矩阵论讲义

许老师

2022 年 6 月 27 日

目录

关于

课程名称：矩阵论
学时：40
描述：
教材 [1] 及参考资料：[2, 1, 3, 4].

0.1 感谢

感谢认真听课的同学们，没有他们此文稿不会存在.

0.2 记号

有些记号不一定能用上，但我们先列出来.

- $A := B$ 或者 $A \overset{\text{def}}{=} B$，表示 A 定义为 B.
- “命题 A 是真的，由定义，如果命题 B 是真的.” 可以用 $A :\iff B$ 表示. 例如，$A \subseteq B$ 是指 $\forall x \in A$ 有$x \in B$，可以写成 $A \subseteq B :\iff \forall x \in A, x \in B$.

设 E 是某个性质，$E(x)$，表示 x 具有性质 E.

- 记号 $\{x \in X; E(x)\}$ 或者 $\{x \in X \mid E(x)\}$ 表示集合 X 中具有性质 E 的元素 x 构成的集合. 我们习惯用竖线这种表示方式 $\{x \in X \mid E(x)\}$.
- $\exists x \in X : E(x)$ 或者 $\exists x \in X, E(x)$，表示存在 $x \in X$ 使得 x 有性质 E.
- $\forall x \in X : E(x)$ 或者 $\forall x \in X, E(x)$，表示对于任意 $x \in X$，x 都有性质 E，$E(x)$ 是真.

- “对于所有的 $x \in X$，$E(x)$ 成立”，这句话，也可以用“$E(x), \quad x \in X$”表示.

- $\mathbb{N} := \{1, 2, \ldots\}$ 自然数集.

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ 分别是整数集、有理数集、实数集、复数集.

- $\mathbb{Z}_{\geq 0} := \{0, 1, 2, \ldots\}$ 非负整数集.

- $\mathbb{F}$ 表示实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$.

- $\mathbb{F}^n$ 表示域 $\mathbb{F}$ 上的 n 维列向量全体.

- $\mathbb{F}^{m \times n}$ 表示域 $\mathbb{F}$ 上 $m \times n$ 矩阵全体. 那么 $A = (a_{ij})_{m \times n}$ (其中 $a_{ij} \in \mathbb{F}$) 就是 $\mathbb{F}^{m \times n}$ 中的一个矩阵.

- $M_n(\mathbb{F})$ 表示 $\mathbb{F}^{n \times n}$，$M_{m,n}(\mathbb{F})$ 表示 $\mathbb{F}^{m \times n}$. 如果背景域 $\mathbb{F}$ 在上下文中是清楚的，可以省略的写成 M_n 以及 $M_{m,n}$.

- $\mathbb{F}[t]$ 表示的是以 t 为不定元的系数属于 $\mathbb{F}$ 的一元多项式全体构成的集合.

第一章　线性空间知识

1.1　向量空间

线性空间的定义及例子

定义 1.1. 本书中域 $\mathbb{F}$ 指的是实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$. 通常情况下，我们会直接在复数域 $\mathbb{C}$ 下工作. 关于“域”的严格定义需要参考近世代数/抽象代数的相关知识.

定义 1.2. (向量空间/线性空间) 设V 是一个非空集合，域 $\mathbb{F}$ 为实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$. 集合 V 上赋予一种运算，称为加法，$+: V\times V \to V, (x,y)\mapsto x+y$. 集合 V 与域 $\mathbb{F}$ 之间赋予了一种运算，称为数乘，$\bullet: \mathbb{F}\times V \to V, (a,x)\mapsto ax$.

- 加法运算满足：
  1. 结合律：$\forall x,y,z\in V$, 有$(x+y)+z = x+(y+z)$.
  2. 交换律：$\forall x,y\in V$, 有$x+y=y+x$.
  3. 有零元：$\exists\theta\in V, \forall x\in V$, 有$\theta+x=x+\theta=x$.
     (可以证明上述 θ 是唯一的，通常将 θ 记为 0.)
  4. 有负元：$\forall x\in V, \exists y\in V$ 使得 $x+y=y+x=0$.
     (可以证明 y 是唯一的，记为 $-x$.)
- 数乘运算满足：$\forall a,b\in\mathbb{F}, \forall x\in V$ 有
  5. “结合律” $(ab)x = a(bx)$.
  6. “分配律” $a(x+y) = ax+ay$
  7. “分配律” $(a+b)x = ax+bx$
  8. “酉性”：$1x = x$.

$(V, \mathbb{F}, +, \cdot)$ 一起考虑，称 V 为 $\mathbb{F}$ 上的一个线性空间，V 中的元素称为向量. 当 $\mathbb{F} = \mathbb{R}$ 时，称为实线性空间. 当 $\mathbb{F} = \mathbb{C}$ 时，称为复线性空间.

例 1.1. $\mathbb{F}^n$ 是 $\mathbb{F}$ 上的线性空间，其中 $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$,
$a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}$.

例 1.2. $\mathbb{F}^{m\times n}$ 按照矩阵的加法，按照数与矩阵的乘法，构成 $\mathbb{F}$ 上的一个线性空间. $\mathbb{R}^{m\times n}$ 称为 $m \times n$ 实矩阵空间. $\mathbb{C}^{m\times n}$ 称为 $m \times n$ 复矩阵空间.

例 1.3. $\mathbb{F}[x]$ 表示以 x 为不定元的系数在域 $\mathbb{F}$ 上多项式全体，按照多项式的加法，数与多项式的乘法，构成线性空间.

例 1.4. $\mathbb{F}[x]_n = \{f(x) \mid f(x) = a_0 + a_1 x + \cdots + a_n x^n, a_i \in \mathbb{F}\}$ 按照多项式的加法，数与多项式的乘法，构成线性空间.

例 1.5. $\mathcal{C}[a, b]$ 表示闭区间 $[a, b]$ 上的所有实值连续函数的集合，按函数乘法，按数与函数的数量乘法，构成线性空间.

例 1.6. (重要例子) 设 $A \in \mathbb{F}^{m\times n}$，齐次线性方程组的解集：

$$W = \{x \in \mathbb{F}^n \mid Ax = 0\}$$

按照 $\mathbb{F}^n$ 中的加法和数量乘法，W 构成 $\mathbb{F}$ 上的线性空间，称 W 为齐次线性方程组 $Ax = 0$ 的解空间. 也称为矩阵 A 的零空间 **(Nullspace)** 或核空间 **(Kernel)**，记为 $\mathcal{N}(A)$ 或者 $\mathrm{Ker}(A)$.

例 1.7. (重要例子) 设 $A \in \mathbb{F}^{m\times n}$，

$$V = \{y \in \mathbb{F}^m \mid y = Ax, x \in \mathbb{F}^n\}$$

按照 $\mathbb{F}^m$ 中的加法和数量乘法，V 构成 $\mathbb{F}$ 上的线性空间. V 称为 A 的列空间或像空间，记为 $\mathcal{R}(A), \mathrm{Im}(A)$.

线性空间有以下基本性质:

1. 零元唯一，负元唯一.
2. $a0 = 0$, $0x = 0$, $(-1)x = -x$.
3. 如果 $ax = 0$，那么或者 $a = 0$ 或者 $x = 0$.

子空间

定义 1.3. (子空间) 域 $\mathbb{F}$ 上的一个线性空间 V 的一个非空子集 W，如果 W 对于 V 的两种运算也构成域 $\mathbb{F}$ 上的线性空间，则称 W 是 V 的一个线性子空间.

定理 1.1. *设 V 为域 $\mathbb{F}$ 上的一个线性空间，W 是 V 的非空子集. 那么 W 能成为子空间当且仅当 W 关于 V 的两种运算是封闭的，即*

1. $\forall x, y \in W$，有 $x + y \in W$.

2. $\forall x \in W, \forall a \in \mathbb{F}$，有 $ax \in W$.

例 1.8. 线性空间 V 中，由单个零向量构成的子集是一个线性子空间，叫做零子空间. 线性空间 V 本身也是 V 的一个子空间. 这两个子空间通常称为 V 的平凡子空间，其它子空间叫做非平凡子空间.

例 1.9. 子空间的交是子空间. 假设 W_j $(j \in J)$ 都是线性空间 V 的子空间，那么 $\cap_{j\in J} W_j$ 也是 V 的子空间.

定义 1.4. 设 V 是域 $\mathbb{F}$ 上的线性空间，S 是 V 的子集合. 由 S张成/生成的子空间，记号为 $\operatorname{span}(S)$，是指 V 中包含集合 S 的所有子空间的交. 那么 $\operatorname{span}(S)$ 恰好是包含 S 的 V 的最小子空间. 如果 $S = \emptyset$,那么 $\operatorname{span}(S) = \{0\}$. 如果 $S \neq \emptyset$，那么

$$\operatorname{span}(S) = \{a_1x_1 + \cdots + a_kx_k \,|\, x_1, \ldots, x_k \in S, a_1, \ldots, a_k \in \mathbb{F}, k = 1, 2, \ldots\}.$$

最后，我们说 V 由 S 张成，如果 $\operatorname{span}(S) = V$，此时也说 S 是 V 的生成集.

例 1.10. 设 $A \in \mathbb{R}^{m\times n}$，那么 A 的所有列向量生成的空间恰好就是 A 的像空间 $\mathcal{R}(A)$，也叫做 A 的列向量空间$\operatorname{col}(A)$.

定义 1.5. 设 V 是域 $\mathbb{F}$ 上的线性空间，k 为正整数，$a_1, \ldots, a_k \in \mathbb{F}$，而 $v_1, \ldots, v_k \in V$，那么表达式 $a_1v_1 + a_2v_2 + \cdots + a_kv_k$ 称为 $v_1, \ldots, v_k$ 的一个线性组合. 因此由非空子集 S 张成的子空间 $\operatorname{span}(S)$ 中的每个向量恰好就是

S 中有限多个向量构成的线性组合. 我们说线性组合 $a_1v_1+a_2v_2+\cdots+a_kv_k$ 是平凡的，如果 $a_1=a_2=\cdots=a_k=0$；否则称为非平凡的.

定义 1.6. 设 V 是域 $\mathbb{F}$ 上的线性空间，V_1, V_2 是 V 的子空间.

$$V_1+V_2:=\{x+y\,|\,x\in V_1, y\in V_2\}$$

称为 V_1, V_2 的和空间. 显然有 $V_1+V_2=\mathrm{span}(V_1\cup V_2)$. 如果和空间 V_1+V_2 每个元素 z 的分解式

$$z=x+y\,(x\in V_1, y\in V_2)$$

是唯一的，那么称和 V_1+V_2 是直和，此时记为 $V_1\oplus V_2$. 进一步，如果 $V=V_1\oplus V_2$，称 V_1 是 V_2 在 V 中的一个直和补空间，称 V_2 是 V_1 在 V 中的一个直和补空间，简单称 V_1, V_2 为在 V 中的直和互补空间.

类似地可以给出有限个子空间的和 $\sum_{i=1}^k V_i$ 的定义，以及和 $\sum_{i=1}^k V_i$ 为直和 $\oplus_{i=1}^k V_i$ 的定义.

注 1.1. 注意！“子空间的交”与“集合之间的交”是一致的，但是“子空间的和”与“集合的并”的概念是不一致的.

定理 1.2. *设 V 是域 $\mathbb{F}$ 上的线性空间，V_1, V_2 是 V 的子空间. 那么以下等价：*

1. *V_1+V_2 是直和.*
2. *零表示法唯一：若 $0=v_1+v_2, v_1\in V_2, v_2\in V_2$，则 $v_1=0, v_2=0$.*
3. *$V_1\cap V_2=\{0\}$.*

定义 1.7. 设 V 是域 $\mathbb{F}$ 上的线性空间，$V_1, V_2, \ldots, V_s$ 是 V 的子空间. 那么它们的和空间定义为 $\sum_{k=1}^s V_k:=\{\sum_{k=1}^s v_k \mid v_k\in V_k, k=1,\ldots,s\}$, 它是 V 的子空间. 和空间 $\sum_{k=1}^s V_k$ 称为直和如果 $\sum_{k=1}^s V_k$ 中的每个向量的分解式是唯一的：若

$$\begin{aligned} v&=v_1+v_2+\cdots+v_s\,(v_k\in V_k, k=1,\ldots,s)\\ &=u_1+u_2+\cdots+u_s\,(u_k\in V_k, k=1,\ldots,s) \end{aligned}$$

则 $v_1=u_1, v_2=u_2,\ldots,v_s=u_s$. 当 $\sum_{k=1}^s V_k$ 是直和时，我们用记号 $\oplus_{k=1}^s V_k$ 表示. 容易证明 $\sum_{k=1}^s V_k=\oplus_{k=1}^s V_k$ 当且仅当零表示法是唯一的，即，如果 $0=v_1+v_2+\cdots+v_s\,(v_k\in V_k, k=1,\ldots,s)$，那么 $v_1=v_2=\cdots=v_s=0$.

线性相关性

定义 1.8. 设 V 是域 $\mathbb{F}$ 上的线性空间，$v_1, v_2, \ldots, v_k \in V$. 如果存在不全为零的数 $a_1, a_2, \ldots, a_k \in \mathbb{F}$ 使得 $a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$，就称 $v_1, v_2, \ldots, v_k$ 是线性相关的. 如果只有平凡的线性组合才能是零向量，就称 $v_1, v_2, \ldots, v_k$ 是线性无关的.

设 $S \subseteq V$，说 S 是线性相关集，如果 S 中有有限个向量是线性相关的. 说 S 是线性无关集，如果 S 中任意有限个向量线性无关. 空集 $\emptyset$ 是线性无关集.

定义 1.9. 设 V 是域 $\mathbb{F}$ 上的线性空间，$x_1, x_2, \ldots, x_s \in V$，$y_1, y_2, \ldots, y_t \in V$ 称 $x_1, x_2, \ldots, x_s$ 能被 $y_1, y_2, \ldots, y_t$ 线性表示/线性表出，如果每个 x_k 都是 $y_1, y_2, \ldots, y_t$ 的线性组合. 称 $x_1, x_2, \ldots, x_s$ 和 $y_1, y_2, \ldots, y_t$ 是线性等价的，如果它们能相互线性表出.

更一般的，假设 $S \subseteq V, T \subseteq V$，我们说 S 可由 T 线性表示，如果 $S \subseteq \mathrm{span}(T)$；称 S 和 T 是线性等价的，如果 $\mathrm{span}(S) = \mathrm{span}(T)$.

常用结论：

1. 单个向量 x 线性相关当且仅当 $x = 0$. 两个以上的向量线性相关当且仅当其中有一个向量能被其余向量线性表出.

2. 设 $x_1, \ldots, x_r$ 线性无关，且 $x_1, \ldots, x_r$ 可以由 $y_1, \ldots, y_s$ 线性表出，那么 $r \leq t$.

3. 设 $x_1, \ldots, x_r$ 线性无关，且 $x_1, \ldots, x_r, y$ 线性相关，那么 y 可以由 $x_1, \ldots, x_r$ 线性表示且表法唯一.

例 1.11. 实多项式 $\mathbb{R}[t]$ 空间，$\{1, t, t^2, \ldots\}$ 构成线性无关集.

线性空间的基

定义 1.10. 设 V 是域 $\mathbb{F}$ 上的线性空间. 若 S 是 V 中的线性无关子集且 $\mathrm{span}(S) = V$，那么称 S 是 V 的基.

关于线性空间的基有以下性质.

定理 1.3. *设 V 是域 $\mathbb{F}$ 上的线性空间.*

1. S 是基当且仅当它是极大线性无关集. 子集 A 是极大线性无关集是指满足以下性质的集合:

   - A 是线性无关的.
   - 如果 $A \subseteq B \subseteq V$ 且 B 是线性无关集，则必有 $B = A$.

2. S 是基当且仅当它是极小生成集. 子集 A 是极小生成集是指满足以下性质的集合:

   - A 是生成集，即 $span(A) = V$.
   - 如果 $B \subseteq A$ 且 $span(B) = V$，则必有 $B = A$.

3. (基存在定理) 线性空间必存在基.

4. 如果 V 存在一个基由 n (这里 n 是一个非负整数) 个元素构成，那么 V 的任意一个基都是由 n 个元素构成. 此时我们说 V 是有限维的且维数是 n，记为 $\dim V = n$，简称 V 是 n 维线性空间. 如果 V 不是有限维的，就称 V 是无限维的.

5. (基扩充定理) 假设 A 是 V 的一个线性无关集，那么存在基 B 使得 $A \subseteq B$. 也就是说从一个线性无关集出发，可以将其扩充为 V 的一个基，特别的，由于空集是线性无关集，从空集出发扩充成 V 的一个基，这就得到了基存在定理.

6. $\dim V = 0$ 当且仅当 $V = \{0\}$. 此时 V 只有一个基：$\emptyset$.

7. 如果 $\dim V = n$，设 W 是 V 的任意一个子空间，则 W 是有限维的且 $\dim W \leq \dim V$.

注 1.2. 对于有限维空间 U，我们说 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 U 的一个基，通常是指有序基 $(\varepsilon_1, \ldots, \varepsilon_n)$ 也就是说与基中向量的顺序是有关的.

1.2 线性映射

本节中 $U, V, W, \ldots$ 都是给定域 $\mathbb{F}$ 上的线性空间.

概念与基本定理

定义 1.11. 映射 $T: U \to V$ 称为线性映射，如果它满足：

1. 保持加法：$T(x+y) = T(x) + T(y)$ 对任意 $x, y \in U$ 成立.
2. 保持数乘：$T(kx) = kT(x)$ 对任意 $k \in \mathbb{F}, x \in U$ 成立.

如果需要明确指明背景域 $\mathbb{F}$ 时，可以说 T 是 $\mathbb{F}-$ 线性的. 从线性空间 U 到自己的线性映射 $S: U \to U$ 通常也称做 U 上的线性变换.

例 1.12. 设 $A \in \mathbb{F}^{m\times n}$，那么可以自然地诱导出一个线性映射 $\varphi_A: \mathbb{F}^n \to \mathbb{F}^m, x \mapsto Ax$.

例 1.13. 考虑积分运算 $S: \mathcal{C}[a,b] \to \mathcal{C}[a,b], f \mapsto S(f)(x) = \int_a^x f(t)dt$ 是实线性映射，这里 $\mathcal{C}[a,b]$ 是指闭区间 $[a,b]$ 上的连续实值函数全体构成的实线性空间.

例 1.14. 矩阵的共轭转置运算 $T: \mathbb{C}^{m\times n} \to \mathbb{C}^{n\times m}, A \mapsto A^H$ 不是 $\mathbb{C}-$ 线性的，但它是 $\mathbb{R}-$ 线性的. 也就是说 $\mathbb{C}^{m\times n}, \mathbb{C}^{n\times m}$ 看成是 $\mathbb{R}$ 线性空间时，T 是线性的，但它们做为 $\mathbb{C}-$ 线性空间时，T 不是线性的.

定理 1.4. *(简单性质) 设 $T: U \to V$ 是线性映射，那么*

1. *$T(0) = 0$.*
2. *$T(a_1x_1 + \cdots + a_mx_m) = a_1T(x_1) + \cdots + a_mT(x_m)$.*
3. *设 $T(x_1), \ldots, T(x_m)$ 是 V 中的线性无关向量，那么 $x_1, x_2, \ldots, x_m$ 是 U 中的线性无关向量.*

注 1.3. 如果 x 是 $x_1, \ldots, x_m \in U$ 的线性组合,那么 $T(x)$ 就是 $T(x_1), \ldots, T(x_m)$ 的线性组合，从而如果知道 U 的一个基 $\varepsilon_1, \ldots, \varepsilon_n$(假设 U 是 n 维的) 在线性映射 T 下的像，那么 U 中任意一个向量 x 的像 $T(x)$ 也就确定了. 换句话说 n 维线性空间 U 到令一个线性空间 V 的线性映射 T 完全由 U 的一个给定基下的像所决定.

定义 1.12. 设 $T: U \to V$ 是线性映射，如果 T 是双射，那么其逆映射 T^{-1} 也是线性的，此时我们称 T 是从 U 到 V 的线性同构，称 U 和 V 是同构的.

定义 1.13. 假设 U 是有限维空间，$\varepsilon_1, \ldots, \varepsilon_n$ 是 U 的一个基，那么对于每个取定的 $u \in U$，存在唯一的一组数 $a_1, \ldots, a_n$ 使得 $u = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$，我们称 $(a_1, \ldots, a_n)^T \in \mathbb{F}^n$ 为 u 在基 $\varepsilon_1, \ldots, \varepsilon_n$ 下的坐标向量.

将 U 中的每个向量对应到其坐标向量的映射就是一个从 U 到 $\mathbb{F}^n$ 的同构，所以 n 维线性空间总是与 $\mathbb{F}^n$ 是同构的.

定义 1.14. 设 $T: U \to V$ 是线性映射，记号 $\mathrm{Ker}(T), \mathcal{N}(T)$ 定义为 U 的子空间 $\mathrm{Ker(T)} := \{x \in U \mid T(x) = 0\}$，称为 T 的核空间. 记号 $\mathrm{Im}(T), \mathrm{Range}(T), \mathcal{R}(T)$ 表示 T 的像空间，它是 V 的子空间，定义为 $\mathrm{Im}(T) := \{Tx \mid x \in U\}$.

在 U, V 都是有限维时，$\mathrm{rank}(T) := \dim \mathcal{R}(T)$ 称为 T 的秩；而 $\mathrm{nullity}(T) := \dim \mathcal{N}(T)$ 称为 T 的零度.

定理 1.5. （*rank-nullity* 定理）设 $T: U \to V$ 为线性映射，且 U, V 是有限维的，那么

$$\dim \mathcal{R}(T) + \dim \mathcal{N}(T) = \dim U.$$

证明. 取 $\mathcal{N}(T)$ 的一个基 $x_1, \ldots, x_k$ 将其扩充为 U 的一个基 $x_1, \ldots, x_k, x_{k+1}, \ldots, x_n$，那么 $T(x_1), \ldots, T(x_n)$ 张成 $\mathcal{R}(T)$，由于 $T(x_1) = \cdots = T(x_k) = 0$，于是 $T(x_{k+1}), \ldots, T(x_n)$ 已经张成 $\mathcal{R}(T)$. 下面我们来证明 $T(x_{k+1}), \ldots, T(x_n)$ 是线性无关的，从而 $T(x_{k+1}), \ldots, T(x_n)$ 是 $\mathcal{R}(T)$ 的一个基，定理同时得到了证明.

假设 $a_{k+1}, \ldots, a_n \in \mathbb{F}$ 使得 $a_{k+1}T(x_{k+1}) + \cdots + a_nT(x_n) = 0$，那么 $T(a_{k+1}x_{k+1} + \cdots + a_nx_n) = 0$. 所以 $a_{k+1}x_{k+1} + \cdots + a_nx_n \in \mathcal{N}(T)$，又因为 $x_1, \ldots, x_k$ 是 $\mathcal{N}(T)$ 的一个基，从而存在 $a_1, \ldots, a_k \in \mathbb{F}$ 使得 $a_{k+1}x_{k+1} + \cdots + a_nx_n = a_1x_1 + \cdots + a_kx_k$，但是 $x_1, \ldots, x_k, x_{k+1}, \ldots, x_n$ 是线性无关的，因此所有这些系数都是零，所以 $T(x_{k+1}), \ldots, T(x_n)$ 是线性无关的. □

例 1.15. n 元齐次线性方程组 $Ax = 0$ 的解空间的维数是 $n - \mathrm{rank}(A)$ 这一结论是上述定理的特例.

例 1.16. (Frobenius 秩不等式)

$$\mathrm{rank}(AB) + \mathrm{rank}(BC) \le \mathrm{rank}(ABC) + \mathrm{rank}(B)$$

证明. 若 V, W 是域 $\mathbb{F}$ 上的有限维空间，U 为 V 的子空间，$\varphi: V \to W$ 是线性映射，则 $\dim \mathrm{Ker}\varphi|_U \le \dim \mathrm{Ker}\varphi$. 现在取 $U = \mathrm{Im}(BC), V = \mathrm{Im}(B), \varphi = A|_{\mathrm{Im}(B)}$，那么 $\dim \mathrm{Ker}A|_{\mathrm{Im}(BC)} \le \dim \mathrm{Ker}A|_{\mathrm{Im}(B)}$，于是 $\mathrm{rank}(BC) - \mathrm{rank}(ABC) \le \mathrm{rank}(B) - \mathrm{rank}(AB)$. 可以利量用下面的交换图简单捋顺思路：

$$\begin{tikzcd}
\mathrm{Im}(B) \arrow[r, "A|_{\mathrm{Im}(B)}"] & \mathrm{Im}(AB) \\
\mathrm{Im}(BC) \arrow[u, "嵌入"'] \arrow[r, "A|_{\mathrm{Im}(BC)}"'] & \mathrm{Im}(ABC) \arrow[u, "嵌入"']
\end{tikzcd}$$

□

例 1.17. (Sylvester 不等式)

$$\operatorname{rank}(A)+\operatorname{rank}(B)\leq n+\operatorname{rank}(AB)$$

其中 $A_{m\times n},B_{n\times m}$ 为矩阵.

证明. 在 Frobenius 不等式中让 B 等于单位阵即可. □

例 1.18. (维数公式) 设 U,V 是 W 的子空间，W 是有限维的，那么 $\dim U+\dim V=\dim(U+V)+\dim(U\cap V)$.

证明. $U\times V=\{(u,v)\,|\,u\in U,v\in V\}$ 可以自然的成为线性空间，加法运算按对应分量做加，数乘运算是按数乘分量得到. 设 $u_1,\ldots,u_n\in U$ 构成 U 的一个基，$v_1,\ldots,v_m$ 构成 V 的一个基，那么 $(u_1,0),\ldots,(u_n,0),(0,v_1),\ldots,(0,v_m)$ 构成 $U\times V$ 的一个基本，于是 $\dim(U\times V)=\dim U+\dim V$.

现在考虑 $T:U\times V\to W,(u,v)\mapsto u-v$. 那么 T 是线性映射，且 $\operatorname{Im}(T)=U+V,\operatorname{Ker}(T)=\{(x,x)\in U\times V\,|\,x\in U\cap V\}$. 显然 $\dim\operatorname{Ker}(T)=\dim(U\cap V)$ (因为 $\operatorname{Ker}(T)$ 和 $U\cap V$ 是同构的). 所以根据 $\dim(U\times V)=\dim\operatorname{Im}(T)+\dim\operatorname{Ker}(T)$ 可以得到

$$\dim U+\dim V=\dim(U+V)+\dim(U\cap V).$$

□

线性映射的运算

我们考虑从线性空间 U 到线性空间 V 的线性映射全体构成的集合 $\mathcal{L}(U,V)$. 对于 $T_1,T_2\in\mathcal{L}(U,V)$，定义 $T_1+T_2:U\to V,x\mapsto T_1(x)+T_2(x)$，那么 T_1+T_2 也是线性的，即 $T_1+T_2\in\mathcal{L}(U,V)$. 对于数 $a\in\mathbb{F},T\in\mathcal{L}(U,V)$，定义 $aT:U\to V,x\mapsto a(T(x))$，那么 aT 也是线性的. 集合 $\mathcal{L}(U,V)$ 按照上述两种运算能构成 $\mathbb{F}$ 上的线性空间.

