> 25~26学年线性代数第三次辅学讲义
>
> 授课：鲁亦智

矩阵的运算和相抵标准形

1 矩阵的基本运算

1.1 加法和数乘

1.1.1 定义

教材 p123 定义4.3

设$A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{m\times n}$，$A,B\in M_{m\times n}(F)$，$\lambda\in F$，定义：

$$\begin{aligned}A+B&=(a_{ij}+b_{ij})_{m\times n}\\ \lambda A&=(\lambda a_{ij})_{m\times n}.\end{aligned}$$

矩阵的加法和数乘是按元素操作（element-wise）的。

1.1.2 性质

教材 p123

- 加法和数乘在$M_{m\times n}(F)$封闭
- 加法满足结合律，交换律
- 数乘满足$1A=A$，$\lambda(\mu A)=(\lambda\mu)A$，$(\lambda+\mu)A=\lambda A+\mu A$，$\lambda(A+B)=\lambda A+\lambda B$

以上性质均可由定义推出。

1.1.3 与线性映射的关系

教材 p122

设$\sigma,\tau\in L(V_1,V_2)$，取$V_1$的一组基$B_1$，$V_2$的一组基$B_2$，记$\sigma,\tau$关于$B_1$和$B_2$的矩阵为$M(\sigma),M(\tau)$，$\lambda\in F$，则

$$\begin{aligned}M(\sigma+\tau)&=M(\sigma)+M(\tau)\\ M(\lambda\sigma)&=\lambda M(\sigma).\end{aligned}$$

即矩阵的和（数乘）所对应的线性映射是矩阵所对应的线性映射的和（数乘）。

1.2 乘法

1.2.1 定义

LALU p191 定义7.4

设$A=(a_{ij})_{p\times m},B=(b_{ij})_{m\times n}$，我们定义$A$与$B$的乘积矩阵$C=AB=(c_{ij})_{p\times n}$是一个$p\times n$矩阵，其中它的第$i$行第$j$列元素为矩阵$A$的第$i$行与矩阵$B$的第$j$列对应位置元素相乘后求和的结果，即

$$c_{ij}=\sum_{k=1}^{m}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{im}b_{mj}\ \ (i=1,\ldots,p,\ \ j=1,\ldots,n).$$

1.2.2 性质

LALU p193

矩阵的乘法有如下基本性质：

- $(AB)C = A(BC)$（结合律）
- $\lambda(AB) = (\lambda A)B = A(\lambda B)$，$\lambda \in \mathbf{F}$
- $A(B+C) = AB + AC$（左分配律）
- $(B+C)P = BP + CP$（右分配律）

使用映射的结合律、线性性和分配律直接证明对应的几何版本的正确性，或者直接暴力设出矩阵元素然后暴力计算证明等号两边对应位置（如第i行第j列元素）相等也是可行的.

例题1

LALU p192 例7.3

设

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -3 \end{pmatrix}, B = \begin{pmatrix} 0 & 3 \\ 1 & 2 \\ 3 & 1 \end{pmatrix},$$

求AB和BA.

> 这个例子表明矩阵的乘法（大多数情况下）不能交换。

事实上矩阵乘法有很多和数的乘法重要的不同：

1. 矩阵乘法不一定满足交换律（即AB不一定等于BA，事实上随手写两个矩阵，很大的概率就是不交换的，甚至交换过来不可乘）． 因此实数的完全平方公式代入矩阵不一定成立，即很多时候 $(A+B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$；
2. 但是注意数量矩阵（即对角线上元素都相等，其余均为0，单位矩阵是其特例）和任何同阶的矩阵相乘都是可交换的，这一点在矩阵求幂时很有用；
3. $A \neq O$且$B \neq O$不能推出$AB \neq O$． 例如线性方程组$AX = 0$有非零解，若B的各列均为方程非零解，则 $AB = O$.
4. 消去律也不一定满足：即$AB = AC$不一定$B = C$． 原因在于$AB = AC \implies A(B-C) = O$，可知不一定$B = C$.

1.2.3 与线性映射的关系

设线性空间 $V_1(F), V_2(F), V_3(F)$ 的基分别为

$$B_1 = \{\varepsilon, \varepsilon_2, \ldots, \varepsilon_n\}, B_2 = \{\zeta_1, \zeta_2, \ldots, \zeta_m\}, B_3 = \{\eta_1, \eta_2, \ldots, \eta_p\}$$

$\sigma \in \mathcal{L}(V_1, V_2), \tau \in \mathcal{L}(V_2, V_3)$，且$\sigma, \tau$ 分别关于基 B_1 和 B_2 及基 B_2 和 B_3 的矩阵为$B = (b_{i,j})_{m\times n}$和 $A = (a_{ij})_{p\times m}$，即：

$$M(\tau) = A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix},$$

$$M(\sigma)=B=\begin{pmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ b_{m1} & b_{m2} & \cdots & b_{mn}\end{pmatrix}.$$

则$\tau\sigma\in\mathcal{L}(V_1,V_3)$关于基 B_1 和 B_3的矩阵$C=(c_{ij})_{p\times n}$中第j列元素$c_{1j},c_{2j},\ldots,c_{pj}$是 $\tau\sigma(\varepsilon_j)$在基$B_3$ 下的坐标. 于是有:

$$\begin{aligned}\tau\sigma(\varepsilon_j)&=\tau(\sigma(\varepsilon_j))\\ &=\tau\left(\sum_{k=1}^{m}b_{kj}\zeta_k\right)=\sum_{k=1}^{m}b_{kj}\tau(\zeta_k)=\sum_{k=1}^{m}b_{kj}\left(\sum_{i=1}^{p}a_{ik}\eta_i\right)=\sum_{i=1}^{p}\left(\sum_{k=1}^{m}a_{ik}b_{kj}\right)\eta_i\end{aligned}$$

