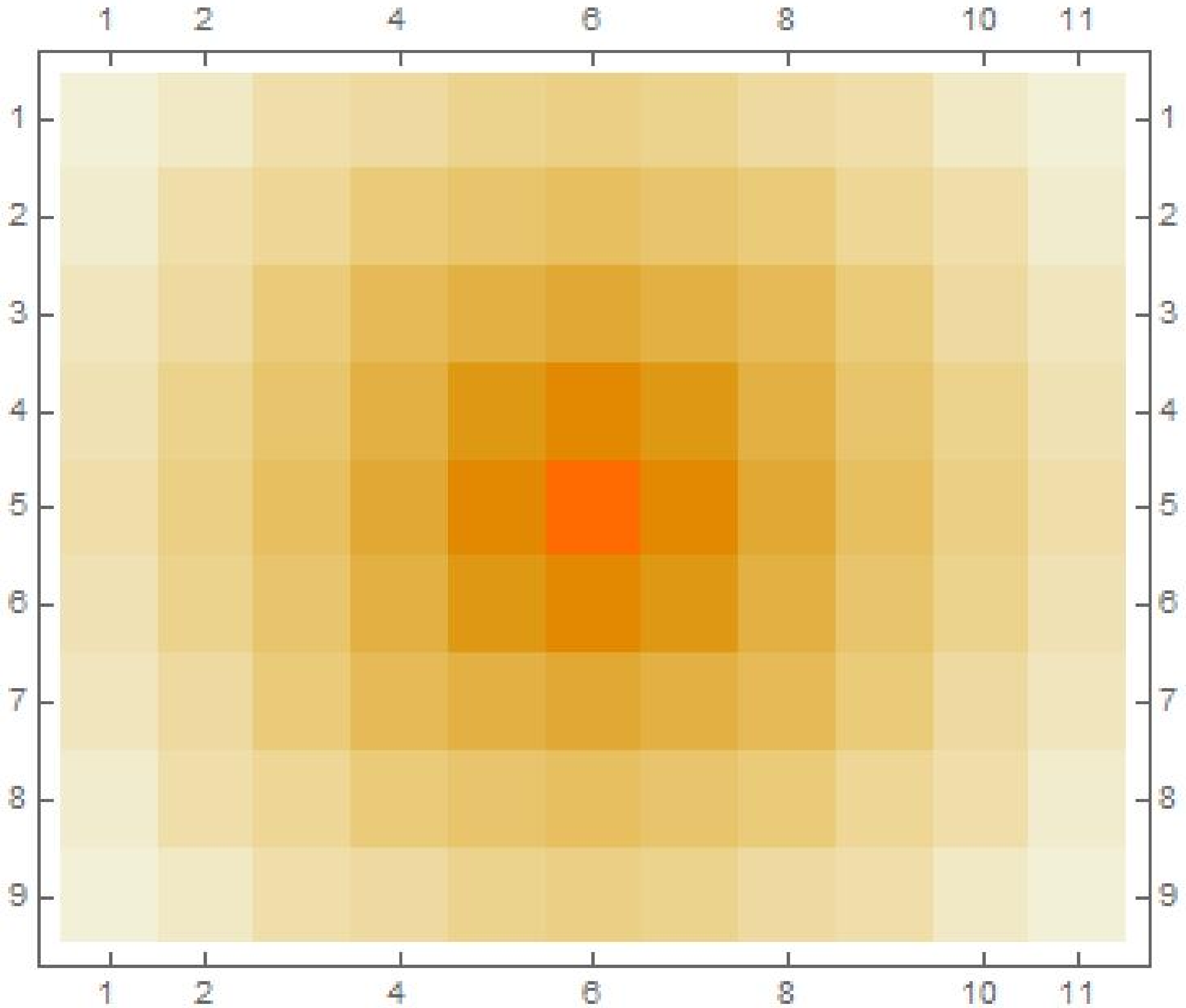


线性代数
习题课讲义

作者：刘国瑞 郑滕飞

时间：January 15, 2024

目录

* 第二、四、七、八章为刘国瑞制作，第一、三、五、六、九、十章为郑滕飞制作

第 1 章 当我们谈论矩阵时，我们在谈论什么

作为第一次习题课，在大家算了两周行列式后，这节课也并不会有什么新的知识，主要是对过去的知识以一个特定脉络梳理，并进行一定的补充。

本章的标题已经说明了脉络：我们将探索几种视角下的矩阵，并且根据每种视角添加对应的基础知识，进行适当的拓展以更好完成后续的课程。

1.1 数与数域

1.1.1 数环、数域、完备性

关于矩阵的最浅显视角，自然是**满足某些运算关系的、一堆元素排成的矩形**。然而，哪怕这个简单的视角下，仍然存在问题：这些元素究竟是什么？

在线性代数 B1 中，我们主要了解的是两类矩阵，**实矩阵** [即所有元素是实数] 与**复矩阵** [即所有元素是复数]。为限定元素的范围，我们给出如下定义：

定义 1.1 (数环、数域)

若复数 $\mathbb{C}$ 的某个包含 1 的子集 A 满足：

1. $x \in A, y \in A \Rightarrow x \pm y \in A$
2. $x \in A, y \in A \Rightarrow xy \in A$

则 A 称为一个数环。

若还满足 $x \in A, x \neq 0 \Rightarrow \frac{1}{x} \in A$，则 A 称为一个数域。

♣

直接验证有结论：

命题 1.2 (数环、数域-例子)

有理数 $\mathbb{Q}$、实数 $\mathbb{R}$、复数 $\mathbb{C}$ 都是数域。整数 $\mathbb{Z}$ 是数环而不是数域。

♠

当然，还有形态更复杂的数环和数域，例如所有 $a + bi$ 在 $a, b \in \mathbb{Z}$ 时构成数环，$a, b \in \mathbb{Q}$ 时构成数域。

对于究竟怎样的集合是数环/数域，有个简单的结论：

命题 1.3 (最小数环/数域)

任何数环包含 $\mathbb{Z}$，任何数域包含 $\mathbb{Q}$。

♠

下面考虑某个数域 $\mathbb{F}$，当矩阵 A 的元素都在 $\mathbb{F}$ 中时，称 A 是 $\mathbb{F}$ 上的矩阵。由于迹、行列式的计算只涉及加减乘，可得 $\det(A) \in \mathbb{F}, \operatorname{tr}(A) \in \mathbb{F}$，换而言之，det 与 tr 都可以看作 $\mathbb{F}^{n\times n} \to \mathbb{F}$ 的函数。

关于矩阵间的运算，同上可知 A, B 均为 $\mathbb{F}$ 上矩阵时，$A + B, A - B, AB$ 若可进行 (行列数符合要求)，则结果都是 $\mathbb{F}$ 上的矩阵。而对于矩阵求逆，利用伴随方阵，可知可逆等价于行列式非零，而计算逆也可仅通过加减乘除进行，从而 A 可逆时 A^{-1} 各个元素也在 $\mathbb{F}$ 中。从而，考虑**某个数域上的矩阵**已经能使得常用的大部分操作可以进行、大部分运算封闭了。

我们常用的行列变换都可以看作乘以矩阵，而对于数域内的矩阵，消元一般采取数域中的系数进行变换，因此通常能保持在数域中。[例如，对有理数上的矩阵进行行列变换时一般不会把某行乘 $\sqrt{2}$ 倍，这

样变换完后仍然是有理数上的矩阵。]

然而，之所以说大部分操作，是因为我们还漏掉了一个重要的矩阵属性：特征值。先给出结论：

命题 1.4 (完备性)

$\mathbb{C}$ 上的方阵 A 的特征值均在 $\mathbb{C}$ 中，而 $\mathbb{R},\mathbb{Q}$ 上均不满足此性质。

这也就意味着，哪怕开始时讨论的数域不涉及复数，只要涉及特征值、特征向量的问题，一般必须放到最大的数域 $\mathbb{C}$ 里讨论——其中最普适的 Jordan 标准形同样是在 $\mathbb{C}$ 上的矩阵定义的。

之所以说一般，是因为有理数的完备化并不需要到 $\mathbb{C}$，只需要全部代数数，而这是一个非常复杂的结构，在线性代数中几乎不单独讨论，于是仍然直接考虑 $\mathbb{C}$ 即可。

不过，如果只关乎相似，不显式出现特征值、特征向量，则有结论：

命题 1.5 (数域上的相似)

对数域 $\mathbb{F}$ 上的方阵 A,B，若存在 $\mathbb{F}$ 上可逆方阵 P 使得 $A=P^{-1}BP$，则称 A,B 在 $\mathbb{F}$ 上相似。只要 $\mathbb{F}$ 上的方阵 A,B 在 $\mathbb{C}$ 上相似，它们一定在 $\mathbb{F}$ 上相似。

1.1.2 单位根

关于复数，本节介绍复数中的一个最重要的循环结构——单位根。

将复数写为极坐标表示 $r(\cos\theta+\mathrm{i}\sin\theta)$ 后，利用**棣莫佛公式**，复数的乘法即对应旋转与缩放。而任何一个单位复数 $\cos\theta+\mathrm{i}\sin\theta$ 都可以看作绕原点逆时针旋转 θ 的变换。在此基础上，我们定义：

定义 1.6 (n 次单位根)

n 次单位根是指

$$\cos\frac{2k\pi}{n}+\mathrm{i}\sin\frac{2k\pi}{n}, k=,0\dots,n-1$$

这 n 个复数，记 $\omega_n=\cos\frac{2\pi}{n}+\mathrm{i}\sin\frac{2\pi}{n}$，则它们可以写为 $\omega_n^0,\omega_n^1,\dots,\omega_n^{n-1}$。

之所以叫单位根，是因为其满足：

命题 1.7 (单位根性质)

n 次单位根是方程 $x^n=1$ 的全部根。

根据**因式定理**，由此可以推出

$$x^n-1=\prod_{k=0}^{n-1}(x-\omega_n^k)$$

对比各项系数可以得到一些恒等式。事实上最常用的为：

命题 1.8 (单位根恒等式)

1.
$$\sum_{k=0}^{n-1}(\omega_n^k)^m=\begin{cases} n & n\mid m \\ 0 & n\nmid m \end{cases}$$

2.

$$\omega_n^a = \omega_n^{a\pm n}$$

3.

$$\bar{\omega}_n^a = \omega_n^{-a}$$

4.

$$\omega_{mn}^{mk} = \omega_n^k$$

第二个恒等式就是最终要的**循环**性质。出于循环，它可以将一些东西按照模 n 的余数分类。例如：

例题 1.1 若已知多项式 $f(x) = a_0 + a_1 x + \cdots + a_m x^m$，如何计算 $\sum_{i \equiv t \mod n} a_i$? [进一步地，取 $f(x) = (1+x)^m$，$a_i = C_m^i$，这就算出了模 n 特定余数的组合数的和。]

提示：代入 $f(1), \dots, f(\omega^{n-1})$ 列出方程，记方阵 Ω 的元素 ω_{ij} 为 $\omega_n^{(i-1)(j-1)}$，并证明 $\Omega^{-1} = \frac{1}{n}\bar{\Omega}$。此处的 Ω 是对称阵，与离散傅里叶变换密切相关，是一个十分重要的矩阵。可以发现，它是 Vandermonde 阵的一种特殊情况。

回归矩阵，矩阵里是否有类似的循环结构呢？对于 $A^n = I$，矩阵上的求解是可能性极多的，但存在一些共同性质。例如，利用相似三角化可以得到 A^n 的特征值为 A 的特征值对应 n 次方，因此 A 的每个特征值 λ 都要满足 $\lambda^n = 1$，即 λ 是 n 次单位根。对 n 阶方阵，其中或许是最特殊的 A 为

$$\begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix}$$

下面记 n 阶情况为 W_n。有结论：

命题 1.9 (W_n 的性质)

1. $W_n^n = I$；
2. $\sum_{k=0}^{n-1} W_n^k = \mathbf{1}$，这里 $\mathbf{1}$ 代表所有元素全为 1 的方阵；
3. $W_n^T W_n = I$；
4. W_n 的特征值恰为 $1, \omega_n, \dots, \omega_n^{n-1}$。

根据这些结论，某种意义上，W_n 可以看作 n 阶矩阵中的"单位根"，具有很好的循环性。由此，我们重新看待循环方阵问题：

例题 1.2 对满足 $a_{ij} = a_{[i-j \mod n]}$ 的循环方阵 A，试计算其特征值。

提示：

$$A = \sum_{k=0}^{n-1} a_k W_n^k$$

利用相似三角化可知 W_n 的多项式的特征值是 W_n 的特征值的多项式，从而即能算出结果。

更进一步地，所有每行每列恰有一个 1，其他为 0 的方阵称为**置换阵**，也具有一些特殊的性质，会在后续介绍。

1.1.3 整数上的矩阵

刚才一直在讨论的是数域上的矩阵，那么，如果在数环上谈论矩阵呢？考虑最简单的数环，也即**整数环**上的矩阵。

若要将矩阵所有元素都限制在整数，加法、减法、乘法、行列式、迹都是可以计算的，然而，原本可逆的矩阵可能会变得**未必可逆**，例如一阶方阵 (2) 在整数环中并不可逆。为了表示区分，在数环上谈论矩阵时，我们不再用“可逆”去描述，而是用**模方阵**去谈论：

定义 1.10 (模方阵)

对一个数环 R，若 A 是 R 上的方阵，且 A^{-1} 也是 R 上的方阵，则称 A 为 R 上的模方阵。♣

在此定义下，容易发现，一阶的模方阵只有 (1) 与 (-1)。

例题 1.3 试找出整数上的所有二阶模方阵。

提示：利用伴随方阵。

事实上可证明：

命题 1.11 (模方阵)

A 是数环 R 上的模方阵当且仅当 $\det A$ 是 R 上的可逆元 (即 $\frac{1}{\det A}$ 在 R 中)。♠

就像我们通过可逆阵定义相抵，通过模方阵也可以定义**模相抵**：

定义 1.12 (模相抵)

对数环 R 上的矩阵 A, B，若存在 R 上模方阵 P, Q 使得 $A = PBQ$，则称 A, B 模相抵。♣

稍微进行一些推广：除了数以外，可以发现，所有实系数多项式/复系数多项式也满足上述“环”的性质，于是多项式方阵中也可以定义模方阵、模相抵。其更深入的性质会在课程后续中学到。一个现在可以发现的事情是，求特征值的过程事实上是计算了一个多项式上的方阵 $\lambda I - A$ 的行列式 (仍为关于 λ 的多项式)，并计算其根。

值得一提的是，这里的“模”指的是一个有趣的代数结构：**模**，可以简单看作环上的线性空间。此时，模方阵即代表两个模的线性同构映射。

1.2 代数结构

对于矩阵的另一个重要视角，是它可以看成有限维线性空间中的线性映射。具体来说，$A \in \mathbb{R}^{m\times n}$ 与 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射 $\mathcal{A}(x) = Ax$ 完全等价。不过，为了更好描述这种看法的意义，我们需要先观察最简单的代数结构——集合。

1.2.1 集合

关于集合的严谨公理化定义详见附录，此处保留一个直观的感受即可。对两个集合 A, B 映射 f 将 A 中的每个元素对应到 B 中的某个元素上。回顾一些基本定义：

定义 1.13 (映射相关定义)

对映射 $fA \to B$，A 的子集 A_0 的像集

$$f(A_0) = \{b \in B \mid \exists a \in A_0, f(a) = b\}$$

B 的子集 B_0 的原像

$$f^{-1}(B_0) = \{a \in A \mid \exists b \in B_0, f(a) = b\}$$

f 是满射是指 $f(A) = B$，f 是单射是指 $\forall a_1, a_2 \in A, a_1 \neq a_2, f(a_1) \neq f(a_2)$。$f$ 是双射是指 f 既是单射又是满射。 ♣

命题 1.14 (复合、逆)

对于映射 $fA \to B$ 与 $gB \to C$，记其复合映射为 $g \circ fA \to C$，满足 $(g \circ f)(a) = g(f(a))$。

映射复合满足结合律，即 $(h \circ g) \circ f = h \circ (g \circ f)$。

映射 $fA \to A, f(a) = a$ 称为恒等映射，记作 Id_A。

考虑 $fA \to B$。若其为单射，则存在 $g_1 B \to A$ 使得 $g_1 \circ f = \mathrm{Id}_A$，称为 f 的左逆；若其为满射，则存在 $g_2 B \to A$ 使得 $f \circ g_2 = \mathrm{Id}_B$，称为 f 的右逆。

左逆唯一、右逆唯一、左逆右逆都存在、f 是双射四者等价，此时左逆右逆相等，称为 f 的逆，记作 f^{-1}。 ♠

在这些复杂的定义与性质之后，有一个更基本的问题，也即**这些定义是为了什么**？想要回答这个问题，就必须关注作为一个代数结构的集合。

集合上有什么结构呢？这个问题其实也就等价于，什么时候两个集合是**可区分**的。由于不附加任何其他结构的集合并不具有特殊性，只要两个集合的**元素个数相同**——更严谨来说，存在双射，它们就是不可区分的。由此可以有这样的理解：

集合 A 与 B 之间的映射意味着其结构间的联系，A 到 B 存在单射意味着 A 的结构被 B 蕴含，存在满射则意味着 A 的结构蕴含 B 的结构，若存在双射，即代表 A, B 的结构相同。

此外，我们还有一些生成集合的方式：对集合 A，我们可以谈论它的**子集** (子结构)，对两个集合 A, B，可以考虑包含它们的最小集合 $A \cup B$ (生成结构)、被它们包含的最大集合 $A \cap B$ (交结构) 与它们的**笛卡尔积** $A \times B$(积结构)。

当然，这两段话单看都显得十分奇怪，因为它们在将一个已经被我们直观理解的东西抽象化。然而，当与更多代数结构进行对比时，它们的含义就变得重要了。作为本段的结尾，我们介绍一个较为隐蔽的构造集合的方式，集合中的商结构，**商集**：

定义 1.15 (商集)

对集合 A，若一族非空集合 $A_i, i \in I$ 满足两两交为空且 $\cup_{i \in I} A_i = A$，则称它们是集合 A 的一个划分。这些 A_i 所构成的集合 $\{A_i, i \in I\}$ 称为集合 A 的一个商集。 ♣

例如，对 $\{1, 2, 3\}$，$\{\{1, 2\}, \{3\}\}$ 就是它的一个商集。不过，既然叫商，说明它的确商去了某个东西，这个“东西”就是**等价关系**，我们将在习题课的最后一部分加以说明。

1.2.2 群、环、域

数学总是从特殊走向一般。当我们习惯了 $\mathbb{C}$ 上熟悉的乘法后，我们用它自然定义出了多项式的乘法，也在线性代数中衍生出了矩阵的乘法。然而，到底什么是“乘法”运算呢？为了回答这个问题，我们需要抽象出一个只包含乘法的结构，这就是群：

定义 1.16 (群)

对集合 A 与乘法映射 $\mathrm{mul} A \times A \to A$[为了符合我们习惯的二元运算形式，$\mathrm{mul}(a,b)$ 也记作 $a \cdot b$，无歧义时也可直接写成 ab]，若满足：

1. $\forall a,b,c \in A, (ab)c = a(bc)$，即结合律；
2. 存在唯一 $e \in A$ 使得 $\forall a \in A, ae = ea = a$，此 e 称为单位元；
3. $\forall a \in A$，存在唯一 $b \in A, ab = ba = e$，此 b 记作 a^{-1}，称为 a 的逆元；

则附加了乘法的集合 $(A, \cdot)$ 称为一个群。

若还有 $\forall a, b \in A, ab = ba$，则此群称为交换群或 Abel 群。♣

例题 1.4

1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ 在加法下构成 Abel 群，$\mathbb{Q}, \mathbb{R}$ 去掉 0 后在乘法下构成 Abel 群；
2. 所有 $m \times n$ 阶矩阵在加法下构成 Abel 群，所有 n 阶可逆方阵在乘法下构成群；
3. 一个集合 A 上的所有双射在复合下构成群，特别地，当 A 为有限集合时，此群称为置换群。

就像集合，一个孤立的集合可以谈论子集，一个群也可以谈论子群：

定义 1.17 (子群)

若 G 在运算 $\cdot$ 下构成群，H 是 G 的子集，且 H 在 $\cdot$ 下也是群，则 H 称为 G 的子群。

子群的等价定义为，H 是 G 的子集，且 H 满足 $e \in H$、$a \in H \Rightarrow a^{-1} \in H$ 与 $a \in H, b \in H \Rightarrow ab \in H$。

子群的另一个等价定义即 H 非空且 $a \in H, b \in H \Rightarrow a^{-1}b \in H$。♣

这时子群作为子结构的特性就非常明显了。下面考虑群中的生成结构与交结构，注意到，正如我们谈论特定集合一般需要在某个“万有集合”中，由于结构的限制，谈论群的交与生成一定要在某个“万有群”中，即：

定义 1.18 (交群、生成群)

对于 G 的两个子群 H 与 K，$H \cap K$ 也是一个子群，称为两个子群的交群，而记 $\langle H, K \rangle$ 为两个子群的生成子群，即同时包含它们的最小子群。♣

集合中谈论最小性是指，任何一个满足条件的集合都包含此集合，也即若生成子群为 S，任何满足 $H \subset A, K \subset A$ 的 G 的子群 A 都有 $S \subset A$。

而对于两个无关的群则可以谈论它们的积群：

定义 1.19 (积群)

对两个群 G, H，在 $G \times H$ 上定义运算 $(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2)$，则 $G \times H$ 在此运算下构成一个群。♣

回到与集合的类比，在这时，这些生成新的群的方式就有了明确的意义——它代表着生成一些**保**

持原有代数结构的“东西”。

当然，只有群是并不完善的，就像集合之间需要映射，我们需要在群之间建立结构上的关联。然而，因为元素地位的差别，这样的结构关联只通过映射是无法保持的，我们需要添加新的要求：

定义 1.20 (群同态)

对两个群 G, H，映射 $fG \to H$ 若满足 $f(g_1)f(g_2) = f(g_1g_2)$，则称为群同态。

群同态若为单射，则称为单同态；若为满射，则称为满同态；若为双射，则称为同构。

就像存在双射意味着两个集合在结构上相同，存在同构也意味着两个群在结构上相同，我们就以之前提到的单位根与置换阵作为例子：

例题 1.5 $\{1, \omega_n, \ldots, \omega_n^{n-1}\}$ 在乘法下构成一个群，$\{I, W_n, \ldots, W_n^{n-1}\}$ 在矩阵乘法下构成一个群，且两个群都同构于 $\{0, 1, \ldots, n-1\}$ 在模 n 同余意义的加法下构成群，

提示：这里写出元素的顺序即为需要构造的同构映射。

例题 1.6 线性代数 A1 期中，22 春 n 元集合上所有双射在复合下构成一个群，所有置换方阵 (定义见上一节) 在矩阵乘法下构成一个群，且两个群同构。

提示：将 n 元集合写成向量 $x = (1, 2, 3, \ldots, n)$，双射可以对应 $x \to Ax$。

在谈论群时，同构的群一般可以视为同一个，上方两个例子中的群分别称为 n 阶循环群 $\mathbb{Z}_n$ 与 n 阶置换群 S_n。这时再看集合中“映射意味着集合间联系”一段话，就能与“同态意味着群间联系”相对应。

群与接下来讨论的环也有商结构，但理论相对复杂，定义可见参考资料。

在群的基础上，若是我们抽象出“具有两种运算的结构”，就有了环和域。值得注意的是，如果两种运算要同时存在于一个结构上，它们必须具有某种**相容性**。对于乘法和加法，我们早在小学时就学过了它们的相容性：$a(b+c) = ab + ac$，这就是**分配律**。

定义 1.21 (域)

考虑集合 $\mathbb{F}$ 上的运算 $+$ 与 $\times$ [此处均为任意二元运算，为方便记为加法、乘法]，若 $\mathbb{F}$ 在加法下构成一个 Abel 群，且单位元记为 0，$\mathbb{F}$ 去掉 0 形成的集合 $\mathbb{F}^*$ 在乘法下构成一个 Abel 群，且加法、乘法满足分配律

$$a(b+c) = ab + ac$$

则 $(\mathbb{F}, +, \times)$ 称为一个域。

若乘法下 $\mathbb{F}^*$ 只需要构成群而非 Abel 群，则称为体。

上方域的定义基本是同一个集合上满足某些相容性的两个群，而环的定义则要减少一个要求：

定义 1.22 (环)

考虑集合 R 上的运算 $+$ 与 $\times$ [此处均为任意二元运算，为方便记为加法、乘法]，若 R 在加法下构成一个 Abel 群，乘法满足结合律、乘法单位元存在，且加法、乘法满足左右分配律

$$a(b+c) = ab + ac, (a+b)c = ac + bc$$

则 $(R, +, \times)$ 称为一个环。

环与域中 a 的加法逆元一般记作 $-a$，而乘法逆元若存在则记作 a^{-1}。比起域，环并不要求非零元

可逆，也不要求乘法可交换 (因此分配律必须左右都保证)。

命题 1.23 (数环/数域的等价定义)

数环/数域即为通常加法、乘法下构成环/域的复数集合。

除了它们以外，我们还将常见如下的环与域：

例题 1.7

1. 若 R 为环，R 上所有多项式的集合 $R[x]$ 在多项式加法、乘法下构成环。特别地，R 为域时也如此。
2. 所有 n 阶方阵构成环 M_n。
3. $\{0,1,\ldots,n-1\}$ 在模 n 同余意义的加法、乘法下构成环，且此环为域当且仅当 n 是素数，称为 p 元域。

域上的多项式环与环上的多项式环存在本质差别，后续课程中将会学习。

仿照群，环与域同样具有子结构、交结构、生成结构与积结构，也仍可以类似讨论同态与同构。此时，保持结构的映射要求 $f(a+b)=f(a)+f(b), f(ab)=f(a)f(b)$。某种意义上，保持结构的映射也就意味着其**与运算的相容性**，这里体现为先映射后运算与先运算后映射的结果相同。

1.2.3 线性空间

在线性代数 B1 中，我们定义线性空间都是在数上，但事实上，在任何域 $\mathbb{F}$ 上都有线性空间，定义如下：

定义 1.24 (域上的线性空间)

对域 $\mathbb{F}$、Abel 群 V [这意味着 V 已经包含了交换结合的加法运算，且存在单位元与逆元]，与它们之间的数乘运算 $\mathbb{F}\times V\to V$，若数乘运算满足

1. $1\alpha=\alpha$，这里 1 为 $\mathbb{F}$ 中乘法单位元；
2. $\lambda(\mu\alpha)=(\lambda\mu)\alpha$，数乘结合律；
3. $(\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha$，左分配律；
4. $\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$，右分配律；

则 V 称为一个 $\mathbb{F}$ 上的线性空间。

此处值得关注的是，左分配律的加法为 $\mathbb{F}$ 作为域上的加法，而右分配律中的加法为 V 作为 Abel 群上的加法。当 $\mathbb{F}$ 取为 $\mathbb{R}$ 或 $\mathbb{C}$ 时，这就是我们熟悉的线性空间，但对其他的域，可能具有其他性质，例如 $\mathbb{F}$ 取二元域时，线性空间中任何元素满足 $\alpha+\alpha=(1+1)\alpha=\mathbf{0}$。

回顾子空间、交空间与和空间的定义，它们事实上与集合或群时一致：

定义 1.25 (生成新线性空间的方式)

对域 $\mathbb{F}$ 上的线性空间 V，其子集 U 在相同的运算下构成线性空间时称为 V 的子空间。

其子空间 U,W 的交 $U\cap W$ 仍为线性空间，称为交空间。

包含其子空间 U,W 的最小线性空间为 $U+W$，即 $\{u+w\mid u\in U,w\in W\}$，称为两者的和空间。

当然，两个线性空间也可以作直积成为新的线性空间，不过直积并不如直和运用广泛，这和线性空间本质上的**有限性**——任何向量都能被有限个基向量生成——密切相关，我们会在之后深入介绍。

线性空间有着群或环并不具有的性质，也即它可以被一组基完全确定。承认选择公理成立时，任何线性空间都存在一组基 $\alpha_i, i \in I$，而线性空间的任何元素可以唯一写为 $\sum_i \lambda_i\alpha_i, i \in I, \lambda_i \in \mathbb{F}$，这里 λ_i 只有有限个非零 [因此可以严谨定义求和]。

为了保持线性空间的结构，仿照前文，我们定义**线性映射**：

定义 1.26 (线性映射)

对 $\mathbb{F}$ 上的线性空间 U, V，映射 f 满足 $f(a+b) = f(a) + f(b), f(\lambda a) = \lambda f(a)$，则称为它们之间的一个线性映射。

若线性映射还是双射，则称为线性同构。

♣

由于只需要知道一组基 α_i 的像 $\beta_i = f(\alpha_i)$，若 U, V 均为有限维，U 的一组基为 $\alpha_1, \dots, \alpha_n$，$V$ 的一组基为 $Mathematica_1, \dots, \gamma_m$，则 U 中任何元素可以用一个 m 维列向量表示 [即代表每个基前的系数]，V 中任何元素可以用一个 n 维列向量表示。

对线性映射 $\mathcal{A}$，假设 $\mathcal{A}(\alpha_i) = \sum_j a_{ij}\gamma_j$，则其可以用矩阵 (a_{ij}) 表示，任何 U 中元素的像对应其向量表示作矩阵乘法后的向量。

1.2.4 视作线性映射的矩阵

在向量表示下，矩阵可以代表一个**有限维线性空间之间保持结构的映射**。其最重要的对应性在于：

命题 1.27 (矩阵乘法与线性映射复合)

假设 U, V, W 分别为 m, n, p 维线性空间，其一组基分别为 $\{u_i\}, \{v_j\}, \{w_k\}$，考虑 $V \to W$ 的线性映射 $\mathcal{A}$ 与 $U \to V$ 的线性映射 $\mathcal{B}$，在对应基下矩阵表示分别为 A 与 B，则 $U \to W$ 的线性映射 $\mathcal{A} \circ \mathcal{B}$ 在对应基下的矩阵表示为 AB。

♠

事实上，这才是矩阵乘法如此定义的原因，它满足结合律也是由映射复合满足结合律自然得到的。

某种意义上，所有有结合律的地方都可以看成映射复合。

而对于矩阵的一个重要概念**秩**，映射角度我们也有结论：

命题 1.28 (秩)

若矩阵行满秩，其对应的映射为满线性映射；若矩阵列满秩，其对应的映射为单线性映射；若矩阵可逆，其对应的映射为线性同构。

♠

事实上，秩可决定线性映射的核与像的维数：

命题 1.29 (秩与维数)

像空间 $\operatorname{Im}\mathcal{A} = \mathcal{A}(V)$ 的维数为 $\operatorname{rank} A$，核空间 $\operatorname{Ker}\mathcal{A} = \{v \mid \mathcal{A}(v) = 0\}$ 的维数为 $\dim V - \operatorname{rank} A$。

♠

直接验证可知二者为线性空间。值得注意的是，维数定义、最大可逆子矩阵定义与向量组秩定义都可对任何域 $\mathbb{F}$ 上的矩阵成立。

于是从线性映射的角度，秩基本决定了映射对两边结构的“保持性”，这与信息角度秩代表矩阵的“信息量”是关联密切的，后者我们会在之后提及。

正因为从线性映射角度可以清晰看出秩的含义，很多秩相关的等式与不等式用线性映射角度都易于证明，如：

命题 1.30 (秩相关性质)

1. $\operatorname{rank}(AB) \le \operatorname{rank}(A)$，当 B 行满秩时等号成立；
2. $|\operatorname{rank}(A) - \operatorname{rank}(B)| \le \operatorname{rank}(A+B) \le \operatorname{rank}(A) + \operatorname{rank}(B)$；
3. $\operatorname{rank}(A) \le \operatorname{rank}(A\ B) \le \operatorname{rank}(A) + \operatorname{rank}(B)$。

当然，秩相关的问题将矩阵视为向量组也是常用的解决方式，而考虑基的像可知视为线性映射与视为列向量组是完全**等价**的，进行转置就得到行向量组与线性映射亦等价。

此外，由于无穷维空间也可以存在线性映射，此时未必可以用矩阵良好表示，这种看法下意味着线性映射就是矩阵的**推广**。有时，对于可数无穷维空间，我们也会将线性映射写成矩阵的形式——矩阵力学就用了类似的想法。

写成无穷维后并不需要担心乘积收敛的问题，因为根据线性空间的要求，矩阵每列必然只有至多有限个非零分量。

1.3 仿射

1.3.1 高维空间几何

在刚才的例子里，矩阵作为了**抽象**的 n 维到 m 维的线性空间之间的线性映射。但事实上，对于 $\mathbb{R}$ 上的矩阵，我们完全可以将它看成一个**具体**、简单的线性映射，也即 $\mathcal{A}\mathbb{R}^n \to \mathbb{R}^m, \mathcal{A}(x) = Ax$。对于方阵，这个映射 $\mathcal{A}$ 就成为了 $\mathbb{R}^n$ 上的一个变换 [此处变换定义为 $\mathbb{R}^n$ 到自身的映射]。在这样的视角下，我们可以用高维空间几何的方式去看待矩阵。

首先回顾向量的模长与夹角：

定义 1.31 (内积)

两个 n 维向量 x, y 的内积 $x \cdot y = \sum_{i=1}^n x_i y_i$，一个向量的模长定义为 $\|x\| = \sqrt{x \cdot x}$，两个向量的夹角定义为 $\theta = \frac{x \cdot y}{\|x\|\|y\|}$。

夹角定义的严谨性需要通过柯西不等式证明。

以及自然推广的向量位置关系：

定义 1.32 (平行、正交)

两向量 x, y 平行当且仅当存在不全为 0 的 λ, μ 使得 $\lambda x + \mu y = \mathbf{0}$，正交当且仅当 $x \cdot y = 0$。

虽然之后的讨论都是直接在 n 维空间中进行的，若难以想像，二维和三维可以提供很多有效的例子。在二维、三维中，我们熟悉平移、旋转、翻折等种种变换，这三种变换统称为**等距变换**，因为它们保持了**任意两点之间的距离不变**。作为高维的推广，我们先来简单探究一下高维空间中平移的推广与表达方式。根据定义，平移保持任何两点的距离与两点间的方向都不变，假设原点变为点 b，容易推出：

命题 1.33 (平移的表示)

n 维空间中任何平移变换都可以写为映射 $f(x) = x + b$。

1.3.2 仿射变换与矩阵

为了探究上方的 $\mathcal{A}$ 对应的变换性质，我们定义仿射变换：

定义 1.34 (仿射变换)

$\mathbb{R}^n$ 上的变换 f 称为仿射变换，当对任何满足 AB 平行于 CD，AB, CD 不为零向量的的 A, B, C, D 都有 $f(A)f(B)$ 平行于 $f(C)f(D)$，且

$$\|f(A)f(B)\| = \|f(C)f(D)\| = 0 \text{ or } \frac{\|f(A)f(B)\|}{\|f(C)f(D)\|} = \frac{\|AB\|}{\|CD\|}$$

♣

对此有结论

命题 1.35 (仿射变换与矩阵乘法)

任何 $\mathbb{R}^n$ 中的变换是仿射变换当且仅当其可以写成 $f(x) = Ax + b$，其中 A 为 n 阶方阵，进而任何将原点映射到原点的变换是仿射变换当且仅当其可以写成 $f(x) = Ax$。♠

想要证明这件事，我们先考虑将原点映射到原点的仿射变换。先取出 $\{e_i\}$，这里 e_i 表示只有第 i 个分量为 1，其他为 0 的向量。根据仿射变换的性质可以证出

$$f\left(\sum_i \lambda_i e_i\right) = \sum_i \lambda_i f(e_i)$$

这就说明了 f 是线性映射，因此是 $f(x) = Ax$。对于一般的仿射变换，只要复合平移 [易验证平移仿射变换] 使得原点回到原点，就可以证明存在 b' 使得 $f(x) - b = Ax$，从而 $f(x) = Ax + b$。另一方面，在给出 f 的形式时验证其为仿射变换是不难的。

由此可见，$\mathcal{A}(x) = Ax$ 实质上是一个**保原点、保平行、保平行线段比例**的变换。在这个视角下，**行列式**具有很明显的几何意义：

命题 1.36 (行列式-几何意义)

记所有 e_i 张成的高维立方体 (如二维时为正方形，三维时为立方体) 为 K，行列式的绝对值表示 $\mathcal{A}(K)$ 的体积 [同样也是任何 n 维体变换后的体积与变换前之比]，若非零，其为负表示反转定向。♠

这里绝对值表示体积变化率在代入二维、三维情况是容易理解的，若行列式为 0，表示其不再是 n 维体，也即这个仿射变换是**退化**的，并不可逆。而反转定向的定义则相对难以理解——简单来说，二维的两种定向分别是 x 轴顺时针转到 y 轴更近与 x 轴逆时针转到 y 轴更近 [若一样近，xy 轴重合，是退化情况，无需考虑定向]，而三维则是左手螺旋与右手螺旋。

利用行列式的几何意义可以给一些问题直观的看法，如：

例题 1.8 记最后一列为 1 的 n 阶方阵 A 的行列式为 Δ，向量 $x = (x_1, \ldots, x_{n-1}, 1)$，将 x 分别替换 A 的每一行，得到的方阵的行列式记为 Δ_n，则

$$\Delta = \sum_{i=1}^{n} \Delta_n$$

我们可以用二维、三维的情况给这个结论一个直观的理解，也即本质为有向面积/体积的加法。然而，虽然几何理解是容易的，想要通过几何严谨写出证明是极为困难的 (因为这个问题里涉及正负定向)。这意味着，**几何直观在矩阵中的运用往往是猜测结论，而非证明结论**。

根据特征方程 $Ax=\lambda x$，似乎特征值、特征向量的几何意义是 Ax 与 x 平行的向量与它们的长度比例，但需要注意谈论特征值需要在 $\mathbb{C}$ 中，因此这里的平行、长度都是在 $\mathbb{C}^n$ 中的，而非简单的 n 维欧氏空间，几何直观并不明显。

1.3.3 幂等、对合、正交

话虽如此，对一些特殊的矩阵，的确可以找到它的几何特性，从而更好理解性质，这里举出三个例子：

定义 1.37 (几种特殊矩阵)

对 $A\in\mathbb{R}^{n\times n}$：

1. 若 $A^2=A$，则 A 称为幂等阵；
2. 若 $A^2=I$；则 A 称为对合阵；
3. 若 $A^TA=I$，则 A 称为正交阵。

首先，对于幂等阵，有结论：

命题 1.38 (幂等-几何性质)

A 是幂等阵等价于 $\mathcal{A}(x)$ 是对某个 $\mathbb{R}^n$ 子空间 V 的投影，这里 x 到 V 的投影 $P_V(x)$ 是指使得 $\|x-v\|$ 最小的 V 中元素 v。设 V 的一组基是 $v_1,\ldots,v_m$，添加 $u_{m+1},\ldots,u_n$ 扩充为空间的一组基，则

$$P_V(x)=\sum_{i=1}^m\lambda_iv_i$$

其中 $\sum_{i=1}^m\lambda_iv_i+\sum_{i=m+1}^n\lambda_iu_i=x$。

根据几何，由于投影到的空间是平直的，它的确是一个线性变换，且原点必然映射到原点，再由投影的定义可知投影对应的矩阵是幂等的；反之，对于幂等矩阵 A，我们记其列向量生成空间为 V，通过幂等可知 $\forall v\in V, Av=v$，又由于 $\forall x, Ax\in V$，即可以找到一组 $\mathbb{R}^n$ 的基使得有些映射为自身，其他映射为 0，从而验证投影性。不过，此性质还有一种更简单的看法：投影后再次投影对结果是无影响的，这与矩阵的幂等性质天然接近。

