数学分析(第三版)修订计划

梅加强

2022.2.14

章节次序调整: 前六章次序不变

第七章欧氏空间

第八章多元函数的微分

第九章多元函数的积分

第十章曲线积分与曲面积分

第十一章微分形式的积分

第十二章数项级数

第十三章函数项级数

第十四章 Fourier 分析

第十五章含参变量的积分

记号调整: Gauss 函数用 $\lfloor\ \rfloor$ 表示? 集合的内部用 int 表示, 直径从 diam 换成 d, 体积换成容积, 立方体换成方体. 集合的减法改用 $\setminus$, 补集尽量用减法表示.

第一章　引言

1.1　从求和谈起

1.1 节: 增加 3 维球体的体积公式

用矩形等规则图形去逼近曲边图形的方法简称“以直化曲”. 这种方法也可以应用在空间问题中.

问题 4 三维球体的体积是多少?

三维球体的体积公式古已有之. 不过, 刘徽发现古书中体积公式并不正确, 要想得到正确的公式则要计算所谓的“牟合方盖”(两个同样半径的圆柱体沿中心线垂直相交所得几何体)的体积. 正确的体积公式直到祖暅才终于推导出来.

如果承认圆的面积公式和圆柱的体积公式, 则我们可以用“以直化曲”的方法来推导球体的体积公式. 半径为 R 的三维球体的体积记为 $V_3(R)$. 以球心为原点建立直角坐标系, 则该球体可以表示为 $\{x^2+y^2+z^2\le R^2\}$. 设 n 为大于 1 的正整数, 记

$$C_i=\Big\{x^2+y^2\le R^2-\frac{i^2}{n^2}R^2,\ \ \frac{i-1}{n}R\le z\le\frac{i}{n}R\Big\},\ \ i=1,2,\cdots,n-1.$$

C_i 均为包含在上半球体中的圆柱体, 因此

$$\begin{aligned}\frac{1}{2}V_3(R)>\sum_{i=1}^{n-1}\pi\Big[R^2-\frac{i^2}{n^2}R^2\Big]\frac{R}{n}&=\sum_{i=1}^{n}\frac{\pi R^3}{n}\Big[1-\frac{i^2}{n^2}\Big]\\&=\pi R^3\Big[\frac{2}{3}-\frac{1}{2n}-\frac{1}{6n^2}\Big].\end{aligned}$$

另一方面, 记

$$\bar{C}_i=\Big\{x^2+y^2\le R^2-\frac{i^2}{n^2}R^2,\ \ \frac{i}{n}R\le z\le\frac{i+1}{n}R\Big\},\ \ i=0,1,2,\cdots,n-1.$$

这 n 个圆柱体并起来包含了上半球体, 因此

$$\frac{1}{2}V_3(R)<\sum_{i=0}^{n-1}\pi\Big[R^2-\frac{i^2}{n^2}R^2\Big]\frac{R}{n}=\pi R^3\Big[\frac{2}{3}+\frac{1}{2n}-\frac{1}{6n^2}\Big].$$

当 n 很大时, 与 $\frac{1}{n}$, $\frac{1}{n^2}$ 相关的误差项可以忽略不计, 由此可以得出 $V_3(R)=\frac{4}{3}\pi R^3$.

增加习题:

1. (1) 设 m 为正整数, 当 $i \geq 0$ 时, 证明

$$(m+1)i^m < (i+1)^{m+1} - i^{m+1} < (m+1)(i+1)^m.$$

(2) 利用"以直化曲"的方法以及 (1) 中的估计求函数 $y = x^m$ 的图像与 x 轴以及直线 $x = 1$ 在第一象限所围成区域的面积.

1.2 迭代与收敛

我们来考虑两个具体的计算问题. 边长为 1 的正方形, 其对角线的长度记为 ℓ. 由勾股定理可知 $\ell^2 = 2$. 古人早就知道 ℓ 不是有理数, 也希望能用有理数(特别是有限位小数)尽可能准确地表示它. 基本的想法是用一个迭代的过程去不断算出 ℓ 的近似值.

比如先取 ℓ 的一个近似值 ℓ_0, 记 ℓ_0 和 ℓ 之间的误差为 h, 即 $\ell = \ell_0 + h$, 则

$$2 = \ell^2 = (\ell_0 + h)^2 = \ell_0^2 + 2\ell_0 h + h^2.$$

关键的想法是, 若 $|h|$ 很小, 则 h^2 更小, 不妨忽略不计. 这样可以从上式中解出 $h \approx \frac{1}{\ell_0} - \frac{\ell_0}{2}$, 从而

$$\ell = \ell_0 + h \approx \frac{1}{\ell_0} + \frac{\ell_0}{2}.$$

记 $\ell_1 = \frac{1}{\ell_0} + \frac{\ell_0}{2}$, 则 ℓ_1 是 ℓ 的一个新的近似值, 它是由 ℓ_0 迭代而来的.

一般地, 当 $n \geq 0$ 时, 记

$$\ell_{n+1} = \frac{1}{\ell_n} + \frac{\ell_n}{2},$$

这就是所谓的**Newton 迭代公式**, 它给出了 ℓ 的一列近似值.

当 $x > 0$ 时, 记 $f(x) = \frac{1}{x} + \frac{x}{2}$, 则 $\ell_{n+1} = f(\ell_n)$. 为了从直观上探讨 $\{\ell_n\}$ 的变化趋势, 我们可以观察平面上的点列 $\{(\ell_n, f(\ell_n)) = (\ell_n, \ell_{n+1})\}$. 取 $\ell_0 = 1.5$, 从下图可以看出, 当 n 变大时, 这一列点很快地趋于 $y = x$ 和 $y = f(x)$ 的交点 (ℓ, ℓ).

我们也可以给出 ℓ_n 和 ℓ 之间的误差估计. 首先, 注意到

$$\ell_n - \ell = \frac{(\ell_{n-1} - \ell)^2}{2\ell_{n-1}},$$

结合 $0 < \ell_0 - \ell < \frac{1}{10}$ 与数学归纳法不难得出

$$0 < \ell_n - \ell < \frac{1}{10^{2^n}}.$$

这说明, 当 n 较大时, 有理数 ℓ_n 的确是 ℓ 很好的近似值. 比如

$$\ell \approx \ell_2 = \frac{577}{408} = 1.414\,215\cdots.$$

我们再来考虑所谓的黄金分割比. 将一直线段分为两部分, 若较短部分与较长部分长度之比等于较长部分与总长度之比, 则此比例称为黄金分割比, 记为 φ. 记 $g(x)=\frac{1}{1+x}$, 则 $g(\varphi)=\varphi$. 当 $n\geq 1$ 时, 令 $g_{n+1}(x)=g\big(g_n(x)\big)$, 其中 $g_1(x)=g(x)$. 记 $h(x)=g_5(x)$, 则 $h(x)=\frac{3x+5}{5x+8}$, 且 $h(\varphi)=\varphi$. 受此启发, 我们用如下迭代公式给出 φ 的一列有理近似值:

$$\varphi_0=0.6,\ \ \varphi_{n+1}=h(\varphi_n)=\frac{3\varphi_n+5}{5\varphi_n+8},\ \ \forall\, n\geq 0.$$

观察平面上的点列 $\{(\varphi_n,h(\varphi_n))=(\varphi_n,\varphi_{n+1})\}$, 从下图可以看出, 当 n 变大时, 这一列点很快地趋于 $y=x$ 和 $y=h(x)$ 的交点 (φ,φ).

我们也可以给出 φ_n 和 φ 之间的误差估计. 易见 $\varphi_n\geq 0.6$, 且

$$\varphi_n-\varphi=h(\varphi_{n-1})-h(\varphi)=\frac{\varphi-\varphi_{n-1}}{(8+5\varphi_{n-1})(8+5\varphi)}.$$

由归纳法可得

$$|\varphi_n-\varphi|<\frac{\varphi-0.6}{11^{2n}},\ \ \forall\, n\geq 1.$$

这说明, 当 n 较大时, 有理数 φ_n 的确是 φ 很好的近似值. 比如

$$\varphi\approx\varphi_2=\frac{377}{610}=0.618\,03\cdots.$$

习题:

1. 设 $c>0$, 令

$$a_0>0,\ \ a_{n+1}=\frac{ca_n+2}{a_n+c},\ \ \forall\, n\geq 0.$$

请通过作图了解 $\{a_n\}$ 的变化趋势.

2. 记 $f(x)=\frac{x+2}{x+1}$. 当 $n\geq 1$ 时, 令 $f_{n+1}(x)=f\big(f_n(x)\big)$, 其中 $f_1(x)=f(x)$.

(1) 证明: $f_n(x)=\frac{a_nx+2b_n}{b_nx+a_n}$, 其中

$$a_1=b_1=1,\ \ a_{n+1}=a_n+2b_n,\ \ b_{n+1}=a_n+b_n,\ \ \forall\, n\geq 1.$$

(2) 证明: $a_n^2-2b_n^2=\pm 1$. 再由此说明当 n 很大时 $\frac{a_n}{b_n}\approx\ell$.

3. 考虑如下迭代数列

$$a_0=\frac{17}{12},\ \ a_{n+1}=\frac{17a_n+24}{12a_n+17},\ \ \forall\, n\geq 0.$$

证明: 当 $n\geq 1$ 时 $0<a_n-\ell<10^{-3n}(a_0-\ell)$.

4. 令 $a_0=1$, $n\geq 0$ 时 $a_{n+1}=1+\frac{a_n}{1+a_n}$. 请通过作图了解 $\{a_n\}$ 的变化趋势.

5. 记 $g(x)=\frac{1}{1+x}$, 当 $n\geq 1$ 时, 令 $g_{n+1}(x)=g\big(g_n(x)\big)$, 其中 $g_1(x)=g(x)$. 证明: $g_n(x)=\frac{F_{n-1}x+F_n}{F_nx+F_{n+1}}$, 其中

$$F_0=0,\ \ F_1=1,\ \ F_{n+1}=F_{n-1}+F_n,\ \ \forall\, n\geq 1.$$

并说明当 n 很大时 $\frac{F_{n-1}}{F_n}\approx\varphi$.

1.3 比较与估计

通过研究面积、体积、长度、比例等问题, 我们可以体会分析学的一些基本特点. 分析学的基本手法包括定性分析和定量分析, 也就是定性估计和定量估计. 做估计也就是去证明某种不等式. 下面我们介绍几个常用的等式和不等式, 它们是做估计的基本工具.

设 n 为正整数, $u \ge 0$, 利用 $u^n - 1 = (u^{n-1} + u^{n-2} + \cdots + u + 1)(u-1)$ 可得

$$n(u-1) \le u^n - 1 \le nu^{n-1}(u-1) \le nu^n(u-1). \tag{1.1}$$

当 $x \ge -1$ 时, 在上式左边取 $u = 1 + x$ 可得如下 Bernoulli 不等式:

$$(1+x)^n \ge 1 + nx. \tag{1.2}$$

上式也可以对 n 用归纳法证明. 用归纳法还可得到广义 Bernoulli 不等式: 设 $x_i \ge -1$, $i = 1, 2, \cdots, n$. 如果 $\{x_i\}_{i=1}^n$ 同号, 则

$$(1+x_1)(1+x_2)\cdots(1+x_n) \ge 1 + x_1 + x_2 + \cdots + x_n. \tag{1.3}$$

设 $a, b \ge 0$, 由 $(a-b)^2 \ge 0$ 可得

$$a^2 + b^2 \ge 2ab, \text{ 或 } \Big(\frac{a+b}{2}\Big)^2 \ge ab. \tag{1.4}$$

上式可以推广到 n 个数的情形.

例 1.3.1 (平均值不等式)**.** 设 $a_1, a_2, \cdots, a_n \ge 0$, 则

$$\Big(\frac{a_1 + a_2 + \cdots + a_n}{n}\Big)^n \ge a_1 a_2 \cdots a_n. \tag{1.5}$$

证明 用归纳法. $n = 1$ 时不等式自然成立. 设 $n = m$ 时不等式成立, 欲证 $n = m + 1$ 时不等式也成立. 记 $A = \frac{1}{m}\sum\limits_{i=1}^m a_i$, 不妨设 $A > 0$, 则

$$\begin{aligned}\frac{1}{m+1}\sum_{i=1}^{m+1} a_i &= \frac{mA + a_{m+1}}{m+1} = A + \frac{a_{m+1} - A}{m+1}\\ &= A\Big[1 + \frac{a_{m+1} - A}{(m+1)A}\Big].\end{aligned}$$

由 Bernoulli 不等式可得

$$\Big(\frac{1}{m+1}\sum_{i=1}^{m+1} a_i\Big)^{m+1} \ge A^{m+1}\Big[1 + (m+1)\frac{a_{m+1} - A}{(m+1)A}\Big] = A^m a_{m+1}.$$

由归纳假设, $A^m \ge a_1 a_2 \cdots a_m$. 这说明 $n = m + 1$ 时欲证不等式的确成立. 于是不等式对所有正整数 n 都成立.

在做估计时, 经常不加申明地使用绝对值不等式, 即

$$|a+b| \leq |a|+|b|, \ \forall\, a,b \in \mathbb{R}. \tag{1.6}$$

它还有一些别的表现形式, 比如

$$\big||a|-|b|\big| \leq |a-b|, \ \forall\, a,b \in \mathbb{R}. \tag{1.7}$$

如下三角不等式也是它的另一表现形式:

$$|a-b| \leq |a-c|+|c-b|, \ \forall\, a,b,c \in \mathbb{R}. \tag{1.8}$$

在分析学中, 我们要研究当某些量变动时, 与之相关联的量如何变动. 比如, 我们考虑 n 次函数如何随自变量变化而变化. 设 $u,v \in \mathbb{R}$, n 为正整数, 则有如下因式分解:

$$u^n - v^n = (u-v)\sum_{i=1}^{n} u^{n-i}v^{i-1}, \tag{1.9}$$

其中我们规定 $u^0 = v^0 = 1$. 特别地, 当 $u,v \geq 0$ 时, 从上式可得

$$nv^{n-1}(u-v) \leq u^n - v^n \leq nu^{n-1}(u-v). \tag{1.10}$$

下面的例子也属于这种类型结果.

例 1.3.2. *设 $a,b,c,d \subset [-M,M]$. 如果 $|a-c|<\varepsilon$, $|b-d|<\varepsilon$, 则 $|ab-cd| \leq 2M\varepsilon$.*

证明 为了估计 ab 和 cd 之间的差异, 我们在这两个数之间插入 bc, 再利用三角不等式:

$$\begin{aligned} |ab-cd| &\leq |ab-bc|+|bc-cd| = |b||a-c|+|c||c-d| \\ &\leq M\varepsilon + M\varepsilon = 2M\varepsilon. \end{aligned}$$

我们再来看一个例子, 它也使用了在中间插入某个数的技巧.

例 1.3.3. *设 m,n 为正整数, $u \geq 0$. 如果 $m \geq n$, 则*

$$n(u^m-1) \geq m(u^n-1). \tag{1.11}$$

证明 不妨设 $m>n$. 我们先将 u^m-1 改写为

$$u^m - 1 = u^n\big(u^{m-n}-1\big) + \big(u^n-1\big),$$

再利用 (1.1) 可得

$$\begin{aligned} u^m-1 &\geq u^n(m-n)(u-1) + \big(u^n-1\big) \\ &\geq \frac{m-n}{n}\big(u^n-1\big) + \big(u^n-1\big), \end{aligned}$$

整理以后即得欲证不等式.

处理与实数相关的问题时, 有时可以先考虑有理数. 下面的例子表明, 可以用有理数去逼近给定的实数.

例 1.3.4. 设 $v \in \mathbb{R}$, n 为正整数, 则存在整数 k, 使得

$$\frac{k}{n} \le v < \frac{k+1}{n}. \tag{1.12}$$

证明 当 $x \in \mathbb{R}$ 时, 我们用 $\lfloor x \rfloor$ 表示不超过 x 的最大整数, $\lfloor x \rfloor$ 称为 Gauss 函数. 根据定义, 我们有

$$\lfloor x \rfloor \le x < \lfloor x \rfloor + 1. \tag{1.13}$$

在上式中取 $x = nv$, $k = \lfloor nv \rfloor$, 则 $k \le nv < k+1$.

通过对等式中的某些项做放缩也可以得到不等式. 例如, 当 n 为正整数时, 我们有 Newton 二项式展开式

$$(a+b)^n = \sum_{i=0}^{n} C_n^i a^i b^{n-i}, \ \ \forall\, a, b \in \mathbb{R}.$$

如果 $a, b \ge 0$, $n \ge 2$, 则有

$$(a+b)^n \ge a^n + na^{n-1}b + \frac{1}{2}n(n-1)a^{n-2}b^2. \tag{1.14}$$

增加习题:

1. 用归纳法证明广义 Bernoulli 不等式.
2. 设 $a > 0$, $a + b \ge 0$, n 为正整数, 则 $(a+b)^n \ge a^n + na^{n-1}b$.
3. 设 $\{a_i, b_i\}_{i=1}^n \subset [-M, M]$. 如果 $|a_i - b_i| \le \varepsilon$ 对每一个 i 都成立, 则

$$\Big|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i\Big| \le nM^{n-1}\varepsilon.$$

4. 设 $a_i, b_i \ge M > 0$, $|a_i - b_i| \le \varepsilon$, $\forall\ i = 1, 2, \cdots, n$. 证明:

$$\Big|\prod_{i=1}^n \frac{1}{a_i} - \prod_{i=1}^n \frac{1}{b_i}\Big| \le \frac{n\varepsilon}{M^{n+1}}.$$

5. 设 $i = 1, 2\cdots, n$ 时 $b_i > 0$, $\frac{a_i}{b_i} \in [a, b]$. 证明: $\frac{a_1+a_2+\cdots+a_n}{b_1+b_2+\cdots+b_n} \in [a, b]$.
6. 设 $a, b \in \mathbb{R}$, 证明:

$$\max\{a, b\} = \frac{1}{2}[a + b + |a-b|], \ \ \min\{a, b\} = \frac{1}{2}[a + b - |a-b|].$$

1.4 确界与开方

中学数学中已经介绍过开方、对数、指数等概念. 不过, 如果认真追究一下, 就会发现这些概念实际上并没有严格地定义过. 以开平方为例. 设 $\beta > 0$, 我们要

解方程 $x^2=\beta$, 它的唯一正解记为 $\sqrt{\beta}$. 问题在于**存在性**: 为什么这个方程一定**存在**正解?

中学数学中默认了解的存在性. 因为从几何直观上看, 平面直线 $y=\beta$ 与曲线 $y=x^2$ 在第一象限中交点的横坐标就应该是方程的正解. 不过, 几何直观不能代替严格证明!

设 m 为大于 1 的正整数, 下面我们研究方程 $x^m=\beta$ 的正解. 由 (1.9) 可知正解具有唯一性, 关键在于存在性. 不妨设 $\beta>1$. 我们用递归的方式构造一列数 $\{a_n\}$, 使得每一个 a_n 都是方程的近似解, 希望 n 很大时 a_n 趋于一个真正的解.

先取 $a_0=1$. 由 $\beta>1$ 可知 $a_0^m<\beta$. 若已找到一个近似解 a_n, 满足条件 $a_n>0$ 且 $a_n^m<\beta$, 我们来选取适当的 $\delta>0$, 使得 $(a_n+\delta)^m<\beta$ 仍成立. 此时再记 $a_{n+1}=a_n+\delta$, 这样就递归地定义了一列数 $\{a_n\}$, 希望得到的数列 $\{a_n\}$ 满足要求.

δ 应该怎么取呢? 首先, 由 $(a_n+\delta)^m<\beta<\beta^m$ 可知 $a_n+\delta<\beta$ 必须成立. 其次, 应用不等式 (1.10) 可得

$$(a_n+\delta)^m\le a_n^m+m(a_n+\delta)^{m-1}\delta<a_n^m+m\beta^{m-1}\delta. \tag{1.15}$$

受此启发, 取 $\delta=\frac{\beta-a_n^m}{m\beta^{m-1}}$, 则由不等式 (1.10) 可得

$$0<\delta<\frac{\beta^m-a_n^m}{m\beta^{m-1}}\le\beta-a_n$$

记 $a_{n+1}=a_n+\delta$, 由 (1.15) 可知 $a_{n+1}^m<\beta$ 仍成立.

总之, 我们得到数列 $\{a_n\}$, 它满足的迭代公式为

$$a_0=1,\ \ a_{n+1}=a_n+\frac{\beta-a_n^m}{m\beta^{m-1}},\ \ \forall\, n\ge 0. \tag{1.16}$$

由上述讨论和归纳法可以得出: $a_n>0$, $a_n^m<\beta$ 总成立, 于是 a_n 不断变大, 但都小于 β.

从直观上看, $\{a_n\}$ 这一列数最终应该趋于某个固定的位置. 为了说明这个固定位置的**存在性**, 我们需要引进新的概念.

设 $A\subset\mathbb{R}$. 如果存在 $M\in\mathbb{R}$, 使得 $A\subset(-\infty,M]$, 则称 A 有上界, M 是 A 的一个上界; 如果存在 $m\in\mathbb{R}$, 使得 $A\subset[m,+\infty)$, 则称 A 有下界, m 是 A 的一个下界. 既有上界, 又有下界的子集称为有界子集.

确界原理 设 $A\subset\mathbb{R}$ 为非空子集. 如果 A 有上界, 则它有一个最小的上界, 称为 A 的**上确界**, 记为 $\sup A$; 如果 A 有下界, 则它有一个最大的下界, 称为 A 的**下确界**, 记为 $\inf A$.

确界原理是实数所满足的一条公理, 我们不加证明地承认它. 回到数列 $\{a_n\}$, 记 $A=\{a_n\mid n\ge 1\}$, 则 β 是 A 的一个上界. 根据确界原理, A 有上确界, 记为 α, 我们要说明 $\alpha^m=\beta$.

首先, 由递推公式 (1.16) 可知

$$a_n = a_{n+1} + \frac{a_n^m - \beta}{m\beta^{m-1}} \leq \alpha + \frac{\alpha^m - \beta}{m\beta^{m-1}},$$

这说明 $\alpha + \frac{\alpha^m-\beta}{m\beta^{m-1}}$ 也是 A 的上界, 由 α 为最小上界可知 $\alpha^m \geq \beta$.

其次, 利用 (1.10) 可得

$$a_n \leq \alpha - \frac{\alpha^m - a_n^m}{m\alpha^{m-1}} < \alpha + \frac{\beta - \alpha^m}{m\alpha^{m-1}},$$

同理就可得出 $\alpha^m \leq \beta$.

总之, $\alpha^m = \beta$. 我们将 α 称为 β 的 m 次根, 记为 $\sqrt[m]{\beta}$ 或 $\beta^{\frac{1}{m}}$.

例 1.4.1. *验证 $(0,1)$ 的上确界为 1, 下确界为 0.*

证明 以上确界为例. 显然, 1 是 $(0,1)$ 的上界. 我们要说明小于 1 的数不是 $(0,1)$ 的上界, 从而 1 是 $(0,1)$ 的最小上界(上确界).

设 $a < 1$. 若 $a \leq 0$, 则 $a < 0.5$, 而 $0.5 \in (0,1)$, 因此 a 不是 $(0,1)$ 的上界; 若 $0 < a < 1$, 则 $a < \frac{1}{2}(a+1)$, 而 $\frac{1}{2}(a+1) \in (0,1)$, 因此 a 也不是 $(0,1)$ 的上界.

例 1.4.2 (确界的刻画)**.** *设 $A \subset \mathbb{R}$ 为非空子集. 则 M 为 A 的上确界当且仅当 M 为 A 的上界, 且任给 $\varepsilon > 0$, 均存在 $a \in A$, 使得 $a > M - \varepsilon$.*

证明 (必要性) 若 M 是 A 的上确界, 则 M 当然是 A 的上界. 任给 $\varepsilon > 0$, 因为 $M - \varepsilon < M$, 它不再是 A 的上界, 故存在 $a \in A$, 使得 $a > M - \varepsilon$.

(充分性) 若 M 满足题设条件, 则只要说明小于 M 的数都不是 A 的上界即可. 设 $m < M$, 取 $\varepsilon = M - m$, 则 $\varepsilon > 0$. 由题设, 存在 $a \in A$, 使得 $a > M - \varepsilon = m$, 这说明 m 不是 A 的上界.

下确界有类似的刻画. 它们常用来判断某个数是不是非空子集的确界.

习题:

1. 设 $A \subset \mathbb{R}$ 为非空子集. 证明: 如果 A 中有最大数, 则最大数就是 A 的上确界; 如果 A 中有最小数, 则最小数就是 A 的下确界.

2. 设 $A \subset B$. 证明: 当 B 有下界时 $\inf A \geq \inf B$; 当 B 有上界时 $\sup A \leq \sup B$.

3. 设 m 为大于 1 的正整数, $\beta > 1$. 令

$$a_0 = \beta, \;\; a_{n+1} = a_n + \frac{\beta - a_n^m}{m a_n^{m-1}}, \;\; \forall\, n \geq 0.$$

证明: $\{a_n\}$ 满足 $a_n > 0$, $a_n^m > \beta$, 且 $\{a_n \mid n \geq 1\}$ 的下确界 α 满足 $\alpha^m = \beta$.

4. 记 $\alpha_0 = 1$, 当 $n \geq 1$ 时, 令

$$\alpha_n = \alpha_{n-1} + \frac{1}{2} - \frac{\alpha_{n-1}^2}{4}.$$

证明: $\{\alpha_n\}$ 满足 $1 \le \alpha_n < \sqrt{2}$, 且

$$0 < \sqrt{2} - \alpha_n \le \frac{1}{2^n}\big(\sqrt{2} - \alpha_0\big) < \frac{1}{2^{n+1}}.$$

5. 设 $a, b \ge 0$, n 为正整数, 证明 $|\sqrt[n]{a} - \sqrt[n]{b}\,| \le \sqrt[n]{|a-b|}$.

6. 设 n 为正整数, A 为非空集合且 $A \subset [0, M]$. 记 $\sqrt[n]{A} = \{\sqrt[n]{a} \mid a \in A\}$. 证明 $\sup \sqrt[n]{A} = \sqrt[n]{\sup A}$

1.5 附录: 实数系的构造

1.5 附录: 实数系的构造

在**定义 1.4.1** 中修改 $\alpha^c = \mathbb{Q} \setminus \alpha$.

在上界定义之前增加例子:

1.5.3(Gauss 函数) 设 $\alpha \in \mathbb{R}$, 则存在整数 n, 使得 $n^* \le \alpha < (n+1)^*$.

证明 取 $r \in \alpha$, $s \in \alpha^c$, 根据**定义 1.5.1**和**命题 1.5.1** 可知 $r^* < \alpha \le s^*$. 取整数 k, l, 使得 $k \le r$, $s \le l$, 则 $k^* < \alpha \le l^*$. 考虑集合 $A = \{n \in \mathbb{Z} \mid k \le n,\ n^* \le \alpha\}$, 则 $k \in A$, 且 A 是有限集, 其最大数就满足我们的要求.

注 满足题设要求的整数 n 常记为 $\lfloor \alpha \rfloor$.

例 1.5.4(有理数的稠密性) 设 $\alpha, \beta \in \mathbb{R}$. 如果 $\alpha < \beta$, 则存在有理数 r, 使得 $\alpha < r^* < \beta$.

证明 取 $s \in \beta \setminus \alpha$, 则 $\alpha \le s^* < \beta$. 再取 $r \in \beta$, 使得 $r > s$, 则 r 满足我们的要求.

在分配律后增加最后一个例子:

例 1.5.7 实数的小数表示.

在中学数学中经常用十进制小数表示实数, 它的好处是比 Dedekind 分割更加直观. 不过, 既然我们已经构造出了实数并定义了它的各种运算, 自然也可以将实数表示为小数. 设 p 是大于 1 的固定正整数, 我们以 $0 < \alpha < 1$ 为例讨论其 p 进制小数表示. 记 $\alpha_1 = \lfloor p\alpha \rfloor$, 由 $0 < p\alpha < p$ 可知 $0 \le \alpha_1 \le p - 1$, 且

$$0 \le \alpha - \frac{\alpha_1}{p} = \frac{p\alpha - \lfloor p\alpha \rfloor}{p} < \frac{1}{p}.$$

再记 $\alpha_2 = \lfloor p^2\alpha - p\alpha_1 \rfloor$, 则 $0 \le \alpha_2 \le p - 1$, 且

$$0 \le \alpha - \frac{\alpha_1}{p} - \frac{\alpha_2}{p^2} < \frac{1}{p^2}.$$

以此类推, 可以归纳地定义一列非负整数 $\{\alpha_n\}$, 使得 $0 \le \alpha_n \le p - 1$, 且

$$0 \le \alpha - \sum_{k=1}^{n} \frac{\alpha_k}{p^k} < \frac{1}{p^n}.$$

记 $\alpha =_p 0.\alpha_1\alpha_2\cdots\alpha_n\cdots$, 称为 α 的 p 进制小数表示.

去掉 Archimedes 原理等.

第二章　极限

2.1　数列极限

给定数列 $\{a_n\}$, 我们想了解当 n 变化时 a_n 如何变化. 比如, 当 n 变大时, 看看 a_n 是否会逐渐趋于某个固定的数. 所谓 a_n 逐渐趋于某个固定的数 α, 是指 a_n 和 α 之间的误差 $a_n-\alpha$ 逐渐趋于零. 此时, $\{a_n\}$ 应该具有如下性质:

$(*)$ 给定正数 ε, 从某项开始, $\{a_n\}$ 满足 $|a_n-\alpha|<\varepsilon$.

当然, 如果 $\{a_n\}$ 只是对某一个 ε 满足上述性质, 那还不足以说明它逐渐趋于 α. 比如, 考虑 $a_n=\frac{(-1)^n}{2}$, 取 $\alpha=0$, $\varepsilon=1$, 则 $|a_n-\alpha|<\varepsilon$ 总成立, 但从直观上看 a_n 并不是逐渐趋于 α 的.

关键的观察在于, 若 a_n 逐渐趋于 α, 则对每一个正数 ε, $(*)$ 都应该成立才行! 总结上述讨论, 我们就得到了数列极限的概念.

定义 2.1.1 (数列极限)**.** *设 $\{a_n\}$ 为数列, $\alpha\in\mathbb{R}$. 如果任给 $\varepsilon>0$, 均存在正整数 $N=N(\varepsilon)$, 使得当 $n>N$ 时, $|a_n-\alpha|<\varepsilon$ 总成立, 则称 $\{a_n\}$ 收敛于 α, 或称 $\{a_n\}$ 以 α 为极限, 记为 $a_n\to\alpha$ $(n\to\infty)$.*

命题 2.1.1. *如果数列 $\{a_n\}$ 既收敛于 α, 也收敛于 β, 则 $\alpha=\beta$.*

证明. 按照定义, 任给 $\varepsilon>0$, 存在 N_1, N_2, 使得当 $n>N_1$ 时 $|a_n-\alpha|<\varepsilon$; 当 $n>N_2$ 时 $|a_n-\beta|<\varepsilon$. 因此, 取 $n>\max\{N_1,N_2\}$, 则有

$$|\alpha-\beta|\le|a_n-\alpha|+|a_n-\beta|<2\varepsilon.$$

由 ε 的任意性可知 $\alpha=\beta$.

注. 若 $\alpha\neq\beta$, 则选取 $\varepsilon=|\alpha-\beta|/2$ 就可导出矛盾.

此命题告诉我们, 若数列收敛, 则其极限具有唯一性. 我们常把数列 $\{a_n\}$ 的极限记为 $\lim\limits_{n\to\infty}a_n$.

例 2.1.1 用定义证明: $\lim\limits_{n\to\infty}\frac{1}{n}=0$.

证明 任给 $\varepsilon>0$, 取正整数 $N>\frac{1}{\varepsilon}$, 则当 $n>N$ 时, 有

$$\left|\frac{1}{n}-0\right|=\frac{1}{n}<\frac{1}{N}<\varepsilon,$$

由数列极限的定义可知 $\lim\limits_{n\to\infty}\frac{1}{n}=0$.

注. 给定了 ε 以后, 只要能找到一个满足要求的 N 就可以. 对于本例来说, N 可以取为 $\left\lfloor\frac{1}{\varepsilon}\right\rfloor+1$.

如果数列不以任何实数为极限, 则称它不收敛或发散.

例 2.1.2: 原**例 2.1.1**.

例 2.1.3 设 $|q|<1$, 则 $\lim\limits_{n\to\infty}q^n=0$.

证明 $q=0$ 时结论显然成立. 当 $0<|q|<1$ 时, 记 $p=\frac{1}{q}$, 则 $|p|>1$. 由 Bernoulli 不等式可得

$$|p|^n\geq 1+n\big(|p|-1\big)>n\big(|p|-1\big),$$

任给 $\varepsilon>0$, 取 $N>\frac{1}{\varepsilon(|p|-1)}$, 当 $n>N$ 时就有

$$|q^n-0|=|q|^n=\frac{1}{|p|^n}<\frac{1}{n(|p|-1)}<\frac{1}{N(|p|-1)}<\varepsilon,$$

由数列极限的定义可知 $\lim\limits_{n\to\infty}q^n=0$.

例 2.1.4 设 m 为大于 1 的正整数, 则 $\lim\limits_{n\to\infty}\frac{1}{\sqrt[m]{n}}=0$.

证明 任给 $\varepsilon>0$, 取 $N>\frac{1}{\varepsilon^m}$, 则当 $n>N$ 时

$$0<\frac{1}{\sqrt[m]{n}}<\frac{1}{\sqrt[m]{N}}<\varepsilon,$$

由数列极限的定义可知 $\lim\limits_{n\to\infty}\frac{1}{\sqrt[m]{n}}=0$.

例 2.1.5: 原**例 2.1.3**, 里面引用的**例 2.1.2** 应改为**例 2.1.3**.

例 2.1.6 设 m 为大于 1 的正整数, 则 $\lim\limits_{n\to\infty}\left[\sqrt[m]{n+1}-\sqrt[m]{n}\right]=0$.

证明 当 $n\geq 1$ 时, 有

$$0<\sqrt[m]{n+1}-\sqrt[m]{n}=\sqrt[m]{n}\left[\sqrt[m]{1+\frac{1}{n}}-1\right]<\sqrt[m]{n}\,\frac{1}{n}\leq\frac{1}{\sqrt{n}},$$

由夹逼原理可知 $\lim\limits_{n\to\infty}\left[\sqrt[m]{n+1}-\sqrt[m]{n}\right]=0$.

例 2.1.7 设 $a>0$, 证明 $\lim\limits_{n\to\infty}\sqrt[n]{a}=1$.

证明 在 (1.1) 式中取 $u=\sqrt[n]{a}$ 可得

$$\frac{a-1}{na}\leq\sqrt[n]{a}-1\leq\frac{a-1}{n}, \tag{2.1}$$

由夹逼原理可知 $\lim\limits_{n\to\infty}\sqrt[n]{a}=1$.

例 2.1.8: 原**例 2.1.9**.

例 2.1.9 设 $q>1$, 则 $\lim\limits_{n\to\infty}\frac{n}{q^n}=0$.

证明 注意到 $q^n=\big[1+(q-1)\big]^n$, 由 (1.14) 式可得

$$q^n\geq 1+n(q-1)+\frac{1}{2}n(n-1)(q-1)^2>\frac{1}{2}n(n-1)(q-1)^2,$$

当 $n>1$ 时, 就有

$$0<\frac{n}{q^n}<\frac{2}{(n-1)(q-1)^2},$$

由夹逼原理可知 $\lim\limits_{n\to\infty}\frac{n}{q^n}=0$.

例 2.1.10 证明 $\lim\limits_{n\to\infty}\sqrt[n]{n}=1$.

证明 任给 $\varepsilon>0$, 由前例可知 $\lim\limits_{n\to\infty}\frac{n}{(1+\varepsilon)^n}=0$. 这说明存在 N, 使得当 $n>N$ 时 $\frac{n}{(1+\varepsilon)^n}<1$. 于是当 $n>N$ 时

$$1\le\sqrt[n]{n}<1+\varepsilon,$$

由数列极限的定义可得 $\lim\limits_{n\to\infty}\sqrt[n]{n}=1$.

例 2.1.11 设 $\lim\limits_{n\to\infty}a_n=\alpha$. 记 $b_n=\frac{1}{n}\sum\limits_{k=1}^{n}a_k$, 则 $\lim\limits_{n\to\infty}b_n=\alpha$.

证明 先设 $\alpha=0$, 则任给 $\varepsilon>0$, 存在 N, 当 $k>N$ 时 $|a_k|<\frac{1}{2}\varepsilon$. 记 $C=\sum\limits_{k=1}^{N}|a_k|$. 当 $n>N$ 时, 由绝对值不等式可得

$$|b_n|\le\frac{1}{n}\sum_{k=1}^{n}|a_k|=\frac{C}{n}+\frac{1}{n}\sum_{k=N+1}^{n}|a_k|<\frac{C}{n}+\frac{1}{n}\sum_{k=N+1}^{n}\frac{1}{2}\varepsilon<\frac{C}{n}+\frac{1}{2}\varepsilon.$$

取 $N'>\max\{N,\ 2C\varepsilon^{-1}\}$, 当 $n>N'$ 时, 由上式可得 $|b_n|<\varepsilon$. 由数列极限的定义即知 $\lim\limits_{n\to\infty}b_n=0$.

一般地, 如果 $\lim\limits_{n\to\infty}a_n=\alpha$, 则 $\lim\limits_{n\to\infty}(a_n-\alpha)=0$. 由前一段的证明可知 $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n}(a_k-\alpha)=0$, 即 $\lim\limits_{n\to\infty}b_n=\alpha$.

注. 反之, $\{b_n\}$ 收敛时 $\{a_n\}$ 未必收敛. 比如当 $a_n=(-1)^n$ 时,

$$b_n=\begin{cases}-\frac{1}{n}, & n\text{ 为奇数},\\ 0, & n\text{ 为偶数}.\end{cases}$$

显然, $\{b_n\}$ 收敛于 0, 但 $\{a_n\}$ 发散. 进一步的结论见 §2.3 **例 2.3.2**.

例 2.1.12 每一个实数都是有理数列的极限.

证明 设 α 为实数. 当 n 为正整数时, 由**例 1.2.4** 可知

$$\frac{\lfloor n\alpha\rfloor}{n}\le\alpha<\frac{\lfloor n\alpha\rfloor+1}{n}.$$

记 $a_n=\frac{\lfloor n\alpha\rfloor}{n}$, 则 $\{a_n\}$ 为有理数列, 且

$$\alpha-\frac{1}{n}<a_n\le\alpha,$$

由夹逼原理可知 $\lim\limits_{n\to\infty}a_n=\alpha$.

原**例 2.1.12** 去掉.

例 2.1.13 设 m 为正整数, 则 $\lim\limits_{n\to\infty}\sqrt[m]{n}=+\infty$.

证明 任给 $\alpha > 0$, 取 $N > \alpha^m$, 则当 $n > N$ 时, 有

$$\sqrt[m]{n} > \sqrt[m]{N} > \alpha,$$

由定义可知 $\lim\limits_{n\to\infty} \sqrt[m]{n} = +\infty$.

例 2.1.14 设 $|q| > 1$, 研究数列 $\{q^n\}$ 的敛散性.

解 当 $q > 1$ 时, 由**例 2.1.9**的证明可知 $\lim\limits_{n\to\infty} q^n = +\infty$. 当 $q < -1$ 时, $|q|^n \to +\infty$, 因此 $q^n \to \infty\ (n\to\infty)$.

例 2.1.15 设 m 为大于 1 的正整数, $q > 1$, 则 $\lim\limits_{n\to\infty} \frac{n^m}{q^n} = 0$.

证明 当 $n > 1$ 时, 由**例 2.1.9**的证明可得

$$0 < \frac{n^m}{q^n} = \left[\frac{n}{\left(\sqrt[m]{q}\,\right)^n}\right]^m < \frac{2^m}{(n-1)^m\left(\sqrt[m]{q}-1\right)^{2m}},$$

由夹逼原理即得欲证结论.

例 2.1.16 设 $q \in \mathbb{R}$, 则 $\lim\limits_{n\to\infty} \frac{q^n}{n!} = 0$.

证明 取 $m = \lfloor |q| \rfloor$, 则

$$\frac{|q|}{m+1} < 1,\ \frac{|q|}{m+2} < 1,\ \cdots.$$

这说明, 当 $n > m$ 时,

$$\left|\frac{q^n}{n!}\right| = \frac{|q|^m}{m!}\frac{|q|}{m+1}\cdots\frac{|q|}{n-1}\frac{|q|}{n} \le \frac{|q|^m}{m!}\frac{|q|}{n},$$

由夹逼原理即得欲证结论.

习题 2.1 调整:

11. 去掉 (1);

12. 设 $\lim\limits_{n\to\infty} na_n = 0$, 利用 Bernoulli 不等式证明: $\lim\limits_{n\to\infty}(1+a_n)^n = 1$.

13. 设 m 为正整数, $a_n \ge 0$ 且 $\lim\limits_{n\to\infty} a_n = \alpha$. 证明: $\lim\limits_{n\to\infty} \sqrt[m]{a_n} = \sqrt[m]{\alpha}$.

2.2 单调数列的极限

2.2.1 确界原理和单调数列

例2.2.1换掉:

例 2.2.1 设 $0 \le c \le 1$, $a_1 = 0$, 当 $n \ge 1$ 时 $a_{n+1} = \frac{1}{2}(c + a_n^2)$, 研究数列 $\{a_n\}$ 的极限.

解. 利用归纳法可知 $0 \le a_n \le 1$. 注意到 $a_2 = \frac{1}{2}c \ge a_1$. 当 $n > 1$ 时

$$a_{n+1} - a_n = \frac{1}{2}(c + a_n^2) - \frac{1}{2}(c + a_{n-1}^2) = \frac{1}{2}(a_n - a_{n-1})(a_n + a_{n-1}),$$

由此再用归纳法即知 $\{a_n\}$ 单调递增, 从而收敛. 其极限记为 α, 则 $0 \le \alpha \le 1$, 且

$$\alpha = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{c+a_n^2}{2} = \frac{c+\alpha^2}{2},$$

由此解出 $\alpha = 1-\sqrt{1-c}$.

注. 当 n 充分大时 $\sqrt{1-c} \approx 1-a_n$, 古巴比伦人曾用此来做开方运算.

2.2.2 自然常数和自然对数

下面我们讨论本课程中的一个重要极限. 在中学数学中我们知道, 利用对数可以把求幂和开方两种运算转换为乘法和除法, 也可以把乘法和除法转换为加法和减法. 在实际应用中往往利用事先编制好的对数表来做近似计算. 那么, 选取哪一个数作为基底来编制对数表呢? 无论是什么基底, 1 的对数总是 0. 人们希望, 若 ω 是一个很小的改变量, 则 $1+\omega$ 的对数约等于 ω. 设 q 是满足这个条件的基底, 则

$$q^\omega \approx 1+\omega. \tag{2.2}$$

设 k 整数, 当 $|k|$ 充分大时, 取 $\omega = \frac{1}{k}$, 则由上式可得

$$q \approx \left(1+\frac{1}{k}\right)^k.$$

受此启发, 我们来考虑下面的两个重要数列

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

我们先来说明这两个数列是单调有界数列. 为此, 先将 Bernoulli 不等式做一点推广. 设 $m>n$, $x \ge -1$. 在 (1.11) 中取 $u = \sqrt[n]{1+x}$ 可得

$$(1+x)^{\frac{m}{n}} \ge 1+\frac{m}{n}x, \tag{2.3}$$

等号成立当且仅当 $x=0$. 于是

$$\left(1+\frac{1}{m}\right)^{\frac{m}{n}} > 1+\frac{m}{n}\frac{1}{m} = 1+\frac{1}{n},$$

即 $a_m > a_n$. 同理有

$$\left(1-\frac{1}{m}\right)^{\frac{m}{n}} > 1-\frac{m}{n}\frac{1}{m} = 1-\frac{1}{n},$$

即 $\{b_n\}$ 严格单调递减. 总之可得

$$2 = a_1 \le a_n < a_{n+1} < b_{n+1} < b_n \le b_1 = 4.$$

由定理 2.2.1 可知 $\{a_n\}$ 和 $\{b_n\}$ 均收敛, 且它们的极限相等, 记为 e, 称为自然常数. 由严格单调性可得

$$\left(1+\frac{1}{n}\right)^n < \mathrm{e} < \left(1+\frac{1}{n}\right)^{n+1},\ \ \forall\, n \ge 1. \tag{2.4}$$

在 §5.6 节中我们将证明自然常数是无理数, 计算表明(关于近似计算, 请参考本书 §8.4 节)

$$\mathrm{e} = 2.718\,281\,828\,459\,0\cdots.$$

例 2.2.4 记 $e_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$, 其中 0! 规定为 1, 则 $\{e_n\}$ 收敛于 e.

证明 当 $n=1$ 时 $e_1 = 2 = a_1$. 当 $n>1$ 时, 根据 Newton 二项式展开可得

$$a_n = \sum_{k=0}^{n} C_n^k \frac{1}{n^k} = 1 + \frac{1}{1!} + \sum_{k=2}^{n} \Big(1-\frac{1}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big)\frac{1}{k!}, \tag{2.5}$$

这说明 $a_n < e_n$. 另一方面, 利用 (1.3) 式可得

$$\Big(1-\frac{1}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big) \geq 1 - \sum_{i=1}^{k-1}\frac{i}{n} = 1 - \frac{k(k-1)}{2n}.$$

代入 (2.5) 式可得

$$a_n \geq e_n - \frac{1}{2n}\sum_{k=2}^{n}\frac{1}{(k-2)!} > e_n - \frac{1}{2n}e_n,$$

于是

$$a_n < e_n < \frac{2n}{2n-1}a_n,$$

根据夹逼原理即知 $\{e_n\}$ 收敛于 e.

注 因为 $\{e_n\}$ 是严格单调递增数列, 实际上就有 $a_n < e_n < e$.

例 2.2.5

下面我们来定义以自然常数为基底的对数函数. 设 $x>0$, 我们要解以 y 为未知量的方程 $\mathrm{e}^y = x$, 其中 e^x 是预想中的以自然常数为基底的指数函数. 设 n 为充分大的正整数, 由 (2.2) 式可得

$$\sqrt[n]{x} = \mathrm{e}^{\frac{y}{n}} \approx 1 + \frac{y}{n},$$

这说明 $y \approx n\big(\sqrt[n]{x}-1\big)$.

由 (2.1) 可知

$$\frac{x-1}{x} \leq n\big(\sqrt[n]{x}-1\big) \leq x-1. \tag{2.6}$$

这说明 $\{n(\sqrt[n]{x}-1)\}$ 是有界数列. 进一步, 设 $m>n$, 在 (1.11) 中取 $u = \sqrt[mn]{x}$ 可得

$$n\big(\sqrt[n]{x}-1\big) \geq m\big(\sqrt[m]{x}-1\big),$$

这说明数列 $\{n(\sqrt[n]{x}-1)\}$ 单调递减, 再结合有界性就知道它是收敛的.

命题 2.2.2 当 $x>0$ 时, 记 $\ln x = \lim\limits_{n\to\infty} n(\sqrt[n]{x}-1)$, 称为自然对数函数, 它有如下简单性质:

(1) $\ln 1 = 0$, $\ln \mathrm{e} = 1$;

(2) $\frac{x-1}{x} \le \ln x \le x-1$;

(3) $\ln(xx') = \ln x + \ln x'$;

(4) $\ln x$ 是严格单调递增函数.

证明. (1) 只需证明 $\ln \mathrm{e} = 1$. 由 (2.4) 式可得

$$1 < n(\sqrt[n]{\mathrm{e}} - 1) < n\left[\left(1+\frac{1}{n}\right)\sqrt[n]{1+\frac{1}{n}} - 1\right].$$

由 Bernoulli 不等式或 (2.6) 式可知 $\sqrt[n]{1+\frac{1}{n}} \le 1 + \frac{1}{n^2}$. 代入上式可得

$$1 < n(\sqrt[n]{\mathrm{e}} - 1) < 1 + \frac{1}{n} + \frac{1}{n^2}.$$

由夹逼原理即得 $\ln \mathrm{e} = 1$.

(2) 这可以由 (2.6) 式立即得到.

(3) 根据定义, 我们有

$$\begin{aligned}\ln(xx') &= \lim_{n\to\infty} n(\sqrt[n]{xx'} - 1)\\ &= \lim_{n\to\infty} n(\sqrt[n]{x} - 1)\sqrt[n]{x'} + \lim_{n\to\infty} n(\sqrt[n]{x'} - 1)\\ &= \ln x + \ln x'.\end{aligned}$$

(4) 由 (3) 和 (1) 可知 $\ln \frac{1}{x} + \ln x = \ln 1 = 0$. 这说明

$$\ln x' - \ln x = \ln x' + \ln \frac{1}{x} = \ln \frac{x'}{x}.$$

当 $x' > x$ 时, 再利用 (2) 即知 $\ln x' > \ln x$.

例 2.2.6 平均值不等式回顾.

设 $a_1, a_2, \cdots, a_n \ge 0$, 记

$$A = \frac{a_1 + a_2 + \cdots + a_n}{n}, \quad G = (a_1 a_2 \cdots a_n)^{\frac{1}{n}},$$

A, G 分别称为 $\{a_i\}_{i=1}^n$ 的算术平均与几何平均. 在例1.2.1中我们已经证明过 $A \ge G$. 现在我们给出另一证明.

不妨设 $G > 0$. 记 $a_i' = \frac{a_i}{G}$, 则 $a_1' a_2' \cdots a_n' = 1$. 利用不等式 $x \ge 1 + \ln x$ 可得

$$\frac{1}{n}\sum_{i=1}^n a_i' \ge \frac{1}{n}\sum_{i=1}^n (1 + \ln a_i') = 1 + \frac{1}{n}\ln(a_1' a_2' \cdots a_n') = 1,$$

这说明 $A \ge G$ 的确成立.

Euler 常数

例 2.2.7

2.2.3 自然指数函数和幂函数

下面我们考虑自然对数函数的反函数. 给定实数 x, 要找 $y>0$, 使得 $\ln y=x$. 根据 $\ln y$ 的定义, 当 n 充分大时,

$$n(\sqrt[n]{y}-1)\approx \ln y=x,$$

因此 $y\approx\big(1+\frac{x}{n}\big)^n$. 我们有

引理 2.2.3 设 $x\in\mathbb{R}$, 则数列 $\{\big(1+\frac{x}{n}\big)^n\}$ 收敛.

证明. 与之前的讨论类似, 当 $n\geq -x$ 时, 利用 (2.3) 式可以得出该数列单调递增. 我们再来说明该数列有上界. 取正整数 $m\geq x$, 由 Bernoulli 不等式可得

$$1+\frac{x}{n}\leq 1+\frac{m}{n}\leq\Big(1+\frac{1}{n}\Big)^m.$$

于是当 $n\geq -x$ 时, 有

$$\Big(1+\frac{x}{n}\Big)^n\leq\Big(1+\frac{1}{n}\Big)^{nm}<\mathrm{e}^m,$$

引理得证.

命题 2.2.4 当 $x\in\mathbb{R}$ 时, 记 $\exp(x)=\lim\limits_{n\to\infty}\big(1+\frac{x}{n}\big)^n$, 称为自然指数函数, 它有如下简单性质:

(1) $\exp(0)=1$, $\exp(1)=\mathrm{e}$;

(2) $\exp(x)\geq 1+x$;

(3) $\exp(x)\exp(x')=\exp(x+x')$;

(4) $\exp(x)$ 是严格单调递增函数.

证明. (1) 按照定义都是显然的.

(2) 当 $n\geq -x$ 时, 根据 Bernoulli 不等式可得

$$\Big(1+\frac{x}{n}\Big)^n\geq 1+n\,\frac{x}{n}=1+x,$$

取极限即得欲证不等式.

(3) 在不等式 (1.10) 中取 $u=\big(1+\frac{x}{n}\big)\big(1+\frac{x'}{n}\big)$, $v=1+\frac{x+x'}{n}$, 令 $n\to\infty$ 即可.

(4) 由 (1) 和 (3) 可知 $\exp(-x)\exp(x)=1$. 特别地, $\exp(x)\neq 0$. 再次利用 (3) 可知 $\exp(x)=\big[\exp(x/2)\big]^2>0$. 当 $x'>x$ 时, 利用 (2) 可得

$$\frac{\exp(x')}{\exp(x)}=\exp(x'-x)\geq 1+(x'-x)>1,$$

这说明 $\exp(x)$ 是严格单调递增函数.

下面的结果说明自然对数函数和自然指数函数互为反函数.

