# 数 学 分 析 入 门

（第二册）

陆亚明

# 目 录

第九章

# 数项级数

> 发散级数是魔鬼的发明，把不管什么样的证明都建立在发散级数的基础上是一种耻辱。利用发散级数人们想要什么结论就可以得到什么结论，这也是为什么发散级数已经产生如此多的谬论和悖论的原因。
>
> —— N. H. Abel

## §9.1 级数的概念及收敛性

**【定义 1.1】** 设 $\{u_n\}$ 是一个数列，我们引入形式的记号

$$\sum_{n=1}^{\infty} u_n = u_1 + u_2 + \cdots + u_n + \cdots, \tag{9.1}$$

并称之为无穷级数，或简称为级数 (series)。称 $u_n$ 为该级数的通项。

为了说明 (9.1) 中形式的求和是否有确切的意义，我们下面给出级数收敛的概念。

**【定义 1.2】** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是一个级数，我们记

$$S_n = \sum_{k=1}^{n} u_k,$$

并称之为级数 $\sum\limits_{n=1}^{\infty} u_n$ 的第 $n$ 个部分和 (partial sum)。若部分和序列 $\{S_n\}$ 收敛于 $S$，则称级数 $\sum\limits_{n=1}^{\infty} u_n$ 收敛 (converge)，并称 $S$ 为该级数的和 (sum)，记作

$$\sum_{n=1}^{\infty} u_n = S.$$

此时称

$$r_n = S - S_n = \sum_{k=n+1}^{\infty} u_k$$

为该级数的余和 (tail)。此外，若 $\{S_n\}$ 发散，我们就称级数 $\sum\limits_{n=1}^{\infty} u_n$ 发散 (diverge)。

由上述定义可以看出，在级数中添加或删除有限多项（但不改变原有的项的顺序）不会改变其敛散性。

若部分和 $S_n$ 趋于 $+\infty$（相应地，$-\infty$），我们则称该级数发散于 $+\infty$（相应地，$-\infty$），此时为方便起见也称该级数的和为 $+\infty$（相应地，$-\infty$），并记作

$$\sum_{n=1}^{\infty} u_n = +\infty \qquad \text{（相应地，} \sum_{n=1}^{\infty} u_n = -\infty\text{）}.$$

**【例 1.3】**形如 $\sum\limits_{n=1}^{\infty} q^n$ 的级数被称为几何级数 (geometric series)，由于部分和

$$S_n = \sum_{k=1}^{n} q^k = \begin{cases} q \cdot \dfrac{1-q^n}{1-q}, & \text{若 } q \neq 1, \\ n, & \text{若 } q = 1. \end{cases}$$

因此级数 $\sum\limits_{n=1}^{\infty} q^n$ 当 $|q| < 1$ 时收敛，其和等于 $\dfrac{q}{1-q}$；当 $|q| \geqslant 1$ 时发散。

特别地，取 $q = \dfrac{1}{2}$ 知级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{2^n}$ 收敛，且其和为 1。取 $q = -1$ 知级数 $\sum\limits_{n=1}^{\infty} (-1)^n$ 发散。

乍一看，级数的敛散性本质上就是数列的敛散性，似乎没有必要多加研究。然而令人难以始料的是，级数的这种形式在很多具体问题中拥有极大的优越性，级数理论也在分析学中占有极重要的地位。

下面介绍收敛级数的一些基本性质。

**【命题 1.4】**设 $\sum\limits_{n=1}^{\infty} u_n$ 与 $\sum\limits_{n=1}^{\infty} v_n$ 均收敛，则对任意的 $\alpha, \beta \in \mathbb{R}$，$\sum\limits_{n=1}^{\infty} (\alpha u_n + \beta v_n)$ 收敛，且

$$\sum_{n=1}^{\infty} (\alpha u_n + \beta v_n) = \alpha \sum_{n=1}^{\infty} u_n + \beta \sum_{n=1}^{\infty} v_n.$$

证明. 以 $S_n$，$S_n'$ 和 $T_n$ 分别表示级数 $\sum\limits_{n=1}^{\infty} u_n$，$\sum\limits_{n=1}^{\infty} v_n$ 与 $\sum\limits_{n=1}^{\infty} (\alpha u_n + \beta v_n)$ 的部分和，则

$$T_n = \sum_{k=1}^{n} (\alpha u_k + \beta v_k) = \alpha \sum_{k=1}^{n} u_k + \beta \sum_{k=1}^{n} v_k = \alpha S_n + \beta S_n'.$$

故由 $\{S_n\}$ 和 $\{S'_n\}$ 的收敛性知 $\{T_n\}$ 收敛，且 $\lim\limits_{n\to\infty} T_n = \alpha \lim\limits_{n\to\infty} S_n + \beta \lim\limits_{n\to\infty} S'_n$。 □

**【命题 1.5】** 若 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\lim\limits_{n\to\infty} u_n = 0$。

证明. 以 $S_n$ 表示该级数的部分和，并记 $\lim\limits_{n\to\infty} S_n = S$，于是由 $u_n = S_n - S_{n-1}$ 知

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0.$$

□

上述命题的逆并不成立，例如级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{\sqrt{n}}$ 的通项趋于 0，但其部分和

$$S_n = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \geqslant n \cdot \frac{1}{\sqrt{n}} = \sqrt{n},$$

因此该级数发散。

**【命题 1.6】** 若 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则对其项任意加括号所得的级数

$$(u_1 + \cdots + u_{n_1}) + (u_{n_1+1} + \cdots + u_{n_2}) + \cdots \tag{9.2}$$

仍收敛，且和不变。换句话说，若 $\{n_k\}$ 是一个单调递增的正整数序列，并记

$$v_k = \sum_{j=n_{k-1}+1}^{n_k} u_j, \qquad \forall\, k \geqslant 1,$$

其中 $n_0 = 0$，则级数 $\sum\limits_{k=1}^{\infty} v_k$ 收敛，且

$$\sum_{n=1}^{\infty} u_n = \sum_{k=1}^{\infty} v_k. \tag{9.3}$$

证明. 以 $S_n$ 和 $T_k$ 分别表示 $\sum\limits_{n=1}^{\infty} u_n$ 和 $\sum\limits_{k=1}^{\infty} v_k$ 的部分和，则

$$T_k = \sum_{\ell=1}^{k} v_\ell = \sum_{\ell=1}^{k} \sum_{j=n_{\ell-1}+1}^{n_\ell} u_j = \sum_{j=1}^{n_k} u_j = S_{n_k}.$$

因此 $\{T_k\}$ 是 $\{S_n\}$ 的一个子列，从而由 $\{S_n\}$ 的收敛性知 $\{T_k\}$ 收敛，且有

$$\lim_{n\to\infty} S_n = \lim_{k\to\infty} T_k,$$

即 (9.3) 成立。 □

上述命题的逆不成立，例如对级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}$ 按如下方式添加括号

$$(1-1)+(1-1)+\cdots,$$

则添加括号后的级数收敛，但原级数发散。然而，如果同一个括号中的项均同号，那么上述命题的逆是成立的。

**【命题 1.7】设 (9.2) 式的同一个括号中的项均有相同的符号，那么当 (9.2) 收敛时原级数 $\sum\limits_{n=1}^{\infty} u_n$ 也收敛，且这两个级数的和相同。**

证明. 仍记

$$v_k = \sum_{j=n_{k-1}+1}^{n_k} u_j, \qquad \forall\, k \geqslant 1,$$

并用 $S_n$ 和 $T_k$ 分别表示 $\sum\limits_{n=1}^{\infty} u_n$ 和 $\sum\limits_{k=1}^{\infty} v_k$ 的部分和。注意到在这种情况下，对任意的 $k \geqslant 1$，当 $n \in [n_k, n_{k+1}]$ 时 $S_n$ 单调，从而有

$$T_k \leqslant S_n \leqslant T_{k+1} \qquad \text{或} \qquad T_{k+1} \leqslant S_n \leqslant T_k, \qquad \forall\, n \in [n_k, n_{k+1}].$$

因此若 $\sum\limits_{k=1}^{\infty} v_k$ 的和为 $T$，则对任意的 $\varepsilon > 0$，均存在 $k_\varepsilon \in \mathbb{Z}_{>0}$，使得当 $k > k_\varepsilon$ 时有 $T-\varepsilon < T_k < T+\varepsilon$。再结合上式便知，当 $n > n_{k_\varepsilon}$ 时就有

$$T-\varepsilon < S_n < T+\varepsilon.$$

从而 $\sum\limits_{n=1}^{\infty} u_n$ 收敛且和也等于 $T$。 □

**【定理 1.8】（Cauchy 收敛准则）** 级数 $\sum\limits_{n=1}^{\infty} u_n$ 收敛的充要条件是：对任意的 $\varepsilon > 0$，均存在 $n_\varepsilon \in \mathbb{Z}_{>0}$，使得对任意的 $m > n > n_\varepsilon$ 均有

$$|S_m - S_n| = \left|\sum_{k=n+1}^{m} u_k\right| < \varepsilon,$$

其中 $S_n$ 表示该级数的部分和。

证明. 这其实就是数列的 Cauchy 收敛准则。 □

**【例 1.9】**我们来研究级数 $\sum_{n=1}^{\infty}\frac{1}{n^s}$ $(s\in\mathbb{R})$ 的敛散性。分以下两种情况讨论：

(1) $s\leqslant 1$。此时，由于对任意的 $n\in\mathbb{Z}_{>0}$ 有

$$\Big|\sum_{k=n+1}^{2n}\frac{1}{k^s}\Big|\geqslant n\cdot\frac{1}{2n}=\frac{1}{2},$$

故而级数发散。

(2) $s>1$。对任意的正整数 $m>n>2$，必存在 $u,v\in\mathbb{Z}_{>0}$ 使得 $2^u<n\leqslant 2^{u+1}$ 以及 $m\leqslant 2^v$。于是

$$\begin{aligned}\Big|\sum_{k=n+1}^{m}\frac{1}{k^s}\Big|&\leqslant\sum_{k=2^u+1}^{2^v}\frac{1}{k^s}=\sum_{j=u}^{v-1}\sum_{k=2^j+1}^{2^{j+1}}\frac{1}{k^s}\\&\leqslant\sum_{j=u}^{v-1}2^j\cdot\frac{1}{(2^j)^s}=\sum_{j=u}^{v-1}\frac{1}{(2^{s-1})^j}=\frac{\dfrac{1}{(2^{s-1})^u}-\dfrac{1}{(2^{s-1})^v}}{1-\dfrac{1}{2^{s-1}}}\\&<\frac{2^{s-1}}{2^{s-1}-1}\cdot\frac{1}{(2^u)^{s-1}}\leqslant\frac{4^{s-1}}{2^{s-1}-1}\cdot\frac{1}{n^{s-1}}.\end{aligned}$$

故由 Cauchy 收敛准则知此时级数收敛。

当 $s>1$ 时，我们把级数 $\sum_{n=1}^{\infty}\frac{1}{n^s}$ 的和记作为 $\zeta(s)$，并称之为 Riemann $\zeta$ 函数 (Riemann zeta function)。

## 习 题 9.1

1. 如果级数 $\sum_{n=1}^{\infty}u_n$ 与 $\sum_{n=1}^{\infty}v_n$ 均发散，那么 $\sum_{n=1}^{\infty}(u_n+v_n)$ 是否发散?
2. 如果级数 $\sum_{n=1}^{\infty}u_n$ 与 $\sum_{n=1}^{\infty}v_n$ 中一个收敛而另一个发散，那么 $\sum_{n=1}^{\infty}(u_n+v_n)$ 的敛散性如何？
3. 如果 $\sum_{n=1}^{\infty}u_{2n-1}$ 和 $\sum_{n=1}^{\infty}u_{2n}$ 均收敛，那么 $\sum_{n=1}^{\infty}u_n$ 是否收敛？
4. 由 $\sum_{n=1}^{\infty}(u_{2n-1}+u_{2n})$ 收敛能否推出 $\sum_{n=1}^{\infty}u_n$ 收敛。
5. 设级数 $\sum_{n=1}^{\infty}u_n$ 满足 $\lim\limits_{n\to\infty}u_n=0$ 且 $\sum_{n=1}^{\infty}(u_{2n-1}+u_{2n})$ 收敛，证明 $\sum_{n=1}^{\infty}u_n$ 收敛？

6. 判断下列级数的敛散性：

(1) $\sum\limits_{n=1}^{\infty} \frac{1}{\sqrt[n]{\log n}}$； (2) $\sum\limits_{n=1}^{\infty} \frac{1}{n(n+1)}$；

(3) $\sum\limits_{n=1}^{\infty} \frac{n^3}{3^n-1}$； (4) $\sum\limits_{n=1}^{\infty} \sin n$；

(5) $1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots$。

7. 证明下列级数收敛并计算级数的和：

(1) $\sum\limits_{n=1}^{\infty} \left(\frac{1}{2^n}+\frac{(-1)^n}{3^n}\right)$； (2) $\sum\limits_{n=1}^{\infty} (\sqrt{n+2}-2\sqrt{n+1}+\sqrt{n})$；

(3) $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$； (4) $\sum\limits_{n=2}^{\infty} \frac{1}{n^2-1}$；

(5) $\sum\limits_{n=1}^{\infty} \frac{1}{(2n-1)2n(2n+1)}$； (6) $\sum\limits_{n=1}^{\infty} \arctan\frac{1}{n^2}$。

8. 设 $u_1=2$，$u_{n+1}=u_n^2-u_n+1$ $(\forall\ n\geqslant 1)$，证明级数 $\sum\limits_{n=1}^{\infty}\frac{1}{u_n}$ 收敛并求其和。

9. 把级数 $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ 中分母含有数字 9 的项全部删去，证明所得新级数收敛，且其和不超过 80。

10. 设 $A$ 是正整数集的一个无限子集。对任意的实数 $x\geqslant 1$，以 $f(x)$ 表示集合 $A\cap[1,x]$ 中的元素个数。如果存在正常数 $c>0$，使得 $\lim\limits_{x\to+\infty}\frac{f(x)}{x}=c$，证明级数 $\sum\limits_{a\in A}\frac{1}{a}$ 发散，这里 $\sum\limits_{a\in A}$ 表示将 $A$ 中元素从小到大排列后进行求和。

11. 证明对任意的 $x>1$ 有

$$\sum_{n=0}^{\infty}\frac{1}{x+2^n}=\frac{\log x}{x\log 2}+O\Big(\frac{1}{x}\Big).$$

12. 把全体有理数排成一列，记作 $r_1, r_2, \cdots, r_n, \cdots$，并令

$$f(x)=\sum_{\substack{n\\ r_n<x}}\frac{1}{2^n},\qquad \forall\ x\in\mathbb{R}.$$

证明 $f(x)$ 是在无理点连续、在有理点间断的严格单调递增函数[1]。

[1] 可以证明，单调函数至多有可数个间断点，因此本题给出了一个极端的例子。

## §9.2 正项级数

在上节中我们给出了判定级数敛散性的 Cauchy 收敛准则，但同时也通过看到了这一判定法则在实际应用中是颇不方便的，因此我们需针对一些常见的情况给出较简易的判定方法。在本节中我们来研究一种最简单的情形，即通项保持一定符号的级数。

**【定义 2.1】** 若 $u_n \geqslant 0\ (\forall\ n \geqslant 1)$，则称 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数。

设 $a$ 是一个正整数，我们首先考虑由区间 $[a,+\infty)$ 上的非负连续函数 $f(x)$ 所给出的级数 $\sum\limits_{n=a}^{\infty} f(n)$。

**【定理 2.2】（Cauchy 积分判别法）** 设 $a$ 是一个正整数，$f(x)$ 是 $[a,+\infty)$ 上的单调递减的非负连续函数，则级数 $\sum\limits_{n=a}^{\infty} f(n)$ 与数列 $\{A_n\}_{n\geqslant a}$ 同敛散，其中

$$A_n = \int_a^n f(x)\,\mathrm{d}x, \qquad \forall\ n \in \mathbb{Z}_{\geqslant a}.$$

证明. 由 $f(x)$ 的单调性及非负性知，对任意的 $k \in \mathbb{Z}_{>a}$ 有

$$f(k) \leqslant \int_{k-1}^{k} f(x)\,\mathrm{d}x \leqslant f(k-1). \tag{9.4}$$

于是对 $k$ 求和便得

$$\sum_{k=a+1}^{n} f(k) \leqslant \int_a^n f(x)\,\mathrm{d}x \leqslant \sum_{k=a+1}^{n} f(k-1) = \sum_{k=a}^{n-1} f(k),$$

若记 $S_n = \sum\limits_{k=a}^{n} f(k)$，那么上式也即

$$S_n - f(a) \leqslant A_n \leqslant S_{n-1}, \qquad \forall\ n \in \mathbb{Z}_{>a}. \tag{9.5}$$

注意到 $\{S_n\}$ 与 $\{A_n\}$ 均是单调递增的，故若 $\{S_n\}$ 收敛，则其必有上界，从而由 (9.5) 中第二个不等式知 $\{A_n\}$ 也有上界，再利用单调有界收敛原理便知 $\{A_n\}$ 收敛。同理可由 $\{A_n\}$ 的收敛性得到 $\{S_n\}$ 收敛。 □

【例 2.3】我们先回到上节最后一个例子。对级数 $\sum_{n=1}^{\infty}\frac{1}{n^s}$ 而言，由于

$$\int_1^n \frac{\mathrm{d}x}{x^s}=\begin{cases}\log n, & 若\ s=1,\\ \dfrac{n^{1-s}-1}{1-s}, & 若\ s\neq 1,\end{cases}$$

故该级数当 $s\leqslant 1$ 时发散，当 $s>1$ 时收敛。

【例 2.4】再来考察级数 $\sum_{n=2}^{\infty}\frac{1}{n(\log n)^{\alpha}}$，由于

$$\int_2^n \frac{\mathrm{d}x}{x(\log x)^{\alpha}}=\begin{cases}\log\log n-\log\log 2, & 若\ \alpha=1,\\ \dfrac{(\log n)^{1-\alpha}-(\log 2)^{1-\alpha}}{1-\alpha}, & 若\ \alpha\neq 1,\end{cases}$$

故该级数当 $\alpha\leqslant 1$ 时发散，当 $\alpha>1$ 时收敛。

从 (9.4) 可以看出，定理 2.2 其实是在把所需要判定的级数与以 $\int_{n-1}^{n} f(x)\,\mathrm{d}x$ 为通项的级数进行比较，而证明的关键在于数列的单调有界收敛原理，由这一原理我们知道正项级数收敛当且仅当其部分和有界，在此基础上我们可以建立如下更一般的比较判别法 (comparison test)。

**【命题 2.5】(比较判别法)** 设 $\sum_{n=1}^{\infty}u_n$ 和 $\sum_{n=1}^{\infty}v_n$ 是两个正项级数，且存在正常数 $c$ 使得自某项开始有 $u_n\leqslant cv_n$。那么，

(1) 若 $\sum_{n=1}^{\infty}v_n$ 收敛，则 $\sum_{n=1}^{\infty}u_n$ 收敛；

(2) 若 $\sum_{n=1}^{\infty}u_n$ 发散，则 $\sum_{n=1}^{\infty}v_n$ 发散。

证明. 因为删去有限多项不会改变级数的敛散性，故不妨设对任意的正整数 $n$ 均有 $u_n\leqslant cv_n$。现以 $S_n$ 和 $T_n$ 分别表示 $\sum_{n=1}^{\infty}u_n$ 和 $\sum_{n=1}^{\infty}v_n$ 的部分和，于是有

$$0\leqslant S_n\leqslant cT_n,\qquad \forall\, n\in\mathbb{Z}_{>0}.$$

若 $\sum_{n=1}^{\infty}v_n$ 收敛，则 $\{T_n\}$ 有上界，从而 $\{S_n\}$ 也有上界，进而 $\sum_{n=1}^{\infty}u_n$ 收敛，这样就证明了 (1)。(2) 可由 (1) 得到。 □

比较判别法的下述极限形式在一些实际应用中更为方便。

**【命题 2.6】**（比较判别法的极限形式） 设 $\sum\limits_{n=1}^{\infty} u_n$ 和 $\sum\limits_{n=1}^{\infty} v_n$ 是两个正项级数，若

$$\lim_{n\to\infty} \frac{u_n}{v_n} \in (0,+\infty),$$

则这两个级数同时收敛或同时发散。

<u>证明</u>. 记 $\lim\limits_{n\to\infty} \frac{u_n}{v_n} = \ell$，则存在 $n_0 \in \mathbb{Z}_{>0}$，使得当 $n > n_0$ 时有

$$\left|\frac{u_n}{v_n} - \ell\right| < \frac{\ell}{2},$$

也即

$$\frac{\ell}{2} v_n < u_n < \frac{3\ell}{2} v_n, \qquad \forall\ n > n_0.$$

于是由命题 2.5 知结论成立。 □

下述命题也是比较判别法的一个变形。

**【命题 2.7】**设 $\sum\limits_{n=1}^{\infty} u_n$ 和 $\sum\limits_{n=1}^{\infty} v_n$ 是两个正项级数，且存在正整数 $n_0$ 使得对任意的 $n \geqslant n_0$ 有

$$\frac{u_{n+1}}{u_n} \leqslant \frac{v_{n+1}}{v_n}.$$

那么，

(1) 若 $\sum\limits_{n=1}^{\infty} v_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} u_n$ 收敛；

(2) 若 $\sum\limits_{n=1}^{\infty} u_n$ 发散，则 $\sum\limits_{n=1}^{\infty} v_n$ 发散。

<u>证明</u>. 不妨设 $n_0 = 1$，于是对任意的 $k \geqslant 1$ 有

$$\frac{u_{k+1}}{u_k} \leqslant \frac{v_{k+1}}{v_k}.$$

现对变量 $k$ 求乘积，则可得到

$$\prod_{k=1}^{n-1} \frac{u_{k+1}}{u_k} \leqslant \prod_{k=1}^{n-1} \frac{v_{k+1}}{v_k},$$

也即 $\frac{u_n}{u_1} \leqslant \frac{v_n}{v_1}$，于是由命题 2.5 知结论成立。 □

有了上述几个判别法，我们就可以把需判定的级数与一些已知级数相比较，从而得到其敛散性。我们首先来看将需判定的级数与几何级数（参见例 1.3）相比较的情形，这就给出了下面的 Cauchy 判别法和 d'Alembert 判别法。

**【定理 2.8】(Cauchy 判别法)** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数。若自某项起有

$$\sqrt[n]{u_n} \leqslant q < 1,$$

其中 $q$ 是一个正常数，则该级数收敛；若自某项起有 $\sqrt[n]{u_n} \geqslant 1$，则该级数发散。

<u>证明</u>. 注意到 $\sqrt[n]{u_n} \leqslant q$ 与 $\sqrt[n]{u_n} \geqslant 1$ 分别等价于

$$u_n \leqslant q^n \qquad \text{及} \qquad u_n \geqslant 1,$$

故由命题 2.5 知结论成立。 □

该判别法也有一个便于应用的极限形式。

**【定理 2.9】(Cauchy 判别法的极限形式)** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数，并记

$$r = \overline{\lim_{n\to\infty}} \sqrt[n]{u_n}.$$

那么，该级数当 $r < 1$ 时收敛，当 $r > 1$ 时发散。

<u>证明</u>. 若 $r < 1$，则存在 $\varepsilon_0 > 0$ 使得 $r + \varepsilon_0 < 1$。因此由第三章定理 6.4 知存在正整数 $n_0$，使得当 $n > n_0$ 时有

$$\sqrt[n]{u_n} < r + \varepsilon_0 < 1,$$

于是由定理 2.8 可得级数 $\sum\limits_{n=1}^{\infty} u_n$ 收敛。

若 $r > 1$，则由第三章定理 6.4 知存在 $n_k$ $(k = 1, 2, \cdots)$ 使得

$$(u_{n_k})^{\frac{1}{n_k}} > 1, \qquad \forall\ k \in \mathbb{Z}_{>0},$$

也即 $u_{n_k} > 1$ $(\forall\ k \in \mathbb{Z}_{>0})$，这说明级数 $\sum\limits_{n=1}^{\infty} u_n$ 的通项不趋于 0，从而发散。 □

应用命题 2.7 来把所需要判定的级数与几何级数比较，可得如下 d'Alembert 判别法。

**【定理 2.10】(d'Alembert 判别法)** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数。若自某项起有

$$\frac{u_{n+1}}{u_n} \leqslant q < 1,$$

其中 $q$ 是一个正常数，则该级数收敛；若自某项起有 $\dfrac{u_{n+1}}{u_n} \geqslant 1$，则该级数发散。

D'Alembert 判别法也有下述便于应用的极限形式，其证明与定理 2.9 的证明类似，我们把它留作练习。

**【定理 2.11】(d'Alembert 判别法的极限形式)** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数，并记

$$\overline{r} = \varlimsup_{n\to\infty} \frac{u_{n+1}}{u_n}, \qquad \underline{r} = \varliminf_{n\to\infty} \frac{u_{n+1}}{u_n}.$$

那么，该级数当 $\overline{r} < 1$ 时收敛，当 $\underline{r} > 1$ 时发散。

下面我们来看看 Cauchy 判别法与 d'Alembert 判别法孰优孰劣。由 §3.6 习题 7 知

$$\varliminf_{n\to\infty} \frac{u_{n+1}}{u_n} \leqslant \varliminf_{n\to\infty} \sqrt[n]{u_n} \leqslant \varlimsup_{n\to\infty} \sqrt[n]{u_n} \leqslant \varlimsup_{n\to\infty} \frac{u_{n+1}}{u_n}, \tag{9.6}$$

因此对于那些能用 d'Alembert 判别法判断敛散性的级数，用 Cauchy 判别法也一定能判断其敛散性，但反之不然。例如，对于以

$$u_n = \left(\frac{2+(-1)^n}{4}\right)^n$$

为通项的级数而言，因为

$$\varlimsup_{n\to\infty} \sqrt[n]{u_n} = \varlimsup_{n\to\infty} \frac{2+(-1)^n}{4} = \frac{3}{4} < 1,$$

故由 Cauchy 判别法知其收敛；但是，

$$\frac{u_{n+1}}{u_n} = \frac{1}{4} \cdot \frac{(2+(-1)^{n+1})^{n+1}}{(2+(-1)^n)^n} = \begin{cases} \dfrac{1}{4} \cdot \dfrac{1}{3^{2k}}, & \text{若 } n = 2k, \\ \dfrac{1}{4} \cdot 3^{2k}, & \text{若 } n = 2k-1, \end{cases}$$

因此

$$\varlimsup_{n\to\infty} \frac{u_{n+1}}{u_n} = +\infty, \qquad \varliminf_{n\to\infty} \frac{u_{n+1}}{u_n} = 0,$$

从而由 d'Alembert 判别法无法判断 $\sum\limits_{n=1}^{\infty} u_n$ 的敛散性。上述讨论说明 Cauchy 判别法的适用范围要比 d'Alembert 判别法更广，但在一些情况下，d'Alembert 判别法会更方便些。下面我们来看几个例子。

**【例 2.12】** 设 $a > 0$。对级数 $\sum\limits_{n=1}^{\infty} \left(\dfrac{a}{n}\right)^n$ 而言，由于

$$\lim_{n\to\infty} \sqrt[n]{\left(\frac{a}{n}\right)^n} = \lim_{n\to\infty} \frac{a}{n} = 0 < 1,$$

故由 Cauchy 判别法的知级数收敛。

**【例 2.13】**对级数 $\sum\limits_{n=1}^{\infty}\dfrac{n!}{n^n}$ 而言，若要用 Cauchy 判别法，需计算极限

$$\varlimsup_{n\to\infty}\frac{\sqrt[n]{n!}}{n},$$

这并不容易。另一方面，由于

$$\lim_{n\to\infty}\frac{(n+1)!/(n+1)^{n+1}}{n!/n^n}=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{-n}=\frac{1}{e}<1,\ ②$$

所以由 d'Alembert 判别法知级数收敛。

如果级数 $\sum\limits_{n=1}^{\infty}u_n$ 能用 Cauchy 判别法或 d'Alembert 判别法确定其收敛，则必存在 $q\in(0,1)$ 使得对充分大的 $n$ 有

$$\sqrt[n]{u_n}\leqslant q \qquad 或 \qquad \frac{u_{n+1}}{u_n}\leqslant q,$$

无论是哪种情况，只需取 $r=\dfrac{q+1}{2}$ 就有 $u_n=o(r^n)$，因此 $\sum\limits_{n=1}^{\infty}u_n$ 比几何级数收敛得快③，而这正是 Cauchy 判别法和 d'Alembert 判别法的适用范围。换句话说，对于那些比几何级数收敛得慢的正项级数，这两个判别法均是失效的。所以为了获得更有效的判别法，一个自然的想法是寻找一些比几何级数收敛得慢的级数，将它们作为"标准"去与需判定的级数作比较，而我们首先想到的就是形如 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^s}$ $(s>1)$ 的级数④，这就引出了下面的 Raabe 判别法。

**【定理 2.14】（Raabe 判别法）** 设 $\sum\limits_{n=1}^{\infty}u_n$ **是正项级数。若自某项起有**

$$n\left(\frac{u_n}{u_{n+1}}-1\right)\geqslant q>1,$$

**其中 $q$ 是一个正常数，则该级数收敛；若自某项起有**

$$n\left(\frac{u_n}{u_{n+1}}-1\right)\leqslant 1,$$

---

②由此及 (9.6) 可顺便得知 $\lim\limits_{n\to\infty}\dfrac{\sqrt[n]{n!}}{n}=\dfrac{1}{e}$。

③对两个收敛的正项级数 $\sum\limits_{n=1}^{\infty}u_n$ 和 $\sum\limits_{n=1}^{\infty}v_n$ 而言，若 $u_n=o(v_n)$（当 $n\to\infty$ 时），则称 $\sum\limits_{n=1}^{\infty}u_n$ 比 $\sum\limits_{n=1}^{\infty}v_n$ 收敛得快。

④由第三章例 2.10 知这个级数比几何级数收敛得慢。

则该级数发散。

证明. $n\left(\dfrac{u_n}{u_{n+1}}-1\right)\geqslant q$ 意味着 $\dfrac{u_{n+1}}{u_n}\leqslant\left(1+\dfrac{q}{n}\right)^{-1}$，因此由 Taylor 公式知对充分大的 $n$ 有

$$\frac{u_{n+1}}{u_n}\leqslant 1-\frac{q}{n}+O\Big(\frac{1}{n^2}\Big).$$

另一方面，对任意的 $s>1$ 有

$$\frac{\dfrac{1}{(n+1)^s}}{\dfrac{1}{n^s}}=\Big(1+\frac{1}{n}\Big)^{-s}=1-\frac{s}{n}+O\Big(\frac{1}{n^2}\Big).$$

当 $q>1$ 时，取 $s=\dfrac{q+1}{2}$ 便有 $1<s<q$，从而可对这一 $s$ 及充分大的 $n$ 得到

$$\frac{u_{n+1}}{u_n}<\frac{\dfrac{1}{(n+1)^s}}{\dfrac{1}{n^s}}.$$

因为 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^s}$ 收敛，故由命题 2.7 知 $\sum\limits_{n=1}^{\infty}u_n$ 收敛。

此外，若自某项起有 $n\left(\dfrac{u_n}{u_{n+1}}-1\right)\leqslant 1$，则对充分大的 $n$ 有

$$\frac{u_{n+1}}{u_n}\geqslant\frac{n}{n+1}=\frac{\dfrac{1}{n+1}}{\dfrac{1}{n}},$$

于是由 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n}$ 发散及命题 2.7 知此时 $\sum\limits_{n=1}^{\infty}u_n$ 发散。 □

下面是 Raabe 判别法的一个便于应用的极限形式，我们也将其证明留作练习。

**【定理 2.15】（Raabe 判别法的极限形式）** 设 $\sum\limits_{n=1}^{\infty}u_n$ 是正项级数，并记

$$r=\lim_{n\to\infty}n\left(\frac{u_n}{u_{n+1}}-1\right).$$

则级数当 $r>1$ 时收敛，当 $r<1$ 时发散。

**【例 2.16】**对级数 $\sum_{n=1}^{\infty}\frac{(2n-1)!!}{(2n)!!}\cdot\frac{1}{2n+1}$ 而言，若将其通项记作 $u_n$，则有

$$\frac{u_{n+1}}{u_n}=\frac{(2n+1)^2}{(2n+2)(2n+3)},$$

结合 (9.6) 知此时

$$\lim_{n\to\infty}\sqrt[n]{u_n}=\lim_{n\to\infty}\frac{u_{n+1}}{u_n}=1.$$

因而 Cauchy 判别法和 d'Alembert 判别法均失效。但注意到

$$\lim_{n\to\infty}n\left(\frac{u_n}{u_{n+1}}-1\right)=\lim_{n\to\infty}\frac{n(6n+5)}{(2n+1)^2}=\frac{3}{2}>1,$$

故由 Raabe 判别法知级数收敛。

**【例 2.17】**我们来考察级数 $\sum_{n=1}^{\infty}\left|\binom{\alpha}{n}\right|$，其中

$$\binom{\alpha}{n}=\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$$

如第六章例 2.5 中所定义。当 $\alpha\in\mathbb{Z}_{\geqslant 0}$ 时该级数的通项自某项开始为 0，此时级数当然收敛。现设 $\alpha\notin\mathbb{Z}_{\geqslant 0}$，并记其通项为 $u_n$，则有

$$\frac{u_{n+1}}{u_n}=\frac{|\alpha-n|}{n+1}.$$

因为当 $n\to\infty$ 时上式趋于 1，所以由 Cauchy 判别法和 d'Alembert 判别法均无法判断其敛散性。但是

$$\lim_{n\to\infty}n\left(\frac{u_n}{u_{n+1}}-1\right)=\lim_{n\to\infty}n\left(\frac{n+1}{n-\alpha}-1\right)=\lim_{n\to\infty}\frac{n(1+\alpha)}{n-\alpha}=1+\alpha,$$

故由 Raabe 判别法知级数当 $\alpha>0$ 时收敛，当 $\alpha<0$ 时发散。

当定理 2.15 中的 $r$ 等于 1 时，由 Raabe 判别法便无法判定级数的敛散性，因此还需寻找更有效的判别法。因为 Raabe 判别法是把所需要判断的级数与 $\sum\limits_{n=1}^{\infty}\frac{1}{n^s}$ 作比较，故为了得到更有效的判别法，需要去寻找比 $\sum\limits_{n=1}^{\infty}\frac{1}{n^s}$ $(s>1)$ 收敛得更慢的级数并将其作为“标准”来使用比较判别法。这样的级数很多，但 $\sum\limits_{n=1}^{\infty}\frac{1}{n(\log n)^{\alpha}}$ $(\alpha>1)$ 是一个颇为自然的选择，这就引出了如下的 Gauss 判别法。

**【定理 2.18】（Gauss 判别法）** 设 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数，且当 $n \to \infty$ 时有

$$\frac{u_n}{u_{n+1}} = 1 + \frac{1}{n} + \frac{\beta}{n\log n} + o\Big(\frac{1}{n\log n}\Big).$$

那么，该级数当 $\beta > 1$ 时收敛，当 $\beta < 1$ 时发散。

<u>证明</u>. 对任意的 $\alpha \in \mathbb{R}$ 有

$$\begin{aligned}
\frac{\dfrac{1}{n(\log n)^\alpha}}{\dfrac{1}{(n+1)(\log(n+1))^\alpha}} &= \left(1+\frac{1}{n}\right)\left(\frac{\log(n+1)}{\log n}\right)^\alpha \\
&= \left(1+\frac{1}{n}\right)\left[1+\frac{1}{\log n}\log\left(1+\frac{1}{n}\right)\right]^\alpha \\
&= \left(1+\frac{1}{n}\right)\left[1+\frac{1}{\log n}\left(\frac{1}{n}+O\Big(\frac{1}{n^2}\Big)\right)\right]^\alpha \\
&= \left(1+\frac{1}{n}\right)\left[1+\frac{1}{n\log n}+O\Big(\frac{1}{n^2\log n}\Big)\right]^\alpha \\
&= \left(1+\frac{1}{n}\right)\left[1+\frac{\alpha}{n\log n}+O\Big(\frac{1}{n^2\log n}\Big)\right] \\
&= 1+\frac{1}{n}+\frac{\alpha}{n\log n}+O\Big(\frac{1}{n^2\log n}\Big).
\end{aligned}$$

若 $\beta > 1$，则当取 $\alpha = \dfrac{\beta+1}{2}$ 时就有 $1 < \alpha < \beta$，从而对充分大的 $n$ 有

$$\frac{u_n}{u_{n+1}} > \frac{\dfrac{1}{n(\log n)^\alpha}}{\dfrac{1}{(n+1)(\log(n+1))^\alpha}},$$

也即

$$\frac{u_{n+1}}{u_n} < \frac{\dfrac{1}{(n+1)(\log(n+1))^\alpha}}{\dfrac{1}{n(\log n)^\alpha}}.$$

注意到此时级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{n(\log n)^\alpha}$ 收敛，故由命题 2.7 知 $\sum\limits_{n=1}^{\infty} u_n$ 收敛。

另一方面，若 $\beta < 1$，则对充分大的 $n$ 有

$$\frac{u_{n+1}}{u_n} > \frac{\dfrac{1}{(n+1)\log(n+1)}}{\dfrac{1}{n\log n}}.$$

于是由 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n\log n}$ 发散及命题 2.7 知 $\sum\limits_{n=1}^{\infty}u_n$ 发散。 □

**【例 2.19】（Gauss）** 考虑超几何级数 (hypergeometric series)[5]

$$F(\alpha,\beta,\gamma,x) = 1 + \sum_{n=1}^{\infty}\frac{\alpha(\alpha+1)\cdots(\alpha+n-1)\cdot\beta(\beta+1)\cdots(\beta+n-1)}{n!\gamma(\gamma+1)\cdots(\gamma+n-1)}x^n,$$

其中 $\alpha, \beta, \gamma, x$ 均是正实数。若将其通项记作 $u_n$，则当 $n \geqslant 1$ 时有

$$\begin{aligned}\frac{u_n}{u_{n+1}} &= \frac{(n+1)(\gamma+n)}{(\alpha+n)(\beta+n)}\cdot\frac{1}{x} = \frac{\left(1+\dfrac{1}{n}\right)\left(1+\dfrac{\gamma}{n}\right)}{\left(1+\dfrac{\alpha}{n}\right)\left(1+\dfrac{\beta}{n}\right)}\cdot\frac{1}{x}\\ &= \frac{1}{x}\left(1+\frac{1+\gamma-\alpha-\beta}{n}+O\left(\frac{1}{n^2}\right)\right).\end{aligned}$$

现分以下几种情况来讨论:

(1) 若 $x \neq 1$，则由 d'Alembert 判别法知级数当 $0 < x < 1$ 时收敛，当 $x > 1$ 时发散。

(2) 若 $x = 1$ 且 $\alpha+\beta \neq \gamma$，则由 Raabe 判别法知级数当 $x = 1$ 且 $\alpha+\beta < \gamma$ 时收敛，当 $x = 1$ 且 $\alpha+\beta > \gamma$ 时发散。

(3) 若 $x = 1$ 且 $\alpha+\beta = \gamma$，由 Gauss 判别法知级数发散。

若在定理 2.18 中 $\beta = 1$，则由 Gauss 判别法亦无法判断级数的敛散性，因此很自然地会想到用比 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n(\log n)^{\alpha}}$ $(\alpha > 1)$ 收敛得更慢的级数作为“标准”来使用比较判别法。或者甚至有人会想到，是否能找一个收敛得“最慢”的级数来作为比较的“标准”，从而一劳永逸地解决正项级数的敛散性问题。然而遗憾的是，所谓收敛得“最慢”的级数是不存在的，关于这一点，可参见习题 15。

在本节结束之前我们再举一个例子，它说明在一些情况下通过研究级数通项的渐近性态从而获得其敛散性是一个更为直接有效的做法。

---

[5] 事实上，这只是一种特殊的超几何级数。所谓超几何级数，是指 $\dfrac{u_{n+1}}{u_n}$ 为关于 $n$ 的有理函数的级数 $\sum\limits_{n=1}^{\infty}u_n$，这是几何级数的一个自然推广。

**【例 2.20】** 设 $u_n=\dfrac{1}{n^s}\left(1-\dfrac{x\log n}{n}\right)^n$，试讨论级数 $\sum\limits_{n=1}^{\infty}u_n$ 的敛散性。

解. 我们有

$$\begin{aligned}
u_n&=\frac{1}{n^s}\exp\left[n\log\left(1-\frac{x\log n}{n}\right)\right]=\frac{1}{n^s}\exp\left[-n\left(\frac{x\log n}{n}+O\left(\frac{x^2(\log n)^2}{n^2}\right)\right)\right]\\
&=\frac{1}{n^s}\exp\left[-x\log n+O\left(\frac{x^2(\log n)^2}{n}\right)\right]=\frac{1}{n^{s+x}}\exp\left[O\left(\frac{x^2(\log n)^2}{n}\right)\right]\\
&=\frac{1}{n^{s+x}}\left[1+O\left(\frac{x^2(\log n)^2}{n}\right)\right].
\end{aligned}$$

因此级数当 $s+x>1$ 时收敛，当 $s+x\leqslant 1$ 时发散。 □

## 习 题 9.2

1. 证明定理 2.11 和 2.15。
2. 判断下列级数的敛散性：

(1) $\displaystyle\sum_{n=2}^{\infty}\frac{\log^{\beta}n}{n^{\alpha}}$； (2) $\displaystyle\sum_{n=10}^{\infty}\frac{1}{n(\log n)(\log\log n)^s}$；

(3) $\displaystyle\sum_{n=1}^{\infty}\frac{n!}{n^{n-1}}$； (4) $\displaystyle\sum_{n=1}^{\infty}\frac{2019^n}{n!}$；

(5) $\displaystyle\sum_{n=2}^{\infty}\frac{n^{\log n}}{\log^n n}$； (6) $\displaystyle\sum_{n=1}^{\infty}\frac{\sqrt[3]{n+1}-\sqrt[3]{n-1}}{n^{\alpha}}$；

(7) $\displaystyle\sum_{n=1}^{\infty}\frac{n^{n+\frac{1}{n}}}{\left(n+\frac{1}{n}\right)^n}$； (8) $\displaystyle\sum_{n=1}^{\infty}\frac{(2n)!}{2^{2n}(n!)^2(2n+1)}x^{2n}$；

(9) $\displaystyle\sum_{n=1}^{\infty}\frac{n!e^n}{n^{n+p}}$； (10) $\displaystyle\sum_{n=1}^{\infty}\frac{a(a+1)\cdots(a+n-1)}{n!}\cdot\frac{1}{n^b}\quad(a,b>0)$。

3. 已知正项级数 $\sum\limits_{n=1}^{\infty}u_n$ 与 $\sum\limits_{n=1}^{\infty}v_n$ 均发散，试讨论下列级数的敛散性：

(1) $\displaystyle\sum_{n=1}^{\infty}\max(u_n,\,v_n)$； (2) $\displaystyle\sum_{n=1}^{\infty}\min(u_n,\,v_n)$。

4. 设 $\sum\limits_{n=1}^{\infty}u_n$ 是一个收敛的正项级数，证明级数 $\sum\limits_{n=1}^{\infty}\dfrac{u_n}{1+u_n}$ 也收敛。

5. 设正项级数 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，证明 $\sum\limits_{n=1}^{\infty} u_n^2$ 也收敛。这一命题的逆命题成立吗?

6. 设级数 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} c_n$ 均收敛，且对任意的正整数 $n$ 有 $a_n \leqslant b_n \leqslant c_n$，证明级数 $\sum\limits_{n=1}^{\infty} b_n$ 也收敛。

7. 设正项级数 $\sum\limits_{n=1}^{\infty} u_n$ 的通项单调递减，且 $\sum\limits_{n=1}^{\infty} u_{2n}$ 收敛，证明 $\sum\limits_{n=1}^{\infty} u_n$ 收敛。

8. 设正项级数 $\sum\limits_{n=1}^{\infty} u_n$ 收敛且 $\{u_n\}$ 单调递减，证明 $u_n = o\Big(\dfrac{1}{n}\Big)$ 。

9. 试构造一个满足 $u_n \neq o\Big(\dfrac{1}{n}\Big)$ 的收敛的正项级数 $\sum\limits_{n=1}^{\infty} u_n$。

10. 设 $\sum\limits_{n=1}^{\infty} u_n$，$\sum\limits_{n=1}^{\infty} v_n$ 是两个正项级数，$\sum\limits_{n=1}^{\infty} u_n$ 收敛且 $\sum\limits_{n=1}^{\infty} v_n$ 发散，问是否一定有 $u_n = O(v_n)$?

11. 设 $\{u_n\}$ 是单调递减的正数列，证明级数 $\sum\limits_{n=1}^{\infty} u_n$ 与数列 $\left\{\dfrac{1}{u_{n+1}} - \dfrac{1}{u_n}\right\}$ 不可能同时收敛。

12. 设正项级数 $\sum\limits_{n=1}^{\infty} u_n$ 满足

$$\lim_{n\to\infty} n \log \frac{u_n}{u_{n+1}} = r.$$

证明该当 $r > 1$ 时 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，当 $r < 1$ 时 $\sum\limits_{n=1}^{\infty} u_n$ 发散。

13. （Kummer 判别法）设 $u_n, v_n > 0$ $(\forall\, n \in \mathbb{Z}_{>0})$。那么

(1) 若存在正实数 $\kappa$，使得

$$\frac{v_n}{v_{n+1}} u_n - u_{n+1} \geqslant \kappa, \qquad \forall\, n \in \mathbb{Z}_{>0},$$

则级数 $\sum\limits_{n=1}^{\infty} v_n$ 收敛;

(2) 若级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{u_n}$ 发散，且

$$\frac{v_n}{v_{n+1}} u_n - u_{n+1} \leqslant 0, \qquad \forall\, n \in \mathbb{Z}_{>0},$$

则级数 $\sum\limits_{n=1}^{\infty} v_n$ 发散。

14. （Лобачевский 判别法）设正项级数 $\sum\limits_{n=1}^{\infty} u_n$ 的通项单调递减趋于 0，并记

$$p_m = \max\left\{n \in \mathbb{Z}_{>0} : u_n \geqslant \frac{1}{2^m}\right\}, \qquad \forall\, m \in \mathbb{Z}_{>0}.$$

证明：级数 $\sum\limits_{n=1}^{\infty} u_n$ 与 $\sum\limits_{m=1}^{\infty} \dfrac{p_m}{2^m}$ 具有相同的敛散性。

15. 设 $\sum\limits_{n=1}^{\infty} u_n$ 是一个收敛的正项级数，记 $r_n = \sum\limits_{k=n+1}^{\infty} u_k$ 为它的第 $n$ 个余和，并记

$$v_n = \sqrt{r_{n-1}} - \sqrt{r_n}, \qquad \forall\, n \in \mathbb{Z}_{>0}.$$

证明级数 $\sum\limits_{n=1}^{\infty} v_n$ 也是一个收敛的正项级数，且有 $u_n = o(v_n)$。

16. 设 $\sum\limits_{n=1}^{\infty} u_n$ 是一个发散的正项级数，记 $S_n = \sum\limits_{k=1}^{n} u_k$ 为其部分和，并记

$$v_n = \frac{u_n}{S_n}, \qquad \forall\, n \in \mathbb{Z}_{>0}.$$

证明级数 $\sum\limits_{n=1}^{\infty} v_n$ 也是一个发散的正项级数，且有 $v_n = o(u_n)$。

17. 设 $\gamma - \beta > 1$，证明 $F(1, \beta, \gamma, 1) = \dfrac{\gamma - 1}{\gamma - \beta - 1}$，其中 $F$ 如例 2.19 所定义。

## § 9.3
## 任意项级数

在上节中我们讨论了正项级数的敛散性问题，接下来的两节将转而研究通项可正可负的那些级数。在本节中我们将对形如 $\sum\limits_{n=1}^{\infty} a_n b_n$ 的级数给出两个判定收敛的方法，即 Abel 判别法和 Dirichlet 判别法，而它们的基础是以挪威数学家 N. H. Abel 的名字命名的求和公式。

**【定理 3.1】（Abel 求和公式）** 设 $(a_n)_{n\in\mathbb{Z}}$ 和 $(b_n)_{n\in\mathbb{Z}}$ 是复数集 $\mathbb{C}$ 的两个元素族，则对任意的 $M \in \mathbb{Z}$ 以及 $N \in \mathbb{Z}_{>0}$，有

$$\sum_{M<n\leqslant M+N} a_n b_n = a_{M+N} B_{M+N} + \sum_{M<n\leqslant M+N-1} (a_n - a_{n+1}) B_n, \tag{9.7}$$

其中 $B_n = \sum\limits_{M<k\leqslant n} b_k$。特别地，若 $(a_n)_{n\in\mathbb{Z}}$ 是 $\mathbb{R}$ 的一个单调的元素族且

$$\sup_{M<n\leqslant M+N} |B_n| \leqslant \rho,$$

则

$$\Big|\sum_{M<n\leqslant M+N} a_n b_n\Big| \leqslant \rho\big(|a_{M+1}|+2|a_{M+N}|\big). \tag{9.8}$$

证明. 由 $b_n = B_n - B_{n-1}$ 知

$$\begin{aligned}\sum_{M<n\leqslant M+N} a_n b_n &= \sum_{M<n\leqslant M+N} a_n(B_n - B_{n-1}) \\ &= \sum_{M<n\leqslant M+N} a_n B_n - \sum_{M-1<n\leqslant M+N-1} a_{n+1}B_n \\ &= a_{M+N}B_{M+N} + \sum_{M<n\leqslant M+N-1}(a_n - a_{n+1})B_n,\end{aligned}$$

此即 (9.7)。若 $(a_n)_{n\in\mathbb{Z}}$ 单调且 $\sup\limits_{M<n\leqslant M+N}|B_n| \leqslant \rho$，则有

$$\begin{aligned}\Big|\sum_{M<n\leqslant M+N} a_n b_n\Big| &\leqslant |a_{M+N}B_{M+N}| + \sum_{M<n\leqslant M+N-1}\big|(a_n - a_{n+1})B_n\big| \\ &\leqslant \rho\bigg(|a_{M+N}| + \sum_{M<n\leqslant M+N-1}\big|(a_n - a_{n+1})\big|\bigg) \\ &= \rho\bigg(|a_{M+N}| + \Big|\sum_{M<n\leqslant M+N-1}(a_n - a_{n+1})\Big|\bigg) \\ &\leqslant \rho\big(|a_{M+1}| + 2|a_{M+N}|\big).\end{aligned}$$

□

**【例 3.2】**在上述定理中取 $a_n = \log n$，$b_n = (-1)^n$ 便可对任意的 $x \geqslant 2$ 得到

$$\sum_{n\leqslant x}(-1)^n \log n \ll \log x.$$

我们来将 (9.7) 与分部积分公式作一对比。若取 $a_n = f(n)$，$b_n = g(n)$，并将求和号与积分号对应起来，将差 $a_{n+1} - a_n$ 与微分 $\mathrm{d}f(x)$ 对应，再将部分和 $B_n$ 与积分 $G(x) = \displaystyle\int_M^x g(t)\,\mathrm{d}t$ 对应，那么我们就会发现 (9.7) 与分部积分公式

$$\int_M^{M+N} f(x)g(x)\,\mathrm{d}x = f(M+N)G(M+N) - \int_M^{M+N} G(x)\,\mathrm{d}f(x)$$

非常相似（注意到当 $g$ 连续时 $g(x) = G'(x)$）。因此人们通常将 (9.7) 和 (9.8) 称为分部求和 (summation by parts)。

**【定理 3.3】（Abel 判别法）** 如果数列 $\{a_n\}$ 单调且有界，级数 $\sum\limits_{n=1}^{\infty} b_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} a_n b_n$ 收敛。

<u>证明</u>. 由 $\{a_n\}$ 有界知存在 $A > 0$ 使得 $|a_n| \leqslant A\ (\forall\ n \geqslant 1)$。此外，因为 $\sum\limits_{n=1}^{\infty} b_n$ 收敛，故由 Cauchy 收敛准则知，对任意的 $\varepsilon > 0$，存在正整数 $n_\varepsilon$，使得对任意的 $m > n > n_\varepsilon$ 均有

$$\Big|\sum_{n<k\leqslant m} b_k\Big| < \varepsilon.$$

于是由 (9.8) 知，对任意的 $m > n > n_\varepsilon$ 有

$$\Big|\sum_{n<k\leqslant m} a_k b_k\Big| < \varepsilon\big(|a_{n+1}| + 2|a_m|\big) \leqslant 3A\varepsilon.$$

因此 $\sum\limits_{n=1}^{\infty} a_n b_n$ 收敛。 □

**【定理 3.4】（Dirichlet 判别法）** 若数列 $\{a_n\}$ 单调趋于 0，且级数 $\sum\limits_{n=1}^{\infty} b_n$ 的部分和有界，则 $\sum\limits_{n=1}^{\infty} a_n b_n$ 收敛。

<u>证明</u>. 因为 $\sum\limits_{n=1}^{\infty} b_n$ 的部分和 $B_n$ 有界，所以存在 $\rho > 0$ 使得 $|B_n| \leqslant \rho\ (\forall\ n \geqslant 1)$。又由 $\{a_n\}$ 趋于 0 知，对任意的 $\varepsilon > 0$，存在 $n_\varepsilon \in \mathbb{Z}_{>0}$ 使得当 $n > n_\varepsilon$ 时有 $|a_n| < \varepsilon$。注意到 $\{a_n\}$ 单调，故由 (9.8) 知对任意的 $m > n > n_\varepsilon$ 均有

$$\Big|\sum_{n<k\leqslant m} a_k b_k\Big| < 2\rho\big(|a_{n+1}| + 2|a_m|\big) < 6\rho\varepsilon.$$

因此由 Cauchy 收敛准则可得 $\sum\limits_{n=1}^{\infty} a_n b_n$ 收敛。 □

下面来看一些例子。

**【例 3.5】** 设数列 $\{a_n\}$ 单调趋于 0，$x \neq 2k\pi\ (k \in \mathbb{Z})$。证明级数

$$\sum_{n=1}^{\infty} a_n \sin nx \qquad \text{与} \qquad \sum_{n=1}^{\infty} a_n \cos nx$$

均收敛[6]。

[6] 事实上，第一个级数对任意的实数 $x$ 均收敛，这是因为当 $x = 2k\pi\ (k \in \mathbb{Z})$ 时其通项恒等于 0。

证明. 由于当 $x\neq 2k\pi$ $(k\in\mathbb{Z})$ 时 $e^{\mathrm{i}x}\neq 1$（这里 i 是虚数单位），故而

$$\sum_{k=1}^{n} e^{\mathrm{i}kx}=\frac{e^{\mathrm{i}x}(1-e^{\mathrm{i}nx})}{1-e^{\mathrm{i}x}}=\frac{e^{\mathrm{i}\frac{x}{2}}(1-e^{\mathrm{i}nx})}{e^{-\mathrm{i}\frac{x}{2}}-e^{\mathrm{i}\frac{x}{2}}},$$

因此

$$\left|\sum_{k=1}^{n} e^{\mathrm{i}kx}\right|\leqslant\frac{2}{\left|e^{-\mathrm{i}\frac{x}{2}}-e^{\mathrm{i}\frac{x}{2}}\right|}=\frac{1}{\left|\sin\dfrac{x}{2}\right|}. \tag{9.9}$$

上面最后一步用到了 Euler 公式 $e^{\mathrm{i}\theta}=\cos\theta+\mathrm{i}\sin\theta$，再次使用该公式可得

$$\left|\sum_{k=1}^{n}\sin kx\right|\leqslant\frac{1}{\left|\sin\dfrac{x}{2}\right|}\qquad\text{及}\qquad\left|\sum_{k=1}^{n}\cos kx\right|\leqslant\frac{1}{\left|\sin\dfrac{x}{2}\right|},$$

这说明级数 $\sum\limits_{n=1}^{\infty}\sin nx$ 与 $\sum\limits_{n=1}^{\infty}\cos nx$ 的部分和均有界。于是由 Dirichlet 判别法知 $\sum\limits_{n=1}^{\infty}a_n\sin nx$ 与 $\sum\limits_{n=1}^{\infty}a_n\cos nx$ 均收敛。 □

**【例 3.6】** 由上例知级数 $\sum\limits_{n=1}^{\infty}\dfrac{\sin nx}{n}$ 收敛，此外我们知道数列 $\left\{\left(1+\dfrac{1}{n}\right)^n\right\}$ 单调有界，故由 Abel 判别法可得级数 $\sum\limits_{n=1}^{\infty}\dfrac{\sin nx}{n}\left(1+\dfrac{1}{n}\right)^n$ 收敛。

值得一提的是，Dirichlet 判别法蕴含 Abel 判别法。事实上，若定理 3.3 的两个条件满足，那么由 $\{a_n\}$ 单调有界知其收敛，不妨将其极限记作 $a$；另外，由 $\sum\limits_{n=1}^{\infty}b_n$ 收敛知其部分和有界，从而由 Dirichlet 判别法知级数

$$\sum_{n=1}^{\infty}(a_n-a)b_n$$

收敛，注意到 $a_nb_n=(a_n-a)b_n+ab_n$，故而级数 $\sum\limits_{n=1}^{\infty}a_nb_n$ 收敛。

作为 Dirichlet 判别法的一个推论，我们给出下述的 Leibniz 判别法。

**【命题 3.7】（Leibniz 判别法）** 如果数列 $\{u_n\}$ 单调递减趋于 0，则级数 $\sum\limits_{n=1}^{\infty}(-1)^n u_n$ 收敛。

证明. 这可由 $\sum\limits_{n=1}^{\infty}(-1)^n$ 的部分和有界及 Dirichlet 判别法推出。 □

如果每个 $u_n$ 均不为 0 且具有相同的符号，那么就称形如 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}u_n$ 的级数为交错级数 (alternating series)，因此 Leibniz 判别法说明了一类特殊的交错级数的收敛性。

由 Leibniz 判别法可立即得到级数 $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n^s}$ $(s>0)$ 收敛。

## 习 题 9.3

1. 判定下列级数的敛散性：

(1) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{\log n}{n}$；

(2) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n\cos n}{n}$；

(3) $\displaystyle\sum_{n=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}-\log n\right)\sin\frac{\pi}{4}n$；

(4) $\displaystyle\sum_{n=1}^{\infty}(-1)^n(\sqrt[n]{n}-1)$；

(5) $\displaystyle\sum_{n=1}^{\infty}\sin\left(\pi\sqrt{n^2+1}\right)$；

(6) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n}{\sqrt{n}}\left(1+\frac{1}{n}\right)^{n+1}$；

(7) $\displaystyle\sum_{n=2}^{\infty}\frac{(-1)^n}{\log^2 n}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right)$；

(8) $\displaystyle\sum_{n=1}^{\infty}\frac{\sin n\,\sin n^2}{n}$；

(9) $\displaystyle\sum_{n=2}^{\infty}\frac{(-1)^n}{(n+(-1)^n)^p}$；

(10) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{[\sqrt{n}]}}{n^p}$。

2. 如果级数 $\sum\limits_{n=1}^{\infty}u_n$ 与 $\sum\limits_{n=1}^{\infty}v_n$ 均收敛，那么 $\sum\limits_{n=1}^{\infty}u_nv_n$ 是否一定收敛？

3. 设数列 $\{u_n\}$ 与 $\{v_n\}$ 满足 $\lim\limits_{n\to\infty}\dfrac{u_n}{v_n}=1$。若级数 $\sum\limits_{n=1}^{\infty}u_n$ 收敛，是否能得出级数 $\sum\limits_{n=1}^{\infty}v_n$ 也收敛?

4. 设级数 $\sum\limits_{n=1}^{\infty}a_n$ 收敛，$\{b_n\}$ 为严格单调递增的正数列且 $\lim\limits_{n\to\infty}b_n=+\infty$。证明当 $n\to\infty$ 时有
$$\sum_{k=1}^{n}a_kb_k=o(b_n).$$

5. 设正数列 $\{u_n\}$ 满足 $u_n=O\Big(\dfrac{1}{\log^2 n}\Big)$，证明级数 $\sum\limits_{n=1}^{\infty}(u_{n+1}-u_n)\log n$ 收敛。

6. 对正偶数 $N$ 证明

$$\sum_{n\leqslant N}(-1)^n\log n \gg \log N.$$

这意味着在不考虑 $\ll$ 常数的情况下例 3.2 中的估计是最优的。

7. 设 $x\in(0,\pi]$，证明对任意的 $N\in\mathbb{Z}_{\geqslant 2}$ 有

$$\sum_{n=1}^{N}(\log n)\sin nx \ll (\log N)\cdot\min\left(N^2x,\frac{1}{x}\right).$$

8. 设 $t\in\mathbb{R}$，试对 $x\in(0,1]$ 讨论级数 $\sum\limits_{n=1}^{\infty}\dfrac{\cos(t\log n)}{n^x}$ 的敛散性。

## § 9.4 绝对收敛与条件收敛

**【定义 4.1】**若 $\sum\limits_{n=1}^{\infty}|u_n|$ 收敛，则称 $\sum\limits_{n=1}^{\infty}u_n$ 绝对收敛 (absolutely convergent)。

**【命题 4.2】**绝对收敛级数必收敛，但反之不然。

证明. 若 $\sum\limits_{n=1}^{\infty}u_n$ 绝对收敛，则由 Cauchy 收敛准则知：对任意的 $\varepsilon>0$，均存在 $n_\varepsilon\in\mathbb{Z}_{>0}$，使得对任意的 $m>n>n_\varepsilon$ 有

$$\sum_{n<k\leqslant m}|u_k|<\varepsilon.$$

于是由三角形不等式知

$$\left|\sum_{n<k\leqslant m}u_k\right|<\varepsilon.$$

进而由 Cauchy 收敛准则知 $\sum\limits_{n=1}^{\infty}u_n$ 收敛。

此外，收敛的级数未必绝对收敛，例如 $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n}$。 □

**【定义 4.3】**若级数 $\sum\limits_{n=1}^{\infty}u_n$ 收敛但不绝对收敛，则称之条件收敛 (conditionally convergent)。

**【例 4.4】**证明级数 $\sum_{n=1}^{\infty} \frac{\sin n}{n}$ 条件收敛。

<u>证明</u>. 由例 3.4 知该级数收敛，下证其不绝对收敛。我们有

$$\left|\frac{\sin n}{n}\right| \geqslant \frac{\sin^2 n}{n} = \frac{1-\cos 2n}{2n}.$$

由于 $\sum_{n=1}^{\infty} \frac{1}{2n}$ 发散且 $\sum_{n=1}^{\infty} \frac{\cos 2n}{2n}$ 收敛，故 $\sum_{n=1}^{\infty} \frac{1-\cos 2n}{2n}$ 发散。再由比较判别法知级数 $\sum_{n=1}^{\infty} \left|\frac{\sin n}{n}\right|$ 发散。 □

由于加法满足交换律和结合律，故当我们把有限多个数相加时，任意改变求和项的次序不会影响求和结果。但级数的情形就不同了，在级数收敛的情况下，改变求和项的顺序是有可能改变级数的性态的。在这一点上，绝对收敛级数与条件收敛级数显示出了巨大的差异，我们下面来逐一进行讨论。

首先证明一个引理，在这里我们对实数 $x$ 采用记号

$$x^+ = \max(x,0), \qquad x^- = -\min(x,0).$$

于是

$$x^+ = \frac{|x|+x}{2} = \begin{cases} x, & \text{若 } x>0, \\ 0, & \text{若 } x \leqslant 0. \end{cases}$$

$$x^- = \frac{|x|-x}{2} = \begin{cases} -x, & \text{若 } x<0, \\ 0, & \text{若 } x \geqslant 0. \end{cases}$$

**【引理 4.5】**(1) 若 $\sum_{n=1}^{\infty} u_n$ 绝对收敛，则 $\sum_{n=1}^{\infty} u_n^+$ 与 $\sum_{n=1}^{\infty} u_n^-$ 均收敛。

(2) 若 $\sum_{n=1}^{\infty} u_n$ 条件收敛，则 $\sum_{n=1}^{\infty} u_n^+$ 与 $\sum_{n=1}^{\infty} u_n^-$ 均发散于 $+\infty$。

<u>证明</u>. (1) 若 $\sum_{n=1}^{\infty} u_n$ 绝对收敛，则由

$$0 \leqslant u_n^+, u_n^- \leqslant |u_n|$$

及比较判别法知 $\sum_{n=1}^{\infty} u_n^+$ 与 $\sum_{n=1}^{\infty} u_n^-$ 均收敛。

(2) 反设 $\sum_{n=1}^{\infty} u_n^+$ 收敛，则由 $\sum_{n=1}^{\infty} u_n$ 收敛及 $u_n^- = u_n^+ - u_n$ 知 $\sum_{n=1}^{\infty} u_n^-$ 收敛。再由 $|u_n| = u_n^+ + u_n^-$ 知 $\sum_{n=1}^{\infty} u_n$ 绝对收敛，从而与假设矛盾，故 $\sum_{n=1}^{\infty} u_n^+$ 发散。同理可证

$\sum\limits_{n=1}^{\infty} u_n^-$ 发散。此外，由于 $u_n^+$ 与 $u_n^-$ 均非负，所以这两个级数的部分和序列均单调递增，从而这两个级数均发散于 $+\infty$。 □

有了上述准备工作，我们首先来讨论绝对收敛级数的情形，下面的结果是由 Dirichlet 于 1837 年给出的。

**【定理 4.6】(Dirichlet) 设 $\sum\limits_{n=1}^{\infty} u_n$ 绝对收敛，$\sum\limits_{n=1}^{\infty} u'_n$ 是将 $\sum\limits_{n=1}^{\infty} u_n$ 的项重排后所得的级数，则 $\sum\limits_{n=1}^{\infty} u'_n$ 也绝对收敛，且它的和与原级数的和相同。**

证明. 证明分两步进行。

(1) 先讨论 $\sum\limits_{n=1}^{\infty} u_n$ 是正项级数的情形。分别以 $S_n$ 和 $S'_n$ 表示 $\sum\limits_{n=1}^{\infty} u_n$ 和 $\sum\limits_{n=1}^{\infty} u'_n$ 的部分和。由于 $\sum\limits_{n=1}^{\infty} u'_n$ 是将 $\sum\limits_{n=1}^{\infty} u_n$ 的项重排后所得，故存在双射 $\sigma: \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0}$ 使得 $u'_n = u_{\sigma(n)}$。对任意的正整数 $k$，当 $n > \max\limits_{1 \leqslant j \leqslant k} \sigma(j)$ 时有

$$S'_k = \sum_{j=1}^{k} u_{\sigma(j)} \leqslant \sum_{i=1}^{n} u_i = S_n.$$

因此令 $n \to \infty$ 便得 $S'_k \leqslant \sum\limits_{n=1}^{\infty} u_n$，于是由单调有界收敛原理知 $\sum\limits_{n=1}^{\infty} u'_n$ 收敛，且

$$\sum_{n=1}^{\infty} u'_n \leqslant \sum_{n=1}^{\infty} u_n.$$

同时，我们也可把 $\sum\limits_{n=1}^{\infty} u_n$ 看作是将 $\sum\limits_{n=1}^{\infty} u'_n$ 的项重排后所得的级数，所以同理可得

$$\sum_{n=1}^{\infty} u_n \leqslant \sum_{n=1}^{\infty} u'_n.$$

因而这两个级数的和相等。

(2) 再来讨论一般的情形。由引理 4.5 知 $\sum\limits_{n=1}^{\infty} u_n^+$ 与 $\sum\limits_{n=1}^{\infty} u_n^-$ 均收敛。因为 $\sum\limits_{n=1}^{\infty} u_n$ 绝对收敛，故由 (1) 中的讨论知 $\sum\limits_{n=1}^{\infty} |u'_n|$ 收敛，于是再由引理 4.5 可得级数 $\sum\limits_{n=1}^{\infty} u_n'^+$ 与 $\sum\limits_{n=1}^{\infty} u_n'^-$ 也均收敛。注意到重排并不会改变项的符号，故而 $\sum\limits_{n=1}^{\infty} u_n'^+$ 必是 $\sum\limits_{n=1}^{\infty} u_n^+$ 的重排，且 $\sum\limits_{n=1}^{\infty} u_n'^-$ 必是 $\sum\limits_{n=1}^{\infty} u_n^-$ 的重排，从而由 (1) 知

$$\sum_{n=1}^{\infty} u_n'^+ = \sum_{n=1}^{\infty} u_n^+, \qquad \sum_{n=1}^{\infty} u_n'^- = \sum_{n=1}^{\infty} u_n^-.$$

进而有
$$\sum_{n=1}^{\infty} u_n' = \sum_{n=1}^{\infty} u_n'^{+} - \sum_{n=1}^{\infty} u_n'^{-} = \sum_{n=1}^{\infty} u_n^{+} - \sum_{n=1}^{\infty} u_n^{-} = \sum_{n=1}^{\infty} u_n.$$
至此定理得证。 □

由上述定理知，对于绝对收敛级数而言，任意改变求和次序不会影响求和结果，这与有限多个数求和的情形相同，是一个非常好的性质。但这对条件收敛级数并不成立，事实上，Riemann 在其 1854 年的论文《Ueber die Darstellbarkeit einer function durch eine trigonometrische reihe （论函数的三角级数表示）》中给出了如下结论，人们常把它称作 Riemann 重排定理 (Riemann rearrangement theorem)。

**【定理 4.7】（Riemann）** 设级数 $\sum\limits_{n=1}^{\infty} u_n$ 条件收敛，则对任意的 $A$ （可以是实数，$+\infty$ 或 $-\infty$），均可将 $\sum\limits_{n=1}^{\infty} u_n$ 的项进行重排，使得重排后的级数的和为 $A$。

证明. 将 $\{u_n\}$ 中的全部非负项依次取出成为一个数列，记作 $\{a_n\}$；将 $\{u_n\}$ 中的全部负项依次取出成为一个数列，记作 $\{b_n\}$，由引理 4.5 (2) 知
$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (-b_n) = +\infty. \tag{9.10}$$

(1) 先考虑 $A \in \mathbb{R}$ 的情形，不妨设 $A \geqslant 0$。基本想法是：先依次取若干个非负项，使得其和刚好大于 $A$，再依次取若干个负项，使得这些负项与之前所取的非负项之和刚好小于 $A$，反复如此操作，注意到级数的通项趋于 0，故可使重排后的级数收敛且和为 $A$。具体来说，由 (9.10) 知存在正整数 $n_1$ 使得
$$\sum_{k \leqslant n_1 - 1} a_k \leqslant A < \sum_{k \leqslant n_1} a_k,$$
换句话说，也即
$$0 < \sum_{k \leqslant n_1} a_k - A \leqslant a_{n_1}.$$
再由 (9.10) 知存在正整数 $m_1$ 使得
$$\sum_{j \leqslant m_1 - 1} (-b_j) \leqslant \sum_{k \leqslant n_1} a_k - A < \sum_{j \leqslant m_1} (-b_j),$$
也即
$$b_{m_1} \leqslant \left( \sum_{k \leqslant n_1} a_k + \sum_{j \leqslant m_1} b_j \right) - A < 0.$$

当然，还存在正整数 $n_2 > n_1$ 使得

$$\left(\sum_{k\leqslant n_1} a_k + \sum_{j\leqslant m_1} b_j\right) - A + \sum_{n_1<k\leqslant n_2-1} a_k \leqslant 0$$

且

$$\left(\sum_{k\leqslant n_1} a_k + \sum_{j\leqslant m_1} b_j\right) - A + \sum_{n_1<k\leqslant n_2} a_k > 0$$

也即

$$0 < \left(\sum_{k\leqslant n_1} a_k + \sum_{j\leqslant m_1} b_j + \sum_{n_1<k\leqslant n_2} a_k\right) - A \leqslant a_{n_2}.$$

以此类推可得两列严格递增的正整数 $\{n_\ell\}$ 和 $\{m_\ell\}$，其对任意的 $\ell \geqslant 1$ 有

$$b_{m_\ell} \leqslant \left(\sum_{k\leqslant n_1} a_k + \sum_{j\leqslant m_1} b_j + \cdots + \sum_{n_{\ell-1}<k\leqslant n_\ell} a_k + \sum_{m_{\ell-1}<j\leqslant m_\ell} b_j\right) - A < 0.$$

以及

$$0 < \left(\sum_{k\leqslant n_1} a_k + \sum_{j\leqslant m_1} b_j + \cdots + \sum_{n_{\ell-1}<k\leqslant n_\ell} a_k + \sum_{m_{\ell-1}<j\leqslant m_\ell} b_j + \sum_{n_\ell<k\leqslant n_{\ell+1}} a_k\right) - A \leqslant a_{n_{\ell+1}}.$$

因为 $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = 0$，所以级数

$$(a_1 + \cdots + a_{n_1}) + (b_1 + \cdots + b_{m_1}) + (a_{n_1+1} + \cdots + a_{n_2}) + (b_{m_1+1} + \cdots + b_{m_2}) + \cdots$$

收敛且其和为 $A$。注意到上式同一个括号中的项均有相同的符号，故由命题 1.7 知，级数

$$a_1 + \cdots + a_{n_1} + b_1 + \cdots + b_{m_1} + a_{n_1+1} + \cdots + a_{n_2} + b_{m_1+1} + \cdots + b_{m_2} + \cdots$$

作为 $\sum\limits_{n=1}^{\infty} u_n$ 的一个重排是收敛的，且和为 $A$。

(2) 对于 $A$ 是 $+\infty$ 或 $-\infty$ 的情形，证明与 $A \in \mathbb{R}$ 的情形类似，在这里我们只简述其方法。以 $A = +\infty$ 为例，我们可以先依次选出若干非负项，使得其和刚好大于 1，在这些项后面添加一个负项；接着再依次添加若干非负项使得总和大于 2，然后在它们后面再添加一个负项。以此类推就可将原级数重排使得其和为 $+\infty$。 □

## 习 题 9.4

1. 讨论下列级数的绝对收敛性和条件收敛性:

(1) $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n\log n}$;　(2) $\sum\limits_{n=1}^{\infty}\dfrac{n^2\sin n}{2^n}$;

(3) $\sum\limits_{n=1}^{\infty}\log\left(1+\dfrac{\sin n}{n^2}\right)$;　(4) $\sum\limits_{n=1}^{\infty}(-1)^n\dfrac{n-1}{n+1}\cdot\dfrac{1}{\sqrt[2019]{n}}$;

(5) $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n^{p+\frac{1}{n}}}$;　(6) $\sum\limits_{n=1}^{\infty}(-1)^n(\sqrt[n]{n}-1)$。

2. 设 $\sum\limits_{n=1}^{\infty}u_n$ 条件收敛，$\sum\limits_{n=1}^{\infty}v_n$ 绝对收敛，证明 $\sum\limits_{n=1}^{\infty}(u_n+v_n)$ 条件收敛。

3. （Cauchy−Schwarz 不等式）设级数 $\sum\limits_{n=1}^{\infty}u_n^2$ 与 $\sum\limits_{n=1}^{\infty}v_n^2$ 均收敛，证明 $\sum\limits_{n=1}^{\infty}u_nv_n$ 绝对收敛，且有

$$\left(\sum_{n=1}^{\infty}|a_nb_n|\right)^2\leqslant\left(\sum_{n=1}^{\infty}a_n^2\right)\left(\sum_{n=1}^{\infty}b_n^2\right).$$

4. 设级数 $\sum\limits_{n=1}^{\infty}u_n^2$ 收敛，证明 $\sum\limits_{n=1}^{\infty}\dfrac{u_n}{n}$ 绝对收敛。

5. 设级数 $\sum\limits_{n=1}^{\infty}u_n$ 条件收敛，记

$$S_n^+=\sum_{k=1}^{n}u_k^+\quad 及\quad S_n^-=\sum_{k=1}^{n}u_k^-.$$

证明 $\lim\limits_{n\to\infty}\dfrac{S_n^+}{S_n^-}=1$。

6. 研究级数 $1-\dfrac{1}{2^\alpha}+\dfrac{1}{3^\beta}-\dfrac{1}{4^\alpha}+\dfrac{1}{5^\beta}-\dfrac{1}{6^\alpha}+\cdots$ 的敛散性，并在收敛时讨论其条件收敛和绝对收敛性。

7. 将级数 $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n-1}}{n^\alpha}$ 的项重排，依序先排 $p$ 个正项，接着依序排 $q$ 个负项，之后再依序排 $p$ 个正项和 $q$ 个负项，如此交替下去。证明:

(1) 若 $\alpha>1$，则重排级数的和与原级数的和相同;

(2) 若 $\alpha=1$，则重排级数的和为 $\log\left(2\sqrt{\dfrac{p}{q}}\right)$;

(3) 若 $0<\alpha<1$，则重排级数仅在 $p=q$ 时收敛。

8. 假设对任意无穷小量 $\{x_n\}$ 而言级数 $\sum\limits_{n=1}^{\infty} x_n u_n$ 均收敛，则 $\sum\limits_{n=1}^{\infty} u_n$ 绝对收敛。

9. 设级数 $\sum\limits_{n=1}^{\infty} u_n$ 条件收敛，其和为 $S$，又设 $\tau : \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0}$ 是双射，且重排级数 $\sum\limits_{n=1}^{\infty} u_{\tau(n)}$ 收敛于 $S' \neq S$。证明：对于任意的 $\rho > 0$，均存在正整数 $n$，使得 $|\tau(n) - n| > \rho$。

10. 若正项级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{u_n}$ 收敛，证明：

    (1) 若 $u_n$ 单调递增，则级数 $\sum\limits_{n=1}^{\infty} \dfrac{n}{u_1 + u_2 + \cdots + u_n}$ 收敛；

    (2) 对一般情形证明级数 $\sum\limits_{n=1}^{\infty} \dfrac{n}{u_1 + u_2 + \cdots + u_n}$ 收敛。

## §9.5 级数的乘法

在 §9.1 命题 1.4 中我们介绍了级数的线性运算，在本节中我们将介绍级数的乘法。

众所周知，对于形如

$$(a_1 + \cdots + a_n)(b_1 + \cdots + b_m)$$

的两个有限和相乘的情形，我们可以用分配律将其展开。但级数乘法的情况就不同了，因为这里有无穷多项，所以无法直接展开，从而我们需要像命题 1.4 中相加的情形一样，将那些所有可能出现的项 $a_i b_j$ $(i \geqslant 1, j \geqslant 1)$ 按某种方式进行合并以作为乘积级数的通项。常用的合并方式有两种，它们分别给出了级数的 Cauchy 乘积和 Dirichlet 乘积。

首先来介绍 Cauchy 乘积，为了说明清楚其由来，我们先给出形式幂级数的概念。

**【定义 5.1】**关于一个字母 $X$ 的<u>形式幂级数 (formal power series)</u> 是指如

$$\sum_{n=0}^{\infty} a_n X^n$$

的形式表达式，其中 $a_n$ $(n \geqslant 0)$ 均为实数。

下面来对形式幂级数的乘积作一个启发式的讨论。若两个形式幂级数 $\sum\limits_{n=0}^{\infty} a_n X^n$ 与 $\sum\limits_{n=0}^{\infty} b_n X^n$ 的乘积也是一个形式幂级数，记作 $\sum\limits_{n=0}^{\infty} c_n X^n$，那么 $c_n$ 应当等于多少呢？一个简单的想法是：把级数 $\sum\limits_{n=0}^{\infty} a_n X^n$ 的每一项与 $\sum\limits_{n=0}^{\infty} b_n X^n$ 的每一项相乘，并将 $X$ 的同次幂的项合并，如此一来应当有

$$c_n X^n = \sum_{k+\ell=n} a_k X^k \cdot b_\ell X^\ell = \Big(\sum_{k+\ell=n} a_k b_\ell\Big) X^n,$$

