 它是 X 中的简单闭曲线. 任给简单闭曲线 $L \subset X$, 怎么判断 L 是否是某个 "洞" 的边界?

(1) 如果能找到 X 中的二维区域 D 使得 $\partial D = L$, 则可断定 L 不是 "洞" 的边界;

(2) 如果不存在这样的 D, 则 L 是某个 "洞" 的边界.

对于给定的 L, 要判断是否存在 $D \subset X$ 使得 $\partial D = L$, 这通常是非常困难的问题. Stokes 定理提供了一个好的办法.

取三个光滑函数 P, Q, R, 考虑曲线积分 $\oint_L Pdx + Qdy + Rdz$. 如果存在 D 使得 $\partial D = L$, 则由 Stokes 定理, 有:

$$\oint_L Pdx + Qdy + Rdz = \iint_D d(Pdx + Qdy + Rdz).$$

如果进一步还有 $d(Pdx + Qdy + Rdz) \equiv 0$, 则可得

$$\oint_L Pdx + Qdy + Rdz = 0.$$

这样, 如果我们找到某组 P, Q, R 满足 $d(Pdx + Qdy + Rdz) \equiv 0$ 且 $\oint_L Pdx + Qdy + Rdz \neq 0$, 则可证明 L 不是 X 中任何二维区域的边界, 从而说明 L 是某个 "洞" 的边界.

总结一下, 我们称满足条件

$$d(Pdx + Qdy + Rdz) \equiv 0$$

的 $Pdx + Qdy + Rdz$ 为一个 "1 维洞的探测器"(数学中称之为 1 维 cocycle). 所有 "1 维洞的探测器" 构成集合

$$Z^1(X) = \{Pdx + Qdy + Rdz | d(Pdx + Qdy + Rdz) \equiv 0\}.$$

为了验证简单闭曲线 $L \subset X$ 是否是 "洞" 的边界, 只要对每个 "探测器"$Pdx + Qdy + Rdz \in Z^1(X)$ 计算积分

$$\oint_L Pdx + Qdy + Rdz,$$

如果发现所有这些积分中有一个非 0, 则 L 一定是 "洞" 的边界.

例 3.14.1. 设 $X = \mathbf{R}^2 - \{(0,0)\}$. 利用上述办法, 可以证明单位圆周 C 代表一个 1 维 "洞" 的边界. 事实上, 考虑 cocycle

$$\frac{-ydx + xdy}{x^2 + y^2} \in Z^1(X),$$

以前我们计算过

$$\oint_{C^+} \frac{-ydx + xdy}{x^2 + y^2} = 2\pi,$$

则 Stokes 定理表明 C 一定是某个 "洞" 的边界.

细心的同学会发现, 如果空间 X 的 1 维 cocycle 有无数个 (一般来说, 的确如此), 则上述判断方法基本上是不可用的, 因为我们不可能把每个 cocycle 对应的积分值都算出来. 解决这

个问题的关键是注意到, 虽然一般情况下都有无限多个 cocylcle, 但很多 cocycle 的用途是完全一样的. 比如, 任取一个函数 W, 它给出一个 cocycle:

$$dW = \frac{\partial W}{\partial x}dx + \frac{\partial W}{\partial y}dy + \frac{\partial W}{\partial z}dz,$$

由 Newton-Leibniz 公式, 这个 cocycle 在任何 (定向的) 简单闭曲线 L 上的积分为

$$\oint_L \frac{\partial W}{\partial x}dx + \frac{\partial W}{\partial y}dy + \frac{\partial W}{\partial z}dz = 0.$$

这表明 dW 是一个无效的 "探测器"! 人们把这种形式的 cocycle 称为 coboundary, 记

$$B^1(X) = \{dW | W是X上的光滑函数\}.$$

进一步, 对任何两个 cocycle:

$$Pdx + Qdy + Rdz, \quad P'dx + Q'dy + R'dz \in Z^1(X),$$

如果它们的差为某个 coboundary

$$(Pdx + Qdy + Rdz) - (P'dx + Q'dy + R'dz) = dW$$

则它们是完全等效的 "探测器". 因为, 对任何 (定向的) 简单闭曲线 L, 总有

$$\oint_L Pdx + Qdy + Rdz = \oint_L P'dx + Q'dy + R'dz,$$

基于这个原因, 我们可以把所有的 cocycle 分类, 定义两个 cocycle 属于同一类, 如果它们的差为某个 coboundary, 即

$$Pdx+Qdy+Rdz \sim P'dx+Q'dy+R'dz \iff \exists W, (Pdx+Qdy+Rdz)-(P'dx+Q'dy+R'dz) = dW$$

这样, 在验证某个闭曲线 L 是否代表 "洞" 的边界时, 并不需要对所有的 cocycle 计算积分, 而只需要在每一类中选一个 cocycle 计算积分就行. 令

$$H^1(X) = \{所有cocycle的等价类\} = \frac{Z^1(X)}{B^1(X)},$$

为 X 上所有 (1 维)cocycle 类构成的集合. 人们发现, 对稍微好一点的空间, $H^1(X)$ 是有限维的线性空间. 这样, 如果要判断 L 是否代表 "洞" 的边界, 只需要计算有限几个积分.

上面的办法是 de Rham 发现的, $H^1(X)$ 称为 X 的 1 维 de Rham 上同调群. 粗略的说, 它是 X 上所有 "1 维洞的探测器"(的等价类) 的集合.

3.14.2

类似的, 如何探测 X 中的 2 维 "洞" 呢? 人们只能看到 "洞" 的边界, 它是 X 中的封闭曲面 S.