命题 2.2.5 当 $x\in\mathbb{R}$ 时 $\ln\exp(x)=x$; 当 $x>0$ 时 $\exp(\ln x)=x$.

证明. 先证明 $\ln\exp(x)=x$. 根据引理 2.2.3 可知, 当 n 充分大时 $\big(1+\frac{x}{n}\big)^n\leq\exp(x)$. 这也说明

$$\exp(x)=\frac{1}{\exp(-x)}\leq\Big(1-\frac{x}{n}\Big)^{-n}.$$

于是

$$x \le n\big(\sqrt[n]{\exp(x)} - 1\big) \le \frac{x}{1 - \frac{x}{n}},$$

由夹逼原理即得 $\ln \exp(x) = x$.

当 $x > 0$ 时, 记 $x' = \exp(\ln x)$, 则 $x' > 0$. 根据刚才的证明, $\ln x' = \ln x$. 由自然对数函数的严格单调性即知 $x' = x$.

我们再简单地讨论一下一般的指数函数和对数函数. 设 $a > 0$, 当 $x \in \mathbb{R}$ 时, 定义 $a^x = \exp(x \ln a)$, 称为以 a 为基底的指数函数. 特别地, $\mathrm{e}^x = \exp(x)$. 利用命题 2.2.4 中的 (1) 和 (3) 以及归纳法不难说明, 当 x 是有理数时, $\exp(x)$ 确实等于 e 的 x 次幂.

设 $a > 0$, 当 $x > 0$ 时, 定义 $\log_a x = \frac{\ln x}{\ln a}$, 称为以 a 为基底的对数函数. 特别地, $\log_{\mathrm{e}} x = \ln x$. 以 10 为基底的对数函数也常记为 $\lg x$.

类似地可以讨论幂函数. 设 α 是固定的数, 我们想要定义 x 的 α 次方, 记为 x^α. 分情况讨论:

(1) 当 $x > 0$ 时, 定义 $x^\alpha = \exp(\alpha \ln x)$;

(2) 当 $x = 0$, $\alpha > 0$ 时, 定义 $x^\alpha = 0$;

(3) 当 $x < 0$ 时, 定义 $x^\alpha = (-1)^\alpha(-x)^\alpha$, 其中 α 只能是有理数, 将它表示为既约分数 $\alpha = \frac{p}{q}$, 其中 p, q 为互素整数, 则

$$(-1)^\alpha = \begin{cases} -1, & p, q \text{ 均为奇数}, \\ 1, & p \text{ 为偶数}, q \text{ 为奇数}. \end{cases}$$

当 q 为偶数时 $(-1)^\alpha$ 暂时不能定义.

习题 2.2 调整:

1(3) 与 2 互换位置;

5, 8, 9 去掉;

7. (1), (2) 改为:

(1) $a_1 > 2$, $a_{n+1} = \frac{a_n^2}{2(a_n - 1)}$; (2) $a_1 > \sqrt{3}$, $a_{n+1} = \frac{2a_n^3}{3(a_n^2 - 1)}$;

11, 12, 13 放到习题 2.3 中.

补充习题:

9. 本题利用数列极限研究自然对数函数.

(1) 设 $\beta \in \mathbb{R}$, $a_1 = \mathrm{e}^{\lfloor \beta \rfloor}$. 当 $n \ge 1$ 时 $a_{n+1} = a_n[1 + \beta - \ln a_n]$. 证明 $\{a_n\}$ 单调递增且有上界.

(2) 利用 (1) 说明自然对数函数 $\ln : (0, +\infty) \to (-\infty, +\infty)$ 是满射.

10. 当 $x > 0$ 时, 记 $\phi(x) = \ln x - x + 1$.

(1) 当 $x \ne y$ 时, 证明: $\phi\big(\frac{x+y}{2}\big) > \frac{1}{2}[\phi(x) + \phi(y)]$;

(2) 利用 $\phi(x) \le 0$, $\phi(1) = 0$ 以及 (1) 证明: 当 $x > 0$, $x \ne 1$ 时 $\frac{x-1}{x} < \ln x < x - 1$;

(3) 模仿上述做法, 证明: 当 $x \neq 0$ 时 $\mathrm{e}^x > 1 + x$.

11. 设 $a_1, a_2, \cdots, a_n \geq 0$, $\lambda_1, \lambda_2, \cdots, \lambda_n \geq 0$ 且 $\sum\limits_{i=1}^{n} \lambda_i = 1$. 记

$$A = \sum_{i=1}^{n} \lambda_i a_i,\ \ G = \prod_{i=1}^{n} a_i^{\lambda_i}\ \ (\text{规定 } 0^0 = 1),$$

A, G 分别称为 $\{a_i\}_{i=1}^n$ 的加权算术平均与加权几何平均. 证明加权平均值不等式: $A \geq G$, 并用上一题研究等号成立的条件.

12. 设数列 $\{a_n\}$ 由函数 ψ 迭代而来, 即 $a_{n+1} = \psi(a_n)$. 证明:

(1) 若 ψ 为单调递增函数, 则 $\{a_n\}$ 必为单调数列;

(2) 若 ψ 为单调递减函数, 则 $\{a_n\}$ 的奇子列和偶子列分别为单调数列.

13. 设 f 是在 $[a, b]$ 中定义的函数, 且存在常数 $L > 0$, 使得

$$|f(x) - f(y)| \leq L\,|x - y|,\ \ \forall\, x, y \in [a, b].$$

设 $f(a) < \mu < f(b)$, 令 $c_1 = a$, 当 $n \geq 1$ 时

$$c_{n+1} = c_n + \frac{1}{L}\big[\mu - f(c_n)\big].$$

证明 $\{c_n\}$ 为 $[a, b]$ 中单调递增数列, 且其极限 c 满足 $f(c) = \mu$.

14. 设 f 同上一题. 当 $n \geq 1$ 时, 记 $c_n = \sup\big\{x \in [a, b] \mid f(x) > M - \frac{1}{n}\big\}$, 其中 $M = \sup\{f(x) \mid x \in [a, b]\}$. 证明: (1) $f(c_n) \geq M - \frac{1}{n}$; (2) $\{c_n\}$ 单调递减, 且其极限 c 满足 $f(c) = M$.

2.3 上极限和下极限

原定理 2.2.3 变为定理 2.3.1.

我们知道, 数列极限具有四则运算性质、保序性质等. 我们来看看下极限和上极限是否具有类似的性质. 考虑如下数列

$$\{a_n\} = \{0, 1, 0, 1, \cdots, 0, 1, \cdots\},\quad \{b_n\} = \{1, 0, 1, 0, \cdots, 1, 0, \cdots\}.$$

显然, $\varliminf\limits_{n\to\infty} a_n = \varliminf\limits_{n\to\infty} b_n = 0$, $\varlimsup\limits_{n\to\infty} a_n = \varlimsup\limits_{n\to\infty} b_n = 1$. 而 $a_n + b_n \equiv 1$, 故

$$\varliminf_{n\to\infty}(a_n + b_n) \neq \varliminf_{n\to\infty} a_n + \varliminf_{n\to\infty} b_n,\ \ \varlimsup_{n\to\infty}(a_n + b_n) \neq \varlimsup_{n\to\infty} a_n + \varlimsup_{n\to\infty} b_n.$$

不过, 我们有

命题 2.3.2 设 $\{a_n\}$ 为有界数列.

(1) 若 $c \in \mathbb{R}$, 则 $\varliminf\limits_{n\to\infty}(c + a_n) = c + \varliminf\limits_{n\to\infty} a_n$, $\varlimsup\limits_{n\to\infty}(c + a_n) = c + \varlimsup\limits_{n\to\infty} a_n$.

(2) 若 $\lambda > 0$, 则 $\varliminf\limits_{n\to\infty} \lambda a_n = \lambda \varliminf\limits_{n\to\infty} a_n$, $\varlimsup\limits_{n\to\infty} \lambda a_n = \lambda \varlimsup\limits_{n\to\infty} a_n$; 若 $\lambda < 0$, 则 $\varliminf\limits_{n\to\infty} \lambda a_n = \lambda \varlimsup\limits_{n\to\infty} a_n$, $\varlimsup\limits_{n\to\infty} \lambda a_n = \lambda \varliminf\limits_{n\to\infty} a_n$.

(3) 若 $\{a_n\}$ 有正下界, 则 $\varliminf\limits_{n\to\infty} a_n \varlimsup\limits_{n\to\infty} \frac{1}{a_n} = 1$.

证明 (1), (2) 留作练习, 我们来证明 (3). 记 $\alpha_n = \inf\{a_k \mid k \geq n\}$, $\beta_n = \sup\left\{\frac{1}{a_k} \mid k \geq n\right\}$. 由题设可知 $\alpha_n > 0$, 且当 $k \geq n$ 时 $a_k \geq \alpha_n$, 此时 $\frac{1}{a_k} \leq \frac{1}{\alpha_n}$. 由上确界为最小上界可知 $\beta_n \leq \frac{1}{\alpha_n}$. 同理, 当 $k \geq n$ 时 $\frac{1}{a_k} \leq \beta_n$, 此时 $a_k \geq \frac{1}{\beta_n}$, 由下确界为最大下界可知 $\alpha_n \geq \frac{1}{\beta_n}$. 总之, $\alpha_n\beta_n = 1$. 令 $n \to \infty$ 即得欲证等式.

例 2.3.1 设 $a_0 > 0$, $a_{n+1} = \frac{1}{1+a_n}$, 研究 $\{a_n\}$ 的极限.

证明 用归纳法可见 $a_n > 0$, 且当 $n \geq 1$ 时 $0 < a_n < 1$, 当 $n \geq 2$ 时 $\frac{1}{2} < a_n < 1$. 记 $\{a_n\}$ 的上下极限分别为 β, α. 利用上下极限的基本性质可得

$$\beta = \varlimsup_{n\to\infty} a_n = \varlimsup_{n\to\infty} a_{n+1} = \frac{1}{1 + \varliminf\limits_{n\to\infty} a_n} = \frac{1}{1+\alpha}.$$

同理可得 $\alpha = \frac{1}{1+\beta}$, 由此可以解出 $\alpha = \beta = \frac{\sqrt{5}-1}{2}$ (舍去负根), 于是 $\{a_n\}$ 收敛于黄金分割比.

命题 2.3.3 设 $\{a_n\}$, $\{b_n\}$ 为有界数列.

(1) 若 $n > N$ 时 $a_n \geq b_n$, 则 $\varliminf\limits_{n\to\infty} a_n \geq \varliminf\limits_{n\to\infty} b_n$, $\varlimsup\limits_{n\to\infty} a_n \geq \varlimsup\limits_{n\to\infty} b_n$.

(2) $\varliminf\limits_{n\to\infty} (a_n + b_n) \geq \varliminf\limits_{n\to\infty} a_n + \varliminf\limits_{n\to\infty} b_n$, $\varlimsup\limits_{n\to\infty} (a_n + b_n) \leq \varlimsup\limits_{n\to\infty} a_n + \varlimsup\limits_{n\to\infty} b_n$.

(3) $\varliminf\limits_{n\to\infty} (a_n + b_n) \leq \varlimsup\limits_{n\to\infty} a_n + \varliminf\limits_{n\to\infty} b_n \leq \varlimsup\limits_{n\to\infty} (a_n + b_n)$.

证明 (1) 暂时固定 $n > N$. 当 $k \geq n$ 时 $a_k \geq b_k \geq \underline{b}_n$. 由下确界为最大下界可得 $\underline{a}_n \geq \underline{b}_n$. 再令 $n \to \infty$ 即得欲证结论. 上极限的情形可类似证明.

(2) 对 $\underline{a}_n + \underline{b}_n \leq a_n + b_n \leq \overline{a}_n + \overline{b}_n$ 利用 (2) 即可.

(3) 由 (2) 可得

$$\varliminf_{n\to\infty} b_n = \varliminf_{n\to\infty} \left[(b_n + a_n) + (-a_n)\right] \geq \varliminf_{n\to\infty} (a_n + b_n) + \varliminf_{n\to\infty} (-a_n),$$

再用前一命题中的 (2) 即可得到左侧欲证不等式, 右侧同理可证.

例 2.3.2 设数列 $\{a_n\}$ 满足以下条件:

(i) 任给 $\varepsilon > 0$, 均存在 $\delta > 0$ 以及 $N \geq 1$, 使得当 $N < n \leq m \leq (1+\delta)n$ 时 $a_m - a_n \geq -\varepsilon$;

(ii) $\lim\limits_{n\to\infty} b_n = \alpha$, 其中 $b_n = \dfrac{1}{n}(a_1 + a_2 + \cdots + a_n)$,

则 $\lim\limits_{n\to\infty} a_n = \alpha$.

证明 记 $K = \lfloor(1+\delta)n\rfloor$. 由 (i) 可知, 当 $N < n \leq m \leq K$ 时 $a_m \geq a_n - \varepsilon$. 此时

$$\frac{Kb_K - nb_n}{K-n} = \frac{1}{K-n}\sum_{m=n+1}^{K} a_m \geq a_n - \varepsilon.$$

令 $n \to \infty$, 由 $\dfrac{n}{K} \to \dfrac{1}{1+\delta}$ 和 (ii) 可得 $\alpha \geq \varlimsup\limits_{n\to\infty} a_n - \varepsilon$, 由 ε 的任意性即得 $\varlimsup\limits_{n\to\infty} a_n \leq \alpha$.

另一方面, 记 $L=\left\lfloor\frac{m}{1+\delta}\right\rfloor$. 当 $m>(1+\delta)N$ 且 $L<n\le m$ 时, $n\le m\le(1+\delta)n$. 此时由 (i) 可知 $a_n\le a_m+\varepsilon$, 于是

$$\frac{mb_m-Lb_L}{m-L}=\frac{1}{m-L}\sum_{n=L+1}^{m}a_n\le a_m+\varepsilon.$$

令 $m\to\infty$ 可得 $\alpha\le\varliminf_{m\to\infty}a_m+\varepsilon$. 由 ε 的任意性即得 $\varliminf_{m\to\infty}a_m\ge\alpha$.

综上可得 $\lim_{n\to\infty}a_n=\alpha$.

注 如果 $\{n(a_n-a_{n-1})\}$ 有下界, 则 $\{a_n\}$ 满足上述条件 (i). 事实上, 设 $n(a_n-a_{n-1})\ge -C$, 其中 $C>0$, 则当 $m>n$ 时, 有

$$a_m-a_n=\sum_{k=n+1}^{n}(a_k-a_{k-1})\ge\sum_{k=n+1}^{m}\frac{-C}{k}>(m-n)\frac{-C}{n},$$

由此容易看出 $\{a_n\}$ 满足上述条件 (i).

例 2.3.3 设 $a_n>0$, $\lim_{n\to\infty}a_n=\alpha$. 记 $b_n=\sqrt[n]{a_1a_2\cdots a_n}$, 则 $\lim_{n\to\infty}b_n=\alpha$.

证明 由题设, 任取 $\varepsilon>0$, 存在 N, 当 $n>N$ 时 $0<a_n<\alpha+\varepsilon$. 此时有

$$b_n\le(\alpha+\varepsilon)\sqrt[n]{a_1a_2\cdots a_N(\alpha+\varepsilon)^{-N}},$$

令 $n\to\infty$, 得 $\varlimsup_{n\to\infty}b_n\le\alpha+\varepsilon$. 当 $\alpha>0$ 时同理可证 $\varliminf_{n\to\infty}b_n\ge\alpha-\varepsilon$, 当 $\alpha=0$ 此不等式当然也成立. 由 ε 的任意性可知

$$\varlimsup_{n\to\infty}b_n\le\alpha\le\varliminf_{n\to\infty}b_n\le\varlimsup_{n\to\infty}b_n,$$

这说明 $\{b_n\}$ 收敛到 α.

例 2.3.4 设 $b_n>0$, $\lim_{n\to\infty}\frac{b_{n+1}}{b_n}=\alpha$, 则 $\lim_{n\to\infty}\sqrt[n]{b_n}=\alpha$.

证明 令 $a_1=b_1$, $n\ge1$ 时 $a_{n+1}=\frac{b_{n+1}}{b_n}$. 由题设, $\{a_n\}$ 收敛到 α. 由上例可得

$$\lim_{n\to\infty}\sqrt[n]{b_n}=\lim_{n\to\infty}\sqrt[n]{a_1a_2\cdots a_n}=\alpha.$$

例 2.3.5 设非负数列 $\{a_n\}$ 满足条件 $a_{m+n}\le a_m+a_n$, $\forall\, m,n\ge1$. 证明数列 $\left\{\frac{a_n}{n}\right\}$ 收敛.

证明 由归纳法可见 $0\le a_n\le na_1$, 因此 $\left\{\frac{a_n}{n}\right\}$ 为有界数列. 暂时固定正整数 k, 当 $n\ge k$ 时, n 可以表示为 $n=mk+l$, 其中 $0\le l\le k-1$. 于是

$$\frac{a_n}{n}\le\frac{a_{mk}+a_l}{n}\le\frac{ma_k+la_1}{n}\le\frac{a_k}{k}+\frac{(k-1)a_1}{n},$$

令 $n\to\infty$, 取上极限可得 $\varlimsup_{n\to\infty}\frac{a_n}{n}\le\frac{a_k}{k}$. 再令 $k\to\infty$, 取下极限可知 $\left\{\frac{a_n}{n}\right\}$ 的上下极限一定相等, 从而该数列收敛.

例 2.3.6 设 m 为正整数, $\{a_n\}$ 为非负的有界数列, 其下极限和上极限分别为 α, β, 则 $\varliminf_{n\to\infty}\sqrt[m]{a_n}=\sqrt[m]{\alpha}$, $\varlimsup_{n\to\infty}\sqrt[m]{a_n}=\sqrt[m]{\beta}$.

证明 记 $b_n = \sqrt[m]{a_n}$, 不难说明 $\underline{b}_n = \sqrt[m]{\underline{a}_n}$, $\bar{b}_n = \sqrt[m]{\bar{a}_n}$. 令 $n \to \infty$ 可得欲证结论.

例 2.3.7 设 $a_0, a_1 > 0$, 当 $n \geq 1$ 时 $a_{n+1} = \sqrt{a_n} + \sqrt{a_{n-1}}$. 证明 $\{a_n\}$ 收敛.

证明 先看看 $\{a_n\}$ 是否为有界数列. 我们希望找到 M, 使得由 $a_n \leq M$ 和 $a_{n-1} \leq M$ 能推出

$$a_{n+1} \leq \sqrt{M} + \sqrt{M} \leq M.$$

令 $M = \max\{a_0, a_1, 4\}$, 根据上述分析可知 M 的确为 $\{a_n\}$ 的上界. 同理, 令 $m = \min\{a_0, a_1, 4\}$, 则 m 为 $\{a_n\}$ 的下界.

于是 $\{a_n\}$ 为有界数列, 其下极限和上极限分别记为 α, β. 则

$$\alpha = \underline{\lim}_{n\to\infty} a_{n+1} \geq \underline{\lim}_{n\to\infty} \sqrt{a_n} + \underline{\lim}_{n\to\infty} \sqrt{a_{n-1}} = 2\sqrt{\alpha}.$$

同理, 取上极限可得 $\beta \leq 2\sqrt{\beta}$. 这说明 $4 \leq \alpha \leq \beta \leq 4$, 因此 $\{a_n\}$ 收敛于 4.

例 2.3.8 设 $\{a_n\}$ 为有界数列, 且 $\{a_{2n} + 2a_n\}$ 收敛. 证明 $\{a_n\}$ 也收敛.

证明 $\{a_n\}$ 的下极限和上极限分别记为 λ, μ. 由**命题 2.3.3** 可得

$$\underline{\lim}_{n\to\infty} (a_{2n} + 2a_n) \leq \overline{\lim}_{n\to\infty} a_{2n} + \underline{\lim}_{n\to\infty} 2a_n \leq \mu + 2\lambda.$$

同理, 取上极限可得

$$\overline{\lim}_{n\to\infty} (a_{2n} + 2a_n) \geq \underline{\lim}_{n\to\infty} a_{2n} + \overline{\lim}_{n\to\infty} 2a_n \geq \lambda + 2\mu.$$

由题设和**定理 2.3.1**可知 $\lambda + 2\mu \leq \mu + 2\lambda$, 这说明 $\lambda = \mu$, 从而 $\{a_n\}$ 收敛.

习题 2.3

1. 给出**命题 2.3.2**中 (1) 和 (2) 的证明.

2. 设 $\{a_n\}$ 为有界数列, α, β 分别是其下极限和上极限. 证明: 任给 $\varepsilon > 0$, 均存在 N, 使得当 $n > N$ 时 $\alpha - \varepsilon < a_n < \beta + \varepsilon$.

3. (1) 设 $\{a_n\}$ 收敛于 α, 模仿**例 2.3.3**中的方法重新证明

$$\lim_{n\to\infty} \frac{1}{n}(a_1 + a_2 + \cdots + a_n) = \alpha.$$

(2) 设 $\{a_n\}$ 为有界数列, 记 $b_n = \frac{1}{n}\sum\limits_{k=1}^{n} a_k$, 则

$$\underline{\lim}_{n\to\infty} a_n \leq \underline{\lim}_{n\to\infty} b_n \leq \overline{\lim}_{n\to\infty} b_n \leq \overline{\lim}_{n\to\infty} a_n.$$

4. 设 $\lim\limits_{n\to\infty} a_n = \alpha$, $\lim\limits_{n\to\infty} b_n = \beta$, 则

$$\lim_{n\to\infty} \frac{1}{n}(a_1 b_n + a_2 b_{n-1} + \cdots + a_n b_1) = \alpha\beta.$$

5. 设 $\{a_n\}$ 为有界数列, $\{b_n\}$ 收敛于 β, 则

$$\underline{\lim}_{n\to\infty} (a_n + b_n) = \underline{\lim}_{n\to\infty} a_n + \beta, \quad \overline{\lim}_{n\to\infty} (a_n + b_n) = \overline{\lim}_{n\to\infty} a_n + \beta.$$

6. 设 $\{a_n\}$ 收敛于非负实数 α, $\{b_n\}$ 为有界数列, 则

$$\varliminf_{n\to\infty} a_nb_n = \alpha \varliminf_{n\to\infty} b_n,\quad \varlimsup_{n\to\infty} a_nb_n = \alpha \varlimsup_{n\to\infty} b_n.$$

7. 设 $\{a_n\}$ 为有界数列. 如果 α 是 $\{a_n\}$ 的某个收敛子列的极限, 则称 α 为 $\{a_n\}$ 的极限点. 极限点的全体记为 A, 则 $\min A = \varliminf\limits_{n\to\infty} a_n$, $\max A = \varlimsup\limits_{n\to\infty} a_n$.

8. 设 $a_0, a_1 \geq 0$, 当 $n \geq 1$ 时 $0 \leq a_{n+1} \leq \frac{1}{2}(a_n + a_{n-1})$. 证明 $\{a_n\}$ 收敛.

9. 设 $a_0, a_1 > 0$, 当 $n \geq 1$ 时 $a_{n+1} = \frac{2}{a_n + a_{n-1}}$. 证明 $\{a_n\}$ 收敛.

2.4 Cauchy 准则

我们借助单调子列来说明 Cauchy 数列等价于收敛数列.

引理 2.4.1 每一个数列都必有单调子列.

证明 设 $\{a_n\}$ 为数列. 我们知道, 若 $\{a_n\}$ 单调递增(减), 则每一项 a_n 都不大(小)于它后面的那些项. 受此启发, 记 $I = \{n \mid m \geq n \text{ 时 } a_m \geq a_n\}$. 分情况讨论:

(1) I 没有上界. 取 $n_1 \in I$, 因为 n_1 不是 I 的上界, 故存在 $n_2 \in I$, 使得 $n_2 > n_1$. 此时 $a_{n_2} \geq a_{n_1}$. 同理, 存在 $n_3 \in I$, 使得 $n_3 > n_2$, 此时 $a_{n_3} \geq a_{n_2}$. 以此类推, 可归纳地定义出 $\{a_n\}$ 中一个单调递增的子列 $\{a_{n_k}\}$.

(2) I 有上界, 设 N 是其上界. 则任给 $n > N$, 均存在 $m > n$, 使得 $a_m < a_n$. 特别地, 取 $n_1 = N+1$, 存在 $n_2 > n_1$, 使得 $a_{n_2} < a_{n_1}$. 同理, 存在 $n_3 > n_2$, 使得 $a_{n_3} < a_{n_2}$. 以此类推, 可归纳地定义出 $\{a_n\}$ 中一个严格单调递减的子列 $\{a_{n_k}\}$.

命题 2.4.2 Cauchy 数列必为有界数列.

$M = \max\{|a_1|, \cdots, |a_N|, 1 + |a_{N+1}|\}$

定理 2.4.3(Cauchy 准则)

充分性: 设 $\{a_n\}$ 为 Cauchy 数列. 由前一命题, $\{a_n\}$ 是有界数列. 再由前一引理可知 $\{a_n\}$ 必有单调子列 $\{a_{n_k}\}$, 它是有界的, 必然收敛. 其极限记为 α, 我们要说明 $\{a_n\}$ 也收敛于 α. 事实上, 任给 $\varepsilon > 0$, 由已知结论和题设条件可知, 存在 N_0, N_1, 使得当 $k > N_0$ 时 $|a_{n_k} - \alpha| < \frac{\varepsilon}{2}$; 当 $m, n > N_1$ 时 $|a_m - a_n| < \frac{\varepsilon}{2}$. 取 $N > \max\{N_0, N_1\}$, 则 $n_N \geq N > N_1$. 于是当 $n > N$ 时, 有

$$|a_n - \alpha| \leq |a_n - a_{n_N}| + |a_{n_N} - \alpha| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

按照数列极限的定义即得欲证结论.

注 也可以利用上下极限证明充分性: $\{a_n\}$ 的上极限记为 β, 我们要说明 $\{a_n\}$ 收敛于 β. 事实上, 任取 $\varepsilon > 0$, 存在 N, 当 $m, n > N$ 时

$$a_m - \varepsilon < a_n < a_m + \varepsilon.$$

在上式中暂时固定 m, 利用**命题 2.3.2**可得

$$a_m - \varepsilon \le \beta \le a_m + \varepsilon.$$

这说明当 $m > N$ 时 $|a_m - \beta| \le \varepsilon$. 根据数列极限的定义即得欲证结论.

2.5 Stolz 公式

去掉**引理 2.5.1**.

在 **1.1 节**中我们曾用裂项的方法求和. 这个方法也可以用来研究数列的极限. 比如, 设 $\{a_n\}$ 为数列, 当 $n > N$ 时, 有

$$a_n = \sum_{i=N+1}^{n} (a_i - a_{i-1}) + a_N,$$

如果我们对相邻两项之差 $a_i - a_{i-1}$ 有比较好的了解, 就能更好地认识 $\{a_n\}$.

定理 2.5.1 证明修改:

(1)

此时, 任给 $\varepsilon > 0$, 存在 N, 当 $n > N$ 时 $|x_n - x_{n-1}| < \varepsilon(y_n - y_{n-1})$, 从而由三角不等式可得

$$|x_n - x_N| \le \sum_{i=N+1}^{n} |x_i - x_{i-1}| < \sum_{i=N+1}^{n} \varepsilon(y_i - y_{i-1}) = \varepsilon(y_n - y_N).$$

再取 $N' \ge N$, 使得当 $n > N'$ 时 $y_n > |y_N| + \frac{1}{\varepsilon}|x_N|$, 此时

$$\left|\frac{x_n}{y_n}\right| \le \frac{|x_n - x_N| + |x_N|}{y_n} < \frac{\varepsilon(y_n - y_N) + |x_N|}{y_n} < 2\varepsilon.$$

由数列极限的定义即知 $\lim\limits_{n\to\infty} \frac{x_n}{y_n} = 0$.

(2)

且

$$x_n - x_N = \sum_{i=N+1}^{n} (x_i - x_{i-1}) > \sum_{i=N+1}^{n} (y_i - y_{i-1}) = y_n - y_N.$$

例 2.5.4 设 $\{a_n\}$ 为数列, 且 $\{a_{n-1} + 2a_n\}$ 收敛, 则 $\{a_n\}$ 也收敛.

证明 不妨设 $\{a_{n-1} + 2a_n\}$ 收敛于零, 我们要说明 $\{a_n\}$ 也收敛于零. 记 $b_n = (-1)^n a_n$, 则

$$\lim_{n\to\infty} \frac{2^n b_n - 2^{n-1} b_{n-1}}{2^n - 2^{n-1}} = \lim_{n\to\infty} (2b_n - b_{n-1}) = \lim_{n\to\infty} (-1)^n (2a_n + a_{n-1}) = 0.$$

由 Stolz 公式可得 $\lim\limits_{n\to\infty} \frac{2^n b_n}{2^n} = 0$, 即 $\{b_n\}$ 收敛于零, 从而 $\{a_n\}$ 也收敛于零.

习题 2.5:

10. (1) 去掉.

2.6　实数系的基本性质

可数集移至闭区间套原理后.

在闭集的定义后面添加例子:

例 2.6.2 设 A 是有上界的集合, 以 A 的所有上界为元素构成的集合记为 B, 则 B 为闭集.

证明 我们要证明 B 的补集 B^c 是开集. 设 $x \in B^c$, 即 x 不是 A 的上界, 则存在 $a \in A$, 使得 $x < a$. 此时易见 $(-\infty, a) \subset B^c$, 按照定义即知 B^c 为开集.

定理 2.6.3(Bolzano) 定理证明修改:

结合**引理 2.4.1**和**定理 2.2.1**可立即得到欲证结论. 下面我们用有限覆盖定理给出另一证明.

习题 2.6：

1. 补充条件 $\lim\limits_{n\to\infty}(b_n - a_n) = 0$.

2.7　附录: 三角函数与 Euler 公式

Warning: 首次阅读本书时, 本节内容可以跳过.

多项式函数是我们非常熟悉的一类函数, 它们可以由常数和 x 经过有限次加、减、乘法运算得到. 两个多项式相除得到的函数称为有理函数. 我们在前面又分别引进了开方运算与求数列极限的运算, 利用这两个新的运算可以得到新型函数. 比如在 §2.2 节中我们用数列极限定义了自然对数函数和自然指数函数.

中学数学中还直观地介绍过简单的三角函数和反三角函数, 其中涉及的弧长(角度)和圆周率并未严格定义. 下面我们用极限来严格地定义 $\sin x$, $\cos x$ 以及圆周率等. 这个过程略微复杂一点, 初学者可以先从几何的观点来看待它们, 到一定的阶段以后再回头学习(也可参见 §4.4 节中的部分内容).

我们假定读者了解复数的定义和基本运算规则, 复数的全体记为 $\mathbb{C}$. 先把实的指数函数拓展为复的指数函数. 为此, 引进复数列收敛的概念: 设 $\{z_n\}$ 为一列复数, 如果存在复数 z_0, 使得 $\lim\limits_{n\to\infty}|z_n - z_0| = 0$, 则称 $\{z_n\}$ 收敛, 其极限为 z_0, 记为 $\lim\limits_{n\to\infty} z_n = z_0$. 容易看出, $\{z_n\}$ 收敛当且仅当其实部和虚部分别构成的实数列收敛. 与实数不同, 复数一般无法比大小, 因此一般不能用单调性研究复数列的极限. 不过 Cauchy 列的概念仍然有用: $\{z_n\} \subset \mathbb{C}$ 为 Cauchy 列是指任给 $\varepsilon > 0$, 存在 N, 当 $m, n > N$ 时均有 $|z_m - z_n| < \varepsilon$. Cauchy 准则对复数列仍然成立.

我们首先有如下结果, 它有时可以代替 Bernoulli 不等式去做一些简单的估计.

引理 2.7.1 设 $w \in \mathbb{C}$, n 为正整数, 则

$$\big|(1+w)^n - 1 - nw\big| \le \frac{1}{2}n(n-1)(1+|w|)^{n-2}|w|^2. \tag{2.7}$$

证明. 不妨设 $n>2$. 根据 Newton 二项式展开我们有

$$\begin{aligned}\big|(1+w)^n-1-nw\big| &= \Big|\sum_{k=2}^{n} C_n^k w^k\Big| \le \sum_{k=2}^{n} C_n^k |w|^k \\ &\le \sum_{k=2}^{n} C_n^2 C_{n-2}^{k-2} |w|^{k-2}|w|^2 \\ &= C_n^2(1+|w|)^{n-2}|w|^2.\end{aligned}$$

引理 2.7.2 设 $z\in\mathbb{C}$, 则复数列 $\{\big(1+\frac{z}{n}\big)^n\}$ 收敛.

证明. 只要验证该数列为 Cauchy 列即可. 注意到

$$\Big|\Big(1+\frac{z}{n}\Big)^n-\Big(1+\frac{z}{n-1}\Big)^{n-1}\Big| = \Big|1+\frac{z}{n-1}\Big|^n\left|\left(1-\frac{\frac{z}{n(n-1)}}{1+\frac{z}{n-1}}\right)^n-\Big(1+\frac{z}{n-1}\Big)^{-1}\right|,$$

在引理 2.7.1 中令 $w=-\dfrac{\frac{z}{n(n-1)}}{1+\frac{z}{n-1}}$, 然后代入上式并化简可得

$$\begin{aligned}\Big|\Big(1+\frac{z}{n}\Big)^n-\Big(1+\frac{z}{n-1}\Big)^{n-1}\Big| &\le \frac{|z|^2}{2n(n-1)}\Big(\Big|1+\frac{z}{n-1}\Big|+\frac{|z|}{n(n-1)}\Big)^{n-2} \\ &\le \frac{|z|^2}{2n(n-1)}\Big(1+\frac{|z|(n+1)}{n(n-1)}\Big)^{n-2} \\ &\le \frac{|z|^2}{2n(n-1)}\exp\Big(|z|\frac{(n+1)(n-2)}{n(n-1)}\Big) \\ &\le \frac{|z|^2\exp(|z|)}{2n(n-1)}.\end{aligned}$$

由此容易看出我们要考虑的复数列的确是 Cauchy 列.

命题 2.7.3 设 $z\in\mathbb{C}$, 记 $\exp(z)=\lim\limits_{n\to\infty}\big(1+\frac{z}{n}\big)^n$, 称为复的自然指数函数, 它有如下简单性质:

(1) $\big|\exp(z)-1-z\big|\le\frac{1}{2}|z|^2\exp(|z|)$;

(2) $\exp(z)\exp(z')=\exp(z+z')$;

(3) 当 $x\in\mathbb{R}$ 时 $\exp(\mathrm{i}x)\in S^1=\{z\in\mathbb{C}\mid |z|=1\}$.

证明. (1) 利用引理 2.7.1 可得

$$\Big|\Big(1+\frac{z}{n}\Big)^n-1-z\Big|\le\frac{1}{2}n(n-1)\Big(1+\frac{|z|}{n}\Big)^{n-2}\frac{|z|^2}{n^2}\le\frac{n-1}{2n}|z|^2\exp(|z|). \tag{2.8}$$

令 $n\to\infty$ 即得欲证结论.

(2) 这可模仿实的情形的证明, 只需注意当 $u,v\in\mathbb{C}$ 时

$$|u^n-v^n|\le n|u-v|\max\{|u|^{n-1},\ |v|^{n-1}\}.$$

(3) 当 $x\in\mathbb{R}$ 时

$$|\exp(\mathrm{i}x)|^2=\lim_{n\to\infty}\Big|1+\frac{\mathrm{i}x}{n}\Big|^{2n}=\lim_{n\to\infty}\Big(1+\frac{x^2}{n^2}\Big)^n.$$

在 (2.8) 中代入 $z=\frac{x^2}{n}$ 不难得出欲证结论.

推论 2.7.4 设 $x\in\mathbb{R}$, $\exp(\mathrm{i}x)$ 的实部和虚部分别记为 $C(x)$, $S(x)$, 则 $C(0)=1$, $S(0)=0$, 且

(1) $C(-x)=C(x)$, $S(-x)=-S(x)$, $C^2(x)+S^2(x)\equiv 1$;

(2) $|S(x)-x|\le \frac{1}{2}x^2\exp(|x|)$;

(3) $C(x+x')=C(x)C(x')-S(x)S(x')$, $S(x+x')=S(x)C(x')+C(x)S(x')$.

证明. (1) 根据前一命题的 (2) 和 (3) 可知 $\exp(-\mathrm{i}x)=\overline{\exp(\mathrm{i}x)}$, 比较两边的实部虚部即知 $C(x)$ 为偶函数, $S(x)$ 为奇函数. 由 $|\exp(\mathrm{i}x)|=1$ 可知 $C^2(x)+S^2(x)\equiv 1$.

(2) 在前一命题的 (1) 中代入 $z=\mathrm{i}x$ 即可得到欲证估计.

(3) 在前一命题的 (2) 中代入 $z=\mathrm{i}x$， $z'=\mathrm{i}x'$, 再比较两边的实部虚部即可.

特别地, 当 $x=x'$ 时由 (3) 可得 $S(2x)=2S(x)C(x)$ 以及

$$C(2x)=C^2(x)-S^2(x)=2C^2(x)-1=1-2S^2(x).$$

从 (3) 也不难得到如下和差化积公式:

$$C(x)-C(x')=-2S\Big(\frac{x-x'}{2}\Big)S\Big(\frac{x+x'}{2}\Big),\ S(x)-S(x')=2S\Big(\frac{x-x'}{2}\Big)C(\frac{x+x'}{2}\Big).$$

和差化积公式可以用来对 $C(x)$ 和 $S(x)$ 做进一步的估计: 如果对 $S(x)$ 有一定的了解, 则利用和差化积公式就可以对 $C(x)$ 有较好的了解, 进而对 $S(x)$ 有更好的了解. 下面我们就用这种想法来做几个有用的估计.

首先, 由 (1) 可知 $|C(x)|\le 1$, $|S(x)|\le 1$. 进一步, 我们还有

$$|S(x)|\le |x|,\ \ \forall\, x\in\mathbb{R}. \tag{2.9}$$

事实上, 当 n 为正整数时, 利用 (3) 可得

$$|S(nx)|\le |S((n-1)x)|\,|C(x)|+|S(x)|\,|C((n-1)x)|\le |S((n-1)x)|+|S(x)|,$$

利用归纳法可得 $|S(nx)|\le n|S(x)|$. 将 x 改写为 $\frac{x}{n}$, 再令 $n\to\infty$ 并利用 (2) 即可得到 (2.9) 式. 由和差化积公式和 (2.9) 式可得

$$|C(x)-C(x')|\le |x-x'|,\ \ |S(x)-S(x')|\le |x-x'|. \tag{2.10}$$

其次, 利用 $1-C(x)=2S^2\big(\frac{x}{2}\big)$ 就有

$$C(x)\ge 1-\frac{1}{2}x^2,\ \ \forall\, x\in\mathbb{R}. \tag{2.11}$$

进一步, 我们还有

$$\frac{S(x)}{x}\ge 1-\frac{1}{6}x^2,\ \ \forall\, x\in\mathbb{R}\setminus\{0\}. \tag{2.12}$$

事实上, 当 n 为正整数时,

$$S(x)=\sum_{j=1}^{n}\Big[S\Big(\frac{j}{n}x\Big)-S\Big(\frac{j-1}{n}x\Big)\Big]=\sum_{j=1}^{n}2S\Big(\frac{x}{2n}\Big)C\Big(\frac{2j-1}{2n}x\Big).$$

当 $x\neq 0$ 且 n 充分大时, 利用 (2.9) 式可得

$$\frac{S(x)}{2nS\big(\frac{x}{2n}\big)}=\frac{1}{n}\sum_{j=1}^{n}C\Big(\frac{2j-1}{2n}x\Big)\geq\frac{1}{n}\sum_{j=1}^{n}\Big[1-\frac{1}{2}\Big(\frac{2j-1}{2n}x\Big)^2\Big],$$

令 $n\to\infty$ 并利用 (2) 即可得到 (2.12) 式.

有了上述估计, 我们可以得出如下重要结论:

命题 2.7.5 存在 $\tau\in(1,2)$, 使得 $C(\tau)=0$, $S(\tau)=1$, 且

(1) $C(x)$ 在 $[0,2\tau]$ 中严格单调递减, 在 $[2\tau,4\tau]$ 中严格单调递增; $S(x)$ 在 $[0,\tau]$ 和 $[3\tau,4\tau]$ 中分别严格单调递增, 在 $[\tau,3\tau]$ 中严格单调递减;

(2) $C(x)$, $S(x)$ 均为周期函数, 4τ 是它们的最小正周期.

证明. 当 $x\in[0,1]$ 时, 由 (2.11) 式可知 $C(x)\geq\frac{1}{2}$, 由 (2.12) 式可知 $S(x)\geq\frac{5}{6}x$. 特别地,

$$C(2)=1-2S^2(1)\leq 1-2\Big(\frac{5}{6}\Big)^2=-\frac{7}{18}<0$$

由**例 1.4.1** 的证明可知 $C(x)$ 在 $(1,2)$ 中有零点(也可见本节习题 2), 记为 τ. 由 $C^2(\tau)+S^2(\tau)=1$ 可知 $S(\tau)=\pm1$. 我们来说明 $S(\tau)>0$, 从而只能有 $S(\tau)=1$.

事实上, 当 $x\in[0,2]$ 时, $\frac{x}{2}\in[0,1]$, 因此前一段中的估计表明

$$S(x)=2S\Big(\frac{x}{2}\Big)C\Big(\frac{x}{2}\Big)\geq\frac{5}{12}x.$$

特别地, $S(\tau)>0$.

于是我们有 $\exp(\mathrm{i}\tau)=\mathrm{i}$, $\exp(\mathrm{i}2\tau)=\mathrm{i}^2=-1$, $\exp(\mathrm{i}4\tau)=(-1)^2=1$. 进一步, 当 $x\in\mathbb{R}$ 时, 就有

$$\exp\big(\mathrm{i}(x+\tau)\big)=\mathrm{i}\exp(\mathrm{i}x),\ \exp\big(\mathrm{i}(x+2\tau)\big)=-\exp(\mathrm{i}x),\ \exp\big(\mathrm{i}(x+4\tau)\big)=\exp(\mathrm{i}x).$$

这些等式当然可以用 $C(x)$, $S(x)$ 表示出来, 特别地我们可以得知 $C(x)$, $S(x)$ 都是周期函数, 4τ 是它们的周期. 为了说明这是最小正周期, 我们再来了解 $C(x)$ 和 $S(x)$ 的单调性.

首先, 前面已说明 $S(x)$ 在 $(0,2)$ 中恒为正, 利用和差化积公式就可以看出 $C(x)$ 在 $[0,2]$ 中严格单调递减. 于是 $C(x)$ 在 $[0,\tau)$ 中恒为正. 进一步, 由 $S(x)=2S\big(\frac{x}{2}\big)C\big(\frac{x}{2}\big)$ 可知 $S(x)$ 在 $(0,2\tau)$ 中恒为正, $S(0)=S(2\tau)=0$. 再次利用和差化积公式就可以看出 $C(x)$ 在 $[0,2\tau]$ 中严格单调递减. 由 $C(x+2\tau)=-C(x)$ 可知 $C(x)$ 在 $[2\tau,4\tau]$ 中严格单调递增.

至于 $S(x)$ 的单调性则可以由 $C(x+\tau)=-S(x)$ 得出. 有了这些单调性就不难看出 4τ 是 $C(x)$ 和 $S(x)$ 的最小正周期.

命题 2.7.6 当 $x\in\mathbb{R}$ 时, 记 $e(x)=\exp(\mathrm{i}x)$, 则

(1) $e\big([0,4\tau)\big)=S^1$;

(2) $e(x)=e(x')$ 当且仅当存在整数 k, 使得 $x-x'=4k\tau$.

证明. (1) 我们已经知道 $e(\mathbb{R})\subset S^1$. 反之, $1=e(0)=e(4\tau)$, $\mathrm{i}=e(\tau)$, $-1=e(2\tau)$, $-\mathrm{i}=e(3\tau)$, 这四个点将单位圆周 S^1 分为四个部分, 我们来分情况讨论. 设 $z\in S^1$, 其实部和虚部分别为 a, b.

若 $0<a,b<1$, 由 $C(0)=1$, $C(\tau)=0$ 可知, 存在 $x\in(0,\tau)$, 使得 $C(x)=a$. 此时由 $S(x)>0$ 可知

$$S(x)=\sqrt{1-C^2(x)}=\sqrt{1-a^2}=b,$$

即 $e(x)=z$.

若 $-1<a<0$, $0<b<1$, 则存在 $x\in(0,\tau)$, 使得 $C(x)=-a$. 此时 $e(2\tau-x)=z$. 若 $-1<a,b<0$, 则存在 $x\in(0,\tau)$, 使得 $C(x)=-a$. 此时 $e(2\tau+x)=z$. 若 $0<a<1$, $-1<b<0$, 则存在 $x\in(0,\tau)$, 使得 $C(x)=a$. 此时 $e(4\tau-x)=z$.

(2) 注意到 $e(t)=1$ 意味着 $C(x+t)=C(x)$, $S(x+t)=S(x)$, 从而 t 必为最小正周期 4τ 的整数倍.

记 $\pi=2\tau$, 称为圆周率(以后可以说明 4τ 是单位圆周 S^1 的周长). 当 $x\in\mathbb{R}$ 时, 分别将 $C(x)$, $S(x)$ 改记为 $\cos x$, $\sin x$, 这就是我们熟悉的余弦函数和正弦函数. $\exp(z)$ 也常记为 e^z. 根据定义, 我们有

$$\mathrm{e}^{\mathrm{i}x}=\cos x+\mathrm{i}\sin x, \tag{2.13}$$

这就是有名的 Euler 公式.

设 z 为非零复数, 记 $r=|z|$, 它是 z 到原点的距离. 此时 $\frac{1}{r}z\in S^1$, 因此存在 $\theta\in[0,2\pi)$, 使得 $\frac{1}{r}z=\mathrm{e}^{\mathrm{i}\theta}$, 或 $z=r\mathrm{e}^{\mathrm{i}\theta}$. θ 称为 z 的辐角, (r,θ) 称为 z 的极坐标.

例 2.7.1 研究数列 $\{\sin(n^2)\}$ 的敛散性.

解 我们来说明该数列是发散的. (反证法) 若它收敛, 其极限记为 β, 则

$$\cos(2n^2)=1-2\sin^2(n^2)\to 1-2\beta^2,\quad (n\to\infty).$$

同理有

$$\cos(2n)^2=2\cos^2(2n^2)-1\to\alpha=2(1-2\beta^2)^2-1,\quad (n\to\infty).$$

由 Euler 公式即知 $\lim\limits_{n\to\infty}\mathrm{e}^{\mathrm{i}(2n)^2}=\alpha+\mathrm{i}\beta\in S^1$. 于是

$$\lim_{n\to\infty}\mathrm{e}^{\mathrm{i}[(2n)^2-(2n-2)^2]}=\lim_{n\to\infty}\frac{\mathrm{e}^{\mathrm{i}(2n)^2}}{\mathrm{e}^{\mathrm{i}(2n-2)^2}}=\frac{\alpha+\mathrm{i}\beta}{\alpha+\mathrm{i}\beta}=1.$$

即 $\lim\limits_{n\to\infty}\mathrm{e}^{\mathrm{i}8n}=\mathrm{e}^{\mathrm{i}4}$. 同理有

$$\lim_{n\to\infty}\mathrm{e}^{\mathrm{i}[8n-8(n-1)]}=\frac{\mathrm{e}^{\mathrm{i}4}}{\mathrm{e}^{\mathrm{i}4}}=1,$$

即 $e^{i8}=1$. 于是存在正整数 k, 使得 $8=2k\pi$, 但这与 $\pi=2\tau\in(2,4)$ 相矛盾.

习题:

1. 证明和差化积公式.

2. 设 $a_1=1$, 当 $n\geq 1$ 时 $a_{n+1}=a_n+C(a_n)$. 证明 $\{a_n\}$ 单调递增且以 2 为上界, 其极限 $\tau\in(1,2)$ 且 $C(\tau)=0$. (提示: $x+C(x)$ 是单调递增函数.)

3. 设 $a\in(0,1)$, 对 $C(x)-a$ 运用上一题中的方法, 去说明存在 $x\in(0,\tau)$, 使得 $C(x)=a$.

4. 当 $x\in(0,\tau)$ 时, 证明 $\frac{S(x)}{x}\geq C(x)$.

5. 设 $z\in\mathbb{C}$, n 为正整数, 证明

$$\left|\left(1+\frac{z}{n}\right)^n-\exp(z)\right|\leq\frac{|z^2|}{2n}\exp(|z|).$$

6. 设 $z\in\mathbb{C}$, n 为正整数, 记 $e_n(z)=1+z+\frac{z^2}{2!}+\cdots+\frac{z^n}{n!}$. 证明

(1) $\left|\left(1+\frac{z}{n}\right)^n-e_n(z)\right|\leq\frac{|z|^2}{2n}\exp(|z|)$;

(2) $\lim\limits_{n\to\infty}e_n(z)=\exp(z)$.

7. 设 $\{z_n\}$ 为复数列, 如果 $\lim\limits_{n\to\infty}\left(z_1+z_2+\cdots+z_n\right)=w$, 则记 $w=\sum\limits_{n=1}^{\infty}z_n$.

(1) 设 $z\in\mathbb{C}$. 当 $|z|<1$ 时, 证明: $\frac{1}{1-z}=1+\sum\limits_{n=1}^{\infty}z^n$.

(2) 设 $r,x\in\mathbb{R}$. 当 $-1<r<1$ 时, 证明:

$$\frac{1-r^2}{1+r^2-2r\cos x}=1+\sum_{n=1}^{\infty}2r^n\cos nx,\quad \frac{\sin x}{1+r^2-2r\cos x}=\sum_{n=1}^{\infty}r^{n-1}\sin nx.$$

8. 利用和差化积公式和 $\sin x$ 的估计证明 $\cos x\leq 1-\frac{x^2}{2!}+\frac{x^4}{4!}$, 再由此说明 $\pi<3.2$.

9. 利用 Euler 公式证明数列 $\{\sin(n^3)\}$ 必定发散.

第三章　连续函数

第三章: 求经典极限时尽量采用定量估计的方法

3.1　函数的极限

例 3.1.3 设 n 为正整数, $f(x)=\sqrt[n]{x}$, $x\in[0,+\infty)$. 研究 f 的函数极限.

解 若 $x_0=0$, 则任给 $\varepsilon>0$, 取 $\delta=\varepsilon^n$, 当 $0<x<\delta$ 时就有 $0\le\sqrt[n]{x}<\varepsilon$. 这说明 $\lim\limits_{x\to0^+}\sqrt[n]{x}=0$.

若 $x_0>0$, 在 (2.6) 中取 $a=\frac{x}{x_0}$ 可得

$$\frac{n\sqrt[n]{x_0}}{x}(x-x_0)\le\sqrt[n]{x}-\sqrt[n]{x_0}\le\frac{n\sqrt[n]{x_0}}{x_0}(x-x_0). \tag{3.1}$$

由夹逼定理可知 $\lim\limits_{x\to x_0}\sqrt[n]{x}=\sqrt[n]{x_0}$.

例 3.1.4 设 $q>0$, $q\ne1$, $f(x)=\log_q x$, $x\in(0,+\infty)$. 研究 f 的函数极限.

解 先看自然对数函数 $\ln x$. 设 $x_0>0$, 将**命题 2.2.2**(2) 中的 x 换成 $\frac{x}{x_0}$ 可得

$$\frac{x-x_0}{x}\le\ln x-\ln x_0\le\frac{x-x_0}{x_0}. \tag{3.2}$$

由夹逼定理可知 $\lim\limits_{x\to x_0}\ln x=\ln x_0$. 再由 $\log_q x=\frac{\ln x}{\ln q}$ 可知 $\lim\limits_{x\to x_0}\log_q x=\log_q x_0$.

例 3.1.5 设 $a>0$, $f(x)=a^x$, $x\in(-\infty,+\infty)$. 研究 f 的函数极限.

解 先看自然指数函数 e^x. 设 $x,x_0\in\mathbb{R}$, 由**命题 2.2.4** 可得

$$\mathrm{e}^x-\mathrm{e}^{x_0}=\mathrm{e}^{x_0}\big(\mathrm{e}^{x-x_0}-1\big)\ge\mathrm{e}^{x_0}(x-x_0).$$

交换 x 和 x_0 的位置时上式当然也成立, 于是有

$$\mathrm{e}^{x_0}(x-x_0)\le\mathrm{e}^x-\mathrm{e}^{x_0}\le\mathrm{e}^x(x-x_0). \tag{3.3}$$

由 e^x 的单调性可知它在 x_0 附近有界, 于是由夹逼定理可知 $\lim\limits_{x\to x_0}\mathrm{e}^x=\mathrm{e}^{x_0}$. 再由 $a^x=\mathrm{e}^{x\ln a}$ 可知

$$a^{x_0}(x-x_0)\ln a\le a^x-a^{x_0}\le a^x(x-x_0)\ln a, \tag{3.4}$$

这说明 $\lim\limits_{x\to x_0} a^x = a^{x_0}$.

