A Lillte More Analysis

致每一个阳光下闪烁着七彩光芒的泡沫

To every bubble glittering with colorful lights under the sun

序

TODO

缘起

王普
浙江大学计算机科学与技术学院
@zju.edu.cn
2023 年 10 月

致读者

TODO

目录

1 线性代数与微积分

1.1 微分学

为了得到定义在 $\mathbb{R}^n$ 上的函数的导数与微分的定义, 回顾一下一元函数微分学的一些内容, 看看怎样把 $n=1$ 的情形解释一下, 进而可以自然地推广到 $n>1$ 的情况.

设 f 是定义在 $(a,b)\subset\mathbb{R}$ 上的函数,$x\in(a,b)$, 那么 f 在 x 处的导数定义为

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

为了更好地处理多元函数的微分学, 我们先考虑一下线性映射的范数:

定义 1.1

线性映射的范数 对于线性映射 $A\in L(\mathbb{R}^n,\mathbb{R}^m)$ 与 $x\in\mathbb{R}^n$, 定义 A 的**范数** $\|A\|$ 为 $\|A\|=\sup\limits_{|x|\leqslant 1}|Ax|$.

这样定义的范数其实反映出了线性映射伸张向量的能力, 根据

$$\|A\|=\sup_{x\neq 0}|A\frac{x}{|x|}|=\sup_{x\neq 0}\frac{|Ax|}{|x|},$$

以及

$$|Ax|=|x||A\frac{x}{|x|}|\leqslant|x|\|A\|,$$

这样就可以定义线性映射 A 在 x 方向的**伸张系数**$\frac{|Ax|}{|x|}$, 并且 $\|A\|$ 的几何意义是一切方向的伸张系数的上确界.

注意到, 对于一般的线性空间 X, 如果范数选取得当, 就算是 X 是无穷维的, 我们对线性映射定义的范数也是有意义的, 但我们敏锐的嗅觉告诉我们, 既然加上了“无穷维”这个

不容易处理的条件, 无穷维线性空间之间的线性映射的性质就会变得更加复杂, 我们暂且不对其继续深入讨论.

作为范数, 首先要满足下面四条性质：

定理 1.1

设 $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $B \in L(\mathbb{R}^l, \mathbb{R}^n)$, 则

1. $\|A\| \geqslant 0$, 且 $\|A\| = 0$ 当且仅当 $A = 0$;
2. 若 $\lambda \in R$,$\|\lambda A\| = |\lambda|\|A\|$;
3. $\|A + B\| \leqslant \|A\| + \|B\|$;
4. $\|AB\| \leqslant \|A\|\|B\|$.

证明

1. 显然;
2. $\|\lambda A\| = \sup\limits_{|x|\leqslant 1} |\lambda Ax| = |\lambda| \sup\limits_{|x|\leqslant 1} |Ax| = |\lambda|\|A\|$;
3. $\|A+B\| = \sup\limits_{|x|\leqslant 1} |(A+B)x| = \sup\limits_{|x|\leqslant 1} |Ax+Bx| \leqslant \sup\limits_{|x|\leqslant 1} |Ax| + \sup\limits_{|x|\leqslant 1} |Bx| = \|A\|+\|B\|$;
4. $\|AB\| = \sup\limits_{|x|\leqslant 1} |ABx| = \sup\limits_{|x|\leqslant 1} |A(Bx)| \leqslant \sup\limits_{|x|\leqslant 1} \|A\||Bx| \leqslant \|A\| \sup\limits_{|x|\leqslant 1} |Bx| = \|A\|\|B\|$.

□

例 1.1

另一种范数 对于线性映射 $A \in L(\mathbb{R}^m, \mathbb{R}^n)$, 其矩阵表示为 $M(A)$, 定义 A 的范数为

$$\|A\|_2 = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

请验证其满足范数的四条性质.

证明

□

线性映射终归也是一个“线性的函数”, 所以自然可以考虑它的有界性和连续性：

定义 1.2

赋范线性空间上的有界映射 设 A 是赋范线性空间 X 到赋范线性空间 Y 的线性映射, 如果存在常数 $M \leqslant 0$ 使得对于任意 $x \in X$, 都有

$$\|Ax\| \leqslant M|x|,$$

那么称 A 是**有界映射**.

定理 1.2

对于线性赋范空间 X 和 Y, 从 X 到 Y 的线性映射 A 是有界映射的充分必要条件是 A 将所有有界集映射为有界集.

证明

设 A 将所有有界集映射为有界集, 那么 A 将单位球面 $S = \{x \in X | \|x\| = 1\}$ 映射为一个有界集, 那么存在常数 $M \leqslant 0$ 使得对于 $y \in S$ 有 $\|Ay\| \leqslant M$, 当 $x = 0$ 的时候,$\|Ax\| \leqslant M|x|$ 自然成立, 当 $x \neq 0$ 的时候, 取 $y = \frac{x}{|x|}$, 那么

$$\frac{\|Ax\|}{|x|} = \|A\frac{x}{|x|}\| = \|Ay\| \leqslant M.$$

这就说明了 A 为有界映射.

反过来, 如果 A 是有界映射, 设 B 是 X 中的有界集, 那么存在常数 N 使得对于任意 $x \in B$ 都有 $|x| \leqslant N$, 那么对于任意 $x \in B$, 有

$$\|Ax\| \leqslant M|x| \leqslant MN,$$

这就说明了 A 将有界集映射为有界集. □

定义 1.3

距离空间上的连续映射 假设 (X, d_X) 和 (Y, d_Y) 是两个距离空间,$f : X \to Y$ 是这两个距离空间的映射. 假设 $x_0 \in X$,$y_0 = f(x_0) \in Y$. 如果对于任意的 $\varepsilon > 0$, 都存在 $\delta > 0$ 使得对于任意满足 $d_X(x, x_0) < \delta$ 的 $x \in X$, 都有 $d_Y(f(x), y_0) < \varepsilon$, 那么称 f 在 x_0 处**连续**. 如果 f 在 X 的每一点都连续, 那么称 f 是**连续映射**.

赋范线性空间 $(X, \|\cdot\|)$ 当然是度量空间, 对于 $x, y \in X$, 只需要定义 $d(x, y) = \|x - y\|$ 就可以得到一个度量, 而线性映射相比于一般的映射, 理应具有更强的性质, 我们下面证明这一点:

定理 1.3

赋范线性空间上的连续映射 设 A 是赋范线性空间 X 到赋范线性空间 Y 上的线性映射, 假设 A 在某一点 $x_0 \in X$ 处连续, 那么 A 是 X 上的连续映射.

证明

□

定理 1.4

有界性和连续性的等价性 从赋范线性空间 X 到赋范线性空间 Y 上的线性映射 A 是有界映射的充分必要条件是 A 是连续映射.

更进一步的, 如果 $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ 那么可以直接得到 A 是有界且一致连续的. 当我们取 $\mathbb{R}^n$ 上的一组标准基 $\{e_1, e_2, \cdots, e_n\}$, 设 $x \in \mathbb{R}^N$,$|x| \leqslant 1$, 那么 x 就可以表示为 $x = \sum\limits_{i=1}^{n} x_i e_i$,$|x_i| \leqslant 1$. 于是

$$|Ax| = |\sum_{i=1}^{n} x_i A e_i| \leqslant \sum_{i=1}^{n} |x_i||Ae_i| \leqslant \sum_{i=1}^{n} |Ae_i|.$$

所以

$$\|A\| \leqslant \sum_{i=1}^{n} |Ae_i| < \infty.$$

并且因为当 $x, y \in \mathbb{R}^n$ 时,$|Ax - Ay| \leqslant \|A\||x - y|$, 所以 A 是一致连续的.

这下我们在 $L(\mathbb{R}^n, \mathbb{R}^m)$ 上有了范数, 自然可以定义度量、开集、连续等概念, 下面对于可逆线性算子的刻画就利用了这些概念:

定理 1.5

设 Ω 为 $\mathbb{R}^n$ 上所有可逆线性算子的集合.

1. 若 $A \in \Omega$,$B \in L(\mathbb{R}^n)$, 并且

$$\|B - A\| \cdot \|A^{-1}\| < 1,$$

那么 $B \in \Omega$;

2. Ω 是 $L(\mathbb{R}^n)$ 的开集, 映射 $A \to A^{-1}$ 是 Ω 上的连续映射.

证明

1.

2.

$\square$

定义 1.4

多元函数的微分

$$\lim_{h\to 0}\frac{|f(x_0+h)-f(x_0)-Ah|}{|h|}=0.$$

回忆一下从一元函数到多元向量值函数的推广过程：在一元函数, 我们研究的主要是导数, 导数的实际意义是函数的变化率, 一元函数的自变量只有一个, 表示自变量的点只能在直线上变动, 移动的方向也只有左右两个方向; 而当我们想在多元函数上讨论函数的变化率的时候, 我们忽然发现, 表示自变量的点可以在一个区域内任意移动了, 不仅可以移动距离, 而且可以按任意方向移动同一段距离. 因此函数的变化不仅与移动的距离有关, 还与移动的方向有关, 所以我们尝试考虑某一个方向上函数的变化率, 当方向被限制在坐标轴方向的时候, 我们就得到**偏导数**, 当我们考虑某一个方向 $\boldsymbol{v}$ 上的变化率的时候, 我们就得到**方向导数**. 可是我们不想仅限于此, 我们注意到可以将方向 $\boldsymbol{v}$ 分解为坐标轴方向的向量的和, 而我们恰恰有函数在坐标轴方向上的变化率——偏导数, 所以我们可以将函数在方向 $\boldsymbol{v}$ 上的变化率表示为函数在坐标轴方向上的变化率的线性组合, 这也就是微分的想法.

定理 1.6

方向导数与微分 对函数 $f:\mathbb{R}^n\to\mathbb{R}$, 设 f 在 x_0 处可微, 那么 f 在 x_0 处的方向导数都存在, 并且有

$$\frac{\partial f}{\partial u}(x_0)=\mathrm{d}f(x_0)(u).$$

并且 f 在 x_0 的偏导数 $f_{x_i}(x_0)$ 都存在, 记 f 在 x_0 处的**梯度**为

$$\nabla f(x_0)=(f_{x_1}(x_0),f_{x_2}(x_0),\cdots,f_{x_n}(x_0)),$$

那么就有

$$\mathrm{d}f(x_0)(u)=\nabla f(x_0)\cdot u.$$

这就表明了梯度其实是微分的矩阵表示.

证明

□

既然函数就是两个集合之间的映射, 迄今为止我们考虑的一直是 $\mathbb{R}^n$ 上的实值函数, 既然线性映射都可以是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的, 那么函数当然也可以这样, 这就是**向量值函数**.

向量值函数只是只是对于函数到达空间的简单推广, 甚至在一般的度量空间之间的映射的视角, 这仅仅只是一个特例而已. 我们知道, 对于一元函数与多元函数在某一点处的微分其实是一个线性映射 L, 使得其是在这点处的最佳线性逼近. 换句话说就是

$$f(x) - f(x_0) = L(x - x_0) + o(|x - x_0|) \quad (f \to x_0).$$

我们按照这样的思路, 可以定义多元向量值函数的微分：

定义 1.5

向量值函数的微分 设 $D \subset \mathbb{R}^n$ 是开集,$f : D \to \mathbb{R}^m$ 是一个向量值的多元函数函数,$x_0 \in D$. 如果存在一个线性映射 $L \in L(\mathbb{R}^n, \mathbb{R}^m)$ 使得在 x_0 附近成立

$$f(x) - f(x_0) = L(x - x_0) + o(|x - x_0|) \quad (f \to x_0),$$

那么称 f 在 x_0 处**可微**, 并且称 L 为 f 在 x_0 处的**微分**, 记作 $df(x_0)$.

从经验上看, 向量值函数与其微分的出现是自然而必然的, 对于一个一般的多元函数 $f : \mathbb{R}^n \to \mathbb{R}$, 我们定义了它的微分是一个线性映射 $\mathrm{d}f \in L(\mathbb{R}^n, \mathbb{R})$, 按照一维的理论, 对一次导数求导就得到了二次导数, 所以我们应该对 $x \mapsto \mathrm{d}f$ 求微分, 而这实际上是一个 $\mathbb{R}^n \to L(\mathbb{R}^\times, \mathbb{R}) \simeq \mathbb{R}^n$ 的函数. 你看, 这就是向量值函数！所以我们自然要定义向量值函数与其微分.

再者, 从换元的角度, 对于一个二元函数 $f(x, y)$, 我们可以随便找一个换元 $x = x(u, v)$ 与 $y = y(u, v)$, 那么 $f(x, y)$ 就变成了 $f(x(u, v), y(u, v))$, 将换元后的函数视作一个复合函数 $f \circ \varphi$, 那么 φ 必须是从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的, 也就是 $\varphi(u, v) = (x(u, v), y(u, v))$. 这就是向量值函数的另一个实例.

从一元函数到多元函数, 对微分的定义其实迈上了一个新的台阶, 原因其实是自变量的每一个分量“杂糅”在一起, 我们需要单独将每一个分量“拎出来”单独讨论, 所以这就是偏导数与全微分的来源. 但是对于向量值函数而言, 它的每一个分量并没有什么特别强的联系——这是因为我们有了坐标, 并且这严重依赖于坐标系的选取——所以我们可以将从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的向量值函数拆成 m 个分量, 每一个分量都是从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的函数. 在这个角度

上, 我们可以直接但是不严谨地得出：向量值函数 f 在 x_0 处可微当且仅当其每一个分量 f_i 在 x_0 处都可微, 并且其微分为

$$\mathrm{d}f(x_0)=Jf(x_0)=\begin{pmatrix}\nabla f_1(x_0)\\ \nabla f_2(x_0)\\ \vdots\\ \nabla f_n(x_0)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0)\\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0)\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0)\end{pmatrix}.$$

其中矩阵 $Jf(x_0)$ 是 f 在 x_0 的微分的矩阵表示, 被称为 Jacobi 矩阵. 虽然这非常符合直觉, 但是我们还是需要严谨的证明：

定理 1.7

微分的计算 假设 $V=\mathbb{R}^n$ 和 $W=\mathbb{R}^m$, 我们分别取坐标系 $\{x_i\}_{i=1,2,\cdots,n}$ 和 $\{y_j\}_{j=1,2,\cdots,m}$, 考虑函数

$$f:V\to W,\ \ x\mapsto f(x)=(f_1(x_1,x_2,\cdots,x_n),\cdots,f_m(x_1,x_2,\cdots,x_n)).$$

那么 f 在 x_0 处可微当且仅当每一个 f_i 在 x_0 处都可微, 并且微分 $\mathrm{d}f(x_0)\in L(\mathbb{R}^n,\mathbb{R}^m)$ 的矩阵表示为

$$\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{\substack{i=1,2,\cdots,m\\ j=1,2,\cdots,n}}.$$

证明

$\square$

另外需要注意的是,

定理 1.8

链式法则 设 D 和 Δ 分别为 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 上的开集,$f:D\to\mathbb{R}^m$ 和 $g:\Delta\to\mathbb{R}^l$ 为向量值函数, 且 $f(D)\subset\Delta$. 如果 f 在 $x_0\in D$ 中可微,g 在 $y_0=f(x_0)$ 处可微, 那么复合函数 $h=g\circ f$ 在 x_0 处可微, 并且有

$$Jh(x_0)=Jg(y_0)Jf(x_0).$$

（画两个交换图）

证明

□

推论 1.1

反函数的微分 给定区域 $D \subset \mathbb{R}^n$ 和 $\Delta \subset \mathbb{R}^m$ 与向量值函数 $f: D \to \Delta$, 如果 f 是一一映射并且其逆映射 $f^{-1}: \Delta \to D$ 是可微的, 那么有

(1) $m = n$;

(2) $Jf(x)$ 是可逆的, 并且有

$$J(f^{-1})(y) = [Jf(x)|_{x=f^{-1}(y)}]^{-1}.$$

证明

我们令 $\Omega = D$,$g = f^{-1}: \Omega \to \Delta$, 那么有 $g \circ f = \mathrm{id}_\Omega$, 根据链式法则有

$$\mathrm{Id}_\Omega = I_n = J(g \circ f)(x_0) = Jg(y_0) \cdot Jf(x_0),$$

反过来用 f^{-1} 替代 f 可得另一侧

$$\mathrm{Id}_\Delta = I_m = J(f \circ g)(y_0) = Jf(x_0) \cdot Jg(y_0).$$

根据线性代数的知识, 很容易可以看出 $Jf(x)$ 和 $Jg(g)$ 都需要是一个方阵, 这就是 (1). 并且也足以说明 $Jf(x)$ 是可逆的, 也就是 $Jg(y) = (Jf(x))^{-1}$. □

与一元函数的微分中值定理相似, 我们也有多元函数的微分中值定理:

定理 1.9

微分中值定理 设 $D \subset \mathbb{R}^n$ 为凸域, 函数 $f: D \to \mathbb{R}$ 在 D 中处处可微, 则任给 $x, y \in D$, 存在 $\theta \in (0,1)$ 使得

$$f(y) - f(x) = \nabla f(\xi) \cdot (y - x), \quad \xi = \theta x + (1-\theta)y.$$

证明

□

一个很自然的问题就是微分中值定理可不可以推广到向量值函数？设 $D \subset \mathbb{R}^n$ 为凸域,$f: D \to \mathbb{R}^m$ 是一个向量值函数,$x, y \in D$, 对 f 的每一个分量 f_i 应用微分中值定理可以得到

$$f_i(x) - f_i(y) = \nabla f_i(\xi_i) \cdot (x - y),$$

其中 $\xi_i \in D$, 注意到这些 xi_i 未必相同, 比如对于函数 $f: \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t^2, t^3)$, 取 $x = 1$,$y = 0$, 通过计算得知 $\xi_1 = \dfrac{1}{2}$,$\xi_2 = \pm\dfrac{1}{\sqrt{3}}$, 因此 $\xi_1 \neq \xi_2$. 这就表明我们不能指望着 $f(x) - f(y) = Jf(\xi)(x - y)$ 对于某一个 ξ 成立. 但是我们却可以有另一个不错的估计, 这就是**拟微分中值定理**.

定理 1.10

拟微分中值定理 设 $D \subset \mathbb{R}^n$ 为凸域,$f: D \to \mathbb{R}^m$ 在 D 中处处可微, 则任给 $x, y \in D$, 存在 $\xi \in D$ 使得

$$|f(x) - f(y)| \leqslant \|Jf(\xi)\| \cdot |x - y|.$$

证明的基本想法是对 f 的分量的线性组合应用微分中值定理, 然后进行放缩.