从而

$$c_n = \sum_{k+\ell=n} a_k b_\ell. \tag{9.11}$$

若取 $X=1$，那么由上述讨论知将由 (9.11) 式所定义的级数 $\sum\limits_{n=0}^{\infty} c_n$ 当作 $\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 的乘积是一个很自然的选择。

**【定义 5.2】**我们称由 (9.11) 所定义的级数 $\sum\limits_{n=0}^{\infty} c_n$ 为 $\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 的 Cauchy 乘积。

按照上面的讨论，我们自然期望 $\sum\limits_{n=0}^{\infty} c_n$ 收敛且有

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big). \tag{9.12}$$

但是一般来说，$\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 收敛并不能保证它们的 Cauchy 乘积的收敛性。例如取

$$a_n = b_n = \frac{(-1)^n}{\sqrt{n}}, \qquad \forall\, n \in \mathbb{Z}_{>0},$$

那么由 Leibniz 判别法知这两个级数均收敛，此时对 $n \geqslant 2$ 有

$$c_n = \sum_{k+\ell=n} a_k b_\ell = (-1)^n \sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}},$$

注意到由算术平均—几何平均不等式可得

$$|c_n| = \sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}} \geqslant \sum_{k=1}^{n-1} \frac{2}{k+(n-k)} = \frac{2(n-1)}{n} \geqslant 1, \qquad \forall\, n \geqslant 2,$$

因此 $\sum\limits_{n=2}^{\infty} c_n$ 发散。然而值得庆幸的是，我们有如下的 Cauchy 定理。

**【定理 5.3】（Cauchy）** 若 $\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 均绝对收敛，且它们的和分别为 $A$ 与 $B$，则将 $a_ib_j$ $(i,j\geqslant 0)$ 按任意方式排列所得的级数皆绝对收敛，且和为 $AB$。

证明. 设 $\sum\limits_{n=0}^{\infty} u_n$ 是将 $a_ib_j$ $(i,j\geqslant 0)$ 按某种方式排列后所得的级数，不妨记 $u_n=a_{i_n}b_{j_n}$，并将 $\sum\limits_{n=0}^{\infty}|u_n|$ 的部分和记作 $S_n$。那么，若令

$$N=\max(i_1,\cdots,i_n,j_1,\cdots,j_n),$$

则有

$$S_n=\sum_{k=0}^{n}|a_{i_k}b_{j_k}|\leqslant\Big(\sum_{i=0}^{N}|a_i|\Big)\Big(\sum_{j=0}^{N}|b_j|\Big)\leqslant\Big(\sum_{i=0}^{\infty}|a_i|\Big)\Big(\sum_{j=0}^{\infty}|b_j|\Big),$$

这说明部分和序列 $\{S_n\}$ 有上界，从而 $\sum\limits_{n=0}^{\infty}|u_n|$ 收敛。

下证级数 $\sum\limits_{n=0}^{\infty} u_n$ 的和为 $AB$。由定理 4.6 以及命题 1.6 知，对于任意一个把 $a_ib_j$ $(i,j\geqslant 0)$ 排列后所得的级数而言，将其项任意加括号后所得的级数均收敛于同一值，因此若记

$$v_n=\Big(\sum_{i=0}^{n}a_i\Big)\Big(\sum_{j=0}^{n}b_j\Big)-\Big(\sum_{i=0}^{n-1}a_i\Big)\Big(\sum_{j=0}^{n-1}b_j\Big)$$

（也即是说，$v_n$ 是 $i$ 与 $j$ 中至少有一个为 $n$ 的那些 $a_ib_j$ 之和），则 $\sum\limits_{n=0}^{\infty} v_n$ 收敛且 $\sum\limits_{n=0}^{\infty} u_n=\sum\limits_{n=0}^{\infty} v_n$。注意到级数 $\sum\limits_{n=0}^{\infty} v_n$ 的部分和为

$$\Big(\sum_{i=0}^{n}a_i\Big)\Big(\sum_{j=0}^{n}b_j\Big),$$

而这恰是 $\sum\limits_{n=0}^{\infty} a_n$ 的部分和与 $\sum\limits_{n=0}^{\infty} b_n$ 的部分和之积，所以

$$\sum_{n=0}^{\infty}v_n=\Big(\sum_{n=0}^{\infty}a_n\Big)\Big(\sum_{n=0}^{\infty}b_n\Big)=AB.$$

至此定理得证。 □

事实上，若只针对 Cauchy 乘积，那么上述定理的条件尚可减弱。

**【定理 5.4】(Mertens)** 设 $\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 均收敛，且两者中至少有一个绝对收敛，那么按 (9.11) 所定义的 Cauchy 乘积 $\sum\limits_{n=0}^{\infty} c_n$ 收敛，且有

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big).$$

<u>证明</u>. 不妨设 $\sum\limits_{n=0}^{\infty} a_n$ 绝对收敛。分别用 $A_n$，$B_n$ 和 $C_n$ 表示级数 $\sum\limits_{n=0}^{\infty} a_n$，$\sum\limits_{n=0}^{\infty} b_n$ 与 $\sum\limits_{n=0}^{\infty} c_n$ 的部分和，又将 $\sum\limits_{n=0}^{\infty} |a_n|$ 的部分和记作 $A_n^*$。再记

$$\sum_{n=0}^{\infty} a_n = A, \qquad \sum_{n=0}^{\infty} b_n = B, \qquad \sum_{n=0}^{\infty} |a_n| = A^*.$$

那么

$$\begin{aligned} C_n &= \sum_{k=0}^{n} c_k = \sum_{k=0}^{n} \sum_{i+j=k} a_i b_j = \sum_{i=0}^{n} a_i \sum_{j=0}^{n-i} b_j = \sum_{i=0}^{n} a_i B_{n-i} \\ &= B\sum_{i=0}^{n} a_i + \sum_{i=0}^{n} a_i(B_{n-i} - B) = A_n B + \sum_{i=0}^{n} a_i(B_{n-i} - B). \end{aligned}$$

因此只需证明 $\lim\limits_{n\to\infty} \sum\limits_{i=0}^{n} a_i(B_{n-i} - B) = 0$ 即可。

由 $\{B_n\}$ 收敛知该数列有界，即存在 $M > 0$ 使得 $|B_n| \leqslant M$ $(\forall\, n \in \mathbb{Z}_{\geqslant 0})$。于是

$$\begin{aligned} \Big|\sum_{i=0}^{n} a_i(B_{n-i} - B)\Big| &= \Big|\sum_{0\leqslant i\leqslant \frac{n}{2}} a_i(B_{n-i} - B) + \sum_{\frac{n}{2}<i\leqslant n} a_i(B_{n-i} - B)\Big| \\ &\leqslant \Big(\max_{0\leqslant i\leqslant \frac{n}{2}} |B_{n-i} - B|\Big) \cdot \sum_{0\leqslant i\leqslant \frac{n}{2}} |a_i| + \sum_{\frac{n}{2}<i\leqslant n-1} |a_i|(M + |B|) \\ &\leqslant A^* \cdot \Big(\max_{0\leqslant i\leqslant \frac{n}{2}} |B_{n-i} - B|\Big) + \big(A_n^* - A_{[\frac{n}{2}]}^*\big)(M + |B|). \end{aligned}$$

因为当 $n \to \infty$ 时上式右侧趋于 0，故而 $\lim\limits_{n\to\infty} \sum\limits_{i=0}^{n} a_i(B_{n-i} - B) = 0$。 □

值得一提的是，$\sum\limits_{n=0}^{\infty} a_n$ 与 $\sum\limits_{n=0}^{\infty} b_n$ 均发散亦不能推出它们的 Cauchy 乘积发散（参见习题 4）。

最后，或许有人会问：如果 $\sum\limits_{n=0}^{\infty} c_n$ 是 $\sum\limits_{n=0}^{\infty} a_n$ 和 $\sum\limits_{n=0}^{\infty} b_n$ 的 Cauchy 乘积，且这三个级数均收敛，那么 (9.12) 是否成立呢？Abel 对此给出了肯定的答复，我们将其证明放在 §10.4。

接下来介绍 Dirichlet 乘积。与 Cauchy 乘积类似，我们仍先作一个启发式的讨论，但在此之前，我们需要了解 Dirichlet 级数的概念。

**【定义 5.5】** 设 $s \in \mathbb{R}$。形如 $\sum\limits_{n=1}^{\infty} \dfrac{a_n}{n^s}$ 的级数被称作 Dirichlet 级数 (Dirichlet series)。

姑且不去研究 Dirichlet 级数的敛散性[7]，我们现在考虑的问题是：如果两个 Dirichlet 级数 $\sum\limits_{n=1}^{\infty} \dfrac{a_n}{n^s}$ 与 $\sum\limits_{n=1}^{\infty} \dfrac{b_n}{n^s}$ 的乘积也是一个 Dirichlet 级数，记作 $\sum\limits_{n=1}^{\infty} \dfrac{c_n}{n^s}$，那么 $c_n$ 应当是多少呢？与 Cauchy 乘积的情形类似，我们也会考虑合并同类项，因此一个自然的选择是将那些使得 $\dfrac{a_k}{k^s} \cdot \dfrac{b_\ell}{\ell^s}$ 的分母相同的项合并在一起，换句话说，

$$\frac{c_n}{n^s} = \sum_{k\ell=n} \frac{a_k}{k^s} \cdot \frac{b_\ell}{\ell^s} = \left( \sum_{k\ell=n} a_k b_\ell \right) \frac{1}{n^s},$$

也即

$$c_n = \sum_{k\ell=n} a_k b_\ell. \text{[8]} \tag{9.13}$$

若取 $s = 0$，那么由上述讨论知，把由 (9.13) 式所确定的级数 $\sum\limits_{n=1}^{\infty} c_n$ 作为 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} b_n$ 的积也是一个合理的选择。

**【定义 5.6】** 设 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} b_n$ 是两个级数，我们把由 (9.13) 式所确定的级数 $\sum\limits_{n=1}^{\infty} c_n$ 称为 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} b_n$ 的 Dirichlet 乘积。

两个级数均收敛亦不能保证它们的 Dirichlet 乘积收敛，这样的一个例子来源于数论。设 $q \in \mathbb{Z}_{>1}$，如果函数 $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$ 满足

(1) $\chi$ 以 $q$ 为周期；

(2) $\chi(n) \neq 0$ 当且仅当 $(n, q) = 1$[9]；

---

[7] Dirichlet 级数的敛散性会在 §10.4 习题 $11 \sim 15$ 中讨论。

[8] 我们通常将如此定义的 $\{c_n\}$ 称为 $\{a_n\}$ 与 $\{b_n\}$ 的 Dirichlet 卷积。

[9] 这里的 $(n, q)$ 表示 $n$ 与 $q$ 的最大公因数。

(3) 对任意的整数 $m$，$n$ 均有 $\chi(mn)=\chi(m)\chi(n)$，

则称它是一个Dirichlet 特征 (Dirichlet character)。可以证明 $|\chi(n)|\leqslant 1$，并且若存在与 $q$ 互素的 $n$ 使得 $\chi(n)\neq 1$，那么

$$\sum_{n\leqslant q}\chi(n)=0.$$ [10]

结合周期性便知对任意的正实数 $x$ 均有 $\left|\sum\limits_{n\leqslant x}\chi(n)\right|\leqslant\dfrac{q}{2}$。由此及 Dirichlet 判别法知级数

$$\sum_{n=2}^{\infty}\frac{\chi(n)}{\log n}$$

收敛。因此若用 $c_n$ 表示上述级数与自身的 Dirichlet 乘积之通项，则有

$$|c_n|=\left|\sum_{k\ell=n}\frac{\chi(k)}{\log k}\cdot\frac{\chi(\ell)}{\log\ell}\right|\geqslant\frac{|\chi(n)|}{(\log n)^2}\sum_{k\ell=n}1=\frac{|\chi(n)|\tau(n)}{(\log n)^2},$$

其中 $\tau(n)=\sum\limits_{k\ell=n}1$ 被称为除数函数 (divisor function)。可以证明，对任意的 $A>0$，均存在无穷多个 $n$ 使得 $|\chi(n)|\tau(n)>(\log n)^A$ 成立[11]，因此级数 $\sum\limits_{n=1}^{\infty}c_n$ 的通项不趋于 0，从而发散。

## 习 题 9.5

1. 试求 $\sum\limits_{n=1}^{\infty}a_n$ 与 $\sum\limits_{n=1}^{\infty}b_n$ 的 Cauchy 乘积，并研究其敛散性:

   (1) $a_n=b_n=q^n$，其中 $|q|<1$;

   (2) $a_n=x^n$，$b_n=(-x)^n$，其中 $|x|<1$;

   (3) $a_n=\dfrac{1}{n!}$，$b_n=\dfrac{(-1)^n}{n!}$;

   (4) $a_n=\dfrac{1}{n^\alpha}$，$b_n=\dfrac{1}{n^\beta}$，其中 $\alpha,\beta>0$;

2. 记 $E(x)=\sum\limits_{n=0}^{\infty}\dfrac{x^n}{n!}$，证明这一级数对任意的 $x\in\mathbb{R}$ 收敛且有

$$E(x+y)=E(x)E(y),\qquad\forall\,x,y\in\mathbb{R}.$$

[10] 参见 [35] §6.5 定理 3。

[11] 参见 [35] §9.3 定理 1。

3. 记 $S(x)=\sum\limits_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}$，$C(x)=\sum\limits_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}$。证明：
   (1) 这两个级数对任意的 $x\in\mathbb{R}$ 均绝对收敛；
   (2) $2S(x)C(x)=S(2x)$；
   (3) $2C^2(x)-1=C(2x)$。
4. 假设 $a_0=b_0=1$，当 $n\geqslant 1$ 时

$$a_n=-\left(\frac{3}{2}\right)^n,\qquad\qquad b_n=\left(\frac{3}{2}\right)^{n-1}\left(2^n+\frac{1}{2^{n+1}}\right).$$

   证明 $\sum\limits_{n=0}^{\infty}a_n$ 与 $\sum\limits_{n=0}^{\infty}b_n$ 均发散，但是它们的 Cauchy 乘积绝对收敛。
5. 对给定的实数 $\nu$ 记 $\sigma_\nu(n)=\sum\limits_{d|n}d^\nu$ [12]，证明

$$\sum_{n=1}^{\infty}\frac{\sigma_\nu(n)}{n^s}=\zeta(s)\zeta(s-\nu),\qquad \forall\ s>\max(1,\nu+1).$$

   特别地，取 $\nu=0$ 可得 $\sum\limits_{n=1}^{\infty}\frac{\tau(n)}{n^s}=\zeta(s)^2$ ($\forall\ s>1$)。

## §9.6 无穷乘积

【**定义 6.1**】设 $\{u_n\}$ 是一个数列。类似于级数，我们将形式的记号

$$\prod_{n=1}^{\infty}u_n=u_1\cdot u_2\cdots u_n\cdots$$

称为无穷乘积。

为了说明上面定义中形式的乘积是否有确切的意义，我们给出无穷乘积收敛的概念。

【**定义 6.2**】对任意的正整数 $n$，称

$$P_n=\prod_{k=1}^{n}u_k$$

[12] 这里 $\sum\limits_{d|n}$ 表示对 $n$ 的全部正因子求和。

为无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 的部分乘积。若数列 $\{P_n\}$ 收敛于某个不等于 0 的实数 $P$，则称无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 收敛，并称 $P$ 为该无穷乘积的积，记作

$$\prod_{n=1}^{\infty} u_n = P.$$

否则就称 $\prod\limits_{n=1}^{\infty} u_n$ 发散。

**【注 6.3】**必须注意的是，当部分乘积序列 $\{P_n\}$ 收敛于 0 时，我们仍将无穷乘积视作发散。特别地，若数列 $\{u_n\}$ 中有某一项等于 0，则无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 发散。

由定义 6.2 知，在无穷乘积中添加或删去有限多个异于 0 的项（但不改变原有的项的顺序）不会改变无穷乘积的敛散性。

**【例 6.4】**对无穷乘积 $\prod\limits_{n=2}^{\infty}\left(1-\dfrac{1}{n}\right)$ 而言，由于

$$P_n = \prod_{k=2}^{n}\left(1-\frac{1}{k}\right) = \prod_{k=2}^{n}\frac{k-1}{k} = \frac{1}{n},$$

故 $\lim\limits_{n\to\infty} P_n = 0$，从而该无穷乘积发散。

**【例 6.5】**对无穷乘积 $\prod\limits_{n=2}^{\infty}\left(1-\dfrac{1}{n^2}\right)$ 而言，由于

$$\begin{aligned} P_n &= \prod_{k=2}^{n}\left(1-\frac{1}{k^2}\right) = \prod_{k=2}^{n}\frac{(k-1)(k+1)}{k^2} = \prod_{k=2}^{n}\frac{k-1}{k}\cdot\prod_{k=2}^{n}\frac{k+1}{k} \\ &= \frac{1}{n}\cdot\frac{n+1}{2} = \frac{n+1}{2n}, \end{aligned}$$

故 $\lim\limits_{n\to\infty} P_n = \dfrac{1}{2}$，从而该无穷乘积收敛，且

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right) = \frac{1}{2}.$$

**【例 6.6】**设 $\theta\in(0,\pi)$，对无穷乘积 $\prod\limits_{n=1}^{\infty}\cos\dfrac{\theta}{2^n}$ 而言，

$$P_n = \prod_{k=1}^{n}\cos\frac{\theta}{2^k} = \frac{\sin\theta}{2^n\sin\frac{\theta}{2^n}},$$

因此 $\lim\limits_{n\to\infty} P_n = \dfrac{\sin\theta}{\theta}$，从而该无穷乘积收敛，且

$$\prod_{n=1}^{\infty} \cos\frac{\theta}{2^n} = \frac{\sin\theta}{\theta}.$$

无穷乘积有如下类似于级数的性质，我们将其证明留作练习。

**【命题 6.7】**设无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 收敛，那么

(1) $\lim\limits_{n\to\infty} u_n = 1$。

(2) 若记

$$\pi_n = \prod_{k=n+1}^{\infty} u_k,$$

则 $\lim\limits_{n\to\infty} \pi_n = 1$。

由于上述命题中的 (1)，我们往往只去研究各项均大于 0 的无穷乘积，因此**在下面的讨论中我们总假设无穷乘积的通项大于** 0。

因为通过取对数可将乘积化为求和，所以我们很自然地把无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 与级数 $\sum\limits_{n=1}^{\infty} \log u_n$ 联系起来[13]。

**【命题 6.8】**无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 收敛当且仅当级数 $\sum\limits_{n=1}^{\infty} \log u_n$ 收敛。

证明. 记

$$P_n = \prod_{k=1}^{n} u_k, \qquad S_n = \sum_{k=1}^{n} \log u_k,$$

则 $P_n = e^{S_n}$。故由指数函数的连续性知，若 $\lim\limits_{n\to\infty} S_n = S$，则有 $\lim\limits_{n\to\infty} P_n = e^S$。同样由对数函数的连续性知，如果 $\lim\limits_{n\to\infty} P_n = P > 0$，则 $\lim\limits_{n\to\infty} S_n = \log P$。从而命题得证。 □

**【命题 6.9】**$\prod\limits_{n=1}^{\infty} u_n$ 发散于 0 当且仅当 $\sum\limits_{n=1}^{\infty} \log u_n$ 发散于 $-\infty$。

证明. 沿用上述命题证明过程中的记号，我们有 $S_n = \log P_n$。显然 $P_n$ 趋于 0 当且仅当 $S_n$ 趋于 $-\infty$。 □

因为由命题 6.7 (1) 知无穷乘积 $\prod\limits_{n=1}^{\infty} u_n$ 收敛的必要条件是 $\lim\limits_{n\to\infty} u_n = 1$，故而在下面的讨论中记 $u_n = 1 + a_n$。

首先来研究 $a_n$ 保持一定符号的情形。

---

[13] 由此可以看出，把那些部分乘积趋于 0 的无穷乘积定义为发散的是一个方便且合理的做法。

**【命题 6.10】**假设自某项开始有 $a_n>0$（或自某项开始有 $a_n<0$），那么无穷乘积 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 与级数 $\sum\limits_{n=1}^{\infty}a_n$ 同敛散。

<u>证明</u>. 若 $\{a_n\}$ 不趋于 0，则 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 与 $\sum\limits_{n=1}^{\infty}a_n$ 均发散。下设 $\lim\limits_{n\to\infty}a_n=0$。此时有

$$\lim_{n\to\infty}\frac{\log(1+a_n)}{a_n}=1,$$

因此利用正项级数的比较判别法可得 $\sum\limits_{n=1}^{\infty}\log(1+a_n)$ 与 $\sum\limits_{n=1}^{\infty}a_n$ 同敛散，再由命题 6.8 知 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 与 $\sum\limits_{n=1}^{\infty}a_n$ 同敛散。 □

**【推论 6.11】**假设对任意的正整数 $n$ 均有 $a_n\in(-1,0]$，且级数 $\sum\limits_{n=1}^{\infty}a_n$ 发散，则 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 发散于 0。

<u>证明</u>. 按照比较判别法，由 $\sum\limits_{n=1}^{\infty}a_n$ 发散可推出级数 $\sum\limits_{n=1}^{\infty}\log(1+a_n)$ 发散。注意到 $\log(1+a_n)\leqslant 0$，故 $\sum\limits_{n=1}^{\infty}\log(1+a_n)$ 必发散于 $-\infty$，再由命题 6.9 知无穷乘积 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 发散于 0。 □

**【例 6.12】**由于 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^s}$ 当 $s>1$ 时收敛，当 $s\leqslant 1$ 时发散，故无穷乘积

$$\prod_{n=1}^{\infty}\left(1+\frac{1}{n^s}\right)$$

也在 $s>1$ 时收敛，在 $s\leqslant 1$ 时发散。

接下来讨论 $a_n$ 可正可负的情形。

**【命题 6.13】**若 $\sum\limits_{n=1}^{\infty}a_n^2$ 收敛，则无穷乘积 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 与级数 $\sum\limits_{n=1}^{\infty}a_n$ 同敛散。

<u>证明</u>. 由 $\sum\limits_{n=1}^{\infty}a_n^2$ 收敛知 $\lim\limits_{n\to\infty}a_n=0$，因此

$$\lim_{n\to\infty}\frac{a_n-\log(1+a_n)}{a_n^2}=\frac{1}{2}. \tag{9.14}$$

于是利用正项级数的比较判别法可得 $\sum\limits_{n=1}^{\infty}\big[a_n-\log(1+a_n)\big]$ 收敛，故 $\sum\limits_{n=1}^{\infty}\log(1+a_n)$ 与 $\sum\limits_{n=1}^{\infty}a_n$ 同敛散，再结合命题 6.8 知结论成立。 □

【**例 6.14**】我们来研究无穷乘积 $\prod_{n=1}^{\infty}\left(1+\frac{\sin n}{n^s}\right)$ 在 $s>0$ 时的敛散性。

(1) 当 $s>\frac{1}{2}$ 时

$$\sum_{n=1}^{\infty}\frac{\sin n}{n^s} \qquad 与 \qquad \sum_{n=1}^{\infty}\frac{\sin^2 n}{n^{2s}}$$

均收敛，故由命题 6.13 知此时无穷乘积收敛。

(2) 当 $0<s\leqslant\frac{1}{2}$ 时，为方便起见记 $a_n=\frac{\sin n}{n^s}$，则 $\lim\limits_{n\to\infty}a_n=0$，于是 (9.14) 成立。进而由 $\sum\limits_{n=1}^{\infty}a_n^2$ 发散知级数 $\sum\limits_{n=1}^{\infty}\big[a_n-\log(1+a_n)\big]$ 发散。注意到 $\sum\limits_{n=1}^{\infty}a_n$ 收敛，故 $\sum\limits_{n=1}^{\infty}\log(1+a_n)$ 发散于 $-\infty$，所以当 $0<s\leqslant\frac{1}{2}$ 时无穷乘积 $\prod\limits_{n=1}^{\infty}\left(1+\frac{\sin n}{n^s}\right)$ 发散于 0。

无穷乘积也有绝对收敛的概念。

【**定义 6.15**】若 $\prod\limits_{n=1}^{\infty}(1+|a_n|)$ 收敛，则称无穷乘积 $\prod\limits_{n=1}^{\infty}(1+a_n)$ <u>绝对收敛</u>。

例如当 $\frac{1}{2}<s\leqslant 1$ 时无穷乘积 $\prod\limits_{n=1}^{\infty}\left(1+\frac{(-1)^n}{n^s}\right)$ 收敛，但不绝对收敛。

## 习 题 9.6

1. 证明命题 6.7。
2. 判断下列无穷乘积的敛散性:

   (1) $\prod\limits_{n=1}^{\infty}\frac{2n-1}{n^2+1}$；　　(2) $\prod\limits_{n=1}^{\infty}a^{\frac{(-1)^n}{n}}\quad(a>0)$；

   (3) $\prod\limits_{n=1}^{\infty}\left(1+\frac{x^n}{2^n}\right)$；　　(4) $\prod\limits_{n=1}^{\infty}\sqrt[n^2]{n}$；

   (5) $\prod\limits_{n=1}^{\infty}\sqrt[n]{1+\frac{1}{n}}$；　　(6) $\prod\limits_{n=1}^{\infty}e^{-\frac{1}{n}}\cdot\sqrt{1+\frac{2}{n}}$。

3. 假设对任意的正整数 $n$ 有 $u_n\in(0,1)$，证明无穷乘积 $\prod\limits_{n=1}^{\infty}\frac{\sin u_n}{u_n}$ 与级数 $\sum\limits_{n=1}^{\infty}u_n^2$ 同敛散。
4. 证明 $\prod\limits_{n=1}^{\infty}\frac{1}{e}\left(1+\frac{1}{n}\right)^n=0$，并由此得出 $\lim\limits_{n\to\infty}\left(\frac{n}{e}\right)^n\cdot\frac{1}{n!}=0$。

5. （Wallis 公式）利用第八章例 5.12 证明

$$\prod_{n=1}^{\infty}\left(1-\frac{1}{(2n+1)^2}\right)=\frac{\pi}{4} \qquad \text{与} \qquad \prod_{n=1}^{\infty}\left(1-\frac{1}{4n^2}\right)=\frac{2}{\pi}.$$

6. 假设 $\sum\limits_{n=1}^{\infty} b_n$ 绝对收敛，且正数列 $\{a_n\}$ 满足 $\dfrac{a_n}{a_{n+1}}=1+\dfrac{1}{n}+O(|b_n|)$。证明级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散。

7. 设

$$a_n=\begin{cases}-\dfrac{1}{\sqrt{k}}, & \text{若 } n=2k-1,\\ \dfrac{1}{\sqrt{k}}+\dfrac{1}{k}+\dfrac{1}{k\sqrt{k}}, & \text{若 } n=2k.\end{cases}$$

证明 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} a_n^2$ 均发散，但 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 收敛。

## § 9.7 应用

作为本章的结束，我们通过两个例子来略微展示一下级数与无穷乘积的应用。

**【例 7.1】**容易看出，对任意的 $x\in(-1,1)$，无穷乘积 $\prod\limits_{k=1}^{\infty}\left(1+x^{2^{k-1}}\right)$ 与级数 $\sum\limits_{n=1}^{\infty} x^{n-1}$ 均收敛。注意到

$$(1-x)\prod_{k=1}^{N}\left(1+x^{2^{k-1}}\right)=1-x^{2^N},$$

因此

$$\prod_{k=1}^{N}\left(1+x^{2^{k-1}}\right)=\frac{1-x^{2^N}}{1-x}=\sum_{n=1}^{2^N} x^{n-1},$$

令 $N\to\infty$ 便得

$$\prod_{k=1}^{\infty}\left(1+x^{2^{k-1}}\right)=\sum_{n=1}^{\infty} x^{n-1}.$$

上面两个式子的算术意义是：每个自然数均可用二进制唯一表示。

**【例 7.2】**在这个例子中，让我们来看看 Euler 是如何证明“存在无穷多个素数”的。首先需要建立如下恒等式（称为 Euler 恒等式）：

$$\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{p}\left(1-\frac{1}{p^s}\right)^{-1},\qquad \forall\, s>1, \tag{9.15}$$

其中 $\prod\limits_{p}$ 表示乘积通过所有素数。记 $P_n=\prod\limits_{p\leqslant n}\left(1-\dfrac{1}{p^s}\right)^{-1}$，则

$$P_n=\prod_{p\leqslant n}\left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots\right),$$

其中，由 $s>1$ 知对每个给定的 $p$ 而言，级数

$$1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots$$

绝对收敛，又注意到每个大于 1 的整数均可用唯一的方式写成素数的乘积（在不计乘积次序的前提下）[14]，故而

$$\sum_{k=1}^{n}\frac{1}{k^s}\leqslant P_n\leqslant\sum_{k=1}^{\infty}\frac{1}{k^s},$$

因此 (9.15) 成立。现在可以来证明我们的命题了，反设只有有限多个素数，那么 (9.15) 右侧是有限多项相乘，因此令 $s\to 1^+$ 便得

$$\lim_{s\to 1^+}\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{p}\left(1-\frac{1}{p}\right)^{-1}<+\infty. \tag{9.16}$$

然而当 $s>1$ 时，对任意的正整数 $N$ 有

$$\begin{aligned}\sum_{n\leqslant 2^N}\frac{1}{n^s}&=1+\sum_{j=0}^{N-1}\sum_{2^j<n\leqslant 2^{j+1}}\frac{1}{n^s}\geqslant 1+\sum_{j=0}^{N-1}2^j\cdot\frac{1}{2^{s(j+1)}}\\&>\frac{1}{2^s}\sum_{j=0}^{N-1}\frac{1}{\left(2^{s-1}\right)^j}=\frac{1}{2^s}\cdot\frac{1-\left(2^{1-s}\right)^N}{1-2^{1-s}},\end{aligned}$$

令 $N\to\infty$ 便知

$$\sum_{n=1}^{\infty}\frac{1}{n^s}\geqslant\frac{1}{2^s-2},$$

[14] 这一结论被称作算术基本定理 (fundamental theorem of arithmetic)。

因此

$$\lim_{s\to 1^+}\sum_{n=1}^{\infty}\frac{1}{n^s}=+\infty,$$

这与 (9.16) 式矛盾。这样就证明了存在无穷多个素数。

值得一提的是，Euler 的上述证明是数学历史上第一次运用分析工具去处理算术问题。

## 习 题 9.7

1. 设定义在正整数集上的函数 $\mu(n)$ 满足[15]

$$\mu(n)=\begin{cases}1, & 若\ n=1,\\ (-1)^k, & 若\ n=p_1p_2\cdots p_k\ 且\ p_1,p_2,\cdots,p_k\ 为两两不同的素数;\\ 0, & 其它.\end{cases}$$

证明：对任意的 $s>1$，有

$$\prod_p\left(1-\frac{1}{p^s}\right)=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}.$$

2. 设 $\mu(n)$ 如上题所定义，证明当 $s>1$ 时有

$$\left(\sum_{n=1}^{\infty}\frac{1}{n^s}\right)\left(\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}\right)=1.$$

3. 设 $z\in\mathbb{R}_{\geqslant 2}$，用 $P(z)$ 表示小于 $z$ 的全部素数的乘积。证明：当 $s>1$ 时有

$$\sum_{\substack{n=1\\(n,P(z))=1}}^{\infty}\frac{1}{n^s}=\zeta(s)\prod_{p<z}\left(1-\frac{1}{p^s}\right).$$

[15] $\mu(n)$ 被称为 Möbius 函数 (Möbius function)。

第十章

# 函数列与函数项级数

对幂级数可以应用四则运算，并进行逐项微分与积分，这一事实的证明以及通过这种证明而获得的对幂级数用处的认识，大大促进了整个分析的简化。

—— D. Hilbert

我们在第三章和第九章中分别学习了数列和数项级数的理论，在那里通项均是常数。在本章中我们将讨论更一般的情形，即通项均是函数的情形。

## § 10.1 引言

【**定义 1.1**】设 $A \subseteq \mathbb{R}$。若对任意的正整数 $n$，$f_n(x)$ 均是定义在 $A$ 上的函数，则称 $\{f_n(x)\}$ 是一个定义在 $A$ 上的函数列 (sequence of functions)。此时，我们称 $A$ 的子集

$$E = \big\{x \in A : \{f_n(x)\} \text{ 收敛}\big\}$$

为函数列 $\{f_n(x)\}$ 的收敛集 (convergence set)。当 $E \neq \varnothing$ 时，在 $E$ 上可通过

$$x \longmapsto \lim_{n\to\infty} f_n(x)$$

定义函数 $f$，我们称 $f(x)$ 为函数列 $\{f_n(x)\}$ 的极限函数 (limiting function)，也称 $\{f_n(x)\}$ 逐点收敛 (pointwise convergent)（或简称收敛）于 $f(x)$。按照定义，

$$f(x) = \lim_{n\to\infty} f_n(x), \qquad \forall\, x \in E.$$

【例 1.2】考虑定义在 $\mathbb{R}$ 上的函数列 $\left\{\left(1+\dfrac{x}{n}\right)^n\right\}$，容易对任意的 $x\in\mathbb{R}$ 证明

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n=e^x,$$

因此该函数列的收敛集为 $\mathbb{R}$，且极限函数为 $e^x$。

类似可给出函数项级数的相关定义。

【定义 1.3】设 $A\subseteq\mathbb{R}$，且对任意的正整数 $n$，$u_n(x)$ 均是定义在 $A$ 上的函数。我们在 $A$ 上引入形式的记号

$$\sum_{n=1}^{\infty}u_n(x)=u_1(x)+u_2(x)+\cdots+u_n(x)+\cdots,$$

并称之为一个定义在 $A$ 上的函数项级数 (series of functions)。此时，称

$$S_n(x)=\sum_{k=1}^{n}u_k(x)$$

为该函数项级数的第 $n$ 个部分和，称 $\{S_n(x)\}$ 的收敛集 $E$ 为 $\sum\limits_{n=1}^{\infty}u_n(x)$ 的收敛集。若 $E\neq\varnothing$ 且 $S(x)$ 是 $\{S_n(x)\}$ 的极限函数，那么我们就说 $\sum\limits_{n=1}^{\infty}u_n(x)$ 在 $E$ 上逐点收敛于 $S(x)$，并将 $S(x)$ 称为 $\sum\limits_{n=1}^{\infty}u_n(x)$ 的和函数，记作

$$\sum_{n=1}^{\infty}u_n(x)=S(x),\qquad\forall\ x\in E.$$

众所周知，有限多个连续（相应地，可导、可积）函数的和仍是连续（相应地，可导、可积）的，并且在可导或可积的情形，和函数的导数（相应地，积分）是每一个函数的导数（相应地，积分）之和，换句话说，在求和个数有限的情形，微分号和积分号均可与求和号交换。人们最关心的问题是：这一性质是否对级数（即求和个数无限的情形）也成立？鉴于定义 1.3，我们可将这一问题转而对函数列来讨论。具体而言，设定义在区间 $I$ 上的函数列 $\{f_n(x)\}$ 在该区间上逐点收敛于 $f(x)$，那么当每个 $f_n(x)$ 均连续（相应地，可导、可积）时，$f(x)$ 是否也在 $I$ 上连续（相应地，可导、可积）呢？此外，若每个 $f_n(x)$ 及 $f(x)$ 均在 $I$ 上可导，那么是否有

$$\lim_{n\to\infty}f_n'(x)=f'(x)\tag{10.1}$$

成立呢？同样，当每个 $f_n(x)$ 及 $f(x)$ 均在 $I$ 上可积时，等式

$$\lim_{n\to\infty}\int_I f_n(x)\,\mathrm{d}x=\int_I f(x)\,\mathrm{d}x \tag{10.2}$$

是否成立？换句话说，微分号和积分号能否与极限号交换？

下面我们将举例说明上述问题均未必有肯定的答案。

**【例 1.4】**考虑定义在区间 $[0,1]$ 上的函数列 $\{x^n\}$，容易看出它在 $[0,1]$ 上逐点收敛于

$$f(x)=\begin{cases}0, & 若\ x\in[0,1),\\ 1, & 若\ x=1.\end{cases}$$

对任意的 $n$，函数 $x^n$ 在 $[0,1]$ 上可导（当然也在 $[0,1]$ 上连续），但极限函数在 $[0,1]$ 上不连续，当然也在该区间上不可导。

**【例 1.5】**在区间 $[0,1]$ 上考虑 $f_n(x)=\lim\limits_{k\to\infty}\cos^{2k}(n!\pi x)$，则有

$$f_n(x)=\begin{cases}1, & 若\ n!\,x\in\mathbb{Z},\\ 0, & 若\ n!\,x\notin\mathbb{Z}.\end{cases}$$

因此对任意给定的 $n$，$f_n(x)$ 仅在区间 $[0,1]$ 上的有限多个点处取值不等于 0，从而 $f_n(x)$ 在 $[0,1]$ 上可积。但容易证明 $\{f_n(x)\}$ 逐点收敛于 Dirichlet 函数 $D(x)$，而 $D(x)$ 在 $[0,1]$ 上不可积。

**【例 1.6】**设 $f_n(x)=\dfrac{\sin nx}{n}$，则 $\{f_n(x)\}$ 在 $\mathbb{R}$ 上逐点收敛于 $f(x)=0$，但由 $f_n'(x)=\cos nx$ 知 (10.1) 在 $\mathbb{R}$ 上不成立。

**【例 1.7】**设 $f_n(x)=2n^2xe^{-n^2x^2}$，则 $\{f_n(x)\}$ 在区间 $[0,1]$ 上逐点收敛于 $f(x)=0$，但是

$$\lim_{n\to\infty}\int_0^1 f_n(x)\,\mathrm{d}x=\lim_{n\to\infty}(1-e^{-n^2})=1\neq\int_0^1 f(x)\,\mathrm{d}x,$$

这说明 (10.2) 未必成立。

以上四个例子说明，为了保证我们之前提出的问题有肯定的答案，必须对函数列或函数项级数添加更多的限制条件，这些限制条件中至关重要的一个就是“一致收敛性”，我们将在接下来的两节中对其作细致的讨论。

在历史上，A. L. Cauchy 在其 1821 年的著作《Cours d'Analyse》中提到通项为连续函数的收敛级数的和函数也连续，他还给出了这一结果的“证明”。但是 N.

H. Abel 在 1826 年给出了一个反例，他证明了等式（参见第十八章例 2.10）

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n}=\frac{\pi-x}{2},\qquad \forall\, x\in(0,2\pi),$$

注意到 $\sin nx$ 以 $2\pi$ 为周期，所以上式左边的级数在 $\mathbb{R}$ 上的和函数不是 $\mathbb{R}$ 上的连续函数，而形如 $2k\pi$ $(k\in\mathbb{Z})$ 的点均是其间断点。事实上，Cauchy 只是对一致收敛的函数项级数证明了该结论，但他本人并未注意到这一点。一致收敛最早出现在 1842 年 K. Weierstrass 关于幂级数的一篇论文中，之后，P. L. von Seidel 和 G. G. Stokes 也于 1848 年在各自的文章中涉及到了这一概念。

## 习 题 10.1

1. 求函数项级数 $\sum\limits_{n=1}^{\infty}\log\left(1+\dfrac{(-1)^n}{n^x}\right)$ 的收敛集。

2. 记 $f_n(x)=\lim\limits_{k\to\infty}\cos^{2k}(n!\pi x)$，证明 $\{f_n(x)\}$ 在 $\mathbb{R}$ 上逐点收敛于 Dirichlet 函数

$$D(x)=\begin{cases}1, & 若\ x\in\mathbb{Q},\\ 0, & 若\ x\notin\mathbb{Q}.\end{cases}$$

3. 设 $\{f_n(x)\}$ 在 $E$ 上逐点收敛于 $f(x)$，且对任意的 $n\in\mathbb{Z}_{>0}$ 而言，$f_n(x)$ 均在 $E$ 上有界，问 $f(x)$ 是否在 $E$ 上有界？

4. 设 $\{f_n(x)\}$ 在有限闭区间 $[a,b]$ 上逐点收敛于 $f(x)$，且对任意的 $n\in\mathbb{Z}_{>0}$ 而言，$f_n(x)$ 均是 $[a,b]$ 上的连续函数，问 $f(x)$ 是否在 $[a,b]$ 上有界？

5. 设 $\{f_n(x)\}$ 在 $E$ 上逐点收敛于 $f(x)$。若对任意的 $M>0$，均存在 $n\in\mathbb{Z}_{\geqslant 1}$ 及 $x\in E$ 使得 $|f_n(x)|>M$，问 $f(x)$ 是否在 $E$ 上无界？

6. （de la Vallée Poussin）证明当 $s>1$ 时有

$$\sum_{n=1}^{\infty}\frac{1}{n^s}=\frac{1}{s-1}+\sum_{n=1}^{\infty}\left[\frac{1}{n^s}+\frac{1}{s-1}\left(\frac{1}{(n+1)^{s-1}}-\frac{1}{n^{s-1}}\right)\right].$$

并证明上式右边的级数在 $0<s<1$ 时也收敛，这意味着我们可以把 $\zeta(s)$ 延拓为 $(0,1)\cup(1,+\infty)$ 上的函数。

## §10.2 一致收敛

首先讨论函数列的情形。

**【定义 2.1】**设 $\{f_n(x)\}$ 在集合 $E$ 上逐点收敛于 $f(x)$。若对任意的 $\varepsilon > 0$，存在至多依赖于 $\varepsilon$ 的正整数 $N$，使得当 $n > N$ 时有

$$|f_n(x) - f(x)| < \varepsilon, \qquad \forall\, x \in E,$$

则称 $\{f_n(x)\}$ 在 $E$ 上一致收敛 (uniformly convergent) 于 $f(x)$。

由上述定义知，所谓 $\{f_n(x)\}$ 在 $E$ 上不一致收敛于 $f(x)$ 是指：存在 $\varepsilon_0 > 0$，使得对任意的正整数 $N$，均存在 $n > N$ 及 $x \in E$ 满足

$$|f_n(x) - f(x)| \geqslant \varepsilon_0.$$

**【例 2.2】**设 $f_n(x) = \dfrac{x}{1+n^2x^2}$，则 $\{f_n(x)\}$ 在 $\mathbb{R}$ 上逐点收敛于 0。对任意的 $\varepsilon > 0$。取 $N = \left[\dfrac{1}{2\varepsilon}\right] + 1$，则当 $n > N$ 时，由 $1 + n^2x^2 \geqslant 2n|x|$ 知

$$|f_n(x) - 0| = \frac{|x|}{1+n^2x^2} \leqslant \frac{1}{2n} < \frac{1}{2N} < \varepsilon.$$

因此 $\{f_n(x)\}$ 在 $\mathbb{R}$ 上一致收敛于 0。

**【例 2.3】**现在回到上节例 1.4，在那里我们研究了函数列 $\{x^n\}$，它在 $[0,1]$ 上逐点收敛于

$$f(x) = \begin{cases} 0, & 若\ x \in [0,1), \\ 1, & 若\ x = 1. \end{cases}$$

由于对任意的 $n \in \mathbb{Z}_{\geqslant 2}$ 有

$$\left|\left(1-\frac{1}{n}\right)^n - f\left(1-\frac{1}{n}\right)\right| = \left(1-\frac{1}{n}\right)^n \geqslant \left(1-\frac{1}{2}\right)^2 = \frac{1}{4},\ ①$$

故 $\{x^n\}$ 在 $[0,1]$ 上不一致收敛。

由定义可直接得出下述命题。

---

①这是因为数列 $\left\{\left(1-\dfrac{1}{n}\right)^n\right\}$ 单调递增，参见第三章例 3.5 的证明。

**【命题 2.4】**设 $\{f_n(x)\}$ 在 $E$ 上逐点收敛于 $f(x)$。记

$$\|f_n - f\| = \sup_{x\in E} |f_n(x) - f(x)|,$$

则 $\{f_n(x)\}$ 在 $E$ 上一致收敛当且仅当 $\lim\limits_{n\to\infty} \|f_n - f\| = 0$。

**【例 2.5】**对于在 $\mathbb{R}$ 上定义的函数 $f_n(x) = \dfrac{nx}{1+n^2x^2}$ 而言，对任意的 $x \in \mathbb{R}$ 有 $\lim\limits_{n\to\infty} f_n(x) = 0$。由于

$$\|f_n - 0\| = \sup_{x\in\mathbb{R}} |f_n(x)| \geqslant \left| f_n\Big(\frac{1}{n}\Big) \right| = \frac{1}{2},$$

故由命题 2.4 知 $\{f_n(x)\}$ 在 $\mathbb{R}$ 上不一致收敛。

**【例 2.6】**接下来我们回到上节例 1.7 中的函数 $f_n(x) = 2n^2xe^{-n^2x^2}$，由它所确定的函数列在 $(0,1)$ 上逐点收敛于 0。因为

$$\|f_n - 0\| = \sup_{x\in(0,1)} |f_n(x)| \geqslant \left| f_n\Big(\frac{1}{n}\Big) \right| = 2ne^{-1},$$

故 $\{f_n(x)\}$ 在 $(0,1)$ 上不一致收敛。但是，对 $(0,1)$ 的任一闭子区间 $[a,b]$ 而言，

$$\sup_{x\in[a,b]} |f_n(x) - f(x)| \leqslant 2n^2e^{-n^2a^2},$$

故而 $\lim\limits_{n\to\infty} \sup\limits_{x\in[a,b]} |f_n(x)| = 0$，这说明 $\{f_n(x)\}$ 在 $(0,1)$ 的任一闭子区间上一致收敛。由此引出如下定义。

**【定义 2.7】**若 $\{f_n(x)\}$ 在区间 $I$ 的任一闭子区间上一致收敛，则称它在 $I$ 上<u>内闭一致收敛</u>。

容易看出一致收敛蕴含内闭一致收敛，但反之不然（参见例 2.6）。然而由内闭一致收敛可推出逐点收敛。

下面我们对函数项级数给出相应的定义。

**【定义 2.8】**若函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的部分和序列 $\{S_n(x)\}$ 在集合 $E$ 上一致收敛于 $S(x)$，则称 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $E$ 上<u>一致收敛</u>于 $S(x)$。

**【定义 2.9】**若 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的部分和序列 $\{S_n(x)\}$ 在区间 $I$ 上内闭一致收敛，则称 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $I$ 上<u>内闭一致收敛</u>。

**【例 2.10】** 级数 $\sum\limits_{n=1}^{\infty} x^{n-1}$ 在开区间 $(-1,1)$ 上逐点收敛于 $S(x)=\dfrac{1}{1-x}$，该级数的部分和为

$$S_n(x)=\sum_{k=1}^{n} x^{k-1}=\frac{1-x^n}{1-x}.$$

由于

$$\sup_{x\in(-1,1)}|S_n(x)-S(x)|=\sup_{x\in(-1,1)}\frac{|x|^n}{1-x}=+\infty,$$

故 $\sum\limits_{n=1}^{\infty} x^{n-1}$ 在 $(-1,1)$ 上不一致收敛。但是对任意的 $[a,b]\subseteq(-1,1)$，由

$$\sup_{x\in[a,b]}|S_n(x)-S(x)|=\sup_{x\in[a,b]}\frac{|x|^n}{1-x}\leqslant\frac{1}{1-b}\max(|a|^n,\,|b|^n)$$

知 $\lim\limits_{n\to\infty}\sup\limits_{x\in[a,b]}|S_n(x)-S(x)|=0$，因此 $\sum\limits_{n=1}^{\infty} x^{n-1}$ 在 $(-1,1)$ 上内闭一致收敛。

当函数列或函数项级数的形式比较复杂时，从定义出发去判定其一致收敛性通常是极不方便的，下面我们来介绍一些判定方法。

**【定理 2.11】（Cauchy 收敛准则）** 设 $\{f_n(x)\}$ 是一个定义在 $E$ 上的函数列，那么它在 $E$ 上一致收敛的充要条件是：对任意的 $\varepsilon>0$，存在正整数 $N$，使得对任意的 $m,n>N$ 有

$$|f_n(x)-f_m(x)|<\varepsilon,\qquad \forall\, x\in E. \tag{10.3}$$

<u>证明</u>. 必要性：若 $\{f_n(x)\}$ 在 $E$ 上一致收敛于 $f(x)$，则对任意的 $\varepsilon>0$，存在正整数 $N$，使得当 $n>N$ 时有

$$|f_n(x)-f(x)|<\frac{\varepsilon}{2},\qquad \forall\, x\in E.$$

因此对任意的 $m,n>N$ 及 $x\in E$ 有

$$|f_n(x)-f_m(x)|\leqslant|f_n(x)-f(x)|+|f_m(x)-f(x)|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

充分性：此时由数列的 Cauchy 收敛准则知 $\{f_n(x)\}$ 在 $E$ 上逐点收敛，记其极限函数为 $f(x)$，那么在 (10.3) 中令 $m\to\infty$ 知，对任意的 $n>N$ 及 $x\in E$ 有

$$|f_n(x)-f(x)|\leqslant\varepsilon.$$

因此 $\{f_n(x)\}$ 在 $E$ 上一致收敛于 $f(x)$。 □

由这一定理可立即得到判定函数项级数一致收敛性的 Cauchy 准则。

**【定理 2.12】（Cauchy 收敛准则）** 设 $\sum_{n=1}^{\infty} u_n(x)$ 是一个定义在 $E$ 上的函数项级数，那么它在 $E$ 上一致收敛的充要条件是：对任意的 $\varepsilon > 0$，存在正整数 $N$，使得对任意的 $m > n > N$ 有

$$\left|\sum_{k=n+1}^{m} u_k(x)\right| < \varepsilon, \qquad \forall\, x \in E.$$

下面我们把讨论的重点转向函数项级数。首先给出一致收敛的一个必要条件，我们把它的证明留作练习。

**【命题 2.13】** 如果函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 $E$ 上一致收敛，那么 $\{u_n(x)\}$ 在 $E$ 上一致收敛于 0。

接下来介绍应用颇为广泛的 Weierstrass 判别法。

**【定理 2.14】（Weierstrass 判别法）** 假设存在正整数 $N$，使得当 $n > N$ 时对任意的 $x \in E$ 均有

$$|u_n(x)| \leqslant a_n. \tag{10.4}$$

如果数项级数 $\sum_{n=1}^{\infty} a_n$ 收敛，那么 $\sum_{n=1}^{\infty} u_n(x)$ 在 $E$ 上一致收敛。

<u>证明</u>. 因为 $\sum_{n=1}^{\infty} a_n$ 收敛，故对任意的 $\varepsilon > 0$，存在正整数 $M$，使得当 $m > n > M$ 时有

$$\sum_{k=n+1}^{m} a_k = \left|\sum_{k=n+1}^{m} a_k\right| < \varepsilon.$$

于是由 (10.4) 知，当 $m > n > \max(M, N)$ 时，对任意的 $x \in E$ 有

$$\left|\sum_{k=n+1}^{m} u_k(x)\right| \leqslant \sum_{k=n+1}^{m} |u_k(x)| \leqslant \sum_{k=n+1}^{m} a_k < \varepsilon,$$

再由定理 2.12 便知 $\sum_{n=1}^{\infty} u_n(x)$ 在 $E$ 上一致收敛。 □

在上面的定理中，由于函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 的通项被数项级数 $\sum_{n=1}^{\infty} a_n$ 的通项所控制，故而常将 $\sum_{n=1}^{\infty} a_n$ 称作 $\sum_{n=1}^{\infty} u_n(x)$ 的<u>优势级数 (majorant series)</u>。

**【例 2.15】** 若 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛，那么 $\sum\limits_{n=1}^{\infty} a_n \sin nx$ 与 $\sum\limits_{n=1}^{\infty} a_n \cos nx$ 均在 $\mathbb{R}$ 上一致收敛。

细心的读者一定意识到了 Weierstrass 判别法所需要的条件太强了。事实上，由 Weierstrass 判别法的条件可以看出 $\sum\limits_{n=1}^{\infty} |u_n(x)|$ 一致收敛。然而确实存在级数 $\sum\limits_{n=1}^{\infty} u_n(x)$，其一致收敛，但 $\sum\limits_{n=1}^{\infty} |u_n(x)|$ 不一致收敛，甚至于有可能 $\sum\limits_{n=1}^{\infty} |u_n(x)|$ 不收敛（即 $\sum\limits_{n=1}^{\infty} u_n(x)$ 不绝对收敛），关于这一点可参见本节末尾的讨论。因此，我们有必要对一些特殊情形给出更加细致的判别法。

**【定理 2.16】（Abel 判别法）** 假设

(1) $\sum\limits_{n=1}^{\infty} b_n(x)$ 在 $E$ 上一致收敛；

(2) 对任意给定的 $x \in E$，$\{a_n(x)\}$ 单调。并且 $\{a_n(x)\}$ 在 $E$ 上一致有界 (uniformly bounded)，也即存在常数 $M > 0$ 使得

$$|a_n(x)| \leqslant M, \qquad \forall\, n \in \mathbb{Z}_{>0} \text{ 及 } x \in E.$$

那么 $\sum\limits_{n=1}^{\infty} a_n(x) b_n(x)$ 在 $E$ 上一致收敛。

证明. 由 $\sum\limits_{n=1}^{\infty} b_n(x)$ 的一致收敛性知，对任意的 $\varepsilon > 0$，存在正整数 $N$，使得对于任意的 $m > n > N$ 及 $x \in E$ 有

$$\left| \sum_{k=n+1}^{m} b_k(x) \right| < \frac{\varepsilon}{3M}.$$

于是由分部求和（参见第九章定理 3.1）可得

$$\left| \sum_{k=n+1}^{m} a_k(x) b_k(x) \right| \leqslant \frac{\varepsilon}{3M} \cdot \big(|a_{n+1}(x)| + 2|a_m(x)|\big) \leqslant \frac{\varepsilon}{3M} \cdot 3M = \varepsilon,$$

再由 Cauchy 收敛准则知命题成立。 □

**【定理 2.17】（Dirichlet 判别法）** 假设

(1) $\sum\limits_{n=1}^{\infty} b_n(x)$ 的部分和序列 $\{B_n(x)\}$ 在 $E$ 上一致有界，即存在常数 $M > 0$ 使得

$$|B_n(x)| \leqslant M, \qquad \forall\, n \in \mathbb{Z}_{>0} \text{ 及 } x \in E;$$

(2) 对任意给定的 $x \in E$，$\{a_n(x)\}$ 单调。并且 $\{a_n(x)\}$ 在 $E$ 上一致收敛于 0。

那么 $\sum\limits_{n=1}^{\infty} a_n(x)b_n(x)$ 在 $E$ 上一致收敛。

<u>证明</u>. 由 $\{a_n(x)\}$ 在 $E$ 上一致收敛于 0 知，对任意的 $\varepsilon>0$，存在正整数 $N$，使得对任意的 $n>N$ 及 $x\in E$ 有

$$|a_n(x)|<\frac{\varepsilon}{6M}.$$

现设 $m>n>N$，因为对任意的 $x\in E$ 有

$$\left|\sum_{k=n+1}^{m} b_k(x)\right| = |B_m(x)-B_n(x)| \leqslant 2M,$$

所以对任意的 $x\in E$，由分部求和可得

$$\left|\sum_{k=n+1}^{m} a_k(x)b_k(x)\right| \leqslant 2M\big(|a_{n+1}(x)|+2|a_m(x)|\big) < 2M\Big(\frac{\varepsilon}{6M}+2\cdot\frac{\varepsilon}{6M}\Big)=\varepsilon,$$

再由 Cauchy 收敛准则便知命题成立。 □

**【例 2.18】**若 $\{a_n\}$ 单调趋于 0，则 $\sum\limits_{n=1}^{\infty} a_n \sin nx$ 与 $\sum\limits_{n=1}^{\infty} a_n\cos nx$ 均在 $(0,2\pi)$ 上内闭一致收敛，但它们未必在 $(0,2\pi)$ 上一致收敛。

<u>证明</u>. 为了说明内闭一致收敛性，我们只需证明这两个函数项级数均在形如 $[\delta,2\pi-\delta]$ $(0<\delta<\pi)$ 的区间上一致收敛即可。注意到对任意的 $x\in[\delta,2\pi-\delta]$ 有

$$\begin{aligned}\left|\sum_{k=1}^{n} e^{\mathrm{i}kx}\right| &= \left|\frac{e^{\mathrm{i}x}(1-e^{\mathrm{i}nx})}{1-e^{\mathrm{i}x}}\right| \leqslant \frac{2}{|1-e^{\mathrm{i}x}|} = \frac{2}{\left|e^{\mathrm{i}\frac{x}{2}}-e^{-\mathrm{i}\frac{x}{2}}\right|}\\ &= \frac{1}{\left|\sin\dfrac{x}{2}\right|} \leqslant \frac{1}{\left|\sin\dfrac{\delta}{2}\right|},\end{aligned}$$

取实部和虚部可分别得到

$$\left|\sum_{k=1}^{n} \sin nx\right| \leqslant \frac{1}{\left|\sin\dfrac{\delta}{2}\right|} \qquad 及 \qquad \left|\sum_{k=1}^{n}\cos nx\right| \leqslant \frac{1}{\left|\sin\dfrac{\delta}{2}\right|},$$

于是由 Dirichlet 判别法知命题成立。

此外，这两个函数项级数均未必在 $(0,2\pi)$ 上一致收敛。例如

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n}$$

就不在 $(0,2\pi)$ 上一致收敛，这是因为对任意的 $n\in\mathbb{Z}_{\geqslant 1}$ 有

$$\sup_{x\in(0,2\pi)}\left|\sum_{k=n+1}^{2n}\frac{\sin kx}{k}\right|\geqslant\left|\sum_{k=n+1}^{2n}\frac{\sin k\cdot\frac{1}{2n}}{k}\right|\geqslant\frac{1}{2}\sin\frac{1}{2}.$$

□

最后值得一提的是，函数项级数的绝对收敛性与一致收敛性互相不能蕴含，这是两个互不相干的概念。例如，$\sum\limits_{n=1}^{\infty}x^{n-1}$ 在 $(-1,1)$ 上绝对收敛，但由例 2.10 知它在这一区间上不一致收敛。又如，由例 2.18 知 $\sum\limits_{n=1}^{\infty}\frac{\sin nx}{n}$ 在 $[1,2]$ 上一致收敛，但它在该区间上不绝对收敛。此外，$\sum\limits_{n=1}^{\infty}u_n(x)$ 在某集合上的绝对收敛性和一致收敛性不能保证 $\sum\limits_{n=1}^{\infty}|u_n(x)|$ 在该集合上一致收敛，参见习题 9。

## 习 题 10.2

1. 证明命题 2.4。
2. 证明命题 2.13。
3. 研究下述函数列 $\{f_n(x)\}$ 在指定集合上的一致收敛性及内闭一致收敛性：

   (1) $f_n(x)=\dfrac{1}{x+n}$，$x\in(0,+\infty)$；　(2) $f_n(x)=\dfrac{\sin nx}{n}$，$x\in\mathbb{R}$；

   (3) $f_n(x)=\left(1+\dfrac{x}{n}\right)^n$，$x\in\mathbb{R}$；　(4) $f_n(x)=\dfrac{x^n}{1+x^n}$，$x\in[0,1)$；

   (5) $f_n(x)=\sqrt{x^2+\dfrac{1}{n^2}}$，$x\in\mathbb{R}$；　(6) $f_n(x)=\dfrac{x}{n}\log\dfrac{x}{n}$，$x\in(0,1]$；

   (7) $f_n(x)=\left(\dfrac{\sin x}{x}\right)^n$，$x\in(0,1)$；

   (8) $f_n(x)=\dfrac{\log(1+e^{nx})}{n}$，$x\in(0,+\infty)$；

   (9) $f_n(x)=n(\sqrt[n]{x}-1)$，$x\in[1,+\infty)$。
4. 研究下述函数项级数在指定集合上的一致收敛性及内闭一致收敛性：

   (1) $\displaystyle\sum_{n=1}^{\infty}\frac{n^2x}{1+n^4x^2}$，$x\in(0,1)$；　(2) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{(x+n)(x+n+1)}$，$x\in(0,+\infty)$；

   (3) $\displaystyle\sum_{n=1}^{\infty}\frac{\sin nx}{x^2+n^2}$，$x\in\mathbb{R}$；　(4) $\displaystyle\sum_{n=1}^{\infty}2^n\sin\frac{1}{3^nx}$，$x\in(0,+\infty)$；

(5) $\sum\limits_{n=1}^{\infty} x^2 e^{-nx}$，$x \in (0, +\infty)$；　　(6) $\sum\limits_{n=1}^{\infty} \frac{1}{n^x}$，$x \in (1, +\infty)$；

(7) $\sum\limits_{n=1}^{\infty} \frac{\sin x \, \sin nx}{\sqrt{x+n}}$，$x \in [0, +\infty)$；

(8) $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}+x}\left(1+\frac{1}{n+x}\right)^n$，$x \in (0, +\infty)$。

5. 设 $f$ 是定义在集合 $E$ 上的一个函数，记 $f_n(x) = \frac{[nf(x)]}{n}$，其中 $[t]$ 表示不超过 $t$ 的最大整数。证明函数列 $\{f_n(x)\}$ 在 $E$ 上一致收敛。

6. 设 $\{f_n(x)\}$ 与 $\{g_n(x)\}$ 均在 $E$ 上一致收敛，证明对任意的 $\alpha, \beta \in \mathbb{R}$，函数列 $\{\alpha f_n(x) + \beta g_n(x)\}$ 也在 $E$ 上一致收敛。

7. 设 $\{f_n(x)\}$ 在 $E$ 上一致收敛，且每个 $f_n(x)$ 在 $E$ 上均有界，证明 $\{f_n(x)\}$ 在 $E$ 上一致有界。

8. 设 $\{f_n(x)\}$ 与 $\{g_n(x)\}$ 均在 $E$ 上一致收敛，
   (1) 举例说明 $\{f_n(x)g_n(x)\}$ 未必在 $E$ 上一致收敛；
   (2) 如果对每个给定的 $n$，$f_n(x)$ 与 $g_n(x)$ 均在 $E$ 上有界，那么 $\{f_n(x)g_n(x)\}$ 在 $E$ 上一致收敛。

9. 记 $u_n(x) = (-1)^n(1-x)x^n$，证明 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $[0,1]$ 上绝对收敛且一致收敛，但 $\sum\limits_{n=1}^{\infty} |u_n(x)|$ 在 $[0,1]$ 上不一致收敛。

10. 设 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $[a,b)$ 上一致收敛，且 $\sum\limits_{n=1}^{\infty} u_n(b)$ 收敛，证明 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上一致收敛。

11. 设 $f$ 在 $\mathbb{R}$ 上连续。对任意的正整数 $n$，记
$$f_n(x) = \frac{1}{n}\sum_{i=0}^{n-1} f\left(x+\frac{i}{n}\right).$$
证明 $\{f_n(x)\}$ 在任一有界闭区间上一致收敛。

12. 设级数 $\sum\limits_{n=1}^{\infty} \frac{1}{|a_n|}$ 收敛。证明函数项级数
$$\sum_{n=1}^{\infty} \frac{1}{x-a_n}$$
在不包含 $a_n$ $(n=1,\,2,\,\cdots)$ 的任一有界闭区间上绝对收敛且一致收敛。

13. 设 $\{a_n\}$ 是单调递减的非负数列，证明 $\sum\limits_{n=1}^{\infty} a_n \sin nx$ 在 $\mathbb{R}$ 上一致收敛的充要条件是 $\lim\limits_{n\to\infty} na_n = 0$。

## § 10.3 极限函数的性质

本节的目的是在一定前提下回答 §10.1 中所提出的问题。

**【命题 3.1】** 设函数列 $\{f_n(x)\}$ 在集合 $E$ 上一致收敛于 $f(x)$。又设 $x_0 \in E$ 是 $E$ 的一个聚点，并且对任意给定的正整数 $n$，极限 $\lim\limits_{\substack{x\to x_0\\ x\in E}} f_n(x)$ [2]均存在，那么

$$\lim_{n\to\infty} \lim_{\substack{x\to x_0\\ x\in E}} f_n(x) = \lim_{\substack{x\to x_0\\ x\in E}} f(x) = \lim_{\substack{x\to x_0\\ x\in E}} \lim_{n\to\infty} f_n(x). \tag{10.5}$$

<u>证明</u>. 对任意的 $n$ 记 $\lim\limits_{\substack{x\to x_0\\ x\in E}} f_n(x) = a_n$，我们来证明 $\{a_n\}$ 收敛。因为 $\{f_n(x)\}$ 在 $E$ 上一致收敛，所以对任意的 $\eta > 0$，存在 $M \in \mathbb{Z}_{>0}$，使得对任意的 $m > n > M$ 及 $x \in E$ 有

$$|f_m(x) - f_n(x)| < \eta.$$

在上式中令 $x \to x_0$ 可得 $|a_m - a_n| \leqslant \eta$，进而由 Cauchy 收敛准则知 $\{a_n\}$ 收敛。如果记 $\lim\limits_{n\to\infty} a_n = a$，那么 (10.5) 左边就等于 $a$，因此只需证明

$$\lim_{\substack{x\to x_0\\ x\in E}} f(x) = a.$$

首先，由 $\{f_n(x)\}$ 的一致收敛性知，对任意的 $\varepsilon > 0$，存在 $N_1 \in \mathbb{Z}_{>0}$，使得对任意的 $n > N_1$ 及 $x \in E$ 有

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

其次，由 $\lim\limits_{n\to\infty} a_n = a$ 知存在 $N_2 \in \mathbb{Z}_{>0}$，使得当 $n > N_2$ 时有

$$|a_n - a| < \frac{\varepsilon}{3}.$$

---

[2] 我们用 $\lim\limits_{\substack{x\to x_0\\ x\in E}}$ 表示 $x$ 在 $E$ 中取值且趋于 $x_0$，参见第十二章定义 2.1。

现取 $N > \max(N_1, N_2)$，因为 $\lim\limits_{\substack{x\to x_0\\x\in E}} f_N(x) = a_N$，所以存在 $\delta > 0$，使得

$$|f_N(x) - a_N| < \frac{\varepsilon}{3}, \qquad \forall\, x \in E \cap \Big((x_0-\delta, x_0+\delta) \setminus \{x_0\}\Big).$$

于是对任意的 $x \in E \cap \Big((x_0-\delta, x_0+\delta) \setminus \{x_0\}\Big)$ 有

$$\begin{aligned}|f(x) - a| &\leqslant |f(x) - f_N(x)| + |f_N(x) - a_N| + |a_N - a|\\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

至此命题得证。 □

作为上述命题的一个特殊情形，我们立即得到如下具有广泛应用的结论。

**【命题 3.2】**设 $\{f_n(x)\}$ 在区间 $I$ 上一致收敛于 $f(x)$，且对任意的正整数 $n$，$f_n(x)$ 均在 $I$ 上连续，那么 $f(x)$ 也在 $I$ 上连续。

把上面两个命题中的 $f_n(x)$ 取成函数项级数的部分和就可得到下面两个结论。

**【命题 3.3】**设 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在集合 $E$ 上一致收敛。又设 $x_0 \in E$ 是 $E$ 的一个聚点，并且对任意给定的正整数 $n$，极限 $\lim\limits_{\substack{x\to x_0\\x\in E}} u_n(x)$ 均存在，那么

$$\sum_{n=1}^{\infty}\Big(\lim_{\substack{x\to x_0\\x\in E}} u_n(x)\Big) = \lim_{\substack{x\to x_0\\x\in E}}\Big(\sum_{n=1}^{\infty} u_n(x)\Big).$$

**【命题 3.4】**设 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在区间 $I$ 上一致收敛于 $S(x)$，且对任意的正整数 $n$，$u_n(x)$ 均在 $I$ 上连续，那么 $S(x)$ 也在 $I$ 上连续。

事实上，上述命题的逆命题在一定条件下也是成立的，这就是所谓的 Dini 定理，它可被用于去判定一些特殊的函数项级数的一致收敛性。

**【定理 3.5】（Dini）** 假设对任意的正整数 $n$ 而言，$u_n(x)$ 均在有界闭区间 $I$ 上连续且非负，又设 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的和函数 $S(x)$ 也在 $I$ 上连续，那么 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $I$ 上一致收敛于 $S(x)$。

<u>证明</u>. 反设 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 $I$ 上不一致收敛，则存在 $\varepsilon_0 > 0$，使得对于任意的正整数 $N$，均存在 $n > N$ 及 $x \in I$ 使得

$$|S_n(x) - S(x)| \geqslant \varepsilon_0,$$

其中 $S_n(x)$ 表示 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的部分和。换句话说，存在 $\mathbb{Z}_{>0}$ 的一个子列 $\{n_k\}$ 及包含于 $I$ 内的数列 $\{x_{n_k}\}$ 使得

$$|S_{n_k}(x_{n_k}) - S(x_{n_k})| \geqslant \varepsilon_0, \qquad \forall\ k \geqslant 1. \tag{10.6}$$

按照 Bolzano−Weierstrass 定理，我们不妨假设 $\{x_{n_k}\}$ 收敛于 $a$，而由 $I$ 是闭区间知 $a \in I$。

现任取正整数 $m$，则必存在 $K \in \mathbb{Z}_{>0}$，使得当 $k > K$ 时有 $n_k \geqslant m$，注意到对任意的 $n$ 而言 $u_n(x)$ 均非负，故当 $k > K$ 时有

$$S(x_{n_k}) - S_m(x_{n_k}) \geqslant S(x_{n_k}) - S_{n_k}(x_{n_k}) \geqslant 0,$$

结合 (10.6) 式便得

$$S(x_{n_k}) - S_m(x_{n_k}) \geqslant \varepsilon_0, \qquad \forall\ k > K.$$

现令 $k \to \infty$，由 $S$ 及 $S_m$ 的连续性知

$$S(a) - S_m(a) \geqslant \varepsilon_0.$$

注意到 $m$ 的任意性，故在上式中令 $m \to \infty$ 可得 $0 \geqslant \varepsilon_0$，矛盾。 □

我们再来考虑积分号和微分号同极限号交换的问题。

**【命题 3.6】设 $\{f_n(x)\}$ 在 $[a,b]$ 上一致收敛于 $f(x)$，且对任意的正整数 $n$ 而言 $f_n(x)$ 均在 $[a,b]$ 上连续，则**

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x. \tag{10.7}$$

**此外，变上限积分 $\int_a^x f_n(t)\,\mathrm{d}t$ 也在 $[a,b]$ 上一致收敛于 $\int_a^x f(t)\,\mathrm{d}t$。**

证明. 由命题 3.2 知 $f(x)$ 在 $[a,b]$ 上连续，从而在 $[a,b]$ 上可积。因为 $\{f_n(x)\}$ 一致收敛，故对任意的 $\varepsilon > 0$，存在正整数 $N$，使得对任意的 $n > N$ 及 $x \in [a,b]$ 有

$$|f_n(x) - f(x)| < \varepsilon.$$

于是当 $n > N$ 时有

$$\left|\int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x\right| \leqslant \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x < \varepsilon(b-a),$$

因此 (10.7) 成立。

此外，将上式中的积分上限换为 $x$，那么当 $x \in [a,b]$ 时该式仍成立，这说明 $\int_a^x f_n(t)\,\mathrm{d}t$ 在 $[a,b]$ 上一致收敛于 $\int_a^x f(t)\,\mathrm{d}t$。 □

【注 3.7】上述命题的条件可以被弱化，具体参见习题 6 及其脚注。

【命题 3.8】假设对任意的正整数 $n$，$f_n(x)$ 均在 $[a,b]$ 上连续可导。又设 $\{f_n(x)\}$ 在 $[a,b]$ 上逐点收敛于 $f(x)$，且 $\{f_n'(x)\}$ 在 $[a,b]$ 上一致收敛于 $g(x)$。那么 $f'(x)=g(x)$，也即

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\lim_{n\to\infty}f_n(x)\Big)=\lim_{n\to\infty}\frac{\mathrm{d}}{\mathrm{d}x}f_n(x).$$

此外，$\{f_n(x)\}$ 在 $[a,b]$ 上也一致收敛。

证明. 由命题 3.6 知，对任意的 $x\in[a,b]$ 有

$$\int_a^x g(t)\,\mathrm{d}t=\lim_{n\to\infty}\int_a^x f_n'(t)\,\mathrm{d}t=\lim_{n\to\infty}\big(f_n(x)-f_n(a)\big)=f(x)-f(a),$$

也即是说，

$$f(x)=f(a)+\int_a^x g(t)\,\mathrm{d}t.$$

注意到由命题 3.2 可得 $g(x)$ 在 $[a,b]$ 上连续，故而上式中的变上限积分是 $g(x)$ 的原函数，因此由上式知 $f(x)$ 可导且有 $f'(x)=g(x)$。此外，上式及命题 3.6 也蕴含了 $\{f_n(x)\}$ 在 $[a,b]$ 上一致收敛于 $f(x)$。 □

将上面两个命题中的 $\{f_n(x)\}$ 换成函数项级数的部分和序列，则可得到如下两个结论。

【命题 3.9】设 $\sum\limits_{n=1}^{\infty}u_n(x)$ 在 $[a,b]$ 上一致收敛于 $S(x)$，且对任意的正整数 $n$ 而言 $u_n(x)$ 均在 $[a,b]$ 上连续，则

$$\sum_{n=1}^{\infty}\int_a^b u_n(x)\,\mathrm{d}x=\int_a^b S(x)\,\mathrm{d}x=\int_a^b\Big(\sum_{n=1}^{\infty}u_n(x)\Big)\,\mathrm{d}x,$$

也即是说，可以逐项积分。

【命题 3.10】假设对任意的正整数 $n$，$u_n(x)$ 均在 $[a,b]$ 上连续可导。又设 $\sum\limits_{n=1}^{\infty}u_n(x)$ 在 $[a,b]$ 上逐点收敛于 $S(x)$，且 $\sum\limits_{n=1}^{\infty}u_n'(x)$ 在 $[a,b]$ 上一致收敛于 $g(x)$。那么 $S'(x)=g(x)$，也即

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\sum_{n=1}^{\infty}u_n(x)\Big)=\sum_{n=1}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}u_n(x).$$

也即是说，可以逐项求导。此外，$\sum\limits_{n=1}^{\infty}u_n(x)$ 也在 $[a,b]$ 上一致收敛于 $S(x)$。

【例 3.11】考虑定义在 $(1,+\infty)$ 上的 Riemann $\zeta$ 函数

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}.$$

容易由 Weierstrass 判别法得出 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^s}$ 在 $(1,+\infty)$ 上内闭一致收敛，因此由命题 3.4 知 $\zeta(s)$ 在 $(1,+\infty)$ 的任一闭子区间上连续，进而在 $(1,+\infty)$ 上连续。此外，$\left(\dfrac{1}{n^s}\right)'=-\dfrac{\log n}{n^s}$，且同样由 Weierstrass 判别法可得 $\sum\limits_{n=1}^{\infty}\dfrac{\log n}{n^s}$ 在 $(1,+\infty)$ 上内闭一致收敛，所以由命题 3.8 可推出在 $(1,+\infty)$ 的任一闭子区间上有

$$\zeta'(s)=-\sum_{n=1}^{\infty}\frac{\log n}{n^s},$$

进而可知上式对任意的 $s\in(1,+\infty)$ 成立。类似可对任意的正整数 $k$ 得到

$$\zeta^{(k)}(s)=\sum_{n=1}^{\infty}\frac{(-\log n)^k}{n^s},\qquad \forall\, s\in(1,+\infty).$$

## 习 题 10.3

1. 计算极限：

   (1) $\lim\limits_{x\to 0}\sum\limits_{n=1}^{\infty}\dfrac{\cos nx}{(2+x)^n}$；　　(2) $\lim\limits_{x\to 1}\sum\limits_{n=1}^{\infty}\dfrac{1}{(x+n)(x+n+1)}$；

   (3) $\lim\limits_{n\to\infty}\displaystyle\int_0^1\frac{\sqrt{x}\log n}{e^{nx^2}}\,\mathrm{d}x$；　　(4) $\lim\limits_{n\to\infty}\displaystyle\int_0^{\frac{\pi}{2}}\left(1+\frac{x}{n}\right)^n\sin x\,\mathrm{d}x$。

2. 证明 $\sum\limits_{n=0}^{\infty}\dfrac{x^2}{(1+x^2)^n}$ 在包含原点的任一有限闭区间上不一致收敛。

3. 证明函数项级数 $\sum\limits_{n=0}^{\infty}e^{-\pi n^2 x}$ 在 $(0,+\infty)$ 上收敛，若将其和函数记作 $f(x)$，证明 $f(x)$ 在 $(0,+\infty)$ 上任意阶可导。

4. 证明零阶 Bessel 函数

$$J_0(x)=\sum_{n=0}^{\infty}\frac{(-1)^n}{[(2n)!!]^2}x^{2n}$$

   在 $\mathbb{R}$ 上满足微分方程 $xJ_0''(x)+J_0'(x)+xJ_0(x)=0$。

5. 试举例说明：由在区间 $I$ 上处处不连续的函数所构成的函数列可以一致收敛于某个在 $I$ 上连续的函数。

6. 设函数列 $\{f_n(x)\}$ 在区间 $I$ 上一致收敛于 $f(x)$，并且每个 $f_n(x)$ 均在 $I$ 上一致连续，证明 $f(x)$ 也在 $I$ 上一致连续。

7. 设 $\{f_n(x)\}$ 在 $[a,b]$ 上一致收敛于 $f(x)$，且对任意的正整数 $n$ 而言 $f_n(x)$ 均在 $[a,b]$ 上可积。证明 $f(x)$ 在 $[a,b]$ 上可积，且有
$$\lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$ [3]

8. 把 $(0,1)$ 内的全体有理数排成一列，记作 $r_1, r_2, \cdots, r_n, \cdots$，并令
$$f(x) = \sum_{n=1}^{\infty}\frac{|x-r_n|}{2^n}, \qquad \forall\ x\in(0,1).$$
证明 $f(x)$ 在 $(0,1)$ 上连续，并且在无理点处可导、在有理点处不可导。

## § 10.4 幂级数

幂级数 (power series) 是形如
$$\sum_{n=0}^{\infty} a_n(x-x_0)^n$$
的函数项级数，其部分和是关于 $x-x_0$ 的多项式。当 $a_n = 0\ (\forall\ n > N)$ 时上述级数就退化成了一个多项式，因此幂级数可以被看作是多项式的推广。

幂级数虽然只是一种特殊的函数项级数，但是其简洁明了的形式以及非常良好的性质使得它成为了极其重要的一类函数项级数，它在微积分学的形成和发展过程中起到了不可替代的作用。

作变量替换 $x-x_0 \longmapsto x$ 可将上述幂级数化为
$$\sum_{n=0}^{\infty} a_n x^n$$
的形式，因此我们后面的讨论主要围绕这种特殊类型的幂级数展开。

首先需要解决的是幂级数的敛散性问题。

---

[3] 这个命题的条件可被进一步弱化，事实上 C. Arzela 于 1885 年证明了如下结论：设 $\{f_n(x)\}$ 在 $[a,b]$ 上一致有界且逐点收敛于 $f(x)$，若每个 $f_n(x)$ 以及 $f(x)$ 均在 $[a,b]$ 上可积，那么
$$\lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

**【定理 4.1】（Abel 第一定理）** 若 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $x=\xi\ (\xi \neq 0)$ 处收敛，则其在 $(-|\xi|, |\xi|)$ 上绝对收敛；若 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $x=\eta$ 处发散，则其在 $|x|>|\eta|$ 上发散。

<u>证明</u>. 由 $\sum\limits_{n=0}^{\infty} a_n \xi^n$ 收敛知 $\lim\limits_{n\to\infty} a_n\xi^n = 0$，因此 $a_n\xi^n = O(1)$，于是当 $|x|<|\xi|$ 时有

$$|a_n x^n| = |a_n \xi|^n \cdot \left|\frac{x}{\xi}\right|^n = O\left(\left|\frac{x}{\xi}\right|^n\right).$$