关于对合阵的讨论则更为有趣，根据定义，其对应所有原点不变且逆为自身的仿射变换。回顾作业题：

例题 1.9 找到所有二阶对合方阵。

可以发现，对合阵对应的变换不仅有我们熟悉的翻折与旋转 π，还有些并不直观的结果。在几何上，有如下结论：

命题 1.39 (对合-几何性质)

A 是对合阵等价于 $\mathbb{R}^n$ 可以分解为两个子空间 U,V [此处分解指 $U\cap V=\mathbf{0}, U+V=\mathbb{R}^n$]，使得 $\forall u\in U, Au=u, \forall v\in V, Av=v$。

从几何性质推出对合是容易的，考虑 U,V 的各一组基即可，而对合推出几何性质则较为复杂。证明梗概：先定义出 U,V，若 $U+V$ 不是 $\mathbb{R}^n$，假设它们之外的一组基构成线性空间 W。对任何 $w\in W$，可以证明 $Aw\in W$，否则 $A(Aw)\notin W$。记它们连线为 l，由仿射变换性质 $\mathcal{A}(l)=l$，而根据介值定理可

以得到 w 与 Aw 连线上会有属于 U 的非零点 (若此点为 $\mathbf{0}$，w, Aw 共线，由条件可以推出 $w = \pm Aw$)，与它们都在 W 中矛盾。[这基本是 22 秋几何学基础的期末最后一题。]

对于二维的情况，这个几何性质可以直观理解：当 U 是二维时，意味着这个变换是恒等变换；当 V 是二维时，意味着这个变换是绕原点旋转 π；而若两者都是一维，当且仅当两条直线**垂直**时其可以看成翻折，否则会成为怪异的结果。

例题 1.10 分析三维时对合阵何时能对应旋转、翻折。进一步说明对合阵对应翻折当且仅当 U 是 $n-1$ 维，V 是一维，且两者相互垂直。

于是，通过对合阵，我们得到了保持原点不动的翻折的矩阵表示特点，最后，让我们回到等距变换：

命题 1.40 (正交-几何性质)

A 是正交阵等价于 A 是保持原点不动的等距变换。

在 n 维空间中的 n 个相互垂直的单位向量称为**标准正交基**，例如 $\{e_i\}$ 就是一组标准正交基。而几何角度，A 是正交阵等价于其将任何一组标准正交基映射到另一组标准正交基，由此可进一步说明它与等距变换的等价性。

若等距变换额外要求保定向，则称为刚体变换，而保原点不变的刚体变换就是转轴过原点的旋转，对应行列式为 1 的正交阵。记所有 n 维正交阵构成 O_n，所有 n 维行列式为 1 的正交阵构成 SO_n，则它们都在矩阵乘法下构成群，分别称为**正交群**与**特殊正交群**。

正交方阵还可以衍生出许多几何性质，例如对称阵就是某组标准正交基下的对角阵，后续课程中将会学习。

1.4 行列变换技巧

比起前三节的标题，这节的标题看起来颇为无趣，但却是操作性最强的一种矩阵看法：将矩阵看作一**系列行变换或列变换的组合**。作为计算数学人，这样的视角事实上是最常用的，也对应做题的一种简单粗暴的方式—— “打洞”。

1.4.1 初等变换阵

记 E_{ii} 为只有第 i 行第 j 列为 1，其他为 0 的方阵：

定义 1.41 (初等变换阵)

以下三种方阵合称初等变换阵：

1. $S_{ij} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}, i \neq j$
2. $D_i(\lambda) = I + (\lambda - 1)E_{ii}, \lambda \neq 0$
3. $T_{ij}(\lambda) = I + \lambda E_{ij}, i \neq j$

根据之前的定义，S_{ij} 符合置换阵的要求，且其对应的恰好是两个元素的交换 [称为**对换**]；$D_i(\lambda)$ 是只有一个元素非 1 的对角阵，为了方便以后的讨论，我们用 $D_i(0)$ 表示 $\lambda = 0$ 的情况，虽然此时其不可逆，并非初等变换阵；$T_{ij}(\lambda)$ 则是对角线之外只有一个元素非零的矩阵，其必然是三角阵。

对某个矩阵 X，左乘矩阵可以视为进行一些行变换，而右乘矩阵可以视为进行一些列变换，对于初等变换阵，这样的行/列变换都有简单的表达：

1. 左乘 S_{ij} 为交换 X 的 i,j 两行，右乘 S_{ij} 为交换 X 的 i,j 两列；
2. 左乘 $D_i(\lambda)$ 为将 X 的第 i 行放大 λ 倍，右乘 $D_i(\lambda)$ 为将 X 的第 i 列放大 λ 倍；
3. 左乘 $T_{ij}(\lambda)$ 为将 X 的第 i 行增加第 j 行的 λ 倍，右乘 $T_{ij}(\lambda)$ 为将 X 的第 j 列增加第 i 列的 λ 倍。

注意 $T_{ij}(\lambda)$ 的左右乘改变的行/列是相反的。

之所以称它们是初等变换阵，是因为它们对应求解线性方程组时进行的初等变换。从这个角度来说，A 可逆时求解 $Ax = b$ 的过程即是用一系列初等变换阵合成 A^{-1}，随即得到 $x = A^{-1}b$ [此看法就是高斯消元法求解线性方程组的过程]。此外，初等变换阵之间并不是独立的，S_{ij} 可以被其他初等变换组合而成，即 $S_{ij} = D_i(-1)T_{ij}(-1)T_{ji}(1)T_{ij}(-1)$。

此组合其实与交换两整型变量值时利用代码 $a = a + b; b = a - b; a = a - b;$ 完全类似。

为了更好刻画它们的性质，我们先来计算它们的逆与转置：

命题 1.42 (初等变换阵的封闭性)

1. $S_{ij}^{-1} = S_{ij}^T = S_{ij}$
2. $D_i(\lambda)^T = D_i(\lambda), D_i(\lambda)^{-1} = D_i(\frac{1}{\lambda})$
3. $T_{ij}(\lambda)^T = T_{ji}(\lambda), T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$

对于可逆矩阵看作一系列变换，从数学上是：

命题 1.43 (可逆矩阵的分解)

任何可逆矩阵可以分解为一系列初等变换阵的乘积。

从而，Hermite 标准形也代表着，任何矩阵都可以通过初等行、列变换得到 $\begin{pmatrix} I & O \\ O & O \end{pmatrix}$ 的形式。另一方面，出于可逆矩阵的分解，相似也可以通过行列变换一步步进行，下方的以 P 相似指的是 $P^{-1}XP$：

1. 以 S_{ij} 相似，相当于交换 ij 两行、ij 两列；
2. 以 $D_i(\lambda)$ 相似，相当于第 i 行缩小 λ 倍，将第 i 列放大 λ 倍；
3. 以 $T_{ij}(\lambda)$ 相似，相当于第 i 行减第 j 行的 λ 倍，第 j 列加第 i 列的 λ 倍。

虽然用文字写出这些变换看起来并不简洁，对一些具体矩阵，这样的直观操作是用处很大的，我们以**相合**作为例子：

定义 1.44 (相合)

对 n 阶方阵 A, B，若存在可逆矩阵 P 使得 $B = P^TAP$，则称 A, B 相合。

例题 1.11 任何 n 阶对称阵都可以相合为对角元只有 $1, -1, 0$ 的对角阵，且 $1, -1, 0$ 个数唯一确定。1 的个数称为**正惯性指数**，-1 的个数称为**负惯性指数**。

提示：直接通过操作将非对角元全部消除，然后归一化对角阵。

1.4.2 分块矩阵与打洞

前面介绍的是对单独行列的操作，而实际上，我们最常用的是对分块矩阵进行行列变换。就以分块二阶矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 为例，在 A, D 都是方阵时，它对应的“分块初等变换阵”可以是：

$$\begin{pmatrix} M & O \\ O & I \end{pmatrix}, \begin{pmatrix} I & O \\ O & M \end{pmatrix}, \begin{pmatrix} I & M \\ O & I \end{pmatrix}, \begin{pmatrix} I & O \\ M & I \end{pmatrix}$$

由于 A, D 阶数未必一致，S_{ij} 类型的交换一般难以进行，而后两者运用较多，尤其是 $T_{ij}(\lambda)$ 类型的分块初等变换阵，它可以用来像消去元素一样消去分块。

例如，当 A 可逆时，有

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ O & I \end{pmatrix} = \begin{pmatrix} A & O \\ C & D - CA^{-1}B \end{pmatrix}$$

因此利用 Laplace 展开可知 $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det(D - CA^{-1}B)$。

运用分块矩阵进行乘法有时可以省去对行列变换的繁琐说明，但也可能导致过程更加抽象。

例题 1.12 任何幂等方阵都可以写成 $P^{-1} \begin{pmatrix} I & O \\ O & O \end{pmatrix} P$ 的形式。

提示：先写为 $P_0 \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q_0$，再根据幂等条件分块构造。

例题 1.13 $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times m}$，则

$$\lambda^{n-m} \det(\lambda I - AB) = \det(\lambda I - BA)$$

从纯矩阵论角度证明 Jordan 标准形时的最重要一步也可以由行列变换直接得到：

命题 1.45 (幂零阵的性质)

若存在 k 使得 $A^k = O$，则 A 称为幂零阵。对任何幂零方阵 A 都存在 P 使得 $P^{-1}AP$ 能写成分块矩阵 (B_{ij})，其中 B_{ii} 是方阵，且仅当 $j = i + 1$ 时 $B_{ij} \neq O$。

当然，分块矩阵方法是一种非常操作性的方法，这也就导致其中存在各种技巧。在掌握线性空间的语言后，很多矩阵论的技巧就显得没有那么重要了，然而，作为基本操作方式，行列变换的熟练度仍然是十分关键的。

1.4.3 低秩分解

在这一节，我们来讨论一种特殊的矩阵，低秩矩阵。秩在某种意义上可以代表矩阵的信息量，是由于：

命题 1.46 (低秩分解)

若矩阵 $A \in \mathbb{R}^{m\times n}$ 的秩为 r，则存在 $P \in \mathbb{R}^{m\times r}, Q \in \mathbb{R}^{r\times n}$，使得 $A = PQ$。

这意味着，若 r 相比 m, n 很小，只需要存储 $(m+n)r$ 个元素就能刻画一个 mn 个元素的矩阵。当 $r = 1$ 时，A 即可以分解为两个向量的乘积。

这个结论还有更多的用途，例如，当 A 为方阵时，假设其可以分解为 PQ，结合上一部分的定理立刻得到 A 的特征值恰为 QP 的特征值增添 $n-r$ 个 0。下面观察一个课上说过的例子。

例题 1.14 计算 $a_{ij}=b_i+c_j$ 的方阵 A 的特征值。

提示：$\operatorname{rank}(A)=2$，找到合适的矩阵即可进行分解。

低秩分解还有些其他的作用，如下方的计算例子。

例题 1.15 计算 $a_{ij}=b_ic_j$ 的方阵 A 的次方 A^n。

提示：$\operatorname{rank}(A)=1$，因此可直接分解为向量。

由于低秩矩阵是易于计算、存储的，在一个矩阵上加上低秩矩阵的行为称为**低秩校正**。篇幅所限，我们此处只考虑它的一个简单的性质：

命题 1.47 (低秩校正的逆)

A 可逆且 $1-v^TA^{-1}u\neq 0$ 时

$$(A+uv^T)^{-1}=A^{-1}-\frac{A^{-1}uv^TA^{-1}}{1-v^TA^{-1}u}$$

这说明**秩一校正的逆是逆的秩一校正**。

事实上由于 $E=-(A+E)\big((A+E)^{-1}-A^{-1}\big)A$ 有：

命题 1.48 (低秩校正的性质)

当 A 与 $A+E$ 可逆时，$\operatorname{rank}\big((A+E)^{-1}-A^{-1}\big)\leq \operatorname{rank}E$。

1.5 矩阵学习指南

1.5.1 等价关系与标准形思想

在前面的部分中，我们留下了一个问题——商集是商去了什么？对这个问题的解答关乎着矩阵论中最重要的思想之一，因此放到这一总结性的环节中。

我们先定义集合上的等价关系：

定义 1.49 (关系、等价关系)

对集合 A，$A\times A$ 的任何一个子集 R 称为集合 A 上的一个关系。若 $(a,b)\in R$，记作 aRb，否则记作 $a\bar{R}b$。

若一个关系满足

1. 自反性 $\forall a\in A, aRa$；
2. 对称性 $\forall a,b\in a, aRb\Rightarrow bRa$；
3. 传递性 $\forall a,b,c\in a, aRb, bRc\Rightarrow aRc$；

则其称为集合 A 上的一个等价关系。

事实上两个集合间也可以类似定义 $A\times B$ 的任何一个子集为关系，其可以看作映射在某种程度上的推广，不过我们此处只关心某个集合上的等价关系。

顾名思义，等价关系代表着某种意义上的**等价性**。

例题 1.16

- 任何集合上，相等是一种等价关系。
- 平面上所有三角形集合的全等、相似是等价关系。
- 对任何正整数 n，整数集合模 n 同余是等价关系。

对矩阵来说，我们有：

命题 1.50 (矩阵中的等价关系)

对 n 阶方阵集合，相抵、相似、相合都是等价关系。

等价关系的最大作用是用于**分类**，准确来说，我们的任何分类都是在定义等价关系：

定义 1.51 (等价类)

对集合 A，根据任何一个等价关系 R，都可以唯一确定一个商集 $\{A_i\}, i \in I$，使得

$$\forall a, b \in A_i, aRb; \forall a \in A_i, b \in A_j, i \neq j, a\bar{R}b$$

此外，反之则有：

命题 1.52 (商集确实等价关系)

对集合 A 的任何一个商集 $\{A_i\}, i \in I$，定义关系 R 满足

$$\forall a, b \in A_i, aRb; \forall a \in A_i, b \in A_j, i \neq j, a\bar{R}b$$

则其构成等价关系。

于是，等价关系与商集一一对应，根据等价关系得到的商集称为它划分出的**等价类**，元素 a 所在的等价类记作 $[a]$，从每个等价类中各取出一个元素称为这个等价类的一族**代表元**。

例题 1.17

1. 对于整数上模 n 同余的等价关系，其等价类的一族代表元可以是 $0, 1, \ldots, n-1$，从而等价类可写成 $[0], [1], \ldots, [n-1]$。
2. 对于 n 阶方阵阵上相抵的等价关系，其等价类的一族代表元可以是 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}, r = 0, 1, \ldots, n$。
3. 对于平面三角形全等的等价关系，其等价类的一族代表元可以是所有满足 $0 < a \leq b \leq c, a+b > c$ 的向量 (a, b, c) 所确定的三角形各一个。

在我们讨论很多矩阵的问题时，我们其实都只关注它的某些属性，从而同一等价类的元素都可以使用，例如如果我们只关心矩阵的秩，其整个相抵等价类都是不影响的，而如果我们只关心特征值，整个相似等价类都满足特征值一致。

于是，我们总会试图**在等价类中找到形式最简单的代表元**，这就是矩阵的标准形。在相抵等价类下，最简单的代表元就是 Hermite 标准形，而相似下就成了 Jordan 标准形。之前提到的实对称方阵能唯一相合成的 $-1, 1, 0$ 为对角元的对角阵，也就称为实对称方阵的相合标准形。

对于具体的题目，分析其关注矩阵的哪些信息，然后选取合适的标准形，往往可以大幅简化过程。

例题 1.18 讨论与某可对角化矩阵 A 可交换的复矩阵构成的复线性空间维数。

1.5.2 暴力计算 vs 技巧性

对于这个话题，我们先来看一个有趣的题目，也是线性代数 B1 的经典习题。

例题 1.19 对 n 阶方阵 A, B，证明 $AB - BA = I$ 无解。

注意到 tr 的可交换性，这题可以用一句话说完。然而，这是一个高度技巧性的证明，而技巧往往是精妙却难以想到的。如果不是这题的特殊构造，或许问题就不会那么简单，如：

例题 1.20 对怎样的方阵 X，$AB - BA = X$ 有解？

分析可以发现，哪怕对二阶的情况，这也并不是一个平凡的问题。我们就二阶情况进行一些分析：由于 A, B, X 同时以 P 相似不会影响结论，不妨**直接考虑 X 的 Jornda 标准形**。又由于 $\mathrm{tr}(X)$ 必须为 0，事实上只可能是 $\begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix}$ 或 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$。前者得到的方程组是 (由于迹为 0 可减少一个方程)

$$\begin{cases} a_{12}b_{21} - a_{21}b_{12} = x \\ alpha_{11}b_{12} + a_{12}b_{22} - a_{12}b_{11} - a_{22}b_{12} = 0 \\ alpha_{21}b_{11} + a_{22}b_{21} - a_{11}b_{21} - a_{21}b_{22} = 0 \end{cases}$$

只需 $A = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ 即可构造出解。类似地，后者取 $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 即可。也就是说，二阶时只要 $\mathrm{tr}(X) = 0$，解即存在。对高阶时的分析仍然需要通过这样的方程组进行，此处不详细讨论。

从这个问题中可以看出，技巧往往是在题目具有特殊性时可以大幅简化，因此是必要的。但对于一般的情况，或是没有头绪的问题，暴力计算进行探索——哪怕是被大家吐槽多年的“打洞”——仍然是首要的处理方式。尤其是对于偏数值的方向可能与遇到的一些估计问题，就更需要精确的计算。3Blue1Brown 的一期视频很好地说明了这个观点[1]，这里推荐大家去看一看。

1.5.3 当我们谈论矩阵时

在之前的四节中，我们分别探讨了分析角度关注元素的矩阵、代数角度作为同态或同构的矩阵、几何角度作为仿射的矩阵与数值计算角度作为行列变换的矩阵。这些都是关注矩阵的不同视角，也能对不同的问题提供直观。

当我们谈论矩阵时，这些观点事实上都在我们谈论的范畴内，而关注它们的统一性则是理解矩阵的很好方式。本学期中，对不同类型的特殊矩阵，用不同视角进行处理，尤其是进行几何的感受，是非常重要的。当然，更多空间的问题并不是有限维，也无法建立矩阵的表示，但用有限维的矩阵例子感受直观仍是一个不错的具象方式——后续的**泛函分析**中，我们会接触更加分析化的无穷维空间，也将面临更抽象的问题，不过已经不在本学期讨论的范畴了。

1.6 补充阅读：集合论基础

1.6.1 集合悖论

介绍公理化的集合论时，我们需要回答的第一个问题是，集合是什么？

在高中时，对集合的定义只有很含糊的“描述下的某个确定集体”(或者说，任何一个公式 P，都存在形如 $\{x \mid x\text{满足}P\}$ 的集合)，但这样的定义会导致很多问题，例如著名的**罗素悖论**：

$$S = \{A \mid A \notin A\}$$

[1] www.bilibili.com/video/BV1yP4y1P7mo

假设集合 S 是所有不属于自身的集合构成的集合（注意集合也可以构成集合），它是否属于自身呢？[2]若 $S\in S$，根据定义可知 $S\notin S$；而若 $S\notin S$，根据定义可知 $S\in S$，这就导致了悖论的产生。

另一个经典悖论是**康托悖论**，假设 V 是所有集合构成的集合，V 应该具有最大的势，但 V 的幂集会比 V 具有更大的势，从而导致矛盾[3]。这也就意味着，所有集合构成的集合**并不存在**。

为了解决这些悖论，我们必须对集合进行更严格的规定。**策梅洛**和**弗伦克尔**提出了一套公理系统以解决这些问题，这就是常说的 **ZF 公理系统**。ZF 公理系统共有 8 条，接下来便从这些公理开始重新看待集合系统的生成。

1.6.2 集合的生成

数学中，“公理”指的是直接承认而不加证明的结论。从每一个公理出发，又可以得到一系列的定理，从而建立出整个理论体系。反过来说，如果要进行公理出发的推导，需要先抛弃“理所应当”的印象。一个简单的例子是：若 A,B 是集合，假设没有公理，我们没有充分的理由断言 $A\cap B$ 也是集合，因为它来自定义 $\{x\mid x\in A \text{ and } x\in B\}$，而由于我们已经否定了“一切描述性的定义都是集合”，这个描述未必能代表集合。

接下来，我们便从公理出发，重建熟悉的集合体系。

公理 1.53 (外延公理 ZF1)

两集合 A,B 相等，当且仅当

$$\forall x, x\in A\leftrightarrow x\in B$$

♡

可以看出，这个公理就是我们习惯的“元素相等代表集合相等”，此外，它也保证了集合中元素的唯一性，因为根据定义，$\{1,1\}$ 和 $\{1\}$ 表示的集合是完全相同的 (但注意，从第一个公理中并不能看出 $\{1\}$ 真的是一个集合)。

类似我们给出包含和真包含的定义：

定义 1.54 (包含)

若 A,B 是集合，$\forall x, x\in A\rightarrow x\in B$，则称 A 包含于 B，记作 $A\subset B$；而若还有 $A\neq B$，则称 A 真包含于 B，记作 $A\subsetneq B$。

♣

公理 1.55 (分离公理 ZF2)

若 A 是集合，对任何公式[a] P，存在集合 B 使得 $B=\{x\in A\mid x\text{满足}P\}$。

[a]此公理更严谨的叫法是分离公理模式，这是一个数理逻辑中的概念。这里的公式同理，可以理解为任何一句判断。

♡

这个公理与之前含糊的“描述”非常类似，只是增加了必须在某个集合子集中的规定。这个定义事实上意味着，若 A 是集合，A 的任何一部分也都是集合，这也是我们熟悉的“子集”。通过这个公理，我们也能得出空集存在：

[2]此悖论的另一个名字是理发师悖论，因为这个问题可以完全对应“假设一个理发师只为不给自己理发的人理发，它是否给自己理发”。

[3]这里涉及的名词的详细解释见本章后续。

定义 1.56 (空集)

对任何集合 A，定义空集 $\varnothing = \{x \in A \mid x \neq x\}$。♣

练习 1.1 若 A, B 是集合，记 $\varnothing_A = \{x \in A \mid x \neq x\}, \varnothing_B = \{x \in B \mid x \neq x\}$，证明 $\varnothing_A = \varnothing_B$（这也就说明了**空集是唯一的**）。

解 由于 $x \in \varnothing_A \Leftrightarrow x \in A, x \neq x$，对任何元素 x，$x \in \varnothing_A$ 恒假，因此 $\varnothing_A \subset \varnothing_B$，同理 $\varnothing_B \subset \varnothing_A$，于是 $\varnothing_A = \varnothing_B$。[4]

此外，交集运算和补集也可以进行：

定义 1.57 (交集)

对有限个集合 $A_1, \ldots, A_n$，定义它们的交集 $A_1 \cap \cdots \cap A_n$ 为

$$\{x \in A_1 \mid x \in A_2, \ldots, x \in A_n\}$$

♣

定义 1.58 (补集/差集)

对两个集合 A, B，定义 B 对 A 的补集 (或 A 对 B 的差集)A

$beta = \{x \in A \mid x \notin B\}$。♣

很遗憾的是，分离公理虽然很强大，但只给了我们“缩小”集合的方式。想要“扩大”集合，还需要接下来的三条公理：

公理 1.59 (无序对公理 ZF3)

若 A, B 是集合，则 $\{A, B\}$ 也是集合。♡

练习 1.2 若 A 是集合，则 $\{A\}$ 也是集合。

解 由无序对公理 $\{A, A\}$ 是集合，而由外延公理 $\{A, A\} = \{A\}$，从而得证。

通过无序对公理，可以生成非常重要的**有序对**：

定义 1.60 (有序对)

若 A, B 是集合，定义 A, B 的有序对 $(A, B) = \{\{A\}, \{A, B\}\}$。♣

根据高中熟悉的知识，有序对应该满足 $(A, B) = (C, D)$ 当且仅当 $A = B$ 且 $C = D$。下面验证这个结论：

定理 1.61 (有序对的相等)

$(A, B) = (C, D) \Leftrightarrow A = C$ and $B = D$ ♡

证明 当 $A \neq B$ 时，(A, B) 是一个两元素的集合，且两元素是元素个数分别为 1 或 2 的集合；当 $A = B$ 时，根据外延公理 $(A, B) = \{\{A\}\}$，是一个单元素的集合。由此可以先得到 $A = B \Leftrightarrow C = D$，再分两种情况讨论即可。

除了无序对外，第二种扩大集合的方式是并集：

[4] 这个证明中涉及了 x 假则 $x \to P$ 为真，本质仍然是是数理逻辑的知识。集合的定义中也大量包含一阶谓词逻辑的推导，限于篇幅忽略此部分，用我们习惯的语言进行直觉性的描述。

公理 1.62 (并集公理 ZF4)

任给集合 A，$\{x \mid \exists B, x \in B, B \in A\}$ 也是集合，记作 $\cup A$。

我们用一个练习来了解并集公理的含义：

练习 1.3 若 A, B, C 是集合，证明 $\cup\{A, B, C\}$ 是集合，且其代表 $\{x \mid x \in A \text{ or } x \in B \text{ or } x \in C\}$。

解 由无序对公理，$\{A, B\}$ 与 $\{B, C\}$ 是集合，从而 $\{\{A, B\}, \{B, C\}\}$ 是集合。记 $S = \cup\{\{A, B\}, \{B, C\}\}$，则 S 中所有元素为：$A \in \{A, B\}$、$B \in \{A, B\}$、$B \in \{B, C\}$ 与 $C \in \{B, C\}$，从而 $S = \{A, B, C\}$，而再次根据并集公理，$\cup\{A, B, C\}$ 存在，类似推导可知 $\cup\{A, B, C\}$ 即为我们熟悉 A, B, C 之并。

我们可以发现，$\cup A$ 即为 A 的所有**元素**的并集，而用上方类似的方法，从无序对公理与并集公理出发，假设 $A_1, \ldots, A_n$ 是集合，可以得到 $\{A_1, \ldots, A_n\}$ 也是一个集合，于是有定义

定义 1.63 (并集)

对有限个集合 $A_1, \ldots, A_n$，定义它们的并集 $A_1 \cup \cdots \cup A_n$ 为 $\cup\{A_1, \ldots, A_n\}$。

最后，第三种扩大集合的方式是幂集，定义如下：

公理 1.64 (幂集公理 ZF5)

若 A 是集合，$\{B \mid B \subset A\}$ 也是一个集合，称为 A 的幂集，记作 $\mathcal{P}(A)$。

幂集公理也就代表着一个集合的所有子集也构成一个集合，如 $\{1, 2, 3\}$ 的幂集为

$$\{\varnothing, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$$

至此，我们熟悉的有限交、并、补与幂集运算都得到了严格的定义。但是，问题还远远没有解决。

1.6.3 存在与无穷

在介绍下一条公理前，我们先进行一个简单的练习：

练习 1.4 若一个集合 A 有 n 个元素，则 $\mathcal{P}(A)$ 有 2^n 个元素。

解 对任何子集 B，A 的每个元素有在其中出现与不在其中出现两种可能，因此所有子集的个数为 2^n，这就是幂集的元素个数。

由此可以看出，若一个集合元素个数有限，则其幂集元素个数也有限，于是可以先通过分离公理得到所有子集，再利用无序对公理与并集公理得到以所有子集为元素的集合。这样一来，**有限集幂集的存在性无需幂集公理**。

另一个颇有趣的问题是，我们上面的所有公理都是在**存在集合**的前提下进行讨论的。哪怕是空集的定义，都是先随意找了一个集合，再进行定义。然而，如果“一个集合都不存在”的话，整套讨论都是没有意义的。自然，我们可以用公理规定集合的存在性（这有时记为公理 ZF0），不过，由于下一条公理规定了一个更复杂集合的存在，ZF0 是可以被其他公理推出的——这个“更复杂的集合”就是自然数集。

定义 1.65 (自然数-集合论视角)

定义 0 为 $\varnothing$，且对任何 n，定义 $n+1$ 为 $n \cup \{n\}$（根据无序对公理与并集公理，这确实是一个集合）。

此定义事实上给每个自然数“赋予”了一个集合，可以逐一写出它们为 $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}$ 等等。在本节中我们写出 $0, 1, \dots$ 时，均表示这些集合。

公理 1.66 (无穷公理 ZF6)

满足 $\varnothing \in X$ 且 $x \in X \Rightarrow x \cup \{x\} \in X$ 的集合 X 存在。

根据之前给自然数赋予的集合，这个集合 X 事实上就代表着所有自然数。此外，有两件值得一提的事：第一，若不事先假定集合的存在性，无穷公理与自然数定义中的 $\varnothing$ 只是作为一个符号，只有在确定其他集合存在后，空集才得以定义；第二，无穷公理的关键是规定了“**所有**自然数”构成的集合存在，因为用之前的任何公理都只能推出有限自然数集合的存在性，而无法跨越至无穷。

在有了自然数集后，我们已经可以基本构造出所有想象中的数的集合了——从自然数出发，我们可以一步步生成实数乃至复数。接下来，我们可以通过一条公理解决悖论：

公理 1.67 (正则公理 ZF7)

任何非空集合 A，一定存在 $x \in A$，使得 $A \cap x = \varnothing$。

为了解释这条公理，我们来观察一个最常见的应用：

练习 1.5 对任何集合 A，$A \notin A$。

解 若 A 是空集，显然 $A \notin A$，否则假设 $A \in A$，根据之前推导，$\{A\}$ 也是集合，而 $\{A\}$ 的元素只有 A，但 $A \cap \{A\} = A \neq \varnothing$，与正则公理矛盾。

也就是说，正则公理的存在是为了消灭某些奇异的情况，保证了罗素悖论或康托悖论所规定出的描述根本不可能对应一个真实的集合。

至此，我们的集合公理体系似乎已经完美了。可是，仅有数的无穷集合仍然并不是我们想要的，例如，我们甚至无法说明熟悉的 $\{1, x, x^2, x^3, \dots\}$ 或者 $\{x, y, z\}$ 是一个集合——因为此处的符号 x 无法看作自然数集的元素或子集，而我们又没有无穷公理之外的方式扩展集合。由此，我们必须有某种“迁移”的方式，来真正得到各种有意义集合，这种“迁移”某种意义上就是映射。ZF 公理体系的最后一条公理[5]即与映射相关：

公理 1.68 (替换公理模式 ZF8)

对任何集合 A，定义对应规则 φ 使得任何 a 对应到 $\varphi(a)$，则所有 $\varphi(a)$ 构成一个集合。

值得注意的是，这里并没有要求 $\varphi(a)$ 是某个集合的子集，因此 φ 只是代表一套对应规则。此处暗示了，对应规则本身就具有一定的结构传递性。对群来说，对一个群 G 与映射 $fG \to H$，H 是任何一个集合，若 f 满足当 $f(g_1) = f(g_2), f(g_3) = f(g_4)$ 时 $f(g_1 g_2) = f(g_3 g_4)$，在 $f(G)$ 中定义乘法 $f(g_1)f(g_2) = f(g_1 g_2)$，则 $f(G)$ 构成一个群。

[5] ZF 公理体系共有八条公理，而现实常用的集合论公理体系事实上是 ZFC 公理体系，多出来的 C 取自 Choice，代表著名的**选择公理**。不过，由于这个公理相对复杂且网络上有充分资源，讲义中略去对其的介绍。

第 2 章 矩阵与行列式

2.1 习题讲解

2.1.1 第一周

作业 2.1 (P54 T10)

证明以下等式：

1. $$\begin{vmatrix} a & b & c & d \\ a & a+b & a+b+c & a+b+c+d \\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d \\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d \end{vmatrix} = a^4;$$
2. 设 $a_{ij} = a_{i,j-1} + a_{i-1,j}, i, j = 1, 2, \cdots, n$, 则
$$\begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 1 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = 1$$

♡

解

1.
$$\text{原式} \xlongequal{-r_1+r_2,-r_1+r_3,-r_1+r_4} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & 2a & 3a+2b & 4a+3b+2c \\ 0 & 3a & 6a+3b & 10a+6b+3c \end{vmatrix}$$
$$\xlongequal{-2r_2+r_3,-3r_2+r_4} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & 0 & a & 2a+b \\ 0 & 0 & 3a & 7a+3b \end{vmatrix} \xlongequal{-3r_3+r_4} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & 0 & a & 2a+b \\ 0 & 0 & 0 & a \end{vmatrix} = a^4$$

2.
$$\begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 1 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \xlongequal{-r_{i-1}+r_i,-c_{i-1}+c_i,2\leq i\leq n} \begin{vmatrix} 1 & a_{12}-1 & a_{13}-a_{12} & \cdots & a_{1n}-a_{1,n-1} \\ 0 & 1 & a_{12} & \cdots & a_{1,n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix}$$
$$= \cdots = \begin{vmatrix} 1 & a_{12}-1 & a_{13}-a_{12} & \cdots & a_{1n}-a_{1,n-1} \\ 0 & 1 & a_{12}-1 & \cdots & a_{1,n-1}-a_{1,n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1$$

作业 2.2 (P65-66 T2)

设 A, B, C 和 D 依次是由下表

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

中删去第 1、2、3、4 列而得到的三阶行列式，证明：

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 & 0 & 0 \\ a_2 & b_2 & c_2 & d_2 & 0 & 0 \\ a_3 & b_3 & c_3 & d_3 & 0 & 0 \\ 0 & 0 & a_1 & b_1 & c_1 & d_1 \\ 0 & 0 & a_2 & b_2 & c_2 & d_2 \\ 0 & 0 & a_3 & b_3 & c_3 & d_3 \end{vmatrix} = AD - BC$$

♡

解 将行列式按前三行作 Laplace 展开，得到原式为

$$\sum_{1\le j_1<j_2<j_3\le 6} (-1)^{1+2+3+j_1+j_2+j_3} A\begin{pmatrix} 1 & 2 & 3 \\ j_1 & j_2 & j_3 \end{pmatrix} A\begin{pmatrix} 4 & 5 & 6 \\ j_4 & j_5 & j_6 \end{pmatrix}$$

$$\begin{aligned} =&(-1)^{1+2+3+1+2+3}\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix} + (-1)^{1+2+4+1+2+3}\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} \cdot \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix} \\ =&AD - BC \end{aligned}$$

作业 2.3 (P65-66 T5)

给定 n 阶方阵 $A = (a_{ij})$，证明

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{21}-a_{11} & a_{22}-a_{12} & \cdots & a_{2n}-a_{1n} \\ a_{31}-a_{11} & a_{32}-a_{12} & \cdots & a_{3n}-a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1}-a_{11} & a_{n2}-a_{12} & \cdots & a_{nn}-a_{1n} \end{vmatrix} = \sum_{1\le i,j\le n} A_{ij}$$

其中 A_{ij} 是行列式 $\det A$ 中元素 a_{ij} 的代数余子式，$1 \le i, j \le n$。

♡

解 在不影响行列式值时升阶为：

$$\begin{vmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & a_{21}-a_{11} & a_{22}-a_{12} & \cdots & a_{2,n-1}-a_{1,n-1} & a_{2n}-a_{1n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n-1,1}-a_{11} & a_{n-1,2}-a_{12} & \cdots & a_{n-1,n-1}-a_{1,n-1} & a_{n-1,n}-a_{1n} \\ 0 & a_{n1}-a_{11} & a_{n2}-a_{12} & \cdots & a_{n,n-1}-a_{1,n-1} & a_{nn}-a_{1n} \end{vmatrix}$$

$$\xlongequal{r_1+r_i,3\leq i\leq n}\begin{vmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 1 & a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix}$$

将此时的行列式按第 1 列 Laplace 展开，再按第二行 Laplace 展开后即可得证.

作业 2.4 (P72 T6)

设线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases}$$

的系数行列式 $\Delta\neq 0$。利用 $n+1$ 阶行列式

$$\begin{vmatrix} b_i & a_{i1} & a_{i2} & \cdots & a_{in} \\ b_1 & a_{11} & a_{12} & \cdots & a_{1n} \\ b_2 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=0,\quad i=1,2,\cdots,n$$

证明：$(\frac{\Delta_1}{\Delta},\frac{\Delta_2}{\Delta},\cdots,\frac{\Delta_n}{\Delta})$ 是此方程组的解.

