

考研线代

枫聆

2021 年 8 月 5 日

目录

1 高等代数的研究对象	3
2 行列式	7
2.1 行列式的基本定义	7
2.2 行列式的基本性质	9
2.3 按行 (列) 展开	13
2.4 克拉默法则	16
2.5 Laplace 展开	18
2.6 行列式的计算	22
3 线性空间	25
3.1 线性空间的基本定义	25
3.2 线性空间的基本性质	25
3.3 子空间	25
3.4 线性组合和线性表出	25
3.5 n 元线性方程的线性表示	26
3.6 线性相关和线性无关	26
3.7 线性相关和线性无关的性质	27
3.8 极大线性无关组	27
3.9 向量组等价	27
3.10 向量组的秩	28
3.11 有限维线性空间的基和维数	29
3.12 矩阵的秩	29

3.13	方程组有解的充要条件	30
3.14	齐次方程组解集结构	30
3.15	非齐次线性方程组解集结构	31
4	矩阵运算	33
4.1	矩阵运算的定义	33
4.2	矩阵运算的性质	33
4.3	矩阵乘法中的线性组合	34
4.4	特殊矩阵	35
4.5	可逆矩阵	39
4.6	分块矩阵	42
4.7	分块矩阵的初等变换	44
4.8	矩阵乘积的行列式	47
4.9	矩阵运算中秩的性质	49
5	矩阵相似及二次型	52
5.1	内积和正交定义	52
5.2	正交矩阵	54
5.3	正交变换	55
5.4	矩阵的特征值和特征向量	55
5.5	等价矩阵	60
5.6	相似矩阵	61
5.7	矩阵对角化的条件	62
5.8	实对称矩阵的对角化	64
6	二次型	67
6.1	二次型及其标准型	67
6.2	规范形	72
6.3	正定二次型与正定矩阵	73

高等代数的研究对象

linear algebra 在研究多元一次方程组的解的过程中逐渐形成的一门学科. 如有这样一个方程组

$$\begin{aligned}x_1 + 3x_2 + x_3 &= 2, \\3x_1 + 4x_2 + 2x_3 &= 9, \\-x_1 - 5x_2 + 4x_3 &= 10, \\2x_2 + 7x_3 &= 1\end{aligned}$$

我们求解方程组的解过程通常是把用某一行的 n 倍加上或者减去另一行, 这种情况下我们其实是在对每一行方程的变元前面的系数做运算, 只要我们排列好对应系数的位置就可以避免每一次都写变元, 自然而然地, 多元一次方程组对应的系数矩阵和增广矩阵就诞生了, 例如上面方程组的增广矩阵如下

$$\begin{bmatrix} 1 & 3 & 1 & 2 \\ 3 & 4 & 2 & 9 \\ -1 & -5 & 4 & 10 \\ 2 & 7 & 1 & 1 \end{bmatrix}$$

这个矩阵也可以称为 4 阶矩阵 (4×4). 去掉最后常数项列就是对应的系数矩阵. 我们通过研究增广矩阵来研究原方程的解, 那么研究方程组的解可以从两个方向出发, 把解直接解出来或者判别方程组解的情况.



上图是整个 linear algebra 研究过程的发展, 值得关注是线性空间是向量空间的一般推广; 矩阵不光可以表示 n 维线性方程, 同样也是可以用来表示线性映射; 线性空间到自身的线性映射叫做线性变换; 从线性空间到线性度量空间需要用到一个双线性函数来描述两个其空间两个元素的度量;

Definition 1.1. **阶梯型矩阵** 若给定矩阵满足下面条件

1. 0 行全在下方;
2. **主元**(非零行第一个非零元素) 的列指标随着行指标的增加而严格增大.

则称其为行阶梯形矩阵. i.e.

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

同理还有列阶梯形矩阵, 把上面这个矩阵顺时针旋转 90° .

$$\begin{pmatrix} 2 & -3 & 6 & 0 \\ 1 & -1 & 3 & 0 \\ 3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Definition 1.2. **(简化阶梯型矩阵)** 若给定行阶梯形满足下面条件

1. 主元都是 1.
2. 主元所在列的其余元素都是 0.

则称其为简化行阶梯形矩阵. i.e.

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

Definition 1.3. 下面几种矩阵变换操作称之为矩阵的**初等行变换**

1. 把一行的倍数加到另一行上;
2. 两行互换;
3. 一行乘以一个非零数.

Theorem 1.4. **(初等行变换的性质)** 矩阵的初等行变换得到的方程组与原来的方程组同解.

Proposition 1.5. **(线性方程组的判别方法)** n 元线性方程组解的情况只有三种可能, 其对应的行阶梯形矩阵的特征如下

1. 若非零行个数等于未知数个数，则有唯一解；
2. 若非零行个数小于未知数个数，则有无穷多个解；
3. 若存在非零行对应等式 $0 = d$ ，其中 d 不等于零，则无解。

有一个很形象的理解方式类比到平面上两条直线，它们可能存在三种关系：平行，相交（相交于一点），重合。

证明. 其中无解对应的情况是显然的，那么我们在确保不会出现 $0 = d$ 的情况下有解，且为什么只会出现唯一解或者无穷多个解，而不会出现有两个解的这种情况呢？若给定 n 元线性方程组的增广矩阵，经过初等行变换化成行阶梯形矩阵 J . 设 J 有 r 个非零行， J 有 $n+1$ 列，其中 $n+1$ 列表示等式右边的常数项。

$$J = \begin{pmatrix} \cdots & \cdots & c_1 \\ \vdots & \vdots & \vdots \\ \cdots & b_{rt} & \cdots \\ 0 & & \end{pmatrix}$$

设 r 行的主元为 b_{rt} ，即它在第 t 列，那么必有 $t \leq n$. 根据行阶梯形的定义，列指标是随着行指标的增加而严格增大的，因此也有 $t \geq r$ ，同时 $t \leq n$ ，因此可以得到 $r \leq n$ ，即非零行的个数是不会超过 n 的. 此刻再把 J 通过初等行变换变成简化行阶梯形矩阵 J_1 ，那么 J_1 也有 r 非零行，即有 r 个主元，下面分两种情况分别讨论：当 $r = n$ 时， J_1 就有 n 个主元，那么按照行阶梯形主元排列顺序，

$$J_1 = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & c_1 \\ 0 & 1 & \cdots & \cdots & c_2 \\ 0 & 0 & 1 & \cdots & c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & 1 & c_n \end{pmatrix}$$

即 n 个主元，要分别放到 n 列，那么就是每列都有一个主元，即原方程组有唯一解. 当 $r < n$ 时，此时简化行阶梯形矩阵为

$$J_1 = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & c_1 \\ \cdots & 1 & \cdots & \cdots & \cdots & c_2 \\ \cdots & \cdots & 1 & \cdots & \cdots & c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & 1 & \cdots & c_r \end{pmatrix}$$

那么现在我们可以把上面这 r 个主元对应的未知数表示出来

$$\begin{aligned}x_1 &= b_{1,1}x_{i_1} + b_{1,2}x_{i_2} + \cdots + b_{1,n-r}x_{n-r} + d_1 \\x_{j_2} &= b_{j_2,1}x_{i_1} + b_{j_2,2}x_{i_2} + \cdots + b_{j_2,n-r}x_{n-r} + d_{j_2} \\&\vdots \\x_{j_r} &= b_{j_r,1}x_{i_1} + b_{j_r,2}x_{i_2} + \cdots + b_{j_r,n-r}x_{n-r} + d_{j_r}\end{aligned}$$

其中 $x_1, x_{j_2}, \cdots, x_{j_r}$ 表示对应列的 r 个主元, 其余的自由变量有 $n-r$ 个用 $x_{i_1}, x_{i_2}, \cdots, x_{n-r}$ 表示, 常数 b_{ij} . 那么只要任意确定一组自由变量的取值, 也就确定了一组解, 即这种情况下只有无穷多个解, 上面这个形式也称为一般解. \square

Definition 1.6. n 元齐次线性方程组如下

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\&\vdots \\a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0\end{aligned}$$

显然 $(0, 0, \cdots, 0)$ 是它的一个解, 称为零解; 其余的解 (如果有) 称为非零解.

Lemma 1.7. (齐次线性方程组有非零解的充要条件) n 元齐次线性方程组有非零解当且仅当系数矩阵经过初等行变换得到的阶梯形矩阵的非零行个数 $r < n$.

Lemma 1.8. (齐次线性方程族有非零解的充分条件) n 元齐次线性方程组的方程个数 $m < n$, 则它有非零解.

行列式

行列式的基本定义

Annotation 2.1. 行列式的意义 在用矩阵解 n 元线性方程组的过程中, 我们需要将对应的增广矩阵转换为阶梯型矩阵, 然后向上回溯求解. 如果对于一个 n 元线性方程组, 我们只想关注它的解结构或者更明确一点, 是否有解就够了, 那我们是否可以在不把具体解解出来的情况下, 回答这个问题呢? 也就是说我们是否可以在不进行初等行变换化阶梯形矩阵的情况下, 来回答这个问题?

从二元线性方程组作为例子出发来进行讨论, 给定一个二元线性方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= c_1 \\ a_{21}x_1 + a_{22}x_2 &= c_2 \end{aligned}$$

其中 a_{11}, a_{21} 不全为 0, 不妨设 a_{11} 不等于 0. 其对应的增广矩阵通过初等行变换化作阶梯形矩阵如下

$$\begin{pmatrix} a_{11} & a_{12} & c_1 \\ 0 & a_{22} - a_{12}\frac{a_{21}}{a_{11}} & c_2 - c_1\frac{a_{21}}{a_{11}} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & c_1 \\ 0 & \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}} & \frac{c_2a_{11} - c_1a_{21}}{a_{11}} \end{pmatrix}$$

若 $a_{11}a_{22} - a_{12}a_{21} \neq 0$, 则原方程组有唯一解; 反之, 若 $a_{11}a_{22} - a_{12}a_{21} = 0$, 则原方程组有无穷多个解或者无解. 可以很快发现 $a_{11}a_{22} - a_{12}a_{21}$ 就是 $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ 的行列式, 称其为**二阶行列式**. 基于这个直觉就将其推广到 n 元方程组对应的系数矩阵行列式来判别解的情况. 观察 $x_2 = \frac{c_2a_{11} - c_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$, 实际上它就是两个行列式做除法

$$x_2 = \frac{\begin{vmatrix} a_{11} & c_1 \\ a_{21} & c_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{|B_2|}{|A|}.$$

其中 $|B_1|$ 就是将其系数矩阵的第 2 列换成常数项, 这也是后面会提到**克拉默法则**的一个特殊情况.

Example 2.2. 3 阶行列式的记忆法

Definition 2.3. (逆序数的概率) 若给定一个排列 $j_1j_2 \cdots j_n$, 取其中任意两个不同的数 j_p 和 j_q , 使得 $1 \leq p < q \leq n$. 若 $j_p > j_q$, 就称 j_p 和 j_q 构成一个**逆序**. 一个排列中所有逆序的总和称为这个排列的**逆序数**. 特别地, 如果一个排列的逆序数是偶数, 则称这个排列为**偶排列**, 否则称为**奇排列**.

Definition 2.4. (n 阶行列式的概念) n 阶行列式的完全展开式如下

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}.$$

其中 j_1, j_2, \dots, j_n 表示一组列指标的排列, 即 $a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ 表示取自不同行不同列的 n 个元素的乘积. 用 $\tau(j_1 j_2 \cdots j_n)$ 表示 $j_1 j_2 \cdots j_n$ 的逆序数.

Definition 2.5. 记 $|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$, $|\mathbf{A}^T| = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$ 称为 $|\mathbf{A}|$ 的转置行列式.

行列式的基本性质

Lemma 2.6. 一个排列中任意两个元素对换，其逆序数的奇偶性改变.

证明. 设给定一个排列为 $j_1 j_2 \cdots j_n$ ，先考虑两个相邻的数交换，即

$$j_1 j_2 \cdots j_i j_{i+1} \cdots j_n \longrightarrow j_1 j_2 \cdots j_{i+1} j_i \cdots j_n$$

其中 $1 \leq i < n$. 很显然相邻的两个数交换，并不会影响它们两边的数的逆序个数，再分两种情况讨论，若 $j_i > j_{i+1}$ ，那么变换之后的排列的逆序数减 1；若 $j_i < j_{i+1}$ ，那么变换之后的排列数的逆序数加 1 即，因此做一次相邻交换其奇偶性改变. 再来考虑交换任意两个数 j_i 和 j_k ，我们可以把这个交换拆解为多个相邻数的交换，即

$$j_1 j_2 \cdots j_i \cdots j_k \cdots j_n \longrightarrow j_1 j_2 \cdots j_{i-1} \cdots j_i j_k \cdots j_n \longrightarrow j_1 j_2 \cdots j_{i-1} j_k \cdots j_i \cdots j_n$$

那么第一个箭头做了 $k - i - 1$ 次两两交换，第二个箭头做了 $k - i$ 次两两交换，那么一共做了 $2(k - i) - 1$ 次，因此这个排列的逆序数的奇偶性一定会改变. \square

Lemma 2.7. 任意一个排列 $j_1 j_2 \cdots j_n$ 的逆序数与它经过一系列对换变到自然序 $123 \cdots n$ 的次数的奇偶性相同.

证明. 首先自然序的逆序数等于 0 是一个偶数，那么分两种情况来考虑. 若 $j_1 j_2 \cdots j_n$ 逆序数是一个奇数，那么它必须经过奇数次相邻变换才能得到一个逆序数是偶数的排列，同理若 $j_1 j_2 \cdots j_n$ 逆序数是一个偶数，那么它必须经过偶数次相邻变换才能得到一个逆序数是偶数的排列. \square

Lemma 2.8. 任取 n 阶行列式中一项 $(-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ 经过 s 次两个元素位置互换，使其变成了 $a_{i_1 k_1} a_{i_2 k_2} \cdots a_{i_n k_n}$ ，则

$$(-1)^{\tau(j_1 j_2 \cdots j_n)} = (-1)^{\tau(i_1 i_2 \cdots i_n) + \tau(k_1 k_2 \cdots k_n)}.$$

元素位置交换得到新排列的行指标逆序数与列指标逆序数的和与原来排列逆序数的奇偶性相同. 这个 lemma 还是比较重要的，因为它可以改写行列式通项公式的，就是说不一定行指标要按自然序排列，即选定行指标 i_1, i_2, \cdots, i_n 之后，有

$$\sum_{k_1, k_2, \cdots, k_n} (-1)^{\tau(i_1 i_2 \cdots i_n) + \tau(k_1 k_2 \cdots k_n)} a_{i_1 k_1} a_{i_2 k_2} \cdots a_{i_n k_n}.$$

证明. 分别考虑行指标和列指标的变换，则有

$$\begin{aligned} (-1)^{\tau(i_1 i_2 \cdots i_n)} &= (-1)^s (-1)^{\tau(123 \cdots n)} \\ (-1)^{\tau(k_1 k_2 \cdots k_n)} &= (-1)^s (-1)^{\tau(j_1 j_2 \cdots j_n)} \end{aligned}$$

两式左右两边相乘就可以得到命题结果. \square

Proposition 2.9. 行列式的基本性质如下

1. 行列式和它的转置行列式相等, 即 $|\mathbf{A}^T| = |\mathbf{A}|$;
2. 两行互换位置, 行列式的值变号;
3. 某行若有公因子 k , 则可以把 k 提出行列式记号之外;
4. 如果行列式某行 (或列) 是两个元素之和, 则可将行列式拆成两个行列式之和;

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \cdots & a_{in} + b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

5. 把某行 (或者列) 的 k 加到另一行 (或者列), 行列式的值不变 (契合初等变换中最重要的一个性质).

证明. 给定 n 阶行列式为

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

(1) 设 $|\mathbf{A}^T| = \sum_{p_1, p_2, \dots, p_n} (-1)^{\tau(p_1 p_2 \dots p_n)} b_{1p_1} b_{2p_2} \cdots b_{np_n} = \sum_{p_1, p_2, \dots, p_n} (-1)^{\tau(p_1 p_2 \dots p_n)} a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}$, 其中 $b_{ij} = a_{ji}$. 要证明 $|\mathbf{A}| = |\mathbf{A}^T|$, 我们需要证明对于 $|\mathbf{A}^T|$ 代数和中的任意一项都有 $|\mathbf{A}|$ 代数和的一项与其对应, 反之亦然. 那么任取 $|\mathbf{A}^T|$ 中一项 $\sum (-1)^{\tau(p_1 p_2 \dots p_n)} a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}$, 我们可以让其通过有限次相邻元素交换变成 $a_{1q_1} a_{2q_2} \cdots a_{nq_n}$ 之后, 现在这个排列就是按行取元素的一个形式, 但是还不确定是其逆序数的情况, 记其相邻交换的次数为 s . 这里我们同时考察行指标和列指标变化, 即有

$$\begin{aligned} (-1)^{\tau(q_1 q_2 \dots q_n)} &= (-1)^s (-1)^{\tau(123 \dots n)} \\ (-1)^{\tau(123 \dots n)} &= (-1)^s (-1)^{\tau(p_1 p_2 \dots p_n)}. \end{aligned}$$

等式两边相乘我们就可以得到 $(-1)^{\tau(q_1 q_2 \dots q_n)} = (-1)^{\tau(p_1 p_2 \dots p_n)}$, 所以经过变换之后确实找到了 $|\mathbf{A}|$ 代数和中的一项, 实际上这是 lemma 2.8 中的一个特殊情况. 反之任取 $|\mathbf{A}|$ 代数和中的一项 $\sum_{p_1, p_2, \dots, p_n} (-1)^{\tau(p_1 p_2 \dots p_n)} a_{1p_1} a_{2p_2} \cdots a_{np_n}$, 也可以在 $|\mathbf{A}^T|$ 代数和中找到一项 $\sum_{q_1, q_2, \dots, q_n} (-1)^{\tau(q_1 q_2 \dots q_n)} a_{q_1 1} a_{q_2 2} \cdots a_{q_n n}$ 与之对应. 这个性质更本质的意义是行列式中的行和列具有同等的地位.