注意到此时 $\sum\limits_{n=0}^{\infty} \left|\frac{x}{\xi}\right|^n$ 收敛，所以 $\sum\limits_{n=0}^{\infty} a_n x^n$ 绝对收敛。

另一方面，若存在 $x_0$ 使得 $|x_0|>|\eta|$ 且 $\sum\limits_{n=0}^{\infty} a_n x_0^n$ 收敛，则由上一段刚证明的结论知 $\sum\limits_{n=0}^{\infty} a_n \eta^n$ 绝对收敛，这与假设矛盾。 □

现在我们假设存在 $\xi \neq 0$ 使得 $\sum\limits_{n=0}^{\infty} a_n \xi^n$ 收敛，那么 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $\mathbb{R}$ 上的收敛集 $E$ 就不会只包含 0。如果记

$$R = \sup E,$$

则幂级数 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $(-R, R)$ 上绝对收敛，当 $|x|>R$ 时发散。事实上，一方面，当 $|x|<R$ 时，由上确界的定义知存在 $y \in E$ 使得 $|x|<y<R$，因为 $y\in E$，所以 $\sum\limits_{n=0}^{\infty} a_n y^n$ 收敛，再由定理 4.1 可得 $\sum\limits_{n=0}^{\infty} a_n x^n$ 绝对收敛；另一方面，当 $R \neq +\infty$ 且 $|x|>R$ 时，$\frac{|x|+R}{2} \notin E$，从而 $\sum\limits_{n=0}^{\infty} a_n \left(\frac{|x|+R}{2}\right)^n$ 发散，再由定理 4.1 便得 $\sum\limits_{n=0}^{\infty} a_n x^n$ 发散。

通过变量替换 $x \longmapsto x - x_0$ 便知 $\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ 在 $(x_0-R, x_0+R)$ 上绝对收敛，在 $|x-x_0|>R$ 上发散。因此，$\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ 的收敛集必是一个区间（这个区间也可能退化为单元素集 $\{x_0\}$），我们称之为<u>收敛区间 (convergence interval)</u>，并称 $R$ 为 $\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ 的<u>收敛半径 (radius of convergence)</u>。

**【例 4.2】** 考虑幂级数 $\sum\limits_{n=1}^{\infty} \frac{x^n}{n}$，由于它在 $x=-1$ 处收敛，故而在 $(-1,1)$ 上绝对收敛；又由于它在 $x=1$ 处发散，从而在 $|x|>1$ 上发散。所以该幂级数的收敛区间为 $[-1,1)$，收敛半径为 1。

显然，一旦确定了幂级数 $\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$ 的收敛半径 $R$，我们就知道它在 $(x_0-R, x_0+R)$ 上绝对收敛，当 $|x-x_0|>R$ 时发散，因此只需进一步考察该级数

在 $x=x_0-R$ 及 $x=x_0+R$ 处的敛散性就能确定出收敛区间，所以现在最迫切的问题是如何方便地去计算幂级数的收敛半径。

**【定理 4.3】（Cauchy–Hadamard）** 幂级数 $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 的收敛半径为

$$R=\frac{1}{\varlimsup\limits_{n\to\infty}\sqrt[n]{|a_n|}},$$

其中，当 $\varlimsup\limits_{n\to\infty}\sqrt[n]{|a_n|}=0$ 时取 $R=+\infty$，当 $\varlimsup\limits_{n\to\infty}\sqrt[n]{|a_n|}=+\infty$ 时取 $R=0$。

<u>证明</u>. 记 $\ell=\varlimsup\limits_{n\to\infty}\sqrt[n]{|a_n|}$，则

$$\varlimsup_{n\to\infty}\sqrt[n]{|a_n(x-x_0)^n|}=|x-x_0|\cdot\ell.$$

按照 Cauchy 判别法（参见第九章定理 2.9），$\sum\limits_{n=0}^{\infty}|a_n(x-x_0)^n|$ 在上式 $<1$ 时收敛，在上式 $>1$ 时发散。由于当 $\ell=0$ 时这一上极限总等于 0，而当 $\ell=+\infty$ 时该上极限小于 1 当且仅当 $x=x_0$，所以它们分别对应于 $R=+\infty$ 和 $R=0$。当 $\ell\in(0,+\infty)$ 时我们有

$$\{x:|x-x_0|\cdot\ell<1\}=(x_0-R,x_0+R)=\{x:|x-x_0|<R\},$$

故而 $R=\dfrac{1}{\ell}$。 □

**【例 4.4】** 由 $\lim\limits_{n\to\infty}\sqrt[n]{n^2}=1$ 知，幂级数 $\sum\limits_{n=1}^{\infty}n^2x^n$ 的收敛半径为 1，又因为这一幂级数在 $x=1$ 和 $x=-1$ 时均发散，所以其收敛区间为 $(-1,1)$。

**【例 4.5】** 求幂级数 $\sum\limits_{n=1}^{\infty}\dfrac{(2+(-1)^n)^n}{n^2}x^n$ 的收敛半径和收敛区间。

<u>解</u>. 利用 Cauchy–Hadamard 公式可求得收敛半径

$$R=\left(\varlimsup_{n\to\infty}\frac{2+(-1)^n}{\sqrt[n]{n^2}}\right)^{-1}=\frac{1}{3}.$$

当 $x=\pm\dfrac{1}{3}$ 时

$$\left|\frac{(2+(-1)^n)^n}{n^2}x^n\right|\leqslant\frac{1}{n^2},$$

故而该级数在 $x=\pm\dfrac{1}{3}$ 处绝对收敛，从而其收敛区间为 $\left[-\dfrac{1}{3},\dfrac{1}{3}\right]$。 □

**【注 4.6】**当极限 $\lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|$ 存在时，由 (9.6) 式知其值与 $\lim\limits_{n\to\infty}\sqrt[n]{|a_n|}$ 相同，因此在这种情况下收敛半径

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|.$$

**【例 4.7】**设 $\alpha\in\mathbb{R}\setminus\mathbb{Z}_{\geqslant 0}$。对幂级数 $\sum\limits_{n=0}^{\infty}\dbinom{\alpha}{n}x^n$ 而言，由于

$$\lim_{n\to\infty}\left|\frac{\binom{\alpha}{n}}{\binom{\alpha}{n+1}}\right|=\lim_{n\to\infty}\left|\frac{n+1}{\alpha-n}\right|=1,$$

故其收敛半径为 1，且在 $(-1,1)$ 上绝对收敛。

下面来讨论幂级数的一致收敛性。

**【定理 4.8】（Abel 第二定理）** 设 $\sum\limits_{n=0}^{\infty}a_n(x-x_0)^n$ 的收敛半径为 $R$，那么

(1) 该级数在 $(x_0-R,x_0+R)$ 上内闭一致收敛。

(2) 若这一级数在 $x=x_0+R$ 处收敛，则它在 $(x_0-R,x_0+R]$ 上内闭一致收敛。特别地，

$$\sum_{n=0}^{\infty}a_nR^n=\lim_{x\to R^-}\left(\sum_{n=0}^{\infty}a_nx^n\right).$$

(3) 若这一级数在 $x=x_0-R$ 处收敛，则它在 $[x_0-R,x_0+R)$ 上内闭一致收敛。特别地，

$$\sum_{n=0}^{\infty}a_n(-R)^n=\lim_{x\to -R^+}\left(\sum_{n=0}^{\infty}a_nx^n\right).$$

证明. 利用变量替换 $x-x_0\longmapsto x$ 可将讨论限于 $x_0=0$ 的情形。

(1) 对任意的 $[a,b]\subseteq(-R,R)$，记 $c=\max(|a|,|b|)$，则 $\sum\limits_{n=0}^{\infty}a_nc^n$ 绝对收敛。因为对任意的 $x\in[a,b]$ 有 $|a_nx^n|\leqslant|a_nc^n|$，故由 Weierstrass 判别法知 $\sum\limits_{n=0}^{\infty}a_nx^n$ 在 $[a,b]$ 上一致收敛。

(2) 利用 (1) 的结论，我们只需证 $\sum\limits_{n=0}^{\infty}a_nx^n$ 在 $[0,R]$ 上一致收敛即可。注意到

$$a_nx^n=\frac{x^n}{R^n}\cdot a_nR^n,$$

其中 $\left\{\dfrac{x^n}{R^n}\right\}$ 在 $[0,R]$ 上一致有界，且对任意的 $x\in[0,R]$，$\left\{\dfrac{x^n}{R^n}\right\}$ 单调，故由 Abel 判别法知 $\sum\limits_{n=0}^{\infty}a_nx^n$ 在 $[0,R]$ 上一致收敛。

此外，由内闭一致收敛性及命题 3.4 知 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $(-R,R]$ 上连续，特别地，在 $x_0+R$ 处的连续性意味着

$$\sum_{n=0}^{\infty} a_n R^n = \lim_{x\to R^-}\left(\sum_{n=0}^{\infty} a_n x^n\right).$$

类似可证明 (3)。 □

【**注 4.9**】一般来说，Abel 第二定理的逆命题不成立，也即是由极限

$$\lim_{x\to R^-}\left(\sum_{n=0}^{\infty} a_n x^n\right)$$

的存在性无法推出该极限值等于 $\sum\limits_{n=0}^{\infty} a_n R^n$，事实上我们甚至不能保证这个级数的收敛性。但是如果给 $a_n$ 再添加一些条件，那么是可以证明 Abel 第二定理的逆命题成立的，具体参见习题 5，6 和 7。

将定理 4.8 与上节的命题 3.4 及命题 3.9 结合知，$\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 在 $(x_0-R,x_0+R)$ 上连续并且对任意的 $[a,b]\subseteq(x_0-R,x_0+R)$ 有

$$\int_a^b\left(\sum_{n=0}^{\infty} a_n(x-x_0)^n\right)\mathrm{d}x = \sum_{n=0}^{\infty} a_n\int_a^b (x-x_0)^n\,\mathrm{d}x.$$

此外，

$$\frac{\mathrm{d}}{\mathrm{d}x}\big[a_n(x-x_0)^n\big] = na_n(x-x_0)^{n-1},$$

而 $\overline{\lim\limits_{n\to\infty}}\sqrt[n]{n|a_n|} = \overline{\lim\limits_{n\to\infty}}\sqrt[n]{|a_n|}$，因此 $\sum\limits_{n=0}^{\infty} na_n(x-x_0)^{n-1}$ 与 $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 有相同的收敛半径，进而知 $\sum\limits_{n=0}^{\infty} na_n(x-x_0)^{n-1}$ 也在 $(x_0-R,x_0+R)$ 上内闭一致收敛，于是可对 $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 逐项求导。类似可进一步得知，若 $S(x)=\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$，则 $S(x)$ 在 $(x_0-R,x_0+R)$ 上任意阶可导，且

$$S^{(k)}(x) = \sum_{n=k}^{\infty}\frac{n!}{(n-k)!}a_n(x-x_0)^{n-k},\qquad \forall\, x\in(x_0-R,x_0+R). \tag{10.8}$$

**【例 4.10】** 对任意的 $x \in (-1,1)$ 有

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n,$$

对任意的 $x \in (-1,1)$，通过逐项积分可得

$$\log(1+x) = \int_0^x \frac{1}{1+t}\,\mathrm{d}t = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n.$$

注意到上式右边级数在 $x=1$ 处收敛，故由定理 4.8 (2) 知

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2.$$

**【例 4.11】** 计算 $\sum\limits_{n=1}^{\infty} \dfrac{n}{2^n}$。

<u>解</u>. 考虑幂级数 $\sum\limits_{n=0}^{\infty} x^n$，其收敛半径为 1，且在 $(-1,1)$ 上有

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

逐项求导可得

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}, \qquad \forall\, x \in (-1,1),$$

两边同乘 $x$ 便得

$$\sum_{n=1}^{\infty} n x^{n} = \frac{x}{(1-x)^2}, \qquad \forall\, x \in (-1,1).$$

再将 $x = \dfrac{1}{2}$ 代入即知 $\sum\limits_{n=1}^{\infty} \dfrac{n}{2^n} = 2$。 □

在本节结束前，我们应用定理 4.8 来证明如下关于 Cauchy 乘积的结论。

**【定理 4.12】（Abel）** 设 $\sum\limits_{n=0}^{\infty} c_n$ 是 $\sum\limits_{n=0}^{\infty} a_n$ 和 $\sum\limits_{n=0}^{\infty} b_n$ 的 Cauchy 乘积，且这三个级数均收敛，则有

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big). \tag{10.9}$$

<u>证明</u>. 对 $x \in (-1,1]$ 记

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad g(x) = \sum_{n=0}^{\infty} b_n x^n, \qquad h(x) = \sum_{n=0}^{\infty} c_n x^n,$$

这一定义的合理性可由级数 $\sum\limits_{n=0}^{\infty} a_n$，$\sum\limits_{n=0}^{\infty} b_n$ 及 $\sum\limits_{n=0}^{\infty} c_n$ 的收敛性得到。按照 Cauchy 乘积的定义，

$$h(x) = f(x)g(x).$$

由定理 4.8，$f(x)$，$g(x)$ 以及 $h(x)$ 均在 $x=1$ 处左连续，因此在上式中令 $x \to 1^-$ 即得 (10.9)。 □

## 习 题 10.4

1. 求下列函数项级数的收敛区间：

(1) $\displaystyle\sum_{n=1}^{\infty} \frac{3^n + (-2)^n}{n} x^n$；　　(2) $\displaystyle\sum_{n=1}^{\infty} \Big(1+\frac{1}{n}\Big)^{n^2} x^n$；

(3) $\displaystyle\sum_{n=1}^{\infty} \Big(1+\frac{1}{2}+\cdots+\frac{1}{n}\Big) x^n$；　　(4) $\displaystyle\sum_{n=1}^{\infty} \frac{2-\cos n}{n} x^n$；

(5) $\displaystyle\sum_{n=1}^{\infty} \frac{x^{n^2}}{2^n}$；　　(6) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n+1}\Big(\frac{1-x}{1+x}\Big)^n$。

2. 在下列级数的收敛区间内求它们的和：

(1) $\displaystyle\sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}$；　　(2) $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$；

(3) $\displaystyle\sum_{n=0}^{\infty} (-1)^n n^2 x^n$；　　(4) $\displaystyle\sum_{n=1}^{\infty} n(n+2) x^n$。

3. 求幂级数 $1+\displaystyle\sum_{n=1}^{\infty} \frac{(2n-1)!!}{(2n)!!} x^n$ 的收敛区间以及它的和。

4. 试求幂级数 $\sum\limits_{n=0}^{\infty} \dbinom{\alpha}{n} x^n$ 的收敛区间。

5. 令 $f(x) = \begin{cases} \dfrac{\log(1+x)}{x}, & x>0, \\ 1, & x=0. \end{cases}$ 证明 $\displaystyle\int_0^1 f(x)\,\mathrm{d}x = \frac{\zeta(2)}{2}$。

6. 设幂级数 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $(-1,1)$ 上收敛，证明对任意的 $x \in (-1,1)$ 有

$$\sum_{n=0}^{\infty} a_n x^n = \frac{a_0}{1-x} + \frac{x}{1-x}\sum_{n=0}^{\infty} (a_{n+1}-a_n) x^n.$$

7. 证明当 $x \to 1^-$ 时有

$$\sum_{n=0}^{\infty}(\log n)x^n \sim \frac{1}{1-x}\log\frac{1}{1-x}.$$

8. 举例说明极限 $\lim\limits_{x\to R^-}\left(\sum\limits_{n=0}^{\infty} a_n x^n\right)$ 的存在性无法保证级数 $\sum\limits_{n=0}^{\infty} a_n R^n$ 收敛。

9. 设幂级数 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $(-1,1)$ 上收敛，且 $a_n \geqslant 0$ $(\forall\ n \geqslant 1)$。又设

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty} a_n x^n = S,$$

证明 $\sum\limits_{n=0}^{\infty} a_n$ 收敛于 $S$。

10. （Tauber 定理）设幂级数 $\sum\limits_{n=0}^{\infty} a_n x^n$ 在 $(-1,1)$ 上收敛于 $f(x)$，且 $\lim\limits_{n\to\infty} na_n = 0$。如果 $\lim\limits_{x\to 1^-}\left(\sum\limits_{n=0}^{\infty} a_n x^n\right) = S$，证明 $\sum\limits_{n=0}^{\infty} a_n$ 收敛于 $S$。

**从第 11 题到第 15 题是一组题，讨论 Dirichlet 级数的敛散性。**

11. 证明存在唯一的 $\sigma_a$（可以是实数，$+\infty$ 或 $-\infty$），使得级数 $\sum\limits_{n=1}^{\infty}\left|\frac{a_n}{n^s}\right|$ 当 $s > \sigma_a$ 时收敛，当 $s < \sigma_a$ 时发散。我们称 $\sigma_a$ 为 Dirichlet 级数 $\sum\limits_{n=1}^{\infty}\frac{a_n}{n^s}$ 的<u>绝对收敛横标 (abscissa of absolute convergence)</u>。

12. 证明存在唯一的 $\sigma_c$（可以是实数，$+\infty$ 或 $-\infty$），使得级数 $\sum\limits_{n=1}^{\infty}\frac{a_n}{n^s}$ 当 $s > \sigma_c$ 时收敛，当 $s < \sigma_c$ 时发散。$\sigma_c$ 被称为 Dirichlet 级数 $\sum\limits_{n=1}^{\infty}\frac{a_n}{n^s}$ 的<u>收敛横标 (abscissa of convergence)</u>。

13. 设 $\sigma_a, \sigma_c \in \mathbb{R}$，证明 $0 \leqslant \sigma_a - \sigma_c \leqslant 1$，并举例说明这一不等式不能被改进。

14. 假设级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散，并记 $S_n = \sum\limits_{k=1}^{n} a_k$。证明

$$\sigma_c = \overline{\lim_{n\to\infty}}\frac{\log|S_n|}{\log n}.$$

15. 假设级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛，并记 $r_n = \sum\limits_{k=n+1}^{\infty} a_k$。证明

$$\sigma_c = \overline{\lim_{n\to\infty}}\frac{\log|r_n|}{\log n}.$$

## § 10.5
## 函数的幂级数展开

多项式是我们认识到的最简单的一类初等函数，作为其推广，幂级数不仅形式简单，而且具有良好的性质，因此我们自然希望把一般的函数与幂级数联系起来，这样就能通过研究幂级数得到该函数的相应性质。

**【定义 5.1】**设函数 $f(x)$ 在点 $x_0$ 的某邻域 $U$ 内有定义。如果存在幂级数 $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 在 $U$ 上收敛于 $f(x)$，则称 $f(x)$ 在 $x_0$ 处解析 (analytic)。

设 $E \subseteq \mathbb{R}$，如果 $f(x)$ 是定义在 $E$ 上的一个函数，且 $f(x)$ 在 $E$ 上每一点处均解析，则称 $f(x)$ 是 $E$ 上的解析函数 (analytic function)。

我们首先来看看什么样的函数能写成幂级数，以及这样的幂级数应该具有什么样的形式。假设在 $(x_0-\delta, x_0+\delta)$ 上有

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n, \tag{10.10}$$

那么由上节的讨论知 $f(x)$ 在 $(x_0-\delta, x_0+\delta)$ 上任意阶可导，且通过逐项积分可得（参见 (10.8) 式）

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n(x-x_0)^{n-k}, \qquad \forall\, x \in (x_0-\delta, x_0+\delta).$$

在上式中取 $x = x_0$ 即得

$$a_k = \frac{1}{k!} f^{(k)}(x_0).$$

因此，若 $f(x)$ 能在 $x_0$ 的某邻域内展开成幂级数，那么它必在 $x_0$ 的某邻域内任意阶可导，且其所对应的幂级数必然是

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

我们把这一级数称为 $f(x)$ 在 $x_0$ 处的 Taylor 级数 (Taylor series) 或 Taylor 展式 (Taylor expansion)，当 $x_0 = 0$ 时也称其为 $f(x)$ 的 Maclaurin 级数 (Maclaurin series) 或 Maclaurin 展式 (Maclaurin expansion)。由上面的讨论亦可看出，若 $f$ 在 $x_0$ 处的 Taylor 级数存在，则必唯一。我们把这些结论总结如下。

**【命题 5.2】如果函数 $f(x)$ 在 $x_0$ 处解析，则其必在 $x_0$ 的某邻域内任意阶可导，且它的各阶导数均在 $x_0$ 处解析。此外，$f(x)$ 在 $x_0$ 处有唯一的 Taylor 展式。**

注意到上述推导是在 (10.10) 成立的条件下进行的，因此现在我们需要去研究反方向的问题：如果 $f(x)$ 在 $x_0$ 的某邻域内任意阶可导，那么 $f(x)$ 是否在 $x_0$ 处解析？换句话说，是否在该邻域内有

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \tag{10.11}$$

成立呢？这个问题的答案是否定的，一个经典的例子是

$$f(x)=\begin{cases}e^{-\frac{1}{x^2}}, & 若\ x\neq 0,\\ 0, & 若\ x=0.\end{cases}$$

由 §5.3 习题 7 知它在 $\mathbb{R}$ 上任意阶可导且 $f^{(n)}(0)=0$ $(\forall\ n\in\mathbb{Z}_{\geqslant 0})$，因而 $f(x)$ 的 Maclaurin 级数为 0，于是 (10.11) 不成立。所以我们需要去讨论在什么情况下 $f(x)$ 才在 $x_0$ 处解析。

(10.11) 式右侧幂级数的形状让我们联想到 Taylor 公式，因此一个很自然的想法是利用 Taylor 公式作为工具来进行讨论。

现设 $f(x)$ 在 $(x_0-\delta,x_0+\delta)$ 上任意阶可导，那么 (10.11) 成立当且仅当

$$\lim_{N\to\infty}\left(f(x)-\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\right)=0.$$

若记

$$f(x)=\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+r_N(x),$$

则 (10.11) 成立的充要条件是对任意的 $x\in(x_0-\delta,x_0+\delta)$ 均有 $\lim\limits_{N\to\infty}r_N(x)=0$。所以我们需要对余项 $r_N(x)$ 有所了解。回忆起在 §6.2 中我们给出了 Lagrange 余项

$$r_N(x)=\frac{f^{(N+1)}(\xi)}{(N+1)!}(x-x_0)^{N+1}$$

和 Cauchy 余项

$$r_N(x)=\frac{f^{(N+1)}(\eta)}{N!}(x-\eta)^N(x-x_0),$$

其中 $\xi$ 和 $\eta$ 位于 $x$ 和 $x_0$ 之间；在 §8.5 中我们给出了积分余项

$$r_N(x)=\frac{1}{N!}\int_{x_0}^{x}f^{(N+1)}(t)(x-t)^N\,\mathrm{d}t.$$

下面我们来对一些特殊函数研究满足 $\lim\limits_{N\to\infty}r_N(x)=0$ 的 $x$ 的范围。

**【例 5.3】** $e^x$ 的 Maclaurin 展式。

由第六章例 2.14 知，存在位于 0 与 $x$ 之间的 $\xi$ 使得

$$e^x = \sum_{n=0}^{N} \frac{x^n}{n!} + \frac{e^\xi}{(N+1)!} x^{N+1}.$$

一方面，幂级数 $\sum\limits_{n=0}^{\infty} \dfrac{x^n}{n!}$ 的收敛区间为 $\mathbb{R}$；另一方面，

$$|r_N(x)| = \left| \frac{e^\xi}{(N+1)!} x^{N+1} \right| \leqslant \max(1, e^x) \cdot \frac{|x|^{N+1}}{(N+1)!},$$

然而 $\lim\limits_{n\to\infty} \dfrac{|x|^{N+1}}{(N+1)!} = 0$（参见第三章例 3.3），故 $\lim\limits_{N\to\infty} r_N(x) = 0$。所以有

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad \forall\, x \in \mathbb{R}.$$

**【例 5.4】** $\sin x$ 和 $\cos x$ 的 Maclaurin 展式。

因为

$$(\sin x)^{(n)} = \sin\Big(x + \frac{n\pi}{2}\Big), \qquad (\cos x)^{(n)} = \cos\Big(x + \frac{n\pi}{2}\Big),$$

故由带 Lagrange 余项的 Taylor 公式知，存在位于 0 与 $x$ 之间的 $\xi$ 及 $\eta$ 使得

$$\sin x = \sum_{n=0}^{N} \frac{(-1)^n}{(2n+1)!} x^{2n+1} + \frac{x^{2N+3}}{(2N+3)!} \sin\Big(\xi + \frac{2N+3}{2}\pi\Big)$$

及

$$\cos x = \sum_{n=0}^{N} \frac{(-1)^n}{(2n)!} x^{2n} + \frac{x^{2N+2}}{(2N+2)!} \cos\Big(\eta + \frac{2N+2}{2}\pi\Big).$$

于是类似于例 5.3 可得

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad \text{及} \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \qquad \forall\, x \in \mathbb{R}.$$

**【例 5.5】** $\log(1+x)$ 的 Maclaurin 展式。

在上节例 4.10 中我们已经看到了

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n, \qquad \forall\, x \in (-1, 1].$$

【例 5.6】$\arctan x$ 的 Maclaurin 展式。

对 $\dfrac{1}{1+x^2}=\sum\limits_{n=0}^{\infty}(-1)^n x^{2n}$ $(-1<x<1)$ 逐项积分可得

$$\arctan x=\int_0^x \frac{1}{1+t^2}\,\mathrm{d}t=\sum_{n=0}^{\infty}\int_0^x (-1)^n t^{2n}\,\mathrm{d}t=\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}x^{2n+1}.$$

注意到上式右边级数在 $x=\pm 1$ 处收敛，故由定理 4.8 知

$$\arctan x=\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}x^{2n+1},\qquad \forall\ x\in[-1,1].$$

特别地，取 $x=1$ 可得 Leibniz 公式

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots+\frac{(-1)^n}{2n+1}+\cdots.$$

【例 5.7】$(1+x)^{\alpha}$ $(\alpha\in\mathbb{R}\setminus\mathbb{Z}_{\geqslant 0})$ 的 Maclaurin 展式。

由带积分余项的 Taylor 公式知

$$(1+x)^{\alpha}=\sum_{n=0}^{N}\binom{\alpha}{n}x^n+\frac{\alpha(\alpha-1)\cdots(\alpha-N)}{N!}\int_0^x (1+t)^{\alpha-N-1}(x-t)^N\,\mathrm{d}t.$$

一方面，由上节例 4.7 知幂级数 $\sum\limits_{n=0}^{\infty}\dbinom{\alpha}{n}x^n$ 在 $(-1,1)$ 上绝对收敛。另一方面，对于 $x\in(-1,1)$，容易验证当 $t$ 位于 $0$ 与 $x$ 之间时有 $\left|\dfrac{x-t}{1+t}\right|\leqslant|x|$，因此

$$\begin{aligned}
|r_N(x)|&=\left|\frac{\alpha(\alpha-1)\cdots(\alpha-N)}{N!}\int_0^x (1+t)^{\alpha-1}\Big(\frac{x-t}{1+t}\Big)^N\,\mathrm{d}t\right|\\
&\leqslant\left|\frac{\alpha(\alpha-1)\cdots(\alpha-N)}{N!}x^N\int_0^x (1+t)^{\alpha-1}\,\mathrm{d}t\right|\\
&=\left|\binom{\alpha-1}{N}x^N\cdot\big[(1+x)^{\alpha}-1\big]\right|.
\end{aligned}$$

现记

$$y_n=\left|\binom{\alpha-1}{n}x^n\right|,$$

则

$$\frac{y_{n+1}}{y_n}=\left|\frac{\alpha-1-n}{n+1}\cdot x\right|,$$

因此 $\lim\limits_{n\to\infty}\dfrac{y_{n+1}}{y_n}=|x|\in[0,1)$，这说明数列 $\{y_n\}$ 自某项开始单调递减，从而收敛。若设 $\lim\limits_{n\to\infty}y_n=h$，则由上式知 $h=|x|h$，从而 $h=0$。于是 $\lim\limits_{N\to\infty}r_N(x)=0$。这样我们就得到了

$$(1+x)^\alpha=\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n,\qquad \forall\, x\in(-1,1).$$

## 习 题 10.5

1. 设 $f(x)$ 在 $(-R,R)$ 上有幂级数展式 $f(x)=\sum\limits_{n=0}^{\infty}a_nx^n$，证明：

   (1) 若 $f$ 是奇函数，则对任意的非负整数 $n$ 有 $a_{2n}=0$；

   (2) 若 $f$ 是偶函数，则对任意的非负整数 $n$ 有 $a_{2n+1}=0$。

2. 设 $a$ 是一个实数，证明 $f(x)=\dfrac{1}{x-a}$ 在 $\mathbb{R}\setminus\{a\}$ 上解析。

3. 证明对任意的 $x\in\mathbb{R}$ 有（我们默认 $\dfrac{\sin t}{t}$ 在 $t=0$ 处取值为 1）

$$\int_0^x\frac{\sin t}{t}\,\mathrm{d}t=\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!(2n+1)}x^{2n+1}.$$

4. 求 $f(x)=\arcsin x$ 的 Maclaurin 展式，并对正整数 $n$ 求 $f^{(n)}(0)$。

5. 求 $f(x)=\log(x+\sqrt{1+x^2})$ 的 Maclaurin 展式。

6. 求使得 $\dfrac{1}{2}(e^x+e^{-x})\leqslant e^{\lambda x^2}$ 对任意的实数 $x$ 均成立的 $\lambda$ 的值。

我们可将

$$\sin x=\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n+1}\quad 及\quad \cos x=\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}x^{2n}\tag{10.12}$$

作为正弦函数和余弦函数的严格定义，并由此出发得到那些我们所熟知的与之相关的结论。从第 7 题到第 17 题是一组题，让我们来逐步完成以上工作。

7. 证明按上述方式定义的 $\sin x$ 与 $\cos x$ 是在 $\mathbb{R}$ 上任意阶可导的函数，且满足 $(\sin x)'=\cos x$，$(\cos x)'=-\sin x$。

8. 证明存在 $x_0>0$，使得 $\cos x$ 在 $[0,x_0]$ 上取值恒大于 0。

9. 证明 $\cos x$ 在 $(0,+\infty)$ 上有零点，且 $\cos x$ 有最小正零点，我们将 $\cos x$ 的最小正零点记作 $\dfrac{\pi}{2}$。[4]

---

[4] 这是 $\pi$ 的一种定义方式。

10. 对任意的 $x \in \mathbb{R}$ 证明 $\cos^2 x + \sin^2 x = 1$，进而得出 $\sin \dfrac{\pi}{2} = 1$。
11. 证明 $\cos x$ 和 $\sin x$ 分别在 $\left[0, \dfrac{\pi}{2}\right]$ 上严格单调递减和严格单调递增。
12. 对任意的 $x, y \in \mathbb{R}$ 证明
$$\sin(x+y) = \sin x \cos y + \sin y \cos x,$$
$$\cos(x+y) = \cos x \cos y - \sin x \sin y.$$
13. 证明 $\cos x$ 与 $\sin x$ 均是以 $2\pi$ 为最小正周期的周期函数。
14. 证明 $\varphi : \theta \longmapsto (\cos\theta, \sin\theta)$ 是从 $[0, 2\pi)$ 到单位圆 $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ 的双射。
15. 提前使用弧长的计算公式 (16.5) 证明单位圆的周长为 $2\pi$，这意味着第 9 题中所给出的 $\pi$ 的定义符合我们在小学所给出的几何定义。
16. 设 $\varphi$ 如第 14 题所定义，对于单位圆上任意一点 $M(x,y)$，我们称 $\varphi^{-1}((x,y))$ 为射线 $\overrightarrow{OM}$ 与 $x$ 轴正向的<u>夹角</u>。证明 $\cos\varphi^{-1}((x,y)) = x$，$\sin\varphi^{-1}((x,y)) = y$。这说明由 (10.12) 所定义的正弦函数及余弦函数符合我们在中学所给出的几何定义。
17. 证明在 $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$ 上有 $|\sin x| \leqslant |x| \leqslant \dfrac{|\sin x|}{|\cos x|}$，并且等号成立均当且仅当 $x = 0$。

## §10.6 幂级数的运算

为了方便起见，在本节中我们只针对形如
$$\sum_{n=0}^{\infty} a_n x^n$$
的幂级数来进行讨论。

假设 $f(x)$ 与 $g(x)$ 均在原点解析，从而在某个公共区间上可展开为 Maclaurin 级数
$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad g(x) = \sum_{n=0}^{\infty} b_n x^n, \tag{10.13}$$
我们需要去了解 $f \pm g$，$f \cdot g$，$f/g$ 及 $g \circ f$ 是否也都在原点解析。首先来看前两种情形。

【命题 6.1】假设在 $(-R,R)$ 上有 (10.13) 成立，则在 $(-R,R)$ 上有

$$f(x)\pm g(x)=\sum_{n=0}^{\infty}(a_n\pm b_n)x^n$$

$$f(x)g(x)=\sum_{n=0}^{\infty}c_nx^n,$$

其中 $c_n=\sum\limits_{k+\ell=n}a_kb_\ell$，且上两式右边幂级数的收敛半径均 $\geqslant R$。

证明. 第一个式子是第九章命题 1.4 的直接推论，第二个式子可由 $\sum\limits_{n=0}^{\infty}a_nx^n$ 和 $\sum\limits_{n=0}^{\infty}b_nx^n$ 在 $(-R,R)$ 上的绝对收敛性及第九章定理 5.3 得到。此外，这两个式子右边的幂级数均在 $(-R,R)$ 上绝对收敛，故它们的收敛半径均 $\geqslant R$。 □

【注 6.2】在上述命题的条件下，利用归纳法知，对任意的正整数 $k$ 有

$$f(x)^k=\sum_{n=0}^{\infty}c_n^{(k)}x^n,\qquad \forall\, x\in(-R,R),$$

其中

$$c_n^{(k)}=\sum_{n_1+\cdots+n_k=n}a_{n_1}\cdots a_{n_k}. \tag{10.14}$$

下面我们要讨论复合运算，在此之前先证明一个关于交换求和号的结论。

【引理 6.3】假设

$$\sum_{j=1}^{\infty}|a_{ij}|=b_i,\qquad \forall\, i\in\mathbb{Z}_{>0},$$

且 $\sum\limits_{i=1}^{\infty}b_i$ 收敛，则

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}.$$

证明. 令 $E=\left\{\dfrac{1}{n}:n\in\mathbb{Z}_{>0}\right\}\cup\{0\}$，并在 $E$ 上定义函数 $f_i$ 及 $g$ 如下：对任意的 $i,n\in\mathbb{Z}_{>0}$，

$$f_i(0)=\sum_{j=1}^{\infty}a_{ij},\qquad f_i\Big(\frac{1}{n}\Big)=\sum_{j=1}^{n}a_{ij},$$

于是对任意的 $i$ 有 $\lim\limits_{n\to\infty} f_i\Big(\dfrac{1}{n}\Big) = f(0)$。由于对任意的 $x \in E$ 而言 $|f_i(x)| \leqslant b_i$，故由 Weierstrass 判别法知 $\sum\limits_{i=1}^{\infty} f_i(x)$ 在 $E$ 上一致收敛，再利用命题 3.3 可得

$$\begin{aligned}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} f_i(0) = \sum_{i=1}^{\infty}\lim_{n\to\infty} f_i\Big(\frac{1}{n}\Big) &= \lim_{n\to\infty}\sum_{i=1}^{\infty} f_i\Big(\frac{1}{n}\Big)\\ &= \lim_{n\to\infty}\sum_{i=1}^{\infty}\sum_{j=1}^{n} a_{ij} = \lim_{n\to\infty}\sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij}\\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}.\end{aligned}$$

□

【**命题 6.4**】假设在 $(-R,R)$ 上有 $g(x) = \sum\limits_{n=0}^{\infty} b_n x^n$，又设 $f(0)=0$ 且 $f(x)$ 在 0 的某邻域内可展开成 Maclaurin 级数

$$f(x) = \sum_{n=1}^{\infty} a_n x^n.$$

那么，必存在 $r>0$ 使得 $g(f(x))$ 在 $(-r,r)$ 上可展开成 Maclaurin 级数

$$g(f(x)) = b_0 + \sum_{n=1}^{\infty}\bigg(\sum_{k=1}^{n} b_k c_n^{(k)}\bigg)x^n, \tag{10.15}$$

其中 $c_n^{(k)}$ 如 (10.14) 所定义，且上式右侧幂级数的收敛半径 $\geqslant r$。

<u>证明</u>. 由定理 4.1 知幂级数 $\sum\limits_{n=1}^{\infty} |a_n x^n|$ 在 0 的某邻域内收敛，再由定理 4.8 知 $\sum\limits_{n=1}^{\infty} |a_n x^n|$ 在这一邻域内内闭一致收敛，从而其和函数在这一邻域内连续，于是存在 $r>0$ 使得

$$\sum_{n=1}^{\infty} |a_n x^n| < R, \qquad \forall\, x \in (-r,r). \tag{10.16}$$

因此当 $x \in (-r,r)$ 时，由注 6.2 可得

$$\begin{aligned} g(f(x)) = \sum_{k=0}^{\infty} b_k (f(x))^k &= b_0 + \sum_{k=1}^{\infty} b_k \bigg(\sum_{n=1}^{\infty} a_n x^n\bigg)^k\\ &= b_0 + \sum_{k=1}^{\infty} b_k \sum_{n=1}^{\infty} c_n^{(k)} x^n. \end{aligned} \tag{10.17}$$

现记 $y=\sum\limits_{n=1}^{\infty}|a_nx^n|$，那么

$$\begin{aligned}\sum_{n=1}^{\infty}|c_n^{(k)}x^n| &= \sum_{n=1}^{\infty}\left|\sum_{n_1+\cdots+n_k=n}a_{n_1}\cdots a_{n_k}\right|\cdot|x|^n\\ &\leqslant \sum_{n=1}^{\infty}\left(\sum_{n_1+\cdots+n_k=n}|a_{n_1}\cdots a_{n_k}|\right)\cdot|x|^n\\ &= \left(\sum_{n=1}^{\infty}|a_nx^n|\right)^k = y^k.\end{aligned}$$

而由 (10.16) 知 $y\in[0,R)$，因此级数 $\sum\limits_{k=1}^{\infty}|b_k|y^k$ 收敛，于是由引理 6.3 知 (10.17) 右边求和号可以交换，即

$$g(f(x)) = b_0+\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}b_kc_n^{(k)}x^n = b_0+\sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty}b_kc_n^{(k)}\right)x^n.$$

再注意到当 $k>n$ 时 $c_n^{(k)}=0$，从而 (10.15) 得证。 □

**【例 6.5】** 在 $(-1,1)\setminus\{0\}$ 上有

$$\begin{aligned}(1+x)^{\frac{1}{x}} &= \exp\left(\frac{1}{x}\log(1+x)\right) = \exp\left(1-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\cdots\right)\\ &= e\cdot\left[1+\left(-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\cdots\right)+\frac{1}{2!}\left(-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\cdots\right)^2\right.\\ &\qquad\left.+\frac{1}{3!}\left(-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\cdots\right)^3+\cdots\right]\\ &= e\left(1-\frac{1}{2}x+\frac{11}{24}x^2-\frac{7}{16}x^3+\cdots\right).\end{aligned}$$

最后我们再来研究除法的情形，因为已经在命题 6.1 中讨论了乘法，所以我们只需说明清楚倒数运算即可，而这其实是复合运算的一个简单应用。

假设在 $(-R,R)$ 上有 $f(x)=\sum\limits_{n=0}^{\infty}a_nx^n$ 且 $a_0\neq 0$，于是由 $f$ 在 $(-R,R)$ 上的连续性知存在 $r>0$，使得对任意的 $x\in(-r,r)$ 有 $|f(x)-f(0)|<|a_0|$，也即 $|f(x)-a_0|<|a_0|$。因此若记 $\widetilde{f}(x)=a_0^{-1}(f(x)-a_0)$，则有

$$\left|\widetilde{f}(x)\right| = \left|\sum_{n=1}^{\infty}\frac{a_n}{a_0}x^n\right| < 1, \qquad \forall\, x\in(-r,r).$$

注意到如果令 $g(y)=\dfrac{1}{1+y}$ $(-1<y<1)$，那么

$$\frac{1}{f(x)}=\frac{1}{a_0}\cdot\frac{1}{1+a_0^{-1}(f(x)-a_0)}=\frac{1}{a_0}\cdot\frac{1}{1+\widetilde{f}(x)}=\frac{1}{a_0}\cdot g\big(\widetilde{f}(x)\big).$$

当 $x\in(-r,r)$ 时，命题 6.4 保证了上式右边的 $g\big(\widetilde{f}(x)\big)$ 可展开成 Maclaurin 级数。

**【例 6.6】** 求 $\sec x$ 的 Maclaurin 展式至 $x^6$ 项。

解. 在 $x=0$ 的附近有

$$\begin{aligned}\sec x&=\frac{1}{1-(1-\cos x)}=\sum_{n=0}^{\infty}(1-\cos x)^n\\&=1+\left(\frac{x^2}{2}-\frac{x^4}{24}+\frac{x^6}{720}+\cdots\right)+\left(\frac{x^2}{2}-\frac{x^4}{24}+\frac{x^6}{720}+\cdots\right)^2\\&\quad+\left(\frac{x^2}{2}-\frac{x^4}{24}+\frac{x^6}{720}+\cdots\right)^3+\cdots\\&=1+\left(\frac{x^2}{2}-\frac{x^4}{24}+\frac{x^6}{720}+\cdots\right)+\left(\frac{x^4}{4}-\frac{x^6}{24}+\cdots\right)+\left(\frac{x^6}{8}+\cdots\right)+\cdots\\&=1+\frac{1}{2}x^2+\frac{5}{24}x^4+\frac{61}{720}x^6+\cdots.\end{aligned}$$

□

## 习 题 10.6

1. 求下列函数的 Maclaurin 展式：

(1) $e^{-x^2}$；　　(2) $\sin^2 x$；

(3) $\dfrac{x^3}{(1-x)^2}$；　　(4) $\dfrac{12-5x}{6-5x-x^2}$。

2. 求下列函数的 Maclaurin 展式至少到 $x^6$ 项：

(1) $x^2\sqrt{1-x}+\dfrac{1}{1-x}$；　　(2) $e^x\log(1-x)$；

(3) $\log\cos x$。

3. 设 $s>2$，证明

$$\sum_{n=1}^{\infty}\left(\zeta(s)-\frac{1}{1^s}-\frac{1}{2^s}-\cdots-\frac{1}{n^s}\right)=\zeta(s-1)-\zeta(s).$$

4. 证明 $\sum_{n=2}^{\infty}(-1)^n\frac{\zeta(n)}{n}=\gamma$，其中 $\gamma$ 是 Euler 常数。

5. 设 $\frac{1}{1-x-x^2}$ 的 Maclaurin 展式为 $\sum_{n=0}^{\infty}a_nx^n$，证明

$$\sum_{n=0}^{\infty}\frac{a_{n+1}}{a_n\,a_{n+2}}=2.$$

6. 设 $B_n$ 由下式所定义：

$$\frac{x}{e^x-1}=\sum_{n=0}^{\infty}\frac{B_n}{n!}x^n,\qquad x\neq 0$$

我们称 $B_n$ 为 Bernoulli 数 (Bernoulli number)。

(1) 证明：当 $n\geqslant 2$ 时有 $B_n=\sum_{j=0}^{n}\binom{n}{j}B_j$；

(2) 计算 $B_n$ $(0\leqslant n\leqslant 10)$ 的值；

(3) 证明对任意的正整数 $n$ 有 $B_{2n+1}=0$。

## § 10.7 Weierstrass 逼近定理

在 §10.5 中我们看到了可将满足一定条件的函数展开成幂级数，然而这儿所需要的条件是极强的，即需要该函数在给定的区间上任意阶可导，但一般的连续函数并不具有这一性质，所以我们只能退而求其次，希望能用多项式去逼近连续函数。

**【定理 7.1】（Weierstrass）** 设 $f(x)$ 在有界闭区间 $[a,b]$ 上连续，则存在多项式列 $\{P_n(x)\}$，它在 $[a,b]$ 上一致收敛于 $f(x)$。

证明.（Бернштейн）这是一个构造性的证明方法。首先讨论 $[a,b]=[0,1]$ 的情形。对正整数 $n$ 定义

$$B_n(f,x)=\sum_{j=0}^{n}\binom{n}{j}f\Big(\frac{j}{n}\Big)x^j(1-x)^{n-j},\qquad \forall\, x\in[0,1],\tag{10.18}$$

这个多项式被称作 $n$ 阶 Бернштейн 多项式 (Bernstein polynomial)，它是由 С. Н. Бернштейн[5] 于 1912 年引入的。下面证明 $\{B_n(f,x)\}$ 在 $[0,1]$ 上一致收敛于 $f(x)$。

由 $f$ 在 $[0,1]$ 上的连续性可得其在该区间上的有界性和一致连续性，从而一方面存在 $M>0$ 使得对任意的 $x\in[0,1]$ 有 $|f(x)|\leqslant M$；另一方面，对任意的 $\varepsilon>0$，

存在 $\delta > 0$，使得对于 $[0,1]$ 中满足 $|x_1 - x_2| < \delta$ 的任意两点 $x_1$ 与 $x_2$ 均有

$$|f(x_1) - f(x_2)| < \varepsilon.$$

现设 $x \in [0,1]$，利用二项式定理可得

$$\begin{aligned}|B_n(f,x) - f(x)| &= \left|\sum_{j=0}^{n}\binom{n}{j}f\left(\frac{j}{n}\right)x^j(1-x)^{n-j} - f(x)\sum_{j=0}^{n}\binom{n}{j}x^j(1-x)^{n-j}\right| \\ &\leqslant \sum_{j=0}^{n}\left|f\left(\frac{j}{n}\right) - f(x)\right|\binom{n}{j}x^j(1-x)^{n-j} \\ &= \sum\nolimits_1 + \sum\nolimits_2\end{aligned}$$

其中 $\sum_1$ 表示对满足 $\left|\dfrac{j}{n} - x\right| < \delta$ 的 $j$ 求和，$\sum_2$ 表示对满足 $\left|\dfrac{j}{n} - x\right| \geqslant \delta$ 的 $j$ 求和。于是容易得出

$$\sum\nolimits_1 < \varepsilon\sum_{j=0}^{n}\binom{n}{j}x^j(1-x)^{n-j} = \varepsilon.$$

为了估计 $\sum_2$，我们需要解除求和条件中的不等式 $\left|\dfrac{j}{n} - x\right| \geqslant \delta$ 对求和变量的限制，注意到 $\left|\dfrac{j-nx}{n\delta}\right| \geqslant 1$，所以可以对求和的项乘以 $\left(\dfrac{j-nx}{n\delta}\right)^2$ 并将求和条件中的上述不等式删去[⑤]，也即

$$\begin{aligned}\sum\nolimits_2 &= \sum_{|\frac{j}{n}-x|\geqslant\delta}\left|f\left(\frac{j}{n}\right) - f(x)\right|\binom{n}{j}x^j(1-x)^{n-j} \\ &\leqslant \sum_{j=0}^{n}\left(\frac{j-nx}{n\delta}\right)^2\left|f\left(\frac{j}{n}\right) - f(x)\right|\binom{n}{j}x^j(1-x)^{n-j} \\ &\leqslant \frac{2M}{n^2\delta^2}\sum_{j=0}^{n}(j-nx)^2\binom{n}{j}x^j(1-x)^{n-j}.\end{aligned}$$

可以证明（留作练习）

$$\begin{cases}\displaystyle\sum_{j=0}^{n} j\binom{n}{j}x^j(1-x)^{n-j} = nx, \\ \displaystyle\sum_{j=0}^{n} j^2\binom{n}{j}x^j(1-x)^{n-j} = n(n-1)x^2 + nx.\end{cases} \tag{10.19}$$

---

⑤这是一个常用的技术性手段，它被称为 Rankin 的技巧 (Rankin's trick)。

于是

$$\sum\nolimits_2 \leqslant \frac{2M}{n^2\delta^2}\big[(n(n-1)x^2+nx)-2nx\cdot nx+n^2x^2\big]=\frac{2M}{n\delta^2}x(1-x)\leqslant\frac{M}{2n\delta^2},$$

上面最后一步用到了算术平均—几何平均不等式。综上便得

$$|B_n(f,x)-f(x)|<\varepsilon+\frac{M}{2n\delta^2},$$

因此当 $n>\dfrac{M}{2\varepsilon\delta^2}$ 时便可对任意的 $x\in[0,1]$ 得到 $|B_n(f,x)-f(x)|<2\varepsilon$。这就证明了 $\{B_n(f,x)\}$ 在 $[0,1]$ 上一致收敛于 $f(x)$。

对于一般的区间 $[a,b]$，我们记

$$g(x)=f(a+(b-a)x),$$

则当 $f\in C([a,b])$ 时 $g\in C([0,1])$。因此 $\{B_n(g,x)\}$ 在 $[0,1]$ 上一致收敛于 $g(x)$，从而 $\left\{B_n\Big(g,\dfrac{x-a}{b-a}\Big)\right\}$ 在 $[a,b]$ 上一致收敛于 $g\Big(\dfrac{x-a}{b-a}\Big)=f(x)$。 □

## 习 题 10.7

1. 证明 (10.19)。
2. 求函数 $f(x)=x^2$ $(0\leqslant x\leqslant 1)$ 的 $n$ 阶 Бернштейн 多项式。
3. 求函数 $f(x)=e^x$ $(0\leqslant x\leqslant 1)$ 的 $n$ 阶 Бернштейн 多项式。
4. 设 $f\in C([a,b])$ 且对任意的正整数 $n$ 有

$$\int_a^b f(x)x^n\,\mathrm{d}x=0.$$

证明 $f$ 在 $[a,b]$ 上恒等于 0。

5. 设多项式序列 $\{P_n(x)\}$ 在 $\mathbb{R}$ 上一致收敛于 $f(x)$，证明 $f(x)$ 是多项式。
6. 设 $0<a<b<1$ 且 $f\in C([a,b])$。证明存在在 $[a,b]$ 上一致收敛于 $f$ 的整系数多项式序列。

**从第 7 题到第 9 题是一组题，给出了与 Бернштейн 多项式相关的一些结果。**

7. 对任意的 $n\in\mathbb{Z}_{\geqslant 1}$ 及 $k\in\mathbb{Z}_{\geqslant 0}$ 记

$$P_{n,k}(x)=\sum_{j=0}^{n}(j-nx)^k\binom{n}{j}x^j(1-x)^{n-j},\qquad x\in[0,1].$$

证明

$$P_{n,k+2}(x)=x(1-x)\big[P'_{n,k+1}(x)+n(k+1)P_{n,k}(x)\big].$$

8. 设 $P_{n,k}(x)$ 如上题所设，证明

$$P_{n,k}(x) \ll n^{[\frac{k}{2}]}, \qquad \forall\, x \in [0,1],$$

其中 $\ll$ 常数仅与 $k$ 有关。

9. 设 $f$ 在 $[0,1]$ 上有界，$x_0 \in (0,1)$，且 $f$ 在 $x_0$ 处二阶可导，证明

$$B_n(f,x_0) = f(x_0) + \frac{f''(x_0)}{2n} x_0(1-x_0) + o\Big(\frac{1}{n}\Big).$$

## § 10.8 应用

在本节中，我们将通过两个例子来展示函数项级数理论的应用。

**【例 8.1】**首先来介绍 $\Gamma$ 函数。该函数有多种引入方法，但我更喜欢用无穷乘积来定义它，因为这种方式能让人看清楚问题的起源。

$\Gamma$ 函数是由 L. Euler 在 1729 年 10 月 13 日写给 C. Goldbach 的一封信中提出的，这解决了 D. Bernoulli 和 Goldbach 所提出的问题：能否把在正整数集上定义的函数 $f(n) = n!$ 延拓成为正实数集上的一个连续函数？Euler 的做法是这样的，先假设 $x$ 与 $n$ 均为正整数，那么

$$x!(x+1)\cdots(x+n) = (x+n)! = n!(n+1)\cdots(n+x),$$

因此

$$\begin{aligned} x! &= \frac{n!(n+1)\cdots(n+x)}{(x+1)\cdots(x+n)} = (n+1)\cdots(n+x)\prod_{k=1}^{n}\Big(1+\frac{x}{k}\Big)^{-1} \\ &= \frac{(n+1)\cdots(n+x)}{(n+1)^x}\cdot\prod_{k=1}^{n}\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}. \end{aligned} \tag{10.20}$$

一方面，$\displaystyle\lim_{n\to\infty}\frac{(n+1)\cdots(n+x)}{(n+1)^x} = 1$；另一方面，无穷乘积

$$\prod_{k=1}^{\infty}\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}$$

对任意不等于负整数的 $x\in\mathbb{R}$ 均收敛，这是因为当 $k\to\infty$ 时有

$$\begin{aligned}\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}&=\bigg(1-\frac{x}{k}+O\Big(\frac{1}{k^2}\Big)\bigg)\bigg(1+\frac{x}{k}+O\Big(\frac{1}{k^2}\Big)\bigg)\\&=1+O\Big(\frac{1}{k^2}\Big).\end{aligned}\tag{10.21}$$

所以我们有理由相信函数

$$\Pi(x)=\prod_{k=1}^{\infty}\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}$$

应该满足我们的要求，下面就来证明这件事。

首先，当 $x$ 是正整数时，由 (10.20) 知

$$\prod_{k=1}^{n}\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}=\frac{(n+1)^x}{(n+1)\cdots(n+x)}\cdot x!,$$

令 $n\to\infty$ 便得 $\Pi(x)=x!$。事实上容易对任意的 $x\notin\mathbb{Z}_{<0}$ 证明

$$\Pi(x+1)=(x+1)\Pi(x).$$

其次，我们来证明 $\Pi(x)$ 在 $\mathbb{R}\setminus\mathbb{Z}_{<0}$ 上连续。为此，只需说明 $\Pi(x)$ 在任意一个不包含负整数的区间 $[a,b]$ 上连续即可。由 (10.21) 知

$$\log\Big(1+\frac{x}{k}\Big)^{-1}\Big(1+\frac{1}{k}\Big)^{x}=O\Big(\frac{1}{k^2}\Big),$$

因此级数 $\sum\limits_{n=1}^{\infty}\log\Big(1+\dfrac{x}{n}\Big)^{-1}\Big(1+\dfrac{1}{n}\Big)^{x}$ 在 $[a,b]$ 上一致收敛，从而由它所定义的函数 $\log\Pi(x)$ 在 $[a,b]$ 上连续，进而可得 $\Pi(x)$ 在 $[a,b]$ 上连续。

我们记 $\Gamma(x)=\Pi(x-1)=x^{-1}\Pi(x)$，并称之为 <u>$\Gamma$ 函数 (Gamma function)</u>，因此

$$\frac{1}{\Gamma(x)}=x\prod_{n=1}^{\infty}\Big(1+\frac{x}{n}\Big)\Big(1+\frac{1}{n}\Big)^{-x},\qquad\forall\ x\notin\mathbb{Z}_{\leqslant 0}.\tag{10.22}$$

**【例 8.2】**利用函数项级数我们可以构造出在 $\mathbb{R}$ 上连续但是处处不可导的函数。

在历史上，人们曾经在几何直观的基础上认为连续性在某种程度上蕴含可微性，例如企图论证连续函数在除去一些“例外点”后可导，直到 K. Weierstrass 建

立了处处不可导的连续函数的例子⑥才彻底断绝了这类想法。Weierstrass 所给出的例子是

$$f(x)=\sum_{n=0}^{\infty} b^n \cos(a^n \pi x),$$

其中 $0<b<1$ 且 $a$ 是满足 $ab>1+\dfrac{3}{2}\pi$ 的奇数。随后，人们构造了更多这样的例子，其中最为简单的或许当属由 B. L. van der Waerden[48] 于 1930 年所构造的如下函数：

$$f(x)=\sum_{n=0}^{\infty} \frac{\|10^n x\|}{10^n}, \tag{10.23}$$

其中 $\|t\|=\min\limits_{n\in\mathbb{Z}}|t-n|$ 表示 $t$ 与离它最近的整数之间的距离。

易见 $0\leqslant\|t\|\leqslant\dfrac{1}{2}$，故 (10.23) 中的级数以 $\sum\limits_{n=0}^{\infty}\dfrac{1}{10^n}$ 为优势级数，从而由 Weierstrass 判别法知它在 $\mathbb{R}$ 上一致收敛。于是由函数 $\|t\|$ 的连续性知 $f(x)$ 在 $\mathbb{R}$ 上连续。

下面来证明 $f(x)$ 在 $\mathbb{R}$ 上处处不可导。因为函数 $\|t\|$ 以 1 为周期，故只需证明 $f(x)$ 在 $[0,1)$ 上处处不可导即可。任取 $x_0\in[0,1)$，我们将其写成

$$x_0=0.a_1a_2\cdots a_n\cdots$$

的十进制小数形式。若 $x_0$ 是有限小数，就在其小数表示的尾部添加一串 0。按照定义

$$\|10^n x_0\|=\begin{cases}0.a_{n+1}a_{n+2}\cdots, & \text{若 } 0.a_{n+1}a_{n+2}\cdots\leqslant\dfrac{1}{2},\\[2mm] 1-0.a_{n+1}a_{n+2}\cdots, & \text{若 } 0.a_{n+1}a_{n+2}\cdots>\dfrac{1}{2}.\end{cases}$$

现对任意的正整数 $m$ 定义 $h_m$ 如下：

$$h_m=\begin{cases}-10^{-m}, & \text{若 } a_m=4 \text{ 或 } 9,\\ 10^{-m}, & \text{其他情形}.\end{cases} \tag{10.24}$$

则 $\lim\limits_{m\to\infty}h_m=0$，并且

$$\frac{f(x_0+h_m)-f(x_0)}{h_m}=\sum_{n=0}^{\infty}\frac{1}{h_m}\left(\frac{\|10^n(x_0+h_m)\|}{10^n}-\frac{\|10^n x_0\|}{10^n}\right)$$

⑥这个例子由 Weierstrass 在 1874 年写信告知 P. Du Bois–Reymond，并由后者于 1875 年发表[16]。

$$= \sum_{n=0}^{\infty} \pm 10^m \cdot \frac{\|10^n(x_0 \pm 10^{-m})\| - \|10^n x_0\|}{10^n}, \tag{10.25}$$

其中正负号分别对应于 (10.24) 式中的两种情况。容易验证，

$$\|10^n(x_0 \pm 10^{-m})\| - \|10^n x_0\| = \begin{cases} 0, & \text{当 } n \geqslant m \text{ 时}, \\ 10^{n-m} \text{ 或 } -10^{n-m}, & \text{当 } 0 \leqslant n < m \text{ 时}. \end{cases}$$

因此 (10.25) 中的级数的项当 $n \geqslant m$ 时为 0，当 $0 \leqslant n < m$ 时为 $\pm 1$，从而

$$\frac{f(x_0 + h_m) - f(x_0)}{h_m}$$

是一个整数，且与 $m$ 具有相同的奇偶性。这说明数列 $\left\{\dfrac{f(x_0 + h_m) - f(x_0)}{h_m}\right\}$ 是由奇偶相间的整数组成，它当然不可能收敛。这就证明了 $f(x)$ 在 $x_0$ 处不可导。

## 习 题 10.8

1. 证明对任意的 $x \notin \mathbb{Z}_{<0}$ 有 $\Pi(x+1) = (x+1)\Pi(x)$，进而对 $x \notin \mathbb{Z}_{\leqslant 0}$ 得到 $\Gamma(x+1) = x\Gamma(x)$。

2. 设 $\alpha > 0$，证明对任意的 $x \in (-1, 1)$ 有

$$\frac{1}{(1-x)^\alpha} = \sum_{n=0}^{\infty} \frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)} x^n.$$

3. 设 $x \notin \mathbb{Z}_{\leqslant 0}$，证明

$$\frac{1}{\Gamma(x)} = x e^{\gamma x} \prod_{n=1}^{\infty} \Big(1 + \frac{x}{n}\Big) e^{-\frac{x}{n}},$$

   其中 $\gamma$ 是 Euler 常数。

4. 证明在 $\mathbb{R}_{>0}$ 上有

$$-\frac{\Gamma'(x)}{\Gamma(x)} = \frac{1}{x} + \gamma - \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+x}\right),$$

   并由此得到 $\Gamma'(1) = -\gamma$。

5. 证明 $\log \Gamma(x)$ 是 $\mathbb{R}_{>0}$ 上的凸函数。

6. 设 $x \notin \mathbb{Z}_{\leqslant 0}$，证明 Legendre 的<u>二倍公式 (duplication formula)</u>

$$\Gamma\Big(\frac{1}{2}\Big)\Gamma(2x) = 2^{2x-1}\Gamma(x)\Gamma\Big(x + \frac{1}{2}\Big).$$

7. 提前使用 §18.2 习题 8 证明余元公式 (reflection formula)

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}, \qquad \forall\, x \notin \mathbb{Z}.$$

8. 证明 van der Waerden 所构造的连续函数 (10.23) 在 $\mathbb{R}$ 的任一子区间上均不单调。

利用函数项级数我们可以构造出一条覆盖整个正方形的连续曲线，因为这种类型的曲线最早是由 G. Peano 于 1890 年提出的，所以也被称为 Peano 曲线 (Peano curve)。从第 9 题到 11 题是一组题，介绍由 I. J. Schoenberg[42] 于 1938 年所给出的 Peano 曲线的例子。

9. 设 $f$ 是定义在 $\mathbb{R}$ 上的以 2 为周期的偶函数，且

$$f(t) = \begin{cases} 0, & 若\ t \in \left[0, \dfrac{1}{3}\right), \\ 3t-1, & 若\ t \in \left[\dfrac{1}{3}, \dfrac{2}{3}\right), \\ 1, & 若\ t \in \left[\dfrac{2}{3}, 1\right]. \end{cases}$$

对任意的 $t \in [0,1]$，令

$$x(t) = \sum_{n=1}^{\infty} \frac{f(3^{2n-2}t)}{2^n}, \qquad y(t) = \sum_{n=1}^{\infty} \frac{f(3^{2n-1}t)}{2^n}.$$

证明由参数方程 $t \longmapsto (x(t), y(t))$ 所给出的曲线 $C$ 是位于正方形 $[0,1]^2$ 内的连续曲线。

10. 对任意的点 $(a,b) \in [0,1]^2$，设 $a$ 与 $b$ 的二进制表示分别为

$$a = \sum_{n=1}^{\infty} \frac{a_n}{2^n}, \qquad b = \sum_{n=1}^{\infty} \frac{b_n}{2^n},$$

其中 $a_n, b_n \in \{0,1\}$ $(\forall\, n \geqslant 1)$。证明

$$t_0 = \frac{2a_1}{3} + \frac{2b_1}{3^2} + \frac{2a_2}{3^3} + \frac{2b_2}{3^4} + \cdots + \frac{2a_n}{3^{2n-1}} + \frac{2b_n}{3^{2n}} + \cdots$$

属于 $[0,1]$。

11. 沿用上题中的记号，对任意的 $n \geqslant 1$ 证明 $f(3^{2n-2}t_0) = a_n$ 与 $f(3^{2n-1}t_0) = b_n$，进而得到 $x(t_0) = a$，$y(t_0) = b$。这样就说明了第 9 题中定义的曲线 $C$ 覆盖了正方形 $[0,1]^2$。

附录 常用函数的幂级数展开式表

| | |
|---|---|
| $e^x$ | $\sum\limits_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}$ |
| $\sin x$ | $\sum\limits_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad x \in \mathbb{R}$ |
| $\cos x$ | $\sum\limits_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad x \in \mathbb{R}$ |
| $\log(1+x)$ | $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n, \quad x \in (-1, 1]$ |
| $\arcsin x$ | $\sum\limits_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!\,(2n+1)} x^{2n+1}, \quad x \in (-1, 1)$ |
| $\arctan x$ | $\sum\limits_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}, \quad x \in [-1, 1]$ |
| $\frac{1}{1-x}$ | $\sum\limits_{n=0}^{\infty} x^n, \quad x \in (-1, 1)$ |
| $\frac{1}{1+x}$ | $\sum\limits_{n=0}^{\infty} (-1)^n x^n, \quad x \in (-1, 1)$ |
| $(1+x)^\alpha,\ \alpha \in \mathbb{R} \setminus \mathbb{Z}_{\geqslant 0}$ | $\sum\limits_{n=0}^{\infty} \binom{\alpha}{n} x^n, \quad x \in (-1, 1)$ |

第十一章

# 广义积分

数学家是这样一种人，公式

$$\int_{-\infty}^{+\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$$

对他们来说就如同二二得四对你来说一样明显。Liouville 是一个数学家。

—— Lord Kelvin

在第八章中，我们介绍了定义在有界闭区间上的有界函数 $f(x)$ 在该区间上的 Riemann 积分。在本章中我们将在两个方面对此进行拓展，一是考虑无界区间上的积分，二是研究有界区间上的无界函数的积分，这两类积分被统称为广义积分 (improper integral)[①]。在本章的讨论中我们将会看到，广义积分与无穷级数有着密切的联系。

## § 11.1 无界区间上的积分

【**定义 1.1**】设 $f(x)$ 在 $[a,+\infty)$ 上有定义，并且对任意的 $A>a$ 而言 $f(x)$ 在 $[a,A]$ 上可积。若极限

$$\lim_{A\to+\infty}\int_a^A f(x)\,\mathrm{d}x$$

存在，则称 $f(x)$ 在 $[a,+\infty)$ 上可积，并称广义积分 $\displaystyle\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛于上述极限值，记作

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{A\to+\infty}\int_a^A f(x)\,\mathrm{d}x.$$

① 也被称作反常积分或瑕积分。

否则就称 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 发散。

类似可以定义广义积分 $\int_{-\infty}^a f(x)\,\mathrm{d}x$ 的收敛与发散。若对任意的 $a \in \mathbb{R}$ 而言 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 与 $\int_{-\infty}^a f(x)\,\mathrm{d}x$ 均收敛，则称广义积分 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ 收敛，并记

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{a} f(x)\,\mathrm{d}x + \int_{a}^{+\infty} f(x)\,\mathrm{d}x.$$[2]

**【例 1.2】** 设 $a, p \in \mathbb{R}$，$a > 0$，我们来研究广义积分 $\int_a^{+\infty} \frac{\mathrm{d}x}{x^p}$ 的敛散性。由于对任意的 $A > a$ 有

$$\int_a^A \frac{\mathrm{d}x}{x^p} = \begin{cases} \dfrac{A^{1-p} - a^{1-p}}{1-p}, & \text{若 } p \neq 1, \\ \log A - \log a, & \text{若 } p = 1, \end{cases}$$

故而

$$\lim_{A \to +\infty} \int_a^A \frac{\mathrm{d}x}{x^p} = \begin{cases} \dfrac{a^{1-p}}{p-1}, & \text{若 } p > 1, \\ +\infty, & \text{若 } p \leqslant 1. \end{cases}$$

因此 $\int_a^{+\infty} \frac{\mathrm{d}x}{x^p}$ 当 $p > 1$ 时收敛，当 $p \leqslant 1$ 时发散。

**【例 1.3】** 由于对任意的 $A > a$ 有

$$\int_a^A \sin x\,\mathrm{d}x = \cos a - \cos A,$$

所以极限 $\lim\limits_{A \to +\infty} \int_a^A \sin x\,\mathrm{d}x$ 不存在，从而 $\int_a^{+\infty} \sin x\,\mathrm{d}x$ 发散。当然 $\int_{-\infty}^{+\infty} \sin x\,\mathrm{d}x$ 也发散。

我们接下来的讨论均是针对形如 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 的广义积分而言的，对于形如 $\int_{-\infty}^a f(x)\,\mathrm{d}x$ 与 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ 的广义积分也有类似结果。为方便起见，如果没有专门提及，我们均假设相应的函数在 $[a, A]$ $(\forall\, A > a)$ 上可积。

---

[2] 容易证明，这一定义与 $a$ 的选择无关。

**【命题 1.4】**设 $f(x)$ 与 $g(x)$ 均在 $[a,+\infty)$ 上可积，则

(1) 对任意的 $b>a$，$f(x)$ 在 $[b,+\infty)$ 上可积；

(2) 对任意的实数 $\alpha$ 和 $\beta$，函数 $\alpha f(x)+\beta g(x)$ 在 $[a,+\infty)$ 上可积，且

$$\int_a^{+\infty}\big[\alpha f(x)+\beta g(x)\big]\,\mathrm{d}x=\alpha\int_a^{+\infty}f(x)\,\mathrm{d}x+\beta\int_a^{+\infty}g(x)\,\mathrm{d}x.$$

<u>证明</u>. 这可直接由定义得到。 □

**【定理 1.5】(Cauchy 收敛准则)** $\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 收敛的充要条件是：对任意的 $\varepsilon>0$，均存在 $A>0$，使得对任意的 $A_1,A_2>A$ 有

$$\left|\int_{A_1}^{A_2}f(x)\,\mathrm{d}x\right|<\varepsilon.$$

<u>证明</u>. 这可由函数极限的 Cauchy 收敛准则得到。 □

接下来介绍无限区间上的广义积分收敛性的一些特殊判别法，与级数的情形类似，首先来考虑非负函数的广义积分。

**【命题 1.6】(比较判别法)** 设 $f(x)$ 与 $g(x)$ 均是定义在 $[a,+\infty)$ 上的非负函数，且存在 $A\geqslant a$ 使得在区间 $[A,+\infty)$ 上有 $f(x)\leqslant g(x)$，那么

(1) 若 $\displaystyle\int_a^{+\infty}g(x)\,\mathrm{d}x$ 收敛，则 $\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 收敛；

(2) 若 $\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 发散，则 $\displaystyle\int_a^{+\infty}g(x)\,\mathrm{d}x$ 发散。

<u>证明</u>. 只需证明 (1)。由 $f(x)$ 的非负性知，$\displaystyle\int_a^{A}f(x)\,\mathrm{d}x$ 是关于 $A\ (>a)$ 的递增函数，并且它也是关于 $A$ 的有界函数，这是因为

$$0\leqslant\int_a^{A}f(x)\,\mathrm{d}x\leqslant\int_a^{A}g(x)\,\mathrm{d}x\leqslant\int_a^{+\infty}g(x)\,\mathrm{d}x.$$

因此极限 $\displaystyle\lim_{A\to+\infty}\int_a^{A}f(x)\,\mathrm{d}x$ 必存在，从而广义积分 $\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 收敛。 □

**【推论 1.7】(比较判别法的极限形式)** 设 $f(x)$ 与 $g(x)$ 均是定义在 $[a,+\infty)$ 上的非负函数，且

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=C,$$

那么

(1) 当 $C\in(0,+\infty)$ 时，$\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 与 $\displaystyle\int_a^{+\infty}g(x)\,\mathrm{d}x$ 同敛散；

(2) 当 $C=0$ 时，$\int_a^{+\infty} g(x)\,\mathrm{d}x$ 收敛蕴含 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛；

(3) 当 $C=+\infty$ 时，$\int_a^{+\infty} g(x)\,\mathrm{d}x$ 发散蕴含 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 发散。

我们可以把比较判别法中的 $g(x)$ 选成一些特殊函数，从而相应地得到判断 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 敛散性的判别法。例如取 $g(x)=\dfrac{1}{x^p}$ 可得如下的 Cauchy 判别法。

**【定理 1.8】（Cauchy 判别法）** 设 $a>0$ 且 $f(x)$ 是 $[a,+\infty)$ 上的非负函数，那么

(1) 若存在 $p>1$ 使得在 $[a,+\infty)$ 上有 $f(x)\ll\dfrac{1}{x^p}$，则 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛；

(2) 若存在 $p\leqslant 1$ 使得在 $[a,+\infty)$ 上有 $f(x)\gg\dfrac{1}{x^p}$，则 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 发散。

Cauchy 判别法也有如下极限形式。

**【定理 1.9】（Cauchy 判别法的极限形式）** 设 $a>0$ 且 $f(x)$ 是定义在 $[a,+\infty)$ 上的非负函数。如果

$$\lim_{x\to+\infty} x^p f(x)=A,$$

那么

(1) 当 $A\in[0,+\infty)$ 且 $p>1$ 时，$\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛；

(2) 当 $A\in(0,+\infty)$ 或 $A=+\infty$ 且 $p\leqslant 1$ 时，$\int_a^{+\infty} f(x)\,\mathrm{d}x$ 发散。

**【例 1.10】** (1) 因为当 $x\in[1,+\infty)$ 时

$$\frac{|\cos x|}{x\sqrt{1+x^2}}\leqslant\frac{1}{x^2},$$

故而 $\int_1^{+\infty}\dfrac{|\cos x|}{x\sqrt{1+x^2}}\,\mathrm{d}x$ 收敛。

(2) 由于当 $x\to+\infty$ 时

$$\frac{\sqrt{x^3+1}}{x^2+2x+2}\sim\frac{1}{x^{\frac{1}{2}}},$$

所以 $\int_1^{+\infty}\dfrac{\sqrt{x^3+1}}{x^2+2x+2}\,\mathrm{d}x$ 发散。

**【例 1.11】** 设 $a>1$，下面来讨论广义积分 $\displaystyle\int_a^{+\infty}\frac{1}{x^\alpha(\log x)^\beta}\,\mathrm{d}x$ 的敛散性。分以下几种情况：

(1) 若 $\alpha>1$，则由 $\dfrac{\alpha+1}{2}>1$ 及

$$\lim_{x\to+\infty}x^{\frac{\alpha+1}{2}}\cdot\frac{1}{x^\alpha(\log x)^\beta}=\lim_{x\to+\infty}\frac{1}{x^{\frac{\alpha-1}{2}}(\log x)^\beta}=0$$

知 $\displaystyle\int_a^{+\infty}\frac{1}{x^\alpha(\log x)^\beta}\,\mathrm{d}x$ 收敛。

(2) 若 $\alpha<1$，则由 $\dfrac{\alpha+1}{2}<1$ 及

$$\lim_{x\to+\infty}x^{\frac{\alpha+1}{2}}\cdot\frac{1}{x^\alpha(\log x)^\beta}=\lim_{x\to+\infty}\frac{x^{\frac{1-\alpha}{2}}}{(\log x)^\beta}=+\infty$$

知 $\displaystyle\int_a^{+\infty}\frac{1}{x^\alpha(\log x)^\beta}\,\mathrm{d}x$ 发散。

(3) 若 $\alpha=1$，则对任意的 $A>a$ 有

$$\int_a^A\frac{1}{x(\log x)^\beta}\,\mathrm{d}x=\begin{cases}\log\log A-\log\log a, & \text{若 }\beta=1,\\ \dfrac{(\log A)^{1-\beta}-(\log a)^{1-\beta}}{1-\beta}, & \text{若 }\beta\neq1.\end{cases}$$

由此可以看出 $\displaystyle\int_a^{+\infty}\frac{1}{x(\log x)^\beta}\,\mathrm{d}x$ 在 $\beta>1$ 时收敛，在 $\beta\leqslant1$ 时发散。

无限区间上的广义积分使人不由得联想到无穷级数，下述命题将这二者联系了起来。

**【命题 1.12】** $\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 收敛的充要条件是：对于任意一个趋于 $+\infty$ 且满足 $A_1\geqslant a$ 的单调递增数列 $\{A_n\}$ 而言，级数

$$\sum_{n=1}^{\infty}\int_{A_{n-1}}^{A_n}f(x)\,\mathrm{d}x\qquad（记 A_0=a）\tag{11.1}$$

均收敛。

<u>证明</u>. 按照定义，$\displaystyle\int_a^{+\infty}f(x)\,\mathrm{d}x$ 存在是指极限

$$\lim_{A\to+\infty}\int_a^A f(x)\,\mathrm{d}x$$

存在，而由 Heine 归结原理知，上述极限存在当且仅当对趋于 $+\infty$ 的任意递增数列 $\{A_n\}$ $(A_1 \geqslant a)$ 而言极限

$$\lim_{n\to\infty}\int_a^{A_n} f(x)\,\mathrm{d}x \tag{11.2}$$

均存在。注意到

$$\int_a^{A_n} f(x)\,\mathrm{d}x = \sum_{k=1}^{n}\int_{A_{k-1}}^{A_k} f(x)\,\mathrm{d}x,$$

所以极限 (11.2) 存在等价于级数 (11.1) 收敛，从而命题得证。 □

由 Heine 归结原理我们知道，$\int_a^{+\infty} f(x)\,\mathrm{d}x$ 存在当且仅当对于任意一个趋于 $+\infty$ 且满足 $A_1 \geqslant a$ 的单调递增数列 $\{A_n\}$ 而言，级数 (11.1) 均收敛于同一个和，此时，该和也即是广义积分 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 的值。因此一方面，若能找到一个满足命题 1.12 中的条件的数列 $\{A_n\}$ 使得级数 (11.1) 发散，或找到两个数列使得它们所对应的级数 (11.1) 不收敛于同一个和，则可判定 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 发散。另一方面，要使用命题 1.12 来验证某广义积分收敛是比较困难的，然而，当 $f(x)$ 是非负函数时，我们可将命题 1.12 的条件弱化为如下便于应用的形式。

**【命题 1.13】** 设 $f(x)$ 是定义在 $[a,+\infty)$ 上的非负函数，则 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛的充要条件是：**存在**一个趋于 $+\infty$ 且满足 $A_1 \geqslant a$ 的单调递增数列 $\{A_n\}$ 使得级数 (11.1) 收敛[3]。

证明. 只需证充分性即可。由 $f$ 非负知

$$\int_a^{A} f(x)\,\mathrm{d}x$$

是关于 $A$ 的单调递增函数，且其值不超过级数 (11.1) 的和，因此极限

$$\lim_{A\to+\infty}\int_a^{A} f(x)\,\mathrm{d}x$$

必存在，从而 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛。 □

---

[3] 若 $f(x)$ 不是非负函数，则这一命题未必成立，例如由例 1.3 知 $\int_{\frac{\pi}{2}}^{+\infty} \sin x\,\mathrm{d}x$ 发散，但是级数

$$\sum_{n=1}^{\infty}\int_{(n-\frac{1}{2})\pi}^{(n+\frac{1}{2})\pi} \sin x\,\mathrm{d}x$$

收敛。然而如果对 $f(x)$ 及 $\{A_n\}$ 再增加一些限制条件，是可以保证得出所需结论的，参见习题 6。

**【例 1.14】** 证明广义积分 $\displaystyle\int_0^{+\infty}\frac{\mathrm{d}x}{1+x^3\sin^2 x}$ 收敛。

<u>证明</u>. 按照命题 1.13，我们只需证明级数

$$\sum_{n=0}^{\infty}\int_{n\pi}^{(n+1)\pi}\frac{\mathrm{d}x}{1+x^3\sin^2 x} \tag{11.3}$$

收敛即可。注意到

$$\begin{aligned}\int_{n\pi}^{(n+1)\pi}\frac{\mathrm{d}x}{1+x^3\sin^2 x}&=\int_0^{\pi}\frac{\mathrm{d}x}{1+(x+n\pi)^3\sin^2 x}\\&\leqslant\int_0^{\pi}\frac{\mathrm{d}x}{1+n^3\sin^2 x}=2\int_0^{\frac{\pi}{2}}\frac{\mathrm{d}x}{1+n^3\sin^2 x}\\&=2\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}x}{1+n^3\sin^2 x}+2\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}\frac{\mathrm{d}x}{1+n^3\sin^2 x},\end{aligned}$$

其中

$$\begin{aligned}\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}x}{1+n^3\sin^2 x}&=\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}x}{\cos^2 x+(n^3+1)\sin^2 x}=\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}(\tan x)}{1+(n^3+1)\tan^2 x}\\&=\frac{1}{\sqrt{n^3+1}}\arctan\sqrt{n^3+1}<\frac{\pi}{2n^{\frac{3}{2}}},\end{aligned}$$

而

$$\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}\frac{\mathrm{d}x}{1+n^3\sin^2 x}\leqslant\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}\frac{\mathrm{d}x}{\frac{1}{2}n^3}=\frac{\pi}{2n^3}.$$