解 将该行列式按第 1 行 Laplace 展开可得其为

$$\begin{aligned} & b_i\Delta+a_{i1}(-1)^{1+2}\Delta_1+a_{i2}(-1)^{1+3+1}\Delta_2+\cdots+a_{in}(-1)^{1+(n+1)+(n-1)}\Delta_n \\ &= b_i\Delta-a_{i1}\Delta_1-a_{i2}\Delta_2-\cdots-a_{in}\Delta_n \end{aligned}$$

又由其有两行相同，行列式为 0，即

$$a_{i1}\frac{\Delta_1}{\Delta}+a_{i2}\frac{\Delta_2}{\Delta}+\cdots+a_{in}\frac{\Delta_n}{\Delta}=b_i$$

由此 $(\frac{\Delta_1}{\Delta},\frac{\Delta_2}{\Delta},\cdots,\frac{\Delta_n}{\Delta})$ 是方程组的解。

作业 2.5 (P83 T1(5))

计算行列式

$$\begin{vmatrix} x & a & a & a & \cdots & a \\ -a & x & a & a & \cdots & a \\ -a & -a & x & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & a \\ -a & -a & \cdots & -a & x & a \\ -a & -a & \cdots & -a & -a & x \end{vmatrix}$$

解 解法一:

$$\Delta_n=\begin{vmatrix} x & a & a & a & \cdots & a\\ -a & x & a & a & \cdots & a\\ -a & -a & x & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & a\\ -a & -a & \cdots & -a & x & a\\ -a & -a & \cdots & -a & -a & x \end{vmatrix}=\begin{vmatrix} x & a & a & a & \cdots & a+0\\ -a & x & a & a & \cdots & a+0\\ -a & -a & x & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & a+0\\ -a & -a & \cdots & -a & x & a+0\\ -a & -a & \cdots & -a & -a & a+(x-a) \end{vmatrix}$$

$$=\begin{vmatrix} x & a & a & a & \cdots & a\\ -a & x & a & a & \cdots & a\\ -a & -a & x & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & a\\ -a & -a & \cdots & -a & x & a\\ -a & -a & \cdots & -a & -a & a \end{vmatrix}+\begin{vmatrix} x & a & a & a & \cdots & 0\\ -a & x & a & a & \cdots & 0\\ -a & -a & x & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & 0\\ -a & -a & \cdots & -a & x & 0\\ -a & -a & \cdots & -a & -a & x-a \end{vmatrix}$$

对上式右端第一个行列式，用 -1 乘以它的第 n 行，然后加到其它各行；对第二个行列式的第 n 列作 Laplace 展开，得到

$$\begin{aligned}\Delta_n&=\begin{vmatrix} x+a & 2a & 2a & 2a & \cdots & 0\\ 0 & x+a & 2a & 2a & \cdots & 0\\ 0 & 0 & x+a & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & 0\\ 0 & 0 & \cdots & 0 & x+a & 0\\ -a & -a & \cdots & -a & -a & a \end{vmatrix}+(x-a)\Delta_{n-1}\\ &=a(x+a)^{n-1}+(x-a)\Delta_{n-1}\end{aligned}$$

考虑 Δ_n 的转置行列式，同理得到

$$\Delta_n=-a(x-a)^{n-1}+(x+a)\Delta_{n-1}$$

解得

$$\Delta_n=\frac{1}{2}\big((x+a)^n+(x-a)^n\big)$$

解法二:

记

$$D_n(t)=\begin{vmatrix} x+t & a+t & a+t & a+t & \dots & a+t\\ -a+t & x+t & a+t & a+t & \dots & a+t\\ -a+t & -a+t & x+t & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \ddots & a+t\\ -a+t & -a+t & \cdots & -a+t & x+t & a+t\\ -a+t & -a+t & \cdots & -a+t & -a+t & x+t \end{vmatrix}$$

可以通过添加一行 $t,t,\dots,t$ 再添加一列 $1,0,\dots,0$ 说明其是 t 的一次函数，又因为 $D_n(0)=\Delta_n$，有

$D_n(s) = Ms + \Delta_n$。$t = a$ 时，$D_n(a) = (x+a)^n$；$t = -a$ 时，$D_n(-a) = (x-a)^n$，故

$$\begin{cases} \Delta_n + Ma = (x+a)^n \\ \Delta_n - Ma = (x-a)^n \end{cases}$$

即得 $\Delta_n = \frac{1}{2}\big((x+a)^n + (x-a)^n\big)$。

作业 2.6 (P83 T1(6))

计算行列式

$$\begin{vmatrix} a_1 & -a_2 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & -a_3 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{n-1} & -a_n \\ 1 & 1 & \cdots & 1 & 1 & 1+a_n \end{vmatrix}$$

解

$$\begin{vmatrix} a_1 & -a_2 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & -a_3 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{n-1} & -a_n \\ 1 & 1 & \cdots & 1 & 1 & 1+a_n \end{vmatrix} \xlongequal[]{\frac{a_{i+1}}{a_i}c_i + c_{i+1}, 1\le i\le n-1}$$

$$\begin{vmatrix} a_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{n-1} & 0 \\ 1 & \sum_{k=1}^{2}\frac{a_2}{a_k} & \cdots & \sum_{k=1}^{n-2}\frac{a_{n-2}}{a_k} & \sum_{k=1}^{n-1}\frac{a_{n-1}}{a_k} & a_n + \sum_{k=1}^{n}\frac{a_n}{a_k} \end{vmatrix} = \prod_{k=1}^{n} a_k\left(1 + \sum_{k=1}^{n}\frac{1}{a_k}\right)$$

作业 2.7 (P83 T1(10))

计算行列式

$$\begin{vmatrix} 1+a_1+x_1 & a_1+x_2 & \cdots & a_1+x_n \\ a_2+x_1 & 1+a_2+x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}+x_n \\ a_n+x_1 & \cdots & a_n+x_{n-1} & 1+a_n+x_n \end{vmatrix}$$

解

$$\begin{vmatrix} 1+a_1+x_1 & a_1+x_2 & \cdots & a_1+x_n \\ a_2+x_1 & 1+a_2+x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}+x_n \\ a_n+x_1 & \dots & a_n+x_{n-1} & 1+a_n+x_n \end{vmatrix}$$

$$= \left| I_n + \begin{pmatrix} a_1+x_1 & a_1+x_2 & \dots & a_1+x_n \\ a_2+x_1 & a_2+x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}+x_n \\ a_n+x_1 & \dots & a_n+x_{n-1} & a_n+x_n \end{pmatrix} \right|$$

$$= \left| I_n + \begin{pmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix} \right| = \left| I_2 + \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} \right|$$

$$= \begin{vmatrix} 1+\sum_{k=1}^n a_k & n \\ \sum_{k=1}^n a_k x_k & 1+\sum_{k=1}^n x_k \end{vmatrix} = \left(1+\sum_{k=1}^n a_k\right)\left(1+\sum_{k=1}^n x_k\right) - n\sum_{k=1}^n a_k x_k$$

作业 2.8 (P83 T1(18))

计算行列式

$$\begin{vmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix}$$

♡

解 原式加边后得到

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1 & 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix}$$

于是进一步有

$$\xlongequal{-c_1+c_i,2\le i\le n} \begin{vmatrix} 1 & -1 & -1 & \cdots & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} \xlongequal{\text{按第 1 行展开}} \begin{vmatrix} 2 & 0 & 0 & \cdots & 0 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} - \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix}$$

$$= \left(2\prod_{k=1}^n x_k - \prod_{k=1}^n (x_k-1)\right) \prod_{1\le i<j\le n} (x_j - x_i)$$

这里最后一步是在使用范德蒙德行列式的结论后合并化简的结果。

2.1.2 第二周

作业 2.9 (P66 T8)

给定 n 阶方阵 $A=(a_{ij})$，证明：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & z \end{vmatrix} = z\det A - \sum_{1\le i,j\le n} A_{ij}x_iy_j$$

其中 A_{ij} 是行列式 $\det A$ 的元素 a_{ij} 的代数余子式, $1\le i,j\le n$。

解 将行列式依次按第 $n+1$ 行与第 n 列进行 Laplace 展开得到原式为

$$y_1(-1)^{n+2}\begin{vmatrix} a_{12} & \cdots & a_{1n} & x_1 \\ a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n2} & \cdots & a_{nn} & x_n \end{vmatrix} + \cdots + y_n(-1)^{2n+1}\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & x_1 \\ a_{21} & \cdots & a_{2,n-1} & x_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & x_n \end{vmatrix} + z\det A$$

再对最后一列展开即得结论。

作业 2.10 (P99 T8)

找出所有 2 阶幂零方阵.

解 设

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

注意到 $A^k, k=1,2,\dots$ 的秩单调递减，且由书上结论，某一位置 $A^k=A^{k+1}$ 时，后续均有 $\operatorname{rank}(A^n)=\operatorname{rank}(A^k), \forall n>k$。

又 A 的秩若为 2，则 $\det A$ 非 0，从而 $\operatorname{rank}(A^k)=2, \forall k$。故 A 的秩只能为 0 或者 1，并且 $\operatorname{rank}(A^2)=0$，即 $A^2=O$。故

$$A^2=\begin{pmatrix} a_{11}^2+a_{12}a_{21} & a_{12}(a_{11}+a_{22}) \\ a_{21}(a_{11}+a_{22}) & a_{12}a_{21}+a_{22}^2 \end{pmatrix}=O$$

故 $a_{22}^2=a_{11}^2$。

若 $a_{11}=a_{22}$：

1. 若 $a_{11}=a_{22}=0$，满足 $a_{11}=-a_{22}$ 条件，见下方讨论。
2. 若 $a_{11}=a_{22}\neq 0$，则 $a_{12}=a_{21}=0$，此时由 $a_{11}^2+a_{12}a_{21}=0$ 得 $a_{11}=0$，矛盾。

若 $a_{11}=-a_{22}$，只需满足 $a_{11}^2+a_{12}a_{21}=0$ 即可。综上可得到

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix}, a_{11}^2+a_{12}a_{21}=0$$

作业 2.11 (P99 T9)

如果方阵 A 适合 $A^2=I$, 则 A 称为对合。求出所有 2 阶对合方阵.

解 设

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

则有

$$A^2=\begin{pmatrix} a_{11}^2+a_{12}a_{21} & a_{12}(a_{11}+a_{22}) \\ a_{21}(a_{11}+a_{22}) & a_{12}a_{21}+a_{22}^2 \end{pmatrix}=I$$

故 $a_{22}^2=a_{11}^2$。

若 $a_{11}=a_{22}$：

1. 若 $a_{11}=a_{22}=0$，满足 $a_{11}=-a_{22}$ 条件，见下方讨论。
2. 若 $a_{11}=a_{22}\neq 0$，则 $a_{12}=a_{21}=0$, 此时由 $a_{11}^2+a_{12}a_{21}=1$, 得出 $a_{11}=a_{22}=\pm 1$。

若 $a_{11}=-a_{22}$，只需满足 $a_{11}^2+a_{12}a_{21}=1$ 即可。

综上，$A=\pm I$ 或

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix}, a_{11}^2+a_{12}a_{21}=1$$

作业 2.12 (P105 T7)

适合 $AA^T=I_{(n)}=A^TA$ 的 n 阶实方阵 A 称为正交的。证明：

1. 正交方阵的行列式等于 ± 1；
2. 位于正交方阵的 k 个行上的所有 k 阶子式的平方和等于 1，$k=1,2,\cdots,n$。

解

1.
$$\det(AA^T)=\det(I_{(n)})=1=\det A\cdot\det A^T=(\det A)^2\Rightarrow \det A=\pm 1$$

2. 记 $C=AA^T=I_{(n)}$，有
$$1=C\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}=\sum_{1\le m_1<m_2<\cdots<m_k\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ m_1 & m_2 & \cdots & m_k \end{pmatrix}^2$$

作业 2.13 (P98 T3)

求出所有和方阵 A 可交换的方阵：

1. $A=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{pmatrix}$
2. $A=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

解 直接列方程组求解可得结果为：

1. $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}-3a_{12} & 0 \\ 3a_{13} & a_{13} & a_{11}+a_{13} \end{pmatrix}$
2. $\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{14} & b_{11} & b_{12} & b_{13} \\ b_{13} & b_{14} & b_{11} & b_{12} \\ b_{12} & b_{13} & b_{14} & b_{11} \end{pmatrix}$

这里的 a_{1i} 与 b_{1i} 可取任何实数。

作业 2.14 (P98 T17)

对角元素都相等的对角方阵称为纯量方阵。证明：和所有 n 阶方阵都可交换的方阵一定是纯量方阵。

解 直接取 $B=E_{ij}$，E_{ij} 为第 i 行第 j 列元素为 1，其余元素为 0 的矩阵，由 $AB=BA$ 可得 A 为对角方阵。

其次，取

$$B=\begin{pmatrix} 1 & 2 & \cdots & n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

可进一步得到 A 的对角线系数均相等，故 A 一定是纯量方阵。

作业 2.15 (P98 T18)

设 n 阶方阵 A 的每一行上都恰有 2 个元素为 1，而其他元素为零，J 是元素全为 1 的 n 阶方阵。求出所有适合 $A^2+2A=2J$ 的 n 阶方阵 A。

解 不妨设 $A=(a_{ij})_{n\times n}, J=(b_{ij})_{n\times n}$。由条件

$$b_{ii}=2a_{ii}+\sum_{l=1}^{n}a_{il}a_{li}=3a_{ii}+\sum_{l\neq i}a_{il}a_{li}=2$$

故 a_{ii} 不等于 1，即 $a_{ii}=0$。

假设 $a_{1i_1}=a_{1i_2}=1$，根据上方有 $a_{i_11}=a_{i_21}=1$，且存在 k_1,k_2 使得 $a_{i_1k_1}=a_{i_2k_2}=1$，故 A^2 每行元素之和为 4，即 A^2 所有元素之和为 $4n$。而 $2A$ 元素之和为 $4n$，$2J$ 元素之和为 $2n^2$，则 $8n=2n^2$，解得 $n=4$，即 A 为 4 阶矩阵，进一步讨论得只能为：

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix},\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

2.2 补充内容

线性代数和别的领域一个很大不同是其对于基础框架（矩阵、向量空间、行列式等）的推导并不唯一，有非常多条道路可以展开整个知识体系，如不同书中对行列式的定义过程都不相同。笔者见过最奇葩最绕的推导过程就是线代 B1 的推导，极尽所能的复杂化整个过程。

2.2.1 行列式

前两周老师主要讲了行列式定义和一些定理，以及行列式的计算。其中行列式计算是考试中比较常出现的题目，它的基础在于敏锐的使用三种行列式基本变换 + 熟练掌握各种技巧和模型的特点，下述是行列式定义，定理，以及计算技巧和经典模型的整理：

1. 行列式定义：
   (a). 代数余子式归纳（递推）
   (b). 看作函数，为 $\mathbb{F}^n$ 上的 n 重、反对称、规范、线性的函数，并规定行列式的性质，证明满足这些性质的行列式函数存在且唯一（公理化）
   (c). 完全展开
2. 行列式的 Laplace 展开
3. Cramer 法则
4. Binet-Cauchy 公式
5. 行列式的计算技巧
   (a). 化为三角行列式
   (b). 建立递推公式
   (c). 加边
   (d). 拆行
   (e). 视行列式为某些元素的多项式（参数化）
   (f). 对某一行或者某一列 Laplace 展开
6. 一些经典行列式
   (a). 三对角行列式
   (b). 鸭爪形行列式
   (c). 每项可以写成 $f(a_i)+g(b_j)$ 或 $f(a_i)g(b_j)$ 形式的行列式

问题 2.1 计算行列式

$$\begin{vmatrix} x_1-1 & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2-2 & x_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n-n \end{vmatrix}$$

问题 2.2 计算行列式

$$\begin{vmatrix} x & 0 & 0 & \cdots & a_n \\ -1 & x & 0 & \cdots & a_{n-1} \\ 0 & -1 & x & \ddots & a_{n-2} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & x+a_1 \end{vmatrix}$$

2.2.2 矩阵

1. 矩阵的代数运算
2. 可逆矩阵（如何计算 A^{-1}?）
   (a). 算 det(A)+ 算每个代数余子式
   (b). $(A, I_n)->(I, A^{-1})$ 做初等变换
   (c). 求出零化多项式
   (d). 用聪明的脑袋直接构造出来
3. 分块矩阵-> 打洞
4. 矩阵的秩和相抵
   (a). $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$
   (b). $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$
   (c). $\text{rank}(ABC) \geq \text{rank}(AB) + \text{rank}(BC) - \text{rank}(B)$
   (d). $\text{rank}(AC) \geq \text{rank}(A) + \text{rank}(C) - n$
   (e). $\text{rank}(A) \geq \text{rank}(A^2) \geq \text{rank}(A^3) \geq \cdots \text{rank}(A^k) = \text{rank}(A^{k+1}) = \cdots \text{rank}(A^n) = \cdots$
   (f). $\text{rank}(A, B) \leq \text{rank}(A) + \text{rank}(B)$

问题 2.3 证明 $\lambda^n \det(\lambda I_m - AB) = \lambda^m \det(\lambda I_n - BA)$。

2.2.3 附加题

前两周的内容没有什么很困难的题目，这里补充几个比较有趣和有用的题目：

命题 2.16

n 阶方阵 A 的附属方阵记为 A^*，证明：

1. $\text{rank}\, A^* = n$ 的充分必要条件是 $\text{rank}\, A = n$；
2. $\text{rank}\, A^* = 1$ 的充分必要条件是 $\text{rank}\, A = n-1$；
3. $\text{rank}\, A^* = 0$ 的充分必要条件是 $\text{rank}\, A < n-1$；
4. 当 $n > 2$ 时，$(A^*)^* = (\det A)^{n-2}A$；当 $n = 2$ 时，$(A^*)^* = A$。

证明

1. 当 $\text{rank}\, A = n$, 有 $\det A \neq 0$, 由 $AA^* = (\det A)I_{(n)}$ 知 $\det A^* = (\det A)^{n-1} \neq 0$, 所以 $\text{rank}\, A^* = n$。反过来，若 $\text{rank}\, A^* = n$，有 $\det A^* \neq 0$，由 $AA^* = (\det A)I_{(n)}$ 知 $\det A = \sqrt[n-1]{\det A^*} \neq 0$，所以 $\text{rank}\, A = n$。

2. 当 $\operatorname{rank} A = n-1$ 时，$\det A = 0$，从而 $AA^* = O$。又由于 $0 = \operatorname{rank}(AA^*) \geq \operatorname{rank} A + \operatorname{rank} A^* - n$，必有 $\operatorname{rank} A^* \leq 1$。而根据 $\operatorname{rank} A = n-1$，$A^*$ 中至少有一个 $A_{ij} \neq 0$，因此 $\operatorname{rank} A^* = 1$。反过来，若 $\operatorname{rank} A^* = 1$，由 $AA^* = (\det A)I_{(n)}$ 可知 $\det A = 0$ 且 $\operatorname{rank} A \leq n-1$，而由 $\operatorname{rank} A^* = 1$ 可知 A^* 中至少有一个 $A_{ij} \neq 0$，故 $\operatorname{rank} A \geq n-1$，因此 $\operatorname{rank} A = n-1$。
3. 当 $\operatorname{rank} A < n-1$ 时，A^* 中任意 $A_{ij} = 0$，所以 $A^* = 0$，则 $\operatorname{rank} A^* = 0$。反过来，若 $\operatorname{rank} A^* = 0$，则有 $A^* = 0$，因此任意 $A_{ij} = 0$，说明 $\operatorname{rank} A < n-1$。
4. 若 $\det A \neq 0$，即 A 可逆时，由

$$AA^* = \det AI_n \Rightarrow \det A^* = (\det A)^{n-1}, A^* = \det A \cdot A^{-1}$$

可知

$$A^*(A^*)^* = \det A^* I_n \Rightarrow (A^*)^* = \det A^*(A^*)^{-1} = (\det A)^{n-2}A$$

无论 n 是否为 2 都满足结论。

若 $\det A = 0$，即 A 不可逆时，当 $n > 2$ 时，由 $\operatorname{rank} A^* \leq 1 < n-1$ 可知 $\operatorname{rank}(A^*)^* = 0$，即 $(A^*)^* = O$，因此 $(A^*)^* = (\det A)^{n-2}A = O$。

当 $n = 2$，因为 $\operatorname{rank} A \leq 1$，设

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, ad - bc = 0$$

则有

$$A^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, (A^*)^* = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

综上可知，当 $n > 2$ 时 $(A^*)^* = (\det A)^{n-2}A$，当 $n = 2$ 时，$(A^*)^* = A$。

命题 2.17

设 $a_1, a_2, \cdots, a_n$ 是正整数。证明：n 阶行列式

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

能被 $1^{n-1}2^{n-2}\cdots(n-2)^2(n-1)$ 整除.

证明 对行列式做列变换可得其等于

$$\begin{vmatrix} 1 & a_1 & a_1(a_1-1) & \cdots & a_1(a_1-1)(a_1-2)\cdots(a_1-n+2) \\ 1 & a_2 & a_2(a_2-1) & \cdots & a_2(a_2-1)(a_2-2)\cdots(a_2-n+2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n(a_n-1) & \cdots & a_n(a_n-1)(a_n-2)\cdots(a_n-n+2) \end{vmatrix}$$

根据数论结论有 $k! \mid a_1(a_i-1)(a_i-2)\cdots(a_i-k+1)$，故行列式第 k 列元素均为 $k!$ 倍数，提取出来即有结论成立。

命题 2.18

设 A 和 B 为 n 阶方阵，证明：

$$\operatorname{rank}(AB - I_n) \le \operatorname{rank}(A - I_n) + \operatorname{rank}(B - I_n)$$

证明

$$\begin{pmatrix} I_n & A \\ 0 & I_n \end{pmatrix} \begin{pmatrix} A - I_n & 0 \\ 0 & B - I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ I_n & I_n \end{pmatrix} = \begin{pmatrix} AB - I_n & AB - A \\ B - I_n & B - I_n \end{pmatrix},$$

所以

$$\begin{aligned} \operatorname{rank}(AB - I_n) &\le \operatorname{rank} \begin{pmatrix} AB - I_n & AB - A \\ B - I_n & B - I_n \end{pmatrix} = \operatorname{rank} \begin{pmatrix} A - I_n & 0 \\ 0 & B - I_n \end{pmatrix} \\ &= \operatorname{rank}(A - I_n) + \operatorname{rank}(B - I_n) \end{aligned}$$

第 3 章 相抵与相合

3.1 习题讲解

3.1.1 第三周

作业 3.1 (P113 T5)

证明：

1. 正交方阵一定可逆，且其逆正交；
2. 酉方阵一定可逆，且其逆为酉方阵；
3. 可逆对称方阵的逆仍对称；
4. 可逆斜对称方阵的逆仍斜对称。

♡

解

1. 设其为 A，由定义 $AA^T = A^TA = I$ 可知其可逆，逆为 A^T，而 $(A^T)^T = A$，由此即可验证逆正交。
2. 设其为 A，由定义 $AA^H = A^HA = I$ 可知其可逆[1]，逆为 A^H，而 $(A^H)^H = A$，由此即可验证逆为酉方阵。
3. 设其为 A，由于 $(A^{-1})^TA = (A^{-1})^TA^T = (AA^{-1})^T = I$，有 $(A^{-1})^T = A^{-1}$，从而逆仍对称。
4. 设其为 A，由于 $(A^{-1})^TA = -(A^{-1})^TA^T = -(AA^{-1})^T = -I$，有 $(A^{-1})^T = -A^{-1}$，从而逆仍斜对称。

作业 3.2 (P113 T10)

设 A_{ij} 是 n 阶方阵 A 相对 a_{ij} 的代数余子式，证明 $\forall 1 \le i < j \le n, 1 \le k < l \le n$ 有

$$\begin{vmatrix} A_{ik} & A_{jk} \\ alpha_{il} & A_{jl} \end{vmatrix} = (-1)^{i+j+k+l} A \begin{pmatrix} 1 & \dots & i-1 & i+1 & \dots & j-1 & j+1 & \dots & n \\ 1 & \dots & k-1 & k+1 & \dots & l-1 & l+1 & \dots & n \end{pmatrix} \det A$$

♡

解 当 $\det A = 0$ 时，根据下一道习题可知 $\operatorname{rank} A^* \le 1$，从而左侧必然为 0，得证。

当 $\det A \ne 0$，若 $i \ne 1$，考虑 $A' = S_{1i}A$，这里 S_{1i} 代表交换 1 与 i 的初等方阵。这样变换后满足 $A'_{1k} = A_{ik}(-1)^{i+1}$，$A'_{1l} = A_{il}(-1)^{i+1}$，而右侧第一项 $(-1)^{i+j+k+l} = (-1)^{1+j+k+l}(-1)^{i+1}$，后两项由 $\det S_{1i} = -1$ 可知均乘 -1，由此可知原问题等价于

$$\begin{vmatrix} A'_{1k} & A'_{jk} \\ alpha'_{1l} & A'_{jl} \end{vmatrix} = (-1)^{1+j+k+l} A' \begin{pmatrix} & & & 2 & \dots & j-1 & j+1 & \dots & n \\ 1 & \dots & k-1 & k+1 & \dots & l-1 & l+1 & \dots & n \end{pmatrix} \det A'$$

再对 j, k, l 类似操作，可使得 $j = l = 2, k = 1$，也即只需要考虑 $i = k = 1, j = l = 2$ 的情况。

将 A^* 与 A 按照左上 2×2 右下 $(n-2) \times (n-2)$ 分块，对应记为 $\begin{pmatrix} A_1^* & A_2^* \\ A_3^* & A_4^* \end{pmatrix}$，$\begin{pmatrix} A_1 & A_2 \\ alpha_3 & A_4 \end{pmatrix}$，则

---

[1] 上标 H 代表共轭转置。

要证的问题即化为 $\det A_1^* = \det A_4 \det A$。根据伴随矩阵性质，$AA^* = (\det A)I$，可得

$$\det A_1^* = \det \begin{pmatrix} A_1^* & O \\ alpha_3^* & I \end{pmatrix} = \det A^{-1} A \begin{pmatrix} A_1^* & O \\ alpha_3^* & I \end{pmatrix} = (\det A)^{-1} \det \begin{pmatrix} (\det A)I & A_2 \\ O & A_4 \end{pmatrix}$$

进一步化简即 $\det A_1^* = \det A \det A_4$，得证。

这题也可以通过直接对应计算的方法求解，但本质仍然是利用 $AA^* = (\det A)I$ 作分块的拆分。

作业 3.3 (P122 T2)

求 λ 使如下矩阵 A 秩最小：

$$\begin{pmatrix} 3 & 1 & 1 & 4 \\ \lambda & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix}$$

解 计算得原矩阵可相抵为 $\operatorname{diag}(1, 1, \lambda, 0)$[2]，当且仅当 $\lambda = 0$ 时秩最小，为 2，否则为 3。

作业 3.4 (P124 T13)

n 阶方阵 A 的附属方阵记为 A^*，证明：

1. $\operatorname{rank} A^* = n$ 的充分必要条件是 $\operatorname{rank} A = n$；
2. $\operatorname{rank} A^* = 1$ 的充分必要条件是 $\operatorname{rank} A = n - 1$；
3. $\operatorname{rank} A^* = 0$ 的充分必要条件是 $\operatorname{rank} A < n - 1$；
4. 当 $n > 2$ 时，$(A^*)^* = (\det A)^{n-2} A$；当 $n = 2$ 时，$(A^*)^* = A$。

解 见上章附加题中的命题 2.16。

作业 3.5 (P143 T6)

对齐次线性方程组 $Ax = \mathbf{0}$，$A \in \mathbb{R}^{m\times n}, \operatorname{rank}(A) = n - 1$，说明该方程任意两解成比例。

解 设 $A = P \begin{pmatrix} I_{n-1} & \mathbf{0} \\ O & \mathbf{0} \end{pmatrix} Qx$，$P, Q$ 可逆，则分析得所有解可写为 $Q^{-1}(0, \dots, 0, t)^T, t \in \mathbb{R}$，此即 $tQ^{-1}e_n$，于是任何两解相差转置 (e_n 表示第 n 个分量为 1，其他为 0 的向量)。

作业 3.6 (P151 T2)

设 $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{p\times q}, C \in \mathbb{R}^{m\times q}$，未知量 $X \in \mathbb{R}^{n\times p}$，证明 $AXB = C$ 有解等价于 $(I - AA^-)C = C(I - B^-B) = O$，且有解时通解为

$$X = A^-CB^- + (I - A^-A)Y + Z(I - B^-B) + (I - AA^-)W(I - BB^-)$$

$Y, Z, W \in \mathbb{R}^{n\times p}$ 可任取。

解 若 $(I - AA^-)C = C(I - B^-B) = O$，代入通解形式计算验证可知均为解。

方程有解时，代入 $C = AXB$ 可验证 $(I - AA^-)C = C(I - B^-B) = O$。

对任何解 X_0，取 $Y = Z = X_0, W = -X_0$，代入 $C = AX_0B$ 可计算验证

$$X_0 = A^-CB^- + (I - A^-A)Y + Z(I - B^-B) + (I - AA^-)W(I - BB^-)$$

[2]此处记号 $\operatorname{diag} \alpha$ 指以 α 的元素顺次为对角元的对角阵。

从而任何解可写为通解形式。

“通解”指此形式可以包含全部解，因此最后一部分证明是必要的。这部分证明中的 Y, Z, W 可待定后解出，下题同理。

作业 3.7 (P151 T3)

设 $A \in \mathbb{R}^{m\times p}, B \in \mathbb{R}^{q\times n}, C \in \mathbb{R}^{m\times n}$，未知量 $X \in \mathbb{R}^{p\times n}, Y \in \mathbb{R}^{m\times q}$，证明 $AX - YB = C$ 有解等价于 $(I - AA^-)C(I - B^-B) = O$，且有解时通解为

$$X = A^-C + A^-ZB + (I - A^-A)W$$

$$Y = -(I - AA^-)CB^- + Z - (I - AA^-)ZBB^-$$

$W \in \mathbb{R}^{p\times n}, Z \in \mathbb{R}^{m\times q}$ 可任取。

解 若 $(I - AA^-)C(I - B^-B) = O$，代入通解形式计算验证可知均为解。

方程有解时，代入 $C = AX - YB$ 可验证 $(I - AA^-)C(I - B^-B) = O$。

对任何解 X_0, Y_0，取 $W = X_0, Z = Y_0$，代入 $C = AX_0 - Y_0B$ 可计算验证

$$X_0 = A^-C + A^-ZB + (I - A^-A)W$$

$$Y_0 = -(I - AA^-)CB^- + Z - (I - AA^-)ZBB^-$$

从而任何解可写为通解形式。

思考：广义逆并不唯一，这两题中提到的广义逆是指存在还是任意？

作业 3.8 (P151 T4)

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

$$\operatorname{rank}\begin{pmatrix} A & C \\ O & B \end{pmatrix} = \operatorname{rank} A + \operatorname{rank} B + \operatorname{rank}(I - AA^-)C(I - B^-B)$$

解 引理：对矩阵 $D \in \mathbb{R}^{a\times b}$ 与其广义逆 $D^- \in \mathbb{R}^{b\times a}$，若 $\operatorname{rank} A = r$，必然存在可逆方阵 P, Q 与某个 $R \in \mathbb{R}^{(b-r)\times(a-r)}$ 使得

$$D = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q, D^- = Q^{-1}\begin{pmatrix} I_r & O \\ O & R \end{pmatrix}P^{-1}$$

引理证明：考虑 $D = P_0\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q_0$，由 $DD^-D = D$ 可解出必有 $D^- = Q_0^{-1}\begin{pmatrix} I_r & D_2 \\ D_3 & D_4 \end{pmatrix}P_0^{-1}$。由于

$$\begin{pmatrix} I & O \\ -D_3 & I \end{pmatrix}\begin{pmatrix} I_r & D_2 \\ D_3 & D_4 \end{pmatrix}\begin{pmatrix} I & -D_2 \\ O & I \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & D_4 - D_3D_2 \end{pmatrix}$$

令

$$P = P_0\begin{pmatrix} I & D_2 \\ O & I \end{pmatrix}, Q = \begin{pmatrix} I & O \\ D_3 & I \end{pmatrix}Q_0, R = D_4 - D_3D_2$$

可计算验证成立。

原题证明：利用引理，假设 A,B 秩为 a,b，对应的拆分为 P_A,Q_A,R_A,P_B,Q_B,R_B，则计算得

$$(I-AA^-)C(I-B^-B)=P_A\begin{pmatrix}O&O\\O&I_{m-a}\end{pmatrix}P_A^{-1}CQ_B^{-1}\begin{pmatrix}O&O\\O&I_{n-b}\end{pmatrix}Q_B$$

记 $C_0=P_A^{-1}CQ_B^{-1}$，则右侧为

$$a+b+\begin{pmatrix}O&O\\O&I_{m-a}\end{pmatrix}C_0\begin{pmatrix}O&O\\O&I_{n-b}\end{pmatrix}$$

设第三项为 c，即为 C_0 右下角 $(m-a)\times(n-b)$ 阶子矩阵的秩，记此矩阵为 C_4。

另一方面，左侧等于

$$\operatorname{rank}\begin{pmatrix}P_A^{-1}&O\\O&P_B^{-1}\end{pmatrix}\begin{pmatrix}A&C\\O&B\end{pmatrix}\begin{pmatrix}Q_A^{-1}&O\\O&Q_B^{-1}\end{pmatrix}=\operatorname{rank}\begin{pmatrix}\operatorname{diag}(I_a,O)&C_0\\O&\operatorname{diag}(I_b,O)\end{pmatrix}$$

利用行列变换，I_a 所在行的其他部分与 I_b 所在列的其他部分可消去，因此左侧化为

$$\operatorname{rank}\begin{pmatrix}I_a&&&\\&O_{m-a}&&C_4\\&&I_b&\\&&&O_{n-b}\end{pmatrix}$$

考虑列秩可得到其秩即为 $a+b+c$，得证。

总有些题没什么特殊技巧，需要硬打洞处理，这时就可见基本功了。

3.1.2 第四周

虽然作业题目里使用了 $A>0,A>B$ 这样的记号，但由于可能存在歧义，尽量不要使用。

作业 3.9 (P365 T2)

对 n 阶对称方阵 S，证明存在唯一 n 阶对称方阵 S_1 满足 $S_1^3=S$。

解 设 S 正交相似对角化为 $P^T\operatorname{diag}(\lambda_1,\dots,\lambda_n)P$，且 $\lambda_1\ge\dots\ge\lambda_n$，则 $S_1=P^T\operatorname{diag}(\sqrt[3]{\lambda_1},\dots,\sqrt[3]{\lambda_n})P$ 满足方程，因此存在。

为说明唯一性，若有 S_2 正交相似对角化为 $Q^T\operatorname{diag}(\mu_1,\dots,\mu_n)Q$ 符合条件，根据 S_2^3 与 S 特征值相同可知 $\mu_1^3,\dots,\mu_n^3$ 必然为 $\lambda_1,\dots,\lambda_n$ 的一个排列，于是通过置换相似可不妨设 $\mu_i=\sqrt[3]{\lambda_i}$。

此时，由于 $Q^T\operatorname{diag}(\lambda_1,\dots,\lambda_n)Q=P^T\operatorname{diag}(\lambda_1,\dots,\lambda_n)P$，同时左乘 P 右乘 Q^T 可知 PQ^T 与 $\operatorname{diag}(\lambda_1,\dots,\lambda_n)$ 可交换。由大小关系假设，必有 $\operatorname{diag}(\lambda_1,\dots,\lambda_n)=\operatorname{diag}(a_1I_1,\dots,a_kI_k)$，其中 I_i 的阶数和为 n，a_i 互不相等。

计算可知，与 $\operatorname{diag}(a_1I_1,\dots,a_kI_k)$ 可交换当且仅当方阵可写为 $\operatorname{diag}(B_1,\dots,B_k)$ 的形式，B_i 与 I_i 阶数相等为任意方阵。由于 $x\to x^3$ 在实数上是单射，$\operatorname{diag}(\sqrt[3]{\lambda_1},\dots,\sqrt[3]{\lambda_n})=\operatorname{diag}(\sqrt[3]{a_1}I_1,\dots,\sqrt[3]{a_k}I_k)$，与 $\operatorname{diag}(\lambda_1,\dots,\lambda_n)$ 可交换同与 $\operatorname{diag}(\sqrt[3]{\lambda_1},\dots,\sqrt[3]{\lambda_n})$ 可交换等价，因此有

$$Q^T\operatorname{diag}(\sqrt[3]{\lambda_1},\dots,\sqrt[3]{\lambda_n})Q=P^T\operatorname{diag}(\sqrt[3]{\lambda_1},\dots,\sqrt[3]{\lambda_n})P$$

即 $S_1=S_2$，得证。

由此出发，任何对称阵的奇数次方根是可定义的，而任何半正定阵的任意算术次方根是可定义的 (从而可定义任何非负分数次方，再由极限定义非负实数次方)。对正定阵，由于 -1 次方存在，任何实数次方都有意义，这就类似我们对正数、0、负数的不同幂次系统。

作业 3.10 (P365 T3)

若 A,B 正定，证明 AB 所有特征值都是正的。

解 我们介绍三种方法：

1. 根据正定阵存在唯一正定平方根，有 $AB=A^{1/2}A^{1/2}BA^{1/2}A^{-1/2}$ 相似于 $A^{1/2}BA^{1/2}$，由 A 对称，$A^{1/2}BA^{1/2}$ 与 B 相合，从而正定，其特征值均正，因此 AB 特征值均正。
2. 设 $A=P^TP$，P 可逆，则 $AB=PP^TBPP^{-1}$ 相似于 P^TBP，P^TBP 与 B 相合，从而正定，其特征值均正，因此 AB 特征值均正。
3. 设 AB 相对特征值 λ 的特征向量为 x，则 $ABx=\lambda x\Rightarrow x^HBx=\lambda x^HA^{-1}x$，由于正定阵必为 Hermite 阵[3]，$x$ 非零，可知 $\lambda=\frac{x^HBx}{x^HA^{-1}x}>0$。

作业 3.11 (P365 T6)

设 n 阶对称方阵 S 前 $n-1$ 个顺序主子式大于 0，且行列式非负，证明其半正定。

解 根据书上定理，设 S 分块为 $\begin{pmatrix} S_{n-1} & \alpha \\ \alpha^T & s_{nn} \end{pmatrix}$，则 S_{n-1} 正定。由于

$$\begin{pmatrix} I & \mathbf{0} \\ -\alpha^TS_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} S_{n-1} & \alpha \\ \alpha^T & s_{nn} \end{pmatrix}\begin{pmatrix} I & -S_{n-1}^{-1}\alpha \\ \mathbf{0} & 1 \end{pmatrix}=\begin{pmatrix} S_{n-1} & \mathbf{0} \\ \mathbf{0}^T & s_{nn}-\alpha^TS_{n-1}^{-1}\alpha \end{pmatrix}$$

S 相合于 $\begin{pmatrix} S_{n-1} & \alpha \\ \alpha^T & s_{nn}-\alpha^TS_{n-1}^{-1}\alpha \end{pmatrix}$，且 $\det S=(s_{nn}-\alpha^TS_{n-1}^{-1}\alpha)\det S_{n-1}$。根据 $\det S\geq 0,\det S_1>0$ 可知 $s_{nn}-\alpha^TS_{n-1}^{-1}\alpha\geq 0$，从而计算特征值可知 $\begin{pmatrix} S_{n-1} & \mathbf{0} \\ \mathbf{0}^T & s_{nn}-\alpha^TS_{n-1}^{-1}\alpha \end{pmatrix}$ 半正定，因此 S 半正定。

作业 3.12 (P365 T7)

设 A,B 半正定，证明 $\det(A+B)\geq\det A+\det B$。

解 若 A,B 均非正定，则右侧为 0，而 $x^T(A+B)x=x^TAx+x^TBx\geq 0$，且 $(A+B)^T=A+B$，因此 $A+B$ 半正定，$\det(A+B)\geq 0=\det A+\det B$。

否则，不妨设 A 正定，可写为 P^TP，P 可逆，则右侧大于 0，有

$$\frac{\det(A+B)}{\det A+\det B}=\frac{\det P^T\det(I+P^{-T}BP^{-1})\det P}{\det P^T\det P+\det P^T\det(P^{-T}BP^{-1})\det P}$$

记 $B_0=P^{-T}BP^{-1}$，其与 B 相合，因此仍半正定，假设 $B_0=Q^T\operatorname{diag}(d_1,\dots,d_n)Q$ 为其正交相似对角化，此分式可写为

$$\frac{\det(I+Q^T\operatorname{diag}(d_1,\dots,d_n)Q)}{1+\det Q^T\operatorname{diag}(d_1,\dots,d_n)Q}=\frac{\det Q^T(I+\operatorname{diag}(d_1,\dots,d_n))Q}{1+d_1\dots d_n}=\frac{(1+d_1)\dots(1+d_n)}{1+d_1\dots d_n}$$

由于分子为分母展开中的两项，$d_i\geq 0$，此分数必然 ≥ 1，即得证。

作业 3.13 (P366 T12)

对 n 阶正定阵 S 与 n 维列向量 x,y[a]，证明：

$$(x^TSy)^2\leq(x^TSx)(y^TSy)$$

[3] 见本次补充内容。

且等号成立当且仅当 x 与 y 成比例。

[a]原题表述为行向量，但写成列向量更加自然。

解 记 $S=P^TP$，P 可逆，并令 $x_0=Px,y_0=Py$，则原不等式化为

$$(x_0^Ty_0)^2\le x_0^Tx_0y_0^Ty_0$$

按分量展开发现此即为柯西不等式，取等当且仅当 x_0,y_0 成比例，即存在不全为 0 的 λ,μ 使得 $\lambda x_0+\mu y_0=\mathbf{0}$，同时左乘 P^{-1} 得此等价于 $\lambda x+\mu y=\mathbf{0}$，于是得证。

作业 3.14 (P366 T13)

设 n 阶实方阵 A 极分解唯一，证明 A 可逆。

解 证明逆否命题，即假设 A 不可逆证明极分解不唯一。根据书上定理，极分解 $A=SO$ 中半正定阵 S 唯一。若 S 正定由 O 可逆可知 A 可逆，因此 S 不正定，设其正交相似对角化为 $P^T\operatorname{diag}(s_1,\dots,s_{n-1},0)P$，考虑

$$O'=P^T\operatorname{diag}(1,\dots,1,-1)PO$$

由于乘积中每个分量都是正交阵，O' 仍正交，且计算可验证 $SO'=SO$，从而不唯一。

作业 3.15 (P375 T3)

设 A 为 n 阶正定阵，$B\in\mathbb{R}^{n\times m}$ 列满秩，计算

$$C=\begin{pmatrix}A & B\\ B^T & O\end{pmatrix}$$

的逆。

解 行列变换可知

$$\begin{pmatrix}I & O\\ -B^TA^{-1} & I\end{pmatrix}\begin{pmatrix}A & B\\ B^T & O\end{pmatrix}\begin{pmatrix}I & -A^{-1}B\\ O & I\end{pmatrix}=\begin{pmatrix}A & O\\ O & -(B^TA^{-1}B)\end{pmatrix}$$