证明

不妨设 $f(x) \neq f(y)$, 任意取定 $\mathbb{R}^m$ 中的单位向量 $u = (u_1, u_2, \cdots, u_m)$, 记

$$g = u \cdot f = \sum_{i=1}^{m} u_i f_i.$$

则 g 为 D 中的可微函数, 并且对 x, y 存在 $\xi \in D$ 使得

$$g(x) - g(y) = \nabla g(\xi) \cdot (x - y).$$

且

$$\nabla g(\xi) = \sum_{i=1}^{m} u_i \nabla f_i(\xi).$$

利用 Cauchy-Schwarz 不等式, 有

$$|\nabla g(\xi)| \leqslant \sum_{i=1}^{m} |u_i||\nabla f_i(\xi)| \leqslant \left(\sum_{i=1}^{m} |u_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{m} |\nabla f_i(\xi)|^2\right)^{\frac{1}{2}} = \|Jf(\xi)\|.$$

由 $g(x) - g(y) = u \cdot (f(x) - f(y))$, 我们有

$$|u \cdot [f(x) - f(y)]| = |g(x) - g(y)| \leqslant |\nabla g(\xi)| \cdot |x - y| \leqslant \|Jf(\xi)\| \cdot |x - y|.$$

取 $u=\dfrac{f(x)-f(y)}{|f(x)-f(y)|}$ 就完成了证明. □

拟微分中值定理告诉我们, 雅各布矩阵的范数可以用来估计向量值函数的改变量. 下面的结果给出了向量值函数与其线性化映射之间的误差估计.

推论 1.2

设 $D\subset\mathbb{R}^n$ 为开集,$f:D\to\mathbb{R}^m$ 在 D 中处处可微, 且 $Jf(x)$ 关于 x 连续, 如果 $C\subset D$ 为紧凸集, 则任给 $\varepsilon>0$, 存在 $\delta>0$ 使得对于任意 $u,v\in C$, 只要 $|u-v|<\delta$, 就有

$$|f(u)-f(v)-Jf(v)(u-v)|\leqslant\varepsilon|u-v|.$$

证明

□

既然我们可以对一个函数的反函数求微分, 回想起对于一元函数, 如果它可微并且导数处处非零, 那么该函数存在反函数并且反函数可微. 那么一个自然的问题就是：在什么样的条件下, 多元向量值函数存在反函数并且反函数可微呢？这就是**反函数定理**. 在考虑反函数定理之前, 我们先看一个小小的比喻.

例 1.2

设 A 是一个 n 阶方阵, 如果 $\|A\|<1$, 则 I_n-A 可逆.

证明

设 $u\in\mathbb{R}^n$, 如果 $(I_n-A)u=0$, 则 $|u|=|Au|\leqslant\|A\||u|<|u|$, 那么 u 只能等于 0, 所以 I_n-A 是单射, 所以其可逆. □

以分析学的视角看：给定 $v\in\mathbb{R}^n$, 解方程

$$(I_n-A)x=v,\ \text{亦即}\ x=Ax+u.$$

我们先考虑映射 $\varphi(x)=Ax+v$, 当 $x,y\in\mathbb{R}^n$ 时,

$$|\varphi(x)-\varphi(y)|=|A(x-y)|\leqslant\|A\|\cdot|x-y|.$$

这个性质很不错, 回忆一下**压缩映射原理**：

定理 1.11

压缩映射原理 设 (X,d) 是一个完备的度量空间,C 是一个非空闭子集,$f:C\to C$ 是一个压缩映射, 即存在常数 $0\leqslant k<1$ 使得对于任意 $x,y\in C$ 都有

$$d(f(x),f(y))\leqslant kd(x,y),$$

那么 f 在 C 上有唯一的不动点. 即存在唯一的 $x_*\in C$ 使得 $f(x_*)=x_*$.

证明

任取 $a_0\in C$, 当 $n\geqslant 1$ 时, 利用 $a_n=f(a_{n-1})$ 可以递归定义 C 中的数列 $\{a_n\}$, 我们来说明 $\{a_n\}$ 为 Cauchy 列. 根据题设, 存在 $0\leqslant L<1$, 使得当 $n\geqslant 1$ 时,

$$d(a_{n+1},a_n)=d(f(a_n),f(a_{n-1}))\leqslant Ld(a_n,a_{n-1}).$$

由上式与归纳法可知 $d(a_n,a_{n-1})\leqslant L^{n-1}d(a_1,a_0)$ 对 $n\geqslant 1$ 均成立, 当 $m>n$ 时, 有

$$\begin{aligned}d(a_m,a_n)&\leqslant d(a_m,a_{m-1})+d(a_{m-1},a_{m-2})+\cdots+d(a_{n+1},a_n)\\&\leqslant (L^{m-1}+L^{m-2}+\cdots+L^n)d(a_1,a_0)\\&\leqslant \frac{L^n}{1-L}d(a_1,a_0).\end{aligned}$$

由于 $L<1$, 所以当 $n\to\infty$ 时,$L^n\to 0$, 所以 $\{a_n\}$ 为 Cauchy 列, 其极限记为 a_*, 则 $a_*\in C$, 由 $a_n=f(a_{n-1})$ 与 f 的连续性可知 $f(a_*)=a_*$.

下面证明唯一性：若另有 $a\in C$ 使得 $f(a)=a$, 则

$$d(a,a_*)=d(f(a),f(a_*))\leqslant Ld(a,a_*),$$

由于 $0\leqslant L<1$, 所以 $d(a,a_*)=0$, 即 $a=a_*$. □

所以这是一个压缩映射, 因此 φ 有唯一的不动点, 即存在唯一的 $x_*\in\mathbb{R}^n$ 使得 $x_*=Ax_*+v$, 即 $(I_n-A)x_*=v$, 所以 I_n-A 为单射, 即可逆.

对于恒同映射 I_n 来讲, 范数小于 1 的映射相对于它是一个微小扰动, 上面的例子告诉我们, 对于一个像恒同映射这样的可逆映射施加一个微小扰动得到的映射是可逆的. 在一般的向量值函数中, 微小扰动的例子丰富极了, 因为每一个在某一点 x_0 可微的向量值函数在该点都可以将其线性化, 也就是分解成 $L(x-x_0)+o(|x-x_0|)$, 将后一项视作微小扰动 $\varepsilon(x-x_0)$, 所以我们可以猜测：如果向量值函数的微分在某一点 x_0 可逆, 那么存在 x_0 的邻域 U 与 $y_0=f(x_0)$ 的邻域 V 使得 $f|_U:U\to V$ 是可逆的. 不过我们需要对这个函数添加一定的可微性条件. 一个向量值函数是 C^k 的是指其每个分量都是一个 k 次连续可微函

数. 这就是**反函数定理**的内容.

定理 1.12

反函数定理 设 $D \subset \mathbb{R}^n$ 为开集,$f: D \to \mathbb{R}^n$ 为 C^k 的映射 $(k \geqslant 1)$,$x_0 \in D$. 如果 $\det Jf(x_0) \neq 0$, 那么存在 x_0 的邻域 $U \subset D$ 与 $f(x_0)$ 的邻域 $V \subset D$ 使得 $f|_U : U \to V$ 是可逆的, 并且其逆也是 C^k 的.

证明

由于平移不改变函数的可微性与可逆性, 所以我们不失一般性地设 $(x_0, y_0) = (0, 0)$, 记 L 为 f 在 $x_0 = 0$ 处的微分, 则 L 可逆, 那么根据链式法则,$L^{-1} \circ f$ 在 x_0 处的微分是恒同映射. 如果欲证结论对于 $L^{-1} \circ f$ 成立, 那么套上一个可逆的线性映射 L 之后, 对于 f 也一定成立. 因此我们可以直接就假设 $Jf(x_0) = I_n$.

在 $x_0 = 0$ 附近,g 是恒同映射的小扰动:

$$g(x) = f(x) - x, \quad Jg(0) = 0.$$

扰动项 g 自然也是 C^k 的, 因此存在 $\delta > 0$, 使得

$$\|Jg(x)\| \leqslant \frac{1}{2}, \quad \forall x \in \overline{B(0, \delta)} \subset D.$$

由拟微分中值定理可知:

$$|g(x_1) - g(x_2)| \leqslant \frac{1}{2}|x_1 - x_2|, \quad \forall x_1, x_2 \in \overline{B(0, \delta)}.$$

给定 $y \in B(0, \frac{\delta}{2})$, 我们在 $B(0, \delta)$ 中解方程

$$f(x) = y, \quad \text{亦即} \quad x = y - g(x).$$

特别地, 取 $x_2 = 0$, 那么

$$|g(x_1)| \leqslant \frac{1}{2}|x_1|.$$

记 $\varphi(x) = y - g(x)$, 当 $x \in \overline{B(0, \delta)}$ 时

$$|\varphi(x)| \leqslant |y| + |g(x)| < \frac{\delta}{2} + \frac{1}{2} \leqslant \delta.$$

这说明 $\varphi(\overline{B(0, \delta)}) \subset B(0, \delta)$. 当 $x_1, x_2 \in \overline{B(0, \delta)}$ 时

$$|\varphi(x_1) - \varphi(x_2)| = |g(x_2) - g(x_1)| \leqslant \frac{1}{2}|x_2 - x_1|.$$

根据压缩映射原理,$f(x)=y$ 或 $x=y-g(x)$ 在 $\overline{B(0,\delta)}$ 中有唯一解, 记作 x_y. 且 $|x_y|\leqslant\delta$, 也就是 $x_y\in B(0,\delta)$. 令

$$U=f^{-1}(B(0,\frac{\delta}{2}))\cap B(0,\delta),\ \ V=B(0,\frac{\delta}{2}),$$

则我们已经证明了 $f|_U:U\to V$ 是一一映射, 其逆映射 $h(y)=x_y$ 满足

$$y-g(h(y))=h(y)$$

下面逐步证明 h 是 C^k 的.

(1) $h:V\to U$ 是连续映射：当 $y_1,y_2\in V$ 时,

$$|h(y_1)-h(y_2)|\leqslant|y_1-y_2|+|g(h(y_1))-g(h(y_2))|\leqslant|y_1-y_2|+\frac{1}{2}|h(y_1)-h(y_2)|,$$

这说明 $|h(y_1)-h(y_2)|\leqslant 2|y_1-y_2|$, $\forall y_1,y_2\in V$. 所以 h 是连续的.

(2) $h:V\to U$ 是可微映射：设 $y_0\in V$, 则对 $y\in V$ 有

$$\begin{aligned}h(y)-h(y_0)&=(y-y_0)-[g(h(y))-g(h(y_0))]\\&=(y-y_0)-Jg(h(y_0))(h(y)-h(y_0))+o(|h(y)-h(y_0)|).\end{aligned}$$

利用 $Jf=I_n=Jg$ 与 (1), 上式可以改写为

$$Jf(h(y_0))\cdot(h(y)-h(y_0))=y-y_0+o(|y-y_0|),$$

所以

$$h(y)-h(y_0)=Jf(h(y_0))^{-1}\cdot(y-y_0)+o(|y-y_0|)$$

因而 h 在 y_0 处可微.

(3) $h:V\to U$ 是 C^k 的：由 (2) 有 $Jh(y)=Jf(h(y))^{-1}$, $\forall y\in V$. 由于 f 是 C^k 的知 Jf 是 C^{k-1} 的. 由上知 Jh 连续, 所以 h 是 C^2 的, 以此类推就可得到 h 是 C^k 的.

这样就完成了证明. □

利用反函数定理, 我们甚至可以利用它研究隐函数. 所谓“隐”函数, 对于一个函数 $f(x,y)$ 来说, 就是方程 $f(x,y)=c$ 实际上隐含的将 y 定义为 x 的函数, **隐函数定理**具体将这个函数构造了出来, 下面是一个很有启发性的例子：

例 1.3

设 $f:\mathbb{R}^2\to\mathbb{R}^2$ 为 C^k 函数,$k\geqslant 1$, $\frac{\partial f}{\partial y}(x^0,y^0)\neq 0$, 在 (x^0,y^0) 附近解方程 $f(x,y)=f(x_0,y_0)$.

解

显然 (x_0,y_0) 是方程的解, 利用反函数定理, 我们可以在 (x_0,y_0) 附近找到别的解, 因此构造函数

$$F:\mathbb{R}^2\to\mathbb{R}^2,\ \ F(x,y)=(x,f(x,y)),$$

在 (x^0,y^0) 处, 有

$$\det JF(x^0,y^0)=\begin{vmatrix} 1 & 0 \\ \dfrac{\partial f}{\partial x}(x^0,y^0) & \dfrac{\partial f}{\partial y}(x^0,y^0)\end{vmatrix}=\frac{\partial f}{\partial y}(x^0,y^0)\neq 0.$$

由反函数定理, 在 (x^0,y^0) 附近 F 为可逆映射. 于是当 x 在 x^0 附近时, 记

$$F^{-1}(x,f(x^0,y^0))=(\varphi(x),\psi(x)),$$

则 $\varphi(x)$ 与 $\psi(x)$ 均为 C^k 函数. 根据 F 的定义可得

$$(\varphi(x),f(\varphi(x),\psi(x)))=(x,f(x^0,y^0)).$$

这说明 $\varphi(x)=x$, 并且 $f(x,\psi(x))=f(x^0,y^0)$, 所以 $\psi(x)$ 就是我们要找的隐函数. 对 x 求导还可以得到

$$\frac{\partial f}{\partial x}(x,\psi(x))+\frac{\partial f}{\partial y}(x,\psi(x))\psi'(x)=0,$$

从而有

$$\psi'(x)=-[\frac{\partial f}{\partial y}(x,\psi(x))]^{-1}\frac{\partial f}{\partial x}(x,\psi(x)).$$

这就给出了对 $y=\psi(x)$ 的刻画.

这个例子可以很轻松地推广到一般形式, 只需要将偏导数换成雅各布矩阵, 这就是隐函数定理的内容. 但在正式介绍隐函数定理之前, 我们看一下它线性的说法:

例 1.4

如果 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 而 $y=(y_1,y_2,\cdots,y_m)\in\mathbb{R}^m$, 我们将一个点（或者说向量）$(x_1,\cdots,x_n,y_1,\cdots,y_m)\in\mathbb{R}^{n+m}$ 记作 (x,y), 那么每一个线性映射 $A\in L(\mathbb{R}^{n+m},\mathbb{R}^m)$ 可以被表示成两个线性映射 A_x 和 A_y 的“拼接”, 分别由

$$A_xh=A(h,0),\ \ A_yk=A(0,k)$$

确定, 其中 $h \in \mathbb{R}^n$,$k \in \mathbb{R}^m$, 并且 $A_x \in L(\mathbb{R}^n, \mathbb{R}^m)$,$A_y \in L(\mathbb{R}^m)$. 并且有

$$A(h,k) = A_x h + A_y k.$$

如果 A_y 是可逆的, 那么对于每一个 $x \in \mathbb{R}^n$, 存在唯一的 $y \in \mathbb{R}^m$ 使得 $A(x,y) = c$, 这里的 $c \in \mathbb{R}^m$. 更进一步地, 这个 y 可以利用公式

$$y = -[A_y]^{-1} A_x(x) + [A_y]^{-1} c$$

确定.

这个例子直观且很容易证明, 所以证明就此略去. 进行一些略微不严谨的思考：对于方程 $A(x,y) = c$ 来说, 我们将其视为一个线性的隐函数, 等式右边可以确定一个 m 维的空间中的点, 所以“提供了 m 维的信息”, 因此左侧就是一个“余 m 维”式子, 因此可以根据 n 维的点来确定另外一个 m 维的点. 现在来看隐函数定理, 你会惊奇地发现隐函数定理的形式与隐函数为线性时的情况是如此相似.

定理 1.13

隐函数定理 设 $W \subset \mathbb{R}^{n+m}$ 为开集,W 中的点用 (x,y) 表示, 其中 $x = (x_1, x_2, \cdots, x^n)$, $y = (y_1, y_2, \cdots, y_m)$,$f : W \to \mathbb{R}^m$ 为 C^k 映射, 用分量表示为

$$f(x,y) = (f_1(x,y), f_2(x,y), \cdots, f_m(x,y)).$$

设 $(x^0 . y^0) \in W$, 且 $\det J_y(x^0, y^0) \neq 0$, 其中 $J_y(x,y) = \left(\frac{\partial f_i}{\partial x_j}(x,y)\right)_{m \times n}$. 则存在 x^0 的开邻域 $V \subset \mathbb{R}^n$ 以及唯一的 C^k 映射 $\psi : V \to \mathbb{R}^m$, 使得

(1) $\psi(x^0) = y^0$, $f(x, \psi(x)) = f(x^0, y^0)$, $\forall x \in V$;

(2) $J\psi(x) = -[J_y f(x, \psi(x))]^{-1} J_x f(x, \psi(x))$, 其中 $J_x f(x,y) = \left(\frac{\partial f_i}{\partial x_j}(x,y)\right)_{m \times n}$.

反函数定理与隐函数定理对函数在 (x^0, y^0) 时的微分 $J_* f(x^0, y^0)$（这里的 $*$ 表示隐函数对 y 求微分）有着很高的要求, 亦即要求它是可逆也就是**满秩**的, 那么如果这个微分不满秩呢？下面的**秩定理**. 我们还是先回忆一点关于线性变换的事实.

定理 1.14

射影 对于一个线性空间 X 上的线性变换 $P \in L(X)$ 满足 $P^2 = P$, 那么称 P 为 X 里的一个**射影**. 射影满足下面的性质：

(1) 每一个 $x \in X$ 都可以唯一表示成 $x = x_1 + x_2$ 的形式, 其中 $x_1 \in \mathrm{im} A$,$x_2 \in \ker A$;

(2) 如果 X 是有限维的线性空间,X_1 是 X 内的一个线性子空间, 那么在 X 中存在一个射影 P 使得 $\mathrm{im}P = X_1$.

证明

(1) 令 $x_1 = Px$,$x_2 = x - Px$, 则 $Px_2 = Px - Px_1 = Px - P^2x = 0$, 所以 $x_2 \in \ker P$. 将 P 作用在 $x = x_1 + x_2$ 上, 有 $Px = Px_1 + Px_2 = Px_1$, 所以 $x_1 = Px$, 这就证明了表示的唯一性.

(2) 如果 $X = \{0\}$, 那么这是显然的. 于是我们假设 $\dim X_1 = k > 0$, 取 X 的某个基 $\{u_1, u_2, \cdots, u_n\}$ 使得 $\{u_1, u_2, \cdots, u_k\}$ 是 X_1 的基, 那么我们可以定义 P 使得 $P(u_1) = u_1, P(u_2) = u_2, \cdots, P(u_k) = u_k$, 并且 $P(u_{k+1}) = P(u_{k+2}) = \cdots = P(u_n) = 0$, 这样就构造出了一个射影.

□

定理 1.15

秩定理

证明

□

1.2 积分学

多元函数的积分学中的核心定理主要是 Fubini 定理与重积分换元法.