例 3.1.7 设 $x_0 \in \mathbb{R}$, 则 $\lim\limits_{x\to x_0} \sin x = \sin x_0$, $\lim\limits_{x\to x_0} \cos x = \cos x_0$.

证明 由和差化积公式可得

$$|\sin x - \sin x_0| = \left|2\sin\frac{x-x_0}{2}\right|\left|\cos\frac{x+x_0}{2}\right| \le 2\left|\frac{x-x_0}{2}\right| = |x-x_0|. \tag{3.5}$$

同理可得

$$|\cos x - \cos x_0| = \left|-2\sin\frac{x-x_0}{2}\right|\left|\sin\frac{x+x_0}{2}\right| \le 2\left|\frac{x-x_0}{2}\right| = |x-x_0|. \tag{3.6}$$

由夹逼定理即知欲证结论成立.

例 3.1.9 研究函数 $f(x) = \left(1+\frac{1}{x}\right)^x$ 在无穷远处的极限.

解 当 $x>0$ 或 $x<-1$ 时, 利用**命题 2.2.2**(2)可得

$$\frac{1}{1+x} \le \ln\left(1+\frac{1}{x}\right) \le \frac{1}{x}. \tag{3.7}$$

这说明

$$x>0 \text{ 时}, \ \frac{x}{x+1}\mathrm{e} \le \left(1+\frac{1}{x}\right)^x \le \mathrm{e}; \quad x<-1 \text{ 时}, \ \mathrm{e} \le \left(1+\frac{1}{x}\right)^x \le \frac{x}{x+1}\mathrm{e}.$$

总之, $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x = \mathrm{e}$.

例 3.1.10 验证 $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}} = \mathrm{e}$.

证明 作变量替换 $y=\frac{1}{x}$, 当 $x\to 0$ 时 $y\to\infty$, 从而欲证结论可从前例得出. 如果取自然对数, 还可得到

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = 1, \tag{3.8}$$

它当然也可以从 (3.2) 直接得出.

例 3.1.11 设 $a>0$, 验证 $\lim\limits_{x\to 0}\frac{a^x-1}{x} = \ln a$.

证明 不妨设 $a\ne 1$. 由 (3.3) 可得

$$\lim_{x\to 0}\frac{\mathrm{e}^x-1}{x} = 1. \tag{3.9}$$

于是有

$$\lim_{x\to 0}\frac{a^x-1}{x} = \lim_{x\to 0}\frac{\mathrm{e}^{x\ln a}-1}{x\ln a}\ln a = \ln a. \tag{3.10}$$

例 3.1.12 设 $a\in\mathbb{R}$, 验证 $\lim\limits_{x\to 0}\frac{(1+x)^a-1}{x} = a$.

证明 不妨设 $a\ne 0$. 当 $x\to 0$ 时 $y=a\ln(1+x)\to 0$, 于是有

$$\lim_{x\to 0}\frac{(1+x)^a-1}{x} = \lim_{x\to 0}\frac{\mathrm{e}^y-1}{y}\frac{a\ln(1+x)}{x} = a. \tag{3.11}$$

习题 3.1 修改:

2. (7): $\lim\limits_{x\to1}\frac{1-x^a}{1-x}$.
4. $(x\in\mathbb{R},\,a,b>0)$
(6): $\lim\limits_{n\to\infty}\left(\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2}\right)^n$.
6. 设 $x,x_0>0$, $\alpha<0$. 证明:

$$\alpha x_0^{\alpha-1}(x-x_0)\le x^\alpha-x_0^\alpha\le\alpha x^{\alpha-1}(x-x_0);$$

3.2 无穷小(大)量

例 3.2.3 设 $\alpha>0$, 则 $\lim\limits_{x\to+\infty}\dfrac{\ln x}{x^\alpha}=0$, 即 $\ln x=o(x^\alpha)$ $(x\to+\infty)$.

证明 当 $y>0$ 时我们知道 $\ln y\le y-1<y$. 因此, 当 $x>1$ 时

$$0<\frac{\ln x}{x^\alpha}=\frac{2}{\alpha}\frac{\ln x^{\frac{\alpha}{2}}}{x^\alpha}<\frac{2}{\alpha}\frac{x^{\frac{\alpha}{2}}}{x^\alpha}=\frac{2}{\alpha}\frac{1}{x^{\frac{\alpha}{2}}},$$

由 $\alpha>0$ 可知欲证结论成立.

例 3.2.4 设 $\alpha>0$, $a>1$ 则 $\lim\limits_{x\to+\infty}\dfrac{x^\alpha}{a^x}=0$, 即 $x^\alpha=o(a^x)$ $(x\to+\infty)$.

证明 注意到 $\dfrac{x^\alpha}{a^x}=\left(\dfrac{x}{b^x}\right)^\alpha$, 其中 $b=a^{\frac{1}{\alpha}}>1$. 只要证明 $\lim\limits_{x\to+\infty}\dfrac{x}{b^x}=0$ 即可. 记 $y=b^x$, 则当 $x\to+\infty$ 时 $y\to+\infty$. 因此

$$\lim_{x\to+\infty}\frac{x}{b^x}=\lim_{y\to+\infty}\frac{\ln y}{y\ln b}=0.$$

这里我们用了前例中的结论.

3.3 连续函数

3.3 节: **例 3.3.3** 中 Gauss 函数记号改为 $\lfloor x\rfloor$.

3.4 连续函数的整体性质

3.4 节:

最值定理后增加一个例子:

例 3.4.1 设 $f\in C^0[a,b]$ 且满足如下条件:

$$f\Big(\frac{x+y}{2}\Big)\le\frac{1}{2}[f(x)+f(y)],\ \ \forall\,x,y\in[a,b].$$

证明: f 必在 a 或 b 处达到最大值.

证明 根据最值定理, f 必在某一点 x_0 处达到最大值 M. 分情况讨论:

(1) $a\le x_0\le\frac{1}{2}(a+b)$. 此时 $x_1=2x_0-a\in[a,b]$. 由题设可得

$$M=f(x_0)\le\frac{1}{2}[f(a)+f(x_1)]\le\frac{1}{2}[f(a)+M],$$

即 $M \le f(a)$, 于是 $M = f(a)$.

(2) $\frac{1}{2}(a+b) \le x_0 \le b$. 此时 $x_2 = 2x_0 - b \in [a,b]$. 由题设可得

$$M = f(x_0) \le \frac{1}{2}[f(x_2) + f(b)] \le \frac{1}{2}[M + f(b)],$$

即 $M \le f(b)$, 于是 $M = f(b)$.

例 3.4.5:

解 由 (3.5) 可知, 任给 $\varepsilon > 0$, 取 $\delta = \varepsilon$, 当 $|x-y| < \delta$ 时

$$|\sin x - \sin y| \le |x-y| < \delta = \varepsilon,$$

这说明 $\sin x$ 在 $(-\infty, +\infty)$ 中一致连续. 同理, $\cos x$ 也一致连续.

例 3.4.6 设 $0 < \alpha < 1$, 则 $f(x) = x^\alpha$ 是 $[0, +\infty)$ 中的 α 阶 Hölder 函数.

证明 我们来说明, 当 $x, y \ge 0$ 时 $|x^\alpha - y^\alpha| \le |x-y|^\alpha$. 不妨设 $x > y > 0$. 记 $t = \frac{y}{x}$, 则 $0 < t < 1$, 且

$$\begin{aligned}(x-y)^\alpha + y^\alpha &= x^\alpha\big[(1-t)^\alpha + t^\alpha\big]\\ &\ge x^\alpha\big[(1-t) + t\big] = x^\alpha,\end{aligned}$$

这说明欲证结论成立.

习题 3.4:

9. (1): 去掉.

19. 设 $f \in C^0[a,b]$ 且满足如下条件:

$$f\Big(\frac{x+y}{2}\Big) = \frac{1}{2}[f(x) + f(y)],\ \ \forall\, x, y \in [a,b].$$

证明: 存在常数 α, β, 使得 $f(x) = \alpha x + \beta$.

3.5 连续函数的积分

3.5 节:

例 3.5.2

由积分的定义和自然对数的定义可得

$$\int_0^1 a^x \,\mathrm{d}x = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} a^{\frac{i}{n}} = \lim_{n\to\infty} \frac{a-1}{n\big(\sqrt[n]{a}-1\big)} = \frac{a-1}{\ln a}.$$

例 3.5.3

当 $\alpha = -1$ 时, 由上式及自然对数的定义可得

$$\int_a^b \frac{\mathrm{d}x}{x} = \lim_{n\to\infty} n(q-1) = \ln\frac{b}{a}.$$

积分的保序性质可以用来证明不等式. 我们先回顾平均值不等式.

例 3.5.7 用积分证明 $\mathrm{e}^x \geq 1+x$, 等号成立当且仅当 $x=0$.

证明 利用之前例子中的方法可得 $\displaystyle\int_0^x \mathrm{e}^t\,\mathrm{d}t = \mathrm{e}^x - 1$. 若 $x>0$, 由积分的保序性质可得

$$\begin{aligned}\mathrm{e}^x &= 1+x+\int_0^x \left(\mathrm{e}^t-1\right)\mathrm{d}t \geq 1+x+\int_{\frac{x}{2}}^x \left(\mathrm{e}^t-1\right)\mathrm{d}t\\ &\geq 1+x+\left(\mathrm{e}^{\frac{x}{2}}-1\right)\left(x-\frac{x}{2}\right) > 1+x.\end{aligned}$$

$x<0$ 的情形可以类似地证明.

原关于平均值不等式的内容删除.

再看 Bernoulli 不等式.

例 3.5.8 设 $x \geq -1$, $\alpha \geq 1$, 则 $(1+x)^\alpha \geq 1+\alpha x$. 等号成立当且仅当 $\alpha=1$ 或 $x=0$.

证明 不妨设 $\alpha>1$. 当 $x>0$ 时, 由之前的例子和积分的保序性质可得

$$\frac{1}{\alpha}\left[(1+x)^\alpha-1\right] = \int_1^{1+x} y^{\alpha-1}\,\mathrm{d}y = x+\int_1^{1+x}\left(y^{\alpha-1}-1\right)\mathrm{d}y > x.$$

$-1 \leq x < 0$ 的情形可以类似地证明.

例 3.5.9(Young 不等式) 设 $a,b>0$, $p,q>1$ 且 $\frac{1}{p}+\frac{1}{q}=1$, 则

$$ab \leq \frac{1}{p}a^p+\frac{1}{q}b^q,$$

等号成立当且仅当 $a^p=b^q$.

证明 当 $x>0$ 时, 由 Bernoulli 不等式可得

$$\frac{1}{p}x^p = \frac{1}{p}\left[1+(x-1)\right]^p \geq \frac{1}{p}\left[1+p(x-1)\right] = x-\frac{1}{q},$$

在上式中代入 $x=ab^{1-q}$ 可得

$$\frac{1}{p}a^p b^{-q} \geq ab^{1-q}-\frac{1}{q},$$

整理以后即得欲证结论.

原**例 3.5.9**移到选读材料中.

习题 3.5 增加:

1. 设 $x, x_0>0$. 用 Bernoulli 不等式证明: 当 $\alpha \geq 1$ 时,

$$\alpha x_0^{\alpha-1}(x-x_0) \leq x^\alpha - x_0^\alpha \leq \alpha x^{\alpha-1}(x-x_0);$$

当 $0<\alpha<1$ 时,

$$\alpha x^{\alpha-1}(x-x_0) \leq x^\alpha - x_0^\alpha \leq \alpha x_0^{\alpha-1}(x-x_0).$$

第四章　微积分基本公式

4.1　导数和微分

例 4.1.4 对数函数和指数函数的导数.

由 (3.2) 式可知 $(\ln x)' = \frac{1}{x}$. 当 $q > 0$, $q \neq 1$ 时, 由 $\log_q x = \frac{\ln x}{\ln q}$ 可知 $(\log_q x)' = \frac{1}{x \ln q}$.

由 (3.3) 式可知 $\left(\mathrm{e}^x\right)' = \mathrm{e}^x$. 当 $a > 0$ 时, 由 (3.4) 式可知 $\left(a^x\right)' = a^x \ln a$.

在**定义 4.1.2**(微分)之前, 回顾 Newton 迭代:

为了体会这种基本手法, 我们回顾一下 Newton 方法. 设 $P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$ 为 n 次多项式, ξ 是它的一个根. 我们想求 ξ 的近似值. 若已知 x_i 为 ξ 的粗略的近似值, 其误差记为 $h = \xi - x_i$. 则

$$0 = P(\xi) = P(x_i + h) = P(x_i) + (na_0 x_i^{n-1} + \cdots)h + \text{含 } h \text{ 的高阶项}.$$

微分学的基本想法是, 当 $|h|$ 很小时, 高阶项可以省略掉, 只保留一次项:

$$0 = P(\xi) \approx P(x_i) + P'(x_i)h, \quad \text{即} \quad h \approx -\frac{P(x_i)}{P'(x_i)}.$$

记 $x_{i+1} = x_i - \frac{P(x_i)}{P'(x_i)}$, 则 x_{i+1} 是 ξ 的一个新的近似值(通常是更好的近似值). 在一定的条件下, 这个迭代公式给出的数列 $\{x_i\}$ 将会很快地收敛于 ξ.

定义 4.1.2(微分) 设函数 f 在 x_0 附近有定义. 如果存在 $\alpha \in \mathbb{R}$, 使得

$$f(x_0 + h) = f(x_0) + \alpha h + o(h), \ \ (h \to 0),$$

则称 f 在 x_0 处可微, 线性函数 $y = \alpha h$ (h 是自变量) 称为 f 在 x_0 处的微分, 记为 $df(x_0)$.

命题 4.1.3 的证明改为:

证明 设 f 在 x_0 处可导, 记 $R(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$, 则

$$\lim_{h \to 0} \frac{R(h)}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) = 0,$$

即 $R(h) = o(h)$ ($h \to 0$), 从而 f 在 x_0 处可微.

反之, 设 f 在 x_0 处可微, $f(x_0+h)=f(x_0)+\alpha h+o(h)$ $(h\to 0)$, 则

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}=\alpha+\lim_{h\to 0}\frac{o(h)}{h}=\alpha,$$

这说明 f 在 x_0 处可导, 且 $f'(x_0)=\alpha$.

习题 4.1:

21. 设 $a_{ij}(x)$ 均为可导函数, 记 $f(x)=\det A(x)$, 其中 $A(x)=\big(a_{ij}(x)\big)_{n\times n}$. 若 $f(x_0)\neq 0$, 证明 $f'(x_0)=f(x_0)\mathrm{tr}\big(A^{-1}(x_0)A'(x_0)\big)_{n\times n}$, 其中 $A'(x_0)=\big(a'_{ij}(x_0)\big)_{n\times n}$.

4.2 Newton-Leibniz 公式

命题 4.2.4

证明

任取 $x_0\in[a,b]$. 由积分关于区间的可加性以及积分中值定理可得

$$g(x)=g(x_0)+\int_{x_0}^{x}f(t)\,\mathrm{d}t=g(x_0)+f(\xi_x)(x-x_0),$$

其中 $\xi_x=x_0+\theta(x-x_0)$, $\theta\in[0,1]$. 当 $x\to x_0$ 时, $\xi_x\to x_0$, 由 f 在 x_0 处连续可知 $g'(x_0)=f(x_0)$.

例 4.2.6 设 $x>0$ 且 $x\neq 1$, 则 $\frac{x-1}{x}<\ln x<x-1$.

证明 记 $f(x)=\ln x-x+1$, 则 $f'(x)=\frac{1}{x}-1$. 当 $0<x<1$ 时 $f'(x)>0$; 当 $x>1$ 时 $f'(x)<0$. 因此 f 在 $(0,1]$ 中严格单调递增, 在 $[1,+\infty)$ 中严格单调递减, 这说明 $f(x)\leq f(1)=0$, 等号仅在 $x=1$ 处成立.

另一方面, 也有 $\ln x=-\ln\frac{1}{x}>1-\frac{1}{x}=\frac{x-1}{x}$.

习题 4.2:

5. 用本节中的方法证明 $\mathrm{e}^x\geq 1+x+\frac{1}{2}x^2+\frac{1}{6}x^3$.

4.3 积分的计算方法

4.4 简单的微分方程

例 4.4.5 设 $\omega>0$ 为常数, 求解二次微分方程 $\varphi''=-\omega^2\varphi$.

注记 易见 $C_1=\varphi(0)$, $C_2=\frac{1}{\omega}\varphi'(0)$, 因此 φ 完全由 $\varphi(0)$ 和 $\varphi'(0)$ 的值所确定. 在物理学中, φ 常表示简谐振动, 我们考虑其总能量 $E=\frac{1}{2}(\varphi')^2+\frac{1}{2}\omega^2\varphi^2$, 求导可得

$$E'=\varphi'\varphi''+\omega^2\varphi\varphi'=\varphi'(\varphi''+\omega^2\varphi)=0,$$

这代表的是能量守恒. 特别地, 若 $\varphi(0)=\varphi'(0)=0$, 则 $\varphi\equiv 0$.

上述方法常用来研究微分方程解的唯一性, 我们再做一点简单介绍.

引理 4.4.1 设 f 在区间 (a,b) 中可导且 $|f'| \le M|f|$, 其中 $M \ge 0$ 为常数. 如果存在 $t_0 \in (a,b)$, 使得 $f(t_0) = 0$, 则 $f \equiv 0$.

证明 记 $E = f^2$, 则 $|E'| = |2ff'| \le 2ME$. 于是

$$\big[\mathrm{e}^{2Mt}E(t)\big]' = \mathrm{e}^{2Mt}\big[2ME(t) + E'(t)\big] \ge 0,$$

这说明 $\mathrm{e}^{2Mt}E(t)$ 单调递增. 特别地, 当 $t < t_0$ 时

$$0 \le \mathrm{e}^{2Mt}E(t) \le \mathrm{e}^{2Mt_0}E(t_0) = 0,$$

这说明此时 $f(t) = 0$. 同理, 计算表明 $\mathrm{e}^{-2Mt}E(t)$ 单调递减, 于是可说明当 $t > t_0$ 时也有 $f(t) = 0$.

定理 4.4.2 设 $P_1(t), P_2(t), \cdots, P_n(t)$ 均在区间 (a,b) 中连续, $f(t)$ 满足方程

$$f^{(n)}(t) + P_1(t)f^{(n-1)}(t) + \cdots + P_{n-1}(t)f'(t) + P_n(t)f(t) = 0. \tag{4.1}$$

如果存在 $t_0 \in (a,b)$, 使得 $f(t_0) = f'(t_0) = \cdots = f^{(n-1)}(t_0) = 0$, 则 $f(t) \equiv 0$.

证明 任取 (a,b) 中包含 t_0 的闭区间 I. 根据已知条件, 存在常数 C, 使得在 I 中 $|P_i| \le C$ 均成立. 此时, 在 I 中就有

$$|f^{(n)}| \le C\big[|f| + |f'| + \cdots + |f^{(n-1)}|\big]. \tag{4.2}$$

记 $E = f^2 + (f')^2 + \cdots + (f^{(n-1)})^2$, 则 $E(t_0) = 0$, 且

$$E' = 2ff' + 2f'f'' + \cdots + 2f^{(n-2)}f^{(n-1)} + 2f^{(n-1)}f^{(n)}$$

由 (4.2) 以及不等式 $2ab \le a^2 + b^2$ 可得 $|E'| \le ME$, 其中 M 是只与 n, C 有关的常数. 由前一引理可知 E 在 I 中恒等于零, 从而 f 在 I 中恒为零. 由此容易看出 f 在 (a,b) 中也恒为零.

下面三角函数的分析学定义中解的唯一性就可以省略一些了:

(1) 解的唯一性和有界性.

根据上述讨论可知方程 (4.22), (4.23) 的解具有唯一性, 且 $f^2 + g^2 \equiv 1$, 特别地, f, g 均为有界函数.

4.5 附录: 有理函数的分解

计算有理函数积分时往往需要先将有理函数分解为简单的真分式的和. 下面我们来补充说明为什么有理函数总是可以做这样的分解.

为了方便起见, 先考虑复系数复值多项式. 设 $p(z) = \sum\limits_{k=0}^{n} a_k z^k$ 是 n 次多项式, 其中 $a_0, a_1, \cdots, a_n \in \mathbb{C}$, $a_n \ne 0$. 记 $p'(z) = \sum\limits_{k=1}^{n} k a_k z^{k-1}$, 这是 $n-1$ 次多项式, 称为 p 的导数. 如果 $p(z)$, $q(z)$ 均为多项式, 则有理函数 $\frac{q(z)}{p(z)}$ 的导数定义为

$$\Big(\frac{q}{p}\Big)'(z) = \frac{p(z)q'(z) - p'(z)q(z)}{p^2(z)}.$$

不难验证此定义的恰当性. 同样地可以定义有理函数的高阶导数, 且 Leibniz 公式之类的求导法则仍然成立.

引理 4.5.1. 设 $p(z)$ 是次数不超过 n 的多项式, $z_0 \in \mathbb{C}$, 则

$$p(z) = p(z_0) + \sum_{k=1}^{n} \frac{p^{(k)}(z_0)}{k!}(z - z_0)^k. \tag{4.3}$$

证明 只要对 $p(z) = 1, z, \cdots, z^n$ 验证 (4.3) 式成立就可以了. $p(z) = 1$ 的情形是显然的. 若 $p(z) = z^m (1 \le m \le n)$, 则由 Newton 二项式展开可得

$$z^m = \big[(z - z_0) + z_0\big]^m = z_0^m + \sum_{k=1}^{m} C_m^k z_0^{m-k}(z - z_0)^k.$$

注意到当 $k \le m$ 时 $(z^m)^k(z_0) = k! C_m^k z_0^{m-k}$; 当 $k > m$ 时 $(z^m)^{(k)} = 0$. 于是 (4.3) 式对 $p(z) = z^m$ 的确成立.

推论 4.5.2. 设 $p(z)$ 是 n 次多项式, 则 z_0 是 $p(z)$ 的 k 重根当且仅当

$$p(z_0) = p'(z_0) = \cdots = p^{(k-1)}(z_0) = 0, \ \ p^{(k)}(z_0) \ne 0. \tag{4.4}$$

证明 若 z_0 是 $p(z)$ 的 k 重根, 则 $p(z) = (z - z_0)^k q(z)$, 其中 $q(z_0) \ne 0$. 求导并利用 Leibniz 公式即知 (4.4) 式成立.

反之, 若 (4.4) 式成立, 则由 (4.3) 式可得

$$p(z) = (z - z_0)^k \Big[\frac{p^{(k)}(z_0)}{k!} + \frac{p^{(k+1)}(z_0)}{(k+1)!}(z - z_0) + \cdots + \frac{p^{(n)}(z_0)}{n!}(z - z_0)^{n-k}\Big],$$

这说明 z_0 是 $p(z)$ 的 k 重根.

推论 4.5.3. 设 $r(z)$ 是 n 次多项式, $r(z_0) \ne 0$, $1 \le m \le n$. 记

$$f(z) = r(z)\sum_{s=0}^{m-1} \frac{c_s}{s!}(z - z_0)^s - 1,$$

其中 $c_s = \big(\frac{1}{r}\big)^{(s)}(z_0)$, 则 z_0 是 $f(z)$ 的根, 其重数至少为 m.

证明 $f(z)$ 可以改写为 $r(z)g(z)$, 其中

$$g(z) = \sum_{s=0}^{m-1} \frac{c_s}{s!}(z - z_0)^s - \frac{1}{r(z)}.$$

由 c_s 的定义可知 $g(z_0) = g'(z_0) = \cdots = g^{(m-1)}(z_0) = 0$. 于是利用 Leibniz 公式可知 $f(z_0) = f'(z_0) = \cdots = f^{(m-1)}(z_0) = 0$. 由前一推论即知 z_0 的重数至少为 m.

设 $q(z)$, $p(z)$ 均为多项式, 我们想将有理函数 $\frac{q(z)}{p(z)}$ 分解为若干简单有理函数的和. 先考虑 $\frac{1}{p(z)}$ 的分解. 首先, 若 $z_1 \neq z_2$, 则

$$\frac{1}{(z-z_1)(z-z_2)} = \frac{1}{z_1-z_2}\frac{1}{z-z_1} + \frac{1}{z_2-z_1}\frac{1}{z-z_2}.$$

其次, 若 z_1, z_2, z_3 互不相等, 则由上式可得

$$\frac{1}{(z-z_1)(z-z_2)(z-z_3)} = \frac{1}{z_1-z_2}\frac{1}{z-z_1}\frac{1}{z-z_3} + \frac{1}{z_2-z_1}\frac{1}{z-z_2}\frac{1}{z-z_3},$$

其中上式右端两项都可以如上继续分解. 总结起来就有

引理 4.5.4. 设 $p(z)$ 是 n 次多项式. 若 $\{z_k\}_{k=1}^n$ 是 $p(z)$ 的 n 个互不相同的根, 则

$$\frac{1}{p(z)} = \sum_{k=1}^n \frac{1}{p'(z_k)}\frac{1}{z-z_k}. \tag{4.5}$$

证明 由以上讨论和归纳法可知, 存在常数 $\{c_k\}_{k=1}^n$, 使得

$$\frac{1}{p(z)} = \sum_{k=1}^n \frac{c_k}{z-z_k}.$$

下面我们来确定这些常数的值. 记 $p(z) = (z-z_k)q_k(z)$, 求导可得 $q_k(z_k) = p'(z_k)$. 另一方面, 注意到

$$\frac{1}{q_k(z)} = c_k + \sum_{l\neq k} c_l \frac{z-z_k}{z-z_l}.$$

在上式中令 $z = z_k$ 即得 $\frac{1}{p'(z_k)} = c_k$.

由此还可以得到一个有趣的结论:

推论 4.5.5. 设 $p(z)$ 是 n 次多项式. 若 $\{z_k\}_{k=1}^n$ 是 $p(z)$ 的 n 个互不相同的根, 则当 $n > 1$ 时必有

$$\sum_{k=1}^n \frac{1}{p'(z_k)} = 0. \tag{4.6}$$

证明 在 (4.5) 两边乘以 z, 再令 $z \to \infty$ 即可.

我们再来看 $p(z)$ 有重根的情形. 若 $z_0 \neq z_1$, $k > 1$, 则

$$\frac{1}{(z-z_0)(z-z_1)^k} = \frac{1}{z_1-z_0}\frac{1}{(z-z_1)^k} + \frac{1}{z_0-z_1}\frac{1}{(z-z_0)(z-z_1)^{k-1}}.$$

若 $k > 2$, 上式右端第二项还可以继续分解.

一般地, 我们有

定理 4.5.6. 设 $p(z)$ 是 n 次多项式, $\{z_i\}_{i=1}^k$ 是其单根, $\{w_j\}_{j=1}^l$ 是其重根, 其中 w_j 的重数是 m_j, 则

$$\frac{1}{p(z)} = \sum_{i=1}^k \frac{c_i}{z-z_i} + \sum_{j=1}^l \sum_{s=0}^{m_j-1} \frac{c_{j,s}}{(z-w_j)^{m_j-s}}, \tag{4.7}$$

其中

$$c_i = \frac{1}{p'(z_i)},\ 1 \le i \le k;\ \ c_{j,s} = \frac{1}{s!}\Big[\frac{(z-w_j)^{m_j}}{p(z)}\Big]^{(s)}(w_j),\ 0 \le s < m_j,\ 1 \le j \le l. \tag{4.8}$$

证明 我们先假设 (4.7) 式成立. 此时, 用引理 4.5.4 中的方法可以确定 $c_i = \frac{1}{p'(z_i)}$. 当 $1 \le j \le l$ 时,

$$\frac{(z-w_j)^{m_j}}{p(z)} - \Big[\sum_{i=1}^{k}\frac{c_i}{z-z_i} + \sum_{t\neq j}\sum_{s=0}^{m_t-1}\frac{c_{t,s}}{(z-w_t)^{m_t-s}}\Big](z-w_j)^{m_j} = \sum_{s=0}^{m_j-1} c_{j,s}(z-w_j)^s.$$

注意上式左边第二项在 w_j 处的 $s(s<m_j)$ 阶导数均为零, 于是由引理 4.5.1 即知 $c_{j,s}$ 必定由 (4.8) 式所确定.

设 c_i, $c_{j,s}$ 由 (4.8) 式所确定, 我们来说明 (4.7) 式确实成立. 为此, 令

$$h(z) = 1 - p(z)\Big[\sum_{i=1}^{k}\frac{c_i}{z-z_i} + \sum_{j=1}^{l}\sum_{s=0}^{m_j-1}\frac{c_{j,s}}{(z-w_j)^{m_j-s}}\Big],$$

则 $h(z)$ 是次数不超过 $n-1$ 的多项式. 根据前一段的讨论以及推论 4.5.3 可知, $p(z)$ 的根均为 $h(z)$ 的根(含重数). 但次数不超过 $n-1$ 的非零多项式最多只有 $n-1$ 个根, 这说明 $h(z) \equiv 0$.

对于一般的有理函数 $\frac{q(z)}{p(z)}$, 利用带余除法, 不妨设 $q(z)$ 的次数小于 $p(z)$ 的次数, 且 $q(z)$ 和 $p(z)$ 无公共根. 若 $p(z)$ 如上述定理, 则

$$\frac{q(z)}{p(z)} = \sum_{i=1}^{k}\frac{c_i}{z-z_i} + \sum_{j=1}^{l}\sum_{s=0}^{m_j-1}\frac{c_{j,s}}{(z-w_j)^{m_j-s}}, \tag{4.9}$$

其中

$$c_i = \frac{q(z_i)}{p'(z_i)},\ 1 \le i \le k;\ \ c_{j,s} = \frac{1}{s!}\Big[\frac{q(z)(z-w_j)^{m_j}}{p(z)}\Big]^{(s)}(w_j),\ 0 \le s < m_j,\ 1 \le j \le l.$$

证明方法与上述定理完全类似, 我们留给读者.

如果 $q(z), p(z)$ 为实系数多项式, 则 $p(z)$ 可能会有单实根、重实根、单复根、重复根等. 不过, 实系数多项式的复根必定成对出现, 因此可以在 (4.9) 式中将成对出现的项予以合并. 比如, 若 $z_0 = \lambda + \mu\mathrm{i}$ 和 $\bar{z}_0 = \lambda - \mu\mathrm{i}$ 是 $p(z)$ 的一对单复根, 则当 $z = t \in \mathbb{R}$ 时

$$\frac{c}{z-z_0} + \frac{\bar{c}}{z-\bar{z}_0} = \frac{at-b}{(t-\lambda)^2+\mu^2},$$

其中 $a = c+\bar{c} \in \mathbb{R}$, $b = c\bar{z}_0 + \bar{c}z_0 \in \mathbb{R}$. 若 $z_0, \bar{z}_0$ 是一对 m 重复根, 则当 $z = t \in \mathbb{R}$, $1 \le j \le m$ 时,

$$\frac{c}{(z-z_0)^j} + \frac{\bar{c}}{(z-\bar{z}_0)^j} = \frac{r(t)}{\big[(t-\lambda)^2+\mu^2\big]^j},$$

其中 $r(t)=c(t-\bar{z}_0)^j+\bar{c}(t-z_0)^j$ 是实系数多项式. 若 $r(t)$ 的次数大于 1, 则由带余除法可知

$$r(t)=r_1(t)\big[(t-\lambda)^2+\mu^2\big]+r_0(t),$$

其中 $r_0(t)$ 的次数不超过 1. 此时

$$\frac{r(t)}{\big[(t-\lambda)^2+\mu^2\big]^j}=\frac{r_1(t)}{\big[(t-\lambda)^2+\mu^2\big]^{j-1}}+\frac{r_0(t)}{\big[(t-\lambda)^2+\mu^2\big]^j}.$$

若 $r_1(t)$ 的次数仍大于 1, 则可以继续分解. 总之, 有理函数 $\frac{q(t)}{p(t)}$ 可以分解为如下几种简单函数的和:

$$\text{多项式},\ \ \frac{b}{(t-a)^i},\ \ \frac{ct+d}{\big[(t-\lambda)^2+\mu^2\big]^j},$$

其中 b,c,d 等均为常数, a 对应着 $p(t)$ 的实根, $\lambda\pm\mu\mathrm{i}$ 对应着 $p(t)$ 的复根.

第五章 微分学的应用

5.1 函数的极值

例 5.1.4: 最后一句中“见 (3.15) 式”改为“见**例 3.5.9**”.

习题 5.1:

8. 设 n 为正整数, $0 < x < \pi$, 证明 $0 < \sum\limits_{k=1}^{n} \frac{\sin kx}{k} < \pi$. (提示: 用归纳法研究函数的最值.)

5.2 微分中值定理

习题 5.2:

13. 设 f 在 $[a, +\infty)$ 中可导且 $f(a) = 0$, $|f'| \le 2\sqrt{|f|}$. 证明 $|f(x)| \le (x-a)^2$.

5.3 凸函数

命题 5.3.6 设 $f \in C^0(I)$, 如果 f 具有中点凸性, 则 f 为凸函数.

证明 设 $a, b \in I$, $a < b$. 记 ℓ 为满足条件 $\ell(a) = f(a)$, $\ell(b) = f(b)$ 的线性函数. 显然, 函数 $g(x) = f(x) - \ell(x)$ 仍然具有中点凸性. 由**例 3.4.1** 可知

$$g(x) \le \max\{g(a), g(b)\} = 0, \ \ \forall\, x \in [a, b].$$

这说明 f 为凸函数.

5.4 函数和曲线作图

引理 5.4.1 φ 可导的条件不够, 要改成 C^1.

5.5 L'Hospital 法则

习题 5.5:

9. 设 f 在 x_0 附近 n 阶可导, 在 x_0 处 $n+1$ 可导, 记

$$g(x)=\begin{cases}\frac{f(x)-f(x_0)}{x-x_0}, & x\neq x_0,\\ f'(x_0), & x=x_0.\end{cases}$$

证明

$$g^{(n)}(x)=\begin{cases}\frac{(-1)^n n!}{(x-x_0)^{n+1}}\Big[f(x)-f(x_0)+\sum\limits_{k=1}^{n}(-1)^k\frac{f^{(k)}(x)}{k!}(x-x_0)^k\Big], & x\neq x_0,\\ \frac{f^{(n+1)}(x_0)}{n+1}, & x=x_0.\end{cases}$$

进一步, 若 f 在 x_0 附近 $n+1$ 可导, 则

$$g^{(n)}(x)=\frac{f^{(n+1)}\big(x_0+\theta(x-x_0)\big)}{n+1},$$

其中 $\theta=\theta(x)\in(0,1)$; 若 $f^{(n+1)}$ 连续, 则

$$g^{(n)}(x)=\frac{1}{(x-x_0)^{n+1}}\int_{x_0}^{x}(t-x_0)^n f^{(n+1)}(t)\,\mathrm{d}t.$$

5.6 Taylor 展开

定理 5.6.3 后增加(恢复)一个例子:

例 5.6.3 积分余项的一个应用.

设 $a,b\in\mathbb{R}$, $n\geq 0$, 考虑函数 $f(t)=(t-a)^{2n+1}$. 显然, 当 $k\leq n$ 时 $f^{(k)}(a)=0$. 由 (5.13) 式可得

$$(b-a)^{2n+1}=\int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}\,\mathrm{d}t,$$

整理以后可得

$$\int_a^b (t-a)^n(b-t)^n\,\mathrm{d}t=\frac{(n!)^2}{(2n+1)!}(b-a)^{2n+1}.$$

习题 5.6:

5. 设 $m,n\geq 0$, 利用积分余项计算积分 $\displaystyle\int_a^b (t-a)^m(b-t)^n\,\mathrm{d}t$.

6. 用 Lagrange 余项证明 Bernoulli 不等式: 设 $x\geq -1$, 则 $0<\alpha<1$ 时 $(1+x)^\alpha\leq 1+\alpha x$; $\alpha<0$ 或 $\alpha>1$ 时 $(1+x)^\alpha\geq 1+\alpha x$.

9. 记 $a_n=\tan^{(n)}(0)$, 利用 $(\tan x)'=1+\tan^2 x$ 和 Taylor 公式系数的唯一性证明

$$a_1=1,\ \ a_{2n+1}=\sum_{k=1}^{n}C_{2n}^{2k-1}a_{2k-1}a_{2(n-k)+1}.$$

5.7 进一步应用举例

第六章　积分的推广和应用

6.1　Riemann 积分

6.2　Riemann 积分的性质

命题 6.2.8[阶梯逼近] 设 $f \in R[a,b]$, 则存在两列阶梯函数 φ_n, ψ_n, 使得

$$\varphi_n \le f \le \psi_n, \ \int_a^b \big[\psi_n(x) - \varphi_n(x)\big]\,\mathrm{d}x < \frac{1}{n}, \tag{6.1}$$

且每一个 φ_n, ψ_n 均均介于 f 的上下确界之间. 此外, 任给 $g \in R[a,b]$, 均有

$$\lim_{n\to\infty}\int_a^b \varphi_n(x)g(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_a^b \psi_n(x)g(x)\,\mathrm{d}x = \int_a^b f(x)g(x)\,\mathrm{d}x. \tag{6.2}$$

证明. 由题设, 任给 $n \ge 1$, 存在 $[a,b]$ 的分割

$$\pi:\ a = x_0 < x_1 < x_2 < \cdots < x_m = b,$$

使得 $\sum\limits_{i=1}^{m}(M_i - m_i)\Delta x_i < 1/n$, 其中 M_i, m_i 分别是 f 在 $[x_{i-1}, x_i]$ 中的上确界和下确界.

在 $[a,b]$ 中定义阶梯函数 φ_n, ψ_n 如下: 当 $x \in [x_{i-1}, x_i)$ 时 $\varphi_n(x) = m_i$, $\psi_n(x) = M_i$; 规定 $\varphi_n(b) = m_n$, $\psi_n(b) = M_n$. 显然, φ_n, ψ_n 均介于 f 的上下确界之间, 且满足 (6.1) 式.

当 $g \in R[a,b]$ 时, 可设 $|g| \le M$. 此时

$$|\varphi_n g - fg| \le M(\psi_n - \varphi_n), \ |\psi_n g - fg| \le M(\psi_n - \varphi_n),$$

这说明 (6.2) 式成立.

习题增加：

15. 设 $f \in R[a,b]$. 证明: 任给 $\varepsilon > 0$, 均存在多项式 P, Q, 使得

$$P \le f \le Q, \ \int_a^b \big[Q(x) - P(x)\big]\,\mathrm{d}x < \varepsilon.$$

(提示: 先将命题 6.2.8 中的阶梯函数改造为分段线性的连续函数, 再用 Weierstrass 逼近定理.)

6.3 广义积分

命题 6.3.2[广义分部积分] 设 f 在 $[a,b]$ 中的瑕积分收敛, $g\in R[a,b]$, 则 fG 在 $[a,b]$ 中的瑕积分也收敛, 其中

$$F(x)=\int_a^x f(t)\,\mathrm{d}t+C_1,\ \ G(x)=\int_a^x g(t)\,\mathrm{d}t+C_2,$$

此时有

$$\int_a^b f(x)G(x)\,\mathrm{d}x=F(x)G(x)\Big|_a^b-\int_a^b g(x)F(x)\,\mathrm{d}x.$$

证明 不妨设 a 是 f 唯一可能的瑕点. 当 $a'\in(a,b)$ 时, 由 **定理 6.2.10**可知

$$\int_{a'}^b f(x)G(x)\,\mathrm{d}x=F(x)G(x)\Big|_{a'}^b-\int_{a'}^b g(x)F(x)\,\mathrm{d}x.$$

由题设可知 $F,G\in C^0[a,b]$, 于是 $gF\in R[a,b]$. 在上式中令 $a'\to a^+$ 即得欲证结论.

习题 6.3: 替换 8 为

8. (1) 设 f 在 $[a,b]$ 中的瑕积分收敛, g 是 $[a,b]$ 中定义的单调函数, 则 fg 在 $[a,b]$ 中的瑕积分也收敛.

(2) 请将积分第二中值定理推广到广义积分的情形.

(3) 请说明, 若 f,g 在 $[a,b]$ 中的瑕积分都收敛且 $|f|$, $|g|$ 中至少有一个瑕积分收敛, 则**命题 6.3.2**的结论仍然成立.

6.4 广义积分的收敛判别法

定理 6.4.2 后注记 (3) 改为:

(3) 类似地, 考虑以 a 为瑕点的瑕积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$, 有:

(i) 如果 $p<1$, 且存在常数 $C>0$, 使得 $0\le f(x)\le C(x-a)^{-p}$, 则瑕积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛;

(ii) 如果 $p\ge 1$, 且存在常数 $C>0$, 使得 $f(x)\ge C(x-a)^{-p}$, 则瑕积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 发散;

我们往往通过求极限去寻找 C. 如果 $\displaystyle\lim_{x\to a^+}(x-a)^p f(x)=l$ 存在, 则

(iii) 如果 $p<1$, $0\le l<+\infty$, 则瑕积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 收敛;

(iv) 如果 $p\ge 1$, $0<l\le+\infty$, 则瑕积分 $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ 发散.

6.5 积分的几何应用

6.6 进一步的例子

P200 页第二行中, "下一章"改为"数项级数".

增加习题:

7. 设 $a \geq 0$, $b > 0$, 计算下列积分:

(1) $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{x^4 + ax^2 + b}$; (2) $\displaystyle\int_0^{+\infty} \frac{x^2\,\mathrm{d}x}{x^4 + ax^2 + b}$; (3) $\displaystyle\int_0^{+\infty} \frac{\ln x\,\mathrm{d}x}{(x+a)^2 + b^2}$.

第七章　欧氏空间和连续映射

从这一章开始我们要研究多个变量的函数. 我们知道, 一元函数的诸多性质都依赖于实数的性质. 为了研究多个变量的函数, 我们要首先研究其定义域的基本性质.

7.1　内积和外积

一元函数是指以 $\mathbb{R}$ 的子集为定义域的函数. 当变化因素不只一个时, 我们就要考虑定义在一般欧氏空间中的函数. 设 n 为正整数, 记

$$\mathbb{R}^n = \{(a_1, \cdots, a_n) \mid a_i \in \mathbb{R},\ i = 1, \cdots, n\},$$

它是由 n 元有序实数组构成的集合. 以显然的方式, $\mathbb{R}^n$ 成为 $\mathbb{R}$ 上的向量空间, 称为 n 维欧氏空间. 以 $\mathbb{R}^n$ 的子集为定义域的函数称为 n 元函数, $n > 1$ 时统称多元函数.

为了研究多元函数, 我们首先要研究 $\mathbb{R}^n$ 的性质, 这可以与 $\mathbb{R}$ 的性质做对比. 比如, 为了比较两个实数的大小, 我们可以考虑它们之间差的绝对值. 从几何上看, 这个绝对值可以认为是这两个实数在直线上所代表的两个点之间的距离. 为了方便起见, $\mathbb{R}^n$ 中的向量(n 元实数组)也称为 $\mathbb{R}^n$ 中的点. 给定 $\mathbb{R}^n$ 中的两个点, 怎样定义它们之间的距离? 除了距离之外, 我们还可以问怎样定义 $\mathbb{R}^n$ 中的两个向量之间的夹角等等问题. 为此, 我们引入如下概念.

定义 7.1.1 (内积). *设 V 是实数域 $\mathbb{R}$ 上的向量空间, 如果函数 $g : V \times V \to \mathbb{R}$ 满足以下条件*

(1) *任给 $v \in V$, 均有 $g(v, v) \geq 0$, 且 $g(v, v) = 0 \iff v = 0$, (正定性)*

(2) *任给 $u, v \in V$, 均有 $g(u, v) = g(v, u)$, (对称性)*

(3) *任给 $\alpha, \beta \in \mathbb{R}$, $u, v, w \in V$, 均有*

$$g(\alpha u + \beta v, w) = \alpha g(u, w) + \beta g(v, w), \text{ (线性性)}$$

则称 g 为 V 中的内积, (V, g) 称为内积空间.

我们常用 $\langle , \rangle$ 表示内积, 比如 $\langle u, v\rangle = g(u, v)$ 表示 u, v 之间的内积. 有了内积就可以定义向量的长度和向量之间的夹角. 设 $u \in V$, 记 $\|u\| = \sqrt{\langle u, u\rangle}$, 称为 u 的范数或长度.

定理 7.1.1 (Schwarz 不等式). 设 $(V, \langle , \rangle)$ 为内积空间, $u, v \in V$, 则

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|,$$

等号成立当且仅当 u, v 线性相关.

证明. 当 $u = 0$(或 $v = 0$) 时, 由内积的线性性可得

$$\langle 0,\ v\rangle = \langle 0 \cdot 0,\ v\rangle = 0\langle 0,\ v\rangle = 0,$$

此时不等式自然成立. 下设 $u, v \neq 0$, 当 $t \in \mathbb{R}$ 时, 根据内积的定义, 我们有

$$\langle u, u\rangle - 2t\langle u, v\rangle + t^2\langle v, v\rangle = \langle u - tv, u - tv\rangle \geq 0,$$

上式是关于 t 的一元二次函数, 因此其判别式非正:

$$\Delta = 4\langle u, v\rangle^2 - 4\langle u, u\rangle\langle v, v\rangle \leq 0,$$

等号成立的条件留作练习.

注. 不考虑等号成立的条件, 只要 $\langle , \rangle$ 具有非负性, Schwarz 不等式就仍然成立.

根据 Schwarz 不等式, 当 u, v 为非零向量时, 可以取 $\theta(u, v) \in [0, \pi]$, 使得

$$\cos\theta(u, v) = \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|},$$

$\theta(u, v)$ 称为 u, v 之间的夹角, 也记为 $\angle(u, v)$.

例 7.1.1. $\mathbb{R}^n$ 中的标准内积.

设 $u, v \in \mathbb{R}^n$, 记 $u = (u_1, \cdots, u_n)$, $v = (v_1, \cdots, v_n)$, 定义

$$\langle u, v\rangle = u \cdot v = \sum_{i=1}^{n} u_i v_i,$$

容易看出 $\langle , \rangle$ 为 $\mathbb{R}^n$ 中的内积, 称为标准内积或欧氏内积. 此时, u 的范数为

$$\|u\| = \left(u_1^2 + u_2^2 + \cdots + u_n^2\right)^{1/2}.$$

当 $n = 1$ 时, 范数也就是绝对值.

例 7.1.2. 矩阵空间中的内积.

用 $M_{m\times n}$ 表示 $m\times n$ 型的实矩阵的全体构成的向量空间. 设 $A,B\in M_{m\times n}$, 定义

$$\langle A,B\rangle=\operatorname{tr}\big(AB^{\top}\big),$$

不难验证 $\langle,\rangle$ 为 $M_{m\times n}$ 中的内积. 若 $A=\big(a_{ij}\big)_{m\times n}$, 则其范数可以表示为

$$\|A\|=\Big(\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^{2}\Big)^{\frac{1}{2}}.$$

下面的结果建立了矩阵和向量范数之间的某种联系.

引理 7.1.2 (范数不等式)**.** 设 $A=\big(a_{ij}\big)_{m\times n}$, $b\in\mathbb{R}^n$, 则 $\|Ab\|\le\|A\|\,\|b\|$.

证明 我们把矩阵 A 的第 i 行记为 a_i, 则 Ab 用列向量表示为

$$Ab=\big(a_1\cdot b,\ a_2\cdot b,\cdots,\ a_m\cdot b\big)^{\top}.$$

根据 Cauchy 不等式可得

$$\|Ab\|^2=\sum_{i=1}^{m}(a_i\cdot b)^2\le\sum_{i=1}^{m}\|a_i\|^2\,\|b\|^2=\|A\|^2\,\|b\|^2.$$

有了范数就可以定义距离. 设 $(V,\ \langle,\rangle)$ 为内积空间, $u,v\in V$. 记 $d(u,v)=\|u-v\|$, 称为 u,v 之间的距离.

命题 7.1.3 (三角不等式)**.** 设 $(V,\ \langle,\rangle)$ 为内积空间, $x,y,z\in V$, 则有

$$d(x,z)\le d(x,y)+d(y,z).$$

证明. 我们先证明绝对值不等式: 任给 $u,v\in V$, 均有

$$\|u+v\|\le\|u\|+\|v\|.\tag{7.1}$$

事实上, 根据 Schwarz 不等式, 有

$$\begin{aligned}\|u+v\|^2&=\langle u+v,u+v\rangle=\langle u,u\rangle+2\langle u,v\rangle+\langle v,v\rangle\\&\le\|u\|^2+2\|u\|\cdot\|v\|+\|v\|^2\\&=(\|u\|+\|v\|)^2.\end{aligned}$$

于是, 当 $x,y,z\in V$ 时, 利用绝对值不等式可得

$$\begin{aligned}d(x,z)&=\|x-z\|=\|(x-y)+(y-z)\|\\&\le\|x-y\|+\|y-z\|=d(x,y)+d(y,z).\end{aligned}$$

设 $u,v\in\mathbb{R}^n$, 若 $u\cdot v=0$, 则称 u,v 互相垂直(正交), 记为 $u\perp v$. 第 i 个位置为 1, 其余位置为零的向量记为 e_i, 我们将 $\{e_i\}_{i=1}^{n}$ 称为 $\mathbb{R}^n$ 中的标准正交基.

在 $\mathbb{R}^n$ 中给定 $n-1$ 个向量 $v^1, \cdots, v^{n-1}$, 我们要构造另一向量 w, 使得 w 与 v^i 均垂直. 为此, 记 $v_i = (v_1^i, \cdots, v_n^i)$, 考虑线性函数 $\ell: \mathbb{R}^n \to \mathbb{R}$:

$$\ell(x) = \begin{vmatrix} x_1 & \cdots & x_n \\ v_1^1 & \cdots & v_n^1 \\ \vdots & \ddots & \vdots \\ v_1^{n-1} & \cdots & v_n^{n-1} \end{vmatrix}.$$

按照行列式的定义, $\ell(x)$ 可以表示为

$$\ell(x) = x \cdot w = \sum_{i=1}^n x_i w_i, \tag{7.2}$$

其中 $w = (w_1, \cdots, w_n)$, 且

$$w_i = (-1)^{i-1} \begin{vmatrix} v_1^1 & \cdots & v_{i-1}^1 & v_{i+1}^1 & \cdots & v_n^1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_1^{n-1} & \cdots & v_{i-1}^{n-1} & v_{i+1}^{n-1} & \cdots & v_n^{n-1} \end{vmatrix}, \quad i = 1, \cdots, n. \tag{7.3}$$

由 $\ell(x)$ 的定义易见 $\ell(v^i) = 0$ $(1 \le i \le n-1)$. 再由 (7.2) 可知 w 的确与 v_i 都垂直. 我们称 w 是 $v^1, \cdots, v^{n-1}$ 的外积, 记为

$$w = v^1 \times \cdots \times v^{n-1}.$$

外积还具有下列性质:

- $\{v^i\}_{i=1}^{n-1}$ 线性相关当且仅当 $v^1 \times \cdots \times v^{n-1} = 0$. 一方面, 如果 $\{v^i\}_{i=1}^{n-1}$ 线性相关, 则 $\ell(x) \equiv 0$, 从而 $\|w\|^2 = \ell(w) = 0$; 另一方面, 如果 $\{v^i\}_{i=1}^{n-1}$ 线性无关, 则可以将它们扩充为 $\mathbb{R}^n$ 中的一组基 $\{u, v_i \mid 1 \le i \le n-1\}$. 此时 $\ell(u) \ne 0$, 即 $u \cdot w \ne 0$, 特别地 $w \ne 0$.

- 外积关于关于每一个分量 v^i 都是线性的, 即任给 $\lambda, \mu \in \mathbb{R}$ 以及向量 $v^i, \tilde{v}^i$, 有

$$v^1 \times \cdots \times (\lambda v^i + \mu \tilde{v}^i) \times \cdots \times v^{n-1} = \lambda v^1 \times \cdots \times v^i \times \cdots \times v^{n-1} + \mu v^1 \times \cdots \times \tilde{v}^i \times \cdots \times v^{n-1}.$$

- 当 $i < j$ 时, 交换 v^i 和 v^j 的位置以后外积变一个符号, 即

$$v^1 \times \cdots \times v^j \times \cdots \times v^i \times \cdots \times v^{n-1} = -v^1 \times \cdots \times v^i \times \cdots \times v^j \times \cdots \times v^{n-1}.$$

  一般地, 设 σ 为 $\{1, \cdots, n-1\}$ 的置换, 则

$$v^{\sigma(1)} \times \cdots \times v^{\sigma(n-1)} = (-1)^\sigma v^1 \times \cdots \times v^{n-1},$$

  其中 σ 为偶置换时 $(-1)^\sigma = 1$, σ 为奇置换时 $(-1)^\sigma = -1$.

