破局五　矩阵乘法与初等变换、逆矩阵与伴随矩阵

本部分内容对应课本 3.2~3.5.

对于矩阵的初等变换，我们或多或少已经有了一点了解. 至少在解线性方程组的时候，矩阵的初等变换用了不少吧？我们已经知道，矩阵乘法实际上是一种线性映射，可以实现向量的变换. 那么矩阵的初等变换是不是也可以和矩阵乘法起联系呢？

另外，我还会谈一谈逆矩阵和伴随矩阵的事儿. 逆矩阵为什么叫“逆”矩阵？既然矩阵乘法是一种线性映射，那么逆矩阵是不是和逆映射有关呢？

1. 逆矩阵

在数的乘法中，我们用到过倒数的概念. 如果两个数 a,b 满足 $ab=1=ba$，那么 a 和 b 互为倒数，可以说 a 是 b 的倒数，同时 b 也是 a 的倒数. a 的倒数记作 $\frac{1}{a}$ 或 a^{-1}.

将其推广到矩阵，我们给出逆矩阵的定义. 如果两个矩阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{E}=\boldsymbol{BA}$，那么 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 互为逆矩阵，可以说 $\boldsymbol{A}$ 是 $\boldsymbol{B}$ 的逆矩阵，同时 $\boldsymbol{B}$ 也是 $\boldsymbol{A}$ 的逆矩阵. $\boldsymbol{A}$ 的逆矩阵记作 $\boldsymbol{A}^{-1}$.

由于数的乘法满足交换律，即 $ab=ba$，所以在倒数的定义中，$ab=1=ba$ 可以简化，只需要 $ab=1$ 即可. 那么逆矩阵是否也满足这样？

虽然矩阵乘法并不满足交换律，但是在逆矩阵的定义中，$\boldsymbol{AB}=\boldsymbol{E}=\boldsymbol{BA}$ 可以简化成 $\boldsymbol{AB}=\boldsymbol{E}$ 或 $\boldsymbol{BA}=\boldsymbol{E}$. 这里不作证明，可参考课本 P49 定理 4.

一个重要定理：对 n 阶方阵 $\boldsymbol{A}$ 而言，有

$$\boldsymbol{A}\text{ 可逆} \Leftrightarrow |\boldsymbol{A}|\neq 0 \Leftrightarrow r(\boldsymbol{A})=n.$$

我们同样略去了证明，证明过程可参考课本 P48 定理 3. 关于逆矩阵的运算性质，这里不作赘述，可参考课本 P49.

2. 用矩阵的观点看初等变换

我们知道，任何一个矩阵乘上一个单位阵 $\boldsymbol{E}$ 后都是它本身，那么是否可以对单位阵 $\boldsymbol{E}$ 做一些小小的改动，让它对矩阵产生某种效果？

例 1　已知 $\boldsymbol{A}=(a_{ij})_{3\times3}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$，$\boldsymbol{E}_{12}=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$，求 $\boldsymbol{E}_{12}\boldsymbol{A}$ 和 $\boldsymbol{AE}_{12}$.

解　由矩阵乘法的定义易得，$\boldsymbol{E}_{12}\boldsymbol{A}=\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$，$\boldsymbol{A}\boldsymbol{E}_{12}=\begin{pmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{pmatrix}$.

首先分析一下 $\boldsymbol{E}_{12}$ 是怎么来的. 它可以看成单位阵 $\boldsymbol{E}$ 交换了第 1 行和第 2 行的结果，也可以看成是单位阵 $\boldsymbol{E}$ 交换了第 1 列和第 2 列的结果. 将 $\boldsymbol{E}_{12}$ 乘在 $\boldsymbol{A}$ 的左边，实现了对 $\boldsymbol{A}$ 的第 1 行和第 2 行的互换操作；而将 $\boldsymbol{E}_{12}$ 乘在 $\boldsymbol{A}$ 的右边，就实现了对 $\boldsymbol{A}$ 的第 1 列和第 2 列的互换操作.

例 2　已知 $\boldsymbol{A}=(a_{ij})_{3\times3}$，$\boldsymbol{E}_2(c)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$，求 $\boldsymbol{E}_2(c)\boldsymbol{A}$ 和 $\boldsymbol{A}\boldsymbol{E}_2(c)$.

解　$\boldsymbol{E}_2(c)\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ ca_{21} & ca_{22} & ca_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$，$\boldsymbol{A}\boldsymbol{E}_2(c)=\begin{pmatrix} a_{11} & ca_{12} & a_{13} \\ a_{21} & ca_{22} & a_{23} \\ a_{31} & ca_{32} & a_{33} \end{pmatrix}$.