(1) 如果能找到 X 中的 3 维区域 Ω 使得 $\partial\Omega = S$, 则可断定 S 不是 "洞" 的边界;

(2) 如果不存在这样的 Ω, 则 S 是某个 "洞" 的边界.

为了说明 S 的确是 "洞" 的边界, 由 Stokes 定理可尝试计算积分

$$\iint_S Pdydz + Qdzdx + Rdxdy,$$

其中 $d(Pdydz + Qdzdx + Rdxdy) \equiv 0$. 称满足这个条件的 $Pdydz + Qdzdx + Rdxdy$ 为 X 上的 2 维 cocyle, 记

$$Z^2(X) = \{Pdydz + Qdzdx + Rdxdy | d(Pdydz + Qdzdx + Rdxdy) \equiv 0\}.$$

同样, 注意到有很多 cocycle 实际上是无效的 "探测器", 令

$$B^2(X) = \{d(Adx + Bdy + Cdz) | A, B, C\text{是光滑函数}\}$$

为 2 维 coboundary 的集合, 则真正有效的 "探测器" 的集合为

$$H^2(X) = \frac{Z^2(X)}{B^2(X)},$$

$H^2(X)$ 称为 X 的 2 维 de Rham 上同调群. 粗略的说, 它是 X 上所有 "2 维洞的探测器"(的等价类) 的集合.

例 3.14.2. 设 $X = \mathbf{R}^3 - \{(0,0,0)\}$. 利用上述办法, 可以证明单位球面 S 代表一个 2 维 "洞" 的边界. 事实上, 考虑 cocycle

$$\frac{xdydz + ydzdx + zdxdy}{(x^2+y^2+z^2)^{3/2}} \in Z^2(X),$$

以前我们计算过

$$\iint_{S^+} \frac{xdydz + ydzdx + zdxdy}{(x^2+y^2+z^2)^{3/2}} = 4\pi,$$

Gauss 公式/Stokes 定理表明 S 代表某个 2 维 "洞" 的边界.

第四章　常微分方程初步

4.1　常微分方程

物理规律是用微分方程描述的, 比如 Newton 力学定律,

$$m\frac{d^2\mathbf{r}(t)}{dt^2} = \mathbf{F};$$

Maxwell 方程

$$\begin{cases} \nabla \cdot \mathbf{E} = 0, \\ \nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} = 0, \\ \nabla \times \mathbf{B} = \dfrac{\partial \mathbf{E}}{\partial t}; \end{cases}$$

Schrodinger 方程

$$ih\frac{\partial \psi}{\partial t} = -\frac{h^2}{2m}\nabla^2\psi.$$

Newton 方程只涉及函数的各阶导函数, Maxwell 方程, Schrodinger 方程涉及函数的各阶偏导数. 我们把前者称为常微分方程, 把后者称为偏微分方程.

定义 4.1.1 (n 阶 ODE)**.** 给定映射 $F : \mathbf{R}^{n+2} \to \mathbf{R}$, 求函数 $y : I \to \mathbf{R}$, 使得

$$F(x, y(x), y^{(1)}(x), ..., y^{(n)}(x)) = 0, \quad \forall x \in I,$$

人们把这种问题称为区间 I 上的 n 阶常微分方程.

例 4.1.2. n 阶常微分方程 $y^{(n)}(x) = 0$ 有很多解. 对任何常数 $C_0, ..., C_{n-1}$, 函数

$$\phi(x) = C_0 + C_1 x + ... + C_{n-1}x^{n-1}$$

都是微分方程的解. 实际上, 可以证明这就是全部解.

找出物理规律后, 如果已知系统的初始状态, 则可以完全了解系统的演化过程. 这相当于求解微分方程

$$F(x,y,y^{(1)},...,y^{(n)})=0$$

且满足初值条件

$$\begin{cases} y(x_0)=y_0 \\ y^{(1)}(x_0)=y_1 \\ ... \\ y^{(n-1)}(x_0)=y_{n-1} \end{cases}$$

这里 $y_0,...,y_{n-1}$ 是事先给定的常数. 人们把这种求解问题称为初值问题, 或者 Cauchy 问题.

定义 4.1.3 (初值问题)**.** 给定映射 $F:\mathbf{R}^{n+2}\to\mathbf{R}$ 与 $x_0;y_0,y_1,...,y_{n-1}$, 求函数 $y:[x_0,x_1]\to\mathbf{R}$, 使得

$$\begin{cases} F(x,y(x),y^{(1)}(x),...,y^{(n)}(x))=0, \quad \forall x\in[x_0,x_1]; \\ y(x_0)=y_0, \\ y^{(1)}(x_0)=y_1, \\ ... \\ y^{(n-1)}(x_0)=y_{n-1}. \end{cases}$$

人们把这种问题称为初值问题.

4.2 微分方程的初等解法

4.2.1 分离变量法

形如

$$\frac{dy}{dx}=f(x)g(y)$$

的微分方程称为变量分离的, 因为可以把变量 x,y 分开写在等式两边:

$$\frac{dy}{g(y)}=f(x)dx. \tag{4.1}$$

这种微分方程的解法如下. 设 $\frac{1}{g}$ 有原函数 G, 则

$$\frac{dG(y(x))}{dx}=G'(y(x))y'(x)=\frac{1}{g(y(x))}y'(x)=f(x),$$

这表明 $G(y(x))$ 是 $f(x)$ 的一个原函数, 即存在常数 C 使得:

$$G(y(x)) = \int f(x)dx + C.$$

可以形式化的看成对 (1) 两边积分,

$$\int \frac{dy}{g(y)} = \int f(x)dx + C.$$

例 4.2.1.

$$y' - ky = 0.$$

例 4.2.2. 设 a, b, c 为常数, 求解

$$\frac{dy}{dx} = f(ax + by + c).$$

例 4.2.3. 求解微分方程

$$\frac{dy}{dx} = f(x, y),$$

其中 $f(x, y)$ 为 0 次齐次函数, 即对任何 λ, 有 $f(\lambda x, \lambda y) = \lambda^0 f(x, y)$.