- $\|v^1\times\cdots\times v^{n-1}\| = \left[\det\left(v^i\cdot v^j\right)_{(n-1)\times(n-1)}\right]^{1/2}$. 事实上, 当$\{v^i\}_{i=1}^{n-1}$ 线性相关时 $w=0$, $\det\left(v^i\cdot v^j\right)_{(n-1)\times(n-1)}=0$; 当$\{v^i\}_{i=1}^{n-1}$ 线性无关时, 利用 $\|w\|^2=\ell(w)$ 以及 $v^i\cdot w=0$ 可得

$$\begin{aligned}\|w\|^4 &= \det\begin{pmatrix} w\\ v^1\\ \vdots\\ v^{n-1}\end{pmatrix}\begin{pmatrix} w\\ v^1\\ \vdots\\ v^{n-1}\end{pmatrix}^\top = \det\begin{pmatrix} w\cdot w & 0\\ 0 & \left(v^i\cdot v^j\right)_{(n-1)\times(n-1)}\end{pmatrix}\\ &= \|w\|^2\det\left(v^i\cdot v^j\right)_{(n-1)\times(n-1)},\end{aligned}$$

  这说明 $\|w\|^2=\det\left(v^i\cdot v^j\right)_{(n-1)\times(n-1)}$.

- 若 $\{e_i\}_{i=1}^n$ 为 $\mathbb{R}^n$ 中的标准正交基, 则 $e_1\times\cdots\times e_{i-1}\times e_{i+1}\times\cdots\times e_n=(-1)^{i-1}e_i$.

习题 7.1

1. 我们用 $C[a,b]$ 表示 $[a,b]$ 中连续函数的全体构成的空间. 设 $f,g\in C[a,b]$, 定义
$$\langle f,g\rangle=\int_a^b f(x)g(x)\,\mathrm{d}x,$$
验证 $\langle,\rangle$ 是 $C[a,b]$ 中的内积.

2. 设 $\langle,\rangle$ 为向量空间 V 中的内积, $u,v\in V$, $v\neq 0$. 在不等式
$$\langle u-tv,u-tv\rangle\geq 0$$
中代入 $t=\frac{\langle u,v\rangle}{\langle v,v\rangle}$ 从而证明 Schwarz 不等式.

3. 设 $(V,\langle,\rangle)$ 为内积空间, 证明如下的平行四边形公式:
$$\|u+v\|^2+\|u-v\|^2=2(\|u\|^2+\|v\|^2),\ \ \forall\, u,\ v\in V.$$

4. 设 $u\in\mathbb{R}^n$, 证明: $\|u\|=\max\{u\cdot v\mid v\in S^{n-1}\}$, 其中 $S^{n-1}=\{x\in\mathbb{R}^n\mid \|x\|=1\}$.

5. 设 A,B 为 n 阶实方阵, 证明: $\|AB\|\leq\|A\|\,\|B\|$.

6. 设 A 为 n 阶实方阵. 若 $\|A\|<1$, 证明: I_n-A 可逆, 其中 I_n 是 n 阶单位阵.

7. 设 $u,v,w\in\mathbb{R}^3$, 证明: $(u\times v)\times w+(v\times w)\times u+(w\times u)\times v=0$.

7.2 欧氏空间的拓扑

通过一元函数的学习, 我们知道无论是函数的极限还是连续性等性质都与实数的基本性质密切相关. 下面我们把实数的基本性质推广到欧氏空间中.

设 $x \in \mathbb{R}^n$, $r > 0$. 记 $B_r(x) = \{y \in \mathbb{R}^n \mid d(y,x) = \|y-x\| < r\}$, 称为以 x 为中心(球心), r 为半径的开球.

开集: 设 $U \subset \mathbb{R}^n$. 如果任给 $x \in U$, 均存在 $r > 0$, 使得 $B_r(x) \subset U$, 则称 U 为 $\mathbb{R}^n$ 中的开集. 注意, 空集也是开集. 当然, 整个 $\mathbb{R}^n$ 是开集.

闭集: 如果一个集合在 $\mathbb{R}^n$ 中的补集是开集, 则称该集合为 $\mathbb{R}^n$ 中的闭集. 空集以及整个 $\mathbb{R}^n$ 都是闭集.

例 7.2.1. 开集的简单例子.

我们先来说明开球 $B_r(x)$ 是开集. 事实上, 若 $y \in B_r(x)$, 记 $r' = r - d(y,x)$, 则 $r' > 0$. 当 $z \in B_{r'}(y)$ 时, 由三角不等式可得

$$d(z,x) \le d(z,y) + d(y,x) < r' + d(y,x) = r,$$

即 $z \in B_r(x)$, 这说明 $B_{r'}(y) \subset B_r(x)$, 由开集的定义即知 $B_r(x)$ 是开集.

设 (a_i, b_i) $(i = 1, 2, \cdots, n)$ 均为 $\mathbb{R}$ 中的开区间, 记

$$I = \prod_{i=1}^{n} (a_i, b_i) = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \mid a_i < x_i < b_i,\ 1 \le i \le n\},$$

称为 n 维开矩形. 我们来说明开矩形是开集. 事实上, 若 $x \in I$, 记

$$r = \min\{x_i - a_i,\ b_i - x_i \mid 1 \le i \le n\},$$

则 $r > 0$. 当 $y = (y_1, y_2, \cdots, y_n) \in B_r(x)$ 时, 由

$$|y_i - x_i| \le d(y,x) < r \le \min\{x_i - a_i,\ b_i - x_i\}$$

可知 $a_i < y_i < b_i$, $\forall\, 1 \le i \le n$. 这说明 $y \in I$, 即 $B_r(x) \subset I$, 于是开集的定义即知 I 是开集.

例 7.2.2. 闭集的简单例子.

设 $x \in \mathbb{R}^n$, $r > 0$. 记 $\bar{B}_r(x) = \{y \in \mathbb{R}^n \mid d(y,x) = \|y-x\| \le r\}$, 称为以 x 为中心(球心), r 为半径的闭球. 我们来说明闭球是闭集. 事实上, 若 $y \notin \bar{B}_r(x)$, 记 $r' = d(y,x) - r$, 则 $r' > 0$, 由三角不等式可知 $B_{r'}(y) \subset \mathbb{R}^n \setminus \bar{B}_r(x)$, 由定义可知 $\mathbb{R}^n \setminus \bar{B}_r(x)$ 为开集, 即 $\bar{B}_r(x)$ 是闭集.

设 $[a_i, b_i]$ $(i = 1, 2, \cdots, n)$ 均为 $\mathbb{R}$ 中的闭区间, 记

$$I = \prod_{i=1}^{n} [a_i, b_i] = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \mid a_i \le x_i \le b_i,\ 1 \le i \le n\},$$

称为 n 维(闭)矩形. 记 $\nu(I)=\prod_{i=1}^{n}(b_i-a_i)$, 称为 I 的容积. 记

$$d(I)=\left[(b_1-a_1)^2+(b_2-a_2)^2+\cdots+(b_n-a_n)^2\right]^{\frac{1}{2}},$$

称为 I 的直径. 容易验证 $d(I)=\max\{d(x,y)\mid x,y\in I\}$.

我们来说明 I 是闭集. 事实上, 若 $y=(y_1,y_2,\cdots,y_n)\notin I$, 则存在 i, 使得 $y_i\notin[a_i,b_i]$. 不妨设 $y_i<a_i$, 记 $r=a_i-y_i$, 则 $r>0$. 当 $z=(z_1,z_2,\cdots,z_n)\in B_r(y)$ 时,

$$a_i-z_i=r+y_i-z_i\geq r-\|y-z\|>0$$

即 $z\notin I$, 这说明 $B_r(y)\subset\mathbb{R}^n\setminus I$, 因此 $\mathbb{R}^n\setminus I$ 是开集, I 是闭集.

为了体会开集和闭集概念的用处, 我们再引进极限的概念: 设 $\{u_i\}_{i=1}^{\infty}\subset\mathbb{R}^n$, $u\in\mathbb{R}^n$. 如果 $\lim\limits_{i\to\infty}d(u_i,u)=\lim\limits_{i\to\infty}\|u_i-u\|=0$, 则称点列 $\{u_i\}$ 收敛于 u, 或 $\{u_i\}$ 以 u 为极限. 由三角不等式可知, 极限具有唯一性.

命题 7.2.1. (1) *设 $\{u_i\}$ 收敛于 u, 若 u 属于开集 U, 则存在 N, 使得当 $i>N$ 时 u_i 也属于 U.*

(2) *C 为闭集$\Longleftrightarrow$ C 中任何收敛点列的极限仍属于 C.*

证明 (1) 由 U 为开集可知, 存在 $r>0$, 使得 $B_r(u)\subset U$. 由极限的定义可知, 存在 N, 使得当 $i>N$ 时 $d(u_i,u)<r$, 即 $u_i\in B_r(u)\subset U$.

(2) "$\Longrightarrow$" 设 $\{u_i\}$ 为闭集 C 中的收敛点列, 其极限为 u. 若 $u\notin C$, 则由 C 为闭集可知, 存在 $r>0$, 使得 $B_r(u)\subset\mathbb{R}^n\setminus C$. 由 (1) 可知, 当 i 充分大时 $u_i\in\mathbb{R}^n\setminus C$, 这与 $\{u_i\}\subset C$ 相矛盾.

"$\Longleftarrow$" 我们来说明 $\mathbb{R}^n\setminus C$ 是开集. 设 $u\in\mathbb{R}^n\setminus C$, 我们要说明存在 $i\geq 1$, 使得 $B_{1/i}(u)\subset\mathbb{R}^n\setminus C$. (反证法) 若不然, 则对每一个正整数 i, 均存在 $u_i\in B_{1/i}(u)\cap C$. 此时 $\{u_i\}$ 收敛于 u, 这与题设相矛盾.

下面我们将数列极限的重要结果推广到点列极限的情形.

定理 7.2.2 (Bolzano)**.** *设 $\{u_i\}\subset\mathbb{R}^n$, 如果 $\{\|u_i\|\}$ 有界, 则 $\{u_i\}$ 必有收敛子列.*

证明 以 $n=2$ 为例. 记 $u_i=(x_i,y_i)$, 则由 $|x_i|,|y_i|\leq\|u_i\|$ 可知 $\{x_i\}$, $\{y_i\}$ 均为有界数列. 于是 $\{x_i\}$ 存在收敛子列, 不妨设它本身收敛 x_0. 注意 $\{y_i\}$ 也有收敛子列, 比如 $\{y_{i_j}\}$ 收敛于 y_0. 记 $u_0=(x_0,y_0)$, 由

$$\|u_{i_j}-u_0\|=\sqrt{(x_{i_j}-x_0)^2+(y_{i_j}-y_0)^2}\leq|x_{i_j}-x_0|+|y_{i_j}-y_0|$$

即知 $\{u_{i_j}\}$ 收敛于 u_0.

引理 7.2.3 (Lebesgue 数引理)**.** *设 $\{U_\alpha\}$ 为 $\mathbb{R}^n$ 中的一族开集, 若它们的并集包含闭矩形 I, 则存在 $\lambda>0$, 使得 I 中每一个直径小于 λ 的矩形均包含于某一个 U_α 中.*

证明 (反证法) 若结论不对, 则对每一个正整数 i, 均存在直径小于 $\frac{1}{i}$ 矩形 $I_i \subset I$, 使得 I_i 不包含于任何一个 U_α 中. 取 $u_i \in I_i$, 则 $\{u_i\} \subset I$, 从而存在收敛子列. 不妨设它本身收敛于 u_0, 则 $u_0 \in I$. 由题设, u_0 属于某个 U_α. 由于 U_α 为开集, 故存在 $r > 0$, 使得 $B_r(u_0) \subset U_\alpha$. 取 $N > \frac{2}{r}$, 使得当 $i > N$ 时, $d(u_i, u_0) < \frac{r}{2}$. 此时, 任给 $u \in I_i$, 有

$$d(u, u_0) \le d(u, u_i) + d(u_i, u_0) \le d(I_i) + \frac{r}{2} < \frac{1}{i} + \frac{r}{2} < r,$$

即 $u \in B_r(u_0)$, 从而 $I_i \subset B_r(u_0) \subset U_\alpha$, 这与 I_i 的选取相矛盾.

为了叙述起来方便起见, 我们引进覆盖的概念. 若一族集合 $\{U_\alpha\}$ 的并集包含集合 S, 则称 $\{U_\alpha\}$ 是 S 的一个覆盖; 每一个 U_α 都是开集的覆盖称为开覆盖.

定理 7.2.4 (有限覆盖定理)**.** *设 $\{U_\alpha\}$ 为闭矩形 I 的开覆盖, 则可以从这一族开集中选取有限个 U_α, 使得它们的并集仍然包含 I.*

证明 设 λ 是前一引理中的正数(称为 Lebesgue 数). 取正整数 $m > \frac{d(I)}{\lambda}$. 将 I 等分为 m^n 个小矩形, 每一个的直径均为 $\frac{d(I)}{m}$, 它小于 λ. 于是由前一引理可知, 每一个小矩形均包含于某一个 U_α 中, 因此 I 包含于不超过 m^n 个 U_α 的并集中.

设 $S \subset R^n$ 为非空子集, 记 $d(S) = \sup\{d(x, y) \mid x, y \in S\}$, 称为 S 的直径. 当 $d(S) < +\infty$ 时称 S 为有界子集.

定理 7.2.5 (闭集套原理)**.** *设 $\{C_i\}$ 是 $\mathbb{R}^n$ 中一列非空有界闭集, 如果 $C_i \supset C_{i+1}$ 对每一个 $i \ge 1$ 都成立, 则 $\{C_i\}$ 必有公共点. 进一步, 如果 $\lim\limits_{i\to\infty} d(C_i) = 0$, 则 $\{C_i\}$ 有且仅有一个公共点.*

证明 先来说明 $\{C_i\}$ 必有公共点. (反证法) 若不然, 记 $U_i = \mathbb{R}^n \setminus C_i$, 则开集族 $\{U_i\}$ 的并集是整个 $\mathbb{R}^n$. 取矩形 $I \supset C_1$, 由前一定理, 存在 m, 使得

$$I \subset \bigcup_{i=1}^{m} U_i = U_m,$$

这与 $C_m \subset C_1 \subset I$ 相矛盾.

如果 $\lim\limits_{i\to\infty} d(C_i) = 0$, 且 x, y 均为 $\{C_i\}$ 的公共点, 则 $d(x, y) \le d(C_i)$ 对每一 i 都成立, 令 $i \to \infty$ 可知 $d(x, y) = 0$, 即 $x = y$.

研究数列极限时可以考虑单调性. 在一般的欧氏空间中没有单调点列的概念, 不过 Cauchy 准则仍然成立.

Cauchy 列: 设 $\{u_i\}$ 为 $\mathbb{R}^n$ 中的点列. 如果任给 $\varepsilon > 0$, 均存在 N, 使得当 $i, j > N$ 时 $d(u_i, u_j) < \varepsilon$, 则称 $\{u_i\}$ 为 Cauchy 列或基本列.

例 7.2.3. *设 $A \in M_{n\times n}$, 记 $e_i(A) = \sum\limits_{k=0}^{i} \frac{A^k}{k!}$, 其中 $A^0 = I_n$, 则 $\{e_i(A)\}$ 为 Cauchy 列.*

证明 $M_{n\times n}$ 可视为 n^2 维的欧氏空间. 当 $j>i$ 时, 根据矩阵范数的性质可得

$$d\big(e_j(A),e_i(A)\big)=\|e_j(A)-e_i(A)\|\le\sum_{k=i+1}^{j}\frac{\|A\|^k}{k!}\le\frac{\|A\|^{i+1}}{(i+1)!}\mathrm{e}^{\|A\|},$$

再由 $\lim\limits_{i\to\infty}\frac{\|A\|^i}{i!}=0$ 即知 $\{e_i(A)\}$ 为 Cauchy 列.

定理 7.2.6 (Cauchy 准则)**.** *设 $\{u_i\}$ 为 $\mathbb{R}^n$ 中的点列. 则 $\{u_i\}$ 收敛$\Longleftrightarrow$ $\{u_i\}$ 为 Cauchy 列.*

证明 若 $\{u_i\}$ 收敛于 u, 则利用三角不等式 $d(u_i,u_j)\le d(u_i,u)+d(u_j,u)$ 易见 $\{u_i\}$ 为 Cauchy 列.

反之, 若 $\{u_i\}$ 为 Cauchy 列, 则与 Cauchy 数列类似, $\{u_i\}$ 必为有界点列, 从而存在收敛子列. 与数列的情形类似, 容易看出 $\{u_i\}$ 本身也是收敛的.

设 A 为 n 阶实方阵, 根据前例我们就知道 $\{e_i(A)\}$ 收敛, 其极限记为 $\exp(A)$ 或 e^A. 称 exp 为矩阵的指数映射, 关于它的进一步讨论可参见本节习题.

定理 7.2.7 (压缩映射原理)**.** *设 C 为 $\mathbb{R}^n$ 中的闭集, $f:C\to C$ 为映射. 如果存在常数 $\lambda\in[0,1)$, 使得*

$$d\big(f(x),f(y)\big)\le\lambda\,d(x,y),\ \ \forall\,x,y\in C,$$

则存在唯一的 $c\in C$, 使得 $f(c)=c$.

证明 任取 $c_0\in C$, 当 $i\ge1$ 时用 $c_i=f(c_{i-1})$ 递归地定义 C 中一列点 $\{c_i\}$, 我们来说明它是 Cauchy 列.

事实上, 当 $i\ge1$ 时, 根据题设可得

$$d(c_{i+1},c_i)=d\big(f(c_i),f(c_{i-1})\big)\le\lambda\,d\big(c_i,c_{i-1}\big),$$

由归纳法可以得出 $d(c_{i+1},c_i\big)\le\lambda^i d(c_1,c_0)$. 因此, 当 $j>i$ 时, 由三角不等式可得

$$d\big(c_j,c_i\big)\le\sum_{k=i+1}^{j}d(c_k,c_{k-1})\le\sum_{k=i+1}^{j}\lambda^{k-1}d(c_1,c_0)\le\frac{\lambda^i}{1-\lambda}d(c_1,c_0),$$

由 $\lambda\in[0,1)$ 可知 $\lim\limits_{i\to\infty}\lambda^i=0$, 这说明 $\{c_i\}$ 为 Cauchy 列, 其极限记为 c, 由 C 为闭集可知 $c\in C$. 由 $c_i=f(c_{i-1})$ 和三角不等式可得

$$d\big(f(c),c\big)\le d\big(f(c),f(c_{i-1})\big)+d(c_i,c)\le\lambda\,d(c,c_{i-1})+d(c_i,c),$$

令 $i\to\infty$ 可知 $d\big(f(c),c\big)=0$, 即 $f(c)=c$.

若另有 $c'\in C$, 使得 $f(c')=c'$, 则

$$d(c',c)=d\big(f(c'),f(c)\big)\le\lambda\,d(c',c),$$

由 $\lambda \in [0,1)$ 可知 $d(c',c)=0$, 即 $c'=c$.

满足上述定理要求的映射 f 称为压缩映射, c 称为 f 的不动点. 下一章我们将用压缩映射研究多元函数的反函数定理.

例 7.2.4. 设 $A \in M_{n\times n}$, 若 $\|A\|<1$, 则存在 $A_* \in M_{n\times n}$, 使得 $A_*^2 = I_n - A$.

证明 记 $\alpha = 1-\sqrt{1-\|A\|}$, 则 $\alpha \in [0,1)$. 记 $C=\{X \in M_{n\times n} \mid \|X\| \le \alpha\}$, 则 C 为 $M_{n\times n}$ 中的闭集. 当 $X \in C$ 时, 令 $f(X)=\frac{1}{2}(A+X^2)$. 由矩阵范数的性质可得

$$\|f(X)\| \le \frac{1}{2}\big(\|A\|+\|X\|^2\big) \le \frac{1}{2}\big(\|A\|+\alpha^2\big)=\alpha,$$

这说明 $f(X)\in C$. 当 $X,Y \in C$ 时, 有

$$\|f(X)-f(Y)\| = \frac{1}{2}\|(X-Y)X+Y(X-Y)\| \le \frac{1}{2}\|X-Y\|(\|X\|+\|Y\|) \le \alpha\|X-Y\|,$$

这说明 $f: C \to C$ 为压缩映射. f 的不动点记为 B, 则 $(I_n-B)^2 = I_n - A$.

习题 7.2

1. 证明: $\mathbb{R}^n$ 中的有限点集是闭集; 坐标均为整数的点构成的集合 $\mathbb{Z}^n$ 是闭集; 坐标均为有理数的点构成的集合 $\mathbb{Q}^n$ 既不是开集, 也不是闭集.

2. 证明: 一族开集的并集仍为开集, 有限个开集的交集仍为开集; 一族闭集的交集仍为闭集, 有限个闭集的并集仍为闭集.

3. 证明: 有限覆盖定理中的矩形 I 可以换成 $\mathbb{R}^n$ 的有界闭集, 结论仍然成立. (提示: 设 C 为有界闭集, 取矩形 $I \supset C$, 则 $\mathbb{R}^n \setminus C$ 和给定的 U_α 一起成为 I 的开覆盖.)

4. 设 $A \in M_{n\times n}$, P 是 n 阶可逆矩阵, 证明: $P\exp(A)P^{-1} = \exp\big(PAP^{-1}\big)$.

5. 设 $A \in M_{n\times n}$, 证明: $\det\exp(A) = \exp(\mathrm{tr}A)$.

6. 设 $A \in M_{n\times n}$, 当 $t \in \mathbb{R}$ 时记 $\tau(t)=\exp(tA)$, 证明: $\tau'(t) = A\tau(t) = \tau(t)A$.

7. 本题研究如下常系数线性常微分方程

$$x^{(n)}(t)+a_{n-1}x^{(n-1)}(t)+\cdots+a_1x'(t)+a_0x(t)=0$$

在初始条件 $x(t_0)=c_0$, $x'(t_0)=c_1$, $\cdots$, $x^{(n-1)}(t_0)=c_{n-1}$ 下解的存在性.

(1) 记 $X(t)=\big(x(t),x'(t),\cdots,x^{(n-1)}(t)\big)^\top$, 证明: 上述方程可以转化为 $X'(t) = AX(t)$, 其中 $A \in M_{n\times n}$.

(2) 证明: $X(t)=\exp\big[(t-t_0)A\big](c_0,c_1,\cdots,c_{n-1})^\top$ 满足 $X'(t)=AX(t)$.

8. 设 $A \in M_{n\times n}$ 且 $\|A\| < 1$. 递归地定义一列方阵 $\{B_i\}$: $B_0 = I_n$, $i \geq 1$ 时 $B_i = I_n + AB_{i-1}$. 证明 $\{B_i\}$ 收敛, 且其极限 B 满足 $(I_n - A)B = B(I_n - A) = I_n$.

9. 设 C 为 $\mathbb{R}^n$ 中的闭集, $f : C \to C$ 为映射. 如果存在正整数 m, 使得 f 和自身复合 m 次以后是压缩映射, 则 f 有唯一的不动点.

7.3 连续映射

我们想将一元连续函数的性质推广到多元函数. 设 $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}^m$ 为映射. $n > 1$ 时 f 常称为多元向量值函数, $m = 1$ 时简称多元函数. 设 $u_0 \in S$, 如果任给 $\varepsilon > 0$, 均存在 $\delta > 0$, 使得

$$f\big(B_\delta(u_0) \cap S\big) \subset B_\varepsilon\big(f(u_0)\big), \tag{7.4}$$

则称 f 在 u_0 处连续. 如果 f 处处连续, 则称 f 为 S 上的连续映射.

命题 7.3.1. *映射 f 在 u_0 处连续$\Longleftrightarrow$任给 S 中收敛于 u_0 的点列 $\{u_i\}$, 均有 $\lim\limits_{i\to\infty} f(u_i) = f(u_0)$.*

证明 "$\Longrightarrow$" 若 f 在 u_0 处连续, 则任给 $\varepsilon > 0$, 存在 $\delta > 0$ 使得 (7.4) 成立. 当 S 中的点列 $\{u_i\}$ 收敛于 u_0 时, 存在 N, 使得当 $i > N$ 时 $u_i \in B_\delta(u_0)$. 此时 $f(u_i) \in B_\varepsilon\big(f(u_0)\big)$. 这说明 $\{f(u_i)\}$ 收敛于 $f(u_0)$.

"$\Longleftarrow$" (反证法) 若 f 在 u_0 处不连续, 则存在 $\varepsilon_0 > 0$ 以及点列 $\{u_i\} \subset S$, 使得

$$d(u_i, u_0) < \frac{1}{i}, \ \ d\big(f(u_i), f(u_0)\big) \geq \varepsilon_0,$$

此时 $\{u_i\}$ 收敛于 u_0, 但 $\{f(u_i)\}$ 不收敛于 $f(u_0)$, 这就导出了矛盾.

下面命题的证明与一元函数类似, 我们略去证明.

命题 7.3.2 (连续函数的四则运算性质)**.** *设 $f, g : S \to \mathbb{R}$ 为连续函数, 则*

(1) 当 $\lambda, \mu \in \mathbb{R}$ 时, $\lambda f + \mu g$ 也是连续函数;

(2) fg 为连续函数;

(3) 当 $g \neq 0$ 时, $\frac{f}{g}$ 为连续函数.

为了更好地刻画连续映射, 我们引进相对开集和相对闭集的概念. 设 $u \in S$, $r > 0$, 记 $B_r^S(u) = B_r(u) \cap S$, 称为 S 中的相对开球. 设 $U \subset S$, 如果任给 $u \in U$, 均存在 $r > 0$, 使得 $B_r^S(u) \subset S$, 则称 U 为 S 中的相对开集. 设 $C \subset S$, 若 $S \setminus C$ 为 S 中的相对开集, 则称 C 为 S 中的相对闭集.

命题 7.3.3 (连续映射的刻画)**.** *映射 $f : S \to \mathbb{R}^m$ 连续 $\Longleftrightarrow$ 开集的原像为 S 中的相对开集 $\Longleftrightarrow$ 闭集的原像为 S 中的相对闭集.*

证明 以开集为例. "$\Longrightarrow$" 设 f 连续, V 为 $\mathbb{R}^m$ 中的开集. 任取 $u_0 \in f^{-1}(V)$, 则 $f(u_0) \in V$, 由于 V 为开集, 故存在 $\varepsilon > 0$, 使得 $B_\varepsilon\big(f(u_0)\big) \subset V$. 由 f 在 u_0 处连续可知, 存在 $\delta > 0$, 使得 (7.4) 成立. 这说明 $B_\delta^S(u_0) \subset f^{-1}(V)$, 因此 $f^{-1}(V)$ 为 S 中的相对开集.

"$\Longleftarrow$" 任取 $u_0 \in S$ 以及 $\varepsilon > 0$. 于是 $U = f^{-1}\big(B_\varepsilon\big(f(u_0)\big)\big)$ 为 S 中包含 u_0 的相对开集, 从而存在 $\delta > 0$, 使得 $B_\delta^S(u_0) \subset U$. 即 $f\big(B_\delta^S(u_0)\big) \subset B_\varepsilon\big(f(u_0)\big)$, 于是 f 在 u_0 处连续.

利用连续映射的刻画可以得出连续映射的复合仍为连续映射, 我们略去证明.

命题 7.3.4 (连续映射的复合性质)**.** *设 $f : S \to \mathbb{R}^m$, $g : T \to \mathbb{R}^k$ 均为连续映射, 其中 $f(S) \subset T \subset \mathbb{R}^m$, 则 $g \circ f : S \to \mathbb{R}^k$ 也是连续映射.*

例 7.3.1. *行列式函数.*

$M_{n\times n}$ 可视为 n^2 维的欧氏空间, 此时矩阵的行列式 $\det : M_{n\times n} \to \mathbb{R}$ 是多元函数, 它是具有 n^2 个变量的多元多项式, 因此根据连续函数的四则运算性质可知 det 是连续函数.

记 $GL(n,\mathbb{R}) = \{A \in M_{n\times n} \mid \det A \neq 0\}$, 则 $GL(n,\mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\})$. 因为 $\mathbb{R} \setminus \{0\}$ 为 $\mathbb{R}$ 中的开集, 由 det 连续可知 $GL(n,\mathbb{R})$ 是 $M_{n\times n}$ 中的开集. 直观地看, 若 A 是非退化方阵, 则 A 附近的方阵也是非退化的.

下面我们来研究连续映射的整体性质, 先看介值性, 它与所谓的道路连通性密切相关.

设 $S \subset \mathbb{R}^n$. 如果任给 $u_0, u_1 \in S$, 均存在连续映射 $\sigma : [0,1] \to \mathbb{R}^n$, 使得 $\sigma(0) = u_0$, $\sigma(1) = u_1$, 且 $\sigma\big([0,1]\big) \subset S$, 则称 S 是道路连通子集. 此时 σ 称为 S 中连接 u_0 和 u_1 的一条道路.

例 7.3.2. *设 $n \geq 2$, 则 $\mathbb{R}^n \setminus \{0\}$ 是 $\mathbb{R}^n$ 中的道路连通子集.*

证明 设 u_0, u_1 都不是原点. 若连接 u_0 和 u_1 的直线不经过原点, 则令 $\sigma(t) = (1-t)u_0 + tu_1$, 此时 σ 是连接 u_0 和 u_1 的一条道路, 且不经过原点.

若连接 u_0 和 u_1 的直线经过原点, 由于 $n \geq 2$, 我们就可以在该直线之外任取一点 u. 令

$$\sigma(t) = \begin{cases} (1-2t)u_0 + 2tu, & 0 \leq t \leq \frac{1}{2}, \\ 2(1-t)u + (2t-1)u_1, & \frac{1}{2} < t \leq 1, \end{cases}$$

则 σ 是连接 u_0 和 u_1 的一条道路, 且不经过原点.

定理 7.3.5 (介值定理)**.** *设 S 为 $\mathbb{R}^n$ 中道路连通子集, $f : S \to \mathbb{R}^m$ 为连续映射, 则 $f(S)$ 为 $\mathbb{R}^m$ 中道路连通子集.*

证明 任取 $v_0, v_1 \in f(S)$, 则存在 $u_0, u_1 \in S$, 使得 $f(u_0) = v_0$, $f(u_1) = v_1$. 设 $\sigma : [0,1] \to \mathbb{R}^n$ 是 S 中连接 u_0, u_1 的道路, 记 $\tau(t) = f\big(\sigma(t)\big)$, 则 τ 就是 $f(S)$ 中连接 v_0, v_1 的道路, 这说明 $f(S)$ 是道路连通子集.

定理 7.3.6 (零值定理)**.** *设 S 为 $\mathbb{R}^n$ 中道路连通子集, $f : S \to \mathbb{R}$ 为连续函数. 若 $f(u_0) < 0$, $f(u_1) > 0$, 则存在 $\zeta \in S$, 使得 $f(\zeta) = 0$.*

证明 设 $\sigma : [0,1] \to \mathbb{R}^n$ 是 S 中连接 u_0, u_1 的道路, 记 $\tau(t) = f\big(\sigma(t)\big)$, 则 $\tau : [0,1] \to \mathbb{R}$ 为连续函数, 且 $\tau(0) < 0$, $\tau(1) > 0$. 根据一元连续函数的零值定理, 存在 $\xi \in (0,1)$, 使得 $\tau(\xi) = 0$. 记 $\zeta = \sigma(\xi)$, 则 $f(\zeta) = 0$.

例 7.3.3. *设 $f : S^1 \to \mathbb{R}$ 为定义在单位圆周上的连续函数, 则存在 $u \in S^1$, 使得 $f(-u) = f(u)$.*

证明 显然, 单位圆周是道路连通的. 考虑函数 $g(x) = f(x) - f(-x)$, 则 g 也是 S^1 上的连续函数, 且

$$g(x)g(-x) = [f(x) - f(-x)][f(-x) - f(x)] = -[f(x) - f(-x)]^2 \le 0,$$

根据零值定理, 存在 $u \in S^1$, 使得 $g(u) = 0$, 此时 $f(-u) = f(u)$.

例 7.3.4. *记 $GL^+(n,\mathbb{R}) = \{A \in M_{n\times n} \mid \det A > 0\}$, 则 $GL^+(n,\mathbb{R})$ 是 $M_{n\times n}$ 中的道路连通子集.*

证明 用归纳法. 当 $n = 1$ 时结论显然成立. 设 $GL^+(n-1,\mathbb{R})$ 道路连通. 任给 $A \in GL^+(n,\mathbb{R})$, 我们要在 $GL^+(n,\mathbb{R})$ 中找到连接 A 和单位矩阵 I_n 的道路. 分情况讨论:

(1) A 没有负特征值. 记 $\sigma(t) = tI_n + (1-t)A$, 则当 $t \in [0,1]$ 时 $\sigma(t) \in GL(n,\mathbb{R})$. 对 $\det\sigma(t)$ 应用零值定理还可以知道 $\sigma(t) \in GL^+(n,\mathbb{R})$, 因此 σ 是 $GL^+(n,\mathbb{R})$ 中连接 A 和 I_n 的一条道路.

(2) A 有负特征值, 比如 λ. 设 v_1 是对应的一个特征向量, 将 v_1 扩充为 $\mathbb{R}^n$ 的一组基 $\{v_1, v_2, \cdots, v_n\}$, 以这一组基为列向量的 n 阶方阵记为 P, 则 $A = P\mathrm{diag}\big(\lambda, B\big)P^{-1}$, 其中 B 是 $n-1$ 阶方阵, 其行列式小于零. 根据归纳假设, 存在道路 $\tau : [0,1] \to GL^+(n-1,\mathbb{R})$, 使得 $\tau(0) = \mathrm{diag}\big(-1, 1, \cdots, 1\big)B$, $\tau(1) = I_{n-1}$. 当 $t \in [0, 1/2]$ 时, 令

$$\sigma(t) = P\mathrm{diag}\Big(-2t + (1-2t)\lambda,\ \mathrm{diag}\big(-1, 1, \cdots, 1\big)\tau(2t)\Big)P^{-1};$$

当 $t \in [1/2, 1]$ 时, 令

$$\sigma(t) = P\mathrm{diag}\left(\begin{pmatrix}\cos(2\pi t) & -\sin(2\pi t)\\ \sin(2\pi t) & \cos(2\pi t)\end{pmatrix},\ I_{n-2}\right)P^{-1}.$$

则 σ 就是 $GL^+(n,\mathbb{R})$ 中连接 A 和 I_n 的一条道路.

我们再看连续映射的一致连续性质和最值性质, 它们都与有界闭集密切相关. 设 $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}^m$ 为映射. 如果任给 $\varepsilon > 0$, 均存在 $\delta > 0$, 使得对每一个 $u \in S$, 均有 $f\big(B^S_\delta(u)\big) \subset B_\varepsilon\big(f(u)\big)$, 则称 f 一致连续. 显然, 一致连续意味着连续.

定理 7.3.7 (一致连续性)**.** *设 $f : S \to \mathbb{R}^m$ 为连续映射. 若 S 为有界闭集, 则 f 一致连续.*

证明 (反证法) 若不然, 则存在 $\varepsilon_0 > 0$, 使得对每一个正整数 i, 均存在 $u_i, v_i \in S$, 满足

$$d(u_i, v_i) < \frac{1}{i},\ \ d\big(f(u_i), f(v_i)\big) \geq \varepsilon_0.$$

因为 S 为有界闭集, $\{u_i\}$ 就存在收敛子列. 不妨设它本身收敛于 u_0, 则 $u_0 \in S$. 由三角不等式可知 $\{v_i\}$ 也收敛于 u_0, 且

$$\varepsilon_0 \leq d\big(f(u_i), f(u_0)\big) + d\big(f(u_0), f(v_i)\big).$$

令 $i \to \infty$, 由 f 连续可知上式右边趋于零, 这就导出了矛盾.

定理 7.3.8 (最值定理)**.** *设 $f : S \to \mathbb{R}^m$ 为连续映射. 若 S 为有界闭集, 则 $f(S)$ 为 $\mathbb{R}^m$ 中的有界闭集. 特别地, 当 $m = 1$ 时, f 达到最小值和最大值.*

证明 先证明 $f(S)$ 有界. (反证法) 若不然, 则对每一个正整数 i, 均存在 $u_i \in S$, 使得 $\|f(u_i)\| > i$. 因为 S 为有界闭集, $\{u_i\}$ 就存在收敛子列. 不妨设它本身收敛于 u_0, 则 $u_0 \in S$. 由 f 连续可知 $\{f(u_i)\}$ 收敛于 $f(u_0)$, 则 $\{f(u_i)\}$ 就是有界点列, 这就导出了矛盾.

我们用命题 7.2.1 来说明 $f(S)$ 是闭集. 设 $\{f(u_i)\}$ 收敛于 v_0, 我们要说明 $v_0 \in f(S)$. 类似于前一段的论证, 不妨设 $\{u_i\}$ 收敛于 $u_0 \in S$. 由 f 连续可知 $\{f(u_i)\}$ 收敛于 $f(u_0)$. 由极限的唯一性即知 $v_0 = f(u_0) \in f(S)$.

当 $m = 1$ 时, $f(S)$ 是 $\mathbb{R}^1$ 中的有界闭集, 其下确界和上确界分别记为 α, β. 因为 α 是最大下界, 对每一个正整数 i, 就存在 $\alpha_i \in f(S)$, 使得

$$\alpha \leq \alpha_i < \alpha + \frac{1}{i}.$$

这说明 $\{\alpha_i\}$ 是 $f(S)$ 中收敛点列, 由 $f(S)$ 为闭集即知其极限 $\alpha \in f(S)$. 同理, $\beta \in f(S)$. 这说明 f 达到了最小值 α 和最大值 β.

下面我们介绍最值定理的简单应用. 设 $A = \big(a_{ij}\big)_{n\times n}$ 为 n 阶对称实方阵, 考虑函数 $Q : \mathbb{R}^n \to \mathbb{R}$,

$$Q(x) = \langle x, Ax\rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j.$$

若 $Q(x) \geq 0$ 恒成立, 则称 A 半正定; 若 A 半正定, 且 $Q(x) = 0$ 仅在 $x = 0$ 处成立, 则称 A 正定.

例 7.3.5. 设 A 为 n 阶正定矩阵, 则存在正数 λ, μ, 使得

$$\lambda\|x\|^2 \le Q(x) \le \mu\|x\|^2, \ \forall\, x \in \mathbb{R}^n.$$

证明 记 $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. 由题设可知 $Q(x)$ 在 S^{n-1} 上恒为正. 注意到 $Q(x)$ 是连续函数, S^{n-1} 是有界闭集, 由最值定理, $Q(x)$ 在 S^{n-1} 上达到(正的)最小值 λ 和最大值 μ. 当 $x \neq 0$ 时, 记 $x' = \frac{x}{\|x\|}$, 则 $Q(x) = \|x\|^2 Q(x')$, 于是

$$\lambda \le \frac{Q(x)}{\|x\|^2} = Q(x') \le \mu.$$

这就得到了欲证结论.

事实上, λ, μ 分别是 A 的最小特征值和最大特征值.

引理 7.3.9. $Q(x)$ 在 S^{n-1} 上的最小值必为矩阵 A 的特征值.

证明 根据最值定理, 存在 $u \in S^{n-1}$, 使得 $Q(u) = \lambda$. 此时 $Q(x) \ge \lambda\|x\|^2$, $\forall\, x \in \mathbb{R}^n$. 特别地, 对任意 $y \in \mathbb{R}^n$, $t \in \mathbb{R}$, 有

$$\varphi(t) = Q(u + ty) - \lambda\|u + ty\|^2 \ge 0.$$

$\varphi(t)$ 关于 t 为光滑函数, $t = 0$ 时取到最小值 0, 故 $\varphi'(0) = 0$. 即

$$\langle y, Au\rangle + \langle u, Ay\rangle - \lambda\big(\langle u, y\rangle + \langle y, u\rangle\big) = 0.$$

由 A 为对称方阵可得 $\langle y,\, Au - \lambda u\rangle = 0$. 取 $y = Au - \lambda u$, 由内积的正定性可得 $Au - \lambda u = 0$. 这说明 λ 为 A 的特征值.

这个引理的证明可以推广如下: 设 $V \subset \mathbb{R}^n$ 为子向量空间, 如果 $AV \subset V$, 则称 V 为 A 的一个不变子空间. 令

$$\lambda' = \inf\{Q(x) \,|\, x \in V,\ \|x\| = 1\},$$

则完全类似的证明可以推出 λ' 为 A 的特征值.

如果 V 为不变子空间, 则其正交补

$$V^{\perp} = \{y \in \mathbb{R}^n \,|\, \langle y, x\rangle = 0,\ \forall\, x \in V\}$$

也是 A 的不变子空间; 而如果 λ' 为特征值, 则特征子空间

$$V(\lambda') = \{x \in \mathbb{R}^n \,|\, Ax = \lambda' x\}$$

是不变子空间. 因此, 如果重复上面的证明过程, 我们就可以得到 A 的所有特征值

$$\lambda = \lambda_0 < \lambda_1 < \cdots < \lambda_r, \quad r \le n.$$

其中

$$\lambda_i = \min\{Q(x) \,\big|\, x \perp V(\lambda_0) \oplus V(\lambda_1) \oplus \cdots \oplus V(\lambda_{i-1}),\ \|x\| = 1\}, \quad i \ge 1.$$

从而有

推论 7.3.10. 设 A 为 n 阶实对称方阵, 则
(1) A 的特征值全为实数;
(2) A 为正定矩阵 $\Longleftrightarrow$ A 的特征值都是正实数.

证明. (1) 以及 (2) 的 "$\Longrightarrow$" 部分已证. (2) 的 "$\Longleftarrow$"部分可由引理 7.3.9 得出.

引理 7.3.11. 设 A 为 n 阶实对称方阵, 其正特征值个数 (含重数) 为 k, 如果 V 为 $\mathbb{R}^n$ 的子向量空间, 且对任意 $x \in V \setminus \{0\}$, 均有 $Q(x) > 0$, 则 $\dim V \le k$.

证明. (反证法) 设 $\dim V > k$, 记 A 的正特征值对应的特征子空间的直和为 V_+, 则 $\dim V_+ = k$. 此时

$$\dim V + \dim V_+^{\perp} = \dim V + (n-k) > n,$$

于是 $V \cap V_+^{\perp} \neq \{0\}$. 设 x 是 $V \cap V_+^{\perp}$ 中的非零向量, 由 $x \in V_+^{\perp}$ 可知 $Q(x) \le 0$, 这与 $x \in V$ 相矛盾.

现在我们就得到了矩阵正定性的如下判别法:

定理 7.3.12. $A = \big(a_{ij}\big)_{n\times n}$ 为正定方阵 $\Longleftrightarrow \det\big(a_{ij}\big)_{\substack{1\le i\le k\\1\le j\le k}} > 0$, $\forall\, 1 \le k \le n$.

证明. "$\Longrightarrow$" 如果 A 正定, 则显然 $\big(a_{ij}\big)_{\substack{1\le i\le k\\1\le j\le k}}$ 正定, 其特征值全为正实数, 故行列式为正 (行列式为特征值之积).

"$\Longleftarrow$" 对 n 用归纳法. $n=1$ 显然. 设命题对 $n-1$ 成立, 则对于 n 阶方阵, 由归纳假设, $\big(a_{ij}\big)_{\substack{1\le i\le n-1\\1\le j\le n-1}}$ 正定, 这说明 A 在子向量空间 $\mathbb{R}^{n-1} = \{x \in \mathbb{R}^n \,|\, x_n = 0\}$ 上正定. 由引理 7.3.11, A 至少有 $n-1$ 个正的特征值. 又因为 $\det A > 0$, 故所有特征值均为正数, 由推论 7.3.10 即知 A 是正定的.

注. 用微分学的方法判断多元函数极值时, 需要用到正定矩阵的部分知识.

习题 7.3

1. 证明: 多元向量值函数为连续映射当且仅当它的每一个分量均为连续函数.

2. 设 $S \subset \mathbb{R}^n$. 证明: 若 U 为 S 中的相对开集, 则存在 $\mathbb{R}^n$ 中的开集 W, 使得 $U = W \cap S$; 若 C 为 S 中的相对闭集,则存在 $\mathbb{R}^n$ 中的闭集 D, 使得 $C = D \cap S$.

3. 记 $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. 证明: 当 $n \ge 2$ 时 S^{n-1} 是道路连通子集.

4. 记 $SL(n,\mathbb{R}) = \{A \in M_{n\times n} \mid \det A = 1\}$. 证明: $SL(n,\mathbb{R})$ 是 $M_{n\times n}$ 中的闭集.

5. 记 $O(n) = \{A \in M_{n\times n} \mid AA^{\top} = I_n\}$. 证明: $O(n)$ 是 $M_{n\times n}$ 中的有界闭集.

6. 设 $f:\mathbb{R}^2\to\mathbb{R}$ 为连续函数, 证明: 存在 $u_0\neq u_1$, 使得 $f(u_0)=f(u_1)$.

7. 设 U 为 $\mathbb{R}^n$ 中道路连通子集, $u_0,u_1\in U$. 证明: 当 $u_0\neq u_1$ 时, 存在 U 中连接 u_0 和 u_1 的折线段.

8. 设 A 是 n 阶正定矩阵, $u\in\mathbb{R}^n$, $c\in\mathbb{R}$. 记
$$f(x)=\langle x,Ax\rangle+\langle u,x\rangle+c,\ \ \forall\, x\in\mathbb{R}^n.$$
证明: $\lim\limits_{x\to\infty}f(x)=+\infty$, 并求 f 的最小值.

9. 记 $SO(n)=\{A\in O(n)\mid \det A=1\}$. 证明: 当 $n\geq 2$ 时 $SO(n)$ 是 $M_{n\times n}$ 中的道路连通子集.

7.4 Lipschitz 映射和零测集

我们知道, 定义在闭区间中的一元有界函数可积当且仅当其间断点集为零测集. 在研究多元函数积分时, 我们将首先研究 n 维闭矩形中定义的有界函数的积分, 同样地要给出可积函数间断点集的刻画.

我们先看一类非常特殊的函数. 设 $S\subset\mathbb{R}^n$, 定义函数 $\chi_S:\mathbb{R}^n\to\mathbb{R}$ 如下:
$$\chi_S(x)=\begin{cases}1, & x\in S,\\ 0, & x\in\mathbb{R}^n\setminus S.\end{cases}$$
χ_S 称为子集 S 的特征函数. 不难看出, χ_S 在 u 处连续时, 要么存在 $r>0$, 使得 $B_r(u)\subset S$; 要么存在 $r>0$, 使得 $B_r(u)\subset\mathbb{R}^n\setminus S$. 为了方便起见, 我们再引进几个与子集有关的概念.

内点: 设 $u\in S$. 如果存在 $r>0$, 使得 $B_r(u)\subset S$, 则称 u 为 S 的内点. 内点的全体记为 $\mathrm{int}S$, 称为 S 的内部.

外点: 设 $u\notin S$. 如果存在 $r>0$, 使得 $B_r(u)\subset\mathbb{R}^n\setminus S$, 则称 u 为 S 的外点. 外点的全体记为 $\mathrm{ext}S$, 称为 S 的外部.

边界点: 设 $u\in\mathbb{R}^n$. 如果任给 $r>0$, $B_r(u)$ 中既有属于 S 中的点, 也有不属于 S 中的点, 则称 u 为 S 的边界点. 边界点的全体记为 ∂S, 称为 S 的边界. ∂S 就是特征函数 χ_S 的间断点(不连续点)集.

从定义不难看出, $\mathrm{int}S$ 是包含于 S 的“最大”开集, S 的外点就是 $\mathbb{R}^n\setminus S$ 的内点. 整个空间 $\mathbb{R}^n$ 可以分解为
$$\mathbb{R}^n=\mathrm{int}S\cup\partial S\cup\mathrm{ext}S,$$
上式右边的三个子集互不相交. 我们再列举几条今后会用到的性质:

- ∂S 为闭集, 这是因为它的补集等于 $\mathrm{int}S\cup\mathrm{ext}S$.

- $\text{int}S \cup \partial S$ 也是闭集, 记为 $\bar{S}$, 称为 S 的闭包. 此外, $\bar{S} = S \cup \partial S$ 也成立. 这是因为
$$\bar{S} \subset S \cup \partial S \subset (\text{int}S \cup \partial S) \cup \partial S = \bar{S}.$$

- S 为闭集当且仅当 $\partial S \subset S$, 即 $S = \bar{S}$. 事实上, 若 S 是闭集, 则 $\mathbb{R}^n \setminus S$ 为开集, 它里面的点不可能是边界点, 即 $\partial S \subset S$. 反之, 若 $\partial S \subset S$, 则 $S = \bar{S}$ 是闭集.

- 当 $S \subset T$ 时, $\bar{S} \subset \bar{T}$. 这是因为, 此时 $\mathbb{R}^n \setminus S \supset \mathbb{R}^n \setminus T$, 从而 $\text{ext}S \supset \text{ext}T$, 即 $\bar{S} \subset \bar{T}$ 成立.

例 7.4.1. *球的边界.*

设 $u \in \mathbb{R}^n$, $r > 0$. 由于 $B_r(u)$ 和 $\mathbb{R}^n \setminus \bar{B}_r(u)$ 均为开集, 故
$$\partial B_r(u) = \left\{x \in \mathbb{R}^n \,\middle|\, |x - u| = r\right\},$$
称为以 u 为中心, r 为半径的 $n-1$ 维球面.

下面我们将 $\mathbb{R}$ 中零测集的概念推广到 $\mathbb{R}^n$ 中.

零测集: 设 $S \subset \mathbb{R}^n$. 如果任给 $\varepsilon > 0$, 存在至多可数个矩形 $\{I_i\}$, 使得
$$C \subset \bigcup_{i \geq 1} I_i, \ \ \sum_{i \geq 1} \nu(I_i) < \varepsilon,$$
则称 C 为 $\mathbb{R}^n$ 中的零测集.

和一维的情形类似, 不难看出:有限点集均为零测集; 零测集的子集仍为零测集; 可数个零测集之并仍为零测集.

设 $u = (u_1, u_2, \cdots, u_n) \in \mathbb{R}^n$, $r > 0$, 记 $I_r(u) = \prod\limits_{i=1}^{n} [u_i - r, u_i + r]$, 称为以 u 为中心, $2r$ 为边长的 n 维(闭)方体. 在零测集中定义中, “矩形”也可以换成“方体”, 这可以从如下简单引理看出来.

引理 7.4.1. *设 I 为 n 维矩形, 则 I 可被容积之和不超过 $2^{n-1}\nu(I)$ 的有限个方体所覆盖.*

证明 不妨设 $n \geq 2$. 基本的想法是将 I 的各边延长为最短边长度的整数倍, 然后将其分割为若干个 n 维方体. 设 $I = \prod\limits_{i=1}^{n} [a_i, b_i]$, 记 $\ell = \min\{b_i - a_i \mid 1 \leq i \leq n\}$, 不妨设 $b_1 - a_1 = \ell$. 记 $k_1 = 0$, $c_1 = b_1$. 当 $i > 1$ 时, 记
$$k_i = \left\lfloor \frac{b_i - a_i}{\ell} \right\rfloor, \ \ c_i = a_i + (k_i + 1)\ell.$$
令 $J = \prod\limits_{i=1}^{n} [a_i, c_i]$, 则 $I \subset J$. 显然, J 可以分割为 $\prod\limits_{i=1}^{n} (k_i + 1)$ 个边长为 ℓ 的 n 维方体, 其容积之和满足
$$\left[\prod_{i=1}^{n} (k_i + 1)\right] \ell^n \leq \left[\prod_{i=2}^{n} (2k_i)\right] \ell^n \leq 2^{n-1}\nu(I).$$

例 7.4.2. 设 I 为 n 维矩形, $f: I \to \mathbb{R}^m$ 为连续映射. 记

$$G_f = \{(x, f(x)) \in \mathbb{R}^{n+m} \mid x \in I\},$$

则 G_f 是 $\mathbb{R}^{n+m}$ 中的零测集.

