2021 春季学期线性代数

课堂笔记

杨 春雨

USTC | 安徽省合肥市金寨路 96 号

目录

Chapter Ⅰ 向量与复数

向量的基本定义

单位向量：$\boldsymbol{a}^0=\frac{\boldsymbol{a}}{|\boldsymbol{a}|}$

在仿射空间中点和向量是不一样的。

定义 1.6

给定三个向量 $\mathbf{a},\mathbf{b},\mathbf{c}$，称 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 为 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的*混合积*，它是一个数量。

混合积常用于证明三个向量共面，其几何含义是以$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$为棱的平行六面体的**有向体积**。
当$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$构成右手系时，结果是正的，当$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$构成左手系时，结果是负的。
且混合积具有**轮换对称性**：$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}=(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}$

记$\boldsymbol{a}=(a_1,a_2,a_3),\boldsymbol{b}=(b_1,b_2,b_3),\boldsymbol{c}=(c_1,c_2,c_3)$,则$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\begin{vmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{vmatrix}$

定义 1.7

给定三个向量 $\mathbf{a},\mathbf{b},\mathbf{c}$，称 $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 为这三个向量的二重外积。

命题 1.7

对任意向量 $\mathbf{a},\mathbf{b},\mathbf{c}$，有

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}.$$

Pf:
引理：$\boldsymbol{a}^2\boldsymbol{b}^2=(\boldsymbol{a}\times\boldsymbol{b})^2+(\boldsymbol{a}\cdot\boldsymbol{b})^2$
考虑$\boldsymbol{a},\boldsymbol{b},\boldsymbol{a}\times\boldsymbol{b}$三个向量显然不共面，所以$\boldsymbol{a},\boldsymbol{b},\boldsymbol{a}\times\boldsymbol{b}$可以作为空间中的一组基，
记$\boldsymbol{c}=\alpha\boldsymbol{a}+\beta\boldsymbol{b}+\gamma(\boldsymbol{a}\times\boldsymbol{b})$
$\therefore(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=(\boldsymbol{a}\times\boldsymbol{b})\times\big(\alpha\boldsymbol{a}+\beta\boldsymbol{b}+\gamma(\boldsymbol{a}\times\boldsymbol{b})\big)=\alpha(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{a}+\beta(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}$
我们先考虑这个公式的一个退化情形：$\boldsymbol{b}=\boldsymbol{c}$时应该有$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}=(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{b}-\boldsymbol{b}^2\boldsymbol{a}$
下面我们来证明这个公式：

显然$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}$和$\boldsymbol{a},\boldsymbol{b}$是共面的，不妨设$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}=\lambda\boldsymbol{a}+\mu\boldsymbol{b}$

等式两边同时点乘上$\boldsymbol{a},\boldsymbol{b}$可以得到：$\begin{cases}\lambda\boldsymbol{a}^2+\mu\boldsymbol{a}\cdot\boldsymbol{b}=-(\boldsymbol{a}\times\boldsymbol{b})^2\\ \lambda\boldsymbol{a}\cdot\boldsymbol{b}+\mu\boldsymbol{b}^2=0\end{cases}$

利用 Cramer 规则求解此方程组可得：$\begin{cases}\lambda=-\boldsymbol{b}^2\\ \mu=\boldsymbol{a}\cdot\boldsymbol{b}\end{cases}$

代入$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}=\lambda\boldsymbol{a}+\mu\boldsymbol{b}$得$(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}=(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{b}-\boldsymbol{b}^2\boldsymbol{a}$

利用此公式可以得到

$$
\begin{aligned}
&\alpha(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{a}+\beta(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{b}\\
&=\alpha[\boldsymbol{a}^2\boldsymbol{b}-(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{a}]+\beta[(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{b}-\boldsymbol{b}^2\boldsymbol{a}]\\
&=[\alpha\boldsymbol{a}^2+\beta(\boldsymbol{a}\cdot\boldsymbol{b})]\boldsymbol{b}-\alpha[(\boldsymbol{a}\cdot\boldsymbol{b})+\boldsymbol{b}^2]\boldsymbol{a}\\
&=\boldsymbol{a}\cdot(\alpha\boldsymbol{a}+\beta\boldsymbol{b})\boldsymbol{b}-\boldsymbol{b}\cdot(\alpha\boldsymbol{a}+\beta\boldsymbol{b})\boldsymbol{a}\\
&=(\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b}-(\boldsymbol{b}\cdot\boldsymbol{c})\boldsymbol{a}
\end{aligned}
$$

证毕

复数和复数的四则运算：

复数 $z=x+iy,\ x,y\in\mathbb{R}$

x,y 分别称为复数 z 的**实部**与**虚部**，记作 Rez 和 Imz

$z_1=x_1+iy_1$，$z_2=x_2+iy_2$，$\lambda\in\mathbb{R}$

加法运算：$z_1+z_2=(x_1+x_2)+i(y_1+y_2)$

数乘运算：$\lambda z_1=\lambda x_1+i\lambda y_1$

乘法运算：$z_1z_2=(x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1)$

除法运算：$\dfrac{z_1}{z_2}=\dfrac{(x_1x_2+y_1y_2)+i(x_1y_2-x_2y_1)}{x_2^2+y_2^2},\ z_2\neq0$

欧拉公式：$e^{i\theta}=\cos\theta+i\sin\theta$

复数的除法是有定义的，且根据共轭复数有$\frac{z_1}{z_2}=\frac{z_1\bar{z}_2}{z_2\bar{z}_2}$

定义 1.10

复数 z 所对应的向量 $\overrightarrow{OP}$ 与 x 轴正向的夹角 θ 称为复数 z 的**幅角**，记作 $\arg z$

复数的辐角并不唯一，可能彼此相差2π的整数倍。我们在实际应用中取$0\leq\arg z<2\pi$，

称为辐角主值。

复数的三角表示：

$$z = x + iy = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta)$$

$$z_1 = r_1(\cos\theta_1 + i\sin\theta_1),\ z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$$

$$z_1 z_2 = r_1 r_2\left[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right]$$

$$|z_1 z_2| = |z_1||z_2|,\ \arg(z_1 z_2) = \arg z_1 + \arg z_2$$

利用欧拉公式

$$z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1+\theta_2)}$$

封闭的定义：设F是一个数集，在F中任取两个数做某种运算，如果其结果仍然在F中，则称数集F对这种运算时封闭的。

定义 1.11

设数集 F 至少包含两个不同的元素，称 F 为**数域**，如果 F 对数的加减乘除运算是封闭的，即当 $a, b \in F$ 时，$a \pm b,\ ab,\ \dfrac{a}{b}\ (b \neq 0) \in F$。

容易得到，任意一个数域中一定会有0,1两个元素。

代数基本定理

任何一个一元复系数方程式都至少有一个复数根，即复数域是代数封闭的。

证明：
https://baike.baidu.com/item/%E4%BB%A3%E6%95%B0%E5%9F%BA%E6%9C%AC%E5%AE%9A%E7%90%86/18104?anchor=3

推论：任意一个非零的一元n次复系数多项式，有且仅有n个复数根（包括重根）

多重求和的性质：

多重求和：从内向外对每个求和符号逐次求和，例如：

$$\begin{aligned}\sum_{j=1}^{m}\sum_{i=1}^{n}a_{ij} &= \sum_{j=1}^{m}(a_{1j}+a_{2j}+\cdots+a_{nj}) \\ &= \sum_{j=1}^{m}a_{1j}+\sum_{j=1}^{m}a_{2j}+\cdots+\sum_{j=1}^{m}a_{nj}.\end{aligned}$$

容易看出：$\displaystyle\sum_{j=1}^{m}\sum_{i=1}^{n}a_{ij}=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}$

性质：在多重求和时，可以交换求和符号的次序。

Chapter Ⅱ 线性方程组

线性方程组的定义和高斯消元法

定义 2.2

若将 $x_1 = c_1, x_2 = c_2, \cdots, x_n = c_n$ 代入上述方程等式都成立，则称 $(c_1, c_2, \cdots, c_n)$ 为该方程组的**一组解**。线性方程组解的全体称为该方程组的**解集**。如果解集非空，则称线性方程组是**相容**的; 否则，称线性方程组**不相容**.

采用矩阵的表示，线性方程组可写成：

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

称为线性方程组的**增广矩阵**

利用 Gauss 消元法将矩阵转化为最简形式 a.k.a 标准形式：

最简形式或标准形式

$$\begin{pmatrix} c_{11} & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & c_{1n} & d_1 \\ 0 & \ldots & 0 & c_{2j_2} & \ldots & \ldots & \ldots & \ldots & c_{2n} & d_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & c_{rj_r} & \ldots & c_{rn} & d_r \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & d_{r+1} \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

其中 $c_{11}, c_{2j_2}, \ldots, c_{rj_r}$ 均非零。

定理 2.2

线性方程组 (1) 的解的属性如下：

- 当 $d_{r+1} \neq 0$ 时，线性方程组 (1) 无解；
- 当 $d_{r+1} = 0$ 且 $r = n$ 时，线性方程组 (1) 有唯一解；
- 当 $d_{r+1} = 0$ 且 $r < n$ 时，线性方程组 (1) 有多解。

通解

$$\begin{cases} x_1 = & \alpha_{11}t_1 + \ldots + \alpha_{1,n-r}t_{n-r} + \beta_1 \\ x_2 = & \alpha_{21}t_1 + \ldots + \alpha_{2,n-r}t_{n-r} + \beta_2 \\ \quad\vdots & \\ x_n = & \alpha_{n1}t_1 + \ldots + \alpha_{n,n-r}t_{n-r} + \beta_n \end{cases}$$

其中 $t_1, \ldots, t_{n-r}$ 为参数，

$\alpha_{ij}, \beta_i \in F$, $i = 1, \ldots, n$, $j = 1, \ldots, n-r$。

写成向量形式为$\boldsymbol{x} = t_1\boldsymbol{\alpha_1} + t_2\boldsymbol{\alpha_2} + t_{n-r}\boldsymbol{\alpha_{n-r}} + \boldsymbol{\beta}$

对于齐次线性方程组，总是有$d_{r+1} = 0$，所以齐次线性方程组总是有解，特别的有$x_1 = 0, x_2 = 0, \ldots, x_n = 0$总是齐次线性方程组的一组解，我们称之为**零解**或者**平凡解**。齐次线性方程组的非零解称之为非平凡解。

推论 2.1

齐次线性方程组有非零解的充要条件为 $r < n$。

Chapter Ⅲ 矩阵与行列式

矩阵和矩阵的运算

对角元：对于矩阵的一个元素a_{ij}，当$i=j$时，称a_{ij}是矩阵的对角元
数量阵：对角元是a，其他元素都是0的矩阵称为数量阵，记作aI
对角阵：除了对角线上的元素都是0的矩阵称为对角阵，可以记作：

$$A=diag(a_{11},a_{22},\dots,a_{nn})\text{或者}A=\begin{pmatrix}a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn}\end{pmatrix}$$

- 若方阵 $A=(a_{ij})_{n\times n}$ 满足 $a_{ij}=a_{ji}$ 对所有 i,j 成立，则 A 称为**对称阵**。
- 若方阵 $A=(a_{ij})_{n\times n}$ 满足 $a_{ij}=-a_{ji}$ 对所有 i,j 成立，则 A 称为**反对称阵**。
- 若矩阵 A 的元素都是整数、有理数、实数、复数、多项式等，则 A 分别称为整数矩阵、有理数矩阵、实矩阵、复矩阵、多项式矩阵。一般地，若 A 的元素都取自某个数域 F，则 A 称为**数域 F 上的矩阵**。数域 F 上的 $m\times n$ 矩阵的全体，记作 $F^{m\times n}$。

第(i,j) 元素是1，其它元素全是0的$m\times n$矩阵称为**基本矩阵**，记作E_{ij}.
容易验证每一个$m\times n$矩阵$A=(a_{ij})_{m\times n}$都可以表示为：

$$A=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}E_{ij}$$

我们规定：$A^0=I$

矩阵的转置满足：$(AB)^T=B^TA^T$

分块矩阵

给定矩阵 $A=(a_{ij})_{m\times n}$，可将 A 用水平线与竖直线分成若干块，

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \cdots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}$$

称为**分块矩阵**，记作 $A=(A_{ij})_{r\times s}$，每个 $A_{ij}\in F^{m_i\times n_j}$ 称为 A 的**子块**。

普通矩阵可以看作分块矩阵的一个特殊情况，即分块矩阵中的每一个块A_{ij}中都只有一个元素a_{ij}.

分块可以是**任意**的，但一般情形下，都是带有目的地——**发掘矩阵的特殊结构**，譬如零矩阵块，单位阵块，数量阵块，对称块等。

若 $A_{ij}=O$ 对所有 $i\neq j$ 成立，则称 A 为**准对角阵**，记作

$$A=\begin{pmatrix} A_{11} & & \\ & \ddots & \\ & & A_{rr} \end{pmatrix} \quad \text{或} \quad A=\mathrm{diag}(A_{11},A_{22},\cdots,A_{rr}).$$

若 $A_{ij}=O$ 对所有 $i>j$ 成立，则 A 称为**准上三角阵**。

若 $A_{ij}=O$ 对所有 $i<j$ 成立，则称 A 为**准下三角阵**。

准上三角阵和准下三角阵统称为**准三角阵**。

子矩阵

由 A 的第 $i_1,i_2,\cdots,i_r$ 行和第 $j_1,j_2,\cdots,j_s$ 列上的元素依次排列组成的 $r\times s$ 矩阵称为 A 的**子矩阵**，记作

$$A\begin{pmatrix} i_1 i_2\cdots i_r \\ j_1 j_2\cdots j_s \end{pmatrix}=\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_s} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_s} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_s} \end{pmatrix}.$$

对矩阵的分块运算取转置矩阵：$A^T=({A_{ji}}^T)_{s\times r}$

初等方阵

对单位方阵施行初等变换，得到的方阵称为**初等方阵**。

左乘是行变换，右乘是列变换

① 交换矩阵的第i,j行（或交换第i,j列）

$$
S_{ij} = \begin{array}{c}\left(\begin{array}{ccccccc}
1 & & & & & & \\
& \ddots & & & & & \\
& & 0 & & 1 & & \\
& & & \ddots & & & \\
& & 1 & & 0 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{array}\right)\end{array}
\begin{array}{l} \\ \\ \leftarrow 第\ i\ 行 \\ \\ \leftarrow 第\ j\ 行 \\ \\ \\ \end{array}
$$

② 将单位阵的第i行（或第i列）乘以非零常数λ

$$
D_i(\lambda) = \left(\begin{array}{ccccccc}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & & \lambda & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{array}\right)
\begin{array}{l} \\ \\ \\ \leftarrow 第\ i\ 行 \\ \\ \\ \\ \end{array} \qquad (2)
$$

③ 将单位阵的第j行的λ倍加到第i行（或将第i列的λ倍加到第j列）

$$
T_{ij}(\lambda) = \left(\begin{array}{ccccccc}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & \lambda & & \\
& & & \ddots & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{array}\right)
\begin{array}{l} \\ \\ \leftarrow 第\ i\ 行 \\ \\ \leftarrow 第\ j\ 行 \\ \\ \\ \end{array} \qquad (3)
$$

使用**矩阵方法**处理斐波那契数列：

$$F_n = F_{n-1} + F_{n-2}, F_1 = 0, F_2 = 1$$

不难得到：$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix}$；

所以$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} F_2 \\ F_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

引理任何一个矩阵我们都可以写成$C = ABA^{-1}$，从而，$C^n = AB^nA^{-1}$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$$

所以$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^{n-2} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

对于方阵而言“满秩”和“可逆”是等价的概念。

矩阵的迹有如下性质：

$$tr(A+B) = tr(A) + tr(B)$$
$$tr(AB) = tr(BA)$$

如果对矩阵施行一系列初等行变换与初等列变换，目标是使 得矩阵具有尽可能简单的形式，那么最终的形式是怎样的？

定理 3.8

对任意矩阵 $A = (a_{ij})_{m\times n}$，存在一系列 m 阶初等方阵 $P_1, \ldots, P_s$ 和 n 阶初等方阵 $Q_1, \ldots, Q_t$，使得

$$P_s \cdots P_1 \, A \, Q_1 \cdots Q_t = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix},$$

其中 r 为非负整数。

行列式

行列式有很多等价定义：递归定义，展开式定义，几何定义……

几何定义方式：把有向面积、有向体积所具有的本征性质推广到一般的 n 维数组空间。

有向面积、有向体积的本征性质：多重线性性、反对称性、规范性。

D1：多重线性性

若记 $\Delta_2 = \det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$，则

$$\det(\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\beta}, \boldsymbol{\alpha}_2) = \lambda \det(\boldsymbol{\alpha}, \boldsymbol{\alpha}_2) + \mu \det(\boldsymbol{\beta}, \boldsymbol{\alpha}_2),$$
$$\det(\boldsymbol{\alpha}_1, \lambda\boldsymbol{\alpha} + \mu\boldsymbol{\beta}) = \lambda \det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}) + \mu \det(\boldsymbol{\alpha}_1, \boldsymbol{\beta}),$$

对任意的 $\lambda, \mu \in F$ 都成立。

推广：$\det(\boldsymbol{\alpha}_1, \dots, \lambda\boldsymbol{\eta}_i + \mu\boldsymbol{\xi}_i, \dots, \boldsymbol{\alpha}_n) =$
$\lambda \det(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\eta}_i, \dots, \boldsymbol{\alpha}_n) + \mu \det(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\xi}_i, \dots, \boldsymbol{\alpha}_n)$，$\lambda, \mu \in F$ 对 $i = 1, 2, \dots, n$ 都成立。

D2：反对称性

$$\det(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) = -\det(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) \quad \Longleftrightarrow \quad \det(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 0,\ \forall \boldsymbol{\alpha}$$

