

多元微积分

$$W \times f$$

目录

1	极限与连续性	1
1.1	多元函数的例子	1
1.2	\mathbb{R}^m 中的距离与点列极限	3
1.3	多元函数的连续性	11
1.4	多元函数的极限	17
1.5	习题讨论课1	27
2	多元函数的微分	29
2.1	导数、微分与方向导数	29
2.2	坐标系与偏导数	37
2.3	内积空间中可微函数的梯度与方向导数	46
2.4	高阶偏导数和 Taylor 展开	50
2.5	函数凸凹性、函数极值	56
2.6	应用: Newton 法	64
2.7	习题讨论课2	66
2.8	习题讨论课3	68
3	隐函数与逆映射定理	69
3.1	隐函数定理	69
3.2	逆映射定理	75
3.3	应用: 隐函数和反函数求导法	77
3.4	应用: 曲面, 切平面和法向量	78
3.5	应用: 条件极值与 Lagrange 乘子法	86
3.6	附: 隐函数定理的另一个证明	92
3.7	附: 条件极值类型判定	94
3.8	习题讨论课4	96
4	含参数的积分与广义积分	97
4.1	含参数的积分关于参数的连续性、可积性、可微性	97

4.2	含参数的广义积分	104
4.3	应用：积分变换	113
4.4	习题讨论课5	118
5	重积分	119
5.1	重积分的概念	119
5.2	Jordan 可测集和 Riemann 积分	122
5.3	重积分的计算：累次积分	125
5.4	重积分的计算：换元公式	128
5.5	重积分应用的几个例子	132
5.6	习题讨论课6	136
6	曲线与曲面上函数的积分	137
6.1	第一型曲线积分	137
6.2	二维曲面上的第一型曲面积分	141
6.3	\mathbb{R}^m 中任意维数曲面上的函数的积分	144
6.4	附：行列式与体积	147
6.5	习题讨论课7	149
7	向量场的微积分	151
7.1	第二型曲线积分	151
7.2	有势场、保守场和无旋向量场	156
7.3	平面向量场的旋度与散度，Green 公式	160
7.4	第二型曲面积分的概念	164
7.5	空间向量场的旋度与散度，Gauss 公式与 Stokes 公式	173
7.6	向量场的微积分	178
7.7	习题讨论课8	183
8	级数	185
8.1	级数的概念	185
8.2	绝对收敛与比较法	187
8.3	条件收敛的级数及其收敛判别法	192
8.4	级数的交换律、结合律	195
8.5	无穷乘积	198
8.6	习题讨论课：数项级数	200
9	函数项级数与幂级数	203
9.1	一致收敛与极限函数的性质	203
9.2	函数项级数的一致收敛性与和函数的性质	207
9.3	幂级数	212

10 Fourier 级数初步	225
10.1 Fourier 级数的背景	225
10.2 Fourier 级数的定义与计算	228
10.3 平方平均误差与 Fourier 级数的最小二乘性质	232
10.4 Fourier 级数的收敛性	235
10.5 Fourier 级数应用	238
10.6 附: Parseval 等式的证明	242
10.7 习题讨论课	245

第 1 章 极限与连续性

1.1 多元函数的例子

如果一个变量 y 的值由另外一些变量 x_1, \dots, x_m 的值所唯一确定, 则我们说 y 是关于 x_1, \dots, x_m 的一个 m 元函数, 称 y 为因变量或函数值变量, 称 x_1, \dots, x_m 为自变量。

例如, 地球上每个点的位置可以由经度 x_1 和纬度 x_2 描述, 该点处的地表海拔高度 $h(x_1, x_2)$ 是关于经度 x_1 和纬度 x_2 的一个二元函数。

又如, 一张黑白图片, 每个像素点的位置可以由一对横纵坐标 (x_1, x_2) 表示, 而对每个位置 (x_1, x_2) , 有一个灰度值 $f(x_1, x_2)$ 与之对应。

再如, 我们知道单摆运动由微分方程

$$\theta'' + \frac{g}{l} \sin \theta = 0$$

刻画, 而这个二阶微分方程的解由初始条件 $(\theta, v) = (\theta(x_0), \theta'(x_0))$ 确定。而机械能

$$E = \frac{v^2}{2} + \frac{g}{l}(1 - \cos \theta)$$

是关于相点 (θ, v) 的二元函数, 相点组成的集合称为相空间。

上面这三个例子, 第三个我们写出了函数的精确表达式, 但前两个恐怕没有简单的精确表达式。所以表达式只是了解函数的途径之一。

了解函数 $f: D \rightarrow \mathbb{R}$ 的另一途径是函数 f 的图像, 也就是集合

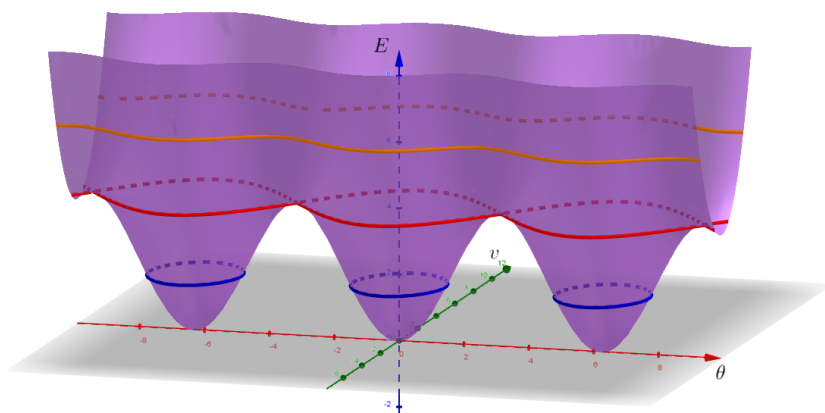
$$\{(\mathbf{x}, f(\mathbf{x})) | \mathbf{x} \in D\}.$$

在某些情况下, 这是比表达式更直观的描述方式。

例如地球表面就是高度函数 $h(x_1, x_2)$ 的图像, 这时我们把定义域 D 看成海平面高度的一个标准的球面。

一张黑白图片本身也就是灰度值函数 $f(x_1, x_2)$ 的图像, 所以你目前看到的这一页就是一个函数的图像。

对单摆的机械能而言, 我们可以借助计算机软件画出机械能函数 $E(\theta, v)$ 的图像。

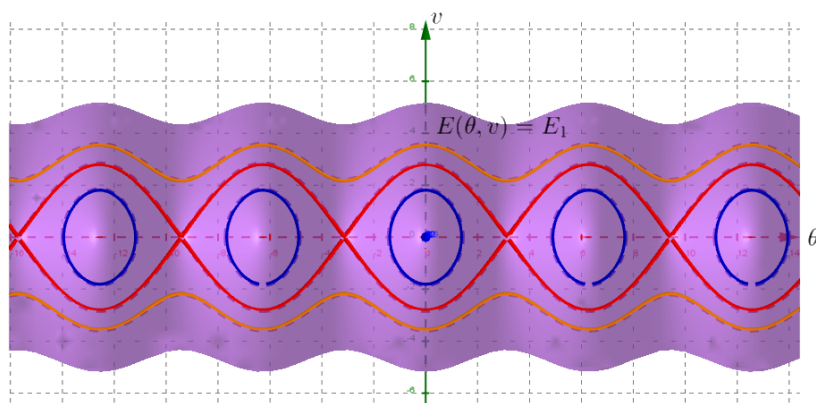


然而，对自变量个数大于 2 的函数，我们需要在至少 4 维空间(至少 3 个自变量和 1 个函数值)中才能描述函数图象。这对于生活在 3 维空间中的人类来说不太现实。

解决这个问题一个办法是利用截面，即把一部分自变量视为参数。比如一部黑白无声电影可以看成是屏幕上的点(由一对坐标 (x_1, x_2))和时间 t 所确定的灰度值函数 $f(x_1, x_2, t)$ 。电影中的每一帧画面就是给定 t 时，关于 (x_1, x_2) 的函数 $f(x_1, x_2, t)$ ，这时把 t 当参数，把 (x_1, x_2) 当成自变量。无声电影的制作过程就是在产生一本连环画(电影胶片)，而放映过程就是把电影胶片变成我们在屏幕上看到的连续画面。

另一个方法，是把函数值反映到自变量空间中，这就是等值集：即对给定的值 C ，相应的集合

$$f^{-1}(C) := \{\mathbf{x} | f(\mathbf{x}) = C\}.$$



如果我们在自变量空间中画出 f 的所有的等值集，我们也就了解了 f 。从等值集的分布情况和形状特征，我们可以了解一个函数的重要性质。例如，在

地图中通过等高线, 我们可以了解海拔高度函数 $h(x_1, x_2)$ 的变化, 从而知道哪里是山峰, 哪里是山谷。又如, 由于机械能守恒, 所以在单摆的相平面上, 能量函数 E 的等值集包含了单摆方程的相曲线 $(\theta(t), \theta'(t))$ 。我们会发现在平衡点 $(0, 0)$, $(\pi, 0)$, 以及其他点附近, 能量函数的等值集的分布情况是不一样的, 从中我们可以了解它们运动性态上的差别。

除了多元函数, 还有多元映射, 即多个变量 y_1, \dots, y_n 的值由一组变量 x_1, \dots, x_m 所确定

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_m) \\ \vdots \\ f_n(x_1, \dots, x_m) \end{pmatrix}, \quad \text{即} \quad \begin{cases} y_1 = f_1(x_1, \dots, x_m), \\ \vdots \\ y_n = f_n(x_1, \dots, x_m). \end{cases}$$

例如, 电脑屏幕上的一张彩色画面是每个位置 (x_1, x_2) 对应一组红绿蓝色彩值 (r, g, b) 。又如, 在某一瞬间, 地球在经纬度 (x_1, x_2) 的地方的温度 T 、气压 P 和风速大小 V 。再如, 平面直角坐标 (x, y) 和极坐标 (r, θ) 之间由二元映射

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}, \quad \text{即} \quad \begin{cases} x = r \cos \theta, \\ y = r \sin \theta \end{cases}$$

相互联系。

有一种多元映射称为向量场, 它在空间的每个点 \mathbf{x} 处指定一个向量 $\mathbf{V}(\mathbf{x})$ 。例如, 单摆方程在它的相空间中每个点 (θ, v) 处对应一个向量 $(v, -\frac{g}{l} \sin \theta)$ 。

习题1.1

1. 举出更多日常生活中遇到的多元函数、多元映射和向量场, 包括有具体表达式的例子和没有具体表达式的例子。并说明它们的图像和等值集。
2. 举一个至少有四个自变量的多元函数的例子, 并描述它的样子。
3. 用 Geogebra 或 Wolframalpha 画以下二元函数的图像和它的等值集。
(a) $x^2 + 2y^2$; (b) $x^2 - y^2$; (c) $x^2 + xy + y^2$; (d) xy ; (e) $x + 2y$
4. 计算函数 $f(x, y) = \frac{x^2 + y^2}{x + y}$ 的等值集, 并据此画出 f 的函数图像。
5. 计算函数 $g(x, y, z) = \frac{x^2 + y^2 + z^2}{x + y + z}$ 的等值集, 并据此想象 g 的函数图像。
6. 如何在一个两个自变量两个函数值的映射的定义域中画出这个映射的样子?

1.2 \mathbb{R}^m 中的距离与点列极限

微积分的目的是通过在局部用比较简单的函数来代替原来的函数来研究与函数有关的性质。比如, 函数的连续性、函数的极限、积分都是用局部常值函

数近似；为了了解函数在一点附近的变化，我们用线性函数(微分)、二次函数或多项式(Taylor 展开)近似。这些都依赖于对逼近程度的度量，所以让我们从距离和极限的概念开始。

由 m 个实数 x_1, x_2, \dots, x_m 组成一个有序实数组，

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = (x_1, \dots, x_m)^T,$$

这些实数组构成一个 m 维(实)线性空间

$$\mathbb{R}^m = \{\mathbf{x} = (x_1, \dots, x_m)^T \mid x_1, \dots, x_m \in \mathbb{R}\}$$

中的元素我们有时称为“点”，有时称为“向量”， x_k 称为 \mathbf{x} 的第 k 个坐标。在 \mathbb{R}^m 中可以引入加法和数乘运算：对 $\mathbf{x} = (x_1, \dots, x_m)^T, \mathbf{y} = (y_1, \dots, y_m)^T \in \mathbb{R}^m$ 及 $\lambda \in \mathbb{R}$,

$$\begin{aligned} \mathbf{x} + \mathbf{y} &:= (x_1 + y_1, \dots, x_m + y_m)^T, \\ \lambda \mathbf{x} &:= (\lambda x_1, \dots, \lambda x_m)^T. \end{aligned}$$

我们知道

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

是 \mathbb{R}^m 的一组基，即每个 $\mathbf{x} \in \mathbb{R}^m$ 可以唯一地写为这组基的线性组合

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_m \mathbf{e}_m =: (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix},$$

这个线性组合中的系数 x_1, x_2, \dots, x_m 就是 \mathbf{x} 在这组基下的坐标。

在点和点之间我们可以引入距离的概念，用来刻画它们之间的远近程度。

定义 1.2.1. 称二元函数 $d: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个距离，如果它满足

1. 对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \geq 0$;
2. 对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, $d(\mathbf{x}, \mathbf{y}) = 0$ 当且仅当 $\mathbf{x} = \mathbf{y}$;

3. 对任意 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ 。

对 $r > 0$, 记

$$B(\mathbf{x}, r) := \{\mathbf{y} \in \mathbb{R}^m \mid d(\mathbf{x}, \mathbf{y}) < r\},$$

称为以 \mathbf{x} 为中心, 以 r 为半径的开球。

称 \mathbb{R}^m 上的距离 d 是平移不变的, 如果对任意 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$,

$$d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) = d(\mathbf{x}, \mathbf{y}).$$

对于向量, 我们可以引入“范数”来度量它的长度。

定义 1.2.2. 称函数 $\|\cdot\|: \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个范数, 如果

1. 对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;
2. 对任意 $\mathbf{x} \in \mathbb{R}^m$, $\|\mathbf{x}\| \geq 0$; $\|\mathbf{x}\| = 0$ 当且仅当 $\mathbf{x} = \mathbf{0}$;
3. 对任意 $\mathbf{x} \in \mathbb{R}^m$ 和任意 $\lambda \in \mathbb{R}$, $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$ 。

在向量空间中还可以引进“内积”, 它不仅定量地刻画两个向量在方向上的差别, 还能帮助我们度量向量的长度。

定义 1.2.3. 称二元函数 $\langle \cdot, \cdot \rangle: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个内积, 如果它满足

1. (对称) 对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;
2. (双线性) 对任意 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ 以及任意 $\alpha, \beta \in \mathbb{R}$,

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle;$$

3. (正定) 对任意 $\mathbf{x} \in \mathbb{R}^m$, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$; 且 $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ 当且仅当 $\mathbf{x} = \mathbf{0}$ 。

注 1.2.4. 称 $f: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个双线性函数, 如果对任意 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ 以及任意 $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) &= \alpha f(\mathbf{x}, \mathbf{z}) + \beta f(\mathbf{y}, \mathbf{z}); \\ f(\mathbf{z}, \alpha\mathbf{x} + \beta\mathbf{y}) &= \alpha f(\mathbf{z}, \mathbf{x}) + \beta f(\mathbf{z}, \mathbf{y}). \end{aligned}$$

当 f 是对称(即 $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$)或反对称(即 $f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$)时, 上述线性关系等式只须验证其中一个成立即可。

定理 1.2.5. 1. \mathbb{R}^m 上的范数 $\|\cdot\|$ 给出一个平移不变的距离 d :

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

2. \mathbb{R}^m 上的内积 $\langle \cdot, \cdot \rangle$ 确定了一个范数 $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, 内积和由它确定的范数满足 *Cauchy-Schwarz* 不等式:

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

例 1.2.6. (1) 对任意 $p \geq 1$,

$$\|\mathbf{x}\|_p := (|x_1|^p + \cdots + |x_m|^p)^{\frac{1}{p}}$$

定义了 \mathbb{R}^m 上一个范数, 它给出 \mathbb{R}^m 上一个距离:

$$d_p(\mathbf{x}, \mathbf{y}) := (|x_1 - y_1|^p + \cdots + |x_m - y_m|^p)^{\frac{1}{p}}.$$

当 $p = 2$ 时, 范数 $\|\cdot\|_2$ 和距离 d_2 分别称为欧氏范数和欧氏距离, 它们由内积

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_m y_m$$

导出。

(2)

$$\|\mathbf{x}\|_\infty := \max\{|x_1|, \dots, |x_m|\}$$

定义了 \mathbb{R}^m 上一个范数, 它给出 \mathbb{R}^m 上一个距离:

$$d_\infty(\mathbf{x}, \mathbf{y}) := \max\{|x_1 - y_1|, \dots, |x_m - y_m|\}.$$

(3) 对任意 $p, q \geq 1$,

$$\begin{aligned} \|\mathbf{x}\|_\infty &\leq \|\mathbf{x}\|_p \leq m^{\frac{1}{p}} \|\mathbf{x}\|_\infty, \\ \frac{1}{m^{\frac{1}{q}}} \|\mathbf{x}\|_q &\leq \|\mathbf{x}\|_p \leq m^{\frac{1}{p}} \|\mathbf{x}\|_q. \end{aligned}$$

定义 1.2.7. 称集合 $E \subset \mathbb{R}^m$ 有界, 如果存在 $M > 0$ 使得对任意 $\mathbf{x} \in E$, $\|\mathbf{x}\| \leq M$ 。

定义 1.2.8. 设 $\{\mathbf{x}_n\}$ 是 $\mathbf{x}_n \in \mathbb{R}^m$ 中的点列。

1. 称 $\{\mathbf{x}_n\}$ **有界**, 如果存在 $M > 0$ 使得对任意正整数 n , $\|\mathbf{x}_n\| \leq M$ 。
2. 称 $\{\mathbf{x}_n\}$ **收敛**, 如果存在 $A \in \mathbb{R}^m$ 使得对任意 $\varepsilon > 0$, 存在正整数 $N(\varepsilon) > 0$ 使得对任意正整数 n , 只要 $n \geq N(\varepsilon)$ 就有 $\|\mathbf{x}_n - A\| < \varepsilon$ 。称这样的 A 是点列 $\{\mathbf{x}_n\}$ 的一个**极限**。记 $A = \lim_{n \rightarrow +\infty} \mathbf{x}_n$ 。
3. 称 $\{\mathbf{x}_n\}$ 是一个 **Cauchy 点列**, 如果对任意 $\varepsilon > 0$, 存在正整数 $N(\varepsilon) > 0$ 使得对任意正整数 p, q , 只要 $p, q \geq N(\varepsilon)$ 就有 $\|\mathbf{x}_p - \mathbf{x}_q\| < \varepsilon$ 。

因为

$$|x_k| \leq \|\mathbf{x}\|_\infty = \max_{1 \leq k \leq m} |x_k| \leq \sum_{k=1}^m |x_k|,$$

所以以下命题1.2.9、定理1.2.10、定理1.2.11在 $\|\cdot\|_\infty$ 下成立, 即点列的有界性、收敛性、Cauchy 性质可以通过坐标分量来检查。

命题 1.2.9. 设 $\{\mathbf{x}_n\}$ 是 $\mathbf{x}_n \in \mathbb{R}^m$ 中的点列, $\mathbf{x}_n = (x_1^{(n)}, \dots, x_m^{(n)})^T$ 。则

1. $\{\mathbf{x}_n\}$ 是有界点列当且仅当对任意 $1 \leq k \leq m$, 数列 $\{x_k^{(n)}\}_{n \geq 1}$ 是有界数列。
2. $\{\mathbf{x}_n\}$ 是 Cauchy 点列当且仅当对任意 $1 \leq k \leq m$, 数列 $\{x_k^{(n)}\}_{n \geq 1}$ 是 Cauchy 数列。
3. $\lim_{n \rightarrow +\infty} \mathbf{x}_n = A$ 当且仅当对任意 $1 \leq k \leq m$, 数列 $\lim_{n \rightarrow +\infty} x_k^{(n)} = A_k$ 。

□

于是利用实数集 \mathbb{R} 的性质可以知道: (在 $\|\cdot\|_\infty$ 下)

定理 1.2.10. \mathbb{R}^m 中的一个点列收敛当且仅当它是一个 Cauchy 点列。 □

定理 1.2.11. \mathbb{R}^m 中任何有界点列都有收敛的子点列。 □

定理 1.2.12. \mathbb{R}^m 上任何范数 $\|\cdot\|$ 都与范数 $\|\cdot\|_\infty$ 等价, 即存在正数 $C > 1$ 使得对任意 $\mathbf{x} \in \mathbb{R}^m$,

$$\frac{1}{C} \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq C \|\mathbf{x}\|_\infty.$$

因而 \mathbb{R}^m 上所有范数彼此等价。

证明. 记 $\mathbf{e}_k = (0, \dots, 1, \dots, 0)^T$, 即 \mathbb{R}^m 中第 k 个坐标为 1, 其他坐标都为 0 的向量。

一方面, 记 $C_1 = \|\mathbf{e}_1\| + \cdots + \|\mathbf{e}_m\|$. 则

$$\|\mathbf{x}\| = \|x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m\| \leq |x_1|\|\mathbf{e}_1\| + \cdots + |x_m|\|\mathbf{e}_m\| \leq C_1\|\mathbf{x}\|_\infty.$$

另一方面, 假设对任意正整数 n , 都存在 $\mathbf{y}_n \in \mathbb{R}^m$ 使得

$$\|\mathbf{y}_n\|_\infty > n\|\mathbf{y}_n\|.$$

则 $\mathbf{y}_n \neq \mathbf{0}$, 取 $\mathbf{x}_n = \frac{\mathbf{y}_n}{\|\mathbf{y}_n\|_\infty}$, 则 $\|\mathbf{x}_n\|_\infty = 1$, $\|\mathbf{x}_n\| < \frac{1}{n}$.

因为 $\|\mathbf{x}_n\|_\infty = 1$, 所以根据定理1.2.11知存在子列 \mathbf{x}_{n_k} (在 $\|\cdot\|_\infty$ 下)收敛, 极限为 \mathbf{x}_0 . 因此

$$\|\mathbf{x}_0\| \leq \|\mathbf{x}_{n_k}\| + \|\mathbf{x}_{n_k} - \mathbf{x}_0\| \leq \frac{1}{n_k} + C_1\|\mathbf{x}_{n_k} - \mathbf{x}_0\|_\infty \rightarrow 0, \quad k \rightarrow +\infty.$$

从而 $\mathbf{x}_0 = \mathbf{0}$. 然而

$$1 = \|\mathbf{x}_n\|_\infty \leq \|\mathbf{x}_0\|_\infty + \|\mathbf{x}_0 - \mathbf{x}_{n_k}\|_\infty = \|\mathbf{x}_0 - \mathbf{x}_{n_k}\|_\infty \rightarrow 0, \quad k \rightarrow +\infty.$$

矛盾. 所以存在正整数 N 使得对任意 $\mathbf{x} \in \mathbb{R}^m$, $\|\mathbf{x}\|_\infty \leq N\|\mathbf{x}\|$. 取 $C = C_1 + N$, 则对任意 $\mathbf{x} \in \mathbb{R}^m$,

$$\frac{1}{C}\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq C\|\mathbf{x}\|_\infty.$$

□

于是命题1.2.9、定理1.2.10、定理1.2.11在任何范数 $\|\cdot\|$ 下成立. 在 \mathbb{R}^m 中讨论涉及有界、收敛以及后面引入的连续、可微等概念的问题时, 所得结论与选用的范数无关.

定义 1.2.13. 称集合 $E \subseteq \mathbb{R}^m$ 是一个**闭集**, 如果 E 中任何收敛点列 $\{\mathbf{x}_n\}$ 的极限 $\lim_{n \rightarrow +\infty} \mathbf{x}_n$ 也在 E 中.

定理 1.2.14. $K \subseteq \mathbb{R}^m$ 是有界闭集 $\Leftrightarrow K$ 中任何点列有在 K 中收敛的子列.

证明. (\Leftarrow) 设 K 满足: K 中任何点列有在 K 中收敛的子列. 假设 K 无界. 则对任意正整数 n , 存在 $\mathbf{x}_n \in K$ 使得 $\|\mathbf{x}_n\| > n$. 于是点列 $\{\mathbf{x}_n\}$ 的任何子列也无界, 从而不收敛. 这与 K 中任何点列有在 K 中收敛的子列矛盾.

假设 K 不是闭集. 则存在 K 中点列 $\{\mathbf{x}_n\}$ 使得 $A = \lim_{n \rightarrow +\infty} \mathbf{x}_n$ 存在但 $A \notin K$. 然而 $\{\mathbf{x}_n\}$ 的任何子列也以 A 为极限, 这与 K 中任何点列有在 K 中收敛的子列矛盾.

所以 K 是有界闭集.

(\Rightarrow) 设 K 是有界闭集, 则 K 中任何点列 $\{\mathbf{x}_n\}$ 有界, 因此有在 $\{\mathbb{R}^m\}$ 中收敛的子列 $A = \lim_{k \rightarrow +\infty} \mathbf{x}_{n_k}$. 由于 K 是闭集, 所以 $A \in K$. □

习题1.2

1. 使用 Geogebra 软件在 \mathbb{R}^2 中画出以下集合的图形:

$$\{(x, y) \mid |x| + |y| = 1\}, \quad \{(x, y) \mid |x|^2 + |y|^2 = 1\}, \quad \{(x, y) \mid \max\{|x|, |y|\} = 1\}.$$

观察

$$\{(x, y) \mid |x|^p + |y|^p = 1\}$$

的图形随参数 p 的变化, 并说明为什么 $0 < p < 1$ 时 $(|x|^p + |y|^p)^{\frac{1}{p}}$ 不是范数。结合上述图形, 理解定理1.2.14。

2. 称 $C \subset \mathbb{R}^m$ 是一个**凸集**, 如果对任意 $\mathbf{x}, \mathbf{y} \in C$ 以及任意 $0 < t < 1$, 都有 $(1-t)\mathbf{x} + t\mathbf{y} \in C$ 。称 $C \subset \mathbb{R}^m$ 是**对称的**, 如果对任意 $\mathbf{x} \in C$, 都有 $-\mathbf{x} \in C$ 。

- (a) 若 $\|\cdot\|$ 是 \mathbb{R}^m 上的一个范数, 则 $C = \{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x}\| \leq 1\}$ 是一个对称的闭的凸集。
- (b) 设 \mathbb{R}^m 上的函数 $\|\cdot\|$ 满足正定性和正齐次性。证明 $\|\cdot\|$ 是范数当且仅当 $C = \{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x}\| \leq 1\}$ 是一个凸集。
- (c) 设 $C \subset \mathbb{R}^m$ 是一个对称凸集, 满足: 对任意 $\mathbf{x} \in \mathbb{R}^m$, 存在 $\lambda_{\mathbf{x}} > 0$ 使得 $\lambda_{\mathbf{x}}^{-1}\mathbf{x} \in C$ 。试讨论集合 C 的 Minkowski 泛函

$$\|\mathbf{x}\|_C := \inf\{\lambda > 0 \mid \lambda^{-1}\mathbf{x} \in C\}$$

是否在 \mathbb{R}^m 上定义了一个范数。

3. 设 $\|\cdot\|^\#$ 和 $\|\cdot\|^\ast$ 分别是 \mathbb{R}^m 和 \mathbb{R}^n 上的两个范数。证明对 $p \geq 1$,

$$\begin{aligned} \|(\mathbf{x}, \mathbf{y})\|_p &= \left(\|\mathbf{x}\|^\#{}^p + \|\mathbf{y}\|^\ast{}^p \right)^{\frac{1}{p}}, \\ \|(\mathbf{x}, \mathbf{y})\|_\infty &= \max\{\|\mathbf{x}\|^\#, \|\mathbf{y}\|^\ast\}, \end{aligned}$$

都是 $\mathbb{R}^m \times \mathbb{R}^n$ 上的范数。

4. 设 A 是一个 n 行 m 列的矩阵, A 的 m 个列向量线性无关, $\|\cdot\|$ 是 \mathbb{R}^n 上的一个范数。对 $\mathbf{x} \in \mathbb{R}^m$, 令

$$\|\mathbf{x}\|_A = \|A\mathbf{x}\|.$$

证明: $\|\cdot\|_A$ 是 \mathbb{R}^m 上的一个范数。

5. 设 $\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_m$ 是 \mathbb{R}^m 的任意一组基。则对任意 $\mathbf{x} \in \mathbb{R}^m$, 存在唯一一组实数 ξ_1, \dots, ξ_m 使得

$$\mathbf{x} = \xi_1 \tilde{\mathbf{e}}_1 + \dots + \xi_m \tilde{\mathbf{e}}_m.$$

记

$$\|\mathbf{x}\| = \max\{|\xi_1|, \dots, |\xi_m|\}.$$

证明: $\|\cdot\|$ 是 \mathbb{R}^m 上的一个范数。再由 \mathbb{R}^m 上范数等价知, \mathbb{R}^m 中涉及极限、连续等概念的结论与 \mathbb{R}^m 中坐标系选取无关。

6. (Lagrange 恒等式) 证明对任意实数 $a_1, \dots, a_n, b_1, \dots, b_n$,

$$\left(\sum_{i=1}^n |a_i|^2\right) \left(\sum_{i=1}^n |b_i|^2\right) - \left|\sum_{i=1}^n a_i b_i\right|^2 = \sum_{1 \leq i < j \leq n} |a_i b_j - a_j b_i|^2.$$

由这个恒等式不难得到传统的 Cauchy-Schwarz 不等式

$$\left(\sum_{i=1}^n |a_i|^2\right) \left(\sum_{i=1}^n |b_i|^2\right) \geq \left|\sum_{i=1}^n a_i b_i\right|^2,$$

同时它也给出了计算 \mathbb{R}^n 中以向量 $\mathbf{a} = (a_1, \dots, a_n)^T$ 和 $\mathbf{b} = (b_1, \dots, b_n)^T$ 为邻边所形成的平行四边形的面积的办法, 等式右端的那些二阶行列式的绝对值 $|a_i b_j - a_j b_i|$ 是该平行四边形在各个二维坐标平面上的投影面积。

7. 设 $\|\cdot\|$ 是 \mathbb{R}^m 上的一个范数。证明存在 \mathbb{R}^m 上的内积 $\langle \cdot, \cdot \rangle$ 使得对任意 $\mathbf{x} \in \mathbb{R}^m$, $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, 当且仅当对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$,

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$$

(提示: $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2}{2}$) 并根据这个结论判断是否存在内积使得 $\|\mathbf{x}\|_\infty = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ 。

8. 记 $\mathcal{M}_{m,n}$ 为由所有 m 行 n 列的实数方阵组成的线性空间, 定义

$$\langle A, B \rangle = \text{tr}(A^T B), \quad A, B \in \mathcal{M}_{m,n},$$

其中 $\text{tr}(C)$ 表示矩阵 C 的对角线元素之和(称为 C 的迹)。证明 $\langle A, B \rangle$ 是 $\mathcal{M}_{m,n}$ 上的一个内积。这实际上是 $\mathbb{R}^{m \times n}$ 上的标准内积。

9. (压缩不动点定理) 设 $E \subseteq \mathbb{R}^m$ 是一个非空闭集, $F: E \rightarrow \mathbb{R}^m$ 满足

(a) $F(E) \subseteq E$;

(b) 存在 $0 < \lambda < 1$ 使得 $\|F(\mathbf{x}) - F(\mathbf{y})\| \leq \lambda \|\mathbf{x} - \mathbf{y}\|$ 。

证明 F 有唯一不动点 \mathbf{x}^* (即 $F(\mathbf{x}^*) = \mathbf{x}^*$), 且对任意 $\mathbf{x} \in E$, $\lim_{n \rightarrow +\infty} F^n(\mathbf{x}) = \mathbf{x}^*$, 事实上,

$$\|F^n(\mathbf{x}) - \mathbf{x}^*\| \leq \frac{\lambda^n}{1 - \lambda} \|F(\mathbf{x}) - \mathbf{x}\|, \quad \forall \mathbf{x} \in E, \forall n \geq 1.$$

10. 定义

$$d_1(x, y) = \left| x - \frac{1}{2} \sin x - y + \frac{1}{2} \sin y \right|,$$

$$d_2(x, y) = |e^x - e^y|,$$

- (a) 证明: d_1, d_2 都是 \mathbb{R} 上的一个距离, 但它们都不是平移不变的。
 (b) 这两个距离是否与 $|x - y|$ 定义的距离等价?

1.3 多元函数的连续性

定义 1.3.1. 设 $E \subseteq \mathbb{R}^m$. 称 $f: E \rightarrow \mathbb{R}^p$ 在 $\mathbf{x}_0 \in E$ 处连续, 如果对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 $\mathbf{x} \in E$, 只要 $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ 就有 $\|f(\mathbf{x}) - f(\mathbf{x}_0)\|_* < \varepsilon$. 这里 $\|\cdot\|, \|\cdot\|_*$ 分别是 f 的定义域空间和值域空间中的范数。

称 f 在 $E_0 \subseteq E$ 上连续, 如果 f 在每个 $\mathbf{x}_0 \in E_0$ 处连续。

称 f 是一个连续映射, 如果 f 在它的定义域 E 上连续。

定理 1.3.2. $f: E \rightarrow \mathbb{R}^p$ 在 $\mathbf{x}_0 \in E$ 处连续当且仅当对 E 中任意点列 $\{\mathbf{x}_n\}_{n \geq 1}$, 只要 $\lim_{n \rightarrow +\infty} \mathbf{x}_n = \mathbf{x}_0$, 就有 $\lim_{n \rightarrow +\infty} f(\mathbf{x}_n) = f(\mathbf{x}_0)$. \square

易见, 连续性与空间中等价范数的选取无关。

命题 1.3.3. 设 $f: E \rightarrow \mathbb{R}^p$ 连续, $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))^T$. 则 $f: E \rightarrow \mathbb{R}^p$ 连续当且仅当每个 $f_k: E \rightarrow \mathbb{R}$ ($1 \leq k \leq p$) 都连续。

证明. 充分性时对 $f(\mathbf{x})$ 用范数 $\|\cdot\|_1$; 必要性时对 $f(\mathbf{x})$ 用范数 $\|\cdot\|_\infty$. \square

例 1.3.4. \mathbb{R}^m 上任何常值映射是连续的。

例 1.3.5. \mathbb{R}^m 上任何范数是连续的。

证明. $|\|\mathbf{x}\| - \|\mathbf{x}_0\|| \leq \|\mathbf{x} - \mathbf{x}_0\|$. \square

例 1.3.6. \mathbb{R}^m 上任何线性映射是连续的。

证明. 设 $L: \mathbb{R}^m \rightarrow \mathbb{R}^p$ 是线性映射。于是对

$$\mathbf{x} = x_1 \mathbf{e}_1 + \cdots + x_m \mathbf{e}_m, \quad \mathbf{y} = y_1 \mathbf{e}_1 + \cdots + y_m \mathbf{e}_m,$$

由 L 的线性知

$$\begin{aligned} \|L(\mathbf{x}) - L(\mathbf{y})\| &= \|(x_1 - y_1)L(\mathbf{e}_1) + \cdots + (x_m - y_m)L(\mathbf{e}_m)\| \\ &\leq |x_1 - y_1| \|L(\mathbf{e}_1)\| + \cdots + |x_m - y_m| \|L(\mathbf{e}_m)\| \\ &\leq \|\mathbf{x} - \mathbf{y}\|_\infty (\|L(\mathbf{e}_1)\| + \cdots + \|L(\mathbf{e}_m)\|). \end{aligned}$$

\square

例 1.3.7. 任何双线性映射 $B: \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^p$, 即 $\forall \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x} \in \mathbb{R}^m, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y} \in \mathbb{R}^n$

$$B(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2, \mathbf{y}) = \alpha_1 B(\mathbf{x}_1, \mathbf{y}) + \alpha_2 B(\mathbf{x}_2, \mathbf{y}),$$

$$B(\mathbf{x}, \beta_1 \mathbf{y}_1 + \beta_2 \mathbf{y}_2) = \beta_1 B(\mathbf{x}, \mathbf{y}_1) + \beta_2 B(\mathbf{x}, \mathbf{y}_2),$$

都是连续的。

证明.

$$\begin{aligned} \|B(\mathbf{x}, \mathbf{y})\| &= \left\| \sum_{i=1}^m \sum_{j=1}^n x^i y^j B(\mathbf{e}_i, \mathbf{f}_j) \right\| \\ &\leq \sum_{i=1}^m \sum_{j=1}^n |x^i| |y^j| \|B(\mathbf{e}_i, \mathbf{f}_j)\| \\ &\leq \|\mathbf{x}\|_\infty \|\mathbf{y}\|_\infty \sum_{i=1}^m \sum_{j=1}^n \|B(\mathbf{e}_i, \mathbf{f}_j)\| \leq M \|\mathbf{x}\| \|\mathbf{y}\|. \end{aligned}$$

所以

$$\begin{aligned} &\|B(\mathbf{x}, \mathbf{y}) - B(\mathbf{x}_0, \mathbf{y}_0)\| \\ &= \|B(\mathbf{x} - \mathbf{x}_0, \mathbf{y} - \mathbf{y}_0) + B(\mathbf{x} - \mathbf{x}_0, \mathbf{y}_0) + B(\mathbf{x}_0, \mathbf{y} - \mathbf{y}_0)\| \\ &\leq \|B(\mathbf{x} - \mathbf{x}_0, \mathbf{y} - \mathbf{y}_0)\| + \|B(\mathbf{x} - \mathbf{x}_0, \mathbf{y}_0)\| + \|B(\mathbf{x}_0, \mathbf{y} - \mathbf{y}_0)\| \\ &\leq M \|\mathbf{x} - \mathbf{x}_0\| \|\mathbf{y} - \mathbf{y}_0\| + M \|\mathbf{x} - \mathbf{x}_0\| \|\mathbf{y}_0\| + M \|\mathbf{x}_0\| \|\mathbf{y} - \mathbf{y}_0\|. \end{aligned}$$

因此对任意 $\varepsilon > 0$, 当

$$\|(\mathbf{x}, \mathbf{y}) - (\mathbf{x}_0, \mathbf{y}_0)\| = \max\{\|\mathbf{x} - \mathbf{x}_0\|, \|\mathbf{y} - \mathbf{y}_0\|\} < \min\left\{1, \frac{\varepsilon}{M(1 + \|\mathbf{x}_0\| + \|\mathbf{y}_0\|)}\right\}$$

时, $\|B(\mathbf{x}, \mathbf{y}) - B(\mathbf{x}_0, \mathbf{y}_0)\| < \varepsilon$. \square

类似可得

例 1.3.8. 任何多重线性映射都是连续映射。

作为上述例题的推论,

例 1.3.9 (向量运算的连续性). \mathbb{R}^m 中的向量的加法、数乘、内积、实数乘法都是连续映射。

证明. 加法 $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y}$ 是一个线性映射。

数乘 $(\lambda, \mathbf{x}) \mapsto \lambda \mathbf{x}$, 内积 $(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$, 和实数乘法 $(x, y) \mapsto xy$ 都是双线性映射. \square

例 1.3.10. 投影 $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \cdots \times \mathbb{R}^{m_k} \rightarrow \mathbb{R}^{m_i}, (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) \mapsto \mathbf{x}_i$ 是线性映射, 从而是连续映射。

定理 1.3.11 (复合映射的连续性). 设 $E \subseteq \mathbb{R}^m$, $F \subseteq \mathbb{R}^p$, $f: E \rightarrow \mathbb{R}^p$ 在 $\mathbf{x}_0 \in E$ 处连续, $g: F \rightarrow \mathbb{R}^q$ 在 $\mathbf{y}_0 = f(\mathbf{x}_0) \in F$ 连续, 则 $g \circ f: E \cap f^{-1}(F) \rightarrow \mathbb{R}^q$ 在 \mathbf{x}_0 处连续. \square

例 1.3.12. 行列式是关于矩阵的连续函数。

证明. 记 A_i 为矩阵 A 的第 i 列向量, 则 $A \mapsto A_i$ 是连续函数, 而 $\det(A)$ 是关于 (A_1, A_2, \dots, A_m) 的 m 重线性函数, 所以 \det 连续. \square

例 1.3.13. 设 $E \subseteq \mathbb{R}^m$, $f: E \rightarrow \mathbb{R}$ 在 $\mathbf{x}_0 \in E$ 处连续, $f(\mathbf{x}_0) \neq 0$, 则 $\frac{1}{f}$ 在 \mathbf{x}_0 处连续。

证明. 取 $g(y) = \frac{1}{y}$, 于是 g 在 $\mathbb{R} \setminus \{0\}$ 连续, $\frac{1}{f} = g \circ f$ 在 \mathbf{x}_0 连续. \square

例 1.3.14. 考虑 \mathbb{R}^2 上的函数

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, & (x_1, x_2) \neq (0, 0), \\ A, & (x_1, x_2) = (0, 0). \end{cases}$$

问 f 是否为连续函数?

解: 投影

$$L_k(x_1, x_2) = x_k, \quad k = 1, 2$$

都是连续的。

$$h(y) = \frac{1}{y}, \quad y \neq 0$$

是连续函数。所以当 $x_1^2 + x_2^2 \neq 0$ 时,

$$f(\mathbf{x}) = L_1(\mathbf{x})L_2(\mathbf{x})h(L_1(\mathbf{x})L_1(\mathbf{x}) + L_2(\mathbf{x})L_2(\mathbf{x}))$$

是连续函数。

假设 f 在 $\mathbf{x} = 0$ 处连续。记

$$\mathbf{x}(t) = \begin{cases} (t \cos \frac{1}{t}, t \sin \frac{1}{t}), & t \neq 0, \\ (0, 0), & t = 0. \end{cases}$$

则 $\mathbf{x}(t)$ 是连续映射, 所以 $f(\mathbf{x}(t))$ 在 $t = 0$ 连续。而

$$f(\mathbf{x}(t)) = \begin{cases} \frac{1}{2} \sin \frac{2}{t}, & t \neq 0, \\ A, & t = 0, \end{cases}$$

无论 A 取何值, $f(\mathbf{x}(t))$ 在 $t = 0$ 处都不连续, 从而 f 在 $\mathbf{x} = 0$ 处都不连续. \square

定理 1.3.15. 设 $E \subseteq \mathbb{R}^m$, $f: E \rightarrow \mathbb{R}^p$ 连续。则

1. 对 E 的任何有界闭子集 K , $f(K)$ 也是一个有界闭集。
2. 若 $p = 1$, 则对 E 的任何非空有界闭子集 K , f 在 K 上有最大值和最小值。

证明. 任给 $f(K)$ 中点列 $\{f(\mathbf{x}_n)\}_{n \geq 1}$, 其中 $\mathbf{x}_n \in K$ 。

因为 K 是有界闭集, 所以点列 $\{\mathbf{x}_n\}$ 有收敛子列 $\lim_{k \rightarrow +\infty} \mathbf{x}_{n_k} = \mathbf{x}_0 \in K$ 。

因为 f 在 \mathbf{x}_0 连续, 所以 $\lim_{k \rightarrow +\infty} f(\mathbf{x}_{n_k}) = f(\mathbf{x}_0) \in f(K)$ 。

因此 $f(K)$ 是有界闭集。

若 $p = 1$, 则 $f(K)$ 是 \mathbb{R} 中的非空有界闭集, 从而 $f(K)$ 有最大值和最小值。 □

例 1.3.16. 设 $\langle \cdot, \cdot \rangle$ 是 \mathbb{R}^n 上的内积, A 是个 n 阶对称方阵, 即

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

证明: \mathbb{R}^n 中存在 n 个两两正交的单位向量 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 以及实数 $\lambda_1, \dots, \lambda_n$ 使得

$$A\mathbf{v}_k = \lambda_k \mathbf{v}_k, \quad k = 1, \dots, n.$$

证明. 定义

$$f(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle.$$

则 f 是连续函数。考虑 $K = \{\mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{v}\| = 1\}$ 。则 K 是 \mathbb{R}^n 的有界闭子集。 f 在 K 上取得最大值 $\lambda_1 = f(\mathbf{v}_1) = \langle A\mathbf{v}_1, \mathbf{v}_1 \rangle$ 。

任取与 \mathbf{v}_1 正交的单位向量 $\hat{\mathbf{w}}_1$, 记

$$\mathbf{u}(\theta) = \cos \theta \mathbf{v}_1 + \sin \theta \hat{\mathbf{w}}_1, \quad g(\theta) = f(\mathbf{u}(\theta)).$$

则 $\langle \mathbf{u}(\theta), \mathbf{u}(\theta) \rangle = 1$, 即 $\mathbf{u}(\theta) \in K$, 从而 $\theta = 0$ 是 g 的最大值点。

注意到 $\mathbf{u}(\theta) = \mathbf{v}_1 + \theta \hat{\mathbf{w}}_1 + o(\theta)$ ($\theta \rightarrow 0$), 所以

$$\begin{aligned} g(\theta) &= \langle A\mathbf{u}(\theta), \mathbf{u}(\theta) \rangle \\ &= \langle A\mathbf{v}_1 + \theta A\hat{\mathbf{w}}_1 + o(\theta), \mathbf{v}_1 + \theta \hat{\mathbf{w}}_1 + o(\theta) \rangle \\ &= g(0) + 2\theta \langle A\mathbf{v}_1, \hat{\mathbf{w}}_1 \rangle + o(\theta), \quad \theta \rightarrow 0. \end{aligned}$$

因为 $\theta = 0$ 是 g 的最大值点, 所以 $\langle A\mathbf{v}_1, \hat{\mathbf{w}}_1 \rangle = \frac{g'(0)}{2} = 0$, $A\mathbf{v}_1$ 与 $\hat{\mathbf{w}}_1$ 正交。

因此 $A\mathbf{v}_1$ 与 \mathbf{v}_1 共线, 从而 $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ 。

如果 V 是 A 不变子空间, 记 $V^\perp = \{\mathbf{w} \in \mathbb{R}^n \mid \langle \mathbf{w}, \mathbf{v} \rangle = 0, \forall \mathbf{v} \in V\}$, 则

$$\langle A\mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{w}, A\mathbf{v} \rangle = 0, \quad \forall \mathbf{w} \in V^\perp, \forall \mathbf{v} \in V,$$

所以对任意 $\mathbf{w} \in V^\perp$, $A\mathbf{w} \in V^\perp$, 即 V^\perp 是 A 不变子空间。

对特征向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ($k < n$), 它们张成的线性子空间 V_k 及其正交补空间 V_k^\perp 都是 A 不变子空间。在 V_k^\perp 上考虑 f 在 $K \cap V_k^\perp$ 上的最大值点, 得到特征向量 \mathbf{v}_{k+1} 和相应的特征值 λ_{k+1} 。这样最终得到两两正交的单位特征向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 和相应的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。□

定理 1.3.17. 设 $E \subseteq \mathbb{R}^m$, $f: E \rightarrow \mathbb{R}^p$ 连续。则对 E 的任何有界闭子集 K , f 在 K 上是一致连续的, 即对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 $\mathbf{x} \in K$ 以及任意 $\mathbf{y} \in E$, 只要 $\|\mathbf{x} - \mathbf{y}\| < \delta$, 就有 $\|f(\mathbf{x}) - f(\mathbf{y})\|_* < \varepsilon$ 。这里 $\|\cdot\|, \|\cdot\|_*$ 分别是 f 的定义域空间和值域空间中的范数。

证明留作练习。

定义 1.3.18. 称 $E \subseteq \mathbb{R}^m$ 是一个道路连通集合, 如果对任意 $\mathbf{x}, \mathbf{y} \in E$, 存在连续映射 $f: [0, 1] \rightarrow \mathbb{R}^m$ 使得 $f(0) = \mathbf{x}, f(1) = \mathbf{y}$ 且对任何 $t \in [0, 1]$, $f(t) \in E$ 。这样的连续映射 f 被称为是在 E 中从 x 到 y 的一条道路。

定理 1.3.19. 设 $E \subseteq \mathbb{R}^m$ 是一个道路连通集合, $f: E \rightarrow \mathbb{R}^p$ 连续。则 $f(E)$ 是一个道路连通集合。

推论 1.3.20 (介值性质). 设 $E \subseteq \mathbb{R}^m$ 是一个道路连通集合, $f: E \rightarrow \mathbb{R}$ 是连续函数。则 $f(E)$ 是区间。

证明留作练习。

习题1.3

1. 记 $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ 为所有 \mathbb{R}^m 到 \mathbb{R}^n 的线性映射的组成的线性空间。

(a) 对任何 $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, 证明最大值 $\max_{\|\mathbf{x}\| \leq 1} \|A\mathbf{x}\|$ 存在:

(b) 对 $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, 记 $\|A\| := \max_{\|\mathbf{x}\| \leq 1} \|A\mathbf{x}\|$ (其中 $\|\mathbf{x}\|$ 和 $\|A\mathbf{x}\|$ 分别是 \mathbf{x} 在 \mathbb{R}^m 中以及 $A\mathbf{x}$ 在 \mathbb{R}^n 中的范数), 证明这样的 $\|A\|$ 给出 $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ 上一个范数, 并且满足对任何 $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ 以及 $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$,

$$\|BA\| \leq \|B\|\|A\|.$$

(c) 对 2 阶矩阵 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 请用 a, b, c, d 表示 $\max_{\|\mathbf{x}\|_\infty \leq 1} \|A\mathbf{x}\|_\infty$ 以及 $\max_{\|\mathbf{x}\|_1 \leq 1} \|A\mathbf{x}\|_1$ 。

2. 设 $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ 满足 $\|A\| < 1$ 。证明存在唯一的 $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ 使得 $(I-A)(I+B) = I$ 。并且证明 $\|B\| \leq \frac{\|A\|}{1-\|A\|}$ 。(提示: 证明 $F(B) = A+AB$ 是一个压缩映射)

3. 证明: 若 $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ 可逆, 则存在 $\delta > 0$ 使得对任意 $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, 只要 $\|B - A\| < \delta$, 就有 B 可逆. 并且映射 $B \mapsto B^{-1}$ 是连续的.
4. 设 $E \subseteq \mathbb{R}^m$, $f: E \rightarrow \mathbb{R}^p$ 连续. 证明: 对 E 的任何有界闭子集 K , f 在 K 上是一致连续的, 即对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 $\mathbf{x} \in K$ 以及任意 $\mathbf{y} \in E$, 只要 $d(\mathbf{x}, \mathbf{y}) < \delta$, 就有 $d(f(\mathbf{x}), f(\mathbf{y})) < \varepsilon$.
5. (连续性与偏连续性) 设 $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$.

- (a) 如果 f 是连续函数, 那么是否成立: 对任意 $y \in [c, d]$, $f(\cdot, y): [a, b] \rightarrow \mathbb{R}$ 连续; 对任意 $x \in [a, b]$, $f(x, \cdot): [c, d] \rightarrow \mathbb{R}$ 连续?
- (b) 如果对任意 $y \in [c, d]$, $f(\cdot, y): [a, b] \rightarrow \mathbb{R}$ 连续; 对任意 $x \in [a, b]$, $f(x, \cdot): [c, d] \rightarrow \mathbb{R}$ 连续, 那么 f 是否为连续函数? 考虑

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

- (c) 设 $x_0 \in [a, b], y_0 \in [c, d]$, f 满足: $f(\cdot, y_0): [a, b] \rightarrow \mathbb{R}$ 连续; 并且对任意 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$ 使得对任意 $x \in [a, b]$ 以及任意 $y \in [c, d]$, 只要 $|y - y_0| < \delta(\varepsilon)$, 就有 $|f(x, y) - f(x, y_0)| < \varepsilon$. 证明 f 在 (x_0, y_0) 连续.
- (d) 设 $f(x, y)$ 分别对每个变量 x 和 y 是连续的, 且对 y 是单调不减的, 证明 f 关于 (x, y) 是连续的.
6. (a) 如果 f 沿着每条射线 $\mathbf{x}(t) = t\mathbf{x}_0$ 在 $t = 0$ 处都连续, 那么 f 是否在 $\mathbf{x} = \mathbf{0}$ 处连续? 考虑以下函数

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

- (b) 如果 E 是道路连通集, f 沿着 E 中每条经过点 \mathbf{x}_0 的连续曲线 $\mathbf{x}(t)$ 在 \mathbf{x}_0 处都连续, 那么 f 是否在 \mathbf{x}_0 处连续? 为什么?
7. 设 $f: E \rightarrow \mathbb{R}^p$ 是连续映射, 其中 $E \subseteq \mathbb{R}^m$ 是闭集. 证明对任意闭集 $K \subseteq \mathbb{R}^p$, $f^{-1}(K)$ 是 \mathbb{R}^m 的闭子集.
8. 设 A 是对称正定矩阵. 任给单位向量 \mathbf{v}_1 , 记

$$\mathbf{v}_{k+1} = \frac{A\mathbf{v}_k}{\|A\mathbf{v}_k\|}.$$

很明显, 如果极限 $\lim_{k \rightarrow +\infty} \mathbf{v}_k$ 存在且非零, 则该极限是 A 的一个特征向量. 试讨论极限 $\lim_{k \rightarrow +\infty} \mathbf{v}_k$ 是否存在, 以及你能否给出一个计算 A 的所有特征值和特征向量的办法.

9. 对实数 a, b, μ 和正数 σ , 记

$$f(\mu, \sigma, a, b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

讨论函数 f 的连续性。

10. (连续情形的隐函数定理) 设 $F: \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ 是一个连续函数, 满足 $F(\mathbf{x}_0, y_0) = 0$, 在 (\mathbf{x}_0, y_0) 的某个邻域 U 中, 对所有 \mathbf{x} , $F(\mathbf{x}, y)$ 关于 y 都是严格增的。证明存在 \mathbf{x}_0 的一个邻域 W 以及一个连续函数 $g: W \rightarrow \mathbb{R}$, 使得对任意 $\mathbf{x} \in W$, $(\mathbf{x}, y) \in U$ 满足 $F(\mathbf{x}, y) = 0$ 当且仅当 $y = g(\mathbf{x})$ 。
11. 证明 \mathbb{R}^3 上的函数 $f(x, y, z) = \frac{x \sin(xy+z)}{x^2+2y^2+3z^2+1}$ 有正的最大值和负的最小值。
12. 设连续函数 $f: \mathbb{R}^m \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ 满足 $f(t\mathbf{x}) = f(\mathbf{x})$ ($\forall t > 0, \mathbf{x} \neq \mathbf{0}$)。证明 f 有最大值和最小值。
13. 设 $m > 1$, 证明 $\{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x}\| = 1\}$ 是一个道路连通集合。
14. 设 $m > 1$, 证明 $\{A \mid A \text{ 是 } m \text{ 阶方阵, } \det A > 0\}$ 是一个道路连通集合。其中 $\|A\| = \max_{1 \leq i, j \leq m} |a_{ij}|$ 。

1.4 多元函数的极限

极限的定义

定义 1.4.1. 设 $E \subseteq \mathbb{R}^m$ 。称 $\mathbf{x}_0 \in \mathbb{R}^m$ 是 E 的一个聚点, 如果对任意 $\varepsilon > 0$, 存在 $\mathbf{x} \in E$ 使得 $0 < d(\mathbf{x}, \mathbf{x}_0) < \varepsilon$ 。

设 E 是映射 f 的定义域的一个子集, $\mathbf{x}_0 \in \mathbb{R}^m$ 是 E 的一个聚点。称 $A \in \mathbb{R}^p$ 是 $f(\mathbf{x})$ 当 $\mathbf{x} \in E$ 趋于 \mathbf{x}_0 时的极限, 记为 $\lim_{E \ni \mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = A$, 如果对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 $\mathbf{x} \in E$, 只要 $0 < d_1(\mathbf{x}, \mathbf{x}_0) < \delta$ 就有 $d_2(f(\mathbf{x}), A) < \varepsilon$ 。这里 d_1, d_2 分别是 f 的定义域空间和值域空间中的距离。

当 E 是映射 f 的定义域时, $\lim_{E \ni \mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ 简记为 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ 。

注 1.4.2. (1) E 的聚点不一定是 E 的点, E 的点也未必都是 E 的聚点。在极限的定义中, \mathbf{x} 一定不能是 \mathbf{x}_0 。极限 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ 是否存在以及它的值与 f 是否在 \mathbf{x}_0 处有定义以及 $f(\mathbf{x}_0)$ 的值无关。

(2) 聚点和极限的概念与空间中等价距离的选择无关。

例 1.4.3. 求极限 $\lim_{(x,y) \rightarrow (0,0)} \frac{e^{x^3+y^3} - 1}{x^2 + y^2}$ 。

解: 因为 $e^t - 1 = t + o(t)$ ($t \rightarrow 0$) 且连续, 所以存在 δ_0 使得当 $|t| < \delta_0$ 时,

$$|e^t - 1| \leq 2|t|.$$

对 $\varepsilon > 0$ 和 $\mathbf{x} = (x, y)$, 当 $0 < \|\mathbf{x}\|_\infty < \min\{(\frac{\delta_0}{2})^{1/3}, \frac{\varepsilon}{4}\}$ 时,

$$\begin{aligned} |x^3 + y^3| &\leq |x|^3 + |y|^3 \leq 2\|\mathbf{x}\|_\infty^3 < \delta_0, \\ x^2 + y^2 &\geq \|\mathbf{x}\|_\infty^2 > 0, \end{aligned}$$

从而

$$\left| \frac{e^{x^3+y^3} - 1}{x^2 + y^2} \right| \leq \frac{2|x^3 + y^3|}{x^2 + y^2} \leq \frac{4\|\mathbf{x}\|_\infty^3}{\|\mathbf{x}\|_\infty^2} = 4\|\mathbf{x}\|_\infty < \varepsilon,$$

所以 $\lim_{(x,y) \rightarrow (0,0)} \frac{e^{x^3+y^3} - 1}{x^2 + y^2} = 0$ 。 \square

定理 1.4.4. 设 \mathbf{x}_0 是 E 的一个聚点, 记

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \mathbf{x} \in E \setminus \{\mathbf{x}_0\}; \\ A, & \mathbf{x} = \mathbf{x}_0. \end{cases}$$

则 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = A$ 当且仅当 \tilde{f} 在 \mathbf{x}_0 连续。特别地, f 在 \mathbf{x}_0 连续当且仅当 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$ 。

根据这个定理, 连续函数的运算性质(四则运算、向量运算、内积运算)同样适用于极限。

注 1.4.5. 对前述例题, 以下推理是否成立? 为什么?

$$\lim_{(x,y) \rightarrow (0,0)} \frac{e^{x^3+y^3} - 1}{x^2 + y^2} = \lim_{(x,y) \rightarrow (0,0)} \frac{e^{x^3+y^3} - 1}{x^3 + y^3} \lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = 0$$

定理 1.4.6 (复合函数极限). 设 E 为复合函数 f 的定义域, $\lim_{E \ni \mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$,

$\lim_{\mathbf{y} \rightarrow \mathbf{y}_0} g(\mathbf{y}) = A$ 。记

$$\tilde{g}(\mathbf{y}) = \begin{cases} g(\mathbf{y}), & \mathbf{y} \neq \mathbf{y}_0; \\ A, & \mathbf{y} = \mathbf{y}_0. \end{cases}$$

则 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \tilde{g}(f(\mathbf{x})) = A$ 。特别地, 若 g 在 \mathbf{y}_0 连续, 则 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(f(\mathbf{x})) = A$ 。

例 1.4.7. 求极限 $\lim_{(x,y) \rightarrow (1,0)} (x+y) \frac{x+y+1}{x+y-1}$ 。

解: 记 $D = \{(x, y) | x + y > 0, x + y \neq 1\}$,

$$g(t) = \begin{cases} \frac{t+2}{t} \ln(1+t), & t > -1 \text{ 且 } t \neq 0; \\ 2, & t = 0, \end{cases}$$

$$f(x, y) = x + y - 1.$$

则 f, g 连续, 从而 $g \circ f$ 连续。而

$$(x+y)^{\frac{x+y+1}{x+y-1}} = e^{g(f(x,y))}, \quad (x,y) \in D.$$

从而

$$\lim_{(x,y) \rightarrow (1,0)} (x+y)^{\frac{x+y+1}{x+y-1}} = \lim_{(x,y) \rightarrow (1,0)} e^{g(f(x,y))} = e^{g(f(1,0))} = e^2.$$

□

另解: 因为

$$\lim_{t \rightarrow 0} e^{\frac{t+2}{t} \ln(1+t)} = e^{\lim_{t \rightarrow 0} \frac{t+2}{t} \ln(1+t)} = e^2,$$

所以对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 t , 当 $0 < |t| < \delta$ 时,

$$\left| e^{\frac{t+2}{t} \ln(1+t)} - e^2 \right| < \varepsilon.$$

于是当 (x,y) 满足 $x+y \neq 1$ 且 $0 < |x-1| + |y| < \delta$ 时,

$$0 < |x+y-1| \leq |x-1| + |y| < \delta,$$

从而

$$\left| (x+y)^{\frac{x+y+1}{x+y-1}} - e^2 \right| = \left| e^{\frac{x+y+1}{x+y-1} \ln(x+y)} - e^2 \right| < \varepsilon.$$

所以 $\lim_{(x,y) \rightarrow (1,0)} (x+y)^{\frac{x+y+1}{x+y-1}} = e^2$ 。 □

例 1.4.8. 讨论函数 $f(x,y) = e^{-\frac{x^2}{\sqrt{y}}}$ 在其定义域的聚点处的极限。

解: f 的定义域为

$$D = \{(x,y) | y > 0\}$$

对 $(x_0, y_0) \in D$, f 在 (x_0, y_0) 连续, 所以 $\lim_{(x,y) \rightarrow (x_0, y_0)} e^{-\frac{x^2}{\sqrt{y}}} = e^{-\frac{x_0^2}{\sqrt{y_0}}}$ 。

$(x_0, 0)$ 是 D 的聚点。

当 $x_0 \neq 0$ 时, $\lim_{D \ni (x,y) \rightarrow (x_0, 0)} \frac{\sqrt{y}}{x^2} = 0$, 且 $\frac{\sqrt{y}}{x^2} > 0$ (对 D 中所有足够接近 $(x_0, 0)$ 的 (x,y))。令 $g(t) = e^{-\frac{1}{t}}$, 则 $\lim_{t \rightarrow 0^+} g(t) = 0$,

$$\lim_{(x,y) \rightarrow (x_0, 0)} f(x,y) = \lim_{(x,y) \rightarrow (x_0, 0)} g\left(\frac{\sqrt{y}}{x^2}\right) = \lim_{t \rightarrow 0^+} g(t) = 0.$$

当 $x_0 = 0$ 时, 取 $x(t) = \sqrt{t}$, $y(t) = t^2 / (2 + \sin \frac{1}{t})^2$, 则对 $t > 0$, $(x(t), y(t)) \in D$, 且 $\lim_{t \rightarrow 0^+} (x(t), y(t)) = (0, 0)$, 然而

$$f(x(t), y(t)) = e^{-(2 + \sin \frac{1}{t})},$$

在 $t \rightarrow 0$ 时无极限, 所以 $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$ 不存在。 □

除了在聚点处的极限, 我们还可以讨论在无穷远时的极限。

例 1.4.9. 求极限 $\lim_{x \rightarrow \infty, y \rightarrow a} \left(1 + \frac{1}{x}\right)^{\frac{x^2}{x+y}}$ 。

解: 令 $t = \frac{1}{x}$ 。则对任意 $N > 0$, $|x| > N$ 当且仅当 $0 < |t| < \frac{1}{N}$ 。所以

$$\begin{aligned} \lim_{x \rightarrow \infty, y \rightarrow a} \left(1 + \frac{1}{x}\right)^{\frac{x^2}{x+y}} &= \lim_{x \rightarrow \infty, y \rightarrow a} e^{\frac{x^2}{x+y} \ln\left(1 + \frac{1}{x}\right)} \\ &= \lim_{t \rightarrow 0, y \rightarrow a} e^{\frac{1}{t+y} \ln(1+t)} \\ &= \lim_{t \rightarrow 0, y \rightarrow a} e^{\frac{1}{1+y} \frac{\ln(1+t)}{t}} = e. \end{aligned}$$

□

累次极限

关于多元函数, 除了上述形式的极限(称为重极限, 即多个变量同时变化)外, 还有累次极限的概念, 即按照特定的先后顺序逐次对每个变量取极限。

例 1.4.10. 对 $f(x, y) = e^{-\frac{x^2}{\sqrt{y}}}$, 求 $\lim_{x \rightarrow a} \left(\lim_{y \rightarrow b} f(x, y)\right)$ 和 $\lim_{y \rightarrow b} \left(\lim_{x \rightarrow a} f(x, y)\right)$ 。

解: 若 $b > 0$, 则

$$\lim_{y \rightarrow b} e^{-\frac{x^2}{\sqrt{y}}} = e^{-\frac{x^2}{\sqrt{b}}},$$

从而

$$\lim_{x \rightarrow a} \left(\lim_{y \rightarrow b} e^{-\frac{x^2}{\sqrt{y}}}\right) = \lim_{x \rightarrow a} e^{-\frac{x^2}{\sqrt{b}}} = e^{-\frac{a^2}{\sqrt{b}}};$$

同理

$$\lim_{y \rightarrow b} \left(\lim_{x \rightarrow a} e^{-\frac{x^2}{\sqrt{y}}}\right) = e^{-\frac{a^2}{\sqrt{b}}}.$$

若 $b = 0$, 则对 $x \neq 0$,

$$\lim_{y \rightarrow 0^+} e^{-\frac{x^2}{\sqrt{y}}} = 0,$$

从而对任意 a , 都有

$$\lim_{x \rightarrow a} \left(\lim_{y \rightarrow 0^+} e^{-\frac{x^2}{\sqrt{y}}}\right) = 0.$$

对任意 $y > 0$,

$$\lim_{x \rightarrow a} e^{-\frac{x^2}{\sqrt{y}}} = e^{-\frac{a^2}{\sqrt{y}}}.$$

当 $a \neq 0$ 时,

$$\lim_{y \rightarrow 0^+} \left(\lim_{x \rightarrow a} e^{-\frac{x^2}{\sqrt{y}}}\right) = \lim_{y \rightarrow 0^+} e^{-\frac{a^2}{\sqrt{y}}} = 0;$$

当 $a = 0$ 时,

$$\lim_{y \rightarrow 0^+} \left(\lim_{x \rightarrow 0} e^{-\frac{x^2}{\sqrt{y}}}\right) = \lim_{y \rightarrow 0^+} 1 = 1.$$

□

关于多元函数(重)极限与累次极限之间的关系, 见课后习题。
 O 和 o

定义 1.4.11. 对两个函数或映射 f, g ,

1. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = O(g)$, 如果存在 $\delta_0 > 0$ 和 $M > 0$ 使得

$$\|f(\mathbf{x})\| \leq M\|g(\mathbf{x})\|, \quad \forall \mathbf{x} : 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_0.$$

2. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, f 与 g 同阶, 如果 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时 $f = O(g)$ 且 $g = O(f)$, 即存在 $\delta_0 > 0$ 和 $M > 1$ 使得

$$\frac{1}{M}\|g(\mathbf{x})\| \leq \|f(\mathbf{x})\| \leq M\|g(\mathbf{x})\|, \quad \forall \mathbf{x} : 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_0.$$

3. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = o(g)$, 如果对任意 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$ 使得

$$\|f(\mathbf{x})\| \leq \varepsilon\|g(\mathbf{x})\|, \quad \forall \mathbf{x} : 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta(\varepsilon).$$

4. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, f 与 g 等价, 如果 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = g + o(g)$ 。

5. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, f 是一个无穷小, 如果 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|f(\mathbf{x})\| = 0$ 。

6. 称 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, f 是一个无穷大, 如果 $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{1}{\|f(\mathbf{x})\|} = 0$ 。

易见, 上述定义与等价范数的选取无关。

例 1.4.12. 若 $f, g : E \rightarrow \mathbb{R}^p$ 满足 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = O(h)$ 且 $g = O(h)$, 则

$$\lambda f + \mu g = O(h), \quad \mathbf{x} \rightarrow \mathbf{x}_0.$$

即

$$O(h)_{E \rightarrow \mathbb{R}^p} = \{f : E \rightarrow \mathbb{R}^p \mid f = O(h), \mathbf{x} \rightarrow \mathbf{x}_0\}$$

是一个线性空间。类似的,

$$o(h)_{E \rightarrow \mathbb{R}^p} = \{f : E \rightarrow \mathbb{R}^p \mid f = o(h), \mathbf{x} \rightarrow \mathbf{x}_0\}$$

也是一个线性空间, 并且它是 $O(h)_{E \rightarrow \mathbb{R}^p}$ 的一个子空间。

例 1.4.13. 设 $f : E \rightarrow \mathbb{R}^p$, $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))$ 。则 $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = O(g)$ 当且仅当对每个 $k : 1 \leq k \leq p$, $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f_k = O(g)$; $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f = o(g)$ 当且仅当对每个 $k : 1 \leq k \leq p$, $\mathbf{x} \rightarrow \mathbf{x}_0$ 时, $f_k = o(g)$ 。

证明.

$$|f_k(\mathbf{x})| \leq \|f(\mathbf{x})\|_1 = |f_1(\mathbf{x})| + \cdots + |f_p(\mathbf{x})|.$$

□

例 1.4.14. 任何线性映射 $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$ 满足 $\mathbf{x} \rightarrow \mathbf{0}$ 时, $A(\mathbf{x}) = O(\|\mathbf{x}\|)$ 。

证明.

$$\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|.$$

□

例 1.4.15. 任何线性映射 $A: \mathbb{R}^m \rightarrow \mathbb{R}^m$ 满足 $\mathbf{x} \rightarrow \mathbf{0}$ 时, $\langle A(\mathbf{x}), \mathbf{x} \rangle = O(\|\mathbf{x}\|^2)$ 。

证明. 由 Cauchy-Schwarz 不等式,

$$|\langle A(\mathbf{x}), \mathbf{x} \rangle| \leq \|A(\mathbf{x})\|\|\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|^2.$$

□

例 1.4.16. (1) 当 $(x, y) \rightarrow (0, 0)$ 时, $ax^2 + bxy + cy^2$ ($b^2 < 4ac$) 与 $x^2 + y^2$ 同阶, 但 $x^2 - y^2$ 不与 $x^2 + y^2$ 同阶。

(2) 二次多项式 xy 满足当 $(x, y) \rightarrow (0, 0)$ 时, $xy = O(x^2 + y^2)$, 但 xy 不与 $x^2 + y^2$ 同阶。

(3) 对正数 k , $xy = o((x^2 + y^2)^k)$ (当 $(x, y) \rightarrow (0, 0)$ 时) 当且仅当 $k < 1$ 。

证明. (1) 因为 $b^2 < 4ac$, 所以 a, c 同号. 取正数 $\epsilon > 0$ 满足

$$0 < \epsilon < \min \left\{ |a|, \frac{4ac - b^2}{4(|a| + |c|)} \right\}.$$

则

$$|ax^2 + bxy + cy^2| \leq (|a| + |b| + |c|)(x^2 + y^2),$$

所以当 $(x, y) \rightarrow (0, 0)$ 时, $ax^2 + bxy + cy^2 = O(x^2 + y^2)$ 。

另一方面,

$$\begin{aligned} & |ax^2 + bxy + cy^2| \\ & \geq |a|x^2 - |b||xy| + |c|y^2 \\ & = \epsilon(x^2 + y^2) + (|a| - \epsilon)x^2 - |b||xy| + (|c| - \epsilon)y^2 \\ & = \epsilon(x^2 + y^2) \\ & \quad + (|a| - \epsilon) \left(|x| - \frac{|b|}{2(|a| - \epsilon)} |y| \right)^2 + \frac{4(|a| - \epsilon)(|c| - \epsilon) - |b|^2}{4(|a| - \epsilon)} |y|^2 \\ & \geq \epsilon(x^2 + y^2), \end{aligned}$$

所以当 $(x, y) \rightarrow (0, 0)$ 时, $x^2 + y^2 = O(ax^2 + bxy + cy^2)$ 。因此, $ax^2 + bxy + cy^2$ ($b^2 < 4ac$) 与 $x^2 + y^2$ 同阶。

对 $x^2 - y^2$,

$$|x^2 - y^2| \leq x^2 + y^2,$$

所以 $x^2 - y^2 = O(x^2 + y^2)$ 。但是沿直线 $y = x$ 趋于原点时,

$$x^2 - y^2 = 0 < x^2 + y^2.$$

所以 $x^2 + y^2 = O(x^2 - y^2)$ 不成立, 从而 $x^2 + y^2$ 与 $x^2 - y^2$ 不同阶。

(2)

$$|xy| \leq x^2 + y^2.$$

所以 $(x, y) \rightarrow (0, 0)$ 时, $xy = O(x^2 + y^2)$ 。

但是对 $(x, y) = (t, 0)$, $x^2 + y^2 = t^2 > 0 = xy$, 所以不存在常数 $M > 0$ 以及 $\delta_0 > 0$ 使得对任意 $0 < \sqrt{x^2 + y^2} < \delta_0$ 时,

$$|x^2 + y^2| \leq M|xy|.$$

因此 xy 不与 $x^2 + y^2$ 同阶。

(3)

$$\left| \frac{xy}{(x^2 + y^2)^k} \right| = r^{2-2k} |\cos \theta \sin \theta| \leq r^{2(1-k)}, \quad r = \sqrt{x^2 + y^2},$$

所以当 $k < 1$ 时, $\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{(x^2 + y^2)^k} = 0$ 。当 $k \geq 1$ 时, 对 $(x, y) = (t, t)$,

$$\left| \frac{xy}{(x^2 + y^2)^k} \right| = \frac{1}{2^k} t^{2-2k},$$

所以 $\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{(x^2 + y^2)^k} = 0$ 不成立。因此 $xy = o((x^2 + y^2)^k)$ (当 $(x, y) \rightarrow (0, 0)$ 时) 当且仅当 $k < 1$ 。□

这个例子说明, 在多元情形, 即使次数相同的齐次多项式也未必同阶, 因此多元函数比一元函数复杂, 其中很重要的因素是由于定义域空间维数增加导致的: 比如在直线上, 一个动点连续移动趋于一个定点而始终不与动点重合, 只有左右两个方向可以选择; 但是在平面上, 一个动点可以沿任何连续曲线趋于一个动点, 这样的方式有无穷多种。

习题1.4

1. 极限的一般定义涉及滤子的概念。称一个集合族 $\mathcal{B} = \{B_\alpha\}_\alpha$ 是一个滤子基, 如果它满足

(a) $\emptyset \notin \mathcal{B}$;

(b) 对任意 $B_1, B_2 \in \mathcal{B}$, 存在 $B_3 \in \mathcal{B}$ 使得 $B_3 \subseteq B_1 \cap B_2$ 。

记

$$\mathcal{F} = \{F \mid \text{存在 } B \in \mathcal{B} \text{ 使得 } B \subseteq F\}.$$

则 \mathcal{F} 满足

- (1) 若 $F \in \mathcal{F}$, $F \subset G$, 则 $G \in \mathcal{F}$;
- (2) 若 $F_1, F_2 \in \mathcal{F}$, 则 $F_1 \cap F_2 \in \mathcal{F}$;
- (3) $\emptyset \notin \mathcal{F}$ 。

称满足以上三条性质的子集族 \mathcal{F} 为一个滤子。

证明:

(1) 对集合 $E \subseteq \mathbb{R}^m$ 及 E 的聚点 $\mathbf{x}_0 \in \mathbb{R}^m$, $\mathcal{B} = \{U_E(\mathbf{x}_0, r) \mid r > 0\}$ 是一个滤子基, 其中

$$U_E(\mathbf{x}_0, r) = \{\mathbf{x} \in E \mid 0 < d(\mathbf{x}, \mathbf{x}_0) < r\}.$$

这个滤子基对应于 $E \ni \mathbf{x} \rightarrow \mathbf{x}_0$ 。

(2) 对集合 $E \subseteq \mathbb{R}^m$, $\mathcal{B} = \{V_E(r) \mid r > 0\}$ 是一个滤子基, 其中

$$V_E(r) = \{\mathbf{x} \in E \mid \|\mathbf{x}\| > r\}.$$

这个滤子基对应于 $E \ni \mathbf{x} \rightarrow \infty$ 。

(3) 滤子基 \mathcal{B} 由如下形式集合组成

$$\{(x, y) \mid x < -N, y > M\}, \quad M, N > 0.$$

这个滤子基对应于 $x \rightarrow -\infty, y \rightarrow +\infty$ 。

(4) 滤子基 \mathcal{B} 由如下形式集合组成

$$\{(x, y) \mid 0 < x < \varepsilon, y > N\}, \quad \varepsilon, N > 0.$$

这个滤子基对应于 $x \rightarrow 0^+, y \rightarrow +\infty$ 。

2. 极限 $\lim_{\mathcal{F}} f(\mathbf{x}) = A$ 表示: 对任意 $\varepsilon > 0$, 存在 $F \in \mathcal{F}$ 使得对任意 $\mathbf{x} \in F$, $d(f(\mathbf{x}), A) < \varepsilon$ 。证明以下复合函数极限定理:
 设滤子 \mathcal{F}, \mathcal{G} 满足对任意 $G \in \mathcal{G}$, 存在 $F \in \mathcal{F}$ 使得对任意 $\mathbf{x} \in F$ 都有 $f(\mathbf{x}) \in G$, $\lim_{\mathcal{G}} g(\mathbf{y}) = A$, 则 $\lim_{\mathcal{F}} g(f(\mathbf{x})) = A$ 。
3. (1) 证明: 当 $(x, y) \rightarrow (0, 0)$ 时, $|x| + |y|$ 是与 $\sqrt{x^2 + y^2}$ 同阶的无穷小。
 (2) 是否存在正数 k 使得当 $(x, y) \rightarrow (0, 0)$ 时, $x + y$ 与 $(\sqrt{x^2 + y^2})^k$ 是同阶的无穷小?
 (3) 你是否可以对多元函数给出一个“阶”的定义, 使得任何单项式 $x_1^{m_1} \cdots x_n^{m_n}$ 在 $\mathbf{x} \rightarrow \mathbf{0}$ 时是 $m_1 + \cdots + m_n$ 阶的无穷小?

4. 证明: 若 $f + o(f) = g + o(g)$ ($\mathbf{x} \rightarrow \mathbf{x}_0$ 时), 则 $f = g + o(g)$ ($\mathbf{x} \rightarrow \mathbf{x}_0$ 时)。由此知道“等价”是一个对称的关系。
5. 对线性映射 $A: \mathbb{R}^m \rightarrow \mathbb{R}^m$, 证明: 当 $\mathbf{x} \rightarrow \mathbf{0}$ 时 $A\mathbf{x}$ 与 \mathbf{x} 同阶当且仅当 A 可逆。
6. 设 $m > 1$, $A: \mathbb{R}^m \rightarrow \mathbb{R}^m$ 是对称线性映射, 证明: 当 $\mathbf{x} \rightarrow \mathbf{0}$ 时 $\langle A\mathbf{x}, \mathbf{x} \rangle$ 与 $\|\mathbf{x}\|^2$ 同阶当且仅当 A 是正定的或者负定的。
7. 设 $U \subseteq \mathbb{R}^2$ 是开集, $(a, b) \in U$, f 在 $U \setminus \{(a, b)\}$ 上有定义, 且满足极限

$$A = \lim_{(x,y) \rightarrow (a,b)} f(x,y), \quad B = \lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x,y)$$

都存在。证明 $A = B$ 。由此知道, 如果

$$\lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x,y) \neq \lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x,y),$$

则极限 $\lim_{(x,y) \rightarrow (a,b)} f(x,y)$ 不存在。

8. 设

(a) 极限 $H(y) = \lim_{x \rightarrow a} f(x,y)$ 对任意 $y \neq b$ 存在;

(b) 极限 $G(x) = \lim_{y \rightarrow b} f(x,y)$ 对任意 x 存在, 且存在 $\delta_0 > 0$ 使得: 对任意 $\varepsilon > 0$, 存在 $\delta_2(\varepsilon) > 0$ 使得对任意 $0 < |y - b| < \delta_2(\varepsilon)$ 以及任意 $0 < |x - a| < \delta_0$ 都有 $|f(x,y) - G(x)| < \varepsilon$;

则极限 $\lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x,y)$, $\lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x,y)$ 和 $\lim_{x \rightarrow a, y \rightarrow b} f(x,y)$ 都存在, 且

$$\lim_{x \rightarrow a, y \rightarrow b} f(x,y) = \lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x,y) = \lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x,y).$$

9. 设

$$f(x,y) = \begin{cases} x \sin \frac{1}{y}, & y \neq 0, \\ 0, & y = 0. \end{cases}$$

讨论极限 $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$, $\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x,y)$ 以及 $\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x,y)$ 。

10. 二元函数 $x^y (x > 0, y \in \mathbb{R})$ 可以最大扩充到什么范围仍为连续函数?

11. 设 $f(x,y) = x^4 + y^4 + x^2 + xy - y$ 。证明

(a) $\lim_{(x,y) \rightarrow \infty} f(x,y) = +\infty$;

(b) f 有最小值;

(c) 对任意实数 C , $\{(x,y) | f(x,y) = C\}$ 是有界闭集。

12. 设 $f(x, y) > g(x) + h(y)$, 其中 g, h 是连续函数且满足

$$\lim_{x \rightarrow \infty} g(x) = \lim_{y \rightarrow \infty} h(y) = +\infty.$$

证明 $\lim_{(x, y) \rightarrow \infty} f(x, y) = +\infty$.

13. 设 $f(\mathbf{x}) = f_k(\mathbf{x}) + f_{k-1}(\mathbf{x}) + \cdots + f_1(\mathbf{x}) + C$, 其中 $\mathbf{x} \in \mathbb{R}^m$, f_j 是连续函数且是 j 次齐次的, 即

$$f_j(\lambda \mathbf{x}) = \lambda^j f_j(\mathbf{x}), \quad \forall \lambda > 0.$$

(a) 证明: 若 $\lim_{\|\mathbf{x}\| \rightarrow \infty} f_k(\mathbf{x}) = +\infty$, 则

i. $\lim_{\|\mathbf{x}\| \rightarrow \infty} f(\mathbf{x}) = +\infty$;

ii. f 有最小值;

iii. 对任意实数 C , $\{\mathbf{x} | f(\mathbf{x}) = C\}$ 是有界闭集.

(b) 证明: 若 $f(\mathbf{x}) = o(\|\mathbf{x}\|^k)$ ($\mathbf{x} \rightarrow \mathbf{0}$), 则 $f(\mathbf{x}) = 0$. (这个结论可以用来证明 Taylor 多项式的唯一性)

14. 设 $K \subset \mathbb{R}^k$ 是有界闭子集, $f: \mathbb{R}^m \times K \rightarrow \mathbb{R}$ 是连续函数. 记 $g(\mathbf{x}) = \max_{\mathbf{y} \in K} f(\mathbf{x}, \mathbf{y})$. 证明 $g: \mathbb{R}^m \rightarrow \mathbb{R}$ 是连续函数.

15. 对任意正整数 $1 \leq k \leq m$, 记

$$V_k = \{(\mathbf{v}_1, \dots, \mathbf{v}_k) | \mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^m\},$$

$$O_k = \{(\mathbf{v}_1, \dots, \mathbf{v}_k) \in V_k | \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j}, \forall 1 \leq i \leq j \leq k\},$$

其中, $\langle \cdot, \cdot \rangle$ 是 \mathbb{R}^m 上的内积, $\delta_{i,j} = \begin{cases} 0, & i \neq j; \\ 1, & i = j \end{cases}$. 记

$$\|(\mathbf{v}_1, \dots, \mathbf{v}_k)\| = \max_{1 \leq i \leq k} \|\mathbf{v}_i\|.$$

(a) 证明 V_k 是一个 km 维线性空间, O_k 是 V_k 中一个有界闭集.

(b) 对 m 行 n 矩阵 A 和 $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^m$, 记

$$f_k(A, \mathbf{v}_1, \dots, \mathbf{v}_k) = \det((\langle A\mathbf{v}_i, A\mathbf{v}_j \rangle)_{k \times k}).$$

证明 $\max_{(\mathbf{v}_1, \dots, \mathbf{v}_k) \in O_k} f_k(A, \mathbf{v}_1, \dots, \mathbf{v}_k)$ 是 A 的最大的 k 个奇异值的乘积.

(c) 证明矩阵 A 的奇异值 (按从大到小的顺序排列) 关于 A 是连续函数.