证明 先看一个最简单的情形: $f \equiv c$, 其中 $c \in \mathbb{R}^m$. 任给 $\varepsilon > 0$, 设 δ 为待定正数. 设 $I_\delta(c)$ 是以 c 为中心的 m 维方体, 记 $J = I \times I_\delta(c)$, 则 J 为 $n+m$ 维矩形. 显然, $G_f \subset J$, 且当 δ 充分小时, J 的容积满足

$$\nu(J) = \nu(I)\nu(I_\delta(c)) = \nu(I)(2\delta)^m < \varepsilon.$$

按照定义, G_f 为 $\mathbb{R}^{n+m}$ 中的零测集.

我们再看一般情形. 任给 $\varepsilon > 0$, 设 δ 为待定正数. 根据定理 7.3.7, 我们可以将 I 等分为有限个小矩形 $\{I_i\}$, 使得每一个 $f(I_i)$ 的直径均小于 δ. 取 $c_i \in f(I_i)$, 则 $f(I_i) \subset I_\delta(c_i)$. 于是 G_f 包含于有限个 $n+m$ 维矩形 $\{I_i \times I_\delta(c_i)\}$ 的并集之中. 当 δ 充分小时, 它们的容积之和满足

$$\sum_i \nu(I_i)\nu(I_\delta(c_i)) = \sum_i \nu(I_i)(2\delta)^m = \nu(I)(2\delta)^m < \varepsilon.$$

按照定义, G_f 为 $\mathbb{R}^{n+m}$ 中的零测集.

下面我们再简单介绍一下零测集的不变性质, 这将用于多元积分的变量替换.

Lipschitz 映射: 设 $S \subset \mathbb{R}^n$, $f: S \to \mathbb{R}^m$ 为映射. 如果存在非负常数 L, 使得

$$d(f(x), f(y)) \le L\, d(x, y), \ \ \forall\, x, y \in S, \tag{7.5}$$

则称 f 为 Lipschitz 映射. 显然, Lipschitz 映射是一致连续的.

例 7.4.3. 设 $S \subset \mathbb{R}^n$ 为非空子集, 当 $x \in \mathbb{R}^n$ 时, 定义

$$d(x, S) = \inf\{d(x, u) \mid u \in S\},$$

则 $d(x, S)$ 是 *Lipschitz* 函数, 且 $d(x, S) = 0$ 当且仅当 $x \in \bar{S}$.

证明 任取 $x, y \in \mathbb{R}^n$, 由三角不等式可知

$$d(x, u) \le d(x, y) + d(y, u), \ \ \forall\, u \in S.$$

在上式中关于 $u \in S$ 取下确界可得 $d(x, S) \le d(x, y) + d(y, S)$. 交换 x, y 的位置时此不等式当然也成立, 这说明 $|d(x, S) - d(y, S)| \le d(x, y)$.

显然, $d(x, S) \ge 0$. 若 $d(x, S) = 0$, 则存在 S 中一列点 $\{u_i\}$, 使得 $\lim\limits_{i\to\infty} d(x, u_i) = 0$, 此时 $x \in \bar{S}$. 反之, 若 $x \in \bar{S}$, 则存在 S 中一列点 $\{u_i\}$, 使得 $\lim\limits_{i\to\infty} d(x, u_i) = 0$, 此时 $d(x, S) = 0$.

$d(x, S)$ 称为到 S 的距离函数, 它的一个用处体现在下例中.

例 7.4.4. 设 C, D 是 $\mathbb{R}^n$ 中两个不相交的非空闭集, 则存在连续函数 $\rho : \mathbb{R}^n \to \mathbb{R}$, 使得

$$0 \le \rho \le 1, \ \ \rho\big|_C \equiv 1, \ \rho\big|_D \equiv 0.$$

证明 根据前例, 当 $x \in \mathbb{R}^n$ 时均有 $d(x,C) + d(x,D) > 0$. 令

$$\rho(x) = \frac{d(x,D)}{d(x,C) + d(x,D)}, \ \ \forall\, x \in \mathbb{R}^n,$$

则 ρ 是满足要求的连续函数.

受距离函数定义的启发, 我们还有如下结果.

引理 7.4.2. *设 S 为 $\mathbb{R}^n$ 中的子集, $f : S \to \mathbb{R}$ 满足条件 $|f(x) - f(y)| \le L\, d(x,y)$, 则存在函数 $F : \mathbb{R}^n \to \mathbb{R}$, 满足条件 $F|_S = f$ 且 $|F(x) - F(y)| \le L\, d(x,y)$.*

证明 若 F 是预想中的函数, 则当 $x \in \mathbb{R}^n$ 时, 应该有

$$F(x) \le F(u) + L\, d(x,u) = f(u) + L\, d(x,u), \ \ \forall\, u \in S.$$

受此启发, 我们干脆就定义 $F(x) = \inf\{f(u) + L\, d(x,u) \mid u \in S\}$. 我们首先说明 $F(x)$ 有意义: 固定一个 $v \in S$, 则当 $x \in \mathbb{R}^n$ 时

$$f(u) + L\, d(x,u) \ge f(v) - L\, d(u,v) + L\, d(x,u) \ge f(v) - L\, d(x,v),$$

这说明 $\{f(u) + L\, d(x,u) \mid u \in S\}$ 有下界, 从而有下确界.

其次, 我们验证 $F|_S = f$. 设 $v \in S$, 由题设可得

$$f(v) \le f(u) + L\, d(v,u), \ \ \forall\, u \in S,$$

且 $u = v$ 时等号成立. 这说明 $F(v) = f(v)$.

最后, 我们验证 F 是 Lipschitz 函数. 当 $x, y \in \mathbb{R}^n$ 时, 由三角不等式可得

$$f(u) + L\, d(x,u) \le f(u) + L\, d(y,u) + L\, d(x,y),$$

在上式中关于 $u \in S$ 取下确界可得 $F(x) \le F(y) + L\, d(x,y)$. 交换 x, y 的位置不等式当然也成立, 这说明 $|F(x) - F(y)| \le L\, d(x,y)$.

F 称为 f 的 Lipschitz 延拓. 利用上述引理, 我们还可以将向量值的 Lipschitz 函数做类似的延拓.

推论 7.4.3. *设 S 为 $\mathbb{R}^n$ 中的子集, $f : S \to \mathbb{R}^m$ 为 Lipschitz 映射, 则存在 Lipschitz 映射 $F : \mathbb{R}^n \to \mathbb{R}^m$, 使得 $F|_S = f$.*

证明 设 f 满足条件 7.5. 记 $f(x) = \big(f_1(x), f_2(x), \cdots, f_m(x)\big)$, 则

$$|f_i(x) - f_i(y)| \le d\big(f(x), f(y)\big) \le L\, d(x,y), \ \ \forall\, 1 \le i \le m.$$

由前一引理, 存在 $F_i : \mathbb{R}^n \to \mathbb{R}$, 使得 $F_i|_S = f_i$, 且 $|F_i(x) - F_i(y)| \le L\, d(x, y)$. 记 $F(x) = \big(F_1(x), F_2(x), \cdots, F_m(x)\big)$, 则

$$d\big(F(x), F(y)\big) = \|F(x) - F(y)\| \le \sqrt{m} L\, d(x, y).$$

注. 1934 年, Kirszbraun 证明了更强的结论: 可以适当地选取延拓 F, 使得 $d\big(F(x), F(y)\big) \le L\, d(x, y)$.

最后我们说明 Lipschitz 映射与零测集之间的关系.

命题 7.4.4. *设 S 为 $\mathbb{R}^n$ 中的零测集, $f : S \to \mathbb{R}^n$ 为 Lipschitz 映射, 则 $f(S)$ 也是 $\mathbb{R}^n$ 中的零测集.*

证明 根据上述推论, 不妨设 f 定义在整个 $\mathbb{R}^n$ 中. 根据零测集的定义以及引理 7.4.1, 任给 $\varepsilon > 0$, 存在至多可数个方体 $\{I_i = I_{r_i}(u_i)\}$, 使得

$$S \subset \bigcup_{i\ge1} I_i,\ \ \sum_{i\ge1} \nu(I_i) = \sum_{i\ge1} (2r_i)^n < \varepsilon.$$

设 $\|f(x) - f(y)\| \le L\,\|x - y\|$, 则当 $x \in I_i$ 时,

$$\|f(x) - f(u_i)\| \le L\,\|x - u_i\| \le \sqrt{n} L r_i.$$

这说明 $f(I_i) \subset B_{\delta_i}\big(f(u_i)\big)$, 其中 $\delta_i = \sqrt{n} L r_i$. 记 $J_i = I_{\delta_i}\big(f(u_i)\big)$, 则 $J_i \supset f(I_i)$, 于是

$$f(S) \subset \bigcup_{i\ge1} f(I_i) \subset \bigcup_{i\ge1} J_i,$$

且

$$\sum_{i\ge1} \nu(J_i) = \sum_{i\ge1} (2\sqrt{n} L r_i)^n < (\sqrt{n} L)^n \varepsilon,$$

这说明 $f(S)$ 为 $\mathbb{R}^n$ 中的零测集.

设 $A \in M_{n\times n}$, $b \in \mathbb{R}^n$, 记 $\varphi(x) = Ax + b$, 映射 $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ 称为仿射变换. 由范数不等式可知

$$\|\varphi(x) - \varphi(y)\| = \|A(x - y)\| \le \|A\|\,\|x - y\|,$$

这说明仿射变换是 Lipschitz 映射, 因此将零测集映为零测集.

习题 7.4

1. 设 S, T 为 $\mathbb{R}^n$ 中的子集, 证明:

$$\chi_{S\cap T} = \chi_S\, \chi_T,\ \ \chi_{S\cup T} = \chi_S + \chi_T - \chi_S\, \chi_T.$$

2. 设 $S \subset \mathbb{R}^n$, $u \in S$. 如果存在 $r > 0$, 使得 $B_r(u) \cap S = \{u\}$, 则称 u 为 S 的孤立点. 证明: 孤立点必为边界点.

3. 设 $S \subset \mathbb{R}^n$, $u \in \mathbb{R}^n$. 如果任给 $r > 0$, $B_r(u) \setminus \{u\}$ 中均含有 S 中的点, 则称 u 为 S 的聚点. 聚点的全体记为 S'. 证明: S' 为闭集, 且 $S' \cup \partial S = \bar{S}$.

4. 设 $S \subset \mathbb{R}^n$, $u \in \bar{S}$. 证明: S 中存在收敛于 u 的点列.

5. 设 $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}^m$ 为映射, u 为 S 的聚点. 如果存在 $c \in \mathbb{R}^m$, 使得任给 $\varepsilon > 0$, 均存在 $\delta > 0$, 使得 $f\big(B_\delta^S(u) \setminus u\big) \subset B_\varepsilon(c)$, 则称 f 在 u 处以 c 为极限, 记为 $\lim\limits_{x\to u} f(x) = c$. 证明: f 在 u 处连续当且仅当 $\lim\limits_{x\to u} f(x) = f(u)$.

6. 设 S 为 $\mathbb{R}^n$ 中的零测集, 证明: $S \times \mathbb{R}^m$ 为 $\mathbb{R}^{n+m}$ 中的零测集.

7. 设 I 为 $\mathbb{R}^n$ 中的 n 维矩形, 证明: ∂I 为 $\mathbb{R}^n$ 中的零测集.

8. 设 V 为 $\mathbb{R}^n$ 中的仿射子向量空间. 证明: 若 $V \neq \mathbb{R}^n$, 则 V 为 $\mathbb{R}^n$ 中的零测集.

9. 证明: S^{n-1} 是 $\mathbb{R}^n$ 中零测集.

7.5 凸集和凸函数

Warning: 本节内容为选读材料.

我们把一元凸函数的概念和基本性质推广到多元函数. 设 $x, y \in \mathbb{R}^n$, 记 $[x, y] = \{(1-t)x + ty \in \mathbb{R}^n \mid t \in [0,1]\}$, 它是连接 x, y 的直线段.

凸集: 设 $C \subset \mathbb{R}^n$. 如果任给 $x, y \in C$, 均有 $[x, y] \subset C$, 则称 C 为 $\mathbb{R}^n$ 中的凸集. 显然, 凸集都是道路连通集.

凸函数: 设 C 为 $\mathbb{R}^n$ 中的凸集, $f : C \to \mathbb{R}$ 为定义在 C 中的函数. 如果任给 $x, y \in C$ 以及 $t \in [0,1]$, 均有

$$f\big((1-t)x + ty\big) \le (1-t)f(x) + tf(y),$$

则称 f 为凸函数. 如果 $-f$ 为凸函数, 则称 f 为凹函数. 从定义可以看出, f 为凸(凹)函数当且仅当它限制在定义域中的每一条直线段上均为一元凸(凹)函数.

例 7.5.1. 凸集的例子.

显然, $\mathbb{R}^n$ 中的开球和闭球, 开长方体和闭长方体都是凸集. 设 $\alpha \in \mathbb{R}^n \setminus \{0\}$, $\beta \in \mathbb{R}$, 则 $\mathbb{R}^n$ 中的仿射超平面 $\{x \in \mathbb{R}^n \mid \alpha \cdot x + \beta = 0\}$ 是闭的凸集, 而 $\{x \in \mathbb{R}^n \mid \alpha \cdot x + \beta > 0\}$ 和 $\{x \in \mathbb{R}^n \mid \alpha \cdot x + \beta < 0\}$ 均为开的凸集. 一般地, 若 f 为凸函数, $c \in \mathbb{R}$, 则 $\{x \mid f(x) < c\}$ 和 $\{x \mid f(x) \le c\}$ 均为凸集.

例 7.5.2. 设 $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, 记 $f(x) = \ln\left(\mathrm{e}^{x_1} + \mathrm{e}^{x_2} + \cdots + \mathrm{e}^{x_n}\right)$, 则 f 为凸函数.

证明 因为 f 连续, 只需要验证 f 满足中点凸性:

$$f\Big(\frac{x+y}{2}\Big) \le \frac{1}{2}\big[f(x)+f(y)\big],\ \ \forall\, x,y\in\mathbb{R}^n.$$

事实上, 利用 Cauchy 不等式可得

$$\begin{aligned} f\Big(\frac{x+y}{2}\Big) &= \ln\Big(\sum_{i=1}^n \mathrm{e}^{\frac{x_i+y_i}{2}}\Big) = \ln\Big(\sum_{i=1}^n \sqrt{\mathrm{e}^{x_i}}\sqrt{\mathrm{e}^{x_i}}\Big) \\ &\le \ln\Big(\sum_{i=1}^n \mathrm{e}^{x_i}\Big)^{\frac12}\Big(\sum_{i=1}^n \mathrm{e}^{y_i}\Big)^{\frac12} = \frac12\big[f(x)+f(y)\big]. \end{aligned}$$

例 7.5.3. 记 $\mathbb{R}^n_{++} = \{(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n \mid x_i>0,\ 1\le i\le n\}$. 在 $\mathbb{R}^n_{++}$ 中定义几何平均值函数 $f(x) = \big(x_1x_2\cdots x_n\big)^{\frac1n}$, 则 f 为凹函数.

证明 因为 f 连续, 只需验证它满足中点凹性:

$$\Big[\prod_{i=1}^n \frac{x_i+y_i}{2}\Big]^{\frac1n} \ge \frac12\Big[\Big(\prod_{i=1}^n x_i\Big)^{\frac1n} + \Big(\prod_{i=1}^n y_i\Big)^{\frac1n}\Big].$$

记 $\lambda = \big(x_1x_2\cdots x_n\big)^{\frac1n}$, $\mu = \big(y_1y_2\cdots y_n\big)^{\frac1n}$. 利用归一化的方法, 不妨设 $\lambda+\mu=1$. 于是, 由 $\ln t$ 为一元凹函数可得

$$\ln(x_i+y_i) = \ln\Big(\lambda\,\frac{x_i}{\lambda} + \mu\,\frac{y_i}{\mu}\Big) \ge \lambda\,\ln\Big(\frac{x_i}{\lambda}\Big) + \mu\,\ln\Big(\frac{y_i}{\mu}\Big),$$

求和可得

$$\sum_{i=1}^n \ln(x_i+y_i)\ge 0,\ \ \text{或}\ \ \prod_{i=1}^n (x_i+y_i)\ge 1,$$

由此即得欲证结论.

例 7.5.4. 设 A 为 n 阶半正定方阵, $\alpha\in\mathbb{R}^n$, $\beta\in\mathbb{R}$. 在 $\mathbb{R}^n$ 中定义函数 $f(x) = xAx^\top + \alpha\cdot x + \beta$, 则 f 为凸函数.

证明 任取 $x,y\in\mathbb{R}^n$, 只需验证 $\phi(t) = f\big((1-t)x+ty\big)$ 是关于 t 的一元凸函数即可. 简单的计算表明

$$\phi''(t) = 2(y-x)A(y-x)^\top \ge 0,\ \ \forall\, t\in\mathbb{R},$$

这说明 $\phi(t)$ 的确是凸函数.

下面我们将一元凸函数的一些简单性质推广到多元凸函数, 先看连续性.

命题 7.5.1. 设 C 为 $\mathbb{R}^n$ 中的凸集, $f: C\to\mathbb{R}$ 为凸函数. 则 f 在 C 的内部连续.

证明 我们对维数 n 进行归纳. $n=1$ 时我们有**命题 5.3.3**. 设结论对 $n-1$ 元函数成立, 往证结论对 n 元函数也成立. 任取 C 的内点 u, 则存在 $r>0$, 使得 $I_{2r}(u)$ 包含于 C 的内部. 注意 $I_r(u)$ 和 $I_{2r}(u)$ 的边界分别由 2^n 个 $n-1$ 维方体组成, f 限制在每一个 $n-1$ 维方体上均为凸函数, 由归纳假设可知它在上述边界上连续.

记 $\omega = \max\{|f(v)-f(w)| \mid v\in\partial I_r(u),\ w\in\partial I_{2r}(u)\}$. 任取 $x\neq y\in I_r(u)$, 将直线段 $[x,y]$ 分别向两端延长, 分别交 $\partial I_r(u)$ 于 x',y', 交 $\partial I_{2r}(u)$ 于 x'',y''. f 限制在直线段 $[x'',y'']$上是一元凸函数, 根据**命题 5.3.3** 的证明可得

$$\frac{|f(x)-f(y)|}{d(x,y)} \le \max\left\{\frac{|f(x')-f(x'')|}{d(x',x'')},\ \frac{|f(y')-f(y'')|}{d(y',y'')}\right\} \le \frac{\omega}{r},$$

这说明 f 是 $I_r(u)$ 中的 Lipschitz 函数, 特别地它在 u 处连续.

我们再将一元凸函数支撑线的概念推广到多元凸函数. 设 C 为 $\mathbb{R}^n$ 中的凸集, $f: C\to\mathbb{R}$ 为凸函数.

次梯度: 设 $x\in C$. 如果存在 $\alpha\in\mathbb{R}^n$, 使得

$$f(y)\ge f(x)+\alpha\cdot(y-x),\ \ \forall\, y\in C,$$

则称 f 在 x 处存在次梯度, α 称为 f 在 x 处的一个次梯度, x 处的所有次梯度构成的集合记为 $\partial f(x)$.

例如, 一元凸函数 $\phi(t)=|t|$ 在原点处的不可导但有次梯度, 事实上 $\partial\phi(0)=[-1,1]$. 一般地, 一元凸函数在定义域内部的每一点处都有支撑线, 支撑线的斜率就是次梯度. 我们想对多元凸函数证明类似的结论, 先从几何的角度简单地分析一下. 从定义可知, $\alpha\in\partial f(x)$ 当且仅当 f 的图像位于仿射超平面 $\{(u,t)\in\mathbb{R}^{n+1} \mid u\in\mathbb{R}^n,\ t=f(x)+\alpha\cdot u\}$ 的上方, 此时将这样的仿射超平面称为 f 的一个支撑面. 我们可以把次梯度的存在性问题转化为支撑面的存在性问题.

为了寻找支撑面, 我们注意到不仅 f 的图像位于支撑面的上方, 若记 $\mathrm{epi}(f)=\{(x,t)\in\mathbb{R}^{n+1} \mid x\in C,\ t\ge f(x)\}$, 则 $\mathrm{epi}(f)$ 也位于支撑面的上方. 容易验证 $\mathrm{epi}(f)$ 是凸集.

引理 7.5.2. *设 C 为 $\mathbb{R}^n$ 中闭的非空凸集, $u\notin C$. 则存在唯一的 $v\in C$, 使得*

$$d(u,v)=d(u,C),\ \ C\subset\{x\in\mathbb{R}^n \mid (x-v)\cdot(u-v)\le 0\}.$$

此时 C 和 u 分别位于仿射超平面 $\{x\in\mathbb{R}^n \mid (x-v)\cdot(u-v)=0\}$ 的两边.

证明 取 C 中点列 $\{v_i\}$, 使得 $\lim\limits_{i\to\infty} d(u,v_i)=d(u,C)$. 由平行四边形公式和 C 为凸集可得

$$\begin{aligned}0\le\left\|\frac{v_i-v_j}{2}\right\|^2 &= \frac{1}{2}\big[\|v_i-u\|^2+\|v_j-u\|^2\big]-\left\|\frac{v_i+v_j}{2}-u\right\|^2\\ &\le \frac{1}{2}\big[\|v_i-u\|^2+\|v_j-u\|^2\big]-d^2(u,C),\end{aligned}$$

这说明 $\{v_i\}$ 为 Cauchy 列, 其极限记为 v, 由 C 为闭集可知 $v \in C$. 此时 $d(u,v) = d(u,C)$. 若另有 $v' \in C$, 使得 $d(u,v') = d(u,C)$, 则将上述平行四边形公式中的 v_i, v_j 分别换成 v, v' 后可立即知道 $v - v' = 0$.

任给 $x \in C$, 由 C 为凸集可知 $[v,x] \subset C$, 因此函数 $\phi(t) = d^2\big(u,\, v + t(x-v)\big)$ 在 $t = 0$ 处达到最小值, 这说明

$$0 \leq \lim_{t\to 0^+} \frac{1}{t}\big[\|u - v - t(x-v)\|^2 - \|u-v\|^2\big] = -2(x-v)\cdot(u-v).$$

推论 7.5.3. *设 C 为 $\mathbb{R}^n$ 中的闭凸集, $u \in \partial C$, 则存在 $\alpha \in \mathbb{R}^n \setminus \{0\}$, 使得 $C \subset \{x \in \mathbb{R}^n \mid \alpha\cdot(x-u) \geq 0\}$.*

证明 取 $\mathbb{R}^n \setminus C$ 中点列 $\{u_i\}$, 使得 $\lim\limits_{i\to\infty} u_i = u$. 根据上述引理, 存在 $v_i \in C$, 使得

$$d(u_i, v_i) = d(u_i, C),\ \ C \subset \{x \in \mathbb{R}^n \mid (x - v_i)\cdot(u_i - v_i) \leq 0\}.$$

由三角不等式可得

$$d(v_i, u) \leq d(v_i, u_i) + d(u_i, u) = d(u_i, C) + d(u_i, u) \leq 2\,d(u_i, u),$$

这说明 $\lim\limits_{i\to\infty} v_i = u$. 记 $\alpha_i = \frac{v_i - u_i}{\|v_i - u_i\|}$, 则 $\{\alpha_i\} \subset S^{n-1}$, 因此它有收敛子列, 不妨设 $\{\alpha_i\}$ 本身收敛于 α. 注意到

$$C \subset \{x \in \mathbb{R}^n \mid \alpha_i\cdot(x - v_i) \geq 0\},$$

令 $i \to \infty$ 即得欲证结论.

推论 7.5.4. *设 $f: \mathbb{R}^n \to \mathbb{R}$ 为凸函数, 则任给 $x \in \mathbb{R}^n$, 均有 $\partial f(x) \neq \emptyset$.*

证明 由命题 7.5.1 可知 f 连续, 此时不难验证 $\mathrm{epi}(f)$ 是 $\mathbb{R}^{n+1}$ 中的闭凸集, 其边界就是 f 的图像. 任给 $x \in \mathbb{R}^n$, 根据前一推论, 存在不同时为零的 $\alpha' \in \mathbb{R}^n$ 以及 $\beta \in \mathbb{R}$, 使得

$$\alpha'\cdot(y-x) + \beta\big(t - f(x)\big) = (\alpha', \beta)\cdot\big(y - x, t - f(x)\big) \geq 0,\ \ \forall\, y \in \mathbb{R}^n,\ t \geq f(y).$$

在上式中代入 $y = x - \alpha'$, $t = |f(x)| + |f(y)| + 1$ 可知 $\beta > 0$. 记 $\alpha = -\frac{\alpha'}{\beta}$, 则 $\alpha \in \partial f(x)$.

命题 7.5.5. *设 C 为 $\mathbb{R}^n$ 中的凸集, $f: C \to \mathbb{R}$ 为凸函数. 若 x 为 C 的内点, 则 $\partial f(x) \neq \emptyset$.*

证明 我们想办法将问题转化为上述推论中情形. 为此, 任给 $v \in \mathbb{R}^n$, 当 $|t|$ 充分小时, $x + tv \in C$, 此时 $\phi(t) = \phi(x + tv)$ 为一元凸函数. 由**命题 5.3.3**可知 $\phi'_+(0)$ 存在, 记为 $\ell_x(v)$, 则 $\ell_x(v) \in \partial\phi(0)$. 显然, $\ell_x(0) = 0$.

不难验证 $\ell_x:\mathbb{R}^n\to\mathbb{R}$ 为凸函数. 由上述推论, $\partial\ell_x(0)\neq\emptyset$. 取 $\alpha\in\partial\ell_x(0)$, 当 $y\in C$ 时, 取 $v=y-x$ 可得

$$f(y)=\phi(1)\geq\phi(0)+\phi'_+(0)=f(x)+\ell_x(y-x)\geq f(x)+\alpha\cdot(y-x),$$

这说明 $\alpha\in\partial f(x)$.

注. 从上述证明不难看出 $\partial\ell_x(0)=\partial f(x)$. 还可以看出, 若 x 为 f 的极小值点(即局部最小值点), 则 x 必为 f 的最小值点. 事实上, 此时 ℓ_x 非负, 于是 $0\in\partial\ell_x(0)=\partial f(x)$, 这说明 $f(y)\geq f(x)$ 对每一个 $y\in C$ 均成立.

对于一元凸函数, 我们知道它在 x 处可导当且仅当它在 x 处的左右导数相等, 即它在 x 处具有唯一的次梯度. 为了将此结论推广到多元函数, 我们先讨论 $\partial f(x)$ 如何随 x 变化而变化. 设 $S\subset\mathbb{R}^n$, $\delta>0$, 记 $U_\delta(S)=\{y\in\mathbb{R}^n\mid d(y,S)<\delta\}$. 我们有

引理 7.5.6. *设 f 和 x 如前一命题, 则 $\partial f(x)$ 是有界的闭凸集, 且任给 $\varepsilon>0$, 存在 $\delta>0$, 使得只要 $d(y,x)<\delta$, 就有 $\partial f(y)\subset U_\delta\big(\partial f(x)\big)$.*

证明 由次梯度的定义不难看出 $\partial f(x)$ 是闭的凸集. 为了说明其有界性, 取 $r>0$, 使得 $\overline{B_{2r}(x)}$ 包含于 C 的内部. 由命题 7.5.1 可知 f 在 $\overline{B_{2r}(x)}$ 中连续, 记 $\omega=\max\{|f(y)-f(z)|\mid y,z\in\overline{B_{2r}(x)}\}$. 当 $y\in B_r(x)$ 时, 由三角不等式可知 $\overline{B_r(y)}\subset B_{2r}(x)$. 此时, 任给 $z\in\overline{B_r(y)}$, $\beta\in\partial f(y)$, 有

$$\omega\geq f(z)-f(y)\geq\beta\cdot(z-y),$$

由此可知 $\|\beta\|\leq\frac{\omega}{r}$, 这说明 $\partial f(y)$ 有界.

(反证法)若剩下的欲证结论不成立, 则存在 $\varepsilon_0>0$ 以及点列 $\{y_i\}$, $\beta_i\in\partial f(y_i)$, 使得 $d(y_i,x)<\frac{r}{i}$ 且 $d\big(\beta_i,\partial f(x)\big)\geq\varepsilon_0$. 由前一段的论证可知 $\{\beta_i\}$ 是有界点列, 从而有收敛子列. 不妨设 $\{\beta_i\}$ 本身收敛于 β. 一方面, 在不等式

$$f(y)\geq f(y_i)+\beta_i\cdot(y-y_i)$$

中令 $i\to\infty$ 可知 $\beta\in\partial f(x)$. 另一方面, 又有

$$d\big(\beta,\partial f(x)\big)=\lim_{i\to\infty}d\big(\beta_i,\partial f(x)\big)\geq\varepsilon_0>0,$$

这就导出了矛盾.

命题 7.5.7. *设 f 和 x 如前一命题, 则 $\partial f(x)=\{\alpha\}\iff$ 任给 $\varepsilon>0$, 存在 $\delta>0$, 使得只要 $\|h\|<\delta$, 就有*

$$|f(x+h)-f(x)-\alpha\cdot h|\leq\varepsilon\|h\|.$$

此时称 f 在 x 处可微, α 称为 f 在 x 处的梯度, 记为 $\nabla f(x)$.

证明 "$\Longrightarrow$" 由前一引理可知, 任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使得只要 $\|h\| < \delta$, $\beta \in \partial f(x+h)$, 就有 $\|\beta - \alpha\| < \varepsilon$. 此时有

$$0 \le f(x+h) - f(x) - \alpha \cdot h \le (\beta - \alpha) \cdot h \le \varepsilon \|h\|.$$

"$\Longleftarrow$" 若另有 $\alpha' \in \partial f(x)$, 则任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使得只要 $\|h\| < \delta$, 就有

$$(\alpha' - \alpha) \cdot h \le f(x+h) - f(x) - \alpha \cdot h \le \varepsilon \|h\|,$$

这说明 $\|\alpha' - \alpha\| \le \varepsilon$. 由 ε 的任意性即知 $\alpha' = \alpha$.

习题 7.5

1. 设 C 为 $\mathbb{R}^n$ 中的凸集, 则 $\bar{C}$ 也是凸集; 设 C 为 $\mathbb{R}^n$ 中的闭集, 如果任给 $x, y \in C$, 均有 $\frac{x+y}{2} \in C$, 则 C 为凸集.

2. 记 $S^n_{++} = \{A \in GL(n, \mathbb{R}) \mid A \text{ 正定对称}\}$. 证明: S^n_{++} 为 $M_{n\times n}$ 中的凸集, 且函数 $f(A) = \ln\det A$ 是 S^n_{++} 中定义的凹函数.

3. 设 $\phi(t)$ 为一元凸函数, 当 $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ 时, 记 $f(x) = \ln\big(\mathrm{e}^{\phi(x_1)} + \mathrm{e}^{\phi(x_2)} + \cdots + \mathrm{e}^{\phi(x_n)}\big)$, 证明 f 为凸函数.

4. 设 $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n_{++}$, $(y_1, y_2, \cdots, y_n) \in \mathbb{R}^n_{++}$.

   (1) 当 $p > 1$ 时, 利用归一法和 t^p 为凸函数证明如下 Minkowski 不等式:

   $$\Big[\sum_{i=1}^n (x_i + y_i)^p\Big]^{\frac{1}{p}} \le \Big[\sum_{i=1}^n x_i^p\Big]^{\frac{1}{p}} + \Big[\sum_{i=1}^n y_i^p\Big]^{\frac{1}{p}}.$$

   (2) 当 $0 < p < 1$ 时, 利用归一法和 t^p 为凹函数证明如下反向 Minkowski 不等式:

   $$\Big[\sum_{i=1}^n (x_i + y_i)^p\Big]^{\frac{1}{p}} \ge \Big[\sum_{i=1}^n x_i^p\Big]^{\frac{1}{p}} + \Big[\sum_{i=1}^n y_i^p\Big]^{\frac{1}{p}}.$$

5. 设 $p \ne 0$, 当 $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n_{++}$ 时, 记 $f(x) = \big(\frac{x_1^p + x_2^p + \cdots + x_n^p}{n}\big)^{\frac{1}{p}}$. 证明: 当 $p < 0$ 或 $p \ge 1$ 时 f 为凸函数, 当 $0 < p < 1$ 时 f 为凹函数.

6. 设 C 为 $\mathbb{R}^n$ 中的非空闭凸集, 任给 $x \in \mathbb{R}^n$, 记 $P(x)$ 是 C 中满足 $d\big(x, P(x)\big) = d(x, C)$ 的唯一点. 证明: $d\big(P(x), P(y)\big) \le d(x, y)$.

7. 设 C, D 是 $\mathbb{R}^n$ 中的非空凸集, 记 $C - D = \{x - y \in \mathbb{R}^n \mid x \in C,\ y \in D\}$. 验证 $C - D$ 是凸集; 若 $C \cap D = \emptyset$, 则存在 $\alpha \in \mathbb{R}^n \setminus \{0\}$, 使得任给 $x \in C$, $y \in D$, 均有 $\alpha \cdot x \le \alpha \cdot y$.

8. 沿用命题 7.5.5 中记号, 证明 $\ell_x(v) = \sup\{\alpha \cdot v \mid \alpha \in \partial f(x)\}$.

9. 设 C 为 $\mathbb{R}^n$ 中的凸集, $f: C \to \mathbb{R}$ 为凸函数. 记

$$C^* = \{x^* \in \mathbb{R}^n \mid \{x^* \cdot x - f(x) \mid x \in C\} \text{ 有上界}\}.$$

则 C^* 为凸集. 当 $x^* \in C^*$ 时, 定义 $f^*(x^*) = \sup\{x^* \cdot x - f(x) \mid x \in C\}$, 证明 f^* 为凸函数, 且若 x 为 C 的内点, 则 $(f^*)^*(x) = f(x)$.

7.6 附录: 距离空间

从前面几节的讨论中我们可以看到, 欧氏空间中的内积给出了向量的范数, 进而给出了两点之间的距离, 许多重要的结论都来源于距离函数的性质. 距离的概念可以做如下推广.

定义 7.6.1 (距离空间)**.** *设 X 为非空集合, $d: X \times X \to \mathbb{R}$ 为函数. 如果 d 满足以下条件*

(1) *任给 $x, y \in X$, 均有 $d(x, y) \geq 0$, 且 $d(x, y) = 0 \Longleftrightarrow x = y$, (正定性)*

(2) *任给 $x, y \in X$, 均有 $d(x, y) = d(y, x)$, (对称性)*

(3) *任给 $x, y, z \in X$, 均有 $d(x, z) \leq d(x, y) + d(y, z)$, (三角不等式)*

则称 d 为 X 中的距离或度量, (X, d) 称为距离空间或度量空间, $d(x, y)$ 称为 x, y 之间的距离.

例如, 在欧氏空间 $\mathbb{R}^n$ 中, $d(x, y) = \|x - y\|$ 就是满足上述条件的距离. 如果 $S \subset \mathbb{R}^n$, 同样的公式也给出了 S 中的距离, 这时 (S, d) 称为欧氏空间的子距离空间. 我们再来介绍几个例子.

例 7.6.1. *$\mathbb{R}^n$ 中的 L^p 距离.*

设 $p \geq 1$, $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, 记

$$|x|_p = \left[|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right]^{\frac{1}{p}}.$$

当 $x, y \in \mathbb{R}^n$ 时, 定义 $d_p(x, y) = |x - y|_p$. 显然, $d_2(x, y) = \|x - y\|$. 我们来验证 d_p 都是 $\mathbb{R}^n$ 中的距离, 只需要验证 d_p 满足三角不等式即可. 当 $p = 1$ 时, 三角不等式可以由绝对值不等式给出. 下设 $p > 1$, 我们先来证明所谓的 Minkowski 不等式:

$$|u + v|_p \leq |u|_p + |v|_p, \ \forall\, u, v \in \mathbb{R}^n. \tag{7.6}$$

事实上, 记 $u = (u_1, u_2, \cdots, u_n)$, $v = (v_1, v_2, \cdots, v_n)$, 则

$$|u + v|_p^p = \sum_{i=1}^n |u_i + v_i|\, |u_i + v_i|^{p-1} \leq \sum_{i=1}^n |u_i|\, |u_i + v_i|^{p-1} + \sum_{i=1}^n |v_i|\, |u_i + v_i|^{p-1}.$$

对上式右边两项分别应用 Hölder 不等式可得

$$|u + v|_p^p \leq |u|_p\, |u + v|_p^{p-1} + |v|_p\, |u + v|_p^{p-1},$$

整理以后即得 (7.6). 有了此式, 若 $x, y, z \in \mathbb{R}^n$, 则

$$|x-z|_p = |(x-y)+(y-z)|_p \leq |x-y|_p + |y-z|_p,$$

这说明 d_p 满足三角不等式. d_p 称为 $\mathbb{R}^n$ 中的 L^p 距离.

令 $p \to \infty$, 由数列极限可知 $|x|_p \to |x|_\infty = \max\{|x_i| \mid 1 \leq i \leq n\}$, 于是 $d_\infty(x,y) = |x-y|_\infty$ 也定义了 $\mathbb{R}^n$ 中的一个距离, 称为 L^∞ 距离或最大模距离.

例 7.6.2. $C[a,b]$ 中的 L^p 距离.

设 $p \geq 1$, $f \in C[a,b]$, 记

$$|f|_p = \Big(\int_a^b |f(x)|^p \,\mathrm{d}x\Big)^{\frac{1}{p}}.$$

当 $f, g \in C[a,b]$ 时, 定义 $d_p(f,g) = |f-g|_p$. 与前例类似, 可以说明 d_p 是 $C[a,b]$ 中的距离, 称为 L^p 距离. 令 $p \to \infty$ 还可以得到 L^∞ 距离或最大模距离:

$$d_\infty(f,g) = \max\big\{|f(x)-g(x)| \,\big|\, x \in [a,b]\big\}.$$

例 7.6.3. $M_{n\times n}$ 中的 d_m 距离.

设 $A \in M_{n\times n}$, 令 $|A|_m = \max\{\|Ax\| \mid x \in S^{n-1}\}$. 当 $A, B \in M_{n\times n}$ 时, 记 $d_m(A,B) = |A-B|_m$, 不难验证 d_m 是 $M_{n\times n}$ 中的距离.

例 7.6.4. 离散距离空间.

设 X 为任意一个非空集合, 定义函数 $d: X\times X \to \mathbb{R}$ 如下:

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

不难验证 d 为 X 中的距离, (X,d) 称为离散距离空间.

欧氏空间以及连续映射的许多重要性质都可以平行地推广到距离空间中. 凡是用距离定义的概念、结果都可以移植到距离空间中, 如开集、闭集、连续映射、道路连通子集等, 这些我们不再重复叙述. 不过, 有两个重要的地方值得注意: 一是距离空间中的有界点列未必有收敛子列, 二是距离空间中的 Cauchy 列未必收敛. 比如, 作为 $\mathbb{R}$ 的子距离空间, 开区间 $(0,2)$ 中的点列 $\big\{\frac{1}{n}\big\}$ 是 Cauchy 列, 但它在 $(0,2)$ 中并不收敛. 为此, 我们再引进几个概念. 下设 (X,d) 为距离空间.

预紧集: 设 $S \subset X$. 如果 S 中的每一个点列均有收敛子列(不要求极限仍在 S 中), 则称 S 为 X 中的预紧集.

紧集: 设 $S \subset X$. 如果 S 的每一个开覆盖均存在有限子覆盖(即此覆盖中的有限个开集组成的覆盖), 则称 S 为 X 的紧集.

完备性: 若 X 中的每一个 Cauchy 列均收敛, 则称 (X,d) 为完备距离空间.

我们来看这几个概念之间的联系以及它们的用处. 首先, 如果回顾 Lebesgue 数引理的证明, 我们会发现对于既是预紧集又是闭集的子集, Lebesgue 数引理仍然成立. 其次, 我们有

定理 7.6.1 (紧集的刻画)**.** *设 $S \subset X$, 则 S 为紧集 $\Longleftrightarrow S$ 既是预紧集又是闭集.*

证明 "$\Longleftarrow$" 设 S 既是预紧集又是闭集, $\{U_\alpha\}$ 是 S 的开覆盖, 我们要找它的有限子覆盖. 设 λ 是给定开覆盖的 Lebesgue 数, 记 $\delta = \frac{1}{3}\lambda$, 则对每一个 $u \in S$, $B_\delta(u)$ 的直径都小于 λ, 从而包含于某个 U_α 中. 如果能找有限个 u_i, 使得 $\{B_\delta(u_i)\}$ 覆盖了 S, 则对应的有限个 U_α 就是要找的子覆盖.

为此, 任取 $u_1 \in S$. 若 $S \setminus B_\delta(u_1) \neq \emptyset$, 再任取 $u_2 \in S \setminus B_\delta(u_1)$. 此时 $d(u_2, u_1) \geq \delta$. 若 $S \setminus \big(B_\delta(u_1) \cup B_\delta(u_2)\big) \neq \emptyset$, 继续取 $u_3 \in S \setminus \big(B_\delta(u_1) \cup B_\delta(u_2)\big)$, 此时 $d(u_3, u_1) \geq \delta$, $d(u_3, u_2) \geq \delta$. 依次类推, 我们断言经过有限次步骤之后 $\{B_\delta(u_i)\}$ 就覆盖了 S. (反证法) 若不然, 则在 S 中可找到两两之间距离不小于 δ 的点列 $\{u_i\}$, 它显然不存在收敛子列, 从而与 S 为预紧集相矛盾.

"$\Longrightarrow$" 设 S 为紧集, 我们先来说明它是闭集. 任取 $u \notin S$, 对每一正整数 i, 记 $U_i = \{x \in X \mid d(x, u) > \frac{1}{i}\}$, 则 $\{U_i\}$ 为 S 的开覆盖. 因此存在 k, 使得

$$S \subset \bigcup_{i=1}^{k} U_i = U_k,$$

这说明 $B_{\frac{1}{k}}(u) \subset X \setminus S$, 因此 $X \setminus S$ 是开集, S 是闭集.

设 $\{u_i\}$ 为 S 中的点列, 我们断言存在 $u \in S$, 使得对每个正整数 i, $B_{\frac{1}{i}}(u)$ 中均含有无限项 u_i, 此时容易看出 $\{u_i\}$ 存在收敛于 u 的子列. (反证法) 若对每一 $u \in S$, 均存在 $i = i(u)$, 使得 $B_{\frac{1}{i}}(u)$ 中只含有有限项 u_i. 则 $\{B_{\frac{1}{i}}(u)\}_{u \in S}$ 是 S 的开覆盖, 它存在有限子覆盖. 但这意味着 S 只含有有限项 u_i, 与 $\{u_i\} \subset S$ 相矛盾. 因此 S 是预紧集.

注. 若 S 为预紧集, 不难看出 $\bar{S}$ 也是预紧集, 从而是紧集.

定理 7.6.2 (一致连续性)**.** *设 (Y, ρ) 也是距离空间, $f : S \to Y$ 为连续映射, 其中 S 为 X 中的紧集. 则任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对每一个 $u \in S$, 均有 $f\big(B^S_\delta(u)\big) \subset B_\varepsilon\big(f(u)\big)$.*

证明 任给 $y \in Y$, 由 f 连续可知, 存在 X 中的开集 U_y, 使得 $U_y \cap S = f^{-1}(B_{\frac{\varepsilon}{2}}(y))$. 于是 $\{U_y\}_{y \in Y}$ 是 S 的开覆盖. 设 λ 是此覆盖的 Lebesgue 数, 取 $\delta = \frac{1}{3}\lambda$, 则对每一个 $u \in S$, $B_\delta(u)$ 的直径小于 λ, 它就必包含于某个 U_y 中. 此时由三角不等式可得

$$f\big(B_\delta(u)\big) \subset B_{\frac{\varepsilon}{2}}(y) \subset B_\varepsilon\big(f(u)\big).$$

定理 7.6.3 (最值定理)**.** *设 (Y,ρ) 也是距离空间, $f: X \to Y$ 为连续映射. 若 S 为 X 中的预紧集, 则 $f(S)$ 是 Y 的预紧集; 若 S 为 X 中的紧集, 则 $f(S)$ 是 Y 的紧集.*

证明 设 S 为预紧集, $\{f(u_i)\}$ 是 $f(S)$ 中的点列, 则 $\{u_i\}$ 是 S 中的点列, 因而有收敛子列. 由 f 连续即知 $\{f(u_i)\}$ 中对应的子列也收敛. 这说明 $f(S)$ 是预紧集.

设 S 为紧集, 任给 $f(S)$ 的开覆盖 $\{V_\alpha\}$, 由 f 连续可知存在开集族 U_α, 使得 $U_\alpha \cap S = f^{-1}(V_\alpha)$. 此时 $\{U_\alpha\}$ 是 S 的开覆盖, 从而存在有限子覆盖 $\{U_{\alpha_i}\}_{i=1}^k$. 此时 $\{V_{\alpha_i} = f(U_{\alpha_i})\}_{i=1}^k$ 是 $f(S)$ 的覆盖, 这说明 $f(S)$ 为紧集.

若 X 本身是紧集, 则称 (X,d) 为紧致距离空间.

命题 7.6.4. *紧致距离空间必定是完备距离空间.*

证明 设 (X,d) 为紧致距离空间, $\{u_i\}$ 为 Cauchy 列. 由紧集的刻画可知 X 是预紧集, 这说明 $\{u_i\}$ 存在收敛子列, 由此容易看出 $\{u_i\}$ 本身也是收敛的.

我们来研究前面例子中距离空间 $(C[a,b], d_\infty)$ 的完备性, 读者可以自行检查其余的例子中的距离空间是否完备.

命题 7.6.5. *$(C[a,b], d_\infty)$ 是完备距离空间.*

证明 设 $\{f_i\}$ 为 Cauchy 列, 则对每一个 $x \in [a,b]$, 由 $|f_i(x) - f_j(x)| \le d_\infty(f_i, f_j)$ 可知 $\{f_i(x)\}$ 是 Cauchy 数列, 从而收敛, 其极限记为 $f(x)$.

我们先来说明 $f \in C[a,b]$. 由题设, 任给 $\varepsilon > 0$, 存在 N, 当 $i, j > N$ 时 $d_\infty(f_i, f_j) < \varepsilon$. 暂时固定 $i > N$, 当 $j > N$ 时, 就有

$$|f_i(x) - f_j(x)| \le d_\infty(f_i, f_j) < \varepsilon, \ \forall\, x \in [a,b].$$

令 $j \to \infty$ 可得

$$|f_i(x) - f(x)| \le \varepsilon, \ \forall\, x \in [a,b]. \tag{7.7}$$

由 f_i 一致连续可知, 存在 $\delta > 0$, 使得只要 $|x-y| < \delta$, 就有 $|f_i(x) - f_i(y)| < \varepsilon$. 此时有

$$|f(x) - f(y)| \le |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| < 3\varepsilon,$$

这说明 f 一致连续.

(7.7) 表明 $d_\infty(f_i, f) \le \varepsilon$, 它对每一个 $i > N$ 都成立, 这说明 Cauchy 列 $\{f_i\}$ 收敛于 f.

下面的有用结果刻画了 $(C[a,b], d_\infty)$ 中的预紧集.

定理 7.6.6 (Arzelà-Ascoli)**.** *设 $S \subset C[a,b]$, 则 S 为 $(C[a,b], d_\infty)$ 中的预紧集 $\Longleftrightarrow$ S 满足下面两个条件:*

(i) *(一致有界) 存在 $M \geq 0$, 使得 $|f| \leq M$ 对每一个 $f \in S$ 都成立;*

(ii) *(等度连续) 任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使得只要 $|s-t| < \delta$, $|f(s)-f(t)| < \varepsilon$ 就对每一个 $f \in S$ 都成立.*

证明 "$\Longrightarrow$" 设 S 为预紧集, 容易看出 S 是有界集, 这说明 (i) 成立.

任给 $\varepsilon > 0$, 根据紧集刻画的证明, 存在 S 中的有限集 $\{f_i\}_{i=1}^k$, 使得当 $i \neq j$ 时 $d_\infty(f_i, f_j) \geq \frac{\varepsilon}{3}$, 且 $\{B_{\frac{\varepsilon}{3}}(f_i)\}_{i=1}^k$ 为 S 的覆盖. 因为 f_i 一致连续, 存在 $\delta_i > 0$, 使得只要 $|s-t| < \delta_i$, 就有 $|f_i(s) - f_i(t)| < \frac{\varepsilon}{3}$. 记 $\delta = \min\{\delta_i \mid 1 \leq i \leq k\}$, 当 $|s-t| < \delta$, $f \in S$ 时, 设 $f \in B_{\frac{\varepsilon}{3}}(f_i)$, 则

$$\begin{aligned}|f(s) - f(t)| &\leq |f(s) - f_i(s)| + |f_i(s) - f_i(t)| + |f_i(t) - f(t)| \\ &\leq d_\infty(f, f_i) + \frac{\varepsilon}{3} + d_\infty(f_i, f) < \varepsilon.\end{aligned}$$

这就证明了 (ii).

"$\Longleftarrow$" 设 $\{f_i\} \subset S$, 我们要找到一个收敛子列.

断言: 任给 $\varepsilon > 0$, $\{f_i\}$ 均存在子列, 使得该子列中任意两项之间的距离都小于 ε.

先承认此断言, 我们可以从 $\{f_i\}$ 中找出一个 Cauchy 子列, 再由前一命题即知它是收敛子列. 事实上, 先取 $\varepsilon = \frac{1}{2}$, 则存在 $\{f_i\}$ 的子列, 使得任意两项之间的距离都小于 $\frac{1}{2}$. 在该子列中任取一项 f_{i_1}. 再取 $\varepsilon = \frac{1}{2^2}$, 上述子列就存在一个子列, 使得任意两项之间的距离都小于 $\frac{1}{2^2}$. 在该子列的子列中取一项 f_{i_2}, 使得 $i_2 > i_1$. 依次类推, 可以找到 $\{f_i\}$ 的子列 $\{f_{i_j}\}$, 使得

$$d_\infty(f_{i_j}, f_{i_k}) < \frac{1}{2^k}, \ \forall\, j \geq k \geq 1.$$

显然, $\{f_{i_j}\}$ 是 Cauchy 列.

最后我们给出上述断言的证明. 任给 $\varepsilon > 0$, 根据 (ii), 存在 $\delta > 0$, 使得当 $|s-t| < \delta$ 时 $|f(s)-f(t)| < \frac{\varepsilon}{3}$ 对每一个 $f \in S$ 都成立. 取正整数 $m > \frac{b-a}{\delta}$, 将 $[a,b]$ 作 m 等分, 分点为 $t_\ell = a + \frac{\ell}{m}(b-a)$, $\ell = 0, 1, \cdots, m$. 定义映射 $\Phi: S \to \mathbb{R}^{m+1}$ 如下:

$$\Phi(f) = \big(f(t_0), f(t_1), \cdots, f(t_m)\big), \ \forall\, f \in S.$$

由 (i) 可知 $\Phi(S)$ 是 $\mathbb{R}^{m+1}$ 中的有界子集, 因此 $\{\Phi(f_i)\}$ 有收敛子列. 特别地, 存在子列 $\{f_{i_j}\}$, 使得

$$|f_{i_j}(t_\ell) - f_{i_k}(t_\ell)| \leq \|\Phi(f_{i_j}) - \Phi(f_{i_k})\| < \frac{\varepsilon}{3}, \ \forall\, j, k \geq 1, \ 0 \leq \ell \leq m.$$

任取 $t \in [a,b]$, 则存在 t_ℓ, 使得 $|t - t_\ell| < \frac{b-a}{m} < \delta$, 因此

$$\begin{aligned}|f_{i_j}(t) - f_{i_k}(t)| &\leq |f_{i_j}(t) - f_{i_j}(t_\ell)| + |f_{i_j}(t_\ell) - f_{i_k}(t_\ell)| + |f_{i_k}(t_\ell) - f_{i_k}(t)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,\end{aligned}$$

这说明 $d_\infty(f_{i_j}, f_{i_k}) < \varepsilon$, 断言证毕.

注. (1) 上述定理中"一致有界"的条件可以改成: 存在 $t_0 \in [a,b]$, 使得 $\{f(t_0) \mid f \in S\}$ 为有界集. 此时定理结论仍成立.

(2) 上述定理可以推广: 若 (X,d) 为紧致度量空间, 记 $C(X,\mathbb{R}^n) = \{f : X \to \mathbb{R}^n \mid f \text{ 为连续映射}\}$. 当 $f, g \in C(X,\mathbb{R}^n)$ 时, 定义

$$d_\infty(f,g) = \max\big\{\|f(x) - g(x)\| \big| x \in X\big\},$$

则 $\big(C(X,\mathbb{R}^n), d_\infty\big)$ 是完备距离空间, 相应的 Arzelà-Ascoli 定理仍成立.

我们再来看几个相关的应用. 注意到在完备距离空间中, 压缩映射原理仍然成立.