即得:$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{im}b_{mj}$. 从而我们从一般的线性映射出发,证明了表示矩阵的乘积等于线性映射复合的表示矩阵. 如果用 $\mathbf{M}_{B_1,B_2}(\sigma)$ 表示 σ 在基 B_1 和 B_2 的下的矩阵表示,那么我们的推导得到了如下结论:

$$\mathbf{M}_{B_2,B_3}(\tau)\mathbf{M}_{B_1,B_2}(\sigma)=\mathbf{M}_{B_1,B_3}(\tau\sigma)$$

例题2

LALU p193 例7.4

考虑先旋转θ_1,然后旋转θ_2对应的两个变换σ_1,σ_2的复合$\sigma_2\sigma_1$,实际上就是旋转$\theta_1+\theta_2$角度,其矩阵表示为$M_{\theta_1+\theta_2}$,而矩阵乘法

$$\begin{aligned}M_{\theta_2}M_{\theta_1}&=\begin{pmatrix}\cos\theta_2 & -\sin\theta_2\\ \sin\theta_2 & \cos\theta_2\end{pmatrix}\begin{pmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{pmatrix}\\ &=\begin{pmatrix}\cos\theta_2\cos\theta_1-\sin\theta_2\sin\theta_1 & -\cos\theta_2\sin\theta_1-\sin\theta_2\cos\theta_1\\ \sin\theta_2\cos\theta_1+\cos\theta_2\sin\theta_1 & -\sin\theta_2\sin\theta_1+\cos\theta_2\cos\theta_1\end{pmatrix}\\ &=\begin{pmatrix}\cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2)\\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2)\end{pmatrix}=M_{\theta_1+\theta_2},\end{aligned}$$

这表明矩阵乘法$M_{\theta_2}M_{\theta_1}$的结果确实与$\sigma_2\sigma_1$的矩阵表示$M_{\theta_1+\theta_2}$一致.

例题3

LALU p199 例7.8

设$A=\begin{pmatrix}1 & 0 & 2\\ -1 & 2 & 1\\ 1 & 2 & 5\end{pmatrix}$为两个三维线性空间之间的线性映射$\sigma$对应的矩阵,求$\sigma$的像空间和核空间.

例题4

LALU p199 例7.9

已知3阶矩阵$A=\begin{pmatrix}1 & 0 & 1\\ 0 & -1 & 0\\ -1 & 1 & -1\end{pmatrix}$. 定义$\mathbf{R}^{3\times3}$上的线性变换$\sigma(X)=AX,\quad X\in\mathbf{R}^{3\times3}$. 求$\sigma$的像和核.

1.2.4 矩阵多项式

我们在线性空间中已经介绍过，一般用$\mathbf{F}[x]_{m+1}$表示数域$\mathbf{F}$上的次数最高为m的多项式全体，其中的元素记为

$$p(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0,\ \ a_i\in\mathbf{F}\ \ (i=1,2,\ldots,m)$$

我们发现这里的自变量不一定需要是一个数，也可以是线性变换或者方阵，因为只要可以和自己相乘就能定义乘方.

例如线性映射$\sigma:V\to V$构成的m次多项式可以记为

$$p(\sigma)=a_m\sigma^m+a_{m-1}\sigma^{m-1}+\cdots+a_1\sigma+a_0I$$

其中σ^i表示σ复合i次，I表示恒等映射.

我们很容易说明当σ在V的一组基下矩阵表示为A时，$p(\sigma)$在同一组基下的矩阵表示为

$$p(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E,$$

其中E表示单位矩阵.

由此我们便得到了矩阵多项式的定义，有如下几点需要强调：

1. 这里我们要求σ是线性变换（即出发空间和到达空间一致），因为只有满足这一条件才能复合. 对于矩阵而言，其作用在标准 n 维向量空间 $\mathbf{F}^n$ 上，所以矩阵可求幂即要求出发空间和到达空间维数相同即可，这样才能保证矩阵的幂次可以定义（即A和A可乘，因此A的行列数一致）；
2. 上面的定义隐含：$\sigma^0=I,A^0=E$;
3. $A^kA^m=A^{k+m}$，$(A^k)^m=A^{km}$，其中A为方阵，k,m为任意正整数. 当 A 可逆时这一式可以拓展到全体整数，负整数对应于逆矩阵的情况，接下来可逆的部分会作进一步解释.

例题5

LALU p198 例7.7

设$f(x),g(x)\in\mathbf{F}[x],\ \ A,B\in\mathbf{M}_n(\mathbf{F})$.
证明：

1. $f(A)g(A)=g(A)f(A)$;
2. 如果$AB=BA$，则$f(A)g(B)=g(B)f(A)$;

1.2.5 矩阵可交换问题

一般来说此类问题直接设可交换矩阵的每一个元素都是未知数即可.