线性映射的矩阵表示

假设 U 是 n 维线性空间，一个给定基为 $\varepsilon:\varepsilon_1,\ldots,\varepsilon_n$；假设 V 是 m 维线性空间，一个给定的基是 $\eta:\eta_1,\ldots,\eta_m$；假设 $T:U\to V$ 是线性映射.

那么存在唯一的矩阵 $A \in \mathbb{F}^{m\times n}$ 使得下图交换，其中图里上下水平的箭头映射分别表示在给定基下取坐标的映射：

$$\begin{tikzcd} U \arrow[r, "x\mapsto [x]_\varepsilon"'] \arrow[d, "T"] & \mathbb{F}^n \arrow[d, dotted, "A"] \\ V \arrow[r, "y\mapsto [y]_\eta"'] & \mathbb{F}^m \end{tikzcd}$$

这个矩阵 A 称为 T 在基 ε, η 下的**表示矩阵**，矩阵 A 的第 k 列恰好是 $T(\varepsilon_k)$ 在 $\eta_1, \ldots, \eta_m$ 下的坐标.

注 1.4. 矩阵 A 可以按照下面等式得到：

$$\begin{cases} T(\varepsilon_1) = a_{11}\eta_1 + a_{21}\eta_2 + \cdots + a_{m1}\eta_m \\ T(\varepsilon_2) = a_{12}\eta_1 + a_{22}\eta_2 + \cdots + a_{m2}\eta_m \\ \vdots \\ T(\varepsilon_n) = a_{1n}\eta_1 + a_{2n}\eta_2 + \cdots + a_{mn}\eta_m \end{cases} \tag{1.1}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

利用形式上的矩阵乘法，我们可以将上述事情表示成：

$$T(\varepsilon_1, \ldots, \varepsilon_n) = (\eta_1, \ldots, \eta_m)A.$$

另外一方面，假设我们不知道矩阵的乘法，假设我们运气好一上来就凑巧从1.1式子中提取出系数阵 A(即现在这种“转置”了之后的形式)，那么我们可以通过上面的交换图中虚线箭头的映射 $\mathbb{F}^n \to \mathbb{F}^m$ 来定义矩阵 A 与列向量的乘法，如果是这样的话，那么对于 $u \in U$，则 Tu 的坐标恰好是矩阵 A 与 u 的坐标之乘积.

定理 1.6. 假设线性空间 U, V 是有限维的，$\dim U = n, \dim V = m$，那么线性空间 $\mathcal{L}(U,V) \cong \mathbb{F}^{m\times n}$，因此 $\dim \mathcal{L}(U,V) = \dim U \cdot \dim V$.

证明. 取定 U 和 V 的基之后，每个线性映射对应于它的表示矩阵矩阵是从 $\mathcal{L}(U,V)$ 到 $\mathbb{F}^{m\times n}$ 的线性同构. □

如果我们有了矩阵与列向量的乘法，那么矩阵乘法的定义就应该对应于线性映射的复合.

定理 1.7. 假设 $U \xrightarrow{T} V \xrightarrow{S} W$ 是有限维线性空间中的线性映射，在 U, V, W 中取定基后 (设 U 取的基为 $u_1, u_2, \ldots$)，我们有下面的交换图

$$\begin{tikzcd} U \arrow[r, "T"] \arrow[d, "{[\cdot]}"] & V \arrow[r, "S"] \arrow[d, "{[\cdot]}"] & W \arrow[d, "{[\cdot]}"] \\ \mathbb{F}^n \arrow[r, "A"] & \mathbb{F}^m \arrow[r, "B"] & \mathbb{F}^p \end{tikzcd}$$

其中竖着的箭头的映射 $[\cdot]$ 表示取坐标映射，A, B 分别为对应的表示矩阵. 那么复合映射 $S \circ T$ 对应的矩阵 C 的第 k 列 C_k 恰好是 B 乘以 A 的第 k 列向量 A_k 得到，这就完全刻画了 C，而矩阵 B 与矩阵 A 的乘积 BA 就定义为 C. 由于映射有结合律，从而矩阵乘法也有结合律.

证明. 做图追踪，有 $C_k = C[u_k] = [(S \circ T)(u_k)] = [S(Tu_k)] = B[Tu_k] = B(A[u_k]) = BA_k$. □

定义 1.15. 设 n 维线性空间 U 给定了两个基 $\varepsilon : \varepsilon_1, \varepsilon_2, \ldots$ 和 $\eta : \eta_1, \eta_2 \ldots$，存在矩阵 P 使得下图交换：

$$\begin{tikzcd} U \arrow[r, "I"] \arrow[d, "{[\cdot]_\eta}"] & U \arrow[d, "{[\cdot]_\varepsilon}"] \\ \mathbb{F}^n \arrow[r, dotted, "\exists P"] & \mathbb{F}^n \end{tikzcd}$$

其中 I 是恒等映射. 矩阵 P 称为从基 ε 到基 η 的过渡阵，P 的第 k 列恰好是 η_k 在基 ε 下的坐标，显然 P 是可逆矩阵. 利用形式上的矩阵乘法，上述事实可以表示为

$$\eta = \varepsilon P$$

(这是从基ε到基η的过渡阵的定义方式，而不是从 η 到 ε 的). 但是这样的话取坐标映射的关系却是：

$$[\cdot]_\eta = P^{-1}[\cdot]_\varepsilon$$

.

定理 1.8. 设 $T : U \to V$ 为有限维线性空间之间的线性映射,设 $\varepsilon : \varepsilon_1, \ldots, \varepsilon_n$ 以及 $\varepsilon' : \varepsilon'_1, \ldots, \varepsilon'_n$ 为 U 的两个给定基，设 $\eta : \eta_1, \ldots, \eta_m$ 和 $\eta' : \eta'_1, \ldots, \eta'_m$

为 V 的两个给定基，设 T 在 ε,η 下的矩阵是 A 在 ε',η' 下的矩阵是 B，基 ε 到基 ε' 的过渡阵是 P，基 η 到基 η' 的过渡阵是 Q，那么有以下交换图：

$$\begin{tikzcd}
U \arrow[rr, "T"] \arrow[dr, "{[\cdot]_{\varepsilon'}}"] \arrow[dd, "{[\cdot]_{\varepsilon}}"'] & & V \arrow[dr, "{[\cdot]_{\eta'}}"] \arrow[dd, dotted, "{[\cdot]_{\eta}}"'] & \\
& \mathbb{F}^n \arrow[rr, "B"'] \arrow[dl, "P"] & & \mathbb{F}^m \arrow[dl, "Q"] \\
\mathbb{F}^n \arrow[rr, "A"] & & \mathbb{F}^m &
\end{tikzcd}$$

矩阵 B 与 A 的关系为：

$$B = Q^{-1}AP.$$

证明. 由于 $AP[\cdot]_{\varepsilon'} = A[\cdot]_{\varepsilon} = [\cdot]_{\eta}T = Q[\cdot]_{\eta'}T = QB[\cdot]_{\varepsilon'}$，而 $[\cdot]_{\varepsilon'}$ 是同构，于是 $AP = QB$，所以 $B = Q^{-1}AP$. □

定理 1.9. 设 $T : U \to U$ 为有限维线性空间上的线性变换，设 $\varepsilon : \varepsilon_1, \ldots, \varepsilon_n$ 以及 $\varepsilon' : \varepsilon'_1, \ldots, \varepsilon'_n$ 为 U 的两个给定基，设 T 在 ε 下的矩阵是 A，在 ε' 下的矩阵是 B，基 ε 到基 ε' 的过渡阵是 P，那么有以下交换图：

$$\begin{tikzcd}
U \arrow[rr, "T"] \arrow[dr, "{[\cdot]_{\varepsilon'}}"] \arrow[dd, "{[\cdot]_{\varepsilon}}"'] & & U \arrow[dr, "{[\cdot]_{\varepsilon'}}"] \arrow[dd, dotted, "{[\cdot]_{\varepsilon}}"'] & \\
& \mathbb{F}^n \arrow[rr, "B"'] \arrow[dl, "P"] & & \mathbb{F}^n \arrow[dl, "P"] \\
\mathbb{F}^n \arrow[rr, "A"] & & \mathbb{F}^n &
\end{tikzcd}$$

矩阵 B 与 A 的关系为：

$$B = P^{-1}AP.$$

证明. 该定理是定理1.8的特殊版本. □

1.3 内积空间

定义和例子

定义 1.16. 设 V 是 $\mathbb{R}$ 上的线性空间，二元函数 $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ 称为 V 上的内积，如果它满足：对于任意 $x, y, z \in V$，任意的 $a \in \mathbb{R}$ 有

1. 正定性：$\langle x,x\rangle \geq 0$；等号成立当且仅当 $x=0$.

2. 对称性：$\langle x,y\rangle = \langle y,x\rangle$

3. 关于第一变元线性：

   - $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$
   - $\langle ax,y\rangle = a\langle x,y\rangle$

此时称 $(V,\langle\cdot,\cdot\rangle)$ 是实内积空间，有限维实内积空间通常也称为**欧式空间**.

定义 1.17. 设 V 是 $\mathbb{C}$ 上的线性空间，二元函数 $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{C}$ 称为 V 上的内积，如果它满足：对于任意 $x,y,z\in V$，任意的 $a\in\mathbb{C}$ 有

1. 正定性：$\langle x,x\rangle \geq 0$；等号成立当且仅当 $x=0$.

2. 共轭对称性：$\langle x,y\rangle = \overline{\langle y,x\rangle}$

3. 关于第一变元线性：

   - $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$
   - $\langle ax,y\rangle = a\langle x,y\rangle$

此时称 $(V,\langle\cdot,\cdot\rangle)$ 是复内积空间，有限维复内积空间通常也称为**酉空间**.

注 1.5. 内积的符号通常也直接用 $(\cdot,\cdot)$ 表示. 在物理学中，有的作者在内积的定义中要求关于第二变元是线性的而第一变元是共轭线性的. 内积的其它记号有 $(\cdot|\cdot),\langle\cdot|\cdot\rangle$ 等等.

注 1.6. 由于复数域不是有序域，当写一个复数 $w\geq 0$ 时，指的是 w 此时是实数且 $w\geq 0$. 所以在复内积空间中“正定性”：$\langle x,x\rangle$ 是实数且 $\langle x,x\rangle \geq 0$. 当然，如果用共轭对称性也是能说明 $\langle x,x\rangle$ 是实数.

内积的一些简单的性质：

1. 实内积的情形下，

   (a) $\langle 0,x\rangle = \langle x,0\rangle = 0$

   (b) 关于第二变元也是线性的.

2. 复内积的情形下，

   (a) $\langle 0,x\rangle = \langle x,0\rangle = 0$

(b) 关于第二变元是共轭线性的：
- $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$
- $\langle x, ay\rangle = \overline{a}\,\langle x, y\rangle$

3. 内积空间 U 的线性子空间 V 继承空间 U 中的内积自动也成为内积空间.

例 1.19. n 维实空间 $\mathbb{R}^n$ 中,对于任意的 $x=(a_1,\ldots,a_n)^T, y=(b_1,\ldots,b_n)^T$，规定 $\langle x,y\rangle = y^Tx=\sum_{k=1}^n a_kb_k$，那么 $\langle\cdot,\cdot\rangle$ 是 $\mathbb{R}^n$ 上的内积. 这个内积我们称为 $\mathbb{R}^n$ 上的标准/通常内积.

例 1.20. n 维实空间 $\mathbb{R}^n$ 中,对于任意的 $x=(a_1,\ldots,a_n)^T, y=(b_1,\ldots,b_n)^T$，规定 $\langle x,y\rangle=\sum_{k=1}^n kx_ky_k$. 那么 $\langle\cdot,\cdot\rangle$ 也是 $\mathbb{R}^n$ 上的内积.

例 1.21. $\mathcal{C}[0,1]$ 上，$\langle f,g\rangle := \int_0^1 f(x)g(x)dx$，也是内积.

例 1.22. 实空间 $\mathbb{R}^{m\times n}$ 中，$\langle A,B\rangle := \mathrm{tr}(B^TA)$ 是内积.

例 1.23. 设 $x=(a_1,\ldots,a_n)^T\in\mathbb{C}^n, y=(b_1,\ldots,b_n)^T\in\mathbb{C}^n$，令 $\langle x,y\rangle := y^Hx=\sum_{k=1}^n a_k\overline{b_k}$. 那么 $\langle\cdot,\cdot\rangle$ 是 $\mathbb{C}^n$ 上的内积，称为标准内积/通常内积.

例 1.24. 复空间 $\mathbb{C}^{m\times n}$ 中，$\langle A,B\rangle := \mathrm{tr}(B^HA)$ 是内积.

定义 1.18. 设 V 是内积空间，$x\in V$，则 x 的长度/由内积导出的范数，$\|x\|$，定义为 $\sqrt{\langle x,x\rangle}$. 由内积导出的范数也常用记号 $\|\cdot\|_2$ 表示.

显然 $\|x\|=0 \iff x=0$. 如果 $\|x\|=1$，则称 x 是单位向量.

内积的基本性质

定义 1.19. 设 V 是内积空间，$x,y\in V$，如果 $\langle x,y\rangle=0$，则称 x 正交/垂直于 y，记为 $x\perp y$. 显然 $x\perp y \iff y\perp x$.

定理 1.10. *(勾股定理)* 假设 $x\perp y$，那么 $\|x+y\|^2=\|x\|^2+\|y\|^2$. 进一步，如果 $v_1,\ldots,v_m$ 两两垂直，那么 $\|\sum_{k=1}^m v_k\|^2=\sum_{k=1}^m\|v_k\|^2$.

定理 1.11. *(Cauchy-Bunyakovsky-Schwarz 不等式)* 内积空间中，

$$|\langle x,y\rangle| \le \|x\|\cdot\|y\|$$

等号成立当且仅当 x 和 y 是线性相关的.

证明. 若 $y=0$，不等式显然成立.

若 $y \neq 0$，对 x 做正交分解：$x = u + v$，其中 $u = \frac{\langle x,y\rangle}{\langle y,y\rangle}y$ 而 $v = x - \frac{\langle x,y\rangle}{\langle y,y\rangle}y$. 那么 $v \perp y, v \perp u$. 于是根据勾股定理得，$\|x\| \geq \|u\|$，所以有 $|\langle x,y\rangle| \leq \|x\| \cdot \|y\|$.

在不等式取等号的假设下，如果 $y=0$ 自然有 x 与 y 是线性相关的；如果 $y \neq 0$，那么 $\|x\| = \|u\|$，此时 $v=0$，所以 $x=u$ 与 y 线性相关. 反之，如果 x 与 y 是线性相关的，此时显然有 $|\langle x,y\rangle| = \|x\| \cdot \|y\|$. □

定理 1.12. *内积导出的范数有以下性质：*

1. *正定性：$\forall x \in V$ 有 $\|x\| \geq 0$ 等号成立当且仅当 $x=0$.*
2. *齐次性：$\forall x \in V, \forall a \in \mathbb{F}$, 有 $\|ax\| = |a|\|x\|$.*
3. *三角不等式：$\forall x,y \in V$ 有 $\|x+y\| \leq \|x\| + \|y\|$.*

证明. 利用 Cauchy-Schwarz 不等式，可以得到三角不等式的证明. 其余两条性质是显然的. □

复内积中的常用结论

定理 1.13. *(简单但经常使用的结论) 设 $A \in \mathbb{C}^{m\times n}$，$x \in \mathbb{C}^n$，$y \in \mathbb{C}^m$，那么复内积 $\langle Ax,y\rangle = \langle x, A^H y\rangle$.*

证明. 直接根据内积的定义计算即可. □

定理 1.14. *设 $A \in \mathbb{C}^{m\times n}$，若复内积 $\langle Ax,y\rangle = 0$ 对任意 $x \in \mathbb{C}^n$ 和任意的 $y \in \mathbb{C}^m$ 成立，则 $A=0$.*

证明. 任取定 $x \in \mathbb{C}^n$，然后取 $y = Ax$，由 $\langle Ax,y\rangle = 0$ 得知 $Ax=0$. 那么 $Ax=0$ 对任意的 $x \in \mathbb{C}^n$ 成立，故 $A=0$. □

下面的定理仅对复数域的情形是成立的：

定理 1.15. *设 A 为 n 阶复方阵，$\langle Ax,x\rangle = 0$ 对所有 $x \in \mathbb{C}^n$ 成立，则 $A=0$.*

证明. 由条件，$\forall x,y \in \mathbb{C}^n$，$\forall k \in \mathbb{C}$，$\langle A(x+ky), x+ky\rangle = 0$，展开后再利用已知条件化简后可得：$\overline{k}\langle Ax,y\rangle + k\langle Ay,x\rangle = 0$. 等式 $\overline{k}\langle Ax,y\rangle + k\langle Ay,x\rangle = 0$ 中，固定 x 和 y，分别取 $k=1,i$，那么有 $\langle Ax,y\rangle = \langle Ay,x\rangle = 0$. 所以 $\langle Ax,y\rangle = 0$ 对任意的 $x,y \in \mathbb{C}^n$ 成立，于是 $A=0$. □

定义 1.20. 设 A 为 n 阶复方阵，如果 $A^H = A$，则称 A 是 Hermite 的，厄米的.

定理 1.16. *设 A 为 n 阶复方阵, 则 A 是 Hermite 的当且仅当对任意 $x \in \mathbb{C}^n$ 有 $\langle Ax, x\rangle \in \mathbb{R}$.*

证明. 显然，如果 A 是厄米的，那么 $\forall x \in \mathbb{C}^n$ 有 $\langle Ax, x\rangle \in \mathbb{R}$.

反之，假设 $\forall x \in \mathbb{C}^n$ 有 $\langle Ax, x\rangle \in \mathbb{R}$，那么 $\forall x \in \mathbb{C}^n$ 有 $\langle Ax, x\rangle = \langle x, Ax\rangle$. 于是 $\forall x \in \mathbb{C}^n$ 有 $\langle (A - A^H)x, x\rangle = 0$，故 $A^H = A$，矩阵 A 是 Hermite 的. □

酉阵

定义 1.21. 设向量组 $u_1, u_2, \ldots, u_m$ 是内积空间 U 中的不含有零向量的向量组，

- 若它们两两正交，则称之为**正交组**. 如果进一步，它们还构成基，则称为**正交基**.
- 若它们两两正交且每个都是单位向量，则称之为**标准/规范正交组**. 如果进一步，它们还构成基，则称为**标准正交基**，简称**标正基**.

显然，根据定义，内积空间 U 中的向量 $u_1, \ldots, u_m$ 构成标准交组当且仅当它们满足：$\langle u_j, u_k\rangle = \delta_{jk}, j, k = 1, 2, \ldots, m$，其中 $\delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$

符号 δ_{jk} 称为克罗内克符号，它是一个常用符号，例如 n 阶单位矩阵 I_n 可以表示为 $(\delta_{ij})_{n\times n}$

定理 1.17. *设 $u_1, u_2, \ldots, u_m$ 为内积空间 U 的一个正交组，那么它们是线性无关组.*

定义 1.22. 设 $Q \in \mathbb{R}^{n\times n}$，如果 $Q^TQ = I$，则称 Q 为实正交阵. 设 $U \in \mathbb{C}^{n\times n}$，如果 $U^HU = I$，则称 U 为酉阵.

定理 1.18. *(酉阵的刻画) 设 U 为 n 阶复方阵，则以下等价：*

1) U 为酉阵.

2) U 是可逆阵，且 $U^H = U^{-1}$.

3) $UU^H = I$.

4) U^H 是酉阵.
5) U 的 n 个列向量是单位正交向量组.
6) U 的 n 个行向量是单位正交向量组.
7) (保内积) $\forall x, y \in \mathbb{C}^n$，$\langle Ux, Uy\rangle = \langle x, y\rangle$.
8) (保长度) 对于 $\forall x \in \mathbb{C}^n$，$\|x\| = \|Ux\|$.

证明. 1)，2)，3)，4) 四者等价是显然的. 1)与5) 等价是显然的. 4)与6)等价是显然的.

1) $\Longrightarrow$ 7) $\Longrightarrow$ 8)是显然的.

8) $\Longrightarrow$ 1)：利用定理1.15. □

注 1.7. 实正交阵也有对应的版本，但是定理1.15对实内积的情形不再成立，此时需要一点点技巧可以直接证明“保长度”蕴含“保内积”，证明剩余部分是相对直接的.

例 1.25. 若 A 和 B 都是 n 阶酉阵，则 AB 也是酉阵.

例 1.26. 若 U 为酉阵，则 $|det(U)| = 1$.

Schmidt 正交化

下面介绍 Gram-Schmidt 正交化过程.

定理 1.19. *(Gram-Schmidt 正交化) 假设 $u_1, \ldots, u_m$ 是内积空间 U 中的一个线性无关组，那么可以逐步求出 $v_1, v_2, \ldots, v_m$ 使得它们构成标正组并且满足：*

$$span(u_1, \ldots, u_k) = span(v_1, \ldots, v_k), k = 1, 2, \ldots, m.$$

证明. 首先取 $v_1 = \frac{u_1}{\|u_1\|}$. 一般的，假设已经求出了 $v_1, \ldots, v_k$，它们是单位正交的且具有性质：$\mathrm{span}(u_1, \ldots, u_j) = \mathrm{span}(v_1, \ldots, v_j), j = 1, 2, \ldots, k$.

下求 v_{k+1}. 因为 $\mathrm{span}(u_1, \ldots, u_k) = \mathrm{span}(v_1, \ldots, v_k)$ 且 $u_1, \ldots, u_k, u_{k+1}$ 线性无关,所以 $u_{k+1} \notin \mathrm{span}(v_1, \ldots, v_k)$. 那么向量 $w_{k+1} = u_{k+1} - \langle u_{k+1}, v_1\rangle v_1 - \langle u_{k+1}, v_2\rangle v_2 - \cdots - \langle u_{k+1}, v_k\rangle v_k \neq 0$，且 $w_{k+1} \perp v_j\,(j = 1, \ldots, k)$，且 $\mathrm{span}(v_1, \ldots, v_k, w_{k+1}) = \mathrm{span}(v_1, \ldots, v_k, u_{k+1}) = \mathrm{span}(u_1, \ldots, u_k, u_{k+1})$，最后令 $v_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|}$，则 $v_1, \ldots, v_{k+1}$ 是标正组且 $\mathrm{span}(u_1, \ldots, u_{k+1}) = \mathrm{span}(v_1, \ldots, v_{k+1})$.

由归纳原理，结论得到证明. □

注 1.8. Gram-Schmidt 正交化过程也可以分两步走：

1. 正交化：归纳地让

$$\begin{aligned}
y_1 &= u_1 \\
y_2 &= u_2 - \frac{\langle u_2, y_1 \rangle}{\langle y_1, y_1 \rangle} y_1 \\
&\vdots \\
y_m &= u_m - \frac{\langle u_m, y_1 \rangle}{\langle y_1, y_1 \rangle} y_1 - \frac{\langle u_m, y_2 \rangle}{\langle y_2, y_2 \rangle} y_2 - \cdots - \frac{\langle u_m, y_{m-1} \rangle}{\langle y_{m-1}, y_{m-1} \rangle} y_{m-1}
\end{aligned}$$

   那么 $y_1, y_2, \ldots, y_m$ 是正交组且

$$\operatorname{span}(u_1, \ldots, u_k) = \operatorname{span}(y_1, \ldots, y_k), k = 1, 2, \ldots, m.$$

2. 单位化：令

$$\begin{aligned}
v_1 &= \frac{y_1}{\|y_1\|} \\
v_2 &= \frac{y_2}{\|y_2\|} \\
&\vdots \\
v_m &= \frac{y_m}{\|y_m\|}
\end{aligned}$$

   那么 $v_1, v_2, \ldots, v_m$ 是标准正交组且

$$\operatorname{span}(u_1, \ldots, u_k) = \operatorname{span}(v_1, \ldots, v_k), k = 1, 2, \ldots, m.$$

定理 1.20. *(QR 分解)* 设 $A \in \mathbb{C}^{m\times n}$ 是列满秩矩阵，那么存在 $Q \in \mathbb{C}^{m\times n}$ 以及上三角阵 $R \in \mathbb{C}^{n\times n}$ 使得

$$A = QR,$$

其中 $Q^H Q = I_n$，即 Q 是由 n 个两两正交的单位列向量构成的矩阵. 此分解式称为矩阵的 QR 分解.

假设有两个这样的分解 $A = Q_1 R_1 = Q_2 R_2$,则存在 $D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$, $|d_i| = 1, i = 1, \ldots, n$，使得 $R_1 = DR_2$，$Q_1 = Q_2 D^{-1}$.

在 QR 分解中，可以要求上三角阵 R 的对角元全为正数，那么在这个要求下 QR 分解式是唯一的.

证明. 对 A 的列向量组 $\alpha_1,\dots,\alpha_n$ 进行 Gram-Schmidt 正交化，让

$$\begin{aligned}
\beta_1 &= \alpha_1 \\
\beta_2 &= \alpha_2 - \frac{\langle\alpha_2,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1 \\
&\vdots \\
\beta_n &= \alpha_n - \frac{\langle\alpha_n,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1 - \frac{\langle\alpha_n,\beta_2\rangle}{\langle\beta_2,\beta_2\rangle}\beta_2 - \cdots - \frac{\langle\alpha_n,\beta_{n-1}\rangle}{\langle\beta_{n-1},\beta_{n-1}\rangle}\beta_{n-1}
\end{aligned}$$

以及

$$\begin{aligned}
\gamma_1 &= \frac{\beta_1}{\|\beta_1\|} \\
\gamma_2 &= \frac{\beta_2}{\|\beta_2\|} \\
&\vdots \\
\gamma_n &= \frac{\beta_n}{\|\beta_n\|}
\end{aligned}$$

那么

$$\begin{aligned}
A &= (\beta_1,\dots,\beta_n)\begin{bmatrix} 1 & \frac{\langle\alpha_2,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} & * & * & \frac{\langle\alpha_n,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} \\ & 1 & \cdots & * & * \\ & & \ddots & \vdots & \vdots \\ & & & 1 & \frac{\langle\alpha_n,\beta_{n-1}\rangle}{\langle\beta_{n-1},\beta_{n-1}\rangle} \\ & & & & 1 \end{bmatrix} \\
&= (\gamma_1,\dots,\gamma_n)\begin{bmatrix} \|\beta_1\| & & & & \\ & \|\beta_2\| & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \|\beta_n\| \end{bmatrix}\begin{bmatrix} 1 & \frac{\langle\alpha_2,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} & * & * & \frac{\langle\alpha_n,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} \\ & 1 & \cdots & * & * \\ & & \ddots & \vdots & \vdots \\ & & & 1 & \frac{\langle\alpha_n,\beta_{n-1}\rangle}{\langle\beta_{n-1},\beta_{n-1}\rangle} \\ & & & & 1 \end{bmatrix}
\end{aligned}$$

令 $Q=(\gamma_1,\dots,\gamma_n)$,$R=\begin{bmatrix}\|\beta_1\| & & & & \\ & \|\beta_2\| & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \|\beta_n\|\end{bmatrix}\begin{bmatrix}1 & \frac{\langle\alpha_2,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} & * & * & \frac{\langle\alpha_n,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle} \\ & 1 & \cdots & * & * \\ & & \ddots & \vdots & \vdots \\ & & & 1 & \frac{\langle\alpha_n,\beta_{n-1}\rangle}{\langle\beta_{n-1},\beta_{n-1}\rangle} \\ & & & & 1\end{bmatrix}$,

则 $A=QR$ 是 QR 分解且 R 的对角元都是正实数.

