3 矩阵计算　矩阵的秩　相抵标准形

3.1 课前准备

3.1.1 矩阵乘法基础

矩阵乘法 AB 表现为 A 的行和 B 的列的点乘，我们可以给出矩阵乘法的正式定义：

定义 3.1

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

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

这一定义带给我们的感受与我们在上一讲定义矩阵加法和数乘时的直观不同，如果脱离了映射复合的背景，在我们初看这一定义时必然会产生一个疑惑：为什么矩阵乘法定义得如此复杂，为什么不定义成两个矩阵对应元素相乘就可以了呢？现在我们给出最一般的从线性映射复合出发定义矩阵乘法的方法.

1. 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 σ,τ 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：

$$M(\tau)=A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\vdots&\vdots&\ddots&\vdots\\a_{p1}&a_{p2}&\cdots&a_{pm}\end{pmatrix},M(\sigma)=B=\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&\ddots&\vdots\\b_{m1}&b_{m2}&\cdots&b_{mn}\end{pmatrix}.$$

2. 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. 如果我们用 $\mathbf{M}_{B_1,B_2}(\sigma)$ 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:

$$\mathbf{M}_{B_2,B_3}(\tau)\mathbf{M}_{B_1,B_2}(\sigma)=\mathbf{M}_{B_1,B_3}(\tau\sigma)\tag{3.1}$$

需要注意的是，在两个映射矩阵表示的基底选择中，V_2 选取的基底应当是相同的，否则前面的推导过程就会失败，交换图也会失效.

接下来我们再给出两个线性映射的例子来熟悉矩阵乘法. 在给出例子前，我们需要强调一个事实，即矩阵相乘时左侧矩阵的列数应该等于右边矩阵的行数:

1. 一方面，从矩阵相乘的运算方式来看，若考虑矩阵乘法 $C=AB$，那么 C 的第 i 行第 j 列元素 c_{ij} 是 A 的第 i 行与 B 的第 j 列对应位置元素相乘后求和的结果，因此这一定义要求 A 每一行的元素个数与 B 每一列的元素个数相等，即 A 的列数等于 B 的行数.

2. 回顾线性映射的复合，若复合 $\sigma_1\sigma_2$ 符合定义，则必须有 σ_2 的到达空间恰好是 σ_1 的出发空间，故两空间维数一致，那么 σ_1 对应的矩阵 A 的列数和 σ_2 对应的矩阵 B 的行数一致，这也是我们要求两个矩阵 A,B 可乘的重要条件的来源. 而最后乘法的结果行数等于 A 的行数，列数等于 B 的列数，这也与 $\sigma_1\sigma_2$ 出发空间为 σ_2 出发空间（对应于 B 的列数），到达空间为 σ_1 到达空间（对应于 A 的行数）一致.

接下来我们来看两个例子. 第一个例子是最直接的矩阵乘法的计算. 这一例子中我们将看到，矩阵乘法不满足交换律，即 AB 不一定等于 BA（这一点之后会进一步介绍）:

例 3.1

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.

第二个例子需要上一节提到的旋转 θ 角线性映射对应的矩阵 $M_\theta=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$.

例 3.2

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.

接下来给出矩阵运算的几个基本性质：

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

证明方法十分简单：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.

实际上，由矩阵加法和乘法满足的运算律可知，全体 n 阶方阵构成的集合 $\mathbf{F}^{n\times n}$ 关于矩阵加法和乘法构成环.

事实上矩阵乘法有很多和数的乘法重要的不同，我们在此特别指出供读者参考：

(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 $AB = O$ 在环上这种元素通常被称为零因子.

(4) 消去律也不一定满足：即 $AB = AC$ 不一定 $B = C$. 原因在于 $AB = AC \implies A(B-C) = O$，由 (3) 可知不一定 $B = C$.

我们在线性空间中已经介绍过，我们一般用 $\mathbf{F}[x]_{m+1}$ 表示数域 $\mathbf{F}$ 上的次数最高为 m 的多项式全体，其中的元素我们一般记为

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0,\ a_i \in \mathbf{F}\ \ (i = 1, 2, \ldots, m)$$

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

$$p(\sigma) = a_m \sigma^m + a_{m-1} \sigma^{m-1} + \cdots + a_1 \sigma + a_0 I$$

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

$$p(A) = a_m A^m + a_{m-1} A^{m-1} + \cdots + a_1 A + a_0 E,$$

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，我们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. $A^k A^m = A^{k+m}$，$(A^k)^m = A^{km}$，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 3.3

展开矩阵多项式 $(A + \lambda E)^n$.

例 3.4

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 $f(A)g(B) = g(B)f(A)$;

最后，我们介绍三类经典的矩阵，它们可以视为三个矩阵多项式的“零点”:

1. 幂等矩阵：满足 $A^2 = A$，即 A 的平方等于自身，满足 $A^2 - A = O$. 显然的性质是对于任意的 $k \in \mathbf{N}^+$ 都有 $A^k = A$，对 k 做数学归纳法即可证明;
2. 幂零矩阵：存在 $k \in \mathbf{N}^+$ 使得 $A^k = O$，即 A 的 k 次幂为零矩阵. 一个显然的性质是对于任意的 $m \in \mathbf{N}$ 都有 $A^{m+k} = O$;
3. 对合矩阵：满足 $A^2 = E$，即 A 的平方等于单位矩阵，满足 $A^2 - E = O$. 显然的性质是对于任意的 $k \in \mathbf{N}^+$ 都有 $A^{2k} = E$，$A^{2k+1} = A$.

3.1.2 矩阵的逆基础

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，因此由同构的等价条件，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. 回顾式 3.1，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 $AB = \mathbf{M}_{B_1,B_1}(I_{V_1}) = E_n = \mathbf{M}_{B_2,B_2}(I_{V_2}) = BA$，由此我们从可逆映射的角度引入矩阵的逆的概念:

定义 3.2 矩阵的逆

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**逆矩阵**，记作 $B = A^{-1}$.

在一些比较经典的教材中可逆矩阵也被称为非奇异矩阵，不可逆矩阵被称为**奇异矩阵**.

注意，因为逆映射是同构，故要求 V_1 和 V_2 的维数相同，故逆矩阵定义必定也基于方阵，非方阵没有上述逆矩阵（在讨论了矩阵的秩之后我们会给出另一个方面的解释）. 广义逆矩阵允许非方阵，但那是另一个定义，我们不需要掌握. 对于可逆矩阵，以下两个定理是基本的：

定理 3.1

可逆矩阵 A 的逆矩阵唯一.

注意这个唯一性的证明，我们在证明群的单位元唯一时使用了完全一致的思想.

定理 3.2

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

即对于方阵而言，我们有 $AB = E \iff BA = E \iff A, B$ 可逆.

下面是矩阵的逆的基本性质：

(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) $(AB)^{-1} = B^{-1}A^{-1}$，$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$；注意这一点和 (2) 的证明都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;

注意，这种验证逆的相关性质思想（即直接验证相乘是否为单位矩阵，然后利用逆的唯一性的方法）在之后的讨论中也是非常常见的，希望读者掌握.

(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.

需要强调的是，我们之后讨论运算性质的时候都是循着类似的思路，考虑加法、数乘、乘法（2 个相乘，n 个相乘，矩阵的幂）、逆、转置、共轭等，所以虽然每个地方给出的性质都很多，但实际上大致研究思路是一致的.

在介绍完性质后我们非常关心如何给定一个具体的矩阵求出它的逆的问题，这里我们给出第一种基本方法，即基于解方程的方法. 事实上，我们在矩阵乘法一节中就将 $AX = b$ 和 $\sigma(a) = \beta$ 联系在一起，其中 σ 在基 B_1, B_2 下表示矩阵为 A，b 是 β 在基 B_2 下的坐标. 回顾本讲开头引入可逆矩阵的过程，可逆矩阵 A 应当是可逆线性映射 σ 关于某组基的表示矩阵. 对于可逆映射而言，首先必须是单射，因此 $\sigma(a) = \beta$ 只能有唯一解，因此 $AX = b$ 只能有唯一解.

事实上我们可以很简便地表达出这个解. 我们在 $AX = b$ 左右同时左乘 A^{-1}（矩阵乘法不可交换，所以必须在同一侧乘），有 $A^{-1}AX = A^{-1}b$，即 $X = A^{-1}b$. 因此，当 A 可逆时，对于任意的 b 线性方程组都有唯一解，且解可以被表示为 $X = A^{-1}b$ 的形式. 因此我们可以通过解线性方程组的方法求解逆矩阵. 我们将通过下面这个例子详细介绍这种方法的计算过程：

例 3.5

用解方程法求矩阵 $A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}$ 的逆矩阵.

此后我们还会有使用伴随矩阵求解逆矩阵的方法.

3.1.3 矩阵转置基础

此前的讨论我们都是先介绍矩阵运算与线性映射的关联，然而转置背后的关联可能有些复杂，因此我们仅给出更具体的计算的讨论.

定义 3.3 矩阵的转置

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（$k = 1, 2, \ldots, n$）；它的第 r 列正好是 A 的第 r 行（$r = 1, 2, \ldots, m$）.

根据定义，矩阵的转置显然等价于第 i 行第 j 列的元素变成转置后矩阵的第 j 行第 i 列的元素. 一个简单的例子是，令 $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$，则 $A^{\mathrm{T}} = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix}$.

事实上，在求转置的时候，我们只需要把第一行保持顺序变成第一列，第二行保持顺序变成第二列，以此类推. 因为这样第一行第 i 列的元素就变成了第 i 行第一列的元素，符合转置的要求，其它行的元素也是类似的，从上面的例子也能看出这一点. 下面是一些基本性质：

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：

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；

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；

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 $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$；

与矩阵转置相关的一个重要的概念是对阵矩阵和反对称矩阵. 我们给出定义：

定义 3.4

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.

由定义易知 A 为对称矩阵的充要条件为 $A = A^{\mathrm{T}}$，A 为反对称矩阵的充要条件为 $A = -A^{\mathrm{T}}$.

例 3.6

证明以下几点性质：

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.

下面的例子讨论了对称矩阵和反对称矩阵构成线性空间的性质：

例 3.7

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 $V = V_1 \oplus V_2$.

3.1.4 初等矩阵基础

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.