先证明 $B^TA^{-1}B$ 可逆：由 A 正定可知 A^{-1} 正定。由 B 列满秩，$Bx=\mathbf{0}$ 的解仅有 $x=\mathbf{0}$，于是 $x\ne\mathbf{0}$ 时 $x^TB^TA^{-1}Bx=(Bx)^TA^{-1}Bx>0$，从而 $B^TA^{-1}B$ 正定，必然可逆。

于是，原矩阵的逆可写为

$$\begin{pmatrix}I & -A^{-1}B\\ O & I\end{pmatrix}\begin{pmatrix}A^{-1} & O\\ O & -(B^TA^{-1}B)^{-1}\end{pmatrix}\begin{pmatrix}I & O\\ -B^TA^{-1} & I\end{pmatrix}$$

作业 3.16 (P376 T13)

若 A 与 $B-A$ 半正定，求证 $B^{1/2}-A^{1/2}$ 半正定。

解 利用半正定定义可知 B 半正定。由于 $B^{1/2}$ 难以处理，我们考虑本题等价的描述：若 C,D 半正定，且 C^2-D^2 半正定，求证 $C-D$ 半正定。

设 $C=P^T\operatorname{diag}(c_1,\dots,c_n)P,D=P^T\operatorname{diag}(d_1,\dots,d_n)P$ 为 C,D 的同时相合对角化，则 C^2-D^2 半正定等价于 $S=P^{-T}(C^2-D^2)P^{-1}$ 半正定，即

$$S=\operatorname{diag}(c_1,\dots,c_n)PP^T\operatorname{diag}(c_1,\dots,c_n)-\operatorname{diag}(d_1,\dots,d_n)PP^T\operatorname{diag}(d_1,\dots,d_n)$$

半正定。

记 $F=PP^T$，则其正定，于是 $f_{ii}=e_i^TFe_i>0$，而根据 S 半正定得 $e_i^TSe_i=(c_i^2-d_i^2)f_{ii}\geq 0$，从而 $c_i^2\geq d_i^2$，再根据 C,D 半正定，c_i,d_i 非负，即有 $c_i\geq d_i$，于是 $C-D=P^T\operatorname{diag}(c_1-d_1,\dots,c_n-d_n)P$ 半正定。

事实上 C,D 半正定时，存在正整数 n 使 C^n-D^n 半正定即可推出 $C-D$ 半正定，但证明较为复杂。

3.2 补充内容

3.2.1 广义逆

在书上，我们介绍的 Moore-Penrose 广义逆时共有四个条件：

$$\begin{cases} AA^+A=A & \text{P1} \\ A^+AA^+=A^+ & \text{P2} \\ (AA^+)^H=AA^+ & \text{P3} \\ (A^+A)^H=A^+A & \text{P4} \end{cases}$$

书中定义满足 P1 的矩阵称为广义逆，记作 A^-。另一个常用的广义逆定义为满足 P1 与 P2 的矩阵，这里称为强广义逆[4]，记作 $A^=$。比起书上的广义逆，强广义逆具有相互性，因此也称为自反广义逆。为了简化说明，以下的讨论均在实矩阵范畴。

定理 3.17 (强广义逆的表示)

对矩阵 $D\in\mathbb{R}^{a\times b}$ 与其强广义逆 $D^=\in\mathbb{R}^{b\times a}$，若 $\operatorname{rank}A=r$，必然存在可逆方阵 P,Q 使得

$$D=P\begin{pmatrix} I_r & O \\ O & O\end{pmatrix}Q,D^==Q^{-1}\begin{pmatrix} I_r & O \\ O & O\end{pmatrix}P^{-1}$$

证明 由于强广义逆一定为广义逆，根据作业 3.8 的引理可得能写为 $Q^{-1}\begin{pmatrix} I_r & O \\ O & R\end{pmatrix}P^{-1}$ 的形式，再由 $D^=DD^==D^=$ 可计算得 $R=O$，从而得证。

由此有结论：强广义逆的秩与原矩阵相同。根据作业 3.8 的形式可验证反之也成立，于是有：

定理 3.18 (强广义逆的等价定义)

A 的广义逆 A^- 是 A 的强广义逆当且仅当 $\operatorname{rank}A^-=\operatorname{rank}A$。

此外，强广义逆对解有更限定的刻画，即：

定理 3.19 (非齐次线性方程组的解)

方程 $Ax=b,b\neq 0$ 的通解为 $x=A^=b$，这里 $A^=$ 为 A 的任何强广义逆。

证明 注意到书上对通解是 A^-b 的证明中将 A^- 取为了 $Q^{-1}\begin{pmatrix} I & O \\ C & O\end{pmatrix}P^{-1}$，可计算得其秩与 A 相同，因此是强广义逆。

[4]有教材直接将强广义逆作为广义逆的定义，这里用上标 $=$ 同样也是为了区分。

之所以书上提到“广义逆”在计算数学领域有较多作用，很大程度是因为它可以直接刻画方程组的解，下面就给出一个经典的例子：

定理 3.20 (最小二乘问题的求解)

给定矩阵 $X\in\mathbb{R}^{n\times q},T\in\mathbb{R}^{n\times m}$，求使得 $\|Y-XB\|_F$[a]最小的 $B\in\mathbb{R}^{q\times m}$ 的问题称为最小二乘问题。对任何符合 P1、P2、P3 的 X 的广义逆 $X^{1,2,3}$，$B=X^{1,2,3}Y$ 为最小二乘问题的一个解。特别地，X^+Y 为一个解。

[a]此处 F 指 Frobenius 范数，即所有元素平方和的平方根。

♡

证明 由于 Forbenius 范数可以分离，只需要单独考虑 Y 的每一列即可拼接，也即只需说明 $q=1$ 时满足条件，记此时为 y 与 b。直接计算 $(y-Xb)^T(y-Xb)$ 对应的二次型并求导可知等号成立当且仅当 $X^TXb=X^Ty$。

根据之前定理，无论 X^Ty 是否为 0，$b=(X^TX)^=X^Ty$ 均为此方程的解，因此只需证明符合 P2 与 P3 等价于可写为 $(X^TX)^=X^T$。

将 X 作奇异值分解 $P\operatorname{diag}(\Sigma,O)Q$，其中 P,Q 正交，Σ 为对角元素大于 0 的对角阵，则 $X^TX=Q^T\operatorname{diag}(\Sigma^2,O)Q$，设 $(X^TX)^==Q^TZQ$，可得 $Z=\begin{pmatrix}\Sigma^{-2} & U\\ V & V\Sigma^2U\end{pmatrix}$，$U,V$ 任意，从而

$$(X^TX)^=X^T=Q^T\begin{pmatrix}\Sigma^{-1} & O\\ V\Sigma & O\end{pmatrix}P^T$$

反之，在奇异值分解下设满足 P1、P2、P3 的某 $X^{1,2,3}$ 写为 $Q^TX'P^T$，并对应分块为 $\begin{pmatrix}X_1' & X_2'\\ X_3' & X_4'\end{pmatrix}$，由 P3 可知 X_1' 对称，$X_3'=O$，再利用 P1、P2 即有

$$X^{1,2,3}=Q^T\begin{pmatrix}\Sigma^{-1} & O\\ W & O\end{pmatrix}P^T$$

由于 Σ 可逆，V,W 均任意，此形式与上方完全相同。

可以验证 $X^{2,3}$ 满足 P1，也即 P2 与 P3 可推出 P1。

从上面的过程中可以发现，当考虑 P3 与 P4 时，通过相抵分解成 Hermite 标准形已经不够了，需要进行更进一步的奇异值分解。Moore-Penrose 广义逆在奇异值分解下有非常优美的表示：

定理 3.21 (M-P 广义逆等价表示)

设 A 的奇异值分解为 $P\operatorname{diag}(\Sigma,O)Q$，其中 P,Q 正交，Σ 为对角元素大于 0 的对角阵，则

$$A^+=Q^T\operatorname{diag}(\Sigma^{-1},O)P^T$$

♡

证明 在上一个定理证明中已经得到了 $X^{2,3}$ 的表示，再结合 P4 即可得到最终结果。

3.2.2 复矩阵

由于课本是从 $\mathbb{R}^n$ 出发定义正定性的，我们课上与习题中的讨论基本都在对实矩阵进行，事实上，对复矩阵可以进行类似的讨论：

定义 3.22 (Hermite 阵的正定性)

满足 $A^H = A$ 的方阵称为 Hermite 阵，若 $x^H Ax \geq 0$ 恒成立，则称该 Hermite 阵半正定，若还有 $x \neq \mathbf{0}$ 时 $x^H Ax > 0$，则称其正定。

可以说明如下类似实方阵时的结论 (注意以下的是正数蕴含了是实数)：

命题 3.23 (酉相似与相合对角化)

对 Hermite 阵 A，存在酉方阵 U 使得 $A = U^H DU$，D 为对角阵，且对角元均为实数；存在可逆方阵 P 使得 $A = P^H DP$，D 为对角元只有 $0, \pm 1$ 的对角阵。

命题 3.24 (正定的等价定义)

对 n 阶 Hermite 阵 S，以下命题等价：

1. S 正定；
2. S 特征值均正；
3. 存在正定阵 S_1 使得 $S = S_1^2$；
4. 存在可逆阵 P 使得 $S = P^H P$；
5. S 的主子式均为正；
6. S 的顺序主子式均为正；
7. 对每个 k，S 的所有 k 阶主子式之和为正。

命题 3.25 (半正定的等价定义)

对 n 阶 Hermite 阵 S，以下命题等价：

1. S 半正定；
2. S 特征值均非负；
3. 存在半正定阵 S_1 使得 $S = S_1^2$；
4. 存在矩阵 P 使得 $S = P^H P$，特别地，可取 P 为 $\operatorname{rank} P = \operatorname{rank} S$ 的方阵；
5. S 的主子式均非负；
6. 对每个 k，S 的所有 k 阶主子式之和非负。

命题 3.26 (平方根唯一性)

对半正定阵 S，存在唯一半正定阵 S_1 使得 $S_1^2 = S$。

命题 3.27 (同时相合对角化)

对 n 阶 Hermite 阵 A, B，若 A 正定或 A, B 均半正定，则存在可逆阵 P 使得 $P^T AP, P^T BP$ 均为对角阵。

证明 A 正定时证明类似书 7.9 节例 1。对后一种情况，考虑 $A+B$ 的相合对角化为 $Q^H(A+B)Q = \operatorname{diag}(I_r, O)$，设

$$Q^H AQ = \begin{pmatrix} A_1 & A_2 \\ alpha_2^H & A_3 \end{pmatrix}, Q^H BQ = \begin{pmatrix} I_r - A_1 & -A_2 \\ -A_2^H & -A_3 \end{pmatrix}$$

由半正定性可知 $\pm A_3$ 均半正定，因此 $A_3 = O$。若 A_2 不为 O，构造向量 $\begin{pmatrix} u \\ \lambda A_2^H u \end{pmatrix}$ 在 $\lambda \to -\infty$ 可验证 A 不半正定，矛盾。最后考虑 A_1 的酉相似对角化，两边同时对左上角进行即得结论。

由此可定义出复矩阵的奇异值分解与极分解：

定理 3.28 (奇异值分解)

对任何复矩阵 A，存在酉方阵 U, V 使得

$$A = U\,\mathrm{diag}(\mathrm{diag}(\sigma_1, \ldots, \sigma_k), O)V$$

且满足 $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_k \geq 0$，它们称为矩阵 A 的奇异值，唯一确定。

定理 3.29 (谱分解)

对任何复方阵 A，存在半正定 Hermite 阵 S 与酉方阵 U 使得 $A = SU$，且 S 唯一确定。U 唯一确定当且仅当 A 可逆。

比起秩，特征值与奇异值更受计算数学人的喜爱，因为它们满足对矩阵 A 的连续性。在计算数学中，奇异值分解也被称为 SVD 分解，改进其算法是一个重要的课题。

最后，留一个有趣的问题：

作业 3.30

对于一个矩阵，我们可以用幂级数的形式定义其函数，如

$$\mathrm{e}^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

$$\ln A = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (A - I)^n$$

此处极限均指按各个分量收敛。由此，我们自然可以定义 $A^\alpha = \mathrm{e}^{\alpha \ln A}$。

1. 对任何 A，证明 e^A 存在且特征值均为正；当 A 可相似对角化时，证明 $\ln A$ 存在当且仅当 A 特征值均正[a]。
2. 当 A 为 Hermite 阵时，证明如此定义的 A^α 存在当且仅当 A 正定，当 $\alpha = \frac{p}{q}$ 为有理数时，其为 $X^q = A^p$ 的唯一半正定解。
3. 当 A 为半正定 Hermite 阵且 $\alpha > 0$ 时，定义 $A^\alpha = \lim_{\lambda \to 0^+} (\lambda I + A)^\alpha$，证明此极限存在，且当 $\alpha = \frac{p}{q}$ 为有理数时，其为 $X^q = A^p$ 的唯一半正定解。若 $\alpha < 0$，证明此极限不存在。
4. 计算 $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{\mathrm{i}\pi}$。

[a] 后续学到 Jordan 标准形时我们将重新研究一般情况的初等函数存在性。

第 4 章 多项式

4.1 习题讲解

4.1.1 第六周

作业 4.1 (P7 T3)

设非零的实系数多项式 $f(x)$（即系数都是实数的多项式）满足 $f(f(x)) = f(x)^k$，其中 k 是给定的正整数。求多项式 $f(x)$。

解 引理：设多项式 $f(x)$ 与 $g(x)$ 的次数都不超过 n。如果有 $n+1$ 个不同的值 a 使 $f(a) = g(a)$，则多项式 $f(x)$ 与 $g(x)$ 相等。

引理证明：根据代数基本定理，若 $f(a) - g(a)$ 非零，则其至多有 n 个零点，矛盾。

原题证明：当 $f(x)$ 是常数 a 时，取 $a = a^k$ 即可。其他情况下，设 $\deg f(x) = n \geq 1$。注意到 $f(\beta) = \beta^k$ 对所有的 $\beta = f(\alpha)$ 成立，只要 β 可取得足够多不同的值就可知 $f(x) = x^k$。对每一个 β 值，使 $f(\alpha) = \beta$ 的 α 值至多 n 个。因此，当 α 取了 nm 个不同的值时至少能得到 m 个不同的 β 值，又 α 可以取得无穷多个不同的值，于是有无穷多个不同的 $\beta = f(\alpha)$ 使 $f(\beta) = \beta^k$。由恒等定理就有 $f(x) = x^k$。

此引理虽然简单，但在判定多项式相等时十分有用。

作业 4.2 (P7 T4)

设非零的实系数多项式 $f(x)$ 满足 $f(x^2) = f^2(x)$，求多项式 $f(x)$。

解 设所求的多项式为

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

其中 a_n 非 0，我们证明 $a_{n-1} = \cdots = a_1 = a_0 = 0$。假设 $a_i(i = 0, 1, \cdots, n-1)$ 不全为 0，设 $k < n$ 是使得 a_k 非 0 的最大下标。由 $f(x^2) = f^2(x)$，得

$$a_n x^{2n} + a_{n-1}x^{2k} + \cdots + a_1 x^2 + a_0 = (a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_o)^2$$

比较 x^{n+k} 得系数，得出 $0 = 2a_n a_k$, 这与 $a_n \neq 0, a_k \neq 0$ 矛盾。因此上述的断言正确, 即 $f(x) = a_n x^n$。再由 $f(x^2) = f^2(x)$ 即知 $a_n = 1$, 所以 $f(x) = x^n$.

作业 4.3 (P16 T1)

设多项式 $g(x) = x^2 - 2ax + 2$ 整除多项式 $f(x) = x^4 + 3x^2 + ax + b$，求实数 a 和 b。

解 由 $f(x) = x^4 + 3x^2 + ax + b = (x^2 + 2ax + 4a^2 + 1)(x^2 - 2ax + 2) + (8a^3 - a)x + b - 8a^2 - 2$

可知 $8a^3 - a = b - 8a^2 - 2 = 0$。因此有 $a = 0, b = 2$ 或者 $a = \pm\frac{\sqrt{2}}{4}, b = 3$.

作业 4.4 (P16 T3)

设 m 为自然数，证明当 $n = 6m+5$ 时，多项式 $x^2 + xy + y^2$ 整除多项式 $(x+y)^n - x^n - y^n$；当 $n = 6m+1$ 时，多项式 $(x^2 + xy + y^2)^2$ 整除多项式 $(x+y)^n - x^n - y^n$。

解 当 $n=6m+5$ 时，由

$$\begin{aligned}(x+y)^n-x^n-y^n&=(x+y)^{6m+5}-x^{6m+5}-y^{6m+5}\\&=(x^2+xy+y^2+xy)^{3m+2}(x+y)-x^{6m+5}-y^{6m+5}\end{aligned}$$

及二项式定理, 只需考察 $(xy)^{3m+2}(x+y)-x^{6m+5}-y^{6m+5}$ 与 x^2+xy+y^2 的整除性.

$$\begin{aligned}&(xy)^{3m+2}(x+y)-x^{6m+5}-y^{6m+5}\\&=x^{3m+3}y^{3m+2}+x^{3m+2}y^{3m+3}-x^{6m+5}-y^{6m+5}\\&=(x^{3m+3}-y^{3m+3})(-x^{3m+2}+y^{3m+2})\end{aligned}$$

故 $(x^{3m+3}-y^{3m+3})|(xy)^{3m+2}(x+y)-x^{6m+5}-y^{6m+5}$。又 $x^3-y^3|(x^{3m+3}-y^{3m+3})$, $(x^2+xy+y^2)|x^3-y^3$, 所以 $x^2+xy+y^2|(xy)^{3m+2}(x+y)-x^{6m+5}-y^{6m+5}$, 进而 x^2+xy+y^2 整除 $(x+y)^n-x^n-y^n$.

事实上，我们有一种更为简便的方法。注意到 $x^2+x+1=(x-\alpha)(x-\beta)$, 其中 $\alpha=-\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i},\beta=-\frac{1}{2}+\frac{\sqrt{3}}{2}\mathrm{i}$, 令 $f(x)=(x+1)^n-x^n-1$, 则当 $n=6m+5$ 时,

$$\begin{aligned}f(\alpha)&=(\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i})^n-(-\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i})^n-1\\&=(\cos(-\frac{\pi}{3})+\mathrm{i}\sin(-\frac{\pi}{3}))^n-(\cos\frac{4\pi}{3}+\sin\frac{4\pi}{3}\mathrm{i})^n-1\\&=\cos(-\frac{n\pi}{3})+\mathrm{i}\sin(-\frac{n\pi}{3})-(\cos\frac{4n\pi}{3}+\sin\frac{4n\pi}{3}\mathrm{i})-1\\&=\cos(-\frac{(6m+5)\pi}{3})+\mathrm{i}\sin(-\frac{(6m+5)\pi}{3})\\&-(\cos\frac{4(6m+5)\pi}{3}+\sin\frac{4(6m+5)\pi}{3}\mathrm{i})-1\\&=\cos(-\frac{5\pi}{3})+\mathrm{i}\sin(-\frac{5\pi}{3})-(\cos\frac{2\pi}{3}+\sin\frac{2\pi}{3}\mathrm{i})-1\\&=\frac{1}{2}+\frac{\sqrt{3}}{2}\mathrm{i}-(-\frac{1}{2}+\frac{\sqrt{3}}{2}\mathrm{i})-1=0\end{aligned}$$

同理可知 $f(\beta)=0$, 故 $f(x)=(x+1)^n-x^n-1$ 可被 $x-\alpha$ 与 $x-\beta$ 整除, 从而被 $(x-\alpha)(x-\beta)=x^2+x+1$ 整除。令 $x=\frac{x}{y}$, 可知 x^2+xy+y^2 整除 $(x+y)^n-x^n-y^n$.

又 $f'(x)=n(x+1)^{n-1}-nx^{n-1}$, 则当 $n=6m+1$ 时,

$$\begin{aligned}f'(\alpha)&=n(\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i})^{n-1}-n(-\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i})^{n-1}\\&=n[(\cos(-\frac{\pi}{3})+\mathrm{i}\sin(-\frac{\pi}{3}))^{n-1}-(\cos\frac{4\pi}{3}+\sin\frac{4\pi}{3}\mathrm{i})^{n-1}]\\&=n[\cos(-\frac{(n-1)\pi}{3})+\mathrm{i}\sin(-\frac{(n-1)\pi}{3})-(\cos\frac{4(n-1)\pi}{3}+\sin\frac{4(n-1)\pi}{3}\mathrm{i})]\\&=n[\cos(-\frac{6m\pi}{3})+\mathrm{i}\sin(-\frac{6m\pi}{3})-(\cos\frac{4\times 6m\pi}{3}+\sin\frac{4\times 6m\pi}{3}\mathrm{i})]\\&=n[1--1]=0\end{aligned}$$

同理可知 $f'(\beta)=0$ 且 $f(\alpha)=f(\beta)=0$, 故 $(x-\alpha)^2(x-\beta)^2=(x^2+x+1)^2$ 整除。令 $x=\frac{x}{y}$, 可知 $(x^2+xy+y^2)^2$ 整除 $(x+y)^n-x^n-y^n$.

很多同学对于 $n=6m+1$ 情况都想尝试归纳法证明，但是单重的归纳法应该是不可行的（助教没看到有做对的），但是 hyw 同学尝试了双重归纳法来证明结论是可行的，具体方法在习题课上讲。

作业 4.5 (P16 T4(1))

求多项式 $f(x)$ 与 $g(x)$ 的最大公因式:

$f(x)=x^4+x^3-3x^2-4x-1, g(x)=x^3+x^2-x-1$

解 对多项式 $f(x)$ 和 $g(x)$ 用辗转相除法如下:

$$
\begin{aligned}
& & f(x) &= x^4+x^3-3x^2-4x-1\\
& & g(x) &= x^3+x^2-x-1\\
q_1(x) &= x & r_1(x) &= -2x^2-3x-1\\
q_2(x) &= -\tfrac{1}{2}x+\tfrac{1}{4} & r_2(x) &= -\tfrac{3}{4}x-\tfrac{3}{4}\\
q_3(x) &= \tfrac{8}{3}x+\tfrac{4}{3} & r_3(x) &= 0
\end{aligned}
$$

故最大公因式为 $x+1$。

作业 4.6 (P16 T5(1)(3))

求多项式 $u(x)$ 与 $v(x)$, 使得 $f(x)u(x)+g(x)v(x)=d(x)$, 这里 $d(x)$ 是多项式 $f(x)$ 与 $g(x)$ 的最大公因式。

(1) $f(x)=x^4+2x^3-x^2-4x-2, g(x)=x^4+x^3-x^2-2x-2$

(3) $f(x)=3x^3-2x^2+x+2, g(x)=x^2-x+1$

解 (1) 对多项式 $f(x)$ 和 $g(x)$ 用辗转相除法如下:

$$
\begin{aligned}
f(x) &= x^4+2x^3-x^2-4x-2\\
g(x) &= x^4+x^3-x^2-2x-2\\
f(x) &= g(x)+(x^3-2x)\\
g(x) &= (x+1)(x^3-2x)+(x^2-2)\\
x^3-2x &= x(x^2-2)
\end{aligned}
$$

将以上各式整理后可得

$$
f(x)(-x-1)+g(x)(x+2)=x^2-2,
$$

故 $u(x)=-x-1, v(x)=x+2, d(x)=x^2-2$。

(3) 对多项式 $f(x)$ 和 $g(x)$ 用辗转相除法如下:

$$
\begin{aligned}
f(x) &= 3x^3-2x^2+x+2\\
g(x) &= x^2-x+1\\
f(x) &= (3x+1)g(x)-(x-1)\\
g(x) &= x(x-1)+1\\
x-1 &= (x-1)\cdot 1
\end{aligned}
$$

将以上各式整理后可得

$$f(x)x+g(x)(-3x^2-x+1)=1,$$

故 $u(x)=x, v(x)=-3x^2-x+1, d(x)=1$。

作业 4.7 (P16 T6(1)(3))

用待定系数法确定多项式 $u(x)$ 与 $v(x)$，使得 $f(x)u(x)+g(x)v(x)=1$, 其中 $f(x)$ 与 $g(x)$ 如下：

(1) $f(x)=x^3, g(x)=(1-x)^2$

(3) $f(x)=x^4-4x^3+1, g(x)=x^3-3x^2+1$

解 (1) 设 $u(x)=a_0+a_1x+a_2x^2, v(x)=b_0+b_1x+b_2x^2+b_3x^3$ 满足

$$x^3u(x)+(1-x)^2v(x)=1$$

我们有

$$\begin{aligned}&b_0+(b_1-2b_0)x+(b_0-2b_1+b_2)x^2+(a_0+b_1-2b_2+b_3)x^3\\&+(a_1+b_2-2b_3)x^4+(a_2+b_3)x^5=1\end{aligned}$$

取 $b_3=0$ 即可解得 $u(x)=-3x+4, v(x)=3x^2+2x+1$.

(3) 设 $u(x)=a_0+a_1x+a_2x^2, v(x)=b_0+b_1x+b_2x^2+b_3x^3$ 满足

$$(x^4-4x^3+1)u(x)+(x^3-3x^2+1)v(x)=1$$

类似得

$$\begin{cases}a_0+b_0=1\\a_1+b_1=0\\a_2-3b_0+b_2=0\\-4a_0+b_0-3b_1+b_3=0\\a_0-4a_1+b_1-3b_2=0\\a_1-4a_2+b_2-3b_3=0\\a_2+b_3=0\end{cases}\Rightarrow\begin{cases}a_0=\frac{26}{3}\\a_1=\frac{37}{3}\\a_2=-\frac{16}{3}\\b_0=-\frac{23}{3}\\b_1=-\frac{37}{3}\\b_2=-\frac{53}{3}\\b_3=\frac{16}{3}\end{cases}$$

于是 $u(x)=-\frac{16}{3}x^2+\frac{37}{3}x+\frac{26}{3}, v(x)=\frac{16}{3}x^3-\frac{53}{3}x^2-\frac{37}{3}x-\frac{23}{3}$。

作业 4.8 (P16 T7(1))

求次数最低的多项式 $u(x)$ 与 $v(x)$，使得

$$(x^4-2x^3-4x^2+6x+1)u(x)+(x^3-5x-3)v(x)=x^4$$

解 对多项式 $f(x)$ 和 $g(x)$ 用辗转相除法如下：

$$
\begin{aligned}
f(x) &= x^4 - 2x^3 - 4x^2 + 6x + 1 \\
g(x) &= x^3 - 5x - 3 \\
f(x) &= (x-2)g(x) + (x^2 - x - 5) \\
g(x) &= (x+1)(x^2 - x - 5) + (x+2) \\
x^2 - x - 5 &= (x-3)(x+2) + 1 \\
x + 2 &= (x+2) \cdot 1
\end{aligned}
$$

整理后得到

$$
f(x)(x^2 - 2x - 2) + g(x)(-x^3 + 4x^2 - 3x - 1) = 1
$$

为了使得 $u(x), v(x)$ 次数最低，故令

$$
\begin{aligned}
u(x) &= x^4(x^2 - 2x - 2) - h(x)(x^3 - 5x - 3) \\
v(x) &= -x^4(x^3 - 4^2 + 3x + 1) + h(x)(x^4 - 2x^3 - 4x^2 + 6x + 1)
\end{aligned}
$$

其中取 $h(x)$ 使得 $u(x), v(x)$ 次数尽可能小，即约去尽可能多的高次项，得到 $h(x) = x^3 - 2x^2 + 3x - 7$，于是

$$
\begin{aligned}
u(x) &= 9x^2 - 26x - 21 \\
v(x) &= -9x^3 + 44x^2 - 39x - 7
\end{aligned}
$$

若 $u(x), v(x)$ 不是次数最低的多项式，则存在 $u'(x), v'(x)$，使得 $\deg u > \deg u' \deg(v) > \deg(v')$，不妨 $\deg u > \deg u'$，则

$$
(x^4 - 2x^3 - 4x^2 + 6x + 1)(u(x) - u'(x)) + (x^3 - 5x - 3)(v(x) - v'(x)) = 0
$$

由于 $(x^4 - 2x^3 - 4x^2 + 6x + 1, x^3 - 5x - 3) = 1$，故 $x^3 - 5x - 3 | u(x) - u'(x)$，但 $\deg(u - u') = 2$ 且 $u(x) - u'(x) \neq 0$，不可能！所以所求 $u(x), v(x)$ 即为次数最低的满足条件的多项式。

作业 4.9 (P16 T9)

求次数最低的多项式 $f(x)$，使得 $f(x)$ 被多项式 $x^4 - 2x^3 - 2x^2 + 10x - 7$ 除时余式为 $x^2 + x + 1$，被多项式 $x^4 - 2x^3 - 3x^2 + 13x - 10$ 除时余式为 $2x^2 - 3$。

解 设

$$
\begin{aligned}
f(x) &= g(x)(x^4 - 2x^3 - 2x^2 + 10x - 7) + (x^2 + x + 1) \\
&= h(x)(x^4 - 2x^3 - 3x^2 + 13x - 10) + (2x^2 - 3)
\end{aligned}
$$

于是 $g(x)(x^4 - 2x^3 - 2x^2 + 10x - 7) - h(x)(x^4 - 2x^3 - 3x^2 + 13x - 10) = x^2 - x - 4$。对两多项式用辗转相除法如下：

$$
\begin{aligned}
x^4-2x^3-2x^2+10x-7&=(x^4-2x^3-3x^2+13x-10)+(x^2-3x+3)\\
x^4-2x^3-3x^2+13x-10&=(x^2+x-3)(x^2-3x+3)+(x-1)\\
x^2-3x+3&=(x-2)(x-1)+1\\
x-1&=(x-1)\cdot 1
\end{aligned}
$$

我们得到

$$(x^3-x^2-5x+7)(x^4-2x^3-2x^2+10x-7)-(x^3-x^2-4x+5)(x^4-2x^3-3x^2+13x-10)=1$$

于是

$$
\begin{aligned}
&(x^5-2x^4-8x^3+16x^2+13x-28)(x^4-2x^3-2x^2+10x-7)\\
&-(x^5-2x^4-7x^3+13x^2+11x-20)(x^4-2x^3-3x^2+13x-10)=x^2-x-4
\end{aligned}
$$

从而 $g(x)=x^5-2x^4-8x^3+16x^2+13x-28, h(x)=x^5-2x^4-7x^3+13x^2+11x-20$, 而

$$
\begin{aligned}
&f_0(x)=g(x)(x^4-2x^3-2x^2+10x-7)+(x^2+x+1)\\
=&(x^5-2x^4-8x^3+16x^2+13x-28)(x^4-2x^3-2x^2+10x-7)+(x^2+x+1)\\
=&x^9-4x^8-6x^7+46x^6-30x^5-152x^4+246x^3+75x^2-370x+197
\end{aligned}
$$

而 $(x^4-2x^3-2x^2+10x-7)(x^4-2x^3-3x^2+13x-10)=x^8-4x^7-x^6+33x^5-57x^4-22x^3+171x^2-191x+70$。最后得到次数最低的多项式

$$
\begin{aligned}
f(x)&=f_0(x)-x(x^4-2x^3-2x^2+10x-7)(x^4-2x^3-3x^2+13x-10)\\
&=-5x^7+13x^6+27x^5-130x^4+75x^3+266x^2-440x+197
\end{aligned}
$$

接着证明得到的函数次数最低。

设 $f_1(x)$ 也是满足题意的多项式，可得

$$
\begin{cases}
x^4-2x^3-2x^2+10x-7 \mid f(x)-f_1(x)\\
x^4-2x^3-3x^2+13x-10 \mid f(x)-f_1(x)
\end{cases}
$$

则

$$[x^4-2x^3-2x^2+10x-7, x^4-2x^3-3x^2+13x-10] \mid f(x)-f_1(x)$$

又因为

$$(x^4-2x^3-2x^2+10x-7, x^4-2x^3-3x^2+13x-10)=1,$$

所以

$$(x^4-2x^3-2x^2+10x-7)(x^4-2x^3-3x^2+13x-10) \mid f(x)-f_1(x)$$

所以 $\deg(f(x)-f_1(x))\geq 8$ 或 $f(x)=f_1(x)$。由于 $\deg f(x)=7$, 所以 $\deg f(x)\leq \deg f_1(x)$.

作业 4.10 (P16 T8)

求次数最低的多项式 $f(x)$，使得 $f(x)$ 被多项式 $(x-1)^2$ 除时余式为 $2x$，被多项式 $(x-2)^3$ 除时余式为 $3x$。

解 设 $f(x)=g(x)(x-1)^2+2x=h(x)(x-2)^3+3x$, 则有

$$g(x)(x-1)^2-h(x)(x-2)^3=x$$

对两多项式用辗转相除法如下：

$$\begin{aligned}(x-2)^3&=(x-4)(x-1)^2+(3x-4)\\(x-1)^2&=(\frac{1}{3}x-\frac{2}{9})(3x-4)+\frac{1}{9}\\3x-4&=(27x-36)\cdot\frac{1}{9}\end{aligned}$$

整理后得到

$$(3x^2-14x+17)(x-1)^2-(3x-2)(x-2)^3=1$$

于是

$$(3x^3-14x^2+17x)(x-1)^2-(3x^2-2x)(x-2)^3=x$$

从而 $g(x)=3x^3-14x^2+17x, h(x)=3x^2-2x$, 而 $f(x)=g(x)(x-1)^2+2x=3x^5-20x^4+48x^3-48x^2+19x$.

接着,

$$\begin{aligned}&3x^5-20x^4+48x^3-48x^2+19x-3(x-1)^2(x-2)^3\\=&4x^4-27x^3+66x^2-65x+24\end{aligned}$$

即为所要求次数最低的多项式 $f(x)$.

作业 4.11 (P104 T1)

设 $A\in\mathbb{F}^{2\times 2}$。证明：$A^2-\mathrm{tr}(A)A+(\det A)I_2=0$.

解 直接运算即可，注意到

$$\begin{aligned}A^2&=\begin{pmatrix}a_{11}^2+a_{12}a_{21} & a_{11}a_{12}+a_{12}a_{22}\\a_{11}a_{21}+a_{21}a_{22} & a_{12}a_{21}+a_{22}^2\end{pmatrix},\\\mathrm{tr}(A)A&=\begin{pmatrix}a_{11}^2+a_{11}a_{22} & a_{11}a_{12}+a_{12}a_{22}\\a_{11}a_{21}+a_{21}a_{22} & a_{11}a_{22}+a_{22}^2\end{pmatrix},\\(\det A)I_n&=\begin{pmatrix}a_{11}a_{22}-a_{12}a_{21} & 0\\0 & a_{11}a_{22}-a_{12}a_{21}\end{pmatrix}\end{aligned}$$

验证即可

作业 4.12 (P104 T2)

证明：不存在 $A, B \in \mathbb{R}^{n\times n}$，使得 $AB - BA = I_n$。

解 假设存在满足题意的 A, B，则

$$n = \mathrm{tr}(I_n) = \mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0,$$

矛盾。

4.1.2 第七周

作业 4.13 (P22 T1(2)(4))

把下列复系数多项式分解为一次因式的乘积：

(2) $(x+1)^n + (x-1)^n$

(4) $x^{2n} + C_{2n}^2 x^{2n-2}(x^2-1) + C_{2n}^4 x^{2n-4}(x^2-1)^2 + \cdots + (x^2-1)^n$

解 (2) 该多项式的点必为虚数，故有

$$(\frac{x+1}{x-1})^n = -1 = e^{i\pi}$$

于是多项式的根为

$$\frac{e^{\frac{i(2k+1)\pi}{n}} + 1}{e^{\frac{i(2k+1)\pi}{n}} - 1}, k = 0, \cdots, n-1$$

处理得多项式得根为

$$i \tan \frac{2k+1}{2n}\pi, k = 0, \cdots, n-1$$

注意 $\tan \pi/2$ 无意义，故最后多项式分解为

$$(x+1)^n + (x-1)^n = \begin{cases} 2\prod_{k=0}^{n-1}(x - \mathrm{i}\tan(\frac{2k+1}{2n}\pi)), & \text{当} n = 2l \\ 2x\prod_{k=0,k\neq l-1}^{n-1}(x - \mathrm{i}\tan(\frac{2k+1}{2n}\pi)), & \text{当} n = 2l-1 \end{cases}, l \in N_+$$

(4) 注意到

$$\begin{aligned} & x^{2n} + C_{2n}^2 x^{2n-2}(x^2-1) + C_{2n}^4 x^{2n-4}(x^2-1)^2 + \cdots + (x^2-1)^n \\ = & \frac{1}{2}[(x + \sqrt{x^2-1})^{2n} + (x - \sqrt{x^2-1})^{2n}] \\ = & \frac{1}{2}(x^2-1)[(\frac{x}{\sqrt{x^2-1}} + 1)^{2n} + (\frac{x}{\sqrt{x^2-1}} - 1)^{2n}] \end{aligned}$$

由 (2) 可知方程的解为 $\frac{x}{\sqrt{x^2-1}} = \mathrm{i}\tan(\frac{2k+1}{4n}\pi)$, 即

$$x = \pm\sin(\frac{2k+1}{4n}\pi), k = 0, 1, 2, \cdots, n-1$$

于是

$$\begin{aligned}&x^{2n}+C_{2n}^{2}x^{2n-2}(x^2-1)+C_{2n}^{4}x^{2n-4}(x^2-1)^2+\cdots+(x^2-1)^n\\&=2^{2n-1}\prod_{k=0}^{n-1}(x-\sin\frac{2k+1}{4n}\pi)(x+\sin\frac{2k+1}{4n}\pi)\end{aligned}$$

这里一定一定要注意系数问题，不能求出来根了之后直接写。

作业 4.14 (P22 T2(2)(4)(6))

把下列实系数多项式分解为实的不可约因式的乘积：

(2) x^6+27

(4) $x^{2n}-2x^n+2$

(6) $x^{2n}+x^n+1$

解 (2)

$$\begin{aligned}x^6+27&=(x^2+3)(x^4-3x^2+9)\\&=(x^2+3)(x^4-3x^2+9)\\&=(x^2+3)[(x^4+6x^2+9)-9x^2]\\&=(x^2+3)[(x^2+3)^2-(3x)^2]\\&=(x^2+3)(x^2-3x+3)(x^2+3x+3)\end{aligned}$$

(4) 令

$$x^{2n}-2x^n+2=(x^n-1)^2+1=0,$$

解得 $x^n=1+\mathrm{i}=\sqrt{2}(\cos\frac{\pi}{4}+\mathrm{i}\sin\frac{\pi}{4})=\sqrt{2}e^{\mathrm{i}\frac{\pi}{4}}$ 或 $x^n=1-\mathrm{i}=\sqrt{2}(\cos(-\frac{\pi}{4})+\mathrm{i}\sin(-\frac{\pi}{4}))=\sqrt{2}e^{-\mathrm{i}\frac{\pi}{4}}$。我们得到 $x=2^{\frac{1}{2n}}e^{\mathrm{i}(\frac{\pi}{4n}+\frac{2k\pi}{n})}$ 或 $x=2^{\frac{1}{2n}}e^{\mathrm{i}(-\frac{\pi}{4n}+\frac{2k\pi}{n})}$, 其中 $k=0,1,\cdots,n-1$ 于是

$$x^{2n}-2x^n+2=\prod_{k=0}^{n-1}(x^2-(2^{\frac{2n+1}{2n}}\cos(\frac{\pi}{4n}+\frac{2k\pi}{n}))x+2^{\frac{1}{n}})$$

(6) 由

$$x^{2n}+x^n+1=(x^n+\frac{1}{2})^2+\frac{3}{4}=0$$

解得

$$x^n=-\frac{1}{2}-\mathrm{i}\frac{\sqrt{3}}{2}=e^{\mathrm{i}\frac{4\pi}{3}},x^n=-\frac{1}{2}+\mathrm{i}\frac{\sqrt{3}}{2}=e^{\mathrm{i}\frac{2\pi}{3}},$$

其中 $k=0,1,2,\cdots,n-1$。故

$$x=e^{\mathrm{i}(\frac{4\pi}{3n}+\frac{2k\pi}{n})},x=e^{\mathrm{i}(\frac{2\pi}{3n}+\frac{2k\pi}{n})}$$

于是

$$x^{2n}+x^n+1=\prod_{k=0}^{n-1}(x^2-(2\cos(\frac{2\pi}{3n}+\frac{2k\pi}{n}))x+1)$$

作业 4.15 (P22 T3)

证明复系数多项式 $f(x)$ 对所有实数 x 恒取正值的充分必要条件是：存在没有实数根的复系数多项式 $\phi(x)$，使得 $f(x)=|\phi(x)|^2$.