1.5 习题讨论课1

1. 判断以下极限的存在性；若存在，求出极限值。

$$(1) \lim_{(x,y) \rightarrow (0,0)} \frac{\arcsin(x^2 + y^2)}{x^2 + y^2} \quad (2) \lim_{(x,y) \rightarrow (0,0)} \frac{xy - \sin(xy)}{xy - xy \cos(xy)}$$

$$(3) \lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2)e^{x-y} \quad (4) \lim_{(x,y) \rightarrow (0,0)} \frac{xy}{\sqrt{x^2 + y^2}}$$

$$(5) \lim_{(x,y) \rightarrow (0,0)} \frac{x+y}{|x|+|y|} \quad (6) \lim_{(x,y) \rightarrow (0,0)} \frac{x^3 - y^3}{x+y}$$

$$(7) \lim_{(x,y) \rightarrow \infty} \frac{x+y}{x^2 + xy + y^2} \quad (8) \lim_{(x,y) \rightarrow (0,0)} x^y$$

$$(9) \lim_{(x,y) \rightarrow \infty} \left(\frac{|xy|}{x^2 + xy + y^2} \right)^{x^2} \quad (10) \lim_{x \rightarrow \infty, y \rightarrow \infty} \left(\frac{|xy|}{x^2 + xy + y^2} \right)^{x^2}$$

2. 设 $f: E \rightarrow \mathbb{R}$ 。证明： $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = A$ 当且仅当对 E 中满足 $\mathbf{x}_n \neq \mathbf{a}$ ($\forall n \geq 1$) 以及 $\lim_{n \rightarrow +\infty} \mathbf{x}_n = \mathbf{a}$ 的任意点列 $\{\mathbf{x}_n\}$ 都有 $\lim_{n \rightarrow +\infty} f(\mathbf{x}_n) = A$

3. 设 $m > 1$, $\|\cdot\|$ 是 \mathbb{R}^m 上一个范数, 记 $S = \{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x}\| = 1\}$ 。证明对任何连续函数 $f: S \rightarrow \mathbb{R}$, $f(S)$ 是一个有界闭区间。

4. 设 $E \subseteq \mathbb{R}^m$ 是闭集, $U \subseteq \mathbb{R}^p$, $F: U \times E \rightarrow \mathbb{R}^m$ 满足

(a) 对任意 $\mathbf{x} \in U$ 及任意 $\mathbf{y} \in E$, 都有 $F(\mathbf{x}, \mathbf{y}) \in E$;

(b) 存在 $0 < \lambda < 1$ 使得对任意 $\mathbf{x} \in U$ 及任意 $\mathbf{y}_1, \mathbf{y}_2 \in E$, 都有

$$\|F(\mathbf{x}, \mathbf{y}_1) - F(\mathbf{x}, \mathbf{y}_2)\| \leq \lambda \|\mathbf{y}_1 - \mathbf{y}_2\|.$$

证明:

(1) 对任意 $\mathbf{x} \in U$, 存在唯一的 $\mathbf{y}(\mathbf{x}) \in E$ 使得 $F(\mathbf{x}, \mathbf{y}(\mathbf{x})) = \mathbf{y}(\mathbf{x})$ 。

(2) 若 $F(\mathbf{x}, \mathbf{y})$ 关于 \mathbf{x} 连续, 则 $\mathbf{y}(\mathbf{x})$ 关于 \mathbf{x} 连续。

5. 设 $\Omega \subseteq \mathbb{R}^m$ 是一个非空集合。

(a) 证明函数 $f: \mathbb{R}^m \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \inf_{\mathbf{y} \in \Omega} \|\mathbf{y} - \mathbf{x}\|$, 是一个连续函数。

(b) 若 Ω 是闭集, 证明对任意 $\mathbf{x} \in \mathbb{R}^m \setminus \Omega$, 存在 $\mathbf{y}^* \in \Omega$ 使得 $f(\mathbf{x}) = \|\mathbf{y}^* - \mathbf{x}\|$, 并且 $\mathbf{y}^* \in \partial\Omega$ (即 \mathbf{y}^* 的任意邻域中都有不属于 Ω 的点)。

(c) 若 Ω_1, Ω_2 是闭集且其中至少一个有界, 证明存在 $\mathbf{x}_k^* \in \Omega_k$ 使得

$$\|\mathbf{x}_1^* - \mathbf{x}_2^*\| = \inf_{\mathbf{x}_1 \in \Omega_1, \mathbf{x}_2 \in \Omega_2} \|\mathbf{x}_1 - \mathbf{x}_2\|.$$

(d) 举例说明上述“闭”、“有界”的条件是不能或缺的。

第2章 多元函数的微分

2.1 导数、微分与方向导数

微分的本质就是在一点附近用线性映射代替非线性映射。

一元函数的导数

$$f'(t_0) = \lim_{h \rightarrow 0} \frac{f(t_0 + h) - f(t_0)}{h}$$

无法直接推广到多元函数，因为当 t_0 和 h 是向量时，上式右端极限中的分式没有意义。但是，上式可以转换为如下写法：

$$\lim_{h \rightarrow 0} \frac{f(t_0 + h) - f(t_0) - f'(t_0)h}{h} = 0,$$

即

$$f(t_0 + h) - f(t_0) - f'(t_0)h = o(h), \quad h \rightarrow 0.$$

最后这个等式也适用于 f 是多元函数(即 t_0, h 是向量)的情况。其中 $f'(t_0)h$ 是一个以 h 为自变量的线性函数，它就是一元函数 f 在 t_0 处的微分，而传统的导数 $f'(t_0)$ 成为表达这个线性函数所需要的比例系数。从函数逼近的角度看，这个线性函数远比这个比例系数在概念上更重要也更本质。这启发我们给出多元函数和多元映射的导数与微分的定义。

定义 2.1.1. 称 \mathbf{x}_0 是集合 $E \subseteq \mathbb{R}^m$ 的一个内点，如果存在正数 $\delta_{\mathbf{x}_0} > 0$ 使得

$$B(\mathbf{x}_0, \delta_{\mathbf{x}_0}) \subseteq E.$$

即 \mathbf{x}_0 及其附近所有的点都在 E 中。 E 的所有内点组成的集合称为 E 的内部，记为 $\text{int}E$ 。

称 E 是一个开集，如果 $E = \text{int}E$ ，即 E 的每个点都是内点。

定义 2.1.2. 设 \mathbf{x}_0 是集合 $E \subseteq \mathbb{R}^m$ 的一个内点, 称 $f: E \rightarrow \mathbb{R}^n$ 在 \mathbf{x}_0 处可微, 如果存在线性映射 $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$ 使得

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + A(\mathbf{v}) + o(\|\mathbf{v}\|), \quad \mathbf{v} \rightarrow \mathbf{0},$$

此时称线性映射 A 为 f 在 \mathbf{x}_0 处的微分或导数, 记为

$$\partial f(\mathbf{x}_0): \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad \partial f(\mathbf{x}_0)(\mathbf{v}) = A(\mathbf{v}).$$

当 $n = 1$ 即 $f: E \rightarrow \mathbb{R}$ 是函数时, 微分 $\partial f(\mathbf{x}_0): \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个线性函数, 人们通常用 $df(\mathbf{x}_0)$ 表示函数 f 在 \mathbf{x}_0 处的微分。

注 2.1.3. (1) 通过平移 $\mathbf{v} = \mathbf{x} - \mathbf{x}_0, \mathbf{w} = \mathbf{y} - f(\mathbf{x}_0)$, 在以 $(\mathbf{x}_0, f(\mathbf{x}_0))$ 为原点、以 (\mathbf{v}, \mathbf{w}) 为向量的线性空间中, f 呈现为以下形式

$$\mathbf{w} = f(\mathbf{x}_0 + \mathbf{v}) - f(\mathbf{x}_0) = A\mathbf{v} + o(\mathbf{v}), \quad \mathbf{v} \rightarrow \mathbf{0}.$$

所以导数 A 是 f 在 $(\mathbf{x}_0, f(\mathbf{x}_0))$ 附近的局部线性化映射。

(2) 可微性和导数的概念无需自变量空间和函数值空间具有坐标系, 但这两个空间要具备范数, 但可微性和导数与等价范数的选择无关。

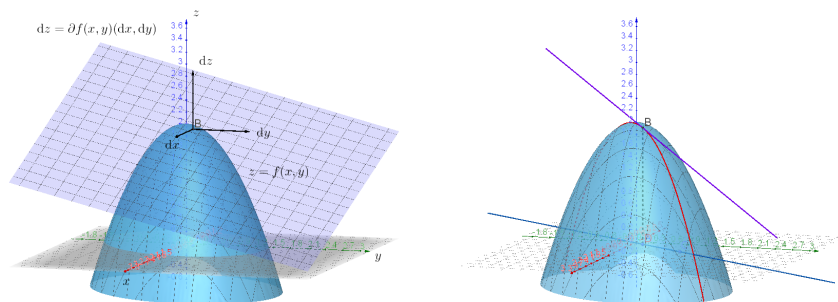


图 2.1: 微分和沿向量的导数

例 2.1.4 (一元映射的导数与微分, 曲线和它的速度向量). 当 $m = 1$ 时, $E = (a, b) \subseteq \mathbb{R}$, 一元映射 $f: (a, b) \rightarrow \mathbb{R}^n$ 相当于 \mathbb{R}^n 中一条曲线, 如果我们把自变量 t 看作时间, 那么 f 也可以看成是曲线上的点沿曲线的运动。极限

$$f'(t_0) = \lim_{h \rightarrow 0} \frac{f(t_0 + h) - f(t_0)}{h}.$$

相当于这个曲线上点的速度向量，它存在等价于

$$f(t_0 + h) = f(t_0) + f'(t_0)h + o(h), \quad h \rightarrow 0,$$

即 f 在 t_0 可微。而 $\partial f(t_0) : h \mapsto f'(t_0)h$ 是 $\mathbb{R} \rightarrow \mathbb{R}^n$ 的一个线性映射，即 f 在 t_0 瞬间相当于一个匀速直线运动， $f'(t_0) = \partial f(t_0)(1)$ 是这个运动的瞬间速度，它是值域空间 \mathbb{R}^n 中一个向量。□

为了便于使用一元微积分的工具研究多元函数，一个常用手法是把多元函数限制在定义域中的一条直线或曲线上，就像我们在下述定义中做的那样。

定义 2.1.5 (沿向量的导数、方向导数). 设 $f : E \rightarrow \mathbb{R}^n$ 和点 $\mathbf{x}_0 \in E$ 以及向量 $\mathbf{v} \in \mathbb{R}^m$ 满足：存在 $\delta > 0$ 使得对任意 $0 \leq t < \delta$ ，都有 $\mathbf{x}_0 + t\mathbf{v} \in E$ ，且

$$\lim_{t \rightarrow 0^+} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

存在，则记它为 $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ 或 $\partial_{\mathbf{v}} f(\mathbf{x}_0)$ ，称为 f 在 \mathbf{x}_0 处沿向量 \mathbf{v} 的导数。

特别地，当 $\mathbf{v} \in \mathbb{R}^m$ 是单位向量(即 $\|\mathbf{v}\| = 1$)时，称它为 f 在 \mathbf{x}_0 处沿方向 \mathbf{v} 的导数。

注 2.1.6. 函数沿向量的导数不要求自变量空间具有范数，也不要求自变量空间具有坐标系。

例 2.1.7 (沿所有向量都有导数，但不连续的函数的例子：一捆乱柴)。

$$f(x, y) = \begin{cases} p^2 x, & \text{若 } xy \neq 0 \text{ 且 } \frac{y}{x} \text{ 是有理数, } \frac{y}{x} = \frac{p}{q} \text{ 是既约分数;} \\ 0, & \text{否则.} \end{cases}$$

f 在 $(0, 0)$ 处沿所有向量有方向导数，但 f 在 $(0, 0)$ 处不连续。

证明. 由定义知 $f(0, 0) = 0$ 。取有理数数列 s_n 使 $\lim_{n \rightarrow +\infty} s_n = \pi$ 。设 $s_n = \frac{p_n}{q_n}$ 是既约分数，其中 p_n, q_n 是正整数。则 $q_n \rightarrow +\infty$ 。于是 $p_n \rightarrow +\infty$ 。取 $x_n = \frac{1}{p_n}$ ， $y_n = \frac{1}{q_n}$ ，则 $\lim_{n \rightarrow +\infty} (x_n, y_n) = (0, 0)$ 。

$$f(x_n, y_n) = p_n^2 x_n = p_n \rightarrow +\infty,$$

所以 f 在 $(0, 0)$ 附近无界，因此 f 在 $(0, 0)$ 处不连续。

任取 $\mathbf{v} = (a, b) \neq (0, 0)$ 。

若 $ab = 0$ 或 b/a 是无理数，则 $f(t\mathbf{v}) = f(ta, tb) = 0$ ，从而 $\frac{\partial f}{\partial \mathbf{v}}(0, 0) = 0$ 。

若 $ab \neq 0$ 且 $b/a = \frac{p}{q}$ 是既约分数，则 $f(t\mathbf{v}) = f(ta, tb) = p^2 ta$ ，从而 $\frac{\partial f}{\partial \mathbf{v}}(0, 0) = p^2 a$ 。

总之， f 在 $(0, 0)$ 处沿所有向量有导数。□

由以下定理知, 上例中的函数 f 虽然在 $(0, 0)$ 处沿所有向量都有导数, 但它在 $(0, 0)$ 处不可微。

定理 2.1.8. 若 f 在 \mathbf{x}_0 可微, 则 f 在 \mathbf{x}_0 连续。 \square

定理 2.1.9. 若 f 在 \mathbf{x}_0 可微, 则 f 在 \mathbf{x}_0 处沿所有向量 \mathbf{v} 都有导数, 且

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

此时, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ 关于 \mathbf{v} 是线性的。 \square

由这个定理知以下函数虽然连续, 且沿每个向量都有导数, 但它不可微。

例 2.1.10 (沿所有向量都有方向导数, 连续但不可微的函数的例子). 对 $f(x, y) = \sqrt{x^2 + y^2}$, 在 $(0, 0)$ 处, 对 $\mathbf{v} = (a, b)$,

$$f(at, bt) = \sqrt{a^2t^2 + b^2t^2} = t\sqrt{a^2 + b^2}, \quad t \geq 0.$$

因此

$$\frac{\partial f}{\partial \mathbf{v}}(0, 0) = \sqrt{a^2 + b^2} = \|\mathbf{v}\|.$$

所以 $\frac{\partial f}{\partial \mathbf{v}}(0, 0)$ 关于 \mathbf{v} 不是线性的, 所以 f 在 $(0, 0)$ 处不可微。 \square

上面的例子说明一元函数与多元函数的可微性虽然定义形式相同, 但有所不同。

接下来, 我们看一些可微函数的例子。

例 2.1.11. (1) 任何常值映射 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ 是可微的, 而且对于任何 $\mathbf{x}_0 \in \mathbb{R}^m$, $\partial f(\mathbf{x}_0) = 0$ 。

(2) 任何线性映射 $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$ 是可微的, 而且对于任何 $\mathbf{x}_0 \in \mathbb{R}^m$, $\partial A(\mathbf{x}_0) = A$ 。

例 2.1.12. 设 $U = \{A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \mid \|A - I\| < 1\}$ 。证明映射

$$f: U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \quad f(A) = A^{-1}$$

是可微的, 并求 $\partial f(A)(B)$ 。

解: 先考虑 $A = I$ 的情形。

对任意满足 $\|B\| < 1$ 的 $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $I + B$ 可逆。因为

$$I = (I + B)(I - B) + B^2 = (I + B) [I - B + (I + B)^{-1}B^2],$$

所以

$$(I + B)^{-1} = I - B + (I + B)^{-1}B^2,$$

其中

$$\begin{aligned} \|(I+B)^{-1}B^2\| &\leq \|(I+B)^{-1}\| \|B\|^2 \\ &\leq (\|(I+B)^{-1} - I\| + \|I\|) \|B\|^2 \\ &\leq \left(\frac{\|B\|}{1-\|B\|} + 1 \right) \|B\|^2 \\ &\leq 2\|B\|^2, \quad \text{当 } \|B\| < \frac{1}{2} \text{ 时.} \end{aligned}$$

因此 $(I+B)^{-1}B^2 = o(B)$ ($B \rightarrow 0$), 从而

$$\partial f(I)(B) = -B.$$

对任意 $A \in U$, 由上述结论知 A 可逆. 当 $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ 满足 $\|B\| < 1/\|A^{-1}\|$ 时, $\|A^{-1}B\| < 1$, 从而 $I+A^{-1}B$ 可逆, 进而 $A+B = A(I+A^{-1}B)$ 可逆, 并且

$$\begin{aligned} (A+B)^{-1} &= (I+A^{-1}B)^{-1}A^{-1} \\ &= [I - A^{-1}B + o(\|B\|)] A^{-1} \\ &= A^{-1} - A^{-1}BA^{-1} + o(\|B\|), \quad \|B\| \rightarrow 0. \end{aligned}$$

其中 $-A^{-1}BA^{-1}$ 关于 B 是线性的, 所以 f 在 A 可微, 并且

$$\partial f(A)(B) = -A^{-1}BA^{-1}.$$

□

注 2.1.13. 随着人工智能的发展, 在机器学习领域中越来越多地涉及到对矩阵的求导运算. 读者可以参考 K. B. Petersen 和 M. S. Pedersen 的讲义 *The Matrix Cookbook*.

定理 2.1.14 (复合函数导数的链索法则). 设 f 在 \mathbf{x}_0 可微, g 在 $\mathbf{y}_0 = f(\mathbf{x}_0)$ 处可微, 则 $g \circ f$ 在 \mathbf{x}_0 处可微, 且复合的微分等于微分的复合:

$$\partial(g \circ f)(\mathbf{x}_0) = \partial g(\mathbf{y}_0) \circ \partial f(\mathbf{x}_0) = \partial g(f(\mathbf{x}_0)) \circ \partial f(\mathbf{x}_0).$$

注 2.1.15. 注意 $\partial g(f(\mathbf{x}_0))$ 是 g 在 $f(\mathbf{x}_0)$ 处的微分, $\partial(g \circ f)(\mathbf{x}_0)$ 是 $g \circ f$ 在 \mathbf{x}_0 处的微分, 它们含义不同.

链索法则的大意是: 接连进行一系列非线性映射, 从线性近似的角度看, 可以近似为依次进行各阶段相应的线性映射. 即粗略地说, 如果

$$\begin{aligned} f(\mathbf{x}_0) &= \mathbf{y}_0, \\ f(\mathbf{x}) &= f(\mathbf{x}_0) + A(\mathbf{x} - \mathbf{x}_0), \\ g(\mathbf{y}) &= g(\mathbf{y}_0) + B(\mathbf{y} - \mathbf{y}_0), \end{aligned}$$

则

$$g(f(\mathbf{x})) = g(\mathbf{y}_0) + B(f(\mathbf{x}) - f(\mathbf{x}_0)) = g(\mathbf{y}_0) + BA(\mathbf{x} - \mathbf{x}_0).$$

严格的证明主要是对在这里被完全忽略了的 o 余项所产生的误差做更细致的分析, 具体证明与我们对一元函数的证明是一样的(见一元微积分讲义)。

由链索法则, 我们可以得到如下结论, 它把一个映射的可微性转换为其分量函数的可微性。

推论 2.1.16. 映射 $f: E \rightarrow \mathbb{R}^n, f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix}$ 在 $\mathbf{x}_0 \in E$ 处可微, 当且仅当 f 的每个坐标分量函数 $f_k: E \rightarrow \mathbb{R}$ 在 \mathbf{x}_0 可微, 且

$$\partial f(\mathbf{x}_0)(\mathbf{v}) = \begin{pmatrix} df_1(\mathbf{x}_0)(\mathbf{v}) \\ \vdots \\ df_n(\mathbf{x}_0)(\mathbf{v}) \end{pmatrix}.$$

证明. (必要性) 取 L_k 为坐标投影函数

$$L_k(y_1, \dots, y_n) = y_k,$$

则 $L_k: \mathbb{R}^n \rightarrow \mathbb{R}$ 是线性函数, 从而可微, 于是由链索法则 $f_k = L_k \circ f$ 在 \mathbf{x}_0 可微, 并且 $df_k(\mathbf{x}_0) = dL_k(\mathbf{y}_0) \circ \partial f(\mathbf{x}_0) = L_k \circ \partial f(\mathbf{x}_0)$ 。所以

$$\partial f(\mathbf{x}_0)(\mathbf{v}) = \begin{pmatrix} df_1(\mathbf{x}_0)(\mathbf{v}) \\ \vdots \\ df_n(\mathbf{x}_0)(\mathbf{v}) \end{pmatrix}.$$

(充分性) 设每个 $f_k: E \rightarrow \mathbb{R}$ 在 \mathbf{x}_0 可微, 则

$$f_k(\mathbf{x}_0 + \mathbf{v}) = f_k(\mathbf{x}_0) + \partial f_k(\mathbf{x}_0)(\mathbf{v}) + \alpha_k(\mathbf{v}), \quad \alpha_k(\mathbf{v}) = o(\mathbf{v}), \quad \mathbf{v} \rightarrow \mathbf{0},$$

所以

$$\begin{pmatrix} f_1(\mathbf{x}_0 + \mathbf{v}) \\ \vdots \\ f_n(\mathbf{x}_0 + \mathbf{v}) \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x}_0) \\ \vdots \\ f_n(\mathbf{x}_0) \end{pmatrix} + \begin{pmatrix} df_1(\mathbf{x}_0)(\mathbf{v}) \\ \vdots \\ df_n(\mathbf{x}_0)(\mathbf{v}) \end{pmatrix} + \begin{pmatrix} \alpha_1(\mathbf{v}) \\ \vdots \\ \alpha_n(\mathbf{v}) \end{pmatrix},$$

其中

$$\|\alpha(\mathbf{v})\| \leq C[|\alpha_1(\mathbf{v})| + \dots + |\alpha_n(\mathbf{v})|] = o(\mathbf{v}), \quad \mathbf{v} \rightarrow \mathbf{0},$$

$\begin{pmatrix} df_1(\mathbf{x}_0)(\mathbf{v}) \\ \vdots \\ df_n(\mathbf{x}_0)(\mathbf{v}) \end{pmatrix}$ 关于 \mathbf{v} 是线性映射, 所以 f 在 \mathbf{x}_0 可微。 □

推论 2.1.17. 若 U, V 都是 \mathbb{R}^m 中的开集, $f: U \rightarrow V$ 是一个可逆映射, 且 f, f^{-1} 都可微, 则

$$\partial(f^{-1})(\mathbf{y}) = (\partial f(\mathbf{x}))^{-1}, \quad \mathbf{y} = f(\mathbf{x}).$$

证明. 恒等式

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

两边对 \mathbf{x} 求导, 由链索法则知

$$\text{id} = \partial(f^{-1})(\mathbf{y}) \circ \partial f(\mathbf{x}).$$

因此

$$\partial(f^{-1})(\mathbf{y}) = (\partial f(\mathbf{x}))^{-1}, \quad \mathbf{y} = f(\mathbf{x}).$$

□

这大意上是说, 如果两个互逆的非线性映射都可以用线性映射来逼近, 那么这两个线性映射就必然互逆。

微分关于加法、乘法、内积等一些基本运算的算法则见本节的习题, 这里不赘述。

例 2.1.18. 直角坐标关于极坐标的可微性。证明 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ 是 $(r, \theta) \in (0, +\infty) \times \mathbb{R}$ 到 $(x, y) \in \mathbb{R}^2$ 的可微映射, 并求它的微分。

解: 根据上述例题结论, 只需证明 $x(r, \theta) = r \cos \theta$ 和 $y(r, \theta) = r \sin \theta$ 分别是可微函数。根据本节后面的习题结论知, 因为 r 和 $\cos \theta, \sin \theta$ 都是关于 (r, θ) 的可微函数, 且

$$dr(\rho, \phi) = \rho, \quad d \cos \theta(\rho, \phi) = -\phi \sin \theta, \quad d \sin \theta(\rho, \phi) = \phi \cos \theta,$$

并且 $r \cos \theta$ 和 $r \sin \theta$ 都是关于 (r, θ) 的可微函数。

$$dx(r, \theta)(\rho, \phi) = r d \cos \theta(\rho, \phi) + \cos \theta dr(\rho, \phi) = -r \sin \theta \cdot \phi + \cos \theta \cdot \rho.$$

同理 $dy(r, \theta)(\rho, \phi) = r \cos \theta \cdot \phi + \sin \theta \cdot \rho$ 。所以

$$\partial \begin{pmatrix} x(r, \theta) \\ y(r, \theta) \end{pmatrix} (\rho, \phi) = \begin{pmatrix} -r \phi \sin \theta + \rho \cos \theta \\ r \phi \cos \theta + \rho \sin \theta \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \begin{pmatrix} \rho \\ \phi \end{pmatrix}.$$

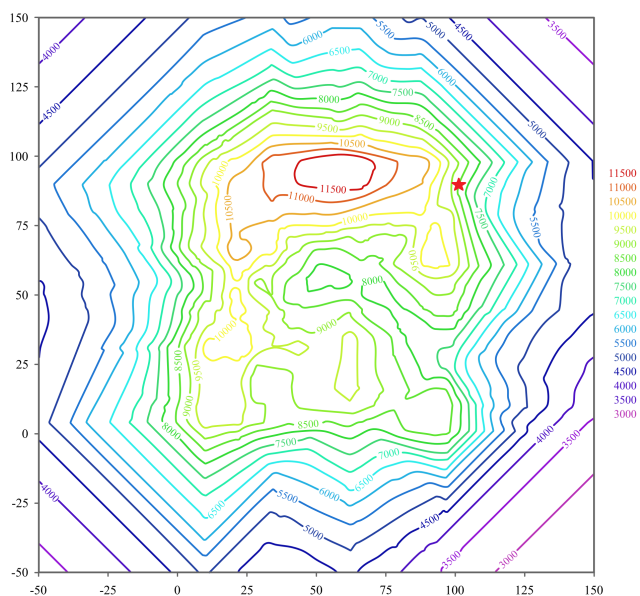
平面极坐标对直角坐标的可微性留作习题。

□

习题2.1

1. 在等高线图¹中所标记的五角星处, 高度函数沿哪些方向的方向导数是正的? 沿哪些方向的方向导数为零?

¹此图来自于维基百科



2. 如果集合 $E \subseteq \mathbb{R}^n$ 和 \mathbf{x}_0 满足: 对任意 $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 都存在 $\delta(\mathbf{v}) > 0$ 使得对任意 $t \in (-\delta(\mathbf{v}), \delta(\mathbf{v}))$, $\mathbf{x}_0 + t\mathbf{v} \in E$, 那么 \mathbf{x}_0 是否为 E 的内点?

3. 证明

(a) 任意一组开集的并集是开集, 有限多个开集的交集是开集。

(b) 集合 $E \subseteq \mathbb{R}^m$ 是开集当且仅当 $\mathbb{R}^m \setminus E$ 是闭集。

(c) 若 $E \subseteq \mathbb{R}^m$ 是开集, 则 $f: E \rightarrow \mathbb{R}^n$ 连续当且仅当对任何开集 $U \subset \mathbb{R}^n$, $f^{-1}(U)$ 是开集。

(d) 若 $E \subseteq \mathbb{R}^m$ 是闭集, 则 $f: E \rightarrow \mathbb{R}^n$ 连续当且仅当对任何闭集 $K \subset \mathbb{R}^n$, $f^{-1}(K)$ 是闭集。

4. 证明: 如果 f 在 \mathbf{x}_0 可微, 则微分 $\partial f(\mathbf{x}_0)$ 是唯一的, 即存在唯一的线性映射 A 使得 $f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + A\mathbf{v} + o(\mathbf{v})(\mathbf{v} \rightarrow \mathbf{0})$ 。

5. 设 $f, g: E \rightarrow \mathbb{R}$ 可微。证明

(a) (线性和乘积的 Leibniz 公式) $f + g, fg$ 可微, 且

$$\begin{aligned} d(f + g)(\mathbf{x}) &= df(\mathbf{x}) + dg(\mathbf{x}), \\ d(fg)(\mathbf{x}) &= f(\mathbf{x})dg(\mathbf{x}) + g(\mathbf{x})df(\mathbf{x}); \end{aligned}$$

(b) 若 $f(\mathbf{x}) \neq 0$, 则 $\frac{1}{f}$ 可微, 且

$$d\left(\frac{1}{f}\right)(\mathbf{x}) = -\frac{1}{f(\mathbf{x})^2}df(\mathbf{x})$$

6. (a) 设 $F, G: E \rightarrow \mathbb{R}^m$ 可微。证明

$$d(F, G)(\mathbf{x})(\mathbf{h}) = \langle F(\mathbf{x}), \partial G(\mathbf{x})\mathbf{h} \rangle + \langle G(\mathbf{x}), \partial F(\mathbf{x})\mathbf{h} \rangle.$$

- (b) 对任何 m 行 n 列矩阵 A , 定义 $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \mathbf{x}^T A^T A \mathbf{x}$, 求 $df(\mathbf{x})(\mathbf{h})$ 。

7. 设 $f: E \rightarrow \mathbb{R}$ 可微, 则 $\hat{f}: E \times F \rightarrow \mathbb{R}, \hat{f}(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})$ 可微, 且

$$d\hat{f}(\mathbf{x}, \mathbf{y})(\mathbf{h}, \mathbf{k}) = df(\mathbf{x})(\mathbf{h}).$$

8. 对任何 m 行 n 列矩阵 A , 定义 $f(A) = A^T A$, 求 $\partial f(A)(B)$ 。

9. 设 \mathcal{M}_n 是所有 n 阶实数矩阵组成的线性空间。证明: 行列式作为函数

$$\det: \mathcal{M}_n \rightarrow \mathbb{R}, \quad A \mapsto \det(A)$$

是可微函数, 并求 $d \det(A)(B)$ 。

10. 极坐标关于直角坐标的可微性。设 A 是直角坐标平面上一点, 且不是原点, A 的直角坐标为 (x_0, y_0) 。设正数 r_0 和实数 θ_0 满足

$$x_0 = r_0 \cos \theta_0, \quad y_0 = r_0 \sin \theta_0.$$

证明: 对任何满足 $\|B - A\| < r_0$ 的点 B , 存在唯一一对正数 $r = r(x, y)$ 和实数 $\theta = \theta(x, y)$ 满足它们关于 (x, y) 连续, 且

$$x = r \cos \theta, \quad y = r \sin \theta,$$

以及 $r(x_0, y_0) = r_0, \theta(x_0, y_0) = \theta_0$, 并且 $r(x, y)$ 和 $\theta(x, y)$ 关于 (x, y) 是可微函数, 并求它们的微分。这里 (x, y) 是点 B 的直角坐标。

11. 我们知道: 如果一元函数在一点处存在左右导数, 则它在该点连续; 如果左右导数相等, 则它在该点可微。问: 是否存在这样的函数: 它在一点沿每个方向的方向导数都为零, 但它在这点并不连续? 这样的例子表明多元函数与一元函数有本质不同。

2.2 坐标系与偏导数

在上节中我们引进的导数、微分、沿向量或方向的方向导数这些概念都不依赖于(自变量)空间的坐标系。笛卡尔引入坐标系的概念后, 复杂的推理可以由计算来代替。如果在有限维空间中引入坐标系, 那么点和向量都可以用一组数(即坐标)来表示, 这会使计算变得更容易操作。但是, 坐标系的引进也会产生一些问题, 同一对象在不同的坐标系下的表现形式不一样, 在一定程度上这会掩盖事物的本质特征。所以坐标系的引入既有有利的一面, 也有不利的一面。

定义 2.2.1 (偏导数). 设 (x_1, x_2, \dots, x_m) 为 E 中点的坐标, 在 E 的内点 $\mathbf{x}_0 = (a_1, \dots, a_m)$ 处 m 元函数 $f: E \rightarrow \mathbb{R}$ 存在以下极限

$$\lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_k + t, \dots, a_m) - f(a_1, \dots, a_k, \dots, a_m)}{t},$$

则把这个极限的值记为 $\frac{\partial f}{\partial x_k}(\mathbf{x}_0)$, $f'_{x_k}(\mathbf{x}_0)$ (有时甚至写成 $f_{x_k}(\mathbf{x}_0)$), $\partial_{x_k} f(\mathbf{x}_0)$ 或 $\partial_k f(\mathbf{x}_0)$, 并称为 f 关于坐标 x_k 的偏导数。

注 2.2.2. (1) 由上述定义知 f 对坐标 x_k 的偏导数正是 f 在 \mathbf{x}_0 处沿基底向量 $\mathbf{e}_k = (0, \dots, 1, \dots, 0)^T$ (即第 k 位是 1, 其他是 0) 的导数:

$$\frac{\partial f}{\partial x_k}(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_k) - f(\mathbf{x}_0)}{t} = \frac{\partial f}{\partial \mathbf{e}_k}(\mathbf{x}_0) = -\frac{\partial f}{\partial (-\mathbf{e}_k)}(\mathbf{x}_0).$$

(2) 定义偏导数时, 无需空间具备范数。

(3) 偏导数的概念与坐标系有关, 所以物理学家更倾向于使用形如

$$\left(\frac{\partial f}{\partial x_1} \right)_{x_2, \dots, x_n}(\mathbf{x}_0)$$

的符号, 它表明 f 是 n 个独立变量 x_1, x_2, \dots, x_n 的函数, 这个导数是当 x_2, \dots, x_n 固定时, 仅对变化的 x_1 求的导数。

例 2.2.3 (有偏导数, 但并非沿所有向量都有导数的函数的例子).

$$f(x, y) = \begin{cases} 1, & \text{若 } xy = 0; \\ 0, & \text{否则.} \end{cases}$$

偏导数 $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, 但沿其他向量没有导数。 \square

定理 2.2.4. 设 m 元函数 f 在 \mathbf{x}_0 处可微, 则对于任何 $\mathbf{v} = (\xi_1, \dots, \xi_m)^T \in \mathbb{R}^m$,

$$\begin{aligned} df(\mathbf{x}_0)(\mathbf{v}) &= \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \xi_1 \frac{\partial f}{\partial x_1}(\mathbf{x}_0) + \dots + \xi_m \frac{\partial f}{\partial x_m}(\mathbf{x}_0) \\ &= \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_m}(\mathbf{x}_0) \right) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}. \end{aligned}$$

因此函数 f 的微分为

$$df(\mathbf{x}_0) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0)dx_1 + \dots + \frac{\partial f}{\partial x_m}(\mathbf{x}_0)dx_m,$$

其中

$$dx_i: \mathbb{R}^m \rightarrow \mathbb{R}, \quad dx_i(\mathbf{v}) = \xi_i,$$

是坐标函数 $x_i: \mathbb{R}^m \rightarrow \mathbb{R}, \quad (x_1, \dots, x_m) \mapsto x_i$ 的微分。

证明.

$$\begin{aligned} df(\mathbf{x}_0)(\mathbf{v}) &= df(\mathbf{x}_0)(\xi_1 \mathbf{e}_1 + \cdots + \xi_m \mathbf{e}_m) \\ &= \xi_1 df(\mathbf{x}_0)(\mathbf{e}_1) + \cdots + \xi_m df(\mathbf{x}_0)(\mathbf{e}_m) \\ &= \xi_1 \frac{\partial f}{\partial x_1}(\mathbf{x}_0) + \cdots + \xi_m \frac{\partial f}{\partial x_m}(\mathbf{x}_0) \\ &= \frac{\partial f}{\partial x_1}(\mathbf{x}_0) dx_1(\mathbf{v}) + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{x}_0) dx_m(\mathbf{v}). \end{aligned}$$

□

注 2.2.5. 给定 m 维线性空间 V 的一组基底向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$, 则对任何 $\mathbf{v} \in \mathbb{R}^m$, 都存在唯一一组实数 x_1, x_2, \dots, x_m 使得

$$\mathbf{v} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_m \mathbf{v}_m.$$

于是 V 与 \mathbb{R}^m 线性同构. 这组实数 x_1, x_2, \dots, x_m 称为 $\mathbf{v} \in \mathbb{R}^m$ 在基底向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ 下的坐标. 映向量 \mathbf{v} 为其坐标 x_i 的函数是 V 上的一个线性函数, 记为 dx_i , 所以对任何线性函数 L ,

$$L(\mathbf{v}) = x_1 L(\mathbf{v}_1) + x_2 L(\mathbf{v}_2) + \cdots + x_m L(\mathbf{v}_m) = \sum_{i=1}^m L(\mathbf{v}_i) dx_i(\mathbf{v}),$$

从而

$$L = L(\mathbf{v}_1) dx_1 + L(\mathbf{v}_2) dx_2 + \cdots + L(\mathbf{v}_m) dx_m.$$

所以在 \mathbb{R}^m 上的所有线性函数组成的线性空间中, dx_1, dx_2, \dots, dx_m 组成一组基底, 我们称它是基底向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ 的对偶基. 在这组基底, 我们把线性函数 L 对应于行向量

$$(L(\mathbf{v}_1), L(\mathbf{v}_2), \dots, L(\mathbf{v}_m)),$$

因而行向量

$$\left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_m}(\mathbf{x}_0) \right)$$

就是在坐标系 (x_1, x_2, \dots, x_m) 的对偶基底 $(dx_1, dx_2, \dots, dx_m)$ 下, 微分 $df(\mathbf{x}_0)$ 的坐标.

推论 2.2.6. 若 m 元映射 $F: E \rightarrow \mathbb{R}^n$,

$$F(x_1, \dots, x_m) = (f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m))^T$$

在 \mathbf{x}_0 处可微, 则对于任何 $\mathbf{v} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$,

$$\partial F(\mathbf{x}_0)(\mathbf{v}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_m}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_n}{\partial x_m}(\mathbf{x}_0) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}.$$

因此映射 f 的微分的坐标表示为

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_m}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_n}{\partial x_m}(\mathbf{x}_0) \end{pmatrix},$$

记这个矩阵为 $JF(\mathbf{x}_0) = \left. \frac{\partial(y_1, \dots, y_m)}{\partial(x_1, \dots, x_m)} \right|_{\mathbf{x}_0}$, 称为 F 在 \mathbf{x}_0 处的 **Jacobi 矩阵**. $JF(\mathbf{x}_0)$ 的行列式 $\det JF(\mathbf{x}_0)$ 称为 F 在 \mathbf{x}_0 处的 **Jacobi 行列式**. \square

定理 2.2.7. 设 $E \subseteq \mathbb{R}^n$, 映射 $f: E \rightarrow \mathbb{R}^m$ 在 \mathbf{x}_0 处可微, $\mathbf{v}_1, \dots, \mathbf{v}_n$ 是 \mathbb{R}^n 的一组基底向量, (x_1, \dots, x_n) 是相应的坐标系, $\mathbf{w}_1, \dots, \mathbf{w}_m$ 是 \mathbb{R}^m 的一组基底向量, (y_1, \dots, y_m) 是相应的坐标系, 即

$$f(\mathbf{x}_0 + x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n) = f(\mathbf{x}_0) + y_1(x_1, \dots, x_n)\mathbf{w}_1 + \cdots + y_m(x_1, \dots, x_n)\mathbf{w}_m,$$

则对向量 $\mathbf{v} = \xi_1\mathbf{v}_1 + \cdots + \xi_n\mathbf{v}_n$,

$$\partial f(\mathbf{x}_0)(\mathbf{v}) = (\mathbf{w}_1, \dots, \mathbf{w}_m) \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

称矩阵

$$Jf(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

为 f 在 \mathbf{x}_0 处关于基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 和 $\mathbf{w}_1, \dots, \mathbf{w}_m$ 或关于坐标系 (x_1, \dots, x_n) 和 (y_1, \dots, y_m) 的 **Jacobi 矩阵**.

证明.

$$\begin{aligned} \partial f(\mathbf{x}_0)(\mathbf{v}) &= \sum_{j=1}^n \xi_j \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = \sum_{j=1}^n \xi_j \sum_{i=1}^m \frac{\partial y_i}{\partial x_j}(\mathbf{x}_0) \mathbf{w}_i = \sum_{i=1}^m \left(\sum_{j=1}^n \frac{\partial y_i}{\partial x_j}(\mathbf{x}_0) \xi_j \right) \mathbf{w}_i \\ &= (\mathbf{w}_1, \dots, \mathbf{w}_m) \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}. \end{aligned}$$

\square

推论 2.2.8 (链索法则的矩阵形式). 设 G 在 \mathbf{x}_0 处可微, F 在 $\mathbf{y}_0 = G(\mathbf{x}_0)$ 处可微, 则

$$J(F \circ G)(\mathbf{x}_0) = JF(\mathbf{y}_0) \cdot JG(\mathbf{x}_0).$$

设 $\mathbf{y} = G(\mathbf{x})$, $\mathbf{u} = F(\mathbf{y})$, 则上式的分量形式为

$$\partial_j(F \circ G)_i(\mathbf{x}_0) = \sum_k \partial_k F_i(\mathbf{y}_0) \partial_j G_k(\mathbf{x}_0) = \sum_k \partial_k F_i(G(\mathbf{x}_0)) \partial_j G_k(\mathbf{x}_0),$$

用传统符号写为

$$\frac{\partial u_i}{\partial x_j}(\mathbf{x}_0) = \sum_k \frac{\partial u_i}{\partial y_k}(\mathbf{y}_0) \frac{\partial y_k}{\partial x_j}(\mathbf{x}_0) = \sum_k \frac{\partial u_i}{\partial y_k}(G(\mathbf{x}_0)) \frac{\partial y_k}{\partial x_j}(\mathbf{x}_0).$$

如果可微映射 F 的逆映射 F' 也是可微的, 则

$$J(F^{-1})(\mathbf{y}_0) = (JF(\mathbf{x}_0))^{-1}.$$

□

设 $\mathbf{u} = G(\mathbf{x})$ 是从坐标系 (x_1, \dots, x_n) 空间到坐标系 (u_1, \dots, u_n) 空间的一个可微变换(即 G 可逆, 且 G 和逆映射 G^{-1} 都可微)。 $y = f(\mathbf{u})$ 是坐标系 (u_1, \dots, u_n) 空间上的一个可微函数, 则根据链索法则, $y = f(G(\mathbf{x}))$ 可微, 并且

$$\begin{aligned} d(f \circ G)(\mathbf{x}_0)(\mathbf{v}) &= \sum_{k=1}^n \frac{\partial(f \circ G)}{\partial x_k}(\mathbf{x}_0) \xi_k \\ &= \sum_{k=1}^n \sum_{j=1}^n \left[\frac{\partial f}{\partial u_j}(\mathbf{u}_0) \frac{\partial u_j}{\partial x_k}(\mathbf{x}_0) \right] \xi_k \\ &= \sum_{j=1}^n \left[\frac{\partial f}{\partial u_j}(\mathbf{u}_0) \sum_{k=1}^n \frac{\partial u_j}{\partial x_k}(\mathbf{x}_0) \xi_k \right] \\ &= \sum_{j=1}^n \frac{\partial f}{\partial u_j}(\mathbf{u}_0) du_j(\mathbf{x}_0)(\mathbf{v}), \end{aligned}$$

所以

$$d(f \circ G)(\mathbf{x}_0) = \sum_{j=1}^n \frac{\partial f}{\partial u_j}(\mathbf{u}_0) du_j(\mathbf{x}_0),$$

即

$$dy = \sum_{k=1}^n \frac{\partial y}{\partial x_k} dx_k = \sum_{j=1}^n \frac{\partial y}{\partial u_j} du_j.$$

最后这个等式叫做一阶微分的形式不变性, 它表明可微函数的微分与自变量空间中坐标系选取无关。

例 2.2.9. 对直角坐标系 (x, y) 下的二元可微函数 $u = f(x, y)$, 求它在极坐标系 (r, θ) 下的偏导数 $\frac{\partial u}{\partial r}, \frac{\partial u}{\partial \theta}$ 。

解: (利用微分的形式不变性) $x = r \cos \theta$, $y = r \sin \theta$, 所以

$$\begin{aligned} dx &= \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta = \cos \theta dr - r \sin \theta d\theta, \\ dy &= \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta = \sin \theta dr + r \cos \theta d\theta, \end{aligned}$$

另由一阶微分的形式不变性得到

$$\begin{aligned} \frac{\partial u}{\partial r} dr + \frac{\partial u}{\partial \theta} d\theta &= du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \\ &= \frac{\partial u}{\partial x} [\cos \theta dr - r \sin \theta d\theta] + \frac{\partial u}{\partial y} [\sin \theta dr + r \cos \theta d\theta] \\ &= \left[\frac{\partial u}{\partial x} \cos \theta + \frac{\partial u}{\partial y} \sin \theta \right] dr + \left[r \cos \theta \frac{\partial u}{\partial y} - r \sin \theta \frac{\partial u}{\partial x} \right] d\theta. \end{aligned}$$

所以

$$\begin{aligned} \frac{\partial u}{\partial r}(r, \theta) &= \cos \theta \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta), \\ \frac{\partial u}{\partial \theta}(r, \theta) &= r \cos \theta \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) - r \sin \theta \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta). \end{aligned}$$

□

另解： (利用链索法则的Jacobi矩阵形式，但注意不能随意交换矩阵前后顺序)

$$\begin{aligned} \left(\frac{\partial u}{\partial r}, \frac{\partial u}{\partial \theta} \right) &= \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} \\ &= \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \\ &= \left(\cos \theta \frac{\partial u}{\partial x} + \sin \theta \frac{\partial u}{\partial y}, -r \sin \theta \frac{\partial u}{\partial x} + r \cos \theta \frac{\partial u}{\partial y} \right) \end{aligned}$$

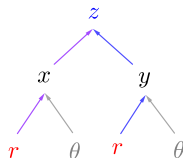
或者

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial r}, \quad \frac{\partial u}{\partial \theta} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial \theta}.$$

□

另解： 多元复合函数求导总显得形式比较复杂，初学者往往不容易掌握。这里我们一种叫做“变量树”的办法。

引入变量 $z = f(x, y)$, $x = x(r, \theta)$, $y = y(r, \theta)$, 并把因变量、中间变量和自变量依次作为根节点和叶节点，按函数嵌套顺序分层排列，形成一棵变量的树状图。为了计算偏导数 $\frac{\partial z}{\partial r}$, 需要完成以下步骤：



1. 从根节点 z 到最终叶节点 r 总共有两条道路： $z \leftarrow x \leftarrow r$ 和 $z \leftarrow y \leftarrow r$ 。

2. 沿每条道路把相邻两个变量之间的偏导数依次相乘:

$$\frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial r}, \quad \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial r}.$$

3. 把所有道路的结果相加, 就得到

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial r}.$$

我们利用 $x = r \cos \theta$ 和 $y = r \sin \theta$ 得到

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

于是

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

读者可以用这个办法自己计算 $\frac{\partial z}{\partial \theta}$. □

我们业已看到, 函数的可微性以及微分是与坐标系选取无关的性质, 但是作为线性函数的微分毕竟显得比较抽象且不利于操作. 在给定坐标系后, 偏导数更方便计算, 我们可以利用关于坐标系的偏导数来表示微分. 然而, 存在偏导数只是函数可微的必要而非充分条件. 下面的定理说, 如果所有一阶偏导数都连续, 那么函数一定是可微的. 偏导数连续是比可微定义更容易验证的条件.

定理 2.2.10. 若所有一阶偏导函数 $\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x})$ 都连续, 则 f 可微.

证明. (仅以二元函数为例) 设 $f(x, y)$ 在含 (a, b) 的一个开集上定义, 且 $\frac{\partial f}{\partial x}(a, b)$ 存在, $\frac{\partial f}{\partial y}(x, y)$ 在 (a, b) 的一个邻域里存在, 且在 (a, b) 处连续.

对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得当 $|x - a| < \delta$ 且 $|y - b| < \delta$ 时,

$$\left| f(x, b) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) \right| \leq \varepsilon|x - a|,$$

且

$$\left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(a, b) \right| < \varepsilon.$$

于是

$$\begin{aligned} & \left| f(x, y) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) - \frac{\partial f}{\partial y}(a, b)(y - b) \right| \\ & \leq \left| f(x, y) - f(x, b) - \frac{\partial f}{\partial y}(a, b)(y - b) \right| + \left| f(x, b) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) \right| \\ & \leq \left| \frac{\partial f}{\partial y}(x, b + t(y - b))(y - b) - \frac{\partial f}{\partial y}(a, b)(y - b) \right| + \varepsilon|x - a| \\ & = \left| \frac{\partial f}{\partial y}(x, b + t(y - b)) - \frac{\partial f}{\partial y}(a, b) \right| |y - b| + \varepsilon|x - a| \\ & \leq \varepsilon|y - b| + \varepsilon|x - a| \end{aligned}$$

所以当 $(x, y) \rightarrow (a, b)$ 时,

$$f(x, y) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) - \frac{\partial f}{\partial y}(a, b)(y - b) = o(|x - a| + |y - b|).$$

因此 f 在 (a, b) 可微。 \square

习题2.2

- 关于偏导数的符号。传统上,人们习惯不区分变量和函数。例如在函数 $z = f(x, y)$ 中,人们把因变量 z 等同于函数 f ,甚至写成 $z = z(x, y)$ 的形式,因而 $\frac{\partial z}{\partial x}$ 也就是 $\frac{\partial f}{\partial x}$ 。但这在复合函数时会导致歧义,比如对 $z = z(x, xy)$,符号 $\frac{\partial z}{\partial x}$ 表达的是什么含义,谈谈你的理解以及就如何避免这种歧义的发生说说你的看法。

- 设 $f: \mathbb{R}^m \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ 是可微函数。证明 f 是 k 次齐次函数,即

$$f(t\mathbf{x}) = t^k f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}, \forall t > 0,$$

当且仅当

$$\sum_{i=1}^m x_i \frac{\partial f}{\partial x_i}(\mathbf{x}) = k f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}.$$

- 设 \mathcal{M}_n 是所有 n 阶实数矩阵组成的线性空间。求 $\frac{\partial \det}{\partial a_{i,j}}(A)$, 并证明:行列式作为函数

$$\det: \mathcal{M}_n \rightarrow \mathbb{R}, \quad A \mapsto \det(A)$$

是可微函数, 并求 $d \det(A)(B)$ 。

- 设 A 是 n 阶正交矩阵, $\mathbf{y} = A\mathbf{x}$, 证明对任何可微函数 $u = u(\mathbf{y})$,

$$\left(\frac{\partial u}{\partial y_1}\right)^2 + \cdots + \left(\frac{\partial u}{\partial y_n}\right)^2 = \left(\frac{\partial \tilde{u}}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial \tilde{u}}{\partial x_n}\right)^2,$$

其中 $\tilde{u}(\mathbf{x}) = u(A\mathbf{x})$ 。这说明在欧氏度量下,微分的大小与单位正交基底的选取无关。

- 对实数 a, b, μ 和正数 σ , 记

$$f(\mu, \sigma, a, b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

证明 f 是可微函数, 并求它的微分。

- 理想气体状态方程

$$PV = nRT$$

描述了气体三个宏观物理量压强 P ，体积 V 和温度 T 之间的关系，其中 n （气体分子数）和 R （理想气体普适常数）都可视为常数。于是这三个物理量中的任何一个都可以看成是由另外两个物理量决定的函数。证明：

$$\left(\frac{\partial P}{\partial V}\right)_T \cdot \left(\frac{\partial V}{\partial P}\right)_T \cdot \left(\frac{\partial T}{\partial P}\right)_V = -1.$$

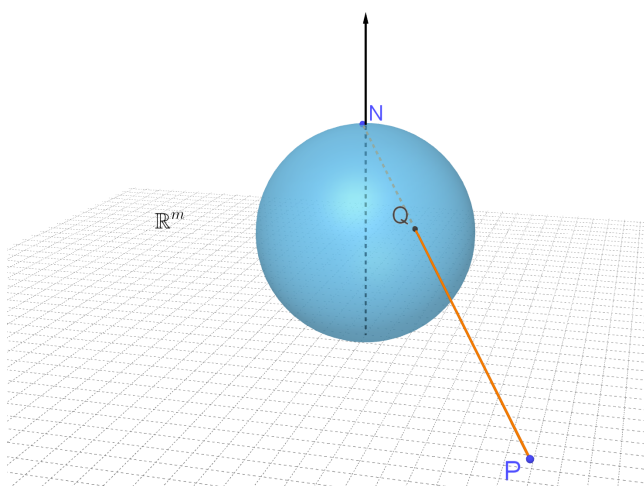
这个例子表明，偏导数符号是一个整体，不能望文生义地理解为分式。另外这里 $\left(\frac{\partial z}{\partial x}\right)_y$ 是物理学家在热力学计算时使用的符号，它表示 z 是以 x, y 为独立自变量的函数，当 y 固定时， z 对 x 的(偏)导数。

7. 设 $f: \mathbb{R} \rightarrow \mathbb{R}$ 是 C^1 函数，定义

$$F: \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad F(\mathbf{x}) = f(\|\mathbf{x}\|_2)\mathbf{x}.$$

求 $\det \partial F(\mathbf{x})$ 。

8. 球极投影。记



$$\mathbb{R}^m \times \mathbb{R} = \{(X^1, \dots, X^m, X^{m+1}) \mid X^k \in \mathbb{R}\},$$

$$\mathbb{R}^m = \{(x^1, \dots, x^m) \mid x^k \in \mathbb{R}\}.$$

记 $N = (0, \dots, 0, 2) \in \mathbb{R}^m \times \mathbb{R}$, $S^m = \{(X^1, \dots, X^m, X^{m+1}) \mid (X^1)^2 + \dots + (X^m)^2 + (X^{m+1} - 1)^2 = 1\}$ 。对 $(x^1, \dots, x^m) \in \mathbb{R}^m$ ，连接 $(x^1, \dots, x^m, 0)$ 与 N 的直线交 S^m 于 $(X^1, \dots, X^m, X^{m+1})$ 。证明映射

$$(X^1(x^1, \dots, x^m), \dots, X^m(x^1, \dots, x^m), X^{m+1}(x^1, \dots, x^m))$$

是可微映射，并求它的 Jacobi 矩阵。

2.3 内积空间中可微函数的梯度与方向导数

对线性函数和可微函数，我们可以引进梯度向量的概念，它的作用相当于传统一元微积分时作为系数的导数。

定理 2.3.1 (Riesz表示定理). 设 $\langle \cdot, \cdot \rangle$ 是 \mathbb{R}^n 上一个内积。则对 \mathbb{R}^n 上任何一个线性函数 $L: \mathbb{R}^n \rightarrow \mathbb{R}$ ，都存在唯一一个向量 $\nabla L \in \mathbb{R}^n$ 使得

$$L(\mathbf{v}) = \langle \mathbf{v}, \nabla L \rangle, \quad \forall \mathbf{v} \in \mathbb{R}^n.$$

从而 ∇L 与 $\text{Ker}L$ 正交，且 $\|\nabla L\| = \|L\| = \max_{\|\mathbf{v}\|=1} L(\mathbf{v})$ ，单位向量 \mathbf{v} 满足 $L(\mathbf{v}) = \|\nabla L\|$ 当且仅当 \mathbf{v} 是 ∇L 的方向向量。称这个唯一的向量 ∇L 为线性函数 L 的梯度向量。

证明. 对 $L \equiv 0$ ，易见 $\nabla L = \mathbf{0}$ 是唯一满足定理结论中等式的向量。

下设 $L \neq 0$ 。于是取单位向量 \mathbf{u}^* 使得 $\mathbf{u}^* \perp \text{Ker}L$ ，取 $\nabla L = L(\mathbf{u}^*)\mathbf{u}^*$ ，则对任意 $\mathbf{v} \in \mathbb{R}^n$ ，易见 $\mathbf{w} := \mathbf{v} - \frac{L(\mathbf{v})}{L(\mathbf{u}^*)}\mathbf{u}^* \in \text{Ker}L$ ，

$$\langle \mathbf{v}, \nabla L \rangle = \langle \mathbf{v}, L(\mathbf{u}^*)\mathbf{u}^* \rangle = \left\langle \frac{L(\mathbf{v})}{L(\mathbf{u}^*)}\mathbf{u}^*, L(\mathbf{u}^*)\mathbf{u}^* \right\rangle = L(\mathbf{v}).$$

请读者自己证明唯一性。对任意单位向量 \mathbf{v} ，由 Cauchy-Schwarz 不等式，

$$L(\mathbf{v}) = \langle \mathbf{v}, \nabla L \rangle \leq \|\mathbf{v}\| \|\nabla L\| \leq \|\nabla L\|,$$

其中等式成立当且仅当 \mathbf{v} 是 ∇L 的方向向量。所以 $\|L\| = \|\nabla L\|$ 。 \square

定义 2.3.2. 设 E 是内积空间的子集，函数 $f: E \rightarrow \mathbb{R}$ 在 \mathbf{x}_0 处可微。记 $df(\mathbf{x}_0)$ 的梯度向量为 $\text{grad}f(\mathbf{x}_0)$ 或 $\nabla f(\mathbf{x}_0)$ ，称为 f 在 \mathbf{x}_0 处的梯度向量。于是

$$df(\mathbf{x}_0)(\mathbf{v}) = \langle \mathbf{v}, \nabla f(\mathbf{x}_0) \rangle.$$

例 2.3.3. 设 $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是线性映射。求 $f(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle$ 的微分与梯度。

解:

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{v}) &= \langle A\mathbf{x}_0 + A\mathbf{v}, \mathbf{x}_0 + \mathbf{v} \rangle = \langle A\mathbf{x}_0, \mathbf{x}_0 \rangle + \langle A\mathbf{x}_0, \mathbf{v} \rangle + \langle A\mathbf{v}, \mathbf{x}_0 \rangle + \langle A\mathbf{v}, \mathbf{v} \rangle \\ &= f(\mathbf{x}_0) + \langle (A + A^T)\mathbf{x}_0, \mathbf{v} \rangle + o(\|\mathbf{v}\|), \quad \mathbf{v} \rightarrow 0, \end{aligned}$$

所以 $df(\mathbf{x}_0)(\mathbf{v}) = \langle (A + A^T)\mathbf{x}_0, \mathbf{v} \rangle$ ， $\nabla f(\mathbf{x}_0) = (A + A^T)\mathbf{x}_0$ 。

特别当 $A = I$ 时， $df(\mathbf{x}_0)(\mathbf{v}) = \langle 2\mathbf{x}_0, \mathbf{v} \rangle$ ， $\nabla f(\mathbf{x}_0) = 2\mathbf{x}_0$ 。 \square

定理 2.3.4. 设 f 在 \mathbf{x}_0 处可微, 则对于任意向量 $\mathbf{v} \in \mathbb{R}^n$,

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = df(\mathbf{x}_0)(\mathbf{v}) = \langle \mathbf{v}, \nabla f(\mathbf{x}_0) \rangle.$$

从而对任何单位向量 \mathbf{v} ,

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) \leq \|\nabla f(\mathbf{x}_0)\|,$$

等号成立当且仅当 \mathbf{v} 是 f 在 \mathbf{x}_0 处的梯度方向。所以 f 的梯度方向是 f 增长最快的方向。□

注 2.3.5. 参照上述定理的最终结论, 我们可以在一般范数情形下定义 f 的梯度向量。设 $\|\cdot\|$ 是 \mathbb{R}^m 上一个范数, $f: \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个可微函数。于是 f 的微分 $df(\mathbf{x}_0)$ 是 \mathbb{R}^m 上一个线性函数。函数 f 在 \mathbf{x}_0 处的梯度为这样一个向量, $\nabla f(\mathbf{x}_0) = \lambda \mathbf{v}^*$, 其中 $\|\mathbf{v}^*\| = 1$ 且

$$\lambda = df(\mathbf{x}_0)\mathbf{v}^* = \max_{\|\mathbf{v}\|=1} df(\mathbf{x}_0)\mathbf{v}.$$

我们后面会知道, 从 \mathbf{x}_0 出发沿梯度 $\nabla f(\mathbf{x}_0)$ 的方向, 函数值 $f(\mathbf{x})$ 上升最快, 函数值的最大增长率就是梯度的大小。但是, 在没有内积的情况下, 我们无法借助梯度计算函数沿其他方向的导数。

定理 2.3.6. 在直角坐标系 (x_1, \dots, x_m) 下, 可微函数 f 的梯度为

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_m} \right)^T.$$

证明. 设 $\mathbf{e}_1, \dots, \mathbf{e}_m$ 是这个坐标系对应的单位正交基底, 并且

$$\nabla f = c_1 \mathbf{e}_1 + \dots + c_m \mathbf{e}_m.$$

则对任意 $\mathbf{v} = \xi_1 \mathbf{e}_1 + \dots + \xi_m \mathbf{e}_m$,

$$\sum_{i=1}^m \frac{\partial f}{\partial x_i}(\mathbf{x}) \xi_i = df(\mathbf{x})(\mathbf{v}) = \langle \nabla f, \mathbf{v} \rangle = \left\langle \sum_{i=1}^m c_i \mathbf{e}_i, \sum_{j=1}^m \xi_j \mathbf{e}_j \right\rangle = \sum_{i=1}^m c_i \xi_i,$$

所以 $c_i = \frac{\partial f}{\partial x_i}(\mathbf{x})$, 从而

$$\nabla f(\mathbf{x}) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\mathbf{x}) \mathbf{e}_i.$$

□

注 2.3.7. (1) 读者需要注意的是: 与作为线性函数的微分不同, 梯度是函数定义域所在空间中的一个向量, 所以在给定坐标系下梯度是一个列向量。

(2) 因为函数的微分和空间中的内积都与坐标系无关, 所以梯度本质上是与坐标系无关的一个向量。但在不同的坐标系中, 梯度的坐标会不同。在一般坐标系下, 梯度 $\nabla f(\mathbf{x}_0)$ 的坐标表达式见本节后的习题。

习题2.3

1. (a) 设 L 是内积空间上一个非零的线性函数, 非零向量 $\mathbf{n} \in (\text{Ker}L)^\perp$ 满足 $L(\mathbf{n}) = \|\mathbf{n}\|^2$, 证明 $\nabla L = \mathbf{n}$.
- (b) 设 $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ 是 \mathbb{R}^n 中的 $n-1$ 个(列)向量, 记

$$L(\mathbf{v}) = \det(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{v}).$$

求线性函数 L 的梯度向量 ∇L .

2. 设 $f, g: E \rightarrow \mathbb{R}$ 可微。证明

(a) (线性和乘积的 Leibniz 公式) $f+g, fg$ 可微, 且

$$\begin{aligned}\nabla(f+g)(\mathbf{x}) &= \nabla f(\mathbf{x}) + \nabla g(\mathbf{x}), \\ \nabla(fg)(\mathbf{x}) &= f(\mathbf{x})\nabla g(\mathbf{x}) + g(\mathbf{x})\nabla f(\mathbf{x});\end{aligned}$$

(b) 若 $f(\mathbf{x}) \neq 0$, 则 $\frac{1}{f}$ 可微, 且

$$\nabla\left(\frac{1}{f}\right)(\mathbf{x}) = -\frac{1}{f(\mathbf{x})^2}\nabla f(\mathbf{x}).$$

3. 设 $F, G: E \rightarrow \mathbb{R}^m$ 可微。证明

$$\nabla\langle F, G\rangle(\mathbf{x}) = (\partial G(\mathbf{x}))^T F(\mathbf{x}) + (\partial F(\mathbf{x}))^T G(\mathbf{x}).$$

4. 设 $f: \mathbb{R}^m \rightarrow \mathbb{R}$ 是可微函数, $\mathbf{v}_1, \dots, \mathbf{v}_m$ 是 \mathbb{R}^m 的一组基, (x_1, \dots, x_m) 是对应这组基的坐标系。证明在这个坐标系下, $\nabla f(\mathbf{x}_0)$ 的坐标为

$$\begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_m \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{v}_m, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_m, \mathbf{v}_m \rangle \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{pmatrix}.$$

5. 设 $f: \mathbb{R}^m \rightarrow \mathbb{R}$ 是可微函数, $G: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是可微映射, 求 $f \circ G$ 的梯度。
6. 可微函数 $f(x, y)$ 在平面极坐标系 (r, θ) 下的梯度。
7. 设 M_n 为所有 n 阶实数矩阵组成的线性空间, $\det: M_n \rightarrow \mathbb{R}(A \mapsto \det A, \det A$ 为矩阵 A 的行列式)。

(a) 证明 \det 是可微函数, 并求 \det 在 $A \in M_n$ 处的微分。

(b) 证明 $\langle A, B \rangle = \text{tr}(A^T B)$ 是 M_n 上一个内积, 并求 \det 在 $A \in M_n$ 处的梯度。

8. 分别求线性函数 $f(x, y) = 2x + 3y$ 在以下范数下的梯度向量。

(a) $\sqrt{x^2 + y^2}$ (b) $\max\{|x|, |y|\}$ (c) $|x| + |y|$

9. 设 $E \subset \mathbb{R}^m$ 是开集, 并且包含从 \mathbf{a} 到 \mathbf{b} 的直线段, $f: E \rightarrow \mathbb{R}^p$ 是可微映射。

(a) 若 $p = 1$, 证明在 \mathbf{a} 到 \mathbf{b} 的直线段上存在 \mathbf{c} 使得

$$f(\mathbf{b}) - f(\mathbf{a}) = df(\mathbf{c})(\mathbf{b} - \mathbf{a}) = \langle \nabla f(\mathbf{c}), \mathbf{b} - \mathbf{a} \rangle.$$

(b) 若 $p > 1$, 证明对任意向量 $\mathbf{u} \in \mathbb{R}^p$, 在 \mathbf{a} 到 \mathbf{b} 的直线段上存在 \mathbf{c}_u 使得

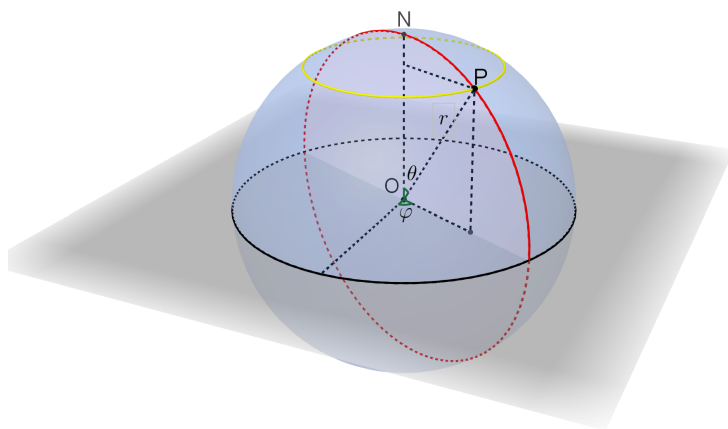
$$\langle \mathbf{u}, f(\mathbf{b}) - f(\mathbf{a}) \rangle = \langle \mathbf{u}, \partial f(\mathbf{c}_u)(\mathbf{b} - \mathbf{a}) \rangle.$$

(c) 若 $\|\partial f(\mathbf{x})\| \leq M (\forall \mathbf{x} \in E)$, 则

$$\|f(\mathbf{b}) - f(\mathbf{a})\| \leq M \|\mathbf{b} - \mathbf{a}\|.$$

(提示: 在(b)中取 $\mathbf{u} = f(\mathbf{b}) - f(\mathbf{a})$.)

10. \mathbb{R}^3 中的球坐标系。其中极径 r , 纬度 θ 和经度 φ 如图所示。



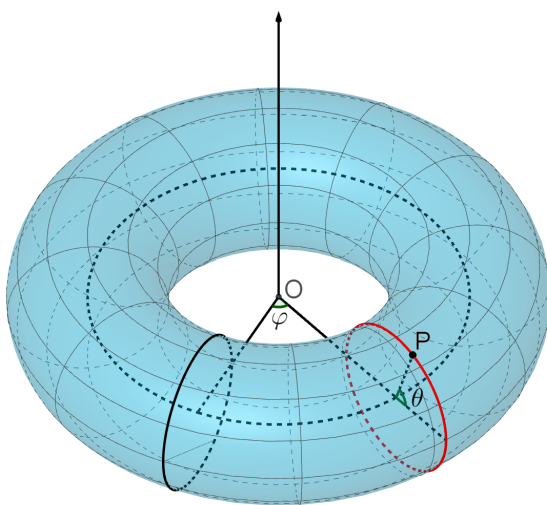
(a) 试写出用球坐标 (r, θ, φ) 表达的直角坐标 $(x(r, \theta, \varphi), y(r, \theta, \varphi), z(r, \theta, \varphi))$ 的表达式, 并证明它是可微映射, 求出该映射的 Jacobi 矩阵。该矩阵是否可逆? 为什么?

(b) 记

$$\mathbf{e}_r = \begin{pmatrix} \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial r} \end{pmatrix}, \quad \mathbf{e}_\theta = \begin{pmatrix} \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \theta} \end{pmatrix}, \quad \mathbf{e}_\varphi = \begin{pmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{pmatrix},$$

求它们给出的度量矩阵。

- (c) 计算 $(x(r, \theta, \varphi), y(r, \theta, \varphi), z(r, \theta, \varphi))$ 的 Jacobi 矩阵的行列式, 并对结果给出几何解释。
- (d) 对可微函数 $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, 求在球坐标系下 f 的梯度 ∇f (用偏导数 $\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \varphi}$ 表示)。
- (e) 记 $S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$ 。如果 $f(\theta, \varphi)$ 是 S^2 上的可微函数, 求 ∇f (用偏导数 $\frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \varphi}$ 表示)。
11. 环面上, 纬度 θ 和经度 φ 如图所示。设环面轴径为 R , 经线半径为 b , $0 < b < a$ 。



- (a) 试写出用环面坐标 (θ, φ) 表达的直角坐标 $(x(\theta, \varphi), y(\theta, \varphi), z(\theta, \varphi))$ 的表达式, 并证明它是可微映射, 求出该映射的 Jacobi 矩阵。该矩阵是否可逆? 为什么?

- (b) 记

$$\mathbf{e}_\theta = \begin{pmatrix} \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \theta} \end{pmatrix}, \quad \mathbf{e}_\varphi = \begin{pmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{pmatrix},$$

求它们给出的度量矩阵。

- (c) 如果 $f(\theta, \varphi)$ 是环面上的可微函数, 求 ∇f (用偏导数 $\frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \varphi}$ 表示)。

2.4 高阶偏导数和 Taylor 展开

导数和微分为我们提供了一个可微函数或映射的线性近似, 而这些非线性函数和映射的非线性特征则需要通过高阶导数来体现。Taylor 公式为我们提供了通过函数的高阶导数用多项式形式近似非线性可微函数的一个途径。

定义 2.4.1. 对 $i_1, \dots, i_k \in \{1, 2, \dots, n\}$, 记

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(\mathbf{x}_0) = \frac{\partial}{\partial x_{i_k}} \left(\cdots \frac{\partial f}{\partial x_{i_1}} \right) (\mathbf{x}_0).$$

称为 f 在 \mathbf{x}_0 处的一个 k 阶偏导数。

高阶偏导数有时也用这样的符号: $\partial_{x_{i_k}, \dots, x_{i_1}}^k f$, $\partial_{i_k, \dots, i_2, i_1}^k f$, $f_{x_{i_1}, \dots, x_{i_k}}^{(k)}$, 甚至 $f_{x_{i_1}, \dots, x_{i_k}}$ 。

如果对任何 $i_1, \dots, i_k \in \{1, 2, \dots, n\}$, 偏导函数 $\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(\mathbf{x})$ 都连续, 即 f 的任何 k 阶偏导数都连续, 则称 f 是一个 C^k 函数, 记 $f \in C^k$ 。

如果对任何正整数 k , $f \in C^k$, 则记 $f \in C^\infty$ 。

定理 2.4.2. 对任何 C^k 函数 f , 对任何 $i_1, \dots, i_k \in \{1, 2, \dots, n\}$ 以及 $1, 2, \dots, k$ 的任何排列 π ,

$$\frac{\partial^k f}{\partial x_{i_{\pi(k)}} \cdots \partial x_{i_{\pi(1)}}}(\mathbf{x}) = \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(\mathbf{x}).$$

即高阶偏导数的值与求导顺序无关。

这个定理可以通过对 k 做数学归纳法得到证明, 而证明的核心是以下关于二元函数的结论。

引理 2.4.3. 若偏导函数 $\frac{\partial^2 f}{\partial x \partial y}$ 和 $\frac{\partial^2 f}{\partial y \partial x}$ 在 (a, b) 处连续, 则 $\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$ 。

由于 C^k 函数的高阶导数与求导顺序无关, 所以我们也采取对每个变量记重数的方式, 比如 $\frac{\partial^3 f}{\partial x^2 \partial y}$ 可以表示下面这些偏导数中的任何一个

$$\frac{\partial^3 f}{\partial x \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x \partial x}.$$

类似地, 你应该不难理解

$$\frac{\partial^p f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_m^{\alpha_m}} \quad (\alpha_1 + \alpha_2 + \cdots + \alpha_m = p)$$

所代表的偏导数, 这里 p 是一个正整数, 是函数 f 被求导的累计次数, 而非负整数 $\alpha_1, \alpha_2, \dots, \alpha_m$ 分别是对每个自变量求导的次数, 所以 $\partial x_k^{\alpha_k}$ 是

$$\underbrace{\partial x_k \cdots \partial x_k}_{\alpha_k \uparrow}$$

的缩记形式。

定理 2.4.4. (1) 若函数 f, g 都是 C^k 的, 则 $f + g, fg$ 也都是 C^k 的。

(2) 若复合映射 $g \circ f$ 中 f, g 都是 C^k 的, 则映射 $g \circ f$ 也是 C^k 的。

证明. 只证明第2个结论。

对 k 做数学归纳法。由链索法则

$$\frac{\partial(g \circ f)}{\partial x_i}(\mathbf{x}) = \sum_k \frac{\partial g}{\partial y_k}(f(\mathbf{x})) \frac{\partial f_k}{\partial x_i}(\mathbf{x})$$

知上式右端中每一项都是连续函数, 所以 $k=1$ 时结论成立。

假设结论对 k 成立。设 f, g 都是 C^{k+1} 的, 则上式右端中的每一项都是 C^k 的。所以 $g \circ f$ 是 C^{k+1} 的。 \square

例 2.4.5. 把直角坐标系下 C^2 函数 $u = u(x, y)$ 满足的方程

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

写成极坐标形式, 并求它的各向同性解 $u = u(r)$ 。 Δu 读作“ u 的 Laplacian”。

解: 由 $x = r \cos \theta, y = r \sin \theta$ 及链索法则得到

$$\begin{aligned} \frac{\partial}{\partial r} &= \frac{\partial x}{\partial r} \frac{\partial}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial}{\partial y} = \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y}, \\ \frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial}{\partial y} = -r \sin \theta \frac{\partial}{\partial x} + r \cos \theta \frac{\partial}{\partial y}, \end{aligned}$$

这是关于求导算子 $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ 的线性方程组, 由此解得

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

这样就把直角坐标系下的求导算子表示为极坐标下的求导算子, 所以

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

相加, 然后作用于 u 得到

$$0 = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{1}{r} \frac{\partial u}{\partial r}.$$

该方程的各向同性解 $u = u(r)$ 满足 $u'' + \frac{u'}{r} = 0$, 这是一个 Euler 方程。解得 $u(r) = A \ln r + B$ 是所有的各向同性解。请读者验证

$$u(x, y) = \frac{A}{2} \ln(x^2 + y^2) + B$$

是原方程的解。 \square

高阶微分

函数 f 的一阶微分是关于一个向量 \mathbf{v} 一个线性函数

$$\partial f(\mathbf{x})(\mathbf{v}) = \sum_i \frac{\partial f}{\partial x_i}(\mathbf{x})v_i = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}).$$

它对 \mathbf{x} 再次微分, 就得到 f 的二阶微分, 这是一个关于一对向量 \mathbf{v}, \mathbf{w} 的双线性函数

$$d^2 f(\mathbf{x})(\mathbf{v}, \mathbf{w}) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})v_i w_j = \frac{\partial}{\partial \mathbf{w}} \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}).$$

一般地, 函数 f 的 k 阶微分是关于 k 向量 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ 的 k 重线性函数

$$\begin{aligned} d^k f(\mathbf{x})(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) &= \sum_{i_1, i_2, \dots, i_k} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}(\mathbf{x})v_{1, i_1} v_{2, i_2} \cdots v_{k, i_k} \\ &= \frac{\partial^k f}{\partial \mathbf{v}_k \cdots \partial \mathbf{v}_1}(\mathbf{x}). \end{aligned}$$

当 $\mathbf{v}_1 = \mathbf{v}_2 = \cdots = \mathbf{v}_k = \mathbf{v}$ 时,

$$d^k f(\mathbf{x})(\mathbf{v}, \mathbf{v}, \dots, \mathbf{v}) = \sum_{i_1, i_2, \dots, i_k} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}(\mathbf{x})v_{i_1} v_{i_2} \cdots v_{i_k} = \frac{\partial^k f}{\partial \mathbf{v}^k}(\mathbf{x})$$

是关于向量 \mathbf{v} 的一个 k 次形。

Taylor 公式

对 C^k 函数 f , 考虑

$$g(t) = f(\mathbf{x}_0 + t\mathbf{v}).$$

则 g 是 C^k 的,

$$g'(t) = \partial f(\mathbf{x}_0 + t\mathbf{v})(\mathbf{v}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{v})\xi_i = \sum_{i=1}^n \xi_i \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{v}).$$

$$g''(t) = \sum_{1 \leq i, j \leq n} \xi_i \xi_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t\mathbf{v}) = \left(\sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i} \right)^2 f(\mathbf{x}_0 + t\mathbf{v}).$$

一般地,

$$g^{(k)}(t) = \sum_{1 \leq i_1, \dots, i_k \leq n} \xi_{i_1} \cdots \xi_{i_k} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(\mathbf{x}_0 + t\mathbf{v}) = \left(\sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i} \right)^k f(\mathbf{x}_0 + t\mathbf{v}).$$

所以

$$\begin{aligned}
 f(\mathbf{x}_0 + t\mathbf{v}) &= g(t) = \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} t^j + \frac{g^{(k)}(t\theta)}{k!} t^k \quad (0 < \theta < 1) \\
 &= \sum_{j=0}^{k-1} \frac{\left(\sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i}\right)^j f(\mathbf{x}_0)}{j!} t^j + \frac{\left(\sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i}\right)^k f(\mathbf{x}_0 + t\theta\mathbf{v})}{k!} t^k \\
 &= \sum_{j=0}^{k-1} \frac{\sum_{1 \leq i_1, \dots, i_j \leq n} \xi_{i_1} \cdots \xi_{i_j} \frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}(\mathbf{x}_0)}{j!} t^j \\
 &\quad + \frac{\sum_{1 \leq i_1, \dots, i_k \leq n} \xi_{i_1} \cdots \xi_{i_k} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(\mathbf{x}_0 + t\theta\mathbf{v})}{k!} t^k.
 \end{aligned}$$

对任意向量 \mathbf{v} , 设 $\mathbf{v} = \|\mathbf{v}\|\mathbf{v}_1$, 则

$$\begin{aligned}
 f(\mathbf{x}_0 + \mathbf{v}) &= \sum_{j=0}^{k-1} \frac{1}{j!} \sum_{1 \leq i_1, \dots, i_j \leq n} \frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}(\mathbf{x}_0) \xi_{i_1} \cdots \xi_{i_j} \\
 &\quad + \frac{1}{k!} \sum_{1 \leq i_1, \dots, i_k \leq n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(\mathbf{x}_0 + \theta\mathbf{v}) \xi_{i_1} \cdots \xi_{i_k} \\
 &= \sum_{j=0}^k \frac{1}{j!} \sum_{1 \leq i_1, \dots, i_j \leq n} \frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}(\mathbf{x}_0) \xi_{i_1} \cdots \xi_{i_j} + o(\|\mathbf{v}\|^k) \\
 &= \sum_{j=0}^k \sum_{\alpha_1 + \cdots + \alpha_n = j} \frac{\partial^j f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\mathbf{x}_0) \frac{\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}}{\alpha_1! \cdots \alpha_n!} + o(\|\mathbf{v}\|^k)
 \end{aligned}$$

定理 2.4.6. 设 f 在 \mathbf{x}_0 处 k 阶连续可微。则对 $\mathbf{v} = (\xi_1, \dots, \xi_n)$,

$$\begin{aligned}
 f(\mathbf{x}_0 + \mathbf{v}) &= \sum_{j=0}^k \frac{1}{j!} d^j f(\mathbf{x}_0)(\mathbf{v}, \dots, \mathbf{v}) + o(\|\mathbf{v}\|^k) \\
 &= \sum_{j=0}^k \sum_{\alpha_1 + \cdots + \alpha_n = j} \frac{\partial^j f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\mathbf{x}_0) \frac{\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}}{\alpha_1! \cdots \alpha_n!} + o(\|\mathbf{v}\|^k), \quad \|\mathbf{v}\| \rightarrow 0.
 \end{aligned}$$

另外, 只要 f 在连接 \mathbf{x}_0 和 $\mathbf{x}_0 + \mathbf{v}$ 的线段上一直有定义, 就存在 $0 < \theta < 1$ 使得

$$\begin{aligned}
 f(\mathbf{x}_0 + \mathbf{v}) &= \sum_{j=0}^{k-1} \sum_{\alpha_1 + \cdots + \alpha_n = j} \frac{\partial^j f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\mathbf{x}_0) \frac{\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}}{\alpha_1! \cdots \alpha_n!} \\
 &\quad + \sum_{\alpha_1 + \cdots + \alpha_n = k} \frac{\partial^k f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\mathbf{x}_0 + \theta\mathbf{v}) \frac{\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}}{\alpha_1! \cdots \alpha_n!}.
 \end{aligned}$$

注 2.4.7. Taylor 公式与等价范数的选取无关。

定理 2.4.8 (Taylor 展开的唯一性). 设 f 在 \mathbf{x}_0 处 k 阶可微。则 k 次多项式 P 满足

$$f(\mathbf{x}_0 + \mathbf{v}) = P(\mathbf{v}) + o(\|\mathbf{v}\|^k), \quad \mathbf{v} \rightarrow \mathbf{0},$$

当且仅当 P 是 f 在 \mathbf{x}_0 处的 k 次 Taylor 多项式。

例 2.4.9. 求 $\sin(x + y^2)$ 在 $(0, 0)$ 处的 3 次 Taylor 展开。

解: \sin 和多项式 $x + y^2$ 都是 C^∞ 函数, 所以 $\sin(x + y^2)$ 也是 C^∞ 函数。

由一元函数 \sin 的 Taylor 展开知

$$\begin{aligned}\sin(x + y^2) &= (x + y^2) - \frac{(x + y^2)^3}{3!} + o((x + y^2)^3), \quad x + y^2 \rightarrow 0 \\ &= x + y^2 - \frac{x^3}{6} - \frac{3x^2y^2 + 3xy^4 + y^6}{6} + o((x + y^2)^3), \quad x + y^2 \rightarrow 0.\end{aligned}$$

因为当 $r = \max\{|x|, |y|\} < 1$ 时,

$$\begin{aligned}|x + y^2| &\leq |x| + |y^2| \leq 2 \max\{|x|, |y|\} = 2r, \\ |x + y^2|^3 &\leq 8 \max\{|x|, |y|\}^3 = 8r^3, \\ |x^m y^n| &\leq \max\{|x|, |y|\}^{m+n} = r^{m+n},\end{aligned}$$

所以由 Taylor 展开的唯一性定理, $x + y^2 - \frac{x^3}{6}$ 是 $\sin(x + y^2)$ 在 $(0, 0)$ 处的三次 Taylor 多项式。 \square

习题 2.4

1. 证明引理 2.4.3。
2. 试着用变量树方法计算复合函数的高阶导数, 看看会出现哪些问题, 想想该如何解决。
3. 证明线性变换求逆运算 $A \mapsto A^{-1}$ 是 C^∞ 映射。
4. 设常数 A, B, C 满足 $B^2 > 4AC$, C^2 函数 $z = z(x, y)$ 满足

$$A \frac{\partial^2 z}{\partial x^2} + B \frac{\partial^2 z}{\partial x \partial y} + C \frac{\partial^2 z}{\partial y^2} = 0.$$

(a) 若 $A = C = 0$, 求上述微分方程的解 $z = z(x, y)$ 。

(b) 若 $A \neq 0$, 证明存在实数 α, β 使得在变量替换 $\begin{cases} u = \alpha x + y, \\ v = \beta x + y \end{cases}$ 下,

上述方程变成 $\frac{\partial^2 z}{\partial u \partial v} = 0$ 。

5. 平面极坐标系下 Laplace 方程为 $\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} = 0$ 。求分离变量形式的解 $u(r, \theta) = X(r)Y(\theta)$ 。
6. 求直角坐标系下的 Laplace 方程 $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$ 在空间球坐标系下的表达式。

7. (a) 如果多项式 $P(\mathbf{x})$ 满足: 对任意 $\lambda > 0$, $P(\lambda\mathbf{x}) = \lambda^k P(\mathbf{x})$, 并且 $P(\mathbf{x}) = o(\|\mathbf{x}\|^k)(\mathbf{x} \rightarrow \mathbf{0})$. 证明 $P(\mathbf{x}) = 0$.
- (b) 证明 Taylor 展开的唯一性定理。
8. 求 $\partial^2 \det(A)$ 以及 \det 在 I 处的 2 阶 Taylor 展开式。(提示: $AA^{*T} = \det(A)I$.)