例 4.2.4. 设 $a_i, b_i, c_i, i = 1, 2$ 为常数, 求解

$$\frac{dy}{dx} = f(\frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}).$$

4.2.2 一阶线性微分方程

$$\frac{dy}{dx} + P(x)y = Q(x),$$

其中 $P(x), Q(x)$ 是给定的连续函数. 上述方程左边关于 $y, \frac{dy}{dx}$ 都是线性的.

设 $y(x) = e^{-\int_{x_0}^{x} P(t)dt}u(x)$, 则 u 满足微分方程:

$$u'(x) = Q(x)e^{\int_{x_0}^{x} P(t)dt},$$

解得

$$u(x) = \int_{x_0}^{x} Q(s)e^{\int_{x_0}^{s} P(t)dt}ds + C.$$

例 4.2.5 (伯努利方程)**.**

$$\frac{dy}{dx} + p(x)y = q(x)y^n.$$

4.2.3 全微分方程与积分因子

考虑微分方程

$$\frac{dy}{dx} = -\frac{P(x,y)}{Q(x,y)}.$$

(1) 如果 $Pdx + Qdy$ 是全微分, 即存在二元函数 $u(x,y)$ 使得 $Pdx + Qdy = du$, 或等价的, $P = \frac{\partial u}{\partial x}, \quad Q = \frac{\partial u}{\partial y}$, 则可解出微分方程

$$P(x,y(x)) + Q(x,y(x))\frac{dy}{dx} = 0 \iff \frac{du(x,y(x))}{dx} = 0 \iff u(x,y(x)) = C.$$

问题 4.2.6. *如何判断 $Pdx + Qdy$ 是否为全微分?*

命题 4.2.7. *设平面区域 D 是拓扑圆盘, $P(x,y), Q(x,y)$ 是 D 上的光滑函数, 则存在光滑函数 u 使得 $Pdx + Qdy = du$ 的充分必要条件是*

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \equiv 0.$$

进一步, 原函数 u 由下式给出

$$u(x,y) = u(x_0,y_0) + \int_L Pdx + Qdy,$$

其中 L 是从 (x_0,y_0) 到 (x,y) 的任意光滑道路.

(2) 如果 $Pdx + Qdy$ 不是全微分, 但存在函数 $\mu(x,y) \neq 0$ 使得 $\mu \cdot (Pdx + Qdy) = du$ 是全微分, 则称 $\mu(x,y)$ 是该微分方程的积分因子. 这种情形下同样可得

$$u(x,y(x)) = C.$$

问题 4.2.8. *积分因子是否存在? 如何求积分因子?*

拓扑圆盘 D 上, μ 是积分因子的充分必要条件是

$$\frac{\partial(\mu Q)}{\partial x} - \frac{\partial(\mu P)}{\partial y} \equiv 0,$$

即

$$\frac{\partial \mu}{\partial x}Q - \frac{\partial \mu}{\partial y}P = \mu \cdot (\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}),$$

这个方程一般不容易解, 我们只看特殊情况. 若 μ 只是 x 的函数, 则有

$$\frac{1}{\mu}\frac{d\mu}{dx} = \frac{\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}}{Q},$$

特别的, 这要求 $\frac{\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}}{Q}$ 与 y 无关.

反之, 假如 $\frac{\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}}{Q}=R(x)$ 只是 x 的函数, 则可解出积分因子

$$\mu(x)=C_0exp\left(\int_{x_0}^{x}R(t)dt\right).$$

命题 4.2.9. 如果 $\frac{\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}}{Q}=R(x)$ 是 x 的函数, 则有积分因子

$$\mu(x)=C_0exp\left(\int_{x_0}^{x}R(t)dt\right).$$

类似的, 如果 $\frac{\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}}{-P}=T(y)$ 是 y 的函数, 则有积分因子

$$\lambda(y)=C_0exp\left(\int_{y_0}^{y}T(z)dz\right).$$

例 4.2.10. 解微分方程 $-(1-x^2+y)dx+xdy=0$.

4.2.4 降阶

处理高阶微分方程的一个办法是降阶: 转化成阶数更小的微分方程.

(1) 不显含 y 的微分方程, 例如:

$$F(x,y',y'')=0.$$

令 $z(x)=y'(x)$, 则 z 满足一阶微分方程

$$F(x,z,z')=0.$$

设由此解出 z 满足普通方程

$$\Phi(x,z;C)=0,$$

则 y 满足一阶微分方程

$$\Phi(x,y';C)=0.$$

(2) 不显含自变量 x 的方程, 例如:

$$F(y,y',y'')=0.$$

设 $y=y(x)$ 满足上述微分方程, 则在 $y'(x)\neq 0$ 的区域上, 它有反函数 $x=x(y)$. 令 $p(y)=y'(x(y))$, 则有:

$$\frac{dp}{dy}=y''(x(y))\frac{dx}{dy}=y''(x(y))\frac{1}{y'(x(y))},$$

由此得 $y''(x(y)) = p(y)\frac{dp}{dy}$. 代回原方程, 可知 p 满足微分方程:

$$F(y, p, p\frac{dp}{dy}) = 0.$$

这是一阶微分方程, 设由此解出 p 满足普通方程

$$\Phi(y, p; C) = 0,$$

则 $y = y(x)$ 满足一阶微分方程

$$\Phi(y, y'; C) = 0.$$

例 4.2.11. 求解初值问题

$$\begin{cases} 1 + (y')^2 = 2yy'' \\ y(1) = 1, \quad y'(1) = -1 \end{cases}$$

解. 令 $p = y'$, 视为 y 的函数, 则它满足微分方程

$$1 + p^2 = 2yp\frac{dp}{dy}.$$

可分离变量

$$\frac{dy}{y} = \frac{2pdp}{1 + p^2},$$

两边积分得

$$\int \frac{dy}{y} = \int \frac{2pdp}{1 + p^2}.$$

由初值条件 $p(1) = -1$, 可知

$$1 + p(y)^2 = 2y,$$

在 $y = 1$ 附近有

$$p(y) = -\sqrt{2y - 1}.$$

这样, $y = y(x)$ 满足

$$y'(x) = -\sqrt{2y(x) - 1},$$

分离变量后两边积分可得 $y(x) = \frac{x^2 - 4x + 5}{2}$.

□

4.3 一阶微分方程初值问题解的存在唯一性

问题 4.3.1. *一阶微分方程的初值问题*

$$\begin{cases} y' = f(x,y) \\ y(x_0) = y_0 \end{cases}$$

是否有解? 解是否唯一?(其中 f 是连续函数)

首先注意到上述初值问题等价于积分方程

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.$$

直接解这个积分方程是困难的, 但可以得到一列近似解:

(0) 常值函数

$$\phi_0(x) \equiv y_0$$

是积分方程的 0 阶近似解.

(1) 代回上述积分方程, 得 1 阶近似解:

$$\phi_1(x) = y_0 + \int_{x_0}^{x} f(t, \phi_0(t))dt,$$

(2) 继续迭代得 2 阶近似解:

$$\phi_2(x) = y_0 + \int_{x_0}^{x} f(t, \phi_1(t))dt,$$

....

(n) 以此类推, 得 n 阶近似解:

$$\phi_n(x) = y_0 + \int_{x_0}^{x} f(t, \phi_{n-1}(t))dt,$$

这样得到一列越来越精确的近似解 $\{\phi_n\}_{n=1}^{\infty}$, 称为 Picard 序列. 如果 Picard 序列收敛到某个极限

$$\lim_{n\to\infty} \phi_n = \phi,$$

则可 ϕ 是上述积分方程的精确解.

定理 4.3.2 (Picard-Lindelof 定理)**.** 如果函数 $f(x,y)$ 在闭区域 $[x_0-a,x_0+a]\times[y_0-b,y_0+b]$ 上连续, 且满足 *Lipschitz* 条件, 即存在常数 L 使得

$$|f(x,y_1)-f(x,y_2)|\le L|y_1-y_2|,\quad \forall x,y_1,y_2,$$

则初值问题

$$\begin{cases} y'=f(x,y) \\ y(x_0)=y_0 \end{cases}$$

在区间 $[x_0-h,x_0+h]$ 上有唯一解, 其中

$$M=\max\{|f(x,y)|:(x,y)\in[x_0-a,x_0+a]\times[y_0-b,y_0+b]\},\quad h=\min\{a,\frac{b}{M},\frac{1}{2L}\}.$$

推论 4.3.3. 如果 $f(x,y)$ 是 C^1 光滑函数, 则它在有界闭区域上满足 *Lipschitz* 条件. 从而对任何 (x_0,y_0), 在它的某个邻域内初值问题有唯一解.

一般的, 人们把这种通过不停迭代找不动点的方法称为:

定理 4.3.4 (压缩映像定理)**.** 设 (X,d) 是完备的度量空间, d 是度量. 设映射 $T:X\to X$ 是压缩的, 即存在常数 $0<k<1$ 使得 $\forall x,y\in X$, 总有

$$d(Tx,Ty)\le kd(x,y),$$

则 T 有唯一的不动点.

解释: 为什么需要 f 满足 Lipschitz 条件?

令 $\mathcal{C}$ 为区间 $[x_0-h,x_0+h]$ 上所有连续函数构成的空间. 考虑映射 $T:\mathcal{C}\to\mathcal{C}$,

$$(Ty)(x)=y_0+\int_{x_0}^{x}f(t,y(t))dt,$$

积分方程的解一一对应成 T 的不动点 $Ty=y$.

$\mathcal{C}$ 上有度量

$$d(\phi,\psi)=\max_{x\in[x_0-h,x_0+h]}|\phi(x)-\psi(x)|.$$

为了利用压缩映像定理, 只要验证 T 是压缩的. 为此需要估计

$$|T\phi(x)-T\psi(x)|=\int_{x_0}^{x}|f(t,\phi(t))-f(t,\psi(t))|dt,$$

如果 f 满足 Lipschitz 条件, 则

$$\int_{x_0}^{x}|f(t,\phi(t))-f(t,\psi(t))|dt\le\int_{x_0}^{x}L|\phi(t)-\psi(t)|dt\le L|x-x_0|d(\phi,\psi)\le Lh\cdot d(\phi,\psi).$$

由此可得

$$d(T\phi, T\psi) \leq Lh \cdot d(\phi, \psi),$$

特别的, 当 $h \leq \frac{1}{2L}$ 时 T 是压缩的.

注 4.3.5. 这里我们简要的说明了初值问题在区间

$$[x_0 - h, x_0 + h], \quad h = \min\{a, \frac{b}{M}, \frac{1}{2L}\}$$

上有唯一解. 实际上, 可以证明在稍微大一点的区间

$$[x_0 - k, x_0 + k], \quad k = \min\{a, \frac{b}{M}\}$$

上该微分方程有唯一解, 这是我们课本上叙述的定理.