定义 3.5

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：

$$E_i(c)=\left[\begin{array}{ccccc}1&&&&\\&\ddots&&&\\&&c&&\\&&&\ddots&\\&&&&1\end{array}\right]\begin{array}{l}\\ \\ \text{第 } i \text{ 行}\\ \\ \\ \end{array}$$

$$E_{ij}(c)=\left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&\ddots&&&\\&&c&&1&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\begin{array}{l}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{array}$$

$$E_{ij}=\left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&0&&1&&\\&&&\ddots&&&\\&&1&&0&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\begin{array}{l}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{array}$$

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. 我们很容易通过计算验证，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 $P_1,P_2,\dots,P_k$，则有

$$P_kP_{k-1}\cdots P_2P_1A=U.$$

大家非常关心为什么初等矩阵左乘代表行变换，右乘代表列变换. 事实上读者只需要回顾矩阵乘法一小节中说明的如下性质："矩阵 A 与 B 相乘，乘积的每一列都是矩阵 A 各列的线性组合，每一行都是矩阵 B 各行的线性组合" 即可. 左乘的时候情况对应于 A 是初

等矩阵，它乘在 B 的左边，那么乘积 AB 的每一行都是 B 的各行的线性组合，即 A 左乘 B 后相对于对 B 进行了行变换. 右乘的情况对应于 B 为初等矩阵，结果 AB 的每一列都是 A 的各列的线性组合，即 B 右乘 A 后相当于对 A 进行了列变换.

接下来我们还有几个细节需要讨论：

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.

初等矩阵的可逆性结合可逆矩阵的乘积性质表明，任何可逆矩阵乘以初等矩阵后仍然是可逆矩阵. 下面这个例题综合了初等矩阵的性质以及此前学到的矩阵运算性质，是一个有价值的练习：

例 3.8

已知矩阵 $A=\begin{pmatrix} a & b & c\\ d & e & f\\ h & x & y\end{pmatrix}$ 的逆是 $A^{-1}=\begin{pmatrix} -1 & -2 & -1\\ 2 & 1 & 0\\ 0 & -3 & -1\end{pmatrix}$，

$B=\begin{pmatrix} a-2b & b-3c & -c\\ d-2e & e-3f & -f\\ h-2x & x-3y & -y\end{pmatrix}$. 求矩阵 X 满足：

$$X+\left(B(A^{\mathrm{T}}B^2)^{-1}A^{\mathrm{T}}\right)^{-1}=X\left(A^2(B^{\mathrm{T}}A)^{-1}B^{\mathrm{T}}\right)^{-1}(A+B)$$

初等矩阵的可逆性是平凡的结论，但反过来，可逆矩阵是否都可以表达为初等矩阵的乘积则显得更为有趣. 事实上答案是肯定的，我们将其写在下述定理中，这是一个相当重要的定理，在之后很多讨论和解题中都扮演关键角色：

定理 3.3

任意可逆矩阵都可以被表示为若干个初等矩阵的乘积.

本节我们将基于上述对初等矩阵的讨论给出逆矩阵求解中另一种基本且更常用的方法. 我们首先给出一个引理:

引理 3.1

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 $P_1, P_2, \ldots, P_k$ 后 A 变为 E 时, E 变为 A^{-1}.

证明

由题意有 $P_kP_{k-1}\cdots P_2P_1A = E$, 即 $P_kP_{k-1}\cdots P_2P_1 = A^{-1}$, 因此对 E 做相同的初等行变换有 $P_kP_{k-1}\cdots P_2P_1E = A^{-1}E = A^{-1}$. □

我们可以将上述过程写成 $\left(\begin{array}{c:c} A & E \end{array}\right) \xrightarrow{\text{初等行变换}} \left(\begin{array}{c:c} E & A^{-1} \end{array}\right)$. 事实上初等列变换也有类似过程:

$$\begin{pmatrix} A \\ \hdashline E \end{pmatrix} \xrightarrow{\text{初等列变换}} \begin{pmatrix} E \\ \hdashline A^{-1} \end{pmatrix}$$

原因是对 A 做列变换 $P_1, P_2, \ldots, P_k$ 后, A 变为 E, 这一过程可以写为 $AP_1P_2\cdots P_k = E$, 因此 $P_1P_2\cdots P_k = A^{-1}$, 因此对 E 做相同的列变换有 $EP_1P_2\cdots P_k = EA^{-1} = A^{-1}$.

注意, 上面我们在行变换时将 A 和 E 放在一行是为了方便我们实际操作的时候, 我们可以对 A 和 E 同时做行变换, 列变换放在一列表示的原因同理, 只是为了实际操作的时候更容易看清, 在后面的例子中我们可以仔细体会到这一点.

注意, 基于初等变换的方法是非常重要的, 我们很多时候使用的方法就是初等行变换(列变换也可以用但一般更习惯行变换). 我们将通过下面这个例子详细介绍这种方法的计算过程:

例 3.9

用上述方法求矩阵 $A = \begin{pmatrix} 0 & 2 & -1 \\ 1 & 1 & 2 \\ -1 & -1 & -1 \end{pmatrix}$ 的逆矩阵.

例 3.10

1. 设 A 为三阶矩阵，将 A 的第二列加到第一列得到矩阵 B，再对调 B 的 2、3 行得到单位矩阵. 令 $P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$，试用 P_1 和 P_2 表示 A.

2. 设 A 为可逆矩阵，将 A 的第 i 行和第 j 行对调得到矩阵 B，证明矩阵 B 可逆并求 AB^{-1}.

3. 设 A 为三阶可逆矩阵，且 $P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$，其中 $P = (\alpha_1, \alpha_2, \alpha_3)$，令 $Q = (\alpha_1 + \alpha_2, \alpha_2, \alpha_3)$，求 $Q^{-1}AQ$.

3.1.5 矩阵分块基础

矩阵分块在矩阵计算中是非常核心的一种手段，这可以使得我们将大矩阵分为更容易处理的小矩阵，结合并行计算等工具能大大加速矩阵计算. 除此之外，基于分块矩阵的初等变换也是研究矩阵求逆、矩阵的秩以及矩阵分解等多个问题的重要工具.

定义 3.6

一般地，对于 $m \times n$ 矩阵 A，如果在行的方向分成 s 块，在列的方向分成 t 块，就得到 A 的一个 $s \times t$ **分块矩阵**，记作 $A = (A_{kl})_{s\times t}$，其中 A_{kl} $(k = 1, \dots, s,\ \ l = 1, \dots, t)$ 称为 A 的子块.

实际上上述表示方法就是将一般矩阵表示 $A = (a_{ij})_{m\times n}$ 中的 a_{ij} 替换为了小块矩阵，字母含义并无变化，内层代表索引，外层代表总行列数（只是分块矩阵是块索引和块数）. 我们接下来考察分块矩阵的运算性质.

1. 分块矩阵的加法：设分块矩阵 $A = (A_{kl})_{s\times t}$，$B = (B_{kl})_{s\times t}$. 如果 A 与 B 对应的子块 A_{kl} 和 B_{kl} 都是同型矩阵，则

$$A + B = (A_{kl} + B_{kl})_{s\times t}$$

由此我们看到分块矩阵加法要求小块形状和行列分块数都一致，实际上回顾一般矩阵加法要求矩阵完全同型即可理解这一要求.

2. 分块矩阵的数乘：设分块矩阵 $A = (A_{kl})_{s\times t}$，λ 是一个数，则

$$\lambda A = (\lambda A_{kl})_{s\times t}$$

实际上数乘最好理解，因为如此计算的效果相当于一般矩阵数乘的效果，即给每个元素都乘以一个常数 λ.

3. 分块矩阵的乘法：设 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{n\times p}$，如果把 A,B 分别分块为 $r\times s$ 和 $s\times t$ 分块矩阵，且 A 的列分块法与 B 的行分块法相同（注意这些条件始终保证可乘性成立），则

$$AB=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1s}\\ A_{21} & A_{22} & \cdots & A_{2s}\\ \vdots & \vdots & \ddots & \vdots\\ A_{r1} & A_{r2} & \cdots & A_{rs}\end{pmatrix}\begin{pmatrix}B_{11} & B_{12} & \cdots & B_{1t}\\ B_{21} & B_{22} & \cdots & B_{2t}\\ \vdots & \vdots & \ddots & \vdots\\ B_{s1} & B_{s2} & \cdots & B_{st}\end{pmatrix}=C=(C_{kl})_{r\times t}$$

其中 C 是 $r\times t$ 分块矩阵，且 C_{kl} 与一般矩阵计算类似，即为 A 第 k 行块 B 的 l 列块对应元素相乘后相加，即

$$C_{kl}=A_{k1}B_{1l}+A_{k2}B_{2l}+\cdots+A_{ks}B_{sl},\ \ k=1,\ldots,r,\ \ l=1,\ldots,t$$

4. 分块矩阵的转置：大、小矩阵都要转置，这是分块矩阵与普通矩阵的一大性质差异；即 $s\times t$ 分块矩阵 $A=(A_{kl})_{s\times t}$ 转置后 $A^{\mathrm{T}}=(B_{lk})_{t\times s}$ 为 $t\times s$ 分块矩阵，且 $B_{lk}=A_{kl}^{\mathrm{T}}$. 例如 $\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}^{\mathrm{T}}=\begin{pmatrix}A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}}\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix}$.

3.1.6 线性代数的核心问题

上课揭晓答案！

3.2 特殊矩阵

3.2.1 对角矩阵

我们一般记主对角矩阵为 $\mathrm{diag}(d_1, d_2, \ldots, d_n)$，准对角矩阵为 $\mathrm{diag}(A_1, A_2, \ldots, A_n)$.

定理 3.4

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 $\mathrm{diag}(d_1, d_2, \ldots, d_n)$ 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 $\mathrm{diag}(d_1, d_2, \ldots, d_s)$ 相当于给 A 的第 $i(i = 1, 2, \ldots, s)$ 行元素都乘以 d_i.

定理 3.5

对角矩阵和分块对角矩阵的性质：

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

对角矩阵，且 $AB = \mathrm{diag}(a_1b_1, a_2b_2, \ldots, a_nb_n)$.

对于乘方运算，有 $A^k = \mathrm{diag}(a_1^k, a_2^k, \ldots, a_n^k)$.

4. 两个准对角矩阵 $A = \mathrm{diag}(A_1, A_2, \ldots, A_n)$，$B = \mathrm{diag}(B_1, B_2, \ldots, B_n)$ 中 A_i 和 B_i 是同级方阵，则乘积仍然是准对角矩阵，且 $AB = \mathrm{diag}(A_1B_1, A_2B_2, \ldots, A_nB_n)$.

3.2.2 上（下）三角矩阵

定理 3.6

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 3.11

已知 $A_1, \ldots, A_n$ 是 n 个对角元都为 0 的上三角矩阵，证明：$A_1A_2\cdots A_n = O$.

3.2.3 自然基与基本矩阵

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. 其运算性质我们可以总结如下，读者应当不难验证：

定理 3.7

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) $f_i^{\mathrm{T}}Ae_j$ 是 A 的第 i 行第 j 列元素.