2.5 函数凸凹性、函数极值

凸凹性和极值是非线性函数区别于线性函数的性质。就像我们在一元微积分课程中已经了解的那样, 研究可微函数的凸凹性和极值需要用到函数的二阶导数。

定义 2.5.1. f 在 \mathbf{x}_0 处的二阶微分

$$d^2 f(\mathbf{x}_0)(\mathbf{v}, \mathbf{v}) = \mathbf{v}^T H_f(\mathbf{x}_0) \mathbf{v} = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) v_i v_j$$

是一个对称双线性形, 其系数矩阵

$$H_f(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

称为函数 f 在 \mathbf{x}_0 处的 **Hesse 矩阵**。

利用 Hesse 矩阵, 我们可以更方便地表达函数的二阶 Taylor 展开式, 这是我们研究凸凹性和极值问题的基本工具。

定理 2.5.2 (二阶 Taylor 展开). 设 f 在 \mathbf{x}_0 处 2 阶连续可微. 则对 $\mathbf{v} = (\xi_1, \dots, \xi_n)$,

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{v}) + \frac{1}{2} \mathbf{v}^T H_f(\mathbf{x}_0) \mathbf{v} + o(\|\mathbf{v}\|^2), \quad \|\mathbf{v}\| \rightarrow 0.$$

另外, 只要 f 在连接 \mathbf{x}_0 和 $\mathbf{x}_0 + \mathbf{v}$ 的线段上一直有定义, 就存在 $0 < \theta < 1$ 使得

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{v}) + \frac{1}{2} \mathbf{v}^T H_f(\mathbf{x}_0 + \theta \mathbf{v}) \mathbf{v}.$$

□

定义 2.5.3. 称 $C \subseteq \mathbb{R}^n$ 是一个**凸集**, 如果对任意 $\mathbf{x}, \mathbf{y} \in C$ 以及任意 $0 \leq t \leq 1$, $(1-t)\mathbf{x} + t\mathbf{y} \in C$.

设 $C \subseteq \mathbb{R}^n$ 是一个凸集. 称 $f: C \rightarrow \mathbb{R}$ 是一个**凸函数**, 如果对任意 $\mathbf{x}, \mathbf{y} \in C$ 以及任意 $0 \leq t \leq 1$,

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}).$$

称 $f: C \rightarrow \mathbb{R}$ 是一个**凹函数**, 如果 $-f$ 是凸函数.

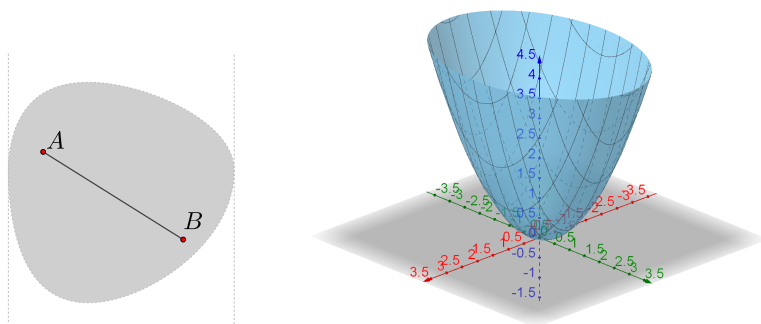


图 2.2: 凸集 (左) 和凸函数 (右)

定理 2.5.4. 如果 \mathbf{x}_0 是严格凸/严格凹函数 f 的极小值点, 则 \mathbf{x}_0 是 f 的最小值/最大值点, 且它是唯一的最小值/最大值点.

证明. 设 f 是严格凸函数.

假设存在 $\mathbf{x}_1 \neq \mathbf{x}_0$ 使得 $f(\mathbf{x}_1) \leq f(\mathbf{x}_0)$. 设 U 是 \mathbf{x}_0 的邻域, 使得对任意 $\mathbf{x} \in U$, $f(\mathbf{x}) \geq f(\mathbf{x}_0)$. 则当 $t > 0$ 足够小时, $(1-t)\mathbf{x}_0 + t\mathbf{x}_1 \in U$, 同时

$$f((1-t)\mathbf{x}_0 + t\mathbf{x}_1) < (1-t)f(\mathbf{x}_0) + tf(\mathbf{x}_1) \leq f(\mathbf{x}_0).$$

但这与 $f(\mathbf{x}_0)$ 是极小值矛盾. 这同时说明了最小值点是唯一的. \square

定理 2.5.5. C^2 函数 f 是凸/凹函数当且仅当 f 在各点处的 Hesse 矩阵都是半正定/半负定矩阵.

如果 f 在各点处的 Hesse 矩阵都是正定/负定矩阵, 则 f 是严格凸/严格凹函数.

证明. 考虑

$$g(t) = f((1-t)\mathbf{x} + t\mathbf{y}) - (1-t)f(\mathbf{x}) - tf(\mathbf{y}).$$

则 $g(0) = g(1) = 0$, 且

$$g''(t) = (\mathbf{y} - \mathbf{x})^T H_f((1-t)\mathbf{x} + t\mathbf{y})(\mathbf{y} - \mathbf{x}) \geq 0.$$

所以 g 是凸函数, 从而 $g(t) \leq (1-t)g(0) + tg(1) = 0$. \square

定义 2.5.6. 称 \mathbf{x}_0 是函数 f 的一个临界点, 如果对任意 $\mathbf{v} \in \mathbb{R}^n$, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = 0$.

称 \mathbf{x}_0 是 C^2 函数 f 的一个非退化临界点, 如果 \mathbf{x}_0 是函数 f 的一个临界点, 且 Hesse 矩阵 $H_f(\mathbf{x}_0)$ 是可逆矩阵.

定理 2.5.7. 若 f 在 \mathbf{x}_0 可微, 则 \mathbf{x}_0 是 f 的一个临界点当且仅当

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = \frac{\partial f}{\partial x_2}(\mathbf{x}_0) = \cdots = \frac{\partial f}{\partial x_m}(\mathbf{x}_0) = 0,$$

也即 \mathbf{x}_0 是 f 的梯度向量场 ∇f 的一个零点. \square

定理 2.5.8. 若 \mathbf{x}_0 是可微函数 f 定义域的一个内点, 且 \mathbf{x}_0 是 f 的一个极值点, 则 \mathbf{x}_0 是 f 的一个临界点.

证明. 对任意 $\mathbf{v} \in \mathbb{R}^n$, 考虑满足以下条件的可微曲线 $\mathbf{x}(t)$,

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

因为 \mathbf{x}_0 是 f 的极值点, 所以 $t = 0$ 是函数 $f(\mathbf{x}(t))$ 的极值点. 因此

$$0 = \left. \frac{\partial}{\partial t} f(\mathbf{x}(t)) \right|_{t=0} = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0),$$

即 \mathbf{x}_0 是 f 的一个临界点. \square

定理 2.5.9. 若函数 f 可微并且是凸函数, \mathbf{x}_0 是 f 的一个临界点, 则 \mathbf{x}_0 是 f 的最小值点.

证明. 因为 f 是凸函数, 所以它的 Hesse 矩阵是半正定的. 由 Taylor 展开知

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

\square

定理 2.5.10. 设 \mathbf{x}_0 是 C^2 函数 f 的一个临界点. 则

1. 若 $H_f(\mathbf{x}_0)$ 正定 (即所有特征值都是正数), 则 \mathbf{x}_0 是 f 的一个极小值点;
2. 若 $H_f(\mathbf{x}_0)$ 负定 (即所有特征值都是负数), 则 \mathbf{x}_0 是 f 的一个极大值点;

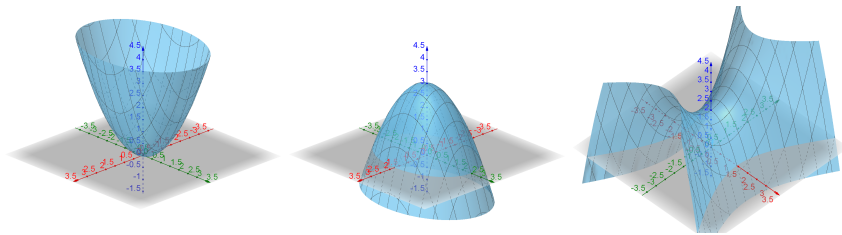


图 2.3: 极小、极大、鞍点

3. 若 $H_f(\mathbf{x}_0)$ 既有正特征值又有负特征值, 则 \mathbf{x}_0 不是 f 极值点。若 $H_f(\mathbf{x}_0)$ 可逆但既不正定也不负定, 则称 \mathbf{x}_0 是 C^2 函数 f 的一个鞍型临界点, 鞍型临界点不是极值点。

证明. (1) 因为 $H_f(\mathbf{x})$ 在 \mathbf{x}_0 处连续并且 $H_f(\mathbf{x}_0)$ 正定, 所以 f 在 \mathbf{x}_0 附近是严格凸的。

(2) f 在 \mathbf{x}_0 附近是严格凹的。

(3) 设 \mathbf{v} 是 $H_f(\mathbf{x}_0)$ 的特征向量, 对应特征值 $\lambda < 0$ 。则 $|t|$ 充分小时,

$$f(\mathbf{x}_0 + t\mathbf{v}) = f(\mathbf{x}_0) + \frac{t^2}{2} \mathbf{v}^T H_f(\mathbf{x}_0) \mathbf{v} + o(t^2) = f(\mathbf{x}_0) + \frac{\lambda t^2}{2} \mathbf{v}^T \mathbf{v} + o(t^2) < f(\mathbf{x}_0).$$

□

例 2.5.11. 求函数 $f(x, y) = xy \ln(x^2 + y^2)$ ($(x, y) \neq (0, 0)$) 的极值。

解: 令 $g(r, \theta) = f(r \cos \theta, r \sin \theta)$ 。则 $g(r, \theta) = r^2 \ln r \cdot \sin 2\theta$,

$$\begin{cases} \frac{\partial g}{\partial r} = r(2 \ln r + 1) \sin 2\theta, \\ \frac{\partial g}{\partial \theta} = 2r^2 \ln r \cdot \cos 2\theta, \end{cases} \quad \begin{cases} \frac{\partial^2 g}{\partial r^2} = (2 \ln r + 3) \sin 2\theta, \\ \frac{\partial^2 g}{\partial r \partial \theta} = 2r(2 \ln r + 1) \cos 2\theta, \\ \frac{\partial^2 g}{\partial \theta^2} = -4r^2 \ln r \cdot \sin 2\theta. \end{cases}$$

由 $\frac{\partial g}{\partial r} = \frac{\partial g}{\partial \theta} = 0$ 解得以下临界点

$$(1) \begin{cases} r = \frac{1}{\sqrt{e}}, \\ \theta = \pm \frac{\pi}{4}, \end{cases} \quad (2) \begin{cases} r = \frac{1}{\sqrt{e}}, \\ \theta = \pm \frac{3\pi}{4}, \end{cases} \quad (3) \begin{cases} r = 1, \\ \theta = 0, \end{cases} \quad (4) \begin{cases} r = 1, \\ \theta = \pi, \end{cases} \quad (5) \begin{cases} r = 1, \\ \theta = \pm \frac{\pi}{2}. \end{cases}$$

临界点	$(\frac{1}{\sqrt{e}}, \pm \frac{\pi}{4})$	$(\frac{1}{\sqrt{e}}, \pm \frac{3\pi}{4})$	$(1, 0)$ $(1, \pi)$	$(1, \pm \frac{\pi}{2})$
Hesse	$\begin{pmatrix} \pm 2 & 0 \\ 0 & \pm \frac{1}{e} \end{pmatrix}$	$\begin{pmatrix} \mp 2 & 0 \\ 0 & \mp \frac{1}{e} \end{pmatrix}$	$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}$
极值点	极小(极大)	极大(极小)	鞍点	鞍点
极值	$\mp \frac{1}{2e}$	$\pm \frac{1}{2e}$		

所以 $(\frac{1}{\sqrt{e}}, \frac{1}{\sqrt{e}})$ 和 $(-\frac{1}{\sqrt{e}}, -\frac{1}{\sqrt{e}})$ 是 f 的极小值点, 极小值为 $-\frac{1}{2e}$; $(\frac{1}{\sqrt{e}}, -\frac{1}{\sqrt{e}})$ 和 $(-\frac{1}{\sqrt{e}}, \frac{1}{\sqrt{e}})$ 是 f 的极大值点, 极大值为 $\frac{1}{2e}$. \square

另解: 函数 f 满足如下对称性: $f(-y, x) = -f(x, y)$. 所以当我们采用极坐标时, 只须考虑 $\theta \in [0, \frac{\pi}{2})$ 或 $\theta \in (-\frac{\pi}{4}, \frac{\pi}{4})$ 的情况. 令 $g(r, \theta) = f(r \cos \theta, r \sin \theta)$. 则

$$g(r, \theta) = r^2 \ln r \sin 2\theta.$$

根据 g 的梯度向量场

$$\nabla g = \left(\frac{\partial g}{\partial r}, \frac{\partial g}{\partial \theta} \right)^T = (r(2 \ln r + 1) \sin 2\theta, 2r^2 \ln r \cdot \cos 2\theta)^T$$

的方向 (见图2.5), 可以确定 g 的取值变化情况.

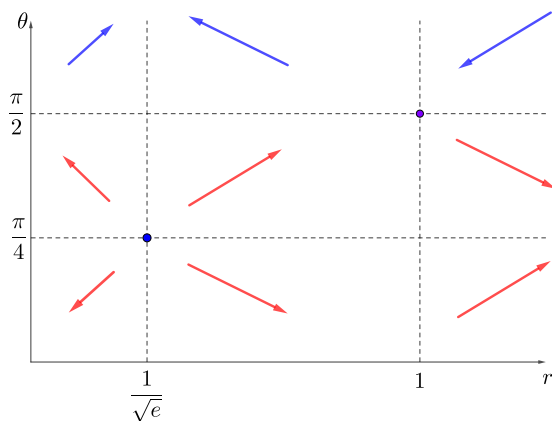


图 2.4: g 的梯度向量场 ∇g

具体而言, 由 $\frac{\partial g}{\partial r} = r(2 \ln r + 1) \sin 2\theta$ 知: 对 $0 < \theta < \frac{\pi}{2}$, 当 $0 < r < \frac{1}{\sqrt{e}}$ 时, $\frac{\partial g}{\partial r} < 0$; 当 $r > \frac{1}{\sqrt{e}}$ 时, $\frac{\partial g}{\partial r} > 0$. 所以 $g(r, \theta)$ 的临界点必然满足 $r = \frac{1}{\sqrt{e}}$, 且 $g(r, \theta)$ 关于 r 在 $r = \frac{1}{\sqrt{e}}$ 处取得最小值 $g(\frac{1}{\sqrt{e}}, \theta) = -\frac{1}{2e} \sin 2\theta$, 后者在 $\theta = \frac{\pi}{4}$ 处取得最小值 $-\frac{1}{2e}$. 因此 $(\frac{1}{\sqrt{e}}, \frac{\pi}{4})$ 为 g 的极小值点. 从而 $(\frac{1}{\sqrt{e}}, \frac{1}{\sqrt{e}})$ 为 f 的极小值点. 再由 f 的对称性知 $(-\frac{1}{\sqrt{e}}, \frac{1}{\sqrt{e}})$ 和 $(\frac{1}{\sqrt{e}}, -\frac{1}{\sqrt{e}})$ 为 f 的极大值点, $(-\frac{1}{\sqrt{e}}, -\frac{1}{\sqrt{e}})$ 为 f 的极小值点. f 的极小值为 $-\frac{1}{2e}$, 极大值为 $\frac{1}{2e}$.

因为 $\lim_{r \rightarrow +\infty} g(r, \frac{\pi}{4}) = +\infty$, 所以 g , 从而 f , 无上界, 进而没有最大值. 再由 f 的对称性知, f 也没有下界, 因而没有最小值.

对 $-\frac{\pi}{4} < \theta < \frac{\pi}{4}$, 当 $0 < r < 1$ 时, $\frac{\partial g}{\partial \theta} < 0$; 当 $r > 1$ 时, $\frac{\partial g}{\partial \theta} > 0$. 所以 $g(r, \theta)$ 的临界点必然满足 $r = 1$. 但 $\frac{\partial g}{\partial r}(1, \theta) = \sin 2\theta$ 在 $\theta = 0$ 两侧有相反的正负号, 所以 g 在此区域内的唯一的临界点 $(1, 0)$ 不是 g 的极值点. 因此 $(\pm 1, 0)$ 和 $(0, \pm 1)$ 是 f 的临界点但不是极值点.

你能根据 g 的梯度向量场 ∇g 的大致走向想象函数 g (进而函数 f) 的水平集和函数图像的大致样貌吗? \square

例 2.5.12. (最小二乘法) 给定矩阵 A 和向量 \mathbf{b} , 求向量 \mathbf{x} 使得 $\|A\mathbf{x} - \mathbf{b}\|_2$ 最小。

解: 令 $f(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|_2^2$. 则 f 连续且(只要 $A \neq 0$)当 $\|\mathbf{x}\| \rightarrow +\infty$ 时 $f(\mathbf{x}) \rightarrow +\infty$, 所以 f 必有最小值。

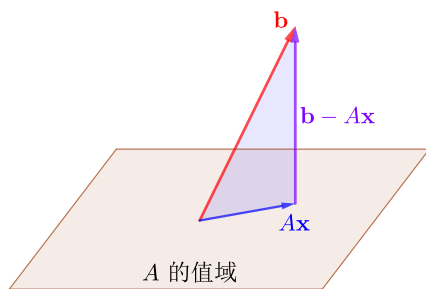


图 2.5: 最小二乘法

$$f(\mathbf{x}) = \langle A\mathbf{x} - \mathbf{b}, A\mathbf{x} - \mathbf{b} \rangle = \langle A^T A\mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, A^T \mathbf{b} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle.$$

所以 f 可微, 其最小值点必然是极值点, 从而是临界点。

$$df(\mathbf{x})(\mathbf{v}) = 2\langle A^T A\mathbf{x}, \mathbf{v} \rangle - 2\langle \mathbf{v}, A^T \mathbf{b} \rangle = 2\langle \mathbf{v}, A^T A\mathbf{x} - A^T \mathbf{b} \rangle.$$

\mathbf{x} 是 f 的临界点当且仅当对任意 \mathbf{v} , $df(\mathbf{x})(\mathbf{v}) = 0$, 即 $A^T A\mathbf{x} = A^T \mathbf{b}$ 。

最后这个等式称为最小二乘问题的正交方程, 其几何含义是 $A\mathbf{x} - \mathbf{b}$ 与 A 的每个列向量正交, 也就是 $A\mathbf{x}$ 是 \mathbf{b} 在 A 的列向量空间中的正交投影, 它是 A 的列向量空间中与 \mathbf{b} 距离最近的向量。

当 A 满列秩时, $A^T A$ 可逆, 此时解得唯一的临界点

$$\mathbf{x} = (A^T A)^{-1} A\mathbf{b},$$

它是 f 的最小值点。2-范数也就是欧氏范数是用坐标的平方计算的, 所以这个优化问题叫做最小二乘(二乘就是指平方)。 \square

注 2.5.13. 在一些实际应用中, 仅仅找到最小二乘问题的最优解往往还不够。我们还需要说明按特定模型使用最小二乘法的合理性, 此时人们常取

$$R^2 = 1 - \frac{\|A\mathbf{x}^* - \mathbf{b}\|_2^2}{\|\mathbf{b}\|_2^2},$$

R^2 的值越大, \mathbf{b} 就越倾向于成为 A 的列向量的线性组合。

最小二乘法是数据处理的一个重要方法。例如，我们想用父母的身高 x, y 来预测他们的孩子成年后的身高 z 。如果使用线性模型

$$z = ax + by + c$$

我们需要知道系数 a, b, c 。当然影响身高的因素有很多，所以这只是个近似模型。当我们获得了大量实际数据 (x_k, y_k, z_k) 以后，我们希望 a, b, c 满足

$$x_k a + y_k b + c - z_k = 0$$

这是关于 a, b, c 三个数的方程，但我们可以有成百上千甚至更多的数据。如果只是从大量数据中选取三组数据，虽然可以得到 a, b, c 的值，但很可能它们与其他数据是矛盾的。通常不可能有这样的 a, b, c 的值适合所有数据。Gauss 在处理大地测量的问题也遇到了同样的问题，他提出了最小二乘法，就是寻找使得平均平方误差

$$\frac{1}{n} \sum_{k=1}^n (ax_k + by_k + c - z_k)^2$$

达到最小时的参数值 a, b, c 。此时我们取

$$A = \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

此时，正交方程为

$$\begin{pmatrix} \overline{x\bar{x}} & \overline{x\bar{y}} & \overline{x\bar{z}} \\ \overline{x\bar{y}} & \overline{y\bar{y}} & \overline{y\bar{z}} \\ \overline{x\bar{z}} & \overline{y\bar{z}} & \overline{z\bar{z}} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \overline{x\bar{z}} \\ \overline{y\bar{z}} \\ \overline{z\bar{z}} \end{pmatrix}, \quad \overline{uv} = \frac{\sum_{k=1}^n u_k v_k}{n}, \quad \overline{u} = \frac{\sum_{k=1}^n u_k}{n}$$

除了线性模型，最小二乘法还可适用于多项式模型

$$y = a_0 + a_1 x + \dots + a_m x^m.$$

此时

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

习题2.5

1. 称向量场 \mathbf{F} 是一个保守场, 如果存在函数 P 使得 $\mathbf{F}(\mathbf{x}) = -\nabla P(\mathbf{x})$ 。物理中称这样的函数 P 称为向量场 \mathbf{F} 的一个势能函数。证明: 如果 \mathbf{F} 是一个保守场, P 是它的一个势能函数, 则对于微分方程

$$\mathbf{x}' = \mathbf{F}(\mathbf{x})$$

的每一条解曲线 $\mathbf{x}(t)$, 函数 $P(\mathbf{x}(t))$ 关于 t 都是单调不减函数。

2. 若 $f: C \rightarrow \mathbb{R}$ 是严格凸函数, \mathbf{x}_0 是 f 的临界点, 证明 \mathbf{x}_0 是 f 的唯一临界点, 并且是 f 的最小值点。

3. 求以下函数的极值和最大值

(1) $u = \sin x + \sin y + \sin z - \sin(x + y + z) (0 \leq x, y, z \leq \pi)$

(2) $f(\mathbf{x}) = \frac{\langle A\mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} (\mathbf{x} \neq \mathbf{0})$, 其中 A 是一个对称矩阵。

(3) $x^y - (x-1)y + (x-1)^2 (x > 0, y \in \mathbb{R})$ 。

4. 设 γ 是由抛物线 $y = c^2 - x^2 (-c \leq x \leq c)$ 和 x 轴上的线段组成的简单封闭曲线。求 γ 的内接四边形的最大面积。

5. 证明: $f(x, y) = -y + x^2 + xy + x^4 + y^4$ 有唯一极值点, 并判断该极值点的类型。

6. 设 $F: U \rightarrow \mathbb{R}^n$ 是 \mathbf{x}_0 的一个开邻域 $U \subseteq \mathbb{R}^m$ 上的一个 C^2 映射, f 在 $\mathbf{y}_0 = F(\mathbf{x}_0)$ 的一个邻域上的 C^2 函数, $U \subseteq \mathbb{R}^m$ 是开集。求 $f \circ F$ 在 \mathbf{x}_0 处的 Hesse 矩阵。

7. 如果用下述模型指数型模型拟合数据 (x_k, y_k) , 请给出一个计算模型系数的方法。

(1) $y = ax^b$,

(2) $y = Ae^{\lambda x}$,

(3) $y = \frac{a}{1+be^{-\lambda x}}$ 。

8. Hölder 不等式。设正数 p, q 满足 $\frac{1}{p} + \frac{1}{q} = 1$ 。证明对任意正数 x_1, x_2, y_1, y_2 都成立不等式

$$(x_1^p + x_2^p)^{\frac{1}{p}} (y_1^q + y_2^q)^{\frac{1}{q}} \geq x_1 y_1 + x_2 y_2.$$

(提示: 视 y_1, y_2 是给定的常数, 考虑函数 $f(x_1, x_2) = (x_1^p + x_2^p)^{\frac{1}{p}} (y_1^q + y_2^q)^{\frac{1}{q}} - x_1 y_1 - x_2 y_2$)

9. 如果二元函数 $f(x, y)$ 沿着过原点的每条直线都在原点取极小值, 问原点是否为 f 的极小值点?

2.6 应用: Newton 法

设 $f: E \rightarrow \mathbb{R}^m$ 是可微映射。设 \mathbf{x}^* 是方程 $f(\mathbf{x}) = 0$ 的解。假设 \mathbf{x}_n 是 \mathbf{x}^* 的一个近似解。在 \mathbf{x}_n 处把函数 $f(\mathbf{x})$ 局部线性化:

$$\mathbf{y} = f(\mathbf{x}_n) + \partial f(\mathbf{x}_n)(\mathbf{x} - \mathbf{x}_n).$$

由线性方程

$$\mathbf{0} = f(\mathbf{x}_n) + \partial f(\mathbf{x}_n)(\mathbf{x} - \mathbf{x}_n)$$

解得

$$\mathbf{x} = \mathbf{x}_n - (\partial f(\mathbf{x}_n))^{-1} f(\mathbf{x}_n).$$

这样得到 Newton 迭代

$$\mathbf{x}_{n+1} = \mathbf{x}_n - (\partial f(\mathbf{x}_n))^{-1} f(\mathbf{x}_n).$$

我们证明

定理 2.6.1. 如果 $f: E \rightarrow \mathbb{R}^m$ 是 C^1 可微映射, $f(\mathbf{x}^*) = 0$ 且 $\partial f(\mathbf{x}^*)$ 可逆, 则 Newton 迭代是局部收敛的。即存在 \mathbf{x}^* 的一个邻域 U 使得对任意 $\mathbf{x}_0 \in U$, 以 \mathbf{x}_0 为初值的 Newton 迭代点列 \mathbf{x}_n 都满足 $\lim_{n \rightarrow +\infty} \mathbf{x}_n = \mathbf{x}^*$ 。

如果进一步 f 是 C^2 可微的, 则 Newton 迭代还是二阶收敛的, 即存在常数 $M > 0$ 使得在 \mathbf{x}^* 的一个邻域 U 中, Newton 迭代点列 \mathbf{x}_n 都满足 $\|\mathbf{x}_{n+1} - \mathbf{x}^*\| \leq M\|\mathbf{x}_n - \mathbf{x}^*\|^2$ 。

证明. 记

$$T(\mathbf{x}) = \mathbf{x} - (\partial f(\mathbf{x}))^{-1} f(\mathbf{x}), \quad A = \|(\partial f(\mathbf{x}^*))^{-1}\| + 1.$$

存在 $\delta > 0$, 使得当 $\|\mathbf{x} - \mathbf{x}^*\| < \delta$ 时,

$$\|\partial f(\mathbf{x}) - \partial f(\mathbf{x}^*)\| < \frac{1}{4A}, \quad \|(\partial f(\mathbf{x}))^{-1}\| \leq A.$$

记 $\mathbf{v} = \mathbf{x} - \mathbf{x}^*$, 则

$$\begin{aligned} \|T(\mathbf{x}) - \mathbf{x}^*\| &= \|\mathbf{x} - \mathbf{x}^* - (\partial f(\mathbf{x}))^{-1} f(\mathbf{x}) + (\partial f(\mathbf{x}))^{-1} f(\mathbf{x}^*)\| \\ &\leq \|(\partial f(\mathbf{x}))^{-1}\| \|\partial f(\mathbf{x})(\mathbf{x} - \mathbf{x}^*) - f(\mathbf{x}) + f(\mathbf{x}^*)\| \\ &\leq A \left\| \int_0^1 \frac{\partial}{\partial t} (\partial f(\mathbf{x})(t\mathbf{v}) - f(\mathbf{x}^* + t\mathbf{v})) dt \right\| \\ &= A \left\| \int_0^1 \partial f(\mathbf{x})(\mathbf{v}) - \partial f(\mathbf{x}^* + t\mathbf{v})(\mathbf{v}) dt \right\| \\ &\leq A \int_0^1 \|\partial f(\mathbf{x}) - \partial f(\mathbf{x}^* + t\mathbf{v})\| \|\mathbf{v}\| dt \\ &\leq A \int_0^1 \frac{1}{2A} \|\mathbf{v}\| dt \leq \frac{1}{2} \|\mathbf{x} - \mathbf{x}^*\| < \delta, \end{aligned}$$

所以

$$\|T^k(\mathbf{x}) - \mathbf{x}^*\| \leq \frac{1}{2^k} \|\mathbf{x} - \mathbf{x}^*\| \rightarrow 0, \quad k \rightarrow +\infty.$$

因此 Newton 法是局部收敛的。

如果 $f: E \rightarrow \mathbb{R}^n$ 是 C^2 可微映射, 则不妨设当 $\|\mathbf{x} - \mathbf{x}^*\| < \delta$ 时, $\|D^2 f(\mathbf{x})\| < M_2$, 则

$$\begin{aligned} \|T(\mathbf{x}) - \mathbf{x}^*\| &\leq A \int_0^1 \|\partial f(\mathbf{x}) - \partial f(\mathbf{x}^* + t\mathbf{v})\| \|\mathbf{v}\| dt \\ &= A \int_0^1 \left\| \int_t^1 \frac{d}{ds} \partial f(\mathbf{x}^* + s\mathbf{v}) \mathbf{v} ds \right\| \|\mathbf{v}\| dt \\ &\leq AM_2 \|\mathbf{v}\|^2 \int_0^1 \int_t^1 ds dt = \frac{AM_2}{2} \|\mathbf{x} - \mathbf{x}^*\|^2, \end{aligned}$$

因此 Newton 法是二阶局部收敛的。 \square

注 2.6.2. 虽然 Newton 法具有很好的收敛速度, 但是在每次迭代时都需要重新计算 $\partial f(x_n)^{-1}$, 这反而影响了算法的效率。广义 Newton 法采用一系列更容易计算逆矩阵的可逆矩阵 A_n (甚至是固定的矩阵 A_0) 代替 $\partial f(x_n)$,

$$\mathbf{x}_{n+1} = \mathbf{x}_n - A_n^{-1} f(\mathbf{x}_n),$$

虽然每步迭代变成一阶收敛, 但由于 A_n^{-1} 节省了计算成本, 效率反而提高了。

关于 Newton 法、广义 Newton 法以及 Newton-Kantorovich 定理, 感兴趣的读者可以参考 P.G.Ciarlet 著 Linear and Nonlinear Functional Analysis with Applications (中译本, 《线性与非线性泛函分析及其应用》下册, 秦铁虎译, 高等教育出版社 2017 年第一版)

习题 2.6

1. 为了得到可逆矩阵 A 的逆矩阵, 对 $F(X) = X^{-1} - A = 0$ 使用 Newton 迭代。请写出迭代的递推关系式。

2.7 习题讨论课2

1. 设 $\Omega \subset \mathbb{R}^2$ 是有界区域, 可微函数 $f(x, y)$ 满足

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) + \frac{\partial f}{\partial y}(x, y) = kf(x, y), & (x, y) \in \Omega, \\ f(x, y) = 0, & (x, y) \in \partial\Omega. \end{cases}$$

证明 $f(x, y) = 0 (\forall (x, y) \in \Omega)$.

2. 解方程
$$\begin{cases} \frac{\partial z}{\partial t} = a \frac{\partial z}{\partial x} + b \frac{\partial z}{\partial y}, \\ z(x, y, 0) = f(x, y). \end{cases}$$

3. 设 f 可微, 证明对任意 $t > 0$, 存在 (ξ, η) 满足 $\xi^2 + \eta^2 = t$, 且

$$\eta \frac{\partial f}{\partial x}(\xi, \eta) = \xi \frac{\partial f}{\partial y}(\xi, \eta).$$

4. 设 $\begin{cases} x = e^t \cos \theta, \\ y = e^t \sin \theta. \end{cases}$ 求 Laplace 方程 $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ 在 (t, θ) 坐标系下的形式。

5. 在平面直角坐标系 (x, y) 下, 对平面向量场 $(X(x, y), Y(x, y))$, 定义其旋度和散度依次为

$$\text{rot}(X, Y) = \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}, \quad \text{div}(X, Y) = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}.$$

对平面数量场(即函数) $f(x, y)$, 其梯度向量场为

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)^T.$$

求

(a) $\text{rot} \nabla f$ 和 $\text{div} \nabla f$.

(b) 考虑平面线性向量场

$$\begin{pmatrix} X \\ Y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

求这个向量场的旋度和散度。讨论旋度和散度分别为零时 A 的特征。

6. 设 M_n 为所有 n 阶实数矩阵组成的线性空间, $\det : M_n \rightarrow \mathbb{R} (A \mapsto \det A, \det A$ 为矩阵 A 的行列式)。

- (a) 证明 \det 是可微函数, 并求 \det 在 $A \in M_n$ 处的微分。
- (b) 证明 $\langle A, B \rangle = \text{tr}(A^T B)$ 是 M_n 上一个内积, 并求 \det 在 $A \in M_n$ 处的梯度。

2.8 习题讨论课3

1. 求Laplace方程

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = 0$$

的各向同性解 $u = u(r)$ (其中 $r = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$).

2. 设有界闭区域 Ω 的连续函数 u 是边值问题

$$\begin{cases} \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = u, & \mathbf{x} \in \text{int}\Omega, \\ u(x_1, x_2, \dots, x_n) = 0, & \mathbf{x} \in \partial\Omega \end{cases}$$

的解。证明 $u \equiv 0$ 。

3. 求以下函数在指定点的 Peano 余项 Taylor 展开

(a) x^y 在 $(x, y) = (1, 0)$ 处展开到二阶;

(b) $\lambda x + e^{-(x^2+y^2)}$ 在它的临界点处展开到二阶。

4. 对参数 λ , 讨论函数 $\lambda x + e^{-(x^2+y^2)}$ 的极值情况。

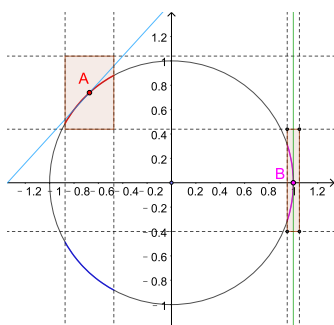
5. 求 $u = \sin x_1 + \sin x_2 + \sin x_3 - \sin(x_1 + x_2 + x_3)$ 的所有临界点, 判别其类型, 并求 u 的最大值和最小值。

6. 证明: 沿过原点的任何直线 $u = (y - x^2)(y - 3x^2)$ 都在原点处取严格极小值, 但原点不是这个二元函数的极值点。

第3章 隐函数与逆映射定理

3.1 隐函数定理

很多数学问题最终都归结为解方程。一个方程是否有解，是否有唯一解，以及如果方程中含有参数，那么方程的解对参数的变化有怎样的响应，这些就是隐函数定理讨论的问题。从几何的角度说，即函数 $F(x, y)$ 的水平集 $F(x, y) = 0$ 是否可以（局部）看成是一个可微函数 $y = f(x)$ 的图像，如何确保 f 具有适当的光滑程度，以及如何得到保 f 的近似表达式。可以说，隐函数定理在多元微分学乃至非线性分析中都发挥着一个基本而且重要的作用。



让我们从最简单的方程——二元一次方程开始。

$$ax + by + c = 0,$$

当 $b \neq 0$ 时，上述方程有唯一解

$$y = b^{-1}(-c - ax).$$

当 $b = 0$ 且 $a = c = 0$ 时，上述方程对每个 x 都有无穷多解；当 $b = 0$ 且 $a = 0$ 且 $c \neq 0$ 时，上述方程对每个 x 都无解；当 $b = 0$ 且 $a \neq 0$ 时，上述方程对 $x = -\frac{c}{a}$ 有无穷多解，对 $x \neq -\frac{c}{a}$ 无解。总之，上述方程关于 y 有唯一解当且仅当 $b \neq 0$ 。

我们在考察非线性方程, 例如

$$x^2 + y^2 = 1.$$

对任何 (x_0, y_0) , 如果 $x_0^2 + y_0^2 = 1$ 且 $|x_0| < 1$, 则取

$$\delta = \min\{1 - x_0, x_0 + 1\},$$

于是 $(x_0 - \delta, x_0 + \delta) \subset (-1, 1)$,

高维时, 考虑线性方程

$$A\mathbf{x} + B\mathbf{y} + \mathbf{c} = 0.$$

对每个 \mathbf{x} 上述方程有唯一解当且仅当 B 是可逆矩阵, 此时

$$\mathbf{y} = B^{-1}(-\mathbf{c} - A\mathbf{x}).$$

如果 B 不是列满秩矩阵, 则可能对某些 \mathbf{x} , $A\mathbf{x} + \mathbf{c}$ 不在 B 的列空间中, 此时方程无解。如果 B 列满秩矩阵, 但列数大于它的秩, 则方程有无穷多解。

以下隐函数定理是多元微分学中最重要结论, 它给出了一般的非线性方程解的存在唯一性的一个充分条件。

定理 3.1.1 (隐函数定理). 设 $F: W \rightarrow \mathbb{R}^n$ 在开集 $W \subseteq \mathbb{R}^m \times \mathbb{R}^n$ 内是 C^k 可微映射, $(\mathbf{x}_0, \mathbf{y}_0) \in W$ 满足 $F(\mathbf{x}_0, \mathbf{y}_0) = 0$ 且 $\partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)$ 可逆。则存在 \mathbf{x}_0 的邻域 $U \subset \mathbb{R}^m$ 和 \mathbf{y}_0 的邻域 V 以及 C^k 映射 $f: U \rightarrow \mathbb{R}^n$ 使得:

- (1) 对任意 $\mathbf{x} \in U$, $F(\mathbf{x}, f(\mathbf{x})) = 0$, 并且
- (2) 若 $(\mathbf{x}, \mathbf{y}) \in U \times V$ 满足 $F(\mathbf{x}, \mathbf{y}) = 0$, 则 $\mathbf{y} = f(\mathbf{x})$ 。

注 3.1.2. (1) 隐函数定理的目的是在一个特解 $(\mathbf{x}_0, \mathbf{y}_0)$ 附近找到非线性方程的所有解。把非线性方程在特解 $(\mathbf{x}_0, \mathbf{y}_0)$ 处线性化,

$$\begin{aligned} F(\mathbf{x}, \mathbf{y}) &= F(\mathbf{x}_0, \mathbf{y}_0) + \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + \partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + h.o.t. \\ &= \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + \partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + h.o.t. \end{aligned}$$

得到线性方程

$$\partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + \partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) = 0.$$

隐函数定理中的非退化条件 “ $\partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)$ 可逆” 是为了确保这个线性方程有唯一解。这个线性方程的解

$$\mathbf{y} = \mathbf{y}_0 - (\partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0)$$

正是非线性方程 $F(\mathbf{x}, \mathbf{y}) = 0$ 的解的线性近似:

$$\mathbf{y} = \mathbf{y}_0 - (\partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + h.o.t.$$

(2) 如果隐函数 f 可微, 则对恒等式

$$F(\mathbf{x}, f(\mathbf{x})) = 0$$

两边求导, 得到

$$\partial_{\mathbf{x}}F(\mathbf{x}, f(\mathbf{x})) + \partial_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x}))\partial_{\mathbf{x}}f(\mathbf{x}) = 0.$$

由于 F 是 C^1 , f 连续, $\partial_{\mathbf{y}}F(\mathbf{x}_0, f(\mathbf{x}_0))$ 可逆, 所以对 \mathbf{x}_0 附近的所有 \mathbf{x} , 线性映射 $\partial_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x}))$ 可逆, 而线性映射求逆的运算是 C^∞ 的, 因此

$$\partial_{\mathbf{x}}f(\mathbf{x}) = -(\partial_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x})))^{-1}\partial_{\mathbf{x}}F(\mathbf{x}, f(\mathbf{x}))$$

且它是连续的, 从而 f 是 C^1 的。

用数学归纳法不难证明, 如果 F 是 C^k 的, 则 f 也是 C^k 的。

(3) 如果隐函数 f 连续, 我们证明 f 可微。

记

$$\begin{aligned} A &= \partial_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x})), \\ B &= \partial_{\mathbf{x}}F(\mathbf{x}, f(\mathbf{x})), \\ M &= \|A^{-1}B\| + \|A^{-1}\|. \end{aligned}$$

由于 F 可微, 所以

$$\begin{aligned} 0 &= F(\mathbf{x} + \mathbf{v}, f(\mathbf{x} + \mathbf{v})) \\ &= F(\mathbf{x}, f(\mathbf{x})) + B\mathbf{v} + A(f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})) + \alpha(\mathbf{x}, \mathbf{v}) \\ &= B\mathbf{v} + A(f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})) + \alpha(\mathbf{x}, \mathbf{v}), \end{aligned}$$

其中 $\alpha(\mathbf{x}, \mathbf{v})$ 满足: 对任意 $\varepsilon > 0$, 存在 $\delta_\varepsilon > 0$ 使得当 $\|\mathbf{v}\| + \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\| < \delta_\varepsilon$ 时,

$$\|\alpha(\mathbf{x}, \mathbf{v})\| \leq \varepsilon (\|\mathbf{v}\| + \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\|).$$

由于 f 连续, 所以存在 $\tilde{\delta}_\varepsilon > 0$, 使得当 $\|\mathbf{v}\| < \tilde{\delta}_\varepsilon$ 时, $\|\mathbf{v}\| + \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\| < \delta_\varepsilon$ 。

$$\begin{aligned} \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\| &= \|A^{-1}(B\mathbf{v} + \alpha(\mathbf{x}, \mathbf{v}))\| \\ &\leq M\|\mathbf{v}\| + M\varepsilon(\|\mathbf{v}\| + \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\|) \end{aligned}$$

所以

$$\|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\| \leq \frac{M + M\varepsilon}{1 - M\varepsilon} \|\mathbf{v}\|,$$

因此

$$\|\alpha(\mathbf{x}, \mathbf{v})\| \leq \varepsilon (\|\mathbf{v}\| + \|f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})\|) \leq \varepsilon \left(1 + \frac{M + M\varepsilon}{1 - M\varepsilon}\right) \|\mathbf{v}\|,$$

即当 $\mathbf{v} \rightarrow 0$ 时, $\alpha(\mathbf{x}, \mathbf{v}) = o(\|\mathbf{v}\|)$ 。

$$f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) + A^{-1}B\mathbf{v} = -A^{-1}\alpha(\mathbf{x}, \mathbf{v}) = o(\|\mathbf{v}\|), \quad \mathbf{v} \rightarrow 0.$$

所以 f 可微。

(4) 如果隐函数 f 在 \mathbf{x}_0 处连续, 则对 \mathbf{x}_0 附近的所有 \mathbf{x} , $F(\mathbf{x}, f(\mathbf{x})) = 0$ 且 $\partial_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x}))$ 可逆。于是隐函数定理的证明关键之处在于证明 \mathbf{x}_0 附近隐函数 f 的局部的存在性和唯一性, 以及 f 在 \mathbf{x}_0 连续性。

证明. 记 $A = \partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0)$, $B = \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)$ 。令

$$G(\mathbf{u}, \mathbf{v}) = A^{-1}F(\mathbf{x}_0 + \mathbf{u}, \mathbf{y}_0 - A^{-1}B\mathbf{u} + \mathbf{v}).$$

则 $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$ 和 $\mathbf{y} = \mathbf{y}_0 - A^{-1}B\mathbf{u} + \mathbf{v}$ 满足 $F(\mathbf{x}, \mathbf{y}) = 0$ 当且仅当 $G(\mathbf{u}, \mathbf{v}) = 0$ 。

G 是 C^1 映射, 满足 $G(\mathbf{0}, \mathbf{0}) = \mathbf{0}$, $\partial_{\mathbf{u}}G(\mathbf{0}, \mathbf{0}) = 0$, $\partial_{\mathbf{v}}G(\mathbf{0}, \mathbf{0}) = I$ 。

定义

$$T(\mathbf{u}, \mathbf{v}) = \mathbf{v} - G(\mathbf{u}, \mathbf{v}).$$

则 T 是 C^1 映射, $T(\mathbf{0}, \mathbf{0}) = \mathbf{0}$, $\partial_{\mathbf{u}}T(\mathbf{0}, \mathbf{0}) = 0$, $\partial_{\mathbf{v}}T(\mathbf{0}, \mathbf{0}) = 0$ 。于是 \mathbf{v} 是 $G(\mathbf{u}, \mathbf{v}) = 0$ 的解当且仅当 \mathbf{v} 是 $T(\mathbf{u}, \cdot)$ 的不动点。

任给 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 (\mathbf{u}, \mathbf{v}) , 只要 $\|\mathbf{u}\| \leq \delta$ 及任意 $\|\mathbf{v}\| \leq \delta$, 就有 $(\mathbf{x}_0 + \mathbf{u}, \mathbf{y}_0 - A^{-1}B\mathbf{u} + \mathbf{v}) \in W$, 且

$$\|\partial_{\mathbf{v}}T(\mathbf{u}, \mathbf{v})\| < \frac{1}{2}, \quad \|\partial_{\mathbf{u}}T(\mathbf{u}, \mathbf{v})\| = \|\partial_{\mathbf{u}}G(\mathbf{u}, \mathbf{v})\| < \varepsilon,$$

于是对任意 $\|\mathbf{u}\| \leq \delta$ 及任意 $\|\mathbf{v}_1\| \leq \delta$, $\|\mathbf{v}_2\| \leq \delta$,

$$\|T(\mathbf{u}, \mathbf{v}_2) - T(\mathbf{u}, \mathbf{v}_1)\| \leq \frac{1}{2}\|\mathbf{v}_2 - \mathbf{v}_1\|.$$

由此知, 如果 $T(\mathbf{u}, \mathbf{v}_1) = \mathbf{v}_1$ 且 $T(\mathbf{u}, \mathbf{v}_2) = \mathbf{v}_2$, 则必有 $\mathbf{v}_1 = \mathbf{v}_2$ 。即对任意 $\mathbf{u} (\|\mathbf{u}\| \leq \delta)$, $T(\mathbf{u}, \mathbf{v})$ 关于 $\mathbf{v} (\|\mathbf{v}\| \leq \delta)$ 至多只有一个不动点, 即方程 $G(\mathbf{u}, \mathbf{v}) = 0$ 关于 $\mathbf{v} (\|\mathbf{v}\| \leq \delta)$ 至多只有一个解。

另外

$$\|T(\mathbf{u}, \mathbf{v})\| \leq \|T(\mathbf{u}, \mathbf{v}) - T(\mathbf{u}, \mathbf{0})\| + \|T(\mathbf{u}, \mathbf{0}) - T(\mathbf{0}, \mathbf{0})\| \leq \varepsilon\|\mathbf{u}\| + \frac{1}{2}\|\mathbf{v}\|.$$

取

$$E = \{(\mathbf{u}, \mathbf{v}) \mid \|\mathbf{u}\| \leq \frac{\delta}{1+2\varepsilon}, \|\mathbf{v}\| \leq 2\varepsilon\|\mathbf{u}\|\}.$$

则 E 是闭集, 对任意 $(\mathbf{u}, \mathbf{v}) \in E$, $\|\mathbf{v}\| \leq \delta$,

$$\|T(\mathbf{u}, \mathbf{v})\| \leq \varepsilon\|\mathbf{u}\| + \frac{1}{2}\|\mathbf{v}\| \leq \varepsilon\|\mathbf{u}\| + \varepsilon\|\mathbf{u}\| \leq 2\varepsilon\|\mathbf{u}\|.$$

故 $(\mathbf{u}, T(\mathbf{u}, \mathbf{v})) \in E$ 。所以存在唯一的 $\mathbf{v}^* = f(\mathbf{u})$ 使得

1. $(\mathbf{u}, f(\mathbf{u})) \in E$,
2. 对满足 $\|\mathbf{u}\| \leq \frac{\delta}{1+2\varepsilon}$ 和 $\|\mathbf{v}\| \leq \delta$ 的任意 (\mathbf{u}, \mathbf{v}) : $T(\mathbf{u}, \mathbf{v}) = \mathbf{v}$ 当且仅当 $\mathbf{v} = f(\mathbf{u})$ 。

因为 $f(\mathbf{0}) = \mathbf{0}$ 且

$$\|f(\mathbf{u})\| \leq 2\varepsilon\|\mathbf{u}\|,$$

所以 f 在 $\mathbf{u} = \mathbf{0}$ 处可微, 且 $\partial f(\mathbf{0}) = \mathbf{0}$ 。从而

$$\mathbf{y}(\mathbf{x}) = \mathbf{y}_0 - (\partial_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} \partial_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|),$$

故 $\mathbf{y}(\mathbf{x})$ 在 \mathbf{x}_0 处可微。 □

例 3.1.3. 易见 $(x, y, u, v) = (1, 1, 3, 2)$ 是方程组

$$\begin{cases} x^2 + xy + y^2 - u = 0, \\ x^2 - xy + 2y^2 - v = 0 \end{cases}$$

的解。证明在对 $(3, 2)$ 附近的任何 (u, v) , 上述方程组关于 (x, y) 在 $(1, 1)$ 附近有唯一解 $(x, y) = f(u, v)$, 并求 f 在 $(u, v) = (1, 1)$ 处的二阶 Taylor 多项式。

解: 记 $\xi = x - 1, \eta = y - 1, \varphi = u - 3, \theta = v - 2$, 则原方程改写为

$$\begin{cases} (1 + \xi)^2 + (1 + \xi)(1 + \eta) + (1 + \eta)^2 - (3 + \varphi) = 0, \\ (1 + \xi)^2 - (1 + \xi)(1 + \eta) + 2(1 + \eta)^2 - (2 + \theta) = 0, \end{cases}$$

即

$$\begin{cases} 3\xi + 3\eta - \varphi + \xi^2 + \xi\eta + \eta^2 = 0, \\ \xi + 3\eta - \theta + \xi^2 - \xi\eta + 2\eta^2 = 0. \end{cases} \quad (*)$$

解得

$$\begin{cases} \xi = \frac{\varphi - \theta}{2} - \xi\eta + \frac{\eta^2}{2}, \\ \eta = \frac{-\varphi + 3\theta}{6} - \frac{\xi^2}{2} + \frac{2\xi\eta}{3} - \frac{5\eta^2}{6}. \end{cases}$$

方程组 (*) 在 $(\xi, \eta, \varphi, \theta) = (0, 0, 0, 0)$ 处的线性化方程组有唯一解

$$\begin{cases} \xi = \frac{\varphi - \theta}{2}, \\ \eta = \frac{-\varphi + 3\theta}{6}. \end{cases}$$

所以根据隐函数定理, 方程组 (*) 在 $(\xi, \eta, \varphi, \theta) = (0, 0, 0, 0)$ 附近有形如

$$\begin{cases} \xi = \frac{\varphi - \theta}{2} + z(\varphi, \theta), \\ \eta = \frac{-\varphi + 3\theta}{6} + w(\varphi, \theta) \end{cases}$$

的解。代回方程组 (*) 得到

$$\begin{cases} z = -\xi\eta + \frac{\eta^2}{2} \\ \quad = -\left(\frac{\varphi-\theta}{2}\right)\left(\frac{-\varphi+3\theta}{6}\right) + \frac{1}{2}\left(\frac{-\varphi+3\theta}{6}\right)^2 + o(r^2), \\ w = -\frac{\xi^2}{2} + \frac{2\xi\eta}{3} - \frac{5\eta^2}{6} \\ \quad = -\frac{1}{2}\left(\frac{\varphi-\theta}{2}\right)^2 + \frac{2}{3}\left(\frac{\varphi-\theta}{2}\right)\left(\frac{-\varphi+3\theta}{6}\right) - \frac{5}{6}\left(\frac{-\varphi+3\theta}{6}\right)^2 + o(r^2) \end{cases}$$

其中 $r = \sqrt{\varphi^2 + \theta^2}$ 。

因此

$$\begin{cases} x = 1 + \frac{\varphi-\theta}{2} - \left(\frac{\varphi-\theta}{2}\right)\left(\frac{-\varphi+3\theta}{6}\right) + \frac{1}{2}\left(\frac{-\varphi+3\theta}{6}\right)^2 + o(r^2), \\ y = 1 + \frac{-\varphi+3\theta}{6} - \frac{1}{2}\left(\frac{\varphi-\theta}{2}\right)^2 + \frac{2}{3}\left(\frac{\varphi-\theta}{2}\right)\left(\frac{-\varphi+3\theta}{6}\right) - \frac{5}{6}\left(\frac{-\varphi+3\theta}{6}\right)^2 + o(r^2), \end{cases}$$

其中 $r = \sqrt{\varphi^2 + \theta^2}$, $\varphi = u - 3, \theta = v - 2$ 。 □

习题3.1

1. 设 $F(x, y, z)$ 是一个可微的三元函数, 满足

$$\frac{\partial F}{\partial x} \cdot \frac{\partial F}{\partial y} \cdot \frac{\partial F}{\partial z} \neq 0.$$

证明: 对任意 (x_0, y_0, z_0) , 在 (x_0, y_0, z_0) 的一个邻域里, 方程

$$F(x, y, z) = F(x_0, y_0, z_0)$$

唯一地确定了三个可微函数 $x = x(y, z)$, $y = y(z, x)$, $z = z(x, y)$, 并且

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1.$$

2. 设 C^1 函数 $F: U \times (y_0 - \delta_0, y_0 + \delta_0) \rightarrow \mathbb{R}$ 满足:

- (a) $F(\mathbf{x}_0, y_0) = 0$,
 (b) 对任意 $\mathbf{x} \in U$, $F(\mathbf{x}, y)$ 关于 y 严格增。

证明:

- (a) 存在 $\delta_1, \delta_2 > 0$ 使得, 对任意 $\mathbf{x}: \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$, 存在唯一的 $y(\mathbf{x})$:
 $|y(\mathbf{x}) - y_0| < \delta_2$ 使得 $F(\mathbf{x}, y(\mathbf{x})) = 0$;
 (b) $y(\mathbf{x})$ 关于 \mathbf{x} 在 x_0 处连续。
 (c) 若 $\frac{\partial F}{\partial y}(\mathbf{x}_0, y_0) > 0$, 则 $y(\mathbf{x})$ 在 \mathbf{x}_0 是 C^1 的。

3. 利用上个习题的结论证明隐函数定理。
4. 证明 $-y + x^2 + xy + x^4 + y^4 = 0$ 在 $(0, 0)$ 的一个邻域内定义了一个 C^∞ 隐函数 $y = h(x)$, 并求 h 在 $x = 0$ 处的四阶泰勒展开式。
5. 设 $f : U \times V \rightarrow \mathbb{R}$ 是 C^2 函数, \mathbf{x}_0 是 $f(\cdot, \lambda_0)$ 的非退化临界点, 即 $f(\mathbf{x}, \lambda_0)$ 关于 \mathbf{x} 在 \mathbf{x}_0 处的 Hesse 矩阵是可逆的。证明对足够接近 λ_0 的所有 λ , $f(\mathbf{x}, \lambda)$ 关于 \mathbf{x} 在 \mathbf{x}_0 附近有唯一的临界点, 并且该临界点的类型与 \mathbf{x}_0 作为 $f(\cdot, \lambda_0)$ 的临界点的类型是一样的。
6. 设 $\varepsilon \in \mathbb{R}$, 讨论 $f_\varepsilon(x, y) = \varepsilon x + x^2 e^{-y^2}$ 的极值点情况。
7. 设 λ_0 是矩阵 A_0 的一个单重特征值(即 $p(\lambda_0) = 0, p'(\lambda_0) \neq 0$, 其中 $p(\lambda) = \det(\lambda I - A_0)$), \mathbf{x}_0 是 A_0 对应于 λ_0 的一个单位特征向量。证明存在 $\delta > 0$ 使得对任意矩阵 A , 只要 $\|A - A_0\| < \delta$, A 就有唯一的特征值 $\lambda(A)$ 和相应的单位特征向量 $\mathbf{x}(A)$ 使得 $\lambda(A_0) = \lambda_0, \mathbf{x}(A_0) = \mathbf{x}_0$, 并且 $\lambda(A)$ 和 $\mathbf{x}(A)$ 关于 A 是 C^∞ 的。

3.2 逆映射定理

以下反函数定理(也称为逆映射定理)说, 如果一个可微映射在一点处的线性近似是可逆的, 那么这个可微映射在该点附近就是局部可逆的。

定理 3.2.1. 设 $f : U \rightarrow \mathbb{R}^n$ 在开集 U 上 C^k 可微, $\partial f(\mathbf{x}_0)$ 可逆, 则存在 \mathbf{x}_0 的邻域 V 和 $f(\mathbf{x}_0)$ 的邻域 W 使得 $f : V \rightarrow W$ 有 C^k 的逆映射。

证明. 考虑 $F(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}$ 。则

$$\partial_{\mathbf{x}} F(\mathbf{x}, \mathbf{y}) = \partial f(\mathbf{x}), \quad \partial_{\mathbf{y}} F(\mathbf{x}, \mathbf{y}) = -I.$$

所以 F 是 C^k 映射。 $F(\mathbf{x}_0, f(\mathbf{x}_0)) = 0$, 因为 $\partial_{\mathbf{x}} F(\mathbf{x}_0, f(\mathbf{x}_0)) = \partial f(\mathbf{x}_0)$ 可逆, 所以存在 \mathbf{x}_0 的邻域 V 以及 $f(\mathbf{x}_0)$ 的邻域 W 和 C^k 映射 $g : W \rightarrow V$ 使得对任意 $\mathbf{y} \in W$ 以及 $\mathbf{x} \in V$,

$$F(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = g(\mathbf{y}),$$

即

$$f(\mathbf{x}) = \mathbf{y} \iff \mathbf{x} = g(\mathbf{y}).$$

□

例 3.2.2. 从平面极坐标 (r, θ) 到直角坐标 (x, y) 由 C^∞ 映射

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

确定。它的 Jacobi 矩阵为

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix},$$

其行列式为 r 。所以根据逆映射定理，在除原点以外的任何一点的某个邻域内，上述映射有 C^∞ 的逆映射，它从直角坐标得到极坐标。 \square

定义 3.2.3. 称 $f: U \rightarrow V$ 是一个 C^k 微分同胚，如果 f 是 C^k 映射，并且存在 C^k 的逆映射 $g: V \rightarrow U$ ，即

$$g(f(\mathbf{x})) = \mathbf{x}, \quad f(g(\mathbf{y})) = \mathbf{y}, \quad \forall \mathbf{x} \in U, \mathbf{y} \in V.$$

定理 3.2.4. 设 $U \subseteq \mathbb{R}^n$ 是开集， $f: U \rightarrow \mathbb{R}^n$ 是 C^k 映射。则

$$V = f(U) \text{ 是开集且 } f: U \rightarrow V \text{ 是 } C^k \text{ 微分同胚}$$

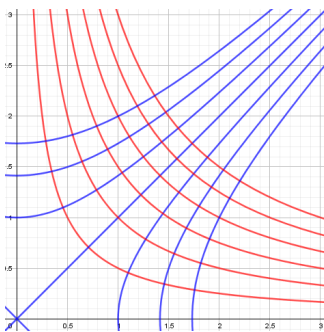
当且仅当

f 是单射，且对任意 $\mathbf{x} \in U$ ， $\partial f(\mathbf{x})$ 是可逆的线性映射。

例 3.2.5.

$$f(x_1, x_2) = (x_1^2 - x_2^2, 2x_1x_2)$$

是 $(0, +\infty) \times (0, +\infty)$ 到 $\mathbb{R} \times (0, +\infty)$ 的微分同胚。



证明.

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}, \quad \det \frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = 4(x_1^2 + x_2^2) > 0, \quad \forall \mathbf{x} \neq \mathbf{0}.$$

再证明 f 是单射。假设

$$(x_1^2 - x_2^2, 2x_1x_2) = (u_1^2 - u_2^2, 2u_1u_2).$$

则

$$(x_1^2 + x_2^2)^2 = (x_1^2 - x_2^2)^2 + 4x_1^2x_2^2 = (u_1^2 - u_2^2)^2 + 4u_1^2u_2^2 = (u_1^2 + u_2^2)^2,$$

所以 $x_1^2 + x_2^2 = u_1^2 + u_2^2$ 。从而 $x_1^2 = u_1^2$, $x_2^2 = u_2^2$, 因此 $x_1 = u_1$, $x_2 = u_2$ 。

事实上, 如果取 $z = x_1 + ix_2$, 则 $z^2 = (x_1^2 - x_2^2) + i2x_1x_2$ 。把映射 $\mathbb{C} \ni z \mapsto z^2 \in \mathbb{C}$ 局限在第一象限就是 f 。□

借助于微分同胚, 我们可以摆脱整体线性坐标系, 从而不再仅限于平直的线性空间, 而在具有局部的曲线坐标系的弯曲的几何空间(曲线、曲面, 更一般地称为流形, manifold)上研究微积分。

习题3.2

1. 设

$$f(x_1, x_2) = (x_1x_2, (1 - x_1)x_2).$$

证明 f 是 $(0, 1) \times (0, 1)$ 到 $\{(y_1, y_2) | y_1 > 0, y_2 > 0, y_1 + y_2 < 1\}$ 的微分同胚。

2. 记 \mathcal{S}_n 为所有 n 阶对称矩阵组成的线性空间, \mathcal{P}_n 为所有 n 阶对称正定矩阵组成的集合。证明

(a) \mathcal{P}_n 是 \mathcal{S}_n 的开子集;

(b) 对任意 $A \in \mathcal{P}_n$, 存在唯一的 $A^{1/2} \in \mathcal{P}_n$ 使得 $(A^{1/2})^2 = A$, 并且 $A^{1/2}$ 关于 A 是 C^∞ 的。

3.3 应用：隐函数和反函数求导法

例 3.3.1 (王竹溪《热力学》第二版1957年, 第一章习题). 设三函数 f, g, h 都是二独立变数 x, y 的函数, 证

$$\left(\frac{\partial f}{\partial g}\right)_h = 1 / \left(\frac{\partial g}{\partial f}\right)_h, \quad \left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial h}\right)_f \left(\frac{\partial h}{\partial f}\right)_g = -1.$$

证明.

$$df = \left(\frac{\partial f}{\partial g}\right)_h dg + \left(\frac{\partial f}{\partial h}\right)_g dh,$$

又

$$dg = \left(\frac{\partial g}{\partial f}\right)_h df + \left(\frac{\partial g}{\partial h}\right)_f dh,$$

代后式入前式，整理得到

$$\begin{aligned} df &= \left(\frac{\partial f}{\partial g}\right)_h \left[\left(\frac{\partial g}{\partial f}\right)_h df + \left(\frac{\partial g}{\partial h}\right)_f dh \right] + \left(\frac{\partial f}{\partial h}\right)_g dh \\ &= \left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial f}\right)_h df + \left[\left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial h}\right)_f + \left(\frac{\partial f}{\partial h}\right)_g \right] dh \end{aligned}$$

视 df 与 dh 独立，比较系数得到

$$\begin{aligned} \left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial f}\right)_h &= 1, \\ \left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial h}\right)_f + \left(\frac{\partial f}{\partial h}\right)_g &= 0, \end{aligned}$$

于是

$$\begin{aligned} \left(\frac{\partial f}{\partial g}\right)_h &= 1 / \left(\frac{\partial g}{\partial f}\right)_h, \\ \left(\frac{\partial f}{\partial g}\right)_h \left(\frac{\partial g}{\partial h}\right)_f \left(\frac{\partial h}{\partial f}\right)_g &= - \left(\frac{\partial h}{\partial f}\right)_g \left(\frac{\partial f}{\partial h}\right)_g = -1. \end{aligned}$$

□

习题3.3

1. 设 f, g, h 都是自变量 x, y 的函数，证明

$$\begin{aligned} \text{(a)} \quad \left(\frac{\partial f}{\partial g}\right)_x &= \left(\frac{\partial f}{\partial y}\right)_x / \left(\frac{\partial g}{\partial y}\right)_x; \\ \text{(b)} \quad \left(\frac{\partial y}{\partial x}\right)_f &= - \left(\frac{\partial f}{\partial x}\right)_y / \left(\frac{\partial f}{\partial y}\right)_x; \\ \text{(c)} \quad \left(\frac{\partial f}{\partial x}\right)_g &= \left(\frac{\partial f}{\partial x}\right)_y + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_g. \end{aligned}$$

2. 设 $f_1, \dots, f_n; u_1, \dots, u_n$ 都是独立变量 x_1, \dots, x_n 的函数。证明

$$\begin{aligned} \text{(a)} \quad \det \frac{\partial(f_1, \dots, f_n)}{\partial(u_1, \dots, u_n)} &= \det \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} / \det \frac{\partial(u_1, \dots, u_n)}{\partial(x_1, \dots, x_n)}; \\ \text{(b)} \quad \left(\frac{\partial f_1}{\partial x_1}\right)_{f_2, \dots, f_n} &= \det \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} / \det \frac{\partial(f_2, \dots, f_n)}{\partial(x_2, \dots, x_n)}; \\ \text{(c)} \quad \left(\frac{\partial f_1}{\partial u_1}\right)_{u_2, \dots, u_n} &= \det \frac{\partial(f_1, u_2, \dots, u_n)}{\partial(x_1, \dots, x_n)} / \det \frac{\partial(u_1, \dots, u_n)}{\partial(x_1, \dots, x_n)}. \end{aligned}$$

3.4 应用：曲面，切平面和法向量

定义 3.4.1. 设 $1 \leq m \leq n-1$. 称 $\Sigma \subset \mathbb{R}^n$ 是一个 m 维 C^k 曲面, 如果对任意 $\mathbf{x}_0 \in \Sigma$, 在 \mathbf{x}_0 的一个邻域中 Σ 是一个 m 元 C^k 映射的图像。

当 $m = n-1$ 时, 称 Σ 是一个 C^k 超曲面。

当 $m = 1$ 时, Σ 是一条 C^k 曲线。

例 3.4.2. (1) [函数图像] C^k 映射 $f: U \rightarrow \mathbb{R}^n$ ($U \subseteq \mathbb{R}^m$ 是区域) 的图像

$$\Sigma = \{(\mathbf{x}, f(\mathbf{x})) | \mathbf{x} \in U\}$$

是 \mathbb{R}^{m+n} 中一个 m 维的 C^k 曲面。

(2) [正则水平集] 如果 a 是 C^k 映射 $F: U \rightarrow \mathbb{R}^n$ ($U \subseteq \mathbb{R}^m$ 是区域) 的一个正则值, 即对满足 $F(\mathbf{x}) = a$ 的任意 $\mathbf{x} \in U$, $\partial F(\mathbf{x})$ 满行秩, 则 F 的水平集

$$\Sigma = \{\mathbf{x} \in U | F(\mathbf{x}) = a\}$$

是 \mathbb{R}^m 中一个 $m-n$ 维的 C^k 曲面。

(3) [参数曲面] 设 $U \subseteq \mathbb{R}^m$ 是开集, C^k 单射 $f: U \rightarrow \mathbb{R}^n$ ($\mathbf{x} = f(u_1, \dots, u_m)$) 满足对任意 $\mathbf{u} \in U$, $\text{rank}(\partial f(\mathbf{u})) = m$, 即 $\partial f(\mathbf{u})$ 满列秩, 则 $f(U)$ 是 \mathbb{R}^n 中一个 m 维 C^k 曲面。 f 称为这个曲面的一个 C^k 参数表示。

证明. (2) 对 $\mathbf{x}_0 \in \Sigma$, 不妨设 $\partial_{\mathbf{v}} F(\mathbf{x}_0)$ 可逆, 其中

$$\mathbf{x} = (\mathbf{u}, \mathbf{v}), \quad \mathbf{u} = (x_1, \dots, x_{m-n}), \quad \mathbf{v} = (x_{m-n+1}, \dots, x_n).$$

于是由隐函数定理存在 \mathbf{x}_0 的邻域以及 C^k 映射 h 使得在该邻域中, $F(\mathbf{x}) = a$ 当且仅当 $\mathbf{v} = h(\mathbf{u})$ 。因此 \mathbf{x}_0 的这个邻域中, Σ 为 h 的图像。

(3) 不妨设在 \mathbf{u}_0 处, $\frac{\partial(x_1, \dots, x_m)}{\partial(u_1, \dots, u_m)}(\mathbf{u}_0)$ 可逆。于是由反函数定理, 在 \mathbf{u}_0 的一个邻域里, $(x_1(\mathbf{u}), \dots, x_m(\mathbf{u}))$ 有 C^k 反函数 $\mathbf{u} = h(x_1, \dots, x_m)$, 于是

$$f(\mathbf{u}) = (x_1, \dots, x_m, x_{m+1}(h(x_1, \dots, x_m)), \dots, x_n(h(x_1, \dots, x_m))).$$

所以 $f(U)$ 是 C^k 曲面。 □

函数图像形式的曲面可以看成是参数方程形式的曲面的特例。

例 3.4.3. 我们以三维欧氏空间中的球面为例。

$$\Sigma: x^2 + y^2 + z^2 = R^2$$

设 $P(x_0, y_0, z_0)$ 是球面 Σ 上一点。则 $x_0^2 + y_0^2 + z_0^2 = R^2 > 0$ 。

(1) 记 $F(x, y, z) = x^2 + y^2 + z^2 - R^2$ 。对球面上任意一点 $P(x_0, y_0, z_0)$,

$$\left(\frac{\partial F}{\partial x}(P), \frac{\partial F}{\partial y}(P), \frac{\partial F}{\partial z}(P) \right) = (2x_0, 2y_0, 2z_0) \neq \mathbf{0},$$

所以 0 是 F 的正则值。根据隐函数定理, 在 P 附近, 球面上点的坐标 x, y, z 中总有一个可以写为另外两个坐标的 C^∞ 函数, 所以 Σ 是个 C^∞ 曲面。

(2) 事实上, 对球面而言, 我们可以直接求解方程, 而不必借助隐函数定理。当 $z_0 > 0$ 时, 曲面 Σ 位于 P 附近的点都满足

$$z = \sqrt{R^2 - x^2 - y^2},$$

当 $z_0 < 0$ 时, 曲面 Σ 位于 P 附近的点都满足

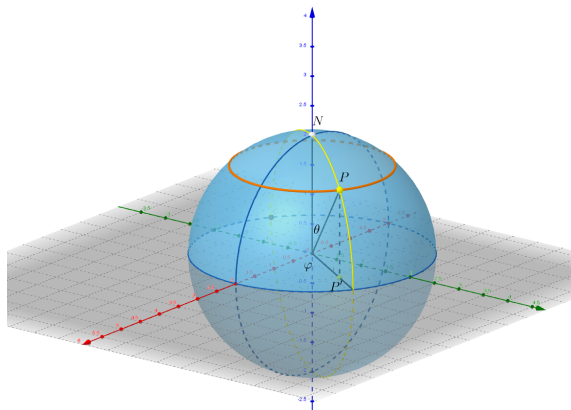
$$z = -\sqrt{R^2 - x^2 - y^2},$$

若 $z_0 = 0$, 则 x_0, y_0 中必有一个非零, 此时我们可以构造相应的 C^∞ 二元函数使得 Σ 在 P 点附近都可以看成为该函数的图像, 所以 Σ 是一个 C^∞ 曲面, 而这些函数给出了曲面 Σ 的局部参数表达。

(3) 另外, 我们可以引入球面的经、纬度参数表达:

$$\begin{cases} x = R \sin \theta \cos \varphi, \\ y = R \sin \theta \sin \varphi, \\ z = R \cos \theta, \end{cases} \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi < 2\pi.$$

θ 是数学纬度, 从北极沿经线旋转至南极; φ 是经度, 沿赤道是从零度经线旋



转一圈。这个参数方程的 Jacobi 矩阵为

$$\begin{pmatrix} \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} R \cos \theta \cos \varphi & -R \sin \theta \sin \varphi \\ R \cos \theta \sin \varphi & R \sin \theta \cos \varphi \\ -R \sin \theta & 0 \end{pmatrix}$$

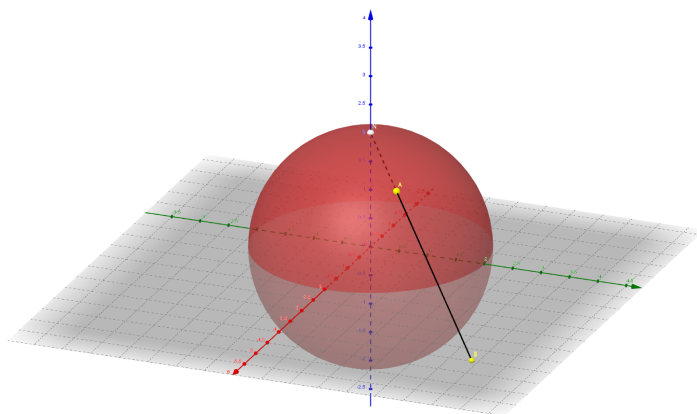
当 $0 < \theta < \pi$ 时, 这个矩阵满列秩。去掉南北两极后, 球面上的点与 $(0, \pi) \times [0, 2\pi)$ 中的 (θ, φ) 一一对应。为了表达南北两极, 可以另选一对对径点再用相应的经纬度即可。因为参数方程的 Jacobi 矩阵满列秩, 所以 x, y, z 中必然有两

个变量与 θ, φ 之间存在着互逆且可微的函数关系，这时第三个变量就成了这两个变量的函数，这再次印证了 Σ 是曲面。

(4) 我们还可以通过球极投影得到球面的参数方程。过 (x, y) 坐标平面上的点 $(u, v, 0)$ 与北极 $(0, 0, R)$ 所连的直线 $(tu, tv, (1-t)R)$ 与球面的交点满足

$$(tu)^2 + (tv)^2 + (R(1-t))^2 = R^2,$$

解得



$$t = \frac{2R^2}{u^2 + v^2 + R^2}$$

所以球面上的点为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{2R^2 u}{u^2 + v^2 + R^2} \\ \frac{2R^2 v}{u^2 + v^2 + R^2} \\ R \left(1 - \frac{2R}{u^2 + v^2 + R^2} \right) \end{pmatrix}$$

这样也得到球面的一个参数方程，这个方程可以刻画除北极以外的球面上的每一个点。请读者自己验证这个参数方程的 Jacobi 矩阵是满列秩的。 \square

问题 3.4.4. 是否存在球面的一个可微参数表示，它可以刻画球面上每一个点？

定义 3.4.5. 设 Σ 是 \mathbb{R}^n 中一个 C^k 曲面， $\mathbf{x}_0 \in \Sigma$ 。 $\gamma : (a, b) \rightarrow \mathbb{R}^n$ 是 Σ 上一条过 \mathbf{x}_0 的 C^1 曲线，即对任意 $t \in (a, b)$ ， $\gamma(t) \in \Sigma$ ，且 $\gamma(t_0) = \mathbf{x}_0$ 。此时称 $\gamma'(t_0) = \partial\gamma(t_0)(1) \in \mathbb{R}^n$ 为 Σ 在 \mathbf{x}_0 处的一个切向量。称 Σ 在 \mathbf{x}_0 处的所有切向量组成的集合为 Σ 在 \mathbf{x}_0 处的切空间，记为 $T_{\mathbf{x}_0}\Sigma$ 。

曲面 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的切平面为 $\mathbf{x}_0 + T_{\mathbf{x}_0}\Sigma$ 。

曲面 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的法空间为 $(T_{\mathbf{x}_0}\Sigma)^\perp$ ，其中的向量称为 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的法向量。

曲面 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的法线/法平面为 $\mathbf{x}_0 + (T_{\mathbf{x}_0}\Sigma)^\perp$ 。

定理 3.4.6. 设 $\Sigma \subseteq \mathbb{R}^n$ 是一个 m 维 C^k 曲面, 则对于任意 $\mathbf{x}_0 \in \Sigma$, 切空间 $T_{\mathbf{x}_0}\Sigma$ 是 \mathbb{R}^n 的一个 m 维线性子空间, 且

1. 若在 $\mathbf{x}_0 \in \Sigma$ 的一个邻域中, Σ 为 C^k 函数 f 的图像: $\{(\mathbf{u}, f(\mathbf{u})) | \mathbf{u} \in U\}$, $\mathbf{x}_0 = (\mathbf{u}_0, f(\mathbf{u}_0))$, 则 $T_{\mathbf{x}_0}\Sigma$ 是线性映射 $\partial f(\mathbf{u}_0)$ 的图像,

$$T_{\mathbf{x}_0}\Sigma = \{(\mathbf{v}, \partial f(\mathbf{u}_0)(\mathbf{v})) | \mathbf{v} \in \mathbb{R}^m\}.$$

切平面方程

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_0 \\ f(\mathbf{u}_0) \end{pmatrix} + \begin{pmatrix} \mathbf{v} \\ \partial f(\mathbf{u}_0)\mathbf{v} \end{pmatrix} \quad \text{或} \quad \mathbf{w} = f(\mathbf{u}_0) + \partial f(\mathbf{u}_0)(\mathbf{u} - \mathbf{u}_0).$$

2. 若 Σ 是一个 C^k 映射 F 的正则水平集, $\Sigma = \{\mathbf{x} \in \mathbb{R}^n | F(\mathbf{x}) = a\}$, 则

$$T_{\mathbf{x}_0}\Sigma = \text{Ker} \partial F(\mathbf{x}_0) = \{\mathbf{v} \in \mathbb{R}^n | \partial F(\mathbf{x}_0)(\mathbf{v}) = 0\}.$$

切平面方程

$$\partial F(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0,$$

梯度向量 $\nabla F(\mathbf{x}_0)$ 是曲面 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的一个法向量。

3. 若 $f: U \rightarrow \mathbb{R}^n$ 是 Σ 的一个 C^k 参数化表示, 则对 $\mathbf{x}_0 = f(\mathbf{u}_0)$,

$$T_{\mathbf{x}_0}\Sigma = \text{Range} \partial f(\mathbf{u}_0) = \{\partial f(\mathbf{u}_0)(\mathbf{v}) \in \mathbb{R}^n | \mathbf{v} \in \mathbb{R}^m\}.$$

切平面方程

$$\mathbf{x} = \mathbf{x}_0 + \partial f(\mathbf{u}_0)(\mathbf{v}), \quad \mathbf{v} \in \mathbb{R}^m,$$

曲面 Σ 在 $\mathbf{x}_0 \in \Sigma$ 处的法空间为 $\text{Ker}((\partial f(\mathbf{x}_0))^T)$ 。

证明. (1) 设 $\gamma(t) = (\mathbf{u}(t), \mathbf{y}(t))$ 是 Σ 上一条 C^1 曲线。则 $\mathbf{y}(t) = f(\mathbf{u}(t))$ 。于是

$$\gamma'(t_0) = (\mathbf{u}'(t_0), \mathbf{y}'(t_0)) = (\mathbf{u}'(t_0), \partial f(\mathbf{u}(t_0))\mathbf{u}'(t_0)).$$

另一方面, 对任意 $\mathbf{v} \in \mathbb{R}^m$,

$$\gamma(t) = (\mathbf{u}_0 + t\mathbf{v}, f(\mathbf{u}_0 + t\mathbf{v}))$$

是 Σ 上一条 C^1 曲线,

$$\gamma'(0) = (\mathbf{v}, \partial f(\mathbf{u}_0)\mathbf{v}).$$

所以

$$T_{\mathbf{x}_0}\Sigma = \{(\mathbf{v}, \partial f(\mathbf{u}_0)\mathbf{v}) | \mathbf{v} \in \mathbb{R}^m\}$$

是一个 m 维线性空间。

(2) 由隐函数定理, 不妨设在 \mathbf{x}_0 附近, $F(\mathbf{x}) = 0$ 当且仅当 $\mathbf{x} = (\mathbf{u}, f(\mathbf{u}))$, 其中 $f: U \rightarrow \mathbb{R}^{n-m}$ 是 C^k 映射。于是由(1)知

$$T_{\mathbf{x}_0}\Sigma = \{(\mathbf{v}, \partial f(\mathbf{u}_0)\mathbf{v}) \mid \mathbf{v} \in \mathbb{R}^m\}.$$

而

$$\partial f(\mathbf{u}_0) = -\partial_{\mathbf{y}}F(\mathbf{x}_0)^{-1}\partial_{\mathbf{u}}F(\mathbf{x}_0),$$

所以

$$0 = \partial F(\mathbf{x}_0)(\mathbf{u}, \mathbf{y}) = \partial_{\mathbf{u}}F(\mathbf{x}_0)\mathbf{u} + \partial_{\mathbf{y}}F(\mathbf{x}_0)\mathbf{y}$$

当且仅当

$$0 = \partial_{\mathbf{y}}F(\mathbf{x}_0)^{-1}\partial_{\mathbf{u}}F(\mathbf{x}_0)\mathbf{u} + \mathbf{y} = -\partial f(\mathbf{u}_0)\mathbf{u} + \mathbf{y}.$$

因此

$$T_{\mathbf{x}_0}\Sigma = \text{Ker}\partial F(\mathbf{x}_0).$$

设 $\gamma(t)$ 是 Σ 上一条 C^1 曲线, 满足 $\mathbf{x}_0 = \gamma(t_0)$, $\mathbf{v} = \gamma'(t_0)$ 。则 $F(\gamma(t)) = 0$, 于是

$$0 = \left. \frac{\partial F(\gamma(t))}{\partial t} \right|_{t=t_0} = \partial F(\mathbf{x}_0)\mathbf{v}.$$

所以 $\mathbf{v} \in \text{Ker}\partial F(\mathbf{x}_0)$, $T_{\mathbf{x}_0}\Sigma \subseteq \text{Ker}\partial F(\mathbf{x}_0)$ 。

(3) 证明留给读者完成。 \square

对参数曲面 $f: U \rightarrow \mathbb{R}^n$ 而言,

$$(u_1, \dots, u_m) \mapsto f(\mathbf{u}_0 + u_1\mathbf{e}_1 + \dots + u_m\mathbf{e}_m)$$

在曲面 Σ 上局部引入了一个曲线坐标系。

$$\frac{\partial f}{\partial u_1}(\mathbf{u}_0), \dots, \frac{\partial f}{\partial u_m}(\mathbf{u}_0)$$

是切空间 $T_{\mathbf{x}_0}\Sigma$ 的一组基底向量, 与它们都正交的向量就是曲面的法向量。当这个曲面是超曲面时, 法空间维数是 1。

如果 $(x_1, x_2, \dots, x_{m+1})$ 是直角坐标系, 曲面参数方程为

$$\begin{cases} x_1 = x_1(u_1, \dots, u_m), \\ x_2 = x_2(u_1, \dots, u_m), \\ \vdots \\ x_{m+1} = x_{m+1}(u_1, \dots, u_m), \end{cases}$$

则

$$\mathbf{n} = \begin{pmatrix} (-1)^0 \frac{\partial \hat{x}_1}{\partial (u_1, \dots, u_m)} \\ (-1)^1 \frac{\partial \hat{x}_2}{\partial (u_1, \dots, u_m)} \\ \vdots \\ (-1)^m \frac{\partial \hat{x}_{m+1}}{\partial (u_1, \dots, u_m)} \end{pmatrix}$$

为曲面的法向量, 其中 \hat{x}_k 表示从 $(x_1, x_2, \dots, x_{m+1})$ 中删除 x_k 后所剩的 m 个坐标。这是因为

$$\left\langle \frac{\partial \mathbf{x}}{\partial u_k}, \mathbf{n} \right\rangle = \begin{vmatrix} \frac{\partial x_1}{\partial u_k} & \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_m} \\ \frac{\partial x_2}{\partial u_k} & \frac{\partial x_2}{\partial u_1} & \cdots & \frac{\partial x_2}{\partial u_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_{m+1}}{\partial u_k} & \frac{\partial x_{m+1}}{\partial u_1} & \cdots & \frac{\partial x_{m+1}}{\partial u_m} \end{vmatrix} = 0.$$

例 3.4.7. 我们仍以三维欧氏空间中的球面为例, 计算曲面的切平面和法线。

$$\Sigma : x^2 + y^2 + z^2 = R^2$$

设 $P(x_0, y_0, z_0)$ 是球面 Σ 上一点。则 $x_0^2 + y_0^2 + z_0^2 = R^2 > 0$ 。

(1) 若 $z_0 > 0$, 则 $x_0^2 + y_0^2 < R^2$, 此时曲面 Σ 位于 P 附近的点都满足

$$z = \sqrt{R^2 - x^2 - y^2},$$

其微分为

$$\partial z(x_0, y_0) = \frac{-2x_0 dx - 2y_0 dy}{2\sqrt{R^2 - x_0^2 - y_0^2}} = -\frac{x_0}{z_0} dx - \frac{y_0}{z_0} dy.$$

所以切平面方程为

$$z - z_0 = -\frac{x_0}{z_0}(x - x_0) - \frac{y_0}{z_0}(y - y_0),$$

即

$$x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0) = 0,$$

后者也适用于 $z_0 \leq 0$ 的情况。由此得到曲面的法向量

$$\mathbf{n} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix},$$

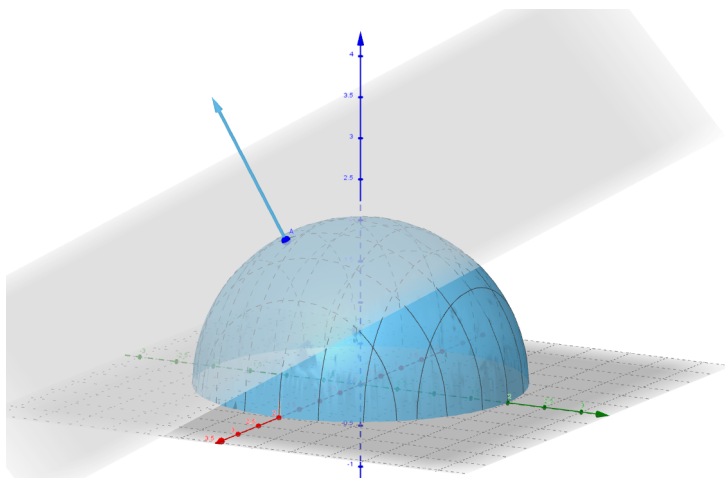
法线为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P + t\mathbf{n} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

这也适用于 $z_0 \leq 0$ 的情况。

(2) 对 $F(x, y, z) = x^2 + y^2 + z^2 - R^2$, F 在 $P(x_0, y_0, z_0)$ 的梯度

$$\nabla F(P) = \begin{pmatrix} \frac{\partial F}{\partial x}(P) \\ \frac{\partial F}{\partial y}(P) \\ \frac{\partial F}{\partial z}(P) \end{pmatrix} = \begin{pmatrix} 2x_0 \\ 2y_0 \\ 2z_0 \end{pmatrix},$$



它就是曲面 $F = 0$ 在 P 的法向量，因此曲面的切平面方程为

$$2x_0(x - x_0) + 2y_0(y - y_0) + 2z_0(z - z_0) = 0.$$

(3) 对参数方程，

$$\begin{cases} x = R \sin \theta \cos \varphi, \\ y = R \sin \theta \sin \varphi, \\ z = R \cos \theta, \end{cases} \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi < 2\pi.$$

Σ 的一对切向量为

$$\begin{pmatrix} \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \theta} \end{pmatrix} = \begin{pmatrix} R \cos \theta \cos \varphi \\ R \cos \theta \sin \varphi \\ -R \sin \theta \end{pmatrix}, \quad \begin{pmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} -R \sin \theta \sin \varphi \\ R \sin \theta \cos \varphi \\ 0 \end{pmatrix},$$

切平面方程为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} R \sin \theta \cos \varphi \\ R \sin \theta \sin \varphi \\ R \cos \theta \end{pmatrix} + t \begin{pmatrix} R \cos \theta \cos \varphi \\ R \cos \theta \sin \varphi \\ -R \sin \theta \end{pmatrix} + s \begin{pmatrix} -R \sin \theta \sin \varphi \\ R \sin \theta \cos \varphi \\ 0 \end{pmatrix}, \quad t, s \in \mathbb{R}.$$

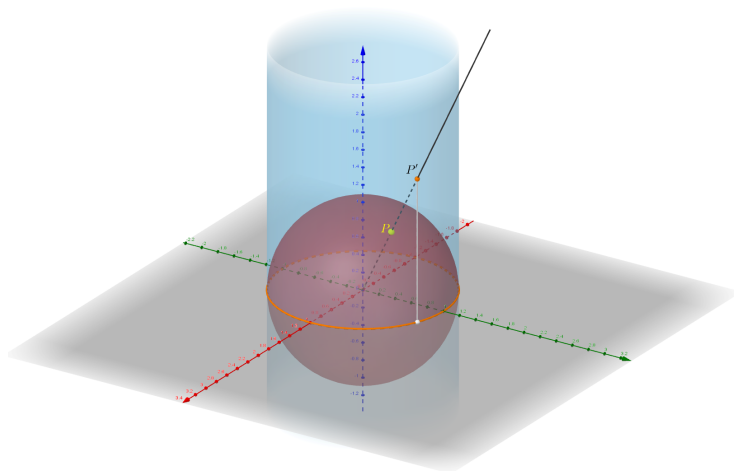
法向量为

$$\mathbf{n} = \begin{pmatrix} \det \frac{\partial(y,z)}{\partial(\theta,\varphi)} \\ \det \frac{\partial(z,x)}{\partial(\theta,\varphi)} \\ \det \frac{\partial(x,y)}{\partial(\theta,\varphi)} \end{pmatrix} = R^2 \sin \theta \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}.$$

□

习题3.4

1. 证明定理3.4.6。
2. 证明 $\{(x, y) \in \mathbb{R}^2 \mid -y + x^2 + xy + x^4 + y^4 = 0\}$ 是一条有界闭的 C^∞ 曲线。
3. 绘制平面地图有很多方法, Mercator 投影是其中之一。如图所示, 它把球面上的点 $P(x, y, z)$ 映为从球心出发的射线 OP 与一个圆柱面的交点 $P'(\lambda x, \lambda y, \lambda z)$, 然后在柱面上再做一个变换得到点 $Q(\lambda x, \lambda y, f(\lambda z))$, 该圆柱面与球面相切于赤道。这样除南北两极外, 地球上的点 P 与柱面上的点 Q 一一对应, 把圆柱面沿母线剪开就得到一张地图。求函数 f 的表



达式使得 Mercator 地图上所画的两条相交道路的夹角与实际道路的夹角大小相等。两条相交曲线的夹角被定义为它们的切线所形成的夹角(锐角或直角)。(提示: 线性变换 $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 保角当且仅当它把正交且长度相等的向量映为正交且长度相等的向量, 从而存在常数 $C > 0$ 使得内积 $\langle A\mathbf{u}, A\mathbf{v} \rangle = C\langle \mathbf{u}, \mathbf{v} \rangle, \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$)

4. 设 P_0 是曲面 Σ 上一点, P 是曲面 Σ 上靠近 P_0 的一点。 $P - P_0 = \mathbf{v} + \mathbf{w}$, 其中 \mathbf{v} 是曲面的切向量, \mathbf{w} 是曲面的法向量。证明当 $P \rightarrow P_0$ 时, $\|\mathbf{w}\| = o(\|\mathbf{v}\|)$ 。

3.5 应用: 条件极值与 Lagrange 乘子法

条件极值问题就是在满足一定的约束条件下求一个目标函数的极值。

例如, 根据连续函数的性质, 在一个有界闭区域 Ω 上, 连续函数 f 有最大值和最小值。如果这些最值点在 Ω 的内部, 那么它们是 f 的临界点。在 Ω 的内点 \mathbf{x}_0 附近, \mathbf{x} 可以从 \mathbf{x}_0 出发往各个方向自由地移动一段小距离, 不受约束, 这样的极值点称为无条件极值点。但是, f 的最大值或最小值也可能在 Ω 的边

界 $\partial\Omega$ 处达到。在 Ω 的边界点 \mathbf{x}_0 附近，沿 $\partial\Omega$ 移动的点 \mathbf{x} 会受到限制，这样的极值点称为条件极值点。

下面我们考虑带约束条件的极值问题

$$\begin{cases} \max \text{ 或 } \min & f(\mathbf{x}), \\ \text{s.t.} & g_1(\mathbf{x}) = 0, \\ & g_2(\mathbf{x}) = 0, \\ & \vdots \\ & g_r(\mathbf{x}) = 0. \end{cases}$$

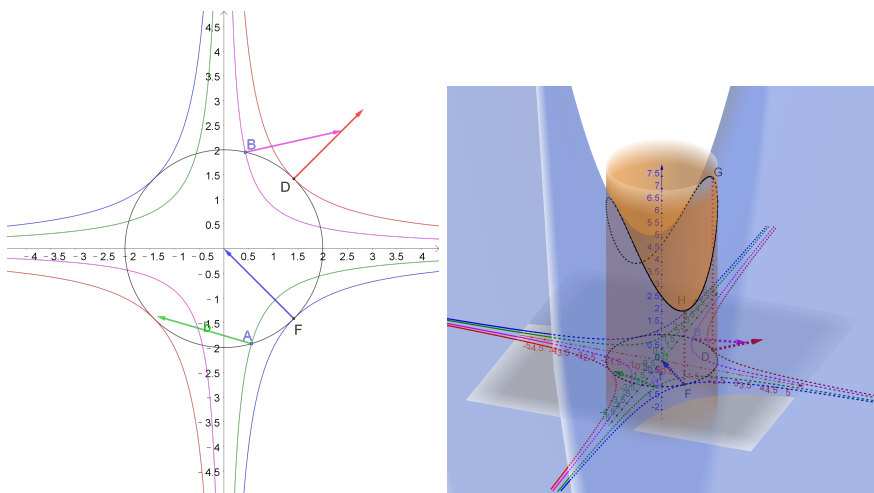
其中 f 是 C^1 函数， $g = (g_1, \dots, g_r)^T$ 是 C^1 映射。

设 \mathbf{x}^* 是该条件极值问题的解，假设 $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_r(\mathbf{x}^*)$ 线性无关（即约束条件之间没有必然联系）。根据隐函数定理，此时 $g = 0$ 在 \mathbf{x}^* 附近确定了一个 C^1 曲面 Σ 。如果 $\nabla f(\mathbf{x}^*)$ 在 Σ 的切空间中有非零投影 \mathbf{v} ，则存在 Σ 中过 \mathbf{x}^* 的 C^1 曲线 $\mathbf{x}(t)$ ，使得 $\mathbf{x}(0) = \mathbf{x}^*$ ， $\mathbf{x}'(0) = \mathbf{v}$ ，于是

$$\left. \frac{\partial f(\mathbf{x}(t))}{\partial t} \right|_{t=0} = \langle \nabla f(\mathbf{x}^*), \mathbf{v} \rangle \neq 0.$$

因此 f 沿着曲线 $\mathbf{x}(t)$ 不可能在 \mathbf{x}^* 取得极值，这与 \mathbf{x}^* 是 f 的条件极值点矛盾。所以， $\nabla f(\mathbf{x}^*)$ 只能在 Σ 的法空间中，而 $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_r(\mathbf{x}^*)$ 是 Σ 的法空间的一组基，所以存在 $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ 使得

$$\nabla f(\mathbf{x}^*) = \lambda_1 \nabla g_1(\mathbf{x}^*) + \dots + \lambda_r \nabla g_r(\mathbf{x}^*).$$



例 3.5.1. 求在约束条件 $x^2 + y^2 = 4$ 下，函数 $f(x, y) = xy + 5$ 的最大值和最小值。

解: 约束条件是圆周, 可以取参数方程 $x = 2 \cos \theta, y = 2 \sin \theta$, 于是

$$f(2 \cos \theta, 2 \sin \theta) = 4 \cos \theta \sin \theta + 5 = 5 + 2 \sin 2\theta.$$

所以 f 的最大值为 7, 在 $\theta = \frac{\pi}{4}$ 或 $\theta = \frac{5\pi}{4}$, 即 $(x, y) = (\sqrt{2}, \sqrt{2})$ 和 $(x, y) = (-\sqrt{2}, -\sqrt{2})$ 时取得. f 的最小值为 3, 在 $\theta = \frac{3\pi}{4}$ 或 $\theta = \frac{7\pi}{4}$, 即 $(x, y) = (-\sqrt{2}, \sqrt{2})$ 和 $(x, y) = (\sqrt{2}, -\sqrt{2})$ 时取得.