(2) 如果我们交换第 i 和 k 行得到的矩阵我们记为 A_2 , 如下

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \longrightarrow \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

那么

$$\begin{aligned} \det(\mathbf{A}_2) &= \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} b_{1j_1} b_{2j_2} \cdots b_{nj_n} \\ &= \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \cdots \mathbf{j}_i \cdots \mathbf{j}_k \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{kj_k} \cdots a_{ij_i} \cdots a_{nj_n} \\ &= \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \cdots \mathbf{j}_k \cdots \mathbf{j}_i \cdots j_n) + 1} a_{1j_1} a_{2j_2} \cdots a_{ij_i} \cdots a_{kj_k} \cdots a_{nj_n} \quad (\text{性质 1}) \\ &= (-1) \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \cdots \mathbf{j}_k \cdots \mathbf{j}_i \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{ij_i} \cdots a_{kj_k} \cdots a_{nj_n} \end{aligned}$$

这里有一个推论是若行列式有两行 (或者列) 完全相同, 那么此行列式等于零. 这是因为 $|\mathbf{A}| = -|\mathbf{A}|$, 所以 $|\mathbf{A}| = 0$.

(3) 考虑下面行列式

$$|\mathbf{B}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

按照完全展开式

$$\begin{aligned} |\mathbf{B}| &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots ka_{ij_i} \cdots a_{nj_n} \\ &= k \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots a_{ij_i} \cdots a_{nj_n} \end{aligned}$$

(4) 按照完全展开式

$$\begin{aligned} |\mathbf{A}| &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots (a_{ij_i} + b_{ij_i}) \cdots a_{nj_n} \\ &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots a_{ij_i} \cdots a_{nj_n} + \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots b_{ij_i} \cdots a_{nj_n} \end{aligned}$$

(5) 若考虑把第 i 行的 k 倍加第到 q 行, 那么当且的完全展开式如下

$$\begin{aligned}
 |\mathbf{A}_1| &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots (a_{qj_q} + k a_{ij_i}) \cdots a_{nj_n} \\
 &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots (a_{qj_q}) \cdots a_{nj_n} + k \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots a_{ij_i} \cdots (a_{ij_i}) \cdots a_{nj_n} \\
 &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} \cdots (a_{qj_q}) \cdots a_{nj_n}
 \end{aligned}$$

□

Annotation 2.10. 行列式的几何意义 给定两个向量 $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ 及 $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, 且这两个向量不共线, 则行列式

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

表示这两个向量张成的平行四边形的定向面积.

按行 (列) 展开

Definition 2.11. 在 n 阶行列式中划去 a_{ij} 所在的第 i 行和第 j 列, 留下来的 $n-1$ 阶行列式叫做 a_{ij} 的余子式, 记为 M_{ij} ; 称 $(-1)^{i+j}M_{ij}$ 为代数余子式, 记为 A_{ij} .

Lemma 2.12. 给定如下行列式

$$\det(D) = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

则 $\det(D) = a_{11}A_{11}$.

证明. 这其实是很 trivial 的, 其完全展开式如下

$$\begin{aligned} \det(D) &= \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1 j_2 \dots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \\ &= \sum_{j_1=1, \dots, j_n} (-1)^{\tau(j_1 j_2 \dots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \\ &= a_{11} \sum_{k_2, \dots, k_n} (-1)^{\tau(k_2 \dots k_n)} a_{2k_2} \cdots a_{nk_n} \\ &= a_{11}(-1)^{1+1} M_{11} \\ &= a_{11}A_{11}, \end{aligned}$$

其中 $\tau(1j_2j_3 \cdots j_n) = \tau(j_2j_3 \cdots j_n)$, k_2, k_3, \dots, k_n 均不为 1. □

Lemma 2.13. 若给定 n 行列式 D 存在元素 a_{ij} 所在行元素除了它自己以外全为零, 如下

$$\det(D) = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

则 $\det(D) = a_{ij}A_{ij}$

证明. 考虑将第 i 行先向上移到第一行, 再将第 j 列向左移动到第一列, 还原到 Lemma 2.8 的情况下我们设为 D_1 . 注意这里是相邻移动, 而不是直接将第一行和第 i 行的对换, 第一列和第 j 列的替换, 这样做的目的就是在将 a_{ij} 移动到 $(1,1)$ 了之后, 对应的 M_{11} 是和原来的 M_{ij} 相同的. 所以现在我们只需要考虑在移动过程中对 D 的影响, 向上移动了 $i-1$, 向左移动了 $j-1$ 次, 那么一共移动了 $i+j-2$ 次, 即

$$D = (-1)^{i+j-2} D_1 = (-1)^{i+j-2} a_{ij} M_{ij} = a_{ij} A_{ij}.$$

□

Theorem 2.14. (按行 (列) 展开定理) n 阶行列式等于它的任一行 (列) 的各元素与其对应的代数余子式乘积的代数和, 即

$$D = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}, \quad i = 1, 2, \cdots, n$$

$$D = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}, \quad j = 1, 2, \cdots, n.$$

证明. 将 D 按第 i 行拆成 n 个行列式.

$$\begin{aligned} D &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} + 0 + \cdots + 0 & 0 + a_{i2} + \cdots + 0 & \cdots & 0 + 0 + \cdots + a_{in} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{in} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}. \end{aligned}$$

□

Theorem 2.15. n 阶行列式的任一行 (列) 元素与另一行 (列) 元素的代数余子式乘积的代数和等于零, 即

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0, \quad (i \neq j).$$

证明. 我们把给定行列式 D 的第 j 行特别地标注出来

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_1 & \cdots & b_n \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

其中 b_1, \cdots, b_n 表示第 j 行, 那么此时

$$D = b_1A_{j1} + \cdots + b_nA_{jn}.$$

我们再第 j 行用第 i 行换掉得到 D_1 , 那么此时第 j 行各元素的代数余子式是没有发生变化的, 且现在 D_1 是有两行相同的元素, 所以

$$D_1 = a_{i1}A_{j1} + \cdots + a_{in}A_{jn} = 0.$$

同理也可以对第 j 列做上述操作.

□

克拉默法则

Proposition 2.16. **行列式作为线性方程组解的判定条件** 给定 n 个 n 元线性方程的方程组，它有唯一解的充要条件其对应的系数矩阵的行列式不为零. 反之若其系数矩阵的行列式为零，则有无穷多个解或者有非零解.

Corollary 2.17. 给定 n 个 n 元线性方程的齐次线性方程组，有零解当且仅当系数矩阵的行列式等于零. 反之若其系数矩阵行列式等于零则有非零解.

Theorem 2.18. 给定 n 个 n 元线性方程的方程组的系数矩阵如下

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

及其常数项为 $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. 若 $|\mathbf{A}| \neq 0$ ，那么此线性方程组有唯一解，解可以表示为

$$x_1 = \frac{|\mathbf{B}_1|}{|\mathbf{A}|}, x_2 = \frac{|\mathbf{B}_2|}{|\mathbf{A}|}, \cdots, x_n = \frac{|\mathbf{B}_n|}{|\mathbf{A}|},$$

其中 $|\mathbf{B}_j|$ 是把行列式 $|\mathbf{A}|$ 第 j 列换成方程常数项之后得到的行列式，即

$$|\mathbf{B}_j| = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}, j = 1, 2, \cdots, n.$$

证明. 下面将 $(\frac{|\mathbf{B}_1|}{|\mathbf{A}|}, \frac{|\mathbf{B}_2|}{|\mathbf{A}|}, \cdots, \frac{|\mathbf{B}_n|}{|\mathbf{A}|})$ 带入原方程组，验证它的确是一个解. 首先将原方程组简写为

$$\sum_{j=1}^n a_{ij} x_j = b_i, i = 1, 2, \cdots, n.$$

那么

$$\begin{aligned} \sum_{j=1}^n a_{ij} \frac{|\mathbf{B}_j|}{|\mathbf{A}|} &= \frac{1}{|\mathbf{A}|} \sum_{j=1}^n a_{ij} |\mathbf{B}_j| \\ &= \frac{1}{|\mathbf{A}|} \sum_{j=1}^n a_{ij} \sum_{s=1}^n b_s A_{sj} \\ &= \frac{1}{|\mathbf{A}|} \sum_{s=1}^n \sum_{j=1}^n b_s a_{ij} A_{sj} \quad \text{theorem 2.15} \\ &= \frac{1}{|\mathbf{A}|} b_i |\mathbf{A}| = b_i. \end{aligned}$$



Laplace 展开

Definition 2.19. 在 n 阶行列式 $|\mathbf{A}|$ 中任意选定 k 行 k 列 ($k \leq n$), 位于这些行和列的交点上的 k^2 个元素按原来的次序组成的 k 阶行列式 M , 称为行列式 $|\mathbf{A}|$ 的 **k 阶子式**, 记做

$$\mathbf{A} \begin{pmatrix} i_1, i_2, \dots, i_k \\ j_1, j_2, \dots, j_k \end{pmatrix},$$

其中 $i_1 < i_2 < \dots < i_k$, $j_1 < j_2 < \dots < j_k$. 当 $k < n$ 时, 在 $|\mathbf{A}|$ 中划去这 k 行 k 列后余下的元素按照原来的次序组成的 $n - k$ 阶行列式 M' 称为 **k 阶子式 M 的余子式**, 令

$$\begin{aligned} \{i'_1, i'_2, \dots, i'_{n-k}\} &= \{1, 2, \dots, n\} \setminus \{i_1, i_2, \dots, i_k\}, \\ \{j'_1, j'_2, \dots, j'_{n-k}\} &= \{1, 2, \dots, n\} \setminus \{j_1, j_2, \dots, j_k\}, \end{aligned}$$

其中 $i'_1 < i'_2 < \dots < i'_{n-k}$, $j'_1 < j'_2 < \dots < j'_{n-k}$, 则 M' 记为

$$\mathbf{A} \begin{pmatrix} i'_1, i'_2, \dots, i'_{n-k} \\ j'_1, j'_2, \dots, j'_{n-k} \end{pmatrix}.$$

M 的代数余子式为

$$(-1)^{(i_1+i_2+\dots+i_k)+(j_1+j_2+\dots+j_k)} M'.$$

Lemma 2.20. n 阶行列式 $|\mathbf{A}|$ 的任一个子式 M 与它的代数余子式 A 的乘积中的每一项都是行列式 $|\mathbf{A}|$ 的展开式中的一项, 而且符号也一致.

证明. 首先考虑 M 位于 $|\mathbf{A}|$ 的左上角时, 即

$$\left| \begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ \hline a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots & \vdots & M' & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{n,n} \end{array} \right|$$

此时 M 的代数余子式为

$$A = (-1)^{(1+2+\dots+k)+(1+2+\dots+k)} M' = M'.$$

M 中的每一项记为

$$(-1)^{\tau(\alpha_1 \alpha_2 \dots \alpha_k)} a_{1\alpha_1} a_{1\alpha_2} \cdots a_{k\alpha_k},$$

其中 $\alpha_1\alpha_2\cdots\alpha_k$ 表示 $1, 2, \dots, k$ 的一个排列. M' 中的每一项记为

$$(-1)^{\tau(\beta_{k+1}\beta_{k+2}\cdots\beta_{n-k})}a_{k+1,\beta_{k+1}}a_{k+2,\beta_{k+2}}\cdots a_{n,\beta_{n-k}},$$

其中 $\beta_{k+1}\beta_{k+2}\cdots\beta_{n-k}$ 表示 $k+1, k+2, \dots, n$ 的一个排列. 那么它们的乘积的每一项可以表示为

$$\begin{aligned} &= (-1)^{\tau(\alpha_1\alpha_2\cdots\alpha_k)+\tau(\beta_{k+1}\beta_{k+2}\cdots\beta_{n-k})}a_{1\alpha_1}\cdots a_{k\alpha_k}a_{k+1,\beta_{k+1}}\cdots a_{n\beta_{n-k}} \\ &= (-1)^{\tau(\alpha_1\alpha_2\cdots\alpha_k\beta_{k+1}\beta_{k+2}\cdots\beta_{n-k})}a_{1\alpha_1}\cdots a_{k\alpha_k}a_{k+1,\beta_{k+1}}\cdots a_{n\beta_{n-k}}, \end{aligned}$$

其中 $\alpha_1, \alpha_2, \dots, \alpha_k$ 均小于 $\beta_{k+1}, \beta_{k+2}, \dots, \beta_{n-k}$, 因此乘积的每一项对应上了 $|\mathbf{A}|$ 中的一项.

下面来证明一般情况. 设 M 位于 $|\mathbf{A}|$ 的第 i_1, i_2, \dots, i_k 行, 第 j_1, j_2, \dots, j_k 列, 这里有

$$i_1 < i_2 < \cdots < i_k, j_1 < j_2 < \cdots < j_k.$$

我们现将 M 重新变换到 $|\mathbf{A}|$ 的左上角, 即将第 i_1 行向上移动至第 1 行, 依次再将剩下的第 i_2, \dots, i_k 也上移, 同理也将第 j_1 列向左移动至第一列, 依次再将剩下的第 j_2, \dots, j_k 也左移. 那么总共行和列移动的次数为

$$\begin{aligned} &= (i_1 - 1) + (i_2 - 2) + \cdots + (i_k - k) + (j_1 - 1) + (j_2 - 2) + \cdots + (j_k - k) \\ &= (i_1 + i_2 + \cdots + i_k) + (j_1 + j_2 + \cdots + j_k) + k(k+1), \end{aligned}$$

将移动之后的行列式记为 $|\mathbf{A}'|$, 那么

$$|\mathbf{A}'| = (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)}|\mathbf{A}|,$$

即 $|\mathbf{A}'|$ 和 $|\mathbf{A}|$ 中的对应项都只差一个符号. 同样现在 $M\mathbf{A}$ 的每一项和 $|\mathbf{A}'|$ 的一项 s 对应, 那么就是和 $|\mathbf{A}|$ 中的 $(-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)}s$ 一项对应. \square

Theorem 2.21. Laplace 定理 设在 n 阶行列式 $|\mathbf{A}|$ 中任意取定了 k ($1 \leq k \leq n-1$) 个行. 由这 k 行元素所构成一切 k 阶子式与它们的代数余子式的乘积的和等于行列式 $|\mathbf{A}|$.

证明. 设取定 k 行之后得到的所有可能的子式为 M_1, M_2, \dots, M_t , 它们的余子式分别为 A_1, A_2, \dots, A_t , 其中 $t = C_n^k$. 定理要求证明

$$|\mathbf{A}| = M_1A_1 + M_2A_2 + \cdots + M_tA_t.$$

根据 lemma 2.20, M_iA_i 中每一项都是 $|\mathbf{A}|$ 中的一项, 而且 M_iA_i 和 M_jA_j ($j \neq i$) 无公共项. 因此我们只需要证明上述等式两边项数相等即可. 左边已知有 $n!$ 项, 而右边每一项 M_iA_i 中有 $(k)!(n-k)!$, 那么一共有

$$C_n^k(k)!(n-k)! = \frac{n!}{k!(n-k)!}(k)!(n-k)! = n!,$$

即定理得证.

若我们舍弃上面构造性的证明手法，从行列式的完全展开式出发去理解 laplace 展开. 首先我们取定 $i_1 < i_2 < \cdots < i_k, i'_1 < i_2 < \cdots < i'_{n-k}$, 其中 $\{i'_1, i'_2, \cdots, i'_{n-k}\} = \{1, 2, \cdots, n\} \setminus \{i_1, i_2, \cdots, i_k\}$, 那么可以将完全展开式写成

$$\sum_{\alpha_1, \alpha_2, \cdots, \alpha_k, \beta_1, \cdots, \beta_{n-k}} (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_k-k)+\tau(\alpha_1\alpha_2\cdots\alpha_k\beta_1\beta_2\cdots\beta_{n-k})} a_{i_1\alpha_1} a_{i_2\alpha_2} \cdots a_{i_k\alpha_k} a_{i'_1\beta_1} a_{i'_2\beta_2} \cdots a_{i'_{n-k}\beta_{n-k}}.$$

接下来我们将一个 \sum 分解成三个. 我们考虑先从选出第 j_1, j_2, \cdots, j_k 行, 且 $j_1 < j_2 < \cdots < j_k$, 那么剩下第 $j'_1, j'_2, \cdots, j'_{n-k}$ 行, 且 $j'_1 < j'_2 < \cdots < j'_{n-k}$. 然后我们从 j_1, j_2, \cdots, j_k 里面确定一个排列 $\alpha_1\alpha_2\cdots\alpha_k$, 再从 $j'_1, j'_2, \cdots, j'_{n-k}$ 中确定一个排列 $\beta_1\beta_2\cdots\beta_{n-k}$. 这样一来依然确定了原式中的一个 $\alpha_1\alpha_2\cdots\alpha_k\beta_1\beta_2\cdots\beta_{n-k}$ 排列, 即此时有

$$\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq n} \sum_{\alpha_1, \alpha_2, \cdots, \alpha_k} \sum_{\beta_1, \beta_2, \cdots, \beta_{n-k}} S$$

其中 $S = (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_k-k)+\tau(\alpha_1\alpha_2\cdots\alpha_k\beta_1\beta_2\cdots\beta_{n-k})} a_{i_1\alpha_1} a_{i_2\alpha_2} \cdots a_{i_k\alpha_k} a_{i'_1\beta_1} a_{i'_2\beta_2} \cdots a_{i'_{n-k}\beta_{n-k}}$. 下面再将指示符号拆开

$$\begin{aligned} (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_k-k)+\tau(\alpha_1\alpha_2\cdots\alpha_k\beta_1\beta_2\cdots\beta_{n-k})} &= (-1)^{(i_1+i_2+\cdots+i_k)-\frac{k(k+1)}{2}+\tau(\alpha_1\alpha_2\cdots\alpha_k\beta_1\beta_2\cdots\beta_{n-k})} \\ &= (-1)^{(i_1+i_2+\cdots+i_k)-\frac{k(k+1)}{2}} \cdot \\ &\quad (-1)^{\tau(\alpha_1\alpha_2\cdots\alpha_k)} (-1)^{\tau(j_1j_2\cdots j_k\beta_1\beta_2\cdots\beta_{n-k})} \\ &= (-1)^{(i_1+i_2+\cdots+i_k)-\frac{k(k+1)}{2}} \cdot \\ &\quad (-1)^{(j_1+j_2+\cdots+j_k)-\frac{k(k+1)}{2}} \\ &\quad (-1)^{\tau(\alpha_1\alpha_2\cdots\alpha_k)} (-1)^{\tau(\beta_1\beta_2\cdots\beta_{n-k})}. \end{aligned}$$

最终展开式就可以写成

$$\begin{aligned} &\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq n} (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)} \\ &\quad \sum_{\alpha_1, \alpha_2, \cdots, \alpha_k} (-1)^{\tau(\alpha_1\alpha_2\cdots\alpha_k)} a_{i_1\alpha_1} a_{i_2\alpha_2} \cdots a_{i_k\alpha_k} \\ &\quad \sum_{\beta_1, \beta_2, \cdots, \beta_{n-k}} (-1)^{\tau(\beta_1\beta_2\cdots\beta_{n-k})} a_{i'_1\beta_1} a_{i'_2\beta_2} \cdots a_{i'_{n-k}\beta_{n-k}} \end{aligned}$$

非常美妙!