`http://en.wikipedia.org/wiki/Picard-Lindelöf_theorem`

例 4.3.6 (不一定有长期解)**.** 设 $f(x, y) = -y^2$, 考虑微分方程的初值问题

$$\begin{cases} y' = -y^2, \\ y(x_0) = y_0, \end{cases}$$

其中 x_0, y_0 是给定的正数. 直接求解, 得

$$y(x) = \frac{1}{x - (x_0 - \frac{1}{y_0})},$$

其定义域是 $(x_0 - \frac{1}{y_0}, +\infty)$. 这个例子表明, 一般的, 我们只能保证初值问题在 x_0 的某个邻域中有解, 并不能保证长期解存在.

4.4 二阶线性微分方程

形如

$$y^{(n)} + p_1(x)y^{(n-1)} + ... + p_{n-1}(x)y^{(1)} + p_n(x)y = f(x)$$

的微分方程称为 n-阶线性微分方程, 因为方程中只出现 $y, y^{(1)}, ..., y^{(n)}$ 的一次项. 如果 $f(x) \equiv 0$, 称之为 n-阶线性齐次方程; 否则称为 n-阶线性非齐次方程.

我们来考虑 2-阶线性微分方程初值问题解的存在唯一性.

定理 4.4.1. *设函数 $p(x), q(x), f(x)$ 在区间 $[a, b]$ 上连续, 则初值问题 (对给定的 $x_0 \in (a, b)$)*

$$\begin{cases} y'' + p(x)y' + q(x)y = f(x) \\ y(x_0) = y_0, y'(x_0) = y_1 \end{cases}$$

在区间 $[a, b]$ 上存在唯一解 $y(x)$.

解释: 化成一阶微分方程

$$\begin{cases} y' = z \\ z' = f(x) - p(x)z - q(x)y \end{cases}$$

把它看成矢量值函数的微分方程

$$\frac{d}{dx}(y,z) = (z, f(x) - p(x)z - q(x)y),$$

再用 Picard-Lindelof 定理的某种推广形式.

类似的, 我们可以得到 n-阶线性微分方程初值问题解的存在唯一性.

定理 4.4.2 (Picard-Lindelof 定理)**.** *如果函数 $f(x,y)$ 在闭区域 $[x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$ 上连续, 且满足 Lipschitz 条件, 即存在常数 L 使得*

$$|f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2|, \quad \forall x, y_1, y_2,$$

则初值问题

$$\begin{cases} y' = f(x,y) \\ y(x_0) = y_0 \end{cases}$$

在区间 $[x_0 - h, x_0 + h]$ 上有唯一解, 其中

$$M = \max\{|f(x,y)| : (x,y) \in [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]\}, \quad h = \min\{a, \frac{b}{M}, \frac{1}{2L}\}.$$

推论 4.4.3. *如果 $f(x,y)$ 是 C^1 光滑函数, 则它在有界闭区域上满足 Lipschitz 条件. 从而对任何 (x_0, y_0), 在它的某个邻域内初值问题有唯一解.*

例 4.4.4 (不一定有长期解)**.** 设 $f(x,y) = -y^2$, 考虑微分方程的初值问题

$$\begin{cases} \quad y' = -y^2, \\ y(x_0) = y_0, \end{cases}$$

其中 x_0, y_0 是给定的正数. 直接求解, 得

$$y(x) = \frac{1}{x - (x_0 - \frac{1}{y_0})},$$

其定义域是 $(x_0 - \frac{1}{y_0}, +\infty)$. 这个例子表明, 一般的, 我们只能保证初值问题在 x_0 的某个邻域中有解, 并不能保证长期解存在.

形如
$$y^{(n)}+p_1(x)y^{(n-1)}+...+p_{n-1}(x)y^{(1)}+p_n(x)y=f(x)$$
的微分方程称为 n-阶线性微分方程, 因为方程中只出现 $y,y^{(1)},...,y^{(n)}$ 的一次项. 如果 $f(x)\equiv 0$, 称之为 n-阶线性齐次方程; 否则称为 n-阶线性非齐次方程.

我们来考虑 2-阶线性微分方程初值问题解的存在唯一性.

定理 4.4.5. *设函数 $p(x),q(x),f(x)$ 在区间 $[a,b]$ 上连续, 则初值问题 (对给定的 $x_0\in(a,b)$)*
$$\begin{cases}y''+p(x)y'+q(x)y=f(x)\\ y(x_0)=y_0,y'(x_0)=y_1\end{cases}$$
在区间 $[a,b]$ 上存在唯一解 $y(x)$.

解释: 化成一阶微分方程
$$\begin{cases}y'=z\\ z'=f(x)-p(x)z-q(x)y\end{cases}$$
把它看成矢量值函数的微分方程
$$\frac{d}{dx}(y,z)=(z,f(x)-p(x)z-q(x)y),$$
再用 Picard-Lindelof 定理的某种推广形式.

类似的, 我们可以得到 n-阶线性微分方程初值问题解的存在唯一性.

4.4.1 二阶线性微分方程解的结构

命题 4.4.6 (解的叠加)**.** *设 $y_i(x)$ 是微分方程*
$$y''+p(x)y'+q(x)y=f_i(x)$$
的解 $(i=1,2)$. 则 $y=y_1+ky_2$ 是微分方程
$$y''+p(x)y'+q(x)y=f_1(x)+kf_2(x)$$
的解.

推论 4.4.7. *设 $y^*(x)$ 是微分方程*
$$y''+p(x)y'+q(x)y=f(x)$$

的一个解, 则该微分方程的任何解 $y(x)$ 都能表示成

$$y = y^* + y_0,$$

其中 y_0 是微分方程

$$y'' + p(x)y' + q(x)y = 0$$

的解.

