前言

To be added.

郭学军
2011 年 5 月于南京

目　录

第一章　多项式

在本章里, 我们将学习整数，数域和多项式的基本知识.

§1.1　整数

自然数起源于计数。例如一棵树,两只眼睛,三只羊,⋯。这样便有了自然数

$$1,\ 2,\ 3, \ldots.$$

这个过程看似简单，实际上是人类认识上的一个巨大的飞跃。这是一个从具体到抽象的过程。有了这个进步，人们可以看到10个苹果和10个人之间共同的东西。这对人类认识周围的世界, 管理资源分配等活动都有实在的指导意义。而且从具体到抽象，也使人摆脱了实物的局限性，使得$1,\ 2,\ 3, \ldots.$可以指代任何具体的事物，是文明的发展必不可少的一步。

0的出现比$1,\ 2,\ 3, \ldots.$等要晚很多，而0真正的像$1,\ 2,\ 3, \ldots.$等一样平等的参与数学运算也不过只有几百年的历史。现在一般都把0也算作是自然数。自然数的集合用$\mathbb{N}$表示。

有了自然数以后，用0减去自然数，便有了负整数。自然数和负整数统称为整数。整数的集合用$\mathbb{Z}$表示。考虑两个非零整数的商，便有了有理数。有理数的集合用$\mathbb{Q}$表示。

我们知道$\mathbb{Z}$中的元素都满足下面的法则。

(A1) 加法结合律:$(a+b)+c=a+(b+c)$;

(A2) 加法交换律:$a+b=b+a$;

(A3) 有元素0，满足：$a+0=0+a=a$;

(A4) 对每个元素a，总有元素b，使得$a+b=b+a=0$;

(M1) 乘法结合律: $(a\cdot b)\cdot c=a\cdot(b\cdot c)$;

(M2) 乘法交换律: $ab=ba$;

(M3) 有元素1，满足性质，$1\cdot a=1\cdot a=a$;

(D) 加乘分配律$a\cdot(b+c)=a\cdot b+a\cdot c$.

而$\mathbb{Q}$中的元素除了满足上面的8条法则之外，还满足

(I) 有逆元。对任意非零的有理数a, 都存在一个有理数b使得$ab=ba=1$.

两千多年的毕达哥拉斯学派曾经以为世界上所有的数都是有理数。但是后来希伯斯证明了$\sqrt{2}$不是有理数。这使得$\sqrt{2}$成了被发现第一个无理数。因为$\sqrt{2}$是边长为1的正方形的对角线的长度，所以$\sqrt{2}$ 是一个相当自然的对象。

其实不但$\sqrt{2}$是个无理数，只要n不是一个完全平方数，$\sqrt{n}$就是无理数。有理数和无理数构成了全体实数$\mathbb{R}$.

假设$b,\ c \in \mathbb{Z}$，非零$m \in \mathbb{N}$。如果$m|(b-c)$, 那么我们称$b,\ c$模m同余，记为

$$b \equiv c(\mod m).$$

m叫做模，b叫做c模m的剩余。例如：

$$-9 \equiv 16(\mod 5).$$

注意这里的$b,\ c$是整数，而不能是分数。但是如果

$$bc \equiv 1(\mod m),$$

我们也常常把b记为$1/c$.

给定a, a模m的所有的剩余（我们称之为一个剩余类）都形如

$$a+km,\ k \in \mathbb{Z}.$$

定理 1.1.1. 假设m是一个正整数，$A,\ a$是任意两个整数，则

$$a,\ a+1,\ ...,\ a+m-1$$

中有且恰有一个模m同余于A.

因此对任一整数A, 都有它的一个模m剩余位于集合

$$\{0,\ 1,\ 2,\ ...,\ m-1\}$$

中。这样的集合我们成称为是模m的一个完全剩余系。任何一个含有m个元素的整数集合$\{a_0,\ ...,\ a_{m-1}\}$，如果其中的元素模m两两不同。我们都称它是一个完全剩余系。

命题 1.1.2. 如果$A \equiv a(\mod m)$, $B \equiv b(\mod m)$，那么$AB \equiv ab(\mod m)$. 假设$f(x)$是一个整系数的多项式。如果$a \equiv b(\mod m)$, 那么$f(a) \equiv f(b)(\mod m)$.

假设m是一个正整数，我们定义$\varphi(m)$为$1,\ 2,\ ...,\ m-1$中与m互素的元素的个数。我们称$\varphi(m)$为欧拉函数。如果一个含有$\varphi(m)$个元素的整数集合$S=\{a_1,\ ...,\ a_{\varphi(m)}\}$中的元素模$m$ 两两不同，而且都和m互素，我们就称S是模m的一个缩系。

定理 1.1.3. 假设m是一个正整数, a和m互素. 则$a^{\varphi(m)} \equiv 1(\mod m)$.

证明 假设$S=\{a_1, ..., a_{\varphi(m)}\}$是模$m$的一个缩系。那么$T=\{aa_1, ..., aa_{\varphi(m)}\}$也是模$m$的一个缩系。所以两个缩系中的元素乘起来以后模$m$是相同的。所以$a^{\varphi(m)}\equiv 1(\mod m)$. □

只能被1和其自身整除的整数称为是素数。例如

$$2, 3, 5, 7, 11, ...$$

注意1不是素数，普林斯顿大学的著名数学家Conway认为应该修改素数的定义，使得

$$-1$$

是一个素数。同时注意2是一个素数，2是唯一的一个偶素数。

定理 1.1.4. *素数有无限多个。*

证明 $n=p_1\cdots p_k+1$. □

定理 1.1.5. *模4余3的素数有无限多个。*

证明 假设模4余3的素数只有$p_1, \cdots, p_k$ $n=p_1^2\cdots p_k^2+2$模4余3，所以必然被某个模4余3的素数整除，可是这个素数显然不在$p_1, \cdots, p_k$中，矛盾. □

定理 1.1.6. *设p是一个素数，a, b是两个正整数，而且$p>a$, $p>b$. 则$p\nmid ab$.*

证明 我们首先固定a, 然后来证明不存在正整数$b<p$使得$p|ab$. 我们下面用反证法来证明。假设这样的b存在，其中一定有最小的。假设这个最小的是x. 也就是说任何满足$b<p$且$p|ab$的正整数b都大于等于x. 我们来推出矛盾。

首先$x>1$. 所以我们知道一定存在一个正整数k使得$kx\le p<(k+1)x$. 我们假设$p=kx+c$，这里的c是一个正整数。所以我们有

$$ac=ap-axk.$$

注意到右边两项都是p的倍数。所以左边也是p的倍数。但是c比x更小。这就引起了矛盾。

□

定理 1.1.7. *设p是一个素数，a, b是两个正整数。如果$p|ab$, 则$p|a$或者$p|b$.*

证明 假设$p\nmid b$，而且$p\nmid b$. 那么取a, b模p的最小的正剩余c, d, 我们知道c, d都是小于p 的正整数。而且这时有$p|cd$。这就与上一个定理矛盾。

□

定理 1.1.8. *设p是一个素数。如果$p|a_1\cdots a_n$, 则存在i使得$p|a_i$.*

证明 归纳即可。由$p|(a_1\cdots a_{n-1})a_n$ 可知$p|(a_1\cdots a_{n-1})$或者$p|a_n$。依次类推。□

定理 1.1.9 (算术基本定理). *每个大于1的自然数均可写为素数的积，而且这些素因子按大小排列之后，写法仅有一种方式。*

证明 首先每个大于1的自然数n均可写为素数的积。如果n本身是一个素数，毋庸赘言。否则它能够被素数p整除，即可以写成$n=pm$的形式，m比n小. 以此类推，有限步之后，就可以把n写为素数的积。

我们只需要证明这种写法的唯一性。

假设

$$n=p_1^{a_1}\cdots p_n^{a_n}=p_1^{b_1}\cdots p_n^{b_n},$$

其中p_1，...，p_n是互不相同的素数。这里的指数a_i 和b_i都有可能是0, 我们假设它们不同时为0. 如果对某个i来说, $a_i>b_i$；那么两边同时除以$p_i^{b_i}$以后，就有

$$p_1^{a_1}\cdots p_i^{a_i-b_i}\cdots p_n^{a_n}=p_1^{b_1}\cdots 1\cdots p_n^{b_n}.$$

这样子p_i能够整除上式的左边，可是根据定理4.3，p_i 不能够整除上式的右边，引起矛盾。所以$a_i>b_i$不成立。同理$a_i<b_i$也不成立。所以$a_i=b_i$总是成立的。所以n写成素数乘积的写法是惟一的。 □

例 1.1.1. *$\sqrt{2}$是无理数。我们可以利用反证法很简单的证明$\sqrt{2}$是个无理数. 假设$\sqrt{2}$是有理数. 那么肯定存在两个互素的整数使得*

$$\sqrt{2}=\frac{a}{b}.$$

两边平方可知，

$$2b^2=a^2.$$

因为明显左边是偶数，所以右边也是偶数，所以a是偶数，能够写成$a=2n$的形式。因为我们已经假设了a，b互素，所以b是奇数。所以

$$2b^2=a^2\Longrightarrow b^2=2n^2.$$

注意到b是奇数，这是不可能的.

例 1.1.2. *$\lg 2=\log_{10}2$是无理数。*

在中国古代著名数学著作《孙子算经》中，有这样一道叫做“物不知数”的题目，

“有物不知其数，三三数之剩二，五五数之剩三，七七数之剩二。问物几何？”

中国数学家秦九韶于1247年做出了完整的解答，口诀如下：

“三人同行七十希，五树梅花廿一支，七子团圆正半月，除百零五便得知”

定理 1.1.10 (中国剩余定理). 假设n_1, ..., n_k是两两互素的正整数，则对于任意的整数a_1, ..., a_k, 同余方程

$$\begin{cases} x \equiv a_1 (\mod n_1), \\ \vdots \\ x \equiv a_k (\mod n_k). \end{cases}$$

有模$N = n_1 \cdots n_k$惟一的解。

其推广的形式为

定理 1.1.11. 假设n_1, ..., n_k是正整数，则对于任意的整数a_1, ..., a_k, 同余方程

$$\begin{cases} x \equiv a_1 (\mod n_1), \\ \vdots \\ x \equiv a_k (\mod n_k). \end{cases}$$

有解当且仅当最大公因子

$$(n_i,\ n_j)|(a_i - a_j).$$

对任意的$i \neq j$成立。有解的时候，解模$N = [n_1,\ ...,\ n_k]$（最小公倍数）惟一。

注意我们这里的解都是一次的，如果是高次的同余方程，那么求解非常困难，而且没有固定的方法。

§1.2 复数

复数域$\mathbb{C}$是由实数添加了虚数单位

$$i = \sqrt{-1}$$

得到的。复数$\mathbb{C} = \{x + yi | x,\ y \in \mathbb{R}\}$. 假设$\alpha = x + yi,\ x,\ y \in \mathbb{R}$. 我们称$x$为$\alpha$的实部, y为α的虚部, $r = \sqrt{x^2 + y^2}$叫做α的模长, 记为$|\alpha|$. $\overline{\alpha} = x - yi$叫做$\alpha$的的共轭. 对$\alpha = x + yi$来说, 我们还可以把$\alpha$写成下面的形式

$$\alpha = r(\frac{x}{r} + \frac{y}{r} i),\ r = \sqrt{x^2 + y^2}.$$

由于$(\frac{x}{r})^2 + (\frac{y}{r})^2 = 1$, 存在角度$\theta$使得

$$\cos\theta = \frac{x}{r},\ \sin\theta = \frac{y}{r},\ 0 \leq \theta < 2\pi.$$

因此
$$\alpha = r(\cos\theta + i\sin\theta),$$
这是复数α的极坐标的写法, θ称为是α的幅角, 记为$\arg\theta$.

对任意的$\alpha = x + yi$, $\beta = z + wi$, 我们定义
$$\alpha + \beta = (x + z) + (y + w)i,$$
$$\alpha\beta = (xz - yw) + (xw - yz)i.$$
不难证明这样定义的加法和乘法满足通用的交换律, 结合律, 分配律等运算法则. 两个复数相等当且仅当它们的实部和虚部都相等. 即$x + yi = z + wi$当且仅当$x = z$而且$y = w$. 对任意的非零$\alpha = x + yi$, 存在唯一的β使得$\alpha\beta = 1$, 记$\beta = \alpha^{-1}$. 实际上,
$$\beta = \frac{x - yi}{x^2 + y^2}.$$
另外, 也不难验证
$$\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}.$$

命题 1.2.1. *假设r_1, ..., r_n是大于等于0的实数, θ_1, ..., θ_n是大于等于0, 小于2π的实数. 则*
$$\prod_{k=1}^{n} r_k(\cos\theta_k + i\sin\theta_k) = (\prod_{k=1}^{n} r_k)(\cos(\theta_1 + \cdots + \theta_n) + i\sin(\theta_1 + \cdots + \theta_n)).$$

证明 只需要对$n = 2$的情形验证即可.

由于
$$\begin{aligned}
&(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2)\\
=&(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)\\
=&\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2).
\end{aligned}$$
所以命题成立. □

由该命题可以推出De Moive定理成立

定理 1.2.2.
$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

命题 1.2.3. *假设n是一个正整数, 方程$x^n - 1$的所有的根是*
$$\cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n},\ k = 0,\ 1,\ ...,\ n - 1.$$

证明 假设$\alpha = r(\cos\theta + i\sin\theta)$是$x^n - 1$的根, 其中$r$是大于等于0的实数, θ是大于等于0, 小于2π的实数. 那么$\alpha^n = 1$当且仅当$r = 1$, 而且$n\theta$ 是2π的整数倍. 由于θ是大于等于0, 小于2π的实数, 所以θ共有n种可能, 即

$$\theta = \frac{2k\pi}{n},\ k = 0,\ 1,\ ...,\ n-1.$$

所以方程$x^n - 1$的所有的根是$\cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n},\ k = 0,\ 1,\ ...,\ n-1.$

□

定义 1.2.1 (数域). 假设F是复数域$\mathbb{C}$的包含$\mathbb{Q}$的子集。我们称F是一个数域，如果它满足下面的条件。

1. 如果$a,\ b \in F$，则$a + b \in F$;
2. 如果$a,\ b \in F$，则$ab \in F$;
3. 如果非零$a \in F$，则$a^{-1} \in F$.

例 1.2.1. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$都是数域。

我们下面来看一看如何从一个数域出发，通过添加元素得到新的数域。

定义 1.2.2 (构造新的数域). 假设$\alpha \in \mathbb{C}$. 我们令

$$F(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} | \text{其中} f \text{ 和} g \text{ 都是} F \text{ 上的多项式而且} g(\alpha) \neq 0 \right\}.$$

很容易看出来这样定义的集合满足数域的定义条件。所以它构成了一个数域。我们称$F(\alpha)$为由F添加一个元素α得到的数域. 只要α不在数域F中。我们这样得到的就是一个新的数域。我们递归的定义添加两个元素得到的数域

$$F(\alpha,\ \beta) = (F(\alpha))(\beta).$$

以此类推，我们可以定义添加任意多个元素得到的数域$F(\alpha_1,\ ...,\ \alpha_n)$.

例 1.2.2 (求证$\mathbb{Q}(\sqrt{2},\ \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$).

证明 首先有定义可知$\mathbb{Q}(\sqrt{2},\ \sqrt{3}) \supseteq \mathbb{Q}(\sqrt{2}+\sqrt{3})$. 因此我们只需要证明$\mathbb{Q}(\sqrt{2},\ \sqrt{3}) \subseteq \mathbb{Q}(\sqrt{2} + \sqrt{3})$即可. 为此只要证明$\sqrt{2} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$ 和$\sqrt{3} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$ 即可。因为$\sqrt{3} - \sqrt{2}$是$\sqrt{3} + \sqrt{2}$的倒数，所以$\sqrt{3} - \sqrt{2} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$。又因为数域对加法封闭，所以$2\sqrt{3} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$。再利用数域对于乘法封闭，即可得$\sqrt{3} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$。同样的方法可证$\sqrt{2} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$。 □

习题 1.2

1. 求复数α使得$\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{2},\ \sqrt{3},\ \sqrt{5})$.

§1.3 四元数

下面我们简单的介绍一下四元数。

”在那个月早些时候，每天我下来吃早饭时，你的小哥哥和你总是喜欢问我’哎，爸爸，你那个三元数能乘了吗？’我总是不得不摇摇头说’还不行，只能加加减减。’在那个月的16号，那天正好是星期一，爱尔兰皇家科学院要开会，我还是主持人呢。我和你妈妈沿着皇家运河走着，她也许还有点急。尽管她时不时的和我说着什么，我的脑子里却在想着别的东西。这头脑深处的潜流最终有了一个结果，毫不夸张的说，我立刻便认识到了它的重要性。电路接通了，火花打着了，信使来报信了！之后多年的思想和工作，有我的，也有别人的，都和这个灵光乍现直接相关。这些后续发展，我当时就预见到了。如果这些工作有什么不足的地方，那是我的责任。这些工作无论如何是要靠别人来发展的。要是我能活得再长一点就好了，那样我就能和别人再交流交流我的成果了。我当时无法遏制自己的冲动，完全失去了理性，当我们经过布卢姆桥的时候，我用一把刀子，把那个基本公式刻在了桥墩的一块石头上，就是

$$i^2 = j^2 = k^2 = ijk = -1.$$

”

摘自1865年8月5号哈密尔顿写给第二个儿子阿奇博尔德.哈密尔顿的信。这封信写完后不到一个月，9月2号，哈密尔顿就去世了。哈密尔顿信中提到的”那个月”是1843 年的10月。哈密尔顿当时发现的是人类历史上第一个不交换的代数－四元数代数。

我们用$\mathbb{R}$表示实数集合, $\mathbb{C}$表示复数集合，$\mathbb{H}$表示四元数集合。四元数集合$\mathbb{H} = \mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$, $\mathbb{H}$中的运算法则就是上面哈密顿在信中写的。从哈密顿的公式可以推出

$$ij = -ji = k,$$
$$jk = -kj = i,$$
$$ki = -ik = j.$$

任何一下四元数都可以写成

$$\alpha = d + ai + bj + ck$$

的形式，其中a, b, c, d都是实数。d叫做α的实部，$ai + bj + ck$ 叫做α 的虚部。共轭定义为

$$\overline{\alpha} = d - ai - bj - ck.$$

两个四元数$\alpha = d + ai + bj + ck$和$\widetilde{\alpha} = \widetilde{d} + \widetilde{a}i + \widetilde{b}j + \widetilde{c}k$ 相等当且仅当它们对应的系

数相等，即$a=\widetilde{a},\ b=\widetilde{b},\ c=\widetilde{c},\ d=\widetilde{d}$. 特别的，$\alpha=d+ai+bj+ck=0$当且仅当$a=b=c=d=0$.

定义 1.3.1. *假设*$\alpha=d+ai+bj+ck$*是一个不等于0的四元数。令*

$$\beta=\frac{1}{d^2+a^2+b^2+c^2}\overline{\alpha}=\frac{1}{d^2+a^2+b^2+c^2}(d-ai-bj-ck),$$

则$\alpha\beta=\beta\alpha=1$. *我们称*$\beta$*是*$\alpha$*的逆。根据定义，*$\alpha$*的逆是由*$\alpha$*唯一确定的，而且如果*$\beta$*是*$\alpha$*的逆，那么*$\alpha$*也是*$\beta$*的逆。*

三维空间中的一个向量都可以写成

$$\alpha=ai+bj+ck$$

的形式，其中$a,\ b,\ c$分别是$x,\ y,\ z$轴的坐标。模仿上面二维的情形。我们定义两个向量$\alpha,\ \beta$的点积,

$$\alpha\cdot\beta:=\frac{1}{2}(\alpha\overline{\beta}+\beta\overline{\alpha})=-\frac{1}{2}(\alpha\beta+\beta\alpha),$$

因为$\overline{\alpha}=-\alpha,\ \overline{\beta}=-\beta$.

类似的，我们还可以定义两个向量的叉积，

$$\alpha\times\beta:=\frac{1}{2}(\alpha\beta-\beta\alpha).$$

根据定义$\alpha\times\beta=-\beta\times\alpha$.

根据定义, 我们知道如果$\alpha=x_1i+y_1j+z_1k,\ \beta=x_2i+y_2j+z_2k$, 则

$$\alpha\cdot\beta=x_1x_2+y_1y_2+z_1z_2\in\mathbb{R},$$
$$\alpha\times\beta=(y_1z_2-y_2z_1)i+(x_2z_1-x_1z_2)j+(x_1y_2-x_2y_1)k\in V.$$

假设$\alpha=xi+yj+zk$, 则$\alpha\cdot\alpha=x^2+y^2+z^2$, 我们把这个数的平方根$\sqrt{x^2+y^2+z^2}$叫做$\alpha$的长度，或者绝对值。$\alpha=x_1i+y_1j+z_1k,\ \beta=x_2i+y_2j+z_2k$，$\theta$是$\alpha,\ \beta$的夹角。我们有下面的公式：

$$\alpha\cdot\beta=|\alpha||\beta|\cos\theta$$

。所以我们有下面的定理(都可以通过把α分解成$\alpha=\alpha_1+\alpha_2$, 其中α_1和β平行，α_2和β垂直来证明)

定理 1.3.1 (两个向量的内积夹角公式)**.** *假设*θ*是*$\alpha,\ \beta$*之间的夹角。则*θ*满足*

$$\alpha\cdot\beta=|\alpha||\beta|\cos\theta.$$

定理 1.3.2 (两个向量的外积夹角公式)**.** *假设*θ*是*$\alpha,\ \beta$*之间的夹角。则*θ*满足*

$$|\alpha\times\beta|=|\alpha||\beta|\sin\theta.$$

定理 1.3.3 (拉格朗日公式)**.** 对任意的向量$\alpha,\ \beta,\ \gamma$, 有

$$\alpha\times(\beta\times\gamma)=(\alpha\cdot\gamma)\beta-(\alpha\cdot\beta)\gamma.$$

证明 注意到一个实数乘以一个向量，跟这个向量乘以这个实数是相等的。所以

$$\begin{aligned}
\text{右边}&=\frac{1}{2}(\beta(\alpha\cdot\gamma)-\gamma(\alpha\cdot\beta)+(\alpha\cdot\gamma)\beta-(\alpha\cdot\beta)\gamma)\\
&=-\frac{1}{4}(\underbrace{\beta\alpha\gamma}+\beta\gamma\alpha\underline{-\gamma\alpha\beta}-\gamma\beta\alpha+\alpha\gamma\beta\underline{+\gamma\alpha\beta}-\alpha\beta\gamma\underbrace{-\beta\alpha\gamma})\\
&=-\frac{1}{4}(\beta\gamma\alpha-\gamma\beta\alpha+\alpha\gamma\beta-\alpha\beta\gamma)\\
&=\frac{1}{2}(\alpha(\frac{1}{2}(\beta\gamma-\gamma\beta))+(\frac{1}{2}(\beta\gamma-\gamma\beta))\alpha)\\
&=\text{左边}.
\end{aligned}$$

□

定理 1.3.4 (雅可比等式)**.**

$$\alpha\times(\beta\times\gamma)+\beta\times(\gamma\times\alpha)+\gamma\times(\alpha\times\beta)=0.$$

证明 这就是要证明

$$(\alpha\beta\gamma-\alpha\gamma\beta-\beta\gamma\alpha+\gamma\beta\alpha)+(\beta\gamma\alpha-\beta\alpha\gamma-\gamma\alpha\beta+\alpha\gamma\beta)+(\gamma\alpha\beta-\gamma\beta\alpha-\alpha\beta\gamma+\beta\alpha\gamma)=0.$$

这总共有十二项。三个向量的排列只有六种，再加上正负号，正好是十二项。所以两两抵消掉了，正好是零。 □

如果记$[x,\ y]=xy-yx$, 则上面的雅可比等式就是

$$[\alpha,\ [\beta,\ \gamma]]+[\beta,\ [\gamma,\ \alpha]]+[\gamma,\ [\alpha,\ \beta]]=0.$$

这就和李代数中的雅可比等式从内容到形式完全一致了。

定理 1.3.5 (混合积公式, 我们总是先算叉后算点)**.**

$$\alpha\times\beta\cdot\gamma=\beta\times\gamma\cdot\alpha=\gamma\times\alpha\cdot\beta.$$

证明 验证完全是形式主义的, 极其容易的。我们这里只验证第一个等式。

$$\begin{aligned}
-4(\alpha\times\beta\cdot\gamma-\beta\times\gamma\cdot\alpha)&=\underline{\alpha\beta\gamma}-\beta\alpha\gamma+\gamma\alpha\beta\underbrace{-\gamma\beta\alpha}-\beta\gamma\alpha+\underbrace{\gamma\beta\alpha}\underline{-\alpha\beta\gamma}+\alpha\gamma\beta\\
&=-\beta(\alpha\gamma+\gamma\alpha)+(\alpha\gamma+\gamma\alpha)\beta\\
&=0\quad(\text{上行括号内的东西其实是实数})
\end{aligned}$$

□

定理 1.3.6 (拉格朗日恒等式)**.** *对任意的四个向量α, β, γ, δ, 有*

$$(\alpha\times\beta)\cdot(\gamma\times\delta)=(\alpha\cdot\gamma)(\beta\cdot\delta)-(\alpha\cdot\delta)(\beta\cdot\gamma)$$

证明

$$\begin{aligned}(\alpha\times\beta)\cdot(\gamma\times\delta)&=\alpha\cdot\beta\times(\gamma\times\delta)(\text{ 混合积公式})\\&=\alpha\cdot((\beta\cdot\delta)\gamma-(\beta\cdot\gamma)\delta)(\text{ 拉格朗日公式})\\&=(\alpha\cdot\gamma)(\beta\cdot\delta)-(\alpha\cdot\delta)(\beta\cdot\gamma).\end{aligned}$$

□

1928年，四元数被物理学家狄拉克在研究微分算子的时候重新发现。狄拉克由此出发，预言了正电子的存在。1932年，安德森做实验观察到了正电子的存在。1933年，狄拉克获得了诺贝尔奖，年仅31岁。

§1.4 p元域(模p的剩余类域)

假设p是一个固定的素数。我们在这一节中只考虑由集合$\{0,\ 1,\ 2,\ ...,\ p-1\}$构成的有限域$\mathbb{F}_p$。作为集合来看，$\mathbb{F}_p=\{0,\ 1,\ 2,\ ...,\ p-1\}$。但是$\mathbb{F}_p$跟单纯的集合相比，它里面多了两种运算，一种是加法，一种是乘法。任给$a,\ b\in\{0,\ 1,\ 2,\ ...,\ p-1\}$, $a,\ b$之和有可能大于等于p, 不在$\{0,\ 1,\ 2,\ ...,\ p-1\}$里面了。假设$x$是$a,\ b$之和再除以$p$的余数。在有限域$\mathbb{F}_p$中，我们定义$a+b$等于这个余数$x$。同样的任给$a,\ b\in\{0,\ 1,\ 2,\ ...,\ p-1\}$, $a,\ b$之乘积有可能大于等于p, 不在$\{0,\ 1,\ 2,\ ...,\ p-1\}$ 里面了。假设y是$a,\ b$之积再除以p的余数。在有限域$\mathbb{F}_p$中，我们定义ab等于这个余数y。

注意在有限域$\mathbb{F}_p$中，$-a=p-a$, $p=0$。任取一个非零元素$a\in\mathbb{F}_p$。因为a不能被p整除，所以由欧几里得除法可知，存在整数$x,\ y$使得$ax+py=1$. 所以在有限域$\mathbb{F}_p$中，$ax=1$. 也就是说，有限域$\mathbb{F}_p$中，任何非零元素都可逆。显然这个逆元素也是唯一确定的。我们记$x=a^{-1}$.

例 1.4.1. *例如在$\mathbb{F}_5$中，$3+4=2\in\mathbb{F}_5$, $4\cdot4=1\in\mathbb{F}_5$. $3^{-1}=2\in\mathbb{F}_5$*

有限域$\mathbb{F}_p$的所有的元素加起来等于$(p-1)p/2$. 如果$p>2$, 则$p-1$是偶数, 所以$(p-1)p/2$能够被p 整除，所以在有限域$\mathbb{F}_p$中$(p-1)p/2=0$。我们再看看有限域$\mathbb{F}_p$中所有的非零元素乘起来等于多少。

定理 1.4.1 (威尔逊定理)**.** $\displaystyle\prod_{\text{非零}x\in\mathbb{F}_p}x=1\times2\times3\times\cdots\times(p-1)=-1.$

证明 有限域$\mathbb{F}_p$中任何一个非零元素和它自己的逆相乘都等于1. 而除了1和$p-1$以外，其他的元素都不等于自身的逆。所以$1\times2\times3\times\cdots\times(p-1)=p-1=-1$。

□

定理 1.4.2 (费马小定理). 任给非零元素$a \in \mathbb{F}_p$, 则$a^{p-1} = 1$.

证明 注意到

$$\mathbb{F}_p = \{0,\ 1,\ 2,\ ...,\ p-1\} = \{0,\ a,\ 2a,\ ...,\ (p-1)a\},$$

所以

$$1 \times 2 \times 3 \times \cdots \times (p-1) = a^{p-1}(1 \times 2 \times 3 \times \cdots \times (p-1)).$$

因此$a^{p-1} = 1$. □

定义 1.4.1. 有了有限域$\mathbb{F}_p$的定义以后，我们可以考虑以$\mathbb{F}_p$中的元素为系数的多项式。这些多项式相加或者相乘的运算法则由有限域$\mathbb{F}_p$的加法和乘法运算法则决定。我们用$\mathbb{F}_p[x_1,\ ...,\ x_n]$表示以$\mathbb{F}_p$中的元素为系数的多项式全体形成的集合。

例 1.4.2. 例如在$\mathbb{F}_5[x_1,\ x_2]$中，$(3x_1 - 4x_2)(2 + x_1) = 3x_1^2 + x_1 + 2x_2 + x_1x_2$.

跟我们前面讲的类似，我们可以定义有限域的直积$\mathbb{F}_p^n = \mathbb{F}_p \times \cdots \times \mathbb{F}_p$。其中的加法和乘法分别定义成分量之间的加法和乘法。

同样的有限域的直积$\mathbb{F}_p^n = \mathbb{F}_p \times \cdots \times \mathbb{F}_p$中，所有的元素加起来也是0.

对任意的一个以有限域$\mathbb{F}_p$中的元素为系数的多项式$f(x) \in \mathbb{F}_p[x]$, 只要$f(a) = 0$, 就能推出$x - a$是$f(x)$的一个因子。从$f(x)$中除掉这个因子，$f(x)$的次数就降低了1。这样归纳下去可知，m次多项式最多只能有m个根。所以$x^k = 1$在有限域$\mathbb{F}_p$中最多只有k个根。注意到有限域$\mathbb{F}_p$中含有$p-1$个非零元素，所以如果$k < p-1$, 我们一定能找到$\mathbb{F}_p$ 中的一个非零的元素g使得$g^k \neq 1$.

实际上可以做的更好，可以证明存在一个$g \in \mathbb{F}_p$使得对任意的$k < p-1$，都有$g^k \neq 1$. 这样的g 叫做"原根"。

定义 1.4.2 (阶). 假设m为正整数，$(a, m) = 1$, 则满足$a^n \equiv 1 \pmod m$的最小的正整数n叫做a 模m的阶。

引理 1.4.3. 假设m是一个正整数，a是一个与m互素的整数。假设a模m的阶为k。如果$a^n \equiv 1 \pmod m$, 则$k|n$.

证明 否则，存在正整数α, β使得$n = k\alpha + \beta$，$\beta < k$. 由$a^n \equiv 1 \pmod m$而且$a^k \equiv 1 \pmod m$可得$a^\beta \equiv 1 \pmod m$. 这和a模m的阶为k矛盾。 □

引理 1.4.4. 假设m是一个正整数，a, b是两个与m互素的整数。假设a, b模m的阶分别为k, l。如果k, l互素，则ab 模m阶为kl.

证明 假设ab模m阶为n. 因为$(ab)^{kl} \equiv 1 \pmod m$, 所以$n|kl$. 又$(ab)^{kn} \equiv 1 \pmod m$而$a^{kn} \equiv 1 \pmod m$，所以$b^{kn} \equiv 1 \pmod m$. 因此$l|nk$. 但是$l$, k互素，所以$l|n$. 同理$k|n$. 所以$kl|n$. 所以$n = kl$. □

定理 1.4.5 (原根的存在性). 假设p是一个素数，则存在一个整数g使得g模p的阶等于$p-1$. $\Longleftrightarrow$存在$g\in\mathbb{F}_p$使得$\mathbb{F}_p=\{0,\ g,\ g^2,\ ...,\ g^{p-1}\}$. $\Longleftrightarrow$存在一个整数$g\in\mathbb{Z}$, 使得$\{g,\ g^2,\ ...,\ g^{p-1}\}$构成了模$p$的一个缩系。这样的$g$ 叫做"原根"。

证明 假设$p-1=p_1^{a_1}\cdots p_k^{a_k}$. 断言存在一个整数$g_i$使得$g_i$模$p$的阶等于$p_i^{a_i}$, $1\le i\le k$. 否则，注意到$x^{p_i^{a_i}}=1$在$\mathbb{F}_p$中恰好有$p_i^{a_i}$个解。如果这些解的阶都小于$p_i^{a_i}$，则这些解都满足方程

$$x^{p_i^{a_i-1}}=1.$$

但是我们知道上面这个方程的解最多只有$p_i^{a_i-1}$个，引起矛盾。

令$g=g_1\cdots g_k$，则g是一个"原根"。

□

§1.5 一元多项式

我们假设F是一个域, 可以是数域, 也可以是p元域. x是一个变量. 那么所有的形如

$$ax^m,\ a\in F,\ m\in\mathbb{Z}$$

形式的式子都称为是单项式. 我们称若干个单项式的和

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\ a_n\neq 0$$

为多项式. 如果$a_n=1$, 那么我们称$f(x)$是一个首一的多项式. 我们记F上的一元多项式的全体为$F[x]$, 称之为F上的一元多项式环.

在上面的表达式中的n称为是$f(x)$的次数, 记为

$$\deg f=n,$$

即f的次数就是它的表达式中单项式的最高次数. 如果$f(x)=0$, 我们称其为零多项式. 我们一般把零多项式的次数定义为$-\infty$.

定义 1.5.1 (多项式的加法和乘法). 假设$f=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $g=b_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $n\ge m$. 那么f与g的和为

$$f+g=(a_n+b_n)x^n+\cdots+(a_{n-1}+b_{n-1})x^{n-1}+\cdots+(a_1+b_1)x+(a_0+b_0),$$

这里的$b_n=b_{n-1}=\cdots=b_{m+1}=0$是为了记号的整齐对称添加的.

f与g的积为

$$fg=\sum_{k=0}^{mn}c_kx^k,$$

这里的

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

注意我们这里公式中会有诸如a_{n+1}，a_{n+2}，... 或者b_{m+1}，b_{m+1}，... 之类的没有定义的记号出现. 这些记号都是为了公式的写法统一方便而添加的, 它们都等于0.

两个多项式f与g的差如果是0, 我们就称f与g 是相等的多项式, 记为$f(x) = g(x)$. 所以两个多项式相等当且仅当它们每一个单项式的系数都相等. 我们很容易看出, 如果两个多项式f与g相等的话, 那么对任意的$a \in F$, 都有

$$f(a) = g(a)$$

成立. 由于我们本书中讨论的数域都是复数域的子集, 对这样的数域来说, 对任意的$a \in F$, 都有

$$f(a) = g(a)$$

成立. 也能够推出$f(x) = g(x)$ 成立. 对于更一般的系数域(例如p元域来说, 在每一个点处取值都相等, 并不能推出两个多项式是相等的. 两个多项式相等要求在形式上完全一样, 而不仅仅作为域上的函数相等.

定理 1.5.1 (多项式的带余除法). *假设F是一个域. $f(x)$, $g(x) \in F[x]$. 假设$\deg f \geq \deg g$, 那么一定存在多项式$q(x)$, $r(x)$使得*

$$f(x) = q(x)g(x) + r(x)$$

而且$\deg r < \deg q$.

定理 1.5.2. *多项式$f(x)$能够被$x-\alpha$整除当且仅当$f(\alpha)=0$, 即α是$f(x)$的一个根.*

证明 利用多项式的带余除法，我们知道存在多项式$g(x)$和常数a使得

$$f(x) = (x-\alpha)g(x) + a.$$

多项式$f(x)$能够被$x-\alpha$整除当且仅当$a=0$, 即$f(\alpha)=0$。 □

推论 1.5.3. *数域F上的n次多项式在F中的根数不超过n个。*

推论 1.5.4. *数域F上的两个次数不超过n的多项式如果在$n+1$个点处的取值相同，那么它们就是相同的多项式。*

定理 1.5.5 (拉格朗日差值多项式). *任给n对数(x_1, y_1), ..., (x_{n+1}, y_{n+1}). 我们假设x_1, ..., x_{n+1}互不相同。则一定存在一个唯一的n次多项式$f(x)$使得*

$$f(x_1) = y_1,\ \ldots,\ f(x_{n+1}) = y_{n+1}.$$

证明 我们首先看着这个多项式如果存在的话，那么由上面的推论可以知道，这样的多项式一定是唯一的。因此关键是证明这样的多项式的存在性。

我们先来说明一定可以构造出n次多项式f_i使得

$$f_i(x_j) = \begin{cases} 0, & i \neq j \\ y_i, & i = j. \end{cases}$$

其实只要令

$$h_i(x) = (x - x_1)(x - x_2) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_{n+1})$$

可知$h_i(x)$出了在x_i处不等于0, 在其余的n个点处都等于0. 因此我们令

$$g_i(x) = \frac{h_i(x) y_i}{h_i(x_i)}.$$

令

$$f(x) = f_1(x) + \cdots + f_{n+1}(x)$$

即可。 □

上面定理中的多项式叫做是拉格朗日差值多项式。

§1.6 多项式的因式分解定理

本节中F表示一个域, 可以是数域或者p元域.

定义 1.6.1 (整除). *如果f, g, $h \in F[x]$, 而且$f = gh$, 那么我们称g整除f.*

定义 1.6.2 (不可约多项式). *如果$f \in F[x]$, 而且而且f不能写成两个次数较低的多项式的乘积, 那么就称f是以后各不可约多项式.*

定理 1.6.1 (多项式的因式分解以及唯一性定理). *如果$f \in F[x]$是一个次数大于等于1的首一多项式, 那么f 一定可以分解成首一不可约多项式的乘积*

$$f(x) = p_1(x)^{a_1} \cdots p_t^{a_t},$$

这里的p_1, ..., p_t是互不相同的首一不可约多项式. 而且如果不考虑排列次序的话, 上面的分解表达式是唯一的.

定义 1.6.3 (重因式). *如果不可约多项式$p(x)$满足$p(x)^k | f(x)$, 但是$p(x)^{k+1} \nmid f(x)$. 就称p 是f的k重因式.*

定义 1.6.4 (导数)**.** 假设多项式$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. 我们定义$f$的导数为

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$$

定理 1.6.2. 假设$p(x)$是一个不可约多项式. 如果$p(x)$是$f(x)$的k重因式, 则p是f, f', ..., $f^{(k-1)}$的因式, 但不是$f^{(k)}$的因式.

定理 1.6.3 (爱森斯坦判别法)**.**

定义 1.6.5 (多项式的同余)**.** 假设$f_1(x)$, $f_2(x)$, $g(x)$, $h(x)$是四个多项式, 满足关系式

$$f_1(x) - f_2(x) = g(x)h(x).$$

我们称$f_1(x)$, $f_2(x)$模$g(x)$同余, 记为

$$f_1(x) \equiv f_2(x)(\mod g(x)).$$

定理 1.6.4 (多项式的中国剩余定理)**.** 假设$f_1(x)$, ..., $f_k(x)$是两两互素的多项式, 则对于任意的多项式g_1, ..., g_k, 同余方程

$$\begin{cases} f \equiv g_1(\mod f_1), \\ \vdots \\ f \equiv g_k(\mod f_k). \end{cases}$$

有模$N = f_1 \cdots f_k$惟一的解。

推论 1.6.5 (拉格朗日差值公式)**.** 假设a_1, ..., a_k是两两不同的数, 则对于任意k个数b_1, ..., b_k, 存在唯一的一个次数小于等于$k-1$的多项式f 使得

$$f(a_1) = b_1, \ ..., \ f(a_k) = b_k.$$

证明 考虑同余方程组

$$\begin{cases} f \equiv b_1(\mod (x - a_1)), \\ \vdots \\ f \equiv b_k(\mod (x - a_k)). \end{cases}$$

这个同余方程组的解, 多项式f模$(x-a_1)\cdots(x-a_k)$惟一。而模$(x-a_1)\cdots(x-a_k)$的余项是个次数小于等于$k-1$的多项式. □

类似于秦九韶的口诀, 这个多项式可以具体的构造出来. 对任意的$i = 1, 2, ..., k$, 我们需要找$g_i(x)$使得它满足

$$\begin{cases} g_i(x) \equiv 0(\mod (x-a_1)), \\ \vdots \\ g_i(x) \equiv 0(\mod (x-a_{i-1})), \\ g_i(x) \equiv 1(\mod (x-a_i)), \\ g_i(x) \equiv 0(\mod (x-a_{i+1})), \\ \vdots \\ g_i(x) \equiv 0(\mod (x-a_k)). \end{cases}$$

然后令

$$f = b_1 g_1(x) + \cdots + b_k g_k(x)$$

即可. 这样的$g_i(x)$很容易构造, 首先它有因式

$$(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_k).$$

其次这个多项式在a_i处取值为b_i, 所以我们可以假设

$$g_i(x) = \alpha(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_k),$$

其中α是个待定的常数. 易见这个常数为

$$\alpha = \frac{(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_k)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_k)}.$$

综上可知, 满足$f(a_1) = b_1, \ ..., \ f(a_k) = b_k$的多项式

$$f = \sum_{i=1}^{n} \frac{(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_k)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_k)} + (x-a_1)\cdots(x-a_k)h(x),$$

其中$h(x)$为任意的多项式.

§1.7 Mason-Stothers 定理

对任意的多项式F, 我们用$n_0(F)$表示F的互不相同的根的个数, 即如果有重根的话, 则不考虑其重数, 都只记一次.

定理 1.7.1. 假设$f, \ g, \ h$是3个互素的多项式,而且$f + g = h$. 那么

$$\mathrm{Max}\{\deg(f), \ \deg(g), \ \deg(h))\} \le n_0(fgh) - 1.$$

证明 令$F=f/h,\ G=g/h$, 则$F+G=1$. 两边求导可得$F'+G'=0$. 所以

$$\frac{F'}{F}F+\frac{G'}{G}G=0,$$

即

$$\frac{g}{f}=\frac{G}{F}=-\frac{F'/F}{G'/G}.$$

假设

$$f(x)=c_1\prod(x-\alpha_i)^{a_i},\ g(x)=c_2\prod(x-\beta_j)^{b_i},\ h(x)=c_3\prod(x-\gamma_k)^{c_i},$$

则

$$\frac{g}{f}=-\frac{F'/F}{G'/G}=-\frac{\sum\frac{a_i}{x-\alpha_i}-\prod\frac{c_k}{x-\gamma_k}}{\sum\frac{b_j}{x-\beta_j}-\prod\frac{c_k}{x-\gamma_k}}.$$

因为$n_0(fgh)=\prod(x-\alpha_i)\prod(x-\beta_j)\prod(x-\gamma_k)$, 所以

$$F_1=n_0(fgh)F'/F,\ G_1=n_0(fgh)G'/G,$$

都是多项式. 又因为

$$\frac{g}{f}=-\frac{F_1}{G_1},$$

所以$f,\ g$的次数都要小于等于$n_0(fgh)-1$. 所以

$$\mathrm{Max}\{\deg(f),\ \deg(g),\ \deg(h))\}\le n_0(fgh)-1.$$

□

推论 1.7.2. *如果$n\ge 3$, 那么不存在互素的多项式$f,\ g,\ h$ 使得*

$$f^n+g^n=h^n.$$

注意如果$n=2$. 这样的多项式是存在的. 例如

$$(k^2-1)^2+(2k)^2=(k^2+1)^2.$$

上面的这个等式, 当k跑遍所有的有理数的时候,其实可以据此给出所有的勾股数组的通解.

§1.8 多元多项式

定理 1.8.1 (唯一因式分解定理)**.**

定义 1.8.1 (初等对称多项式)**.**

定理 1.8.2 (对称多项式是初等对称多项式的多项式)**.**

§1.9 形式幂级数

习题 1.9

1. 一个系数都是有理数的多项式在所有的整数处都取值为整数。那么这个多项式是不是整系数的?

2. 当已知$f(x)$是一个系数为有理数的多项式，而且对任意大于等于10000的正整数n, $f(n)$都是整数。求证对任意整数n, $f(n)$都是整数。

3. 假设a_1, ..., a_n是n个不同的整数. 求证, 多项式

$$f(x) = (x - a_1)(x - a_2) \cdots (x - a_n) - 1$$

在$\mathbb{Q}$上不可约.

证明 假设$f(x)$在$\mathbb{Q}$上可约, 那么它一定在$\mathbb{Z}$上可约. 因此, 存在两个次数都比f低的整系数多项式$g(x)$和$h(x)$使得

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

由于$f(a_i) = g(a_i)h(a_i) = -1$, 所以$|g(a_i)| = |h(a_i)| = 1$ 而且$g(a_i) + h(a_i) = 0$. 所以多项式

$$g(x) + h(x)$$

至少有n个互不相同的根. 可是因为$g(x)+h(x)$的次数小于n, 所以$g(x)+h(x)$必须是零多项式. 因此$f(x) = -g(x)^2$. 这和$f(x)$的首相系数为1 矛盾. 因此$f(x)$在$\mathbb{Q}$上不可约.

□

4. 假设a_1, ..., a_n是$n \geq 5$个不同的整数. 求证, 多项式

$$f(x) = (x - a_1)(x - a_2) \cdots (x - a_n) + 1$$

在$\mathbb{Q}$上不可约.

证明 假设$f(x)$在$\mathbb{Q}$上可约, 那么它一定在$\mathbb{Z}$上可约. 因此, 存在两个次数都比f低的整系数多项式$g(x)$和$h(x)$使得

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

由于$f(a_i)=g(a_i)h(a_i)=1$, 所以$|g(a_i)|=|h(a_i)|=1$ 而且$g(a_i)-h(a_i)=0$. 所以多项式

$$g(x)-h(x)$$

至少有n个互不相同的根. 可是因为$g(x)-h(x)$的次数小于n, 所以$g(x)-h(x)$必须是零多项式. 因此$f(x)=g(x)^2$. 所以

$$(x-a_1)(x-a_2)\cdots(x-a_n)=g(x)^2-1=(g(x)+1)(g(x)-1).$$

所以n一定是偶数$n=2k$, 我们不妨设

$$g(x)+1=(x-a_1)(x-a_2)\cdots(x-a_k),\ a_1<a_2<\cdots<a_k$$
$$g(x)-1=(x-a_{k+1})(x-a_{k+2})\cdots(x-a_n),\ a_{k+1}<a_{k+2}<\cdots<a_n.$$

则$-2=g(a_1)-1=(a_1-a_{k+1})(a_1-a_{k+2})\cdots(a_1-a_n)$. 左边这几个数只能是2, 1, -1. 所以

$$k=3, a_1-a_4=2,\ a_1-a_5=1,\ a_1-a_6=-1.$$

同理还有

$$a_2-a_4=2,\ a_2-a_5=1,\ a_2-a_6=-1.$$

这是不可能的. 所以$f(x)$在$\mathbb{Q}$上不可约.

□

注意在这个题目中的$n\geq 5$的条件是必不可少的. 如果$n=4$, 那么命题是不成立的. 反例如下:

$$x(x-1)(x-2)(x-3)+1=x(x-3)\cdot(x-1)(x-2)+1=(x^2-3x+1)^2.$$

5. 假设n是正整数, 多项式

$$f(x)=1+x+\frac{x(x+1)}{2!}+\frac{x(x+1)(x+2)}{3!}+\cdots+\frac{x(x+1)\cdots(x+n)}{(n+1)!}.$$

(1) 求$f(x)=0$的所有的有理数根.

(2) 求证$f(x)-\frac{1}{(n+1)!}$在有理数域上不可约.

解 注意到

$$1+x+\frac{x(x+1)}{2!}=\frac{(x+2)(x+1)}{2!},$$
$$1+x+\frac{x(x+1)}{2!}+\frac{x(x+1)(x+2)}{3!}=\frac{(x+1)(x+2)(x+3)}{3!}.$$

可以猜测

$$f(x) = \frac{(x+1)\cdots(x+n)(x+n+1)}{(n+1)!}.$$

利用数学归纳法, 很容易证明上面的式子成立. 因此$f(x)=0$的所有的有理数根就是-1, -2, ..., $-n-1$.

如果$f(x)-\frac{1}{(n+1)!}$在有理数域上可约, 那么$(x+1)(x+2)\cdots(x+n)(x+n+1)-1$在有理数域上可约. 我们在前面的习题里已经证明过了这是不可能的.

□

6. 假设a_1, ..., a_n是n个不同的整数. 求证, 多项式

$$f(x) = (x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2+1$$

在$\mathbb{Q}$上不可约.

证明 假设$f(x)$在$\mathbb{Q}$上可约, 那么它一定在$\mathbb{Z}$上可约. 因此, 存在两个次数都比f低的首一的整系数多项式$g(x)$ 和$h(x)$ 使得

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

注意我们假设了$g(x)$和$h(x)$ 是首一的整系数多项式. 所以当x是非常大的实数的时候, $g(x)$和$h(x)$取值肯定都是正数. 而又由于$f(x)$在实数上的取值永远都是正数. 所以$g(x)$和$h(x)$在实数上的取值也一定都是正数.

我们不妨假设g的次数大于等于h的次数. 由于$f(a_i)=g(a_i)h(a_i)=1$, 所以$|g(a_i)|=|h(a_i)|=1$ 而且$g(a_i)-h(a_i)=0$. 所以多项式

$$g(x)-h(x)$$

至少有n个互不相同的根. 这时候可能有两种情况. 第一种是$g(x)-h(x)$是零多项式. 第二种是$g(x)-h(x)$的次数大于n, 且有因子$(x-a_1)(x-a_2)\cdots(x-a_n)$. 注意由于$g$, h都是首一的, 因此它们不可能都是n次的. 所以我们就排除了$g(x)-h(x)$ 的次数等于n 的情形.

在第一种情形, 我们有$f(x)=g(x)^2$. 所以

$$(x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2 = g(x)^2-1 = (g(x)+1)(g(x)-1).$$

所以n一定是偶数$n=2k$. 注意到$g(x)+1$和$g(x)-1$是有理数域上的两个互素的多项式(因为它们能组合出1 来), 所以$g(x)+1$和$g(x)-1$没有公共根. 我们不妨设

$$g(x)+1 = (x-a_1)^2(x-a_2)^2\cdots(x-a_k)^2,\ a_1<a_2<\cdots<a_k$$
$$g(x)-1 = (x-a_{k+1})^2(x-a_{k+2})^2\cdots(x-a_n)^2,\ a_{k+1}<a_{k+2}<\cdots<a_n.$$

则$-2 = g(a_1) - 1 = (a_1 - a_{k+1})^2(a_1 - a_{k+2})^2 \cdots (a_1 - a_n)^2$ 是不可能的. 所以$f(x)$在$\mathbb{Q}$上不可约.

在第二种情形. 我们假设$g(x)$的次数等于$n+t$, $h(x)$的次数等于$n-t$. 那么由于

$$1 = f(a_i) = (g(a_i) - h(a_i) + h(a_i))h(a_i) = h(a_i)^2$$

可知$h(a_i)^2 = 1$ 对任意的$1 \le i \le n$ 成立. 我们在第一段已经说过了, $h(x)$的取值永远是正数. 所以$h(a_i) = 1$ 对任意的$1 \le i \le n$ 成立. 这和h的次数小于n是矛盾的.

□

7. 假设n, m 都是正整数. 求x^n+1和x^m+1的最大公因式.

解 我们知道

$$x^n + 1 = \prod_{i=1}^{n}(x - \zeta_{2n}^{2i-1}),\ x^m + 1 = \prod_{j=1}^{m}(x - \zeta_{2m}^{2j-1}).$$

这里我们使用了记号$\zeta_k = \cos\frac{2\pi}{k} + i\sin\frac{2\pi}{k}$. 假设$(n,\ m) = d$, $n' = n/d$, $m' = m/d$. 由于

$$\zeta_{2n}^{2i-1} = \zeta_{2m}^{2j-1} \iff \frac{2i-1}{2n} = \frac{2j-1}{2m} \iff (2j-1)n' = (2i-1)m', \qquad (*)$$

所以如果m' 或者n'是偶数, 则不存在i, j使得$(*)$成立. 所以x^n+1和x^m+1的最大公因式为1. 注意, 由于m' 和n' 互素, 所以它们不可能都是偶数.