我们已经讨论了自然基的运算性质，这是只有一个元素为 1，其余元素全为 0 的向量，接下来我们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}，它们具有如下性质（可以联系左右乘对应行列变换进行记忆）:

定理 3.8

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. $E_{ij}AE_{kl}=a_{jk}E_{il}$.

例 3.12

求证: 任一 n 阶矩阵均可表示为形如 $I_n+a_{ij}E_{ij}$ 这样的矩阵之积，其中 E_{ij} 是 n 阶基本矩阵.

3.2.4 其他矩阵

其他特殊矩阵如正交矩阵、幂等矩阵、幂零矩阵等，我们将在后续讲义合适的位置描述它们的性质，那时我们的讨论不局限于本节的运算性质，会有更多的其它性质. 除此之外，在 LALU 的习题中我们也会看到一些特殊矩阵的进阶性质，感兴趣的读者可以尝试.

3.2.5 初等矩阵的补充例子

例 3.13

解答有关于初等变换的三个问题:

1. 证明: 对换变换可以通过若干次倍加和倍减变换实现;
2. 将二阶矩阵 $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ 表示为若干个倍加初等矩阵的乘积;
3. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $ad - bc = 1$, 证明: A 可以表示为若干个倍加初等矩阵的乘积.

3.3 矩阵乘法

3.3.1 矩阵乘法的进一步认识

我们需要探讨四个非常细节且重要的问题:

1. 在有了矩阵乘法的定义后, 高斯-若当消元法中我们将线性方程组简记为 $AX = b$ 实际上是符合矩阵乘法定义的. 除此之外, 我们将向量坐标表示为列向量的形式, 例如将向量 v 对基 α 分解

$$v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

这也是符合矩阵形式乘法定义的一种习惯, 尽管基一般不是数域 $\mathbf{F}$ 或者向量空间 $\mathbf{F}^m$ 中的元素.

2. 事实上, 在这里我们可以看出求解线性方程组和线性映射之间的关联. 我们设 $AX = b$ 的解为

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

由 $AX=b$ 和线性映射矩阵表示，我们有

$$(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\ldots,\sigma(\varepsilon_n))\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=(\alpha_1,\alpha_2,\ldots,\alpha_m)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\beta \tag{3.2}$$

即 b 是 β 在基 $\alpha_1,\alpha_2,\ldots,\alpha_m$ 下的坐标. 由上式可知 $x_1\sigma(\varepsilon_1)+x_2\sigma(\varepsilon_2)+\cdots+x_n\sigma(\varepsilon_n)=\beta$，即

$$\sigma(x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n)=\beta. \tag{3.3}$$

由此我们将线性方程组的求解问题和找到线性映射到达空间中某个向量在出发空间中原像的坐标联系起来了，即将求解 $AX=b$ 和求解 $\sigma(a)=\beta$ 联系起来了，只是我们求解后者时通常是对于一般的向量空间而言的.

若前述 $b=0$，则我们将齐次线性方程组的解空间与线性映射的核空间联系起来了，即线性映射的核空间中元素在一组基下的向量就是这一线性映射在这组基下的矩阵表示作为系数矩阵的线性方程组的解. 这一联系将在线性方程一章中有更深入的讨论.

3. 我们可以更进一步理解矩阵乘法. 假设矩阵 $A=(a_{ij})_{m\times n}$ 与 $B=(b_{ij})_{n\times l}$ 相乘，我们有如下结论：

(1) 乘积的第 k 列等于 A 乘以 B 的第 k 列，乘积的第 j 行等于 A 的第 j 行乘以 B，这一点根据矩阵乘法计算方式显然，我们可以利用这一性质证明下面例子的结论：

例 3.14

证明：有一行元素或一列元素全为 0 的 n 阶方阵必定不可逆.

例 3.15

设 A,B 都是由非负实数组成的矩阵且 AB 有一行等于 0，证明：或者 A 有一行为 0，或者 B 有一行为 0.

(2) 乘积的每一列都是矩阵 A 各列的线性组合，每一行都是矩阵 B 各行的线性组合. 我们简要说明前者，后者理由类似. 我们考察乘积的每一列，由 1 可知乘积的第 k 列等于 A 乘以 B 的第 k 列，我们展开写乘积矩阵 $C=(c_{ij})_{m\times l}$ 第 k 列的结

果:

$$
\begin{aligned}
c_{1k} &= a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk} \\
c_{2k} &= a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk} \\
&\vdots \\
c_{mk} &= a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk}
\end{aligned}
$$

我们将上面的行进行组合，写成列向量形式，即

$$
\begin{pmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{pmatrix} = b_{1k}\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_{2k}\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + b_{nk}\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}
$$

由此我们将乘积的列表示成了矩阵 A 各列的线性组合.

4. 在不同的教材中我们可能会看见两种记号，即

$$
\begin{aligned}
(\sigma(\varepsilon_1), \sigma(\varepsilon_2), \ldots, \sigma(\varepsilon_n)) &= (\alpha_1, \alpha_2, \ldots, \alpha_m)A \\
\sigma(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) &= (\alpha_1, \alpha_2, \ldots, \alpha_m)A
\end{aligned}
$$

事实上两个记号是等价的，只有记号上的差别，含义完全相同. 有时我们还会看到一个很特别的书写方式

$$
(\sigma(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n))B = \sigma((\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)B)
$$

这容易导致读者的困惑，因此我们这里简要说明它们的确是成立的，从而接下来读者可以放心地自由使用这一结论. 根据上述的第一个性质可知，我们只需要证明对 B 的某一列上式成立即可，因为乘法结果是列与列对应的. 我们设 B 的第 k 列为

$$
B_k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}
$$

则

$$
\begin{aligned}
(\sigma(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n))\begin{pmatrix} b_{1k}\\ b_{2k}\\ \vdots\\ b_{nk}\end{pmatrix} &= (\sigma(\varepsilon_1),\sigma(\varepsilon_2),\dots,\sigma(\varepsilon_n))\begin{pmatrix} b_{1k}\\ b_{2k}\\ \vdots\\ b_{nk}\end{pmatrix}\\
&= b_{1k}\sigma(\varepsilon_1)+b_{2k}\sigma(\varepsilon_2)+\cdots+b_{nk}\sigma(\varepsilon_n)\\
&= \sigma(b_{1k}\varepsilon_1+b_{2k}\varepsilon_2+\cdots+b_{nk}\varepsilon_n)\\
&= \sigma((\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)\begin{pmatrix} b_{1k}\\ b_{2k}\\ \vdots\\ b_{nk}\end{pmatrix})
\end{aligned}
$$

故得证.

3.3.2 矩阵可交换问题

首先我们需要强调一点：一般来说在本课程中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. 我们来看下面的例子：

例 3.16

求所有与 A 可交换的矩阵，其中

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

对于一些矩阵直接设未知数计算比较复杂，这里我们讨论一个基本的技巧，即利用

$$
\forall t,\ AB=BA \iff (A-tE)B=B(A-tE).
$$

这一等式成立是显然的. 运用时难点主要在决定 t 的值，我们要根据矩阵的对角线上元素来决定，原则是使得 B 与 $A-tE$ 相乘的计算过程更为简单（一般是使得 0 元素更多），这样解方程也会更轻松. 我们看一个简单的例子来体会：

例 3.17

求与矩阵 $A=\begin{pmatrix}3&0&0\\-1&3&0\\0&-1&3\end{pmatrix}$ 可交换的矩阵.

例 3.18

1. 已知矩阵 $A=\begin{pmatrix}1&0&4\\0&1&2\\0&1&2\end{pmatrix}$，求证：所有与 A 可交换的矩阵构成 $\mathbf{M}_3(\mathbf{R})$ 的一个子空间，并求子空间的一组基.
2. 已知矩阵 $A=\begin{pmatrix}1&0&1\\0&2&0\\1&0&1\end{pmatrix}$，
   (1) 求所有与 A 可交换的矩阵；
   (2) 若 $AB+E=A^2+B$，求 B.
3. 设 $A\in\mathbf{F}^{n\times n}$，令 $C(A)=\{B\in\mathbf{F}^{n\times n}\mid AB=BA\}$.
   (1) 证明：$C(A)$ 为 $\mathbf{F}^{n\times n}$ 的一个子空间；
   (2) 求 $C(E)$；
   (3) 当 A 为对角线上元素互不相等的对角阵时，求 $C(A)$ 的维数和一组基.

例 3.19

证明以下两个命题：

1. 与矩阵 $I=\begin{pmatrix}0&1&&&\\&&1&&\\&&&\ddots&\\&&&&1\\1&&&&0\end{pmatrix}$ 可交换的矩阵 A 都可以写成 I 的一个多项式，即 $A=a_{11}E+a_{12}I+a_{13}I^2+\cdots+a_{1n}I^{n-1}$；

2. 与矩阵 $J = \begin{pmatrix} 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & 0 \end{pmatrix}$ 可交换的矩阵 A 都可以写成 J 的一个多项式，即 $A = a_{11}E + a_{12}J + a_{13}J^2 + \cdots + a_{1n}J^{n-1}$.

事实上，我们有如下关于可交换矩阵更一般的结论：

定理 3.9

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.

3.3.3 矩阵的幂

1. 找规律

在矩阵的转置例 3.48 中我们已经见识了一种找规律的方式，下面是一种类似的题型：

例 3.20

计算 $(PAQ)^k$，其中

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

本质而言此类题目只需要发现中间多次出现的乘积 QP 是很容易处理的矩阵（例 3.48 中甚至是一个数）即可解决.

例 3.21

设 $A=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&-1\end{pmatrix}$，$P^{-1}AP=B$，求 $B^{2004}-2A^2$.

还有一种找规律基于幂等矩阵，显然幂等矩阵的任意次方都与其本身相等是很好的性质，另一种找规律基于对合矩阵，偶数次方等于单位矩阵，奇数次方等于本身，这显然也可以直接利用. 这里主要与大家分享两类特殊矩阵的幂的性质，这些性质在矩阵的幂的计算中是非常有用的.

例 3.22

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 $1\leqslant k\leqslant n$.

下面的例子应用了上述结论：

例 3.23

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

我们采用将矩阵分为 $A=aE+B$ 的方式，因为这样矩阵 B 为上三角矩阵且对角线上全为 0，甚至是若当块矩阵——我们已经证明这是典型的幂零矩阵.

2. 数学归纳法

例 3.24

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

这一问题对应我们常见的旋转变换（所以建议要求读者记忆这一矩阵形式），n 次方就是旋转 n 次. 当然这是直观而言的结论，严谨说明可以通过数学归纳法证：

例 3.25

证明 $\begin{pmatrix} a & c\\ 0 & b\end{pmatrix}^n=\begin{pmatrix} a^n & (a^{n-1}+a^{n-2}b+\cdots+b^{n-1})c\\ 0 & b^n\end{pmatrix}$.