在介绍重积分换元法之前, 我们先考虑重积分换元法的核心——坐标变换. 坐标变换以链式法则为前提, 我们在很早之前就接触到了坐标变换, 比如极坐标与 Descartes 坐标系之间的转换, 我们首先使用映射的语言描述极坐标：

例 1.5

坐标变换 对于定义区域 Ω_1 和 Ω_2 如下：

$$\Omega_1 = \mathbb{R}^2 \setminus \{(x,y)|x \geqslant 0, y = 0\},\ \ \Omega_2 = \mathbb{R}_{>0} \times (0, 2\pi) = \{(r,\vartheta)|r > 0, \vartheta \in (0, 2\pi)\}.$$

我们熟悉的坐标变换 $x = r\cos\vartheta$,$y = r\sin\vartheta$ 就可以写成

$$\Phi : \Omega_2 \to \Omega_1,\ \ \Phi(r,\vartheta) = (r\cos\vartheta, r\sin\vartheta).$$

（这里要有一个图）由于在 Ω_1 上我们给定了 (x,y) 作为坐标, 在 Ω_2 上我们给定了 (r,ϑ) 作为坐标, 所以我们可以使用 Jacobi 矩阵表示上述映射的微分

$$\mathrm{d}\Phi = J\Phi = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \vartheta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \vartheta} \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -r\sin\vartheta \\ \sin\vartheta & r\cos\vartheta \end{pmatrix}$$

根据反函数定理, 这个映射当然是可逆的.

这个例子告诉我们：坐标变换允许我们将被积区域进行变换, 比如将圆或球转化成一个矩形, 这样就可以极大简化运算. 更确切来说, 对于原被积区域 Ω 上的函数 f, 我们可以定义一个映射 $\varphi:\Sigma\to\Omega$, $\varphi(\Sigma)=\Omega$, 使得我们只需要在现被积区域 Σ 对复合函数 $f\circ\varphi$ 进行积分, 函数的复合保证了积分区域转换的合法性, 但是我们并不可以草率进行 $\displaystyle\int_\Omega f\,\mathrm{d}x = \int_\Sigma f\circ\varphi\,\mathrm{d}x$ 的计算, 因为对于 Ω 上的某一块体积元 σ, 对与 f 与坐标变换 φ 下对应的 Σ 上的体积元 σ' 的体积并不一定相等, 就好像对于上面例子中, 将圆转化为矩形一样. 那么体积究竟变化了多少呢？这就是下面重积分换元法将要讨论的事情了.

下面, 我们首先考虑坐标变换为线性映射的情况, 再使用微分学的基本手法对一般的坐标变换做线性化并且估计误差.

(1) 平移变换. 设 $\boldsymbol{v_0}$ 为一个固定的向量,

(2) 伸缩变换.

引理 1.1

第一覆盖引理 （这里需要一个图图）

定理 1.16

(3) 正交变换

引理 1.2

第二覆盖引理

定理 1.17

(4) 一般的线性变换

定理 1.18

引理 1.3

引理 1.4

引理 1.5

定理 1.19

重积分换元法

内容总结

习题

A 组

1.

B 组

1.

C 组

1.

2

实数系的构造和数列极限

2.1 Review: 基本定义和定理

2.1.1 数列极限

定义 2.1 数列极限的定义

设 $\{x_n\}$ 是一个给定数列,a 是一个实常数, 如果对于任意给定的 $\epsilon > 0$, 可以找到正整数 N, 使得当 $n > N$ 时,

$$|x_n - a| < \epsilon$$

成立, 则称数列 $\{x_n\}$ 收敛于 a,（或者 a 是数列 $\{x_n\}$ 的极限）, 记为

$$\lim_{n\to\infty} x_n = a$$

如果我们知晓邻域的概念, 那么我们就可以引入下面的阐述, 这样就可以给数列极限一个可视化的理解：

定义 2.2 数列极限的几何阐述

设 $\{x_n\}$ 是一个给定的数列,a 是一个实常数, 如果对于任意给定的 $\epsilon > 0$, 可以找到正整数 N, 使得当 $n > N$ 时,$x_n \in O(a, \epsilon)$, 则

$$\lim_{n\to\infty} x_n = a$$

另外, 读者可以给出数列发散（即不收敛）的严谨定义（即练习 1.1）

进一步的, 我们可以对具有特殊极限或者有广义极限（参见如下说明）的数列给予命

名, 这样我们就有了无穷小量和无穷大量的概念, 这两个概念很好地帮助了我们理解了数列的极限.

定义 2.3 无穷大量和无穷小量

1. 如果数列 $\{x_n\}$ 收敛于 0, 则我们称数列 $\{x_n\}$ 是无穷小量;
2. 如果对于任意给定的 $G>0$, 可以找到正整数 N, 使得当 $n>N$ 时,
$$|x_n|>G$$
成立, 则称数列 x_n 是无穷大量, 记为
$$\lim_{n\to\infty} x_n=\infty.$$
如果无穷大量 $\{x_n\}$ 最终恒正（或者恒负）, 则称其为正无穷大量（或者负无穷大量）.
并且, 如果数列 $\{x_n\}$ 是无穷大量, 我们则认为数列 $\{x_n\}$ 有着广义极限.

另外, 我们经常能在提题目中看见类似下面的表述：$\lim\limits_{n\to\infty} x_n=a\ (-\infty<a<+\infty)$, 这表明数列 $\{x_n\}$ 收敛于某个常数, 而并非无穷大量.

利用无穷小量这一概念, 我们可以给出数列极限的两个十分好用的等价定义, 但是需要先讨论一下无穷大量和无穷小量的性质.

性质 2.1 无穷大量和无穷小量的性质

1.

推论 2.1 数列极限的等价定义

1. 如果数列 $\{x_n-a\}$ 是无穷小量, 则数列 $\{x_n\}$ 收敛于 a;
2. 设 $\{x_n\}$ 是一个给定数列,a 和 $K\ (K>0)$ 是两个实常数, 如果对于任意给定的 $\epsilon>0$, 可以找到正整数 N, 使得当 $n>N$ 时,
$$|x_n-a|<K\epsilon$$
成立, 则
$$\lim_{n\to\infty} x_n=a$$

我们需要格外注意并且理解第二条推论, 这允许我们更容易证明数列收敛于某个极限. 下面开始复习数列极限的一些性质：

性质 2.2 数列极限的性质

1.（极限的唯一性）收敛数列的极限必唯一;
2.（数列的有界性）收敛数列必有界;
3.（数列的保序性）
4.（夹逼定理）

性质 2.3 数列极限的四则运算

定理 2.1 Stolz 定理

设 $\{y_n\}$ 是严格单调递增的正无穷大量, 且

$$\lim_{n\to\infty}\frac{x_n - x_{n-1}}{y_n - y_{n-1}} = a\ (a \in \mathbb{R}\cup\{-\infty, +\infty\})$$

则有:

$$\lim_{n\to\infty}\frac{x_n}{y_n} = a$$

Stolz 定理的证明是十分经典的先处理极限是 **0** 的情况, 再利用这种容易证明并且容易推广的情况帮助完成后续证明的例子.

2.1.2 实数系完备性定理

定义 2.4 有界性

我们称一个集合 $\boldsymbol{A}$ 有上界, 当且仅当 $\exists\ M \in \mathbb{R}, \forall x \in \boldsymbol{A}, x < M$

类似地, 我们可以定义集合有下界. 当一个集合既有上界又有下界, 我们称这个集合有界.

我们记 $\boldsymbol{U}$ 是 $\boldsymbol{A}$ 的全体上界所组成的集合, 则显然 $\boldsymbol{U}$ 没有最大数, 但是当 $\boldsymbol{U}$ 有最小数 β 的时候, 我们称 β 是 $\boldsymbol{A}$ 的上确界, 记作

$$\beta = \sup\ \boldsymbol{A}$$

按这种方式, 我们也可以定义下确界.

容易看出, 确界有两个性质, 我们这里以上确界举例：

性质 2.4 上确界的性质

对于数集 S 的上界 β:

1. β 是数集 S 的上界：$\forall x \in S$, 有 $x \leqslant \beta$.
2. 任何小于 β 的数不是数集 S 的上界：$\forall \epsilon > 0, \exists\, x \in S$, 使得 $s > \beta - \epsilon$.

定理 2.2 实数系连续性定理——确界存在定理

非空有上界的实数集必定有上确界, 非空有下界的实数集必定有下确界

教材上本定理的证明是由无穷小数法给出的, 比较符合直观、比较容易理解. 倘若以公理法定义实数集, 本条定理会作为实数系的连续性公理存在. 而后续我们会给出由 Dedekind 分割以及由 Contor 基本列给出的两种不同的证明, 这两种体系均能够很好地建立起完备且连续实数系.

定理 2.3 单调有界数列收敛定理

单调有界数列必定收敛, 更确切地说, 数列一定收敛于他的确界.

这个证明通过确界存在定理以及确界的性质可以很容易给出, 这里留作练习.

定理 2.4 闭区间套定理

如果一组闭区间 $\{[a_n, b_n]\}$ 满足条件:

1. $[a_{n+1}, b_{n+1}] \subset [a_n, b_n],\ n = 1, 2, 3 \cdots;$
2. $\lim\limits_{n\to\infty} (n_n - a_n) = 0;$

则称这列闭区间形成一个闭区间套.

对于闭区间套 $\{[a_n, b_n]\}$, 存在唯一的实数 ξ 属于所有的闭区间 $[a_n, b_n]$, 且有 $\xi = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$

利用闭区间套定理可以得到一个似乎和闭区间套定理完全不搭边的一个定理, 但是这个定理是刻画实数系性质的一个很重要的定理:

定理 2.5

实数集是不可列集.

证明

用反证法：假设实数集 $\mathbb{R}$ 是可数集, 则可以找到一种排列规律使得

$$\mathbb{R}=\{x_1,x_2,\cdots,x_n,\cdots\},$$

则任取一个闭区间 $[a_1,b_1]\subset\mathbb{R}$ 且 $x_1\notin[a_1,b_1]$, 然后将此区间三等分：则在等分后的区间：

$$\left[a_1,\frac{2a_1+b_1}{3}\right],\left[\frac{2a_1+b_1}{3},\frac{a_1+2b_1}{3}\right],\left[\frac{a_1+2b_1}{3},\frac{b_1}{3}\right]$$

中, 必存在一个区间, 这个区间不包含 x_2, 则将这个区间记作 $[a_2,b_2]$, 则继续之前的操作 (这是可以一直做下去的), 则我们可以得到一个闭区间套 $\{[a_n,b_n]\}$, 满足：

$$x_n\notin[a_n,b_n]$$

则根据闭区间套定理, 存在一个实数 ξ 属于所有的闭区间, 换言之：

$$\xi\neq x_n\ (n=1,2,3,\cdots)$$

而这恰恰与集合 $\{x_1,x_2,\cdots,x_n\cdots\}$ 表示实数集 $\mathbb{R}$ 矛盾! □

这个定理的证明很经典, 各位可以复习一下.

下面这个定理在证明一个数列的发散与否很有用, 这一点可以类似于判断方向导数的存在性.

定理 2.6

如果数列 $\{x_n\}$ 收敛于 a, 则它的任何子列 $\{x_{n_k}\}$ 也收敛于 a.

定理 2.7 Bolzano-Weierstrass 定理

有界数列必有收敛子列.

在数列无界的时候, 我们也有类似的定理

定理 2.8 Bolzano-Weierstrass 定理 (无界版)

若数列 $\{x_n\}$ 是一个无界数列, 则必存在一个子列 $\{x_{n_k}\}$ 使得 $\lim\limits_{k\to\infty} x_{n_k} = \infty$

Bolzano-Weierstrass 定理的另外一种表述如下：

定理 2.9 Bolzano-Weierstrass 定理

实数轴上的任何有界无限点集 S 至少存在一个聚点.

定义 2.5 基本列

如果数列 $\{x_n\}$ 满足：对于任意给定的 $\epsilon > 0$, 总存在一个自然数 N, 使得当 $m, n > N$ 时, 有：

$$|x_m - x_n| < \epsilon$$

则称数列 $\{x_n\}$ 是一个基本列.

定理 2.10 Cauchy 收敛定理

数列 $\{x_n\}$ 收敛的充分必要条件是 $\{x_n\}$ 是基本列.

Cauchy 收敛定理表明了很重要的一件事: 由实数组成的基本列 $\{x_n\}$ 必存在实数极限, 这个性质被称为实数的完备性. 我们在引入了点集拓扑等更深刻的理论之后会介绍另外一条实数系完备性定理——有限覆盖定理, 这是实数系完备性定理的最后一块拼图, 更好地刻画了实数系的性质.

定理 2.11 压缩映射定理

2.1.3 例题与练习

例 2.1

证明：任何数列都有单调子列.

证明

首先，对于无上界或者无下界的数列，根据无界的定义，我们能很容易地证明这个数列存在单调子列 (更确切地说，无上界的数列必有递增子列，无下界的数列必有递减子列).

我们考虑有界的数列，分以下两种情况：

1. 如果 $\forall k \in \mathbb{N}$, 数列 $\{a_{k+n}\}$ 存在最大数，则我们按下列方式构造单调递减数列 $\{a_{m_n}\}$：对于 $k=1$, 取数列 $\{a_{1+n}\}$ 的最大数 a_{m_1}, 接着考虑数列 $\{a_{m_1+n}\}$, 其必有最大数 a_{m_2} 且 $a_{m_2} < a_{m_1}$, 继续考虑数列 $\{a_{m_2+n}\}\cdots$ 这是可以一直做下去的，这样，我们就得到了一个单调递减的子列 $\{a_{m_n}\}$.
2. 如果至少存在一个 k, 令数列 $\{a_{k+n}\}$ 不存在最大值，则我们按下列方式构造单调递增数列 $\{a_{k+m_n}\}$：对于 $m_1 = 1$, 由于此数列不存在最大值，所以可以取出 $m_2 > m_1$, 令 $a_{k+m_1} < a_{k+m_2}\cdots$ 这是可以一直做下去的，这样，我们就得到了一个单调递增的子列 $\{a_{k+m_n}\}$.

$\square$

通过这个例题和单调有界数列收敛定理，我们可以直接得到 Bolzano-Weierstrass 定理.

例 2.2

如果数列 $\{x_n\}$ 满足 $\lim\limits_{n\to\infty} x_n = a\ (-\infty < a < +\infty)$, 则证明：

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{x_k}{n} = a.$$

例 2.3

设 $0<\lambda<1$, $\lim\limits_{n\to\infty} a_n = a$, 证明：

$$\lim_{n\to\infty}\left(a_n + \lambda a_{n-1} + \lambda^2 a_{n-2} + \cdots + \lambda^n a_n\right) = \frac{a}{1-\lambda}$$

习题

We can never 'reach' infinity, we must develop methods which allow us to prove statements about infinitely many function values 'near infinity'.

——Herbert Amann

1. 计算:

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n}\sqrt{1+\frac{k}{n^2}}-n\right)$$

2.2 数列的上下极限

我们首先考虑下面两个数列：

$$a_n=\inf\{x_k|k\geqslant n,k\in\mathbb{Z}\}\quad b_n=\sup\{x_k|k\geqslant n,k\in\mathbb{Z}\},$$

由于 a_n 不减,b_n 不增, 亦即

$$a_1\leqslant a_2\leqslant\cdots\leqslant a_n\leqslant\cdots\leqslant b_n\leqslant\cdots\leqslant b_2\leqslant b_1,$$

由于单调有界数列收敛, 所以 $\{a_n\}$ 和 $\{b_n\}$ 的极限存在, 分别称之为数列 $\{x_n\}$ 的下极限和上极限, 记之为

$$\varliminf_{n\to\infty}x_n=\lim_{n\to\infty}a_n\qquad\varlimsup_{n\to\infty}x_n=\lim_{n\to\infty}b_n.$$

某些时候, 我们也会见到 $\liminf\limits_{n\to\infty}x_n$ 和 $\limsup\limits_{n\to\infty}x_n$ 的写法, 我们只需要清楚这仅仅是记号的不同而已.

2.3 点集拓扑初步

定理 2.12 Heine-Borel 有限覆盖定理

3 函数极限与连续性

3.1 Review: 基本定义和定理

3.1.1 函数极限

定义 3.1 函数极限：$\epsilon - \delta$ 语言

设函数 $y = f(x)$ 在点 x_0 的某个去心邻域中有定义, 即存在 $\rho > 0$, 使

$$O(x_0, \rho) \backslash \{x_0\} \subset D_f$$

如果存在实数 A 使得对于任意给定的 $\epsilon > 0$, 可以找到 $\delta > 0$, 使得当 $0 < |x - x_0| < \delta$ 时, 成立

$$|f(x) - f(x_0)| < \epsilon$$

则称 A 是函数 $f(x)$ 在 x_0 点的极限, 记作 $\lim\limits_{x \to x_0} f(x) = A$.
如果不存在满足上述性质的常数 A, 则函数在 x_0 点的极限不存在.

数列极限和函数极限同为极限, 下面的 Heine 定理建立起了函数极限和数列极限之间的关系, 它能将函数值数列的极限**归结到**函数的极限上, 所以我们一般将 Heine 定理称为 Heien 归结原理.

定理 3.1 Heine 归结原理

$\lim\limits_{x \to x_0} f(x) = A$ 的充分必要条件是：对于任意满足条件 $\lim\limits_{n \to \infty} x_n = x_0, x_n \neq x_0$ $(n =$

$1, 2, 3, \cdots$) 的数列 $\{x_n\}$, 相应的函数值数列 $\{f(x_n)\}$ 成立：

$$\lim_{x\to\infty} f(x_n) = A.$$

事实上, 当我们只需要判断函数在某一点 (如 x_0) 的敛散性时, 我们可以使用比较" 宽松" 一点的 Heine 定理：

定理 3.2

函数极限 $\lim\limits_{x\to x_0} f(x)$ 存在的充分必要条件是：对于任意满足条件 $\lim\limits_{n\to\infty} x_n = x_0$ 且 $x_n \neq x_0 (n = 1, 2, 3, \cdots)$ 的数列 $\{x_n\}$, 相应的函数值数列 $\{f(x_n)\}$ 收敛.

和数列极限相似, 通过 Heine 定理, 我们可以证明函数极限下的 Cauchy 收敛原理：

定理 3.3 Cauchy 收敛原理

函数极限 $\lim\limits_{x\to x_0} f(x)$ 存在且有限的充分必要条件是：对于任意给定的 $\epsilon > 0$, 存在 $\delta > 0$, 使得对于任意满足条件 $|x_1 - x_0| < \delta, |x_2 - x_0| < \delta$ 的 x_1, x_2, 都有：

$$|f(x_1) - f(x_2)| < \epsilon$$

性质 3.1 函数极限的性质

1. (极限的唯一性) 设 A 和 B 都是函数 $f(x)$ 在 x_0 处的极限, 则有 $A = B$.