例题6

LALU p255 例9.2

可交换矩阵1 求所有与A可交换的矩阵，其中

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

更一般地，我们有如下结论：

1. 与主对角元两两互异的对角矩阵可交换的方阵只能是对角矩阵；
2. 准对角矩阵A每个对角分块为对角矩阵且每个分块内对角线元素相同，但不同对角块之间不同，则与A可交换

的矩阵只能是准对角矩阵；

3. 与所有n级可逆矩阵可交换的矩阵为数量矩阵；
4. 与所有n级矩阵可交换的矩阵为数量矩阵.

1.3 逆

1.3.1 定义

LALU p202 定义7.5

设 $A \in \mathbf{M}_n(\mathbf{F})$. 若存在 $B \in \mathbf{M}_n(\mathbf{F})$ 使得 $AB = BA = E_n$（不刻意强调时可以省略 n），则称矩阵 A 可逆，并把 B 称为 A 的，记作 $B = A^{-1}$.

> 可逆矩阵一定是方阵。

由定义可推出如下事实：

- 可逆矩阵A的逆矩阵唯一.
- 设$A, B \in \mathbf{M}_n(\mathbf{F})$，则$AB = E \iff A$与$B$互为逆矩阵.

1.3.2 性质

1. 主对角元都是非零数的对角矩阵一定可逆，且逆矩阵就是对角线上元素取倒数（单位矩阵即为特例，其逆矩阵是其自身）；
2. 注意没有加法性质（例如A可逆（则$-A$也可逆），但$A+(-A)=O$不可逆），对于数乘有$(\lambda A)^{-1} = \lambda^{-1}A^{-1}$；
3. $(AB)^{-1} = B^{-1}A^{-1}$，$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$；以上几个性质证明只需要直接验证结果即可，即因为$ABB^{-1}A^{-1} = AA^{-1} = E$，所以根据逆的唯一性可知$(AB)^{-1} = B^{-1}A^{-1}$一定成立；
4. $(A^k)^{-1} = (A^{-1})^k$，$A^kA^m = A^{k+m}$，$(A^k)^m = A^{km}$；注意这里的k和m不一定需要非负，事实上负数就是逆矩阵的幂次或幂次的逆，如$A^{-2} = (A^{-1})^2 = (A^2)^{-1}$；
5. 若 A 可逆，则消去律成立，即 $AB = AC \implies B = C$ 成立，我们只需在$AB = AC$ 的等式两边同时左乘 A^{-1} 即可证明（$BA = CA$的情况也是成立的，只需要等式两边同时右乘 A^{-1} 即可证明）.这个结论的一个显然的推论是，若 A 可逆且 $AB = O$（或$BA = O$）可以推出 $B = O$（令 $C = O$ 即可）.更进一步地，回忆在不可逆矩阵的情况下，即使 $A \neq O$ 且$B \neq O$，我们也可能有 $AB = O$，但当 A（或B）可逆时，根据前面的结论可知 B（或A）必然为零矩阵，因此不可能存在这样的情况.

1.3.3 与线性映射的关系

现在我们已经知道，矩阵加法、数乘以及乘法的定义来源于线性映射的加法、数乘以及复合，所以在讨论矩阵的逆的时候，我们自然会想到线性映射的逆.

回忆线性映射的逆的定义，令 $\sigma \in \mathcal{L}(V_1, V_2)$为可逆映射，若 $\tau \in \mathcal{L}(V_2, V_1)$ 使得$\sigma\tau = I_{V_2}$ 且 $\tau\sigma = I_{V_1}$，则称 τ 为σ 的逆映射. 其中 I_{V_1} 和 I_{V_2} 分别是 V_1 和 V_2上的恒等映射. 需要注意的是，我们知道可逆等价于同构，因此我们要求 V_1 和 V_2的维数相同，设为 n.

我们知道线性映射的复合对应矩阵乘法，取 V_1 的一组基 B_1 和 V_2的一组基 B_2，σ 关于 B_1, B_2 的矩阵为 $A = \mathbf{M}_{B_1,B_2}(\sigma)$，$\tau$ 关于 B_2, B_1 的矩阵为$B = \mathbf{M}_{B_2,B_1}(\tau)$，则 A 和 B 均为 n 阶方阵. 回顾，我们有

$$\begin{aligned}\mathbf{M}_{B_2,B_2}(\sigma\tau) = \mathbf{M}_{B_1,B_2}(\sigma)\mathbf{M}_{B_2,B_1}(\tau) = AB,\\ \mathbf{M}_{B_1,B_1}(\tau\sigma) = \mathbf{M}_{B_2,B_1}(\tau)\mathbf{M}_{B_1,B_2}(\sigma) = BA.\end{aligned}$$

不难验证恒等映射在出发空间和到达空间取同一组基下的矩阵表示一定为单位矩阵，故

$$AB = \mathbf{M}_{B_1,B_1}(I_{V_1}) = E_n = \mathbf{M}_{B_2,B_2}(I_{V_2}) = BA.$$

例题7

LALU p260 例9.7

设A为非零矩阵，且$A^3 = O$，证明：$E + A$和$E - A$都可逆.

例题8

LALU p260 例9.8

若X, Y是两个列向量，且$X^{\mathrm{T}}Y = 2$，证明：

1. $(XY^{\mathrm{T}})^k = 2^{k-1}(XY^{\mathrm{T}})$;
2. 如果$A = E + XY^{\mathrm{T}}$，则A可逆，并求其逆矩阵.

例题9

LALU p261 例9.9

设 A, B 分别是 $n \times m$ 和 $m \times n$ 的矩阵，且 $E_n \pm AB$ 可逆，则 $E_m \pm BA$ 可逆.