□

Corollary 2.22. 下式成立

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \vdots & \mathbf{A} & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1r} \\ \vdots & \mathbf{C} & \vdots & \vdots & \mathbf{B} & \vdots \\ c_{r1} & \cdots & c_{rk} & b_{r1} & \cdots & b_{rr} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{rr} \end{vmatrix} (|\mathbf{A}| \cdot |\mathbf{B}|).$$

证明. 利用 laplace 开展, 选择前 k 行, 那么只有在选择前 k 列的时候, 其 k 阶子式才没有零列, 此时余子式为右下方的 $|B|$. \square

Corollary 2.23. 下式成立

$$\begin{vmatrix} 0 & \cdots & 0 & a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots & \vdots & \mathbf{A} & \vdots \\ 0 & \cdots & 0 & a_{k1} & \cdots & a_{kk} \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1r} \\ \vdots & \mathbf{C} & \vdots & \vdots & \mathbf{B} & \vdots \\ c_{r1} & \cdots & c_{rk} & b_{r1} & \cdots & b_{rr} \end{vmatrix} = (-1)^{(n+1)k} \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{rr} \end{vmatrix} ((-1)^{(n+1)k} |\mathbf{A}| \cdot |\mathbf{B}|).$$

行列式的计算

Proposition 2.24. (几种特殊的行列式)

1. 上(下)三角形行列式

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

2. 对角行列式

$$\begin{vmatrix} a_{11} & & & \\ & a_{22} & & 0 \\ & & \ddots & \\ 0 & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

3. 副对角线

$$\begin{vmatrix} & 0 & & a_{1n} \\ & & a_{2,n-1} & a_{2n} \\ & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_{1n}a_{2,n-1}\cdots a_{n1}.$$

Example 2.25. 计算 n 阶行列式 ($n \geq 2$)

$$\begin{vmatrix} k & \lambda & \lambda & \cdots & \lambda \\ \lambda & k & \lambda & \cdots & \lambda \\ \lambda & \lambda & k & \cdots & \lambda \\ \vdots & \vdots & \vdots & \cdots & \lambda \\ \lambda & \lambda & \lambda & \cdots & k \end{vmatrix}$$

hints 化阶梯型.

$$\begin{vmatrix} k + (n-1)\lambda & \lambda & \lambda & \cdots & \lambda \\ k + (n-1)\lambda & k & \lambda & \cdots & \lambda \\ k + (n-1)\lambda & \lambda & k & \cdots & \lambda \\ \vdots & \vdots & \vdots & \cdots & \lambda \\ k + (n-1)\lambda & \lambda & \lambda & \cdots & k \end{vmatrix} = [k + (n-1)\lambda] \begin{vmatrix} 1 & \lambda & \lambda & \cdots & \lambda \\ 1 & k & \lambda & \cdots & \lambda \\ 1 & \lambda & k & \cdots & \lambda \\ \vdots & \vdots & \vdots & \cdots & \lambda \\ 1 & \lambda & \lambda & \cdots & k \end{vmatrix}$$

$$= [k + (n-1)\lambda] \begin{vmatrix} 1 & \lambda & \lambda & \cdots & 0 \\ 0 & k - \lambda & 0 & \cdots & 0 \\ 0 & 0 & k - \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & k - \lambda \end{vmatrix}$$

$$= [k + (n-1)\lambda] (k - \lambda)^{n-1}$$

Example 2.26. 计算 3 阶行列式

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix}$$

hints 按第一列展开.

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1) \begin{vmatrix} 1 & 1 \\ a_2 & a_3 \end{vmatrix} \\ = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2).$$

Example 2.27. 计算 n 阶范德蒙行列式 ($n \geq 2$)

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

hints 数学归纳法. 显然在 $n = 2$ 时成立, 假设 $n = k$ 时成立. 那么当 $n = k + 1$ 时, 我们把第一列变成除了 a_{k+1} 其余全为 0, 变的方法是从最后一行开始, 用第 i 行乘 a_1 消去第 $i + 1$ 的 a^i , 则有

$$\begin{aligned} &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & a_3 & \cdots & a_k & a_{k+1} \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_k^2 & a_{k+1}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^{k-1} & a_2^{k-1} & a_3^{k-1} & \cdots & a_k^{k-1} & a_{k+1}^{k-1} \\ 0 & a_2^k - a_1 a_2^{k-1} & a_3^k - a_1 a_3^{k-1} & \cdots & a_k^k - a_1 a_k^{k-1} & a_{k+1}^k - a_1 a_{k+1}^{k-1} \end{vmatrix} \\ &\vdots \\ &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_k - a_1 & a_{k+1} - a_1 \\ 0 & a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_k(a_k - a_1) & a_{k+1}(a_{k+1} - a_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_2^{k-2}(a_2 - a_1) & a_3^{k-2}(a_3 - a_1) & \cdots & a_k^{k-2}(a_k - a_1) & a_{k+1}^{k-2}(a_{k+1} - a_1) \\ 0 & a_2^{k-1}(a_2 - a_1) & a_3^{k-1}(a_3 - a_1) & \cdots & a_k^{k-1}(a_k - a_1) & a_{k+1}^{k-1}(a_{k+1} - a_1) \end{vmatrix} \\ &= (a_2 - a_1)(a_3 - a_1) \cdots (a_k - a_1)(a_{k+1} - a_1) \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_2 & a_3 & \cdots & a_k & a_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2^{k-2} & a_3^{k-2} & \cdots & a_k^{k-2} & a_{k+1}^{k-2} \\ a_2^{k-1} & a_3^{k-1} & \cdots & a_k^{k-1} & a_{k+1}^{k-1} \end{vmatrix} \\ &= (a_2 - a_1)(a_3 - a_1) \cdots (a_k - a_1)(a_{k+1} - a_1) \prod_{2 \leq i < j \leq k+1} (a_j - a_i) \\ &= \prod_{1 \leq i < j \leq k+1} (a_j - a_i) \end{aligned}$$

线性空间

线性空间的基本定义

Definition 3.1. 域 F 作用在一个阿贝尔群 G 上.

线性空间的基本性质

Proposition 3.2. 设 V 是数域 F 上的一个线性空间, 则 V 的零元唯一.

Proposition 3.3. 设 V 是数域 F 上的一个线性空间, 则 $v \in V$ 都有唯一的逆元.

Proposition 3.4. 设 V 是数域 F 上的一个线性空间, 则 $v \in V$ 都有 $0v = 0$.

Proposition 3.5. 设 V 是数域 F 上的一个线性空间, 给定 $a \in F$ 和 $v \in V$. 若 $av = 0$, 则 $a = 0$ 或者 $v = 0$.

Proposition 3.6. 设 V 是数域 F 上的一个线性空间, 则 $v \in V$ 都有 $(-1)v = -v$.

子空间

Definition 3.7. 设 V 是数域 F 上的一个线性空间, 给定 U 是 V 的一个子集. 若 U 中的元素在向量加法和数量下封闭, 则称 U 是 V 的一个子空间 (实际上这是一个用于判定子空间的方法, 让你不用考虑整个线性空间的定义).

线性组合和线性表出

Definition 3.8. 给定向量组 $v_1, v_2, \dots, v_n \in V$, 任取 F 上一组数 a_1, a_2, \dots, a_n , 称向量

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

为 v_1, v_2, \dots, v_n 的一个线性组合. 取定 $v \in V$, 若存在 c_1, c_2, \dots, c_n 使得

$$v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n,$$

则称 v 可以由 v_1, v_2, \dots, v_n 线性表出.

n 元线性方程的线性表示

Definition 3.9. 给定 F 上 n 元线性方程组

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

则可以写成

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

即 $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$, 其中 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 表示该线性方程组系数矩阵的列向量组, β 是由其常数项组成的列向量.

Annotation 3.10. 研究线性方程组解存在性的重要思路转换 n 元线性方程组有解等价于 β 可以被 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出. 更深刻一点, 我们需要探究下述等式成立的条件

$$\beta \stackrel{?}{\in} \text{span}(\alpha_1, \alpha_2, \cdots, \alpha_n)$$

其中 $\text{span}(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 是 $F^{m,1}$ 的一个子空间. 自然地, 我们需要子空间的一些结构和性质.

线性相关和线性无关

Definition 3.11. 给定向量空间 V 上一组向量 v_1, v_2, \cdots, v_n . 若存在一组不全为 0 的数 $a_1, a_2, \cdots, a_n \in F$ 使得

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = \mathbf{0},$$

则称 v_1, v_2, \cdots, v_n 是线性相关的. 反之只有在 a_1, a_2, \cdots, a_n 均为 0 上式成立, 则称 v_1, v_2, \cdots, v_n 是线性无关的.

线性相关和线性无关的性质

Proposition 3.12. 若 v_1, v_2, \dots, v_n 线性相关, 则下面结论均成立.

1. **内部线性表出** v_1, v_2, \dots, v_n 至少有一个向量可以由其余的向量组表出;
2. **齐次线性方程组** 若 v_1, v_2, \dots, v_n 表示齐次线性方程的系数矩阵的列向量组, 则该齐次方程组有非零解;
3. **行列式** 以 v_1, v_2, \dots, v_n 为行 (列) 向量组成的矩阵的行列式等于零;
4. **线性表出任意向量** 若取定另外一个向量 $v \in V$ 可以被 v_1, v_2, \dots, v_n 线性表出, 则表出方式有无穷种.
5. **向量组与部分组的关系** 若取定另外一个向量 $v \in V$, 则 v_1, v_2, \dots, v_n, v 也是线性相关.
6. **缩短向量组** 若将 v_1, v_2, \dots, v_n 中每个向量都去掉相同位置的 m 个分量, 则得到的新的向量组也是线性相关. 相对于缩小了原来这 n 个向量生成的向量空间维数, n 个向量在较大维数的向量空间是线性相关的, 那么在更小维数的向量空间一定还是线性相关的.

Proposition 3.13. 若 v_1, v_2, \dots, v_n 线性无关, 则下面结论均成立.

1. **内部线性表出** v_1, v_2, \dots, v_n 中的每一个向量都不能被其余的向量组表出;
2. **齐次线性方程组** 若 v_1, v_2, \dots, v_n 表示齐次线性方程的系数矩阵的列向量组, 则该齐次方程组只有零解;
3. **行列式** 以 v_1, v_2, \dots, v_n 为行 (列) 向量组成的矩阵的行列式不等于零;
4. **线性表出任意向量** 若取定另外一个向量 $v \in V$ 可以被 v_1, v_2, \dots, v_n 线性表出, 则表出方式唯一.
5. **向量组与部分组的关系** 部分向量组 $v_{i_1}, v_{i_2}, \dots, v_{i_s}$ 也是线性无关.
6. **延伸向量组** 若将 v_1, v_2, \dots, v_n 中每个向量都新增相同位置的 m 个分量, 则得到的新的向量组也是线性无关.

极大线性无关组

Definition 3.14. 给定某个向量组, 其基数 (向量个数) 最大的线性无关的部分向量组称为极大线性无关组.

向量组等价

Definition 3.15. 若 v_1, \dots, v_n 的每一个向量都可以由向量组 w_1, \dots, w_m 线性表出, 那么就称向量组 v_1, \dots, v_n 可以由向量组 w_1, \dots, w_m 线性表出.

Definition 3.16. 等价关系定义

Definition 3.17. 若向量组 v_1, \dots, v_n 和向量组 w_1, \dots, w_m 可以相互线性表出, 则称向量组 v_1, \dots, v_n 和向量组 w_1, \dots, w_m 等价, 记做

$$\{v_1, \dots, v_n\} \cong \{w_1, \dots, w_m\}.$$

Proposition 3.18. 向量组与它的极大线性组等价.

向量组的秩

Definition 3.19. 向量组的极大线性无关组中所含向量的个数被称为该向量组的秩. 全体由零向量组成的向量组的值规定为 0.

Lemma 3.20. 设向量组 v_1, v_2, \dots, v_n 可以由 v'_1, v'_2, \dots, v'_m 线性表出, 如果 $n > m$, 则 v_1, v_2, \dots, v_n 线性相关.

证明. 如何优雅的证明这个 lemma? 如果 v_1, v_2, \dots, v_n 可以由 v'_1, v'_2, \dots, v'_m 线性表出, 那么

$$\text{span}(v_1, v_2, \dots, v_n) \subseteq \text{span}(v'_1, v'_2, \dots, v'_m)$$

于是 $\dim \text{span}(v_1, v_2, \dots, v_n) \leq \dim \text{span}(v'_1, v'_2, \dots, v'_m) \leq m$. 因此当 $n > m$ 时, v_1, v_2, \dots, v_n 是线性相关的. \square

Lemma 3.21. (3.20的逆否命题) 设向量组 v_1, v_2, \dots, v_n 可以由 v'_1, v'_2, \dots, v'_m 线性表出, 如果 v_1, v_2, \dots, v_n 线性无关, 则 $n \leq m$.

Proposition 3.22. 若向量组 I 可以由向量组 II 线性表出, 那么

$$\text{rank } I \leq \text{rank } II.$$

证明. $\text{span}(I) \subseteq \text{span}(II)$ \square

Corollary 3.23. 等价的线性无关的向量组所含向量的个数相等.

Corollary 3.24. 向量组的任意两个极大线性无关组所含向量相等.

Proposition 3.25. 线性无关组的秩和其向量组个数相等.

有限维线性空间的基和维数

Definition 3.26. 给定向量空间 V ，取其空间上一组向量 v_1, v_2, \dots, v_n 若满足

1. v_1, v_2, \dots, v_n 线性无关,
2. $\forall v \in V$, v 都可以被 v_1, v_2, \dots, v_n 线性表出,

则称 v_1, v_2, \dots, v_n 为 V 的一个基.

Proposition 3.27. 向量空间上所有基所含向量的个数均相等.

Definition 3.28. 向量空间一个基所含向量的个数被称为该向量空间的维数.

矩阵的秩

Definition 3.29. 矩阵的行向量生成的向量空间的维数为矩阵的行秩，同理矩阵的列向量生成的向量空间的维数为矩阵的列秩.

Lemma 3.30. 阶梯型矩阵的行秩和列秩相同.

Lemma 3.31. 初等行变换不改变矩阵的行秩和列秩.

Proposition 3.32. 矩阵的行秩和列秩相同.

Definition 3.33. 矩阵的行秩及列秩都统称为矩阵的秩.

Corollary 3.34. 矩阵通过初等变换化成行阶梯形矩阵的非零个数就是矩阵的秩.

Theorem 3.35. 矩阵的秩等于它的不为零的子式的最高阶数.

Corollary 3.36. 极大线性无关组的一个求法 矩阵的不为零的最高阶数的子式所在的行和列是原矩阵的行向量和列向量的极大线性无关组.

Annotation 3.37. 化行阶梯形之后在去找最高阶不为 0 的子式更方便一点.

Definition 3.38. n 阶矩阵 A 的秩等于 n ，则称 A 是一个满秩矩阵.

Theorem 3.39. n 阶矩阵 A 的秩等于 n 当且仅当 $|A| \neq 0$.

方程组有解的充要条件

Theorem 3.40. 给定任意 n 线性方程组 (任意方程个数)

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$$

有解的充分必要条件是: 它的系数矩阵与增广矩阵的秩相同.

证明. 解的判定是看是否 $\beta \in \text{span}(\alpha_1, \alpha_2, \cdots, \alpha_n)$. □

Theorem 3.41. 给定 n 元线性方程组有解时, 如果它的系数矩阵 A 的秩等于 n , 那么原方程组有唯一解; 如果 A 的秩小于 n , 那么原方程组有无穷多个解.

Corollary 3.42. 齐次线性方程组有非零解的条件是其系数矩阵的秩小于未知数的个数.

Proposition 3.43. 非齐次线性方程组无解当且仅当系数矩阵的秩的增广矩阵的秩减 1, 即

$$\text{rank}(A) + 1 = \text{rank}(\bar{A}).$$

其中 \bar{A} 表示原方程组的增广矩阵.

齐次方程组解集结构

Proposition 3.44. n 元齐次线性方程组的解集 $\left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\}$ 构成 F^n 的一个子空间.

Definition 3.45. 齐次线性方程组有非零解时, 如果它的有限多个解 $\eta_1, \eta_2, \cdots, \eta_t$ 满足

1. $\eta_1, \eta_2, \cdots, \eta_t$ 线性无关;
2. 该齐次线性方程组每一个解都可以由 $\eta_1, \eta_2, \cdots, \eta_t$ 线性表出,

那么就称 $\eta_1, \eta_2, \cdots, \eta_t$ 是该齐次线性方程组的一个基础解系.

Theorem 3.46. n 元齐次线性方程组的解空间的维数为

$$\dim W = n - \text{rank}(A),$$

其中 A 是它的系数矩阵.

证明. 分两种情况: 若 $\text{rank}(\mathbf{A}) = n$, 即原方程组只有零解, 那么 $W = \mathbf{0}$. 自然地, 命题中等式成立; 若 $\text{rank}(\mathbf{A}) < n$.

(step 1) 把 \mathbf{A} 通过初等行变换化成简化行阶梯形矩阵 J , 设 J 有 r 个主元, 不妨设它们分别在 $1, 2, \dots, r$ 列. 于是该齐次线性方程组的一般解为

$$\begin{cases} x_1 = -b_{1,r+1}x_{r+1} - b_{1,r+2}x_{r+2} - \dots - b_{1n}x_n \\ x_2 = -b_{2,r+1}x_{r+1} - b_{2,r+2}x_{r+2} - \dots - b_{2n}x_n \\ \vdots \\ x_r = -b_{r,r+1}x_{r+1} - b_{r,r+2}x_{r+2} - \dots - b_{rn}x_n \end{cases}$$

其中 $x_{r+1}, x_{r+2}, \dots, x_n$ 是自由变量.