目标函数 f 和约束函数 g 的梯度为

$$\nabla f = (y, x), \quad \nabla g = (2x, 2y),$$

它们共线当且仅当 $x^2 = y^2 = 2$, 此时 f 的水平集与约束条件 $g = 0$ 相切. \square

因为约束条件 $g = 0$ 通常不容易求解, 所以 Lagrange 提出可以考虑如下扩展函数

$$F(\mathbf{x}, \lambda_1, \dots, \lambda_r) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \dots - \lambda_r g_r(\mathbf{x}).$$

F 的临界点满足

$$\begin{cases} \nabla_{\mathbf{x}} F(\mathbf{x}, \lambda_1, \dots, \lambda_r) = \nabla f(\mathbf{x}) - \lambda_1 \nabla g_1(\mathbf{x}) - \dots - \lambda_r \nabla g_r(\mathbf{x}) = 0, \\ \frac{\partial F}{\partial \lambda_i}(\mathbf{x}, \lambda_1, \dots, \lambda_r) = -g_i(\mathbf{x}) = 0, \quad i = 1, \dots, r. \end{cases}$$

其中第一个等式对应于条件极值点的必要条件, 后面的等式恰好是约束条件. 这样, 就把条件极值问题转化为扩展函数 F 的无约束极值问题. 这个方法叫做“Lagrange 乘子法”, 其中 λ 叫做 Lagrange 乘子.

另解: 连续函数 f 在有界闭集 $g = 0$ 上必然有最大值和最小值. 考虑扩展函数

$$F(x, y, \lambda) = xy + 5 - \lambda(x^2 + y^2 - 4).$$

求导得到

$$\begin{cases} \frac{\partial F}{\partial x}(x, y, \lambda) = y - 2\lambda x = 0, \\ \frac{\partial F}{\partial y}(x, y, \lambda) = x - 2\lambda y = 0, \\ \frac{\partial F}{\partial \lambda}(x, y, \lambda) = -(x^2 + y^2 - 4) = 0. \end{cases}$$

因此

$$y^2 = 2\lambda xy = x^2 = 2,$$

解得 $|x| = |y| = \sqrt{2}$. 而

$$F(\sqrt{2}, \sqrt{2}) = F(-\sqrt{2}, -\sqrt{2}) = 7, \quad F(-\sqrt{2}, \sqrt{2}) = F(\sqrt{2}, -\sqrt{2}) = 3,$$

所以 7 是 f 的最大值, 3 是 f 的最小值. \square

例 3.5.2. 设 A 是对称矩阵, 求

$$\begin{cases} \max & \langle A\mathbf{x}, \mathbf{x} \rangle, \\ \text{s.t.} & \langle \mathbf{x}, \mathbf{x} \rangle = 1 \end{cases}$$

的解。

解: 因为 $K = \{\mathbf{x} | \langle \mathbf{x}, \mathbf{x} \rangle = 1\}$ 是 \mathbb{R}^n 中的有界闭集, $\langle A\mathbf{x}, \mathbf{x} \rangle$ 关于 \mathbf{x} 连续, 所以在 K 上, $\langle A\mathbf{x}, \mathbf{x} \rangle$ 有最大值。

令 $F(\mathbf{x}, \lambda) = \langle A\mathbf{x}, \mathbf{x} \rangle - \lambda(\langle \mathbf{x}, \mathbf{x} \rangle - 1)$ 。则

$$\begin{cases} \nabla_{\mathbf{x}} F(\mathbf{x}, \lambda) = 2[A - \lambda I]\mathbf{x} = 0, \\ \frac{\partial F}{\partial \lambda}(\mathbf{x}, \lambda) = 1 - \langle \mathbf{x}, \mathbf{x} \rangle = 0. \end{cases}$$

所以条件极值点 \mathbf{x} 是 A 的单位特征向量, 相应的 Lagrange 乘子 λ 是 A 的相应的特征值。原条件极值的解为 A 的最大特征值对应的单位特征向量。 \square

例 3.5.3. 证明对任何 $p > 1$ 以及非负实数 $x_1, y_1, \dots, x_n, y_n$,

$$(x_1^p + \dots + x_n^p)^{\frac{1}{p}} + (y_1^p + \dots + y_n^p)^{\frac{1}{p}} \geq [(x_1 + y_1)^p + \dots + (x_n + y_n)^p]^{\frac{1}{p}},$$

等式成立当且仅当 (x_1, \dots, x_n) 与 (y_1, \dots, y_n) 线性相关。

证明. 考虑

$$\begin{cases} \max & (x_1 + y_1)^p + \dots + (x_n + y_n)^p, \\ \text{s.t.} & x_1^p + \dots + x_n^p = \alpha, \\ & y_1^p + \dots + y_n^p = \beta. \end{cases}$$

$x_1^p + \dots + x_n^p$, $y_1^p + \dots + y_n^p$ 和 $(x_1 + y_1)^p + \dots + (x_n + y_n)^p$ 都是关于 (\mathbf{x}, \mathbf{y}) 的连续函数, 所以 $K = \{(\mathbf{x}, \mathbf{y}) | x_1^p + \dots + x_n^p = \alpha, y_1^p + \dots + y_n^p = \beta\}$ 是有界闭集, $(x_1 + y_1)^p + \dots + (x_n + y_n)^p$ 在 K 上存在最大值。

令

$$F(\mathbf{x}, \mathbf{y}, \lambda, \mu) = \sum_{k=1}^n (x_k + y_k)^p - \lambda \left[\sum_{k=1}^n x_k^p - \alpha \right] - \mu \left[\sum_{k=1}^n y_k^p - \beta \right].$$

则由

$$\begin{cases} \frac{\partial F}{\partial x_k} = p(x_k + y_k)^{p-1} - \lambda p x_k^{p-1} = 0, \\ \frac{\partial F}{\partial y_k} = p(x_k + y_k)^{p-1} - \mu p y_k^{p-1} = 0, \\ \frac{\partial F}{\partial \lambda} = \alpha - \sum_{k=1}^n x_k^p = 0, \\ \frac{\partial F}{\partial \mu} = \beta - \sum_{k=1}^n y_k^p = 0, \end{cases}$$

得到:

情形1:

$$x_k + y_k = 0, \quad k = 1, \dots, n,$$

即所有 $x_k = y_k = 0$, 此时 $\sum_{k=1}^n (x_k + y_k)^p = 0$;

或者情形2: 至少有一个 k_0 使得 $x_{k_0} + y_{k_0} > 0$, 此时 $\lambda, \mu > 0$ 且对任意 k :

$$x_k = \frac{x_k + y_k}{\lambda^{\frac{1}{p-1}}}, \quad y_k = \frac{x_k + y_k}{\mu^{\frac{1}{p-1}}}, \quad \frac{1}{\lambda^{\frac{1}{p-1}}} + \frac{1}{\mu^{\frac{1}{p-1}}} = 1,$$

$$\alpha = \sum_{k=1}^n \frac{(x_k + y_k)^p}{\lambda^{\frac{p}{p-1}}}, \quad \beta = \sum_{k=1}^n \frac{(x_k + y_k)^p}{\mu^{\frac{p}{p-1}}}.$$

$$\alpha^{\frac{1}{p}} + \beta^{\frac{1}{p}} = \left(\sum_{k=1}^n (x_k + y_k)^p \right)^{\frac{1}{p}} \left(\frac{1}{\lambda^{\frac{1}{p-1}}} + \frac{1}{\mu^{\frac{1}{p-1}}} \right) = \left(\sum_{k=1}^n (x_k + y_k)^p \right)^{\frac{1}{p}}.$$

易见情形2时, $(\alpha^{\frac{1}{p}} + \beta^{\frac{1}{p}})^p$ 是条件极值问题的最大值, 所以

$$\left(\sum_{k=1}^n (x_k + y_k)^p \right)^{\frac{1}{p}} \leq \alpha^{\frac{1}{p}} + \beta^{\frac{1}{p}},$$

其中等号成立当且仅当 (x_1, \dots, x_n) 与 (y_1, \dots, y_n) (正)线性相关。 \square

习题3.5

1. 设 α, β, γ 都是正数, 求实数 t 使得函数

$$f(x, y, z) = \frac{x^\alpha y^\beta z^\gamma}{(x + y + z)^t}$$

在集合 $D = \{(x, y, z) | x > 0, y > 0, z > 0\}$ 上有上界, 并求 f 的上确界的值。

2. 求 $xy^3 - x$ 在闭区域 $\{(x, y) | x^2 + y^2 \leq 1\}$ 上的最大值和最小值。
3. 设 A 是由 n 个 n 维单位列向量 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 组成的 n 阶矩阵, 证明 $(\det A)^2 \leq 1$, 其中等号成立当且仅当 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 是 \mathbb{R}^n 的一组单位正交基。
4. (1) Hölder不等式。设正数 p, q 满足 $\frac{1}{p} + \frac{1}{q} = 1$ 。证明对任何非负实数 $x_1, y_1, \dots, x_n, y_n$,

$$(x_1^p + \dots + x_n^p)^{\frac{1}{p}} (y_1^q + \dots + y_n^q)^{\frac{1}{q}} \geq x_1 y_1 + \dots + x_n y_n,$$

并讨论等式成立的条件。

- (2) 证明对任何非负实数 $x_1, y_1, \dots, x_n, y_n$,

$$(x_1 + \dots + x_n) \max\{y_1, \dots, y_n\} \geq x_1 y_1 + \dots + x_n y_n,$$

并讨论等式成立的条件。

- (3) 对线性函数 $L: \mathbb{R}^n \rightarrow \mathbb{R}$ 及范数 $\|\cdot\|$, 设 \mathbf{x}^* 是 $L(\mathbf{x})$ 在单位球面 $S = \{\mathbf{x} | \|\mathbf{x}\| = 1\}$ 上的最大值点, 我们称 $L(\mathbf{x}^*)\mathbf{x}^*$ 为 L 在范数 $\|\cdot\|$ 下的梯度向量。对 $p > 1$ 或 $p = +\infty$, 求线性函数 $L(x, y) = ax + by$ 在 p 范数 $\|\cdot\|_p$ 下的梯度向量。
5. (熵) 求 $-\sum_{i=1}^n p_i \ln p_i$ 的最大值, 其中 p_1, \dots, p_n 是正数且满足 $\sum_{i=1}^n p_i = 1$ 。

6. 将长为 L 的铁丝分成三段，依次围成圆、正方形、正三角形，三个图形面积总和是否存在最小值？若存在，求最小值。
7. 设 (x, y) 满足 $-y + x^2 + xy + x^4 + y^4 = 0$ ，求 y 的最小值。
8. 设 a, b, c 为一个三角形的三条边的边长，求 $\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b}$ 的取值范围。
9. 求抛物线 $y = x^2$ 上一点 P 使得从 $A(0, 1)$ 经 P 到 $B(1/2, 1)$ 的折线长度最小。
10. 如何确定条件极值点的极值类型？

3.6 附：隐函数定理的另一个证明

对有限维空间，以下证明用数学归纳法和解方程组的消元法，比较初等。

定理 3.6.1. 设 $W \subseteq \mathbb{R}^n \times \mathbb{R}$ 是开集， $F: W \rightarrow \mathbb{R}$ 是 C^1 映射，

$$F(\mathbf{x}_0, y_0) = 0, \quad \frac{\partial F}{\partial y}(\mathbf{x}_0, y_0) \neq 0,$$

则存在 $\delta_1 > 0$ 和 $\delta_2 > 0$ 以及 C^1 函数 $f: \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\| < \delta_1\} \rightarrow (y_0 - \delta_2, y_0 + \delta_2)$ 使得：对任意 \mathbf{x}, y ，只要 $\|\mathbf{x} - \mathbf{x}_0\| < \delta_1$ 且 $|y - y_0| < \delta_2$ ，就有

$$F(\mathbf{x}, y) = 0 \text{ 当且仅当 } y = f(\mathbf{x}).$$

证明. 不妨设 $\partial_{\mathbf{x}} F(\mathbf{x}_0, y_0) = 0, \frac{\partial F}{\partial y}(\mathbf{x}_0, y_0) = 1$ ，否则可以考虑

$$\tilde{F}(\mathbf{x}, y) = \frac{1}{\frac{\partial F}{\partial y}(\mathbf{x}_0, y_0)} [F(\mathbf{x}, y) - \partial_{\mathbf{x}} F(\mathbf{x}_0, y_0)(\mathbf{x} - \mathbf{x}_0)].$$

由于 F 是 C^1 映射，所以存在 $\delta_3, \delta_2 > 0$ 使得对任意 \mathbf{x}, y ，只要 $\|\mathbf{x} - \mathbf{x}_0\| < \delta_3$ 且 $|y - y_0| < \delta_2$ ，就有 $\frac{\partial F}{\partial y}(\mathbf{x}, y) > 0$ ，所以 $F(\mathbf{x}, y)$ 关于 y 是严格增函数。

任取 $y_1 \in (y_0 - \delta_2, y_0)$ 和 $y_2 \in (y_0, y_0 + \delta_2)$ ，则

$$F(\mathbf{x}_0, y_1) < F(\mathbf{x}_0, y_0) = 0 < F(\mathbf{x}_0, y_2).$$

由于 F 连续，所以存在 $\delta_1(y_1, y_2) > 0$ 使得 $0 < \delta_1 < \delta_3$ ，且对任意 \mathbf{x} ，只要 $\|\mathbf{x} - \mathbf{x}_0\| < \delta_1$ ，就有

$$F(\mathbf{x}, y_1) < 0 < F(\mathbf{x}, y_2).$$

于是由介值定理和 F 对 y 的单调性，存在唯一的 $y = f(\mathbf{x}) \in (y_1, y_2)$ 使得 $F(\mathbf{x}, f(\mathbf{x})) = 0$ 。事实上， $f(\mathbf{x})$ 是 $F(\mathbf{x}, y)$ 关于 y 在 $(y_0 - \delta_2, y_0 + \delta_2)$ 中的唯一零点，于是 $f(\mathbf{x}_0) = y_0$ 。

对任意 $0 < \varepsilon < \delta_3$ ，取 $\delta_1(\varepsilon) = \delta_1(y_0 - \varepsilon, y_0 + \varepsilon)$ ，则对任意 $\|\mathbf{x} - \mathbf{x}_0\| < \delta_1(\varepsilon)$ ， $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \varepsilon$ 。从而 f 在 \mathbf{x}_0 连续。

注意到 $F(\mathbf{x}_0, y_0) = 0, \partial_{\mathbf{x}} F(\mathbf{x}_0, y_0) = 0, \frac{\partial F}{\partial y}(\mathbf{x}_0, y_0) = 1$ ，所以 F 在 (\mathbf{x}_0, y_0) 的 Taylor 展开为

$$F(\mathbf{x}, y) = (y - y_0) + \alpha(\mathbf{x}, y),$$

其中 $\alpha(\mathbf{x}, y) = o(\|\mathbf{x} - \mathbf{x}_0\| + |y - y_0|)$ 。于是

$$0 = F(\mathbf{x}, f(\mathbf{x})) = (f(\mathbf{x}) - f(\mathbf{x}_0)) + \alpha(\mathbf{x}, f(\mathbf{x})),$$

从而对任意 $\varepsilon > 0$ ，存在 $\delta > 0$ 使得当 $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ 时，

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = |-\alpha(\mathbf{x}, f(\mathbf{x}))| \leq \varepsilon (\|\mathbf{x} - \mathbf{x}_0\| + |f(\mathbf{x}) - f(\mathbf{x}_0)|),$$

因此

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| \leq \frac{\varepsilon}{1 - \varepsilon} \|\mathbf{x} - \mathbf{x}_0\|,$$

即 $f(\mathbf{x}) = f(\mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$, 从而 f 在 \mathbf{x}_0 处可微。

把上述论述应用于 $(\mathbf{x}, f(\mathbf{x}))$, 由 f 的唯一性可得 f 在 \mathbf{x} 处可微。再由 $F(\mathbf{x}, f(\mathbf{x})) = 0$ 求导得到

$$\partial_{\mathbf{x}} f(\mathbf{x}) = -\frac{\partial_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x}))}{\frac{\partial F}{\partial y}(\mathbf{x}, f(\mathbf{x}))}$$

所以 $\partial_{\mathbf{x}} f$ 连续, 因此 f 是 C^1 函数。 \square

现在看高维情形。

定理 3.6.2. 设 $W \subseteq \mathbb{R}^n \times \mathbb{R}^m$ 是开集, $F: W \rightarrow \mathbb{R}^m$ 是 C^k 映射,

$$F(\mathbf{x}_0, \mathbf{y}_0) = 0, \quad \partial_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0) \text{ 可逆},$$

则存在 $(\mathbf{x}_0, \mathbf{y}_0)$ 的邻域 U, V 以及 C^k 映射 $f: U \rightarrow V$ 使得当 $(\mathbf{x}, \mathbf{y}) \in U \times V$ 时,

$$F(\mathbf{x}, \mathbf{y}) = 0 \text{ 当且仅当 } \mathbf{y} = f(\mathbf{x}).$$

证明. 先考虑 $k = 1$. 对 m 做数学归纳法。前述定理表明现在这个定理在 $m = 1$ 时成立。假设 $m - 1$ 时定理成立。

考虑

$$G(\mathbf{x}, \mathbf{y}) = (\partial_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0))^{-1} F(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \tilde{G}(\mathbf{x}, \mathbf{u}, v) \\ G_m(\mathbf{x}, \mathbf{u}, v) \end{pmatrix}, \quad \mathbf{y} = (\mathbf{u}, v) \in \mathbb{R}^{m-1} \times \mathbb{R}.$$

则 $F(\mathbf{x}, \mathbf{y}) = 0$ 当且仅当 $G(\mathbf{x}, \mathbf{y}) = 0$.

$$G(\mathbf{x}_0, \mathbf{u}_0, v_0) = 0, \quad \partial_{\mathbf{u}} \tilde{G}(\mathbf{x}_0, \mathbf{u}_0, v_0) = I_{m-1},$$

$$\partial_{\mathbf{u}} G_m(\mathbf{x}_0, \mathbf{u}_0, v_0) = 0, \quad \frac{\partial G_m}{\partial v}(\mathbf{x}_0, \mathbf{u}_0, v_0) = 1.$$

由归纳假设, 存在 C^1 映射 \tilde{f} 使得在 $(\mathbf{x}_0, \mathbf{u}_0, v_0)$ 附近, $\tilde{G}(\mathbf{x}, \mathbf{u}, v) = 0$ 当且仅当 $\mathbf{u} = \tilde{f}(\mathbf{x}, v)$. 因此 $G(\mathbf{x}, \mathbf{u}, v) = 0$ 当且仅当 $G_m(\mathbf{x}, \tilde{f}(\mathbf{x}, v), v) = 0$.

$$\begin{aligned} & \left. \frac{\partial}{\partial v} \left(G_m(\mathbf{x}, \tilde{f}(\mathbf{x}, v), v) \right) \right|_{(\mathbf{x}_0, \tilde{f}(\mathbf{x}_0, v_0), v_0)} \\ &= \partial_{\mathbf{u}} G_m(\mathbf{x}_0, \mathbf{u}_0, v_0) \frac{\partial \tilde{f}}{\partial v}(\mathbf{x}_0, v_0) + \frac{\partial G_m}{\partial v}(\mathbf{x}_0, \mathbf{u}_0, v_0) = 1, \end{aligned}$$

所以由前述定理, 存在 C^1 函数 f_m 使得在 (\mathbf{x}_0, v_0) 附近, $G_m(\mathbf{x}, \tilde{f}(\mathbf{x}, v), v) = 0$ 当且仅当 $v = f_m(\mathbf{x})$. 取 $f(\mathbf{x}) = (\tilde{f}(\mathbf{x}, f_m(\mathbf{x})), f_m(\mathbf{x}))$, 则在 $(\mathbf{x}_0, \mathbf{y}_0)$ 附近, $G(\mathbf{x}, \mathbf{y}) = 0$ 当且仅当 $\mathbf{y} = f(\mathbf{x})$.

对 $F(\mathbf{x}, f(\mathbf{x})) = 0$ 求导, 解得

$$\partial f(\mathbf{x}) = -(\partial_{\mathbf{y}} F(\mathbf{x}, f(\mathbf{x})))^{-1} \partial_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x})),$$

用数学归纳法可以证明: 当 F 是 C^k 映射时, f 也是 C^k 的。 \square

3.7 附：条件极值类型判定

考虑目标函数 $f(\mathbf{x})$ 在约束条件 $g_1(\mathbf{x}) = 0, \dots, g_m(\mathbf{x}) = 0$ 下的条件极值问题, 其中 f, g_1, \dots, g_m 都是 C^2 函数。记

$$\Sigma = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) = 0, \dots, g_m(\mathbf{x}) = 0\}.$$

设 $\mathbf{x}(t)$ 是 Σ 上的一条 C^2 曲线, 即 $\mathbf{x}(t) \in \Sigma (\forall t)$ 。记 $\mathbf{x}(0) = \mathbf{x}^*$, $\mathbf{x}'(0) = \mathbf{v} \in \mathbb{R}^n$, $\mathbf{x}''(0) = \mathbf{a} \in \mathbb{R}^n$ 。则 $\mathbf{x}(t) = \mathbf{x}^* + t\mathbf{v} + \frac{t^2}{2}\mathbf{a} + o(t^2) (t \rightarrow 0)$ 。

由 $g_j(\mathbf{x}(t)) = 0$ 以及 Taylor 展开

$$\begin{aligned} 0 &= g_j(\mathbf{x}(t)) = g_j(\mathbf{x}^*) + \left\langle \nabla g_j(\mathbf{x}^*), t\mathbf{v} + \frac{t^2}{2}\mathbf{a} \right\rangle + \frac{t^2}{2}\mathbf{v}^T H_{g_j}(\mathbf{x}^*)\mathbf{v} + o(t^2), \quad t \rightarrow 0 \\ &= t\langle \nabla g_j(\mathbf{x}^*), \mathbf{v} \rangle + \frac{t^2}{2} [\mathbf{v}^T H_{g_j}(\mathbf{x}^*)\mathbf{v} + \langle \nabla g_j(\mathbf{x}^*), \mathbf{a} \rangle] + o(t^2), \quad t \rightarrow 0 \end{aligned}$$

得到

$$\begin{aligned} \langle \nabla g_j(\mathbf{x}^*), \mathbf{v} \rangle &= 0, \\ \langle \nabla g_j(\mathbf{x}^*), \mathbf{a} \rangle + \mathbf{v}^T H_{g_j}(\mathbf{x}^*)\mathbf{v} &= 0. \end{aligned}$$

类似地,

$$f(\mathbf{x}(t)) = f(\mathbf{x}^*) + t\langle \nabla f(\mathbf{x}^*), \mathbf{v} \rangle + \frac{t^2 [\mathbf{v}^T H_f(\mathbf{x}^*)\mathbf{v} + \langle \nabla f(\mathbf{x}^*), \mathbf{a} \rangle]}{2} + o(t^2), t \rightarrow 0$$

于是, 若 f 沿 $\mathbf{x}(t)$ 在 $t = 0$ 时取得极值, 则对满足 $\langle \nabla g_j(\mathbf{x}^*), \mathbf{v} \rangle = 0$ 的任何向量 \mathbf{v} , 都成立 $\langle \nabla f(\mathbf{x}^*), \mathbf{v} \rangle = 0$ 。因此 $\nabla f(\mathbf{x}^*)$ 在由 $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$ 所张成的线性空间中, 即存在实数 $\lambda_1, \dots, \lambda_m$ 使得

$$\nabla f(\mathbf{x}^*) = \lambda_1 \nabla g_1(\mathbf{x}^*) + \dots + \lambda_m \nabla g_m(\mathbf{x}^*).$$

$(\mathbf{x}, \lambda_1, \dots, \lambda_m)$ 是方程组

$$\begin{cases} \nabla f(\mathbf{x}^*) = \lambda_1 \nabla g_1(\mathbf{x}^*) + \dots + \lambda_m \nabla g_m(\mathbf{x}^*), \\ g_1(\mathbf{x}^*) = 0, \\ \vdots \\ g_m(\mathbf{x}^*) = 0 \end{cases}$$

的解, 当且仅当 $(\mathbf{x}, \lambda_1, \dots, \lambda_m)$ 是 Lagrange 扩充函数

$$F(\mathbf{x}, \lambda_1, \dots, \lambda_m) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \dots - \lambda_m g_m(\mathbf{x})$$

的临界点。这就是 Lagrange 乘子法。

$$\begin{aligned}
\mathbf{v}^T H_f(\mathbf{x}^*) \mathbf{v} + \langle \nabla f(\mathbf{x}^*), \mathbf{a} \rangle &= \mathbf{v}^T H_f(\mathbf{x}^*) \mathbf{v} + \sum_{j=1}^m \lambda_j \langle \nabla g_j(\mathbf{x}^*), \mathbf{a} \rangle \\
&= \mathbf{v}^T H_f(\mathbf{x}^*) \mathbf{v} - \sum_{j=1}^m \lambda_j \mathbf{v}^T H_{g_j}(\mathbf{x}^*) \mathbf{v} \\
&= \mathbf{v}^T \left(H_f(\mathbf{x}^*) - \sum_{j=1}^m \lambda_j H_{g_j}(\mathbf{x}^*) \right) \mathbf{v}.
\end{aligned}$$

所以, 当 $H_f - \sum_{j=1}^m \lambda_j H_{g_j}(\mathbf{x}^*)$ 在 $T_{\mathbf{x}^*} \Sigma$ 上正定时, \mathbf{x}^* 是 $f|_{\Sigma}$ 的极小值点; 当 $H_f - \sum_{j=1}^m \lambda_j H_{g_j}(\mathbf{x}^*)$ 在 $T_{\mathbf{x}^*} \Sigma$ 上负定时, \mathbf{x}^* 是 $f|_{\Sigma}$ 的极大值点。

例 3.7.1. 求 $f(x, y, z) = xy + yz + zx$ 在约束 $xyz = 1$ 下的条件极值。

解: 令 $F(x, y, z, \lambda) = xy + yz + zx - \lambda(xyz - 1)$ 。则

$$\begin{cases} \frac{\partial F}{\partial x} = y + z - \lambda yz, \\ \frac{\partial F}{\partial y} = z + x - \lambda zx, \\ \frac{\partial F}{\partial z} = x + y - \lambda xy, \\ \frac{\partial F}{\partial \lambda} = 1 - xyz, \end{cases}$$

由 $\frac{\partial F}{\partial \lambda} = 0$ 即 $xyz = 1$ 知 $x, y, z \neq 0$ 。从而 $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = \frac{\partial F}{\partial z} = \frac{\partial F}{\partial \lambda} = 0$ 解得

$$\lambda = \frac{1}{y} + \frac{1}{z} = \frac{1}{z} + \frac{1}{x} = \frac{1}{x} + \frac{1}{y},$$

从而 $x^* = y^* = z^* = 1, \lambda^* = 2$ 。在 $(x^*, y^*, z^*) = (1, 1, 1)$ 处, 曲面 $xyz = 1$ 的切向量 $\mathbf{v} = (\xi, \eta, \zeta)^T$ 满足方程

$$0 = \partial(xyz - 1)(\mathbf{v}) = y^* z^* dx(\mathbf{v}) + z^* x^* \partial y(\mathbf{v}) + x^* y^* \partial z(\mathbf{v}) = \xi + \eta + \zeta.$$

F 关于 (x, y, z) 的 Hesse 矩阵为

$$H_F = \begin{pmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{pmatrix}_{(x^*, y^*, z^*, \lambda^*)} = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

于是对曲面 $xyz = 1$ 的任意切向量 \mathbf{v} ,

$$\begin{aligned}
\mathbf{v}^T H_F \mathbf{v} &= -2\eta\zeta - 2\zeta\xi - 2\xi\eta \\
&= -2\eta(-\xi - \eta) - 2(-\xi - \eta)\xi - 2\xi\eta \\
&= 2(\xi^2 + \xi\eta + \eta^2) > 0.
\end{aligned}$$

所以 $(x^*, y^*, z^*) = (1, 1, 1)$ 是 f 的条件极小值点, $f(1, 1, 1) = 3$ 是条件极小值。

□

3.8 习题讨论课4

1. 考虑

$$\Sigma = \{(x, y, z) | x(1 + yz) + e^{x+y+z} - 1 = 0\}.$$

- (1) 证明: Σ 是一个 C^∞ 曲面。
 (2) 求曲面 Σ 在点 $(0, 0, 0)$ 处的切平面和法线方程。
 (3) 证明: 曲面 Σ 在点 $(0, 0, 0)$ 附近确定了一个 C^∞ 的隐函数 $z = f(x, y)$, 并把这个函数在 $(0, 0)$ 点 Taylor 展开到二阶。

2. 我们已经知道

$$f(x_1, x_2) = (x_1^2 - x_2^2, 2x_1x_2)$$

是 $(0, +\infty) \times (0, +\infty)$ 到 $\mathbb{R} \times (0, +\infty)$ 的微分同胚。设 $g(y_1, y_2)$ 是 f 的逆映射。求 g 在点 $(0, 2)$ 处的二阶 Taylor 展开式。

3. (1) 证明 $\gamma = \{(x, y) | xy + e^{x+y} = 0, x > 0\}$ 是一个 C^∞ 函数 $f: (0, +\infty) \rightarrow \mathbb{R}$ 的图像, 从而是一条 C^∞ 曲线。
 (2) 求上述函数 f 的最大值。

4. 求直角坐标系中原点到曲面

$$x^2y + e^{2x} + z = 0$$

的最短距离。

5. 证明矩阵方程

$$X(t)^2 + tAX(t) = I$$

有满足 $X(0) = I$ 的 C^∞ 解, 并求 $X(t)$ 在 $t = 0$ 处的二阶 Taylor 展开。

6. 设 Σ 是 \mathbb{R}^3 中直角坐标系 (x, y, z) 下由参数方程

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \\ z = z(u, v), \end{cases} \quad (u, v) \in U \subseteq \mathbb{R}^2$$

给出的一个的二维曲面。求它的切平面和法向量。

一般地, 如果 Σ 是 \mathbb{R}^n 中直角坐标系 (x_1, \dots, x_n) 下由参数方程

$$\begin{cases} x_1 = x_1(u_1, \dots, u_{n-1}), \\ x_2 = x_2(u_1, \dots, u_{n-1}), \\ \vdots \\ x_n = x_n(u_1, \dots, u_{n-1}) \end{cases} \quad (u_1, \dots, u_{n-1}) \in U \subseteq \mathbb{R}^{n-1}$$

给出的一个的 $n - 1$ 维曲面。求它的切平面和法向量。

第4章 含参数的积分与广义积分

4.1 含参数的积分关于参数的连续性、可积性、可微性

微积分的研究对象是函数，而初等函数往往不能满足实际问题的需要。构造新函数的基本手段是运算，这不仅包含加减乘除这样的算术运算，也包括函数的复合运算，还会涉及跟求解方程相联系的隐函数、反函数。除此之外，还有分析运算，也就是取极限，这包括连续变量的极限也包括离散变量的极限(函数列极限和级数)，也包括积分这样的运算。在这一章我们集中考察由积分给出的函数，即含有参数的积分或广义积分。

设函数 $f : [a, b] \times U \rightarrow \mathbb{R}$ ($U \subseteq \mathbb{R}^n$) 满足：对任意 $\mathbf{y} \in U$, $f(\cdot, \mathbf{y}) : [a, b] \rightarrow \mathbb{R}$ Riemann 可积，此时称 $\int_a^b f(x, \mathbf{y}) dx$ 为以 \mathbf{y} 为参数的含参数的定积分。

由含参数的积分 $\int_a^b f(x, \mathbf{y}) dx$ 给出了关于 \mathbf{y} 的一个函数

$$F(\mathbf{y}) = \int_a^b f(x, \mathbf{y}) dx.$$

这提供了除初等函数以外，构造新函数的一个途径。我们这里要研究这个函数 F 的连续性、可微性和可积性，以及它的极限、导数和积分。

连续性

定理 4.1.1. 设函数 $f : [a, b] \times U \rightarrow \mathbb{R}$ ($U \subseteq \mathbb{R}^n$) 满足：

1. 对任意 $\mathbf{y} \in U$, Riemann 积分 $F(\mathbf{y}) = \int_a^b f(x, \mathbf{y}) dx$ 存在,
2. $f(x, \mathbf{y})$ 关于 \mathbf{y} 在 $\mathbf{y}_0 \in U$ 连续, 且这连续性对积分变量 $x \in [a, b]$ 一致: 即对任意 $\varepsilon > 0$, 存在 $\delta(\mathbf{y}_0, \varepsilon) > 0$, 使得对任意 $\mathbf{y} \in U$, 只要 $\|\mathbf{y} - \mathbf{y}_0\| < \delta(\mathbf{y}_0, \varepsilon)$, 就有对任意 $x \in [a, b]$, $|f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| < \varepsilon$.

则 F 在 \mathbf{y}_0 连续, 即

$$\lim_{\mathbf{y} \rightarrow \mathbf{y}_0} \int_a^b f(x, \mathbf{y}) dx = \int_a^b f(x, \mathbf{y}_0) dx = \int_a^b \lim_{\mathbf{y} \rightarrow \mathbf{y}_0} f(x, \mathbf{y}) dx.$$

证明. 当 $\|\mathbf{y} - \mathbf{y}_0\| < \delta(\mathbf{y}_0, \varepsilon)$ 时,

$$\left| \int_a^b f(x, \mathbf{y}) dx - \int_a^b f(x, \mathbf{y}_0) dx \right| \leq \int_a^b |f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| dx < \varepsilon(b-a).$$

□

注 4.1.2. 考虑区间 $[a, b]$ 上的所有 Riemann 可积函数组成的线性空间 $R[a, b]$, 对 $f \in R[a, b]$ 定义

$$\|f\| = \sup_{x \in [a, b]} |f(x)|,$$

则 $\|\cdot\|$ 是 $R[a, b]$ 上的一个范数。

$$I(f) = \int_a^b f(x) dx.$$

则 $I: R[a, b] \rightarrow \mathbb{R}$ 是一个线性函数, 满足

$$|I(f)| \leq \|f\|(b-a), \quad \forall f \in R[a, b].$$

从而 I 是 $R[a, b]$ 上的连续函数。上述定理的条件表明 $f(x, \mathbf{y})$ 给出一个映射

$$g: U \rightarrow R[a, b], \quad g(\mathbf{y}) = f(\cdot, \mathbf{y}) \in R[a, b],$$

并且 g 在 \mathbf{y}_0 连续, 于是 $F(\mathbf{y}) = I(g(\mathbf{y}))$ 在 \mathbf{y}_0 连续。

推论 4.1.3. 若 $U \subset \mathbb{R}^n$ 是闭集或者开集, $f: [a, b] \times U \rightarrow \mathbb{R}$ 连续, 则 F 在 U 上连续。

证明. 设 $\mathbf{y}_0 \in U$ 。若 U 是开集, 则取 $\delta_0 > 0$ 使得 $\text{cls}B(\mathbf{y}_0, \delta_0) \subset U$; 若 U 是闭集, 则任取 $\delta_0 > 0$ 。

总有 $K = U \cap \text{cls}B(\mathbf{y}_0, \delta_0)$ 是 U 的子集, 且为有界闭集。于是连续函数 f 在 $[a, b] \times K$ 上一致连续。从而上述定理的条件成立, 因此 F 在 \mathbf{y}_0 连续。 □

推论 4.1.4. 设 $U \subseteq \mathbb{R}^n$ 是开集, 函数 $f: [a, b] \times U \rightarrow \mathbb{R}$ 满足:

1. 对任意 $\mathbf{y} \in U$, Riemann 积分 $F(\mathbf{y}) = \int_a^b f(x, \mathbf{y}) dx$ 存在,
2. 偏导函数 $\frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_n}$ 在 $[a, b] \times U$ 上有界,

则 F 在 U 上连续。

证明. 设 $\mathbf{y} = (y_1, \dots, y_n)$, $\mathbf{y}_0 = (y_1^0, \dots, y_n^0)$. 对 $k = 1, 2, \dots, n$, 取 $\mathbf{y}_k = \mathbf{y}_{k-1} + (y_k - y_k^0)\mathbf{e}_k$, 则

$$|f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| \leq \sum_{k=1}^n |f(x, \mathbf{y}_k) - f(x, \mathbf{y}_{k-1})| \leq \sum_{k=1}^n M|y_k - y_k^0| = M\|\mathbf{y} - \mathbf{y}_0\|_1,$$

其中 $M > 0$ 是 $\left| \frac{\partial f}{\partial y_1} \right|, \dots, \left| \frac{\partial f}{\partial y_n} \right|$ 的上界. □

例 4.1.5. $f(x, y) = \cos(xy)$ 关于 (x, y) 是连续函数. 则当 $y \neq 0$ 时,

$$\int_0^1 \cos(xy) dx = \int_0^y \cos u \partial_u \left(\frac{u}{y} \right) = \frac{\sin y}{y}.$$

当 $y = 0$ 时,

$$\int_0^1 \cos(xy) dx = \int_0^1 dx = 1.$$

由上述定理知

$$\lim_{y \rightarrow 0} \int_0^1 \cos(xy) dx = \int_0^1 dx,$$

这结论与极限 $\lim_{y \rightarrow 0} \frac{\sin y}{y} = 1$ 是吻合的. □

可微性

定理 4.1.6. 设 $U \subseteq \mathbb{R}^n$ 是开集, 函数 $f: [a, b] \times U \rightarrow \mathbb{R}$ 满足:

1. 对任意 $\mathbf{y} \in U$, $f(x, \mathbf{y})$ 关于 x 在区间 $[a, b]$ 上 Riemann 可积;
2. 对任意 $x \in [a, b]$, $f(x, \mathbf{y})$ 关于 $\mathbf{y} \in U$ 是可微的;
3. 对任意 $k = 1, 2, \dots, n$, $\frac{\partial f}{\partial y_k}(x, \mathbf{y})$ 关于 (x, \mathbf{y}) 连续.

则 $F(\mathbf{y}) = \int_a^b f(x, \mathbf{y}) dx$ 关于 \mathbf{y} 是 C^1 的, 且

$$\frac{\partial F}{\partial y_k}(\mathbf{y}) = \frac{\partial}{\partial y_k} \int_a^b f(x, \mathbf{y}) dx = \int_a^b \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx.$$

证明. 任取 $\mathbf{y} \in U$, 取 $\delta_{\mathbf{y}} > 0$ 使得 $B(\mathbf{y}, 2\delta_{\mathbf{y}}) \subset U$, 取 $K = \text{cls}B(\mathbf{y}, \delta_{\mathbf{y}})$. 则 K 是有界闭集, $K \subset U$. 而 $\frac{\partial f}{\partial y_k}$ 在 $[a, b] \times K$ 上一致连续 ($k = 1, 2, \dots, n$), 所以对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对任意 $x_1, x_2 \in [a, b]$ 以及任意 $\mathbf{y}_1, \mathbf{y}_2 \in K$, 只要 $|x_1 - x_2| + \|\mathbf{y}_1 - \mathbf{y}_2\| < \delta$, 就有

$$\left| \frac{\partial f}{\partial y_k}(x_1, \mathbf{y}_1) - \frac{\partial f}{\partial y_k}(x_2, \mathbf{y}_2) \right| < \varepsilon, \quad k = 1, 2, \dots, n.$$

因此当 $|t| < \delta$ 时,

$$\begin{aligned} & \left| F(\mathbf{y} + t\mathbf{e}_k) - F(\mathbf{y}) - t \int_a^b \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx \right| \\ &= \left| \int_a^b f(x, \mathbf{y} + t\mathbf{e}_k) dx - \int_a^b f(x, \mathbf{y}) dx - t \int_a^b \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx \right| \\ &\leq \int_a^b \left| f(x, \mathbf{y} + t\mathbf{e}_k) - f(x, \mathbf{y}) - t \frac{\partial f}{\partial y_k}(x, \mathbf{y}) \right| dx \\ &= |t| \int_a^b \left| \frac{\partial f}{\partial y_k}(x, \mathbf{y} + \theta t\mathbf{e}_k) - \frac{\partial f}{\partial y_k}(x, \mathbf{y}) \right| dx \quad 0 < \theta = \theta_{x, \mathbf{y}, t, k} < 1 \\ &\leq \varepsilon(b-a)|t|, \end{aligned}$$

所以

$$\frac{\partial F}{\partial y_k}(\mathbf{y}) = \int_a^b \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx, \quad k = 1, 2, \dots, n.$$

再由连续性定理知道 F 的所有一阶偏导数是连续的, 从而 F 是 C^1 的。 \square

推论 4.1.7. 如果上述定理中, f 对 \mathbf{y} 的任何 k 阶偏导数关于 (x, \mathbf{y}) 是连续的, 则 F 是 C^k 的, 并且所有不超过 k 阶的偏导数运算与积分运算可交换。 \square

可积性

定理 4.1.8 (累次积分交换顺序). 设函数 $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ 连续, 则

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

证明. 令

$$F(t) = \int_a^b g(x, t) dx = \int_c^t \left(\int_a^b f(x, y) dx \right) dy,$$

其中 $g(x, t) = \int_c^t f(x, y) dy$ 。因为 f 连续, 所以

- $\int_a^b f(x, y) dx$ 关于 y 连续, 从而 $\int_c^t \left(\int_a^b f(x, y) dx \right) dy$ 关于 t 可导, 且

$$\frac{\partial}{\partial t} \int_c^t \left(\int_a^b f(x, y) dx \right) dy = \int_a^b f(x, t) dx.$$

- $g(x, t)$ 关于 x 连续, 关于 t 可微, 且 $\frac{\partial g}{\partial t}(x, t) = f(x, t)$ 连续, 从而 $\int_a^b g(x, t) dx$ 关于 t 可微, 且

$$\frac{\partial}{\partial t} \int_a^b g(x, t) dx = \int_a^b \frac{\partial g}{\partial t}(x, t) dx = \int_a^b f(x, t) dx.$$

因此 $F'(t) = 0$, 又 $F(c) = 0$, 所以 $F(d) = 0$ 。 \square

上述定理结论中的积分称为**累次积分**，它是特定顺序把一个多元函数按其自变量依次积分。

例 4.1.9. 设二元函数 $f(x, y)$ 满足：

1. $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ 关于 (x, y) 连续；
2. $\frac{\partial f}{\partial x}(x, y)$ 关于 x 连续，
3. 对某个 x_0 ， $\frac{\partial f}{\partial y}(x_0, y)$ 存在。

则 $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ 存在，且 $\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y)$ 。

证明. 因为 $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ 关于 (x, y) 连续，所以

$$\begin{aligned} \int_{x_0}^x \left[\int_{y_0}^y \frac{\partial^2 f}{\partial y \partial x}(u, v) dv \right] du &= \int_{y_0}^y \left[\int_{x_0}^x \frac{\partial^2 f}{\partial y \partial x}(u, v) du \right] dv, \\ \int_{y_0}^y \frac{\partial^2 f}{\partial y \partial x}(u, v) dv &= \frac{\partial f}{\partial x}(u, y) - \frac{\partial f}{\partial x}(u, y_0). \end{aligned}$$

因为 $\frac{\partial f}{\partial x}(x, y)$ 关于 x 连续，所以

$$\begin{aligned} f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0) &= \int_{x_0}^x \left[\frac{\partial f}{\partial x}(u, y) - \frac{\partial f}{\partial x}(u, y_0) \right] du \\ &= \int_{x_0}^x \left[\int_{y_0}^y \frac{\partial^2 f}{\partial y \partial x}(u, v) dv \right] du \\ &= \int_{y_0}^y \left[\int_{x_0}^x \frac{\partial^2 f}{\partial y \partial x}(u, v) du \right] dv. \end{aligned}$$

因为 $f(x_0, y)$ 关于 y 可导，所以 $f(x, y)$ 关于 y 可导，且

$$\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y) = \int_{x_0}^x \frac{\partial^2 f}{\partial y \partial x}(u, y) du.$$

从而 $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ 存在，且 $\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y)$ 。 □

例 4.1.10. 设 $0 < a < b$ ，求 $\int_0^1 \frac{x^b - x^a}{\ln x} dx$ 的值。

解: 令 $F(b) = \int_0^1 \frac{x^b - x^a}{\ln x} dx$ 。

$$\lim_{x \rightarrow 0^+} \frac{x^b - x^a}{\ln x} = 0, \quad \lim_{x \rightarrow 1^-} \frac{x^b - x^a}{\ln x} = b - a,$$

所以

$$f(x, b) = \begin{cases} \frac{x^b - x^a}{\ln x}, & 0 < x < 1, \\ b - a, & x = 1, \\ 0, & x = 0 \end{cases}$$

关于 x 在 $[0, 1]$ 上连续。

$$\frac{\partial f}{\partial b}(x, b) = x^b, \quad 0 \leq x \leq 1$$

关于 (x, b) 连续。所以

$$F'(b) = \int_0^1 \frac{\partial}{\partial b} \left(\frac{x^b - x^a}{\ln x} \right) dx = \int_0^1 x^b dx = \frac{1}{b+1},$$

所以

$$F(b) = F(a) + \int_a^b F'(t) dt = \int_a^b \frac{1}{t+1} dt = \ln(b+1) - \ln(a+1) = \ln \frac{b+1}{a+1}.$$

□

注 4.1.11. 当正数 a, b 满足 $a \neq b$ 时, $\frac{x^b - x^a}{\ln x}$ 没有初等原函数, 所以上述积分不能用 Newton-Leibniz 公式直接求得。通过对参数 b 求导, 我们得到了这一族积分之间的联系, 再借助 $a = b$ 时的特殊情况, 最终求得上述积分的结果。

例 4.1.12 (变分法).¹ 给定 \mathbb{R}^m 中两点 \mathbf{a} 和 \mathbf{b} , 对连接 \mathbf{a} 和 \mathbf{b} 的 C^∞ 曲线 $\mathbf{x}(t)$ ($0 \leq t \leq 1$, $\mathbf{x}(0) = \mathbf{a}$, $\mathbf{x}(1) = \mathbf{b}$), 其长度为

$$L = \int_0^1 \|\mathbf{x}'(t)\| dt = \int_0^1 \sqrt{\langle \mathbf{x}'(t), \mathbf{x}'(t) \rangle} dt.$$

我们希望找到长度最短的 C^∞ 曲线。

考虑 C^∞ 映射

$$\mathbf{y} : [0, 1] \times (-\delta, \delta) \rightarrow \mathbb{R}^m,$$

满足 $\mathbf{y}(t, 0) = \mathbf{x}(t)$ ($0 \leq t \leq 1$), $\mathbf{y}(0, s) = \mathbf{a}$, $\mathbf{y}(1, s) = \mathbf{b}$ 。即对每个 s , $\mathbf{y}(t, s)$ 关于参数 t 是一条连接 \mathbf{a} 和 \mathbf{b} 的 C^∞ 曲线。

于是

$$L(s) = \int_0^1 \left\| \frac{\partial \mathbf{y}}{\partial t}(t, s) \right\| dt = \int_0^1 \sqrt{\left\langle \frac{\partial \mathbf{y}}{\partial t}(t, s), \frac{\partial \mathbf{y}}{\partial t}(t, s) \right\rangle} dt$$

如果 $L(s)$ 在 $s = 0$ 时取最小值, 则 $L'(0) = 0$ 。由

$$\begin{aligned} L'(0) &= \frac{\partial}{\partial s} \int_0^1 \sqrt{\left\langle \frac{\partial \mathbf{y}}{\partial t}(t, s), \frac{\partial \mathbf{y}}{\partial t}(t, s) \right\rangle} dt \Bigg|_{s=0} \\ &= \int_0^1 \frac{\partial}{\partial s} \sqrt{\left\langle \frac{\partial \mathbf{y}}{\partial t}(t, s), \frac{\partial \mathbf{y}}{\partial t}(t, s) \right\rangle} dt \Bigg|_{s=0} \quad (\text{含参积分对参数求导}) \\ &= \int_0^1 \frac{\left\langle \frac{\partial \mathbf{y}}{\partial t}(t, 0), \frac{\partial^2 \mathbf{y}}{\partial t \partial s}(t, 0) \right\rangle}{\sqrt{\left\langle \frac{\partial \mathbf{y}}{\partial t}(t, 0), \frac{\partial \mathbf{y}}{\partial t}(t, 0) \right\rangle}} dt \\ &= \int_0^1 \left\langle \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}, \frac{\partial^2 \mathbf{y}}{\partial t \partial s}(t, 0) \right\rangle dt, \end{aligned}$$

¹关于变分法, 初学的读者可以阅读 Peter Olver 写的一份讲义 The Calculus of Variations, <https://www-users.math.umn.edu/~olver/lv-cv.pdf>

知

$$\begin{aligned} L'(0) &= \int_0^1 \left\langle \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}, \frac{\partial^2 \mathbf{y}}{\partial t \partial s}(t, 0) \right\rangle dt \\ &= \left\langle \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}, \frac{\partial \mathbf{y}}{\partial s}(t, 0) \right\rangle \Big|_{t=0}^{t=1} - \int_0^1 \left\langle \frac{\partial}{\partial t} \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}, \frac{\partial \mathbf{y}}{\partial s}(t, 0) \right\rangle dt \quad (\text{分部积分}) \\ &= - \int_0^1 \left\langle \frac{\partial}{\partial t} \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}, \frac{\partial \mathbf{y}}{\partial s}(t, 0) \right\rangle dt \quad (\text{因为 } \frac{\partial \mathbf{y}}{\partial s}(0, 0) = \frac{\partial \mathbf{y}}{\partial s}(1, 0) = \mathbf{0}). \end{aligned}$$

任取 $\varepsilon \in (0, 1)$, 我们取

$$\mathbf{y}(t, s) = \mathbf{x}(t) + st(1-t) \frac{\partial}{\partial t} \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|},$$

则

$$0 = L'(0) = - \int_0^1 t(1-t) \left\| \frac{\partial}{\partial t} \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|} \right\|^2 dt,$$

所以对任意 $t \in [0, 1]$ 都成立:

$$\frac{\partial}{\partial t} \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|} = \mathbf{0},$$

因此存在常值单位向量 \mathbf{c} 使得 $\mathbf{x}'(t) = \|\mathbf{x}'(t)\| \mathbf{c}$, 因此

$$\mathbf{x}(t) = \mathbf{a} + \mathbf{c} \int_0^t \|\mathbf{x}'(s)\| ds$$

是直线段。 □

你觉得这个证明对吗?

习题4.1

1. 对函数序列 $f_n : [a, b] \rightarrow \mathbb{R}$, 提出一个适当的条件使得

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow +\infty} f_n(x) dx$$

成立, 证明你的结论。

2. 考虑 $f(x, y) = \frac{2xy^2}{(x^2+y^2)^2}$ 。

(a) 求 $\int_0^1 f(x, y) dx$;

(b) 求 $\lim_{y \rightarrow 0} \int_0^1 f(x, y) dx$ 并讨论这个收敛对 x 的一致性;

(c) 证明 $\lim_{y \rightarrow 0} \int_0^1 f(x, y) dx \neq \int_0^1 \lim_{y \rightarrow 0} f(x, y) dx$ 。

3. 证明球面上连接两点的最短曲线是一段大圆弧。

4. 设 $L: \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ 是一个 C^∞ 函数, $L(t, q, p)$.

(1) 证明: 在所有连接点 \mathbf{a} 和点 \mathbf{b} 的 C^∞ 曲线中, 使得以下积分

$$\int_{\alpha}^{\beta} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$$

取得最小值的曲线 $\mathbf{x}(t)$ ($\alpha \leq t \leq \beta$, $\mathbf{x}(\alpha) = \mathbf{a}$, $\mathbf{x}(\beta) = \mathbf{b}$) 必然满足

$$\frac{\partial L}{\partial q}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) - \frac{\partial}{\partial t} \frac{\partial L}{\partial p}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0.$$

最后这个等式称为上述变分问题的 Euler-Lagrange 方程。

(2) 对 $L(t, q, p) = \sqrt{\langle p, p \rangle}$, 写出相应的 Euler-Lagrange 方程。

4.2 含参数的广义积分

对函数 $f: [a, +\infty) \times U \rightarrow \mathbb{R}$ ($U \subseteq \mathbb{R}^n$), 称 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 是一个以 \mathbf{y} 为参数的含参数的广义积分。

收敛的含参广义积分定义了以 \mathbf{y} 为自变量的函数

$$F(\mathbf{y}) = \int_a^{+\infty} f(x, \mathbf{y}) dx.$$

我们希望了解这个函数的连续性、可微性和可积性, 为此我们需要引进“一致收敛”的概念。

一致收敛的定义

定义 4.2.1. 称 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 关于 \mathbf{y} 在 U 上一致收敛, 若存在 $F: U \rightarrow \mathbb{R}$ 使得: 对任意 $\varepsilon > 0$, 存在 $A(\varepsilon) > a$ 使得: 当 $A > A(\varepsilon)$ 时, 对任意 $\mathbf{y} \in U$, 都有

$$\left| \int_a^A f(x, \mathbf{y}) dx - F(\mathbf{y}) \right| < \varepsilon.$$

一致收敛的含参广义积分的性质

(1) 连续性, 积分与函数极限交换次序

定理 4.2.2. 设函数 $f: [a, +\infty) \times U \rightarrow \mathbb{R}$ ($U \subseteq \mathbb{R}^n$) 满足:

1. 广义积分 $F(\mathbf{y}) = \int_a^{+\infty} f(x, \mathbf{y}) dx$ 在 U 上一致收敛,
2. 对于任何 $x \geq a$, $f(x, \mathbf{y})$ 关于 \mathbf{y} 在 $\mathbf{y}_0 \in U$ 连续, 且这连续性对 x 在任何有界闭区间 $[a, b]$ 上一致: 即对任意 $\varepsilon > 0$, 存在 $\delta(\mathbf{y}_0, \varepsilon) > 0$, 使得对任意 $\mathbf{y} \in U$, 只要 $\|\mathbf{y} - \mathbf{y}_0\| < \delta(\mathbf{y}_0, \varepsilon)$, 就有对任意 $x \in [a, b]$, $|f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| < \varepsilon$.

则 F 在 \mathbf{y}_0 连续, 即

$$\lim_{\mathbf{y} \rightarrow \mathbf{y}_0} \int_a^{+\infty} f(x, \mathbf{y}) dx = \int_a^{+\infty} f(x, \mathbf{y}_0) dx = \int_a^{+\infty} \lim_{\mathbf{y} \rightarrow \mathbf{y}_0} f(x, \mathbf{y}) dx.$$

证明. 对任意 $\varepsilon > 0$, 取 $A(\varepsilon) > a$ 使得: 对任意 $\mathbf{y} \in U$, 都有

$$\left| \int_a^{A(\varepsilon)} f(x, \mathbf{y}) dx - F(\mathbf{y}) \right| < \varepsilon.$$

对区间 $[a, A(\varepsilon)]$, 存在 $\delta(\mathbf{y}_0, \varepsilon) > 0$ 使得对任意 $\mathbf{y} \in U$, 只要 $\|\mathbf{y} - \mathbf{y}_0\| < \delta(\mathbf{y}_0, \varepsilon)$, 就有对任意 $x \in [a, A(\varepsilon)]$,

$$|f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| < \frac{\varepsilon}{A(\varepsilon) - a}.$$

于是

$$\begin{aligned} |F(\mathbf{y}) - F(\mathbf{y}_0)| &\leq \left| F(\mathbf{y}) - \int_a^{A(\varepsilon)} f(x, \mathbf{y}) dx \right| + \left| F(\mathbf{y}_0) - \int_a^{A(\varepsilon)} f(x, \mathbf{y}_0) dx \right| \\ &\quad + \left| \int_a^{A(\varepsilon)} f(x, \mathbf{y}) dx - \int_a^{A(\varepsilon)} f(x, \mathbf{y}_0) dx \right| \\ &\leq 2\varepsilon + \int_a^{A(\varepsilon)} |f(x, \mathbf{y}) - f(x, \mathbf{y}_0)| dx \\ &< 2\varepsilon + \int_a^{A(\varepsilon)} \frac{\varepsilon}{A(\varepsilon) - a} dx = 3\varepsilon. \end{aligned}$$

所以 F 在 \mathbf{y}_0 连续。 □

(2) 可微性, 积分与求导交换次序

定理 4.2.3. 设 $U \subseteq \mathbb{R}^n$ 是开集, $\mathbf{y}_0 \in U$, 函数 $f: [a, +\infty) \times U \rightarrow \mathbb{R}$ 满足:

1. 广义积分 $\int_a^{+\infty} f(x, \mathbf{y}_0) dx$ 收敛,
2. 对每个 $k = 1, 2, \dots, n$, $\frac{\partial f}{\partial y_k}(x, \mathbf{y})$ 连续, 且广义积分 $\int_a^{+\infty} \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx$ 关于 \mathbf{y} 在 U 中一致收敛。

则 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 在 \mathbf{y}_0 的一个邻域中一致收敛, 关于 \mathbf{y} 是 C^1 的, 且

$$\frac{\partial}{\partial y_k} \int_a^{+\infty} f(x, \mathbf{y}) dx = \int_a^{+\infty} \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx, \quad k = 1, 2, \dots, n.$$

证明. 记

$$\begin{aligned} F_A(\mathbf{y}) &= \int_a^A f(x, \mathbf{y}) dx, & F(\mathbf{y}) &= \int_a^{+\infty} f(x, \mathbf{y}) dx, \\ G_{k,A}(\mathbf{y}) &= \int_a^A \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx, & G_k(\mathbf{y}) &= \int_a^{+\infty} \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx. \end{aligned}$$

对任意 $\varepsilon > 0$, 存在 $A_1(\varepsilon) > a$ 使得对任意 $k = 1, 2, \dots, n$,

$$|G_{k,A}(\mathbf{y}) - G_k(\mathbf{y})| < \varepsilon, \quad \forall \mathbf{y} \in U, \quad \forall A > A_1(\varepsilon).$$

另外, 对任意 $\mathbf{y}_0 \in U$, 存在 $A(\varepsilon) > A_1(\varepsilon)$ 使得

$$|F_A(\mathbf{y}_0) - F(\mathbf{y}_0)| < \varepsilon, \quad \forall A > A(\varepsilon).$$

取 $\delta_0 > 0$ 使得 $B(\mathbf{y}_0, \delta_0) \subseteq U$, 则对任意 $k = 1, 2, \dots, n$ 以及任意 $|t| < \delta_0$,

$$\begin{aligned} & \left| F_A(\mathbf{y}_0 + t\mathbf{e}_k) - F(\mathbf{y}_0) - \int_0^t G_k(\mathbf{y}_0 + s\mathbf{e}_k) ds \right| \\ & \leq \left| F_A(\mathbf{y}_0 + t\mathbf{e}_k) - F_A(\mathbf{y}_0) - \int_0^t G_k(\mathbf{y}_0 + s\mathbf{e}_k) ds \right| + |F_A(\mathbf{y}_0) - F(\mathbf{y}_0)| \\ & \leq \left| \int_0^t \frac{\partial F_A(\mathbf{y}_0 + s\mathbf{e}_k)}{\partial s} - G_k(\mathbf{y}_0 + s\mathbf{e}_k) ds \right| + \varepsilon \\ & \leq \int_0^t |G_{k,A}(\mathbf{y}_0 + s\mathbf{e}_k) - G_k(\mathbf{y}_0 + s\mathbf{e}_k)| ds + \varepsilon \\ & \leq \varepsilon(t+1) \leq \varepsilon(\delta_0 + 1), \end{aligned}$$

所以 $\lim_{A \rightarrow +\infty} F_A(\mathbf{y}_0 + t\mathbf{e}_k)$ 存在, 且关于 $|t| \leq \delta_0$ 一致收敛, 并且

$$F(\mathbf{y}_0 + t\mathbf{e}_k) = F(\mathbf{y}_0) + \int_0^t G_k(\mathbf{y}_0 + s\mathbf{e}_k) ds.$$

由此可知 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 在 \mathbf{y}_0 的一个邻域中关于 \mathbf{y} 一致收敛, 且对任意 k ,

$$\frac{\partial}{\partial t} \int_a^{+\infty} f(x, \mathbf{y}_0 + t\mathbf{e}_k) dx = \int_a^{+\infty} \frac{\partial f}{\partial y_k}(x, \mathbf{y}_0 + t\mathbf{e}_k) dx,$$

从而

$$\frac{\partial}{\partial y_k} \int_a^{+\infty} f(x, \mathbf{y}) dx = \int_a^{+\infty} \frac{\partial f}{\partial y_k}(x, \mathbf{y}) dx.$$

因此 $F(\mathbf{y})$ 的所有一阶偏导数都连续, 从而 F 是 C^1 函数。 \square

(3) 可积性, 累次积分交换顺序

定理 4.2.4. 设 $f : [c, +\infty) \times [a, b] \rightarrow \mathbb{R}$ 连续, 广义积分 $\int_c^{+\infty} f(x, y) dx$ 关于 y 在 $[a, b]$ 上一致收敛。则广义积分 $\int_c^{+\infty} \int_a^b f(x, y) dy dx$ 收敛, 且

$$\int_a^b \int_c^{+\infty} f(x, y) dx dy = \int_c^{+\infty} \int_a^b f(x, y) dy dx.$$

证明. 对任意 $\varepsilon > 0$, 存在 $N(\varepsilon) > c$ 使得对任意 $D > C > N(\varepsilon)$ 以及任意 $y \in [a, b]$,

$$\left| \int_C^D f(x, y) dx \right| < \varepsilon,$$

从而

$$\left| \int_C^D \int_a^b f(x, y) dy dx \right| = \left| \int_a^b \int_C^D f(x, y) dx dy \right| \leq \int_a^b \varepsilon dy = (b-a)\varepsilon,$$

因此 $\int_c^{+\infty} \int_a^b f(x, y) dy dx$ 收敛。

在等式

$$\int_c^{1/t} \int_a^b f(x, y) dy dx = \int_a^b \int_c^{1/t} f(x, y) dx dy$$

两边让 $t \rightarrow 0^+$, 左边极限为 $\int_c^{+\infty} \int_a^b f(x, y) dy dx$ 。

因为 $\int_c^{+\infty} f(x, y) dx$ 关于 y 在 $[a, b]$ 上一致收敛, 所以函数

$$g(y, t) = \begin{cases} \int_c^{1/t} f(x, y) dx, & t > 0; \\ \int_c^{+\infty} f(x, y) dx, & t = 0 \end{cases}$$

关于 y 连续, 且关于 t 连续, 且在 $t = 0$ 处关于 t 的连续性对 $y \in [a, b]$ 一致。所以

$$\lim_{t \rightarrow 0^+} \int_a^b \int_c^{1/t} f(x, y) dx dy = \int_a^b \lim_{t \rightarrow 0^+} \int_c^{1/t} f(x, y) dx dy.$$

因此

$$\int_c^{+\infty} \int_a^b f(x, y) dy dx = \int_a^b \int_c^{+\infty} f(x, y) dx dy.$$

□

定理 4.2.5. 设 $f: [c, +\infty) \times [a, +\infty) \rightarrow \mathbb{R}$ 连续, 满足

1. 广义积分 $\int_c^{+\infty} f(x, y) dx$ 关于 y 在任何有界闭区间 $[a, b]$ 上一致收敛。
2. 广义积分 $\int_a^{+\infty} f(x, y) dy$ 关于 x 在任何有界闭区间 $[c, d]$ 上一致收敛。
3. $\int_a^{+\infty} \int_c^{+\infty} f(x, y) dx dy$ 和 $\int_c^{+\infty} \int_a^{+\infty} f(x, y) dy dx$ 中至少有一个存在。

则

$$\int_a^{+\infty} \int_c^{+\infty} f(x, y) dx dy = \int_c^{+\infty} \int_a^{+\infty} f(x, y) dy dx.$$

证明. 设 $\int_a^{+\infty} \int_c^{+\infty} f(x, y) dx dy$ 存在, 即极限

$$\lim_{b \rightarrow +\infty} \int_a^b \int_c^{+\infty} f(x, y) dx dy$$

存在。

因为广义积分 $\int_c^{+\infty} f(x, y) dx$ 关于 y 在任何有界闭区间 $[a, b]$ 上一致收敛, 所以

$$\int_a^b \int_c^{+\infty} f(x, y) dx dy = \int_c^{+\infty} \int_a^b f(x, y) dy dx.$$

考虑

$$g(t, x) = \begin{cases} \int_a^{\frac{1}{t}} f(x, y) dy, & t > 0, \\ \int_a^{+\infty} f(x, y) dy, & t = 0. \end{cases}$$

则由条件2知 $g(t, x)$ 连续, 且对 t 在 $t = 0$ 处连续, 并且这连续性对 $x \in [c, d]$ 上一致。于是

$$\lim_{t \rightarrow 0^+} \int_c^{+\infty} g(t, x) dx = \int_c^{+\infty} g(0, x) dx,$$

所以

$$\begin{aligned} \int_c^{+\infty} \int_a^{+\infty} f(x, y) dy dx &= \lim_{t \rightarrow 0^+} \int_c^{+\infty} \int_a^{\frac{1}{t}} f(x, y) dy dx \\ &= \lim_{t \rightarrow 0^+} \int_a^{\frac{1}{t}} \int_c^{+\infty} f(x, y) dx dy \\ &= \int_a^{+\infty} \int_c^{+\infty} f(x, y) dx dy. \end{aligned}$$

□

一致收敛的判别: 一致 Cauchy 准则

定理 4.2.6. 设函数 $f(x, \mathbf{y})$ 关于 x 在任何有界区间 $I \subset [a, +\infty)$ 上 Riemann 可积。则广义积分 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 关于 $\mathbf{y} \in U$ 一致收敛当且仅当它满足一致 Cauchy 条件: 对任意 $\varepsilon > 0$, 存在 $A(\varepsilon) > a$ 使得

$$\left| \int_{A_1}^{A_2} f(x, \mathbf{y}) dx \right| < \varepsilon, \quad \forall A_2 > A_1 \geq A(\varepsilon), \quad \forall \mathbf{y} \in U.$$

证明. 一致收敛 \Rightarrow 一致 Cauchy 是显然的。下面证明: 一致 Cauchy \Rightarrow 一致收敛。令 $F_n(\mathbf{y}) = \int_a^{a+n} f(x, \mathbf{y}) dx$ 。则对任意 $\mathbf{y} \in U$, $\{F_n(\mathbf{y})\}_{n=1}^{\infty}$ 是 Cauchy 数列。于是存在极限 $\lim_{n \rightarrow +\infty} F_n(\mathbf{y}) = F(\mathbf{y})$ 。

对任意 $\varepsilon > 0$, 存在 $A(\varepsilon) > a$ 使得

$$\left| \int_{A_1}^{A_2} f(x, \mathbf{y}) dx \right| < \varepsilon, \quad \forall A_2 > A_1 \geq A(\varepsilon), \quad \forall \mathbf{y} \in U.$$

另外, 存在 $N = N(\varepsilon, \mathbf{y}) > A_1 - a$ 使得

$$|F_N(\mathbf{y}) - F(\mathbf{y})| < \varepsilon.$$

于是

$$\left| F(\mathbf{y}) - \int_a^{A_1} f(x, \mathbf{y}) dx \right| \leq |F(\mathbf{y}) - F_N(\mathbf{y})| + \left| \int_{A_1}^{N+a} f(x, \mathbf{y}) dx \right| < 2\varepsilon,$$

所以广义积分 $\int_a^{+\infty} f(x, \mathbf{y}) dx$ 关于 $\mathbf{y} \in U$ 一致收敛。 □

定理 4.2.7. 设函数 $f(x, \mathbf{y}), g(x)$ 满足:

$$|f(x, \mathbf{y})| \leq g(x), \quad \forall x \geq a, \forall \mathbf{y} \in U,$$

且 $\int_a^{+\infty} g(x)dx$ 收敛, 则 $\int_a^{+\infty} f(x, \mathbf{y})dx$ 关于 $\mathbf{y} \in U$ 一致(绝对)收敛。

证明.

$$\left| \int_A^B f(x, \mathbf{y})dx \right| \leq \int_A^B |f(x, \mathbf{y})| dx \leq \int_A^B g(x)dx.$$

□

例 4.2.8 (Γ 函数).

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

关于 $\alpha > 0$ 是 C^∞ 函数。

证明. 任取正整数 k 。

任取 $0 < \delta < 1$, 函数 $x^\delta (\ln x)^k$ 在区间 $(0, 1)$ 上有界, 取 $M_1 > 0$ 使得对任意 $x \in (0, 1)$ 都有 $|x^\delta (\ln x)^k| \leq M_1$ 。

对任意 $0 < x < 1$, 对任意 $\alpha \in [2\delta, +\infty)$,

$$\left| \frac{\partial^k (x^{\alpha-1} e^{-x})}{\partial \alpha^k} \right| = |x^{\alpha-1} (\ln x)^k e^{-x}| \leq x^{\delta-1} |x^\delta (\ln x)^k| e^{-x} \leq M_1 x^{\delta-1} e^{-x},$$

因为 $\int_0^1 x^{\delta-1} e^{-x} dx$ 收敛, 所以 $\int_0^1 \frac{\partial^k (x^{\alpha-1} e^{-x})}{\partial \alpha^k} dx$ 关于 α 在 $[2\delta, +\infty)$ 上一致收敛。

函数 $\frac{(\ln x)^k}{x}$ 在区间 $(1, +\infty)$ 上有界, 取 $M_2 > 0$ 使得对任意 $x \in (1, +\infty)$ 都有 $\left| \frac{(\ln x)^k}{x} \right| \leq M_2$ 。对任意 $x > 1$, 对任意 $\alpha \in (0, \frac{2}{\delta}]$,

$$\left| \frac{\partial^k (x^{\alpha-1} e^{-x})}{\partial \alpha^k} \right| = |x^{\alpha-1} (\ln x)^k e^{-x}| \leq M_2 x^\alpha e^{-x} \leq M_2 x^{\frac{2}{\delta}} e^{-x},$$

因为 $\int_1^{+\infty} x^{\frac{2}{\delta}} e^{-x} dx$ 收敛, 所以 $\int_1^{+\infty} \frac{\partial^k (x^{\alpha-1} e^{-x})}{\partial \alpha^k} dx$ 关于 α 在 $(0, \frac{2}{\delta}]$ 上一致收敛。

因此 $\int_0^{+\infty} \frac{\partial^k (x^{\alpha-1} e^{-x})}{\partial \alpha^k} dx$ 关于 α 在 $[2\delta, \frac{2}{\delta}]$ 上一致收敛。所以 Γ 函数在 $[2\delta, \frac{2}{\delta}]$ 上是 C^k 函数, 又因为 $\delta \in (0, 1)$ 是任意的, 所以 Γ 函数在 $(0, +\infty)$ 上是 C^k 函数。再由 k 的任意性知 Γ 函数在 $(0, +\infty)$ 上是 C^∞ 函数。□

例 4.2.9 (Beta 函数).

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx, \quad \alpha > 0, \beta > 0.$$

是 C^∞ 函数, 满足

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

证明. 因为

$$x^{\alpha-1}(1-x)^{\beta-1} = x^{\alpha-1} + o(x^{\alpha-1}), \quad x \rightarrow 0,$$

所以 $\int_0^{1/2} x^{\alpha-1}(1-x)^{\beta-1} dx$ 收敛当且仅当 $\int_0^{1/2} x^{\alpha-1} dx$ 收敛, 后者当且仅当 $\alpha > 0$. 同理可证 $\int_0^{1/2} x^{\alpha-1}(1-x)^{\beta-1} dx$ 收敛当且仅当 $\int_{1/2}^1 x^{\alpha-1}(1-x)^{\beta-1} dx$ 收敛当且仅当 $\beta > 0$.

取 $x = \frac{1}{1+y}$ ($y \in (0, +\infty)$), 则

$$B(\alpha, \beta) = \int_{+\infty}^0 \left(\frac{1}{1+y}\right)^{\alpha-1} \left(\frac{y}{1+y}\right)^{\beta-1} d\frac{1}{1+y} = \int_0^{+\infty} \frac{y^{\beta-1}}{(1+y)^{\alpha+\beta}} dy.$$

注意到对 $y > 0$,

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx = \int_0^{+\infty} (xy)^{\alpha-1} e^{-xy} d(xy) = y^\alpha \int_0^{+\infty} x^{\alpha-1} e^{-xy} dx,$$

所以

$$\Gamma(\alpha + \beta) = (1+y)^{\alpha+\beta} \int_0^{+\infty} x^{\alpha+\beta-1} e^{-x(1+y)} dx,$$

从而

$$\begin{aligned} \Gamma(\alpha + \beta)B(\alpha, \beta) &= \int_0^{+\infty} \frac{\Gamma(\alpha + \beta)y^{\beta-1}}{(1+y)^{\alpha+\beta}} dy \\ &= \int_0^{+\infty} \int_0^{+\infty} y^{\beta-1} x^{\alpha+\beta-1} e^{-x(1+y)} dx dy \\ &= \int_0^{+\infty} \int_0^{+\infty} y^{\beta-1} x^{\alpha+\beta-1} e^{-x(1+y)} dy dx \\ &= \int_0^{+\infty} x^{\alpha-1} e^{-x} \int_0^{+\infty} (xy)^{\beta-1} e^{-xy} d(xy) dx \\ &= \Gamma(\alpha)\Gamma(\beta). \end{aligned}$$

因此

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

所以 $B(\alpha, \beta)$ 是 C^∞ 函数. □

例 4.2.10 (Euler-Poisson积分). 求广义积分 $\int_0^{+\infty} e^{-x^2} dx$ 以及 $\Gamma\left(\frac{1}{2}\right)$ 的值.