同样，这里的 $\boldsymbol{E}_2(c)$ 既可以看成是单位阵 $\boldsymbol{E}$ 的第 2 行乘了一个常数 c，也可以看成是单位阵 $\boldsymbol{E}$ 的第 2 列乘了一个常数 c. 将 $\boldsymbol{E}_2(c)$ 乘在 $\boldsymbol{A}$ 的左边，对 $\boldsymbol{A}$ 的第 2 行进行了倍乘；而将 $\boldsymbol{E}_2(c)$ 乘在 $\boldsymbol{A}$ 的右边，就对 $\boldsymbol{A}$ 的第 2 列进行了倍乘.

例 3　已知 $\boldsymbol{A}=(a_{ij})_{3\times3}$，$\boldsymbol{E}_{13}(c)=\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$，求 $\boldsymbol{E}_{13}(c)\boldsymbol{A}$ 和 $\boldsymbol{A}\boldsymbol{E}_{13}(c)$.

解　$\boldsymbol{E}_{13}(c)\boldsymbol{A}=\begin{pmatrix} a_{11}+ca_{31} & a_{12}+ca_{32} & a_{13}+ca_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$，$\boldsymbol{A}\boldsymbol{E}_{13}(c)=\begin{pmatrix} a_{11}+ca_{13} & a_{12} & a_{13} \\ a_{21}+ca_{23} & a_{22} & a_{23} \\ a_{31}+ca_{33} & a_{32} & a_{33} \end{pmatrix}$.

同样，这里的 $\boldsymbol{E}_{13}(c)$ 可以看成 $\boldsymbol{E}$ 经过了 R_1+cR_3（行）的倍加，也可以看成 $\boldsymbol{E}$ 经过了 C_1+cC_3（列）的倍加. 将 $\boldsymbol{E}_{13}(c)$ 乘在 $\boldsymbol{A}$ 的左边，就对 $\boldsymbol{A}$ 进行了 R_1+cR_3（行）的倍加；而将 $\boldsymbol{E}_{13}(c)$ 乘在 $\boldsymbol{A}$ 的右边，就对 $\boldsymbol{A}$ 进行了 C_1+cC_3（列）的倍加.

上面三个例题中出现的 $\boldsymbol{E}_{12}$（互换），$\boldsymbol{E}_2(c)$（倍乘），$\boldsymbol{E}_{13}(c)$（倍加）都是单位阵 $\boldsymbol{E}$ 经过一次初等变换得到的，因此叫初等矩阵. 将初等矩阵乘在矩阵 $\boldsymbol{A}$ 的左边，相当于对 $\boldsymbol{A}$ 进行了一次初等行变换；而将初等矩阵乘在矩阵 $\boldsymbol{A}$ 的右边，就相当于对 $\boldsymbol{A}$ 进行了一次初等列变换.

3. 用初等行变换求逆矩阵

任意给定一个满秩的方阵 $\boldsymbol{A}$，你能用初等行变换把它化成单位阵 $\boldsymbol{E}$ 吗？这是很容易的吧？如果你一下子反应不过来，那就想想，$\boldsymbol{E}$ 不是一种非常特殊的阶梯型矩阵吗？把一个矩阵化成阶梯型矩阵，想必你已经很熟练了吧？

如果用矩阵乘法表示上面所说的这个过程，那就是

$$\boldsymbol{L}_1\boldsymbol{L}_2\ldots\boldsymbol{L}_k\boldsymbol{A}=\boldsymbol{E}. \qquad (5.1)$$

因为 $\boldsymbol{A}$ 满秩，所以 $\boldsymbol{A}$ 可逆，根据逆矩阵的定义，$\boldsymbol{L}_1\boldsymbol{L}_2\ldots\boldsymbol{L}_k=\boldsymbol{A}^{-1}$. 在等号左边乘一个单位阵 $\boldsymbol{E}$，等号依然成立，所以

$$\boldsymbol{L}_1\boldsymbol{L}_2\ldots\boldsymbol{L}_k\boldsymbol{E}=\boldsymbol{A}^{-1}. \qquad (5.2)$$

利用分块矩阵，将式（5.1）和（5.2）写成一个式子

$$\boldsymbol{L}_1\boldsymbol{L}_2\ldots\boldsymbol{L}_k(\boldsymbol{A}\mid\boldsymbol{E})=(\boldsymbol{E}\mid\boldsymbol{A}^{-1}).$$

也就是说

$$(\boldsymbol{A}\mid\boldsymbol{E})\xrightarrow{\text{初等行变换}}(\boldsymbol{E}\mid\boldsymbol{A}^{-1}).$$

初等列变换同理，这里不再赘述.