2. (局部保序性) 若 $\lim\limits_{x\to x_0} f(x) = A, \lim\limits_{x\to x_0} g(x) = B$, 且 $A < B$, 则存在 $\delta > 0$, 当 $0 < |x - x_0| < \delta$ 时, 成立
$$f(x) > g(x).$$

3. (局部有界性) 若 $\lim\limits_{x\to x_0} f(x) = A$, 则存在 $\delta > 0$, 使得 $f(x)$ 在 $O(x_0, \delta)\backslash\{x_0\}$ 中有界.

4. (夹逼定理) 若存在 $r > 0$, 使得当 $0 < |x - x_0| < r$ 时, 成立:
$$g(x) \leqslant f(x) \leqslant h(x),$$
且 $\lim\limits_{x\to x_0} f(x) = \lim\limits_{x\to x_0} h(x) = A$, 则 $\lim\limits_{x\to x_0} f(x) = A$.

5. (四则运算) 设 $\lim\limits_{x\to x_0} f(x) = A, \lim\limits_{x\to x_0} g(x) = B$, 则:

   (1) $\lim\limits_{x\to x_0} (\alpha f(x) + \beta g(x)) = \alpha A + \beta B$;

(2) $\lim\limits_{x\to x_0}(f(x)g(x)) = AB$;

(3) $\lim\limits_{x\to x_0}\frac{f(x)}{g(x)} = \frac{A}{B}\ (B\neq 0)$.

定义 3.2 单侧极限

设函数 $f(x)$ 在 $(x_0-\rho, x_0)$ 有定义 $(\rho>0)$. 如果存在实数 B, 对于任意给定的 $\epsilon>0$, 可以找到 $\delta>0$, 使得当 $-\delta<x-x_0<0$ 时, 成立

$$|f(x)-B|<\epsilon,$$

则称 B 是函数 $f(x)$ 在点 x_0 处的左极限, 记为

$$\lim_{x\to x_0^-} f(x) = f(x_0^-) = B.$$

类似地, 如果函数 $f(x)$ 在 $(x_0, x_0+\rho)$ 有定义 $(\rho>0)$. 如果存在实数 C, 对于任意给定的 $\epsilon>0$, 可以找到 $\delta>0$, 使得当 $0<x-x_0<\delta$ 时, 成立

$$|f(x)-C|<\epsilon,$$

则称 C 是函数 $f(x)$ 在点 x_0 处的右极限, 记为

$$\lim_{x\to x_0^+} f(x) = f(x_0^+) = C.$$

事实上, 自变量的极限过程可以分为六种情况：$x\to x_0, x\to x_0^-, x\to x_0^+, x\to\infty, x\to -\infty, x\to+\infty$. 函数值的极限有四种情况：$f(x)\to A, f(x)\to\infty, f(x)\to-\infty, f(x)\to+\infty$. 这样, 我们就可以将函数极限的定义扩充到很宽的情况.

命题 3.1 单调函数单侧极限存在定理

设 $f(x)$ 在区间 (a,b) 上单调, 则 $\lim\limits_{x\to b^-} f(x) = f(b^-)$ 一定有意义, 即若函数 $f(x)$ 单调增加时, 如 $f(x)$ 在 (a,b) 有上界, 则 $f(b^-) = \sup\{y|\ \exists\, x\in(a,b), y=f(x)\}$, 否则 $f(b^-)=+\infty$, 反之亦然.

3.1.2 函数的连续性

定义 3.3 连续函数

设函数 $f(x)$ 在点 x_0 的某个去心邻域中有定义, 并且成立

$$\lim_{x\to x_0} = f(x_0),$$

则称函数 $f(x)$ 在点 $\boldsymbol{x_0}$ 连续, 而称 x_0 是函数 $f(x)$ 的连续点.
如果函数 $f(x)$ 在区间 (a,b) 的每一点都连续, 则称函数 $f(x)$ 在开区间 (a,b) 上连续.

定义 3.4 振幅

设 $f(x)$ 在 x_0 的一个开邻域内有定义, 称

$$\omega_f(x_0,r) = \sup\{|f(x')-f(x'')||x',x''\in(x_0-r,x_0+r)\}\ (r>0)$$

为 f 在区间 (x_0-r,x_0+r) 上的振幅. 显然,$\omega_f(x_0,r)$ 关于 $r\to 0^+$ 单调递减, 因此

$$\omega_f(x_0) = \lim_{r\to 0^+}\omega_f(x_0,r)$$

存在 (不一定有限), 称为 f 在 x_0 处的振幅.
为了用振幅来刻画一致连续性, 设 f 定义在区间 I 中,$r>0$. 令

$$\omega_f(r) = \sup\{|f(x')-f(x'')||x',x''\in I,|x'-x''|<r\},$$

则 $\omega_f(r)$ 关于 $r\to 0^+$ 单调递减.

定理 3.4

函数 $f(x)$ 在 x_0 处连续当且仅当 $\omega_f(x_0)=0$.
函数 $f(x)$ 在 I 中一致连续当且仅当 $\lim\limits_{r\to 0^+}\omega_f(r)=0$.

这种用振幅来刻画函数的连续性的想法也出现在刻画函数的可积性上.

定义 3.5 三类不连续点

1. 第一类不连续点：函数 $f(x)$ 在 x_0 左、右极限都存在但是不相等, 即 $f(x^+)\neq f(x^-)$. 第一类不连续点又称为跳跃点, 右极限和左极限之差 $f(x^+)-f(x^-)$ 称为函数 $f(x)$ 在 x_0 处的跃度;

2. 第二类不连续点：函数 $f(x)$ 在 x_0 的左、右极限中至少一个不存在;

3. 第三类不连续点：函数 $f(x)$ 在 x_0 的左、右极限都存在且相等, 但是不等于 $f(x_0)$

或者函数 $f(x)$ 在 x_0 处无定义. 第三类不连续点又称为可去不连续点.

性质 3.2

在区间 (a,b) 上的单调函数的不连续点是第一类不连续点, 即左右极限都存在但不相等.

证明

设 $x_0 \in (a,b)$ 是任意一点, 集合 $\{f(x)|x \in (a,x_0)\}$ 非空且有上界, 则一定存在上确界 α,

$$\alpha = \sup\{f(x)|x \in (a,x_0)\}.$$

对一切 $x \in (a,x_0), f(x) \leqslant \alpha$, 且有 $\forall \epsilon > 0, \exists\ x^{'} \in (a,x_0)$, 使得 $\alpha - f(x^{'}) < \epsilon$, 取 $\delta = x_0 - x^{'}$, 则当 $-\delta < x - x_0 < 0$ 时,$-\epsilon < f(x^{'}) - f(x_0) \leqslant f(x) - f(x_0) \leqslant 0$, 这样有 $\lim\limits_{x \to x_{0-}} f(x) = \alpha$, 同理 $\lim\limits_{x \to x_{0+}} f(x) = \beta$, 其中 $\beta = \inf\{f(x)|x \in (x_0,b)\}$.

□

定理 3.5 连续函数的四则运算

定理 3.6 反函数的存在性

若函数 $y = f(x), x \in D_f$ 是单调的, 则它的反函数 $x = f^{-1}(y), y \in R_f$ 存在, 而且 $f^{-1}(y)$ 的单调性和 $f(x)$ 相同.

定理 3.7 反函数的连续性

设函数 $f(x)$ 在闭区间 $[a,b]$ 连续且严格单调增加,$f(a) = \alpha, f(b) = \beta$, 则它的反函数 $f^{-1}(y)$ 在 $[\alpha,\beta]$ 连续且严格单调增加.

定理 3.8 复合函数的连续性

若 $u = g(x)$ 在点 x_0 连续,$g(x_0) = u_0$, 又 $y = f(u)$ 在点 u_0 连续, 则复合函数 $y = f \circ g(x)$ 在点 x_0 连续.

定理 3.9 初等函数的连续性

我们认为初等函数是指的是六类基本初等函数经过有限次四则运算和复合运算所产生的函数.
初等函数在其定义区间上连续.

我们利用实数系的完备性定理可以证明以下连续函数的性质,

定理 3.10 闭区间上连续函数的性质

1. (有界性定理)
2. (最值定理)
3. (介值定理)
4. (零点存在定理)

推论 3.1

1. 设 $f(x)$ 是 $I=[a,b]$ 上的连续函数, 则 $f(I)=[m,M]$ 其中 m,M 分别是 f 在 $[a,b]$ 上的最小值和最大值.
2. 设 $f(x)$ 是区间 I 上的连续函数, 则 $f(I)$ 也是一个区间 (可以退化成一个点).
3. 设 $f(x)$ 是区间 I 上的连续函数, 则 $f(x)$ 可逆当且仅当 $f(x)$ 是严格单调函数.

定义 3.6 一致连续

定理 3.11 一致连续性的判断

设函数 $f(x)$ 在区间 X 上有定义, 则 $f(x)$ 在 X 上一致连续的充分必要条件是: 对任何点列 $\{x_n^{'}\}$ $(x_n^{'}\in X)$ 和 $\{x_n^{''}\}$ $(x_n^{''}\in X)$, 只要满足 $\lim\limits_{n\to\infty}(x_n^{''}-x_n^{''})=0$, 就成立 $\lim\limits_{n\to\infty}(f(x_n^{'})-f(x_n^{''}))=0$.

证明

首先证明必要性:

由于函数 $f(x)$ 的一致连续性,

$$\forall \epsilon > 0, \exists \delta > 0, \forall x^{'}, x^{''} \in X, (|x^{'} - x^{''}| < \delta) : |f(x^{'}) - f(x^{''})| < \epsilon.$$

□

定理 3.12 Contor 定理

闭区间上的连续函数在此区间一致连续.
换句话说：若函数 $f(x)$ 在闭区间 $[a, b]$ 上连续, 则它在 $[a, b]$ 上一致连续.

定理 3.13

函数 $f(x)$ 在有限开区间 (a, b) 上连续, 则 $f(x)$ 在 (a, b) 上一致连续的充分必要条件是：$f(a+)$ 与 $f(b-)$ 存在.

3.1.3 无穷大量和无穷小量的阶

定义 3.7 无穷小量

定义 3.8 无穷大量

定义 3.9 等价量

定理 3.14 替换定理

设 $f(x), g(x), f_1(x)$ 在 x_0 的某个去心邻域 U 中有定义, 且 $f(x) \sim f_1(x)$ $(x \to x_0)$, 那么：

1. 当 $\lim\limits_{x \to x_0} f_1(x) g(x) = A$ 时, 有 $\lim\limits_{x \to x_0} f(x) g(x) = A$;
2. 当 $\lim\limits_{x \to x_0} \frac{g(x)}{f_1(x)} = A$ 时, 有 $\lim\limits_{x \to x_0} \frac{g(x)}{f(x)} = A$.

例 3.1 常见的等价量

1. $\ln(1+x) \sim x \ (x \to 0)$
2. $\sin x \sim x \ (x \to 0)$
3. $\cos x \sim 1 - \frac{x^2}{2} \ (x \to 0)$
4. $e^x \sim 1 + x \ (x \to 0)$

3.1.4 例题与练习

例 3.2 函数极限的换元法

设 $\lim\limits_{x\to a} g(x) = A$, $\lim\limits_{y\to A} f(y) = B$ 成立, 且在点 a 的某个邻域上 $g(x) = y$. 如果满足以下条件之一:

1. 存在点 a 的一个去心邻域 $O_{\delta_0}(a) - a$, 在其中 $g(x) \neq A$;
2. $\lim\limits_{y\to A} f(y) = f(A)$;
3. $A = \infty$, 且 $\lim\limits_{x\to A} f(y)$ 有意义;

则成立

$$\lim_{x\to a} f(g(x)) = \lim_{y\to A} f(y) = B.$$

例 3.3

设 $\lim\limits_{x\to a} g(x) = A$, $\lim\limits_{y\to A} f(y) = B$, 证明：极限 $\lim\limits_{x\to a} f(g(x))$ 只有三种可能:
(1): $\lim\limits_{x\to a} f(g(x)) = B$; (2): $\lim\limits_{x\to a} f(g(x)) = f(A)$; (3): 极限 $\lim\limits_{x\to a} f(g(x))$ 不存在.

例 3.4

设 $f(x)$ 是定义在区间 I 上的函数, 如果存在 $0 < \alpha \leqslant 1$, 以及常数 M, 使得

$$|f(x_1) - f(x_2)| \leqslant M|x_1 - x_2|^\alpha, \ \forall x_1, x_2 \in I$$

则称 $f(x)$ 是 I 中的 α 阶 Holder 函数, 当 $\alpha = 1$ 时也称为 Lipschitz 函数.
证明：Holder 函数都是一致连续的.

例 3.5 Brouwer 不动点定理

函数 $f : [0,1] \to [0,1]$ 是连续映射, 那么 f 有不动点, 即存在 $x \in [0,1]$, 使得 $f(x) = x$.

例 3.6

证明：不存在函数 $f : \mathbb{R} \to \mathbb{R}$ 在所有无理点不连续, 而在所有有理点连续.

3.2 e 和 e^x 的构造

命题 3.2

证明: 数列 $\{(1+\frac{1}{n})^n\}$ 单调递增, 数列 $\{(1+\frac{1}{n})^{n+1}\}$ 单调递减.

定理 3.15 e 的构造

数列 $\{(1+\frac{1}{n})^n\}$ 和数列 $\{(1+\frac{1}{n})^{n+1}\}$ 收敛于相同的极限, 记此极限为 e.

例 3.7

证明

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} = e$$

例 3.8

记 $b_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n$, 证明：数列 $\{b_n\}$ 收敛.

例 3.9

证明：

$$\lim_{n\to\infty}\left[1-\frac{1}{2}+\frac{1}{3}-\cdots+(-1)^{n+1}\frac{1}{n}\right]=\ln 2.$$

证明

$$\left(\frac{p}{q}\right)=\begin{cases}1 & if\ n^2\equiv q(modp)for\ integer\ r\\ -1 & otherwise\end{cases}$$

□

4 一元函数微分学

4.1 Review：Back to Newton

4.1.1 导数与微分

4.1.2 微分的逆运算：不定积分

5

一元函数积分学

5.1 Review: 从 Riemann 到 Lebesgue

5.1.1 函数的可积性

5.1.2 Riemann 积分的一般性质与应用

利用黎曼和可以解决某些本难以解决的级数问题, 这里浅举几例. 解决这类问题的关键在于凑出黎曼和的形式, 并且我们一般都取等距分割.

例 5.1

计算或证明下面的极限:

1.
$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{\sqrt{n^2+n+k^2}}.$$

2.
$$\lim_{n\to\infty}\sum_{k=1}^{n}\left[f(x+\frac{k}{n^2+k^2})-f(x)\right]=f'(x)\frac{\ln 2}{2}.$$

证明

□

例 5.2

非负函数 $f \in C[a,b]$, 证明：

$$\lim_{n\to\infty}\left(\int_a^b f^n(x)dx\right)^{\frac{1}{n}} = \max\{f(x) \mid x \in [a,b]\}.$$

5.1.3 变限积分

变限积分的主要结果是下面两个命题.

定理 5.1

1. 设 $f \in R[a,b]$, 则 $F(x) = \int_a^x f(t)dt$ 与 $G(x) = \int_x^a f(t)dt$ 在 $[a,b]$ 上连续.
2. 设 $f \in R[a,b]$,$x \in [a,b]$ 是 f 的连续点, 则

$$\frac{d}{dx}\int_a^x f(t)dt = f(x).$$

由上面的定理可以得出原函数存在的一个充分条件. 并且, 变限积分的一个副产品是微积分基本定理, 亦即 Newton-Leibniz 公式.

定理 5.2 原函数存在定理

设 $f \in C[a,b]$, 则 f 在 $[a,b]$ 上存在原函数.

例 5.3

设 $f \in R[A,B]$,$a,b \in [A,B]$ 是 f 的两个连续点, 证明：

$$\lim_{h\to 0}\int_a^b \frac{f(x+h)-f(x)}{h}dx = f(b)-f(a).$$

证明

要点是通过换元改变积分限, 然后将极限看作求导.

$$
\begin{aligned}
\lim_{h\to 0}\int_a^b \frac{f(x+h)-f(x)}{h}dx &= \lim_{h\to 0}\frac{1}{h}\left(\int_a^b f(x+h)dx-\int_a^b f(x)dx\right)\\
&= \lim_{h\to 0}\frac{1}{h}\left(\int_{a+h}^{b+h} f(x)dx-\int_a^b f(x)dx\right)\\
&= \lim_{h\to 0}\frac{1}{h}\left(\int_{b}^{b+h} f(x)dx-\int_a^{a+h} f(x)dx\right)\\
&= f(b)-f(a)
\end{aligned}
$$

□

Tip: 这个题目不能使用下面的证法, 下面的证明每一步都是错误的.

$$
\begin{aligned}
\lim_{h\to 0}\int_a^b \frac{f(x+h)-f(x)}{h}dx &= \int_a^b \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}dx\\
&= \int_a^b f'(x)dx = f(b)-f(a)
\end{aligned}
$$

首先, 没有依据就把求极限和积分号交换是不对的, 我们需要在级数那章讨论求极限和积分号交换的条件; 其次, 对差商求导的时候忘记了题目只给了 f 在两个点处的连续条件, 甚至没有给可导的条件; 最后, 即使 $f(x)$ 在区间 $[a,b]$ 可导,$f'(x)$ 也不一定可积.

5.2 积分的逼近性质

在本节, 我们主要说明积分的逼近性质, 即, 如果一个函数是 Riemann 可积的, 那么它可以被一列简单函数 (通常是阶梯函数或者分段线性函数) 逼近. 这一性质在证明积分的一些性质时是非常有用的.

定理 5.3 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

$$\phi_n\leqslant f\leqslant \psi_n, \int_a^b\left[\psi_n(x)-\phi_n(x)\right]dx<\frac{1}{n},$$

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

$$\lim_{n\to\infty}\int_a^b \phi_n(x)g(x)dx=\lim_{n\to\infty}\int_a^b \psi_n(x)g(x)dx=\lim_{n\to\infty}\int_a^b f(x)g(x)dx.$$

证明

□

定理 5.4 分段线性逼近

设 $f \in R[a,b]$, 则存在一列连续的分段线性函数 f_n, 使得 $f_n(a)=f(a), f_n(b)=f(b)$,

$$\lim_{n\to\infty}\int_a^b |f_n(x)-f(x)|\,dx=0,$$

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

$$\lim_{n\to\infty}\int_a^b f_n(x)g(x)dx=\lim_{n\to\infty}\int_a^b f(x)g(x)dx.$$

证明

□

下面是一个利用逼近解决的问题, 我们给出阶梯逼近和分段线性逼近两种做法.