解: 对 $y > 0$,

$$I = \int_0^{+\infty} e^{-x^2} dx = y \int_0^{+\infty} e^{-(xy)^2} dx,$$

从而

$$\begin{aligned} I^2 &= I \int_0^{+\infty} e^{-y^2} dy = \int_0^{+\infty} e^{-y^2} \left(y \int_0^{+\infty} e^{-(xy)^2} dx \right) dy \\ &= \int_0^{+\infty} \left(\int_0^{+\infty} ye^{-y^2} e^{-(xy)^2} dy \right) dx \\ &= \frac{1}{2} \int_0^{+\infty} \left(\int_0^{+\infty} e^{-(1+x^2)u} du \right) dx \quad (u = y^2) \\ &= \frac{1}{2} \int_0^{+\infty} \frac{1}{1+x^2} dx = \frac{\arctan x|_0^{+\infty}}{2} = \frac{\pi}{4}, \end{aligned}$$

因此 $I = \frac{\sqrt{\pi}}{2}$ 。

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} x^{-\frac{1}{2}} e^{-x} dx = \int_0^{+\infty} \frac{1}{u} e^{-u^2} du^2 = 2 \frac{\sqrt{\pi}}{2} = \sqrt{\pi}.$$

□

例 4.2.11. 求 $D(\alpha) = \int_0^{+\infty} \frac{\sin \alpha x}{x} dx$ 。

解: 易见 $D(-\alpha) = -D(\alpha)$, 另外对 $\alpha > 0$,

$$D(\alpha) = \int_0^{+\infty} \frac{\sin \alpha x}{x} dx = \int_0^{+\infty} \frac{\sin t}{t} dt = D(1), \quad (t = \alpha x).$$

为求 $D(1)$ 的值, 我们考虑

$$g(y) = \int_0^{+\infty} e^{-yx} \frac{\sin x}{x} dx, \quad y \geq 0.$$

任取 $\delta > 0$ 。

当 $y \geq \delta$ 时, $\int_0^{+\infty} e^{-\delta x} dx$ 收敛且

$$\begin{aligned} \left| e^{-yx} \frac{\sin x}{x} \right| &\leq e^{-\delta x}, \quad \forall x > 0, \\ \left| x e^{-yx} \frac{\sin x}{x} \right| &\leq e^{-yx} \leq e^{-\delta x}, \quad \forall x > 0, \end{aligned}$$

所以广义积分 $\int_0^{+\infty} e^{-yx} \frac{\sin x}{x} dx$ 和 $\int_0^{+\infty} x e^{-yx} \frac{\sin x}{x} dx$ 对 $y \in [\delta, +\infty)$ 一致收敛。

从而 $g(y)$ 在 $(0, +\infty)$ 上连续、可微, 且

$$g'(y) = - \int_0^{+\infty} e^{-yx} \sin x dx = -1 + y^2 \int_0^{+\infty} e^{-yx} \sin x dx = -1 - y^2 g'(y).$$

解得 $g(y) = -\arctan y + C$ 。

因为

$$-\frac{\pi}{2} + C = \lim_{y \rightarrow +\infty} g(y) = \int_0^{+\infty} \lim_{y \rightarrow +\infty} e^{-yx} \frac{\sin x}{x} dx = 0,$$

所以 $C = \frac{\pi}{2}$ 。最后我们证明 $g(y)$ 在 $y = 0$ 连续, 从而 $g(0) = C = \frac{\pi}{2}$, 即

$$\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}, \quad (\text{Dirichlet 积分})$$

从而 $\int_0^{+\infty} \frac{\sin \alpha x}{x} dx = \frac{\pi}{2} \operatorname{sgn} \alpha$ 。对任意 $\varepsilon > 0$, 对任意 $B > A > \frac{8}{\varepsilon}$ 以及任意 $y \in [0, \frac{1}{2}]$,

$$\begin{aligned} \left| \int_A^B e^{-xy} \frac{\sin x}{x} dx \right| &= \left| \int_A^B e^{-xy} \frac{1}{x} d \cos x \right| \\ &\leq \left| e^{-xy} \frac{\cos x}{x} \cos x \right|_A^B + \left| \int_A^B \cos x d \frac{e^{-xy}}{x} \right| \\ &\leq \frac{1}{A} + \frac{1}{B} + \left| \int_A^B \frac{-xy e^{-xy} - e^{-xy}}{x^2} \cos x dx \right| \\ &\leq \frac{2}{A} + y \left| \int_A^B \frac{e^{-xy}}{x} \cos x dx \right| + \int_A^B \frac{1}{x^2} dx \\ &\leq \frac{3}{A} + y \left| \int_A^B \frac{e^{-xy}}{x} \cos x dx \right| \\ &\leq \frac{3}{A} + \frac{3y}{A} + y^2 \left| \int_A^B e^{-xy} \frac{\sin x}{x} dx \right| \\ &\leq \frac{6}{A} + \frac{1}{4} \left| \int_A^B e^{-xy} \frac{\sin x}{x} dx \right|, \end{aligned}$$

所以

$$\left| \int_A^B e^{-xy} \frac{\sin x}{x} dx \right| \leq \frac{1}{1 - \frac{1}{4}} \frac{6}{A} = \frac{8}{A} < \varepsilon,$$

所以 $g(y)$ 在 $[0, \frac{1}{2}]$ 上一致收敛, 从而 $g(y)$ 在 $[0, \frac{1}{2}]$ 上连续。因此从而 $g(0) = \frac{\pi}{2}$ 。□

对于条件收敛的含参数广义积分的一致收敛性, 有以下判别法。

定理 4.2.12. 含参数的广义积分 $\int_0^{+\infty} f(x, y)g(x, y)dx$ 关于参数 $y \in U$ 一致收敛, 如果 f, g 满足如下两个条件中的一个:

1. (Dirichlet)

- (a) $F(t, y) = \int_a^t f(x, y)dx$ 对所有 $t \in [a, +\infty)$ 和 $y \in U$ 一致有界;
- (b) 对所有 $y \in U$, $g(x, y)$ 关于 x 单调减;
- (c) $\lim_{x \rightarrow +\infty} g(x, y) = 0$ 且关于 $y \in U$ 一致。

2. (Abel)

- (a) 广义积分 $\int_a^{+\infty} f(x, y)dx$ 关于参数 $y \in U$ 一致收敛;

(b) 对所有 $y \in U$, $g(x, y)$ 关于 x 单调减;

(c) $g(x, y)$ 有界, 且关于 $y \in U$ 一致。

4.3 应用: 积分变换

我们说的积分变换, 是指形如

$$T: f \mapsto T(f) = g, \quad g(y) = \int_{-\infty}^{+\infty} K(x, y)f(x)dx$$

的变换, 它把一个函数 f 变换为另一个函数 g , 其中 $K(x, y)$ 称为这个积分变换的积分核。它对于的离散化形式为

$$b_n = \sum_m K_{n,m} a_m,$$

它把数列 $\{a_n\}$ 变成数列 $\{b_n\}$ 。它们可以认为是由矩阵确定的线性变换 $\mathbf{b} = \mathbf{K}\mathbf{a}$ 的推广。

卷积

形如

$$(K * f)(y) = \int_{-\infty}^{+\infty} K(y-x)f(x)dx,$$

称 $K * f$ 是 K 与 f 的卷积。

如果 $K(t)$ 是可微函数, 则(只要下述广义积分一致收敛, 就有)

$$(K * f)'(y) = \int_{-\infty}^{+\infty} K'(y-x)f(x)dx = (K' * f)(x),$$

例 4.3.1 (Laplace 方程的边值问题). 考虑 $K(x, y) = K_y(x) = \frac{y}{\pi(x^2+y^2)}$, 不难验证

$$\begin{aligned} \int_{-\infty}^{+\infty} K(x, y)dx &= 1, \\ \frac{\partial K}{\partial x}(x, y) &= -2\pi K(x, y)K(y, x) = \frac{\partial K}{\partial x}(y, x), \\ \frac{\partial K}{\partial y}(x, y) &= \pi [K(y, x)^2 - K(x, y)^2] = -\frac{\partial K}{\partial y}(y, x), \\ \frac{\partial^2 K}{\partial x^2}(x, y) &= 2\pi \left[-K(y, x)\frac{\partial K}{\partial x}(x, y) - K(x, y)\frac{\partial K}{\partial y}(y, x) \right], \\ \frac{\partial^2 K}{\partial y^2}(x, y) &= 2\pi \left[K(y, x)\frac{\partial K}{\partial x}(y, x) - K(x, y)\frac{\partial K}{\partial y}(x, y) \right], \end{aligned}$$

从而 $K, \frac{\partial K}{\partial x}, \frac{\partial K}{\partial y}$ 都是有界函数, 并且

$$\Delta K = \frac{\partial^2 K}{\partial x^2} + \frac{\partial^2 K}{\partial y^2} = 0,$$

对有界函数 f , 令

$$u(x, y) = K_y * f = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{y}{(x-t)^2 + y^2} f(t) dt.$$

则

$$\Delta u = \Delta(K_y * f) = (\Delta K) * f = 0.$$

则

$$\begin{aligned} u(x, y) - f(x) &= \frac{1}{\pi} \int_{|t-x| < \delta} \frac{y}{(x-t)^2 + y^2} (f(t) - f(x)) dt \\ &\quad + \frac{1}{\pi} \int_{|t-x| \geq \delta} \frac{y}{(x-t)^2 + y^2} (f(t) - f(x)) dt. \end{aligned}$$

其中

$$\begin{aligned} \left| \int_{|t-x| \geq \delta} \frac{y}{(x-t)^2 + y^2} (f(t) - f(x)) dt \right| &\leq 2M \int_{|t-x| \geq \delta} \frac{y}{(x-t)^2 + y^2} dt \\ &= 2M \int_{|s| \geq \frac{\delta}{y}} \frac{1}{s^2 + 1} ds \quad s = \frac{t-x}{y} \\ &= 2M \left(\pi - 2 \arctan \frac{\delta}{y} \right) \rightarrow 0, \quad y \rightarrow 0^+. \end{aligned}$$

$$\left| \int_{|t-x| < \delta} \frac{y}{(x-t)^2 + y^2} (f(t) - f(x)) dt \right| \leq \varepsilon \int_{|t-x| < \delta} \frac{y}{(x-t)^2 + y^2} dt \leq \pi \varepsilon.$$

所以 $\lim_{y \rightarrow 0^+} u(x, y) = f(x)$ 。因此 $u(x, y) = K_y * f$ 是

$$\begin{cases} \Delta u = 0, & y > 0; \\ u(x, 0) = f(x) \end{cases}$$

的解。 □

例 4.3.2 (独立随机变量的和的概率分布). 独立抛掷两个骰子, 得到的点数分别是随机变量 X, Y , 则 $X + Y$ 的概率分布为

$$P(X + Y = n) = \sum_{k=1}^6 P(X = k, Y = n - k) = \sum_{k=1}^6 P(X = k)P(Y = n - k),$$

这就是离散情形的卷积。对连续的随机变量, 其概率分布由概率密度刻画, 如果 X 的概率密度是 f , Y 的概率密度是 g , 那么当它们独立时, $X + Y$ 的概率密度就是 f 和 g 的卷积

$$\int_{-\infty}^{+\infty} g(x-t)f(t)dt.$$

比如, 正态分布 $N(\mu, \sigma^2)$ 的概率密度为

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

若 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 且它们独立, 则 $X + Y$ 的概率密度为

$$\begin{aligned} & \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-t-\mu_2)^2}{2\sigma_2^2}} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(t-\mu_1)^2}{2\sigma_1^2}} dt \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_2^2}} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-t-\mu_2)^2}{2\sigma_2^2} - \frac{(t-\mu_1)^2}{2\sigma_1^2}} dt \\ &= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(\sigma_1^2+\sigma_2^2)}{2\sigma_1^2\sigma_2^2} \left[\left(t - \frac{\sigma_1^2(x-\mu_2)+\sigma_2^2\mu_1}{\sigma_1^2+\sigma_2^2} \right)^2 + \frac{\sigma_1^2\sigma_2^2(x-\mu_1-\mu_2)^2}{(\sigma_1^2+\sigma_2^2)^2} \right]} dt \\ &= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{(x-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}} e^{-\frac{(\sigma_1^2+\sigma_2^2)}{2\sigma_1^2\sigma_2^2} \left(t - \frac{\sigma_1^2(x-\mu_2)+\sigma_2^2\mu_1}{\sigma_1^2+\sigma_2^2} \right)^2} dt \\ &= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{(x-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}}. \end{aligned}$$

所以 $X + Y$ 服从正态分布 $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. □

Laplace变换

$$L[f](p) = \int_0^{+\infty} e^{-px} f(x) dx = F(p),$$

称 F 是 f 的 **Laplace 变换**。

如果 f 满足 $|f(x)| \leq Me^{\alpha x}$, 则对任意 $\delta > 0$, 上述广义积分以及广义积分

$$\int_0^{+\infty} e^{-px} x^k f(x) dx$$

对 p 在区间 $[\alpha + \delta, +\infty)$ 上一致收敛, 从而 F 在区间 $(\alpha, +\infty)$ 上是 C^∞ 函数。

除了线性以外, Laplace 变换还满足

$$L[f(\alpha x)](p) = \frac{1}{\alpha} L[f]\left(\frac{p}{\alpha}\right),$$

$$L[e^{\alpha x} f(x)](p) = L[f](p - \alpha),$$

$$\begin{aligned} L[f'](p) &= \int_0^{+\infty} e^{-px} f'(x) dx = e^{-px} f(x) \Big|_0^{+\infty} + p \int_0^{+\infty} e^{-px} f(x) dx \\ &= -f(0) + pL[f](p), \end{aligned}$$

$$\frac{\partial^k}{\partial p^k} L\left[\frac{f(x)}{x^k}\right](p) = (-1)^k L[f](p)$$

例 4.3.3.

$$\begin{aligned}
L[x^\alpha] &= \int_0^{+\infty} x^\alpha e^{-px} dx = \frac{1}{p^{\alpha+1}} \int_0^{+\infty} u^\alpha e^{-u} du = \frac{\Gamma(\alpha+1)}{p^{\alpha+1}}, \\
L[e^{\alpha x}] &= L[1](p-\alpha) = \frac{1}{p-\alpha}, \\
L[\sin \omega x] &= \int_0^{+\infty} e^{-px} \sin \omega x dx = -\frac{1}{\omega} \int_0^{+\infty} e^{-px} d \cos \omega x \\
&= \frac{1}{\omega} - \frac{p}{\omega} \int_0^{+\infty} e^{-px} \cos \omega x dx = \frac{1}{\omega} - \frac{p^2}{\omega^2} \int_0^{+\infty} e^{-px} \sin \omega x dx \\
&= \frac{1}{\omega} \frac{1}{1 + \frac{p^2}{\omega^2}} = \frac{\omega}{p^2 + \omega^2}, \\
L[\cos \omega x] &= L\left[\frac{1}{\omega} \sin' \omega x\right] = p \cdot \frac{1}{\omega} \cdot \frac{\omega}{p^2 + \omega^2} = \frac{p}{p^2 + \omega^2}, \\
L[\sinh ax] &= L\left[\frac{e^{ax} - e^{-ax}}{2}\right] = \frac{1}{2} \left(\frac{1}{p-a} - \frac{1}{p+a} \right) = \frac{a}{p^2 - a^2}, \\
L[\cosh ax] &= \frac{p}{p^2 - a^2}.
\end{aligned}$$

例 4.3.4. 求初值问题

$$\begin{cases} y'' + 4y = 4x, \\ y(0) = 1, \\ y'(0) = 5 \end{cases}$$

的解。

解:

$$L[y''] + 4L[y] = 4L[x]$$

故

$$-y'(0) + p(-y(0) + pL[y]) + 4L[y] = \frac{4}{p^2}$$

所以

$$L[y] = \frac{1}{p^2 + 4} \left(\frac{4}{p^2} + 5 + p \right) = \frac{p}{p^2 + 4} + \frac{4}{p^2 + 4} + \frac{1}{p^2} = L[\cos 2x + 2 \sin 2x + x]$$

所以

$$y = \cos 2x + 2 \sin 2x + x$$

是初值问题的解。 □

关于 Laplace 变换的更多内容和有趣应用, 建议读者阅读 George F. Simmons 的经典著作 *Differential Equations With Applications And Historical Notes*, 那里还有一篇短小而有趣的关于 Laplace 生平的介绍。

习题 4.3

1. 记 $K_t(x) = K(x, t) = \frac{1}{2\sqrt{\pi t}} e^{-\frac{x^2}{4t}}$ 。证明

$$\int_{-\infty}^{+\infty} K(x, t) dx = 1.$$

并且对有界的连续函数 $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$u(x, t) = (K_t * f)(x) = \int_{-\infty}^{+\infty} \frac{1}{2\sqrt{\pi t}} e^{-\frac{(x-\xi)^2}{4t}} f(\xi) d\xi$$

是如下热传导方程的初值问题的解

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0, \\ u(x, 0) = f(x), \quad (\text{即 } \lim_{t \rightarrow 0^+} u(x, t) = f(x).) \end{cases}$$

4.4 习题讨论课5

1. (1) 利用换元
- $\theta = 2 \arctan t$
- , 证明

$$\int_0^\pi \frac{d\theta}{1+p \cos \theta} = \frac{\pi}{\sqrt{1-p^2}}, \quad |p| < 1.$$

- (2) 证明

$$\int_0^\pi \frac{\ln(1+x \cos \theta) d\theta}{\cos \theta} = \pi \arcsin x, \quad |x| < 1.$$

- (3) 设
- $0 < \theta < 1$
- ,
- $f(x, y) = \frac{1}{1-\theta^2 y^2 \sin^2 x}$
- (
- $x \in [0, \frac{\pi}{2}], y \in [0, 1]$
-). 证明

$$\int_0^{\frac{\pi}{2}} f(x, y) dx = \frac{\pi}{2\sqrt{1-\theta^2 y^2}}, \quad \int_0^1 f(x, y) dy = \frac{1}{2\theta \sin x} \ln \frac{1+\theta \sin x}{1-\theta \sin x},$$

并且

$$\int_0^{\frac{\pi}{2}} \frac{1}{\sin x} \ln \frac{1+\theta \sin x}{1-\theta \sin x} dx = \pi \arcsin \theta.$$

2. 计算
- $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_1^2 e^y \sin\left(\frac{x}{y}\right) dy dx$
- .

3. 计算
- $\int_{-\infty}^{+\infty} x e^{-\frac{1}{2}x^2} \sin(tx) dx$
- .

4. 已知
- $\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$
- . 设
- $F(x) = \int_{-\infty}^{+\infty} \frac{\sin(xt)}{t(1+t^2)} dt$
- . 证明

(1) $F(x) - F''(x) = \frac{\pi}{2} (x > 0)$;

(2) $F(x) = \frac{\pi}{2} (1 - e^{-x})$;

(3) $\int_{-\infty}^{+\infty} \frac{\cos(xt)}{1+t^2} dt = \frac{\pi}{2} e^{-x} (x > 0)$.

5. 证明对任意
- $x \geq 0$
- ,
- $\int_{-\infty}^{+\infty} \frac{\ln(1+x^2 t^2)}{1+t^2} dt = \pi \ln(1+x)$
- .

6. 记
- $K_t(x) = K(x, t) = \frac{1}{2\sqrt{\pi t}} e^{-\frac{x^2}{4t}}$
- . 证明

$$\int_{-\infty}^{+\infty} K(x, t) dx = 1.$$

并且对有界的连续函数 $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$u(x, t) = (K_t * f)(x) = \int_{-\infty}^{+\infty} \frac{1}{2\sqrt{\pi t}} e^{-\frac{(x-\xi)^2}{4t}} f(\xi) d\xi$$

是如下热传导方程的初值问题的解

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0, \\ u(x, 0) = f(x), \quad (\text{即 } \lim_{t \rightarrow 0^+} u(x, t) = f(x).) \end{cases}$$

第5章 重积分

5.1 重积分的概念

定义 5.1.1 (矩形区域上的积分). 称 \mathbb{R}^m 的形如

$$\begin{aligned} R &= \{(x_1, x_2, \dots, x_m) \in \mathbb{R}^m \mid a_k \leq x_k \leq b_k, k = 1, 2, \dots, m\} \\ &=: [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_m, b_m] \end{aligned}$$

的子集为 \mathbb{R}^m 中的一个 m 维矩形, 记

$$\mu_m(R) := (b_1 - a_1)(b_2 - a_2) \cdots (b_m - a_m)$$

称为 R 的体积。在不导致混淆时, 我们也记 $\mu(R) = \mu_m(R)$ 。

称 $\{R_k\}_{k=1}^p$ 为矩形 R 的一个分割, 如果每个 R_k 都是 \mathbb{R}^m 中的一个矩形, 并且 $\bigcup_{k=1}^p R_k = R$, $\text{int}R_i \cap \text{int}R_j = \emptyset$ (对任意 $1 \leq i < j \leq p$)。

给定函数 $f: R \rightarrow \mathbb{R}$, 对于矩形 R 的一个分割 $\mathcal{P} = \{R_k\}_{k=1}^p$ 以及相应于这个分割的一组标志点 $\xi = \{\xi_k\}_{k=1}^p$ (对每个 k , $\xi_k \in R_k$), 记

$$S(f, \mathcal{P}, \xi) = \sum_{k=1}^p f(\xi_k) \mu(R_k)$$

称之为 f 在 R 上关于分割 \mathcal{P} 和标志点组 ξ 的 Riemann 和。

我们称 f 在 R 上 **Riemann 可积**, 如果存在 $I \in \mathbb{R}$ 使得: 对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得对 R 的任意分割 \mathcal{P} , 只要 $\|\mathcal{P}\| = \max_{1 \leq k \leq p} \left(\sup_{\mathbf{x}, \mathbf{y} \in R_k} \|\mathbf{x} - \mathbf{y}\| \right) < \delta$, 就有对相对于 \mathcal{P} 任意标志点组 ξ , 都有 $|S(f, \mathcal{P}, \xi) - I| < \varepsilon$ 。此时记

$$I = \int_R f(x) d\mu$$

称它为 f 在 R 上 **Riemann 积分**。

定义 5.1.2 (矩形区域上的积分, Darboux 的定义). 设 $f: R \rightarrow \mathbb{R}$ 矩形 R 的一个有界函数, $\mathcal{P} = \{R_k\}_{k=1}^p$ 是 R 的一个分割。记

$$S(f, \mathcal{P}) = \sum_{k=1}^p \sup_{x \in R_k} f(x) \mu(R_k), \quad s(f, \mathcal{P}) = \sum_{k=1}^p \inf_{x \in R_k} f(x) \mu(R_k),$$

分别称它们为 f 在 R 上关于分割 \mathcal{P} 的 **Darboux 上和** 与 **Darboux 下和**。

称 R 的分割 $\mathcal{Q} = \{\tilde{R}_j\}_{j=1}^q$ 是 $\mathcal{P} = \{R_k\}_{k=1}^p$ 的一个**加细**, 如果 \mathcal{P} 中每个矩形 R_k 都是 \mathcal{Q} 中某些矩形的并集。

易见: 对矩形 R 的任何两个分割 $\mathcal{P}_1, \mathcal{P}_2$, 都存在它们的一个共同加细分割 $\mathcal{P}_1 \vee \mathcal{P}_2$, 且

$$s(f, \mathcal{P}_1) \leq s(f, \mathcal{P}_1 \vee \mathcal{P}_2) \leq S(f, \mathcal{P}_1 \vee \mathcal{P}_2) \leq S(f, \mathcal{P}_2).$$

记 \mathfrak{P} 为矩形 R 的所有分割组成的集合。称

$$S(f) = \inf_{\mathcal{P} \in \mathfrak{P}} S(f, \mathcal{P}), \quad s(f) = \sup_{\mathcal{P} \in \mathfrak{P}} s(f, \mathcal{P})$$

分别为 f 在 R 上的 **Darboux 上积分** 和 **Darboux 下积分**。

如果 $S(f) = s(f)$, 则称 f 在 R 上 **Darboux 可积**, 它们共同的值记为

$$I = \int_R f(x) d\mu$$

称它为 f 在 R 上 **Darboux 积分**。

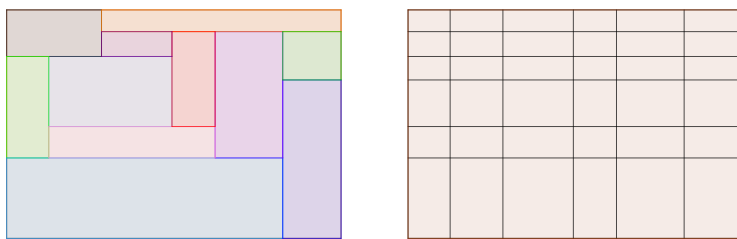


图 5.1: 矩形的分割 \mathcal{P} (左) 和它的一个加细分割 \mathcal{Q} (右)

定理 5.1.3. 以下结论等价:

1. f 在矩形 R 上 *Riemann* 可积;
2. f 在矩形 R 上有界且 *Darboux* 可积;

3. f 在矩形 R 上有界且 f 在矩形 R 中的间断点集是一个体积为零的集合。

4. 对任意 $\varepsilon > 0$, 存在为矩形 R 的一个分割 \mathcal{P} 使得 $S(f, \mathcal{P}) - s(f, \mathcal{P}) < \varepsilon$ 。

注 5.1.4. 称一个集合 $A \subset \mathbb{R}^m$ 体积为零是指: 对任意 $\varepsilon > 0$, 存在可数无穷多个矩形 R_1, R_2, \dots , 使得 $A \subset R_1 \cup R_2 \cup \dots$ 并且 R_1, R_2, \dots 的体积总和小于 ε 。

推论 5.1.5. 连续函数在任何有界闭的矩形区域上都是 *Riemann* 可积的。

例 5.1.6. 设 R 是一个矩形区域, 且存在 R 的一个稠密子集 A 使得 $R \setminus A$ 也是 R 的一个稠密子集, 则

$$f(x) = \begin{cases} 1, & x \in A; \\ 0, & x \in R \setminus A, \end{cases}$$

在 R 上不是 *Riemann* 可积函数。这是因为 f 在 R 上到处都是间断的。 \square

例 5.1.7. 设 R_1, R_2 是两个矩形区域, $R_1 \subset R_2$ 。

如果 f 在 R_2 上是 *Riemann* 可积函数, 则 f 在 R_1 上也是 *Riemann* 可积函数。

如果 f 在 R_1 上是 *Riemann* 可积函数, 则

$$\tilde{f}(x) = \begin{cases} f(x), & x \in R_1; \\ 0, & x \in R_2 \setminus R_1 \end{cases}$$

在 R_2 上也是 *Riemann* 可积函数, 且 $\int_{R_2} \tilde{f}(x) d\mu = \int_{R_1} f(x) d\mu$ 。 \square

定理 5.1.8. 设 $R \subset \mathbb{R}^m$ 是一个 m 维矩形, 记 $\mathfrak{R}(R)$ 为所有在 R 上 *Riemann* 可积函数组成的集合, $C(R)$ 为所有在 R 上连续函数组成的集合。则

1. (线性) $\mathfrak{R}(R)$ 是一个线性空间, $C(R)$ 是 $\mathfrak{R}(R)$ 的一个线性子空间, 并且对任意 $f, g \in \mathfrak{R}(R)$ 以及任意 $\alpha, \beta \in \mathbb{R}$,

$$\int_R (\alpha f(x) + \beta g(x)) d\mu = \alpha \int_R f(x) d\mu + \beta \int_R g(x) d\mu;$$

2. (保序性) 若 $f, g \in \mathfrak{R}(R)$ 满足 $f(x) \leq g(x)$ ($\forall x \in R$), 则 $\int_R f(x) d\mu \leq \int_R g(x) d\mu$;

3. $f \in \mathfrak{R}(R)$ 当且仅当 $f_+, f_- \in \mathfrak{R}(R)$, 其中

$$f_+(x) = \max\{f(x), 0\} = \frac{|f(x)| + f(x)}{2},$$

$$f_-(x) = \max\{-f(x), 0\} = \frac{|f(x)| - f(x)}{2},$$

从而 $f = f_+ - f_-$, $|f| = f_+ + f_-$.

并且 $|\int_R f(x) d\mu| \leq \int_R |f(x)| d\mu$;

4. (Cauchy-Schwarz 不等式) 对任意 $f, g \in \mathfrak{R}(R)$, $fg \in \mathfrak{R}(R)$, 且

$$\left| \int_R f(x)g(x)d\mu \right| \leq \sqrt{\int_R |f(x)|^2 d\mu} \cdot \sqrt{\int_R |g(x)|^2 d\mu};$$

5. (积分平均值定理) 若 $f \in C(R)$, $g \in \mathfrak{R}(R)$ 且 $g(x) \geq 0$ ($\forall x \in R$), 则存在 $\eta \in R$ 使得

$$\int_R f(x)g(x)d\mu = f(\eta) \int_R g(x)d\mu.$$

习题5.1

1. 设 $\{R_k\}_{k=1}^p$ 是矩形 R 的一个分割, 则 $\mu(R) = \mu(R_1) + \cdots + \mu(R_p)$ 。
2. 对任何矩形 R , $\mu(R) = \int_R 1d\mu$ 。
3. 设 $A = \{(x, y) \in [0, 1] \times [0, 1] | x, y \in \mathbb{Q}\}$ 。证明 A 在 $R = [0, 1] \times [0, 1]$ 中稠密, 但 A 是面积(2维体积)为零的集合。
4. 证明 \mathbb{R}^m 的任何低维线性子空间都是 m 维体积为零的集合。

5.2 Jordan 可测集和 Riemann 积分

定义 5.2.1. 设 $D \subset \mathbb{R}^m$ 是非空的有界子集, 对函数 $f: D \rightarrow \mathbb{R}$, 我们取矩形 $R \subset \mathbb{R}^m$ 覆盖 D , 即 $D \subseteq R$, 并定义

$$\tilde{f}(x) = \begin{cases} f(x), & x \in D; \\ 0, & x \in R \setminus D, \end{cases}$$

如果 \tilde{f} 在 R 上 Riemann 可积, 则称 f 在 D 上 Riemann 可积, 并且记

$$\int_D f(x)d\mu = \int_R \tilde{f}(x)d\mu.$$

很明显这与覆盖 D 的矩形 R 的选取无关。

另外, 我们希望在 D 上的所有连续函数应该是 Riemann 可积函数。特别是

$$1_D(x) = \begin{cases} 1, & x \in D; \\ 0, & x \in \mathbb{R}^m \setminus D \end{cases}$$

应该是 Riemann 函数。然而这并非对每个集合 $D \subset \mathbb{R}^m$ 都成立。 1_D 是有界函数, 它的间断点集是 D 的边界 ∂D 。所以, 1_D 是 Riemann 可积函数当且仅当 ∂D 是体积为零的集合。

定义 5.2.2. 称 $D \subset \mathbb{R}^m$ 是一个 **Jordan 可测集**, 如果 ∂D 是体积为零的集合。

如果 D 是有界的 Jordan 可测集, 则记

$$\mu(D) = \int_D 1_D(x) d\mu,$$

称为 D 的 m 维体积。

例 5.2.3. 设 $D_1 \subset \mathbb{R}^m$ 是一个有界闭的 Jordan 可测集, $a, b : D_1 \rightarrow \mathbb{R}$ 是连续函数, 则

$$D = \{(\mathbf{x}, y) \in \mathbb{R}^{m+1} \mid \mathbf{x} \in D_1, a(\mathbf{x}) \leq y \leq b(\mathbf{x})\}$$

是 \mathbb{R}^{m+1} 中一个有界闭的 Jordan 可测集。

证明. 因为 a, b 在有界闭集 D_1 上连续, 所以 D 是闭集, 并且存在 $M > 0$ 使得对任意 $\mathbf{x} \in D_1$, $-M \leq a(\mathbf{x}) \leq b(\mathbf{x}) \leq M$ 。从而 $D \subseteq D_1 \times [-M, M]$, 因此 D 是有界集合。

$$\begin{aligned} \partial D &= \{(\mathbf{x}, y) \mid \mathbf{x} \in \partial D_1, a(\mathbf{x}) \leq y \leq b(\mathbf{x})\} \\ &\cup \{(\mathbf{x}, y) \mid \mathbf{x} \in D_1, y = a(\mathbf{x})\} \cup \{(\mathbf{x}, y) \mid \mathbf{x} \in D_1, y = b(\mathbf{x})\}. \end{aligned}$$

对任意 $\varepsilon > 0$, 取矩形 $\tilde{R}_1, \tilde{R}_2, \dots \subset \mathbb{R}^m$, 使得 $\partial D_1 \subset \tilde{R}_1 \cup \tilde{R}_2 \cup \dots$, 且

$$\sum_{k \geq 1} \mu_m(\tilde{R}_k) < \frac{\varepsilon}{6M}.$$

取 $R_k = \tilde{R}_k \times [-M, M]$, 则 R_k 是 \mathbb{R}^{m+1} 中的矩形, 且 $\partial D_1 \times [-M, M] \subset R_1 \cup R_2 \cup \dots$, 且

$$\sum_{k \geq 1} \mu_{m+1}(R_k) = \sum_{k \geq 1} \mu_m(\tilde{R}_k) 2M < \frac{\varepsilon}{3}.$$

另外, a, b 在有界闭集 D_1 上连续, 从而一致连续。所以存在 $\delta > 0$ 使得: 当 $\mathbf{x}_1, \mathbf{x}_2 \in D_1$ 满足 $\|\mathbf{x}_1 - \mathbf{x}_2\| < \delta$ 时,

$$|a(\mathbf{x}_1) - a(\mathbf{x}_2)| + |b(\mathbf{x}_1) - b(\mathbf{x}_2)| < \frac{\varepsilon}{6\mu_m(D_1)}.$$

存在有限多个矩形 $\hat{R}_1, \dots, \hat{R}_N$ 使得

1. $D_1 \subset \hat{R}_1 \cup \dots \cup \hat{R}_N$, 且 $\sum_{k \geq 1} \mu_m(\hat{R}_k) < 2\mu_m(D_1)$,

2. 对 $k = 1, \dots, N$, $D_1 \cap \hat{R}_k \neq \emptyset$ 且对任意 $\mathbf{x}_1, \mathbf{x}_2 \in \hat{R}_k$, $\|\mathbf{x}_1 - \mathbf{x}_2\| < \delta$ 。

因此 $\{(\mathbf{x}, y) | \mathbf{x} \in D_1, y = a(\mathbf{x})\}$ 和 $\{(\mathbf{x}, y) | \mathbf{x} \in D_1, y = b(\mathbf{x})\}$ 分别被 N 个 $m+1$ 维矩形所覆盖, 这些矩形的 $m+1$ 维体积总和小于 $\frac{2\varepsilon}{3}$. 因此 ∂D 是 \mathbb{R}^{m+1} 中一个体积为零的集合, 因此 D 是 Jordan 可测集. \square

例 5.2.4. 设 $A, B \subset \mathbb{R}^m$ 是 Jordan 可测集, 则 $A \cup B, A \cap B, A \setminus B$ 都是 Jordan 可测集.

定理 5.2.5. 设 $D \subset \mathbb{R}^m$ 是一个有界的 Jordan 可测集. 则

1. 函数 $f: D \rightarrow \mathbb{R}$ 是 Riemann 可积函数当且仅当 f 有界, 且 f 的间断点集是一个零体积集合.
2. 如果 D 还是闭集, 则 D 上的所有连续函数都是 Riemann 可积函数.

定理 5.2.6. 若函数 f 在有界的 Jordan 可测集 D_1 和 D_2 上都是 Riemann 可积的, 且 $D_1 \cap D_2 \subseteq \partial D_1 \cap \partial D_2$, 则 f 在 $D_1 \cup D_2$ 上也是 Riemann 可积的, 且 $\int_{D_1 \cup D_2} f(x) d\mu = \int_{D_1} f(x) d\mu + \int_{D_2} f(x) d\mu$.

习题 5.2

1. 记

$$D = \{(x, y) \in [0, 1]^2 | x, y \in \mathbb{Q}\}.$$

证明: D 不是 Jordan 可测集, 1_D 不是 Riemann 可积函数, $R: D \rightarrow \mathbb{R}$ 是 Riemann 可积函数, 其中

$$R: D \rightarrow \mathbb{R}, \quad R(x, y) = \frac{1}{p+m}, \text{ 若 } x = \frac{q}{p}, y = \frac{n}{m} \text{ 是既约分数.}$$

2. 设 D 是有界矩形 R 的一个闭子集. $\{R_i\}_{i \in I}$ 是有限多个矩形, 满足 $D \subseteq \bigcup_{i \in I} R_i$. 记 \mathcal{U} 为 D 的这种矩形覆盖的全体. 定义 D 的 Jordan 外测度为

$$\mu^*(D) = \inf_{\{R_i\}_{i \in I} \in \mathcal{U}} \sum_{i \in I} \mu(R_i).$$

证明: 函数 1_D 的 Darboux 上积分等于 $\mu^*(D)$, 1_D 的 Darboux 下积分等于 $\mu(R) - \mu^*(R \setminus D)$. 因此 D 是 Jordan 可测集当且仅当

$$\mu^*(D) + \mu^*(R \setminus D) = \mu(R).$$

当 D 是 Jordan 可测集时,

$$\mu^*(D) = \int_D 1 d\mu.$$

3. 设 D_1, D_2 分别是 \mathbb{R}^m 和 \mathbb{R}^n 中的有界闭的 Jordan 可测集, 证明 $D_1 \times D_2$ 是 \mathbb{R}^{m+n} 中的有界闭的 Jordan 可测集.

4. 设 D_2 是 \mathbb{R}^n 中的有界闭的 Jordan 可测集, $\varphi: [a, b] \times D_2 \rightarrow \mathbb{R}$ 连续, 且对任意 $y \in D_2$, $\varphi(x, y)$ 关于 $x \in [a, b]$ 严格减,

$$h: [a, b] \times D_2 \rightarrow \mathbb{R}^{1+n}, \quad h(x, y) = (\varphi(x, y), y),$$

证明 $h([a, b] \times D_2)$ 是 \mathbb{R}^{1+n} 中的有界闭的 Jordan 可测集。当 $\varphi(x, y)$ 关于 $x \in [a, b]$ 严格增时, 此结论也成立。(可以考虑 $h_1(x, y) = (-x, y)$ 以及 $h_2(x, y) = (\varphi(-x, y), y)$)

5.3 重积分的计算: 累次积分

定理 5.3.1. 如果 R_1, R_2 分别是 \mathbb{R}^m 和 \mathbb{R}^n 中的矩形, $f: R_1 \times R_2 \rightarrow \mathbb{R}$ 是 Riemann 可积函数, 则对任意 $y \in R_2$, $f(x, y)$ 关于 x 的 Darboux 上积分和 Darboux 下积分

$$\overline{\int}_{R_1} f(x, y) d\mu_x, \quad \underline{\int}_{R_1} f(x, y) d\mu_x$$

都是 Riemann 可积函数, 且

$$\int_{R_1 \times R_2} f(x, y) d(\mu_x \times \mu_y) = \int_{R_2} \left(\overline{\int}_{R_1} f(x, y) d\mu_x \right) d\mu_y = \int_{R_2} \left(\underline{\int}_{R_1} f(x, y) d\mu_x \right) d\mu_y.$$

定理 5.3.2. 如果 $D \subset \mathbb{R}^{m+n}$ 是一个有界闭的 Jordan 可测集, 则

$$D_2 = \{y \in \mathbb{R}^n \mid \text{存在 } x \in \mathbb{R}^m \text{ 使得 } (x, y) \in D\}$$

是 \mathbb{R}^n 中的一个有界闭的 Jordan 可测集, 且对于任意 $y \in D_2$,

$$D_y = \{x \in \mathbb{R}^m \mid (x, y) \in D\}$$

是 \mathbb{R}^m 中的一个有界闭的 Jordan 可测集。

对任何连续函数 $f: D \rightarrow \mathbb{R}$,

$$\int_D f(x, y) d(\mu_x \times \mu_y) = \int_{D_2} \int_{D_y} f(x, y) d\mu_x d\mu_y.$$

例 5.3.3. 求 $\int_0^1 \left(\int_x^1 e^{-y^2} dy \right) dx$ 。

解:

$$D = \{(x, y) \mid 0 \leq x \leq 1, x \leq y \leq 1\}$$

是有界闭的 Jordan 可测集, e^{-y^2} 是 D 上的连续函数。所以

$$\begin{aligned} \int_0^1 \left(\int_x^1 e^{-y^2} dy \right) dx &= \int_D e^{-y^2} dx dy \\ &= \int_{\mathbb{R}^2} e^{-y^2} 1_{0 \leq x \leq y \leq 1} dx dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-y^2} 1_{0 \leq x \leq y \leq 1} dx \right) dy \\ &= \int_{\mathbb{R}} \left(\int_0^y e^{-y^2} 1_{0 \leq y \leq 1} dx \right) dy = \int_0^1 e^{-y^2} \left(\int_0^y dx \right) dy \\ &= \int_0^1 ye^{-y^2} dy = -\frac{1}{2}e^{-y^2} \Big|_0^1 = \frac{1-e^{-1}}{2}. \end{aligned}$$

□

例 5.3.4. 求由圆柱面 $x^2 + y^2 = 1$, $y^2 + z^2 = 1$ 和 $z^2 + x^2 = 1$ 围成的有界区域的体积。

解:

$$\begin{aligned} D &= \{(x, y, z) | x^2 + y^2 \leq 1, y^2 + z^2 \leq 1, z^2 + x^2 \leq 1\}. \\ \mu(D) &= \int_{x^2+y^2 \leq 1, y^2+z^2 \leq 1, z^2+x^2 \leq 1} d\mu \\ &= \int_{x^2+y^2 \leq 1} \left(\int_{|z| \leq \min\{\sqrt{1-y^2}, \sqrt{1-x^2}\}} dz \right) dx dy \\ &= 2 \int_{x^2+y^2 \leq 1} \min\{\sqrt{1-y^2}, \sqrt{1-x^2}\} dx dy \\ &= 2 \int_{x^2+y^2 \leq 1, |x| \leq |y|} \sqrt{1-y^2} dx dy + 2 \int_{x^2+y^2 \leq 1, |y| < |x|} \sqrt{1-x^2} dx dy \end{aligned}$$

其中

$$\begin{aligned} &\int_{x^2+y^2 \leq 1, |x| \leq |y|} \sqrt{1-y^2} dx dy \\ &= \int_{|x| \leq |y| \leq \sqrt{1-x^2}} \sqrt{1-y^2} dx dy = \int_{|x| \leq \sqrt{1-x^2}} \left(\int_{|x| \leq |y| \leq \sqrt{1-x^2}} \sqrt{1-y^2} dy \right) dx \\ &= 4 \int_0^{\frac{1}{\sqrt{2}}} \left(\int_x^{\sqrt{1-x^2}} \sqrt{1-y^2} dy \right) dx \\ &= \left(4x \int_x^{\sqrt{1-x^2}} \sqrt{1-y^2} dy \right) \Big|_0^{\frac{1}{\sqrt{2}}} - 4 \int_0^{\frac{1}{\sqrt{2}}} x dx \left(\int_x^{\sqrt{1-x^2}} \sqrt{1-y^2} dy \right) \\ &= 4 \int_0^{\frac{1}{\sqrt{2}}} \frac{x}{\sqrt{1-x^2}} dx, \quad x = \sin \theta \\ &= 4 \int_0^{\frac{\pi}{4}} \frac{\sin \theta}{\sqrt{1-\sin^2 \theta}} d \sin \theta = 4 - 2\sqrt{2}, \end{aligned}$$

所以 $\mu(D) = 16 - 8\sqrt{2}$.

□

例 5.3.5. 求 m 维欧氏空间中半径为 R 的球的体积。

解:

$$D = \{(x_1, \dots, x_m) \in \mathbb{R}^m \mid x_1^2 + \dots + x_m^2 \leq R^2\}$$

是有界闭的 Jordan 可测集。

$$\begin{aligned} V_m(R) &= \mu_m(D) = \int_D d\mu_m \\ &= \int_{-R}^R \int_{x_1^2 + \dots + x_{m-1}^2 \leq R^2 - x_m^2} d\mu_{m-1} dx_m \\ &= \int_{-R}^R V_{m-1}(\sqrt{R^2 - x_m^2}) dx_m \end{aligned}$$

利用这个递推关系以及 $V_1(R) = 2R$, 可以猜测并用数学归纳法证明对任意正整数 m , $V_m(R) = A_m R^m$, 从而

$$\begin{aligned} A_m &= V_m(1) = \int_{-1}^1 A_{m-1}(\sqrt{1-t^2})^{m-1} dt = A_{m-1} \int_0^1 s^{-\frac{1}{2}}(1-s)^{\frac{m-1}{2}} ds \\ &= A_{m-1} B\left(\frac{1}{2}, \frac{m+1}{2}\right) = A_{m-1} \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m+2}{2}\right)} \\ &= \left(\Gamma\left(\frac{1}{2}\right)\right)^{m-1} \frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{m+2}{2}\right)} A_1 = \frac{\pi^{\frac{m}{2}}}{\Gamma\left(\frac{m+2}{2}\right)}, \end{aligned}$$

因此 $V_m(R) = \frac{\pi^{\frac{m}{2}}}{\Gamma\left(\frac{m+2}{2}\right)} R^m$. □

总结: 把积分区域 D 用不等式(组)的形式表达, 固定其中一些变量 x , 解出另一些变量 y 的范围 $D(x)$, 再取 $D_1 = \{x \mid D(x) \neq \emptyset\}$. 于是

$$\int_D f(x, y) d(\mu_x \times \mu_y) = \int_{D_1} \left(\int_{D(x)} f(x, y) d\mu_y \right) d\mu_x.$$

习题5.3

1. 设 R 是 \mathbb{R}^m 中的矩形, L 是对角元素全是 1 的 m 阶下三角矩阵, U 是对角元素全是 1 的 m 阶上三角矩阵, P 是 m 阶排列矩阵,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{pmatrix}.$$

- (a) 证明 $L(R), U(R), P(R), \Lambda(R)$ 都是 \mathbb{R}^m 中的有界闭的 Jordan 可测集;
 (b) 求 $\mu(L(R)), \mu(U(R)), \mu(P(R)), \mu(\Lambda(R))$ 的值;
 (c) 证明对任何 m 阶可逆矩阵 A , $A(R)$ 是 \mathbb{R}^m 中的有界闭的 Jordan 可测集, 并且 $\mu(A(R)) = |\det A| \mu(R)$.

2. 在长度为 1 的线段上随机取两点, 求它们的距离的平均值。

5.4 重积分的计算：换元公式

定理 5.4.1. 设 $D \subset \mathbb{R}^m$ 是一个有界闭的 Jordan 可测集, $\varphi: D \rightarrow \Omega = \varphi(D) \subseteq \mathbb{R}^m$ 是一个 C^1 微分同胚. 则对任意 $f \in \mathfrak{R}(\Omega)$,

$$\int_{\Omega} f(u) d\mu_{\Omega} = \int_D f(\varphi(x)) |\det D\varphi(x)| d\mu_D.$$

用传统符号表达, 上式为

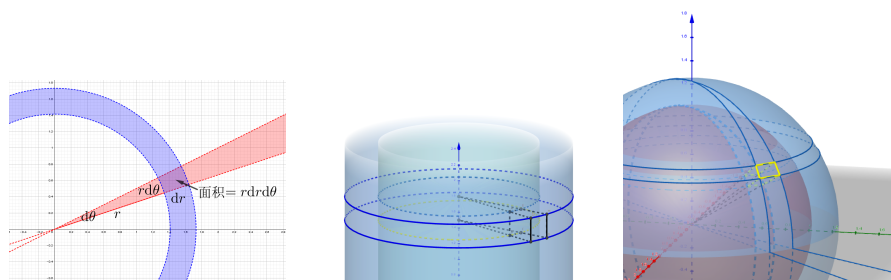
$$\int_{\Omega} f(u) du_1 \cdots du_m = \int_D f(\varphi(x)) \left| \det \frac{\partial(u_1, \dots, u_m)}{\partial(x_1, \dots, x_m)} \right| dx_1 \cdots dx_m.$$

注 5.4.2. (1) 积分换元的目的是为了把积分区域或者被积函数变得简单, 使积分计算得到简化.

(2) 微分同胚 φ 相当于一个非线性的坐标变换, 它把 D 中的 Jordan 可测集 R 变成 Ω 中的 Jordan 可测集 $\varphi(R)$. 在点 x 附近, φ 可以近似看成线性映射 $\partial\varphi(x)$, 它把 x 附近的小矩形 R 变成 $\varphi(R)$, 后者近似为一个平行多面体 $\partial\varphi(x)(R)$, 从而

$$\mu_{\Omega}(\varphi(R)) \approx \mu_{\Omega}(\partial\varphi(x)(R)) = |\det \partial\varphi(x)| \mu_D(R).$$

(3) 当积分区域或被积函数具有某种对称性时, 通常会采用适合这种对称性的坐



标变换. 例如对平面上的极坐标系 $x = r \cos \theta, y = r \sin \theta$,

$$dx dy = \left| \det \frac{\partial(x, y)}{\partial(r, \theta)} \right| dr d\theta = r dr d\theta.$$

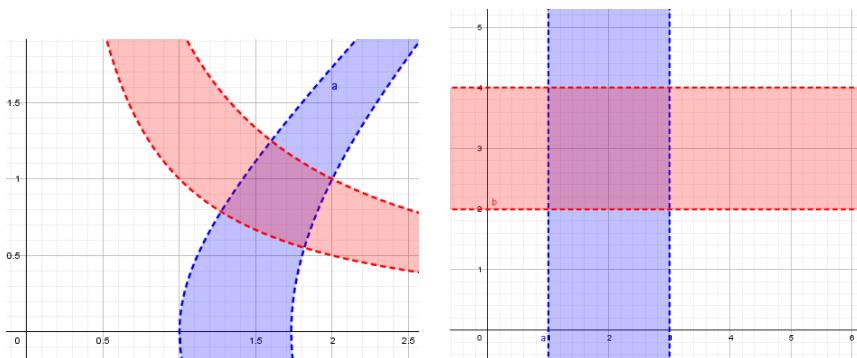
在三维空间中绕 z 轴旋转的柱面坐标 $x = r \cos \theta, y = r \sin \theta, z = z$,

$$dx dy dz = \left| \det \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| dr d\theta dz = r dr d\theta dz.$$

在三维空间中以原点为中心的球面坐标 $x = R \sin \theta \cos \varphi, y = R \sin \theta \sin \varphi, z = R \cos \theta$,

$$dx dy dz = \left| \det \frac{\partial(x, y, z)}{\partial(R, \theta, \varphi)} \right| dr d\theta dz = R^2 \sin \theta dR d\theta d\varphi.$$

例 5.4.3. 设 $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid 1 \leq x_1^2 - x_2^2 \leq 3, 2 \leq 2x_1x_2 \leq 4\}$. 求 $\int_D (x_1^2 + x_2^2) dx_1 dx_2$.



解: 令 $(y_1, y_2) = \varphi(x_1, x_2) = (x_1^2 - x_2^2, 2x_1x_2)$ 。则 φ 是 $(0, +\infty) \times (0, +\infty)$ 到 $(0, +\infty) \times \mathbb{R}$ 的微分同胚, 且 $\varphi(D) = \{(y_1, y_2) \in \mathbb{R}^2 | 1 \leq y_1 \leq 3, 2 \leq y_2 \leq 4\}$,

$$\det \frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{vmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{vmatrix} = 4(x_1^2 + x_2^2) > 0, \quad \forall (x_1, x_2) \neq (0, 0),$$

$$y_1^2 + y_2^2 = (x_1^2 + x_2^2)^2 \quad (\mathbf{y} = \varphi(\mathbf{x})).$$

$$\begin{aligned} \int_D (x_1^2 + x_2^2) dx_1 dx_2 &= \int_{\varphi(D)} \sqrt{y_1^2 + y_2^2} \left| \det \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right| dy_1 dy_2, \quad \mathbf{x} = \varphi^{-1}(\mathbf{y}) \\ &= \int_{[1,3] \times [2,4]} \sqrt{y_1^2 + y_2^2} \left| \frac{1}{4\sqrt{y_1^2 + y_2^2}} \right| dy_1 dy_2 \\ &= 1. \end{aligned}$$

□

例 5.4.4. 求 $\int_D \frac{x+y}{x^2+y^2} dx dy$, 其中

$$D = \{(x, y) | x^2 + y^2 \leq x + y\}.$$

解: 对 $0 < t < 1$, $2t^2 < 2t$, 所以 $(t, t) \in D$,

$$\lim_{t \rightarrow 0^+} \frac{t+t}{t^2+t^2} = +\infty,$$

被积函数无界, 所以上述积分不是 Riemann 积分。

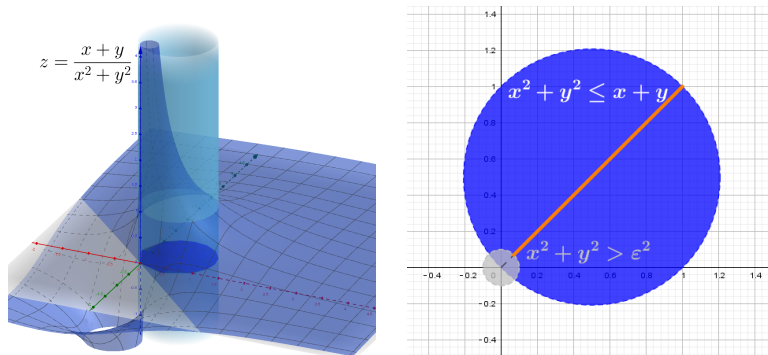
对 $\varepsilon > 0$, 记

$$D_\varepsilon = \{(x, y) | \varepsilon^2 \leq x^2 + y^2 \leq x + y\}.$$

则 $f(x, y) = \frac{x+y}{x^2+y^2}$ 在 D_ε 上是连续函数, 从而 Riemann 可积。

令 $x = r \cos(\theta - \frac{\pi}{4})$, $x = r \sin(\theta - \frac{\pi}{4})$, 则

$$x^2 + y^2 = r^2, \quad x + y = \sqrt{2}r \sin \theta, \quad \frac{x+y}{x^2+y^2} = \frac{\sqrt{2} \sin \theta}{r},$$



于是 $(x, y) \in D_\varepsilon \Leftrightarrow \varepsilon \leq r \leq \sqrt{2} \sin \theta$, 在 (r, θ) 坐标系下, 积分区域对应的不等式以及被积函数都是变量分离的。

$$\begin{aligned}
 \int_{D_\varepsilon} \frac{x+y}{x^2+y^2} dx dy &= \int_{\varepsilon \leq r \leq \sqrt{2} \sin \theta} \frac{r\sqrt{2} \sin \theta}{r^2} r dr d\theta \\
 &= \int_{\varepsilon \leq \sqrt{2} \sin \theta} \left(\int_{\varepsilon}^{\sqrt{2} \sin \theta} \sqrt{2} \sin \theta dr \right) d\theta \\
 &= \int_{\arcsin \frac{\varepsilon}{\sqrt{2}}}^{\pi - \arcsin \frac{\varepsilon}{\sqrt{2}}} 1 - \cos 2\theta - \sqrt{2}\varepsilon \sin \theta d\theta \\
 &= \pi - 2 \arcsin \frac{\varepsilon}{\sqrt{2}} + \sin \left(2 \arcsin \frac{\varepsilon}{\sqrt{2}} \right) - 2\sqrt{2}\varepsilon \cos \arcsin \frac{\varepsilon}{\sqrt{2}} \\
 &\rightarrow \pi, \quad \varepsilon \rightarrow 0^+.
 \end{aligned}$$

所以

$$\int_D \frac{x+y}{x^2+y^2} dx dy = \pi.$$

□

例 5.4.5. 求 $\int_D f(x_1 + x_2 + \cdots + x_m) dx_1 dx_2 \cdots dx_m$, 其中

$$D = \{(x_1, \dots, x_m) | x_1, \dots, x_m \geq 0, x_1 + x_2 + \cdots + x_m \leq a\}.$$

解: 令

$$\begin{cases} y_1 = x_1, \\ y_2 = x_1 + x_2, \\ \vdots \\ y_m = x_1 + x_2 + \cdots + x_m. \end{cases}$$

则 $\varphi(\mathbf{x}) = \mathbf{y}$ 满足 $\varphi(D) = \Omega = \{(y_1, y_2, \dots, y_m) | 0 \leq y_1 \leq y_2 \leq \cdots \leq y_m \leq a\}$,

$\det \partial\varphi(\mathbf{x}) = 1$ 。于是

$$\begin{aligned} & \int_D f(x_1 + x_2 + \cdots + x_m) dx_1 dx_2 \cdots dx_m \\ &= \int_{\Omega} f(y_m) \left| \det \frac{\partial(x_1, x_2, \dots, x_m)}{\partial(y_1, y_2, \dots, y_m)} \right| dy_1 dy_2 \cdots dy_m \\ &= \int_0^a f(t) \left(\int_{0 \leq y_1 \leq y_2 \leq \cdots \leq y_{m-1} \leq t} dy_1 dy_2 \cdots dy_{m-1} \right) dt \\ &= \int_0^a f(t) \frac{t^{m-1}}{(m-1)!} dt \end{aligned}$$

□

例 5.4.6. 设 Σ 是 m 阶对称正定矩阵, $\mu \in \mathbb{R}^m$, 证明

$$\int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)} dx_1 \cdots dx_m = 1.$$

解: 取可逆矩阵 A 使得 $\Sigma = AA^T$ 。令 $\mathbf{y} = A^{-1}(\mathbf{x} - \mu)$ 。则

$$\det \Sigma = \det A \det A^T = (\det A)^2, \quad A^T \Sigma^{-1} A = I,$$

于是

$$\begin{aligned} & \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)} dx_1 \cdots dx_m \\ &= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det \Sigma}} e^{-\frac{1}{2}(A\mathbf{y})^T \Sigma^{-1}(A\mathbf{y})} \left| \det \frac{\partial(x_1, \dots, x_m)}{\partial(y_1, \dots, y_m)} \right| dy_1 \cdots dy_m \\ &= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det \Sigma}} e^{-\frac{1}{2}\mathbf{y}^T A^T \Sigma^{-1} A \mathbf{y}} |\det A| dy_1 \cdots dy_m \\ &= \int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m}} e^{-\frac{1}{2}\mathbf{y}^T \mathbf{y}} dy_1 \cdots dy_m = \prod_{k=1}^m \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_k^2} dy_k. \end{aligned}$$

当 $m = 2$ 时,

$$\int_{\mathbb{R}^2} \frac{1}{2\pi} e^{-\frac{y_1^2 + y_2^2}{2}} dy_1 dy_2 = \int_{r>0, 0 \leq \theta \leq 2\pi} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta = -e^{-\frac{r^2}{2}} \Big|_0^{+\infty} = 1.$$

所以 $\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt = 1$, 从而

$$\int_{\mathbb{R}^m} \frac{1}{\sqrt{(2\pi)^m \det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)} dx_1 \cdots dx_m = 1.$$

需要说明的是, 这里的积分不是 Riemann 积分, 是收敛的广义重积分。 □

应用中(例如物理和概率统计中)常常遇到广义重积分, 想深入了解广义重积分的数学理论的读者可以阅读数学分析教材。

习题5.4

1. 求 $\int_D \left| \frac{y}{x} \right| dx dy$, 其中 $D = \{(x, y) | 1 \leq x^2 + y^2 \leq 2x\}$ 。
2. 求 $\int_{x_1^2 + x_2^2 + \dots + x_m^2 \leq R^2} f(\sqrt{x_1^2 + x_2^2 + \dots + x_m^2}) dx_1 dx_2 \dots dx_m$ 。
3. 证明一元微积分中给出的旋转体体积公式。
4. 在边长为 1 的正方形区域中随机取两个点, 求它们之间距离的平均值。
5. 求 n 维锥体的体积公式。

5.5 重积分应用的几个例子

例 5.5.1. 求质量均匀分布的 m 维实心半球的质心, 其中半径为 R 。

解: 设密度为 $\rho > 0$ 。则质量为

$$M = \int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} \rho dx_1 \dots dx_m.$$

质心为

$$\bar{\mathbf{x}} = \frac{1}{M} \int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} \mathbf{x} \rho dx_1 \dots dx_m = \frac{\int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} \mathbf{x} dx_1 \dots dx_m}{\int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} dx_1 \dots dx_m}.$$

由对称性知 $\bar{x}_1 = \dots = \bar{x}_{m-1} = 0$ 。我们已经知道

$$\begin{aligned} \int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} dx_1 \dots dx_m &= \frac{\pi^{\frac{m}{2}}}{2\Gamma\left(\frac{m+2}{2}\right)} R^m. \\ \int_{x_1^2 + \dots + x_m^2 \leq R^2, x_m \geq 0} x_m dx_1 \dots dx_m &= \int_0^R x_m \left(\int_{x_1^2 + \dots + x_{m-1}^2 \leq R^2 - x_m^2} dx_1 \dots dx_{m-1} \right) dx_m \\ &= \int_0^R t \frac{\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m+1}{2}\right)} (R^2 - t^2)^{\frac{m-1}{2}} dt \\ &= \frac{\pi^{\frac{m-1}{2}}}{2\Gamma\left(\frac{m+1}{2}\right)} R^{m+1} \int_0^1 s^{\frac{m-1}{2}} ds, \quad s = 1 - t^2 \\ &= \frac{\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m+1}{2}\right)} R^{m+1} \frac{1}{m+1} = \frac{\pi^{\frac{m-1}{2}}}{2\Gamma\left(\frac{m+3}{2}\right)} R^{m+1}, \end{aligned}$$

因此

$$\bar{x}_m = \frac{\frac{\pi^{\frac{m-1}{2}}}{2\Gamma\left(\frac{m+3}{2}\right)} R^{m+1}}{\frac{\pi^{\frac{m}{2}}}{2\Gamma\left(\frac{m+2}{2}\right)} R^m} = \frac{\Gamma\left(\frac{m+2}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m+3}{2}\right)} R.$$

□

例 5.5.2. 在区间 $[0, 1]$ 中独立地随机抽取 n 个数, 求其中的最小值的平均值(概率里称为“期望”或“数学期望”).

解: 区间 $[0, 1]$ 上均匀分布的密度为 $f(x) = 1_{0 \leq x \leq 1}$, 独立地随机抽取 n 个数, 概率密度为

$$f(x_1)f(x_2) \cdots f(x_n) = 1_{(x_1, \dots, x_n) \in [0, 1]^n}.$$

最小值的平均值为

$$\begin{aligned} & \int_{[0, 1]^n} \min_{1 \leq k \leq n} x_k \cdot f(x_1)f(x_2) \cdots f(x_n) dx_1 \cdots dx_n \\ &= \int_{[0, 1]^n} \sum_{\sigma} 1_{x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}} \min_{1 \leq k \leq n} x_k dx_1 \cdots dx_n \\ & \quad (\sigma \text{ 取遍 } 1, 2, \dots, n \text{ 的所有排列}) \\ &= \sum_{\sigma} \int_{0 < x_{\sigma(1)} < \cdots < x_{\sigma(n)} < 1} x_{\sigma(1)} \cdot \left| \det \frac{\partial(x_1, \dots, x_n)}{\partial(x_{\sigma(1)}, \dots, x_{\sigma(n)})} \right| dx_{\sigma(1)} \cdots dx_{\sigma(n)} \\ &= \sum_{\sigma} \int_{0 < y_1 < \cdots < y_n < 1} y_1 dy_1 \cdots dy_n \\ &= n! \int_0^1 y_1 \left(\int_{y_1 < y_2 < \cdots < y_n < 1} dy_2 \cdots dy_n \right) dy_1 \\ &= n \int_0^1 y_1 \left(\int_{(y_2, \dots, y_n) \in (y_1, 1)^{n-1}} dy_2 \cdots dy_n \right) dy_1 \\ &= n \int_0^1 y_1 (1 - y_1)^{n-1} dy_1 = nB(2, n) = n \frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} = \frac{1}{n+1}. \end{aligned}$$

□

物理中的质心、概率中的期望都是计算加权平均。计算质心时, 被平均的对象是点的位置, 相应的权重是各点处的密度, 它刻画了质量在空间中的分布情况。计算期望时, 被平均的对象是随机变量的取值, 相应的权重是各个取值对应的概率密度, 它刻画了概率在随机变量取值空间中的分布。

一般地, 如果 $f(x)$ 是被平均的对象, 权重为 $g(x) > 0$, 则 $f(x)$ 的加权平均值为

$$\bar{f} = \frac{\int_D f(x)g(x)d\mu}{\int_D g(x)d\mu}.$$

例 5.5.3. 求密度为 ρ 半径为 R 的三维球体对球外质点 m 的万有引力。

解: 取直角坐标系使球心位于原点, 而质点 m 位于 x_3 轴坐标为 a 处。

$$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_3^2 \leq R^2\}.$$

$$d\mathbf{F}(\mathbf{x}) = \frac{Gm\rho d\mu}{\|\mathbf{x} - a\mathbf{e}_3\|^2} \frac{\mathbf{x} - a\mathbf{e}_3}{\|\mathbf{x} - a\mathbf{e}_3\|},$$

由对称性, 球对质点 m 的引力位于 x_3 轴上, 从 m 指向球心,

$$\begin{aligned}
 \mathbf{F} &= \mathbf{e}_3 \int_D \frac{Gm\rho(x_3 - a)}{\|\mathbf{x} - a\mathbf{e}_3\|^3} d\mu \\
 &= \mathbf{e}_3 \int_{-R}^R \left(\int_{x_1^2 + x_2^2 \leq R^2 - x_3^2} \frac{Gm\rho(x_3 - a)}{(x_1^2 + x_2^2 + (a - x_3)^2)^{3/2}} dx_1 dx_2 \right) dx_3 \\
 &= \mathbf{e}_3 \int_{-R}^R \left(\int_{0 \leq r \leq \sqrt{R^2 - x_3^2}, 0 \leq \theta \leq 2\pi} \frac{Gm\rho(x_3 - a)}{(r^2 + (a - x_3)^2)^{3/2}} r dr d\theta \right) dx_3 \\
 &\quad (x_1 x_2 \text{ 坐标平面中用极坐标 } x_1 = r \cos \theta, x_2 = r \sin \theta, \det \frac{\partial(x_1, x_2)}{\partial(r, \theta)} = r) \\
 &= \mathbf{e}_3 \int_{-R}^R -\frac{2\pi Gm\rho(x_3 - a)}{\sqrt{r^2 + (a - x_3)^2}} \Big|_0^{\sqrt{R^2 - x_3^2}} dx_3 \\
 &= 2\pi Gm\rho \mathbf{e}_3 \int_{-R}^R -1 - \frac{(x_3 - a)}{\sqrt{R^2 + a^2 - 2ax_3}} dx_3 \\
 &= -\frac{Gm\rho \frac{4\pi R^3}{3}}{a^2} \mathbf{e}_3.
 \end{aligned}$$

结论: 球体对质点 m 的引力相当于把球体的全部质量集中于球心时对质点 m 的引力。□

例 5.5.4 (Kepler 第二定律与有心力场). 设 $\mathbf{u} : J \rightarrow \mathbb{R}^2$ 是一个平面运动, 其中 $J \subseteq \mathbb{R}$ 是一个区间。记 Ω_J 为连接原点和 $\mathbf{u}(t)$ 线段在时段 J 内扫过的区域。则 Ω_J 中的点可以用 $(s, t) \in [0, 1] \times J$ 来表示: $(x_1(s, t), x_2(s, t)) = s\mathbf{u}(t) = (su_1(t), su_2(t))$ 。于是

$$\begin{aligned}
 \mu(\Omega_J) &= \int_{\Omega_J} dx_1 dx_2 = \int_{[0,1] \times J} \left| \det \frac{\partial(x_1, x_2)}{\partial(s, t)} \right| ds dt \\
 &= \int_{[0,1] \times J} |\det(\mathbf{u}(t), s\mathbf{u}'(t))| ds dt \\
 &= \int_0^1 s ds \int_J |\det(\mathbf{u}(t), \mathbf{u}'(t))| dt \\
 &= \frac{1}{2} \int_J |\det(\mathbf{u}(t), \mathbf{u}'(t))| dt.
 \end{aligned}$$

Kepler 第二定律断言行星运动满足: $\frac{\mu(\Omega_J)}{|J|}$ 是常值。这等价于 $|\det(\mathbf{u}(t), \mathbf{u}'(t))|$ 是常数。又因为 $\det(\mathbf{u}(t), \mathbf{u}'(t))$ 关于 t 连续, 所以上述结论等价于 $\det(\mathbf{u}(t), \mathbf{u}'(t))$ 是常数, 即

$$\frac{d}{dt} \det(\mathbf{u}(t), \mathbf{u}'(t)) = 0.$$

而

$$\frac{d}{dt} \det(\mathbf{u}(t), \mathbf{u}'(t)) = \det(\mathbf{u}'(t), \mathbf{u}'(t)) + \det(\mathbf{u}(t), \mathbf{u}''(t)) = \det(\mathbf{u}(t), \mathbf{u}''(t)),$$

所以 Kepler 第二定律等价于 $\det(\mathbf{u}(t), \mathbf{u}''(t)) = 0$, 即 $\mathbf{u}''(t)$ 总是与 $\mathbf{u}(t)$ 在同一条直线上, 也就是当且仅当质点受到的力总是沿位置向径的。□

习题5.5

1. 关于球体对球外质点的万有引力的结论是否与空间维数有关?
2. 设 γ 为平面曲线, 它在极坐标系下的方程为 $r = a(1 + \cos \theta)$ ($a > 0$)。
 - (a) 求 γ 所围成的平面有界区域 D 的面积和形心坐标。
 - (b) 求平面区域 D 绕 x 轴旋转所得有界区域的体积和形心坐标。

5.6 习题讨论课6

1. 设 γ 是平面中一条轴对称的简单封闭曲线, D 是 γ 所围成的平面有界区域, Σ 是 γ 绕旋转轴旋转一周形成的曲面, Ω 是 Σ 所围成的空间有界区域。试讨论曲线 γ 、区域 D 、曲面 Σ 、区域 Ω 的形心坐标的关系。

第6章 曲线与曲面上函数的积分

6.1 第一型曲线积分

曲线的弧长

称 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^m$ 是一个正则的 C^1 参数曲线, 如果它是单射并且对任何 $t \in [a, b]$, $\|\mathbf{x}'(t)\| > 0$ 。称

$$\int_{[a,b]} \|\mathbf{x}'(t)\| dt = \int_a^b \|\mathbf{x}'(t)\| dt$$

为正则的 C^1 参数曲线 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^m$ 的弧长。我们把 t 想象成时间, 把 $\mathbf{x}(t)$ 理解为空间 \mathbb{R}^m 中一个运动着的点在时刻 t 时所处的位置, 于是 $\|\mathbf{x}'(t)\|$ 就是这个动点在时刻 t 时的速率, 所以弧长就是这个动点在时间段 $[a, b]$ 中的运动路程。

如果 $t = t(s)$ 是 C^1 函数, 满足 $t'(s) \neq 0 (\forall s)$, $t[\alpha, \beta] = [a, b]$, 则

$$\begin{aligned} \int_{[\alpha,\beta]} \left\| \frac{d}{ds} \mathbf{x}(t(s)) \right\| ds &= \int_{[\alpha,\beta]} \|\mathbf{x}'(t(s))t'(s)\| ds = \int_{[\alpha,\beta]} \|\mathbf{x}'(t(s))\| |t'(s)| ds \\ &= \int_{[a,b]} \|\mathbf{x}'(t)\| dt, \end{aligned}$$

最后这个等号是利用了(无向)积分换元。因此曲线的弧长与曲线的 C^1 参数化形式的选择无关。也就是说一个动点沿曲线运动产生的路程与它的运动速度和运动方向都没有关系, 这是曲线本身的一个几何属性。

记 $\gamma = \{\mathbf{x}(t) | t \in [a, b]\}$, 称

$$L(\gamma) = \int_{[a,b]} \|\mathbf{x}'(t)\| dt$$

为曲线 γ 的弧长, 称

$$l = l(t) = \int_{[a,t]} \|\mathbf{x}'(s)\| ds,$$

为曲线 γ 的弧长参数。于是

$$dl(t) = \|\mathbf{x}'(t)\|dt.$$

由一阶微分的形式不变性(即链索法则), dl 与曲线 γ 的参数化形式无关。如果 $\mathbf{x}(l)$ 是 γ 的以弧长参数 l 为自变量的参数表示, 则 $\|\mathbf{x}'(l)\| = 1$ 。

函数沿曲线的积分

设 $f: \gamma \rightarrow \mathbb{R}$ 为连续函数, $\mathbf{x}(t)$ 为曲线 γ 的任何一个 C^1 参数化表示, 记

$$\int_{\gamma} f dl = \int_{[a,b]} f(\mathbf{x}(t))\|\mathbf{x}'(t)\|dt,$$

称为 f 沿 γ 的积分。由积分的换元公式知, 上述积分的值与曲线 γ 的参数化形式的选择无关。

例 6.1.1. 证明 \mathbb{R}^m 中连接两点的 C^2 曲线中以直线段长度最短。

证明. 设 γ 是连接 P, Q 两点的 C^2 曲线, 参数方程为 $\mathbf{x}: [0, l] \rightarrow \mathbb{R}^m$, 满足 $\mathbf{x}(0) = P$, $\mathbf{x}(l) = Q$, 其中 l 是 γ 的弧长。则

$$L(\gamma) = \int_{\gamma} dl = \int_{[0,l]} \|\mathbf{x}'(t)\|dt.$$

任取 C^2 映射 $\mathbf{y}: [0, l] \rightarrow \mathbb{R}^m$ 满足 $\mathbf{y}(0) = \mathbf{y}(l) = 0$ 。记 γ_s 为曲线 $\mathbf{x}(t) + s\mathbf{y}(t)$, 于是

$$L(\gamma_s) = \int_{[0,l]} \|\mathbf{x}'(t) + s\mathbf{y}'(t)\|dt = \int_{[0,l]} \sqrt{\langle \mathbf{x}'(t) + s\mathbf{y}'(t), \mathbf{x}'(t) + s\mathbf{y}'(t) \rangle} dt.$$

从而

$$\begin{aligned} \left. \frac{d}{ds} L(\gamma_s) \right|_{s=0} &= \left. \frac{d}{ds} \int_{[0,l]} \sqrt{\langle \mathbf{x}'(t) + s\mathbf{y}'(t), \mathbf{x}'(t) + s\mathbf{y}'(t) \rangle} dt \right|_{s=0} \\ &= \int_{[0,l]} \left. \frac{\partial}{\partial s} \sqrt{\langle \mathbf{x}'(t) + s\mathbf{y}'(t), \mathbf{x}'(t) + s\mathbf{y}'(t) \rangle} \right|_{s=0} dt \\ &= \int_{[0,l]} \frac{\langle \mathbf{x}'(t), \mathbf{y}'(t) \rangle}{\sqrt{\langle \mathbf{x}'(t), \mathbf{x}'(t) \rangle}} dt = \int_{[0,l]} \langle \mathbf{x}'(t), \mathbf{y}'(t) \rangle dt, \end{aligned}$$

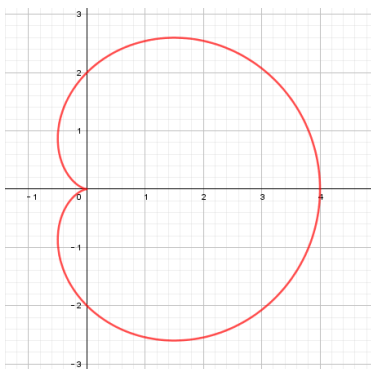
最后一个等式利用了弧长参数下 $\|\mathbf{x}'(t)\| = 1$ 。从而

$$\begin{aligned} \left. \frac{d}{ds} L(\gamma_s) \right|_{s=0} &= \int_{[0,l]} \langle \mathbf{x}'(t), \mathbf{y}'(t) \rangle dt \\ &= \int_{[0,l]} d\langle \mathbf{x}'(t), \mathbf{y}(t) \rangle - \int_{[0,l]} \langle \mathbf{x}''(t), \mathbf{y}(t) \rangle dt \\ &= \langle \mathbf{x}'(t), \mathbf{y}(t) \rangle \Big|_{t=0}^{t=l} - \int_{[0,l]} \langle \mathbf{x}''(t), \mathbf{y}(t) \rangle dt \\ &= - \int_{[0,l]} \langle \mathbf{x}''(t), \mathbf{y}(t) \rangle dt. \end{aligned}$$

曲线 γ 使 $L(\gamma)$ 取最小值, 当且仅当对满足 $\mathbf{y}(0) = \mathbf{y}(l) = 0$ 的任意 C^2 映射 $\mathbf{y} : [0, l] \rightarrow \mathbb{R}^m$, $\int_{[0, l]} \langle \mathbf{x}''(t), \mathbf{y}(t) \rangle dt = 0$, 因此对任意 t , $\mathbf{x}''(t) = 0$, 所以 γ 是直线段。 \square

读者可以考虑下面这个问题: 在一个曲面上, 连接两点且长度最短的曲线是什么?

例 6.1.2. 设平面曲线 γ 在极坐标系下的方程为 $r = a(1 + \cos \theta)$ ($a > 0$)。求 γ 的弧长和它的形心坐标。



解: 曲线 γ 在直角坐标系下的参数方程为

$$\begin{cases} x(\theta) = r \cos \theta = a(1 + \cos \theta) \cos \theta = a \cos \theta + \frac{a}{2}(\cos 2\theta + 1), \\ y(\theta) = r \sin \theta = a(1 + \cos \theta) \sin \theta = a \sin \theta + \frac{a}{2} \sin 2\theta. \end{cases}$$

于是

$$\sqrt{x'(\theta)^2 + y'(\theta)^2} = a\sqrt{2}\sqrt{1 + \cos \theta} = 2a \left| \cos \frac{\theta}{2} \right|.$$

于是

$$L = \int_0^{2\pi} \sqrt{x'(\theta)^2 + y'(\theta)^2} d\theta = 2a \int_0^{2\pi} \left| \cos \frac{\theta}{2} \right| d\theta = 8a.$$

形心位于 (\bar{x}, \bar{y}) , 则易见 $\bar{y} = 0$, 且

$$\bar{x} = \frac{\int_0^{2\pi} x(\theta) \sqrt{x'(\theta)^2 + y'(\theta)^2} d\theta}{L} = \frac{a^2 \int_0^{2\pi} (2 \cos \theta + \cos 2\theta + 1) \left| \cos \frac{\theta}{2} \right| d\theta}{L} = \frac{4a}{5}.$$

所以 γ 形心位于 $(\frac{4a}{5}, 0)$ 。 \square

不同坐标系下的弧长

在欧氏空间直角坐标系下,

$$dl = \|\mathbf{x}'(t)\| dt = \sqrt{\langle \mathbf{x}'(t), \mathbf{x}'(t) \rangle} dt = \sqrt{\sum_{k=1}^m x'_k(t)^2} dt = \sqrt{\sum_{k=1}^m (dx_k)^2}.$$

如果考虑坐标变换 $\mathbf{x} = \varphi(\mathbf{u})$, 则

$$\begin{aligned} dl &= \|\mathbf{x}'(t)\|dt = \sqrt{\langle \mathbf{x}'(t), \mathbf{x}'(t) \rangle} dt \\ &= \sqrt{\langle \partial\varphi(\mathbf{u}(t))\mathbf{u}'(t), \partial\varphi(\mathbf{u}(t))\mathbf{u}'(t) \rangle} dt \\ &= \sqrt{\left\langle \sum_{i=1}^m \frac{\partial\mathbf{x}}{\partial u_i}(\mathbf{u}(t))u'_i(t), \sum_{j=1}^m \frac{\partial\mathbf{x}}{\partial u_j}(\mathbf{u}(t))u'_j(t) \right\rangle} dt \\ &= \sqrt{\sum_{i,j=1}^m \left\langle \frac{\partial\mathbf{x}}{\partial u_i}(\mathbf{u}(t)), \frac{\partial\mathbf{x}}{\partial u_j}(\mathbf{u}(t)) \right\rangle du_i du_j} \\ &= \sqrt{(du_1, \dots, du_m) \left(\left\langle \frac{\partial\mathbf{x}}{\partial u_i}(\mathbf{u}(t)), \frac{\partial\mathbf{x}}{\partial u_j}(\mathbf{u}(t)) \right\rangle \right)_{m \times m} \begin{pmatrix} du_1 \\ \vdots \\ du_m \end{pmatrix}} \end{aligned}$$

这是一个只与度量(即内积)有关而与坐标无关的形式, 它适用于各种坐标系 (u_1, \dots, u_m) , 矩阵

$$\left(\left\langle \frac{\partial\mathbf{x}}{\partial u_i}(\mathbf{u}), \frac{\partial\mathbf{x}}{\partial u_j}(\mathbf{u}) \right\rangle \right)_{m \times m}$$

称为度量矩阵也称为 **Gram 矩阵**。特别是当 $\partial\varphi(\mathbf{u})$ 是正交变换时, 度量矩阵为单位矩阵, 从而坐标系 (u_1, \dots, u_m) 中的微弧长与坐标系 (x_1, \dots, x_m) 中的微弧长相等。

例如对平面极坐标 (r, θ) ,

$$\frac{\partial\mathbf{x}}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \frac{\partial\mathbf{x}}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix},$$

度量矩阵为

$$\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

所以

$$dl = \sqrt{(dr)^2 + (rd\theta)^2}.$$

请对最后这个结果给出直接的几何解释。

习题6.1

1. 分别求柱面坐标系和球坐标系下曲线的弧长公式。
2. 球面上连接两点的最短曲线是什么?
3. 求曲线 $\begin{cases} z = \sqrt{4 - x^2 - y^2}, \\ (x - 1)^2 + y^2 = 1 \end{cases}$ 的长度和形心。

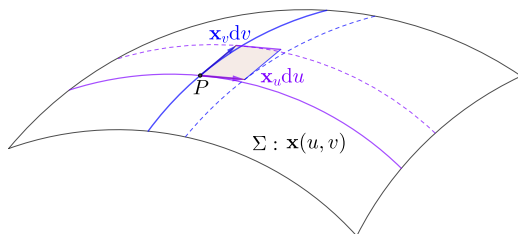
6.2 二维曲面上的第一型曲面积分

曲面的面积微元

\mathbb{R}^m ($m \geq 3$)中的二维曲面 Σ 的 C^1 正则参数表示是一个 C^1 单射

$$\mathbf{x} : D \rightarrow \mathbb{R}^m, \quad \mathbf{x}(u_1, u_2) = \begin{pmatrix} x_1(u_1, u_2) \\ x_2(u_1, u_2) \\ \vdots \\ x_m(u_1, u_2) \end{pmatrix}, \quad (u_1, u_2) \in D \subseteq \mathbb{R}^2$$

满足 $\frac{\partial(x_1, x_2, \dots, x_m)}{\partial(u_1, u_2)}$ 是满秩矩阵, 也即向量 $\frac{\partial \mathbf{x}}{\partial u_1}$ 和 $\frac{\partial \mathbf{x}}{\partial u_2}$ 是 Σ 线性无关, 它们张成了曲面 Σ 在 \mathbf{x} 处的切空间。



视 $\frac{\partial \mathbf{x}}{\partial u_1} du_1$ 和 $\frac{\partial \mathbf{x}}{\partial u_2} du_2$ 是一对无穷小的切向量, 记它们的夹角为 θ , 则它们形成的无穷小平行四边形的面积为

$$\begin{aligned} d\sigma &= \left\| \frac{\partial \mathbf{x}}{\partial u_1} du_1 \right\| \left\| \frac{\partial \mathbf{x}}{\partial u_2} du_2 \right\| \sin \theta = \left\| \frac{\partial \mathbf{x}}{\partial u_1} du_1 \right\| \left\| \frac{\partial \mathbf{x}}{\partial u_2} du_2 \right\| \sqrt{1 - \cos^2 \theta} \\ &= \sqrt{\left\| \frac{\partial \mathbf{x}}{\partial u_1} du_1 \right\|^2 \left\| \frac{\partial \mathbf{x}}{\partial u_2} du_2 \right\|^2 - \left\langle \frac{\partial \mathbf{x}}{\partial u_1} du_1, \frac{\partial \mathbf{x}}{\partial u_2} du_2 \right\rangle^2} \\ &= \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)_{2 \times 2}} du_1 du_2 \\ &= \sqrt{EG - F^2} du_1 du_2, \end{aligned}$$

其中

$$E = \left\langle \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_1} \right\rangle = \left\| \frac{\partial \mathbf{x}}{\partial u_1} \right\|^2, \quad F = \left\langle \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2} \right\rangle, \quad G = \left\langle \frac{\partial \mathbf{x}}{\partial u_2}, \frac{\partial \mathbf{x}}{\partial u_2} \right\rangle = \left\| \frac{\partial \mathbf{x}}{\partial u_2} \right\|^2.$$

我们称 $d\sigma = \sqrt{EG - F^2} du_1 du_2$ 为曲面 Σ 的面积微元, 定义曲面 Σ 的面积为

$$\sigma(\Sigma) = \int_{\Sigma} d\sigma = \int_D \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)_{2 \times 2}} du_1 du_2.$$

参数变换下的曲面面积微元

如果 $u_1 = u_1(t_1, t_2), u_2 = u_2(t_1, t_2)$ ($(t_1, t_2) \in \Omega$) 是正则的参数变换, 即它是 Ω 到 D 的微分同胚, 则 $\tilde{\mathbf{x}}(t_1, t_2) = \mathbf{x}(u_1(t_1, t_2), u_2(t_1, t_2))$ 满足

$$\begin{aligned} & \sqrt{\det \left(\left\langle \frac{\partial \tilde{\mathbf{x}}}{\partial t_i}, \frac{\partial \tilde{\mathbf{x}}}{\partial t_j} \right\rangle \right)_{2 \times 2}} dt_1 dt_2 \\ &= \sqrt{\det \left(\left\langle \sum_{k=1}^2 \frac{\partial \mathbf{x}}{\partial u_k} \frac{\partial u_k}{\partial t_i}, \sum_{l=1}^2 \frac{\partial \mathbf{x}}{\partial u_l} \frac{\partial u_l}{\partial t_j} \right\rangle \right)_{2 \times 2}} dt_1 dt_2 \\ &= \sqrt{\det \left(\sum_{k=1}^2 \sum_{l=1}^2 \frac{\partial u_k}{\partial t_i} \left\langle \frac{\partial \mathbf{x}}{\partial u_k}, \frac{\partial \mathbf{x}}{\partial u_l} \right\rangle \frac{\partial u_l}{\partial t_j} \right)_{2 \times 2}} dt_1 dt_2 \\ &= \sqrt{\det \left(\frac{\partial(u_1, u_2)}{\partial(t_1, t_2)} \right)^T \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_k}, \frac{\partial \mathbf{x}}{\partial u_l} \right\rangle \right)_{2 \times 2} \frac{\partial(u_1, u_2)}{\partial(t_1, t_2)}} dt_1 dt_2 \\ &= \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_k}, \frac{\partial \mathbf{x}}{\partial u_l} \right\rangle \right)_{2 \times 2}} \left| \det \frac{\partial(u_1, u_2)}{\partial(t_1, t_2)} \right| dt_1 dt_2 \\ &= \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_k}, \frac{\partial \mathbf{x}}{\partial u_l} \right\rangle \right)_{2 \times 2}} du_1 du_2, \end{aligned}$$

所以 $d\sigma$ 与曲面 Σ 的 C^1 正则表示选择无关。

例 6.2.1. 设 $D \subset \mathbb{R}^2$ 是一个有界闭区域, $f: D \rightarrow \mathbb{R}$ 是 C^1 函数, 求 f 的图像 $\Sigma = \{(x, y, f(x, y)) | (x, y) \in D\}$ 的面积。

解: 曲面 Σ 的两个切向量

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ f_x(x, y) \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ f_y(x, y) \end{pmatrix},$$

因此

$$E = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 1 + (f_x)^2, \quad F = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = f_x f_y, \quad G = \langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 1 + (f_y)^2,$$

面积微元

$$\begin{aligned} d\sigma &= \sqrt{EG - F^2} dx dy = \sqrt{1 + (f_x(x, y))^2 + (f_y(x, y))^2} dx dy \\ &= \sqrt{1 + \|\nabla f(x, y)\|^2} dx dy, \end{aligned}$$

曲面面积为 $\sigma(\Sigma) = \int_D \sqrt{1 + \|\nabla f(x, y)\|^2} dx dy$ 。

根据梯度的定义我们不难知道, $\|\nabla f(x, y)\|$ 的大小是曲面 $z = f(x, y)$ 的切平面与 xy 坐标平面的夹角正切值。 \square

例 6.2.2. 求直角坐标系中函数 $y = f(x)$ ($a \leq x \leq b, f(x) > 0$) 的图像绕 x 轴旋转形成的旋转体的侧面积。

解：曲面参数方程为

$$\begin{cases} x = x, \\ y = f(x) \cos \theta, \\ z = f(x) \sin \theta, \end{cases} \quad a \leq x \leq b, \quad 0 \leq \theta \leq 2\pi.$$

它的两个切向量

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ f'(x) \cos \theta \\ f'(x) \sin \theta \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -f(x) \sin \theta \\ f(x) \cos \theta \end{pmatrix},$$

因此

$$E = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 1 + f'(x)^2, \quad F = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0, \quad G = \langle \mathbf{v}_2, \mathbf{v}_2 \rangle = f(x)^2,$$

面积微元

$$d\sigma = \sqrt{EG - F^2} dx d\theta = f(x) \sqrt{1 + f'(x)^2} dx d\theta,$$

曲面面积为 $\sigma(\Sigma) = 2\pi \int_{[a,b]} f(x) \sqrt{1 + f'(x)^2} dx = \int_{\gamma} 2\pi y dl$. \square

例 6.2.3. 设 Σ 为球面 $x^2 + y^2 + z^2 = R^2$ 上纬度 θ 介于 θ_1, θ_2 之间, 经度 φ 介于 φ_1, φ_2 之间的部分. 求 Σ 的面积.