3. 利用秩为 1 的矩阵

这一方法的核心是利用上一讲中例 3.83 的结论，我们来看下面的例子进行体会：

例 3.26

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

针对本例我们可以给出一个更加具体的例子供读者练习，读者可以从中感受观察矩阵特点然后凑出秩为 1 矩阵的技巧.

例 3.27

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

解答中看出秩 1 矩阵的分解需要一定的灵光一现与经验，日后行列式中我们还会遇到很多需要将矩阵分解为两个向量或矩阵乘积的情况，这里只是一个开始.

例 3.28

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：$A^2 = O$.

事实上，将来我们讨论幂零矩阵的时候将会进一步推广本例的结论.

4. 利用初等矩阵的性质

例 3.29

设 A 为三阶矩阵，P 为三阶可逆矩阵，$P^{-1}AP = B$，其中

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

求 A^{2024}.

事实上本题一个关键的洞察在于我们很容易看出 B 这一非常简单的矩阵作为初等矩阵复合的情况（交换 1、3 行以及每一行都乘以 -1），因此其平方为 E.

5. 利用对角化和若当标准形：我们将在后续相应章节中讲解.

例 3.30

已知数列 $\{a_n\},\{b_n\}$ 满足 $a_0 = -1$，$b_0 = 3$，且

$$\begin{cases} a_n = 3a_{n-1} + b_{n-1} + 2^{n-1} \\ b_n = 2a_{n-1} + 4b_{n-1} + 2^n \end{cases}$$

求 $\{a_n\},\{b_n\}$ 的通项公式.

3.4 矩阵的逆

3.4.1 基本技巧

第一个基本技巧是利用 $AB = E$ 可以推出 $BA = E$ 这一容易被遗忘的性质:

例 3.31

若 A, B 为两个 n 阶矩阵且满足 $A + B = AB$，证明:

1. $A - E$ 和 $B - E$ 均可逆;
2. $AB = BA$;
3. $r(A) = r(B)$.

下面这一例子也非常经典，在学习了特征值后会有进一步的解释:

例 3.32

设 A 为 n 阶矩阵，A 的每行各元素之和都等于 k，证明:

1. 若 A 可逆，则 $k \neq 0$ 且 A^{-1} 的每行各元素之和都等于 $\dfrac{1}{k}$;
2. A^i（其中 i 为正整数）每行元素之和为 k^i.

例 3.33

判断: 设 n 阶方阵 A 的每一行元素之和是 10，则 $2A^3 + A + 9E_n$ 的每一行元素之和是 2019.

例 3.34

设 A, B, C 为二阶复方阵, 且 A, B, C 在 $\mathbf{M}_2(\mathbf{C})$ 中线性无关. 证明: 存在 $z_1, z_2, z_3 \in \mathbf{C}$ 使得 $z_1A + z_2B + z_3C$ 为可逆矩阵.

3.4.2 凑因子法

此类题目出现比较频繁，实际上就是通过一些初中所学的因式分解等基本变换，得到需要求逆的矩阵与另一个矩阵相乘等于单位矩阵 (的一个倍数), 这样就相当于凑出了逆矩阵. 我们来看一个简单的例子:

例 3.35

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

例 3.36

1. 判断：设 n 阶方阵 A 满足 $A^2-5A+5E_n=0$，则对所有的有理数 r，$A+rE_n$ 都是可逆阵；
2. 设方阵 A 满足 $A^2-A-2E=O$，证明：
   (1) A 和 $E-A$ 都是可逆矩阵，并求它们的逆矩阵；
   (2) $A+E$ 和 $A-2E$ 不可能同时可逆.
3. (1) 若 n 阶方阵 A 满足 $A^2=E$，若 A 还满足 $E+A$ 可逆，证明：$A=E$；
   (2) 若 n 阶方阵 A 满足 $A^2=A$，证明：$E+A$ 可逆.

例 3.37

1. 设 A 是 n 阶矩阵，$A^k=O$ 对某个正整数 k 成立，求证下列方阵可逆，并求它们的逆：
   (1) $E+A$；
   (2) $E-A$；
   (3) $E+A+\dfrac{1}{2!}A^2+\cdots+\dfrac{1}{(k-1)!}A^{k-1}$.
2. 设 $A=\begin{pmatrix}1&\cdots&1\\ \vdots&\ddots&\vdots\\ 1&\cdots&1\end{pmatrix}_{n\times n}$，求：
   (1) 一个二次实系数多项式 $f(x)=ax^2+bx$，使得 $f(A)=O$；
   (2) A^{100}；
   (3) $(A+E)^3$；
   (4) $(A+E)^{-1}$.

例 3.38

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.

凑因子法可以更具技巧性，下面的例子展示了这一点：

例 3.39

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.

这一例子我们将多次使用：例 3.43是完全一致的，例 3.65 是 A,B 为方阵时的特例. 但使用凑因子法不仅可以证明逆矩阵存在，还可以直接给出逆矩阵的表达式.

有了这一结果后，我们可以证明一个计算数学中非常常见的公式：Sherman-Morrison-Woodbury 公式.

例 3.40

设 A 为 n 阶可逆矩阵，考虑 $B=A+XRY$，其中 X 是 $n\times r$ 矩阵，Y 是 $r\times n$ 矩阵，而 R 是 $r\times r$ 的可逆矩阵. 若 $R^{-1}+YA^{-1}X$ 可逆，则

$$B^{-1}=A^{-1}-A^{-1}X(R^{-1}+YA^{-1}X)^{-1}YA^{-1}.$$

初看这一公式想必一定会发出“这么复杂的公式为什么有用”的感叹. 事实上，我们考虑 r 比 n 小很多的情况，那么这个时候对 R 和 $R^{-1}+YA^{-1}X$ 求逆会更容易，因为它们的阶数会低很多，计算量也小很多. 考虑极端情况，若 x 和 y 是非零列向量，取 $X=x$，$Y=y^{\mathrm{T}}$，$R=(1)$，且满足 $1+y^{\mathrm{T}}A^{-1}x\neq 0$，我们可以得到 Sherman-Morrison-Woodbury 公式的一个特殊情形（称为 Sherman-Morrison 公式）：

$$(A+xy^{\mathrm{T}})^{-1}=A^{-1}-(1+y^{\mathrm{T}}A^{-1}x)^{-1}A^{-1}xy^{\mathrm{T}}A^{-1}.$$

这也称为“秩 1 校正”，因为 xy^{T} 是一个秩 1 矩阵，在最优化理论中是常见的.

例 3.41

求 n 阶矩阵 $A = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{pmatrix}$ 的逆矩阵.

3.4.3 求逆的分式思想

虽然矩阵没有除法运算，但是我们如果将 $(E-A)^{-1}$ 写成 $\dfrac{E}{E-A}$，再类比泰勒展开

$$\frac{1}{1-x} = 1 + \sum_{n=1}^{\infty} x^n \qquad x \in (-1,1)$$

我们可以得到（不严谨！只能用来解题的时候当作初步的思路！）

$$(E-A)^{-1} = \frac{E}{E-A} = E + A + A^2 + \cdots$$

例 3.42

已知方阵 A 满足 $A^k = O$，其中 k 是一个正整数，求 $E-A$ 的逆.

例 3.43

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.

3.4.4 提逆思想

这一思想的来源是矩阵逆没有加减相关的运算法则（即没有 $(A+B)^{-1} = A^{-1} + B^{-1}$ 这样的性质），因此我们需要提逆产生一些乘积项来解决问题.

例 3.44

设 $A, B, A+B$ 都是可逆矩阵，求证：$A^{-1} + B^{-1}$ 也是可逆矩阵.

例 3.45

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

最后，除了上述介绍的方法之外，在接下来介绍了分块矩阵初等变换之后，我们还会介绍其它与逆矩阵相关的一些技巧.

3.4.5 矩阵方程

本节我们将讨论矩阵方程这一概念，即矩阵作为未知量的方程.

1. 设 A, B, C, X 为矩阵，且 A, B 可逆，考虑以下情形：

   (1) $AX = B \implies X = A^{-1}B$, $XA = B \implies X = BA^{-1}$;

   (2) $AXB = C \implies X = A^{-1}CB^{-1}$;

2. 考虑以下情形：$AX = B$ 但 A 不可逆（X 不一定是列向量），根据矩阵乘法的性质 "A 和 X 乘积的第 k 列等于 A 乘以 X 的第 k 列" 直接对 B 逐列解方程即可；

3. 考虑以下求解方式的合理性：

   (1) 若求 A^{-1}，只需对 $\left(\begin{array}{c:c} A & E \end{array}\right)$ 只做初等行变换，可以得到 $\left(\begin{array}{c:c} E & A^{-1} \end{array}\right)$;

   (2) 若求 $A^{-1}B$，只需对 $\left(\begin{array}{c:c} A & B \end{array}\right)$ 只做初等行变换，可以得到 $\left(\begin{array}{c:c} E & A^{-1}B \end{array}\right)$;

   (3) 若求 BA^{-1}，只需对 $\begin{pmatrix} A \\ \hdashline B \end{pmatrix}$ 只做初等列变换，可以得到 $\begin{pmatrix} E \\ \hdashline BA^{-1} \end{pmatrix}$;

   (4) 对 $\begin{pmatrix} A & E_n \\ E_n & O \end{pmatrix}$ 的前 n 行与 n 列做相同的行列变换，可以得到 $\begin{pmatrix} P^{\mathrm{T}}AP & P^{\mathrm{T}} \\ P & O \end{pmatrix}$.

例 3.46

设 $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$，求矩阵 X 满足：

$$AXA + BXB = AXB + BXA + A(A - B)$$

例 3.47

已知矩阵 $A = \begin{pmatrix} 1 & 2 & a \\ 1 & 3 & 0 \\ 2 & 7 & -a \end{pmatrix}$ 可以通过初等列变换转化为矩阵 $B = \begin{pmatrix} 1 & a & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

1. 求常数 a;

2. 求满足 $AP = B$ 的可逆矩阵 P.

3.5 矩阵转置

3.5.1 基本技巧

关于转置有一个经典的例题需要读者掌握，是一种矩阵求幂的基本技巧：

例 3.48

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

因此本题的关键在于，求矩阵的幂的时候，中间的项可以直接变成常数. 事实上，在将来讲解矩阵运算技巧时我们还会大量运用本题的技巧.

例 3.49

设 α,β 为三维列向量，且 $\alpha\beta^{\mathrm{T}}=\begin{pmatrix}-1&2&1\\1&-2&-1\\2&-4&-2\end{pmatrix}$，求 $\alpha^{\mathrm{T}}\beta$.