假设 $A=Q_1R_1=Q_2R_2$ 是两个 QR 分解，那么 $Q_2^HQ_1=R_2R_1^{-1}$ 且 $Q_1^HQ_2=R_1R_2^{-1}$. 所以 $Q_2^HQ_1$ 是酉阵,而 $Q_2^HQ_1=R_2R_1^{-1}$ 上三角阵,从而只能是对角的酉阵. 让 $D=Q_2^HQ_1=R_2R_1^{-1}$,则对角阵 D 的对角元 d_k 是绝对值为 1 的复数，所以 $R_2=DR_1$，$Q_1R_1=Q_2R_2=Q_2DR_1$，$Q_1=Q_2D$. □

注 1.9.

1. 在复数域的版本下，QR 分解有时也叫做 UR 分解.
2. 通常我们会要求 A 是可逆方阵，这时候，QR 分解中的 Q 就是我们通常意义下的酉阵.
3. 矩阵 A 是可逆实矩阵的情况也有类似的版本，此时 QR 分解中的 Q 是实正交阵，R 是实的可逆上三角阵.

定理 1.21. *(存在标正基) 有限维内积空间存在标准正交基.*

证明. 取定一个基，利用 Gram-Schmidt 正交化，得到一个标正基. □

定理 1.22. *n 限维内积空间中的任一标正组 $\varepsilon_1,\dots,\varepsilon_k$ 可以扩充为一个标正基 $\varepsilon_1,\dots,\varepsilon_k,\varepsilon_{k+1},\dots,\varepsilon_n$.*

定理 1.23. *n 维实 (复) 内积空间 U 同构于内积空间 $\mathbb{R}^n(\mathbb{C}^n)$.*

证明. 设 U 是 n 维实内积空间，那么存在标正基 $\varepsilon_1,\dots,\varepsilon_n$. 考虑向量在这个基下的坐标映射 $\sigma:U\to\mathbb{R}^n$，首先取坐标映射是 U 到 $\mathbb{R}^n$ 的线性同构，其次 $\langle\sum_{k=1}^n a_k\varepsilon_k,\sum_{k=1}^n b_k\varepsilon_k\rangle_U=\sum_{k=1}^n a_kb_k$，所以 $\forall x,y\in U$, 有 $\langle x,y\rangle_U=\langle\sigma(x),\sigma(y)\rangle_{\mathbb{R}^n}$，这就是说 σ 是保持内积的线性同构.

复的版本是类似的证明. □

正交补空间

定义 1.23. 设 $(V,\langle\cdot,\cdot\rangle)$ 是内积空间，$x\in V$，S 是 V 的非空子集. 如果 $\forall y\in S_2$ 都有 $x\perp y$，则称 x 与 S 是正交的，记为 $x\perp S$.

定义 1.24. 设 $(V,\langle\cdot,\cdot\rangle)$ 是内积空间，设 S_1 和 S_2 是 V 的非空子集. 如果 $\forall x\in S_1,\forall y\in S_2$ 有 $x\perp y$，则称 S_1 与 S_2 是正交的，记为 $S_1\perp S_2$.

定义 1.25. 设 $(V,\langle\cdot,\cdot\rangle)$ 是内积空间，S 是 V 的非空子集，

$$S^{\perp}:=\{y\in V\,|\,\forall x\in S,y\perp x\}$$

是子空间，称为 S 的正交补空间.

定理 1.24. *(正交直和分解)* 设 $(V,\langle\cdot,\cdot\rangle)$ 是内积空间，W 是 V 的有限维子空间，那么 $V=W\oplus W^{\perp}$.

证明. 由于 W 是有限维的，W 存在标正基 $\varepsilon_1,\ldots,\varepsilon_k$. 现在任取 $x\in V$，令 $y=x-\langle x,\varepsilon_1\rangle\varepsilon_1-\cdots-\langle x,\varepsilon_k\rangle\varepsilon_k$，则 $y\perp\varepsilon_j,j=1,\ldots,k$，所以 $y\in W^{\perp}$. 于是 $x=(\langle x,\varepsilon_1\rangle\varepsilon_1+\cdots+\langle x,\varepsilon_k\rangle\varepsilon_k)+y\in W+W^{\perp}$，故 $V=W+W^{\perp}$.

设 $z\in W\cap W^{\perp}$，那么 $z\perp z$，所以 $z=0$，故 $W\cap W^{\perp}=\{0\}$.

综上，$V=W\oplus W^{\perp}$. □

定理 1.25. *(最佳逼近元)* 设 $(V,\langle\cdot,\cdot\rangle)$ 是内积空间，W 是 V 的有限维子空间. 那么对于 $x\in V$，存在唯一的 $y\in W$，使得

$$\|x-y\|=\min_{z\in W}\|x-z\|.$$

该定理中的 y 称为在 W 上对于 x 的最佳逼近元.

证明. 由定理1.24，$V=W\oplus W^{\perp}$，所以 $\forall x\in V$，$\exists y\in W,w\in W^{\perp}$ 使得

$$x=y+w.$$

现在 W 中任取一向量 z,则 $x-z=(x-y)+(y-z)$,而 $x-y\in W^{\perp},y-z\in W$，所以 $\|x-z\|^2=\|x-y\|^2+\|y-z\|^2$，故 $\|x-z\|\geq\|x-y\|$，该不等式取等号时当且仅当 $z=y$. 证毕. □

第二章　矩阵的相似变换

2.1　特征值与特征向量

定义与概念

对于一个 n 阶复方阵 A，我们总是有两个观点：一方面它是一个数表，另一方面将它看成是 $\mathbb{C}^n$ 上的线性变换 $\mathbb{C}^n \to \mathbb{C}^n, x \mapsto Ax$.

本节给出特征值和特征向量的一些基本性质.

本节记号约定：本节中 T 总是指 n 维复线性空间 V 上的一个线性变换，A 总是指 n 阶复方阵，$f(t), f$ 总是指复系数一元多项式.

定义 2.1. 设 T 是 n 维复线性空间 V 上的一个线性变换，$\lambda_0 \in \mathbb{C}$ 和 $0 \neq v_0 \in V$ 满足 $Tv_0 = \lambda_0 v_0$. 则称 λ_0 是 T 的一个特征值，v_0 是 T 的对应于特征值 λ_0 的特征向量.

为了说话方便，我们也直接说 (λ_0, v_0) 是 T 的一个特征对 (eigenpair).

假设 λ_0 是一个特征值，称 $V_{\lambda_0} := \{v \in V \mid Tv = \lambda_0 v\}$ 为 T 关于 λ_0 的特征子空间，$\dim V_{\lambda_0}$ 称为 λ_0 的**几何重数**.

任意取定 V 的一个基 $\varepsilon : \varepsilon_1, \ldots, \varepsilon_n$，设 T 在该基下的矩阵为 A，用 $[\cdot]$ 表示向量在该基下的取坐标映射. 那么 $Tx = y \iff A[x] = [y]$. 所以 $Tv_0 = \lambda_0 v_0 \iff A[v_0] = \lambda_0 [v_0]$. 我们得到结论：

1. λ_0 是 T 的特征值 $\iff \lambda_0$ 是 A 的特征值.
2. v_0 是 T 的属于 λ_0 的特征向量 $\iff [v_0]$ 是 A 的属于特征值 λ_0 的特征向量.

我们将这个结果写成定理形式：

定理 2.1. 设 T 是 n 维复线性空间 V 上的一个线性变换，那么 $\lambda_0 \in \mathbb{C}$ 是 T 的特征值当且仅当 λ_0 是 T 在任一基下的矩阵 A 的特征值；而 $v_0 \in V$ 是 T 的属于 λ_0 的特征向量当且仅当 v_0 在任意一个取定基下的坐标向量 $[v_0]$ 是 T 在该基下的矩阵 A 的属于特征值 λ_0 的特征向量.

显然，λ_0为 n 阶方阵 A 的特征值.

$\iff \exists x_0 \neq 0$，使得x_0满足x的方程$Ax = \lambda_0 x$.

$\iff (\lambda_0 I - A)x = 0$有非零解.

$\iff \det(\lambda_0 I - A) = 0$.

$\iff \lambda_0$为 n 次多项式$\varphi(\lambda) = \det(\lambda I - A)$的一个零点.

定义 2.2. 设 T 是 n 维复线性空间 V 上的一个线性变换，V 中取定一个基后，T 在该基下的矩阵为 A，我们称关于 t 的 n 次多项式 $\varphi(t) = \det(tI - A)$ 为 T 的特征多项式.

当我们直接说 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为 T 的 n 个特征值时，我们总是指特征多项式可以分解为 $\varphi(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$.

注 2.1. 注意到线性变换在不同基下的表示阵是相似的，如果方阵 $B = P^{-1}AP$，那么 $\det(tI - B) = \det(P^{-1}(tI - A)P) = \det(tI - A)$，所以定义2.2中 $\varphi(t)$ 与基的选择无关.

定义 2.3. 设 λ_0 是线性变换 T (或者方阵 $A_{n\times n}$) 的特征值，那么 λ_0 做为 T (或者 A) 的特征多项式的根，其根的重数 k 称为特征值 λ_0 的**代数重数**.

对于方阵 A 的特征值 λ_0 的几何重数 k，由定义它是齐次线性方程 $(A - \lambda_0 I)x = 0$ 解空间维数，即任一个基础解系所含向量元素个数，因此 k 等于 $n - \mathrm{rank}(A - \lambda_0 I)$.

代数重数和几何重数有下面的结果：

定理 2.2. 几何重数 $\leq$ 代数重数.

证明. 假设 T 的特征值 λ_0 的几何重数为 k. 在特征子空间 V_{λ_0} 取一个基 $x_1, \ldots, x_k$，将其扩充为 V 的一个基 $x_1, \ldots, x_k, x_{k+1}, \ldots, x_n$. 那么存在适当矩阵 C, D 使得

$$T(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) \begin{bmatrix} \lambda_0 I_k & C_{k\times(n-k)} \\ 0 & D_{n-k} \end{bmatrix},$$

于是 T 在该基下的矩阵为 $\begin{bmatrix}\lambda_0 I_k & C\\ 0 & D\end{bmatrix}$，$T$ 的特征多项式为 $\varphi(t) = (t-\lambda_0)^k \det(tI_{n-k}-D)$，于是 λ_0 的代数重数 $\geq k$. 证毕. □

基本性质

定义 2.4. 设 $f(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_0 \in \mathbb{C}[\lambda], A \in \mathbb{C}^{n\times n}$，规定 $f(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I$ 称为 A 的一个多项式. 也可以记为 $f(A) = \sum_{i=0}^m a_i A^i$, 此时我们约定 $A^0 = I$.

对于线性变换 T，定义 $f(T) := a_m T^m + a_{m-1}T^{m-1} + \cdots + a_1 T + a_0 E$，其中 E 是恒等变换.

定理 2.3. *设 (λ, x) 为 A 的一个特征对，那么 $(f(\lambda), x)$ 为 $f(A)$ 的一个特征对.*

注 2.2. 以后我们会得到一个更强的结论：$A \in \mathbb{C}^{n\times n}$ 的 n 个特征值为 $\lambda_1, \ldots, \lambda_n$，$f(t) \in \mathbb{C}[t]$. 那么 $f(A)$ 的 n 个特征值为 $f(\lambda_1), \ldots, f(\lambda_n)$.

定理 2.4. *方阵的不同特征值对应的特征向量线性无关.*

证明. 设 $(\lambda_1, v_1), \ldots, (\lambda_s, v_s)$ 为 A 的特征对且 $\lambda_1, \ldots, \lambda_s$ 各不相同. 设 $a_1 v_1 + \cdots + a_s v_s = 0$，任取 $g(t) \in \mathbb{C}[t]$，有 $0 = g(A)(\sum_{k=1}^s a_k v_k) = \sum_{k=1}^s a_k g(\lambda_k) v_k$. 对于任意取定的 $k \in \{1, 2, \ldots, s\}$，存在多项式 g_k 使得 $\begin{cases} g_k(\lambda_k) \neq 0, \\ g_k(\lambda_j) = 0, \quad 对于\ j \neq k\ 时 \end{cases}$，由于 $v_k \neq 0$，所以 $a_k = 0$. 这就说明 $v_1, \ldots, v_s$ 是线性无关的.

对于 g_k 的选择，例如可以取 $g_k(t) = \prod_{j\neq k}(t - \lambda_j)$. □

定理 2.5. *设 $\lambda_1, \lambda_2 \ldots, \lambda_s$ 是 T 的不同特征值，$V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_s}$ 是对应的特征子空间，那么 $\sum_{k=1}^s V_{\lambda_k}$ 是直和 $\oplus_{k=1}^s V_{\lambda_k}$.*

证明. 假设 $v_k \in V_{\lambda_k}, k = 1, 2, \ldots, s, 0 = v_1 + v_2 + \cdots + v_s$. $\forall k \in \{1, 2, \ldots, s\}$，让 $g_k(t) = \prod_{j\neq k}(t - \lambda_j)$，那么 $0 = g_k(T)\left(\sum_{j=1}^s v_j\right) = g_k(\lambda_k) v_k$，由于 $g_k(\lambda_k) \neq 0$，所以 $v_k = 0$.

这就说明 $\sum_{k=1}^s V_{\lambda_k} = \oplus_{k=1}^s V_{\lambda_k}$. □

定理 2.6. *由不同特征值对应的线性无关特征向量组并起来的向量组也是线性无关的.*

定理 2.7. 设 A 为 n 阶方阵，其 n 个特征值为 $\lambda_1, \ldots, \lambda_n$，那么 A^T 的 n 个特征值为 $\lambda_1, \ldots, \lambda_n$，而 A^H 的 n 个特征值为 $\overline{\lambda}_1, \ldots, \overline{\lambda}_n$.

证明. 如果

$$\det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n),$$

则

$$\det(\lambda I - A^H) = \det(\lambda I - \overline{A}) = \lambda^n + \overline{a_{n-1}}\lambda^{n-1} + \cdots + \overline{a_0} = (\lambda - \overline{\lambda_1})\cdots(\lambda - \overline{\lambda_n}).$$

□

定理 2.8. $Tr(A_{m\times n}B_{n\times m}) = Tr(BA)$.

定理 2.9. A 为 n 阶复方阵，$\lambda_1, \ldots, \lambda_n$ 为 n 个特征值，那么 $tr(A) = \lambda_1 + \cdots + \lambda_n, \det(A) = \lambda_1 \cdots \lambda_n$.

定义 2.5. 设 T 是 n 维复线性空间 V 上的一个线性变换，我们也可以定义 T 的行列式 $\det T$ 和 T 的迹 $\mathrm{Tr}(T)$ 分别为 n 个特征值的乘积以及 n 个特征值之和，它们分别等于线性变换 T 在任意选定基下的矩阵的行列式与迹.

Sylvester 定理

定理 2.10. *(Sylvester 定理)* 设 $A_{m\times n}, B_{n\times m}$，则

$$\lambda^n|\lambda I_m - AB| = \lambda^m|\lambda I_n - BA|.$$

证明. 通过考虑 $\begin{bmatrix} \lambda I_n & B_{n\times m} \\ \lambda A_{m\times n} & \lambda I_m \end{bmatrix}$ 的行列式可以得到证明.

或者根据 $\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix}\begin{bmatrix} AB & 0 \\ B & 0_n \end{bmatrix}\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} 0_m & 0 \\ B & BA \end{bmatrix}$ 得到 $\begin{bmatrix} AB & 0 \\ B & 0_n \end{bmatrix}$ 与 $\begin{bmatrix} 0_m & 0 \\ B & BA \end{bmatrix}$ 相似，从而得到证明. □

注 2.3. 上述定理中，如果假设 $m \le n$，那么 BA 的 n 个特征值恰好是 AB 的 m 个特征值再添加 $n - m$ 个零特征值.

例 2.1. 假设 $A \in M_n$ 可以分解为 $A = XY^T$ 其中 $X, Y \in M_{n,r}$ 且 $r < n$，那么 A 的 n 个特征值由 Y^TX 的 r 个特征值和 $n - r$ 个零特征值构成.

例 2.2. 让 $A = xy^T$，$x, y \in \mathbb{C}^n$，那么 A 的 n 个特征值为：$n-1$ 个零和 y^Tx. 让 $A = xy^T + zw^T = \begin{bmatrix} x & z \end{bmatrix} \begin{bmatrix} y & w \end{bmatrix}^T$ 其中 $x, y, z, w \in \mathbb{C}^n$，那么 A 的 n 个特征值由 $n-2$ 个零特征值与 $\begin{bmatrix} y & w \end{bmatrix}^T \begin{bmatrix} x & z \end{bmatrix} = \begin{bmatrix} y^Tx & y^Tz \\ w^Tx & w^Tz \end{bmatrix}$ 的两个特征值构成.

例 2.3. $A = [i+j]_{i,j=1}^n = \begin{bmatrix} 2 & 3 & 4 & \cdots \\ 2 & 4 & 5 & \cdots \\ 4 & 5 & 6 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = ve^T + ev^T = \begin{bmatrix} v & e \end{bmatrix} \begin{bmatrix} e & v \end{bmatrix}^T$，其中 $v = \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix}^T, e = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$. 那么 A 的特征值由 $n-2$ 个零与 $B = \begin{bmatrix} e & v \end{bmatrix}^T \begin{bmatrix} v & e \end{bmatrix} = \begin{bmatrix} \frac{n(n+1)}{2} & n \\ \frac{n(n+1)(2n+1)}{6} & \frac{n(n+1)}{2} \end{bmatrix}$ 的两个特征值构成.

2.2 相似对角化

可相似对角化条件

定义 2.6. 设 $A, B \in \mathbb{C}^{n\times n}$. 若存在可逆阵 $P \in \mathbb{C}^{n\times n}$ 使得 $P^{-1}AP = B$，称 A 相似于 B.

在矩阵相似的定义中，由 $AP = PB$，可知 A 在由 P 的列构成的基下的矩阵是 B. 反之，如果 A 在 $\mathbb{C}^n$ 的某个基下的矩阵为 B，那么 A 相似于 B. 所以 A 能否相似于 B 等价于问能否找到一个合适的基使得 A 在该基下的表示矩阵为 B.

方阵的相似有以下简单的性质：

1. 矩阵的相似关系是一个数学上的等价关系 (见定义8.9).
   从而 A 相似于 B，我们也称 A 与 B 相似.

2. 相似的矩阵具有相同的特征多项式.

3. 若 A 与 B 相似，则矩阵多项式 $f(A)$ 与 $f(B)$ 相似.

定义 2.7. 如果一个方阵可以相似于一个对角矩阵，就称矩阵**可对角化**.

如果一个线性变换 T 在某基下的矩阵是对角阵，就称 T 可对角化.

定义 2.8. 设 T 是 n 维复线性空间 V 上的一个线性变换，W 为 V 的子空间. 如果 $\forall w \in W$，有 $Tw \in W$，则称 W 是 $T-$ **不变子空间**.

例 2.4. 显然，$x \in V$ 是 T 的特征向量当且仅当 $W = \mathbb{C}x := \{ax \mid a \in \mathbb{C}\}$ 为 T 的一维不变子空间.

例 2.5. 设 $A \in \mathbb{C}^{n\times n}$，$W \subseteq \mathbb{C}^n$ 是 k 维 $A-$ 不变子空间，那么在 W 中选定一个基 $x_1, \ldots, x_k$ 并将其扩充为 $\mathbb{C}^n$ 的一个基 $x_1, \ldots, x_n$，则存在矩阵 $B_{k\times k}, C_{k\times(n-k)}, D_{(n-k)\times(n-k)}$ 使得 $A(x_1, \ldots, x_n) = (x_1, \ldots, x_n)\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$,所以 A 相似于分块上三角阵 $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$.

反之，如果 A 可以相似于分块上三角阵 $\begin{bmatrix} B_{k\times k} & C_{k\times(n-k)} \\ 0 & D_{(n-k)\times(n-k)} \end{bmatrix}$，那么 A 有 k 维的不变子空间.

例 2.6. 设 $A \in \mathbb{C}^{n\times n}$,$W_1 \subseteq \mathbb{C}^n$ 是 k 维 $A-$ 不变子空间,$W_2 \subseteq \mathbb{C}^n$ 是 $n-k$ 维 $A-$ 不变子空间,且 $\mathbb{C}^n = W_1 \oplus W_2$,那么在 W_1 中取一个基 $x_1, \ldots, x_k$,在在 W_2 中取一个基 $y_1, \ldots, y_{n-k}$,由此构成 $\mathbb{C}^n$ 的一个基 $x_1, \ldots, x_k, y_1, \ldots, y_{n-k}$,则 A 在该基下的矩阵为分块对角阵的形式 $\begin{bmatrix} B_{k\times k} & \\ & C_{(n-k)\times(n-k)} \end{bmatrix}$. 反之如果 A 相似于分块对角阵 $\begin{bmatrix} B_{k\times k} & \\ & C_{(n-k)\times(n-k)} \end{bmatrix}$，那么 $\mathbb{C}^n$ 是 A 的某个 k 维不变子空间和某个 $n-k$ 维不变子空间的直和.

记号: 如果 $B_1 \in M_{n_1}, B_2 \in M_{n_2}, \ldots, B_s \in M_{n_s}$,那么分块对角阵 $\begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_s \end{bmatrix}$ 也记为 $B_1 \oplus B_2 \oplus \cdots \oplus B_s$.

定理 2.11. *设 A 为 n 阶方阵, $\lambda_1, \ldots, \lambda_s$ 是 A 的全部不同特征值, $\lambda_1, \ldots, \lambda_s$ 的几何重数分别为 $g_1, \ldots, g_s$，代数重数分别为 $m_1, \ldots, m_s$，则以下等价:*

1) A 可对角化.

2) A 具有 n 个线性无关的特征向量.

3) $\mathbb{C}^n$ 是 n 个一维 $A-$ 不变子空间的直和.

4) 特征值 λ_k 的代数重数与其几何重数相等.

5) $\mathbb{C}^n$ *是* A *的特征值子空间的直和：*$\mathbb{C}^n = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s}$.

证明. 1) $\Longrightarrow$ 2) $\Longrightarrow$ 3) 是显然的.

3) $\Longrightarrow$ 4): 设 $\mathbb{C}^n = \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \cdots \oplus \mathbb{C}x_n$，其中 $x_1, \ldots, x_n$ 都是 A 的特征向量. 根据直和的性质知 $x_1, \ldots, x_n$ 线性无关. 设 $\{x_1, \ldots, x_n\}$ 中属于 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 的特征向量分别有 $k_1, k_2, \ldots, k_s$ 个，那么 $k_1 + k_2 + \cdots + k_s = n$. 因为 $k_j \le g_j \le m_j, j = 1, 2, \ldots, s$，又有 $n = \sum_{j=1}^s k_j \le \sum_{j=1}^s g_j \le \sum_{j=1}^s m_j = n$，从而迫使 $k_j = g_j = m_j, j = 1, \ldots, s$.

4) $\Longrightarrow$ 5):由于 $g_j = m_j, j = 1, \ldots, s$ 且 $n = \sum_{j=1}^s m_j$,所以 $\dim(\oplus_{j=1}^s V_{\lambda_j}) = \sum_{j=1}^s g_j = n$. 线性空间 $U = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s}$ 是 $\mathbb{C}^n$ 的子空间且 U 的维数是 n，只能 $U = \mathbb{C}^n$.

5) $\Longrightarrow$ 1)：在每个特征子空间 V_{λ_j} 取一个基 $x_{j1}, x_{j2}, \ldots, x_{jg_j}$，由此构成 $\mathbb{C}^n$ 的一个基 $x_{11}, \ldots, x_{1g_1}, \ldots, x_{s1}, \ldots, x_{sg_s}$，那么 A 在该基下的矩阵为 $\lambda_1 I_{g_1} \oplus \cdots \oplus \lambda_s I_{g_s}$，所以 A 可相似对角化. □

推论 2.1. *如果* n *阶方阵有* n *个不同特征值，那么可以对角化.*

同时对角化

本节研究矩阵同时对角化问题.

例 2.7. 设 n 阶方阵 $A = \lambda_1 I_{n_1} \oplus \lambda_2 I_{n_2} \oplus \cdots \oplus \lambda_s I_{n_s}$，其中 $\lambda_1, \ldots, \lambda_s$ 为各不相同的复数. 假设 n 阶方阵 B 与 A 可以交换：$AB = BA$. 那么存在 $B_k \in \mathbb{C}^{n_k \times n_k}, k = 1, \ldots, s$，使得 $B = B_1 \oplus \cdots \oplus B_s$.

证明. 将 B 进行分块 $B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix}$，其中 B_{ij} 是 $n_i \times n_j$ 矩阵，$i, j = 1, \ldots, s$. 由于 $AB = BA$，于是 $\lambda_i B_{ij} = \lambda_j B_{ij}$，当 $i \ne j$ 时，$\lambda_i \ne \lambda_j$，$B_{ij} = 0$. 故 B 是分块对角阵 $B_{11} \oplus \cdots \oplus B_{ss}$. □

定理 2.12. *设* $B_1, B_2, \ldots, B_d$ *分别是* $n_1, n_2, \ldots, n_d$ *阶复方阵，*$B = B_1 \oplus \cdots \oplus B_d$*，那么* B *可以对角化当且仅当* $B_1, B_2, \ldots, B_d$ *中的每个矩阵都可以对角化.*

证明. 如果 $B_1,\ldots,B_d$ 中的每个矩阵都可以对角化,那么分别存在 $n_1,\ldots,n_d$ 阶可逆阵 $S_1,\ldots,S_d$，使得 $S_1^{-1}B_1S_1,\ldots,S_d^{-1}B_dS_d$ 都是对角阵，令 $S=S_1\oplus\cdots\oplus S_d$，则 $S^{-1}BS=(S_1^{-1}B_1S_1)\oplus\cdots\oplus(S_d^{-1}B_dS_d)$ 是对角阵.