解：曲面参数方程为

$$\begin{cases} x = R \sin \theta \cos \varphi, \\ y = R \sin \theta \sin \varphi, \\ z = R \cos \theta, \end{cases} \quad 0 \leq \theta \leq \theta_0, \quad 0 \leq \varphi \leq 2\pi.$$

它的两个切向量

$$\mathbf{x}_\varphi = \begin{pmatrix} -R \sin \theta \sin \varphi \\ R \sin \theta \cos \varphi \\ 0 \end{pmatrix}, \quad \mathbf{x}_\theta = \begin{pmatrix} R \cos \theta \cos \varphi \\ R \cos \theta \sin \varphi \\ -R \sin \theta \end{pmatrix},$$

因此

$$E = \langle \mathbf{x}_\varphi, \mathbf{x}_\varphi \rangle = R^2 \sin^2 \theta, \quad F = \langle \mathbf{x}_\varphi, \mathbf{x}_\theta \rangle = 0, \quad G = \langle \mathbf{x}_\theta, \mathbf{x}_\theta \rangle = R^2,$$

面积微元

$$d\sigma = \sqrt{EG - F^2} d\varphi d\theta = R^2 \sin \theta d\varphi d\theta,$$

(你能给这个面积微元的结果一个直接的几何解释吗?) 于是, 曲面面积为

$$\sigma(\Sigma) = \int_{\varphi_1}^{\varphi_2} \left(\int_{\theta_1}^{\theta_2} R^2 \sin \theta d\varphi \right) d\theta = R^2 (\cos \theta_1 - \cos \theta_2) (\varphi_2 - \varphi_1).$$

□

曲面上函数的积分

设 Σ 是一个 C^1 正则曲面, $\mathbf{x}: D \rightarrow \mathbb{R}^m$ 是它的正则参数表示. $f: \Sigma \rightarrow \mathbb{R}$ 是一个连续函数, 则记

$$\int_{\Sigma} f d\sigma = \int_D f(\mathbf{x}) \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)_{2 \times 2}} du_1 du_2$$

称它为 f 在曲面 Σ 上的积分. 不难证明这个积分的值与曲面 Σ 的正则参数表示的选择无关.

利用这个概念我们可以计算给定质量分布(面密度)的曲面的质量和质心等物理问题.

习题6.2

1. 设 Σ 为球面 $x^2 + y^2 + z^2 = R^2$ 上纬度 θ 小于 θ_0 的部分. 求 Σ 的形心坐标.
2. 设 γ 为平面曲线, 它在极坐标系下的方程为 $r = a(1 + \cos \theta)$ ($a > 0$). 求 γ 绕 x 轴旋转所得曲面的面积和形心坐标.
3. 求轮胎的表面积和体积.
4. 求由圆柱面 $x^2 + y^2 = 1$, $y^2 + z^2 = 1$ 和 $z^2 + x^2 = 1$ 围成的有界区域的表面积.
5. 求区域 $\Omega = \{(x, y, z) | 0 \leq z \leq \sqrt{4 - x^2 - y^2}, (x - 1)^2 + y^2 \leq 1\}$ 的表面积.
6. 求质量均匀分布的球壳 $x^2 + y^2 + z^2 = R^2$ 对不在球壳上的质点 m 的万有引力.
7. 设 γ 是平面上的有界 C^1 曲线, $f: \gamma \rightarrow \mathbb{R}$ 是连续函数. 记

$$\Sigma = \{(x, y, f(x, y)) | (x, y) \in \gamma\}.$$

求 Σ 的面积.

6.3 \mathbb{R}^m 中任意维数曲面上的函数的积分

设 Σ 是一个 k 维的 C^1 正则曲面, $\mathbf{x}: D \rightarrow \mathbb{R}^m$ 是它的正则参数表示, 其中 $D \subseteq \mathbb{R}^k$ 是一个区域, $\frac{\partial \mathbf{x}}{\partial u_1}, \dots, \frac{\partial \mathbf{x}}{\partial u_k}$ 是线性无关的切向量. 则 Σ 的 (k 维) 体积微元为

$$d\sigma = \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)_{k \times k}} du_1 du_2 \cdots du_k.$$

这个结果在 $k = 1$ 和 $k = 2$ 时我们都曾经得到过, 一般结论留给读者验证。

设 $f: \Sigma \rightarrow \mathbb{R}$ 是一个连续函数, 则记

$$\int_{\Sigma} f d\sigma = \int_D f(\mathbf{x}) \sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)_{k \times k}} du_1 du_2 \cdots du_k$$

称它为 f 在曲面 Σ 上的积分。当 $f = 1$ 时, 得到曲面 Σ 的面积 $\int_{\Sigma} d\sigma$ 。

不难证明上述积分的值与曲面 Σ 的正则参数表示的选择无关。

例 6.3.1. 设 Σ 为 m 元函数 f 的图像 $\{(x_1, \dots, x_m, f(x_1, \dots, x_m)) | (x_1, \dots, x_m) \in D\}$, 则

$$d\sigma = \sqrt{1 + \|\nabla f(\mathbf{x})\|^2} dx_1 \cdots dx_m.$$

证明. 曲面 Σ 的 m 个线性无关切向量分别为

$$\mathbf{v}_k = \mathbf{e}_k + \frac{\partial f}{\partial x_k}(\mathbf{x}) \mathbf{e}_{m+1}, \quad k = 1, 2, \dots, m.$$

其中 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \mathbf{e}_{m+1}$ 是 \mathbb{R}^m 的标准基。于是

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j} + \frac{\partial f}{\partial x_i}(\mathbf{x}) \frac{\partial f}{\partial x_j}(\mathbf{x}), \quad i, j = 1, 2, \dots, m.$$

因此矩阵 $(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{m \times m} = I_m + \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^T$ 。

任取 \mathbb{R}^m 中与 $\nabla f(\mathbf{x})$ 垂直的向量 \mathbf{v} , 则

$$(I + \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^T) \mathbf{v} = \mathbf{v} + \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^T \mathbf{v} = \mathbf{v}.$$

取 $\nabla f(\mathbf{x})$ 的单位向量 \mathbf{u} , 则

$$(I + \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^T) \mathbf{u} = \mathbf{u} + \|\nabla f(\mathbf{x})\|^2 \mathbf{u} \mathbf{u}^T \mathbf{u} = (1 + \|\nabla f(\mathbf{x})\|^2) \mathbf{u}.$$

这样就得到矩阵 $(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{m \times m}$ 的全部特征值: 1 ($m - 1$ 重) 和 $1 + \|\nabla f(\mathbf{x})\|^2$ 。所以

$$d\sigma = \sqrt{\det(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)} dx_1 \cdots dx_m = \sqrt{1 + \|\nabla f(\mathbf{x})\|^2} dx_1 \cdots dx_m.$$

□

下面我们用这个结果来解决一个具体问题。

例 6.3.2. 求 \mathbb{R}^{m+1} 中 m 维球面 $\Sigma: x_1^2 + \cdots + x_m^2 + x_{m+1}^2 = R^2$ 的表面积。

解: 记 $\hat{\mathbf{x}} = (x_1, \dots, x_m)^T$, $\Sigma_{\pm}: x_{m+1} = \pm \sqrt{R^2 - \|\hat{\mathbf{x}}\|^2}$ ($\|\hat{\mathbf{x}}\| \leq R$)。则 $\Sigma = \Sigma_+ \cup \Sigma_-$ 。

记 $f(\hat{\mathbf{x}}) = \sqrt{R^2 - \|\hat{\mathbf{x}}\|^2}$, 由上面的例子知

$$d\sigma = \sqrt{1 + \|\nabla f(\hat{\mathbf{x}})\|^2} dx_1 dx_2 \cdots dx_m = \frac{R}{\sqrt{R^2 - \|\hat{\mathbf{x}}\|^2}} dx_1 dx_2 \cdots dx_m.$$

所以

$$\int_{\Sigma_+} d\sigma = \int_{\|\hat{\mathbf{x}}\| \leq R} \frac{R}{\sqrt{R^2 - \|\hat{\mathbf{x}}\|^2}} dx_1 dx_2 \cdots dx_m.$$

因此

$$\begin{aligned} A_m(R) &= \int_{\Sigma_+} d\sigma + \int_{\Sigma_-} d\sigma = 2 \int_{\|\hat{\mathbf{x}}\| \leq R} \frac{R}{\sqrt{R^2 - \|\hat{\mathbf{x}}\|^2}} dx_1 dx_2 \cdots dx_m \\ &= 2 \int_{-R}^R \left(\int_{\|\hat{\mathbf{u}}\| \leq \sqrt{R^2 - x_m^2}} \frac{R}{\sqrt{R^2 - x_m^2 - \|\hat{\mathbf{u}}\|^2}} dx_1 dx_2 \cdots dx_{m-1} \right) dx_m \\ &= 2 \int_0^R \frac{R}{\sqrt{R^2 - x_m^2}} A_{m-1}(\sqrt{R^2 - x_m^2}) dx_m. \end{aligned}$$

已知 m 维球体的体积

$$V_m(R) = 2 \int_0^R V_{m-1}(\sqrt{R^2 - x_m^2}) dx_m.$$

对 R 求导得到

$$V'_m(R) = 2 \int_0^R \frac{R}{\sqrt{R^2 - x_m^2}} V'_{m-1}(\sqrt{R^2 - x_m^2}) dx_m.$$

所以 $A_m(R)$ 与 $V'_m(R)$ 满足相同的递推关系。

而 $V_2(R) = \pi R^2$, $V'_2(R) = 2\pi R = A_1(R)$, 所以

$$A_m(R) = V'_{m+1}(R) = \frac{(m+1)\pi^{\frac{m+1}{2}}}{\Gamma\left(\frac{m+3}{2}\right)} R^m.$$

□

例 6.3.3. 设 $\mathbf{x} = \varphi(\mathbf{u})$ 是 m 维有界区域 Ω 到 m 维有界区域 Ω 的微分同胚。则

$$\sqrt{\det \left(\left\langle \frac{\partial \mathbf{x}}{\partial u_i}, \frac{\partial \mathbf{x}}{\partial u_j} \right\rangle \right)} = \sqrt{\det (D\varphi(\mathbf{u})^T D\varphi(\mathbf{u}))} = |\det D\varphi(\mathbf{u})|,$$

因此 $d\sigma = |\det D\varphi(\mathbf{u})| du_1 \cdots du_m$, 这就是重积分时的积分换元公式。所以曲线曲面上函数的积分是线性空间上 Riemann 积分的推广。

习题6.3

1. 证明内积空间中 k 个向量 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 所生成的平行多面体的体积为

$$\sqrt{\det (\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{k \times k}}.$$

2. 求质量均匀分布的半径为 R 的 m 维半球壳的质心坐标。

3. 设 $f: \mathbb{R}^m \rightarrow \mathbb{R}$ 是 C^1 函数, 满足对某个 $t_0 \in \mathbb{R}$,

$$\Omega(t_0) = \{\mathbf{x} \in \mathbb{R}^m | f(\mathbf{x}) \leq t_0\}$$

是一个有界闭区域, 且对任意 $\mathbf{x} \in \partial\Omega(t_0)$, $\nabla f(\mathbf{x}) \neq \mathbf{0}$ 。证明

(a) 对 t_0 的某个邻域里的任意 t , $\Omega(t)$ 都是有界闭区域, 且对任意 $\mathbf{x} \in \partial\Omega(t)$, $\nabla f(\mathbf{x}) \neq \mathbf{0}$ 。

(b) 记 $V(t)$ 为 $\Omega(t)$ 的体积, 则

$$\frac{dV(t)}{dt} = \int_{\partial\Omega(t)} \frac{1}{\|\nabla f(\mathbf{x})\|} d\sigma.$$

(c) 考虑 $f(x, y, z) = (\sqrt{x^2 + y^2} - R)^2 + z^2$ 。求 $\Omega(r^2)$ 的体积以及 $\partial\Omega(r^2)$ 的表面积, 其中 $0 < r < R$ 。

6.4 附：行列式与体积

对 \mathbb{R}^m 中的 k 个向量 $\mathbf{v}_1, \dots, \mathbf{v}_k$, 记

$$\Gamma(\mathbf{v}_1, \dots, \mathbf{v}_k) = (\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{k \times k} = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_k \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_2, \mathbf{v}_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{v}_k, \mathbf{v}_1 \rangle & \langle \mathbf{v}_k, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle \end{pmatrix}$$

为 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 的度量矩阵或者 **Gram** 矩阵, 称它的行列式为 **Gram** 行列式。

定理 6.4.1. $\sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_k)}$ 是由 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 张成的平行多面体的 k 维体积。

证明. 对 k 做数学归纳法。

当 $k = 1$ 时, $\sqrt{\Gamma(\mathbf{v}_1)} = \sqrt{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \|\mathbf{v}_1\|$, 结论成立。

假设 k 时结论成立。

设

$$\mathbf{v}_{k+1} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k + \mathbf{v}_{k+1}^\perp,$$

其中 \mathbf{v}_{k+1}^\perp 与 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 中每个向量都正交。取

$$C = \begin{pmatrix} I & -\mathbf{a} \\ 0 & 1 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix},$$

则

$$\begin{aligned}
 & C^T \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_k \rangle & \langle \mathbf{v}_1, \mathbf{v}_{k+1} \rangle \\ \vdots & \ddots & \vdots & \vdots \\ \langle \mathbf{v}_k, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle & \langle \mathbf{v}_k, \mathbf{v}_{k+1} \rangle \\ \langle \mathbf{v}_{k+1}, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_{k+1}, \mathbf{v}_k \rangle & \langle \mathbf{v}_{k+1}, \mathbf{v}_{k+1} \rangle \end{pmatrix} C \\
 &= \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_k \rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle \mathbf{v}_k, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle & 0 \\ 0 & \cdots & 0 & \langle \mathbf{v}_{k+1}^\perp, \mathbf{v}_{k+1}^\perp \rangle \end{pmatrix}
 \end{aligned}$$

所以

$$\sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_{k+1})} = \sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_k)} \|\mathbf{v}_{k+1}^\perp\|.$$

根据归纳假设, $\sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_k)}$ 是由 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 张成的平行多面体 P_k 的 k 维体积, 而 $\|\mathbf{v}_{k+1}^\perp\|$ 是以 P_k 为底面时的平行多面体 P_{k+1} 的高, 所以

$$\sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_{k+1})}$$

是由 $\mathbf{v}_1, \dots, \mathbf{v}_{k+1}$ 张成的平行多面体 P_{k+1} 的 $k+1$ 维体积. \square

推论 6.4.2. 设 A 是可逆方阵. 则 $|\det A|$ 是 A 的所有列向量 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 张成的平行多面体的 m 维体积.

证明. 因为

$$A^T A = \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_m),$$

所以

$$|\det A| = \sqrt{(\det A)^2} = \sqrt{\det(A^T) \det A} = \sqrt{\det(A^T A)} = \sqrt{\det \Gamma(\mathbf{v}_1, \dots, \mathbf{v}_m)}.$$

\square

定义 6.4.3. 称 \mathbb{R}^m 的一组基底向量 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 是正向的(相应地, 负向的), 如果依次以 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 为列向量的矩阵 A 满足 $\det A > 0$ (相应地, $\det A < 0$). 称 $\det A$ 为依次以 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 为棱的平行多面体的有向体积.

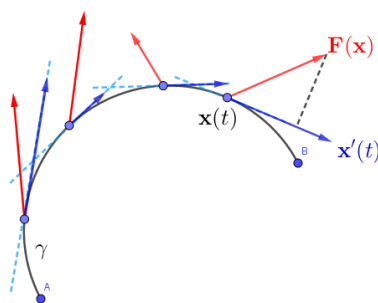
6.5 习题讨论课7

1. 设 γ 是平面中一条轴对称的简单封闭曲线, D 是 γ 所围成的平面有界区域, Σ 是 γ 绕旋转轴旋转一周形成的曲面, Ω 是 Σ 所围成的空间有界区域。试讨论曲线 γ 、区域 D 、曲面 Σ 、区域 Ω 的形心坐标的关系。

第 7 章 向量场的微积分

7.1 第二型曲线积分

例 7.1.1 (变力做功). 考虑一个力对一个运动质点所做的功, 质点的运动由一个路径 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^m$ 来刻画, 这时 $t \in [a, b]$ 是时间, $\mathbf{x}(t)$ 是质点在时刻 t 所在的位置。



$\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ 是一个力场, 它在空间 \mathbb{R}^m 的每个点 \mathbf{x} 处给出一个力 $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^m$. 力是向量, 所以我们也称 \mathbf{F} 是一个向量场. 为了区别位置和向量, 我们分别使用了 \mathbb{R}^m 和 \mathbf{R}^m .

力场 \mathbf{F} 对质点做的功为

$$W = \int_a^b \langle \mathbf{F}(\mathbf{x}), \mathbf{x}'(t) \rangle dt.$$

这称为向量场 \mathbf{F} 沿路径 $\mathbf{x}(t)$ 的积分。 □

例 7.1.2 (流速场环量). 如果 γ 是 \mathbb{R}^m 中一条 C^1 曲线, $\mathbf{T} : \gamma \rightarrow \mathbf{R}^m$ 是 γ 的一个连续单位切向量场, 则称

$$\int_{\gamma} \langle \mathbf{F}(\mathbf{x}), \mathbf{T}(\mathbf{x}) \rangle dl$$

为向量场 \mathbf{F} 沿曲线 γ 的单位切向量场 \mathbf{T} 的环量。

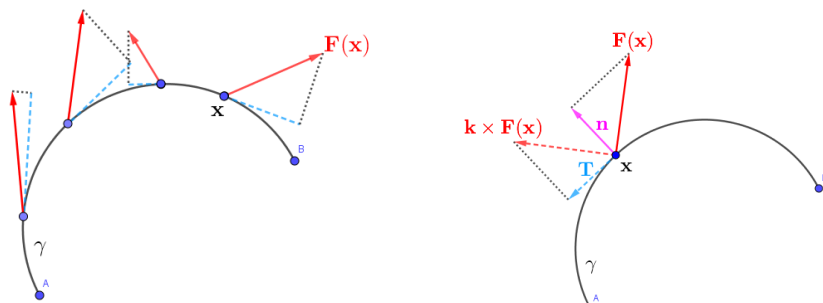
当 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^m$ 是 γ 的一个 C^1 正则参数表示, 且 $\mathbf{T}(\mathbf{x}(t)) = \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$ 时,

$$\int_{\gamma} \langle \mathbf{F}(\mathbf{x}), \mathbf{T}(\mathbf{x}) \rangle dl = \int_{[a,b]} \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{T}(\mathbf{x}(t)) \rangle \|\mathbf{x}'(t)\| dt = \int_a^b \langle \mathbf{F}(\mathbf{x}), \mathbf{x}'(t) \rangle dt.$$

如果取 $\mathbf{F}(\mathbf{x}) = f(\mathbf{x})\mathbf{T}(\mathbf{x})$, 则

$$\int_{\gamma} \langle \mathbf{F}(\mathbf{x}), \mathbf{T}(\mathbf{x}) \rangle dl = \int_{\gamma} f(\mathbf{x}) dl,$$

所以函数 $f(\mathbf{x})$ 沿曲线 γ 的第一型曲线积分可以转化为向量场 $f(\mathbf{x})\mathbf{T}$ 沿曲线 γ 的第二型曲线积分。 \square



例 7.1.3 (平面流速场通量). 如果 γ 是平面上一条 C^1 的简单封闭曲线, $\mathbf{n} : \gamma \rightarrow \mathbb{R}^2$ 是 γ 的单位法向量场, 我们称

$$\int_{\gamma} \langle \mathbf{F}, \mathbf{n} \rangle dl = \int_{\gamma} \langle \mathbf{F}(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle dl$$

为向量场 \mathbf{F} 沿曲线 γ 的法方向 \mathbf{n} 的通量。

当 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^2$ 是 γ 的一个 C^1 正则参数表示, 且 $\mathbf{T}(\mathbf{x}(t)) = \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$ 满足 $\mathbf{k} \times \mathbf{n} = \mathbf{T}$ (这里 \mathbf{k} 是平面的单位法向量场, $\mathbf{k} \times \mathbf{n} = \mathbf{T}$ 意味着 \mathbf{n} 逆时针旋转 90° 得到 \mathbf{T}) 时,

$$\int_{\gamma} \langle \mathbf{F}, \mathbf{n} \rangle dl = \int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{T}(\mathbf{x}(t)) \times \mathbf{k} \rangle \|\mathbf{x}'(t)\| dt = \int_a^b \langle \mathbf{k} \times \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt.$$

\square

定义 7.1.4. 向量场 $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ 沿分段 C^1 的连续路径 $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^m$ 的积分为

$$\int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt.$$

与路径定向的关系

上述积分与曲线的同向正则参数表示的选择无关。设可微函数 $t: [\alpha, \beta] \rightarrow [a, b]$ 满足 $t(\alpha) = a$, $t(\beta) = b$, $\tilde{\mathbf{x}}(s) = \mathbf{x}(t(s))$, 则

$$\int_{\alpha}^{\beta} \langle \mathbf{F}(\tilde{\mathbf{x}}(s)), \tilde{\mathbf{x}}'(s) \rangle ds = \int_{\alpha}^{\beta} \langle \mathbf{F}(\mathbf{x}(t(s))), \mathbf{x}'(t(s))t'(s) \rangle ds = \int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt.$$

另外, 易见对曲线的反向参数表示, 即可微函数 $t: [\alpha, \beta] \rightarrow [a, b]$ 满足 $t(\alpha) = b$, $t(\beta) = a$, $\tilde{\mathbf{x}}(s) = \mathbf{x}(t(s))$, 则积分结果为相反数,

$$\int_{\alpha}^{\beta} \langle \mathbf{F}(\tilde{\mathbf{x}}(s)), \tilde{\mathbf{x}}'(s) \rangle ds = - \int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt.$$

另外, 需要特别指出的是这里并未要求 $t(s)$ 是单调函数。这也是沿路径积分与沿曲线积分的不同。

一阶微分形式

在直角坐标系中, $\mathbf{F} = (F_1, \dots, F_m)^T$, $\mathbf{x} = (x_1, \dots, x_m)^T$, 则

$$\int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt = \int_a^b F_1(\mathbf{x}(t)) dx_1(t) + \dots + F_m(\mathbf{x}(t)) dx_m(t).$$

称

$$\omega = F_1(\mathbf{x}) dx_1 + \dots + F_m(\mathbf{x}) dx_m$$

为一个一阶微分形式。

一个一阶微分形式 ω 是空间中的一个线性函数场, 对空间中的切向量场 \mathbf{G} , 设其坐标为 $(G_1(\mathbf{x}), \dots, G_m(\mathbf{x}))^T$, 则 $dx_k(\mathbf{G}) = G_k(\mathbf{x})$, 因此

$$\omega(\mathbf{G})(\mathbf{x}) = F_1(\mathbf{x})G_1(\mathbf{x}) + \dots + F_m(\mathbf{x})G_m(\mathbf{x}).$$

对一条 C^1 曲线 $\gamma: \mathbf{x}(t) = (x_1(t), \dots, x_m(t))^T$, 其速度向量场

$$\mathbf{x}'(t) = (x_1'(t), \dots, x_m'(t))^T$$

是曲线的一个切向量场,

$$\omega(\gamma) = F_1(\mathbf{x}(t))x_1'(t) + \dots + F_m(\mathbf{x}(t))x_m'(t).$$

从而

$$\int_{\gamma} \omega = \int_{\gamma} F_1 dx_1 + \dots + F_m dx_m = \int_a^b F_1(\mathbf{x}(t))x_1'(t) + \dots + F_m(\mathbf{x}(t))x_m'(t) dt.$$

定理 7.1.5. 1. $\int_{\gamma} \langle \mathbf{F}, \mathbf{T} \rangle dl$ 和 $\int_{\gamma} \omega$ 关于向量场 \mathbf{F} 和一阶微分形式 ω 都是线性的。

$$\begin{aligned} \int_{\gamma} \langle \alpha \mathbf{F}_1 + \beta \mathbf{F}_2, \mathbf{T} \rangle dl &= \alpha \int_{\gamma} \langle \mathbf{F}_1, \mathbf{T} \rangle dl + \beta \int_{\gamma} \langle \mathbf{F}_2, \mathbf{T} \rangle dl, \\ \int_{\gamma} \alpha \omega_1 + \beta \omega_2 &= \alpha \int_{\gamma} \omega_1 + \beta \int_{\gamma} \omega_2. \end{aligned}$$

2. 如果 γ_1 和 γ_2 是收尾相连的两个分段 C^1 的连续路径(即 γ_1 的终点就是 γ_2 的起点), 则

$$\int_{\gamma_1+\gamma_2} \langle \mathbf{F}, \mathbf{T} \rangle dl = \int_{\gamma_1} \langle \mathbf{F}, \mathbf{T} \rangle dl + \int_{\gamma_2} \langle \mathbf{F}, \mathbf{T} \rangle dl,$$

$$\int_{\gamma_1+\gamma_2} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega,$$

其中 $\gamma_1 + \gamma_2$ 是连接 γ_1 和 γ_2 而成的分段 C^1 的连续路径。

例 7.1.6. 设平面简单封闭曲线 γ 在极坐标系下的方程为 $r = r(\theta)$ ($0 \leq \theta \leq 2\pi$), $r(2\pi) = r(0)$ 。证明

$$\int_{\gamma} x dy = - \int_{\gamma} y dx = \frac{1}{2} \int_{\gamma} -y dx + x dy$$

且积分值为 γ 所围的有界区域 D 的面积。

证明.

$$\int_{\gamma} y dx + x dy = \int_0^{2\pi} d(x(\theta)y(\theta)) = x(2\pi)y(2\pi) - x(0)y(0) = 0.$$

所以 $\int_{\gamma} x dy = - \int_{\gamma} y dx$ 。另一方面,

$$\begin{aligned} \frac{1}{2} \int_{\gamma} -y dx + x dy &= \frac{1}{2} \int_0^{2\pi} -r(\theta) \sin \theta d(r(\theta) \cos \theta) + r(\theta) \cos \theta d(r(\theta) \sin \theta) \\ &= \frac{1}{2} \int_0^{2\pi} r(\theta)^2 d\theta = \int_0^{2\pi} \int_0^{r(\theta)} \rho d\rho d\theta \\ &= \int_D \rho d\rho d\theta = \int_D dx dy, \end{aligned}$$

所以 $\int_{\gamma} x dy = - \int_{\gamma} y dx = \frac{1}{2} \int_{\gamma} -y dx + x dy$ 且为由曲线 γ 区域 D 的面积。 \square

例 7.1.7. 求 $\int_{\gamma} z dx + x dy + y dz$, 其中 γ 为曲线

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

按右手法则绕 z 轴正半轴旋转一周。

解: 把 $z = -(x + y)$ 代入 $x^2 + y^2 + z^2 = R^2$ 后, 配方整理得到

$$\left(x + \frac{y}{2}\right)^2 + \frac{3y^2}{4} = \frac{R^2}{2}.$$

所以 γ 在坐标平面 $z = 0$ 中的投影是一个绕原点逆时针旋转的椭圆。因为积分与路径参数表达无关, 所以令 $y = \frac{R \sin \theta}{\sqrt{2}} \sqrt{\frac{4}{3}}$, $x + \frac{y}{2} = \frac{R \cos \theta}{\sqrt{2}}$, $0 \leq \theta \leq 2\pi$, 得

到参数方程

$$\begin{cases} x = \frac{R}{\sqrt{2}} \cos \theta - \frac{R}{\sqrt{6}} \sin \theta, \\ y = \frac{2R}{\sqrt{6}} \sin \theta, \\ z = -\frac{R}{\sqrt{2}} \cos \theta - \frac{R}{\sqrt{6}} \sin \theta, \end{cases}, \quad 0 \leq \theta \leq 2\pi,$$

参数 θ 增加的方向与 γ 绕 z 轴旋转的方向一致。于是

$$\begin{aligned} & zdx + xdy + ydz \\ &= (-x - y)dx + xdy + yd(-x - y) = (-x - 2y)dx + (x - y)dy \\ &= \left(-\frac{R}{\sqrt{2}} \cos \theta - \frac{3R}{\sqrt{6}} \sin \theta\right) d\left(\frac{R}{\sqrt{2}} \cos \theta - \frac{R}{\sqrt{6}} \sin \theta\right) \\ &\quad + \left(\frac{R}{\sqrt{2}} \cos \theta - \frac{3R}{\sqrt{6}} \sin \theta\right) d\frac{2R}{\sqrt{6}} \sin \theta \\ &= \frac{\sqrt{3}R^2}{2} d\theta \end{aligned}$$

所以

$$\int_{\gamma} zdx + xdy + ydz = \int_0^{2\pi} \frac{\sqrt{3}R^2}{2} d\theta = \sqrt{3}\pi R^2.$$

□

另解：曲线 γ 可以表示为 $\begin{cases} z = -x - y, \\ x^2 + y^2 + (-x - y)^2 = 2x^2 + 2xy + 2y^2 = R^2. \end{cases}$

沿 γ ,

$$\begin{aligned} zdx + xdy + ydz &= (-x - y)dx + xdy + yd(-x - y) \\ &= d\left(-\frac{x^2}{2} - \frac{y^2}{2} + xy\right) - 3ydx, \end{aligned}$$

因为 γ 是闭曲线，所以 $\int_{\gamma} d\left(-\frac{x^2}{2} - \frac{y^2}{2} + xy\right) = 0$ ，因此

$$\int_{\gamma} zdx + xdy + ydz = \int_{\gamma} -3ydx = \int_{\Gamma} -3ydx,$$

这里 Γ 是 γ 在 xy 坐标平面中的投影曲线 $x^2 + xy + y^2 = \frac{R^2}{2}$ 。

根据上一例题的结论， $\int_{\Gamma} -ydx$ 是平面曲线 Γ 所围有界区域 $D = \{(x, y) | x^2 + xy + y^2 \leq \frac{R^2}{2}\}$ 的面积，所以

$$3 \int_{\Gamma} -ydx = 3 \iint_{x^2 + xy + y^2 \leq \frac{R^2}{2}} dx dy = \int_{-\frac{\sqrt{2}}{3}R}^{\frac{\sqrt{2}}{3}R} \int_{\frac{-x - \sqrt{2R^2 - 3x^2}}{2}}^{\frac{-x + \sqrt{2R^2 - 3x^2}}{2}} dy dx = \sqrt{3}\pi R^2.$$

□

习题7.1

1. 在 C^1 坐标变换 $\mathbf{y} = \Phi(\mathbf{x})$ 下, 路径 γ 变为

$$\tilde{\gamma} : \mathbf{y}(t) = \Phi(\mathbf{x}(t)),$$

向量场 $\mathbf{F}(\mathbf{x})$ 变为

$$\partial\Phi(\mathbf{x})\mathbf{F}(\mathbf{x}) = \partial\Phi(\Phi^{-1}(\mathbf{y}))\mathbf{F}(\Phi^{-1}(\mathbf{y})) = \tilde{\mathbf{F}}(\mathbf{y}),$$

试讨论 $\int_{\tilde{\gamma}} \langle \tilde{\mathbf{F}}, \tilde{\mathbf{T}} \rangle_{\mathbf{y}} dl$ 与 $\int_{\gamma} \langle \mathbf{F}, \mathbf{T} \rangle_{\mathbf{x}} dl$ 的关系。

7.2 有势场、保守场和无旋向量场

函数 f 的微分

$$df = \frac{\partial f}{\partial x_1}(\mathbf{x})dx_1 + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{x})dx_m$$

是一个一阶微分形式, 它沿任何路径 γ 的积分

$$\begin{aligned} \int_{\gamma} df &= \int_a^b \frac{\partial f}{\partial x_1}(\mathbf{x}(t))x'_1(t) + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{x}(t))x'_m(t) dt = \int_a^b \frac{d}{dt} f(\mathbf{x}(t)) dt \\ &= f(\mathbf{x}(b)) - f(\mathbf{x}(a)). \end{aligned}$$

此时, 积分结果只与路径的起点和终点有关, 与路径选择无关。

定义 7.2.1. 称一阶微分形式 ω 是一个全微分, 如果存在 C^1 函数 f 使得 $\omega = df$, 此时称 f 为 ω 的一个原函数。

称一个向量场 \mathbf{F} 为有势场, 如果存在 C^1 函数 f 使得 $\mathbf{F} = \nabla f$, 此时称 f 为 \mathbf{F} 的一个势函数。

称一个向量场 \mathbf{F} 为保守场, 如果对任何路径 γ , 积分 $\int_{\gamma} \langle \mathbf{F}, \mathbf{T} \rangle dl$ 的值都只与 γ 的起点和终点有关, 而与 γ 无关。

定理 7.2.2 (微积分基本定理). 对任何 C^1 函数 $f : U \rightarrow \mathbb{R}$ 以及任意 C^1 路径 $\gamma \subset U$,

$$\int_{\gamma} df = \int_{\gamma} \langle \nabla f, \mathbf{T} \rangle dl = f(B) - f(A),$$

其中 A, B 分别是路径 γ 的起点和终点, \mathbf{T} 是 γ 的单位切向量场。所以梯度向量场(也就是有势场)都是保守场。

如果 γ 是一条封闭曲线, 则 $\int_{\gamma} df = \int_{\gamma} \langle \nabla f, \mathbf{T} \rangle dl = 0$ 。

例 7.2.3. 力场 $\mathbf{F}(\mathbf{x}) = \frac{-\mathbf{x}}{\|\mathbf{x}\|^3}$ 是 $\mathbb{R}^m \setminus \{\mathbf{0}\}$ 中的一个有势场。求它的一个势函数, 以及 \mathbf{F} 对沿路径 γ 运动的单位质量的质点所做的功。

解:

$$F_1(\mathbf{x})dx_1 + \cdots + F_m(\mathbf{x})dx_m = -\sum_{k=1}^m \frac{x_k dx_k}{\|\mathbf{x}\|^3} = -\frac{d(\|\mathbf{x}\|^2)}{2\|\mathbf{x}\|^3} = d\frac{1}{\sqrt{\|\mathbf{x}\|^2}},$$

所以

$$\mathbf{F}(\mathbf{x}) = \nabla \left(\frac{1}{\|\mathbf{x}\|} \right).$$

所以 $\frac{1}{\|\mathbf{x}\|}$ 是 \mathbf{F} 的一个势函数。于是

$$\int_{\gamma} \langle \mathbf{F}, \mathbf{T} \rangle dl = \frac{1}{\|B\|} - \frac{1}{\|A\|},$$

其中 A, B 分别是 γ 的起点和终点。有势力场对运动质点做的功等于质点势能的该变量。 \square

例 7.2.4. 求 $\int_{\gamma} (e^y + \sin x)dx + (xe^y - \cos y)dy$, 其中 γ 是沿圆弧 $(x-\pi)^2 + y^2 = \pi^2$ 从 $(0,0)$ 逆时针到 (π, π) 。

解: 采用分组凑微分的办法,

$$\begin{aligned} (e^y + \sin x)dx + (xe^y - \cos y)dy &= e^y dx + xde^y - d\cos x - d\sin y \\ &= d(xe^y - \cos x - \sin y), \end{aligned}$$

这是全微分, 所以积分与路径无关, 因此利用原函数 $xe^y - \cos x - \sin y$ 直接求得

$$\begin{aligned} \int_{\gamma} (e^y + \sin x)dx + (xe^y - \cos y)dy &= (xe^y - \cos x - \sin y) \Big|_{(0,0)}^{(\pi,\pi)} \\ &= \pi e^{\pi} + 1 - (-1) = 2 + \pi e^{\pi}. \end{aligned}$$

\square

例 7.2.5. 求 $\int_{\gamma} (x^2 - yz)dx + (y^2 - zx)dy + (z^2 - xy)dz$, 其中 γ 是螺旋线 $x = a \cos t, y = a \sin t, z = bt$ ($0 \leq t \leq 2\pi$)。

解: 观察知 $\frac{x^3 + y^3 + z^3}{3} - xyz$ 是原函数,

$$\begin{aligned} &\int_{\gamma} (x^2 - yz)dx + (y^2 - zx)dy + (z^2 - xy)dz \\ &= \left(\frac{x^3 + y^3 + z^3}{3} - xyz \right) \Big|_{(a,0,0)}^{(a,0,2\pi b)} = \frac{8\pi^3 b^3}{3}. \end{aligned}$$

\square

例 7.2.6. 求 $\int_{\gamma} \frac{x dx + y dy}{\sqrt{x^2 + y^2 - 1}}$, 其中 γ 是连接 $(2,0)$ 到 $(0,3)$ 一条 C^1 路径, 且中途不与 $x^2 + y^2 = 1$ 相交。

解: 观察知 $\sqrt{x^2 + y^2 - 1}$ 是原函数,

$$\int_{\gamma} \frac{xdx + ydy}{\sqrt{x^2 + y^2 - 1}} = \sqrt{0^3 + 3^2 - 1} - \sqrt{2^2 + 0^2 - 1} = 2\sqrt{2} - \sqrt{3}.$$

□

定理 7.2.7. 开区域中的连续的保守场都是有势场。

证明. 设向量场 $\mathbf{F} : U \rightarrow \mathbf{R}^m$ 是开区域 $U \subseteq \mathbf{R}^m$ 中的一个连续的保守场。因为开区域是道路连通集, 所以任取 $A_0, A \in U$, 存在 C^1 曲线 γ 使得 $\mathbf{x} : [a, b] \rightarrow U$ 使得 $\mathbf{x}(0) = A_0$, $\mathbf{x}(b) = A$, 令

$$f(A) = \int_{\gamma} \langle \mathbf{F}, \mathbf{T} \rangle dl = \int_a^b \langle \mathbf{F}(\mathbf{x}(t)), \mathbf{x}'(t) \rangle dt.$$

因为 \mathbf{F} 是保守场, 所以 $f(A)$ 完全由 $A_0, A \in U$ 确定, 与 γ 的选择无关。

取 $\delta_0 > 0$ 使得以 A 为中心 δ_0 为半径的开球 $B(A, \delta_0) \subset U$ 。考虑分段 C^1 的连续曲线 $\tilde{\gamma}$:

$$\mathbf{y}(t) = \begin{cases} \mathbf{x}(t), & a \leq t \leq b; \\ A + (t - b)\mathbf{e}_k. & b \leq t \leq b + \delta_0. \end{cases}$$

由上述唯一性知

$$\begin{aligned} f(A + (s - b)\mathbf{e}_k) &= \int_{\tilde{\gamma}} \langle \mathbf{F}, \mathbf{T} \rangle dl \\ &= f(A) + \int_b^s \langle \mathbf{F}(\mathbf{y}(t)), \mathbf{y}'(t) \rangle dt \\ &= f(A) + \int_b^s \langle \mathbf{F}(A + (t - b)\mathbf{e}_k), \mathbf{e}_k \rangle dt \\ &= f(A) + \int_b^s F_k(A + (t - b)\mathbf{e}_k) dt. \end{aligned}$$

从而

$$\frac{\partial f}{\partial x_k}(A) = \left. \frac{d}{ds} f(A + (s - b)\mathbf{e}_k) \right|_{s=b} = F_k(\mathbf{x}(b)) = F_k(A),$$

所以 $f : U \rightarrow \mathbf{R}$ 是 C^1 函数, 且 $\nabla f = \mathbf{F}$ 。从而 \mathbf{F} 是有势场。 □

推论 7.2.8. 若 C^1 向量场 $\mathbf{F} : U \rightarrow \mathbf{R}^m$ 是一个有势场, 则

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}, \quad 1 \leq i < j \leq m.$$

满足后面这个条件的向量场称为无旋向量场。所以有势场都是无旋向量场。

证明. 设 f 是 \mathbf{F} 的势函数, 则 f 是 C^2 函数, 且 $\frac{\partial f}{\partial x_i} = F_i$, 从而

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial F_j}{\partial x_i}, \quad 1 \leq i < j \leq m.$$

因此 \mathbf{F} 是无旋向量场。 □

但无旋向量场未必都是有势场, 请看:

例 7.2.9. 考虑向量场

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right),$$

以及路径

$$\gamma_k : (x(\theta), y(\theta)) = (\cos \theta, \sin \theta), \quad 0 \leq \theta \leq 2k\pi.$$

则 γ_k 是闭路径, 而

$$\int_{\gamma_k} \langle \mathbf{F}, \mathbf{T} \rangle dl = \int_0^{2k\pi} \frac{-y(\theta)}{x^2(\theta) + y^2(\theta)} x'(\theta) + \frac{x(\theta)}{x^2(\theta) + y^2(\theta)} y'(\theta) d\theta = 2k\pi \neq 0,$$

所以 \mathbf{F} 不是保守场, 因而也不是有势场。

另一方面,

$$\frac{\partial}{\partial y} \left(\frac{-y}{x^2 + y^2} \right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right),$$

所以 \mathbf{F} 是无旋向量场。

虽然

$$\nabla \left(\arctan \frac{y}{x} \right) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) = \nabla \left(-\arctan \frac{x}{y} \right),$$

但是 $\arctan \frac{y}{x}$ 只在 $x \neq 0$ 时有定义, $-\arctan \frac{x}{y}$ 只在 $y \neq 0$ 时有定义。所以 \mathbf{F} 在 $x > 0$ 的半平面和 $x < 0$ 的半平面以及 $y > 0$ 的半平面和 $y < 0$ 的半平面中有势函数, 甚至

$$f(x, y) = \begin{cases} \arctan \frac{y}{x}, & x > 0; \\ \frac{\pi}{2}, & x = 0, y > 0; \\ \pi + \arctan \frac{y}{x}, & x < 0 \end{cases}$$

是 \mathbf{F} 在 $\mathbb{R}^2 \setminus \{(0, y) | y \leq 0\}$ 中的势函数, 但 \mathbf{F} 在 $\mathbb{R}^2 \setminus \{(0, 0)\}$ 中没有势函数。

另外, 我们也可以考虑极坐标。如果 γ 的极坐标形式为 $r = r(t), \theta = \theta(t)$, 则

$$\begin{aligned} \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy &= \frac{-\sin \theta(t)}{r(t)} d(r(t) \cos \theta(t)) + \frac{\cos \theta(t)}{r(t)} d(r(t) \sin \theta(t)) \\ &= d\theta(t), \end{aligned}$$

所以

$$\int_{\gamma} \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = \theta(b) - \theta(a),$$

即曲线 γ 绕原点旋转的角度(以逆时针方向为正)。

□

注 7.2.10. 在复平面 \mathbb{C} 中, $z = x + iy$, 于是

$$\frac{dz}{z} = \frac{dx + idy}{x + iy} = \frac{(x - iy)(dx + idy)}{x^2 + y^2} = d \ln \sqrt{x^2 + y^2} + i \frac{-y dx + x dy}{x^2 + y^2}.$$

例 7.2.11. 上例中向量场 \mathbf{F} 对应的一阶微分形式

$$\omega = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy,$$

也可以看成是 \mathbb{R}^3 中关于 x, y, z 的一个一阶微分形式。这时只要 γ 是一条不经过 z 轴的曲线 γ , 则

$$\int_{\gamma} \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

就是曲线绕 z 轴旋转的角度(以 xy 平面逆时针旋转方向为正)。 \square

\mathbb{R}^3 中的向量场

$$\mathbf{F}(x, y, z) = \left(\frac{-y}{x^2 + y^2} + \frac{z}{z^2 + x^2}, \frac{x}{x^2 + y^2} + \frac{-z}{y^2 + z^2}, \frac{-x}{z^2 + x^2} + \frac{y}{y^2 + z^2} \right)$$

是一个无旋向量场。

例 7.2.12. 沿上半椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 (y \geq 0)$ 从 $A(a, 0)$ 到 $B(-a, 0)$ 计算

$$\int_{\gamma} (1 + ye^x) dx + (x + e^x) dy.$$

解:

$$(1 + ye^x) dx + (x + e^x) dy = x dy + d(x + ye^x),$$

$$\int_{\gamma} d(x + ye^x) = (x + ye^x)|_A^B = -2a,$$

取 $x = a \cos \theta, y = b \sin \theta (0 \leq \theta \leq \pi)$, 所以

$$\int_{\gamma} x dy = \int_0^{\pi} a \cos \theta db \sin \theta = ab \int_0^{\pi} \cos^2 \theta d\theta = \frac{\pi}{2} ab,$$

因此

$$\int_{\gamma} (1 + ye^x) dx + (x + e^x) dy = \frac{\pi}{2} ab - 2a.$$

\square

7.3 平面向量场的旋度与散度, Green 公式

定义 7.3.1. 对平面 C^1 向量场

$$\mathbf{F}(x, y) = \begin{pmatrix} X(x, y) \\ Y(x, y) \end{pmatrix},$$

记

$$\begin{aligned}\operatorname{rot} \mathbf{F} &= \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}, \\ \operatorname{div} \mathbf{F} &= \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = \operatorname{tr} \frac{\partial(X, Y)}{\partial(x, y)},\end{aligned}$$

分别称它们为 \mathbf{F} 的旋度和散度。

旋度为零的向量场称为无旋向量场, 散度为零的向量场称为无源向量场。

例 7.3.2. 设

$$\mathbf{F}_A(x, y) = A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}.$$

则

$$\begin{aligned}\operatorname{rot} \mathbf{F}_A &= a_{21} - a_{12}, \\ \operatorname{div} \mathbf{F}_A &= \operatorname{tr} A = a_{11} + a_{22}.\end{aligned}$$

因此, \mathbf{F}_A 是无旋向量场当且仅当 A 是对称阵, \mathbf{F}_A 是无源向量场当且仅当 A 的迹为零。

对任何二阶矩阵 A ,

$$A = A_s + A_a, \quad A_s = \frac{A + A^T}{2}, \quad A_a = \frac{A - A^T}{2}.$$

A_s 是对称阵对应无旋向量场 \mathbf{F}_{A_s} , A_s 有一对彼此正交的特征方向, 它们限制了向量场 \mathbf{F}_{A_s} 的积分曲线绕原点旋转; A_a 是反对称阵, 对角线为零, 是无源向量场 \mathbf{F}_{A_a} , \mathbf{F}_{A_a} 的积分曲线是以原点为中心的圆周, 这些圆周使向量场的积分曲线无法穿进或穿出。

$\mathbf{F}_A = \mathbf{F}_{A_s} + \mathbf{F}_{A_a}$, 并且

$$\operatorname{rot} \mathbf{F}_A = \operatorname{rot} \mathbf{F}_{A_a}, \quad \operatorname{div} \mathbf{F}_A = \operatorname{div} \mathbf{F}_{A_s}.$$

□

定理 7.3.3 (Green公式). 对平面 C^1 向量场

$$\mathbf{F}(x, y) = \begin{pmatrix} X(x, y) \\ Y(x, y) \end{pmatrix},$$

和平面区域 Ω ,

$$\int_{\partial\Omega} X(x, y)dx + Y(x, y)dy = \int_{\Omega} \left(\frac{\partial Y(x, y)}{\partial x} - \frac{\partial X(x, y)}{\partial y} \right) dx dy,$$

写成向量场形式

$$\int_{\partial\Omega} \langle \mathbf{F}, \mathbf{T} \rangle dl = \int_{\Omega} \operatorname{rot} \mathbf{F} \, dx dy, \quad \text{沿区域边界的环量等于区域内部旋度的积分}$$

$$\int_{\partial\Omega} \langle \mathbf{F}, \mathbf{n} \rangle dl = \int_{\Omega} \operatorname{div} \mathbf{F} \, dx dy, \quad \text{沿区域边界的通量等于区域内部散度的积分.}$$

其中 $\partial\Omega$ 为区域 Ω 的边界, 其定向为按区域外法向的左手法则, 即当你沿 $\partial\Omega$ 前进时区域总位于你的左侧。

证明. 我们先证明

$$\int_{\partial\Omega} X(x, y) dx = \int_{\Omega} \left(-\frac{\partial X(x, y)}{\partial y} \right) dx dy.$$

考虑形如

$$D = \{(x, y) | a \leq x \leq b, f(x) \leq y \leq g(x)\}$$

的平面区域。

$$\begin{aligned} \int_{\partial D} X(x, y) dx &= \int_{a \leq x \leq b, y=f(x)} X(x, y) dx + \int_{x=b, f(b) \leq y \leq g(b)} X(x, y) dx \\ &\quad - \int_{a \leq x \leq b, y=g(x)} X(x, y) dx - \int_{x=a, f(a) \leq y \leq g(a)} X(x, y) dx \\ &= \int_a^b X(x, f(x)) dx - \int_a^b X(x, g(x)) dx \\ &= - \int_a^b (X(x, g(x)) - X(x, f(x))) dx \\ &= - \int_a^b \int_{f(x)}^{g(x)} \frac{\partial X(x, y)}{\partial y} dy dx \\ &= - \int_D \frac{\partial X(x, y)}{\partial y} dx dy. \end{aligned}$$

把 Ω 分解为有限多个形如 D 的区域的并, 于是

$$\int_{\partial\Omega} X(x, y) dx = \int_{\Omega} \left(-\frac{\partial X(x, y)}{\partial y} \right) dx dy.$$

类似可证

$$\int_{\partial\Omega} Y(x, y) dy = \int_{\Omega} \frac{\partial Y(x, y)}{\partial x} dx dy.$$

因此

$$\int_{\partial\Omega} X(x, y) dx + Y(x, y) dy = \int_{\Omega} \left(\frac{\partial Y(x, y)}{\partial x} - \frac{\partial X(x, y)}{\partial y} \right) dx dy.$$

于是

$$\int_{\partial\Omega} -Y(x, y) dx + X(x, y) dy = \int_{\Omega} \left(\frac{\partial X(x, y)}{\partial x} + \frac{\partial Y(x, y)}{\partial y} \right) dx dy.$$

上述两式的向量场形式即为旋度公式和散度公式。 \square

对一阶微分形式

$$\omega = X(x, y)dx + Y(x, y)dy,$$

记

$$d\omega = \left(\frac{\partial Y(x, y)}{\partial x} - \frac{\partial X(x, y)}{\partial y} \right) dx \wedge dy.$$

因此 Green 公式可以写成

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$

定义 7.3.4. 称集合 $A \subset \mathbb{R}^m$ 是单连通的, 如果 A 中任何闭曲线都可以在 A 中连续变形为一个点。

例 7.3.5. 平面上的圆盘是单连通的, 圆环、圆周不是单连通的; 球面是单连通的, 地幔是单连通的, 圆柱面不是单连通的, 轮胎面不是单连通的。

推论 7.3.6. 如果 Ω 是平面上的单连通区域, 则 Ω 上的任何无旋的 C^1 向量场都是保守场。

证明. 设 $(X, Y)^T$ 是 Ω 上的一个无旋的 C^1 向量场。为证它是保守场, 只需证明它沿任何闭曲线 γ 积分为零。不妨设 γ 是简单闭曲线。

取 D 为由 γ 围城的区域, 则 $D \subset \Omega$ 。因为 Ω 单连通, 所以 $\partial D = \gamma$, 从而

$$\int_{\gamma} Xdx + Ydy = \int_D \text{rot}(X, Y)^T d\sigma = 0.$$

所以 $(X, Y)^T$ 是 Ω 上保守向量场。 \square

例 7.3.7. 求由曲线 $|x|^p + |y|^p = 1$ ($p > 0$) 围成的有界区域 Ω 的面积。

解: 记 $\gamma_1 : x^p + y^p = 1 (x, y \geq 0)$, $\gamma_2 : (-x)^p + y^p = 1 (x \leq 0 \leq y)$, $\gamma_3 : (-x)^p + (-y)^p = 1 (x, y < 0)$, $\gamma_4 : x^p + (-y)^p = 1 (y \leq 0 \leq x)$ 。于是 $\partial\Omega = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ 。

对 γ_1 , 设 $x = (\cos\theta)^{\frac{2}{p}}$, $y = (\sin\theta)^{\frac{2}{p}}$ 。则

$$\begin{aligned} \frac{1}{2} \int_{\gamma_1} -ydx + xdy &= \frac{1}{p} \int_0^{\frac{\pi}{2}} (\sin\theta)^{\frac{2}{p}-1} (\cos\theta)^{\frac{2}{p}-1} d\theta \\ &= \frac{1}{2p} \int_0^{\frac{\pi}{2}} (\sin\theta)^{\frac{2}{p}-2} (\cos\theta)^{\frac{2}{p}-2} d\sin^2\theta \\ &= \frac{1}{2p} \int_0^1 t^{\frac{1}{p}-1} (1-t)^{\frac{1}{p}-1} dt \\ &= \frac{1}{2p} B\left(\frac{1}{p}, \frac{1}{p}\right) = \frac{1}{2p} \frac{\Gamma\left(\frac{1}{p}\right)^2}{\Gamma\left(\frac{2}{p}\right)}. \end{aligned}$$

对 γ_2 , 考虑 $\Phi(u, v) = (-v, u) = (x, y)$, 则 $\Phi(\gamma_1) = \gamma_2$,

$$\frac{1}{2} \int_{\gamma_2} -y dx + x dy = \frac{1}{2} \int_{\gamma_1} -u d(-v) + (-v) du = \frac{1}{2} \int_{\gamma_1} -v du + u dv,$$

同理可证

$$\frac{1}{2} \int_{\gamma_k} -y dx + x dy = \frac{1}{2} \int_{\gamma_1} -y dx + x dy, \quad k = 2, 3, 4.$$

于是

$$\mu(\Omega) = \int_{\Omega} dx dy = \frac{1}{2} \int_{\gamma} -y dx + x dy = \frac{2}{p} \frac{\Gamma\left(\frac{1}{p}\right)^2}{\Gamma\left(\frac{2}{p}\right)}.$$

□

例 7.3.8. 沿上半椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 (y \geq 0)$ 从 $A(a, 0)$ 到 $B(-a, 0)$ 计算

$$\int_{\gamma} (1 + ye^x) dx + (x + e^x) dy.$$

解: 取 $\Omega = \{(x, y) | \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1, y \geq 0\}$. 则 $\partial\Omega = \gamma + \tilde{\gamma}$, 其中 $\tilde{\gamma}$ 是沿 x 轴从 B 到 A . 所以

$$\int_{\partial\Omega} (1 + ye^x) dx + (x + e^x) dy = \int_{\Omega} -e^x + 1 + e^x dx dy = \int_{\Omega} dx dy = \frac{ab\pi}{2}.$$

于是

$$\begin{aligned} \int_{\gamma} (1 + ye^x) dx + (x + e^x) dy &= \frac{ab\pi}{2} - \int_{\tilde{\gamma}} (1 + ye^x) dx + (x + e^x) dy \\ &= \frac{ab\pi}{2} - \int_{-a}^a dx = \frac{ab\pi}{2} - 2a. \end{aligned}$$

□

7.4 第二型曲面积分的概念

定向曲面

称二维曲面 $\Sigma \subset \mathbb{R}^3$ 是一个可定向曲面, 如果存在连续的单位法向量场 $\mathbf{n}: \Sigma \rightarrow \mathbb{R}^3$, 即对任意 $\mathbf{x} \in \Sigma$, $\|\mathbf{n}(\mathbf{x})\| = 1$ 且 $\mathbf{n}(\mathbf{x}) \perp T_{\mathbf{x}}\Sigma$. 称指定了一个连续的单位法向量场的曲面为定向曲面。

例 7.4.1. 球面 $S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = R^2\}$ 是一个可定向曲面,

$$\mathbf{n}(x, y, z) = \frac{(x, y, z)^T}{R}$$

是一个连续的单位法向量场。

例 7.4.2. 如果曲面 Σ 有一个 C^1 正则的参数表示

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \\ z = z(u, v), \end{cases} \quad (u, v) \in D \subseteq \mathbb{R}^2,$$

是一个单射，且其中

$$\mathbf{u} = \begin{pmatrix} \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial u} \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial v} \end{pmatrix}$$

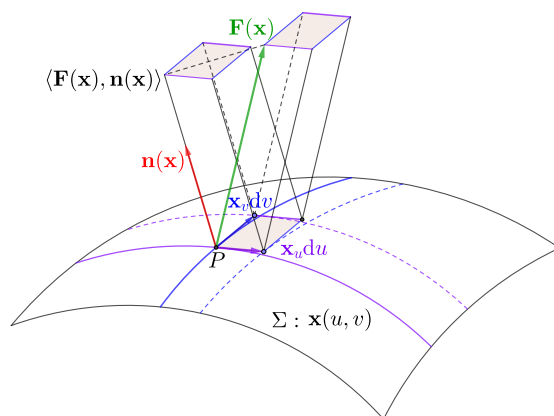
线性无关。则 $\frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|}$ 是 Σ 的一个连续的单位法向量场。

第二型曲面积分：向量场在定向曲面上的积分

定义 7.4.3. 设 (Σ, \mathbf{n}) 是一个定向曲面，对连续向量场 $\mathbf{F} : \Sigma \rightarrow \mathbf{R}^3$ ，称

$$\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma$$

为向量场 \mathbf{F} 沿曲面 Σ 的法向量场 \mathbf{n} 的积分，称为向量场 \mathbf{F} 关于有向曲面 Σ 的
通量。



定理 7.4.4. 对 C^1 正则的参数曲面 Σ :

$$\begin{cases} x = x(u_1, u_2), \\ y = y(u_1, u_2), \\ z = z(u_1, u_2), \end{cases} \quad (u_1, u_2) \in D \subseteq \mathbb{R}^2,$$

满足 $\frac{\partial \mathbf{x}}{\partial u_1} \times \frac{\partial \mathbf{x}}{\partial u_2}$ 与法向量场 \mathbf{n} 的方向一致。则对任何连续向量场 $\mathbf{F}: \Sigma \rightarrow \mathbf{R}^3$,

$$\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = \int_D \det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2} \right) du_1 du_2,$$

其中 $d\sigma$ 为曲面 Σ 的(无向)面积微元。

证明. 由6.4节的结论知

$$\det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2} \right) du_1 du_2$$

是向量 \mathbf{F} 以及曲面上两个无穷小切向量 $\frac{\partial \mathbf{x}}{\partial u_1} du_1$ 和 $\frac{\partial \mathbf{x}}{\partial u_2} du_2$ 张成的平行多面体的三维体积, 所以它等于 $\langle \mathbf{F}, \mathbf{n} \rangle d\sigma$. 于是这就完成了这个定理的证明. \square

另一个证明. 记

$$\mathbf{u}_k = \frac{\partial \mathbf{x}}{\partial u_k} = \begin{pmatrix} \frac{\partial x}{\partial u_k} \\ \frac{\partial y}{\partial u_k} \\ \frac{\partial z}{\partial u_k} \end{pmatrix}, \quad k = 1, 2; \quad \mathbf{v} = \begin{pmatrix} \det \frac{\partial(y, z)}{\partial(u_1, u_2)} \\ \det \frac{\partial(z, x)}{\partial(u_1, u_2)} \\ \det \frac{\partial(x, y)}{\partial(u_1, u_2)} \end{pmatrix},$$

则 $\mathbf{u}_1, \mathbf{u}_2$ 是 Σ 的两个线性无关的切向量, 且对 $k = 1, 2$,

$$\langle \mathbf{u}_k, \mathbf{v} \rangle = \begin{pmatrix} \frac{\partial x}{\partial u_k} & \frac{\partial y}{\partial u_k} & \frac{\partial z}{\partial u_k} \end{pmatrix} \begin{pmatrix} \det \frac{\partial(y, z)}{\partial(u_1, u_2)} \\ \det \frac{\partial(z, x)}{\partial(u_1, u_2)} \\ \det \frac{\partial(x, y)}{\partial(u_1, u_2)} \end{pmatrix} = \det \begin{pmatrix} \frac{\partial x}{\partial u_1} & \frac{\partial x}{\partial u_2} & \frac{\partial x}{\partial u_k} \\ \frac{\partial y}{\partial u_1} & \frac{\partial y}{\partial u_2} & \frac{\partial y}{\partial u_k} \\ \frac{\partial z}{\partial u_1} & \frac{\partial z}{\partial u_2} & \frac{\partial z}{\partial u_k} \end{pmatrix} = 0,$$

从而 \mathbf{v} 是 Σ 的法向量, 且

$$\begin{aligned} \|\mathbf{v}\| &= \frac{\|\mathbf{v}\|^2}{\|\mathbf{v}\|} = \frac{1}{\|\mathbf{v}\|} \left[\left(\det \frac{\partial(y, z)}{\partial(u_1, u_2)} \right)^2 + \left(\det \frac{\partial(z, x)}{\partial(u_1, u_2)} \right)^2 + \left(\det \frac{\partial(x, y)}{\partial(u_1, u_2)} \right)^2 \right] \\ &= \frac{1}{\|\mathbf{v}\|} \det(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}) \\ &= \frac{1}{\|\mathbf{v}\|} \sqrt{\det(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v})^T \det(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v})} \\ &= \frac{1}{\|\mathbf{v}\|} \sqrt{\det \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle & \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \langle \mathbf{v}, \mathbf{u}_1 \rangle & \langle \mathbf{v}, \mathbf{u}_2 \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix}} \\ &= \frac{1}{\|\mathbf{v}\|} \sqrt{\det \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & 0 \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle & 0 \\ 0 & 0 & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix}} = \sqrt{\det \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle \end{pmatrix}} \end{aligned}$$

所以

$$\mathbf{n} = \frac{\mathbf{v}}{\sqrt{\det \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle \end{pmatrix}}}$$

是 Σ 的单位法向量。

注意到

$$d\sigma = \sqrt{\det \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle \end{pmatrix}} du_1 du_2,$$

因此

$$\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = \int_D \langle \mathbf{F}, \mathbf{v} \rangle du_1 du_2 = \int_D \det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2} \right) du_1 du_2.$$

□

推论 7.4.5. 对 m 维 C^1 正则的参数曲面 $\Sigma \subset \mathbb{R}^{m+1}$:

$$\mathbf{x} = \mathbf{x}(u_1, u_2, \dots, u_m), \quad (u_1, u_2, \dots, u_m) \in D \subseteq \mathbb{R}^m,$$

以及连续向量场 $\mathbf{F} : \Sigma \rightarrow \mathbb{R}^{m+1}$,

$$\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = \int_D \det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) du_1 du_2 \cdots du_m,$$

其中 $d\sigma$ 为曲面 Σ 的 m 维(无向)体积微元。

特别当 $m = 1$ 时,

$$\begin{aligned} \int_{\gamma} \langle \mathbf{F}, \mathbf{n} \rangle dl &= \int_a^b \det \left(\mathbf{F}(\mathbf{x}(t)), \frac{\partial \mathbf{x}}{\partial t} \right) dt = \int_a^b \begin{vmatrix} X(x(t), y(t)) & x'(t) \\ Y(x(t), y(t)) & y'(t) \end{vmatrix} dt \\ &= \int_{\gamma} -Y(x, y) dx + X(x, y) dy, \end{aligned}$$

其中 $\mathbf{F}(x, y) = (X(x, y), Y(x, y))^T$ 。

□

注 7.4.6. 上述定理结论的几何含义。由行列式的性质知, 如果按曲面 Σ 的切方向和法方向把 \mathbf{F} 分解为直和 $\mathbf{F} = \mathbf{F}^t + \mathbf{F}^\perp$, 则

$$\begin{aligned} &\det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) \\ &= \det \left(\mathbf{F}^t, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) + \det \left(\mathbf{F}^\perp, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) \\ &= \det \left(\mathbf{F}^\perp, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right). \end{aligned}$$

于是

$$\begin{aligned} & \left| \det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) \right| \\ &= \sqrt{\det \left(\mathbf{F}^\perp, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right)^T \det \left(\mathbf{F}^\perp, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right)} \\ &= \|\mathbf{F}^\perp\| \sqrt{\det (\langle \mathbf{u}_i, \mathbf{u}_j \rangle)}. \end{aligned}$$

这是一个以 \mathbf{F}^\perp 为高的 $m+1$ 维平行多面体的体积, 其底面由曲面 Σ 的切向量 $\frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m}$ 张成. 所以 $\det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2}, \dots, \frac{\partial \mathbf{x}}{\partial u_m} \right) du_1 \cdots du_m$ 是由曲面面积微元沿 \mathbf{F} 形成的无穷小柱体的有向体积. 物理上它是单位时间中沿流速向量 \mathbf{F} 穿过曲面上一个无穷小区域的通量.

线性函数的楔积与二阶微分形式

对两个线性函数 L_1, L_2 和两个向量 \mathbf{u}, \mathbf{v} , 记

$$(L_1 \wedge L_2)(\mathbf{u}, \mathbf{v}) = \det \begin{pmatrix} L_1(\mathbf{u}) & L_1(\mathbf{v}) \\ L_2(\mathbf{u}) & L_2(\mathbf{v}) \end{pmatrix}.$$

因此 $L_1 \wedge L_2$ 满足

1. $L_1 \wedge L_2$ 关于 (\mathbf{u}, \mathbf{v}) 是一个双线性函数: 对 \mathbf{u} 和 \mathbf{v} 分别是线性的.
2. $L_1 \wedge L_2$ 关于 (\mathbf{u}, \mathbf{v}) 是反对称的:

$$(L_1 \wedge L_2)(\mathbf{u}, \mathbf{v}) = -(L_1 \wedge L_2)(\mathbf{v}, \mathbf{u}).$$

3. $L_1 \wedge L_2$ 关于 L_1 和 L_2 分别是线性的, 对 (L_1, L_2) 是反对称的: $L_1 \wedge L_2 = -L_2 \wedge L_1$.

由反对称性知 $L_1 \wedge L_1 = 0$.

对于 C^1 函数 $x = x(u, v), y = y(u, v)$,

$$\begin{aligned} dx \wedge dy &= \left(\frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv \right) \wedge \left(\frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv \right) \\ &= \frac{\partial x}{\partial u} \frac{\partial y}{\partial u} du \wedge du + \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} du \wedge dv + \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} dv \wedge du + \frac{\partial x}{\partial v} \frac{\partial y}{\partial v} dv \wedge dv \\ &= \left(\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right) du \wedge dv \\ &= \det \frac{\partial(x, y)}{\partial(u, v)} du \wedge dv. \end{aligned}$$

对 C^1 正则的参数曲面 Σ :

$$\begin{cases} x = x(u_1, u_2), \\ y = y(u_1, u_2), \\ z = z(u_1, u_2), \end{cases} \quad (u_1, u_2) \in D \subseteq \mathbb{R}^2,$$

以及连续向量场 $\mathbf{F} : \Sigma \rightarrow \mathbf{R}^3$, $\mathbf{F}(x, y, z) = (X(x, y, z), Y(x, y, z), Z(x, y, z))^T$,

$$\begin{aligned} & \det \left(\mathbf{F}, \frac{\partial \mathbf{x}}{\partial u_1}, \frac{\partial \mathbf{x}}{\partial u_2} \right) du_1 \wedge du_2 \\ &= X(x, y, z) \det \frac{\partial(y, z)}{\partial(u_1, u_2)} du_1 \wedge du_2 + Y(x, y, z) \det \frac{\partial(z, x)}{\partial(u_1, u_2)} du_1 \wedge du_2 \\ & \quad + Z(x, y, z) \det \frac{\partial(x, y)}{\partial(u_1, u_2)} du_1 \wedge du_2 \\ &= X(x, y, z) dy \wedge dz + Y(x, y, z) dz \wedge dx + Z(x, y, z) dx \wedge dy. \end{aligned}$$

记

$$\omega = X(x, y, z) dy \wedge dz + Y(x, y, z) dz \wedge dx + Z(x, y, z) dx \wedge dy,$$

称它为一个二阶微分形式。

对 (x, y, z) 处的两个向量 $\mathbf{u} = (\xi_1, \xi_2, \xi_3)$ 和 $\mathbf{v} = (\eta_1, \eta_2, \eta_3)$,

$$\omega(\mathbf{u}, \mathbf{v}) = \det \begin{pmatrix} X & \xi_1 & \eta_1 \\ Y & \xi_2 & \eta_2 \\ Z & \xi_3 & \eta_3 \end{pmatrix}$$

是一个反对称的双线性函数。

$$dx \wedge dy(\mathbf{u}, \mathbf{v}) = \det \begin{pmatrix} 0 & \xi_1 & \eta_1 \\ 0 & \xi_2 & \eta_2 \\ 1 & \xi_3 & \eta_3 \end{pmatrix} = \det \begin{pmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{pmatrix}$$

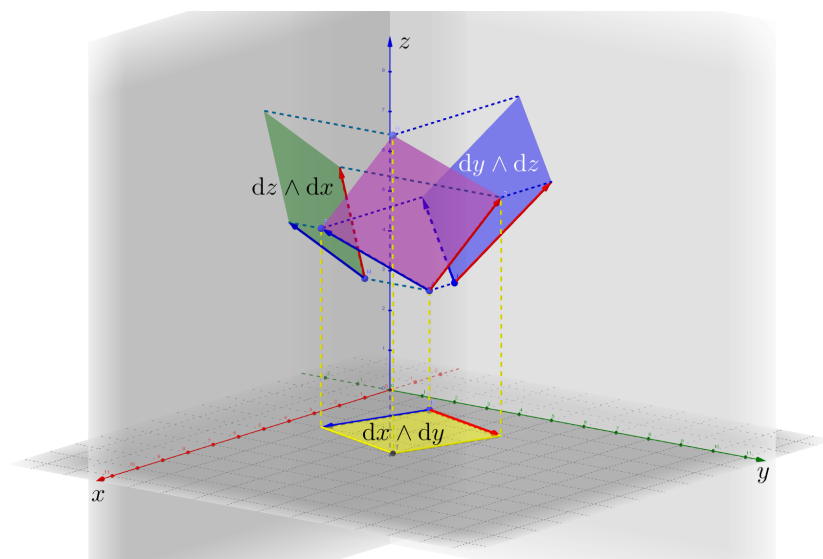
是 \mathbf{u}, \mathbf{v} 在 xy 坐标平面中的投影向量所张成的平行四边形的有向面积, 也就是由 \mathbf{u}, \mathbf{v} 在 \mathbf{R}^3 中张成的有向平行四边形在 xy 坐标平面上的投影面积。所以在有向曲面 Σ 上, $dx \wedge dy, dy \wedge dz, dz \wedge dx$ 的作用是把 Σ 的面积微元在相应的二维坐标平面上的有向投影面积, 面积的正负号取决于曲面 Σ 的指定法向量在各坐标平面的法向上投影的正负。

定理 7.4.7. 1. $\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma$ 和 $\int_{\Sigma} \omega$ 关于连续向量场 \mathbf{F} 和二阶微分形式 ω 是线性的。

2. $\int_{-\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = -\int_{\Sigma} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma$, $\int_{-\Sigma} \omega = -\int_{\Sigma} \omega$ 。

3. 若两个有向曲面 Σ_1 和 Σ_2 满足 $\Sigma_1 \cap \Sigma_2 \subset \partial\Sigma_1 \cap \partial\Sigma_2$, 且它们可以拼接为一个有向曲面 $\Sigma_1 + \Sigma_2$, 则

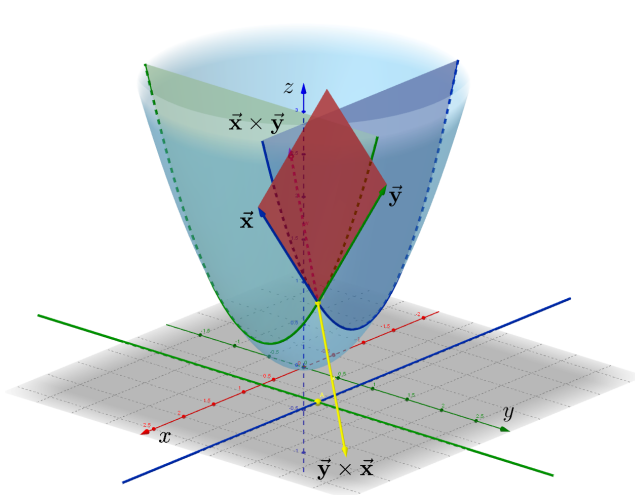
$$\begin{aligned} \int_{\Sigma_1 + \Sigma_2} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma &= \int_{\Sigma_1} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma + \int_{\Sigma_2} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma, \\ \int_{\Sigma_1 + \Sigma_2} \omega &= \int_{\Sigma_1} \omega + \int_{\Sigma_2} \omega. \end{aligned}$$



例 7.4.8. 设有向曲面 Σ 为函数 $z = x^2 + y^2$ 图像在 $x^2 + y^2 \leq 1$ 的部分, 方向朝 z 轴负半轴。求

$$\int_{\Sigma} x dy \wedge dz.$$

解: 因为 $x dy \wedge dz = x dy \wedge dz + 0 dz \wedge dx + 0 dx \wedge dy$, 所以向量场 $\mathbf{F} = (x, 0, 0)^T$ 。



曲面参数可取 x, y 。由于曲面正向朝向 z 轴负方向, 所以参数顺序为 (y, x) , 即 $dy \wedge dx = dx dy$ 。曲面相应的切向量依次为

$$\begin{pmatrix} 0 \\ 1 \\ 2y \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix}, \quad \text{即曲面参数方程} \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix} \text{依次对 } y, x \text{ 求导得到.}$$

所以

$$\int_{\Sigma} x dy \wedge dz = \int_{x^2+y^2 \leq 1} \begin{vmatrix} x & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2y & 2x \end{vmatrix} dx dy = 2 \int_{x^2+y^2 \leq 1} x^2 dx dy$$

由重积分换元 $(x, y) \rightarrow (y, x)$ 知

$$\int_{x^2+y^2 \leq 1} x^2 dx dy = \int_{x^2+y^2 \leq 1} y^2 dx dy,$$

因此

$$\int_{\Sigma} x dy \wedge dz = \int_{x^2+y^2 \leq 1} (x^2 + y^2) dx dy = 2\pi \int_0^1 r^3 dr = \frac{\pi}{2}.$$

□

另解: 以 (θ, r) 为曲面参数, 曲面参数方程为

$$\begin{cases} x = r \cos \theta, \\ y = r \sin \theta, \\ z = r^2, \end{cases} \quad 0 \leq \theta \leq 2\pi, 0 \leq r \leq 1.$$

则

$$\begin{aligned} \int_{\Sigma} x dy \wedge dz &= \int_D r \cos \theta d(r \sin \theta) \wedge d(r^2) = \int_D 2r^3 \cos^2 \theta d\theta \wedge dr \\ &= \int_D 2r^3 \cos^2 \theta dr d\theta \quad (\text{曲面 } \Sigma \text{ 的定向使 } d\theta \wedge dr = dr d\theta) \\ &= \frac{\pi}{2}. \end{aligned}$$

□

例 7.4.9. 设 Σ 为椭球面 $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ 的外侧, 求

$$\int_{\Sigma} x dy \wedge dz + y dz \wedge dx + z dx \wedge dy.$$

解: 椭球坐标

$$x = a \sin \theta \cos \varphi, y = b \sin \theta \sin \varphi, z = c \cos \theta, (D: 0 \leq \varphi \leq 2\pi, 0 \leq \theta \leq \frac{\pi}{2}).$$

椭球面上 (θ, φ) 坐标系与外法向一致, 于是

$$\begin{aligned} &\int_{\Sigma} x dy \wedge dz + y dz \wedge dx + z dx \wedge dy \\ &= \int_D \begin{vmatrix} a \sin \theta \cos \varphi & a \cos \theta \cos \varphi & -a \sin \theta \sin \varphi \\ b \sin \theta \sin \varphi & b \cos \theta \sin \varphi & b \sin \theta \cos \varphi \\ c \cos \theta & -c \sin \theta & 0 \end{vmatrix} d\theta d\varphi \\ &= abc \int_{[0, \frac{\pi}{2}] \times [0, 2\pi]} \sin \theta d\theta d\varphi \\ &= 4\pi abc. \end{aligned}$$

□

另解：球坐标 $(X, Y, Z) : X^2 + Y^2 + Z^2 = 1$,

$$x = aX, y = bY, z = cZ, (S^2 : X^2 + Y^2 + Z^2 = 1).$$

于是

$$\begin{aligned} & \int_{\Sigma} xdy \wedge dz + ydz \wedge dx + zdx \wedge dy \\ &= \int_{S^2} aXd(bY) \wedge d(cZ) + bYd(cZ) \wedge d(aX) + cZd(aX) \wedge d(bY) \\ &= abc \int_{S^2} XdY \wedge dZ + YdZ \wedge dX + ZdX \wedge dY \\ &= abc \int_{S^2} \langle \mathbf{n}, \mathbf{n} \rangle d\sigma, \quad ((X, Y, Z)^T \text{ 是 } S^2 \text{ 的单位外法向量 } \mathbf{n}) \\ &= abc \cdot 4\pi \end{aligned}$$

□

例 7.4.10. 设 Σ 为曲面 $z = 1 - x^2 - y^2$ ($x^2 + y^2 \leq 1$), 方向向上. 求

$$\int_{\Sigma} (x^2 - z)dx \wedge dy + (z^2 - y)dz \wedge dx.$$

解：由曲面定向知 $dx \wedge dy = dx dy$.

$$\begin{aligned} I &= \int_{\Sigma} (x^2 - z)dx \wedge dy + (z^2 - y)dz \wedge dx \\ &= \int_{\Sigma} [x^2 - (1 - x^2 - y^2)]dx \wedge dy + [(1 - x^2 - y^2)^2 - y]d(1 - x^2 - y^2) \wedge dx \\ &= \int_{\Sigma} (2x^2 + y^2 - 1)dx \wedge dy + [(1 - x^2 - y^2)^2 - y](-2y)dy \wedge dx \\ &= \int_{\Sigma} (2x^2 - y^2 - 1)dx \wedge dy \quad (dy \wedge dx = -dx \wedge dy, \text{ 且删除关于 } y \text{ 的奇次项}) \\ &= \int_{x^2+y^2 \leq 1} (2x^2 - y^2 - 1)dx dy = -\frac{3\pi}{4}. \end{aligned}$$

□

习题 7.4

1. 设 k 是正整数, L_1, \dots, L_k 是 k 个线性函数, $\mathbf{v}_1, \dots, \mathbf{v}_k$ 是 k 个向量, 记

$$(L_1 \wedge \dots \wedge L_k)(\mathbf{v}_1, \dots, \mathbf{v}_k) = \begin{vmatrix} L_1(\mathbf{v}_1) & L_1(\mathbf{v}_2) & \cdots & L_1(\mathbf{v}_k) \\ L_2(\mathbf{v}_1) & L_2(\mathbf{v}_2) & \cdots & L_2(\mathbf{v}_k) \\ \vdots & \vdots & \ddots & \vdots \\ L_k(\mathbf{v}_1) & L_k(\mathbf{v}_2) & \cdots & L_k(\mathbf{v}_k) \end{vmatrix}.$$

证明 $(L_1 \wedge \dots \wedge L_k)(\mathbf{v}_1, \dots, \mathbf{v}_k)$ 关于 $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ 是反对称 k 重线性函数, $L_1 \wedge \dots \wedge L_k$ 关于 (L_1, \dots, L_k) 也是反对称 k 重线性的。

2. 试给出 k 阶微分形式

$$\sum_{\sigma \text{ 是 } 1, 2, \dots, m \text{ 的排列}} X_{\sigma}(x_1, x_2, \dots, x_m) dx_{\sigma(1)} \wedge dx_{\sigma(2)} \wedge \dots \wedge dx_{\sigma(k)}$$

的一个定义。

7.5 空间向量场的旋度与散度, Gauss 公式与 Stokes 公式

定义 7.5.1. 对 \mathbb{R}^3 中 (直角坐标系下) 的向量场 $\mathbf{F} = (X, Y, Z)^T$, 记

$$\begin{aligned} \operatorname{rot} \mathbf{F} &= \nabla \times \mathbf{F} = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}, \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}, \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right)^T, \\ \operatorname{div} \mathbf{F} &= \nabla \cdot \mathbf{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = \operatorname{tr} \frac{\partial(X, Y, Z)}{\partial(x, y, z)}, \end{aligned}$$

分别称它们为 \mathbf{F} 的旋度和散度。

旋度为零的向量场称为无旋向量场, 散度为零的向量场称为无源向量场。

例 7.5.2. 向量场 $\mathbf{F}(\mathbf{x}) = (X_1(x_1, x_2, x_3), X_2(x_1, x_2, x_3), X_3(x_1, x_2, x_3))^T$ 在以 O 为中心的小球 $B(O, r)$ 的表面 $\partial B(O, r)$ 上产生的向外的通量按球体积的平均值为

$$\begin{aligned} & \lim_{r \rightarrow 0^+} \frac{1}{\frac{4}{3}\pi r^3} \int_{\partial B(O, r)} \langle \mathbf{F}(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle d\sigma \\ &= \lim_{r \rightarrow 0^+} \frac{1}{\frac{4}{3}\pi r^3} \int_{\partial B(O, r)} \sum_{i=1}^3 \frac{x_i}{r} X_i d\sigma \\ &= \lim_{r \rightarrow 0^+} \frac{1}{\frac{4}{3}\pi r^3} \int_{\partial B(O, r)} \sum_{i=1}^3 \frac{x_i}{r} \left[X_i(O) + \sum_{j=1}^3 x_j \frac{\partial X_i}{\partial x_j} X_i(O) + o(r) \right] d\sigma \\ &= \lim_{r \rightarrow 0^+} \frac{1}{\frac{4}{3}\pi r^4} \sum_{i=1}^3 \frac{\partial X_i}{\partial x_i}(O) \int_{\partial B(O, r)} x_i^2 d\sigma \quad (\text{关于 } x_i \text{ 的奇数次项积分为零}) \\ &= \lim_{r \rightarrow 0^+} \frac{1}{\frac{4}{3}\pi r^4} \sum_{i=1}^3 \frac{\partial X_i}{\partial x_i}(O) \int_{\partial B(O, r)} \frac{x_1^2 + x_2^2 + x_3^2}{3} d\sigma \\ &= \frac{1}{\frac{4}{3}\pi r^4} \sum_{i=1}^3 \frac{\partial X_i}{\partial x_i}(O) \int_{\partial B(O, r)} \frac{r^2}{3} d\sigma = \sum_{i=1}^3 \frac{\partial X_i}{\partial x_i}(O) = \operatorname{div} \mathbf{F}(O). \end{aligned}$$

例 7.5.3. 对线性向量场

$$\mathbf{A}(x, y, z) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

则

$$\operatorname{rot} \mathbf{A} = (a_{32} - a_{23}, a_{13} - a_{31}, a_{21} - a_{12})^T$$

$$\operatorname{div} \mathbf{A} = \operatorname{tr} A = a_{11} + a_{22} + a_{33}.$$

因此, \mathbf{A} 是无旋向量场当且仅当矩阵 A 是对称的, \mathbf{A} 是无源向量场当且仅当矩阵 A 的迹为零。

对任何三阶矩阵 A ,

$$A = A_s + A_a, \quad A_s = \frac{A + A^T}{2}, \quad A_a = \frac{A - A^T}{2}.$$

A_s 是对称阵对应无旋向量场 \mathbf{A}_s , A_s 有三个彼此正交的特征方向, 它们限制了向量场 \mathbf{A}_s 的积分曲线绕原点旋转; A_a 是反对称阵, 对角线为零, 是无源向量场 \mathbf{A}_a , \mathbf{A}_a 的积分曲线是以原点为中心的圆周, 这些圆周使向量场的积分曲线无法穿进或穿出。

$\mathbf{A} = \mathbf{A}_s + \mathbf{A}_a$, 并且

$$\operatorname{rot} \mathbf{A} = \operatorname{rot} \mathbf{A}_a, \quad \operatorname{div} \mathbf{A} = \operatorname{div} \mathbf{A}_s.$$

□

对二阶微分形式

$$\omega = X(x, y, z)dy \wedge dz + Y(x, y, z)dz \wedge dx + Z(x, y, z)dx \wedge dy,$$

定义 $d\omega$ 是如下三阶微分形式

$$d\omega = dX(x, y, z) \wedge dy \wedge dz + dY(x, y, z) \wedge dz \wedge dx + dZ(x, y, z) \wedge dx \wedge dy.$$

于是

$$\begin{aligned} d\omega &= dX(x, y, z) \wedge dy \wedge dz + dY(x, y, z) \wedge dz \wedge dx + dZ(x, y, z) \wedge dx \wedge dy \\ &= \left(\frac{\partial X}{\partial x} dx + \frac{\partial X}{\partial y} dy + \frac{\partial X}{\partial z} dz \right) \wedge dy \wedge dz \\ &\quad + \left(\frac{\partial Y}{\partial x} dx + \frac{\partial Y}{\partial y} dy + \frac{\partial Y}{\partial z} dz \right) \wedge dz \wedge dx \\ &\quad + \left(\frac{\partial Z}{\partial x} dx + \frac{\partial Z}{\partial y} dy + \frac{\partial Z}{\partial z} dz \right) \wedge dx \wedge dy \\ &= \frac{\partial X}{\partial x} dx \wedge dy \wedge dz + \frac{\partial Y}{\partial y} dy \wedge dz \wedge dx + \frac{\partial Z}{\partial z} dz \wedge dx \wedge dy \\ &= \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right) dx \wedge dy \wedge dz. \end{aligned}$$

定理 7.5.4 (Gauss 公式). 设 $\Omega \subset \mathbb{R}^3$ 是有界区域, Ω 的边界 $\partial\Omega$ 是分片 C^1 曲面, 方向指向 Ω 外侧。则

1. 对任何 C^1 向量场 $\mathbf{F} : \Omega \rightarrow \mathbf{R}^3$,

$$\int_{\partial\Omega} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = \int_{\Omega} \operatorname{div} \mathbf{F} d\mu,$$

即向量场沿区域边界的向外的净通量等于在区域内部的散度的积分。这里 μ 是三维体积。

2. 对任何 C^1 二阶微分形式 ω ,

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$

证明. 把 Ω 分解成有限多个形如

$$D = \{(x, y, z) | (x, y) \in U, f(x, y) \leq z \leq g(x, y)\}$$

的区域的并, 每个都取外法向, 于是对相邻两个区域, 它们的共同边界具有相反的方向, 积分相互抵消。

$$\begin{aligned} \int_{\partial D} Z(x, y, z) dx \wedge dy &= \int_U Z(x, y, g(x, y)) dx dy - \int_U Z(x, y, f(x, y)) dx dy \\ &= \int_U \int_{f(x, y)}^{g(x, y)} \frac{\partial Z}{\partial z}(x, y, z) dx dy dz \\ &= \int_D \frac{\partial Z}{\partial z}(x, y, z) dx dy dz. \end{aligned}$$

因此

$$\int_{\partial\Omega} Z(x, y, z) dx \wedge dy = \int_{\Omega} \frac{\partial Z}{\partial z}(x, y, z) dx dy dz.$$

同理可证

$$\begin{aligned} \int_{\partial\Omega} X(x, y, z) dy \wedge dz &= \int_{\Omega} \frac{\partial X}{\partial x}(x, y, z) dx dy dz, \\ \int_{\partial\Omega} Y(x, y, z) dz \wedge dx &= \int_{\Omega} \frac{\partial Y}{\partial y}(x, y, z) dx dy dz. \end{aligned}$$

上述三式相加得到

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$

□

推论 7.5.5. 设 $\Omega \subset \mathbb{R}^{m+1}$ 是有界区域, Ω 的边界 $\partial\Omega$ 是分片 C^1 m 维曲面, 方向指向 Ω 外侧。则

1. 对任何 C^1 向量场 $\mathbf{F} : \Omega \rightarrow \mathbf{R}^{m+1}$,

$$\int_{\partial\Omega} \langle \mathbf{F}, \mathbf{n} \rangle d\sigma = \int_{\Omega} \operatorname{div} \mathbf{F} d\mu,$$

即向量场沿区域边界的向外的净通量等于在区域内部的散度的积分。这里 μ 是 $m+1$ 维体积。

2. 对任何 C^1 的 m 阶微分形式

$$\omega = \sum_{k=1}^{m+1} (-1)^{\#\sigma(k)} X_k dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_{k+1} \wedge \cdots \wedge dx_{m+1},$$

(其中 $\#\sigma(k)$ 是 $k, 1, 2, \dots, k-1, k+1, \dots, m+1$ 作为 $1, 2, \dots, m+1$ 的一个排列的奇偶性), 都成立

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega,$$

其中

$$d\omega = \operatorname{tr} \frac{\partial(X_1, X_2, \dots, X_{m+1})}{\partial(x_1, x_2, \dots, x_{m+1})} dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{m+1}.$$

□

例 7.5.6. 设 Σ 为曲面 $z = 1 - x^2 - y^2$ ($x^2 + y^2 \leq 1$), 方向向上。求

$$\int_{\Sigma} (x^2 - z) dx \wedge dy + (z^2 - y) dz \wedge dx.$$

解: 记 $\omega = (x^2 - z) dx \wedge dy + (z^2 - y) dz \wedge dx$ 。

取 $\Omega = \{(x, y, z) | 0 \leq z \leq 1 - x^2 - y^2\}$, $\partial\Omega$ 方向朝外。则由 Gauss 公式

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega = \int_{\Omega} -2 dx dy dz = -2 \int_{x^2+y^2 \leq 1} (1 - x^2 - y^2) dx dy.$$

另一方面

$$\begin{aligned} \int_{\partial\Omega} \omega &= \int_{\Sigma} \omega + \int_{x^2+y^2 \leq 1, z=0 \text{ 下侧}} \omega \\ &= \int_{\Sigma} \omega + \int_{x^2+y^2 \leq 1, z=0 \text{ 下侧}} x^2 dx \wedge dy \\ &= \int_{\Sigma} \omega - \int_{x^2+y^2 \leq 1} x^2 dx dy, \end{aligned}$$

所以

$$\begin{aligned} \int_{\Sigma} \omega &= \int_{\partial\Omega} \omega + \int_{x^2+y^2 \leq 1} x^2 dx dy \\ &= -2 \int_{x^2+y^2 \leq 1} (1 - x^2 - y^2) dx dy + \int_{x^2+y^2 \leq 1} x^2 dx dy \\ &= \int_{x^2+y^2 \leq 1} (3x^2 + 2y^2 - 2) dx dy = I, \end{aligned}$$

$$2I = \int_{x^2+y^2 \leq 1} (5x^2 + 5y^2 - 4) dx dy = -\frac{3\pi}{2},$$

因此 $\int_{\Sigma} \omega = I = -\frac{3\pi}{4}$. □

对一阶微分形式

$$\omega = X(x, y, z)dx + Y(x, y, z)dy + Z(x, y, z)dz,$$

令

$$d\omega = dX(x, y, z) \wedge dx + dY(x, y, z) \wedge dy + dZ(x, y, z) \wedge dz.$$

于是利用 \wedge 的双线性和反对称性, 得到

$$\begin{aligned} d\omega &= dX \wedge dx + dY \wedge dy + dZ \wedge dz \\ &= \left(\frac{\partial X}{\partial x} dx + \frac{\partial X}{\partial y} dy + \frac{\partial X}{\partial z} dz \right) \wedge dx + \left(\frac{\partial Y}{\partial x} dx + \frac{\partial Y}{\partial y} dy + \frac{\partial Y}{\partial z} dz \right) \wedge dy \\ &\quad + \left(\frac{\partial Z}{\partial x} dx + \frac{\partial Z}{\partial y} dy + \frac{\partial Z}{\partial z} dz \right) \wedge dz \\ &= \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z} \right) dy \wedge dz + \left(\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x} \right) dz \wedge dx + \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) dx \wedge dy. \end{aligned}$$

定理 7.5.7 (Stokes 公式). 设 $\Sigma \subset \mathbb{R}^3$ 是有界的分片 C^1 有向曲面, Σ 的边界 $\partial\Sigma$ 是分段 C^1 有向曲线, 使得按曲面指定法向站立并沿 $\partial\Sigma$ 前进时, 曲面总在左手一侧. 则

1. 对任何 C^1 向量场 $\mathbf{F} : \Sigma \rightarrow \mathbf{R}^3$,

$$\int_{\partial\Sigma} \langle \mathbf{F}, \mathbf{T} \rangle dl = \int_{\Sigma} \langle \text{rot } \mathbf{F}, \mathbf{n} \rangle d\sigma,$$

即向量场沿曲面边界的环量等于在曲面上旋度的积分。

2. 对任何 C^1 一阶微分形式 ω ,

$$\int_{\partial\Sigma} \omega = \int_{\Sigma} d\omega.$$

证明. 我们考察 $\int_{\partial\Sigma} X dx_1$.