推广：
$\det(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_j, \dots, \boldsymbol{\alpha}_i, \dots, \boldsymbol{\alpha}_n) = -\det(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_i, \dots, \boldsymbol{\alpha}_j, \dots, \boldsymbol{\alpha}_n)$，
对任意的 $i, j = 1, 2, \dots, n$ 都成立。

$$\Longleftrightarrow \quad \det(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}, \dots, \boldsymbol{\alpha}, \dots, \boldsymbol{\alpha}_n) = 0,\ \forall \boldsymbol{\alpha}$$

D3：规范性

$$\det(\mathbf{e}_1, \mathbf{e}_2) = 1$$

推广：$\det(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$，其中

$\mathbf{e}_i = (0, \cdots, \overset{\overset{i}{\downarrow}}{1}, \cdots, 0)$，$i = 1, 2, \dots, n$。

定义 3.7

设 F^n 为 n 维数组空间，$\det : \underbrace{F^n \times \cdots \times F^n}_{n} \mapsto F$ 且满足 D_1, D_2, D_3 性质，则称 $\det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n)$ 为 n 阶行列式。

行列式是F^n上的规范反对称n重线性函数（可证明这样的函数存在且唯一）

定理 3.9

行列式具有以下性质：

(1) D4：若某一向量 $\boldsymbol{\alpha}_j = \mathbf{0}$，则 $\det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = 0$；

(2) D5：若 $\boldsymbol{\alpha}_j = k\boldsymbol{\alpha}_i \ (i \neq j)$，则 $\det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = 0$；

(3) D6：对 $\forall i \neq j$ 有，$\det(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i + \lambda\boldsymbol{\alpha}_j, \ldots, \boldsymbol{\alpha}_j, \ldots, \boldsymbol{\alpha}_n)$
$= \det(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i, \ldots, \boldsymbol{\alpha}_j, \ldots, \boldsymbol{\alpha}_n)$。

定义 3.8

将 n 个两两不同的正整数 $s_1, s_2, \ldots, s_n$ 按顺序排成的一个有序数组称为一个*排列*，记做 $s = (s_1, s_2, \ldots, s_n)$。满足 $s_1 < s_2 < \ldots < s_n$ 的排列称为*顺序排列*。满足 $i < j$ 且 $s_i > s_j$ 的一对数 (s_i, s_j) 称为 s 的一个*逆序*。s 的逆序的个数称为 s 的*逆序数*，记作 $\tau(s)$。逆序数为奇数的排列称为*奇排列*；逆序数为偶数的排列称为*偶排列*。

命题 3.10

对换具有以下性质：

(1) 任一排列经过一次对换，必改变其奇偶性；

(2) 任意 n 元排列 $(s_1, s_2, \ldots, s_n)$ 都可经过有限次对换变成标准排列 $(1, 2, \ldots, n)$；同一排列 $(s_1, s_2, \ldots, s_n)$ 变成标准排列所经历的对换次数 s 不唯一，但是 s 的奇偶性是确定的且与排列的奇偶性相同。

证明：任意一个排列在经历过一次对换之后，一定会改变其奇偶性。
给定任意一个排序$(a_1, a_2, \ldots, a_i, \ldots, a_j, \ldots, a_{n-1}, a_n)$，不妨设$a_i < a_j$，假设$a_i$到$a_j$之间一共有$k$个元素，假设有$A$个元素比$a_i$小，则有$k - A$个元素比$a_i$大，故对于$a_i$逆序数变化为$k - 2A$；对于$a_j$，假设有$B$个元素比$a_j$大，则有$k - B$个元素比$a_j$小，所以对于$a_j$逆序数变化为$k - 2B$。

a_i和a_j对换造成的逆序数变化为1，所以逆序数总的变化是$2k - 2(A + B) + 1$，是一个奇数，

所以会使得排列的奇偶性发生改变。
证毕

对每个排列 $(j_1,j_2,\ldots,j_n)$ 引入奇偶性符合:

$$\operatorname{sgn}(j_1,j_2,\ldots,j_n) = (-1)^{\tau(j_1,j_2,\cdots,j_n)}$$
$$= \begin{cases} 1, & 当(j_1,j_2,\ldots,j_n)为偶排列; \\ -1, & 当(j_1,j_2,\ldots,j_n)为奇排列. \end{cases}$$

设 $\boldsymbol{\beta}_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i$ $(j=1,2,\ldots,n)$（视为列向量），则

$$\det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n) = \det(\sum_{i_1=1}^{n} a_{i_1 1}\mathbf{e}_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2}\mathbf{e}_{i_2},\ldots, \sum_{i_n=1}^{n} a_{i_n n}\mathbf{e}_{i_n})$$
$$= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}\det(\mathbf{e}_{i_1},\mathbf{e}_{i_2},\ldots,\mathbf{e}_{i_n})$$
$$= \sum_{(i_1,i_2,\ldots,i_n)} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}\det(\mathbf{e}_{i_1},\mathbf{e}_{i_2},\ldots,\mathbf{e}_{i_n}),$$

再利用 $\det(\mathbf{e}_{i_1},\mathbf{e}_{i_2},\ldots,\mathbf{e}_{i_n}) = (-1)^{\tau(i_1,i_2,\ldots,i_n)}$，可得

$$\det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n) = \sum_{(i_1,i_2,\ldots,i_n)} (-1)^{\tau(i_1,i_2,\ldots,i_n)} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}.$$

不要把$\det(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\ldots,\boldsymbol{e}_{i_n})$和前面的一堆求和混到一起。

定义 3.9

设方阵 $A=(a_{ij})_{n\times n}$，称 A 的行向量组（列向量组）张成 n 平行多面体的有向体积为矩阵 A 的**行列式**，记作

$$\det(A) = \det(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)^{\mathsf{T}} = \det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)$$
$$= |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$
$$= \sum_{(i_1,i_2,\ldots,i_n)} (-1)^{\tau(i_1,i_2,\ldots,i_n)} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}$$
$$= \sum_{(j_1,j_2,\ldots,j_n)} (-1)^{\tau(j_1,j_2,\ldots,j_n)} a_{1 j_1}a_{2 j_2}\cdots a_{n j_n}$$

定理 3.11

方阵的行列式具有以下性质：

(1) $\det(A^{\mathrm{T}}) = \det(A)$；

(2) 当方阵 $A = (a_{ij})_{n\times n}$ 为上角阵时，则

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn};$$

(3) 当准上角阵 $A = (A_{ij})_{k\times k}$ 的每个对角块 A_{ii} 都是方阵时，则

$$\det(A) = \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ & A_{22} & \cdots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{vmatrix} = \det(A_{11})\det(A_{22})\cdots\det(A_{kk})。$$

定理 3.12

对于任意两个 n 阶方阵 A 和 B，有 $\det(AB) = \det(A)\det(B)$。

证明：容易验证，任意两个上三角阵相乘，结果仍然是上三角阵，且每个对角元正是原来两个矩阵相应对角元的乘积。

使用类似高斯消元法的变换可以把A,B转变为上三角阵A',B'（不妨设$PA = A', BQ = B'$）且矩阵的行列式不变，而$\det(A'B') = \det(A')\det(B')$

再将同样的变换作用于AB，则有$PABQ = A'B'$

所以有$\det(AB) = \det(A'B') = \det(A')\det(B') = \det(A)\det(B)$

（Binet-Cauchy 公式）

设 $A \in F^{m\times n}$, $B \in F^{n\times m}$，则 $\det(AB) =$

$$\begin{cases} \displaystyle\sum_{1\leqslant j_1<\cdots<j_m\leqslant n} \det\left(A\begin{pmatrix} 12\cdots m \\ j_1j_2\cdots j_m \end{pmatrix}\right)\det\left(B\begin{pmatrix} j_1j_2\cdots j_m \\ 12\cdots m \end{pmatrix}\right), & \text{当} m < n; \\ \det(A)\det(B), & \text{当} m = n; \\ 0, & \text{当} m > n. \end{cases}$$

证明：（邱维声高等代数学习指导 P180）

行列式的计算

技巧 1

利用初等变换，将方阵化为上三角阵、下三角阵或对角阵。

例 3.1

计算 n 阶行列式

$$\begin{vmatrix} x & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & x \end{vmatrix}.$$

将原式第一行加上其余各行可得

$$\begin{vmatrix} x+n-1 & x+n-1 & \dots & x+n-1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & x \end{vmatrix} = (x+n-1)\begin{vmatrix} 1 & 1 & \dots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & x \end{vmatrix}$$

$$= (x+n-1)\begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & x-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & x-1 \end{vmatrix} = (x+n-1)(x-1)^{n-1}$$

例 3.2

计算 Vandermonde 行列式

$$\Delta_n(a_1, a_2, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}.$$

从行列式的第n列开始依次把各列减去前一列的a_n倍，这就得到：

$$\Delta_n(a_1, a_2, \dots, a_n) = \begin{vmatrix} 1 & a_1 - a_n & \dots & a_1^{n-2}(a_1 - a_n) \\ \vdots & \vdots & & \vdots \\ 1 & a_{n-1} - a_n & \dots & a_{n-1}^{n-2}(a_{n-1} - a_n) \\ 1 & 0 & 0 & 0 \end{vmatrix}$$

$$= (-1)^{n+1}\begin{vmatrix} a_1-a_n & a_1(a_1-a_n) & \dots & a_1^{n-2}(a_1-a_n) \\ a_2-a_n & a_2(a_2-a_n) & \dots & a_1^{n-2}(a_2-a_n) \\ \vdots & \vdots & & \vdots \\ a_{n-1}-a_n & a_{n-1}(a_{n-1}-a_n) & \dots & a_{n-1}^{n-2}(a_{n-1}-a_n) \end{vmatrix}$$

$$= \prod_{i=1}^{n-1}(a_n-a_i)\cdot\Delta_{n-1}(a_1,a_2,\dots,a_{n-1})$$

容易得到$\Delta_1=1,\Delta_2=a_2-a_1$，利用数学归纳法可得：

$$\Delta_n(a_1,a_2,\dots,a_n)=\begin{cases} 1\,, n=1 \\ \prod_{j=2}^{n}\prod_{i=1}^{j-1}(a_n-a_i)=\prod_{1\le i<j\le n}^{n}(a_n-a_i),\ n\ge 2 \end{cases}$$

技巧 2

利用行列式的Laplace 展开，将高阶行列式 $\Rightarrow$ 低阶行列式。

定义 3.10

任意矩阵 $A=(a_{ij})_{m\times n}$ 的 k 阶子矩阵 $A\binom{i_1i_2\cdots i_k}{j_1j_2\cdots j_k}$ 的行列式，即 $\left|A\binom{i_1i_2\cdots i_k}{j_1j_2\cdots j_k}\right|$，称为 A 的k **阶子式**，简称**子式**。

定义 3.11

删去 n 阶方阵 $A=(a_{ij})_{n\times n}$ 的第 i 行和第 j 列之后，剩下的 $n-1$ 阶子矩阵的行列式 M_{ij} 称为元素 a_{ij} 的**余子式**。

$A_{ij}=(-1)^{i+j}M_{ij}$ 称为元素 a_{ij} 的**代数余子式**。

定理 3.13

（Laplace 展开定理）设 n 阶方阵 $A=(a_{ij})_{n\times n}$，则

$\det(A)=\sum_{i=1}^{n}a_{ij}A_{ij},\ \forall j=1,2,\dots n$（按某一列展开）；

$\det(A)=\sum_{j=1}^{n}a_{ij}A_{ij},\ \forall i=1,2,\dots n$（按某一行展开）。

Laplace 展开有更一般的形式：

定理 3.14

（Laplace 展开定理）设 n 阶方阵 $A=(a_{ij})_{n\times n}$，取定行指标 $1\leqslant i_1<\ldots<i_p\leqslant n$（列展开情形类似），则

$$\det(A)=\sum_{1\leqslant j_1<\ldots<j_p\leqslant n}\left|A\begin{pmatrix}i_1\cdots i_p\\ j_1\cdots j_p\end{pmatrix}\right|\left[(-1)^{i_1+\cdots+i_p+j_1+\cdots+j_p}\left|A\begin{pmatrix}i_{p+1}\cdots i_n\\ j_{p+1}\cdots j_n\end{pmatrix}\right|\right].$$

定理 3.15

设 n 阶方阵 $A=(a_{ij})_{n\times n}$，则某一行（列）与另一行（列）相应元素的代数余子式乘积之和等于 0，即

$\sum_{k=1}^{n}a_{ik}A_{jk}=0,\ \forall i\neq j$（按某一行展开）；

$\sum_{k=1}^{n}a_{ki}A_{kj}=0,\ \forall i\neq j$（按某一列展开）。

证明是显然的（相当于求有一行（列）是相同的矩阵的行列式）

定义 3.12

设 $A=(a_{ij})$ 为 n 阶方阵，引入

$$A^*:=(A_{ji})_{n\times n}=\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\cdots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}$$

其中 A_{ij} 是 a_{ij} 的代数余子式，称 A^* 为 A 的**伴随方阵**。

需要主义的是是A_{ji}而不是A_{ij}.

$A^*A=AA^*=\det(A)\,I$

利用定理 3.15 和 Laplace 展开定理容易得到这个结论

例 3.3

设 $n \geq 2$，计算 n 阶行列式

$$\Delta_n = \begin{vmatrix} x & & & & a_0 \\ -1 & x & & & a_1 \\ & -1 & \ddots & & \vdots \\ & & \ddots & x & a_{n-2} \\ & & & -1 & x+a_{n-1} \end{vmatrix}.$$

将此行列式按第n列展开得：$\Delta_n = \sum_{k=1}^{n-1}(-1)^{n+k}a_{k-1}A_{kn} + (x+a_{n-1})A_{nn}$

$A_{kn} = (-1)^{n-k}x^{k-1}, k = 1,2,\ldots,n$

所以$\Delta_n = \sum_{k=1}^{n-1}a_{k-1}x^{k-1} + (x+a_{n-1})x^{n-1} = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$

例 3.4

设 $n \geq 2$，计算 n 阶行列式

$$\Delta_n = \begin{vmatrix} \alpha+\beta & \alpha & & & \\ \beta & \alpha+\beta & \alpha & & \\ & \ddots & \ddots & \ddots & \\ & & \beta & \alpha+\beta & \alpha \\ & & & \beta & \alpha+\beta \end{vmatrix}.$$

将这个行列式按第一行展开容易得到$\Delta_n = (\alpha+\beta)\Delta_{n-1} - \alpha\beta\Delta_{n-2}$

利用特征根法当$\alpha \neq \beta$容易得到$\Delta_n = C_1\alpha^n + C_2\beta^n$

又因为$\Delta_1 = \alpha+\beta, \Delta_2 = \alpha^2+\beta^2+\alpha\beta$，容易得到$C_1 = \frac{\alpha}{\alpha-\beta}, C_2 = -\frac{\beta}{\alpha-\beta}$

所以$\Delta_n = \frac{\alpha^2-\beta^2}{\alpha-\beta}$

当$\alpha = \beta$时，容易得到$\Delta_n = C_1\alpha^n + C_2n\alpha^n$，此时$\Delta_1 = 2\alpha, \Delta_2 = 3\alpha^2$，所以$C_1 = C_2 = 1$

故$\Delta_n = (n+1)\alpha^n$

技巧 3

将方阵分解成若干矩阵的乘积，利用 Binet-Cauchy 公式或一些恒等式来计算行列式。

例 3.5

计算 n 阶行列式

$$\Delta_n = \begin{vmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ & \cdots & \cdots & \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{vmatrix}.$$

例 3.6

设 A 是 $m \times n$ 矩阵，B 是 $n \times m$ 矩阵，证明:

$$\det(I_n - BA) = \det(I_m - AB).$$

完整的式子是$\det(I_n - BA) = \det\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} = \det(I_m - AB)$

更一般的，有$\lambda^m \det(\lambda I_n - BA) = \lambda^n \det(\lambda I_m - AB)$

延伸：计算下列 n 阶行列式

$$\Delta_n = \begin{vmatrix} 1 + x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & 1 + x_2y_2 & \cdots & x_2y_n \\ & \cdots & \cdots & \\ x_ny_1 & x_ny_2 & \cdots & 1 + x_ny_n \end{vmatrix}.$$

技巧 4

视行列式 $\det(A)$ 为某些元素的多项式，确定 $\det(A)$ 的次数 m，求出 $\det(A)$ 的 m 个根 $r_1, r_2, \ldots, r_m$，再利用 $\det(A) = c(x - r_1)(x - r_2)\cdots(x - r_m)$，确定常数 c 即可。

例 3.7

计算 n 阶行列式

$$\Delta_n = \begin{vmatrix} x & a_1 & a_2 & \cdots & a_{n-1} \\ a_1 & x & a_2 & \cdots & a_{n-1} \\ & \cdots & \cdots & \cdots & \\ a_1 & a_2 & \cdots & \cdots & x \end{vmatrix}.$$

观察的得到$x = a_1, a_2, \dots, a_{n-1}, -\sum_{i=1}^{n-1} a_i$是方程$\det(A) = 0$的根，
所以$\det(A) = c(x - a_1)(x - a_2) \dots (x - a_{n-1})(x + \sum_{i=1}^{n-1} a_i)$
（也可以利用例 3.1 的方法）