例题10

LALU p264 例9.14

设A是n阶方阵，且$E - A$，$E + A$和A都可逆，证明：$(E - A^{-1})^{-1} + (E - A)^{-1} = E$.

1.3.4 可逆的等价条件

LALU p224 定理8.3

设$A \in \mathbf{M}_n(\mathbf{F})$，则下列命题等价：

1. A可逆;
2. $r(A) = n$;
3. A的n个行（列）向量线性无关;
4. 齐次线性方程组$AX = 0$只有零解.

例题11

LALU p225 例8.1

设 $A = (a_{ij})_{n\times n}$ 是一个 n 阶矩阵，若 $|a_{ii}| > \sum\limits_{j\neq i} |a_{ij}|$（即满足对角占优条件），证明 A 是可逆矩阵.

1.4 转置

1.4.1 定义

LALU p20 定义1.12

设$A = (a_{ij})_{m\times n}$，则A的转置矩阵是一个$n \times m$矩阵，记作A^T，它的第k行正好是A的第k列（$k = 1, 2, \dots, n$）；它的第r列正好是A的第r行（$r = 1, 2, \dots, m$）.

1.4.2 性质

1. $(A^{\mathrm{T}})^{\mathrm{T}}=A$
2. $(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}$
3. $(\lambda A)^{\mathrm{T}}=\lambda A^{\mathrm{T}},\ \ \lambda\in\mathbf{F}$
4. $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}},\ \ (A_1A_2\cdots A_n)^{\mathrm{T}}=A_n^{\mathrm{T}}\cdots A_2^{\mathrm{T}}A_1^{\mathrm{T}},\ \ (A^{\mathrm{T}})^m=(A^m)^{\mathrm{T}}$
5. $(A^{\mathrm{T}})^{-1}=(A^{-1})^{\mathrm{T}}$

关于上述性质我们有如下说明：

- 性质1从计算角度来看是显然的，简而言之就是矩阵第i行变成第i列后又变回了第i行，因此矩阵不变；
- 性质2~4考虑从计算角度验证只需暴力计算即可，至于4的n个矩阵的情况只需要从两个相乘的情况出发数学归纳即可，最后的幂的性质实际上将$(A_1A_2\cdots A_n)^{\mathrm{T}}=A_n^{\mathrm{T}}\cdots A_2^{\mathrm{T}}A_1^{\mathrm{T}}$中的$A_i$全部取成$A$即可；
- 性质5请不要忘记验证逆的运算性质的一般方法，我们只需要看到$(A^{-1})^{\mathrm{T}}A^{\mathrm{T}}=(AA^{-1})^{\mathrm{T}}=E$，这里第一个等号运用了上面第4点转置乘法的性质. 从这一式中我们看到$(A^{-1})^{\mathrm{T}}$是A^{T}的逆矩阵，因此利用逆的唯一性即可得到$(A^{\mathrm{T}})^{-1}=(A^{-1})^{\mathrm{T}}$；

例题12

LALU p206 例7.11

设$\alpha=(1,-1,2)^{\mathrm{T}},\ \ \beta=(3,1,-2)^{\mathrm{T}},\ \ A=\alpha\beta^{\mathrm{T}}$，求$A^n$.

> 事实上利用秩为1的矩阵是计算矩阵的幂次时常用的一个技巧。

例题13

LALU p209 例7.14

a,b,c,d是四个实数. 证明$\begin{cases}a^2+b^2=1\\c^2+d^2=1\\ac+bd=0\end{cases}$成立的充分必要条件是$\begin{cases}a^2+c^2=1\\b^2+d^2=1\\ab+cd=0\end{cases}$.

1.4.3 对称矩阵和反对称矩阵

设$A=(a_{ij})_{n\times n}$，如果$\forall i,j\in\{1,2,\ldots,n\}$均有$a_{ij}=a_{ji}$，则称$A$为对称矩阵. 若均有$a_{ij}=-a_{ji}$，则称$A$为反对称矩阵.

它们具有如下性质：

1. 反对称矩阵主对角元均为 0；
2. AA^{T} 和 $A^{\mathrm{T}}A$ 均为对称矩阵；
3. 设 A,B 分别为 n 阶对称和反对称矩阵，则 $AB+BA$ 是反对称矩阵；
4. 对称矩阵的乘积不一定对称；
5. 可逆的对称（反对称）矩阵的逆矩阵也是对称（反对称）矩阵.

例题14

LALU p208 例7.13

数域$\mathbf{F}$上所有n阶方阵组成的线性空间$V=\mathbf{M}_n(\mathbf{F})$，$V_1$表示所有对称矩阵组成的集合，$V_2$表示所有反对称矩阵组成的集合. 证明：$V_1,V_2$都是$V$的子空间，且$V=V_1\oplus V_2$.

1.5 特殊矩阵

1.5.1 对角矩阵

我们一般记主对角矩阵为 $diag(d_1,d_2,\ldots,d_n)$，准对角矩阵（对角线由方阵构成，其余元素为0）为 $diag(A_1,A_2,\ldots,A_n)$.