因此

$$\int_{n\pi}^{(n+1)\pi}\frac{\mathrm{d}x}{1+x^3\sin^2 x}<\pi\left(\frac{1}{n^{\frac{3}{2}}}+\frac{1}{n^3}\right),$$

从而级数 (11.3) 收敛，进而原广义积分收敛。 □

下面我们转而讨论被积函数在积分区间上可正可负的情形。与级数类似，相应地也有 Abel 判别法和 Dirichlet 判别法，它们的基础是下面的积分第二中值定理，这是一个与 Abel 求和公式（第九章定理 3.1）类似的结果，

**【定理 1.15】（积分第二中值定理）** 设 $f(x)$ 在 $[a,b]$ 上可积，$g(x)$ 在 $[a,b]$ 上单调且非负，

(1) 若 $g(x)$ 单调递减，则存在 $\xi\in[a,b]$ 使得

$$\int_a^b f(x)g(x)\,\mathrm{d}x=g(a)\int_a^{\xi}f(x)\,\mathrm{d}x;$$

(2) 若 $g(x)$ 单调递增，则存在 $\xi\in[a,b]$ 使得

$$\int_a^b f(x)g(x)\,\mathrm{d}x = g(b)\int_\xi^b f(x)\,\mathrm{d}x.$$

证明. (1) 由 $f(x)$ 在 $[a,b]$ 上可积知其在 $[a,b]$ 上有界，即存在 $\rho>0$ 使得

$$|f(x)|\leqslant\rho,\qquad \forall\ x\in[a,b].$$

此外，由 $g(x)$ 在 $[a,b]$ 上单调知其在 $[a,b]$ 上可积，于是对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于 $[a,b]$ 的满足 $\max\limits_i \Delta x_i<\delta$ 的任意一组分点

$$a=x_0<x_1<\cdots<x_n=b \tag{11.4}$$

均有

$$\sum_{i=1}^n \omega_i\Delta x_i<\varepsilon,$$

其中 $\omega_i$ 是 $g(x)$ 在 $[x_{i-1},x_i]$ 上的振幅。现在我们任取一组满足 $\max\limits_i \Delta x_i<\delta$ 的分点 (11.4)。则有

$$\begin{aligned}
\int_a^b f(x)g(x)\,\mathrm{d}x &= \sum_{i=1}^n\int_{x_{i-1}}^{x_i} f(x)g(x)\,\mathrm{d}x\\
&= \sum_{i=1}^n g(x_{i-1})\int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x + \sum_{i=1}^n\int_{x_{i-1}}^{x_i} f(x)\big(g(x)-g(x_{i-1})\big)\,\mathrm{d}x\\
&= \sum_{i=1}^n g(x_{i-1})\int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x + O\bigg(\sum_{i=1}^n\int_{x_{i-1}}^{x_i}\rho\omega_i\,\mathrm{d}x\bigg)\\
&= \sum_{i=1}^n g(x_{i-1})\int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x + O\bigg(\rho\sum_{i=1}^n\omega_i\Delta x_i\bigg)\\
&= \sum_{i=1}^n g(x_{i-1})\big(F(x_i)-F(x_{i-1})\big)+O(\rho\varepsilon),
\end{aligned} \tag{11.5}$$

其中

$$F(x)=\int_a^x f(x)\,\mathrm{d}x.$$

由第八章定理 4.8 知 $F(x)$ 在 $[a,b]$ 上连续，从而 $F(x)$ 在该区间上能取到最大值和最小值，我们记 $M=\max\limits_{x\in[a,b]}F(x)$，$m=\min\limits_{x\in[a,b]}F(x)$。利用分部求和（第九章定理 3.1）

可得（注意到 $F(x_0)=F(a)=0$）

$$\begin{aligned}
&\sum_{i=1}^{n} g(x_{i-1})\big(F(x_i)-F(x_{i-1})\big)\\
&=g(x_{n-1})\big(F(x_n)-F(x_0)\big)+\sum_{i=1}^{n-1}\big(g(x_{i-1})-g(x_i)\big)\big(F(x_i)-F(x_0)\big)\\
&=g(x_{n-1})F(x_n)+\sum_{i=1}^{n-1}\big(g(x_{i-1})-g(x_i)\big)F(x_i).
\end{aligned}$$

因为 $g(x)$ 在 $[a,b]$ 上单调递减且非负，故而

$$\begin{aligned}
\sum_{i=1}^{n} g(x_{i-1})\big(F(x_i)-F(x_{i-1})\big)&\leqslant M\left[g(x_{n-1})+\sum_{i=1}^{n-1}\big(g(x_{i-1})-g(x_i)\big)\right]\\
&=Mg(x_0)=Mg(a).
\end{aligned}$$

同样，

$$\begin{aligned}
\sum_{i=1}^{n} g(x_{i-1})\big(F(x_i)-F(x_{i-1})\big)&\geqslant m\left[g(x_{n-1})+\sum_{i=1}^{n-1}\big(g(x_{i-1})-g(x_i)\big)\right]\\
&=mg(a).
\end{aligned}$$

将以上两式代入 (11.5) 即得

$$mg(a)+O(\rho\varepsilon)\leqslant\int_a^b f(x)g(x)\,\mathrm{d}x\leqslant Mg(a)+O(\rho\varepsilon),$$

再令 $\varepsilon\to0^+$，我们最终得到

$$mg(a)\leqslant\int_a^b f(x)g(x)\,\mathrm{d}x\leqslant Mg(a).$$

因此存在 $\eta\in[m,M]$ 使得

$$\int_a^b f(x)g(x)\,\mathrm{d}x=\eta g(a).$$

注意到 $F(x)$ 在 $[a,b]$ 上连续，故由介值定理知存在 $\xi\in[a,b]$ 使得 $F(\xi)=\eta$，于是

$$\int_a^b f(x)g(x)\,\mathrm{d}x=g(a)F(\xi)=g(a)\int_a^\xi f(x)\,\mathrm{d}x.$$

(2) 当然，我们可以用类似的方法去证明 $g(x)$ 单调递增的情形，但是如下变量替换的方式显得更简洁。容易看出

$$\int_a^b f(x)g(x)\,\mathrm{d}x = \int_a^b f(a+b-t)g(a+b-t)\,\mathrm{d}t.$$

由 $g(x)$ 在 $[a,b]$ 上单调递增可得 $g(a+b-t)$ 在 $[a,b]$ 上单调递减，于是由 (1) 知存在 $\xi \in [a,b]$ 使得

$$\int_a^b f(x)g(x)\,\mathrm{d}x = g(b)\int_a^\xi f(a+b-t)\,\mathrm{d}t = g(b)\int_{a+b-\xi}^b f(x)\,\mathrm{d}x.$$

注意到 $a+b-\xi \in [a,b]$，从而 (2) 得证。 □

如果去除掉 $g$ 的非负性的限制，则有如下一般形式的结论。

**【定理 1.16】（积分第二中值定理）** 设 $f(x)$ 在 $[a,b]$ 上可积，$g(x)$ 在 $[a,b]$ 上单调，则存在 $\xi \in [a,b]$ 使得

$$\int_a^b f(x)g(x)\,\mathrm{d}x = g(a)\int_a^\xi f(x)\,\mathrm{d}x + g(b)\int_\xi^b f(x)\,\mathrm{d}x. \tag{11.6}$$

证明. 若 $g(x)$ 单调递减，则 $g(x)-g(b)$ 也单调递减且非负，于是由定理 1.15 (1) 知存在 $\xi \in [a,b]$ 使得

$$\int_a^b f(x)[g(x)-g(b)]\,\mathrm{d}x = [g(a)-g(b)]\int_a^\xi f(x)\,\mathrm{d}x,$$

此即 (11.6)。当 $g(x)$ 单调递增时我们可也用定理 1.15 (2) 给出类似证明。 □

有了以上准备工作，我们可以来介绍 Abel 判别法和 Dirichlet 判别法了。

**【定理 1.17】（Abel 判别法）** 设 $f(x)$ 在 $[a,+\infty)$ 上可积，$g(x)$ 在 $[a,+\infty)$ 上单调且有界，则 $\displaystyle\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$ 收敛。

证明. 因为 $f(x)$ 在 $[a,+\infty)$ 上可积，故由 Cauchy 收敛准则知，对任意的 $\varepsilon > 0$，存在 $A \geqslant a$，使得当 $A_2 > A_1 > A$ 时有

$$\left|\int_{A_1}^{A_2} f(x)\,\mathrm{d}x\right| < \varepsilon.$$

于是对任意的 $B_2 > B_1 > A$，在区间 $[B_1,B_2]$ 上应用积分第二中值定理知，存在 $\xi \in [B_1,B_2]$ 使得

$$\int_{B_1}^{B_2} f(x)g(x)\,\mathrm{d}x = g(B_1)\int_{B_1}^\xi f(x)\,\mathrm{d}x + g(B_2)\int_\xi^{B_2} f(x)\,\mathrm{d}x = O(\varepsilon),$$

再由 Cauchy 收敛准则便可推出 $\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$ 收敛。 □

**【定理 1.18】(Dirichlet 判别法)** 假设函数 $F(A)=\int_a^A f(x)\,\mathrm{d}x$ 在 $[a,+\infty)$ 上有界，$g(x)$ 在 $[a,+\infty)$ 上单调且 $\lim\limits_{x\to+\infty} g(x)=0$，那么广义积分 $\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$ 收敛。

<u>证明</u>. 由 $\lim\limits_{x\to+\infty} g(x)=0$ 知，对任意的 $\varepsilon>0$，存在 $A\geqslant a$，使得当 $x>A$ 时有 $|g(x)|<\varepsilon$。于是对任意的 $A_2>A_1>A$，由积分第二中值定理知，存在 $\xi\in[A_1,A_2]$ 使得

$$\begin{aligned}\int_{A_1}^{A_2} f(x)g(x)\,\mathrm{d}x &= g(A_1)\int_{A_1}^{\xi} f(x)\,\mathrm{d}x+g(A_2)\int_{\xi}^{A_2} f(x)\,\mathrm{d}x\\ &= g(A_1)\big(F(\xi)-F(A_1)\big)+g(A_2)\big(F(A_2)-F(\xi)\big)=O(\varepsilon).\end{aligned}$$

因此由 Cauchy 收敛准则可推出 $\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$ 收敛。 □

**【例 1.19】** 设 $\alpha>0$，证明 $\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 与 $\int_1^{+\infty}\frac{\cos x}{x^\alpha}\,\mathrm{d}x$ 均收敛。

<u>证明</u>. 对任意的 $A\geqslant 1$ 有

$$\left|\int_1^A \sin x\,\mathrm{d}x\right| = |\cos A-\cos 1|\leqslant 2,$$

故由 Dirichlet 判别法知 $\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 收敛。同理可证 $\int_1^{+\infty}\frac{\cos x}{x^\alpha}\,\mathrm{d}x$ 收敛。 □

广义积分也有绝对收敛和条件收敛的概念。

**【定义 1.20】** 设 $f(x)$ 在 $[a,+\infty)$ 上有定义，若 $\int_a^{+\infty}|f(x)|\,\mathrm{d}x$ 收敛，则称广义积分 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ <u>绝对收敛</u>。如果 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛但不绝对收敛，则称之为<u>条件收敛</u>的。

**【命题 1.21】** 若 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 绝对收敛，则其必收敛。

<u>证明</u>. 这可由

$$\left|\int_{A_1}^{A_2} f(x)\,\mathrm{d}x\right|\leqslant\int_{A_1}^{A_2}|f(x)|\,\mathrm{d}x,\qquad \forall\ A_2>A_1\geqslant a$$

及 Cauchy 收敛准则得到。 □

【**例 1.22**】设 $\alpha>0$，判断 $\displaystyle\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 的绝对收敛性和条件收敛性。

解. 当 $\alpha>1$ 时，由

$$\frac{|\sin x|}{x^\alpha}\leqslant\frac{1}{x^\alpha}$$

及比较判别法知 $\displaystyle\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 绝对收敛。

现设 $0<\alpha\leqslant 1$。由例 1.19 知 $\displaystyle\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 收敛，下证积分 $\displaystyle\int_1^{+\infty}\frac{|\sin x|}{x^\alpha}\,\mathrm{d}x$ 发散。因为

$$\frac{|\sin x|}{x^\alpha}\geqslant\frac{\sin^2 x}{x^\alpha}=\frac{1}{2x^\alpha}-\frac{\cos 2x}{2x^\alpha},$$

并且一方面由 Dirichlet 判别法知 $\displaystyle\int_1^{+\infty}\frac{\cos 2x}{x^\alpha}\,\mathrm{d}x$ 收敛，另一方面由例 1.2 知当 $0<\alpha\leqslant 1$ 时 $\displaystyle\int_1^{+\infty}\frac{1}{x^\alpha}\,\mathrm{d}x$ 发散，故而 $\displaystyle\int_1^{+\infty}\frac{|\sin x|}{x^\alpha}\,\mathrm{d}x$ 发散。所以当 $0<\alpha\leqslant 1$ 时积分 $\displaystyle\int_1^{+\infty}\frac{\sin x}{x^\alpha}\,\mathrm{d}x$ 条件收敛。 □

在本节最后，我们来介绍广义积分主值的概念。

【**定义 1.23**】设 $f(x)$ 在 $\mathbb{R}$ 上有定义，且在任一有限闭区间上可积。如果

$$\lim_{A\to+\infty}\int_{-A}^{A}f(x)\,\mathrm{d}x$$

存在，则将上述极限值称为广义积分 $\displaystyle\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x$ 的 Cauchy 主值，记作

$$\mathrm{P.V.}\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x.$$

若 $\displaystyle\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x$ 收敛，当然有

$$\mathrm{P.V.}\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x.$$

但在一般情况下，Cauchy 主值存在并不蕴含广义积分收敛。例如 $\displaystyle\int_{-\infty}^{+\infty}\sin x\,\mathrm{d}x$ 发散，但是

$$\mathrm{P.V.}\int_{-\infty}^{+\infty}\sin x\,\mathrm{d}x=\lim_{A\to+\infty}\int_{-A}^{A}\sin x\,\mathrm{d}x=0.$$

## 习 题 11.1

1. 证明推论 1.7。

2. 判断下列广义积分的敛散性:

(1) $\displaystyle\int_0^{+\infty} \frac{\sqrt{x}}{x^2+x+2}\,\mathrm{d}x$; (2) $\displaystyle\int_1^{+\infty} \frac{\arctan x}{x\sqrt{1+x^2}}\,\mathrm{d}x$;

(3) $\displaystyle\int_1^{+\infty} \frac{\log(1+x)}{x^\alpha}\,\mathrm{d}x$; (4) $\displaystyle\int_0^{+\infty} x^\alpha e^{-x}\,\mathrm{d}x \quad (\alpha \geqslant 0)$;

(5) $\displaystyle\int_3^{+\infty} \frac{1}{(\log x)^{\log\log x}}\,\mathrm{d}x$; (6) $\displaystyle\int_1^{+\infty} \left(\frac{1}{\sqrt{x}} - \sqrt{\log\frac{1+x}{x}}\right)\mathrm{d}x$;

(7) $\displaystyle\int_1^{+\infty} \log\left(\sin\frac{1}{x} + \cos\frac{1}{x}\right)\mathrm{d}x$; (8) $\displaystyle\int_1^{+\infty} \frac{1}{1+x|\sin x|}\,\mathrm{d}x$;

(9) $\displaystyle\int_1^{+\infty} \frac{x-[x]-\frac{1}{2}}{\sqrt{x}}\,\mathrm{d}x$; (10) $\displaystyle\int_1^{+\infty} \frac{\sin x}{x^{1-\frac{1}{x}}}\,\mathrm{d}x$。

3. 求下列广义积分的 Cauchy 主值:

(1) $\displaystyle\int_{-\infty}^{+\infty} \frac{1+x}{1+x^2}\,\mathrm{d}x$; (2) $\displaystyle\int_{-\infty}^{+\infty} \arctan x\,\mathrm{d}x$。

4. 设 $f$ 是定义在 $\mathbb{R}$ 上的非负函数，且 Cauchy 主值 P.V. $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ 存在，证明 $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ 收敛且

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \mathrm{P.V.} \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$$

5. 证明 $\displaystyle\int_1^{+\infty} \left(\frac{1}{[x]} - \frac{1}{x}\right)\mathrm{d}x = \gamma$，其中 $\gamma$ 是 Euler 常数。

6. 设 $f$ 是定义在 $[a,+\infty)$ 上的函数，且对任意的 $A>a$，$f$ 在 $[a,A]$ 上可积。若存在趋于 $+\infty$ 且满足 $A_1 \geqslant a$ 的单调递增数列 $\{A_n\}$ 使得级数

$$\sum_{n=1}^{\infty} \int_{A_n}^{A_{n+1}} f(x)\,\mathrm{d}x$$

收敛，且对每个 $n \in \mathbb{Z}_{\geqslant 1}$，函数 $f$ 在 $(A_n, A_{n+1})$ 上不变号。证明 $\displaystyle\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛。

7. 证明 $\displaystyle\int_1^{+\infty} \frac{(-1)^{[\sqrt{x}]}}{x}\,\mathrm{d}x$ 收敛。

8. 设 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 绝对收敛，那么 $\int_a^{+\infty} f(x)^2\,\mathrm{d}x$ 是否一定收敛？

9. 设 $f$ 在 $[a,+\infty)$ 上非负连续且 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛，举例说明 $\lim\limits_{x\to+\infty} f(x)=0$ 未必成立。

10. 设 $f$ 在 $[a,+\infty)$ 上一致连续且 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛，证明 $\lim\limits_{x\to+\infty} f(x)=0$。

11. 设 $f$ 在 $[a,+\infty)$ 上单调递减且非负，证明 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 与 $\int_a^{+\infty} f(x)\sin^2 x\,\mathrm{d}x$ 同敛散。

12. 设 $f$ 在 $[a,+\infty)$ 上单调且 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 收敛，证明 $\lim\limits_{x\to+\infty} xf(x)=0$。

13. 假设函数 $g(x)$ 在 $[a,b]$ 上可积，$f(x)$ 在 $[a,b]$ 上可导，$\dfrac{g(x)}{f'(x)}$ 在 $[a,b]$ 上单调且 $\left|\dfrac{f'(x)}{g(x)}\right| \geqslant m>0$。证明

$$\left|\int_a^b g(x)\cos f(x)\,\mathrm{d}x\right| \leqslant \frac{2}{m}.$$

14. 设 $f$ 是定义在 $[a,+\infty)$ 上的函数，且对任意的 $A>a$，$f$ 在 $[a,A]$ 上可积。记

$$f^+(x)=\frac{|f(x)|+f(x)}{2},\qquad f^-(x)=\frac{|f(x)|-f(x)}{2}.$$

证明：

(1) 若 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 绝对收敛，则 $\int_a^{+\infty} f^+(x)\,\mathrm{d}x$ 与 $\int_a^{+\infty} f^-(x)\,\mathrm{d}x$ 均收敛，且

$$\int_a^{+\infty} f(x)\,\mathrm{d}x=\int_a^{+\infty} f^+(x)\,\mathrm{d}x-\int_a^{+\infty} f^-(x)\,\mathrm{d}x.$$

(2) 若 $\int_a^{+\infty} f(x)\,\mathrm{d}x$ 条件收敛，则 $\int_a^{+\infty} f^+(x)\,\mathrm{d}x$ 与 $\int_a^{+\infty} f^-(x)\,\mathrm{d}x$ 均发散于 $+\infty$，且当 $x\to+\infty$ 时有

$$\int_a^x f^+(t)\,\mathrm{d}t \sim \int_a^x f^-(t)\,\mathrm{d}t.$$

## §11.2 有界区间上的无界函数的积分

如果函数 $f(x)$ 在点 $x_0$ 的任意去心邻域（或左邻域，或右邻域）内无界，那么我们就称 $x_0$ 是 $f(x)$ 的一个奇点 (singularity) 或瑕点。例如 0 是 $f(x)=\dfrac{1}{x}$ 的奇点。在本节中我们将在这种情形下讨论相应的广义积分。

**【定义 2.1】**设 $f(x)$ 在区间 $(a,b]$ 上有定义，$a$ 是 $f(x)$ 的奇点，且对任意的 $c\in(a,b]$ 而言 $f(x)$ 在 $[c,b]$ 上可积。若极限

$$\lim_{c\to a^+}\int_c^b f(x)\,\mathrm{d}x \tag{11.7}$$

存在，则称广义积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛，并将上述极限值称作这一广义积分的值，记作

$$\int_a^b f(x)\,\mathrm{d}x=\lim_{c\to a^+}\int_c^b f(x)\,\mathrm{d}x.$$

若 (11.7) 中的极限不存在，就称广义积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 发散。

类似地，若 $f(x)$ 在 $[a,b)$ 的任一闭子区间上可积且 $b$ 是 $f(x)$ 的奇点，则用极限

$$\lim_{c\to b^-}\int_a^c f(x)\,\mathrm{d}x$$

的存在与否来定义广义积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 的收敛和发散。此外，如果 $f(x)$ 在 $[a,c)$ 和 $(c,b]$ 的任一闭子区间上可积且 $c$ 是 $f(x)$ 的奇点，那么当 $\displaystyle\int_a^c f(x)\,\mathrm{d}x$ 与 $\displaystyle\int_c^b f(x)\,\mathrm{d}x$ 均收敛时我们称 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛，并记

$$\int_a^b f(x)\,\mathrm{d}x=\int_a^c f(x)\,\mathrm{d}x+\int_c^b f(x)\,\mathrm{d}x.$$

**【例 2.2】**讨论广义积分 $\displaystyle\int_a^b\frac{\mathrm{d}x}{(x-a)^p}$ $(p>0)$ 的敛散性。

解. 在这里 $a$ 是奇点。由于对任意的 $c\in(a,b]$ 有

$$\int_c^b\frac{\mathrm{d}x}{(x-a)^p}=\begin{cases}\log(b-a)-\log(c-a), & \text{若 } p=1,\\ \dfrac{(b-a)^{1-p}-(c-a)^{1-p}}{1-p}, & \text{若 } p\neq 1,\end{cases}$$

因此

$$\lim_{c\to a^+}\int_c^b \frac{\mathrm{d}x}{(x-a)^p} = \begin{cases} \dfrac{(b-a)^{1-p}}{1-p}, & \text{若 } 0<p<1, \\ +\infty, & \text{若 } p\geqslant 1. \end{cases}$$

故而广义积分 $\displaystyle\int_a^b \frac{\mathrm{d}x}{(x-a)^p}$ 当 $0<p<1$ 时收敛，当 $p\geqslant 1$ 时发散。 □

本节中所讨论的广义积分拥有与无界区间上的广义积分类似的性质，在此仅将相应结论写出而不给证明。为方便起见，在下面的讨论中，如果没有专门提及，我们均假设相应的函数在 $(a,b]$ 的任意闭子区间上可积，且 $x=a$ 是该函数仅有的奇点。

**【命题 2.3】** 设 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 与 $\displaystyle\int_a^b g(x)\,\mathrm{d}x$ 均收敛，那么

(1) 对任意的 $b'\in(a,b)$，$\displaystyle\int_a^{b'} f(x)\,\mathrm{d}x$ 收敛；

(2) 对任意的实数 $\alpha$ 和 $\beta$，$\displaystyle\int_a^b \big[\alpha f(x)+\beta g(x)\big]\,\mathrm{d}x$ 收敛，且

$$\int_a^b \big[\alpha f(x)+\beta g(x)\big]\,\mathrm{d}x = \alpha\int_a^b f(x)\,\mathrm{d}x + \beta\int_a^b g(x)\,\mathrm{d}x.$$

**【定理 2.4】（Cauchy 收敛准则）** $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛的充要条件是：对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对任意的 $\eta,\eta'\in(0,\delta)$ 有

$$\left|\int_{a+\eta}^{a+\eta'} f(x)\,\mathrm{d}x\right| < \varepsilon.$$

**【命题 2.5】（比较判别法）** 设在 $a$ 的某个右邻域内有 $0\leqslant f(x)\leqslant g(x)$，那么

(1) $\displaystyle\int_a^b g(x)\,\mathrm{d}x$ 收敛蕴含 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛；

(2) $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 发散蕴含 $\displaystyle\int_a^b g(x)\,\mathrm{d}x$ 发散。

**【推论 2.6】（比较判别法的极限形式）** 设 $f(x)$ 与 $g(x)$ 均在 $(a,b]$ 上非负，且

$$\lim_{x\to a^+}\frac{f(x)}{g(x)} = C,$$

那么

(1) 当 $C \in (0, +\infty)$ 时，$\int_a^b f(x)\,\mathrm{d}x$ 与 $\int_a^b g(x)\,\mathrm{d}x$ 同敛散；

(2) 当 $C = 0$ 时，$\int_a^b g(x)\,\mathrm{d}x$ 收敛蕴含 $\int_a^b f(x)\,\mathrm{d}x$ 收敛；

(3) 当 $C = +\infty$ 时，$\int_a^b g(x)\,\mathrm{d}x$ 发散蕴含 $\int_a^b f(x)\,\mathrm{d}x$ 发散。

将 $(a, b]$ 上的非负函数与 $\dfrac{1}{(x-a)^p}$ 比较可得下面的 Cauchy 判别法。

**【定理 2.7】（Cauchy 判别法）** 设 $f(x)$ 是定义在 $(a, b]$ 上的非负函数。

(1) 若存在 $p < 1$ 使得 $f(x) \ll \dfrac{1}{(x-a)^p}$，则 $\int_a^b f(x)\,\mathrm{d}x$ 收敛；

(2) 若存在 $p \geqslant 1$ 使得 $f(x) \gg \dfrac{1}{(x-a)^p}$，则 $\int_a^b f(x)\,\mathrm{d}x$ 发散；

(3) 若 $\lim\limits_{x \to a^+} (x-a)^p f(x) = A$，那么当 $A \in [0, +\infty)$ 且 $p < 1$ 时 $\int_a^b f(x)\,\mathrm{d}x$ 收敛；当 $A \in (0, +\infty)$ 或 $A = +\infty$，且 $p \geqslant 1$ 时 $\int_a^b f(x)\,\mathrm{d}x$ 发散。

**【例 2.8】** 讨论广义积分 $\displaystyle\int_1^2 \frac{\log x}{(x-1)^p}\,\mathrm{d}x$ $(p > 0)$ 的敛散性。

<u>解</u>. $x = 1$ 是被积函数仅有的奇点。由于当 $x \to 1^+$ 时 $\dfrac{\log x}{(x-1)^p} \sim \dfrac{1}{(x-1)^{p-1}}$，故而该积分在 $p < 2$ 时收敛，在 $p \geqslant 2$ 时发散。 □

**【例 2.9】** 判断 $\displaystyle\int_0^\infty \frac{\mathrm{d}x}{x^p |1-x|^q}$ $(p > 0,\ q > 0)$ 的敛散性。

<u>解</u>. 分别考虑

$$\int_0^2 \frac{\mathrm{d}x}{x^p |1-x|^q} \qquad 与 \qquad \int_2^\infty \frac{\mathrm{d}x}{x^p |1-x|^q}. \tag{11.8}$$

由于当 $x \to +\infty$ 时 $\dfrac{1}{x^p (x-1)^q} \sim \dfrac{1}{x^{p+q}}$，故 (11.8) 中第二个积分当 $p + q > 1$ 时收敛，当 $p + q \leqslant 1$ 时发散。

下面考虑 (11.8) 中的第一个积分，被积函数有两个奇点，即 $x = 0$ 和 $x = 1$。由于当 $x \to 0^+$ 时 $\dfrac{1}{x^p |1-x|^q} \sim \dfrac{1}{x^p}$，当 $x \to 1$ 时 $\dfrac{1}{x^p |1-x|^q} \sim \dfrac{1}{|1-x|^q}$，故而这一积分收敛当且仅当 $p < 1$ 且 $q < 1$。

综上，$\displaystyle\int_0^\infty \frac{\mathrm{d}x}{x^p |1-x|^q}$ $(p > 0,\ q > 0)$ 当 $p < 1$，$q < 1$ 且 $p + q > 1$ 时收敛，而在其余情形发散。 □

【定理 2.10】(Abel 判别法) 设 $f(x)$ 与 $g(x)$ 均在 $(a,b]$ 上有定义，$x=a$ 是 $f(x)$ 的奇点。若 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛且 $g(x)$ 在 $(a,b]$ 上单调有界，那么 $\displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x$ 收敛。

【定理 2.11】(Dirichlet 判别法) 设 $f(x)$ 与 $g(x)$ 均是定义在 $(a,b]$ 上的函数，$x=a$ 是 $f(x)$ 的奇点。如果函数 $F(c)=\displaystyle\int_c^b f(x)\,\mathrm{d}x$ 在 $(a,b]$ 上有界，且 $g(x)$ 在 $(a,b]$ 上单调并满足 $\displaystyle\lim_{x\to a^+} g(x)=0$，那么 $\displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x$ 收敛。

【定义 2.12】设 $f(x)$ 是定义在 $(a,b]$ 上的函数，$x=a$ 是 $f(x)$ 的奇点。若 $\displaystyle\int_a^b |f(x)|\,\mathrm{d}x$ 收敛，则称 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 绝对收敛。若 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛但不绝对收敛，则称之为条件收敛的。

容易证明，绝对收敛的广义积分必收敛。

【定义 2.13】设 $f(x)$ 在 $[a,c)$ 和 $(c,b]$ 上有定义，且 $c$ 是 $f(x)$ 的奇点。若

$$\lim_{\eta\to 0^+}\left[\int_a^{c-\eta} f(x)\,\mathrm{d}x+\int_{c+\eta}^b f(x)\,\mathrm{d}x\right]$$

存在，则将上述极限值称为广义积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 的 Cauchy 主值，记作

$$\mathrm{P.V.}\int_a^b f(x)\,\mathrm{d}x.$$

若 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛，当然有

$$\mathrm{P.V.}\int_a^b f(x)\,\mathrm{d}x=\int_a^b f(x)\,\mathrm{d}x.$$

但在一般情况下，Cauchy 主值存在并不蕴含广义积分收敛。

## 习 题 11.2

1. 判断下列广义积分的敛散性：

(1) $\displaystyle\int_0^{\frac{1}{2}}\frac{\mathrm{d}x}{\log x}$;　　(2) $\displaystyle\int_0^1 x^p\log^q x\,\mathrm{d}x\quad(p,\,q\in\mathbb{R})$;

(3) $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}x}{\sin^p x \cos^q x} \quad (p>0,\ q>0)$；　(4) $\displaystyle\int_0^{+\infty} x^{p-1}e^{-x}\,\mathrm{d}x \quad (p\in\mathbb{R})$；

(5) $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{|x-1|^p\cdot|x-2|^q} \quad (p>0,\ q>0)$；(6) $\displaystyle\int_0^{+\infty} \frac{\sin x}{x^p}\,\mathrm{d}x \quad (p>0)$。

2. 求下列广义积分的 Cauchy 主值：

(1) $\displaystyle\int_0^3 \frac{\mathrm{d}x}{1-x^2}$；　(2) $\displaystyle\int_{\frac{1}{2}}^2 \frac{\mathrm{d}x}{x\log x}$。

3. 设 $f(x)$ 是 $(0,1]$ 上的单调函数，且 $x=0$ 是 $f$ 唯一的奇点。证明 $\displaystyle\int_0^1 f(x)\,\mathrm{d}x$ 收敛当且仅当极限 $\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n f\Big(\frac{k}{n}\Big)$ 存在，并且当这一广义积分收敛时有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n f\Big(\frac{k}{n}\Big)=\int_0^1 f(x)\,\mathrm{d}x.$$

4. 令 $f(x)=\dfrac{1}{x}-\dfrac{1}{1-x}$，证明 $f(x)$ 在 $(0,1)$ 上单调，且极限 $\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1} f\Big(\frac{k}{n}\Big)$ 存在，但是 $\displaystyle\int_0^1 f(x)\,\mathrm{d}x$ 发散。注意将本题与上题作比较。

5. 设 $f(x)$ 是 $(0,1]$ 上的单调函数，$x=0$ 是 $f$ 唯一的奇点，且 $\displaystyle\int_0^1 f(x)\,\mathrm{d}x$ 收敛，那么是否对任意选取的 $\xi_k\in\Big(\dfrac{k-1}{n},\dfrac{k}{n}\Big]$ $(1\leqslant k\leqslant n)$ 均有

$$\int_0^1 f(x)\,\mathrm{d}x=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n f(\xi_k)\ ?$$

## § 11.3
## 广义积分的计算

本节的主要目的是介绍广义积分的一些基本计算方法。

**【定理 3.1】（Newton–Leibniz 公式）** 设 $(a,b)$ 是一个区间（$a$ 可以是 $-\infty$，$b$ 可以是 $+\infty$），$f$ 在 $(a,b)$ 的任意闭子区间上可积且广义积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛。又设 $f(x)$ 在 $(a,b)$ 上有原函数 $F(x)$，且极限 $\displaystyle\lim_{x\to a^+}F(x)$ 与 $\displaystyle\lim_{x\to b^-}F(x)$ 均存在，那么

$$\int_a^b f(x)\,\mathrm{d}x=\lim_{x\to b^-}F(x)-\lim_{x\to a^+}F(x).$$

证明. 对任意的 $[c,d]\subseteq(a,b)$，由 Newton–Leibniz 公式（第八章定理 5.1）知

$$\int_c^d f(x)\,\mathrm{d}x=F(d)-F(c).$$

令 $c\to a^+$，$d\to b^-$ 即得结论。 □

在定理 3.1 的条件下，为方便起见，通常记

$$\lim_{x\to b^-}F(x)-\lim_{x\to a^+}F(x)=F(x)\Big|_a^b.$$

**【例 3.2】**计算 $\displaystyle\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{1+x^2}$。

解. 由定理 3.1 知

$$\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{1+x^2}=\arctan x\Big|_{-\infty}^{+\infty}=\pi.$$

□

**【命题 3.3】** 设 $(a,b)$ 是一个区间（$a$ 可以是 $-\infty$，$b$ 可以是 $+\infty$），$\varphi$ 是定义在 $(a,b)$ 上的严格单调且可导的函数。现记

$$\alpha=\lim_{t\to a^+}\varphi(t),\qquad\qquad \beta=\lim_{t\to b^-}\varphi(t),$$

并假设 $f$ 在 $(\alpha,\beta)$ 的任意闭子区间上可积，那么 $\displaystyle\int_\alpha^\beta f(x)\,\mathrm{d}x$ 与 $\displaystyle\int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t$ 同敛散，且在二者皆收敛时有

$$\int_\alpha^\beta f(x)\,\mathrm{d}x=\int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

证明. 由第八章命题 5.4 知，对 $(a,b)$ 的任意闭子区间 $[c,d]$ 有

$$\int_{\varphi(c)}^{\varphi(d)}f(x)\,\mathrm{d}x=\int_c^d f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

由假设条件知 $t\to a^+$ 当且仅当 $\varphi(t)\to\alpha$，$t\to b^-$ 当且仅当 $\varphi(t)\to\beta$，因此令 $c\to a^+$，$d\to b^-$ 即得结论。 □

**【例 3.4】**设 $b>a$，计算 $\displaystyle\int_a^b\frac{\mathrm{d}x}{\sqrt{(x-a)(b-x)}}$。

解. 作变量替换 $x=a\cos^2 t+b\sin^2 t$ ($t\in[0,\frac{\pi}{2}]$) 可得

$$\int_a^b\frac{\mathrm{d}x}{\sqrt{(x-a)(b-x)}}=\int_0^{\frac{\pi}{2}}\frac{(b-a)\sin 2t\,\mathrm{d}t}{\sqrt{[(b-a)\sin^2 t]\cdot[(b-a)\cos^2 t]}}=2\int_0^{\frac{\pi}{2}}\,\mathrm{d}t=\pi.$$

□

**【命题 3.5】**（分部积分法） 设 $(a,b)$ 是一个区间（$a$ 可以是 $-\infty$，$b$ 可以是 $+\infty$），$f(x)$ 与 $g(x)$ 均在 $(a,b)$ 上连续可导，且极限

$$\lim_{x\to a^+} f(x)g(x) \qquad \text{与} \qquad \lim_{x\to b^-} f(x)g(x)$$

均存在。那么当广义积分

$$\int_a^b f(x)g'(x)\,\mathrm{d}x \qquad \text{与} \qquad \int_a^b f'(x)g(x)\,\mathrm{d}x$$

中有一个收敛时，另一个必收敛，且有

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x,$$

其中

$$f(x)g(x)\Big|_a^b = \lim_{x\to b^-} f(x)g(x) - \lim_{x\to a^+} f(x)g(x).$$

证明. 对任意的 $[c,d]\subseteq(a,b)$，由第八章命题 5.10（分部积分法）知

$$\int_c^d f(x)g'(x)\,\mathrm{d}x = f(x)g(x)\Big|_c^d - \int_c^d f'(x)g(x)\,\mathrm{d}x.$$

令 $c\to a^+$，$d\to b^-$ 即得结论。 □

**【例 3.6】**计算 $\displaystyle\int_0^1 \frac{\log x}{\sqrt{x}}\,\mathrm{d}x$。

解. 由分部积分可得

$$\begin{aligned}\int_0^1 \frac{\log x}{\sqrt{x}}\,\mathrm{d}x &= 2\int_0^1 \log x\,\mathrm{d}\sqrt{x} = 2\sqrt{x}\log x\Big|_0^1 - 2\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x\\ &= -2\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = -4.\end{aligned}$$

□

最后我们通过计算三个特殊的广义积分来结束本节。

**【例 3.7】**计算 Euler 积分

$$I = \int_0^{\frac{\pi}{2}} \log\sin x\,\mathrm{d}x.$$

解. 作变量替换 $x=2t$，并利用 $\sin 2t = 2\sin t\cos t$ 可得

$$I = 2\int_0^{\frac{\pi}{4}} \log\sin 2t\,\mathrm{d}t = \frac{\pi}{2}\log 2 + 2\int_0^{\frac{\pi}{4}} \log\sin t\,\mathrm{d}t + 2\int_0^{\frac{\pi}{4}} \log\cos t\,\mathrm{d}t. \tag{11.9}$$

由于
$$\int_0^{\frac{\pi}{4}} \log\cos t\,\mathrm{d}t = \int_{\frac{\pi}{2}}^{\frac{\pi}{4}} \log\cos\left(\frac{\pi}{2}-t\right)\mathrm{d}\left(\frac{\pi}{2}-t\right) = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \log\sin t\,\mathrm{d}t,$$
代入 (11.9) 可得
$$I = \frac{\pi}{2}\log 2 + 2\int_0^{\frac{\pi}{2}} \log\sin t\,\mathrm{d}t = \frac{\pi}{2}\log 2 + 2I,$$
因此 $I = -\dfrac{\pi}{2}\log 2$。 □

**【例 3.8】**设 $a$, $b$ 是两个正数，计算 Froullani 积分
$$\int_0^{+\infty} \frac{f(ax)-f(bx)}{x}\,\mathrm{d}x,$$
其中 $f(x)$ 在 $[0,+\infty)$ 上连续，且 $\lim\limits_{x\to+\infty} f(x) = A$。

解. 对任意的 $[\varepsilon, M] \subseteq (0,+\infty)$ 有
$$\begin{aligned}\int_\varepsilon^M \frac{f(ax)-f(bx)}{x}\,\mathrm{d}x &= \int_\varepsilon^M \frac{f(ax)}{x}\,\mathrm{d}x - \int_\varepsilon^M \frac{f(bx)}{x}\,\mathrm{d}x \\ &= \int_{a\varepsilon}^{aM} \frac{f(x)}{x}\,\mathrm{d}x - \int_{b\varepsilon}^{bM} \frac{f(x)}{x}\,\mathrm{d}x \\ &= \int_{a\varepsilon}^{b\varepsilon} \frac{f(x)}{x}\,\mathrm{d}x - \int_{aM}^{bM} \frac{f(x)}{x}\,\mathrm{d}x.\end{aligned}$$
由积分第一中值定理知，存在位于 $a\varepsilon$ 与 $b\varepsilon$ 之间的 $\xi$ 以及位于 $aM$ 与 $bM$ 之间的 $\eta$ 使得
$$\int_{a\varepsilon}^{b\varepsilon} \frac{f(x)}{x}\,\mathrm{d}x = f(\xi)\int_{a\varepsilon}^{b\varepsilon} \frac{\mathrm{d}x}{x} = f(\xi)\log\frac{b}{a}$$
以及
$$\int_{aM}^{bM} \frac{f(x)}{x}\,\mathrm{d}x = f(\eta)\int_{aM}^{bM} \frac{\mathrm{d}x}{x} = f(\eta)\log\frac{b}{a}.$$
于是
$$\int_\varepsilon^M \frac{f(ax)-f(bx)}{x}\,\mathrm{d}x = [f(\xi)-f(\eta)]\log\frac{b}{a}.$$
现令 $\varepsilon \to 0^+$，$M \to +\infty$，则有 $\xi \to 0^+$ 及 $\eta \to +\infty$，从而得到
$$\int_0^{+\infty} \frac{f(ax)-f(bx)}{x}\,\mathrm{d}x = [f(0)-A]\log\frac{b}{a}.$$
□

上例中要求极限 $\lim\limits_{x\to+\infty} f(x)$ 存在，这对于某些连续函数 $f$ 而言是无法满足的。当这一极限不存在时，我们也可以对 $f$ 补充一些条件使得 Froullani 积分可以被计算出来，参见习题 8。

**【例 3.9】**证明：对任意的 $s>0$ 有

$$\Gamma(s)=\int_0^{+\infty} e^{-x}x^{s-1}\,\mathrm{d}x. \tag{11.10}$$

<u>证明</u>. 容易验证 (11.10) 右边的广义积分在 $s>0$ 时收敛，下面来证明它等于 $\Gamma(s)$。在第十章例 8.1 中我们定义了 $\Gamma$ 函数

$$\frac{1}{\Gamma(s)}=s\prod_{n=1}^{\infty}\Big(1+\frac{s}{n}\Big)\Big(1+\frac{1}{n}\Big)^{-s},\qquad \forall\ s\notin\mathbb{Z}_{\leqslant 0},$$

因此，当 $s>0$ 时有

$$\begin{aligned}\Gamma(s)&=\frac{1}{s}\prod_{n=1}^{\infty}\Big(1+\frac{s}{n}\Big)^{-1}\Big(1+\frac{1}{n}\Big)^{s}=\frac{1}{s}\cdot\lim_{N\to\infty}\prod_{n=1}^{N}\Big(1+\frac{s}{n}\Big)^{-1}\Big(1+\frac{1}{n}\Big)^{s}\\&=\frac{1}{s}\cdot\lim_{N\to\infty}\frac{N!}{(s+1)\cdots(s+N)}\cdot N^s.\end{aligned}$$

然而通过反复的分部积分可得

$$\frac{N!}{s(s+1)\cdots(s+N)}\cdot N^s=N^s\int_0^1(1-x)^Nx^{s-1}\,\mathrm{d}x=\int_0^N\Big(1-\frac{x}{N}\Big)^Nx^{s-1}\,\mathrm{d}x,$$

于是

$$\Gamma(s)=\lim_{N\to\infty}\int_0^N\Big(1-\frac{x}{N}\Big)^Nx^{s-1}\,\mathrm{d}x. \tag{11.11}$$

由于对任意的实数 $t$ 有 $e^t\geqslant 1+t$，故当 $x\in[0,N]$ 时有

$$e^x\geqslant\Big(1+\frac{x}{N}\Big)^N.$$

此外，由 Bernoulli 不等式（§2.4 习题 2）知当 $0\leqslant t\leqslant 1$ 时有 $(1-t)^N\geqslant 1-Nt$，故而

$$\begin{aligned}0&\leqslant e^{-x}-\Big(1-\frac{x}{N}\Big)^N=e^{-x}\Big[1-e^x\Big(1-\frac{x}{N}\Big)^N\Big]\leqslant e^{-x}\Big[1-\Big(1-\frac{x^2}{N^2}\Big)^N\Big]\\&\leqslant\frac{e^{-x}x^2}{N}.\end{aligned}$$

进而得到

$$\left|\int_0^N e^{-x}x^{s-1}\,\mathrm{d}x-\int_0^N\Big(1-\frac{x}{N}\Big)^N x^{s-1}\,\mathrm{d}x\right|\leqslant\int_0^N\frac{e^{-x}x^{s+1}}{N}\,\mathrm{d}x$$
$$<\frac{1}{N}\int_0^{+\infty}e^{-x}x^{s+1}\,\mathrm{d}x.$$

现令 $N\to\infty$，结合 (11.11) 便可得出 (11.10)。 □

## 习 题 11.3

1. 计算下列广义积分：

(1) $\displaystyle\int_{-1}^1\frac{\mathrm{d}x}{\sqrt{1-x^2}}$；　(2) $\displaystyle\int_0^{+\infty}e^{-ax}\cos bx\,\mathrm{d}x\quad(a>0)$；

(3) $\displaystyle\int_0^{+\infty}x^2e^{-x}\,\mathrm{d}x$；　(4) $\displaystyle\int_0^1 x\log x\,\mathrm{d}x$；

(5) $\displaystyle\int_1^2\frac{\mathrm{d}x}{x\sqrt{x^2-1}}$；　(6) $\displaystyle\int_0^{+\infty}\frac{\sqrt{x}}{x^2+1}\,\mathrm{d}x$；

(7) $\displaystyle\int_0^{\frac{\pi}{2}}\frac{x}{\tan x}\,\mathrm{d}x$；　(8) $\displaystyle\int_0^{+\infty}\frac{\arctan 10x-\arctan x}{x}\,\mathrm{d}x$。

2. 假设 $s>0$，利用 (11.10) 证明 $\Gamma(s+1)=s\Gamma(s)$，并由此对正整数 $n$ 得出 $\Gamma(n)=(n-1)!$。

3. 设 $\alpha\in\mathbb{R}$，计算 $\displaystyle\int_0^{+\infty}\frac{1}{(1+x^2)(1+x^\alpha)}\,\mathrm{d}x$。

4. 证明 $\displaystyle\int_0^{+\infty}\frac{\sin\Big(x-\dfrac{1}{x}\Big)}{x}\,\mathrm{d}x=0$。

5. 设 $a,b>0$，且广义积分 $\displaystyle\int_0^{+\infty}f\Big(ax+\frac{b}{x}\Big)\,\mathrm{d}x$ 收敛，证明

$$\int_0^{+\infty}f\Big(ax+\frac{b}{x}\Big)\,\mathrm{d}x=\frac{1}{a}\int_0^{+\infty}f\big(\sqrt{x^2+4ab}\big)\,\mathrm{d}x.$$

6. 设 $p,q>0$，讨论积分 $\displaystyle\int_0^{+\infty}\frac{\sin x^p}{x^q}\,\mathrm{d}x$ 的敛散性。

7. 设 $p>0$，研究广义积分 $\displaystyle\int_0^{+\infty}\frac{\sin\Big(x+\dfrac{1}{x}\Big)}{x^p}\,\mathrm{d}x$ 的绝对收敛性和条件收敛性。

8. 设 $f(x)$ 在 $[0,+\infty)$ 上连续，且广义积分 $\displaystyle\int_1^{+\infty}\frac{f(x)}{x}\,\mathrm{d}x$ 收敛，证明：对任意的正实数 $a, b$ 有

$$\int_0^{+\infty}\frac{f(ax)-f(bx)}{x}\,\mathrm{d}x=f(0)\log\frac{b}{a}.$$

并利用这一结果计算 $\displaystyle\int_0^{+\infty}\frac{\cos ax-\cos bx}{x}\,\mathrm{d}x$。

9. 证明 $\displaystyle\int_0^1\frac{\log(1-x)}{x}\,\mathrm{d}x=-\zeta(2)$。

10. 证明 $\displaystyle\int_0^1 x^{-x}\,\mathrm{d}x=\sum_{n=1}^{\infty}\frac{1}{n^n}$。

## § 11.4 求和与积分之间的联系

在第八章中我们知道，定积分是通过 Riemann 和来定义的，因此一个很自然的想法是将求和与积分联系起来，确切地说，对于那些具有良好性态的函数 $f$ 而言，$\displaystyle\sum_{a<n\leqslant b}f(n)$ 与 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 应该相差不多。本节的目的就是去建立这两者之间的联系。

我们首先给出分部求和的一个非常有用的积分形态，这一结论也被称作分部求和。

**【定理 4.1】** 设 $(a_n)_{n\in\mathbb{Z}}$ 是 $\mathbb{R}$ 的一个元素族。又设 $a<b$，并对任意的 $t\in[a,b]$ 记 $S(t)=\displaystyle\sum_{a<n\leqslant t}a_n$。则对任意的 $f\in C^1([a,b])$ 有

$$\sum_{a<n\leqslant b}a_nf(n)=S(b)f(b)-\int_a^b S(t)f'(t)\,\mathrm{d}t. \tag{11.12}$$

<u>证明</u>. 记 $[a]=M$，$[b]=N$，则由 Abel 求和公式（第九章定理 3.1）知

$$\begin{aligned}\sum_{a<n\leqslant b}a_nf(n)&=\sum_{M<n\leqslant N}a_nf(n)=S(N)f(N)+\sum_{M<n\leqslant N-1}S(n)(f(n)-f(n+1))\\&=S(N)f(N)-\sum_{M\leqslant n\leqslant N-1}S(n)\int_n^{n+1}f'(t)\,\mathrm{d}t\\&=S(N)f(N)-\sum_{M\leqslant n\leqslant N-1}\int_n^{n+1}S(t)f'(t)\,\mathrm{d}t\end{aligned}$$

$$= S(N)f(N) - \int_M^N S(t)f'(t)\,\mathrm{d}t.$$

由 $S(t)$ 的定义知，当 $t \in [M,a]$ 时 $S(t)=0$，故

$$\int_M^a S(t)f'(t)\,\mathrm{d}t = 0.$$

另一方面，

$$\int_N^b S(t)f'(t)\,\mathrm{d}t = S(b)\int_N^b f'(t)\,\mathrm{d}t = S(b)(f(b)-f(N)) = S(b)f(b) - S(N)f(N).$$

综上便得定理。 □

**【推论 4.2】**设 $\{a_n\}$ 是一个数列，并记 $S(t)=\sum\limits_{n\leqslant t} a_n$。又设 $x\geqslant 1$。则对任意的 $f\in C^1([0,x])$ 有

$$\sum_{n\leqslant x} a_n f(n) = S(x)f(x) - \int_1^x S(t)f'(t)\,\mathrm{d}t. \tag{11.13}$$

证明. 不妨设 $x>1$。在定理 4.1 中取 $a=1$，$b=x$ 即得

$$\begin{aligned}\sum_{n\leqslant x} a_n f(n) &= a_1 f(1) + \sum_{1<n\leqslant x} a_n f(n)\\ &= a_1 f(1) + (S(x)-a_1)f(x) - \int_1^x (S(t)-a_1)f'(t)\,\mathrm{d}t\\ &= S(x)f(x) - \int_1^x S(t)f'(t)\,\mathrm{d}t.\end{aligned}$$

□

定理 4.1 实质上已经建立了求和与积分之间的联系。但我们可以通过选取特殊的 $a_n$ 将结果写的更加明确一些。

**【定理 4.3】（Euler 求和公式）** 设 $a<b$，则对任意的 $f\in C^1([a,b])$ 有

$$\sum_{a<n\leqslant b} f(n) = \int_a^b f(t)\,\mathrm{d}t + \int_a^b f'(t)\psi(t)\,\mathrm{d}t + f(a)\psi(a) - f(b)\psi(b), \tag{11.14}$$

其中 $\psi(x)=x-[x]-\dfrac{1}{2}$，且 $[x]$ 表示不超过 $x$ 的最大整数。

证明. 在定理 4.1 中取 $a_n=1$ $(\forall\, n>a)$，则 $S(t)=[t]-[a]$，于是

$$\sum_{a<n\leqslant b} f(n) = ([b]-[a])f(b) - \int_a^b ([t]-[a])f'(t)\,\mathrm{d}t$$

$$
\begin{aligned}
&= [b]f(b) - [a]f(a) - \int_a^b [t]f'(t)\,\mathrm{d}t \\
&= f(a)\psi(a) - f(b)\psi(b) + \int_a^b f'(t)\psi(t)\,\mathrm{d}t \\
&\qquad - \Big(a - \frac{1}{2}\Big)f(a) + \Big(b - \frac{1}{2}\Big)f(b) - \int_a^b \Big(t - \frac{1}{2}\Big)f'(t)\,\mathrm{d}t.
\end{aligned}
$$

由分部积分知

$$
-\int_a^b \Big(t - \frac{1}{2}\Big)f'(t)\,\mathrm{d}t = \Big(a - \frac{1}{2}\Big)f(a) - \Big(b - \frac{1}{2}\Big)f(b) + \int_a^b f(t)\,\mathrm{d}t,
$$

从而定理得证。 □

**【例 4.4】**对任意的 $x > 1$，由 Euler 求和公式知

$$
\begin{aligned}
\sum_{n\leqslant x}\frac{1}{n} &= 1 + \sum_{1<n\leqslant x}\frac{1}{n} = 1 + \int_1^x \frac{\mathrm{d}t}{t} - \int_1^x \frac{\psi(t)}{t^2}\,\mathrm{d}t - \frac{\psi(x)}{x} - \frac{1}{2} \\
&= \log x - \int_1^x \frac{\psi(t)}{t^2}\,\mathrm{d}t + \frac{1}{2} - \frac{\psi(x)}{x} \\
&= \log x + \frac{1}{2} - \int_1^\infty \frac{\psi(t)}{t^2}\,\mathrm{d}t + O\Big(\frac{1}{x}\Big) = \log x + \gamma + O\Big(\frac{1}{x}\Big),
\end{aligned}
$$

其中

$$
\gamma = \frac{1}{2} - \int_1^\infty \frac{\psi(t)}{t^2}\,\mathrm{d}t
$$

是一个常数，也即是 Euler 常数。

**【例 4.5】**在这个例子中，我们来给出 $n!$ 的渐近公式。

由 Euler 求和公式知，对任意的 $n \geqslant 2$ 有

$$
\begin{aligned}
\log n! = \sum_{1<k\leqslant n}\log k &= \int_1^n \log t\,\mathrm{d}t + \int_1^n \frac{\psi(t)}{t}\,\mathrm{d}t + \frac{1}{2}\log n \\
&= n\log n - n + 1 + \frac{1}{2}\log n + \int_1^n \frac{\psi(t)}{t}\,\mathrm{d}t.
\end{aligned}
$$

因为 $\psi(t)$ 以 1 为周期，且

$$
\int_0^1 \psi(t)\,\mathrm{d}t = \int_0^1 \Big(t - [t] - \frac{1}{2}\Big)\,\mathrm{d}t = \int_0^1 \Big(t - \frac{1}{2}\Big)\,\mathrm{d}t = 0,
$$

所以 $F(A)=\int_1^A\psi(t)\,\mathrm{d}t$ 在 $[1,+\infty)$ 上有界，于是由 Dirichlet 判别法（定理 1.18）知 $\int_1^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t$ 收敛。因此

$$\log n!=n\log n-n+1+\frac{1}{2}\log n+\int_1^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t-\int_n^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t.$$

对任意的 $N>n$，由积分第二中值定理知存在 $\xi\in[n,N]$ 使得

$$\int_n^N\frac{\psi(t)}{t}\,\mathrm{d}t=\frac{1}{n}\int_n^{\xi}\psi(t)\,\mathrm{d}t\ll\frac{1}{n},$$

令 $N\to+\infty$ 即得

$$\int_n^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t=O\Big(\frac{1}{n}\Big).\tag{11.15}$$

故而

$$\log n!=n\log n-n+1+\frac{1}{2}\log n+\int_1^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t+O\Big(\frac{1}{n}\Big).$$

若记 $\log A=2+2\int_1^{+\infty}\frac{\psi(t)}{t}\,\mathrm{d}t$，那么

$$n!=n^ne^{-n}\sqrt{An}\bigg(1+O\Big(\frac{1}{n}\Big)\bigg).\tag{11.16}$$

下面来计算 $A$ 的值。回忆起在第八章例 5.12 中我们得到了

$$I_n=\int_0^{\frac{\pi}{2}}\sin^n x\,\mathrm{d}x=\begin{cases}\dfrac{(2k-1)!!}{(2k)!!}\cdot\dfrac{\pi}{2}, & 若\ n=2k,\\[2ex] \dfrac{(2k-2)!!}{(2k-1)!!}, & 若\ n=2k-1.\end{cases}$$

一方面，由 (11.16) 可得

$$\begin{aligned}I_{2k}&=\frac{(2k)!}{[(2k)!!]^2}\cdot\frac{\pi}{2}=\frac{(2k)!}{2^{2k}(k!)^2}\cdot\frac{\pi}{2}=\frac{(2k)^{2k}e^{-2k}\sqrt{2Ak}}{2^{2k}k^{2k}e^{-2k}Ak}\cdot\frac{\pi}{2}\bigg(1+O\Big(\frac{1}{n}\Big)\bigg)\\&=\sqrt{\frac{2}{Ak}}\cdot\frac{\pi}{2}\bigg(1+O\Big(\frac{1}{n}\Big)\bigg).\end{aligned}$$

以及

$$I_{2k-1}=\sqrt{\frac{A}{8k}}\bigg(1+O\Big(\frac{1}{n}\Big)\bigg).$$

因此

$$\lim_{k\to\infty}\frac{I_{2k}}{I_{2k-1}}=\frac{2\pi}{A}.\tag{11.17}$$

另一方面，由 $I_n$ 的定义知 $I_n \leqslant I_{n-1}$，因此结合 (8.31) 式有

$$\frac{n}{n+1} = \frac{I_{n+1}}{I_{n-1}} \leqslant \frac{I_n}{I_{n-1}} \leqslant 1,$$

故而 $\lim\limits_{n\to\infty} \dfrac{I_n}{I_{n-1}} = 1$。注意到 (11.17)，我们有 $A = 2\pi$。因此最终得到

$$n! = e^{n\log n - n}\sqrt{2\pi n}\Big(1 + O\Big(\frac{1}{n}\Big)\Big). \tag{11.18}$$

这个公式被称作 Stirling 公式。

## 习 题 11.4

1. 利用命题 1.12 证明 (11.15)。

2. 设 $\{a_n\}$ 是一个数列，$A(x) = \sum\limits_{n\leqslant x} a_n$。证明：对任意的 $x > 1$ 有

$$\sum_{n\leqslant x} a_n \log\frac{x}{n} = \int_1^x \frac{A(t)}{t}\,\mathrm{d}t.$$

3. 设 $x \geq 2$，证明：

$$\sum_{n\leq x} \frac{\log n}{n} = \frac{1}{2}\log^2 x + A + O\Big(\frac{\log x}{x}\Big),$$

   其中 $A$ 是一个常数。

4. 证明

$$\sum_{n=1}^{\infty}(-1)^n\frac{\log n}{n} = \gamma\log 2 - \frac{\log^2 2}{2},$$

   其中 $\gamma$ 是 Euler 常数。

5. 证明：当 $s \to 1^+$ 时有

$$\log\zeta(s) = -\log(s-1) + O(s-1).$$

6. （Stirling 公式）证明：当 $x \geqslant 1$ 时有

$$\log\Gamma(x) = \Big(x - \frac{1}{2}\Big)\log x - x + \frac{1}{2}\log 2\pi + O\Big(\frac{1}{x}\Big).$$

从第 7 题到第 11 题是一组题，介绍 Euler–Maclaurin 求和公式及其应用。

7. 设 $B_n(x)$ 是以 1 为周期的函数，且满足

$$\begin{aligned} &B_0(x) = 1, \quad \forall\, x \in [0,1), \\ &B'_{n+1}(x) = (n+1)B_n(x), \quad \forall\, x \in [0,1),\, n \geq 0, \\ &\int_0^1 B_n(x)\,\mathrm{d}x = 0, \quad \forall\, n \geq 1. \end{aligned}$$

我们称 $B_n(x)$ 为 Bernoulli 多项式。证明

$$\begin{aligned} B_1(x) &= \{x\} - \frac{1}{2}, \\ B_2(x) &= \{x\}^2 - \{x\} + \frac{1}{6}, \\ B_3(x) &= \{x\}^3 - \frac{3}{2}\{x\}^2 + \frac{1}{2}\{x\}, \\ B_4(x) &= \{x\}^4 - 2\{x\}^3 + \{x\}^2 - \frac{1}{30}. \end{aligned}$$

其中 $\{x\} = x - [x]$ 表示 $x$ 的小数部分。

8. 设 $x \in [0,1)$。证明在 $t = 0$ 的某去心邻域内

$$\sum_{n=0}^{\infty} \frac{B_n(x)}{n!} t^n$$

收敛于 $\dfrac{te^{tx}}{e^t - 1}$，并由此推出 $B_n(0)$ 也即是 §10.6 习题 6 中所定义的 Bernoulli 数 $B_n$。

9. （Euler–Maclaurin 求和公式）设 $a, b$ 均是整数，$f \in C^{k+1}([a,b])$ $(k \in \mathbb{Z}_{\geqslant 0})$，证明

$$\begin{aligned} \sum_{a<n\leqslant b} f(n) = \int_a^b f(x)\,\mathrm{d}x + \sum_{j=1}^{k+1} (-1)^j \frac{B_j(0)}{j!}\big(f^{(j-1)}(b) - f^{(j-1)}(a)\big) \\ + \frac{(-1)^k}{(k+1)!}\int_a^b B_{k+1}(x) f^{(k+1)}(x)\,\mathrm{d}x. \end{aligned}$$

10. 设 $N$ 是一个正整数，证明存在 $\theta \in [0,1]$ 使得

$$\sum_{n\leqslant N} \frac{1}{n} = \log N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \frac{\theta}{60N^4},$$

11. 利用上题结论计算 Euler 常数 $\gamma$ 至小数点后五位。

第十二章

# 多元函数的极限

从具体到抽象是数学发展的一条重要大道，因此具体的例子往往是抽象概念的源泉，而所用的方法也往往是高深数学里所用的方法的依据。仅仅熟读了抽象的定义和方法而不知道他们具体来源的数学工作者是没有发展前途的，这样的人要搞深刻研究是可能会遇到无法克服的难关的。

—— 华罗庚

## § 12.1 $\mathbb{R}^n$ 中的点集

按照 (1.2) 式，

$$\mathbb{R}^n = \{(x_1, \cdots, x_n) : x_j \in \mathbb{R}\ (1 \leqslant j \leqslant n)\},$$

但有时出于习惯考虑，也常将 $\mathbb{R}^n$ 中的元素按照列向量 $(x_1, \cdots, x_n)^{\mathrm{T}}$ 的形式写出。在线性代数中我们学习了 $\mathbb{R}^n$ 的代数结构，本节的目的是去初步了解其拓扑结构。

对任意的 $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$，记

$$|\boldsymbol{x}| = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}},$$

这被称作 $\boldsymbol{x}$ 的范数 (norm)。我们可以用 $|\boldsymbol{x} - \boldsymbol{y}|$ 表示 $\mathbb{R}^n$ 中两点 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 之间的距离 (distance)[1]，这是我们常用的直线或平面上两点之间距离的推广。特别地，$|\boldsymbol{x}|$ 就

[1]无论是“范数”还是“距离”均有其严格的数学定义，而我们在这里的设置均是满足这些数学定义的。对此我们不作详细说明，读者只需承认这样的称谓即可。当然，不满足于此的读者可查阅本节习题中“距离空间”的部分。

表示 $\boldsymbol{x}$ 与 $\mathbf{0}=(0,\cdots,0)$ 之间的距离。容易看出

$$|\lambda\boldsymbol{x}|=|\lambda|\cdot|\boldsymbol{x}|,\qquad\qquad \forall\ \lambda\in\mathbb{R}.$$

由 Minkowski 不等式（第二章推论 4.4）知，对任意的 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{R}^n$ 有

$$|\boldsymbol{x}-\boldsymbol{z}|\leqslant|\boldsymbol{x}-\boldsymbol{y}|+|\boldsymbol{y}-\boldsymbol{z}|,$$

这被称作是 $\mathbb{R}^n$ 中的三角形不等式。此外，如果用 $\langle\boldsymbol{x},\boldsymbol{y}\rangle$ 表示 $\boldsymbol{x}=(x_1,\cdots,x_n)$ 与 $\boldsymbol{y}=(y_1,\cdots,y_n)$ 的内积，也即

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle=\sum_{j=1}^n x_jy_j,$$

那么由 Cauchy–Schwarz 不等式（第二章定理 4.3）可得

$$|\langle\boldsymbol{x},\boldsymbol{y}\rangle|\leqslant|\boldsymbol{x}|\cdot|\boldsymbol{y}|.$$

读者需要牢记上面的这些关系式，以后它们会被经常使用，但我们不会再重复说明其出处。

### 12.1.1 邻域，开集

**【定义 1.1】** 设 $\boldsymbol{a}\in\mathbb{R}^n$，$\varepsilon$ 是一个正实数，我们称集合

$$\{\boldsymbol{x}\in\mathbb{R}^n:|\boldsymbol{x}-\boldsymbol{a}|<\varepsilon\}$$

为 $\boldsymbol{a}$ 的 $\varepsilon$-邻域，记作 $B(\boldsymbol{a},\varepsilon)$。称

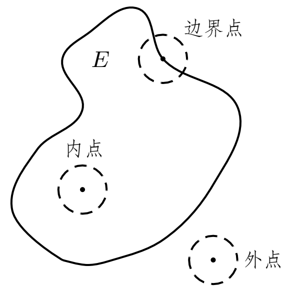
$$B(\boldsymbol{a},\varepsilon)\setminus\{\boldsymbol{a}\}=\{\boldsymbol{x}\in\mathbb{R}^n:0<|\boldsymbol{x}-\boldsymbol{a}|<\varepsilon\}$$

为 $\boldsymbol{a}$ 的去心邻域。

**【定义 1.2】** 设 $E\subseteq\mathbb{R}^n$ 且 $\boldsymbol{a}\in E$。若存在 $\varepsilon>0$ 使得 $B(\boldsymbol{a},\varepsilon)\subseteq E$，则称 $\boldsymbol{a}$ 为 $E$ 的内点 (interior point)。$E$ 的全体内点所成之集被称作 $E$ 的内部 (interior)，记作 $E^\circ$。

**【定义 1.3】** 设 $E\subseteq\mathbb{R}^n$。若 $\boldsymbol{a}$ 是 $E^c$ 的内点，则称 $\boldsymbol{a}$ 为 $E$ 的外点 (exterior point)。$E$ 的全体外点所成之集被称作 $E$ 的外部 (exterior)。

显然，$\boldsymbol{a}$ 是 $E$ 的外点当且仅当存在 $\varepsilon>0$ 使得 $B(\boldsymbol{a},\varepsilon)\cap E=\varnothing$。



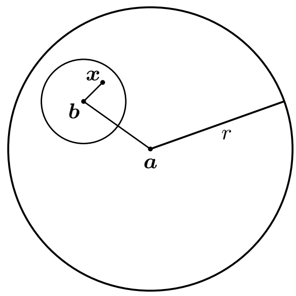

【定义 1.4】设 $E \subseteq \mathbb{R}^n$。若 $\boldsymbol{a}$ 既不是 $E$ 的内点也不是 $E$ 的外点，则称 $\boldsymbol{a}$ 为 $E$ 的边界点 (boundary point)。$E$ 的全体边界点所成之集被称作 $E$ 的边界 (boundary)，记作 $\partial E$。

易见，$\boldsymbol{a} \in \partial E$ 当且仅当对任意的 $\varepsilon > 0$ 有

$$B(\boldsymbol{a},\varepsilon) \cap E \neq \varnothing \qquad \text{且} \qquad B(\boldsymbol{a},\varepsilon) \cap E^c \neq \varnothing.$$

并且由此可以看出 $\partial E = \partial E^c$。

【定义 1.5】设 $G \subseteq \mathbb{R}^n$，若 $G$ 中每个点均是内点，则称 $G$ 是 $\mathbb{R}^n$ 中的开集 (open set)[2]。换句话说，$G$ 是开集当且仅当 $G = G^\circ$。




【例 1.6】设 $\boldsymbol{a} \in \mathbb{R}^n$，$r > 0$，证明 $B(\boldsymbol{a},r)$ 是开集。特别地，$\mathbb{R}$ 中的每个有界开区间均是开集。

证明. 对任意的 $\boldsymbol{b} \in B(\boldsymbol{a},r)$ 有 $|\boldsymbol{b}-\boldsymbol{a}| < r$。现取 $\delta < r - |\boldsymbol{b}-\boldsymbol{a}|$，由三角形不等式知，对任意的 $\boldsymbol{x} \in B(\boldsymbol{b},\delta)$ 有

$$|\boldsymbol{x}-\boldsymbol{a}| \leqslant |\boldsymbol{x}-\boldsymbol{b}| + |\boldsymbol{b}-\boldsymbol{a}| < \delta + |\boldsymbol{b}-\boldsymbol{a}| < r,$$

因此 $B(\boldsymbol{b},\delta) \subseteq B(\boldsymbol{a},r)$。这说明 $B(\boldsymbol{a},r)$ 是开集。

此外，由于 $\mathbb{R}$ 中的有限开区间 $(a,b)$ 也即是 $B\left(\dfrac{a+b}{2},\dfrac{b-a}{2}\right)$，所以它是 $\mathbb{R}$ 中的开集。 □

【命题 1.7】我们有

(1) $\varnothing$ 和 $\mathbb{R}^n$ 都是开集。

(2) 设 $(G_\lambda)_{\lambda\in L}$ 是一族开集，则 $\bigcup\limits_{\lambda\in L} G_\lambda$ 也是开集。

(3) 设 $G_1,\cdots,G_n$ 是开集，则 $\bigcap\limits_{j=1}^{n} G_j$ 也是开集。

证明. (1) 是显然的。

(2) 对任意的 $\boldsymbol{a} \in \bigcup\limits_{\lambda\in L} G_\lambda$，必存在 $\lambda_0 \in L$ 使得 $\boldsymbol{a} \in G_{\lambda_0}$。由于 $G_{\lambda_0}$ 是开集，故存在 $\varepsilon > 0$ 使得 $B(\boldsymbol{a},\varepsilon) \subseteq G_{\lambda_0}$，从而 $B(\boldsymbol{a},\varepsilon) \subseteq \bigcup\limits_{\lambda\in L} G_\lambda$，这说明 $\boldsymbol{a}$ 是 $\bigcup\limits_{\lambda\in L} G_\lambda$ 的内点。再由 $\boldsymbol{a}$ 的任意性知 $\bigcup\limits_{\lambda\in L} G_\lambda$ 是开集。

[2] 我们常用 $G$ 表示开集，它是德文“Gebiet”的首字母。

(3) 按照数学归纳法，只需证 $n=2$ 的情形即可。对任意的 $\boldsymbol{a}\in G_1\cap G_2$，我们有 $\boldsymbol{a}\in G_1$ 且 $\boldsymbol{a}\in G_2$。由于 $G_1$ 与 $G_2$ 都是开集，故存在正实数 $\varepsilon_1$ 和 $\varepsilon_2$ 使得 $B(\boldsymbol{a},\varepsilon_1)\subseteq G_1$ 且 $B(\boldsymbol{a},\varepsilon_2)\subseteq G_2$，现取 $\varepsilon=\min(\varepsilon_1,\varepsilon_2)$，则

$$B(\boldsymbol{a},\varepsilon)\subseteq B(\boldsymbol{a},\varepsilon_1)\subseteq G_1 \qquad \text{且} \qquad B(\boldsymbol{a},\varepsilon)\subseteq B(\boldsymbol{a},\varepsilon_2)\subseteq G_2,$$

从而 $B(\boldsymbol{a},\varepsilon)\subseteq G_1\cap G_2$，这就证明了 $\boldsymbol{a}$ 是 $G_1\cap G_2$ 的内点。再由 $\boldsymbol{a}$ 的任意性知 $G_1\cap G_2$ 是开集。 □

需注意的是，无穷多个开集的交未必是开集。例如，对任意的正整数 $n$ 而言 $\left(-\dfrac{1}{n},\dfrac{1}{n}\right)$ 均是 $\mathbb{R}$ 中的开集，但 $\bigcap\limits_{n=1}^{\infty}\left(-\dfrac{1}{n},\dfrac{1}{n}\right)=\{0\}$ 不是 $\mathbb{R}$ 中的开集。

**【定义 1.8】**设 $E\subseteq\mathbb{R}^n$，若开集 $G$ 满足 $E\subseteq G$，则称 $G$ 是 $E$ 的一个邻域。特别地，当 $E=\{\boldsymbol{a}\}$ 时我们称 $G$ 是 $\boldsymbol{a}$ 的一个邻域。

### 12.1.2　聚点，闭集

**【定义 1.9】**设 $F\subseteq\mathbb{R}^n$。若 $F^c$ 是 $\mathbb{R}^n$ 中的开集，则称 $F$ 是 $\mathbb{R}^n$ 中的闭集 (closed set)[③]。

**【命题 1.10】**我们有

(1) $\varnothing$ 和 $\mathbb{R}^n$ 都是闭集。

(2) 设 $(F_\lambda)_{\lambda\in L}$ 是一族闭集，则 $\bigcap\limits_{\lambda\in L}F_\lambda$ 也是闭集。

(3) 设 $F_1,\cdots,F_n$ 是闭集，则 $\bigcup\limits_{j=1}^{n}F_j$ 也是闭集。

证明. 这可由命题 1.7 及 de Morgan 律（第一章定理 3.13）得到。 □

同样需要注意的是，无穷多个闭集的并未必是闭集。

**【定义 1.11】**设 $E\subseteq\mathbb{R}^n$，$\boldsymbol{a}\in\mathbb{R}^n$。若对任意的 $\varepsilon>0$ 均有

$$(B(\boldsymbol{a},\varepsilon)\setminus\{\boldsymbol{a}\})\cap E\neq\varnothing,$$

则称 $\boldsymbol{a}$ 是 $E$ 的聚点 (cluster point) 或极限点 (limit point)。称 $E$ 的全体聚点所成之集为 $E$ 的导集 (derived set)，记作 $E'$。如果 $\boldsymbol{b}\in E\setminus E'$，则称 $\boldsymbol{b}$ 是 $E$ 的孤立点 (isolated point)。此外，称 $E\cup E'$ 为 $E$ 的闭包 (closure)，记作 $\overline{E}$。

---

③我们常用 $F$ 表示闭集，它是法文“fermé”的首字母。

容易看出，$\boldsymbol{a}$ 是 $E$ 的聚点当且仅当对于 $\boldsymbol{a}$ 的任一邻域 $U$ 均有

$$(U \setminus \{\boldsymbol{a}\}) \cap E \neq \varnothing.$$

**【例 1.12】** (1) 若 $E = \left\{\dfrac{1}{n} : n \in \mathbb{Z}_{>0}\right\} \subseteq \mathbb{R}$，则 $E' = \{0\}$，且 $E$ 中每个点均是孤立点。

(2) $\overline{\mathbb{Q}} = \mathbb{R}$。

(3) $\overline{B(\boldsymbol{a},\varepsilon)} = \{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x} - \boldsymbol{a}| \leqslant \varepsilon\}$。

**【命题 1.13】设 $E \subseteq \mathbb{R}^n$，则 $E$ 是闭集当且仅当 $E = \overline{E}$。**

证明. 由 $\overline{E}$ 的定义知 $E = \overline{E}$ 当且仅当 $E' \subseteq E$，这又等价于 $E' \cap E^c = \varnothing$。因为 $E'$ 是 $E$ 的全体聚点所成之集，故 $E' \cap E^c = \varnothing$ 当且仅当对任意的 $\boldsymbol{a} \in E^c$，均存在 $\varepsilon > 0$ 使得 $B(\boldsymbol{a},\varepsilon) \cap E = \varnothing$，也即 $B(\boldsymbol{a},\varepsilon) \subseteq E^c$，而这等价于 $E^c$ 是开集。 □

**【命题 1.14】设 $E \subseteq \mathbb{R}^n$，则 $\overline{E} = E^\circ \cup \partial E$。**

证明. 首先证 $\overline{E} \subseteq E^\circ \cup \partial E$。为此，只需对任意的 $\boldsymbol{a} \in \overline{E} \setminus E^\circ$ 说明 $\boldsymbol{a} \in \partial E$ 即可。一方面，由 $a \in \overline{E}$ 知，对任意的 $\varepsilon > 0$ 有 $B(\boldsymbol{a},\varepsilon) \cap E \neq \varnothing$；另一方面，由 $\boldsymbol{a} \notin E^\circ$ 又知，对任意的 $\varepsilon > 0$ 而言 $B(\boldsymbol{a},\varepsilon)$ 均不会包含于 $E$，也即 $B(\boldsymbol{a},\varepsilon) \cap E^c \neq \varnothing$。因此 $\boldsymbol{a} \in \partial E$。

其次证 $\overline{E} \supseteq E^\circ \cup \partial E$。注意到 $E^\circ \subseteq E \subseteq \overline{E}$，故只需证明 $\partial E \setminus E \subseteq \overline{E}$ 即可。现设 $\boldsymbol{a} \in \partial E \setminus E$，由 $\boldsymbol{a} \in \partial E$ 知，对任意的 $\varepsilon > 0$ 有 $B(\boldsymbol{a},\varepsilon) \cap E \neq \varnothing$，注意到 $\boldsymbol{a} \notin E$，因此 $\big(B(\boldsymbol{a},\varepsilon) \setminus \{a\}\big) \cap E \neq \varnothing$，所以 $a \in E'$，进而 $a \in \overline{E}$。 □

**【定义 1.15】若 $\mathbb{R}^n$ 的子集 $E$ 满足 $\overline{E} = \mathbb{R}^n$，则称 $E$ 在 $\mathbb{R}^n$ 中**稠密 (dense)。

**【例 1.16】** $\mathbb{Q}$ 及 $\mathbb{R} \setminus \mathbb{Q}$ 均在 $\mathbb{R}$ 中稠密。

若 $\boldsymbol{a}$ 是 $E$ 的聚点，则必存在 $E$ 中的序列 $\{\boldsymbol{x}_m\}$，使得对任意的 $\varepsilon > 0$，均存在正整数 $N$ 满足

$$|\boldsymbol{x}_m - \boldsymbol{a}| < \varepsilon, \qquad \forall\, m > N.$$

此时我们称 $\{\boldsymbol{x}_m\}$ 收敛于 $\boldsymbol{a}$，记作 $\lim\limits_{m\to\infty} \boldsymbol{x}_m = \boldsymbol{a}$ 或 $\boldsymbol{x}_m \to \boldsymbol{a}$（当 $n \to \infty$ 时）。

若 $\mathbb{R}^n$ 中的序列 $\{\boldsymbol{x}_m\}$ 满足：对任意的 $\varepsilon > 0$，均存在正整数 $N$ 使得

$$|\boldsymbol{x}_\ell - \boldsymbol{x}_m| < \varepsilon, \qquad \forall\, \ell, m > N,$$

则称 $\{\boldsymbol{x}_m\}$ 是 Cauchy 列。

现记 $\boldsymbol{x}_m=(x_1^{(m)},\cdots,x_n^{(m)})$，那么容易证明 $\{\boldsymbol{x}_m\}$ 收敛于 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 当且仅当对任意的 $1\leqslant j\leqslant n$ 有 $\lim\limits_{m\to\infty}x_j^{(m)}=a_j$。结合数列的 Cauchy 收敛准则便可立即得到下面的定理。

**【定理 1.17】（Cauchy 收敛准则）** $\mathbb{R}^n$ 中的序列 $\{\boldsymbol{x}_m\}$ 收敛当且仅当它是 Cauchy 列。

**【例 1.18】（压缩影像原理）** 设 $E$ 是 $\mathbb{R}^n$ 中的闭集，$f:E\longrightarrow E$。如果存在 $\theta\in(0,1)$ 使得

$$|f(\boldsymbol{x})-f(\boldsymbol{y})|\leqslant\theta|\boldsymbol{x}-\boldsymbol{y}|,\qquad\forall\ \boldsymbol{x},\boldsymbol{y}\in E,$$