定义 3.13

设 A 是一个 n 阶方阵，如果存在 n 阶方阵 X 满足

$$XA = AX = I,$$

则称 A **可逆**，并称 X 为 A 的**逆矩阵**，记作 A^{-1}。

可逆方阵也成为**非奇异方阵**，不可逆方阵称为**奇异方阵**

定理 3.16

方阵 A 可逆的充分必要条件是 $\det(A) \neq 0$。当 A 可逆时，A 有唯一的逆矩阵 $\frac{1}{\det(A)} A^*$。

定理 3.17

对任意 n 阶可逆方阵 A, B，都有

(1) $(A^{-1})^{-1} = A$；

(2) $(\lambda A)^{-1} = \lambda^{-1} A^{-1}$, $\lambda \neq 0$；

(3) $(AB)^{-1} = B^{-1} A^{-1}$；

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

(5) 当 $A_1, \cdots, A_r$ 都可逆时，
$(\mathrm{diag}(A_1, \dots, A_r))^{-1} = \mathrm{diag}(A_1^{-1}, \dots, A_r^{-1})$。

定理 3.18

初等方阵具有下列性质：

(1) S_{ij} 为对称方阵，且 $S_{ij}^{-1}=S_{ij}$；

(2) $D_i(\lambda)$ 为对角方阵，且 $D_i(\lambda)^{-1}=D_i(\lambda^{-1})$；

(3) $T_{ij}(\lambda)$ 为三角方阵，且 $T_{ij}(\lambda)^{-1}=T_{ij}(-\lambda)$。

定理 3.19

方阵 A 可逆的充分必要条件是 A 可以分解为一系列初等方阵的乘积。

技巧 1

利用恒等式$AA^*=A^*A=\det(A)I \ \Rightarrow \ A^{-1}=\frac{1}{\det(A)}A^*$。

例 3.8

设 n 方阵

$$A=\begin{pmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2 & 1 \\ & & & 1 & 2 \end{pmatrix},$$

求 A^{-1}。

根据例 3.4 结论$\det(A)=n+1$

技巧 2

利用初等变换：

$$\left(A \vdots I\right) \xrightarrow{\text{初等行变换}} \left(I \vdots A^{-1}\right)$$

或求解 n 个线性方程组 $A\mathbf{x}_j=\mathbf{e}_j$ $(j=1,2,\ldots,n)$。

例 3.9

设 n 方阵

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

求 A^{-1}。

利用技巧 2 容易得到：

$$A^{-1}=\begin{vmatrix}\frac{2-n}{n-1}&\frac{1}{n-1}&\frac{1}{n-1}&\cdots&\frac{1}{n-1}\\\frac{1}{n-1}&\frac{2-n}{n-1}&\frac{1}{n-1}&\cdots&\frac{1}{n-1}\\\vdots&\vdots&\vdots&\ddots&\vdots\\\frac{1}{n-1}&\frac{1}{n-1}&\frac{1}{n-1}&\cdots&\frac{2-n}{n-1}\end{vmatrix}$$

技巧 3

利用矩阵的分块运算。

例 3.10

设有分块矩阵

$$M=\begin{pmatrix}I&A\\B&I\end{pmatrix},$$

求 M 可逆的条件并算出 M^{-1}。

Schur 公式

设 $A\in F^{r\times r}$, $B\in F^{r\times(n-r)}$, $C\in F^{(n-r)\times r}$, $D\in F^{(n-r)\times(n-r)}$，并且 A 可逆，则

(1) $\begin{pmatrix}I_r&\mathbf{0}\\-CA^{-1}&I_{n-r}\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A&B\\\mathbf{0}&D-CA^{-1}B\end{pmatrix}$;

(2) $\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}I_r&-A^{-1}B\\\mathbf{0}&I_{n-r}\end{pmatrix}=\begin{pmatrix}A&\mathbf{0}\\C&D-CA^{-1}B\end{pmatrix}$;

(3) $\begin{pmatrix}I_r&\mathbf{0}\\-CA^{-1}&I_{n-r}\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}I_r&-A^{-1}B\\\mathbf{0}&I_{n-r}\end{pmatrix}=\begin{pmatrix}A&\mathbf{0}\\\mathbf{0}&D-CA^{-1}B\end{pmatrix}$。

技巧 4

构造矩阵恒等式，通过矩阵运算直接求出 A^{-1}。

$A^m = 0$（幂零方阵） $\Rightarrow$ $I - A$ 可逆，并且
$(I-A)(I+A+\cdots+A^{m-1}) = I - A^m = I$
$\Rightarrow (I-A)^{-1} = I + A + \cdots + A^{m-1}.$

$A^2 = A$（幂等方阵） $\Rightarrow$ $I + A$ 可逆，并且
$(A+I)(A-2I) = -2I$
$\Rightarrow (I+A)^{-1} = I - \frac{1}{2}A.$

例 3.9

设 n 方阵

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

求 A^{-1}。

（利用技巧 4 求）

定理 3.20 (Cramer 法则)

当系数矩阵 $A = (a_{ij})_{n\times n}$ 的行列式不等于零时，线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\cdots \qquad \cdots \qquad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

存在唯一解 $(x_1, x_2, \ldots, x_n)^{\mathrm{T}} = \left(\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \ldots, \frac{\Delta_n}{\Delta}\right)^{\mathrm{T}}$，其中
$\Delta = \det(A)$，Δ_i 是将 A 的第 i 列换成 $\mathbf{b} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$ 后所得方阵的行列式，$i = 1, 2, \ldots, n$。

例 3.11

设 $t_1, t_2, \ldots, t_n$ 为各不相同的常数，$y_1 = f(t_1), y_2 = f(t_2), \ldots$, $y_n = f(t_n)$，试构造多项式 $l_1(t), l_2(t), \ldots, l_n(t)$ 使得

$$f(t) = y_1 l_1(t) + y_2 l_2(t) + \cdots + y_n l_n(t),$$

对任意次数不超过 $n-1$ 次的多项式 $f(t)$ 都成立。

秩与相抵

秩与相抵的基本定义

定义 3.14

设 A, B 是 $m \times n$ 矩阵，如果存在可逆方阵 P 和 Q 使得

$$B = PAQ,$$

则称 A 和 B **相抵**。

A 与 B 相抵 $\Leftrightarrow$ 可以通过一系列初等变换将 A 化为 B。

等价关系

在集合 M 上定义二元关系 $\sim$，即 $E \subset M \times M$，若其满足性质：

(1) （自反性）$x \sim x$;

(2) （对称性）$x \sim y \Rightarrow y \sim x$;

(3) （传递性）$x \sim y, y \sim z \Rightarrow x \sim z$,

则称 $\sim$ 为集合 M 上的**等价关系**。

第一个等价关系——矩阵相似

等价类

利用等价关系 $\sim$，可以对集合 M 进行分类，

$$\bar{x} = \{y \in M \mid y \sim x\} \subset M,$$

称集合 $\bar{x}$ 为包含元素 x 的等价类。任意元素 $y \in \bar{x}$ 都称为等价类 $\bar{x}$ 的代表元。

命题 3.21

由关系 $\sim$ 确定的等价类的集合 $X = \{\bar{x} \mid x \in M\}$ 是集合 M 的一个划分，即

$$M = \bigcup_{\bar{x} \in X} \bar{x} \text{ 且 } \bar{x} \bigcap \bar{y} = \emptyset, \forall\, \bar{x}, \bar{y} \in X,$$

也就是说 M 是这些子集的不交并。

等价类的基本问题：

(1) 等价类的代表元是什么？（譬如最简形式的代表元）

(2) 等价类的不变量有哪些？（不变量指该等价类中所有元素都相等的量，是判别等价的必要条件，主要用于判别不等价）

(3) 等价类的全系不变量有哪些？（全系不变量指该等价类中所有元素都相等的量，且不同等价类的全系不变量必不相等，是判别等价的充要条件）

定义 3.15

设 A 是 $m \times n$ 矩阵，式（4）中的矩阵 $\mathrm{diag}(I_r, \mathbf{0})$ 称为 A 的*相抵标准形*。整数 r 称为矩阵 A 的*秩*，记为 $\mathrm{rank}(A)$ 或 $r(A)$。若 $r = m$，则 A 称为是*行满秩*的；若 $r = n$，则 A 称为是*列满秩*的。

显然有$0 \leq rank(A) \leq \min\{m, n\}$

定理 3.23

设 $A, B \in F^{m\times n}$，则 A 与 B 相抵 $\Leftrightarrow$ $\operatorname{rank}(A) = \operatorname{rank}(B)$。

定理 3.22

设 A 是 $m \times n$ 矩阵，则 A 相抵于 $m \times n$ 矩阵 $\operatorname{diag}(I_r, \mathbf{0})$，即存在 m 阶可逆方阵 P 和 n 阶可逆方阵 Q 使得

$$PAQ = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{4}$$

其中非负整数 r 由 A 唯一决定。

定理 3.23

设 $A, B \in F^{m\times n}$，则 A 与 B 相抵 $\Leftrightarrow$ $\operatorname{rank}(A) = \operatorname{rank}(B)$。

矩阵的秩是相抵关系下的不变量，且是全系不变量。

定理 3.24

设 A 是 $m \times n$ 矩阵，P, Q 分别是 m, n 阶可逆方阵，则 $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$。

初等变换不改变矩阵的秩

引理 3.1

设 A 是 $m \times n$ 方阵，P, Q 分别是 m, n 阶初等方阵。若 A 的所有 k 阶子式都为零，则 PA 与 AQ 的所有 k 阶子式也为零。

提示：P , Q 是初等变换矩阵可以分别考虑不同的初等变换。

定理 3.25

矩阵 A 有 k 阶非零子式的充分必要条件是 $\operatorname{rank}(A) \geqslant k$。

李炯生《线性代数》对矩阵的秩的定义就是：

定义 $m \times n$ 矩阵 A 的所有非零子式的最高阶数称为矩阵 A 的秩，记为 $\operatorname{rank} A$.

矩阵A的秩等于矩阵A的非零子式的最大阶数，也就是说矩阵A至少有一个非零的r阶子式，且矩阵A的所有$r+1$阶子式都是0，那么就可以说矩阵A的秩等于r.

定理 3.26

矩阵的秩具有以下性质：

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

(2) $\text{rank}(A \vdots B) \geqslant \max\{\text{rank}(A), \text{rank}(B)\}$；

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

(4) $\text{rank}\begin{pmatrix} A & C \\ \mathbf{0} & B \end{pmatrix} \geqslant \text{rank}\begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$；

(5) $\text{rank}(A+B) \leqslant \text{rank}(A \vdots B) \leqslant \text{rank}(A) + \text{rank}(B)$。

性质（1）是显然的（$(PAQ)^T = Q^T A^T P^T$）

性质（2）也同样显然

性质（3）：

不妨设$P_1AQ_1 = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$，$P_2BQ_2 = \begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$故有：

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} & O \\ O & \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} \end{pmatrix}$$

再进行一次初等变换可得$\begin{pmatrix} I_{r+s} & O \\ O & O \end{pmatrix}$

故$rank\begin{pmatrix} A & O \\ O & B \end{pmatrix} = rank(A) + rank(B)$

性质（4）经过同性质（3）的初等变换之后

$\begin{pmatrix} A & C \\ O & B \end{pmatrix}$可变换为$\begin{pmatrix} U_{r+s} & M \\ O & N \end{pmatrix}$其中$U_{r+s}$是一个上三角阵，所以我们可以认为$\begin{pmatrix} U_{r+s} & M \\ O & N \end{pmatrix}$的最小非 0 子式的阶是大于$r+s$的。

对于性质（5）我们容易发现$rank(A+B) \le \max\{rank(A), rank(B)\}$

或者利用$rank(A+B) = rank\begin{pmatrix} A+B & O \\ O & O \end{pmatrix} \le rank\begin{pmatrix} A+B & B \\ O & O \end{pmatrix} = rank(A{:}B) \le$

$rank\begin{pmatrix} A & B \\ O & B \end{pmatrix} = rank\begin{pmatrix} A & O \\ O & B \end{pmatrix} = rank(A) + rank(B)$

（根据秩的定义$rank(A{:}B) \le rank\begin{pmatrix} A & B \\ O & B \end{pmatrix}$是显然的）

秩的计算

技巧 1

利用定义，找矩阵的最高阶非零子式。

例 3.12

计算 n 阶方阵 A 的秩，这里

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

$\det(A)=1+(-1)^{n+1}$，所以n为奇数的时候$\det(A)=2$，此时$rank(A)=n$；当n为偶数的时候$\det(A)=0$，此时取$A_{11}=1$，所以矩阵的秩为$n-1$.

技巧 2

利用初等变换，将矩阵化为上三角阵、下三角阵或对角阵，求出矩阵的秩。

例 3.13

计算 n 阶方阵 A 的秩，这里

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

参见前面对行列式的计算，$rank(A)=n$

例 3.14

每个秩为 r 的矩阵都可以写成 r 个秩是 1 的矩阵之和。

不妨设矩阵A的秩为r，所以$A=P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q$，考虑$C=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q$，显然，$C$只有 r行，

C可以写成r个秩为1的矩阵之和（即截取C的一行成为一个矩阵），所以有

$C=F_1+F_2+\cdots+F_r$其中$F_k,k=1,2,3,\ldots,r$的秩为1，

所以$A=PC=P(F_1+F_2+\cdots+F_r)=PF_1+PF_2+\cdots+PF_r$，其中的$PF_k$的秩都为1

这就完成了证明

例 3.15

若 $m\times n$ 矩阵 A 是列满秩的，则 A 是某个 m 阶可逆方阵的前 n 列。

令$A=\begin{pmatrix} (R_1)_{n\times n} \\ (R_2)_{(m-n)\times n} \end{pmatrix}$，只要取$B=\begin{pmatrix} (R_1)_{n\times n} & O \\ (R_2)_{(m-n)\times n} & I \end{pmatrix}$

容易得到$\det(B)\neq 0$，所以B是一个m阶可逆方阵。

例 3.16

（Frobenius 秩不等式）设矩阵 $A\in F^{m\times n}$，$B\in F^{n\times l}$，$C\in F^{l\times p}$，则 $\operatorname{rank}(ABC)\geqslant\operatorname{rank}(AB)+\operatorname{rank}(BC)-\operatorname{rank}(B)$.

要证明：$rank(ABC)\geq rank(AB)+rank(BC)-rank(B)$，

只要证明$rank(ABC)+rank(B)\geq rank(AB)+rank(BC)$，

$$\begin{pmatrix} ABC & O \\ O & B \end{pmatrix}\to\begin{pmatrix} ABC & AB \\ O & B \end{pmatrix}\to\begin{pmatrix} O & AB \\ -BC & B \end{pmatrix}\to\begin{pmatrix} -BC & B \\ O & AB \end{pmatrix}$$

所以：

$$rank(ABC)+rank(B)=rank\begin{pmatrix} ABC & O \\ O & B \end{pmatrix}=rank\begin{pmatrix} -BC & B \\ O & AB \end{pmatrix}\geq rank(AB)+rank(BC)$$

例 3.17

设矩阵 $A\in F^{m\times n}$，$B\in F^{n\times l}$，证明：

$$\operatorname{rank}(A)+\operatorname{rank}(B)-n\leqslant\operatorname{rank}(AB)\leqslant\min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$$

对于右边不妨设$rank(A)=r, rank(B)=s$，

存在P_1,Q_1,P_2,Q_2，使得$A=P_1\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}Q_1$，$B=P_2\begin{bmatrix} I_s & O \\ O & O \end{bmatrix}Q_2$

故$AB=P_1\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}Q_1P_2\begin{bmatrix} I_s & O \\ O & O \end{bmatrix}Q_2$

不妨记$Q_1P_2=\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$

故$rank(AB)=rank\left(\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}\right)=rank\begin{pmatrix} R_{11} & O \\ O & O \end{pmatrix}=rank(R_{11})$

R_{11}是$r\times s$的矩阵，其秩最大为$\min\{r,s\}$
我们也可以从**矩阵的行秩与列秩**的观点考量：

$$rank(AB)=rank\left(P\begin{bmatrix}I_r & O\\ O & O\end{bmatrix}QB\right)=rank\left(\begin{bmatrix}I_r & O\\ O & O\end{bmatrix}\tilde{B}\right)\le r-(n-s)=r+s-n$$

对于右边我们利用 Frobenius 秩不等式很容易证明

例 3.18

设矩阵 A 满足 $A^2=A$，证明：$\mathrm{rank}(A)=\mathrm{tr}(A)$。

我们可设$A=P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}Q$，因为$A^2=A$，所以$A^2=P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}QP\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}Q$

设$QP=\begin{pmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{pmatrix}$，$\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}QP\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}=\begin{pmatrix}R_{11} & O\\ O & O\end{pmatrix}$

所以$A^2=P\begin{pmatrix}R_{11} & O\\ O & O\end{pmatrix}Q=A=P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}Q$，故$R_{11}=I_r$

$QP=\begin{pmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{pmatrix}=\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}$，故$Q=\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}P^{-1}$

所以$tr(A)=tr\left(P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}P^{-1}P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}P^{-1}\right)$

$$=tr\left(\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}\begin{pmatrix}I_r & R_{12}\\ R_{21} & R_{22}\end{pmatrix}\right)=tr\begin{pmatrix}I_r & R_{12}\\ O & O\end{pmatrix}=tr(I_r)$$

而$tr(I_r)=r=rank(A)$，证毕

例 5.17

设 n 阶方阵 A 满足 $\mathrm{rank}(A+I)+\mathrm{rank}(A-I)=n$，证明 $A^2=I$。

证明：$A^2-I=0\Leftrightarrow(A+I)(A-I)=0$
设$rank(A+I)=r$
$\Longrightarrow(A+I)\vec{x}=\vec{0}$的解空间维数为$n-r$，设$V_1=\{x_{r+1},x_{r+2},\dots,x_n\}$
$\Longrightarrow(A-I)\vec{x}=\vec{0}$的解空间维数为$r$，设$V_2=\{x_1,x_2,\dots,x_r\}$
令$X=(\vec{x_1},\vec{x_2},\dots,\vec{x_n})$

那么我们有$AX=X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}\Rightarrow A=X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}X^{-1}$

$\Longrightarrow A^2=X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}X^{-1}$

$X=\begin{pmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{pmatrix}\Rightarrow X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}=\begin{pmatrix}R_{11} & -R_{12}\\ R_{21} & -R_{22}\end{pmatrix}$

$\Longrightarrow X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}=\begin{pmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{pmatrix}=X$

所以$A^2=X\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}\begin{pmatrix}I_r & O\\ O & -I_{n-r}\end{pmatrix}X^{-1}=XX^{-1}=I$

Chapter Ⅳ 线性空间

线性空间的定义

点、直线、平面、空间 $\Rightarrow$ 一般的 n 维线性空间!