由推论知, 解线性非齐次微分方程归结成如下两方面:

(1) 找一个特解;

(2) 研究线性齐次微分方程的解.

这两方面都不容易, 我们先考虑后者, 即研究线性齐次微分方程

$$y'' + p(x)y' + q(x)y = 0$$

的解.

由解的叠加性, 如果找到两个解 $\phi_1(x), \phi_2(x)$, 则它们的任何线性组合

$$k_1\phi_1(x) + k_2\phi_2(x)$$

也都是解, 其中 k_1, k_2 是任何常数. 这一族解恰好由两个常数标记, 它们取遍所有解吗? 当然要排除显然的例外: ϕ_1 与 ϕ_2 成比例时, 上述这族解可由一个常数标记.

定义 4.4.8. 设函数 ϕ_1, ϕ_2 定义在区间 $[a,b]$ 上, 如果存在常数 $(k_1, k_2) \neq (0,0)$ 使得

$$k_1\phi_1(x) + k_2\phi_2(x) = 0, \forall x \in [a,b]$$

则称 ϕ_1, ϕ_2 在区间 $[a,b]$ 上线性相关, 否则称它们在区间 $[a,b]$ 上线性无关.

如何判断线性无关?

命题 4.4.9. 设 $y_1(x), y_2(x)$ 是微分方程

$$y'' + p(x)y' + q(x)y = 0$$

在区间 $[a,b]$ 上的两个解. 则如下命题等价:

(1) y_1, y_2 在区间 $[a,b]$ 上线性相关.

(2) 它们的 *Wronski* 行列式

$$W(x) = \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix}$$

在 $[a,b]$ 上恒等于 0.

(3) 它们的 *Wronski* 行列式在某一点的值为 0.

定理 4.4.10 (线性无关解生成所有解)**.** 设 $y_1(x), y_2(x)$ 是微分方程

$$y'' + p(x)y' + q(x)y = 0$$

在区间 $[a,b]$ 上的两个线性无关解, 则该微分方程的任何解 $y(x)$ 都能表示成

$$y = k_1 y_1 + k_2 y_2,$$

其中 k_1, k_2 是适当选择的常数.

4.4.2 常系数线性齐次方程

设 p, q 是常数, 解 2 阶常系数线性齐次方程

$$y'' + py' + qy = 0.$$

根据上述定理, 只需找到两个线性无关解就可以生成所有解. 考虑形如 $y(x) = e^{\lambda x}$ 的解.

$$y = e^{\lambda x} \text{ 满足方程 } \iff \lambda^2 + p\lambda + q = 0$$

称方程 $\lambda^2 + p\lambda + q = 0$ 为该微分方程的特征方程, 特征方程的根称为特征根.

(1) 特征方程有两个不同的实根 λ_1, λ_2, 则微分方程有两个特解

$$y_i = e^{\lambda_i x}, i = 1, 2.$$

它们的 Wronski 矩阵为

$$\begin{pmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{pmatrix}$$

其行列式处处不等于 0, 因此它们是线性无关解. 由前述定理, 微分方程的所有解都可表示成

$$y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}.$$

(2) 特征方程有 2 重实根 λ, 则微分方程有两个特解

$$y_1 = e^{\lambda x}, y_2 = xe^{\lambda x} = \lim_{\lambda_2 \to \lambda} \frac{e^{\lambda x} - e^{\lambda_2 x}}{\lambda - \lambda_2},$$

它们的 Wronski 行列式 $e^{2\lambda x}$ 处处不等于 0. 所以该微分方程的解都可以表示成

$$y(x) = C_1 e^{\lambda x} + C_2 x e^{\lambda x}.$$

(3) 特征方程有两个共轭复根 $\alpha \pm \beta i$, 则微分方程有两个特解

$$y_{\pm} = e^{(\alpha \pm \beta i)x}.$$

它们生成该微分方程的所有复值函数解,

$$y(x) = e^{\alpha x}[C_1 e^{\beta x i} + C_2 e^{-\beta x i}], C_1, C_2 \in \mathbf{C},$$

如果要求实值函数解, 则要求 $C_1 = \bar{C}_2$, 得所有实值函数解

$$y(x) = e^{\alpha x}(a \cos \beta x + b \sin \beta x), a, b \in \mathbf{R}.$$

4.4.3 常数变易法

考虑线性非齐次方程

$$y'' + p(x)y' + q(x)y = f(x).$$

设已经找到线性齐次方程

$$y'' + p(x)y' + q(x)y = 0$$

的两个线性无关解 $\phi_1(x), \phi_2(x)$, 则它们生成 $y'' + p(x)y' + q(x)y = 0$ 的所有解.

下面我们用 ϕ_1, ϕ_2 构造出

$$y'' + p(x)y' + q(x)y = f(x)$$

的一个特解.

命题 4.4.11 (常数变易法)**.** *设 $\phi_1(x), \phi_2(x)$ 是线性齐次方程*

$$y'' + p(x)y' + q(x)y = 0$$

的两个线性无关解. 如果函数 $C_1(x), C_2(x)$ 满足方程:

$$\begin{cases} C_1'(x)\phi_1(x) + C_2'(x)\phi_2(x) = 0 \\ C_1'(x)\phi_1'(x) + C_2'(x)\phi_2'(x) = f(x) \end{cases}$$

则 $C_1(x)\phi_1(x) + C_2(x)\phi_2(x)$ 是线性非齐次方程

$$y'' + p(x)y' + q(x)y = f(x)$$

的解.