3.5.2 对称矩阵与反对称矩阵

事实上，关于对称矩阵和反对称矩阵的性质还有很多，我们将它们放在习题中供读者作为练习. 经过前面的讨论，我们已经看到转置和对称矩阵之间的关联，因此我们在之后在处理一些对称性很强的问题时，实际上都可以考虑利用转置来解决，例如：

例 3.50

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 $\begin{cases}a^2+c^2=1\\b^2+d^2=1\\ab+cd=0\end{cases}$.

例 3.51

1. 证明以下两个命题：

    (1) 证明：任一 n 阶方阵都可以表示为一个对称矩阵与一个反对称矩阵的和.

    (2) 设 A 是 n 阶复矩阵，若 $\overline{A}^{\mathrm{T}}=A$，则称 A 是一个 Hermite 矩阵. 若 $\overline{A}^{\mathrm{T}}=-A$，则称 A 是一个斜 Hermite 矩阵. 证明：任一 n 阶复矩阵都可以表示为一个 Hermite 矩阵与一个斜 Hermite 矩阵的和.

2. 证明以下两个命题:

(1) 设 A 为 n 阶对称矩阵，证明：A 是零矩阵的充要条件为对任意的 n 维向量 α，都有 $\alpha^{\mathrm{T}}A\alpha=0$.

(2) 设 A 为 n 阶方阵，证明：A 为反对称矩阵的充要条件为对任意的 n 维向量 α，都有 $\alpha^{\mathrm{T}}A\alpha=0$.

3. 设 A,B 为 n 阶方阵，证明:

(1) 若 A,B 为对称矩阵，则 AB 为对称矩阵的充要条件为 $AB=BA$，AB 为反对称矩阵的充要条件为 $AB=-BA$.

(2) 若 A 为对称矩阵，B 为反对称矩阵，则 AB 为反对称矩阵的充要条件为 $AB=BA$，AB 为对称矩阵的充要条件为 $AB=-BA$.

例 3.52

求矩阵 $\begin{pmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{pmatrix}$ 的逆.

3.5.3 矩阵的迹

我们首先定义方阵的迹，迹有很多优良的性质，并且能帮助我们解决一些有趣的问题，因此本节我们将对其展开讨论.

定义 3.7 迹

$A=(a_{ij})_{n\times n}$ 是 n 阶方阵，A 的主对角线上的元素之和称为 A 的**迹**，记为 $\mathrm{tr}(A)$，即

$$\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$$

定理 3.10

设 A 和 B 为 n 阶方阵，则

1. $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$;

2. $\mathrm{tr}(kA)=k\,\mathrm{tr}(A)$;

3. $\mathrm{tr}(A^{\mathrm{T}}) = \mathrm{tr}(A)$;
4. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

上述性质分别被称为迹的线性性(第 1、2 条)、对称性(第 3 条)和交换性(第 4 条). 其中交换性非常重要, 将来讨论矩阵相似的时候我们会讨论: $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(PP^{-1}A) = \mathrm{tr}(A)$(第一个等号将 $P^{-1}A$ 视为一个整体, 与 P 交换即可得到), 这里 P 是可逆矩阵, 这表明矩阵的迹在相似变换下保持不变. 下面我们给出另一个应用:

例 3.53

证明: 不存在 n 阶矩阵 A, B 使得 $AB - BA = kE_n$, 其中 $k \neq 0$.

矩阵的迹结合矩阵转置、对称矩阵可以得到一些有趣的性质, 我们来看一个例子:

例 3.54

证明以下两个命题:

1. 设 A 为 $m \times n$ 阶实矩阵, 则 $A^{\mathrm{T}}A = O$ 的充要条件为 $A = O$;
2. 设 A 为 $m \times n$ 阶复矩阵, 则 $\overline{A^{\mathrm{T}}}A = O$ 的充要条件为 $A = O$.

从这一例子的证明过程中我们可以得到一个显然的推论, 它表明了迹的正定性:

推论 3.1

1. 设 A 为 $m \times n$ 阶实矩阵, 则 $\mathrm{tr}(A^{\mathrm{T}}A) \geqslant 0$, 且 $\mathrm{tr}(A^{\mathrm{T}}A) = 0$ 当且仅当 $A = O$;
2. 设 A 为 $m \times n$ 阶复矩阵, 则 $\mathrm{tr}(\overline{A^{\mathrm{T}}}A) \geqslant 0$, 且 $\mathrm{tr}(\overline{A^{\mathrm{T}}}A) = 0$ 当且仅当 $A = O$.

下面这一例子综合运用了前面推导的迹的性质, 得到的结论也是简单而有趣的:

例 3.55

设 A 为 n 阶实矩阵, 证明: $\mathrm{tr}(A^2) \leqslant \mathrm{tr}(A^{\mathrm{T}}A)$, 等号成立当且仅当 A 为对称矩阵.

例 3.56

设 A, B 是两个 n 阶矩阵, 且 $\mathrm{tr}(ABC) = \mathrm{tr}(CBA)$ 对任意的 n 阶矩阵 C 成立, 证明 $AB = BA$.

例 3.57

1. 证明如下两个命题:

   (1) 设 A 是实对称矩阵，若 $A^2=O$，则 $A=O$；

   (2) 设 $A_1,A_2,\cdots,A_k$ 是实对称矩阵且 $A_1^2+A_2^2+\cdots+A_k^2=O$, 则 $A_1=A_2=\cdots=A_k=O$.

2. 证明如下两个命题:

   (1) 设 n 阶实矩阵 A 满足 $A^{\mathrm{T}}=-A$，如果存在同阶实矩阵 B 使得 $AB=B$，则 $B=O$；

   (2) 设 n 阶复矩阵 A 满足 $\overline{A^{\mathrm{T}}}=-A$，如果存在同阶复矩阵 B 使得 $AB=B$，则 $B=O$.

例 3.58

本题我们给出迹的另一种公理化的刻画，即我们可以从迹的性质出发，定义迹. 设 f 是数域 $\mathbf{F}$ 上 n 阶矩阵构成的空间到 $\mathbf{F}$ 的一个映射，满足

1. 对任意的 n 阶矩阵 A,B，有 $f(A+B)=f(A)+f(B)$;
2. 对任意的 n 阶矩阵 A 和数域中的数 k，有 $f(kA)=kf(A)$;
3. 对任意的 n 阶矩阵 A,B，有 $f(AB)=f(BA)$;
4. $f(E_n)=n$.

证明：存在唯一的映射 f 满足上述性质，且 $f(A)=\mathrm{tr}(A)$（提示：根据线性性，我们确定了 f 在一组基上的像即可确定 f）.

3.6 分块矩阵初等变换

3.6.1 分块矩阵的补充事项

1. 常见的行列分块方法：将矩阵按行/列分块，注意 $A(\beta_1,\ldots,\beta_n)=(A\beta_1,\ldots,A\beta_n)$ 成立，但当 A 在右侧时并不可乘，因为 β 是列向量，只有当 A 为行向量时才能使 βA 乘法是有意义的. 事实上按行分块也有对称的结论，即写成

$$A=\begin{pmatrix}A_1\\ \vdots\\ A_s\end{pmatrix}$$

时，我们有

$$AB = \begin{pmatrix} A_1B \\ \vdots \\ A_sB \end{pmatrix}.$$

2. 分块矩阵求逆通常有两种方法，其一直接使用设未知数的方式完成，我们下面将给出例子，当然也可以利用后续介绍的分块矩阵初等变换进行解决：

例 3.59

设 n 阶矩阵 A 分块为 $A = \begin{pmatrix} B & O \\ C & D \end{pmatrix}$，其中 B, D 分别为 k 阶、m 阶矩阵，求当 B, D 可逆时的 A^{-1}.

3. 分析分块矩阵与普通矩阵的运算性质的异同：

   (1) 分块矩阵转置需要注意大矩阵小分块都要转置；

   (2) 分块矩阵每一块不一定是数，而是矩阵，因此小分块中出现 $^{-1}$ 表示小分块求逆，但如果是一般矩阵就是矩阵元素直接求倒数即可；

   (3) 分块矩阵加法乘法一定要保证块大小对应，否则不可加、不可乘；

   (4) 其他很多性质都是将单个元素推广为一块，例如满足可加、可乘后的加法、乘法计算.

例 3.60

设

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 2 & 3 & 0 \\ 1 & 2 & 0 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

利用分块矩阵的方法，求 A^2，AB，A^{T}，A^{-1}.

在这个例子中我们可以得到一个很关键的经验：分块对角矩阵求逆实际上就是对每一个分块求逆.

例 3.61

设 B 为 n 阶可逆矩阵，$C=\begin{pmatrix}1 & 2 & \cdots & n\\ 0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\end{pmatrix}$，求一矩阵 A 使得 $A\begin{pmatrix}B\\ C\end{pmatrix}=E_n$.

例 3.62

设 A 是数域 $\mathbf{F}$ 上的 n 阶可逆矩阵，把 A 和 A^{-1} 做如下分块：

$$A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix},\quad A^{-1}=\begin{pmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{pmatrix}$$

其中 A_{11} 是 $l\times k$ 矩阵，B_{11} 是 $k\times l$ 矩阵，l，k 是小于 n 的正整数. 用 W 表示 $A_{12}X=0$ 的解空间，U 表示 $B_{12}Y=0$ 的解空间，其中 X 和 Y 为列向量，证明 W 与 U 同构.

3.6.2 分块矩阵初等变换的基本概念

分块矩阵的初等变换实际上可以视为一般矩阵初等变换的推广，实际上也有三种相应的推广形式：

1. 交换分块矩阵两分块行（列）(实际上对应于交换原矩阵若干行/列)；
2. 对分块的某一分块行（列），左（右）乘一个可逆矩阵（对应于普通矩阵初等变换就是对普通矩阵的一行乘以非零数)；
3. 将分块矩阵中的某一分块行（列），左（右）乘矩阵后加到另一分块行（对应于普通矩阵初等变换就是将一行乘以非零数加到另一行）.

回顾一般矩阵的初等矩阵，就是对单位矩阵做了一次初等变换得到的. 在分块初等矩阵中，我们记 E 为分块单位矩阵（事实上就是对单位矩阵做了分块，每个分块是阶数更小的单位矩阵)，于是三类分块矩阵初等变换对应的分块初等矩阵分别就是：

1. 对调 E 的第 i 个分块行（列）与第 j 个分块行（列）得到的矩阵；
2. 以可逆矩阵 C 左（右）乘 E 的第 i 个分块行（列）得到的矩阵；
3. 以矩阵 B 左乘 E 的第 i 个分块行加到第 j 个分块行得到的矩阵，或者以矩阵 B 右乘 E 的第 j 个分块列加到第 i 个分块列得到的矩阵.