平面上的二维向量可通过 (x, y) 表示，空间中的三维向量可通过 (x, y, z) 表示，那么一般的 n 维线性空间中的向量是否可以通过 $(x_1, x_2, \ldots, x_n)$ 来表示?

n 维线性空间的典型代表：n 维数组空间

定义 4.1

设 F 是数域。定义了加法与数乘运算的 n 维数组向量的全体

$$\{(a_1, a_2, \ldots, a_n) \mid a_i \in F\}$$

称为 *n 维数组空间*，记为 F^n。

F^n + 加法 + 数乘 + 八条运算规律 $\Longrightarrow$ 数域F上的n维线性空间

线性相关与线性无关

含有0向量的任何向量组一定线性相关

定理 4.1

设向量组 $S_1 = \{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$ 是向量组 $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ 的一个子集。那么，如果 S_1 线性相关，则 S 也线性相关；如果 S 线性无关，则 S_1 也线性无关。

从求解方程组的观点出发这个结论是显然的。

定理 4.2

设 $\mathbf{a}_i=(a_{i1},\ldots,a_{in})\in F^n$, $i=1,\ldots,m$，用 A 表示以 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 为行构成的 $m\times n$ 阶矩阵，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性相关 $\Leftrightarrow$ 关于 $\lambda_1,\ldots,\lambda_m$ 的齐次线性方程组 $A^T\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_m\end{pmatrix}=\mathbf{0}$ 有非零解 $\Leftrightarrow \operatorname{rank}(A)<m$。

推论 4.1

设 $\mathbf{a}_1,\cdots,\mathbf{a}_m\in F^n$ 是一组数组向量，则

(1) 若 $m>n$，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 必然线性相关；

(2) 若 $m=n$，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性相关 $\Leftrightarrow \det(A)=0$；

(3) 若 $m<n$，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性相关 $\Leftrightarrow$ 矩阵 A 的所有 m 阶子式为零。

定理 4.3

设 $\mathbf{a}_i=(a_{i1},\ldots,a_{ir})\in F^r$, $i=1,2,\ldots,m$，它们的加长向量组为 $\mathbf{b}_i=(a_{i1},\ldots,a_{ir},\ldots,a_{in})\in F^n$ $(n>r)$, $i=1,2,\ldots,m$，则有

(1) 若 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性无关，则 $\mathbf{b}_1,\ldots,\mathbf{b}_m$ 也线性无关；

(2) 若 $\mathbf{b}_1,\ldots,\mathbf{b}_m$ 线性相关，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 也线性相关。

定理 4.3

设 $\mathbf{a}_i=(a_{i1},\ldots,a_{ir})\in F^r$, $i=1,2,\ldots,m$，它们的加长向量组为 $\mathbf{b}_i=(a_{i1},\ldots,a_{ir},\ldots,a_{in})\in F^n$ $(n>r)$, $i=1,2,\ldots,m$，则有

(1) 若 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性无关，则 $\mathbf{b}_1,\ldots,\mathbf{b}_m$ 也线性无关；

(2) 若 $\mathbf{b}_1,\ldots,\mathbf{b}_m$ 线性相关，则 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 也线性相关。

从求解方程组的观点出发这个结论是显然的。

定义 4.4

设 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in F^n$，若 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 线性无关，且任加一个其它向量 $\mathbf{a}_{i_{r+1}}$ 后 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}, \mathbf{a}_{i_{r+1}}$ 线性相关，则称 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 为 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的极大无关组。

定理 4.4

设 S 是某一向量空间中的向量组，$M = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ 是 S 的线性无关子集，则 M 是 S 的极大线性无关组 $\Leftrightarrow$ S 中所有的向量都是 M 中元素的线性组合。

证明是显然的

定理 4.5

设 $\mathbf{a}_1, \cdots, \mathbf{a}_m \in F^n$ 为一组列向量，$A = (\mathbf{a}_1, \cdots, \mathbf{a}_m)$ 是以 $\mathbf{a}_1, \cdots, \mathbf{a}_m$ 为列构成的 $n \times m$ 阶矩阵。若 A 经一系列的初等行变换变为矩阵 $B = (\mathbf{b}_1, \cdots, \mathbf{b}_m)$，则

(1) $\mathbf{a}_1, \cdots, \mathbf{a}_m$ 线性相关（无关）$\Leftrightarrow$ $\mathbf{b}_1, \cdots, \mathbf{b}_m$ 线性相关（无关）；

(2) $\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_r}$ 为 $\mathbf{a}_1, \cdots, \mathbf{a}_m$ 的极大无关组 $\Leftrightarrow$ $\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_r}$ 为 $\mathbf{b}_1, \cdots, \mathbf{b}_m$ 的极大无关组。

证明(1) 当$(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$线性相关时，考虑方程组$A\boldsymbol{x} = 0$，其中$\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$，显然对$A$做初等行变换不会改变方程组的解，所以$\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$也是方程组$B\boldsymbol{x} = 0$的解，所以当$(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$线性相关可以得到$(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$也线性相关；当$(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$线性无关考虑方程组$A\boldsymbol{x} = 0$只有0解，所以容易得到$rank(A) = m$，由于初等变换不改变矩阵的秩，所以我们容易得到$rank(B) = m$，从而得到$B\boldsymbol{x} = 0$只有0解，所以$(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$也线性无关。

证明(2) 记$A' = (\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_r}), B' = (\boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_r})$根据上述证明可得，$A'$和$B'$的解是完全相同的，$A'\boldsymbol{x} = 0$ 无解 $\Longrightarrow B'\boldsymbol{x} = 0$ 无解所以$(\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_r})$线性无关$\Rightarrow$ $(\boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_r})$线性无关，在$A'$中任意增加一个向量，$A'\boldsymbol{x} = 0$即有非0解，对应的$B'\boldsymbol{x} = 0$也就有非0解，所以$(\boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_r})$是$(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$的极大无关组。

向量组和秩

定义 4.5

如果向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 中的每一个向量都可以用向量组 $\mathbf{b}_1, \ldots, \mathbf{b}_l$ 线性表示，则称向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 可以由向量组 $\mathbf{b}_1, \cdots, \mathbf{b}_l$ 线性表示。如果两个向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 与 $\mathbf{b}_1, \ldots, \mathbf{b}_l$ 可以相互线性表示，则称 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 与 $\mathbf{b}_1, \ldots, \mathbf{b}_l$ 等价，记为 $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \sim \{\mathbf{b}_1, \ldots, \mathbf{b}_l\}$。

向量组等价是等价关系

定理 4.6

一组向量组与它的任何一组极大无关组等价。

推论 4.2

向量组的任何两个极大无关组彼此等价。

$\Rightarrow$ 向量组的极大无关组不唯一，彼此之间相互等价，但是所含向量的个数是否相等呢？容易验证，极大无关组的所含向量个数是相等的

定理 4.7

两个线性无关向量组 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 和 $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 等价，则 $r = s$。

推论 4.3

设 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 和 $\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_s}$ 分别为 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的两个极大无关组，则 $r = s$。

向量组的极大无关组元素的个数是等价关系下的不变量。

定义 4.6

向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的极大无关组元素的个数称为向量组的秩，记为 $\operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_m)$，或 $r(\mathbf{a}_1, \ldots, \mathbf{a}_m)$。

定理 4.8

设向量 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in F^n$，则有

(1) $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性无关 $\Leftrightarrow \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_m) = m$；
(2) $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性相关 $\Leftrightarrow \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_m) < m$；
(3) $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 可以用 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 线性表示 $\Rightarrow$ $\operatorname{rank}(\mathbf{b}_1, \ldots, \mathbf{b}_s) \leq \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_r)$；
(4) $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 与 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 等价 $\Rightarrow$ $\operatorname{rank}(\mathbf{b}_1, \ldots, \mathbf{b}_s) = \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_r)$；
(5) $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 可以用 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 线性表示，且 $\mathbf{b}_1, \ldots, \mathbf{b}_s$ 线性无关 $\Rightarrow s \leq r$；
(6) 向量 $\mathbf{b}$ 可以表示成 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的线性组合 $\Leftrightarrow$ $\operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_m) = \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b})$。

向量组的秩与矩阵的秩之间有什么关系?

$$
A = \begin{array}{c@{}l}
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix}
&
\begin{array}{l}
\leftarrow \mathbf{a}_1 \\
\leftarrow \mathbf{a}_2 \\
\quad \vdots \\
\leftarrow \mathbf{a}_m
\end{array} \\
\begin{array}{cccc}
\uparrow & \uparrow & & \uparrow \\
\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n
\end{array} &
\end{array}
$$

矩阵的行秩 — $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$
矩阵的列秩 — $\operatorname{rank}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$
矩阵的秩 — $\operatorname{rank}(A)$

定理 4.9

任何矩阵的行秩 = 矩阵的列秩 = 矩阵的秩。

子空间和基

定义 4.7

设 $V \subset F^n$ 为非空向量集合，它满足

(1) $\forall \mathbf{a}, \mathbf{b} \in V \;\Rightarrow\; \mathbf{a}+\mathbf{b} \in V$;

(2) $\forall \mathbf{a} \in V, \lambda \in F \;\Rightarrow\; \lambda\mathbf{a} \in V$,

则称 V 为 F^n 的**子空间**。

$\Rightarrow$ 对任意 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in V, \lambda_1, \ldots, \lambda_m \in F$，都有 $\sum\limits_{i=1}^{m} \lambda_i \mathbf{a}_i \in V$ 成立

$\Rightarrow$ 集合 V 对线性运算（加法与数乘）运算封闭!

F^n 的平凡子空间：F^n, $\{\mathbf{0}\}$。

定义 4.8

设 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in F^n$ 是一组向量，称集合

$$\langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle := \left\{ \sum_{i=1}^{m} \lambda_i \mathbf{a}_i \mid \lambda_i \in F, i = 1, \ldots, m \right\}$$

为向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ **生成的子空间**，$\mathbf{a}_1, \ldots, \mathbf{a}_m$ 称为生成子空间的**生成元**。

给定一个矩阵 $A \in F^{m\times n}$，由 A 的行向量生成的子空间称为**行空间**；由 A 的列向量生成的子空间称为列空间。

定理 4.10

设 $\mathbf{a}_i, \mathbf{b}_j, \mathbf{b} \in F^n$，则下列结论成立：

(1) 向量 $\mathbf{b}_1, \ldots, \mathbf{b}_l \in \langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle \Rightarrow \langle \mathbf{b}_1, \ldots, \mathbf{b}_l \rangle \subset \langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle$；

(2) 向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 与 $\mathbf{b}_1, \ldots, \mathbf{b}_l$ 等价 $\Leftrightarrow$ $\langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle = \langle \mathbf{b}_1, \ldots, \mathbf{b}_l \rangle$；

(3) $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性相关 $\Leftrightarrow$ 存在 i 使得 $\langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle = \langle \mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_m \rangle$；

(4) $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 线性无关 $\Leftrightarrow$ 对任意 i 都有 $\langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle \neq \langle \mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_m \rangle$；

(5) $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 是 $\mathbf{a}_1, \ldots, \mathbf{a}_m$ 的极大无关组 $\Leftrightarrow$ $\langle \mathbf{a}_1, \ldots, \mathbf{a}_m \rangle = \langle \mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r} \rangle$ 且 $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ 线性无关；

定理 4.11

设 V 是 F^n 的子空间，则存在线性无关的向量组 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 使得 $V = \langle \mathbf{a}_1, \ldots, \mathbf{a}_r \rangle$。

F^n上的子空间一定是生成子空间。

定义 4.9

设 $V \subset F^n$ 是子空间，V 中一组向量 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 称为 V 的一组基，如果它满足

(1) 对任意向量 $\mathbf{a} \in V$, $\mathbf{a}$ 可唯一表示成 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 的线性组合 $\mathbf{a} = \lambda_1 \mathbf{a}_1 + \cdots + \lambda_r \mathbf{a}_r$；

(2) $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 线性无关，

则称 $(\lambda_1, \ldots, \lambda_r)$ 为向量 $\mathbf{a}$ 在基 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 下的坐标。

定义 4.10

设 $V \subset F^n$ 为子空间，称 V 的一组基的向量个数为 V 的维数，记为 $\dim V = \operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r)$。

定理 4.12

在 F^n 空间下列结论成立：

(1) 设 $V \subset F^n$ 为 r 维子空间 $\Rightarrow V$ 中任意 $r+1$ 个向量线性相关；

(2) 设 $V \subset F^n$ 为 r 维子空间 $\Rightarrow V$ 中任意 r 个线性无关向量为 V 的一组基；

(3) 设 U 与 V 为 F^n 的子空间，且 $U \subseteq V \Rightarrow \dim U \leq \dim V$；

(4) 设 U 与 V 为 F^n 的子空间，且 $U \subseteq V$，$\dim U = \dim V$ $\Leftrightarrow U = V$。

定理 4.13

设 $V \subset F^n$ 为 r 维子空间，$\mathbf{a}_1, \ldots, \mathbf{a}_s \in V$ 是 $s\ (s < r)$ 个线性无关的向量，则存在 V 中的向量 $\mathbf{a}_{s+1}, \ldots, \mathbf{a}_r$ 使得 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 构成 V 的一组基。

$\Rightarrow$ 称 $a_1, \ldots, a_r$ 为线性无关组 $a_1, \ldots, a_s$ 的一组**扩充基**

设线性空间 V 有两组基：$\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$，$\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$

$$\Rightarrow \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_r \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_r \end{pmatrix} \begin{array}{c} \begin{pmatrix} t_{11} & \cdots & t_{1r} \\ \vdots & & \vdots \\ t_{r1} & \cdots & t_{rr} \end{pmatrix} \\ \begin{matrix} \uparrow & & \uparrow \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{matrix} \end{array}$$

记 $T = (t_{ij})_{r \times r}$，称矩阵 T 为从基 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 到基 $\mathbf{b}_1, \ldots, \mathbf{b}_r$ 的过渡矩阵。

含义：T 的第 j 列是向量 $\mathbf{b}_j$ 在基 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 下的坐标

定理 4.14

设 $A \in F^{m \times n}$ 为 $m \times n$ 阶矩阵，$\mathbf{b} \in F^m$ 为 m 维列向量，则线性方程组 $A\mathbf{x} = \mathbf{b}$ 有解 $\Leftrightarrow \operatorname{rank}(A) = \operatorname{rank}(A, \mathbf{b})$。线性方程组有唯一解 $\Leftrightarrow \operatorname{rank}(A) = \operatorname{rank}(A, \mathbf{b}) = n$。

线性方程组和解空间

推论 4.4

齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 一定有解。齐次线性方程组有非零解 $\Leftrightarrow$ $\operatorname{rank}(A)<n$。特别地，若 A 为 n 阶方阵，齐次线性方程组有非零解 $\Leftrightarrow \det(A)=0$。

记齐次线性方程组 $A\boldsymbol{x}=0$ 的解集的全体 $V:=\{\boldsymbol{x}\in F^n \mid A\boldsymbol{x}=0\}$，其中$V$称为齐次线性方程组 $A\boldsymbol{x}=0$的解空间。

定理 4.15

齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 的解集 V 是 F^n 的子空间，并且 $\dim V=n-\operatorname{rank}(A)$。

定义 4.11

齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 的解集 V 称为*解空间*，解空间的一组基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{n-r}$ 称为齐次线性方程组的一个*基础解系*。

齐次线性方程组的通解为：$\mathbf{x}=t_1\boldsymbol{\alpha}_1+\ldots+t_{n-r}\boldsymbol{\alpha}_{n-r}$

例 4.3

设 A 是 n 阶奇异方阵，且 A 的 (i,j) 元素的代数余子式 $A_{ij}\neq 0$。证明：$\boldsymbol{\alpha}=(A_{i1},\ldots,A_{in})^T$ 是齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 的基础解系。

因为A是n阶奇异方阵，所以$\det(A)=0$，而A的代数余子式$A_{ij}\neq 0$，所以$rank(A)=n-1$

所以齐次方程组$A\boldsymbol{x}=0$的解空间只有一维。
而$A\boldsymbol{\alpha}=\sum_{k=1}^{n}a_{kj}A_{ij}=\delta_{ki}\det(A)=0$，所以$\boldsymbol{\alpha}=(A_{i1},\ldots,A_{in})$是齐次方程组$A\boldsymbol{x}=0$的解。由于$\boldsymbol{\alpha}$是非零向量，所以它是基础解系。

例 4.5

设 $A\in F^{n\times n}$ 且满足

$$|a_{ii}|>\sum_{1\leqslant j\neq i\leqslant n}|a_{ij}|,\ i=1,2,\ldots,n.$$

证明：齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 只有零解。

例 4.6

设 n 阶方阵 A 满足 $A^2 = A$。证明:

$$\mathrm{rank}(A) + \mathrm{rank}(I - A) = n。$$

例 4.7

设 A 是任意矩阵，证明：$\mathrm{rank}(A^{\mathrm{T}}A) = \mathrm{rank}(AA^{\mathrm{T}}) = \mathrm{rank}(A)$。