例 7.6.5. *高阶线性常微分方程解的存在性.*

设 $a_0, a_1, \cdots, a_{n-1} \in C[a,b]$, 考虑如下方程

$$x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \cdots + a_1(t)x'(t) + a_0(t)x(t) = 0 \tag{7.8}$$

在初始条件 $x(t_0) = c_0$, $x'(t_0) = c_1$, $\cdots$, $x^{(n-1)}(t_0) = c_{n-1}$ 下解的存在性.

记 $X(t) = \big(x(t), x'(t), \cdots, x^{(n-1)}(t)\big)^\top$, 则上述方程可以转化为 $X'(t) = A(t)X(t)$, 其中 $A \in C\big([a,b], M_{n\times n}\big)$. 我们进一步将其转化为如下积分方程

$$X(t) = \int_{t_0}^t A(\tau)X(\tau)\,\mathrm{d}\tau + c^\top, \tag{7.9}$$

其中 $c = (c_0, c_1, \cdots, c_{n-1}) \in \mathbb{R}^n$, 右端的积分是指对每一个分量都关于 τ 积分. 我们通过找不动点来说明此积分方程有解. 为此, 定义映射 $\Phi : C\big([a,b], M_{n\times n}\big) \to C\big([a,b], M_{n\times n}\big)$ 如下: 当 $X \in C\big([a,b], M_{n\times n}\big)$ 时, 令

$$\Phi(X)(t) = \int_{t_0}^t A(\tau)X(\tau)\,\mathrm{d}\tau + c^\top, \ \ \forall\, t \in [a,b].$$

若 X 是映射 Φ 的不动点, 则 X 满足上述积分方程 (7.9), 进而满足微分方程.

为了说明不动点存在, 我们先来对 Φ 做一点估计. 记 $M = \max\{\|A(t)\| \mid x \in [a,b]\}$, 当 $X, Y \in C\big([a,b], M_{n\times n}\big)$ 时, 由 Φ 的定义可得

$$\|\Phi(X)(t) - \Phi(Y)(t)\| \le M\Big|\int_{t_0}^t \|X(\tau) - Y(\tau)\|\,\mathrm{d}\tau\Big|.$$

由此可得

$$\|\Phi(X)(t) - \Phi(Y)(t)\| \le M|t - t_0| d_\infty(X,Y).$$

当 i 为正整数时, Φ 与自身复合 i 次得到的映射记为 Φ^i. 于是有

$$\begin{aligned}\|\Phi^2(X)(t)-\Phi^2(Y)(t)\| &\le M\Big|\int_{t_0}^t \|\Phi(X)(\tau)-\Phi(Y)(\tau)\|\,\mathrm{d}\tau\Big|\\ &\le M^2\Big|\int_{t_0}^t |\tau-t_0| d_\infty(X,Y)\,\mathrm{d}\tau\Big|\\ &=\frac{1}{2}M^2(t-t_0)^2 d_\infty(X,Y).\end{aligned}$$

一般地, 利用归纳法容易得出

$$\|\Phi^i(X)(t)-\Phi^i(Y)(t)\|\le \frac{1}{i!}M^i|t-t_0|^i d_\infty(X,Y),$$

从而也有

$$d_\infty\big(\Phi^i(X),\Phi^i(Y)\big)\le \frac{1}{i!}M^i(b-a)^i d_\infty(X,Y).$$

注意到 $\lim\limits_{i\to\infty}\frac{1}{i!}M^i(b-a)^i=0$, 我们固定一个正整数 i, 使得 $M^i(b-a)^i<i!$, 则 Φ^i 为压缩映射, 它有唯一的不动点 X_*. 由 $\Phi^i\circ\Phi=\Phi^{i+1}=\Phi\circ\Phi^i$ 可知 $\Phi(X_*)$ 也是 Φ^i 的不动点, 由不动点的唯一性即知 $\Phi(X_*)=X_*$. 当然, 因为 Φ 的不动点也是 Φ^i 的不动点, Φ 也只有 X_* 这一个不动点.

例 7.6.6. *一阶常微分方程解的局部存在性.*

设 $\delta>0$, $r>0$, $x_0\in\mathbb{R}^n$, $f\in C\big([-\delta,\delta]\times\bar{B}_r(x_0),\ \mathbb{R}^n\big)$, 考虑如下微分方程

$$x'(t)=f\big(t,x(t)\big),\ \ x(0)=x_0. \tag{7.10}$$

与前例类似, 它可以转化为积分方程:

$$x(t)=\int_0^t f\big(\tau,x(\tau)\big)\,\mathrm{d}\tau+x_0. \tag{7.11}$$

不过, 因为 f 只满足连续性条件, 不方便利用压缩映射原理去求解这个积分方程. 下面我们使用 Tonelli 的技巧构造一列近似解, 然后用 Arzelà-Ascoli 定理导出一个真正的解.

为此, 记 $M=\max\{\|f(t,x)\|\mid (t,x)\in[-\delta,\delta]\times\bar{B}_r(x_0)\}$, 取 $\varepsilon=\min\big\{\delta,\ \frac{r}{M}\big\}$. 当 $i\ge 2$ 时, 定义

$$x_i(t)=\begin{cases}x_0, & t\in[0,\frac{\varepsilon}{i}),\\ x_0+\int_0^{t-\frac{\varepsilon}{i}} f\big(\tau,x_i(\tau)\big)\,\mathrm{d}\tau, & t\in[\frac{\varepsilon}{i},\varepsilon].\end{cases} \tag{7.12}$$

在某种意义上, x_i 是用迭代的方式定义的: 它在 $\big[\frac{\ell}{i}\varepsilon,\frac{\ell+1}{i}\varepsilon\big]$ 中的值由 $\big[0,\frac{\ell}{i}\varepsilon\big]$ 中的值所确定. 不难验证 $x_i\in C\big([0,\varepsilon],\ \mathbb{R}^n\big)$, 且

$$\|x_i(t)-x_0\|<r,\ \ \|x_i(t)-x_i(t')\|\le M|t-t'|,\ \ \forall\, t,t'\in[0,\varepsilon].$$

根据 Arzelà-Ascoli 定理, $\{x_i\}$ 在 $\big(C([0,\varepsilon],\,\mathbb{R}^n),\,d_\infty\big)$ 中有收敛子列, 设 $\{x_{i_j}\}$ 收敛于 x_*. 我们要说明 x_* 满足积分方程 (7.11). 根据定义, 有

$$x_{i_j}(t)=x_0+\int_0^t f\big(\tau,x_{i_j}(\tau)\big)\,\mathrm{d}\tau-\int_{t-\frac{\varepsilon}{i_j}}^t f\big(\tau,x_{i_j}(\tau)\big)\,\mathrm{d}\tau,$$

其中

$$\Big\|\int_{t-\frac{\varepsilon}{i_j}}^t f\big(\tau,x_{i_j}(\tau)\big)\,\mathrm{d}\tau\Big\|\le M\frac{\varepsilon}{i_j}.$$

令 $j\to\infty$, 利用 $\{x_{i_j}\}$ 收敛于 x_* 以及 f 一致连续可得

$$x_*(t)=x_0+\int_0^t f\big(\tau,x_*(\tau)\big)\,\mathrm{d}\tau,$$

这就说明 $x_*:[0,\varepsilon]\to\mathbb{R}^n$ 满足方程 (7.11), 自然也满足 (7.10).

注. 这个结果称为 Peano 存在性定理. 需要注意的是, 上述方程的解一般并不唯一. 比如, 除了恒为零的常值函数之外, $x(t)=\frac{1}{4}t^2$ $(t\ge 0)$ 也是 $x'(t)=\sqrt{|x(t)|}$ 在初始条件 $x(0)=0$ 下的解.

例 7.6.7. *无处可导连续函数的存在性.*

我们在 $C[0,1]$ 中取最大模距离, 此时它是完备距离空间. 记

$$C=\{f\in C[0,1]\,|\,f(0)=0,\ f(1)=1\},$$

易见 C 为 $C[0,1]$ 中的闭集. 定义映射 $F:C\to C$ 如下: 当 $f\in C$ 时, $F(f)$ 定义为

$$F(f)(x)=\begin{cases}\frac{3}{4}f(3x), & x\in[0,\ 1/3),\\ \frac{1}{4}+\frac{1}{2}f(2-3x), & x\in[1/3,\ 2/3),\\ \frac{1}{4}+\frac{3}{4}f(3x-2), & x\in[2/3,\ 1].\end{cases}$$

不难看出 F 为压缩映射, 因此存在唯一的不动点 $f_*\in C$.

断言: f_* 处处不可导. 我们分两步来证明这一点.

(i) 利用归纳法不难得出, 当 $k=1,2,\cdots,3^n$ 时, 有

$$\big|f_*(3^{-n}(k-1))-f_*(3^{-n}k)\big|\ge 2^{-n}. \tag{7.13}$$

(ii) 任取 $x_0\in[0,1]$. 当 $n\ge 1$ 时, 取 $k=\lfloor 3^n x_0\rfloor+1$, 则

$$3^{-n}(k-1)\le x_0<3^{-n}k.$$

由 (i) 可知,

$$\begin{aligned}2^{-n}&\le\big|f_*(3^{-n}(k-1))-f_*(3^{-n}k)\big|\\&\le\big|f_*(3^{-n}(k-1))-f_*(x_0)\big|+\big|f_*(x_0)-f_*(3^{-n}k)\big|\\&\le 2\max\big\{\big|f_*(3^{-n}(k-1))-f_*(x_0)\big|,\ \big|f_*(3^{-n}k)-f_*(x_0)\big|\big\}.\end{aligned}$$

于是可以选取 $x_n = 3^{-n}(k-1)$ 或 $3^{-n}k$, 使得

$$|f_*(x_n) - f_*(x_0)| \geq 2^{-n-1}.$$

此时, $0 < |x_n - x_0| \leq 3^{-n}$, 且当 $n \to \infty$ 时

$$\Big|\frac{f_*(x_n) - f_*(x_0)}{x_n - x_0}\Big| \geq \frac{2^{-n-1}}{3^{-n}} = \frac{1}{2}\Big(\frac{3}{2}\Big)^n \to +\infty,$$

因此 f_* 在 x_0 处不可导.

以上我们找到了一个处处连续但无处可导的函数. 实际上, 在连续函数空间中这种函数多得很. 为了说明这一点, 我们再引入两个结果.

定理 7.6.7 (闭集套原理)**.** *设 (X,d) 为完备度量空间, $\{C_i\}$ 为 X 中的一列非空闭集, 满足条件*

$$C_1 \supset C_2 \supset \cdots \supset C_i \supset \cdots, \quad \lim_{i\to\infty} d(C_i) = 0,$$

则 $\{C_i\}$ 有且仅有一个公共点.

证明 对每一 $i \geq 1$, 取 $u_i \in C_i$. 当 $j > i$ 时, 由 $C_j \subset C_i$ 可知 $u_j \in C_i$, 因此

$$d(u_j, u_i) \leq d(C_i), \ \ \forall\, j \geq i.$$

由题设即知 $\{u_i\}$ 为 Cauchy 列, 从而收敛, 其极限记为 u. 由 C_i 为闭集可知 $u \in C_i$, 这说明 u 是 $\{C_i\}$ 的公共点.

另一方面, 如果 u' 也是 $\{C_i\}$ 的公共点, 则由 $u, u' \in C_i$ 可知 $d(u', u) \leq d(C_i)$. 令 $i \to \infty$ 可得 $d(u', u) = 0$, 这说明 $u' = u$.

注. 闭集套原理可以用来刻画完备性: 如果闭集套原理对度量空间 (X,d) 成立, $\{u_i\}$ 为 X 中 Cauchy 列, 则可以找出它的一个收敛子列(因此它本身也是收敛的). 事实上, 由 Cauchy 列的定义可知, $\{u_i\}$ 存在子列 $\{u_{i_j}\}$, 使得

$$d(u_{i_j}, u_{i_k}) < 2^{-k}, \ \ \forall\, j \geq k \geq 1.$$

当 $j \geq 1$ 时, 记 $C_j = \{u \in X \mid d(u, u_{i_j}) \leq 2^{1-j}\}$, 则 $\{C_j\}$ 为一列闭集. 当 $u \in C_{j+1}$ 时,

$$d(u, u_{i_j}) \leq d(u, u_{i_{j+1}}) + d(u_{i_{j+1}}, u_{i_j}) < 2^{-j} + 2^{-j} = 2^{1-j},$$

这说明 $u \in C_j$, 即 $C_{j+1} \subset C_j$. 显然 $\{C_j\}$ 的直径趋于零. 若闭集套原理成立, 则 $\{C_j\}$ 有一个公共点, 此时 $\{u_{i_j}\}$ 收敛于该公共点.

定理 7.6.8 (Baire 纲定理)**.** *设 (X,d) 为完备度量空间, $\{C_i\}$ 为 X 中的一列闭集. 如果每一个 C_i 都没有内点, 则 $\{C_i\}$ 的并集也没有内点.*

证明 (反证法) 设 u_0 为 $\{C_i\}$ 的并集的内点, 则存在 $r_0 > 0$, 使得 $\bar{B}_{r_0}(u_0)$ 包含于 $\{C_i\}$ 的并集之中. 因为 C_1 是没有内点的闭集, 故存在 u_1 以及 r_1, 使得

$$0 < r_1 < 1,\ \ \bar{B}_{r_1}(u_1) \subset B_{r_0}(u_0) \setminus C_1.$$

同理, 由 C_2 没有内点且为闭集可知, 存在 u_2 以及 r_2, 使得

$$0 < r_2 < \frac{1}{2},\ \ \bar{B}_{r_2}(u_2) \subset B_{r_1}(u_1) \setminus C_2.$$

以次类推, 可以找到一列闭球 $\{\bar{B}_{r_i}(u_i)\}$, 使得

$$0 < r_i < \frac{1}{i},\ \ \bar{B}_{r_i}(u_i) \subset B_{r_{i-1}}(u_{i-1}) \setminus C_i.$$

根据闭集套原理, 这些闭球有一个公共点, 记为 u. 由上式可知 u 不属于任何一个 C_i, 但这与 $u \in B_{r_0}(u_0)$ 相矛盾.

例 7.6.8. 无处可导连续函数的丰富性.

我们在 $C[a,b]$ 中取最大模距离, 此时它是完备距离空间. 记

$$E = \{f \in C[a,b] \mid f \text{ 在某一点处可导}\}.$$

我们要说明 E 不含内点, 因此可以认为“大部分”连续函数都是处处不可导的.

当 $i \geq 1$ 时, 令

$$C_i = \{f \in C[a,b] \mid \text{存在 } x_0,\ \text{使得 } |f(x) - f(x_0)| \leq i|x - x_0|,\ \ \forall\, x \in [a,b]\}.$$

易见 C_i 为闭集, 且 E 包含于 $\{C_i\}$ 的并集之中. 我们来说明每一个 C_i 都不含内点, 于是由 Baire 纲定理即知 E 也不含内点.

为此, 任取 $g \in C_i$ 以及 $r > 0$, 我们要说明 $B_r(g)$ 中含有不属于 C_i 的连续函数, 其中

$$B_r(g) = \{f \in C[a,b] \mid |f(x) - g(x)| < r,\ \ \forall\, x \in [a,b]\}.$$

首先, 取分段线性函数 h, 使得 $|h - g| < \frac{r}{2}$ 在 $[a,b]$ 中处处成立. 其次, 当 $m \geq 1$ 时, 考虑连续函数 $h_m : \mathbb{R} \to \mathbb{R}$, 其中

$$h_m(x) = \frac{1}{m} d(m^2 x, \mathbb{Z}) = \frac{1}{m} \min\{|m^2 x - k| \mid k \in \mathbb{Z}\}.$$

h_m 是 $\mathbb{R}$ 中周期为 m^{-2} 的分段线性函数, 每一段的斜率为 $\pm m$. 于是当 m 充分大时 $h + h_m \in B_r(g) \setminus C_i$, 这说明 C_i 的确不含内点.

第八章　多元函数的微分

8.1　方向导数和微分

在微分的定义之前先引进相对无穷小:

设 $D \subset \mathbb{R}^n$ 为开集, $f, g : D \to \mathbb{R}^m$ 为向量值函数, $x^0 \in D$. 如果任给 $\varepsilon > 0$, 均存在 $\delta > 0$, 使得当 $x \in B_\delta(x^0) \subset D$ 时 $\|f(x)\| \le \varepsilon \|g(x)\|$, 则记 $f(x) = o\big(g(x)\big)$ $(x \to x^0)$.

定义 8.1.2(梯度和微分) 设 $D \subset \mathbb{R}^n$ 为开集, $f : D \to \mathbb{R}$ 为 n 元函数, $x^0 \in D$. 如果存在 $\alpha \in \mathbb{R}^n$, 使得

$$f(x^0 + h) = f(x^0) + \alpha \cdot h + o(\|h\|) \quad (h \to 0),$$

则称 f 在 x^0 处可微, α 称为 f 在 x^0 处的梯度, 记为 $\nabla f(x^0)$. $\mathbb{R}^n$ 上的线性函数 $u \mapsto \alpha \cdot u$ 称为 f 在 x^0 处的微分, 记为 $\mathrm{d}f(x^0)$.

命题 8.1.2(可微必可导)设 f 在 x^0 处可微, 则 f 在 x^0 处的方向导数都存在, 并且有

$$\frac{\partial f}{\partial u}(x^0) = \nabla f(x^0) \cdot u. \tag{8.1}$$

证明 设 u 为单位向量. 根据题设, 有

$$f(x^0 + tu) - f(x^0) = t\alpha \cdot u + o(|t|) \quad (t \to 0),$$

这说明 $\frac{\partial f}{\partial u}(x^0) = \alpha \cdot u$, 即 (8.1) 式成立.

注. 方向导数是函数沿某个方向的变化率. 从 (8.1) 式可以看出, 若 $\nabla f(x^0) \neq 0$, 则在 x^0 处 f 沿梯度方向增长最快.

当 f 在 x^0 处可微时, $\nabla f(x^0) \cdot e_i = \frac{\partial f}{\partial e_i}(x^0) = f_{x_i}(x^0)$, 即梯度的分量就是偏导数. 通常用行向量表示梯度:

$$\nabla f(x^0) = \big(f_{x_1}(x^0), \cdots, f_{x_n}(x^0)\big). \tag{8.2}$$

定理 8.1.4(求导次序的可交换性) 设 $f : D \to \mathbb{R}$ 为二元函数, $z^0 \in D$. 如果 f_{xy} 和 f_{yx} 在 z^0 处都存在, 且其中一个在 z^0 处连续, 则 $f_{xy}(z^0) = f_{yx}(z^0)$.

证明 不妨设 $z^0=(0,0)$, 且 f_{yx} 在 z^0 处连续. 由偏导数的定义可知

$$f_{xy}(0,0)=\lim_{y\to 0}\frac{1}{y}\big[f_x(0,y)-f_x(0,0)\big]=\lim_{y\to 0}\lim_{x\to 0}\frac{1}{yx}\big[\phi(x,y)-\phi(x,0)\big], \tag{8.3}$$

其中 $\phi(x,y)=f(x,y)-f(0,y)$. 利用一元函数的微分中值定理可得

$$\phi(x,y)-\phi(x,0)=y\phi_y(x,\theta y)=y\big[f_y(x,\theta y)-f_y(0,\theta y)\big]=yxf_{yx}(\theta' x,\theta y),$$

其中 $\theta,\theta'\in(0,1)$. 将上式代回 (8.3), 利用题设条件不难得出 $f_{xy}(0,0)=f_{yx}(0,0)$.

习题 8.1:

11. 设 $D\subset\mathbb{R}^2$ 为开集, $(x_0,y_0)\in D$. 证明: 若 $f:D\to\mathbb{R}$ 在 (x_0,y_0) 处连续, 且极限 $\lim\limits_{y\to y_0}\lim\limits_{x\to x_0}f(x,y)$ 存在, 则该极限必定等于 $f(x_0,y_0)$.

12. 设 $D\subset\mathbb{R}^n$ 为开集, 函数 $f:D\to\mathbb{R}$ 为 Lipschitz 函数. 如果 f 在 x^0 处的每一个方向导数都存在, 且存在 $\alpha\in\mathbb{R}^n$, 使得 $\frac{\partial f}{\partial u}(x^0)=\alpha\cdot u$, $\forall\ u\in S^{n-1}$. 证明: f 在 x^0 处可微.

8.2　切线和切面

删去 328 页“(见本章最后一节补充材料)”

习题 8.2:

3. 设 p 是平面椭圆 $\left\{\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\right\}$ 上的一点, 求 p 处的切线 ℓ. 若椭圆的焦点记为 f_1,f_2, 而 q 是 ℓ 上异于 p 的点, 证明

$$d(f_1,q)+d(f_2,q)>d(f_1,p)+d(f_2,p),$$

由此说明从一个焦点发出的光线经椭圆反射后一定到达另一个焦点.

4. 将上一题的结论推广到平面上的双曲线和抛物线上.

8.3　链式法则

定义 8.3.1 (微分)**.** *设 $D\subset\mathbb{R}^n$ 为开集, $f:D\to\mathbb{R}^m$ 为向量值的多元函数, $x^0\in D$. 如果存在线性映射 $L:\mathbb{R}^n\to\mathbb{R}^m$, 使得*

$$f(x^0+h)=f(x^0)+L(h)+o(\|h\|)\quad(h\to 0),$$

则称 f 在 x^0 处可微, L 称为 f 在 x^0 处的微分, 也记为 $\mathrm{d}f(x^0)$.

从定义不难看出, f 可微时它的每一个分量 f_i 均可微, 反之亦然. 此时有

$$f_i(x^0+h)=f_i(x^0)+\nabla f_i(x^0)\cdot h+o(\|h\|)\quad(h\to 0),$$

它们可以用列向量表示为

$$f(x^0+h)=f(x^0)+Jf(x^0)h+o(\|h\|)\ \ (h\to 0),$$

这说明 $df(x^0)(h)=Jf(x^0)h$, 其中

$$Jf(x^0)=\Big(\frac{\partial f_i}{\partial x_j}(x^0)\Big)_{m\times n}. \tag{8.4}$$

$Jf(x^0)$ 称为 f 在 x^0 处的 Jacobi 矩阵, $m=n$ 时其行列式称为 Jacobi 行列式. 利用三角不等式和矩阵的范数不等式还可以得到

$$\|f(x^0+h)-f(x^0)\|\le\big[\|Jf(x^0)\|+o(1)\big]\|h\|\quad(h\to 0). \tag{8.5}$$

我们可以利用微分研究向量值函数的复合求导.

定理 8.3.1 (复合求导)**.** *设 D 和 Δ 分别为 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 中的开集, $f:D\to\mathbb{R}^m$ 和 $g:\Delta\to\mathbb{R}^l$ 为向量值函数, 且 $f(D)\subset\Delta$. 如果 f 在 $x^0\in D$ 处可微, g 在 $y^0=f(x^0)$ 处可微, 则复合函数 $h=g\circ f$ 在 x^0 处可微, 且*

$$dh(x^0)=dg(y^0)\circ df(x^0),\ \ 或\ \ Jh(x^0)=Jg(y^0)\cdot Jf(x^0). \tag{8.6}$$

证明 记 $u=f(x^0+v)-f(x^0)$, 当 $v\to 0$ 时, 由 (8.5) 可知 $\|u\|\le C\|v\|$, 其中 C 是常数. 当 $v\to 0$ 时, 由题设条件可知

$$g(y^0+u)-g(y^0)=dg(y^0)(u)+o(1)\|u\|=dg(y^0)\circ df(x^0)(v)+dg(y^0)\big(o(\|v\|)\big)+o(1)\|v\|,$$

利用三角不等式和矩阵的范数不等式, 上式可以改写为

$$h(x^0+v)-h(x^0)=dg(y^0)\circ df(x^0)(v)+o(\|v\|)\quad(v\to 0),$$

这说明 $dh(x^0)=dg(y^0)\circ df(x^0)$.

例 8.3.1 源文件中注意 φ'.

8.4 拟微分中值定理

设 $D\subset\mathbb{R}^n$ 为开集, 如果 D 也是凸集, 则称 D 为凸域. 设 $f:D\to\mathbb{R}$ 为可微函数, 当 $x,y\in D$ 时, 我们想估计 $f(y)-f(x)$ 的大小. 记 $\phi(t)=f\big(x+t(y-x)\big)$, 由**推论 8.3.2**可得

$$\phi'(t)=\nabla f\big(x+t(y-x)\big)\cdot(y-x). \tag{8.7}$$

如果 f 为凸函数, 则 $\phi(t)$ 为一元凸函数, 因此有

$$f(y)=\phi(1)\ge\phi(0)+\phi'(0)=f(x)+\nabla f(x)\cdot(y-x). \tag{8.8}$$

如果 f 的梯度场有好的性质, 则对 $f(y)-f(x)$ 有更好的估计.

引理 8.4.1. 设 D 为 $\mathbb{R}^n$ 中的凸域, $f: D \to \mathbb{R}$ 为可微函数. 如果存在 $L \geq 0$, 使得

$$\|\nabla f(x) - \nabla f(y)\| \leq L\,\|x-y\|,\ \ \forall\, x, y \in D, \tag{8.9}$$

则

$$|f(y) - f(x) - \nabla f(x)\cdot(y-x)| \leq \frac{1}{2}L\|x-y\|^2,\ \ \forall\, x, y \in D. \tag{8.10}$$

证明 由 (8.7) 和 Newton-Leibniz 公式可得

$$f(y) - f(x) = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,\mathrm{d}t = \int_0^1 \nabla f\big(x + t(y-x)\big)\cdot(y-x)\,\mathrm{d}t.$$

利用 Cauchy-Schwarz 不等式和题设条件可得

$$\begin{aligned}|f(y) - f(x) - \nabla f(x)\cdot(y-x)| &= \Big|\int_0^1 \langle \nabla f\big(x+t(y-x)\big) - \nabla f(x),\ (y-x)\rangle\,\mathrm{d}t\Big| \\ &\leq \int_0^1 Lt\|x-y\|^2\,\mathrm{d}t = \frac{1}{2}L\|x-y\|^2.\end{aligned}$$

设 $f: \mathbb{R}^n \to \mathbb{R}$ 为可微凸函数. 若 $\nabla f(x^*) = 0$, 则由 (8.8) 可知 $f(x^*)$ 是 f 的最小值. 在实际应用中, 人们往往用梯度下降法去求近似的最小值. 其基本想法是, 若 x^i 是已知的近似最小值点, 则通过适当地选取 $h \in \mathbb{R}^n$, 使得 $x^{i+1} = x^i + h$ 为更好的近似最小值点. 为此, 设 f 满足条件 (8.9), 其中 $L > 0$, 则

$$0 \leq f(x^{i+1}) - f(x^i) - \nabla f(x^i)\cdot h \leq \frac{1}{2}L\|h\|^2.$$

取 $h = -\frac{1}{L}\nabla f(x^i)$, 则

$$-\frac{1}{L}\|\nabla f(x^i)\|^2 \leq f(x^{i+1}) - f(x^i) \leq -\frac{1}{2L}\|\nabla f(x_i)\|^2. \tag{8.11}$$

我们希望这样得到的一列近似最小值 $\{f(x^i)\}$ 最终收敛于真正的最小值 $f(x^*)$. 为此, 设 x^0 是最初的一个近似最小值点, 且存在 $M > 0$, 使得

$$\{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\} \subset \overline{B_M(x^*)}.$$

根据 (8.11) 可知 $f(x^i) \leq f(x^0)$ 总成立, 从而有 $\|x^i - x^*\| \leq M$. 若记 $R_i = f(x^i) - f(x^*)$, 则

$$0 \leq R_{i+1} \leq R_i - \frac{1}{2L}\|\nabla f(x^i)\|^2.$$

另一方面, 由 (8.8) 可得

$$0 \leq R_i \leq \nabla f(x^i)\cdot(x^i - x^*) \leq \|\nabla f(x^i)\|\,\|x^i - x^*\| \leq M\|\nabla f(x^i)\|,$$

即 $\|\nabla f(x^i)\| \geq \frac{R_i}{M}$. 于是我们有

$$0 \leq R_{i+1} \leq R_i - \frac{1}{2LM^2}R_i^2.$$

由此可以看出 $\{R_i\}$ 单调递减趋于零. 若 $R_{i+1} > 0$, 则还有

$$\frac{1}{R_{i+1}} \geq \frac{1}{R_i} + \frac{1}{2LM^2 - R_i} \geq \frac{1}{R_i} + \frac{1}{2LM^2},$$

由此再由归纳法可得

$$0 \leq R_i \leq \frac{R_0}{1+\beta i}, \text{ 其中 } \beta = \frac{R_0}{2LM^2}, \, i \geq 0.$$

这就给出了近似最小值的误差估计.

下面我们再看一元函数微分中值定理的多元推广.

定理 8.4.2 (微分中值定理)**.** *设 $D \subset \mathbb{R}^n$ 为凸域, $f : D \to \mathbb{R}$ 为可微函数. 则任给 $x, y \in D$, 存在 $\theta \in (0,1)$, 使得*

$$f(y) - f(x) = \nabla f(\xi) \cdot (y - x), \quad \text{其中 } \xi = (1-\theta)x + \theta y. \tag{8.12}$$

证明 利用 (8.7) 式, 对一元函数 $\phi(t) = f\big(x + t(y-x)\big)$ 运用微分中值定理即可.

习题 8.4

1. 设 D 为 $\mathbb{R}^n$ 中的区域, $f : D \to \mathbb{R}$ 为可微函数. 若 $\nabla f \equiv \alpha$, 则存在 $\beta \in \mathbb{R}$, 使得 $f(x) = \alpha \cdot x + \beta$.

2. 设 D 为 $\mathbb{R}^n$ 中的凸域, $f : D \to \mathbb{R}$ 为可微函数. 证明: f 为凸函数 $\Longleftrightarrow$ $f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$, $\forall\, x, y \in D$.

3. 设 D 为 $\mathbb{R}^n$ 中的凸域, $f : D \to \mathbb{R}$ 为可微函数. 证明: f 为凸函数 $\Longleftrightarrow$ $\langle \nabla f(y) - \nabla f(x),\, y - x \rangle \geq 0$, $\forall\, x, y \in D$.

4. 设 $f : \mathbb{R}^n \to \mathbb{R}^n$ 为可微映射. 当 $x \in \mathbb{R}^n$ 时, 矩阵 $Jf(x)(Jf)^{\top}(x)$ 的最大特征值记为 $\Lambda(x)$. 若 $\Lambda = \sup\limits_{x \in \mathbb{R}^n} \Lambda(x) < \infty$, 证明

$$\|f(x) - f(y)\| \leq \sqrt{\Lambda}\, \|x - y\|, \;\; \forall\, x, y \in \mathbb{R}^n.$$

5. 设 $D \subset \mathbb{R}^n$ 为开集, $f : D \to \mathbb{R}^m$ 处处可微且 $Jf(x)$ 关于 x 连续. 如果 $\overline{B_r(x^0)} \subset D$, 则任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $x, y \in \overline{B_r(x^0)}$ 且 $\|x - y\| < \delta$ 时

$$\|f(y) - f(x) - Jf(x)(y - x)\| \leq \varepsilon \|x - y\|.$$

6. 设 $D \subset \mathbb{R}^n$ 为凸域, $f : D \to \mathbb{R}^m$ 处处可微, 且存在常数 L, 使得

$$\|Jf(x) - Jf(y)\| \leq L\|x - y\|, \;\; \forall\, x, y \in D.$$

证明

$$\|f(y) - f(x) - Jf(x)(y - x)\| \leq \frac{1}{2}L\|x - y\|^2, \;\; \forall\, x, y \in D.$$

7. 设 $f:\mathbb{R}^n\to\mathbb{R}$ 为 n 元函数, 如果存在常数 $M\ge 0$, 使得 $g_\pm(x)=\frac{1}{2}M\|x\|^2\pm f(x)$ 均为凸函数, 证明: f 可微, 且 $\|\nabla f(x)-\nabla f(y)\|\le M\,\|x-y\|$. (提示: 利用凸函数的次梯度去说明 f 可微.)

8.5　逆映射定理和隐映射定理

习题 8.5:

1. 设 A 为 n 阶方阵, 记 $|A|_m=\max\{\|Ax\|\mid x\in S^{n-1}\}$. 证明: 若 $|A|_m<1$, 则 I_n-A 可逆, 且 $|(I_n-A)^{-1}|_m\le(1-|A|_m)^{-1}$.

8.6　多元函数的极值

例 8.6.1: 驻点坐标错了:

$$b=c^{-1}\Big(\sum_{i=1}^n y_i\sum_{i=1}^n x_i^2-\sum_{i=1}^n x_i\sum_{i=1}^n x_iy_i\Big),$$

P346 页: 删除凸函数的定义

命题 8.6.2 设 $D\subset\mathbb{R}^n$ 为凸域, $f:D\to\mathbb{R}$ 为 C^2 函数, 则 f 为凸函数 $\Longleftrightarrow$ $\nabla^2 f\ge 0$ (半正定).

证明 “$\Longleftarrow$” 任取 $x,y\in D$, 记 $\phi(t)=f\big(x+t(y-x)\big)$, 则

$$\phi''(t)=(y-x)^\top\nabla^2 f\big(x+t(y-x)\big)(y-x)\ge 0,$$

这说明 $\phi(t)$ 为一元凸函数. 于是 f 为 n 元凸函数.

“$\Longrightarrow$” 若 f 为 n 元凸函数, 则上述一元函数 $\phi(t)$ 是 C^2 的一元凸函数, 从而 $\phi''(0)=(y-x)^\top\nabla^2 f(x)(y-x)\ge 0$. 取 $r>0$, 使得 $B_r(x)\subset D$. 任给单位向量 $u\in\mathbb{R}^n$, 取 $y=x+\frac{r}{2}v\in B_r(x)$ 可得 $v^\top\nabla^2 f(x)v\ge 0$, 这说明 $\nabla^2 f(x)$ 是半正定的.

习题 8.6

5. 证明 $f(x,y)=\frac{x^2}{y}$ 在 $D=\{(x,y)\in\mathbb{R}^2\mid x\in\mathbb{R},y>0\}$ 中是凸函数.

6. 设 $D\subset\mathbb{R}^n$ 为凸域, $f:D\to\mathbb{R}$ 为 C^2 函数. 如果存在 $\lambda>0$, 使得 $\nabla^2 f\ge\lambda I_n$ 处处成立, 证明:

$$\|\nabla f(y)-\nabla f(x)\|\ge\lambda\,\|y-x\|,\ \ \forall\, x,y\in D.$$

8.7　Lagrange 乘数法

Lagrange 乘子法补充说明:

若函数 f 在曲面 S 上某点 x^0 处达到极值, 则 S 在 x^0 处与 f 的等值面相切: 如若不然, 在 x^0 附近总能找到 S 上的点 x^1, x^2, 使得 $f(x^1) < f(x^0)$, $f(x^2) > f(x^0)$.

习题 8.7:

10. 设 $A = \left(a_{ij}\right)_{n\times n}$ 为对称矩阵, 记 $Q(x) = \sum\limits_{i,j=1}^{n} a_{ij}x_ix_j$. 请用 Lagrange 乘数法解释 $Q(x)$ 在单位球面 S^{n-1} 的最值必为 A 的特征值.

8.8 多元函数微分的补充材料

8.8.1 最小二乘法回顾

我们先简单回顾平面上的直线拟合问题: 设 $\{(a_i,b_i)\}_{i=1}^m$ 是平面上的 m 个点, 我们希望找一条经过这 m 个点的直线 $y = \lambda x + \mu$, 即要解方程

$$\begin{pmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_m & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

这种方程可以推广: 设 $A \in M_{m\times n}$, $b \in \mathbb{R}^m$, 我们想在 $\mathbb{R}^n$ 中解方程 $Ax = b$. 我们知道此方程未必有解. 在实际问题中, 我们退而求其次, 希望能找"最佳"近似解. 为此, 将 A 视为从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射, 从线性代数可知, 方程 $Ax = b$ 有解当且仅当 $b \in \mathrm{Im}A = A(\mathbb{R}^n)$. 如果 $b \notin \mathrm{Im}A$, 则可以去寻找 $\mathrm{Im}A$ 中距离 b 最近的点 $A\hat{x}$, 此时 $\hat{x}$ 可以认为是"最佳"近似解. 这样就将问题转化成了求函数的最值.

定理 8.8.1 (最小二乘法)**.** *设 A, b 如上. 记 $f(x) = \|Ax - b\|^2$, 则 f 在 $\mathbb{R}^n$ 中存在最小值, 且 $\hat{x}$ 为最小值点当且仅当 $A^\top A\hat{x} = A^\top b$.*

证明 注意到 $\mathrm{Im}A$ 是 $\mathbb{R}^n$ 中的闭集, 若 b 在 $\mathrm{Im}A$ 上的正交投影为 $\hat{b}$, 则 $\|\hat{b}-b\|^2$ 就是 f 的最小值.

根据 Fermat 定理, 若 $\hat{x}$ 为 f 的最小值点, 则 $\nabla f(\hat{x}) = 0$. 简单的计算表明, $\nabla f(x) = 2(A^\top Ax - A^\top b)^\top$, 这说明 $A^\top A\hat{x} = A^\top b$.

反之, 若 $A^\top A\hat{x} = A^\top b$, 则当 $x \in \mathbb{R}^n$ 时 $(x-\hat{x})^\top A^\top(A\hat{x} - b) = 0$, 即 $A(x-\hat{x}) \perp (A\hat{x} - b)$, 这说明

$$f(x) = \|Ax - b\|^2 = \|A(x-\hat{x}) + (A\hat{x} - b)\|^2 = \|A(x-\hat{x})\|^2 + \|A\hat{x} - b\|^2,$$

因此 $\hat{x}$ 的确是 f 的最小值点.

注. 若 A 的秩为 n, 则 $A^\top A$ 可逆, 此时 f 的唯一最小值点为 $\hat{x} = (A^\top A)^{-1}A^\top b$. 读者可与**例 8.6.1**中的结果对照起来看.

上述定理证明中的基本思想还可以推广, 我们来看下面的例子.

例 8.8.1. 设 D 为 $\mathbb{R}^n$ 中的开集, $f: D \to \mathbb{R}^m$ 为可微映射. 如果 Jf 的秩恒为 m, 则 $f(D)$ 为 $\mathbb{R}^m$ 中的开集.

证明 设 $v_0 = f(u_0) \in f(D)$, 我们要说明 v_0 是 $f(D)$ 的内点. 也就是说, 我们需要证明当 $\|v - v_0\|$ 充分小时, 方程 $f(x) = v$ 有解. 基本的想法是去找目标函数 $\|f(x) - v\|^2$ 的最小值点. 为此, 我们先做一点准备工作.

记 $(Jf)^\top(u_0)Jf(u_0)$ 的最小特征值为 λ, 由题设可知 $\lambda > 0$. 此时

$$\|Jf(u_0)(x - u_0)\| \geq \sqrt{\lambda}\,\|x - u_0\|, \ \ \forall\, x \in \mathbb{R}^n.$$

于是, 由 f 在 u_0 处可微可知, 存在 $\delta > 0$, 使得 $\overline{B_\delta(u_0)} \subset D$, 且当 $x \in \overline{B_\delta(u_0)} \subset D$ 时, 有

$$\|f(x) - f(u_0)\| \geq \|Jf(u_0)(x - u_0)\| - \|o(\|x - u_0\|)\| \geq \frac{1}{2}\sqrt{\lambda}\,\|x - u_0\|.$$

特别地, 当 $\|x - u_0\| = \delta$ 时 $\|f(x) - f(u_0)\| \geq \frac{1}{2}\sqrt{\lambda}\,\delta$.

记 $r = \frac{1}{4}\sqrt{\lambda}\,\delta$. 任取 $v \in B_r(v_0)$, 考虑函数

$$g(x) = \|f(x) - v\|^2, \ \ x \in \overline{B_\delta(u_0)}.$$

设 g 在 u 处达到最小值, 我们断言 $u \in B_\delta(u_0)$. 事实上, 若 $\|u - u_0\| = \delta$, 则由三角不等式可知

$$\|f(u) - v\| \geq \|f(u) - f(u_0)\| - \|v_0 - v\| \geq 2r - \|v_0 - v\| > \|v_0 - v\|,$$

这说明 $g(u) > g(u_0)$, 从而导出了矛盾.

根据 Fermat 定理, $0 = \nabla g(u) = 2\big(f(u) - v\big)Jf(u)$, 由题设可知 $f(u) - v = 0$. 这说明 $B_r(v_0) \subset f\big(B_\delta(u_0)\big)$, 即 v_0 的确是 $f(D)$ 的内点.

8.8.2 隐映射定理回顾

隐映射定理是多元函数微分学的中重要结果. 在前面我们先证明了逆映射定理, 再用它推出隐映射定理. 也可以反过来做, 下面我们介绍隐映射定理的一种直接证明方法.

定理 8.8.2 (隐映射定理)**.** 设 W 为 $\mathbb{R}^{n+m}$ 中的开集, W 中的点用 (x, y) 表示, 其中 $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_m)$. 设 $f: W \to \mathbb{R}^m$ 为 C^1 映射, 用分量表示为

$$f(x, y) = \big(f_1(x, y), \cdots, f_m(x, y)\big).$$

如果 $f(x^0, y^0) = 0$ 且 $J_y f(x^0, y^0)$ 非退化, 则存在 x^0 的开邻域 $V \subset \mathbb{R}^n$ 以及 C^1 映射 $\psi: V \to \mathbb{R}^m$, 使得

$$y^0 = \psi(x^0), \ \ f(x, \psi(x)) \equiv 0,$$

且 $J_x\psi(x) = -\big[J_y f(x, \psi(x))\big]^{-1} J_x f(x, \psi(x))$.

证明 不妨设 $(x^0, y^0) = (0,0)$. 根据已知条件, 存在 $\mu > 0$, 使得

$$\|J_y f(0,0)v\| \geq \mu\|v\|,\ \forall\, v \in \mathbb{R}^m.$$

由 $f \in C^1$ 可知, 存在 $C > 0$, $\delta > 0$, 使得当 $\|x\| \leq \delta$, $\|y\| \leq \delta$ 时

$$\|J_x f(x,y)\| \leq C,\ \ \|J_y f(x,y) - J_y f(0,0)\| \leq \frac{1}{2}\mu. \tag{8.13}$$

此时, 任给 $v \in \mathbb{R}^m$, 有

$$\begin{aligned}\|J_y f(x,y)v\| &\geq \|J_y f(0,0)v\| - \|[J_y f(x,y) - J_y f(0,0)]v\|\\ &\geq \mu\|v\| - \frac{1}{2}\mu\|v\| = \frac{1}{2}\mu\|v\|,\end{aligned}$$

特别地, $J_y f(x,y)$ 非退化.

由 (8.13) 和拟微分中值定理还可以得出

$$\|f(\bar{x},y) - f(x,y)\| \leq C\|\bar{x} - x\|,\ \ \forall\, \|\bar{x}\|, \|x\|, \|y\| \leq \delta. \tag{8.14}$$

我们还有

$$\|f(x,\bar{y}) - f(x,y)\| \geq \frac{1}{2}\mu\|\bar{y} - y\|,\ \ \forall\, \|x\|, \|\bar{y}\|, \|y\| \leq \delta. \tag{8.15}$$

事实上, 记 $R(y) = f(x,y) - J_y f(0,0)y$, 则 $J_y R = J_y f(x,y) - J_y(0,0)$ 由 (8.13) 和拟微分中值定理可知

$$\|R(\bar{y}) - R(y)\| \leq \frac{1}{2}\mu\|\bar{y} - y\|.$$

于是

$$\begin{aligned}\|f(x,\bar{y}) - f(x,y)\| &= \|J_y f(0,0)(\bar{y} - y) + R(\bar{y}) - R(y)\|\\ &\geq \|J_y f(0,0)(\bar{y} - y)\| - \|R(\bar{y}) - R(y)\|\\ &\geq \frac{1}{2}\mu\|\bar{y} - y\|.\end{aligned}$$

有了上述准备工作, 我们对给定的 x 去寻找适当的 y, 使得 $f(x,y) = 0$. 为此, 考虑函数 $E(y) = \|f(x,y)\|^2$, 我们来研究当 $\|y\| \leq \delta$ 时 $E(y)$ 的最小值. 当 $\|y\| = \delta$ 时,

$$\begin{aligned}\sqrt{E(y)} &= \|[f(x,y) - f(x,0)] + [f(x,0) - f(0,0)]\|\\ &\geq \|[f(x,y) - f(x,0)]\| - \|[f(x,0) - f(0,0)]\|\\ &\geq \frac{1}{2}\mu\|y\| - C\|x\|.\end{aligned}$$

而 $y = 0$ 时

$$\sqrt{E(0)} = \|f(x,0)\| = \|f(x,0) - f(0,0)\| \leq C\|x\|.$$

这说明, 当给定的 x 满足 $\|x\| < \frac{\mu\delta}{4C}$ 时, $E(y)$ 的最小值一定在内部达到. 在最小值点处,

$$0 = \nabla E = 2f(x,y)J_y f(x,y).$$

由 $J_y f(x,y)$ 非退化可知 $f(x,y)=0$. 由 (8.15) 可知这样的 y 是由 x 所唯一确定的, 记为 $y=\psi(x)$.

我们先来说明 ψ 关于 x 连续. 事实上, 设 $\|\bar{x}\|, \|x\| < \frac{\mu\delta}{4C}$, 由 (8.15) 可知

$$\|f(\bar{x},\psi(x))\| = \|f(\bar{x},\psi(\bar{x})) - f(\bar{x},\psi(x))\| \geq \frac{1}{2}\mu\|\psi(\bar{x})-\psi(x)\|.$$

再由 (8.14) 可得

$$\|f(\bar{x},\psi(x))\| = \|f(x,\psi(x)) - f(\bar{x},\psi(x))\| \leq C\|\bar{x}-x\|.$$

于是有

$$\|\psi(\bar{x})-\psi(x)\| \leq \frac{2C}{\mu}\|\bar{x}-x\|. \tag{8.16}$$

最后我们来看 ψ 关于 x 的可微性. 由 f 可微可知, 当 $\bar{x}\to x$ 时

$$\begin{aligned}
0 &= f(\bar{x},\psi(\bar{x})) - f(x,\psi(x)) \\
&= J_x f(x,\psi(x))(\bar{x}-x) + J_y f(x,\psi(x))\big(\psi(\bar{x})-\psi(x)\big) \\
&\quad + o\big(\|(\bar{x}-x, \psi(\bar{x})-\psi(x))\|\big) \\
&= J_x f(x,\psi(x))(\bar{x}-x) + J_y f(x,\psi(x))\big(\psi(\bar{x})-\psi(x)\big) \\
&\quad + o\big(\|\bar{x}-x\|\big).
\end{aligned}$$

这说明

$$\psi(\bar{x})-\psi(x) = -\big[J_y f(x,\psi(x))\big]^{-1} J_x f(x,\psi(x))(\bar{x}-x) + o\big(\|\bar{x}-x\|\big).$$

即 ψ 可微且 $J_x\psi(x) = -\big[J_y f(x,\psi(x))\big]^{-1} J_x f(x,\psi(x))$. 由此还可以看出, 若 $f(x,y)\in C^k$ $(k\geq 1)$, 则 $\psi\in C^k$.

8.8.3 函数的相关性和独立性

第九章　多元函数的积分

9.1　二重 Riemann 积分

删去矩形的直径、面积定义. 面积记号换成 ν.

Darboux 定理前引入内缩矩形:

设 $I = [a,b] \times [c,d]$, δ 是充分小的正数, 记

$$I^\delta = (a+\delta,\ b-\delta) \times (c+\delta, d-\delta),$$

显然 $0 < \nu(I) - \nu(I^\delta) < 2(b-a)\delta + 2(d-c)\delta$.

定理 9.1.3 证明修改:

任给 $\varepsilon > 0$, 存在分割 $\pi' = \{I_{ij}\}$, 使得 $S(\pi') < S(f) + \frac{\varepsilon}{2}$. 设 δ 是充分小的正数, 记 $J_\delta = I \setminus \bigcup\limits_{ij} I_{ij}^\delta$, 我们要求 $\nu(J_\delta) < \varepsilon(2M+1)^{-1}$.

将原零测集的定义和性质删去.

定理 9.1.8(Lebesgue) 设 f 为矩形 I 中定义的有界函数, 则 f 可积 "$\Longleftrightarrow$" f 的间断点集 D_f 为零测集.

证明 "$\Longrightarrow$" 只要证明对每一个 $\eta > 0$, D_η 均为零测集即可. 任给 $\varepsilon > 0$, 根据 Riemann 定理, 可取 I 的分割 $\{I_{ij}\}$, 使得振幅满足 $\omega_{ij} \geq \eta$ 的小矩形的面积之和小于 $\frac{\varepsilon}{2}$.

设 $\delta > 0$, 记 $J_\delta = I \setminus \bigcup\limits_{ij} I_{ij}^\delta$, 我们要求 $\nu(J_\delta) < \frac{\varepsilon}{2}$. 注意到当 $I_{ij}^\delta \cap D_\eta \neq \emptyset$ 时, $\omega_{ij} \geq \eta$. 这说明 D_η 包含在 J_δ 和 $\{I_{ij} \mid I_{ij}^\delta \cap D_\eta \neq \emptyset\}$ 的并集之中, 且它们的面积之和小于 ε. 由零测集的定义即知 D_η 为零测集.

"$\Longleftarrow$" 设 D_f 为零测集. 任给 $\varepsilon, \eta > 0$, 存在至多可数个开矩形 $\{I_i\}$, 使得

$$D_f \subset \bigcup_{i\geq 1} I_i,\ \ \sum_{i\geq 1} \nu(I_i) < \varepsilon.$$

任取 $x \in I \setminus D_f$, 由 f 在 x 处连续可知, 存在 $r > 0$, 使得 f 在 $B_r(x) \cap I$ 中的振幅小于 η. 显然, $\{B_r(x) \mid x \in I \setminus D_f\}$ 和 $\{I_i\}$ 形成了 I 的开覆盖, 设 λ 是它的 Lebesgue 数. 设 $\{I_{ij}\}$ 是 I 的分割, 当它的模小于 λ 时, 每一个小矩形要么完全包含于某一个 $B_r(x)$ 中, 此时 f 在此小矩形中的振幅小于 η; 要么小矩形完全包

含于某一个 I_i 中. 这说明

$$\sum_{\{(ij)|\omega_{ij}\geq\eta\}}\nu(I_{ij})\leq\sum_{i\geq 1}\nu(I_i)<\varepsilon,$$

由 Riemann 定理即知 f 可积.

习题 9.1:

删去原 9,10,11

9.2　多重积分及其性质

删去矩形的直径和体积.

习题 9.2:

删去原 1, 2.

增加:

11. 设 f 是 n 维矩形 I 中定义的函数, 则 f 可积当且仅当任给 $\varepsilon>0$, 均存在连续函数 g,h, 使得

$$g\leq f\leq h,\quad \int_I(h-g)\,\mathrm{d}x<\varepsilon.$$

9.3　重积分的计算

9.4　重积分的变量替换

与一元函数的积分类似, 多元函数的积分也有相应的变量替换公式. 我们先考虑仿射变换, 再利用微分学的基本手法对一般的变换做线性化并估计误差.

9.4.1　仿射变换

设 $\varphi(x)=Px$ 为线性变换, I 为 $\mathbb{R}^n$ 中的 n 维矩形. 若 P 退化, 则 $\varphi(\mathbb{R}^n)$ 是 $\mathbb{R}^n$ 中的零测集, 此时 $\varphi(I)$ 的容积为零. 下设 P 非退化, 则由 §7.4 中的讨论可知 $\partial\varphi(I)=\varphi(\partial I)$ 为零测集, 这说明 $\varphi(I)$ 可求容积.

根据线性代数, 非退化线性变换 $\varphi(x)=Px$ 可以分解为有限个如下初等线性变换的复合:

(1) $(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n)\mapsto(x_1,\cdots,x_j,\cdots,x_i,\cdots,x_n)$, 其中 $i<j$.

(2) $(x_1,\cdots,x_i,\cdots,x_n)\mapsto(x_1,\cdots,\lambda x_i,\cdots,x_n)$, 其中 $1\leq i\leq n$, $\lambda\neq 0$.

(3) $(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n)\mapsto(x_1,\cdots,x_i,\cdots,x_j+x_i,\cdots,x_n)$, 其中 $i<j$.