设A是一个$s\times n$矩阵，把A写成列向量与行向量的形式，分别为

$$A=\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}=\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_s\end{pmatrix}$$

则

$$\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{pmatrix}=\begin{pmatrix}d_1\alpha_1 & d_2\alpha_2 & \cdots & d_n\alpha_n\end{pmatrix}$$
$$\begin{pmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_s\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_s\end{pmatrix}=\begin{pmatrix}d_1\beta_1\\ d_2\beta_2\\ \vdots\\ d_s\beta_s\end{pmatrix}$$

即A右乘对角矩阵 $diag(d_1,d_2,\ldots,d_n)$ 相当于给 A 的第$j(j=1,\ldots,n)$ 列元素都乘以 d_j，A 左乘对角矩阵 $diag(d_1,d_2,\ldots,d_s)$ 相当于给 A 的第 $i(i=1,2,\ldots,s)$行元素都乘以 d_i.

对角矩阵具有如下性质：

1. 对角矩阵$diag(d_1,d_2,\ldots,d_n)$可逆当且仅当对角线上元素均不为0，且此时逆矩阵为 $diag(d_1^{-1},d_2^{-1},\ldots,d_n^{-1})$.
2. 分块对角矩阵$diag(A_1,A_2,\ldots,A_n)$可逆当且仅当每个分块A_i可逆，且此时逆矩阵为 $diag(A_1^{-1},A_2^{-1},\ldots,A_n^{-1})$.
3. 两个对角矩阵$A=diag(a_1,a_2,\ldots,a_n)$，$B=diag(b_1,b_2,\ldots,b_n)$的乘积仍然是对角矩阵，且 $AB=diag(a_1b_1,a_2b_2,\ldots,a_nb_n)$.

   对于乘方运算，有$A^k=diag(a_1^k,a_2^k,\ldots,a_n^k)$.
4. 两个准对角矩阵$A=diag(A_1,A_2,\ldots,A_n)$，$B=diag(B_1,B_2,\ldots,B_n)$中$A_i$和$B_i$是同级方阵，则乘积仍然是准对角矩阵，且$AB=diag(A_1B_1,A_2B_2,\ldots,A_nB_n)$.

以上性质可以通过简单计算验证。

1.5.2 上（下）三角矩阵

已知A,B都是上三角矩阵，且设A的主对角元素分别为$a_{11},\ldots,a_{nn}$，B的主对角元素分别为$b_{11},\ldots,b_{nn}$，则

1. $A^{\mathrm{T}},B^{\mathrm{T}}$都是下三角矩阵；
2. AB仍然是上三角矩阵，且AB的主对角元素为$a_{11}b_{11},\ldots,a_{nn}b_{nn}$；
3. A可逆的充要条件是其主对角元均不为0，且A可逆时，A^{-1}也是上三角矩阵，并且A^{-1}的主对角元素分别为

$a_{11}^{-1}, \ldots, a_{nn}^{-1}$.

例题15

LALU p253 例9.1

已知$A_1, \ldots, A_n$是n个对角元都为0的上三角矩阵，证明：$A_1 A_2 \cdots A_n = O$.

1.5.3 自然基和基本矩阵

回顾 $\mathbf{R}^n$ 中的自然基 $e_1, e_2, \ldots, e_n$，我们知道 e_i 是一个 n 维列向量，其第 i 个元素为 1，其余元素为 0.

自然基的运算性质：

1. $e_i^{\mathrm{T}} e_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$;
2. 若 A 是 $m \times n$ 矩阵，$\mathbf{R}^n$ 中的自然基记为 $e_1, e_2, \ldots, e_n$，$\mathbf{R}^m$ 中的自然基记为 $f_1, f_2, \ldots, f_m$，则
    1. Ae_j 是 A 的第 j 列；
    2. $f_i^{\mathrm{T}} A$ 是 A 的第 i 行；
    3. $f_i^{\mathrm{T}} A e_j$ 是 A 的第 i 行第 j 列元素.

接下来考虑只有一个元素为 1，其余元素全为 0 的矩阵，即基本矩阵的运算性质. 第i行第j列元素为1的基本矩阵记为 E_{ij}，基本矩阵计算具有如下性质（接下来的 A 均表示 n 阶矩阵$(a_{ij})_{n\times n}$）：

1. $A = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} a_{ij} E_{ij}$;
2. $E_{ik} E_{lj} = \begin{cases} E_{ij} & k = l \\ O & k \neq l \end{cases}$;
3. AE_{ij} 的结果就是把 A 的第 i 列移到第 j 列的位置，其余元素都为 0 的矩阵；
4. $E_{ij}A$ 的结果就是把 A 的第 j 行移到第 i 行的位置，其余元素都为 0 的矩阵；
5. $E_{ij} A E_{kl} = a_{jk} E_{il}$.

1.6 矩阵的幂

1.6.1 找规律

例题16

LALU p264 例9.15

计算$(PAQ)^k$，其中

$$P = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \ A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \ Q = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}.$$

例题17

LALU p265 例9.17

证明如下两个矩阵的幂的性质：

1. 设矩阵 A 为 n 阶基础循环矩阵，即 $A=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\1&0&0&\cdots&0\end{pmatrix}$，则 $A^k=\begin{pmatrix}O&E_{n-k}\\E_k&O\end{pmatrix}$，其中 $1\leqslant k\leqslant n$.
2. 设矩阵 A 为上（下）三角矩阵且主对角线元素全为 0，则 $A^n=O$.

   特别地，若 A 为 n 阶若当块矩阵，即 $A=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}$，则 $A^k=\begin{pmatrix}O&E_{n-k}\\O&O\end{pmatrix}$，其中 $1\leqslant k\leqslant n$.

例题18

LALU p267 例9.18

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

1.6.2 数学归纳法

例题19

LALU p268 例9.19

求$A=\begin{pmatrix}\cos\alpha&\sin\alpha\\-\sin\alpha&\cos\alpha\end{pmatrix}^n$.