$A\mathbf{x} = \mathbf{b} \Rightarrow \begin{cases} A\mathbf{y} = \mathbf{0} \\ A\mathbf{x}_0 = \mathbf{b} \end{cases} \Rightarrow A(\mathbf{y} + \mathbf{x}_0) = \mathbf{b}$，即齐次线性方程组任意一个解 $\mathbf{y}$ + 非齐次线性方程组的特解 $\mathbf{x}_0 \Rightarrow$ 非齐次线性方程组的一个解；

反之，$\begin{cases} A\mathbf{x} = \mathbf{b} \\ A\mathbf{x}_0 = \mathbf{b} \end{cases} \Rightarrow A(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{y} + \mathbf{x}_0$, 即非齐次线性方程组任意一个解 $\mathbf{x}$ 一定可以分解为齐次线性方程组任意一个解 $\mathbf{y}$ + 非齐次线性方程组的特解 $\mathbf{x}_0$。

一般线性空间

定义 4.12

设 V 是一个非空集合，F 是一个数域。对 V 中的元素定义两种运算

(1) *加法*：对 V 中的任意两个元素 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 组成的有序对 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$，存在 V 中唯一的一个元素 $\boldsymbol{\gamma}$ 与之对应，简记为 $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\gamma}$；

(2) *数乘*：对任意常数 $\lambda \in F$ 及向量 $\boldsymbol{\alpha} \in V$，存在 V 中唯一的一个元素 $\boldsymbol{\gamma}$ 与之对应，简记为 $\lambda\boldsymbol{\alpha} = \boldsymbol{\gamma}$。

加法与数乘运算满足下列运算规律

(A1) $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha}$ 对任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$ 成立;

(A2) $(\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma} = \boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma})$ 对任意 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in V$ 成立;

(A3) 存在元素 $\boldsymbol{\theta} \in V$ 使得 $\boldsymbol{\alpha} + \boldsymbol{\theta} = \boldsymbol{\theta} + \boldsymbol{\alpha} = \boldsymbol{\alpha}$ 对任意 $\boldsymbol{\alpha} \in V$ 成立。$\boldsymbol{\theta}$ 称为零元素。在不致混淆的情况下，一般线性空间中的零元素也常简记为 $\mathbf{0}$;

(A4) 对任意 $\boldsymbol{\alpha} \in V$，存在 $\boldsymbol{\beta} \in V$ 使得 $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha} = \mathbf{0}$。$\boldsymbol{\beta}$ 称为 $\boldsymbol{\alpha}$ 的负元素，简记为 $-\boldsymbol{\alpha}$，并且定义 $\boldsymbol{\beta} - \boldsymbol{\alpha}a = \boldsymbol{\beta} + (-\boldsymbol{\alpha})$;

(D1) $\lambda(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \lambda\boldsymbol{\alpha} + \lambda\boldsymbol{\beta}$ 对任意 $\lambda \in F$ 及 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$ 成立;

(D2) $(\lambda + \mu)\boldsymbol{\alpha} = \lambda\boldsymbol{\beta} + \mu\boldsymbol{\alpha}$ 对任意 $\lambda, \mu \in F$ 及 $\boldsymbol{\alpha} \in V$ 成立;

(M1) $\lambda(\mu\boldsymbol{\alpha}) = (\lambda\mu)\boldsymbol{\alpha}$ 对任意 $\lambda, \mu \in F$ 及 $\boldsymbol{\alpha} \in V$ 成立;

(M2) $1\boldsymbol{\alpha} = \boldsymbol{\alpha}$ 对任意 $\boldsymbol{\alpha} \in V$ 成立,

则称 V 是数域 F 上的*线性空间*，简记为 $V(F)$ 或 V。线性空间 V 中的元素称为*向量*。

定义 4.14

设 V_1, V_2 是数域 F 上两个线性空间，如果存在一一映射 $\sigma: V_1 \to V_2$ 满足

(1) $\sigma(\boldsymbol{x}+\boldsymbol{y}) = \sigma(\boldsymbol{x}) + \sigma(\boldsymbol{y}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in V_1$；

(2) $\sigma(\lambda \boldsymbol{x}) = \lambda \sigma(\boldsymbol{x}), \quad \forall \lambda \in F, \boldsymbol{x} \in V_1$，

则称线性空间 V_1 与 V_2 **同构**，记为 $V_1 \sim V_2$，σ 称为**同构映射**。当 $V_1 = V_2$ 时，称 σ 为**自同构**。

Chapter V 线性变换

线性变换的定义与性质

定义 5.1

设 V, V' 为数域 F 上的两个线性空间，若映射 $\mathcal{A}: V \to V'$ 满足：对任意 $\boldsymbol{x}, \boldsymbol{y} \in V$, $\lambda \in F$，都有

$$\mathcal{A}(\boldsymbol{x}+\boldsymbol{y}) = \mathcal{A}(\boldsymbol{x}) + \mathcal{A}(\boldsymbol{y}), \tag{1}$$

$$\mathcal{A}(\lambda \boldsymbol{x}) = \lambda \mathcal{A}(\boldsymbol{x}), \tag{2}$$

则称 $\mathcal{A}$ 为从线性空间 V 到线性空间 V' 的**线性映射**。特别地，如果 $V' = V$，则称 $\mathcal{A}$ 为线性空间 V 上的一个**线性变换**。

命题 5.1

设 V 是数域 F 上的线性空间，$\mathcal{A}$ 是 V 上的线性变换。$\mathcal{A}$ 具有以下性质

(1) $\mathcal{A}(\mathbf{0})=\mathbf{0}$；

(2) $\mathcal{A}(-\boldsymbol{\alpha})=-\mathcal{A}(\boldsymbol{\alpha}),\ \boldsymbol{\alpha}\in V$；

(3) $\mathcal{A}(\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_n\boldsymbol{\alpha}_n)=\lambda_1\mathcal{A}(\boldsymbol{\alpha}_1)+\cdots+\lambda_n\mathcal{A}(\boldsymbol{\alpha}_n)$；

(4) 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 为线性空间 V 的一组基，则
$\boldsymbol{\alpha}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_n\boldsymbol{\alpha}_n\Rightarrow\mathcal{A}(\boldsymbol{\alpha})=\lambda_1\mathcal{A}(\boldsymbol{\alpha}_1)+\cdots+\lambda_n\mathcal{A}(\boldsymbol{\alpha}_n)$；

(5) 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为 V 中线性相关的向量，则
$\mathcal{A}(\boldsymbol{\alpha}_1),\mathcal{A}(\boldsymbol{\alpha}_2),\cdots,\mathcal{A}(\boldsymbol{\alpha}_m)$ 也线性相关。

线性变换的矩阵

设 V 为 n 维线性空间，$\mathcal{A}:V\to V$ 为线性变换，在 V 中取定一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$

$$\Rightarrow\mathcal{A}(\boldsymbol{\alpha}_i)\in V,\ i=1,2,\ldots,n$$

$$\Rightarrow\begin{cases}\mathcal{A}(\boldsymbol{\alpha}_1)=a_{11}\boldsymbol{\alpha}_1+a_{21}\boldsymbol{\alpha}_2+\cdots+a_{n1}\boldsymbol{\alpha}_n\\ \mathcal{A}(\boldsymbol{\alpha}_2)=a_{12}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{n2}\boldsymbol{\alpha}_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathcal{A}(\boldsymbol{\alpha}_n)=a_{1n}\boldsymbol{\alpha}_1+a_{2n}\boldsymbol{\alpha}_2+\cdots+a_{nn}\boldsymbol{\alpha}_n\end{cases}$$

$$\Rightarrow(\mathcal{A}(\boldsymbol{\alpha}_1),\ldots,\mathcal{A}(\boldsymbol{\alpha}_n))=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\qquad(3)$$

定义 5.2

设 $\mathcal{A}: V(F) \to V(F)$ 为 n 维线性空间 $V(F)$ 上的线性变换，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 为 $V(F)$ 的一组基。如果数域 F 上的 n 阶方阵 $A = (a_{ij})_{n\times n}$ 满足

$$(\mathcal{A}(\boldsymbol{\alpha}_1), \ldots, \mathcal{A}(\boldsymbol{\alpha}_n)) = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)A,$$

则称方阵 A 为线性变换 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 下的**表示矩阵**，简称**矩阵**。

容易看出：A的第i列为向量$\mathcal{A}(\boldsymbol{\alpha}_i)$在基$\alpha_1, \alpha_2, \ldots, \alpha_n$下的坐标，$i = 1,2,\ldots,n$

定理 5.2

设线性变换 $\mathcal{A}: V \to V$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 下的矩阵为 A。设 $\boldsymbol{x}, \boldsymbol{y} \in V$ 且 $\boldsymbol{y} = \mathcal{A}(\boldsymbol{x})$，若 $\boldsymbol{x}, \boldsymbol{y}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 下的坐标分别为 $X, Y \in F^n$，则 $Y = AX$。

定理 5.3

设 V 为数域 F 上的 n 维线性空间，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 为 V 的一组基，则存在一一映射 $\Phi: L(V) \to M_n(F)$，使得 $\forall \mathcal{A} \in L(V)$，$\Phi(\mathcal{A})$ 为 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 下的矩阵。

$L(V)$：数域 F 上 n 维线性空间 V 上的全体线性变换所构成的集合
$M_n(F)$：数域 F 上的 n 阶方阵构成的集合

定理 5.4

设 $\Phi: L(V) \to M_n(F)$ 为定理5.3中定义的映射，则

(1) $\Phi(\mathcal{A} + \mathcal{B}) = \Phi(\mathcal{A}) + \Phi(\mathcal{B})$;
(2) $\Phi(\lambda\mathcal{A}) = \lambda\Phi(\mathcal{A})$;
(3) $\Phi(\mathcal{B} \circ \mathcal{A}) = \Phi(\mathcal{B}) \cdot \Phi(\mathcal{A})$,

对 $\forall \mathcal{A}, \mathcal{B} \in L(V)$，$\lambda \in F$ 成立。(1) 与 (2) $\Rightarrow$ Φ 为线性同构映射。

同一线性变换在不同基下的表示之间有什么关系?

设线性空间 V 有两组基: $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$,

线性变换 $\mathcal{A}$ 在这两组基下的表示分别为矩阵 A 与 B

$\Rightarrow (\mathcal{A}(\boldsymbol{\alpha}_1),\ldots,\mathcal{A}(\boldsymbol{\alpha}_n))=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)A$

$(\mathcal{A}(\boldsymbol{\beta}_1),\ldots,\mathcal{A}(\boldsymbol{\beta}_n))=(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)B$

而 $(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)T$,从而有

$\Rightarrow (\mathcal{A}(\boldsymbol{\beta}_1),\ldots,\mathcal{A}(\boldsymbol{\beta}_n))=\mathcal{A}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)=\mathcal{A}[(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)T]$

$=[\mathcal{A}(\boldsymbol{\alpha}_1),\ldots,\mathcal{A}(\boldsymbol{\alpha}_n)]T=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)AT$

$\Rightarrow (\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)B=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)AT$

$\Rightarrow (\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)TB=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)AT$

$\Rightarrow TB=AT\Rightarrow B=T^{-1}AT$

定理 5.5

设线性变换 $\mathcal{A}:V\to V$ 在 V 的两组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$ 下的矩阵分别为 A 和 B。设基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$ 的过渡矩阵为 T,即 $(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)T$ 则 $B=T^{-1}AT$。

定义 5.3

设 A,B 为数域 F 上的两个 n 阶方阵,如果存在数域 F 上的 n 阶可逆方阵 T,使得 $B=T^{-1}AT$,则称 A 与 B(在数域 F 上)**相似**,记为 $A\sim B$。

命题 5.6

矩阵的相似关系为等价关系,即满足以下三个条件

(1) (反身性) $A\sim A$;

(2) (对称性) $A\sim B\Rightarrow B\sim A$;

(3) (传递性) $A\sim B$, $B\sim C\Rightarrow A\sim C$。

矩阵的相似

- 代数上：一个线性变换在不同基下的矩阵是相似的，反过来属于该相似类的所有方阵，都是该线性变换在不同基下对应的矩阵。
- 几何上：一个线性空间上的线性变换的性质与该空间的基的选取没有关系 ⇒ 相似的矩阵都具有一些相同的性质，即相似不变量。

特征值和特征向量

一个线性变换在不同基下的矩阵是相似的，选取适当的基可使线性变换的矩阵变得简单，已知一个方阵，如何找到一个尽量简单的方阵与之相似呢？

问题：矩阵相似于对角矩阵的条件?

设 $A \sim \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \Rightarrow$ 存在 n 阶可逆方阵 X，使得 $A = T\Lambda T^{-1}$。记 $T = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n)$，则

$$AT = T\Lambda \Rightarrow A(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n) = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n)\Lambda = (\lambda_1\boldsymbol{t}_1, \ldots, \lambda_n\boldsymbol{t}_n)$$

$$\Rightarrow A\boldsymbol{t}_i = \lambda_i\boldsymbol{t}_i,\ i = 1, 2, \ldots, n$$

关键：找到 n 个满足 $A\boldsymbol{x} = \lambda\boldsymbol{x}$ 的数 $\lambda_1, \ldots, \lambda_n$ 和向量 $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$。

定义 5.4

设 A 为数域 F 上的 n 阶方阵，如果存在 $\lambda \in F$ 及非零列向量 $\mathbf{x} \in F^n$，使得

$$A\boldsymbol{x} = \lambda\boldsymbol{x},$$

则称 λ 为方阵 A 的一个**特征值**，而称 $\boldsymbol{x}$ 为属于特征值 λ 的一个**特征向量**。

几何解释：向量$\boldsymbol{x}$ 在线性变换A下保持方向不变（相同或相反），长度伸缩λ倍。线性变换$\mathcal{A}$的特征值与特征向量也就是矩阵 A 的特征值与特征向量。

命题 5.7

特征向量有如下性质：

(1) 若 α 是线性变换 $\mathcal{A}$ 属于特征值 λ 的特征向量 $\Rightarrow \mu\alpha$ 亦是线性变换 $\mathcal{A}$ 属于特征值 λ 的特征向量，$\forall \mu \neq 0 \in F$；

(2) 若 α 与 β 是线性变换 $\mathcal{A}$ 属于特征值 λ 的特征向量 $\Rightarrow \alpha + \beta$ 亦是线性变换 $\mathcal{A}$ 属于特征值 λ 的特征向量；

(3) 线性变换 $\mathcal{A}$ 属于不同特征值的特征向量线性无关。

(1) 和(2)是显然的，下面给出(3)的证明（反证法）：
设$\lambda_1,\lambda_2,\dots,\lambda_s$是$\mathcal{A}$的特征值$s$个不同的特征值，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s$分别是属于$\lambda_1,\lambda_2,\dots,\lambda_s$的特征向量。假设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s$线性相关。不妨设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s$的线性无关组为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_r$，那么存在$k_1,k_2,\dots,k_r,k_{r+1}$，使得：

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{r+1}\boldsymbol{\alpha}_{r+1}=0 \qquad *$$

*式左右两边同时乘上$\mathcal{A}$得

$$k_1\lambda_1\boldsymbol{\alpha}_1+k_2\lambda_2\boldsymbol{\alpha}_2+\cdots+k_{r+1}\lambda_{r+1}\boldsymbol{\alpha}_{r+1}=0 \qquad \#$$

*式左右两边同时乘上λ_{r+1}得

$$k_1\lambda_{r+1}\boldsymbol{\alpha}_1+k_2\lambda_{r+1}\boldsymbol{\alpha}_2+\cdots+k_{r+1}\lambda_{r+1}\boldsymbol{\alpha}_{r+1}=0 \qquad **$$

$\# - **$可以得到

$$k_1(\lambda_1-\lambda_{r+1})\boldsymbol{\alpha}_1+k_2(\lambda_2-\lambda_{r+1})\boldsymbol{\alpha}_2+\cdots+k_{r+1}(\lambda_r-\lambda_{r+1})\boldsymbol{\alpha}_r=0$$

因为$\lambda_1,\lambda_2,\dots,\lambda_s$是$\mathcal{A}$的特征值$s$个不同的特征值，且$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_r$那么我容易得到$k_1=k_2=\cdots=k_r=0$，所以$k_{r+1}=0$这和$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1}$线性相关矛盾。
证毕。

特征子空间

设 λ 是方阵 A 的特征值，引入

$$V_A(\lambda)=\{\boldsymbol{x}\in F^n\,|\,A\boldsymbol{x}=\lambda\boldsymbol{x}\}.$$

易知 $V_A(\lambda)$ 是 F^n 的子空间，称为矩阵 A 的属于特征值 λ 的**特征子空间**。

特征子空间：特征向量 + 零向量

λ 为方阵 A 的特征值
$\Leftrightarrow$ 齐次线性方程组 $(\lambda I-A)\boldsymbol{x}=0$ 有非零解
$\Leftrightarrow$ $\det(\lambda I-A)=0$

定义 5.6

设 A 是数域 F 上的 n 阶方阵，$\lambda\in F$，称 $\det(\lambda I-A)$ 为矩阵 A 的**特征多项式**，记为 $p_A(\lambda)$。

求解特征值与特征向量的算法：

(1) 计算特征多项式 $p_A(\lambda) = \det(\lambda I - A)$；

(2) 计算 $p_A(\lambda)$ 的根 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 及重数 $n_1, n_2, \ldots, n_s$，即

$$p_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s};$$

(3) 对每个特征值 λ_i $(i = 1, 2, \ldots, s)$，求齐次线性方程组

$$(\lambda_i I - A)\boldsymbol{x} = \mathbf{0}$$

的基础解系：$\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{im_i}$，即 $V_{\lambda_i} = \langle \boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{im_i} \rangle$。