设 $\mathbf{x}(u_1, u_2) = (x_1(u_1, u_2), x_2(u_1, u_2), x_3(u_1, u_2))$ ($(u_1, u_2) \in D \subset \mathbb{R}^2$) 是曲面的 C^1 参数化. $\mathbf{u}(t) = (u_1(t), u_2(t))$ 是 ∂D 的参数方程. $\mathbf{x}(\mathbf{u}(t))$ 是 $\partial\Sigma$ 的参

数方程。

$$\begin{aligned}
 \int_{\partial\Sigma} X dx_1 &= \int_a^b X(\mathbf{x}(\mathbf{u}(t))) dx_1(\mathbf{u}(t)) \\
 &= \int_a^b \left(X(\mathbf{x}(\mathbf{u}(t))) \frac{\partial x_1}{\partial u_1}(\mathbf{u}(t)) u_1'(t) + X(\mathbf{x}(\mathbf{u}(t))) \frac{\partial x_1}{\partial u_2}(\mathbf{u}(t)) u_2'(t) \right) dt \\
 &= \int_{\partial D} X(\mathbf{x}(u_1, u_2)) \frac{\partial x_1}{\partial u_1}(u_1, u_2) du_1 + X(\mathbf{x}(u_1, u_2)) \frac{\partial x_1}{\partial u_2}(u_1, u_2) du_2 \\
 &= \int_D \frac{\partial}{\partial u_1} \left(X(\mathbf{x}(\mathbf{u})) \frac{\partial x_1}{\partial u_2}(\mathbf{u}) \right) - \frac{\partial}{\partial u_2} \left(X(\mathbf{x}(\mathbf{u})) \frac{\partial x_1}{\partial u_1}(\mathbf{u}) \right) du_1 du_2 \\
 &= \int_D \sum_{k=1}^3 \frac{\partial X}{\partial x_k} \left[\frac{\partial x_k}{\partial u_1} \frac{\partial x_1}{\partial u_2} - \frac{\partial x_k}{\partial u_2} \frac{\partial x_1}{\partial u_1} \right] du_1 du_2 \\
 &= \int_D \frac{\partial X}{\partial x_3} \det \frac{\partial(x_3, x_1)}{\partial(u_1, u_2)} - \frac{\partial X}{\partial x_2} \det \frac{\partial(x_1, x_2)}{\partial(u_1, u_2)} du_1 du_2 \\
 &= \int_{\Sigma} \frac{\partial X}{\partial x_3} dx_3 \wedge dx_1 - \frac{\partial X}{\partial x_2} dx_1 \wedge dx_2
 \end{aligned}$$

同理可证

$$\begin{aligned}
 \int_{\partial\Sigma} Y dx_2 &= \int_{\Sigma} \frac{\partial Y}{\partial x_1} dx_1 \wedge dx_2 - \frac{\partial Y}{\partial x_3} dx_2 \wedge dx_3, \\
 \int_{\partial\Sigma} Z dx_3 &= \int_{\Sigma} \frac{\partial Z}{\partial x_2} dx_2 \wedge dx_3 - \frac{\partial Z}{\partial x_1} dx_3 \wedge dx_1,
 \end{aligned}$$

上述三式相加得到

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$

□

7.6 向量场的微积分

梯度、旋度和散度, nabla 算子和 Laplace 算子

对 C^1 函数 f , 梯度向量场和全微分

f (函数)	=====	f (0 阶微分形式)
梯度 ∇ ↓ 内积空间, 直角坐标系		d ↓ 任意空间, 任意坐标系
$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$		$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$

对 C^1 向量场 $\mathbf{F} = (X(x, y, z), Y(x, y, z), Z(x, y, z))^T$ 以及一阶微分形式 $\omega = X(x, y, z)dx + Y(x, y, z)dy + Z(x, y, z)dz$,

$$\begin{array}{ccc} \mathbf{F} = (X, Y, Z)^T & & \omega = Xdx + Ydy + Zdz \\ \text{旋度 rot} = \nabla \times \downarrow \text{内积空间, 直角坐标系} & & d \downarrow \text{任意空间, 任意坐标系} \\ \text{rot } \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X & Y & Z \end{vmatrix} & & d\omega = \begin{vmatrix} dy \wedge dz & dz \wedge dx & dx \wedge dy \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X & Y & Z \end{vmatrix} \end{array}$$

对 C^1 向量场 $\mathbf{F} = (X(x, y, z), Y(x, y, z), Z(x, y, z))^T$ 以及二阶微分形式 $\omega = X(x, y, z)dy \wedge dz + Y(x, y, z)dz \wedge dx + Z(x, y, z)dx \wedge dy$,

$$\begin{array}{ccc} \mathbf{F} = (X, Y, Z)^T & & \omega = Xdy \wedge dz + Ydz \wedge dx + Zdx \wedge dy \\ \text{散度 div} = \nabla \cdot \downarrow \text{内积空间, 直角坐标系} & & d \downarrow \text{任意空间, 任意坐标系} \\ \text{div } \mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} & & d\omega = \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right) dx \wedge dy \wedge dz \end{array}$$

对任何 C^2 函数 f ,

$$\nabla \times \nabla f = 0, \quad d(df) = 0.$$

这是因为

$$\nabla \times \nabla f = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} = 0.$$

因此所有梯度向量场都是无旋场。

对 C^2 函数 f , 称 f 的梯度向量场的散度为 f 的 **Laplacian**, 记为

$$\Delta f = \text{div } \nabla f = \nabla \cdot (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

满足 Laplace 方程 $\Delta f = 0$ 的函数称为调和函数。

对任何 C^2 向量场 \mathbf{F} 以及任何 C^2 一阶微分形式 ω ,

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0, \quad d(d\omega) = 0.$$

这是因为

$$\nabla \cdot (\nabla \times \mathbf{F}) = \nabla \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X & Y & Z \end{vmatrix} = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X & Y & Z \end{vmatrix} = 0.$$

因此任何旋度向量场都是无源场。

可以证明对任何 C^2 微分形式 ω , 都成立 $d(d\omega) = 0$ 。

梯度、旋度、散度的运算性质

在三维空间中,

$$\nabla \text{ 数量场} = \text{向量场}, \quad \nabla \times \text{向量场} = \text{向量场}, \quad \nabla \cdot \text{向量场} = \text{数量场},$$

并且满足

1. 线性: 对常数 α, β , C^1 函数 f, g 和 C^1 向量场 \mathbf{F}, \mathbf{G} ,

$$\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g,$$

$$\nabla \times (\alpha \mathbf{F} + \beta \mathbf{G}) = \alpha \nabla \times \mathbf{F} + \beta \nabla \times \mathbf{G},$$

$$\nabla \cdot (\alpha \mathbf{F} + \beta \mathbf{G}) = \alpha \nabla \cdot \mathbf{F} + \beta \nabla \cdot \mathbf{G}.$$

2. Leibniz性质: 对 C^1 函数 f, g 和 C^1 向量场 \mathbf{F}, \mathbf{G} ,

$$\nabla(fg) = f\nabla g + g\nabla f,$$

$$\nabla \times (g\mathbf{F}) = \nabla g \times \mathbf{F} + g\nabla \times \mathbf{F},$$

$$\nabla \cdot (g\mathbf{F}) = \nabla g \cdot \mathbf{F} + g\nabla \cdot \mathbf{F},$$

$$\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$$

微分形式与微分方程

一阶常微分方程

$$\frac{dy}{dx} = g(x, y)$$

可以改写为

$$g(x, y)dx - dy = 0$$

的形式。

一般地,

$$X(x, y)dx + Y(x, y)dy = 0$$

定义了一个一阶微分方程。它的解曲线 $(x(t), y(t))$ 满足

$$X(x(t), y(t))x'(t) + Y(x(t), y(t))y'(t) = 0, \quad \forall t.$$

在直角坐标系中这意味着向量场 $(X, Y)^T$ 总是与这些解曲线正交。

如果微分形式 $X(x, y)dx + Y(x, y)dy = df$ 是全微分, 则上述方程的解曲线就是 $f(x, y)$ 的水平集(即等值线)。这时我们称 $f(x, y) = C$ (C 是任意常数)是微分方程 $X(x, y)dx + Y(x, y)dy = 0$ 的通解。

如果微分形式 $X(x, y)dx + Y(x, y)dy$ 不是全微分, 但非零函数 $\mu(x, y)$ 满足 $\mu(x, y)X(x, y)dx + \mu(x, y)Y(x, y)dy$ 是全微分, 我们称 $\mu(x, y)$ 为微分方程 $X(x, y)dx + Y(x, y)dy = 0$ 的一个积分因子。此时, $\mu X dx + \mu Y dy$ 的原函数 f 给出方程的通解 $f(x, y) = C$ (C 是任意常数)。

例 7.6.1. 解方程 $(\ln y - \frac{y}{x}) dx + (\frac{x}{y} - \ln x) dy = 0$ 。

解: 记 $\omega = (\ln y - \frac{y}{x}) dx + (\frac{x}{y} - \ln x) dy$ 。则

$$d\omega = \left[-\frac{\partial}{\partial y} \left(\ln y - \frac{y}{x} \right) + \frac{\partial}{\partial x} \left(\frac{x}{y} - \ln x \right) \right] dx \wedge dy = 0.$$

所以由Green公式知, ω 在第一象限内是一个全微分。

我们可以靠观察得到 ω 的原函数

$$x \ln y - y \ln x.$$

也可以通过沿折线 $(1, 1) \rightarrow (x_0, 1) \rightarrow (x_0, y_0)$ 将 ω 积分

$$\begin{aligned} \int_1^{x_0} \omega(x, 1) + \int_1^{y_0} \omega(x_0, y) &= \int_1^{x_0} -\frac{1}{x} dx + \int_1^{y_0} \frac{x_0}{y} - \ln x_0 dy \\ &= -\ln x_0 + x_0 \ln y_0 - (y_0 - 1) \ln x_0 \\ &= x_0 \ln y_0 - y_0 \ln x_0 \end{aligned}$$

这也可以得到 ω 的上述原函数。

因此原方程通解为 $x \ln y - y \ln x = C$ 。 □

例 7.6.2. 解方程 $(x + y)dx + (y - x)dy = 0$ 。

解: 记 $\omega = (x + y)dx + (y - x)dy$ 。则 $d\omega = -2dx \wedge dy \neq 0$ 。所以 ω 不是全微分。

设 $\mu(x, y)$ 是 ω 的一个积分因子。则

$$\begin{aligned} 0 = d(\mu\omega) &= \left[\frac{\partial}{\partial x} (\mu(x, y)(y - x)) - \frac{\partial}{\partial y} (\mu(x, y)(x + y)) \right] dx \wedge dy \\ &= \left[(y - x) \frac{\partial \mu}{\partial x} - (x + y) \frac{\partial \mu}{\partial y} - 2\mu \right] dx \wedge dy, \end{aligned}$$

因此 $\mu(x, y)$ 是 ω 的一个积分因子, 当且仅当

$$(y - x) \frac{\partial \mu}{\partial x} - (x + y) \frac{\partial \mu}{\partial y} = 2\mu,$$

这是一个一阶线性偏微分方程。我们打算去求解这个偏微分方程。

原方程关于 (x, y) 是齐次的, 即对任意常数 λ , $(\lambda x, \lambda y)$ 也满足同样的微分方程。因此考虑极坐标系 $x = r \cos \theta, y = r \sin \theta$, 代入方程得到

$$r(\cos \theta + \sin \theta)(\cos \theta dr - r \sin \theta d\theta) + r(\sin \theta - \cos \theta)(\sin \theta dr + r \cos \theta d\theta) = 0,$$

整理化简得到

$$dr - r d\theta = 0.$$

从这里容易看到 $\mu = \frac{1}{r}$ 是它的一个积分因子,

$$0 = \frac{1}{r} dr - d\theta = d(\ln r - \theta).$$

此方程通解为 $\ln r = \theta + C$, 原方程通解为 $x = Ae^\theta \cos \theta, y = Ae^\theta \sin \theta$, 这是指数螺旋线。 □

7.7 习题讨论课8

1. 设 $a > 1$, 空间中的有向曲线 γ 为上半球面 $x^2 + y^2 + z^2 = 2ax (z \geq 0)$ 与圆柱面 $x^2 + y^2 = 2x$ 的交线, 它在 (x, y) 坐标平面中的投影曲线逆时针旋转. 求积分 $\int_{\gamma} (y^2 + z^2)dx + (z^2 + x^2)dy + (x^2 + y^2)dz$ 的值.

解: 沿曲线 γ , $y^2 + z^2 = 2ax - x^2$, $z^2 + x^2 = 2ax - y^2$, $x^2 + y^2 = 2x$, $z^2 = 2ax - 2x$, 所以

$$\begin{aligned}\int_{\gamma} (y^2 + z^2)dx &= \int_{\gamma} (2ax - x^2)dx = 0, \\ \int_{\gamma} (z^2 + x^2)dy &= \int_{\gamma} (2ax - y^2)dy = 2a \int_{\gamma} xdy, \\ \int_{\gamma} (x^2 + y^2)dz &= \int_{\gamma} 2xdy = 2a \int_{\gamma} x d\sqrt{2(a-1)x} = 0,\end{aligned}$$

因此

$$\int_{\gamma} (y^2 + z^2)dx + (z^2 + x^2)dy + (x^2 + y^2)dz = 2a \int_{\gamma} xdy,$$

取参数方程

$$\begin{cases} x = 1 + \cos \theta, \\ y = \sin \theta, \\ z = \sqrt{2ax - 2x} = \sqrt{2(a-1)(1 + \cos \theta)}, \end{cases} \quad 0 \leq \theta \leq 2\pi.$$

从而

$$2a \int_{\gamma} xdy = 2a \int_0^{2\pi} (1 + \cos \theta) d \sin \theta = 2\pi a.$$

也可以用Stokes公式

$$2a \int_{\gamma} xdy = 2a \int_{\Sigma} dx \wedge dy = 2a \int_{x^2 + y^2 \leq 2x} dx dy = 2\pi a.$$

□

2.

第8章 级数

8.1 级数的概念

定义 8.1.1. 给定线性空间 V 中一个序列 $\{a_k\}_{k=1}^{\infty}$, 称 $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots$ 是一个级数, 称 a_n 为这个级数的第 n 项。对正整数 N , $\sum_{k=1}^N a_k = a_1 + \cdots + a_N$ 称为这个级数的前 N 项和, 记为 S_N 。

如果存在 $S \in V$ 使得 $\lim_{N \rightarrow +\infty} S_N = S$, 则称级数 $\sum_{k=1}^{\infty} a_k$ 收敛, 且称 S 为级数 $\sum_{k=1}^{\infty} a_k$ 的和, 记 $S = \sum_{k=1}^{\infty} a_k$ 。

这里 $\lim_{N \rightarrow +\infty} S_N = S$ 的含义是: 对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $N > N(\varepsilon)$, $\|S_N - S\| < \varepsilon$ 。其中 $\|\cdot\|$ 是 V 中的范数。

不收敛的级数称为是发散的级数。

注 8.1.2. (1) 当 $V = \mathbb{R}$ 或 \mathbb{C} 时, 可以取绝对值 $|\cdot|$ 为范数, 此时级数称为“数项级数”。

(2) 对于 \mathbb{R}^m 的有界闭集 K , 考虑 $V = C(K)$, 即 K 上所有连续函数所组成的线性空间, 取范数为

$$\|f\|_{\infty} = \max_{\mathbf{x} \in K} |f(\mathbf{x})|,$$

此时的级数是“函数项级数”, 并且称收敛的级数为在 K 上一致收敛。

(3) 对于 $V = R[a, b]$, 即区间 $[a, b]$ 上所有 Riemann 可积函数所组成的线性空间, 取内积和相应的范数为

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx, \quad \|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx},$$

其中 $\overline{g(x)}$ 是复数 $g(x)$ 共轭。在此意义下收敛的函数项级数称为在区间 $[a, b]$ 上均方收敛。

例 8.1.3 (几何级数). 设 $|x| < 1$, 则

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots.$$

证明.

$$(1-x)(1+x+x^2+\cdots+x^{N-1}) = 1-x^N,$$

因为 $|x| < 1$, 所以 $1-x$ 可逆, 且

$$|S_N - (1-x)^{-1}| = |(1-x)^{-1}(-x^N)| \leq |(1-x)^{-1}||x|^N \rightarrow 0, \quad N \rightarrow +\infty,$$

因此

$$\lim_{N \rightarrow +\infty} S_N = (1-x)^{-1}.$$

□

注: 上述结论对任意实数或复数 x (此时 $|x|$ 是 x 的绝对值)、任意方阵 x (此时 $|x|$ 是矩阵 x 的算子范数) 都成立。

例 8.1.4.

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1.$$

证明.

$$\sum_{k=1}^N \frac{1}{k(k+1)} = \sum_{k=1}^N \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{N+1} \rightarrow 1, \quad N \rightarrow +\infty.$$

□

定理 8.1.5. 设级数 $\sum_{k=1}^{\infty} a_k$ 和 $\sum_{k=1}^{\infty} b_k$ 都收敛, 则

1. $\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k)$ 收敛, 且

$$\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{\infty} a_k + \beta \sum_{k=1}^{\infty} b_k.$$

2. 如果 a_k, b_k 都是实数, 且对所有 $k \geq 1$, $a_k \leq b_k$, 则 $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$ 。

一般而言, 能够用上面这样的初等方法求和的级数是少而又少的。虽然通过求导和积分我们能够得到更多一些级数的和, 但能精确得到级数的和仍是一个奢望。如何在无法获知级数的和的情况下判定级数的收敛性是一个重要的问题。

定义 8.1.6. 称级数 $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots$ 满足 **Cauchy 性质**, 如果对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $m > n > N(\varepsilon)$, $\left\| \sum_{k=n}^m a_k \right\| < \varepsilon$.

定理 8.1.7 (Cauchy 收敛准则). 如果线性空间 V 在范数 $\|\cdot\|$ 下是完备的(即任意 Cauchy 序列都收敛), 则级数 $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots$ 收敛当且仅当它满足 Cauchy 性质。

推论 8.1.8. 任何实数(或复数)级数 $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots$, 它收敛当且仅当它满足 Cauchy 性质。

推论 8.1.9. 如果级数 $\sum_{k=1}^{\infty} a_k$ 收敛, 则 $\lim_{n \rightarrow +\infty} \|a_n\| = 0$.

例 8.1.10 (几何级数). 对 $x \in \mathbb{C}$, $\sum_{n=0}^{\infty} x^n$ 收敛当且仅当 $|x| < 1$. □

例 8.1.11 (调和级数). 我们即将知道 $\sum_{n=1}^{\infty} \frac{1}{n}$ 是发散的。这说明 $\lim_{n \rightarrow +\infty} \|a_n\| = 0$ 仅仅是级数收敛的必要条件, 而不是充分条件。 □

8.2 绝对收敛与比较法

以下我们总假设空间是完备的。

定义 8.2.1. 称级数 $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots$ **绝对收敛**, 若级数 $\sum_{k=1}^{\infty} \|a_k\|$ 收敛。

定理 8.2.2. 绝对收敛的级数都是收敛的。

证明. 设 $\sum_{k=1}^{\infty} \|a_k\|$ 收敛, 则 $\sum_{k=1}^{\infty} \|a_k\|$ 具有 Cauchy 性质, 因此对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $m > n > N(\varepsilon)$, $\sum_{k=n}^m \|a_k\| < \varepsilon$. 从而

$$\left\| \sum_{k=n}^m a_k \right\| \leq \sum_{k=n}^m \|a_k\| < \varepsilon,$$

从而 $\sum_{k=1}^{\infty} a_k$ 具有 Cauchy 性质, 因而收敛。 □

比较法是判别级数是否绝对收敛的最常用的办法。

定理 8.2.3 (比较法). 设 $a_k = O(b_k)$ ($k \rightarrow +\infty$), 即存在 K_0 和 $M > 0$ 使得 $\|a_k\| \leq M\|b_k\|$ ($\forall k \geq K_0$).

1. 如果 $\sum_{k=1}^{\infty} \|b_k\|$ 收敛, 则 $\sum_{k=1}^{\infty} \|a_k\|$ 收敛。

2. 如果 $\sum_{k=1}^{\infty} \|a_k\|$ 发散, 则 $\sum_{k=1}^{\infty} \|b_k\|$ 发散。

证明. (1) 因为 $\sum_{k=1}^{\infty} \|b_k\|$ 收敛, 所以对任意 $\varepsilon > 0$, 存在 $N(\varepsilon) > K_0$ 使得对任意正整数 $m > n > N(\varepsilon)$, $\sum_{k=n}^m \|b_k\| < \frac{\varepsilon}{M}$. 从而

$$\sum_{k=n}^m \|a_k\| \leq \sum_{k=n}^m M\|b_k\| < \varepsilon,$$

从而 $\sum_{k=1}^{\infty} \|a_k\|$ 收敛。

(2) 这是(1)的逆否命题。 □

推论 8.2.4. 若 $\lim_{k \rightarrow +\infty} \frac{\|a_k\|}{\|b_k\|} < +\infty$, 则上述定理结论成立。

因此当我们希望利用比较法证明一个级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛时, 适当选取或构造一个参照级数 $\sum_{k=1}^{\infty} b_k$ 就是核心任务。以下的 D'Alembert 和 Cauchy 的判别法就是利用几何级数作为比较的参照对象。

定理 8.2.5. 1. (D'Alembert 分式判别法) 如果存在 $0 < r < 1$ 以及 K_0 使得对任意 $k \geq K_0$,

$$\frac{\|a_{k+1}\|}{\|a_k\|} < r,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛。

如果

$$\lim_{k \rightarrow +\infty} \frac{\|a_{k+1}\|}{\|a_k\|} < 1,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛。

2. 如果

$$\lim_{k \rightarrow +\infty} \frac{\|a_{k+1}\|}{\|a_k\|} > 1,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 发散。

3. (Cauchy 根式判别法) 如果存在 $0 < r < 1$ 以及 K_0 使得对任意 $k \geq K_0$,

$$\|a_k\|^{\frac{1}{k}} < r,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛。

如果

$$\lim_{k \rightarrow +\infty} \|a_k\|^{\frac{1}{k}} < 1,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛。

4. 如果存在无穷多个 k 使得

$$\|a_k\|^{\frac{1}{k}} \geq 1,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 发散。

如果

$$\lim_{k \rightarrow +\infty} \|a_k\|^{\frac{1}{k}} > 1,$$

则级数 $\sum_{k=1}^{\infty} a_k$ 发散。

证明. (1) 对任意 $k > K_0$,

$$\frac{\|a_k\|}{\|a_{K_0}\|} = \frac{\|a_{K_0+1}\|}{\|a_{K_0}\|} \frac{\|a_{K_0+2}\|}{\|a_{K_0+1}\|} \cdots \frac{\|a_k\|}{\|a_{k-1}\|} < r^{k-K_0},$$

从而 $a_k = O(r^k)$ ($k \rightarrow +\infty$)。因为 $0 < r < 1$, 所以 $\sum_{k=1}^{\infty} r^k$ 收敛。于是由比较法

知级数 $\sum_{k=1}^{\infty} a_k$ 绝对收敛。

其他部分的证明留给读者完成。 \square

例 8.2.6. 对任意 x , 级数

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

绝对收敛。

证明. 记 $a_k = \frac{x^k}{k!}$ 。则 $\|a_k\| \leq \frac{\|x\|^k}{k!} = b_k$,

$$\frac{b_{k+1}}{b_k} = \frac{\|x\|}{k+1} < \frac{1}{2}, \quad \forall k > 2\|x\|.$$

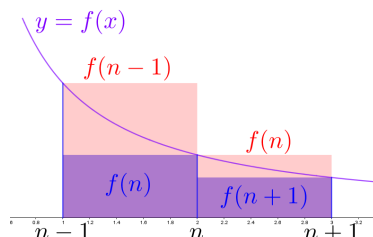
所以 $\sum_{k=0}^{\infty} b_k$ 收敛, 从而 $\sum_{k=0}^{\infty} a_k$ 绝对收敛。 \square

定理 8.2.7 (积分判别法). 设 $f: (0, +\infty) \rightarrow \mathbb{R}$ 非负、单调减, 则 $\int_1^{+\infty} f(x)dx$ 收敛当且仅当 $\sum_{k=1}^{\infty} f(k)$ 收敛。

证明. 记 $b_k = \int_{k-1}^k f(x)dx$. 则

$$f(k-1) \geq b_k \geq f(k), \quad \forall k \geq 1.$$

因此 $\sum_{k=1}^{\infty} f(k)$ 收敛当且仅当 $\sum_{k=1}^{\infty} b_k$ 收敛。



由

$$b_n + b_{n+1} + \cdots + b_m = \int_{n-1}^m f(x)dx$$

易知 $\int_1^{+\infty} f(x)dx$ 满足Cauchy性质当且仅当 $\sum_{k=1}^{\infty} b_k$ 满足Cauchy性质。

从而 $\int_1^{+\infty} f(x)dx$ 收敛当且仅当 $\sum_{k=1}^{\infty} f(k)$ 收敛。 \square

例 8.2.8. 对任意实数 p , 级数

$$\sum_{n=1}^{\infty} \frac{1}{n^p}, \quad \sum_{n=1}^{\infty} \frac{1}{n(\ln n)^p}, \quad \sum_{n=1}^{\infty} \frac{1}{n \ln(\ln n)^p}, \quad \cdots$$

中任何一个收敛都是当且仅当 $p > 1$ 。

证明. 考虑 $\int_1^{\infty} \frac{1}{x^p} dx, \int_3^{\infty} \frac{1}{x(\ln x)^p} dx, \cdots$. \square

另外, 我们还可以利用函数的渐近展开来研究级数的收敛性, 这样的办法称为比阶法, 它不仅可用于判别级数的绝对收敛性, 也可用于判别一般的收敛和发散性质。

例 8.2.9. 判断 $\sum_{n=1}^{\infty} \left(n \ln \frac{2n+1}{2n-1} - 1 \right)$ 的敛散性。

解:

$$\begin{aligned} n \ln \frac{2n+1}{2n-1} - 1 &= n \ln \frac{1 + \frac{1}{2n}}{1 - \frac{1}{2n}} - 1 \\ &= n \left[\ln \left(1 + \frac{1}{2n} \right) - \ln \left(1 - \frac{1}{2n} \right) \right] - 1 \\ &= n \left[\frac{1}{2n} - \frac{1}{4n^2} + O\left(\frac{1}{n^3}\right) - \left(-\frac{1}{2n} - \frac{1}{4n^2} + O\left(\frac{1}{n^3}\right) \right) \right] - 1 \\ &= O\left(\frac{1}{n^2}\right), \quad n \rightarrow +\infty, \end{aligned}$$

因为 $\sum_{n=1}^{\infty} \frac{1}{n^2}$ 收敛, 所以 $\sum_{n=1}^{\infty} \left(n \ln \frac{2n+1}{2n-1} - 1 \right)$ 绝对收敛。 \square

借助 $\sum_{n=1}^{\infty} \frac{1}{n^p}$ 收敛性, 可以结合比较法得到一些相应的收敛判别法。
对 $a_n = \frac{1}{n^p}$,

$$\frac{a_n}{a_{n+1}} = \left(1 + \frac{1}{n} \right)^p = 1 + \frac{p}{n} + o\left(\frac{1}{n}\right), \quad n \rightarrow +\infty.$$

所以

$$p = n \left(\frac{a_n}{a_{n+1}} - 1 \right) + o(1), \quad n \rightarrow +\infty.$$

定理 8.2.10 (Raabe判别法). 如果存在正整数 N_0 和实数 $p > 1$ 使得当 $n > N_0$ 时

$$n \left(\frac{\|a_n\|}{\|a_{n+1}\|} - 1 \right) > p,$$

则级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛。

如果存在正整数 N_0 和实数 $p < 1$ 使得当 $n > N_0$ 时

$$n \left(\frac{\|a_n\|}{\|a_{n+1}\|} - 1 \right) < p,$$

则级数 $\sum_{n=1}^{\infty} \|a_n\|$ 发散。

证明. 取 $q \in (1, p)$ 。则存在 $N_1 > N_0$ 使得对任意 $n \geq N_1$,

$$\frac{(n+1)^q}{n^q} = \left(1 + \frac{1}{n} \right)^q = 1 + \frac{q}{n} + o\left(\frac{1}{n}\right) < 1 + \frac{p}{n} < \frac{\|a_n\|}{\|a_{n+1}\|}.$$

从而

$$\|a_n\| \leq \frac{\|a_{N_1}\| N_1^q}{n^q}.$$

因为 $\sum_{n=1}^{\infty} \frac{1}{n^q}$ 收敛, 所以 $\sum_{n=1}^{\infty} a_n$ 绝对收敛。 \square

例 8.2.11. 证明 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^\alpha + (-1)^{n-1}}$ 绝对收敛当且仅当 $\alpha > 1$ 。

证明.

$$\begin{aligned} n \left(\frac{|a_n|}{|a_{n+1}|} - 1 \right) &= n \left(\frac{(n+1)^\alpha + (-1)^n}{n^\alpha + (-1)^{n-1}} - 1 \right) \\ &= n \left(\frac{\left(1 + \frac{1}{n}\right)^\alpha + \frac{(-1)^n}{n^\alpha}}{1 - \frac{(-1)^n}{n^\alpha}} - 1 \right) = \alpha + o(1), \quad n \rightarrow +\infty \end{aligned}$$

所以由 Raabe 判别法, 当 $\alpha > 1$ 时, 级数绝对收敛; 当 $\alpha < 1$ 时, 级数不是绝对收敛。当 $\alpha = 1$ 时, $\left| \frac{(-1)^{n-1}}{n+(-1)^{n-1}} \right| \geq \frac{1}{n+1}$, 级数不是绝对收敛。 \square

8.3 条件收敛的级数及其收敛判别法

定义 8.3.1. 称级数 $\sum_{k=1}^{\infty} a_k$ 条件收敛, 如果它收敛但不是绝对收敛。

不变(正负)号的实数项级数必然是绝对收敛的。因此, 条件收敛的实数项级数, 它的项必然总是不断变号的, 即对任意 $N \geq 1$, $\{a_n\}_{n \geq N}$ 中总是既有正数又有负数。

定理 8.3.2. 如果以下三组条件之一成立, 则级数 $\sum_{n=1}^{\infty} a_n b_n$ 收敛。

1. (Hardy)

$$(a) \sum_{n=1}^N a_n \text{ 有界};$$

$$(b) \sum_{n=2}^{\infty} |b_n - b_{n-1}| \text{ 收敛且 } \lim_{n \rightarrow +\infty} b_n = 0.$$

2. (Dirichlet)

$$(a) \sum_{n=1}^N a_n \text{ 有界};$$

$$(b) b_n \text{ 单调且 } \lim_{n \rightarrow +\infty} b_n = 0.$$

3. (Abel)

$$(a) \sum_{n=1}^{\infty} a_n \text{ 收敛};$$

$$(b) b_n \text{ 单调且有界}.$$

证明. (1) 证明 Hardy 判别法。记 $A_N = \sum_{n=1}^N a_n$, 于是存在 $M > 0$ 使得对任意正整数 N , 都有 $|A_N| \leq M$ 。对任意 $\varepsilon > 0$, 存在 N_ε 使得对任意 $N \geq N_\varepsilon$, 则 $|b_N| < \frac{\varepsilon}{3M}$, $\sum_{n=N+1}^{N+p} |b_n - b_{n-1}| < \frac{\varepsilon}{3M}$ 。于是对任意正整数 p ,

$$\begin{aligned} \left| \sum_{n=N+1}^{N+p} a_n b_n \right| &= \left| \sum_{n=N+1}^{N+p} (A_n - A_{n-1}) b_n \right| \\ &= \left| A_{N+p} b_{N+p+1} - A_N b_N - \sum_{n=N+1}^{N+p+1} A_{n-1} (b_n - b_{n-1}) \right| \\ &\leq M |b_{N+p+1}| + M |b_N| + \sum_{n=N+1}^{N+p+1} M |b_{n-1} - b_n| < \varepsilon, \end{aligned}$$

因为 $\lim_{n \rightarrow +\infty} b_n = 0$, 所以级数 $\sum_{n=1}^{\infty} a_n b_n$ 满足 Cauchy 条件, 从而收敛。

(2) 证明 Hardy \Rightarrow Dirichlet。不妨设 b_n 单调减。则

$$\sum_{n=2}^N |b_n - b_{n-1}| = \sum_{n=2}^N (b_{n-1} - b_n) = b_1 - b_N \rightarrow b_1.$$

所以 $\sum_{n=2}^N |b_n - b_{n-1}|$ 收敛。

(3) 证明 Dirichlet \Rightarrow Abel。记 $b = \lim_{n \rightarrow +\infty} b_n$, $c_n = b_n - b$ 。因此 c_n 单调且 $\lim_{n \rightarrow +\infty} c_n = 0$ 。因为 $\sum_{n=1}^{\infty} a_n$ 收敛, 所以 $\sum_{n=1}^N a_n$ 有界。因此由 Dirichlet 判别法知级数 $\sum_{n=1}^{\infty} a_n c_n$ 收敛。由

$$\sum_{n=1}^N a_n b_n = \sum_{n=1}^N a_n (c_n + b) = \sum_{n=1}^N a_n c_n + b \sum_{n=1}^N a_n$$

知级数 $\sum_{n=1}^{\infty} a_n b_n$ 收敛。 \square

注 8.3.3. (1) 上述证明的方法与广义积分收敛的 Dirichlet-Abel 判别法的证明类似, 使用了离散形式的“分部积分”。

(2) Dirichlet-Abel 判别法适用于 a_n 是复数甚至是向量或矩阵的情形, 然而由于单调性假设, b_n 必须是实数。

(3) 如果 $\sum_{n=1}^N \|a_n\|$ 有界, 则因为它关于 N 单调不减, 所以级数 $\sum_{n=1}^{\infty} \|a_n\|$ 收敛。此时因为 $|b_n| \leq K (\forall n \geq 1)$, 所以

$$\|a_n b_n\| \leq K \|a_n\|, \quad \forall n \geq 1.$$

所以级数 $\sum_{n=1}^{\infty} a_n b_n$ 绝对收敛。因此 Dirichlet-Abel 判别法通常用于当 $\sum_{n=1}^{\infty} a_n$ 不是绝对收敛时级数 $\sum_{n=1}^{\infty} a_n b_n$ 的收敛性。

例 8.3.4. 对复数 z , 级数 $\sum_{n=1}^{\infty} \frac{z^n}{n}$ 收敛当且仅当 $|z| \leq 1$ 且 $z \neq 1$ 。

证明. 根据 Cauchy 根式判别法, 因为

$$\lim_{n \rightarrow +\infty} \left| \frac{z^n}{n} \right|^{1/n} = |z|,$$

所以当 $|z| < 1$ 时, 级数 $\sum_{n=1}^{\infty} \frac{z^n}{n}$ 绝对收敛; 当 $|z| > 1$ 时, 级数 $\sum_{n=1}^{\infty} \frac{z^n}{n}$ 发散。

当 $z = 1$ 时, $\sum_{n=1}^{\infty} \frac{z^n}{n} = \sum_{n=1}^{\infty} \frac{1}{n}$ 发散。

若 $|z| = 1$ 且 $z \neq 1$, 则

$$\left| \sum_{n=1}^N z^n \right| = |(1-z)^{-1}(z - z^{N+1})| \leq 2|(1-z)^{-1}|,$$

所以由 Dirichlet 判别法知级数 $\sum_{n=1}^{\infty} \frac{z^n}{n}$ 收敛, 但它不绝对收敛。 \square

以下的交错级数 Leibniz 判别法是 Dirichlet 判别法的一个很特殊的情况。

定义 8.3.5. 称级数 $\sum_{k=1}^{\infty} (-1)^k b_k$ 为一个交错级数, 其中所有 $b_k \geq 0$ 。也就是这个级数的项正负交替出现。

推论 8.3.6 (Leibniz 判别法). 设级数 $\sum_{k=1}^{\infty} (-1)^k b_k$ 满足: $b_1 \geq b_2 \geq \dots \geq b_n \geq \dots \geq 0$ 。则该级数收敛当且仅当 $\lim_{k \rightarrow +\infty} b_k = 0$ 。 \square

例 8.3.7. 设 $\alpha > 0$, 讨论 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha} + (-1)^{n-1}}$ 的收敛性。

解: 这是一个交错级数, 但不满足 Leibniz 判别法中的单调性条件。我们对通项进行渐近展开,

$$a_n = \frac{(-1)^{n-1}}{n^{\alpha}} \frac{1}{1 + \frac{(-1)^{n-1}}{n^{\alpha}}} = \frac{(-1)^{n-1}}{n^{\alpha}} - \frac{1 + o(1)}{n^{2\alpha}} = \frac{(-1)^{n-1}}{n^{\alpha}} - b_n.$$

而

$$\lim_{n \rightarrow +\infty} \frac{b_n}{\frac{1}{n^{2\alpha}}} = 1,$$

所以, 级数 $\sum_{n=1}^{\infty} b_n$ (绝对) 收敛当且仅当级数 $\sum_{n=1}^{\infty} \frac{1}{n^{2\alpha}}$ (绝对) 收敛, 后者当且仅当 $\alpha > \frac{1}{2}$ 。

当 $\alpha > 1$ 时, 级数 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha}}$, $\sum_{n=1}^{\infty} b_n$ 都是绝对收敛的, 所以原级数绝对收敛。

当 $\frac{1}{2} < \alpha \leq 1$ 时, 级数 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha}}$ 条件收敛, 级数 $\sum_{n=1}^{\infty} b_n$ 绝对收敛, 所以原级数条件收敛。

当 $\alpha \leq \frac{1}{2}$ 时, 级数 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha}}$ 条件收敛, 级数 $\sum_{n=1}^{\infty} b_n$ 发散, 所以原级数发散。 \square

注 8.3.8. 对 $0 < \alpha \leq \frac{1}{2}$, 虽然级数 $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^{\alpha}}$ (条件) 收敛, 并且 $\frac{(-1)^{n-1}}{n^{\alpha} + (-1)^{n-1}} = \frac{(-1)^n}{n^{\alpha}}(1 + o(1))$, 但是 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{\alpha} + (-1)^{n-1}}$ 发散。这说明适用于绝对收敛级数的比较法并不适用于条件收敛的级数。

8.4 级数的交换律、结合律

结合律

加法结合律就是不改变级数项的排列次序，但让其中一些位置相邻的项先加起来，最后的结果不改变。这对有限多个数的加法是成立的，我们看看它是否对级数(无穷多个数排列相加)也成立。

例 8.4.1. $1 + (-1) + 1 + (-1) + \cdots$ ，即级数 $\sum_{n=1}^{\infty} (-1)^{n-1}$ ，它是发散的。但是

$$[1 + (-1)] + [1 + (-1)] + \cdots$$

是收敛的，和为 0。而

$$1 + [(-1) + 1] + [(-1) + 1] + \cdots$$

也是收敛的，和却为 1。

因此不是所有级数都满足结合律。

是否有可能通过适当加括号分组让和取其他值或无穷大呢？事实上，上述级数的部分和数列

$$S_n = \begin{cases} 1, & n \text{ 是奇数,} \\ 0, & n \text{ 是偶数.} \end{cases}$$

定理 8.4.2. 收敛级数满足结合律。

证明. 因为分组加括号后得到的新级数的部分和数列是原来级数部分和数列的子列。□

定理 8.4.3. 存在这样一个级数 $\sum_{n=1}^{\infty} a_n$ ，满足：对任意实数 $x \in \mathbb{R}$ ，存在一个分组方式使得原级数在该分组下的新级数收敛，且和为 x ；同时也存在一个分组方式，使得这个级数在相应分组下的新级数的部分和数列无界(或趋于无穷大)。

证明. 请构造一个满足要求的级数。□

交换律

所谓交换律就是：对正整数集的任何一个排列 $\sigma: \mathbb{N}^* \rightarrow \mathbb{N}^*$ ，是否级数 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 与级数 $\sum_{n=1}^{\infty} a_n$ 总是具有相同的敛散性，且它们收敛时和相等？

级数交换律的问题要远比结合律复杂，因为允许交换无穷多项的次序，所以 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 的部分和数列有可能与 $\sum_{n=1}^{\infty} a_n$ 的部分和数列没有任何关系。

定理 8.4.4. 绝对收敛的级数满足交换律。

证明. 设级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛, 其和为 S , 其部分和 $S_n = \sum_{k=1}^n a_k$. 设 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 是 $\sum_{n=1}^{\infty} a_n$ 的任一重排, 其部分和为 $\tilde{S}_m = \sum_{k=1}^m a_{\sigma(k)}$.

对任意 $\varepsilon > 0$, 因为 $\sum_{n=1}^{\infty} a_n$ 绝对收敛, 所以存在正整数 n_ε 使得 $\sum_{k=n_\varepsilon+1}^{\infty} \|a_k\| < \varepsilon$, 因此 $\|S_{n_\varepsilon} - S\| \leq \sum_{k=n_\varepsilon+1}^{\infty} \|a_k\| < \varepsilon$.

对任意 n , 记 $A_n = \{1, 2, \dots, n\}$, $N_n = \max \sigma^{-1}(A_n)$.

于是对任意 $m \geq N_n$, $\sigma^{-1}(A_n) \subseteq A_{N_n} \subseteq A_m$, 从而 $A_n \subseteq \sigma(A_m)$. 于是,

$$\tilde{S}_m = \sum_{k=1}^m a_{\sigma(k)} = \sum_{j \in \sigma(A_m)} a_j = \sum_{j=1}^n a_j + \sum_{j \in \sigma(A_m) \setminus A_n} a_j = S_n + \sum_{j \in \sigma(A_m) \setminus A_n} a_j,$$

从而对任意 $m \geq N_{n_\varepsilon}$,

$$\begin{aligned} \|\tilde{S}_m - S\| &\leq \|\tilde{S}_m - S_{n_\varepsilon}\| + \|S_{n_\varepsilon} - S\| \\ &\leq \sum_{j \in \sigma(A_m) \setminus A_{n_\varepsilon}} \|a_j\| + \|S_{n_\varepsilon} - S\| \\ &\leq \sum_{j=n_\varepsilon+1}^{\infty} \|a_j\| + \|S_{n_\varepsilon} - S\| < 2\varepsilon. \end{aligned}$$

因此级数 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 收敛, 且 $\sum_{n=1}^{\infty} a_{\sigma(n)} = S$. 对 $\sum_{n=1}^{\infty} \|a_{\sigma(n)}\|$ 使用上述证明, 得到 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 绝对收敛. \square

接下来我们考察无穷多次改变正负号且不是绝对收敛的实数级数.

首先, 不是每个这样的级数都可以重排为收敛级数, 例如 $\sum_{n=1}^{\infty} (-1)^n n$, 这个级数无论怎样排列都是发散的——因为对任意重排 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 都有 $|a_{\sigma(n)}| \geq 1$; 而 $\lim_{n \rightarrow +\infty} a_{\sigma(n)} = 0$ 是级数 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 收敛的必要条件.

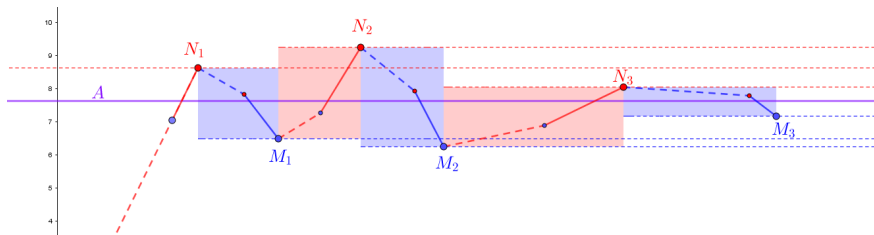
定理 8.4.5. 对任何条件收敛的实数项级数 $\sum_{n=1}^{\infty} a_n$, 以及任何 $A \in \mathbb{R} \cup \{+\infty, -\infty\}$, 都存在重排 $\sum_{n=1}^{\infty} a_{\sigma(n)}$ 使得 $\lim_{N \rightarrow +\infty} \sum_{n=1}^N a_{\sigma(n)} = A$.

证明. 因为级数 $\sum_{n=1}^{\infty} a_n$ 条件收敛, 所以它有无穷多个非负项和无穷多个负项. 数列 a_n 的所有非负项组成子列 a_{p_1}, a_{p_2}, \dots , 所有负数项组成子列 a_{n_1}, a_{n_2}, \dots . 从而 $\lim_{k \rightarrow +\infty} p_k = +\infty$, $\lim_{k \rightarrow +\infty} n_k = +\infty$.

级数 $\sum_{k=1}^{\infty} a_{p_k} = +\infty$, $\sum_{k=1}^{\infty} a_{n_k} = -\infty$. (假若其中有一个收敛, 则另一个必然也收敛, 并且两个级数都是绝对收敛. 这样 $\sum_{n=1}^{\infty} a_n$ 绝对收敛, 但这与已知矛盾.)

任取实数 A 。我们按以下规则每次从上述两个子列中按序列顺序取出一些项求和, 这样得到的级数是 $\sum_{n=1}^{\infty} a_n$ 的一个重排。我们证明最终得到的级数和为 A 。我们记这个级数的部分和为 S_n 。

最开始时令 $S_0 = 0$ 。每当 $S_n < A$ 时, 就从非负项子列再取一些项加入 S_n , 从而 S_n 不断增长, 直到 S_n 大于 A ; 每当 $S_n > A$ 时, 就从负项子列再取一些项加入 S_n , 使得 S_n 不断减少, 直到 S_n 小于 A 。



这样就得到正整数列 $N_1 < M_1 < N_2 < M_2 < \dots$ 使得部分和序列 S_n 在 $[N_k, M_k]$ 上严格减, 在 $[M_k, N_{k+1}]$ 上单调不减, 对任意 $n \in [N_k, M_k)$ 及任意 $m \in [M_k, N_{k+1})$, $S_{N_k} \geq S_n \geq S_{M_{k-1}} > A \geq S_{N_{k+1}-1} \geq S_m \geq S_{M_k}$ 。因此

$$|S_n - A| \leq a_{\sigma(N_k)}, \quad |S_m - A| \leq a_{\sigma(M_k)}.$$

因为 $\lim_{j \rightarrow +\infty} a_{\sigma(j)} = 0$, 所以 $\lim_{n \rightarrow +\infty} S_n = A$ 。

其余证明留给读者完成。 \square

级数的乘积

定理 8.4.6. 如果级数 $\sum_{n=0}^{\infty} a_n$ 和 $\sum_{n=0}^{\infty} b_n$ 都绝对收敛, 则级数 $\sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right)$ 绝对收敛, 且

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) = \left(\sum_{n=0}^{\infty} a_n \right) \cdot \left(\sum_{n=0}^{\infty} b_n \right).$$

例 8.4.7. 我们已经知道级数 $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ 对所有 x 绝对收敛, 所以根据上述定理和 Newton 二项式定理,

$$\left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} \right) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}.$$

即 $\exp(x) \cdot \exp(y) = \exp(x+y)$ 。

这件事对所有复数 x, y 成立, 对任何两个可交换(即满足 $xy = yx$)的矩阵 x, y 也成立。

另外, 易见 $\exp(1) = e$ 。稍后我们会证明对实数 x , $\exp(x) = e^x$ 。 \square

上面的那个定理是以下定理的特殊情形。

定理 8.4.8. 如果级数 $\sum_{n=0}^{\infty} \left(\sum_{i+j=n} |a_{i,j}| \right)$ 收敛, 则对于任何一一对应 $\sigma: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, 级数 $\sum_{n=0}^{\infty} a_{\sigma(n)}$ 绝对收敛, 且

$$\sum_{n=0}^{\infty} a_{\sigma(n)} = \sum_{n=0}^{\infty} \sum_{i+j=n} a_{i,j}.$$

证明. 事实上, 对级数 $\sum_{n=0}^{\infty} \sum_{i+j=n} a_{i,j}$, 存在一个特殊的一一对应 $\tau: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, 使得级数 $\sum_{n=0}^{\infty} \sum_{i+j=n} a_{i,j}$ 就是级数 $\sum_{n=0}^{\infty} a_{\tau(n)}$ 。

于是对于任何一一对应 $\sigma: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, 级数 $\sum_{n=0}^{\infty} a_{\sigma(n)}$ 是 $\sum_{n=0}^{\infty} a_{\tau(n)}$ 的一个重排: $\sigma \circ \tau^{-1}: \mathbb{N} \rightarrow \mathbb{N}$ 。所以根据级数重排定理, 级数 $\sum_{n=0}^{\infty} a_{\sigma(n)}$ 绝对收敛, 且 $\sum_{n=0}^{\infty} a_{\sigma(n)} = \sum_{n=0}^{\infty} a_{\tau(n)}$ 。 \square

考虑 $\sum_{n=0}^{\infty} \sum_{\max\{i,j\}=n} a_i b_j$, 可得上一个定理的结论, 这是因为:

$$\sum_{n=0}^N \sum_{\max\{i,j\}=n} a_i b_j = \sum_{i,j \leq N} a_i b_j = \sum_{i=0}^N a_i \cdot \sum_{j=0}^N b_j.$$

习题8.4

- 对于实数级数 $\sum_{n=1}^{\infty} a_n$, 证明: 若级数 $\sum_{n=1}^{\infty} \left(\sum_{j=k_n}^{k_{n+1}-1} a_j \right)$ ($k_1 = 1 < k_2 < k_3 < \dots < k_n < \dots$) 收敛, 且同一括号中所有 a_j 同号, 则级数 $\sum_{n=1}^{\infty} a_n$ 收敛, 且和不变。

8.5 无穷乘积

形如 $\prod_{n=1}^{\infty} a_n$ 的对象称为无穷乘积。称无穷乘积 $\prod_{n=1}^{\infty} a_n$ 收敛, 如果 $\left\{ \prod_{n=1}^N a_n \right\}_{N \geq 1}$ 收敛于一个非零数。

记 $P_N = \prod_{n=1}^N a_n$ 。则 P_N 收敛的必要条件是 $a_{N+1} = \frac{P_{N+1}}{P_N} \rightarrow 1$ ($N \rightarrow +\infty$)。

因此不妨设所有 $a_n > 0$ 。此时 $\prod_{n=1}^{\infty} a_n$ 收敛当且仅当 $\sum_{n=1}^{\infty} \ln a_n$ 收敛, 且 $\prod_{n=1}^{\infty} a_n = e^{\sum_{n=1}^{\infty} \ln a_n}$ 。

于是, 我们可以把无穷乘积的研究就转变为对相应的级数的研究。

例 8.5.1. 1734年, 时年27岁的 Euler 发现了以下等式

$$\frac{\sin x}{x} = \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2\pi^2}\right).$$

把左边 Taylor 展开, 并两边比较 x^2 的系数, Euler 得到 $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, 从而解决了意大利数学家 Pietro Mengoli 于1644年提出的这个充满挑战性的问题(被称为 Basel 问题), 它曾令 Bernoulli 兄弟、Leibniz 等一众数学家束手无策, 以至于号称无穷级数大师的 Jacob Bernoulli 在他的论文中不得不这样写道“如果谁能解决并告知这个我们无能为力的问题, 我们将不胜感激”。

我们可以用上面所说的办法证明上式右端的无穷乘积是收敛的。

$$\ln \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2\pi^2}\right) = \sum_{n=1}^{\infty} \ln \left(1 - \frac{x^2}{n^2\pi^2}\right) = \sum_{n=1}^{\infty} O\left(\frac{1}{n^2}\right)$$

绝对收敛。在下一章我们会给出等式 $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ 的一个严格证明。 \square

8.6 习题讨论课：数项级数

1. 求和。裂项错位相减

(1) $\sum_{k=1}^{\infty} (b_{k+1} - b_k)$ 收敛当且仅当 $\lim_{k \rightarrow +\infty} b_k$ 收敛。

(2) $\sum_{n=1}^{\infty} \frac{1}{(pn+q)[p(n+1)+q]}$

(3) $\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)}$

(4) $\sum_{n=1}^{\infty} (\sqrt{n+2} - 2\sqrt{n+1} + \sqrt{n})$

(5) 设 $\lim_{n \rightarrow +\infty} nu_n$ 收敛。证明级数 $\sum_{n=1}^{\infty} (n+1)(u_{n+1} - u_n)$ 收敛当且仅当

$\sum_{n=1}^{\infty} u_n$ 收敛。

(6) $\sum_{n=1}^N \sin nx$ (这虽然不是级数, 但在使用 Dirichlet 判别法时需要计算部分和并验证部分和有界)

2. 讨论 $\sum_{n=1}^{\infty} nz^n$ 的收敛性并求和。

3. (Cauchy凝聚定理) 设数列 $\{a_n\}$ 单调不增, 证明 $\sum_{n=1}^{\infty} a_n$ 收敛当且仅当

$\sum_{n=1}^{\infty} 2^n a_{2^n}$ 收敛。用这个办法来检验 $\sum_{n=1}^{\infty} \frac{1}{n^p}$, $\sum_{n=1}^{\infty} \frac{1}{n(\ln n)^p}$ 的收敛性。1689年

Jacob Bernoulli 用这个办法证明了调和级数 $\sum_{n=1}^{\infty} \frac{1}{n}$ 是发散的。

4. (1) 对任意复数 z , 证明

$$C(z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}, \quad S(z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!}$$

都是绝对收敛的。

(2) 证明任意复数 z_1, z_2 ,

$$C(z_1 - z_2) = C(z_1)C(z_2) + S(z_1)S(z_2).$$

从而对任意复数 z ,

$$C(z)^2 + S(z)^2 = 1.$$

(3) 证明对任意复数 z ,

$$C(z) + iS(z) = \exp(iz).$$

5. 讨论级数收敛性(绝对收敛、条件收敛、发散)

(1) $\sum_{n=1}^{\infty} \sin \frac{\pi}{2^n}$

(2)
$$\sum_{n=1}^{\infty} \frac{\ln n}{n^{\frac{3}{2}}}$$

(3)
$$\sum_{n=1}^{\infty} \frac{1}{n} \sin \frac{1}{\sqrt{n}}$$

(4)
$$\sum_{n=1}^{\infty} \left(e^{\frac{1}{\sqrt{n}}} - 1 \right)$$

(5)
$$\sum_{n=1}^{\infty} \frac{n^2}{2^n}$$

(6)
$$\sum_{n=1}^{\infty} \frac{(2n-1)!!}{3^n n!}$$

(7)
$$\sum_{n=1}^{\infty} 2^{n-1} \tan \frac{\pi}{2^n}$$

(8)
$$\sum_{n=1}^{\infty} \frac{1}{3^n} \left(1 + \frac{1}{n} \right)^{n^2}$$

(9)
$$\sum_{n=1}^{\infty} \frac{1}{\sqrt[3]{n+1}} \ln \frac{n+2}{n}$$

(10)
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n - \ln n}$$

(11)
$$\sum_{n=2}^{\infty} (-1)^n \sqrt{\frac{n+1}{n(n-1)}}$$

(12)
$$\sum_{n=1}^{\infty} \sin(\pi \sqrt{n^2 + 1})$$

(13)
$$\sum_{n=1}^{\infty} \frac{\sin n}{n^p + \sin n}$$

6. 我们知道级数 $S = \sum_{n=1}^{\infty} \frac{1}{n^2}$ 收敛, 因此部分和 S_n 是 S 的近似值. 为了得到 S 的更好的近似值, 我们可以考虑对 S_n 进行修正. 比如假设

$$S = S_n + \frac{A}{n} + O\left(\frac{1}{n^2}\right).$$

证明 $A = 1$. 请试做进一步的修正, 并通过数值计算进行观察.

第9章 函数项级数与幂级数

9.1 一致收敛与极限函数的性质

定义 9.1.1. 设 $I \subseteq \mathbb{R}$, 对函数 $f: I \rightarrow \mathbb{R}$, 记

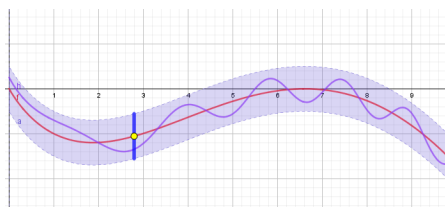
$$\|f\|_{\infty} = \sup_{x \in I} |f(x)|.$$

对函数序列 $f_n: I \rightarrow \mathbb{R}$, 我们称 f_n 在 I 上一致收敛于 f (记为 $f_n \xrightarrow{I} f$), 如果

$$\lim_{n \rightarrow +\infty} \|f_n - f\|_{\infty} = 0,$$

即: 对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $n > N(\varepsilon)$, 对任意 $x \in I$, $|f_n(x) - f(x)| < \varepsilon$. 这里所说的一致是指 $N(\varepsilon)$ 与 $x \in I$ 无关。

若对任意 $x \in I$ 都有 $\lim_{n \rightarrow +\infty} f_n(x) = f(x)$, 则称 f_n 在 I 上逐点收敛于 f , 即对任意 $x \in I$ 以及任意 $\varepsilon > 0$, 存在 $N(x, \varepsilon)$ 使得对任意正整数 $n > N(x, \varepsilon)$, $|f_n(x) - f(x)| < \varepsilon$.



很明显, 若 f_n 在 I 上一致收敛于 f , 则 f_n 在 I 上逐点收敛于 f . 但下面这个例子表明逐点收敛并不蕴涵一致收敛。

例 9.1.2. 当 $n \rightarrow +\infty$ 时, x^n 在区间 $[0, 1]$ 上逐点收敛于函数

$$f(x) = \begin{cases} 0, & 0 \leq x < 1; \\ 1, & x = 1. \end{cases}$$

x^n 在任何区间 $[0, a]$ ($0 < a < 1$) 上一致收敛, 但在区间 $[0, 1]$ 上不是一致收敛。

证明. 当 $0 \leq x < 1$ 时, $f_n(x) = x^n$ 满足 $\lim_{n \rightarrow +\infty} f_n(x) = 0$ 。对 $x = 1$, $f_n(1) = 1$, 所以 $\lim_{n \rightarrow +\infty} f_n(1) = 1$ 。所以 f_n 在 $[0, 1]$ 上逐点收敛于 f 。

对任意 $0 < a < 1$, 在区间 $[0, a]$ 上,

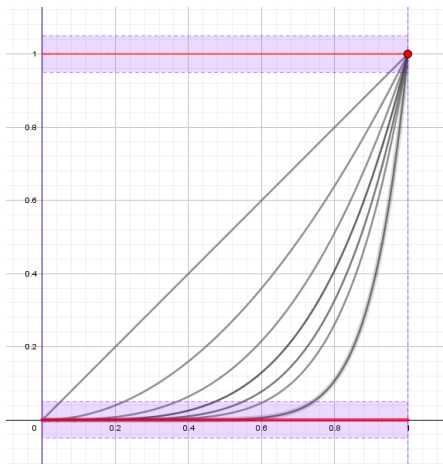
$$\|f_n - f\|_\infty = \sup_{x \in [0, a]} |x^n - 0| = a^n \rightarrow 0, \quad n \rightarrow +\infty.$$

所以 f_n 在 $[0, a]$ 上一致收敛于 f 。

在区间 $[0, 1]$ 上, 对任意正整数 $n \geq 2$, 取 $x_n = \sqrt[n]{1 - \frac{1}{n}}$, 则 $f_n(x_n) = 1 - \frac{1}{n}$, $f(x_n) = 0$, 所以

$$\|f_n - f\|_\infty \geq |f_n(x_n) - f(x_n)| = 1 - \frac{1}{n} \geq \frac{1}{2}.$$

因此 f_n 在 $[0, 1]$ 上不是一致收敛于 f 的。 □



注 9.1.3. 几何上, f_n 在 I 上一致收敛于 f , 意味着当 n 充分大时, f_n 的图像全部落入以 f 的图像为中心上下宽度为 2ε 的带形区域中。很明显, 上面这个例子明显无法做到这一点。

定理 9.1.4. 函数序列 $\{f_n\}$ 在 I 上一致收敛当且仅当 $\{f_n\}$ 在 I 上满足一致 Cauchy 条件: 即对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $m, n > N(\varepsilon)$, $\|f_m - f_n\|_\infty < \varepsilon$, 即对任意 $x \in I$, $|f_n(x) - f_m(x)| < \varepsilon$ 。

证明. 必要性是显然的。下证充分性。设 $\{f_n\}$ 在 I 上满足一致 Cauchy 条件。则对任意 $x \in I$, 数列 $f_n(x)$ 是 Cauchy 数列, 从而存在极限, 记 $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$ 。

由一致 Cauchy 知: 对任意 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $m, n > N(\varepsilon)$, 对任意 $x \in I$, $|f_n(x) - f_m(x)| < \varepsilon$ 。

让 $m \rightarrow +\infty$, 则 $|f_n(x) - f(x)| \leq \varepsilon$ 。这对所有正整数 $n > N(\varepsilon)$ 以及任意 $x \in I$ 都成立, 所以 $\{f_n\}$ 在 I 上一致收敛于 f 。□

记

$$\mathcal{B}(I) = \{f : I \rightarrow \mathbb{R} \mid \|f\|_\infty < +\infty\},$$

即 $\mathcal{B}(I)$ 为在 I 上定义且有界的函数组成的集合。

易见, $\mathcal{B}(I)$ 是一个线性空间, $\|\cdot\|_\infty$ 是它上面的一个范数。当 I 是有界闭集时, 连续函数组成的集合 $\mathcal{C}(I)$ 是 $\mathcal{B}(I)$ 的一个线性子空间。

定理 9.1.5. 设 $f_n \in \mathcal{B}(I)$, f_n 在 I 上一致收敛到 f 。则

1. $f \in \mathcal{B}(I)$, 即 $(\mathcal{B}(I), \|\cdot\|_\infty)$ 是完备的。
2. f_n 在 I 上一致有界: 即存在 $M > 0$ 使得对任意 n 以及任意 $x \in I$ 都有 $|f_n(x)| \leq M$ 。

证明. (1) 因为 f_n 在 I 上一致收敛到 f , 所以存在 N 使得对任意 $n \geq N$ 以及任意 $x \in I$, $|f_n(x) - f(x)| < 1$ 。

从而 $|f(x)| \leq |f_N(x)| + 1 \leq \|f_N\|_\infty + 1$, 因此 $f \in \mathcal{B}(I)$ 。

(2) 对任意正整数 $n \geq N$ 和任意 $x \in I$, $|f_n(x)| \leq |f(x)| + 1 \leq \|f\|_\infty + 1$ 。所以对任意正整数 $n \geq 1$, $\|f_n\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty + \cdots + \|f_N\|_\infty + \|f\|_\infty + 1$ 。因此 f_n 在 I 上一致有界。□

例 9.1.6. 设

$$f_n(x) = \begin{cases} n, & 0 < x \leq \frac{1}{n}, \\ \frac{1}{x}, & \frac{1}{n} < x \leq 1. \end{cases}$$

很明显 f_n 在区间 $(0, 1]$ 上逐点收敛到 $\frac{1}{x}$, 却不是一致收敛。因为, 如果 f_n 在区间 $(0, 1]$ 上一致收敛, 则极限函数必然是 $\frac{1}{x}$ 。每个 f_n 都有界, 但 $\frac{1}{x}$ 是无界函数, 这与上述定理矛盾。另外, 虽然每个 f_n 都有界, 但它们不是一致有界的, 所以根据上述定理也可以知道这收敛不是一致收敛。□

问题 9.1.7. 在数列情形我们知道有界的实数列或复数列都有收敛的子列。现在的问题是, 如果 f_n 在 I 上一致有界, 那么它是否有子列在 I 上一致收敛?

定理 9.1.8. 设 $f_n \in \mathcal{C}(I)$ 并且 $\{f_n\}$ 在 I 上一致收敛到 f , 则

1. $f \in \mathcal{C}(I)$ 。当 I 是有界闭集时, $\mathcal{C}(I)$ 是 $(\mathcal{B}(I), \|\cdot\|_\infty)$ 的闭子集。

2. 当 I 是有界闭集时, f_n 在 I 上等度一致连续: 即对任意 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$ 使得对任意正整数 n 以及任意 $x, y \in I$, 只要 $|x - y| < \delta(\varepsilon)$, 就都有 $|f_n(x) - f_n(y)| < \varepsilon$.

证明. (1) 因为 $\{f_n\}$ 在 I 上一致收敛到 f , 所以对任意 $\varepsilon > 0$, 存在 N_ε 使得对任意正整数 $n \geq N_\varepsilon$, 对任意 $x \in I$, $|f_n(x) - f(x)| < \varepsilon$.

对每个 $x_0 \in I$, 因为 f_{N_ε} 在 x_0 连续, 所以存在 $\delta(x_0, \varepsilon) > 0$ 使得对任意 $x \in I$, 只要 $|x - x_0| < \delta(x_0, \varepsilon)$, 就有 $|f_{N_\varepsilon}(x) - f_{N_\varepsilon}(x_0)| < \varepsilon$. 因此,

$$|f(x) - f(x_0)| \leq |f(x) - f_{N_\varepsilon}(x)| + |f_{N_\varepsilon}(x) - f_{N_\varepsilon}(x_0)| + |f_{N_\varepsilon}(x_0) - f(x_0)| < 3\varepsilon.$$

所以 f 在 x_0 连续.

(2) 当 I 是有界闭集时, f 在 I 上一致连续, 所以存在 $\delta_1(\varepsilon) > 0$ 使得对任意 $x, y \in I$, 只要 $|x - y| < \delta_1(\varepsilon)$, 就都有 $|f(x) - f(y)| < \varepsilon$.

对任意 $n > N_\varepsilon$,

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f(x)| + |f(x) - f(y)| + |f(y) - f_n(y)| \leq 3\varepsilon.$$

对 $n \leq N_\varepsilon$, 存在 $\delta_2(\varepsilon) > 0$ 使得对任意 $x, y \in I$, 只要 $|x - y| < \delta_1(\varepsilon)$, 就有

$$|f_n(x) - f_n(y)| \leq \varepsilon.$$

因此对任意正整数 n 以及任意 $x, y \in I$, 只要 $|x - y| < \min\{\delta_1(\varepsilon), \delta_2(\varepsilon)\}$, 就都有 $|f_n(x) - f_n(y)| < 3\varepsilon$. 所以 f_n 在 I 上等度一致连续. \square

定理 9.1.9. 设 $f_n \in R[a, b]$ 并且 $\{f_n\}$ 在 $[a, b]$ 上一致收敛于 f , 则 $f \in R[a, b]$ 且 $\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \int_a^b f_n(x)dx$. 因此 $R[a, b]$ 是 $(\mathcal{B}(I), \|\cdot\|_\infty)$ 的闭子集, \int_a^b 是 $R[a, b]$ 上的连续函数.

证明. 因为 $f_n \in R[a, b]$, 所以 f_n 有界. 由前面的定理结论知 f 有界.

记 D_n 为 f_n 在 $[a, b]$ 中的间断点组成的集合, 则 D_n 长度为零. 令 $D = \bigcup_{n \geq 1} D_n$, 则 D 长度为零. 对任意 $x \in [a, b] \setminus D$, 每个 f_n 都在 x 连续. 由上述定理的证明知 f 在 x 处连续. 所以 f 在 $[a, b]$ 中的间断点组成的集合长度为零. 因此由 Riemann 可积的 Lebesgue 准则知 $f \in R[a, b]$.