现在假设 B 可以对角化，我们来证明 $B_1,B_2,\ldots,B_d$ 中的每个矩阵都可以对角化.

对个数 d 进行归纳. 当 $d=1$ 时，结论显然成立. 当 $d\geq 2$ 时，$B_1\oplus\cdots\oplus B_d=(B_1\oplus\cdots\oplus B_{d-1})\oplus B_d$. 因此根据归纳法，只需要证明 $d=2$ 的情形即可.

假设 $B=B_1\oplus B_2$，S 是可逆阵，D 为对角阵，使得 $BS=SD$. 将 S 进行分块 $S=\begin{bmatrix}X_{n_1\times(n_1+n_2)}\\ Y_{n_2\times(n_1+n_2)}\end{bmatrix}$，那么 $B_1X=XD$ 且 $B_2Y=YD$. 由于 $B_1X=XD$，所以 X 的每个非零列向量都是 B_1 的特征向量. 因为 $\operatorname{rank}(X)=n_1$，所以 X 中存在 n_1 个线性无关的列向量，所以 B_1 有 n_1 个线性无关的特征向量，故 B_1 可对角化. 类似的，根据 $B_2Y=YD$，$\operatorname{rank}(Y)=n_2$，可以知道 B_2 有 n_2 个线性无关的特征向量，于是 B_2 也可以对角化. □

定义 2.9. 设 $A,B\in\mathbb{C}^{n\times n}$，如果存在一个共同的可逆阵 $S\in\mathbb{C}^{n\times n}$ 使得 $S^{-1}AS$ 和 $S^{-1}BS$ 都是对角阵，则称为 A,B 可以同时对角化.

设 $\emptyset\neq\mathcal{F}\subseteq\mathbb{C}^{n\times n}$ 是一族矩阵，如果存在共同的可逆阵 $S\in\mathbb{C}^{n\times n}$ 使得对于 $\forall A\in\mathcal{F}$ 都有 $S^{-1}AS$ 是对角阵，则称 $\mathcal{F}$ 可以同时对角化.

定理 2.13. *假设 $A,B\in M_n$ 可以对角化. 那么 A 与 B 可同时对角化当且仅当它们可交换.*

证明. 对角阵与对角阵之间是可交换的，我们只需证明当 A,B 可交换时 A,B 可以同时对角化.

设 A 的所有不同特征值为 $\lambda_1,\ldots,\lambda_s$ 其代数重数分别为 $n_1,\ldots,n_s$，由于 A 可以对角化,故存在可逆阵 P 使得 $P^{-1}AP=\lambda_1I_{n_1}\oplus\cdots\oplus\lambda_sI_{n_s}$. 因为 A 与 B 可以交换,所以 $P^{-1}AP$ 与 $P^{-1}BP$ 可以交换,于是 $P^{-1}BP$ 是分块对角阵的形式: 存在 $C_1\in M_{n_1},\ldots,C_s\in M_{n_s}$ 使得 $P^{-1}BP=C_1\oplus\cdots\oplus C_s$. 又因为 B 可以对角化,从而 $C_1\oplus\cdots\oplus C_s$ 可以对角化,根据定理2.12,$C_1,\ldots,C_s$ 都可以对角化，存在可逆的 $T_j\in M_{n_j},j=1,\ldots,s$ 使得 $T_j^{-1}C_jT_j,j=1,\ldots,s$ 是对角阵. 现在让 $S=P(T_1\oplus\cdots\oplus T_s)$，那么 $S^{-1}AS=\lambda_1I_{n_1}\oplus\cdots\oplus\lambda_sI_{n_s}$ 与 $S^{-1}BS=(T_1^{-1}C_1T_1)\oplus\cdots\oplus(T_s^{-1}C_sT_s)$ 都是对角阵. □

同时对角化定理可以推广到任意多个矩阵的情形：

定理 2.14. 设 $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ 是由一些可对角化的矩阵构成，那么 $\mathcal{F}$ 可同时对角化当且仅当 $\mathcal{F}$ 中任意两个矩阵都可交换.

证明. 如果 $\mathcal{F}$ 可同时对角化，显然 $\mathcal{F}$ 中任意两个矩阵可交换.

现在假设 $\mathcal{F}$ 中任意两个矩阵可交换，我们来证明 $\mathcal{F}$ 可同时对角化. 对阶数 n 进行归纳，$n=1$ 时结论显然成立. 假设阶数 $\leq n-1$ 时结论成立，往证阶数为 n 时结论也成立. 如果 $\mathcal{F}$ 中所有矩阵都是数量阵，那么 $\mathcal{F}$ 已经同时对角化. 现假设 $A \in \mathcal{F}$ 是具有两个以上不同特征值的矩阵，设 A 的不同特征值 $\lambda_1, \ldots, \lambda_s$ 的代数重数分别为 $k_1, \ldots, k_s$. 对 $\mathcal{F}$ 中的矩阵进行同一个相似变换后，不妨设 $A = \lambda_1 I_{k_1} \oplus \cdots \lambda_s I_{k_s}$. 由于 $\mathcal{F}$ 中的任意一个矩阵都于 A 交换，所以 $\mathcal{F}$ 中的任意一个矩阵 B 都是分块对角阵的形式：$B_1 \oplus \cdots \oplus B_s$，其中 $B_j \in \mathbb{C}^{k_j \times k_j}, j=1,\ldots,s$. 任取 $\mathcal{F}$ 中的两个矩阵 $C = C_1 \oplus \cdots \oplus C_s$ 和 $D = D_1 \oplus \cdots \oplus D_s$，由于 C 和 D 可交换，所以对于每个 $j \in \{1, \ldots, s\}$，C_j 和 D_j 均可交换，注意到每一块子矩阵的阶数都不超过 $n-1$，由归纳假设存在对应阶数的相似变换可逆阵 $P_1, \ldots, P_s$，使得对于 $\mathcal{F}$ 中的任意一个矩阵 $B = B_1 \oplus \cdots \oplus B_s$，$P_j^{-1} B_j P_j, j = 1, \ldots, s$ 是对角阵. 那么 $P = P_1 \oplus \cdots \oplus P_s$ 对角化 $\mathcal{F}$ 中的每个矩阵.

证毕. □

2.3 Jordan 标准形介绍

问题. 我们如何才能判断两个同阶复方阵何时相似？

例 2.8. 考虑矩阵 $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 与矩阵 $B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$，它们有相同的特征多项式，也有相同的秩，但是 $A^2 = 0$ 而 $B^2 \neq 0$，所以 A 与 B 不相似.

Jordan 标准形定理

定义 2.10. 设 $\lambda_0 \in \mathbb{C}$，k 为自然数，我们将矩阵

$$J_k(\lambda_0) := \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}_{k\times k}$$

称为 k 阶 **Jordan 块**. 当 $k=1,2$ 时，$J_1(\lambda_0)=[\lambda_0]$，$J_2(\lambda_0)=\begin{bmatrix} \lambda_0 & 1 \\ & \lambda_0 \end{bmatrix}$.

由若干个 Jordan 块构成的分块对角矩阵

$$J = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_q}(\lambda_q),\ n_1+n_2+\cdots+n_q=n,$$

称为 **Jordan 矩阵**.

定理 2.15. （*Jordan* 标准形定理）*任何一个 n 阶复方阵 A 都可以相似于一个 Jordan 矩阵 J. 这个矩阵 J 在除 Jordan 块的排列次序外由 A 唯一确定，称 J 为 A 的 Jordan 标准形.*

从这个大定理我们可以得到以下结果.

1. 重数计算在内，J 的所有特征值，从而是 A 的所有特征值，就是 J 的对角元素.

2. 对于给定的特征值 λ_0，它的几何重数等于 $n-\mathrm{rank}(A-\lambda_0 I)$，等于对应于特征值 λ_0 的 Jordan 块的块数.

3. 对于给定的特征值 λ，所有对应于特征值 λ 的 Jordan 块的阶数之和等于特征值 λ 的代数重数.

4. 方阵 A 可以对角化当且仅当对于 A 的每个特征值 λ 其代数重数等于其几何重数.

5. 对应于特征值 λ 的 Jordan 块中的阶数至少是 j 的块数有 $\mathrm{rank}(A-\lambda I)^{j-1}-\mathrm{rank}(A-\lambda I)^j$，因此阶数刚好为 j 的块 $J_j(\lambda)$ 共有 $\mathrm{rank}(A-\lambda I)^{j-1}+\mathrm{rank}(A-\lambda I)^{j+1}-2\mathrm{rank}(A-\lambda I)^j$ 个.

6. 对于特征值 λ, 它在极小多项式中的根的重数等于其对应的 Jordan 块中阶数最大的块的阶数.

我们证明其中一条，记 $N = J_n(0)$，那么 $\operatorname{rank}(N^k) = \begin{cases} n-k, & k \leq n \\ 0, & k \geq n \end{cases}$. 那么

$$\operatorname{rank}(N^{k-1}) - \operatorname{rank}(N^k) = \begin{cases} 1, & 1 \leq k \leq n \\ 0, & k > n \end{cases}.$$

考虑 $\operatorname{rank}(A-\lambda I)^{j-1} - \operatorname{rank}(A-\lambda I)^j$，则

$$\mu_j(\lambda) = \operatorname{rank}(A-\lambda I)^{j-1} - \operatorname{rank}(A-\lambda I)^j = \sum (\operatorname{rank}(J_{r_i}^{j-1}(\lambda_i-\lambda)) - \operatorname{rank}(J_{r_i}^{j}(\lambda_i-\lambda))$$

$$\operatorname{rank}(J_{r_i}^{j-1}(\lambda_i-\lambda)) - \operatorname{rank}(J_{r_i}^{j}(\lambda_i-\lambda) = \begin{cases} 0, & \lambda \neq \lambda_i \\ 1, & \lambda = \lambda_i \text{且} 1 \leq j \leq r_i \\ 0, & \lambda = \lambda_i \text{且} j > r_i \end{cases}$$

于是 $\mu_j(\lambda) = \#\{i | \lambda = \lambda_i \text{且} 1 \leq j \leq r_i\}$ 即 A 中 Jordan 块里面以 λ 为特征值的且块的阶数大于等于 j 的那些 Jordan 块的块数. 最后我们得到 A 中 $J_j(\lambda)$ 的块数为 $\mu_j(\lambda) - \mu_{j+1}(\lambda)$.

Lambda 矩阵介绍

定义 2.11. 设 $A(\lambda) = (a_{ij}(\lambda))_{m\times n}$，其中 $a_{ij}(\lambda)$ 都是以 λ 为不定元的复系数一元多项式，称 $A(\lambda)$ 是 $\lambda-$ 矩阵. $m \times n$ 阶 $\lambda-$ 阵全体构成的集合，可以用 $M_{m,n}(\mathbb{C}[\lambda])$ 表示. 类似的，用 $M_n(\mathbb{C}[\lambda])$ 表示 n 阶 λ 阵全体构成的集合.

注 2.4. 从记号上来说，$A(\lambda)$ 只是为了强调我们是在 λ 阵里面工作.

方阵 A 的特征矩阵 $\lambda I - A$ 就是一个 $\lambda-$ 矩阵. $\lambda-$ 矩阵包括数字矩阵. $\lambda-$ 矩阵也可以定义加法和乘法，它们与数字矩阵的运算有相同的运算规律. $\lambda-$ 矩阵也可以定义行列式 $|A(\lambda)|$ (此时行列式的值是一个多项式)，它与数字矩阵的行列式有类似的性质.

多项式环 $\mathbb{C}[\lambda]$ 在代数结构上与整数环 $\mathbb{Z}$ 有许多类似的地方，对应地，矩阵环 $M_n(\mathbb{C}[\lambda])$ 与 $M_n(\mathbb{Z})$ 也有许多类似的性质. 本节中许多概念和结论都可以平行地放在 $\mathbb{Z}$-矩阵上考虑，当你这么做时，你要注意的是你只在整数范围里面工作.

定义 2.12. 设 $A(\lambda)$ 为 $\lambda-$ 矩阵，那么 $A(\lambda)$ 的最高阶非零子式的阶数就定义为 $A(\lambda)$ 的秩，记为 $\operatorname{rank}(A(\lambda))$.

注意到 $A(\lambda)$ 的元素都是多项式，所以 $A(\lambda)$ 的任何一个子式也是多项式，我们说非零子式指的是该子式不是零多项式.

例 2.9. 设 A 为 n 阶复方阵，那么 $\lambda I - A$ 做为 $\lambda-$ 阵它的秩为 n，因为 $\det(\lambda I - A)$ 是关于 λ 的 n 次多项式不是零多项式.

定义 2.13. 设 $A(\lambda)$ 为 n 阶 λ 阵，如果存在 n 阶 λ 阵 $B(\lambda)$ 使得

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I,$$

则称 $A(\lambda)$ 是 (在 λ 阵中) **可逆的**，其逆为 $B(\lambda)$，记号为 $A^{-1}(\lambda) = B(\lambda)$.

例 2.10. n 阶 $\lambda-$ 阵 $A(\lambda)$ 在 λ 阵中是可逆的当且仅当 $\det A(\lambda)$ 是一个非零常数.

证明. 设 $A(\lambda)$ 可逆，那么存在 n 阶 $\lambda-$ 阵 $B(\lambda)$ 使得 $A(\lambda)B(\lambda) = I$，两边取行列式，得 $\det A(\lambda) \cdot \det B(\lambda) = 1$，故 $\det A(\lambda) \in \mathbb{C} \setminus \{0\}$.

反之如果 $c = \det A(\lambda) \in \mathbb{C} \setminus \{0\}$，那么让 $B = \frac{1}{c}\mathrm{adj}(A(\lambda))$，则 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I$，所以 $A(\lambda)$ 是可逆的. □

例 2.11. 请你给出整数矩阵在整数范围内可逆的定义，并尝试研究 n 阶整数阵 A 在 $M_n(\mathbb{Z})$ 可逆的充要条件.

λ 矩阵经过初等变换可以化为 Smith 标准形.

定理 2.16. *设 λ 矩阵 $A(\lambda)$ 是秩为 r 的 $m \times n$ 矩阵，则可通过初等变换化为如下形式的矩阵*

$$S(\lambda) = \begin{bmatrix} d_1(\lambda) & & & & \\ & d_2(\lambda) & & & \\ & & \ddots & & \\ & & & d_r(\lambda) & \\ & & & & 0_{(m-r)\times(n-r)} \end{bmatrix}$$

其中 $d_i(\lambda)$ 都是首一多项式，且 $d_i | d_{i+1}$，$i = 1, 2, \ldots, r$. 这个 $S(\lambda)$ 是由 $A(\lambda)$ 唯一确定的，称为 Smith 标准形，d_i 称为 $A(\lambda)$ 的第 i 个不变因子.

按以下步骤可以求出复方阵 A 的 Jordan 标准形.

1. 将 $\lambda I - A$ 化为 Smith 标准形，求出 $d_1, d_2, \ldots, d_n$，称为 A 的**不变因子**.

2. 将 A 中每个次数大于零的不变因子进行分解因式分解成一次因式方幂的乘积，全部这些一次方幂合在一起称为 A 的**初等因子** (组)，设 A 的初等因子 (组) 为

$$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\ldots,(\lambda-\lambda_s)^{r_s}$$

其中 $\lambda_1,\lambda_2,\ldots,\lambda_s$ 可能相同，$r_1+r_2+\cdots+r_s=n$.

3. 对于每个初等因子 $(\lambda-\lambda_i)^{r_i}$ 写出对应的 Jordan 块 $J_{r_i}(\lambda_i)$

4. A 的 Jordan 标准形就是 $J_{r_1}(\lambda_1)\oplus J_{r_2}(\lambda_2)\oplus\cdots\oplus J_{r_s}(\lambda_s)$

行列式因子法

定义 2.14. 设 λ 矩阵 $A(\lambda)$ 的秩为 r，对于正整数 $k(1\le k\le r)$，$A(\lambda)$ 的全部 k 阶行列式的首一的最大公因式 $D_k(\lambda)$ 称为 $A(\lambda)$ 的 k 阶行列式因子.

定理 2.17. *设 $A(\lambda)$ 是秩为 r 的 $m\times n$ 矩阵，则 $A(\lambda)$ 的行列式因子 $D_k(\lambda)$ 为*

$$D_k=d_1d_2\cdots d_k$$

其中 d_i 是 $A(\lambda)$ 的第 i 个不变因子.

对于一个 n 阶复方阵 A，我们称 $\lambda I-A$ 的不变因子为 A 的不变因子，$\lambda I-A$ 的行列式因子为 A 的行列式因子. 注意到这个时候 $\lambda I-A$ 的秩是 n，$\lambda I-A$ 的 n 个行列式因子 $D_1,D_2,\ldots,D_n$ 就决定了 A 的 n 个不变因子，从而确定了 A 的初等因子，确定出它的 Jordan 标准形.

λ 矩阵的理论是很强的理论，对于域 F 上的方阵 A 和 B 有下面定理

定理 2.18. *设 A,B 为域 F 上的两个 n 阶方阵，则 A 与 B 相似 (即存在 F 上的 n 阶可逆矩阵 P 使得 $P^{-1}AP=B$) 当且仅当 A,B 具有相同的不变因子或者具有相同的行列式因子.*

定理 2.19. $J_n(\lambda_0)=\begin{pmatrix}\lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0\end{pmatrix}_{n\times n}$ *，则*

$$J^k=\sum_{i=0}^{\infty}\frac{k(k-1)\cdots(k-i+1)}{i!}\lambda_0^{k-i}N^i=\sum_{i=0}^{\infty}\frac{1}{i!}(\lambda^k)^{(i)}N^i|_{\lambda=\lambda_0}$$

于是对于一个多项式 $f(\lambda)$ 而言，$f(J_n(\lambda_0)) = \sum_{i=0}^{\infty} \frac{f^{(i)}(\lambda_0)}{i!} N^i$，其中 $N = J_n(0)$.

定理 2.20. 设 A, B 是 n 阶方阵，并且是上三角矩阵，则 AB 也是上三角矩阵, 且 AB 对角线元素为 A, B 对角线元素之积.

定理 2.21. 设 A 的 n 个特征值为 $\lambda_1, \ldots, \lambda_n$，$f(t)$ 为一个多项式，则 $f(A)$ 的 n 个特征值为 $f(\lambda_1), \ldots, f(\lambda_n)$.

2.4 Hamilton-Cayley 定理

定义 2.15. 设 $A \in \mathbb{C}^{n \times n}$，$f(\lambda)$ 是一个多项式. 如果有 $f(A) = 0$，则称 f 为**零化** A 的一个多项式.

定理 2.22. 设 $A \in \mathbb{C}^{n \times n}$，特征多项式 $\varphi(\lambda) = det(\lambda I - A)$，则 $\varphi(A) = 0$，即 φ 零化 A.

例 2.12. 已知矩阵 $A = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$, 计算 $A^7 - A^5 - 19A^4 + 28A^3 + 6A - I$, A^{-1},A^{100}.

定义 2.16. 设 $A \in \mathbb{C}^{n \times n}$，在零化 A 的多项式中，次数最低且首一的多项式称为 A 的极小多项式，记为 $m_A(\lambda)$.

定理 2.23. 方阵 A 的极小多项式 $m_A(\lambda)$ 整除任何一个零化 A 的多项式，且极小多项式是唯一的.

定理 2.24. 方阵 A 的特征值都是其极小多项式 $m_A(\lambda)$ 的零点，即若 λ_0 是 A 的特征值，那么 $m_A(\lambda_0) = 0$.

定理 2.25. 相似的矩阵具有相同的极小多项式.

定理 2.26. 方阵 A 的极小多项式 $m_A(\lambda) = d_n(\lambda)$，其中 d_n 为 A 的第 n 个不变因子.

定理 2.27. 复方阵可相似对角化当且仅当它的极小多项式没有重根.

例 2.13. 例如 n 阶幂等阵 A, 满足 $A^2 = I$, 那么 A 必然相似于 $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$, 其中 r 为 A 的特征值 $\lambda = 1$ 的代数重数, $s = n - r$.

2.5 酉相似下的标准形

定理 2.28. *(Schur)* 设 A 为 n 阶复方阵，则 A 可以酉相似于上三角矩阵.

定义 2.17. 设 $A \in \mathbb{C}^{n\times n}$, 若 A 满足

$$AA^H = A^H A$$

则称 A 为正规矩阵.

例 2.14. 酉阵，实正交阵，Hermite 矩阵，反 Hermite 矩阵 ($A^H = -A$)，对角阵都是正规阵. 设 A 是正规阵，那么 A 的多项式也正规阵.

例 2.15. 根据正规阵的定义，只要 A^H 是和 A 可以交换的矩阵，那么 A 就是正规阵，例如 A^H 是 A 的多项式时，A 是正规阵. 另外 A 是正规阵时，A^H 也是正规阵.

以下定理是正规阵的刻画定理.

定理 2.29. 设 $A \in \mathbb{C}^{n\times n}$，那么 A 可以酉相似于对角阵当且仅当 A 是正规阵.

显然，对于正规阵 A 而言，$A = 0$ 当且仅当 A 的所有特征值都是零.

定理 2.30. 设 $A \in \mathbb{C}^{n\times n}$，则 A 为正规阵 $\Leftrightarrow$ 对于 $\forall x \in \mathbb{C}^n$，有 $||Ax|| = ||A^H x||$.

定理 2.31. A 正规，则对任何数 k，$kI - A$ 是正规阵, A^H 也是正规阵.

定理 2.32. 设 A 正规，$Ax_0 = \lambda x_0$, $x_0 \neq 0$，则有 $A^H x_0 = \overline{\lambda} x_0$.

推论 2.2. *Hermite* 矩阵的特征值都是实数，反 *Hermite* 矩阵的特征值为零或者是纯虚数.

推论 2.3. 设 A 正规，则不同特征值对应的特征向量是正交的.

证明. $\lambda_1(x_1, x_2) = (Ax_1, x_2) = (x_1, A^H x_2) = (x_1, \overline{\lambda_2} x_2) = \lambda_2(x_1, x_2)$. □

定义 2.18. 设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 矩阵. 如果对于任意 $0 \neq x \in \mathbb{C}^n$ 都有

$$x^H Ax > 0 (\geq 0)$$

就称 A 是 Hermite 正定矩阵（半正定矩阵）

需要 (Ax, x) 是实数，所以需要 A 是 Hermite 的.

定理 2.33. *设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 矩阵. 则以下条件等价:*

1. *A 为 Hermite 正定矩阵*
2. *A 的特征值全都是正实数*
3. *存在可逆阵 P 使得 $A = P^H P$*

推论 2.4. *Hermite 正定矩阵的行列式大于零.*

定理 2.34. *设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 矩阵. 则以下条件等价:*

1. *A 为 Hermite 半正定矩阵*
2. *A 的特征值全都是非负实数*
3. *存在 $P \in \mathbb{C}^{n\times n}$ 使得 $A = P^H P$*

定理 2.35. *设 $A \in \mathbb{C}^{m\times n}$，则*

1. *$A^H A$ 和 AA^H 的特征值全为非负实数. 它们是半正定矩阵.*
2. *$A^H A$ 和 AA^H 的非零特征值相同.*
3. *$rank(A^H A) = rank(AA^H) = rank(A)$.*

定理 2.36. *设 $A \in \mathbb{C}^{n\times n}$，$A$ 为 Hermite 矩阵. 则 A 正定当且仅当它的所有顺序主子式都大于零.*

第三章　矩阵分解

3.1　QR 分解

定义 3.1. 设 $u \in \mathbb{C}^n$ 是单位向量，即 $||u||_2 = 1$，称

$$H = I - 2uu^H$$

为 Householder 矩阵或初等反射阵.

定理 3.1. *设 H 为 n 阶 Householder 矩阵，则有以下性质：*

1. $H^H = H$（*Hermite*）
2. $H^H H = I$（酉阵）
3. $H^2 = I$
4. $H^{-1} = H$
5. $\begin{bmatrix} I_r & \\ & H \end{bmatrix}$ 是 $n+r$ 阶 *Householder* 阵.
6. $\det H = -1$.

定理 3.2. *设 $z \in \mathbb{C}^n (n \geq 2)$ 是单位向量，则对任意 $x \in \mathbb{C}^n$，对满足 $\alpha x^H z \in \mathbb{R}$ 且 $|\alpha| = ||x||_2$ 的复数 α，存在 Householder 阵 H 使得 $Hx = \alpha z$. (其中条件 $\alpha x^H z \in \mathbb{R}$ 且 $|\alpha| = ||x||_2$ 是为了保证 $x = \frac{1}{2}(x + \alpha z) + \frac{1}{2}(x - \alpha z)$ 是正交直和分解）*

第四章　矩阵的特殊乘积

4.1　Kronecker 积

定义 4.1. 设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{p\times q}$, 称如下的矩阵

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

为 A 与 B 的直积或 Kronecker 积. 它是 $mp\times nq$ 矩阵，以 $a_{ij}B$ 为子块构成的分块矩阵.

定义 4.2. (vec 拉直算子) 矩阵 $A_{m\times n}=(A_1,A_2,\ldots,A_n)$，定义按列拉直向量 $\mathrm{vec}(A)=\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}$.

例 4.1. $A=\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$，$\mathrm{vec}(A)=\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$.

例 4.2. $\mathrm{vec}(A^T)$ 是将 A 按行拉直后得到的行向量再转置.

4.2 Hadamard 积

定义 4.3. $A=(a_{ij}),B=(b_{ij})\in\mathbb{C}^{m\times n}$，$A$ 与 B 的 Hadamard 积 $A\circ B$ 定义为 $A\circ B:=(a_{ij}b_{ij})_{m\times n}$.

Hadamard 积具有以下性质：

1. $A\circ B=B\circ A$，$k(A\circ B)=(kA)\circ B=A\circ(kB)$.

2. $A\circ(B+C)=A\circ B+A\circ C$，$(A\circ B)\circ C=A\circ(B\circ C)$.

3. $(A\circ B)^T=A^T\circ B^T$，$(A\circ B)^H=A^H\circ B^H$.

4. $A_{m\times n}$，$B_{m\times n}$，C 为 m 阶对角阵，D 为 n 阶对角阵，那么 $C(A\circ B)=(CA)\circ B=A\circ(CB)$，$(A\circ B)D=A\circ(BD)=(AD)\circ B$.

5. 设 $x,y,z,w\in\mathbb{C}^n$，那么 $(xy^H)\circ(zw^H)=(x\circ z)(y\circ w)^H$.