解 证法一：

充分性：我们记 $\varphi(x)=a(x)+ib(x)$ 其中 $a(x),b(x)$ 均为实系数多项式，则有：

$$f(x)=(a(x))^2+(b(x))^2$$

倘若 $f(x)$ 可以取到 0，那么 $a(x),b(x)$ 有一个公共实根 x_0, 这使得 $\varphi(x_0)=0$ 和它没有实数根矛盾。

必要性：首先记 $f(x)=g(x)+ih(x)$ 其中 $g(x),h(x)$ 均为实系数多项式，那么倘若 $h(x)$ 不为 0，避开它的零点任取一实数 x_0, 就有 $f(x_0)$ 不为正值，从而 $h(x)\equiv 0$，下面做实数域上的唯一分解：

$$f(x)=a\prod_{k=1}^{m}(x-x_k)^{\alpha_k}\prod_{k=1}^{n}(x^2+b_kx+c_k),b_k^2-4c_k<0,k=1,2,\cdots,n$$

则首先 $a>0,\alpha_k$ 均为偶数，否则可以取到负值，下面注意到：

$$x^2+b_kx+c_k=(x+\frac{b_k}{2})^2+(\sqrt{c_k-\frac{b_k^2}{4}})^2$$

利用公式：

$$(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2$$

归纳可得：

$$\prod_{k=1}^{n}(x^2+b_kx+c_k)=(s(x))^2+(t(x))^2$$

从而有：

$$f(x)=(\sqrt{a}s(x)\prod_{k=1}^{m}(x-x_k)^{\alpha_k/2})^2+(\sqrt{2}t(x)\prod_{k=1}^{m}(x-x_k)^{\alpha_k/2})^2$$

$$\Leftrightarrow f(x)=\left|\sqrt{a}s(x)\prod_{k=1}^{m}(x-x_k)^{\alpha_k/2}+i\sqrt{a}t(x)\prod_{k=1}^{m}(x-x_k)^{\alpha_k/2}\right|^2$$

证法二：设 $f(x)=f_1(x)+\mathrm{i}f_2(x)$, 这里 $f_1(x)f_2(x)$ 属于 $\mathrm{R}[x]$, 那么对任意 x_0 属于 R, $f(x_0)=f_1(x_0)+if_2(x_0)>0$, 说明 $f_2(x)$ 恒等于 0，即 f 其实为实系数多项式。因其没有实根，且恒为正，则

$$f(x)=c(x-c_1)(x-c_1')(x-c_2)(x-c_2')\cdots(x-c_n)(x-c_n'),$$

其中 c_i 与 c_i' 为共轭复根且 $c>0$。记 $(x-c_1)(x-c_2)\cdots(x-c_n)=g_1(x)+\mathrm{i}g_2(x)$

$$(x-c_1')(x-c_2')\cdots(x-c_n')=g_1(x)-\mathrm{i}g_2(x),$$

则

$$f(x)=c(g_1^2(x)+g_2^2(x))=\left|\sqrt{c}(g_1(x)+\mathrm{i}g_2(x))\right|^2,$$

取

$$\phi(x)=\sqrt{c}(g_1(x)+\mathrm{i}g_2(x))$$

作业 4.16 (P22 T4)

证明实系数多项式 $f(x)$ 对所有实数 x 恒取非负实数值的充分必要条件是：存在实系数多项式 $\varphi(x)$ 和 $\psi(x)$, 使得 $f(x)=[\varphi(x)]^2+[\psi(x)]^2$。

解 必要性。因实系数多项式 $f(x)$ 对所有实数恒取非负实数值。不妨让 $f(x)$ 在 $\mathbb{R}$ 上分解因式

$$f(x)=a(x-x_1)^{l_1}\cdots(x-x_s)^{l_s}(x^2+a_1x+b_1)^{e_1}\cdots(x^2+a_kx+b_k)^{e_k},$$

其中 $l_1+\cdots+l_s+2(e_1+\cdots+e_k)=\deg f(x)$ 且二次因式在 $\mathbb{R}$ 上不可约, $x_i(1\le i\le s)$ 为实根。这里可设 $x_1<x_2<\cdots<x_s$。注意到 $x^2+a_jx+b_j>0(1\le j\le k)$ 即 $(x^2+a_1x+b_1)^{e_1}\cdots(x^2+a_kx+b_k)^{e_k}$, 而

$$\frac{f(x)}{(x^2+a_1x+b_1)^{e_1}\cdots(x^2+a_kx+b_k)^{e_k}}=a(x-x_1)^{l_1}\cdots(x-x_s)^{l_s}$$

在 $\mathbb{R}$ 上非负。取 $x=x_0<x_1$, 则得 $a>0$; 取 $x=\frac{x_1+x_2}{2}$, 则得 l_1 为偶数。同理可得 $l_m(1\le m\le s)$ 为偶数。考虑 $g(x)=(x^2+a_1x+b_1)^{e_1}\cdots(x^2+a_kx+b_k)^{e_k}$, 因其没有实根, 则只有共轭复根, 不妨设

$$g(x)=[(x-z_1)(x-\overline{z_1})]^{e_1}\cdots[(x-z_k)(x-\overline{z_k})]^{e_k},$$

记

$$g_1(x)=(x-z_1)^{e_1}\cdots(x-z_k)^{e_k}=h(x)+\text{ is }(x),$$

则

$$\overline{g_1(x)}=(x-\overline{z_1})^{e_1}\cdots(x-\overline{z_k})^{e_k}=h(x)-\text{ is }(x),$$

其中 $h(x),s(x)\in\mathbb{R}[x]$。故

$$g(x)=g_1(x)\overline{g_1(x)}=h^2(x)-s^2(x),$$

所以

$$f(x)=[\sqrt{a}(x-x_1)^{l_1/2}\cdots(x-x_s)^{l_s/2}h(x)]^2+[\sqrt{a}(x-x_1)^{l_1/2}\cdots(x-x_s)^{l_s/2}s(x)]^2,$$

这里记 $\varphi(x)=\sqrt{a}(x-x_1)^{l_1/2}\cdots(x-x_s)^{l_s/2}h(x),\phi(x)=\sqrt{a}(x-x_1)^{l_1/2}\cdots(x-x_s)^{l_s/2}s(x)$.

充分性是显然的, 另外, 用此题结论便可证明第三题.

作业 4.17 (P26 T1(2)(3)(5))

利用 Eisenstein 判别准则判定下述整系数多项式的不可约性:

(2) x^4-x^3+2x+1;

(3) x^4+1;

(5) $\sum_{i=0}^{p-1}(x+1)^i$, 其中 p 是素数。

♡

解 (2) 令 $x=y+1$, 则 $x^4-x^3+2x+1=(y+1)^4-(y+1)^3+2(y+1)+1=y^4+3y^3+3y^2+3y+3$。取 $p=3$, 显然 3 不能整除 $y^4+3y^3+3y^2+3y+3$ 的首项系数, $p^2=9$ 不能整除常数项, 而 $p=3$ 整除首项系数外其他各项, 根据 Eisenstein 判别准则, 多项式在 $\mathbb{Z}$ 上不可约.

(3) 令 $x=y+1$, 我们有

$$x^4+1=(y+1)^4+1=y^4+4y^3+6y^2+4y+2,$$

取 $p=2$, 显然 2 不能整除 $y^4+4y^3+6y^2+4y+2$ 的首项系数, $p^2=4$ 不能整除常数项, 而 $p=2$ 整除首项系数外其他各项, 根据 Eisenstein 判别准则, 多项式在 $\mathbb{Z}$ 上不可约.

(5) 由于

$$\begin{aligned}f(x)&=\sum_{i=0}^{p-1}(x+1)^i=\frac{(x+1)^p-1}{x}\\&=x^{p-1}+C_p^1x^{p-2}+\cdots+C_p^{k-1}x^{p-k}+\cdots+C_p^{p-2}x+p\end{aligned}$$

显然 p 不能整除 $f(x)$ 的首项系数, p^2 不能整除常数项, 但 p 整除首项系数外其他各项, 根据 Eisenstein 判别准则, 多项式在 $\mathbb{Z}$ 上不可约.

作业 4.18 (P26 T2)

设 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$ 是整系数多项, 且素数 p 满足: $p\nmid a_0,p\nmid a_1,\ldots,p\nmid a_k,p\mid a_i,i=k+1,k+2,\ldots,n$, 而 $p^2\nmid a_n$。证明 $f(x)$ 具有次数不低于 $n-k$ 的整系数不可约因式。

♡

解 实际上, 该题可以加强, 将条件改为 $p\nmid a_0,p\mid a_i,\quad i=k+1,k+2,\cdots,n,p^2\nmid a_n$, 我们通过类似归纳的方式叙述该题的证明:

首先叙述这样一个论断: 若 f 分解为整系数多项式 g 和 h 的乘积

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

则 g 与 h 之一的次数 $\geq n-k$, 并且该多项式 (对同样的 p 和 $n-k$) 满足类似的三个的条件

我们来证明这一论断。设

$$g(x)=b_0x^s+b_1x^{s-1}+\ldots+b_{s-1}x+b_s$$
$$h(x)=c_0x^r+c_1x^{r-1}+\ldots+c_{r-1}x+c_r$$

其中 $s+r=n$。因为 $p\,|a_n=b_sc_r, p^2\big|\,a_n=b_sc_r$, 所以, 不妨设 $p\mid b_s, p^2\nmid b_s, p\nmid c_r$。又 $p\nmid a_0=b_0c_0$, 所以, 不妨设 b_{s-m} 是 $b_s, b_{s-1},\cdots,b_0$ 之中第一个不被 p 整除者, 考察

$$a_{n-m}=a_{s+r-m}=b_{s-m}c_r+b_{s-m+1}c_{r-1}+\cdots+b_sc_{r-m}$$

可知 $p\nmid a_{n-m}$, 因而 $n-m\le k$, 即 $m\ge n-k$, 这时 g 的次数大于 $n-k$。我们确认多项式 g 满足以下三个条件:

(i) $p\nmid b_0$;

(ii) $p\mid b_j(j=s,s-1,\cdots,s+k+1-n)$;

(iii) $p^2\nmid b_s$.

若 g 不可约, 则定理的结论已证实, 否则可重复类似讨论, 也就是继续对 $g(x)$ 进行分解，同样可以得到一个比 $g(x)$ 次数更小的，满足条件的多项式，次数是严格递减的。直到得出一个次数 $\ge n-k$ 的不可约因式。证毕.

当然，也可以直接对问题讨论，将 $g(x),h(x)$ 直接设出来考虑每一项关于 p 的整除性，具体的方式在习题课上讲。

作业 4.19 (P26 T6)

设整系数多项式 $f(x)$ 在 x 的 4 个不同整数值上都取值为 1，则 $f(x)$ 在 x 的其它整数值上的值不可能是 -1。

解 将这四个整数分别记为 a_1,a_2,a_3,a_4, 设 $f(x)-1=g(x)$, 则有 $g(a_1)=g(a_2)=g(a_3)=g(a_4)=0$。又因为 a_1,a_2,a_3,a_4 互不相同, 故 $g(x)=(x-a_1)(x-a_2)(x-a_3)(x-a_4)h(x)$，所以

$$g(n)=(n-a_1)(n-a_2)(n-a_3)(n-a_4)h(n)=f(n)-1$$

由题意, $n-a_1,n-a_2,n-a_3,n-a_4$ 这四个数彼此不同, 其中最多只能有两个分别取作 $1,-1$, 而其他两个不能在 $-1,1$ 这两个数中取值, 故取得的 $|g(n)|$ 的最小值只能为 $(-1)\times1\times(-2)\times2=4$, 从而 $f(n)-1=g(n)$ 不为 -2, $f(n)\neq-1$.

作业 4.20 (P26 T7)

证明，设正整数 $n\ge12$，并且 n 次整系数多项式 $f(x)$ 在 x 的 $[\frac{n}{2}]+1$ 个以上的整数值上取值为 ±1，则 $f(x)$ 在 $\mathbb{Q}$ 不可约。次数 n 的下界 12 是否还可缩小？

解 当 $n\ge12$ 时, $[\frac{n}{2}]+1\ge7$, 则其中必有不少于 4 个整数取函数值 1 或 -1 。由上题结论可知, 若 $[\frac{n}{2}]+1$ 个函数值中有 4 个 1，则函数值不可能为 -1，同理可知, 若其中有 4 个 -1，则函数值不可能为 1 。设 $m=[\frac{n}{2}]+1$ 个函数值全为 1，分别记为 $a_1,a_2,\cdots,a_m$, 设 $f(x)$ 在 $\mathbb{Q}$ 上可约, 则 $f(x)$ 在 $\mathbb{Z}$ 上可约, 因此, $f(x)=g(x)h(x)$, 其中 $g(x),h(x)\in\mathbb{Z}[x]$, 且 $\deg g(x)<\deg f(x), \deg h(x)<\deg f(x)$.

显然

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_m)k(x)+1=g(x)h(x)$$

那么有 $g(a_i)h(a_i)=1$, 则 $g(a_i)=h(a_i)=1$ 或 $g(a_i)=h(a_i)=-1$, 同理可知 $g(a_i)$ 要么全为 1，要么全为 -1 。这样 $g(x)-1$ 或 $g(x)+1$ 至少有 $[\frac{n}{2}]+1$ 个根。同理可知 $h(x)-1$ 或 $h(x)+1$ 至少有 $[\frac{n}{2}]+1$ 个根。因此 $\min\{\deg g(x),\deg h(x)\}\geq[\frac{n}{2}]+1$, 然而

$$n=\deg f(x)=\deg g(x)+\deg h(x)\geq 2[\frac{n}{2}]+2>n,$$

矛盾。

次数 n 的下界事实上可以缩小到 8，下证当 $n=8,9,10,11$ 时命题均成立，而当 $n\geq 7$ 时均有反例说明不可约。

我们一次性用反证法证明 $n=8,9,10,11$ 情况。由上述证明我们可以知道，假设 $f(x)$ 可约，$f(x)=g(x)h(x)$，则若 $g(x)$ 或者 $h(x)$ 在 4 个以上整点满足全为 1 或者全为-1，则可以推出 $f(x)$ 在 $\mathbb{Q}$ 上不可约，从而矛盾。而 $[\frac{8}{2}]+1=5$，故 $g(x),h(x)$ 在至少 5 个整点上为 1 或者-1，不妨设 $g(x)$ 在 3 个整点 x_1,x_2,x_3 为 1，在 x_4,x_5 上为-1，不妨设 $x_1<x_2<x_3,x_4<x_5$，这五个点互不相同，则可以设

$$g(x)=a(x)(x-x_1)(x-x_2)(x-x_3)+1$$

从而

$$a(x_4)(x_4-x_1)(x_4-x_2)(x_4-x_3)=-2$$

故 $a(x_4)=\pm 1$，x_1,x_2,x_3,x_4 必为连续的四个整数点，且 $x_1<x_4<x_3$ 但

$$a(x_5)(x_5-x_1)(x_5-x_2)(x_5-x_3)=-2$$

从而 x_1,x_2,x_3,x_5 也必须为连续的四个整数点，且 $x_1<x_5<x_3$，不可能！故假设不成立，原命题成立。

对于 $n\leq 7$，我们给出所有反例如下：

1. $n=7$ $f(x)=(x(x-2)(x-3)+1)(x(x-1)(x-2)(x-3)+1)$
2. $n=6$ $f(x)=(x(x-2)(x-3)+1)^2$
3. $n=5$ $f(x)=(2x^2-1)(x(x-1)(x-2)+1)$
4. $n=4$ $f(x)=(2x^2-1)^2$
5. $n=3$ $f(x)=x^3$
6. $n=2$ $f(x)=x^2$

4.2 补充内容

这两周的作业以及未来很多课程作业都可以使用 Mathematica 进行计算，一方面可以检查正确与否，一方面可以提高完成作业的效率，是一个非常强大的计算软件。本次习题课会讲一下 Mathematica 如何下载使用，以及基础的一些函数和如何使用它来检验作业答案。

4.2.1 安装与使用

科大买了 Mathematica 正版软件，可以直接下载使用，无需从网上购买盗版的。点击链接进入科大正版软件页面下载 Mathematica：https://software.ustc.edu.cn/zbh.php。下载好后需要激活注册，点击**邮箱激活 (可校外)**，遵照指示步骤激活即可使用。

Mathematica 的编程非常简单，只需要会 C 语言即可上手。需要注意的是，其中的函数均为大写开头，参数使用中括号，如 Sin[]、Cos[]。Mathematica 基本能实现所有目前出现的运算，包括微分方程中的 ODE 和 PDE 问题，实现一些不会写的计算可以上网搜，咨询 GPT，我建议学会查询使用手册。打开方法为：点击菜单栏-帮助 (H)-Wolfram 参考资料 (D) 进入，在其中搜索想要的函数即可。

图 4.1: 正版软件安装

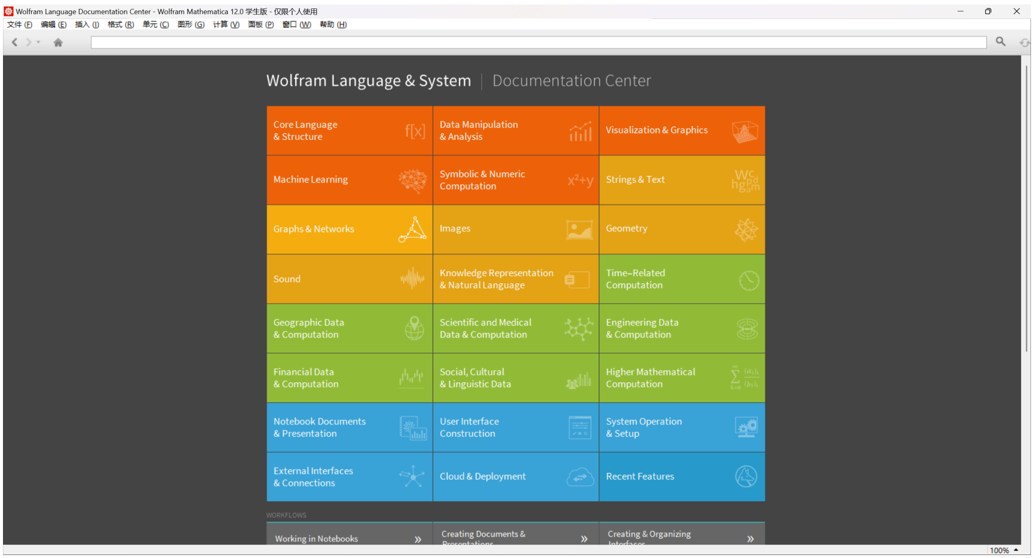


图 4.2: 帮助文档

4.2.2 例子

给出了求解待定系数问题和因式分解问题的例子。展示了如何求 ODE 和 PDE 数值解方式，以及对于矩阵做运算的方法，求矩阵逆，秩，特征值和特征向量，Jordan 分解的方法，具体内容见下方展示：

习题课

解线性方程组

P16 T6(3)

In[ ]:=
```
Solve[{a0 + b0 == 1, a1 + b1 == 0, a2 - 3 b0 + b2 == 0,
解方程
  -4 a0 + b0 - 3 b1 + b3 == 0, a0 - 4 a1 + b1 - 3 b2 == 0,
  a1 - 4 a2 + b2 - 3 b3 == 0, a2 + b3 == 0},
 {a0, a1, a2, b0, b1, b2, b3}]
```

Out[ ]=
$$\left\{\left\{a0 \to \frac{26}{3},\ a1 \to \frac{37}{3},\ a2 \to -\frac{16}{3},\ b0 \to -\frac{23}{3},\ b1 \to -\frac{37}{3},\ b2 \to -\frac{53}{3},\ b3 \to \frac{16}{3}\right\}\right\}$$

验证

```
f[x_] = x^4 - 4 x^3 + 1;
g[x_] = x^3 - 3 x^2 + 1;
u[x_] = -16 / 3 x^2 + 37 / 3 x + 26 / 3;
v[x_] = 16 / 3 x^3 - 53 / 3 x^2 - 37 / 3 x - 23 / 3;
result = f[x] × u[x] + v[x] × g[x]
Expand[result]
展开
```

Out[ ]=
$$\left(1-3x^2+x^3\right)\left(-\frac{23}{3}-\frac{37x}{3}-\frac{53x^2}{3}+\frac{16x^3}{3}\right)+\left(\frac{26}{3}+\frac{37x}{3}-\frac{16x^2}{3}\right)\left(1-4x^3+x^4\right)$$

Out[ ]=
1

因式分解

In[ ]:=
```
Clear[f]
清除
n = 10;
f[x_] = (x + 1) ^n + (x - 1) ^n
Factor[f[x]]
因式分解
TraditionalForm[Factor[f[x]]]
传统格式         因式分解
```

Out[ ]= $(-1 + x)^{10} + (1 + x)^{10}$

Out[ ]= $2 \left(1 + x^2\right) \left(1 + 44 x^2 + 166 x^4 + 44 x^6 + x^8\right)$

Out[ ]//TraditionalForm=

$$2 \left(x^2 + 1\right) \left(x^8 + 44 x^6 + 166 x^4 + 44 x^2 + 1\right)$$

求解微分方程问题

12. 求解下列微分方程并画出解函数。

(1) $\begin{cases} y''(x)+y(x)=\cos x \\ y(0)=0 \\ y'(0)=0 \end{cases} \quad (x\in[0,20])$

(2) $\begin{cases} \dfrac{du}{dt}=0.09u\left(1-\dfrac{u}{20}\right)-0.45uv \\ \dfrac{dv}{dt}=0.06v\left(1-\dfrac{v}{15}\right)-0.001uv \\ u(0)=1.6 \\ v(0)=1.2 \end{cases}$

(3) 热方程

$$\frac{\partial u(t,x)}{\partial t}=\frac{\partial^2 u(t,x)}{\partial x^2}$$

$$u(0,x)=0,\quad u(t,0)=\sin t,\quad u(t,5)=0$$

$$t\in[0,10],\quad x\in[0,5]$$

(4) 波动方程

$$\frac{\partial^2 u(t,x)}{\partial t^2}=\frac{\partial^2 u(t,x)}{\partial x^2}$$

$$u(0,x)=e^{-x^2},\quad u(t,-10)=u(t,10),\quad \left.\frac{\partial u(t,x)}{\partial t}\right|_{t=0}=0$$

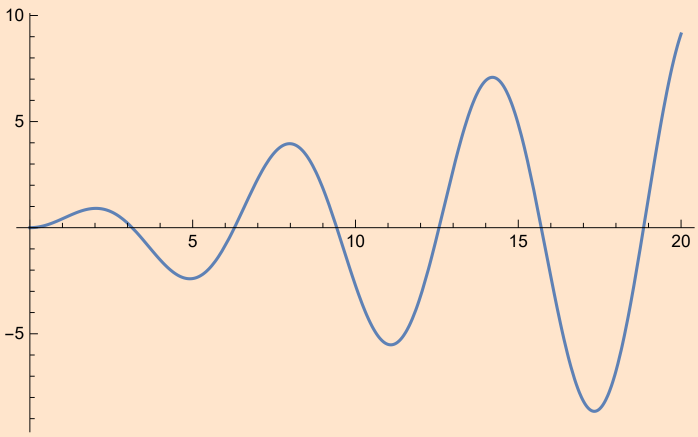
$$t\in[0,40],\quad x\in[-10,10]$$

(1)

In[∘]:=

```
S1 = NDSolve[{y''[x] + y[x] == Cos[x],
        数值求解微分方程组                余弦
     y[0] == 0, y'[0] == 0}, y, {x, 0, 20}];
f = S1[[1, 1, 2]];
Plot[f[x], {x, 0, 20}]
绘图
```

Out[∘]=

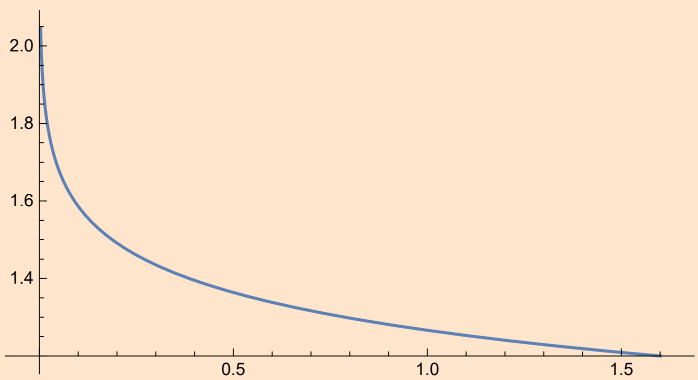


(2)

In[∘]:=

```
S2 = NDSolve[{u'[t] == 0.09 u[t] (1 - u[t] / 20) - 0.45 u[t] × v[t],
        数值求解微分方程组
     v'[t] == 0.06 v[t] (1 - v[t] / 15) - 0.001 u[t] × v[t],
     u[0] == 1.6, v[0] == 1.2}, {u, v}, {t, 0, 10}];
u = S2[[1, 1, 2]];
v = S2[[1, 2, 2]];
ParametricPlot[{u[t], v[t]}, {t, 0, 10}]
绘制参数图
```

Out[∘]=

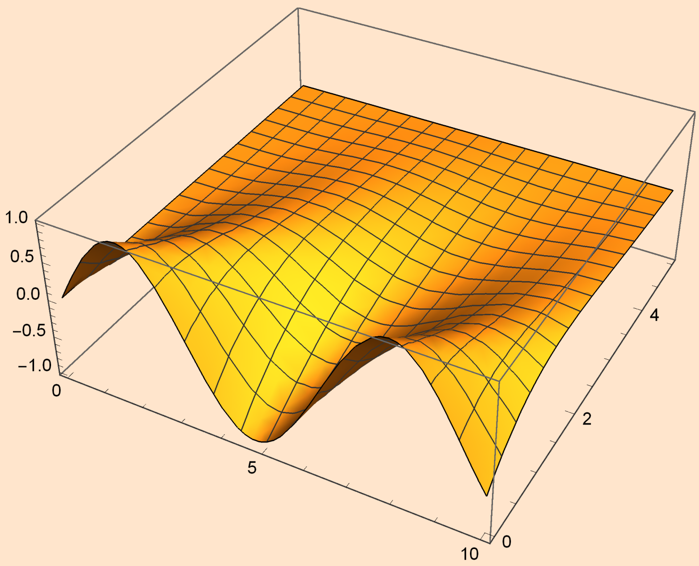


(3)

In[ ]:=
```
Remove[u];
NDSolve[{D[u[t, x], t] == D[u[t, x], x, x], u[0, x] == 0,
  u[t, 0] == Sin[t], u[t, 5] == 0}, u, {t, 0, 10}, {x, 0, 5}]
Plot3D[Evaluate[u[t, x] /. %], {t, 0, 10},
 {x, 0, 5}, PlotRange → All]
```
去除 数值求… 偏导 偏导 正弦 绘… 计算 绘制范围 全部

Out[ ]= {{u → InterpolatingFunction[Domain: {{0., 10.}, {0., 5.}} Output: scalar]}}

Out[ ]=



(4)

In[ ]:=

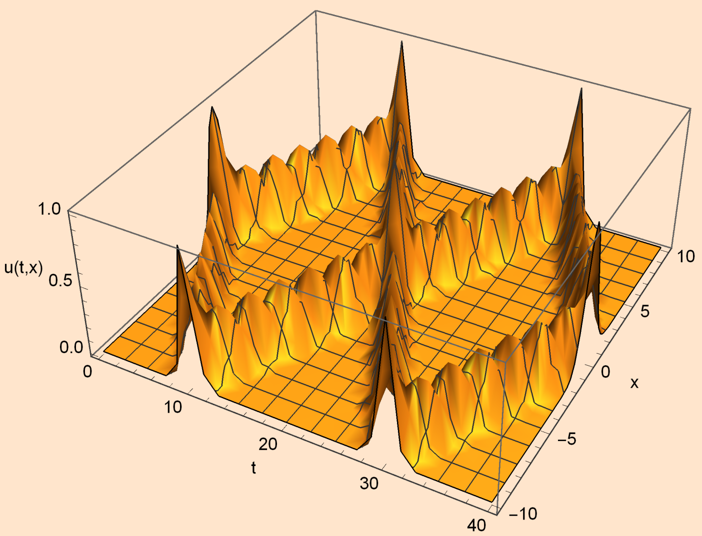
```
Remove[u];
去除
sol = NDSolve[{D[u[t, x], {t, 2}] == D[u[t, x], {x, 2}],
      数值求…   偏导                   偏导
   u[0, x] == E^(-x^2), u[t, -10] == u[t, 10],
             自然常数
   Derivative[1, 0][u][0, x] == 0}, u, {t, 0, 40}, {x, -10, 10}]
   导数

Plot3D[u[t, x] /. sol, {t, 0, 40}, {x, -10, 10},
绘制三维图形
 PlotRange → All, AxesLabel → {"t", "x", "u(t,x)"}]
 绘制范围     全部  坐标轴标签
```

Out[ ]=

{{u → InterpolatingFunction[Domain: {{0., 40.}, {-10., 10.}} Output: scalar]}}

Out[ ]=



矩阵运算

基本运算，求秩

In[ ]:=
```
A = {{1, 0, -1}, {0, 2, 3}};
B = {{2, -1, 4}, {1, 0, -2}, {0, 3, 1}};
Dot[A, B] // MatrixForm
                矩阵格式
Dot[B, B] // MatrixForm
                矩阵格式
MatrixRank[B]
矩阵的秩
```

Out[ ]//MatrixForm=

$$\begin{pmatrix} 2 & -4 & 3 \\ 2 & 9 & -1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} 3 & 10 & 14 \\ 2 & -7 & 2 \\ 3 & 3 & -5 \end{pmatrix}$$

Out[ ]= 3

求逆

In[ ]:=

```
n = 10;
A = Table[If[i ≤ j, j - i + 1, 0], {i, n}, {j, n}];
     表格  如果
A // MatrixForm
     矩阵格式
Inverse[A] // MatrixForm
逆              矩阵格式
```

Out[ ]//MatrixForm=

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

特征值和特征向量

In[ ]:=
```
A = Table[If[i + j == 4, 1, 0], {i, 3}, {j, 3}];
      表格   如果
A // MatrixForm
       矩阵格式
Eigenvalues[A] // MatrixForm
特征值                矩阵格式
v1 = Eigenvectors[A][[1]] // MatrixForm
     特征向量                矩阵格式
v2 = Eigenvectors[A][[2]] // MatrixForm
     特征向量                矩阵格式
v3 = Eigenvectors[A][[3]] // MatrixForm
     特征向量                矩阵格式
```

Out[ ]//MatrixForm=

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Jordan分解

In[ ]:=

```
A = {{3, -4, 0, 2}, {4, -5, -2, 4}, {0, 0, 3, 2}, {0, 0, 2, -1}};
{P, J} = JordanDecomposition[A];
            约旦分解
P // MatrixForm
      矩阵格式
J // MatrixForm
      矩阵格式
```

Out[ ]//MatrixForm=

$$\begin{pmatrix} 1 & \frac{1}{4} & \frac{-1+3\sqrt{2}}{-3+2\sqrt{2}} & \frac{1+3\sqrt{2}}{3+2\sqrt{2}} \\ 1 & 0 & \frac{1+2\sqrt{2}}{-3+2\sqrt{2}} & \frac{-1+2\sqrt{2}}{3+2\sqrt{2}} \\ 0 & 0 & 1-\sqrt{2} & 1+\sqrt{2} \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Out[ ]//MatrixForm=

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1-2\sqrt{2} & 0 \\ 0 & 0 & 0 & 1+2\sqrt{2} \end{pmatrix}$$

第 5 章 特征方阵

5.1 习题讲解

5.1.1 第八周

作业 5.1 (P29 T1)

设多项式 $f(x_1,\dots,x_n),g(x_1,\dots,x_n)\in\mathbb{F}[x_1,\dots,x_n]$，证明 $fg=0\Rightarrow f=0$ 或 $g=0$。♡

解 若 f,g 均非零，考虑字典序下，f 的首项为 $a\prod_i {x_i}^{m_i}$，g 的首项为 $b\prod_i {x_i}^{n_i}$，其中 $a,b\neq 0$。分析次数大小关系可知 fg 首项为 $ab\prod_i {x_i}^{m_i+n_i}$，于是非零，矛盾。

作业 5.2 (P35 T2)

证明三次实系数方程 $x^3+ax^2+bx+c=0$ 的每个根实部都是负数的充要条件为

$$a>0,ab-c>0,c>0$$

♡

解 记三根为 x_1,x_2,x_3，由 Vièta 定理计算知此三条件等价于

$$x_1+x_2+x_3<0,(x_1+x_2)(x_2+x_3)(x_1+x_3)<0,x_1x_2x_3<0$$

必要性：若三根实部都是负数，不妨设 $x_1<0$ 为实根，无论 x_2,x_3 是实根还是共轭复根都可算得 $x_2+x_3<0,x_2x_3>0,(x_1+x_2)(x_1+x_3)>0$，因此三条件均成立。

充分性：分类讨论，当三根均为实根时，由第三个条件知必须两正一负或全负，若两正一负，不妨设 x_1,x_2 为正，则由第一个条件可知 $-x_3>x_1+x_2$，从而 $x_1+x_3<0,x_2+x_3<0$，与第二个条件矛盾；当有一个实根与一对共轭复根时，不妨设 x_1 为实，由 $x_2x_3\geq 0$ 与第三个条件可知 $x_1<0$，再结合 $(x_1+x_2)(x_1+x_3)\geq 0$ 与第二个条件即有 x_2,x_3 实部为负。

作业 5.3 (P35 T3)

设方程 $x^3+ax^2+bx+c=0$ 的三个根是某个三角形内角的正弦，证明

$$a(4ab-a^3-8c)=4c^2$$

♡

解 设三个根为 $\sin A,\sin B,\sin C$，代入可得到要证

$$-\sin^4 A-\sin^4 B-\sin^4 C+2\sin^2 A\sin^2 B+2\sin^2 B\sin^2 C+2\sin^2 A\sin^2 C=4\sin^2 A\sin^2 B\sin^2 C$$

利用 $\sin C=\sin(A+B)$ 展开化简可得证，或由正弦定理、余弦定理得到

$$\sin^2 C=\sin^2 A+\sin^2 B-2\sin A\sin B\cos C$$

两侧平方后化简亦可。

作业 5.4 (补充题)

矩阵 $A=(a_{ij})$ 满足

$$a_{ij}=\begin{cases}1 & j=i+1\\ x & j=i\\ -1 & j=i-1\\ 0 & \text{otherwise.}\end{cases}$$

计算 $2\det A$。

解 对 n 阶的情况记 $\det A=\Delta_n$，则对首行首列 Laplace 展开可知 $\Delta_n=x\Delta_{n-1}+\Delta_{n-2}$。利用特征方程考虑数列递推可解得

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

作业 5.5 (P35 T1(2)(4)(6))

把下列对称多项式化为基本对称多项式的多项式：

(2) $(2x_1-x_2-x_3)(2x_2-x_1-x_3)(2x_3-x_1-x_2)$；

(4) $\sum_{1\le i<j\le n}(x_i-x_j)^2$；

(6) $\sum_{1\le i<j\le n,k\ne i,j}(x_i+x_j-x_k)^2$。

解 (2) 直接利用教材算法可得为 $2\sigma_1^3-9\sigma_1\sigma_2+27\sigma_3$。

(4) 由次数知可设其为 $A\sigma_1^2+B\sigma_2$。代入 x_i 均为 1 可知 $n^2A+\frac{n(n-1)}{2}B=0$；代入 $x_1=1$，其余为 0 可知 $A=n-1$，从而解得其为 $(n-1)\sigma_1^2-2n\sigma_2$。

(6) 由次数知可设其为 $A\sigma_1^2+B\sigma_2$。代入 $x_1=1$，其余为 0，则 $i=1$ 有 $(n-1)(n-2)$ 种可能，$k=1$ 有 $\frac{(n-1)(n-2)}{2}$ 种可能，每种可能各贡献一个 1，因此 $A=\frac{3}{2}(n-1)(n-2)$。代入 x_i 均为 1，则考虑项数可知 $n^2A+\frac{n(n-1)}{2}B=\frac{n(n-1)}{2}(n-2)$。综合两式解得其为 $\frac{3}{2}(n-1)(n-2)\sigma_1^2-(n-2)(3n-1)\sigma_2$。

作业 5.6 (P35 T4)

设 $x_1,\dots,x_n$ 是 $f(x)=a_0+\cdots+a_{n-1}x^{n-1}+x^n$ 的 n 个根，证明关于 $x_2,\dots,x_n$ 的多项式可以表示为关于 x_1 的多项式。

解 由对称多项式的表示，只需说明 $x_2,\dots,x_n$ 的基本对称多项式 $\sigma_1',\dots,\sigma_{n-1}'$ 可用 x_1 的多项式表示。根据 Vièta 定理，$x_1,\dots,x_n$ 的基本对称多项式 $\sigma_k=(-1)^ka_{n-k}$ 为常数，而由定义算得

$$\sigma_1'=\sigma_1-x_1,\sigma_k'=\sigma_k-x_1\sigma_{k-1}',k=2,\dots,n-1$$

从而归纳可得结论。

作业 5.7 (P36 T10)

求多项式 $f(x)=x^n+\sum_{i=0}^{n-1}(a^{n-i}+b^{n-i})x^i$ 的根的等幂和 $s_1,\dots,s_n$。

解 我们归纳证明

$$s_k = \begin{cases} -(a^k + b^k) & k = 2l+1 \\ -(a^l - b^l)^2 & k = 2l \end{cases}$$

$k=1$ 时直接知成立，由牛顿恒等式

$$s_k = \sigma_1 s_{k-1} - \sigma_2 s_{k-2} + \cdots + (-1)^k \sigma_{k-1} s_1 + (-1)^{k+1} k\sigma_k$$

而由 Viète 定理 $\sigma_k = (-1)^k(a^k + b^k)$，合并可得

$$s_k = -\big((a+b)s_{k-1} + (a^2+b^2)s_{k-2} + \cdots + (a^{k-1}+b^{k-1})s_1 + k(a^k+b^k)\big)$$

利用归纳假设可发现右侧只有形如 $a^m b^{k-m}$ 的项，直接对每个 m 对比系数得结论。

此题最困难的步骤或许是猜测出结果。

5.1.2 第九周

作业 5.8 (P287 T1(1)(3)(5)(6))