如果m' 和n' 都是奇数, 那么由于

$$x^n + 1 = y^{n'} + 1,\ x^m + 1 = y^{m'} + 1$$

可知道x^n+1和x^m+1有公因式$y+1$, 即x^d+1. 又因为m' 和n' 互素, 所以$y^{n'}+1 = 0$ 和$y^{m'}+1 = 0$只有一个公共根$y = 1$. 所以$x^n+1 = (x^d)^{n'}$和$x^m+1 = (x^d)^{m'}$只有d个公共根. 所以

□

8. 假设a_1, ..., a_n是n个不同的整数. 求证, 多项式

$$f(x) = (x-a_1)^4(x-a_2)^4 \cdots (x-a_n)^4 + 1$$

在$\mathbb{Q}$上不可约.

证明 太复杂, 见数学分析中的问题和定理第二册414页.

□

9. 已知f是一个多项式, 而且$f(x^2+1) = f(x)^2+1$, $f(0) = 0$. 求$f(x)$. **解** 因为$f(0)=0$, 所以

$$f(1) = f(0^2+1) = f(0)^2+1 = 1.$$

同理可得$f(2)=2$, $f(5)=5$, $f(26)=26$, ... 所以$f(x)-x$在无限多个整数处取值为零. 所以$f(x)=x$.

□

10. 假设整系数多项式$f(x) = a_nx^n+\cdots+a_1x+a_0$, p是一个素数, 而且$p\nmid a_n$, ..., $p\nmid a_{n-k}$, $p|a_{n-k-1}$, ..., $p|a_0$, $p^2\nmid a_n$. 求证$f(x)$具有次数大于等于$n-k$的整系数不可约因式.

证明 假设$f(x)$分解成整系数不可约多项式的乘积为$f(x) = f_1\cdots f_k$,没有次数大于等于$n-k$的整系数不可约因式. 那么由于$p^2\nmid a_n$, 所以只有一个f_i的常数项是p的倍数, 而且这个f_i 的次数是小于$n-k$的. 假设

$$f_i = b_mx^m+\cdots+b_tx^t+p(b_{t-1}x^{t-1}+\cdots+b_0),$$

其中b_m, ..., b_t都不是p的倍数. ‘那么由于$f(x) = f_1\cdots f_k$, 而除了f_i之外, 其余的常数项都不是p 的倍数. 所以$f_1\cdots f_k$ 中x^t 的系数不是p的倍数. 由于$t\le m<n-k$. 这和题设的$p|a_{n-k-1}$, ..., $p|a_0$, $p^2\nmid a_n$矛盾.

□

§1.10 课本上的习题解答

22. 证明x_0是$f(x)$的k重根的充分必要条件是$f(x_0) = f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0$, 而$f^{(k)}(x_0)\neq 0$.

证明 假设$g(x) = f(x+x_0)$. 则问题就变为了0是$g(x)$的k重根的充分必要条件是$g(0) = g'(0) = \cdots = g^{(k-1)}(0) = 0$, 而$g^{(k)}(0)\neq 0$. 所以我们不妨在一开始就假设$x_0=0$.

x_0是$f(x)$的k重根的充分必要条件是$f(x)$的次数最低的项是x^k项. 显然$f(x)$的次数最低的项是x^k项充分必要条件是$f(x_0) = f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0$, 而$f^{(k)}(x_0)\neq 0$. □

23. 0是x的一重根, 但是0不是$\frac{1}{2}x^2+1$的2重根.

24. 如果$(x-1)|f(x^n)$, 那么$(x^n-1)|f(x^n)$.

证明

假设$f(x) = \prod(x-\alpha_i)$. 那么$f(x^n) = \prod(x^n-\alpha_i)$. 如果$(x-1)|f(x^n)$, 那么$f(1) = \prod(1-\alpha_i) = 0$, 所以某个$\alpha_i = 1$. 所以$f(x^n) = \prod(x^n-\alpha_i)$ 有因子(x^n-1).
□

25. 如果$(x^2+x+1)|f_1(x^3)+xf_2(x)$, 那么

$$(x-1)|f_1(x),\ (x-1)|f_2(x).$$

证明 假设$\zeta=\frac{-1+\sqrt{3}i}{2}$. 则由$(x^2+x+1)|f_1(x^3)+xf_2(x)$可得

$$f_1(1)+\zeta f_2(1)=0$$
$$f_1(1)+\zeta^2 f_2(1)=0$$

解方程可知$f_1(1)=f_2(1)=0$. 所以$(x-1)|f_1(x),\ (x-1)|f_2(x)$.. □

§1.11 课本上的补充习题解答

9. 证明$f(x)=x^n+ax^{n-m}+b$不能有不为零的重数大于2 的根.

证明 如果c是一个不为零的重数大于2 的根, 那么$(x-c)|f(x)$,和f', f''. 这个结论可以看前面的习题22. 所以我们得到

$$c^n+ac^{n-m}+b=0$$
$$nc^{n-1}+a(n-m)c^{n-m-1}=0$$
$$n(n-1)c^{n-2}+a(n-m)(n-m-1)c^{n-m-2}=0$$

把a, b看成是未知数, 容易看出后面的两个方程式互相矛盾的. 所以没有解. □

10 如果$f(x)|f(x^n)$, 那么$f(x)=0$的根只能是0或者是单位根.

证明 假设$f(x)=\prod(x-a_i)$. 则$f(x^n)=\prod(x^n-a_i)=\prod_i(\prod_{j=1}^n(x-\zeta_n^j a_i^{1/n}))$. 如果$f(x)|f(x^n)$, 那么对每个$i$, 都有$k$, j使得

$$a_i=\zeta_n^j a_k^{1/n}.$$

所以$a_i^n=a_k$. 所以作为集合来看, 下面这两个有限集合相等.

$$\{a_1,\ ...\}=\{a_1^n,\ ...\}$$

再把每一个元素取n次方, 它们当然没也还相等. 如果某个a_i不是0或者是单位根, 那个a_i, a_i^n, $a_i^{n^2}$, ...就会是一个无穷序列, 引起矛盾. □

11. 如果$f(x)$是一个n次多项式, 而且$f'|f$. 求证$f(x)=0$有n重根.

证明 因为$f'|f$, 所以我们假设$f=f'(x-a)$. 则两边求导可知$f^{(2)}(x)(x-a)=0$, 所以$f^{(2)}=0$. 所以$f(x)$只能是一次多项式. 所以命题自然成立. □

14. 如果$f(x)$是一个整系数多项式, 而且$f(0)$, $f(1)$都是奇数, 求证$f(x)$不可能有整数根.

证明 我们用反证法. 假设$f(a)=0$. 则因为$f(0)$是奇数, 所以$f(x)$的常数项是奇数. 所以a不可能是偶数. 所以a一定是奇数. 假设$f(x)=\sum a_ix^i$. 则

$$f(a)-f(1)=\sum a_i(a^i-1).$$

注意上面的等式右边的括号内的每一项都是偶数. 所以右边是个偶数, 因此左边也是偶数.这样的话, $f(a)$必须跟$f(1)$一样都是奇数, 自然也就不可能是0了. $\square$

§1.12 难题

1. 第一个系数不是正负1的分圆多项式是$n=105$对应的分圆多项式, 这是一个48次的多项式.

2. 两个整系数不可约多项式相乘, 得到的乘积多项式的项数反而变少了.

   这里肯定要求是整系数不可约的不可约多项式. 因为

   $$(x^4+4)=(x^2+2-2x)(x^2+2+2x).$$

3. 假设$f_n(x)=\frac{x^{n-2}(x^3-x-1)+(x^3+x^2-1)}{x-1}$. 记$\zeta_k$为$k$次本原单位根. 求证如果$f_n(\zeta_k)=0$, 则$k\le 180$. **解** 见Gross "Cyclotomic factors of Coxeter polynomials" $\square$

4. 假设

   $$P=a_0a_1\cdots a_n,\ 0\le a_i\le 9,\ a_0\ne 0,$$

   是一个用十进制表示的素数. 求证多项式

   $$a_0x^m+a_1x^{m-1}+\cdots+a_m$$

   在$\mathbb{Q}$上不可约.

第二章　矩阵和行列式

矩阵的思想萌芽很早, 但是成形却非常晚, 至今只有一百多年的历史. 高斯关于二次二次型的合成律的研究里已经把两个线性方程组的系数像矩阵那样进行运算了, 矩阵似乎已经呼之欲出了. 后来德国天才数学家爱森斯坦在1842年19岁的时候看到了高斯的名著《算术研究》, 就有了矩阵的想法, 1843年, 他到爱尔兰都柏林拜访了哈密顿, 注意到1843年就是哈密顿发现四元数的年份. 哈密顿非常欣赏爱森斯坦, 给了他一本书, 是阿贝尔关于解代数方程的著名论文. 高斯和阿贝尔的著作激发了爱森斯坦的研究兴趣, 他在1844年写到两个线性方程组的复合, 可以从这个线性方程组的系数方程的运算得到. 这里面只涉及到纯粹的数字运算, 不涉及任何的变量在里面. 但是他并没有写成现在的矩阵的形式. 爱森斯坦的命运和阿贝尔非常像, 也是一生极度穷困, 性格高傲. 他表露的这些特征引起了当时的很有名的善于发现年轻数学天才的克列尔, 洪堡等人的警觉. 这些人都知道阿贝尔和伽罗瓦的悲剧, 非常担心同样的悲剧在爱森斯坦身上重演, 以至于克列尔常常要求爱森斯坦把还没有完成的数学手稿寄给他发表, 以至于爱森斯坦的一些论文会让觉得不完整. 爱森斯坦二十多就死于肺结核, 和阿贝尔死因相同.

现在这种形式的矩阵理论是西尔维斯特在十九世纪的就是年代才发展起来的.

§2.1　矩阵和初等变换

定义 2.1.1. 数域F上的一个$m\times n$的矩阵A是指由mn个数a_{ij}, $i=1$, ..., m; $j=1$, 2, ..., n 按照下面的方式写成的数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{n2} & \cdots & a_{mn} \end{pmatrix}$$

其中的a_{ij}称为是矩阵的$(i,\ j)$元素, 有时候就简单的称其为元素, i 称为是元素a_{ij}的行指标, j 称为是元素a_{ij}的列指标. m称为是矩阵A的行数, n称为是矩阵A的列数. 我们把$m\times n$叫做是矩阵的尺寸. 两个矩阵相等当且仅当它们的尺寸相等而且每个位置的元素都相等. 即如果$A=(a_{ij})_{m\times n}$, $B=(b_{kl})_{s\times t}$, 那么$A=B$当且仅当

$$m=k,\ n=l,\ a_{ij}=b_{ij},\ 1\le i\le m,\ 1\le j\le n.$$

我们把矩阵的每一行都称为是它的一个n 维行向量, 每一列称为是它的一个m为维列向量. 为了能清楚的区分行向量的不同位置的元素, 我们经常用逗号加以分

隔. 例如矩阵的第i个行向量经常表示成

$$\alpha_i = (a_{i1},\ a_{i2},\ \ldots,\ \ a_{in}).$$

而矩阵的第j个列向量经常表示成

$$\beta_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

的形式.

如果$m = n$, 我们简称A为数域F上的一个n阶方阵. 如果a_{ij}全部都是零, 我们就称A是零矩阵. 对于零矩阵, 我们经常简单的用0来表示. 要根据上下文, 看出这个零矩阵的尺寸. 不同尺寸的零矩阵是不相等的. 例如

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

如果

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

我们就称A是n阶单位矩阵, 我们经常用I_n表示单位矩阵. 在n比较清楚的情况下, 我们经常省略, 用I代表单位矩阵.

我们经常把A简记为$A = (a_{ij})$. 在相同尺寸的两个矩阵$A = (a_{ij})$, $B = b_{ij}$之间, 我们可以定义矩阵的加法:

$$A + B = (a_{ij} + b_{ij}).$$

上面的这个等式的意思就是说矩阵A+B的$(i,\ j)$位置的元素就是a_{ij}+b_{ij}. 域F中的元素和矩阵之间还有数乘运算, 对任意的数$k \in F$和矩阵$A = (a_{ij})$, 我们定义$tA = (ta_{ij})$.

定义 2.1.2 (矩阵的乘法)**.** 如果A是一个$m \times n$矩阵, 而B是一个$n \times r$矩阵. 那么A和B的乘积是一个按照如下的方式定义的一个$m \times r$阶的矩阵

$$AB = (c_{ij}),\ c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

可以直接验证矩阵的乘法满足结合律(留做习题), 即如果A是一个$m \times n$矩阵, B是一个$n \times r$矩阵, C是一个$r \times s$的矩阵, 那么

$$(AB)C = A(BC).$$

我们可以对矩阵做一下的三种初等变换:

1. 交换矩阵的两行(列).

2. 把矩阵的某一行(列)乘上一个非零的常数t.

3. 把矩阵的某一行(列)乘上一个常数t加到另外一行(列)上去.

第一种初等变换可以由后两种初等变换复合得到, 交换矩阵的第i行和第j行可以由下面的四步复合得到.

1. 把矩阵的第i行加到第j行上去.

2. 再把矩阵的第j行乘上一个-1加到第i行上去.

3. 再把矩阵的第i行加到第j行上去.

4. 最后把矩阵的第i行乘上-1.

假设I_n是n 阶单位矩阵

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

如果某个n 阶矩阵能够由n 阶单位矩阵经由一次初等变换得到. 我们称其为初等矩阵. 根据定义初等矩阵有三类, 分别对应于三类初等变换. 我们把n阶单位矩阵交换第i行和第j行得到的矩阵记为是P_{ij}. 如果记

$$e_i = (0,\ ...,\ 0,\ 1,\ 0,\ ...,\ 0)$$

为除了第i个位置为1, 其余的位置都为0的向量, 那么P_{ij}相当于单位矩阵做下面的

变换,

$$I_n = \begin{pmatrix} e_1 \\ \vdots \\ e_i \\ \vdots \\ e_j \\ \vdots \\ e_n \end{pmatrix} \longrightarrow P_{ij} = \begin{pmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i \\ \vdots \\ e_n \end{pmatrix}$$

我们把n阶单位矩阵的第i行乘上常数t得到的矩阵记为是$S_i(t)$, 那么$S_i(t)$相当于单位矩阵做下面的变换,

$$I_n = \begin{pmatrix} e_1 \\ \vdots \\ e_i \\ \vdots \\ e_n \end{pmatrix} \longrightarrow S_i(t) = \begin{pmatrix} e_1 \\ \vdots \\ te_i \\ \vdots \\ e_n \end{pmatrix}$$

我们把n阶单位矩阵的第i行乘上常数t加到第j行得到的矩阵记为是$E_{ij}(t)$, 那么$E_{ij}(t)$相当于单位矩阵做下面的变换,

$$I_n = \begin{pmatrix} e_1 \\ \vdots \\ e_i \\ \vdots \\ e_j \\ \vdots \\ e_n \end{pmatrix} \longrightarrow E_{ij}(t) = \begin{pmatrix} e_1 \\ \vdots \\ e_i \\ \vdots \\ te_i + e_j \\ \vdots \\ e_n \end{pmatrix}$$

注意n阶单位矩阵交换第i列和第j列得到的矩阵也是P_{ij}. 把n阶单位矩阵的第i列乘上常数t得到的矩阵也是$S_i(t)$. 把n 阶单位矩阵的第j 行乘上常数t加到第i行得到的矩阵记也是$E_{ij}(t)$.

命题 2.1.1. 矩阵

$$A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

可以写成$E_{ij}(t)$型的矩阵的乘积.

证明 这其实只要说明只经过第三种初等变换就可以把A变成单位矩阵即可. 首先把A的第一行乘上a^{-1}加到第二行上去, 得到

$$A=\begin{pmatrix} a & 0 \\ 1 & a^{-1} \end{pmatrix}.$$

再把第二行乘上$-a$加到第一行上去, 得到

$$\begin{pmatrix} 0 & -1 \\ 1 & a^{-1} \end{pmatrix}.$$

再把第一行乘上a^{-1}加到第二行上来, 得到

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

再把第一行乘上-1记到第二行, 得到

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

再把第二行加第一行, 得到.

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

再把第一行乘-1加第二行即可得到单位矩阵. □

§2.2 分块矩阵

我们在做关于矩阵的运算的时候, 常常会将矩阵进行适当的分块. 例如我们知道

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}=\begin{pmatrix} A_1B_1+A_2B_3 & A_1B_2+A_2B_4 \\ A_3B_1+A_4B_3 & A_3B_2+A_4B_4. \end{pmatrix}$$

如果这里的A_i, B_i是大小适当的矩阵的话, 上面的乘法依然是有意义的. 例如如果A_i, B_i的大小如下

$$A_1:\ m_1\times n_1,\ B_1:\ n_1\times k_1$$

$$A_2:\ m_1\times n_2,\ B_3:\ n_2\times k_1$$

$$A_3:\ m_2\times n_1,\ B_2:\ n_1\times k_2$$

$$A_4:\ m_2\times n_2,\ B_4:\ n_2\times k_2$$

§2.3 行列式

二阶矩阵的行列式定义为

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

其绝对值等于由向量$(a,\ b), (c,\ d)$张成的平行四边形的面积.

三阶矩阵的行列式定义为

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh.$$

其绝对值等于由向量$(a,\ b,\ c)$, $(d,\ e,\ f)$, $(g,\ h,\ i)$张成的平行六面体的体积.

我们下面来定义一般的行列式. 我们用$F^{n\times n}$表示F上的所有的$n\times n$的矩阵的全体.

定义 2.3.1 (排列)**.**

定义 2.3.2 (排列的逆序数)**.**

定义 2.3.3 (奇排列和偶排列)**.**

定义 2.3.4 (行列式)**.** $A=(a_{ij})$的行列式等于

$$|A| = \sum_{j_1, j_2, \dots,\ j_n} (-1)^{\delta(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

求和号中$j_1 j_2 \cdots j_n$跑遍1, 2, ..., n的所有排列,

$$\delta(j_1 j_2 \cdots j_n) = \begin{cases} 1, & j_1 j_2 \cdots j_n\text{是一个偶排列}, \\ -1, & j_1 j_2 \cdots j_n\text{是一个奇排列}, \end{cases}$$

定义 2.3.5 (余子式和代数余子式)**.** n阶行列式中$|A| = |(a_{ij})|$划掉第i行和第j列之后剩下的$n{-}1$阶行列式称为是a_{ij}的余子式, 通常记为D_{ij}. 我们称$(-1)^{i+j}D_{ij}$为a_{ij}的代数余子式, 通常记为A_{ij}.

定理 2.3.1 (拉普拉斯定理, 按照一行的展开)**.** 记n阶行列式中$|A| = |(a_{ij})|$的a_{ij}的代数余子式为A_{ij}. 则我们有下面的等式

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \begin{cases} |A|, & i = k; \\ 0, & i \neq k, \end{cases}$$

$$\sum_{i=1}^{n} a_{ij} A_{ik} = \begin{cases} |A|, & j = k; \\ 0, & j \neq k. \end{cases}$$

定理 2.3.2 (拉普拉斯定理, 按照一列的展开)**.** 记n阶行列式中$|A| = |(a_{ij})|$的a_{ij}的代数余子式为A_{ij}. 则我们有下面的等式

$$\sum_{i=1}^{n} a_{ij}A_{ik} = \begin{cases} |A|, & j = k; \\ 0, & j \neq k, \end{cases}$$
$$\sum_{i=1}^{n} a_{ij}A_{ik} = \begin{cases} |A|, & j = k; \\ 0, & j \neq k. \end{cases}$$

记n阶矩阵$A = (a_{ij})$的伴随矩阵为

$$A^* = \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}.$$

则上面的拉普拉斯定理也可以写成下面的形式.

$$AA^* = A^*A = \begin{pmatrix} |A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A| \end{pmatrix} = |A|I_n,$$

其中I_n是n阶单位矩阵.

从伴随矩阵的定义可知, 如果A是奇数阶的矩阵, 那么A和$-A$的伴随矩阵是相同的.

定理 2.3.3 (拉普拉斯定理, 按照若干行的展开)**.**

引理 2.3.4. 假设A, B是数域F 上的两个n阶矩阵, 那么

$$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = |A||B|.$$

证明

□

推论 2.3.5. 准三角矩阵的行列式等于对角线上的行列式的乘积.

定理 2.3.6.

$$|AB| = |A||B|$$

定理 2.3.7. 假设$A = (a_{ij})$是一个n阶矩阵, A^*是其伴随矩阵. 那么$|A^*| = |A|^{n-1}$.

证明 在等式$AA^*=A^*A=|A|I_n$的左右两边取行列式可以得到

$$|A||A^*|=|A|^n.$$

如果$|A|\neq 0$. 那么两边消掉一个$|A|$, 就可以得到$|A^*|=|A|^{n-1}$.

如果$|A|=0$, 那么可以利用微小扰动法来证明命题成立.

假设$A_\lambda=A+\lambda I_n$, 则对无限多个$\lambda\in F$ 都有

$$|A_\lambda|\neq 0.$$

所以对无限多个$\lambda\in F$ 都有

$$|A_\lambda^*|=|A_\lambda|^{n-1}.$$

注意上面的等式两边都是λ的多项式, 它们在无限多个地方取值相等. 所以这两个多项式就是相同的多项式. 也就是说这两个多项式的每一个系数都相等, 所以它们在所有的地方都取相同的值. 特别的, 它们在$\lambda=0$的地方也取值相同. 所以在$|A_\lambda^*|=|A_\lambda|^{n-1}$的左右两边令$\lambda=0$就可证出命题成立.

□

定理 2.3.8 (Binet-Cauchy 定理)**.** *假设A是一个$m\times n$的矩阵, B是一个$n\times m$的矩阵, 那么*

$$|AB|=\begin{cases}0, & m>n;\\ \sum_{1\le j_1<\cdots<j_m\le n}A\binom{12\cdots m}{j_1j_2\cdots j_m}B\binom{j_1j_2\cdots j_m}{12\cdots m}, & m\le n\end{cases}$$

这里的$A\binom{12\cdots m}{j_1j_2\cdots j_m}$是$A$中由第1, 2, ..., m行和第j_1, j_2, ..., j_2m 列组成的子式, $B\binom{j_1j_2\cdots j_m}{12\cdots m}$是$B$中由第$j_1$, j_2, ..., j_m行和第1, 2, ..., m列组成的子式.

证明 当$m>n$时,

$$AB=\begin{pmatrix}A_{m\times n} & 0_{m\times(m-n)}\end{pmatrix}\begin{pmatrix}B_{n\times m}\\ 0_{(m-n)\times m}\end{pmatrix}.$$

所以AB等于有零行零列的两个方阵的乘积, 因此$|AB|=0$.

考虑下面的行列式

$$\Delta=\begin{vmatrix}0 & A\\ B & I\end{vmatrix}$$

利用第三种初等变换不改变行列式的性质可知,

$$\begin{vmatrix}0 & A\\ B & I\end{vmatrix}=\begin{vmatrix}-AB & A\\ 0 & I\end{vmatrix}=(-1)^m|AB|.$$

我们记$C(j_1j_2\cdots j_m)$为把单位矩阵I_n中划掉第j_1，j_2，...，j_m列后剩下来的$n\times(n-m)$ 阶子矩阵. 把行列式Δ 按照前m行展开, 可知

$$\Delta=\sum_{1\le j_1<\cdots<j_m\le n}A\begin{pmatrix}12\cdots m\\ j_1j_2\cdots j_m\end{pmatrix}\begin{vmatrix}B & C(j_1j_2\cdots j_m)\end{vmatrix}(-1)^{1+2+\cdots+m+(m+j_1)+\cdots+(m+j_m)},$$

根据$C(j_1j_2\cdots j_m)$的定义, 我们通过对

$$\begin{pmatrix}B & C(j_1j_2\cdots j_m)\end{pmatrix}$$

做初等变换可知,

$$\begin{vmatrix}B & C(j_1j_2\cdots j_m)\end{vmatrix}=B\begin{pmatrix}j_1j_2\cdots j_m\\ 12\cdots m\end{pmatrix}(-1)^{1+2+\cdots+m+j_1+\cdots+j_m}.$$

所以

$$\Delta=(-1)^m\sum_{1\le j_1<\cdots<j_m\le n}A\begin{pmatrix}12\cdots m\\ j_1j_2\cdots j_m\end{pmatrix}B\begin{pmatrix}j_1j_2\cdots j_m\\ 12\cdots m\end{pmatrix}.$$

所以

$$|AB|=\begin{cases}0, & m>n;\\ \sum_{1\le j_1<\cdots<j_m\le n}A\binom{12\cdots m}{j_1j_2\cdots j_m}B\binom{j_1j_2\cdots j_m}{12\cdots m}, & m\le n\end{cases}$$

□

§2.4 矩阵的逆和Cramer 法则

Cramer 是一位十八世纪的瑞士数学家, 18岁就获得博士学位. Cramer 法则是一种明确给出线性方程组解的方法.

我们在上一节看到, 对域F上的n阶矩阵A来说, 拉普拉斯定理也可以写成下面的形式.

$$AA^*=A^*A=\begin{pmatrix}|A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A|\end{pmatrix}=|A|I_n,$$

其中I_n是n阶单位矩阵. 因此如果$|A|\ne 0$, 令

$$B=|A|^{-1}A^*,$$

则$AB=BA=I_n$. 这时我们称B为A的逆矩阵, 简称A的逆, 记为A^{-1}. 我们看到

$$A^*=|A|A^{-1}.$$

定义 2.4.1 (矩阵的逆)**.** . 对域F上的n阶矩阵A来说, 如果存在n阶矩阵B使得

$$AB = BA = I_n,$$

我们称B为A的逆矩阵, 简称A的逆.

定理 2.4.1. *假设A是域F上的$m \times n$矩阵, 下面几条是等价的*

1. *$m = n$, 存在n阶矩阵B使得$AB = BA = I_n$.*

2. *$m = n$, 存在n阶矩阵B使得$AB = I_n$.*

3. *$m = n$, $|A| \neq 0$.*

4. *存在一个域F上的$n \times m$矩阵B, 使得$AB = I_m$, $BA = I_n$.*

因此以上的任何一条都可以作为可逆矩阵的等价定义.

定理 2.4.2. *假设A是域F上的n阶可逆矩阵, 则A可以写成有限个初等矩阵的乘积, 即A可以经过有限次初等变换变为单位矩阵.*

证明 我们用数学归纳法来证明. 当$n = 1$时, 命题显然成立. 假设命题对于$n-1$阶的矩阵都成立. 那么由于A可逆, 所以A的第一列一定不全为0, 我们可以通过初等变换把这个非零的元素调整到左上角, 然后变成1. 再通过把第一行乘上适当的元素依次加到下面的各行, 可以把第一列中除了第一行之外其余的位置都变成0. 再通过把第一列乘上适当的元素加到后面的各列, 可以把第一行除了第一列之外所以的位置都变为0. 这样矩阵A就通过有限次初等变换变为了下面的样子

$$\begin{pmatrix} 1 & 0 \\ 0A_1 & \end{pmatrix}$$

其中A_1是一个$n-1$阶的可逆矩阵. 由归纳假设, A_1可以通过初等变换变为$n-1$阶单位矩阵. 定理得证. □

根据上面这个定理, 可以给一种求矩阵的逆的方法. 假设

$$A = P_1 \cdots P_k,$$

其中的P_i是初等矩阵. 那么

$$A^{-1} = P_k^{-1} \cdots P_1^{-1}.$$

所以对下面的$n \times 2n$矩阵

$$(A, I_n)$$

进行初等行变换相等于在左边乘上一个可逆矩阵P.

$$(A, I_n) \longrightarrow P(A, I_n) = (PA, P).$$

如果$PA = I_n$, 那么$P = A^{-1}$. 所以我们只需要对$n \times 2n$矩阵(A, I_n)进行初等行变换, 当左半边的A变成单位矩阵的时候, 右半边的I_n就变成了A^{-1}. 这个方法对利用伴随矩阵求逆的方法要快捷很多.

假设A是一个n阶的可逆矩阵,

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

那么线性方程组

$$Ax = b$$

的解为

$$x = A^{-1}b.$$

定理 2.4.3 (Cramer 法则)**.** *如果A是一个可逆矩阵, 那么线性方程组*

$$Ax = b$$

的解为

$$\begin{cases} x_1 = \frac{|A_1|}{|A|} \\ \vdots \\ x_n = \frac{|A_n|}{|A|} \end{cases}$$

其中A_j是把A的第j列换成b之后得到的矩阵.

证明 由前面的讨论可知,

$$x = A^{-1}b = |A|^{-1}A^*b.$$

所以

$$x_j = |A|^{-1}(\sum_{i=1}^{n} A_{ij}b_i).$$

由拉普拉斯定理可知$\sum_{i=1}^{n} A_{ij}b_i$是把$|A|$的第j列换成b之后得到的行列式. 所以$x_j = \frac{|A_j|}{|A|}$. □

Cramer 法则在理论上非常明晰漂亮. 但是在实际计算线性方程组的解的时候, 经常采用后面我们要介绍的高斯消元法.

§2.5 LU分解和PLU分解

如果一个矩阵A能写成一个下三角矩阵L和一个上三角U的乘积, 我们就称

$$A = LU$$

是A的LU分解. 注意并不是每个矩阵都有LU分解. 即使是可逆矩阵, 也不一定有LU分解. 例如

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

就没有LU分解.

如果一个矩阵A能写成一个置换矩阵P, 一个下三角矩阵L和一个上三角U的乘积, 我们就称

$$A = PLU$$

是A的PLU分解. 每个可逆矩阵都有PLU分解.

§2.6 行列式的第二种定义

定义 2.6.1 (行列式的第二种定义). *如果一个映射*

$$\begin{aligned} \det: \quad F^{n\times n} &\longrightarrow \quad\quad F, \\ A=(a_{ij}) &\longmapsto \quad \det(A). \end{aligned}$$

满足下面的两条:

1. $\det(AB) = \det(A)\det(B)$.
2. $\det$*是一个域*F*上的关于*n^2*个变量*a_{ij}*的非常值多元多项式, 而且是满足上面一条的次数最低的多项式,*

那么我们就把$\det(A)$*称为是*A*的行列式.*

我们下面来证明满足上面两条性质的映射如果存在的话, 一定是唯一的, 而且和行列式的第一种定义相一致, 即

$$\det(A) = |A|.$$

也就是说上面的两条性质就可以完全确定了行列式.

定理 2.6.1 (行列式的两种定义的等价性). *假设映射*$\det$*满足行列式的第二种定义, 那么它一定满足下面的性质:*

1. 单位矩阵的det是1.

2. 零矩阵的det是0.

3. $\det(E_{ij}(t)) = 1$.

4. $\det(S_i(t)) = t$.

5. $\det(P_{ij}) = -1$.

6. det关于每一行都是线性的. 即如果把A除了第i行之外的位置都取固定的常数, 那么det可以看做是关于第i行n个变量a_{i1}, ..., a_{in}的n元函数. 我们记为$g = g(a_{i1}, ..., a_{in})$. 这时如果

$$a_{ij} = aa_{ij}^{(1)} + ba_{ij}^{(2)},$$

那么

$$g(a_{i1}, ..., a_{in}) = ag(a_{i1}^{(1)}, ..., a_{in}^{(2)}) + bg(a_{i1}^{(2)}, ..., a_{in}^{(2)}).$$

7. $A = (a_{ij})$的det等于

$$\sum_{j_1, j_2, \ldots, j_n} (-1)^{\delta(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} = |A|$$

求和号中$j_1 j_2 \cdots j_n$跑遍1, 2, ..., n的所有排列,

$$\delta(j_1 j_2 \cdots j_n) = \begin{cases} 1, & j_1 j_2 \cdots j_n \text{是一个偶排列}, \\ -1, & j_1 j_2 \cdots j_n \text{是一个奇排列}, \end{cases}$$

证明 我们假设满足第二种定义的det是存在的. 那么

1. 选一个det不等于零的A, 由$I_n A = A$, 可知单位矩阵的det是1.

2. 由$0A = 0$可知零矩阵的det是0.

3. 由$E_{ij}(t)E_{ij}(-t) = I_n$可知$\det(E_{ij}(t))\det(E_{ij}(-t)) = 1$. 两个多项式相乘等于1, 只能是每一个都是零次的多项式即常数. 既然是常数, 就和t 的取值无关, 因此令$t = 0$即可知$\det(E_{ij}(t)) = 1$.

4. 因为$S_i(t)S_i(1/t) = 1$, 所以多项式$S_i(t)$只能是个单项式at^k的形式. 由$|S_i(1)| = 1$可知$a = 1$. 所以$|S_i(t)| = t^k$. 由于我们要求det 的次数最低, 这一点可以保证$k = 1$. 实际上只要关于a_{ij}的非常值多元多项式θ满足$\theta(AB) = \theta(A)\theta(B)$, 那么$\theta$的任意次幂也满足这个条件. 所以仅仅要求$\theta(AB) = \theta(A)\theta(B)$不足以完全

确定θ, 对次数的极小性的要求是必须的. 这里的$k \geq 1$. 我们下面会看到存在关于a_{ij}的非常值多元多项式θ, 使得$\theta(AB) = \theta(A)\theta(B)$, 而且$\theta(S_i(t)) = t$. 因此$k = 1$.

由$\det(S_i(t)) = t$能够得出矩阵的某一行是零行, 那么其det一定是零.

5. 由于$P_{ij} = S_i(-1)E_{ij}(1)E_{ji}(-1)E_{ij}(1)$, 所以$\det(P_{ij}) = -1$.

6. 如果把det 除了第i行之外的位置都取固定的常数, 即把det看做是关于第i行n个变量a_{i1}, ..., a_{in}的n元函数$g = g(a_{i1}, ..., a_{in})$. 那么由于$|S_i(t)| = t$, 所以$g$是关于$a_{ii}$的一次多项式. 因为$|S_j(t)| = t$和$|P_{ij}| = -1$, 所以$g$是关于$a_{ij}$的一次多项式. 所以$g = g(a_{i1}, ..., a_{in})$是关于$a_{i1}$, ..., a_{in}的n元一次多项式. 又因为$g(0, ..., 0) = 0$, 所以g是关于a_{i1}, ..., a_{in}的n元一次齐次多项式. 所以如果
$$a_{ij} = aa_{ij}^{(1)} + ba_{ij}^{(2)},$$
那么
$$g(a_{i1}, ..., a_{in}) = ag(a_{i1}^{(1)}, ..., a_{in}^{(2)}) + bg(a_{i1}^{(2)}, ..., a_{in}^{(2)}).$$

7. 记e_i为第i个位置是1, 其余位置是0的行向量. 则矩阵A的第i行向量为
$$a_{i1}e_1 + \cdots + a_{in}e_n.$$
由于$\det(P_{ij}) = -1$, 所以
$$\det\begin{pmatrix} e_{j_1} \\ e_{j_2} \\ \vdots \\ e_{j_n} \end{pmatrix} = \begin{cases} 1, & j_1j_2\cdots j_n\text{是一个偶排列}, \\ -1, & j_1j_2\cdots j_n\text{是一个奇排列}, \end{cases}.$$

把$\det(A) = det(a_{ij})$按照线性的运算法则展开可得
$$\det(A) = |(a_{ij})| = \det\begin{pmatrix} \sum_j a_{1j}e_j \\ \sum_j a_{2j}e_j \\ \vdots \\ \sum_j a_{nj}e_j \end{pmatrix}$$
$$= \sum_{j_1, j_2, ..., j_n} (-1)^{\delta(j_1j_2\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}$$
求和号中$j_1j_2\cdots j_n$跑遍1, 2, ..., n的所有排列,
$$\delta(j_1j_2\cdots j_n) = \begin{cases} 1, & j_1j_2\cdots j_n\text{是一个偶排列}, \\ -1, & j_1j_2\cdots j_n\text{是一个奇排列}, \end{cases}$$

□

到此为止, 我们证明了如果行列式是存在的. 只能是定理中最后一条的表达式, 是唯一的. 我们把这个多项式记为是$\theta(A)$. 我们还要证明这个多项式的确满足行列式的定义. 由第四条的证明可知, 次数的极小性是满足的. 因此只需证明$\theta(AB) = \theta(A)\theta(B)$.

§2.7 矩阵的秩

定义 2.7.1. *如果$m \times n$ 矩阵A存在一个k阶子式不等于0, 而所有的$k+1$阶的子式都等于0. 我们就说A的秩是k.*

定理 2.7.1. *初等变换不改变矩阵的秩.*

证明 由于初等变换都是可逆的. 所以如果我们能证明初等变换不会增加矩阵的秩, 也就相当于说明了初等变换不会减少矩阵的秩.

由于行列的对称性, 我们只需要说明初等的行变换不会增加矩阵的秩即可. 假设$m \times n$ 矩阵A的秩是r. 我们下面来说明对A做一次初等行变换, A的秩不会增加, 即A的所有的$k+1$阶的子式仍然都等于0.

把矩阵的某一行乘上一个非零的常数后, A的$k+1$阶子式即使有变化, 也只是变为原来的t倍, 所以仍然都等于0. 把矩阵的第i行乘上一个t后, 加到A的第j行上去. 那么A的某个$k+1$阶子式即使有变化, 那么有变化的那一行是原来矩阵的两行的一个组合. 把$k+1$阶子式的有变化的那一行按照拉普拉斯定理对行列式展开可知, 变化后的$k+1$阶子式能够写成原来的两个$k+1$阶子式的一个组合. 所以仍然还是0. 这样第二种和三种初等变换都不会增加矩阵的秩. 由于第一种初等变换能够由第二种初等变换和第三种初等变换复合得到, 所以第一种初等变化也不会增加矩阵的秩.

因此初等变换不改变矩阵的秩. □

定理 2.7.2. *对任意的秩为k的$m \times n$ 矩阵A, 都存在一个m阶可逆矩阵P和一个n阶可逆矩阵Q使得*

$$PAQ = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}.$$

其中I_k是一个r阶的单位矩阵, 而其余的位置都是0.

证明 根据假设A有一个k阶子式不等于0, 即A有一个$k \times k$的子矩阵A_1是可逆的. 我们可以通过换行和换列把这个k阶子矩阵调整到左上角. 这样子相当于说存在一个m阶可逆矩阵P_1和一个n阶可逆矩阵Q_1使得

$$P_1AQ_1 = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}.$$

然后我们通过左乘和右乘得到

$$\begin{pmatrix} I_k & 0 \\ -A_3A_1^{-1} & I_{m-k} \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\begin{pmatrix} I_k & -A_1^{-1}A_3 \\ 0 & I_{n-k} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_4 - A_3A_1^{-1}A_2 \end{pmatrix}.$$

如果$A_4 - A_3A_1^{-1}A_2 \neq 0$, 那么肯定有非零的元素. 我们假设这个非零的元素α 位于$m \times n$ 矩阵的$(i,\ j)$位置. 那么在P_1AQ_1中我们考虑由1, ..., k, i行和1, ..., k, j 列所组成的$k+1$ 阶矩阵B. 矩阵B经过初等变换之后得到了

$$\begin{pmatrix} I_k & 0 \\ 0 & \alpha \end{pmatrix}$$

是一个$k+1$阶可逆矩阵. 所以B也是可逆矩阵. 所以在A中存在一个$k+1$阶可逆子矩阵. 这个假设矛盾. 所以$A_4 - A_3A_1^{-1}A_2 = 0$. 所以命题成立. □

定理 2.7.3. *假设A是$s \times n$的矩阵, B是$n \times m$的矩阵. 求证*

$$r(AB) \geq r(A) + r(B) - n.$$

证明 首先可以利用矩阵秩的定义来证明

$$r(A) + r(B) = r\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \leq r\begin{pmatrix} A & 0 \\ I_n & B \end{pmatrix}.$$

然后利用矩阵的初等变换可以知道

$$r\begin{pmatrix} A & 0 \\ I_n & B \end{pmatrix} = r\begin{pmatrix} A & -AB \\ I_n & 0 \end{pmatrix} = r\begin{pmatrix} 0 & -AB \\ I_n & 0 \end{pmatrix} = n + r(AB).$$

□

习题 2.7

1. 计算行列式

$$\begin{vmatrix} a & a+h & a+2h & \cdots & a+(n-2)h & a+(n-1)h \\ -a & a & 0 & \cdots & 0 & 0 \\ 0 & -a & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 0 \\ 0 & 0 & 0 & \cdots & -a & a \end{vmatrix}$$

解 从行列式的最后一列开始, 逐列往前一列加,, 就可以得到一个上三角的行列式. 因此行列式等于

$$a^{n-1}(na+\frac{n(n-1)}{2}h).$$

□

2. 求行列式

$$\Delta=\begin{vmatrix}1+a_1+x_1 & a_1+x_2 & \cdots & a_1+x_n\\ a_2+x_1 & 1+a_2+x_2 & \cdots & a_2+x_n\\ \vdots & \vdots & \ddots & \vdots\\ a_n+x_1 & a_n+x_2 & \cdots & 1+a_n+x_n\end{vmatrix}$$

解

$$\begin{aligned}\Delta&=\begin{vmatrix}1 & x_1 & x_2 & \cdots & x_n\\ 0 & 1+a_1+x_1 & a_1+x_2 & \cdots & a_1+x_n\\ 0 & a_2+x_1 & 1+a_2+x_2 & \cdots & a_2+x_n\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & a_n+x_1 & a_n+x_2 & \cdots & 1+a_n+x_n\end{vmatrix}\\
&=\begin{vmatrix}1 & x_1 & x_2 & \cdots & x_n\\ -1 & 1+a_1 & a_1 & \cdots & a_1\\ -1 & a_2 & 1+a_2 & \cdots & a_2\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ -1 & a_n & a_n & \cdots & 1+a_n\end{vmatrix}\\
&=\begin{vmatrix}1 & x_1 & x_2-x_1 & \cdots & x_n-x_1\\ -1 & 1+a_1 & -1 & \cdots & -1\\ -1 & a_2 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ -1 & a_n & 0 & \cdots & 1\end{vmatrix}=\begin{vmatrix}A & B\\ C & I_{n-1}\end{vmatrix}=\begin{vmatrix}A-BC & 0\\ C & I_{n-1}\end{vmatrix}\\
&=|A-BC|=\begin{vmatrix}1-nx_1+\sum\limits_{i=1}^n x_i & x_1+x_1\sum\limits_{i=1}^n a_i-\sum\limits_{i=1}^n x_ia_i\\ -n & 1+\sum\limits_{i=1}^n a_i\end{vmatrix}\\
&=(1+\sum_{i=1}^n a_i)(1+\sum_{i=1}^n x_i)-n(\sum_{i=1}^n a_ix_i).\end{aligned}$$

□

3. 假设在矩阵

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

中删掉第一, 二, 三, 四列所得到的行列式是A, B, C, D. 求证

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

证明 按照前三行展开. 根据拉普拉斯定理, 应该得到二十项. 但是前三行的子式不能包含后两列中的任何一列, 而后三行的子式不能包含前两列. 所以前三行的子式必须包含前两列, 而后三行的子式必须包含后两列. 因为这样下来, 总共就只有两项非零了. 就是题目中的$AD - BC$. □

4. 假设

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}, \ B = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix},$$

$$D = \begin{pmatrix} x_1a_1 & x_1b_1 & x_2a_1 & x_2b_1 & x_3a_1 & x_3b_1 \\ x_1a_2 & x_1b_2 & x_2a_2 & x_2b_2 & x_3a_2 & x_3b_2 \\ y_1a_1 & y_1b_1 & y_2a_1 & y_2b_1 & y_3a_1 & y_3b_1 \\ y_1a_2 & y_1b_2 & y_2a_2 & y_2b_2 & y_3a_2 & y_3b_2 \\ z_1a_1 & z_1b_1 & z_2a_1 & z_2b_1 & z_3a_1 & z_3b_1 \\ z_1a_2 & z_1b_2 & z_2a_2 & z_2b_2 & z_3a_2 & z_3b_2 \end{pmatrix}.$$

求证$|D| = |A|^3 \cdot |B|^2$.

证明 可以把矩阵D写成下面的两个矩阵相乘.

$$D = \begin{pmatrix} x_1I_2 & x_2I_2 & x_3I_2 \\ y_1I_2 & y_2I_2 & y_3I_2 \\ z_1I_2 & z_2I_2 & z_3I_2 \end{pmatrix} \begin{pmatrix} A & & \\ & A & \\ & & A \end{pmatrix}$$

这里的I_2是二阶单位矩阵. 其中矩阵

$$\begin{pmatrix} x_1I_2 & x_2I_2 & x_3I_2 \\ y_1I_2 & y_2I_2 & y_3I_2 \\ z_1I_2 & z_2I_2 & z_3I_2 \end{pmatrix}$$

经过适当的换行换列之后就可以变成

$$\begin{pmatrix} B & \\ & B \end{pmatrix}$$

的形式. 其行列式是$|B|^2$, 而

$$\begin{pmatrix} A & & \\ & A & \\ & & A \end{pmatrix}$$

的行列式显然是$|A|^3$. 所以$|D| = |A|^3 \cdot |B|^2$. □

5. 假设$A = (a_{ij})$是一个$n \times n$的矩阵, a_{ij}的代数余子式记为A_{ij}, 则

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

证明 题中的行列式可以通过添行填列写成

$$\begin{vmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & a_{21} - a_{11} & a_{22} - a_{12} & \cdots & a_{2n} - a_{1n} \\ 0 & a_{31} - a_{11} & a_{32} - a_{12} & \cdots & a_{3n} - a_{1n} \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} - a_{11} & a_{n2} - a_{12} & \cdots & a_{nn} - a_{1n} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 1 & 1 & \cdots & 1 \\ 1 & a_{21} & a_{22} & \cdots & a_{2n} \\ 1 & a_{31} & a_{32} & \cdots & a_{3n} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

先按照第一列做拉普拉斯展开, 再按照第二行做拉普拉斯展开即可. □

6. 如果n阶行列式Δ的某一行或者列的所有的元素都是1, 则Δ等于所有元素的代数余子式的和.

证明 假设行列式的第i行的元素都是1. 那么按照这一行做拉普拉斯展开, 可知Δ等于这一行的代数余子式之和. 现在看对第j 行做拉普拉斯展开. 那么这第j 行的代数余子式之和等于把第j 行换成全是1之后的行列式. 可是一旦把第j行都换成了1 , 那么行列式中就有了两行全是1, 所以行列式就是零. 所以除了第i行之外其余的行的代数余子式之和都等于零. □

7. 计算三对角行列式

$$\begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix}$$

证明 我们把k阶情形的行列式记为Δ_k, 那么利用拉普拉斯展开定理可以知道有下面的关系式

$$\Delta_n = a\Delta_{n-1} - bc\Delta_{n-2},$$

然后我们知道$\Delta_1 = a$, $\Delta_2 = a^2 - bc$. 所以可以得到特征方程为

$$x^2 - ax + bc = 0.$$

这个方程的解我们假设为α, β 那么我们知道

$$\Delta_n = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}.$$

如果两个根α, β相等, 那么上式去极限值$(n+1)\alpha^n = (n+1)(a/2)^n$. □

8. 用x_1, ..., x_{n-1}, 1替换

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1n-1} & 1 \\ a_{12} & \cdots & a_{2n-1} & 1 \\ \vdots & \ddots & \vdots & 1 \\ a_{n1} & \cdots & a_{nn-1} & 0 \end{vmatrix}$$

的第i行, 得到的行列式记为Δ_i. 求证

$$\Delta = \Delta_1 + \cdots \Delta_n.$$

证明

注意要证的等式的右边是个关于x_1, ..., x_{n-1}的多项式, 而且其常数项是Δ. 把右边对每个x_i求导, 可知右边对每个变量x_i的导数都是0. 所以右边只有常数项. □

9. 假设$A = (a_{ij})$ 是一个$n \times n$的矩阵, $X = (x_1, .., x_n)^T$是一个n维列向量, $Y = (y_1, .., y_n)$ 是一个行向量n 维. 求证

$$\begin{vmatrix} A & X \\ Y & z \end{vmatrix} = z|A| - \sum_{ij} A_{ij} x_i y_j.$$

证明 [1] 将题目中的行列式先按照最后一列, 再按照最后一行做拉普拉斯展开即可.

□

证明 [2] 如果A可逆, 则

$$\begin{vmatrix} A & X \\ Y & z \end{vmatrix} = \begin{vmatrix} A & X \\ 0 & z - YA^{-1}X \end{vmatrix} = z|A| - YA^*X = z|A| - \sum_{ij} A_{ij} x_i y_j.$$

如果A不可逆, 则使用微小攝动法即可转化为A可逆的情形. □

10. 用化为三角形的方法计算行列式.

$$\begin{vmatrix} a & a+d & \cdots & a+(n-2)d & a+(n-1)d \\ a+d & a+2d & \cdots & a+(n-1)d & a \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a+(n-2)d & a+(n-1)d & \cdots & a+(n-4)d & a+(n-3)d \\ a+(n-1)d & a & \cdots & a+(n-3)d & a+(n-2)d \end{vmatrix}$$

证明 把所有的列到加到第一列得到

$$
\begin{aligned}
&\begin{vmatrix}
na+\frac{n(n-1)}{2}d & a+d & \cdots & a+(n-2)d & a+(n-1)d \\
na+\frac{n(n-1)}{2}d & a+2d & \cdots & a+(n-1)d & a \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
na+\frac{n(n-1)}{2}d & a+(n-1)d & \cdots & a+(n-4)d & a+(n-3)d \\
na+\frac{n(n-1)}{2}d & a & \cdots & a+(n-3)d & a+(n-2)d
\end{vmatrix} \\
=&(na+\frac{n(n-1)}{2}d)\begin{vmatrix}
1 & a+d & \cdots & a+(n-2)d & a+(n-1)d \\
1 & a+2d & \cdots & a+(n-1)d & a \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
1 & a+(n-1)d & \cdots & a+(n-4)d & a+(n-3)d \\
1 & a & \cdots & a+(n-3)d & a+(n-2)d
\end{vmatrix} \\
=&(na+\frac{n(n-1)}{2}d)\begin{vmatrix}
1 & a+d & \cdots & a+(n-2)d & a+(n-1)d \\
0 & d & \cdots & d & -(n-1)d \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & d & \cdots & d & d \\
0 & -(n-1)d & \cdots & d & d
\end{vmatrix} \\
=&(na+\frac{n(n-1)}{2}d)d^{n-1}\begin{vmatrix}
1 & \cdots & 1 & -(n-1) \\
1 & \cdots & -(n-1) & 1 \\
\vdots & \ddots & \vdots & \vdots \\
1 & \cdots & 1 & 1 \\
-(n-1) & \cdots & 1 & 1
\end{vmatrix}_{n-1} \\
=&(na+\frac{n(n-1)}{2}d)d^{n-1}\begin{vmatrix}
-1 & 1 & \cdots & 1 & -(n-1) \\
-1 & 1 & \cdots & -(n-1) & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
-1 & -(n-1) & \cdots & 1 & 1 \\
-1 & 1 & \cdots & 1 & 1
\end{vmatrix}_{n-1} \\
=&(na+\frac{n(n-1)}{2}d)d^{n-1}\begin{vmatrix}
-1 & 1 & \cdots & 1 & -(n-1) \\
0 & 0 & \cdots & -n & n \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & -n & \cdots & 0 & 0 \\
0 & n & \cdots & 0 & 0
\end{vmatrix}_{n-1} \\
=&(-1)^{-1+[\frac{n-2}{2}]}(na+\frac{n(n-1)}{2}d)\begin{vmatrix}
n & 0 & \cdots & 0 \\
-n & n & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & -n & n
\end{vmatrix}_{n-2} \\
=&(-1)^{-1+[\frac{n-2}{2}]}(a+\frac{(n-1)}{2}d)(nd)^{n-1}
\end{aligned}
$$

□

11. 爱丽丝和鲍勃两个人玩游戏. 他们轮流往一个十乘十的网格里填数字. 爱丽丝先填一个位置, 鲍勃再在剩下的一个位置上填上一个数字. 两人轮流填下去, 直到把这个十乘十的网格的一百个位置都填满. 如果最后得到的十阶行列式不等于零, 那么爱丽丝胜. 否则就是鲍勃胜. 请问他们两个人谁有必胜的策略? 这个策略是什么?

    解 鲍勃有必胜的策略. 他能够做到使得行列式的前两行一模一样, 从而行列式是零. 如果爱丽丝在第一行的某个位置填上了一个数, 那么鲍勃就在第二行相同的位置上填上相同的数. 如果爱丽丝在第二行的某个位置上填上了一个数, 那么鲍勃就在第一行的相同位置上填一样的数. 这样子就可以保证填出来的行列式有两行是完全相同的, 这样行列式等于零. 鲍勃就赢了. □

12. (Burnside定理) 假设$A = (a_{ij})$是一个偶数阶的反对称矩阵, 其中的a_{ij} $(i > j)$是未知数. 求证存在一个关于$a_{ij}(i > j)$ 的多项式f, 使得A 的行列式等于f^2. 这里的f称为是A的Pfaffian.