那么存在唯一的 $\boldsymbol{a}\in E$ 使得 $f(\boldsymbol{a})=\boldsymbol{a}$。我们称 $\boldsymbol{a}$ 为 $f$ 的<u>不动点 (fixed point)</u>。

<u>证明</u>. 存在性：任取 $\boldsymbol{x}_0\in E$，按照 $\boldsymbol{x}_m=f(\boldsymbol{x}_{m-1})$ $(m\in\mathbb{Z}_{>0})$ 递归地定义 $E$ 中的点列 $\{\boldsymbol{x}_m\}$，那么对任意的正整数 $m$ 有

$$|\boldsymbol{x}_{m+1}-\boldsymbol{x}_m|=|f(\boldsymbol{x}_m)-f(\boldsymbol{x}_{m-1})|\leqslant\theta|\boldsymbol{x}_m-\boldsymbol{x}_{m-1}|,$$

从而可归纳地证得

$$|\boldsymbol{x}_{m+1}-\boldsymbol{x}_m|\leqslant\theta^m|\boldsymbol{x}_1-\boldsymbol{x}_0|,\qquad\forall\ m\in\mathbb{Z}_{>0}.$$

于是对任意的 $m>k$ 有

$$|\boldsymbol{x}_m-\boldsymbol{x}_k|\leqslant\sum_{j=k}^{m-1}|\boldsymbol{x}_{j+1}-\boldsymbol{x}_j|\leqslant\sum_{j=k}^{m-1}\theta^j|\boldsymbol{x}_1-\boldsymbol{x}_0|\leqslant\frac{\theta^k}{1-\theta}|\boldsymbol{x}_1-\boldsymbol{x}_0|,$$

由此可知 $\{\boldsymbol{x}_m\}$ 是 Cauchy 列，进而收敛。现记 $\lim\limits_{m\to\infty}\boldsymbol{x}_m=\boldsymbol{a}$，则由 $E$ 是闭集知 $\boldsymbol{a}\in E$。从而 $|f(\boldsymbol{x}_m)-f(\boldsymbol{a})|\leqslant\theta|\boldsymbol{x}_m-\boldsymbol{a}|$，这说明 $\lim\limits_{m\to\infty}f(\boldsymbol{x}_m)=f(\boldsymbol{a})$。现对等式 $f(x_{m-1})=\boldsymbol{x}_m$ 两边令 $m\to\infty$ 即得 $f(\boldsymbol{a})=\boldsymbol{a}$。

唯一性：若 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 均是 $f$ 在 $E$ 中的不动点，则

$$|\boldsymbol{a}-\boldsymbol{b}|=|f(\boldsymbol{a})-f(\boldsymbol{b})|\leqslant\theta|\boldsymbol{a}-\boldsymbol{b}|,$$

于是必有 $\boldsymbol{a}=\boldsymbol{b}$。 □

类似于闭区间套定理，我们有下面的闭矩形套定理。在引入这一定理之前，先介绍两个概念。我们称形如 $[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ 的集合为 $\mathbb{R}^n$ 中的<u>闭矩形</u>。此外，对 $\mathbb{R}^n$ 的任一非空子集 $E$，我们记

$$\operatorname{diam}(E)=\sup_{\boldsymbol{x},\boldsymbol{y}\in E}|\boldsymbol{x}-\boldsymbol{y}|,$$

并称之为 $E$ 的<u>直径 (diameter)</u>。如果 $E=B(\boldsymbol{a},r)$，那么容易验证 $\operatorname{diam}(E)=2r$。

**【定理 1.19】（闭矩形套定理）** 设闭矩形列 $\{I_m\}$ 满足 $I_{m+1} \subseteq I_m$ $(\forall\, m \in \mathbb{Z}_{>0})$ 以及 $\lim\limits_{m\to\infty} \operatorname{diam}(I_m) = 0$，那么存在唯一的 $\boldsymbol{a} \in \mathbb{R}^n$ 使得

$$\bigcap_{m=1}^{\infty} I_m = \{\boldsymbol{a}\}.$$

我们把这一定理的证明留作练习。

以上定理 1.17 和 1.19 均是 $\mathbb{R}^n$ 的完备性的体现。

### 12.1.3　紧集

**【定义 1.20】** 设 $E \subseteq \mathbb{R}^n$ 且 $(G_\lambda)_{\lambda\in L}$ 是 $\mathbb{R}^n$ 中的一族开集。若

$$E \subseteq \bigcup_{\lambda\in L} G_\lambda,$$

则称 $(G_\lambda)_{\lambda\in L}$ 是 $E$ 的一个开覆盖。此外，若 $L' \subseteq L$ 满足

$$E \subseteq \bigcup_{\lambda\in L'} G_\lambda,$$

则称 $(G_\lambda)_{\lambda\in L'}$ 是一个子覆盖。

**【定义 1.21】** 我们称 $\mathbb{R}^n$ 的子集 $K$ 是一个紧集 (compact set)[④]，如果 $K$ 的每个开覆盖均有有限子覆盖。

**【命题 1.22】** $\mathbb{R}^n$ 中的闭矩形 $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ 是紧集。

证明. 记 $I_1 = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$。反设存在 $I_1$ 的一个开覆盖 $(G_\lambda)_{\lambda\in L}$，它没有有限子覆盖。现用 $n$ 个超平面[⑤] $x_j = \dfrac{a_j + b_j}{2}$ $(1 \leqslant j \leqslant n)$ 把 $I_1$ 等分成 $2^n$ 个闭矩形，那么其中至少有一个不能用 $(G_\lambda)_{\lambda\in L}$ 中有限个开集覆盖，我们把这个闭矩形记作 $I_2$。类似将 $I_2$ 等分成 $2^n$ 个闭矩形，并将其中不能用 $(G_\lambda)_{\lambda\in L}$ 中有限个开集覆盖的某个闭矩形记作 $I_3$。如此下去，我们得到了一个闭矩形列 $\{I_m\}$，其满足：

(1) 对任意的 $m$ 均有 $I_{m+1} \subseteq I_m$；

(2) $\operatorname{diam}(I_m) = 2^{-m+1}\operatorname{diam}(I_1) \to 0$（当 $m \to \infty$ 时）；

(3) 每个 $I_m$ 均不能被有限多个 $G_\lambda$ 覆盖。

---

[④] 我们常用 $K$ 表示紧集，它是德文“kompakt”的首字母。

[⑤] 所谓 $\mathbb{R}^n$ 中的超平面，是指维数为 $n-1$ 的子空间。

由闭矩形套定理知存在 $\boldsymbol{a} \in \mathbb{R}^n$ 使得

$$\bigcap_{m=1}^{\infty} I_m = \{\boldsymbol{a}\}.$$

显然有 $\boldsymbol{a} \in \bigcup_{\lambda \in L} G_\lambda$，从而存在 $\lambda_0 \in L$ 使得 $\boldsymbol{a} \in G_{\lambda_0}$。因为 $G_{\lambda_0}$ 是开集，所以存在 $\varepsilon > 0$ 使得 $B(\boldsymbol{a}, \varepsilon) \subseteq G_{\lambda_0}$。注意到 $\lim_{m\to\infty} \operatorname{diam}(I_m) = 0$，故存在正整数 $\ell$ 使得 $\operatorname{diam}(I_\ell) < \varepsilon$。由于 $\boldsymbol{a} \in I_\ell$，所以对任意的 $\boldsymbol{x} \in I_\ell$ 均有

$$|\boldsymbol{x} - \boldsymbol{a}| \leqslant \operatorname{diam}(I_\ell) < \varepsilon,$$

这说明 $I_\ell \subseteq B(a, \varepsilon)$，进而 $I_\ell \subseteq G_{\lambda_0}$，但这与 (3) 矛盾。 □

在给出 $\mathbb{R}^n$ 中紧集的一个确切描述之前，我们先介绍有界集的概念。

**【定义 1.23】设 $E \subseteq \mathbb{R}^n$。若存在 $M > 0$，使得对任意的 $\boldsymbol{x} \in E$ 均有 $|\boldsymbol{x}| \leqslant M$，则称 $E$ 是有界的 (bounded)。**

**【定理 1.24】设 $K \subseteq \mathbb{R}^n$，则 $K$ 是紧集当且仅当它是有界闭集。**

证明. 必要性：我们假设 $K$ 是紧集。首先，由于 $(B(\mathbf{0}, m))_{m \in \mathbb{Z}_{>0}}$ 是 $\mathbb{R}^n$ 的一个开覆盖，自然也是 $K$ 的一个开覆盖，从而存在 $m_1, \cdots, m_\ell$ 使得

$$K \subseteq \bigcup_{j=1}^{\ell} B(\mathbf{0}, m_j).$$

记 $M = \max(m_1, \cdots, m_\ell)$，则有 $K \subseteq B(\mathbf{0}, M)$。这就证明了 $K$ 是有界集。

其次，对任意的 $\boldsymbol{a} \in K^c$，我们来考虑以 $\boldsymbol{a}$ 为中心，以 $\dfrac{1}{m}$ 为半径的闭球 $\overline{B\Big(\boldsymbol{a}, \dfrac{1}{m}\Big)}$。容易证明

$$\bigcap_{m=1}^{\infty} \overline{B\Big(\boldsymbol{a}, \frac{1}{m}\Big)} = \{\boldsymbol{a}\},$$

于是由 de Morgan 律知

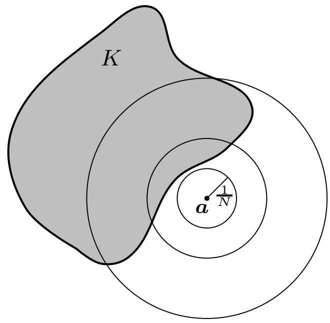
$$\bigcup_{m=1}^{\infty} \overline{B\Big(\boldsymbol{a}, \frac{1}{m}\Big)}^{c} = \mathbb{R}^n \setminus \{\boldsymbol{a}\}.$$

注意到 $\boldsymbol{a} \notin K$，故由上式可得 $\left(\overline{B\Big(\boldsymbol{a}, \dfrac{1}{m}\Big)}^{c}\right)_{m \in \mathbb{Z}_{>0}}$ 是 $K$ 的一个开覆盖，再由 $K$ 是紧集知存在 $m_1, \cdots, m_k$ 使得

$$K \subseteq \bigcup_{j=1}^{k} \overline{B\Big(\boldsymbol{a}, \frac{1}{m_j}\Big)}^{c},$$

也即

$$\bigcap_{j=1}^{k} \overline{B\Big(\boldsymbol{a}, \frac{1}{m_j}\Big)} \subseteq K^c.$$




现记 $N = \max(m_1, \cdots, m_k)$，由上式可得 $B\Big(\boldsymbol{a}, \dfrac{1}{N}\Big) \subseteq K^c$，这说明 $\boldsymbol{a}$ 是 $K^c$ 的内点，再由 $\boldsymbol{a}$ 的任意性知 $K^c$ 为开集，进而 $K$ 是闭集。

充分性：现设 $K$ 是有界闭集。由有界性知存在 $M > 0$ 使得对任意的 $\boldsymbol{x} \in E$ 均有 $|\boldsymbol{x}| \leqslant M$，从而 $K \subseteq [-M, M]^n$。现考虑 $K$ 的任意一个开覆盖 $(G_\lambda)_{\lambda \in L}$，由 $K^c$ 是开集知

$$\{G_\lambda : \lambda \in L\} \cup \{K^c\}$$

形成 $[-M, M]^n$ 的一个开覆盖。注意到在命题 1.22 中我们已经证明了 $[-M, M]^n$ 是紧集，因而在上述集族中能选出有限个开集覆盖 $[-M, M]^n$。若这有限个开集中有 $K^c$，那么把它剔除掉后可得 $(G_\lambda)_{\lambda \in L}$ 关于 $K$ 的一个有限子覆盖。这就证明了 $K$ 是紧集。 □

最后，我们来介绍 Bolzano−Weierstrass 定理。

**【定理 1.25】（Bolzano−Weierstrass 定理）** $\mathbb{R}^n$ 的任意一个有界无限子集必有聚点。

<u>证明</u>. 设 $E$ 是 $\mathbb{R}^n$ 的一个有界的无限子集。由有界性知存在 $M > 0$ 使得 $E \subseteq [-M, M]^n$。若 $E$ 无聚点，那么对任意的 $\boldsymbol{x} \in [-M, M]^n$，存在 $\boldsymbol{x}$ 的邻域 $U_{\boldsymbol{x}}$，它至多含有 $E$ 中一个点（若 $\boldsymbol{x} \in E$，那么这个点就是 $\boldsymbol{x}$；否则 $U_{\boldsymbol{x}} \cap E = \varnothing$）。注意到

$$(U_{\boldsymbol{x}})_{\boldsymbol{x} \in [-M, M]^n}$$

是 $[-M, M]^n$ 的一个开覆盖，但由 $E$ 是无限集知上面集族中任意有限个开集均不能覆盖 $E$，当然也不能覆盖 $[-M, M]^n$，这与 $[-M, M]^n$ 是紧集矛盾。 □

### 12.1.4 连通集

**【定义 1.26】** 设 $A \subseteq E \subseteq \mathbb{R}^n$。如果存在 $\mathbb{R}^n$ 中的开集（相应地，闭集）$S$ 使得 $A = E \cap S$，则称 $A$ 是 $E$ 的一个<u>开子集</u>（相应地，<u>闭子集</u>）。

**【例 1.27】** (1) $[1,2)$ 是 $[1,5]$ 的开子集。

(2) $E$ 和 $\varnothing$ 既是 $E$ 的开子集又是 $E$ 的闭子集。

(3) 如果 $\boldsymbol{a}$ 是 $E$ 的孤立点，那么 $\{\boldsymbol{a}\}$ 既是 $E$ 的开子集又是 $E$ 的闭子集。

**【命题 1.28】** 设 $E \subseteq \mathbb{R}^n$，$A, B \subseteq E$，那么

(1) $A$ 是 $E$ 的开子集的充要条件是：对任意的 $\boldsymbol{a} \in A$，存在 $\boldsymbol{a}$ 的邻域 $U$ 使得 $E \cap U \subseteq A$;

(2) $B$ 是 $E$ 的闭子集当且仅当 $E \setminus B$ 是 $E$ 的开子集。

证明. (1) 必要性：如果 $A$ 是 $E$ 的开子集，则存在 $\mathbb{R}^n$ 中的开集 $G$ 使得 $A = E \cap G$。于是对任意的 $\boldsymbol{a} \in A$，由 $\boldsymbol{a} \in G$ 知 $G$ 是 $\boldsymbol{a}$ 的邻域，从而必要性得证。

充分性：若对任意的 $\boldsymbol{a} \in A$，均存在 $\boldsymbol{a}$ 的邻域 $U_{\boldsymbol{a}}$ 使得 $E \cap U_{\boldsymbol{a}} \subseteq A$，那么 $\bigcup\limits_{\boldsymbol{a} \in A} U_{\boldsymbol{a}}$ 是 $\mathbb{R}^n$ 中的开集。并且一方面显然有 $A \subseteq E \cap \Big(\bigcup\limits_{\boldsymbol{a} \in A} U_{\boldsymbol{a}}\Big)$，另一方面

$$E \cap \left(\bigcup_{\boldsymbol{a} \in A} U_{\boldsymbol{a}}\right) = \bigcup_{\boldsymbol{a} \in A} (E \cap U_{\boldsymbol{a}}) \subseteq A,$$

故而

$$A = E \cap \left(\bigcup_{\boldsymbol{a} \in A} U_{\boldsymbol{a}}\right),$$

因此 $A$ 是 $E$ 的开子集。

(2) 若 $B$ 是 $E$ 的闭子集，则存在 $\mathbb{R}^n$ 中的闭集 $F$ 使得 $B = E \cap F$，于是

$$\begin{aligned} E \setminus B &= E \setminus (E \cap F) = E \cap (E \cap F)^c = E \cap (E^c \cup F^c) \\ &= E \cap F^c. \end{aligned}$$

注意到 $F^c$ 是 $\mathbb{R}^n$ 中的开集，所以 $E \setminus B$ 是 $E$ 的开子集。

反之，若 $E \setminus B$ 是 $E$ 的开子集，则存在 $\mathbb{R}^n$ 中的开集 $G$ 使得 $E \setminus B = E \cap G$。与上面类似可得

$$B = E \setminus (E \cap G) = E \cap G^c.$$

因为 $G^c$ 是 $\mathbb{R}^n$ 中的闭集，所以 $B$ 是 $E$ 的闭子集。 □

**【定义 1.29】** 设 $E \subseteq \mathbb{R}^n$。如果不存在 $E$ 的两个非空开子集 $A$ 和 $B$ 使得 $A \cup B = E$ 且 $A \cap B = \varnothing$，则称 $E$ 是 $\mathbb{R}^n$ 中的连通集 (connected set)。$\mathbb{R}^n$ 中的连通开集被称作区域。

如果 $E$ 是区域，那么也将 $\overline{E}$ 称为闭区域。特别需要注意的是，**闭区域不是区域**。

**【例 1.30】**(1) $E=(1,2)\cup(3,4]$ 不是 $\mathbb{R}$ 中的连通集。

(2) $\mathbb{Q}$ 不是 $\mathbb{R}$ 中的连通集，这是因为它可用如下方式写成两个不相交的非空开子集的并：

$$\mathbb{Q}=\{x\in\mathbb{Q}:x<\sqrt{2}\}\cup\{x\in\mathbb{Q}:x>\sqrt{2}\}.$$

(3) 由例 1.27 (3) 我们知道：如果 $\boldsymbol{a}$ 是 $E$ 的孤立点，那么 $E$ 是连通集当且仅当 $E=\{\boldsymbol{a}\}$。

**【命题 1.31】设 $E$ 是 $\mathbb{R}$ 的非空子集，那么 $E$ 是 $\mathbb{R}$ 中的连通集当且仅当 $E$ 是区间。**

证明. 必要性：设 $E$ 是连通集。若 $E$ 是单元素集，那么它是一个（退化的）区间，下设 $E$ 中至少有两个元素。记 $a=\inf E$，$b=\sup E$（$a$ 可以是 $-\infty$，$b$ 可以是 $+\infty$）。为了证明 $E$ 是区间，只需说明 $(a,b)\subseteq E$ 即可。现任取 $x\in(a,b)$，若 $x\notin E$，那么我们可以通过

$$E=[E\cap(-\infty,x)]\cup[E\cap(x,+\infty)]$$

将 $E$ 写成它的两个不相交的非空开子集的并，这与 $E$ 是连通集矛盾。

充分性：假设 $E$ 是区间。如果 $E$ 不是连通集，则必存在 $E$ 的两个非空开子集 $A$ 和 $B$ 使得 $A\cup B=E$ 且 $A\cap B=\varnothing$。现取 $x\in A$，$y\in B$，并不妨假设 $x<y$，再记

$$z=\sup(A\cap[x,y]),$$

那么当然有 $x\leqslant z\leqslant y$。注意到 $E$ 是区间，故由 $x,y\in E$ 知 $[x,y]\subseteq E$，从而 $z\in E=A\cup B$。下面分情况来讨论：

(1) 若 $z\in A$，则由 $A\cap B=\varnothing$ 可得 $z<y$。由 $A$ 是 $E$ 的开子集知存在 $\mathbb{R}$ 的开集 $G_1$ 使得 $A=E\cap G_1$。因为 $z\in G_1$ 且 $G_1$ 是 $\mathbb{R}$ 中的开集，故存在 $s\in(z,y)$ 使得 $(z,s)\subseteq G_1$。此外 $(z,s)\subseteq[x,y]\subseteq E$，进而有 $(z,s)\subseteq E\cap G_1=A$，这与 $z$ 的定义矛盾。

(2) 若 $z\in B$，则 $z>x$。由 $B$ 是 $E$ 的开子集知存在 $\mathbb{R}$ 的开集 $G_2$ 使得 $B=E\cap G_2$。于是由 $z\in G_2$ 且 $G_2$ 是 $\mathbb{R}$ 中的开集知，存在 $t\in(x,z)$ 使得 $(t,z)\subseteq G_2$。此外 $(t,z)\subseteq[x,y]\subseteq E$，所以 $(t,z)\subseteq E\cap G_2=B$。注意到 $A\cap B=\varnothing$，因此 $(t,z)\cap A=\varnothing$，这也与 $z$ 的定义矛盾。 □

**【命题 1.32】设 $E$ 是 $\mathbb{R}^n$ 中的连通集，且 $E\subseteq S\subseteq\overline{E}$，那么 $S$ 也是 $\mathbb{R}^n$ 中的连通集。特别地，$\overline{E}$ 是 $\mathbb{R}^n$ 中的连通集。**

证明. 反设 $S$ 不是连通集，则存在 $S$ 的非空开子集 $A$ 与 $B$ 满足 $A\cup B=S$ 及 $A\cap B=\varnothing$。$A\cap E$ 与 $B\cap E$ 当然是 $E$ 的开子集，下证 $A\cap E$ 与 $B\cap E$ 都不是空集。

事实上，如果 $A\cap E=\varnothing$，那么由 $E\subseteq S\subseteq\overline{E}=E\cup E'$ 知 $A\subseteq E'$。现取 $\boldsymbol{a}\in A$，则 $\boldsymbol{a}\in E'$。一方面，由命题 1.28 (1) 知存在 $\boldsymbol{a}$ 的邻域 $U$ 使得 $S\cap U\subseteq A$，当然 $E\cap U\subseteq A$；另一方面，由 $\boldsymbol{a}\in E'$ 知 $E\cap U\neq\varnothing$，这与 $A\cap E=\varnothing$ 矛盾。同理可证 $B\cap E\neq\varnothing$。

注意到

$$(A\cap E)\cup(B\cap E)=E\qquad\text{以及}\qquad(A\cap E)\cap(B\cap E)=\varnothing,$$

这说明 $E$ 不是连通集，从而与假设矛盾。 □

## 习题 12.1

1. 设 $\boldsymbol{x}_m=(x_1^{(m)},\cdots,x_n^{(m)})$，证明 $\{\boldsymbol{x}_m\}$ 收敛于 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 当且仅当对任意的 $1\leqslant j\leqslant n$ 有 $\lim\limits_{m\to\infty}x_j^{(m)}=a_j$。

2. 对下列集合 $E$，求 $E^\circ$，$\partial E$ 以及 $\overline{E}$:

   (1) $E=\{\sqrt[n]{n}:n\in\mathbb{Z}_{\geqslant1}\}$;　　(2) $E=\{(x,y)\in\mathbb{R}^2:0<y<x+1\}$;

   (3) $E=\{(x,y)\in\mathbb{R}^2:xy\neq0\}$;　　(4) $E=\{(x,y)\in\mathbb{R}^2:|x|+|y|<1\}$;

   (5) $E=\{(x,y)\in\mathbb{R}^2:x,\,y\in\mathbb{Q}\}$;　　(6) $E=\{(x,y)\in\mathbb{R}^2:1<x^2+y^2<4\}$。

3. 设 $E\subseteq\mathbb{R}^n$，证明

$$E^\circ=\bigcup_{\substack{G\subseteq E\\ G\text{ 是开集}}}G,\qquad\qquad \overline{E}=\bigcap_{\substack{F\supseteq E\\ F\text{ 是闭集}}}F.$$

   因此 $E^\circ$ 是包含于 $E$ 的最大开集；$\overline{E}$ 是包含 $E$ 的最小闭集。

4. 设 $A,\,B\subseteq\mathbb{R}^n$，证明：

   (1) 若 $A\subseteq B$，则 $A^\circ\subseteq B^\circ$ 且 $\overline{A}\subseteq\overline{B}$;

   (2) $(A\cap B)^\circ=A^\circ\cap B^\circ$，$\overline{A\cup B}=\overline{A}\cup\overline{B}$;

   (3) $(A\cup B)^\circ\supseteq A^\circ\cup B^\circ$，$\overline{A\cap B}\subseteq\overline{A}\cap\overline{B}$;

   (4) 举例说明 (3) 中的两个式子均未必有等号成立。

5. 设 $G_1$，$G_2$ 是 $\mathbb{R}^n$ 中的两个不相交的开集，证明 $G_1\cap\overline{G_2}=\varnothing$。

6. 设 $\boldsymbol{a}=(a_1,\cdots,a_n)\in\mathbb{R}^n$，$\boldsymbol{b}=(b_1,\cdots,b_m)\in\mathbb{R}^m$，记

$$(\boldsymbol{a},\boldsymbol{b})=(a_1,\cdots,a_n,b_1,\cdots,b_m)\in\mathbb{R}^{m+n}.$$

若 $G$ 是包含 $(\boldsymbol{a},\boldsymbol{b})$ 的开集，证明必存在 $\mathbb{R}^n$ 中包含 $\boldsymbol{a}$ 的开集 $G_1$ 和 $\mathbb{R}^m$ 中包含 $\boldsymbol{b}$ 的开集 $G_2$ 使得 $G_1\times G_2\subseteq G$。

7. 设 $E\subseteq\mathbb{R}^n$ 且 $E$ 中每个点均是孤立点。证明存在开集 $G$ 和闭集 $F$ 使得 $E=G\cap F$。

8. 设 $E,F\subseteq\mathbb{R}^n$，证明 $E\cup F$，$E\cap F$ 以及 $E\setminus F$ 的边界均是 $\partial E\cup\partial F$ 的子集。

9. 设 $E\subseteq\mathbb{R}^n$，证明 $\partial E=\varnothing$ 的充要条件是 $E=\varnothing$ 或 $E=\mathbb{R}^n$。

10. 设 $E$ 是 $\mathbb{R}^n$ 的一个既开且闭的子集，证明 $E=\varnothing$ 或 $E=\mathbb{R}^n$。

11. 设 $E\subseteq\mathbb{R}^n$，证明 $\partial E$ 是 $\mathbb{R}^n$ 中的闭集。

12. 设 $E\subseteq\mathbb{R}^n$，证明 $\overline{E}=E\cup\partial E$。

13. 设 $E=B(\boldsymbol{a},r)$，证明 $\operatorname{diam}(E)=2r$。

14. 证明定理 1.19。

15. 设 $K$ 是 $\mathbb{R}^n$ 中的紧集，$G$ 是 $\mathbb{R}^n$ 中的开集且 $K\subseteq G$。证明存在开集 $G_1$ 满足

$$K\subseteq G_1\subseteq\overline{G_1}\subseteq G.$$

16. 举例说明：如果把上题中的条件“$K$ 是紧集”换成“$K$ 是闭集”，则结论未必成立。

17. 对 $\mathbb{R}^n$ 的任意两个非空子集 $A$ 和 $B$，记

$$d(A,B)=\inf_{\boldsymbol{x}\in A,\ \boldsymbol{y}\in B}|\boldsymbol{x}-\boldsymbol{y}|.$$

(1) 设 $A$ 是紧集，$B$ 是闭集且 $A\cap B=\varnothing$，证明 $d(A,B)>0$。

(2) 若 $A$ 与 $B$ 均是闭集且 $A\cap B=\varnothing$，是否一定有 $d(A,B)>0$?

18. 沿用上题记号，并当 $A=\{\boldsymbol{x}\}$ 时将 $d(A,B)$ 简记为 $d(\boldsymbol{x},B)$。证明：对任意的 $E\subseteq\mathbb{R}^n$ 有

$$\overline{E}=\{\boldsymbol{x}\in\mathbb{R}^n:d(\boldsymbol{x},E)=0\}.$$

19. 设 $F$ 是 $\mathbb{R}^n$ 中的闭集，$r$ 是一个正常数，证明 $\{\boldsymbol{x}\in\mathbb{R}^n:d(\boldsymbol{x},F)<r\}$ 是开集。

20. 证明 $\mathbb{R}^n$ 中每个闭集均可表为可数个开集的交，每个开集均可表为可数个闭集的并。

21. 假设 $(F_\lambda)_{\lambda\in L}$ 是 $\mathbb{R}^n$ 中的一族闭集，其中至少有一个 $F_\lambda$ 是有界集，并且 $\bigcap\limits_{\lambda\in L} F_\lambda=\varnothing$。证明：存在 $\lambda_1,\lambda_2,\cdots,\lambda_m\in L$ 使得 $\bigcap\limits_{1\leqslant j\leqslant m} F_{\lambda_j}=\varnothing$。

22. （闭集套定理）设 $(F_j)_{j\in\mathbb{Z}_{>0}}$ 是 $\mathbb{R}^n$ 中的一族有界闭集，满足

 (1) 对任意的正整数 $j$ 有 $F_j\supseteq F_{j+1}$；

 (2) $\lim\limits_{j\to\infty}\operatorname{diam}(F_j)=0$。

 证明存在唯一的 $\boldsymbol{a}\in\mathbb{R}^n$ 使得 $\bigcap\limits_{j=1}^{\infty} F_j=\{\boldsymbol{a}\}$。

23. 设 $E$ 是 $\mathbb{R}^n$ 的一个不可数的子集。证明存在 $E$ 的聚点 $\boldsymbol{a}$，使得 $\boldsymbol{a}$ 的任一邻域与 $E$ 的交集均是不可数集。

24. 设 $E\subseteq\mathbb{R}^n$，证明以下命题等价：

 (1) $E$ 是连通集；

 (2) 不存在 $E$ 的非空真子集 $A$，使得它既是 $E$ 的开子集又是 $E$ 的闭子集；

 (3) 不存在 $E$ 的两个非空闭子集 $A$ 和 $B$ 使得 $A\cup B=E$ 且 $A\cap B=\varnothing$；

 (4) 不存在 $E$ 的两个子集 $A$ 和 $B$ 使得 $A\cup B=E$ 且 $A\cap\overline{B}=\overline{A}\cap B=\varnothing$。

25. 设 $E,F$ 均是 $\mathbb{R}^n$ 中的连通集且 $\overline{E}\cap F\neq\varnothing$，证明 $E\cup F$ 也是 $\mathbb{R}^n$ 中的连通集。

26. 设 $(E_\lambda)_{\lambda\in L}$ 是 $\mathbb{R}^n$ 中的一族连通集，并且 $\bigcap\limits_{\lambda\in L} E_\lambda\neq\varnothing$，证明 $\bigcup\limits_{\lambda\in L} E_\lambda$ 也是连通集。

27. 证明 $\mathbb{R}$ 中的每个开集均可表为至多可数个互不相交的开区间的并。

设 $E$ 是一个非空集合，如果存在映射 $d: E^2\longrightarrow\mathbb{R}$ 满足

(1) 对任意的 $x,y\in E$ 均有 $d(x,y)\geqslant 0$，并且 $d(x,y)=0$ 当且仅当 $x=y$；

(2) 对任意的 $x,y\in E$ 均有 $d(x,y)=d(y,x)$；

(3) （三角形不等式）对任意的 $x,y,z\in E$ 有 $d(x,z)\leqslant d(x,y)+d(y,z)$，

则称 $E$ 是一个<u>距离空间 (metric space)</u>，称 $d$ 为 $E$ 上的<u>距离</u>。从第 28 题到第 34 题是一组题，介绍距离空间的一些例子。

28. 设 $E$ 是一个非空集合，对任意的 $x,y\in E$，定义

$$d(x,y)=\begin{cases}0, & \text{若 } x=y,\\ 1, & \text{若 } x\neq y.\end{cases}$$

证明 $d$ 是 $E$ 上的距离。拥有这一距离的集合 $E$ 被称为离散距离空间 (discrete metric space)。

29. 设 $E$ 是以 $d$ 为距离的距离空间，对任意的 $x, y \in E$，定义
$$d_1(x,y)=\frac{d(x,y)}{1+d(x,y)}.$$
证明：$d_1$ 也是 $E$ 上的一个距离。

30. 对 $\mathbb{R}^2$ 中的任意两个元素 $\boldsymbol{x}=(x_1,x_2)$ 和 $\boldsymbol{y}=(y_1,y_2)$，定义
$$d(\boldsymbol{x},\boldsymbol{y})=\max\big\{|x_1-y_1|,\,|x_2-y_2|\big\},$$
证明：$d$ 是 $\mathbb{R}^2$ 上的距离。

31. 设 $E$ 是定义在有界闭区间 $[a,b]$ 上的全体连续函数所成之集，对任意的 $f, g \in E$，定义
$$d(f,g)=\sup_{x\in[a,b]}|f(x)-g(x)|.$$
证明 $d$ 是 $E$ 上的距离。

32. 设 $E$ 是定义在有界闭区间 $[a,b]$ 上的全体连续函数所成之集，对任意的 $f, g \in E$，定义
$$d(f,g)=\int_a^b|f(x)-g(x)|\,\mathrm{d}x.$$
证明 $d$ 是 $E$ 上的距离。

33. 设 $p$ 是一个素数，对任一非零整数 $a$，定义 $v_p(a)$ 为使得 $p^k$ 整除 $a$ 的最大的 $k$，即 $v_p(a)=\max\{k\in\mathbb{Z}_{\geqslant 0}: p^k \mid a\}$。进一步地，对非零有理数 $x=\dfrac{a}{b}$（$a$ 和 $b$ 均为整数），定义 $v_p(x)=v_p(a)-v_p(b)$[⑥]。现对 $x\in\mathbb{Q}$ 记
$$|x|_p=\begin{cases}p^{-v_p(x)}, & 若\ x\neq 0,\\ 0, & 若\ x=0.\end{cases}$$
证明：$d(x,y)=|x-y|_p$ 为 $\mathbb{Q}$ 上的一个距离。

34. 集合 $E$ 上的距离 $d$ 被称为是非 Archimedes 的 (non-Archimedean)，是指对任意的 $x, y, z \in E$，均有
$$d(x,y)\leqslant\max\big\{d(x,z),\,d(y,z)\big\}.$$

⑥这一定义的合理性参见 §1.2 习题 1。

否则，就称该距离是 Archimedes 的。试讨论习题 30 ∼ 33 中的距离是否是非 Archimedes 的。

§12.2

## 多元函数的极限

**【定义 2.1】**设 $E \subseteq \mathbb{R}^n$，$f: E \longrightarrow \mathbb{R}^m$，$\boldsymbol{a}$ 是 $E$ 的聚点。若存在 $\boldsymbol{b} \in \mathbb{R}^m$，使得对任意的 $\varepsilon > 0$，均存在 $\delta > 0$ 满足

$$|f(\boldsymbol{x}) - \boldsymbol{b}| < \varepsilon, \qquad \forall\, \boldsymbol{x} \in \Big(B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}\Big) \cap E,$$

那么就称 $\boldsymbol{b}$ 为 $f$ 沿 $E$ 中元素趋于 $\boldsymbol{a}$ 时的极限。

**【命题 2.2】（极限的唯一性）** 设 $E \subseteq \mathbb{R}^n$，$f: E \longrightarrow \mathbb{R}^m$，$\boldsymbol{a}$ 是 $E$ 的聚点。如果 $\boldsymbol{b}$ 与 $\boldsymbol{c}$ 均是 $f$ 沿 $E$ 中元素趋于 $\boldsymbol{a}$ 时的极限，则 $\boldsymbol{b} = \boldsymbol{c}$。

证明. 若 $\boldsymbol{b} \neq \boldsymbol{c}$，则 $|\boldsymbol{b} - \boldsymbol{c}| \neq 0$。按照定义 2.1，对 $\varepsilon = \dfrac{|\boldsymbol{b} - \boldsymbol{c}|}{2}$，存在正数 $\delta_1$ 与 $\delta_2$ 使得

$$|f(\boldsymbol{x}) - \boldsymbol{b}| < \varepsilon, \qquad \forall\, \boldsymbol{x} \in \Big(B(\boldsymbol{a}, \delta_1) \setminus \{\boldsymbol{a}\}\Big) \cap E$$

以及

$$|f(\boldsymbol{x}) - \boldsymbol{c}| < \varepsilon, \qquad \forall\, \boldsymbol{x} \in \Big(B(\boldsymbol{a}, \delta_2) \setminus \{\boldsymbol{a}\}\Big) \cap E.$$

于是若记 $\delta = \min(\delta_1, \delta_2)$，则对 $\boldsymbol{x} \in \Big(B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}\Big) \cap E$ 有

$$|\boldsymbol{b} - \boldsymbol{c}| \leqslant |f(\boldsymbol{x}) - \boldsymbol{b}| + |f(\boldsymbol{x}) - \boldsymbol{c}| < 2\varepsilon = |\boldsymbol{b} - \boldsymbol{c}|.$$

从而矛盾。 □

鉴于命题 2.2，当 $\boldsymbol{b}$ 是 $f$ 沿 $E$ 中元素趋于 $\boldsymbol{a}$ 时的极限时，我们就记

$$\lim_{\substack{\boldsymbol{x} \to \boldsymbol{a} \\ \boldsymbol{x} \in E}} f(\boldsymbol{x}) = \boldsymbol{b},$$

如果 $\boldsymbol{a}$ 是 $E$ 的内点，则将其简记作 $\lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$。

若记 $\boldsymbol{x} = (x_1, \cdots, x_n)$，$\boldsymbol{a} = (a_1, \cdots, a_n)$，那么 $\boldsymbol{x} \to \boldsymbol{a}$ 当且仅当对 $1 \leqslant j \leqslant n$ 都有 $x_j \to a_j$，因此我们也常将极限式 $\lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$ 写成

$$\lim_{x_j \to a_j,\ 1 \leqslant j \leqslant n} f(x_1, \cdots, x_n) = \boldsymbol{b}.$$

类似可定义诸如 $\lim\limits_{x_j\to+\infty,\ 1\leqslant j\leqslant n} f(x_1,\cdots,x_n)=\boldsymbol{b}$ 的极限式。

现设 $E\subseteq\mathbb{R}^n$ 且 $f:E\longrightarrow\mathbb{R}^m$，则对任意的 $\boldsymbol{x}\in E$，$f(\boldsymbol{x})$ 是一个 $m$ 维向量，从而可写成 $(f_1(\boldsymbol{x}),\cdots,f_m(\boldsymbol{x}))^{\mathrm{T}}$ 的形式[7]，其中 $f_j:E\longrightarrow\mathbb{R}$ $(1\leqslant j\leqslant m)$ 均是定义在 $E$ 上的多元函数。此时我们通常记 $f=(f_1,\cdots,f_m)^{\mathrm{T}}$，并称 $f_j$ 是 $f$ 的第 $j$ 个<u>分量函数</u>。下述结论是比较显然的。

**【命题 2.3】** 设 $E\subseteq\mathbb{R}^n$，$f=(f_1,\cdots,f_m)^{\mathrm{T}}:E\longrightarrow\mathbb{R}^m$ 且 $\boldsymbol{a}$ 是 $E$ 的聚点。又设 $\boldsymbol{b}=(b_1,\cdots,b_m)^{\mathrm{T}}\in\mathbb{R}^m$。那么 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=\boldsymbol{b}$ 当且仅当对 $1\leqslant j\leqslant m$ 有 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f_j(\boldsymbol{a})=b_j$。

下面给出映射极限的一些性质，它们与函数极限类似，因此我们略去它们的证明。

**【命题 2.4】** 设 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})$ 与 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{x})$ 均存在，则对任意的实数 $\alpha,\beta$ 有

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}}\big[\alpha f(\boldsymbol{x})+\beta g(\boldsymbol{x})\big]=\alpha\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})+\beta\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{x}).$$

**【命题 2.5】（复合映射的极限）** 设 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=\boldsymbol{b}$，$\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{y})=\boldsymbol{c}$，且存在 $\delta>0$，使得在 $\Big(B(\boldsymbol{a},\delta)\setminus\{\boldsymbol{a}\}\Big)\cap E$ 内有 $f(\boldsymbol{x})\neq\boldsymbol{b}$，则

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(f(\boldsymbol{x}))=\boldsymbol{c}.$$

**【定理 2.6】（Heine 归结原理）** $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=\boldsymbol{b}$ 的充要条件是：对于 $E$ 中满足 $\lim\limits_{k\to\infty}\boldsymbol{x}_k=\boldsymbol{a}$ 且 $\boldsymbol{x}_k\neq\boldsymbol{a}$ $(\forall\ k)$ 的任一序列 $\{\boldsymbol{x}_k\}$ 均有 $\lim\limits_{k\to\infty} f(\boldsymbol{x}_k)=\boldsymbol{b}$。

**【例 2.7】** 设 $f(x,y)=\dfrac{xy}{x^2+y^2}$，我们来考虑极限 $\lim\limits_{\substack{x\to0\\ y\to0}} f(x,y)$。由于

$$\lim_{k\to\infty} f\Big(\frac{1}{k},\frac{1}{k}\Big)=\frac{1}{2}\qquad \text{且}\qquad \lim_{k\to\infty} f\Big(\frac{2}{k},\frac{1}{k}\Big)=\frac{2}{5},$$

故极限 $\lim\limits_{\substack{x\to0\\ y\to0}} f(x,y)$ 不存在。

---

[7] 在这里我们采用列向量的写法，目的是和下一章一致。

**【定理 2.8】（Cauchy 收敛准则）** $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})$ 存在的充要条件是：对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于任意的 $\boldsymbol{x},\boldsymbol{y}\in\Big(B(\boldsymbol{a},\delta)\setminus\{\boldsymbol{a}\}\Big)\cap E$ 有

$$|f(\boldsymbol{x})-f(\boldsymbol{y})|<\varepsilon.$$

鉴于命题 2.3，下面我们将仅限于讨论多元函数的情形。因为 $\mathbb{R}$ 是 Archimedes 有序域，故有许多良好的性质可以利用。

**【命题 2.9】（保序性）** 设 $f$ 与 $g$ 均是 $n$ 元函数，且

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=A,\qquad \lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{x})=B$$

那么

(1) 若 $A<B$，则存在 $\boldsymbol{a}$ 的某去心邻域，使得在该去心邻域内有 $f(\boldsymbol{x})<g(\boldsymbol{x})$;

(2) 若在 $\boldsymbol{a}$ 的某去心邻域内有 $f(\boldsymbol{x})\leqslant g(\boldsymbol{x})$，则 $A\leqslant B$。

**【命题 2.10】** 设 $f$ 与 $g$ 均是 $n$ 元函数，且 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=A$，$\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{x})=B$，则

(1) $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})g(\boldsymbol{x})=AB$;

(2) 若 $B\neq 0$，则 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} \dfrac{f(\boldsymbol{x})}{g(\boldsymbol{x})}=\dfrac{A}{B}$。

特别常用的是下面的夹逼定理。

**【定理 2.11】（夹逼定理）** 设 $E\subseteq\mathbb{R}^n$，$\boldsymbol{a}$ 是 $E$ 的聚点，$f,g,h$ 均是从 $E$ 到 $\mathbb{R}^m$ 的映射，并且存在 $\delta>0$，使得在 $\Big(B(\boldsymbol{a},\delta)\setminus\{\boldsymbol{a}\}\Big)\cap E$ 内有 $f(\boldsymbol{x})\leqslant g(\boldsymbol{x})\leqslant h(\boldsymbol{x})$。如果

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} f(\boldsymbol{x})=\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} h(\boldsymbol{x})=A,$$

那么 $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in E}} g(\boldsymbol{x})=A$。

**【例 2.12】** 计算 $\lim\limits_{\substack{x\to 0^+\\ y\to 0^+}} \dfrac{\sin(xy)}{x+y}$。

<u>解</u>. 由于当 $x>0$，$y>0$ 时有

$$\left|\frac{\sin(xy)}{x+y}\right|\leqslant\left|\frac{xy}{x+y}\right|\leqslant\frac{x+y}{4},$$

而 $\lim\limits_{\substack{x\to0^+\\y\to0^+}}\dfrac{x+y}{4}=0$，因此 $\lim\limits_{\substack{x\to0^+\\y\to0^+}}\dfrac{\sin(xy)}{x+y}=0$。 □

最后我们再把问题限于二元函数的情形，并研究 $\lim\limits_{\substack{x\to a\\y\to b}}f(x,y)$ 与 $\lim\limits_{y\to b}\lim\limits_{x\to a}f(x,y)$ 以及 $\lim\limits_{x\to a}\lim\limits_{y\to b}f(x,y)$ 之间的关系。其中第一个极限通常被称作<u>二重极限</u>，而后两个极限被称作<u>二次极限</u>或<u>累次极限</u>。在这里，$\lim\limits_{y\to b}\lim\limits_{x\to a}f(x,y)$ 是指对于 $b$ 的某去心邻域内的每个 $y$ 均存在极限 $\lim\limits_{x\to a}f(x,y)=\varphi(y)$，并且 $\lim\limits_{y\to b}\varphi(y)$ 也存在，那么我们就记

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

在一般情形下，二重极限与二次极限并无直接关联。下面我们来作具体说明：

(1) 两个二次极限均不存在，但二重极限仍有可能存在。例如

$$f(x,y)=\begin{cases}x\sin\dfrac{1}{y}+y\sin\dfrac{1}{x}, & \text{若 } x\neq0 \text{ 且 } y\neq0,\\ 0, & \text{若 } x=0 \text{ 或 } y=0,\end{cases}$$

由于当 $y\neq0$ 时 $\lim\limits_{x\to0}y\sin\dfrac{1}{x}$ 不存在，且当 $x\neq0$ 时 $\lim\limits_{y\to0}x\sin\dfrac{1}{y}$ 不存在，故

$$\lim_{x\to0}f(x,y)\qquad\text{与}\qquad\lim_{y\to0}f(x,y)$$

均不存在，当然两个二次极限也不存在。但是由 $|f(x,y)|\leqslant|x|+|y|$ 知

$$\lim_{\substack{x\to0\\y\to0}}f(x,y)=0.$$

(2) 两个二次极限存在但可能不相等。例如

$$f(x,y)=\begin{cases}\dfrac{x^2+x+y^2-y}{x+y}, & \text{若 } x+y\neq0,\\ 0, & \text{若 } x+y=0,\end{cases}$$

由于当 $y\neq0$ 时 $\lim\limits_{x\to0}f(x,y)=y-1$，故

$$\lim_{y\to0}\lim_{x\to0}f(x,y)=-1,$$

同理可得

$$\lim_{x\to0}\lim_{y\to0}f(x,y)=1,$$

(3) 两个二次极限存在且相等，但二重极限仍有可能不存在。例如

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

容易看出 $\lim\limits_{x\to0}\lim\limits_{y\to0}f(x,y)=\lim\limits_{y\to0}\lim\limits_{x\to0}f(x,y)=0$，但由例 2.7 知二重极限 $\lim\limits_{\substack{x\to0\\y\to0}}f(x,y)$ 不存在。

(4) 二重极限和一个二次极限存在，但另一个二次极限仍有可能不存在。例如

$$f(x,y)=\begin{cases}x+y\sin\dfrac{1}{x}, & \text{若 } x\neq0,\\ 0, & \text{若 } x=0,\end{cases}$$

由 $|f(x,y)|\leqslant|x|+|y|$ 知 $\lim\limits_{\substack{x\to0\\y\to0}}f(x,y)=0$，此外还有 $\lim\limits_{x\to0}\lim\limits_{y\to0}f(x,y)=0$，但是 $\lim\limits_{y\to0}\lim\limits_{x\to0}f(x,y)$ 不存在。

综上我们得知，为了建立二重极限与二次极限之间的联系，必须加入更苛刻的条件。

**【命题 2.13】设 $\lim\limits_{\substack{x\to a\\y\to b}}f(x,y)=A$（有限或无穷），且在 $b$ 的某去心邻域内存在有限极限 $\lim\limits_{x\to a}f(x,y)=\varphi(y)$，则**

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

<u>证明</u>. 我们仅讨论 $A\in\mathbb{R}$ 的情形。由 $\lim\limits_{\substack{x\to a\\y\to b}}f(x,y)=A$ 知，对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于满足 $0<\sqrt{(x-a)^2+(y-b)^2}<\delta$ 的任意点 $(x,y)$ 均有

$$|f(x,y)-A|<\varepsilon.$$

现对任意给定的满足 $0<|y-b|<\delta$ 的 $y$，令 $x\to a$ 可得

$$|\varphi(y)-A|\leqslant\varepsilon,$$

这说明 $\lim\limits_{y\to b}\lim\limits_{x\to a}f(x,y)=\lim\limits_{y\to b}\varphi(y)=A$。 □

## 习 题 12.2

1. 用精确的数学语言表述:

(1) $\lim\limits_{\substack{x\to a^+\\y\to b^-}}f(x,y)=A$;　　(2) $\lim\limits_{\substack{x\to a\\y\to+\infty}}f(x,y)=A$;

(3) $\lim\limits_{\substack{x\to\infty\\y\to b}}f(x,y)=-\infty$;　　(4) $\lim\limits_{\substack{x\to+\infty\\y\to-\infty}}f(x,y)=+\infty$;

(5) $\lim\limits_{\substack{x\to a\\y\to b}}f(x,y)\neq A$;　　(6) $\lim\limits_{\substack{x\to\infty\\y\to\infty}}f(x,y)\neq\infty$。

2. 设 $f$ 在 $\mathbb{R}$ 上有定义且在任意有限闭区间上可积，证明 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ 收敛的充要条件是极限

$$\lim_{\substack{A\to-\infty\\B\to+\infty}} \int_A^B f(x)\,\mathrm{d}x$$

存在。

3. 判断下列二重极限是否存在，如果存在，计算其值：

(1) $\lim\limits_{\substack{x\to0\\y\to0}} \dfrac{x^3y^2}{x^4+y^4}$； (2) $\lim\limits_{\substack{x\to+\infty\\y\to+\infty}} \dfrac{x+y}{\sqrt{x^2-xy+y^2}}$；

(3) $\lim\limits_{\substack{x\to0\\y\to0}} \dfrac{x^3+y^3}{x^2+y^2}$； (4) $\lim\limits_{\substack{x\to+\infty\\y\to+\infty}} (x^2+y^2)e^{-(x+y)}$；

(5) $\lim\limits_{\substack{x\to+\infty\\y\to+\infty}} \left(\dfrac{xy}{x^2+y^2}\right)^{x^2}$； (6) $\lim\limits_{\substack{x\to0^+\\y\to0^+}} \dfrac{e^{xy}-1}{x+\sin y}$；

(7) $\lim\limits_{\substack{x\to0^+\\y\to0^+}} x^y$。

4. 证明下列二元函数的两个累次极限 $\lim\limits_{y\to0}\lim\limits_{x\to0} f(x,y)$ 与 $\lim\limits_{x\to0}\lim\limits_{y\to0} f(x,y)$ 均存在且相等，但二重极限 $\lim\limits_{\substack{x\to0\\y\to0}} f(x,y)$ 不存在：

(1) $f(x,y)=\dfrac{\log(1+xy)}{x^2+y^2\cos y}$；

(2) $f(x,y)=\dfrac{x^2y^2}{x^2y^2+(x-y)^2}$。

假设 $\left(a_{(m,n)}\right)_{(m,n)\in\mathbb{Z}^2_{>0}}$ 是 $\mathbb{R}$ 的以 $\mathbb{Z}^2_{>0}$ 为指标集的一个元素族，为方便起见我们把它记作 $(a_{m,n})$。现对任意的正整数 $M$, $N$ 记

$$S_{M,N}=\sum_{m=1}^{M}\sum_{n=1}^{N} a_{m,n},$$

如果二重极限 $\lim\limits_{\substack{M\to\infty\\N\to\infty}} S_{M,N}$ 存在，则称<u>二重级数</u> $\sum\limits_{m,n=1}^{\infty} a_{m,n}$ 收敛，并将这一极限值称为该二重级数的和；如果二重极限 $\lim\limits_{\substack{M\to\infty\\N\to\infty}} S_{M,N}$ 不存在，则称二重级数 $\sum\limits_{m,n=1}^{\infty} a_{m,n}$ 发散。从第 5 题到第 10 题是一组题，介绍二重级数的一些性质和例子。

5. 设二重级数 $\sum\limits_{m,n=1}^{\infty} a_{m,n}$ 收敛，且对任意给定的 $m$，级数 $\sum\limits_{n=1}^{\infty} a_{m,n}$ 均收敛，证

明级数 $\sum\limits_{m=1}^{\infty}\sum\limits_{n=1}^{\infty}a_{m,n}$ 收敛，且有

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}a_{m,n}=\sum_{m,n=1}^{\infty}a_{m,n}$$

6. 设 $a_{m,n}\geqslant 0$ $(\forall\ m,n\geqslant 1)$，证明：当三个级数 $\sum\limits_{m,n=1}^{\infty}a_{m,n}$，$\sum\limits_{m=1}^{\infty}\sum\limits_{n=1}^{\infty}a_{m,n}$ 与 $\sum\limits_{n=1}^{\infty}\sum\limits_{m=1}^{\infty}a_{m,n}$ 中有一个收敛时，其余两个也收敛，且三者有相同的和。
7. 设 $\sum\limits_{m,n=1}^{\infty}|a_{m,n}|$ 收敛[⑧]，证明 $\sum\limits_{m,n=1}^{\infty}a_{m,n}$ 收敛。
8. 设 $\sum\limits_{m,n=1}^{\infty}|a_{m,n}|$ 收敛，那么将 $a_{m,n}$ $(m,n\geqslant 1)$ 按任意方式排成一列所建立的级数都绝对收敛，且所有的级数的和均等于 $\sum\limits_{m,n=1}^{\infty}a_{m,n}$。
9. （Goldbach）设 $\{u_k\}$ 是将形如 $m^n$ $(m,n\in\mathbb{Z}_{>1})$ 的整数排成一列所得的数列，证明

$$\sum_{k=1}^{\infty}\frac{1}{u_k-1}=1.$$

10. 判断二重级数 $\sum\limits_{m,n=1}^{\infty}\dfrac{1}{(m+n^2)^{\alpha}}$ 的敛散性。

## § 12.3 连续映射

【**定义 3.1**】设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}^m$。又设 $\boldsymbol{a}\in E$。若对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对任意的 $\boldsymbol{x}\in E\cap B(\boldsymbol{a},\delta)$ 均有

$$|f(\boldsymbol{x})-f(\boldsymbol{a})|<\varepsilon,$$

则称 $f$ 在 $\boldsymbol{a}$ 处<u>连续</u>。如果 $f$ 在 $E$ 的每一点处均连续，则称 $f$ 在 $E$ 上<u>连续</u>。

【**注 3.2**】按照上述定义，$E$ 上的任一映射 $f$ 在 $E$ 的孤立点处总是连续的。

---

[⑧] 此时我们称二重级数 $\sum\limits_{m,n=1}^{\infty}a_{m,n}$ <u>绝对收敛</u>。

【**例 3.3**】从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射是 $\mathbb{R}^n$ 上的连续映射。

<u>证明</u>. 设 $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ 是线性映射，则必存在 $m\times n$ 阶矩阵 $A=(\alpha_{ij})$ 使得

$$f(\boldsymbol{x})=A\boldsymbol{x}.$$

我们先证明

$$|A\boldsymbol{x}|\leqslant M\sqrt{mn}\cdot|\boldsymbol{x}|, \tag{12.1}$$

其中

$$M=\max\{|\alpha_{ij}|:1\leqslant i\leqslant m,\ 1\leqslant j\leqslant n\}.$$

事实上，

$$\begin{aligned}|A\boldsymbol{x}|&=\left[\left(\sum_{j=1}^{n}\alpha_{1j}x_j\right)^2+\cdots+\left(\sum_{j=1}^{n}\alpha_{mj}x_j\right)^2\right]^{\frac{1}{2}}\\&\leqslant\left[\left(\sum_{j=1}^{n}\alpha_{1j}^2\right)\left(\sum_{j=1}^{n}x_j^2\right)+\cdots+\left(\sum_{j=1}^{n}\alpha_{mj}^2\right)\left(\sum_{j=1}^{n}x_j^2\right)\right]^{\frac{1}{2}}\\&=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_{ij}^2\right)^{\frac{1}{2}}|\boldsymbol{x}|\leqslant M\sqrt{mn}\cdot|\boldsymbol{x}|.\end{aligned}$$

现考虑 $\mathbb{R}^n$ 中任一点 $\boldsymbol{a}=(a_1,\cdots,a_n)^{\mathrm{T}}$。对任意的 $\varepsilon>0$，取 $\delta=\dfrac{\varepsilon}{(M+1)\sqrt{mn}}$，则对满足 $|\boldsymbol{x}-\boldsymbol{a}|<\delta$ 的任一 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 有

$$|f(\boldsymbol{x})-f(\boldsymbol{a})|=|A(\boldsymbol{x}-\boldsymbol{a})|\leqslant M\sqrt{mn}\cdot|\boldsymbol{x}-\boldsymbol{a}|<\varepsilon,$$

故 $f$ 在 $\boldsymbol{a}$ 处连续。再由 $\boldsymbol{a}$ 的任意性知 $f$ 是 $\mathbb{R}^n$ 上的连续映射。 □

下述定理是连续性的一个本质刻画。

【**定理 3.4**】设 $E\subseteq\mathbb{R}^n$ 且 $f:E\longrightarrow\mathbb{R}^m$，则下列命题等价：

(1) $f$ 在 $E$ 上连续；

(2) 对 $\mathbb{R}^m$ 中任意的开集 $G$，$f^{-1}(G)$ 均是 $E$ 的开子集；

(3) 对 $\mathbb{R}^m$ 中任意的闭集 $F$，$f^{-1}(F)$ 均是 $E$ 的闭子集。

<u>证明</u>. (1)⇒(2)：设 $G$ 是 $\mathbb{R}^m$ 中的开集。对任意的 $\boldsymbol{a}\in f^{-1}(G)$ 有 $f(\boldsymbol{a})\in G$，于是存在 $\varepsilon>0$ 使得 $B(f(\boldsymbol{a}),\varepsilon)\subseteq G$。对于这个 $\varepsilon$，由 $f$ 的连续性知存在 $\delta>0$，使得对任意的 $\boldsymbol{x}\in E\cap B(\boldsymbol{a},\delta)$ 均有

$$|f(\boldsymbol{x})-f(\boldsymbol{a})|<\varepsilon,$$

也即 $f(\boldsymbol{x})\in B(f(\boldsymbol{a}),\varepsilon)\subseteq G$。换句话说，当 $\boldsymbol{x}\in E\cap B(\boldsymbol{a},\delta)$ 时有 $\boldsymbol{x}\in f^{-1}(G)$，因此

$$E\cap B(\boldsymbol{a},\delta)\subseteq f^{-1}(G).$$

于是由命题 1.28 (1) 知 $f^{-1}(G)$ 是 $E$ 的开子集。

(2)⇒(1)：设 $\boldsymbol{a}\in E$。对任意的 $\varepsilon>0$，$B(f(\boldsymbol{a}),\varepsilon)$ 是 $\mathbb{R}^m$ 中的开集，因此由 (2) 知 $f^{-1}\big(B(f(\boldsymbol{a}),\varepsilon)\big)$ 是 $E$ 的开子集。注意到 $\boldsymbol{a}\in f^{-1}\big(B(f(\boldsymbol{a}),\varepsilon)\big)$，故由命题 1.28 知存在 $\delta>0$ 使得 $E\cap B(\boldsymbol{a},\delta)\subseteq f^{-1}\big(B(f(\boldsymbol{a}),\varepsilon)\big)$。这意味着对任意的 $x\in E\cap B(\boldsymbol{a},\delta)$ 有 $x\in f^{-1}\big(B(f(\boldsymbol{a}),\varepsilon)\big)$，也即 $f(x)\in B(f(\boldsymbol{a}),\varepsilon)$。这就证明了 $f$ 在 $\boldsymbol{a}$ 处的连续性。再由 $\boldsymbol{a}$ 的任意性知 $f$ 在 $E$ 上连续。

(2)⇔(3)：这可由 $f^{-1}(\mathbb{R}^m\setminus F)=E\setminus f^{-1}(F)$ 及命题 1.28 (2) 得到。 □

此外，由命题 2.3 可直接得到下述结论。

**【命题 3.5】设 $E\subseteq\mathbb{R}^n$ 且 $f=(f_1,\cdots,f_m)^{\mathrm{T}}:E\longrightarrow\mathbb{R}^m$，那么 $f$ 是 $E$ 上的连续映射当且仅当每个 $f_j$ $(1\leqslant j\leqslant m)$ 均是 $E$ 上的连续函数。**

上述命题把判定 $f$ 是否为连续映射的问题转换为讨论分量函数是否为连续函数，而由上节命题 2.4，2.5 及 2.10 知多元连续函数的和、差、积、商以及复合均为连续函数（商的情形要求分母不为 0）。回忆起我们在 §4.5 中证明了初等函数的连续性，因此立刻可以得出诸如

$$\sin(x+y^2),\qquad \log(x+y+\tan z),\qquad \frac{e^{x-y}}{z^2+1}$$

的函数在其定义域上连续。

下面我们来介绍连续映射的性质。

**【定理 3.6】设 $f:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ 是连续映射。若 $K$ 是 $\mathbb{R}^n$ 中的紧集，则 $f(K)$ 是 $\mathbb{R}^m$ 中的紧集。**

证明. 设 $(G_\lambda)_{\lambda\in L}$ 是 $f(K)$ 的一个开覆盖，则由 $f$ 是连续映射知，每个 $f^{-1}(G_\lambda)$ 均是开集。此外，

$$K\subseteq f^{-1}(f(K))\subseteq f^{-1}\left(\bigcup_{\lambda\in L}G_\lambda\right)=\bigcup_{\lambda\in L}f^{-1}(G_\lambda),$$

因此 $(f^{-1}(G_\lambda))_{\lambda\in L}$ 是 $K$ 的一个开覆盖。于是由 $K$ 是紧集知有有限子覆盖，即存在 $\lambda_1,\cdots,\lambda_\ell\in L$ 使得

$$K\subseteq\bigcup_{j=1}^{\ell}f^{-1}(G_{\lambda_j}).$$

进而有

$$f(K) \subseteq f\left(\bigcup_{j=1}^{\ell} f^{-1}(G_{\lambda_j})\right) = \bigcup_{j=1}^{\ell} f(f^{-1}(G_{\lambda_j})) \subseteq \bigcup_{j=1}^{\ell} G_{\lambda_j}.$$

因此 $f(K)$ 的开覆盖 $(G_\lambda)_{\lambda\in L}$ 有有限子覆盖。这就证明了 $f(K)$ 是紧集。 □

由 $\mathbb{R}$ 中的紧集是有界闭集不难得出下面的推论（留作练习）。

**【推论 3.7】**设 $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ 是连续函数。若 $K$ 是 $\mathbb{R}^n$ 中的紧集，则 $f$ 在 $K$ 上有界，且能取到最大值和最小值。

**【定理 3.8】**设 $E$ 是 $\mathbb{R}^n$ 中的连通集，且 $f: E \longrightarrow \mathbb{R}^m$ 是连续映射，则 $f(E)$ 是 $\mathbb{R}^m$ 中的连通集。

证明. 反设 $f(E)$ 不是 $\mathbb{R}^m$ 中的连通集，则存在 $f(E)$ 的两个非空开子集 $A$ 和 $B$ 使得

$$A\cup B = f(E) \qquad \text{且} \qquad A\cap B = \varnothing.$$

于是

$$E = f^{-1}(f(E)) = f^{-1}(A\cup B) = f^{-1}(A)\cup f^{-1}(B). \tag{12.2}$$

首先，由 $A, B$ 均非空知 $f^{-1}(A)$ 与 $f^{-1}(B)$ 均是 $E$ 的非空子集。其次，由 $A$ 是 $f(E)$ 的开子集知存在 $\mathbb{R}^m$ 中的开集 $G$ 使得 $A = f(E)\cap G$，于是

$$f^{-1}(A) = f^{-1}(f(E)\cap G) = f^{-1}(f(E))\cap f^{-1}(G) = E\cap f^{-1}(G),$$

而由 $f$ 连续知 $f^{-1}(G)$ 是 $E$ 的开子集（参见命题 3.4），因此 $f^{-1}(A)$ 也是 $E$ 的开子集，同理 $f^{-1}(B)$ 是 $E$ 的开子集。最后，由 $A\cap B=\varnothing$ 知

$$f^{-1}(A)\cap f^{-1}(B) = f^{-1}(A\cap B) = \varnothing.$$

综上得知，(12.2) 是将 $E$ 写成两个不相交的非空开子集的并的一种方式，这与 $E$ 是连通集矛盾。 □

**【例 3.9】**因为函数 $f(x)=\sin\dfrac{1}{x}$ 是区间 $(0,+\infty)$ 上的连续函数，所以由命题 3.5 知映射

$$g: x \longmapsto \left(x, \sin\frac{1}{x}\right)$$

是从 $(0,+\infty)$ 到 $\mathbb{R}^2$ 的连续映射。注意到 $(0,+\infty)$ 是 $\mathbb{R}$ 中的连通集，故而集合

$$D = \left\{\left(x, \sin\frac{1}{x}\right) : x>0\right\}$$

作为 $(0,+\infty)$ 在 $g$ 下的像是 $\mathbb{R}^2$ 中的连通集。最后，由命题 1.32 知

$$\overline{D}=\left\{\left(x,\sin\frac{1}{x}\right):x>0\right\}\cup\big\{(0,y):-1\leqslant y\leqslant 1\big\}$$

也是 $\mathbb{R}^2$ 中的连通集。

**【例 3.10】** $\mathbb{R}^n$ 的子集 $S$ 被称为是凸集 (convex set)，如果对任意的 $\boldsymbol{x},\boldsymbol{y}\in S$ 均有

$$\{(1-\lambda)\boldsymbol{x}+\lambda\boldsymbol{y}:\lambda\in[0,1]\}\subseteq S.$$

为方便起见，我们将上式左边的集合记作 $[\boldsymbol{x},\boldsymbol{y}]$[9]。证明：$\mathbb{R}^n$ 中的凸集必是连通集。

证明. 设 $S$ 是 $\mathbb{R}^n$ 中的凸集，若它不是连通集，则存在 $S$ 的非空开子集 $A$ 和 $B$ 满足

$$A\cup B=S\qquad 且\qquad A\cap B=\varnothing.$$

现取 $\boldsymbol{a}\in A$ 及 $\boldsymbol{b}\in B$，由 $S$ 是凸集可得 $[\boldsymbol{a},\boldsymbol{b}]\subseteq S$。此外，由 $A$ 是 $S$ 的开子集知存在 $\mathbb{R}^n$ 中的开集 $G$ 使得 $A=S\cap G$，因此 $A\cap[\boldsymbol{a},\boldsymbol{b}]=[\boldsymbol{a},\boldsymbol{b}]\cap G$ 是 $[\boldsymbol{a},\boldsymbol{b}]$ 的非空开子集，同理 $B\cap[\boldsymbol{a},\boldsymbol{b}]$ 也是 $[\boldsymbol{a},\boldsymbol{b}]$ 的非空开子集。并且

$$\big(A\cap[\boldsymbol{a},\boldsymbol{b}]\big)\cup\big(B\cap[\boldsymbol{a},\boldsymbol{b}]\big)=[\boldsymbol{a},\boldsymbol{b}],\qquad \big(A\cap[\boldsymbol{a},\boldsymbol{b}]\big)\cap\big(B\cap[\boldsymbol{a},\boldsymbol{b}]\big)=\varnothing.$$

这说明 $[\boldsymbol{a},\boldsymbol{b}]$ 不是连通集。然而由定理 3.8 知，$[\boldsymbol{a},\boldsymbol{b}]$ 作为 $[0,1]$ 在连续映射 $f(\lambda)=(1-\lambda)\boldsymbol{a}+\lambda\boldsymbol{b}$ 下的像必是连通的，从而得出矛盾。 □

因为 $\mathbb{R}$ 中的连通集必是区间（命题 1.31），所以有下面的介值定理 (intermediate value theorem)。

**【推论 3.11】（介值定理）** 设 $E$ 是 $\mathbb{R}^n$ 中的连通集，且 $f:E\longrightarrow\mathbb{R}$ 是连续函数。若 $\boldsymbol{a},\boldsymbol{b}\in E$ 满足 $f(\boldsymbol{a})<f(\boldsymbol{b})$，则对任意的 $\xi\in(f(\boldsymbol{a}),f(\boldsymbol{b}))$，存在 $\boldsymbol{c}\in E$ 使得 $f(\boldsymbol{c})=\xi$。

最后我们来介绍一致连续映射。

**【定义 3.12】** 设 $E\subseteq\mathbb{R}^n$ 且 $f:E\longrightarrow\mathbb{R}^m$。若对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于满足 $|\boldsymbol{x}-\boldsymbol{y}|<\delta$ 的任意 $\boldsymbol{x},\boldsymbol{y}\in E$ 均有

$$|f(\boldsymbol{x})-f(\boldsymbol{y})|<\varepsilon,$$

---

[9] 这个集合可以理解为连接 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的“线段”。

则称 $f$ 在 $E$ 上一致连续。

显然，$E$ 上的一致连续映射必在 $E$ 上连续。

**【例 3.13】** 由 (12.1) 知，从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射在 $\mathbb{R}^n$ 上一致连续。

类似于 Heine–Cantor 定理（第四章定理 8.3），我们有下述结论。

**【定理 3.14】** 设 $K$ 是 $\mathbb{R}^n$ 中的紧集，$f: K \longrightarrow \mathbb{R}^m$ 是 $K$ 上的连续映射，那么 $f$ 在 $K$ 上一致连续。

证明. 任给 $\varepsilon > 0$，由 $f$ 在 $K$ 上连续知，对任意的 $\boldsymbol{a} \in K$，存在 $\delta_{\boldsymbol{a}} > 0$ 使得

$$|f(\boldsymbol{x}) - f(\boldsymbol{a})| < \frac{\varepsilon}{2}, \qquad \forall\, \boldsymbol{x} \in B(\boldsymbol{a}, \delta_{\boldsymbol{a}}) \cap K. \tag{12.3}$$

注意到

$$K \subseteq \bigcup_{\boldsymbol{a} \in K} B\Big(\boldsymbol{a}, \frac{1}{2}\delta_{\boldsymbol{a}}\Big),$$

故由 $K$ 是紧集知存在 $K$ 中元素 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_\ell$ 使得

$$K \subseteq \bigcup_{j=1}^{\ell} B\Big(\boldsymbol{a}_j, \frac{1}{2}\delta_{\boldsymbol{a}_j}\Big).$$

记 $\delta = \dfrac{1}{2}\min\limits_{1 \leqslant j \leqslant \ell} \delta_{\boldsymbol{a}_j}$，并考虑 $K$ 中满足 $|\boldsymbol{x} - \boldsymbol{y}| < \delta$ 的任意两点 $\boldsymbol{x}$ 与 $\boldsymbol{y}$。由 $\boldsymbol{x} \in K$ 知存在 $j_0 \in [1, \ell]$ 使得 $\boldsymbol{x} \in B\Big(\boldsymbol{a}_{j_0}, \dfrac{1}{2}\delta_{\boldsymbol{a}_{j_0}}\Big)$。再由 $|\boldsymbol{x} - \boldsymbol{y}| < \delta \leqslant \dfrac{1}{2}\delta_{\boldsymbol{a}_{j_0}}$ 可得

$$|\boldsymbol{y} - \boldsymbol{a}_{j_0}| \leqslant |\boldsymbol{y} - \boldsymbol{x}| + |\boldsymbol{x} - \boldsymbol{a}_{j_0}| < \frac{1}{2}\delta_{\boldsymbol{a}_{j_0}} + \frac{1}{2}\delta_{\boldsymbol{a}_{j_0}} = \delta_{\boldsymbol{a}_{j_0}}.$$

因此 $\boldsymbol{x}, \boldsymbol{y}$ 均属于 $B(\boldsymbol{a}_{j_0}, \delta_{\boldsymbol{a}_{j_0}})$，于是由 (12.3) 知

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leqslant |f(\boldsymbol{x}) - f(\boldsymbol{a}_{j_0})| + |f(\boldsymbol{y}) - f(\boldsymbol{a}_{j_0})| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

所以 $f$ 在 $K$ 上一致连续。 □

## 习 题 12.3

1. 证明命题 3.5。
2. 证明推论 3.7。
3. 利用 Bolzano–Weierstrass 定理证明定理 3.14。

4. 计算下列二重极限：

(1) $\lim\limits_{\substack{x\to 1\\ y\to 0}} \dfrac{\log(x+e^y)}{\sqrt{x^2+y^2}}$；   (2) $\lim\limits_{\substack{x\to \infty\\ y\to 2}} \left(1+\dfrac{1}{3x}\right)^{\frac{x^2}{x+y}}$；

(3) $\lim\limits_{\substack{x\to 0\\ y\to 0}} (x^2+y^2)^{x^2y^2}$。

5. 证明 $f:(a,\boldsymbol{x})\longmapsto a\boldsymbol{x}$ 是从 $\mathbb{R}\times\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的连续映射。

6. 设 $A$ 是一个 $n\times n$ 矩阵，证明：对任意的 $\boldsymbol{x}$, $\boldsymbol{y}\in\mathbb{R}^n$ 有 $|\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y}| = O(|\boldsymbol{x}|\cdot|\boldsymbol{y}|)$。

7. 设 $A$ 是一个 $n\times n$ 矩阵，证明由 $\varphi(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 所定义的映射 $\varphi:\mathbb{R}^n\longrightarrow\mathbb{R}$ 是连续的。

8. 设 $\varphi\in C^1((a,b))$。对任意的 $x, y\in(a,b)$，$x\neq y$，定义

$$f(x,y)=\frac{\varphi(x)-\varphi(y)}{x-y}.$$

问能否对 $x\in(a,b)$ 定义 $f(x,x)$ 的值使得 $f$ 在 $(a,b)^2$ 上连续？

9. 设 $L:\mathbb{R}^n\longrightarrow\mathbb{R}^n$ 是非奇异的线性映射[10]，证明 $L$ 将 $\mathbb{R}^n$ 中的开集映为开集，将 $\mathbb{R}^n$ 中的闭集映为闭集。

10. 举例说明一般的连续映射未必把开集映为开集。

11. 设 $E\subseteq\mathbb{R}^n$，$f$ 与 $g$ 均是从 $E$ 到 $\mathbb{R}^m$ 的连续映射，证明集合

$$\{\boldsymbol{x}\in E: f(\boldsymbol{x})=g(\boldsymbol{x})\}$$

是 $E$ 的闭子集。

12. 设 $f$ 与 $g$ 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的两个连续映射，$E$ 是 $\mathbb{R}^n$ 的一个稠密子集，并且对任意的 $\boldsymbol{x}\in E$ 有 $f(\boldsymbol{x})=g(\boldsymbol{x})$。证明 $f=g$。

13. 设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}$。对任意的 $\boldsymbol{a}\in E$ 记

$$\omega(f;\boldsymbol{a})=\inf_{\delta>0}\Big(\sup_{\boldsymbol{x}\in B(\boldsymbol{a},\delta)\cap E} f(\boldsymbol{x})-\inf_{\boldsymbol{x}\in B(\boldsymbol{a},\delta)\cap E} f(\boldsymbol{x})\Big),$$

并称之为 $f$ 在 $\boldsymbol{a}$ 处的<u>振幅</u>。证明 $f$ 在 $\boldsymbol{a}$ 处连续的充要条件是 $\omega(f;\boldsymbol{a})=0$。

14. 设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}^m$。证明 $f$ 在 $E$ 上连续的充要条件是：对任意的 $A\subseteq\mathbb{R}^m$ 有 $\overline{f^{-1}(A)}\cap E\subseteq f^{-1}(\overline{A})$。

[10] 所谓非奇异，是指它在标准基下所对应的矩阵是可逆矩阵。

15. 设 $E$ 是 $\mathbb{R}^n$ 中的开集且 $f: E \longrightarrow \mathbb{R}^m$。证明 $f$ 在 $E$ 上连续的充要条件是：对任意的 $A \subseteq \mathbb{R}^m$ 有 $f^{-1}(A^\circ) \subseteq (f^{-1}(A))^\circ$。

16. 设 $K$ 是 $\mathbb{R}^n$ 中的非空紧集，$f: K \longrightarrow K$ 满足

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| < |\boldsymbol{x} - \boldsymbol{y}|, \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in K,\ \boldsymbol{x} \neq \boldsymbol{y}.$$

证明存在唯一的 $\boldsymbol{a} \in K$ 使得 $f(\boldsymbol{a}) = \boldsymbol{a}$。

17. 设

$$E_1 = \{(x, y) : x \in [0, 1] \cap \mathbb{Q},\ y \in [0, 1]\},$$

$$E_2 = \{(x, y) : x \in [0, 1] \setminus \mathbb{Q},\ y \in [-1, 0]\}.$$

证明 $E_1 \cup E_2$ 是 $\mathbb{R}^2$ 中的连通集。

18. 设 $E$ 是 $\mathbb{R}^n$ 中的凸集，证明 $\overline{E}$ 也是 $\mathbb{R}^n$ 中的凸集。

19. 设 $A$ 是 $\mathbb{R}^n$ 的一个给定的非空子集。对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ 记 $d(\boldsymbol{x}, A) = \inf\limits_{\boldsymbol{a} \in A} |\boldsymbol{x} - \boldsymbol{a}|$。证明 $d(\boldsymbol{x}, A)$ 是 $\mathbb{R}^n$ 上的一致连续函数。

20. 判断下列函数 $f$ 在指定集合 $E$ 上是否一致连续：

(1) $f(x, y) = \sqrt{x^2 + y^2}$，$E = \mathbb{R}^2$。

(2) $f(x, y) = \cos(xy)$，$E = \mathbb{R}^2$。

(3) $f(x, y) = \sin \dfrac{\pi}{1 - x^2 - y^2}$，$E = B(\mathbf{0}, 1)$。

(4) $f(x, y) = x^y$，$E = (0, 1)^2$。

21. 设 $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 在 $\mathbb{R}^n$ 上一致连续，证明对任意的 $\boldsymbol{x} \in \mathbb{R}^n \setminus B(\mathbf{0}, 1)$ 有 $|f(\boldsymbol{x})| = O(|\boldsymbol{x}|)$。

22. 设 $f$ 是从 $\mathbb{Q}^n$ 到 $\mathbb{R}^m$ 的一致连续映射，证明 $f$ 可被延拓为从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的一致连续映射。

# 第十三章

# 多元函数的微分

……微积分学的基本思想，即用线性函数“局部”逼近函数。在微积分学的经典教学中，这种思想被下述偶然的事实直接掩盖着，即在一维向量空间中，线性型与数之间存在一一对应，因此在一点处的导数被定义为数而不是线性型。在处理多变量函数时，不惜一切地墨守数值解释的老调会使情况变得更糟。例如，如此得到的复合函数偏导数的经典公式已失去了所有直观意义的痕迹。

—— J. Dieudonné

## §13.1 微分的定义

在第五章中我们介绍了一元函数导数的概念。函数 $f(x)$ 被称为是在点 $x_0$ 处可导（或可微）的，如果极限

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}$$

存在，并且上述极限值被称为 $f(x)$ 在 $x_0$ 处的导数。该定义方式在推广至高维时会遇到困难，这是因为如果 $f:\mathbb{R}^n\longrightarrow\mathbb{R}^m$，那么对 $\boldsymbol{x}_0,\boldsymbol{h}\in\mathbb{R}^n$，分式

$$\frac{f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)}{\boldsymbol{h}}$$

一般来说是没有意义的，因此我们必须改变一下思路。回忆起一元函数 $f(x)$ 在 $x_0$ 处可导当且仅当存在 $A\in\mathbb{R}$ 使得

$$f(x_0+h)=f(x_0)+Ah+o(h)$$

（参见第五章命题 2.2），也即是说，可以在 $x_0$ 附近用线性函数逼近 $f(x)$。因此对于一般的映射 $f$，我们自然地考虑能否在局部用线性映射去逼近 $f$，这就引出了下述定义。

**【定义 1.1】**设 $E \subseteq \mathbb{R}^n$，$f: E \longrightarrow \mathbb{R}^m$。又设 $\boldsymbol{a}$ 是 $E$ 的一个内点。若存在线性映射 $L: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 使得