由这个命题, 只要求解上述关于 C_1', C_2' 的方程组, 即可得非齐次方程的特解, 再加上齐次方程的一般解, 就给出非齐次方程的所有解. 因为 ϕ_1, ϕ_2 是线性无关的, 它们的 Wronski 矩阵

$$\begin{pmatrix} \phi_1(x) & \phi_2(x) \\ \phi_1'(x) & \phi_2'(x) \end{pmatrix}$$

处处是可逆矩阵, 从而可解出 $C_1'(x), C_2'(x)$.

第五章 Fourier 级数与 Fourier 变换

5.1 Fourier 级数

粗略的说, 所谓的 Fourier 变换就是把一般的函数表示成具有某种特定对称性的函数的线性叠加.

例 5.1.1. 任何函数 $f:\mathbf{R}\to\mathbf{R}$ 可以表示成一个奇函数与一个偶函数的和.

例 5.1.2. 给定正整数 n. 复平面具有旋转 $\frac{2\pi}{n}$ 的对称性.

5.1.1 周期函数

称 $f:\mathbf{R}\to\mathbf{R}$ 是以 T 为周期的函数, 如果

$$f(x+T)=f(x),\quad \forall x\in\mathbf{R}.$$

例 5.1.3. 以下函数是以 T 为周期的函数:

$$1,\quad \sin(\frac{2n\pi x}{T}),\quad \cos(\frac{2n\pi x}{T}),\quad (n=1,2,...)$$

对周期为 T 的函数 f,g, 定义

$$\langle f,g\rangle=\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)g(x)dx.$$

即在一个周期内积分 f 与 g 的乘积.

例 5.1.4. 对正整数 m, n, 有

$$\begin{aligned}
&\int_{-\frac{T}{2}}^{\frac{T}{2}} \sin(\frac{2m\pi x}{T}) \cdot \sin(\frac{2n\pi x}{T}) dx \\
=&\int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{1}{2} \cdot \left(\cos\frac{2(m-n)\pi x}{T} - \cos\frac{2(m+n)\pi x}{T}\right) dx \\
=&\begin{cases} 0, & \text{如果} m \neq n, \\ \frac{T}{2}, & \text{如果} m = n. \end{cases}
\end{aligned}$$

引入 Kronecker 记号

$$\delta_{m,n} = \begin{cases} 0, & \text{如果} m \neq n, \\ 1, & \text{如果} m = n, \end{cases}$$

则可把上述结果记作

$$\langle \sin(\frac{2m\pi x}{T}), \sin(\frac{2n\pi x}{T}) \rangle = \frac{T}{2}\delta_{m,n}, \quad \forall m, n \in \mathbf{Z}_+.$$

完全类似的, 有如下结果.

例 5.1.5. 对正整数 m, n, 有

$$\begin{aligned}
&\langle 1, 1 \rangle = T, \\
&\langle 1, \sin(\frac{2n\pi x}{T}) \rangle = 0, \\
&\langle 1, \cos(\frac{2n\pi x}{T}) \rangle = 0, \\
&\langle \sin(\frac{2m\pi x}{T}), \sin(\frac{2n\pi x}{T}) \rangle = \frac{T}{2}\delta_{m,n}, \\
&\langle \sin(\frac{2m\pi x}{T}), \cos(\frac{2n\pi x}{T}) \rangle = 0, \\
&\langle \cos(\frac{2m\pi x}{T}), \cos(\frac{2n\pi x}{T}) \rangle = \frac{T}{2}\delta_{m,n}.
\end{aligned}$$

5.1.2 Fourier 级数的定义

问题 5.1.6. *设 f 是周期为 T 的函数, 能否用正弦函数和余弦函数表示 f? 具体的说, 是否存在系数 $\{a_n\}_{n=0}^{\infty}, \{b_n\}_{n=1}^{\infty}$, 使得*

$$\begin{aligned}
f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{2n\pi x}{T} + b_n \sin\frac{2n\pi x}{T}\right) \\
&= \frac{a_0}{2} + \lim_{n\to\infty} \sum_{k=1}^{n} \left(a_k \cos\frac{2k\pi x}{T} + b_k \sin\frac{2k\pi x}{T}\right).
\end{aligned}$$

命题 5.1.7. *如果函数级数*

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right)$$

在 $[-\frac{T}{2},\frac{T}{2}]$ *上一致收敛到* f, *则*

$$\begin{aligned}
a_0&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)dx,\\
a_n&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)\cos\frac{2n\pi x}{T}dx,\quad \forall n\in\mathbf{Z}_+,\\
b_n&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)\sin\frac{2n\pi x}{T}dx,\quad \forall n\in\mathbf{Z}_+.
\end{aligned}\tag{5.1}$$

这样, 如果希望把 f 用三角函数表示, 唯一的候选级数是

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right),\tag{5.2}$$

其系数 a_0,a_n,b_n 由 (1) 式给出. 称这些 a_0,a_n,b_n 为 f 的 Fourier 系数, 称对应的函数级数 (2) 为 f 的 Fourier 级数, 记作

$$f(x)\sim\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right).$$

例 5.1.8. 设 f 是周期为 2π 的函数, 则其 Fourier 系数为

$$\begin{aligned}
a_0&=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)dx,\\
a_n&=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nxdx,\quad \forall n\in\mathbf{Z}_+,\\
b_n&=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nxdx,\quad \forall n\in\mathbf{Z}_+.
\end{aligned}$$

问题 5.1.9. *f 的 Fourier 级数是否收敛? 和函数是什么?*

定理 5.1.10 (狄利克雷)**.** *设 $f:\mathbf{R}\to\mathbf{R}$ 是周期为 T 的函数, 满足:*

(1) f 分段单调, 即存在有限个点 $-\frac{T}{2}=x_0<x_1<...<x_m=\frac{T}{2}$, 使得对 $i=1,...,m$, f 在 (x_{i-1},x_i) 上是单调的;

(2) f 没有第二类间断点, 在一个周期 $[-\frac{T}{2},\frac{T}{2}]$ 内至多有限多个第一类型的间断点,

则 f 的 Fourier 级数处处收敛, 且

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right)=\frac{f(x-)+f(x+)}{2},\quad \forall x\in\mathbf{R},$$

其中 $f(x-), f(x+)$ 分别是 f 在点 x 处的左右极限.