    证明 我们的证明中将采用最一般的记号. 也就是我们假设a_{ij} $(i > j)$是未知的变量. 我们要证明存在一个关于$a_{ij}(i > j)$ 的多项式f, 使得A的行列式等于f^2. 这样子的话, 如果A是一个偶数阶的元素都是整数的反对称矩阵, 那么我们立即就可以推出A的行列式是一个完全平方数, 即A的行列式是另外一个整数的平方.

    我们用数学归纳法, 假设命题对$n-2$阶的矩阵成立. 那么对n阶矩阵A, 我们假设

$$A = \begin{pmatrix} B & -C^T \\ C & D \end{pmatrix}.$$

其中

$$B = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$$

是一个2阶反对称矩阵. 而D是一个$n-2$阶的反对称矩阵, C是一个$(n-2)\times 2$的矩阵. 对A做初等变换可以得到

$$A = \begin{pmatrix} B & -C^T \\ C & D \end{pmatrix} \longrightarrow \begin{pmatrix} B & -C^T \\ 0 & D + CB^{-1}C^T \end{pmatrix}$$

不改变行列式. 由于$D+CB^{-1}C^T$也是一个反对称矩阵. 根据归纳假设可知, 一定存在多项式g, h使得$|A| = g^2/h^2$. 注意这里之所以会出现分母, 是因为B^{-1}中

的元素会出现$1/b$, 相当于a_{21}会出现在分母上. 所以计算行列式的时候需要先通分, 才能用归纳假设.

可是我们知道$|A|$是一个多项式, 而多项式有唯一分解定理, 所以h一定会整除g. 所以g/h实际上是一个多项式, 也就是Pfaffian.

□

下面两题是可以把行列式视为某些元素的多项式.

13. 求

$$\Delta_n=\Delta_n(a_1,\ \dots,\ a_{n-1})=\begin{vmatrix} x & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & x & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & a_2 & x & \cdots & a_{n-2} & a_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & x & a_{n-1} \\ a_1 & a_2 & a_3 & \cdots & a_{n-1} & x \end{vmatrix}.$$

解 首先可以看到要求的行列式Δ_n是一个关于x的n次首一多项式. 因此如果我们知道了该多项式所有的根, 我们也就可以把这个多项式确定下来. 我们可以把该行列式的所有的列都加到第一列上去, 这样得到的行列式没有变, 但是第一列的所有元素都变成了

$$x+\sum_{i=1}^{n-1}a_i.$$

因此如果令$x=-\sum\limits_{i=1}^{n-1}a_i$, 则可知行列式为零. 所以$\Delta_n=0$ 有一个根是$-\sum\limits_{i=1}^{n-1}a_i$. 另外令$x=a_i$, 可以发现$\Delta_n$的第$i$列和第$i+1$列是相同的, 因此行列式是零. 所以$x=a_i$也是$\Delta_n=0$ 的根.

所以如果

$$-\sum_{i=1}^{n-1}a_i,\ a_1,\ \dots,\ a_{n-1}$$

是n个互不相同的根, 那么由于n次方程最多只有n个根, 所以Δ_n只能是

$$\Delta_n=(x+\sum_{i=1}^{n-1}a_i)(x-a_1)\cdots(x-a_{n-1}).$$

如果$-\sum\limits_{i=1}^{n-1}a_i,\ a_1,\ \dots,\ a_{n-1}$中有重复的也没有关系. 我们可以用微小摄动来证明上面的式子仍然是成立的. 令

$$b_1=a_1+y,\ \dots,\ b_{n-1}=a_{n-1}+(n-1)y,\ b_n=-\sum_{i=1}^{n-1}b_i$$

我们考虑下面的非零多项式

$$g(y)=\prod_{n\geq i>j\geq 1}(b_i-b_j).$$

这个多项式的根只有有限个. 我们记δ为该多项式的最小的正根, 如果该多项式没有正根, 我们就记$\delta=1$. 那么我们只要取$y\in(0,\ \delta)$, 就可以保证$g(y)\neq 0$. 所以$b_1,\ b_2,\ ...,\ b_n$两两不同. 所以

$$\Delta_n(b_1,\ ...,\ b_{n-1})=(x+\sum_{i=1}^{n-1}b_i)(x-b_1)\cdots(x-b_{n-1}).$$

上式等式两边都是关于y的多项式, 而且对$(0,\ \delta)$内的所有的值都相等. 由于这个区间内有无限多个值. 两个多项式在无限多个点处都取相同的值, 这两个多项式肯定是相同的多项式. 特别的, 它们在$y=0$的地方也相等. 所以

$$\Delta_n=(x+\sum_{i=1}^{n-1}a_i)(x-a_1)\cdots(x-a_{n-1})$$

总是成立的. □

14. 求行列式

$$\Delta_n=\begin{vmatrix}1&1&\cdots&1\\x_1&x_2&\cdots&x_n\\\vdots&\vdots&\vdots&\vdots\\x_1^{k-1}&x_2^{k-1}&\cdots&x_n^{k-1}\\x_1^{k+1}&x_2^{k+1}&\cdots&x_n^{k+1}\\\vdots&\vdots&\vdots&\vdots\\x_1^n&x_2^n&\cdots&x_n^n\end{vmatrix}.$$

解 令

$$\Delta=\begin{vmatrix}1&1&\cdots&1&1\\x_1&x_2&\cdots&x_n&x\\\vdots&\vdots&\vdots&\vdots&\vdots\\x_1^{k-1}&x_2^{k-1}&\cdots&x_n^{k-1}&x^{k-1}\\x_1^{k}&x_2^{k}&\cdots&x_n^{k}&x^{k}\\x_1^{k+1}&x_2^{k+1}&\cdots&x_n^{k+1}&x^{k+1}\\\vdots&\vdots&\vdots&\vdots&\vdots\\x_1^n&x_2^n&\cdots&x_n^n&x^n\end{vmatrix}.$$

则Δ是一个$n+1$阶的范德蒙行列式, 所以

$$\Delta=(x-x_1)\cdots(x-x_n)\prod_{n\geq i>j\geq 1}(x_i-x_j).$$

Δ是关于x的多项式, 通过把Δ的最后一列做拉普拉斯展开可知, 其x^k 项的系数就是

$$(-1)^{k+1+n+1}\Delta_n.$$

而对$\Delta=(x-x_1)\cdots(x-x_n)\prod_{n\geq i>j\geq 1}(x_i-x_j)$.做展开可以其$x^k$ 项的系数是

$$\prod_{n\geq i>j\geq 1}(x_i-x_j)((-1)^{n-k}\sum_{1\leq i_1<\cdots<i_{n-k}\leq n}x_{i_1}\cdots x_{i_{n-k}}).$$

所以

$$\Delta_n=(\sum_{1\leq i_1<\cdots<i_{n-k}}x_{i_1}\cdots x_{i_{n-k}}\leq n)\prod_{n\geq i>j\geq 1}(x_i-x_j).$$

□

15. 假设$P_i(x)$是一个i次多项式, 其x^i项的系数是a_i. 求证

$$\begin{vmatrix} P_0(x_1) & P_0(x_2) & \cdots & P_0(x_n) \\ P_1(x_1) & P_1(x_2) & \cdots & P_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n-1}(x_1) & P_{n-1}(x_2) & \cdots & P_{n-1}(x_n) \end{vmatrix} = a_0\cdots a_{n-1}\prod_{n\geq i>j\geq 1}(x_i-x_j).$$

证明 首先注意到行列式的第一行都是a_0. 所以通过第一行呈上适当的倍数加到第二行到第n行, 可以把这$n-1$ 行的常数项都消掉. 这样, 第二行就变成了

$$(a_1x_1,\ ...,\ a_1x_n).$$

再把这个第二行乘上适当的倍数加到下面的各行, 可以把第3行到第n行的x项都消掉. 依次这样做下去, 就得到行列式

$$\begin{vmatrix} a_0 & a_0 & \cdots & a_0 \\ a_1x_1 & a_1x_2 & \cdots & a_1x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1}x_1^{n-1} & a_{n-1}x_2^{n-1} & \cdots & a_{n-1}x_n^{n-1} \end{vmatrix} = a_0\cdots a_{n-1}\prod_{n\geq i>j\geq 1}(x_i-x_j).$$

□

16. 证明下面的等式

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \cos\theta_1 & \cos\theta_2 & \cdots & \cos\theta_n \\ \vdots & \vdots & \ddots & \vdots \\ \cos(n-1)\theta_1 & \cos(n-1)\theta_2 & \cdots & \cos(n-1)\theta_n \end{vmatrix} = 2^{\frac{(n-2)(n-1)}{2}} \prod_{n\geq i>j\geq 1} (\cos\theta_i - \cos\theta_j),$$

$$\begin{vmatrix} \sin\theta_1 & \sin\theta_2 & \cdots & \sin\theta_n \\ \sin 2\theta_1 & \sin 2\theta_2 & \cdots & \sin 2\theta_n \\ \vdots & \vdots & \ddots & \vdots \\ \sin n\theta_1 & \sin n\theta_2 & \cdots & \sin n\theta_n \end{vmatrix} = 2^{\frac{n(n-1)}{2}} \sin\theta_1 \cdots \sin\theta_n \prod_{n\geq i>j\geq 1} (\cos\theta_i - \cos\theta_j).$$

17. 假设$A = (|i-j|)$, 求$|A|$.

解 从$|A|$中把各列减去第一列, 然后把第一行加到其余的各行, 再按照最后一行做拉普拉斯展开得到:

$$\begin{vmatrix} 0 & 1 & 2 & 3 & \cdots & n-1 \\ 1 & 0 & 1 & 2 & \cdots & n-2 \\ 2 & 1 & 0 & 1 & \cdots & n-3 \\ 3 & 2 & 1 & 0 & \cdots & n-4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n-1 & n-2 & n-3 & n-4 & \cdots & 0 \end{vmatrix}$$

$$= \begin{vmatrix} 0 & 1 & 2 & 3 & \cdots & n-1 \\ 1 & -1 & 0 & 1 & \cdots & n-3 \\ 2 & -1 & -2 & -1 & \cdots & n-5 \\ 3 & -1 & -2 & -3 & \cdots & n-7 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n-1 & -1 & -2 & -3 & \cdots & 1-n \end{vmatrix} = \begin{vmatrix} 0 & 1 & 2 & 3 & \cdots & n-1 \\ 1 & 0 & 2 & 4 & \cdots & 2n-4 \\ 2 & 0 & 0 & 2 & \cdots & 2n-6 \\ 3 & 0 & 0 & 0 & \ddots & 3n-8 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n-1 & 0 & 0 & 0 & \cdots & 0 \end{vmatrix}$$

$$= (-1)^{n-1}(n-1)2^{n-2}.$$

□

18. 求

$$|A| = |a_i j| = \begin{vmatrix} p_1 & a & a & \cdots & a \\ b & p_2 & a & \cdots & a \\ b & b & p_3 & \cdots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & p_n \end{vmatrix}$$

解 假设$f(x) = |a_ij - x|$. 则通过各行减去第一行可知, $f(x)$是一个关于x的一元一次多项式. 再由

$$f(a) = \prod_{i=1}^{n}(p_i - a),\ f(b) = \prod_{i=1}^{n}(p_i - b)$$

可知

$$|A| = \frac{b\prod\limits_{i=1}^{n}(p_i - a) - a\prod\limits_{i=1}^{n}(p_i - b)}{b - a}.$$

□

19. 假设$s_k = \sum\limits_{i=1}^{n} x_i^k,\ k = 0,\ 1,\ 2,\ \ldots.$ 求证

$$\begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} & 1 \\ s_1 & s_2 & s_3 & \cdots & s_n & x \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} & x^2 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ s_{n-1} & s_n & ss_{n+1} & \cdots & s_{2n-2} & x^{n-1} \\ s_n & s_{n+1} & ss_{n+2} & \cdots & s_{2n-1} & x^n \end{vmatrix}$$

证明

注意到

$$\begin{pmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} & 1 \\ s_1 & s_2 & s_3 & \cdots & s_n & x \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} & x^2 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ s_{n-1} & s_n & ss_{n+1} & \cdots & s_{2n-2} & x^{n-1} \\ s_n & s_{n+1} & ss_{n+2} & \cdots & s_{2n-1} & x^n \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n & x \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 & x^2 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ x_1^n & x_2^n & x_2^n & \cdots & x_n^n & x^n \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & 0 \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} & 0 \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

然后再等式两边去行列式即可. □

20.
$$\begin{vmatrix} \binom{m}{k} & \cdots & \binom{m}{k+n-1} \\ \vdots & \ddots & \vdots \\ \binom{m+n-1}{k} & \cdots & \binom{m+n-1}{k+n-1} \end{vmatrix} = \frac{\binom{m+n-1}{n}\cdots\binom{m+n-k}{n}}{\binom{k+n-1}{n}\cdots\binom{n}{n}}.$$

21. 设A是非零方阵, 其伴随矩阵$A^* = A'$, 则A是否一定是可逆矩阵?

证明 在实数域上式成立的. 在复数域上不一定成立. 例如

$$A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

□

第三章　线性方程组

§3.1　线性方程组的古代例子

线性方程组的求解是线性代数的起源, 其历史以追溯到三千多年的古巴比伦时代. 现在用来求解线性方程组的标准做法"高斯消元法" 其实在中国古代的"九章算术"中就有记载.

假设F是一个域, 以F中的元素为系数, 关于n个变量x_1, ..., x_n的一次方程称为是一个线性方程, 形如

$$a_1x_1 + \cdots + a_nx_n = b,\ a_i \in F,\ b \in F.$$

而一组这样的方程称为是线性方程组. 形如

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

例如中国古代的数学著作"孙子算经"中有下面的题目

"今有兽六首四足, 禽四首二足, 上有七十六首, 下有四十六足, 问禽兽各几何?"

翻译成现代的语言就是求解线性方程组

$$\begin{cases} 6x + 4y = 76 \\ 4x + 2y = 46 \end{cases}$$

"孙子算经"中给出的答案是"八兽七禽", 解法是

" 倍足以减首, 余半之即兽, 以四乘兽, 减足, 余半之即禽." 用现在的记号相当于是上面的线性方程组中第二个式子两边乘2, 然后减去第一个式子, 得到

$$2x = 16,$$

所以$x = 8$. 再把$x = 8$带入第二个式子, 解出$y = 7$, 所以线性方程组的解为

$$\begin{cases} x = 8 \\ y = 7 \end{cases}$$

在"九章算术"中记载了一个更复杂的题目, 相当于解下面的线性方程组

$$\begin{cases} 3x+2y+z=39 \\ 2x+3y+z=34 \\ x+2y+3z=26 \end{cases}$$

我们把线性方程组的3个方程依次记为①, ②, ③. 记号s①$+t$ⓙ表示把原来的方程组的第i行变为第i行的s倍再加上第j行的t倍. "九章算术"中提供的解法如下.

首先2②, 3③得到

$$\begin{cases} 3x+2y+z=39 \\ 6x+9y+3z=102 \\ 3x+6y+9z=78 \end{cases}$$

这时用①, ②, ③就是指上面新的方程组的三个方程, 下面的也使用相同的记号. ①, ②, ③永远是指新得到的线性方程组的三个方程. 用② − 2②, ③ − ① 得到

$$\begin{cases} 3x+2y+z=39 \\ 5y+z=24 \\ 4y+8z=39 \end{cases}$$

再5③ − 4②得到

$$36z=99,$$

所以$z=11/4$, 再把$z=11/4$带入②即$5y+z=24$可得$y=17/4$, 最后带入①可得$x=37/4$, 即

$$\begin{cases} x=37/4 \\ y=17/4 \\ z=11/4. \end{cases}$$

孙子算经和九章算术里的例子其实都是采用消元法. 即从线性方程组中依次消去未知元, 最后得到一个只包含一个未知元的线性方程, 从而解出该未知元的具体数值, 然后再依次返回去带入其他的方程, 求出其他的未知元的值.

§3.2 高斯消元法

从第一节中我们看到对线性方程组进行下面的操作不影响方程组的解.

1. 把方程组中的某一个行乘上一个非零倍数, 即把方程的某个等式两边乘上一个相同的倍数.

2. 交换方程组中的两行, 即某两个方程的位置.

3. 把方程组的某一行乘上一个倍数加到另外一行上去.

上面的三种变换, 叫做线性方程组的三种初等变换.

另外我们也看到, 解方程的过程, 只涉及到线性方程组的各项的系数之间的加减乘除等运算. 所以把未知元单独拿出来, 把线性方程组的系数写成矩阵的形式, 对矩阵进行操作即可. 假设

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

解线性方程组就是要解矩阵方程

$$Ax = b.$$

对线性方程组进行三种初等变换就相当于把适当的可逆矩阵P乘到上面的$Ax = b$上去, 将其变成

$$PAx = Pb.$$

下面为了讨论方便我们定义增广矩阵

定义 3.2.1 (增广矩阵)**.** *我们把线性方程组$Ax = b$对应的$m \times (n+1)$ 的矩阵$\begin{pmatrix} A & b \end{pmatrix}$称为是其增广矩阵.*

定义 3.2.2 (约化行梯形矩阵)**.** *一个矩阵称为约化行梯形矩阵如果它满足下面的几条性质*

1. *矩阵的非零行在零行的上方.*

2. *每一个非零行从左到右的第一个非零元素为1, 叫做主1.*

3. *主1所在的列除了主1之外都是0.*

4. *主1所在的列指标随着行指标的递增而严格递增.*

定理 3.2.1. *任何一个矩阵都可以通过有限次初等行变换变为约化行梯形矩阵.*

证明 假设矩阵为M. 首先看第一列中是否有非零元素, 如果没有, 那么这一列就是零列, 可以不管, 转去考虑下一列. 如果第一列有非零元, 那么可以通过初等行

变换把这个非零元调到第一行, 并且将它变成主1. 再通过第一行乘上适当的系数往下面各行加, 可以把第一列中, 除了第一个位置之外都变成0.

然后来考虑第二列. 如果第二列中第二行和第二行以下的各个位置都为零, 那么转为考虑下一列. 否则, 在第二列中第二行和第二行以下的各个位置找一个非零的元素通过初等行变换调整到第二行, 并且将它也变为主1.

再下面依次类推, 可以对第三列, 第四列, ..., 施行同样的操作. 最后得到的矩阵是一个约化行梯形矩阵.

□

推论 3.2.2. *任何一个$m\times n$ 的矩阵A都可以通过左乘一个m阶可逆矩阵P, 使得PA变为约化行梯形矩阵.*

定理 3.2.3 (约化行梯形矩阵的唯一性)**.** *假设$m\times n$ 的矩阵A可以通过左乘一个m阶可逆矩阵P_1, 使得P_1A变为约化行梯形矩阵, 也可以通过左乘一个m阶可逆矩阵P_2, 使得P_2A变为约化行梯形矩阵. 那么$P_1A=P_2A$.*

证明 首先矩阵的秩只和A有关, 也就是主1的个数只和A有关. 其次, 主1所在的列也只和A 有关. 我们来简单的证明一下这个事实. 第一个主1所在的j_1列一定是A的第一个非零列, 当然只和A有关. 第一个主1所在的j_1 列确定了以后, 如果它的下一列j_1+1 列与它线性无关, 那么j_1+1 列也有一个主1. 否则就要跳过这一列, 继续往后找, 直到找到第j_2列, 满足j_2列和j_1列线性无关. 这个j_2的值显然也是只和A有关系. 所以j_1就是A的第一个非零列的列数, j_2就是第一个不能被j_1列线性表示的列. 依次类推可知. 每一个主1所在的列就是不能被该列之前的各列线性表示的列. 这样的列当然是只跟A相关, 而跟具体的初等变换无关了.

因此矩阵的秩和零列的位置只和A有关, 与P_1, P_2无关. 所以我们可以通过右乘同一个可逆的置换矩阵Q使得

$$P_1AQ=\begin{pmatrix} I_r & A_1 \\ 0 & 0 \end{pmatrix},\ P_2AQ=\begin{pmatrix} I_r & A_2 \\ 0 & 0 \end{pmatrix}$$

所以

$$P_2P_1^{-1}\begin{pmatrix} I_r & A_1 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} I_r & A_2 \\ 0 & 0 \end{pmatrix}$$

对$P_2P_1^{-1}$进行适当的分块, 假设

$$P_2P_1^{-1}=\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix},$$

则

$$\begin{pmatrix} M_1 & M_1A_1 \\ M_3 & M_3A_1 \end{pmatrix}=\begin{pmatrix} I_r & A_2 \\ 0 & 0 \end{pmatrix},$$

所以$M_1 = I_r$, $A_1 = A_2$. 因此$P_1A = P_2A$. □

定理 3.2.4. 假设$A = (a_{ij})$是一个$m \times n$的矩阵, $b = (b_1, b_2, ..., b_m)^T$是一个$m$阶列向量, P是一个m 阶可逆矩阵使得

$$P\begin{pmatrix} A & b \end{pmatrix} = \begin{pmatrix} \widetilde{A} & \widetilde{b} \end{pmatrix}$$

是增广矩阵$\begin{pmatrix} A & b \end{pmatrix}$的约化行梯形矩阵. 则约化行梯形矩阵的最后一列没有主1, 当且仅当A的秩等于增广矩阵$\begin{pmatrix} A & B \end{pmatrix}$的秩, 也当且仅当线性方程组$AX = b$有解. 如果$AX = b$有解, 而且$A$的秩等于$n$, 那么$AX = b$有唯一解. 如果$AX = b$有解, 但是$A$的秩小于$n$, 那么$AX = b$的解不唯一. 如果域$F$中有无限多个元素, 方程就有无限多个解.

证明 适当的调换未知元的次序, 我们可以不妨设$\begin{pmatrix} A & b \end{pmatrix}$ 约化行梯形矩阵为

$$\begin{pmatrix} I_r & A_1 & \gamma_1 \\ 0 & 0 & \gamma_2 \end{pmatrix}$$

其中前n列为A的约化行梯形矩阵. 如果最后一列有主1, 那么一定出现在γ_2 中, 这样子原来的方程组经过行初等变换就会出现下面的不可能成立的式子

$$0 = 1.$$

故原方程组无解. 如果最后一列没有主1, 那么$\gamma_2 = 0$. 我们假设

$$A_1 = \begin{pmatrix} c_{1r+1} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{rr+1} & \cdots & c_{rn} \end{pmatrix}, \ \gamma_1 = \begin{pmatrix} d_1 \\ \vdots \\ d_r \end{pmatrix}$$

则线性方程组$Ax = b$和下面的方程组同解

$$\begin{cases} x_1 + c_{1r+1}x_{r+1} + \cdots + c_{1n}x_n = d_1 \\ \qquad\qquad \vdots \\ x_r + c_{rr+1}x_{r+1} + \cdots + c_{rn}x_n = d_r. \end{cases}$$

这个方程组显然是有解的. 实际上对x_{r+1}, ..., x_n任意赋值, 都可以解出对应的x_1, ..., x_r的值. 这个方程组的同解可以写成

$$\begin{cases} x_1 = -c_{1r+1}t_1 - \cdots - c_{1n}t_{n-r} + d_1 \\ \qquad\qquad \vdots \\ x_r = -c_{rr+1}t_1 - \cdots - c_{rn}t_{n-r} + d_r. \end{cases}$$

其中的t_1, ..., t_{n-r}是可以在域F中任意取值的独立的参数.

从上面的通解公式也可以看出, 如果如果$Ax = b$有解, 而且A的秩等于n, 那么$Ax = b$有唯一解. 如果$Ax = b$有解, 但是A的秩小于n, 那么$Ax = b$的解不唯一. 如果域F中有无限多个元素, 方程就有无限多个解. □

注意在上面的证明过程中, 通解公式

$$\begin{cases} x_1 = -c_{1r+1}t_1 - \cdots - c_{1n}t_{n-r} + d_1 \\ \qquad\vdots \\ x_r = -c_{rr+1}t_1 - \cdots - c_{rn}t_{n-r} + d_r. \end{cases}$$

也可以写成

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -c_{1r+1}t_1 - \cdots - c_{1n}t_{n-r} + d_1 \\ \vdots \\ c_{rr+1}t_1 - \cdots - c_{rn}t_{n-r} + d_r \\ t_1 \\ \vdots \\ t_{n-r} \end{pmatrix}$$
$$= \begin{pmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -c_{1r+1} \\ \vdots \\ c_{rr+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + t_{n-r} \begin{pmatrix} -c_{1n} \\ \vdots \\ c_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

如果我们记

$$\beta = \begin{pmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \alpha_1 = \begin{pmatrix} -c_{1r+1} \\ \vdots \\ c_{rr+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ ..., \ \alpha_{n-r} = \begin{pmatrix} -c_{1n} \\ \vdots \\ c_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

那么通解公式就是

$$x = \beta + t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}.$$

§3.3 线性方程组的解的结构

在前一节中, 我们把线性方程组

$$Ax = b$$

的通解写成

$$x = \beta + t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$$

的形式, 其中t_1, ..., t_{n-r}可以取F中任意的常数. 特别的, 我们如果取$t_1 = \cdots = t_{n-r} = 0$, 那么$x = \beta$就是原方程的解. 我们称$\beta$为原来方程的一个特解. 任取

$$\alpha = t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$$

为α_1, ..., a_{n-r}的一个线性组合. 那么

$$A\alpha = 0.$$

反之如果$A\alpha = 0$, 那么由前面的讨论可知α也一定能写成α_1, ..., a_{n-r}的线性组合. 我们把前面所得的结果写成下面的线性方程组的解的结构定理

定理 3.3.1. *线性方程组*

$$Ax = b$$

有解当且仅当A的秩r等于增广矩阵$\begin{pmatrix} A & b \end{pmatrix}$的秩. 这时, 存在$n$维列向量$\beta$, α_1, ..., a_{n-r}, 使得线性方程组的通解为

$$(x_1,\ x_2,\ ...,\ x_n)^T = \beta + t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$$

的形式. 其中β是一个特解, $t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$是线性齐次方程组$Ax = 0$的通解. 我们称$\alpha_1$, ..., a_{n-r}为线性其次方程组$Ax = 0$的一个基础解系.

上面这个定理也可以用线性变换的语言加以重新阐述.

令$V = F^n$为n维列向量空间, $W = F^m$为m维向量空间, 假设A是一个$m \times n$的矩阵. 我们定义映射

$$\mathscr{A}:\ V \longrightarrow\ \ W,$$

$$\alpha \longmapsto A\alpha.$$

很容易验证这是一个线性变换. 求$Ax = 0$的通解等价于求$\mathscr{A}$的核.

把A的列记为γ_1, ..., γ_n, 那么$A\alpha$是γ_1, ..., γ_n的线性组合, 所以$\mathscr{A}$的像空间就是γ_1, ..., γ_n张成的向量空间. 这个空间的秩等于A 的列秩, 也就是A的秩r. 所以根据维数公式, 求$\mathscr{A}$的核是一个$n - r$维的向量空间. 这个向量空间的任何一组基都叫做是$Ax = 0$的一个基础解系.

线性方程组

$$Ax = b$$

的解其实就是列向量b的原像. 我们知道b的原像不是空集当且仅当$b \in Im(\mathscr{A})$, 即b在A的像空间中, 当且仅当b在γ_1, ..., γ_n 张成的向量空间中, 即b能够写成A的列向量的线性组合. 而b能够写成A的列向量的线性组合当且仅当向量组γ_1, ..., γ_n的秩和γ_1, ..., γ_n, b的秩相等. 因此b的原像不是空集当且仅当系数矩阵的秩等于增广矩阵的秩.

定理 3.3.2. *假设A是一个n阶方阵, 如果A^k的秩等于A^{k+1}的秩, 求证A^{k+1}的秩等于A^{k+2}的秩.*

证明 A^k的秩等于A^{k+1}的秩, 当且仅当$A^k x = 0$的和$A^{k+1}x = 0$同解.

$$A^{k+2}x = 0 \Longrightarrow A^{k+1}(Ax) = 0 \Longrightarrow A^k(Ax) = 0$$

所以$A^{k+1}x = 0$的和$A^{k+2}x = 0$同解, 所以A^{k+1}的秩等于A^{k+2}的秩. □

推论 3.3.3. *假设A是一个n阶方阵, 那么A^n的秩等于A^{n+1}的秩.*

证明 如果A^n的秩不等于A^{n+1}的秩, 那么由上面的定理, 必然是

$$n \geq r(A) > r(A^2) > \cdots > r(A^n) > r(A^{n+1}) \geq 0.$$

上面这个式子要成立, 必须是

$$r(A) = n,\ r(A^2) = n-1,\ ...,\ r(A^{n+1}) = 0.$$

这显然是不可能的, 因为$r(A) = n$就能推出A是可逆的, 所以它的多少次幂都是可逆的, 其秩都是n. 所以矛盾. □

定理 3.3.4. *设数域F上的$m \times n$矩阵。A, B和$A+B$的秩分别是r, s和$r+s \leq min\{m,\ n\}$。证明存在可逆矩阵P, Q 使得*

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad PBQ = \begin{pmatrix} 0 & 0 \\ 0 & I_s \end{pmatrix}.$$

这里的I_r, I_s分别表示r, s阶的单位方阵。

证明 我们在这里提供一个纯矩阵的证明.

假设

$$A = P_1Q_1,\ B = P_2Q_2,$$

其中P_1, P_2列满秩, Q_1, Q_2行满秩. 那么

$$A+B=\begin{pmatrix}P_1 & P_2\end{pmatrix}\begin{pmatrix}Q_1\\ Q_2\end{pmatrix}.$$

注意由于A, B和$A+B$的秩分别是r, s和$r+s\le min\{m,\ n\}$, 上面这个等式的右边是个列满秩的矩阵乘以一个行满秩的矩阵. 因此可以通过初等行变换把左边的

$$\begin{pmatrix}P_1 & P_2\end{pmatrix}$$

变为约化行梯形, 即存在可逆矩阵P使得

$$P\begin{pmatrix}P_1 & P_2\end{pmatrix}=\begin{pmatrix}I_r & 0\\ 0 & 0\\ 0 & I_s\end{pmatrix}=\begin{pmatrix}PP_1 & PP_2\end{pmatrix}.$$

同理存在可逆矩阵Q, 使得

$$\begin{pmatrix}Q_1\\ Q_2\end{pmatrix}Q=\begin{pmatrix}I_r & 0 & 0\\ 0 & 0 & I_s\end{pmatrix}=\begin{pmatrix}Q_1Q\\ Q_2Q\end{pmatrix}.$$

所以

$$PAQ=PP_1Q_1Q=\begin{pmatrix}I_r\\ 0\\ 0\end{pmatrix}\begin{pmatrix}I_r & 0 & 0\end{pmatrix}=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix},$$

同理

$$PBQ=\begin{pmatrix}0 & 0\\ 0 & I_s\end{pmatrix}.$$

□

定理 3.3.5. 假设$A=(a_{ij})$是一个n阶矩阵, A^*是其伴随矩阵. 那么

$$r(A^*)=\begin{cases}n, & |A|\neq 0\\ 1, & r(A)=n-1\\ 0, & r(A)<n-1.\end{cases}$$

证明 当$|A|\neq 0$时, A是可逆的, $A^*=|A|A^{-1}$也是可逆的, 所以$r(A^*)=n$. 如果$r(A)=n-1$, 那么至少有一个$n-1$阶的子式不等于0. 所以$r(A^*)\geq 1$. 同时由于线性方程组

$$Ax=0$$

的解空间是一维的, 而A^*的列向量都是该线性方程组的解, 所以A^*的列秩不超过1. 所以$r(A^*)=1$.

如果$r(A)<n-1$, 那么A的所有的$n-1$阶子式都是0, 所以根据伴随矩阵的定义可知$A^*=0$. □

定理 3.3.6. 假设$A=(a_{ij})$是一个n阶矩阵, A^*是其伴随矩阵. 那么

$$r(A^*)=\begin{cases} n, & |A|\neq 0 \\ 1, & r(A)=n-1 \\ 0, & r(A)<n-1. \end{cases}$$

定义 3.3.1 (对角占优矩阵)**.** 假设$A=a_{ij}$是一个n阶复矩阵, 对每个$i\in\{1,\ ...,\ n\}$, 定义

$$R_i=\sum_{j\neq i}|a_{ij}|,\ C_j=\sum_{i\neq j}|a_{ij}|.$$

如果对于每个$i\in\{1,\ ...,\ n\}$, 都有

$$|a_{ii}|\geq R_i,$$

那么我们就称A是一个行对角占优矩阵. 如果$j\in\{1,\ ...,\ n\}$, 都有

$$|a_{jj}|\geq C_j,$$

那么我们就称A是一个列对角占优矩阵. 如果上面的大于等于号$\geq$ 改成严格的大于号$>$. 我们就称A是一个严格行对角占优矩阵或者严格列对角占优矩阵. 行*(*列*)*对角占优矩阵统称为对角占优矩阵, 严格行*(*列*)*对角占优矩阵统称为严格对角占优矩阵.

定理 3.3.7 (Levy - Desplanques定理)**.** 假设$A=a_{ij}$是一个n阶严格对角占优复矩阵. 那么A是可逆矩阵.

证明 不妨假设A是严格行主角占优矩阵. 如果$|A|=0$, 则存在非零的n维列向量$X=(x_1,\ ...,\ x_n)^T$使得

$$AX=0.$$

假设

$$|x_k|=\max\{|x_1|,\ ...,\|x_n|\}.$$

则

$$a_{kk}x_k=-\sum_{j\neq k}a_{kj}x_j,$$

两边取绝对值得

$$|a_{kk}||x_k|=|\sum_{j\neq k}a_{kj}x_j|\leq|x_k|R_k.$$

两边消去$|x_k|$可知$|a_{kk}|\leq R_k$. 与A是严格行主角占优矩阵. □

习题 3.3

1. 假设A是一个$m\times n$的实矩阵, 求证$r(A)=r(AA^T)$.

   证明 因为$Ax=0$当且仅当$(Ax)^TAx=0$当且仅当$x^TA^TAx=0$. 所以$Ax=0$和$A^TAx=0$同解. □

2. 假设A是一个$m\times n$的矩阵, 而且$r(A)=r$. 求证A可以写成r个秩为1 的矩阵的和, 即存在A_1, ..., A_r 使得

$$A=A_1+\cdots+A_r$$

而且$r(A_1)=\cdots=r(A_r)=1$.

证明 由于$r(A)=r$, 所以存在可逆矩阵P, Q使得

$$A=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}Q.$$

由于

$$I_r=\begin{pmatrix}1&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}+\cdots+\begin{pmatrix}0&&&\\&0&&\\&&\ddots&\\&&&1\end{pmatrix}.$$

我们记上式右边的求和中的r个矩阵为B_1, ..., B_r, 则

$$r(B_1)=\cdots=r(B_r)=1.$$

$$\begin{aligned}A&=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}Q\\&=P(\begin{pmatrix}B_1&0\\0&0\end{pmatrix}+\cdots+\begin{pmatrix}B_r&0\\0&0\end{pmatrix})Q.\end{aligned}$$

我们记

$$A_1=P\begin{pmatrix}B_1&0\\0&0\end{pmatrix}Q,\ ...,\ A_r=P\begin{pmatrix}B_r&0\\0&0\end{pmatrix})Q.$$

则$A=A_1+\cdots+A_r$, 而且$r(A_1)=\cdots=r(A_r)=1$. □

3. 假设A是一个$m\times n$的矩阵, 而且$r(A)=1$. 求证存在一个m维的列向量α和一个n 维的行向量β使得

$$A=\alpha\beta.$$

证明 因为$r(A)=1$, 所以存在可逆$P,\ Q$使得

$$A=P\begin{pmatrix}1 & 0_{1\times(n-1)}\\ 0_{(m-1)\times(1)} & 0_{(m-1)\times(n-1)}\end{pmatrix}Q.$$

假设$\tilde{\beta}=(1,\ 0,\ ...,\ 0)$是一个$n$维的行向量, $\tilde{\alpha}=(1,\ 0,\ ...,\ 0)^T$ 是一个m 维的列向量. 则

$$\begin{pmatrix}1 & 0_{1\times(n-1)}\\ 0_{(m-1)\times(1)} & 0_{(m-1)\times(n-1)}\end{pmatrix}=\tilde{\alpha}\tilde{\beta}.$$

所以

$$A=(P\tilde{\alpha})(\tilde{\beta}Q).$$

令$\alpha=P\tilde{\alpha},\ \beta=\tilde{\beta}Q$, 则$\alpha$是$m$阶方阵$P$的第一列, β 是n阶方阵Q的第一行, 而且

$$A=\alpha\beta.$$

□

4. 矩阵方程

$$A_{m\times n}X_{n\times k}=B_{m\times k}$$

有解当且仅当$r(A,\ B)=r(A)$.

证明 直接从线性方程组的解的相容性定理可知.

□

5. 若n阶矩阵A满足$A^2=A$, 求证存在可逆矩阵P使得

$$P^{-1}AP=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}.$$

其中r是A的秩.

证明 假设A的列向量的极大线性无关组为第$i_1,\ ...,\ i_r$列, 我们分别记为$\alpha_1,\ ..,\ \alpha_r$. 那么因为$A^2=A$, 所以

$$A\alpha_1=\alpha_1,\ ..,\ A\alpha_r=\alpha_r.$$

然后我们假设线性方程组

$$Ax=0$$

的一个基础解系为$\beta_1,\ ...,\ b_{n-r}$. 则令

$$P=(\alpha_1,\ ..,\ \alpha_r,\ \beta_1,\ ...,\ b_{n-r}).$$

P是一个n阶矩阵, 而且

$$AP = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

我们只需证明P是一个可逆矩阵即可. 这个只需要证明P的列向量线性无关即可. 假设c_1, ..., c_r, d_1, ..., d_{n-r}是常数使得

$$c_1\alpha_1 + \cdots + c_r\alpha_r + d_1\beta_1 + \cdots + d_{n-r}\beta_{n-r} = 0.$$

那么把上面的式子用A来作用可得

$$A(c_1\alpha_1 + \cdots + c_r\alpha_r + d_1\beta_1 + \cdots + d_{n-r}\beta_{n-r}) = 0.$$

因为β_1, ..., b_{n-r}是$Ax = 0$的基础解系, 所以$A(d_1\beta_1 + \cdots + d_{n-r}\beta_{n-r}) = 0$. 所以

$$A(c_1\alpha_1 + \cdots + c_r\alpha_r) = 0.$$

所以$c_1\alpha_1 + \cdots + c_r\alpha_r = 0$.. 由$\alpha_i$的取法可知, α_1, .., α_r是线性无关的. 所以c_1, ..., c_r都等于0. 因此α_1, .., α_r, β_1, ..., b_{n-r}是线性无关的, 所以P 是一个可逆矩阵, 而且满足

$$P^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

□

6. 若n阶矩阵A满足$A^2 = 0$, 求证存在可逆矩阵P使得

$$P^{-1}AP = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}.$$

其中B是一个行数和列数都小于A的矩阵.

证明

假设β_1, ..., b_{n-r}是$Ax = 0$的基础解系. 那么因为$A^2 = 0$, 所以A的列向量都包含在$Ax = 0$的解空间中. 将β_1, ..., b_{n-r}扩充称n维向量空间的一组基

$$\beta_1,\ ...,\ b_{n-r},\ \alpha_1,\ ..,\ \alpha_r.$$

那么$A\alpha_1$, .., $A\alpha_r$都可以写成β_1, ..., b_{n-r}的线性组合. 令

$$P = (\beta_1,\ ...,\ b_{n-r},\ \alpha_1,\ ..,\ \alpha_r),$$

则P是一个可逆矩阵. 而且

$$AP = P\begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}.$$

所以命题得证 □

7. 与所有的n阶矩阵都可以交换的矩阵只有纯量n阶矩阵.

证明 假设矩阵B个任何一个n阶矩阵都可以交换, 那么我们要证明它一定事故纯量矩阵.

假设A是一个对角线上依次为1, 2, ..., n的对角矩阵. 假设$AB = BA$, 那么立即可以推出B必须是对角矩阵. 再由B和任何初等的置换矩阵可以交换可知, B的对角线上的元素一定是相同的. 所以B是一个纯量矩阵.

□

8. 证明斜对称矩阵的秩是偶数.

证明 假设命题对阶数小于n的矩阵成立. 那么下面我们证明命题对n阶斜对称矩阵A也成立. 我们下面对A做初等变换的时候, 遵循一个基本的原则是对行做什么变换就对列做什么变换, 若我们交换了第i, j行, 我们就要同时交换第i, j列. 如果我们把某一行乘上了一个常数, 我们就要把相同的列也乘上相同的常数. 如果我们把第i行乘上t再加到第j上上去, 我们就要把第i列也乘上常数t也加到第j 列上去. 我们这样做可以保证, 对A经过这样对称的初等变换之后, 变出来的矩阵仍然是斜对称的.

我们不妨假设A的第一列不是零列. 这时我们可以通过换行和换列, 使得A的(0, 1)位置不是0. 然后乘上一个常数使得它变成1. 由于我们对A做的是相同的行列变换, 所以现在的$(2, 1)$ 位置一定是-1. 这样A就变成了下面的样子

$$A \longrightarrow \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

其中

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

C是一个斜对称矩阵. 由归纳假设可知, C的秩是个偶数, 因此A的秩等于C的秩加2也是偶数. □

9. 假设A, B是n阶方阵, $AB = BA = 0$, 并且$r(A)^2 = r(A)$. 求证

$$r(A + B) = r(A) + r(B).$$

证明 由于$r(A)^2=r(A)$, 所以存在存在n阶矩阵C和D使得

$$A^2C=A,\ DA^2=A.$$

所以

$$\begin{aligned}
r(A)+r(B) &= r\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \\
&= r\begin{pmatrix} A & 0 \\ A & B \end{pmatrix} \\
&= r\begin{pmatrix} A & A \\ A & A+B \end{pmatrix} \\
&= r\begin{pmatrix} A-A^2C & A \\ A-A^2C-BAC & A+B \end{pmatrix} \\
&= r\begin{pmatrix} 0 & A \\ 0 & A+B \end{pmatrix} \\
&= r\begin{pmatrix} 0 & A-DA^2-DAB \\ o & A+B \end{pmatrix} \\
&= r\begin{pmatrix} 0 & 0 \\ 0 & A+B \end{pmatrix} \\
&= r(A+B).
\end{aligned}$$

$\square$

第四章　线性空间　 线性映射

§4.1　线性空间

定义 4.1.1. 设F是一个数域，V是一个集合。我们称V是F上的线性空间，如果那么V中有两种运算"加法"和"数乘"，而且"加法"和"数乘"这两种运算满足下面的八条公理. 我们按照习惯, 称F中的元素为纯量, V中的元素为向量.

这八条公理如下:

1. 加法交换律: 对任意向量α, $\beta \in V$, 都有$\alpha + \beta = \beta + \alpha$,

2. 加法结合律: 对任意向量α, β, $\gamma \in V$都有, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$,

3. 存在零元: 对任意向量$\alpha \in V$, 都存在一个向量$0 \in V$使得$\alpha + 0 = \alpha$,

4. 存在负元: 对任意向量$\alpha \in V$, 都存在一个向量α'使得$\alpha' + \alpha = 0$,

5. 乘法对元素加法的分配律（数乘分配律I）:对任意向量α, $beta \in V$, $k \in F$, $k \cdot (\alpha + \beta) = k \cdot \alpha + k \cdot \beta$,

6. 乘法对纯量加法的分配律（数乘分配律II）: 对任意向量$\alpha \in V$, k, $l \in F$, $(k + l) \cdot \alpha = k \cdot \alpha + l \cdot \alpha$,

7. 数乘结合律: 对任意向量$\alpha \in V$, 纯量k, $l \in F$, 都有$(kl) \cdot \alpha = k \cdot (l \cdot \alpha)$,

8. 单位律: 对任意向量$\alpha \in V$, 纯量$1 \in F$, 都有$1 \cdot \alpha = \alpha$.

按照这个定义，加法交换律是不能由后面的7条推出的。

如果把第四条的负元α'写在右边的，这样加法交换律就可以由后面的7条推出了。具体如下：因为$0\cdot\alpha+0\cdot\alpha = (0+0)\cdot\alpha = 0\cdot\alpha$, 所以$0\cdot\alpha+0\cdot\alpha+(0\cdot\alpha)' = 0\cdot\alpha+(0\cdot\alpha)'$, 所以$0 \cdot \alpha + 0 = 0$, 再由第三条可知$0 \cdot \alpha = 0$. 所以$(-1) \cdot \alpha + \alpha = \alpha + (-1) \cdot \alpha = 0$. 又$0 \cdot \alpha + 1 \cdot \alpha = (0 + 1) \cdot \alpha = \alpha$, 所以$0 + \alpha = \alpha$. 然后就是简单的啦，

$$\alpha + \beta + \alpha + \beta = 2(\alpha + \beta) = 2\alpha + 2\beta = \alpha + \alpha + \beta + \beta,$$

所以

$$-\alpha + \alpha + \beta + \alpha + \beta + (-\beta) = -\alpha + \alpha + \alpha + \beta + \beta + (-\beta),$$

这就是

$$\alpha + \beta = \beta + \alpha.$$

例 4.1.1. 我们在解析几何中接触到的坐标平面

$$\mathbb{R}^2 = \{(x,\ y) | x \in \mathbb{R},\ y \in \mathbb{R}\}$$

是实数域上的二维线性空间. 这个线性空间中的加法运算就是我们熟知的对应坐标的相加, 而数乘就是常数乘到每一个分量上去. 即如果$\alpha = (x_1,\ y_1)$, $\beta = (x_2,\ y_2)$,

$k \in \mathbb{R}$, 那么

$$\alpha + \beta = (x_1 + x_2,\ y_1 + y_2),\ k\alpha = (kx_1,\ ky_1).$$

同样的三维空间

$$\mathbb{R}^3 = \{(x,\ y,\ z)|x,\ y,\ z \in \mathbb{R}\}$$

是实数域上的三维线性空间. 其向量之间的加法运算也对应坐标的相加, 数乘是常数乘到每一个分量上去. 即如果$\alpha = (x_1,\ y_1,\ z_1)$, $\beta = (x_2,\ y_2,\ z_2)$, $k \in \mathbb{R}$, 那么

$$\alpha + \beta = (x_1 + x_2,\ y_1 + y_2,\ z_1 + z_2),\ k\alpha = (kx_1,\ ky_1,\ kz_1)).$$

利用同样的方法, 我们可以定义一般的数域F上的n 维线性空间

$$F^n = \{(a_1,\ a_2,\ ...,\ a_n)|a_1,\ ...,\ a_n \in \mathbb{R}\}$$

其向量之间的加法运算同样的对应坐标的相加, 数乘是常数乘到每一个分量上去.

例 4.1.2. 假设n是一个固定的整数. 数域F上的所有的次数小于等于n的多项式的全体组成的集合V是线性空间. 注意由于0的次数是负无穷大, 所以$0 \in V$.

例 4.1.3. 假设$m,\ n$是两个个固定的整数. 数域F上的所有$m \times n$的矩阵的全体组成的集合$F^{m\times n}$是线性空间. 加法是矩阵对应位置相加, 数乘是数乘到矩阵的每一个位置上.

定义 4.1.2 (矩阵的迹)**.** 假设$A = (a_{ij})$是一个$n \times n$的矩阵. 矩阵的对角线上的元素之和

$$a_{11} + \cdots + a_{nn}$$

称为矩阵的迹*(Trace)*, 记为$Tr(A)$.

例 4.1.4. 迹为零的$n \times n$的矩阵的全体形成一个线性空间.

例 4.1.5. 数域F上的两个线性空间之间的线性变换的全体.

例 4.1.6. 以数域F中的元素为系数的多项式全体.

例 4.1.7. 闭区间$[0,\ 1]$上的连续函数的全体.

假设V_1, V_2是数域F上的两个线性空间. 我们可以通过下面定义的直和运算来构造新的线性空间.

定义 4.1.3 (线性空间的形式直和)**.** 假设$V_1,\ V_2$是数域F上的两个线性空间. 我们构造一个新的集合$V = \{(\alpha_1,\ \alpha_2)|\alpha_1 \in V_1,\ \alpha_2 \in V_2\}$. 我们定义$V$的零元为$(0_1,\ 0_2)$, 其中$0_1$是$V_1$的零元, 0_2是V_2的零元. V中元素$(\alpha_1,\ \alpha_2)$的负元定义为$(-\alpha_1,\ -\alpha_2)$.

V中的加法定义为对应的位置相加, 数乘定义为常数平等地乘到每一个位置上. 容易验证这样定义出来的V是一个新的线性空间. 我们称这个线性空间为V_1, V_2的形式直和. 记为

$$V = V_1 \bigoplus V_2.$$

定义 4.1.4 (线性无关和线性相关). 假设V是数域F上的线性空间. α_1, ..., a_n是V中的n个向量. 如果存在不全为零的n个数a_i使得

$$a_1\alpha_1 + \cdots + a_n\alpha_n = 0,$$

就称α_1, ..., a_n是V中线性相关的n个向量. 否则就称α_1, ..., a_n线性无关. 如果无限多个向量中的任意有限个都是线性无关的, 我们称这个无限向量组是线性无关的.

定义 4.1.5. 假设V是数域F 上的线性空间, $S = \alpha_1$, ..., α_n是V的一个向量组. 如果α 能够写成S中的元素的线性组合, 就称α能够被α_1, ..., α_n线性表出. 假设T是一个向量集合, 如果T中的每个向量能够被α_1, ..., α_n线性表出, 我们就称T能够S 线性表出.

定义 4.1.6 (向量组的等价). 如果两个向量组S, T张成的空间是相同的, 就称它们等价. 显然两个向量组S, T等价当且仅当S能够被T线性表出, T 也能被S线性表出.

定义 4.1.7 (极大线性无关组). 假设V是数域F 上的线性空间, S是V的一个向量组. 如果α_1, ..., $\alpha_n \in S$ 是线性无关的n个向量, 而且任取$\beta \in S$, $n+1$个向量

$$\alpha_1,\ ...,\ \alpha_n,\ \beta \in S$$

一定是线性相关的, 那么就称α_1, ..., α_n是V的一个极大线性无关组.

定义 4.1.8 (子空间, 子空间的和与交). 线性空间V的子集V_1称为是V的子空间如果V_1在继承自V的加法和数乘之下也成为一个线性空间. 子空间之间可以相加, 两个子空间的和定义为

$$V_1 + V_2 := \{\alpha_1 + \alpha_2 | \alpha_1 \in V_1,\ \alpha_2 \in V_2\},$$

容易验证这仍然是一个线性空间. 同样的, 也容易验证两个子空间的交$V_1 \cap V_2$也是一个子空间.

定义 4.1.9 (子集张成的子空间). 假设T是V的一个子集, T中的元素经过有限次数乘和加法运算生成的元素集合

$\{a_1t_1 + \cdots + a_kt_k|$, k 是正整数(可以任意大, 但必须是有限的), $a_i \in F,\ t_i \in T.\}$

是一个线性空间, 称为是由T张成的子空间。

定义 4.1.10 (基)**.** *如果线性空间V的子集T中的元素是线性无关的, 而且V能够由T张成，则称T是V的一组基。*

注意我们谈集合的时候，不考虑其中元素的排列次序，而当我们谈基的时候，常常要考虑其中元素的排列次序.

为了说明线性空间都有一组基, 我们先来介绍一下偏序集和佐恩引理.

定义 4.1.11 (偏序集)**.** *如果在集合S中有满足下列性质的二元关系$\leq$:*
(1) 对称性: $a \leq a$;
(2) 反对称性: 如果$a \leq b$和$b \leq a$同时成立, 那么$a = b$;
(3) 传递性: 如果$a \leq b$和$b \leq c$同时成立, 那么$a \leq c$,
我们就称带有二元关系$\leq$的集合S为一个偏序集.

例 4.1.8. *1. S为实数集合, 二元关系$\leq$为通常的实数之间的"小于等于"关系.*

2. S为集合T的全部子集组成的集合, 二元关系$\leq$为"包含于$\subset$"关系.

3. S为正整数集合, 二元关系$\leq$为"整除"关系.

从上面的最后一个例子可以看出偏序集中两个元素a, b之间并非必须要有关系. 例如2, 3之间, 既不是$2 \leq 3$, 也不是$3 \leq 2$. S中的一列元素a_1, a_2, ..., a_n, ..., 如果满足

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq \cdots$$

我们就称a_1, a_2, ..., a_n, ..., 是一条链. 这里要注意实际上链中所包含的元素个数不一定非要是可数的, 只要偏序集的一个子集中任何两个元素之间都可以排序, 都有关系, 即要么$a \leq b$, 要么$b \leq a$, 我们就称它是一条链. 链中的元素之间都可以比较次序, 这样的集合称为是全序集. 设T是S的一个子集, $a \in S$. 如果对任意的$t \in T$, 都有$t \leq s$; 我们就称t是T的一个上界. 注意此处要求T的每个元素t都要满足$t \leq s$, 即每个元素和s之间都要有二元关系. 如果T中存在这样一个元素t, 使得对于任意的$x \in T$, 只要$t \leq x$, 必然蕴含$x = t$, 我们就称t是集合T的最大元. 注意最大元和其他的元素之间不一定非要有二元关系.