定理 4.1. *$rank(A\circ B)\le rank(A)rank(B)$.*

证明. 如果 A,B 之一为零矩阵，则结论显然成立. 下设 $\operatorname{rank}(A)=r_1>0$，$\operatorname{rank}(B)=r_2>0$. 那么存在向量 $x_1,\ldots,x_{r_1}\in\mathbb{C}^m$，$y_1,\ldots,y_{r_1}\in\mathbb{C}^n$ 使得 $A=x_1y_1^T+\cdots+x_{r_1}y_{r_1}^T$. 存在向量 $u_1,\ldots,u_{r_2}\in\mathbb{C}^m$，$v_1,\ldots,v_{r_2}\in\mathbb{C}^n$ 使得 $B=u_1v_1^T+\cdots+u_{r_2}v_{r_2}^T$. 从而 $A\circ B=\sum_{k=1}^{r_2}\sum_{j=1}^{r_1}(x_jy_j^T)\circ(u_kv_k^T)$ 为 r_1r_2 个矩阵之和，我们只需要证明加和项的每个矩阵 $(x_jy_j^T)\circ(u_kv_k^T)$ 的秩不超过 1 即可.

对于 $\alpha=(a_1,\ldots,a_m)^T,\beta=(b_1,\ldots,b_m)^T\in\mathbb{C}^m$ 以及 $\gamma=(c_1,\ldots,c_n)^T,\delta=(d_1,\ldots,d_n)^T\in\mathbb{C}^n$，有

$$(\alpha\gamma^T)\circ(\beta\delta^T)=(a_ic_j)_{m\times n}\circ(b_id_j)_{m\times n}=(a_ib_ic_jd_j)_{m\times n}=(\alpha\circ\beta)(\gamma\circ\delta)^T$$

于是 $(\alpha\gamma^T)\circ(\beta\delta^T)=(\alpha\circ\beta)(\gamma\circ\delta)^T$ 是两个秩不超过 1 的矩阵之乘积，从而 $\operatorname{rank}((\alpha\circ\beta)(\gamma\circ\delta)^T)\le 1$.

所以 $\operatorname{rank}(A\circ B)\le\operatorname{rank}(A)\operatorname{rank}(B)$. □

定理 4.2. *设 A,B 均为 n 阶 Hermite（半）正定矩阵，则 $A\circ B$ 也是 Hermite（半）正定矩阵.*

证明. (1) 因为 A,B 是 Hermite 半正定矩阵，可设 $A=\lambda_1u_1u_1^H+\cdots+\lambda_ru_ru_r^H$，其中 $\lambda_1,\ldots,\lambda_r>0$ 且 $u_1,\ldots,u_r$ 为 $\mathbb{C}^n$ 中两两正交的单位向量.

可设 $B=\delta_1 v_1 v_1^H+\cdots+\delta_t v_t v_t^H$，其中 $\delta_1,\ldots,\delta_t>0$ 且 $v_1,\ldots,v_t$ 为 $\mathbb{C}^n$ 中两两正交的单位向量. 那么

$$A\circ B=\sum \lambda_j\delta_k(u_j u_j^H)\circ(v_k v_k^H)=\sum \lambda_j\delta_k(u_j\circ v_k)(u_j\circ v_k)^H$$

为半正定矩阵.

(2) 如果 A,B 是正定的，那么在 (1) 的证明中矩阵 A,B 的分解假设中 $r=t=n$. 此时

$$A\circ B=\sum_{j,k=1}^{n}\lambda_j\delta_k(u_j\circ v_k)(u_j\circ v_k)^H.$$

. 设 $x=(x_1,\ldots,x_n)^T\in\mathbb{C}^n$，$x^H(A\circ B)x=0$. 那么 $\sum_{j,k=1}^n\lambda_j\delta_k x^H(u_j\circ v_k)(u_j\circ v_k)^H x=0$，从而 $\forall j,k,(u_j\circ v_k)^H x=0$，即

$$\forall j,k,u_j^H(\overline{v_k}\circ x)=0.$$

因为 $u_1,\ldots,u_n$ 构成 $\mathbb{C}^n$ 的一个标准正交基，上式说明 $\forall k,(\overline{v_k}\circ x)=0$，或者等价的，$\forall k,(v_k\circ\overline{x})=0$. 那么有

$$\begin{bmatrix}\overline{x_1}&&\\&\ddots&\\&&\overline{x_n}\end{bmatrix}(v_1,\ldots,v_n)=0,$$

由于 $(v_1,\ldots,v_n)$ 构成可逆矩阵，故 $x=0$. 因此 $A\circ B$ 是正定的. $\square$

在定理 4.2 的正定部分的证明中，我们发现可以对 B 的正定性条件弱化也能得到结论。我们假设 A 是正定，B 是半正定的，那么可以设 $A=\lambda_1 u_1 u_1^H+\cdots+\lambda_n u_n u_n^H$,$B=\delta_1 v_1 v_1^H+\cdots+\delta_t v_t v_t^H$，其中记号约定同上，那么由 $x^H(A\circ B)x=0$ 可得到

$$\begin{bmatrix}\overline{x_1}&&\\&\ddots&\\&&\overline{x_n}\end{bmatrix}(v_1,\ldots,v_t)=0,$$

只需要 $n\times t$ 矩阵 $(v_1,\ldots,v_t)$ 的每一行都是非零行即可得到 $x=0$，从而得到 $A\circ B$ 是正定矩阵. 但是矩阵 $(v_1,\ldots,v_t)$ 每一行都是非零行这一条件与矩阵 B 的对角元是正数这一条件等价. 于是我们有下面的改进结果：

定理 4.3. 如果设 A 为 n 阶 *Hermite* 正定矩阵，设 B 为 n 阶 *Hermite* 半正定矩阵且对角元 $b_{ii} > 0(i = 1, 2, \ldots, n)$，则 $A \circ B$ 也是 *Hermite* 正定矩阵.

定理 4.4. 设 $A \in \mathbb{R}^{n\times n}$ 是对称矩阵，那么 A 是半正定矩阵当且仅当对所有 n 阶半正定矩阵 $B = (b_{ij})_{n\times n}$ 有 $\sum_{i,j=1}^n a_{ij}b_{ij} \geq 0$.

4.3 Geršgorin 定理

定义 4.4. 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ 且 $n \geq 2$. 如果

$$|a_{ii}| > R_i(A) := \sum_{j\neq i} |a_{ij}|, \quad i = 1, 2, \ldots, n,$$

则称 A 为**行严格对角占优阵**.

如果

$$|a_{ii}| > C_i(A) := \sum_{j\neq i} |a_{ji}|, \quad i = 1, 2, \ldots, n,$$

则称 A 为**列严格对角占优阵**.

无特殊说明，严格对角占优阵总是相对行而言的.

定理 4.5. 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ 且 $n \geq 2$ 是严格对角占优阵，则 A 是可逆矩阵.

证明. 如果 A 不可逆，那么存在非零的 $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ 使得 $Ax = 0$. 因为 $x \neq 0$，取下标 k 使得 $|x_k| = \max_j |x_j|$. 考虑 Ax 的第 k 分量 $a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n$，那么 $|a_{kk}||x_k| \leq \sum_{j\neq k} |a_{kj}x_j| = \sum_{j\neq k} |a_{kj}||x_j|$，从而 $|a_{kk}| \leq \sum_{j\neq k} |a_{kj}|\frac{|x_j|}{|x_k|} \leq R_k$ 与 A 是严格对角阵矛盾. 这就说明 A 是可逆矩阵. □

定理 4.6. *(戈氏圆盘定理)* 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ 且 $n \geq 2$，复平面上的圆域 $G_i(A) := \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i(A)\}, i = 1, 2, \ldots, n$, 称为 n 个戈氏圆盘. 那么 A 的 n 个特征值 $\lambda_1, \ldots, \lambda_n$ 落在这 n 个圆盘之并集 $G(A) = \cup_{i=1}^n G_i(A)$ 之中.

进一步，如果这 n 个圆盘中有 k 个圆盘之并集 D_k 和剩余的 $n-k$ 个圆盘之并集 F_{n-k} 不相交，那么这 k 个圆盘之并集 D_k 恰好含有 k 个特征值 *(重数计算在内)*.

证明. 只证明定理的第一部分，第二部分比较难，省去证明.

假设 λ 是 A 的特征值，那么 $\lambda I - A$ 是不可逆矩阵，从而不可能是严格对角占优阵，于是 $\lambda \in \cup_{i=1}^{n} G_i(A)$. □

由于 A 与 A^T 具有相同的特征值，于是戈氏圆盘定理也有列的版本. 请写出上述定理的列版本的形式.

注 4.1. 注意戈氏圆盘定理中的分离性条件是不能去掉的，戈氏圆盘定理说 n 个特征值落在 n 个圆盘之并集 $G(A) = \cup_{i=1}^{n} G_i(A)$ 之中，并不能说每个圆盘里都有特征值.

戈氏圆盘定理列的版本：

推论 4.1. $A = (a_{ij})_{n\times n} \in \mathbb{C}^{n\times n}$ 的 n 个特征值落在 $G(A^T) = \cup_{j=1}^{n}\{z \in \mathbb{C} | |z - a_{jj}| \leq C_j\}$ 中.

推论 4.2. 如果 A 的 n 个戈氏圆盘两两不相交，那么 A 有 n 个不同特征值.

例 4.3. $A = \begin{bmatrix} -4 & -10 \\ 1 & 6 \end{bmatrix}$ 的两个特征值为 $1 \pm \sqrt{15}$，$G_1 = \{z||z+4| \leq 10\}$，$G_2 = \{z||z-6| \leq 1\}$，$G_2$ 中不含有任何特征值.

例 4.4. 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ 且 $n \geq 2$ ，尝试用戈氏圆盘定理证明 $\rho(A) \leq \min\{||A||_\infty, ||A||_1\}$. 回忆:$||A||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$,$||A||_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$.

证明. 设 λ 为 A 的任何一个特征值，由戈氏圆盘定理，$|\lambda - a_{ii}| \leq R_i$，对于任意 i 成立. 由于 $|\lambda| - |a_{ii}| \leq |\lambda - a_{ii}|$，所以 $|\lambda| \leq |a_{ii}| + R_i$，$|\lambda| \leq ||A||_\infty$，这就说明 $\rho(A) \leq ||A||_\infty$.

使用列形式的戈氏圆盘定理，可以证明另外一部分结论. □

例 4.5. 用 Geršgorin 定理证明：矩阵

$$A = \begin{bmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

有至少两个不同的实特征值.

证明. 在复平面内画出区域 $G_1 = \{z||z-9| \le 4\}$，$G_2 = \{z||z-8| \le 2\}$，$G_3 = \{z||z-4| \le 1\}$，$G_4 = \{z||z-1| \le 1\}$. 那么 G_4 与 $G_1 \cup G_2 \cup G_3$ 不相交，G_4 中含有 1 个特征值, $G_1 \cup G_2 \cup G_3$ 中恰好有 3 个特征值. 由于 A 是实矩阵, A 的非实特征值共轭成对出现. 从而 G_4 中的特征值是实数，并且 $G_1 \cup G_2 \cup G_3$ 中至少含有一个实特征值. □

因为对于任何一个可逆阵 P，矩阵 $P^{-1}AP$ 与 A 具有相同的特征值，因此我们可以适当的选择 P，对 $P^{-1}AP$ 应用戈氏圆盘定理可望对特征值得到较好的估计. 选择 P 为对角阵 $D = \mathrm{diag}(p_1, p_2, \ldots, p_n)$，$d_i > 0, i = 1, \ldots, n$. 此时 $D^{-1}AD = (p_i^{-1} a_{ij} p_j)_{n\times n}$.

推论 4.3. 设 $A = (a_{ij})_{n\times n} \in \mathbb{C}^{n\times n}$，$d_{1,}, d_2, \ldots, d_n$ 为正实数，那么 A 的 n 个特征值落在

$$G(D^{-1}AD) = \cup_{i=1}^{n} \{z \in \mathbb{C}||z - a_{ii}| \le \frac{1}{p_i} \sum_{j\neq i} p_j |a_{ij}|\}$$

之中. 进一步，如果 $G_1(D^{-1}AD), \ldots, G_n(D^{-1}AD)$ 这 n 个圆盘中有 k 个圆盘之并集和剩余的 $n-k$ 个圆盘之并集不相交，那么这 k 个圆盘之并集恰好含有 k 个特征值 *(*重数计算在内*)*.

例 4.6. 假设 $A \in \mathbb{C}^{n\times n}$ 的 n 个 Geršgorin 圆盘两两不相交. 如果 A 是实矩阵，证明 A 有 n 个实的不同特征值.

例 4.7. 假设 $A = (a_{ij})_{n\times n}$ 是严格对角占优的，如果 $a_{ii} > 0, i = 1, 2, \ldots, n$，那么 A 的每个特征值都有正的实部. 如果 A 是 Hermite 的且对角元 $a_{ii} > 0, i = 1, \ldots, n$，那么 A 是正定的.

4.4 矩阵广义逆的几何

（TODO）

现在来考虑 Moore-Penrose 方程的几何意义.

我们设 $A_{m\times n}$, 记 $U = \mathbb{C}^n, V = \mathbb{C}^m, V_1 = Im(A), U_0 = Ker(A)$.

定理 4.7. 设 $G \in A\{1\} \iff \forall y \in V_1$, 有 $AGy = y$. 通俗的说，*1*-逆 G 相当于一个挑原像的作用. 设 $G \in A\{1\}$，那么 G 限制在 V_1 上是单射，且 $G(V_1)$ 是 U_0 在 U 中的直和补.

定理 4.8. 设 $G \in A\{1\}$，则有以下性质：

1. AG 和 GA 都是幂等阵.
2. $Im(AG) = Im(A), Ker(GA) = Ker(A)$. 那么 $U = Im(GA) \oplus U_0, V = V_1 \oplus Ker(AG)$. 且限制映射 $G|_{V_1} : V_1 \to G(V_1)$ 与 $A|_{G(V_1)} : G(V_1) \to V_1$ 互为逆，而 $A|_{U_0} = 0$, $Ker(AG) = G^{-1}(U_0)$.
3. 若还有 $G \in A\{1,2\}$,那么 $A \in G\{1\}$,从而还有 $Im(GA) = Im(G), Ker(AG) = Ker(G)$.

定理 4.9. 集合 $A\{1\}$ 与集合 $\Sigma_1 = \{(U_1, LinearMap(V_0, U_0)|U_1 \oplus U_0 = U, V_0 \oplus V_1 = V\}$ 有一一对应关系. 对应关系可以为 $\varphi_1 : G \mapsto \big(G(V_1), G|_{Ker(AG)} : Ker(AG) \to U_0\big)$.

定理 4.10. 集合 $A\{1,2\}$ 与集合 $\Sigma_{12} = \{(U_1, V_0)|U_1 \oplus U_0 = U, V_0 \oplus V_1 = V\}$ 一一对应. 对应关系可以为 $\varphi_{12} : G \mapsto (G(V_1), Ker(G))$.

定理 4.11. $A\{1,3\}$ 与 $\Sigma_{13} = \big\{(U_1, LinearMap(V_1^{\perp}, U_0)|U_1 \oplus U_0 = U\big\}$ 一一对应，对应关系为 $\varphi_1|_{A\{1,3\}}$.

定理 4.12. $A\{1,4\}$ 与 $\Sigma_{14} = \{LinearMap(V_0, U_0)|V_0 \oplus V_1 = V\}$ 一一对应，对应关系为 $\varphi_{14} : G \mapsto (G|_{Ker(AG)} : Ker(AG) \to U_0)$.

现在容易看出，A^+ 实际上是唯一的映射满足 $A^+|_{V_1^{\perp}} = 0$ 和 $A^+|_{V_1}$ 为 $A|_{U_0^{\perp}}$ 的逆.

第五章　范数理论

5.1　向量范数

定义 5.1. 设 V 是域 $\mathbb{F}$ 上的线性空间，这里 $\mathbb{F}$ 是实数域 $\mathbb{R}$ 或者是复数域 $\mathbb{C}$. 函数 $\|\cdot\|: V \to \mathbb{R}$ 称为 V 上的一个范数，如果函数 $\|\cdot\|$ 满足以下三条公理:

1. 正定性：$\forall x \in V, \|x\| \geq 0$，且等号成立当且仅当 $x = 0$.
2. 绝对齐次性：$\forall k \in \mathbb{F}, \forall x \in V$，$\|kx\| = |k| \|x\|$.
3. 三角不等式：$\forall x, y \in V$，$\|x + y\| \leq \|y\| + \|y\|$.

此时称 V 为赋范线性空间.

例 5.1. $\mathbb{C}^n$ 中，对于任意 x，取 $\|x\|_2 = \sqrt{\sum_{k=1}^n |x_k|^2}$，是范数，称为 $2-$ 范数. $2-$ 范数是酉不变范数：即对于任意酉阵 U，$\|Ux\|_2 = \|x\|_2$.

例 5.2. $\mathbb{C}^n$ 中，对于任意 x，取 $\|x\|_1 = \sum_{k=1}^n |x_k|$，是范数，称为 1 范数.

例 5.3. l$\mathbb{C}^n$ 中，对于任意 x，取 $\|x\|_\infty = \max_k |x_k|$，是范数，称为 ∞ 范数.

定义 5.2. $\mathbb{C}^n$ 中 p-范数的定义. 设 $p > 1$，$x \in \mathbb{C}^n$，由等式 $||x||_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ 定义出的范数称为 p 范数.

为证明上述定义中的 p 范数确实是范数，验证 p 范数确实是范数，难点是在于解决三角步等式的证明，该三角不等式也称为 Minkowski 不等式. 通常分析教材会在此处额外介绍两个不等式，Young 不等式与 Holder 不等式，利用 Holder 不等式得到 Minkowski 不等式.

另外一方面，显然，之前定义的 2 范数是 $p = 2$ 时的 p 范数. 而 $\lim_{p\to 1} \|x\|_p = \|x\|_1$ 且 $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$.

引理 5.1. （*Young's Inequality)* 对于任意实数 $\alpha \geq 0$ 和 $\beta \geq 0$，都有

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$$

其中 $p > 1, q > 1$ 且 $\frac{1}{p} + \frac{1}{q} = 1$.

证明. 利用对数函数的凹凸性，立即可以得到结论. □

定理 5.1. （*Hölder)* 设 $x, y \in \mathbb{C}^n$，有

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$$

其中 $p > 1, q > 1$ 且 $\frac{1}{p} + \frac{1}{q} = 1$.

Minkowski 不等式的证明.

我们此处给出另外一种证明.

证明. 欲证 $\|a+b\|_p \leq \|a\|_p + \|b\|_p$，不妨设 $\|a\|_p \neq 0, \|b\|_p \neq 0$. 于是等价于证 $\left\|\frac{a}{\|a\|_p+\|b\|_p} + \frac{b}{\|a\|_p+\|b\|_p}\right\|_p \leq 1$,不等式变形为 $\left\|\frac{\|a\|_p}{\|a\|_p+\|b\|_p} \cdot \frac{a}{\|a\|_p} + \frac{\|b\|_p}{\|a\|_p+\|b\|_p} \cdot \frac{b}{\|b\|_p}\right\|_p \leq 1$. 只需要证明，$\|\lambda x + \mu y\|_p \leq 1$ 其中 $0 \leq \lambda, \mu \leq 1, \lambda + \mu = 1, \|x\|_p = \|y\|_p = 1$. 直接写开，

$$\begin{aligned}
\|\lambda x + \mu y\|_p &= \left(\sum_{k=1}^{n} |\lambda x_k + \mu y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} (\lambda|x_k| + \mu|y_k|)^p\right)^{\frac{1}{p}} \\
&\leq \left(\sum_{k=1}^{n} (\lambda|x_k|^p + \mu|y_k|^p)\right)^{\frac{1}{p}} = 1.
\end{aligned}$$

□

定理 5.2. 设 $A \in \mathbb{C}_n^{m\times n}$，$||\cdot||_a$ 是 $\mathbb{C}^m$ 上的一个向量范数. 对于任意 $x \in \mathbb{C}^n$，我们规定

$$||x||_b = ||Ax||_a$$

则 $||x||_b$ 是 $\mathbb{C}^n$ 中的向量范数.

定义 5.3. 距离空间.

依范数收敛.

有限维赋范线性空间范数等价定理.

5.2 广义矩阵范数

定义 5.4. $\mathbb{C}^{m\times n}$ 上的广义矩阵范数 $||\cdot||$ 就是指通常的满足范数公理的范数.

有了范数，我们可以谈收敛. 应用有限维空间上的范数等价定理，可以知道 $\mathbb{C}^{m\times n}$ 上的任意两个广义的矩阵范数是等价的.

设 $A^{(k)}=[a_{ij}^{(k)}]\in\mathbb{C}^{m\times n}, k=1,2,\cdots$. 如果

$$\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij},\forall i,j$$

则称矩阵序列 $\left\{A^{(k)}\right\}$ 收敛于 $A=[a_{ij}]$,或者它有极限 A,记做 $\lim_{k\to\infty}A^{(k)}=A$.

显然对于任意一个广义矩阵范数 $||\cdot||$, 我们有 $\lim_{k\to\infty}A^{(k)}=A$ 当且仅当 $\lim_{k\to\infty}||A^{(k)}-A||=0$

若 $\lim_{k\to\infty}A^{(k)}=A$，则有 $\lim_{k\to\infty}||A^{(k)}||=||A||$.

定义 5.5. 设 $||\cdot||_\alpha,||\cdot||_\beta,||\cdot||_\gamma$ 分别是 $\mathbb{C}^{m\times n},\mathbb{C}^{n\times p},\mathbb{C}^{m\times p}$ 上的广义矩阵范数. 若对任意的 $A\in\mathbb{C}^{m\times n},B\in\mathbb{C}^{n\times p}$ 有

$$||AB||_\gamma\le||A||_\alpha||B||_\beta$$

我们称 $||\cdot||_\alpha,||\cdot||_\beta,||\cdot||_\gamma$ 是相容的.

这个定义包含了广义矩阵范数与向量范数的相容性定义.

例 5.4. $\mathbb{C}^{m\times n}$ 上的 Frobenius 范数: $||A||_F=(\sum|a_{ij}|^2)^{\frac{1}{2}}=\sqrt{tr(A^HA)}=\sqrt{tr(AA^H)}$.

F 范数是酉不变的，且具有相容性.

例 5.5. $\mathbb{C}^{m\times n}$ 上的 $||A||_{m_1}=\sum_{i,j}|a_{ij}|$ 与向量的 p 范数是相容的.

定义 5.6. 设 $||\cdot||_\delta$ 与 $||\cdot||_\tau$ 分别为 $\mathbb{C}^n$ 与 $\mathbb{C}^m$ 上的向量范数. 由下式定义的非负实值函数 $||\cdot||_{\tau\delta}:\mathbb{C}^{m\times n}\to\mathbb{R}$

$$||A||_{\tau\delta}=\max_{0\neq x\in\mathbb{C}^n}\frac{||Ax||_\tau}{||x||_\delta}=\max_{||x||_\delta=1}||Ax||_\tau=\max_{||x||_\delta\le1}||Ax||_\tau$$

叫做由向量范数 $||\cdot||_\delta$ 与 $||\cdot||_\tau$ 导出的广义矩阵范数或算子范数.

注 5.1. 算子范数有相容性.

定理 5.3. 设 $A = (a_{ij}) \in \mathbb{C}^{m\times n}$，则

1. 若 $||\cdot||_1$ 为 $\mathbb{C}^n$ 与 $\mathbb{C}^m$ 上的 1 范数（l_1 范数），那么它们导出的广义矩阵范数为：
$$||A||_1 = \max_j \sum_i |a_{ij}|$$
也称为列和范数.

2. 若 $||\cdot||_2$ 为 $\mathbb{C}^n$ 与 $\mathbb{C}^m$ 上的 2 范数（l_2 范数），那么它们导出的广义矩阵范数为：
$$||A||_2 = \sqrt{\lambda_1}, \quad \text{其中}\lambda_1\text{为}A^HA\text{的最大特征值}$$

3. 若 $||\cdot||_\infty$ 为 $\mathbb{C}^n$ 与 $\mathbb{C}^m$ 上的 ∞ 范数（l_∞ 范数），那么它们导出的广义矩阵范数为：
$$||A||_\infty = \max_i \sum_j |a_{ij}|$$
也称为行和范数.

5.3 矩阵范数

定义 5.7. 对于 $\mathbb{C}^{n\times n}$ 上的广义矩阵范数（即满足范数的三个公理)，然后再加上一条相容性条件（或者叫次可乘条件）：$\|AB\| \le \|A\|\,\|B\|$，则称 $\|\cdot\|$ 是 $\mathbb{C}^{n\times n}$ 上的矩阵范数.

注 5.2. 当我们看到矩阵范数这四个字，通常我们是要要求有相容性条件的. 我们这里说的是方阵的矩阵范数，相容性条件可以没有含糊的要求给出.

对于由向量范数导数的矩阵范数 $\|\cdot\|$ 而言，单位矩阵 I 的范数 $\|I\| = 1$，但是当 $n > 1$ 时，Frobenius 范数 $\|I\|_F = \sqrt{n}$，从而不是导出的矩阵范数.

定理 5.4. 若 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上的矩阵范数，$S \in \mathbb{C}^{n\times n}$ 非奇异，则如下定义的实值函数
$$\|A\|_S = \left\|S^{-1}AS\right\|, \forall A \in \mathbb{C}^{n\times n}$$
是 $\mathbb{C}^{n\times n}$ 上的矩阵范数.

定理 5.5. 若 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上的矩阵范数，则存在 $\mathbb{C}^n$ 上的向量范数 $\|\cdot\|_v$ 满足：
$$\|Ax\| \le \|A\|\,\|x\|_v$$

定理 5.6. 设 $A \in \mathbb{C}^{n\times n}$, U, V 为 n 阶酉矩阵, 则 $\left\|A^H\right\|_2 = \|A\|_2, \|UAV\|_2 = \|A\|_2$

例 5.6. 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$，规定 $||A||_{m_1} = \sum_{i,j}|a_{ij}|$，则 $||A||_{m_1}$ 是矩阵范数.

例 5.7. 设 $A = (a_{ij}) \in \mathbb{C}^{n\times n}$，规定 $||A||_{m_\infty} = n\max_{i,j}|a_{ij}|$，则 $||\cdot||_{m_\infty}$ 是矩阵范数.

例 5.8. $\mathbb{C}^{m\times n}$ 上，常用的矩阵范数有：

1. $\|\boldsymbol{A}\|_{m_1} = \sum|a_{ij}|$, m_1 范数.
2. $\|A\|_{\mathrm{F}} = \sqrt{\operatorname{tr}(A^H A)}$, F 范数.
3. $\|A\|_M = \max\{m, n\}\max|a_{ij}|$, 最大范数.
4. $\|A\|_G = \sqrt{mn}\max|a_{ij}|$, 几何平均范数.
5. $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$, 列和范数.
6. $\|A\|_2 = \sqrt{\lambda_1(A^H A)}$，谱范数.
7. $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$，行和范数.

5.4 范数应用举例

方阵的谱半径

定义 5.8. 设 $A \in \mathbb{C}^{n\times n}$，$\lambda_1, \ldots, \lambda_n$ 为 A 的 n 个特征值，称

$$\rho(A) = \max_j |\lambda_j|$$

为 A 的谱半径.