具体例子可参考课本 P61 例 16.

4. 矩阵乘法与矩阵的秩

这部分在课本 P63-P65 有详细的讲述，主要就是几个证明过程，本资料从略. 这里只是整理一下这些结论并给出直观理解，帮助记忆.

（1）若 $\boldsymbol{P}$ 为可逆矩阵，则 $r(\boldsymbol{PA})=r(\boldsymbol{A})$，$r(\boldsymbol{AP})=r(\boldsymbol{A})$.

可逆意味着 $\boldsymbol{P}$ 是满秩方阵，而满秩方阵是可以拆成一系列初等矩阵的，所以 $\boldsymbol{PA}$ 就是对 $\boldsymbol{A}$ 进行了初等行变换，而初等行变换不改变矩阵的秩，因此 $r(\boldsymbol{PA})=r(\boldsymbol{A})$. 后式同理.

（2）$r(\boldsymbol{AB})\le\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}$，或者说 $r(\boldsymbol{AB})\le r(\boldsymbol{A})$ 且 $r(\boldsymbol{AB})\le r(\boldsymbol{B})$.

仅仅为了记忆的话，或许可以把两个矩阵的乘法 $\boldsymbol{AB}$ 与两个集合的交集类比?

（3）$r(\boldsymbol{A}+\boldsymbol{B})\le r(\boldsymbol{A})+r(\boldsymbol{B})$.

仅仅为了记忆的话，或许可以把两个矩阵的加法 $\boldsymbol{A}+\boldsymbol{B}$ 与两个集合的并集类比?

（4）$r\begin{vmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{vmatrix}=r(\boldsymbol{A})+r(\boldsymbol{B})$.

为了记忆，就设想 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是阶梯型矩阵好了，这样就可以直接数阶梯头的个数了，显然 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的阶梯头不会互相干扰吧？那就只要把两个阶梯头的数量加起来就行了.

（5）Sylvester 不等式：若 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶方阵，则

$$r(\boldsymbol{AB})\ge r(\boldsymbol{A})+r(\boldsymbol{B})-n.$$

说不出来，就一直觉得和集合里的某些结论有点像.

（6）Frobenius 不等式：若 A,B,C 都是 n 阶方阵，则

$$r(ABC)\geq r(AB)+r(BC)-r(B).$$

特例：当 $B=E$ 时，式（6）就退化成式（5）.

（7）零积秩不等式：若 A,B 都是 n 阶方阵，且 $AB=O$，则 $r(A)+r(B)\leq n$.

这个式子虽然出现在课本 P74 第 59 题，但这个式子用到的场合还是蛮多的. 建议有能力的话可以记一下.

例 4　证明零积秩不等式.

证明　由 Sylvester 不等式，$r(AB)\geq r(A)+r(B)-n$.

又 $AB=O$，所以 $0\geq r(A)+r(B)-n$，整理得 $r(A)+r(B)\leq n$.

---

5. 伴随矩阵

还记得下面这个重要定理吗？

对 n 阶方阵 A 而言：A 可逆 $\Leftrightarrow |A|\neq 0 \Leftrightarrow r(A)=n$.

课本在证明第一个充要条件时用到了这个矩阵：

$$A^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

其中 A_{ij} 是代数余子式.

这个矩阵叫伴随矩阵. 为什么会出现这个矩阵？主要是因为它的性质：

$$AA^*=A^*A=|A|E.$$

这样，当 $|A|\neq 0$ 时，就有

$$A\left(\frac{A^*}{|A|}\right)=\left(\frac{A^*}{|A|}\right)A=E.$$

也就是说，当 $|A|\neq 0$ 时，我们就能找到 A 的逆矩阵 $A^{-1}=\dfrac{A^*}{|A|}$. 这为我们求逆矩阵提供了一种思路. 不过，这种求逆矩阵的方法和行列式的定义式一样，不适用于计算. 想想也知道，如果要求一个 n 阶方阵的伴随矩阵，用这种方式就得求 n^2 个代数余子式，谁愿意呢？所以，它只适合用在证明题里. （个人教训：本人在秋冬学期的某次小测中用这种方法求了一个逆矩阵，但是**行列式忘记除了**，痛失一题的分数）

将上面这个式子稍作变形，得 $A^*=|A|A^{-1}$. 这个式子用处可大了！关于伴随矩阵有很多结论的证明都会用到它. 不过，要是在这里把一堆结论列出来，恐怕只会徒增负担. 所以，下面只举一个例子来展现一下它的作用.

例 5 证明：$\left|A^*\right| = \left|A\right|^{n-1}$.