1.6.3 秩1矩阵

已知M是秩为 1 的矩阵，记$tr(M)=b$，下面讨论$(aE+M)^n$的计算结果.

> $tr(M)$表示M的对角线元素之和，称为M的迹。

事实上，$M^k=b^{k-1}M$，因此

1. 当$b=0$时，$M^k=O(k\geqslant 2)$，因此 $(aE+M)^n=a^nE+na^{n-1}M$.
2. 当$b\neq 0$时有

$$
\begin{aligned}
(aE+M)^n &= \sum_{k=0}^{n} C_n^k a^{n-k} M^k \\
&= \sum_{k=0}^{n} C_n^k a^{n-k} b^{k-1} M \\
&= a^n E + \frac{1}{b}\sum_{k=1}^{n} C_n^k a^{n-k} b^k M \\
&= a^n E + \frac{1}{b}\left(\sum_{k=0}^{n} C_n^k a^{n-k} b^k - a^n\right) M \\
&= a^n E + \frac{(a+b)^n - a^n}{b} M
\end{aligned}
$$

例题20

LALU p270 例9.22

求矩阵的幂$A^n = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}^n$.

例题21

LALU p270 例9.23

已知A是数域P上的一个2阶方阵，且存在正整数l使得$A^l = O$，证明：$A^2 = O$.

1.6.4 利用初等矩阵性质

例题22

LULA p271 例9.24

设A为三阶矩阵，P为三阶可逆矩阵，$P^{-1}AP = B$，其中

$$
P = \begin{pmatrix} 0 & 2 & -1 \\ 1 & 1 & 2 \\ -1 & -1 & -1 \end{pmatrix},\ B = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix},
$$

求A^{2024}.

分块矩阵和打洞法

2 矩阵的秩与相抵标准形

2.1 秩，行秩，列秩

2.1.1 定义和基本性质

设 A 是线性映射 σ 对应的矩阵，定义矩阵 A 的秩为 $r(A) = r(\sigma)$. 此外，将矩阵A的所有行向量组成的秩称为A的行秩，常记为r_{r}；所有列向量组成的向量组的秩称为A的列秩，常记为r_{c}.

我们有如下重要推论：

- *LALU p222 定理8.1* 任意矩阵 $A = (a_{ij})_{m\times n}$ 的秩 = 行秩 = 列秩（进而有$r(A) = r(A^T)$）.

- *LALU p224 定理8.2* 线性映射是单射当且仅当其矩阵表示为列满秩矩阵，线性映射是满射当且仅当其矩阵表示为行满秩矩阵.

2.1.2 秩等式和不等式

1. $r(A)=r(PA)=r(AQ)=r(PAQ)$，其中P,Q可逆.
2. $|r(A)-r(B)|\leqslant r(A\pm B)\leqslant r(A)+r(B)$.
3. $r(AB)\leqslant\min\{r(A),r(B)\}$.
4. $r(A)=r(A^{\mathrm{T}})=r(AA^{\mathrm{T}})=r(A^{\mathrm{T}}A)$.
5. （Sylvester不等式）$A\in\mathbf{F}^{s\times n}$，$B\in\mathbf{F}^{n\times m}$，则$r(AB)\geqslant r(A)+r(B)-n$.
6. （Frobenius不等式）$r(ABC)\geqslant r(AB)+r(BC)-r(B)$.

例题23

LALU p282 例9.31

证明以下矩阵的秩不等式：

1. $r\begin{pmatrix}A & O\\ O & B\end{pmatrix}=r(A)+r(B)$.
2. $r(A)+r(B)+r(D)\geqslant r\begin{pmatrix}A & D\\ O & B\end{pmatrix}\geqslant r(A)+r(B)$，$r(A)+r(B)+r(C)\geqslant r\begin{pmatrix}A & O\\ C & B\end{pmatrix}\geqslant r(A)+r(B)$
   .

例题24

LALU p286 例9.32

若A,B为两个n阶矩阵，则

A. $r(A,B)=r(A^{\mathrm{T}},B^{\mathrm{T}})$

B. $r(A,AB)=r(A)$

C. $r(A,B)=\max\{r(A),r(B)\}$

D. $r(A,BA)=r(A)$

2.2 初等变换和初等矩阵

2.2.1 定义

我们首先给出三种初等矩阵的定义，其与高斯-若当消元法中的三类初等变换对应.

将单位矩阵E做一次初等变换得到的矩阵称为初等矩阵，与三种初等行、列变换对应的三类初等矩阵为：

1. 将单位矩阵第i行（或列）乘c，得到初等倍乘矩阵$E_i(c)$；
2. 将单位矩阵第i行乘c加到第j行，或将第j列乘c加到第i列，得到初等倍加矩阵$E_{ij}(c)$；
3. 将单位矩阵第i,j行（或列）对换，得到初等对换矩阵E_{ij}.