计算下列 λ 矩阵的 Smith 标准形，并求出行列式因子、不变因子、初等因子组：

(1) $\operatorname{diag}\left(\lambda(\lambda+1), \lambda, (\lambda+1)^2\right)$

(3) $\begin{pmatrix} \lambda-2 & -1 & \\ & \lambda-2 & -1 \\ & & \lambda-2 \end{pmatrix}$

(5) $\begin{pmatrix} \lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \lambda & -1 & \\ & & & \lambda & -1 \\ 1 & 2 & 3 & 4 & 5+\lambda \end{pmatrix}$

(6) $a_{ij} = \begin{cases} \lambda & i=j \\ 1 & i<j \\ 0 & i>j \end{cases}$

解 (1) 直接计算行列式因子可知

$$D_1(\lambda) = 1, D_2(\lambda) = \lambda(\lambda+1), D_3(\lambda) = \lambda^2(\lambda+1)^3$$

于是

$$d_1(\lambda) = 1, d_2(\lambda) = \lambda(\lambda+1), d_3(\lambda) = \lambda(\lambda+1)^2$$

由此即得 Smith 标准形，而初等因子组为

$$\lambda, \lambda+1, \lambda, (\lambda+1)^2$$

(3) 直接计算行列式因子可知

$$D_1(\lambda) = 1, D_2(\lambda) = 1, D_3(\lambda) = (\lambda-2)^3$$

于是

$$d_1(\lambda) = 1, d_2(\lambda) = 1, d_3(\lambda) = (\lambda - 2)^3$$

由此即得 Smith 标准形，而初等因子组为 $(\lambda - 2)^3$。

(5) 直接计算行列式因子，由于一到四阶行列式可取右上角 4×4 方阵的主子式，必有 $D_i(\lambda) = 1, i = 1, 2, 3, 4$，而计算行列式有

$$D_5(\lambda) = \lambda^5 + 5\lambda^4 + 4\lambda^3 + 3\lambda^2 + 2\lambda + 1$$

由此 $d_i(\lambda) = D_i(\lambda)$，从而有 Smith 标准形。

可验证 $D_5(\lambda)$ 无有理根，考虑 $\gcd(D_5'(\lambda), D_5(\lambda))$ 可知其无重根，于是记其根为 $\lambda_1, \dots, \lambda_5$，初等因子组为

$$\lambda - \lambda_1, \dots, \lambda - \lambda_5$$

(6) 设其为 n 阶，直接计算行列式因子。由于 1 到 $n-1$ 阶行列式可取右上角 $(n-1) \times (n-1)$ 方阵的主子式，必有 $D_i(\lambda) = 1, i < n$，而 $D_n(\lambda) = \lambda^n$，于是 $d_i(\lambda) = D_i(\lambda)$，从而有 Smith 标准形，初等因子组为 λ^n。

注意初等因子组是每个不变因子分解的结果，可以重复。

作业 5.9 (P287 T3)

对 n 阶 Hermite 阵 H_1, H_2，证明 λ 方阵 $H_1 + \lambda H_2$ 的不变因子都是实系数多项式。

解 记 $A = H_1 + \lambda H_2$，考虑 A 的 k 阶子式

$$f(\lambda) = A\begin{pmatrix} i_1 & \dots & i_k \\ j_1 & \dots & j_k \end{pmatrix}, g(\lambda) = A\begin{pmatrix} j_1 & \dots & j_k \\ i_1 & \dots & i_k \end{pmatrix}$$

由 H_1, H_2 为 Hermite 阵可知 $a_{ij} = \overline{a_{ji}}$ (这里多项式的共轭代表每个系数取共轭)，从而有 $g(\lambda) = \overline{f(\lambda)}$，于是

$$\gcd(f, g) = \gcd(f, \bar{f}) = \gcd(f + \bar{f}, \mathrm{i}(f - \bar{f}))$$

因此 $\gcd(f, g)$ 是实系数多项式。由此，A 的任何一个行列式因子都可以写成若干个实系数多项式的最大公因式，仍然是实系数多项式，由此即得结论。

也可利用分解证明 A^H 的不变因子是 A 的不变因子的共轭，从而直接得结论。

作业 5.10 (P293 T1(2)(3)(6)(9))

求下列方阵的 Jordan 标准形：

(2) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$

(3) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$

(6) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ & & 3 & -2 \\ & & 2 & -1 \end{pmatrix}$

(9) $J_n(0)^2$

解 (2) 直接计算得为 $J_3(1)$。

(3) 直接计算得为 $\mathrm{diag}(J_1(3), J_2(-1))$。

(6) 直接计算得为 $\mathrm{diag}(J_2(1), J_2(-1))$。

(9) 设其为 $A = (a_{ij})$，由于 $a_{ij} = \begin{cases} 1 & j = i+2 \\ 0 & \text{otherwise.} \end{cases}$，考虑将 $1, 3, 5, \dots, 2[(n+1)/2] - 1$ 置换到 $1, 2, \dots, [n+1]/2$，将 $2, 4, \dots, 2[n/2]$ 置换到 $[n+1]/2+1, \dots, n$ 的置换 (注意到 $n - [(n+1)/2] = [n/2]$)，记其为 σ。

将此置换对应的初等变换阵记为 S，右乘其并左乘其逆即为对行列进行相同的置换，分析可得这样相似后即为

$$\mathrm{diag}(J_{[n+1]/2}(0), J_{[n/2]}(0))$$

最后一问也可以分析初等因子组完成，不过此类操作性的思路可以完全解决 $J_n(0)^k$ 的 Jordan 标准形，见补充内容。

作业 5.11 (P294 T3)

证明一组两两可交换的可对角化方阵可被同一个可逆方阵相似对角化。

解 先说明两个方阵 A, B 的情况。由于 $P^{-1}APP^{-1}BP = P^{-1}ABP$，$A, B$ 可交换等价于同时相似后可交换，由此可设 $P^{-1}AP$ 是对角阵 $A_0 = \mathrm{diag}(\lambda_1 I_1, \dots, \lambda_k I_k)$，$\lambda_i$ 互不相同，而 $P^{-1}BP = B_0$。

直接计算可知 B_0 需要可以写为 $\mathrm{diag}(B_1, \dots, B_k)$，其中 B_i 与 I_i 同阶，由 B 可相似对角化，B_0 可相似对角化，因此其每个对角块可相似对角化，不妨设 $Q_i^{-1}B_iQ_i$ 为对角阵 D_i，则令

$$R = P\,\mathrm{diag}(Q_1, \dots, Q_k)$$

验证可知 $R^{-1}AR = A_0$ 为对角阵，且 $R^{-1}BR = \mathrm{diag}(D_1, \dots, D_k)$ 为对角阵。

对多个方阵时，以 A, B, C 为例，类似上不妨设 A, B 已为对角阵，由 C 与 A 可交换可先将 C 分块为 $\mathrm{diag}(C_1, \dots, C_k)$，又由 B, C 可交换知 D_k, C_k 可交换，从而进一步细分可构造出同时将 A, B, C 对角化的过渡矩阵，对任意有限个矩阵时情况类似。

对无穷多个矩阵，注意到 P 若能同时对角化某些矩阵，也能对角化它们的线性组合，即对角化它们在 $\mathbb{R}^{n\times n}$ 中生成的子空间的任何方阵。此外，若一族矩阵两两可交换，它们生成的子空间中任何矩阵两两可交换。由于 $\mathbb{R}^{n\times n}$ 为有限维，其子空间必然有一组有限基，从而通过有限个矩阵的情况可得到结论。

作业 5.12 (P294 T7)

若 A 的特征值全为 1，k 为正整数，证明 A^k 与 A 相似。

解 考虑 A 的 Jordan 标准形知只需证明 $J_n(1)^k$ 与 $J_n(1)$ 相似。直接计算可知 $J_n(1)^k$ 是上三角矩阵，且对

角元为 1，$j=i+1$ 处为 k，于是 $\lambda I-J_n(1)^k$ 的前 $n-1$ 个行列式因子全为 1 (取右上角 $(n-1)\times(n-1)$ 方阵的主子式)，行列式为 $(\lambda-1)^n$，由此即得与 $J_n(1)$ 相似。

5.2 补充内容

5.2.1 多项式矩阵

一般环上矩阵知识点：

1. 模方阵 (环上的可逆方阵) **定义**；
2. **模方阵等价表述**：行列式是环上的可逆元；
3. Smith 标准形：**存在性**、**算法**；
4. 行列式因子、不变因子、初等因子组定义。

多项式矩阵知识点：

1. **特征方阵**定义；
2. 特征方阵相抵**等价**于矩阵相似；
3. 特征方阵计算 Jordan 标准形。

事实上，环上的矩阵与一个特殊的代数结构——**模**——存在密切的联系，由于课程中并不会涉及这一结构，我们一般也不会研究多项式矩阵更深的结论。不过，我们仍然可以观察一些稍显复杂的性质：

定理 5.13 (多项式方程求解)

对域 $\mathbb{F}$ 上的多项式矩阵 A 与向量 b，方程 $Ax=b$ 有多项式解等价于 (A,b) 与 $(A,\mathbf{0})$ 在多项式矩阵的意义下相抵 (也即书中相抵)。♡

证明 思路：通过设出 A 的 Smith 标准形构造分析，转化为 A 是对角阵时的情况。

此结论对整数亦成立。

定理 5.14 (模相抵的等价定理)

对域 $\mathbb{F}$ 上的多项式矩阵 A,B，A,B 相抵等价于存在矩阵 P_1,P_2,Q_1,Q_2 使得

$$A=P_1BP_2, B=Q_1AQ_2$$

♡

证明 左推右直接由 $A=PBQ,B=P^{-1}AQ^{-1}$ 可得。对右推左，假设 $S_1=X_1P_1Y_1,S_2=X_2P_2Y_2$ 是 P_1,P_2 的 Smith 标准形，X_1,Y_1,X_2,Y_2 可逆，则有 $X_1AY_2=S_1Y_1^{-1}BX_2^{-1}S_2$。由 S_1,S_2 对角，X_1AY_2 的子式一定为 $Y_1^{-1}BX_2^{-1}$ 子式的倍数，因此 X_1AY_2 的行列式因子对应为 $Y_1^{-1}BX_2^{-1}$ 行列式因子的倍数，即 A 行列式因子对应为 B 行列式因子倍数。同理 B 行列式因子对应为 A 行列式因子倍数，于是 A,B 非零行列式因子必然相同，即得证。

更多时候，我们将目光放在特征方阵上，希望用它简化 Jordan 标准形的研究，例如：

定理 5.15 (幂的 Jordan 标准形)

对 $\lambda\neq 0$，证明 $J_n(\lambda)^k$ 相似于 $J_n(\lambda^k)$。♡

证明 思路：仿照习题中对 $J_n(1)$ 的幂的计算进行。

例题 5.1 证明 $J_n(0)^k$ 的 Jordan 标准形为 $[n/k]$ 阶与 $[n/k]+1$ 阶 Jordan 块共 k 个 (零阶 Jordan 块视为不存在)。

提示：直接通过置换可找到算法，注意计算过程中的阶数特点。

5.2.2 Jordan 标准形

Jordan 标准形对矩阵论研究有两个最重要的好处：

1. 将完整的矩阵变为准对角阵，从而只需针对 Jordan 块研究；
2. 同时相似不影响多项式与可交换性等，因此可以不妨假设已经具有标准形式。

综合前一部分得到的幂次结论，我们来考虑一个有趣的问题：

例题 5.2 对给定的 k，求所有使得 $A^k=B$ 有解的复方阵 B。

提示：分为非零特征值与零特征值考虑。

为了能进行更深入的计算，我们需要一个 Jordan 块的多项式的基本性质：

定理 5.16 (Jordan 块的多项式)

设 f 为多项式，则

$$f(J_n(\lambda))=\begin{cases}\frac{f^{(j-i)}(\lambda)}{(j-i)!} & i\le j\\ 0 & i>j\end{cases}$$

这里上标括号代表求导，$0!=1$。♡

证明 思路：只需说明 $f(x)=x^n$ 的情况即可，此时可直接归纳计算。

由此，我们可以用多项式的极限定义矩阵的级数，从而建构类似数的分析性质，例如：

例题 5.3 解方程 $\mathrm{e}^A=I$。确定 $\ln A$ 的定义域，并求解 $\ln A=I$。

提示：可不妨设 A 已经相似为了标准形考虑。值得注意的是，$\mathrm{e}^x=0$ 在复数域中解为 $x=2k\pi\mathrm{i},k\in\mathbb{Z}$。

例题 5.4 试求所有的 B 使得存在 A 使得 $\mathrm{e}^A=B$。对 $\ln A=B$ 呢？

* **数值计算视角**：在数值计算的角度下，秩与 Jordan 标准形是“不好”的，因为它们不具有对矩阵的**连续性**。但是，多元线性微分方程的精确求解似乎又与 Jordan 标准形密切相关，这二者会产生矛盾吗？

最后，我们补充一些更困难的习题：

定理 5.17 (幂次自相似)

若 A 可逆，存在正整数 $i\ne j$ 使得 A^i 相似于 A^j，则存在正整数 k 使得 A^k 特征值都是 1。♡

证明 不妨设 $i>j$。由幂的特征值是特征值的幂，条件即 A 特征值的 i 次方与 j 次方相差排列。对任何特征值 λ，假设 $\lambda^i=\lambda_1^j,\lambda_1^i=\lambda_2^j,\dots$，由特征值个数有限必然存在 $\lambda_s,s\le n-1$ 使得 $\lambda_s^i=\lambda^j$，从而计算得 $\lambda^{i^{s+1}}=\lambda^{j^{s+1}}$。

于是，对 A 的任何特征值 λ 必然存在某 $t\le n$ 使得 $\lambda^{i^t-j^t}=1$，取 $k=i^{n!}-j^{n!}$ 即可。

定理 5.18 (Jordan 分解)

对任何方阵 A，存在唯一 B,C 使得 B 可对角化，C 幂零，且 $BC=CB,A=B+C$。♡

证明 不妨设 $A=P^{-1}JP$，J 为 Jordan 标准形，其对角元为 D，则取 $B=P^{-1}DP, C=P^{-1}(J-D)P$ 可验证符合要求。

为说明唯一性，我们先证明上述构造的 B,C 可以写成 A 的多项式。若 $f(A)=B$，则 $f(J)=D$，根据之前 Jordan 块多项式的结论，假设 J 不同的对角元为 $\lambda_1,\dots,\lambda_k$，$\lambda_i$ 对应的最大 Jordan 块为 n_i 阶，则 f 需要满足

$$f(\lambda_i)=\lambda_i, f'(\lambda_i)=0,\dots,f^{(n_i-1)}(\lambda_i)=0$$

对任何 λ_i 成立。

根据数值分析中的 Hermite 插值定理，存在唯一 $n_1+\cdots+n_k-1$ 次多项式满足上述要求，而 $B=f(A)$，则有 $C=A-f(A)$，于是 C 也能写成 A 的多项式。

若还有另一种分解 $A=B'+C'$ 满足条件，由 $B'C'=C'B'$ 可以推出 $B'A=B'(B'+C')=AB'$，同理 $C'A=AC'$，于是任何 A 的多项式都与 B',C' 可交换，即有 $BB'=B'B, CC'=C'C$。

根据习题，由于 B,B' 均可对角化，存在将其同时对角化的 P，于是 $B-B'$ 可对角化。另一方面，由于 $C^n=(C')^n=O$ 且其可交换，一定有 $(C'-C)^{2n}=O$ (其展开后每项至少包含一个 C^n 或 $(C')^n$)，于是 $C'-C$ 是幂零的。注意到 $B-B'=C'-C$，且幂零的可对角化矩阵一定是 O，即得证 $B=B', C=C'$。

此方法来自同学贡献，通过 Hermite 插值结论规避了繁杂处理的过程。然而，提出此方法的同学并没有学过数值分析中的 Hermite 插值，因此卡在了证明 B 是 A 的多项式的部分。我们下面看一个需要更多细节处理的唯一性证明过程：

证明 存在性证明与之前相同，为证唯一性，我们先说明若 $BC=CB$，且 B 可对角化，则存在 P 使得 $P^{-1}BP, P^{-1}CP$ 分别为 B,C 的相似标准形。此结论的证明与习题几乎相同，即先将 B 相似到标准形，接着说明 C 可对应分块，再将 C 的每个分块相似到标准形即可。

对任意方阵 A，设有满足条件的分解 $A=B+C$，用 P 将 B,C 同时相似到标准形后，对比形式可发现 $P^{-1}AP$ 必然为 A 的相似标准形，唯一确定，因此 B,C 的相似标准形 $D, J-D$ 唯一确定。

为进一步说明 B,C 唯一确定，假设 B',C' 也满足要求，且 Q 将 B',C' 同时相似至对应标准形，则有

$$B=PDP^{-1}, B'=QDQ^{-1}, A=PJP^{-1}, A=QJQ^{-1}$$

记 $F=P^{-1}Q$，对 $JF=FJ$ 分析可发现当第 i、j 个特征值不同时 f_{ij} 一定为 0，于是根据与对角阵可交换的条件知 $FD=DF$，也即 $B=B'$，得证。

第 6 章 线性空间

6.1 习题讲解

6.1.1 第十周

作业 6.1 (P294 T6)

设 A 可逆，证明 AB 与 BA 相似，A 不可逆时是否成立？

解 A 可逆时 $A^{-1}ABA = BA$，因此相似。A 不可逆时考虑

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

即可验证不相似。

作业 6.2 (P305 T3)

设 A, B 为 n 阶方阵，且 $\operatorname{diag}(A, A)$ 与 $\operatorname{diag}(B, B)$ 相似，求证 A, B 相似。

解 设 $P^{-1}AP$ 是 A 的相似标准形，则可发现 $\operatorname{diag}(P, P)\operatorname{diag}(A, A)\operatorname{diag}(P^{-1}, P^{-1})$ 是 $\operatorname{diag}(A, A)$ 的相似标准形，因此 $\operatorname{diag}(A, A)$ 的相似标准形为把 A 的每个 Jordan 块复制一遍，由此即能从 $\operatorname{diag}(A, A)$ 与 $\operatorname{diag}(B, B)$ 相似推出 A, B 相似。

作业 6.3 (P305 T4)

已知 A 的特征多项式 $\varphi(\lambda) = (\lambda - 2)^3(\lambda + 7)^2$，最小多项式 $d(\lambda) = (\lambda - 2)^2(\lambda + 7)$，求 Jordan 标准形。

解 A 的特征值 2 对应 Jordan 块共 3 阶，最大 2 阶；-7 对应 Jordan 块共 2 阶，最大一阶，于是标准形为 $\operatorname{diag}(J_2(2), 2, -7, -7)$。

作业 6.4 (P305 T5)

设 A 为 n 阶可逆方阵，且 A 与 A^k 相似，$k > 1$ 为整数，证明 A 的特征值均为单位根。

解 上章附加内容中定理 5.17 为此题的推广。

作业 6.5 (P163 T5)

设向量 $\alpha_1, \ldots, \alpha_k \in \mathbb{F}^n$ 线性相关，$k \geq 2$，说明任取 $\alpha_{k+1} \in \mathbb{F}^n$ 存在不全为 0 的 $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ 使得 $\alpha_1 + \lambda_1\alpha_{k+1}, \ldots, \alpha_k + \lambda_k\alpha_{k+1}$ 线性相关。

解 由题设存在 μ_i 使得 $\sum_i \mu_i\alpha_i = 0$，且 μ_i 不全为 0。我们只要说明存在不全为 0 的 λ_i 使得 $\sum_i \lambda_i\mu_i = 0$，即有

$$\sum_{i=1}^{k}(\alpha_i + \lambda_i\alpha_{k+1}) = 0$$

构造：只要存在 $\mu_t = 0$，可直接取 $\lambda_t = 1, \lambda_i = 0\ (i \neq t)$，否则可取 $\lambda_1 = -\mu_2, \lambda_2 = \mu_1, \lambda_i = 0\ (i >$

2)。

作业 6.6 (P164 T11)

求 $\alpha_1=(4,-1,3,2),\alpha_2=(8,-2,6,4),\alpha_3=(3,-1,4,-2),\alpha_4=(6,-2,8,-4)$ 的极大线性无关组。

解 将向量写成列向量，行变换可得到

$$\begin{pmatrix}4&8&3&6\\-1&-2&-1&-2\\3&6&4&8\\2&4&-2&-4\end{pmatrix}\Longrightarrow\begin{pmatrix}0&0&0&0\\-1&-2&0&0\\0&0&1&2\\0&0&0&0\end{pmatrix}$$

于是极大线性无关组为 $(1,2),(1,3),(2,3),(2,4)$。

作业 6.7 (P164 T13)

若 A,B 都是 $m\times n$ 矩阵，证明 $\operatorname{rank}(A+B)\le\operatorname{rank}A+\operatorname{rank}B$。

解 设 $\operatorname{rank}A=a,\operatorname{rank}B=b$，$A$ 的列向量极大线性无关组为 $\alpha_1,\dots,\alpha_a$，B 的列向量极大线性无关组为 $\beta_1,\dots,\beta_b$，则 A 的每列可写为 $\sum_j\lambda_{ij}\alpha_j$，$B$ 的每列可写为 $\sum_k\mu_{ik}\beta_k$，于是 $A+B$ 每列可写为

$$\sum_j\lambda_{ij}\alpha_j+\sum_k\mu_{ik}\beta_k$$

其可被所有 α_j,β_k 生成，极大线性无关组维数不超过 $a+b$，即得证。

作业 6.8 (P168 T4)

数域上二阶方程构成线性空间中，求一组基使得其中每个都满足 $A^2=A$。

解 尝试知可取 $\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&0\\1&1\end{pmatrix}$。

6.1.2 第十一周

作业 6.9 (P171 T5)

在 $\mathbb{F}^n$ 中，给定 n 个行向量 $\alpha_1,\dots,\alpha_n$ 可作为行拼成矩阵 A，反之亦然。证明 $\alpha_1,\dots,\alpha_n$ 构成 $\mathbb{F}^n$ 一组基等价于 A 可逆。

解 A 可逆等价于 A^T 可逆，根据线性方程知识可知等价于 $A^Tx=\mathbf{0}$ 无非零解。

而 $A^Tx=\sum_i x_i\alpha_i^T$，其无非零解等价于 $\{\alpha_i\}$ 线性无关，而又由空间为 n 维即知 n 个线性无关向量等价于基。

作业 6.10 (P176 T2)

$\mathbb{Q}$ 上线性空间 V_1,V_2 存在一一映射，则它们是否同构？

解 考虑 $\mathbb{Q},\mathbb{Q}^2$，维数不同，因此不同构，但势均为可列，因此存在一一映射。

作业 6.11 (P184 T2)

考虑 $\mathbb{F}$ 上所有方阵构成的线性空间 $\mathbb{F}^{n\times n}$，设对称方阵构成的集合即为 S，反对称方阵构成的集合记为 K，证明 S,K 都是 $\mathbb{F}^{n\times n}$ 的子空间，且 $S+K=\mathbb{F}^{n\times n}, S\cap K=\{O\}$，并求 $\dim S, \dim K$。♡

解 子空间直接通过加法、数乘封闭性验证即可，下面记 E_{ij} 为只有第 i 行第 j 列为 1，其他为 0 的方阵。

S 中元素 A 可写为 $\sum_i a_{ii}E_{ii}+\sum_{i<j}a_{ij}(E_{ij}+E_{ji})$，每个参数后的矩阵线性无关，因此构成一组基，维数为参数个数 $\frac{n(n+1)}{2}$。

K 中元素 A 可写为 $\sum_{i<j}a_{ij}(E_{ij}-E_{ji})$，每个参数后的矩阵线性无关，因此构成一组基，维数为参数个数 $\frac{n(n-1)}{2}$。

根据上方表达式直接计算可知 $S\cap K=\{O\}$，而对 $S+K$，由于任何方阵 $A=\frac{A+A^T}{2}+\frac{A-A^T}{2}$，可验证前者在 S 中，后者在 K 中，因此 $S+K=\mathbb{F}^{n\times n}$。

作业 6.12 (P185 T6)

设 U,V,W 是线性空间 L 的子空间，证明：

1. $U\cap(V+W)=U\cap V+U\cap W$ 未必成立；
2. $U\cap(V+(U\cap W))=U\cap V+U\cap W$ 恒成立。♡

解

1. 考虑 $\mathbb{R}^2$ 中 $U=\{(a,a),a\in\mathbb{R}\}, V=\{(a,0),a\in\mathbb{R}\}, W=\{(0,a),a\in\mathbb{R}\}$。
2. 左包含右：

$$U\cap V\subset U, U\cap W\subset U\Longrightarrow U\cap V+U\cap W\subset U$$

$$U\cap V\subset V\Longrightarrow U\cap V+U\cap W\subset V+U\cap W$$

右包含左：设左任一元素写成 $u=u_1+u_2$，其中 $u_1\in V, u_2\in U\cap W$，则

$$u_2\in U\cap W\Longrightarrow u_2\in U\xRightarrow{u\in U}u_1=(u-u_2)\in U$$
$$u_1\in V\Longrightarrow u_1\in U\cap V\xRightarrow{u_2\in U\cap W}u\in U\cap V+U\cap W$$

作业 6.13 (P185 T8)

设 U,V,W 是线性空间 L 的子空间，证明：

$$(U+V)\cap(U+W)=U+(U+V)\cap W$$ ♡

解 左包含右：

$$(U+V)\cap W\subset W\Longrightarrow U+(U+V)\cap W\subset U+W$$

$$(U+V)\cap W\subset U+V\Longrightarrow U+(U+V)\cap W\subset U+(U+V)=U+V$$

右包含左：设左任意元素写成 $u=u_1+u_2$，其中 $u_1\in U, u_2\in W$，则

$$u\in U+V, u_1\in U\xRightarrow{U\subset U+V}u_2=(u-u_1)\in U+V\xRightarrow{u_2\in W}u_2\in(U+V)\cap W$$
$$u_1\in U, u_2\in(U+V)\cap W\Longrightarrow u=(u_1+u_2)\in U+(U+V)\cap W$$

作业 6.14 (P185 T5)

考虑 $\mathbb{F}^n$ 中以下集合，若是子空间则确定维数并给出一组基；若否，则写出生成的子空间的一组基：

1. 所有分量和为 0 的向量；
2. 所有分量不同时大于 0 或不同时小于 0 的向量；
3. 存在大于 0 分量的向量。

解 记 e_i 表示第 i 个分量为 1，其余为 0 的向量。

1. 是子空间，维数为 $n-1$，基为 $\{e_1-e_i \mid 2\le i\le n\}$。
2. 不是子空间，考虑二维时 $2e_1-e_2, 2e_2-e_1$ 属于此集合，但两者之和 e_1+e_2 不属于。生成空间为 $\mathbb{F}^n$，因为 $\{ne_i-\sum_{j\neq i}e_j \mid 1\le i\le n\}$ 可以构成一组基 (其对应的矩阵为主角占优阵，必然可逆)。
3. 不是子空间，e_1 属于此集合，但 $-e_1$ 不属于。生成空间为 $\mathbb{F}^n$，因为 $\{e_i\}$ 可以构成一组基。

作业 6.15 (P185 T12)

设 α,β 是线性空间 V 中向量，W 为 V 子空间，α 与 W 生成子空间为 U，β 与 W 生成子空间为 K，证明

$$\beta\in U, \beta\notin W \Longrightarrow \alpha\in K$$

解 由 $\beta\in U$，存在 λ 与 $w\in W$ 使得 $\beta=\lambda\alpha+w$，若 $\lambda=0$ 则 $\beta\in W$，矛盾，从而

$$\alpha=\frac{1}{\lambda}\beta-\frac{1}{\lambda}w$$

由此即知 $\alpha\in K$。

作业 6.16 (P191 T2)

考虑 $\mathbb{F}^{2n}$ 的子空间

$$V=\{\alpha \mid \alpha_i=\alpha_{n+i}, 1\le i\le n\},\quad W=\{\alpha \mid \alpha_i=-\alpha_{n+i}, 1\le i\le n\}$$

证明 $\mathbb{F}^{2n}=V\oplus W$。

解 对任何 $\mathbb{F}^{2n}$ 中元素 (α,β)，这里 α,β 各 n 个分量，其存在表示

$$(\alpha,\beta)=\left(\frac{\alpha+\beta}{2},\frac{\alpha+\beta}{2}\right)+\left(\frac{\alpha-\beta}{2},\frac{\beta-\alpha}{2}\right)$$

前后分别在 V,W 中，于是 $V+W=\mathbb{F}^{2n}$。此外，求解 $V\cap W$ 可知其中只有 $\mathbf{0}$，于是得证。

作业 6.17 (P194 T2)

设 $A\in\mathbb{F}^{m\times n}, \beta\in\mathbb{F}^m$，且 $\operatorname{rank}(A,\beta)=\operatorname{rank}A$。记 $Ax=\mathbf{0}$ 的解空间为 V_A，设 α 是 $Ax=\beta$ 的一个特解，证明 $Ax=\beta$ 的所有解为 $\alpha+V_A$。

解 由秩条件可知有解，能取出 α，从而

$$Ax=\beta \Longleftrightarrow A(x-\alpha)=\mathbf{0} \Longleftrightarrow x-\alpha\in V_A \Longleftrightarrow x\in\alpha+V_A$$

作业 6.18 (P194 T3)

设 W 是 $\mathbb{F}^n$ 的子空间，证明存在 $A\in\mathbb{F}^{m\times n}$ 使得 $Ax=\mathbf{0}$ 的解空间为 W。

解 设 W 的一组基为 $\alpha_1,\dots,\alpha_t$，将其扩充 $\beta_{t+1},\dots,\beta_n$ 成为全空间一组基，并假设以这些基作为列向

量拼成的矩阵为 B。取

$$A = \mathrm{diag}(O_t, I_{n-t})B^{-1}$$

则对比 AB 与右侧的每一列可知 $A\alpha_i = 0$，而 $A\beta_i = e_i$。

设 $x = \sum_{i=1}^{t} \lambda_i\alpha_i + \sum_{i=t+1}^{n} \mu_i\beta_i$，则 $Ax = \sum_{i=t+1}^{n} \mu_i e_i$，于是其为 $\mathbf{0}$ 等价于 μ_i 全为 0，即 $x \in W$。

6.2 期中复习

进度截至第十二周，11 月 22 日课程内容。带 * 表示不强制掌握。

1. 矩阵
   (a). 基本定义、运算
   (b). 可逆性、矩阵的逆
   (c). 初等变换与初等变换阵
   (d). * 置换与置换方阵的对应
   (e). * Schur 变换

$$\begin{pmatrix} I & O \\ -CA^{-1} & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I & -A^{-1}B \\ O & I \end{pmatrix} = \begin{pmatrix} A & O \\ O & D - CA^{-1}B \end{pmatrix}$$

2. 行列式
   (a). Laplace 展开、完全展开
   (b). Binet-Cauchy 公式
   (c). 其他计算技巧：递推计算、加边、拆行等
   (d). 行列式刻画逆：伴随方阵、Crammer 法则
   (e). * 几何视角 [有向体积]、代数视角 [公理化定义]
3. 秩
   (a). 定义、等价定义
   (b). **Hermite 标准形**
   (c). 秩不等式
   (d). 线性方程组的解
   (e). 广义逆 $AA^-A = A, A^-AA^- = A^-$[具体定义应看题干]
   (f). 相抵标准形刻画广义逆
   (g). * 奇异值分解与广义逆 $(AA^-)^H = AA^-, (A^-A)^H = (A^-A)$
4. 相合
   (a). 实对称阵**相合标准形**
   (b). 实对称阵**正交相似标准形**
   (c). 相合与正交相似出发的研究
   (d). 二次型、正定与半正定
   (e). 正定与半正定等价定义
   (f). 同时相合对角化
   (g). * 与对角阵可交换的充要条件
   (h). * 极分解、奇异值分解

    (i). * 正定性相关不等式

5. 多项式
    (a). 域上多项式的基本性质
    (b). 带余除法、最大公因式
    (c). 因式分解、可约性
    (d). 有理系数多项式的可约性
    (e). 对称多项式
    (f). 多项式矩阵与 **Smith 标准形**
    (g). * 多项式操作技巧
6. 相似
    (a). 相似的基本性质
    (b). **相似标准形**
    (c). 特征方阵与相似标准形计算
    (d). * 相似标准形的多项式与幂级数
    (e). * 特征多项式、最小多项式
7. 线性空间
    (a). 基本定义
    (b). 线性相关性、秩
    (c). 基、维数、坐标
    (d). 子空间、交空间、和空间
    (e). 商空间、等价类
8. 线性映射
    (a). 定义与表示
    (b). 像与核
    (c). 对偶空间

第 7 章 线性映射

7.1 习题讲解

7.1.1 第十二周

作业 7.1 (P204 T9)

在所有 2 阶实方阵构成的线性空间 $\mathbb{R}^{2\times 2}$ 中, 取 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. 定义映射 $\mathscr{A}:\mathbb{R}^{2\times 2}\to\mathbb{R}^{2\times 2}$ 与 $\mathscr{B}:\mathbb{R}^{2\times 2}\to\mathbb{R}^{2\times 2}$ 如下: 设 $X\in\mathbb{R}^{2\times 2}$, 则令 $\mathscr{A}(X)=AX,\mathscr{B}(X)=XA$. 证明: 映射 $\mathscr{A}$ 与 $\mathscr{B}$ 都是线性的, 并求出它们在基 $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$. 下的方阵.

解 直接验证可知线性，设基为 $E_{11},E_{21},E_{12},E_{22}$

$$\mathscr{A}(E_{11},E_{21},E_{12},E_{22})=(E_{11},E_{21},E_{12},E_{22})\begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}$$

$$\mathscr{B}(E_{11},E_{21},E_{12},E_{22})=(E_{11},E_{21},E_{12},E_{22})\begin{pmatrix} a & 0 & c & 0 \\ 0 & a & 0 & c \\ b & 0 & d & 0 \\ 0 & b & 0 & d \end{pmatrix}$$

作业 7.2 (P204 T10)

$\mathbb{R}^{2\times 2}$ 中所有形如 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ 的方阵集合记为 V, 其中 $a,b\in\mathbb{R}$. 显然 V 是实数域 $\mathbb{R}$ 上线性空间. 视复数域 $\mathbb{C}$ 为实数域 $\mathbb{R}$ 上线性空间. 定义映射 $\mathscr{A}:\mathbb{C}\to V$ 如下: 设 $\alpha=a+\mathrm{i}b\in\mathbb{C},a,b\in\mathbb{R}$, 则令 $\mathscr{A}(\boldsymbol{\alpha})=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. 证明 $\mathscr{A}$ 是实线性空间 $\mathbb{C}$ 到 V 上的可逆线性映射, 并且对任意 $\alpha,\beta\in\mathbb{C},\mathscr{A}(\boldsymbol{\alpha},\boldsymbol{\beta})=\mathscr{A}(\boldsymbol{\alpha})\mathscr{A}(\boldsymbol{\beta})$.

解 线性映射 trivial

可逆 = 单射 + 满射

$$\mathscr{A}(\alpha\beta)=\mathscr{A}(a+\mathrm{i}b)(c+\mathrm{i}d)=\begin{pmatrix} ac-bd & ad+bc \\ -ad-bc & ac-bd \end{pmatrix}=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} c & d \\ -d & c \end{pmatrix}$$

作业 7.3 (P210 T3)

定义映射 $\mathscr{A}:\mathbb{F}[x]\to\mathbb{F}[x]$ 如下: 设 $f(x)\in\mathbb{F}[x]$, 令 $\mathscr{A}(f(x))=xf(x)$. 设 $\mathscr{D}$ 是微商映射. 证明: $\mathscr{A}$ 是线性的; $\mathscr{D}\mathscr{A}-\mathscr{A}\mathscr{D}$ 是单位映射.

解 线性 trivial。

$$\begin{aligned}(\mathscr{D}\mathscr{A}-\mathscr{A}\mathscr{D})(f(x)) &= \mathscr{D}(xf(x))-\mathscr{A}(f'(x)) \\ &= f(x)+xf'(x)-xf'(x) \\ &= f(x)\end{aligned}$$

作业 7.4 (P210 T4)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 m 维与 n 维线性空间. 取定 $\alpha\in U$. 所有满足 $\mathscr{A}(\alpha)=\mathbf{0}$ 的线性映射 $\mathscr{A}:U\to V$ 的集合记为 K. 证明: K 在线性映射的加法以及纯量与线性映射的乘法下成为数域 $\mathbb{F}$ 上的线性空间. 求 $\dim K$.

解 线性映射 trivial。注意讨论 α 情况!!

1. 若 $\alpha=0$，则 $K=L(U,V)$， $\dim K=mn$
2. 若 $\alpha\neq 0$，将 α 扩充为 U 的一组基，再取 V 的一组基，对于 $\mathscr{A}\in K$，有
$$\mathscr{A}(\alpha,\alpha_2,...,\alpha_m)=(\beta_1,...,\beta_n)A$$
注意 $\mathscr{A}(\alpha)=0$，所以
$$\mathscr{A}(\alpha,\alpha_2,...,\alpha_m)=(\beta_1,...,\beta_n)\begin{pmatrix}0 & A_{m\times(n-1)}\end{pmatrix}$$
注意该式为等价条件，且其中 $A_{m\times(n-1)}$ 每个元素都可任取，故 $\dim K=m(n-1)$

作业 7.5 (P210 T5)

设 $\mathscr{A}:V\to V$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性映射. 所有满足 $\mathscr{A}\mathscr{B}=\mathscr{O}$ 的线性映射 $\mathscr{B}:V\to V$ 的集合记为 R. 证明集合 R 在线性映射的加法以及纯量与线性映射的乘法下成为数域 F 上的线性空间. 选择适当的线性映射 $\mathscr{A}$, 使得 $\dim R=0$, 或 n, 或 n^2.

解 线性空间 trivial，但是一定别漏了。

1. 取 $\mathscr{A}=\mathscr{I}$，则 $R=0$， $\dim R=0$
2. 取 $\mathscr{A}(\alpha_1,\alpha_2,...,\alpha_n)=(0,\alpha_2,...,\alpha_n)$，其中 $(\alpha_1,\alpha_2,...,\alpha_n)$ 为空间一组基，则 $\mathscr{B}:V\to<\alpha_1>$，从而 $\mathscr{B}(\alpha_1,\alpha_2,...,\alpha_n)=(c_1\alpha_1,c_2\alpha_1...,c_n\alpha_1)$，可变动 $c_1,c_2,...,c_n\in\mathbb{F}$，故 $\dim R=n$
3. 取 $\mathscr{A}=0$，则 $R=L(U,U)$， $\dim R=n^2$

作业 7.6 (P219 T2)

设 $\mathscr{D}:\mathbb{F}_n[x]\to\mathbb{F}_n[x]$ 是微商映射. 求 $\rho(\mathscr{A})$ 与 $v(\mathscr{A})$. 等式 $\mathbb{F}_n[x]=\mathrm{Im}(\mathscr{D})\oplus\mathrm{Ker}(\mathscr{D})$ 是否成立?

解 这和 P204 T10 完全一样，记号沿用该题

$$\mathscr{A}(E_{11},E_{21},E_{12},E_{22})=(E_{11},E_{21},E_{12},E_{22})\begin{pmatrix}1 & -1 & 0 & 0\\ -4 & 4 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & -4 & 4\end{pmatrix}$$

记矩阵为 A，则 $\rho(\mathscr{A})=\mathrm{rank}(A)=2$

作业 7.7 (P219 T3)

设 $\mathscr{A}: V \to V$ 是数域上 n 维线性空间 V 到自身的线性映射, 且 $\rho\left(\mathscr{A}^2\right) = \rho(\mathscr{A})$. 证明: $\operatorname{Im}(\mathscr{A}) \cap \operatorname{Ker}(\mathscr{A}) = \{0\}$.