定理 5.7. 设 $A \in \mathbb{C}^{n\times n}$，则有

1. $\rho(A^k) = \rho^k(A)$

2. $\rho(A^H A) = \rho(A^H A) = ||A||_2^2$

3. 当 A 是正规阵时，$\rho(A) = ||A||_2$

定理 5.8. 设 $A \in \mathbb{C}^{n\times n}$，则对 $\mathbb{C}^{n\times n}$ 上的任意矩阵范数 $||\cdot||$，有

$$\rho(A) \le ||A||$$

定理 5.9. 设 $A \in \mathbb{C}^{n\times n}$，对于任意给定的正数 ϵ，存在有一个自相容的矩阵范数 $||\cdot||$，使得

$$||A|| \le \rho(A) + \epsilon$$

例 5.9. 设 A 是幂零阵，但是 $A \neq 0$，那么 $\rho(A) = 0$，而不管什么矩阵范数都有 $||A|| > 0$.

定理5.8和5.9说明

$$\rho(A) = \inf\left\{||A|| : ||\cdot||\text{跑遍 } \mathbb{C}^{n\times n} \text{ 上的任意矩阵范数}\right\}$$

5.5 条件数：矩阵逆和线性方程的解

作为矩阵范数和向量范数的应用，我们考虑计算矩阵逆和方程解的误差估计.

假设线性方程 $Ax = b$ 中 A 是可逆的，但是由于机器计算或者观测误差什么我们可能对于真实的 A 以及真实的 b 会有一点点小的扰动 ΔA 和 Δb，现在需要考虑这些扰动会对计算带来什么影响.

具体的，让 $||\cdot||$ 是 $\mathbb{C}^{n\times n}$ 上的一个矩阵范数, $B = A + \Delta A$，我们假设 $||A^{-1}\Delta A|| < 1$, 这个假设是用来保证 B 也是可逆的.

我们需要估计 $\frac{||A^{-1}-B^{-1}||}{||A^{-1}||}$. 推导细节需要注意一个式子 $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$.

引理 5.2. 设 $P \in \mathbb{C}^{n\times n}$，若对于某个矩阵范数 $||\cdot||$ 有 $||P|| < 1$, 那么 $I - P$ 是可逆的.

定理 5.10. 设 $A, \Delta A \in \mathbb{C}^{n\times n}$,$||A^{-1}\Delta A|| < 1$ 对某个矩阵范数成立. 那么

1. $B = A + \Delta A$ 是可逆的.
2. $||B^{-1}|| \le \frac{||A^{-1}||}{1-||A^{-1}\Delta A||}$
3. $\frac{||A^{-1}-B^{-1}||}{||A^{-1}||} \le \frac{||A^{-1}\Delta A||}{1-||A^{-1}\Delta A||} \le \frac{||A^{-1}||\,||A||}{1-||A^{-1}\Delta A||}\frac{||\Delta A||}{||A||}$

令 $\kappa(A)=\begin{cases}||A^{-1}||\,||A|| & 若A是可逆的\\ \infty & 若A是不可逆的\end{cases}$，称为 A 对于矩阵范数 $||\cdot||$ 的条件数. 注意有 $\kappa(A)\geq ||A^{-1}A||\geq 1$. 我们有

$$\frac{||A^{-1}-(A+\Delta A)^{-1}||}{||A^{-1}||}\leq\frac{\kappa(A)}{1-||A^{-1}\Delta A||}\frac{||\Delta A||}{||A||}$$

如果我们加强假设为 $||A^{-1}||\,||\Delta A||<1$，则有

$$\frac{||A^{-1}-(A+\Delta A)^{-1}||}{||A^{-1}||}\leq\frac{\kappa(A)}{1-\kappa(A)\frac{||\Delta A||}{||A||}}\frac{||\Delta A||}{||A||}$$

后面的这个上界我们成为对于计算 A 的逆的相对误差的一个先验上界（*priori bound).*

如果 $||A^{-1}||\,||\Delta A||$ 不仅仅小于 1, 而是比 1 小很多，那么那个上界可以约等于 $\kappa(A)\frac{||\Delta A||}{||A||}$, 只要 $\kappa(A)$ 不是很大, 在这个前提下，这个使我们相信逆的相对误差和原来的相对误差是同阶的. 基于这个原因，我们称呼 A 是病态的 ill conditioned，如果条件数很大. 如果条件数很小就称良好的 well conditioned . 如果 $\kappa(A)=1$，我们说是完美 perfectly conditioned 的. 当然这些称呼都是相对具体给定的矩阵范数 $||\cdot||$ 而言的.

类似的我们考虑 $Ax=b$，$b\neq 0$，其中 b 做一个小扰动 Δb，考虑

$$(A+\Delta A)(x+\Delta x)=b+\Delta b$$

对 $\frac{||\Delta x||_v}{||x||_v}$ 的估计问题，这里设 $||\cdot||_v$ 和矩阵范数 $||\cdot||$ 是相容的.

那么经过一番操作之后我们得到这么一个结论：

$$\frac{||\Delta x||_v}{||x||_v}\leq\frac{\kappa(A)}{1-\kappa(A)\frac{||\Delta A||}{||A||}}\left(\frac{||\Delta b||_v}{||b||_v}+\frac{||\Delta A||}{||A||}\right)$$

第六章 矩阵分析

6.1 矩阵序列

定义 6.1. 设有 $\mathbb{C}^{m\times n}$ 中的矩阵序列 $\{A^{(k)}\}_{k=1}^{\infty}$, 其中 $A^{(k)}=(a_{ij}^{(k)})_{m\times n}$. 若存在 $A=(a_{ij})_{m\times n}$，使得对于每个 (i,j)，有 $\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij}$，就称矩阵序列 $\{A^{(k)}\}_{k=1}^{\infty}$ 收敛于 A，记为 $\lim_{k\to\infty} A^{(k)}=A$. 不收连的矩阵序列称为发散的序列.

根据范数等价定理，有以下定理.

定理 6.1. 设 $\|\cdot\|$ 为任何一个给定的矩阵范数. 则 $\lim A^{(k)}=A \iff \lim\left\|A^{(k)}-A\right\|=0$.

定理 6.2. 设 $\lim A^{(k)}=A$，$\lim B^{(k)}=B$，$\lambda,\mu\in\mathbb{C}$, 那么有

- $\lim_{k\to\infty}(\lambda A^{(k)}+\mu B^{(k)})=\lambda A+\mu B$
- $\lim A^{(k)}B^{(k)}=AB$
- 当所有 $A^{(k)}$ 与 A 都可逆时，$\lim(A^{(k)})^{-1}=A^{-1}$.

定义 6.2. 若 $\lim A^k=0$, 则称 A 为收敛矩阵.

定理 6.3. A 为收敛阵，当且仅当 $\rho(A)<1$.

推论 6.1. 若有某矩阵范数 $\|\cdot\|$，使得 $\|A\|<1$，则 A 为收敛阵.

6.2 矩阵函数

定义 6.3. 设方阵 $A\in\mathbb{C}^{n\times n}$，其极小多项式为

$$m_A(\lambda)=(\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_s)^{l_s}$$

其中 $\lambda_1, \ldots, \lambda_s$ 各不相同. 称集合 $\{(\lambda_i, l_i) | i = 1, 2, \ldots, s\}$ 为 A 的简谱，记为用 Λ_A.

定义 6.4. 设 $A \in \mathbb{C}^{n\times n}$，称复值函数 $f(\lambda)$ 在 A 的谱 Λ_A 上有定义，是指：

$$\begin{cases} f(\lambda_1), f'(\lambda_1), \ldots, f^{(l_1-1)}(\lambda_1) \\ \cdots \\ f(\lambda_s), f'(\lambda_s), \ldots, f^{(l_s-1)}(\lambda_s) \end{cases}$$

均存在. 若函数 $f(\lambda)$ 在 Λ_A 上有定义，则记上述数组为 $f(\Lambda_A)$.

定理 6.4. 设 $A \in \mathbb{C}^{n\times n}$, $f(\lambda), g(\lambda)$ 为复系数多项式, 则有 $f(A) = g(A) \iff f(\Lambda_A) = g(\Lambda_A)$

定义 6.5. 设 $A\in \mathbb{C}^{n\times n}$, $f(\lambda)$ 为一个在 Λ_A 上有定义的复函数. 若存在多项式 $p(A)$，使得 $p(\Lambda_A) = f(\Lambda_A)$，则定义 $f(A) = p(A)$, 并称 p 为 $f(A)$ 的一个定义多项式，我们称 f 在 A 上有定义.

尽管 p 的选择有很多个但是 $f(A)$ 的值是和 p 的选择无关的. 而在约定次数小于 $\deg m_A$ 的多项式中选择，那么 p 是唯一的.

多项式 p 的存在性是与所谓的 Hermite 多项式差值问题紧密相关的. 我们这里就不打算证明了.

根据这种定义方式，我们容易得到一些很好的结果.

定理 6.5. 若 $A = P^{-1}BP$，则任意在 Λ_A 上有定义的函数 $f(\lambda)$ 在 Λ_B 上也有定义，因为 $\Lambda_A = \Lambda_B$. 而且 $f(A) = P^{-1}f(B)P$.

定理 6.6. $f(A_1 \oplus \cdots \oplus A_k) = f(A_1) \oplus \cdots \oplus f(A_k)$.

现在，对于给定方阵 A, 让 A 的 Jordan 分解为 $S(J_1 \oplus \cdots \oplus J_m)S^{-1}$，其中

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{l_i \times l_i}$$

若 f 在 Λ_A 上有定义，则 $f(A) = S(f(J_1) \oplus \cdots \oplus f(J_m))S^{-1}$. 现在考察 J_i 的最小多项式 $(\lambda - \lambda_i)^{l_i}$，令

$$p(\lambda) = f(\lambda_i) + \frac{f'(\lambda_i)}{1!}(\lambda - \lambda_i) + \cdots + \frac{f^{(l_i-1)}(\lambda_i)}{(l_i - 1)!}(\lambda - \lambda_i)^{l_i - 1}$$

，那么显然 $p(\Lambda_{J_i}) = f(\Lambda_{J_i})$，则有

$$f(J_i) = p(J_i) = \sum_{k=0}^{l_i-1} \frac{f^{(k)}(\lambda_i)}{k!} N_{l_i}^k$$

其中 $N_{l_i} = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}_{l_i \times l_i}$

对于形态较好的在某个开圆域上的解析函数，按照收敛的定义得到的矩阵和上述定义是一致的.

定理 6.7. 设复变函数 $f(z)$ 在开圆域 $|z - z_0| < r$ 内有幂级数展开式: $f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)$，若方阵 A 的所有特征值都在此开圆域内，则有 $f(A) = \sum_{k=0}^{\infty} (A - z_0 I)^k$.

定理 6.8. 设 $G(y_1, \ldots, y_l)$ 是 l 元多项式, 函数 $f_1(\lambda), \ldots, f_l(\lambda)$ 都在 Λ_A 上有定义. 记 $g(\lambda) = G(f_1(\lambda), \ldots, f_l(\lambda))$，则有 $g(A) = G(f_1(A), \ldots, f_l(A))$. 特别地，若 $g(\Lambda_A) = 0$，则有 $G(f_1(A), \ldots, f_l(A)) = 0$

定理 6.9. 设复合函数 $h(\lambda) = g(f(\lambda))$, A 为 n 阶阵，为方便描述，我们直接假设 f 在 A 的每个特征值 λ_i 处有 $n-1$ 阶导, g 在每个 $\mu_i = f(\lambda_i)$ 处有 $n-1$ 阶导, 那么显然 f 在 A 上有定义，$B = f(A)$，且 g 在 B 上也有定义, h 在 A 上有定义. 最终有 $h(A) = g(B) = g(f(A))$.

例 6.1. 由 $\cos^2 x + \sin^2 x = 1$ 知 $\cos^2 A + \sin^2 A = I$.

例 6.2. 给定任意可逆方阵 A，对于任何非零复数 p，总存在矩阵 B 使得 $B^p = A$.

定义 6.6. 设 $A \in \mathbb{C}^{n \times n}$，其特征值为 $\lambda_1, \ldots, \lambda_n$，复变函数 $f(\lambda)$ 在 $\lambda_1, \ldots, \lambda_n$ 处解析，那么 $f(A)$ 有定义. 找一个多项式 $p(\lambda)$ 使得 $p^{(k)}(\lambda_i) = f^{(k)}(\lambda_i), k = 0, 1, \ldots, n-1$ 对任意 i 成立. 则可以定义 $f(A) = p(A)$.

上面的定义，是一个简单的版本.

定理 6.10. 设 $A \in \mathbb{C}^{n \times n}$，其特征值为 $\lambda_1, \ldots, \lambda_n$，复变函数 $f(\lambda)$ 在 $\lambda_1, \ldots, \lambda_n$ 处解析，复变函数 $g(\lambda)$ 在 $f(\lambda_1), \ldots, f(\lambda_n)$ 处解析，那么复合函数 $h(\lambda) = g(f(\lambda))$ 在 $\lambda_1, \ldots, \lambda_n$ 处解析，且有 $h(A) = g(f(A))$.

定理 6.11. 设 $A \in \mathbb{C}^{n\times n}$，其特征值为 $\lambda_1,\dots,\lambda_n$，复变函数 $f_1(\lambda),\dots,f_s(\lambda)$ 在 $\lambda_1,\dots,\lambda_n$ 处解析，$q(u_1,\dots,u_s)$ 为一个多项式，那么复合函数 $h(\lambda) = q(f_1(\lambda),\dots,f_s(\lambda))$ 在在 $\lambda_1,\dots,\lambda_n$ 处解析,且有 $h(A) = q(f_1(A),\dots,f_s(A))$.

6.3 矩阵的微积分

这一节的记号我们重新定义. 以至于能够和微积分有一定的相容性. 采取的是 Numerator layout convention.

即:$\frac{\partial y}{\partial \overrightarrow{x}} = [\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \dots, \frac{\partial y}{\partial x_n}]$, $\frac{\partial \overrightarrow{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$, $\frac{\partial \overrightarrow{y}}{\partial \overrightarrow{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$,

$\frac{\partial y}{\partial X} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1q}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2q}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y}{\partial x_{p1}} & \frac{\partial y}{\partial x_{p2}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}$, $\frac{\partial Y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1q}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2q}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y}{\partial x_{p1}} & \frac{\partial y}{\partial x_{p2}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}$

第七章　复习题

关于多项式互素，我们以下结果，可以直接使用.

定理 7.1. *如果多项式* $f(x), g(x) \in \mathbb{C}[x]$ *在* $\mathbb{C}$ *上没有公共零点，那么必然存在多项式* $u(x), v(x) \in \mathbb{C}[x]$ *使得* $u(x)f(x) + v(x)g(x) = 1$.

7.1　习题

这是你需要的复习题

1. 简答以下题目（给出它们的定义，或者给出某种刻画方式，或者给出计算公式）：

   (a) 给出 $\mathbb{C}^n$ 上的向量范数的定义.

   (b) 给出正规矩阵的定义.

   (c) 给出 Moore-Penrose 逆的两种计算办法.

   (d) 写出关于矩阵乘积的特征多项式的 Sylvester 定理的内容.

   (e) 设 $x \in \mathbb{C}^n$，写出 $||x||_\infty, ||x||_1$ 的计算公式.

   (f) 什么是 Hermite 正定矩阵？

   (g) 写出矩阵的 Moore-Penrose 逆的定义.

   (h) 什么是矩阵的奇异值分解？极分解？QR 分解？

   (i) 写出 n 阶方阵 A 的 $||A||_2, ||A||_1||A||_\infty$ 的计算公式以及谱半径 $\rho(A)$ 的定义.

   (j) 什么是方阵的特征值的几何重数？几何重数与代数重数有何关系？

   (k) 什么是矩阵的酉等价？

(l) 两个正规阵的和是否还是正规阵？

(m) 两个正规阵的乘积是否是正规阵？

(n) 酉阵的张量积是否还是酉阵？

(o) 线性空间 V 是两个线性子空间 V_1 和 V_2 的直和 $V = V_1 \oplus V_2$ 是什么意思？

(p) 范数等价定理是什么？

(q) 戈氏圆盘定理所说的内容？

(r) Householder 矩阵的定义？

(s) 矩阵可以相似对角化有什么刻画？其中有一个结果与极小多项式有关，是什么？

2. 设 A 和 B 为 n 阶复方阵，如果 $A^2 = B$，则称 A 为 B 的一个平方根. 假设 $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$，试问 B 是否有平方根？

3. 设 $(V, \langle \cdot, \cdot \rangle)$ 是内积空间，T 是 V 上的线性变换，如果 $\langle Tx, y \rangle = 0$ 对于任意的 $x, y \in V$ 成立，请证明 $T = 0$.

4. 设复数域上的 n 阶方阵 A 满足 $A^3 = 3A^2 + A - 3I$，判断 A 是否可以对角化.

5. 已知复方阵 $A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$，求

(a) 矩阵 A 的 Jordan 标准形.

(b) 矩阵 A 的极小多项式.

6. 已知 $A = J_{2020}(0)$, 求 A^2 的 Jordan 标准形.

7. 设 $A \in \mathbb{C}^{m \times n}, x \in \mathbb{C}^n, y \in \mathbb{C}^m$. 假设 $Ax = 0$. 证明 $x \perp A^+ y$.

8. 已知 $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$，求 e^{At}.

9. 设 $A=\begin{bmatrix}2&0\\2&0\end{bmatrix}$，$B=\begin{bmatrix}5&2\\0&0\end{bmatrix}$，求 Kronecker 积 $A\otimes B$ 与 Hadamard 积 $A\circ B$.

10. 设 $A(t)=\begin{bmatrix}1&0\\\cos t&t\end{bmatrix}, t\in\mathbb{R}$，求 $\frac{\mathrm{d}}{\mathrm{d}t}A(t)$ 和 $\int_0^{\pi}A(t)\mathrm{d}t$.

11. 设 $A^H=A^2$，且 A 是可逆矩阵. 请证明 $A^3=I$.

12. 已知 $A=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，求 e^{At}.

13. 求方程组

$$\begin{cases} x_1+2x_2=1\\ x_3=1\\ 2x_1+4x_2=3\end{cases}$$

的极小范数最小二乘解.

14. 设 $A,B\in\mathbb{C}^{n\times n}$，$A$ 是正定的，B 是 Hermite 的，证明：存在可逆阵 S 和实对角阵 C，使得 $A=SS^H$ 且 $B=SCS^H$.

15. 设 A 为 n 阶复方阵，$A^H+A=2E$，且 $\det A=1$，证明 $A=E$.

16. 设 $A\in\mathbb{C}^{m\times n}, G\in A\{1\}$，且 $\mathrm{rank}(G)=\mathrm{rank}(A)$，证明 $G\in A\{1,2\}$.

17. 设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上的一个自相容的矩阵范数，$A\in\mathbb{C}^{n\times n}$，$m$ 为正整数. 请证明 $\rho(A)\leq\sqrt[m]{\|A^m\|}$.

18. 设 $A=\begin{bmatrix}2022&1&1\\2&520&1\\0&1&1314\end{bmatrix}$，证明 A 有 3 个正的特征值. (戈氏圆盘)

19. 用 Geršgorin 定理证明：矩阵

$$A=\begin{bmatrix}9&1&2&1\\0&8&1&1\\1&0&4&0\\1&0&0&1\end{bmatrix}$$

有至少两个不同的实特征值.

20. 设 $u_1, u_2, \ldots, u_s \in \mathbb{C}^n$ 是 s 个线性无关的向量，$v_1, \ldots, v_t \in \mathbb{C}^m$ 是 t 个线性无关的向量，证明 $u_i \otimes v_j (i = 1, \ldots, s; j = 1, \ldots, t)$ 是 $\mathbb{C}^{nm}$ 中的 st 个线性无关的向量.

21. 用 Geršgorin 定理证明：矩阵

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{3}{2} & i & 0 \\ 0 & \frac{i}{2} & 5 & \frac{i}{2} \\ 1 & 0 & 0 & 5i \end{bmatrix}$$

至少有一个非实的特征值.

一些其它题

1. 设 A 为 n 阶复方阵，证明：存在 $\delta > 0$，使得当 $\epsilon \in \mathbb{C}$ 且 $0 < |\epsilon| < \delta$ 时，$A + \epsilon I$ 是可逆阵.

2. 假设 $A_{n\times n}$ 的元素是由数 0 和数 1 构成的方阵，假设 A 的所有特征值都是正实数，证明 A 的所有特征值都是 1.

3. 设 A 和 B 为 n 阶复方阵且 $AB = BA$，证明 A 和 B 有公共的特征向量.

4. 设矩阵 A 为 3 阶复方阵，α_1, α_2 分别为 A 的属于特征值 $-2, -3$ 的特征向量，且向量 α_3 满足 $A\alpha_3 = 2\alpha_1 + 3\alpha_2 - 2\alpha_3$，证明 $\alpha_1, \alpha_2, \alpha_3$ 线性无关.

5. 设 $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{n\times p}$，证明：

   (a) $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$

   (b) $\mathcal{N}(B) \subseteq \mathcal{N}(AB)$.

7.2 习题提示

1. 习题2: $\det(A) \geq 1$. 利用均值不等式，$n \geq \mathrm{tr}(A) = \frac{1}{n}(\lambda_1 + \cdots + \lambda_n)n \geq n(\lambda_1 \cdots \lambda_n)^{\frac{1}{n}} = n(\det A)^{\frac{1}{n}} \geq n$.

第八章　附录

8.1　集合运算

关于集合的一些直观的基础概念我们是已知的. 我们仅对一部分知识进行复习或学习.

定义 8.1. 设 $A, B, X, Y, \ldots$ 等等都是集合.

- 集合的差 $A \setminus B := \{a \in A | a \notin B\}$. 称 B 在 A 中的补集. 当我们总是考虑集合 X 的子集时，我们也称 X 为全集，子集 E 在 X 的补集 $X \setminus E$ 也记为 E^c.
- 幂集 $\mathcal{P}(X) := \{S | S \subseteq X\}$.
- 卡氏积集 $X \times Y := \{(x, y) | x \in X, y \in Y\}$ 是由所有的有序二元对 (x, y) 其中 $x \in X, y \in Y$ 构成的集合. 类似地，可以（归纳地) 定义 n 个集合的卡氏积 $X_1 \times X_2 \times \cdots \times X_n := (X_1 \times \cdots \times X_{n-1}) \times X_n$. 而 $X \times X$ 可以简记为 X^2.
- 一族集合的交：设 $\{X_i\}_{i \in I}$ 为集合 X 的一族子集合，这里下标集 I 不为空集，则 $\cap_{i \in I} X_i := \{x \in X | \forall i \in I : x \in X_i\}$.
- 一族集合的并：设 $\{Y_j\}_{j \in J}$ 为集合 Y 的一族子集合，则 $\cup_{j \in J} Y_j := \{y | \exists j \in J : y \in Y_j\}$. 注意：如果此时下标集 J 为空集, 则并集也为 (约定为) 空集.

注 8.1. “空真”, “空集拥有一切性质”. 逻辑上，设 E 是某个性质，$E(x)$ 表示 x 有性质 E. 那么我们认为，命题 $[x \in \emptyset \implies E(x)]$ 对于所有 $x \in X$ 总是真的.

例 8.1. 让 $A_x=(x,x+1]\subseteq\mathbb{R}$ 其中 $x\in\mathbb{R}$，那么我们得到一族集合 $\{A_x\}_{x\in\mathbb{R}}$，那么

$$\bigcup_{x\in\mathbb{R}} A_x=\mathbb{R},\bigcap_{x\in\mathbb{R}} A_x=\emptyset.$$

命题 8.1. $X\times Y=\emptyset \iff X=\emptyset$或$Y=\emptyset$.

命题 8.2. 让 $A_i(i\in I)$ 和 $B_j(j\in J)$ 都是 X 的子集. 那么

1. $(\bigcap_{i\in I} A_i)\cap(\bigcap_{j\in J} B_j)=\bigcap_{(i,j)\in I\times J}(A_i\cap B_j)$.
   $(\bigcup_{i\in I} A_i)\cup(\bigcup_{j\in J} B_j)=\bigcup_{(i,j)\in I\times J}(A_i\cup B_j)$. *(结合律)*
2. $(\bigcap_{i\in I} A_i)\cup(\bigcap_{j\in J} B_j)=\bigcap_{(i,j)\in I\times J}(A_i\cup B_j)$.
   $(\bigcup_{i\in I} A_i)\cap(\bigcup_{j\in J} B_j)=\bigcup_{(i,j)\in I\times J}(A_i\cap B_j)$. （分配律）
3. $(\bigcap_{i\in I} A_i)^c=\bigcup_i A_i^c$.
   $(\bigcup_{i\in I} A_i)^c=\bigcap_{i\in I} A_i^c$. *(de Morgan 律)*

映射的定义

定义 8.2. 设 X 和 Y 是两个集合，如果集合 X 中的每个元素 x 按照某种规则 f 与集合 Y 中的唯一一个元素 y 相对应，我们就说有一个从 X 到 Y 的**映射** f. 集合 X 称为映射的**定义域** (映射的出发域)，而 Y 称为映射的**靶域** (也叫作映射的到达域)，X,f,Y 三者一起考虑才构成映射的定义. 用记号 $f:X\to Y,x\mapsto f(x)$ 表示之.

映射，函数，映照等词语经常是表示同一个意思. 我们也不做区分.

注 8.2. 上述定义严格来说并不能称为定义，因为“按照某种规则”这几个字的含义是不清晰的. 事实上，可以先定义关系，再定义*函数关系*/函数. 若 $R\subseteq X\times Y$，则称 R 是 X 到 Y 的一个*关系*. 而 $(x,y)\in R$ 就称为 x 对 y 有关系 R，记为 xRy. 如果关系 R 具有性质：对于每个 $x\in X$，有且只有一个 $y\in Y$，使得 xRy. 那么称 R 为 X 到 Y 的一个函数关系. 通常用 $f:X\to Y$ 来表示函数关系，而 $y=f(x)$ 则代替 xfy.

例 8.2. 记号 1_A 表示集 A 上的**恒等映射**，$1_A:A\to A,a\mapsto a$.

例 8.3. 设 X,Y 为非空集，$b\in Y$，那么 $X\to Y,x\mapsto b$ 称为**常值映射**.

例 8.4. 设 $f:X\to Y$，$A\subseteq X$. 那么 $f|_A:A\to Y,x\mapsto f(x)$ 称为 f 在子集 A 上的**限制映射**.

例 8.5. 设 $A \subseteq X$，$g : A \to Y$，$f : X \to Y$. 如果 $f|_A = g$，称 f 是 g 的一个**扩张/延拓**.

例 8.6. 集合 X，$A \subseteq X$. 映射 $\chi_A : X \to \{0,1\}, x \mapsto \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ 称为 A 的**示性函数**.

定义 8.3. 映射的合成/复合. 设 $f : X \to Y$，$g : Y \to Z$. 那么 $h : X \to Z, x \mapsto g(f(x))$ 称为 g 与 f 的**复合映射**，记为 $g \circ f$.