定理 4.1.1 (佐恩引理)**.** *如果偏序集S中的每条链都有上界, 那么S中一定有最大元(可能不唯一).*

定理 4.1.2 (线性空间存在基)**.** *如果线性空间V的子集S中的元素是线性无关的, 则一定存在$T \subset V$, 使得T是V的一组基而且$S \subset T$.*

证明 我们考虑集合

$$\mathcal{S}=\{S\subset V|S \text{ 中的元素是线性无关的}\}.$$

这个集合中的元素在包含关系下成为一个偏序集. 由佐恩引理, 这个集合$\mathcal{S}$存在一个极大元T. 根据定义, T中的元素是线性无关的, 我们记T张成的空间为W. 则$W=V$. 否则的话存在$\alpha\in V$, 但是$\alpha\notin W$. 那么$T\cup\{\alpha\}\in\mathcal{S}$, 而且$T\cup\{\alpha\}\neq T$. 这与$T$是极大元矛盾. 所以则$W=V$. 所以$T$ 即为所求.

□

定理 4.1.3 (Steinitz替换定理)**.** *假设V是数域F 上的线性空间, α_1, ..., α_m是V中的m个线性无关的向量. 如果β_1, ..., β_n这n个向量能够张成V. 那么$m\leq n$, 而且适当的重新排列β_i的次序后, 可以使得α_1, ..., α_m, β_{m+1}, ..., β_n 仍然能够张成V. 特别的, $m\leq n$.*

证明 首先由于β_1, ..., β_n张成V, 而α_1是V中的一个向量, 所以α_1能写成β_1, ..., β_n的线性组合

$$\alpha_1=\sum_{i=1}^{n}a_i\beta_i.$$

不妨设$a_1\neq 0$, 否则的话, 适当的重新排列β_i的次序即可. 则我们可以把β_1换成α_1, 我们断言这样替换之后的α_1, β_2, ..., β_n 依然可以张成V. 这是因为$a_1\neq 0$, 所以β_1能够写成α_1, β_2, ..., β_n的线性组合. 因此任何β_1, ..., β_n的线性组合都可以改写为α_1, β_2, ..., β_n的线性组合. 所以只要β_1, ..., β_n能够张成V, α_1, β_2, ..., β_n 就可以张成V.

接下来，把α_2写成α_1, β_2, ..., β_n的线性组合$\alpha_2=b_1\alpha_1+b_2\beta_2+\cdots+b_n\beta_n$. 注意不可能$b_2=\cdots=b_n=0$, 否则$\alpha_1$ 就会和α_2线性相关了. 所以必有某个$i>1$使得$b_i\neq 0$. 不妨设$b_2\neq 0$, 则我们可以把S中的β_2换成α_2, 使得这样替换之后的α_1, α_2, β_3, ..., β_n 依然能够张成V。

依次做下去，可以看出α_1, .., α_m, β_{m+1}, ..., $\beta_n\}$是V的一个基. 定理得证. □

从Steinitz替换定理可以看出来, 如果V能由有限多个元素β_1, ..., β_n张成，那么V中最多有n个线性无关的向量. 在这种情况下, 我就称V是有限维的，它的维数是n. 否则就称V是无限维的.

推论 4.1.4. *假设V是数域F 上的有限维线性空间，α_1, ..., α_m是V的一组基, 而且β_1, ..., β_n 也是V的一组基. 则$m=n$.*

证明 由于β_1, ..., β_n也是V的一组基, 所以能张成V. 利用Steinitz替换定理可知$m\leq n$. 同理可得$n\leq m$. 所以$m=n$. □

因此维数是线性空间固有的一个不变量, 跟基的选取无关, 选取的基可以千差万别, 但是基当中的向量的个数都是一样的, 这个数目就是空间的维数. 我们有下面的定义.

定义 4.1.12. 假设V是数域F上的线性空间, $\{e_1, ..., e_n\}$是V的一个基. 则称V的维数是n, 记为$\mathrm{dim}V = n$.

在这里要注意线性空间的维数跟数域是有关系的. 如果数域改变了, 即使线性空间作为一个集合没有改变, 但是维数也会发生相应的改变. 例如复数$\mathbb{C}$是复数域上的一维线性空间. 如果我们把数域从复数域$\mathbb{C}$改成实数域$\mathbb{R}$, 虽然$\mathbb{C}$作为集合来说没有变, 但是作为线性空间来说已经改变了, 它现在是实数域$\mathbb{R}$上的2 维线性空间了.

引理 4.1.5. 假设V是数域F上的线性空间, S是V的一个向量组, W是S中的向量张成的线性空间. 那么S的任何一个极大线性无关组都是W的基. 因此S的极大线性无关组中向量的个数等于S中的向量张成的线性空间W的维数.

证明

假设W是S中的向量张成的线性空间, $\beta_1, ..., \beta_r$是W的一组基, $\alpha_1, ..., \alpha_k$是S的一个极大线性无关组. 那么由Steinitz替换定理, $r \geq k$. 如果$r > k$, 那么同样由Steinitz替换定理, 适当改变$\beta_1, ..., \beta_r$的次序, 不妨设b_r不能被$\alpha_1, ..., \alpha_k$线性表示. 由于$\alpha_1, ..., \alpha_k$是S的一个极大线性无关组, 所以S中的任何一个向量都能被$\alpha_1, ..., \alpha_k$线性表示, 所以W中任何一个向量都能被$\alpha_1, ..., \alpha_k$线性表示, 引起矛盾. 所以$r = k$.

因此S的任何一个极大线性无关组都是W的基. 因此S的极大线性无关组中向量的个数等于S中的向量张成的线性空间W的维数.

□

由上面的引理可知S的极大线性无关组中向量的个数只与向量组S有关, 和极大线性无关组的选取无关.

定义 4.1.13 (向量组的秩)**.** 假设V是数域F上的线性空间, S是V的一个向量组. 我们称S的任一个极大线性无关组中的向量个数为S的秩.

引理 4.1.6. 假设V是数域F上的线性空间, $S = \alpha_1, ..., \alpha_n$是$V$的一个线性无关向量组. 如果$\alpha$不能够被$\alpha_1, ..., \alpha_n$线性表示, 那么$\alpha_1, ..., \alpha_n, \alpha$也线性无关.

证明 如果$\alpha_1, ..., \alpha_n, \alpha$线性相关, 那么一定存在不全为零的$c_1, ..., c_n, c$使得

$$c_1\alpha_1 + \cdots + c_n\alpha_n + c\alpha = 0.$$

如果$c=0$, 那么c_1, ..., c_n全为零, 因此α_1, ..., α_n线性相关, 矛盾. 所以$c\neq 0$, 所以

$$\alpha=-c^{-1}(c_1\alpha_1+\cdots+c_n\alpha_n),$$

即α能够被α_1, ..., α_n线性表示, 所以也矛盾.

因此α_1, ..., α_n, α也线性无关. □

命题 4.1.7. *假设V是数域F上的有限维线性空间, W是V的子空间. 则W的维数都小于等于整个空间V的维数. 如果$\dim V=\dim W$, 则$V=W$. 如果$\dim V=n>\dim W=m$, 而α_1, ..., α_m是W的一组基, 那么一定存在V中的向量α_{m+1}, ..., α_n使得α_1, ..., α_n 是V的一组基.*

证明 只需要证明命题中的最后一句话, 即子空间W的任何一组基都可以扩充称V 的一组基即可. 如果子空间$W=V$, 那么根据定义二者的维数相等. 如果$W\neq V$, 假设α_1, ..., α_m是W的一组基, 那么V中一定存在向量α_{m+1}不能被α_1, ..., α_m 线性表示. 所以α_1, ..., α_m, α_{m+1}是线性无关的.

如果α_1, ..., α_m, α_{m+1}能够张成V, 那么α_1, ..., α_m, α_{m+1}就是V 的一组基. 否则可以继续找到$\alpha_{m+2},\ldots$. 由于V的维数有限, 所以这个过程有限步以后就终止了. 这时找到的α_1, ..., α_n就是V的一组基. □

从维数和基的定义里面不难看出下面的命题成立。

命题 4.1.8 (子空间和的维数公式)**.** *假设V是数域F上的有限维线性空间. 如果V_1, V_2是V的两个子空间, 则*

$$\mathrm{dim V_1}+\mathrm{dim V_2}=\mathrm{dim(V_1+V_2)}+\mathrm{dim(V_1\cap V_2)}.$$

证明 假设β_1, ..., β_k是$V_1\cap V_2$的基. 由于$V_1\cap V_2$是V_1的一个子空间, 我们可以把β_1, ..., β_k扩充为V_1的一组基α_1, ..., α_s, β_1, ..., β_k. 同样的, 我们可以把β_1, ..., β_k扩充为V_2的一组基β_1, ..., β_k, γ_1, ..., γ_t.

首先我们说明α_1, ..., α_s, β_1, ..., β_k, γ_1, ..., γ_t是线性无关的. 如果它们线性相关, 那么一定存在不全为0的a_1, ..., a_s, b_1, ..., b_k, c_1, ..., c_t, 使得

$$a_1\alpha_1+\cdots+a_s\alpha_s+b_1\beta_1+\cdots+b_k\beta_k=c_1\gamma_1+\cdots+c_t\gamma_t.$$

上式的左边在V_1中, 右边在V_2中. 因此, $c_1\gamma_1+\cdots+c_t\gamma_t\in V_1\cap V_2$. 由$\gamma_1$, ..., γ_t的取法可知, γ_1, ..., γ_t都不在$V_1\cap V_2$ 中, 因此$c_1=\cdots=c_t=0$. 同样的道理可以知道a_1, ..., a_s, b_1, ..., b_k也必须都是0. 所以α_1, ..., α_s, β_1, ..., β_k, γ_1, ..., γ_t是线性无关的.

其次我们说明α_1, ..., α_s, β_1, ..., β_k, γ_1, ..., γ_t能够张成V_1+V_2. 这是因为α_1, ..., α_s, β_1, ..., β_k 能张成V_1, β_1, ..., β_k, γ_1, ..., γ_t能够张成V_2. 所以它们合在一起能够张成V_1+V_2.

因此α_1, ..., α_s, β_1, ..., β_k, γ_1, ..., γ_t是V_1+V_2的一组基. 所以

$$\mathrm{dim(V_1+V_2)=s+k+t=dimV_1+dimV_2-dim(V_1\cap V_2)}.$$

□

定义 4.1.14 (子空间的直和)**.** 假设V是数域F上的有限维线性空间. 如果V_1, V_2是V的两个子空间. 如果$V_1\bigcap V_2=\{0\}$, 那么我们称这两个子空间的和V_1+V_2为直和, 记为

$$V_1+V_2=V_1\bigoplus V_2.$$

我们递归的定义多个子空间的直和. 假设我们已经给出了V中$k-1$个子空间的直和的定义, 而V_1, ..., V_k是V的k个子空间. 如果V_1, ..., V_{k-1}这$k-1$子空间的和是直和, 即$V_1+\cdots+V_{k-1}=V_1\bigoplus\cdots\bigoplus V_{k-1}$, 而且

$$V_k\bigcap(V_1\bigoplus\cdots\bigoplus V_{k-1})=0,$$

那么我们就称V_1, ..., V_k的和为直和, 记为

$$V_1+\cdots+V_k=V_1\bigoplus\cdots\bigoplus V_k.$$

命题 4.1.9. *假设V是数域F上的有限维线性空间, V_1, ..., V_k是V的k个子空间. 则下面几条等价.*

1. *V_1, ..., V_k的和为直和.*

2. *任取向量$\alpha_1\in V_1$, ..., $\alpha_k\in V_k$, 如果$\alpha_1+\cdots+\alpha_k=0$, 那么*

$$\alpha_1=\cdots=\alpha_k=0.$$

3. *任取向量$\alpha\in V$, 如果存在向量$\alpha_1\in V_1$, ..., $\alpha_k\in V_k$使得*

$$\alpha=\alpha_1+\cdots+\alpha_k,$$

那么这种表示法是唯一的.

4. *对任意的$1\leq i\leq k$, V_1, ..., V_{i-1}, V_{i+1}, ..., V_k的和是直和, 而且*

$$(V_1\bigoplus\cdots\bigoplus V_{i-1}\bigoplus V_{i+1}\bigoplus\cdots\bigoplus V_k)\bigcap V_i=0.$$

5. *对任意的$1\leq i\leq k$, 都有*

$$(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_k)\bigcap V_i=0.$$

6.

$$\dim V_1 + \cdots + \dim V_k = \dim(V_1 + \cdots + V_k).$$

证明 我们用数学归纳法, 当$k=2$时, 命题显然成立. 假设命题对$k-1$的情形是成立的. 我们下面证明对k的情形也成立.

"$1 \Longrightarrow 2$". 假设V_1, ..., V_k的和为直和. 如果向量$\alpha_1 \in V_1$, ..., $\alpha_k \in V_k$, 而且$\alpha_1 + \cdots + \alpha_k = 0$, 那么

$$\alpha_k = -\alpha_1 - \cdots - \alpha_{k-1} \in V_k \bigcap (V_1 \bigoplus \cdots \bigoplus V_{k-1}) = 0.$$

所以

$$\alpha_k = -\alpha_1 - \cdots - \alpha_{k-1} = 0.$$

又因为前$k-1$个子空间的和为直和, 所以$\alpha_1 = \cdots = \alpha_{k-1} = 0$.

"$2 \Longrightarrow 3$". 如果表示法不唯一. 那么还存在向量$\beta_1 \in V_1$, ..., $\beta_k \in V_k$使得

$$\alpha = \beta_1 + \cdots + \beta_k,$$

那么

$$0 = (\alpha_1 - \beta_1) + \cdots + (\alpha_k - \beta_k).$$

因为$\alpha_1 - \beta_1 \in V_1$, ..., $\alpha_k - \beta_k \in V_k$. 由2可知$\alpha_1 = \beta_1$, ..., $\alpha_k = \beta_k$.

"$3 \Longrightarrow 4$". 不妨假设$i < k$. 根据归纳假设V_1, ..., V_{i-1}, V_{i+1}, ..., V_{k-1}的和是直和, 如果

$$(V_1 \bigoplus \cdots \bigoplus V_{i-1} \bigoplus V_{i+1} \bigoplus \cdots \bigoplus V_{k-1}) \bigcap V_k \neq 0,$$

那么在它们的交中取一个非零向量α, 这时候零向量就有两种不同的表示法

$$0 = 0 + 0 = \alpha - \alpha.$$

引起矛盾. 所以V_1, ..., V_{i-1}, V_{i+1}, ..., V_k的和是直和. 如果

$$(V_1 \bigoplus \cdots \bigoplus V_{i-1} \bigoplus V_{i+1} \bigoplus \cdots \bigoplus V_k) \bigcap V_i \neq 0,$$

同样的方法可证不可能发生这样的情况.

"$4 \Longrightarrow 5$". 显然.

"$5 \Longrightarrow 6$". 首先在5 中, 取$i = k$, 由维数公式可知

$$\dim(V_1 + \cdots + V_k) = \dim(V_1 + \cdots + V_{k-1}) + \dim V_k.$$

其次由于对任意的$1 \le i \le k-1$,

$$(V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_{k-1}) \bigcap V_i \subset (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_k) \bigcap V_i = \{0\},$$

由归纳假设$\dim V_1+\cdots+\dim V_{k-1}=\dim(V_1+\cdots+V_{k-1})$. 所以

$$\dim V_1+\cdots+\dim V_k=\dim(V_1+\cdots+V_k).$$

"$6\Longrightarrow 1$". 因为

$$\begin{aligned}\dim V_1+\cdots+\dim V_k&=\dim(V_1+\cdots+V_k)\\&=\dim(V_1+\cdots+V_{k-1})+\dim V_k-\dim(V_1+\cdots+V_{k-1})\bigcap V_k,\end{aligned}$$

所以

$$\begin{aligned}\dim V_1+\cdots+\dim V_{k-1}&=\dim(V_1+\cdots+V_{k-1})-\dim(V_1+\cdots+V_{k-1})\bigcap V_k\\&\le\dim V_1+\cdots+\dim V_{k-1}-\dim(V_1+\cdots+V_{k-1})\bigcap V_k,\end{aligned}$$

所以$(V_1+\cdots+V_{k-1})\bigcap V_k=\{0\}$. 即$V_1+\cdots+V_{k-1}$与$V_i$之和为直和, 又因为

$$\dim V_1+\cdots+\dim V_{k-1}=\dim(V_1+\cdots+V_{k-}).$$

由归纳假设可知V_1, ..., V_k的和为直和. 因此V_1, ..., V_k的和为直和.

□

命题 4.1.10. 假设V是F上的n维线性空间。V_1, ...,V_k是V的真子空间。求证$V\neq V_1\bigcup\cdots\bigcup V_k$.

证明 对k归纳。假设命题对$k-1$成立。取$\alpha\in V_k$, 而且$\alpha\notin V_1\bigcup\cdots\bigcup V_{k-1}$。如果不存在这样的$\alpha$, 根据归纳假设命题成立。任取$\beta\notin V_k$. 考虑$\gamma_\lambda=\alpha+\lambda\beta$, $\lambda\in F^*$不等于零. 则$\gamma_\lambda\notin V_k$. 然而数域中有无限多个元素，所以有无限多个λ使得$\gamma_\lambda\notin V_k$。如果这无限多个都在$V_1\bigcup\cdots\bigcup V_{k-1}$中。根据鸽子笼原理，必有两个在同一个$V_i$中。假设$\gamma_{\lambda_1}$, $\gamma_{\lambda_2}\in V_i$. 所以$\frac{1}{\lambda_1}\gamma_{\lambda_1}-\frac{1}{\lambda_2}\gamma_{\lambda_2}\in V_i$. 所以$(\frac{1}{\lambda_1}-\frac{1}{\lambda_2})\alpha\in V_i$. 所以$\alpha\in V_i$. 矛盾!

所以至少有一个λ使得$\gamma_\lambda\notin V_1\bigcup\cdots\bigcup V_{k-1}\bigcup V_k$. 所以$V\neq V_1\bigcup\cdots\bigcup V_k$. □

另一种证明, 这个证明的思路来自于著名的希尔伯特零点定理。

证明 不妨设V_1, ...,V_k都是V的$n-1$维子空间。如果有限个$n-1$维子空间的并不等于V, 那么一般的真子空间都包含在某个$n-1$维子空间中，他们的并自然就更不等于V了。

我们知道$n-1$维的子空间都是n 元一次方程的解. 假设V_i是$f_i(x_1, ..., x_n)=0$的解。那么$V_1\bigcup\cdots\bigcup V_k$ 就是$f_1\cdots f_k=0$的解. 我们的命题转化成了证明$f_1\cdots f_k$不能在任何一点处都等于0, 这里有个概念上的东西需要澄清. 我们说一个多项式不等于0, 是说这个多项式的系数不都为0. 但是就空间上的多项式函数来说, 即使是非

零的多项式也有可能在这个空间上是零函数, 即取值都为0的函数. 例如如果F是二元域$\mathbb{F}_2$, 而V就是F本身. 那么

$$x^2+x$$

这个多项式在V上就是零函数, 但是这个多项式本身并不是零多项式. 在这里$f_1\cdots f_k$作为多项式来说是非零多项式没有问题, 有问题的是这个非零多项式到底是不是零函数. 从上面的例子可以看出来, 这与域本身的性质有关. 所以"$f_1\cdots f_k$不能在任何一点处都等于0"这个结论是需要证明的.

我们用数学归纳法。假设如果$n-1$个变量的多项式取值恒等于0, 则这个多项式是零多项式。

我们不妨假设多项式的乘积$f_1\cdots f_k$展开并且合并同类项以后有某个单项式包含x_1。那么$f_1\cdots f_k$ 可以看成是关于x_1的多项式，其余的变量$x_2,\ ...,\ x_n$我们都看成常数。这个多项式的最高项的系数是一个关于$x_2,\ ...,\ x_n$的多项式$g(x_2,\ ...,\ x_n)$。这个多项式不是零多项式。根据归纳假设，它取值不能恒等于0。所以存在$a_2,\ ...,\ a_n$使得$g(a_2,\ ...,\ a_n)\neq 0$.

这时$f_1(x_1,\ a_2,\ ...,\ a_n)\cdots f_k(x_1,\ a_2,\ ...,\ a_n)$就成为了关于$x_1$的非零多项式. 这个多项式当然只能有有限个解。特别的存在a_1, 使得

$$f_1(a_1,\ a_2,\ ...,\ a_n)\cdots f_k(a_1,\ a_2,\ ...,\ a_n)\neq 0.$$

□

定理 4.1.11. *假设V是n维的实空间。$W_1,\ W_2,\ ...,$是其可数多个真子空间。求证一定存在一个向量α使得它不属于任何一个W_i.*

证明 假设$\alpha_1,\ \alpha_2,\ ...,\ \alpha_n$是$V$的一组基。记

$$U=\{\gamma(k)=\alpha_1+k\alpha_2+...+k^{n-1}\alpha_n|k\in\mathbb{R}\},$$

则对任意的i,

$$U\bigcap W_i$$

是一个有限集合，元素个数小于n。（否则存在$\gamma(k_1),\ \gamma(k_2),\ ...,\ \gamma(k_n)\in W_i$，由范德蒙行列式的性质可知，$\alpha_1,\ \alpha_2,\ ...,\ \alpha_n$必须全部都在$W_i$中，与$W_i$是真子空间矛盾。）

所以

$$U\bigcap(\bigcup_i W_i)=\bigcup_i(U\bigcap W_i)$$

是可数个有限集合的并，仍然是可数的。如果$\bigcup\limits_i W_i=V$，那么$U=U\bigcap V$可数，矛盾，因为由U的定义，U的元素个数和$\mathbb{R}$相同，显然是个不可数集合.

□

§4.2 线性映射

定义 4.2.1 (线性映射)**.** 从线性空间V到W的映射$\mathscr{A}$称为是线性映射如果对任意的α, $\beta \in V$, a, $b \in F$ 都有下面的等式成立.

$$\mathscr{A}(a\alpha + b\beta) = a\mathscr{A}(\alpha) + b\mathscr{A}(\beta).$$

若$\mathscr{A}$恒等于零, 我们就称之为零映射. 如果线性空间$V = W$, 而且对于任意的向量$\alpha \in V$, 都有$\mathscr{A}(\alpha) = \alpha$. 我们就称之为是恒等映射, 记为$\mathscr{I}_V$, 或者简单的记为$\mathscr{I}$. 如果线性空间$V = W$, 而且存在常数$\lambda \in F$, 使得对于任意的向量$\alpha \in V$, 都有$\mathscr{A}(\alpha) = \lambda\alpha$, 我们就称$\mathscr{A}$是一个数乘变换. 如果$\mathscr{A}$是从线性空间$V$到$W$的线性映射, V_1是V的子空间, 那么$\mathscr{A}$在V_1上的限制

$$\mathscr{A}_{|V_1}: \ V_1 \longrightarrow W$$

也是一个线性映射.

命题 4.2.1. *假设$\mathscr{A}$是从线性空间V到W的线性映射, 则*

1. $\mathscr{A}(0) = 0$.
2. 对于任意的向量$\alpha \in V$, 都有$\mathscr{A}(-\alpha) = -\mathscr{A}(\alpha)$.
3. W中的零向量0的原像$\mathscr{A}^{-1}(0) = \{\alpha \in V | \mathscr{A}(\alpha) = 0\}$是$V$的一个子空间, 叫做$\mathscr{A}$的核, 记为$\mathrm{Ker}\mathscr{A}$.
4. $\mathscr{A}$的像集合$\{w \in W |$ 存在$\alpha \in V$ 使得$\mathscr{A}(\alpha) = w\}$是$W$的一个子空间, 叫做$\mathscr{A}$的像, 记为$\mathrm{Im}\mathscr{A}$ 或者$\mathscr{A}(V)$.
5. 如果V_1是V的子空间, $\mathscr{A}_{|V_1}$是$\mathscr{A}$在V_1上的限制, 那么$\mathrm{Ker}\mathscr{A}_{|V_1} \subseteq \mathrm{Ker}\mathscr{A}$, $\mathrm{Im}\mathscr{A}_{|V_1} \subseteq \mathrm{Im}\mathscr{A}$.

证明 留做练习.

□

定义 4.2.2 (线性映射的加法和数乘)**.** 假设$\mathscr{A}$, $\mathscr{B}$是从线性空间V到W的两个线性映射. 我们定义它们的和为

$$(\mathscr{A} + \mathscr{B})(\alpha) := \mathscr{A}(\alpha) + \mathscr{B}(\alpha).$$

对任意的常数$\lambda \in F$, 我们定义

$$(\lambda\mathscr{A})(\alpha) := \lambda\mathscr{A}(\alpha).$$

容易验证, 这样定义的加法和数乘使得从线性空间V到W的线性映射的全体称为一个线性空间.

命题 4.2.2. 假设V, W是域F上的有限维线性空间, α_1, ..., α_n是V的一组基. 那么对W中的任意n个向量β_1, ..., β_n, 都存在唯一的线性映射$\mathscr{A}$使得

$$\mathscr{A}(\alpha_1)=\beta_1,\ ...,\ \mathscr{A}(\alpha_n)=\beta_n.$$

证明 首先这样的线性映射是存在的, 对任意的向量$\alpha=t_1\alpha_1+\cdots+t_n\alpha_n$, 我们定义

$$\mathscr{A}(\alpha):=t_1\beta_1+\cdots+t_n\beta_n,$$

容易看出如此定义的映射是线性的, 而且满足条件.

其次我们说明唯一性. 如果

$$\mathscr{A}(\alpha_1)=\beta_1,\ ...,\ \mathscr{A}(\alpha_n)=\beta_n,$$

那么由线性映射的性质, 必然会导致

$$\mathscr{A}(t_1\alpha_1+\cdots+t_n\alpha_n)=t_1\beta_1+\cdots+t_n\beta_n,$$

由于V的元素写成基向量的表达形式是唯一的, 所以$\mathscr{A}$也是唯一的.

□

定义 4.2.3. 线性映射$\mathscr{A}: V\longrightarrow W$称为是同构映射(有时简单的称为同构), 如果存在线性映射$\mathscr{B}: W\longrightarrow V$ 使得$\mathscr{A}\mathscr{B}=\mathscr{I}_W$, $\mathscr{B}\mathscr{A}=\mathscr{I}_V$. 如果$\mathscr{A}$是从$V$到$W$的同构映射, 我们称$V$同构于$W$.

容易看出，如果V同构于W, 那么W也同构于V.

定理 4.2.3. 两个有限维线性空间同构当且仅当它们的维数相等.

证明 如果两个有限维的线性空间V和W同构, 即存在线性映射$\mathscr{A}: V\longrightarrow W$和$\mathscr{B}: W\longrightarrow V$使得$\mathscr{A}\mathscr{B}=\mathscr{I}_W, \mathscr{B}\mathscr{A}=\mathscr{I}_V$. 则我们知道$V$的基$\alpha_1$, ..., α_n在W中的像$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_n)$ 一定是线性无关的. 因为如果不是这样, 那么一定存在不全为零的常数c_1, ..., c_n使得

$$c_1\mathscr{A}(\alpha_1)+\cdots+c_n\mathscr{A}(\alpha_n)=0.$$

两边用$\mathscr{B}$作用可得

$$c_1\mathscr{B}\mathscr{A}(\alpha_1)+\cdots+c_n\mathscr{B}\mathscr{A}(\alpha_n)=0.$$

但是$\mathscr{B}\mathscr{A}(\alpha_1)=\alpha_1$, ..., $\mathscr{B}\mathscr{A}(\alpha_n)=\alpha_n$. 所以

$$c_1\alpha_1+\cdots+c_n\alpha_n=0.$$

这与α_1, ..., α_n线性无关矛盾. 所以$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_n)$一定是线性无关的. 即W中有n个线性无关的向量. 所以W的维数大于等于V的维数. 反之亦然, 因此两个空间的维数相等.

如果V和W的维数相等, 那么假设α_1, ..., α_n为V的基, β_1, ..., β_n为W的基. 根据上面的命题, 存在唯一的线性映射$\mathscr{A}$使得

$$\mathscr{A}(\alpha_1)=\beta_1,\ ...,\ \mathscr{A}(\alpha_n)=\beta_n,$$

也存在唯一的线性映射$\mathscr{B}$使得

$$\mathscr{B}(\beta_1)=\alpha_1,\ ...,\ \mathscr{B}(\beta_n)=\alpha_n.$$

显然$\mathscr{A}\mathscr{B}=\mathscr{I}_W$, $\mathscr{B}\mathscr{A}=\mathscr{I}_V$. 所以$V$同构于$W$.

□

定义 4.2.4. *假设V有两组基$\{\alpha_1,\ ...,\ \alpha_n\}$, $\{\beta_1,\ ...,\ \beta_n\}$. 我们知道这两组基作为向量组来说是等价的, 即存在一个可逆矩阵$C=(c_{ij})$使得*

$$(\beta_1,\ ...,\ \beta_n)=(\alpha_1,\ ...,\ \alpha_n)C.$$

上面这个式子具体写下来就是

$$\begin{cases}\beta_1 & =c_{11}\alpha_1+\cdots+c_{n1}\alpha_n,\\ & \vdots\\ \beta_n & =c_{1n}\alpha_1+\cdots+c_{nn}\alpha_n.\end{cases}$$

我们称这个矩阵C为从基$\{\alpha_1,\ ...,\ \alpha_n\}$到基$\{\beta_1,\ ...,\ \beta_n\}$的过渡矩阵. 任给$V$的一组基$\{\alpha_1,\ ...,\ \alpha_n\}$, 和一个可逆矩阵$C=(c_{ij})$. 我们也可以通过上面的公式给出$V$的另外一组基$\{\beta_1,\ ...,\ \beta_n\}$.

假设V有两组基$\{\alpha_1,\ ...,\ \alpha_n\}$, $\{\beta_1,\ ...,\ \beta_n\}$. 则下面的线性映射

$$\begin{aligned}\mathscr{A}:\ V &\longrightarrow \quad V\\ \sum x_i\alpha_i &\longmapsto \quad \sum x_i\beta_i\end{aligned}$$

是同构。满射和单射都很显然。

定理 4.2.4 (维数公式)**.** *从有限维线性空间V到线性空间W的线性映射$\mathscr{A}$满足*

$$\dim(\mathrm{Ker}(\mathscr{A}))+\dim(\mathrm{Im}(\mathscr{A}))=\mathrm{dimV}$$

证明 选$\alpha_,$ ..., $\alpha_r \in V$使得$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$ 为$\mathrm{Im}(\mathscr{A})$的一组基。设$\alpha_,$ ..., α_r生成的子空间是V_1, 则我们宣称$V = V_1 \bigoplus \mathrm{Ker}(\mathscr{A})$.

这相当于要证明两条, 一是$V = V_1 + \mathrm{Ker}(\mathscr{A})$, 二是$V_1 \cap \mathrm{Ker}(\mathscr{A}) = \{0\}$.

我们先来证明第一条. 任取$\alpha \in V$, 因为$\mathscr{A}(\alpha) \in \mathscr{A}(V)$, 所以存在$\widetilde{\alpha} \in V_1$ 使得$\mathscr{A}(\alpha) = \mathscr{A}(\widetilde{\alpha})$. 所以$\mathscr{A}(\alpha - \widetilde{\alpha}) = 0$. 因此$\alpha - \widetilde{\alpha} \in \mathrm{Ker}(\mathscr{A})$. 所以$V = V_1 + \mathrm{Ker}(\mathscr{A})$.

第二条易证.

□

下面我们看一些线性映射的例子

例 4.2.1. 令$\mathscr{A}$: $\mathbb{R}^4 \longrightarrow \mathbb{R}^2$为下面的映射

$$\mathscr{A}(\alpha) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \alpha,$$

这里的α是$\mathbb{R}^4$中的一个四维向量. 容易验证这是一个线性映射. 如果$\alpha = (a, b, c, d)'$, 那么

$$\mathscr{A}(\alpha) = (a + 2b + 3c + 4d,\ 5a + 6b + 7c + 8d)'.$$

容易看出$\mathscr{A}$的像空间是

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

由A的列向量张成的, 而$\mathscr{A}$的核空间是线性方程组

$$AX = 0$$

的解空间.

例 4.2.2 (射影变换)**.** 假设V是一个线性空间, W是V的子空间. 我们知道一定存在另外一个子空间U(不唯一)使得

$$V = U \bigoplus W.$$

所以任何一个向量$\alpha \in V$都可以唯一地写成

$$\alpha = \alpha_W + \alpha_U,\ \alpha_W \in W,\ \alpha_U \in U$$

的形式. 我们称映射

$$V \longrightarrow V, \alpha \mapsto \alpha_W$$

为从V到W的射影. 注意这个射影和子空间U的选取有关系, 如果U选的是另外一个. 那么这个射影也会相应的改变.

我们在上一章学过, 线性方程组

$$Ax = b$$

有解当且仅当A的秩r等于增广矩阵$\begin{pmatrix} A & b \end{pmatrix}$的秩. 这时, 存在$n$维列向量$\beta, \alpha_1, ..., a_{n-r}$, 使得线性方程组的通解为

$$(x_1,\ x_2,\ ...,\ x_n)^T = \beta + t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$$

的形式. 其中β是一个特解, $t_1\alpha_1 + \cdots + t_{n-r}\alpha_{n-r}$是线性齐次方程组$Ax = 0$的通解. 我们称$\alpha_1,\ ...,\ a_{n-r}$为线性其次方程组$Ax = 0$的一个基础解系.

我们现在也可以用线性映射的语言加以重新阐述.

令$V = F^n$为n维列线性空间, $W = F^m$为m维线性空间, 假设A是一个$m \times n$的矩阵. 我们定义映射

$$\begin{aligned} \mathscr{A} : V \quad &\longrightarrow W, \\ \alpha \quad &\longmapsto A\alpha. \end{aligned}$$

很容易验证这是一个线性映射. 求$Ax = 0$的通解等价于求$\mathscr{A}$的核.

把A的列记为$\gamma_1,\ ...,\ \gamma_n$, 那么$A\alpha$是$\gamma_1,\ ...,\ \gamma_n$的线性组合, 所以$\mathscr{A}$的像空间就是$\gamma_1,\ ...,\ \gamma_n$张成的线性空间. 这个空间的秩等于$A$ 的列秩, 也就是A的秩r. 所以根据维数公式, 求$\mathscr{A}$的核是一个$n - r$维的线性空间. 这个线性空间的任何一组基都叫做是$Ax = 0$的一个基础解系.

线性方程组

$$Ax = b$$

的解其实就是列向量b的原像. 我们知道b的原像不是空集当且仅当$b \in Im(\mathscr{A})$, 即b在A的像空间中, 当且仅当b在$\gamma_1,\ ...,\ \gamma_n$ 张成的线性空间中, 即b能够写成A的列向量的线性组合. 而b能够写成A的列向量的线性组合当且仅当向量组$\gamma_1, ..., \gamma_n$的秩和$\gamma_1, ..., \gamma_n, b$的秩相等. 因此$b$的原像不是空集当且仅当系数矩阵的秩等于增广矩阵的秩.

§4.3 线性映射对应的矩阵

假设线性空间V有一组基是$\alpha_1, ..., \alpha_n$, 线性空间W有一组基是$\beta_1, ..., \beta_m$. 令V_1为从V 到W的线性映射全体. 令$V_2 = F^{m\times n}$为F上所有的$m \times n$矩阵全体. 则$V_1,\ V_2$都是F上的线性空间. 任取α_j, 则一定存在$a_{1j},\ ...,\ a_{mj}$ 使得

$$\mathscr{A}(\alpha_j) = a_{1j}\beta_1 + \cdots + a_{mj}\beta_m.$$

我们按照如下的方式定义一个从V_1到V_2的映射:

$$\begin{aligned} \varphi:\ V_1 &\longrightarrow V_2, \\ \mathscr{A} &\longmapsto A = (a_{ij}) \end{aligned}$$

容易看出来这是一个线性映射, 而且是个双射, 所以V_1, V_2是F上的两个同构的线性空间. 由于$V_2 = F^{m\times n}$ 是F上所有的$m\times n$矩阵集合, 它的维数是mn. 因此V_1的维数也是mn.

我们可以把线性映射$\mathscr{A}$在基向量上α_1, ..., α_n上的的作用简单的记为

$$\mathscr{A}(\alpha_1,\ ...,\ \alpha_n) = (\beta_1,\ ...,\ \beta_m)A.$$

我们称A为在给定线性空间V的基α_1, ..., α_n和线性空间W的基β_1, ..., β_m 下, 线性映射$\mathscr{A}$ 对应的矩阵, $\mathscr{A}$为A对应的线性映射.

定理 4.3.1. 设$\mathscr{A}$一个为从线性空间V到W的线性映射. 那么一定存在V的一组基α_1, ..., α_n, 和线性空间W的一组基β_1, ..., β_m, 使得$\mathscr{A}$在这两组基下的矩阵为

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

证明 取β_1, ..., β_r为$Im(\mathscr{A})$的一组基, 然后扩充为W的一组基β_1, ..., β_m. 再取α_1, ..., $\alpha_r \in V$ 使得

$$\mathscr{A}(\alpha_1) = \beta_1,\ ...,\ \mathscr{A}(\alpha_r) = \beta_r.$$

然后再取α_{r+1}, ..., α_n为$\ker\mathscr{A}$的一组基. 这时α_1, ..., α_n就是V的一组基. 而且$\mathscr{A}$在这两组基下的矩阵为

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

□

定理 4.3.2. 假设A是数域F上的$m\times n$矩阵, 那么一定存在一个m阶的可逆矩阵P和一个n阶的矩阵Q使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

证明 令$V = F^n$为n维的列线性空间, $W = F^m$为m维的列线性空间. 我们给V取一组标准基e_1, ..., e_n, 其中e_i 是第i个位置为1, 其余的位置为0 的n维列向量. 同样的我们给W也取一组标准基f_1, ..., f_m. 那么我们可以构造一个从V到W的线性映射

$$\begin{aligned} \mathscr{A}:\ V \longrightarrow&\quad W, \\ \alpha \mapsto&\ A\alpha, \end{aligned}$$

即

$$\mathscr{A}(e_1,\ ...,\ e_n) = (f_1,\ ...,\ f_m)A.$$

容易看出这确实是一个线性映射. 根据上面的定理, 我们知道一定存在V的一组基α_1, ..., α_n, 和线性空间W的一组基β_1, ..., β_m, 使得$\mathscr{A}$在这两组基下的矩阵为

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

即

$$\mathscr{A}(\alpha_1, ..., \alpha_n) = (\beta_1, ..., \beta_m)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

我们假设从β_1, ..., β_m到f_1, ..., f_m的过渡矩阵是P, 即

$$(f_1, ..., f_m) = (\beta_1, ..., \beta_m)P,$$

从e_1, ..., e_n到α_1, ..., α_n的过渡矩阵是Q, 即

$$(\alpha_1, ..., \alpha_n) = (e_1, ..., e_n)Q,$$

或者等价地

$$(\alpha_1, ..., \alpha_n)Q^{-1} = (e_1, ..., e_n).$$

因此

$$\mathscr{A}(e_1, ..., e_n) = (f_1, ..., f_m)A.$$

等价于

$$\mathscr{A}(\alpha_1, ..., \alpha_n)Q^{-1} = (\beta_1, ..., \beta_m)PA,$$

即

$$\mathscr{A}(\alpha_1, ..., \alpha_n) = (\beta_1, ..., \beta_m)PAQ.$$

但是我们知道

$$\mathscr{A}(\alpha_1, ..., \alpha_n) = (\beta_1, ..., \beta_m)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

所以

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

□

上面这个定理, 我们以前使用矩阵的初等变换的语言来证明的. 现在我们是用线性映射的语言来证明的. 二者实质上式等价的, 只是所用的语言不同. 从上面的讨论中, 我们可以看出来矩阵的秩其实是矩阵所对应的线性映射的像的维数.

推论 4.3.3. 假设$\mathscr{A}$是从V_1到V_2的一个线性映射, α_1, ..., α_n是V_1的基, β_1, ..., β_m是V_2的基, $\mathscr{A}$在这两组基下的矩阵为A. 则$\mathscr{A}$是满射当且仅当A 是行满秩的, $\mathscr{A}$是单射当且仅当A是列满秩的.

证明 $\mathscr{A}$是满射当且仅当像的维数等于m, 当且仅当矩阵A的秩等于m, 而A恰有m行, 所以这当且仅当A是行满秩的. 类似的, 也可以证明$\mathscr{A}$是单射当且仅当A是列满秩的.

□

假设$\mathscr{A}$是从V_1到V_2的一个线性映射, $\mathscr{B}$是从V_2到V_3的一个线性映射, 那么容易验证复合映射

$$\mathscr{B}\mathscr{A}:\ V_1 \longrightarrow V_3$$

也是一个线性映射. 如果$\alpha_1,\ ...,\ \alpha_n$是$V_1$的基, $\beta_1,\ ...,\ \beta_m$是V_2的基, $\gamma_1,\ ...,\ \gamma_k$是$V_3$的基. 如果在这些取定了的基下,

$$\mathscr{A}(\alpha_1,\ ...,\ \alpha_n) = (\beta_1,\ ...,\ \beta_m)A$$

$$\mathscr{B}(\beta_1,\ ...,\ \beta_m) = (\gamma_1,\ ...,\ \gamma_k)B$$

那么复合映射$\mathscr{B}\mathscr{A}$对应的矩阵为

$$\mathscr{B}\mathscr{A}(\alpha_1,\ ...,\ \alpha_n) = (\gamma_1,\ ...,\ \gamma_k)BA$$

.

因此线性映射的复合对应于矩阵的乘法. 下面我们利用线性映射的语言来证明两个矩阵秩的不等式.

定理 4.3.4 (Sylvester秩不等式)**.** *假设A是$m\times n$的矩阵, B是$n\times k$的矩阵. 求证*

$$r(AB) \geq r(A) + r(B) - n.$$

证明 假设$V_1 = F^k$, $V_2 = F^n$, $V_3 = F^m$分别为$k,\ n,\ m$维的线性空间. 我们给这3个线性空间各自选定一组标准基. 那么A就决定一个线性映射

$$\mathscr{A}:\ V_2 \longrightarrow V_3.$$

B 也决定了一个线性映射

$$\mathscr{B}:\ V_1 \longrightarrow V_2.$$

我们可以用下面的图表来表示

$$V_1 \xrightarrow{\mathscr{B}} V_2 \xrightarrow{\mathscr{A}} V_3.$$

所以

$$r(AB) = \dim(\mathscr{A}\mathscr{B}(V_1)) = \dim(\mathscr{A}(\mathscr{B}(V_1))).$$

我们考虑限制映射

$$\mathscr{A}_{|\mathscr{B}(V_1)}:\ \mathscr{B}(V_1) \longrightarrow V_3,$$

显然有

$$\mathscr{A}_{|\mathscr{B}(V_1)}(\mathscr{B}(V_1)) = \mathscr{A}\mathscr{B}(V_1).$$

所以我们有

$$\begin{aligned} r(AB) &= \dim(\mathscr{A}\mathscr{B}(V_1)) \\ &= \dim(\mathscr{A}_{|\mathscr{B}(V_1)}(\mathscr{B}(V_1))) \\ &= \dim(\mathscr{B}(V_1)) - \mathrm{Ker}(\mathscr{A}_{|\mathscr{B}(V_1)}) \\ &\geq \dim(\mathscr{B}(V_1)) - \mathrm{Ker}(\mathscr{A}) \\ &= r(B) - (n - r(A)) \\ &= r(B) + r(A) - n. \end{aligned}$$

□

定理 4.3.5 (Frobenius秩不等式)**.** 假设A是$m \times n$的矩阵, B是$n \times k$的矩阵, C是$k \times \ell$的矩阵. 求证

$$r(ABC) \geq r(AB) + r(BC) - r(B).$$

证明 仿照上面, 我们可以构造线性映射

$$V_1 \xrightarrow{\mathscr{D}} V_2 \xrightarrow{\mathscr{B}} V_3 \xrightarrow{\mathscr{A}} V_4.$$

令$V = \mathscr{B}(V_2)$, 则我们有下面的图表

$$V_1 \xrightarrow{\mathscr{B}\mathscr{D}} V \xrightarrow{\mathscr{A}_{|V}} V_4.$$

由上面的Sylvester秩不等式可知

$$\begin{aligned} r(ABC) &= \dim(\mathscr{A}\mathscr{B}\mathscr{D}(V_1)) \\ &= \dim(\mathscr{A}_{|V}\mathscr{B}\mathscr{D}(V_1)) \\ &\geq \dim(\mathscr{B}\mathscr{D}(V_1)) + \dim(\mathscr{A}_{|V}(V)) - \dim(V) \\ &= \dim(\mathscr{B}\mathscr{D}(V_1)) + \dim(\mathscr{A}\mathscr{B}(V_2)) - \dim(\mathscr{B}(V_2)) \\ &= r(BC) + r(AB) - r(B). \end{aligned}$$

□

定理 4.3.6. 任何一个矩阵都可以写成一个列满秩矩阵乘一个行满秩矩阵.

证明 [证明一] 根据上面的讨论, 该定理相当于说任何一个线性映射$\mathscr{A}: V \longrightarrow W$都可以分解为先是一个满射, 再接一个单射. 这是显然的. 因为$\mathscr{A}$可以分解为

$$\mathscr{A}: V \longrightarrow Im(\mathscr{A}) \longrightarrow W.$$

□

证明 [证明二]

$$A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = P\begin{pmatrix} I_r \\ 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \end{pmatrix} Q = UV,$$

其中

$$U = P\begin{pmatrix} I_r \\ 0 \end{pmatrix}$$

是可逆矩阵P的前r列, 所以是列满秩的,

$$V = \begin{pmatrix} I_r & 0 \end{pmatrix} Q$$

是可逆矩阵Q的前r行, 所以是行满秩的.

□

定理 4.3.7. 设数域F上的$m \times n$矩阵。A, B和$A+B$的秩分别是r, s和$r+s \leq min\{m, n\}$。证明存在可逆矩阵P, Q 使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad PBQ = \begin{pmatrix} 0 & 0 \\ 0 & I_s \end{pmatrix}.$$

这里的I_r, I_s分别表示r, s阶的单位方阵。

证明 设V, W分别是n, m维的线性空间，随便取它们的基，则A, B, $A+B$决定了$V \longrightarrow W$的三个线性映射$\mathscr{A}$, $\mathscr{B}$, $\mathscr{A}+\mathscr{B}$。只需证明存在V, W的新的基$\{\alpha_1, .., \alpha_n\}$, $\{\beta_1, .., \beta_m\}$，使得$\mathscr{A}$, $\mathscr{B}$在这组新的基下的矩阵为

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 \\ 0 & I_s \end{pmatrix}$$

即可。P就是从W的旧的基到新的基$\{\beta_1, ..., \beta_m\}$ 的过渡矩阵，Q是从V的新的基$\{\alpha_1, ..., \alpha_n\}$到旧的基的过渡矩阵。

设$\mathscr{A}$, $\mathscr{B}$, $\mathscr{A}+\mathscr{B}$的像空间分别是W_1, W_2, W_3，其维数分别是r, s, $r+s$。因为$W_3 \subset W_1 + W_2$，所以$W_1 + W_2$是直和，即W_1, W_2的交集为0。下面就简单了. 设$\mathscr{A}$的核空间为V_1, $\mathscr{B}$的核空间为V_2, $\mathscr{A}+\mathscr{B}$的核空间为V_3. 由于$V_1 \cap V_2 \subset V_3$, 所以只要能说明右边的维数小于等于左边的维数，就可以说明$V_3 = V_1 \cap V_2$. 而这一点不难加以说明. 因为$\dim(V_3) = n - r - s$. 而

$$\begin{aligned} \dim(V_1 \cap V_2) &= \dim(V_1) + \dim(V_2) - \dim(V_1 + V_2) \\ &\geq n + r + n - s - n \\ &= n - r - s. \end{aligned}$$

所以$V_3 = V_1 \cap V_2 \subset V_3$. 而且从中可以看出$V_1 + V_2 = V$.

所以取V_3的基α_{r+1}, ..., α_{n-s}，向前扩为V_2的基α_1, ..., α_{n-s}, 向后扩为V_1的基α_{r+1}, ..., α_n. 这些向量α_1, ..., α_n构成了V的基.

又注意到$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$, $\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$ 必然线性无关. 在这里我们简单的说明一下为什么$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$, $\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$ 是线性无关的. 首先$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$是线性无关的, 否则α_1, ..., α_r将会有一个非平凡的线性组合是在V_1中, 矛盾. 同样的道理$\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$也是线性无关的. 如果$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$的一个非平凡的线性组合等于$\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$的一个非平凡的线性组合, 那么这将会制造出$W_1 \cap W_2$中的非零向量, 也会引起矛盾. 所以$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$, $\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$ 必然线性无关. 这些向量还无法构成W的基, 还需要进行扩成才行. 我们把$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}(\alpha_r)$排在最前面, 记为β_1, ..., β_r, 把$\mathscr{B}(\alpha_{n-s+1})$, ..., $\mathscr{B}(\alpha_n)$排在最后面, 记为β_{m-s+1}, ..., β_m. 把扩充出来的那些向量β_{r+1}, ..., β_{m-s}排在中间. 这些向量β_1, ..., β_m合在一起构成W的基.

则$\mathscr{A}$, $\mathscr{B}$在这个新的基下的矩阵为

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 \\ 0 & I_s \end{pmatrix}.$$

□

习题 4.3

1. 证明d维的线性空间中任意的$d+1$个向量是线性相关的。

2. 假设A是域F上的n阶矩阵, 求证存在一个域F上的非零多项式$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ 使得

$$f(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I_n = 0$$

证明 把所有的n阶矩阵看做是一个线性空间, 这个空间的维数等于n^2, 因此, 任取n^2+1个元素一定是线性相关的. 我们可以取这n^2+1 个元素为

$$A^{n^2},\ \ldots,\ A,\ I_n.$$

则一定存在不全为零的常数a_{n^2}, a_{n^2-1}, ..., a_0使得

$$a_{n^2}A^{n^2} + a_{n^2-1}A^{n^2-1} + \cdots + a_1 A + a_0 I_n = 0.$$

假设a_{n^2}, a_{n^2-1}, ..., a_0中从左向右数, 第一个不等于零的数是a_m, 则令

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

即为符合要求的多项式.

注意此处我们找到的多项式有可能是n^2次的. 我们后面要学到的Hamiton-Cayley定理说的是这个可以找到一个次数不超过n的多项式满足题目的要求. □

3. 假设A是一个n阶方阵, α是一个n维列向量, $A^{r+1}\alpha=0$, 但是$A^r\alpha\neq 0$. 求证

$$\alpha,\ A\alpha,\ ...,\ A\alpha^r$$

线性无关.

证明 假设

$$\alpha,\ A\alpha,\ ...,\ A\alpha^r$$

线性相关. 那么存在不全是零的常数c_0, ..., c_r使得

$$c_0\alpha+c_1A\alpha+\cdot+c_rA\alpha^r=0.$$

两边用A^r来作用, 可以得到$c_0=0$, ... 因此上面那个等式就变成了

$$c_1A\alpha+\cdot+c_rA\alpha^r=0.$$

两边再用A^{r-1}作用, 得到$c_1=0$. 依次做下去, 可以知道每一个系数都是0, 这就得到了矛盾. □

4. 设$A=\begin{pmatrix}1&0&0\\0&2&0\\0&1&2\end{pmatrix}$, 另$C(A)=\{B\in P^{3\times 3}|AB=BA\}$. 则

（1）证明：$C(A)$是$P^{3\times 3}$的子空间；

（2）求$C(A)$的维数和一组基.

证明 (1)证明是子空间很简单.

(2) 假设

$$B=\begin{pmatrix}a&b&c\\d&e&f\\h&i&j\end{pmatrix},$$

则

$$AB=\begin{pmatrix}a&b&c\\2d&2e&2f\\d+2h&e+2i&f+2j\end{pmatrix},\ BA=\begin{pmatrix}a&2b+c&2c\\d&2e+f&2f\\h&2i+j&2j\end{pmatrix}$$

由$AB=BA$可知, $b=c=d=h=f=0$, $e=j$. 所以自由变量只有三个, 所以维数是3.

$\square$

第五章　相似和Jordan标准型

§5.1　矩阵的相似

定义 5.1.1 (矩阵的特征多项式)**.** 假设A是域F上的n阶方阵. 行列式

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

是一个关于λ的多项式, 称为是A的特征多项式.