证明 若 A 可逆，则 $A^* = |A|A^{-1}$，于是 $\left|A^*\right| = \left||A|A^{-1}\right| = |A|^n\left|A^{-1}\right| = |A|^n|A|^{-1} = |A|^{n-1}$.

若 A 不可逆，则 $|A| = 0$，于是 $|A|^{n-1}$. 接下来只要证 $\left|A^*\right| = 0$.

我们用反证法，假设 $\left|A^*\right| \neq 0$，则 A^* 可逆.

于是，由 $AA^* = |A|E = O$ 可得 $AA^*\left(A^*\right)^{-1} = O$，于是 $AE = O$，即 $A = O$.

（其实这里已经得出荒唐的结论了，因为由 $|A| = 0$ 是不能得出 $A = O$ 的，你可以很容易地找到一个 A 作为反例来说明 $A = O$ 的错误，因此假设不成立. 当然，你也可以选择继续写下去，得出与条件矛盾的结论）

于是 $A^* = O$. （A 的所有元素都为 0 了，所有代数余子式当然都是 0）

所以 $\left|A^*\right| = 0$，与假设 $\left|A^*\right| \neq 0$ 矛盾. 因此假设不成立，$\left|A^*\right| = 0$ 得证.

看到 $A^* = |A|A^{-1}$ 的重要性了吗？很多关于伴随矩阵的证明题，一开始看的时候可能会一头雾水，但是如果能用好这个式子的话，应该不成难题.

从上面的例子中，我们还能看到伴随矩阵和逆矩阵的关系. **任何方阵都有伴随矩阵，但不一定有可逆矩阵**. 所以，只有当方阵可逆时，我们才能用 $A^* = |A|A^{-1}$ 得到它的伴随矩阵. 如果不可逆，就得另寻出路了.

6. 逆矩阵与逆映射

*破局四*花了大量的篇幅讲述了线性映射，现在我们来回顾一下.

将一个矩阵 M 作用在列向量 a 上，得到一个新的列向量 Ma，这就是一种线性映射. 现在将 M 的逆矩阵 M^{-1} 作用在 Ma 上，得到 $M^{-1}Ma = a$. 发现了什么？变回 a 了！可见，逆矩阵其实是一种逆映射，它的作用是将线性映射后的向量还原.

例 6 已知 $a = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$，$M = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$，$b = \begin{bmatrix} x \\ y \end{bmatrix} = Ma$.

（1）若 $x_0^2 + y_0^2 = 1$，求 x 和 y 应满足的关系式.

（2）若 $x^2 + \dfrac{y^2}{4} = 1$，求 x_0 和 y_0 应满足的关系式.

解 （1）$b = Ma \Rightarrow \begin{bmatrix} x_0 \\ 2y_0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} x \\ y/2 \end{bmatrix}$，

$x_0^2 + y_0^2 = 1 \Rightarrow x^2 + \dfrac{y^2}{4} = 1$.

（2）$\boldsymbol{b}=\boldsymbol{Ma}\Rightarrow\boldsymbol{a}=\boldsymbol{M}^{-1}\boldsymbol{b}\Rightarrow\begin{bmatrix}x\\y/2\end{bmatrix}=\begin{bmatrix}x_0\\y_0\end{bmatrix}\Rightarrow\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x_0\\2y_0\end{bmatrix}$,

$x^2+\dfrac{y^2}{4}=1\Rightarrow x_0^2+y_0^2=1$.

通过矩阵 $\boldsymbol{M}$，我们将一个圆映射成了椭圆；通过它的逆矩阵 $\boldsymbol{M}^{-1}$，我们将椭圆还原成了圆.

那如果 $\boldsymbol{M}$ 不可逆呢？不妨参考一下破局四例 8. 我们发现，$\boldsymbol{M}=\begin{bmatrix}1&1\\2&2\end{bmatrix}$将所有的列向量 $\boldsymbol{a}$ 都映射到了直线 $y=2x$ 上，进行了“降维打击”. 这样你还能根据映射后的向量求出原来的向量吗？

举个很简单的例子，如果 $\boldsymbol{a}=\begin{bmatrix}u\\v\end{bmatrix}$，那么 $\boldsymbol{Ma}=\begin{bmatrix}u+v\\2(u+v)\end{bmatrix}$；如果 $\boldsymbol{b}=\begin{bmatrix}u+v\\0\end{bmatrix}$，那么 $\boldsymbol{Mb}=\begin{bmatrix}u+v\\2(u+v)\end{bmatrix}$. 发现问题了吗？不同的向量可以映射成同样的向量！在这种情况下，你当然是没法将映射后的向量还原的.

关于分块矩阵，这里不做讲解. 分块矩阵的运算法则与一般矩阵的运算法则有许多是可以类比的.