(step 2) 让自由未知量 $x_{r+1}, x_{r+2}, \dots, x_n$ 分别取下述 $n-r$ 组数

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

则由一般解得到原方程组的 $n-r$ 解, 设其为 $\eta_1, \eta_2, \dots, \eta_{n-r}$, 其中

$$\eta_1 = \begin{pmatrix} -b_{1,r+1} \\ -b_{2,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \eta_{n-r} = \begin{pmatrix} -b_{1n} \\ -b_{2n} \\ \vdots \\ -b_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

上面两步是求一个基础解系的方法. 下面需要证明 $\eta_1, \eta_2, \dots, \eta_{n-r}$ 可以线性表出原线性方程组的所有解. \square

非齐次线性方程组解集结构

Lemma 3.47. 给定 n 元非齐次线性方程组, 设它的解集为 U . 若将其常数项都换成 0 变为齐次线性方程组, 称其为原非齐次线性方程的导出组, 设它的解集为 W .

1. 对任意的 $u_1, u_2 \in U$, 有 $u_2 - u_1 \in W$.

2. 对任意的 $u \in U$ 和 $w \in W$, 有 $u + w \in U$.

Proposition 3.48. affine set 若 n 元非齐次线性方程组有解, 那么它的解集 U 为

$$U = \{u_0 + w \mid w \in W\},$$

其中 u_0 是取定 U 中的一个特解, W 是原非齐次线性方程的导出组的解集.

Annotation 3.49. 求非齐次线性方程组的解集的步骤如下

1. 求出非齐次线性方程组的一般解. 让自由未知量都取 0, 得到一个特解 γ_0 ;
2. 求导出组的一个基础解系 η_1, \dots, η_t ;
3. 写出非齐次线性方程组的解集 U

$$U = \{\gamma_0 + k_1\eta_1 + \dots + k_t\eta_t \mid k_1, \dots, k_t \in F\}.$$

矩阵运算

矩阵运算的定义

Definition 4.1. 给定 F 上 $m \times n$ 矩阵 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$, 则它们的加法定义为

$$A + B = (a_{ij} + b_{ij})_{m \times n}.$$

Definition 4.2. 给定 F 上 $m \times n$ 矩阵 $A = (a_{ij})_{m \times n}$ 和数 $\lambda \in F$, 则它们的数量乘法定义为

$$\lambda A = (\lambda a_{ij})_{m \times n}.$$

Proposition 4.3. 所有 $m \times n$ 矩阵的集合记为 $F^{m,n}$, 则 $F^{m,n}$ 构成一个向量空间.

Definition 4.4. 给定 F 上 $m \times s$ 矩阵 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$, 则它们的乘法定义为

$$AB = C,$$

其中 $C = (c_{ij})_{m \times n}$ 是一个 $m \times n$ 矩阵, C 的第 i 行和第 j 列为

$$c_{ij} = \sum_{r=1}^s a_{ir} b_{rj}.$$

矩阵运算的性质

Proposition 4.5. 矩阵乘法的基本性质性质

1. 结合性 设 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times t}$, $C = (c_{ij})_{t \times m}$, 则 $(AB)C = A(BC)$.
2. 分配性 设 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$, $C = (c_{ij})_{s \times n}$, 则 $A(B + C) = AB + AC$.

Definition 4.6. 主对角线上元素都是 1, 其余元素均为 0 的 n 级矩阵称为 n 级单位矩阵, 记为 I_n , 或者简记为 I .

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

Definition 4.7. n 级矩阵的非负整数次幂定义为

$$A^m = \underbrace{A \cdot A \cdot A \cdots A}_{m \uparrow}, m \in \mathbb{Z}^+;$$
$$A^0 = I.$$

Proposition 4.8. 矩阵加法, 数量乘法和乘法与矩阵转置的关系如下

1. $(A + B)^T = A^T + B^T$;
2. $(kA)^T = kA^T$;
3. $(AB)^T = B^T A^T$.

证明. (3) 设 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$, 则

$$\begin{aligned}(AB)^T(i; j) &= AB(j; i) = \sum_{r=1}^s a_{jr} b_{ri} \\ (B^T A^T)(i; j) &= \sum_{r=1}^s b'_{ir} a'_{rj} = \sum_{r=1}^s b_{ri} a_{jr},\end{aligned}$$

其中 b'_{ij} 表示 $B^T(i; j)$. □

Proposition 4.9. 设 $A = (a_{ij})_{m \times n}$, 若有等式

$$A^T A = 0,$$

则 $A = 0$.

证明. 若有命题前提条件, 则

$$(A^T A)(i; j) = \sum_{r=1}^m b_{ir} a_{rj} = \sum_{r=1}^m a_{ri} a_{rj} = 0.$$

当 $i = j$ 时, 显然有

$$a_{ri} = 0, r = 1, 2, \dots, m.$$

把 i 取遍 $1, 2, \dots, n$, 就得到了 $A = 0$. □

Annotation 4.10. $AB = 0$ 并不能确定 $A = 0$ 或者 $B = 0$. 但若 A 是可逆的, 则可推出 $B = 0$.

矩阵乘法中的线性组合

Proposition 4.11. 设 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$, 把 A 列向量组记为 $\alpha_1, \alpha_2, \dots, \alpha_s$, 则

$$AB = (\alpha_1, \alpha_2, \dots, \alpha_s) \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{s1} & b_{s2} & \cdots & b_{sn} \end{pmatrix} = (b_{11}\alpha_1 + b_{21}\alpha_2 + \cdots + b_{s1}\alpha_s, \dots, b_{1n}\alpha_1 + b_{2n}\alpha_2 + \cdots + b_{sn}\alpha_s).$$

即 AB 每一个列向量都是 A 列向量组的线性组合.

Proposition 4.12. 设 $A = (a_{ij})_{m \times s}$, $B = (b_{ij})_{s \times n}$, 把 B 行向量组记为 $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{pmatrix}$, 则

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{s2} & \cdots & a_{ms} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{pmatrix} = \begin{pmatrix} a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1s}\beta_s \\ a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2s}\beta_s \\ \vdots \\ a_{m1}\beta_1 + a_{m2}\beta_2 + \cdots + a_{ms}\beta_s \end{pmatrix}.$$

即 AB 每一个行向量都是 B 行向量组的线性组合.

Proposition 4.13. 乘积矩阵的秩的重要上界 设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$, 则

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.$$

证明. 由4.11和4.12, 我们 AB 的列向量组可以由 A 的列向量组线性表出, AB 的行向量组可以由 B 的列向量线性表出, 自然地就得到了命题的结论. \square

特殊矩阵

Definition 4.14. 只有一个元素是 1, 其余元素均为 0 的矩阵我们称为基本矩阵. (i, j) 元为 1 的基本矩阵记为 E_{ij}

Definition 4.15. 除对主角线以外的元素均为 0 的方阵称为对角矩阵.

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

记为 $\text{diag}\{d_1, d_2, \cdots, d_n\}$.

Definition 4.16. 主对角线均为 λ 的对角矩阵称为数量矩阵.

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix},$$

记为 λI_n .

Definition 4.17. 主对角线以下 (上) 均为 0 的方阵称为上 (下) 三角矩阵.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix},$$

Definition 4.18. 由单位矩阵经过一次初等行 (列) 变换得到的矩阵称为初等矩阵, 其对应 3 种初等变换如下 (箭头之上表示初等行变换, 箭头之下表示初等列变换)

1. $\mathbf{I} \xrightarrow{\oplus+k\oplus} P(j, i(k))$ 或者 $\mathbf{I} \xrightarrow[\oplus+k\oplus]{} P(j, i(k)),$

$$i \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ j & & k & \cdots & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

2. $\mathbf{I} \xrightarrow{(\oplus, \oplus)} P(i, j)$ 或者 $\mathbf{I} \xrightarrow[(\oplus, \oplus)]{} P(i, j),$

$$i \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & \cdots & 1 \\ & & \vdots & \ddots & \vdots \\ j & & 1 & \cdots & 0 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

3. $\mathbf{I} \xrightarrow{\oplus \cdot \lambda} P(i(\lambda))$ 或者 $\mathbf{I} \xrightarrow[\oplus \cdot \lambda]{} P(j(\lambda))$,

$$i \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & \lambda & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & \ddots \\ j & & & & & & 1 \end{pmatrix}$$

Theorem 4.19. 用初等矩阵左 (右) 乘一个矩阵 \mathbf{A} , 就相当于 \mathbf{A} 作了一次相应的初等行 (列) 变换.

证明. 设矩阵 $\mathbf{A} = (a_{ij})_{m \times n}$ 的行向量组为 $\alpha_1, \alpha_2, \dots, \alpha_m$, 列向量组为 $\gamma_1, \gamma_2, \dots, \gamma_n$. 则对 \mathbf{A} 左乘 $P(j, i(k))$ 有

$$P(j, i(k))\mathbf{A} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & k & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_j + k\alpha_i \\ \vdots \\ \alpha_n \end{pmatrix},$$

同理对 \mathbf{A} 右乘 $P(j, i(k))$ 有

$$\mathbf{A}P(j, i(k)) = (\gamma_1, \gamma_2, \dots, \gamma_m) \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & k & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} = (\gamma_1, \gamma_2, \dots, \gamma_i + k\gamma_j, \dots, \gamma_m).$$

其余两种变换都比较显然. □

Example 4.20.

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & 0 & 0 \\ 0 & a_2 b_2 & 0 \\ 0 & 0 & a_3 b_3 \end{pmatrix}$$

Example 4.21.

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}^n = \begin{pmatrix} a_1^n & 0 & 0 \\ 0 & a_2^n & 0 \\ 0 & 0 & a_3^n \end{pmatrix}$$

Example 4.22.

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{a_1} & 0 & 0 \\ 0 & \frac{1}{a_2} & 0 \\ 0 & 0 & \frac{1}{a_3} \end{pmatrix}$$

Definition 4.23. 一个矩阵 \mathbf{A} 如果满足 $\mathbf{A}^T = \mathbf{A}$, 那么就称 \mathbf{A} 是**对称矩阵**.

Proposition 4.24. 对称矩阵是一个方阵, 并且 n 级矩阵 \mathbf{A} 是对称矩阵当且仅当

$$A(i; j) = A(j; i), i, j = 1, 2, \dots, n.$$

Proposition 4.25. 若 \mathbf{A}, \mathbf{B} 都是 n 级对称矩阵, 则 \mathbf{AB} 为对称矩阵当且仅当 $\mathbf{AB} = \mathbf{BA}$.

证明.

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T = \mathbf{BA}.$$

□

Definition 4.26. 一个矩阵 \mathbf{A} 如果满足 $\mathbf{A}^T = -\mathbf{A}$, 那么称 \mathbf{A} 是**斜对角矩阵**.

Proposition 4.27. 斜对角矩阵是一个方阵, 并且 n 级矩阵 \mathbf{A} 是斜对角矩阵当且仅当

$$A(i; j) = -A(j; i), i, j = 1, 2, \dots, n.$$

Proposition 4.28. 奇数级斜对角矩阵的行列式等于 0.

证明. 若 $\mathbf{A}^T = -\mathbf{A}$, 那么 $|\mathbf{A}^T| = |\mathbf{A}| = |-\mathbf{A}| = (-1)^n |\mathbf{A}|$. 所以当 n 为奇数时, 有 $|\mathbf{A}| = -|\mathbf{A}| = 0$

□

可逆矩阵

Definition 4.29. 给定矩阵 A , 若存在矩阵 B , 使得

$$AB = BA = I,$$

那么称 A 是一个可逆矩阵(非奇异矩阵). 若 A 是一个可逆矩阵, 满足上式的 B 记为 A^{-1} .

Proposition 4.30. 可逆矩阵一定是方阵. 这里的意思是上面的 I 是同一个.

Proposition 4.31. n 级矩阵 A 可逆当且仅当它的秩等于 n (满秩矩阵), 即 $|A| \neq 0$.

证明. (\Rightarrow)

$$n = \text{rank}(I) = \text{rank}(AA^{-1}) \leq \text{rank}(A) \leq n.$$

(\Leftarrow)

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix},$$

其中 A_{ij} 表示 $|A|$ 的第 i 阶行, 第 j 列的代数余子式. □

Definition 4.32. 命题4.31中矩阵

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

被称为 A 的伴随矩阵, 记为 A^* . 即有

$$AA^* = |A|I \text{ 和 } A^*A = |A|I.$$

若 A 可逆时, 那么 $|A|$ 的逆矩阵记为

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

Proposition 4.33. 伴随矩阵相关性质如下

1. $AA^* = |A|I$;
2. 若 $|A| \neq 0$, 则 $(A^*)^{-1} = \frac{A}{|A|}$;

3. $(k\mathbf{A})^* = k^{n-1}\mathbf{A}^*$;
4. $(\mathbf{A}^*)^T = (\mathbf{A}^T)^*$;
5. $|\mathbf{A}^*| = |\mathbf{A}|^{n-1}$;
6. $(\mathbf{A}^*)^* = |\mathbf{A}|^{n-2}\mathbf{A}$, $n \geq 2$.

证明. (3)

$$\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nn} \end{pmatrix} \begin{pmatrix} k^{n-1}A_{11} & k^{n-1}A_{21} & \cdots & k^{n-1}A_{n1} \\ k^{n-1}A_{12} & k^{n-1}A_{22} & \cdots & k^{n-1}A_{n2} \\ \vdots & \vdots & & \vdots \\ k^{n-1}A_{1n} & k^{n-1}A_{2n} & \cdots & k^{n-1}A_{nn} \end{pmatrix} = \begin{pmatrix} k^n|\mathbf{A}| & 0 & \cdots & 0 \\ 0 & k^n|\mathbf{A}| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & k^n|\mathbf{A}| \end{pmatrix}$$

(4)

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

□

(5)

$$|\mathbf{A}\mathbf{A}^*| = |\mathbf{A}||\mathbf{A}^*| = |\mathbf{A}|^n.$$

(6)

$$\mathbf{A}\mathbf{A}^*(\mathbf{A}^*)^* = |\mathbf{A}|(\mathbf{A}^*)^* = \mathbf{A} \begin{pmatrix} |\mathbf{A}|^{n-1} & 0 & \cdots & 0 \\ 0 & |\mathbf{A}|^{n-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |\mathbf{A}|^{n-1} \end{pmatrix},$$

因此 $(\mathbf{A}^*)^* = |\mathbf{A}|^{n-2}\mathbf{A}$.

Proposition 4.34. 设 \mathbf{A} 与 \mathbf{B} 都是 n 级矩阵, 若满足

$$\mathbf{A}\mathbf{B} = \mathbf{I},$$

则 \mathbf{A} 和 \mathbf{B} 互为逆矩阵. 这个命题告诉我们判定 n 阶方阵的只需要判定一个条件即可.

证明. 首先由

$$n = \text{rank}(\mathbf{I}) = \text{rank}(\mathbf{AB}) \leq \mathbf{A} \leq n$$

所以 \mathbf{A} 是一个满秩矩阵, 同理可证 \mathbf{B} 也是一个满秩矩阵. 即 \mathbf{A} 和 \mathbf{B} 均可逆. 命题等式两边同时乘上 \mathbf{A}^{-1}

$$\mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{I},$$

即有 $\mathbf{B} = \mathbf{A}^{-1}$, 同理可证 $\mathbf{A} = \mathbf{B}^{-1}$. □

Proposition 4.35. 初等矩阵都是可逆矩阵.

证明.

$$P(j, i(-k))P(j, i(k)) = \mathbf{I},$$

$$P(i, j)P(i, j) = \mathbf{I},$$

$$P(i(\frac{1}{\lambda}))P(i(\lambda)) = \mathbf{I}$$

□

Proposition 4.36. 可逆矩阵基本性质

1. 单位矩阵 \mathbf{I} 可逆, $\mathbf{I}^{-1} = \mathbf{I}$;
2. 如果 \mathbf{A} 可逆, 那么 \mathbf{A}^{-1} 也可逆, 且 $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$;
3. 如果 n 级矩阵 \mathbf{A}, \mathbf{B} 可逆, 那么 \mathbf{AB} 也可逆, 且 $(\mathbf{AB})^{-1} = (\mathbf{B}^{-1}\mathbf{A}^{-1})$. 该性质可以推广到 2 个以上 n 级矩阵.
4. 如果 \mathbf{A} 可逆, 那么 \mathbf{A}^T 也可逆, 并且 $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

Lemma 4.37. 重要的变换性质 可逆矩阵经过初等矩阵行变换成的简化梯形矩阵一定是单位矩阵.

证明. 可逆矩阵是满秩矩阵, 因此一定是单位矩阵. □

Theorem 4.38. 矩阵 \mathbf{A} 可逆当且仅当它可以表示成一些初等矩阵的乘积.

证明. (\Rightarrow) 由 lemma 4.37, 可知存在初等矩阵 P_1, P_2, \dots, P_t , 使得

$$P_t \cdots P_2 P_1 \mathbf{A} = \mathbf{I}.$$

因此由 proposition 4.34 可知 $(P_t \cdots P_2 P_1)$ 和 \mathbf{A} 可逆, 再有 proposition 4.35 和 proposition 4.36 中的 (3) 有

$$\mathbf{A} = P_1^{-1} P_2^{-1} \cdots P_t^{-1}.$$

(\Leftarrow) 若 $\mathbf{A} = P_1' P_2' \cdots P_t'$, 由 proposition 4.35 和 proposition 4.36 中的 (3) 可知 $(P_1' P_2' \cdots P_t')$ 可逆, 即 \mathbf{A} 可逆. □

Proposition 4.39. 重要的变换性质 用一个可逆矩阵左 (右) 乘一个矩阵 \mathbf{A} , 不改变 \mathbf{A} 的秩.

证明. 可逆矩阵可以分解成多个初等矩阵的乘积, 那么每个初等矩阵作用在 \mathbf{A} 上等价于对 \mathbf{A} 做初等变换, 而初等变换不改变矩阵的秩. \square

分块矩阵

Definition 4.40. 由矩阵 \mathbf{A} 的若干行和若干列的交叉位置元素按原来顺序排成的矩阵称为 \mathbf{A} 的子矩阵.

Definition 4.41. 把一个矩阵 \mathbf{A} 的行分成若干组, 列也分成若干组, 从而 \mathbf{A} 被分成了若干子矩阵, 把 \mathbf{A} 看成是由这些子矩阵组成的, 这称为矩阵的分块. 这种由子矩阵组成的矩阵被称为分块矩阵.