例 5.9

设 $\boldsymbol{x}$ 是方阵 A 的属于 λ 的特征向量，则 $\boldsymbol{x}$ 也是 kA，A^2，$aA + bI$，A^m，$f(A)$，A^{-1}，A^* 分别属于于特征值 $k\lambda$，λ^2，$a\lambda + b$，λ^m，$f(\lambda)$，λ^{-1}，$\frac{|A|}{\lambda}$ 的特征向量，其中 $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$。

证明是显然的。

命题 5.8

相似的矩阵具有相同的特征多项式和特征值。

若 $A \sim B \Rightarrow$ 存在可逆方阵 T，使得 $B = T^{-1}AT$

$$\begin{aligned} \text{设 } p_A(\lambda) = \det(\lambda I - A) \Rightarrow p_B(\lambda) &= \det(\lambda I - B) = \det(\lambda I - T^{-1}AT) \\ &= \det[T^{-1}(\lambda I - A)T] = \det(\lambda I - A) \\ &= p_A(\lambda) \end{aligned}$$

特征多项式和特征值是相似不变量，但不为全系不变量

例 5.10

设 $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$，则 A 与 B 的特征值相同，但 A 与 B 不相似。

（特征向量不是相似不变量）。

关于特征多项式有

$$
\begin{aligned}
\text{记 } p_A(\lambda) &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \\
&= \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n \\
&= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\
\Rightarrow c_1 &= -\sum_{i=1}^{n} a_{ii},\ c_n = (-1)^n \det(A).
\end{aligned}
$$

更一般地我们有

$$
p_A(\lambda) = \det(\lambda I - A) = \lambda^n - \sum_{i=1}^{n} a_i\,\lambda^{n-1} + \Big(\sum_{1\le i_1<i_2\le n} \left|A\begin{pmatrix} i_1 i_2 \\ i_1 i_2 \end{pmatrix}\right|\Big) + \cdots + (-1)^k\Big(\sum_{1\le i_1<i_2<\cdots<i_k\le n} \left|A\begin{pmatrix} i_1 i_2 \ldots i_k \\ i_1 i_2 \ldots i_k \end{pmatrix}\right|\Big) + \cdots + (-1)^n \det(A)
$$

命题 5.9

设 $A = (a_{ij})$ 为数域 $\mathbb{C}$ 上的一个 n 阶方阵，$\lambda_1, \lambda_2, \ldots, \lambda_n$ 为 A 的 n 个特征值，则

(1) $\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$；

(2) $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$。

证明(1)：考虑矩阵的特征多项式$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$
容易发现其中λ^{n-1}的系数为$\sum_{i=1}^{n} \lambda_i$，考虑到：

$$
p_A(\lambda) = \det(\lambda I - A) = \lambda^n - \sum_{i=1}^{n} a_i\,\lambda^{n-1} + \Big(\sum_{1\le i_1<i_2\le n} \left|A\begin{pmatrix} i_1 i_2 \\ i_1 i_2 \end{pmatrix}\right|\Big) + \cdots + (-1)^k\Big(\sum_{1\le i_1<i_2<\cdots<i_k\le n} \left|A\begin{pmatrix} i_1 i_2 \ldots i_k \\ i_1 i_2 \ldots i_k \end{pmatrix}\right|\Big) + \cdots + (-1)^n \det(A)
$$

根据λ^{n-1}的系数相同，所以$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_i$，即$tr(A) = \sum_{i=1}^{n} \lambda_i$

证明(2)：在$p_A(\lambda) = \lambda^n - \sum_{i=1}^{n} a_i\,\lambda^{n-1} + \left(\sum_{1\le i_1<i_2\le n} \left|A\begin{pmatrix} i_1 i_2 \\ i_1 i_2 \end{pmatrix}\right|\right) + \cdots + (-1)^k\left(\sum_{1\le i_1<i_2<\cdots<i_k\le n} \left|A\begin{pmatrix} i_1 i_2 \ldots i_k \\ i_1 i_2 \ldots i_k \end{pmatrix}\right|\right) + \cdots + (-1)^n \det(A)$令$\lambda = 0$，就可以得到
$p_A(0) = (-1)^n \det(A) = (-\lambda_1)(-\lambda_2)\ldots(-\lambda_n) = (-1)^n\lambda_1\lambda_2\ldots\lambda_n$
所以$\det(A) = \lambda_1\lambda_2\ldots\lambda_n$

推论 5.1

n 阶方阵 A 可逆 $\Leftrightarrow$ A 的 n 个特征值均不为零。

Cayley-Hamilton 定理

设 $A \in F^{n\times n}$ 的特征多项式为 $p_A(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n$，
则 $p_A(A) = A^n + c_1A^{n-1} + \cdots + c_nI = \mathbf{0}$。

证明：
首先容易得到下面的引理：

以F[λ]中的多项式为元素的矩阵称为λ-矩阵，记为A(λ)，B(λ)，…，例如

$$A(\lambda) = \begin{bmatrix} \lambda^3+\lambda^2-1 & \lambda^2-3 \\ 2\lambda+4 & 2\lambda^3+1 \end{bmatrix}$$

是一个λ-矩阵，按照矩阵加法，数量乘法的定义，A(λ)可以写成

$$\begin{aligned} A(\lambda) &= \begin{bmatrix} \lambda^3 & 0 \\ 0 & 2\lambda^3 \end{bmatrix} + \begin{bmatrix} \lambda^2 & \lambda^2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2\lambda & 0 \end{bmatrix} + \begin{bmatrix} -1 & -3 \\ 4 & 1 \end{bmatrix} \\ &= \lambda^3\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + \lambda^2\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \lambda\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} -1 & -3 \\ 4 & 1 \end{bmatrix} \end{aligned}$$

其中λ^i的系数都是F上的矩阵，显然这种表示法是惟一的。一般地，一个m×n的λ-矩阵A(λ)可以惟一表示成多项式：

$A(\lambda)=\lambda^m A_0+\lambda^{m-1} A_1+\ldots+\lambda A_{m-1}+A_m$，（1）

其中 $A_0,\ldots,A_m$ 是数域F上m×n矩阵，(1)式右端称为矩阵多项式，若$A_0\neq 0$,m称为矩阵多项式的次数。

设$B^*(\lambda)$是矩阵$\lambda I - A$的伴随矩阵，那么容易得到

$$B^*(\lambda)(\lambda I - A) = p_A(\lambda)I$$

考虑到$B^*(\lambda)$的元素都是$\lambda I - A$的代数余子式，所以$B^*(\lambda)$的元素是次数不超过$n-1$的λ的多项式。
那么$B^*(\lambda)$可以写成：

$$B^*(\lambda) = \lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + \lambda B_{n-2} + B_{n-1}$$

所以

$$\begin{aligned} B^*(\lambda)(\lambda I - A) &= (\lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + \lambda B_{n-2} + B_{n-1})(\lambda I - A) \\ &= \lambda^n B_0 + \lambda^{n-1}(B_1 - B_0A) + \cdots + \lambda(B_{n-1} - B_{n-2}A) - B_{n-1}A \quad * \end{aligned}$$

设$p_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$，所以有

$$p_A(\lambda)I = \lambda^n I + a_1\lambda^{n-1}I + \cdots + a_nI \quad \#$$

比较*和#容易得到

$$\begin{cases} B_0 = I, \\ B_1 - B_0A = a_1I, \\ B_2 - B_1A = a_2I, \\ \vdots \\ B_{n-1} - B_{n-2}A = a_{n-1}I, \\ -B_{n-1}A = a_nI. \end{cases} \quad **$$

我们依次用$A^n, A^{n-1}, \ldots, A, I$右乘上**式的第1,2,..,$n+1$式，可以得到：

$$\begin{cases} B_0A^n = A^n, \\ B_1A^{n-1} - B_0A^n = a_1A^{n-1}, \\ B_2A^{n-2} - B_1A^{n-1} = a_2A^{n-2}, \\ \vdots \\ B_{n-1}A - B_{n-2}A^2 = a_{n-1}A, \\ -B_{n-1}A = a_nI. \end{cases}$$

上式两端分别相加，左边为0，右边为$f(A)$.
这就完成了证明。

命题 5.10

设 $A, B \in F^{n\times n}$，且 $A \sim B$，则

(1) $A^{\mathrm{T}} \sim B^{\mathrm{T}}$，$A^{-1} \sim B^{-1}$（若 A，B 均可逆），$A^* \sim B^*$；

(2) $A^k \sim B^k$；

(3) $p_A(\lambda) = p_B(\lambda)$；

(4) $\mathrm{rank}(A) = \mathrm{rank}(B)$；

(5) $\det(A) = \det(B)$；

(6) $\mathrm{tr}(A) = \mathrm{tr}(B)$。

证明是显然的。

例 5.12

设 n 阶方阵 A 的 n 个特征值分别为 $\lambda_1, \lambda_2, \ldots, \lambda_n$，求 $I + A$ 的特征值及 $\det(I + A)$。

容易验证$I + A$的特征值恰好是A的特征值加1，所以$\det(I + A) = \prod_{i=1}^{n}(\lambda_i + 1)$

例 5.13

设方阵 A 满足 $A^k = \mathbf{0}$，证明：$\det(I - A) = 1$。

例 5.14

设 A 为 n 阶实矩阵满足 $AA^{\mathrm{T}} = I$，且 $|A| < 0$，试求 A 的伴随矩阵 A^* 的一个特征值。

矩阵的相似对角化

定义 5.7

如果一个方阵相似于对角阵，则称该方阵**可对角化**，也称相应的线性变换**可对角化**。

定理 5.11

数域 F 上的 n 阶方阵 A 相似于对角阵 $\Leftrightarrow A$ 有 n 个线性无关的特征向量。

充分性需要考量，只有这n个向量线性无关，它们组成的矩阵（记为P）就可逆，从而可以使得$P^{-1}AP = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$
必要性容易得到$AP = Pdiag(\lambda_1, \lambda_2, \dots, \lambda_n)$

我们能不能对相似对角化条件进行更细致的刻画？

代数重数

设 $A \in F^{n\times n}$ $(F = \mathbb{C})$，A 的特征多项式

$$p_A(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s} \ (n_1 + \cdots + n_s = n),$$

其中 $\lambda_1, \dots, \lambda_s$ 为 A 的所有不同特征值，则称 n_i 为特征值 λ_i 的**代数重数**。

几何重数

特征值 λ_i 的特征子空间 $V_A(\lambda_i)$ 的维数，即线性方程组 $(\lambda_i I - A)\boldsymbol{x} = \mathbf{0}$ 的解空间的维数，称为特征值 λ_i 的**几何重数**。

引理 5.1

设 $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{im_i}$ $(i = 1, 2, \dots, s)$ 是 A 的属于特征值 λ_i 的线性无关的特征向量组，则 $\boldsymbol{x}_{11}, \boldsymbol{x}_{12}, \dots, \boldsymbol{x}_{1m_1}$, $\boldsymbol{x}_{21}, \boldsymbol{x}_{22}, \dots, \boldsymbol{x}_{2m_2}$, $\dots$, $\boldsymbol{x}_{s1}, \boldsymbol{x}_{s2}, \dots, \boldsymbol{x}_{sm_s}$ 也是线性无关的向量组。

证明：

对于任意的$l_{i_1}, l_{i_2}, \dots, l_{i_{m_i}}$我们可以认为$l_{i_1}\boldsymbol{x}_{i_1} + l_{i_2}\boldsymbol{x}_{i_2} + \cdots + l_{i_{m_i}}\boldsymbol{x}_{i_{m_i}} = \boldsymbol{x}_i$，那么我们容易得到

$\boldsymbol{x}_i$是属于λ_i的特征向量，由命题 5.7-(3)知 $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_s$线性无关，那么不存在不全为0的实数组$\mu_1, \mu_2, \dots, \mu_s$使得$\mu_1\boldsymbol{x}_1 + \mu_2\boldsymbol{x}_2 + \cdots + \mu_s\boldsymbol{x}_s = 0$，所以也就不存在不全部为0的实数组$k_{1_1}, \dots, k_{1_{m_1}}, \dots k_{s_1}, \dots, k_{s_{m_s}}$使得$k_{1_1}\boldsymbol{x}_{1_1} + \cdots + k_{1_{m_1}}\boldsymbol{x}_{1_{m_1}} + \cdots + k_{s_1}\boldsymbol{x}_{s_1} + \cdots + k_{s_{m_s}}\boldsymbol{x}_{s_{m_s}} = 0$.

所以$\boldsymbol{x}_{1_1}, \dots, \boldsymbol{x}_{1_{m_1}}, \dots, \boldsymbol{x}_{s_1}, .., \boldsymbol{x}_{s_{m_s}}$线性无关。

引理 5.2

设 λ_i 为 n 阶复方阵 A 的特征值，则它的几何重数不超过它的代数重数，即 $m_i \leqslant n_i$。

证明 设 $\dim V_{\lambda_0} = q$. 取 V_{λ_0} 的基 $\alpha_1, \dots, \alpha_q$, 扩充为 V 的基 $\alpha_1, \dots, \alpha_q, \dots, \alpha_n$. 注意到

$$\mathcal{A}\alpha_i = \lambda_0\alpha_i, \quad i = 1, \dots, q,$$

于是 $\mathcal{A}$ 在基 $\alpha_1, \dots, \alpha_q, \dots, \alpha_n$ 下的矩阵为

$$A = \begin{pmatrix} \lambda_0 & \cdots & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \lambda_0 & * & \cdots & * \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix} = \begin{pmatrix} \lambda_0 E_q & A_{12} \\ O & A_{22} \end{pmatrix}.$$

由此计算即得 $\mathcal{A}$ 的特征多项式

$$|\lambda E - A| = \begin{vmatrix} (\lambda - \lambda_0)E_q & -A_{12} \\ O & \lambda E_{n-q} - A_{22} \end{vmatrix} = (\lambda - \lambda_0)^q |\lambda E_{n-q} - A_{22}|,$$

即得 λ_0 的代数重数 $\geqslant q$.

定理 5.12

复方阵 A 可对角化 $\Leftrightarrow$ A 的每个特征值的几何重数与代数重数相等，即 $m_i = n_i \ (i = 1, 2, \dots, s)$。

证明：如果n阶复方阵A可对角化，那么A一定会有n个线性无关的特征向量，这就要求$\sum_{i=1}^{s} m_i = n$，根据引理 5.2 可知，任何一个特征值的代数重数不超过它的几何重数，即对每一个m_i都有$m_i \le n_i$，而$\sum_{i=1}^{s} n_i = n$，所以$\sum_{i=1}^{s} m_i = n$只有在每一个m_i都等于n_i时才能够成立，证毕。

例 5.17

设 n 阶方阵 A 满足 $\operatorname{rank}(A + I) + \operatorname{rank}(A - I) = n$，证明 $A^2 = I$。

定理 5.13

任何一个 n 阶复方阵 A 都可以相似于一个上三角阵，即

$$A \sim \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 A 的所有特征值。

证明：（利用数学归纳法）

显然，$n=1$ 时成立，那么只需证明假设$n=k-1$ 成立时，有$n=k$ 成立，即证明该定理。

设：

$$P_1 = (x_1, x_2, \cdots, x_k)$$

其中$x_1, x_2, \cdots, x_k$线性无关且x_1是矩阵特征值为λ_1时的特征向量（对其他向量并无要求），那么可得：

$$AP_1 = (Ax_1, Ax_2, \cdots, Ax_k) = ((\lambda_1 x_1, Ax_2, \cdots, Ax_k))$$

以矩阵形式表示可得：

$$\begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1k} \\ 0 & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{k1} & \cdots & b_{kk} \end{bmatrix}$$

则：

$$P_1^{-1}AP_1 = \begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1k} \\ 0 & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{k1} & \cdots & b_{kk} \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix}$$

由假设，矩阵为k-1阶时结论成立，那么对于矩阵A_1，存在可逆矩阵Q，使得：

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & * & \cdots & * \\ & \lambda_3 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_k \end{bmatrix}$$

那么:

$$P_2 = \begin{bmatrix} 1 & 0^T \\ 0 & Q \end{bmatrix}, S = P_1 P_2$$

有:

$$S^{-1}AS = P_2^{-1}P_1^{-1}AP_1P_2 = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_k \end{bmatrix}$$

证毕

但是我们容易发现，这样的上三角阵并不唯一，也就是说这个上三角阵不能够作为相似等价类的代表元。

例 5.18

求与矩阵 $A = \begin{pmatrix} 2 & -1 & 1 \\ 2 & 2 & -1 \\ 1 & 2 & -1 \end{pmatrix}$ 相似的上三角阵。

例 5.19

设 x, y, z 都是 t 的函数，求解线性微分方程组

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = 2x - y + z, \\ \dfrac{\mathrm{d}y(t)}{\mathrm{d}t} = 2x + 2y - z, \\ \dfrac{\mathrm{d}z(t)}{\mathrm{d}t} = x + 2y - z. \end{cases}$$

那么相似等价类的代表元是什么呢？
——若当标准型

定义 5.8

设 λ 是任意复数，m 是任意正整数，形如

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ & \lambda & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}_{m\times m}$$

的 m 阶方阵称为*若尔当块*，记作 $J_m(\lambda)$，其中 λ 是对角线元素，也是特征值。因此 $J_m(\lambda)$ 也称为特征值为 λ 的 m 阶若尔当块。

定义 5.9

如果一个方阵是准对角阵，并且每个对角块都是若当块，则称之为*若当形矩阵*。

定理 5.14

任何一个复方阵 A 都相似于一个若当形矩阵 J，即

$$A\sim J=\begin{pmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s) \end{pmatrix},$$

其中 $J_i=\mathrm{diag}(J_{m_{i1}}(\lambda_i),\dots,J_{m_{ik_i}}(\lambda_i))$，$\lambda_1,\lambda_2,\dots,\lambda_s$ 是 A 的 s 个不同特征值。如果不计若当块的排列顺序，则 J 是唯一的。