容易看出, 相似的矩阵一定是等价的, 也就是它们的秩相等. 而且我们后面会证明任何一个矩阵都相似于它的转置. 合同的矩阵一定也是等价的. 相似的矩阵不一定是合同的. 我们可以举反例说明例如

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

我们知道A和B是相似的, 实际上

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

但是由于A是对称矩阵, 所以和它合同的矩阵一定也是对称的. 但是由于B不是对称的, 所以A 不合同于B.

我们假设$f(\lambda) = |\lambda I - A|$是$A$的特征多项式. 根据行列式的性质可知

$$|\lambda I - A| = \lambda^n - s_1\lambda^{n-1} + \cdots + (-1)^n s_n.$$

其中s_k是A的所有的k阶主子式的和. 特别的, s_1就是矩阵A的迹, 而s_n是矩阵的行列式. 在研究矩阵的相似和Jordan标准型时, 特征多项式是一个重要的东西.

定义 5.1.2. 假设A是数域F上的一个n阶矩阵, $f(x) = a_0x^m + \cdots + a_{m-1}x + a_m \in F[x]$. 那么我们定义

$$f(A) = a_0A^m + \cdots + a_{m-1}A + a_mI_n.$$

定理 5.1.1 (Hamilton-Cayley定理)**.** 假设A是域F 上的n阶方阵,

$$f(\lambda) = |\lambda I - A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$$

是A的特征多项式. 那么

$$f(A) = A^n + a_1A^{n-1} + \cdots + a_nI = 0$$

是个零矩阵.

证明 记$B(\lambda) = (\lambda I - A)^*$为$\lambda I - A$的伴随矩阵. 那么$B(\lambda)$的每一个元素都是$\lambda I - A$ 的一个$n-1$阶的子式, 所以是一个次数不超过$n-1$的关于λ的多项式. 我们假设

$$B(\lambda) = (b_{ij}(\lambda)) = \lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + B_{n-1}.$$

其中B_0, ..., B_{n-1}都是F上的数字矩阵. 已知$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$, 所以则

$$f(\lambda)I = \lambda^n I + a_1\lambda^{n-1}I + \cdots + a_nI.$$

而

$$\begin{aligned}(\lambda I - A)B(\lambda) &= (\lambda I - A)(\lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + B_{n-1})\\ &= \lambda^n B_0 + \lambda^{n-1}(B_1 - AB_0) + \cdots + \lambda(B_{n-1} - AB_{n-1}) - AB_{n-1}\end{aligned}$$

同时由于$B(\lambda) = (\lambda I - A)^*$, 所以

$$(\lambda I - A)B(\lambda) = (\lambda I - A)(\lambda I - A)^* = f(\lambda)I = \lambda^n I + a_1\lambda^{n-1}I + \cdots + a_nI.$$

所以比较上面的两个式子可得

$$\begin{cases} B_0 = I, \\ B_1 - AB_0 = a_1I \\ B_2 - AB_1 = a_2I \\ \cdots \\ B_{n-1} - AB_{n-2} = a_{n-1}I \\ -AB_{n-1} = a_nI. \end{cases}$$

用A^n从左边乘以上面的第一个式子, ..., 用A从左边乘以上面的第n个式子(就是倒数第二个式子), 最后一个式子保持不动. 然后把这$n+1$个式子相加得到

$$f(A) = 0.$$

□

另外一种证明方法

证明 假设

$$\lambda I - A = (p_{ij}(\lambda)),\ i.e.,\ p_{ij}(\lambda) = \begin{cases} x - a_{ij}, & i = j; \\ -a_{ij}, & i \neq j. \end{cases}$$

假设e_i是F^n中的标准向量. 则

$$\sum_{j=1}^{n} p_{ij}(A)e_j = (A - a_{ii}I)e_i + \sum_{j\neq i}^{n}(-a_{ij})e_j = 0,$$

即零向量.

假设

$$Q = (q_{ij}(\lambda)) = (\lambda I - A)^*$$

即$\lambda I - A$的伴随矩阵. 则

$$\sum_{i=1}^{n} q_{ki}(A) \sum_{j=1}^{n} p_{ij}(A)e_j = 0.$$

所以

$$\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ki}(A)p_{ij}(A)e_j = 0.$$

因为

$$Q(\lambda I - A) = f(\lambda)I,$$

所以

$$\sum_{i=1}^{n} q_{ki}(A)p_{ij}(A) = \delta_{kj}f(A).$$

所以$f(A)e_j = 0$ 对任意j成立. 所以$f(A) = 0$. □

定理 5.1.2 (Gershgorin圆盘定理)**.** 假设$A = (a_{ij})$是一个复矩阵, 对每个$i \in \{1, ..., n\}$, 定义

$$R_i = \sum_{j\neq i} |a_{ij}|.$$

则A的任意特征值λ都位于下面的n个闭圆盘之并,

$$\lambda \in \bigcup_{i=1}^{n}\{z \in \mathbb{C} | |z - a_{ii}| \leq |R_i|\}$$

证明 因为λ是A的特征值, 所以

$$|\lambda I - A| = 0.$$

所以$|\lambda I - A$不是主角占优的矩阵, 所以由Levy - Desplanques定理, 必然有某个$i \in \{1, ..., n\}$满足

$$\lambda \in \{z \in \mathbb{C} | |z - a_{ii}| \leq |R_i|\}$$

. 所以命题成立.

□

命题 5.1.3. 假设A是一个$m\times n$的矩阵, B是一个$n\times m$的矩阵, $m\geq n$, 则

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

证明 我们知道

$$\begin{aligned}
\lambda^n|\lambda I_m - AB| &= \begin{vmatrix} \lambda I_m - AB & A \\ 0 & \lambda I_n \end{vmatrix} \\
&= \begin{vmatrix} \lambda I_m - AB & A \\ 0 & \lambda I_n \end{vmatrix} \begin{vmatrix} I_m & 0 \\ B & I_n \end{vmatrix} \\
&= \begin{vmatrix} \lambda I_m & A \\ B & \lambda I_n \end{vmatrix} \\
&= \begin{vmatrix} I_m & 0 \\ -B & I_n \end{vmatrix} \begin{vmatrix} \lambda I_m & A \\ B & \lambda I_n \end{vmatrix} \\
&= \begin{vmatrix} \lambda I_m & A \\ 0 & \lambda I_n - BA \end{vmatrix} \\
&= \lambda^m|\lambda I_n - BA|.
\end{aligned}$$

命题得证. □

定义 5.1.3. 假设A, B是域F上的两个n阶方阵. 如果存在一个可逆矩阵C, 使得

$$C^{-1}AC = B,$$

我们就称A相似于B, 或者是A和B是相似的. 容易看出来, 相似是个等价关系.

定理 5.1.4. 相似的矩阵具有相同的特征多项式.

证明 假设A, B是域F上的n阶方阵. 如果A相似于B, 则存在一个可逆矩阵C, 使得

$$C^{-1}AC = B.$$

所以我们有

$$|\lambda I - A| = |C^{-1}(\lambda I - A)C| = |\lambda I - C^{-1}AC| = |\lambda I - B|.$$

□

注意具有相同的特征多项式的矩阵不一定是相似的. 例如令

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

则A和B的特征多项式都是

$$f(\lambda)=\lambda^2.$$

但是A和B的秩不相等, 所以它们不相似.

定义 5.1.4 (线性空间V上的线性变换)**.** 数域F 上的线性空间V到自身的线性映射叫做V上的线性变换.

定义 5.1.5 (线性变换对应的矩阵)**.** 假设$\mathscr{A}$是n维线性空间V上的线性变换, $\{\alpha_1, ..., \alpha_n\}$是$V$的一组基. 假设

$$\begin{cases}\mathscr{A}(\alpha_1) & =a_{11}\alpha_1+\cdots+a_{n1}\alpha_n,\\ & \vdots \\ \mathscr{A}(\alpha_n) & =a_{1n}\alpha_1+\cdots+a_{nn}\alpha_n,\end{cases}$$

则我们称$\mathscr{A}$在基$\{\alpha_1, ..., \alpha_n\}$之下对应的矩阵是$A$, 记为

$$\mathscr{A}(\alpha_1, ..., \alpha_n)=(\alpha_1, ..., \alpha_n)A.$$

定义 5.1.6 (矩阵对应的线性变换)**.** 假设V是n维线性空间, $\{\alpha_1, ..., \alpha_n\}$是$V$的一组基. 假设$A$是$F$上$n$阶矩阵.

线性变换

$$\begin{aligned}\mathscr{A}: V &\longrightarrow \quad V,\\ \alpha=\sum x_i\alpha_i &\mapsto A\end{aligned}$$

称为是矩阵A对应的线性变换.

容易看出如果$\mathscr{A},\mathscr{B}$是n维线性空间V上的两个线性变换, $\{\alpha_1, ..., \alpha_n\}$是$V$的一组基. 如果$\mathscr{A}$在基$\{\alpha_1, ..., \alpha_n\}$之下对应的矩阵是$A$, $\mathscr{B}$在这组基之下对应的矩阵是B, 那么$\mathscr{A}+\mathscr{B}$ 在这组基下对应的矩阵是$A+B$, $\mathscr{A}\mathscr{B}$在这组基下对应的矩阵是AB.

定义 5.1.7. 假设$\mathscr{A}$是n维线性空间V 上的线性变换, $f(x)=a_0x^m+\cdots+a_{m-1}x+a_m\in F[x]$. 那么我们定义

$$f(\mathscr{A})=a_0\mathscr{A}^m+\cdots+a_{m-1}\mathscr{A}+a_m\mathscr{I}$$

定理 5.1.5. 假设$\mathscr{A}$是n维线性空间V上的线性变换, $\{\alpha_1, ..., \alpha_n\}$是$V$的一组基. 假设$\mathscr{A}$在这组基下对应的矩阵是$A$, 即

$$\mathscr{A}(\alpha_1, ..., \alpha_n)=(\alpha_1, ..., \alpha_n)A.$$

假设$\{\beta_1, ..., \beta_n\}$是$V$的另一组基. 假设$\mathscr{A}$在这组基下对应的矩阵是$B$, 即

$$\mathscr{A}(\beta_1, ..., \beta_n) = (\beta_1, ..., \beta_n)B.$$

则A, B是相似的. 如果两组基之间的过渡矩阵是C, 即

$$(\beta_1, ..., \beta_n) = (\alpha_1, ..., \alpha_n)C,$$

则

$$C^{-1}AC = B.$$

证明

$$\begin{aligned}\mathscr{A}(\beta_1, ..., \beta_n) &= \mathscr{A}((\alpha_1, ..., \alpha_n)C)\\ &= \mathscr{A}(\alpha_1, ..., \alpha_n)C\\ &= (\alpha_1, ..., \alpha_n)AC\\ &= (\beta_1, ..., \beta_n)C^{-1}AC\\ &= (\beta_1, ..., \beta_n)B.\end{aligned}$$

□

定义 5.1.8 (线性变换的特征多项式)**.** 假设$\mathscr{A}$是域F上的n维线性空间V上的一个线性变换. 假设V有一组基$\{\alpha_1, ..., \alpha_n\}$, 而且$\mathscr{A}$在这组基下的矩阵表示为$A$. 则我们称$A$的特征多项式为线性变换$\mathscr{A}$的特征多项式.

由于同一个线性变换在不同的基下的矩阵表示是相似的, 而相似的矩阵具有相同的特征多项式. 所以这个定义不依赖于V的基的选取.

§5.2 空间的分解

假设V是数域F上的线性空间, $\mathscr{A}$是V上的一个线性变换.

定义 5.2.1 (不变子空间)**.** 假设W是V的一个线性子空间, 而且$\mathscr{A}(W) \subseteq W$, 我们就称$W$是线性变换$\mathscr{A}$的一个不变子空间. 零空间0和$V$本身叫做$V$的平凡不变子空间.

例 5.2.1. 把$\mathbb{C}$看成数实数域$\mathbb{R}$上的二维的线性空间, 那么复共轭$\mathbb{C}$上的一个线性变换. 它的非平凡的不变子空间只有$\mathbb{R}$.

例 5.2.2. 假设$\mathscr{A}$是二维xy-平面V的旋转变换. 如果$\mathscr{A}$的旋转角度是π的偶数倍, 那么$\mathscr{A}$是恒等变换. 如果$\mathscr{A}$ 的旋转角度是π的奇数倍, 那么$-\mathscr{A}$是恒等变换. 这两种情况下,任何一条过原点的直线都是$\mathscr{A}$ 的非平凡的不变子空间. 如果$\mathscr{A}$ 的旋转角度不是π的整数倍, 那么$\mathscr{A}$没有非平凡的不变子空间.

例 5.2.3. 以过原点的直线ℓ为对称轴的反射有两个非平凡不变子空间, 一个直线ℓ本身, 另一个是过原点与ℓ垂直的直线.

定义 5.2.2 (零化多项式)**.** 假设V是数域F上的线性空间, $\mathscr{A}$是V上的一个线性变换, $f(x)$是一个以F 中的元素为系数的多项式. 如果$f(\mathscr{A})=0$, 我们就称f是$\mathscr{A}$的一个零化多项式. 类似的, 我们可以定义矩阵的零化多项式.

由Hamilton-Cayley定理可知, 线性变换$\mathscr{A}$的特征多项式是它的一个零化多项式.

定理 5.2.1 (Fitting分解)**.** 设$\mathscr{A}$是数域F上n维线性空间V上的线性变换. 则存在k使得

$$\mathrm{Ker}(\mathscr{A}^k)=\mathrm{Ker}(\mathscr{A}^{k+1})=\cdots,\ \mathrm{Im}(\mathscr{A}^k)=\mathrm{Im}(\mathscr{A}^{k+1})=\cdots.$$

记$V_1=\mathrm{Ker}(\mathscr{A}^k)$, $V_2=\mathrm{Im}(\mathscr{A}^k)$, 则

$$V=V_1\bigoplus V_2,$$

而且$\mathscr{A}$限制在V_1上是幂零线性变换, $\mathscr{A}$限制在V_2上是可逆的线性变换.

证明 由于$\mathrm{Ker}(\mathscr{A})\subseteq\mathrm{Ker}(\mathscr{A}^2)\subseteq\cdots$, 而且这个序列有上界$V$是个有限维的线性空间, 所以一定存在$k$使得$\mathrm{Ker}(\mathscr{A}^k)=\mathrm{Ker}(\mathscr{A}^{k+1})$. 此时由维数公式可知$\mathrm{Im}(\mathscr{A}^k)=\mathrm{Im}(\mathscr{A}^{k+1})$. 因此$\mathscr{A}$限制在$V_2$上是满射, 从而是可逆的线性变换. 由于$V_1=\mathrm{Ker}(\mathscr{A}^k)$, 所以$\mathscr{A}$限制在$V_1$定义出来的线性变换$\mathscr{A}_{|V_1}$做$k$次幂之后是零映射. 所以说$\mathscr{A}$限制在$V_1$上是幂零线性变换.

我们下面来说明$V=V_1\bigoplus V_2$. 任取$\alpha\in V$, $\mathscr{A}^k(\alpha)\in\mathrm{Im}(\mathscr{A}^k)=\mathrm{Im}(\mathscr{A}^{2k})$. 所以存在$\beta\in V$使得$\mathscr{A}^k(\alpha)=\mathscr{A}^{2k}(\beta)=\mathscr{A}^k(\mathscr{A}^k(\beta))$. 因此

$$\mathscr{A}^k(\alpha-\mathscr{A}^k(\beta)).$$

所以$\alpha-\mathscr{A}^k(\beta)\in V_1$. 所以

$$\alpha=\alpha-\mathscr{A}^k(\beta)+\mathscr{A}^k(\beta),$$

其中$\alpha-\mathscr{A}^k(\beta)\ inV_1$, $\mathscr{A}^k(\beta)\in V_2$. 因此$V=V_1+V_2$. 我们还需要说明$V_1\cap V_2=0$. 假设存在非零向量$\alpha\in V_1\cap V_2$, 则由于$\mathscr{A}^k$限制在$V_2$上是可逆的线性变换, 所以$\mathscr{A}^k(\alpha)\neq 0$. 同时由于$\alpha\in V_1$, 所以$\mathscr{A}^k(\alpha)=0$引起矛盾. 这就说明了$V_1\cap V_2=0$. 从而$V=V_1\bigoplus V_2$. □

定理 5.2.2. 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换, 假设f_1, ..., f_k是k个两两互素的多项式. 则

$$\mathrm{Ker}f_1\cdots f_k(\mathscr{A})=\mathrm{Ker}f_1(\mathscr{A})\oplus\cdots\oplus\mathrm{Ker}f_k(\mathscr{A}).$$

证明 显然我们只需要对$k=2$的情形证明即可.

任给$\alpha\in \mathrm{Ker} f_1f_2(\mathscr{A})$. 假设$1=f_1g_1+f_2g_2$, 那么

$$\alpha=(f_1g_1+f_2g_2)(\mathscr{A})(\alpha)=g_1(\mathscr{A})(f_1(\mathscr{A})(\alpha))+g_2(\mathscr{A})(f_2(\mathscr{A})(\alpha)).$$

记$\alpha_1=g_1(\mathscr{A})(f_1(\mathscr{A})(\alpha))$, $\alpha_2=g_2(\mathscr{A})(f_2(\mathscr{A})(\alpha))$. 则

$$\alpha_1\in \mathrm{Ker} f_2(\mathscr{A}),\ \alpha_2\in \mathrm{Ker} f_1(\mathscr{A}).$$

所以$\mathrm{Ker} f_1f_2(\mathscr{A})\subseteq \mathrm{Ker} f_1(\mathscr{A})+\mathrm{Ker} f_2(\mathscr{A})$. 假设$\beta\in \mathrm{Ker} f_1(\mathscr{A})\bigcap \mathrm{Ker} f_2(\mathscr{A})$. 那么

$$\beta=g_1(\mathscr{A})(f_1(\mathscr{A})(\beta))+g_2(\mathscr{A})(f_2(\mathscr{A})(\beta))=0.$$

所以它们的和是直和.

左边包含右边是显然的.

□

推论 5.2.3. 假设V是F上的一个线性空间, $\mathscr{A}$ 是V上的一个线性变换, 假设f_1, ..., f_k 是k个两两互素的多项式. 则

$$\mathrm{Ker} f_1(\mathscr{A})+\cdots+\mathrm{Ker} f_k(\mathscr{A})=\mathrm{Ker} f_1(\mathscr{A})\oplus\cdots\oplus \mathrm{Ker} f_k(\mathscr{A}).$$

定义 5.2.3 (特征值和特征向量)**.** 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. 如果存在非零向量$\alpha\in V$和F中的元素$\lambda\in F$, 使得

$$\mathscr{A}(\alpha)=\lambda\alpha.$$

则称α为$\mathscr{A}$的一个特征向量, λ为α对应的特征值; 或者称λ为$\mathscr{A}$的一个特征值, α 为λ对应的特征向量. 注意特征向量一定是非零向量, 而特征值可以等于0.

类似的, 我们可以定义矩阵的特征值和特征向量. 假设A 是一个$n\times n$矩阵, 如果存在非零向量$\alpha\in V$ 和F 中的元素$\lambda\in F$, 使得

$$A\alpha=\lambda\alpha.$$

则称α为A的一个特征向量, λ为α 对应的特征值; 或者称λ为A的一个特征值, α 为λ对应的特征向量.

定理 5.2.4. 假设A是一个n-级矩阵. 则λ为A的一个特征值, 当且仅当即λ是A的特征多项式的根.

证明 假设λ为A的一个特征值. 那么存在一个非零向量$\alpha\in F^n$, 使得

$$A\alpha=\lambda\alpha.$$

所以α是线性齐次方程组

$$(\lambda I - A)X = 0$$

的一个非零解. 所以$|\lambda I - A| = 0$, 即λ是A的特征多项式的根.

反之如果λ是A的特征多项式的根, 那么$|\lambda I - A| = 0$. 所以线性齐次方程组

$$(\lambda I - A)X = 0$$

一定有非零解$X = \alpha$. 所以$A\alpha = \lambda\alpha$.

□

类似的我们可以证明

定理 5.2.5. 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. 则λ为$\mathscr{A}$的一个特征值, 当且仅当即λ是$\mathscr{A}$的特征多项式的根.

定义 5.2.4 (特征子空间)**.** 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换, λ为$\mathscr{A}$ 的一个特征值. $\mathscr{A}$的特征值λ的特征子空间定义为

$$E_\lambda = \{\alpha \in V | \mathscr{A}(\alpha) = \lambda\alpha.\}$$

定义 5.2.5 (根子空间)**.** 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换, λ为$\mathscr{A}$ 的一个特征值. 则$\mathscr{A}$的特征值λ的根子空间定义为

$$R_\lambda = \{\alpha \in V | (\mathscr{A} - \lambda\mathscr{I})^m(\alpha) = 0 \text{ 对某个正整数 } m \text{ 成立}\}.$$

特征子空间E_λ中所有的元素α都满足$(\mathscr{A} - \lambda\mathscr{I})(\alpha) = 0$. 所以特征子空间一定包含在根子空间中. 但是反之未必成立.

注意到根子空间是个有限维的子空间, 所以存在一组只有有限个向量的基α_1, ..., α_t. 根据根子空间的定义, 存在正整数m_1, ..., m_t使得

$$(\mathscr{A} - \lambda\mathscr{I})^{m_1}(\alpha_1) = \cdots = (\mathscr{A} - \lambda\mathscr{I})^{m_t}(\alpha_t) = 0.$$

选一个比m_1, ..., m_t都大的整数m, 就有

$$(\mathscr{A} - \lambda\mathscr{I})^{m}(\alpha_1) = \cdots = (\mathscr{A} - \lambda\mathscr{I})^{m}(\alpha_t) = 0.$$

所以在根子空间定义中, 虽然指数m依赖于α, 但是可以选一个大一点的m, 让它对所有的向量都有效, 即存在一个常数m, 使得根子空间

$$R_\lambda = \mathrm{Ker}(\mathscr{A} - \lambda\mathscr{I})^m.$$

根据定理5.2.2, 这也说明了不同的特征值的根子空间的和为直和.

我们下面说明, 这个m可以取的不超过特征多项式中根λ 的重数.

命题 5.2.6. 假设V是$\mathbb{C}$上的一个线性空间, $\mathscr{A}$是V上的一个线性变换, $\mathscr{A}$的特征多项式是

$$f_{\mathscr{A}}(\lambda)=(\lambda-\lambda_1)^{a_1}\cdots(\lambda-\lambda_k)^{a_k}.$$

则

$$V=\mathrm{Ker}(\mathscr{A}-\lambda_1\mathscr{I})^{a_1}\oplus\cdots\oplus\mathrm{Ker}(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}$$

证明 由定理5.2.2 和哈密顿-凯莱定理可知.

□

定理 5.2.7. 假设V是$\mathbb{C}$上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. 则V是根子空间的直和. 如果$\mathscr{A}$ 的特征多项式是

$$f_{\mathscr{A}}(\lambda)=(\lambda-\lambda_1)^{a_1}\cdots(\lambda-\lambda_k)^{a_k}.$$

则根子空间

$$R_{\lambda_i}=\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i},\ 1\le i\le k.$$

证明 假设$\mathscr{A}$的特征多项式是

$$f_{\mathscr{A}}(\lambda)=(\lambda-\lambda_1)^{a_1}\cdots(\lambda-\lambda_k)^{a_k}.$$

其中的λ_i两两不同. 则由上面的命题可知

$$V=\mathrm{Ker}f_{\mathscr{A}}(\lambda)=\mathrm{Ker}(\mathscr{A}-\lambda_1\mathscr{I})^{a_1}\oplus\cdots\oplus\mathrm{Ker}(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}.$$

由于$\mathrm{Ker}(\mathscr{A}-\lambda_1\mathscr{I})^{a_1}\subseteq R_{\lambda_1}$, ..., $\mathrm{Ker}(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}\subseteq R_{\lambda_k}$, 而且不同的根子空间之和也是直和, 所以$\mathrm{Ker}(\mathscr{A}-\lambda_1\mathscr{I})^{a_1}=R_{\lambda_1}$, ..., $\mathrm{Ker}(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}=R_{\lambda_k}$. 所以则$V$是根子空间的直和. 所以根子空间

$$R_{\lambda_i}=\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i},\ 1\le i\le k.$$

□

这这里要注意, 虽然根子空间

$$R_{\lambda_i}=\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i},\ 1\le i\le k,$$

但是这里的a_i却不一定是使得这个等式成立的最小的正整数. 例如如果$\mathscr{A}$是恒等变换$\mathscr{I}$, 那么特征多项式为$(\lambda-1)^n$, 所以1的根子空间

$$R_1=\mathrm{Ker}(\mathscr{A}-\mathscr{I})^n.$$

但实际上$R_1=\mathrm{Ker}(\mathscr{A}-\mathscr{I})$.

定义 5.2.6 (最小多项式)**.** 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. $\mathscr{A}$的首项系数为1 的次数最低的零化多项式叫做$\mathscr{A}$的最小多项式. 类似的可以定义矩阵的最小多项式.

定理 5.2.8. 最小多项式是唯一的, 而且最小多项式整除零化多项式.

证明

利用辗转相除法可知. □

推论 5.2.9. 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. 如果$\mathscr{A}$的特征多项式为

$$f_{\mathscr{A}}(\lambda) = (\lambda - \lambda_1)^{a_1} \cdots (\lambda - \lambda_k)^{a_k}, a_i \geq 1,\ 1 \leq i \leq k$$

$\mathscr{A}$的最小多项式为

$$f_{\mathscr{A}}(\lambda) = (\lambda - \lambda_1)^{b_1} \cdots (\lambda - \lambda_k)^{b_k},$$

则$1 \leq b_i \leq a_i,\ 1 \leq i \leq k$. 即不仅最小多项式的根是特征多项式的根, 特征多项式的根也都是最小多项式的根.

定义 5.2.7 (可对角化)**.** 假设A是数域F上的一个n阶矩阵, 如果A相似于一个对角矩阵, 我们就称A 是可对角化的矩阵.

定理 5.2.10. A可对角化当且仅当A有n个线性无关的特征向量, 也当且仅当最小多项式没有重根.

证明 假设A可对角化, 那么一定存在可逆矩阵P使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

所以

$$AP = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

把P按照列向量写出来, 即假设

$$P = (\alpha_1,\ ...,\ \alpha_n).$$

则

$$A(\alpha_1,\ ...,\ \alpha_n) = (\alpha_1,\ ...,\ \alpha_n) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = (\lambda_1\alpha_1,\ ...,\ \lambda_n\alpha_n).$$

所以α_1, ..., $\lambda_n\alpha_n$都是特征向量, 又因为它们是可逆矩阵P的列向量, 所以它们是线性无关的.

把以上的步骤逆回去, 即可知如果A有n个线性无关的特征向量, 那么A可对角化.

如果A可以对角化, 我们知道相似的矩阵具有相同的最小多项式, 对角的矩阵的最小多项式容易知道是没有重根的.

反之, 如果最小多项式没有重根, 那么由定理5.2.2 可知, 空间可以分解成特征子空间的直和, 因此A有n个线性无关的特征向量. □

推论 5.2.11. 如果A有n个互不相同的特征值, 则A 是可以对角化的.

定义 5.2.8 (循环子空间). 假设W是V的一个线性子空间, $\mathscr{A}$是V上的一个线性变换. 如果存在$\alpha \in W$使得W可以由α, $\mathscr{A}(\alpha)$, $\mathscr{A}^2(\alpha)$, ...生成, 我们就称W是由α生成的$\mathscr{A}$ 的循环子空间, 记为$W =< \alpha >$. 注意不同的向量可以生成相同的循环子空间.

定义 5.2.9 (幂零变换). 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个线性变换. 如果存在自然数N使得$\mathscr{A}$是V上的零变换, 那么我们就称$\mathscr{A}$是V上的一个幂零变换.

引理 5.2.12. 假设V是F上的一个线性空间, $\mathscr{A}$是V上的一个幂零变换. α, ..., $\mathscr{A}^k(\alpha)$都不是0. 但是$\mathscr{A}^{k+1}(\alpha)=0$, 则$\alpha$, ..., $\mathscr{A}^k(\alpha)$线性无关. 记W为V的由α生成的循环子空间. 则W是$\mathscr{A}$的不变子空间, $\mathscr{A}^k(\alpha)$, $\mathscr{A}^{k-1}(\alpha)$..., α为W的一组基, $\mathscr{A}$在这组基下的矩阵为

$$\mathscr{A}(\mathscr{A}^k(\alpha),\ \mathscr{A}^{k-1}(\alpha),\ ...,\ \alpha) = (\mathscr{A}^k(\alpha),\ \mathscr{A}^{k-1}(\alpha)...,\ \alpha)\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$$

定理 5.2.13. 假设V是F上的一个n维的线性空间, $\mathscr{A}$是V 上的一个幂零变换. 则V可以分解为循环子空间的直和.

证明 我们假设命题对低于n维的线性空间是成立的. 那么我们令$W = \mathscr{A}(V)$, 则W的维数小于n. 而且W也是$\mathscr{A}$ 的不变子空间. 所以W可以写成循环子空间的直和. 我们假设

$$W =< \mathscr{A}(\alpha_1) > \oplus \cdots \oplus < \mathscr{A}(\alpha_k) > .$$

那么容易证明$< \alpha_1 >$, ..., $< \alpha_k >$的和仍然是直和. 否则一定存在不全为零的$\beta_1 \in< \alpha_1 >$, ..., $\beta_k \in< \alpha_k >$ 使得

$$\beta_1 + \cdots + \beta_k = 0.$$

这导致

$$\mathscr{A}(\beta_1)+\cdots+\mathscr{A}(\beta_k)=0.$$

但是由于$<\mathscr{A}(\alpha_1)>, ..., <\mathscr{A}(\alpha_k)>$的和是直和. 所以$\mathscr{A}(\beta_1)=0, ..., \mathscr{A}(\beta_k)=0$. 我们假设$\mathscr{A}^t(\alpha_1)=0$, 但是$\mathscr{A}^{t-1}(\alpha_1)\neq 0$. 由上面的引理可知$\alpha_1$, $\mathscr{A}(\alpha_1)$, $\mathscr{A}^{t-1}(\alpha_1)$线性无关. 再假设$\beta_1=a_1\alpha_1+a_2\mathscr{A}(\alpha_1)+\cdots+a_t\mathscr{A}^{t-1}(\alpha_1)$. 那么由$\mathscr{A}(\beta_1)=0$可知$a_1\mathscr{A}(\alpha_1)+\cdots+a_{t-1}\mathscr{A}^{t-1}(\alpha_1)=0$. 因为$\mathscr{A}(\alpha_1)$, ..., $\mathscr{A}^{t-1}(\alpha_1)$ 线性无关, 所以$a_1=\cdots=a_{t-1}=0$. 所以$\beta_1\in<\mathscr{A}(\alpha_1)>$. 同理可证$\beta_k\in<\mathscr{A}(\alpha_k)>$. 但是由于$<\mathscr{A}(\alpha_1)>, ..., <\mathscr{A}(\alpha_k)>$的和是直和, 所以必须有$\beta_1=\cdots=\beta_k=0$. 矛盾.

记$U=<\alpha_1>\oplus\cdots\oplus<\alpha_k>$. 则任给$\alpha\in V$, 则$\mathscr{A}(\alpha)\in W=\mathscr{A}(V)$, 所以

$$\mathscr{A}(\alpha)\in<\mathscr{A}(\alpha_1)>\oplus\cdots\oplus<\mathscr{A}(\alpha_k)>.$$

所以一定存在元素$\beta\in U=<\alpha_1>\oplus\cdots\oplus<\alpha_k>$使得$\mathscr{A}(\alpha)=\mathscr{A}(\beta)$. 令$\gamma=\alpha-\beta$. 则$\gamma\in\mathrm{Ker}\mathscr{A}$. 所以$V$中任何一个元素都可以写成$\alpha=\beta+(\alpha-\beta)$的形式. 所以

$$U+\mathrm{Ker}\mathscr{A}=V.$$

所以可以选一个$\mathrm{Ker}\mathscr{A}$的子空间V_1使得

$$V=U\oplus V_1.$$

V_1可以这样选. 先取$\mathrm{Ker}\mathscr{A}\bigcap U$的一组基$\gamma_1$, ..., γ_s, 然后将其扩充为$\mathrm{Ker}\mathscr{A}$的一组基γ_1, ..., γ_s, γ_{s+1}, ..., $\gamma_{s+\widetilde{s}}$, 则令V_1为γ_{s+1}, ..., $\gamma_{s+\widetilde{s}}$生成的空间即可.

定理得证.

□

定义 5.2.10 (Jordan块)**.** *形如*

$$\begin{pmatrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1\\ & & & \lambda\end{pmatrix},\ \lambda\in\mathbb{C}$$

形状的矩阵叫做一个Jordan块.

定理 5.2.14 (Jordan标准形)**.** *假设A是复数域上的一个n 阶矩阵. 则A相似于如下形状的矩阵*

$$\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k\end{pmatrix}$$

其中每一个J_i都是一个Jordan块.

证明 记$V = F^n$, e_1, ..., e_n为V的标准基. 则A决定了V上的一个线性变换$\mathscr{A}$. 我们知道V可以分解成根子空间的直和

$$V = R_1 \oplus \cdots \oplus R_s.$$

而限制在每个根子空间上R_i上时, $\mathscr{A}_i = \mathscr{A}_{|R_i}$ 只有一个特征值λ_i. 所以如果记$\mathscr{I}_i$为根子空间上R_i上的恒等变换, 则

$$\mathscr{A}_i - \lambda_i \mathscr{I}_i$$

就是R_i 上的一个幂零变换. 所以分解为循环子空间的直和. 所以R_i 有一组基, 使得$\mathscr{A}_i - \lambda_i \mathscr{I}_i$在这组基下的矩阵是

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix}$$

每一个J_i都形如

$$J_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$$

因此$\mathscr{A}_i$在这组基下的矩阵是

$$\begin{pmatrix} \widetilde{J_1} & & & \\ & \widetilde{J_2} & & \\ & & \ddots & \\ & & & \widetilde{J_t} \end{pmatrix}$$

每一个$\widetilde{J_i}$都形如

$$\widetilde{J_i} = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}$$

把每个根子空间的基拼起来, 便成为V的基, 由此可知定理成立.

□

§5.3 矩阵指数

我们知道$\mathbb{C}^n$中的向量$\alpha = (x_1,\ ...,\ x_n)$的范数定义为

$$\|\alpha\| := \sqrt{|x_1|^2 + \cdots + |x_n|^2}$$

假设A是一个n阶的实矩阵或者复矩阵,我们定义$A = (a_{ij})$的范数为

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

容易验证如此定义的范数满足

$$\|A + B\| \le \|A\| + \|B\|,$$
$$\|AB\| \le \|A\|\|B\|.$$

定义 5.3.1 (柯西序列)**.** *一个矩阵序列A_m称为是柯西序列, 如果任给$\varepsilon > 0$, 都存在N, 使得当$k,\ l > N$时,*

$$\|A_k - A_l\| < \varepsilon.$$

不难看出下面的命题成立.

命题 5.3.1. *如果一个矩阵序列A_m是柯西序列, 则存在唯一的矩阵A, 使得A_m收敛到A.*

任给矩阵A, 令

$$A_m = \sum_{n=0}^{m} \frac{A^n}{n!}.$$

则A_m形成一个柯西序列, 这个序列的极限我们记为

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

例 5.3.1. *如果*

$$A = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix},$$

则

$$e^A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

例 5.3.2. *如果*

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix},$$

则

$$e^A = \begin{pmatrix} 1 & a & b+\frac{1}{2}ac \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

命题 5.3.2. *1. 如果$AB = BA$, 那么$e^A e^B = e^{A+B} = e^B e^A$.*

2. e^A是可逆矩阵, 其逆矩阵是e^{-A}.

3. $|e^A| = e^{Tr(A)}$.

证明 我们只证明第一条.

$$\begin{aligned} e^A e^B &= (\sum_{k=0}^{\infty} \frac{A^k}{k!})(\sum_{l=0}^{\infty} \frac{B^l}{l!}) \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{A^n}{n!} \frac{B^{m-n}}{(m-n)!} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} A^n B^{m-n} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} (A+B)^m \\ &= e^{A+B} = e^B e^A. \end{aligned}$$

□

我们知道$n \times n$可逆若尔当块是形如

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & \frac{1}{\lambda} & & & \\ & 1 & \frac{1}{\lambda} & & \\ & & \ddots & \ddots & \\ & & & 1 & \frac{1}{\lambda} \\ & & & & 1 \end{pmatrix}$$

的矩阵。设

$$A = \begin{pmatrix} 1 & \frac{1}{\lambda} & & & \\ & 1 & \frac{1}{\lambda} & & \\ & & \ddots & \ddots & \\ & & & 1 & \frac{1}{\lambda} \\ & & & & 1 \end{pmatrix}, \qquad B = I_n - A.$$

则我们定义$\log(A) = \log(I_n - B)$为下面的矩阵,

$$\log(A) = \log(I_n - B) := -\sum_{n=1}^{+\infty} \frac{B^n}{n}.$$

注意着虽然看起来是个无穷级数, 但是实际上由于B是个幂零矩阵, $B^n = 0$, 所以这个级数其实是个有限级数，自然是收敛的。而对任意的方阵M, 易知

$$e^M = \sum_{n=0}^{+\infty} \frac{M^n}{n!}$$

收敛。所以

$$A = e^{\log A}.$$

所以

$$\lambda A = e^{xI_n + \log A}.$$

注意对任意的$\lambda \in \mathbb{C}$, 总存在$x \in \mathbb{C}$使得这里的$e^x = \lambda$.

利用展开式，这个公式成立是显然的。这样我们就有了下面的引理

引理 5.3.3. *对任何可逆阵A, 都存在可逆阵B, 使得$A = e^B$.*

推论 5.3.4. *对任何可逆阵A和正整数k, 都存在可逆阵B, 使得$A = B^k$.*

证明 根据上面的引理, 存在可逆阵C, 使得$A = e^C$. 令

$$B = \frac{1}{k}C,$$

即可得$A = B^k$. □

注意上面的这个推论对不可逆的矩阵, 不一定成立. 例如

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

$k = 2$. 这时候不存在矩阵B, 使得$A = B^2$. 因为如果这样的B存在的话, B一定是秩为1 的幂零矩阵. 但是我们知道秩为1 的幂零矩阵的平方一定是零矩阵, 不可能等于A. 假设A是一个n-级矩阵, 对任意的正整数k都存在可逆阵B使得$A = B^k$当且仅当特征值0的代数重数等于几何重数.

习题 5.3

1. 复数域上任何一个方阵都相似于一个上三角矩阵.

   证明 我们用数学归纳法. 假设命题对$n-1$级矩阵成立, 即任何一个$n-1$级矩阵都相似于上三角矩阵. 下面我们来证明任何一个$n \times n$方阵A都相似于一个上三角矩阵.

假设$f_A(\lambda)$是A的特征多项式. 由代数基本定理, 我们知道在复数域上特征多项式一定有根λ_0. 所以

$$|\lambda_0 I - A| = 0,$$

这里的I是$n \times n$单位矩阵. 所以线性方程组

$$(A - \lambda_0 I)X = 0$$

有非零解α. 这个非零向量就是A的特征值λ的特征向量.

我们知道任何一个非零的n维列向量都可以作为一个可逆矩阵的第一列. 所以存在可逆矩阵P使得P的第一列是α, 假设

$$P = (\alpha,\ \alpha_2,\ ...,\ \alpha_n).$$

那么

$$A\alpha = \lambda_0 \alpha,$$

因此

$$A(\alpha,\ \alpha_2,\ ...,\ \alpha_n) = (A\alpha,\ A\alpha_2,\ ...,\ A\alpha_n) = (\alpha,\ \alpha_2,\ ...,\ \alpha_n)\begin{pmatrix} \lambda_0 & \beta \\ 0 & B \end{pmatrix}$$

其中β是个$n-1$维列向量, B是个$n-1$级矩阵. 所以

$$AP = P\begin{pmatrix} \lambda_0 & \beta \\ 0 & B \end{pmatrix}$$

即

$$P^{-1}AP = \begin{pmatrix} \lambda_0 & \beta \\ 0 & B \end{pmatrix}.$$

由归纳假设存在$n-1$级矩阵Q使得

$$Q^{-1}BQ$$

是上三角矩阵. 所以

$$\begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix} P^{-1}AP \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_0 & \beta Q \\ 0 & Q^{-1}BQ \end{pmatrix}$$

是上三角矩阵. □

2. 设A是一个n级非零矩阵,

$$B=\begin{pmatrix} A & -A \\ A & -A \end{pmatrix}$$

求B的最小多项式.

证明

对$|\lambda I_{2n}-B|$进行行列式的初等变换可知,

$$|\lambda I_{2n}-B|=\begin{vmatrix} \lambda I_n-A & A \\ -A & \lambda I_n+A \end{vmatrix}=\begin{vmatrix} \lambda I_n & -\lambda I_n \\ -A & \lambda I_n+A \end{vmatrix}=\begin{vmatrix} \lambda I_n & 0 \\ -A & \lambda I_n \end{vmatrix}=\lambda^{2n}.$$

由于$B\neq 0$,

$$B^2=\begin{pmatrix} A & -A \\ A & -A \end{pmatrix}^2=0,$$

所以B的最小多项式为λ^2.

或者直接验证$B^2=0$也可以知道B的最小多项式为λ^2. 但是前面的过程告诉我们B的最小多项式一定是λ的某个幂次, 提示我们去计算B的幂次组成的序列.

□

3. 复数域上任何一个方阵是否都相似于一个下三角矩阵? (留做作业).

证明 是的. 只要在上一题的证明中把α作为基的最后一个向量即可.

□

4. 假设数域F的n-级矩阵A的迹$a_{11}+\cdots+a_{nn}=0$. 求证A相似于一个对角线上元素都是0的矩阵.

证明 我们用数学归纳法, 假设命题对$n-1$级的矩阵成立.

首先我们假设A不是零矩阵, 否则是平凡的. 因此A不可能是纯量矩阵$kI=diag(k,\ k,\ ...,\ k)$的形式.

由于A不是纯量矩阵, 因此存在一个非零向量α 使得$\alpha,\ A\alpha$是线性无关的向量. 这个结论可以简单的证明如下: 如果对任何非零向量α来说, $\alpha,\ A\alpha$都是线性相关的. 那么设$Ae_i=a_ie_i$, e_i是第i个位置是1, 其余位置是0的列向量. 所以$A=diag(a_1,\ ..,\ a_n)$. 然后如果$a_i\neq a_j$, 考虑$A(e_i+e_j)=a_ie_i+a_je_j$可知和e_i+e_j线性无关, 矛盾.

然后把$\alpha,\ A\alpha$扩充为F^n的一组基$\alpha,\ A\alpha,\ \alpha_3...,\ \alpha_n$. 则

$$P=(\alpha,\ A\alpha,\ \alpha_3...,\ \alpha_n)$$

是一个可逆矩阵, 而且

$$AP = A(\alpha,\ A\alpha,\ \alpha_3...,\ \alpha_n) = (\alpha,\ A\alpha,\ \alpha_3...,\ \alpha_n)\begin{pmatrix}0 & \beta\\ \gamma & B\end{pmatrix} = P\begin{pmatrix}0 & \beta\\ \gamma & B\end{pmatrix}$$

形状的矩阵. 其中B是$n-1$级的方阵, β是$n-1$ 维行向量, $\gamma = (1,\ 0,\ ...,\ 0)'$是个$n-1$ 维的列向量. 所以A相似于

$$\begin{pmatrix}0 & \beta\\ \gamma & B\end{pmatrix}.$$

再对B用归纳假设即可.

□

5. 假设A是数域F上的一个$n(>1)$级矩阵, 空间V 为由数域F上的所有$n\times n$矩阵组成的n^2维线性空间. 定义

$$\varphi_A:\ V \longrightarrow V,\ \varphi_A(X) = AX.$$

假设A的秩等于r, 求证φ_A的秩等于nr.

证明 我们假设V中由除了第i列之外都是0的矩阵组成的子空间为V_i, $1\le i\le n$. 则每个V_i都是φ_A 的不变子空间, 维数为n. 而且

$$V = V_1\bigoplus\cdots\bigoplus V_n.$$

我们取V_1的一组基$e_{11},\ ..,\ e_{n1}$, 其中e_{i1}是除了第i行, 第1列之外都为0的矩阵. 则

$$\varphi_A(e_{11},\ ..,\ e_{n1}) = (e_{11},\ ..,\ e_{n1})A.$$

同理可证可选取V_i的类似的基使得φ_A在这组基下的矩阵也是A. 把这些基合起来以后就是V的基, φ_A在这组基下的矩阵是

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

即由n个A组成的准对角矩阵. 所以φ_A的秩等于nr.

□

6. 假设A是数域F上的一个$n(>1)$级矩阵, 空间V为由数域F上的所有$n\times n$矩阵组成的n^2 维线性空间. 定义

$$\varphi_A:\ V \longrightarrow V,\ \varphi_A(X) = AX - XA.$$

这是V上的一个线性变换.

我们把V上的所有的线性变换全体形成的空间记为W.

我们定义一个从V到W的线性映射

$$\psi:\ V \longrightarrow W,\ \psi(A)=\varphi_A,\ A\in V.$$

(1) 求W的维数.

(2) 说明ψ既不是单射, 也不是满射.

(3) 求ψ的像空间的维数.

证明

(1) W的维数是n^4.

(2) 因为$\psi(I)=0$, 所以不是单射. 因为V的维数的维数小于W的维数, 所以也不是满射.

(3) 只需要求ψ的核空间的维数. $\psi(A)=0$当且仅当$\varphi_A=0$当且仅当对所有的X, 都有$AX=XA$. 这样的矩阵只有纯量阵, 维数是1. 所以ψ的像空间的维数是n^2-1. □

7. 假设A和B是复数域上的两个$n\times n$矩阵, 而且$AB=BA$, 求证存在可逆矩阵P, 使得

$$P^{-1}AP,\ P^{-1}BP$$

都是上三角矩阵.

证明 我们用数学归纳法, 假设命题对级数小于n的矩阵是成立的.

记$V=F^n$为n维线性空间. 那么A和B作为两个$n\times n$矩阵, 定义了两个线性变换$\mathscr{A}$, $\mathscr{B}$使得

$$\mathscr{A}(\alpha)=A\alpha,\ \mathscr{B}(\beta)=B\beta,$$

对任意的向量α, $\beta\in V$. 我们只需要证明存在V的一组基α_1, ..., α_n 使得两个线性变换$\mathscr{A}$, $\mathscr{B}$在这同一组基下的矩阵同时都是上三角即可.

我们设E_λ是线性变换$\mathscr{A}$的特征值λ的特征子空间, 即

$$E_\lambda=\{\alpha\in V|\mathscr{A}(\alpha)=\lambda\alpha\}.$$

那么由于$AB=BA$, 所以

$$\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$$

所以对任何的$\alpha \in E_\lambda$, 我们都有

$$\mathscr{A}(\mathscr{B}(\alpha)) = \mathscr{B}(\mathscr{A}(\alpha)) = \mathscr{B}(\lambda(\alpha)) = \lambda\mathscr{B}(\alpha).$$

所以$\mathscr{B}(\alpha) \in E_\lambda$. 所以$E_\lambda$ 既是$\mathscr{A}$的不变子空间, 也是$\mathscr{B}$ 的不变子空间. 设这个子空间的一组基是α_1, ..., α_k, 然后扩充成V的一组基α_1, ..., α_n. 则两个线性变换$\mathscr{A}$, $\mathscr{B}$ 在这同一组基下的矩阵都是准上三角形状的. 我们假设

$$\mathscr{A}(\alpha_1, \ ..., \ \alpha_n) = (\alpha_1, \ ..., \ \alpha_n)\begin{pmatrix} \lambda I_k & * \\ 0 & A_1 \end{pmatrix},$$

$$\mathscr{B}(\alpha_1, \ ..., \ \alpha_n) = (\alpha_1, \ ..., \ \alpha_n)\begin{pmatrix} B_2 & * \\ 0 & B_1 \end{pmatrix}.$$

注意此时有$A_1B_1 = B_1A_1$. 由归纳假设可以知道命题成立. □

8. 假设$\mathscr{A}$是有限维线性空间V上的可对角化线性变换, W是$\mathscr{A}$的一个不变子空间, 则$\mathscr{A}$在W上也刻意对角化. **证明** $\mathscr{A}$是V上的可对角化线性变换意味着$\mathscr{A}$的绩效多项式没有重根, 而$\mathscr{A}$在W上的限制变换的极小多项式整除$\mathscr{A}$的极小多项式(看作V 上的线性变换). 所以$\mathscr{A}$在W上的限制变换也是可以对角化的.

□

9. 假设A和B是复数域上的两个可对角化的$n \times n$矩阵, 而且$AB = BA$, 求证存在可逆矩阵P, 使得
$$P^{-1}AP, \ P^{-1}BP$$
都是对角矩阵.

证明 我们用数学归纳法, 假设命题对级数小于n的矩阵是成立的.

记$V = F^n$为n维线性空间. 那么A和B作为两个$n \times n$矩阵, 定义了两个线性变换$\mathscr{A}$, $\mathscr{B}$使得
$$\mathscr{A}(\alpha) = A\alpha, \ \mathscr{B}(\beta) = B\beta,$$
对任意的向量α, $\beta \in V$. 我们只需要证明存在V的一组基α_1, ..., α_n 使得两个线性变换$\mathscr{A}$, $\mathscr{B}$在这同一组基下的矩阵同时对角矩阵即可.

我们设E_λ是线性变换$\mathscr{A}$的特征值λ的特征子空间, 即

$$E_\lambda = \{\alpha \in V | \mathscr{A}(\alpha) = \lambda\alpha\}.$$

那么由于$AB = BA$, 所以
$$\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$$

所以对任何的$\alpha \in E_\lambda$, 我们都有

$$\mathscr{A}(\mathscr{B}(\alpha)) = \mathscr{B}(\mathscr{A}(\alpha)) = \mathscr{B}(\lambda(\alpha)) = \lambda\mathscr{B}(\alpha).$$

所以$\mathscr{B}(\alpha) \in E_\lambda$. 所以$E_\lambda$ 既是$\mathscr{A}$的不变子空间, 也是$\mathscr{B}$ 的不变子空间.

由于A可以对角化, 所以V可以分解为$\mathscr{A}$的特征子空间的直和. 我们假设$\lambda_1 = \lambda$

$$V = E_{\lambda_1} \bigoplus \cdots \bigoplus E_{\lambda_k}.$$

则每一个E_{λ_i}都是$\mathscr{B}$的不变子空间, 而且$\mathscr{A}$, $\mathscr{B}$在每一个E_{λ_i}上都是可交换的可对角化的线性变换.

由归纳假设可以知道命题成立. **证明**

10. 设V 为数域P上一个有限维线性空间，$S = \{\alpha_1, ..., \alpha_n\}$为$V$的一组基，$\mathscr{A}: V \to V$ 为V上一个可逆线性变换。如果$\mathscr{A}$的特征值全为1且$\mathscr{A}(S) = S$, 证明$\mathscr{A}$ 为恒等变换。

    证明 $\mathscr{A}(S) = S$ 说明线性变换把这组基还变为这组基, 只是这组基中的向量次序可能有变化. 由于基里面只有有限多个向量, 最多只有有限种排序, 所以一定存在两个正整数t, s 使得$\mathscr{A}^t(S)$和$\mathscr{A}^s(S)$中向量的排列次序是一模一样的. 因此$\mathscr{A}^t = \mathscr{A}^s$. 由于$\mathscr{A}$ 是可逆的.

    所以$\mathscr{A}^{t-s} = I$, 所以A是可以对角化的, 它的特征值又都是1, 所以它的Jordan标准型时单位矩阵, 所以$\mathscr{A}$是恒等变换.

    □

11. 证明：方阵A是幂零的当且仅当A的所有阶主子式之和均为零。

    证明 我们假设$f(\lambda) = |\lambda I - A|$是$A$的特征多项式. 则$A$的特征多项式为

$$|\lambda I - A| = \lambda^n - s_1\lambda^{n-1} + \cdots + (-1)^n s_n,$$

    其中s_k是A的所有的k阶子式的和. 由于A的所有阶主子式之和均为零, 所以A的特征多项式为λ^n。根据哈密顿-凯莱定理, 特征多项式是零化多项式, 所以$A^n = 0$. 所以方阵A是幂零的.

    □

12. 设$\varepsilon_1 = (1, 0)$, $\varepsilon_2 = (0, 1)$ 为平面直角坐标系的标准基，$\alpha = \varepsilon_1 + \varepsilon_2$. 令$\xi_\alpha$为沿$\alpha$做反射所形成的线性变换，求

    (1) ξ_α 在$\varepsilon_1, \varepsilon_2$下的矩阵$A$；

(2) ξ_α的特征值与特征向量;

(3) 判断A是否可对角化.