例 5.4 Riemann-Lebegsue 引理

设 $f \in R[a,b]$, 则

$$\lim_{\lambda\to+\infty}\int_a^b f(x)\sin\lambda x dx=\lim_{\lambda\to+\infty}\int_a^b f(x)\cos\lambda x dx=0.$$

证明

□

我们甚至还有更强的定理, 亦即 Weierrstrass 逼近定理, 它表明了连续函数可以被多项式逼近.

定理 5.5 Weierrstrass 逼近定理

设 $f \in C^0[a,b]$, 则任给 $\epsilon>0$, 存在多项式 $P(x)$, 使得 $|P(x)-f(x)|<\epsilon$ 在 $[a,b]$ 中处处成立.

证明

□

定义 5.1 Bernstein 多项式

例 5.5

设 f 在 $[0,1]$ 中满足条件 $|f(x)-f(y)|<L|x-y|$, 则

$$|B_n(x,f)-f(x)| \leqslant \frac{L}{2\sqrt{n}}, \forall x \in [0,1].$$

例 5.6 上一个例子的加强

定理 5.6 Riemann 引理

设 f 是周期函数, 周期为 T, 如果 f 在闭区间 $[0,T]$ 中可积, 则有以下陈述:

1. f 在任何闭区间中均可积, 并且在任何长度为 T 的区间中的积分都相等.
2. 思考此式的含义并证明
$$\lim_{\lambda\to\infty} \frac{1}{\lambda}\int_0^\lambda f(x)dx = \frac{1}{T}\int_0^T f(x)dx.$$
3. $g \in \boldsymbol{R}[a,b]$, 证明
$$\lim_{\lambda\to\infty}\int_a^b f(\lambda x)g(x)dx = \frac{1}{T}\int_0^T f(x)dx\int_a^b g(x)dx.$$

5.3 积分中值定理

定理 5.7 积分第一中值定理

设 $f,g\in R[a,b]$, 且 g 不变号, 则存在 $\mu\in\left[\inf_{[a,b]}f(x),\sup_{[a,b]}f(x)\right]$, 使得

$$\int_a^b f(x)g(x)dx=\mu\int_a^b g(x)dx.$$

特别地, 若 f 连续, 则存在 $\xi\in[a,b]$, 使得 $\mu=f(\xi)$.

定理 5.8 积分第二中值定理

设 $f\in R[a,b]$,g 是 $[a,b]$ 中的单调函数, 则存在 $\xi\in[a,b]$, 使得

$$\int_a^b f(x)g(x)dx=g(a)\int_a^\xi f(x)dx+g(b)\int_\xi^b f(x)dx.$$

更进一步, 如果 g 是非负函数：

1. 如果 g 在 $[a,b]$ 单调递减, 则存在 $\zeta\in[a,b]$, 使得

$$\int_a^b f(x)g(x)dx=g(a)\int_a^\zeta f(x)dx.$$

2. 如果 g 在 $[a,b]$ 单调递增, 则存在 $\eta\in[a,b]$, 使得

$$\int_a^b f(x)g(x)dx=g(b)\int_\eta^b f(x)dx.$$

5.4 广义积分

5.5 Stieltjes 积分

5.6 重要积分的计算

5.6.1 Euler 积分

本小节计算 Euler 积分：

$$\int_0^{\pi/2}\ln\sin x\mathrm{d}x.$$

5.6.2 Euler-Poisson 积分

本小节计算 Euler-Poisson 积分：

$$\int_0^{+\infty} e^{-x^2}\mathrm{d}x.$$

5.7 练习

5.7.1 基本概念与计算

1.

2.

3.

下面是一些对积分求极限的习题.

练习 5.1

1. 证明：
$$\lim_{n\to\infty}\int_0^{\pi/2}\sin^n x dx = 0.$$
2. $f \in C[-1,1]$, 证明：
$$\lim_{h\to 0^+}\int_{-1}^{1}\frac{h}{h^2+x^2}f(x)dx = \pi f(0).$$
3. $f \in C[-1,1]$, 证明：
$$\lim_{n\to\infty}\int_0^1 nx^n f(x)dx = f(1).$$

下面这些题需要灵活地使用微分中值定理、泰勒公式以及积分中值定理.

练习 5.2

1. 设 $f \in C^2[a,b]$,$M = \sup\limits_{x\in[a,b]}|f''(x)|$, 证明：
$$|\int_a^b f(x)dx - \frac{b-a}{2}(f(a)+f(b))| \leqslant \frac{M}{12}(b-a)^3.$$

2. 若 $a>0$, 且 $f\in C^1[0,a]$, 则证明：

$$|f(0)|\leqslant\frac{1}{a}\int_0^a|f(x)|dx+\int_0^a|f'(x)|dx.$$

3. 若 $f\in C^2[a,b]$, 证明：存在 $\xi\in(a,b)$, 使得：

$$f''(\xi)=\frac{24}{(b-a)^3}\int\limits_a^b\left[f(x)-f(\frac{a+b}{2})\right]dx.$$

4. 若 $f\in C^1[0,1]$, 且 $f(x)>0,x\in[0,1]$, 证明：$\forall n>1,\exists\xi_n\in(0,1)$, 使得：

$$\frac{1}{n}\int_0^1 f(x)dx=\int_0^{\xi_n}f(x)dx+\int_{\xi_n}^1 f(x)dx.$$

并且,

$$\lim_{n\to+\infty}n\xi_n=\frac{1}{f(0)+f(1)}\int_0^1 f(x)dx.$$

有关变上限积分, 我们有下面的习题, 我们尤其需要注意变上限积分的求导问题.

练习 5.3

1. $f\in C[-1,1]$,f 在 $x=0$ 可导,$f(0)=0,f'(0)\neq0$, 求极限：

$$I=\lim_{x\to0^+}\frac{\displaystyle\int_0^x(x^2-t^2)f(t)dt}{\displaystyle\int_0^x tf(x^2-t^2)dt}.$$

2. 设 $f\in C^1[0,a]$,$f(0)=0$, 证明：

$$\int_0^a|f(x)f'(x)|\mathrm{d}x\leqslant\frac{a}{2}\int_0^a(f'(x))^2\mathrm{d}x.$$

5.7.2 积分的逼近性质

1. 设 f 在 $[0,+\infty)$ 上的无穷积分绝对收敛, 证明：

$$\lim_{\lambda\to+\infty}\int_0^{+\infty}f(x)\sin\lambda x\mathrm{d}x=0$$

6 级数

6.1 数项级数

6.1.1 一般的级数

定义 6.1 级数

设 $a_1, a_2, \cdots, a_n, \cdots$ 是无穷多可列个实数, 我们称形式和

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots + a_n + \cdots$$

为无穷级数, 称 a_n 为级数的通项或一般项, $S_n = \sum\limits_{k=1}^{n} a_k$ 为级数的第 n 个部分和. 如果部分和 S_n 的极限存在且收敛于有限数 S, 则称级数 $\sum\limits_{k=1}^{\infty} a_k$ 收敛, 记作 $\sum\limits_{k=1}^{\infty} a_k = S$, 且称它的和为 S, 否则称级数 $\sum\limits_{k=1}^{\infty} a_k$ 发散.

需要注意的是, 级数的敛散性和其有限项的值无关. 从某种意义上而言, 级数的敛散性本质上就是数列的敛散性, 所以根据数列极限的性质还可以得到:

性质 6.1

1. 级数收敛的必要条件：如果级数 $\sum\limits_{n=1}^{\infty}$ 收敛, 那么其通项满足 $\lim\limits_{n\to\infty} a_n = 0$.

2. 级数收敛的充要条件（Cauchy 准则）：级数 $\sum\limits_{n=1}^{\infty}$ 收敛 $\iff$ 任给 $\epsilon > 0$, 存在

$N=N(\epsilon)$, 当 $m>n>N$ 时, 有 $|a_{n+1}+\cdots+a_m|=\left|\sum\limits_{k=n+1}^{m}a_k\right|<\epsilon$.

3.（线性性）设级数 $\sum\limits_{n=1}^{\infty}a_n$ 和 $\sum\limits_{n=1}^{\infty}b_n$ 都收敛,$\lambda,\mu\in\mathbb{R}$, 则级数 $\sum\limits_{n=1}^{\infty}(\lambda a_n+\mu b_n)$ 也收敛, 且

$$\sum_{n=1}^{\infty}(\lambda a_n+\mu b_n)=\lambda\sum_{n=1}^{\infty}a_n+\mu\sum_{n=1}^{\infty}b_n.$$

回忆上一章利用黎曼和解决的很多问题, 我们似乎可以把无限和与积分联系起来, 下面的定理就将级数求和与广义积分联系了起来:

引理 6.1

设 $\{a_n\}$ 是一列实数, 在 $[1,+\infty)$ 中定义函数 $a(x)$ 如下：设 $k\geqslant 1$, 当 $x\in[k,k+1)$ 时,$a(x)=a_k$, 则级数 $\sum\limits_{n=1}^{\infty}a_n$ 收敛当且仅当无穷积分 $\displaystyle\int_1^{+\infty}a(x)\mathrm{d}x$ 收敛, 且收敛的时候级数的和等于无穷积分的值.

利用上面的引理, 我们可以很轻松地把广义积分敛散性的判别法搬到级数上来, 对于这种方法, 我们不赘述证明, 下面内容会按照传统的方式给出证明.

对于一般的级数, 下面的 Dirichlet 判别法和 Abel 判别法是核心. 其证明的关键在于数列的 **Abel** 变换和 **Abel** 引理：

引理 6.2

1.（Abel 变换）设 $\{a_n\},\{b_n\}$ 是两列实数,$B_i=\sum\limits_{k=1}^{i}b_k$, 则

$$\sum_{i=1}^{n}a_ib_i=\sum_{i=1}^{n-1}(a_i-a_{i+1})B_i+a_nB_n.$$

2.（Abel 引理）设 $\{a_n\},\{b_n\}$ 是两列实数,$B_i=\sum\limits_{k=1}^{i}b_k$, 若 $\{a_n\}$ 为单调数列,$\{B_n\}$ 是有界数列,$M=\sup\limits_{1\leqslant k\leqslant n}|B_k|$, 则

$$\sum_{i=1}^{n}a_ib_i\leqslant M(|a_1|+2|a_n|)$$

乍看 Abel 变换, 你可能一时间摸不到头脑, 其实 Abel 变换只是对 $\{a_n\},\{b_n\}$ 分别进行了差分和逆差分, 在这个想法下, 我们只需要特别关注最后的 a_nB_n 项则可, 而整个证明

过程也只不过是裂项相消后对 B_i 求和.Abel 引理的证明则体现了 Abel 变换的强大之处：在经历过差分和逆差分之后, 得到的 B_i 一般都可以使用题目条件来控制,

Dirichlet 判别法和 Abel 判别法仅仅是 Abel 引理的直接推论, 而 Leibniz 判别法仅仅是 Dirichlet 判别法的简单推论, 我们在此一并列举出来：

定理 6.1

1.（Dirichlet 判别法）

2.（Abel 判别法）

3.（Leibniz 判别法）

6.1.2 正项级数

定理 6.2 基本判别法

定理 6.3 比较判别法

定理 6.4 积分判别法

定理 6.5 Kummer 判别法

注（1）和前面一样, 我们仍然可以通过求极限去寻找 λ, 设

$$\lim_{n\to\infty}\left(\frac{1}{b_n}\frac{a_n}{a_{n+1}}-\frac{1}{b_{n+1}}\right)=\lambda,$$

则当 $\lambda>0$ 时,$\sum\limits_{n=1}^{\infty}a_n$ 收敛; 当 $\lambda<0$ 时, 且 $\sum\limits_{n=1}^{\infty}b_n$ 发散时,$\sum\limits_{n=1}^{\infty}a_n$ 发散.

（2）取 $b_n=1$, 从 Kummer 判别法就直接可以得到 d'Alembert 判别法.

推论 6.1

1.（Raabe 判别法）

2.（Gauss 判别法）

各位对下面的 Cauchy 凝聚判别法可能感到有些莫名其妙, 但是它的思想只不过是证明无穷级数发散的方法的一个变形, 只需要胆大心细的分组放缩就可以证明.

定理 6.6 Cauchy 凝聚判别法

6.1.3 无穷乘积

定义 6.2 无穷乘积

6.1.4 级数的乘积

有限个数的和的乘积, 在我们眼中只是“转换了求和方式”, 也就是使用了求和符号的魔法. 举个简单的例子：

$$\sum_{i=0}^{m}\sum_{j=0}^{n}a_ib_j=(a_0+a_1+\cdots+a_m)(b_0+b_1+\cdots+b_n)=\sum_{k=0}^{m+n}\sum_{i+j=k}a_ib_j=\sum_{k=0}^{m+n}c_k.$$

定义 6.3 Cauchy 乘积

定理 6.7 Cauchy

如果 $\sum\limits_{n=0}^{\infty}a_n$ 和 $\sum\limits_{n=0}^{\infty}b_n$ 都绝对收敛, 则他们的乘积级数也绝对收敛, 且

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

如果将 Cauchy 定理的条件减弱, 这时候下面的结果依然成立.

定理 6.8 Mertens

如果 $\sum\limits_{n=0}^{\infty} a_n$ 和 $\sum\limits_{n=0}^{\infty} b_n$ 都收敛, 且其中至少一个绝对收敛, 则他们的乘积级数也收敛.

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

衡量一个定理的强弱, 可以从两方面来看: 一般来说, 定理的条件越弱, 定理越强; 结果越强, 定理显然也越强, 定理 6.8的条件已经足够弱了, 我们只需要其中一个级数绝对收敛, 就可以得到其乘积级数也收敛, 并且这边绝对收敛的条件不能去掉, 比如取 a_n 和 b_n 均为交错级数 $(-1)^{n-1}\frac{1}{\sqrt{n}}$, 那么所得到的乘积级数就是发散的. 但是, 当我们去掉绝对收敛的条件, 并且加上其乘积级数收敛的条件, 我们就能知道其值一定等于两个级数和的乘积, 这就是下面的 **Abel 定理**.

定理 6.9 Abel 定理

设级数 $\sum\limits_{n=0}^{\infty} a_n$, $\sum\limits_{n=0}^{\infty} b_n$ 收敛, 以及他们的乘积 $\sum\limits_{n=0}^{\infty} c_n$ 也收敛, 那么

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

证明上面的定理, 我们需要下面的 **Abel 引理**,Abel 引理为我们提供了一种对发散级数进行求和的想法, 我们在下一节中就会讲到.

引理 6.3 Abel 引理

如果级数 $\sum\limits_{n=0}^{\infty} a_n$ 收敛, 那么

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n.$$

证明

$\square$

有了引理 6.3, Abel 定理就几乎显然了：对 $x\in(0,1)$, 级数 $\sum\limits_{n=0}^{\infty} a_n x^n$ 和 $\sum\limits_{n=0}^{\infty} b_n x^n$ 都是绝对收敛的, 他们的乘积级数 $\sum\limits_{n=0}^{\infty} c_n x^n$ 也绝对收敛, 根据定理 6.7, 有

$$\sum_{n=0}^{\infty} c_n x^n = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} a_n x^n\right).$$

令 $x \to 1^-$, 根据引理 6.3, 我们就得到了 Abel 定理.

6.2 对数项级数的进一步讨论

6.2.1 无穷乘积

定义 6.4 无穷乘积

设 p_n 是一列实数, 我们将形式乘积

$$\prod_{n=1}^{\infty} p_n = p_1 p_2 \cdots p_n \cdots$$

称为**无穷乘积**, 称 $P_n = p_1 p_2 \cdots p_n$ 为无穷乘积的第 n 个**部分积**, 如果部分积数列 P_n 的极限存在, 且极限为实数或者正负无穷, 我们称此极限为无穷乘积的值, 记为

$$\prod_{n=1}^{\infty} p_n = \lim_{n\to\infty} P_n.$$

当此极限为非零实数的时候, 称这个无穷乘积是**收敛**的, 否则称它是**发散**的.

如果某个 p_n 是零, 显然无穷乘积的值就是零, 下面我们假设每个 p_n 都是非零的, 若无穷乘积 $\prod\limits_{n=1}^{\infty} p_n$ 收敛于 P, 则

$$\lim_{n\to\infty} p_n == \lim_{n\to\infty} \frac{P_n}{P_{n-1}} = \frac{P}{P} = 1.$$

特别地, 当 n 充分大的时候, 必有 $p_n > 0$, 并且由于

$$P_n = \prod_{k=1}^{n} p_k = \exp\left(\sum_{k=1}^{n} \ln p_k\right)$$

我们可以将无穷乘积化为无穷级数加以讨论, 我们有:

定理 6.10

设 p_n 均大于零, 记 $p_n = a_n + 1$, 则

1. 无穷乘积 $\prod\limits_{n=1}^{\infty} p_n$ 收敛的充要条件是级数 $\sum\limits_{n=1}^{\infty} \ln p_n$ 收敛, 且

$$\prod_{n=1}^{\infty} p_n = \exp\left(\sum_{n=1}^{\infty} \ln p_n\right);$$

2. 如果 n 充分大的时候 a_n 不变号, 则无穷乘积 $\prod\limits_{n=1}^{\infty} p_n$ 收敛的充要条件是级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛;
3. 如果级数 $\sum\limits_{n=1}^{\infty} a_n$ 和 $\sum\limits_{n=1}^{\infty} a_n^2$ 都收敛, 则无穷乘积 $\prod\limits_{n=1}^{\infty} p_n$ 也收敛.

这些的证明都很基本, 使用 $\ln(1+x) \sim x \ (x \to 0)$ 以及 $\ln(1+x) \sim x - \dfrac{x^2}{2} \ (x \to 0)$ 与比较判别法即可得到.

我们已经知道了, 很多函数可以用级数的方式来表示, 无穷乘积的美妙之处就在于, 我们也可以将这些函数拿无穷乘积表示, 这里以 $\sin x$、$\sinh x$ 以及 Riemann-zeta 函数 $\zeta(x)$ 为例.

6.2.2 交换求和顺序：级数的重排

定理 6.11 Riemann

如果 $\sum\limits_{n=1}^{\infty} a_n$ 为条件收敛的级数, 则可以将其重排为一个收敛级数, 使得重排后的级数的和为任意指定的实数.

证明

□

6.2.3 级数求和与求极限的可交换性

级数的和是其部分和的极限, 也就是一个数列极限, 那么我们考虑这样的问题：如果有一列数项级数, 他们的和是另一列数, 这列数的极限有什么性质？所以我们考虑依赖于双指标 i, j 的实数列 a_{ij}.

定义 6.5 级数的一致收敛

一列收敛级数 $\sum\limits_{j=1}^{\infty} = \alpha_i$ 关于 i 一致收敛是指：任给 $\varepsilon > 0$, 存在与 i 无关的正整数 $N = N(\epsilon)$, 当 $n > N$ 时, 有

$$\left| \sum_{j=1}^{\infty} a_{ij} - \alpha_i \right| < \varepsilon.$$

下面的定理给出了求极限和求和可交换次序的一个充分条件.