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - L\boldsymbol{h}}{|\boldsymbol{h}|} = \boldsymbol{0}, \tag{13.1}$$

则称 $f$ 在 $\boldsymbol{a}$ 处可微 (differentiable)。若 $f$ 在 $E$ 中每个点处均可微，我们就称 $f$ 在 $E$ 上可微。

**【注 1.2】**定义 1.1 中的 $L\boldsymbol{h}$ 表示 $L$ 在 $\boldsymbol{h}$ 上的作用，之前我们通常把它记作 $L(\boldsymbol{h})$，但现在为了方便起见将它写成 $L\boldsymbol{h}$。

值得一提的是，定义 1.1 中的线性映射 $L$ 如果存在，则必唯一。事实上，若 $L_1$ 与 $L_2$ 均是满足 (13.1) 的线性映射，则

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{L_1\boldsymbol{h} - L_2\boldsymbol{h}}{|\boldsymbol{h}|} = \boldsymbol{0},$$

于是对于 $\mathbb{R}^n$ 中的任一给定的非零元 $\boldsymbol{u}$ 有

$$\lim_{k\to 0^+} \frac{L_1(k\boldsymbol{u}) - L_2(k\boldsymbol{u})}{|k\boldsymbol{u}|} = \boldsymbol{0},$$

此即 $\lim\limits_{k\to 0^+} \dfrac{L_1\boldsymbol{u} - L_2\boldsymbol{u}}{|\boldsymbol{u}|} = \boldsymbol{0}$，从而必有 $L_1\boldsymbol{u} = L_2\boldsymbol{u}$。再由 $\boldsymbol{u}$ 的任意性知 $L_1 = L_2$。于是我们可以给出如下定义。

**【定义 1.3】**设 $E \subseteq \mathbb{R}^n$，$f: E \longrightarrow \mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的一个内点。若 $f$ 在 $\boldsymbol{a}$ 处可微，则称 (13.1) 中的线性映射 $L$ 为 $f$ 在 $\boldsymbol{a}$ 处的微分 (differential)[1]，记作 $f'(\boldsymbol{a})$ 或 $Df(\boldsymbol{a})$。

我们用 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 表示从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的全体线性映射所成之集，它是 $\mathbb{R}$ 上的一个线性空间。当 $f$ 在 $E$ 上可微时，$f': \boldsymbol{x} \mapsto f'(\boldsymbol{x})$ 是从 $E$ 到 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 的映射，我们称之为导映射。

由于从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射必对应于一个 $m \times n$ 矩阵，并且在选定了基底的前提下这个矩阵是唯一确定的。因此为了方便起见，我们也把线性映射 $f'(\boldsymbol{a})$ 关于

[1] 注意与 §5.2 中的“微分”相区分。

$\mathbb{R}^n$ 和 $\mathbb{R}^m$ 的标准基[②]的矩阵记作 $f'(\boldsymbol{a})$。换句话说，$f'(\boldsymbol{a})$ 既表示线性映射，也表示线性映射在标准基下对应的矩阵。

**【例 1.4】**设 $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ 为常值映射，则对任意的 $\boldsymbol{x}\in\mathbb{R}^n$，$f'(\boldsymbol{x})$ 均是零映射（即将 $\mathbb{R}^n$ 中每个元素映为 $\mathbb{R}^m$ 中零元的映射），其对应的矩阵为零矩阵。

**【例 1.5】**若 $f(\boldsymbol{x}) = L\boldsymbol{x}+\boldsymbol{b}$，其中 $L\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 且 $\boldsymbol{b}\in\mathbb{R}^m$，则由

$$f(\boldsymbol{x}+\boldsymbol{h}) = f(\boldsymbol{x}) + L\boldsymbol{h}$$

知 $f'(\boldsymbol{x}) = L\ (\forall\ \boldsymbol{x}\in\mathbb{R}^n)$。

下面我们来介绍可微映射的一些基本性质。

**【命题 1.6】**若 $f$ 在 $\boldsymbol{a}$ 处可微，则 $f$ 在 $\boldsymbol{a}$ 处连续。

<u>证明</u>. 若 $f$ 在 $\boldsymbol{a}$ 处可微，则

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - f'(\boldsymbol{a})\boldsymbol{h}}{|\boldsymbol{h}|} = \boldsymbol{0},$$

因此

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \big(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})\big) = \lim_{\boldsymbol{h}\to\boldsymbol{0}} \left(\frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - f'(\boldsymbol{a})\boldsymbol{h}}{|\boldsymbol{h}|}\cdot|\boldsymbol{h}| + f'(\boldsymbol{a})\boldsymbol{h}\right) = \boldsymbol{0},$$

上面最后一步用到了线性映射的连续性（参见第十二章例 3.3）。 □

**【命题 1.7】**设 $E\subseteq\mathbb{R}^n$，$f = (f_1,\cdots,f_m)^{\mathrm{T}}: E\longrightarrow\mathbb{R}^m$。又设 $\boldsymbol{a}$ 是 $E$ 的一个内点。那么 $f$ 在 $\boldsymbol{a}$ 处可微的充要条件是每个分量函数 $f_j$ 均在 $\boldsymbol{a}$ 处可微。此外，当 $f$ 在 $\boldsymbol{a}$ 处可微时有

$$f'(\boldsymbol{a}) = \begin{bmatrix} f_1'(\boldsymbol{a}) \\ \vdots \\ f_m'(\boldsymbol{a}) \end{bmatrix}.$$

<u>证明</u>. 对任给的一个 $m\times n$ 矩阵 $L$，令

$$F(\boldsymbol{h}) = \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - L\boldsymbol{h}}{|\boldsymbol{h}|}.$$

---

[②] 我们记 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 为 $\mathbb{R}^n$ 的标准基，其中 $\boldsymbol{e}_j$ 是第 $j$ 个分量为 1 而其余分量为 0 的向量。

若记 $L = \begin{bmatrix} L_1 \\ \vdots \\ L_m \end{bmatrix}$，那么 $L\boldsymbol{h} = \begin{bmatrix} L_1\boldsymbol{h} \\ \vdots \\ L_m\boldsymbol{h} \end{bmatrix}$。因此

$$F(\boldsymbol{h}) = \begin{bmatrix} F_1(\boldsymbol{h}) \\ \vdots \\ F_m(\boldsymbol{h}) \end{bmatrix},$$

其中

$$F_j(\boldsymbol{h}) = \frac{f_j(\boldsymbol{a}+\boldsymbol{h}) - f_j(\boldsymbol{a}) - L_j\boldsymbol{h}}{|\boldsymbol{h}|}.$$

按照第十二章命题 2.3，$\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}} F(\boldsymbol{h}) = \boldsymbol{0}$ 当且仅当对 $1 \leqslant j \leqslant m$ 有 $\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}} F_j(\boldsymbol{h}) = 0$，从而命题得证。 □

由 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 是线性空间可立即得到如下结论。

**【命题 1.8】**若 $f$ 与 $g$ 均在 $\boldsymbol{a}$ 处可微，则对任意的实数 $\alpha, \beta$ 而言，$\alpha f + \beta g$ 也在 $\boldsymbol{a}$ 处可微，且

$$(\alpha f + \beta g)'(\boldsymbol{a}) = \alpha f'(\boldsymbol{a}) + \beta g'(\boldsymbol{a}).$$

类似于第五章命题 1.16，我们有下述链式法则 (chain rule)。

**【命题 1.9】**（链式法则）设 $E \subseteq \mathbb{R}^n$，$D \subseteq \mathbb{R}^m$，$f : E \longrightarrow \mathbb{R}^m$，$g : D \longrightarrow \mathbb{R}^\ell$，且 $f(E) \subseteq D$。又设 $\boldsymbol{a}$ 是 $E$ 的内点，$\boldsymbol{b} = f(\boldsymbol{a})$ 是 $D$ 的内点。若 $f$ 在 $\boldsymbol{a}$ 处可微且 $g$ 在 $\boldsymbol{b}$ 处可微，则 $g \circ f$ 在 $\boldsymbol{a}$ 处可微，且有

$$(g \circ f)'(\boldsymbol{a}) = g'(\boldsymbol{b}) \circ f'(\boldsymbol{a}),$$

上式右边是两个线性映射的复合。若将 $(g \circ f)'(\boldsymbol{a})$，$g'(\boldsymbol{b})$ 及 $f'(\boldsymbol{a})$ 均看作是线性映射在标准基下的矩阵，那么上式也可写成

$$(g \circ f)'(\boldsymbol{a}) = g'(\boldsymbol{b}) f'(\boldsymbol{a}),$$

此时等号右边表示矩阵的乘积。

证明. 按照假设，$g$ 在 $\boldsymbol{b} = f(\boldsymbol{a})$ 的某邻域 $U \subseteq D$ 内有定义。由于 $f$ 在 $\boldsymbol{a}$ 处可微，故 $f$ 在 $\boldsymbol{a}$ 处连续，因此存在 $\boldsymbol{a}$ 的邻域 $V \subseteq E$，使得对任意的 $\boldsymbol{x} \in V$ 有 $f(\boldsymbol{x}) \in U$。这样一来，复合映射 $g \circ f$ 就在 $\boldsymbol{a}$ 的邻域 $V$ 上有定义。

由 $f$ 在 $\boldsymbol{a}$ 处可微及 $g$ 在 $\boldsymbol{b}$ 处可微知，存在 $s(\boldsymbol{h})$ 与 $t(\boldsymbol{k})$，使得当 $|\boldsymbol{h}|$，$|\boldsymbol{k}|$ 充分小时有

$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+f'(\boldsymbol{a})\boldsymbol{h}+s(\boldsymbol{h}), \tag{13.2}$$

$$g(\boldsymbol{b}+\boldsymbol{k})=g(\boldsymbol{b})+g'(\boldsymbol{b})\boldsymbol{k}+t(\boldsymbol{k}), \tag{13.3}$$

并且

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{s(\boldsymbol{h})}{|\boldsymbol{h}|}=\lim_{\boldsymbol{k}\to\boldsymbol{0}}\frac{t(\boldsymbol{k})}{|\boldsymbol{k}|}=\boldsymbol{0}.$$

由于 $\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}}[f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})]=\boldsymbol{0}$，所以我们可特取 $\boldsymbol{k}=f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})$ 并将其代入 (13.3)，从而得到

$$g(f(\boldsymbol{a}+\boldsymbol{h}))=g(f(\boldsymbol{a}))+g'(\boldsymbol{b})[f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})]+t(\boldsymbol{k}).$$

现在再将 (13.2) 代入上式右侧，便有

$$g(f(\boldsymbol{a}+\boldsymbol{h}))=g(f(\boldsymbol{a}))+\big(g'(\boldsymbol{b})\circ f'(\boldsymbol{a})\big)\boldsymbol{h}+g'(\boldsymbol{b})s(\boldsymbol{h})+t(\boldsymbol{k}).$$

因此

$$\frac{g(f(\boldsymbol{a}+\boldsymbol{h}))-g(f(\boldsymbol{a}))-\big(g'(\boldsymbol{b})\circ f'(\boldsymbol{a})\big)\boldsymbol{h}}{|\boldsymbol{h}|}=\frac{g'(\boldsymbol{b})s(\boldsymbol{h})}{|\boldsymbol{h}|}+\frac{t(\boldsymbol{k})}{|\boldsymbol{h}|}.$$

一方面，由 $\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}}\dfrac{s(\boldsymbol{h})}{|\boldsymbol{h}|}=\boldsymbol{0}$ 及线性映射的连续性（第十二章例 3.3）知

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{g'(\boldsymbol{b})s(\boldsymbol{h})}{|\boldsymbol{h}|}=\boldsymbol{0}.$$

另一方面，

$$\frac{|\boldsymbol{k}|}{|\boldsymbol{h}|}=\frac{|f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})|}{|\boldsymbol{h}|}=\frac{|f'(\boldsymbol{a})\boldsymbol{h}+s(\boldsymbol{h})|}{|\boldsymbol{h}|}\leqslant\frac{|f'(\boldsymbol{a})\boldsymbol{h}|}{|\boldsymbol{h}|}+\frac{|s(\boldsymbol{h})|}{|\boldsymbol{h}|},$$

而由 (12.1) 知 $|f'(\boldsymbol{a})\boldsymbol{h}|=O(|\boldsymbol{h}|)$，因此 $\dfrac{|\boldsymbol{k}|}{|\boldsymbol{h}|}$ 在 $\boldsymbol{h}\to\boldsymbol{0}$ 时有界，于是

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{t(\boldsymbol{k})}{|\boldsymbol{h}|}=\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{t(\boldsymbol{k})}{|\boldsymbol{k}|}\cdot\frac{|\boldsymbol{k}|}{|\boldsymbol{h}|}=\boldsymbol{0}.$$

综上便知

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{g(f(\boldsymbol{a}+\boldsymbol{h}))-g(f(\boldsymbol{a}))-\big(g'(\boldsymbol{b})\circ f'(\boldsymbol{a})\big)\boldsymbol{h}}{|\boldsymbol{h}|}=\boldsymbol{0},$$

于是由微分的定义可得 $(g\circ f)'(\boldsymbol{a})=g'(\boldsymbol{b})\circ f'(\boldsymbol{a})$。 □

**【例 1.10】** 设 $E \subseteq \mathbb{R}^n$，$f : E \longrightarrow \mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的内点。又设 $A$ 是一个 $\ell \times m$ 矩阵，并记 $g(\boldsymbol{x}) = A\boldsymbol{x}$。那么由例 1.5 知 $g'(\boldsymbol{x}) = A$。于是当 $f$ 在 $\boldsymbol{a}$ 处可微时，由链式法则知

$$(Af)'(\boldsymbol{a}) = (g \circ f)'(\boldsymbol{a}) = Af'(\boldsymbol{a}).$$

## 习 题 13.1

1. 利用定义证明函数 $f(x, y) = |x + y|$ 在 $(0, 0)$ 处不可微。
2. 利用定义证明函数 $f(x, y) = |xy|$ 在 $(0, 0)$ 处可微。
3. 设 $A$ 是一个 $n \times n$ 矩阵。对任意的 $\boldsymbol{x} \in \mathbb{R}^n$，令 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} A\boldsymbol{x}$。证明 $f$ 在 $\mathbb{R}^n$ 上可微并求 $f'(\boldsymbol{x})$。
4. 设 $E \subseteq \mathbb{R}^n$，$f : E \longrightarrow \mathbb{R}^m$ 与 $g : E \longrightarrow \mathbb{R}^m$ 均可微，证明 $h(\boldsymbol{x}) = \langle f(\boldsymbol{x}), g(\boldsymbol{x})\rangle$ 也在 $E$ 上可微且对任意的 $\boldsymbol{x} \in E$ 有

$$h'(\boldsymbol{x}) = f(\boldsymbol{x})^{\mathrm{T}} g'(\boldsymbol{x}) + g(\boldsymbol{x})^{\mathrm{T}} f'(\boldsymbol{x}),$$

   这里 $\langle\,,\,\rangle$ 表示向量的内积。
5. 设 $f : \mathbb{R} \longrightarrow \mathbb{R}^m$ 可微，且对任意的 $x \in \mathbb{R}$ 有 $|f(x)| = 1$。证明：对任意的 $x \in \mathbb{R}$ 有 $\langle f(x), f'(x)\rangle = 0$。
6. 设 $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 连续可微，且对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ 而言 $f'(\boldsymbol{x})$ 均非奇异。又设存在 $\boldsymbol{b} \notin f(\mathbb{R}^n)$，并记 $\varphi(\boldsymbol{x}) = |f(\boldsymbol{x}) - \boldsymbol{b}|^2$，证明：$\varphi'(\boldsymbol{x}) \neq \boldsymbol{0}$ $(\forall\, \boldsymbol{x} \in \mathbb{R}^n)$。

## § 13.2 方向导数与偏导数

在上节中我们给出了可微映射的概念及基本性质，然而却有两个非常实际的问题摆在我们面前：一是到目前为止我们仍然没有一个简便的方法去判定一个映射是否可微；二是哪怕由上节命题 1.8 和 1.9 容易得知形如

$$f(x, y) = \log(\cos x + y^2 + 2)$$

的二元函数是可微的，但却无法方便地求出其微分，或是求出其微分所对应的矩阵。本节的目的就是来解决这些问题。

按照 Heine 归结原理，(13.1) 式左侧的极限存在当且仅当对于趋于 $\mathbf{0}$ 但各项不为 $\mathbf{0}$ 的任意序列 $\{\boldsymbol{x}_m\}$，极限

$$\lim_{m\to\infty}\frac{f(\boldsymbol{a}+\boldsymbol{x}_m)-f(\boldsymbol{a})-L\boldsymbol{x}_m}{|\boldsymbol{x}_m|}$$

均存在，这意味着对于任意一种令 $\boldsymbol{h}$ 趋于 $\mathbf{0}$ 的方式而言，(13.1) 左侧的极限均存在。因此令人想到能否通过选择一些特殊的 $\boldsymbol{h}$ 趋于 $\mathbf{0}$ 的方式，使得由这些特殊方式下某些极限的存在性推出 (13.1) 左侧极限的存在性，进而得到 $f$ 的可微性。当然，在所有的趋于 $\mathbf{0}$ 的方式中，沿某给定向量方向趋于 $\mathbf{0}$ 是最简单的情形了。再结合一元函数导数的原始定义，这就自然地引出了下述方向导数的概念。

**【定义 2.1】设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的一个内点。对于 $\mathbb{R}^n$ 中给定的非零向量 $\boldsymbol{u}$，若极限**

$$\lim_{t\to0}\frac{f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})}{t}$$

**存在，我们就称 $f$ 在 $\boldsymbol{a}$ 处<u>沿方向 $\boldsymbol{u}$ 是可微的</u>，并将上述极限称为 $f$ 在 $\boldsymbol{a}$ 处<u>沿方向 $\boldsymbol{u}$ 的方向导数 (directional derivative)</u>，记作 $\dfrac{\partial f}{\partial\boldsymbol{u}}(\boldsymbol{a})$ 或 $D_{\boldsymbol{u}}f(\boldsymbol{a})$。**

**【例 2.2】**对于二元函数 $f(x,y)=x\cos y$ 而言，它在点 $\mathbf{0}=(0,0)^{\mathrm{T}}$ 处沿方向 $\boldsymbol{u}=(1,0)^{\mathrm{T}}$ 的方向导数为

$$\frac{\partial f}{\partial\boldsymbol{u}}(\mathbf{0})=\lim_{t\to0}\frac{f(t,0)-f(0,0)}{t}=\lim_{t\to0}\frac{t}{t}=1.$$

$f$ 在 $\boldsymbol{a}=(1,0)^{\mathrm{T}}$ 处沿方向 $\boldsymbol{v}=(1,1)^{\mathrm{T}}$ 的方向导数为

$$\begin{aligned}\frac{\partial f}{\partial\boldsymbol{v}}(\boldsymbol{a})&=\lim_{t\to0}\frac{f(1+t,t)-f(1,0)}{t}=\lim_{t\to0}\frac{(1+t)\cos t-1}{t}\\&=\lim_{t\to0}\left(\cos t+\frac{\cos t-1}{t}\right)=1.\end{aligned}$$

**【命题 2.3】设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的一个内点。若 $f$ 在 $\boldsymbol{a}$ 处可微，则 $f$ 在 $\boldsymbol{a}$ 处的所有方向导数均存在，并且对于 $\mathbb{R}^n$ 中的任一非零向量 $\boldsymbol{u}$ 有**

$$\frac{\partial f}{\partial\boldsymbol{u}}(\boldsymbol{a})=f'(\boldsymbol{a})\boldsymbol{u}.$$

<u>证明</u>. 由 $f$ 在 $\boldsymbol{a}$ 处可微知

$$\lim_{\boldsymbol{h}\to\mathbf{0}}\frac{f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})-f'(\boldsymbol{a})\boldsymbol{h}}{|\boldsymbol{h}|}=\mathbf{0},$$

特别地有

$$\lim_{t\to 0}\frac{f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})-f'(\boldsymbol{a})(t\boldsymbol{u})}{|t\boldsymbol{u}|}=\boldsymbol{0}.$$

因此 $\lim\limits_{t\to 0}\dfrac{|f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})-f'(\boldsymbol{a})(t\boldsymbol{u})|}{|t\boldsymbol{u}|}=0$。两边同时乘以 $|\boldsymbol{u}|$ 可得

$$\lim_{t\to 0}\left|\frac{f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})-f'(\boldsymbol{a})(t\boldsymbol{u})}{t}\right|=0,$$

即 $\lim\limits_{t\to 0}\dfrac{f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})-f'(\boldsymbol{a})(t\boldsymbol{u})}{t}=\boldsymbol{0}$，故 $\lim\limits_{t\to 0}\dfrac{f(\boldsymbol{a}+t\boldsymbol{u})-f(\boldsymbol{a})}{t}=f'(\boldsymbol{a})\boldsymbol{u}$。 □

容易预见，上述命题的逆并不成立。事实上，$f$ 在 $\boldsymbol{a}$ 处的所有方向导数均存在甚至不能保证 $f$ 在 $\boldsymbol{a}$ 处连续。

**【例 2.4】**设

$$f(x,y)=\begin{cases}\dfrac{x^3y}{x^6+y^2}, & \text{若 } (x,y)\neq(0,0),\\ 0 & \text{若 } (x,y)=(0,0).\end{cases}$$

那么对任意的非零向量 $\boldsymbol{u}=(u_1,u_2)^{\mathrm{T}}$，$f$ 在 $(0,0)^{\mathrm{T}}$ 处沿方向 $\boldsymbol{u}$ 的方向导数为

$$\lim_{t\to 0}\frac{f(tu_1,tu_2)-f(0,0)}{t}=\lim_{t\to 0}\frac{1}{t}\cdot\frac{(tu_1)^3(tu_2)}{(tu_1)^6+(tu_2)^2}=\lim_{t\to 0}\frac{tu_1^2u_2}{t^4u_1^6+u_2^2}=0$$

（上面最后一步可依 $u_2$ 是否为 0 讨论得到），也即是说 $f$ 在 $(0,0)^{\mathrm{T}}$ 处的所有方向导数均存在且相等。但由

$$f(x,x^3)=\frac{1}{2},\qquad \forall\, x\neq 0$$

知 $f$ 在 $(0,0)^{\mathrm{T}}$ 处不连续。

从形式上看，命题 2.3 是在可微的前提下给出了求方向导数的一个方法，但换一个角度来思考，假设我们暂时只知道 $f$ 在 $\boldsymbol{a}$ 处可微却不知道 $f'(\boldsymbol{a})$ 是什么，那么，如果我们能够对 $\mathbb{R}^n$ 的一组基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n\}$ 计算出 $f$ 在 $\boldsymbol{a}$ 处沿诸方向 $\boldsymbol{u}_j$ 的方向导数，就能通过命题 2.3 得到 $f'(\boldsymbol{a})$ 在这组基下的矩阵了。按照这个思路，我们立刻意识到诸方向导数中尤为重要的是沿坐标向量 $\boldsymbol{e}_j$ 方向的方向导数。

**【定义 2.5】**设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的内点。若 $f$ 在 $\boldsymbol{a}$ 处沿 $\boldsymbol{e}_j$ 方向的方向导数存在，则称该方向导数为 $f$ 在 $\boldsymbol{a}$ 处的第 $j$ 个<u>偏导数 (partial derivative)</u>[③]，记作 $\dfrac{\partial f}{\partial x_j}(\boldsymbol{a})$ 或 $f_{x_j}(\boldsymbol{a})$。

---

③在这里我们沿用了名词“偏导数”，但必须注意的是，在一般情况下方向导数与偏导数均未必是一个“数”。

对于二元函数 $f$，由于通常将其自变量记作 $x$ 和 $y$，所以我们常把它的两个偏导数记作 $\dfrac{\partial f}{\partial x}$ 与 $\dfrac{\partial f}{\partial y}$，或 $f_x$ 与 $f_y$。对于三元函数 $f$，由于通常将其自变量记作 $x$，$y$ 和 $z$，故而常把它的三个偏导数记作 $\dfrac{\partial f}{\partial x}$，$\dfrac{\partial f}{\partial y}$ 与 $\dfrac{\partial f}{\partial z}$，或 $f_x$，$f_y$ 与 $f_z$。

设 $E \subseteq \mathbb{R}^n$，$\boldsymbol{a}$ 是 $E$ 的内点。对于多元函数 $f: E \longrightarrow \mathbb{R}$ 而言，若记 $\boldsymbol{a} = (a_1, \cdots, a_n)^{\mathrm{T}}$，那么

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = \lim_{t \to 0} \frac{f(a_1, \cdots, a_{j-1}, a_j + t, a_{j+1}, \cdots, a_n) - f(a_1, \cdots, a_n)}{t}.$$

换句话说，若记 $g(\lambda) = f(a_1, \cdots, a_{j-1}, \lambda, a_{j+1}, \cdots, a_n)$，则 $\dfrac{\partial f}{\partial x_j}(\boldsymbol{a}) = g'(a_j)$。所以偏导数 $\dfrac{\partial f}{\partial x_j}$ 可通过将诸 $x_i$ $(i \neq j)$ 看作常数而对变量 $x_j$ 求导得到。

**【例 2.6】**对 $f(x,y,z) = \log(\sin x + y + e^z)$ 有

$$f_x = \frac{\partial f}{\partial x} = \frac{\cos x}{\sin x + y + e^z},$$

$$f_y = \frac{\partial f}{\partial y} = \frac{1}{\sin x + y + e^z},$$

$$f_z = \frac{\partial f}{\partial z} = \frac{e^z}{\sin x + y + e^z}.$$

因此若记 $\boldsymbol{a} = (0,1,2)^{\mathrm{T}}$，则

$$\frac{\partial f}{\partial x}(\boldsymbol{a}) = \frac{1}{1+e^2}, \qquad \frac{\partial f}{\partial y}(\boldsymbol{a}) = \frac{1}{1+e^2}, \qquad \frac{\partial f}{\partial z}(\boldsymbol{a}) = \frac{e^2}{1+e^2}.$$

由于 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 是 $\mathbb{R}^n$ 的基，故方向导数可用偏导数表示出来。事实上，若 $f$ 在 $\boldsymbol{a} \in \mathbb{R}^n$ 处可微，则对任意的非零向量 $\boldsymbol{u} = (u_1, \cdots, u_n)^{\mathrm{T}}$ 而言，由命题 2.3 可得

$$\begin{aligned}\frac{\partial f}{\partial \boldsymbol{u}}(\boldsymbol{a}) = f'(\boldsymbol{a})\boldsymbol{u} = f'(\boldsymbol{a})\left(\sum_{j=1}^n u_j \boldsymbol{e}_j\right) &= \sum_{j=1}^n u_j f'(\boldsymbol{a})\boldsymbol{e}_j \\ &= \sum_{j=1}^n u_j \frac{\partial f}{\partial x_j}(\boldsymbol{a}). \end{aligned} \tag{13.4}$$

进一步地，若 $f = (f_1, \cdots, f_m)^{\mathrm{T}}$，那么由命题 1.7 知

$$f'(\boldsymbol{a}) = \begin{bmatrix} f_1'(\boldsymbol{a}) \\ \vdots \\ f_m'(\boldsymbol{a}) \end{bmatrix}. \tag{13.5}$$

然而对 $n$ 元函数 $f_i$ $(1 \leqslant i \leqslant m)$ 应用 (13.4) 可得

$$f_i'(\boldsymbol{a})\boldsymbol{u} = \sum_{j=1}^{n} u_j \frac{\partial f_i}{\partial x_j}(\boldsymbol{a}) = \left(\frac{\partial f_i}{\partial x_1}(\boldsymbol{a}), \cdots, \frac{\partial f_i}{\partial x_n}(\boldsymbol{a})\right)\boldsymbol{u},$$

因此

$$f_i'(\boldsymbol{a}) = \left(\frac{\partial f_i}{\partial x_1}(\boldsymbol{a}), \cdots, \frac{\partial f_i}{\partial x_n}(\boldsymbol{a})\right).$$

将上式代入 (13.5)，我们最终得到

$$f'(\boldsymbol{a}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \dfrac{\partial f_1}{\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\boldsymbol{a}) \\ \dfrac{\partial f_2}{\partial x_1}(\boldsymbol{a}) & \dfrac{\partial f_2}{\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\boldsymbol{a}) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \dfrac{\partial f_m}{\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\boldsymbol{a}) \end{bmatrix}. \tag{13.6}$$

这一矩阵被称为 $f$ 在 $\boldsymbol{a}$ 处的 Jacobi 矩阵 (Jacobian matrix)。当 $m = n$ 时，我们把上述矩阵的行列式记作 $\dfrac{D(f_1, \cdots, f_n)}{D(x_1, \cdots, x_n)}$，并称之为 Jacobi 行列式。

$n$ 元函数 $f$ 在 $\boldsymbol{a}$ 处的 Jacobi 矩阵也被称作 $f$ 在 $\boldsymbol{a}$ 处的梯度 (gradient)，记作 $\operatorname{grad} f(\boldsymbol{a})$。按照 (13.4)，对任意的非零向量 $\boldsymbol{u}$ 有

$$\frac{\partial f}{\partial \boldsymbol{u}}(\boldsymbol{a}) = \langle \operatorname{grad} f(\boldsymbol{a}), \boldsymbol{u} \rangle.$$

如果 $\operatorname{grad} f(\boldsymbol{a}) \neq \boldsymbol{0}$，那么当 $\boldsymbol{u}$ 是 $\operatorname{grad} f(\boldsymbol{a})$ 的正数倍（也即 $\boldsymbol{u}$ 与 $\operatorname{grad} f(\boldsymbol{a})$ 同向）时方向导数 $\dfrac{\partial f}{\partial \boldsymbol{u}}(\boldsymbol{a})$ 达到最大值，因此可以认为梯度方向是 $f$ 在 $\boldsymbol{a}$ 处变化率最大的方向。

如果 $f(x_1, \cdots, x_m)$ 是一个 $m$ 元函数，而每个 $x_j$ 均是 $n$ 元函数 $x_j(t_1, \cdots, t_n)$，那么我们也可以把 $f$ 看作变量 $t_1, \cdots, t_n$ 的函数，于是由链式法则及 (13.6) 知

$$\left(\frac{\partial f}{\partial t_1}, \cdots, \frac{\partial f}{\partial t_n}\right) = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_m}\right) \begin{bmatrix} \dfrac{\partial x_1}{\partial t_1} & \cdots & \dfrac{\partial x_1}{\partial t_j} & \cdots & \dfrac{\partial x_1}{\partial t_n} \\ \dfrac{\partial x_2}{\partial t_1} & \cdots & \dfrac{\partial x_2}{\partial t_j} & \cdots & \dfrac{\partial x_2}{\partial t_n} \\ \vdots & & \vdots & & \vdots \\ \dfrac{\partial x_m}{\partial t_1} & \cdots & \dfrac{\partial x_m}{\partial t_j} & \cdots & \dfrac{\partial x_m}{\partial t_n} \end{bmatrix},$$

因此对 $1 \leqslant j \leqslant n$ 有

$$\frac{\partial f}{\partial t_j} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \cdot \frac{\partial x_i}{\partial t_j}. \tag{13.7}$$

这一公式也被称作偏导数的链式法则。

**【例 2.7】** 设 $f(x,y) = e^x + y$，且 $x = s\cos t$，$y = 2s + \log t$，那么

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} = e^x \cos t + 2 = e^{s\cos t}\cos t + 2$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} = e^x \cdot (-s\sin t) + \frac{1}{t} = -e^{s\cos t} s\sin t + \frac{1}{t}.$$

通过以上讨论，我们发现可按下述步骤判定 $f$ 是否在点 $\boldsymbol{a}$ 处可微:

(1) 判断 $f$ 在 $\boldsymbol{a}$ 处的各偏导数是否存在。如果不全都存在，则 $f$ 在 $\boldsymbol{a}$ 处不可微；如果全部存在，则继续下一步骤。

(2) 设 $f$ 在 $\boldsymbol{a}$ 处的 Jacobi 矩阵为 $L$，将 $L$ 代入 (13.1) 看看 (13.1) 是否成立。如果成立，那么 $f$ 在 $\boldsymbol{a}$ 处可微，否则 $f$ 在 $\boldsymbol{a}$ 处不可微。

**【例 2.8】** 判断函数

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

在 $(0,0)^{\mathrm{T}}$ 处是否可微。

解. 由偏导数的定义知

$$f_x(0,0) = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = 0,$$

同理有 $f_y(0,0) = 0$，从而 $f$ 在 $(0,0)^{\mathrm{T}}$ 处的 Jacobi 矩阵 $L$ 为零矩阵。因为

$$\frac{f(x,y)-f(0,0)-L(x,y)^{\mathrm{T}}}{\sqrt{x^2+y^2}} = \frac{xy}{x^2+y^2},$$

故极限 $\lim\limits_{\substack{x\to 0\\ y\to 0}} \dfrac{f(x,y)-f(0,0)-L(x,y)^{\mathrm{T}}}{\sqrt{x^2+y^2}}$ 不存在，所以 $f$ 在 $(0,0)^{\mathrm{T}}$ 处不可微。 $\square$

虽然在上面我们给出了判定 $f$ 在某点处可微性的具体步骤，但这种利用定义直接验证的方法在实际操作中或许会存在困难，特别是以下两种情况均会增大我们的工作量：一是当我们需要讨论 $f$ 在整个集合上的可微性而非在一点处的可微性时；二是定义域空间和像空间的维数都很大的情况。所以有必要去建立一个较为方便的

判定可微性的方法。在介绍这一判定方法之前，作为证明的一个辅助工具，我们先给出 Lagrange 中值定理在高维情形下的推广。当然，这一推广本身也是极有价值的一个结论。

首先回忆起在第十二章例 3.10 中所定义的集合

$$[\boldsymbol{a},\boldsymbol{b}]=\{(1-\lambda)\boldsymbol{a}+\lambda\boldsymbol{b}:0\leqslant\lambda\leqslant 1\}.$$

**【定理 2.9】（中值定理）** 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$\boldsymbol{a}$ 与 $\boldsymbol{b}$ 是 $E$ 中两个不同的点，且 $[\boldsymbol{a},\boldsymbol{b}]\subseteq E$。又设 $f:E\longrightarrow\mathbb{R}$ 在 $E$ 中每一点处沿方向 $\boldsymbol{u}=\boldsymbol{b}-\boldsymbol{a}$ 的方向导数均存在，那么存在 $\boldsymbol{c}\in[\boldsymbol{a},\boldsymbol{b}]$ 使得

$$f(\boldsymbol{b})-f(\boldsymbol{a})=\frac{\partial f}{\partial\boldsymbol{u}}(\boldsymbol{c}).\tag{13.8}$$

证明. 令 $g(t)=f\big((1-t)\boldsymbol{a}+t\boldsymbol{b}\big)$，则 $g:[0,1]\longrightarrow\mathbb{R}$ 满足 $g(0)=f(\boldsymbol{a})$ 以及 $g(1)=f(\boldsymbol{b})$。此外

$$\begin{aligned}g'(t)&=\lim_{h\to0}\frac{g(t+h)-g(t)}{h}=\lim_{h\to0}\frac{f\big([1-(t+h)]\boldsymbol{a}+(t+h)\boldsymbol{b}\big)-f\big((1-t)\boldsymbol{a}+t\boldsymbol{b}\big)}{h}\\&=\lim_{h\to0}\frac{f\big((1-t)\boldsymbol{a}+t\boldsymbol{b}+h\boldsymbol{u}\big)-f\big((1-t)\boldsymbol{a}+t\boldsymbol{b}\big)}{h}=\frac{\partial f}{\partial\boldsymbol{u}}\big((1-t)\boldsymbol{a}+t\boldsymbol{b}\big).\end{aligned}$$

因此对 $g$ 应用 Lagrange 中值定理知，存在 $\xi\in[0,1]$ 使得

$$f(\boldsymbol{b})-f(\boldsymbol{a})=g(1)-g(0)=g'(\xi)=\frac{\partial f}{\partial\boldsymbol{u}}\big((1-\xi)\boldsymbol{a}+\xi\boldsymbol{b}\big).$$

于是记 $\boldsymbol{c}=(1-\xi)\boldsymbol{a}+\xi\boldsymbol{b}$ 即得 (13.8)。 □

这一定理的形式似乎与 Lagrange 中值定理相差甚远，那么下面的推论将使我们看得更明显一些。

**【推论 2.10】** 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f:E\longrightarrow\mathbb{R}$ 在 $E$ 上可微，则对 $E$ 中任意两个不同的点 $\boldsymbol{a}$, $\boldsymbol{b}$，存在 $\boldsymbol{c}\in[\boldsymbol{a},\boldsymbol{b}]$ 使得

$$f(\boldsymbol{b})-f(\boldsymbol{a})=f'(\boldsymbol{c})(\boldsymbol{b}-\boldsymbol{a}).$$

证明. 这是因为此时由命题 2.3 知

$$\frac{\partial f}{\partial\boldsymbol{u}}(\boldsymbol{c})=f'(\boldsymbol{c})\boldsymbol{u}=f'(\boldsymbol{c})(\boldsymbol{b}-\boldsymbol{a}).$$

□

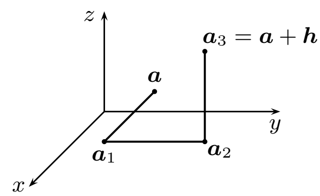
需注意的是，上述结论对一般的映射 $f:E\longrightarrow\mathbb{R}^m$ 未必成立，参见习题 16。

【定理 2.11】设 $E \subseteq \mathbb{R}^n$，$f = (f_1, \cdots, f_m)^{\mathrm{T}} : E \longrightarrow \mathbb{R}^m$，且 $\boldsymbol{a}$ 是 $E$ 的内点。如果存在 $\boldsymbol{a}$ 的邻域 $U$，使得 $f$ 的各分量函数在 $U$ 内有连续的偏导数 $\dfrac{\partial f_i}{\partial x_j}$ $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$，那么 $f$ 在 $\boldsymbol{a}$ 处可微。

证明. 按照命题 1.7，只需证明 $m = 1$ 的情形即可，此时 $f : E \longrightarrow \mathbb{R}$。我们记 $\boldsymbol{h} = (h_1, \cdots, h_n)^{\mathrm{T}}$，$\boldsymbol{a}_0 = \boldsymbol{a}$ 以及

$$\boldsymbol{a}_j = \boldsymbol{a}_{j-1} + h_j \boldsymbol{e}_j, \qquad 1 \leqslant j \leqslant n,$$

那么 $\boldsymbol{a}_n = \boldsymbol{a} + \boldsymbol{h}$（右图给出了 $n = 3$ 的情形）。




因为 $U$ 是 $\boldsymbol{a}$ 的邻域，故不妨设 $U$ 是以 $\boldsymbol{a}$ 为中心的开球。于是当 $|\boldsymbol{h}|$ 充分小时，对 $0 \leqslant j \leqslant n$ 均有 $\boldsymbol{a}_j \in U$。现在考虑 $f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})$，我们有

$$f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) = f(\boldsymbol{a}_n) - f(\boldsymbol{a}_0) = \sum_{j=1}^{n} \big[ f(\boldsymbol{a}_j) - f(\boldsymbol{a}_{j-1}) \big]. \tag{13.9}$$

注意到对任意的 $\lambda \in [0, 1]$，

$$|\lambda \boldsymbol{a}_{j-1} + (1-\lambda)\boldsymbol{a}_j - \boldsymbol{a}| \leqslant \lambda |\boldsymbol{a}_{j-1} - \boldsymbol{a}| + (1-\lambda)|\boldsymbol{a}_j - \boldsymbol{a}|,$$

故而 $[\boldsymbol{a}_{j-1}, \boldsymbol{a}_j] \subseteq U$，于是由定理 2.9 知存在 $\boldsymbol{c}_j \in [\boldsymbol{a}_{j-1}, \boldsymbol{a}_j]$ 使得

$$f(\boldsymbol{a}_j) - f(\boldsymbol{a}_{j-1}) = \frac{\partial f}{\partial (h_j \boldsymbol{e}_j)}(\boldsymbol{c}_j) = h_j \cdot \frac{\partial f}{\partial x_j}(\boldsymbol{c}_j).$$

把这一式子代入 (13.9) 即得

$$f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) = \sum_{j=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(\boldsymbol{c}_j).$$

因此若记

$$L = \left( \frac{\partial f}{\partial x_1}(\boldsymbol{a}), \cdots, \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \right),$$

那么

$$\frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - L\boldsymbol{h}}{|\boldsymbol{h}|} = \sum_{j=1}^{n} \frac{h_j}{|\boldsymbol{h}|} \cdot \left[ \frac{\partial f}{\partial x_j}(\boldsymbol{c}_j) - \frac{\partial f}{\partial x_j}(\boldsymbol{a}) \right].$$

现令 $\boldsymbol{h} \to \boldsymbol{0}$，那么 $\boldsymbol{c}_j \to \boldsymbol{a}$ $(1 \leqslant j \leqslant n)$，于是由偏导数的连续性知

$$\frac{\partial f}{\partial x_j}(\boldsymbol{c}_j) - \frac{\partial f}{\partial x_j}(\boldsymbol{a}) \to 0, \qquad 1 \leqslant j \leqslant n.$$

注意到 $\dfrac{h_j}{|\boldsymbol{h}|}$ 有界，故而

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})-L\boldsymbol{h}}{|\boldsymbol{h}|}=0.$$

这就证明了 $f$ 在 $\boldsymbol{a}$ 处可微。 □

必须指出的是，上述定理陈述的并不是可微的充要条件。事实上，存在可微的映射，它没有连续的偏导数（参见习题 18）。

**【定义 2.12】设 $E$ 是 $\mathbb{R}^n$ 中的开集，且 $f:E\longrightarrow\mathbb{R}^m$ 的每个分量函数在 $E$ 上的各个偏导数均存在且连续，那么就称 $f$ 在 $E$ 上连续可微，记作 $f\in C^1(E)$。**

细心的读者一定已经意识到了上述定义颇为突兀，这是因为在一元微分学里，我们称一个一元函数连续可导是指其导函数连续，因此在多元的情形下连续可微自然应该指的是使得导映射 $f':E\longrightarrow\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 连续的那些 $f$。事实上，在习题 25 中我们会看到这与定义 2.12 是等价的。

现设 $E$ 是 $\mathbb{R}^n$ 中的开集，并考虑 $E$ 上的 $n$ 元函数 $f:E\longrightarrow\mathbb{R}$。若 $f$ 的诸偏导 $\dfrac{\partial f}{\partial x_j}$ 在 $E$ 上存在，那么这些偏导也是定义在 $E$ 上的 $n$ 元函数，从而也可以考虑它们在 $E$ 上的偏导，例如

$$\frac{\partial}{\partial x_k}\Big(\frac{\partial f}{\partial x_j}\Big),$$

我们称之为 $f$ 在 $E$ 上的二阶混合偏导 (second order mixed partial derivative)，或简称为二阶偏导，并将上式记作 $\dfrac{\partial^2 f}{\partial x_k\partial x_j}$ 或 $f_{x_jx_k}$。类似可定义高阶偏导 (higher order partial derivative)。我们常将 $\dfrac{\partial^2 f}{\partial x_j\partial x_j}$ 简记作 $\dfrac{\partial^2 f}{\partial x_j^2}$，将 $\dfrac{\partial^3 f}{\partial x_j\partial x_j\partial x_j}$ 简记作 $\dfrac{\partial^3 f}{\partial x_j^3}$，等等。

如果 $n$ 元函数 $f$ 的每个 $\leqslant r$ 阶的偏导均在 $E$ 上存在且连续，我们就称 $f$ 在 $E$ 上是 $C^r$ 类 (class $C^r$) 的，记作 $f\in C^r(E)$。

**【例 2.13】** 设 $f(x,y)=x^2\log y$，则

$$\frac{\partial f}{\partial x}=2x\log y,\qquad \frac{\partial f}{\partial y}=\frac{x^2}{y}.$$

进而可得诸二阶偏导

$$\frac{\partial^2 f}{\partial x^2}=2\log y,\qquad \frac{\partial^2 f}{\partial y^2}=-\frac{x^2}{y^2},\qquad \frac{\partial^2 f}{\partial y\partial x}=\frac{\partial^2 f}{\partial x\partial y}=\frac{2x}{y}.$$

在上述例子中我们看到了 $\dfrac{\partial^2 f}{\partial y\partial x}=\dfrac{\partial^2 f}{\partial x\partial y}$，但这并不是总成立的，下面是一个反例。

**【例 2.14】**设

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

那么

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

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

于是

$$\frac{\partial f}{\partial y\partial x}(0,0)=\lim_{y\to 0}\frac{1}{y}\left(\frac{\partial f}{\partial x}(0,y)-\frac{\partial f}{\partial x}(0,0)\right)=\lim_{y\to 0}\frac{-y}{y}=-1,$$
$$\frac{\partial f}{\partial x\partial y}(0,0)=\lim_{x\to 0}\frac{1}{x}\left(\frac{\partial f}{\partial y}(x,0)-\frac{\partial f}{\partial y}(0,0)\right)=\lim_{x\to 0}\frac{x}{x}=1.$$

**【定理 2.15】(Schwarz)** 设 $E$ 是 $\mathbb{R}^n$ 中的开集，$f: E\longrightarrow\mathbb{R}$。若 $f\in C^2(E)$，则对任意的 $\boldsymbol{a}\in E$ 及 $1\leqslant i,\,j\leqslant n$ 有

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{a})=\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{a}). \tag{13.10}$$

证明. 因为在计算上式两侧的二阶偏导数时，可以将除 $x_i$ 和 $x_j$ 的其余变量视作常数，故不妨设 $n=2$，$i=1$，$j=2$。又记 $\boldsymbol{a}=(a_1,a_2)^{\mathrm{T}}$。下面我们用两种方法来计算

$$\Delta f=f(a_1+h,a_2+k)-f(a_1+h,a_2)-f(a_1,a_2+k)+f(a_1,a_2).$$

一方面，若记 $\varphi(x)=f(x,a_2+k)-f(x,a_2)$，那么

$$\Delta f=\varphi(a_1+h)-\varphi(a_1).$$

于是由 Lagrange 中值定理知存在位于 $a_1$ 和 $a_1+h$ 之间的某个 $\xi$ 使得

$$\Delta f=\varphi'(\xi)h=\left[\frac{\partial f}{\partial x_1}(\xi,a_2+k)-\frac{\partial f}{\partial x_1}(\xi,a_2)\right]h.$$

再次应用 Lagrange 中值定理知，存在位于 $a_2$ 和 $a_2+k$ 之间的某个 $\eta$ 使得

$$\Delta f=\frac{\partial^2 f}{\partial x_2\partial x_1}(\xi,\eta)\cdot hk.$$

另一方面，若记 $\psi(y)=f(a_1+h,y)-f(a_1,y)$，那么

$$\Delta f=\psi(a_2+k)-\psi(a_2).$$

同理，使用两次 Lagrange 中值定理可得，存在位于 $a_1$ 和 $a_1+h$ 之间的某个 $\zeta$ 以及位于 $a_2$ 和 $a_2+k$ 之间的某个 $\mu$ 使得

$$\Delta f=\frac{\partial^2 f}{\partial x_1\partial x_2}(\zeta,\mu)\cdot hk.$$

因此

$$\frac{\partial^2 f}{\partial x_2\partial x_1}(\xi,\eta)=\frac{\partial^2 f}{\partial x_1\partial x_2}(\zeta,\mu).$$

现令 $(h,k)^{\mathrm{T}}\to(0,0)^{\mathrm{T}}$，那么 $(\xi,\eta)^{\mathrm{T}}$ 与 $(\zeta,\mu)^{\mathrm{T}}$ 均趋于 $\boldsymbol{a}$，于是由二阶偏导的连续性知 (13.10) 成立。 □

如果 $f\in C^r(E)$，则可将 $f$ 的 $r$ 阶偏导数写成 $\dfrac{\partial^r f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ 的形式，其中 $\alpha_1+\cdots+\alpha_n=r$。

## 习 题 13.2

1. 对下列函数 $f$ 计算其在指定点 $\boldsymbol{a}$ 处沿指定方向 $\boldsymbol{u}$ 的方向导数:

   (1) $f(x,y)=x^2+y^2$，$\boldsymbol{a}=(1,1)^{\mathrm{T}}$，$\boldsymbol{u}=(1,-2)^{\mathrm{T}}$;

   (2) $f(x,y)=x+\log(y+1)$，$\boldsymbol{a}=(0,0)^{\mathrm{T}}$，$\boldsymbol{u}=(1,1)^{\mathrm{T}}$;

   (3) $f(x,y,z)=e^{x+y+z^2}$，$\boldsymbol{a}=(1,0,0)^{\mathrm{T}}$，$\boldsymbol{u}=(0,0,1)^{\mathrm{T}}$;

   (4) $f(x,y,z)=\dfrac{x+y^2}{z}$，$\boldsymbol{a}=(1,1,1)^{\mathrm{T}}$，$\boldsymbol{u}=(1,0,1)^{\mathrm{T}}$。

2. 对于下列函数 $f(x,y)$，试求所有的 $\boldsymbol{u}$ ($|\boldsymbol{u}|=1$)，使得 $f$ 在 $(0,0)^{\mathrm{T}}$ 点沿方向 $\boldsymbol{u}$ 的方向导数存在:

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

(2) $f(x,y)=\sqrt{|x^2-y^2|}$。

3. 证明函数 $f(x_1,\cdots,x_n)=|x_1|+\cdots+|x_n|$ 在 $\mathbf{0}$ 处沿任何方向的方向导数均不存在。

4. 对下列映射 $f$ 计算其 Jacobi 矩阵：

(1) $f(x,y)=x^2\log y+xy+2$；

(2) $f(x,y)=\dfrac{x-y}{x^2+y^2+1}$；

(3) $f(x,y)=(x+y,xy)$；

(4) $f(r,\theta)=(r\cos\theta,r\sin\theta)$；

(5) $f(x,y)=(y\sin x,x\cos y)$；

(6) $f(x,y,z)=(e^{x+z}+1,z^2\arctan y)$；

(7) $f(r,\theta,z)=(r\cos\theta,r\sin\theta,z)$；

(8) $f(r,\varphi,\theta)=(r\sin\varphi\cos\theta,r\sin\varphi\sin\theta,r\cos\varphi)$。

5. 对下列函数 $f$ 计算其各二阶偏导数：

(1) $f(x,y)=xy+\dfrac{y}{x}$；

(2) $f(x,y)=\arctan\dfrac{y}{x}$；

(3) $f(x,y)=\log(x^2+y^2)$；

(4) $f(x,y)=x^y$。

6. 设 $f$ 是 $C^2$ 类的，计算下列偏导数：

(1) $u=f(x^2+y^2)$，计算 $\dfrac{\partial u}{\partial x}$，$\dfrac{\partial^2 u}{\partial x^2}$ 及 $\dfrac{\partial^2 u}{\partial x\partial y}$。

(2) $u=f(e^x\log y)$，计算 $\dfrac{\partial u}{\partial y}$ 与 $\dfrac{\partial^2 u}{\partial x\partial y}$。

(3) $u=f\Big(x,\dfrac{x}{y}\Big)$，计算 $\dfrac{\partial^2 u}{\partial x^2}$ 与 $\dfrac{\partial^2 u}{\partial y^2}$。

7. 设 $f(x,y)=e^x\sin y$，对正整数 $m,n$ 计算 $\dfrac{\partial^{m+n}}{\partial x^m\partial y^n}$。

8. 设 $f(x,y,z)=\dfrac{1}{\sqrt{x^2+y^2+z^2}}$，证明在 $\mathbb{R}^3\setminus\{0\}$ 上有

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=0.$$

9. 设 $a,b$ 为常数，$a\neq 0$，令 $f(x,t)=\dfrac{1}{2a\sqrt{\pi t}}e^{-\frac{(x-b)^2}{4a^2t}}$，证明在 $\mathbb{R}\times\mathbb{R}_{>0}$ 上有

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

10. 设 $\varphi$ 与 $\psi$ 均二阶可导，$z=\varphi\Big(\dfrac{y}{x}\Big)+x\psi\Big(\dfrac{y}{x}\Big)$，证明

$$x^2\frac{\partial^2 z}{\partial x^2}+2xy\frac{\partial^2 z}{\partial x\partial y}+y^2\frac{\partial^2 z}{\partial y^2}=0.$$

11. （Euler）假设 $E \subseteq \mathbb{R}^n$，且对任意的 $\boldsymbol{x} \in E$ 及 $t > 0$ 有 $t\boldsymbol{x} \in E$。如果函数 $f : E \longrightarrow \mathbb{R}$ 满足

$$f(t\boldsymbol{x}) = t^\alpha f(\boldsymbol{x}), \qquad \forall\, \boldsymbol{x} \in E,\ t > 0,$$

其中 $\alpha$ 是一个给定的实数，则称 $f$ 是 $E$ 上的一个 <u>$\alpha$ 次齐次函数</u>。对 $E$ 上可微的 $\alpha$ 次齐次函数 $f$ 证明

$$\sum_{j=1}^{n} x_j \cdot \frac{\partial f}{\partial x_j}(\boldsymbol{x}) = \alpha f(\boldsymbol{x}), \qquad \forall\, \boldsymbol{x} = (x_1, \cdots, x_n) \in E. \tag{13.11}$$

12. 设 $E \subseteq \mathbb{R}^n$，且对任意的 $\boldsymbol{x} \in E$ 及 $t > 0$ 有 $t\boldsymbol{x} \in E$。又设 $f : E \longrightarrow \mathbb{R}$ 可微且满足 (13.11)。证明 $f$ 是 $E$ 上的 $\alpha$ 次齐次函数。

13. 设 $E \subseteq \mathbb{R}^n$，$f$ 与 $g$ 均是定义在 $E$ 上的可微函数，证明在 $E$ 上有

$$\operatorname{grad} fg = f \cdot \operatorname{grad} g + g \cdot \operatorname{grad} f.$$

14. 研究函数 $f(x, y) = \sqrt{|xy|}$ 在点 $(0, 0)^{\mathrm{T}}$ 处的可微性。

15. 研究函数

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

在点 $(0, 0)^{\mathrm{T}}$ 处的可微性。

16. 设 $g(x, y)$ 在 $(0, 0)^{\mathrm{T}}$ 的某邻域内连续，证明：$f(x, y) = |x - y| \cdot g(x, y)$ 在 $(0, 0)^{\mathrm{T}}$ 处可微的充要条件是 $g(0, 0) = 0$。

17. 设 $f : \mathbb{R} \longrightarrow \mathbb{R}^2$ 满足 $f(x) = (x, x^2)^{\mathrm{T}}$，证明 $f$ 在 $\mathbb{R}$ 上可微，但不存在 $\xi \in [0, 1]$ 使得 $f(1) - f(0) = f'(\xi)$。

18. 设

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

证明 $f$ 在 $(0, 0)^{\mathrm{T}}$ 处可微，但是 $\dfrac{\partial f}{\partial x}$ 与 $\dfrac{\partial f}{\partial y}$ 均在 $(0, 0)^{\mathrm{T}}$ 处不连续。

19. 设 $E \subseteq \mathbb{R}^n$，$f : E \longrightarrow \mathbb{R}^m$，$\boldsymbol{a}$ 是 $E$ 的一个内点。又设存在 $\boldsymbol{a}$ 的邻域 $U \subseteq E$，使得 $f$ 在 $U$ 上存在各偏导数，且这些偏导数在 $U$ 上有界。证明 $f$ 在 $\boldsymbol{a}$ 处连续。

20. 证明例 2.14 中的函数 $f(x,y)$ 在 $\mathbb{R}^2$ 上连续可微，因此这个例子说明连续可微且各二阶偏导存在的函数其两个二阶偏导 $\dfrac{\partial^2 f}{\partial x\partial y}$ 与 $\dfrac{\partial^2 f}{\partial y\partial x}$ 未必相等。

21. 假设在 $\mathbb{R}^2$ 上定义的函数 $f(x,y)$ 满足

(1) $\dfrac{\partial^2 f}{\partial y\partial x}$ 存在且在 $(x_0,y_0)^{\mathrm{T}}$ 处连续；

(2) 对任意的 $x$，$\dfrac{\partial f}{\partial y}(x,y_0)$ 均存在。

证明 $\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)$ 存在并且 $\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)=\dfrac{\partial^2 f}{\partial y\partial x}(x_0,y_0)$。

22. 设 $f$ 是定义在区域 $D$ 上的可微函数，且在 $D$ 的每一点处 $f$ 的各偏导均等于 0，证明 $f$ 是 $D$ 上的常值函数。

23. 设 $D$ 是 $\mathbb{R}^n$ 中的区域，$f: D\longrightarrow\mathbb{R}^m$ 在 $E$ 上可微。若 $f'$ 是常值映射，证明存在 $L\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 及 $\boldsymbol{a}\in\mathbb{R}^m$ 使得对任意的 $\boldsymbol{x}\in D$ 有 $f(\boldsymbol{x})=L\boldsymbol{x}+\boldsymbol{a}$。

24. 设 $L\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$，我们称

$$\|L\|=\sup_{|\boldsymbol{h}|=1}|L\boldsymbol{h}|$$

为 $L$ 的<u>范数</u>。证明：

(1) 对任意的 $\boldsymbol{x}\in\mathbb{R}^n$ 有 $|L\boldsymbol{x}|\leqslant\|L\|\cdot|\boldsymbol{x}|$；

(2) 对任意的 $\lambda\in\mathbb{R}$，$\|\lambda L\|=|\lambda|\cdot\|L\|$。特别地，$\|-L\|=\|L\|$；

(3) 对任意的 $L_1, L_2\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 有 $\|L_1+L_2\|\leqslant\|L_1\|+\|L_2\|$；

(4) 对任意的 $L_1\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 及 $L_2\in\mathscr{L}(\mathbb{R}^m,\mathbb{R}^\ell)$ 有

$$\|L_2\circ L_1\|\leqslant\|L_2\|\cdot\|L_1\|.$$

25. 设 $E$ 是 $\mathbb{R}^n$ 中的开集，$f: E\longrightarrow\mathbb{R}^m$ 在 $E$ 上可微。证明 $f$ 在 $E$ 上连续可微的充要条件是：对任意的 $\boldsymbol{a}\in E$ 及任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于满足 $|\boldsymbol{x}-\boldsymbol{a}|<\delta$ 的任意的 $\boldsymbol{x}\in E$ 均有

$$\|f'(\boldsymbol{x})-f'(\boldsymbol{a})\|<\varepsilon.$$

## §13.3 有限增量定理与 Taylor 公式

在上节中我们对多元函数得到了中值定理（定理 2.9 和推论 2.10），但同时也指出对于一般的可微映射没有类似结论。然而值得庆幸的是，在一定条件下我们可以用 $|\boldsymbol{b}-\boldsymbol{a}|$ 去控制 $|f(\boldsymbol{b})-f(\boldsymbol{a})|$，这在许多实际应用中已然足够了。为了描述这种控制，我们需要用到上节习题 24 中所给出的线性映射的范数，现将其定义重述如下。

**【定义 3.1】** 设 $L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$，定义 $L$ 的范数 (norm) $\|L\|$ 为

$$\|L\| = \sup_{|\boldsymbol{h}|=1} |L\boldsymbol{h}|.$$

范数的一些基本性质包含在了上节习题 24 中，特别需要提及的是，对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ 有

$$|L\boldsymbol{x}| \leqslant \|L\| \cdot |\boldsymbol{x}|.$$

**【定理 3.2】（有限增量定理）** 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f: E \longrightarrow \mathbb{R}^m$ 在 $E$ 上可微，且存在 $M > 0$ 使得对任意的 $\boldsymbol{x} \in E$ 均有 $\|f'(\boldsymbol{x})\| \leqslant M$。那么对任意的 $\boldsymbol{a}, \boldsymbol{b} \in E$ 有

$$|f(\boldsymbol{b})-f(\boldsymbol{a})| \leqslant M|\boldsymbol{b}-\boldsymbol{a}|.$$

证明. 记 $\boldsymbol{u} = f(\boldsymbol{b})-f(\boldsymbol{a})$，那么 $h(\boldsymbol{x}) = \langle \boldsymbol{u}, f(\boldsymbol{x})\rangle = \boldsymbol{u}^{\mathrm{T}} f(\boldsymbol{x})$ 是一个实值函数，并且由例 1.10 可得 $h'(\boldsymbol{x}) = \boldsymbol{u}^{\mathrm{T}} f'(\boldsymbol{x})$。于是由中值定理知，存在 $\boldsymbol{\xi} \in [\boldsymbol{a}, \boldsymbol{b}]$ 使得

$$\langle \boldsymbol{u}, f(\boldsymbol{b})-f(\boldsymbol{a})\rangle = h(\boldsymbol{b})-h(\boldsymbol{a}) = h'(\boldsymbol{\xi})(\boldsymbol{b}-\boldsymbol{a}) = \boldsymbol{u}^{\mathrm{T}} f'(\boldsymbol{x})(\boldsymbol{b}-\boldsymbol{a}).$$

因此由 Cauchy−Schwarz 不等式知

$$\begin{aligned}|\boldsymbol{u}|^2 = \langle \boldsymbol{u}, f(\boldsymbol{b})-f(\boldsymbol{a})\rangle = \langle \boldsymbol{u}, f'(\boldsymbol{x})(\boldsymbol{b}-\boldsymbol{a})\rangle &\leqslant |\boldsymbol{u}| \cdot |f'(\boldsymbol{x})(\boldsymbol{b}-\boldsymbol{a})| \\ &\leqslant |\boldsymbol{u}| \cdot M|\boldsymbol{b}-\boldsymbol{a}|.\end{aligned}$$

故而 $|f(\boldsymbol{b})-f(\boldsymbol{a})| = |\boldsymbol{u}| \leqslant M|\boldsymbol{b}-\boldsymbol{a}|$。 □

接下来我们把一元函数的 Taylor 公式推广到多元函数的情形。为方便起见，对非负整数 $n$ 元组 $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$ 记

$$|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n, \qquad \boldsymbol{\alpha}! = \alpha_1! \cdots \alpha_n!$$

以及
$$\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}=\frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}},$$
并对 $\boldsymbol{h}=(h_1,\cdots,h_n)^{\mathrm{T}}\in\mathbb{R}^n$ 记 $\boldsymbol{h}^{\boldsymbol{\alpha}}=h_1^{\alpha_1}\cdots h_n^{\alpha_n}$。

**【定理 3.3】（带 Lagrange 余项的 Taylor 公式）** 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f:E\longrightarrow\mathbb{R}$。如果 $f\in C^{r+1}(E)$，则对 $E$ 中任意两点 $\boldsymbol{a}$ 和 $\boldsymbol{a}+\boldsymbol{h}$，存在 $\theta\in(0,1)$ 使得
$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+\sum_{1\leqslant|\boldsymbol{\alpha}|\leqslant r}\frac{1}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a})\boldsymbol{h}^{\boldsymbol{\alpha}}+\sum_{|\boldsymbol{\alpha}|=r+1}\frac{1}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^{\boldsymbol{\alpha}}.\tag{13.12}$$

<u>证明</u>. 令 $g(t)=f(\boldsymbol{a}+t\boldsymbol{h})$，那么由链式法则知 $g$ 在 $[0,1]$ 上可导且
$$g'(t)=f'(\boldsymbol{a}+t\boldsymbol{h})\boldsymbol{h}=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(\boldsymbol{a}+t\boldsymbol{h})\cdot h_i.$$
同理
$$\begin{aligned}g''(t)&=\sum_{j=1}^n\frac{\partial}{\partial x_j}\left(\sum_{i=1}^n\frac{\partial f}{\partial x_i}(\boldsymbol{a}+t\boldsymbol{h})\cdot h_i\right)\cdot h_j\\&=\sum_{j=1}^n\sum_{i=1}^n\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{a}+t\boldsymbol{h})\cdot h_ih_j\end{aligned}$$
一般地，由于 $f\in C^{r+1}(E)$，故可归纳地证明 $g$ 在 $[0,1]$ 上 $r+1$ 阶可导，并且对 $k\leqslant r+1$ 有
$$g^{(k)}(t)=\sum_{j_1=1}^n\cdots\sum_{j_k=1}^n\frac{\partial^k f}{\partial x_{j_1}\cdots\partial x_{j_k}}(\boldsymbol{a}+t\boldsymbol{h})\cdot h_{j_1}\cdots h_{j_k}.\tag{13.13}$$
注意到由 $f\in C^{r+1}(E)$ 及定理 2.15 知 $\dfrac{\partial^k f}{\partial x_{j_1}\cdots\partial x_{j_k}}$ 必可写成如 $\dfrac{\partial^k f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ 的形式，其中 $\alpha_1+\cdots+\alpha_n=k$。而对给定的一组满足 $\alpha_1+\cdots+\alpha_n=k$ 的 $\alpha_\ell$ $(1\leqslant\ell\leqslant k)$，使得 $\dfrac{\partial^k f}{\partial x_{j_1}\cdots\partial x_{j_k}}=\dfrac{\partial^k f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ 的 $(j_1,\cdots,j_k)$ 的个数为
$$\binom{k}{\alpha_1}\binom{k-\alpha_1}{\alpha_2}\binom{k-\alpha_1-\alpha_2}{\alpha_3}\cdots\binom{k-\alpha_1-\cdots-\alpha_{n-1}}{\alpha_n}=\frac{k!}{\alpha_1!\cdots\alpha_n!},$$
因此对 (13.13) 右边合并同类项可得
$$g^{(k)}(t)=\sum_{|\boldsymbol{\alpha}|=k}\frac{k!}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+t\boldsymbol{h})\boldsymbol{h}^{\boldsymbol{\alpha}}.$$

现对 $g$ 使用带 Lagrange 余项的 Taylor 公式，我们有

$$g(1)=\sum_{k=0}^{r}\frac{g^{(k)}(0)}{k!}+\frac{g^{(r+1)}(\theta)}{(r+1)!},$$

其中 $\theta\in(0,1)$，此即

$$\begin{aligned}f(\boldsymbol{a}+\boldsymbol{h})&=f(\boldsymbol{a})+\sum_{k=1}^{r}\frac{1}{k!}\sum_{|\boldsymbol{\alpha}|=k}\frac{k!}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}}f}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a})\boldsymbol{h}^{\boldsymbol{\alpha}}\\&\quad+\frac{1}{(r+1)!}\sum_{|\boldsymbol{\alpha}|=r+1}\frac{(r+1)!}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}}f}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^{\boldsymbol{\alpha}}\\&=f(\boldsymbol{a})+\sum_{1\leqslant|\boldsymbol{\alpha}|\leqslant r}\frac{1}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}}f}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a})\boldsymbol{h}^{\boldsymbol{\alpha}}+\sum_{|\boldsymbol{\alpha}|=r+1}\frac{1}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}}f}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^{\boldsymbol{\alpha}},\end{aligned}$$

从而定理得证。 □

(13.12) 右边 Taylor 多项式的前三项是很重要的，它们是

$$\begin{aligned}&f(\boldsymbol{a})+\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(\boldsymbol{a})h_j+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{a})h_ih_j\\&=f(\boldsymbol{a})+f'(\boldsymbol{a})\boldsymbol{h}+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{a})h_ih_j.\end{aligned}\tag{13.14}$$

若记

$$H_f(\boldsymbol{a})=\begin{bmatrix}\dfrac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_1\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}(\boldsymbol{a})\\ \dfrac{\partial^2 f}{\partial x_2\partial x_1}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n}(\boldsymbol{a})\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial^2 f}{\partial x_n\partial x_1}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_n\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(\boldsymbol{a})\end{bmatrix}.$$

那么 (13.14) 也即

$$f(\boldsymbol{a})+f'(\boldsymbol{a})\boldsymbol{h}+\frac{1}{2}\boldsymbol{h}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{h}.$$

我们称 $H_f(\boldsymbol{a})$ 为 $f$ 在 $\boldsymbol{a}$ 处的 <u>Hesse 矩阵</u>。

利用定理 3.3 可得如下带 Peano 余项的 Taylor 公式。

**【定理 3.4】（带 Peano 余项的 Taylor 公式）** 假设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f: E \longrightarrow \mathbb{R}$。如果 $f \in C^r(E)$，则对任意的 $\boldsymbol{a} \in E$，当 $\boldsymbol{h} \to \boldsymbol{0}$ 时有

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{1 \leqslant |\boldsymbol{\alpha}| \leqslant r} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}) \boldsymbol{h}^{\boldsymbol{\alpha}} + o(|\boldsymbol{h}|^r).$$

证明. 由定理 3.3 知存在 $\theta \in (0,1)$ 使得

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{1 \leqslant |\boldsymbol{\alpha}| \leqslant r-1} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}) \boldsymbol{h}^{\boldsymbol{\alpha}} + \sum_{|\boldsymbol{\alpha}| = r} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+\theta\boldsymbol{h}) \boldsymbol{h}^{\boldsymbol{\alpha}}. \tag{13.15}$$

而由 $r$ 阶偏导的连续性知当 $\boldsymbol{h} \to \boldsymbol{0}$ 时

$$\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}+\theta\boldsymbol{h}) = \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}) + o(1).$$

将这代入 (13.15) 可得

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{1 \leqslant |\boldsymbol{\alpha}| \leqslant r} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a}) \boldsymbol{h}^{\boldsymbol{\alpha}} + o\Big(\sum_{|\boldsymbol{\alpha}|=r} |\boldsymbol{h}^{\boldsymbol{\alpha}}|\Big).$$

注意到当 $|\boldsymbol{\alpha}| = r$ 时

$$|\boldsymbol{h}^{\boldsymbol{\alpha}}| = |h_1^{\alpha_1} \cdots h_n^{\alpha_n}| \leqslant |\boldsymbol{h}|^{\alpha_1 + \cdots + \alpha_n} = |\boldsymbol{h}|^r,$$

从而立刻得到定理结论。 □

**【注 3.5】** 当 $\boldsymbol{a} = \boldsymbol{0}$ 时，上面两个定理中的 Taylor 公式也被称作 Maclaurin 公式。

**【例 3.6】** 设 $f(x,y) = \dfrac{1}{1-x-y}$，则由归纳法容易验证

$$\frac{\partial^{i+j} f}{\partial x^i \partial y^j} = \frac{(i+j)!}{(1-x-y)^{i+j+1}},$$

于是 $\dfrac{\partial^{i+j} f}{\partial x^i \partial y^j}(0,0) = (i+j)!$，所以当 $x \to 0$，$y \to 0$ 时有

$$\frac{1}{1-x-y} = \sum_{0 \leqslant i+j \leqslant n} \frac{(i+j)!}{i!\,j!} x^i y^j + o\big((x^2+y^2)^{\frac{n}{2}}\big).$$

**【注 3.7】**值得一提的是，如果在上例中把 $x+y$ 看作一个整体，那么由一元函数的 Taylor 公式知，当 $x\to 0$，$y\to 0$ 时有

$$\begin{aligned}\frac{1}{1-x-y}&=\sum_{k=0}^{n}(x+y)^k+o(|x+y|^n)\\&=\sum_{k=0}^{n}\sum_{i+j=k}\frac{k!}{i!\,j!}x^iy^j+o\big((x^2+y^2)^{\frac{n}{2}}\big)\\&=\sum_{0\leqslant i+j\leqslant n}\frac{(i+j)!}{i!\,j!}x^iy^j+o\big((x^2+y^2)^{\frac{n}{2}}\big).\end{aligned}$$

于是我们得到了与例 3.6 相同的结果。这并非巧合，而是 Taylor 展开式唯一性的自然体现（参见习题 3）。

## 习 题 13.3

1. 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f:E\longrightarrow E$ 可微，且存在 $\theta\in(0,1)$ 使得对任意的 $\boldsymbol{x}\in E$ 均有 $\|f'(\boldsymbol{x})\|\leqslant\theta$。证明：对任意的 $\boldsymbol{x}_0\in E$，由 $\boldsymbol{x}_{k+1}=f(\boldsymbol{x}_k)$ $(\forall\, k\geqslant 0)$ 所定义的序列 $\{\boldsymbol{x}_k\}$ 收敛。

2. 设 $\boldsymbol{a}\in\mathbb{R}^n$，$f$ 是从 $\boldsymbol{a}$ 的邻域 $U$ 到 $\mathbb{R}^m$ 的连续映射。对 $U$ 中任一异于 $\boldsymbol{a}$ 的点 $\boldsymbol{x}$，$f$ 均在 $\boldsymbol{x}$ 处可微。又设存在 $L\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$，使得对任意的 $\varepsilon>0$，存在 $\delta>0$ 满足
$$\|f'(\boldsymbol{x})-L\|<\varepsilon,\qquad \forall\,\boldsymbol{x}\in B(\boldsymbol{a},\delta)\setminus\{\boldsymbol{a}\}.\text{[4]}$$
证明 $f$ 在 $\boldsymbol{a}$ 处可微且 $f'(\boldsymbol{a})=L$。

3. （Taylor 展开式的唯一性）设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f$ 是定义在 $E$ 上的函数且 $f\in C^r(E)$。又设 $\boldsymbol{a}\in E$，且当 $\boldsymbol{h}\to\boldsymbol{0}$ 时有
$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+\sum_{1\leqslant|\boldsymbol{\alpha}|\leqslant r}A_{\boldsymbol{\alpha}}\boldsymbol{h}^{\boldsymbol{\alpha}}+o(|\boldsymbol{h}|^r).$$
证明对任意的 $\boldsymbol{\alpha}$ 有 $A_{\boldsymbol{\alpha}}=\dfrac{1}{\boldsymbol{\alpha}!}\dfrac{\partial^{\boldsymbol{\alpha}}f}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{a})$。

4. 设 $f(x,y)=g(x)h(y)$，其中 $g\in C^r((a,b))$，$h\in C^r((c,d))$。证明：对任意的 $\boldsymbol{x}_0\in(a,b)\times(c,d)$，$f$ 在 $\boldsymbol{x}_0$ 处直至第 $r$ 次的带 Peano 余项的 Taylor 公式可由 $g(x)$ 和 $h(y)$ 各自的带 Peano 余项的 Taylor 公式相乘得到。

---

[4] 这一条件也被记作 $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}}f'(\boldsymbol{x})=L$。

5. 写出下列函数带 Peano 余项的 Maclaurin 公式：

(1) $\sin(x^2+y^2)$；　　(2) $\sqrt{1+x+y}$；

(3) $e^x\sin y$；　　(4) $\log(1+x)\log(1+y)$。

6. 求下列函数 $f$ 在指定点 $\boldsymbol{a}$ 处的 Taylor 多项式的前三项：

(1) $f(x,y)=x^{\alpha}y^{\beta}$，$\boldsymbol{a}=(1,1)$；

(2) $f(x,y)=\dfrac{1}{1+\sin x-y}$，$\boldsymbol{a}=\left(\dfrac{\pi}{2},0\right)$；

(3) $f(x,y,z)=\cos(x+2y+3z)$，$\boldsymbol{a}=(1,-2,1)$；

(4) $f(x,y,z)=\dfrac{\log(y+z^2)}{1+x}$，$\boldsymbol{a}=(0,1,0)$。

7. 证明：当 $|x|$ 和 $|y|$ 充分小时有

$$\arctan\frac{1+x+y}{1-x+y}\approx\frac{\pi}{4}+x-xy.$$

8. 计算下列二重极限：

(1) $\displaystyle\lim_{\substack{x\to0\\y\to0}}\frac{(1+x)^y-1}{\sin xy}$；　　(2) $\displaystyle\lim_{\substack{x\to0^+\\y\to0^+}}\frac{e^x\sin y-y}{\sqrt{x^2+xy+y^2}}$。

## §13.4 反函数定理

在本节中我们将给出一元函数的反函数求导法则在 $\mathbb{R}^n$ 中的推广。

**【定理 4.1】**（反函数定理） 设 $E$ 是 $\mathbb{R}^n$ 中的开集，$f:E\longrightarrow\mathbb{R}^n$ 且 $f\in C^1(E)$。又设 $\boldsymbol{a}\in E$。若 $f'(\boldsymbol{a})$ 非奇异，那么必存在 $\boldsymbol{a}$ 的邻域 $U$ 使得 $V=f(U)$ 是 $\mathbb{R}^n$ 中的开集，且 $f|_U:U\longrightarrow V$ 是双射。此外，若用 $g$ 表示 $f|_U$ 的逆映射，则 $g\in C^1(V)$，并且对任意的 $\boldsymbol{y}\in V$ 有

$$g'(\boldsymbol{y})=f'(g(\boldsymbol{y}))^{-1}. \tag{13.16}$$

<u>证明</u>. 记 $f=(f_1,\cdots,f_n)^{\mathrm{T}}$。证明分以下五步进行。

(1) 先将问题转化为 $f'(\boldsymbol{a})$ 是恒等映射的情形。

记 $L=f'(\boldsymbol{a})$，由例 1.10 知 $(L^{-1}\circ f)'(\boldsymbol{a})=L^{-1}\circ f'(\boldsymbol{a})$ 是恒等映射，而如果命题对 $L^{-1}\circ f$ 成立，容易验证它也对 $f$ 成立（留作习题）。因此不妨假设 $f'(\boldsymbol{a})$ 是恒等映射。

(2) 证明 $f$ 是局部双射。

由于对任意的 $\boldsymbol{x}\in E$ 有

$$f'(\boldsymbol{x})=\begin{bmatrix}\dfrac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\boldsymbol{x})\\ \vdots & & \vdots\\ \dfrac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\boldsymbol{x})\end{bmatrix},$$

并且每个偏导 $\dfrac{\partial f_i}{\partial x_j}$ 均在 $E$ 上连续，故存在 $\boldsymbol{a}$ 的邻域 $U$，使得对任意的 $\boldsymbol{x}\in U$ 及 $1\leqslant i,\, j\leqslant n$ 有

$$\left|\frac{\partial f_i}{\partial x_j}(\boldsymbol{x})-\frac{\partial f_i}{\partial x_j}(\boldsymbol{a})\right|<\frac{1}{2n}.$$

因此当 $\boldsymbol{x}\in U$ 时，由 (12.1) 可得

$$\|f'(\boldsymbol{x})-I_n\|=\|f'(\boldsymbol{x})-f'(\boldsymbol{a})\|\leqslant\frac{1}{2n}\cdot n=\frac{1}{2}.$$

于是对 $f(\boldsymbol{x})-\boldsymbol{x}$ 应用有限增量定理（定理 3.2）可得

$$|(f(\boldsymbol{x}_1)-\boldsymbol{x}_1)-(f(\boldsymbol{x}_2)-\boldsymbol{x}_2)|\leqslant\frac{1}{2}|\boldsymbol{x}_1-\boldsymbol{x}_2|,\qquad \forall\, x_1,\, x_2\in U. \tag{13.17}$$

进而由三角形不等式得到

$$|f(\boldsymbol{x}_1)-f(\boldsymbol{x}_2)|\geqslant\frac{1}{2}|\boldsymbol{x}_1-\boldsymbol{x}_2|,\qquad \forall\, x_1,\, x_2\in U, \tag{13.18}$$

这说明 $f$ 在 $U$ 上是单射，从而 $f|_U: U\longrightarrow V=f(U)$ 是双射。

(3) 证明 $V$ 是 $\mathbb{R}^n$ 中的开集。

为此，我们需要对任意的 $\boldsymbol{y}_0\in V$ 证明存在 $r>0$ 使得 $B(\boldsymbol{y}_0,r)\subseteq V$。由 $\boldsymbol{y}_0\in V=f(U)$ 知存在 $\boldsymbol{x}_0\in U$ 使得 $\boldsymbol{y}_0=f(\boldsymbol{x}_0)$。由于 $U$ 是开集，故存在 $\varepsilon>0$ 使得 $B(\boldsymbol{x}_0,3\varepsilon)\subseteq U$，进而 $\overline{B(\boldsymbol{x}_0,2\varepsilon)}\subseteq U$。下面来证明 $B(\boldsymbol{y}_0,\varepsilon)\subseteq V$。对任意的 $\boldsymbol{y}\in B(\boldsymbol{y}_0,\varepsilon)$，我们考虑由

$$F(\boldsymbol{x})=\boldsymbol{y}+\boldsymbol{x}-f(\boldsymbol{x})$$

所定义的映射 $F$，那么由 (13.17) 知

$$|F(\boldsymbol{x}_1)-F(\boldsymbol{x}_2)|\leqslant\frac{1}{2}|\boldsymbol{x}_1-\boldsymbol{x}_2|,\qquad \forall\, \boldsymbol{x}_1,\, \boldsymbol{x}_2\in\overline{B(\boldsymbol{x}_0,2\varepsilon)} \tag{13.19}$$