由

$$\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| \leq \int_a^b |f_n(x) - f(x)|dx \leq (b-a)\|f_n - f\|$$

知 $\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \int_a^b f_n(x)dx$ \square

例 9.1.10. 把区间 $[0, 1]$ 中的所有有理数写成一个数列 r_n , 令

$$f_n(x) = \begin{cases} 1, & x \in \{r_1, \dots, r_n\}, \\ 0, & x \in [0, 1] \setminus \{r_1, \dots, r_n\}. \end{cases}$$

则 $f_n \in R[0, 1]$ 且 $\int_0^1 f_n(x)dx = 0$ 。 f_n 逐点收敛到 Dirichlet 函数, 但后者不是 Riemann 可积函数。 \square

例 9.1.11. 考虑函数序列

$$f_n(x) = \begin{cases} 2^{2n+2}x, & 0 \leq x \leq \frac{1}{2^{n+1}}, \\ 2^{2n+2} \left(\frac{1}{2^n} - x\right), & \frac{1}{2^{n+1}} \leq x \leq \frac{1}{2^n}, \\ 0 & \frac{1}{2^n} \leq x \leq 1. \end{cases}$$

则 $f_n \in R[0, 1]$ 且 $\int_0^1 f_n(x)dx = 1$ 。 f_n 逐点收敛到常值函数 $f = 0$, 后者是 Riemann 可积函数, 但是 $\int_0^1 f(x)dx = 0$ 。 \square

定理 9.1.12. 设 $f_n \in C^1[a, b]$, $\{f_n(a)\}$ 收敛, $\{f'_n\}$ 在 $[a, b]$ 上一致收敛于 g , 则 $\{f_n\}$ 在 $[a, b]$ 上一致收敛, 其极限 f 满足 $f \in C^1[a, b]$ 且 $f' = g$ 。

证明. 由连续性定理知, $g \in C[a, b]$ 。

令 $A = \lim_{n \rightarrow +\infty} f_n(a)$, $f(x) = A + \int_a^x g(t)dt$ 。 则 $f \in C^1[a, b]$, $f(a) = A$, $f' = g$ 。

对每个正整数 n , $f_n(x) = f_n(a) + \int_a^x f'_n(t)dt$ 。

于是对任意 $x \in [a, b]$,

$$\begin{aligned} |f_n(x) - f(x)| &= \left| f_n(a) - f(a) + \int_a^x f'_n(t) - g(t)dt \right| \\ &\leq |f_n(a) - f(a)| + (b-a)\|f'_n - g\|, \end{aligned}$$

所以 f_n 一致收敛于 f 。 \square

习题9.1

1. 设 I 是 \mathbb{R}^m 的非空有界闭子集。 $\mathcal{K} \subset C(I)$ 满足一致有界、等度一致连续。证明 \mathcal{K} 中任何序列 $\{f_n\}_{n \geq 1}$ 都含有一致 Cauchy 的子列。

9.2 函数项级数的一致收敛性与和函数的性质

定义 9.2.1. 称函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 $I \subseteq \mathbb{R}$ 上一致收敛, 如果它的部分和序列 $S_N(x) = \sum_{n=1}^N u_n(x)$ 在 $I \subseteq \mathbb{R}$ 上一致收敛。

于是把上一节的结论应用到函数项级数得到

定理 9.2.2. 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 $I \subseteq \mathbb{R}$ 上一致收敛。则

1. 如果每个 $u_n \in \mathcal{B}(I)$, 则和函数 $\sum_{n=1}^{\infty} u_n \in \mathcal{B}(I)$ 。
2. 如果每个 $u_n \in \mathcal{C}(I)$, 则和函数 $\sum_{n=1}^{\infty} u_n \in \mathcal{C}(I)$, 从而对任意 $x_0 \in I$,

$$\lim_{x \rightarrow x_0} \sum_{n=1}^{\infty} u_n(x) dx = \sum_{n=1}^{\infty} \lim_{x \rightarrow x_0} u_n(x).$$

3. 如果 I 是有界闭区间且每个 $u_n \in R(I)$, 则和函数 $\sum_{n=1}^{\infty} u_n \in R(I)$, 并且

$$\int_I \sum_{n=1}^{\infty} u_n(x) dx = \sum_{n=1}^{\infty} \int_I u_n(x) dx.$$

定理 9.2.3. 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在区间 $I \subseteq \mathbb{R}$ 上满足

1. 每个 $u_n \in C^1(I)$, 且 $\sum_{n=1}^{\infty} u'_n$ 在 I 上一致收敛;
2. 存在 $x_0 \in I$ 使得 $\sum_{n=1}^{\infty} u_n(x_0)$ 收敛,

则和函数 $\sum_{n=1}^{\infty} u_n$ 在 I 上一致收敛, 并且

$$\left(\sum_{n=1}^{\infty} u_n(x) \right)' = \sum_{n=1}^{\infty} u'_n(x).$$

关于函数项级数的一致收敛性, 有以下重要结论。

定理 9.2.4. 函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 $I \subseteq \mathbb{R}$ 上一致收敛当且仅当它在 I 上满足一致 Cauchy 条件: 即对于任何 $\varepsilon > 0$, 存在 $N(\varepsilon)$ 使得对任意正整数 $n \geq N(\varepsilon)$

和正整数 $p \geq 1$, $\left\| \sum_{k=n+1}^{n+p} u_n \right\|_{\infty} < \varepsilon$, 即对任意 $x \in I$, 都有

$$\left| \sum_{k=n+1}^{n+p} u_k(x) \right| < \varepsilon.$$

判别函数项级数一致绝对收敛, 最常用的方法是 Weierstrass 强级数法。

定理 9.2.5 (Weierstrass). 如果存在收敛的数项级数 $\sum_{n=1}^{\infty} a_n$ 使得对任意 $n \geq 1$

$$|u_n(x)| \leq a_n, \quad \forall x \in I,$$

即 $\|u_n\|_{\infty} \leq a_n$, 则函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 I 上一致绝对收敛。

例 9.2.6. 讨论级数 $\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$ ($x \in \mathbb{R}$) 的一致收敛性。

解: 对任何 $\delta > 0$, 当 $|x| \geq \delta$ 时, 对任意正整数 n ,

$$0 \leq \frac{x^2}{(1+x^2)^n} \leq \frac{1}{(1+x^2)^{n-1}} \leq \frac{1}{(1+\delta^2)^{n-1}},$$

级数 $\sum_{n=1}^{\infty} \frac{1}{(1+\delta^2)^{n-1}}$ 收敛, 所以 $\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$ 在 $\{x \in \mathbb{R} | |x| \geq \delta\}$ 上一致收敛。

对任意正整数 $n \geq 2$, 函数 $u_n(x) = \frac{x^2}{(1+x^2)^n}$ 在 $x_n = \frac{1}{\sqrt{n-1}}$ 处取最大值。对 $n < k \leq 2n$,

$$u_k(x_n) = \frac{\frac{1}{n-1}}{\left(1 + \frac{1}{n-1}\right)^k} > \frac{\frac{1}{n}}{\left(1 + \frac{2}{n}\right)^{2n}} > \frac{1}{e^{4n}},$$

因此

$$\|u_{n+1} + \cdots + u_{2n}\|_{\infty} \geq \sum_{k=n+1}^{2n} u_k(x_n) > \frac{1}{e^4},$$

因此级数 $\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$ 在任何以 0 为聚点的区间上都不是一致收敛的。 \square

判别函数项级数一致条件收敛的常用方法是如下的 Hardy-Dirichlet-Abel 判别法。

定理 9.2.7 (Hardy-Dirichlet-Abel). 如果以下三组条件之一成立, 则级数 $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ 在 I 上一致收敛。

1. (Hardy)

(a) $\sum_{n=1}^N u_n$ 在 I 上一致有界;

(b) $\sum_{n=1}^{\infty} |v_{n+1} - v_n|$ 在 I 上一致收敛, 且 $v_n \xrightarrow{I} 0$ 。

2. (Dirichlet)

(a) $\sum_{n=1}^N u_n$ 在 I 上一致有界;

(b) v_n 关于 n 单调, 且单调性对 $x \in I$ 相同, 且 $v_n \xrightarrow{I} 0$ 。

3. (Abel)

(a) $\sum_{n=1}^{\infty} u_n$ 在 I 上一致收敛;

(b) v_n 关于 n 单调且有界, 且对 $x \in I$ 一致。

例 9.2.8. 1830年, Bolzano 构造了一个处处连续但处处不可微的函数。1860年, Weierstrass 构造了如下函数

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x).$$

很明显, 当 $0 < a < 1$ 时, 该级数在 \mathbb{R} 上一致收敛, 从而 f 是连续函数。Weierstrass 证明了, 当 $ab > 1 + \frac{3\pi}{2}$ 时, f 处处不可微。

随后, 数学家们给出了越来越多这样的函数的例子, 并且证明了绝大多数的连续函数都是处处不可微的。关于这个话题的历史详见 Johan Thim 的硕士论文 Continuous Nowhere Differentiable functions。

例 9.2.9. 证明

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2}, \quad 0 < x < 2\pi,$$

且上述级数在区间 $(0, 2\pi)$ 中内闭一致收敛。

证明. 对任意 $0 < \delta < \pi$, 当 $\delta \leq x \leq 2\pi - \delta$ 时,

$$\begin{aligned} \sum_{n=1}^N \frac{\sin nx}{n} &= \sum_{k=1}^N \int_{\pi}^x \cos kt dt = \int_{\pi}^x \sum_{n=1}^N \cos nt dt \\ &= \int_{\pi}^x \frac{\sin(N + \frac{1}{2})t - \sin \frac{t}{2}}{2 \sin \frac{t}{2}} dt = \int_{\pi}^x \frac{\sin(N + \frac{1}{2})t}{2 \sin \frac{t}{2}} dt - \frac{x - \pi}{2}, \end{aligned}$$

而

$$\begin{aligned} \left| \int_{\pi}^x \frac{\sin(N + \frac{1}{2})t}{2 \sin \frac{t}{2}} dt \right| &= \left| -\frac{\cos(N + \frac{1}{2})t}{(2N + 1) \sin \frac{t}{2}} \Big|_{\pi}^x - \int_{\pi}^x \frac{\cos(N + \frac{1}{2})t}{(2N + 1)} \frac{\cos \frac{t}{2}}{2 \sin^2 \frac{t}{2}} dt \right| \\ &\leq \frac{1}{(2N + 1) \sin \frac{x}{2}} + \frac{1}{(2N + 1) \sin \frac{\pi}{2}} + \frac{1}{2N + 1} M |x - \pi| \\ &\leq \frac{1}{2N + 1} \left[\frac{1}{\sin \frac{\delta}{2}} + 1 + M |\pi - \delta| \right], \end{aligned}$$

其中 $M = \max_{\delta \leq t \leq 2\pi - \delta} \left| \frac{1}{2 \sin^2 \frac{t}{2}} \right| = \frac{1}{2 \sin^2 \frac{\delta}{2}}$ 。因此 $\lim_{N \rightarrow +\infty} \int_{\pi}^x \frac{\sin(N + \frac{1}{2})t}{2 \sin \frac{t}{2}} dt = 0$, 且这极限对 $x \in [\delta, 2\pi - \delta]$ 一致收敛。所以级数 $\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ 在区间 $[\delta, 2\pi - \delta]$ 上一致收敛于 $\frac{x - \pi}{2}$ 。

所以对任意 $0 < x < 2\pi$, 都有

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2}.$$

然而在 $x = 0$ 附近和 $x = 2\pi$ 附近, 这级数都不是一致收敛的。想一想这是为什么。 \square

例 9.2.10. 证明级数

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$$

对任意实数 x 收敛, 并求它的和函数在区间 $[0, 2\pi]$ 上的表达式。

证明. 由 Weistrass 判别法知级数 $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$ 对任何实数 x 一致收敛, 则其和函数 $F(x)$ 连续. 由上例知, 在区间 $(0, 2\pi)$ 上 $F(x)$ 可导, 并且

$$F'(x) = -\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{x - \pi}{2},$$

从而

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2} = F(x) = F(0) + \int_0^x \frac{t - \pi}{2} dt = F(0) + \frac{x^2 - 2\pi x}{4}.$$

把它在区间 $[0, 2\pi]$ 上积分, 得到

$$0 = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^{2\pi} \cos nx dx = 2\pi F(0) - \frac{\pi^3}{3},$$

所以

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = F(0) = \frac{\pi^2}{6}.$$

从而

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2} = \frac{\pi^2}{6} + \frac{x^2 - 2\pi x}{4}, \quad 0 \leq x \leq 2\pi.$$

□

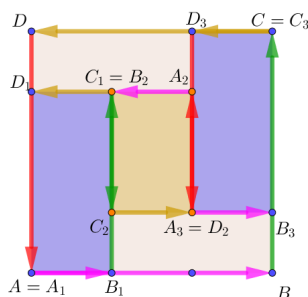
习题9.2

1. 记 $A = \{f \in C[0, 1] \mid f(0) = 0, f(1) = 1, f[0, 1] \subseteq [0, 1]\}$. 对 $f \in A$, 定义

$$Tf(x) = \begin{cases} \frac{3}{4}f(3x), & 0 \leq x \leq \frac{1}{3}; \\ \frac{1}{4} - \frac{1}{2}f(2-3x), & \frac{1}{3} < x < \frac{2}{3}; \\ \frac{1}{4} + \frac{3}{4}f(3x-2), & \frac{2}{3} \leq x \leq 1. \end{cases}$$

证明:

- $A \subset C[0, 1]$ 是闭集,
- 对任意 $f \in A$, $Tf \in A$,
- $T: A \rightarrow A$ 关于 $\|f\| = \sup_{x \in [0, 1]} |f(x)|$ 是压缩映射,
- T 的不动点 $f^* \in A$ 处处不可微。



9.3 幂级数

幂级数的样子就像是一个无穷多次的多项式，在 Newton 和 Leibniz 创建微积分的年代，它是当时的数学家们最得心应手的计算工具，人们利用它可以很方便地计算导数、积分以及求解微分方程。

定义 9.3.1. 形如 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 的级数称为幂级数。

对 C^∞ 函数 f ，称

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

为 f 在 $x=a$ 处的 **Taylor 级数**，当 $a=0$ 时，Taylor 级数

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

又称为 f 的 **Maclaurin 级数**。

引理 9.3.2. (1) 若幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $x_0 \neq 0$ 收敛，则幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| < |x_0|\}$ 中内闭一致绝对收敛，即对任何 $0 < r < |x_0|$ ，幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| \leq r\}$ 上一致绝对收敛。从而在 $\{x \in \mathbb{C} \mid |x| < |x_0|\}$ 内处处收敛。

(2) 若幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $x_0 \neq 0$ 收敛，则 $\sum_{n=0}^{\infty} a_n (tx_0)^n$ 关于 t 在区间 $[0, 1]$ 上一致收敛。

(3) 如果幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 x_0 处发散，则它在 $\{x \in \mathbb{C} \mid |x| > |x_0|\}$ 中处处发散。

证明. (1) 因为 $\sum_{n=0}^{\infty} a_n x_0^n$ 收敛, 所以 $\lim_{n \rightarrow +\infty} |a_n x_0^n| = 0$, 因此数列 $|a_n x_0^n|$ 有界。

$$|a_n x^n| \leq |a_n| |x|^n \leq |a_n| r^n = |a_n x_0^n| \left(\frac{r}{|x_0|} \right)^n \leq M \left(\frac{r}{|x_0|} \right)^n,$$

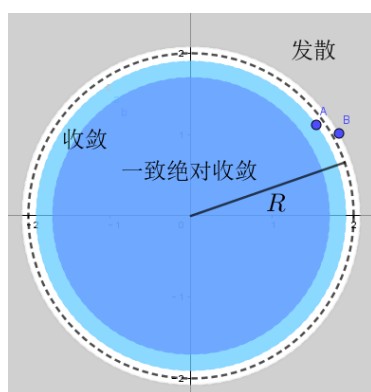
因为 $\sum_{n=0}^{\infty} \left(\frac{r}{|x_0|} \right)^n$ 收敛, 所以由 Weierstrass 判别法知 $\sum_{n=0}^{\infty} a_n x^n$ 在

$$\{x \in \mathbb{C} \mid |x| \leq r\}$$

上一致绝对收敛。

(2) 因为 $\sum_{n=0}^{\infty} a_n x_0^n$ 收敛, t^n 关于 t 在区间 $[0, 1]$ 上单调不增、一致有界, 所以由 Abel 判别法 $\sum_{n=0}^{\infty} a_n (tx_0)^n = \sum_{n=0}^{\infty} a_n x_0^n t^n$ 关于 t 在区间 $[0, 1]$ 上一致收敛。

(3) 这是(1)的逆否命题。 \square



定义 9.3.3. 对幂级数 $\sum_{n=0}^{\infty} a_n x^n$, 记

$$R = \sup \left\{ |x| \mid \sum_{n=0}^{\infty} a_n x^n \text{ 收敛} \right\},$$

称 R 为幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径。

定理 9.3.4. 设 R 为幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径, 则

$$1. R = \frac{1}{\limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n|}}.$$

2. 幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| > R\}$ 处处发散, 在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内闭一致绝对收敛, 其和函数在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内处处连续。
3. 如果幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 x_0 处收敛, 则级数 $\sum_{n=0}^{\infty} a_n (tx_0)^n$ 关于 $t \in [0, 1]$ 关于 $t \in [0, 1]$ 连续。
4. 如果幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 x_0 处收敛, 则

$$\int_0^{x_0} \sum_{n=0}^{\infty} a_n x^n dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x_0^{n+1}.$$

5. 幂级数 $\sum_{n=1}^{\infty} n a_n x^{n-1}$ 的收敛半径也是 R 。如果 $R > 0$, 则幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的和函数 $S(x)$ 在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内是 C^∞ 的, 并且可以逐项求任意阶导数

$$\left(\sum_{n=0}^{\infty} a_n x^n \right)^{(k)} = \sum_{n=k}^{\infty} n \cdots (n-k+1) a_n x^{n-k},$$

从而

$$a_n = \frac{S^{(n)}(0)}{n!}, \quad n = 0, 1, 2, 3, \dots,$$

即任何(具有正的收敛半径的)幂函数, 它是其和函数的 *Maclaurin* 级数。

证明. (1) 任取非负实数 r, s 满足

$$r < \frac{1}{\limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n|}} < s.$$

则 $\limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n r^n|} < 1$, $\limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n s^n|} > 1$ 。根据 Cauchy 根式判别法, 级数 $\sum_{n=0}^{\infty} a_n r^n$ 收敛, $\sum_{n=0}^{\infty} a_n s^n$ 发散, 所以 $r \leq R \leq s$ 。因此

$$R = \frac{1}{\limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n|}}.$$

(2) 由收敛半径的定义知, 幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| > R\}$ 处处发散。

任取有界闭集 K 使得 $K \subset \{x \in \mathbb{C} \mid |x| < R\}$ 。取 $r = \max_{x \in K} |x|$ 。则 $r < R$, 于是存在 x_1 使得 $r < |x_1|$, 级数 $\sum_{n=0}^{\infty} a_n x_1^n$ 收敛。所以根据此前的引理, 幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| < r\}$ 内一致绝对收敛, 从而在 K 上一致绝对收敛。

(3) 因为

$$\sum_{n=0}^{\infty} a_n (tx_0)^n = \sum_{n=0}^{\infty} a_n x_0^n t^n$$

关于 $t \in [0, 1]$ 一致收敛, 所以和函数关于 $t \in [0, 1]$ 连续.

(4)

$$\begin{aligned} \int_0^{x_0} \sum_{n=0}^{\infty} a_n x^n dx &= \int_0^1 \sum_{n=0}^{\infty} a_n x_0^{n+1} t^n dt = x_0 \int_0^1 \sum_{n=0}^{\infty} a_n x_0^n t^n dt \\ &= x_0 \sum_{n=0}^{\infty} \int_0^1 a_n x_0^n t^n dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x_0^{n+1} \end{aligned}$$

(5) 因为

$$\limsup_{n \rightarrow +\infty} \sqrt[n]{|na_n|} = \limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n|} \lim_{n \rightarrow +\infty} \sqrt[n]{n} = \limsup_{n \rightarrow +\infty} \sqrt[n]{|a_n|},$$

所以幂级数 $\sum_{n=1}^{\infty} na_n x^{n-1}$ 的收敛半径也是 R .

由(2)知幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 和 $\sum_{n=1}^{\infty} na_n x^{n-1}$ 在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内闭一致绝对收敛, 所以根据函数项级数可微性定理, 幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内可微, 且

$$\left(\sum_{n=0}^{\infty} a_n x^n \right)' = \sum_{n=1}^{\infty} na_n x^{n-1}.$$

对 $\sum_{n=1}^{\infty} na_n x^{n-1}$ 再用这个结论, 可以知道原幂级数二阶可微. 由此可知原幂级数在 $\{x \in \mathbb{C} \mid |x| < R\}$ 内是 C^∞ 的, 并且可以任意次逐项求导

$$\left(\sum_{n=0}^{\infty} a_n x^n \right)^{(k)} = \sum_{n=k}^{\infty} n \cdots (n-k+1) a_n x^{n-k}.$$

在上式中取 $x=0$ 得到 $S^{(k)}(0) = k!a_k$. 从而 $a_k = \frac{S^{(k)}(0)}{k!}$. □

例 9.3.5. 证明对任何实数 x ,

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}, \\ \sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \quad (\text{Newton, 1669}). \end{aligned}$$

证明. 幂级数 $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ 的收敛半径为 $R = +\infty$. 所以其和函数 $\exp(x)$ 对任何实数 x 有定义, 且是 C^∞ 函数, 并且

$$\exp'(x) = \sum_{n=1}^{\infty} \left(\frac{x^n}{n!} \right)' = \sum_{n=1}^{\infty} \frac{x^{(n-1)}}{(n-1)!} = \exp(x), \quad \exp(0) = 1.$$

即 $\exp(x)$ 是线性常微分方程初值问题 $\begin{cases} y' = y, \\ y(0) = 1 \end{cases}$ 的解, 而 e^x 也是这个初值问题的解。所以由这个初值问题解的唯一性(可用常数变易法证明), 我们得到 $\exp(x) = e^x$ 。

用类似办法可得其他两个等式(Euler, 1775)。 \square

利用幂级数, 我们可以把上面三个函数的定义域扩充到复数 $z \in \mathbb{C}$, 不难发现

$$\exp(iz) = \cos z + i \sin z, \quad \forall z \in \mathbb{C}.$$

例 9.3.6 (John Bernoulli, 1697). 求积分 $\int_0^1 x^x dx$ 的值。

解: 因为对任意 $x \in (0, 1]$, $x \ln x \in [-1, 0]$, 所以级数

$$x^x = e^{x \ln x} = 1 + x \ln x + \frac{(x \ln x)^2}{2!} + \cdots + \frac{(x \ln x)^n}{n!} + \cdots,$$

在 $(0, 1]$ 上一致收敛。在以下积分中取 $s = -\frac{\ln x}{n+1}$, 则

$$\int_0^1 (x \ln x)^n dx = \int_0^{+\infty} e^{-(n+1)t} (-t)^n dt = (-1)^n \int_0^{+\infty} e^{-s} s^n ds = \frac{(-1)^n n!}{(n+1)^{n+1}},$$

所以

$$\int_0^1 x^x dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^{n+1}}.$$

由于交错级数

$$\left| \sum_{n=9}^{\infty} \frac{(-1)^n}{(n+1)^{n+1}} \right| < \frac{1}{10^{10}},$$

所以

$$\int_0^1 x^x dx \approx \sum_{n=0}^9 \frac{(-1)^n}{(n+1)^{n+1}} = 0.7834305107.$$

\square

例 9.3.7 (Mercator 1668). 证明对任何实数 $-1 < x \leq 1$,

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n}.$$

由此知

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots.$$

证明. 先计算幂级数 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n}$ 的收敛半径。因为

$$\lim_{n \rightarrow +\infty} \left| \frac{(-1)^{n-1} x^n}{n} \right|^{\frac{1}{n}} = |x|,$$

所以由 Cauchy 根式判别法知, $R = 1$ 。

原级数在 $x = -1$ 发散, 在 $x = 1$ 处是收敛的交错级数(Leibniz 判别法)。所以级数收敛域为 $(-1, 1]$ 。从而该幂级数的和函数 $S(x)$ 在区间 $(-1, 1]$ 上连续, 在区间 $(-1, 1)$ 内是 C^∞ 的, 并且可以逐项求导, 于是

$$S'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1} = \sum_{k=0}^{\infty} (-x)^k = \frac{1}{1+x}, \quad -1 < x < 1.$$

因此对任意 $-1 < x \leq 1$,

$$S(x) = S(0) + \int_0^x \frac{1}{1+t} dt = \ln(1+x).$$

代入 $x = 1$ 得到

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots.$$

如果是求右端级数的和, 那么这个等式给出了完美的答案。但是如果用这个级数计算 $\ln 2$ 的近似值, 那么这个级数收敛得太慢了——要算到第 72 项才能准确得到一位精确小数, 到第 159 项才得到两位精确小数, 要得到三位精确小数需要算到第 3397 项!

如果使用

$$\ln 2 = -\ln\left(1 - \frac{1}{2}\right) = \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{1}{2}\right)^n,$$

或者使用

$$\ln 2 = \ln\left(1 + \frac{1}{3}\right) - \ln\left(1 - \frac{1}{3}\right) = \sum_{n=1}^{\infty} \frac{6}{2n-1} \left(\frac{1}{9}\right)^n,$$

都可以更快地得到 $\ln 2$ 的近似值。Euler 在《无穷分析引论》一书的第 123 小节中用后面这个办法计算了 1 到 10 的自然对数, 精确到 25 位小数。建议读者通过数值计算进行观察。□

如果用幂级数的办法对复数 z 定义函数 $\ln(1+z)$, 其收敛范围是什么?

例 9.3.8 (Gregory, 1671). 求反正切函数 $\arctan x$ 的 Maclaurin 级数, 并讨论其收敛范围。

解:

$$(\arctan x)' = \frac{1}{1+x^2},$$

当 $|x| < 1$ 时,

$$\frac{1}{1+x^2} = 1 + (-x^2) + (-x^2)^2 + \cdots + (-x^2)^n + \cdots +,$$

所以对 $-1 < x < 1$,

$$\arctan x = \arctan 0 + \int_0^x \sum_{n=0}^{\infty} (-x^2)^n dx = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}.$$

这个幂级数的收敛半径为 1, 且该幂级数在 $x = \pm 1$ 都收敛, 所以其和函数在 ± 1 连续, 因此对任意 $-1 \leq x \leq 1$,

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}.$$

取 $x = 1$ 得到

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots.$$

1674年, Leibniz用另外的办法也得到了这个结论, 但作为计算 π 的近似值的一个算法, 这远远不是个好方法——它收敛得太慢了. Euler 当年用

$$\frac{\pi}{4} = 5 \arctan \frac{1}{7} + 2 \arctan \frac{3}{79} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \left[5 \left(\frac{1}{7} \right)^{2n+1} + 2 \left(\frac{3}{79} \right)^{2n+1} \right]$$

得到了 π 的20位小数近似值, 他说“全部计算花费的时间仅约为1小时”. 此前, 荷兰人 Ludolph van Ceulun 花了26年的时间计算圆外切和内接正 2^{30} 边形的周长, 得到了 π 的20位小数近似值, 随后他又用了14年的时间在他生命的最后一年(1610年)得到了正 2^{62} 边形的周长得到了 π 的35位小数近似值. 在敬佩 Ludolph 执着精神与惊人毅力的同时, 我们不禁感叹正确的数学对生命的重要性. \square

注 9.3.9.

$$\arctan \frac{x}{y} - \arctan \frac{u}{v} = \arctan \left(\frac{\frac{x}{y} - \frac{u}{v}}{1 + \frac{xu}{yv}} \right) = \arctan \frac{xv - yu}{xu + yv}$$

例 9.3.10. 求 $(1+x)^\alpha$ 的 Maclaurin 级数, 并讨论其收敛范围.

解: 记 $y = (1+x)^\alpha$, 则 $y' = \alpha(1+x)^{\alpha-1}$, 从而

$$\begin{cases} (1+x)y' = \alpha y, \\ y(0) = 1. \end{cases}$$

设 $y = \sum_{n=0}^{\infty} a_n x^n$, 则 $y(0) = a_0 = 1$,

$$\begin{aligned} (1+x)y' &= (1+x) \sum_{n=1}^{\infty} n a_n x^{n-1} \\ &= \sum_{n=0}^{\infty} [n a_n + (n+1) a_{n+1}] x^n \\ &= \sum_{n=0}^{\infty} \alpha a_n x^n, \end{aligned}$$

所以

$$a_0 = 1, \quad na_n + (n+1)a_{n+1} = \alpha a_n, \quad \forall n \geq 0.$$

从而

$$a_{n+1} = \frac{(\alpha - n)a_n}{n+1}.$$

于是

$$a_n = \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-(n-1))}{n!}, \quad \forall n \geq 0.$$

这样得到幂级数

$$\sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-(n-1))}{n!} x^n = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n.$$

Newton 就知道这个幂级数展开, 它推广了正整数次数的二项式公式, 我们称作广义 Newton 二项式。只不过 Newton 从来不关心(也没意识到还有)收敛的问题。

利用 D'Alembert 分式判别法,

$$\frac{\left| \frac{(\alpha-n)(\alpha-n+1)\cdots\alpha}{(n+1)!} x^{n+1} \right|}{\left| \frac{(\alpha-n+1)\cdots\alpha}{n!} x^n \right|} = \frac{|\alpha-n||x|}{n+1} \rightarrow |x|, \quad n \rightarrow +\infty,$$

所以该幂级数收敛半径为 1, 因此该幂级数的和函数 $y(x)$ 与 $(1+x)^\alpha$ 都是以下线性微分方程初值问题的解,

$$\begin{cases} (1+x)y' = \alpha y, & -1 < x < 1 \\ y(0) = 1 \end{cases}$$

所以由该初值问题解的唯一性得到 $y(x) = (1+x)^\alpha$, 即上述幂级数是 $(1+x)^\alpha$ 的 Maclaurin 级数。

$$a_n = \frac{(\alpha - (n-1)) \cdots \alpha}{n!} = \frac{(-1)^n (n-1-\alpha) \cdots (-\alpha)}{n!}$$

因此级数 $\sum_{n=0}^{\infty} a_n$ 最终成为交错级数。

当 $\alpha \leq -1$ 时, $|a_n| \geq 1$, 所以上述幂级数在 $|x| = 1$ 时发散, 其收敛范围是 $(-1, 1)$ 。

当 $\alpha > -1$ 时,

$$n \left(\frac{|a_n|}{|a_{n+1}|} - 1 \right) = n \left(\frac{n+1}{n-\alpha} - 1 \right) = (\alpha+1) \frac{n}{n-\alpha} \rightarrow \alpha+1, \quad n \rightarrow +\infty.$$

由 Raab 判别法知:

1. 当 $\alpha+1 > 1$ 即 $\alpha > 0$ 时, 级数 $\sum_{n=0}^{\infty} |a_n|$ 收敛, 于是原幂级数在 $x = \pm 1$ 收敛, 其收敛范围是 $[-1, 1]$;

2. 当 $\alpha + 1 < 1$ 即 $\alpha < 0$ 时, 级数 $\sum_{n=0}^{\infty} |a_n|$ 发散, 于是原幂级数在 $x = -1$ 发散。

当 $-1 < \alpha < 0$ 时, 取 $0 < s < r < \alpha + 1$ 。于是存在 N 使得对任意 $n \geq N$,

$$r < n \left(\frac{|a_n|}{|a_{n+1}|} - 1 \right).$$

于是数列 $|a_n|$ 最终单调减, 且存在 $M > N$ 使得对任意 $n > M$,

$$\left(1 + \frac{1}{n} \right)^s < 1 + \frac{r}{n} < \frac{|a_n|}{|a_{n+1}|}.$$

于是存在 $C > 0$ 使得 $|a_n| < \frac{C}{n^s}$, 所以 $\lim_{n \rightarrow +\infty} |a_n| = 0$ 。由交错级数收敛判别法, 该幂级数在 $x = 1$ 收敛。

总结如下:

1. 当 $\alpha \leq -1$ 时, 二项级数收敛范围是 $(-1, 1)$;
2. 当 $-1 < \alpha < 0$ 时, 二项级数收敛范围是 $(-1, 1]$;
3. 当 $\alpha > 0$ 时, 级数收敛范围是 $[-1, 1]$ 。

□

例 9.3.11. 求 $\arcsin x$ 的 Maclaurin 级数, 以及它的收敛范围。

解:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-\frac{1}{2}}.$$

Newton 于是利用广义二项式展开得到 $\arcsin x$ 导数的幂级数, 然后再逐项积分得到了 $\arcsin x$ 导数的幂级数。请读者自己追随 Newton 的脚步完成后续计算。

□

例 9.3.12. 求椭圆

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a > b > 0$$

的周长。

解: 取 $e = \sqrt{1 - \frac{b^2}{a^2}}$, 则

$$\begin{aligned} L &= 4 \int_0^{\pi/2} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = 4a \int_0^{\pi/2} \sqrt{1 - e^2 \cos^2 t} dt \\ &= 4a \int_0^{\pi/2} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-e^2 \cos^2 t)^n dt \quad u = \cos^2 t, du = -2 \cos t \sin t dt \\ &= 4a \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n e^{2n} \int_0^1 u^n \frac{du}{2u^{\frac{1}{2}} \sqrt{1-u}} \\ &= 2a \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n e^{2n} B\left(n + \frac{1}{2}, \frac{1}{2}\right) \\ &= 2a\pi - \sum_{n=1}^{\infty} \frac{(2n)!^2}{(2^n n!)^4 (2n-1)} e^{2n}. \end{aligned}$$

□

例 9.3.13. 给定一个数列 $\{a_n\}_{n=0}^{\infty}$, 我们称

$$g(z) = \sum_{n=0}^{\infty} a_n z^n$$

为数列 $\{a_n\}_{n=0}^{\infty}$ 的母函数(generating function)。

对于一个取值为非负整数的随机变量 X , 其概率分布列 $\{P(X = n)\}_{n \geq 0}$ 对应的母函数

$$g_X(z) = \sum_{n=0}^{\infty} P(X = n) z^n = E(z^X)$$

称为 X 的概率母函数。这个幂级数对 $|z| \leq 1$ 是一致收敛的。易见概率母函数与概率分布相互唯一确定。

例如对服从 Poisson 分布 $\text{Poisson}(\lambda)$ 的随机变量 X ,

$$P(X = n) = \frac{e^{-\lambda} \lambda^n}{n!}, n = 0, 1, 2, \dots$$

其概率母函数为

$$g_X(z) = \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} z^n = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda z)^n}{n!} = e^{\lambda(z-1)}.$$

$$g'_X(z) = \sum_{n=1}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} n z^{n-1}$$

也在 $|z| \leq 1$ 上一致收敛, 所以 X 存在数学期望, 且

$$EX = \sum_{n=1}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} n = g'_X(1) = \lambda e^{\lambda(z-1)} \Big|_{z=1} = \lambda.$$

如果 X, Y 相互独立, 分布服从 Poisson 分布 $\text{Poisson}(\lambda)$ 和 $\text{Poisson}(\mu)$, 则

$$g_{X+Y}(z) = E(z^{X+Y}) = E(z^X)E(z^Y) = e^{\lambda(z-1)}e^{\mu(z-1)} = e^{(\lambda+\mu)(z-1)},$$

所以 $X + Y$ 服从 Poisson 分布 $\text{Poisson}(\lambda + \mu)$ 。 \square

例 9.3.14. 函数

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

的 Maclaurin 级数恒为零。所以并非所有的 C^∞ 函数的 Taylor 级数都收敛到函数自身。

可以证明, 每个幂级数都是一个函数的 Taylor 级数。而这个例子也说明, 每一个幂级数都是无穷多个函数的 Taylor 级数。

C^∞ 函数 $f(x)$ 的 Taylor 级数收敛到其自身当且仅当其 Taylor 余项 $R_n(x)$ 收敛到零。 \square

习题9.3

1. 证明幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径为 $R = \sup \{ |x| \mid \{a_n x^n\}_{n \geq 1} \text{ 有界} \}$ 。
2. 如果把

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

作为 \sin, \cos 的定义,

- (a) 证明 $\sin' = \cos, \cos' = -\sin$ 。
- (b) 对于任意 $0 \leq x \leq 2$, 利用 $1 - \frac{x^2}{n(n+1)} \geq 1 - \frac{x^2}{6} > 0$ ($\forall n \geq 2$) 证明 $\sin x > 0, \cos 2 < 0$ 。
- (c) 证明 \cos 在区间 $[0, 2]$ 中单调减且有唯一零点 x^* , 从而当 $0 \leq x < x^*$ 时, $\cos x > 0$ 。
- (d) 记 $\pi = 2x^*$ 。证明 $2 < \pi < 4$ 。
- (e) 证明 $\sin \frac{\pi}{2} = 1$ 。
- (f) 证明对于任意实数 $x, \cos(x + \frac{\pi}{2}) = -\sin x$ 。
- (g) 证明 $\cos(\pi) = -1, \sin(\pi) = 0$ 。
- (h) 证明对于任意实数 $x, \cos(x + \pi) = -\cos x, \cos(x + 2\pi) = \cos x, \sin(x + \frac{\pi}{2}) = \cos x, \sin(x + \pi) = -\sin x, \sin(x + 2\pi) = \sin x$ 。
- (i) 证明 \sin 在区间 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上严格增, 在区间 $[\frac{\pi}{2}, \frac{3\pi}{2}]$ 上严格减; \cos 在区间 $[-\pi, 0]$ 上严格增, 在区间 $[0, \pi]$ 上严格减。所以 2π 是 \sin, \cos 最小正周期。

(j) 证明 \sin, \cos 的值域都是区间 $[-1, 1]$ 。

3. (a) 证明幂级数 $(z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} + \dots + \frac{(-1)^{n-1}(z-1)^n}{n} + \dots$ 在 $\{z \in \mathbb{C} \mid |z-1| \leq 1, z \neq 0\}$ 上收敛, 记它的和函数为 $\ln z$ 。

(b) 证明或反驳:

$$\ln(z_1 z_2) = \ln z_1 + \ln z_2, \quad \forall z_1, z_2 \in \{z \in \mathbb{C} \mid |z-1| \leq 1, z \neq 0\}.$$

(c) 证明或反驳: $\forall z \in \{z \in \mathbb{C} \mid |z-1| \leq 1, z \neq 0\}, \exp(\ln z) = z$ 。

(d) 求 $\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ 的值, 其中 x 是实数。

(e) 求 $\ln \frac{1+z}{1-z}$ 在 $z=0$ 处的幂级数展开。

(f) 对任意正数 a , 取 $z = \frac{a-1}{a+1}$, 则 $-1 < z < 1$ 且 $\ln a = \ln \frac{1+z}{1-z}$ 。证明

$$\ln 2 = \frac{2}{3} \sum_{n=0}^{\infty} \frac{1}{9^n (2n+1)}.$$

4. (Euler) 计算积分 $\int_0^1 \frac{\sin \ln x}{\ln x} dx$ 。

5. 单摆运动满足机械能守恒 $E = mgl(1 - \cos \theta) + \frac{1}{2}ml^2(\theta')^2$ 。当机械能 $E < mgl$ 时, 单摆是周期运动, 试把周期 $T(E)$ 展成为能量 E 的幂级数。

6. 求 Bessel 方程 $x^2 y'' + xy' + (x^2 - n^2)y = 0$ 的幂级数形式的解。

7. (a) 把 $\int_0^x \frac{t^{k-1}}{1-t^8} dt$ 展成关于 x 的幂级数;

(b) 计算 $\int_0^{1/\sqrt{2}} \frac{4\sqrt{2}(1-x^4) - 8x^3(1+x^2)}{1-x^8} dx$;

(c) (Bailey-Borwein-Plouffe, 1997) 证明

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right).$$

利用这个结果, Bailey, Borwein 和 Plouffe 提出了计算圆周率 π 的 BBP 算法, 该算法可以直接得到 π 的 16 进制展开中的第 n 位数字, 而无需计算之前的各位数字。

8. 设

$$f(z) = \begin{cases} \frac{z}{e^z - 1}, & z \neq 0, \\ 1, & z = 0. \end{cases}$$

(a) 证明当 $|z| < 1$ 时, $f(z)$ 可以写成一个收敛的幂级数

$$f(z) = B_0 + \frac{B_1}{1!}z + \frac{B_2}{2!}z^2 + \dots$$

(b) 上述系数 $B_0, B_1, B_2, \dots, B_n, \dots$ 叫做 **Bernoulli 数**, 求 B_0, B_1, B_2, B_3 的值。

(c) 比较等式

$$\left(B_0 + \frac{B_1}{1!}z + \frac{B_2}{2!}z^2 + \cdots \right) \left(\frac{1}{1!} + \frac{z}{2!} + \frac{z^2}{3!} + \cdots \right) = 1$$

两边 z^{n-1} 的系数, 证明

$$\binom{n}{0}B_0 + \binom{n}{1}B_1 + \cdots + \binom{n}{n-1}B_{n-1} = 0, \quad \forall n \geq 2.$$

(d) 证明当 $n > 1$ 是奇数时, $B_n = 0$ 。

(e) 证明

$$\tan x = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{2^{2n}(2^{2n}-1)B_{2n}}{(2n)!} x^{2n-1}.$$

第 10 章 Fourier 级数初步

10.1 Fourier 级数的背景

例 10.1.1 (简谐振动). 弹簧振子的运动服从胡克定律

$$my'' = -ky,$$

它的通解是

$$y(x) = A \cos \sqrt{\frac{k}{m}}x + B \sin \sqrt{\frac{k}{m}}x$$

是周期函数。它可以看成是匀速圆周运动在一维直线上的投影，这样的振动是最简单的周期运动。

例 10.1.2 (弦振动与波动方程). 一条细弦，质量均匀分布，线密度为 ρ 。它因被拽离平衡位置(视为水平线)而产生振动，记 $u(x, t)$ 为弦上水平坐标为 x 的点在时刻 t 时偏离平衡位置的位移。

作为近似，把弦看成一串离散质点，相邻距离为 h 。第 n 个质点在 t 时刻位于 $(x_n, u(x_n, t))$ 处，它受到与其相邻的质点对它的张力。记 $y_n(t) = u(x_n, t)$ 。张力大小与 $\frac{(y_{n+1}-y_n)+(y_{n-1}-y_n)}{h}$ 成正比，比例系数为 τ 。因此

$$\rho h y_n'' = \tau \frac{(y_{n+1} - y_n) + (y_{n-1} - y_n)}{h},$$

从而

$$\rho \frac{\partial^2 u}{\partial t^2} = \tau \frac{u(x_n + h, t) + u(x_n - h, t) - 2u(x_n, t)}{h^2} \rightarrow \tau \frac{\partial^2 u}{\partial x^2}, \quad h \rightarrow 0,$$

其中上式等号右端分子中把 $u(x, t)$ 对 x 在 x_n 处 Taylor 展开到二阶得到最终极限。于是得到波动方程

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad c = \sqrt{\frac{\tau}{\rho}}.$$

给定初始条件(初始位置和初始速度): $u(x, 0) = f(x)$, $\frac{\partial u}{\partial t}(x, 0) = g(x)$, 我们求解上述微分方程。

利用换元 $\xi = x + ct, \eta = x - ct$, 则

$$du = \frac{\partial u}{\partial \xi} d\xi + \frac{\partial u}{\partial \eta} d\eta = \left(\frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta} \right) dx + c \left(\frac{\partial u}{\partial \xi} - \frac{\partial u}{\partial \eta} \right) dt,$$

所以

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}, \quad \frac{\partial}{\partial t} = c \left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta} \right),$$

所以

$$c^2 \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} = c^2 \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right)^2 - c^2 \left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta} \right)^2 = 4c^2 \frac{\partial^2}{\partial \xi \eta},$$

所以解得波动方程的解为

$$u(x, t) = \Phi(\xi) + \Psi(\eta) = \Phi(x + ct) + \Psi(x - ct).$$

对照初始条件, 得到

$$u(x, 0) = \Phi(x) + \Psi(x) = f(x), \quad u'_t(x, 0) = c\Phi'(x) - c\Psi'(x) = g(x),$$

解得

$$\Phi'(x) = \frac{cf'(x) + g(x)}{2c}, \quad \Psi'(x) = \frac{cf'(x) - g(x)}{2c}.$$

取 g 的原函数 G , 则

$$\begin{aligned} \Phi(x) &= \frac{cf(x) + G(x) + A}{2c}, & \Psi(x) &= \frac{cf(x) - G(x) - A}{2c}, \\ u(x, t) &= \frac{cf(x + ct) + G(x + ct)}{2c} + \frac{cf(x - ct) - G(x - ct)}{2c} \\ &= \frac{f(x + ct) + f(x - ct)}{2} + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds. \end{aligned}$$

波动方程形如 $u(x, t) = \Phi(x + ct)$ (或 $u(x, t) = \Psi(x - ct)$) 的解叫做行波解(**traveling wave**), 它是由 D'Alembert 于 1747 年得到的。

求解波动方程的另一个办法是分离变量法: 即求形如 $u(x, t) = U(x)V(t)$ 的解。此时波动方程为

$$U(x)V''(t) = c^2 U''(x)V(t).$$

于是

$$c^2 \frac{U''(x)}{U(x)} = \frac{V''(t)}{V(t)},$$

它等价于方程组

$$\begin{cases} U''(x) = \lambda U(x), \\ V''(t) = c^2 \lambda V(t). \end{cases}$$

第二个方程有非常值的有界解当且仅当 $\lambda < 0$ 。设 $\lambda = -m^2$, 则第二个方程的解为

$$V(t) = A \cos(mct) + B \sin(mct).$$

第一个方程的解为

$$U(x) = \alpha \cos mx + \beta \sin mx.$$

如果弦的两端始终固定(这样的解称为驻波, **standing wave**), 且相距为 π (取适当的长度单位即可), 则

$$U(0) = U(\pi) = 0,$$

从而 $\alpha = 0$ 且 m 是正整数。因此波动方程的解为

$$u(x, t) = \sin mx (A_m \cos(mct) + B_m \sin(mct)).$$

波动方程是线性方程, 满足叠加原理, 所以驻波的通解为

$$\begin{aligned} u(x, t) &= \sum_{m=1}^{\infty} \sin mx (A_m \cos(mct) + B_m \sin(mct)) \\ &= \sum_{m=1}^{\infty} \left(\frac{A_m \sin m(x+ct) - B_m \cos m(x+ct)}{2} \right. \\ &\quad \left. + \frac{A_m \sin m(x-ct) + B_m \cos m(x-ct)}{2} \right). \end{aligned}$$

其中的系数 A_m, B_m 决定于初始条件:

$$\begin{cases} f(x) = u(x, 0) = \sum_{m=1}^{\infty} A_m \sin mx, \\ g(x) = u'_t(x, 0) = -c \sum_{m=1}^{\infty} m B_m \sin mx. \end{cases}$$

所以问题归结于是否存在以及如何找到满足这两个条件的系数 A_m, B_m , 或者说 f, g 是否可以写成上述形式的级数。

例 10.1.3 (热传导方程和 Laplace 方程). 热由热源向周围扩散的过程可以用以下微分方程刻画(平面情形时)

$$\frac{1}{c^2} \frac{\partial u}{\partial t} = \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

当扩散过程达到平衡态时, 温度场 $u(x, y, t)$ 不再随时间改变, 此时的温度场 $u(x, y)$ 满足 Laplace 方程

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

考虑单位圆盘区域上的 Laplace 方程以及给定的边界条件 $u = f(\theta)$ 。我们考虑极坐标系 (r, θ) 下的分离变量形式的解

$$u(r, \theta) = U(r)V(\theta).$$

则

$$U''(r)V(\theta) + \frac{1}{r}U'(r)V(\theta) + \frac{1}{r^2}U(r)V''(\theta) = 0,$$

即

$$\frac{r^2 U''(r) + rU'(r)}{U(r)} = -\frac{V''(\theta)}{V(\theta)},$$

这等价于方程组

$$\begin{cases} r^2 U''(r) + rU'(r) - \lambda U(r) = 0, \\ V''(\theta) + \lambda V(\theta) = 0. \end{cases}$$

第一个方程是 Euler 方程, 可以利用换元 $W(s) = U(e^s)$ 变为

$$W''(s) = U''(e^s)e^{2s} + U'(e^s)e^s = \lambda W(s),$$

第二个方程有 2π 周期解当且仅当 $\lambda = m^2$ (m 是非负整数)。于是

$$V(\theta) = A_m \cos m\theta + B_m \sin m\theta, \quad W(s) = \alpha_m e^{ms} + \beta_m e^{-ms}.$$

从而

$$U(r) = \alpha_m r^m + \beta_m r^{-m}.$$

考虑到 $u(x, y)$ 在原点连续, 所以 $\beta_m = 0$ 。再由叠加原理得到

$$u(r, \theta) = \sum_{m=0}^{\infty} r^m (A_m \cos m\theta + B_m \sin m\theta).$$

此时边界条件为

$$f(\theta) = u(1, \theta) = \sum_{m=0}^{\infty} (A_m \cos m\theta + B_m \sin m\theta).$$

因此, 只要我们能从边界条件 f 获得相应的系数 A_m, B_m , 我们就能得到 Laplace 方程对应于边界条件 f 的解。于是问题转化为: f 是否可以写成上述形式的级数?

10.2 Fourier 级数的定义与计算

为了研究一般的周期函数, 我们试图用最简单的一些周期函数进行逼近。

考虑以 $T = 2l$ 为周期的函数 $f(x)$, 以及由简谐振动(或者匀速圆周运动)产生的 T 周期函数

$$\cos \frac{2n\pi x}{T}, \quad \sin \frac{2n\pi x}{T}, \quad n = 1, 2, \dots$$

我们希望有以下表达

$$f(x) = A_0 + A_1 \cos \frac{2\pi x}{T} + B_1 \sin \frac{2\pi x}{T} + \dots + A_n \cos \frac{2n\pi x}{T} + B_n \sin \frac{2n\pi x}{T} + \dots$$

如果存在这样的表达方式, 我们希望用 f 来计算其中的系数 A_0, A_1, B_1, \dots 。

为此我们注意到如下事实

$$P = \int_{-l}^l \cos \frac{2m\pi x}{T} \cos \frac{2n\pi x}{T} dx = \begin{cases} 2l, & m = n = 0; \\ l, & m = n \text{ 是正整数}; \\ 0, & m \neq n. \end{cases}$$

$$Q = \int_{-l}^l \sin \frac{2m\pi x}{T} \sin \frac{2n\pi x}{T} dx = \begin{cases} l, & m = n \text{ 是正整数}; \\ 0, & \text{其他情形}. \end{cases}$$

$$R = \int_{-l}^l \sin \frac{2m\pi x}{T} \cos \frac{2n\pi x}{T} dx = 0. \quad (\text{被积函数是奇函数})$$

(上述结果可以通过计算 $P \pm Q$ 得到)。

于是对正整数 n ,

$$\int_{-l}^l f(x) \cos \frac{2n\pi x}{T} dx = lA_n,$$

$$\int_{-l}^l f(x) \sin \frac{2n\pi x}{T} dx = lB_n,$$

另外

$$\int_{-l}^l f(x) dx = 2lA_0,$$

因此

$$A_0 = \frac{1}{2l} \int_{-l}^l f(x) dx,$$

$$A_n = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{2n\pi x}{T} dx, \quad n \geq 1,$$

$$B_n = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{2n\pi x}{T} dx, \quad n \geq 1.$$

这些称为 f 的 **Euler-Fourier 系数**。这里积分区间 $[-l, l]$ 可以换成任何长度是 $T = 2l$ 的周期区间。此时我们记

$$f(x) \sim A_0 + A_1 \cos \frac{2\pi x}{T} + B_1 \sin \frac{2\pi x}{T} + \cdots + A_n \cos \frac{2n\pi x}{T} + B_n \sin \frac{2n\pi x}{T} + \cdots,$$

称后者为 f 的 **Fourier 级数**。

正弦级数和余弦级数

如果 f 是奇函数, 则所有 $A_n = 0$, 于是 f 的 Fourier 级数为

$$f(x) \sim B_1 \sin \frac{2\pi x}{T} + \cdots + B_n \sin \frac{2n\pi x}{T} + \cdots,$$

这称为一个 **正弦级数**。

如果 f 是偶函数, 则所有 $B_n = 0$, 于是 f 的 Fourier 级数为

$$f(x) \sim A_0 + A_1 \cos \frac{2\pi x}{T} + \cdots + A_n \cos \frac{2n\pi x}{T} + \cdots,$$

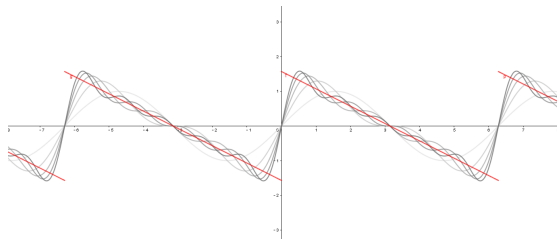
这称为一个余弦级数。

它们分别反映了函数 f 的对称性。

当我们被要求把一个区间上的函数展成正弦级数或余弦级数时，我们首先要把它对称延拓到一个对称区间上，然后再进行 Fourier 展开。

例 10.2.1. 考虑 2π 周期函数 $f(x)$ ，它在区间 $[0, 2\pi]$ 上满足 $f(x) = \frac{\pi-x}{2}$ ，我们计算它的 Fourier 级数。

不难看出 f 是奇函数，所以它的 Fourier 级数是正弦级数。



$$b_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi-x}{2} \sin nx dx = \frac{1}{2\pi} \int_0^{2\pi} -x \sin nx dx = \frac{1}{n},$$

所以 f 的 Fourier 级数为

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n},$$

记 $S(x)$ 为它的和函数，在例9.2.9中我们曾经证明了 $S(x) = \frac{\pi-x}{2}$ ($0 < x < 2\pi$)。易见

$$S(x) = \begin{cases} \frac{\pi-x}{2}, & 0 < x < 2\pi, \\ 0, & x = 0. \end{cases}$$

□

例 10.2.2. 在例9.2.10中我们曾证明了

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2} = \frac{\pi^2}{6} + \frac{x^2 - 2\pi x}{4}, \quad 0 \leq x \leq 2\pi.$$

记上式右端函数为 $f(x)$ 。由于上述级数一致收敛，所以 $f(x)$ 的 Fourier 系数可由逐项积分得到：

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx dx = \sum_{k=1}^{\infty} \frac{1}{k^2 \pi} \int_0^{2\pi} \cos kx \sin nx dx = 0, \\ a_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx dx = \sum_{k=1}^{\infty} \frac{1}{k^2 \pi} \int_0^{2\pi} \cos kx \cos nx dx = \frac{1}{n^2}, \\ a_0 &= 0. \end{aligned}$$

所以上述级数就是 f 的 Fourier 级数。

□

例 10.2.3. 考虑 2π 周期函数 $f(x) = x^2$ ($-\pi \leq x < \pi$), 求它的 Fourier 级数.

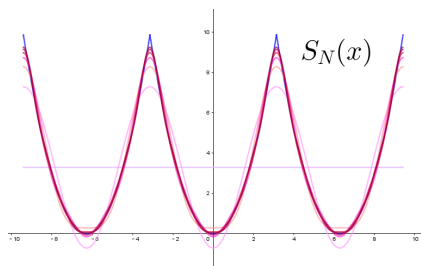
解: f 是偶函数, 所以它的 Fourier 级数是余弦级数,

$$A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{\pi^2}{3}, \quad A_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos nx dx = (\text{分部积分}) = \frac{4(-1)^n}{n^2}.$$

所以 f 的 Fourier 级数为

$$\frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^2} \cos nx.$$

由 Weierstrass 判别法, 该级数对任意 x 一致绝对收敛, 其和函数 $S(x)$ 连续.



因为(例9.2.10)

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2} = \frac{\pi^2}{6} + \frac{x^2 - 2\pi x}{4}, \quad 0 \leq x \leq 2\pi,$$

所以对 $-\pi \leq x \leq \pi$,

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nx = \sum_{n=1}^{\infty} \frac{\cos n(x+\pi)}{n^2} = \frac{\pi^2}{6} + \frac{(x+\pi)^2 - 2\pi(x+\pi)}{4} = \frac{3x^2 - \pi^2}{12}.$$

从而

$$\frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^2} \cos nx = x^2.$$

分别取 $x = 0$ 和 $x = \pi$, 得到

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} = \frac{\pi^2}{12}, \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2 - \frac{\pi^2}{3}}{4} = \frac{\pi^2}{6},$$

进而得到

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{\pi^2}{8}.$$

你还可以取 $x = \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}$ 看看可以得到哪些结论。 □

10.3 平方平均误差与 Fourier 级数的最小二乘性质

区间 $[a, b]$ 上两个函数 f, g , 其平方平均误差被定义为

$$\int_a^b |f(x) - g(x)|^2 dx.$$

为此, 我们需要 f, g 在区间 $[a, b]$ 上平方可积, 即

$$\int_a^b |f(x)|^2 dx < +\infty, \quad \int_a^b |g(x)|^2 dx < +\infty.$$

由 Cauchy-Schwarz 不等式, 我们知道有界区间 $[a, b]$ 上的平方可积函数都是绝对可积的,

$$\int_a^b |f(x)| dx \leq \sqrt{\int_a^b |f(x)|^2 dx \int_a^b 1^2 dx} = \sqrt{b-a} \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}} < +\infty.$$

记 $\mathcal{L}^2[-l, l]$ 为区间 $[-l, l]$ 上所有平方可积函数组成的集合. 对 $f, g \in \mathcal{L}^2[-l, l]$ 定义

$$\langle f, g \rangle = \int_{-l}^l f(x)g(x) dx.$$

这是一个半正定的双线性函数.

$$\langle f, f \rangle = \int_{-l}^l |f(x)|^2 dx \geq 0.$$

在 $\mathcal{L}^2[-l, l]$ 中, 如果 $\int_{-l}^l |f(x)|^2 dx = 0$, 我们就认为 $f = 0$. 于是上述双线性函数是正定的, 它定义了 $\mathcal{L}^2[-l, l]$ 上的一个内积, 该内积导致了一个范数 $\|\cdot\|_2$, 此时 f, g 的平方平均误差就是 $\|f - g\|_2^2$.

我们称函数列 f_n 在 \mathcal{L}^2 意义下收敛于 f , 如果

$$\lim_{n \rightarrow +\infty} \|f_n - f\|_2 = 0,$$

即

$$\lim_{n \rightarrow +\infty} \int_{-l}^l |f_n(x) - f(x)|^2 dx = 0.$$

在上述内积下,

$$\mathbf{c}_0 = \frac{1}{\sqrt{2l}}, \quad \mathbf{c}_n = \frac{1}{\sqrt{l}} \cos \frac{n\pi x}{l}, \quad \mathbf{s}_n = \frac{1}{\sqrt{l}} \sin \frac{n\pi x}{l}, \quad n = 1, 2, \dots$$

成为 $\mathcal{L}^2[-l, l]$ 中一列单位正交向量, f 的 Euler-Fourier 系数

$$\begin{aligned} A_0 &= \frac{1}{2l} \int_{-l}^l f(x) dx = \frac{1}{\sqrt{2l}} \langle f, \mathbf{c}_0 \rangle, \\ A_n &= \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi x}{l} dx = \frac{1}{\sqrt{l}} \langle f, \mathbf{c}_n \rangle, \\ B_n &= \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi x}{l} dx = \frac{1}{\sqrt{l}} \langle f, \mathbf{s}_n \rangle, \end{aligned}$$

所以 f 的 Fourier 级数的前 $2N + 1$ 项的部分和

$$\begin{aligned} f_N(x) &= A_0 + \sum_{n=1}^N \left(A_n \cos \frac{2n\pi x}{T} + B_n \sin \frac{2n\pi x}{T} \right) \\ &= \langle f, \mathbf{c}_0 \rangle \mathbf{c}_0 + \sum_{n=1}^N (\langle f, \mathbf{c}_n \rangle \mathbf{c}_n + \langle f, \mathbf{s}_n \rangle \mathbf{s}_n), \end{aligned}$$

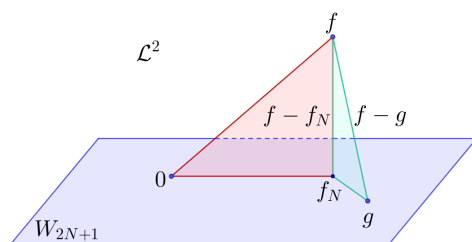
它是 f 在

$$\mathbf{c}_0, \mathbf{c}_1, \mathbf{s}_1, \dots, \mathbf{c}_N, \mathbf{s}_N$$

所张成的 $2N + 1$ 维空间 W_{2N+1} 中的正交投影, 这也是这个 $2N + 1$ 维空间中与 f 平均平方误差最小的一个, 即

$$\|f - f_N\|_2 = \min_{g \in W_{2N+1}} \|f - g\|_2.$$

另外, 由于 $f - f_N \perp f_N$, 所以



$$\|f\|_2^2 = \|f - f_N\|_2^2 + \|f_N\|_2^2.$$

于是, 一方面

$$\|f_N\|_2^2 \leq \|f\|_2^2,$$

即

定理 10.3.1 (Bessel 不等式). 对任意 $f \in \mathcal{L}^2[-l, l]$, f 的 Euler-Fourier 系数满足

$$A_0^2 + \frac{1}{2} \sum_{n=1}^N (A_n^2 + B_n^2) \leq \frac{1}{2l} \int_{-l}^l |f(x)|^2 dx.$$

推论 10.3.2. 任意 $f \in \mathcal{L}^2[-l, l]$ 的 Euler-Fourier 系数满足

$$A_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2 + B_n^2)$$

必然是收敛的。

例 10.3.3. 证明 $\sum_{n=1}^{\infty} \frac{\sin nx}{\ln(n+1)}$ 收敛, 但不是任何平方可积函数的 Fourier 级数。

证明. 由 Dirichlet 判别法知 $\sum_{n=1}^{\infty} \frac{\sin nx}{\ln(n+1)}$ 对任何实数 x 收敛。但由于

$$\sum_{n=1}^{\infty} \left(\frac{1}{\ln(n+1)} \right)^2$$

发散, 所以由 Bessel 不等式知 $\sum_{n=1}^{\infty} \frac{\sin nx}{\ln(n+1)}$ 不是任何平方可积函数的 Fourier 级数。从而该级数的和函数不是平方可积函数。□

另一方面, 为证明 $\lim_{N \rightarrow +\infty} \|f - f_N\|_2 = 0$, 只需证明

定理 10.3.4 (Parseval 等式). 对任意 $f \in \mathcal{L}^2[-l, l]$, f 的 Euler-Fourier 系数满足

$$A_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2 + B_n^2) = \frac{1}{2l} \int_{-l}^l |f(x)|^2 dx.$$

其证明见本章最后一节。

推论 10.3.5. 若 $f, g \in \mathcal{L}^2[-l, l]$ 具有完全相同的 Fourier 级数, 则对几乎所有的 x 都有 $f(x) = g(x)$ 。

证明. 当 $N \rightarrow +\infty$ 时,

$$\|f - g\|_2 \leq \|f_N - f\|_2 + \|f_N - g_N\|_2 + \|g_N - g\|_2 = \|f_N - f\|_2 + \|g_N - g\|_2 \rightarrow 0.$$

□

推论 10.3.6.

$$A_0\alpha_0 + \frac{1}{2} \sum_{n=1}^{\infty} (A_n\alpha_n + B_n\beta_n) = \frac{1}{2l} \int_{-l}^l f(x)g(x)dx,$$

其中 A_0, A_n, B_n 和 $\alpha_0, \alpha_n, \beta_n$ 分别是 $f, g \in \mathcal{L}^2[-l, l]$ 的 Euler-Fourier 系数。

证明. 这个等式可以通过

$$\int_{-l}^l f(x)g(x)dx = \frac{1}{4} \left(\int_{-l}^l (f(x) + g(x))^2 dx - \int_{-l}^l (f(x) - g(x))^2 dx \right)$$

得到, 后面这个等式本质上是内积及其对应范数之间的如下结论:

$$\|f + g\|^2 - \|f - g\|^2 = 4\langle f, g \rangle.$$

□

利用 Parseval 等式, 我们可以求得一些特殊的级数的和。

例 10.3.7. 我们已经知道

$$x^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^2} \cos nx, \quad \forall -\pi \leq x \leq \pi.$$

对它用 Parseval 等式, 得到

$$\frac{\pi^4}{5} = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^4 dx = \frac{\pi^4}{9} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{16}{n^4},$$

进而

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

利用逐项积分, 得到

$$\frac{x^3}{3} = \int_0^x t^2 dt = \frac{\pi^2}{3} x + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^3} \sin nx, \quad \forall -\pi \leq x \leq \pi,$$

这样得到 2π 周期函数 $\frac{x^3}{3} - \frac{\pi^2 x}{3}$ ($-\pi \leq x < \pi$) 的 Fourier 级数。取 $x = \frac{\pi}{2}$, 得到

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^3} = \frac{\pi^3}{32}.$$

对这个 Fourier 级数使用 Parseval 等式, 我们可以得到 $\sum_{n=1}^{\infty} \frac{1}{n^6}$ 的值。Euler 知道所有到 $\sum_{n=1}^{\infty} \frac{1}{n^{2m}}$ 的值, 但 $\sum_{n=1}^{\infty} \frac{1}{n^{2m-1}}$ 的值至今无人知晓。□

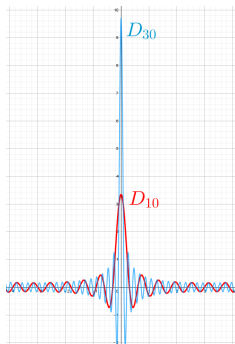
10.4 Fourier 级数的收敛性

1966 年, L.Carleson 证明了任何平方可积函数 f , 其 Fourier 级数几乎处处收敛到 f 自身。1968 年, R.A.Hunt 把这个结论推广到 p ($p > 1$) 次绝对可积函数。然而, 早在 1923/26 年 Kolmogorov 证明了: 存在绝对可积函数 f , 其 Fourier 级数几乎处处/处处发散。更早的时候, 1873 年 Paul du Bois-Reymond 证明存在连续函数 f , 其 Fourier 级数在某处发散。由此可见 Fourier 级数的逐点收敛问题是一个比较难的数学问题。

不妨设 $l = \pi$, 则

$$\begin{aligned} f_N(x) &= \frac{\int_{-\pi}^{\pi} f(t) dt}{2\pi} + \sum_{n=1}^N \frac{\cos nx \int_{-\pi}^{\pi} f(t) \cos ntdt + \sin nx \int_{-\pi}^{\pi} f(t) \sin ntdt}{\pi} \\ &= \int_{-\pi}^{\pi} f(t) \left(\frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^N \cos n(x-t) \right) dt = \int_{-\pi}^{\pi} f(t) D_N(x-t) dt \\ &= \int_{-\pi}^{\pi} f(x-t) D_N(t) dt = \int_{-\pi}^{\pi} f(x+t) D_N(-t) dt \\ &= \int_{-\pi}^{\pi} f(x+t) D_N(t) dt \\ &= \int_{-\pi}^{\pi} \frac{f(x+t) + f(x-t)}{2} D_N(t) dt, \end{aligned}$$

从而



$$f_N(x) - \frac{f(x_-) + f(x_+)}{2} = \int_{-\pi}^{\pi} \frac{f(x+t) + f(x-t) - f(x_+) - f(x_-)}{2} D_N(t) dt,$$

其中 $f(x_-), f(x_+)$ 是 f 在 x 处的左极限和右极限,

$$D_N(t) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^N \cos nt = \frac{\sin(N + \frac{1}{2})t}{2\pi \sin \frac{t}{2}}$$

称为 **Dirichlet 核**, 它满足

$$\int_{-\pi}^{\pi} D_N(t) dt = 1, \quad \lim_{N \rightarrow +\infty} \int_{\delta < |t| \leq \pi} D_N(t) dt = 0.$$

其中第二个积分表明, 虽然 Euler-Fourier 系数是整个周期区间上的与 f 有关的积分, 但函数 f 的 Fourier 级数在点 x 处的值实际上只由 f 在 x 附近的值决定。利用上述结果可以证明

定理 10.4.1. 如果 2π 周期函数 f 在 $[-\pi, \pi]$ 上绝对可积, 并且在 x 处满足以下条件之一

1. (Dini) 瑕积分 $\int_0^{\delta} \frac{|f(x+t) - f(x_+)| + |f(x-t) - f(x_-)|}{t} dt$ 收敛;

2. (Hölder) 存在 $M > 0$ 和 $0 < \alpha \leq 1$ 使得 $|f(x+t) - f(x_+)| \leq Mt^\alpha$,
 $|f(x-t) - f(x_-)| \leq Mt^\alpha$;

3. f' 在 $[-\pi, \pi]$ 上分段连续(间断点个数有限, 且为第一类间断点),

则 f 的 Fourier 级数在 x 处收敛到 $\frac{f(x_+) + f(x_-)}{2}$ 。特别地, 当 f 在 x 处连续时,
 $\lim_{N \rightarrow +\infty} f_N(x) = f(x)$ 。 \square

定理 10.4.2. 任何 2π 周期的 C^1 函数 f 的 Fourier 级数都一致收敛到 f 自身。 \square

推论 10.4.3. 对任何 2π 周期的 C^2 函数 f , f' 的 Fourier 级数可以从 f 的 Fourier 级数逐项得到。

证明. 记

$$A_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx, \quad B_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx.$$

则

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) \cos nx dx = \frac{f(x) \cos nx}{\pi} \Big|_{-\pi}^{\pi} + \frac{n}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = nB_n,$$

同理可得 $\frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) \sin nx dx = -nA_n$ 。所以

$$f'(x) = \sum_{n=1}^{\infty} (nB_n \cos nx - nA_n \sin nx).$$

\square

例 10.4.4. 设 $t \in \mathbb{R} \setminus \mathbb{Z}$ 。求 $\cos tx$ ($-\pi \leq x \leq \pi$) 的 Fourier 级数。

解: $\cos tx$ ($-\pi \leq x \leq \pi$) 是 C^∞ 的偶函数,

$$\begin{aligned} A_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos tx \cos nx dx = \frac{1}{2\pi} \left[\frac{\sin(t+n)x}{t+n} + \frac{\sin(t-n)x}{t-n} \right]_{-\pi}^{\pi} \\ &= \frac{1}{\pi} \left[\frac{\sin(t+n)\pi}{t+n} + \frac{\sin(t-n)\pi}{t-n} \right] = \frac{(-1)^n 2t \sin t\pi}{\pi(t^2 - n^2)}, \\ A_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos tx dx = \frac{\sin t\pi}{\pi t}, \end{aligned}$$

因此

$$\cos tx = \frac{\sin t\pi}{t\pi} + \frac{2t \sin t\pi}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{t^2 - n^2} \cos nx.$$

取 $x = \pi$, 两边除以 $\frac{\sin t\pi}{\pi}$, 移项, 得到

$$\pi \cot \pi t - \frac{1}{t} = \sum_{n=1}^{\infty} \frac{-2t}{n^2 - t^2}. \quad (\text{Euler})$$

上式两边对 t 从 0 到 $x \in (0, 1)$ 积分,

$$\ln \frac{\sin \pi x}{\pi x} = \sum_{n=1}^{\infty} \ln \left(1 - \frac{x^2}{n^2} \right),$$

即

$$\frac{\sin \pi x}{\pi x} = \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2} \right),$$

从而对 $0 < x < \pi$,

$$\frac{\sin x}{x} = \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2 \pi^2} \right). \quad (\text{Euler})$$

这样我们严格地证明了 Euler 当年的重要发现。

Euler 利用这个无穷乘积以及 $\sin x$ 的 Maclaurin 级数并借鉴多项式的根与系数的韦达定理得到了 $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, 这解决了困惑人们几十年的著名难题。

事实上, 我们也可以利用极限, 严格得到这个数项级数的和

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^2} &= \lim_{t \rightarrow 0} \sum_{n=1}^{\infty} \frac{1}{n^2 - t^2} = \lim_{t \rightarrow 0} \left(\frac{1}{2t^2} - \frac{\pi}{2t} \cot(\pi t) \right) \\ &= \lim_{t \rightarrow 0} \frac{\sin(\pi t) - \pi t \cos(\pi t)}{2t^2 \sin(\pi t)} = \frac{\pi^2}{6}. \end{aligned}$$

□

10.5 Fourier 级数应用

等周问题

在给定周长的所有平面曲线中, 哪一个能围成最大的面积? 这是一个很古老的几何问题, 它的提出很可能是人类丈量土地的需要, 而且也很可能知道答案是圆周, 但给出证明却花了人类两千多年的时间。远在公元前二世纪, 古希腊人 Zenodorus 就研究过这个问题。直到十九世纪, 符合现代数学要求的证明才正式被提出。

定理 10.5.1 (等周不等式). 对于平面中任意一条简单封闭 C^2 曲线 γ , 其周长 L 和它所围成的有界区域的面积 A 都满足 $4\pi A \leq L^2$, 其中等号成立当且仅当 γ 是一个圆周。

证明. 设 $\gamma: (x(t), y(t))$ 是平面中一个 C^2 简单封闭曲线, $t \in [0, L]$ 是它的弧长参数, L 是 γ 的弧长。因为 γ 是封闭曲线, 所以我们可以视 $x(t), y(t)$ 是 L -周期函数。设

$$\begin{cases} x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{2n\pi t}{L} + b_n \sin \frac{2n\pi t}{L}, \\ y(t) = \alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos \frac{2n\pi t}{L} + \beta_n \sin \frac{2n\pi t}{L}. \end{cases}$$

则

$$\begin{cases} x'(t) = \frac{2\pi}{L} \sum_{n=1}^{\infty} -na_n \sin \frac{2n\pi t}{L} + nb_n \cos \frac{2n\pi t}{L}, \\ y'(t) = \frac{2\pi}{L} \sum_{n=1}^{\infty} -n\alpha_n \sin \frac{2n\pi t}{L} + n\beta_n \cos \frac{2n\pi t}{L}. \end{cases}$$

由于是弧长参数, 所以 $x'(t)^2 + y'(t)^2 = 1$, 从而

$$\begin{aligned} L^2 &= L \int_0^L x'(t)^2 + y'(t)^2 dt = L \left(\langle x'(t), x'(t) \rangle + \langle y'(t), y'(t) \rangle \right) \\ &= 2\pi^2 \sum_{n=1}^{\infty} n^2 (a_n^2 + b_n^2 + \alpha_n^2 + \beta_n^2), \end{aligned}$$

曲线 γ 围成的平面区域的面积为

$$\begin{aligned} A &= \frac{1}{2} \int_{\gamma} xdy - ydx = \frac{1}{2} \int_0^L x(t)y'(t) - y(t)x'(t) dt \\ &= \frac{1}{2} \left(\langle x(t), y'(t) \rangle - \langle y(t), x'(t) \rangle \right) = \pi \sum_{n=1}^{\infty} n(-\alpha_n b_n + a_n \beta_n) \\ &\leq \frac{\pi}{2} \sum_{n=1}^{\infty} n(\alpha_n^2 + b_n^2 + a_n^2 + \beta_n^2) \leq \frac{\pi}{2} \sum_{n=1}^{\infty} n^2 (\alpha_n^2 + b_n^2 + a_n^2 + \beta_n^2) = \frac{L^2}{4\pi}. \end{aligned}$$

等号成立当且仅当

$$\alpha_1 = -b_1, \quad \beta_1 = a_1, \quad a_n = b_n = \alpha_n = \beta_n = 0, \quad n \geq 2.$$

此时

$$\begin{cases} x(t) = a_0 + a_1 \cos \frac{2\pi t}{L} + b_1 \sin \frac{2\pi t}{L}, \\ y(t) = \alpha_0 - b_1 \cos \frac{2\pi t}{L} + a_1 \sin \frac{2\pi t}{L}, \end{cases}$$

这是一个圆周。 □

常微分方程周期边值问题

例 10.5.2. 求以下二阶微分方程满足周期边界条件的解。

$$\begin{cases} y'' + \lambda y = \sin x, \\ y(0) = y(2\pi), \\ y'(0) = y'(2\pi). \end{cases}$$

解: 我们先证明满足上述边值条件的解 $y(x)$ 是 2π 周期函数。

令 $u(x) = y(x + 2\pi) - y(x)$ 。则

$$u''(x) + \lambda u(x) = y''(x + 2\pi) - y''(x) + \lambda[y(x + 2\pi) - y(x)] = \sin(x + 2\pi) - \sin(x) = 0,$$

并且

$$u(0) = y(2\pi) - y(0) = 0, \quad u'(0) = y'(2\pi) - y'(0) = 0.$$

由微分方程初值问题解的存在唯一性, 对任意 x , $u(x) = 0$ 。因此 $y(x + 2\pi) = y(x)$, 从而 $y(x)$ 是 2π 周期函数。并且由方程知 $y(x)$ 是 C^∞ 函数。

设 $y(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$ 。逐项求导并代入微分方程得到

$$\lambda a_0 + \sum_{n=1}^{\infty} [(\lambda - n^2)a_n \cos nx + (\lambda - n^2)b_n \sin nx] = \sin x.$$

如果 $\lambda = 0$, 则

$$-a_1 = 0, \quad -n^2 a_n = 0, \quad -b_1 = 1, \quad -n^2 b_n = 0, \quad n \geq 2,$$

解得无穷多解

$$y(x) = a_0 - \sin x.$$

如果 $\lambda = k^2$ ($k \geq 2$ 是正整数), 则

$$\begin{cases} a_0 = 0, \\ (\lambda - n^2)a_n = 0, & n \neq k \\ (\lambda - k^2)a_k = 0. \end{cases} \quad \begin{cases} (\lambda - n^2)b_n = 0, & n \neq 1 \\ (\lambda - 1)b_1 = 1, \end{cases}$$

解得无穷多解

$$y(x) = \frac{1}{k^2 - 1} \sin x + a_k \cos kx + b_k \sin kx.$$

如果 $\lambda = 1$, 则

$$(\lambda - 1)b_1 = 1,$$

无解。

如果 $\lambda \neq k^2$ (k 是整数), 则

$$\begin{cases} a_0 = 0, \\ (\lambda - n^2)a_n = 0, \end{cases} \quad \begin{cases} (\lambda - n^2)b_n = 0, & n \neq 1 \\ (\lambda - 1)b_1 = 1, \end{cases}$$

解得唯一解

$$y(x) = \frac{1}{\lambda - 1} \sin x.$$

□

考虑关于 2π 周期 C^2 函数 y 的线性映射

$$Ay = -y''.$$

则对 2π 周期 C^2 函数 y, z ,

$$\begin{aligned} \langle Ay, z \rangle &= \int_{-\pi}^{\pi} -y''(x)z(x)dx \\ &= - \int_{-\pi}^{\pi} z(x)dy'(x) = -y'(x)z(x)|_{-\pi}^{\pi} + \int_{-\pi}^{\pi} y'(x)z'(x)dx \\ &= \langle y', z' \rangle = \langle y, Az \rangle, \end{aligned}$$

所以 A 是对称线性映射。 $Ay = \lambda y$, 即 $-y'' = \lambda y$ 。

当 $\lambda < 0$ 时, 这个方程的通解为 $y(x) = C_1 e^{\sqrt{-\lambda}x} + C_2 e^{-\sqrt{-\lambda}x}$, 这样的解有界当且仅当 $C_1 = C_2 = 0$, 即 $y = 0$ 。

当 $\lambda = 0$ 时, 这个方程的通解为 $y(x) = C_1 + C_2 x$, 这样的解有界当且仅当 $C_2 = 0$, 即 $y = C_1$, 它也是 2π 周期解。

当 $\lambda > 0$ 时, 这个方程的通解为 $y(x) = C_1 \cos(\sqrt{\lambda}x) + C_2 \sin(\sqrt{\lambda}x)$, 这样的解有界。它是 2π 周期解当且仅当 $\sqrt{\lambda}$ 是正整数, 即 $\lambda = n^2$ 。

所以 A 的特征函数为

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx, \dots$$

设 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots$ 是线性映射 A 的特征向量, $\lambda_1, \lambda_2, \dots, \lambda_n, \dots$ 为对应的特征值。对向量 \mathbf{b} 和数 λ , 求解方程

$$(A - \lambda)\mathbf{y} = \mathbf{b},$$

可以先做分解

$$\begin{aligned} \mathbf{y} &= a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_n \mathbf{e}_n + \dots, \\ \mathbf{b} &= b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + \dots + b_n \mathbf{e}_n + \dots, \end{aligned}$$

上述方程称为

$$(\lambda_k - \lambda)a_k = b_k, \quad k = 1, 2, \dots$$

如果 λ 不是 A 的特征值, 则上述方程有唯一解

$$\mathbf{y} = \frac{b_1}{\lambda_1 - \lambda} \mathbf{e}_1 + \frac{b_2}{\lambda_2 - \lambda} \mathbf{e}_2 + \dots + \frac{b_n}{\lambda_n - \lambda} \mathbf{e}_n + \dots$$

如果 $\lambda = \lambda_m$, 则上述方程有解当且仅当 $b_m = 0$, 此时方程有无穷多解

$$\mathbf{y} = a_m \mathbf{e}_m + \sum_{k \neq m} \frac{b_k}{\lambda_k - \lambda_m} \mathbf{e}_k.$$

微分方程的边值问题与初值问题不同, 它把我们从有限维空间代入到无穷维空间, 而 Fourier 级数理论对应于无穷维内积空间中一些对称线性算子的谱性质。从应用的角度讲, Fourier 级数也是一种信息压缩, 它把一个连续变量的函数对应为一个数列 (Fourier 系数)。所以无论在理论上还是在应用中, 它都是重要的工具。

伽利略说, 自然这本大书是由上帝之手用数学的语言写成的。牛顿说, 自然法则是用微分方程表达的。从三百多年前牛顿创立微积分开始, 人们就试图发展各种数学工具来实现牛顿的梦想——通过研究微分方程来理解它描述的自然法则。现代科学正是肇始于此, 并随之逐渐发展壮大, 于是才有了人类今天的文明。

10.6 附: Parseval 等式的证明

定理 10.6.1. 对 Riemann 可积的 2π 周期函数 f ,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = A_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2 + B_n^2),$$

其中 A_n, B_n 是 f 的 Euler-Fourier 系数。

引理 10.6.2. Parseval 等式对所有三角多项式

$$f(x) = A_0 + \sum_{n=1}^N (A_n \cos nx + B_n \sin nx)$$

成立。

证明. 直接计算即可。 \square

推论 10.6.3. 对任何 Riemann 可积函数 $f: [0, 2\pi] \rightarrow \mathbb{R}$ 都有: f 满足 Parseval 等式当且仅当 $\lim_{N \rightarrow +\infty} \int_0^{2\pi} |f(x) - S_N(x)|^2 dx = 0$ 。

证明. 这是因为 $\int_0^{2\pi} |f(x)|^2 dx = \int_0^{2\pi} |f(x) - S_N(x)|^2 dx + \int_0^{2\pi} |S_N(x)|^2 dx$ 。 \square

引理 10.6.4. Parseval 等式对

$$f(x) = \begin{cases} 1, & x \in [\alpha, \beta] \subseteq [0, 2\pi], \\ 0, & x \in [0, 2\pi] \setminus [\alpha, \beta] \end{cases}$$

成立。

证明.

$$\begin{aligned} A_0 &= \frac{1}{2\pi} \int_{\alpha}^{\beta} dx = \frac{\beta - \alpha}{2\pi}, \\ A_n &= \frac{1}{\pi} \int_{\alpha}^{\beta} \cos nx dx = \frac{\sin n\beta - \sin n\alpha}{n\pi}, \\ B_n &= \frac{1}{\pi} \int_{\alpha}^{\beta} \sin nx dx = \frac{\cos n\alpha - \cos n\beta}{n\pi}, \end{aligned}$$

所以

$$\begin{aligned} A_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2 + B_n^2) &= \frac{(\beta - \alpha)^2}{4\pi^2} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{2 - 2\cos n(\beta - \alpha)}{n^2\pi^2} \\ &= \frac{1}{\pi^2} \left[\frac{(\beta - \alpha)^2}{4} + \frac{\pi^2}{6} - \sum_{n=1}^{\infty} \frac{\cos n(\beta - \alpha)}{n^2} \right] \\ &= \frac{1}{\pi^2} \left[\frac{(\beta - \alpha)^2}{4} + \frac{\pi^2}{6} - \frac{\pi^2}{6} - \frac{(\beta - \alpha)^2 - 2\pi(\beta - \alpha)}{4} \right] \\ &= \frac{1}{2\pi} (\beta - \alpha) \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx. \end{aligned}$$

其中第三个等号使用了例9.2.10的结论。□

引理 10.6.5. 设 f_1, f_2 都是 Riemann 可积的 2π 周期函数, 满足

$$\int_0^{2\pi} f_1(x)f_2(x)dx = 0.$$

若 Parseval 等式对 f_1, f_2 都成立, 则 Parseval 等式对 $f_1 + f_2$ 也成立。

证明. 记 $A_n(f_k), B_n(f_k)$ 为 f_k 的 Euler-Fourier 系数。则易见

$$A_n(f_1 \pm f_2) = A_n(f_1) \pm A_n(f_2), \quad B_n(f_1 \pm f_2) = B_n(f_1) \pm B_n(f_2).$$

则

$$\begin{aligned} A_n^2(f_1 + f_2) + A_n^2(f_1 - f_2) &= 2(A_n^2(f_1) + A_n^2(f_2)), \\ B_n^2(f_1 + f_2) + B_n^2(f_1 - f_2) &= 2(B_n^2(f_1) + B_n^2(f_2)). \end{aligned}$$

由假设条件知

$$\frac{1}{2\pi} \int_0^{2\pi} |f_k(x)|^2 dx = A_0^2(f_k) + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2(f_k) + B_n^2(f_k)), \quad k = 1, 2.$$

因此

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} |f_1(x) \pm f_2(x)|^2 dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f_1(x)|^2 dx + \frac{1}{2\pi} \int_0^{2\pi} |f_2(x)|^2 dx \\ &= A_0^2(f_1) + A_0^2(f_2) + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2(f_1) + B_n^2(f_1) + A_n^2(f_2) + B_n^2(f_2)) \\ &= \frac{A_0^2(f_1 + f_2) + A_0^2(f_1 - f_2)}{2} \\ & \quad + \frac{1}{2} \sum_{n=1}^{\infty} \frac{A_n^2(f_1 + f_2) + B_n^2(f_1 + f_2) + A_n^2(f_1 - f_2) + B_n^2(f_1 - f_2)}{2}, \end{aligned}$$

由 Bessel 不等式知

$$A_0^2(f_1 \pm f_2) + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2(f_1 \pm f_2) + B_n^2(f_1 \pm f_2)) \leq \frac{1}{2\pi} \int_0^{2\pi} |f_1(x) \pm f_2(x)|^2 dx,$$

所以上述等式表明

$$A_0^2(f_1 \pm f_2) + \frac{1}{2} \sum_{n=1}^{\infty} (A_n^2(f_1 \pm f_2) + B_n^2(f_1 \pm f_2)) = \frac{1}{2\pi} \int_0^{2\pi} |f_1(x) \pm f_2(x)|^2 dx,$$

即 Parseval 等式对 $f_1 \pm f_2$ 都成立。□

由上述两个引理知

推论 10.6.6. Parseval 等式对任何阶梯函数成立。

证明. 设 $0 = a_0 < a_1 < \cdots < a_n = 2\pi$, $f(x) = f_1(x) + \cdots + f_n(x)$, 其中 $f_k(x) = c_k I_{[a_{k-1}, a_k]}(x)$. 则由第一个引理知, Parseval 等式对 f_1, \dots, f_n 都成立. 另一方面, f_1, \dots, f_n 彼此正交, $\int_0^{2\pi} f_i(x) f_j(x) dx = 0$ ($1 \leq i < j \leq n$), 所以由第二个引理知 Parseval 等式对 $f = f_1 + \cdots + f_n$ 成立. \square

定理的证明. 设 f 是区间 $[0, 2\pi]$ 上的 Riemann 可积函数. 存在常数 $M > 0$, 并且对任意 $\varepsilon > 0$, 存在阶梯函数 g, h 使得 $-M < g(x) \leq f(x) \leq h(x) < M$, 且 $\int_0^{2\pi} |h(x) - g(x)| dx < \varepsilon$.

于是

$$\int_0^{2\pi} |f(x) - g(x)|^2 dx \leq \int_0^{2\pi} |h(x) - g(x)|^2 dx \leq 2M \int_0^{2\pi} |h(x) - g(x)| dx < 2M\varepsilon.$$

记 $S_N(f)(x) = \frac{A_0(f)}{2} + \sum_{n=1}^N [A_n(f) \cos nx + B_n(f) \sin nx]$. 则

$$\begin{aligned} \int_0^{2\pi} |f(x) - S_N(f)(x)|^2 dx &\leq \int_0^{2\pi} |f(x) - S_N(g)(x)|^2 dx \\ &\leq \int_0^{2\pi} |f(x) - g(x)|^2 dx + \int_0^{2\pi} |g(x) - S_N(g)(x)|^2 dx \\ &\leq 2M\varepsilon + \int_0^{2\pi} |g(x) - S_N(g)(x)|^2 dx \rightarrow 0, \end{aligned}$$

所以 $\lim_{N \rightarrow +\infty} \int_0^{2\pi} |f(x) - S_N(f)(x)|^2 dx = 0$. 因此 Parseval 等式对 f 成立. \square

10.7 习题讨论课

1. 求 x^2 ($0 \leq x \leq \pi$) 的正弦级数(周期 2π), 并分别求级数 $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^3}$, $\sum_{n=1}^{\infty} \frac{1}{n^6}$ 的和。

2. 证明 x^4 ($0 \leq x \leq \pi$) 的余弦级数(周期 2π) 为

$$x^4 = \frac{\pi^4}{5} + 8 \sum_{n=1}^{\infty} (-1)^n \frac{\pi^2 n^2 - 6}{n^4} \cos nx,$$

并利用 Parseval 等式求级数 $\sum_{n=1}^{\infty} \frac{1}{n^8}$ 的和。

3. 求 e^x ($x \in [-\pi, \pi)$) 的 Fourier 级数, 并分别求级数 $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$, $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2+1}$ 的和。

4. 设 $f: [a, b] \rightarrow \mathbb{R}$ 连续且分段可微, f' 在 $[a, b]$ 上平方可积。

- (a) 若 $[a, b] = [0, \pi]$, 且 f 满足 $f(0) = f(\pi)$ 或 $\int_0^\pi f(x) dx = 0$, 证明

$$\int_0^\pi f^2(x) dx \leq \int_0^\pi (f')^2(x) dx,$$

其中等号成立当且仅当 $f(x) = a \cos x$ 。

- (b) 若 $[a, b] = [-\pi, \pi]$, 且 f 满足 $f(-\pi) = f(\pi)$ 以及 $\int_{-\pi}^\pi f(x) dx = 0$, 证明

$$\int_{-\pi}^\pi f^2(x) dx \leq \int_{-\pi}^\pi (f')^2(x) dx,$$

其中等号成立当且仅当 $f(x) = a \cos x + b \sin x$ 。

5. 幂级数与 Fourier 级数。利用

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad e^{ix} = \cos x + i \sin x$$

证明

$$\sum_{n=0}^{\infty} \frac{\cos(nx)}{n!} = e^{\cos x} \cos(\sin x), \quad \sum_{n=1}^{\infty} \frac{\sin(nx)}{n!} = e^{\cos x} \sin(\sin x).$$

并求

$$\sum_{n=0}^{\infty} \frac{(-1)^n \cos 2nx}{(2n)!}, \quad \sum_{n=0}^{\infty} \frac{(-1)^n \sin 2nx}{(2n)!},$$

$$\sum_{n=0}^{\infty} \frac{(-1)^n \cos(2n+1)x}{(2n+1)!}, \quad \sum_{n=0}^{\infty} \frac{(-1)^n \sin(2n+1)x}{(2n+1)!}$$

的和函数表达式。