Proposition 4.42. 设 $\mathbf{A} = (a_{ij})_{m \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$, 则

$$\begin{array}{cccccccc} & s_1 & s_2 & \cdots & s_t & n_1 & n_2 & \cdots & n_v \\ \begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_u \end{matrix} & \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & & \vdots \\ A_{u1} & A_{u2} & \cdots & A_{ut} \end{pmatrix} & \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1v} \\ B_{21} & B_{22} & \cdots & B_{2v} \\ \vdots & \vdots & & \vdots \\ B_{t1} & B_{t2} & \cdots & B_{tv} \end{pmatrix} & \begin{matrix} s_1 \\ s_2 \\ \vdots \\ s_t \end{matrix} \end{array}$$

$$= \begin{pmatrix} \sum_{r=1}^t A_{1r} B_{r1} & \cdots & \sum_{r=1}^t A_{1r} B_{rv} \\ \sum_{r=1}^t A_{2r} B_{r1} & \cdots & \sum_{r=1}^t A_{2r} B_{rv} \\ \vdots & & \vdots \\ \sum_{r=1}^t A_{ur} B_{r1} & \cdots & \sum_{r=1}^t A_{ur} B_{rv} \end{pmatrix}$$

证明. (1) 纯代数手法: 直接比较分块矩阵乘法得到的 (i, j) 元和原常规矩阵乘法的 (i, j) 元. 棘手的地方在于你要确定 (i, j) 元在哪两个子矩阵乘积中.

(2) 递归分解:

$$\begin{aligned}
& \begin{matrix} & s_1 & s_2 & \cdots & s_t & n_1 & n_2 & \cdots & n_v \\ m_1 & \left(\begin{array}{c|ccc} A_{11} & A_{12} & \cdots & A_{1t} \end{array} \right) & \left(\begin{array}{cccc} B_{11} & B_{12} & \cdots & B_{1v} \end{array} \right) & s_1 \\ m_2 & \left(\begin{array}{c|ccc} A_{21} & A_{22} & \cdots & A_{2t} \end{array} \right) & \left(\begin{array}{cccc} B_{21} & B_{22} & \cdots & B_{2v} \end{array} \right) & s_2 \\ \vdots & \left(\begin{array}{c|ccc} \vdots & \vdots & & \vdots \end{array} \right) & & & \vdots \\ m_u & \left(\begin{array}{c|ccc} A_{u1} & A_{u2} & \cdots & A_{ut} \end{array} \right) & \left(\begin{array}{cccc} B_{t1} & B_{t2} & \cdots & B_{tv} \end{array} \right) & s_t \end{matrix} \\
&= \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{u1} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1v} \end{pmatrix} + \begin{matrix} & s_2 & \cdots & s_t & n_1 & n_2 & \cdots & n_v \\ m_1 & \left(\begin{array}{c|ccc} A_{12} & \cdots & A_{1t} \end{array} \right) & \left(\begin{array}{cccc} B_{21} & B_{22} & \cdots & B_{2v} \end{array} \right) & s_2 \\ m_2 & \left(\begin{array}{c|ccc} A_{22} & \cdots & A_{2t} \end{array} \right) & \left(\begin{array}{cccc} \vdots & \vdots & & \vdots \end{array} \right) & \vdots \\ \vdots & \left(\begin{array}{c|ccc} \vdots & & & \vdots \end{array} \right) & & & \vdots \\ m_u & \left(\begin{array}{c|ccc} A_{u2} & \cdots & A_{ut} \end{array} \right) & \left(\begin{array}{cccc} B_{t1} & B_{t2} & \cdots & B_{tv} \end{array} \right) & s_t \end{matrix}
\end{aligned}$$

我们通过组合部分子矩阵构造了新的分块矩阵, 所以我们只要证明两个更为简单的分块矩阵乘法即可, 即

$$\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_1 B_1 + A_2 B_2 \end{pmatrix}$$

和

$$\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_u \end{pmatrix} \begin{pmatrix} B_1 & B_2 & \cdots & B_v \end{pmatrix} = \begin{pmatrix} A_1 B_1 & A_1 B_2 & \cdots & A_1 B_v \\ A_2 B_1 & A_2 B_2 & \cdots & A_2 B_v \\ \vdots & \vdots & & \vdots \\ A_u B_1 & A_u B_2 & \cdots & A_u B_v \end{pmatrix}$$

相对来要所要比直接证明一般的分部矩阵乘法要稍微简单一点. □

Example 4.43.

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

Example 4.44.

$$\begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{-1} \end{pmatrix}$$

Example 4.45.

$$\begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{0} & \mathbf{C}^{-1} \\ \mathbf{B}^{-1} & \mathbf{0} \end{pmatrix}$$

分块矩阵的初等变换

Proposition 4.46. 分块矩阵的初等行变换操作

1. 把一个块行的左乘 P (P 是矩阵) 倍加到另一个块行上,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{\mathcal{Q}+P\cdot\mathcal{Q}} \begin{pmatrix} A_1 & A_2 \\ PA_1 + A_3 & PA_2 + A_4 \end{pmatrix}.$$

2. 互换两个块行的位置,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{(\mathcal{Q},\mathcal{Q})} \begin{pmatrix} A_3 & A_4 \\ A_1 & A_2 \end{pmatrix}.$$

3. 用一个可逆矩阵左乘某一块行,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{P\cdot\mathcal{Q}} \begin{pmatrix} PA_1 & PA_2 \\ A_3 & A_4 \end{pmatrix}.$$

Proposition 4.47. 分块矩阵的初等列变换操作

1. 把一个列行的右乘 P (P 是矩阵) 倍加到另一个列行上,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{\mathcal{Q}+\mathcal{Q}\cdot P} \begin{pmatrix} A_1 & A_1P + A_2 \\ A_3 & A_3P + A_4 \end{pmatrix}.$$

2. 互换两个列行的位置,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{(\mathcal{Q},\mathcal{Q})} \begin{pmatrix} A_2 & A_1 \\ A_4 & A_3 \end{pmatrix}.$$

3. 用一个可逆矩阵右乘某一块列,

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \xrightarrow{P\cdot\mathcal{Q}} \begin{pmatrix} A_1 & A_2P \\ A_3 & A_4P \end{pmatrix}.$$

Definition 4.48. 把单位矩阵分块得到的矩阵经过一次分块矩阵的初等行 (列) 变换得到的矩阵被称为分块初等矩阵, 如下

- 1.

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \xrightarrow{\mathcal{Q}+P\cdot\mathcal{Q}} \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$$

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \xrightarrow{\mathcal{Q}+\mathcal{Q}\cdot P} \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$$

2.

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \xrightarrow{(\oplus, \otimes)} \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \xrightarrow{(\oplus, \otimes)} \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}$$

3.

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \xrightarrow{P \cdot \otimes} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{P} \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \xrightarrow{\otimes \cdot P} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & P \end{pmatrix}$$

Theorem 4.49. 用分块初等矩阵左乘一个分块矩阵, 就相当于对这个分块矩阵做了一次相应的分块矩阵的初等行变换. 同理用分块初等矩阵右乘一个分块矩阵, 就相当于对这个分块矩阵做了一次相应的分块矩阵的初等列变换.

证明.

$$\begin{pmatrix} \mathbf{I} & 0 \\ P & \mathbf{I} \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ PA_1 + A_3 & PA_2 + A_4 \end{pmatrix}$$

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ P & \mathbf{I} \end{pmatrix} = \begin{pmatrix} A_1 + A_2P & A_2 \\ A_3 + A_4P & A_4 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} A_3 & A_4 \\ A_1 & A_2 \end{pmatrix}$$

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} = \begin{pmatrix} A_2 & A_1 \\ A_4 & A_3 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{P} \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ PA_3 & PA_4 \end{pmatrix}$$

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{P} \end{pmatrix} = \begin{pmatrix} A_1 & A_2P \\ A_3 & A_4P \end{pmatrix}$$

□

Theorem 4.50. 重要变换性质 分块矩阵的初等行 (列) 变换不改变矩阵的秩.

证明. 初等分块矩阵都是可逆矩阵.

□

矩阵乘积的行列式

Lemma 4.51. 设 $A = (a_{ij})_{n \times s}$, $B = (b_{ij})_{s \times n}$, 若 $n > s$, 则

$$|AB| = 0.$$

证明. 由 proposition 4.13可知

$$\text{rank}(AB) < \text{rank}(B) < s,$$

而 AB 是一个 $n \times n$ 矩阵, 因此它不是满秩矩阵, 即行列式等于 0. □

Lemma 4.52. 设 $A = (a_{ij})_{n \times s}$, $B = (b_{ij})_{s \times n}$, 若 $n = s$, 则

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

证明. 下面分两种情况讨论:

若 $|A| = 0$, 那么 A 不是满秩, 由 proposition 4.13可知 $|AB|$ 也不是满秩, 即 $|AB| = 0$.

若 $|A| \neq 0$, 那么 A 就可逆. 因此由 proposition 4.38有

$$A = P_1 P_2 \cdots P_t I,$$

其中 P_1, P_2, \dots, P_t 是一些初等矩阵. 那么

$$AB = (P_1 P_2 \cdots P_t)B.$$

下面考虑三种左乘初等矩阵对矩阵 B 的行列式改变, 即分三种情况

$$|P(j, i(k))B| = |B|$$

$$|P(i, j)B| = -|B|$$

$$|P(i(k))B| = k|B|$$

. 特别地, 当上述 B 取单位矩阵 I 时, 我们有

$$|P(j, i(k))I| = |P(j, i(k))| = |I| = 1$$

$$|P(i, j)I| = |P(i, j)| = -|I| = -1$$

$$|P(i(k))I| = |P(i(k))| = k|I| = k$$

于是

$$|P(j, i(k))B| = |P(j, i(k))||B|$$

$$|P(i, j)B| = |P(i, j)||B|$$

$$|P(i(k))B| = |P(i(k))||B|$$

. 综上

$$|\mathbf{AB}| = |P_1 P_2 \cdots P_t \mathbf{B}| = |P_1| |P_2| \cdots |P_t| |\mathbf{B}|.$$

特别地, 当 $\mathbf{B} = \mathbf{I}$ 时,

$$|\mathbf{A}| = |P_1| |P_2| \cdots |P_t|,$$

所以 $|\mathbf{AB}| = |\mathbf{A}| |\mathbf{B}|$. □

Theorem 4.53. Binet-Cauchy 公式 设 $\mathbf{A} = (a_{ij})_{n \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$, 若 $n \leq s$, 则

$$|\mathbf{AB}| = \sum_{1 \leq v_1 < v_2 < \cdots < v_n \leq s} \mathbf{A} \begin{pmatrix} 1, & 2, & \cdots, & n \\ v_1, & v_2, & \cdots, & v_n \end{pmatrix} \mathbf{B} \begin{pmatrix} v_1, & v_2, & \cdots, & v_n \\ 1, & 2, & \cdots, & n \end{pmatrix}.$$

即 $|\mathbf{AB}|$ 等于 \mathbf{A} 的所有 n 阶子式与 \mathbf{B} 的相应 n 阶子式的乘积之和.

证明. 考虑特殊分块矩阵的行列值

$$\begin{vmatrix} \mathbf{A} & 0 \\ \mathbf{I}_n & \mathbf{B} \end{vmatrix} = |\mathbf{A}| |\mathbf{B}|.$$

对这个分块矩阵我们把 2 行的 $-\mathbf{A}$ 倍加到第一行上去就可以把 \mathbf{AB} 构造出来, 这个初等变换操作等价于

$$\begin{pmatrix} \mathbf{I}_n & -\mathbf{A} \\ 0 & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{A} & 0 \\ \mathbf{I}_n & \mathbf{B} \end{pmatrix} = \begin{pmatrix} 0 & -\mathbf{AB} \\ \mathbf{I}_n & \mathbf{B} \end{pmatrix}$$

□

矩阵运算中秩的性质

Proposition 4.54.

$$\operatorname{rank}(\mathbf{A}^T \mathbf{A}) = \operatorname{rank}(\mathbf{A})$$

证明. 将 \mathbf{A} 通过左乘初等矩阵变成行阶梯形矩阵 \mathbf{R} , 即

$$P_1 \cdots P_t \mathbf{A}^T = \mathbf{R}.$$

于是

$$\mathbf{A}^T = P_t^{-1} \cdots P_1^{-1} \mathbf{R},$$

也有

$$\mathbf{A} = \mathbf{R}^T (P_t^{-1} \cdots P_1^{-1})^T.$$

那么

$$\operatorname{rank}(\mathbf{A}^T \mathbf{A}) = \operatorname{rank}(P_t^{-1} \cdots P_1^{-1} \mathbf{R} \mathbf{R}^T (P_t^{-1} \cdots P_1^{-1})^T) = \operatorname{rank}(\mathbf{R} \mathbf{R}^T).$$

现在有一个好结论: 行阶梯形和它的转置的乘积的秩是等于原本这个行阶梯形的, 即

$$\operatorname{rank}(\mathbf{R} \mathbf{R}^T) = \operatorname{rank}(\mathbf{R}).$$

我们来证明这个结论. 若 \mathbf{R} 是满秩的, 上面这个结论是显然的; 若 \mathbf{R} 不是满秩的, 设 $\mathbf{R}^T = (\mathbf{R}_1^T, \mathbf{0})$, 其中 \mathbf{R}_1^T 是满秩的. 那么显然

$$\operatorname{rank}(\mathbf{R} \mathbf{R}^T) = \operatorname{rank}(\mathbf{R}_1 \mathbf{R}_1^T) = \operatorname{rank}(\mathbf{R}_1) = \operatorname{rank}(\mathbf{R}).$$

因此最终我们有

$$\operatorname{rank}(\mathbf{R}) = \operatorname{rank}(\mathbf{A}^T) = \operatorname{rank}(\mathbf{A}).$$

□

Corollary 4.55.

$$\operatorname{rank}(\mathbf{A}^T \mathbf{A}) = \operatorname{rank}(\mathbf{A} \mathbf{A}^T).$$

证明. 由4.54, 立即有

$$\operatorname{rank}(\mathbf{A}^T \mathbf{A}) = \operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^T) = \operatorname{rank}(\mathbf{A} \mathbf{A}^T).$$

□

Proposition 4.56.

$$\operatorname{rank}(\mathbf{A} + \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}).$$

证明. 设 \mathbf{A} 的列向量组极大线性无关组为 $\alpha_1, \dots, \alpha_s$, 即 $\text{rank}(\mathbf{A}) = s$; 设 \mathbf{B} 的列向量组极大线性无关组为 β_1, \dots, β_t , 即 $\text{rank}(\mathbf{B}) = s$; 那么 $\mathbf{A} + \mathbf{B}$ 的列向量组可以被 $\alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_t$ 线性表出, 所以有命题结论. \square

Proposition 4.57. 给定矩阵 $\mathbf{A} = (a_{ij})_{m \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$. 若 $\mathbf{AB} = \mathbf{0}$, 则

$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \leq s$$

证明. 设 $\mathbf{B} = (\mathbf{B}_1, \dots, \mathbf{B}_n)$, 那么

$$\mathbf{A}(\mathbf{B}_1, \dots, \mathbf{B}_n) = \mathbf{0}.$$

于是我们得到了 n 个齐次方程组

$$\mathbf{AB}_i = \mathbf{0}, i = 1, \dots, n.$$

即 B_i 是实际上 s 元齐次线性方程组 $\mathbf{AX} = \mathbf{0}$ 的一个解, 所以由3.46我们有

$$\text{rank}(\mathbf{B}) \leq s - \text{rank}(\mathbf{A}),$$

即 $\text{rank}(\mathbf{B}) + \text{rank}(\mathbf{A}) \leq s$. \square

Lemma 4.58. 给定矩阵 $\mathbf{A} = (a_{ij})_{m \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$, 则

$$\text{rank} \left(\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \right) = \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$$

Lemma 4.59. 给定矩阵 $\mathbf{A} = (a_{ij})_{m \times n}$, $\mathbf{B} = (b_{ij})_{m \times s}$, 则

$$\max(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})) \leq \text{rank}((\mathbf{A}, \mathbf{B})) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}).$$

Lemma 4.60. 给定矩阵 $\mathbf{A} = (a_{ij})_{m \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$, 则

$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \leq \begin{pmatrix} \mathbf{A} & * \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \leq \min(\text{rank}(\mathbf{A}) + n, \text{rank}(\mathbf{B}) + m).$$

证明. 由4.58和4.59可知一个不等号. 将中间这个矩阵可以分别看成两个分块矩阵, 即

$$\begin{pmatrix} \mathbf{A} & * \\ \mathbf{0} & \mathbf{B} \end{pmatrix} = \left(\mathbf{A} \mid * \right) = \left(\frac{\mathbf{A} \quad *}{\mathbf{0} \quad \mathbf{B}} \right)$$

\square

Theorem 4.61. **Sylvester inequality** 给定矩阵 $\mathbf{A} = (a_{ij})_{m \times s}$, $\mathbf{B} = (b_{ij})_{s \times n}$, 则

$$\text{rank}(\mathbf{AB}) \geq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) - s.$$

这是一个非常好的乘积矩阵的秩的下界.

证明. 下面的证明的过程完美的利用了分块矩阵的性质. 设

$$\text{rank}(\mathbf{AB}) + s = \text{rank} \left(\begin{pmatrix} \mathbf{I}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{AB} \end{pmatrix} \right)$$

我们可以分块矩阵的初等变换, 使得

$$\begin{pmatrix} \mathbf{I}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{AB} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{I}_s & \mathbf{0} \\ \mathbf{A} & \mathbf{AB} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{I}_s & -\mathbf{B} \\ \mathbf{A} & \mathbf{0} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{B} & \mathbf{I}_s \\ \mathbf{0} & \mathbf{A} \end{pmatrix}$$

根据4.60就有

$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \leq \text{rank} \left(\begin{pmatrix} \mathbf{B} & \mathbf{I}_s \\ \mathbf{0} & \mathbf{A} \end{pmatrix} \right) = \text{rank}(\mathbf{AB}) + s.$$

□

矩阵相似及二次型

内积和正交定义

Definition 5.1. 在 \mathbb{R}^n 中, 任给 $\alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n)$, 规定

$$\langle \alpha, \beta \rangle = a_1 b_1 + a_2 b_2 + \dots + a_n b_n,$$

$\langle \alpha, \beta \rangle$ 是 α 和 β 的**内积**(实际上就是一个二元的线性泛函). 通常可以写作

$$\langle \alpha, \beta \rangle = \alpha \beta^T.$$

Proposition 5.2. 内积的基本性质如下

1. 对称性 $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$;
2. 线性可加 $\langle \alpha + \gamma, \beta \rangle = \langle \alpha, \beta \rangle + \langle \gamma, \beta \rangle$;
3. 线性齐次 $\langle \lambda \alpha, \beta \rangle = \lambda \langle \alpha, \beta \rangle$;
4. 正定性 $\langle \alpha, \alpha \rangle \geq 0$, 等号成立当且仅当 $\alpha = 0$.