特别地，

$$|F(\boldsymbol{x})-F(\boldsymbol{x}_0)|\leqslant\frac{1}{2}|\boldsymbol{x}-\boldsymbol{x}_0|\leqslant\varepsilon,\qquad \forall\, \boldsymbol{x}\in\overline{B(\boldsymbol{x}_0,2\varepsilon)}.$$

此外，

$$|F(\boldsymbol{x}_0)-\boldsymbol{x}_0|=|\boldsymbol{y}-\boldsymbol{y}_0|<\varepsilon.$$

于是当 $\boldsymbol{x}\in\overline{B(\boldsymbol{x}_0,2\varepsilon)}$ 时有

$$|F(\boldsymbol{x})-\boldsymbol{x}_0|\leqslant|F(\boldsymbol{x})-F(\boldsymbol{x}_0)|+|F(\boldsymbol{x}_0)-\boldsymbol{x}_0|<\varepsilon+\varepsilon=2\varepsilon,$$

也即 $F(\boldsymbol{x})\in B(\boldsymbol{x}_0,2\varepsilon)$，因此 $F\left(\overline{B(\boldsymbol{x}_0,2\varepsilon)}\right)\subseteq\overline{B(\boldsymbol{x}_0,2\varepsilon)}$。结合 (13.19)，便可由压缩映像原理（第十二章例 1.18）得知存在 $\boldsymbol{z}\in\overline{B(\boldsymbol{x}_0,2\varepsilon)}$ 使得 $F(\boldsymbol{z})=\boldsymbol{z}$，此即 $\boldsymbol{y}=f(\boldsymbol{z})$，从而 $\boldsymbol{y}\in f(U)=V$。这就证明了 $B(\boldsymbol{y}_0,\varepsilon)\subseteq V$。

(4) 接下来证明逆映射 $g$ 在 $V$ 上可微且 (13.16) 成立。

首先需要注意的是，由 $f'(\boldsymbol{a})$ 的非奇异性及 $f\in C^1(E)$ 知，可适当地选择 $U$，使得对任意的 $y\in V=f(U)$ 而言 $f'(g(\boldsymbol{y}))$ 均是非奇异的。现对任意的 $\boldsymbol{b}\in V$，设 $\boldsymbol{k}\neq\mathbf{0}$ 满足 $\boldsymbol{b}+\boldsymbol{k}\in V$，则由 $g$ 是双射知 $g(\boldsymbol{b}+\boldsymbol{k})\neq g(\boldsymbol{b})$。现记 $\boldsymbol{h}=g(\boldsymbol{b}+\boldsymbol{k})-g(\boldsymbol{b})$，那么

$$\boldsymbol{k}=(\boldsymbol{b}+\boldsymbol{k})-\boldsymbol{b}=f(g(\boldsymbol{b}+\boldsymbol{k}))-f(g(\boldsymbol{b}))=f(g(\boldsymbol{b})+\boldsymbol{h})-f(g(\boldsymbol{b})),$$

从而由 (13.18) 可得 $|\boldsymbol{k}|\geqslant\dfrac{1}{2}|\boldsymbol{h}|$，故当 $\boldsymbol{k}\to\mathbf{0}$ 时 $\boldsymbol{h}\to\mathbf{0}$ 并且 $\dfrac{|\boldsymbol{h}|}{|\boldsymbol{k}|}$ 有界。于是

$$\begin{aligned}\lim_{\boldsymbol{k}\to\mathbf{0}}\frac{g(\boldsymbol{b}+\boldsymbol{k})-g(\boldsymbol{b})-f'(g(\boldsymbol{b}))^{-1}\boldsymbol{k}}{|\boldsymbol{k}|}&=\lim_{\boldsymbol{k}\to\mathbf{0}}\frac{f'(g(\boldsymbol{b}))^{-1}\left[f'(g(\boldsymbol{b}))\boldsymbol{h}-\boldsymbol{k}\right]}{|\boldsymbol{k}|}\\&=\lim_{\boldsymbol{k}\to\mathbf{0}}-f'(g(\boldsymbol{b}))^{-1}\left(\frac{f(g(\boldsymbol{b})+\boldsymbol{h})-f(g(\boldsymbol{b}))-f'(g(\boldsymbol{b}))\boldsymbol{h}}{|\boldsymbol{h}|}\right)\cdot\frac{|\boldsymbol{h}|}{|\boldsymbol{k}|}=\mathbf{0}.\end{aligned}$$

因此 $g$ 在 $\boldsymbol{b}$ 处可微且 $g'(\boldsymbol{b})=f'(g(\boldsymbol{b}))^{-1}$。

(5) 最后证 $g\in C^1(V)$。

记 $g=(g_1,\cdots,g_n)^{\mathrm{T}}$，那么对任意的 $\boldsymbol{y}\in V$，由 (13.16) 知

$$\begin{bmatrix}\dfrac{\partial g_1}{\partial x_1}(\boldsymbol{y}) & \cdots & \dfrac{\partial g_1}{\partial x_n}(\boldsymbol{y})\\ \vdots & & \vdots\\ \dfrac{\partial g_n}{\partial x_1}(\boldsymbol{y}) & \cdots & \dfrac{\partial g_n}{\partial x_n}(\boldsymbol{y})\end{bmatrix}=\begin{bmatrix}\dfrac{\partial f_1}{\partial x_1}(g(\boldsymbol{y})) & \cdots & \dfrac{\partial f_1}{\partial x_n}(g(\boldsymbol{y}))\\ \vdots & & \vdots\\ \dfrac{\partial f_n}{\partial x_1}(g(\boldsymbol{y})) & \cdots & \dfrac{\partial f_n}{\partial x_n}(g(\boldsymbol{y}))\end{bmatrix}^{-1},$$

所以每个 $\dfrac{\partial g_i}{\partial x_j}(\boldsymbol{y})$ 必可由诸 $\dfrac{\partial f_k}{\partial x_\ell}(g(\boldsymbol{y}))$ $(1\leqslant k,\ell\leqslant n)$ 通过四则运算表出，然而由 $f\in C^1(U)$ 及 $g$ 的连续性知 $\dfrac{\partial f_k}{\partial x_\ell}(g(\boldsymbol{y}))$ 在 $V$ 上连续，因此 $g\in C^1(V)$。 □

**【注 4.2】**(1) 在定理 4.1 的条件下我们知道 $f$ 在 $E$ 中每个点处有局部双射，但这并不意味着 $f: E \longrightarrow f(E)$ 是双射。习题 3 给出了一个具体的例子[⑤]。

(2) 设 $E$ 是 $\mathbb{R}^n$ 中的开集，$f: E \longrightarrow \mathbb{R}^n$ 连续可微且 $f'(\boldsymbol{x})$ 对任意的 $\boldsymbol{x} \in E$ 均是非奇异的，那么由反函数定理知，对任意的 $\boldsymbol{x} \in E$，存在 $\boldsymbol{x}$ 的邻域 $U$ 以及 $f(\boldsymbol{x})$ 的邻域 $V \subseteq f(E)$ 使得 $f|_U : U \longrightarrow V$ 是双射。因此 $f(E)$ 必是开集。

(3) 在 (2) 的条件下，若 $E$ 是 $\mathbb{R}^n$ 中的区域，那么 $f(E)$ 也必是 $\mathbb{R}^n$ 中的区域。

(4) 在 (2) 的条件下，如果进一步假设 $f: E \longrightarrow \mathbb{R}^n$ 是单射（从而 $f: E \longrightarrow f(E)$ 是双射），那么一方面我们知道 $f(E)$ 是 $\mathbb{R}^n$ 中的开集；另一方面，由于对任意的 $\boldsymbol{x} \in E$，均存在 $\boldsymbol{x}$ 的一个邻域 $U$，使得 $f$ 的逆映射 $g$ 在 $f(U)$ 上连续可微，因此 $g$ 必在 $f(E)$ 上连续可微。当然，此时 (13.16) 对任意的 $\boldsymbol{y} \in f(E)$ 成立。

(5) 定理 4.1 中的条件“$f \in C^1(E)$”不能减弱为“$f$ 在 $E$ 上可微”，习题 8 给出了一个反例。

## 习 题 13.4

1. 设 $f: (u, v) \longmapsto (x, y)$ 由
$$\begin{cases} x = u + v, \\ y = \dfrac{u}{v} \end{cases}$$
所定义，问 $f$ 在哪些点的附近存在反函数？并求相应的反函数的 Jacobi 矩阵。

2. 设 $f: (\rho, \varphi, \theta) \longmapsto (x, y, z)$ 由
$$\begin{cases} x = \rho \sin\varphi \cos\theta, \\ y = \rho \sin\varphi \sin\theta, \\ z = \rho \cos\varphi, \end{cases}$$
所定义，问 $f$ 在哪些点的附近存在反函数？并求相应的反函数的 Jacobi 矩阵。

3. 设 $E = (0,1) \times (0,a)$，且 $f: E \longrightarrow \mathbb{R}^2$ 满足 $f(x,y) = (x\cos y,\, x\sin y)$。证明 $f$ 在 $E$ 中每个点处有局部双射，但当 $a > 2\pi$ 时 $f: E \longrightarrow f(E)$ 不是双射。

4. 详细写出定理 4.1 证明过程中的第一步。

---

⑤注意，如果 $n = 1$，那么必可推出 $f: E \longrightarrow f(E)$ 是双射，参见第六章注 1.8。

5. 设 $E$ 是 $\mathbb{R}^n$ 中的凸开集，$f: E \longrightarrow \mathbb{R}^n$ 可微，且对任意的 $\boldsymbol{x} \in E$ 而言 $f'(\boldsymbol{x})$ 均是正定的，证明 $f$ 是单射。

6. 设 $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 连续可微，并且存在 $\alpha > 0$，使得对任意的 $\boldsymbol{x}, \boldsymbol{h} \in \mathbb{R}^n$ 均有 $\boldsymbol{h}^{\mathrm{T}} f'(\boldsymbol{x})\boldsymbol{h} \geqslant \alpha|\boldsymbol{h}|^2$。证明：

   (1) 对任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ 有 $\langle f(\boldsymbol{x}) - f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y}\rangle \geqslant \alpha|\boldsymbol{x} - \boldsymbol{y}|^2$；

   (2) $f$ 把闭集映为闭集；

   (3) $f$ 是双射。

7. 设 $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 连续可微，对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ 而言 $f'(\boldsymbol{x})$ 均非奇异，并且若 $K$ 是紧集则 $f^{-1}(K)$ 亦是紧集。证明 $f$ 是满射。

8. 设 $f = (f_1, f_2)^{\mathrm{T}}: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ 满足 $f_1(x, y) = x$ 以及

$$f_2(x,y) = \begin{cases} \left(x^{\frac{4}{3}} + y^{\frac{4}{3}}\right)\sin\dfrac{1}{\sqrt{x^2+y^2}} + y, & \text{若 } x^2+y^2 \neq 0, \\ 0, & \text{若 } x^2+y^2 = 0. \end{cases}$$

   证明 $f'(\mathbf{0})$ 是恒等映射，但 $\dfrac{\partial f_2}{\partial x}$ 与 $\dfrac{\partial f_2}{\partial y}$ 在 $\mathbf{0}$ 处皆不连续，并且在 $\mathbf{0}$ 的任意邻域内均存在 $\boldsymbol{a} \neq \boldsymbol{b}$ 使得 $f(\boldsymbol{a}) = f(\boldsymbol{b})$。

## § 13.5 隐函数定理

我们来考虑一个相当一般化的问题：如果 $E \subseteq \mathbb{R}^{m+n}$ 且 $f: E \longrightarrow \mathbb{R}^m$ 连续可微，则

$$f(x_1, \cdots, x_{m+n}) = \mathbf{0}$$

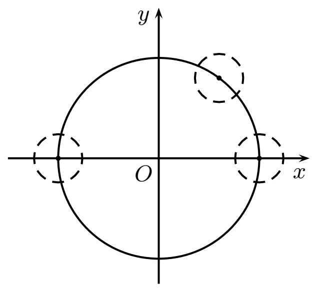
是由 $m$ 个方程组成的方程组，那么能否从 $x_1, \cdots, x_{m+n}$ 中任意指定 $n$ 个未知量，而用它们把其余 $m$ 个未知量唯一表示出来呢？这个问题提的太泛了，因此我们可以先考虑如下的局部化问题。设 $\boldsymbol{a} = (a_1, \cdots, a_n) \in \mathbb{R}^n$，$\boldsymbol{b} = (b_1, \cdots, b_m) \in \mathbb{R}^m$，我们记

$$(\boldsymbol{a}, \boldsymbol{b}) = (a_1, \cdots, a_n, b_1, \cdots, b_m) \in \mathbb{R}^{m+n}$$

如果 $f(\boldsymbol{a}, \boldsymbol{b}) = \mathbf{0}$，那么是否在 $\boldsymbol{a}$ 的某个邻域存在唯一的映射 $g$，使得对于该邻域中的每一个点 $\boldsymbol{x}$ 均有 $f(\boldsymbol{x}, g(\boldsymbol{x})) = \mathbf{0}$ 呢？

如果不对 $f$ 添加更多的条件，那么上述问题的答案在一般情况下是否定的。以二元函数

$$f(x, y) = x^2 + y^2 - 1$$




为例，满足 $f(x,y)=0$ 的 $(x,y)$ 构成了平面上的单位圆。现设 $(a,b)$ 满足 $f(a,b)=0$，那么当 $(a,b)\neq(\pm1,0)$ 时，存在 $a$ 的邻域，使得在该邻域内存在唯一的函数 $y=g(x)$ 使得 $f(x,g(x))=0$，事实上，当 $b>0$ 时可取 $g(x)=\sqrt{1-x^2}$，当 $b<0$ 时可取 $g(x)=-\sqrt{1-x^2}$。然而，对于 $(a,b)=(1,0)$ 或 $(-1,0)$ 而言，这样唯一的函数 $g$ 是不存在的，以 $(a,b)=(1,0)$ 为例，对于 $a=1$ 的左邻域内的任意一点 $x$，均存在两个 $y$ 值与之对应。

我们需要去了解产生上述差异的原因。因为 $\dfrac{\partial f}{\partial y}=2y$，所以当 $(a,b)\neq(\pm1,0)$ 时 $\dfrac{\partial f}{\partial y}(a,b)\neq 0$，而当 $(a,b)=(\pm1,0)$ 时 $\dfrac{\partial f}{\partial y}(a,b)=0$，因此对于 $(a,b)=(\pm1,0)$ 的情形而言，很可能是由于偏导数等于 0 导致了不存在唯一的函数 $y=g(x)$ 使得 $f(x,g(x))=0$。事实上我们也可以换个角度来进行观察。当 $\dfrac{\partial f}{\partial y}(a,b)\neq 0$ 时，必存在 $(a,b)$ 的邻域，使得在该邻域内 $\dfrac{\partial f}{\partial y}$ 保持符号不变，从而 $f$ 相对于变量 $y$ 而言是单调的，因此对这个邻域内的每个 $x$，至多存在一个 $y$ 满足 $f(x,y)=0$。由此可以看出，偏导数不等于 0 是一个非常关键的因素，在此基础上再添加一些条件便可对本节开头提出的局部化问题给出肯定的回答。

**【定理 5.1】（隐函数定理）** **设 $E$ 是 $\mathbb{R}^{m+n}$ 中的开集，$f=(f_1,\cdots,f_m)^{\mathrm{T}}: E\longrightarrow\mathbb{R}^m$ 连续可微。又设 $\boldsymbol{a}\in\mathbb{R}^n$ 及 $\boldsymbol{b}\in\mathbb{R}^m$，使得 $(\boldsymbol{a},\boldsymbol{b})\in E$ 且 $f(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{0}$。现将 $f$ 的 Jacobi 矩阵写成如下分块矩阵**

$$\begin{bmatrix}\dfrac{\partial f}{\partial \boldsymbol{x}} & \dfrac{\partial f}{\partial \boldsymbol{y}}\end{bmatrix}$$

**的形式，其中**

$$\frac{\partial f}{\partial \boldsymbol{x}}=\left(\frac{\partial f_i}{\partial x_j}\right)_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant n},\qquad \frac{\partial f}{\partial \boldsymbol{y}}=\left(\frac{\partial f_i}{\partial x_{n+j}}\right)_{1\leqslant i,\,j\leqslant m}.\text{[⑥]}\tag{13.20}$$

**那么当**

$$\det\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a},\boldsymbol{b})\neq 0$$

**时，存在 $\boldsymbol{a}$ 在 $\mathbb{R}^n$ 中的邻域 $U$ 以及唯一的连续可微映射 $g:U\longrightarrow\mathbb{R}^m$，使得**

**(1) $g(\boldsymbol{a})=\boldsymbol{b}$;**

---

⑥勿将这里的符号 $\dfrac{\partial f}{\partial \boldsymbol{x}}$ 及 $\dfrac{\partial f}{\partial \boldsymbol{y}}$ 与方向导数混淆。

(2) 对任意的 $\boldsymbol{x} \in U$ 有 $f(\boldsymbol{x}, g(\boldsymbol{x})) = \boldsymbol{0}$;

(3) 对任意的 $\boldsymbol{x} \in U$ 有 $\det \dfrac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x}, g(\boldsymbol{x})) \neq 0$，并且

$$g'(\boldsymbol{x}) = -\left(\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x}, g(\boldsymbol{x}))\right)^{-1} \frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{x}, g(\boldsymbol{x})). \tag{13.21}$$

证明. 在下面的讨论中，我们总默认 $\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{y} \in \mathbb{R}^m$。并且为了书写方便，我们不再严格区分列向量与行向量，也即是说，每个向量都有可能以列向量的形式出现，也有可能以行向量的形式出现，但我们相信任何一个有经验的读者均可以通过上下文将其区分。现按照

$$F(\boldsymbol{x}, \boldsymbol{y}) = \begin{bmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}, \boldsymbol{y}) \end{bmatrix} \tag{13.22}$$

定义映射 $F : E \longrightarrow \mathbb{R}^{m+n}$，那么 $F$ 在 $E$ 上连续可微，且

$$F'(\boldsymbol{a}, \boldsymbol{b}) = \begin{bmatrix} I_n & O_{n \times m} \\ \dfrac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) & \dfrac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \end{bmatrix}.$$

于是 $\det F'(\boldsymbol{a}, \boldsymbol{b}) = \det \dfrac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \neq 0$。现对 $F$ 应用反函数定理，则存在 $(\boldsymbol{a}, \boldsymbol{b})$ 在 $\mathbb{R}^{m+n}$ 中的邻域 $W$，使得 $F|_W : W \longrightarrow F(W)$ 是双射且 $F|_W$ 的逆映射在 $F(W)$ 上连续可微。由 §12.1 习题 6 知存在 $\boldsymbol{a}$ 在 $\mathbb{R}^n$ 中的邻域 $U$ 以及 $\boldsymbol{b}$ 在 $\mathbb{R}^m$ 中的邻域 $V$ 使得 $U \times V \subseteq W$，所以 $F|_{U \times V}$ 是双射，并且 $F|_{U \times V}$ 的逆映射 $H : F(U \times V) \longrightarrow U \times V$ 是连续可微的。由 (13.22) 可得

$$H(\boldsymbol{x}, f(\boldsymbol{x}, \boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y}),$$

因此 $H$ 保持前 $n$ 个坐标不变，从而可将上式写成

$$H(\boldsymbol{x}, \boldsymbol{z}) = (\boldsymbol{x}, h(\boldsymbol{x}, \boldsymbol{z})), \qquad \forall\ (\boldsymbol{x}, \boldsymbol{z}) \in F(U \times V),$$

这里 $h : F(U \times V) \longrightarrow \mathbb{R}^m$ 也是连续可微映射。

若令 $\pi : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}^m$ 是由 $\pi(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}$ 所定义的投影映射，则 $f = \pi \circ F$。于是

$$\begin{aligned} f(\boldsymbol{x}, h(\boldsymbol{x}, \boldsymbol{z})) &= f \circ H(\boldsymbol{x}, \boldsymbol{z}) = (\pi \circ F) \circ H(\boldsymbol{x}, \boldsymbol{z}) = \pi \circ (F \circ H)(\boldsymbol{x}, \boldsymbol{z}) \\ &= \pi(\boldsymbol{x}, \boldsymbol{z}) = \boldsymbol{z}. \end{aligned}$$

注意到 $F(\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{a},\boldsymbol{0})$，因此 $(\boldsymbol{a},\boldsymbol{0})\in F(U\times V)$，从而可在上式中特取 $\boldsymbol{z}=\boldsymbol{0}$ 得到

$$f(\boldsymbol{x},h(\boldsymbol{x},\boldsymbol{0}))=\boldsymbol{0}.$$

故在 $U$ 上令 $g(\boldsymbol{x})=h(\boldsymbol{x},\boldsymbol{0})$ 即满足 (1) 与 (2)。

下面来证明 $g$ 的唯一性。若在 $U$ 上还存在一个使得 (1) 与 (2) 都成立的 $\widetilde{g}$，那么对任意的 $x\in U$ 有

$$\begin{aligned}F(\boldsymbol{x},g(\boldsymbol{x}))&=(\boldsymbol{x},f(\boldsymbol{x},g(\boldsymbol{x})))=(\boldsymbol{x},\boldsymbol{0})=(\boldsymbol{x},f(\boldsymbol{x},\widetilde{g}(\boldsymbol{x})))\\&=F(\boldsymbol{x},\widetilde{g}(\boldsymbol{x})).\end{aligned}$$

于是由 $F$ 在 $U\times V$ 上是双射知 $g(\boldsymbol{x})=\widetilde{g}(\boldsymbol{x})$。

最后，由 $\det\dfrac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a},\boldsymbol{b})\neq 0$ 及 $f\in C^1(E)$ 知，可选取 $\boldsymbol{a}$ 的邻域 $U$ 使得在 $U$ 内有

$$\det\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x},g(\boldsymbol{x}))\neq 0.$$

现记 $\varphi(\boldsymbol{x})=\begin{bmatrix}\boldsymbol{x}\\ g(\boldsymbol{x})\end{bmatrix}$，那么 $f(\boldsymbol{x},g(\boldsymbol{x}))=\boldsymbol{0}$ 等价于

$$f\circ\varphi(\boldsymbol{x})=\boldsymbol{0}.$$

利用链式法则对上式求微分可得

$$\begin{aligned}\boldsymbol{0}=f'(\varphi(\boldsymbol{x}))\varphi'(\boldsymbol{x})&=\begin{bmatrix}\dfrac{\partial f}{\partial \boldsymbol{x}}(\varphi(\boldsymbol{x})) & \dfrac{\partial f}{\partial \boldsymbol{y}}(\varphi(\boldsymbol{x}))\end{bmatrix}\cdot\begin{bmatrix}I_n\\ g'(\boldsymbol{x})\end{bmatrix}\\&=\begin{bmatrix}\dfrac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{x},g(\boldsymbol{x})) & \dfrac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x},g(\boldsymbol{x}))\end{bmatrix}\cdot\begin{bmatrix}I_n\\ g'(\boldsymbol{x})\end{bmatrix}\\&=\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{x},g(\boldsymbol{x}))+\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x},g(\boldsymbol{x}))\cdot g'(\boldsymbol{x}).\end{aligned}$$

从而 (13.21) 得证。 □

在上述定理中，映射 $g$ 是“隐藏”在表达式 $f(\boldsymbol{x},g(\boldsymbol{x}))=\boldsymbol{0}$ 中的，因此我们称 $\boldsymbol{y}=g(\boldsymbol{x})$ 是由 $f(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0}$ 所确定的隐映射，当 $m=1$ 时 $g$ 是一个函数，此时也称它是由 $f(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0}$ 所确定的隐函数 (implicit function)。

在定理 5.1 的条件下，由于隐映射 $g$ 存在，并且 (13.21) 等价于

$$\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{x},\boldsymbol{y})+\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y})\cdot g'(\boldsymbol{x})=\boldsymbol{0},$$

因此我们可以从一开始就将表达式 $f(\boldsymbol{x},\boldsymbol{y}) = \mathbf{0}$ 中的 $\boldsymbol{y}$ 看作是 $\boldsymbol{x}$ 的函数并对 $f(\boldsymbol{x},\boldsymbol{y}) = \mathbf{0}$ 运用链式法则求导，当然也可用这样的方式进一步求高阶偏导。我们来看几个例子。

**【例 5.2】**设 $y = y(x)$ 由方程 $\log(x^2+y^2) = xy$ 所确定，则

$$\frac{2x+2yy'}{x^2+y^2} = y + xy'.$$

因此当 $x(x^2+y^2) \neq 2y$ 时有

$$y' = \frac{2x - y(x^2+y^2)}{x(x^2+y^2)-2y}.$$

**【例 5.3】**假设 $z = z(x,y)$ 由方程 $x - z + \sin(y+z) = 0$ 所确定，则分别对 $x$ 和 $y$ 求偏导可得

$$\begin{cases} 1 - \dfrac{\partial z}{\partial x} + \cos(y+z)\cdot\dfrac{\partial z}{\partial x} = 0 \\ -\dfrac{\partial z}{\partial y} + \cos(y+z)\cdot\left(1 + \dfrac{\partial z}{\partial y}\right) = 0, \end{cases}$$

从而当 $\cos(y+z) \neq 1$ 时有

$$\frac{\partial z}{\partial x} = \frac{1}{1-\cos(y+z)}, \qquad \frac{\partial z}{\partial y} = \frac{1}{1-\cos(y+z)} - 1.$$

进一步地，

$$\frac{\partial^2 z}{\partial x^2} = -\frac{\sin(y+z)}{[1-\cos(y+z)]^2}\cdot\frac{\partial z}{\partial x} = -\frac{\sin(y+z)}{[1-\cos(y+z)]^3},$$

$$\frac{\partial^2 z}{\partial y^2} = -\frac{\sin(y+z)}{[1-\cos(y+z)]^2}\cdot\left(1+\frac{\partial z}{\partial y}\right) = -\frac{\sin(y+z)}{[1-\cos(y+z)]^3}.$$

## 习 题 13.5

1. 计算 $\dfrac{\partial z}{\partial x}$ 及 $\dfrac{\partial z}{\partial y}$，其中：

(1) $z^3 - xyz + \sin(x+y) = 10$；　　(2) $z = e^{xyz}$；

(3) $z = \left(\dfrac{x}{y}\right)^z$；　　(4) $z\sin z = xy + \log\dfrac{z}{x}$。

2. 设 $xu-yv=0$，$yu+xv=1$，求 Jacobi 矩阵

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}.$$

3. 设 $x+y=u+v$，$xe^u=yv$，求 Jacobi 矩阵

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}.$$

4. 设 $F:\mathbb{R}^2 \longrightarrow \mathbb{R}$ 可微，证明由方程

$$F\Big(\frac{x-x_0}{z-z_0}, \frac{y-y_0}{z-z_0}\Big)=0$$

可定义函数 $z=z(x,y)$ 满足

$$(x-x_0)\frac{\partial z}{\partial x}+(y-y_0)\frac{\partial z}{\partial y}=z-z_0.$$

5. 设 $F:\mathbb{R}^2 \longrightarrow \mathbb{R}$ 可微，证明由方程

$$F\Big(x+\frac{z}{y}, y+\frac{z}{x}\Big)=0$$

可定义函数 $z=z(x,y)$ 满足

$$x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}=z-xy.$$

6. 设 $x=y+\varphi(y)$，其中 $\varphi(0)=0$ 且当 $y\in[-a,a]$ 时有 $|\varphi'(y)|\leqslant\alpha<1$。证明存在 $\varepsilon>0$，使得当 $x\in(-\varepsilon,\varepsilon)$ 时，存在唯一的可微函数 $y=y(x)$ 满足上面方程且 $y(0)=0$。

7. 设 $f(x,y)$ 是定义在 $[-a,a]\times[-b,b]$ 上的一个函数，满足:

(1) $f(0,0)=0$;

(2) 对给定的 $y\in[-b,b]$，$f(x,y)$ 作为 $x$ 的函数在 $[-a,a]$ 上连续；对给定的 $x\in[-a,a]$，$f(x,y)$ 作为 $y$ 的函数在 $[-b,b]$ 上连续；

(3) 对任意的 $y\in[-b,b]$ 有 $\dfrac{\partial f}{\partial y}(0,y)>0$。

证明存在 $a_0 \in (0, a]$ 以及定义在 $[-a_0, a_0]$ 上的函数 $g(x)$ 使得

$$f(x, g(x)) = 0, \qquad \forall\, x \in [-a_0, a_0].$$

8. 证明方程 $(x^2 + y^2)^2 = a^2(x^2 - y^2)$ $(a > 0)$ 在点 $x = 0$ 的充分小的邻域内能确定两个可微函数 $y = y(x)$。
9. 证明不存在从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的连续可微的双射。

从第 10 题到第 12 题是一组题，讨论全局问题的解。

10. 设 $E \subseteq \mathbb{R}^n$，$F$ 是 $\mathbb{R}^m$ 中的闭集，$f : E \times F \longrightarrow F$ 连续，且存在常数 $\alpha \in (0, 1)$ 使得

$$|f(\boldsymbol{x}, \boldsymbol{y}_1) - f(\boldsymbol{x}, \boldsymbol{y}_2)| \leqslant \alpha |\boldsymbol{y}_1 - \boldsymbol{y}_2|, \qquad \forall\, \boldsymbol{x} \in E,\ \boldsymbol{y}_1, \boldsymbol{y}_2 \in F.$$

证明存在唯一的连续映射 $g : E \longrightarrow F$ 使得对任意的 $\boldsymbol{x} \in E$ 有

$$f(\boldsymbol{x}, g(\boldsymbol{x})) = g(\boldsymbol{x}).$$

11. 设 $E \subseteq \mathbb{R}$，$I$ 是一个区间，$f : E \times I \longrightarrow \mathbb{R}$ 连续，且满足：

(1) 存在 $\alpha > 0$ 使得

$$\frac{\partial f}{\partial y}(x, y) \geqslant \alpha, \qquad \forall\, x \in E,\ y \in I.$$

(2) 对任意的 $x \in E$，均存在 $y \in I$ 使得 $f(x, y) = 0$。

证明存在唯一的一个连续函数 $g : E \longrightarrow I$ 使得对任意的 $x \in E$ 有

$$f(x, g(x)) = 0.$$

并说明当 $I = \mathbb{R}$ 时条件 (2) 是多余的。

12. 证明由方程 $xy = \log(x + y)$ 可唯一地确定 $\mathbb{R}_{\geqslant 2}$ 上的非负连续函数 $y = y(x)$，其满足

$$y = \frac{\log x}{x} + \frac{\log x}{x^3} + O\Big(\frac{\log^2 x}{x^5}\Big), \qquad \text{当 } x \to +\infty \text{ 时}.$$

§ 13.6

# 几何应用

## 13.6.1 空间曲线的切线与法平面

设空间曲线 $C$ 由参数方程

$$\begin{cases} x = x(t) \\ y = y(t) \qquad\qquad t \in [\alpha, \beta] \\ z = z(t) \end{cases}$$

给出，其中 $x(t)$，$y(t)$ 及 $z(t)$ 均在 $(\alpha, \beta)$ 上可导。又设 $t_0 \in (\alpha, \beta)$ 满足

$$x'(t_0)^2 + y'(t_0)^2 + z'(t_0)^2 \neq 0.$$

我们来求 $C$ 上的点 $M_0\big(x(t_0), y(t_0), z(t_0)\big)$ 处的切线方程。对于 $C$ 上任一异于 $M_0$ 的点 $M\big(x(t), y(t), z(t)\big)$ 而言，割线 $M_0M$ 的方向可由向量

$$\big(x(t) - x(t_0),\, y(t) - y(t_0),\, z(t) - z(t_0)\big)$$

表示，当然也可用向量

$$\left(\frac{x(t) - x(t_0)}{t - t_0}, \frac{y(t) - y(t_0)}{t - t_0}, \frac{z(t) - z(t_0)}{t - t_0}\right)$$

表示。现令 $t \to t_0$，那么割线 $M_0M$ 的极限位置便是 $M_0$ 处的切线，从而对上式在 $t \to t_0$ 时求极限知切线的方向可由向量

$$\big(x'(t_0), y'(t_0), z'(t_0)\big) \tag{13.23}$$

确定，我们称这一向量为 $C$ 在 $M_0$ 处的切向量。相应地，$C$ 在 $M_0$ 处的切线方程为

$$\frac{x - x(t_0)}{x'(t_0)} = \frac{y - y(t_0)}{y'(t_0)} = \frac{z - z(t_0)}{z'(t_0)}.$$

如果我们将过 $M_0$ 且与 $M_0$ 处切线垂直的平面称为曲线 $C$ 在 $M_0$ 处的法平面 (normal plane)，那么切向量 (13.23) 自然就是这一法平面的法向量，于是该法平面的方程也即

$$x'(t_0)(x - x(t_0)) + y'(t_0)(y - y(t_0)) + z'(t_0)(z - z(t_0)) = 0.$$

【**例 6.1**】设 $a,\ b>0$。对于圆柱螺旋线

$$\begin{cases} x = a\cos t \\ y = a\sin t \qquad\qquad t\in\mathbb{R} \\ z = bt \end{cases}$$

而言，在参数为 $t$ 的点处的切向量为 $(-a\sin t, a\cos t, b)$，从而在参数为 $t$ 的点处的切线方程为

$$\frac{x-a\cos t}{-a\sin t} = \frac{y-a\sin t}{a\cos t} = \frac{z-bt}{b},$$

相应的法平面方程为

$$-(a\sin t)(x-a\cos t) + (a\cos t)(y-a\sin t) + b(z-bt) = 0.$$

现假设空间曲线 $C$ 是以两个曲面的交线的形式给出的，不妨设其方程为

$$\begin{cases} F(x,y,z) = 0, \\ G(x,y,z) = 0. \end{cases} \tag{13.24}$$

又设 $M_0\,(x_0,y_0,z_0)$ 是 $C$ 上一点，$F$ 与 $G$ 均在 $M_0$ 的某邻域内连续可微，并且 $M_0$ 处的 Jacobi 行列式

$$\left.\frac{D(F,G)}{D(y,z)}\right|_{M_0} = \begin{vmatrix} \dfrac{\partial F}{\partial y}(M_0) & \dfrac{\partial F}{\partial z}(M_0) \\ \dfrac{\partial G}{\partial y}(M_0) & \dfrac{\partial G}{\partial z}(M_0) \end{vmatrix} \neq 0.$$

那么由隐函数定理知，在 $M_0$ 的某邻域内存在唯一的一对函数 $f(x)$，$g(x)$，使得

$$y = f(x), \qquad\qquad z = g(x)$$

满足 (13.24) 式，因而 $C$ 在 $M_0$ 处的切线由方程

$$\frac{x-x_0}{1} = \frac{y-y_0}{f'(x_0)} = \frac{z-z_0}{g'(x_0)} \tag{13.25}$$

所确定。下面来计算 $f'(x_0)$ 与 $g'(x_0)$。为此，在 (13.24) 中对 $x$ 求导可得

$$\begin{cases} \dfrac{\partial F}{\partial x} + \dfrac{\partial F}{\partial y}\cdot f'(x) + \dfrac{\partial F}{\partial z}\cdot g'(x) = 0, \\ \dfrac{\partial G}{\partial x} + \dfrac{\partial G}{\partial y}\cdot f'(x) + \dfrac{\partial G}{\partial z}\cdot g'(x) = 0. \end{cases}$$

由此解出

$$f'(x)=\frac{D(F,G)}{D(z,x)}\bigg/\frac{D(F,G)}{D(y,z)},$$

$$g'(x)=\frac{D(F,G)}{D(x,y)}\bigg/\frac{D(F,G)}{D(y,z)}.$$

将这代入 (12.25)，我们最终得到 $C$ 在 $M_0$ 处的切线方程为

$$\frac{x-x_0}{\left.\dfrac{D(F,G)}{D(y,z)}\right|_{M_0}}=\frac{y-y_0}{\left.\dfrac{D(F,G)}{D(z,x)}\right|_{M_0}}=\frac{z-z_0}{\left.\dfrac{D(F,G)}{D(x,y)}\right|_{M_0}}.$$

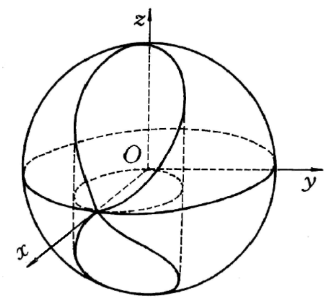
相应地，$C$ 在 $M_0$ 处的法平面方程为

$$\left.\frac{D(F,G)}{D(y,z)}\right|_{M_0}\cdot(x-x_0)+\left.\frac{D(F,G)}{D(z,x)}\right|_{M_0}\cdot(y-y_0)+\left.\frac{D(F,G)}{D(x,y)}\right|_{M_0}\cdot(z-z_0)=0.$$

尽管上述结论是在 $\left.\dfrac{D(F,G)}{D(y,z)}\right|_{M_0}\neq 0$ 的前提下得到的，但由对称性不难看出，只要 $\left.\dfrac{D(F,G)}{D(y,z)}\right|_{M_0}$，$\left.\dfrac{D(F,G)}{D(z,x)}\right|_{M_0}$ 与 $\left.\dfrac{D(F,G)}{D(x,y)}\right|_{M_0}$ 中有一个不等于 0，以上结论就成立。

【**例 6.2**】设 $a>0$。求 Viviani 曲线（如右图所示）

$$\begin{cases}x^2+y^2+z^2=a^2\\ x^2+y^2=ax\end{cases}$$




在点 $M\left(\dfrac{a}{2},\dfrac{a}{2},\dfrac{a}{\sqrt{2}}\right)$ 处的切线方程。

解. 记 $F(x,y,z)=x^2+y^2+z^2-a^2$，$G(x,y,z)=x^2+y^2-ax$，则

$$\frac{D(F,G)}{D(y,z)}=-4yz,\qquad \frac{D(F,G)}{D(z,x)}=2z(2x-a),\qquad \frac{D(F,G)}{D(x,y)}=2ay.$$

因此该曲线在 $M$ 处的切线方程为

$$\frac{x-\frac{a}{2}}{-\sqrt{2}a^2}=\frac{y-\frac{a}{2}}{0}=\frac{z-\frac{a}{\sqrt{2}}}{a^2},$$

也即

$$\begin{cases}x+\sqrt{2}z=\dfrac{3a}{2},\\ y=\dfrac{a}{2}.\end{cases}$$

□

### 13.6.2 曲面的切平面与法线

设曲面 $\Gamma$ 的方程为

$$F(x,y,z)=0. \tag{13.26}$$

$M_0\,(x_0,y_0,z_0)$ 是 $\Gamma$ 上一点，且满足 $\operatorname{grad} F(M_0)\neq\mathbf{0}$。又设 $F$ 在 $M_0$ 的某邻域内连续可微。现考虑曲面 $\Gamma$ 上过 $M_0$ 的任意一条曲线 $C$，设其由参数方程

$$\begin{cases} x=x(t) \\ y=y(t) \qquad\qquad t\in[\alpha,\beta] \\ z=z(t) \end{cases}$$

给出，并设点 $M_0$ 对应于参数 $t_0$。由于 $C$ 位于 $\Gamma$ 上，故而 $F\big(x(t),y(t),z(t)\big)=0$，对这一式子在 $t_0$ 处求导可得

$$\frac{\partial F}{\partial x}(M_0)\cdot x'(t_0)+\frac{\partial F}{\partial y}(M_0)\cdot y'(t_0)+\frac{\partial F}{\partial z}(M_0)\cdot z'(t_0)=0,$$

也即

$$\Big\langle \operatorname{grad} F(M_0),\ \big(x'(t_0),y'(t_0),z'(t_0)\big)^{\mathrm{T}}\Big\rangle=0,$$

这说明 $C$ 在 $M_0$ 处的切向量与 $F$ 在 $M_0$ 处的梯度向量垂直。注意到 $\operatorname{grad} F(M_0)$ 是一个确定的向量，故而曲面 $\Gamma$ 上过 $M_0$ 的所有曲线在 $M_0$ 的切线是共面的，我们把这一平面称为 $\Gamma$ 在 $M_0$ 处的<u>切平面 (tangent plane)</u>，由于 $\operatorname{grad} F(M_0)$ 是它的法向量，所以其方程为

$$\frac{\partial F}{\partial x}(M_0)\cdot(x-x_0)+\frac{\partial F}{\partial y}(M_0)\cdot(y-y_0)+\frac{\partial F}{\partial z}(M_0)\cdot(z-z_0)=0.$$

该切平面在 $M_0$ 处的法线被称作曲面 $\Gamma$ 在 $M_0$ 处的<u>法线 (normal line)</u>，其方程为

$$\frac{x-x_0}{\dfrac{\partial F}{\partial x}(M_0)}=\frac{y-y_0}{\dfrac{\partial F}{\partial y}(M_0)}=\frac{z-z_0}{\dfrac{\partial F}{\partial z}(M_0)}.$$

**【例 6.3】**设 $p>0$。求旋转抛物面 $x^2+y^2=2pz$ 在点 $M_0\,(\sqrt{p},\sqrt{p},1)$ 处的切平面和法线方程。

<u>解</u>. 记 $F(x,y,z)=x^2+y^2-2pz$，那么该曲面方程也即 $F(x,y,z)=0$。容易求得

$$\frac{\partial F}{\partial x}(M_0)=\frac{\partial F}{\partial y}(M_0)=2\sqrt{p}, \qquad \frac{\partial F}{\partial z}(M_0)=-2p,$$

因此 $M_0$ 处的切平面方程为 $x+y-\sqrt{p}z=\sqrt{p}$，法线方程为

$$x-\sqrt{p}=y-\sqrt{p}=-\frac{z-1}{\sqrt{p}}.$$

□

接下来考虑曲面 $\Gamma$ 是由参数方程

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

给出的情形。设 $M_0\,(x_0,y_0,z_0)$ 是 $\Gamma$ 上对应于参数 $(u_0,v_0)$ 的点，关于 $u, v$ 的函数 $x, y, z$ 均在 $M_0$ 的某邻域内连续可微。又设 $\left.\dfrac{D(x,y)}{D(u,v)}\right|_{M_0}$，$\left.\dfrac{D(y,z)}{D(u,v)}\right|_{M_0}$ 及 $\left.\dfrac{D(z,x)}{D(u,v)}\right|_{M_0}$ 中至少有一个不等于 0，不妨设

$$\left.\frac{D(x,y)}{D(u,v)}\right|_{M_0}\neq 0.$$

那么由反函数定理知，利用 $x=x(u,v)$ 和 $y=y(u,v)$ 可在 $M_0$ 的某邻域内唯一地确定两个函数

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

从而

$$z=z(u,v)=z\big(u(x,y),v(x,y)\big),$$

这样就把 $\Gamma$ 用形如 (13.26) 的方程表示了出来。按照之前的讨论，$\Gamma$ 在 $M_0$ 处的切平面存在。我们记 $(a,b,c)^{\mathrm{T}}$ 是该切平面的一个法向量，那么它就应当与 $\Gamma$ 上过 $M_0$ 的任意曲线在 $M_0$ 处的切线垂直。特别地，由

$$\begin{cases} x=x(u,v_0), \\ y=y(u,v_0), \\ z=z(u,v_0) \end{cases}$$

所给出的曲线在 $M_0$ 处的切向量为 $\left(\dfrac{\partial x}{\partial u}(M_0), \dfrac{\partial y}{\partial u}(M_0), \dfrac{\partial z}{\partial u}(M_0)\right)^{\mathrm{T}}$，而由

$$\begin{cases} x=x(u_0,v), \\ y=y(u_0,v), \\ z=z(u_0,v) \end{cases}$$

所给出的曲线在 $M_0$ 处的切向量为 $\left(\dfrac{\partial x}{\partial v}(M_0), \dfrac{\partial y}{\partial v}(M_0), \dfrac{\partial z}{\partial v}(M_0)\right)^{\mathrm{T}}$，因此

$$\begin{cases} \dfrac{\partial x}{\partial u}(M_0)\cdot a + \dfrac{\partial y}{\partial u}(M_0)\cdot b + \dfrac{\partial z}{\partial u}(M_0)\cdot c = 0, \\ \dfrac{\partial x}{\partial v}(M_0)\cdot a + \dfrac{\partial y}{\partial v}(M_0)\cdot b + \dfrac{\partial z}{\partial v}(M_0)\cdot c = 0. \end{cases}$$

由此解得

$$a = c\cdot \left.\frac{D(y,z)}{D(u,v)}\right|_{M_0}\left(\left.\frac{D(x,y)}{D(u,v)}\right|_{M_0}\right)^{-1}, \qquad b = c\cdot \left.\frac{D(z,x)}{D(u,v)}\right|_{M_0}\left(\left.\frac{D(x,y)}{D(u,v)}\right|_{M_0}\right)^{-1}.$$

从而 $\Gamma$ 在 $M_0$ 处的切平面方程为

$$\left.\frac{D(y,z)}{D(u,v)}\right|_{M_0}\cdot(x-x_0) + \left.\frac{D(z,x)}{D(u,v)}\right|_{M_0}\cdot(y-y_0) + \left.\frac{D(x,y)}{D(u,v)}\right|_{M_0}\cdot(z-z_0) = 0,$$

相应地，$\Gamma$ 在 $M_0$ 处的法线方程为

$$\frac{x-x_0}{\left.\dfrac{D(y,z)}{D(u,v)}\right|_{M_0}} = \frac{y-y_0}{\left.\dfrac{D(z,x)}{D(u,v)}\right|_{M_0}} = \frac{z-z_0}{\left.\dfrac{D(x,y)}{D(u,v)}\right|_{M_0}}.$$

## 习 题 13.6

1. 求下列曲线在指定点处的切线和法平面的方程:

   (1) $x = a\sin^2 t$，$y = b\sin t\cos t$，$z = c\cos^2 t$，在 $t = \dfrac{\pi}{4}$ 对应的点处;

   (2) $z = x^2 + y^2$，$x + y = 1$，在点 $\left(\dfrac{1}{2}, \dfrac{1}{2}, \dfrac{1}{2}\right)$ 处;

   (3) $x^2 + y^2 + z^2 = 9$，$z = xy$，在点 $(1, 2, 2)$ 处;

   (4) $x^2 + y^2 = a^2$，$x^2 + z^2 = a^2$，在点 $\left(\dfrac{a}{\sqrt{2}}, \dfrac{a}{\sqrt{2}}, \dfrac{a}{\sqrt{2}}\right)$ 处。

2. 求下列曲面在指定点处的切平面和法线的方程:

   (1) $z = x^2 - 9y^2$，在点 $(4, -1, 5)$ 处;

   (2) $z = \arctan\dfrac{x}{y}$，在点 $\left(1, 1, \dfrac{\pi}{4}\right)$ 处;

   (3) $x = u\cos v$，$y = u\sin v$，$z = av$，在参数 $(u_0, v_0)$ 对应的点处;

   (4) $x = a\cos\varphi\cos\theta$，$y = b\cos\varphi\sin\theta$，$z = c\sin\theta$，在参数 $(\varphi_0, \theta_0)$ 对应的点处。

3. 设 $a>0$，证明曲面 $\sqrt{x}+\sqrt{y}+\sqrt{z}=\sqrt{a}$ 上任一点处的切平面在三个坐标轴上的截距之和是一个常数。
4. 证明曲面 $xyz=a^3$ $(a>0)$ 上任一点处的切平面与三个坐标面围成的四面体体积均为 $\dfrac{9}{2}a^3$。
5. 设 $f$ 在 $\mathbb{R}_{\geqslant 0}$ 上可导，且导数不等于 0。证明旋转面 $z=f\big(\sqrt{x^2+y^2}\,\big)$ 的法线与旋转轴相交。
6. 设 $f\in C^1(\mathbb{R}^2)$，$a$ 和 $b$ 是常数。证明曲面 $f(x-az,y-bz)=0$ 上任一点处的切平面均与某定直线平行。
7. 设 $f\in C^1(\mathbb{R}^2)$，$a$，$b$，$c$ 为常数。证明曲面 $F\left(\dfrac{x-a}{z-c},\dfrac{y-b}{z-c}\right)=0$ 的切平面过一固定点。

## § 13.7 多元函数的极值与条件极值

### 13.7.1 极值

**【定义 7.1】**设 $E\subseteq\mathbb{R}^n$，$f:E\longrightarrow\mathbb{R}$，$\boldsymbol{a}$ 是 $E$ 的内点。如果存在 $\varepsilon>0$，使得 $B(\boldsymbol{a},\varepsilon)\subseteq E$，并且对任意的 $\boldsymbol{x}\in B(\boldsymbol{a},\varepsilon)$ 均有 $f(\boldsymbol{x})\geqslant f(\boldsymbol{a})$（相应地，$f(\boldsymbol{x})\leqslant f(\boldsymbol{a})$），那么就称 $\boldsymbol{a}$ 为 $f$ 的一个极小值点（相应地，极大值点），并称 $f(\boldsymbol{a})$ 是 $f$ 的一个极小值（相应地，极大值）。

极小值点和极大值点被统称为极值点，极小值和极大值被统称为极值。

对应于 Fermat 定理（第六章定理 1.3），我们有下述命题。

**【命题 7.2】**设 $E\subseteq\mathbb{R}^n$，$\boldsymbol{a}$ 是 $E$ 的内点，且 $f:E\longrightarrow\mathbb{R}$ 在 $\boldsymbol{a}$ 处的各偏导数均存在。如果 $\boldsymbol{a}$ 是 $f$ 的极值点，那么

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a})=0,\qquad j=1,\,2,\,\cdots,\,n.\tag{13.27}$$

证明. 不妨设 $\boldsymbol{a}$ 是 $f$ 的极小值点。按照定义，存在 $\varepsilon>0$，使得 $B(\boldsymbol{a},\varepsilon)\subseteq E$，并且

$$f(\boldsymbol{x})\geqslant f(\boldsymbol{a}),\qquad\forall\,\boldsymbol{x}\in B(\boldsymbol{a},\varepsilon).$$

若记 $\boldsymbol{a}=(a_1,\cdots,a_n)^{\mathrm{T}}$，那么对于满足 $|x_j-a_j|<\varepsilon$ 的任意实数 $x_j$ 有

$$f(a_1,\cdots,a_{j-1},x_j,a_{j+1},\cdots,a_n)\geqslant f(\boldsymbol{a}).$$

这说明上式左侧关于 $x_j$ 的函数在 $a_j$ 处取到极小值，从而由 Fermat 定理知这一函数在 $a_j$ 处的导数为 0，也即

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a})=0.$$

□

【注 7.3】若 $\boldsymbol{a}$ 是 $f$ 的极值点且 $f$ 在 $\boldsymbol{a}$ 处可微，则 $f'(\boldsymbol{a})=\boldsymbol{0}$。

我们把满足 (13.27) 的点称为 $f$ 的驻点。按照命题 7.2，$f$ 的极值点只可能是驻点或者至少有一个偏导数不存在的点。

与一元函数的情况类似，对于多元函数而言驻点也未必是极值点，例如 $(0,0)$ 是函数 $f(x,y)=xy$ 的驻点，但由 $z=xy$ 的图形是马鞍面知 $(0,0)$ 不是 $f$ 的极值点。下面我们将利用 Taylor 公式给出判定极值点的一个充分条件，回忆起 Hesse 矩阵

$$H_f(\boldsymbol{a})=\begin{bmatrix}\dfrac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_1\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}(\boldsymbol{a})\\ \dfrac{\partial^2 f}{\partial x_2\partial x_1}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n}(\boldsymbol{a})\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial^2 f}{\partial x_n\partial x_1}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x_n\partial x_2}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(\boldsymbol{a})\end{bmatrix}.$$

再回忆起线性代数中所提到的正定矩阵的概念，那就是对于矩阵 $A$ 而言，如果对任意的 $\boldsymbol{x}\neq\boldsymbol{0}$ 均有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}>0$（相应地，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}<0$），则称 $A$ 是正定矩阵（相应地，负定矩阵）；如果存在非零向量 $\boldsymbol{x},\boldsymbol{y}$ 使得 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}<0<\boldsymbol{y}^{\mathrm{T}}A\boldsymbol{y}$，则称 $A$ 是不定矩阵。

**【命题 7.4】**设 $E$ 是 $\mathbb{R}^n$ 中的开集，$f:E\longrightarrow\mathbb{R}$ 是 $C^2$ 类的，又设 $\boldsymbol{a}\in E$ 是 $f$ 的驻点。那么

(1) 若 $H_f(\boldsymbol{a})$ 是正定的，则 $\boldsymbol{a}$ 是 $f$ 的极小值点；

(2) 若 $H_f(\boldsymbol{a})$ 是负定的，则 $\boldsymbol{a}$ 是 $f$ 的极大值点；

(3) 若 $H_f(\boldsymbol{a})$ 是不定的，则 $\boldsymbol{a}$ 不是 $f$ 的极值点。

证明. 由带 Peano 余项的 Taylor 公式（定理 3.4）知，当 $\boldsymbol{h}\to\boldsymbol{0}$ 时有

$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+f'(\boldsymbol{a})\boldsymbol{h}+\frac{1}{2}\boldsymbol{h}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{h}+o(|\boldsymbol{h}|^2).$$

因为 $\boldsymbol{a}$ 是 $f$ 的驻点，故 $f'(\boldsymbol{a})=\boldsymbol{0}$，从而上式也即

$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+\frac{1}{2}\boldsymbol{h}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{h}+o(|\boldsymbol{h}|^2). \tag{13.28}$$

现在考虑定义在 $\partial B(\mathbf{0},1)=\{\boldsymbol{x}\in\mathbb{R}^n:|\boldsymbol{x}|=1\}$ 上的 $n$ 元函数

$$\varphi(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{x},$$

这是一个连续函数（参见 §12.3 习题 7），故由 $\partial B(\mathbf{0},1)$ 是闭集知 $\varphi(\boldsymbol{x})$ 在 $\partial B(\mathbf{0},1)$ 上能取到最大值和最小值。当 $H_f(\boldsymbol{a})$ 是正定矩阵时，由于 $\varphi(\boldsymbol{x})>0$ $(\forall\ \boldsymbol{x}\neq\mathbf{0})$，因此 $\varphi(\boldsymbol{x})$ 在 $\partial B(\mathbf{0},1)$ 上的最小值 $m>0$，于是对任意的 $\boldsymbol{h}\neq 0$ 有

$$\boldsymbol{h}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{h}=|\boldsymbol{h}|^2\cdot\left(\frac{\boldsymbol{h}}{|\boldsymbol{h}|}\right)^{\mathrm{T}}H_f(\boldsymbol{a})\frac{\boldsymbol{h}}{|\boldsymbol{h}|}\geqslant m|\boldsymbol{h}|^2.$$

从而由 (13.28) 知当 $\boldsymbol{h}\to\mathbf{0}$ 时

$$f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})\geqslant\left(\frac{1}{2}m+o(1)\right)|\boldsymbol{h}|^2>0,$$

所以 $\boldsymbol{a}$ 是 $f$ 的极小值点。同理可证当 $H_f(\boldsymbol{a})$ 是负定矩阵时 $\boldsymbol{a}$ 是 $f$ 的极大值点。

如果 $H_f(\boldsymbol{a})$ 是不定矩阵，那么存在 $\boldsymbol{p},\boldsymbol{q}\in\mathbb{R}^n$ 使得

$$\boldsymbol{p}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{p}<0<\boldsymbol{q}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{q}.$$

于是由 (13.28) 知，当 $\varepsilon\to 0$ 时有

$$f(\boldsymbol{a}+\varepsilon\boldsymbol{p})-f(\boldsymbol{a})=\left(\frac{1}{2}\boldsymbol{p}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{p}\right)\cdot\varepsilon^2+o(\varepsilon^2)<0,$$

$$f(\boldsymbol{a}+\varepsilon\boldsymbol{q})-f(\boldsymbol{a})=\left(\frac{1}{2}\boldsymbol{q}^{\mathrm{T}}H_f(\boldsymbol{a})\boldsymbol{q}\right)\cdot\varepsilon^2+o(\varepsilon^2)>0,$$

故而 $\boldsymbol{a}$ 不是 $f$ 的极值点。 □

特别地，当 $f$ 是 $C^2$ 类的二元函数时，

$$H_f(\boldsymbol{a})=\begin{bmatrix}\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial x\partial y}(\boldsymbol{a})\\[2ex] \dfrac{\partial^2 f}{\partial x\partial y}(\boldsymbol{a}) & \dfrac{\partial^2 f}{\partial y^2}(\boldsymbol{a})\end{bmatrix}.$$

于是 $H_f(\boldsymbol{a})$ 正定当且仅当 $\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a})>0$ 且 $\det H_f(\boldsymbol{a})>0$；$H_f(\boldsymbol{a})$ 负定当且仅当 $\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a})<0$ 且 $\det H_f(\boldsymbol{a})>0$；$H_f(\boldsymbol{a})$ 不定当且仅当 $\det H_f(\boldsymbol{a})<0$。因此有如下推论。

**【推论 7.5】** 设 $E$ 是 $\mathbb{R}^2$ 中的开集，$f: E \longrightarrow \mathbb{R}$ 是 $C^2$ 类的，又设 $\boldsymbol{a} \in E$ 是 $f$ 的驻点。那么

(1) 当 $\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a}) > 0$ 且 $\det H_f(\boldsymbol{a}) > 0$ 时 $\boldsymbol{a}$ 是 $f$ 的极小值点；

(2) 当 $\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a}) < 0$ 且 $\det H_f(\boldsymbol{a}) > 0$ 时 $\boldsymbol{a}$ 是 $f$ 的极大值点；

(3) 当 $\det H_f(\boldsymbol{a}) < 0$ 时 $\boldsymbol{a}$ 不是 $f$ 的极值点。

**【例 7.6】** 求函数 $f(x,y) = x^3 + x^2 - xy + y^2$ 的极值点。

解. 易见

$$\frac{\partial f}{\partial x} = 3x^2 + 2x - y, \qquad \frac{\partial f}{\partial y} = 2y - x,$$

因此令 $\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$ 可求得驻点 $\mathbf{0} = (0,0)^{\mathrm{T}}$ 及 $\boldsymbol{a} = \left(-\dfrac{1}{2}, -\dfrac{1}{4}\right)^{\mathrm{T}}$。进一步，

$$\frac{\partial^2 f}{\partial x^2} = 6x + 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = -1, \qquad \frac{\partial^2 f}{\partial y^2} = 2,$$

所以 $\dfrac{\partial^2 f}{\partial x^2}(\mathbf{0}) = 2$，$\dfrac{\partial^2 f}{\partial x^2}(\boldsymbol{a}) = -1$，并且由 $\det H_f = 12x + 3$ 可得

$$\det H_f(\mathbf{0}) = 3, \qquad \det H_f(\boldsymbol{a}) = -3$$

故由推论 7.5 知 $(0,0)^{\mathrm{T}}$ 是 $f$ 的极小值点，$\left(-\dfrac{1}{2}, -\dfrac{1}{4}\right)^{\mathrm{T}}$ 不是 $f$ 的极值点。 □

### 13.7.2　条件极值

在实际应用中，我们通常需要在一定的限制条件下去求某个函数的极值，例如在满足方程组

$$\begin{cases} g_1(x_1, \cdots, x_{m+n}) = 0 \\ \quad \cdots \ \cdots \ \cdots \ \cdots \\ g_m(x_1, \cdots, x_{m+n}) = 0 \end{cases} \tag{13.29}$$

的点集上去考虑函数 $f(x_1, \cdots, x_{m+n})$ 的极值，这被称作<u>条件极值</u>。如果记 $g = (g_1, \cdots, g_m)^{\mathrm{T}}$，那么上式也可被简记作

$$g(x_1, \cdots, x_{m+n}) = \mathbf{0}.$$

我们首先通过一个简单情形来作启发式的讨论，即去求函数 $f(x,y)$ 在限制条件 $g(x,y) = 0$ 下的极值，其中 $f$ 与 $g$ 均在 $\mathbb{R}^2$ 上连续可微。如果 $\dfrac{\partial g}{\partial y} \neq 0$，那

么由隐函数定理知通过 $g(x,y)=0$ 可将 $y$ 表成 $x$ 的函数，也即存在函数 $h$ 使得 $g(x,h(x))=0$，于是上述条件极值问题就转化为求 $f(x,y)=f(x,h(x))$ 的极值。此时，若 $\boldsymbol{a}=(x_0,y_0)$ 是 $f$ 在条件 $g(x,y)=0$ 下的极值点，那么 $x_0$ 必然是 $f(x,h(x))$ 的极值点，从而 $x_0$ 是 $f(x,h(x))$ 的驻点，进而由链式法则知

$$\frac{\partial f}{\partial x}(\boldsymbol{a})+\frac{\partial f}{\partial y}(\boldsymbol{a})\cdot h'(x_0)=0.$$

此外，按照隐函数定理，

$$\frac{\partial g}{\partial x}(\boldsymbol{a})+\frac{\partial g}{\partial y}(\boldsymbol{a})\cdot h'(x_0)=0.$$

因此

$$\begin{bmatrix}\dfrac{\partial f}{\partial x}(\boldsymbol{a}) & \dfrac{\partial f}{\partial y}(\boldsymbol{a})\\ \dfrac{\partial g}{\partial x}(\boldsymbol{a}) & \dfrac{\partial g}{\partial y}(\boldsymbol{a})\end{bmatrix}\cdot\begin{bmatrix}1\\ h'(x_0)\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

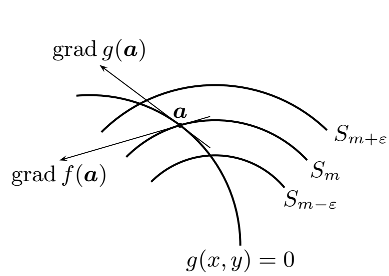
这说明 $\operatorname{grad} f(\boldsymbol{a})=\left(\dfrac{\partial f}{\partial x}(\boldsymbol{a}) \quad \dfrac{\partial f}{\partial y}(\boldsymbol{a})\right)$ 与 $\operatorname{grad} g(\boldsymbol{a})=\left(\dfrac{\partial g}{\partial x}(\boldsymbol{a}) \quad \dfrac{\partial g}{\partial y}(\boldsymbol{a})\right)$ 线性相关，也即存在 $\mu$ 使得

$$\operatorname{grad} f(\boldsymbol{a})=\mu\cdot\operatorname{grad} g(\boldsymbol{a}). \tag{13.30}$$

我们再从几何观点来看待一下上述问题，为此，记

$$S_\alpha=\{(x,y)\in\mathbb{R}^2 : f(x,y)=\alpha\}$$

并称之为 $f$ 的水平集 (level set)。现假设 $\boldsymbol{a}=(x_0,y_0)$ 是函数 $f(x,y)$ 在限制条件 $g(x,y)=0$ 下的极值点，且 $f(\boldsymbol{a})=m$，我们来对充分小的 $\varepsilon>0$ 同时考察水平集 $S_{m-\varepsilon}$，$S_m$ 以及 $S_{m+\varepsilon}$（如右图所示）。因为 $g(\boldsymbol{a})=0$，所以 $g(x,y)=0$ 所对应的曲线与 $S_m$ 在 $\boldsymbol{a}$ 点相交。注意到 $\operatorname{grad} f(\boldsymbol{a})$ 与 $\operatorname{grad} g(\boldsymbol{a})$ 分别表示曲线 $f(x,y)=m$ 和 $g(x,y)=0$ 在 $\boldsymbol{a}$ 处的切向量，因此如果 (13.30) 不成立，那么这两个切向量不平行。于是对于 $\boldsymbol{a}$ 的任一邻域 $U$，我们均可将 $\varepsilon$ 取得足够小，使得 $g(x,y)=0$ 所对应的曲线在 $U$ 内与 $S_{m-\varepsilon}$ 及 $S_{m+\varepsilon}$ 均有交集。换句话说，在 $U$ 内不但存在着同时满足 $g(x,y)=0$ 及 $f(x,y)=m-\varepsilon$ 的点，也存在着同时满足 $g(x,y)=0$ 及 $f(x,y)=m+\varepsilon$ 的点，这意味着 $\boldsymbol{a}$ 不再是 $f$ 在条件



$g(x,y)=0$ 下的极值点，从而与最初的假设矛盾。以上讨论说明，如果 $\boldsymbol{a}$ 是 $f$ 在条件 $g(x,y)=0$ 下的极值点，那么曲线 $g(x,y)=0$ 与 $f(x,y)=m$ 在 $\boldsymbol{a}$ 处的切线必然重合，也即 (13.30) 成立，用通俗的话说，这两条曲线在 $\boldsymbol{a}$ 点处应当拟合得很好。

如果 $\boldsymbol{a}=(x_0,y_0,z_0)$ 是函数 $f(x,y,z)$ 在限制条件 $g_1(x,y,z)=0$ 与 $g_2(x,y,z)=0$ 下的极值点且 $f(\boldsymbol{a})=m$，那么与上面的讨论类似，由

$$\begin{cases} g_1(x,y,z)=0 \\ g_2(x,y,z)=0 \end{cases}$$

所确定的曲线在 $\boldsymbol{a}$ 处的切线 $\ell$ 应当位于 $f(x,y,z)=m$ 在 $\boldsymbol{a}$ 处的切平面上。因为 $\ell$ 也位于 $g_j(x,y,z)=0$ $(j=1,2)$ 在 $\boldsymbol{a}$ 处的切平面上，所以三个曲面 $f(x,y,z)=m$，$g_j(x,y,z)=0$ $(j=1,2)$ 在 $\boldsymbol{a}$ 处的三个切平面的法向量必然共面（因为它们均与 $\ell$ 垂直），注意到这三个法向量分别为 $\operatorname{grad} f(\boldsymbol{a})$ 和 $\operatorname{grad} g_j(\boldsymbol{a})$ $(j=1,2)$，故而存在 $\mu_1$, $\mu_2$ 使得

$$\operatorname{grad} f(\boldsymbol{a})=\mu_1\cdot\operatorname{grad} g_1(\boldsymbol{a})+\mu_2\cdot\operatorname{grad} g_2(\boldsymbol{a}).$$

更一般地，如果 $\boldsymbol{a}$ 是函数 $f(x_1,\cdots,x_{m+n})$ 在条件 (13.29) 下的极值点，那么应当存在 $\mu_1,\cdots,\mu_m$ 使得

$$\operatorname{grad} f(\boldsymbol{a})=\sum_{j=1}^{m}\mu_j\cdot\operatorname{grad} g_j(\boldsymbol{a}),$$

若记 $g=(g_1,\cdots,g_m)^{\mathrm{T}}$ 以及 $\boldsymbol{\mu}=(\mu_1,\cdots,\mu_m)^{\mathrm{T}}$，那么上式也即

$$\operatorname{grad} f(\boldsymbol{a})=\boldsymbol{\mu}^{\mathrm{T}}\cdot g'(\boldsymbol{a}).$$

以上讨论结果可用严格语言描述如次。

**【命题 7.7】** 设 $E$ 是 $\mathbb{R}^{m+n}$ 中的开集，$f:E\longrightarrow\mathbb{R}$ 及 $g:E\longrightarrow\mathbb{R}^m$ 均在 $E$ 上连续可微，且集合 $\{\boldsymbol{z}\in\mathbb{R}^{m+n}:g(\boldsymbol{z})=\boldsymbol{0}\}$ 不是空集。如果 $\boldsymbol{x}_0\in\mathbb{R}^n$，$\boldsymbol{y}_0\in\mathbb{R}^m$ 满足 $\boldsymbol{a}=\begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{y}_0\end{bmatrix}\in E$ 以及 $\det\dfrac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\neq 0$ [7]，并且 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\boldsymbol{0}$ 下的极值点，那么必存在 $\boldsymbol{\mu}\in\mathbb{R}^m$ 使得

$$\operatorname{grad} f(\boldsymbol{a})=\boldsymbol{\mu}^{\mathrm{T}}\cdot g'(\boldsymbol{a}). \tag{13.31}$$

[7] 这里符号 $\dfrac{\partial g}{\partial \boldsymbol{y}}$ 的定义参见 (13.20)。

证明. 沿用 (13.20) 中的记号。由隐函数定理知，存在 $\boldsymbol{x}_0$ 在 $\mathbb{R}^n$ 中的一个邻域 $U$ 以及唯一的连续可微映射 $h : U \longrightarrow \mathbb{R}^m$ 使得

$$h(\boldsymbol{x}_0) = \boldsymbol{y}_0 \qquad \text{及} \qquad g(\boldsymbol{x}, h(\boldsymbol{x})) = \boldsymbol{0}\ (\forall\ x \in U).$$

由于 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$ 下的极值点，故 $\boldsymbol{x}_0$ 是 $f(\boldsymbol{x}, h(\boldsymbol{x}))$ 的极值点。于是由命题 7.2 知 $\boldsymbol{x}_0$ 是 $f(\boldsymbol{x}, h(\boldsymbol{x}))$ 的驻点，从而可由链式法则得到

$$\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{a}) + \frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a}) h'(\boldsymbol{x}_0) = \boldsymbol{0}.$$

注意到由隐函数定理知 $h'(\boldsymbol{x}_0) = -\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\right)^{-1} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a})$，故而上式也即

$$\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{a}) = \frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a})\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\right)^{-1} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a}). \tag{13.32}$$

现记 $\boldsymbol{\mu}^{\mathrm{T}} = \frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a})\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\right)^{-1}$，即

$$\frac{\partial f}{\partial \boldsymbol{y}}(\boldsymbol{a}) = \boldsymbol{\mu}^{\mathrm{T}} \cdot \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a}),$$

那么由 (13.32) 知

$$\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{a}) = \boldsymbol{\mu}^{\mathrm{T}} \cdot \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a}),$$

综合上两式便得 $\operatorname{grad} f(\boldsymbol{a}) = \boldsymbol{\mu}^{\mathrm{T}} \cdot g'(\boldsymbol{a})$。 □

如果在上述命题中记 $\boldsymbol{\lambda} = -\boldsymbol{\mu} = (\lambda_1, \cdots, \lambda_m)^{\mathrm{T}}$，那么 (13.31) 式也即

$$f'(\boldsymbol{a}) + \boldsymbol{\lambda}^{\mathrm{T}} \cdot g'(\boldsymbol{a}) = \boldsymbol{0}.$$

因此若令

$$F(\boldsymbol{z}) = f(\boldsymbol{z}) + \boldsymbol{\lambda}^{\mathrm{T}} \cdot g(\boldsymbol{z}) = f(\boldsymbol{z}) + \sum_{j=1}^{m} \lambda_j g_j(\boldsymbol{z}), \tag{13.33}$$

则 $f$ 在条件 $g(\boldsymbol{z}) = \boldsymbol{0}$ 下的极值点必然是 $F$ 的驻点。换句话说，为了求 $f$ 在条件 $g(\boldsymbol{z}) = \boldsymbol{0}$ 下的极值点，我们可以通过引入一个待定参量 $\boldsymbol{\lambda}$ 来构造形如 (13.33) 的函数 $F$，然后在 $F$ 的驻点中寻求所需的条件极值点，这一方法被称作 Lagrange 乘子法 (Lagrange multiplier method)。我们把它用下述命题总结出来。

**【命题 7.8】** 设 $E$ 是 $\mathbb{R}^{m+n}$ 中的开集，$f: E \longrightarrow \mathbb{R}$ 及 $g=(g_1,\cdots,g_m)^{\mathrm{T}}: E \longrightarrow \mathbb{R}^m$ 均在 $E$ 上连续可微，$\boldsymbol{a} \in E$ 满足 $g(\boldsymbol{a})=\mathbf{0}$，且 $g'(\boldsymbol{a})$ 所对应的矩阵的秩为 $m$[8]。现设

$$F(\boldsymbol{z})=f(\boldsymbol{z})+\boldsymbol{\lambda}\cdot g(\boldsymbol{z})=f(\boldsymbol{z})+\sum_{j=1}^{m}\lambda_j g_j(\boldsymbol{z}),$$

其中 $\boldsymbol{\lambda}=(\lambda_1,\cdots,\lambda_m)^{\mathrm{T}}$ 是 $\mathbb{R}^m$ 中使得 $\boldsymbol{a}$ 是 $F(\boldsymbol{z})$ 的驻点的某个元素。那么

(1) 如果 $\boldsymbol{a}$ 是 $F$ 的极小值点，则它是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极小值点；

(2) 如果 $\boldsymbol{a}$ 是 $F$ 的极大值点，则它是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极大值点。

证明. 这是因为当 $g(\boldsymbol{z})=\mathbf{0}$ 时有 $f(\boldsymbol{z})=F(\boldsymbol{z})$。 □

利用命题 7.4 可得如下结论。

**【推论 7.9】** 在命题 7.8 的假设条件下，

(1) 如果 $H_F(\boldsymbol{a})$ 是正定矩阵，那么 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极小值点；

(2) 如果 $H_F(\boldsymbol{a})$ 是负定矩阵，那么 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极大值点。

当 $H_F(\boldsymbol{a})$ 是不定矩阵时（此时 $\boldsymbol{a}$ 不是 $F$ 的极值点），$\boldsymbol{a}$ 仍有可能是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极值点。事实上，只需对于满足 $g(\boldsymbol{a}+\boldsymbol{h})=0$ 的任意非零元 $\boldsymbol{h}$ 均有 $\boldsymbol{h}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{h}>0$（相应地，$\boldsymbol{h}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{h}<0$）就能保证 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极小值点（相应地，极大值点）。受这一思想启发，我们可得下述结论。

**【命题 7.10】** 在命题 7.8 的假设条件下，

(1) 如果对于满足 $g'(\boldsymbol{a})\boldsymbol{h}=\mathbf{0}$ 的任一非零元 $\boldsymbol{h}$ 均有 $\boldsymbol{h}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{h}>0$，那么 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极小值点；

(2) 如果对于满足 $g'(\boldsymbol{a})\boldsymbol{h}=\mathbf{0}$ 的任一非零元 $\boldsymbol{h}$ 均有 $\boldsymbol{h}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{h}<0$，那么 $\boldsymbol{a}$ 是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极大值点；

(3) 若存在非零元 $\boldsymbol{p},\boldsymbol{q}\in\mathbb{R}^{m+n}$ 使得 $g'(\boldsymbol{a})\boldsymbol{p}=g'(\boldsymbol{a})\boldsymbol{q}=\mathbf{0}$ 并且

$$\boldsymbol{p}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{p}<0<\boldsymbol{q}^{\mathrm{T}}H_F(\boldsymbol{a})\boldsymbol{q},$$

那么 $\boldsymbol{a}$ 不是 $f$ 在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下的极值点。

证明. 因为在条件 $g(\boldsymbol{z})=\mathbf{0}$ 下有 $F(\boldsymbol{z})=f(\boldsymbol{z})$，因此对于满足 $\boldsymbol{a}+\boldsymbol{\delta}\in E$ 及 $g(\boldsymbol{a}+\boldsymbol{\delta})=\mathbf{0}$ 并且位于 $\mathbf{0}$ 的充分小的去心邻域内的 $\boldsymbol{\delta}$，由带 Peano 余项的 Taylor 公式（定理 3.4）知

$$f(\boldsymbol{a}+\boldsymbol{\delta})-f(\boldsymbol{a})=F(\boldsymbol{a}+\boldsymbol{\delta})-F(\boldsymbol{a})$$

---

[8] 我们以后把这一条件简记作 $\operatorname{rank} g'(\boldsymbol{a})=m$。

$$
\begin{aligned}
&= F'(\boldsymbol{a})\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\delta}^{\mathrm{T}} H_F(\boldsymbol{a})\boldsymbol{\delta} + o(|\boldsymbol{\delta}|^2) \\
&= \frac{1}{2}\boldsymbol{\delta}^{\mathrm{T}} H_F(\boldsymbol{a})\boldsymbol{\delta} + o(|\boldsymbol{\delta}|^2). \qquad (13.34)
\end{aligned}
$$

因为 $\operatorname{rank} g'(\boldsymbol{a}) = m$，所以由隐函数定理知，存在 $\boldsymbol{a}$ 的一个邻域 $U$，使得对于满足 $g(\boldsymbol{z}) = \boldsymbol{0}$ 的 $\boldsymbol{z} \in U$，我们可以从 $\boldsymbol{z}$ 的诸分量中选出 $n$ 个，使得它们可把其余 $m$ 个分量用唯一的方式表出。在此不妨记 $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$，其中 $\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{y} \in \mathbb{R}^m$，并假设 $\boldsymbol{y}$ 可由 $\boldsymbol{x}$ 唯一表出。从而存在从 $\mathbb{R}^n$ 的某个开子集到 $\mathbb{R}^m$ 的连续可微映射 $\varphi$，使得 $\boldsymbol{y} = \varphi(\boldsymbol{x})$，也即 $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \varphi(\boldsymbol{x}) \end{bmatrix}$。

现将由 $\boldsymbol{a}$ 和 $\boldsymbol{\delta}$ 的前 $n$ 个分量所构成的向量分别记作 $\boldsymbol{b}$ 和 $\boldsymbol{\xi}$，则

$$
\boldsymbol{a} = \begin{bmatrix} \boldsymbol{b} \\ \varphi(\boldsymbol{b}) \end{bmatrix}, \qquad (\boldsymbol{a} + \boldsymbol{\delta}) = \begin{bmatrix} \boldsymbol{b} + \boldsymbol{\xi} \\ \varphi(\boldsymbol{b} + \boldsymbol{\xi}) \end{bmatrix}.
$$

因此 $\boldsymbol{\delta} \to \boldsymbol{0}$ 当且仅当 $\boldsymbol{\xi} \to \boldsymbol{0}$，并且

$$
\boldsymbol{\delta} = (\boldsymbol{a} + \boldsymbol{\delta}) - \boldsymbol{a} = \begin{bmatrix} \boldsymbol{\xi} \\ \varphi(\boldsymbol{b} + \boldsymbol{\xi}) - \varphi(\boldsymbol{b}) \end{bmatrix} = \begin{bmatrix} I_n \\ \varphi'(\boldsymbol{b}) \end{bmatrix} \boldsymbol{\xi} + o(|\boldsymbol{\xi}|),
$$

上面最后一步可由微分的定义得到。记

$$
\boldsymbol{\eta} = \begin{bmatrix} I_n \\ \varphi'(\boldsymbol{b}) \end{bmatrix} \boldsymbol{\xi}, \tag{13.35}
$$

则有 $\boldsymbol{\delta} = \boldsymbol{\eta} + o(|\boldsymbol{\xi}|) = \boldsymbol{\eta} + o(|\boldsymbol{\eta}|)$，将这代入 (13.34) 就有

$$
f(\boldsymbol{a} + \boldsymbol{\delta}) - f(\boldsymbol{a}) = \frac{1}{2}\boldsymbol{\eta}^{\mathrm{T}} H_F(\boldsymbol{a})\boldsymbol{\eta} + o(|\boldsymbol{\eta}|^2).
$$

注意到由隐函数定理知 $\varphi'(\boldsymbol{b}) = -\left(\dfrac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\right)^{-1} \cdot \dfrac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a})$，故而

$$
\boldsymbol{0} = \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a}) + \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a})\varphi'(\boldsymbol{b}) = \begin{bmatrix} \dfrac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{a}) & \dfrac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{a}) \end{bmatrix} \cdot \begin{bmatrix} I_n \\ \varphi'(\boldsymbol{b}) \end{bmatrix} = g'(\boldsymbol{a}) \begin{bmatrix} I_n \\ \varphi'(\boldsymbol{b}) \end{bmatrix}
$$