证明

(1)

$$\xi_\alpha(\varepsilon_1,\varepsilon_2)=(\varepsilon_2,\varepsilon_1)=(\varepsilon_1,\varepsilon_2)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

$\square$

第六章　λ-矩阵

§6.1　整矩阵

如果$n\times n$矩阵A的元素都是整数, 我们称A为整矩阵. 如果A, B是两个整矩阵, 而且

$$AB=BA=I,$$

我们就称A是可逆整矩阵, B为A的逆, 记为$B=A^{-1}$. 和域上的矩阵一样, 只要$AB=I$, 自然就能推出BA也等于I.

命题 6.1.1. *A是可逆整矩阵当且仅当$|A|=\pm1$.*

证明

假设A是可逆整矩阵, 则存在整数矩阵B, 使得$AB=I$. 两边取行列式可得$|A||B|=1$. 由于整矩阵的行列式也是整数, 所以$|A|$, $|B|$都是整数, 乘积又是1 , 因此A的行列式只能是1 或者是-1.

假设$|A|=\pm1$, 则$B=\frac{1}{|A|}A^*$就是一个整矩阵, 而且$AB=BA=I$. 所以A是可逆整矩阵.

□

一般的, 我们有下面的命题.

命题 6.1.2. *假设A是交换环R上的n-级矩阵, 则A可逆当且仅当它的行列式是R中的可逆元.*

定义 6.1.1 (整矩阵的初等变换)**.** *对整矩阵所做的下面三种变换称为是初等变换:*

1. *交换矩阵的两行(列).*

2. *把矩阵的某一行(列)乘上一个±1.*

3. *在矩阵的某一行(列)上, 加上另外一行(列)的t倍, 这里的t是个整数.*

定义 6.1.2 (整矩阵的等价)**.** *. 如果A, B是两个整矩阵, 而且如果A 可以经过有限次初等变换变为B, 我们就称A, B是两个等价的整矩阵.*

定理 6.1.3 (Smith标准形)**.** *假设A是一个整数矩阵, 则它一定等价于下面形式*

的矩阵

$$\begin{pmatrix} d_1 & & & & & & \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

其中d_i都是正整数, 而且

$$d_i | d_{i+1},\ 1 \le i \le r-1.$$

这种形式的矩阵是由A唯一确定的, 与初等变换的选取无关.

证明 利用初等变换, 可以把A的元素的最大公因子d_1 变出来, 并且放在左上角. 然后利用初等变换, 把A的第一行和第一列的其他元素都变成0. A就变成了下面的这个样子

$$\begin{pmatrix} d_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

其中A_1是一个$n-1$级的整矩阵, 而且它的每个元素都是d_1的倍数. 再对$n-1$级的整矩阵A_1做类似的操作即可看出, A可以最终变为定理中的对角矩阵.

我们下面来证明唯一性. 我们知道初等变换的过程不会改变矩阵元素的最大公因子. 所以d_1是唯一确定的. 其次也不会改变所有的二阶子式的最大公因子. 所以d_2也是唯一确定的. 以此类推, 所有的d_i都是唯一确定的.

□

§6.2 λ-矩阵的定义和基本性质

定义 6.2.1 (λ-矩阵, 可逆)**.** *如果n-级矩阵$A(\lambda)$的每个元素都是一个以λ为变量, 以域F中的元素为系数的多项式, 那么就称$A(\lambda)$ 为数域F 上的一个λ- 矩阵. 假设$A(\lambda)$是一个λ-矩阵, 如果存在另一个λ 矩阵$B(\lambda)$使得*

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n.$$

我们称$A(\lambda)$是可逆矩阵, 称$B(\lambda)$是$A(\lambda)$ 的逆矩阵.

命题 6.2.1. *一个n-级矩阵$A(\lambda)$可逆当且仅当$A(\lambda)$的行列式$|A(\lambda)|$是F中的非零元素.*

证明 如果n-级矩阵$A(\lambda)$可逆, 根据定义, 存在另一个λ矩阵$B(\lambda)$使得

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

两边取行列式可得

$$|A(\lambda)||B(\lambda)| = 1.$$

由于$|A(\lambda)|$, $|B(\lambda)|$都是以λ为变量的多项式, 而且他们的乘积等于1 . 所以他们只能是F 中互逆的数.

反之, 假设行列式$|A(\lambda)|$是F中的非零元素. 记$A(\lambda)^*$为$A(\lambda)$的伴随矩阵, 则根据定义$A(\lambda)^*$仍是一个λ-矩阵, 而且

$$A(\lambda)A(\lambda)^* = A(\lambda)^*A(\lambda) = |A(\lambda)|I_n.$$

所以$\frac{1}{|A(\lambda)|}A(\lambda)^*$是$A(\lambda)$的逆.

□

定义 6.2.2 (λ-矩阵的初等变换). *对λ-矩阵所做的下面三种变换称为是初等变换:*

1. *交换λ-矩阵的两行(列).*
2. *把λ-矩阵的某一行(列)乘上一个F中的非零元素.*
3. *在λ-矩阵的某一行(列)上, 加上另外一行(列)的$f(\lambda)$倍, 这里的$f(\lambda)$是个多项式.*

定义 6.2.3 (初等λ-矩阵). *以上三种初等变换分别对应下面的三种初等λ-矩阵:*

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

2.

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

其中a是域F中的一个非零元素.

3.

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & f(\lambda) & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

其中$f(\lambda)$是一个多项式.

定义 6.2.4 (λ-矩阵的等价)**.** . 如果$A(\lambda)$, $B(\lambda)$是两个λ-矩阵, 而且如果$A(\lambda)$可以经过有限次初等变换变为$B(\lambda)$, 我们就称$A(\lambda)$, $B(\lambda)$是两个等价的λ-矩阵.

定义 6.2.5 (行列式因子, 不变因子)**.** . 假设$A(\lambda)$是一个$n \times n$ λ-矩阵. 对任意的$1 \leq k \leq n$, 我们称所有的k- 级子式的首项系数为1的最大公因式为$A(\lambda)$的k-级行列式因子, 记为$D_k(\lambda)$. 令$D_0(\lambda) = 1$. 则对任意的$1 \leq k \leq n$, 定义$A(\lambda)$的k级不变因子$d_k(\lambda) = D_k(\lambda)/D_{k-1}(\lambda)$.

定理 6.2.2. λ-矩阵的初等变换不改变其各级行列式因子和不变因子.

证明 根据定义, 我们只需证明初等变换不改变各级行列式因子即可.

假设$A(\lambda)$是一个$n \times n$ λ-矩阵. 第一种初等变换最多把$A(\lambda)$的一个k级子式变成另一个k-级子式, 再乘上一个-1. 这种变换自然不改变最大公因式. 第二种初等变换最多把$A(\lambda)$ 的一个k级子式变成另一个k- 级子式的a倍, a是域中的非零元素. 这种变换也不改变最大公因式.

所以我们只需证明第三种初等变换不改变行列式因子即可. 假设这第三种初等变换是在λ-矩阵的第j行上, 加上了第i 行的$f(\lambda)$ 倍, 把$A(\lambda)$变成了$B(\lambda)$. 我们固定k, 看看$B(\lambda)$ 的k级行列式因子和$A(\lambda)$的k级行列式因子是不是一样的. 如果某

个k级子式不包括第j行, 那么这个k级子式在经过这个初等变换后当然是不改变的. 如果这个k级子式既包含第i 行, 也包含第j行, 那么根据行列式的性质可知, 这个k级子式在经过这个初等变换后也是不改变的. 如果如果这个k 级子式不包含第i行, 但包含第j行. 根据行列式的拆行, 我们这个k级子式都能写成$A(\lambda)$的两个k级子式的线性组合, 而且线性组合的系数是±1. 把$A(\lambda)$ 的某个k 级子式加到另一个k-级子式上去, 这种操作当然不会改变最大公因式.

所以初等变换不改变行列式因子, 自然也就不改变不变因子.

□

定理 6.2.3. *假设$A(\lambda)$是一个λ-矩阵, 则它一定等价于下面形式的矩阵*

$$\begin{pmatrix} d_1(\lambda) &&&&&& \\ & d_2(\lambda) &&&&& \\ && \ddots &&&& \\ &&& d_r(\lambda) &&& \\ &&&& 0 && \\ &&&&& \ddots & \\ &&&&&& 0 \end{pmatrix}$$

其中d_i都是首一多项式, 而且

$$d_i|d_{i+1},\ 1 \le i \le r-1.$$

这种形式的矩阵是由$A(\lambda)$唯一确定的, 与初等变换的选取无关, 叫做$A(\lambda)$的标准形, 其中r叫做$A(\lambda)$的秩.

证明 利用初等变换, 可以把$A(\lambda)$的元素的最大公因式$d_1(\lambda)$ 变出来, 并且放在左上角. 然后利用初等变换, 把A 的第一行和第一列的其他元素都变成0. A就变成了下面的这个样子

$$\begin{pmatrix} d_1(\lambda) & 0 \\ 0 & A_1(\lambda) \end{pmatrix}$$

其中$A_1(\lambda)$是一个$n-1$级的λ-矩阵, 而且它的每个元素都能够被$d_1(\lambda)$整除. 再对$n-1$级的λ-矩阵$A_1(\lambda)$做类似的操作即可看出, A可以最终变为定理中的对角矩阵.

我们下面来证明唯一性. 我们知道初等变换的过程不会改变矩阵行列式因子和不变因子. 而对角矩阵的行列式因子依次为

$$d_1(\lambda),\ d_1(\lambda)d_2(\lambda),\ d_1(\lambda)d_2(\lambda)d_3(\lambda),\ ....$$

所以所有的d_i都是唯一确定的.

□

推论 6.2.4. λ-矩阵$A(\lambda)$可逆当且仅当$A(\lambda)$能经过有限次初等变换变成单位矩阵, 即当且仅当A能写成有限个初等矩阵的乘积

证明 我们前面已经看到λ-矩阵$A(\lambda)$可逆当且仅当$|A(\lambda)|$是域F中的非零元素. $|A(\lambda)|$而除以它的首项系数之后就是$A(\lambda)$的$n \times n$行列式因子. 所以$A(\lambda)$的$n \times n$行列式因子等于1. 所以$A(\lambda)$的标准型就是单位矩阵.

□

§6.3 矩阵的相似

这里我们引入一个新的记号. 任何一个n-级λ-矩阵$M(\lambda)$, 都可以写成下面的形式

$$M(\lambda) = M_k\lambda^k + \cdots + M_1\lambda + M_0$$

的形式, 其中M_i是数域F上的n-级多项式. 例如

$$\begin{pmatrix} \lambda^5 + 2\lambda^4 + 3 & \lambda \\ \lambda^2 + 4 & \lambda^5 + \lambda^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \lambda^5 + \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \lambda^4 + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \lambda^3 + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \lambda^2 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \lambda + \begin{pmatrix} 3 & 0 \\ 4 & 0 \end{pmatrix}.$$

对数域F上的任意矩阵N, 我们定义$M(N)$为下面的式子,

$$M(N) := M_kN^k + \cdots + M_1N + M_0.$$

定理 6.3.1. 假设A, B是数域F上的两个n-级矩阵, 则A与B在数域F上相似当且仅当λ-矩阵$\lambda I - A$与$\lambda I - B$等价.

证明 如果A与B相似, 那么就存在数域F上的可逆矩阵P使得$PAP^{-1} = B$, 所以$P(\lambda I - A)P^{-1} = \lambda I - B$. 数域$F$上的可逆矩阵当然是可逆的$\lambda$-矩阵, 所以$\lambda I - A$与$\lambda I - B$ 等价.

反之, 假设$\lambda I - A$与$\lambda I - B$等价, 即存在可逆λ-矩阵$P(\lambda)$, $Q(\lambda)$使得

$$P(\lambda)(\lambda I - A)Q(\lambda) = \lambda I - B. \tag{$*$}$$

假设

$$Q(\lambda) = (q_{ij}(\lambda)) = Q_k\lambda^k + \cdots + Q_1\lambda + Q_0,$$
$$Q(\lambda)^{-1} = (\overline{q}_{ij}(\lambda)) = \overline{Q}_\ell\lambda^\ell + \cdots + \overline{Q}_1\lambda + \overline{Q}_0,$$

其中Q_i, $\overline{Q}_j$都是数域F上的矩阵. 因为$Q(\lambda)^{-1}Q(\lambda) = I_n$, 所以

$$\overline{Q}_\ell Q(\lambda)\lambda^\ell + \cdots + \overline{Q}_1 Q(\lambda)\lambda + \overline{Q}_0 Q(\lambda) = I_n.$$

把上面的式子中的λ换成B, 等式仍然成立. 所以

$$\overline{Q}_\ell Q(B)B^\ell+\cdots+\overline{Q}_1 Q(B)B+\overline{Q}_0 Q(B)=I_n. \tag{$**$}$$

由($*$)可知

$$(\lambda I-A)Q(\lambda)=P^{-1}(\lambda)(\lambda I-B),$$

所以

$$Q(\lambda)\lambda-AQ(\lambda)=P^{-1}(\lambda)(\lambda I-B).$$

假设

$$P^{-1}(\lambda)=(\overline{p}_{ij}(\lambda))=\overline{P}_m\lambda^m+\cdots+\overline{P}_1\lambda+\overline{P}_0,$$

则

$$\begin{aligned}Q(\lambda)\lambda-AQ(\lambda)=&Q_k\lambda^{k+1}+(Q_{k-1}-AQ_k)+\cdots+(Q_0-AQ_1)\lambda-AQ_0\\=&\overline{P}_m\lambda^{m+1}+(\overline{P}_{m-1}-\overline{P}_mB)\lambda^m+\cdots+(\overline{P}_0-\overline{P}_1B)\lambda-\overline{P}_0B\end{aligned}$$

把$\lambda=B$带入上面的等式, 则最右边是0. 所以最左边也是0. 所以

$$Q(B)B=AQ(B). \tag{$***$}$$

由此易知

$$Q(B)B^2=A^2Q(B),\ Q(B)B^3=A^3Q(B),\ \ldots.$$

所以($**$)就变成了

$$\overline{Q}_\ell A^\ell Q(B)+\cdots+\overline{Q}_1 AQ(B)+\overline{Q}_0 Q(B)=I_n,$$

即

$$(\overline{Q}_\ell A^\ell+\cdots+\overline{Q}_1A+\overline{Q}_0)Q(B)=I_n.$$

所以$Q(B)$是可逆的. 由($***$)可知所以A, B相似. □

推论 6.3.2. 域F上的两个n-级矩阵A, B 相似当且仅当$\lambda I-A$和$\lambda I-B$有相同的行列式因子和不变因子.

推论 6.3.3. 已知A, B是两个n-级的数域F上的矩阵. 如果它们在复数域上相似, 则它们在数域F上相似.

证明 因为A, B在复数域上相似, 所以它们在复数域上有着相同的行列式因子. 由于A, B是两个n-级的数域F上的矩阵, 所以这些行列式因子其实是数域F上的多项式. 当把$\lambda I-A$和$\lambda I-B$的行列式因子看成是复系数多项式的时候它们相等, 它们原来当然也是相等的. 所以A, B在数域F上相似. □

定义 6.3.1 (初等因子)**.** 假设$A(\lambda)$是复数域$\mathbb{C}$上的λ-矩阵, 秩为r. 由于复数域上的多项式都可以写成一次因式的乘积, 所以我们可以假设

$$\begin{aligned} d_1(\lambda) &= (\lambda-\lambda_1)^{a_{11}}\cdots(\lambda-\lambda_k)^{a_{1k}} \\ d_2(\lambda) &= (\lambda-\lambda_1)^{a_{21}}\cdots(\lambda-\lambda_k)^{a_{2k}} \\ &\vdots \\ d_r(\lambda) &= (\lambda-\lambda_1)^{a_{r1}}\cdots(\lambda-\lambda_k)^{a_{rk}} \\ d_{r+1}(\lambda) &= \cdots = d_n(\lambda) = 0, \end{aligned}$$

其中λ_1, λ_2, ..., λ_k是多项式$d_r(\lambda)$的所有的不同的根, 指数满足关系式

$$\begin{gathered} 0 \le a_{11} \le \cdots \le a_{r1}, \\ \vdots \\ 0 \le a_{1k} \le \cdots \le a_{rk}. \end{gathered}$$

注意这里面的指数是可以为0的. 如果$a_{ij}>0$, 我们就称$(\lambda-\lambda_j)^{a_{ij}}$是属于$\lambda_j$的一个初等因子, 所有的初等因子合起来叫做是$A(\lambda)$的初等因子组.

如果两个λ矩阵具有相同的不变因子, 那么根据定义, 它们也具有相同的初等因子. 反之, 则未必成立. 例如

$$A(\lambda)=\begin{pmatrix}(\lambda-1)(\lambda-2) & 0\\ 0 & 0\end{pmatrix},\ B(\lambda)=\begin{pmatrix}(\lambda-1) & 0\\ 0 & (\lambda-2)\end{pmatrix}$$

具有相同的初等因子组, 但是他们的秩不同, 因此他们不可能等价, 自然也就不可能有相同的不变因子.

定理 6.3.4. 假设$A(\lambda)$, $B(\lambda)$是复数域$\mathbb{C}$上的两个λ-矩阵, 秩都为r, 而且具有相同的初等因子组, 则$A(\lambda)$, $B(\lambda)$ 是等价的.

证明

首先由于两个矩阵的秩相同, 所以它们的大于r级的不变因子都相等, 都等于0.

我们先来说明$d_r(\lambda)$完全由初等因子决定. 我们知道每个初等因子都是形如$(\lambda-\lambda_j)^*$形式的, 这里的星号表示指数未定. 这个初等因子叫做属于λ_j的初等因子. 我们把属于λ_j的初等因子中次数最高的挑出来, 记为$(\lambda-\lambda_j)^{a_{rj}}$. 那么根据初等因子的定义, 我们知道

$$d_r(\lambda)=(\lambda-\lambda_1)^{a_{r1}}(\lambda-\lambda_2)^{a_{r2}}\cdots(\lambda-\lambda_k)^{a_{rk}}.$$

所以$d_r(\lambda)$完全由初等因子决定, 它就等于把属于不同的λ_j的初等因子中次数最高的乘在一起得到的.

在初等因子组中把$(\lambda-\lambda_1)^{a_{r1}}$, $(\lambda-\lambda_2)^{a_{r2}}$, $\cdots$, $(\lambda-\lambda_k)^{a_{rk}}$去掉.

如果还有剩余的初等因子, 那么在剩下的初等因子中重复上一段的操作,把幂次最高的初等因子乘在一起, 就得到了$d_{r-1}(\lambda)$. 以此类推, 直到所有的初等因子都用完. 假设定出$d_t(\lambda)$以后, 就没有剩余的初等因子了. 那么我们就有

$$d_{t-1}(\lambda)=\cdots=d_1(\lambda)=1.$$

因此可以看出, 所有的不变因子都有初等因子决定.

所以$A(\lambda)$, $B(\lambda)$具有相同的不变因子, 是等价的.

□

至此, 我们可以看出, 对复方阵A来说, $\lambda I-A$在等价关系之下的等价类, 相当于A在相似关系之下的等价类. 我们以后称$\lambda I-A$为方阵A的特征方阵, 把$\lambda I-A$的行列式因子, 不变因子和初等因子简单的称为A的行列式因子, 不变因子和初等因子. 由前面的讨论可知, A的所有的不变因子的乘积等于A的特征多项式.

定理 6.3.5. *假设A, B是复数域$\mathbb{C}$上的两个$n\times n$矩阵. 则下面的命题等价*

1. *A, B相似.*
2. *$\lambda I-A$, $\lambda I-B$等价.*
3. *A, B有相同的行列式因子.*
4. *A, B有相同的不变因子.*
5. *A, B有相同的初等因子.*

证明 注意到$\lambda I-A$, $\lambda I-B$的秩都等于n即可. □

§6.4 Jordan标准形

这一节, 我们要推导出复矩阵的Jordan标准形的理论.

引理 6.4.1. *假设*

$$A(\lambda)=\begin{pmatrix}f_1(\lambda)g_1(\lambda) & 0\\ 0 & f_2(\lambda)g_2(\lambda)\end{pmatrix},\ B(\lambda)=\begin{pmatrix}f_1(\lambda)g_2(\lambda) & 0\\ 0 & f_2(\lambda)g_1(\lambda),\end{pmatrix}$$

而且对任意的$1\le i,\ j\le 2$, 都有$f_i(\lambda)$与$g_j(\lambda)$互素, 则$A(\lambda)$, $B(\lambda)$等价.

证明 容易看出$A(\lambda)$, $B(\lambda)$的2级行列式因子相等, 即它们的行列式相等. 而它们的一级行列式因子都是

$$\begin{aligned}(f_1(\lambda)g_1(\lambda),\ f_2(\lambda)g_2(\lambda))&=(f_1(\lambda)g_2(\lambda),\ f_2(\lambda)g_1(\lambda))\\&=(f_1,\ f_2(\lambda))(g_1(\lambda),\ g_2(\lambda)).\end{aligned}$$

□

定理 6.4.2. 准对角矩阵

$$C(\lambda)=\begin{pmatrix}A(\lambda) & 0\\ 0 & B(\lambda)\end{pmatrix}$$

的初等因子组是由$A(\lambda),B(\lambda)$的初等因子组合并得到的.

证明 我们知道初等变换不改变矩阵的初等因子组. 所以我们不妨设$A(\lambda),B(\lambda)$都是标准形, 而且对角线上的元素都展开成了初等因子的乘积. 然后我们反复使用上面的引理, 让$A(\lambda),B(\lambda)$的初等因子在$2n\times 2n$的矩阵的对角线上移动, 相互交换位置, 最终可以使得属于同一个λ_j的初等因子中幂次越高的, 越在下方. 对角线上的元素, 前面的整除后面的. 所有的零行都在非零行的下方, 所有的零列都在非零列的右边. 按照定义, 满足这个条件的矩阵是一个标准形.

因此我们可以从准对角矩阵$C(\lambda)$出发, 经过上一段中中的初等变换操作, 变成标准形. 在这个过程中我们可以看出来, $A(\lambda),B(\lambda)$的初等因子, 仅仅是移动了位置, 一个也不多, 一个也不少地出现在了$C(\lambda)$的标准形中. 所以$C(\lambda)$的初等因子组是由$A(\lambda),B(\lambda)$的初等因子组合并得到的. □

命题 6.4.3. 假设A, B是域F上的两个n-级矩阵, 则A与B相似当且仅当

$$\begin{pmatrix}A & \\ & A\end{pmatrix} \text{ 与 } \begin{pmatrix}B & \\ & B\end{pmatrix}$$

相似.

证明 必要性很简单, 我们只证充分性. 我们知道

$$\begin{pmatrix}A & \\ & A\end{pmatrix} \text{ 与 } \begin{pmatrix}B & \\ & B\end{pmatrix}$$

的初等因子组分别是A与B的初等因子组重复一遍得到的. A与B的初等因子组重复一遍以后是相同的, 所以A 与B的初等因子组原来也是相同的. 所以A与B 相似.

□

命题 6.4.4. 假设

$$A=\begin{pmatrix}a & 1 & & & \\ & a & 1 & & \\ & & \ddots & \ddots & \\ & & & a & 1\\ & & & & a\end{pmatrix}_{m\times m}$$

为Jordan块. 则$\lambda I-A$的初等因子为$(\lambda-a)^m$.

证明 $\lambda I-A$中能找到行列式为±1的子式, 所以

$$d_1(\lambda)=\cdots=d_{m-1}(\lambda)=1,\ d_m(\lambda)=(\lambda-a)^m.$$

所以$\lambda I-A$的初等因子为$(\lambda-a)^m$. □

定理 6.4.5. *任意一个复矩阵A都相似于一个准对角矩阵, 其中每一个准对角块都是一个Jordan块. 这个准对角矩阵叫做是A 的Jordan 标准形. 在不计对角线上的对角块排列次序的情况下, Jordan标准形是唯一的, 由A完全决定.*

证明 假设A的初等因子为

$$(\lambda-\lambda_1)^{k_1},\ ...,\ (\lambda-\lambda_m)^{k_m},$$

注意这里的$\lambda_1,\ \lambda_2,\ ...,\lambda_m$不一定互不相同.

对任意$1\le i\le m$, 令

$$J_i=\begin{pmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i\end{pmatrix}_{k_i\times k_i},$$

则J_i的初等因子为$(\lambda-\lambda_i)^{k_i}$. 所以

$$\lambda I-\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m\end{pmatrix}$$

的初等因子组为

$$(\lambda-\lambda_1)^{k_1},\ ...,\ (\lambda-\lambda_m)^{k_m}.$$

由于A和

$$\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m\end{pmatrix}$$

具有相同的初等因子组, 所以他们等价, 所以A相似于

$$\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m\end{pmatrix}.$$

由于每一个Jordan块都对应于一个初等因子. 而初等因子组是由A完全决定的(不计初等因子组中初等因子的排列次序), 所以Jordan块也是由A完全决定的(同样的, 不计Jordan块的排列次序).

□

命题 6.4.6. 假设A是域F上的一个n-级矩阵. 则A的n-级不变因子$d_n(\lambda)$等于A的最小多项式$d(\lambda)$.

证明 我们把域从F扩充到复数域$\mathbb{C}$上. 只要证明了到复数域$\mathbb{C}$上, A的n-级不变因子$d_n(\lambda)$等于A 的最小多项式$d(\lambda)$, 那么这个结论在F 上自然也成立. 由于相似变换不改变矩阵的不变因子和最小多项式, 所以我们可以假设A就是一个Jordan 标准形矩阵, 它的准对角块是Jordan块. 我们进一步假设具有相同特征值的Jordan块排列在一起, 即

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix},$$

其中每个A_i都只有一个特征值λ_i(不计重复). 我们假设

$$A_i = \begin{pmatrix} A_{i1} & & & \\ & A_{i2} & & \\ & & \ddots & \\ & & & A_{it_i} \end{pmatrix},$$

其中A_{it_i}的级数最高, 是k_i-级的Jordan块. 则

$$(\lambda I - A_{it_i})^{k_i - 1} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, \ (\lambda I - A_{it_i})^{k_i} = 0.$$

所以A_{it_i}的最小多项式是$(\lambda - \lambda_i)^{k_i}$. 所以$A$的最小多项式等于

$$d(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_m)^{k_m},$$

即$d(\lambda)$由属于不同特征值的初等因子中幂次最高者相乘得到. 由前面讨论过的从初等因子构造不变因子的过程可知, $d(\lambda)$就是n-级不变因子.

□

定理 6.4.7 (Weyl)**.** 假设A, B是两个n-级复方阵. 则A与B相似当且仅当对任意的正整数k和复数λ 都有

$$rank(\lambda I - A)^k = rank(\lambda I - B)^k.$$

证明 显然只需要证明充分性.

假设A, B都是Jordan标准形, . 假设λ_i是A的一个特征值, 则当k充分大时,

$$\mathrm{rank}(\lambda I - A)^k = n - (\lambda_i \text{ 的重数}).$$

由此可知. A和B的特征多项式是相同的. 所以A的属于λ_iJordan块的级数之和等于B的属于λ_i 的Jordan块的级数之和. 我们假设具有相同特征值的Jordan块排列在一起, 即

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}, \ B = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_m \end{pmatrix},$$

其中A_i, B_i的都只有一个特征值λ_i, 且级数相同, 而且

$$\mathrm{rank}(\lambda I - A_i)^k = \mathrm{rank}(\lambda I - B_i)^k$$

对任意正整数k和复数λ成立. 我们主要能证明A_i相似于B_i即可. 由$\mathrm{rank}(\lambda_i I - A_i) = \mathrm{rank}(\lambda_i I - B_i)$可知, 二者的Jordan块的个数相同. 由$\mathrm{rank}(\lambda_i I - A_i)^2 = \mathrm{rank}(\lambda_i I - B_i)^2$可知$A_i$和$B_i$中级数大于等于2的Jordan块的个数相同. 依次做下去, 可知A_i和B_i中的任意级的Jordan块的个数都是相同的, 所以二者相似, 所以A和B相似.

□

注意, 即使对任意的复数λ 都有

$$\mathrm{rank}(\lambda I - A) = \mathrm{rank}(\lambda I - B),$$

也不一定有A与B相似. 反例如

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

容易看出, 对任意的复数λ 都有$\mathrm{rank}(\lambda I - A) = \mathrm{rank}(\lambda I - B)$, 但是$A$和$B$并不相似.

命题 6.4.8. 对任何一个复矩阵A, 下列命题等价:

1. A可以对角化.

2. A的最小多项式$d(\lambda)$没有重根.

3. A的n-级不变因子没有重根.

4. A的所有的不变因子都没有重根.

5. A的所有的初等因子都是一次的.

证明 由前面的讨论, 容易证明.

□

定理 6.4.9. *假设$\mathscr{A}$是n维复线性空间V上的一个线性变换. 则下列命题等价.*
(1) V是$\mathscr{A}$的一个循环子空间.
(2) $\mathscr{A}$的最小多项式是n次的, 即最小多项式等于特征多项式.
(3) $\mathscr{A}$的任意特征值的特征子空间都是1维的.
(4) $\mathscr{A}$的前$n-1$级不变因子$d_1(\lambda)$, ..., $d_{n-1}(\lambda)$都是1.
(5) $\mathscr{A}$的第n个不变因子等于$\mathscr{A}$的特征多项式.
(6) $\mathscr{A}$的前$n-1$级行列式因子$D_1(\lambda)$, ..., $D_{n-1}(\lambda)$都是1.
(7) 与$\mathscr{A}$可交换的所有的线性变换的全体形成的线性空间是n维的.
(8) $\mathscr{A}$的若尔当标准形中任意两个若尔当块的特征值都不相同.

证明 (1)$\Longrightarrow$ (2). 如果$\mathscr{A}$的最小多项式$g(\lambda)$次数$m < n$, 那么$g(\mathscr{A}) = 0$是零变换, 所以$g(\mathscr{A})(\alpha) = 0$对任意的$\alpha \in V$成立, 所以$\alpha$, $\mathscr{A}(\alpha)$, $\mathscr{A}^2(\alpha)$, ..., $\mathscr{A}^m(\alpha)$肯定线性相关, 所以生成的子空间的维数小于$m+1 \le n$. 所以V不可能是一个循环子空间, 导致矛盾.

(2)$\Longrightarrow$ (1). 假设$\mathscr{A}$的最小多项式等于特征多项式, 都为

$$f_{\mathscr{A}}(\lambda) = (\lambda - \lambda_1)^{a_1} \cdots (\lambda - \lambda_k)^{a_k}.$$

则对任意的i, 存在$\alpha_i \in \mathrm{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{a_i}$, 但$\alpha_i \notin \mathrm{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{a_i - 1}$. 所以

$$\alpha_i,\ (\mathscr{A} - \lambda_i \mathscr{I})(\alpha_i),\ ...,\ (\mathscr{A} - \lambda_i \mathscr{I})^{a_i - 1}(\alpha_i)$$

这a_i个向量是线性无关的, 可以作为特征值λ_i的a_i维的根子空间V_i的一组基. 易见这a_i个向量组成的向量组和向量组

$$\alpha_i,\ \mathscr{A}(\alpha_i),\ ...,\ \mathscr{A}^{a_i - 1}(\alpha_i)$$

是等价的. 所以特征值λ_i的a_i维的根子空间是$\mathscr{A}$的一个循环子空间.

令$\alpha=\alpha_1+\cdots+\alpha_k$, 我们断言$V$可由$\alpha$生成. 为此, 只需证明任何一个$\alpha_i$都能由$\alpha$生成即可. 为了符号简单, 我们下面只说明$\alpha_1$能由$\alpha$生成即可, 其余的类似. 记$V_1$为特征值$\lambda_1$的根子空间. 则

$$(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}(\alpha)=(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}(\alpha_1).$$

而V_1是$(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}$的不变子空间, 且$(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k}$在$V_1$上是可逆变换. 由于$(\lambda-\lambda_1)^{a_1}$和$f(\lambda)/(\lambda-\lambda_1)^{a_1}$互素, 因此存在多项式$h_1,\ h_2$使得

$$h_1(\lambda-\lambda_1)^{a_1}+h_2f/(\lambda-\lambda_1)^{a_1}=1.$$

所以

$$(h_1(\mathscr{A})(\mathscr{A}-\lambda_1\mathscr{I})^{a_1}+h_2(\mathscr{A})(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k})(\alpha_1)=\alpha_1.$$

但

$$(h_1(\mathscr{A})(\mathscr{A}-\lambda_1\mathscr{I})^{a_1})(\alpha_1)=0,$$

所以

$$h_2(\mathscr{A})(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k})(\alpha_1)=\alpha_1.$$

所以α_1在由$(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k})(\alpha_1)$生成的$\mathscr{A}$ 的循环子空间中, 而$(\mathscr{A}-\lambda_2\mathscr{I})^{a_2}\cdots(\mathscr{A}-\lambda_k\mathscr{I})^{a_k})(\alpha_1)$ 在由α生成的$\mathscr{A}$的循环子空间中. 所以α_1也在由α生成的$\mathscr{A}$的循环子空间中.

(2)$\Longrightarrow$ (3). 利用

$$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})\subsetneq\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^2\subsetneq\cdots\subsetneq\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i}=V_i$$

这条子空间的严格单调递增链的最后一个维数是a_i. 所以$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$只能是1维的. 所以$\mathscr{A}$的任意特征值的特征子空间都是1维的.

(3)$\Longrightarrow$ (2). 此时$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$是1维的. 我们来看$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^2$的维数. 任取$\alpha_1,\ \alpha_2\in\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^2-\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$. 则$(\mathscr{A}-\lambda_i\mathscr{I})(\alpha_1)\in\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$, $(\mathscr{A}-\lambda_i\mathscr{I})(\alpha_2)\in\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$. 而$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$ 是1维的, 所以存在常数t使得$(\mathscr{A}-\lambda_i\mathscr{I})(\alpha_1)=(\mathscr{A}-\lambda_i\mathscr{I})(t\alpha_2)$. 所以$\alpha_1-t\alpha_2\in\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$. 这说明$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^2$的维数比$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})$ 最多增加1. 依此类推, 可知$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i-1}$的维数最多是$a_i-1$, 不可能是$a_i$ 维的. 所以$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i-1}\neq V_i$, $\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{a_i}=V_i$. 所以最小多项式等于特征多项式.

(2), (4), (5), (6), (8)等价比较简单.

我们下面证明(7), (8)等价. 假设$\mathscr{A}$的若尔当标准型为$A=diag(A_1,\ ...,\ A_k)$, 其中每个A_i都由一个或者多个具有相同特征值的若尔当块组成, 不同的A_i 的特征值

互不相同. 假设$B=(B_{ij})$与A可交换. 那么

$$A_iB_{ij}=B_{ij}A_j,$$

所以$\varphi(A_i)B_{ij} = B_{ij}\varphi(A_j)$, 由于$A_i$和$A_j$的特征值不同, 取$\varphi$为$A_i$的特征多项式可知$B_{ij}=0$. 所以可以假设

$$B=diag(B_1,\ ...,\ B_k).$$

所以我们一开始就可以假设A的特征值都是相同的, 进一步地, 减去一个纯量阵之后, 我们还可以假设A是一个幂零矩阵. 即我们可以不妨假设A_i都是幂零的若尔当块. 容易验证, 跟每个幂零的若尔当块可以交换的矩阵形成的空间维数等于这个若尔当块的维数. 但是使得

$$A_iB_{ij}=B_{ij}A_j,$$

的B_{ij}就不再必须等于0了. 所以总维数一定会大于n.

因此维数等于n当且仅当每个特征值只有一个若尔当块.

□

§6.5 有理标准形

定义 6.5.1. 假设数域F上的首一多项式

$$d(\lambda)=\lambda^n+a_{n-1}\lambda+\cdots+a_1\lambda+a_0.$$

我们称矩阵

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

为多项式$d(\lambda)$的友矩阵.

容易看出$\lambda I-A$的$n-1$级行列式因子为1. 所以它的不变因子为

$$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1,\ d_n(\lambda)=d(\lambda).$$

所以任何一个首一多项式都是某个矩阵的特征多项式. 任何一个首一多项式也是某个矩阵的最小多项式.

定义 6.5.2 (有理标准形矩阵). 如果准对角矩阵

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

的每个对角块A_i都是某个首一多项式$d_i(\lambda)$的友矩阵, 而且对任意的$1\le i\le k-1$都有

$$d_i(\lambda)|d_{i+1}(\lambda),$$

我们就称A是一个有理标准形矩阵.

定理 6.5.1. 数域F上的任何一个矩阵A都相似于唯一的一个有理标准形矩阵.

证明 我们假设$\lambda I-A$的标准形为

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(\lambda) & & \\ & & & & \ddots & \\ & & & & & d_k(\lambda) \end{pmatrix},$$

其中$d_1|d_2|\cdots$. 假设$d_1(\lambda)$的友矩阵为A_1, ..., $d_k(\lambda)$的友矩阵为A_k, 令

$$B=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}.$$

则B也是一个n-级矩阵. 而且因为$\lambda I-A_i$ 等价于

$$B_i(\lambda)=\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d_i(\lambda) \end{pmatrix},$$

所以$\lambda I-B$ 等价于

$$\begin{pmatrix} B_1(\lambda) & & \\ & \ddots & \\ & & B_k(\lambda) \end{pmatrix}.$$

使用换行换列的初等变换, 可以将上面的λ矩阵等价于

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(\lambda) & & \\ & & & & \ddots & \\ & & & & & d_k(\lambda) \end{pmatrix}.$$

所以$\lambda I - A$和$\lambda I - B$具有相同的λ矩阵标准形. 所以A和B相似. 而且由λ矩阵标准形的唯一性可知有理标准形的唯一性. □

1. 已知A, B是两个n-级实数矩阵, 即它们的元素都是实数. 如果它们在复数域上相似(即存在可逆的复矩阵P使得$PAP^{-1} = B$), 则它们在实数域上相似(即存在可逆的实矩阵P使得$PAP^{-1} = B$).

   证明

   假设$P = R + iQ$, 其中R, Q都是实方阵. 则因为

   $$(R + iQ)A = B(R + iQ),$$

   所以

   $$RB = AR,\ QB = AQ,$$

   所以对任意的实数λ, 都有

   $$(R + \lambda Q)A = B(R + \lambda Q).$$

   记$f(\lambda) = \det(R + \lambda Q)$, 则$f(\lambda)$是一个非零多项式(因为$f(i) \neq 0$). 我们知道非零的多项式只有有限多个根, 所以一定存在实数a使得$f(a) \neq 0$. 所以$\det(R + aQ) \neq 0$, 所以$R + aQ$是可逆的实方阵. 所以A, B在实数域上相似.

   □

2. 模仿上一道题目的做法证明: 已知A, B是两个n-级的数域F上的矩阵. 如果它们在复数域上相似, 则它们在数域F上相似. (汤馥羽)

   证明 设可逆的复矩阵P使得$PAP^{-1} = B$, 考虑由P的所有元素张成的F上的线性空间E. 注意E只是F上的有限维线性空间, 不也是F的扩域. 设E作为F上的线性空间有一组基e_1, ..., e_t, 则假设

   $$P = \sum_{i=1}^{t} e_i P_i,$$

其中$P_1, ..., P_t$都是数域F上的方阵. 则因为

$$(\sum_{i=1}^{t} e_i P_i)A = B(\sum_{i=1}^{t} e_i P_i),$$

所以

$$P_i A = BP_i,\ i = 1,\ ...,\ t.$$

所以对任意的$x_1, ..., x_t \in F$, 都有

$$(\sum_{i=1}^{t} x_i P_i)A = B(\sum_{i=1}^{t} x_i P_i).$$

我们需要找合适的$x_1, ..., x_t \in F$使得$(\sum_{i=1}^{t} x_i P_i)$可逆, 即找合适的$x_1, ..., x_t \in F$使得$|\sum_{i=1}^{t} x_i P_i| \neq 0$. 假设

$$f(x_1,\ ...,\ x_t) = |\sum_{i=1}^{t} x_i P_i|.$$

则$f(x_1, ..., x_t)$在复数域中当$x_1 = e_1, ..., x_t = e_t$时, 可以取到非零值. 所以$f(x_1, ..., x_t)$不是零多项式. 我们下面来证明数域$F$上的非零多项式, 一定在数域$F$上能取到非零值.

我们用数学归纳法。假设命题对变量个数小于t的多项式成立, 即如果变量个数小于t的多项式f 取值恒等于0, 则这个多项式是零多项式。

我们不妨假设多项式f有某个单项式包含x_1。那么f 可以看成是关于x_1的多项式，其余的变量$x_2, ..., x_t$我们都看成常数。这个多项式的最高项的系数是一个关于$x_2, ..., x_t$的多项式$g(x_2, ..., x_t)$。这个多项式不是零多项式.根据归纳假设，它取值不能恒等于0。所以存在$a_2, ..., a_t \in F$使得$g(a_2, ..., a_t) \neq 0$.

这时$f(x_1, a_2, ..., a_t)$就成为了关于x_1的非零多项式. 这个多项式当然只能有有限个解。特别的存在$a_1 \in F$, 使得

$$f(a_1,\ a_2,\ ...,\ a_t) \neq 0.$$

所以

$$\sum_{i=1}^{t} x_i P_i$$

是可逆的数域F上的矩阵. 所以A, B在实数域上相似.

□

3. 如果

$$\begin{pmatrix} A & \\ & A \end{pmatrix} \text{ 和 } \begin{pmatrix} B & \\ & B \end{pmatrix}$$

相似, 求证A, B相似.

4. 假设A是$m \times n$的复矩阵, B和C分别是n-级和m-级复方阵, B, C没有公共的特征值, 而且$AB = CA$. 求证$A = 0$.

证明 假设f是B的特征多项式, 则由$AB = CA$可得

$$Af(B) = f(C)A.$$

因为B, C没有公共的特征值, 所以$f(C)$可逆. 由哈密顿-凯莱定理可知, $f(B) = 0$. 所以$A = 0$.

□

5. 假设n1-级矩阵$(n > 1)$,

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix},$$

$AB = BA$. 求证B可以写成A的多项式, 即存在多项式f, 使得$B = f(A)$.

证明 假设$B = (b_{ij})$, 则由$AB = BA$, 可得

$$\begin{pmatrix} b_{21} & b_{22} & \cdots & b_{2n} \\ b_{31} & b_{32} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \\ 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & b_{11} & b_{12} & \cdots & b_{1n-1} \\ 0 & b_{21} & b_{22} & \cdots & b_{2n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & b_{n1} & b_{n2} & \cdots & b_{nn-1} \end{pmatrix},$$

所以我们有

$$\begin{aligned} b_{ij} &= 0, \quad \text{如果} i > j, \\ b_{11} &= \cdots = b_{nn}, \\ b_{12} &= \cdots = b_{n-1n}, \\ &\vdots \\ b_{1n-1} &= b_{2n}, \end{aligned}$$

即

$$B = \begin{pmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ & b_1 & b_2 & \cdots & b_{n-1} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & b_2 \\ & & & & b_1 \end{pmatrix}$$
$$= b_1 I + b_2 A + b_3 A^2 + \cdots + b_n A^{n-1}.$$

令

$$f(x) = b_1 + b_2 x + b_3 x^2 + \cdots + b_n x^{n-1}.$$

□

6. 假设n-级矩阵$(n > 1)$,

$$A = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix},$$

$AB = BA$. 求证B可以写成A的多项式, 即存在多项式f, 使得$B = f(A)$.

证明 由上一题可知, 对$A-\lambda I$, 存在g使得$B = g(A-\lambda I)$. 所以令$f(x) = g(x-\lambda)$, 则$B = f(A)$. □

7. 假设

$$A = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

其中J_i为特征值是λ_i的Jordan块, 而且λ_1, ..., λ_k两两不同, $AB = BA$. 求证B可以写成A的多项式.

证明 我们不妨假设特征值都不为零, 否则把每个在A上加上适当的λI, 就可以做到所有的特征值都不为零, 不影响结论成立.

把B做相同的分块, 由习题4可知, B也是形状相同的准对角矩阵,

$$B=\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}.$$

所以存在多项式f_1, ..., f_k使得

$$B_i=f_i(J_i).$$

由于J_1, ..., J_k的特征多项式g_1, ..., g_k两两互素, 所以由中国剩余定理, 存在多项式f使得

$$\begin{cases} f\equiv f_1 \mod g_1 \\ \vdots \\ f\equiv f_k \mod g_k \end{cases}$$

z这时候就有

$$B_i=f(J_i).$$

□

8. 如果数域F上的n-级矩阵A的特征多项式等于最小多项式, 则与A交换的方阵B可以写成方阵A的多项式.

   证明 我们知道如果B可以写成方阵A的多项式, 那么$AB=BA$. 可以写成方阵A的多项式的矩阵全体构成F上的n-维空间W, 这是因为最小多项式是n次的. 在复数域上这个结论仍是成立的, 即可以写成方阵A的多项式的矩阵全体构成F上的n-维空间W.

   假设数域F上与A交换的方阵B形成的空间为U. 如果U的维数大于n, 那么在复数域上看U的维数也大于n. 这就与前面的结果矛盾了. □

9. 假设数域F上方阵B和每一个与A可交换的矩阵都可交换, 求证方阵B可以写成方阵A的多项式.

   证明 我们不妨假设A是有理标准形矩阵. 否则的话, 我们选一个适当的可逆矩阵P, 用PAP^{-1}, PBP^{-1}代替A, B证明命题即可.

   我们假设$A=diag(A_1, ..., A_r)$, $D=diag(\lambda_1 I, ..., \lambda_r I)$, D的分块方式与A相同, 而且λ_1, ..., λ_r两两不同. 则我们有

$$AD=DA.$$

因此$AB = BA$. 因此A也必须是准对角矩阵, 分块方式与A相同. 即$B = diag(B_1,\ ...,\ B_r)$, 对每个$1 \leq i \leq r$, B_i与每个与A_i交换的矩阵都可以交换, 特别的$A_iB_i = B_iA_i$. 由于A_i是有理标准形矩阵, 所以它的特征多项式等于最小多项式. 由前面的讨论可知, B_i可以写成A_i的多项式, 即$B_i = f_i(A_i)$.

我们现在面临的问题是要找一个统一的f使得$B_i = f(A_i)$对任意i都成立. 这时候就不能像前面那样证明了. 因为这时候$A_1,\ ...,\ A_r$的特征多项式不是互素的. 我们下面来说明要找的这个f其实就是f_r.

我们假设空间V分解成了循环子空间的直和,

$$V = C(\alpha_1) \bigoplus \cdots \bigoplus C(\alpha_r),$$

每个$C(\alpha_i)$都对应于A_i, A_i的特征多项式为d_i, 而且$d_1|d_2|\cdots$. 假设$\mathscr{B}$对应于矩阵B,

$$\mathscr{B}(\alpha_i) = f_i(\mathscr{A})(\alpha_i).$$

假设$d_r = d_ih_i,\ i = 1,\ 2,\ ...,\ r$, 则构造$\mathscr{D}_i : V \longrightarrow V$,

$$\mathscr{D}_i|_{C(\alpha_i)} : g(\mathscr{A})(\alpha_i) \mapsto g(\mathscr{A})(\alpha_i + h_i(\mathscr{A})\alpha_r), \quad \text{for any } g$$
$$\mathscr{D}_i|_{C(\alpha_r)} : g(\mathscr{A})(\alpha_r) \mapsto g(\mathscr{A})(\alpha_i + \alpha_r), \quad \text{for any } g$$
$$\mathscr{D}_i|_{C(\alpha_j)} : identity,\ j \neq i,\ r.$$

这里的"identity"表示恒同变换. 则这是一个线性变换, 而且与$\mathscr{A}$可以交换, 所以和$\mathscr{B}$可以交换. 注意这里首先要说明这个线性变换是良好定义的; 即如果

$$g_1(\mathscr{A})(\alpha_i) = g_2(\mathscr{A})(\alpha_i),$$

那么像是相同的, 这个容易验证. 所以$\mathscr{D}_i\mathscr{B}(\alpha_r) = \mathscr{B}\mathscr{D}_i(\alpha_r)$. 所以

$$f_r(\mathscr{A})(\alpha_i + \alpha_r) = f_i(\mathscr{A})(\alpha_i) + f_r(\mathscr{A})(\alpha_r).$$

所以

$$f_r(\mathscr{A})(\alpha_i) = f_i(\mathscr{A})(\alpha_i).$$

证完.

□

10. 假设方阵A不可逆, 求证$r(A) = r(A^2)$的充分必要条件是方阵A的属于特征值0的初等因子都是一次的. □

第七章　二次型

固定域F, F上的二次齐次多项式称为是二次型. 我们通常把F中的元素0也看做是二次型, 称之为平凡二次型. 二次型的一般形式是

$$f(x_1,\ ...,\ x_n) = \sum_{n \geq i \geq j \geq 1} c_{ij} x_i x_j.$$

为了写成对称的形式, 我们通常把上式改写为

$$f(x_1,\ ...,\ x_n) = \sum_{i,\ j=1}^{n} a_{ij} x_i x_j,$$

其中$a_{ij} = a_{ji}$.

§7.1　二次型的矩阵表示和矩阵的合同

任何一个二次型都可以

$$\begin{aligned} f(x_1,\ ...,\ x_n) &= \sum_{i,\ j=1}^{n} a_{ij} x_i x_j, \\ &= X'AX, \end{aligned}$$

其中$X' = (x_1,\ ..,\ x_n)$是一个由变量组成的n维行向量, $A = (a_{ij})$是一个n阶对称矩阵. 显然在对称矩阵和二次型之间有一一对应的关系.

如果A是一个对角矩阵, 我们称这个二次型是一个标准型. 我们假设$Y = (y_1,\ ..,\ y_n)'$是由另外一组变量组成的行向量, 这两组变量之间的关系式为

$$X = CY,$$

其中$C = (c_{ij})$是一个n阶矩阵. 那么原来的关于$x_1,\ ..,\ x_n$的二次型就变成了新的关于$y_1,\ ..,\ y_n$的二次型

$$g(y_1,\ ..,\ y_n) = Y'DY,$$

其中$D = C'AC$.

我们称$X = CY$为一个线性替换. 如果C可逆, 我们称之为非退化的线性替换, 我们以下考虑的都是非退化的线性替换. 线性替换可以复合, 非退化的线性替换的复合还是非退化的线性替换. 如果两个二次型之间可以用非退化的线性替换从一个变成另一个, 我们就称这两个二次型是等价的.

定理 7.1.1. *任何一个二次型都可以经过非退化的线性替换变成标准型. 等价的说法是, 对任何一个对称矩阵A, 都存在一个可逆矩阵C, 使得$C'AC$是对角矩阵.*

证明 我们用数学归纳法, 假设定理对$n-1$个变量的二次型成立.

首先我们可以假设二次型有平方项. 如果没有的话, 那我们随便选择二次型表达式中出现的一个交叉项x_ix_j. 然后令

$$x_i = y_i + y_j,$$
$$x_j = y_i - y_j$$

就可以制造出平方项来. 因此我们只需要对有平方项的二次型来证明定理成立即可.

现在我们不妨假设假设这个二次型有平方项x_1^2, 而且这一项的系数为$a \neq 0$. 那我们可以假设二次型的其他变量是常数, 这时候二次型可以写成

$$f = ax_1^2 + 2gx_1 + h = a(x_1 + g/a)^2 + h - (g/a)^2,$$

其中g是关于x_2, ..., x_n的一个线性组合, h 是关于x_2, ..., x_n的一个二次型. 我们令

$$y_1 = x_1 + g/a$$
$$y_i = x_i,\ i \geq 2;$$

易见这是一个非退化的线性替换. 这时, 二次型变为

$$f = ay_1^2 + h - (g/a)^2.$$

注意这里的$h-(g/a)^2$原来是关于x_2, ..., x_n的二次型. 变量替换之后变为关于y_2, ..., y_n的二次型. 根据归纳假设, 定理对$n-1$个变量的情形成立. 所以$h-g^2$可以经过非退化的线性替换变为标准型. 因此原来的二次型可以经过非退化的线性替换变成标准型. □

定义 7.1.1. *假设A, B是数域F上的两个n阶矩阵, 如果存在一个n阶的可逆矩阵C使得*

$$A = C'BC,$$

我们就称A合同于B, 或者说A和B是合同的. 合同关系是一个等价关系, 满足自反性, 对称性和传递性.

任何一个矩阵A和它的转置A'都是合同的, 不过这个结论并不容易证明. 合同关系下的标准型理论也比较复杂. 但是对称矩阵的标准型理论很简单. 上面的定理说明任何一个对称矩阵都可以合同于一个对角矩阵.

这里要注意二次型的标准型并不唯一. 两个不相同的对角矩阵完全可以是合同的. 想要找到类似矩阵的秩那样的不变量, 我们需要更精细的规范型理论.