考虑一个若当块$J_m(\lambda)=\begin{vmatrix} \lambda & 1 & 0 & \dots & 0 \\ & \lambda & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ & & & & \lambda \end{vmatrix}$，$J_m(\lambda)-\lambda I=J_m(0)=\begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ & & & & 0 \end{vmatrix}$

$\left(J_m(0)\right)^2=\begin{vmatrix} 0 & 0 & 1 & \dots & 0 \\ & 0 & 0 & \ddots & \vdots \\ & & \ddots & \ddots & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{vmatrix}$，$\left(J_m(0)\right)^{m-1}=\begin{vmatrix} 0 & 0 & 0 & \dots & 1 \\ & 0 & 0 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{vmatrix}$

从上式不难看出$(J_m(\lambda)-\lambda I)^k$中的$k$每增加1，$(J_m(\lambda)-\lambda I)^k$中1所在列就向右上方移动一位，矩阵的秩就减少1

求复数域 $\mathbb{C}$ 上的 n 阶方阵 A 的若尔当标准形的算法:

(1) 求出 A 的特征多项式和全部特征值:
$p_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, $\sum_{i=1}^{s} n_i=n$;

(2) 对每个特征值 λ_i, 计算序列 $A-\lambda_i I,(A-\lambda_i I)^2,(A-\lambda_i I)^3,\ldots,$
记 $r_k=\operatorname{rank}(A-\lambda_i I)^k$，$k\geqslant 0$，约定 $r_0=n$，
$d_k=r_{k-1}-r_k$，$k\geqslant 1$，
$\delta_k=d_k-d_{k+1}$，$k\geqslant 1$，
则 $d_k=J$ 中特征值为 λ_i 的阶大于等于 k 的若尔当块的个数，
$\delta_k=J$ 中特征值为 λ_i 的阶等于 k 的若尔当块的个数;

(3) 依据 δ_k，$k\geqslant 1$，写出 A 的若尔当标准形。

例 5.21

证明：n 阶方阵 A 满足 $A^2=A$ 的充要条件是 A 相似于准对角阵 $\operatorname{diag}(I_r,0)$，这里 $r=\operatorname{rank}(A)$。

证明：

例 5.22

设 x,y,z 都是 t 的函数，求解线性微分方程组

$$\begin{cases}\frac{\mathrm{d}x(t)}{\mathrm{d}t}=-4x+9y-4z,\\ \frac{\mathrm{d}y(t)}{\mathrm{d}t}=-9y+18y-8z,\\ \frac{\mathrm{d}z(t)}{\mathrm{d}t}=-15x+29y-13z.\end{cases}$$

Chapter Ⅵ 欧几里得空间

定义和基本性质

线性空间：集合 + 加法与数乘 + 八条运算规律

$\mathbb{R}^3$：有度量（长度、夹角等）的线性空间 $\Rightarrow$ n 维的有度量线性空间，即 n 维欧几里德空间

关键问题：如何将三维空间的度量推广到一般的 n 维线性空间？

定义 6.1

设 V 是实数域 $\mathbf{R}$ 上的线性空间，如果 V 中任意两个向量 $\mathbf{a}$ 和 $\mathbf{b}$ 都按某一法则对应于一个实数，记作 $(\mathbf{a},\mathbf{b})$，且满足

(1) **对称性**：$\forall \mathbf{a},\mathbf{b}\in V$，有 $(\mathbf{a},\mathbf{b})=(\mathbf{b},\mathbf{a})$;

(2) **线性性**：$\forall \lambda\in\mathbb{R},\mathbf{a},\mathbf{b},\mathbf{c}\in V$，有

$$(\lambda\mathbf{a},\mathbf{b})=\lambda(\mathbf{a},\mathbf{b}),$$
$$(\mathbf{a}+\mathbf{b},\mathbf{c})=(\mathbf{a},\mathbf{c})+(\mathbf{b},\mathbf{c});$$

(3) **正定性**：$\forall \mathbf{a}\in V$，有 $(\mathbf{a},\mathbf{a})\geq 0$，等号成立当且仅当 $\mathbf{a}=\mathbf{0}$。

则称 $(\mathbf{a},\mathbf{b})$ 为向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的**内积**，定义了内积的实数域 $\mathbf{R}$ 上的线性空间 V 称为**欧几里得** (Euclid) **空间**，简称**欧氏空间**。

$\Rightarrow \forall \lambda_1,\lambda_2\in\mathbb{R},\boldsymbol{a},\boldsymbol{b}\in V$，有

$$(\lambda_1\mathbf{a}+\lambda_2\mathbf{b},\mathbf{c})=\lambda_1(\mathbf{a},\mathbf{c})+\lambda_2(\mathbf{b},\mathbf{c})$$
$$(\mathbf{a},\lambda_1\mathbf{b}+\lambda_2\mathbf{c})=\lambda_1(\mathbf{a},\mathbf{b})+\lambda_2(\mathbf{a},\mathbf{c})$$

$$\Rightarrow (\sum_{i=1}^{k}\lambda_i\mathbf{a}_i,\sum_{j=1}^{m}\mu_j\mathbf{b}_j)=\sum_{i=1}^{k}\sum_{j=1}^{m}\lambda_i\mu_j(\mathbf{a}_i,\mathbf{b}_j)$$

$$(\mathbf{a},\mathbf{0})=(\mathbf{a},\lambda\mathbf{0})=\lambda(\mathbf{a},\mathbf{0})$$
$$\Rightarrow (\lambda-1)(\mathbf{a},\mathbf{0})=0,\ \forall\lambda\in\mathbb{R}$$
$$\Rightarrow (\mathbf{a},\mathbf{0})=0$$

定理 6.1 (Cauchy-Schwarz 不等式)

设 V 是欧氏空间，$(\cdot,\cdot)$ 是 V 的内积，则对 V 中的任意两个向量 $\mathbf{a}$ 和 $\mathbf{b}$，有

$$|(\mathbf{a},\mathbf{b})| \leqslant \sqrt{(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})}.$$

$$\begin{aligned}
&\forall\lambda \in \mathbb{R}, \boldsymbol{a},\boldsymbol{b} \in V \Rightarrow (\lambda\boldsymbol{a}+\boldsymbol{b},\lambda\boldsymbol{a}+\boldsymbol{b}) \geqslant 0 \\
&\Rightarrow (\boldsymbol{a},\boldsymbol{a})\lambda^2 + 2(\boldsymbol{a},\boldsymbol{b})\lambda + (\boldsymbol{b},\boldsymbol{b}) \geqslant 0 \\
&\Rightarrow |2(\boldsymbol{a},\boldsymbol{b})|^2 \leqslant 4(\boldsymbol{a},\boldsymbol{a})(\boldsymbol{b},\boldsymbol{b}) \\
&\Rightarrow |(\boldsymbol{a},\boldsymbol{b})| \leqslant \sqrt{(\boldsymbol{a},\boldsymbol{a})(\boldsymbol{b},\boldsymbol{b})}
\end{aligned}$$

在一般线性空间之中我们定义范数，在欧几里得空间中范数就是内积。

定义 6.2

设 V 是欧氏空间，$(\cdot,\cdot)$ 是 V 的内积，对于任意的 $\mathbf{a} \in V$，称

$$|\mathbf{a}| = \sqrt{(\mathbf{a},\mathbf{a})}$$

称为 $\mathbf{a}$ 的**长度或模**。

范数⇒距离，$d(\boldsymbol{a},\boldsymbol{b}) = |\boldsymbol{a}-\boldsymbol{b}|, \forall \boldsymbol{a},\boldsymbol{b} \in V$
距离满足如下性质：

- **对称性**：$d(\mathbf{a},\mathbf{b}) = d(\mathbf{b},\mathbf{a})$；
- **正定性**：$d(\mathbf{a},\mathbf{b}) \geq 0$，等号成立当且仅当 $\mathbf{a}=\mathbf{b}$；
- **三角不等式**：$d(\mathbf{a},\mathbf{b}) \leq d(\mathbf{a},\mathbf{c}) + d(\mathbf{c},\mathbf{b})$。

我们也对向量之间的夹角进行了推广：

定义 6.3

对于欧氏空间 V 中两个非零向量 $\mathbf{a}$ 和 $\mathbf{b}$，定义它们之间的**夹角**为

$$\theta = \arccos\frac{(\mathbf{a},\mathbf{b})}{|\mathbf{a}||\mathbf{b}|}. \tag{1}$$

特别地，当 $(\mathbf{a},\mathbf{b}) = 0$ 时，称向量 $\mathbf{a}$ 和 $\mathbf{b}$ **正交**或**垂直**，记作 $\mathbf{a} \perp \mathbf{b}$。

例 6.1

设 $V=\mathbf{R}^n$，对于任意两个向量

$$\mathbf{a}=(a_1,a_2,\dots,a_n),\quad \mathbf{b}=(b_1,b_2,\dots,b_n)$$

定义

$$(\mathbf{a},\mathbf{b})=a_1b_1+a_2b_2+\cdots+a_nb_n\,,$$

容易验证它是 $\mathbf{R}^n$ 上的一个内积。此时，$\mathbf{R}^n$ 中向量 $\mathbf{a}$ 的模为

$$|\mathbf{a}|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\,,$$

两个向量 $\mathbf{a}$ 与 $\mathbf{b}$ 的夹角为

$$\theta=\arccos\frac{a_1b_1+a_2b_2+\cdots+a_nb_n}{\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}}。$$

对应这种内积，Cauchy-Schwarz 不等式为

$$|a_1b_1+\cdots+a_nb_n|\leqslant\sqrt{a_1^2+\cdots+a_n^2}\sqrt{b_1^2+\cdots+b_n^2}。$$

例 6.2

设 $C[a,b]$ 是闭区间 $[a,b]$ 上实的连续函数的全体。对于任意的 $f,g\in C[a,b]$，定义

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x\,,$$

容易验证，它满足内积定义中的 (1) 和 (2)。如果

$$(f,f)=\int_a^b f^2(x)\mathrm{d}x=0\,,$$

则由连续函数的性质可知 $f\equiv 0$，所以 (3) 也满足。对应的 Cauchy-Schwarz 不等式为

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leqslant\sqrt{\int_a^b f^2(x)\mathrm{d}x}\sqrt{\int_a^b g^2(x)\mathrm{d}x}。$$

特别地，对于 $C[-\pi,\pi]$，容易验证三角函数

$$1,\cos x,\sin x,\cos 2x,\sin 2x,\dots,\cos nx,\sin nx$$

在上述内积之下是两两正交的。

利用三角函数$1,\cos x,\sin x,\cos 2x,\sin 2x,\dots,\cos nx,\sin nx$是两两正交的我们发展出了 Fourier 分析。

内积的表示与标准正交基

和我们之前描述线性变换的思想类似，如果知道内积在线性空间V的一组基上的值，那么内积在任意向量上的取值也就清楚了。

设V是有限维的欧氏空间，$\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\dots,\boldsymbol{\alpha_n}$是它的一组基。任取$V$中的两个向量$\boldsymbol{x},\boldsymbol{y}$，

$$\boldsymbol{x}=x_1\boldsymbol{\alpha_1}+x_2\boldsymbol{\alpha_2}+\cdots+x_n\boldsymbol{\alpha_n}$$
$$\boldsymbol{y}=x_1\boldsymbol{\alpha_1}+x_2\boldsymbol{\alpha_2}+\cdots+y_n\boldsymbol{\alpha_n}$$

$$\Rightarrow (\boldsymbol{x},\boldsymbol{y})=(\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i,\sum_{j=1}^{n}y_j\boldsymbol{\alpha}_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$$

$$=\begin{pmatrix}x_1 & x_2 & \cdots & x_n\end{pmatrix}\begin{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n)\\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n)\\ \vdots & \vdots & \ddots & \vdots\\ (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n)\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$$

记 $G=(g_{ij})_{n\times n}$，其中 $g_{ij}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$, $i,j=1,2,\dots n$，称矩阵 G 是内积 $(\cdot,\cdot)$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_n$ 下的度量矩阵。

利用度量矩阵内积可以表示为$(\boldsymbol{x},\boldsymbol{y})=X^TGY$，其中$X=(x_1,x_2,\dots,x_n)^T,Y=(y_1,y_2,\dots,y_n)$

容易发现度量矩阵是一个**实对称阵**，并且具有如下性质$\boldsymbol{xGx}\geq 0$，等号成立当且仅当$\boldsymbol{x}=\boldsymbol{0}$

线性空间的基是不唯一的，问题：

(1) 欧氏空间的内积在不同基下对应的度量矩阵之间有什么关系?

(2) 是否存在一组特殊的基，使得欧氏空间的内积对应的度量矩阵尽可能的简单?

设欧氏空间V的一组基为：$\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\dots,\boldsymbol{\alpha_n}$，内积$(\cdot,\cdot)$在这组基下面的度量矩阵为$G$，现在有$V$的另外一组基：$\boldsymbol{\eta_1},\boldsymbol{\eta_2},\dots,\boldsymbol{\eta_n}$，内积$(\cdot,\cdot)$在这组基下面的度量矩阵为$\bar{G}$，则对于$\boldsymbol{a},\boldsymbol{b}\in V$有

$$(\boldsymbol{a},\boldsymbol{b})=\begin{pmatrix}a_1 & \cdots & a_n\end{pmatrix}G\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}=\begin{pmatrix}\overline{a}_1 & \cdots & \overline{a}_n\end{pmatrix}\overline{G}\begin{pmatrix}\overline{b}_1\\ \vdots\\ \overline{b}_n\end{pmatrix},$$

其中 $\boldsymbol{a}=a_1\boldsymbol{\alpha}_1+\cdots+a_n\boldsymbol{\alpha}_n=\overline{a}_1\boldsymbol{\eta}_1+\cdots+\overline{a}_n\boldsymbol{\eta}_n$,

$\boldsymbol{b}=b_1\boldsymbol{\alpha}_1+\cdots+b_n\boldsymbol{\alpha}_n=\overline{b}_1\boldsymbol{\eta}_1+\cdots+\overline{b}_n\boldsymbol{\eta}_n$。

记 $\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\dots,\boldsymbol{\alpha_n}$ 到 $\boldsymbol{\eta_1},\boldsymbol{\eta_2},\dots,\boldsymbol{\eta_n}$ 的过度矩阵为 P，即 $(\boldsymbol{\eta_1},\boldsymbol{\eta_2},\dots,\boldsymbol{\eta_n})=(\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\dots,\boldsymbol{\alpha_n})P$，

$$\Rightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = P \begin{pmatrix} \overline{a}_1 \\ \vdots \\ \overline{a}_n \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = P \begin{pmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_n \end{pmatrix}$$

所以

$$(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a_1}\ \boldsymbol{a_2}\ \dots\ \boldsymbol{a_n})G \begin{pmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \\ \vdots \\ \boldsymbol{b_n} \end{pmatrix} = (\overline{\boldsymbol{a_1}}\ \overline{\boldsymbol{a_2}}\ \dots\ \overline{\boldsymbol{a_n}})P^T GP \begin{pmatrix} \overline{\boldsymbol{b_1}} \\ \overline{\boldsymbol{b_2}} \\ \vdots \\ \overline{\boldsymbol{b_n}} \end{pmatrix}$$

$$= (\overline{\boldsymbol{a_1}}\ \overline{\boldsymbol{a_2}}\ \dots\ \overline{\boldsymbol{a_n}})\bar{G} \begin{pmatrix} \overline{\boldsymbol{b_1}} \\ \overline{\boldsymbol{b_2}} \\ \vdots \\ \overline{\boldsymbol{b_n}} \end{pmatrix}$$

容易得到$\bar{G} = P^T GP$

定义 6.4

设 $A, B \in F^{n\times n}$，若存在数域 F 上的 n 阶可逆方阵 P，使得 $B = P^{\mathrm{T}}AP$，则称 A 与 B（在数域 F 上）**相合**，记作 $A \sim B$。

容易验证，相合是一种等价关系。
因此可以得到：欧氏空间内积在不同基下的度量矩阵彼此相合。

定义 6.5

如果 V 中有一组两两正交的非零向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r$，即

$$(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = |\boldsymbol{\alpha}_i||\boldsymbol{\alpha}_j|\delta_{ij},\ i,j = 1, 2, \dots, r,$$

则称该向量组为**正交向量组**。由正交向量组构成的基称为**正交基**。由单位向量组构成的正交基称为**标准正交基**。

$$其中\delta_{ij} = \begin{cases} 1\,, i = j \\ 0\,, i \neq j \end{cases}$$

规定：单独一个非零向量也构成正交向量组。

命题 6.2

欧氏空间中的正交向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r\} \Rightarrow \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r$ **线性无关。**

内积在正交基下的度量矩阵为对角阵，在标准正交基下的度量矩阵为单位阵。

定理 6.3 (Schmidt 正交化)

从 n 维欧氏空间 V 的任意一组基出发，可以构造一组标准正交基。

设欧氏空间 V 的一组基：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n \Rightarrow$ 构造一组标准正交基?

$$\boldsymbol{e}_1 = \frac{\boldsymbol{\alpha}_1}{|\boldsymbol{\alpha}_1|},$$

$$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - (\boldsymbol{\alpha}_2, \boldsymbol{e}_1)\boldsymbol{e}_1,\ \boldsymbol{e}_2 = \frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},$$

$$\vdots$$

$$\boldsymbol{\beta}_k = \boldsymbol{\alpha}_k - \sum_{i=1}^{k-1}(\boldsymbol{\alpha}_k, \boldsymbol{e}_i)\boldsymbol{e}_i,\ \boldsymbol{e}_k = \frac{\boldsymbol{\beta}_k}{|\boldsymbol{\beta}_k|},\ k = 1, 2, \ldots, n,$$

$\Rightarrow \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 为一组标准正交基，上述构造性算法称为 Schmidt 正交化方法