我们先来验证在这三种初等变换下, 均有 $\nu(\varphi(I)) = |\det P|\nu(I)$. 情形 (1) 和 (2) 比较显然, 我们来看情形 (3), 以 $n=2$ 为例. 记 $I=[a,b]\times[c,d]$, $(u,v) = \varphi(x,y) = (x, y+x)$, 则 $\varphi(I)$ 可以表示为

$$\varphi(I) = \{(u,v)\in\mathbb{R}^2 \mid a\le u\le b,\ c+u\le v\le d+u\},$$

这是一个平行四边形, 其容积可以用重积分化累次积分计算:

$$\nu(\varphi(I)) = \int_a^b \mathrm{d}u \int_{c+u}^{d+u} \mathrm{d}v = \int_a^b \mathrm{d}u \int_c^d \mathrm{d}v = \nu(I).$$

一般地, 设 f 是某个 n 维矩形中的可积函数, 在此矩形外规定 f 为零. f 在矩形中的积分可以写成它在整个 $\mathbb{R}^n$ 中的积分. 我们有

引理 9.4.1. *设 f 是某个 n 维矩形中的可积函数, P 是非退化 n 阶方阵, 则*

$$\int_{\mathbb{R}^n} f(Px)\,\mathrm{d}x = |\det P|^{-1}\int_{\mathbb{R}^n} f(x)\,\mathrm{d}x. \tag{9.1}$$

证明 如果 $P=A,B$ 时, 公式 (9.1) 成立, 则由 $\det(BA) = (\det B)(\det A)$ 可知公式 (9.1) 对 $P=BA$ 也成立. 因此, 只要对初等变换验证引理成立即可.

取矩形 I, 使得 f 在 I 之外为零, 设 $\pi=\{I_{ij}\}$ 为 I 的分割, f 在 I_{ij} 中的下确界和上确界分别记为 m_{ij} 和 M_{ij}, 则

$$m_{ij}\,\nu\big(P^{-1}I_{ij}\big) \le \int_{P^{-1}I_{ij}} f(Px)\,\mathrm{d}x \le M_{ij}\,\nu\big(P^{-1}I_{ij}\big).$$

上式关于 i,j 求和, 利用之前的讨论可得

$$|\det P|^{-1}\sum_{ij} m_{ij}\,\nu(I_{ij}) \le \int_{\mathbb{R}^n} f(Px)\,\mathrm{d}x \le |\det P|^{-1}\sum_{ij} M_{ij}\,\nu(I_{ij}),$$

令 π 的模趋于零即得欲证结论.

推论 9.4.2. *设 $\varphi(x) = Px+b$ 为仿射变换, S 为 $\mathbb{R}^n$ 中的可求容积集, 则 $\nu\big(\varphi(S)\big) = |\det P|\,\nu(S)$.*

证明 不妨设 $b=0$ 且 $\det P\ne 0$, 此时 $\varphi(S)$ 可求容积. 在前一引理中, 将 f 取为 $\varphi(S)$ 的特征函数即得欲证结论.

例 9.4.1. *设 $\{v_i\}_{i=1}^n$ 为 $\mathbb{R}^n$ 中的向量, 求如下平行多面体 的容积:*

$$P(v_1,\cdots,v_n) = \Big\{\sum_{i=1}^n x_i v_i \in\mathbb{R}^n \,\Big|\, 0\le x_i\le 1,\ i=1,2,\cdots,n\Big\}.$$

解 记 $\mathbb{R}^n$ 的标准正交基为 $\{e_i\}_{i=1}^n$, 考虑线性变换 $\varphi:\mathbb{R}^n\to\mathbb{R}^n$, $\varphi(e_i)=v_i$. 则显然有

$$\begin{aligned}P(v_1,\cdots,v_n)&=\Big\{\sum_{i=1}^n x_iv_i\in\mathbb{R}^n\,\Big|\,0\le x_i\le 1,\ i=1,2,\cdots,n.\Big\}\\&=\Big\{\varphi\Big(\sum_{i=1}^n x_ie_i\Big)\in\mathbb{R}^n\,\Big|\,0\le x_i\le 1,\ i=1,2,\cdots,n.\Big\}\\&=\varphi(P(e_1,\cdots,e_n)).\end{aligned}$$

显然, $P(e_1,\cdots,e_n)=[0,1]^n$, 因此, 由上述推论可得

$$\nu(P(v_1,\cdots,v_n))=|\det P|\nu([0,1]^n)=|\det P|,$$

其中 P 是以 $v_1,v_2,\cdots,v_n$ 为列向量的方阵. 例如, 当 $n=2$, $v_1=(a_1,b_1)$, $v_2=(a_2,b_2)$ 时,

$$\nu(P(v_1,v_2))=|a_1b_2-a_2b_1|,$$

这是平行四边形的面积公式.

例 9.4.2. *设 $a>0$, 求 n 维单形 $\Delta_n(a)$ 容积, 其中*

$$\Delta_n(a)=\{x\in\mathbb{R}^n\,|\,x_1\ge 0,\cdots,x_n\ge 0,\ x_1+\cdots+x_n\le a\}.$$

解 这个例子我们在前节算过, 现在用递推的办法再算一次. 利用伸缩变换 $\varphi(x)=ax$ 易见 $\nu\big(\Delta_n(a)\big)=\nu\big(\Delta_n(1)\big)a^n$. 记 $\nu\big(\Delta_n(1)\big)=v_n$, 我们只要求出 v_n 就可以了. 利用投影法可得

$$v_{n+1}=\int_0^1\nu\big(\Delta_n(1-x_1)\big)\,\mathrm{d}x_1=v_n\int_0^1(1-x_1)^n\,\mathrm{d}x_1=\frac{1}{n+1}v_n.$$

由 $v_1=1$ 以及上述递推式即得 $v_n=\frac{1}{n!}$, 因此 $\nu\big(\Delta_n(a)\big)=\frac{1}{n!}a^n$, 这和上节的计算结果一致.

例 9.4.3. *设 $\{v_i\}_{i=1}^n$ 为 $\mathbb{R}^n$ 中的向量, 求如下单形的容积:*

$$\Delta(v_1,\cdots,v_n)=\Big\{\sum_{i=1}^n x_iv_i\in\mathbb{R}^n\,\Big|\,x_i\ge 0\ (1\le i\le n),\ x_1+\cdots+x_n\le 1\Big\}.$$

解 考虑线性变换 $\varphi:\mathbb{R}^n\to\mathbb{R}^n$, $\varphi(e_i)=v_i$, 则

$$\Delta(v_1,\cdots,v_n)=\varphi\big(\Delta_n(1)\big),$$

由之前的结果可得

$$\nu\big(\Delta(v_1,\cdots,v_n)\big)=|\det P|\nu\big(\Delta_n(1)\big)=\frac{1}{n!}|\det P|,$$

其中 P 是以 $v_1,v_2,\cdots,v_n$ 为列向量的方阵.

例 9.4.4. *求中心在 $x^0 \in \mathbb{R}^n$, 半径为 r 的 n 维球体 $\overline{B_r(x^0)}$ 的容积, 其中*

$$\overline{B_r(x^0)} = \{x \in \mathbb{R}^n \mid (x - x^0) \cdot (x - x^0) \le r^2\}.$$

解 根据体积的平移不变性, 不妨设 x_0 为原点. 半径为 r 的 n 维球体的容积记为 $\omega_n(r)$. 利用伸缩变换可知 $\omega_n(r) = \omega_n(1)r^n$, 为了方便起见, 记 $\omega_n = \omega_n(1)$. 利用投影法可得

$$\begin{aligned}\omega_{n+1} &= \int_{-1}^{1} \omega_n\big(\sqrt{1 - x_1^2}\big)\,\mathrm{d}x_1 = \omega_n \int_{-1}^{1} (1 - x_1^2)^{\frac{n}{2}}\,\mathrm{d}x_1 \\ &= 2\omega_n \int_0^{\frac{\pi}{2}} \sin^{n+1} t\,\mathrm{d}t = 2\omega_n J_{n+1}.\end{aligned}$$

其中, J_{n+1} 在例 **??** 中已经计算过:

$$J_{2k} = \frac{(2k-1)!!}{(2k)!!}\frac{\pi}{2},\ \ J_{2k+1} = \frac{(2k)!!}{(2k+1)!!},\ \ \forall\, k \ge 0.$$

根据 $\omega_1 = 2$ 以及上述递推公式可得

$$\omega_1 = 2,\ \ \omega_2 = \pi,\ \ \omega_{n+2} = \omega_n \frac{2\pi}{n+2},\ \ n \ge 1.$$

由此进一步得到

$$\omega_{2k} = \frac{(2\pi)^k}{(2k)!!} = \frac{\pi^k}{k!},\ \ \omega_{2k-1} = \frac{2^k \pi^{k-1}}{(2k-1)!!},\ \ k \ge 1. \tag{9.2}$$

我们有如下简单估计:

$$2^n n^{-\frac{n}{2}} \le \omega_n \le 2^n. \tag{9.3}$$

事实上, 矩形 $[-r, r]^n$ 的外接球半径为 $\sqrt{n}r$, 因此 $(2r)^n \le \omega_n(\sqrt{n}r)^n$; 同理, 矩形 $[-r, r]^n$ 的内接球半径为 r, 因此 $\omega_n r^n \le (2r)^n$.

9.4.2 一般的变量替换

设 D 为 $\mathbb{R}^n$ 中的开集, $\varphi : D \to \mathbb{R}^n$ 为 C^1 映射. 根据拟微分中值定理, φ 为局部 Lipschitz 映射, 因此将零测集映为零测集. 假设 $J\varphi$ 处处非退化, 则由逆映射定理可知 φ 将内点映为内点. 这说明如果 S 可求容积且 $\bar{S} \subset D$, 则 $\varphi(S)$ 也可求容积. 为了研究 $\varphi(S)$ 的容积, 我们将 φ 线性化并做误差估计.

取 $\delta > 0$, 使得 $K = \{x \mid d(x, S) \le \delta\} \subset D$. 当 $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ 时, 记 $|x|_m = \max\{|x_1|, |x_2|, \cdots, |x_n|\}$. 若 $u \in \mathbb{R}^n$, $r > 0$, 则以 u 为中心, $2r$ 为边长的方体可以表示为 $I_r(u) = \{x \in \mathbb{R}^n \mid |x - u|_m < r\}$.

引理 9.4.3. *沿用以上记号, 则任给 $\varepsilon > 0$, 存在 $0 < \eta < \frac{\delta}{\sqrt{n}}$, 使得当 $0 < r < \eta$, $u \in S$ 时*

$$\nu\big(\varphi(I_r(u))\big) \le \big[|\det J\varphi(u)| + O(\varepsilon)\big]\nu(I_r(u)).$$

证明 记 $C=\max_K\|[J\varphi]^{-1}\|$. 由 φ 为 C^1 映射可知, 任给 $\varepsilon>0$, 存在 $0<\eta<\frac{\delta}{\sqrt{n}}$, 使得当 $x',x''\in K$ 且 $|x'-x''|_m<\eta$ 时 $\|J\varphi(x')-J\varphi(x'')\|<\varepsilon$.

记 $F(x)=\varphi(x)-\varphi(u)-J\varphi(u)(x-u)$. 当 $0<r<\eta$ 时 $I_r(u)\subset K$. 当 $x\in I_r(u)$ 时, 由拟微分中值定理可得

$$\begin{aligned}\|F(x)\|&=\|F(x)-F(u)\|\le\|JF(\xi)\|\,\|x-u\|\\&=\|J\varphi(\xi)-J\varphi(u)\|\,\|x-u\|<\varepsilon\sqrt{n}r.\end{aligned}$$

考虑仿射变换 $L(y)=\big[J\varphi(u)\big]^{-1}\big(y-\varphi(u)\big)+u$, 则

$$L\circ\varphi(x)=x+\big[J\varphi(u)\big]^{-1}F(x).$$

于是当 $x\in I_r(u)$ 时, 有

$$|L\circ\varphi(x)-u|_m\le|x-u|_m+\big\|\big[J\varphi(u)\big]^{-1}F(x)\big\|\le r+C\varepsilon\sqrt{n}r.$$

这说明

$$\nu\big(L\circ\varphi(I_r(u))\big)\le(1+C\varepsilon\sqrt{n})^n\nu(I_r(u)),$$

再由推论 9.4.2 即得欲证结论.

定理 9.4.4 (重积分的变量替换)**.** *设 $\varphi:D\to\mathbb{R}^n$ 为 C^1 单射, 且 $J\varphi$ 处处非退化. 设 S 可求容积, $\bar{S}\subset D$, f 在 $\varphi(S)$ 中可积, 则*

$$\int_{\varphi(S)}f(y)\,\mathrm{d}y=\int_S f\circ\varphi(x)\,|\det J\varphi(x)|\,\mathrm{d}x.\tag{9.4}$$

特别地,

$$\nu(\varphi(S))=\int_S|\det J\varphi(x)|\,\mathrm{d}x.$$

证明 不妨设 $f\ge0$, 且在 $\varphi(S)$ 外规定 f 为零. 取 $\delta>0$, 使得 $\{x\mid d(x,S)\le\delta\}\subset D$. 用一个大的立方体包含住 S, 将大立方体等分为小方体, 使得每一个小方体的直径小于 δ. 此时, 与 S 有非空交的小方体都包含在 D 中, 我们只要在每一个小方体中验证 (9.4) 式成立就可以了. 因此, 不妨设一开始就设 S 为方体. 任给 S 的等分分割 $\{I_{ij}\}$, 我们有

$$\int_{\varphi(S)}f(y)\,\mathrm{d}y=\sum_{ij}\int_{\varphi(I_{ij})}f(y)\,\mathrm{d}y\le\sum_{ij}\big[\sup_{\varphi(I_{ij})}f\big]\,\nu(\varphi(I_{ij})).$$

当分割的模充分小时, 由前一引理可得

$$\begin{aligned}\int_{\varphi(S)}f(y)\,\mathrm{d}y&\le\sum_{ij}\big[\sup_{I_{ij}}f\circ\varphi\big]|\det J\varphi(\xi_{ij})|\nu(I_{ij})+O(\varepsilon)\\&=\int_S f\circ\varphi(x)\,|\det J\varphi(x)|\,\mathrm{d}x+O(\varepsilon).\end{aligned}$$

令 $\varepsilon \to 0^+$ 可得

$$\int_{\varphi(S)} f(y)\,\mathrm{d}y \le \int_S f\circ\varphi(x)\,|\det J\varphi(x)|\,\mathrm{d}x.$$

根据逆映射定理, $\varphi : D \to \varphi(D)$ 可逆. 对 φ^{-1} 重复上述论证就可得到反过来的不等式, 于是欲证结论成立.

例 9.4.5

例 9.4.6

9.4.3 极坐标变换

习题 9.4:

1. 设 $r > 0$, 利用线性变换求如下椭球的容积:

$$E_r = \{(x, y, z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 + xy + yz + zx \le r^2\}.$$

2. 设 P 为 n 阶正定对称实方阵, 求 $E(P) = \{x \in \mathbb{R}^n \,|\, xPx^\top \le 1\}$ 的容积.

删去 3.

9.5 重积分的应用和推广

第十章　曲线积分与曲面积分

10.1　第一型曲线积分

10.2　第二型曲线积分

10.3　第一型曲面积分

“面积” “体积”改为“容积”？
例 10.3.5 中“第十二章”改为“前一章”

10.4　几类积分之间的联系

10.5　附录: Riemann-Stieltjes 积分

第十一章　微分形式的积分

11.1　欧氏空间中的微分形式

P449 页6行: “第十一章”改为“第八章”

P454 页倒数11-12行: 源文件上是对的, 编辑改错了!

Σ 的面积 $\sigma(\Sigma)$ 定义为:

$$\sigma(\Sigma) = \int_{\Sigma} d\sigma = \int_{D} \sqrt{\det[(J\psi)^{\top} J\psi]}\, \mathrm{d}u_1 \cdots \mathrm{d}u_{n-1}.$$

11.2　微分形式之间的运算

11.3　Gauss-Green 公式

所谓单位分解, 就是将 1 分解为若干个具有紧支集的函数之和. 其中, 函数 f 的支集 $\operatorname{supp} f$ 定义为

$$\operatorname{supp} f = \overline{\{x \mid f(x) \neq 0\}}.$$

我们在 $\mathbb{R}$ 中定义偶函数 $\phi(t)$ 如下:

$$\phi(t) = 1,\ \ t \in [0, 0.5];\ \ \phi(t) = \frac{(1-t)^3}{(t-0.5)^3 + (1-t)^3},\ \ 0.5 < t \leq 1;\ \ \phi(t) = 0,\ \ t \geq 1.$$

当 $t < 0$ 时令 $\phi(t) = \phi(-t)$. 容易验证 ϕ 是 C^2 函数, 称为 $\mathbb{R}$ 中的鼓包函数.

将单位分解引理中的“光滑”换成“C^2”

定理 11.3.5 最后一行: 应为 B_R 中的积分.

11.4　不动点定理和毛球定理

引理 11.4.1. 设 $\psi : D \to \mathbb{R}^n$ 为连续的向量值函数, 且当 $x \in S^{n-1} = \partial D$ 时 $\psi(x) = x$, 则任给 $\varepsilon > 0$, 存在 C^2 的向量值函数 $\rho : \mathbb{R}^n \to \mathbb{R}^n$, 使得

$$\rho(x) = x,\ \ \forall\, \|x\| \geq 1;\ \ \|\rho(x) - \psi(x)\| < \varepsilon,\ \ \forall\, x \in D.$$

证明. 记 $f(x)=\psi(x)-x$, 则 $f\big|_{S^{n-1}}\equiv 0$. 任给 ε, 由 f 在 D 中一致连续可知, 存在 $\delta>0$, 使得只要 $\|x-y\|\le\delta$, 就有 $\|f(x)-f(y)\|<\frac{\varepsilon}{2}$. 令

$$g(x)=\begin{cases} f\big((1+\delta)x\big), & \|x\|\le\frac{1}{1+\delta},\\ 0, & \|x\|>\frac{1}{1+\delta}.\end{cases}$$

则当 $x\in D$ 时 $\|g(x)-f(x)\|<\frac{\varepsilon}{2}$; 当 $\|x-y\|\le\eta=\frac{\delta}{1+\delta}$ 时, $\|g(x)-g(y)\|<\frac{\varepsilon}{2}$.

根据前节单位分解引理的证明, 存在有限点集 $\{x^i\}_{i=1}^k\subset\overline{B_{1-\eta}(0)}$, 以及相应的 C^2 函数 $\{\phi_i\}$, 使得 $\operatorname{supp}\phi_i\subset B_\eta(x^i)$, 且

$$0\le\sum_{i=1}^k\phi_i\le 1;\ \ \sum_{i=1}^k\phi_i(x)=1,\ \ \forall\, x\in\overline{B_{1-\eta}(0)}.$$

当 $x\in D$ 时, 令 $h(x)=\sum\limits_{i=1}^k g(x^i)\phi_i(x)$, 则当 $\|x\|\ge 1$ 时 $h(x)=0$, 且当 $x\in D$ 时

$$h(x)-g(x)=\sum_{i=1}^k\big[g(x^i)-g(x)\big]\phi_i(x).$$

于是

$$\|h(x)-g(x)\|\le\sum_{i=1}^k\|g(x^i)-g(x)\|\,\phi_i(x).$$

上式右端只需要考虑 $\phi_i(x)>0$ 的项, 此时 $x\in B_\eta(x^i)$, 因此 $\|g(x^i)-g(x)\|<\frac{\varepsilon}{2}$, 这表明 $\|h(x)-g(x)\|<\frac{\varepsilon}{2}$. 记 $\rho(x)=x+h(x)$, 则 ρ 是满足要求的 C^2 映射.

本节中的“光滑”一般要改为 C^2.

习题增加:

1. 设 C 为 $\mathbb{R}^n$ 中的有界闭凸集, $f:C\to C$ 为连续映射. 证明 f 必有不动点.

第十二章　数项级数

12.1　级数的收敛与发散

习题 12.1:

删掉 15, 原 16 变为 15

16. 设 $-1 < \alpha < 0$, 证明级数 $\sum\limits_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$ 收敛.

12.2　正项级数的敛散性

判别法做一些调整:

在比较判别法中取 $b_n = q^n$, 可以得到两个特殊的判别法.

定理 12.2.2[Cauchy 判别法] 设 $\sum\limits_{n=1}^{\infty} a_n$ 为正项级数. 如果存在常数 $q \in [0,1)$, $N \geq 1$, 使得当 $n > N$ 时, $\sqrt[n]{a_n} \leq q$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 收敛; 如果存在无穷多个 n, 使得 $\sqrt[n]{a_n} \geq 1$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

证明 若 $\sqrt[n]{a_n} \leq q$, 则 $a_n \leq q^n$. 当 $0 \leq q < 1$ 时, 几何级数 $\sum\limits_{n=1}^{\infty} q^n$ 收敛. 于是由比较判别法即知 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

若 $\sqrt[n]{a_n} \geq 1$, 则 $a_n \geq 1$. 如果有无限项 a_n 大于或等于 1, 则级数 $\sum\limits_{n=1}^{\infty} a_n$ 的通项不趋于零, 从而它是发散级数.

Cauchy 判别法也称为根值判别法, 其极限形式有时也很有用.

定理 12.2.3[Cauchy 判别法的极限形式] 设 $\sum\limits_{n=1}^{\infty} a_n$ 为正项级数, $\varlimsup\limits_{n\to\infty} \sqrt[n]{a_n} = \lambda$, 则 $\lambda < 1$ 时, $\sum\limits_{n=1}^{\infty} a_n$ 收敛; $\lambda > 1$ 时, 级数发散.

证明 若 $\lambda < 1$, 取 $q \in (\lambda, 1)$. 由上极限的定义可知, 存在 $N \geq 1$, 使得当 $n > N$ 时 $\sqrt[n]{a_n} < q$. 由前一定理即知 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

若 $\lambda > 1$, 则有无限项 a_n 大于或等于 1, 此时 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

定理 12.2.4[d'Alembert 判别法] 设 $\sum\limits_{n=1}^{\infty} a_n$ 为正项级数. 如果存在常数 $q \in$

$(0,1)$, $N \geq 1$, 使得当 $n \geq N$ 时, $\frac{a_{n+1}}{a_n} \leq q$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 收敛; 如果 $n \geq N$ 时 $\frac{a_{n+1}}{a_n} \geq 1$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

证明 若 $\frac{a_{n+1}}{a_n} \leq q$, 则 $\frac{a_{n+1}}{q^{n+1}} \leq \frac{a_n}{q^n}$, 这说明当 $n \geq N$ 时 $\left\{\frac{a_n}{q^n}\right\}$ 单调递减, 此时 $\frac{a_n}{q^n} \leq \frac{a_N}{q^N}$. 由比较判别法和几何级数的收敛性即知 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

若 $\frac{a_{n+1}}{a_n} \geq 1$, 则 $a_{n+1} \geq a_n$, 这说明 a_n 不趋于零, 从而 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

d'Alembert 判别法也称为比值判别法, 其极限形式为: 设 $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = \lambda$, 则 $\lambda < 1$ 时级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, $\lambda > 1$ 时 $\sum\limits_{n=1}^{\infty} a_n$ 发散. 其证明留给读者.

需要指出的是, 在根值判别法或比值判别法的极限形式中, 极限 $\lambda = 1$ 时是无法直接判断级数是否收敛或发散的. 这主要是因为用来比较的几何级数收敛的速度比较“快”. 如果与收敛得比较“慢”的 $p-$级数进行比较, 则可以得出精细一些的判别法. 比如我们希望当 n 充分大时 $a_n \leq M\frac{1}{n^p}$, 即希望 $\{n^p a_n\}$ 有上界. 若 $\{n^p a_n\}$ 单调递减, 则它当然有上界, 此时要求

$$\frac{a_n}{a_{n+1}} \geq \Big(1+\frac{1}{n}\Big)^p = 1 + \frac{p}{n} + O\Big(\frac{1}{n^2}\Big), \ \ (n \to \infty). \tag{12.1}$$

受此启发, 我们有

定理 12.2.5[Raabe 判别法] 设 $\sum\limits_{n=1}^{\infty} a_n$ 为正项级数. 如果存在常数 $q > 1$, $N \geq 1$, 使得当 $n \geq N$ 时, $n\big(\frac{a_n}{a_{n+1}} - 1\big) \geq q$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 收敛; 如果 $n \geq N$ 时 $n\big(\frac{a_n}{a_{n+1}} - 1\big) \leq 1$, 则 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

证明 若 $q > 1$, 取 $p \in (1, q)$, 由 (12.1) 式右端可知, 当 n 充分大时

$$\frac{a_n}{a_{n+1}} \geq 1 + \frac{q}{n} > \Big(1+\frac{1}{n}\Big)^p,$$

这说明 $n^p a_n > (n+1)^p a_{n+1}$. 由 $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ 收敛以及比较比较判别法即知 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

若 $n\big(\frac{a_n}{a_{n+1}} - 1\big) \leq 1$, 则 $na_n \leq (n+1)a_{n+1}$. 由调和级数 $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ 发散以及比较判别法即知 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

Raabe 判别法的极限形式为: 设 $\lim\limits_{n\to\infty} n\big(\frac{a_n}{a_{n+1}} - 1\big) = \lambda$, 则 $\lambda > 1$ 时级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, $\lambda < 1$ 时 $\sum\limits_{n=1}^{\infty} a_n$ 发散. 其证明留给读者.

注意到条件 $\lim\limits_{n\to\infty} n\big(\frac{a_n}{a_{n+1}} - 1\big) = \lambda$ 也可以改写为

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\lambda}{n} + o\Big(\frac{1}{n}\Big), \ \ (n \to \infty).$$

当 $\lambda = 1$ 时, Raabe 判别法的极限形式也是失效的. 不过, 如果将用于比较的 $p-$级数换成收敛得更慢的级数, 则可以得出更精细的判别法.

例 12.2.5 研究级数 $\sum\limits_{n=2}^{\infty} \frac{1}{n\ln n}$ 的敛散性.

解 由微分中值定理可知, 当 $k > 1$ 时, 存在 $\xi \in (k, k+1)$, 使得

$$\ln\ln(k+1) - \ln\ln k = \frac{1}{\xi\ln\xi} < \frac{1}{k\ln k},$$

这说明当 $n \to \infty$ 时

$$\sum_{k=2}^{n} \frac{1}{k\ln k} > \sum_{k=2}^{n} \big[\ln\ln(k+1) - \ln\ln k\big] = \ln\ln(n+1) - \ln\ln 2 \to +\infty,$$

因此 $\sum\limits_{n=2}^{\infty} \frac{1}{n\ln n}$ 发散.

定理 12.2.6[Gauss 判别法] 设 $\sum\limits_{n=1}^{\infty} a_n$ 为正项级数. 如果存在常数 λ, 使得

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\lambda}{n} + o\Big(\frac{1}{n\ln n}\Big), \ \ (n \to \infty).$$

则当 $\lambda > 1$ 时级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, $\lambda \le 1$ 时 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

证明 $\lambda > 1$ 或 $\lambda < 1$ 的情形可应用 Raabe 判别法的极限形式, 下设 $\lambda = 1$. 注意到

$$\frac{(n+1)\ln(n+1)}{n\ln n} - \Big(1 + \frac{1}{n}\Big) = \frac{(n+1)\ln\big(1+\frac{1}{n}\big)}{n\ln n} > \frac{1}{n\ln n},$$

由题设可知, 当 n 充分大时,

$$\frac{(n+1)\ln(n+1)}{n\ln n} > \frac{a_n}{a_{n+1}},$$

即 $\big[(n+1)\ln(n+1)\big]a_{n+1} > \big[n\ln n\big]a_n$. 由前例和比较判别即知级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

从以上讨论可以看到, 用于比较的级数收敛得越“慢”, 得出的判别法就越精细. 那么是否存在收敛得“最慢”的级数呢? 下面的例子给出了否定的回答.

例 12.2.6 设正项级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 则存在另一收敛的正项级数 $\sum\limits_{n=1}^{\infty} b_n$, 使得 $\frac{a_n}{b_n} \to +\infty$ $(n \to \infty)$.

证明 当 $n \ge 1$ 时, 记 $r_n = \sum\limits_{k=n}^{\infty} a_k$, 则 $r_n \to 0$ $(n \to \infty)$. 注意到

$$a_n = r_n - r_{n+1} = \big(\sqrt{r_n} - \sqrt{r_{n+1}}\big)\big(\sqrt{r_n} + \sqrt{r_{n+1}}\big),$$

若记 $b_n = \sqrt{r_n} - \sqrt{r_{n+1}}$, 则 $\sum\limits_{n=1}^{\infty} b_n$ 收敛，且

$$\frac{a_n}{b_n} = \frac{1}{\sqrt{r_n} + \sqrt{r_{n+1}}} \to +\infty, \ \ (n \to \infty).$$

从此例可知, 我们不能奢望找到某种终极的比较判别法. 不过, 如果能灵活地选取 b_n, 而并不使用某种固定的 b_n，则比较判别法仍有值得发挥的空间. 比如, 我们可以换一种观点重新证明 Raabe 判别法的前半部分: 若 $n\big(\frac{a_n}{a_{n+1}}-1\big)\geq q$, 则

$$na_n-(n+1)a_{n+1}\geq (q-1)a_{n+1}.$$

当 $q>1$ 时, 从上式可知

$$\sum_{k=N+1}^{n} a_k \leq \frac{1}{q-1}\sum_{k=N+1}^{n}\big[(k-1)a_{k-1}-ka_k\big]=\frac{1}{q-1}\big(Na_N-na_n\big)\leq \frac{1}{q-1}Na_N,$$

这说明 $\sum\limits_{n=1}^{\infty} a_n$ 的部分和有上界, 从而收敛. 一般地, 我们有

定理 12.2.7[Kummer] 设 $\sum\limits_{n=1}^{\infty} a_n$, $\sum\limits_{n=1}^{\infty} b_n$ 为正项级数. 如果存在常数 λ, $N\geq 1$, 使得 $n\geq N$ 大时

$$\frac{1}{b_n}\frac{a_n}{a_{n+1}}-\frac{1}{b_{n+1}}\geq \lambda, \tag{12.2}$$

则 $\sum\limits_{n=1}^{\infty} a_n$ 收敛; 如果 $n\geq N$ 时

$$\frac{1}{b_n}\frac{a_n}{a_{n+1}}-\frac{1}{b_{n+1}}\leq 0, \tag{12.3}$$

且 $\sum\limits_{n=1}^{\infty} b_n$ 发散, 则 $\sum\limits_{n=1}^{\infty} a_n$ 也发散.

证明 条件 (12.2) 可改写为

$$a_{n+1}\leq \frac{1}{\lambda}\Big(\frac{a_n}{b_n}-\frac{a_{n+1}}{b_{n+1}}\Big).$$

这说明当 $n\geq N$ 时

$$\sum_{k=N+1}^{n} a_k\leq \frac{1}{\lambda}\Big(\frac{a_N}{b_N}-\frac{a_n}{b_n}\Big)\leq \frac{1}{\lambda}\frac{a_N}{b_N},$$

这说明 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

条件 (12.3) 说明 $\frac{a_n}{b_n}\leq\frac{a_{n+1}}{b_{n+1}}$, 即 $\big\{\frac{a_n}{b_n}\big\}$ 关于 n 单调递增, 从而 $a_n\geq\frac{a_N}{b_N}b_n$, 由比较判别法立得欲证结论.

习题 12.2:

习题 13 换掉:

13. 设 $\alpha>0$, 证明级数 $\sum\limits_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$ 绝对收敛.

12.3 无穷乘积

习题 12.3

6. 本题给出 (12.6) 式的证明步骤, 请补充细节.

(1) 设 $\beta \in \left(0, \frac{\pi}{2}\right)$, $|\alpha| \le \beta$, 则

$$\frac{\alpha^2}{\beta^2} \le \frac{\sin^2 \alpha}{\sin^2 \beta} \le \frac{\alpha^2}{\beta^2} + \alpha^2.$$

(2) 设 $0 \le \lambda_l \le 1$, $0 \le h \le \lambda_l$, $l = 1, 2, \cdots, n$, 请用广义 Bernoulli 不等式或归纳法证明

$$\prod_{l=1}^{n} (\lambda_l - h) \ge \prod_{l=1}^{n} \lambda_l - nh.$$

(3) 设 $|x| \le \frac{\pi}{2}$, $n \ge 1$, 则

$$\prod_{l=1}^{n} \left[1 - \left(\frac{x}{l\pi}\right)^2\right] - n\left(\frac{x}{2n+1}\right)^2 \le \prod_{l=1}^{n} \left[1 - \frac{\sin^2 \frac{x}{2n+1}}{\sin^2 \frac{l\pi}{2n+1}}\right] \le \prod_{l=1}^{n} \left[1 - \left(\frac{x}{l\pi}\right)^2\right],$$

由此说明当 $|x| \le \frac{\pi}{2}$ 时 (12.6) 式成立.

(4) 记 $S_n(x) = x \prod\limits_{l=1}^{n} \left[1 - \left(\frac{x}{l\pi}\right)^2\right]$, 则当 $n > \frac{x}{\pi}$ 时, $S_n(x+\pi) = -S_n(x) \frac{(n+1)\pi + x}{n\pi - x}$, 由此说明 (12.6) 式对任何 x 都成立.

8. (1) 用 (12.6) 式证明

$$\cos \frac{\pi}{4} x + \sin \frac{\pi}{4} x = (1+x) \prod_{l=1}^{\infty} \left(1 + \frac{x}{4l+1}\right)\left(1 - \frac{x}{4l-1}\right).$$

(2) 用上式证明如下 Euler 的等式:

$$\frac{\pi^3}{32} = 1 + \sum_{l=1}^{\infty} \left[\frac{1}{(4l+1)^3} - \frac{1}{(4l-1)^3}\right] = 1 - \frac{1}{3^3} + \frac{1}{5^3} - \frac{1}{7^3} + \frac{1}{9^3} + \cdots.$$

9. 设 $a_n > 0$ 且 $\prod\limits_{n=1}^{\infty} a_n$ 收敛. 如果 $b_m = o\left(\frac{1}{m}\right)$ $(m \to \infty)$, 则

$$\lim_{m \to \infty} \prod_{n=1}^{m} (a_n + b_m) = \prod_{n=1}^{\infty} a_n.$$

10. 验证 (12.8) 式成立.

原8变为11, 原9变为12.

12.4 数项级数的进一步讨论

12.4.1 级数的乘积

12.4.2 Abel 求和与 Cesàro 求和

1911 年, Littlewood 将 Tauber 第一定理中的条件改进为 $a_n = O(1/n)$. 1912 年, Hardy-Littlewood 提出猜测: 只要 $\{na_n\}$ 单边有界(比如有上界), 则 Tauber

第一定理中的结论仍然成立. 这个猜测被他们在 1914 年证实是对的. 1930 年, Karamata 简化了他们的证明. 1952 年, Wielandt 又进一步将证明打磨得十分精巧. 下面我们就来介绍这个证明.

定理 12.4.8[Hardy-Littlewood, 1914] 设 $\sum\limits_{n=0}^{\infty} a_n = S\ (A)$, 如果 $\{na_n\}$ 有上界, 则 $\sum\limits_{n=0}^{\infty} a_n$ 收敛于 S.

证明 不妨设 $S=0$. 当 $k\geq 1$ 时, 由题设可得 $\lim\limits_{x\to 1^-}\sum\limits_{n=0}^{\infty} a_n x^{kn}=0$. 这说明, 当 $p(t)$ 是常数项为零的多项式时, 总有

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty} a_n p(x^n)=0. \tag{12.4}$$

另一方面, 当 $k\geq 0$ 时,

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} x^n x^{kn}=\lim_{x\to 1^-}\frac{1-x}{1-x^{k+1}}=\frac{1}{k+1}=\int_0^1 t^k\,\mathrm{d}t,$$

因此当 $p(t)$ 为多项式时,

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} x^n p(x^n)=\int_0^1 p(t)\,\mathrm{d}t. \tag{12.5}$$

记 $\chi(t)$ 为 $\left[\mathrm{e}^{-1},\ 1\right]\subset[0,1]$ 的特征函数, 即当 $0\leq t<\mathrm{e}^{-1}$ 时 $\chi(t)=0$; 当 $\mathrm{e}^{-1}\leq t\leq 1$ 时 $\chi(t)=1$. 当 $N\geq 1$, $x=\mathrm{e}^{-\frac{1}{N}}$ 时,

$$\sum_{n=0}^{\infty} a_n\chi(x^n)=\sum_{n\leq N} a_n\chi(x^n)=\sum_{n=0}^{N} a_n.$$

记 $\phi(t)=\chi(t)-t$, 则

$$\sum_{n=0}^{N} a_n-\sum_{n=0}^{\infty} a_n x^n=\sum_{n=0}^{\infty} a_n\phi(x^n)=\sum_{n=0}^{\infty} a_n(1-x^n)x^n\psi(x^n), \tag{12.6}$$

其中

$$\psi(t)=\frac{\phi(t)}{t(1-t)}=\begin{cases}\frac{1}{t-1}, & 0\leq t<\mathrm{e}^{-1},\\ \frac{1}{t}, & \mathrm{e}^{-1}\leq t\leq 1.\end{cases}$$

显然, $\psi\in R[0,1]$. 根据 §6.2 习题 15, 任给 $\varepsilon>0$, 存在多项式 P,Q, 使得

$$P\leq\psi\leq Q,\quad \int_0^1 [Q(t)-P(t)]\,\mathrm{d}t<\varepsilon.$$

记 $\bar{P}(t)=t+t(1-t)P(t)$, $\bar{Q}(t)=t+t(1-t)Q(t)$, 由 (12.6) 式可得

$$\begin{aligned}\sum_{n=0}^{N} a_n&=\sum_{n=0}^{\infty} a_n\bar{P}(x^n)+\sum_{n=0}^{\infty} a_n(1-x^n)x^n\big[\psi-P\big](x^n)\\&=\sum_{n=0}^{\infty} a_n\bar{Q}(x^n)+\sum_{n=0}^{\infty} a_n(1-x^n)x^n\big[\psi-Q\big](x^n).\end{aligned}$$

根据题设, 存在常数 $C>0$, 使得 $na_n\le C$. 此时 $a_n(1-x^n)\le C(1-x)$. 结合 $\psi-P\le Q-P$ 和 $\psi-Q\ge P-Q$ 可得

$$\sum_{n=0}^{\infty}a_n\bar{Q}(x^n)-C(1-x)\sum_{n=0}^{\infty}x^n\big[Q-P\big](x^n)\le\sum_{n=0}^{N}a_n\le\sum_{n=0}^{\infty}a_n\bar{P}(x^n)+C(1-x)\sum_{n=0}^{\infty}x^n\big[Q-P\big](x^n).$$

令 $N\to\infty$, 则 $x\to 1^-$, 由 (12.4) 式和 (12.5) 式可得

$$-C\varepsilon<-C\int_0^1\big[Q(t)-P(t)\big]\,\mathrm{d}t\le\varliminf_{N\to\infty}\sum_{n=0}^{N}a_n\le\varlimsup_{N\to\infty}\sum_{n=0}^{N}a_n\le C\int_0^1\big[Q(t)-P(t)\big]\,\mathrm{d}t<C\varepsilon.$$

由 ε 的任意性即知 $\sum\limits_{n=0}^{\infty}a_n=0$.

类似的方法可以得到

定理 12.4.9[Hardy-Littlewood, 1914] 设 $\lim\limits_{x\to 1^{-1}}(1-x)\sum\limits_{n=0}^{\infty}a_nx^n=S$, 如果 $\{a_n\}$ 有上界或下界, 则 $\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{k=0}^{n}a_k=S$.

证明 当 $k\ge 0$ 时,

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty}a_nx^nx^{kn}=\lim_{x\to 1^-}\frac{1-x}{1-x^{k+1}}\big(1-x^{k+1}\big)\sum_{n=0}^{\infty}a_n\big(x^{k+1}\big)^n=\frac{S}{k+1}=S\int_0^1t^k\,\mathrm{d}t.$$

这说明当 $p(t)$ 为多项式时,

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty}a_nx^np(x^n)=S\int_0^1p(t)\,\mathrm{d}t.$$

不妨设 $a_n\ge 0$. 利用 §6.2 习题 15 就可以知道, 任给 $f\in R[0,1]$, 均有

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty}a_nx^nf(x^n)=S\int_0^1f(t)\,\mathrm{d}t.$$

特别地, 取 f 如下:

$$f(t)=\begin{cases}0, & 0\le t<\mathrm{e}^{-1},\\ t^{-1}, & \mathrm{e}^{-1}\le t\le 1,\end{cases}$$

则当 $n\ge 1$, $x=\mathrm{e}^{-\frac{1}{n}}$ 时, 有

$$(1-x)\sum_{k=0}^{\infty}a_kx^kf(x^k)=\big(1-\mathrm{e}^{-\frac{1}{n}}\big)\sum_{k=0}^{n}a_k.$$

令 $n\to\infty$ 可得

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}a_k=\lim_{n\to\infty}\big(1-\mathrm{e}^{-\frac{1}{n}}\big)\sum_{k=0}^{n}a_k=S\int_0^1f(t)\,\mathrm{d}t=S.$$

12.4.3 级数的重排

226 页行4 例 12.4.4: $\frac{1}{2}$ 应为 $1-\frac{1}{2}$

12.4.4 级数求和与求极限的可交换性

第十三章　函数项级数

13.1　一致收敛

习题 13.1:

1. 当 $t \in [0,1]$ 时, 定义一列多项式 $p_n(t)$: $p_0 \equiv 0$, $n \geq 0$ 时 $p_{n+1}(t) = \frac{1}{2}\big(t + p_n^2(t)\big)$. 证明 $\{p_n(t)\}$ 在 $[0,1]$ 中一致收敛于 $1-\sqrt{1-t}$.

2. 利用上一题的结论证明: 任给 $\varepsilon > 0$, 存在多项式 $p(x)$, 使得

$$\big|p(x) - |x|\big| < \varepsilon, \ \forall\, x \in [-1,1].$$

13.2　求和与求导、积分的可交换性

P228 页倒数3行(源文件中的是对的):

此时, 对 $f_m - f_n$ 应用微分中值定理可得

P240 页, 6行, $[0,\frac{\pi}{2}]$ 应改为 $(0,\pi)$; 9行, “闭区间”应改为“区间”.

习题 13.2:

10. 通过在 $x = \frac{\pi}{4}$ 处对 (13.5) 式不断求导证明

$$\frac{5\pi^5}{1536} = 1 + \sum_{n=1}^{\infty}\Big[\frac{1}{(4n+1)^5} - \frac{1}{(4n-1)^5}\Big] = 1 - \frac{1}{3^5} + \frac{1}{5^5} - \frac{1}{7^5} + \frac{1}{9^5} + \cdots.$$

原10变11, 11变12.

13.3　幂级数

13.3.2 Taylor 展开与幂级数

Bernstein 定理中, 区间 $[a,b]$, (a,b) 统一换成 $[a,b)$.

p274 页 $\sin^{-2}x - x^{-2}$ 例子: 行 6 中 $(-\pi,\pi)$ 换成 $[0,\pi)$

增加 $\tan x$ 的例子

习题 13.3:

9: 去掉“用上一题”

11. 证明: 当 $\alpha > 0$ 时 (13.13) 式对 $x \in [-1,1]$ 成立; 当 $-1 < \alpha < 0$ 时, (13.13) 式对 $x \in (-1,1]$ 成立.

15: Legendre 多项式定义中 $(2n)!$ 应改为 $(2n)!!$.

13.4　函数项级数的进一步讨论

p263 页倒数5行，第2个减号应改为加号

第十四章 Fourier 分析

14.1 Fourier 级数

我们知道, 许多物理现象可以用简谐运动的方程 $u''(x) = -\omega^2 u(x)$ 来描述, 此方程的通解为 $u(x) = C_1 \cos\omega x + C_2 \sin\omega x = A\sin(\omega x + \phi)$, 其中 A 称为振幅, ω 称为频率, ϕ 称为初始相位. 简谐运动是周期性的运动, 周期为 $T = \frac{2\pi}{\omega}$. 当 ω 为整数时, 可得函数列

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos nx,\ \sin nx,\ \cdots$$

称为三角函数系, 它们当中任何两个不同函数的乘积在 $[-\pi, \pi]$ 中的积分都等于零, 这个性质称为三角函数系的正交性.

14.2 Fourier 级数的收敛性

14.3 Parseval 恒等式

设 u 是欧氏空间 $\mathbb{R}^n$ 中的向量, 它在标准正交基 $\{e_k\}_{k=1}^n$ 下可以分解为 $u = \sum\limits_{k=1}^n u_k e_k$, 其中 $u_k = \langle u, e_k\rangle$. 此时

$$\langle u, u\rangle = \sum_{k,l=1}^n u_k u_l \langle e_k, e_l\rangle = \sum_{k=1}^n u_k^2 = \sum_{k=1}^n \langle u, e_k\rangle^2.$$

这是勾股定理的高维推广, 它还可以推广到无穷维空间中. 为此, 设 $f, g \in R[-\pi, \pi]$, 定义

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x.$$

Fourier 展开 $f(x) = \frac{a_0}{2} + \sum\limits_{n=1}^\infty (a_n \cos nx + b_n \sin nx)$ 可以看成是函数 f 在三角函数系下的正交分解, 类比于欧氏空间中的勾股定理, 应该有

$$\langle f, f\rangle = \frac{a_0^2}{4}\langle 1, 1\rangle + \sum_{n=1}^\infty \left[a_n^2 \langle \cos nx, \cos nx\rangle + b_n^2 \langle \sin nx, \sin nx\rangle\right] = \frac{1}{2}a_0^2 + \sum_{n=1}^\infty (a_n^2 + b_n^2),$$

它称为 Parseval 恒等式, 下面我们来讨论它的严格证明.

14.4 Fourier 级数的进一步讨论

14.4.2: Hardy 引理的证明改为参照 §2.3 例 2.3.2 及其注记.

14.4.4 增加应用:

作为应用, 考虑 $[0,1]$ 中的复值函数 $f(x)=\mathrm{e}^{2\pi\mathrm{i}x^2}$. 注意到 $f(0)=f(1)=1$, $f(x)$ 可延拓为周期为 1 的周期函数, 作 Fourier 展开可得

$$f(x)=\sum_{n\in\mathbb{Z}}c_n\mathrm{e}^{2\pi\mathrm{i}nx},\ \text{ 其中 }\ c_n=\int_0^1\mathrm{e}^{2\pi\mathrm{i}(x^2-nx)}\,\mathrm{d}x.$$

在上式中取 $x=0$ 可得

$$1=\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\frac{\pi}{2}\mathrm{i}n^2}\int_0^1\mathrm{e}^{2\pi\mathrm{i}(x-\frac{n}{2})^2}\,\mathrm{d}x=\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\frac{\pi}{2}\mathrm{i}n^2}\int_{-\frac{n}{2}}^{-\frac{n}{2}+1}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y.$$

当 $n=2k$ 为偶数时 $\mathrm{e}^{-\frac{\pi}{2}\mathrm{i}n^2}=1$; 当 $n=2k+1$ 为奇数时 $\mathrm{e}^{-\frac{\pi}{2}\mathrm{i}n^2}=-\mathrm{i}$. 因此

$$\begin{aligned}1&=\sum_{k\in\mathbb{Z}}\int_{-k}^{-k+1}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y+\sum_{k\in\mathbb{Z}}(-\mathrm{i})\int_{-k-\frac{1}{2}}^{-k+\frac{1}{2}}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y-\mathrm{i}\int_{-\infty}^{+\infty}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y,\end{aligned}$$

这说明

$$\int_{-\infty}^{+\infty}\mathrm{e}^{2\pi\mathrm{i}y^2}\,\mathrm{d}y=\frac{1}{1-\mathrm{i}}=\frac{1}{2}(1+\mathrm{i}).$$

分别考虑上式的实部和虚部并作简单的伸缩变换可得如下 Fresnel 积分:

$$\int_{-\infty}^{+\infty}\cos(x^2)\,\mathrm{d}x=\int_{-\infty}^{+\infty}\sin(x^2)\,\mathrm{d}x=\frac{1}{2}\sqrt{2\pi}.\tag{14.1}$$

定理 14.4.17 后半段改用上下极限做:

其次, 任取 $[a,b)\subset[0,1]$, 考虑阶梯函数 $\phi(x)$: $\phi(x)=1$, $x\in[a,b)$; $\phi(x)=0$, $x\notin[a,b)$. 任给 $\varepsilon>0$, 容易找到周期为 1 的分段线性连续函数 f_1,f_2, 使得在 $[0,1]$ 中有

$$f_1(x)\le\phi(x)\le f_2(x),\ \ \int_0^1[f_2(x)-f_1(x)]\,\mathrm{d}x<\frac{\varepsilon}{2}.$$

因为 f_1,f_2 的 Fourier 展开一致收敛到自身, 故可以找三角多项式 P_1,P_2, 使得

$$P_1(x)\le\phi(x)\le P_2(x),\ \ \int_0^1[P_2(x)-P_1(x)]\,\mathrm{d}x<\varepsilon.$$

此时

$$\frac{1}{n}\sum_{k=1}^n P_1(x_k)\le\frac{1}{n}\sum_{k=1}^n\phi(x_k)\le\frac{1}{n}\sum_{k=1}^n P_2(x_k).$$

令 $n \to \infty$ 可得

$$\int_0^1 P_1(x)\,\mathrm{d}x \le \varliminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\phi(x_k) \le \varlimsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\phi(x_k) \le \int_0^1 P_2(x)\,\mathrm{d}x.$$

这说明

$$\int_0^1 \phi(x)\,\mathrm{d}x - \varepsilon \le \varliminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\phi(x_k) \le \varlimsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\phi(x_k) \le \int_0^1 \phi(x)\,\mathrm{d}x + \varepsilon.$$

由 ε 的任意性即得

$$\lim_{n\to\infty} \frac{s_n[a,b)}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}\phi(x_k) = \int_0^1 \phi(x)\,\mathrm{d}x = b-a.$$

第十五章　含参变量的积分

15.1　含参变量的积分

P475 页10行: “§12.3” 改为 §9.3

定理 15.1.4: 标上 [Leibniz 公式].

P478 页，原习题 15.1， 2(1) 分母中的 a 应改为 λ

15.2　含参变量的广义积分

本小节中多数地方的“广义积分”应明确为“无穷积分”.

在定义 15.2.1 前增加一段:

根据无穷积分的定义, $I(y)$ 可以表示为变上限黎曼积分的极限:

$$I(y) = \lim_{A\to+\infty}\int_a^A f(x,y)\,\mathrm{d}x.$$

为了研究 $I(y)$ 关于 y 的连续性、可导性以及积分性质等, 与函数列类似, 我们引进一致收敛的概念.

在例 15.2.1 前增加一个注记:

注 因为积分总可以视为变下限积分或变上限积分的极限, 即便是含参变量的黎曼积分, 也存在着极限是否关于参数一致收敛的问题. 比如, 对于每一个 $y\in[0,1]$, $\displaystyle\int_0^1 \frac{y}{x^2+y^2}\,\mathrm{d}x$ 都是正常的黎曼积分. 不过, 任给正整数 $n\geq 1$, 取 $y=\frac{1}{n}$, 则

$$\int_0^{\frac{1}{n}} \frac{y}{x^2+y^2}\,\mathrm{d}x = \arctan(nx)\Big|_0^{\frac{1}{n}} = \frac{\pi}{4},$$

这说明当 $n\to\infty$ 时 $\displaystyle\int_{\frac{1}{n}}^1 \frac{y}{x^2+y^2}\,\mathrm{d}x$ 虽然收敛于 $\displaystyle\int_0^1 \frac{y}{x^2+y^2}\,\mathrm{d}x$, 但它关于 $y\in[0,1]$ 并不是一致收敛的. 当然, 我们主要关心的是有瑕点的情形. 因此, 谈到积分关于参数一致收敛的时候, 通常需要明确它是无穷积分的一致收敛性还是对于某个瑕点的瑕积分的一致收敛性.

P483 页16行: 注意“定理 8.2.1”(源文件不用改)

习题 15.2:

3. 研究下列无穷积分关于参数的一致收敛性:

4. 研究如下瑕积分关于参数的一致收敛性:

增加习题:

11. 设 $\alpha \geq 0$, 证明:

(1) $\int_0^{+\infty} \mathrm{e}^{-\alpha x^2} \sin(x^2)\,\mathrm{d}x = \sqrt{\frac{\pi}{8}}\sqrt{\frac{\sqrt{1+\alpha^2}-\alpha}{1+\alpha^2}}$; (2) $\int_0^{+\infty} \mathrm{e}^{-\alpha x^2} \cos(x^2)\,\mathrm{d}x = \sqrt{\frac{\pi}{8}}\sqrt{\frac{\sqrt{1+\alpha^2}+\alpha}{1+\alpha^2}}$.