我们下面来看一看针对一个具体的对称矩阵A, 怎样找一个可逆矩阵C, 使得

$$C'AC$$

是一个对角矩阵.

定义 7.1.2 (初等合同变换)**.** *假设A是一个$m\times n$矩阵. 初等合同变换包括以下三种*

1. *交换A的第i, j行, 再交换A的第i, j列.*

2. *把A的第i行乘上非零的常数t, 再把其第i列乘上相同的常数t.*

3. *把A的第i行乘上非零的常数t, 然后加到第j行上去. 再把A的第i列乘上相同的常数t, 然后加到第j 列上去.*

简单的说就是对A的行和列同时做相同的初等变换.

我们前面的定理相当于说任何一个对称的矩阵, 都可以经过一系列的初等合同变化变成对角矩阵. 这给我们提供了一种具体的操作流程. 我们可以对下面这个矩阵进行初等合同变换

$$\begin{pmatrix} A \\ I_n \end{pmatrix}$$

当这个矩阵的上半部分变成对角矩阵D的时候, 我们记下半部分为C. 那么就有

$$C'AC = D$$

为标准型.

别忘了举个例子.

定理 7.1.2. *假设$A=(a_{ij})$是一个反对称矩阵, 则存在一个$f\in F$, 使得$|A|=f^2$.*

证明 显然只要与A合同的矩阵的行列式是域F中的元素的平方, 那么A的行列式一定也是域F中的元素的平方.

首先由于奇数阶的反对称矩阵行列式为零, 所以我们假设n是偶数. 我们用数学归纳法, 当矩阵是2阶时, 结论当然是成立的. 假设命题对$n-2$阶的矩阵成立. 我们可以通过初等合同变换, 将矩阵的前两行和前两列交叉得到的2×2子矩阵变成

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

的样子. 然后再把第二行乘上适当的元素往下加, 可以把第一列中下面的各行都变为零, 然后做同样的列变换把第一行中后面的各列都变为0.

再把第一行乘上适当的元素往下加, 可以把第二列中下面的各行都变为零, 然后做同样的列变换把第二行中后面的各列都变为0. 原来的矩阵A经过这样的合同变换之后就变成了

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$$

的形式, 命题得证. □

推论 7.1.3. 假设$A = (a_{ij})$是一个反对称整数矩阵, 则存在一个$f \in \mathbb{Z}$, 使得$|A| = f^2$.

证明 由上面的定理可知, A的行列式是一个有理数的平方. 但是它同时又是一个整数矩阵, 所以它的行列式是一个整数. 由算术基本定理我们知道一个整数如果是有理数的平方, 一定是整数的平方. □

实际上假设$A = (a_{ij})$是一个反对称矩阵, 那么存在一个系数是整数的以a_{ij} $(i < j)$为变量的多元多项式$Pf(A)$, 使得

$$|A| = Pf(A)^2.$$

$Pf(A)$叫做是A的Pfaffian. 如果A是奇数阶的, 那么$Pf(A) = 0$. 如果

$$A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix},$$

那么$Pf(A) = a$. 如果

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

则$Pf(A) = af - be + dc$.

§7.2 规范型

定义 7.2.1 (二次型的秩)**.** 二次型对应矩阵的秩称为二次型的秩.

定理 7.2.1. 假设$F = \mathbb{C}$, f是F上的一个秩为r的二次型, 那么一定存在一个非退化的线性替换, 把f变为

$$f = y_1^2 + \cdots + y_r^2.$$

这个标准型称为是原来二次型的规范型. 规范型是由f 唯一确定的, 与采用什么非退化的线性替换没有关系.

证明 我们在上一节已经看到了f等价于一个对角型. 所以我们不妨设

$$f = d_1x_1^2 + \cdots + d_rx_r,$$

上式中每一个系数都不等于0. 注意我们在这里已经用到了f的秩等于0的条件. 然后我们令

$$y_1 = \sqrt{d_1}x_1,\ ...,\ y_r = \sqrt{d_r}x_r$$

即可. □

上面的证明中很关键的地方是对每一个系数d_i都可以开平方. 这个操作在复数域生当然没有问题, 但是在实数域上就不行了.

定理 7.2.2. *假设$F = \mathbb{R}$, f是F上的一个秩为r的n元二次型, 那么一定存在一个非退化的线性替换, 把f变为*

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1} - \cdots - y_r^2.$$

这个标准型称为是原来二次型的规范型. 其中的p称为是正惯性指数, 而$q = r - p$称为是负惯性指数. 这里的惯性是指在变换中保持不变的东西. 这个规范型是由f唯一确定的, 与采用什么非退化的线性替换没有关系.

证明 任何一个二次型都可以用非退化的线性替换变成规范型是很容易看出来的, 模仿复数域情形的证明即可, 唯一不同的是要注意只有正数才能开平方. 我们只需要证明正惯性指数的唯一性, 由于秩是唯一确定的, 所以只要正惯性指数由原来的二次型唯一确定, 负惯性指数也必须由原来的二次型唯一确定.

我们用反证法, 假设二次型f即可以通过非退化的线性替换变成

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1} - \cdots - y_r^2,$$

也可以经过非退化的线性替换变成

$$f = z_1^2 + \cdots + z_s^2 - z_{s+1} - \cdots - z_r^2.$$

等价. 而$p < s$.

我们知道所有的y_i, z_i都是x_i的线性组合. 因此我们考虑下面的线性方程组

$$\begin{cases} y_1 & = 0, \\ \vdots & , \\ y_p & = 0 \\ z_{s+1} & = 0, \\ \vdots & , \\ z_n & = 0 \end{cases}$$

上面这个关于x_1, ..., x_n的方程组里面只有$n-s+p$个方程. 根据齐次线性方程组的解的结构定理, 这个方程组有非零解

$$x_1 = a_1,\ ...x_n = a_n.$$

把这个非零解代回原来的二次型. 如果我们用

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1} - \cdots - y_r^2,$$

这个表达式, 会得到

$$f \leq 0.$$

如果我们用

$$f = z_1^2 + \cdots + z_s^2 - z_{s+1} - \cdots - z_r^2.$$

这个表达式, 会得到$f > 0$. 这就引起了矛盾. 为什么这里f一定会是正的呢? 因为如果是0的话, 就会推出

$$z_1 = \cdots = z_s = 0,$$

而我们已经有

$$z_{s+1} = 0,\ ...,\ z_n = 0$$

了. 这必然会导致非零向量$(a_1,\ ...,\ a_n)$在非退化线性替换下变成了零向量这个不可能的结论. □

我们刚才看到了非退化的线性替换会保持正惯性指数不变, 也会保持负惯性指数不变, 这正是它们被称作"惯性"的原因. 因为它们总是倾向于保持不变.

如果我们采用的是退化的线性替换, 那么正负惯性指数是有可能会变化的. 不过不管是正惯性指数, 还是负惯性指数, 都不会变大, 如果有改变, 一定是变小了. 这个可以用和上面类似的方法证出来, 读者可以作为练习. 退化的线性替换由于是不可逆的, 会丢失二次型的信息. 所以在任何情况下都应该避免.

§7.3 正定二次型

我们在本节讨论中, 总是假定数域F就是实数域$\mathbb{R}$. 假设f是n个变量的二次型. 那么f可以看成是n维的实线性空间上的函数, 它把每一个向量都变成一个实数. 如果这个函数的取值总是大于等于0. 我们就称这个二次型是半正定的. 如果半正定二次型f取值等于零当且仅当每个变量都取零, 我们就称这个二次型是正定二次型. 类似的我们可以定义半负定二次型和负定二次型. 如果二次型既可以取正值, 也可以取负值, 我们就称之为不定的二次型. 如果f是正定的二次型, f对应的矩阵A 就

称为是正定矩阵. 类似的, 我们可以定义半正定矩阵, 负定矩阵, 半负定矩阵, 不定矩阵等等.

如果f在P点取值为正, 在Q点取值为负, 我们知道连接P, Q的直线ℓ一定不会经过原点. 由实数的连续性可知, f一定在ℓ上的某个点处取值为零, 这个点不是原点.

A的行列式中前k行和前k列组成的k阶子式叫做A 的顺序主子式.

定理 7.3.1. *假设$f(x_1, ..., x_n) = X'AX$是一个二次型, 那么下面几条等价:*

1. *f是正定二次型,*

2. *f的正惯性指数为n,*

3. *A合同于单位矩阵,*

4. *存在可逆矩阵C, 使得$A = C'C$,*

5. *A的顺序主子式都是正的,*

6. *A的所有的主子式都是正的.*

证明

(1) $\Longrightarrow$ (2). 我们用反证法. 如果f的正惯性指数为$p < n$, 那么一定存在一个非退化的线性替换把f化成规范型

$$y_1^2 + \cdots + y_p^2 + d_{p+1}y_{p+1}^2 + \cdots + d_n y_n^2,$$

其中d_{p+1}, ..., d_n是-1或者是0. 所以我们可以解线性方程组

$$y_{p+1} = \cdots = y_n = 1, \ y_1 = \cdots = y_p = 0.$$

这个方程组一定有非零解. 而f在这个非零解处的取值是小于等于零的. 这就和f是正定二次型矛盾了.

(2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) $\Longrightarrow$ (1)很容易. 所以前四条是等价的.

(1) $\Longrightarrow$ (6). 我们假设命题对变量个数小于n的情形成立. 即如果f是个关于k $(k < n)$个变量的正定二次型, 那么f所对应的k阶矩阵的所有的主子式都是正的. 假设B是A的第i_1, ..., i_t行, 第i_1, ..., i_t列交叉得到的t 阶子矩阵. 我们令$x_{i_1} = y_1$, ..., $x_{i_t} = y_t$,, 其余的变量x_j都取0, 可以得到一个关于t个变量的二次型,

$$g(y_1, ..., y_t) = Y'BY, \ Y = (y_1, ..., y_t)'.$$

如果$t < n$, 那么由归纳假设就可以知道B的行列式是正的. 所以只需要证明$t = n$的情形. 这时, $B = A$, 相当于要证明A 的行列式是正数. 由(4)直接可以看出A的行列式是正数.

(6) $\Longrightarrow$ (5) 显然.

(5) $\Longrightarrow$ (1) . 如果A的顺序主子式都是正的, 那么其中前$n-1$个是正的, 所以如果把A写成分块矩阵的形式

$$\begin{pmatrix} B & Y \\ Y' & a_{nn} \end{pmatrix}$$

那么左上角的$n-1$阶矩阵B是正定的, 当然也是可逆的. 所以我们可以通过第三种初等合同变换把它变为

$$\begin{pmatrix} B & 0 \\ 0 & a_{nn} - Y'B^{-1}Y \end{pmatrix}$$

不改变行列式, 所以$a_{nn} - Y'B^{-1}Y$是个正数. 这是一个分块的对角矩阵, 每一个对角块都是正定的, 整体当然是正定的.

□

对称矩阵A是负定的, 当且仅当$-A$正定, 当且仅当A的奇数阶顺序主子式是负的, 偶数阶顺序主子式是正的.

定理 7.3.2. *假设$f(x_1, \ ..., \ x_n) = X'AX$是一个二次型, 那么下面几条等价:*

1. *f是半正定二次型,*
2. *f的负惯性指数为0,*
3. *A合同于*
$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$
4. *存在矩阵C (不一定可逆), 使得$A = C'C$,*
5. *A的所有的主子式都是非负的.*

证明 前4条等价是显然的. 我们只需要证明和第5条等价.

假设命题对阶数低于n的矩阵是成立的.

如果A是半正定的, 那么它对任意$n-1$个变量都是半正定的. 所以由归纳假设, 任意阶数小于等于$n-1$ 的主子式都是非负的. 所以只需要证明A的行列式非负. 我们用反证法. 如果A的行列式是负的,我们看A的规范型, 肯定对角线上会有-1. 因此二次型肯定能取到负值. 所以A不可能是半正定的.

反之如果A的所有的主子式都是非负的, 我们要证明A是半正定的. 我们考虑多项式

$$|\lambda I_n + A| = \lambda^n + s_1\lambda^{n-1} + \cdots + s_n.$$

这个等式的证明可以利用行列式的拆行, 将等式的左边每一行都拆开, 写成2^n个行列式之和, 其中的每个行列式都是λ的某个幂次乘上一个主子式. 合并同类项以后就可以得到右边的表达形式.

其中s_k是A的所有的k阶子式的和. 所以每一个s_k 都是非负的. 所以对任意的正数λ, 上式都是正数. 同样的道理, $\lambda I_n + A$的所有的顺序主子式都是正的, 所以$\lambda I_n + A$是正定的.

由于对任意小的正数λ, 都有$\lambda I_n + A$是正定的. 所以如果存在非零向量Y, 使得$Y'AY < 0$, 那么$Y'(\lambda I_n + A)Y < 0$对足够小的λ成立. 这与$\lambda I_n + A$是正定的矛盾. 所以不存在非零向量Y, 使得$Y'AY < 0$, 所以A是半正定的.

□

1. 证明秩为r的对称矩阵可以写成r个秩为1的对称矩阵之和.

    证明 我们知道任何一个二次型都在复数域上合同于规范型, 即任何一个对称矩阵都可以写成

$$A = C'\begin{pmatrix} I_r & 0 \\ 0 & o \end{pmatrix} = C'D_1C + \cdots + C'D_rC,$$

    其中D_i是一个n阶的对角矩阵, 它的第$(i,\ i)$位置是1, 其余的位置都是0. 每一个$C'D_1C$ 的秩都是1.

□

2. 一个实二次型可以分解成两个实系数的一次齐次多项式的乘积的充分必要条件是, 它的秩等于2和符号差等于0, 或者秩等于1.

    证明 充分性很显然. 我们只需要证明必要性.

    假设二次型

$$f = X'AX = gh,$$

    $g,\ h$都是一次型,

$$(g,\ h) = (x_1,\ ...,\ x_n)B,$$

    B是一个$n\times 2$的矩阵. 由于B的列数是2, 而且不是零矩阵, 所以它的秩只有0或者1两种可能. 如果它的秩是2, 那么我们可以通过在B的后面添加$n-2$列将B扩充成一个可逆矩阵C. 我们可以考虑非退化的线性替换

$$Y' = X'C,$$

这样的话, 原来的二次型在这个非退化的线性替换之下就变成

$$f = y_1y_2 = (u+v)(u-v) = u^2 - v^2.$$

它的秩等于2和符号差等于0. □

3. 判断二次型

$$f = \sum_{i=1}^{n} x_i^2 + \sum_{1\le i<j\le n} x_ix_j$$

是否是正定的.

证明 是正定的. 因为对应的对称矩阵是

$$\frac{1}{2}\begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix} = \frac{1}{2}(I_n + \alpha\alpha'),$$

这里的$\alpha' = (1,\ 1,\ ...,\ 1)$. 我们知道$I_n+\alpha\alpha'$的行列式等于$\det(1+\alpha'\alpha) = n+1$是个正数. 而它的各阶顺序主子式都是这个样子的, 同样的方法可以求出其行列式都是正数. 所以它是正定的. □

4. 判断二次型

$$f = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_ix_{i+1}$$

是否是正定的.

证明 这个相当于判断一个三对角矩阵的行列式是不是正的. 因为二次型对应的各级顺序主子式都是一个标准的三对角行列式. 我们在上个学期已经学过了三对角行列式的计算方法, 拿过来套用即可知道题目中的三对角行列式是正的, 所以对应的矩阵是正定的. □

5. 假设A是一个n阶的实对称矩阵, 而且行列式小于零, 则一定存在实的n维向量$X \neq 0$ 使得

$$X'AX < 0.$$

证明 证明A合同于规范型, 对角线上有-1, 合理选取X即可.

□

6. 证明二次型

$$f = n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2$$

是半正定的.

证明

我们在中学就学过这样的柯西不等式

$$(a_1b_1+\cdots+a_nb_n)^2\leq(a_1^2+\cdots+a_n^2)(a_1^2+\cdots+a_n^2)$$

让$a_i=x_i,\ b_i=1$即可. 当x_i都相等的时候, 二次型为零, 所以只是半正定, 不是正定的.

其实$f=\sum(x_i-x_j)^2$.

□

7. 假设实二次型

$$f(x_1,\ ...,\ x_n)=\sum_{i=1}^{s}(a_{i1}x_1+\cdots+a_{in}x_n)^2,$$

证明f的秩等于矩阵$A=(a_{ij})$的秩.

证明

$$f=X'A'AX,$$

所以f对应的矩阵是$A'A$, 它的秩等于A的秩. □

8. 假设实二次型

$$f(x_1,\ ...,\ x_n)=l_1^2+\cdots+l_p^2-l_{p+1}^2-\cdots-l_{p+q}^2,$$

其中l_i是$x_1,\ ...,\ x_n$的一次齐次式. 证明f的正惯性指数$\leq p$, 负惯性指数$\leq q$.

证明

我们假设二次型f可以通过非退化的线性替换变成

$$f=y_1^2+\cdots+y_s^2-y_{s+1}^2-\cdots-y_{s+t}^2,$$

而$s>p$.

我们知道所有的$l_i,\ y_i$都是x_i的线性组合. 因此我们考虑下面的线性方程组

$$\begin{cases} l_1 & =0,\\ \vdots & ,\\ l_p & =0\\ y_{s+1} & =0,\\ \vdots & ,\\ y_n & =0 \end{cases}$$

上面这个关于x_1, ..., x_n的方程组里面只有$n-s+p$个方程. 根据齐次线性方程组的解的结构定理, 这个方程组有非零解

$$x_1 = a_1,\ ...,\ x_n = a_n.$$

把这个非零解代回原来的二次型. 如果我们用

$$f = l_1^2 + \cdots + l_p^2 - l_{p+1}^2 - \cdots - l_{p+q}^2,$$

这个表达式, 会得到

$$f \leq 0.$$

如果我们用

$$f = y_1^2 + \cdots + y_s^2 - y_{s+1} - \cdots - y_{s+t}^2$$

这个表达式, 会得到$f > 0$. 这就引起了矛盾. 为什么这里f一定会是正的呢? 因为如果是0的话, 就会推出

$$y_1 = \cdots = y_s = 0,$$

而我们已经有

$$y_{s+1} = 0,\ ...,\ y_n = 0$$

了. 这必然会导致非零向量$(a_1,\ ...,\ a_n)$在非退化线性替换下变成了零向量这个不可能的结论. 所以这就证明了正惯性指数$s \leq p$.

同理可以证明负惯性指数$\leq q$.

□

9. (1) 如果$f(x_1,\ ...,\ x_n) = X'AX$是正定二次型, 那么

$$f(y_1,\ ...,\ y_n) = \begin{vmatrix} A & Y \\ Y' & 0 \end{vmatrix}$$

是负定的, 这里的$X = (x_1,\ ...,\ x_n)'$, $Y = (y_1,\ ...,\ y_n)'$.

证明 因为

$$\begin{pmatrix} I_n & 0 \\ -Y'A^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & Y \\ Y' & 0 \end{pmatrix} \begin{pmatrix} I_n & -A^{-1}Y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & -Y'A^{-1}Y \end{pmatrix}$$

是合同变换

所以二次型

$$f(y_1,\ \dots,\ y_n)$$

等价于二次型

$$-|A|Y'A^{-1}Y.$$

因为$Y'A^{-1}Y$是正定的, 所以$-|A|Y'A^{-1}Y$ 是负定的. 所以$f(y_1,\ \dots,\ y_n)$是负定的.

□

(2) 如果A是正定的, 那么

$$|A| \leq a_{nn}P_{n-1},$$

其中P_{n-1}是A的$n-1$级顺序主子式.

证明 假设

$$A = \begin{pmatrix} A_{n-1} & Y \\ Y' & a_{nn} \end{pmatrix},$$

则

$$\begin{pmatrix} I_n & 0 \\ -Y'A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & Y \\ Y' & a_{nn} \end{pmatrix} \begin{pmatrix} I_n & -A_{n-1}^{-1}Y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - Y'A_{n-1}^{-1}Y \end{pmatrix}.$$

注意到$Y'A_{n-1}^{-1}Y$是个非负数, 所以

$$|A| \leq a_{nn}P_{n-1}.$$

□

(3) 如果A是正定矩阵, 那么

$$|A| \leq a_{11}\cdots a_{nn}.$$

证明 归纳即可.

□

(4). 如果$T=(t_{ij})$是$n\times n$可逆矩阵, 那么

$$|T|^2 \leq \prod_{i=1}^{n}(t_{1i}^2 + \cdots + t_{ni}^2).$$

证明

把(3)中的A换成$T'T$即可.

□

10. 实对称阵A的一阶主子式之和与二阶主子式之和都为0, 求证$A=0$.

证明 假设$A=(a_{ij})$, 则

$$\sum_{i=1}^{n} a_{ii}=0, \sum_{i\neq j}(a_{ii}a_{jj}-a_{ij}^2)=0.$$

所以

$$0=-(\sum_{i=1}^{n} a_{ii})^2+2\sum_{i\neq j}(a_{ii}a_{jj}-a_{ij}^2)=-\sum_{i=1}^{n} a_{ii}^2-2\sum_{i\neq j} a_{ij}^2.$$

由于A的所有的元素都是实数, 所以上面的每一项都是0, 所以$A=0$. □

第八章　欧几里得空间

欧几里得是公元前3-4世纪的古希腊伟大的数学家, 是¡几何原本¿作者. 这本书的名字其实叫¡原本¿, 曾经在两千多年的时间里, 作为标准的数学教科书, 是数学史上的奇迹. 明朝的时候, 徐光启和利玛窦将这本书的前六卷翻译成中文的时候, 命名为¡几何原本¿.

§8.1　内积

定义 8.1.1 (欧几里得空间). 假设V是$\mathbb{R}$上的线性空间。假设f是从集合$V\times V$到$\mathbb{R}$的一个映射，而且满足下面的条件

$$\begin{gathered}
f(\alpha,\ \beta)=f(\beta,\ \alpha)\\
f(\alpha_1+\alpha_2,\ \beta)=f(\alpha_1,\ \beta)+f(\alpha_2,\ \beta)\\
f(k\alpha,\ \beta)=kf(\alpha,\ \beta),\ k\in\mathbb{R},\\
f(\alpha,\ \alpha)\geq 0,\ \ \text{等号成立当且仅当}\alpha=0,
\end{gathered}$$

那么我们称$f(\alpha,\ \beta)$为$\alpha,\ \beta$的内积，以后简单的记为$(\alpha,\ \beta)$. 我们称带有这种内积的空间V 为欧几里得空间。

以下都假设V是有限维的。

定义 8.1.2 (欧几里得空间的同构). 假设$V,\ W$是两个n维的欧几里得空间, $\mathscr{A}: V\longrightarrow W$是线性同构. 如果对任意的$\alpha_1,\ \alpha_2\in V$, 都有

$$(\alpha_1,\ \alpha_2)=(\mathscr{A}(\alpha_1),\ \mathscr{A}(\alpha_2)).$$

我们就称$\mathscr{A}$是从欧几里得空间V到另一个欧几里得空间W的同构. 即欧几里得空间之间的同构不单要是线性同构, 还要是保持内积不变的.

例 8.1.1. 假设$V=\mathbb{R}^n$是实数域$\mathbb{R}$上的n维线性空间. 假设$\alpha=(x_1,\ ...,\ x_n)',\ \beta=(y_1,\ ...,\ y_n)'\in V$, 通常的内积定义为

$$(\alpha,\ \beta)=\alpha'\beta=x_1y_1+\cdots+x_ny_n.$$

带有这个内积的V成为一个欧氏空间. 对向量$\alpha\in V$, 我们称

$$|\alpha|:=\sqrt{(\alpha,\ \alpha)}$$

为向量α的长度. 假设$\alpha,\ \beta\in V$, 它们的夹角为θ, 则

$$\cos\theta=\frac{(\alpha,\ \beta)}{|\alpha||\beta|},$$

或者说

$$\theta = \arccos\frac{(\alpha,\ \beta)}{|\alpha||\beta|},\ 0 \le \theta \le \pi.$$

定理 8.1.1 (Cauchy-Bunyakovsky-Schwarz 不等式)**.** *假设α, β是n-维欧氏空间V中的两个非零向量, 则*

$$(\alpha,\ \beta)^2 \le (\alpha,\ \alpha)(\beta,\ \beta),$$

等号成立当且仅当α, β线性相关.

证明 假设

$$f(t) = (\alpha + t\beta,\ \alpha + t\beta) = (\alpha,\ \alpha)t^2 + 2(\alpha,\ \beta)t + (\beta,\ \beta).$$

这是一个首项系数为正数的一元二次方程, 取值永远大于等于零, 而且$f(t) = 0$当且仅当$\alpha + t\beta = 0$. 因此$f(t)$的判别式

$$\Delta = 4(\alpha,\ \beta)^2 - 4(\alpha,\ \alpha)(\beta,\ \beta) \le 0,$$

等号成立当且仅当$\alpha + t\beta = 0$有解. 所以我们有

$$(\alpha,\ \beta)^2 \le (\alpha,\ \alpha)(\beta,\ \beta),$$

等号成立当且仅当$\alpha + t\beta = 0$有解. 而存在$t \in \mathbb{R}$使得$\alpha + t\beta = 0$当且仅当α, β线性相关. □

由于对是n-维欧氏空间V中的任意两个非零向量α, β, 都有

$$(\alpha + \beta,\ \alpha + \beta) = (\alpha,\ \alpha) + (\beta,\ \beta) + 2(\alpha,\ \beta),$$

所以Cauchy-Bunyakovsky-Schwarz 不等式等价于

$$(\alpha + \beta,\ \alpha + \beta) \le (\alpha,\ \alpha) + (\beta,\ \beta) + 2|\alpha| \cdot |\beta|,$$

两边取平方根, 上面的式子变成

$$|\alpha + \beta| \le |\alpha| + |\beta|.$$

这就是我们中学熟悉的三角不等式. 所以说Cauchy-Bunyakovsky-Schwarz 不等式就是三角不等式.

定义 8.1.3 (Gram矩阵, 度量矩阵)**.** *假设V是n-维欧氏空间, α_1, ..., α_n 是V的一组基. 则V 在这组基下的Gram 矩阵(或者度量矩阵) 定义为*

$$A = \begin{pmatrix} (\alpha_1,\ \alpha_1) & (\alpha_1,\ \alpha_2) & \cdots & (\alpha_1,\ \alpha_n) \\ (\alpha_2,\ \alpha_1) & (\alpha_2,\ \alpha_2) & \cdots & (\alpha_2,\ \alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_n,\ \alpha_1) & (\alpha_n,\ \alpha_2) & \cdots & (\alpha_n,\ \alpha_n) \end{pmatrix}$$

命题 8.1.2. 假设V是n-维欧氏空间, $\alpha_1,\ ...,\ \alpha_n$ 是V的一组基. 则V在这组基下的$Gram$矩阵A 是一个正定矩阵. 假设V在另一组基$\beta_1,\ ...,\ \beta_n$之下的矩阵为B, 从$\alpha_1,\ ...,\ \alpha_n$到$\beta_1,\ ...,\ \beta_n$ 的过渡矩阵为C, 则

$$C'AC = B.$$

证明

任取$\gamma = c_1\alpha_1 + \cdots + c_n\alpha_n \in V$. 令$X = (c_1,\ ...,\ c_n)'$, 则

$$(\gamma,\ \gamma) = X'AX.$$

由欧氏空间内积的定义可知, 对任意非零的γ, 都有$(\gamma,\ \gamma) > 0$, 所以对任意非零的X, 都有$X'AX > 0$. 所以A是正定的.

假设$\beta_1,\ ...,\ \beta_n$ 是V的另一组基. 则存在一个过渡矩阵C使得$\alpha_1,\ ...,\ \alpha_n$化为$\beta_1,\ ...,\ \beta_n$. 即

$$(\alpha_1,\ ...,\ \alpha_n)C = (\beta_1,\ ...,\ \beta_n).$$

因此如果记V在$\alpha_1,\ ...,\ \alpha_n$之下的Gram矩阵为

$$A = \begin{pmatrix} (\alpha_1,\ \alpha_1) & (\alpha_1,\ \alpha_2) & \cdots & (\alpha_1,\ \alpha_n) \\ (\alpha_2,\ \alpha_1) & (\alpha_2,\ \alpha_2) & \cdots & (\alpha_2,\ \alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_n,\ \alpha_1) & (\alpha_n,\ \alpha_2) & \cdots & (\alpha_n,\ \alpha_n) \end{pmatrix},$$

V在$\beta_1,\ ...,\ \beta_n$之下的Gram矩阵为

$$B = \begin{pmatrix} (\beta_1,\ \beta_1) & (\beta_1,\ \beta_2) & \cdots & (\beta_1,\ \beta_n) \\ (\beta_2,\ \beta_1) & (\beta_2,\ \beta_2) & \cdots & (\beta_2,\ \beta_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\beta_n,\ \beta_1) & (\beta_n,\ \beta_2) & \cdots & (\beta_n,\ \beta_n) \end{pmatrix}$$

则

$$C'AC = B.$$

□

定义 8.1.4 (正交, 垂直)**.** 假设V是一个欧几里得空间. 如果$(\alpha,\ \beta) = 0$, 我们就称α, β互相正交或者说它们互相垂直. 对V的任一子空间W, 我们定义W的正交补$W^{\perp} = \{x \in V | (x,\ W) = 0\}$. 假设$V_1$, V_2都是V的子空间, 如果对任意的$\alpha_1 \in V_1$, $\alpha_2 \in V_2$, 都有

$$(\alpha_1,\ \alpha_2) = 0,$$

我们就称V_1与V_2互相正交, 或者说V_1, V_2是互相垂直的.

命题 8.1.3. 假设V是一个欧几里得空间, V_1, V_2都是V的两个互相正交的子空间, 则V_1, V_2的和为直和.

证明 任取$\alpha \in V_1 \cap V_2$. 因为V_1, V_2互相垂直, 所以

$$(\alpha,\ \alpha) = 0.$$

由欧几里得空间的性质可知$\alpha = 0$. 所以V_1, V_2的和为直和.

□

定义 8.1.5 (标准正交基)**.** 假设e_1, ..., e_n 是n-维欧氏空间的一组基, 而且满足

$$(e_i,\ e_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases}$$

我们就称e_1, ..., e_n 是n-维欧氏空间的一组标准正交基.

由命题8.1.2可知, n-维欧氏空间V在任何一组基α_1, ..., α_n下的矩阵都是正定矩阵. 我们知道正定矩阵都合同于单位矩阵, 即存在可逆矩阵C使得

$$CAC' = I.$$

假设α_1, ..., α_n经过C过渡到了另外一组基β_1, ..., β_n, 则V在β_1, ..., β_n 下的矩阵是单位矩阵. 所以β_1, ..., β_n是V的一组标准正交基.

定理 8.1.4. 假设α_1, ..., α_n是n-维欧氏空间V的一组基, 则一定存在V的一组标准正交基β_1, ..., β_n使得对任意的$1 \leq m \leq n$都有

$$L(\alpha_1,\ ...,\ \alpha_m) = L(\beta_1,\ ...,\ \beta_m).$$

证明 显然我们只需要找到一组符合定理中要求的正交基, 然后把每个基向量都除以它的长度就可以得到一组标准正交基.

首先令$\beta_1 = \alpha_1$, 然后令

$$\begin{aligned}
\beta_2 &= \alpha_2 - \frac{(\alpha_2,\ \beta_1)}{(\beta_1,\ \beta_1)}\beta_1 \\
\beta_3 &= \alpha_3 - \frac{(\alpha_3,\ \beta_1)}{(\beta_1,\ \beta_1)}\beta_1 - \frac{(\alpha_3,\ \beta_2)}{(\beta_2,\ \beta_2)}\beta_2 \\
&\cdots \\
\beta_m &= \alpha_m - \frac{(\alpha_m,\ \beta_1)}{(\beta_1,\ \beta_1)}\beta_1 - \cdots - \frac{(\alpha_m,\ \beta_{m-1})}{(\beta_{m-1},\ \beta_{m-1})}\beta_{m-1} \\
&\cdots \\
\beta_n &= \alpha_n - \frac{(\alpha_n,\ \beta_1)}{(\beta_1,\ \beta_1)}\beta_1 - \cdots - \frac{(\alpha_n,\ \beta_{n-1})}{(\beta_{n-1},\ \beta_{n-1})}\beta_{n-1}
\end{aligned}$$

容易验证, 如此构造的是一组符合要求的正交基, 然后再将其除以长度, 就可以使其变成标准正交基. □上面的定理中的把一组基化为标准正交基的过程叫做施密特标准正交化过程.

定理 8.1.5. *假设V是一个欧几里得空间, W是V的子空间, 则$V = W \bigoplus W^{\perp}$.*

证明 我们前面已经看到$W \bigcap W^{\perp} = \{0\}$, 所以二者之和为直和. 所以为了证明$V = W \bigoplus W^{\perp}$ 只需要证明W 和$W^{\perp}$能够张成V, 即$V = W + W^{\perp}$. 任取$\alpha \in V$, 任取W的一组标准正交基$\alpha_1,\ \alpha_2,\ ...,\ \alpha_k$, 则我们令实数$x_1 = (\alpha,\ \alpha_1)$, $x_2 = (\alpha,\ \alpha_2)$, ..., $x_k = (\alpha,\ \alpha_k)$. 则

$$(\alpha - x_1\alpha_1 - x_2\alpha_2 - \cdots - x_k\alpha_k,\ \alpha_i) = 0,$$

对每个$1 \le i \le k$成立. 即如果记

$$\beta = \alpha - x_1\alpha_1 - x_2\alpha_2 - \cdots - x_k\alpha_k,$$

则

$$\alpha = (\alpha - \beta) + \beta,$$

其中$\beta \in W^{\perp}$ 且$\alpha - \beta \in W$. 所以W 和$W^{\perp}$能够张成V.

□

§8.2 正交矩阵和正交变换

定义 8.2.1. *假设O是实数域上的n阶方阵, 如果*

$$O'O = OO' = I_n,$$

那么我们就称O是正交矩阵. 如果$\mathscr{A}$是欧氏空间$\mathbb{R}^n$上的一个线性变换, 而且保持向量的内积不变, 即对任意的向量$\alpha,\ \beta$都有

$$(\alpha,\ \beta) = (\mathscr{A}(\alpha),\ \mathscr{A}(\beta)),$$

我们就称$\mathscr{A}$是正交变换.

注意实际上由$O'O = I_n$, 即可得到$O' = O^{-1}$, 所以$OO' = I_n$, 即O是正交矩阵.

由正交变换和标准正交基的定义可知, n-维欧氏空间的正交变换把标准正交基仍然变为标准正交基.

命题 8.2.1. *1. 正交变换在任何一组标准正交基下的矩阵都是正交阵.*

2. 如果欧氏空间的线性变换在某组标准正交基下的矩阵是正交阵, 那么该变换一定是正交变换.

3. 从标准正交基到标准正交基的过渡矩阵是正交阵.

4. 正交阵的列向量构成了一组标准正交基, 行向量也是.

证明

1. 假设V是n-维欧氏空间, $\mathscr{A}$是V上的一个线性变换, $e_1,\ ...,\ e_n$ 是V的一组标准正交基, $A=(a_{ij})$ 是$\mathscr{A}$在这组标准正交基下的矩阵, 则

$$\mathscr{A}(e_1,\ ...,\ e_n)=(e_1,\ ...,\ e_n)A.$$

所以

$$\mathscr{A}(e_j)=\sum_{i=1}^{n}a_{ij}e_i,$$

因此

$$\begin{aligned}(\mathscr{A}(e_i),\ \mathscr{A}(e_j))&=\sum_{k,\ l=1}^{n}a_{ki}a_{lj}(e_k,\ e_l)\\&=\sum_{k,\ l=1}^{n}a_{ki}a_{lj}\delta_{kl}\\&=\sum_{k=1}^{n}a_{ki}a_{kj}.\end{aligned}$$

假设$A'A=(c_{ij})$, 则

$$c_{ij}=\sum_{k=1}^{n}a_{ki}a_{kj}.$$

2. 所以$\mathscr{A}$是正交变换当且仅当$C=I$, 即$\mathscr{A}$在标准正交基下的矩阵是正交矩阵. 把上面的推理过程逆回去可知, 如果欧氏空间的线性变换在某组标准正交基下的矩阵是正交阵, 那么该变换一定是正交变换.

3, 4 两条容易证明, 留做练习.

□

命题 8.2.2. *假设O是2阶正交矩阵. 如果$|O|=1$, 则一定存在$0\leq\theta<2\pi$使得*

$$O=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

如果$|O|=-1$, 则一定存在$0\leq\theta<2\pi$使得

$$O=\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}.$$

证明

假设

$$O=\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

则由$OO^{'}=I_2$可得

$$a^2+b^2=1$$

$$c^2+d^2=1$$

$$ac+bd=0$$

因为$a^2+b^2=1$, 所以一定存在α使得$a=\cos\alpha,\ b=\sin\alpha,\ 0\leq\alpha<2\pi$. 同样的, 一定存在$\beta$使得$c=\cos\beta,\ d=\sin\beta,\ 0\leq\beta<2\pi$. 因为$ac+bd=0$, 所以

$$\cos\alpha\cos\beta+\sin\alpha\sin\beta=\cos(\alpha-\beta)=0.$$

所以$\alpha-\beta=\pm\pi/2$ 或者$\alpha-\beta=pm3\pi/2$.

如果$\alpha-\beta=-\pi/2$ 或者$3\pi/2$, 则

$$O=\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

这时候$|O|=1$.

如果$\alpha-\beta=\pi/2$或者$-3\pi/2$, 则

$$O=\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

这时候$|O|=-1$.

□

定义 8.2.2 (正交相似). 假设A和B是两个实矩阵. 如果存在一个正交阵O使得

$$A=O'BO,$$

那么就说A正交相似于B. 容易看出, 正交相似是一个等价关系.

例 8.2.1 (正交相似). 假设

$$O=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

则O有两个特征值1, -1, 对应的特征向量分别是

$$\alpha=\begin{pmatrix} 1+\cos\theta \\ \sin\theta \end{pmatrix},\ \beta=\begin{pmatrix} 1-\cos\theta \\ -\sin\theta \end{pmatrix}.$$

容易看出这是两个互相正交的向量, 令

$$C=(\alpha/|\alpha|,\beta/|\beta|)=\begin{pmatrix}(1+\cos\theta)/\sqrt{2+2\cos\theta} & (1-\cos\theta)/\sqrt{2-2\cos\theta}\\ \sin\theta/\sqrt{2+2\cos\theta} & -\sin\theta/\sqrt{2-2\cos\theta}\end{pmatrix}$$

则C是一个正交矩阵. 而且

$$OC=C\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

所以O正交相似于

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

注意两个实矩阵相似, 并不一定正交相似.

命题 8.2.3. 假设$A=(a_{ij})$和$B=(b_{ij})$是两个实矩阵. 如果他们正交相似, 那么

$$\sum_{i,j=1}^{n}a_{ij}^2=\sum_{i,j=1}^{n}b_{ij}^2.$$

证明 如果存在正交矩阵T使得

$$T^{'}AT=B,$$

那么

$$Tr(BB^{'})=Tr(T^{'}AA^{'}T)=Tr(AA^{'}TT^{'})=Tr(AA^{'}).$$

所以命题成立. □

例 8.2.2 (相似但不正交相似)**.**

$$A=\begin{pmatrix}3&1\\-2&0\end{pmatrix},\ B=\begin{pmatrix}1&1\\0&2\end{pmatrix}$$

这个两个矩阵是相似的. 但是由上面的命题可知, 它们不是正交相似的.

从正交相似的定义可以看出, 两个矩阵如果是正交相似的, 那么它们既是相似的, 又是合同用的. 反之却不成立.

例 8.2.3 (相似且合同但不正交相似)**.**

$$A=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ B=\begin{pmatrix}0&2\\0&0\end{pmatrix}$$

这个两个矩阵是相似的, 因为它们具有相同的*Jordan*标准形; 它们是合同的, 因为

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

但是由上面的命题可知, 它们不是正交相似的.

命题 8.2.4. 假设O是n阶正交矩阵. 则O正交相似于准对角矩阵

$$diag(A_1,\ ...,\ A_k)$$

其中的每个A_i是± 1或者二阶的行列式为1的正交矩阵.

证明 我们先选择$\mathbb{R}^n$的一组标准正交基. 正交矩阵A在这组基下决定了一个正交变换$\mathscr{A}$.

首先说明实线性变换一定有一维或者二维的不变子空间W. 如果有实的特征值, 那自然有实的特征向量, 这就张成了一个一维的不变子空间. 如果没有实的特征值, 那肯定有两个互相共轭的复特征根

$$a+bi,\ a-bi,\ a,\ b\in\mathbb{R}.$$

所以也就有两个互相共轭的特征向量

$$\alpha+\beta i,\ \alpha-\beta i,$$

这里的$\alpha,\ \beta$是n维的实向量. 即

$$\mathscr{A}(\alpha+\beta i)=(a+bi)(\alpha+\beta i),\ \mathscr{A}(\alpha-\beta i)=(a-bi)(\alpha-\beta i).$$

所以

$$\mathscr{A}(\alpha)=a\alpha-b\beta,\ \mathscr{A}(\beta)=a\beta+b\alpha.$$

所以α,β 就张成了一个实数域上的二维的不变子空间W. 注意这里的α,β 在实数域上一定是线性无关的. 这是由于$\alpha+\beta i,\ \alpha-\beta i$是两个不同的特征值的特征向量（复数域上）, 所以它们在复数域上是线性无关的. 因此α,β 在实数域上一定是线性无关的.

由于$\mathscr{A}$是一个正交变换, 保持了所有的内积不变. 所以$\mathscr{A}$把和W正交的向量变成和$\mathscr{A}(W)=W$正交的向量. 因此W的正交补$W^{\perp}$也是$\mathscr{A}$的一个不变子空间, 但是维数更低. 对维数归纳即可. □

定义 8.2.3 (镜面反射)**.** 假设$\mathscr{A}$是n维欧氏空间V上的正交变换, 如果$\mathscr{A}$在某组标准正交基下的矩阵表示为

$$\begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix},$$

则称$\mathscr{A}$是是一个镜面反射.

命题 8.2.5. 假设$\mathscr{A}$是n维欧氏空间V上的正交变换, 则$\mathscr{A}$可以写成$n-k$个镜面反射的乘积, 其中k是$\mathscr{A}$的特征值*1*的重数*(*如果1不是$\mathscr{A}$的特征值, 那么$k=0$.*)*

证明 注意到二阶的行列式为1的正交矩阵是两个镜面反射的乘积. 而每个特征值-1都对应于一个镜面反射. 所以$\mathscr{A}$可以写成$n-k$个镜面反射的乘积, 其中k是$\mathscr{A}$的特征值1的重数(如果1不是$\mathscr{A}$的特征值, 那么$k=0$.

□

推论 8.2.6. *假设O是一个行列式为1的三级正交阵, 则一定正交相似于某个*

$$\begin{pmatrix} 1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta \end{pmatrix}$$

证明

首先因为O是一个行列式为1的三级正交阵, 所以一定有一个特征值是1, 假设α是它的一个长度为1的特征向量. 然后将α, 扩充为V 的一组标准正交基$\alpha,\ \beta,\ \gamma$. 记$W=L(\beta,\ \gamma)$, 则W也是$\mathscr{O}$的不变子空间, 所以$\mathscr{O}$在W上也是一个正交变换. 因此存在θ使得

$$\mathscr{O}(\alpha,\ \beta,\ \gamma)=(\alpha,\ \beta,\ \gamma)\begin{pmatrix} 1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta \end{pmatrix}.$$

□

注意上面的推论说明三维空间中的行列式为1 的正交变换$\mathscr{O}$是一个旋转, 是围绕着向量α, 转了θ 角度. 反之也成立, 即旋转是行列式为1 的正交变换. 由于两个行列式为1的三级正交阵的乘积仍是行列式为1的三级正交阵, 这说明了旋转的复合仍是旋转, 即假设$\mathscr{A}_1$是围绕着向量α_1, 转了θ_1角度, $\mathscr{A}_2$ 是围绕着向量α_2, 转了θ_2 角度. 则$\mathscr{A}_2\mathscr{A}_1$仍是一个旋转, 即一定存在$\alpha_3$, θ_3使得$\mathscr{A}_2\mathscr{A}_1$是围绕着向量$\alpha_3$, 转了$\theta_3$ 角度.

物理中描述旋转经常采用欧拉角. 我们下面来简单的描述一下. 假设三位空间的坐标轴xyz, 经过旋转之后变成了XYZ. 我们下面把这个旋转分解为三步完成, 每一步都对应着一个角度, 这三个角度合起来称为是该旋转的欧拉角$(\alpha,\ \beta,\ \gamma)$. 注意我们下面所说的旋转的角度都是按照右手螺旋的方向, 也就是右手的大拇指指向旋转轴的正向, 其余四指弯曲的方向就是旋转的方向.

我们假设xy平面和XY平面的交线为N, N的正向为$z\times Z$的方向.

(1) 保持z轴不动, 旋转xy平面, 使得x转到N的位置, 假设转的角度为α.

(2)保持N不动, 旋转zZ平面, 使得z转到N的位置, 假设转的角度为β. 这时候原来的xy平面被已经变到了和XY平面重合的位置了.

3)保持Z不动, 旋转xy平面, 使得xy转到XY的位置, 假设转的角度为γ.

任何一个旋转都可以分解为这三步完成, 所以任何一个旋转都可以用三个角度$(\alpha,\ \beta,\ \gamma)$来刻画. 这三个角度就是欧拉角.

我们也可以用四元数来刻画行列式为1的三级正交阵.

回忆任何一下四元数都可以写成

$$\alpha = d + ai + bj + ck$$

的形式，其中a, b, c, d都是实数。d叫做α的实部，$ai+bj+ck$ 叫做α的虚部。共轭定义为

$$\overline{\alpha} = d - ai - bj - ck.$$

假设$\alpha = d + ai + bj + ck$是一个不等于0的四元数, 则

$$\beta = \frac{1}{d^2+a^2+b^2+c^2}\overline{\alpha} = \frac{1}{d^2+a^2+b^2+c^2}(d - ai - bj - ck)$$

是α的逆元. 四元数$\alpha = d + ai + bj + ck$的长度定义为

$$|\alpha| = \sqrt{d^2+a^2+b^2+c^2}.$$

如果$d = 0$, 我们就称α为纯四元数. 纯四元数形成的空间我们记成V, 即

$$V = \{ai + bj + ck | a,\ b,\ c \in \mathbb{R}\},$$

这是实数域上的3维线性空间. 在这个实线性空间上我们可以定义内积

$$(\alpha,\ \beta) = (-\alpha\beta - \beta\alpha)/2.$$

这里的向量之间的乘法是前面定义的四元数的乘法. 因此, 对于$\alpha = ai + bj + ck$, $\beta = xi + yj + dk$, 我们定义α, β之间的内积为

$$(\alpha,\ \beta) = ax + by + cz.$$

容易看出这样定义的内积使得V成为了通常的三维欧氏空间.

引理 8.2.7. 任何一个四元数都可以写成$k(\cos\theta + \sin\theta\alpha)$, 其中$k$是一个实数, α是一个长度为1的纯虚的四元数。

证明 假设四元数

$$x = d + ai + bj + ck,$$

取$k = \sqrt{a^2+b^2+c^2+d^2}$为该四元数的长度, 取$\theta$, α 满足

$$\cos\theta = \frac{d}{k}$$
$$\sin\theta = \frac{\sqrt{a^2+b^2+c^2}}{k}$$
$$\alpha = \frac{ai+bj+ck}{\sqrt{a^2+b^2+c^2}}.$$

即可. □

引理 8.2.8. *假设非零$\delta \in \mathbb{H}$, 而且对所有的$x \in V$, 都有$\delta x \delta^{-1} = x$, 则$\delta \in \mathbb{R}$.*

证明 取$x = i$, 可知δ的j, k坐标为0. 取$x = j$, 可知δ的i坐标为0. 所以可知δ是个实数. □

假设$\alpha \in V$是一个向量，把任何一个向量都绕着α按照右手螺旋的方向旋转2θ角, 这是一个旋转变换.下面我们用四元数来刻画这个旋转.

首先可以假设$\alpha^2 = -1$, 即假设向量α的长度等于1, 否则给它乘上一个适当的系数就行了. 然后可以取一个和α垂直的长度为1 的向量β. 这时候$(\alpha,\ \beta) = 0$, 所以$\alpha\beta = -\beta\alpha$. 这时令$\gamma = \alpha\beta$. 则

$$\alpha\gamma = \alpha^2\beta = -\alpha\beta\alpha = -\gamma\alpha.$$

所以

$$(\alpha,\ \gamma) = 0.$$

同理可知

$$(\beta,\ \gamma) = 0.$$

而且通过计算长度可以知道γ的长度等于1. 所以α, β, γ是三个互相垂直的长度为1的向量, 构成了V的一个新的坐标系。令$\delta = \cos\theta + \alpha\sin\theta$, 则$\delta^{-1} = \cos\theta - \alpha\sin\theta$. 考虑线性变换

$$\begin{aligned} \mathscr{A}:\ V &\longrightarrow \quad V \\ x &\longmapsto \quad \delta x \delta^{-1}. \end{aligned}$$

我们有下面的公式。

$$\begin{aligned} (\cos\theta + \alpha\sin\theta)\alpha(\cos\theta - \alpha\sin\theta) &= \alpha, \\ (\cos\theta + \alpha\sin\theta)\beta(\cos\theta - \alpha\sin\theta) &= \cos 2\theta.\beta + \sin 2\theta.\gamma, \\ (\cos\theta + \alpha\sin\theta)\gamma(\cos\theta - \alpha\sin\theta) &= \cos 2\theta.\gamma - \sin 2\theta.\beta. \end{aligned}$$

由上面的公式我们可以看到$\mathscr{A}$是围绕α按照右手螺旋的方向旋转了2θ角度。

定理 8.2.9. *空间中的任何一个旋转相当于一个内自同构，即存在δ使得旋转变换等于下面的映射*

$$x \longmapsto \delta x \delta^{-1}.$$

证明 从我们前面的讨论可以看出旋转是个内自同构. 反之, 如果$\mathscr{A}$是个内自同构, 即存在长度为1的四元数$\delta = \cos\theta + \alpha\sin\theta$ 使得对任意的$\alpha \in V$都有

$$\mathscr{A}(\alpha) = \delta\alpha\delta^{-1},$$

则$\mathscr{A}$相当于是以α为旋转轴, 沿着右手螺旋法则旋转了2θ角. □

定理 8.2.10. 三维空间的旋转群同构于$\mathbb{H}^1/\{\pm 1\}$.

证明 用$\mathbb{H}^1$记长度为1的四元数的全体. 由前面的讨论可知$\mathbb{H}^1/\{\pm 1\}$中的不同元素诱导出不同的旋转。 □

这样的表示法有很大的优点。首先按照上面的公式，δ很容易求得。这样计算两个旋转的复合就很容易，只需要计算δ_1, δ_2两个四元数的乘积即可。要知道用通常的办法计算围绕α_1按照右手螺旋的方向旋转了$2\theta_1$角度, 然后再围绕α_2按照右手螺旋的方向旋转了$2\theta_2$角度之后的复合非常复杂。

群$\mathbb{H}^1/\{\pm 1\} \simeq \mathbb{P}^3_{\mathbb{R}}$相当于把3维球面$a^2+b^2+c^2+d^2=1$的对径点粘起来, 这就是3维的实射影空间.

定理 8.2.11. 一个实对称矩阵的特征值都是实数.

证明 假设A是一个实对称阵, $\lambda \in \mathbb{C}$是A的一个特征值, $\alpha \in \mathbb{C}^n$是A的一个特征向量. 那么我们有

$$A\alpha = \lambda\alpha \Longrightarrow \overline{\alpha'}A\alpha = \lambda\overline{\alpha'}\alpha,$$

两边再取共轭转置可得

$$\overline{\alpha'}A\alpha = \overline{\lambda\alpha'}\alpha.$$

所以

$$\lambda\overline{\alpha'}\alpha = \overline{\lambda\alpha'}\alpha.$$

所以$\lambda = \overline{\lambda}$, 即$\lambda$是实数. □

定理 8.2.12. 实对称矩阵正交相似于对角矩阵.

证明 可以对阶数进行归纳.

假设

$$A\alpha = \lambda\alpha,$$

其中α是单位长度的特征向量, λ是特征值. 那么将α扩成一组标准正交基,

$$\alpha_1 = \alpha,\ \alpha_2,\ \dots,\ \alpha_n.$$

把这n个列向量拼在一起, 就成了一个正交矩阵O. 我们有

$$AO = O\begin{pmatrix} \lambda & * \\ 0 & * \end{pmatrix} \Longrightarrow O'AO = \begin{pmatrix} \lambda & * \\ 0 & * \end{pmatrix}$$

由于A是对称矩阵, 所以右边这个带星号的矩阵也是对称的. 所以A正交相似于

$$\begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}$$

形状的对称阵, 归纳即可. □

定理 8.2.13. 一个实对称矩阵是正定的当且仅当它的特征值都是正数.