解 矩阵方法:

设 $\rho\left(\mathscr{A}^2\right) = \rho(\mathscr{A}) = r$, 取 $\operatorname{Ker}(\mathscr{A})$ 的基 $\{\alpha_{r+1}, \cdots, \alpha_n\}$, 它可以扩充为 U 的基 $\{\alpha_1, \cdots, \alpha_n\}$, 则

$$\mathscr{A}(\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} A_{11} & 0 \\ A_{21} & 0 \end{pmatrix},$$

$$\mathscr{A}^2(\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} A_{11}^2 & 0 \\ A_{11}A_{21} & 0 \end{pmatrix}.$$

由 $r = \rho\left(\mathscr{A}^2\right) = \operatorname{rank}\begin{pmatrix} A_{11}^2 & 0 \\ A_{11}A_{21} & 0 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A_{11} & 0 \\ A_{21} & 0 \end{pmatrix}\begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix} \leq \operatorname{rank}\begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix} = \operatorname{rank} A_{11} \leq r$ 可知 A_{11} 可逆. 若 $\alpha \in \operatorname{Im}(\mathscr{A}) \cap \operatorname{Ker}(\mathscr{A})$, 则由 $\alpha \in \operatorname{Im}\mathscr{A}$ 可知 $\alpha = \mathscr{A}(\beta)$, 且 $\mathscr{A}^2(\beta) = \mathscr{A}(\alpha) = 0$.

设 $\beta = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}$, 则

$$\alpha = \mathscr{A}(\beta) = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} A_{11}X_{11} \\ A_{21}X_{11} \end{pmatrix},$$

$$\mathscr{A}^2(\beta) = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} A_{11}^2X_{11} \\ A_{11}A_{21}X_{11} \end{pmatrix} = 0,$$

即 $A_{11}^2X_{11} = 0$. 由 A_{11} 可逆可知 $X_{11} = 0$, 从而

$$\alpha = A(\beta) = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)\begin{pmatrix} A_{11}X_{11} \\ A_{21}X_{11} \end{pmatrix} = 0.$$

因此 $\operatorname{Im}(\mathscr{A}) \cap \operatorname{Ker}(\mathscr{A}) = \{0\}$.

空间方法:

首先 $\operatorname{Ker}(\mathscr{A}) \subset \operatorname{Ker}(\mathscr{A}^2)$, 且由条件 $\dim Ker(\mathscr{A}^2) = \dim Ker(\mathscr{A})$, 从而 $\operatorname{Ker}(\mathscr{A}) = \operatorname{Ker}(\mathscr{A}^2)$。

设 $\alpha \in \operatorname{Im}(\mathscr{A}) \cap \operatorname{Ker}(\mathscr{A})$, $\alpha = \mathscr{A}(\beta)$, 则 $\mathscr{A}^2(\beta) = 0$, 即 $\beta \in \operatorname{Ker}(\mathscr{A}^2) = \operatorname{Ker}(\mathscr{A})$, 从而 $\alpha = \beta = 0$, 故结论成立

可以看到选择正确的思路对于做题很有帮助, 这题用矩阵做有点弱智。

作业 7.8 (P219 T7)

设 $\mathscr{A}: U \to U$ 是线性映射, $\rho(\mathscr{A}) = 1$. 证明: 存在唯一 $\lambda \in \mathbb{F}$, 使得 $\mathscr{A}^2 = \lambda\mathscr{A}$, 而且当 $\lambda \neq 1$ 时, $\mathscr{E} - \mathscr{A}$ 是可逆线性映射.

解 矩阵方法:

设 $\mathscr{A}$ 对应矩阵为 A, 其秩为 1, 故 $A = \alpha\beta^T$, 其中 $\alpha, \beta \in \mathbb{F}^n$。从而 $A^2 = (\alpha^T\beta)\alpha\beta^T$, 取 $\lambda = \alpha^T\beta$ 即可。唯一性 trivial。

又 $(I-A)*\frac{1}{1-\lambda}A=I$，故 $\mathscr{E}-\mathscr{A}$ 的逆映射为矩阵 $\frac{1}{1-\lambda}A$ 对应线性映射。

空间方法：

由 $\rho(\mathscr{A})=1$，存在 U 的一组基 $(\alpha_1,\alpha_2,...,\alpha_n)$，满足 $\mathscr{A}(\alpha_1,\alpha_2,...,\alpha_n)=(\beta,0,...,0)$，设 $\beta=\sum_{i=1}^n\lambda_i\alpha_i$，则 $\mathscr{A}^2(\alpha_1,\alpha_2,...,\alpha_n)=(\lambda_1\beta,0,...,0)$，取 $\lambda=\lambda_1$ 即可。

可逆线性映射 = 线性 + 单射 ($\mathrm{Ker}(\mathscr{A})=0$)+ 满射 ($\mathrm{Im}(\mathscr{A})=U$ 或 U 的基均在像集中)，线性，单射 trivial。对于任意 $\alpha_i,i>1$，有 $(\mathscr{E}-\mathscr{A})(\alpha_i)=\alpha_i$，又 $(\mathscr{E}-\mathscr{A})(\alpha_1+\frac{1}{1-\lambda}\beta)=\alpha_1$，故满射成立

作业 7.9 (P219 T8)

设 $V_0,V_1,\cdots,V_{n+1}$ 是数域 $\mathbb{F}$ 上有限维线性空间, $V_0=V_{n+1}=\{0\}$. 设 $\mathscr{A}_i:V_i\to V_{i+1}$ 是线性映射, $i=0,1,\cdots,n$, 且 $\mathrm{Ker}\left(\mathscr{A}_{i+1}\right)=\mathrm{Im}\left(\mathscr{A}_i\right),i=0,1,\cdots,n-1$. 证明:

$$\sum_{i=1}^n(-1)^i\dim V_i=0$$

解 $\dim V_i=\dim Ker(\mathscr{A}_i)+\dim Im(\mathscr{A}_i)=\dim Im(\mathscr{A}_i)+\dim Im(\mathscr{A}_{i-1})$ 从而 $\sum_{i=1}^n(-1)^i\dim V_i=\dim Im(\mathscr{A}_0)+(-1)^n\dim Im(\mathscr{A}_n)=0$。这是因为 $\mathrm{Im}(\mathscr{A})\subset V_0=0$，$\mathrm{Im}(\mathscr{A})\subset V_{n+1}=0$

作业 7.10 (补充题目)

设 n 维线性空间 V 的基 $\alpha_1,\alpha_2,...,\alpha_n$ 到基 $\beta_1,\beta_2,...,\beta_n$ 的过度矩阵为 P，并且 $\alpha_1^*,...,\alpha_n^*$ 和 $\beta_1^*,...,\beta_n^*$ 分别为两组基的对偶基。求 V 的对偶空间 $V*$ 的基 $\alpha_1^*,...,\alpha_n^*$ 到 $\beta_1^*,...,\beta_n^*$ 的过渡矩阵

解 利用 $\alpha_i^*(\alpha_j)=\delta_{ij}$，注意

$$\begin{aligned}\begin{pmatrix}\alpha_1^*\\\alpha_2^*\\\vdots\\\alpha_n^*\end{pmatrix}\begin{pmatrix}\beta_1&\beta_2&\cdots&\beta_n\end{pmatrix}&=\begin{pmatrix}\alpha_1^*\\\alpha_2^*\\\vdots\\\alpha_n^*\end{pmatrix}\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_n\end{pmatrix}P\\&=(\alpha_i^*(\alpha_j))_{n\times n}P\\&=P\end{aligned}$$

且有

$$\begin{aligned}\begin{pmatrix}\beta_1^*\\\beta_2^*\\\vdots\\\beta_n^*\end{pmatrix}\begin{pmatrix}\beta_1&\beta_2&\cdots&\beta_n\end{pmatrix}&=(\beta_i^*(\beta_j))_{n\times n}\\&=I\end{aligned}$$

故

$$(\beta_1^*,...,\beta_n^*)=(\alpha_1^*,...,\alpha_n^*)(P^{-1})^T$$

7.1.2 第十三周

作业 7.11 (P223 T3)

设数域 $\mathbb{F}$ 上 n 维线性空间 V 是子空间 U 与 W 的直和, 则对任意 $\alpha \in V$, 存在唯一一对向量 β 与 $\gamma, \beta \in U, \gamma \in W$, 使得 $\alpha = \beta + \gamma$. 定义映射 $\mathscr{A}: V \to V$ 如下: 设 $\alpha \in V$, 则令 $\mathscr{A}(\alpha) = \beta$. 映射 $\mathscr{A}$ 称为 V 沿子空间 W 在 U 上的投影变换. 证明:

(1) 投影变换 $\mathscr{A}$ 是线性变换;

(2) 线性变换 $\mathscr{B}: V \to V$ 为投影变换当且仅当 $\mathscr{B}$ 为幂等变换, 即 $\mathscr{B}$ 满足 $\mathscr{B}^2 = \mathscr{B}$;

(3) 线性变换 $\mathscr{B}: V \to V$ 为投影变换当且仅当 $\mathscr{I} - \mathscr{B}$ 为投影变换, 其中 $\mathscr{I}$ 单位映射.

解

(1) 只需证 $\forall \lambda, \mu \in \mathbb{F}, \forall \alpha_1, \alpha_2 \in V$ 有 $\mathscr{A}(\lambda\alpha_1 + \mu\alpha_2) = \lambda\mathscr{A}(\alpha_1) + \mu\mathscr{A}(\alpha_2)$

(2) 投影变换 $\to$ 幂等变换是 trivial 的，下说明幂等变换 $\mathscr{B}$ 为沿 $\ker\mathscr{B}$ 在 $\Im\mathscr{B}$ 上的投影变换。事实上，由幂等可推 $\mathrm{Ker}(\mathscr{B}) \cap \mathrm{Im}(\mathscr{B}) = 0$，从而 $V = \mathrm{Ker}(\mathscr{B}) \oplus \mathrm{Im}(\mathscr{B}) = 0$。故对任意 $\alpha \in V, \exists! \beta \in \mathrm{Ker}(\mathscr{B}), \gamma \in \mathrm{Im}(\mathscr{B}), \alpha = \beta + \gamma$，故 $\mathscr{B}(\alpha) = \mathscr{B}(\beta) = \beta$

(3) 用第二问验证即可

作业 7.12 (P223 T6)

设 $\mathscr{A}: V \to V$ 是线性变换. 证明: 存在 $f(x) \in \mathbb{F}[x]$, 使得 $f(\mathscr{A}) = 0$.

解

法一：取线性变换对应的矩阵 A($\mathscr{A}(\alpha_1, \alpha_2, ..., \alpha_n) = (\alpha_1, \alpha_2, ..., \alpha_n)A$，及其零化多项式 $\varphi_A(\lambda)$，则 $\varphi_A(\mathscr{A})(\alpha_1, \alpha_2, ..., \alpha_n) = (\alpha_1, \alpha_2, ..., \alpha_n)\varphi_A(A) = 0$，即可得到结论。

法二：仍旧取线性变换对应矩阵 A，A 属于 $M_n(\mathbb{R})$，而 n 阶矩阵空间维度为 n^2，故 $I, A, A^2, ..., A^{n^2}$ 一定线性相关，设 $\lambda_0 I + \lambda_1 A + ... + \lambda_{n^2} A^{n^2} = 0$，即得题目要求多项式。

作业 7.13 (P223 T8)

设 $\mathscr{A}: V \to V$ 与 $\mathscr{B}: V \to V$ 是幕等变换, 即 $\mathscr{A}^2 = \mathscr{A}, \mathscr{B}^2 = \mathscr{B}$. 证明:

(1) $\mathrm{Im}(\mathscr{A}) = \mathrm{Im}(\mathscr{B})$ 的充分必要条件是 $\mathscr{A}\mathscr{B} = \mathscr{B}, \mathscr{B}\mathscr{A} = \mathscr{A}$;

(2) $\mathrm{Ker}(\mathscr{A}) = \mathrm{Ker}(\mathscr{B})$ 的充分必要条件是 $\mathscr{A}\mathscr{B} = \mathscr{A}, \mathscr{B}\mathscr{A} = \mathscr{B}$.

解 (1) 只需证 $\mathscr{A}\mathscr{B} = \mathscr{B} \Leftrightarrow \mathrm{Im}(\mathscr{B}) \subset \mathrm{Im}(\mathscr{A})$。事实上，LHS 推 RHS 显然，若 $\mathrm{Im}(\mathscr{B}) \subset \mathrm{Im}(\mathscr{A})$，则对任意 $x \in V$，$\mathscr{B}(x) \in \mathrm{Im}(\mathscr{B}) \subset \mathrm{Im}(\mathscr{A})$，又由 T3 可知，$\mathscr{A}(\mathrm{Im}(\mathscr{A})) = Im(\mathscr{A})$，故 $\mathscr{A}\mathscr{B}(x) = \mathscr{B}(x)$，即为 LHS

(2) 由 T3，$\mathscr{I} - \mathscr{A}$ 也是幂等变换，并且 $\mathrm{Im}(\mathscr{I} - \mathscr{A}) = \mathrm{Ker}(\mathscr{A})$，对 $\mathscr{I} - \mathscr{A}$ 使用第二问结论即可

作业 7.14 (P223 T11)

设 $A \in \mathbb{F}^{n\times n}$ 满足 $A^2 = 0$. 证明: 方阵 A 相似于方阵 $\begin{pmatrix} 0 & A_{r\times(n-r)} \\ 0 & 0 \end{pmatrix}$.

解

设矩阵 A 对应的线性变换 $\mathscr{A}$，设 $\mathrm{Im}(\mathscr{A})$ 的一组基为 $\alpha_1, \alpha_2, ..., \alpha_r$，其中 r 为 A 的秩，将这组基扩

充成 n 维空间的一组基 $\alpha_1, \alpha_2, ..., \alpha_n$，直接不妨这组基就是 $\mathscr{A}$ 和 A 之间的纽带，桥梁，爱的红线。

故 $\mathscr{A}(\alpha_1, ..., \alpha_n) = (0, 0, ..., \mathscr{A}(\alpha_{r+1})), ..., \mathscr{A}(\alpha_n))$, 注意后 $n-r$ 项属于 $\mathrm{Im}(\mathscr{A})$, 可以被 $\{\alpha_1, ..., \alpha_r\}$ 表示，且前 r 项属于 $\mathrm{Im}(\mathscr{A}) = 0$，故最后有

$$\mathscr{A}(\alpha_1, \alpha_2, ..., \alpha_n) = (\alpha_1, \alpha_2, ..., \alpha_n)\begin{pmatrix} 0 & A_(r \times (n-r)) \\ 0 & 0 \end{pmatrix}$$

从而 $A = \begin{pmatrix} 0 & A_(r \times (n-r)) \\ 0 & 0 \end{pmatrix}$

作业 7.15 (P185 T10)

分别求下列向量 $\alpha_1, \alpha_2, \alpha_3$ 与向量 $\beta_1, \beta_2, \beta_3$ 生成的子空间 W_1 与 W_2 的维数, 并给出子空间 $W_1 \cap W_2$ 与 $W_1 + W_2$ 的一组基:

(1) $\alpha_1 = (1, 2, 1, -2), \alpha_2 = (2, 3, 1, 0), \alpha_3 = (1, 2, 2, -3)$,
$\beta_1 = (1, 1, 1, 1), \beta_2 = (1, 0, 1, -1), \beta_3 = (1, 3, 0, -4)$;
(2) $\alpha_1 = (1, 1, 0, 0), \alpha_2 = (0, 1, 1, 0), \alpha_3 = (0, 0, 1, 1)$,
$\beta_1 = (1, 0, 1, 0), \beta_2 = (0, 2, 1, 1), \beta_3 = (1, 2, 1, 2)$.

♡

解

(1) $\dim W_1 = \dim W_2 = 3$.

设 $\alpha = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = y_1\beta_1 + y_2\beta_2 + y_3\beta_3 \in W_1 \cap W_2$. 由 $x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 - y_1\beta_1 - y_2\beta_2 - y_3\beta_3 = 0$ 得 $(x_1, x_2, x_3, y_1, y_2, y_3) = t_1(5, -1, -2, 0, 0, 1) + t_2(-2, 1, 1, 1, 0, 0)$, 从而 $\alpha = -t_1\beta_3 - t_2\beta_1$, 即 $\{\beta_1, \beta_3\}$ 为 $W_1 \cap W_2$ 的一组基. 由 $\dim(W_1 + W_2) = 4$ 可知 $W_1 + W_2 = \mathbb{F}^4$, 可取 $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$ 作为一组基.

(2) $\dim W_1 = \dim W_2 = 3$.

设 $\alpha = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = y_1\beta_1 + y_2\beta_2 + y_3\beta_3 \in W_1 \cap W_2$. 由 $x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 - y_1\beta_1 - y_2\beta_2 - y_3\beta_3 = 0$ 得 $(x_1, x_2, x_3, y_1, y_2, y_3) = t_1(2, 0, 2, 1, 0, 1) + t_2(1, 1, 1, 1, 1, 0)$, 从而 $\alpha = t_1(\beta_1 + \beta_3) + t_2(\beta_1 + \beta_2)$, 即 $\{\beta_1 + \beta_2, \beta_1 + \beta_3\}$ 为 $W_1 \cap W_2$ 的一组基. 由 $\dim(W_1 + W_2) = 4$ 可知 $W_1 + W_2 = \mathbb{F}^4$, 可取 $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$ 作为一组基.

作业 7.16 (P191 T3)

设 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 都是四维复向量空间 $\mathbb{C}^4$ 中的向量, $\mathbb{C}^4$ 中分别由向量集合 $\{\alpha_1, \alpha_2\}$ 和 $\{\alpha_3, \alpha_4\}$ 生成的子空间记为 V 和 W. 试判断 $\mathbb{C}^4 = V \oplus W$ 是否成立? (1) $\alpha_1 = (0, 1, 0, 1), \alpha_2 = (0, 0, 1, 0), \alpha_3 = (1, 0, 1, 0), \alpha_4 = (1, 1, 0, 0)$;
(2) $\alpha_1 = (-1, 1, 1, 0), \alpha_2 = (0, 1, -1, 1), \alpha_3 = (1, 0, 0, 0), \alpha_4 = (0, 0, 0, 1)$;
(3) $\alpha_1 = (1, 0, 0, 1), \alpha_2 = (0, 1, 1, 0), \alpha_3 = (1, 0, 1, 0), \alpha_4 = (0, 1, 0, 1)$.

♡

解 直接把四个向量拼成矩阵，验证矩阵是否满秩即可

(1)(2) 成立，(3) 不成立

7.2 补充内容

7.2.1 关于矩阵和空间

线性代数 = 矩阵 + 空间，我理解为存在两套理论体系，实际上两者是并存密不可分的，但是做题时可以采用不同的理论体系去进行思考。下面展示一些题目，这些题目均可以使用两套不同体系得到结果。在实操时使用自己喜欢的那套就好了。

例题 7.1 习题 5.4.3 设 $\mathscr{A}: V \to V$ 是数域上 n 维线性空间 V 到自身的线性映射, 且 $\rho\left(\mathscr{A}^2\right) = \rho(\mathscr{A})$. 证明: $\operatorname{Im}(\mathscr{A}) \cap \operatorname{Ker}(\mathscr{A}) = \{0\}$.

解 解答见第十二周习题讲解

例题 7.2 习题 5.5.11 设 $A \in \mathbb{F}^{n\times n}$ 满足 $A^2 = 0$. 证明: 方阵 A 相似于方阵 $\begin{pmatrix} 0 & A_{r\times(n-r)} \\ 0 & 0 \end{pmatrix}$.

解 解答见第十三周习题讲解。

7.2.2 补充题

这两道题是活用空间 Ker 和 Im 技巧性十足的题，比较有难度。讲义中只提供大致的做法，具体的会在习题课上面讲

定理 7.17

$A_1, A_2, ..., A_m$ 是 n 阶方阵，满足 $A_1 + ... + A_m = I_n$，则一下命题等价：

(1) $A_i^2 = A_i$

(2) $\operatorname{rank}(A_1) + ... + \operatorname{rank}(A_m) = n$

(3) $\forall i \neq j, \quad A_iA_j = 0$

解

(1) $\to$ (2): $\operatorname{rank}(A_i) = tr(A_i)$

(2) $\to$ (3): 由 (2)

$$\dim(\operatorname{Ker}(A_k)) = \sum_{i\neq k} dim(\operatorname{Im}(A_i))$$

而右边包含左边，从而两边相等，即

$$\operatorname{Ker}(A_k) = \sum_{i\neq k} Im(A_i)$$

故

$$A_k(\operatorname{Ker}(A_k)) = \sum_{i\neq k} A_k(\operatorname{Im}(A_i)) = 0$$

即对 $\forall x_i, i \neq k$，有 $\sum_{i\neq k} A_kA_ix_i = 0$，由 x_i 任意性，可以得到 $A_kA_i = 0, \forall i \neq k$

(3) $\to$ (1): trivial

定理 7.18

若 $\operatorname{rank}(A) = \operatorname{rank}(A^2)$， $AB = BA$，则 $\operatorname{rank}(A+B) = \operatorname{rank}(A) + \operatorname{rank}(B) \Leftrightarrow AB = 0$

解

$\text{rank}(A) = \text{rank}(A^2)$ 可以推出 $\text{Ker}(A) = \text{Ker}(A^2)$，进而 $V = \text{Ker}(A) \oplus \text{Im}(A)$。

$LHS \to RHS$: 此时

$$\dim((A+B)Im(A)) \leq \dim(\text{Im}(A)) = \text{rank}(A)$$

$$\dim((A+B)Ker(A)) \leq \dim(\text{Im}(B)) = \text{rank}(B)$$

从而

$$\text{rank}(A+B) = dim(\text{Im}(A+B)) \leq \dim((A+B)Im(A)) + \dim((A+B)Ker(A)) \leq \text{rank}(A) + \text{rank}(B)$$

故该式不等号要取等，即 $(A+B)Ker(A) = Im(B)$，即 $B(\text{Ker}(A)) = Im(B)$。最后

$$AB = 0 \Leftrightarrow A(\text{Im}(B)) \Leftrightarrow AB(\text{Ker}(A)) = 0 \Leftrightarrow BA(\text{Ker}(A)) = 0$$

结论成立

$RHS \to LHS$: $AB = 0$ 可以推出 $\text{Im}(B) \subset \text{Ker}(A)$，从而 $\text{Im}(B) \cap \text{Im}(A) = 0$，又 $BA = AB = 0$，从而 $\text{Im}(A) \subset \text{Ker}(B)$，故

$$(A+B)Im(A) = Im(A)$$

因为 $\text{Im}(A+B) \subset \text{Im}(A) + \text{Im}(B)$，而对 $Ax + By$，$Ax + By = (A+B)y + A(x-y) \subset \text{Im}(A+B)$，从而 $\text{Im}(A+B) = Im(A) + \text{Im}(B)$。故 $\text{rank}(A+B) = \text{rank}(A) + \text{rank}(B)$

第 8 章 循环子空间

8.1 习题讲解

8.1.1 第十四周

作业 8.1 (P229 T3)

设 $\mathscr{A}: V \to V$ 是线性变换, $0 \neq \alpha \in V$, 证明: V 中由向量 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^k(\alpha), \cdots$ 生成的子空间 U 是 $\mathscr{A}$ 的不变子空间. 设 $\dim U = r$, 证明: $\{\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha)\}$ 是 U 的基. 求 $\mathscr{A}|_U$ 在这组基下的方阵.

解

设 k 为使得 $\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^{k-1}(\alpha)$ 线性无关的最大值，则 $\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^k(\alpha)$ 线性相关，从而 $\mathscr{A}^k(\alpha) \in \mathrm{Span}\{\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^{k-1}(\alpha)\}$,

设 $\mathscr{A}^k(\alpha) = \lambda_0\alpha + \lambda_1\mathscr{A}(\alpha) + \cdots + \lambda_{k-1}\mathscr{A}^{k-1}(\alpha)$,

则 $\mathscr{A}^k(\alpha) \in \mathrm{Span}\{\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^{k-1}(\alpha)\}$，依次类推，得 $U = \mathrm{Span}\{\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^{k-1}(\alpha)\}$，从而 $k = r$

$$\mathscr{A}|_U \text{ 对应方阵 } \begin{pmatrix} 0 & & & & \lambda_0 \\ 1 & 0 & & & \lambda_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & \vdots \\ & & & 1 & \lambda_{r-1} \end{pmatrix}$$

作业 8.2 (P229 T5)

V 是 n 维复线性空间, $\mathscr{A}: V \to V$ 是线性变换. 证明: V 的每个子空间都是 $\mathscr{A}$ 的不变子空间的充分必要条件是 $\mathscr{A}$ 为纯量变换, 即 $\mathscr{A} = \lambda\mathscr{I}$, 其中 $\lambda \in \mathbb{C}$。

解

$(\Leftarrow)$trivial, 对于 $(\Rightarrow)$, 考虑子空间 <x>,$\forall x \in V$, 可得 $\mathscr{A}(x) = \lambda_x x$, 对另一个 $y \neq x$, 有 $\mathscr{A}(x+y) = \lambda_{x+y}(x+y) = \lambda_x x + \lambda_y y$，由任意性可得 $\lambda_{x+y} = \lambda_x = \lambda_y$，故存在统一的 λ，使得 $\mathscr{A}(x) = \lambda x$

作业 8.3 (P229 T8)

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\mathbb{F}$ 中两两不等的 n 个数, 线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的方阵为 $\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$. 求 $\mathscr{A}$ 的不变子空间的个数.

解

首先任取 1 到 n 中 k 个数 $i_1, \ldots, i_k$，有 $\mathrm{Span}\{\alpha_{i_1}, a_{i_2}, \ldots, \alpha_{i_k}\}$ 为 V 的不变子空间，从而得到 2^n 个不变子空间，下证只有这些。

取 $U \subset V$ 为不变子空间，不妨仅 $\alpha_1, \alpha_2, \ldots, \alpha_r \in U$，且 $\dim U = r + s$，$s = 0$ 则结论成立，若 $s \neq 0$，则设 $U = \mathrm{Span}\{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s\}$，不妨设 $\beta_1 = \mu_{r+1}\alpha_{r+1}, \ldots, \mu_n\alpha_n$，且系数均非 0。

则对任意 k，$\mathscr{A}^k(\beta_1)=\lambda_{r+1}^k\mu_{r+1}\alpha_{r+1}+\cdots+\lambda_n^k\mu_n\alpha_n$，从而

$$(\beta_1,\mathscr{A}(\beta_1),\ldots,\mathscr{A}^{n-r-1}(\beta_1))=(\alpha_{r+1},\ldots,\alpha_n)\operatorname{diag}(\mu_{r+1},\ldots,\mu_n)\begin{pmatrix}1&\lambda_{r+1}&\cdots&\lambda_{r+1}^{n-r-1}\\1&\lambda_{r+2}&\cdots&\lambda_{r+2}^{n-r-1}\\\vdots&\vdots&\cdots&\vdots\\1&\lambda_n&\cdots&\lambda_n^{n-r-1}\end{pmatrix}$$

所以 $\beta_1,\mathscr{A}(\beta_1),\ldots,\mathscr{A}_{n-r-1}(\beta_1)$ 线性无关且均在 U 中，这 n-r 个元素可以和 $\alpha_1,\ldots,\alpha_r$ 组成 U 中一组线性无关集合，从而 $\dim U=n$，即 $U=V$，结论也成立

作业 8.4 (P239 T7)

设 A 与 B 为 n 阶复方阵, 且方阵 A 的特征多项式为 $\varphi(\lambda)$. 证明: 方阵 $\varphi(B)$ 可逆的充分必要条件是, 方阵 A 与 B 没有公共特征值.

解

设 B 的 Jordan 标准形为 J，设 $\varphi(J)=(\lambda_1I-J)^{n_1}\ldots(\lambda_kI-J)^{n_k}$，则 J 不以 $\lambda_1,\ldots,\lambda_k$ 为特征值时，λ_iI-J 可逆，从而 $\varphi(J)$ 可逆。

当 $\exists i,\det(\lambda_iI-j)=0$ 时，则 λ_iI-J 对角元有 0，不可逆，结论成立

作业 8.5 (P239 T8)

设 A 与 B 为 n 阶复方阵, 则关于未知方阵 X 的方阵方程 $AX=XB$ 只有零解的充分必要条件是, 方阵 A 与 B 没有公共特征值.

解

($\Leftarrow$): $AX=XB\Rightarrow\varphi(A)X=X\varphi(B)$，从而 $X\varphi(B)=0$，其中 $\varphi(\lambda)$ 为 A 特征多项式。而由 T7，$\varphi(B)$ 可逆，从而 $AX=XB$ 只有零解

($\Rightarrow$): 若 A，B 有公共特征值，设为 λ，则设 $Ax=\lambda x,y^TB=\lambda y^T,x,y\neq0$（注意 B 和 B 转置特征值相同），从而 $Axy^T=\lambda xy^T=xy^TB$，方程有非零解

作业 8.6 (P239 T11)

证明: 准对角方阵的最小多项式等于每个对角块的最小多项式的最小公倍式.

解

设 $A=\operatorname{diag}(A_1,\ldots,A_k)$，$A_i$ 最小多项式为 $d_i(\lambda)$，A 最小多项式为 $d_A(\lambda)$。注意

$$\varphi(A)=\operatorname{diag}(\varphi(A_1),\ldots,\varphi(A_k))=0$$

等价于

$$\varphi(A_i)=0,\forall i$$

这等价于 $d_i|\varphi,\forall i$，从而 $\operatorname{lcm}(d_1,\ldots,d_k)|d_A$。

注意到 $\operatorname{lcm}(d_1,\ldots,d_k)$ 为零化多项式，从而 $d_A|\operatorname{lcm}(d_1,\ldots,d_k)$，故 $d_A=\operatorname{lcm}(d_1,\ldots,d_k)$

作业 8.7 (P239 T13(1))

求以下方阵的特征多项式与最小多项式:

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ a_0 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}_{n\times n}$$

解

特征多项式: $\lambda^n - a_{n-1}\lambda^{n-1} - \cdots - a_0$

最小多项式: 由 $D_{n-1}(\lambda) = 1 \Rightarrow d(\lambda) = \varphi(\lambda)$

作业 8.8 (P239 T16)

取方阵 $A \in \mathbb{C}^{n\times n}$. 定义映射 $\mathscr{A} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ 如下: 设 $X \in \mathbb{C}^{n\times n}$, 则令 $\mathscr{A}(X) = AX$. 证明: 线性变换 $\mathscr{A}$ 的最小多项式等于方阵 $\mathscr{A}$ 的最小多项式.

解 注意 $f(\mathscr{A})(X) = f(A)(X)$, 故 $f(\mathscr{A}) = 0 \Leftrightarrow f(A) = 0$, 即 $d_{\mathscr{A}}(\lambda) = d_A(\lambda)$

作业 8.9 (P247 T3)

由于方阵 A 的迹 $\mathrm{tr}(A)$ 是方阵在相似下的不变量, 因此可以定义线性变换 $\mathscr{A} : V \to V$ 在 V 的某组基下的方阵 A 的 $\mathrm{tr}(A)$ 为线性变换 $\mathscr{A}$ 的迹 $\mathrm{tr}(\mathscr{A})$. 证明: 如果 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 满足 $\mathrm{tr}(\mathscr{A}) = 0$, 则存在 V 的一组基, 使得线性变换 $\mathscr{A}$ 在这组基下的方阵的主对角元都是零.

解

只需证明 $\forall A \in F^{n\times n}$ 满足 $\mathrm{tr}(A) = 0$, 存在与 A 相似的矩阵 B, 其对角元全为 0。用归纳法做:

首先 n=1 时结论显然成立, 假设 n-1 时成立, 考虑 n 时情况, 设 $A = \begin{pmatrix} A_1 & \beta \\ T & a_{nn} \end{pmatrix}$

1. Case 1: 若 α, β 不全为 0, 则不妨 $a_{1n} > 0$, 设 $\gamma = (-\frac{a_{nn}}{a_{1n}}, 0, \ldots, 0)^T$, 再由归纳假设, 设 $PA_1P^{-1} = B$ 对角元全 0, 得
$$\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I_{n-1} & 0 \\ \gamma^T & 1 \end{pmatrix}\begin{pmatrix} A_1 & \beta \\ \alpha & a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & 0 \\ -\gamma^T & 1 \end{pmatrix}\begin{pmatrix} P^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} B & \beta' \\ \alpha'^T & 0 \end{pmatrix}$$
右侧矩阵对角元全为 0, 与 A 相似
2. Case 2: 若 $\alpha = \beta = 0$, 若 $a_{nn} = 0$trivial, 若非 0, 可以类似的取 γ 将 A_1 中的非零元调换到 α, β 中, 归为 Case 1

作业 8.10 (P249 T6)

设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 可对角化, 且 U 是线性变换 $\mathscr{A}$ 的不变子空间. 证明: 线性变换 $\mathscr{A}$ 在 U 上的限制 $\mathscr{A}|_U$ 也是可对角化的.

解

设 U 的基为 $\{\alpha_1, \ldots, \alpha_r\}$, 扩充为 V 的基 $\{\alpha_1, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}$, 则

$$\mathscr{A}(\alpha_1, \ldots, \alpha_n) = (\alpha_1, \ldots, \alpha_n)\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

对 A_{11} 的特征值 λ_0，有

$$\lambda_0 I - A = \begin{pmatrix} \lambda_0 I - A_{11} & -A_{12} \\ 0 & \lambda_0 I - A_{22} \end{pmatrix}$$

设 λ_0 在 A 中的代数重数为 n，几何重数为 m，在 A_{ii} 中的代数重数为 n_i，几何重数为 m_i，故由 A 可以对角化，$m=n$，且由代数重数定义，$n=n_1+n_2$，从而

$$\begin{aligned} n = \dim \operatorname{Ker}(\lambda_0 I - A) &= n - \operatorname{rank}(\lambda_0 I - A) \\ &\leq n - \operatorname{rank}(\lambda_0 I - A_{11}) - \operatorname{rank}(\lambda_0 I - A_{22}) \\ &= \dim \operatorname{Ker}(\lambda_0 I - A_{11}) + \dim \operatorname{Ker}(\lambda_0 I - A_{22}) \\ &= m_1 + m_2 \\ &\leq n_1 + n_2 = n \end{aligned}$$

从而不等式取等，即 $n_1 = m_1, n_2 = m_2$，即 $\mathscr{A}|_U$ 可以对角化

8.1.2 第十五周

作业 8.11 (P240 T17)

设 $\mathscr{A}: V \to V$ 是线性变换，且 $U_1, U_2, \ldots, U_k$ 是线性变换 $\mathscr{A}$ 的不变子空间，$V = U_1 \oplus U_2 \oplus \cdots \oplus U_k$，证明：线性变换 $\mathscr{A}$ 的最小多项式等于 $\mathscr{A}|_{U_1}, \ldots, \mathscr{A}|_{U_k}$ 的最小多项式的最小公倍式

解

将 $U_1, \ldots, U_k$ 的基并成 V 的一组基（因为直和），可以得到

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \ldots, \alpha_n) \operatorname{diag}(A_1, A_2, \ldots, A_k)$$

由 P239T11 和 P239T16 结论即证

作业 8.12 (P259 T5)

证明: 3 阶复方阵 A 与 B 相似的充分必要条件是, 方阵 A 与 B 具有相同的特征多项式与最小多项式.