Definition 5.3. 向量空间 \mathbb{R}^n 有了内积之后, 就称 \mathbb{R}^n 为一个**欧几里得空间**(Euclidean space).

Definition 5.4. 在欧几里得空间 \mathbb{R}^n 中, 向量 α 的长度 $\|\alpha\|$ 规定为

$$\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle}.$$

长度为 1 的向量称为**单位向量**.

Lemma 5.5.

$$\|\lambda \alpha\| = |\lambda| \|\alpha\|.$$

Definition 5.6. 把非零向量 α 乘上 $\frac{1}{\|\alpha\|}$ 称为把 α **单位化**.

Definition 5.7. 在欧几里得空间 \mathbb{R}^n 中, 两个非零向量 α 和 β 的夹角规定为

$$\frac{\langle \alpha, \beta \rangle}{\|\alpha\| \|\beta\|}.$$

Definition 5.8. 在欧几里得空间 \mathbb{R}^n 中, 如果 $\langle \alpha, \beta \rangle = 0$, 那么称 α 和 β 是**正交**, 记做 $\alpha \perp \beta$. 显然零向量与任何向量都正交.

Definition 5.9. 在欧几里得空间 \mathbb{R}^n 中, 由非零向量组成的向量组如果其中每两个不同的向量都正交, 那么称它们为**正交向量组**. 如果正交向量组的每个向量都是单位向量, 那么称它为**正交单位向量组**.

Proposition 5.10. 在欧几里得空间 \mathbb{R}^n 中, 正交向量组一定是线性无关的.

证明. 设 $\alpha_1, \alpha_2, \dots, \alpha_t$ 是正交向量组. 设有 k_1, k_2, \dots, k_t 使得

$$k_1\alpha_1 + k_2\alpha_2 + \dots + k_t\alpha_t = \mathbf{0},$$

等式两边同时分别取关于 α_i ($1 \leq i \leq t$) 的内积, 当 $i \neq j$ 均有 $\langle \alpha_i, \alpha_j \rangle = 0$, 于是 $k_i \langle \alpha_i, \alpha_i \rangle = 0$, 即 $k_i = 0, 1 \leq i \leq t$. \square

Definition 5.11. 在欧几里得空间 \mathbb{R}^n 中, n 个向量组成的正交向量组一定是 \mathbb{R}^n 的一个基, 称它为**正交基**. n 个单位向量组成的正交向量组称为 \mathbb{R}^n 的一个**标准正交基**.

Theorem 5.12. 施密特 (Schmit) 正交化 设 $\alpha_1, \alpha_2, \dots, \alpha_t$ 是欧几里得空间 \mathbb{R}^n 中的一个线性无关的向量组, 令

$$\begin{aligned} \beta_1 &= \alpha_1 \\ \beta_2 &= \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 \\ &\dots \\ \beta_t &= \alpha_t - \sum_{i=1}^{t-1} \frac{\langle \alpha_t, \beta_i \rangle}{\langle \beta_i, \beta_i \rangle} \beta_i \end{aligned},$$

则 $\beta_1, \beta_2, \dots, \beta_t$ 是正交向量组, 并且 $\beta_1, \beta_2, \dots, \beta_t$ 与 $\alpha_1, \alpha_2, \dots, \alpha_t$ 等价. 自然地, $\frac{\beta_1}{\|\beta_1\|}, \frac{\beta_2}{\|\beta_2\|}, \dots, \frac{\beta_t}{\|\beta_t\|}$ 是正交单位向量组. 其实就是在 α_2 上减去它在 α_1 投影方向的分向量, 余类推.

证明. 对所给定的线性无关的向量组所含向量个数进行归纳. 当 $s = 1$ 时, $\beta_1 = \alpha_1$ 是正交向量组, 且 β_1 和 α_1 等价. 假设在 $s = k$ 时成立, 那么当 $s = k + 1$ 时, 有

$$\beta_{k+1} = \alpha_{k+1} - \sum_{i=1}^k \frac{\langle \alpha_{k+1}, \beta_i \rangle}{\langle \beta_i, \beta_i \rangle} \beta_i, \quad (1)$$

因此当 $1 \leq j \leq k$ 时, 有

$$\langle \beta_{k+1}, \beta_j \rangle = \langle \alpha_{k+1}, \beta_j \rangle - \sum_{i=1}^k \frac{\langle \alpha_{k+1}, \beta_i \rangle}{\langle \beta_i, \beta_i \rangle} \langle \beta_i, \beta_j \rangle = \langle \alpha_{k+1}, \beta_j \rangle - \langle \alpha_{k+1}, \beta_j \rangle = 0.$$

所以 β_{k+1} 和 β_j 均正交. 观察 (1) 式可知 β_{k+1} 是由 $\alpha_1, \alpha_2, \dots, \alpha_t$ 线性表出的, 并且 α_k 的系数为 1, 所以 β_{k+1} 不为 0. 于是 $\beta_1, \dots, \beta_k, \beta_{k+1}$ 为正交向量组. 还是由 (1) 式 $\beta_1, \dots, \beta_k, \beta_{k+1}$ 和 $\alpha_1, \dots, \alpha_k, \alpha_{k+1}$ 显然是等价. \square

正交矩阵

Annotation 5.13. 设平面上两个相互垂直的单位向量 $\alpha = (a_1, a_2), \beta = (b_1, b_2)$, 它们满足

$$a_1^2 + a_2^2 = 1, b_1^2 + b_2^2 = 1, a_1b_1 + a_2b_2 = 0.$$

正好就是它们为行向量构成的矩阵和转置的乘积为单位矩阵, 即

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Definition 5.14. 如果 n 阶矩阵 A 满足

$$AA^T = I,$$

则称 A 是正交矩阵.

Proposition 5.15. 下述性质是相互等价的 (互为充要条件).

1. n 阶矩阵 A 是正交矩阵;
2. $AA^T = I$;
3. A 可逆, 且 $A^{-1} = A^T$;
4. $A^T A = I$.

证明. 从正交矩阵的定义, 由 proposition 4.34 直接可知 A 和 A^T 互为逆矩阵, 因此上面的性质都很显然. 特别地, 设 A 的行向量为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 那么按照正交矩阵的定义

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \begin{pmatrix} \alpha_1^T & \alpha_2^T & \cdots & \alpha_n^T \end{pmatrix} = I.$$

它给定的关系式就是

$$\alpha_i \alpha_j^T = \begin{cases} 1 & i = j, \\ 0 & i \neq j \end{cases}$$

□

Proposition 5.16. 设 $\eta_1, \eta_2, \dots, \eta_n$ 是欧几里得空间 \mathbb{R}^n 的一个标准正交基, 则以它们为行向量的矩阵

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

是一个正交矩阵.

Proposition 5.17. 正交矩阵具有下列性质

1. I 是正交矩阵;
2. 若 A 和 B 都是 n 阶正交矩阵, 则 AB 也是正交矩阵;
3. 若 A 是正交矩阵, 则 A^{-1} 也是正交矩阵;
4. 若 A 是正交矩阵, 则 $|A| = 1$ 或者 -1 .

证明.

□

(1)(3) 是显然的. 若 A 是正交矩阵, 那么

$$|AA^T| = |A||A^T| = |A|^2 = |I|,$$

所以 $|A| = 1$ 或者 $A = -1$, 即 (4) 得证. 若 A 和 B 都是 n 阶矩阵, 那么

$$AB(AB)^T = ABB^T A^T = I,$$

即 (2) 得证.

正交变换

矩阵的特征值和特征向量

Definition 5.18. 设 A 是一个 n 阶矩阵, 如果数 λ 和 n 维非零列向量 α 使得

$$A\alpha = \lambda\alpha \tag{1}$$

成立. 那么称 λ 为矩阵 A 的**特征值**, 非零向量 α 称为矩阵 A 的对应于特征值 λ 的一个**特征向量**. (1) 式可以改写为

$$(A - \lambda I)\alpha = 0, \tag{2}$$

这个等式可以看做一个 n 个未知数的齐次线性方程组 $(\mathbf{A} - \lambda \mathbf{I})\mathbf{X} = 0$, 而 α 是这个齐次线性方程的一个非零解, 这个齐次线性方程被称为 \mathbf{A} 的**特征方程**. 把 $|\mathbf{A} - \lambda \mathbf{I}|$ 称为 \mathbf{A} 的**特征多项式**, $|\mathbf{A} - \lambda \mathbf{I}|$ 写作

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

Lemma 5.19. 若 α 是 \mathbf{A} 的对应于特征值 λ 的一个特征向量, 那么 $k\alpha$ 也是 \mathbf{A} 的对应于特征值 λ 的一个特征向量.

证明.

$$\mathbf{A}(k\alpha) = k(\mathbf{A}\alpha) = k(\lambda\alpha) = \lambda(k\alpha).$$

□

Proposition 5.20. \mathbf{A} 的特征值 λ 就是特征多项式的**根**.

证明. (2) 式要有非零解的充要条件是 $|\mathbf{A} - \lambda \mathbf{I}| = 0$, 而 $|\mathbf{A} - \lambda \mathbf{I}|$ 是关于 λ 的一个多项式, 因此所求 λ 就是特征多项式的根. □

Proposition 5.21. \mathbf{A} 的特征向量 α 是齐次线性方程 $(\mathbf{A} - \lambda \mathbf{I})\mathbf{X} = 0$ 的一个非零解.

Annotation 5.22. 求 \mathbf{A} 的全部特征值和特征向量的方法如下

1. 计算 \mathbf{A} 的特征多项式 $|\mathbf{A} - \lambda \mathbf{I}|$,
2. 如果多项式 $|\mathbf{A} - \lambda \mathbf{I}|$ 在数域 F (实数域和复数域) 中没有根, 那么 \mathbf{A} 没有特征值. 如果多项式 $|\mathbf{A} - \lambda \mathbf{I}|$ 在数域 F 中有根, 那么它在 F 中的全部根就是 \mathbf{A} 的特征值, 此时再做第 3 步.
3. 对 \mathbf{A} 的每一个特征值 λ_i , 求齐次线性方程组 $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{X} = 0$ 的一个基础解系 $\eta_1, \eta_2, \dots, \eta_t$. 于是 \mathbf{A} 的对应于特征值 λ_i 的全部特征向量组成的集合是

$$\{k_1\eta_1 + k_2\eta_2 + \cdots + k_t\eta_t \mid k_1, k_2, \dots, k_t \in F, \text{ 且它们不全为 } 0\}.$$

Definition 5.23. 设 λ 是 n 阶 \mathbf{A} 的一个特征值, 把齐次方程组 $(\mathbf{A} - \lambda \mathbf{I})\mathbf{X} = 0$ 的解空间称为 \mathbf{A} 的对应于 λ 的**特征子空间**.

Proposition 5.24. 设 λ 是 n 阶 \mathbf{A} 的一个特征值, 则下述命题成立

1. λ^m 是 \mathbf{A}^m 的特征值;

2. 当 \mathbf{A} 可逆时, 则 $\frac{1}{\lambda}$ 是 \mathbf{A}^{-1} 的特征值.
3. \mathbf{A} 和 \mathbf{A}^T 的特征值相同;
4. \mathbf{A} 有 n 个特征值 (包含重根, 也可能是复根);
5. \mathbf{A} 的所有特征值之和等于 \mathbf{A} 的对角线之和;
6. \mathbf{A} 的所有特征值之积等于 \mathbf{A} 的行列式;

证明. (1)

$$\mathbf{A}^m \boldsymbol{\alpha} = \mathbf{A}^{m-1} \mathbf{A} \boldsymbol{\alpha} = \lambda \mathbf{A}^{m-1} \boldsymbol{\alpha} = \cdots = \lambda^m \boldsymbol{\alpha}.$$

(2)

$$\mathbf{A} \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha},$$

两边左乘 \mathbf{A}^{-1} , 有

$$\boldsymbol{\alpha} = \lambda \mathbf{A} \boldsymbol{\alpha}.$$

故 $\mathbf{A} \boldsymbol{\alpha} = \frac{1}{\lambda} \boldsymbol{\alpha}$.

(3)

$$|\mathbf{A}^T - \lambda I| = |(\mathbf{A}^T - \lambda I)^T| = |\mathbf{A} - \lambda I|.$$

(5) 记 \mathbf{A} 的特征多项式

$$p(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n) = \lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_0.$$

考虑其展开式中 λ^{n-1} 项

$$-\lambda_1 \lambda^{n-1} - \cdots - \lambda_n \lambda^{n-1} = -(\lambda_1 + \cdots + \lambda_n) \lambda^{n-1}.$$

因此 $c_{n-1} = -(\lambda_1 + \cdots + \lambda_n)$. 再从 $|\lambda \mathbf{I} - \mathbf{A}|$ 的行列式来考虑 c_{n-1} , 即

$$\begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{21} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}$$

按行列式完全展开式, 只有 $(\lambda - a_{11}) \cdots (\lambda - a_{nn})$ 包含 λ^{n-1} 项, 其他项最高次只有 λ^{n-2} 项, 于是 $p(\lambda)$ 可以改写为

$$p(\lambda) = (\lambda - a_{11}) \cdots (\lambda - a_{nn}) + q(\lambda).$$

其中 $q(\lambda)$ 最高次项为 λ^{n-2} , 因此 $c_{n-1} = -(a_{11} + \cdots + a_{nn})$.

(6) 因 $p(0) = (-1)^n \prod_{i=1}^n \lambda_i$, 而 $p(0) = |0 \cdot \mathbf{I} - \mathbf{A}| = (-1)^n |\mathbf{A}|$.

□

Proposition 5.25. 设 A 是 n 阶可逆矩阵, A^* 是 A 的伴随矩阵, λ 是 A 的一个特征值, 则 $\frac{|A|}{\lambda}$ 是 A^* 的特征值.

证明.

$$\begin{aligned} A^* \alpha &= \lambda' \alpha \\ |A| A^{-1} \alpha &= \lambda' \alpha \\ |A| \alpha &= \lambda' A \alpha \\ \frac{|A|}{\lambda'} \alpha &= A \alpha \end{aligned}$$

因此 $\frac{|A|}{\lambda'} = \lambda$, 即 $\lambda' = \frac{|A|}{\lambda}$. □

Definition 5.26. 设 A 是一个 n 阶行列式, λ_1 是 A 的一个特征值. 把 A 的对应于 λ_1 的特征子空间的维数叫做特征值 λ_1 的几何重数. 而把 λ_1 作为 A 的特征多项式的根的重数叫做 λ_1 的代数重数, 代数重数简称重数.

Proposition 5.27. 设 λ_1 是 n 阶矩阵 A 的一个特征值, 则 λ_1 的几何重数不超过它的代数重数.

证明. 证明思路就是从多项式 $|A - \lambda I|$ 里面提个 $(\lambda - \lambda_1)^r$ 出来, 其中 r 就是 λ_1 的几何重数, 因此 λ_1 的代数重数是大于等于 r 的. 详细证明丘高代 p272. □

Proposition 5.28. 设 $\lambda_1, \lambda_2, \dots, \lambda_m$ 是 n 阶矩阵 A 的特征值, $\alpha_1, \alpha_2, \dots, \alpha_m$ 分别是与之对应的特征向量. 如果 $\lambda_1, \lambda_2, \dots, \lambda_m$ 均互不相等, 则 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性无关.

证明. 对 m 进行归纳证明. 当 $m = 1$ 时, α_1 线性无关. 假设在 $m = k$ 时成立. 那么当 $m = k + 1$, 我们考察等式

$$b_1 \alpha_1 + b_2 \alpha_2 + \dots + b_k \alpha_k + b_{k+1} \alpha_{k+1} = \mathbf{0}, \quad (3)$$

上式两边左乘 A , 有

$$b_1 A \alpha_1 + b_2 A \alpha_2 + \dots + b_k A \alpha_k + b_{k+1} A \alpha_{k+1} = \mathbf{0},$$

即

$$b_1 \lambda_1 \alpha_1 + b_2 \lambda_2 \alpha_2 + \dots + b_k \lambda_k \alpha_k + b_{k+1} \lambda_{k+1} \alpha_{k+1} = \mathbf{0}. \quad (4)$$

将 (3) 式乘上 λ_{k+1} 减去 (4) 式, 有

$$b_1 (\lambda_{k+1} - \lambda_1) \alpha_1 + b_2 (\lambda_{k+1} - \lambda_2) \alpha_2 + \dots + b_k (\lambda_{k+1} - \lambda_k) \alpha_k = \mathbf{0}$$

根据假设 $n = k$ 时命题成立, 那么

$$b_1 (\lambda_{k+1} - \lambda_1) = b_2 (\lambda_{k+1} - \lambda_2) = \dots = b_k (\lambda_{k+1} - \lambda_k).$$

其中 $(\lambda_{k+1} - \lambda_i)$ 均不为零 ($1 \leq i \leq k$), 因此 $b_1 = b_2 = \dots = b_k = 0$, 带入 (3) 式, 因 $\alpha_{k+1} \neq \mathbf{0}$, 所以 $b_{k+1} = 0$. 即 $n = k + 1$ 时命题也成立. □

Corollary 5.29. 设 λ_1 和 λ_2 是 n 阶矩阵 \mathbf{A} 的两个不同的特征值, 那么分别对应于 λ_1 及 λ_2 下的两个线性无关特征向量组 $\xi_1, \xi_2, \dots, \xi_s$ 和 $\eta_1, \eta_2, \dots, \eta_t$ 合在一起也是线性无关的.

证明. 由 proposition 5.28, 可以清晰的感觉到不同特征值的特征子空间是 distinct 的, 那么当前这个命题就很显然了. □

等价矩阵

Definition 5.30. 对于数域 F 上的两个 $m \times n$ 矩阵 A 和 B , 若从 A 经过一系列初等行变换和初等列变换能成 B , 那么称 A 与 B 是等价的 (相抵的), 记为 $A \simeq B$.

Annotation 5.31. 矩阵等价是一个等价关系.