证明 由上面的定理可以知道实对称阵正交相似于一个对角矩阵. 所以实对称阵合同于对角线上都是特征值的对角阵. 这个对角阵是正定的当且仅当对角元都是正数. 所以实对称矩阵是正定的当且仅当它的特征值都是正数. □

定义 8.2.4 (对称变换)**.** 假设$\mathscr{A}$是欧氏空间V上的线性变换. 如果对任意的$\alpha, \beta \in V$, 都有

$$(\mathscr{A}(\alpha),\ \beta) = (\alpha,\ \mathscr{A}(\beta)),$$

那么就称$\mathscr{A}$是一个对称变换.

定理 8.2.14. 假设$\mathscr{A}$是n维欧氏空间V上的线性变化. 则它是对称变换当且仅当$\mathscr{A}$在标准正交基下的矩阵表示是对称矩阵. 注意由于同一个线性变换在不同基下的矩阵表示是正交相似的. 所以只要$\mathscr{A}$在某一组标准正交基之下的矩阵是对称阵, 那么它在任意的标准正交基之下的矩阵都是对称阵.

证明 假设$\alpha_1,\ ...,\ \alpha_n$是$n$维欧氏空间$V$的一组标准正交基, $\mathscr{A}$在这组基下的矩阵为$A = (a_{ij})$. 那么$\mathscr{A}$是一个对称变换当且仅当对所有的$1 \le i,\ j \le n$, 都有

$$(\mathscr{A}(\alpha_i),\ \alpha_j) = (\alpha_i,\ \mathscr{A}(\alpha_j)),$$

即

$$a_{ji} = a_{ij}.$$

所以$\mathscr{A}$是一个对称变换当且仅当A是对称矩阵.

□

注意对称变换在非标准正交基下的矩阵不一定是对称矩阵.

定理 8.2.15. $A,\ B$都是正定矩阵, 则AB的特征值都是正数. 如果AB是对称矩阵, 即$AB = BA$, 则AB是正定矩阵.

证明 假设$AB\alpha = \lambda\alpha$, 这里的$\lambda \in \mathbb{C}$, $\lambda \in \mathbb{C}^n$.

则

$$\overline{\alpha'}BAB\alpha = \lambda\overline{\alpha'}B\alpha.$$

由于

$$\overline{\alpha'}BAB\alpha = (\overline{\alpha' B'})A(B\alpha) > 0,\ \overline{\alpha}'B\alpha > 0,$$

所以$\lambda > 0$. 注意上行的不等式可以从正定矩阵都能写成$C'C = \overline{C'}C$的形式很容易的推出来.

□

命题 8.2.16. 假设A, B都是实对称矩阵, B是正定的, AB的特征是都是正数, 则A是正定的.

证明 因为B是正定的, 所以存在可你的实矩阵T使得$B = T'T$, 所以

$$AB = AT'T = T^{-1}(TAT')T.$$

因为AB的特征是都是正数, 所以TAT'的特征值都是正数, 所以TAT'是正定的, 所以A是正定的. □

定理 8.2.17. 假设A, B, C都是正定矩阵. 如果ABC是对称矩阵, 即$ABC = CBA$, 则ABC是正定矩阵.

证明 假设$A = D'D$, 则

$$ABC = D'DBC = D'(DBCD^{-1})D.$$

所以我们只需证明$DBCD^{-1}$是正定的,这等价于证明它的特征值都是正数, 也等价于证明BC的特征值都是正数. 由上面的定理可以知道命题成立.

□

定理 8.2.18. 假设A是一个半正定矩阵, 那么存在唯一的半正定矩阵B使得$B^2 = A$. 如果A是正定的, 那么这里的B 也是正定的.

证明 存在性容易证, 我们这里只证唯一性.

假设

$$A = B^2 = C^2,$$

不失一般性, 我们可以假设

$$B = U'diag(\lambda_1,\ ...,\ \lambda_n)U,\ C = V'diag(\lambda_1,\ ...,\ \lambda_n)V$$

U, V是正交阵. 这是因为B, C是具有相同特征值的半正定矩阵. 所以可以正交相似到对角矩阵, 然后再用同样也是正交矩阵的置换矩阵将对角线上的元素排列次序变得相同即可. 进一步, 我们可以把$diag(\lambda_1,\ ...,\ \lambda_n)$的对角线上元素相同的放在一起, 这样就变成了

$$\Lambda = diag(\lambda_1,\ ...,\ \lambda_1,\ ...,\ \lambda_k,\ ...,\ \lambda_k) = diag(\lambda_1 I_{a_1},\ ..,\ \lambda_1 I_{a_k})$$

$\lambda_1,\ ...,\ \lambda_k$两两不同. 所以

$$U'\Lambda^2 U = V'\Lambda^2 V, \Longrightarrow \Lambda^2 UV' = UV'\Lambda^2,$$

由两个矩阵可交换可以知道, UV'也是准对角的, 分块方式同Λ. 所以

$$\Longrightarrow \Lambda UV' = UV'\Lambda \Longrightarrow B = C.$$

$\square$

定理 8.2.19 (矩阵的极分解). *假设A是一个实矩阵, 则一定存在一个正交矩阵R和一个半正定矩阵P, 使得*

$$A = RP.$$

这里的P是由A唯一决定的. 如果A是可逆的, 那么R也是由A唯一决定的. 如果A是可逆的, 那么P是正定的.

证明 易知$A'A$是一个半正定矩阵. 由上面的定理可知, 存在唯一的半正定矩阵P使得$P^2 = A'A$. 我们假设

$$P = Odiag(\lambda_1,\ ...,\ \lambda_k,\ 0,\ ...,\ 0)O',$$

这里的O是某个适当的正交阵, $\lambda_1 \cdots \lambda_k \neq 0$. 记$B = AO$. 则

$$B'B = diag(\lambda_1,\ ...,\ \lambda_k,\ 0,\ ...,\ 0).$$

所以B的前k列是相互正交的, 线性无关的, B的后$n-k$列都是0. 将将B的前k列进行施密特标准正交化(由于这k 个向量本来就是互相正交的, 所以只要进行标准化把它们的长度都变成1 就行了), 并扩充称$\mathbb{R}^n$的一组标准正交基为

$$\alpha_1,\ ...,\ \alpha_k,\ \alpha_{k+1},\ ...,\ \alpha_n.$$

我们知道PO的前k列也是相互正交的, 线性无关的, PO的后$n-k$列都是0. 将PO的前k列进行施密特标准正交化(同样的, 由于这k个向量本来就是互相正交的, 所以只要进行标准化把它们的长度都变成1 就行了), 并扩充称$\mathbb{R}^n$的一组标准正交基为

$$\beta_1,\ ...,\ \beta_k,\ \beta_{k+1},\ ...,\ \beta_n.$$

则这两组标准正交基之间的过渡矩阵是一个正交阵. 即存在正交阵R使得

$$R\beta_i = \alpha_i,\ 1 \leq i \leq n.$$

则断言$A = RP$. 只需证$AO = RPO$. AO的第一列就是B的第一列, 等于$\lambda_1\alpha_1$. PO的第一列等于$\lambda_1\beta_1$. 由于$R\beta_1 = \alpha_1$,所以AO的第一列等于RPO的第一列. 同理可证AO和RPO的前k列都是相同的, 它们的后$n-k$列都是0. 所以$AO = RPO$.

因为AO的前k列线性无关, 后$n-k$列都是0. 同样的, RPO的后$n-k$列也都是0. 而AO, RPO的前k列相等, 所以$AO=RPO$.

如果A是可逆的, 那么P是正定的.

现在来证明唯一性. 如果$A=RP$, 其中R是一个正交矩阵, P 是一个半正定矩阵. 则因为

$$P^2=PR'RP=A'A$$

我们知道满足这个性质的P是唯一的, 所以R也是唯一的.

□

由上面的定理可知, 也一定存在唯一的矩阵对$(R,\ P)$, 其中R是一个正交矩阵, P 是一个半正定矩阵, 使得

$$A=PR.$$

这个分解方式之所以叫做极分解, 是因为可以类比于复数的极坐标的写法, 任何一个复数都可以写成

$$\alpha=re^{i\theta}$$

的形式. 极分解中的P可以类比于这里的长度r, 正交矩阵可以类比于这里的$e^{i\theta}$.

定理 8.2.20 ($n\times n$ 矩阵的奇异值分解)**.** *假设A是一个$n\times n$ 矩阵, 则存在正交矩阵U, V使得UAV是对角矩阵, UAV的对角元叫做A的奇异值.*

证明 由矩阵的极分解, A可以利用极分解写成PC的形式, 其中P是半正定矩阵, C是正交矩阵,? 然后P可以正交相似于对角矩阵, 即它可以写成$P=USU'$的形式, 其中U是正交矩阵. 所以

$$A=PC=USU'C,$$

其中$U'C$是两个正交阵的乘积, 当然还是正交阵. 令$V=U'C$即可.

□

§8.3 例题

1. 用正交替换把二次型化为标准形

   (1). 写出二次型对应的矩阵A.

   (2). 求出A的特征值λ_1, ..., λ_k.

   (3). 对每个λ_i, 解线性方程组$(A-\lambda_i I)X=0$, 求出λ_i所对应的特征子空间E_{λ_i}的一组基α_{i1}, ..., α_{ia_i}. 由于实对称矩阵是可对角化的, 所以V能写成特征子空间E_{λ_i}的直和, 这些特征子空间E_{λ_i}的基拼起来就是V的基.

(4). 将α_{i1}, ..., α_{ia_i}进行施密特标准正交化得到一组标准正交基β_{i1}, ..., β_{ia_i}, 由于不同的特征值对应的特征向量是正交的, 所以这些特征子空间的正交基拼起来就是V的标准正交基.

(5). β_{11}, ..., β_{1a_1}, ..., β_{k1}, ..., β_{ka_k}是V的标准正交基. 以这组基向量为列向量组成的矩阵是正交阵, 记为T. 则$T'AT$为对角矩阵.

(6). 要求的正交线性替换是$X = TY$.

找个例子.

Rayleigh 设S为n-级实对称矩阵, $S^{n-1} = \{X \in \mathbb{R}^n | x_1^2 + \cdots + x_n^2 = 1\}$为$n-1$维球面, 则

$$\lambda \le X'SX \le \mu,$$

其中λ是S的最小的特征值, μ是A 的最大的特征值. **证明** 由于正交变换把单位球面还变为单位球面, 把实对称矩阵变为对角矩阵, 所以不妨假设S就是对角矩阵$diag(\lambda_1, ..., \lambda_n)$. 所以

$$\lambda \le X'SX = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 \le \mu.$$

注意$x_1^2 + \cdots + x_n^2 = 1$.

□

§8.4 最小二乘法

假设A是一个$m \times n$的矩阵, b是一个m维的列向量. 我们知道线性方程组

$$AX = b$$

有解当且仅当A的秩等于增广矩阵(A, b)的秩. 如果增广矩阵(A, b)的秩大于A的秩, 方程组就无解. 这时候, 我们想找一个向量X_0 使得$|AX_0 - b|$最小. 如果线性方程组有解, 这个最小值当然就是0. 如果没有解, 这个最小值就是个正数. 我们下面看看这个最小值是多少, 以及如何找到达到最小值的X_0.

我们知道$\{AX | X \in \mathbb{R}^n\}$是一个子空间, 记为$L$, L是由A的列向量张成的子空间. 那么对任意的$X \in V = \mathbb{R}^n$, $|AX - b|$ 就是AX到b的距离. 因为问题就转化成求b到子空间L的最小距离. 从几何上很容易想像, 这应该是从b到L做一条垂线, b到子空间L的最小距离就是从b到垂足的距离. 这条垂线必须与L中的每一条直线都垂直. 由于L是由A的列向量张成的子空间, 这当且仅当于$AX_0 - b$和A的列向量都垂直, 即

$$A'(AX_0 - b) = 0, \ i.e., A'AX_0 = A'b.$$

由于$A'A$的秩与增广矩阵$(A'A, A'b)$的秩相等, 所以X_0总是有解的.

这个找X_0的方法就叫做最小二乘法. 最小二乘法有很多实际的应用.

§8.5 酉空间

定义 8.5.1 (酉空间). 假设V是n维的复线性空间. 假设f是从集合$V\times V$到$\mathbb{C}$的一个映射, 而且满足下面的条件

$$f(\alpha,\ \beta)=\overline{f(\beta,\ \alpha)}$$
$$f(\alpha_1+\alpha_2,\ \beta)=f(\alpha_1,\ \beta)+f(\alpha_2,\ \beta)$$
$$f(k\alpha,\ \beta)=kf(\alpha,\ \beta),\ k\in\mathbb{C},$$
$$f(\alpha,\ \alpha)\geq 0,\quad \text{等号成立当且仅当}\alpha=0,$$

那么我们称$f(\alpha,\ \beta)$为$\alpha,\ \beta$的内积, 以后简单的记为$(\alpha,\ \beta)$. 我们称带有这种内积的复空间V为酉*(unitary)*空间。

例 8.5.1. 我们假设$V=\mathbb{C}^n$, 对任意的$\alpha=(a_1,\ ...,\ a_n)'$, $\beta=(b_1,\ ...,\ b_n)'$, 我们定义

$$(\alpha,\ \beta)=a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n},$$

不难验证, 这样定义的内积使得V称为一个酉空间.

在酉空间中, 我们可以类似的定义向量α的长度为

$$|\alpha|=\sqrt{(\alpha,\ \alpha)}.$$

这个长度也满足Cauchy-Bunyakovsky-Schwarz 不等式,

$$(\alpha,\ \beta)^2\leq(\alpha,\ \alpha)(\beta,\ \beta).$$

我们也可以类似的定义正交和正交补等概念.

定义 8.5.2 (酉矩阵). n-级复矩阵A如果满足

$$A\overline{A'}=\overline{A'}A=I,$$

我们就称A是一个酉矩阵.

定义 8.5.3 (埃尔米特(Hermite)矩阵). n-级复矩阵A如果满足

$$A=\overline{A'},$$

我们就称A是一个$Hermite$矩阵.

定义 8.5.4 (酉相似矩阵). 对n-级复矩阵$A,\ B$, 如果存在酉矩阵U使得

$$UA\overline{U'}=B,$$

我们就称A酉相似于B.

容易看出酉相似是一个等价关系. 注意两个矩阵复相似, 它们不一定是酉相似的.

命题 8.5.1. 假设$A=(a_{ij})$和$B=(b_{ij})$是两个复矩阵. 如果他们酉相似, 那么

$$\sum_{i,j=1}^{n}|a_{ij}|^2=\sum_{i,j=1}^{n}|b_{ij}|^2.$$

证明 如果存在酉矩阵T使得

$$\overline{T'}AT=B,$$

那么

$$Tr(B\overline{B'})=Tr(\overline{T'}A\overline{A'}T)=Tr(A\overline{A'}T\overline{T'})=Tr(A\overline{A'}).$$

所以命题成立. □

例 8.5.2 (相似但不酉相似)**.**

$$A=\begin{pmatrix}3&1\\-2&0\end{pmatrix},\ B=\begin{pmatrix}1&1\\0&2\end{pmatrix}$$

这个两个矩阵是相似的. 但是由上面的命题可知, 它们不是酉相似的.

定理 8.5.2. 埃尔米特矩阵的特征值都是实数, 并且属于不同的特征值的特征向量互相正交.

证明 假设A是一个埃尔米特矩阵, $\lambda\in\mathbb{C}$是A的一个特征值, $\alpha\in\mathbb{C}^n$是A的一个特征向量. 那么我们有

$$A\alpha=\lambda\alpha\Longrightarrow\overline{\alpha'}A\alpha=\lambda\overline{\alpha'}\alpha,$$

两边再取共轭转置可得

$$\overline{\alpha'}A\alpha=\overline{\lambda\alpha'}\alpha.$$

所以

$$\lambda\overline{\alpha'}\alpha=\overline{\lambda\alpha'}\alpha.$$

所以$\lambda=\overline{\lambda}$, 即$\lambda$是实数.

假设$A\alpha=\lambda\alpha, A\beta=\mu\beta$, 其中$\alpha$, β是非零向量, λ, μ是互不相同的复数. 则

$$\overline{\beta'}A\alpha=\overline{\beta'}\lambda\alpha=\lambda\overline{\beta'}\alpha.$$

另一方面

$$\overline{\beta'}A\alpha=\overline{(A\beta)'}\alpha=\overline{(\mu\beta)'}\alpha=\mu\overline{\beta'}\alpha.$$

又因为λ, μ是互不相同的复数, 所以$\overline{\beta'}\alpha=0$, 所以他们是互相垂直的. □

定理 8.5.3. 埃尔米特矩阵酉相似于对角矩阵.

证明 可以对阶数进行归纳.

假设

$$A\alpha = \lambda\alpha,$$

其中α是单位长度的特征向量, λ是特征值. 那么将α扩成一组标准正交基,

$$\alpha_1 = \alpha,\ \alpha_2,\ ...,\ \alpha_n.$$

把这n个列向量拼在一起, 就成了一个酉矩阵O. 我们有

$$AO = O\begin{pmatrix}\lambda & * \\ 0 & *\end{pmatrix} \Longrightarrow O'AO = \begin{pmatrix}\lambda & * \\ 0 & *\end{pmatrix}$$

由于A是埃尔米特矩阵, 所以右边这个带星号的矩阵也是对称的. 所以A正交相似于

$$\begin{pmatrix}\lambda & 0 \\ 0 & B\end{pmatrix}$$

形状的埃尔米特阵, 归纳即可. □

定理 8.5.4. 假设A是n-级实矩阵, B是n-级正定矩阵, k是一个正整数, $AB^k = B^kA$. 则$AB = BA$.

证明

由于对A, B做相同的正交变换不影响结论的成立, 而B由于是实对称矩阵, 所以总能正交相似于对角矩阵. 所以我们不妨假设$B = diag(\lambda_1 I_{a_1},\ ...,\ \lambda_t I_{a_t})$, 其中$\lambda_1,\ ...,\ \lambda_t$是两两不同的正数. 则$B^k = diag(\lambda_1^k I_{a_1},\ ...,\ \lambda_t^k I_{a_t})$, 其中$\lambda_1^k,\ ...,\ \lambda_t^k$必然也是两两不同的. 则由于$AB^k = B^kA$可以推知$A$的非对角块都等于零, 所以$B = diag(B_1,\ ...,\ B_t)$是和$A$分块形式相同的准对角阵. 所以$\lambda_i I_{a_i} B_i = B_i \lambda_i I_{a_i}$对任意的$1 \le i \le t$都成立, 所以$AB = BA$.

□

定理 8.5.5. 假设A, B是两个n-级实矩阵且A酉相似于B. 求证A正交相似于B.

证明 因为A酉相似于B, 所以存在酉矩阵$U = P + iQ$, 使得

$$AU = UB,$$

这里的P, Q两个n-级实矩阵, i是虚数单位. P, Q可以看成是U的实部和虚部. 所以

$$AP = PB,\ AQ = QB.$$

由$BU' = U'A$可得

$$BP' = P'A,\ BQ' = Q'A.$$

记$R_t = P + tQ$, t是待定参数. 则一定存在实数t使得R_t是可逆矩阵. 这是因为多项式

$$f(t) = |P + tQ|$$

只能有有限多个实根, 所以只要t不是它的实根就行了. 由于实数有无限多个, 这个当然可以做到. 还需要说明$f(t)$不是零多项式. 这可以从$f(i) = |U| \neq 0$看出来.

注意到R_t满足

$$AR_t = R_tB, R_t'A = BR_t'.$$

因为R_tR_t'正定, 所以存在正定矩阵C使得$C^2 = R_tR_t'$. 记$T = C^{-1}R_t$, 则

$$TT' = C^{-1}R_tR_t'C^{-1} = I$$

是单位矩阵, 所以T是正交矩阵. 由于

$$AC^2 = AR_tR_t' = R_tBR_t' = R_tR_t'A = C^2A,$$

所以$AC = CA$. 所以

$$AT = AC^{-1}R_t = C^{-1}AR_t = C^{-1}R_tB = TB.$$

所以A和B正交相似. □

定义 8.5.5 (正规矩阵)**.** *假设A是一个n级复矩阵, 如果$A\overline{A'} = \overline{A'}A$, 我们就称$A$是一个正规矩阵.*

定理 8.5.6. *假设A是一个n级复矩阵. 则A是一个正规矩阵当且仅当A酉相似于对角矩阵.*

证明 我们只需要证明如果A是一个正规矩阵, 则A酉相似于对角矩阵. 另外一个方向直接由定义可知成立.

记

$$B = \frac{1}{2}(A + \overline{A'}),\ C = -\frac{i}{2}(A - \overline{A'}),$$

则$A = B + iC$. 因为$A\overline{A'} = \overline{A'}A$, 所以

$$(B + iC)(\overline{B'} - i\overline{C'}) = (\overline{B'} - i\overline{C'})(B + iC).$$

因为$\overline{B'} = B$, $\overline{C'} = C$, 所以上面的式子就变成了

$$2i(BC - CB) = 0,$$

由此可知$BC=CB$. 所以B, C是两个可交换的矩阵, 因为他们都是埃尔米特阵, 所以它们都酉相似于对角矩阵. 所以它们可以同时酉相似于对角矩阵, 所以A可以酉相似于对角矩阵.

所以A是可以对角化的. □

推论 8.5.7. *假设A是一个n级实正规矩阵. 则A是对称矩阵当且仅当A的特征值都是实数.*

证明

如果A是实对称矩阵, 那么由于任何一个实对称矩阵都正交相似于对角矩阵, 它的特征值当然都是实数.

如果A的特征值都是实数, 那么由上面的定理可知A酉相似于实对角矩阵. 由于实矩阵之间的酉相似蕴含着正交相似, 所以A也正交相似于一个实对角矩阵, 所以A是对称矩阵.

□

定理 8.5.8 (Schur不等式)**.** *假设$A=(a_{ij})$是一个n级复矩阵, λ_1, ..., λ_n是A的特征值. 则*

$$\sum_{i,j=1}^{n}|a_{ij}|^2\geq|\lambda_1|^2+\cdots+|\lambda_n|^2,$$

等号成立当且仅当A是正规矩阵.

证明 首先我们知道, 对A进行酉相似变换不改变不等式的左边. 由于任何一个复矩阵都酉相似于一个上三角矩阵, 所以我们不妨假设A是上三角矩阵, 对角元为λ_1, ..., λ_n. 所以Schur不等式是成立的.

等号成立当且仅当上三角矩阵是对角矩阵, 所以当且仅当A是正规矩阵.

□

习题 8.5

1. 任何一个复矩阵都酉相似于一个上三角矩阵.

   证明

   任何一个复矩阵A都有一个复特征值λ及其特征向量α, 即$A\alpha=\lambda\alpha$. 我们不妨假设α的长度为1, 将其扩充为酉空间的一组标准正交基, 即α是一个酉矩阵U 的第一列. 这样的话

$$U^{-1}AU=\begin{pmatrix}\lambda & *\\ 0 & A_1\end{pmatrix}$$

就是一个准三角矩阵, 其中A_1是$(n-1)\times(n-1)$的, 对矩阵的阶数归纳即可.

$\square$

2. 上三角的正交矩阵A必为对角矩阵, 且对角元为1或者-1

证明 先看A的第一列, 由于A是正交矩阵, 所以第一列的长度为1, 所以$a_{11}=\pm1$. 再看A的第一行, 由于这一行作为行向量, 长度是1, 所以除了a_{11}之外, 都为0. 依次类推即可.

$\square$

3. 任何一个可逆实矩阵A都可以分解为

$$A=QT$$

的形式, 其中Q是正交矩阵, T是上三角矩阵.

证明 存在性由施密特标准正交化过程可知, 唯一性有上一道题可知.

$\square$

4. 任何一个正定矩阵A都可以写成$A=T'T$的形式, 其中T是一个上三角矩阵.

证明

$$A=C'C=(QT)'QT=T'Q'QT=T'T.$$

$\square$

5. 假设$\mathscr{A}$是一个正交变换, 则$\mathscr{A}$的不变子空间W的正交补也是不变子空间.

证明 这相当于说任取$\alpha\in W^{\perp}$, 仍然有$\mathscr{A}\alpha\in W^{\perp}$. 即任取$\beta\in W$, 都有$(\mathscr{A}\alpha,\ \beta)=0$. 由于$\mathscr{A}$是$W$上的可逆变换, 所以一定有$\gamma\in W$使得$\beta=\mathscr{A}\gamma$. 因此

$$(\mathscr{A}\alpha,\ \beta)=(\mathscr{A}\alpha,\ \mathscr{A}\gamma)=(\alpha,\ \gamma)=0.$$

$\square$

6. 假设η是欧氏空间中的一个单位向量, 定义

$$\mathscr{A}\alpha=\alpha-2(\eta,\ \alpha)\eta.$$

证明

(1) $\mathscr{A}$是第二类正交变换. 这样的$\mathscr{A}$叫做镜面反射.

(2) 如果n维欧氏空间中, 正交变换$\mathscr{A}$有一个特征值1, 而且特征值1的特征子空间的维数为$n-1$, 则$\mathscr{A}$ 一定是镜面反射.

证明

(1)

$$
\begin{aligned}
(\mathscr{A}\alpha,\ \mathscr{A}\alpha) &= (\alpha,\ \alpha) - 4(\eta,\ \alpha)(\alpha,\ \eta) + 4(\eta,\ \alpha)^2(\eta,\ \eta) \\
&= (\alpha,\ \alpha) - 4(\eta,\ \alpha)(\alpha,\ \eta) + 4(\eta,\ \alpha)^2 \\
&= (\alpha,\ \alpha).
\end{aligned}
$$

所以$\mathscr{A}$保持任何一个向量的长度不变, 从而保持内积不变, 所以是正交变换.

从$\mathscr{A}$的定义可知, $\mathscr{A}$在$L(\eta)$的正交补上是恒等变换, 而在$L(\eta)$上的作用相等于添一个负号, 即$\mathscr{A}$正交相似于

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

所以它的行列式是-1, 从而是第二类正交变换.

(2) 假设1的特征子空间是W, 设η是W的正交补中的单位向量. 则$L(\eta)$是正交变换$\mathscr{A}$的不变子空间. 所以

$$\mathscr{A}(\eta) = \pm\eta.$$

如果$\mathscr{A}(\eta) = \eta$. 那么η也是1的特征向量, 从而特征值1的特征子空间的维数为n, 与题设的特征值1的特征子空间的维数为$n-1$矛盾. 所以$\mathscr{A}(\eta) = -\eta$.因此对任意的$\alpha \in V$ 都有

$$\mathscr{A}\alpha = \alpha - 2(\eta,\ \alpha)\eta.$$

所以是一个镜面反射. □

7. 假设$\alpha,\ \beta$是n-维欧氏空间V中的两个单位向量. 则一定存在一个镜面反射$\mathscr{A}$使得$\mathscr{A}(\alpha) = \beta$.

证明 令$\eta = (\alpha - \beta)/|\alpha - \beta|$. 对任意的$\gamma \in V$ 定义

$$\mathscr{A}\gamma = \gamma - 2(\eta,\ \gamma)\eta.$$

则$\mathscr{A}$是镜面反射使得$\mathscr{A}(\alpha) = \beta$.

□

8. 假设$\mathscr{A}$是欧氏空间的一个变换, 如果$\mathscr{A}$保持内积不变, 即对于任意的α, β都有$(\mathscr{A}\alpha,\ \mathscr{A}b) = (\alpha,\ \beta)$, 则它一定是线性的, 因而是正交变换.

证明

(1)

先证明$\mathscr{A}(\alpha_1+\alpha_2) = \mathscr{A}\alpha_1 + \mathscr{A}\alpha_2$,即$\mathscr{A}(\alpha_1+\alpha_2) - \mathscr{A}\alpha_1 - \mathscr{A}\alpha_2 = 0$. 这只需要证明$\mathscr{A}(\alpha_1+\alpha_2) - \mathscr{A}\alpha_1 - \mathscr{A}\alpha_2$的长度等于0即可.

$$
\begin{aligned}
&(\mathscr{A}(\alpha_1+\alpha_2) - \mathscr{A}(\alpha_1) - \mathscr{A}(\alpha_2),\ \mathscr{A}(\alpha_1+\alpha_2) - \mathscr{A}(\alpha_1) - \mathscr{A}(\alpha_2)) \\
&= (\mathscr{A}(\alpha_1+\alpha_2),\ \mathscr{A}(\alpha_1+\alpha_2)) + (\mathscr{A}(\alpha_1),\ \mathscr{A}(\alpha_1)) + (\mathscr{A}(\alpha_2),\ \mathscr{A}(\alpha_2)) \\
&\quad - 2(\mathscr{A}(\alpha_1+\alpha_2),\ \mathscr{A}(\alpha_1)) - 2(\mathscr{A}(\alpha_1+\alpha_2),\ \mathscr{A}(\alpha_2)) + 2(\mathscr{A}(\alpha_1),\ \mathscr{A}(\alpha_2)) \\
&= (\alpha_1+\alpha_2,\ \alpha_1+\alpha_2) + (\alpha_1,\ \alpha_1) + (\alpha_2,\ \alpha_2) \\
&\quad - 2(\alpha_1+\alpha_2,\ \alpha_1) - 2(\alpha_1+\alpha_2,\ \alpha_2) + 2(\alpha_1,\ \alpha_2) \\
&= 0.
\end{aligned}
$$

由于在欧几里得空间中长度为0的向量只有0. 所以$\mathscr{A}(\alpha_1+\alpha_2) - \mathscr{A}\alpha_1 - \mathscr{A}\alpha_2 = 0$. 所以$\mathscr{A}(\alpha_1+\alpha_2) = \mathscr{A}\alpha_1 + \mathscr{A}\alpha_2$.

(2)类似的可以证明

$$\mathscr{A}(k\alpha) = k\mathscr{A}\alpha,$$

所以$\mathscr{A}$一定是线性的, 因而是正交变换.

□

9. 假设$\alpha_1,\ ...,\ \alpha_m$和$\beta_1,\ ...,\ \beta_m$是$n$维欧氏空间$V$中的两个向量组. 证明存在一个正交变换$\mathscr{A}$使得

$$\mathscr{A}\alpha_i = \beta_i,\ i = 1,\ 2,\ ..., m$$

的充分必要条件为

$$(\alpha_i,\ \alpha_j) = (\beta_i,\ \beta_j),\ i,\ j = 1,\ 2,\ ...,\ m.$$

证明 (1) 如果$m = n$ 而且$\alpha_1,\ ...,\ \alpha_n$是$n$维欧氏空间$V$的一组基, 那么定义映射

$$
\begin{aligned}
\mathscr{A} : V &\longrightarrow V \\
\sum_{i=1}^{n} a_i\alpha_i &\longmapsto \sum_{i=1}^{n} b_i\beta_i,
\end{aligned}
$$

易知这是一个映射, 由上一题可知$\mathscr{A}$是线性的, 所以是正交变换.

(2) 如果$\alpha_1,\ ...,\ \alpha_m$能够张成$V$, 我们不妨设其中的$\alpha_1,\ ...,\ \alpha_n$是$V$的一组基. 那么那么定义映射

$$\mathscr{A}: V \longrightarrow V$$
$$\sum_{i=1}^{n} a_i\alpha_i \longmapsto \sum_{i=1}^{n} b_i\beta_i,$$

从(1) 可知, 这是一个正交变换. 我们只需要证明

$$\mathscr{A}\alpha_j = \beta_j,\ j = n+1,\ n+2,\ ..., m.$$

注意到对任意的$\alpha = \sum\limits_{i=1}^{n} a_i\alpha_i$, 和任意的$j = n+1,\ n+2,\ ..., m$.

$$\begin{aligned}(\mathscr{A}\alpha,\ \mathscr{A}\alpha_j) &= (\mathscr{A}(\sum_{i=1}^{n} a_i\alpha_i),\ \mathscr{A}\alpha_j) \\ &= (\sum_{i=1}^{n} a_i\alpha_i,\ \alpha_j) \\ &= (\sum_{i=1}^{n} a_i\alpha_i,\ \alpha_j) \\ &= (\sum_{i=1}^{n} a_i\mathscr{A}\alpha_i,\ \beta_j) \\ &= (\mathscr{A}\alpha,\ \beta_j)\end{aligned}$$

由α的任意性可知$\mathscr{A}\alpha_j = \beta_j$.

(3) 如果$\alpha_1,\ ...,\ \alpha_m$不能够张成$V$. 我们记$W = L(\alpha_1,\ ...,\ \alpha_m)$. 取$W^\perp$的一组标准正交基$\alpha_{m+1},\ ...,\ \alpha_t$. 再记$U = L(\beta_1,\ ...,\ \beta_m)$. 则考虑度量矩阵易知$W$和$U$的维数相同. 所以它们的正交补的维数也相同, 取$U^\perp$的一组标准正交基$\beta_{m+1},\ ...,\ b_t$. 则我们有

$$(\alpha_i,\ \alpha_j) = (\beta_i,\ \beta_j),\ i,\ j = 1,\ 2,\ ...,\ t.$$

由(2)可知命题成立. □

10. 实对称阵A的一阶主子式之和是0, 二阶主子式之和也是0, 求证$A = 0$.

证明 首先要知道实对称阵的特征值都是实数. 假设实对称阵A的特征值为

$$\lambda_1,\ \lambda_2,\ ...,\ \lambda_n.$$

则A的一阶主子式之和为$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$. 二阶主子式之和为

$$\sum_{1\le i\ne j\le n} \lambda_i\lambda_j = 0.$$

所以
$$\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2=0.$$
所以
$$\lambda_1=\lambda_2=\cdots=\lambda_n=0.$$
所以$A=0$.

□

11. 假设A是一个n-级实矩阵, 求证存在一个可逆阵R使得AR是对称阵.

证明 根据矩阵的极分解, 存在正交阵R使得AR是半正定矩阵. □

12. 辛矩阵的行列式为1.

证明 辛矩阵的定义是
$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}\begin{pmatrix}A' & C'\\ B' & D'\end{pmatrix}=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}.$$
这就得到了
$$\begin{cases}BA' & =AB',\\ CD' & =DC',\\ AD'-BC' & =I,\\ CB'-DA' & =-I.\end{cases}$$
如果A是可逆的, 那么
$$\begin{vmatrix}A & B\\ C & D\end{vmatrix}=\begin{vmatrix}A & B\\ 0 & D-CA^{-1}B\end{vmatrix}.$$
我们知道
$$\begin{aligned}|D-CA^{-1}B||A'| &= |DA'-CA^{-1}BA'|\\ &= |DA'-CA^{-1}AB'|\\ &= |DA'-CB'|=1.\end{aligned}$$

如果A不是可逆的, 那么由实矩阵的极分解可知, 存在正交阵R使得BR'是对称阵, 所以
$$BR'=RB'.$$
因为
$$|A+\lambda R|=\lambda^n|R+\frac{1}{\lambda}A|,$$

记$\delta = \frac{1}{\lambda}$. 则当$\delta \longrightarrow 0$时, $|R + \frac{1}{\lambda}A| \longrightarrow |R| = \pm 1$. 所以当$\lambda$足够大的时候, $|A + \lambda R| \neq 0$. 所以$|A + \lambda R|$不是关于λ的零多项式, 所以$|A + \lambda R| = 0$ 只有有限多个根, 所以存在$\varepsilon > 0$使得当$\lambda \in (0,\ \varepsilon)$时, $|A + \lambda R| \neq 0$. 这时有

$$\begin{pmatrix} A + \lambda R & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A' + \lambda R' & C' \\ B' & D' \end{pmatrix} = \begin{pmatrix} 0 & I + \lambda RD' \\ -I - \lambda DR' & 0 \end{pmatrix}.$$

因为当适$\lambda \in (0,\ \varepsilon)$时, $|A + \lambda R| \neq 0$, 所以

$$\begin{vmatrix} A + \lambda R & B \\ C & D \end{vmatrix} = \begin{vmatrix} A + \lambda R & B \\ 0 & D - C(A + \lambda R)^{-1}B \end{vmatrix}.$$

我们知道

$$\begin{aligned} & |D - C(A + \lambda R)^{-1}B||(A + \lambda R)'| \\ = & |DA' + \lambda DR' - C(A + \lambda R)^{-1}B(A + \lambda R)'| \\ = & |DA' + \lambda DR' - C(A + \lambda R)^{-1}(A + \lambda R)B'| \\ = & |DA' - CB' + \lambda DR'| = |I + \lambda DR'|. \end{aligned}$$

即

$$\begin{vmatrix} A + \lambda R & B \\ C & D \end{vmatrix} = |I + \lambda DR'|$$

对任意的$\lambda \in (0,\ \varepsilon)$成立. 由于上面等式的两边都是关于$\lambda$的多项式. 所以等式两边作为关于$\lambda$的多项式是相同的, 所以对任意$\lambda$的值都成立. 特别的对$\lambda = 0$也成立. 所以

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = 1.$$

□

13. 假设V_1, V_2是n维欧氏空间的两个子空间. 如果V_1的维数小于V_2的维数. 求证在V_2中存在一个非零的向量α, 使得$\alpha \perp V_1$.

证明 利用维数公式, 可知V_1的正交补的维数加上V_2的维数大于V的维数, 所以V_1的正交补和V_2的交不等于0, 其中必有非零向量α.

□

14. 假设S是个半正定矩阵, O是一个正交矩阵, 而且SO的特征多项式和S的特征多项式相同, 求证$SO = S$.

证明 SO的每个元素绝对值的平方和等于$Tr(SO(SO)') = Tr(S^2)$, 等于S^2的特征值之和, 即S的特征值的平方和. 因为SO的特征多项式和S的特征多项式相同, 所以SO的每个元素绝对值的平方和等于SO的特征值的平方和. 因此由Schur不

等式, SO是正规矩阵. 因为SO是正规矩阵, 而且特征值都是实数, 所以SO 是对称矩阵, 特征值都大于等于0. 因此SO是半正定的. 所以SO有两种极分解,

$$SO = SO \cdot I = S \cdot O.$$

由极分解的唯一性可知, $SO = S$. $\square$

第九章　双线性函数与辛空间

§9.1　对偶空间

定义 9.1.1 (对偶空间). 假设V是域F上的线性空间. 把F看作是以为F上的一维的线性空间. 那么从V 到F的所有的线性映射全体形成的线性空间叫做是V的对偶空间, 记为V^*.

定义 9.1.2 (对偶基). 假设V是域F上的n维线性空间, 有一组基e_1, ..., e_n. 定义f_i为V^* 中满足下面条件的元素

$$f_i(e_j) = \delta_{ij},\ 1 \le j \le n.$$

则f_1, ..., f_n称为是V^*的对偶基. 注意对无限维的线性空间来说, 按照上面的方式定义的f_i虽然彼此之间是线性无关的, 但却不能张成V^*.

当V是有限维线性空间的时候, V和V^*的维数相同, 所以它们是同构的. 但是这两个线性空间之间却没有一个"典范"同构, 即所有的同构都要借助于基的选取. 没有哪个同构能够说自己是独特的, 大家的地位彼此都是平等的.

我们记V^{**}为V的对偶空间的对偶空间.

定理 9.1.1. 假设V是有限维线性空间. 则V和V^{**}是同构的, 而且它们之间存在一个典范同构.

证明 我们直接把这个典范同构写出来.

$$\varphi:\ V \longrightarrow V^{**},\ \alpha \mapsto \varphi(\alpha)$$

这里的$\varphi(\alpha)$把任意的$f \in V^*$都映为$f(\alpha)$.

易知这是个同构, 且定义方式不依赖于基的选取.

□

定理 9.1.2 (Reisz表示定理). 假设V是一个n维的欧氏空间, 它的内积为(α, β). 那么对V的任何一个元素α, 我们定义$\alpha^* \in V^*$:

$$\alpha^*(\beta) = (\alpha, \beta),\ 任意\beta \in V.$$

则映射

$$\varphi:\ V \longrightarrow V^*,\ \alpha \longmapsto \alpha^*.$$

是一个同构.

证明 首先易知α^*确实是个线性函数, 而φ也是一个线性映射. 其实如果$\alpha \in V$而$\varphi(\alpha) = 0$, 那么$(\alpha, \beta) = 0$对任意的β成立, 特别的, 对$\beta = \alpha$也成立, 由内积的非退化性可知$\alpha = 0$. 所以说φ是单射, 所以φ是个同构.

□

§9.2 双线性函数

定义 9.2.1 (双线性函数)**.** *假设V 是域F上的线性空间, $f(\alpha,\ \beta)$是$V \times V$上的函数, 而且对任何一个分量都是线性的, 我们就称f是V上的一个双线性函数. 我们常常把$f(\alpha,\ \beta)$简单的记为$(\alpha,\ \beta)$*

如果$(\alpha,\ \beta) = (\beta,\ \alpha)$, 我们称$f$是对称双线性函数. 如果$(\alpha,\ \beta) = -f(\beta,\ \alpha)$, 我们称$f$是反对称双线性函数. 如果$F = \mathbb{C}$, $(\alpha,\ \beta) = \overline{f(\beta,\ \alpha)}$, 我们称$f$是Hermitian双线性函数.

定义 9.2.2 (双线性函数的度量矩阵)**.** *假设f是V上的一个双线性函数, $\alpha_1,\ ...,\ \alpha_n$的一组基. 我们称下面的矩阵为$f$在$\alpha_1,\ ...,\ \alpha_n$这组基下的度量矩阵.*

$$\begin{pmatrix} f(\alpha_1,\ \alpha_1) & f(\alpha_1,\ \alpha_2) & \cdots & f(\alpha_1,\ \alpha_n) \\ f(\alpha_2,\ \alpha_1) & f(\alpha_2,\ \alpha_2) & \cdots & f(\alpha_2,\ \alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(\alpha_n,\ \alpha_1) & f(\alpha_n,\ \alpha_2) & \cdots & f(\alpha_n,\ \alpha_n) \end{pmatrix}.$$

容易看出, 双线性函数在不同的基下的度量矩阵是合同的. 如果从$\alpha_1,\ ...,\ \alpha_n$到$\beta_1,\ ...,\ \beta_n$的过渡矩阵是$C$, 即

$$(\alpha_1,\ ...,\ \alpha_n)C = (\beta_1,\ ...,\ \beta_n)$$

则f在$\alpha_1,\ ...,\ \alpha_n$之下的度量矩阵$A$和它在$\beta_1,\ ...,\ \beta_n$之下的度量矩阵$B$之间的关系式为

$$B = C'AC.$$

命题 9.2.1. *假设f是数域F上的线性空间V上的一个双线性函数. 下面论断等价.*

(1)则f在某组基下的度量矩阵是可逆的.

(2)如果(α, β)对所有的$\beta \in V$成立, 则$\alpha = 0$.

(3)如果(α, β)对所有的$\alpha \in V$成立, 则$\beta = 0$.

证明 如果f在某组基下的度量矩阵是不可逆的, 那么行向量的一个非平凡的线性组合是零向量, 把这个非平凡的线性组合的系数都写到度量矩阵元素的第一个分量里, 这就得到了一个非零向量与所有的基向量做内积都是0.

□

假设上面命题中的三个等价的条件成立, 则我们有同构

$$\varepsilon: V \longrightarrow V^*$$
$$\alpha \mapsto \varepsilon(\alpha),$$

其中$\varepsilon(\alpha)(\beta) := (\alpha, \beta)$ 对任意$\beta \in V$. 由于V和V^*的维数相同, 所以由维数公式, 只需要说明ε是单射, 就可说明是同构. 考虑到$\varepsilon(\alpha) = 0$当且仅当对所有的$\beta \in V$, 都有$(\alpha, \beta) = 0$, 说明如果存在这样的非零向量, 就说明了二次空间V 是退化的, 和假设矛盾. 所以ε是单射, 从而是同构.

容易看出如果f是数域F上的线性空间V上的一个双线性函数, 则它在任何基下的度量矩阵是对称矩阵. 如果它是反对称双线性函数, 则它在任何基下的矩阵都是反对称矩阵. 如果它是Hermitian双线性函数, 则它在任何基下的矩阵都是Hermite矩阵.

定义 9.2.3 (二次空间, 二次型). *假设f是数域F上的线性空间V上的一个对称双线性函数, 称函数*

$$Q: V \longrightarrow F, \ \alpha \longmapsto f(\alpha, \alpha)$$

为V上的二次型, 称带有这个二次型的V为二次空间, 记为(V, Q). 在Q的含义明确的情况下, 也常常省略掉Q, 直接说V是二次空间. 如果f是非退化的对称双线性型, 我们就称V是非退化的二次空间, 否则就称V 是退化的二次空间. 我们常常把$Q(\alpha)$简单的记为(α, α),或者$|\alpha|^2$.

定义 9.2.4 (二次空间之间的等距). *假设f是数域F上的两个二次空间(V, Q)到(V, Q')的线性映射, 如果对任意的$\alpha \in V$, 都有*

$$Q'(f(\alpha)) = Q(\alpha),$$

我们就称f是二次空间之间的等距. 如果$(V, Q) = (V, Q')$, 我们称f是V的一个等距变换.

如果(V, Q)是一个非退化的二次空间, f是(V, Q)到(V, Q')的等距, 则f一定是单射. 否则一定存在α使得$f(\alpha) = 0$, 因此对任意的$\beta \in V$都有

$$(\alpha, \beta) = (f(a), f(b)) = 0,$$

这和我们假设V是非退化的矛盾.

假设V是一个二次空间. 如果$(\alpha, \beta) = 0$, 我们就称α, β正交, 和整个空间正交的向量全体叫做f的核, 或者叫做V 的根空间, 记为$rad(V)$. 容易看出f是非退化的当且仅当$rad(V) = 0$.

定义 9.2.5 (迷向向量, 迷向子空间)**.** 假设V是数域F上的二次空间, Q 为其二次型. 如果非零向量α满足$Q(\alpha)=0$, 我们就称α是一个迷向向量. 完全由零向量和迷向向量组成的子空间叫做迷向子空间.

如果我们取定了V的一组基$\alpha_1,\ ...,\ \alpha_n$, 那么$\alpha$可以写成基向量的线性组合

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n.$$

因此

$$Q(\alpha) = \sum_{i,j=1}^{n} f(\alpha_i,\ \alpha_j)x_ix_j$$

就是一个n元二次齐次多项式, 和我们前面学习的二次型的定义一致.

定义 9.2.6 (正交, 正交基)**.** 假设V是数域F上的二次空间. 如果它的基$\alpha_1,\ ...,\ \alpha_n$中, 任何两个不同的基向量都互相正交, 我们就称这是一组正交基.

命题 9.2.2. 数域F上的n维二次空间V一定存在一组正交基.

证明 我们用数学归纳法. 假设命题对维数小于n的二次空间成立.

如果二次空间V是本身迷向的, 那么所有的向量都是迷向的. 所以对任何两个向量$\alpha,\ \beta \in V$都有

$$(\alpha,\ \beta)\frac{1}{2}((\alpha+\beta,\ \alpha+\beta) - (\alpha,\ \alpha) - (\beta,\ \beta)) = 0$$

所以任何一组基都是正交基. 否则的话, 一定会存在非迷向的非零向量$\alpha \in V$.我们考虑$L(\alpha)$的正交补W. 可以证明

$$V = L(\alpha) \bigoplus W$$

由归纳假设W有一组正交基$\alpha_1,\ ...,\ \alpha_{n-1}$. 所以$V$有一组正交基

$$\alpha,\ \alpha_1,\ ...,\ \alpha_{n-1}.$$

□

虽然数域F上的n维二次空间V一定存在一组正交基, 但是不一定能存在由非迷向向量组成的正交基, 这样的基存在当且仅当二次空间是非退化的. 如果二次空间V是非退化的, 那么任何一组正交基的基向量都是非迷向的. 在欧几里得空间那一章, 我们知道任何一组正交的向量都可以扩充成空间的一组正交基. 这个性质对一般的二次空间并不成立, 而且对非退化的二次空间V也不成立. 例如, 我们假设二维的二次空间在基$\alpha_1,\ \alpha_2$ 之下的度量矩阵是

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

则

$$Q(\alpha_1+\alpha_2)=0.$$

假设

$$(x\alpha_1+y\alpha_2,\ \alpha_1+\alpha_2)=0,$$

则$x=y$, 所以和$\alpha_1+\alpha_2$正交的向量一定和$\alpha_1+\alpha_2$线性相关, 因此$\alpha_1+\alpha_2$无法扩充成一组正交基.

这个命题等价于我们前面学过的对称矩阵合同于对角矩阵, 或者二次型可以用非退化的线性替换变成标准形.

类似的, 我们前面学习过任何一个反对称矩阵都合同于

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

这相当于说如果f是反对称双线性函数, 那么一定存在V 的一组基α_1, ..., α_n, 使得f在这组基下面的度量矩阵是上面的形式.

定义 9.2.7 (邻接基). 二次空间V的两个正交基

$$\alpha=(\alpha_1,\ ...,\ \alpha_n),\ \alpha'=(\alpha_1',\ ...,\ \alpha_n')$$

称为是邻接的如果存在i, j使得$\alpha_i=\alpha_j'$.

定理 9.2.3. 任给非退化二次空间V的两个正交基

$$\alpha=(\alpha_1,\ ...,\ \alpha_n),\ \alpha'=(\alpha_1',\ ...,\ \alpha_n'),$$

一定存在一个正交基序列

$$\alpha^{(0)}=\alpha,\ \alpha^{(1)},\ ...,\ \alpha^{m}=\alpha'$$

使得对任意的$0\le i\le m-1$, $\alpha^{(i)}$与$\alpha^{(i+1)}$是邻接的.

证明 (1) 假设$(\alpha_1,\ \alpha_1)(\alpha_1',\ \alpha_1') - (\alpha_1,\ \alpha_1')^2 \neq 0$. 这相当于说是$\alpha_1,\ \alpha_1'$张成了一个非退化的二维平面$P = L(\alpha_1,\ \alpha_1')$. 假设$\varepsilon_2 \in P$和$\alpha_1$正交, $\varepsilon_2' \in P$和α_1'正交. 设$\alpha_3'',\ ...,\ \alpha_n''$为$H$的正交补的正交基. 则

$$\alpha \longrightarrow (\alpha_1,\ \varepsilon_2, \alpha_3'',\ ...,\ \alpha_n'') \longrightarrow (\alpha_1',\ \varepsilon_2', \alpha_3'',\ ...,\ \alpha_n'') \longrightarrow \alpha'.$$

为要求的正交基序列.

(2) 假设$(\alpha_1,\ \alpha_1)(\alpha_2',\ \alpha_2') - (\alpha_1,\ \alpha_2')^2 \neq 0$. 证明类似.

(3) 假设$(\alpha_1,\ \alpha_1)(\alpha_1',\ \alpha_1') - (\alpha_1,\ \alpha_1')^2 = (\alpha_1,\ \alpha_1)(\alpha_2',\ \alpha_2') - (\alpha_1,\ \alpha_2')^2 = 0$.

则我们说一定存在$x \in F$使得$\alpha_x = \alpha_1' + x\alpha_2'$满足

$$(\alpha_x,\ \alpha_x) = (\alpha_1',\ \alpha_1') + x^2(\alpha_2',\ \alpha_2') \neq 0,$$

而且$L(\alpha_1,\ \alpha_x)$是个非退化的平面. 注意到空间是非退化的, α_2'已经和其他的基向量都正交了, 所以不可能和自己也正交. 因此$(\alpha_2',\ \alpha_2') \neq 0$. 域里的元素有无限多个, 所以满足$(\alpha_1',\ \alpha_1') + x^2(\alpha_2',\ \alpha_2') \neq 0$ 的x总是找得到的.

为了满足$L(\alpha_1,\ \alpha_x)$是个非退化的平面, 必须

$$(\alpha_1,\ \alpha_1)(\alpha_x,\ \alpha_x) - (\alpha_1,\ \alpha_x)^2 \neq 0.$$

将α_x的表达式带入上式的变成

$$-2x(\alpha_1, \alpha_1')(\alpha_1, \alpha_2') \neq 0.$$

由于我们第三种情况的假设可以知道(α_1, α_1'), (α_1, α_2')都不等于0, 所以满足这个条件的x总是有的, 只要不等于0即可.

假设$\alpha_x,\ \alpha_2''$是$L(\alpha_1',\ \alpha_2')$的一组正交基, 记

$$\alpha'' = (\alpha_x,\ \alpha_2'',\ \alpha_3',\ ...,\ \alpha_n')$$

为V的正交基. 则$L(\alpha_1,\ \alpha_x)$是一个非退化的平面.

所以我们可以通过邻接基序列把α和α''关联起来, 又因为α'和α''是邻接的, 所以α和α'能关联起来, 定理得证. □

§9.3 辛空间