解

($\Leftarrow$) 显然

($\Rightarrow$): 本质上为证明特征多项式和最小多项式能唯一确定一个 Jordan 标准形，本题使用分类讨论阐述，统一设 $\varphi(\lambda), d(\lambda)$ 为 A，B 的特征多项式和最小多项式

1. $\varphi(\lambda) = (\lambda - \lambda_0)^3$，此时 $d(\lambda) = \lambda - \lambda_0$，则 $J = \operatorname{diag}(J_1(\lambda_0), J_1(\lambda_0), J_1(\lambda_0))$，若 $d(\lambda) = (\lambda - \lambda_0)^2$，则 $J = \operatorname{diag}(J_2(\lambda_0), J_1(\lambda_0))$，$d(\lambda) = (\lambda - \lambda_0)^3$ 情况类似
2. $\varphi(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2), \lambda_1 \neq \lambda_2$，则 $d(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$，此时 $J = \operatorname{diag}(\lambda_1, \lambda_1, \lambda_2)$。若 $d(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2)$，此时 $J = \operatorname{diag}(J_2(\lambda_1), \lambda_2)$。
3. $\varphi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$，此时 $d(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$，$J = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$。

作业 8.13 (P259 T7)

设 $\mathscr{A}$ 是属于 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换. 证明: 存在线性变换 $\mathscr{A}$ 的不变子空间 V_1 和 V_2, 使得 $V = V_1 \oplus V_2$, 并且线性变换 $\mathscr{A}$ 在 V_1 上的限制 $\mathscr{A} \mid V_1$ 是可逆的, 而在 V_2 上的限制 $\mathscr{A}|_{V_2}$ 是幂零的. 简单地说, 任意线性变换 $\mathscr{A}$ 都可以分解为可逆线性变换 $\mathscr{A}|_{V_1}$ 与幂零线性变换 $\mathscr{A}|_{V_2}$ 的直和. (注意, 如果数域 $\mathbb{F}$ 是复数域, 则本题可用第一分解定理加予证明. 这里要求给出一个不用空间第一分解定理的证明.) ♡

解

取 $V_1 = \operatorname{Im}\mathscr{A}^n, V_2 = \operatorname{Ker}\mathscr{A}^n$，注意有性质 $\operatorname{Ker}\mathscr{A}^n = \operatorname{Ker}\mathscr{A}^{n+1} = \cdots = \operatorname{Ker}\mathscr{A}^{2n} = \ldots$

首先，对任意 $\alpha \in V_1 \cap V_2$，设 $\alpha = \mathscr{A}^n\beta$，则 $\mathscr{A}^{2n}\beta = \mathscr{A}^n\alpha = 0$，即 $\beta \in \operatorname{Ker}\mathscr{A}^{2n} = \operatorname{Ker}\mathscr{A}^n$，故 $\alpha = \mathscr{A}^n\beta = 0$，故 $V_1 \cap V_2 = \{0\}$，故 $V = V_1 \oplus V_2$

其次，对任意 $x \in V_2, \mathscr{A}^n x = 0$，故 $(\mathscr{A}|_{V_2})^n = 0$，即幂零

最后，$\mathscr{A}|_{V_1}(V_1) = \operatorname{Im}(\mathscr{A}^{n+1}) = \operatorname{Im}\mathscr{A}^n$，故映射为满射。且对任意 $x \in \operatorname{Ker}(\mathscr{A}|_{V_1})$，设 $x = \mathscr{A}^n y$，则 $\mathscr{A}^{n+1}y = 0 \Rightarrow \mathscr{A}^n y = x = 0$，从而映射为单射，也是双射，故可逆

作业 8.14 (P259 T9)

设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的最小多项式为 $d(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t}$ 其中 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是线性变换 $\mathscr{A}$ 的全部不同特征值. 并设 W_j 是线性变换 $(\mathscr{A}-\lambda_i\mathscr{I})^{m_j}$ 的核, $j = 1, 2, \cdots, t$. 证明:

(1) 根子空间 $W_{\lambda_j} = W_j, j = 1, 2, \cdots, t$;

(2) 线性变换 $\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}|_{W_j}$ 的最小多项式为 $(\lambda-\lambda_j)^{m_j}, j = 1, 2, \cdots, t$. ♡

解

这个题目告诉我们 $W_\lambda = \operatorname{Ker}(\mathscr{A}-\lambda_0 I)^k$ 中的 k 如何求

(1) 设 $W_{\lambda_j} = W_{\lambda_j}^{(n_j)}$, 则 $\mathscr{A}|_{w_{\lambda_j}}$ 最小多项式为 $(\lambda-\lambda_j)^{n_j}$, 由第一分解定理 $V = W_{\lambda_1} \oplus W_{\lambda_2} \cdots \oplus W_{\lambda_t}$, 由 P240T17，$d(\lambda) = (\lambda-\lambda_1)^{n_1}\ldots(\lambda-\lambda_t)^{n_t}$ 从而 $n_j = m_j$，即 $W_{\lambda_j} = W_j$

(2) 由 $\mathscr{A}|_{W_j} = \mathscr{A}|_{W_{\lambda_j}}$ 以及 $n_j = m_j$ 即可得到结论

作业 8.15 (P268 T4)

设 $\mathscr{A}$ 是 n 维复线性空间 V 的 k 次幂零变换, 且 V 分解为循环子空间 $C_1, C_2, \cdots, C_k$ 的直和. $C_1, C_2, \cdots, C_k$ 中维数为 j 的子空间的个数记为 $n_j, j = 1, 2, \cdots, k$. 证明:

(1) $\dim(\operatorname{Im}(\mathscr{A})) = n_2 + 2n_3 + \cdots + (k-1)n_k$;

(2) $n_j = \dim\left(\operatorname{Im}\left(\mathscr{A}^{j+1}\right)\right) + \dim\left(\operatorname{Im}\left(\mathscr{A}^{j-1}\right)\right) - 2\dim\left(\operatorname{Im}\left(\mathscr{A}^{j}\right)\right), j = 1, 2, \cdots, k$;

(3) V 本身是非零向量 α 生成的循环子空间的充分必要条件是, $\dim\left(\operatorname{Im}\left(\mathscr{A}^{j}\right)\right) = n-j, j = 1, 2, \cdots, n$. ♡

解

(1) $V = C_1 \oplus C_2 \cdots \oplus C_k$, 由 C_j 的基并起来构成 V 的基, 可得 $\mathscr{A}$ 对应的方阵为 $\operatorname{diag}(N_{m_1}^T, \ldots, N_{m_k}^T)$, 有

$$\dim\operatorname{Im}\mathscr{A} = \operatorname{rank}\mathscr{A} = n-k = n_2 + n_3 + \cdots + (k-1)n_k$$

(2) 研究 $\operatorname{diag}(N_{m_1}^T,\dots,N_{m_k}^T)$，注意 $\operatorname{rank} N_{m_i}^j=\max\{m_i-j,0\}$，故

$$\dim\operatorname{Im}\mathscr{A}^j=\sum_{i=1}^k\operatorname{rank}(N_{m_i}^j)=\sum_{i=j}^k n_i(i-j)$$

(3)

($\Leftarrow$): 由 (1)，$\dim\operatorname{Im}\mathscr{A}=n-k=n-1$，从而 $k=1$

($\Rightarrow$): 存在 V 的一组基为 $\{\alpha,\mathscr{A}\alpha,\dots,\mathscr{A}^{n-1}\alpha\}$，且 $\mathscr{A}$ 为 n 次幂零变换。$\mathscr{A}$ 对应矩阵为 N^T，N 为上次对角线上全为 1，其余位置为 0 的矩阵，从而 $\dim\operatorname{Im}\mathscr{A}^j=\operatorname{rank}N^j=n-j$

作业 8.16 (P268 T5)

设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换. 如果存在非零向量 $\alpha_0\in V$, 使得由向量 α_0 生成的循环子空间 $C_0\in V$, 则 $\mathscr{A}$ 称为循环变换, 向量 α_0 称为 $\mathscr{A}$ 的循环向量. 证明, $\mathscr{A}$ 为循环变换的充分必要条件是, 存在 V 的基, 使得 $\mathscr{A}$ 在这组基下的方阵为

$$A=\begin{pmatrix}0&\cdots&\cdots&0&-a_0\\1&\ddots&&\vdots&-a_1\\0&\ddots&\ddots&\vdots&\\\vdots&\ddots&1&0&-a_{n-2}\\0&\cdots&0&1&-a_{n-1}\end{pmatrix}.$$

并由此证明, $\mathscr{A}$ 为循环变换的充分必要条件是, $\mathscr{A}$ 的最小多项式等于 $\mathscr{A}$ 的特征多项式. 注: 形如 A 的方阵称为友方阵.

♡

解

$\Leftarrow$: 取 $\{\alpha_1,\dots,\alpha_n\}$，使得 $\mathscr{A}(\alpha_1,\dots,\alpha_n)=(\alpha_1,\dots,\alpha_n)A$。则 $\mathscr{A}\alpha_1=\alpha_2,\dots,\mathscr{A}\alpha_{n-1}=\alpha_n$，故

$$\{\alpha_1,\dots,\alpha_n\}=\{\alpha_1,\mathscr{A}\alpha_1,\dots,\mathscr{A}^{n-1}\alpha_1\}$$

，从而 $V=C_{\alpha_0}$

$\Rightarrow$: 设 $\alpha_0\in V$，使得 $V=C_0$，则 $\dim C_0=n$，由 P229T3 可证

作业 8.17 (P269 T8)

证明: 如果数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换 $\mathscr{A}$ 的二次幂 $\mathscr{A}^2$ 为循环变换, 则 $\mathscr{A}$ 本身也是循环变换. 反之是否成立?

♡

解

由于 $\mathscr{A}^2$ 为循环变换, 所以存在 $\alpha_0\in V$, 使得 $V=\operatorname{Span}\{\alpha_0,\mathscr{A}^2\alpha_0,\dots,\mathscr{A}^{2n-2}\alpha_0\}$, 设 $\dim C_{\alpha_0,\mathscr{A}}=k$，即 $C_{\alpha_0,k}=\operatorname{Span}\{\alpha_0,\mathscr{A}\alpha_0,\dots,\mathscr{A}^{k-1}\alpha_0\}$。

注意对任意 $m\geq k$，均有 $\mathscr{A}^m\alpha_0\in C_{\alpha_0,\mathscr{A}}$，故 $V=C_{\alpha_0,\mathscr{A}^2}\subset C_{\alpha_0,\mathscr{A}}$，故 $C_{\alpha_0,\mathscr{A}}=V$。

反之不成立，反例：取 $\mathscr{A}:\mathbb{F}^2\to\mathbb{F}^2$，$(x_1,x_2)^T\to(0,x_1)^T$ 即可，$\mathbb{F}^2$ 的基为 $\{(1,0)^T,(0,1)^T\}$，则 $\mathscr{A}(e_1,e_2)=(e_2,0)=(e_1,e_2)\begin{pmatrix}0&0\\1&0\end{pmatrix}$，由 P268T5 知 $\mathscr{A}$ 为循环变换，但平方后为 0，不为循环变换

作业 8.18 (P269 T10)

设 $\mathscr{A}$ 与 $\mathscr{B}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的可交换的线性变换, 且 $\mathscr{A}$ 是循环变换. 证明, 存在多项式 $f(\lambda) \in \mathbb{F}[\lambda]$, 使得 $\mathscr{B} = f(\mathscr{A})$.

解 可以使用矩阵思路进行分析，见手写版讲义，下面给出比较简单的做法：

设 $V = \mathrm{Span}\{\alpha_0, \mathscr{A}\alpha_0, \ldots, \mathscr{A}^{n-1}\alpha_0\}$，并设 $\mathscr{B}\alpha_0 = \sum_{i=0}^{n-1} \lambda_i \mathscr{A}^i \alpha_0 = f(\mathscr{A})\alpha_0$，由 $\mathscr{A}, \mathscr{B}$ 可交换性，有对于任意 k，$\mathscr{B}(\mathscr{A}^k\alpha_0) = f(\mathscr{A})(\mathscr{A}^k\alpha_0)$，从而 $\mathscr{B} = f(\mathscr{A})$。

8.2 补充内容

8.2.1 求 Jordan 标准形的过渡矩阵方法

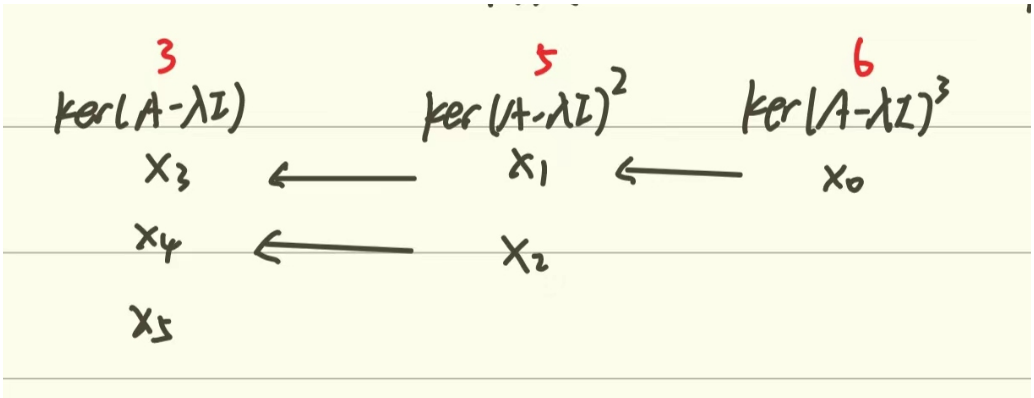
李尚志书中有关于解法的详细说明和几个例子，见书 7.4 节，例子从 P388 开始，电子版群文件中有。以下为我自己研究出来的较为简单的一种算法，供大家参考。

问题 8.1 求 $P^{-1}AP = J$ 中的过渡矩阵 P。

解

Step 1: $\varphi_A(\lambda) = (\lambda - \lambda_1)^{n_1} \ldots (\lambda - \lambda_t)^{n_t}$，对每个特征值考虑

Step 2: 对特征值 λ，找 k 为满足 $\mathrm{Ker}(\mathscr{A} - \lambda I)^k = n$ 的最小正整数，这一步我们需要取 n 个向量，我们不妨特征值 λ 的代数重数为 6，找到的 k 为 3，且 $\dim \mathrm{Ker}(\mathscr{A} - \lambda I) = 3, \dim \mathrm{Ker}(\mathscr{A} - \lambda I)^2 = 5$



1. 我们先取 x_0，满足 $(A - \lambda I)^2 x_0 \neq 0$
2. 取 $x_1 = (A - \lambda I)x_0$，取 $x_2 \in \mathrm{Ker}(A - \lambda I)^2$ 与 x_1 线性无关，且 $(A - \lambda I)x_2 \neq 0$（这里就能看出为什么要设 $(A - \lambda I)^2 x_0 \neq 0$ 了）
3. 取 $x_3 = (A - \lambda I)x_1, x_4 = (A - \lambda I)x_2$，$x_5$ 取与这两个向量线性无关
4. 若此时 x_4, x_3 线性相关，则重新回到步骤 2，直至 x_3, x_4 线性无关即可（这一步通常很难出现，所以整体并不太复杂）
5. $\{x_3, x_1, x_0, x_4, x_2, x_5\}$ 就是我们想要的 6 个向量（注意排序方法），对应的 Jordan 块为

$$\mathrm{diag}(J_3(\lambda), J_2(\lambda), J_1(\lambda))$$

Step 3: 所有找到的向量集合拼起来就得到了矩阵 P，同时也可以得到具体的 Jordan 标准形 J。

大家可以使用该方法试着计算一下李尚志书中的几个例子，这种方法的特殊反例见纸质版讲义，截图如下：

例3 $A=\begin{pmatrix} 2 & 0 & 1 & 2 & 3 \\ 0 & 2 & 0 & 1 & -1 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

① $\varphi_A(\lambda)=(\lambda-2)^5$

② $A-2I=\begin{pmatrix} & & 1 & 2 & 3 \\ & & & 1 & -1 \\ & & & 1 & 2 \\ & & & & \\ & & & & \end{pmatrix}$ $(A-2I)^2=\begin{pmatrix} & & & 1 & 2 \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}$ $(A-2I)^3=0$

$\begin{pmatrix} 0 \\ 8 \\ 3 \end{pmatrix}$

$\begin{matrix} 1 \\ 3 \\ 0 \end{matrix}$

取 $x_0=(0,0,0,0,1)^T$

$x_1=(A-2I)x_0=(3,-1,2,0,0)^T$

$x_3=(A-2I)x_1=(2,0,0,0,0)$

若此时取 $x_2=(0,0,1,0,0)^T \Rightarrow x_4=(1,0,0,0,0)^T$ 与 x_3 相关

故需换一个 x_2，如取 $x_2=(0,0,0,2,-1)^T \Rightarrow x_4=(1,3,0,0,0)^T$.

取 $P=(x_3,x_1,x_0,x_4,x_2)$ 即可

第 9 章 有理标准形

9.1 习题讲解

9.1.1 第十六周

作业 9.1 (P276 T1)

设 A 为 n 阶复方阵，所有不同特征值为 $\lambda_1,\dots,\lambda_t$，证明：

1. 存在 m 使得 $\operatorname{rank} A^m = \operatorname{rank} A^{m+1} = \operatorname{rank} A^{m+2} = \dots$；
2. 设 m_j 为使 $\operatorname{rank}(A-\lambda_j I)^{m_j} = \operatorname{rank}(A-\lambda_j I)^{m_j+1}$ 的最小正整数，则 A 最小多项式为
$$d(\lambda) = \prod_{j=1}^{t}(\lambda-\lambda_j)^{m_j}$$
3. 设 $(\lambda-\lambda_j)^l$ 是 A 属于特征值 λ_j 的初等因子，则 $l \le m_j$；
4. 设 A 属于 λ_j 的初等银子组为 $(\lambda-\lambda_j)^{m_{j1}},\dots,(\lambda-\lambda_j)^{m_{jk_j}}$，并将属于 λ_j 次数为 l 的初等因子记为 n_{jl} (若为 0 则代表无此次数初等因子)，则
$$n_{jl} = \operatorname{rank}(A-\lambda_j I)^{l+1} + \operatorname{rank}(A-\lambda_j I)^{l-1} - 2\operatorname{rank}(A-\lambda_j I)^{l}, \quad 1 \le l \le m_j$$

♡

解

1. 通过 Frobenius 秩不等式有
$$\operatorname{rank} A^m A + \operatorname{rank} A^m A - \operatorname{rank} A^m \le \operatorname{rank} A A^m A$$
结合 $\operatorname{rank} A^{m+2} \le \operatorname{rank} A^{m+1}$ 即有 $\operatorname{rank} A^m = \operatorname{rank} A^{m+1}$ 时 $\operatorname{rank} A^{m+1} = \operatorname{rank} A^{m+2}$，由于 $\operatorname{rank} A^m$ 对 m 单调不增，且取值有限，即得成立。
2. 由于 $\operatorname{rank} J_n(0)^m = \begin{cases} n-m & m<n \\ 0 & m \ge n \end{cases}$，且准对角矩阵的秩为各对角块之和，考虑 A 的 Jordan 标准形即可知 m_j 是 n 最大的 $J_n(\lambda_j)$ 对应的 n，从而即为最小多项式形式。
3. 由于初等因子 $(\lambda-\lambda_j)^l$ 对应 Jordan 块 $J_l(\lambda_j)$ 与上一问即得证。
4. 根据第二问的计算，$\operatorname{rank}(A-\lambda_j I)^l - \operatorname{rank}(A-\lambda_j I)^{l-1}$ 代表所有 λ_j 对应的大于等于 l 阶的 Jordan 块的数量，再做差分即得到 n_{jl}。

作业 9.2 (P277 T2(2)(3)(6))

利用上题方法计算下列矩阵 Jordan 标准形：
$$\begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}, \quad \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{0} & I_{n-1} \\ 1 & \mathbf{0} \end{pmatrix}$$

♡

解 分别为 $\operatorname{diag}(J_2(-1),0)$、$J_3(1)$ 与 $\operatorname{diag}(1,\omega,\dots,\omega^{n-1}), \omega = \mathrm{e}^{2\pi\mathrm{i}/n}$。[第三题事实上可直接通过特征多项式无重根得到结果。]

作业 9.3 (P277 T3)

若 n 维复线性空间 V 上线性变换 $\mathscr{A}$ 所有不同特征值为 $\lambda_1,\dots,\lambda_t$，特征多项式中 $(\lambda-\lambda_j)$ 对应次数为 e_j，最小多项式中为 m_j，证明 $\operatorname{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{m_j}=\operatorname{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{e_j}$。♡

解 由特征多项式知识可知 $m_j\le e_j$。利用 $\operatorname{Ker}(\mathscr{A}-\lambda_j\mathscr{I})^{m_j}\subset\operatorname{Ker}(\mathscr{A}-\lambda_j\mathscr{I})^{e_j}$，只需证明维数相同即可得到两空间相等，而在矩阵表示下根据解空间维数可知这即等价于 $\operatorname{rank}(A-\lambda_jI)^{m_j}=\operatorname{rank}(A-\lambda_jI)^{e_j}$，利用第一题可得结果。

作业 9.4 (P304 T1)

每行每列恰有一个 1，其余为 0 的方阵称为置换方阵，证明其相似于对角形。♡

解 直接计算可知置换方阵的乘积仍然为置换方阵，由于共有 $n!$ 个，对任何置换方阵 A，必存在 $A^m=A^n, m<n$。直接计算可知其行列式为 ±1，从而可逆，于是 $A^{n-m}=I$，也即 $x^{n-m}-1$ 为化零多项式，无重根，最小多项式为其因式，亦无重根，从而可相似对角化。

作业 9.5 (P309 T2)

设 $2k$ 维实线性空间 V 线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)=(\lambda^2+a\lambda+b)^k$，其中 a,b 为实数，且 $a^2-4b<0$，证明存在 V 的一组基使得其在这组基下方阵为

$$J=\begin{pmatrix}B & E & & \\ & B & \ddots & \\ & & \ddots & E\\ & & & B\end{pmatrix},\quad B=\begin{pmatrix}0 & 1\\ -b & -a\end{pmatrix},\quad E=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$$

♡

解 考虑矩阵表示。由本节定理 1，只需说明两方阵相似即可。由于最小多项式次数为 $2k$，A 的相似标准形即为 $\operatorname{diag}(J_k(\lambda_1),J_k(\lambda_2))$，$\lambda_1,\lambda_2$ 为 $\lambda^2+a\lambda+b=0$ 的两根，下面计算 J 的相似标准形。

直接利用准上三角阵行列式计算可知 J 的特征多项式也为 $d(\lambda)$，于是只要说明最小多项式亦为此。由实多项式虚根成对，其不变因子均为实多项式，因此每个均为 $\lambda^2+a\lambda_b$ 的某次方，从而可知最小多项式必然为 $(\lambda^2+a\lambda+b)^t$，于是可直接计算 [这里星号代表任意]

$$J^2+aJ+bI=\begin{pmatrix}O & BE+EB & * & * \\ & O & \ddots & * \\ & & \ddots & BE+EB\\ & & & O\end{pmatrix}$$

由于 $BE+EB$ 的特征值非 0，利用分块矩阵直接计算可知 $(J^2+aJ+bI)^t, t<k$ 不为化零多项式，从而得证。

作业 9.6 (P309 T3)

证明任何实方阵 A 可实相似于准对角形，其对角块为 $J_n(\lambda_0)$ 或

$$K_n(a,b)=\begin{pmatrix} B & I & & \\ & B & \ddots & \\ & & \ddots & I \\ & & & B \end{pmatrix},\quad B=\begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix},\quad I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

解 考虑矩阵表示。将上题结论中的 E 换为 $I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$，类似可证明结论仍然成立。利用不变因子可说明当 $\lambda\notin\mathbb{R}$ 时实方阵的 Jordan 块 $J_n(\lambda)$ 与 $J_n(\bar{\lambda})$ 必然成对出现，从而可以相似为某个 $K_n(a,b)$，即得证。

作业 9.7 (P309 T4)

设 A 是数域 $\mathbb{F}$ 上的方阵，证明其视作数域上方阵与 $\mathbb{C}$ 上方阵时特征多项式与最小多项式不变。

解 由于行列式因子不变，不变因子不随数域变化，从而得证。

9.1.2 第十七周

作业 9.8 (P327 T1)

设 $\alpha_1,\dots,\alpha_k$ 为 n 维 Euclid 空间 V 中两两正交的单位向量，设 $\alpha\in V$，证明

$$\sum_{i=1}^{k}|(\alpha,\alpha_i)|^2\leq\|\alpha\|^2$$

且 $\beta=\alpha-\sum_{i=1}^{k}(\alpha,\alpha_i)\alpha_i$ 与每个 α_i 正交。

解 将 $\alpha_1,\dots,\alpha_k$ 扩充 $\alpha_{k+1},\dots,\alpha_n$ 成为标准正交基，则展开计算可知

$$\|\alpha\|^2=(\alpha,\alpha)=\sum_{i=1}^{n}|(\alpha,\alpha_i)|^2\geq\sum_{i=1}^{k}|(\alpha,\alpha_i)|^2$$

类似直接展开计算可得 β 与每个 α_i 正交。

作业 9.9 (P327 T2)

设 $\alpha_1,\dots,\alpha_n$ 为 n 维 Euclid 空间 V 中一组向量，证明以下三者等价：

1. $\alpha_1,\dots,\alpha_n$ 为标准正交基；
2. $\forall\alpha,\beta\in V,(\alpha,\beta)=\sum_{i=1}^{n}(\alpha,\alpha_i)(\alpha_i,\beta)$；
3. $\forall\alpha\in V,\|\alpha\|^2=\sum_{i=1}^{n}|(\alpha,\alpha_i)|^2$。

解 1 推 2 可以直接计算得到，2 推 3 代入 $\beta=\alpha$ 即可，于是只需证明 3 推 1。

设空间一组标准正交基 $e_1,\dots,e_n$，并设 $(\alpha_i,e_j)=a_{ij}$，$(\alpha,e_j)=x_j$，代入配方计算可得结论。

作业 9.10 (P328 T5)

设 $\xi_1,\ldots,\xi_n$ 为 n 维 Euclid 空间 V 的一组标准正交基，向量 α_i 在这组基下的坐标为向量 x_j，并设以 x_j 为列的矩阵为 X，证明 Gram 方阵 $G(\alpha_1,\ldots,\alpha_n)$ 满足

$$\det G(\alpha_1,\ldots,\alpha_n)=(\det X)^2$$

解 计算得 $(\alpha_i,\alpha_j)=x_i^Tx_j$，从而 $G(\alpha_1,\ldots,\alpha_n)=X^TX$，于是

$$\det G(\alpha_1,\ldots,\alpha_n)=\det(X^TX)=\det X^T\det X=(\det X)^2$$

作业 9.11 (P328 T7)

设 $\alpha_1,\ldots,\alpha_n$ 为 n 维 Euclid 空间 V 中一组向量，证明

$$\det G(\alpha_1,\ldots,\alpha_n)\le\prod_{i=1}^n\|\alpha_i\|^2$$

等号成立当且仅当 $\alpha_1,\ldots,\alpha_n$ 两两正交或有零向量，由此证明对 n 阶实方阵 A 有

$$(\det A)^2\le\prod_{i=1}^n\sum_{j=1}^n a_{ij}^2$$

解 利用上题表示可知本题的两问均等价于证明对方阵 X 有

$$(\det X)^2\le\prod_{i=1}^n x_i^Tx_i$$

考虑 QR 分解可知存在正交阵 Q 使得 $QX=R$，其中 R 为上三角阵，记其列向量为 r_i，根据正交阵性质

$$(\det X)^2=(\det Q)^2(\det X)^2=(\det R)^2$$

$$\prod_{i=1}^n x_i^Tx_i=\prod_{i=1}^n x_i^TQ^TQx_i=\prod_{i=1}^n r_i^Tr_i$$

于是只需说明对上三角阵 R 结论成立，此时左侧为 $\prod_{i=1}^n r_{ii}^2$，右侧为 $\prod_{i=1}^n\sum_{j=1}^n r_{ij}^2$，于是左小于等于右。可计算验证取等当且仅当 r_i 彼此正交或有零向量，计算得到 x_i 满足同样性质，即得结论。

作业 9.12 (P328 T9)

设 O 为 n 阶正交方阵，$A=\operatorname{diag}(a_1,\ldots,a_n)$，证明 OA 的任何特征值 λ_0 满足 $m\le|\lambda_0|\le M$，其中 $m=\min_j|a_j|,M=\max_j|a_j|$。

解 对 $OA\alpha=\lambda_0\alpha$ 两边取模长得到 $\|OA\alpha\|=|\lambda_0|\|\alpha\|$，从而 $\|A\alpha\|=|\lambda_0|\|\alpha\|$。而直接比较系数得 $m\|\alpha\|\le\|A\alpha\|\le M\|\alpha\|$，又由 $\|\alpha\|\ne0$ 得证。

作业 9.13 (P336 T7)

设 $\mathbb{R}^{n\times n}$ 是所有 n 阶实方阵在内积 $(X,Y)=\operatorname{tr}(XY^T)$ 下构成的 Euclid 空间，设 P 为某可逆方阵，计算 $\mathscr{A}_P(X)=P^{-1}XP$ 的伴随变换 $\mathscr{A}_P^*$。

解 由定义须 $\operatorname{tr}((P^{-1}XP)Y^T)=\operatorname{tr}(X\mathscr{A}_P^*(Y))$，由 $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ 得

$$\operatorname{tr}((P^{-1}XP)Y^T)=\operatorname{tr}(XPY^TP^{-1})=\operatorname{tr}(X(P^{-T}YP^T)^T)$$

于是 $\mathscr{A}_P^*(Y)=P^{-T}YP^T$，可验证其为线性变换，从而成立。

作业 9.14 (P342 T6)

设 $\alpha=\beta+\mathrm{i}\gamma$ 是实规范方阵 A 属于特征值 λ 的特征向量，其中 β,γ 为实向量，证明：

1. α 为 A^T 属于特征值 $\bar{\lambda}$ 的特征向量；
2. $\lambda\notin\mathbb{R}$ 时 $\beta^T\gamma=0$ 且 $\|\beta\|=\|\gamma\|$。

解

1. 由 A 为实方阵有
$$(A^T\alpha-\bar{\lambda}\alpha)^H(A^T\alpha-\bar{\lambda}\alpha)=(\alpha^HA-\lambda\alpha^H)(A^T\alpha-\bar{\lambda}\alpha)$$
再通过规范性展开即得其为 0，从而得证。
2. 将实部虚部展开可知此题结论等价于 $\alpha^T\alpha=0$，由 $\lambda\notin\mathbb{R}$ 直接计算可知
$$\lambda\alpha^T\alpha=(\lambda\alpha)^T\alpha=\alpha^TA^T\alpha=\bar{\lambda}\alpha^T\alpha$$
从而由 $\lambda\neq\bar{\lambda}$ 只能 $\alpha^T=\alpha$，即得证。

若想不到这种办法，一个更自然的思路是从正交相似标准形出发考虑。

作业 9.15 (P347 T3)

设 α_1,α_2 与 β_1,β_2 是 n 维 Euclid 空间 V 中两对向量，且 $\|\alpha_1\|=\|\beta_1\|,\|\alpha_2\|=\|\beta_2\|$，$\alpha_1,\alpha_2$ 夹角与 β_1,β_2 夹角相同，证明存在正交变换 $\mathscr{A}$ 使得 $\mathscr{A}(\alpha_1)=\beta_1,\mathscr{A}(\alpha_2)=\beta_2$。

解 考虑矩阵表示。记 $H_u=I-2uu^T,\|u\|=1$，可发现 H_u 正交。先找 H_u 使得 $H_u\alpha_1=\beta_1$，代入可得
$$\alpha_1-2(\alpha_1^Tu)u=\beta_1$$
，于是令 $u=\frac{\beta_1-\alpha_1}{\|\beta_1-\alpha_1\|}$，代入验证可知成立。若此时 $\beta_2=H_u\alpha_2$ 则已找到，否则令 $v=\frac{\beta_2-H_u\alpha_2}{\|\beta_2-H_u\alpha_2\|}$，则可验证 $H_vH_u\alpha_2=\beta_2$，而
$$H_vH_u\alpha_1=H_v\beta_1=\beta_1-\frac{1}{\|\beta_2-H_u\alpha_2\|^2}(\beta_2-H_u\alpha_2)(\beta_2-H_u\alpha_2)^T\beta_1$$
为证其为 β_1，只需 $(\beta_2-H_u\alpha_2)^T\beta_1=0$，由定义 $H_u^T=H_u$，进一步计算验证可知 $H_u^{-1}=H_u$，于是由模长相等、内积相等可知原式为
$$\beta_2^T\beta_1-\alpha_2^TH_u\beta_1=\alpha_2^T(\alpha_1-H_u\beta_1)=\alpha_2^TH_u(H_u\alpha_1-\beta_1)=0$$
从而令 $A=H_vH_u$ 即得证。

此做法事实上有很强的几何意义，详见补充内容。

作业 9.16 (P347 T5(1))

求正交方阵 O 使得 $B=O^TAO$ 是正交方阵
$$A=\frac{1}{3}\begin{pmatrix}2&-1&2\\2&2&-1\\-1&2&2\end{pmatrix}$$
在正交相似下的标准形。

解 直接由上节算法计算得可取

$$B=\begin{pmatrix}\frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{14}} & \frac{3}{\sqrt{14}}\\ \frac{4}{\sqrt{42}} & -\frac{5}{\sqrt{42}} & \frac{1}{\sqrt{42}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}}\end{pmatrix}$$

作业 9.17 (P347 T8)

设 n 阶正交方阵 O 特征值不为 -1，证明，$I+O$ 可逆，$K=(I-O)(I+O)^{-1}$ 反对称，且 $O=(I-K)(I+K)^{-1}$。

解 由特征值定义知可逆，而

$$K^T=(I+O^T)^{-1}-(I+O^T)^{-1}O^T=(I+O^T)^{-1}-(O(I+O^T))^{-1}=(I+O^T)^{-1}-(I+O)^{-1}$$
$$K=(I+O)^{-1}-O(I+O)^{-1}=(I+O)^{-1}-((I+O)O^T)^{-1}=(I+O)^{-1}-(I+O^T)^{-1}$$

从而 K 反对称。

对第二个式子，由于

$$I-K=(I+O)(I+O)^{-1}-(I-O)(I+O)^{-1}=2O(I+O)^{-1}$$
$$I+K=(I+O)(I+O)^{-1}+(I-O)(I+O)^{-1}=2(I+O)^{-1}$$

直接计算即得证。

类似地，若想不到上述方法，由于将 O,K 同时正交相似为 P^TOP,P^TKP 不影响结论，于是考虑 O 的正交相似标准形可知只需对每个对角块说明成立，直接计算验证即可。

9.2 补充内容

9.2.1 初等正交阵

记 E_{ij} 为第 i 行第 j 列为 1，其他为 0 的矩阵，定义 Givens 方阵

$$G_{ij}(\theta)=I+(\cos\theta-1)E_{ii}+(\cos\theta-1)E_{jj}+\sin\theta E_{ij}-\sin\theta E_{ji},\quad i\neq j$$

则计算得其行列式为 1，且 $G_{ij}^{-1}(\theta)=G_{ij}(-\theta)=G_{ij}^T(\theta)$。

对**单位向量** u，定义 Householder 方阵

$$H_u=I-2uu^T$$

则计算得其行列式为 -1，且 $H_u^{-1}=H_u=H_u^T$。

这两种方阵合称**初等正交阵**。

定理 9.18 (几何意义)

右乘列向量的意义下，$G_{ij}(\theta)$ 代表在 ij 平面上绕原点逆时针旋转 θ 的旋转变换，而 H_u 则代表对过原点垂直于 u 的超平面的反射变换。

证明 由于 $G_{ij}(\theta)$ 只改变 i,j 分量，直接根据旋转变换形式可知成立。对 H_u，由于 $H_u\alpha=\alpha-2(u^T\alpha)u$，在 u 为单位向量时 $u^T\alpha$ 即为 α 在 u 上分量，-2 使此分量倒转，即代表反射。

回到作业 9.15，我们的操作事实上是：先通过反射将 α_1 反射到 β_1，这时 α_2 来到 $H_u\alpha_2$，由于

$H_u\alpha_2,\beta_2$ 均与 β_1 夹角相同，β_1 在 $H_u\alpha_2,\beta_2$ 中垂超平面上，因此将 $H_u\alpha_2$ 反射到 β_2 不改变 β_1，从而即找到了变换。

类似可逆阵用初等方阵分解，正交阵也可通过初等正交阵分解，也即：

定理 9.19 (初等正交阵分解)

任何行列式为 1 的 n 阶正交阵可以分解为至多 $\frac{n(n-1)}{2}$ 个 Givens 方阵乘积。

任何 n 阶正交阵可分解为至多 n 个 Householder 方阵的乘积，或至多 $\frac{n(n-1)}{2}$ 个 Givens 方阵与可能存在的 $\mathrm{diag}(I_{n-1},-1)$ 乘积。

证明 对向量 α，记 $\alpha_0=(\|\alpha\|,0,\dots,0)^T$。我们先说明，一个 Householder 阵可使得 α 变为 α_0；维数至少为 2 时至多 $n-1$ 个 Givens 阵可使得 α 变为 α_0。

对 Householder 阵，直接构造 $u=\frac{\alpha-\alpha_0}{\|\alpha-\alpha_0\|}$ 验证得 H_u 符合条件；而对 Givens 阵，利用三角函数知识直接计算可知

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}=\begin{pmatrix}\sqrt{a^2+b^2}\\ 0\end{pmatrix}$$

必然有解，我们依次取 $G_{12},\dots,G_{1n}$，重复此过程即可构造。

对原问题使用归纳，对任何 n 阶正交阵，由于其每列模长 1，可以找到对应的一个 Householder 阵或 $n-1$ 个 Givens 阵使得第一列成为 e_1。利用正交性计算验证可知此时第一行也必须只有首个元素为 1，从而化为 $n-1$ 阶情况。经过 $n-1$ 个 Householder 方阵或 $\frac{n(n-1)}{2}$ 个 Givens 方阵操作后，原矩阵只能成为 I 或 $\mathrm{diag}(I_{n-1},-1)$。

若行列式为 1，则必须为 I，已经得到了分解；若行列式为 -1，对 Givens 阵乘 $\mathrm{diag}(I_{n-1},-1)$ 即为分解；由于 $\mathrm{diag}(I_{n-1},-1)$ 为 Householder 阵 H_{e_n}，这也得到了 n 个 Householder 阵的分解。

在计算数学中，这两种方阵常用来在 SVD 分解等需要正交阵操作的情况作为基本元素组合出正交阵。

9.2.2 有理标准形

注意到，初等因子组可以在**任何域** (甚至未必是数域) 上定义，我们有如下结论：

定理 9.20 (有理标准形)

在某个域上，对矩阵 A 的每个初等因子 $f(\lambda)^n$，f 不可约，设 $f(\lambda)=\lambda^k+a_{k-1}\lambda^{k-1}+\dots+a_0$，定义矩阵

$$J_n(f(\lambda))=\begin{pmatrix}F & E_{n1} & & \\ & F & \ddots & \\ & & \ddots & E_{n1}\\ & & & F\end{pmatrix},\quad F=\begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1\\ -a_{n-1} & \cdots & -a_1 & -a_0\end{pmatrix}$$

这里 F 即为多项式 $f(\lambda)$ 的友方阵。

矩阵 A 的有理标准形定义为将每个初等因子对应的 $J_n(f(\lambda))$ 作为对角块的准对角阵，其可以在域上相似到此标准形，且在对角块次序可交换的意义下唯一。

此结论真正统一了之前对多项式、循环子空间与相似标准形的讨论，给出了任何域上矩阵的相似

标准形，特别地：

- 当域是**代数闭**的，即任何多项式可分解为一次因式 (如 $\mathbb{C}$)，则 $f(\lambda)$ 一定为一次多项式，对应友方阵即 $(-a_0)$，有理标准形即为 Jordan 标准形的形式；
- 当域未必代数闭，有理标准形与 Jordan 标准形不同。例如，对 $\mathbb{R}$，有理标准形即为作业 9.6 确定的形式中将 I 换回作业 9.5 的 E；
- 由于不变因子对扩域不变，两个矩阵在某个域上相似当且仅当它们在其**代数闭包** [包含它的最小代数闭域] 上相似，也即可以通过 Jordan 标准形讨论。

下面给两个有趣的计算例子：

例题 9.1 计算 $\mathbb{F}_2$ 上 $\begin{pmatrix} \mathbf{0} & I_3 \\ 1 & \mathbf{0} \end{pmatrix}$ 的有理标准形。

解 直接计算可知其行列式因子 d_1 至 d_3 为 0，$d_4 = D_4 = \lambda^4 + 1$ [二元域上 ± 1 无区别]，进一步计算可验证

$$\lambda^4 + 1 = (\lambda + 1)^4$$

于是初等因子组为 $(\lambda + 1)^4$，对应有理标准形为

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}$$

例题 9.2 计算 $\mathbb{F}_2$ 上 $\begin{pmatrix} \mathbf{0} & I_2 \\ 1 & \mathbf{0} \end{pmatrix}$ 的有理标准形。

解 直接计算可知其行列式因子 $d_1 = d_2 = 0$，$d_3 = D_3 = \lambda^3 + 1$，进一步计算可验证

$$\lambda^3 + 1 = (\lambda + 1)(\lambda^2 + \lambda + 1)$$

于是初等因子组为 $\lambda + 1, \lambda^2 + \lambda + 1$，对应有理标准形为

$$\begin{pmatrix} 1 & & \\ & 0 & 1 \\ & 1 & 1 \end{pmatrix}$$

对一般的 n，事实上与 $\mathbb{F}_2$ 上的分圆多项式有关。

9.3 期末复习

1. 矩阵 [详见期中复习]
2. 行列式 [详见期中复习]
3. 秩
   (a). 定义、等价定义
   (b). **Hermite 标准形**
   (c). 秩不等式
   (d). 线性方程组的解
   (e). 广义逆 $AA^-A = A, A^-AA^- = A^-$[具体定义应看题干]

   (f). 相抵标准形刻画广义逆
   (g). * 奇异值分解与广义逆 $(AA^-)^H = AA^-, (A^-A)^H = (A^-A)$
   (h). 向量组视角：行秩、列秩
   (i). 线性变换视角：**像空间维数**
4. 相合 [详见期中复习]
5. 正交方阵
   (a). 定义与**几何性质**
   (b). 内积空间与标准正交基
   (c). 正交化过程
   (d). 正交相抵、QR 分解、**奇异值分解**
   (e). 正交相似、规范方阵
   (f). * 正交相似标准形的计算
6. 多项式 [详见期中复习，期末不太可能重点]
7. 相似
   (a). 相似的基本性质
   (b). **相似标准形**
   (c). 特征方阵与相似标准形计算
   (d). 特征多项式、最小多项式
   (e). **实数域上的相似**
   (f). * 相似标准形的多项式与幂级数
   (g). * 一般域上的相似条件
   (h). * 有理标准形
8. 线性空间 [详见期中复习]
9. 线性映射
   (a). 定义与表示
   (b). 像与核
   (c). * 线性空间同构定理
   (d). 对偶空间
   (e). 特征子空间
   (f). 根子空间、**循环子空间**
   (g). 空间视角的 Jordan 标准形

附录 A 参考资料

- zhuanlan.zhihu.com/p/346371552 [ZFC 公理系统]
- zhuanlan.zhihu.com/p/346542841 [ZFC 公理系统]
- zhuanlan.zhihu.com/p/602850933 [正规子群、商群]
- 李炯生《线性代数》网络解答
- 王新茂《线性代数讲义》