三种初等矩阵具体形状如下：

$$E_i(c)=\begin{bmatrix}1&&&&\\&\ddots&&&\\&&c&&\\&&&\ddots&\\&&&&1\end{bmatrix}\begin{array}{l}\\ \\ \text{第}i\text{行}\\ \\ \\ \end{array}$$

$$E_{ij}(c)=\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&\ddots&&&\\&&c&&1&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}\begin{array}{l}\\ \\ \\ \text{第}i\text{行}\\ \text{第}j\text{行}\\ \\ \\ \end{array}$$

$$E_{ij}=\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&0&&1&&\\&&&\ddots&&&\\&&1&&0&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}\begin{array}{l}\\ \\ \\ \text{第}i\text{行}\\ \text{第}j\text{行}\\ \\ \\ \end{array}$$

事实上，初等矩阵的定义就将我们在高斯-若当消元法中使用的初等变换的三种形式对应到了矩阵的形式上. 简而言之，对单位矩阵E做了一次初等变换后得到的矩阵P，乘以其他任何矩阵A的效果就是对A做了和对E做的同样的初等变换.

当我们对矩阵左乘一个初等矩阵时，相当于对矩阵做了对应的初等行变换；右乘一个初等矩阵时，相当于对矩阵做了对应的初等列变换.
所以在高斯-若当消元法中，假设系数矩阵为A，简化阶梯矩阵为U，我们做的初等行变换分别为$P_1, P_2, \ldots, P_k$，则有

$$P_k P_{k-1} \cdots P_2 P_1 A = U.$$

有如下几点细节需要注意：

1. 倍加变化请注意i和j在行列变换的情况下的不同，行变换是第i行乘c加到第j行，列变换是第j列乘c加到第i列；
2. 注意三类矩阵不是三个矩阵，例如倍乘矩阵乘以哪一行/哪一列，以及乘以多少都是不唯一的；
3. 三种初等矩阵都是可逆的，且$E_i^{-1}(c) = E_i(1/c)$，$E_{ij}^{-1}(c) = E_{ij}(-c)$，$E_{ij}^{-1} = E_{ij}$.
   原因非常简单，只需要记住这三类矩阵在单位矩阵基础上做了什么，需要反过来作用什么来抵消就可以理解；
4. 三种初等矩阵的转置：$E_i^{\mathrm{T}}(c) = E_i(c)$，$E_{ij}^{\mathrm{T}}(c) = E_{ji}(c)$，$E_{ij}^{\mathrm{T}} = E_{ij}$，因此初等矩阵转置前后分别对应于同样的行列变换操作.
   例如倍乘行变换表示对第i行乘以c，转置后如果视为列变换则表示对第i列乘以c，行列操作一致，对换也是如此.

而倍加$E_{ij}(c)$在行变换表示将第i行乘c加到第j行，转置后如果视为列变换，$E_{ji}(c)$表示将第i列乘c加到第j列，这两者在行列操作上保持了一致性.

总结而言就是P为初等矩阵，则PA和AP^{T}表示的变换除了行/列的名字外同名.

2.2.2 性质

- *LALU p236 定理8.10* 任意可逆矩阵都可以被表示为若干个初等矩阵的乘积.
- *LALU p237 引理8.1* 设A为n阶可逆矩阵，如果对A和n阶单位矩阵E做相同的初等行变换，即$P_1, P_2, \dots, P_k$后A变为E时，E变为A^{-1}.
- *LALU p238 定理8.11* 初等变换不改变矩阵的秩（包括行变换和列变换）.

例题25

LALU p237 例8.5

求矩阵$A = \begin{pmatrix} 0 & 2 & -1 \\ 1 & 1 & 2 \\ -1 & -1 & -1 \end{pmatrix}$的逆矩阵.

2.3 相抵标准形

2.3.1 定义

LALU p231 定理8.9

设 A 是 $m \times n$ 矩阵，则 $r(A) = r$ 的充要条件为存在可逆矩阵 P和 Q，使得

$$PAQ = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} = U_r$$

其中 E_r 表示 r 阶单位矩阵.

设 A 和 B 是 $m \times n$ 矩阵，如果存在可逆矩阵 P 和 Q，使得 $PAQ = B$，则称 A 和 B 是**相抵**的. 称 $PAQ = U_r$ 中的 U_r 为矩阵 A 的**相抵标准形**，其中 E_r 表示 r 阶单位矩阵，$r = r(A)$.

根据初等变换不改变矩阵的秩的性质，我们有如下等价定义：

LALU p239 定义8.4

我们称两个矩阵相抵当且仅当两个矩阵可以通过一系列初等变换互相转化.

2.3.2 性质

1. 任何矩阵都对应一个相抵标准形，并且所有形状相等（即行列数相等）且秩相等的矩阵有相同的相抵标准形，即$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$.
2. 矩阵A与B相抵 $\Longleftrightarrow$ 存在可逆矩阵P和Q使得$PAQ = B$.
3. 矩阵A与B相抵 $\Longleftrightarrow$ $r(A) = r(B)$.
4. 相抵是矩阵的一个等价关系.

例题26

LALU p240 例8.6

设$A=\begin{pmatrix}1 & 0 & 2 & -4\\ 2 & 1 & 3 & -6\\ -1 & -1 & -1 & 2\end{pmatrix}$. 求

1. A的秩r和相抵标准形;
2. 3 阶可逆矩阵P和 4 阶可逆矩阵Q使得$PAQ=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}$.

例题27

LALU p243 例8.8

设A为n阶方阵，证明:

1. 存在可逆矩阵B和幂等矩阵C（即满足$C^2=C$）使得$A=BC$;
2. 存在对称矩阵B和可逆矩阵C使得$A=BC$.

2.4 基变换

2.4.1 过渡矩阵

设 $B_1=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $B_2=\{\beta_1,\beta_2,\ldots,\beta_n\}$ 是线性空间 $V(\mathbf{F})$ 的任意两组基. 若存在矩阵 A 使得 $B_2=B_1A$，则称矩阵 A 为 B_1 变为基 B_2 的**过渡矩阵**.