容易验证分块初等矩阵都是可逆矩阵，并且矩阵的分块初等行（列）变换就相当于用同类分块初等矩阵左（右）乘以被变换的矩阵. 除此之外，我们还需要强调分块矩阵初等

变换也是不改变矩阵的秩的，实际上这一结论很直观，例如分块行（列）对换，实际上可以看成一次性做了多次普通的行列对换，其它情况也可以类似分析，严谨证明此处略去.

当然，在实际应用时我们一般只会出现 2×2 的情况，因此我们进行详细的讨论. 2×2 分块矩阵对应的三种分块初等矩阵为对单位矩阵

$$\begin{pmatrix} E & O \\ O & E \end{pmatrix}$$

做了三种初等变换得到的矩阵，即：

1. 交换分块矩阵的两行（列），对应的矩阵均为：

$$\begin{pmatrix} O & E \\ E & O \end{pmatrix}$$

该矩阵左（右）乘以分块矩阵相当于对分块矩阵交换两行（列）；

2. 倍乘矩阵：

$$\begin{pmatrix} C & O \\ O & E \end{pmatrix},\ \begin{pmatrix} E & O \\ O & C \end{pmatrix}$$

其中 C 为可逆矩阵，左（右）乘以上述第一个分块矩阵相当于对分块矩阵的第一行（列）乘以 C，第二个矩阵则对应第二行（列）的倍乘；

3. 倍加矩阵：

$$\begin{pmatrix} E & O \\ B & E \end{pmatrix},\ \begin{pmatrix} E & B \\ O & E \end{pmatrix}$$

(1) 左乘第一个矩阵相当于对分块矩阵的第一行乘以 B 后加到第二行，右乘第一个矩阵相当于对分块矩阵的第二列乘以 B 后加到第一列；

(2) 右乘第一个矩阵相当于对分块矩阵的第二列乘以 B 后加到第一列，左乘第一个矩阵相当于对分块矩阵的第一列乘以 B 后加到第二列.

事实上我们并不需要特别记忆，因为这和之前普通的初等变换并无本质区别，只需要注意左乘右乘即可. 除此之外，上述矩阵的逆矩阵也是同理可得的，只需要思考什么样的逆变换能变回单位矩阵即可，此处不再赘述，后面会有例题进行练习.

3.6.3 分块矩阵与逆矩阵求解

分块矩阵初等行变换的一个重要的应用就是“打洞法”，常用于分块矩阵求逆，在之后行列式的一些技巧性处理中也很常见. 这一方法非常形象，因为打洞就是使得矩阵出现一些零分块，成为分块对角矩阵（或者分块三角矩阵）后更加易于处理，例如：

1. 当 A 可逆时，我们可以通过初等行变换消去 C：

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

可以继续做列变换消去 B：

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

这里读者可能第一次接触这样的写法，笔者还是在此进行以下耐心的解释. 第一步我们希望消去 C，对于分块矩阵而言，由于已知 A 可逆，如果我们采用行变换，我们就给第一行左乘 $-CA^{-1}$ 使 A 变为 $-C$ 然后加到第二行，这样第二行的 C 就会变为 O. 这里的思考是很直接的，然后我们就可以根据我们之前想要做的行变换写出初等矩阵左乘在原矩阵上即可. 特别注意这里有两层左乘：第一层是小块内要左乘 $-CA^{-1}$，如果这里思考成了右乘就会错写为乘以 $-A^{-1}C$ 才能使 A 变为 $-C$，第二层是初等矩阵 $\begin{pmatrix} E & O \\ -CA^{-1} & E \end{pmatrix}$ 要左乘原分块矩阵. 第二步的思考是同理的，只是我们使用了列变换，需要注意右乘. 我们每一次的变换都希望将整个分块变为零矩阵，事实上这就像是在矩阵上挖了个洞填 0，因此打洞法这一名字的来由便显得更为清楚了.

2. 特别地，对于对称矩阵 $\begin{pmatrix} A & B \\ B^{\mathrm{T}} & D \end{pmatrix}$，其中 A 和 D 也是对称方阵，则 A 可逆时，可以通过下述变换（称为合同变换）消除 B 和 B^{T}，即

$$\begin{pmatrix} E & -A^{-1}B \\ O & E \end{pmatrix}^{\mathrm{T}}\begin{pmatrix} A & B \\ B^{\mathrm{T}} & D \end{pmatrix}\begin{pmatrix} E & -A^{-1}B \\ O & E \end{pmatrix} = \begin{pmatrix} A & O \\ O & D-B^{\mathrm{T}}A^{-1}B \end{pmatrix}$$

事实上，根据前述内容，我们能利用初等变换得到分块对角矩阵，这对于我们求解分块矩阵的逆十分有帮助. 回顾初等变换求解一般矩阵的逆的过程，主要过程就是

$$\left(\begin{array}{c:c} A & E \end{array}\right) \xrightarrow{\text{初等行变换}} \left(\begin{array}{c:c} E & A^{-1} \end{array}\right),$$

事实上这里的初等行变换换成分块初等行变换也完全没有问题，因为原理都是对 A 和 E 做了相同的初等行变换，当 A 变为 E 时，表明初等行变换综合作用效果为左乘 A^{-1}，因此单位矩阵此时就变成了 A^{-1}. 我们来看一个经典的例子：

例 3.63

当 D 和 $A-BD^{-1}C$ 可逆时，通过分块矩阵初等变换求 $P=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 的逆矩阵.

注意到本例中我们专门给出了条件 $A-BD^{-1}C$ 可逆，然后在结果中也看到这一形式的大量出现. 基于这一性质，我们称其为 P 矩阵中 D 矩阵的 Schur 补（因为此时 D 可逆，若 A 可逆完全可以定义出 A 的 Schur 补），记为 P/D. 于是矩阵 P 的逆可以写成

$$\begin{pmatrix} (P/D)^{-1} & -(P/D)^{-1}BD^{-1} \\ -D^{-1}C(P/D)^{-1} & D^{-1}+D^{-1}C(P/D)^{-1}BD^{-1} \end{pmatrix}.$$

我们在之后讨论分块矩阵行列式的相关运算技巧时还会遇到 Schur 补，由此可见其在矩阵计算中的重要性. 我们会在很多习题中遇到类似于 Schur 补的形式，希望读者都能认出并指明其本质只是为了打洞而做初等变换得到的.

例 3.64

求下列矩阵的可逆的条件与逆矩阵：

1. $\begin{pmatrix} A & B \\ O & D \end{pmatrix}$，其中 A 是 m 阶方阵，D 是 n 阶方阵，B 是 $m\times n$ 矩阵；
2. $\begin{pmatrix} O & B \\ C & D \end{pmatrix}$，其中 B 是 n 阶方阵，C 是 m 阶方阵，D 是 $n\times m$ 矩阵.

例 3.65

设 A,B 为 n 阶矩阵，证明：$E\pm AB$ 可逆 $\Longleftrightarrow$ $E\pm BA$ 可逆.

例 3.66

设 A 为 n 阶矩阵，B,C 分别为 $n\times m$ 和 $m\times n$ 阶矩阵. 证明：$E_m+CA^{-1}B$ 可逆 $\Longleftrightarrow$ $A+BC$ 可逆.

事实上，总结上述两题的解决方法，都是将待证明的一个矩阵构成分块矩阵的一部分，然后利用初等变换变出另一个矩阵，使得这两个矩阵的可逆性相同，从而得到结论.

3.6.4 分块矩阵与数学归纳法

分块矩阵经常运用在数学归纳法中，我们在之后的课程中也会经常用到这样的思想来证明一些结论，这一思想基于以下内容：

对于 $\begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}$，假设 A_1 可逆，我们有

$$\begin{pmatrix} E_{n-1} & O \\ -\beta A_1^{-1} & 1 \end{pmatrix}\begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}=\begin{pmatrix} A_1 & \alpha \\ O & a_{nn}-\beta A_1^{-1}\alpha \end{pmatrix}$$

需要注意的是，通过这一变换我们也可以知道，当 A_1 可逆时矩阵 $\begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}$ 可逆的充要条件为 $a_{nn} \neq \beta A_1^{-1}\alpha$（因为矩阵满秩当且仅当最后一行不等于零）. 我们通过一个例子来体会如何利用上式结合数学归纳法得到一些矩阵分析中的结论：

例 3.67

若 n 阶矩阵 A 的各阶左上角子块矩阵都可逆，则存在主对角元全为 1 的下三角矩阵 L 和上三角矩阵 U，使得 $A = LU$（LU 分解）.

例 3.68

证明：若 n 阶矩阵 A 的各阶左上角子块矩阵都可逆，则存在 n 阶下三角矩阵 B，使得 BA 为上三角矩阵.

3.7 矩阵的秩与相抵标准形

3.7.1 矩阵的秩

首先给出矩阵三个秩的定义：

定义 3.8

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}.

这里矩阵的秩的定义存在两个问题：

1. 同一个矩阵 A 可以是不同线性映射的矩阵表示，那么是否有可能这些线性映射的秩不同？答案是否定的. 回忆线性映射矩阵表示的定义方式，要得到矩阵 A，我们需要将线性映射 σ 通过两个坐标映射转化为 $\mathbf{F}^n \to \mathbf{F}^m$ 的映射 $\sigma_{\mathbf{A}}$. 注意无论是什么坐标映射，都是同构映射，而线性映射的秩是线性映射像空间的维数，在同构变换下像空间维数都不变，因此保证了 σ 和 $\sigma_{\mathbf{A}}$ 的秩相等，故所有可以表示 A 的线性映射都有相同的秩；

2. 同一个映射 σ 可以在不同基下得到不同的矩阵表示，故这里的定义表明，所有的矩阵表示必须有相同的秩. 理由和前一点类似，因为不同基下的矩阵表示无非是使用了不同的坐标映射，不改变导出的 $\mathbf{F}^n \to \mathbf{F}^m$ 映射的秩.

对于以上定义的三个秩，我们有如下定理，这一定理无论是证明还是结果都非常关键：

定理 3.11

任意矩阵 $A=(a_{ij})_{m\times n}$ 的秩 = 行秩 = 列秩.

关于这一定理，我们有以下几点补充说明:

1. 矩阵的秩等于列秩的证明我们复习了同构的性质. 事实上这一结论还可以告诉我们，无论是 σ 在哪组基下的表示矩阵，都有相同的秩；除此之外，这一定理使得我们可以将求矩阵的秩的问题转化为求矩阵行/列极大线性无关向量组的问题；
2. 行秩等于列秩的第一种证明给了我们两个启示:
   (1) 我们在证明过程第二步证明反向不等式的时候直接考察了转置矩阵得出结论，这一思想在将来一些秩的等式/不等式的证明中也是常见的，因为转置就是将行和列互换，所以特别适合于这种证明；
   (2) $r(A)=r(A^{\mathrm{T}})$，即矩阵转置不改变矩阵的秩. 事实上根据这一定理我们有 A^{T} 的行秩 $=A$ 的列秩 $=A$ 的秩 $=A$ 的行秩 $=A^{\mathrm{T}}$ 的列秩 $=A^{\mathrm{T}}$ 的秩.