从而 (13.35) 式中的 $\boldsymbol{\eta}$ 应当满足 $g'(\boldsymbol{a})\boldsymbol{\eta} = \boldsymbol{0}$。接下来的证明过程与命题 7.4 的证明类似，不再赘述。 □

下面来看两个例子，它们的结果事实上均可由 Cauchy–Schwarz 不等式得到，但我们的目的是为了阐述本节的方法。

**【例 7.11】** 设 $a_i > 0\ (i=1,\cdots,n)$。求函数 $f(x_1,\cdots,x_n)=\sum\limits_{j=1}^{n} a_j x_j^2$ 在条件 $\sum\limits_{j=1}^{n} x_j = 1$ 下的极值。

<u>解</u>. 作辅助函数

$$F(x_1,\cdots,x_n)=\sum_{j=1}^{n} a_j x_j^2+\lambda\Bigl(\sum_{j=1}^{n} x_j - 1\Bigr).$$

由方程组

$$\begin{cases}\dfrac{\partial f}{\partial x_j}=2a_jx_j+\lambda=0, \qquad j=1,\cdots,n,\\ \sum\limits_{j=1}^{n} x_j=1\end{cases}$$

可解得

$$x_j=\Bigl(a_j\sum_{k=1}^{n}\frac{1}{a_k}\Bigr)^{-1}, \qquad \lambda=-2\Bigl(\sum_{k=1}^{n}\frac{1}{a_k}\Bigr)^{-1}. \tag{13.36}$$

注意到 $H_F=\begin{bmatrix}2a_1 & & \\ & \ddots & \\ & & 2a_n\end{bmatrix}$ 是正定矩阵，因此 $f$ 在由 (13.36) 所给出的诸 $x_j$ 处取到条件极小值。如果记 $c=\sum\limits_{k=1}^{n}\dfrac{1}{a_k}$，那么该极小值为

$$f\Bigl(\frac{1}{a_1c},\cdots,\frac{1}{a_nc}\Bigr)=\sum_{j=1}^{n} a_j\Bigl(\frac{1}{a_jc}\Bigr)^2=\frac{1}{c^2}\sum_{j=1}^{n}\frac{1}{a_j}=\frac{1}{c}.$$

□

**【例 7.12】** 求函数 $f(x,y,z)=x+2y+4z$ 在条件 $xyz=1$ 下的极值。

<u>解</u>. 记 $g(x,y,z)=xyz-1$，并作辅助函数

$$F(x,y,z)=f(x,y,z)+\lambda g(x,y,z)=x+2y+4z+\lambda(xyz-1).$$

解方程组

$$\begin{cases}\dfrac{\partial f}{\partial x}=1+\lambda yz=0,\\ \dfrac{\partial f}{\partial y}=2+\lambda zx=0,\\ \dfrac{\partial f}{\partial z}=4+\lambda xy=0,\\ xyz=1\end{cases}$$

可得

$$x = 2, \qquad y = 1, \qquad z = \frac{1}{2}, \qquad \lambda = -2.$$

现记 $\boldsymbol{a} = \left(2, 1, \dfrac{1}{2}\right)^{\mathrm{T}}$，那么由

$$H_F = \begin{bmatrix} 0 & \lambda z & \lambda y \\ \lambda z & 0 & \lambda x \\ \lambda y & \lambda x & 0 \end{bmatrix}$$

知

$$H_F(\boldsymbol{a}) = \begin{bmatrix} 0 & -1 & -2 \\ -1 & 0 & -4 \\ -2 & -4 & 0 \end{bmatrix},$$

这是个不定矩阵，故而无法直接使用推论 7.9 来判定 $\boldsymbol{a}$ 是否为条件极值点。

因为 $g'(\boldsymbol{a}) = \left(\dfrac{1}{2}, 1, 2\right)$，所以当非零元 $\boldsymbol{h} = (h_1, h_2, h_3)^{\mathrm{T}}$ 满足 $g'(\boldsymbol{a})\boldsymbol{h} = 0$ 时，我们有 $h_2 = -\dfrac{1}{2}(h_1 + 4h_3)$，对于这样的 $\boldsymbol{h}$，

$$\begin{aligned} \boldsymbol{h}^{\mathrm{T}} H_F(\boldsymbol{a})\boldsymbol{h} &= -2h_1h_2 - 4h_1h_3 - 8h_2h_3 \\ &= h_1(h_1 + 4h_3) - 4h_1h_3 + 4h_3(h_1 + 4h_3) \\ &= h_1^2 + 4h_1h_3 + 16h_3^2 = (h_1 + 2h_3)^2 + 12h_3^2 > 0, \end{aligned}$$

从而由命题 7.10 知 $\boldsymbol{a}$ 是函数 $f$ 在条件 $xyz = 1$ 下的极小值点，并且条件极小值为 $f(\boldsymbol{a}) = 6$。 □

## 习 题 13.7

1. 求下列函数的极值：

   (1) $f(x, y) = x^2 - (y-1)^2$；

   (2) $f(x, y) = x^4 + y^4 - x^2 - 2xy - y^2$；

   (3) $f(x, y) = xy\log(x^2 + y^2)$；

   (4) $f(x, y) = x + \dfrac{y^2}{4x} + \dfrac{z^2}{y} + \dfrac{2}{z}$ $(x,\, y,\, z > 0)$。

2. 求 $f(x, y) = x^2 + y^2$ 在条件 $\dfrac{x}{a} + \dfrac{y}{b} = 1$ 下的极值。

3. 求半径为 $a$ 的球内具有最大体积的内接长方体。

4. （Rolle 定理）设 $K$ 是 $\mathbb{R}^n$ 中的一个紧集，$f$ 是定义在 $K$ 上的一个连续函数，且在 $K^\circ$ 上可微，在 $\partial K$ 上取常值。证明存在 $\boldsymbol{a} \in K^\circ$ 使得 $f'(\boldsymbol{a}) = \boldsymbol{0}$。

5. 求 $f(x,y) = \dfrac{1}{2}(x^n + y^n)$ 在条件 $x + y = a\ (x, y > 0)$ 下的极值，并由此证明 $\left(\dfrac{x+y}{2}\right)^n \leqslant \dfrac{x^n + y^n}{2}$。

6. （Hölder 不等式）设 $x_i$, $y_i$ $(1 \leqslant i \leqslant n)$ 以及 $p$，$q$ 均是正实数，满足 $\dfrac{1}{p} + \dfrac{1}{q} = 1$。证明
$$\sum_{i=1}^n x_i y_i \leqslant \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}}.$$

7. 设 $A$ 是 $n$ 阶对称矩阵，求 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ 在条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1$ 下的最大值和最小值。

8. 设 $f(x,y) = x^3 - 4x^2 + 2xy - y^2$，证明 $f(x,y)$ 在 $D = [-1,4] \times [-1,1]$ 的内部有唯一的极值点，但该极值点并非 $f$ 在 $D$ 上的最值点[9]。

---

[9] 注意把此题与 §6.1 习题 12 作比较。

第十四章

# 含参变量的积分与广义积分

皇家天文学家 Airy 先生指出由波动理论可以导出一个关于光照的公式，这一公式中包含了定积分 $\int_0^\infty \cos\frac{\pi}{2}(w^3-mw)\,\mathrm{d}w$ 的平方，……

经过多次尝试，我终于成功地给出了 Airy 先生的积分的一个表达式，通过它可以在 $m$ 较大时极其方便地对这一积分进行数值计算，……此外，该表达式还在无需数值计算的前提下指出了 $m$ 较大时这一函数的变化规律。

—— G. G. Stokes

## §14.1 含参变量的积分

假设 $f(x,y)$ 是定义在矩形 $[a,b]\times[c,d]$ 上的二元函数，如果对任意给定的 $y\in[c,d]$，$f(x,y)$ 作为 $x$ 的函数均在 $[a,b]$ 上可积，那么就可以通过

$$y\longmapsto\int_a^b f(x,y)\,\mathrm{d}x$$

定义区间 $[c,d]$ 上的一个函数，我们记作

$$I(y)=\int_a^b f(x,y)\,\mathrm{d}x. \tag{14.1}$$

更一般地，我们还可以讨论积分上、下限均是 $y$ 的函数的情况。具体来说，如果 $g$ 和 $h$ 均是从 $[c,d]$ 到 $[a,b]$ 的函数，并且对任意给定的 $y\in[c,d]$，$f(x,y)$ 均在区间 $[g(y),h(y)]$ 上可积，那么就可定义区间 $[c,d]$ 上的一个函数

$$J(y)=\int_{g(y)}^{h(y)} f(x,y)\,\mathrm{d}x. \tag{14.2}$$

我们把 (14.1) 和 (14.2) 右侧的积分均称为含参变量 $y$ 的积分。

对于含参变量积分而言，一个基本任务就是去讨论它关于参变量 $y$ 的连续性、可微性及可积性。

**【命题 1.1】设 $f(x,y)$ 在 $[a,b]\times[c,d]$ 上连续，函数 $g:[c,d]\longrightarrow[a,b]$ 与 $h:[c,d]\longrightarrow[a,b]$ 均在 $[c,d]$ 上连续。那么，由 (14.2) 所定义的 $J(y)$ 在 $[c,d]$ 上连续。特别地，由 (14.1) 所定义的 $I(y)$ 在 $[c,d]$ 上连续。**

证明. 只需对任意的 $y_0\in[c,d]$ 证明 $J(y)$ 在 $y_0$ 处连续即可。

由 $f(x,y)$ 在 $[a,b]\times[c,d]$ 上连续知该矩形上有界，即存在 $M>0$，使得对任意的 $(x,y)\in[a,b]\times[c,d]$ 有 $|f(x,y)|<M$。同时，由连续性还可得到 $f(x,y)$ 在该矩形上的一致连续性，于是对任意的 $\varepsilon>0$，存在 $\delta>0$，使得对于矩形内任意两点 $(x_1,y_1)$ 与 $(x_2,y_2)$，只要满足 $|x_1-x_2|<\delta$ 及 $|y_1-y_2|<\delta$ 就有

$$|f(x_1,y_1)-f(x_2,y_2)|<\varepsilon.$$

此外，因为 $g$ 与 $h$ 均在 $[c,d]$ 上连续，故不妨假设上述 $\delta$ 还使得对任意的 $y\in(y_0-\delta,y_0+\delta)\cap[c,d]$ 有

$$|g(y)-g(y_0)|<\varepsilon \qquad \text{及} \qquad |h(y)-h(y_0)|<\varepsilon.$$

现在将 $J(y)$ 写成如下形式

$$J(y)=\int_{g(y)}^{g(y_0)}f(x,y)\,\mathrm{d}x+\int_{g(y_0)}^{h(y_0)}f(x,y)\,\mathrm{d}x+\int_{h(y_0)}^{h(y)}f(x,y)\,\mathrm{d}x,$$

于是

$$J(y)-J(y_0)=\int_{g(y)}^{g(y_0)}f(x,y)\,\mathrm{d}x+\int_{g(y_0)}^{h(y_0)}[f(x,y)-f(x,y_0)]\,\mathrm{d}x+\int_{h(y_0)}^{h(y)}f(x,y)\,\mathrm{d}x,$$

从而当 $y\in(y_0-\delta,y_0+\delta)\cap[c,d]$ 时有

$$\begin{aligned}|J(y)-J(y_0)|&\leqslant\left|\int_{g(y)}^{g(y_0)}|f(x,y)|\,\mathrm{d}x\right|+\int_a^b|f(x,y)-f(x,y_0)|\,\mathrm{d}x\\&\quad+\left|\int_{h(y_0)}^{h(y)}|f(x,y)|\,\mathrm{d}x\right|\\&\leqslant M|g(y)-g(y_0)|+(b-a)\varepsilon+M|h(y)-h(y_0)|\\&<(b-a+2M)\varepsilon.\end{aligned}$$

这就证明了 $J(y)$ 在 $y_0$ 处连续。 □

【注 1.2】$I(y)$ 在 $[c,d]$ 上的连续性意味着对任意的 $y_0\in[c,d]$ 有

$$\lim_{y\to y_0}\int_a^b f(x,y)\,\mathrm{d}x=\int_a^b\Big(\lim_{y\to y_0}f(x,y)\Big)\,\mathrm{d}x.$$

也即是说，极限号与积分号可以交换。

【命题 1.3】设 $f(x,y)$ 与 $\dfrac{\partial f}{\partial y}$ 均在矩形 $[a,b]\times[c,d]$ 上连续，又设函数 $g:[c,d]\longrightarrow[a,b]$ 与 $h:[c,d]\longrightarrow[a,b]$ 均在 $[c,d]$ 上可导。那么，由 (14.2) 所定义的 $J(y)$ 在 $[c,d]$ 上可导，且

$$J'(y)=f(h(y),y)h'(y)-f(g(y),y)g'(y)+\int_{g(y)}^{h(y)}\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

特别地，

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^b f(x,y)\,\mathrm{d}x=\int_a^b\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x,$$

也即是说，微分运算与积分运算可以交换。

证明. 对任意给定的 $y_0\in[c,d]$，我们将 $J(y)$ 写成如下形式

$$\begin{aligned}J(y)&=\int_{g(y)}^{g(y_0)}f(x,y)\,\mathrm{d}x+\int_{g(y_0)}^{h(y_0)}f(x,y)\,\mathrm{d}x+\int_{h(y_0)}^{h(y)}f(x,y)\,\mathrm{d}x\\&=J_1(y)+J_2(y)+J_3(y).\end{aligned}$$

下面来证明 $J_j$ 在 $y_0$ 处可导。

首先，由 $J_3(y_0)=0$ 知

$$\frac{J_3(y)-J_3(y_0)}{y-y_0}=\frac{J_3(y)}{y-y_0}=\frac{1}{y-y_0}\int_{h(y_0)}^{h(y)}f(x,y)\,\mathrm{d}x,$$

按照积分第一中值定理（第八章定理 4.6），存在位于 $h(y_0)$ 和 $h(y)$ 之间的 $\xi$ 使得

$$\frac{J_3(y)-J_3(y_0)}{y-y_0}=f(\xi,y)\cdot\frac{h(y)-h(y_0)}{y-y_0},$$

于是由 $h$ 的可微性及 $f$ 的连续性知

$$\lim_{y\to y_0}\frac{J_3(y)-J_3(y_0)}{y-y_0}=f(h(y_0),y_0)h'(y_0),$$

这就证明了 $J_3$ 在 $y_0$ 处可导且

$$J_3'(y_0)=f(h(y_0),y_0)h'(y_0).$$

同理可证 $J_1$ 在 $y_0$ 处可导且

$$J_1'(y_0) = -f(g(y_0), y_0)g'(y_0).$$

再来看 $J_2$，我们有

$$\frac{J_2(y) - J_2(y_0)}{y - y_0} = \frac{1}{y - y_0} \int_{g(y_0)}^{h(y_0)} [f(x, y) - f(x, y_0)] \, \mathrm{d}x.$$

由 Lagrange 中值定理知，存在位于 $y_0$ 与 $y$ 之间的 $\eta$ 使得

$$f(x, y) - f(x, y_0) = (y - y_0) \cdot \frac{\partial f}{\partial y}(x, \eta),$$

从而

$$\frac{J_2(y) - J_2(y_0)}{y - y_0} = \int_{g(y_0)}^{h(y_0)} \frac{\partial f}{\partial y}(x, \eta) \, \mathrm{d}x.$$

注意到 $\dfrac{\partial f}{\partial y}$ 的连续性，故由命题 1.1 知

$$\lim_{y \to y_0} \frac{J_2(y) - J_2(y_0)}{y - y_0} = \int_{g(y_0)}^{h(y_0)} \left( \lim_{y \to y_0} \frac{\partial f}{\partial y}(x, \eta) \right) \mathrm{d}x = \int_{g(y_0)}^{h(y_0)} \frac{\partial f}{\partial y}(x, y_0) \, \mathrm{d}x.$$

因此 $J_2$ 在 $y_0$ 处可导且

$$J_2'(y_0) = \int_{g(y_0)}^{h(y_0)} \frac{\partial f}{\partial y}(x, y_0) \, \mathrm{d}x.$$

综上，命题获证。 □

**【例 1.4】**

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}y} \left( \int_1^{y^2} \frac{\sin(xy)}{x} \, \mathrm{d}x \right) &= \frac{2 \sin y^3}{y} + \int_1^{y^2} \cos(xy) \, \mathrm{d}x \\ &= \frac{2 \sin y^3}{y} + \frac{\sin y^3 - \sin y}{y} = \frac{3 \sin y^3 - \sin y}{y}. \end{aligned}$$

最后我们来考虑含参变量积分的可积性问题。设 $f(x, y)$ 在 $[a, b] \times [c, d]$ 上连续，那么由命题 1.1 知

$$I(y) = \int_a^b f(x, y) \, \mathrm{d}x$$

在区间 $[c, d]$ 上连续，从而在 $[c, d]$ 上可积，因此我们可以得到积分

$$\int_c^d \left( \int_a^b f(x, y) \, \mathrm{d}x \right) \mathrm{d}y,$$

上式通常被记作

$$\int_c^d \mathrm{d}y \int_a^b f(x,y)\,\mathrm{d}x.$$

同理还可得到积分

$$\int_a^b \mathrm{d}x \int_c^d f(x,y)\,\mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y\right)\mathrm{d}x.$$

我们需要去了解上面两式中的积分是否相等。

**【命题 1.5】**如果 $f(x,y)$ 在 $[a,b]\times[c,d]$ 上连续，那么

$$\int_c^d \mathrm{d}y \int_a^b f(x,y)\,\mathrm{d}x = \int_a^b \mathrm{d}x \int_c^d f(x,y)\,\mathrm{d}y.$$

证明. 这实质上是重积分理论中的 Fubini 定理的一个简单推论（参见第十五章推论 4.2），但我们也可用下述方式来证明。

记

$$G_1(u) = \int_c^u \mathrm{d}y \int_a^b f(x,y)\,\mathrm{d}x, \qquad G_2(u) = \int_a^b \mathrm{d}x \int_c^u f(x,y)\,\mathrm{d}y.$$

我们来证 $G_1'(u) = G_2'(u)$。一方面，由命题 1.1 知 $I(y) = \displaystyle\int_a^b f(x,y)\,\mathrm{d}x$ 在 $[c,d]$ 上连续，于是

$$G_1'(u) = \frac{\mathrm{d}}{\mathrm{d}u}\left(\int_c^u I(y)\,\mathrm{d}y\right) = I(u) = \int_a^b f(x,u)\,\mathrm{d}x.$$

另一方面，若记 $F(x,u) = \displaystyle\int_c^u f(x,y)\,\mathrm{d}y$，那么由 $\dfrac{\partial F}{\partial u} = f(x,u)$ 知它在 $[a,b]\times[c,d]$ 上连续，于是由命题 1.3 可得

$$G_2'(u) = \frac{\mathrm{d}}{\mathrm{d}u}\left(\int_a^b F(x,u)\,\mathrm{d}x\right) = \int_a^b \frac{\partial F}{\partial u}(x,u)\,\mathrm{d}x = \int_a^b f(x,u)\,\mathrm{d}x.$$

这样就证明了 $G_1'(u) = G_2'(u)$。因而存在常数 $C$ 使得 $G_1(u) = G_2(u) + C$。注意到

$$G_1(c) = 0 = G_2(c),$$

所以 $C=0$，从而 $G_1(u) = G_2(u)$，再取 $u=d$ 便得命题结论。 □

**【例 1.6】**设 $b>a>0$，计算 $I=\displaystyle\int_0^1 \frac{x^b-x^a}{\log x}\,\mathrm{d}x$。

解. 容易看出

$$I=\int_0^1 \mathrm{d}x\int_a^b x^y\,\mathrm{d}y,$$

因此由命题 1.5 知

$$I=\int_a^b \mathrm{d}y\int_0^1 x^y\,\mathrm{d}x=\int_a^b \frac{\mathrm{d}y}{y+1}=\log\frac{b+1}{a+1}.$$

□

## 习 题 14.1

1. 计算下列极限：

(1) $\displaystyle\lim_{n\to\infty}\int_0^1 \frac{\mathrm{d}x}{1+\left(1+\frac{x}{n}\right)^n}$；

(2) $\displaystyle\lim_{n\to\infty} n\int_0^1 \log\left(1+\frac{\sin x}{n}\right)\mathrm{d}x$；

(3) $\displaystyle\lim_{y\to 0}\int_{-1}^1 \sqrt{x^2+y^2}\,\mathrm{d}x$；

(4) $\displaystyle\lim_{y\to 0}\int_y^{1+y} \frac{\mathrm{d}x}{1+x^2+y^2}$。

2. 证明椭圆积分

$$E(k)=\int_0^{\frac{\pi}{2}} \sqrt{1-k^2\sin^2\varphi}\,\mathrm{d}\varphi,\qquad 0<k<1$$

满足微分方程 $E''(k)+\dfrac{1}{k}E'(k)+\dfrac{1}{1-k^2}E(k)=0$。

3. 设 $f\in C([a,b])$，证明定义在 $[a,b]$ 上的函数

$$y(x)=\frac{1}{k}\int_a^x f(t)\sin k(x-t)\,\mathrm{d}t$$

满足微分方程 $y''+k^2y=f(x)$。

4. 设 $n\in\mathbb{Z}$，证明 Bessel 函数

$$J_n(x)=\frac{1}{\pi}\int_0^\pi \cos(n\varphi-x\sin\varphi)\,\mathrm{d}\varphi$$

满足 Bessel 方程 $x^2J_n''(x)+xJ_n'(x)+(x^2-n^2)J_n(x)=0$。

5. 设 $\alpha\in(-1,1)$，计算积分

$$I(\alpha)=\int_0^\pi \log(1+\alpha\cos x)\,\mathrm{d}x.$$

6. 设 $|\alpha| \neq 1$，证明

$$\int_0^{\pi} \log(1 - 2\alpha\cos x + \alpha^2)\,\mathrm{d}x = \begin{cases} 0, & \text{若 } |\alpha| < 1, \\ 2\pi\log|\alpha|, & \text{若 } |\alpha| > 1. \end{cases}$$

7. 设 $|\alpha| < 1$，计算

$$I(\alpha) = \int_0^{\frac{\pi}{2}} \log\frac{1 + \alpha\cos x}{1 - \alpha\cos x} \cdot \frac{1}{\cos x}\,\mathrm{d}x.$$

## § 14.2
## 含参变量的广义积分

与广义积分类似，我们通常遇到的含参变量的广义积分有两种情形：一种是无界区间上的积分，如 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$；另一种是有界区间上的无界函数的积分，例如 $\displaystyle\int_a^b f(x,y)\,\mathrm{d}x$，其中 $a$ 为奇点，且对于某些给定的 $y$ 而言函数 $f(x,y)$ 在 $(a,b)$ 的任一闭子区间上对于变量 $x$ 可积。无论是哪种情况，我们都把 $y$ 称为参变量。当然，我们还可以进一步讨论含多个参变量的情形，但其研究方法与只含一个参变量的类似，因此在本节中我们仅讨论含一个参变量的广义积分。

### 14.2.1　一致收敛

与函数项级数的情形类似，在研究含参变量的广义积分时一个极其重要的概念便是一致收敛性。事实上我们可以在更大的范围内去讨论这一问题，为此，我们引入二元函数关于某个变量一致收敛的概念。

**【定义 2.1】**假设 $X$ 和 $Y$ 是 $\mathbb{R}$ 的两个子集，$f(x,y)$ 是定义在 $X \times Y$ 上的函数，$\varphi(y)$ 是定义在 $Y$ 上的一个函数，$x_0$ 是 $X$ 的一个聚点。若对任意的 $\varepsilon > 0$，存在至多依赖于 $\varepsilon$ 的正实数 $\delta$，使得对任意的 $x \in \Big((x_0 - \delta, x_0 + \delta) \setminus \{x_0\}\Big) \cap X$ 及任意的 $y \in Y$ 均有

$$|f(x,y) - \varphi(y)| < \varepsilon,$$

则称当 $x \to x_0$ 时 $f(x,y)$ 在 $Y$ 上<u>一致收敛</u>于 $\varphi(y)$。

$\varphi(y)$ 被称为当 $x$ 沿 $X$ 中元素趋于 $x_0$ 时 $f(x,y)$ 的<u>极限函数</u>，按照 §12.2 中的记号我们记

$$\varphi(y) = \lim_{\substack{x \to x_0 \\ x \in X}} f(x,y).$$

类似可定义 $x \to +\infty$ 和 $x \to -\infty$ 时 $f(x,y)$ 一致收敛的概念，例如，当 $\sup X = +\infty$ 时，若对任意的 $\varepsilon > 0$，存在至多依赖于 $\varepsilon$ 的正实数 $\rho$，使得对任意的 $x \in (\rho, +\infty) \cap X$ 及 $y \in Y$ 均有

$$|f(x,y) - \varphi(y)| < \varepsilon,$$

则称当 $x \to +\infty$ 时 $f(x,y)$ 在 $Y$ 上一致收敛于 $\varphi(y)$。

例如令 $X = \mathbb{Z}_{\geqslant 1}$，那么对每个给定的 $n \in X$，$f(n,y)$ 是关于 $y$ 的函数，因此上面给出的当 $n \to \infty$ 时 $f(n,y)$ 在 $Y$ 上一致收敛的定义也即是第十章定义 2.1 中所给出的函数列一致收敛的定义，这说明函数列和函数项级数的一致收敛性均是上述二元函数一致收敛性的特殊情形，所以第十章中的一些结果也是本章中即将提到的更一般的结果的特殊情况。在本节的学习中，一定要注意这两者之间的比较。

接下来引入含参变量的广义积分一致收敛性的概念。

**【定义 2.2】** 设 $E \subseteq \mathbb{R}$。若广义积分 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 对任意的 $y \in E$ 均收敛，并且对任意的 $\varepsilon > 0$，存在至多依赖于 $\varepsilon$ 的实数 $A > a$，使得对任意的 $A' > A$ 及 $y \in E$ 均有

$$\left| \int_{A'}^{+\infty} f(x,y)\,\mathrm{d}x \right| < \varepsilon,$$

则称 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。

**【注 2.3】** 当广义积分 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 对任意的 $y \in E$ 均收敛时，我们就可通过

$$y \longmapsto \int_a^{+\infty} f(x,y)\,\mathrm{d}x$$

定义 $E$ 上的一个函数 $F$。现记

$$F(A,y) = \int_a^{A} f(x,y)\,\mathrm{d}x,$$

那么

$$\int_{A'}^{+\infty} f(x,y)\,\mathrm{d}x = F(y) - F(A',y).$$

因此 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛等价于当 $A \to +\infty$ 时 $F(A,y)$ 在 $E$ 上一致收敛于 $F(y)$。

【例 2.4】对任意的 $y>0$ 及 $A>1$ 有

$$\int_A^{+\infty}\frac{\mathrm{d}x}{1+(xy)^2}=\frac{1}{y}\Big(\frac{\pi}{2}-\arctan Ay\Big).$$

因此取 $y=\dfrac{1}{A}$ 即可得上式右边等于 $\dfrac{\pi}{4}A$，从而由定义知 $\displaystyle\int_1^{+\infty}\frac{\mathrm{d}x}{1+(xy)^2}$ 在 $(0,+\infty)$ 上不一致收敛。但对任意给定的 $\delta>0$，当 $y\in[\delta,+\infty)$ 时有

$$\left|\int_A^{+\infty}\frac{\mathrm{d}x}{1+(xy)^2}\right|\leqslant\frac{1}{\delta}\Big(\frac{\pi}{2}-\arctan A\delta\Big)\longrightarrow 0\qquad（当 A\to+\infty 时），$$

这说明 $\displaystyle\int_1^{+\infty}\frac{\mathrm{d}x}{1+(xy)^2}$ 在 $[\delta,+\infty)$ 上一致收敛。

【定义 2.5】设 $E\subseteq\mathbb{R}$，且广义积分 $\displaystyle\int_a^b f(x,y)\,\mathrm{d}x$ 对任意的 $y\in E$ 均收敛，其中 $a$ 是积分区间上仅有的奇点。如果对任意的 $\varepsilon>0$，均存在至多依赖于 $\varepsilon$ 的正数 $\delta$，使得对任意的 $\delta'\in(0,\delta)$ 及任意的 $y\in E$ 均有

$$\left|\int_a^{a+\delta'}f(x,y)\,\mathrm{d}x\right|<\varepsilon,$$

则称 $\displaystyle\int_a^b f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。

类似可定义奇点在积分区间右端点或内部的含参变量广义积分的一致收敛性，此外，还可仿效函数项级数给出内闭一致收敛的概念，读者可自行将它们写出，我们不再一一赘述。

下面来介绍一些一致收敛的判别法，为了明确目标，我们仅介绍与含参变量的广义积分相关的判别法，而把一些与一般二元函数相关的结论放到习题中。此外，为了节约篇幅，在这里我们只讨论无界区间上的积分，读者可以毫不费力地对有界区间上的无界函数的积分给出相应结论。

【定理 2.6】(**Cauchy 收敛准则**) $\displaystyle\int_a^{+\infty}f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛的充要条件是：对任意的 $\varepsilon>0$，均存在至多依赖于 $\varepsilon$ 的实数 $A>a$，使得对任意的 $A',A''>A$ 及任意的 $y\in E$ 有

$$\left|\int_{A'}^{A''}f(x,y)\,\mathrm{d}x\right|<\varepsilon.\tag{14.3}$$

证明. 必要性可直接由定义得出，下证充分性。首先，由条件及第十一章定理 1.5 知 $\displaystyle\int_a^{+\infty}f(x,y)\,\mathrm{d}x$ 在 $E$ 上逐点收敛。进而在 (14.3) 式中令 $A''\to+\infty$ 即得 $\displaystyle\int_a^{+\infty}f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。 □

**【定理 2.7】（Weierstrass 判别法）** 假设对任意的 $x \in [a,+\infty)$ 及任意的 $y \in E$ 有 $|f(x,y)| \leqslant F(x)$，并且广义积分 $\displaystyle\int_a^{+\infty} F(x)\,\mathrm{d}x$ 收敛，那么 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。我们称 $F(x)$ 为 $f(x,y)$ 的优势函数。

证明. 由 $\displaystyle\int_a^{+\infty} F(x)\,\mathrm{d}x$ 收敛知，对任意的 $\varepsilon > 0$，存在 $A > a$，使得对任意的 $A', A'' > A$ 有

$$\left|\int_{A'}^{A''} F(x)\,\mathrm{d}x\right| < \varepsilon.$$

于是对任意的 $y \in E$ 可得

$$\left|\int_{A'}^{A''} f(x,y)\,\mathrm{d}x\right| \leqslant \left|\int_{A'}^{A''} F(x)\,\mathrm{d}x\right| < \varepsilon,$$

从而由 Cauchy 收敛准则知命题成立。 □

**【例 2.8】** 对任意的 $y \geqslant 0$ 有 $\left|\dfrac{\sin xy}{x^2+y}\right| \leqslant \dfrac{1}{x^2}$，故 $\displaystyle\int_1^{+\infty} \frac{\sin xy}{x^2+y}\,\mathrm{d}x$ 在 $[0,+\infty)$ 上一致收敛。

**【定理 2.9】（Abel 判别法）** 假设

(1) $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛；

(2) 对任意给定的 $y \in E$，$g(x,y)$ 是关于 $x$ 的单调函数。并且 $g(x,y)$ 在 $[a,+\infty)\times E$ 上有界，也即存在 $M > 0$，使得对任意的 $x \geqslant a$ 及 $y \in E$ 有 $|g(x,y)| \leqslant M$。

那么 $\displaystyle\int_a^{+\infty} f(x,y)g(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。

证明. 由 (1) 知，对任意的 $\varepsilon > 0$，存在 $A > a$，使得对任意的 $A', A'' > A$ 及 $y \in E$ 有

$$\left|\int_{A'}^{A''} f(x,y)\,\mathrm{d}x\right| < \varepsilon.$$

于是由积分第二中值定理知，对任意的 $y \in E$，存在位于 $A'$ 和 $A''$ 之间的 $\xi(y)$ 使得

$$\begin{aligned}\left|\int_{A'}^{A''} f(x,y)g(x,y)\,\mathrm{d}x\right| &= \left|g(A',y)\int_{A'}^{\xi(y)} f(x,y)\,\mathrm{d}x + g(A'',y)\int_{\xi(y)}^{A''} f(x,y)\,\mathrm{d}x\right| \\ &< 2M\varepsilon,\end{aligned}$$

因此 $\displaystyle\int_a^{+\infty} f(x,y)g(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。 □

**【定理 2.10】（Dirichlet 判别法）** 假设

(1) 对任意给定的 $y \in E$，$g(x,y)$ 是关于 $x$ 的单调函数。并且当 $x \to +\infty$ 时，函数 $g(x,y)$ 在 $E$ 上一致收敛于 0;

(2) $F(A,y) = \int_a^A f(x,y)\,\mathrm{d}x$ 在 $[a,+\infty) \times E$ 上有界，也即存在 $M > 0$，使得对任意的 $A \geqslant a$ 及 $y \in E$ 有 $|F(A,y)| \leqslant M$。

那么 $\int_a^{+\infty} f(x,y)g(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。

证明. 由 (1) 知，对任意的 $\varepsilon > 0$，存在 $A > a$，使得对任意的 $x > A$ 及 $y \in E$ 有 $|g(x,y)| < \varepsilon$。于是由积分第二中值定理知，对任意的 $A', A'' > A$ 及任意的 $y \in E$，存在位于 $A'$ 和 $A''$ 之间的 $\xi(y)$ 使得

$$\begin{aligned}
\left|\int_{A'}^{A''} f(x,y)g(x,y)\,\mathrm{d}x\right| &= \left|g(A',y)\int_{A'}^{\xi(y)} f(x,y)\,\mathrm{d}x + g(A'',y)\int_{\xi(y)}^{A''} f(x,y)\,\mathrm{d}x\right| \\
&= \Big|g(A',y)\big[F(\xi(y),y) - F(A',y)\big] + g(A'',y)\big[F(A'',y) - F(\xi(y),y)\big]\Big| \\
&< 4M\varepsilon,
\end{aligned}$$

因此 $\int_a^{+\infty} f(x,y)g(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。 □

**【例 2.11】** 证明 $\int_0^{+\infty} e^{-xy}\dfrac{\sin x}{x}\,\mathrm{d}x$ 在 $[0,+\infty)$ 上一致收敛。

证明. 由于 $\int_0^{+\infty} \dfrac{\sin x}{x}\,\mathrm{d}x$ 收敛，而对任意给定的 $y \geqslant 0$ 函数 $e^{-xy}$ 是关于 $x$ 的单调函数，并且

$$0 \leqslant e^{-xy} \leqslant 1, \qquad \forall\, x \geqslant 0,\ y \geqslant 0.$$

因此由 Abel 判别法知命题成立。 □

### 14.2.2 含参变量广义积分的性质

**【命题 2.12】** 设 $f(x,y)$ 是定义在 $[a,b] \times E$ 上的一个二元函数，$y_0$ 是 $E$ 的聚点（当 $\sup E = +\infty$ 时，$y_0$ 可以是 $+\infty$；当 $\inf E = -\infty$ 时，$y_0$ 可以是 $-\infty$），满足

(1) 对任意给定的 $y \in E$，$f(x,y)$ 关于变量 $x$ 是 $[a,b]$ 上的连续函数；

(2) 当 $y \to y_0$ 时 $f(x,y)$ 在 $[a,b]$ 上一致收敛于 $\varphi(x)$。

那么 $\varphi(x)$ 在 $[a,b]$ 上连续，且有

$$\lim_{\substack{y\to y_0\\ y\in E}}\int_a^b f(x,y)\,\mathrm{d}x=\int_a^b \varphi(x)\,\mathrm{d}x.\text{ ①} \tag{14.4}$$

<u>证明</u>. 第一部分的证明过程与第十章命题 3.1 的类似。我们仅讨论 $y_0\in\mathbb{R}$ 的情形。

首先来证明 $\varphi\in C([a,b])$。为此，只需对任意的 $x_0\in[a,b]$ 证明 $\varphi(x)$ 在 $x_0$ 处连续即可。由条件 (2) 知，对任意的 $\varepsilon>0$，均存在 $\delta>0$，使得对于任意的 $y\in\Big((y_0-\delta,y_0+\delta)\setminus\{y_0\}\Big)\cap E$ 及任意的 $x\in[a,b]$ 有

$$|f(x,y)-\varphi(x)|<\frac{\varepsilon}{3}. \tag{14.5}$$

现取定 $y_1\in\Big((y_0-\delta,y_0+\delta)\setminus\{y_0\}\Big)\cap E$，于是由 $f(x,y_1)$ 关于变量 $x$ 的连续性知，存在 $\eta>0$，使得当 $x\in(x_0-\eta,x_0+\eta)\cap[a,b]$ 时

$$|f(x,y_1)-f(x_0,y_1)|<\frac{\varepsilon}{3}.$$

进而可对任意的 $x\in(x_0-\eta,x_0+\eta)\cap[a,b]$ 得到

$$\begin{aligned}|\varphi(x)-\varphi(x_0)|&\leqslant|\varphi(x)-f(x,y_1)|+|f(x,y_1)-f(x_0,y_1)|+|f(x_0,y_1)-\varphi(x_0)|\\&<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

这就证明了 $\varphi(x)$ 在 $x_0$ 处连续。

此外，由 (14.5) 还可对任意的 $y\in\Big((y_0-\delta,y_0+\delta)\setminus\{y_0\}\Big)\cap E$ 得到

$$\left|\int_a^b f(x,y)\,\mathrm{d}x-\int_a^b\varphi(x)\,\mathrm{d}x\right|\leqslant\int_a^b|f(x,y)-\varphi(x)|\,\mathrm{d}x<\frac{b-a}{3}\varepsilon.$$

因此 (14.4) 成立。 □

**【注 2.13】**在上述命题取 $E=\mathbb{Z}_{\geqslant 1}$，并考虑极限过程 $y\to+\infty$，则可得到第十章命题 3.2 和 3.6。

下面给出命题 2.12 对应于广义积分的形式。

---

① 为了使得这个式子成立，事实上只需对任意给定的 $y\in E$ 要求 $f(x,y)$ 关于变量 $x$ 在 $[a,b]$ 上可积即可，参见习题 5。

**【命题 2.14】**设 $f(x,y)$ 是定义在 $[a,+\infty)\times E$ 上的一个二元函数，$y_0$ 是 $E$ 的聚点（当 $\sup E=+\infty$ 时，$y_0$ 可以是 $+\infty$；当 $\inf E=-\infty$ 时，$y_0$ 可以是 $-\infty$），满足

(1) 对任意给定的 $y\in E$，$f(x,y)$ 关于变量 $x$ 是 $[a,+\infty)$ 上的连续函数；

(2) 对任意的 $b>a$，当 $y\to y_0$ 时 $f(x,y)$ 在 $[a,b]$ 上一致收敛于 $\varphi(x)$；

(3) $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $E$ 上一致收敛。

那么

$$\lim_{\substack{y\to y_0\\ y\in E}}\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\varphi(x)\,\mathrm{d}x. \tag{14.6}$$

证明. 先证 $\displaystyle\int_a^{+\infty}\varphi(x)\,\mathrm{d}x$ 收敛。由条件 (3) 知，对任意的 $\varepsilon>0$，存在 $A>a$，使得对任意的 $A'$, $A''>A$ 及任意的 $y\in E$ 有

$$\left|\int_{A'}^{A''} f(x,y)\,\mathrm{d}x\right|<\varepsilon.$$

注意到由条件 (2) 知，当 $y\to y_0$ 时 $f(x,y)$ 在 $[A',A'']$ 上一致收敛于 $\varphi(x)$，于是在上式中令 $y\to y_0$ 便可由命题 2.12 得到

$$\left|\int_{A'}^{A''}\varphi(x)\,\mathrm{d}x\right|\leqslant\varepsilon.$$

再由 Cauchy 收敛准则知 $\displaystyle\int_a^{+\infty}\varphi(x)\,\mathrm{d}x$ 收敛。

接下来证明 (14.6)。由 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 的一致收敛性及 $\displaystyle\int_a^{+\infty}\varphi(x)\,\mathrm{d}x$ 的收敛性知，对任意的 $\varepsilon>0$，存在 $A>a$ 使得

$$\left|\int_A^{+\infty} f(x,y)\,\mathrm{d}x\right|<\frac{\varepsilon}{3}\quad(\forall\, y\in E)\qquad\text{及}\qquad\left|\int_A^{+\infty}\varphi(x)\,\mathrm{d}x\right|<\frac{\varepsilon}{3}. \tag{14.7}$$

此外，条件 (2) 意味着存在 $\delta>0$，使得对任意的 $y\in\Big((y_0-\delta,y_0+\delta)\setminus\{y_0\}\Big)\cap E$ 及任意的 $x\in[a,A]$ 有

$$|f(x,y)-\varphi(x)|<\frac{\varepsilon}{3(A-a)},$$

进而对任意的 $y\in\Big((y_0-\delta,y_0+\delta)\setminus\{y_0\}\Big)\cap E$ 有

$$\left|\int_a^A f(x,y)\,\mathrm{d}x-\int_a^A\varphi(x)\,\mathrm{d}x\right|\leqslant\int_A^{+\infty}|f(x,y)-\varphi(x)|\,\mathrm{d}x<\frac{\varepsilon}{3}.$$

结合 (14.7) 便可对任意的 $y \in \Big((y_0-\delta, y_0+\delta) \setminus \{y_0\}\Big) \cap E$ 得到

$$\left|\int_a^{+\infty} f(x,y)\,\mathrm{d}x - \int_a^{+\infty} \varphi(x)\,\mathrm{d}x\right| < \varepsilon.$$

于是命题得证。 □

有了以上准备工作，我们可以来进一步研究含参变量广义积分的相关性质了。

**【命题 2.15】**假设

(1) $f(x,y)$ 在 $[a,+\infty)\times[b,c]$ 上连续；

(2) $F(y)=\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $[b,c]$ 上一致收敛。

那么 $F(y)$ 在 $[b,c]$ 上连续。

证明. 对任意给定的 $A>a$，由命题 1.1 知 $F(A,y)=\displaystyle\int_a^A f(x,y)\,\mathrm{d}x$ 关于变量 $y$ 在区间 $[b,c]$ 上连续。回忆起注 2.3，由条件 (2) 知当 $A\to+\infty$ 时 $F(A,y)$ 在 $[b,c]$ 上一致收敛于 $F(y)$，因此由命题 2.12 可得 $F(y)$ 的连续性。 □

类似于函数项级数的 Dini 定理，我们有下述结论。

**【定理 2.16】(Dini)** 设 $f(x,y)$ 是定义在 $[a,+\infty)\times[b,c]$ 上的非负连续函数，且对任意的 $y\in[b,c]$ 而言，$\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 收敛于 $F(y)$。如果 $F(y)$ 在 $[b,c]$ 上连续，那么 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $[b,c]$ 上一致收敛。

证明. 记 $u_n(y)=\displaystyle\int_{a+n-1}^{a+n} f(x,y)\,\mathrm{d}x$，则

$$F(y)=\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\sum_{n=1}^{\infty} u_n(y).$$

由 $f$ 所满足的条件及命题 1.1 知每个 $u_n(y)$ 均在 $[b,c]$ 上非负连续，从而由函数项级数的 Dini 定理（第十章定理 3.5）知 $\sum\limits_{n=1}^{\infty} u_n(y)$ 在 $[b,c]$ 上一致收敛于 $F(y)$。于是对任意的 $\varepsilon>0$，存在正整数 $N$，使得对任意 $y\in[b,c]$ 有

$$\sum_{n=N}^{\infty} u_n(y) < \varepsilon.$$

进而对任意的 $A>a+N-1$ 及任意的 $y\in[b,c]$ 有

$$0 \leqslant \int_A^{+\infty} f(x,y)\,\mathrm{d}x \leqslant \int_{a+N-1}^{+\infty} f(x,y)\,\mathrm{d}x = \sum_{n=N}^{\infty} u_n(y) < \varepsilon.$$

从而命题得证。 □

下面来看积分号的交换问题。

**【命题 2.17】**假设

(1) $f(x,y)$ 在 $[a,+\infty)\times[b,c]$ 上连续;

(2) $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 在 $[b,c]$ 上一致收敛。

那么

$$\int_b^c \mathrm{d}y \int_a^{+\infty} f(x,y)\,\mathrm{d}x = \int_a^{+\infty} \mathrm{d}x \int_b^c f(x,y)\,\mathrm{d}y.$$

<u>证明</u>. 由命题 1.5 知，对任意的 $A>a$ 有

$$\int_b^c \mathrm{d}y \int_a^{A} f(x,y)\,\mathrm{d}x = \int_a^{A} \mathrm{d}x \int_b^c f(x,y)\,\mathrm{d}y. \tag{14.8}$$

而在注 2.3 中我们提到过，条件 (2) 意味着当 $A\to+\infty$ 时 $F(A,y)=\displaystyle\int_a^A f(x,y)\,\mathrm{d}x$ 在 $[b,c]$ 上一致收敛于 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$。注意到对任意给定的 $y\in[b,c]$，$F(A,y)$ 是关于 $A$ 的连续函数，因此由命题 2.12 可得

$$\lim_{A\to+\infty}\int_b^c F(A,y)\,\mathrm{d}y = \int_b^c \mathrm{d}y \int_a^{+\infty} f(x,y)\,\mathrm{d}x.$$

结合 (14.8) 便知

$$\lim_{A\to+\infty}\int_a^{A} \mathrm{d}x \int_b^c f(x,y)\,\mathrm{d}y = \int_b^c \mathrm{d}y \int_a^{+\infty} f(x,y)\,\mathrm{d}x,$$

从而命题得证。 □

**【命题 2.18】**假设

(1) $f(x,y)$ 在 $[a,+\infty)\times[b,+\infty)$ 上连续;

(2) $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 和 $\displaystyle\int_b^{+\infty} f(x,y)\,\mathrm{d}y$ 分别在 $[b,+\infty)$ 及 $[a,+\infty)$ 的任一有界闭子区间上一致收敛;

(3) $\displaystyle\int_b^{+\infty} \mathrm{d}y \int_a^{+\infty} |f(x,y)|\,\mathrm{d}x$ 和 $\displaystyle\int_a^{+\infty} \mathrm{d}x \int_b^{+\infty} |f(x,y)|\,\mathrm{d}y$ 中至少有一个收敛。

那么

$$\int_b^{+\infty} \mathrm{d}y \int_a^{+\infty} f(x,y)\,\mathrm{d}x = \int_a^{+\infty} \mathrm{d}x \int_b^{+\infty} f(x,y)\,\mathrm{d}y.$$

<u>证明</u>. 不妨假设 $\int_b^{+\infty}\mathrm{d}y\int_a^{+\infty}|f(x,y)|\,\mathrm{d}x$ 收敛。由命题 2.17 知，对任意的 $A>a$ 有

$$\int_b^{+\infty}\mathrm{d}y\int_a^{A}f(x,y)\,\mathrm{d}x=\int_a^{A}\mathrm{d}x\int_b^{+\infty}f(x,y)\,\mathrm{d}y. \tag{14.9}$$

现记 $F(A,y)=\int_a^A f(x,y)\,\mathrm{d}x$，则

$$|F(A,y)|\leqslant\int_a^A|f(x,y)|\,\mathrm{d}x\leqslant\int_a^{+\infty}|f(x,y)|\,\mathrm{d}x.$$

由于预先假设了 $\int_b^{+\infty}\mathrm{d}y\int_a^{+\infty}|f(x,y)|\,\mathrm{d}x$ 收敛，所以由 Weierstrass 判别法知 $\int_b^{+\infty}F(A,y)\,\mathrm{d}y$ 在 $[a,+\infty)$ 上一致收敛。此外，条件 (2) 意味着当 $A\to+\infty$ 时 $F(A,y)$ 在 $[b,+\infty)$ 的任一有限闭子区间上一致收敛于 $\int_a^{+\infty}f(x,y)\,\mathrm{d}x$。于是在 (14.9) 中令 $A\to+\infty$，由命题 2.14 便得

$$\int_b^{+\infty}\mathrm{d}y\int_a^{+\infty}f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_b^{+\infty}f(x,y)\,\mathrm{d}y.$$

□

在上述命题中我们看到，为了保证两个无界区间上的积分可交换，其条件是颇为繁杂的。但是在 $f(x,y)$ 非负的情况下，利用 Dini 定理可将其中的条件 (2) 略微简化，这对于某些实际应用而言便捷了许多（参见例 2.24 和 2.25）。

**【推论 2.19】**假设

(1) $f(x,y)$ 在 $[a,+\infty)\times[b,+\infty)$ 上连续且非负；

(2) 函数

$$F(y)=\int_a^{+\infty}f(x,y)\,\mathrm{d}x\qquad 与\qquad G(x)=\int_b^{+\infty}f(x,y)\,\mathrm{d}y$$

分别在 $[b,+\infty)$ 和 $[a,+\infty)$ 上连续；

(3) $\int_b^{+\infty}\mathrm{d}y\int_a^{+\infty}f(x,y)\,\mathrm{d}x$ 和 $\int_a^{+\infty}\mathrm{d}x\int_b^{+\infty}f(x,y)\,\mathrm{d}y$ 中至少有一个收敛。

那么

$$\int_b^{+\infty}\mathrm{d}y\int_a^{+\infty}f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_b^{+\infty}f(x,y)\,\mathrm{d}y.$$

<u>证明</u>. 由条件 (1)，(2) 及 Dini 定理知 $\int_a^{+\infty}f(x,y)\,\mathrm{d}x$ 和 $\int_b^{+\infty}f(x,y)\,\mathrm{d}y$ 分别在 $[b,+\infty)$ 及 $[a,+\infty)$ 的任一有限闭子区间上一致收敛，从而由命题 2.18 知结论成立。 □

最后来讨论积分号下求导的问题。

**【命题 2.20】**假设

(1) $f(x,y)$ 与 $\dfrac{\partial f}{\partial y}$ 均在 $[a,+\infty)\times[b,c]$ 上连续；

(2) 对任意给定的 $y\in[b,c]$，广义积分 $\displaystyle\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ 收敛；

(3) $\displaystyle\int_a^{+\infty}\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$ 在 $[b,c]$ 上一致收敛。

那么

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x. \tag{14.10}$$

证明. 为方便起见，记

$$F(y)=\int_a^{+\infty} f(x,y)\,\mathrm{d}x,\qquad \psi(y)=\int_a^{+\infty}\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

由命题 2.15 知 $\psi(y)$ 在 $[b,c]$ 上连续，于是变上限积分 $\displaystyle\int_b^y\psi(t)\,\mathrm{d}t$ 在 $[b,c]$ 上可导，且其导数为 $\psi(y)$。此外，由命题 2.17 可得

$$\begin{aligned}\int_b^y\psi(t)\,\mathrm{d}t&=\int_b^y\mathrm{d}t\int_a^{+\infty}\frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_b^y\frac{\partial f}{\partial t}(x,t)\,\mathrm{d}t\\&=\int_a^{+\infty}\big[f(x,y)-f(x,b)\big]\,\mathrm{d}x=F(y)-F(b).\end{aligned}$$

因此 $F(y)$ 也可导且 $F'(y)=\psi(y)$，此即 (14.10)。 □

### 14.2.3 在广义积分计算中的应用

作为上一小节中的理论的应用，我们来计算一些特殊的广义积分，其中例 2.22, 2.24 及 2.25 的结果非常重要，需要牢记。

**【例 2.21】**利用第十一章例 3.8 中所计算的 Froullani 积分可直接得到

$$\int_0^{+\infty}\frac{e^{-ax}-e^{-bx}}{x}\,\mathrm{d}x=\log\frac{b}{a},\qquad\forall\ b>a>0.$$

现在我们换一种方式来计算它。将上式左边记作 $I$，那么

$$I=\int_0^{+\infty}\left(\int_a^b e^{-xy}\,\mathrm{d}y\right)\mathrm{d}x. \tag{14.11}$$

一方面，$e^{-xy}$ 在 $[0,+\infty)\times[a,b]$ 上连续；另一方面，当 $x\geqslant 0$ 且 $y\in[a,b]$ 时

$$0\leqslant e^{-xy}\leqslant e^{-ax},$$

而 $\int_0^{+\infty} e^{-ax}\,\mathrm{d}x$ 收敛，因此由 Weierstrass 判别法知 $\int_0^{+\infty} e^{-xy}\,\mathrm{d}x$ 在 $[a,b]$ 上一致收敛，于是由命题 2.17 知 (14.11) 中的积分号可以交换，从而有

$$I=\int_a^b \mathrm{d}y\int_0^{+\infty} e^{-xy}\,\mathrm{d}x=\int_a^b \frac{1}{y}\,\mathrm{d}y=\log\frac{b}{a}.$$

**【例 2.22】**计算 Dirichlet 积分 $\int_0^{+\infty}\frac{\sin x}{x}\,\mathrm{d}x$。

解. 引入收敛因子 $e^{-\alpha x}$ 并考虑积分

$$I(\alpha)=\int_0^{+\infty} e^{-\alpha x}\frac{\sin x}{x}\,\mathrm{d}x.$$

记

$$f(x,\alpha)=\begin{cases} e^{-\alpha x}\dfrac{\sin x}{x}, & \text{若 } x\neq 0,\\ 1, & \text{若 } x=0,\end{cases}$$

则 $\frac{\partial f}{\partial\alpha}=-e^{-\alpha x}\sin x$，容易验证 $f$ 与 $\frac{\partial f}{\partial\alpha}$ 均在 $[0,+\infty)\times[0,+\infty)$ 上连续。此外，由 $|e^{-\alpha x}\sin x|\leqslant e^{-\alpha x}$ 及 Weierstrass 判别法知，$\int_0^{+\infty}\frac{\partial f}{\partial\alpha}(x,\alpha)\,\mathrm{d}x$ 在 $(0,+\infty)$ 上内闭一致收敛。按照命题 2.20，在 $(0,+\infty)$ 的任一闭子区间上有

$$\begin{aligned} I'(\alpha)&=\int_0^{+\infty}\frac{\partial f}{\partial\alpha}(x,\alpha)\,\mathrm{d}x=-\int_0^{+\infty} e^{-\alpha x}\sin x\,\mathrm{d}x\\ &=\frac{e^{-\alpha x}(\alpha\sin x+\cos x)}{1+\alpha^2}\bigg|_0^{+\infty}=-\frac{1}{1+\alpha^2}.\end{aligned}$$

进而对任意的 $\alpha>0$ 得到

$$I(\alpha)-\lim_{A\to+\infty}I(A)=-\int_\alpha^{+\infty} I'(u)\,\mathrm{d}u=\frac{\pi}{2}-\arctan\alpha. \tag{14.12}$$

注意到对任意的 $\alpha>0$ 有

$$|I(\alpha)|=\left|\int_0^{+\infty} e^{-\alpha x}\frac{\sin x}{x}\,\mathrm{d}x\right|\leqslant\int_0^{+\infty} e^{-\alpha x}\,\mathrm{d}x=\frac{1}{\alpha}$$

因此 $\lim\limits_{A\to+\infty}I(A)=0$。将这代入 (14.12) 便得

$$I(\alpha)=\frac{\pi}{2}-\arctan\alpha,\qquad \forall\ \alpha>0. \tag{14.13}$$

最后，因为利用 Abel 判别法可得 $\displaystyle\int_0^{+\infty} e^{-\alpha x}\frac{\sin x}{x}\,\mathrm{d}x$ 在 $[0,+\infty)$ 上一致收敛（参见例 2.11），故而由命题 2.15 知 $I(\alpha)$ 在 $[0,+\infty)$ 上连续，于是在 (14.13) 的等式中令 $\alpha\to 0^+$ 便得

$$\int_0^{+\infty}\frac{\sin x}{x}\,\mathrm{d}x = I(0) = \lim_{\alpha\to 0^+} I(\alpha) = \frac{\pi}{2}.$$

□

【**注 2.23**】通过变量替换容易推出

$$\int_0^{+\infty}\frac{\sin ax}{x}\,\mathrm{d}x = \begin{cases} \dfrac{\pi}{2}, & \text{若 } a>0, \\ 0, & \text{若 } a=0, \\ -\dfrac{\pi}{2}, & \text{若 } a<0. \end{cases}$$

【**例 2.24**】计算广义积分 $I = \displaystyle\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x$。

解. 利用参数 $y>0$ 作变量替换 $x\longmapsto xy$ 可得

$$I = y\int_0^{+\infty} e^{-x^2y^2}\,\mathrm{d}x.$$

在上式两边同乘 $e^{-y^2}$ 并对 $y$ 从 0 到 $+\infty$ 积分，我们有

$$\begin{aligned} I^2 = I\int_0^{+\infty} e^{-y^2}\,\mathrm{d}y &= \int_0^{+\infty} e^{-y^2}\cdot y\left(\int_0^{+\infty} e^{-x^2y^2}\,\mathrm{d}x\right)\mathrm{d}y \\ &= \int_0^{+\infty}\left(\int_0^{+\infty} e^{-y^2(x^2+1)}y\,\mathrm{d}x\right)\mathrm{d}y. \end{aligned} \tag{14.14}$$

一方面，

$$\int_0^{+\infty} e^{-y^2(x^2+1)}y\,\mathrm{d}x = e^{-y^2}I$$

是关于 $y>0$ 的连续函数；另一方面，对任意的 $\delta>0$，

$$\int_\delta^{+\infty} e^{-y^2(x^2+1)}y\,\mathrm{d}y = \frac{1}{2}\int_\delta^{+\infty} e^{-y^2(x^2+1)}\,\mathrm{d}y^2 = \frac{e^{-\delta^2(x^2+1)}}{2(x^2+1)}$$

是关于 $x\geqslant 0$ 的连续函数，因此由推论 2.19 知

$$\begin{aligned} \int_\delta^{+\infty}\left(\int_0^{+\infty} e^{-y^2(x^2+1)}y\,\mathrm{d}x\right)\mathrm{d}y &= \int_0^{+\infty}\left(\int_\delta^{+\infty} e^{-y^2(x^2+1)}y\,\mathrm{d}y\right)\mathrm{d}x \\ &= \int_0^{+\infty}\frac{e^{-\delta^2(x^2+1)}}{2(x^2+1)}\,\mathrm{d}x. \end{aligned}$$

现令 $\delta \to 0^+$，那么由 (14.14) 及命题 2.15 可得

$$I^2 = \lim_{\delta \to 0^+} \int_0^{+\infty} \frac{e^{-\delta^2(x^2+1)}}{2(x^2+1)} \, \mathrm{d}x = \int_0^{+\infty} \frac{\mathrm{d}x}{2(x^2+1)} = \frac{\pi}{4}.$$

因此 $I = \dfrac{\sqrt{\pi}}{2}$。 □

【**例 2.25**】我们定义 B 函数[②]

$$\mathrm{B}(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} \, \mathrm{d}x, \tag{14.15}$$

则它在 $p > 0$ 且 $q > 0$ 时收敛，并且对任意的 $\varepsilon_1, \varepsilon_2 > 0$，$\mathrm{B}(p,q)$ 在 $[\varepsilon_1, +\infty) \times [\varepsilon_2, +\infty)$ 上一致收敛，故而它在 $\mathbb{R}^2_{>0}$ 上连续。下面来证明关系式

$$\mathrm{B}(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \qquad \forall\, p > 0,\ q > 0. \tag{14.16}$$

回忆起在第十一章例 3.9 中我们对 $s > 0$ 证明了

$$\Gamma(s) = \int_0^{+\infty} e^{-x} x^{s-1} \, \mathrm{d}x.$$

先考虑 $p > 1$ 且 $q > 1$ 的情形，我们有

$$\begin{aligned} \Gamma(p)\Gamma(q) &= \Gamma(q) \int_0^{+\infty} e^{-x} x^{p-1} \, \mathrm{d}x = \int_0^{+\infty} e^{-x} x^{p-1} \left( \int_0^{+\infty} e^{-y} y^{q-1} \, \mathrm{d}y \right) \mathrm{d}x \\ &= \int_0^{+\infty} \left( \int_0^{+\infty} e^{-(x+y)} x^{p-1} y^{q-1} \, \mathrm{d}y \right) \mathrm{d}x. \end{aligned}$$

对内层积分作变量替换 $y \longmapsto xy$ 可得

$$\Gamma(p)\Gamma(q) = \int_0^{+\infty} \left( \int_0^{+\infty} e^{-x(1+y)} x^{p+q-1} y^{q-1} \, \mathrm{d}y \right) \mathrm{d}x. \tag{14.17}$$

现记 $f(x,y) = e^{-x(1+y)} x^{p+q-1} y^{q-1}$，则由 $p > 1$ 且 $q > 1$ 知 $f(x,y)$ 在 $\mathbb{R}^2_{\geqslant 0}$ 上非负且连续。此外，

$$\int_0^{+\infty} f(x,y) \, \mathrm{d}x = \frac{\Gamma(p+q) y^{q-1}}{(1+y)^{p+q}}$$

在 $\mathbb{R}_{\geqslant 0}$ 上关于变量 $y$ 的连续函数；

$$\int_0^{+\infty} f(x,y) \, \mathrm{d}y = \Gamma(q) e^{-x} x^{p-1}$$

[②] 这里的 B 是希腊字母 $\beta$ 的大写。

在 $\mathbb{R}_{\geqslant 0}$ 上关于变量 $x$ 的连续函数。因此由推论 2.19 知 (14.17) 中的积分号可以交换，从而得到

$$\begin{aligned}\Gamma(p)\Gamma(q) &= \int_0^{+\infty} \mathrm{d}y \int_0^{+\infty} e^{-x(1+y)} x^{p+q-1} y^{q-1} \,\mathrm{d}x \\ &= \Gamma(p+q) \int_0^{+\infty} \frac{y^{q-1}}{(1+y)^{p+q}} \,\mathrm{d}y.\end{aligned}$$

因此只需说明

$$\mathrm{B}(p,q) = \int_0^{+\infty} \frac{y^{q-1}}{(1+y)^{p+q}} \,\mathrm{d}y$$

即可，而这可以在 (14.15) 中令 $x = \dfrac{1}{1+y}$ 得到。

接下来证明 $p > 0$，$q > 0$ 的情形。利用分部积分可得（留作练习）

$$\mathrm{B}(p+1,q+1) = \frac{pq}{(p+q+1)(p+q)} \mathrm{B}(p,q).$$

注意到 $p+1>1$ 及 $q+1>1$，故而利用上面一段的结论可推出

$$\begin{aligned}\mathrm{B}(p,q) &= \frac{(p+q+1)(p+q)}{pq} \mathrm{B}(p+1,q+1) \\ &= \frac{(p+q+1)(p+q)}{pq} \cdot \frac{\Gamma(p+1)\Gamma(q+1)}{\Gamma(p+q+2)} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},\end{aligned}$$

其中最后一步用到了 $\Gamma(s+1) = s\Gamma(s)$。

**【例 2.26】**在这个例子中，我们利用 (14.16) 来重新证明

$$\int_0^{+\infty} e^{-x^2} \,\mathrm{d}x = \frac{\sqrt{\pi}}{2}. \tag{14.18}$$

事实上，在 (14.16) 中取 $p = q = \dfrac{1}{2}$ 可得

$$\mathrm{B}\Big(\frac{1}{2},\frac{1}{2}\Big) = \Gamma\Big(\frac{1}{2}\Big)^2.$$

一方面，在 B 函数的定义式 (14.15) 中作变量替换 $x = \cos^2\theta$ 可推出

$$\mathrm{B}\Big(\frac{1}{2},\frac{1}{2}\Big) = \int_0^1 \frac{\mathrm{d}x}{\sqrt{x(1-x)}} = \int_{\frac{\pi}{2}}^0 \frac{-2\cos\theta\sin\theta}{\sqrt{\cos^2\theta\sin^2\theta}} \,\mathrm{d}\theta = \pi.$$

另一方面，

$$\Gamma\Big(\frac{1}{2}\Big) = \int_0^{+\infty} e^{-x} x^{-\frac{1}{2}} \,\mathrm{d}x = 2\int_0^{+\infty} e^{-x} \,\mathrm{d}\sqrt{x} = 2\int_0^{+\infty} e^{-x^2} \,\mathrm{d}x.$$

综上便得 (14.18)。

## 习 题 14.2

1. 设 $X$ 和 $Y$ 是 $\mathbb{R}$ 的两个子集，$f(x,y)$ 是定义在 $X\times Y$ 上的函数，$x_0$ 是 $X$ 的一个聚点。试写出判定 $x$ 沿 $X$ 中元素趋于 $x_0$ 时 $f(x,y)$ 在 $Y$ 上一致收敛的 Cauchy 准则并证明之。

2. 判断下列含参变量广义积分在指定集合上是否一致收敛：

   (1) $\displaystyle\int_1^{+\infty}\frac{y^2-x^2}{(x^2+y^2)^2}\,\mathrm{d}x$，$y\in\mathbb{R}$;

   (2) $\displaystyle\int_0^{1}\frac{y^2-x^2}{(x^2+y^2)^2}\,\mathrm{d}x$，$y\in[0,1]$;

   (3) $\displaystyle\int_0^{+\infty}e^{-xy}\sin x\,\mathrm{d}x$，$y\in[a,+\infty)$，其中 $a>0$;

   (4) $\displaystyle\int_0^{+\infty}e^{-xy}\sin x\,\mathrm{d}x$，$y\in(0,+\infty)$;

   (5) $\displaystyle\int_0^{+\infty}\frac{x\cos xy}{x^2+1}\,\mathrm{d}x$，$y\in[a,+\infty)$，其中 $a>0$;

   (6) $\displaystyle\int_0^{+\infty}\frac{x\cos xy}{x^2+1}\,\mathrm{d}x$，$y\in(0,+\infty)$;

   (7) $\displaystyle\int_0^{+\infty}\sqrt{y}e^{-yx^2}\,\mathrm{d}x$，$y\in(0,+\infty)$;

   (8) $\displaystyle\int_0^{1}\frac{1}{x^y}\sin\frac{1}{x}\,\mathrm{d}x$，$y\in(0,2)$。

3. 设 $f(x)$ 在 $\mathbb{R}_{\geqslant 0}$ 上连续，若积分 $\displaystyle\int_0^{+\infty}x^yf(x)\,\mathrm{d}x$ 在 $y=a$ 和 $y=b$ $(a<b)$ 时均收敛，证明 $\displaystyle\int_0^{+\infty}x^yf(x)\,\mathrm{d}x$ 在 $[a,b]$ 上一致收敛。

4. 证明积分
$$\int_1^{+\infty}e^{-\frac{1}{y^2}(x-\frac{1}{y})^2}\,\mathrm{d}x$$
在 $(0,1)$ 上一致收敛，但不存在与 $y$ 无关的函数作为上述被积函数的优势函数。

5. 设 $f(x,y)$ 是定义在 $[a,b]\times E$ 上的二元函数，$y_0$ 是 $E$ 的聚点，满足：

   (1) 对任意给定的 $y\in E$，$f(x,y)$ 作为 $x$ 的函数在 $[a,b]$ 上可积；

   (2) 当 $y\to y_0$ 时 $f(x,y)$ 在 $[a,b]$ 上一致收敛于 $\varphi(x)$。

证明 $\varphi(x)$ 在 $[a,b]$ 上可积并且

$$\lim_{\substack{y\to y_0\\ y\in E}}\int_a^b f(x,y)\,\mathrm{d}x=\int_a^b \varphi(x)\,\mathrm{d}x.$$

6. 设 $a\neq 0$，$g(y)=\displaystyle\int_0^{+\infty}\frac{1}{x}(1-e^{-xy})\cos ax\,\mathrm{d}x$。证明：

(1) $g(y)$ 在 $\mathbb{R}_{\geqslant 0}$ 上连续；

(2) $g(y)$ 在 $\mathbb{R}_{>0}$ 上可导。

7. 研究函数 $g(y)=\displaystyle\int_0^{+\infty}\frac{\arctan x}{x^y(3+x^3)}\,\mathrm{d}x$ 的定义域及连续性。

8. 计算 $I(\alpha)=\displaystyle\int_0^{+\infty}e^{-x^2}\cos\alpha x\,\mathrm{d}x$。

9. 设 $\alpha\geqslant 0$，计算

$$I(\alpha)=\int_0^{+\infty}\frac{\arctan\alpha x}{x(1+x^2)}\,\mathrm{d}x.$$

10. 设 $a>0$，$b>0$，计算

$$\int_0^{+\infty}\frac{1-\cos ax}{x}e^{-bx}\,\mathrm{d}x.$$

11. 利用 $\dfrac{\arctan x}{x}=\displaystyle\int_0^1\frac{1}{1+x^2y^2}\,\mathrm{d}y$ $(x\neq 0)$ 计算积分

$$I=\int_0^1\frac{\arctan x}{x\sqrt{1-x^2}}\,\mathrm{d}x.$$

12. 利用 $\dfrac{1}{\sqrt{x}}=\dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^{+\infty}e^{-xy^2}\,\mathrm{d}y$，并引入收敛因子来证明 Fresnel 积分

$$\int_{-\infty}^{+\infty}\sin x^2\,\mathrm{d}x=\int_{-\infty}^{+\infty}\cos x^2\,\mathrm{d}x=\sqrt{\frac{\pi}{2}}.$$

13. 利用例 2.22 和 2.24 中的结果计算下列积分：

(1) $\displaystyle\int_0^1\frac{1}{x}\sin\frac{1}{x}\,\mathrm{d}x$；　　(2) $\displaystyle\int_0^{+\infty}\frac{\sin^4 x}{x^2}\,\mathrm{d}x$；

(3) $\displaystyle\int_{-\infty}^{+\infty}e^{-(ax^2+bx+c)}\,\mathrm{d}x\quad(a>0)$；

(4) $\displaystyle\int_{-\infty}^{+\infty}e^{-x^2-\frac{a^2}{x^2}}\,\mathrm{d}x\quad(a>0)$；

(5) $\displaystyle\int_0^{+\infty} \frac{\cos ax - \cos bx}{x^2}\,\mathrm{d}x \quad (0 < a < b)$；

(6) $\displaystyle\int_0^{+\infty} \frac{e^{-ax^2} - e^{-bx^2}}{x^2}\,\mathrm{d}x \quad (0 < a < b)$。

14. 设 $p > 0$，$q > 0$。证明

$$\mathrm{B}(p+1, q+1) = \frac{pq}{(p+q+1)(p+q)}\mathrm{B}(p, q).$$

15. 计算 $\displaystyle\int_0^{+\infty} e^{-x}\log x\,\mathrm{d}x$。

形如 $\displaystyle\int_0^{+\infty} \cos(x^3 - 3xy)\,\mathrm{d}x$ 的积分被称为 Airy 积分，它是由 G. B. Airy 于 1838 年引入的，Airy 对其进行了一些数值计算。关于这一积分的渐近性态的研究始于 G. G. Stokes[45]。从第 16 题到第 20 题是一组题，给出了 $y \to +\infty$ 时这一积分的一个渐近公式，这里所用的方法源于 G. H. Hardy[20]。

16. 设 $y \geqslant 1$，利用变量替换 $t = \dfrac{x - \sqrt{y}}{\sqrt{y}}$ 及 §11.1 习题 13 证明

$$\int_0^{+\infty} \cos(x^3 - 3xy)\,\mathrm{d}x = \sqrt{y}\int_{-1}^{1} \cos\left((t^3 + 3t^2 - 2)y^{\frac{3}{2}}\right)\mathrm{d}t + O\left(\frac{1}{y}\right).$$

17. 令 $s = t^3 + 3t^2$，证明 $\dfrac{1}{3t^2 + 6t} - \dfrac{1}{2\sqrt{3s}}$ 在 $t \in (0, 1]$ 时分段单调且存在常数 $\alpha, \beta < 0$ 使得

$$\alpha \leqslant \frac{1}{3t^2 + 6t} - \frac{1}{2\sqrt{3s}} \leqslant \beta.$$

18. 利用变量替换 $s = t^3 + 3t^2$ 及 §11.1 习题 13 证明

$$\int_0^1 \cos\left((t^3 + 3t^2 - 2)y^{\frac{3}{2}}\right)\mathrm{d}t = \frac{1}{2\sqrt{3}}\int_0^4 \frac{\cos\left((s-2)y^{\frac{3}{2}}\right)}{\sqrt{s}}\,\mathrm{d}s + O(y^{-\frac{3}{2}}).$$

19. 利用习题 12 中的 Fresnel 积分证明

$$\int_0^1 \cos\left((t^3 + 3t^2 - 2)y^{\frac{3}{2}}\right)\mathrm{d}t = \frac{\sqrt{\pi}}{2\sqrt{3}\,y^{\frac{3}{4}}}\sin\left(2y^{\frac{3}{2}} + \frac{\pi}{4}\right) + O(y^{-\frac{3}{2}}).$$

20. 对 $y \geqslant 1$ 证明

$$\int_0^{+\infty} \cos(x^3 - 3xy)\,\mathrm{d}x = \sqrt{\frac{\pi}{3}}\,y^{-\frac{1}{4}}\sin\left(2y^{\frac{3}{2}} + \frac{\pi}{4}\right) + O\left(\frac{1}{y}\right).$$

# 参 考 文 献


1. K. Alladi，C. Defant，*Revisiting the Riemann zeta function at positive even integers*，Int. J. Number Theory，**14** (2018)，1849−1856。
2. E. Artin，*The Gamma function*，Holt，Rinehart and Winston，Inc.，New York，1964。
3. M. Artin，*Algebra*，Prentice-Hall，Inc.，New Jersey，1991。
4. S. Banach，A. Tarski，*Sur la décomposition des ensembles de points en parties respectivement congruentes*，Fund. Math.，**6** (1924)，244−277。
5. S. Bernstein，*Démonstration du théorème de Weierstrass, fondeé sur le calcul des probabilités*，Commun. Soc. Math. Kharkow，**13** (2) (1912−1913)，1−2。
6. N. Bourbaki，*Functions of a real variable*，Springer-Verlag，Berlin，2004。
7. ———，*General topology*，Springer-Verlag，New York，1989。
8. 常庚哲，史济怀，数学分析教程（第三版），中国科学技术大学出版社，合肥，2012。
9. N. G. de Bruijn，*Asymptotic methods in analysis*，North-Holland，Amsterdam，1958。
10. H. Cartan，*Cours de calcul différentiel*，Hermann，Paris，1967。（中译本：微分学，高等教育出版社， 2009。）
11. R. Courant，F. John，*Introduction to calculus and analysis*，Springer-Verlag，New York，1989。
12. A. Deitmar，*A first course in harmonic analysis (2nd ed.)*，Springer-Verlag，Berlin，2005。
13. Б. П. Демидович，Сборник задач и упражнений по математическому анализу，Гостехиздат，1956。（中译本：吉米多维奇，数学分析习题集，人民教育出版社， 1958。）
14. J. Dieudonné，*Foundations of modern analysis*，Academic Press，Inc.，New York，1960。
15. ———，*Calcul infinitésimal*，Hermann，Paris，1980。（中译本：无穷小计算，高等教育出版社， 2012。）
16. P. Du Bois−Reymond，*Versuch einer Classification der willkürlichen Functionen reeller Argumente*，J. für. Math.，**79** (1875)，21−37。
17. Г. М. Фихтенгольц，Курс дифференциального и интегрального исчисления，Наука，1969。（中译本：菲赫金哥尔茨，微积分学教程，高等教育出版社，2006。）

18. I. S. Gradshteyn，I. M. Ryzhik，*Table of integrals, series, and products (7th ed.)*，Elsevier Inc.，2007。

19. G. Green，*An essay on the application of mathematical analysis to the theories of electricity and magnetism*， J. Reine Angew. Math.，**39** (1850)，73–89；**44** (1852)，356–374；**47** (1854)，161–221。

20. G. H. Hardy，*On certain definite integrals considered by Airy and by Stokes*， Quart. J. Math.，**41** (1910)，226–240。

21. G. H. Hardy，M. Riesz，*The general theory of Dirichlet's series*，Cambridge Tracts in Mathematics and Mathematical Physics，vol. 18，Cambridge University Press，Cambridge，1915。

22. 华罗庚，高等数学引论，高等教育出版社，北京，2009。

23. ———，数论导引，科学出版社，北京，1957。

24. A. Hurwitz，*Sur quelques applications géométriques des séries de Fourier*， Annales Scientifiques de l'École Normale Supérieure，**19** (1902)，357–408。

25. J. L. W. V. Jensen，*Sur les fonctions convexes et les inégalités entre les valeurs moyennes*，Acta Math.，**30** (1906)，175–193。

26. G. Klambauer，*Mathematical analysis*，Marcel Dekker，Inc.，New York，1975。

27. ———，*Problems and propositions in analysis*，Marcel Dekker，Inc.，New York，1979。

28. M. Kline，*Mathematical thought from ancient to modern times*，Oxford University Press，1972。（中译本：古今数学思想，上海科学技术出版社，2002。）

29. E. Landau，*Foundations of analysis (3rd ed.)*，Chelsea Publishing Company，New York，1966。

30. 廖可人，李正元，数学分析 3，高等教育出版社，北京，2015。

31. 林源渠，方企勤，李正元，廖可人，数学分析习题集，高等教育出版社，北京，2015。

32. L. H. Loomis，S. Sternberg，*Advanced calculus*，Jones and Bartlett Publishers，Inc.， London，1990。（中译本：高等微积分，高等教育出版社，2005。）

33. G. G. Lorentz，*Bernstein polynomials*，Chelsea Publishing Company，New York，1986。

34. 梅加强，数学分析，高等教育出版社，北京，2011。

35. 闵嗣鹤，严士健，初等数论（第三版），高等教育出版社，北京，2003。

36. J. R. Munkres，*Analysis on manifolds*，Westview Press，1991。

37. 欧阳光中，朱学炎，金福临，陈传璋，数学分析（第三版），高等教育出版社，北京，2007。

38. 潘承洞，于秀源，阶的估计，山东科学技术出版社，济南，1983。

39. C. C. Pugh，*Real mathematical analysis*，Springer-Verlag，New York，2001。

40. B. Riemann，*Gesammelte mathematische werke und wissenschaftlicher nachlass*，Leipzig，1892。（中译本：黎曼全集，高等教育出版社，2016。）

41. W. Rudin，*Principles of mathematical analysis (3rd ed.)*，McGraw-Hill Companies，Inc.，1976。
42. I. J. Schoenberg，*On the Peano curve of Lebesgue*，Bull. Amer. Math. Soc.，**44** (1938)，519。
43. M. Spivak，*Calculus on manifolds*，Westview Press，1995。
44. E. M. Stein，R. Shakarchi，*Fourier analysis: An introduction*，Princeton University Press，2003。
45. G. G. Stokes，*On the numerical calculation of a class of definite integrals and infinite series*，Camb. Phil. Trans.，Vol. ix.，Part I.，1850。
46. T. Tao，*An introduction to measure theory*，Graduate Studies in Mathematics，vol. 126，Amer. Math. Soc.，Providence，2011。
47. E. C. Titchmarsh，*The theory of functions (2nd ed.)*，Oxford University Press，London，1939。
48. B. L. van der Waerden，*Ein einfaches Beispiel einer nicht-differenzierbaren stetigen Funktion*，Math. Z.，**32** (1930)，474–475。
49. 汪林，实分析中的反例，高等教育出版社，北京，2014。
50. E. T. Whittaker，G. N. Watson，*A course of modern analysis (4th ed.)*，Cambridge University Press，1927。
51. 谢惠民，恽自求，易法槐，钱定边，数学分析习题课讲义，高等教育出版社，北京，2003。
52. 杨宗磐，数学分析入门，科学出版社，北京，1958。
53. 周民强，数学分析习题演练（第二版），科学出版社，北京，2010。
54. В. А. Зорич，Математический анализ，МЦНМО.，Москва，2002。（中译本：卓里奇，数学分析，高等教育出版社，2006。）

# 索　引