定理 6.12

设一列级数 $\sum\limits_{j=1}^{\infty} a_{ij} = \alpha_i$ 关于 i 一致收敛, 当 $j \geqslant 1$ 时, 有 $\lim\limits_{i\to\infty} a_{ij} = a_j$, 则极限 $\lim\limits_{i\to\infty} \alpha_i$ 存在, 级数 $\sum\limits_{j=1}^{\infty} a_j$ 收敛, 且

$$\lim_{i\to\infty} \alpha_i = \sum_{j=i}^{\infty} a_j \quad \text{或} \quad \lim_{i\to\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \lim_{n\to\infty} a_{ij}.$$

6.2.4 Abel 求和与 Cesàro 求和

定义 6.6 Abel 求和

我们知道, 级数 $\sum\limits_{n=0}^{\infty} (-1)^n$ 在通常意义下是发散的, 但是它在 Abel 意义下就可以求和了：

定义 6.7 Cesàro 求和

Cesàro 可和比 Abel 求和更强一些, 也就是如果某个级数是 Cesàro 可和的, 那么它一定是 Abel 可和的, 且两种意义下的和相等, 这就是下面的定理.

定理 6.13

6.3 函数项级数

设 I 为区间,$\{g_n(x)\}$ 为 I 之中定义的一列函数, 如果对于每一个 $x \in I$, 数列 $\{g_n(x)\}$ 均收敛, 其极限记为 $g(x)$, 那么我们称 $\{g_n(x)\}$ 收敛于 $g(x)$, 注意到这种形式的定义其实是通过**逐点定义**得到的, 我们称这个函数列**点态收敛**于函数 $g(x)$, 记为 $\lim\limits_{n\to\infty} g_n = g$ 或者 $g_n \to g \ (n \to \infty)$.

点态收敛其实很弱, 甚至连保证连续函数列的极限是连续函数都做不到：考虑函数列 $\{x^n\}$, 任给 $x \in (0,1)$, 都有 $\lim\limits_{n\to\infty} x^n = 0$, 但是当 $x = 1$ 的时候,$x^n = 1$, 所以其收敛于的函数 g 不连续. 所以我们需要更强一些的收敛方式, 这就是**一致收敛**.

定义 6.8　一致收敛

设 I 为区间,$\{g_n(x)\}$ 为 I 之中定义的一列函数, 如果任给 $\varepsilon > 0$, 存在与 x 无关的正整数 $N = N(\varepsilon)$, 当 $n > N$,$x \in I$ 时, 有 $|g_n(x) - g(x)| < \varepsilon$, 则称函数列 $\{g_n(x)\}$ 在 I 上**一致收敛**于函数 $g(x)$, 记为 $g_n(x) \rightrightarrows g(x)$.

显然, 如果函数列 $\{g_n(x)\}$ 在 I 上一致收敛于函数 $g(x)$, 那么它在 I 上点态收敛于 $g(x)$. 并且一致收敛还可以改写为：任给 $\varepsilon > 0$, 存在 N, 当 $n > N$ 时,

$$\sup_{x\in I} |g_n(x) - g(x)| < \varepsilon, \text{ 亦即 } \lim_{n\to\infty} \sup_{x\in I} |g_n(x) - g(x)| = 0.$$

一致收敛保持连续性, 这就是下面的定理：

定理 6.14

设 $\{g_n(x)\}$ 在区间 I 上一致收敛于函数 $g(x)$, 如果每一个函数 $g_n(x)$ 都是连续函数, 则 $g(x)$ 也是连续函数

证明

□

上面的定理其实保证了求极限次序的可交换性, 对于共用一个相同的连续点的函数列, 在这个连续点上, 先对这个点取极限和对函数列取极限的结果是相同的. 但是这个连续的条件真的是必须的吗？我们证明一个一般一点的结论, 在这个结论中, 我们甚至不需要连续性就可以保证取极限顺序的可交换性.

定理 6.15

设 $\{g_n(x)\}$ 在 x_0 的一个空心邻域中一致收敛与函数 g, 如果 g_n 在 x_0 的函数极限为 a_n, 则数列极限 $\lim\limits_{n\to\infty} a_n$ 和函数极限 $\lim\limits_{x\to x_0} g(x)$ 都存在, 并且两个极限相等：

$$\lim_{x\to x_0} \lim_{n\to\infty} g_n(x) = \lim_{n\to+\infty} \lim_{x\to x_0} g_n(x)$$

证明

□

更多地, 我们还有**内闭一致收敛**的定义.

定义 6.9 内闭一致收敛

若对于任意给定的闭区间 $[a,b] \subset I$, 函数列 $\{g_n(x)\}$ 在 $[a,b]$ 上一致收敛于函数 $g(x)$, 则称函数列 $\{g_n(x)\}$ 在 I 上**内闭一致收敛**于函数 $g(x)$.

在 D 上一致收敛的函数列一定在 D 上内闭一致收敛, 但是其逆命题不成立.

6.3.1 函数列一致收敛的判别

定理 6.16

设函数列 $\{g_n(x)\}$ 在集合 D 上点态收敛于函数 $g(x)$, 定义 $g_n(x)$ 与 $g(x)$ 的距离为

$$d(g_n,g) = \sup_{x\in D}|g_n(x) - g(x)|.$$

则 $\{g_n(x)\}$ 在 D 上一致收敛于 $g(x)$ 当且仅当 $\lim\limits_{n\to\infty} d(g_n,g) = 0$.

证明

充分性: 若 $\lim\limits_{n\to\infty} d(g_n,g) = 0$, 则任给 $\varepsilon > 0$, 存在 $N > 0$, 当 $n > N$ 时, 有

$$d(g_n,g) = \sup_{x\in D}|g_n(x) - g(x)| < \varepsilon.$$

这就是一致收敛的定义.

必要性: 若 $\{g_n(x)\}$ 在 D 上一致收敛于 $g(x)$, 则任给 $\varepsilon > 0$, 存在 $N > 0$, 当 $n > N$ 时, 任给 $x \in D$, 都有

$$|g_n(x) - g(x)| < \varepsilon.$$

亦即

$$d(g_n,g) = \sup_{x\in D}|g_n(x) - g(x)| < \varepsilon.$$

加之极限的定义, 我们就得到了 $\lim\limits_{n\to\infty} d(g_n,g) = 0$, 这就完成了证明. □

定理 6.17

设函数列 $\{g_n(x)\}$ 在集合 D 上点态收敛于函数 $g(x)$, 则 $\{g_n(x)\}$ 在集合 D 上一致收

敛的充分必要条件是：对于任意数列 $\{x_n\}$，$x_n \in D$, 成立

$$\lim_{n\to\infty}(g_n(x_n) - g(x_n)) = 0.$$

证明

充分性：我们使用反证法, 下面证明：若 $\{g_n(x)\}$ 在 D 上不一致收敛于 $g(x)$, 则存一定存在一数列 $\{x_n\}$，$x_n \in D$, 使得 $g_(x_n) - g(x_n) \nrightarrow 0(n \to \infty)$.

由于 $\{g_n(x)\}$ 在 D 上不一致收敛于 $g(x)$, 则存在 $\varepsilon_0 > 0$, 对于任意的正整数 $N > 0$, 存在 $m > N$, 存在 $x_0 \in D$ 使得 $|g(x_0) - g_m(x_0)| \geqslant \varepsilon_0$.

取 $N = 1$, 则存在 $m_1 > 1$, 存在 $x_{m_1} \in D$ 使得 $|g(x_{m_1}) - g_{m_1}(x_{m_1})| \geqslant \varepsilon_0$. 取 $N = 2$, 则存在 $m_2 > m_1$, 存在 $x_{m_2} \in D$ 使得 $|g(x_{m_2}) - g_{m_2}(x_{m_2})| \geqslant \varepsilon_0$. 重复这个过程, 我们就得到了一个数列 $\{x_n\}$, 其子列 $\{x_{m_n}\}$ 使得 $|g(x_{m_n}) - g_{m_n}(x_{m_n})| \geqslant \varepsilon_0$, 显然这个数列使得 $g_(x_n) - g(x_n) \nrightarrow 0(n \to \infty)$, 充分性得证 7.

必要性：若 $\{g_n(x)\}$ 在集合 D 上一致收敛, 那么对于任意的 $varepsilon > 0$, 存在 $N > 0$, 对于任意的 $m > N$，$x \in D$

$$|g_n(x) - g(x)| < \varepsilon,$$

而 $x_n \in D$, 则有

$$|g_n(x_n) - g(x_n)| < \varepsilon.$$

这就完成了证明 □

定理 6.18 Cauchy 准则

定义在 I 上的函数列 $\{g_n\}$ 一致收敛当且仅当对任给的 $\varepsilon > 0$, 存在与 x 无关的正整数 $N = N(\varepsilon)$, 当 $m > n > N$，$x \in I$ 时, 有 $|g_m(x) - g_n(x)| < \varepsilon$.

证明

□

定理 6.19 Dini 定理

设 $\{g_n\}$ 是 $[a,b]$ 上的一列非负连续函数, 且对每一个 $x \in [a,b]$, $\{g_n(x)\}$ 关于 n 单调递减趋于 0, 则函数列 $\{g_n\}$ 一致收敛于 0.

证明

任给 $\varepsilon > 0$, 我们要证明存在 $N > 0$, 当 $n > N, x \in [a,b]$ 时, 有 $0 \leqslant g_n(x) < \varepsilon$. 设 $A_m = \{x \in [a,b] | g_n(x) \geqslant \varepsilon\}$, 由于

$$A_1 \supset A_2 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots .$$

我们下面只用证明某一个 A_n 为空集, 这样当 n 充分大的时候, 后面所有的 A_m 就都是空集了.

使用反证法, 假设所有的 A_n 都不是空集, 在每个集合之中都取一个 $x_n \in A_n$, 那么 $\{x_n\}$ 就是 $[a,b]$ 中的有界数列, 根据 Bolzano 定理, 有界数列一定有收敛子列, 这个数列的收敛子列设为 $\{x_{n_i}\}$, 并且其收敛到 x_0, 从上可知

$$A_k \supset A_{n_k} \supset \{x_{n_k}, x_{n_{k+1}}, \cdots\}.$$

由于 $g_k(x)$ 在 x_0 连续, 我们有

$$g_k(x_0) = \lim_{i\to\infty} g_k(x_{n_i}) \geqslant \varepsilon.$$

上式对于每一个 $k \geqslant 1$ 都成立, 但是这就和 $\{g_n(x_0)\}$ 收敛于 0 矛盾了, 这就完成了证明. □

6.3.2 一致收敛级数的判别

现在, 设 $\{f_n(x)\}$ 为一列函数, 形如 $\sum\limits_{n=1}^{\infty} f_n(x)$ 的形式和称为函数项级数, 如果其部分和函数 $S_n(x) = \sum\limits_{k=1}^{n} f_k(x)$ 在某点 x 处收敛, 那么称这个函数项级数在该点收敛, 这个点就是该级数的一个收敛点, 所有收敛点的集合就是这个函数项级数的一个收敛域. 在相应的收敛域上,$S(x) = \sum\limits_{n=1}^{\infty} f_n(x)$ 其实就定义了一个的**和函数**. 同理可以定义一致收敛. 对于函数项级数, 下面的性质与判别法与上面的是一致的.

1. 如果 f_n 都为连续函数, 并且 $\sum\limits_{n=1}^{\infty} f_n(x)$ 一致收敛于 $S(x)$, 则 $S(x)$ 也是连续函数;

2. $\sum\limits_{n=1}^{\infty} f_n(x)$ 一致收敛当且仅当对于任给的 $\varepsilon > 0$, 存在与 x 无关的正整数 $N = N(\varepsilon)$, 当 $m > n > N$, $x \in I$ 时, 有

$$|f_{n+1} + f_{n+2} + \cdots + f_m(x)| < \varepsilon.$$

定理 6.20 Weierstrass 判别法

设函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)\ (x \in D)$ 的每一项 $f_n(x)$ 都满足

$$|f_n(x)| \leqslant a_n,\quad x \in D.$$

并且数项级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 则 $\sum\limits_{n=1}^{\infty} f_n(x)$ 在 D 上一致收敛.

证明

使用 Cauchy 准则, 对于任给的 $\varepsilon > 0$, 存在 $N > 0$, 对于任意的 $m > n > N$, $x \in D$, 有

$$|\sum_{k=m+1}^{n} f_k(x)| \leqslant \sum_{k=m+1}^{n} a_k < \varepsilon.$$

显然这个 N 和 x 的选取无关, 根据 Cauchy 准则, 我们就完成了证明. □

定理 6.21 Abel 判别法

如果函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)\ (x \in D)$ 可以写成 $\sum\limits_{n=1}^{\infty} a_n(x)b_n(x)$ 的形式, 且满足函数列 $\{a_n(x)\}$ 对每一个固定的 $x \in D$ 关于 n 单调, 且 $\{a_n(x)\}$ 在 D 上一致有界

$$|a_n(x)| \leqslant M,\quad x \in D,\quad n \in \mathbb{N}^+,$$

函数项级数 $\sum\limits_{n=1}^{\infty} b_n(x)$ 在 D 上一致收敛, 则函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)$ 在 D 上一致收敛.

证明

□

下面的 Dini 定理其实在定理 6.19中证明过了, 这里给出另外一种证明.

定理 6.22 Dini 定理

设连续函数列 $\{S_n(x)\}$ 在闭区间 $[a,b]$ 上点态收敛于连续函数 $S(x)$, 如果 $\{S_n(x)\}$ 关于 n 单调, 那么 $\{S_n(x)\}$ 在 $[a,b]$ 上一致收敛于 $S(x)$.

证明

□

这个定理解决了定理 6.15的逆命题是否成立的问题：如果一个连续函数列点态收敛与一个连续函数, 那么在大多数情况下这个函数列还不是一致收敛的, 但是如果满足一定的条件, 在这里是满足单调性, 那么一直连续性就满足了, 这就是这种形式的 Dini 定理, 其相应的函数项级数形式在下面.

定理 6.23

设函数项级数

定理 6.24 Dirichlet 判别法

如果函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)$ $(x \in D)$ 可以写成 $\sum\limits_{n=1}^{\infty} a_n(x)b_n(x)$ 的形式, 且满足函数列 $\{a_n(x)\}$ 对每一个固定的 $x \in D$ 关于 n 单调, 且 $\{a_n(x)\}$ 在 D 上一致收敛于 0, 同时函数项级数 $\sum\limits_{n=1}^{\infty} b_n(x)$ 在 D 上一致有界

$$|\sum_{k=1}^{n} b_k(x)| \leqslant M, \quad x \in D, \quad n \in \mathbb{N}^+,$$

则函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)$ 在 D 上一致收敛.

证明

□

定理 6.25

设 $f_n(x)$ 在区间 I 上连续且非负, 如果函数项级数 $\sum\limits_{n=1}^{\infty} f_n(x)$ 在区间 I 上收敛于连续函数 $S(x)$, 则 $\sum\limits_{n=1}^{\infty} f_n(x)$ 在区间 I 上一致收敛于 $S(x)$.

证明

□

6.3.3 一致收敛级数的性质

函数项级数存在三个很自然的基本问题：首先我们知道，对于有限个函数的和，我们可以交换求极限与求和的顺序、求导与求和的顺序、积分与求和的顺序，那么对于无穷个函数的和，我们能否交换次序呢？在什么情况下可以交换次序呢？仅仅有点态收敛的条件显然是不够的.

对于求极限与求和的交换顺序的问题，根据证明过的定理 6.15，我们将里面的函数列看作函数项级数的部分函数，就可以得到下面的定理.

定理 6.26　逐项求极限

设对每个 n，$u_n(x)$ 在 $[a,b]$ 上连续，且函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在区间 $[a,b]$ 上一致收敛于 $S(x)$，则 $S(x)$ 在 $[a,b]$ 上连续，且对于任意 $x_0 \in [a,b]$，成立

$$\lim_{x\to x_0}\sum_{k=1}^{\infty} u_k(x) = \sum_{n=1}^{\infty}\lim_{x\to x_0} u_n(x).$$

即极限运算与无限和可以交换顺序.

定理 6.27

设函数列 $\{S_n(x)\}$ 的每一项 $S_n(x)$ 在 $[a,b]$ 上连续，且在 $[a,b]$ 上一致收敛于 $S(x)$，则 $S(x)$ 在区间 $[a,b]$ 上可积，且

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

证明

□

将上述定理中的 $\{S_n(x)\}$ 看成函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的部分和函数列，对应到函数项级数，我们就得到下面的**逐项积分定理：**

定理 6.28 逐项积分定理

设对每个 n, $u_n(x)$ 在 $[a,b]$ 上连续, 且函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在区间 $[a,b]$ 上一致收敛于 $S(x)$, 则 $S(x)$ 在 $[a,b]$ 上可积, 且

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

更进一步地, 对任意固定的 $x_0 \in [a,b]$, 函数列 $\{\int_{x_0}^{x} S_n(t)\,\mathrm{d}t\}$ 在 $[a,b]$ 上一致收敛于 $\int_{x_0}^{x} S(t)\,\mathrm{d}t$, 同样地, 函数项级数 $\{\sum\limits_{n=1}^{\infty}\int_{x_0}^{x} u_n(t)\,\mathrm{d}t\}$ 在 $[a,b]$ 上一致收敛于 $\int_{x_0}^{x} S(t)\,\mathrm{d}t$.

定理 6.29

如果函数列 $\{S_n(x)\}$ 满足：对每一个 n, $S_n(x)$ 在 $[a,b]$ 上有连续的导函数,$\{S_n(x)\}$ 在 $[a,b]$ 上点态收敛于 $S_n(x)$, 且 $\{S_n^{'}(x)\}$ 在 $[a,b]$ 上一致收敛于 $\sigma(x)$, 则 $S(x)$ 在 $[a,b]$ 上可导, 且 $\sigma(x) = S^{'}(x)$.

证明

□

将上述定理中的 $\{S_n(x)\}$ 看成函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 的部分和函数列, 对应到函数项级数, 我们就得到下面的**逐项求导定理**:

定理 6.30 逐项求导定理

设函数项级数 $\sum\limits_{n=1}^{\infty} u_n(x)$ 满足：对每一个 n, $u_n(x)$ 在 $[a,b]$ 上有连续的导函数,$\{u_n(x)\}$ 在 $[a,b]$ 上点态收敛于函数 $S(x)$, 且 $\{u_n^{'}(x)\}$ 在 $[a,b]$ 上一致收敛于 $\sigma(x)$, 则 $S(x)$ 在 $[a,b]$ 上可导, 且

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

6.3.4 幂级数

定义 6.10 幂级数

我们将形如

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + \cdots$$

这样的函数项级数称为**幂级数**.