根据 $B_2=B_1A$ 的展开式，B_1 变为基 B_2 的过渡矩阵 A 就是将B_2 中的向量在 B_1 下的坐标按列排列. 关于这一定义，我们有以下几点需要强调:

1. 在之后的讨论或者题目中需要特别注意说的是 B_1 变为基 B_2的过渡矩阵还是反过来基 B_2 变为基 B_1 的过渡矩阵;
2. 注意过渡矩阵一定是基与基之间的表示矩阵，一般的向量组之间不称过渡矩阵. 当然同一线性空间中任意两组长度相等的线性无关向量组之间都可以有过渡矩阵，这是因为我们可以将这两组向量组看作是它们张成的空间的两组基;
3. 根据定义，不难验证 B_1 变为基 B_2 的过渡矩阵实际上就是恒等映射I 在两组基 B_2 和基 B_1 下的矩阵表示 $\mathbf{M}_{B_2,B_1}(I)$. 根据矩阵乘法的定义，$\mathbf{M}_{B_2,B_1}(I)\mathbf{M}_{B_1,B_2}(I)=\mathbf{M}_{B_1,B_1}(I)=E$，即 B_1 变为基 B_2 的过渡矩阵可逆，并且逆矩阵就是 B_2 变为基 B_1的过渡矩阵.

2.4.2 换基公式

我们有如下性质:

1. *LALU p227 定理8.4* 设 $S_1=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是线性无关的向量组，且
$$(\beta_1,\beta_2,\ldots,\beta_s)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$
则向量组 $S_2=\{\beta_1,\beta_2,\ldots,\beta_s\}$ 的秩等于矩阵 A的秩.
2. *LALU p228 定理8.5* 若线性空间 V 中的两个向量组 S_1 和 S_2 满足 $S_2=S_1A$，其中 A 可逆，则 S_1 与 S_2 是等价向量组.
3. *LALU p228 定理8.6* 设线性空间V的两组基为B_1和B_2，且基B_1到B_2的变换矩阵（过渡矩阵）为A，如果 $\xi\in V(\mathbf{F})$在B_1和B_2下的坐标分别为X和Y，则$Y=A^{-1}X$.
4. *LALU p229 定理8.7* 设 $\sigma\in\mathcal{L}(V_1,V_2)$，$B_1,B_1'$ 是 V_1 的两组基，B_2,B_2' 是 V_2 的两组基. 设 P 是 B_1 变为 B_1' 的过渡矩阵，Q 是 B_2 变为 B_2' 的过渡矩阵，则 $\mathbf{M}_{B_1',B_2'}(\sigma)=Q^{-1}\mathbf{M}_{B_1,B_2}(\sigma)P$.
5. *LALU p229 定理8.8* 设线性变换$\sigma\in\mathcal{L}(V,V)$，$B_1=\{\alpha_1,\ldots,\alpha_n\}$和$B_2=\{\beta_1,\ldots,\beta_n\}$是线性空间

$V(\mathbf{F})$的两组基，基B_1变为基B_2的过渡矩阵为C. 如果σ在基B_1下的矩阵为A，则σ关于基B_2所对应的矩阵为$C^{-1}AC$.

例题28

LALU p228 例8.2

已知$\beta_1 = \alpha_2 + \alpha_3$，$\beta_2 = \alpha_1 + \alpha_3$，$\beta_3 = \alpha_1 + \alpha_2$，证明$\alpha_1, \alpha_2, \alpha_3$与$\beta_1, \beta_2, \beta_3$等价.

例题29

LALU p230 例8.3

已知三维线性空间 V 的线性变换 σ 关于基$\alpha_1, \alpha_2, \alpha_3$ 所对应的矩阵为

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

1. 求 σ 在基 $\beta_1, \beta_2, \beta_3$ 下对应的矩阵 B，其中:

$$\beta_1 = 2\alpha_1 + \alpha_2 + 3\alpha_3,\ \ \beta_2 = \alpha_1 + \alpha_2 + 2\alpha_3,\ \ \beta_3 = -\alpha_1 + \alpha_2 + \alpha_3;$$

2. 求 σ 的值域 $\sigma(V)$ 和核 $\ker \sigma$;
3. 把 $\sigma(V)$ 的基扩充为 V 的基，并求 σ 在这组基下对应的矩阵;
4. 把 $\ker \sigma$ 的基扩充为 V 的基，并求 σ 在这组基下对应的矩阵.

例题30

LALU p242 例8.7

设$\sigma : V \to W$为线性映射，取V和W的基$\alpha_1, \alpha_2, \ldots, \alpha_n$和$\beta_1, \beta_2, \ldots, \beta_m$，且$\sigma$在这两组基下的矩阵表示为$A = (a_{ij})_{m\times n}$.

1. 若将α_i换为$c\alpha_i$（c为常数），求σ在新基下的矩阵表示;
2. 若将β_i换为$c\beta_i$（c为常数），求σ在新基下的矩阵表示;
3. 若将α_i换为$\alpha_i + k\alpha_j$（k为常数），求σ在新基下的矩阵表示;
4. 若将β_i换为$\beta_i + k\beta_j$（k为常数），求σ在新基下的矩阵表示;
5. 若将α_i和α_j对换，求σ在新基下的矩阵表示;
6. 若将β_i和β_j对换，求σ在新基下的矩阵表示.