   除此之外，我们可以仔细品味以下行秩 = 列秩这一结论. 这表明我们随手写任意一个矩阵，它行向量组的秩和列向量组的秩就一定是相等的——明明是很杂乱无章的数字排列，却有这么一个和谐而美观的性质，着实令人赞叹. 事实上，行秩 = 列秩还有更深层的含义等待我们在后续章节中逐步揭示，届时我们也将给读者一个比较完整的对矩阵转置的理解.

除此之外，我们还应强调以下结论，在后续研究线性方程组解的性质时是常用的:

定理 3.12

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.

我们需要注意，虽然之前证明矩阵的秩 = 列秩时我们将列秩和像空间的维数等同，但这里列满秩是和单射等同的，这是因为列数一定的情况下，只有当列满秩时像空间的维数才能达到秩的最大，不要混淆.

最后，通过矩阵的秩的学习我们可以总结可逆矩阵的几个等价条件，它们非常重要，在未来的讨论以及习题中非常常见:

定理 3.13

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.

事实上，在学完行列式后这一命题还可以增加一个行列式 $|A|\neq 0$ 的等价条件. 下面我们来看一个重要的应用，一方面是这一定理的直接应用，更重要的是我们将引入“对角占优矩阵”，这是一个在实际应用中非常常见的矩阵类型.

例 3.69　对角占优

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.

本题使用了齐次线性方程组只有零解这一等价条件，非常有趣. 实际上这一等价条件是看起来最不能直接应用的，但不得不承认在很多问题中它是最有用的.

例 3.70

设 A 为 n 阶实反对称矩阵，证明：$E-A$ 可逆.

例 3.71

判断：设 A_1，A_2，$\ldots$，A_{n+1} 是任意 $n+1$ 个 n 阶矩阵，必存在不全为 0 的实数 λ_1，λ_2，$\ldots$，λ_{n+1}，使得矩阵 $\lambda_1A_1+\lambda_2A_2+\cdots+\lambda_{n+1}A_{n+1}$ 不可逆.

例 3.72

若 n 阶方阵 $A_1,A_2,\ldots,A_m$ 满足 $A_i^2\neq O\ (i=1,2,\ldots,m)$，且当 $i\neq j$ 时 $A_iA_j=O$，证明：$m\leqslant n$.

例 3.73

证明如下结论：

1. 若存在正整数 m 使得 $r(A^m)=r(A^{m+1})$，则必有 $r(A^m)=r(A^{m+1})=r(A^{m+2})=\cdots$;
2. 设 A 为 n 阶方阵，证明：$r(A^n)=r(A^{n+1})=r(A^{n+2})=\cdots$.

3.7.2 过渡矩阵和基变换

这一讲我们将讨论和基变换相关的几个定理. 在讨论这些定理前我们首先介绍过渡矩阵（变换矩阵）的概念.

考虑两个同一线性空间内的向量组 $S_1=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 和 $S_2=\{\beta_1,\beta_2,\ldots,\beta_m\}$, 且有

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n \\ \beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n \\ \quad\vdots \\ \beta_m = a_{1m}\alpha_1 + a_{2m}\alpha_2 + \cdots + a_{nm}\alpha_n \end{cases}$$

则上述关系可以用矩阵简单表达为

$$(\beta_1,\beta_2,\ldots,\beta_m)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

令 $A=(a_{ij})_{n\times m}$，则上式可以进一步简化为

$$S_2 = S_1 A.$$

当 S_1 和 S_2 是同一线性空间的两组基时，我们称 A 为 S_1 变为 S_2 的过渡矩阵，下面我们正式地叙述这一定义.

定义 3.9

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**变换矩阵**（或**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 $\mathbf{M}_{B_2,B_1}(I)$. 根据式 3.1, $\mathbf{M}_{B_2,B_1}(I)\mathbf{M}_{B_1,B_2}(I) = \mathbf{M}_{B_1,B_1}(I) = E$, 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.

上面的讨论告诉我们, 两组基（更一般地, 两个相同长度的线性无关向量组）之间的过渡矩阵可逆. 下面我们要讨论的是, 如果已知向量组 S_1, S_2 有关系 $S_2 = S_1A$, 其中 A 是可逆矩阵, S_1 线性无关, 那么 S_2 是否线性无关? 答案是肯定的. 我们首先给出一个更一般的定理, 其推论就是这一问题的答案.

定理 3.14

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.

因此若 A 可逆, 此时 $s = n$, 根据定理可知 $r(S_2) = r(A) = n$, 即 S_2 也线性无关. 下面的定理扩展了这一结论, 讨论了两个不一定线性无关的向量组之间的关系.

定理 3.15

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.

我们来看一些例子应用上述定理:

例 3.74

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 3.75

证明: 当 n 为奇数时, $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关的充要条件是 $\alpha_1+\alpha_2, \alpha_2+\alpha_3, \ldots, \alpha_n+\alpha_1$ 线性无关.

例 3.76

设 A 和 B 是两个 $m \times n$ 矩阵, 且存在两个方阵 C, D 使得 $A = BC, B = AD$, 证明: 存在可逆矩阵 M 使得 $B = AM$.

有了上述内容的铺垫, 我们可以开始介绍本节第一个重要定理:

定理 3.16 基的选择对向量坐标的影响

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 $Y = A^{-1}X$.

上述定理描述了**同一个向量在不同基下坐标之间的关系**.

有了过渡矩阵的定义，我们可以介绍第二个重要的定理，也是本节的核心：我们想知道**同一个线性映射在不同的基下的矩阵表示之间有什么关系**，下面的定理回答了这一问题：

定理 3.17 换基公式

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 $\mathbf{M}_{B_1', B_2'}(\sigma) = Q^{-1}\mathbf{M}_{B_1, B_2}(\sigma)P$.

这一换基公式将在后续讨论相抵标准形的时候体现其意义. 下面我们给出一个推论，即一种特殊情况，此时线性映射我们要求是线性变换（即 $V_1 = V_2$），并且 $B_1 = B_1', B_2 = B_2'$，这将引出未来相似标准形的讨论：

定理 3.18 基的选择对变换矩阵的影响

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 3.77

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.

3.7.3 相抵标准形：映射角度

接下来我们将开始讨论矩阵的第一个标准形——相抵标准形，我们将首先从线性映射的角度展开讨论. 事实上，我们讨论标准形的目标就是使得线性映射矩阵表示越简单越好，这样将便于我们的计算与研究. 所谓的简单，就是我们希望矩阵中尽可能多的 0，然后其它的元素也尽可能排列规律. 相抵标准形就是这样一种形式：

定理 3.19

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.

这一定理也就自然给出了相抵以及相抵标准形的定义：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，$r=r(A)$. 此后我们还会基于初等变换讨论矩阵的相抵标准形，届时我们再展开讨论“相抵”这一概念.

基于上述定理可以给出矩阵行秩 = 列秩的第三种证明.

例 3.78

设矩阵空间 $\mathbf{R}^{2\times 2}$ 的子空间为

$$V=\{X=(x_{ij})_{2\times 2}\mid x_{11}+x_{12}+x_{21}=0,\ x_{ij}\in\mathbf{R}\}$$

V 中的线性变换为 $\sigma(X)=X+X^{\mathrm{T}}$，求 V 的一组基，使得 σ 在该基下的矩阵表示为对角矩阵.

3.7.4 从初等变换到相抵标准形

接下来我们将从矩阵的角度得到相抵标准形，我们会首先给出推导的思路，然后从线性映射的角度出发再看初等变换，加深我们的理解.

定理 3.20

初等变换不改变矩阵的秩（包括行变换和列变换）.

定理的证明很简单，只需对各个初等变换逐一通过计算验证即可，由这一定理我们同样可以证明定理 3.19.

根据上面的描述，我们可以给出相抵的等价定义：

定义 3.10

我们称两个矩阵相抵即两个矩阵可以通过一系列初等变换可以互相转化.

根据前面的讨论，我们总结出以下几点：

1. 根据定理 3.19，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 $\iff$ 存在可逆矩阵 P 和 Q 使得 $PAQ=B$. 原因很简单，只需要利用定理 3.3，可逆矩阵一定能拆分成若干初等变换的乘积，因此我们可以将 P 和 Q 拆分，那么上面的结论就转化为矩阵相抵的定义；

3. 矩阵 A 与 B 相抵 $\iff$ $r(A)=r(B)$. 只需利用定理 3.20 就能轻松得到结论；

   我们在这里对初等变换做一个小小的总结. 事实上初等变换只有三个非常重要的性质，即初等变换可逆，可逆矩阵可以写为初等变换的乘积，以及初等变换不改变矩阵的秩，只需牢记这三点就能覆盖几乎全部的证明技巧.

4. 事实上，相抵也被称为等价，相抵标准形也被称为等价标准形，原因就在于相抵是矩阵的一个等价关系；证明如下：

   (1) $E_mAE_n=A$, 即 $A\cong A$;

   (2) 若 $A\cong B$, 则存在可逆矩阵 P 和 Q, 使得 $PAQ=B$, 于是, $P^{-1}BQ^{-1}=A$, 从而 $B\cong A$;

   (3) 若 $A\cong B, B\cong C$, 则存在可逆矩阵 P,S,Q,T, 使得 $PAQ=B, SBT=C$, 于是就有 $(SP)A(QT)=C$, 从而 $A\cong C$.

   这里要强调的是，这一等价关系将矩阵空间 $\mathbf{F}^{m\times n}$ 中的全体元素按秩进行了分类，每一类对应的相抵标准形都是一样的.

例 3.79

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

3.7.5 基变换与初等变换

我们先来看下面这个例子:

例 3.80

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 σ 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.

这一例子给了我们一个很大的启示:我们可以将对于基向量组的上述操作也视为一种初等变换,那么对基的初等变换的作用效果就是表示矩阵也做了初等变换. 其中对出发空间的基做变换相当于对矩阵的列做变换;对目标空间的基做变换相当于对矩阵的行做变换.

因此矩阵的初等变换实际上并不改变背后的线性映射,因此初等变换不改变矩阵的秩也是显然可以达成的,而我们在线性映射证明相抵标准形的过程中也知道,一定存在一组基使得线性映射在这组基下的矩阵表示是相抵标准形,那么我们可以从任意一组基出发,通过基的初等变换得到目标基,在这一过程中矩阵表示也就实现了初等变换得到相抵标准形的目标.