15.3 Euler 积分

15.4 Laplace 变换

[Warning] 本节内容可以作为选读材料.

设 f 是定义在 $(0,+\infty)$ 中的函数. 在计算积分 $\int_0^{+\infty} f(t)\,\mathrm{d}t$ 时, 我们曾用过引入衰减因子 $\mathrm{e}^{-\alpha t}$ 的技巧. 这个技巧还有着其他广泛的用途, 值得进一步介绍.

定义 15.4.1 (Laplace 变换)**.** *设 f 是定义在 $(0,+\infty)$ 中的函数, 记*

$$S = \left\{ s \in \mathbb{R} \,\middle|\, \int_0^{+\infty} f(t)\mathrm{e}^{-st}\,\mathrm{d}t \text{ 收敛} \right\}.$$

若 $S \neq \emptyset$, 记

$$\mathcal{L}(f)(s) = \int_0^{+\infty} f(t)\mathrm{e}^{-st}\,\mathrm{d}t, \quad s \in S,$$

$\mathcal{L}(f)$ 称为 f 的 Laplace 变换.

Laplace 变换是一种所谓的积分变换, 它最早可以追溯到 Euler 在 18 世纪 60 年代的工作. 18 世纪 80 年代, Laplace 在研究概率论时用到了这种积分变换.

Laplace 变换也称为 Laplace 积分, 它与幂级数的关系十分紧密. 首先, 若 $s \in S$, 则根据 Abel 判别法可知 $[s,+\infty) \subset S$, 这说明当 $S \neq \emptyset$ 时, 它一定是一个区间. 这一点与幂级数的收敛域有些类似. 其次, 考虑幂级数 $\sum\limits_{n=0}^{\infty} a_n x^n$. 当 $x > 0$ 时, 记 $x = \mathrm{e}^{-s}$, 则 $\sum\limits_{n=0}^{\infty} a_n x^n = \sum\limits_{n=0}^{\infty} a_n \mathrm{e}^{-ns}$, 它可以看成 Laplace 积分的离散形式. 如果在 $[0,+\infty)$ 中定义函数 f 为: 当 $t \in [n, n+1)$ 时 $f(t) = a_n$, $\forall\, n \geq 0$, 则

$$\sum_{n=0}^{\infty} a_n \mathrm{e}^{-ns} = \sum_{n=0}^{\infty} \frac{s}{1-\mathrm{e}^{-s}} \int_n^{n+1} f(t)\mathrm{e}^{-st}\,\mathrm{d}t = \frac{s}{1-\mathrm{e}^{-s}} \int_0^{\infty} f(t)\mathrm{e}^{-st}\,\mathrm{d}t, \tag{15.1}$$

这几乎就是 Laplace 积分了.

例 15.4.1. *设 $t > 0$ 时 $f(t) \equiv 1$, 求 $\mathcal{L}(f)$.*

解 由定义可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}\mathrm{e}^{-st}\,\mathrm{d}t=\frac{1}{s},\ \ s>0.$$

例 15.4.2. *设 $n\geq 1$, 且 $t>0$ 时 $f(t)=t^n$, 求 $\mathcal{L}(f)$.*

解 由定义可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}t^n\mathrm{e}^{-st}\,\mathrm{d}t=(-1)^n\frac{d^n}{ds^n}\int_0^{+\infty}\mathrm{e}^{-st}\,\mathrm{d}t=(-1)^n\frac{d^n}{ds^n}\frac{1}{s}=\frac{n!}{s^{n+1}},\ \ s>0.$$

例 15.4.3. *设 $\alpha>0$, 且 $t>0$ 时 $f(t)=t^{\alpha-1}$, 求 $\mathcal{L}(f)$.*

解 由定义可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}t^{\alpha-1}\mathrm{e}^{-st}\,\mathrm{d}t=\frac{1}{s^\alpha}\int_0^{+\infty}x^\alpha\mathrm{e}^{-x}\,\mathrm{d}x=\frac{\Gamma(\alpha)}{s^\alpha},\ \ s>0.$$

例 15.4.4. *设 $\alpha\in\mathbb{R}$, 且 $t>0$ 时 $f(t)=\sin(\alpha t)$, 求 $\mathcal{L}(f)$.*

解 由定义和**例 6.6.1**可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}\sin(\alpha t)\mathrm{e}^{-st}\,\mathrm{d}t=\frac{\alpha}{s^2+\alpha^2},\ \ s>0.$$

同理, 若 $f(t)=\cos(\alpha t)$, 则 $\mathcal{L}(f)(s)=\frac{s}{s^2+\alpha^2}$.

例 15.4.5. *设 $\alpha\in\mathbb{R}$, 且 $t>0$ 时 $f(t)=\mathrm{e}^{\alpha t}$, 求 $\mathcal{L}(f)$.*

解 由定义可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}\mathrm{e}^{(\alpha-s)t}\,\mathrm{d}t=\frac{1}{s-\alpha},\ \ s>\alpha.$$

从定义可以看出, 若 $f(t)=\mathrm{e}^{t^2}$, 则 f 的 Laplace 变换不存在. 下面的结果给出了 Laplace 变换存在的充分条件.

定理 15.4.1. *设 f 满足如下条件:*
(i) 任给 $b>0$, f 在 $(0,b]$ 中的积分均存在;
(ii) 存在常数 $\alpha\in\mathbb{R}$ 使得 $f(t)=O\big(\mathrm{e}^{\alpha t}\big)\ (t\to+\infty)$,
则当 $s>\alpha$ 时, $\mathcal{L}(f)(s)$ 存在.

证明 由条件 (ii) 可知, 存在 $M,T>0$, 使得当 $t\geq T$ 时 $|f(t)|\leq M\mathrm{e}^{\alpha t}$. 此时 $|f(t)\mathrm{e}^{-st}|\leq\mathrm{e}^{(\alpha-s)t}$, 这说明当 $s>\alpha$ 时, $f(t)\mathrm{e}^{-st}$ 在 $(T,+\infty)$ 中的无穷积分收敛. 另一方面, 由条件 (i) 和 Abel 判别法又可以知道 $f(t)\mathrm{e}^{-st}$ 在 $(0,T]$ 中的积分也是收敛的. 这就说明当 $s>\alpha$ 时, $\mathcal{L}(f)(s)$ 存在.

需要注意的是, 上述条件只是充分条件, 不一定是必要条件. 比如 $f(t)=2t\mathrm{e}^{t^2}\cos\left(\mathrm{e}^{t^2}\right)$ 虽不满足上述条件, 但当 $s>0$ 时, 其 Laplace 变换存在.

定理 15.4.2 (Laplace 变换的唯一性). 设 f, g 是定义在 $(0, +\infty)$ 中的函数. 如果存在 $\alpha \in \mathbb{R}$, 使得当 $s \geq \alpha$ 时均有 $\mathcal{L}(f)(s) = \mathcal{L}(g)(s)$, 则 f 和 g 几乎处处相等, 即在某个零测集之外一定相等.

证明 记 $h(t) = \big[f(t) - g(t)\big]\mathrm{e}^{-\alpha t}$, 则当 $s \geq 0$ 时, 有

$$\int_0^{\infty} h(t)\mathrm{e}^{-st}\,\mathrm{d}t = \mathcal{L}(f)(s+\alpha) - \mathcal{L}(g)(s+\alpha) = 0.$$

于是对每一个多项式 $P(x) = \sum\limits_{k=0}^{n} c_k x^k$, 均有

$$\int_0^{\infty} h(t)P\big(\mathrm{e}^{-t}\big)\,\mathrm{d}t = \sum_{k=0}^{n} c_k \int_0^{\infty} h(t)\mathrm{e}^{-kt}\,\mathrm{d}t = 0.$$

作变量替换 $t = -\ln x$ 可得

$$\int_0^1 \frac{1}{x}h(-\ln x)P(x)\,\mathrm{d}x = \int_0^{\infty} h(t)P\big(\mathrm{e}^{-t}\big)\,\mathrm{d}t = 0.$$

任给多项式 $Q(x)$, 记

$$P(x) = \int_0^x Q(u)\,\mathrm{d}x,\ \ H(x) = \int_x^1 \frac{1}{u}h(-\ln u)\,\mathrm{d}u,$$

由(推广的)分部积分可得

$$\int_0^1 H(x)Q(x)\,\mathrm{d}x = H(x)P(x)\Big|_0^1 + \int_0^1 \frac{1}{x}h(-\ln x)P(x)\,\mathrm{d}x = 0.$$

注意到 $H \in C[0,1]$, 根据 Weierstrass 逼近定理, 任给 $\varepsilon > 0$, 存在多项式 $Q(x)$, 使得 $|H(x) - Q(x)| < \varepsilon$, $\forall\, x \in [0,1]$. 于是

$$\int_0^1 H^2(x)\,\mathrm{d}x = \int_0^1 H(x)\big[H(x) - Q(x)\big]\,\mathrm{d}x \leq \varepsilon \int_0^1 |H(x)|\,\mathrm{d}x.$$

由 ε 的任意性可知 $\displaystyle\int_0^1 H^2(x)\,\mathrm{d}x = 0$, 从而 $H \equiv 0$. 再由微积分基本公式和 Lebesgue 定理容易得出 h 几乎处处为零.

在介绍 Laplace 变换的应用之前, 我们先讨论它的基本性质. 首先列几条简单的, 其证明可以直接从定义得到, 故略去.

- (线性性质) 设 $c_1, c_2 \in \mathbb{R}$, 如果 $\mathcal{L}(f_1)(s)$ 和 $\mathcal{L}(f_2)(s)$ 都存在, 则

$$\mathcal{L}(c_1 f_1 + c_2 f_2)(s) = c_1\mathcal{L}(f_1)(s) + c_2\mathcal{L}(f_2)(s);$$

- (平移性质) 设 $\mathcal{L}(f)(s)$ 存在, $a \geq 0$. 我们规定, 当 $t < 0$ 时 $f(t) = 0$. 记 $f_a(t) = f(t-a)$, 则 $\mathcal{L}(f_a)(s) = \mathrm{e}^{-as}\mathcal{L}(f)(s)$;

- (伸缩性质) 设 $\mathcal{L}(f)(s)$ 存在, $b>0$, 记 $f^b(t)=f(bt)$, 则 $\mathcal{L}\big(f^b\big)(s)=\frac{1}{b}\mathcal{L}(f)\big(\frac{s}{b}\big)$.

下面的结果对应着幂级数的 Abel 引理, 它给出了用衰减因子计算广义积分的理论依据.

引理 15.4.3. *设广义积分* $\int_0^{+\infty} f(t)\,\mathrm{d}t$ *收敛, 则* $\lim\limits_{s\to 0^+}\mathcal{L}(f)(s)=\int_0^{+\infty} f(t)\,\mathrm{d}t$.

证明 记 $F(t)=\int_0^t f(u)\,\mathrm{d}u$, 则 $F\in C[0,+\infty)$, 且由题设可知 $\lim\limits_{t\to+\infty}F(t)$ 存在, 记为 ℓ. 当 $s>0$ 时, 由分部积分可得

$$\mathcal{L}(f)(s)=\mathrm{e}^{-st}F(t)\Big|_0^{+\infty}+\int_0^{+\infty} s\mathrm{e}^{-st}F(t)\,\mathrm{d}t=\int_0^{+\infty} s\mathrm{e}^{-st}F(t)\,\mathrm{d}t,$$

它可以改写为

$$\mathcal{L}(f)(s)-\ell=\int_0^{+\infty}\mathrm{e}^{-u}\big[F\big(u/s\big)-\ell\big]\,\mathrm{d}u.$$

注意 F 是有界函数, 设 $|F-\ell|\le M$. 由题设, 任给 $\varepsilon>0$, 存在 $A>0$, 使得当 $t>A$ 时均有 $|F(t)-\ell|<\varepsilon$. 于是当 $0<s<\frac{\varepsilon}{A}$ 时, 有

$$\big|\mathcal{L}(f)(s)-\ell\big|\le\int_0^{\varepsilon}\mathrm{e}^{-u}M\,\mathrm{d}u+\int_{\varepsilon}^{+\infty}\mathrm{e}^{-u}\varepsilon\,\mathrm{d}u<M\varepsilon+\varepsilon,$$

这说明欲证结论成立.

命题 15.4.4 (连续性质)**.** *设* $\mathcal{L}(f)(s_0)$ *存在, 则* $\mathcal{L}(f)\in C[s_0,+\infty)$.

证明 $\mathcal{L}(f)$ 在 $s=s_0$ 处的右连续性可由上述引理直接得出. 任给 $s_1>s_0$, 我们要说明 $\mathcal{L}(f)$ 在 $s=s_1$ 处连续. 为此, 记 $\delta=s_1-s_0$, $s'=s-s_0$, $g(t)=f(t)\mathrm{e}^{-s_0t}$. 当 $s>s_0$ 时, 由分部积分可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty} g(t)\mathrm{e}^{-s't}\,\mathrm{d}t=\int_0^{+\infty} G(t)s'\mathrm{e}^{-s't}\,\mathrm{d}t,$$

其中 $G(t)=\int_0^t g(u)\,\mathrm{d}u$. 注意 $G(t)s'\mathrm{e}^{-s't}$ 关于 t,s' 连续, 且无穷积分 $\int_0^{+\infty} G(t)s'\mathrm{e}^{-s't}\,\mathrm{d}t$ 关于参数 $s'\in\big[\delta/2,+\infty\big)$ 一致收敛, 由**引理 15.2.1** 即知该无穷积分关于 s' 在 $s'=\delta$ 处连续, 即 $\mathcal{L}(f)$ 在 $s=s_1$ 处连续.

命题 15.4.5 (可导性质)**.** *设* $\mathcal{L}(f)(s_0)$ *存在, 则* $\mathcal{L}(f)$ *在* $(s_0,+\infty)$ *中可导, 且*

$$\mathcal{L}(f)'(s)=\int_0^{+\infty}-tf(t)\mathrm{e}^{-st}\,\mathrm{d}t,\ \ \forall\, s>s_0. \tag{15.2}$$

若 $\int_0^{+\infty} tf(t)\mathrm{e}^{-s_0t}\,\mathrm{d}t$ *收敛, 则* $\mathcal{L}(f)$ *在* s_0 *处的右导数也存在, 且*

$$\mathcal{L}(f)'_+(s_0)=\int_0^{+\infty}-tf(t)\mathrm{e}^{-s_0t}\,\mathrm{d}t.$$

证明 不妨设 $s_0=0$. 任给 $\delta>0$, 我们先来说明 $\mathcal{L}(f)$ 在 $[\delta,+\infty)$ 中可导. 记 $F(t)=\displaystyle\int_0^t f(u)\,\mathrm{d}u$, 当 $s\geq\delta$ 时, 由分部积分可得

$$\mathcal{L}(f)(s)=\int_0^{+\infty}F(t)s\mathrm{e}^{-st}\,\mathrm{d}t.$$

注意到 $F(t)s\mathrm{e}^{-st}$ 关于 $(t,s)\in[0,+\infty)\times[\delta,+\infty)$ 连续、关于 s 可导且不难验证它满足**定理 15.2.3**的条件. 因此

$$\mathcal{L}(f)'(s)=\int_0^{+\infty}F(t)\frac{d}{ds}\big(s\mathrm{e}^{-st}\big)\,\mathrm{d}t=\int_0^{+\infty}F(t)\frac{d}{dt}\big(t\mathrm{e}^{-st}\big)\,\mathrm{d}t.$$

再利用分部积分即得 (15.2) 式.

设 $\displaystyle\int_0^{+\infty}tf(t)\,\mathrm{d}t$ 收敛, 我们要证明

$$\lim_{s\to0^+}\frac{1}{s}\int_0^{+\infty}f(t)\big[\mathrm{e}^{-st}-1\big]\,\mathrm{d}t=\int_0^{+\infty}-tf(t)\,\mathrm{d}t.$$

记 $g(t)=-tf(t)$, $h(x)=\frac{1}{x}\big[1-\mathrm{e}^{-x}\big]=\displaystyle\int_0^1\mathrm{e}^{-xu}\,\mathrm{d}u$, 则上式可以改写为

$$\lim_{s\to0^+}\int_0^{+\infty}g(t)h(st)\,\mathrm{d}t=\int_0^{+\infty}g(t)\,\mathrm{d}t.$$

这可以用完全类似于引理 15.4.3 的方法证明, 我们留给读者.

由 (15.2) 式可知, Laplace 积分的导数仍为某个 Laplace 积分. 因此, 反复利用上述命题可以得出

推论 15.4.6. *设 $\mathcal{L}(f)(s_0)$ 存在, 则 $\mathcal{L}(f)$ 在 $(s_0,+\infty)$ 中任意次可导, 且其 n 阶导数满足*

$$\big[\mathcal{L}(f)\big]^{(n)}(s)=\int_0^{+\infty}(-t)^nf(t)\mathrm{e}^{-st}\,\mathrm{d}t,\ \ \forall\, s>s_0. \tag{15.3}$$

若上式右端对 $s=s_0$ 也收敛, 则它收敛于 $\mathcal{L}(f)$ 在 $s=s_0$ 处的 n 阶右导数.

例 15.4.6. *复值函数的 Laplace 变换.*

设 f 为复值函数, 其实部和虚部分别为 g,h. 定义 $\mathcal{L}(f)=\mathcal{L}(g)+\mathrm{i}\mathcal{L}(h)$. 设 $z\in\mathbb{C}$, 其实部和虚部分别为 β,α. 由例 15.4.4 可知, 复值函数 $f(t)=\mathrm{e}^{zt}$ 的 Laplace 变换为

$$\mathcal{L}(f)(s)=\frac{s-\beta}{(s-\beta)^2+\alpha^2}+\mathrm{i}\frac{\alpha}{(s-\beta)^2+\alpha^2}=\frac{1}{s-\beta-\mathrm{i}\alpha}=\frac{1}{s-z}.$$

当 $k\geq1$ 时, 利用导数性质还可以知道 $\frac{t^k}{k!}\mathrm{e}^{zt}$ 的 Laplace 变换是 $\frac{1}{(s-z)^{k+1}}$.

下面我们再来考虑导数的 Laplace 积分.

命题 15.4.7. *设 $f \in C(0,+\infty)$, 如果*

(i) f 分段可导, 且任给 $b>0$, f' 在 $(0,b]$ 中的积分都存在;

(ii) $\mathcal{L}(f)(s)$ 存在,

则 $\mathcal{L}(f')(s)$ 存在当且仅当 $\lim\limits_{t\to+\infty} f(t)\mathrm{e}^{-st}=0$, 此时

$$\mathcal{L}(f')(s)=s\mathcal{L}(f)(s)-\lim_{t\to0^+} f(t). \tag{15.4}$$

证明 我们先来说明 f 在 0 处的右极限存在. 事实上, 当 $b>0$ 时, 根据条件 (i) 以及 Newton-Leibniz 公式可得

$$\int_0^b f(t)\,\mathrm{d}t=f(b)-\lim_{t\to0^+} f(t).$$

这说明 f 在 0 处的右极限的确存在. 为了方便起见, 我们规定 $f(0)=\lim\limits_{t\to0^+} f(t)$.

当 $b>0$ 时, 由分部积分可得

$$\int_0^b f'(t)\mathrm{e}^{-st}\,\mathrm{d}t=f(b)\mathrm{e}^{-sb}-f(0)+s\int_0^b f(t)\mathrm{e}^{-st}\,\mathrm{d}t. \tag{15.5}$$

由此可见, $\mathcal{L}(f')(s)$ 存在当且仅当 $\lim\limits_{b\to+\infty} f(b)\mathrm{e}^{-sb}$ 存在. 由条件 (ii) 可知, 此极限存在的话必须为零. 此时, 在 (15.5) 式中令 $b\to+\infty$ 即得 (15.4) 式.

为了将刚才的结果推广到高阶导数, 我们先做一点准备工作.

引理 15.4.8. *设 $f\in C^{n-1}(0,1]$. 如果 $f^{(n-1)}$ 分段可导, 且 $f^{(n)}$ 在 $(0,1]$ 中积分存在, 则 $\lim\limits_{t\to0^+} f(t)$ 存在, 且若定义 $f(0)=\lim\limits_{t\to0^+} f(t)$, 则 $f\in C^{n-1}[0,1]$.*

证明 由题设和 Newton-Leibniz 公式可知 $\lim\limits_{t\to0^+} f^{(n-1)}(t)$ 存在. 对 $f^{(n-2)}$ 应用 Lagrange 微分中值定理, 再由 Cauchy 准则可知 $\lim\limits_{t\to0^+} f^{(n-2)}(t)$ 也存在. 重复这一过程, 最后可得 $\lim\limits_{t\to0^+} f(t)$ 存在.

若定义 $f(0)=\lim\limits_{t\to0^+} f(t)$, 则由**例 5.2.4** 可知 f 在 0 处(右)可导, 且 $f'(0)=\lim\limits_{t\to0^+} f'(t)$. 再对 f' 应用**例 5.2.4** 可知 f 在 0 处存在二阶导数, 且 $f''(0)=\lim\limits_{t\to0^+} f''(t)$. 重复这一过程, 最后可得 $f\in C^{n-1}[0,1]$.

引理 15.4.9. *设 F 在 $(0,+\infty)$ 中 n 阶可导. 如果 $\lim\limits_{t\to+\infty} F(t)$ 存在且 $\lim\limits_{t\to+\infty} F^{(n)}(t)=0$, 则当 $1\le i\le n$ 时, 均有 $\lim\limits_{t\to+\infty} F^{(i)}(t)=0$.*

证明 记 $\ell=\lim\limits_{t\to+\infty} F(t)$. 由题设可知, 任给 $\varepsilon>0$, 存在 $A>0$, 使得当 $t\ge A$ 时, 均有 $\left|F(t)-\ell\right|<\varepsilon$, $\left|F^{(n)}(t)\right|<\varepsilon$.

当 $t\ge A$ 时, 对 $j=1,2,\cdots,n-1$, 由 Taylor 公式可知, 存在 $\xi_j\in(t,t+j)$, 使得

$$F(t+j)-F(t)=\sum_{i=1}^{n-1}\frac{F^{(i)}(t)}{i!}j^i+\frac{F^{(n)}(\xi_j)}{n!}j^n.$$

这说明当 $t \geq A$ 时

$$\Big|\sum_{i=1}^{n-1} \frac{F^{(i)}(t)}{i!} j^i\Big| < 2\varepsilon + \frac{(n-1)^n}{n!}\varepsilon = C\varepsilon.$$

改写为向量和矩阵的形式后, 上式成为

$$\Big(F'(t), \frac{1}{2!}F''(t), \cdots, \frac{1}{(n-1)!}F^{(n-1)}(t)\Big)M = o(1), \ \ (t \to +\infty),$$

其中 $M = \big(j^i\big)_{(n-1)\times(n-1)}$ 是 Vandermonde 矩阵, 其行列式不等于零. 于是有

$$\Big(F'(t), \frac{1}{2!}F''(t), \cdots, \frac{1}{(n-1)!}F^{(n-1)}(t)\Big) = o(1)M^{-1} = o(1), \ \ (t \to +\infty),$$

引理得证.

定理 15.4.10. *设 $f \in C^{n-1}(0, +\infty)$ 且 $\lim\limits_{t\to+\infty} f^{(n-1)}(t)\mathrm{e}^{-st} = 0$. 如果 $f^{(n-1)}$ 分段可导, 且任给 $b > 0$, $f^{(n)}(t)$ 在 $(0, b]$ 中积分均存在, 则*

(1) $\lim\limits_{t\to0^+} f(t)$ 存在, 且若定义 $f(0) = \lim\limits_{t\to0^+} f(t)$, 则 $f \in C^{n-1}[0, +\infty)$;

(2) 当 $\mathcal{L}(f)(s)$ 存在时, $\mathcal{L}(f^{(n)})(s)$ 也存在, 且

$$\mathcal{L}\big(f^{(n)}\big)(s) = s^n\mathcal{L}(f)(s) - f(0)s^{n-1} - f'(0)s^{n-2} - \cdots - f^{(n-2)}(0)s - f^{(n-1)}(0). \tag{15.6}$$

证明 (1) 由引理 15.4.8 即可得到.

(2) $n = 1$ 的情形我们已经证过, 下设 $n > 1$. 我们先来证明, 当 $0 \leq i \leq n-1$ 时, 均有 $\lim\limits_{t\to+\infty} f^{(i)}(t)\mathrm{e}^{-st} = 0$. 若 $s > 0$, 反复应用 L'Hospital 法则即得

$$\lim_{t\to+\infty} f^{(i)}(t)\mathrm{e}^{-st} = \lim_{t\to+\infty} \frac{f^{(i)}(t)}{\mathrm{e}^{st}} = \lim_{t\to+\infty} \frac{f^{(i+1)}(t)}{s\mathrm{e}^{st}} = \cdots = \lim_{t\to+\infty} \frac{f^{(n-1)}(t)}{s^{n-1-i}\mathrm{e}^{st}} = 0.$$

若 $s \leq 0$, 则由题设可知 $\mathcal{L}(f)(0)$ 存在, 且 $\lim\limits_{t\to+\infty} f^{(n-1)}(t) = 0$. 对 $F(t) = \displaystyle\int_0^t f(u)\,\mathrm{d}u$ 应用引理 15.4.9 可知, 当 $0 \leq i \leq n-1$ 时, 均有 $\lim\limits_{t\to+\infty} f^{(i)}(t) = 0$. 于是当 $s < 0$ 时我们仍然可以反复应用 $\frac{0}{0}$ 型的 L'Hospital 法则得到欲证结论.

最后, 反复应用 (15.4) 式即可得到 (15.6) 式.

从 (15.6) 式可以看出, Laplace 变换可以将求导运算转换为代数运算. 正是因为这一点才使得 Laplace 变换成为解微分方程的有力工具.

例 15.4.7. *解初值问题 $y' + 3y = \mathrm{e}^{2t}$, $y(0) = 1$.*

解 我们先假定可以做 Laplace 变换. 记 $Y(s) = \mathcal{L}(y)(s)$, 对方程两边做 Laplace 变换可得

$$sY(s) - 1 + 3Y(s) = \frac{1}{s-2},$$

从中可以解出

$$Y(s)=\frac{1}{5}\frac{1}{s-2}+\frac{4}{5}\frac{1}{s+3}.$$

根据例 15.4.5 可知, $Y(s)=\mathcal{L}(\phi)(s)$, 其中

$$\phi(t)=\frac{1}{5}\mathrm{e}^{2t}+\frac{4}{5}\mathrm{e}^{-3t}.$$

容易检验 $y=\phi(t)$ 的确满足上述初值问题. 由**定理 4.4.2** 还可以知道这也是满足上述初值问题的唯一解.

例 15.4.8. *解初值问题* $y''-y'-2y=\mathrm{e}^{2t}$, $y(0)=0$, $y'(0)=1$.

解 与前例类似, 记 $Y(s)=\mathcal{L}(y)(s)$, 则

$$s^2Y(s)-1-sY(s)-2Y(s)=\frac{1}{s-2},$$

从中可以解出

$$Y(s)=\frac{s-1}{(s+1)(s-2)^2}=-\frac{2}{9}\frac{1}{s+1}+\frac{2}{9}\frac{1}{s-2}+\frac{1}{3}\frac{1}{(s-2)^2}.$$

这说明 $Y(s)=\mathcal{L}(\phi)(s)$, 其中

$$\phi(t)=-\frac{2}{9}\mathrm{e}^{-t}+\frac{2}{9}\mathrm{e}^{2t}+\frac{1}{3}t\mathrm{e}^{2t},$$

可以检验 $y=\phi(t)$ 满足上述初值问题, 且就是唯一的解.

例 15.4.9. *解初值问题* $y''-2y'+2y=\mathrm{e}^{-t}$, $y(0)=0$, $y'(0)=1$.

解 与前例类似, 记 $Y(s)=\mathcal{L}(y)(s)$, 则

$$s^2Y(s)-1-2sY(s)+2Y(s)=\frac{1}{s+1},$$

从中可以解出

$$Y(s)=\frac{s+2}{(s+1)[(s-1)^2+1]}=\frac{1}{5}\frac{1}{s+1}-\frac{1}{5}\frac{s-1}{(s-1)^2+1}+\frac{9}{5}\frac{1}{(s-1)^2+1}.$$

利用我们之前计算过的例子可得 $Y(s)=\mathcal{L}(\phi)(s)$, 其中

$$\phi(t)=\frac{1}{5}\mathrm{e}^{-t}-\frac{1}{5}\mathrm{e}^{t}\cos t+\frac{9}{5}\mathrm{e}^{t}\sin t.$$

可以检验 $y=\phi(t)$ 满足上述初值问题, 且就是唯一的解.

上述解常系数常微分方程的方法可以推广到高阶方程. 为此我们先来介绍 Laplace 变换的卷积性质. 这可以类比于幂级数的乘法性质. 设 $\sum\limits_{n=0}^{\infty}a_nx^n$ 和 $\sum\limits_{n=0}^{\infty}b_nx^n$ 都在 $(-R,R)$ 中收敛, 则

$$\Big(\sum_{n=0}^{\infty}a_nx^n\Big)\Big(\sum_{n=0}^{\infty}b_nx^n\Big)=\sum_{n=0}^{\infty}c_nx^n,$$

其中 $c_n=\sum\limits_{k=0}^{n}a_{n-k}b_k$. 这可以看成离散形式的卷积. 对于函数 f,g 来说, 其卷积 $f*g$ 定义为

$$(f*g)(t)=\int_0^t f(t-u)g(u)\,\mathrm{d}u. \tag{15.7}$$

注意, 若 $f(t)=g(t)=|t-0.5|^{-0.5}$, 则

$$\int_0^1 f(1-u)g(u)\,\mathrm{d}u=\int_0^1 |u-0.5|^{-1}\,\mathrm{d}u,$$

此积分是发散的. 因此, 为了让 (15.7) 式有意义, 需要对 f,g 加一些限制条件.

我们规定, 0 是 f,g 唯一可能的瑕点, 且任给 $b>0$, f,g 均在 $(0,b]$ 中绝对可积. 此时, 将 $f*g$ 改写为 $\varphi+\psi$ 的形式, 其中

$$\varphi(t)=\int_0^{\frac{t}{2}} f(t-u)g(u)\,\mathrm{d}u,\ \ \psi(t)=\int_{\frac{t}{2}}^{t} f(t-u)g(u)\,\mathrm{d}u=\int_0^{\frac{t}{2}} f(u)g(t-u)\,\mathrm{d}u.$$

当 $t>0$ 时, 由于 f,g 在 $\left[\frac{t}{2},t\right]$ 中有界, 容易看出 $\varphi(t)$ 和 $\psi(t)$ 作为积分都是存在的. 进一步还有

命题 15.4.11. *在上述规定下, $f*g\in C(0,+\infty)$.*

证明 根据对称性, 只要证明 φ 的连续性就可以了. 任给 $\delta\in(0,1)$, 当 $n>\frac{2}{\delta}$ 时, 由**命题 6.2.9**以及 Weierstrass 逼近定理可知, 存在多项式 P_n, 使得

$$\int_{\frac{1}{n}}^{\frac{1}{2\delta}}\big|P_n(u)-g(u)\big|\,\mathrm{d}u<\frac{1}{n}.$$

记

$$\varphi_n(t)=\int_{\frac{1}{n}}^{\frac{t}{2}} f(t-u)P_n(u)\,\mathrm{d}u=\int_{\frac{t}{2}}^{t-\frac{1}{n}} f(u)P_n(t-u)\,\mathrm{d}u,$$

由 P_n 为多项式不难看出 $\varphi_n\in C[\delta,+\infty)$.

设 M 为 $|f|$ 在 $\left[\frac{\delta}{2},\frac{1}{\delta}\right]$ 中的上界. 当 $t\in\left[\delta,\frac{1}{\delta}\right]$ 时, 我们有

$$\begin{aligned}|\varphi_n(t)-\varphi(t)|&\le\int_0^{\frac{1}{n}}|f(t-u)g(u)|\,\mathrm{d}u+\int_{\frac{1}{n}}^{\frac{t}{2}}\big|f(t-u)[P_n(u)-g(u)]\big|\,\mathrm{d}u\\&\le M\int_0^{\frac{1}{n}}|g(u)|\,\mathrm{d}u+M\int_{\frac{1}{n}}^{\frac{1}{2\delta}}\big|P_n(u)-g(u)\big|\,\mathrm{d}u\\&<M\int_0^{\frac{1}{n}}|g(u)|\,\mathrm{d}u+M\frac{1}{n},\end{aligned}$$

这说明 $\{\varphi_n\}$ 在 $\left[\delta,\frac{1}{\delta}\right]$ 中一致收敛于 φ, 因此 $\varphi\in C\left[\delta,\frac{1}{\delta}\right]$. 由 δ 的任意性即知 $\varphi\in C(0,+\infty)$.

引理 15.4.12. *设 f,g 满足上述规定, 则当 $b>0$ 时*

$$\int_0^b (f*g)(t)\,\mathrm{d}t=\int_0^b F(b-u)g(u)\,\mathrm{d}u,\ \ \text{其中 } F(u)=\int_0^u f(t)\,\mathrm{d}t.$$

证明 为了避开可能的瑕点, 取 $\delta\in(0,b)$, $\varepsilon\in(0,\delta/2)$. 先考虑二元函数 $f(t-u)g(u)$ 在区域 $D=\{\delta\le t\le b,\ \varepsilon\le u\le t/2\}$ 中的积分, 利用重积分化累次积分可得

$$\begin{aligned}\int_\delta^b\int_\varepsilon^{\frac{t}{2}}f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t&=\int_\delta^b\int_\varepsilon^{\frac{\delta}{2}}f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t+\int_\delta^b\int_{\frac{\delta}{2}}^{\frac{t}{2}}f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t\\&=\int_\varepsilon^{\frac{\delta}{2}}\int_\delta^b f(t-u)g(u)\,\mathrm{d}t\mathrm{d}u+\int_{\frac{\delta}{2}}^{\frac{b}{2}}\int_{2u}^b f(t-u)g(u)\,\mathrm{d}t\mathrm{d}u\\&=\int_\varepsilon^{\frac{\delta}{2}}[F(b-u)-F(\delta-u)]g(u)\,\mathrm{d}u\\&\quad+\int_{\frac{\delta}{2}}^{\frac{b}{2}}[F(b-u)-F(u)]g(u)\,\mathrm{d}u.\end{aligned}$$

先让 $\varepsilon\to0^+$, 再让 $\delta\to0^+$, 利用 f 局部有界, g 绝对可积以及 F 有界可得

$$\int_0^b\int_0^{\frac{t}{2}}f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t=\int_0^{\frac{b}{2}}[F(b-u)-F(u)]g(u)\,\mathrm{d}u.\tag{15.8}$$

同理可得

$$\int_0^b\int_{\frac{t}{2}}^t f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t=\int_0^b\int_0^{\frac{t}{2}}f(u)g(t-u)\,\mathrm{d}u\mathrm{d}t=\int_0^{\frac{b}{2}}[G(b-u)-G(u)]f(u)\,\mathrm{d}u,$$

其中 $G(u)=\displaystyle\int_0^u g(t)\,\mathrm{d}t$. 利用分部积分可得

$$\int_0^b\int_{\frac{t}{2}}^t f(t-u)g(u)\,\mathrm{d}u\mathrm{d}t=\int_0^{\frac{b}{2}}[g(b-u)+g(u)]F(u)\,\mathrm{d}u.\tag{15.9}$$

将 (15.8) 式和 (15.9) 相加即得欲证结论.

下面是 Laplace 变换关于卷积的主要结果, 它对应着幂级数的 Mertens 定理.

定理 15.4.13 (卷积性质)**.** *设 f,g 是在 $(0,+\infty)$ 中定义的函数, 如果它们满足如下条件*

(i) *0 是 f,g 唯一可能的瑕点, 且任给 $b>0$, f,g 均在 $(0,b]$ 中绝对可积;*

(ii) *$\mathrm{e}^{-s_0t}f(t)$ 在 $(0,+\infty)$ 中的广义积分收敛, $\mathrm{e}^{-s_0t}g(t)$ 在 $(0,+\infty)$ 中的广义积分绝对收敛,*

则当 $s\ge s_0$ 时 $\mathcal{L}(f*g)(s)=\mathcal{L}(f)(s)\cdot\mathcal{L}(g)(s)$.

证明 我们先来做一些化简. 首先, 只需要对 $s=s_0$ 证明就可以了. 其次, 注意到

$$\left(\mathrm{e}^{-s_0t}f\right)*\left(\mathrm{e}^{-s_0t}g\right)=\int_0^t\mathrm{e}^{-s_0(t-u)}f(t-u)\mathrm{e}^{-s_0u}g(u)\,\mathrm{d}u=\mathrm{e}^{-s_0t}(f*g),$$

因此不妨设 $s_0=0$. 现在, 欲证结论成为

$$\int_0^{+\infty}(f*g)(t)\,\mathrm{d}t=\int_0^{+\infty}f(t)\,\mathrm{d}t\int_0^{+\infty}g(t)\,\mathrm{d}t.$$

利用刚才的引理, 我们只需证明 $\lim\limits_{b\to+\infty}\displaystyle\int_0^b\big[F(b-u)-\ell\big]g(u)\,\mathrm{d}u=0$, 其中

$$F(u)=\int_0^u f(t)\,\mathrm{d}t,\ \ \ell=\int_0^{+\infty}f(t)\,\mathrm{d}t=\lim_{u\to+\infty}F(u).$$

注意到 F 是有界函数, 设 $|F|\le M$. 任给 $\varepsilon>0$, 存在 $A>0$, 使得当 $u\ge A$ 时 $|F(u)-\ell|<\varepsilon$. 由于 g 绝对可积, 不妨设 $|g|$ 在 $[A,+\infty)$ 中的积分小于 ε. 于是, 当 $b>2A$ 时, 有

$$\begin{aligned}\Big|\int_0^b\big[F(b-u)-\ell\big]g(u)\,\mathrm{d}u\Big|&\le\int_0^{\frac b2}\big|[F(b-u)-\ell]g(u)\big|\,\mathrm{d}u+\int_{\frac b2}^b\big|[F(b-u)-\ell]g(u)\big|\,\mathrm{d}u\\&\le\varepsilon\int_0^{\frac b2}|g(u)|\,\mathrm{d}u+2M\int_{\frac b2}^b|g(u)|\,\mathrm{d}u\\&\le\varepsilon\int_0^{+\infty}|g(u)|\,\mathrm{d}u+2M\int_A^{+\infty}|g(u)|\,\mathrm{d}u\\&<\varepsilon\int_0^{+\infty}|g(u)|\,\mathrm{d}u+2M\varepsilon,\end{aligned}$$

这说明欲证结论的确成立.

推论 15.4.14. *设 0 是函数 f,g,h 唯一可能的瑕点, 且任给 $b>0$, 这些函数均在 $(0,b]$ 中绝对可积. 则 $(f*g)*h=f*(g*h)$.*

证明 设 $b>0$, 当 $t\in(0,b]$ 时, 我们要验证 $[(f*g)*h](t)=[f*(g*h)](t)$. 根据卷积的定义, 不妨设 f,g,h 在 $(b,+\infty)$ 中恒为零. 此时 $f*g$, $g*h$ 在 $[2b,+\infty)$ 中也恒为零. 当 $s\ge0$ 时, 由上述定理可得

$$\mathcal{L}\big((f*g)*h\big)(s)=\mathcal{L}(f*g)(s)\cdot\mathcal{L}(h)(s)=\mathcal{L}(f)(s)\cdot\mathcal{L}(g)(s)\cdot\mathcal{L}(h)(s).$$

同理可以得出

$$\mathcal{L}\big(f*(g*h)\big)(s)=\mathcal{L}(f)(s)\cdot\mathcal{L}(g)(s)\cdot\mathcal{L}(h)(s).$$

由 Laplace 变换的唯一性可知 $(f*g)*h$ 和 $f*(g*h)$ 几乎处处相等. 再由卷积的连续性可知 $(f*g)*h=f*(g*h)$.

例 15.4.10. *B 函数回顾.*

设 $\alpha,\beta>0$, 当 $t>0$ 时, 记 $f(t)=t^{\alpha-1}$, $g(t)=t^{\beta-1}$. 根据定义, 我们有

$$(f*g)(t)=\int_0^t(t-u)^{\alpha-1}u^{\beta-1}\,\mathrm{d}u=t^{\alpha+\beta-1}\int_0^1(1-s)^{\alpha-1}s^{\beta-1}\,\mathrm{d}s=t^{\alpha+\beta-1}B(\alpha,\beta).$$

对上式两边做 Laplace 变换可得

$$\frac{\Gamma(\alpha)}{s^\alpha}\frac{\Gamma(\beta)}{s^\beta} = \mathcal{L}(f)(s)\mathcal{L}(g)(s) = \mathcal{L}(f*g)(s) = \frac{\Gamma(\alpha+\beta)}{s^{\alpha+\beta}}B(\alpha,\beta),$$

这就重新得到了**定理 15.3.1**中的结论.

例 15.4.11. 解积分方程 $y(t) = \mathrm{e}^{-t} + \int_0^t \sin(t-u)y(u)\,\mathrm{d}u.$

解 记 $Y(s) = \mathcal{L}(y)(s)$, 对上述方程两边做 Laplace 变换可得

$$Y(s) = \frac{1}{s+1} + \frac{1}{s^2+1}Y(s),$$

从中解出

$$Y(s) = \frac{s^2+1}{(s+1)s^2} = \frac{2}{s+1} + \frac{1}{s^2} - \frac{1}{s},$$

这说明 $Y(s) = \mathcal{L}(\phi)(s)$, 其中

$$\phi(t) = 2\mathrm{e}^{-t} + t - 1.$$

不难验证 $y = \phi(t)$ 是上述积分方程的唯一解.

最后, 我们简要介绍用 Laplace 变换求解一类常系数高阶常微分方程的一般性做法. 设 $\{a_i\}_{i=0}^{n-1}$ 为一组常数, 考虑如下微分方程

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1y'(t) + a_0y(t) = f(t),\ \ t \in (0,+\infty). \tag{15.10}$$

给定初始条件

$$y(0^+) = y_0,\ y'(0^+) = y_1, \cdots,\ y^{(n-1)}(0^+) = y_{n-1}, \tag{15.11}$$

我们想用 Laplace 变换来求解上述方程. 为此, 需要 f 满足一定的条件. 为了简单起见, 我们假定 $f \in C(0,+\infty)$, 且任给 $b > 0$, f 在 $(0,b]$ 中均绝对可积.

基本的想法是先解所谓的齐次方程

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1y'(t) + a_0y(t) = 0, \tag{15.12}$$

然后再找出非齐次方程 (15.10) 的一个特解即可. 注意到如果 y 满足齐次方程, 则由 a_i 均为常数可知, y', y'', $\cdots$, $y^{(n-1)}$ 等也满足齐次方程. 我们来寻找一个所谓的基本解, 通过求导就可以由它生成齐次方程的通解. 事实上, 考虑上述齐次方程的解, 使得它满足如下初值:

$$y(0) = 0,\ y'(0) = 0, \cdots,\ y^{(n-2)}(0) = 0,\ y^{(n-1)}(0) = 1. \tag{15.13}$$

为了找到这个解, 记 $Y(s) = \mathcal{L}(y)(s)$，对齐次方程做 Laplace 变换, 利用 (15.6) 式可得

$$Y(s)\big[s^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0\big] = 1.$$

记 $p(s)=s^n+a_{n-1}s^{n-1}+\cdots+a_1s+a_0$, 称为齐次方程的特征多项式. 由 (4.7) 可得

$$Y(s)=\frac{1}{p(s)}=\sum_{i=1}^{k}\frac{1}{p'(z_i)}\frac{1}{s-z_i}+\sum_{j=1}^{l}\sum_{\nu=0}^{m_j-1}\frac{c_{j,\nu}}{(s-w_j)^{m_j-\nu}},$$

其中 $\{z_i\}_{i=1}^k$, $\{w_j\}_{j=1}^l$ 分别是 $p(s)$ 的单根和重根, 系数 $c_{j,\nu}$ 由 (4.8) 式所确定. 由例 15.4.6 可得 $Y(s)=\mathcal{L}(g)(s)$, 其中

$$g(t)=\sum_{i=1}^{k}\frac{1}{p'(z_i)}\mathrm{e}^{z_it}+\sum_{j=1}^{l}\sum_{\nu=0}^{m_j-1}\frac{c_{j,\nu}t^{m_j-1-\nu}}{(m_j-1-\nu)!}\mathrm{e}^{w_jt}. \tag{15.14}$$

我们来验证 $y=g(t)$ 的确满足 (15.12) 和 (15.13). 事实上, 由 $p(s)\mathcal{L}(g)(s)=1$ 以及 (15.6) 式可得

$$\begin{aligned}&\mathcal{L}\big(g^{(n)}+a_{n-1}g^{(n-1)}+\cdots+a_1g'+a_0g\big)(s)\\&=1-g(0)s^{n-1}-\big[g'(0)+a_{n-1}g(0)\big]s^{n-2}\\&\quad-\big[g''(0)+a_{n-1}g'(0)+a_{n-2}\big]s^{n-3}-\cdots\\&\quad-\big[g^{(n-1)}+a_{n-1}g^{(n-2)}+\cdots+a_1g(0)\big].\end{aligned}$$

注意到 Laplace 积分在无穷远处的极限必定为零(见本节习题 1), 这说明上式右端的多项式一定恒等于零, 因此 $g(t)$ 满足初始条件 (15.13). 再由 Laplace 变换的唯一性即知 $g(t)$ 满足方程 (15.12).

有了 $g(t)$, 可以令

$$\begin{aligned}z(t)=&\,y_0g^{(n-1)}(t)+(y_1+a_{n-1}y_0)g^{(n-2)}(t)+(y_2+a_{n-1}y_1+a_{n-2}y_0)g^{(n-3)}(t)\\&+\cdots+(y_{n-1}+a_{n-1}y_{n-2}+\cdots+a_1y_0)g(t),\end{aligned}$$

不难验证它是齐次方程 (15.12) 满足初始条件 (15.11) 的唯一解.

我们再来寻找非齐次方程 (15.10) 满足初始条件 $y(0)=y'(0)=\cdots=y^{(n-1)}(0)=0$ 的解. 记 $Y(s)=\mathcal{L}(y)(s)$, 对 (15.10) 做 Laplace 变换可得 $p(s)Y(s)=\mathcal{L}(f)(s)$, 因此

$$Y(s)=\frac{1}{p(s)}\mathcal{L}(f)(s)=\mathcal{L}(g)(s)\cdot\mathcal{L}(f)(s)=\mathcal{L}(g*f)(s),$$

记

$$w(t)=(g*f)(t)=\int_0^t g(t-u)f(u)\,\mathrm{d}u,$$

可以验证 $y=w(t)$ 的确满足非齐次方程. 最后, $y=z(t)+w(t)$ 就是非齐次方程 (15.10) 满足初始条件 (15.11) 的唯一解.

习题 15.4

1. 设 f 在 $(0,+\infty)$ 中的广义积分收敛, 证明:

(1) $\lim\limits_{s\to+\infty}\mathcal{L}(f)(s)=0$;

(2) 进一步, 若 f 在 0 处的右极限 $f(0^+)$ 存在, 则 $\lim\limits_{s\to+\infty}s\mathcal{L}(f)(s)=f(0^+)$.

2. 设 f 在 $(0,+\infty)$ 中有定义, 如果任给 $b>0$, f 在 $(0,b]$ 中的积分均存在, 且 $\lim\limits_{t\to+\infty}f(t)=f(+\infty)\in\mathbb{R}$. 证明: $\lim\limits_{s\to0^+}s\mathcal{L}(f)(s)=f(+\infty)$.

3. 设 f 在 $(0,+\infty)$ 中有定义, 且

(i) 任给 $b>0$, f 在 $(0,b]$ 中的积分均存在;

(ii) $\mathrm{e}^{-t}f(t)$ 在 $(0,+\infty)$ 中的广义积分收敛.

记 $F(t)=\displaystyle\int_0^t f(u)\,\mathrm{d}u$, 证明: $\lim\limits_{t\to+\infty}\mathrm{e}^{-t}F(t)=0$, 且

$$\int_0^{+\infty}\mathrm{e}^{-t}F(t)\,\mathrm{d}t=\int_0^{+\infty}\mathrm{e}^{-t}f(t)\,\mathrm{d}t.$$

4. 设 f 在 $(0,+\infty)$ 中 n 阶可导, 且存在 $A,M_0,M_n>0$, 使得当 $t\geq A$ 时 $|f(t)|\leq M_0$, $\left|f^{(n)}(t)\right|\leq M_n$. 证明: 存在与 f 无关的常数 $C(n)$, 使得当 $t\geq A$ 时

$$\left|f^{(i)}(t)\right|\leq C(n)M_0^{1-\frac{i}{n}}M_n^{\frac{i}{n}},\ \ i=0,1,\cdots,n.$$

5. 解如下初值问题:

$$(1)\ y''-3y'-10y=2,\ y(0)=1,\ y'(0)=2;$$
$$(2)\ y''+y=\cos(2t),\ y(0)=2,\ y'(0)=1.$$

6. 解如下初值问题:

$$y'(t)+z'(t)+y(t)+z(t)=1,\ y'(t)+z(t)=\mathrm{e}^t,\ y(0)=-1,\ z(0)=2.$$

7. 解如下初值问题:

$$(1)\ y''(t)+ty'(t)-2y(t)=2,\ y(0)=y'(0)=0;$$
$$(2)\ y''(t)+ty(t)-2y=4,\ y(0)=-1,\ y'(0)=0.$$

8. 设 $\alpha>0$, 数列 $\{a_n\}$ 满足条件 $a_n=O\big(\alpha^n\big)$ $(n\to\infty)$. 记 $f(t)=\sum\limits_{n=0}^{\infty}\frac{a_n}{n!}t^n$, 证明: 当 $s>\alpha$ 时 $\mathcal{L}(f)(s)=\sum\limits_{n=0}^{\infty}a_ns^{-n-1}$.

9. 记 $J_0(t)=\sum\limits_{n=0}^{\infty}\frac{(-1)^n}{(n!)^2}\left(\frac{t}{2}\right)^{2n}$, 证明 $(J_0*J_0)(t)=\sin t$.

10. 解积分方程 $y(t)=1+\displaystyle\int_0^t\cos(t-u)y(u)\,\mathrm{d}u$.

11. 设 $p(t)$ 是 n 次复系数多项式, $\{z_i\}_{i=1}^n$ 是其 n 个互不相同的根. 当 $n>1$ 时, 利用 (15.14) 式证明

$$\sum_{i=1}^n \frac{z_i^{n-1}}{p'(z_i)}=1,\ \sum_{i=1}^n \frac{z_i^j}{p'(z_i)}=0,\ j=0,1,\cdots,n-2.$$

12. 设 f 在 $(0,+\infty)$ 中可导, g 在 $(0,+\infty)$ 中连续. 如果 0 为 f' 唯一可能的瑕点, 且任给 $b>0$, f' 和 g 均在 $(0,b]$ 中绝对可积, 则 $f*g$ 在 $(0,+\infty)$ 中可导, 且

$$(f*g)'(t)=\int_0^t f'(t-u)g(u)\,\mathrm{d}u+f\big(0^+\big)g(t).$$

15.5 Fourier 变换回顾

P498 页, Fourier 变换的定义后加一段:

Fourier 变换可视为周期函数的 Fourier 系数的连续版本. 如果将 $f(x)$ 看成是时间 x 为变量的函数, $\hat{f}(\omega)$ 就是以频率 ω 为变量的函数. 在工程应用中, 这样的变换常用于频谱分析、信号处理等.

15.6 补充材料

补充参考文献:

Edouard Goursat, A Course in Mathematical Analysis, Vol. I, Translated by Earle Raymond Hedrick, 1904, Ginn.

G.H. Hardy and J.E. Littlewood, Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive, Proc. London Math. Soc. (2) 13 (1914), 174 - 191.

J. Karamata, Über die Hardy-Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes, Math. Z. 32 (1930), 319 - 320.

H. D. Kloosterman, On the convergence of series summable (C, r) and on the magnitude of the derivatives of a function of a real variable, J. London Math. Soc., 15(1940), 91-96.

R. Schmidt, Über divergente Folgen und lineare Mittelbildungen, Math. Z. 22 (1925), 89 - 152.

T. Vijayaraghavan, A Tauberian theorem, J. London Math. Soc. (1) 1 (1926), 113 - 120.

H. Wielandt, Zur Umkehrung des Abelschen Stetigkeitssatzes, J. Reine Angew Math. 56 (1952), 27 - 39.

G. Doetsch, Introduction to the Theory and Applications of the Laplace Transformation, Spring-Verlag, 1974.