Theorem 5.32. 设 $m \times n$ 矩阵 A 的秩为 r . 若 $r > 0$, 那么 A 等价于下述形式的矩阵

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

称上述矩阵是 A 的等价标准形; 若 $r = 0$, 则 A 等于于零矩阵.

Theorem 5.33. 设两个 $m \times n$ 矩阵 A 和 B , 若它们等价当且仅当它们的秩相同.

相似矩阵

Annotation 5.34. 为了求 n 阶矩阵的 n 次幂 A^m , 如果能找到可逆矩阵 P , 使得 $PAP^{-1} = D$, 那么

$$A^m = (PAP^{-1})(PAP^{-1}) \cdots (PAP^{-1}) = PD^mP^{-1}.$$

Definition 5.35. 设 A 和 B 都是 n 阶矩阵, 如果存在一个 n 阶可逆矩阵 P 使得

$$P^{-1}AP = B,$$

那么称 A 和 B 是相似的, 记做 $A \sim B$.

Annotation 5.36. 矩阵相似是一种等价关系, 并且矩阵等价的特例.

Proposition 5.37. 重要的一个必要条件 相似矩阵有相等的特征多项式, 从而特征值亦相同.

证明. 设 $A \sim B$, 有

$$|B - \lambda I| = |P^{-1}AP - \lambda P^{-1}P| = |P^{-1}(A - \lambda I)P| = |A - \lambda I|.$$

□

Proposition 5.38. 若 $A \sim B$, 则有以下性质

1. $A^n \sim B^n$;
2. $A + kI \sim B + kI$;
3. 若 A 可逆, 则 $A^{-1} \sim B^{-1}$;

证明. (1)

$$(P^{-1}AP) \cdots (P^{-1}AP) = P^{-1}A^nP = B^n.$$

(2)

$$P^{-1}(A + kI)P = P^{-1}AP + kP^{-1}IP = B + kI.$$

(3) 由于 A 可逆, 所以 B 可逆, 那么

$$B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}P.$$

□

Annotation 5.39. 可以通过矩阵相似来求秩!

Definition 5.40. 若 n 阶矩阵 A 相似于一个对角矩阵 D , 那么称 A 可对角化, D 是 A 的标准相似性型.

Theorem 5.41. n 阶矩阵 A 可对角化的充分条件是存在 n 个线性无关的列向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 以及 n 个数 $\lambda_1, \lambda_2, \dots, \lambda_n$ (它们之中有些可能相等), 使得

$$A\alpha_i = \lambda_i\alpha_i, i = 1, 2, \dots, n.$$

这时, 令 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 则

$$P^{-1}AP = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}.$$

证明. 充分性 设存在 n 个线性无关的列向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 以及 n 个数 $\lambda_1, \lambda_2, \dots, \lambda_n$, 使得

$$A\alpha_i = \lambda_i\alpha_i, i = 1, 2, \dots, n.$$

于是有

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n),$$

记 $P = (\alpha_1, \alpha_2, \dots, \alpha_n), D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. 因此有

$$AP = PD,$$

由于 P 是满秩矩阵, 所以 P 可逆, 那么即有 $P^{-1}AP = D$.

必要性 若 $A \sim D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 因此存在可逆矩阵 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 使得

$$P^{-1}AP = D$$

$$AP = PD$$

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n).$$

即有对应列向量有 $A\alpha_i = \lambda_i\alpha_i, i = 1, 2, \dots, n$. □

Corollary 5.42. 若 n 阶矩阵 A 与对角矩阵 $D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 相似, 那么 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值.

Corollary 5.43. 若 n 阶矩阵 A 的 n 个特征值互不相等, 则 A 与对角矩阵相似.

矩阵对角化的条件

Theorem 5.44. n 阶矩阵 A 可对角化的充要条件是 A 有 n 个线性无关的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_n$. 令 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 则

$$P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

其中 λ_i 是 α_i 对应的特征值 ($i = 1, 2, \dots, n$). 注意矩阵 P 的构造中的列向量的位置要和对角矩阵的对角线元素对应.

Annotation 5.45. 如何判定 n 阶矩阵 A 有没有 n 个线性无关的特征向量?

1. 求出 A 的全部特征值, 设其为 $\lambda_1, \lambda_2, \dots, \lambda_m$.
2. 对于每一个特征值 λ_i , 求出齐次线性方程组 $(A - \lambda_i I)X = 0$ 的一个基础解系 $\eta_{i1}, \eta_{i2}, \dots, \eta_{ir_i}$.
3. 由 corollary 5.29, 这 m 组线性无关向量合在一起还是线性无关的. 从而如果 $r_1 + r_2 + \dots + r_m = n$, 那么 A 有 n 个线性无关的特征向量, 即 A 可对角化. 反之如果 $r_1 + r_2 + \dots + r_m < n$, 那么 A 没有 n 个线性无关的特征向量, 即 A 不可对角化.

Theorem 5.46. n 阶矩阵 A 可对角化的充分必要条件是 A 的对应不同特征值的特征子空间的维数之和等于 n .

证明. 从前面的 annotation 即可得到. □

Corollary 5.47. 若 n 阶矩阵 A 有 n 个不同的特征值, 那么 A 可对角化.

Theorem 5.48. 数域 F 上的 n 阶矩阵 A 可对角化的充要条件是 A 的特征多项式的全部复根都属于 F , 并且 A 的每个特征值的几何重数等于它的代数重数.

实对称矩阵的对角化

Definition 5.49. 如果对于 n 级实矩阵 A, B , 存在一个 n 级正交矩阵 T , 使得 $T^{-1}AT = B$, 那么称 A 正交相似于 B .

Annotation 5.50. 正交相似是一个等价关系.

Theorem 5.51. n 阶实对称矩阵的特征多项式的每一个复根都是实数, 从而它们都是特征值. (代数基本定理告诉我们 n 次多项式在复数域上都 n 个复根, 所以这里是相对于代数基本定理来说的)

证明. 设 λ 是 n 阶实对称矩阵的特征多项式的任意一个复根, 要证 λ 是一个实数, 则需要证明 $\bar{\lambda} = \lambda$. 首先把 A 看做一个复矩阵, 则 λ 是 A 的一个特征值, 从而存在对应的非零特征向量 $\alpha \in \mathbb{C}^n$, 使得

$$A\alpha = \lambda\alpha. \quad (1)$$

我们定义 \bar{A} 表示对 A 里面的所有元素取共轭复数, 然后对 (1) 式两边同时取共轭复数

$$\begin{aligned} \overline{A\alpha} &= \overline{\lambda\alpha} \\ \bar{A}\bar{\alpha} &= \bar{\lambda}\bar{\alpha} \end{aligned}$$

由于 A 是实对称矩阵, 因此可以得到 $A\bar{\alpha} = \bar{\lambda}\bar{\alpha}$. 对 (1) 式两边取转置, 有

$$\begin{aligned} (A\alpha)^T &= (\lambda\alpha)^T \\ \alpha^T A^T &= \lambda\alpha^T, \end{aligned}$$

由于 A 是对称矩阵, 因此可以得到 $\alpha^T A = \lambda\alpha^T$. 对得到的一个等式两边左乘 α^T 减去第二个等式两边右侧 $\bar{\alpha}$, 可以得到

$$(\bar{\lambda} - \lambda)\alpha^T \bar{\alpha} = 0.$$

由于 $\alpha \neq 0$, 所以 $\alpha^T \bar{\alpha} \neq 0$ (互为共轭的复数的乘积非负), 从而 $\bar{\lambda} = \lambda$. □

Corollary 5.52. 实对称矩阵一定有特征值.

证明. 配合代数基本定理和 theorem 5.51 可知实对称矩阵一定有实数根, 尽管可能是重根. □

Theorem 5.53. 实对称矩阵 A 的对应不同的特征值的特征向量是正交.

证明. 设 λ_1, λ_2 的 A 的两个不同特征值, A 的分别与之对应的两个特征向量为 α_1, α_2 . 要证明 $\langle \alpha_1, \alpha_2 \rangle = 0$, 为此计算

$$\begin{aligned} \lambda_1 \langle \alpha_1, \alpha_2 \rangle &= \langle \lambda_1 \alpha_1, \alpha_2 \rangle = (A\alpha_1)^T \alpha_2 = \alpha_1^T A\alpha_2, \\ \lambda_2 \langle \alpha_1, \alpha_2 \rangle &= \langle \alpha_1, \lambda_2 \alpha_2 \rangle = \alpha_1^T (A\alpha_2) = \alpha_1^T A\alpha_2, \end{aligned}$$

上述两式相减, 得到 $(\lambda_1 - \lambda_2)\langle \alpha_1, \alpha_2 \rangle = 0$. 由于 $\lambda_1 \neq \lambda_2$, 因此 $\langle \alpha_1, \alpha_2 \rangle = 0$. □

Theorem 5.54. 实对称矩阵一定正交相似于对角矩阵.

证明. 对实对称矩阵的阶数 n 进行归纳. 当 $n = 1$ 时, $(1)^{-1}(a)(1) = (a)$, 因此命题成立. 假设对于 $n - 1$ 阶实对称矩阵命题成立 ($n = 2, 3, \dots$), 我们来考察 n 阶实对称矩阵 \mathbf{A} . 由 corollary 5.52 可知 \mathbf{A} 一定存在一个特征值 λ_1 , 那么存在 \mathbf{A} 的对应 λ_1 的一个非零特征向量 α_1 . 我们将 α_1 单位化得到 η_1 , 我们将 η_1 扩充称 \mathbb{R}^n 的一个基, 然后经过施密特正交化和单位化, 得到 \mathbb{R}^n 的一个标准正交基 $\eta_1, \eta_2, \dots, \eta_n$. 由 proposition 5.16 可知 $T_1 = (\eta_1, \eta_2, \dots, \eta_n)$ 是一个 n 级正交矩阵. 于是有

$$T_1^{-1}\mathbf{A}T_1 = T_1^{-1}(\mathbf{A}\eta_1, \mathbf{A}\eta_2, \dots, \mathbf{A}\eta_n) = (T_1^{-1}\lambda_1\eta_1, T_1^{-1}\mathbf{A}\eta_2, \dots, T_1^{-1}\mathbf{A}\eta_n). \quad (2)$$

由于 $T_1^{-1}T_1 = (T_1^{-1}\eta_1, T_1^{-1}\eta_2, \dots, T_1^{-1}\eta_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, 因此 $T_1^{-1}\eta_1 = \varepsilon_1$, 带入 (2) 式. 从而得到

$$T_1^{-1}\mathbf{A}T_1 = \begin{pmatrix} \lambda_1 & \beta \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$$

其中 $\mathbf{0} \in \mathbb{R}^{(n-1),1}$, $\beta \in \mathbb{R}^{1,(n-1)}$, 自然地 \mathbf{B} 是一个 $n - 1$ 阶矩阵. 首先我们考察一下 $(T_1^{-1}\mathbf{A}T_1)^T$, 有

$$(T_1^{-1}\mathbf{A}T_1)^T = (\mathbf{A}T_1)^T(T_1^{-1})^T = T_1^T\mathbf{A}^T(T_1^{-1})^T = T_1^{-1}\mathbf{A}T_1.$$

因此 $T_1^{-1}\mathbf{A}T_1$ 也是一个实对称矩阵, 那么有 $\beta = \mathbf{0}$, \mathbf{B} 也是一个实对称矩阵. 自然地, 我们希望把 $(T_1^{-1}\mathbf{A}T_1)^T$ 中 \mathbf{B} 换成 $T_2^{-1}\mathbf{B}T_2 = \text{diag}\{\lambda_2, \lambda_3, \dots, \lambda_n\}$, 来用假设中 $n - 1$ 阶实对称矩阵的结论. 这里很容易联想到分块矩阵的初等变换, 即

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2 \end{pmatrix}^{-1} T_1^{-1}\mathbf{A}T_1 \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2 \end{pmatrix},$$

所以我们需要重新构造一个新的正交矩阵 T , 令

$$T = T_1 \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2 \end{pmatrix}.$$

显然 T 是一个正交矩阵, 即对 n 阶实对称矩阵, 命题也成立. □

Annotation 5.55. 对于 n 阶实对称矩阵 \mathbf{A} , 如何找到一个正交矩阵 T 使得 $T^{-1}\mathbf{A}T$ 为一个对角矩阵? 步骤如下

1. 计算特征多项式 $|\mathbf{A} - \lambda\mathbf{I}|$, 求出它的全部不同的根 $\lambda_1, \lambda_2, \dots, \lambda_m$, 它们是 \mathbf{A} 全部的特征值.
2. 对于每一个特征值 $\lambda_i (i = 1, 2, \dots, m)$, 求 $(\mathbf{A} - \lambda_i\mathbf{I})X = 0$ 一个基础解系, 并对这个基础解系进行施密特正交化和单位化, 得到 $\eta_{i1}, \eta_{i2}, \dots, \eta_{ir_i}$. 它们都是 \mathbf{A} 的对应于 λ_i 的特征向量.

3. 由 theorem 5.54, 可知 $r_1 + r_2 + \cdots + r_m = n$, 所以可对角化. 令

$$T = (\eta_{11}, \cdots, \eta_{1r_1}, \cdots, \eta_{m1}, \cdots, \eta_{mr_m}),$$

由 theorem 5.53 可知 T 是 n 阶正交矩阵, 且

$$T^{-1}AT = \text{diag}\{\underbrace{\lambda_1, \cdots, \lambda_1}_{r_1}, \cdots, \underbrace{\lambda_m, \cdots, \lambda_m}_{r_m}\}.$$

Theorem 5.56. 如果 n 阶实矩阵 A 正交相似于一个对角矩阵, 那么 A 一定是对称矩阵.

证明. 若 $T^{-1}AT = D$, 那么

$$(A)^T = (TDT^{-1})^T = (T^{-1})^T D T^T = T D T^{-1}.$$

多次用正交矩阵的逆就是它的转置. □

Theorem 5.57. 两个 n 阶实对称矩阵正交相似的充分必要条件是它们相似.

证明. 正交相似显然是相似的. 由 5.54 可知 n 阶实对称矩阵 A 正交相似于对角矩阵 $\text{diag}\{\lambda_1, \cdots, \lambda_n\}$, 其中 $\lambda_1, \cdots, \lambda_n$ 都是 A 的特征值, 再由 5.37 可知两个相似矩阵 A, B 的特征值相同的, 因此与之正交相似的两个对角矩阵是相同的, 根据正交相似的传递性, A 和 B 正交相似. □

二次型

二次型及其标准型

Definition 6.1. 数域 F 上的一个 n 元二次型是系数在 F 中 n 个变量的二次齐次多项式，它的一般形式是

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \dots + 2a_{1n}x_1x_n \\ & + a_{22}x_2^2 + 2a_{23}x_2x_3 + \dots + 2a_{2n}x_2x_n \\ & + \dots \quad \dots \quad \dots \quad \dots \\ & + a_{nn}x_n^2 \end{aligned} \quad (1)$$

(1) 式也可以写成

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j, \quad (2)$$

其中当 $j > i$ 时, $a_{ji} = a_{ij}$.

Definition 6.2. 利用矩阵 (2) 式可表示为

$$\begin{aligned} f(x) = & x_1(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) + x_2(a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) \\ & + \dots + x_n(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) \end{aligned}$$

$$= (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix},$$

$$= (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

记 $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$, 则二次型可用矩阵记做

$$f(x_1, x_2, \dots, x_n) = \mathbf{X}^T \mathbf{A} \mathbf{X}.$$

此时, 称 \mathbf{A} 是二次型 $f(x_1, x_2, \dots, x_n)$ 的矩阵, 它是对称矩阵. 显然 \mathbf{A} 是唯一的: 它的主对角线元素依次是 $x_1^2, x_2^2, \dots, x_n^2$ 的系数; 它的 (i, j) 元素是 $x_i x_j$ 的系数的一半, 其中 $i < j$.

Definition 6.3. 令 $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, 设 C 是数域 F 上的 n 级可逆矩阵, 则关系式

$$\mathbf{X} = \mathbf{C}\mathbf{Y}$$

称为变量 x_1, x_2, \dots, x_n 到变量 y_1, y_2, \dots, y_n 的一个非退化线性替换.

Annotation 6.4. 将 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 经过非退化线性替换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ 变成

$$(\mathbf{C}\mathbf{Y})^T \mathbf{A} (\mathbf{C}\mathbf{Y}) = \mathbf{Y}^T \mathbf{C}^T \mathbf{A} \mathbf{C} \mathbf{Y}, \quad (3)$$

记 $\mathbf{B} = \mathbf{C}^T \mathbf{A} \mathbf{C}$, 那么 (3) 式就可以写作 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$. 这就是变量 y_1, y_2, \dots, y_n 的一个二次型. 因为 \mathbf{A} 是对称矩阵, 所以 \mathbf{B} 也是对称矩阵, 从而 \mathbf{B} 正好是二次型 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ 的矩阵.

Definition 6.5. 给定两个 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 和 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$, 若存在一个非退化线性替换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ 可以把 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 变换成 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$, 那么就称二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 和 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ 等价, 记为 $\mathbf{X}^T \mathbf{A} \mathbf{X} \cong \mathbf{Y}^T \mathbf{B} \mathbf{Y}$.

Definition 6.6. 上述二次型等价中的把 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 变换成 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ 中的变成如何理解? 也就是指上面 annotation 的过程, 即存在一个非退化线性替换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ 使得 $\mathbf{B} = \mathbf{C}^T \mathbf{A} \mathbf{C}$.

Definition 6.7. 给定两个 n 阶矩阵 \mathbf{A} 与 \mathbf{B} , 若存在一个可逆 n 阶矩阵 \mathbf{C} 使得

$$\mathbf{C}^T \mathbf{A} \mathbf{C} = \mathbf{B},$$

那么就称 \mathbf{A} 与 \mathbf{B} 合同, 记为 $\mathbf{A} \simeq \mathbf{B}$. 如果对称矩阵 \mathbf{A} 合同于一个对角矩阵, 那么这个对角矩阵称为 \mathbf{A} 的一个合同标准型.

Annotation 6.8. 矩阵合同是一个等价关系.

Proposition 6.9. 若两个 n 阶矩阵 \mathbf{A} 与 \mathbf{B} 合同, 则满足下述简单性质

1. 若 \mathbf{A} 是对称矩阵, 那么 \mathbf{B} 也是对称矩阵.
2. $|\mathbf{A}| = |\mathbf{B}|$.

Proposition 6.10. 两个 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 和 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ 等价当且仅当 \mathbf{A} 和 \mathbf{B} 合同 (congruent).