定理 8.1. *(复合有结合律)* $f : X \to Y, g : Y \to Z, h : Z \to W$，那么 $(h \circ g) \circ f = h \circ (g \circ f)$.

定义 8.4. 映射 $f : X \to Y$. 称 f 为

- 单射，如果：$x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$. 即不同的元素映成不同的像.
- 满射，如果每个 $y \in Y$ 都能被映上：即 $\forall y \in Y, \exists x \in X,$ 使得$y = f(x)$.
- 双射，如果 f 是既单且满的.

定理 8.2. 映射 $f : A \to B$ 为双射 $\iff$ 存在映射 $g : B \to A$ 使得 $f \circ g = 1_B, g \circ f = 1_A$.

上述引理中的 g 是存在且唯一的，称为 f 的**逆**，记为 f^{-1}.

定理 8.3. 设 $f : X \to Y, g : Y \to Z$ 都是双射. 那么 $g \circ f : X \to Z$ 也是双射并且 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

集函数

定义 8.5. 设 $f : X \to Y$ 为映射，$A \subseteq X$，记号 $f(A) := \{f(x)|x \in A\}$ 称为 A 在映射 f 下的像集. 设 $B \subseteq Y$，集合 $f^{-1}(B) := \{x \in X \mid f(x) \in B\}$ 称为集 B 在映射 f 下的原像集.

而集合 $f(X)$ 就直接称为 f 的像集，一般使用记号 $\mathrm{im}(f)$ 表示，不过因为在复数中 im 有特殊的别的含义，我们就不使用 $\mathrm{im}(f)$ 这个记号来表示 f 的像集了.

定义 8.6. 设 $f: X \to Y$ 是映射. 那么我们可以诱导出以下两个映射:

$$f: \mathcal{P}(X) \to \mathcal{P}(Y), A \mapsto f(A) \text{ 和} f^{-1}: \mathcal{P}(Y) \to \mathcal{P}(X), B \mapsto f^{-1}(B).$$

注 8.3. 以上用了似乎不该使用的符号 f, f^{-1} 来表示诱导出来的映射, 但是这个是上下文清晰的，所以不会产生歧义. 当然，使用别的记号是逻辑上更清晰的，例如，用 $\mathrm{Preim}_f(B) := f^{-1}(B), \mathrm{Im}_f(A) := f(A)$ 等.

当 f 是双射时，$\{f^{-1}(y)\} = f^{-1}(\{y\})$ 对所有 y 成立，在这个等式中，显然我们是清楚左右两边中的 f^{-1} 到底是指哪一个.

定理 8.4. *设 $f: X \to Y$. 考虑诱导的集函数 f 和 f^{-1}. 有以下性质:*

1. *若 $A \subseteq B \subseteq X$，则 $f(A) \subseteq f(B)$.*
2. *设 $\forall i \in I$ 都有$A_i \subseteq X$，则 $f(\cup_i A_i) = \cup_i f(A_i)$.*
3. *设 $\forall i \in I$ 都有$A_i \subseteq X$，则 $f(\cap_i A_i) \subseteq \cap_i f(A_i)$.*
4. *设 $A \subseteq X$, 则 $f(A^c) \supseteq f(X) \setminus f(A)$.*
5. *若 $A \subseteq B \subseteq X$，则 $f^{-1}(A) \subseteq f^{-1}(B)$*
6. *设 $\forall i \in I$ 都有$A_i \subseteq X$，则 $f^{-1}(\cup_i A_i) = \cup_i f^{-1}(A_i)$.*
7. *设 $\forall i \in I$ 都有$A_i \subseteq X$，则 $f^{-1}(\cap_i A_i) = \cap_i f^{-1}(A_i)$.*
8. *设 $A \subseteq X$，则 $f^{-1}(A^c) = (f^{-1}(A))^c$.*

如果 $g: Y \to Z$，那么 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

证明. 留给读者. 容易记忆的是：取原像集函数 f^{-1} 具有良好的性质，它可以和交、并、补都交换. □

定义 8.7. 从集合 X 到集合 Y 的映射的全体构成的集合，我们记为 Y^X. 而 $\mathbb{R}^2$ 可以理解为由从 $\{1,2\}$ 到 $\mathbb{R}$ 的所有的函数构成的集合.

关系

定义 8.8. (关系) 集合 X 上的一个关系是指 $X \times X$ 的一个子集 R. 当 $(x,y) \in R$ 时，我们称 x 与/对 y 有关系 R，记 xRy 或者 $x \underset{R}{\sim} y$.

1. 关系 R 称为**自反的**，如果 $\Delta_X = \{(x,x) | x \in X\} \subseteq R$，即 $\forall x \in X, xRx$.

2. 关系 R 称为**传递的**，如果 $xRy \wedge yRz \implies xRz$.

3. 关系 R 称为**对称的**，如果 $xRy \implies yRx$.

定义 8.9. (等价关系) 集合 X 上的一个关系 $\sim$，如果满足以下三条：

1. 自反性：$\forall x \in X, x \sim x$.

2. 对称性：若 $x \sim y$，则 $y \sim x$.

3. 传递性：若 $x \sim y, y \sim z$，则 $x \sim z$.

那么称关系 $\sim$ 是 X 上的一个**等价关系**. 对于每个 $x \in X$，子集 $[x] := \{y \in X | y \sim x\}$ 称为 x 所在的等价类，每一个 $y \in [x]$ 都是该等价类的**代表元**，即，$\forall y \in [x]$ 都有 $[y] = [x]$. 商集

$$X/\sim := \{[x] | x \in X\}$$

称为 X 模掉等价关系 $\sim$ 后得到的商集，它是由所有等价类构成的集合，显然 $X/\sim$ 是幂集 $\mathcal{P}(X)$ 的一个子集.

定义 8.10. 集合 X. 设 $\mathcal{A} \subseteq \mathcal{P}(X) \setminus \{\emptyset\}$ 满足以下性质：$\forall x \in X$，存在唯一一个 $A \in \mathcal{A}$，使得 $x \in A$. 也就是说$\mathcal{A}$ 是由两两不相交的 X 的非空子集构成并且这些子集的并恰好是 X. 那么我们称 $\mathcal{A}$ 是 X 的一个**分划/划分/分割/partition.**

定理 8.5. *设 $\sim$ 为 X 上的一个等价关系. 则 $X/\sim$ 是 X 的一个划分.*

定理 8.6 (* 等价关系与划分一一对应)**.** *集合 X 上的等价关系全体与其上的划分有自然的一一对应关系.*

证明. 只需要证明 $\sim \mapsto X/\sim$ 是从等价关系集到划分集之间的双射即可. 但需要一点点努力. □

例 8.7. 让 X 是此刻在郑州市的人构成的集合. $x \sim y :\iff x, y$ 有相同的爸爸.

例 8.8. $n \times m$ 实矩阵上的矩阵等价？n 阶实方阵上的相似关系？n 阶实对称矩阵上的合同关系？此例中，哪些等价关系诱导出的商集是有限集？

代数基本定理

定理 8.7. （代数基本定理）关于 t 的复系数多项式方程

$$t^n + c_{n-1}t^{n-1} + \cdots + c_0 = 0$$

在复数域中总有根.

推论 8.1. 设 $f(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_0 \in \mathbb{C}[t]$，那么存在 $t_1, \ldots, t_n \in \mathbb{C}$ 使得 $f(t) = (t - t_1) \cdots (t - t_n)$.

8.2 拓扑知识介绍

距离空间

定义 8.11. 设 X 是集合. 函数 $d : X \times X \to \mathbb{R}$ 称为 X 上的一个距离/度量，如果它满足：

1. 正定性 $d(x, y) \geq 0, \quad x, y \in X$，并且等号成立当且仅当 $x = y$.
2. 对称性 $d(x, y) = d(y, x), \quad x, y \in X$.
3. 三角不等式 $d(x, y) \leq d(x, z) + d(z, y), \quad x, y, z \in X$.

若 d 是 X 上的一个距离/度量，则 (X, d) 称为距离/度量空间. 当距离 d 是上下文清晰时，可以直接称 X 为距离空间. 最后，$d(x, y)$ 称为 x 和 y 的距离.

定理 8.8. (X, d) 是距离空间. 那么对于任意 $a, b, c \in X$ 有

$$d(a, b) \geq |d(a, c) - d(c, b)|.$$

定义 8.12. 设 (X, d) 是距离空间，$a \in X$，$r > 0$. 集合

$$\mathbb{B}(a, r) := \mathbb{B}_X(a, r) := \{x \in X | d(a, x) < r\}$$

称为以 a 为球心半径为 r 的开球. 而

$$\tilde{\mathbb{B}}(a, r) := \tilde{\mathbb{B}}_X(a, r) := \{x \in X | d(a, x) \leq r\}$$

称为以 a 为球心半径为 r 的闭球.[1]

[1]闭球 $\tilde{\mathbb{B}}(a, r)$ 我们采取了上方加波浪号表示，大部分文献是使用 $\overline{\mathbb{B}}(a, r)$ 的记号. 因为在更一般的拓扑的意义下集合上方加一杠通常表示闭包，并不能完全和上述定义一致. 为此区分使用记号.

定义 8.13. 设 (X,d) 是距离空间. 集合 $E \subseteq X$，函数 d 在 $E \times E$ 上的限制 d_E 自动是 E 上的一个距离. 距离空间 $(E, d|_E)$ 称为距离空间 X 的子空间. 通常我们也直接将 $d|_E$ 写成 d.

例 8.9. 实数集 $\mathbb{R}$ 上可以通过绝对值函数 $|\cdot|$ 诱导出一个距离: $d(x,y) := |x-y|$ 使得 $\mathbb{R}$ 是距离空间.

例 8.10. 同样的，复数集 $\mathbb{C}$ 也能自然的成为距离空间.

例 8.11. $\mathbb{R}^n$ 上通常的欧式距离 $d(x,y) := \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$，注意验证三角不等式时需要利用 Cauchy 不等式.

例 8.12. （不重要的例子) 设 X 是非空集. 函数

$$d(x,y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

是 X 上的距离.

定义 8.14. (X,d) 为距离空间，$x \in X$，$N \subseteq X$，若存在 $r > 0$ 使得 $B(x,r) \subseteq N$ 就称 N 是 x 的一个**邻域**，称 x 是 N 的一个**内点**. 开球 $B(x,r)$ 称为 x 的一个（开的）$r-$ 邻域.

定义 8.15. (X,d) 距离空间，$E \subseteq X$，若存在 $x_0 \in X$ 和 $r > 0$ 使得 $E \subseteq B(x_0, r)$，则称 E 是**有界集**.

定义 8.16. (X,d) 距离空间. 子集合 $U \subseteq X$ 称为 X 中的一个**开集**，如果 U 中的每个点 x，U 都是 x 的一个邻域. 特别的，$\emptyset$ 和 X 都是开集.

子集合 $F \subseteq X$ 称为**闭集**，如果 F^c 是开集.

定理 8.9. *开球 $\mathbb{B}(a,r)$ 是开集. 闭球 $\tilde{B}(a,r)$ 是闭集.*

定理 8.10. *(开集公理) 开集满足以下开集公理:*

1. *空集和全集是开集.*
2. *开集的有限交是开集.*
3. *开集的任意并是开集.*

于是，距离空间 (X,d) 中 U 是开集当且仅当 U 是一些 (可以是无穷多个) 开球的并.

证明. 根据“开集公理”若 U 是一些开球的并，由于开球是开集，从而 U 是开集. 反之，假设 U 是开集，那么 $\forall x \in U$，存在 $\delta = \delta(x) > 0$，使得 $B_x = B(x, \delta) \subseteq U$，那么 $U = \cup_{x \in U} B_x$. □

相应地，闭集满足以下性质.

定理 8.11. *(闭集公理)*

1. *空集和全集是闭集.*
2. *闭集的有限并是闭集.*
3. *闭集的任意交是闭集.*

拓扑空间

定义 8.17. 设 X 为集合，$\mathcal{T} \subseteq \mathcal{P}(X)$ ，如果 $\mathcal{T}$ 满足：

1. $\emptyset, X \in \mathcal{T}$
2. $U_1, U_2 \in \mathcal{T} \implies U_1 \cap U_2 \in \mathcal{T}$
3. $U_\alpha \in \mathcal{T}, \alpha \in \Lambda \implies \cup_{\alpha \in \Lambda} U_\alpha \in \mathcal{T}$

则称 $\mathcal{T}$ 为 X 上的一个拓扑，$(X, \mathcal{T})$ 称为拓扑空间，$\mathcal{T}$ 中的元素称为开集.

定义 8.18. 设 $(X, \mathcal{T})$ 为拓扑空间.

1. 子集 F 称为**闭集**如果 F^c 是开集.
2. 设 $E \subseteq X$，$x \in E$. 若存在开集 U 使得 $x \in U \subseteq E$，称 E 为 x 的一个**邻域**，x 是 E 的**内点**. 若 E 本身已经是开集，就称为 x 的一个**开邻域**.

定义 8.19. 设 (X, d) 为距离空间，那么在距离空间中定义的“开集”全体形成的集合构成 X 上的拓扑 $\mathcal{T}_d$，拓扑空间 $(X, \mathcal{T}_d)$ 称为由距离 d 诱导出来的拓扑空间.

定义 8.20. (内部，闭包，边界）设 $(X, \mathcal{T})$ 为拓扑空间，$E \subseteq X$.

1. E 的**内部**，记为 E^o 或 $\mathrm{Int}(E)$，是指含于 E 的最大开子集. 即，

$$E^o = \bigcup_{U\subseteq E,\text{且}U\text{是 } X \text{ 中的开集}} U$$

为含于 E 的所有开子集之并. 它是 E 中所有内点构成的子集合.

2. E 的**闭包**，记为 $\overline{E}, E^-, \mathrm{Cl}_X(E), \mathrm{Cl}(E)$，是指包含 E 的最小闭集. 即

$$E^- = \bigcap_{E\subseteq F\text{且 } F \text{ 是闭集}} F.$$

X 中的点 $x \in E^-$ 当且仅当 x 的任何一个（开）邻域 U 与 E 都有非空交.

3. E 的**边界**，记为 $\partial E, \mathrm{Bd}(E),$，定义为 $\partial E := E^- \setminus E^o$. 一个点 $x \in \partial E$ 当且仅当：对于 x 的任何一个 (开) 邻域 U 都有 $U \cap E \neq \emptyset \text{且} U \cap E^c \neq \emptyset$.

证明. 我们来证明 E^- 中的点恰好是满足性质（*）："点 x 的任何一个开邻域 U 与 E 都相交"的点.

首先，$x \notin E^- \iff x \in E^{-c}$，且 E^{-c} 是开集，又 $E^{-c} \cap E = \emptyset$，所以 $x \notin E^- \implies x$ 没有性质 (*).

如果 x 没有性质 (*)，则存在 x 的开邻域 U 使得 $U \cap E = \emptyset$，从而 $E \subseteq U^c$. 因为 U^c 是闭集，所以 $E^- \subseteq U^c$，这说明 $x \in U \subseteq E^{-c}$，$x \notin E^-$.

综上，$x \notin E^- \iff x$ 不满足性质 (*). 所以 $x \in E^-$ 当且仅当 x 满足性质 (*). □

定义 8.21. （子空间拓扑) 设 $(X, \mathcal{T})$ 是拓扑空间, $Y \subseteq X$. 那么我们可以自然的在 Y 上赋予一个拓扑

$$\mathcal{T}_Y := \mathcal{T} \cap Y := \{U \cap Y \,|\, U \in \mathcal{T}\}$$

称为子空间拓扑/限制拓扑. 拓扑空间 $(Y, \mathcal{T}_Y)$ 称为 $(X, \mathcal{T})$ 的子空间. 上下文清晰时，直接说 Y 是 X 的一个 (拓扑) 子空间.

定义 8.22. (连续映射) 设 X, Y 是拓扑空间. 映射 $f : X \to Y$ 称为**连续映射/连续函数**，如果对于任意 Y 中的开集 V，都有 $f^{-1}(V)$ 是 X 中的开集. 即，"开集反射开集".

注 8.4. 一个映射成不成为连续映射必须是依赖于其上赋予拓扑的.

定理 8.12. *设 X, Y 为拓扑空间, 映射 $f : X \to Y$. 则以下等价*

1. *f 是连续的.*
2. *对于 Y 中的任何闭子集 F，有 $f^{-1}(F)$ 是 X 中的闭子集. 即，“闭集反射闭集”.*
3. *对于每个 $x \in X$，$f(x)$ 的每个邻域 V 而言，$f^{-1}(V)$ 是 x 的邻域.*

证明. 我们证明 1 和 3 是等价的.

$1 \implies 3$: 设 $x \in X$，V 是 $f(x)$ 的一个邻域，那么存在 Y 中的开集 W 使得 $f(x) \in W \subseteq V$，所以 $x \in f^{-1}(W) \subseteq f^{-1}(V)$. 而 f 是连续的，所以 $f^{-1}(W)$ 是开的，从而 $f^{-1}(V)$ 是 x 的一个邻域.

$3 \implies 1$: 设 V 是 Y 中的开集，我们来证 $f^{-1}(V)$ 是 X 中的开集. 任取 $x \in f^{-1}(V)$，有 $f(x) \in V$，从而 V 是 $f(x)$ 的邻域，故 $f^{-1}(V)$ 是 x 的邻域. 所以存在 x 的开邻域 U_x 使得 $U_x \subseteq f^{-1}(V)$. 那么 $f^{-1}(V) = \cup_{x \in f^{-1}(V)} U_x$ 是开集. □

注 8.5. 上述定理中最后一条3，如果对某个特定的 $x_0 \in X$ 成立，就说 f 在 x_0 处连续.

即，设 $x_0 \in X$, 我们说 f 在 x_0 处连续是指：对于 $f(x_0)$ 的任意一个邻域 V，有 $f^{-1}(V)$ 是 x_0 的邻域.

那么该条说的是 f 是连续的当且仅当 f 在每一点处都是连续的.

例 8.13. （重要例子）设 X, Y, Z 是拓扑空间.

1. (常值函数是连续函数) 设 $f : X \to Y, x \mapsto y_0$ 其中 y_0 是 Y 中的一个固定点. 那么 f 是连续映射.
2. (连续映射的复合是连续的) 设 $f : X \to Y$ 和 $g : Y \to Z$ 都是连续的，那么 $g \circ f : X \to Z$ 是连续的.
3. 设 $f : X \to Y$ 是连续的. 设 $A \subseteq X$，$f(A) \subseteq B \subseteq Y$ ，那么 $\tilde{f} : A \to B, a \mapsto f(a)$ 是连续的. (此时 A 和 B 分别是 X 和 Y 的子空间.)
4. 设 Y 是 Z 的子空间，$f : X \to Y$ 是连续的. 那么 $g : X \to Z, x \mapsto f(x)$ 也是连续的.

5. 设 $f: X \to Y$，而 $X = \cup_{\alpha\in\Lambda} U_\alpha$ 且每个 U_α 都是 X 中的开集. 那么 f 是连续的当且仅当: 对所有 $\alpha \in \Lambda$，限制映射 $f_\alpha : U_\alpha \to Y, x \mapsto f(x)$ 都是连续的.

证明. 我们证明第三条. 设 U 是 Y 中开集，那么 $\tilde{f}^{-1}(U\cap B) = f^{-1}(U\cap B)\cap A = f^{-1}(U)\cap f^{-1}(B)\cap A = f^{-1}(U)\cap A$. 由于 $f: X \to Y$ 是连续的，所以 $f^{-1}(U)$ 是 X 中的开集，$f^{-1}(U)\cap A$ 是 A 中的开集. 于是 $\tilde{f}: A \to B, a \mapsto f(a)$ 是连续的.

证明第五条的 ($\Longleftarrow$) 部分. 任取 Y 中的开集 U，注意到 $f^{-1}(U)\cap U_\alpha = f_\alpha^{-1}(U)$，如果 f_α 是连续的，则 $f_\alpha^{-1}(U)$ 是 U_α 中的开集. 由于 U_α 是 X 的开集，从而 $f_\alpha^{-1}(U)$ 是 X 中的开集 (?). 而 $f^{-1}(U) = \cup_\alpha (f^{-1}(U)\cap U_\alpha)$，因此 $f^{-1}(U)$ 是 X 中的开集. 这就证明了 f 是连续的.

我们经常的也使用同一个符号 f，此时指明定义域和靶域来表达这个映射 $f: A \to B$. □

现在我们回到距离空间.

定义 8.23. 让 (X, d_X) 和 (Y, d_Y) 是距离空间，$x_0 \in X$，$f: X \to Y$，我们说 f 在 x_0 处连续，如果：$\forall \epsilon > 0, \exists \delta > 0$, 当 $d_X(x, x_0) < \delta$ 时有 $d_Y(f(x), f(x_0)) < \epsilon$.

可以验证，该定义与拓扑空间之间的映射在一点处连续的定义一致. 见注8.5.

证明. 让 $x_0 \in X$，记 $y_0 = f(x_0)$.

设 f 在 x_0 处按 $\epsilon - \delta$ 定义是连续的. 那么对于任意 y_0 的一个邻域 V，存在 $\epsilon > 0$，使得 $B(y_0, \epsilon) \subseteq V$. 根据 $\epsilon - \delta$ 连续的定义，存在 $\delta > 0$，使得当 $x \in B(x_0, \delta)$ 时，有 $f(x) \in B(y_0, \epsilon) \subseteq V$. 所以 $B(x_0, \delta) \subseteq f^{-1}(V)$，$f^{-1}(V)$ 是 x_0 的一个邻域.

反之，假设 f 在 x_0 是按照"反射邻域的方式"定义连续的. 那么 $\forall \epsilon > 0$，对于 y_0 的邻域 $V = B(y_0, \epsilon)$ 而言，$f^{-1}(V)$ 是 x_0 的一个邻域，故存在 $\delta > 0$，使得 $B(x_0, \delta) \subseteq f^{-1}(V)$. 这恰好就是当 $x \in B(x_0, \delta)$ 时，$f(x) \in B(y_0, \epsilon)$ 的意思. □

对于距离空间之间的连续映射，可以有 $\epsilon - \delta$ 的定义方式.

定理 8.13. 设 (X, d_X) 和 (Y, d_Y) 是距离空间，$f: X \to Y$. 那么 f 是连续的当且仅当 f 在每个点 x_0 处都是连续的. 即，$\forall x_0 \in X, \forall \epsilon > 0, \exists \delta > 0$ 使当 $d_X(x, x_0) < \delta$ 时有 $d_Y(f(x), f(x_0)) < \epsilon$.

注意: 一般说来上述 δ 是与 x_0 以及 ϵ 有关.

8.3 极限

序列极限

定义 8.24. 假设 (X,d) 是距离空间, $a:\mathbb{N}\to X, n\mapsto a_n$ 称为 X 中的一个序列/点列. 序列 a 的第 n 项 (标号为 n) 的值是 a_n，通常也直接说 a_n 是该序列的第 n 项. 序列的记号可以表示为 $\{a_n\}_{n=1}^{\infty}$. 我们说序列 $\{a_n\}_{n=1}^{\infty}$ 有极限 $a_0\in X$，如果 $\forall\epsilon>0$，存在 $N\in\mathbb{N}$，当 $n\geq N$ 时，有 $d(a_n,a_0)<\epsilon$. 记为 $\lim_{n\to\infty}a_n=a_0$. 有极限的序列称为收敛序列，没有极限的序列称为发散序列.

显然，$\lim a_n=a_0 \iff \lim d(a_n,a_0)=0$. (后者是实数列的收敛.)

显然，极限如果存在，那么必然唯一.

注 8.6. 序列也可以从某个整数开始，用 $\mathbb{Z}_{\geq k}$ 表示大于等于 k 的整数，$a:\mathbb{Z}_{\geq k}\to X, n\mapsto a_n$ 也称为 X 中的序列，记号当然可以用 $\{a_n\}_{n=k}^{\infty}$.

显然，改变一个序列的有限多项的值并不影响该序列是否收敛以及不影响序列的极限值 (如果收敛的话). 所以当我们只关心收敛性或者极限值时，可以直接省略的写序列记号 $\{a_n\}$.

注 8.7. 注意，有没有极限不单和序列本身有关，和序列所在的空间也有关. 另外，序列的记号 $\{a_n\}_{n=1}^{\infty}$ 与 $\{a_n|n\in\mathbb{N}\}$ 所表示的内涵略有不同，前者仅仅是“函数”的一种约定的简写而后者则表示函数 a 的像集，例如对于常值序列后者则是一个单点集. 显然，我们觉得这种记号多少是有点点歧义的。为了方便，我们有时也用记号 $\{a_n\}_{n=1}^{\infty}\subseteq X$ 表示，a 是 X 中的一个序列，这种写法，一方面 $\{a_n\}_{n=1}^{\infty}$ 既理解为序列，另一方面又理解为映射 a 的像集 (从而表示取值在 X 中). 不管怎样，这只是记号上的“方便”.

例 8.14. $a_n=\frac{1}{n}, n\in\mathbb{N}$，那么如果将 $\{a_n\}$ 看成是 $\mathbb{R}$ 中的序列/数列，则在 $\mathbb{R}$ 中有极限 $a_0=0$，$\lim a_n=0$. 如果将 $\{a_n\}$ 看成是空间 $(0,2)$ 中的序列，那么它没有极限，因为 $0\notin(0,2)$.

定理 8.14. *设 $f:X\to Y$ 是距离空间之间的映射, $x_0\in X$. 则以下等价:*

1. *f 在 x_0 处连续,*
2. *对于任意的一个收敛于 x_0 的序列 $\{x_n\}$ 都有 $\lim f(x_n)=f(x_0)$.*

证明. 我们证 $2 \implies 1$. 如果 f 在 x_0 处不连续，那么存在 $\epsilon > 0$，对于 $\forall n \in \mathbb{N}$, 考虑 $\delta_n = \frac{1}{n}$，则存在 $x_n \in B(x_0, \delta_n)$ 使得 $f(x_n) \notin B(f(x_0), \epsilon)$. 此时 $\lim x_n = x_0$，但是 $\{f(x_n)\}$ 不收敛于 $f(x_0)$. □

现在我们考虑复数 $\mathbb{C}$ 中的数列收敛问题，此时，$\mathbb{C}$ 中的两点 z, w 的距离为 $|z - w|$.

定理 8.15. *考虑复数列* $\{z_n = x_n + iy_n\}_{n=1}^{\infty}$，$z_0 = x_0 + iy_0 \in \mathbb{C}$，*那么* $\lim z_n = z_0$ *当且仅当* $\lim x_n = x_0$*且* $\lim y_n = y_0$.

证明. 对于 $z = x + iy \in \mathbb{C}$，我们有不等式

$$\frac{|x| + |y|}{\sqrt{2}} \le |z| \le |x| + y|.$$

于是

$$\frac{|x_n - x_0| + |y_n - y_0|}{\sqrt{2}} \le |z_n - z_0| \le |x_n - x_0| + |y_n - y_0|,$$

从这个不等式，容易得到结论. □

类似于实数列中的 Cauchy 列，我们也可以在一般的距离空间上定义 Caucy 序列.