幂级数可以看作是一个无限次多项式, 其部分和函数 $S_n(x)$ 是一个 $n-1$ 次多项式, 为了方便, 我们一般取 $x_0 = 0$, 只需要将结果做一个简单的平移就可以得到我们真正想要讨论的幂级数了. 幂级数的一致敛散性非常好, 这就有很多奇妙的应用.

对于幂级数 $\sum\limits_{n=1}^{\infty} a_n x^n$, 首先有

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \cdot |x|.$$

根据 Cauchy 判别法, 上式小于 1 的时候, 该幂级数绝对收敛, 大于 1 的时候幂级数发散, 如果令 $A = \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$, 我们有以下的 Cauchy-Hadamard 定理.

定理 6.31 Cauchy-Hadamard 定理

幂级数 $\sum\limits_{n=1}^{\infty} a_n x^n$ 的收敛域关于 $x=0$ 对称, 收敛域区间长的一半被称为**收敛半径**R, 并且满足

$$R = \begin{cases} +\infty & \text{当} A = 0; \\ \dfrac{1}{A} & \text{当} A \in (0, +\infty); \\ 0 & \text{当} A = +\infty. \end{cases}$$

当 $A = +\infty$ 的时候, 收敛半径为 0, 幂级数只在 $x = x_0$ 的时候收敛; 当 $R = +\infty$ 时, 对于一切的 x, 幂级数都是收敛的; 在此之外, 当 $|x| < R$ 的时候, 幂级数收敛,$|x| > R$ 的时候, 幂级数发散. 区间的端点需要特别判断.

下面的 d'Alembert 判别法在幂级数敛散性的判别上也很好用.

定理 6.32 d'Alembert 判别法

如果对幂级数 $\sum\limits_{n=1}^{\infty} a_n x^n$, 下面极限存在

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

那么幂级数的收敛半径为 $R=\dfrac{1}{A}$.

证明

□

定理 6.33 Abel 第二定理

定理 6.34 和函数连续性

定理 6.35 逐项可积性

定理 6.36 逐项可微性

Stirling 公式在幂级数中比较常用, 我们复习一下：

$$n!=\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}\ \ (n\to\infty).$$

6.3.5 函数的幂级数展开

6.4 对函数项级数的进一步讨论

如果函数 $f(x)$ 在 x_0 的某个邻域 $O(x_0,r)$ 可以表示成幂级数

$$f(x)=\sum_{n=0}^{\infty}a_n(x-x_0)^n,\ \ x\in O(x_0,r)$$

也就是说 $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ 在 $O(x_0,r)$ 上的和函数为 $f(x)$, 那么根据幂级数的逐项可导性, 我们有

$$a_k=\frac{f^{(k)}(x_0)}{k!}\ \ (k=0,1,2,\cdots).$$

这时幂级数的系数由和函数唯一确定, 我们称这一系数为函数 $f(x)$ 在 x_0 处的 Taylor **系数**, 而幂级数 $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ 称为函数 $f(x)$ 在 x_0 处的 Taylor **级数**.

定理 6.37 Bernstein

6.5 习题：数项级数

练习 6.1 判断下列正项级数的敛散性

1. $\sum_{n=1}^{\infty}\left[\frac{1}{n}-\ln(1+\frac{1}{n})\right]$;
2. $x>0$, $\sum_{n=1}^{\infty} n!(\frac{x}{n})^n$;
3.

6.6 习题：函数项级数

7 $\mathbb{R}^n$ 上的拓扑

从这一章开始, 我们开始对 $\mathbb{R}^n$ 上的多元函数的研究, 我们仍然希望仿照前面的思路, 按照连续性、可微性、可积性来研究多元函数. 但是 $\mathbb{R}^n$ 上的拓扑结构对我们来说是格外陌生的, 我们需要重新对极限、连续等概念进行重新定义, 另一方面, 我们希望在尽可能提升一般性的同时, 保持先前研究的许许多多优美且强大的性质. 所以从范数到度量, 我们希望把极限的定义推广到一半的度量空间上; 我们还要借助点集拓扑的一些工具, 借以研究连续映射的性质, 并且进一步处理先前的实数系基本定理等知识.

7.1 度量空间

定义 7.1 度量

设 X 是非空集合, 如果映射 $d : X \times X \to \mathbb{R}_{\geqslant 0}$ 满足以下条件:

1. (正定性): 任给 $x, y \in X$, 均有 $d(x, y) \geqslant 0$, 且 $d(x, y) = 0$ 当且仅当 $x = y$;
2. (对称性): 任给 $x, y \in X$, 均有 $d(x, y) = d(y, x)$;
3. (三角不等式): 任给 $x, y, z \in X$, 均有 $d(x, z) \leqslant d(x, y) + d(y, z)$.

则称 d 是 X 上的一个度量或者距离, 称二元组 (X, d) 是一个度量空间或者距离空间.

在数学中, 所谓的空间一般指的是配备了某种结构的集合 X. 我们使用 $\mathbb{R}^n$ 上的标准内积来定义距离, 显然 $\mathbb{R}^n$ 是关于 $\mathbb{R}$ 的线性空间, 我们将有限维的内积空间称为欧式空间.

定义 7.2 内积

设 V 是实数域上的线性空间, 如果映射 $g: V\times V \to \mathbb{R}$ 满足以下条件：

1. （正定性）：任给 $x\in V$, 均有 $g(x,x)\geqslant 0$, 且 $g(x,x)=0$ 当且仅当 $x=0$;
2. （对称性）：任给 $x,y\in V$, 均有 $g(x,y)=g(y,x)$;
3. （线性性）：任给 $x,y,z\in V$ 和 $\lambda,\mu\in\mathbb{R}$, 均有 $g(\lambda x+\mu y,z)=\lambda g(x,z)+\mu g(y,z)$.

则称 g 是 V 上的一个内积, 称二元组 (V,g) 是一个内积空间. 我们常常使用 $\langle,\rangle$ 表示内积, 比如 $\langle x,y\rangle$ 表示 x,y 的内积.

对于 $\mathbb{R}^n$ 上的两个点 $u=(a_1,a_2,\ldots,a_n), v=(b_1,b_2,\ldots,b_n)$, 他们的标准内积或者欧式内积定义为：

$$\langle u,v\rangle=\sum_{i=1}^n a_ib_i.$$

定义 7.3 范数

V 是 $\mathbb{R}$(或者 $\mathbb{C}$) 上的线性空间, 如果映射 $\|\cdot\|: V\to\mathbb{R}_{\geqslant 0}$ 满足以下条件：

1. 对任意的 $x\in V,\lambda\in\mathbb{R}$, 有 $\|\lambda x\|=|\lambda|\|x\|$;
2. $\|x\|=0$ 当且仅当 $x=0$;
3. 对任意的 $x,y\in V$, 有 $\|x+y\|\leqslant\|x\|+\|y\|$.

则称 $\|\cdot\|$ 是 V 上的一个范数, 称二元组 $(V,\|\cdot\|)$ 是一个赋范线性空间.

直观来说,$\|v\|$ 就是计算向量 v 的某种长度, 而度量则表示了两个点之间的某距离. 有了内积就可以定义向量的长度（亦即范数）和向量的夹角, 这来自于下面的 **Schwarz** 不等式; 有了范数就可以定义度量, 只需要定义 $d(x,y)=\|x-y\|$ 即可.

定理 7.1 Schwarz 不等式

设 $(V,\langle,\rangle)$ 为内积空间,$u,v\in V$, 则

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

等号成立当且仅当 u,v 线性相关.

证明

□

根据 Schwarz 不等式, 当 u, v 为非零向量的时候, 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 的夹角, 也记为 $\angle(u,v)$.

7.2 基本点集拓扑

定义 7.4 开集和闭集

定理 7.2 开集闭集的基本性质

1. 有限多个开集的交仍为开集, 任意多个开集的并仍为开集;
2. 有限多个闭集的交仍为闭集, 任意多个闭集的并仍为闭集;
3. 集合 A 为闭集当且仅当 A 中的任何收敛点列的极限均在 A 中.

证明

□

上述定理的第三条表明了闭集的性质：闭集关于求极限运算是封闭的.

定义 7.5 连续映射

定义 7.6 连续映射的基本性质

定理 7.3 连续映射的刻画

设 $f: X \to Y$ 是度量空间上的映射, 则 f 为连续映射 $\iff$ 开集的原像仍为开集 $\iff$ 闭集的原像仍为闭集.

8

多元函数微分学

section 方向导数和微分

研究多元函数的常用办法是对函数进行局部的线性化, 而在某处的最佳线性函数就是我们研究的对象, 这就是下面微分的定义.

定义 8.1　微分

设 $D \in \mathbb{R}^n$ 为开集, $f : D \to \mathbb{R}$ 为一个多元函数, $x^0 \in D$, 如果存在一个线性映射 $L : \mathbb{R}^n \to \mathbb{R}$, 使得在 x^0 的附近成立:

$$f(x) - f(x^0) = L(x - x^0) + o(\|x - x^0\|) \ \ (x \to x^0),$$

则称函数 f 在 x^0 处可微, 线性映射 L 称为函数 f 在 x^0 处的微分, 记为 $\mathrm{d}f(x^0)$.

9

多元函数积分学

10

曲线积分和曲面积分

10.1　曲线和曲面

10.1.1　曲面的第一基本形式

一个曲面 Ω 的参数方程为

$$\boldsymbol{r} = \boldsymbol{r}(u, v),\ (u, v) \in D$$

我们假定 $\boldsymbol{r} = \boldsymbol{r}(u, v)$ 连续可微足够多次并且满足正则条件：$\boldsymbol{r}_u \times \boldsymbol{r}_v \neq 0,\ \forall (u, v) \in D$. 考察曲面上的一条曲线 $L : \boldsymbol{r} = \boldsymbol{r}(u(t), v(t)),\ t \in J$, 且 $u(t)$ 和 $v(t)$ 都连续可微. 上式对 t 求导得：

$$\boldsymbol{r}' = \boldsymbol{r}_u \frac{\mathrm{d}u}{\mathrm{d}t} + \boldsymbol{r}_u \frac{\mathrm{d}u}{\mathrm{d}t},\ \ \mathrm{d}\boldsymbol{r} = \boldsymbol{r}_u \mathrm{d}u + \boldsymbol{r}_v \mathrm{d}v.$$

曲线 L 的弧长微元可以表示为:

$$\mathrm{d}s = \|\boldsymbol{r}'\|\mathrm{d}t = \pm\|\boldsymbol{r}'\mathrm{d}t\| = \pm\|\mathrm{d}\boldsymbol{r}\|.$$

进而有:

$$\mathrm{d}s^2 = \|\mathrm{d}\boldsymbol{r}\|^2 = \langle \mathrm{d}\boldsymbol{r}, \mathrm{d}\boldsymbol{r} \rangle = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2.$$

其中:

$$E = \boldsymbol{r}_u^2,\ \ F = \boldsymbol{r}_u \cdot \boldsymbol{r}_v,\ \ G = \boldsymbol{r}_v^2.$$

并且记

$$I(\mathrm{d}u, \mathrm{d}v) = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2.$$

于是曲面 Ω 上的曲线 L 的弧长可以按照这样计算:

$$s = s_0 + \int_{t_0}^{t} \sqrt{E\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 + 2F\frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}v}{\mathrm{d}t} + G\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^2}\mathrm{d}t = s_0 + \int_{t_0}^{t} \sqrt{I\left(\frac{\mathrm{d}u}{\mathrm{d}t}, \frac{\mathrm{d}v}{\mathrm{d}t}\right)}\mathrm{d}t.$$

于是我们将微分 $\mathrm{d}u$ 和 $\mathrm{d}v$ 的二次型

$$I = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + F\mathrm{d}v^2$$

称为曲面 Ω 的第一基本形式. 曲面上的曲线的弧长取决于这个曲面的第一基本形式, 后面将要见到, 曲线块的面积也取决于这个曲面的第一基本形式. 因而有：曲面的第一基本形式决定了曲面的度量性质.

10.2 第一类曲线积分

10.3 第二类曲线积分

10.4 第一类曲面积分

10.5 第二类曲面积分

定义 10.1 第二类曲面积分

设 Σ 为 $\mathbb{R}^3$ 中的可定向曲面,φ 是与给定定向相容的参数表示：

$$\varphi(u,v) = (x(u,v), y(u,v), z(u,v)),\ \ (u,v) \in D.$$

对于定义在 Σ 上的连续向量值函数 $X = (P,Q,R)$, 我们定义 X 在 Σ 上的第二类曲面积分为：

$$\Phi = \int_D \left[P\frac{\partial(y,z)}{\partial(u,v)} + Q\frac{\partial(z,x)}{\partial(u,v)} + R\frac{\partial(x,y)}{\partial(u,v)}\right]\mathrm{d}u\mathrm{d}v.$$

也记为:

$$\Phi = \int_\Sigma P\mathrm{d}y \wedge \mathrm{d}z + Q\mathrm{d}z \wedge \mathrm{d}x + R\mathrm{d}x \wedge \mathrm{d}y.$$

第二类曲面积分的定义来自于下面的物理问题：空间中有流速为 $V = (P,Q,R)$ 的流体, 流体通过曲面 Σ 的流量是多少？要规定流量, 首先就要给曲面指定方向. 我们规定曲面在某点的方向在这点的一个单位法向量的方向. 指定了方向之后, 我们利用微元法：任取 Σ 的一个小片, 其面积记为 $\mathrm{d}\sigma$, 曲面的单位法向量为 $\boldsymbol{n}$, 则流体通过这个小片的流量 $\mathrm{d}\Phi$ 为 $V \cdot \boldsymbol{n}\mathrm{d}\sigma$. 于是整个曲面的流量为

$$\Phi = \int_\Sigma V \cdot \boldsymbol{n}\mathrm{d}\sigma = \int_\Sigma V \cdot \mathrm{d}\boldsymbol{\sigma}. \tag{10.1}$$

记 $\mathrm{d}\boldsymbol{\sigma} = (\mathrm{d}y \wedge \mathrm{d}z, \mathrm{d}z \wedge \mathrm{d}x, \mathrm{d}x \wedge \mathrm{d}y)$, 其中 $\mathrm{d}y \wedge \mathrm{d}z$ 是有向面积元 $\mathrm{d}\boldsymbol{\sigma}$ 在 yz 平面的投影,$\mathrm{d}z \wedge \mathrm{d}x$ 是有向面积元 $\mathrm{d}\boldsymbol{\sigma}$ 在 zx 平面的投影,$\mathrm{d}x \wedge \mathrm{d}y$ 是有向面积元 $\mathrm{d}\boldsymbol{\sigma}$ 在 xy 平面的投

影. 于是上式可写为:

$$\Phi = \int_\Sigma V\cdot \boldsymbol{n}\mathrm{d}\sigma = \int_\Sigma V\cdot \mathrm{d}\boldsymbol{\sigma} = \int_\Sigma P\mathrm{d}y\wedge \mathrm{d}z + Q\mathrm{d}z\wedge \mathrm{d}x + R\mathrm{d}x\wedge \mathrm{d}y. \tag{10.2}$$

这其实就表明了第二类曲面积分可以转化成第一类曲面积分, 并且给出了转换的方法.

另一方面, 流量与曲面方向的选取有关, 方向的变化可导致流量的数值差一个正负号. 如果曲面上存在连续的单位法向量场, 则称该曲面可定向, 否则就称该曲面不可定向. 本节涉及的曲面都是可定向的, 其定向（方向）是指一个连续的单位法向量场 $\boldsymbol{n}$.

为了计算式(10.1), 我们先对曲面选取恰当的参数表示：设 $\varphi: D\to\mathbb{R}^3$ 为 Σ 的参数表示, 其中

$$\varphi(u,v) = (x(u,v), y(u,v), z(u,v)),\ \ (u,v)\in D.$$

记 $N=\varphi_u\times\varphi_v$, 则:

$$N = (y_u z_v - z_u y_v, z_u x_v - x_u z_v, x_u y_v - y_u x_v) = \left(\frac{\partial(y,z)}{\partial(u,v)} + \frac{\partial(z,x)}{\partial(u,v)} + \frac{\partial(x,y)}{\partial(u,v)}\right)$$

N 为曲面的法向量. 如果 $\dfrac{N}{\|N\|} = \boldsymbol{n}$, 则称 φ 是与给定定向相容的参数表示. 我们总是选取与给定定向相容的参数表示. 而从上节的讨论可知, 曲面的面积元可写为 $\mathrm{d}\sigma = \|N\|\mathrm{d}u\mathrm{d}v$, 于是有:

$$\Phi = \int_D V\cdot N\mathrm{d}u\mathrm{d}v = \int_D \left[P\frac{\partial(y,z)}{\partial(u,v)} + Q\frac{\partial(z,x)}{\partial(u,v)} + R\frac{\partial(x,y)}{\partial(u,v)}\right]\mathrm{d}u\mathrm{d}v. \tag{10.3}$$

由此看出式(10.3)和式(10.2)是等价的. 这就是第二类曲面积分的来源.

值得注意的是, 为了书写的简便, 我们有时候也会将 $\mathrm{d}x\wedge\mathrm{d}y, \mathrm{d}y\wedge\mathrm{d}z$ 和 $\mathrm{d}z\wedge\mathrm{d}x$ 等记号简写为 $\mathrm{d}x\mathrm{d}y, \mathrm{d}y\mathrm{d}z$ 和 $\mathrm{d}z\mathrm{d}x$. 比如积分 $\int_\Sigma f(x,y,z)\mathrm{d}x\wedge\mathrm{d}y$ 就可以简写为 $\int_\Sigma f(x,y,z)\mathrm{d}x\mathrm{d}y$.

例 10.1

计算积分 $\Phi = \int_\Sigma x^2\mathrm{d}y\wedge\mathrm{d}z + y^2\mathrm{d}z\wedge\mathrm{d}x + z^2\mathrm{d}x\wedge\mathrm{d}y$, 其中 Σ 是球面 $(x-a)^2+(y-b)^2+(z-c)^2=R^2$, 方向为外侧.