例 5.1.11. 给定函数 $f:[-\frac{T}{2},\frac{T}{2}]\to\mathbf{R}$, 可以把 f 按如下方式延拓成周期为 T 的函数 $\widetilde{f}:\mathbf{R}\to\mathbf{R}$. 对任何 $x\in\mathbf{R}$, 存在唯一的 $m\in\mathbf{Z}$, 使得 $x-mT\in[-\frac{T}{2},\frac{T}{2})$, 定义

$$\widetilde{f}(x)=f(x-mT).$$

我们把 $\widetilde{f}$ 的 Fourier 级数

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right)$$

也称为 f 的 Fourier 级数, 其 Fourier 系数为

$$\begin{aligned}
a_0&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)dx,\\
a_n&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)\cos\frac{2n\pi x}{T}dx,\quad\forall n\in\mathbf{Z}_+,\\
b_n&=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f(x)\sin\frac{2n\pi x}{T}dx,\quad\forall n\in\mathbf{Z}_+.
\end{aligned}$$

由定理5.1.10, 如果 f 在 $[-\frac{T}{2},\frac{T}{2}]$ 上分段单调且只有至多有限多个第一类型的间断点, 则 f 的 Fourier 级数处处收敛, 且有

$$\begin{aligned}
\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right)&=\frac{f(x-)+f(x+)}{2},\quad\forall x\in(-\frac{T}{2},\frac{T}{2}),\\
\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{T}+b_n\sin\frac{2n\pi x}{T}\right)&=\frac{f(\frac{T}{2}-)+f(-\frac{T}{2}+)}{2},\quad\forall x\in\{-\frac{T}{2},\frac{T}{2}\}.
\end{aligned}$$

例 5.1.12 (正弦级数)**.** 设 f 是定义在 $[-\pi,\pi]$ 上的奇函数. 由上例, f 的 Fourier 系数为

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nxdx=0,\quad\forall n\in\mathbf{Z}_{\geq 0},$$

则 f 的 Fourier 级数只包含正弦函数项

$$f(x)\sim\sum_{n=1}^{n}b_n\sin nx,$$

称之为 f 的正弦级数, 其系数为

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nxdx=\frac{2}{\pi}\int_0^{\pi}f(x)\sin nxdx.$$

一般的, 如果 f 是定义在 $[0,\pi]$ 上的函数, 且 $f(0)=0$, 则可以把 f 延拓成 $[-\pi,\pi]$ 上的奇函数

$$\bar{f}(x)=\begin{cases} f(x), & \text{如果} x\in[0,\pi] \\ -f(-x), & \text{如果} x\in[-\pi,0). \end{cases}$$

在不引起混淆的情况下, 我们把 $\bar{f}$ 的正弦级数也称之为 f 的正弦级数, 其系数为

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}\bar{f}(x)\sin nxdx=\frac{2}{\pi}\int_0^{\pi}f(x)\sin nxdx.$$

例 5.1.13 (余弦级数)**.** 设 f 是定义在 $[-\pi,\pi]$ 上的偶函数. 由定义, f 的 Fourier 系数为

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nxdx=0,\quad \forall n\in\mathbf{Z}_+,$$

则 f 的 Fourier 级数只包含余弦函数项

$$f(x)\sim\frac{a_0}{2}+\sum_{n=1}^{n}a_n\cos nx,$$

称之为 f 的余弦级数, 其系数为

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nxdx=\frac{2}{\pi}\int_0^{\pi}f(x)\cos nxdx.$$

一般的, 如果 f 是定义在 $[0,\pi]$ 上的函数, 则可以把 f 延拓成 $[-\pi,\pi]$ 上的偶函数

$$\bar{\bar{f}}(x)=\begin{cases} f(x), & \text{如果} x\in[0,\pi] \\ f(-x), & \text{如果} x\in[-\pi,0). \end{cases}$$

在不引起混淆的情况下, 我们把 $\bar{\bar{f}}$ 的余弦级数也称之为 f 的余弦级数, 其系数为

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}\bar{\bar{f}}(x)\cos nxdx=\frac{2}{\pi}\int_0^{\pi}f(x)\cos nxdx.$$

定理 5.1.14. *设 $f:\mathbf{R}\to\mathbf{R}$ 是周期为 T 的函数, 且 f^2 在 $[-\frac{T}{2},\frac{T}{2}]$ 上可积, 则 f 的 Fourier 系数满足所谓的 Parseval 等式*

$$\frac{a_0^2}{2}+\sum_{n=1}^{\infty}(a_n^2+b_n^2)=\frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}f^2(x)dx.$$

第六章　附录

6.1　volume

定义 6.1.1.

引理 6.1.2. 点 $\mathbf{b}$ 到 k-维子空间 $span\{\mathbf{a}_1, ..., \mathbf{a}_k\}$ 的距离为

$$\sqrt{\mathbf{b}\cdot\mathbf{b} - (\mathbf{a}_1\cdot\mathbf{b}, ..., \mathbf{a}_k\cdot\mathbf{b})(\mathbf{a}_\mathbf{i}\cdot\mathbf{a}_j)^{-1}(\mathbf{a}_1\cdot\mathbf{b}, ..., \mathbf{a}_k\cdot\mathbf{b})^t}$$