定义 8.25. 设 (X, d) 是距离空间，$\{a_n\}_{n=1}^{\infty}$ 是 X 中的序列，如果它满足：$\forall \epsilon > 0$，存在 $N \in \mathbb{N}$，当 $n, m \ge N$ 时，有 $d(a_n, a_m) < \epsilon$. 则称 $\{a_n\}_{n=1}^{\infty}$ 是 Cauchy 列.

类似于实数列，我们也有复数列的 Cauchy 收敛定理.

定理 8.16. *(Cauchy 收敛准则) 设有复数列* $\{z_n\}_{n=1}^{\infty}$，*那么* $\{z_n\}$ *在* $\mathbb{C}$ *中收敛当且仅当它是 Cauchy 列.*

证明. 显然，任何距离空间中的收敛序列都是 Cauchy 列. 我们下面证明对于距离空间 $\mathbb{C}$ 而言，Cauchy 列必然也收敛. 由不等式

$$|x_n - x_m| \le |z_n - z_m|, \ |y_n - y_m| \le |z_n - z_m|$$

知道如果 $\{z_n\}_{n=1}^{\infty}$ 是 Cauchy 列，那么 $\{z_n\}_{n=1}^{\infty}$ 的实部数列 $\{x_n\}_{n=1}^{\infty}$ 和虚部数列 $\{y_n\}_{n=1}^{\infty}$ 是 $\mathbb{R}$ 中的 Cauchy 列，而 $\mathbb{R}$ 中的 Cauchy 列是收敛的，于是存在 $x_0 \in \mathbb{R}$ 和 $y_0 \in \mathbb{R}$，使得 $\lim |x_n - x_0| = \lim |y_n - y_0| = 0$，从而 $\{z_n\}_{n=1}^{\infty}$ 收敛且 $\lim z_n = x_0 + iy_0$. □

和收敛实数列的运算类似, 收敛的复数列也有对应的性质.

定理 8.17. *设复数列 $\{z_n\}$ 收敛于 z_0，$\{w_n\}$ 收敛于 w_0. 那么*

1. $\lim(z_n + w_n) = z_0 + w_0$
2. $\lim z_n w_n = z_0 w_0$
3. *若* $w_0 \neq 0$，*则* $\lim \frac{1}{w_n} = \frac{1}{w_0}$.
4. $\lim |z_n| = |z_0|$.

证明. 证明直接模仿实数列的情形即可. 我们这里指出当 $w_0 \neq 0$ 时，需要说明对于充分大的 n, $w_n \neq 0$, 而 $\lim \frac{1}{w_n}$ 理解为从足够大的 n 以后构成的序列 $\{\frac{1}{w_n}\}$ 的极限.

(如果，极限不是在复数集上考虑，而是在扩充的复平面 $(\hat{\mathbb{C}}, \sigma)$ 上考虑的话，分母不为零的要求可以去掉.) □

经过一点点努力，定理8.15与定理8.16 对于任意 k 维欧式空间 $X = \mathbb{R}^k$ 有类似的结论.

设 $X = \mathbb{R}^k$,其上的距离函数 d 按照通常的约定:即 $x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k) \in \mathbb{R}^k$, 距离 $d(x, y) = \sqrt{\sum_{j=1}^{k}(x_j - y_j)^2}$. 由于 $\mathbb{R}^k$ 中的元素 x 按照通常约定其第 j 分量用 x_j 表示，已经使用过了下标，所以我们用上标形式 $\{x^{(n)}\}$ 来表示 X 中的序列，其中 $x^{(n)} = (x_1^{(n)}, \ldots, x_k^{(n)}) \in \mathbb{R}^k, n \in \mathbb{N}$. 当然如果有更好的记号也是可以的.

定理 8.18. *设 $\{p^{(n)}\}$ 为 $X = \mathbb{R}^n$ 中的一个序列，$p^{(0)} \in X$，那么 $\lim p^{(n)} = p^{(0)}$ 当且仅当 $\lim p_j^{(n)} = p_j^{(0)}, j = 1, 2, \ldots, k$. 即 $\mathbb{R}^k$ 中的序列收敛于某个点当且仅当其分量序列收敛于该点对应的分量.*

定理 8.19. *设 $\{p^{(n)}\}$ 为 $X = \mathbb{R}^n$ 中的一个序列，那么 $\{p^{(n)}\}$ 收敛当且仅它是 Cauchy 列.*

以上两个定理都可以通过以下不等式获得证明: 设 $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$，那么

$$\frac{|x_1| + \cdots + |x_k|}{\sqrt{k}} \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2} \leq |x_1| + |x_2| + \cdots + |x_k|.$$

其中 $\frac{|x_1|+\cdots+|x_k|}{\sqrt{k}} \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2}$ 可以通过 Cauchy-Schwartz 不等式得到. 当然不需要 $\frac{|x_1|+\cdots+|x_k|}{\sqrt{k}} \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2}$ 也可以，只要利

用 $|x_j| \leq \sqrt{x_1^2+x_2^2+\cdots+x_k^2}, j=1,2,\ldots,k$ 以及 $\sqrt{x_1^2+x_2^2+\cdots+x_k^2} \leq |x_1|+|x_2|+\cdots+|x_k|$ 即可模仿之前的证明得到相应的证明.

注 8.8. (小知识) 若距离空间 (X,d) 的任何 Cauchy 列都收敛，那么称这个距离空间是**完备**的距离空间.

所以 $\mathbb{R}^k$ 以及复数 $\mathbb{C}$ 是完备的距离空间. 而 $(0,1)$ 作为距离空间 $\mathbb{R}$ 的子空间，$(0,1)$ 不是完备的距离空间.

对于距离空间 (X,d) 中子集 E 的闭包 E^- 可以利用序列极限来刻画:

定理 8.20. *设 (X,d) 是距离空间，$E \subseteq X$，那么 X 中的点 $a \in E^-$ 当且仅当存在 E 中的序列 $\{a_n\}$ 使得 $\lim a_n = a$.*

证明. 假设 $\{a_n\} \subseteq E$ 且 $\lim a_n = a$，根据极限的定义，a 的任何邻域总包含某 a_n，于是 $a \in E^-$.

现在假设 $a \in E^-$，那么 $\forall n \in \mathbb{N}$，$B(a,1/n)\cap E \neq \emptyset$. 对于每个 n，我们可以从 E 中选择一个元素 $a_n \in B(a,1/n)\cap E$. 则 $\lim a_n = a$. □

级数

定义 8.26. (级数) 设 $\{z_n\}_{n=0}^\infty$ 是复数列,我们构造部分和数列 $S_n=\sum_{k=0}^n z_k, n\in \mathbb{Z}_{\geq 0}$. 如果 $\{S_n\}$ 收敛于 S，我们说“级数 $\sum_{n=0}^\infty z_n$”是收敛的，其和为 S，记为 $\sum_{n=0}^\infty z_n = S$. 如果 $\{S_n\}$ 是发散的，就说“级数 $\sum_{n=0}^\infty z_n$”是发散的.

和复数列收敛的定理类似，级数也有对应的收敛定理.

定理 8.21. *复数项级数 $\sum_{n=0}^\infty z_n = a+bi$ 当且仅当 $\sum_{n=0}^\infty x_n = a$且$\sum_{n=0}^\infty y_n = b$.*

定理 8.22. *(Cauchy 收敛原理) 复数项级数 $\sum z_n$ 收敛，当且仅当它满足：$\forall \epsilon>0, \exists N,$ 当 $m>n\geq N$ 时，$|\sum_n^m z_k|<\epsilon$.*

定义 8.27. (绝对收敛级数) 复数项级数 $\sum z_n$ 称为**绝对收敛的**，如果 $\sum |z_n|$ 是收敛的.

定理 8.23. *绝对收敛的复数项级数是收敛的. 若 $\sum |z_n|$ 收敛，那么 $\sum z_n$ 收敛，且 $|\sum z_n| \leq \sum |z_n|$.*

证明. (证明一) 考虑实部和虚部，利用实数项级数的性质.

由于 $\sum |x_n| \leq \sum |z_n|, \quad \sum |y_n| \leq \sum |z_n|$,于是 $\sum x_n$ 与 $\sum y_n$ 都是绝对收敛的实数项级数,从而它们是收敛的. 于是 $\sum z_n$ 收敛. 显然,$|\sum_{n=0}^m z_n| \leq \sum_{n=0}^m |z_n| \leq \sum_{n=0}^\infty |z_n|$，故 $|\sum z_n| \leq \sum |z_n|$.

(证明二) 利用 Cauchy 收敛定理. 注意到 $|\sum_{k=n}^{m} z_k| \leq \sum_{k=n}^{m} |z_k|$ 对于任意 $m > n$ 成立. 由于 $\sum |z_n|$ 收敛，所以 $\forall \epsilon > 0$，$\exists N$，当 $m > n \geq N$ 时，$|\sum_{k=n}^{m} z_k| \leq \sum_{k=n}^{m} |z_k| < \epsilon$. 故 $\sum z_n$ 收敛. □

定义 8.28. 设有级数 $\sum_{n=1}^{\infty} z_n$. 有一个双射 $k : \mathbb{N} \to \mathbb{N}, n \mapsto k_n$. 那么级数 $\sum_{n=1}^{\infty} z_n$ 按 k 得到的重排级数为 $\sum_{n=1}^{\infty} z_{k_n}$，称级数 $\sum_{n=1}^{\infty} z_{k_n}$ 是级数 $\sum_{n=1}^{\infty} z_n$ 的**一个重排**. 通俗的说，就是对级数的项按照一种新顺序重新排列而得到的新级数称为一个重排级数.

定理 8.24. *（重排定理）设复数项级数 $\sum_{n=1}^{\infty} z_n$ 绝对收敛其和为 s，那么任一重排后的级数 $\sum_{n=1}^{\infty} z_{k_n}$ 也是绝对收敛的，并且和也是 s.*

证明. 对于任取定的 $m \in \mathbb{N}$，我们取 $M \in \mathbb{N}$ 使得 $k_1, k_2, \ldots, k_m \leq M$.

那么 $\sum_{n=1}^{m} |z_{k_n}| \leq \sum_{n=1}^{M} |z_n| \leq \sum_{n=1}^{\infty} |z_n|$，所以 $\sum_{n=1}^{\infty} z_{k_n}$ 是绝对收敛的，从而 $\sum_{n=1}^{\infty} z_{k_n}$ 也是收敛的.

任一取定的 $\epsilon > 0$，由于 $\sum_{n=1}^{\infty} z_n$ 是绝对收敛的，故存在 N，使得 $\sum_{n=N}^{\infty} |z_n| \leq \epsilon$. 取 N' 使得 $j \geq N'$ 时有 $\{1, 2, \ldots, N\} \subseteq \{k_1, k_2, \ldots, k_j\}$，并且 $N' \geq N$. 那么当 $j \geq N'$ 时，$|\sum_{n=1}^{j} z_n - \sum_{n=1}^{j} z_{k_n}| \leq \sum_{n=N}^{\infty} |z_n| \leq \epsilon$. 让 $j \to \infty$，得到 $|s - \sum_{n=1}^{\infty} z_{k_n}| \leq \epsilon$，该式对任意 $\epsilon > 0$ 成立，于是 $\sum_{n=1}^{\infty} z_{k_n} = s$. □

定理 8.25. *(级数乘积定理) 设复数项 $\sum_{n=0}^{\infty} z_n$ 和 $\sum_{n=0}^{\infty} w_n$ 都是绝对收敛的，其和分别为 A 和 B. 以任一种方式遍历 $(n, m) \in \{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots\}$ 得到的级数 $\sum z_n w_m$ 是绝对收敛的，和为 AB. 特别的，有 $\sum_{n=0}^{\infty} (\sum_{k=0}^{n} z_n w_{n-k}) = AB$.*

证明. 设 $\sigma : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}, n \mapsto (\sigma_1(n), \sigma_2(n))$ 是双射. 我们需要证明 $\sum_{n=0}^{\infty} z_{\sigma_1(n)} w_{\sigma_2(n)}$ 是绝对收敛的并且收敛的和是 AB.

事实上对于 $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ 中的任何非空有限子集 F，总可以找到 N，使得 $F \subseteq \{0, 1, \ldots, N\} \times \{0, 1, \ldots, N\}$，那么

$$\sum_{(n,m) \in F} |z_n w_m| \leq \left(\sum_{0}^{N} |z_n| \right) \left(\sum_{0}^{N} |w_m| \right) \leq \left(\sum_{n=0}^{\infty} |z_n| \right) \left(\sum_{m=0}^{\infty} |w_m| \right),$$

于是 $\sum_{n=0}^{\infty} z_{\sigma_1(n)} w_{\sigma_2(n)}$ 是绝对收敛的，由重排定理，$\sum_{n=0}^{\infty} z_{\sigma_1(n)} w_{\sigma_2(n)}$ 的任和一个重排级数都是绝对收敛的且收敛于共同的极限.

现在考虑 $\sum_{n=0}^{\infty} z_{\sigma_1(n)} w_{\sigma_2(n)}$ 的一个特殊的重排级数：

$$
\begin{aligned}
&\ z_0w_0\\
&+z_0w_1+z_1w_1+z_1w_0\\
&\ \vdots\\
&+z_0w_n+z_1w_n+\cdots+z_nw_n+z_nw_{n-1}+\cdots+z_nw_0\\
&+\cdots
\end{aligned}
$$

此级数的部分和序列存在一个子列 $z_0w_0, (z_0+z_1)(w_0+w_1), \ldots, (\sum_{n=0}^{m} z_n)(\sum_{n=0}^{m} w_n), \ldots$ 该子列收敛于 AB. 那么级数 $\sum_{n=0}^{\infty} z_{\sigma_1(n)} w_{\sigma_2(n)} = AB$.

再考虑另外一种特殊的重排级数：

$$
\begin{aligned}
&\ z_0w_0\\
&+z_0w_1+z_1w_0\\
&+z_0w_2+z_1w_1+z_2w_0\\
&\ \vdots\\
&+z_0w_n+z_1w_{n-1}+\cdots+z_nw_0\\
&+\cdots
\end{aligned}
$$

那么该级数的和为 AB，而 $\sum_{n=0}^{\infty}(\sum_{k=0}^{n} z_n w_{n-k})$ 是上述级数添加括号后得到的级数，故也收敛于 AB. □

注 8.9. 级数乘积定理的这个证明中，想法是比较直接的，只是书写起来略微会麻烦一些. 也有另外的处理方式，我们这里不展示了.

函数极限

函数极限的一般理论深入思考稍微会有点复杂. 实际上有序列极限就足够了，不过我们在高数中会学习到函数极限的概念，其中又包含各种情形，单边极限、二元函数沿某直线趋近于原点的极限等等的说法. 很有必要对这些概念进行统一的处理.

函数极限理论的一个显然的用处是用来描述连续性，让我们尝试回忆一下连续的概念，从而引入函数极限的概念. 假设 $f : X \to Y$ 是距离空间之间的映射，$x_0 \in X, y_0 = f(x_0)$，我们说 f 在 x_0 处连续是指：对于 y_0 的任何一个邻域 V，可以找到 x_0 的一个邻域 U，使得 $f(U) \subseteq V$.

如果 x_0 是所谓的孤立点，即单点集 $\{x_0\}$ 本身就是一个开集，那么 f 在 x_0 处自动连续. 例如，$X=\mathbb{Z}$ 是距离空间 $\mathbb{R}$ 的子空间，那么任何一个映射 $f:\mathbb{Z}\to Y$ 都是连续映射. 所以考虑连续性，我们需要关心的是那种叫做“极限点/聚点”的点：也就是单点集 $\{x_0\}$ 不是开集的点 x_0，换句话说，x_0 的任何一个邻域都含有异于 x_0 的点. 对于这样的点，函数在这点处连续与否，和函数在这点 x_0 的“不包括 x_0 点的附近”的函数值以及 y_0 的关系有关.

另外一方面，有时候，我们预先没有知道函数 f 在一点 x_0 处的定义，需要补充在这一点处的函数值 $f(x_0)$ 使得新的函数是原来函数的“一个连续延拓”，此时也需要我们关心函数在“不包括 x_0 点的附近”的性态.

为此，我们需要给出“极限点”的概念.

定义 8.29. 假设 (X,d) 是一个距离空间，E 是 X 的子集，点 $p\in X$ 称为 E 的一个**极限点/聚点**，如果 p 的每个邻域 U 都包含 E 中异于 p 的点 (即 $(U\setminus\{p\})\cap E=U\cap E\setminus\{p\}\neq\emptyset$).

例 8.15. 假设 G 是 $\mathbb{C}$ 中的非空开集，$p\in G$. 那么 p 是 G 的极限点，p 是 $G\setminus\{p\}$ 的极限点.

定理 8.26. *设 (X,d) 是一个距离空间，E 是 X 的子集，点 $p\in X$ 为 E 的极限点，那么 p 点的任何一个邻域都包含 E 中的无穷多个点.*

证明. 留做练习. □

定义 8.30. 设 (X,d) 是距离空间，$p\in X$，如果单点集 $\{p\}$ 是开集，那么称点 p 是**孤立点**.

注 8.10. 显然，定义8.29和8.30可以在一般的拓扑空间上做定义，但对于我们距离空间已经足够用了. 另外，定理8.26只适用于距离空间而对于一般的拓扑空间显然不成立.

定义 8.31. (函数极限) 假设X,Y 是距离空间，$E\subseteq X$，$f:E\to Y$，并且 $p\in X$ 是 E 的极限点. 如果存在 $q\in Y$ 满足：对于任意的 $\epsilon>0$，存在 $\delta>0$，使当 $x\in E$ 且 $0<d_X(x,p)<\delta$ 时，有 $d_Y(f(x),q)<\epsilon$. 称 q 是当 x 在 E 上趋近于 p 时 $f(x)$ 的极限. 记成

$$\lim_{x\to p}f(x)=q$$

或者更明确地写成

$$\lim_{x\to p, x\in E} f(x) = q.$$

注 8.11. 上述定义中“$\epsilon - \delta$ 语言部分”可以等价的表述为：对于 q 的任一给定邻域 V_q，$f^{-1}(V_q)$ 将包含 p 的一个去心邻域 (即存在 p 的邻域 U_p 使得 $f(U_p \setminus \{p\}) \subseteq V_q$.)

推论 8.2. *沿用定义8.31的记号，假设 $F \subseteq E$ 并且 p 是 F 的极限点，假设 $\lim_{x\to p, x\in E} f(x) = q$ 存在，那么 $\lim_{x\to p, x\in F} f(x)$ 也存在且极限值为 q. 极限 $\lim_{x\to p, x\in F} f(x)$ 称为 x 在 F 上趋近于 p 时 $f(x)$ 的极限.*

定理 8.27. *假设 X, Y 是距离空间，$E \subseteq X$，$f : E \to Y$，并且点 $p \in E$. 如果 p 是 E 的极限点，那么 f 在 p 点处连续当且仅当 $\lim_{x\to p} f(x) = f(p)$.*

证明. 显然. □

例 8.16. 假设 $f : \mathbb{R} \to \mathbb{R}$，那么 $\lim_{x\to 0} f(x)$ 的意思是通常的极限的意思. 而单侧极限 $\lim_{x\to 0^+} f(x)$ 的含义就是将 f 定义域限制在 $(0, +\infty)$ 后再看 x 趋近 0 这点的极限，也就是

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0, x>0} f(x).$$

例 8.17. 如果 $X = \hat{\mathbb{C}}$ 为 Riemann 球面，$E = \mathbb{N} \subseteq X$，那么 ∞ 是 E 的极限点，复数列 $\{x_n\}_{n=1}^{\infty}$ 的极限问题可以看成函数 $x : \mathbb{N} \to \mathbb{C}, n \mapsto x_n$ 在 $n \to \infty$ 时的极限问题.

例 8.18. 函数 $f : \mathbb{C} \to \mathbb{C}, z \mapsto z$ 是连续的.

定义 8.32. 假设 f, g 都是定义在某集合 E 上的复值函数，那么函数的和 $f + g$ 定义为 $f + g : E \to \mathbb{C}, p \mapsto f(p) + g(p)$. 类似的，可以定义函数的差 $f - g$，函数乘积 fg. 而函数的商 f/g 可以定义在 $g \neq 0$ 的点集上. 也可以定义 $\mathrm{Re}(f) : E \to \mathbb{R}, p \mapsto \mathrm{Re}(f(p))$ 等等.

定理 8.28. *设 (X, d) 是距离空间，$E \subseteq X$，$p_0 \in X$ 是 E 的极限点，$f : E \to \mathbb{C}, p \mapsto u(p) + iv(p)$，设 $A = u_0 + iv_0 \in \mathbb{C}$. 那么*

$$\lim_{p\to p_0} f(p) = A \iff \lim_{p\to p_0} u(p) = u_0 \text{ 且} \lim_{p\to p_0} v(p) = v_0.$$

证明. 注意到 $|u(p) - u_0|, |v(p) - v_0| \leq |f(p) - A| \leq |u(p) - u_0| + |v(p) - v_0|$，剩余证明是容易的. □

定理 8.29. 设 (X,d) 是距离空间，$E \subseteq X$，$p \in X$ 是 E 的极限点，而 f,g 都是 E 上的复值函数，且

$$\lim_{x\to p} f(x) = A, \quad \lim_{x\to p} g(x) = B.$$

那么

1. $\lim_{x\to p}(f+g)(x) = A+B$
2. $\lim_{x\to p}(fg)(x) = AB$
3. $\lim_{x\to p}(\frac{f}{g})(x) = \frac{A}{B}$，如果 $B \neq 0$.

定理 8.30. 设 X 是距离空间，$f,g : X \to \mathbb{C}$，$p \in X$，并且 f,g 在 p 点处连续，那么 $f+g, f-g, fg$ 都在 p 点处连续. 如果 $g(p) \neq 0$，那么 f/g 在 p 点处也连续. (其中 f/g 是定义在 $g \neq 0$ 上的函数).

证明. 可以利用函数极限来证明，也可以直接根据连续性的定义来证明. 证明略. □

所以复系数多项式函数 $w = P(z) = a_0 + a_1 z + \cdots + a_n z^n$ 是 $\mathbb{C}$ 到 $\mathbb{C}$ 上的连续函数. 而有理函数 $w = \frac{P(z)}{Q(z)}$，其中 P,Q 是复系数多项式，在复平面内分母不为零的点集上是连续的.

我们后面主要考虑的都是复变量复值函数，所以后面我们考虑的距离空间主要是 $\mathbb{C}$ 或者 $\hat{\mathbb{C}}$. 当考虑的距离空间是 $\hat{\mathbb{C}}$ 时，有必要对 $\hat{\mathbb{C}}$ 的拓扑进行研究.

定理 8.31. 距离空间 $\hat{\mathbb{C}}$ 中 U 是 $\hat{\mathbb{C}}$ 的开子集当且仅当

1. 或者 $U \subseteq \mathbb{C}$，此时 U 是 $\mathbb{C}$ 中的开集.
2. 或者 $\infty \in U$，此时 $\hat{\mathbb{C}} \setminus U$ 是 $\mathbb{C}$ 中的有界闭集.

证明. 严格的证明并不难. 但是这里不给出了. 因为 $\hat{\mathbb{C}}$ 的拓扑是直接从 $\mathbb{R}^3$ 中的单位球面上的拓扑通过球极投影搬运过来的，从几何直观上可以得到直接的认识. 例如，$\{z \in \mathbb{C} \| z| > \delta\} \cup \{\infty\}$ 将会是 ∞ 的一个邻域，而 $\{z \in \mathbb{C} \| z| > \delta\}$ 是 ∞ 的一个去心邻域. □

这个定理也有另外一种等价的表述：

定理 8.32. 距离空间 $\hat{\mathbb{C}}$ 中 U 是 $\hat{\mathbb{C}}$ 的开子集当且仅当

1. $U \cap \mathbb{C}$ 是 $\mathbb{C}$ 中的开集.

2. 若 $\infty \in U$，则存在 $r > 0$ 使得 $\{z \in \mathbb{C} | |z| > r\} \subseteq U$.

于是复变函数的极限定义 $\lim_{z \to z_0} f(z) = A$ 可以涉及到 z_0 和 A 都是有限的复数、z_0 和 A 之中至少有一个是 ∞ 的情况. 我们下面直接给出部分情况的定义.

定义 8.33. (复变函数的函数极限的 $\epsilon - \delta$ 语言定义) $\lim_{z \to z_0} f(z) = A$

1. 在 z_0 和 A 都是有限的复数时，是指：$\forall \epsilon > 0, \exists \delta > 0$，当 $0 < |z - z_0| < \delta$ 时，有 $|f(z) - A| < \epsilon$.

2. $z_0 = \infty$，$A \in \mathbb{C}$ 时，是指：$\forall \epsilon > 0$，$\exists \delta > 0$，当 $|z| > \delta$ 时 (且 $z \neq \infty$)，有 $|f(z) - A| < \epsilon$.

3. $z_0 \in \mathbb{C}, A = \infty$ 时，是指：$\forall \epsilon > 0$，$\exists \delta > 0$，当 $0 < |z - z_0| < \delta$ 时，有 $|f(z)| > \epsilon$.
   此时，也可以等价于 $\lim_{z \to z_0} |f(z)| = \infty$.

4. $z_0 = \infty, A = \infty$ 时，是指：$\forall \epsilon > 0$，$\exists \delta > 0$，当 $|z| > \delta$ 时 (且 $z \neq \infty$)，$|f(z)| > \epsilon$.
   此时，也可以等价于 $\lim_{z \to \infty} |f(z)| = \infty$.

无论哪种情形，都要求 z_0 是 f 定义域 E 的极限点.

注 8.12. 定义8.33是通常的复变量复值函数中函数极限的定义，简单的说，$\lim_{z \to z_0} f(z) = A$，是指 A 的任何一邻域 V，总可以找到 z_0 的一个去心邻域 $U \setminus \{z_0\}$，使得 $f^{-1}(V)$ 包含该去心邻域. 粗糙的说就是，无论与 A 多接近的误差 V，当自变量 z 与 z_0 足够接近 (但不取 z_0) 时，函数值 $f(z)$ 与 A 的“接近度”会落在预先给定的误差 V. 而怎么定义“接近度”其抽象的公理性质恰好就是“拓扑”的实质内涵.

参考文献

[1] 徐仲等编著. 矩阵论简明教程（第三版）[M]. 北京：科学出版社，2014.

[2] 王鄂芳，石生明. 高等代数（第四版）[M]. 北京：高等教育出版社，2013.

[3] Horn R A, Johnson C R. Matrix analysis [M]. 2nd ed. Cambridge university press, 2012.

[4] 王松桂，吴密霞，贾忠贞. 矩阵不等式（第二版）[M]. 北京：科学出版社，2006.

索引