10.6 Green 公式、Gauss 公式和 Stokes 公式

10.6.1 Green 公式

定理 10.1 Green 公式

设 Ω 为 $\mathbb{R}^2$ 上的有界区域, 其边界由有限条 C^1 曲线组成, 曲线的定向为诱导定向, 如果 P,Q 为 Ω 上的连续可微函数, 则

$$\int_{\Omega}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y=\int_{\partial\Omega}P\mathrm{d}x+Q\mathrm{d}y.$$

例 10.2

设 Ω 为包含原点的有界区域, 其边界为 C^1 曲线, 方向为诱导定向, 计算积分

$$I=\int_{\partial\Omega}\frac{-y\mathrm{d}x}{x^2+y^2}+\frac{x\mathrm{d}y}{x^2+y^2}.$$

10.6.2 Gauss 公式

定理 10.2 Gauss 公式

设 Ω 为 $\mathbb{R}^3$ 中的有界区域, 其边界由有限个 C^1 曲面组成, 曲面的定向为诱导定向, 如果 P,Q,R 为 Ω 上的连续可微函数, 则

$$\int_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z=\int_{\partial\Omega}P\mathrm{d}y\wedge\mathrm{d}z+Q\mathrm{d}z\wedge\mathrm{d}x+R\mathrm{d}x\wedge\mathrm{d}y.$$

对于以函数 P,Q,R 为分量的 C^1 的向量场 $X=(P,Q,R)$, 其散度 $\nabla\cdot X$（或 $\mathrm{div}X$）定义为

$$\mathrm{div}X=\nabla\cdot X=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}.$$

散度是将 Del 算符 (∇) 点积一个向量场 X, 效果是将一个向量场转换成了一个标量场. 利用散度,Gauss 公式可以写为以下形式:

$$\int_{\Omega}\nabla\cdot X\mathrm{d}x\mathrm{d}y\mathrm{d}z=\int_{\partial\Omega}X\cdot\boldsymbol{n}\mathrm{d}\sigma.$$

其中 $\boldsymbol{n}$ 是边界曲面的单位外法向量, 上式也被称为散度定理.

例 10.3 Green 恒等式

设

10.6.3　Stokes 公式

设 Σ 为 $\mathbb{R}^3$ 中的 C^2 的定向曲面,Ω 为 Σ 中的有界区域, 其边界为 C^1 曲线, 其由右手定则定义的诱导定向如下：边界在曲面上的外法向量与边界的切向量的外积得到的曲面的法向量与决定曲面定向的法向量同向. 即如果用右手从曲线外法向到切向做旋转, 则大拇指所指的方向为定向曲面的法向.

定理 10.3　Stokes 公式

设 Σ 为 $\mathbb{R}^3$ 中的 C^2 的定向曲面,Ω 为 Σ 中的有界区域, 其边界赋以诱导定向, 如果 P, Q, R 为 Ω 附近的连续可微函数, 则

$$\int_{\Sigma}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}y\wedge\mathrm{d}z+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}z\wedge\mathrm{d}x+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\wedge\mathrm{d}y$$
$$=\int_{\partial\Sigma}P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z.$$

证明

我们只证明一个特殊情形：　□

设 $X=(P,Q,R)$ 为 C^1 的向量场, 其旋度场 $\nabla\times X$（或 $\mathrm{rot}X$）定义为

$$\nabla\times X=\mathrm{rot}X=\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z},\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x},\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right).$$

利用旋度,Stokes 公式可以写为以下形式：

$$\int_{\Omega}\nabla\times X\cdot\boldsymbol{n}\mathrm{d}\sigma=\int_{\partial\Omega}X\cdot T\mathrm{d}s.$$

其中 $\boldsymbol{n}$ 是曲面 Ω 的单位外法向量,T 是曲线 $\partial\Omega$ 的单位切向量,$\mathrm{d}s$ 是曲线 $\partial\Omega$ 的弧长参数.

10.7　余面积公式

定理 10.4　余面积公式

设 $f:\mathbb{R}^n\to\mathbb{R}$ 为 C^1 函数, 且 $\|\nabla f\|\neq 0$, 如果 g 为区域 $f^{-1}([a,b])$ 上的连续函数, 则

$$\int_{f^{-1}([a,b])}g(x)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=\int_a^b\mathrm{d}t\int_{f^{-1}(t)}\frac{g}{\|\nabla f\|}\mathrm{d}\sigma.$$

证明

□

使用类似的推导方法, 我们可以证明特殊情形的余面积公式, 留作后面的习题.

10.8 习题

10.8.1 余面积公式

1. 设 $f:\mathbb{R}^n\to\mathbb{R}$ 为 C^1 函数, 且 $\|\nabla f\|\neq 0,\Omega$ 的含义与定理 10.4的证明中的相同, 证明:

$$\nu(\Omega)=\int_a^b \mathrm{d}t\int_{f^{-1}(t)\cap\Omega}\frac{1}{\|\nabla f\|}\mathrm{d}\sigma.$$

提示: 如果将 g 取为常值函数 1, 那么此题显然成立, 但是请使用定理 10.4的证明中的方法重新证明.

2. 使用余面积公式计算积分

$$I=\int_\Sigma\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{z^2}{c^4}\right)^{-\frac{1}{2}}\mathrm{d}\sigma,$$

其中 Σ 为椭球面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$.

11

微分形式与基本场论

本质上讲, 流形就是曲线和曲面在更高维空间上的一般化, 在本章, 我们的目标是将

11.1 线性代数基础

11.1.1 对偶空间

定义 11.1 对偶空间

设 V 是一个 n 维实向量空间, $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组基, 定义

$$V^* = L(V, \mathbb{R})$$

为 V 的**对偶空间**, V^* 中的元素被称为 V 上的**余向量**（covector 或 1-covector）

根据已有的线性代数知识, 我们知道 $\dim V^* = \dim V = n$, 很容易能证得 $\{e^1, e^2, \cdots, e^n\}$ 是 V^* 的一组基, 其中

$$e^i(e_j) = \delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

δ^i_j 是克罗内克符号.

我们可以把这组对偶基理解为坐标的投影, 对于 $v \in V$, 它的坐标是 $\{v^1, v^2, \cdots, v^n\}$, 那么 $e^j v = v^j$. 这就相当于使用 e^j 取出了 v 的第 j 个坐标. 因而, 我们可以使用对偶基将任意一个向量写成这样的形式:

$$v = e^1(v)e_1 + e^2(v)e_2 + \cdots + e^n(v)e_n.$$

11.1.2 置换

11.1.3 多重线性函数

V 是一个实线性空间, 如果函数 $f: V^k \to \mathbb{R}$ 对它 k 个系数中的每一个都是线性的, 亦即满足

$$f(\dots, \lambda v + \mu w, \dots) = \lambda f(\dots, v, \dots) + \mu f(\dots, w, \dots),$$

对任意的 $\lambda, \mu \in \mathbb{R}, v, w \in V$, 则称 f 是 V 上的一个 **k 重线性函数**, 也叫做 V 上的 **k-张量**, 特别地, 2 重线性函数也叫做双线性函数, k 叫做 f 的阶. V 上的所有 k-张量构成一个线性空间, 记作 $L_k(V)$.

定义 11.2 对称与交替

对 V 上的一个 k 重线性函数 $f\colon V^k \to \mathbb{R}$ 来说, 如果对任意的 $\sigma \in S_k$, 都有

$$f(v_{\sigma(1)}, \cdots, v_{\sigma(n)}) = f(v_1, \cdots, v_n)$$

则称 f 是**对称**的; 如果对任意的 $\sigma \in S_k$, 都有

$$f(v_{\sigma(1)}, \cdots, v_{\sigma(n)}) = (\operatorname{sgn}\sigma) f(v_1, \cdots, v_n)$$

则称 f 是**交替**的.

像我们熟悉的内积 $\langle , \rangle$ 是对称的二重线性函数; 行列式 $\det(v_1, \cdots, v_n)$ 是交替的 n 重线性函数, 也叫 **k 阶余向量**. 我们对于交替的多重线性函数尤其感兴趣, 将所有线性空间 V 上的 k 重交替线性函数的集合记为 $A_k(V)$. 特别地, $k = 0$ 时, 我们定义 0 阶交替线性函数为常数, 这样 $A_0(V)$ 就是线性空间 $\mathbb{R}$, 而且 $A_1(V) = L_1(V) = V^*$.

设 $f \in L_k(V)$, $\sigma \in S_k$, 置换 σ 对 f 的**作用**得到一个新的 k 重线性函数, 定义为

$$(\sigma f)(v_1, \dots, v_k) = f(v_{\sigma(1)}, \dots, v_{\sigma(k)}).$$

这样, f 是对称的当且仅当对任意的 $\sigma \in S_k$, $\sigma f = f$, f 是交替的当且仅当对任意的 $\sigma \in S_k$, $\sigma f = (\operatorname{sgn}\sigma) f$.

我们可以把置换对线性函数的作用理解为改变参数的位置, 那么先后交换两次参数的位置和一步到位交换参数的位置的效果是相同的:

$$\begin{aligned}
\tau(\sigma f)(v_1, \dots, v_k) &= (\sigma f)(v_{\tau(1)}, \dots, v_{\tau(k)}) \\
&= f(v_{\tau(\sigma(1))}, \dots, v_{\tau(\sigma(k))}) = f(v_{(\tau\sigma)(1)}, \dots, v_{(\tau\sigma)(k)}) \\
&= (\tau\sigma) f(v_1, \dots, v_k).
\end{aligned}$$

上面的证明只需要注意好置换对应下标的变化就好了. 所以置换对线性函数的作用满足结合律, 即对任意的 $\sigma,\tau\in S_k$, 以及 V 上的 k 重线性函数 f, 都满足

$$(\sigma(\tau f))=((\sigma\tau)f).$$

更一般地来说, 我们可以定义群对集合的作用：设 G 是一个群,X 是一个集合, G 对 X 的**左作用**是一个映射：

$$G\times X\to X,\ \ (\sigma,x)\mapsto\sigma\cdot x.$$

这个映射满足结合律和单位元的性质, 即对任意的 $\sigma,\tau\in G$ 和 $x\in X$, e 是 G 中的单位元：

$$(\sigma\tau)\cdot x=\sigma\cdot(\tau\cdot x),\ \ e\cdot x=x.$$

对于给定的 $x\in X$, 它的**轨道**是所有形如 $\sigma{\cdot}x$ 的元素构成的集合：$Gx\coloneqq\{\sigma{\cdot}x\mid\sigma\in G\}$. 在这些术语下, 我们其实定义了置换群 S_k 作用在线性空间 V 上的 k 重线性函数的空间 $L_k(V)$ 上的左作用. 值得注意的是, k 重线性函数的空间 $L_k(V)$ 上的左作用是线性的, 我们留作习题. 类似地, 我们可以定义**右作用**, 在此不多赘述.

任给一个 k 重线性函数 $f\in L_k(V)$, 我们可以利用它生成一个对称的 k 重线性函数 Sf 和一个交替的 k 重线性函数 Af, 它们分别定义为

$$(Sf)(v_1,\dots,v_n)=\sum_{\sigma\in S_k}f(v_{\sigma(1)},\dots,v_{\sigma(k)}),$$

和

$$(Af)(v_1,\dots,v_n)=\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)f(v_{\sigma(1)},\dots,v_{\sigma(k)}).$$

或者使用更加精简的术语：

$$Sf=\sum_{\sigma\in S_k}\sigma f,\ \ Af=\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)\sigma f.$$

可以证明 Sf 和 Af 确实分别是对称和交替的, 我们只对后者给出证明：对任意的 $\tau\in S_k$, 当 σ 遍历 S_k 时,$\tau\sigma$ 也遍历 S_k, 所以有：

$$\begin{aligned}\tau(Af)&=\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)\tau(\sigma f)\\&=\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)(\tau\sigma)f=\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)^2(\tau\sigma)f\\&=(\operatorname{sgn}\tau)\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma\tau)(\tau\sigma)f\\&=(\operatorname{sgn}\tau)Af.\end{aligned}$$

这样就给出了证明, 对于 Sf 的证明就更容易了, 留作习题.

11.1.4 张量积与楔积

定义 11.3 张量积

f 和 g 分别是线性空间 V 上的 k 重和 l 重线性函数, 它们的**张量积**$f\otimes g$ 是 V 上的 $k+l$ 重线性函数, 定义为

$$(f\otimes g)(v_1,\dots,v_k,v_{k+1},\dots,v_{k+l})=f(v_1,\dots,v_k)g(v_{k+1},\dots,v_{k+l}).$$

对于张量积, 一般不会谈它的交换律, 但是张量积满足结合律, 即对线性空间 V 上的任意的多重线性函数 f,g,h, 都有

$$(f\otimes g)\otimes h=f\otimes(g\otimes h).$$

不失一般性, 我们假设 f,g,h 分别是 k,l,m 重线性函数, 那么对于任意的 $v_1,\dots,v_{k+l+m}\in V$, 都有:

$$\begin{aligned}((f\otimes g)\otimes h)(v_1,\dots,v_{k+l+m})&=(f\otimes g)(v_1,\dots,v_k,v_{k+1},\dots,v_{k+l})h(v_{k+l+1},\dots,v_{k+l+m})\\&=f(v_1,\dots,v_k)g(v_{k+1},\dots,v_{k+l})h(v_{k+l+1},\dots,v_{k+l+m})\\&=f(v_1,\dots,v_k)(g\otimes h)(v_{k+1},\dots,v_{k+l+m})\\&=(f\otimes(g\otimes h))(v_1,\dots,v_{k+l+m}).\end{aligned}$$

例 11.1 张量积视角下的双线性函数

令 $e_1,\dots,e_n$ 是 V 的一组基, 其对偶基为 $e^1,\dots,e^n$, 我们研究 V 上的双线性函数 $\langle,\rangle$. 对于任意的 $v,w\in V$, $v=\sum e^i(v)e_i$, $w=\sum e^i(w)e_i$, 记 $g_{ij}=\langle e_i,e_j\rangle$, 那么我们可以将双线性函数 $\langle,\rangle$ 以张量积的形式表示:

$$\langle v,w\rangle=\sum_{1\leqslant i,j\leqslant n}e^i(v)e^j(w)\langle e_i,e_j\rangle=\sum_{1\leqslant i,j\leqslant n}g_{ij}(e^i\otimes e^j)(v,w).$$

这其实表明双线性函数的取值其实完全由其在基上的取值决定.

如果线性空间 V 上的多重线性函数 f,g 都是交替的, 它们的张量积 $f\otimes g$ 的交替性并不是良好定义的, 这就引导我们定义一种形式的积, 让它具有交替性, 这就引出了下面的**楔积**的概念.

定义 11.4 楔积

楔积定义中的系数 $\dfrac{1}{k!l!}$ 是为了弥补求和中的重复:

另外一种避免求和中的冗余的方法是对求和的方式加以限制：如果 $\sigma(1,\dots,\sigma(k))$ 和 $\sigma(k+1),\dots,\sigma(k+l)$ 都是递增的, 亦即

$$\sigma(1)<\dots<\sigma(k),\ \ \sigma(k+1)<\dots<\sigma(k+l),$$

则称置换 $\sigma\in S_{k+l}$ 是一个 (k,l)-shuffle, 这样楔积的定义就可以重写为：

$$(f\wedge g)(v_1,\dots,v_{k+l})=\sum_{(k,l)-\text{shuffle }\sigma}(\operatorname{sgn}\sigma)f(v_{\sigma(1)},\dots,v_{\sigma(k)})g(v_{\sigma(k+1)},\dots,v_{\sigma(k+l)}).$$

这样定义的楔积只需要对 C_{k+l}^k 个置换求和, 而不是对 $C^{k+l}!$ 个置换求和, 就避免了冗余.

楔积满足斜交换律和结合律, 但是这两个性质的证明都不简单, 我们先证明一个引理：

引理 11.1

如果 f,g 分别是线性空间 V 上的 k 重和 l 重线性函数, 那么:

$$A(A(f)\otimes g)=k!A(f\otimes g),\ \ A(f\otimes A(g))=l!A(f\otimes g).$$

定理 11.1 楔积的斜交换律与结合律

设 f,g,h 分别是线性空间 V 上的 k,l,m 重线性函数, 那么有：

$$f\wedge g=(-1)^{kl}g\wedge f,\ \ (f\wedge g)\wedge h=f\wedge(g\wedge h)$$

在上述证明中, 我们发现了:

$$f\wedge g\wedge h=\frac{1}{k!l!m!}A(f\otimes g\otimes h).$$

我们可以将楔积推广到任意数目的参数上：如果 $f_i\in A_{d_i}(V)$, 那么

$$f_1\wedge\dots\wedge f_k=\frac{1}{d_1!\cdots d_k!}A(f_1\otimes\dots\otimes f_k).$$

11.1.5 副产品：行列式

其实就是练习中的 3.8

11.1.6 外代数

定义 11.5 代数

如果一个线性空间 A 满足下面条件:

定义 11.6 外代数

对于有限维线性空间 V, 若 $\dim V = n$

外代数其实是一个斜交换的分次代数（挖坑）

$A_*(V)$ 作为一个线性空间, 我们应该了解它的最基本结构：维数和基. 而我们只需要研究 $A_k(V)$ 就可以了. 我们首先引入下面的记号：

引理 11.2

定理 11.2

k 重交替线性函数 α^I, $I = (i_1 < \cdots i_k)$, 构成了 $A_k(V)$ 的一组基.

这个定理有两个直接推论：

维数

如果 $k > \dim V$, 那么 $A_k(V) = \{0\}$.

赫尔曼 • 格拉斯曼在十九世纪提出了外代数的概念, 他穷极一生, 建立起了以外代数为基础的大厦, 将向量值的微积分从 $\mathbb{R}^3$ 推广到了 $\mathbb{R}^n$. 然而, 格拉斯曼的成果在他生前并未得到应有的认可与重视, 事实上, 由于当时的领军人物莫比乌斯与库尔默并不能理解他的工作, 格拉斯曼的论文被拒, 更无法得到在大学的职位. 直到十九世纪与二十世纪之交, 外代数的终于在微分几何大师嘉当的手中大放光芒, 成为了微分形式的代数基础.

11.2 欧氏空间的微分形式

我们首先回顾一下梯度场和全微分的概念：设 $D \subset \mathbb{R}^n$ 是一个开集, $f : D \to \mathbb{R}$ 为多元函数, 若 f 在 x^0 处可微, 则 f 在 x^0 处的梯度 $\nabla f(x^0)$ 定义为

$$\nabla f(x^0) = \left(\frac{\partial f}{\partial x_1}(x^0), \frac{\partial f}{\partial x_2}(x^0), \cdots, \frac{\partial f}{\partial x_n}(x^0) \right).$$

并且, f 在 x^0 处的全微分 $\mathrm{d}f(x^0)$ 可以用梯度表示为

$$\mathrm{d}f(x^0) : \mathbb{R}^n \to \mathbb{R},\ \ u \mapsto \nabla f(x^0) \cdot u.$$

我们完全可以使用线性代数中的对偶来解释：

11.3 习题

11.3.1 线性代数基础

1. 证明：k 重线性函数的空间 $L_k(V)$ 上的左作用是线性的. 亦即对于任意的 $f, g \in L_k(V)$, $\sigma \in S_k$ 和 $\lambda \in \mathbb{R}$, 都有:

$$\sigma(f + \lambda g) = \sigma f + \lambda \sigma g.$$

2. 若 f 是线性空间 V 上的 k 重交替线性函数, 证明：

$$Af = (k!)f.$$

12

Fourier 分析