在线性代数的学习中我们会介绍矩阵的三种标准形,每种标准形都是矩阵在某一特定的基下的矩阵表示. 事实上我们可以首先写出矩阵在任意一组基下的矩阵表示,然后通过对矩阵做初等变换得到标准形,同时对基做“初等变换”即可得到标准形对应的基. 日后我们提到矩阵标准形时再进一步讨论这一点,现在可以只留一个印象.

3.7.6 相抵标准形的应用

基于相抵标准形的分解是很重要的技术,也带来了相抵标准形的一些有趣的应用. 事实上将来讨论其它标准形时我们都会讨论分解问题,因为这能在实际问题中大大降低计算

难度，便于我们进一步讨论.

第一种分解是非常自然的分解，将来相似、相合标准形也会基于这一思想进行分解. 我们知道，对于矩阵 A 满足 $r(A)=r$，则存在可逆矩阵 P' 和 Q'，使得

$$P'AQ'=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix},$$

因此我们可以得到 $A=P'^{-1}\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q'^{-1}$，即 A 可以分解成一个可逆矩阵、一个相抵标准形和另一个可逆矩阵的乘积. 记 $P=P'^{-1}$，$Q=Q'^{-1}$，则 $A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$. 我们接下来给出一些经典的的例子供读者体会.

例 3.81

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 $A=BC$.

从本例可以看出，很多问题我们需要首先写出分解，然后利用分解去找到符合题目要求的矩阵来证明（特别利用是方阵的相抵标准形有很多好的性质，如上面的幂等和对称），在习题中我们会看到更多这样的问题.

另一种分解技巧是更进一步的，此时我们不仅对原矩阵分解，还对相抵标准形做进一步的分解. 我们对 $s\times n$ 矩阵 $\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$ 有一种很重要的分解：

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

由此我们可以知道任意一个非零矩阵都可以被分解成一个列满秩矩阵和一个行满秩矩阵的乘积：

$$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q=P\begin{pmatrix} E_r \\ O \end{pmatrix}\begin{pmatrix} E_r & O \end{pmatrix}Q$$

记 $P_1=P\begin{pmatrix} E_r \\ O \end{pmatrix}$，$Q_1=\begin{pmatrix} E_r & O \end{pmatrix}Q$，则 $A=P_1Q_1$，且 P_1 和 Q_1 分别为列满秩、行满秩矩阵.

我们简要解释 P_1 列满秩的原因，Q_1 行满秩类似不再赘述. 由于 $\begin{pmatrix} E_r \\ O \end{pmatrix}$ 是 $s\times r$ 矩阵，且秩为 r，列满秩. P 可逆且为 $s\times s$ 矩阵，因此 P_1 仍然是 $s\times r$ 矩阵. 由于可逆矩阵可

以写成若干初等矩阵乘积，初等变换不改变矩阵的秩，故 $r(P)=r(P_1)=r$，又矩阵列秩 $=$ 秩，故 P_1 列满秩.

例 3.82

设 A,B 分别为 3×2 和 2×3 实矩阵. 若 $AB=\begin{pmatrix} 8 & 0 & -4 \\ -\dfrac{3}{2} & 9 & -6 \\ -2 & 0 & 1 \end{pmatrix}$，求 BA.

接下来来看一个重要的关于矩阵的迹的应用：

例 3.83

已知 n 阶矩阵 A 的秩为 1 ，证明：$A^k=\operatorname{tr}(A)^{k-1}A$，其中 k 为任意正整数.

事实上，本题的解答过程给我们了一个很重要的启示，那就是秩为 1 的矩阵一定可以分解为一个列向量和一个行向量的乘积. 之后我们将利用这一结论来解决一些问题，例如便于矩阵求幂等. 此外，若 $\operatorname{tr}(A)=1$，则 $A^k=A$，这表明秩为 1 且迹为 1 的矩阵一定是幂等矩阵.

除此之外，我们还可以利用相抵标准形解决很多问题，例如部分秩不等式的证明.

例 3.84

1. 利用相抵标准形证明以下结论：

   (1) 设 B_1,B_2 为 $s\times n$ 列满秩矩阵，证明：存在 s 阶可逆矩阵 C 使得 $B_2=CB_1$；

   (2) 设 B_1,B_2 为 $s\times n$ 行满秩矩阵，证明：存在 n 阶可逆矩阵 C 使得 $B_2=B_1C$；

   (3) 任意秩为 r 的矩阵都可以被分解为 r 个秩为 1 的矩阵之和；

   (4) 已知 A 是 n 阶方阵，证明：存在 n 阶方阵 B 使得 $A=ABA$，$B=BAB$.

2. 设 $A\in\mathbf{M}_{m\times n}(\mathbf{F})$，$r(A)=r$，$k$ 是满足条件 $r\leqslant k\leqslant n$ 的任意整数，证明存在 n 阶方阵 B，使得 $AB=O$，且 $r(A)+r(B)=k$.

3. 设 A 是 $m\times n$ 矩阵（$m\leqslant n$），$r(A)=m$，证明：存在 $n\times m$ 矩阵 B 使得 $AB=E$.

4. 设 $A,B\in\mathbf{M}_n(\mathbf{F})$，$r(A)+r(B)\leqslant n$，证明：存在可逆矩阵 M，使得 $AMB=O$.

5. 设 A 为 n 阶实方阵且 $r(A)=r>0$，证明存在秩为 r 的实方阵 B 和 C 使得 $AB=CA$.

3.8 矩阵秩不等式

本节我们将讨论一些秩的等式与不等式，事实上有一定的难度，不仅在于技巧也在于理解. 一般而言，解决较为复杂的秩的问题时，我们可以采用如下方法：

1. 回到线性映射的视角进行考察，证明不等式的线性映射版本；
2. 利用向量组线性相关性：因为行秩和列秩的定义就是基于向量组线性相关性的；
3. 利用线性方程组解的一般理论（将在专题五讲解）；
4. 利用（分块）矩阵初等变换：这一方法基于分块矩阵初等变换也是不改变矩阵的秩这一事实；
5. 利用已知的矩阵秩的等式和不等式；
6. 如果证明的是等式，我们考虑初等变换不改变矩阵的秩（推论就是乘以可逆矩阵也不改变，下面将会证明），也经常用两个不等号夹逼得到等号.

我们首先给出一些最常见的秩相关的不等式或等式，希望读者能熟练推导理解.

1. $r(A) = r(PA) = r(AQ) = r(PAQ)$，其中 P, Q 可逆.

   注：这一结论非常重要，即可逆矩阵乘以（不管左乘还是右乘）任何矩阵都是不改变矩阵的秩的. 基于此，我们可以推导如下分块矩阵秩的相关公式：

   例 3.85

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

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

   注意第二个等号需要实矩阵作为前提条件，等式证明我们将在线性方程中一章中讲解.

5. (Sylvester 不等式) $A \in \mathbf{F}^{s\times n}$, $B \in \mathbf{F}^{n\times m}$, 则 $r(AB) \geqslant r(A) + r(B) - n$

   这一不等式有一个特例，即当 $AB = O$ 时有 $r(A) + r(B) \leqslant n$. 这一结论之后还有其他方法可以给出证明.

6. (Frobenius 不等式) $r(ABC) \geqslant r(AB) + r(BC) - r(B)$

   利用这一不等式，我们还可以得到一种特例，即 A, B, C 相等的特殊情况：

   $$r(A^3) \geqslant 2r(A^2) - r(A).$$

   除此之外，若 $B = E_n$ 即单位矩阵时我们有 $r(AC) \geqslant r(A) + r(C) - n$. 因此只要证明了这一个不等式，很多的结论都只是其推论而已.

在下面的例子以及习题中我们将给出更多的例子供读者熟练上面的证明思想与技巧.

例 3.86

1. 证明：矩阵添加一列（或一行），其秩或不变，或增加 1.
2. 设 A 是 $s\times n$ 矩阵, B 是 A 前 m 行构成的 $m\times n$ 矩阵, 证明: $r(B) \geqslant r(A)+m-s$.

例 3.87

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

例 3.88

设 B 是 3×1 矩阵，C 是 1×3 矩阵，证明：$r(BC) \leqslant 1$. 反之，若 A 是秩为 1 的 3×3 矩阵，证明：存在 3×1 矩阵 B 和 1×3 矩阵 C，使得 $A = BC$.

例 3.89

设 α, β 为 n 维列向量，且 $A = \alpha\alpha^{\mathrm{T}} + \beta\beta^{\mathrm{T}}$.

1. 证明：$r(A) \leqslant 2$;
2. 若 α, β 线性相关，证明：$r(A) \leqslant 1$.

例 3.90

设 A 和 B 是两个 n 阶实方阵，证明：$r(A)=r(AB)$ 当且仅当存在 n 阶实方阵 C 使得 $A=ABC$.

例 3.91

设 $n\geqslant 3$，证明下列矩阵不可逆：

$$A=\begin{pmatrix}\cos(\alpha_1-\beta_1) & \cos(\alpha_1-\beta_2) & \cdots & \cos(\alpha_1-\beta_n)\\ \cos(\alpha_2-\beta_1) & \cos(\alpha_2-\beta_2) & \cdots & \cos(\alpha_2-\beta_n)\\ \vdots & \vdots & \ddots & \vdots\\ \cos(\alpha_n-\beta_1) & \cos(\alpha_n-\beta_2) & \cdots & \cos(\alpha_n-\beta_n)\end{pmatrix},$$

$$B=\begin{pmatrix}1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n\\ 1+x_2y_1 & 1+x_2y_2 & \cdots & 1+x_2y_n\\ \vdots & \vdots & \ddots & \vdots\\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n\end{pmatrix}.$$

例 3.92

1. 已知 A 是一个 $s\times n$ 矩阵，证明：$r(E_n-A^{\mathrm{T}}A)-r(E_s-AA^{\mathrm{T}})=n-s$.
2. 利用打洞法完成以下两个问题（(2) 也可以不使用打洞法，可以思考其他方式解决）：

   (1) 设 n 阶方阵 A,B,C,D 满足 $AC+BD=E$,证明:$r(AB)=r(A)+r(B)-n$;

   (2) n 阶方阵 A,B 满足 $AB=BA$，证明：$r(AB)+r(A+B)\leqslant r(A)+r(B)$.
3. $f(x)=f_1(x)f_2(x)$ 是多项式，且 $f_1(x)$ 与 $f_2(x)$ 互素，则 $f(A)=O$ 的充要条件是 $r(f_1(A))+r(f_2(A))=n$.（注：此题的推论非常多，如 $A^2=A$，$A^n=E$ 等形式的结论都可以利用这个例子推导出）