Definition 6.11. 如果二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 等价于一个只含平方项的二次型 (其对应的矩阵就是一个对角矩阵), 那么这个只含平方项的二次型称为 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的一个**标准型**.

Proposition 6.12. 任给 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 都有一个标准型为

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 \mathbf{A} 的全部特征值. 如果 \mathbf{T} 是正交矩阵, 那么 $\mathbf{X} = \mathbf{T} \mathbf{Y}$ 称为**正交变换**.

证明. 由 theorem 5.54 可知存在 n 阶正交矩阵 \mathbf{T} , 使得

$$\mathbf{T}^T \mathbf{A} \mathbf{T} = \text{diag}(\lambda_1, \cdots, \lambda_n)$$

因此 \mathbf{A} 合同于 $\text{diag}(\lambda_1, \cdots, \lambda_n)$, 那么有非线性退化 $\mathbf{X} = \mathbf{T} \mathbf{Y}$ 使得 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 化成 $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$. \square

Annotation 6.13. 注意正交变换只能化二次型为标准二次型, 平方项的系数即特征值.

Annotation 6.14. **配方法** 将二次型 $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 - x_3^2 + 4x_1x_2 - 4x_1x_3 - 4x_2x_3$ 化为标准型

$$\begin{aligned} f(x_1, x_2, x_3) &= x_1^2 + 4x_1(x_2 - x_3) + [2(x_2 - x_3)]^2 - [2(x_2 - x_3)]^2 + 2x_2^2 - x_3^2 - 4x_2x_3 \\ &= (x_1 + 2x_2 - 2x_3)^2 - (4x_2^2 - 8x_2x_3 + 4x_3^2) + 2x_2^2 - x_3^2 - 4x_2x_3 \\ &= (x_1 + 2x_2 - 2x_3)^2 - 2x_2^2 + 4x_2x_3 - 5x_3^2 \\ &= (x_1 + 2x_2 - 2x_3)^2 - 2(x_2^2 - 2x_2x_3 + x_3^2) - 3x_3^2 \\ &= (x_1 + 2x_2 - 2x_3)^2 - 2(x_2 - x_3)^2 - 3x_3^2 \end{aligned}$$

令

$$\begin{cases} y_1 = x_1 + 2x_2 - 2x_3 \\ y_2 = x_2 - x_3 \\ y_3 = x_3 \end{cases}$$

则 $f(x_1, x_2, x_3) = y_1^2 - 2y_2^2 - 3y_3^2$, 其做的非退化线性替换为

$$\begin{cases} x_1 = y_1 - 2y_2 \\ x_2 = y_2 + y_3 \\ x_3 = y_3 \end{cases}$$

即

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Example 6.15. 将二次型 $f(x_1, x_2, x_3) = x_1x_2 + x_1x_3 - 3x_2x_3$ 化为标准二次型

Proposition 6.16. 三种初等变换的转置的性质:

$$\begin{aligned}P(j, i(k))^T &= P(i, j(k)) \\P(i, j)^T &= P(i, j) \\P(i(b))^T &= P(i(b))\end{aligned}$$

Definition 6.17. 设 A, B 是一个 n 阶矩阵, P 是任意的一个初等矩阵, 若

$$B = P^T A P,$$

那么称 B 是 A 通过一个成对初等行列变换得到的. 用自然语言描述成对初等行列变换三种类型的操作如下

1. $P(i, j(k))AP(j, i(k))$ 先将 A 第 j 行的 k 倍加到 i 行, 再将第 j 列的 k 倍加到 i 列.
2. $P(i, j)AP(i, j)$ 先将 A 的第 i, j 行交换, 再将第 i, j 列交换.
3. $P(i(b))AP(i(b))$ 先将 A 的第 i 行乘上非零数 b , 再将第 i 列乘上非零数 b .

Lemma 6.18. 设 A, B 都是 n 阶矩阵, 若 A 合同于 B 当且仅当 A 是经过一系列成对初等行列变换可以得到 B . 此时对 I 只做其中的列变换即可得到可逆矩阵 C , 使得 $C^T A C = B$.

证明. A 和 B 合同的充要条件是存在可逆矩阵 C , 使得 $C^T A C = B$. 而一个可逆矩阵 C 是可以被分解成多个初等矩阵的乘积, 即 $C = P_1 P_2 \cdots P_t$, 所以前式就可以写成

$$P_t^T \cdots P_2^T P_1^T A P_1 P_2 \cdots P_t = B.$$

□

Theorem 6.19. 任一 n 级对称矩阵一定合同于对角矩阵.

证明. 证明手法比较巧妙使用数学归纳法配合 lemma 6.18. 对对称矩阵的阶数进行归纳.

当 $n = 1$ 时, $(a) \simeq (a)$. 假设 $n - 1$ 阶对称矩阵命题成立 ($n = 2, 3, \cdots$), 那么现在来考察 n 阶矩阵的情况
A. 将 A 表示为

$$\begin{pmatrix} a_{11} & \alpha \\ \alpha^T & D_1 \end{pmatrix}.$$

情形 2 若 $a_{11} \neq 0$.

首先对其做一个成对初等行列操作, 将 α 和 α^T 变为 0 得到 A_2 , 即

$$A_2 = \begin{pmatrix} 1 & 0 \\ -a_{11}\alpha^T & I \end{pmatrix} \begin{pmatrix} a_{11} & \alpha \\ \alpha^T & D_1 \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}\alpha \\ 0 & I \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & D_1 - a_{11}\alpha^T\alpha \end{pmatrix}$$

记 $D_2 = D_1 - a_{11}\alpha^T\alpha$, 这里有 $A \simeq A_2$. 显然 D_2 也是对称矩阵. 因此这里就可以用假设中 $n-1$ 阶对称矩阵的结论, 即存在 $C_1^T A_2 C_1 = B_2$, 其中 B_2 是一个对角矩阵. 可以再通过一次成对初等行列变换把 $A_2 = \begin{pmatrix} a_{11} & 0 \\ 0 & D_2 \end{pmatrix}$ 的第二行左乘 C_1^T 和右乘 C_1 得到的矩阵为 A_3 , 即

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & C^T \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & C^T D_2 C \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & B_2 \end{pmatrix}.$$

因此 $A_2 \simeq A_3$. 所以根据合同等价关系的传递性有 $A \simeq A_3$, 即 n 阶对称矩阵时命题也成立.

情形 2 若 $a_{11} = 0$, 存在 $a_{ii} \neq 0$.

做一次成对行列变换将 A 第 $1, i$ 行和第 $1, i$ 列交换得到的矩阵 A_4 , 其 $(1, 1)$ 元为 $a_{ii} \neq 0$, 所以又可以回到情形 1.

情形 3 若 $a_{11} = a_{22} = \cdots = a_{nn} = 0$, 存在 $a_{ij} \neq 0, i \neq j$.

做一次成对行列变换将 A 第 i 行加到 j 行上, 再将第 i 列加到 j 列上得到矩阵 A_5 , 其 (j, j) 元为 $a_{ij} + a_{ji} = 2a_{ij} \neq 0$, 所以又可以回到情形 1

情形 4 若 $A = 0$.

这种情况下就很显然了. □

Annotation 6.20. 结合 lemma 6.18, 这里就有一种求二次型的标准型的办法: 对于 n 元二次型 $X^T A X$,

$$\begin{pmatrix} A \\ I \end{pmatrix} \xrightarrow[\text{对 } A \text{ 只做其中的初等列变换}]{\text{对 } I \text{ 做成对初等行列变换}} \begin{pmatrix} D \\ C \end{pmatrix}$$

其中 D 是对角矩阵.

Proposition 6.21. n 元二次型 $X^T A X$ 的任一标准形中, 系数不为 0 的平方项个数等于 A 的秩.

证明.

$$C^T A C = \text{diag}\{d_1, d_2, \cdots, d_r, 0, \cdots, 0\}.$$

由 proposition 4.39, 可知 $\text{rank}(A) = r$. □

规范形

Definition 6.22. n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的一个标准形, 若其平方项系数的取值只为 $1, -1, 0$, 那么就称其为一个规范形.

Annotation 6.23. 给定 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的一个标准形为

$$d_1 y_1^2 + \cdots + d_p y_p^2 - d_{p+1} y_{p+1}^2 - \cdots - d_r y_r^2 \quad (1)$$

其中 $d_i > 0, i = 1, 2, \cdots, r$. 那么对 (1) 再做一次非退化线性替换

$$\begin{aligned} y_i &= \frac{1}{\sqrt{d_i}} z_i, \quad i = 1, 2, \cdots, r \\ y_j &= z_j, \quad j = r+1, \cdots, n. \end{aligned}$$

则可以得到一个规范形

$$z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2. \quad (2)$$

二次型的规范性实际上被 r 和 p 所决定了.

Theorem 6.24. 惯性定理 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的规范形是唯一的.

证明. 换句话说就是要证前面提到的 p 是唯一确定的, 即标准型里面系数为正的平方项个数是唯一确定的, 而 r 是 \mathbf{A} 的秩, 所以它总是保持不变的. 假设 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 分别经过两个非退化线性替换 $\mathbf{X} = \mathbf{C} \mathbf{Y}, \mathbf{X} = \mathbf{B} \mathbf{Z}$ 变成了两个规范形

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 \quad (3)$$

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = z_1^2 + z_2^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2 \quad (4)$$

现在来证明 $p = q$, 从而 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的规范形唯一. 由 (3) 式和 (4) 式, 可以构造一个 \mathbf{Z} 到 \mathbf{Y} 的非退化线性替换, 即 $\mathbf{Z} = (\mathbf{B}^{-1} \mathbf{C}) \mathbf{Y}$, 从而有

$$z_1^2 + z_2^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2 = y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2. \quad (5)$$

记 $G = (\mathbf{B}^{-1} \mathbf{C}) = (g_{ij})$. 假设 $p > q$, 我们可以构造一组 y_1, y_2, \cdots, y_n 的取值, 通过 G 的变换得到 z_1, z_2, \cdots, z_n 使得 (5) 两边不等, 造成矛盾. 若我们让 \mathbf{Y}^T 取定

$$\boldsymbol{\beta} = (k_1, k_2, \cdots, k_p, 0, \cdots, 0),$$

因此对应的 \mathbf{Z} 的部分取值可以表示为

$$\begin{cases} z_1 = g_{11}k_1 + \cdots + g_{1p}k_p \\ z_2 = g_{21}k_1 + \cdots + g_{2p}k_p \\ \vdots \\ z_q = g_{q1}k_1 + \cdots + g_{qp}k_p \end{cases}$$

若我们想要 $z_1 = z_2 = \cdots = z_q = 0$, 从而 (5) 式左边肯定是非正的, 并且也希望 (5) 式右边是正, 这样就制造了一个矛盾. 由于这里 $q < p$, 所以我们可以找到一组不全为 0 的 k_1, k_2, \cdots, k_p 使得 $z_1 = z_2 = \cdots = z_q = 0$, 这样就满足了我们的需求. 同理构造从 \mathbf{Y} 到 \mathbf{Z} 的一个非退化线性替换, 并假设当 $p < q$, 也有这样的矛盾产生. 从而 $p = q$. \square

Definition 6.25. 在二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的规范形中, 系数为正的平方项个数 p 被称为 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的**正惯性指数**; 系数为负的平方项个数 $r - p$ 被称为 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的**负惯性指数**. 正惯性指数减去负惯性指数所得的差 $2p - r$ 被称为 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的**符号差**.

正定二次型与正定矩阵

Definition 6.26. 设 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$, 对任意非零的 $\alpha \in \mathbb{R}^n$ 都有 $\alpha^T \mathbf{A} \alpha > 0$, 则称 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是**正定的**.

Theorem 6.27. n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是正定的当且仅当它的正惯性指数等于 n .

证明. 必要性 若 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是正定的, 作非退化线性替换 $\mathbf{X} = \mathbf{C} \mathbf{Y}$, 化为规范形 $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$

$$y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2.$$

其中 \mathbf{B} 就是上述的规范形的系数构成的对角矩阵. 假设 $p < n$, 那么 y_n^2 的系数 b_{nn} 可能为 -1 或者 0 , 因此当 \mathbf{Y} 取值 $\beta = (0, \cdots, 0, 1)^T$ 时, 令 $\alpha = \mathbf{C} \beta$, 有

$$\alpha^T \mathbf{A} \alpha = \beta^T \mathbf{B} \beta = (0, \cdots, 0, 1) \mathbf{B} (0, \cdots, 0, 1)^T = b_{nn} \leq 0.$$

这就矛盾了, 因此 $p \geq n$, 所以 $p = n$.

充分性若 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的正惯性指数等于 n , 于是经过非退化线性替换 $\mathbf{X} = \mathbf{C} \mathbf{Y}$, 化为规范形

$$y_1^2 + y_2^2 + \cdots + y_n^2.$$

因此当 \mathbf{Y} 取值非零 $\beta = (k_1, k_2, \cdots, k_n)$ 时 (这里有一个小问题你要保证所有 \mathbf{X} 的取值能完全被 $\mathbf{C} \mathbf{Y}$ 盖住才行, 这是显然, 因为对任意的 \mathbf{X} 取值, 让 $\mathbf{Y} = \mathbf{C}^{-1} \mathbf{X}$ 即可), 有

$$\mathbf{Y}^T \mathbf{B} \mathbf{Y} = k_1^2 + k_2^2 + \cdots + k_n^2 \geq 0.$$

所以 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是正定的. \square

Definition 6.28. 给定 n 阶对称矩阵 \mathbf{A} , 若它诱导出的 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是正定的, 则称 \mathbf{A} 是**正定矩阵**.

Proposition 6.29. 若 n 阶对称矩阵 \mathbf{A} 是正定的, 那么有如下性质

1. \mathbf{A} 的正惯性指数等于 n ;

2. $A \simeq I$;
3. A 的合同标准形中主对角线元素大于 0;
4. A 的特征值大于 0.

Corollary 6.30. 与正定矩阵合同的对称矩阵也是正定矩阵.

证明. 设 A 是正定的, 存在 C 使得 $C^T A C = B$, 那么

$$\alpha^T B \alpha = \alpha^T C^T A C \alpha = (C \alpha)^T A (C \alpha) > 0.$$

□

Theorem 6.31. n 阶对称矩阵 A 是正定的当且仅当 A 的所有顺序主子式都大于 0

证明. 必要性 若 A 是正定的, 对于 $k \in \{1, 2, \dots, n-1\}$, 把 A 写成分块矩阵

$$\begin{pmatrix} A_k & B_1 \\ B_1^T & B_2 \end{pmatrix},$$

其中 $|A_k|$ 是 A 的 k 阶顺序主子式. 在 \mathbb{R}^k 中取一个列向量 β , 由于 A 是正定的, 于是

$$\begin{pmatrix} \beta \\ 0 \end{pmatrix}^T \begin{pmatrix} A_k & B_1 \\ B_1^T & B_2 \end{pmatrix} \begin{pmatrix} \beta \\ 0 \end{pmatrix} = \beta^T A_k \beta > 0.$$

因此 A_k 也是正定的, 即 $|A_k| > 0$.

充分性 对对称矩阵的阶数进行归纳.

当 $n=1$ 时, 一阶矩阵 (a) , 且 $a > 0$. 显然 (a) 是正定的.

假设命题对 $n-1$ 阶对称矩阵成立 ($n=2, 3, \dots$), 我们来考察 n 阶矩阵 $A = (a_{ij})$. 把 A 写成分块矩阵

$$\begin{pmatrix} A_{n-1} & \beta \\ \beta^T & a_{nn} \end{pmatrix}.$$

由于假设中 A_{n-1} 正定的, 因此 A_{n-1} 满秩可逆, 于是对 A 做两次初等变换, 即

$$\begin{pmatrix} I & 0 \\ -\beta^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \beta \\ \beta^T & a_{nn} \end{pmatrix} \begin{pmatrix} I & -A_{n-1}^{-1} \beta \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - \beta^T A_{n-1}^{-1} \beta \end{pmatrix}. \quad (1)$$

由于

$$(-\beta^T A_{n-1}^{-1})^T = -(A_{n-1}^{-1})^T \beta = -(A_{n-1}^T)^{-1} \beta = -A_{n-1}^{-1} \beta,$$

其中第二个等号用了一下 $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$. 于是 (1) 式得到矩阵 \mathbf{B} 可以写成

$$\mathbf{B} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{n-1}^{-1}\boldsymbol{\beta} \\ \mathbf{0} & 1 \end{pmatrix}^T \begin{pmatrix} \mathbf{A}_{n-1} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^T & a_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{n-1}^{-1}\boldsymbol{\beta} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{0} \\ \mathbf{0} & b \end{pmatrix},$$

其中 $b = a_{nn} - \boldsymbol{\beta}^T \mathbf{A}_{n-1}^{-1} \boldsymbol{\beta}$, 因此 $\mathbf{A} \simeq \mathbf{B}$. 由于 $|\mathbf{A}| = |\mathbf{A}_{n-1}|b$, 所以 $b > 0$. 由于 \mathbf{A}_{n-1} 是正定的, 所以存在可逆矩阵 \mathbf{C} , 使得 $\mathbf{C}^T \mathbf{A}_{n-1} \mathbf{C} = \mathbf{I}_{n-1}$, 再对 \mathbf{B} 做一次成对初等行列变换得到对角矩阵 \mathbf{D} , 即

$$\mathbf{D} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}^T \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{0} \\ \mathbf{0} & b \end{pmatrix} \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1^T \mathbf{A}_{n-1} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & b \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{n-1} & \mathbf{0} \\ \mathbf{0} & b \end{pmatrix}.$$

因此最后有 $\mathbf{A} \simeq \mathbf{B} \simeq \mathbf{D}$, 其中 \mathbf{D} 是正定的, 所以由 corollary 6.30 \mathbf{A} 也是正定的, 那么 n 阶对称矩阵命题也成立. 由数学归纳法原理命题对任意阶对称矩阵均成立. \square

Definition 6.32. 设 n 元二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$, 对任意非零的 $\boldsymbol{\alpha} \in \mathbb{R}^n$ 都有 $\boldsymbol{\alpha}^T \mathbf{A} \boldsymbol{\alpha} \geq 0$, 则称 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是半正定的.

Theorem 6.33. n 阶对称矩阵 \mathbf{A} 是负定的当且仅当它的奇数阶顺序主子式全小于 0, 偶数阶顺序主子式全大于 0.