欧几里德空间的同构

定义 6.6

实数域 $\mathbb{R}$ 上的两个欧氏空间 V 和 V' 称为同构的，如果存在一个从 V 到 V' 的一一映射 $\sigma : V \to V'$，满足

(1) $\sigma(\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\beta}) = \lambda\sigma(\boldsymbol{\alpha}) + \mu\sigma(\boldsymbol{\beta})$;

(2) $(\sigma(\boldsymbol{\alpha}), \sigma(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \boldsymbol{\beta})$,

其中 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V,\ \lambda, \mu \in \mathbb{R}$。

定理 6.4

两个有限维欧氏空间同构的充分必要条件是它们的维数相同。

我们可以认为欧氏空间的结构完全由它的维数决定。

欧几里德空间中的线性变换

定义 6.7

设 V 是一个 n 维的欧氏空间，$\mathcal{A}$ 是 V 上的一个线性变换，如果 $\mathcal{A}$ 保持 V 的内积不变，即 $\forall \boldsymbol{a}, \boldsymbol{b} \in V$ 都有

$$(\mathcal{A}\boldsymbol{a}, \mathcal{A}\boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{b}),$$

则称 $\mathcal{A}$ 是 V 上的**正交变换**。

定理 6.5

设 V 是一个 n 维的欧氏空间，$\mathcal{A}$ 是 V 上的一个线性变换，则 $\mathcal{A}$ 为正交变换当且仅当下列两个条件之一成立

(1) *$\mathcal{A}$ 保持任意向量的模长不变；*

(2) *$\mathcal{A}$ 将标准正交基变换为标准正交基。*

证明是非常显然的。

下面我们来看下正交变换的**代数特征**：

定义 6.8

如果实方阵 A 满足

$$A^{\mathrm{T}}A = I \quad 或 \quad A^{-1} = A^{T},$$

则称方阵 A 为**正交矩阵**。

证明：

欧氏空间V上有一组正交基$\boldsymbol{e_1}, \boldsymbol{e_2}, .., \boldsymbol{e_n}$，$V$上的线性变换$\mathcal{A}$在这组基下面的表示矩阵为$A$

$$(\mathcal{A}\mathbf{e}_1, \mathcal{A}\mathbf{e}_2, \ldots, \mathcal{A}\mathbf{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)A$$

$$\Rightarrow \mathcal{A}\boldsymbol{e}_i = \sum_{l=1}^{n} a_{li}\boldsymbol{e}_l,\ \mathcal{A}\boldsymbol{e}_j = \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k$$

$$\Rightarrow (\mathcal{A}\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_j) = (\sum_{l=1}^{n} a_{li}\boldsymbol{e}_l, \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k) = \sum_{l=1}^{n}\sum_{k=1}^{n} a_{li}\, a_{kj}(\boldsymbol{e}_l, \boldsymbol{e}_k)$$

$$= \sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij},\ \forall i, j = 1, 2, \ldots, n \ \Leftrightarrow\ A^{\mathrm{T}}A = I$$

$A^T A = I \Longrightarrow \det(A) = \pm 1$

n 阶实方阵 A 为正交方阵 $\Leftrightarrow$ A 的行向量组（列向量组）构成 $R^{n\times n}$ 的标准正交基。

容易发现A 的行向量组（列向量组）的模都是1，且相互正交。

定理 6.6

欧氏空间中的线性变换 $\mathcal{A}$ 是正交变换的充分必要条件是 $\mathcal{A}$ 在标准正交基下的矩阵 A 是正交矩阵。

必要性在上文已经完成了证明，下面证明充分性：

记 $(\vec{e}_1, \vec{e}_2, \cdots \vec{e}_n)$ 是一组标准正交基

设：$\vec{\alpha} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$，$\vec{\beta} = y_1\vec{e}_1 + \cdots + y_n\vec{e}_n$

$(\vec{\alpha}, \vec{\beta}) = x_1y_1 + \cdots + x_ny_n$

$\mathcal{A}(\vec{\alpha}) = \vec{\alpha}A$　　$\cancel{\mathcal{A}(\vec{\beta}) = \vec{\beta}A}$　　$\mathcal{A}(\vec{\beta}) = \vec{\beta}A$

$= (x_1, \cdots x_n)\begin{pmatrix}\vec{e}_1\\ \vdots \\ \vec{e}_n\end{pmatrix} A$　　$= (y_1, \cdots y_n)\begin{pmatrix}\vec{e}_1\\ \vdots \\ \vec{e}_n\end{pmatrix} A$

$(\mathcal{A}(\vec{\alpha}), \mathcal{A}(\vec{\beta})) = (x_1, \cdots x_n)\begin{pmatrix}\vec{e}_1A\\ \vdots \\ \vec{e}_nA\end{pmatrix}(A^T\vec{e}_1, \cdots A^T\vec{e}_n)\begin{pmatrix}y_1\\ y_2\\ \vdots \\ y_n\end{pmatrix}$

$= x_1y_1 + \cdots + x_ny_n$

$= (\vec{\alpha}, \vec{\beta})$

$\therefore$ 变换 $\mathcal{A}$ 是正交变换

命题 6.7

设 V 是 n 维欧氏空间，则

(1) 单位变换是正交变换；
(2) 两个正交变换的复合仍然是正交变换；
(3) 正交变换一定可逆，其逆变换也是正交变换。

证明是显然的。

定义 6.9

如果正交变换 $\mathcal{A}$ 在一组基下的矩阵行列式为 1，则称 $\mathcal{A}$ 为第一类变换。如果正交变换 $\mathcal{A}$ 在一组基下的矩阵行列式为 -1，则称 $\mathcal{A}$ 为第二类变换。

在R^2中，旋转变换是第一类正交变换，反射变化是第二类正交变换。

记 $O(V)$ 为 n 维欧氏空间 V 上正交变换全体所构成的集合
$\Rightarrow$ 正交变换群

记 $GL(V)$ 为 n 维欧氏空间 V 上可逆变换全体所构成的集合
$\Rightarrow$ 一般线性群

定理 6.8

设 $\mathcal{A}$ 是欧氏空间 V 上的正交变换，则

(1) $\mathcal{A}$ 的特征值的模长都为 1。特别地，$\mathcal{A}$ 的实特征值（如果存在的话）只能是 1 或者 -1；

(2) 如果 V 的维数是奇数且 $\mathcal{A}$ 是第一类正交变换，则 $\mathcal{A}$ 一定存在值为 1 的特征值。

对于(1)设μ为$\mathcal{A}$的一个特征值，那么有$A\boldsymbol{\alpha}=\mu\boldsymbol{\alpha}$，我们知道正交变换不改变向量的模长所以$|\mu\boldsymbol{\alpha}|=|\boldsymbol{\alpha}|$，由此可以得到$|\mu|=1$。

对于(2)有$A^TA\boldsymbol{\alpha}=A^T\mu\boldsymbol{\alpha}\Rightarrow\boldsymbol{\alpha}=\mu A^T\boldsymbol{\alpha}$，那么容易得到$\left(\frac{1}{\mu}I-A^T\right)\boldsymbol{\alpha}=\mathbf{0}$，所以如果$\mu$是这个变换的特征值，那么$\frac{1}{\mu}$一定也是这个变换的特征值。所以如果$V$的维数是奇数，就要求有一个特征值的导数是它自身，这个特征值只能是1或-1，其余的特征值都成对出现，$\mathcal{A}$是第一类正交变换，所以$\det(A)=1$，所以这个特征值只能是1.

例 6.3

在二维平面上任取一点 O 作为原点，将平面上每个点 P 与向量 $\overrightarrow{OP}$ 对应起来，将平面看成二维欧氏空间 V。设 $\mathcal{A}$ 是 V 上的正交变换，则

(1) 当 $\det(\mathcal{A})=1$ 时，$\mathcal{A}$ 是绕原点的旋转，在 V 的任意一组标准正交基下的表示矩阵 $A=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$；

(2) 当 $\det(\mathcal{A})=-1$ 时，$\mathcal{A}$ 是关于过原点的某条直线 l 的轴对称，在 V 的某一组标准正交基下的表示矩阵 $A=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$。

例 6.4

设 $\mathbb{R}^3$ 是建立了直角坐标系的欧式空间，$\mathcal{A}$ 是 $\mathbb{R}^3$ 上的正交变换且 $\det(\mathcal{A}) = 1$。求证：$\mathcal{A}$ 是绕某条过原点的直线的旋转。

证明：

根据 Thm 6.8-(2)，A一定有一个特征值是 1.

设$\boldsymbol{x_3}$是A的属于特征值1的特征向量，那么$\boldsymbol{\beta_3} = \frac{1}{|x_3|}\boldsymbol{x_3}$也是属于特征值1的特征向量，且$|\boldsymbol{\beta_3}| = 1$.将$\boldsymbol{\beta_3}$扩充为$R^3$上的一组标准正交基$\{\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}\}$，以其中的向量为列组成正交方阵$P$，$\det(P) = \pm 1$，当$\det(P) = -1$时，令$\boldsymbol{\beta_1} = -\boldsymbol{\beta_1}$，此时新的方阵仍然是正交方阵，且行列式为1，此时$P$的三个列向量组成右手系标准正交基$M$，设：

$$AP = A(\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}) = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3})B$$

其中B的第j列是$A\boldsymbol{\beta_j}$在基M上的坐标。特别的，B的第三列应该是$A\boldsymbol{\beta_3} = \boldsymbol{\beta_3}$在基$M$下的坐标，等于$(0,0,1)^T$，因此

$$B = P^{-1}AP = \begin{pmatrix} B_{22} & O \\ O & 1 \end{pmatrix}$$

由于P和A都是正交方阵，所以B

定义 6.10

设 V 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的线性变换。如果 $\mathcal{A}$ 满足

$$(\mathbf{a}, \mathcal{A}\mathbf{b}) = (\mathcal{A}\mathbf{a}, \mathbf{b}), \ \forall \boldsymbol{a}, \boldsymbol{b} \in V,$$

则称 $\mathcal{A}$ 是 V 上的对称变换。

定理 6.9

设 $\mathcal{A}$ 是欧氏空间 V 上的线性变换，则 $\mathcal{A}$ 是对称变换 $\Leftrightarrow$ $\mathcal{A}$ 在 V 的任何一组标准正交基下的表示矩阵 A 都是实对称方阵。

证明 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 V 中的标准正交基, $\boldsymbol{A} = (a_{ij})$ 是对称变换 $\mathscr{A}$ 在这组基下的矩阵, 即

$$(\mathscr{A}\boldsymbol{e}_1, \mathscr{A}\boldsymbol{e}_2, \cdots, \mathscr{A}\boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)\boldsymbol{A},$$

或写成

$$\mathscr{A}\boldsymbol{e}_j = a_{1j}\boldsymbol{e}_1 + \cdots + a_{nj}\boldsymbol{e}_n = \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k.$$

因此

$$(\boldsymbol{e}_i, \mathscr{A}\boldsymbol{e}_j) = \left(\boldsymbol{e}_i, \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k\right) = \sum_{k=1}^{n} a_{kj}(\boldsymbol{e}_i, \boldsymbol{e}_k) = a_{ij},$$

$$(\mathscr{A}\boldsymbol{e}_i, \boldsymbol{e}_j) = \left(\sum_{k=1}^{n} a_{ki}\boldsymbol{e}_k, \boldsymbol{e}_j\right) = \sum_{k=1}^{n} a_{ki}(\boldsymbol{e}_k, \boldsymbol{e}_j) = a_{ji}.$$

以上两式相等得 $a_{ij} = a_{ji}$.

反之, 如果 $\mathscr{A}$ 在标准正交基下的矩阵 $\boldsymbol{A} = (a_{ij})$ 满足 $a_{ij} = a_{ji}$, 则

$$(\boldsymbol{e}_i, \mathscr{A}\boldsymbol{e}_j) = (\mathscr{A}\boldsymbol{e}_i, \boldsymbol{e}_j).$$

因此对任意两个向量

$$\boldsymbol{a} = \sum_{i=1}^{n} a_i\boldsymbol{e}_i, \quad \boldsymbol{b} = \sum_{j=1}^{n} b_j\boldsymbol{e}_j,$$

有

$$(\boldsymbol{a}, \mathscr{A}\boldsymbol{b}) = \sum_{i,j=1}^{n} a_i b_j(\boldsymbol{e}_i, \mathscr{A}\boldsymbol{e}_j) = \sum_{i,j=1}^{n} a_i b_j(\mathscr{A}\boldsymbol{e}_i, \boldsymbol{e}_j) = (\mathscr{A}\boldsymbol{a}, \boldsymbol{b})$$

即 $\mathscr{A}$ 是对称变换. □

定理 6.10

设 $\mathcal{A}$ 是欧氏空间 V 上的对称变换，则 $\mathcal{A}$ 的不同特征值对应的特征向量相互正交。

考虑到$\mathcal{A}(\boldsymbol{x}) = \lambda\boldsymbol{x}, \mathcal{A}(\boldsymbol{y}) = \mu\boldsymbol{y}$，所以根据对称变换的定义（定义 6.10）容易得到

$$(\mathcal{A}(\boldsymbol{x}), \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}(\boldsymbol{y})) \Rightarrow \lambda(\boldsymbol{x}, \boldsymbol{y}) = \mu(\boldsymbol{x}, \boldsymbol{y}) \Rightarrow (\lambda - \mu)(\boldsymbol{x}, \boldsymbol{y}) = 0$$

因为$\lambda \neq \mu$，所以$(\boldsymbol{x}, \boldsymbol{y}) = 0$

推论 6.1

实对称阵 A 的属于不同特征值的特征向量必正交。

命题 6.11

实对称矩阵 A 的特征值都是实数。

设λ为A的特征值，故$A\boldsymbol{x} = \lambda\boldsymbol{x}$，其中$\boldsymbol{x} \neq \boldsymbol{0}$是矩阵$A$对应于$\lambda$的特征向量。

记$\bar{\boldsymbol{x}}$为$\boldsymbol{x}$的共轭向量，则有$\bar{\boldsymbol{x}}^T A\boldsymbol{x} = \bar{\boldsymbol{x}}^T \lambda\boldsymbol{x} = \lambda\bar{\boldsymbol{x}}^T\boldsymbol{x}$

又因为$\bar{\boldsymbol{x}}^T A\boldsymbol{x} = \bar{\boldsymbol{x}}^T A^T\boldsymbol{x} = \bar{\boldsymbol{x}}^T \bar{A}^T\boldsymbol{x} = \overline{\boldsymbol{x}^T A^T}\boldsymbol{x} = \bar{\lambda}\bar{\boldsymbol{x}}^T\boldsymbol{x}$

所以$\lambda\bar{\boldsymbol{x}}^T\boldsymbol{x} = \bar{\lambda}\bar{\boldsymbol{x}}^T\boldsymbol{x} \Longrightarrow (\lambda - \bar{\lambda})\bar{\boldsymbol{x}}^T\boldsymbol{x} \Longrightarrow \lambda - \bar{\lambda} = 0$

所以λ为实数。

定理 6.12

对于任意 n 阶实对称阵 A，存在一个 n 阶正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵。

证明 对矩阵 $\boldsymbol{A}$ 的阶数 n 用数学归纳法. 当 $n=1$ 时, 结论显然成立. 现假设 $n-1$ 时结论成立. 现在看 n 的情形, 由命题 7.3.3 知 $\boldsymbol{A}$ 的特征值都是实数. 设 λ_1 是 $\boldsymbol{A}$ 的一个特征值, $\boldsymbol{x}_1$ 是属于 λ_1 的单位特征向量. 则由 Schmidt 正交化可知, 可以将 $\boldsymbol{x}_1$ 扩充成 F^n 的一组标准正交基 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$.

令 $\boldsymbol{T}_n = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$, 则 $\boldsymbol{T}_n$ 是一个正交矩阵. 注意到 $\boldsymbol{A}$ 是对称矩阵, 则 $\boldsymbol{T}_n^{-1}\boldsymbol{A}\boldsymbol{T}_n$ 是对称的. 利用 $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1$ 以及 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_i$ $(2 \leqslant i \leqslant n)$ 的正交性有

$$\boldsymbol{T}_n^{-1}\boldsymbol{A}\boldsymbol{T}_n = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)^{\mathrm{T}}(\lambda_1\boldsymbol{x}_1, \boldsymbol{A}\boldsymbol{x}_2, \cdots, \boldsymbol{A}\boldsymbol{x}_n) = \begin{pmatrix} \lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{n-1} \end{pmatrix},$$

其中 $\boldsymbol{A}_{n-1}$ 是 $n-1$ 阶实对称矩阵, 由归纳假设, 存在 $n-1$ 阶正交矩阵 $\boldsymbol{T}_{n-1}$, 使得 $\boldsymbol{T}_{n-1}^{-1}\boldsymbol{A}_{n-1}\boldsymbol{T}_{n-1} = \mathrm{diag}\,(\lambda_2, \lambda_3, \cdots, \lambda_n)$.

令 $\boldsymbol{T} = \boldsymbol{T}_n \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1} \end{pmatrix}$, 显然 $\boldsymbol{T}$ 仍然是正交矩阵, 且有

$$\begin{aligned}
\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} &= \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1} \end{pmatrix}^{-1} \boldsymbol{T}_n^{-1}\boldsymbol{A}\boldsymbol{T}_n \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1} \end{pmatrix} \\
&= \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1}^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{n-1} \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_{n-1}^{-1}\boldsymbol{A}_{n-1}\boldsymbol{T}_{n-1} \end{pmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).
\end{aligned}$$

Chapter Ⅶ 实二次型

多元二次多项式：二次项的全体组成一个二次齐次多项式，称之为**二次型**。

目标：利用线性代数的理论，研究二次型的基本理论、性质及其应用等。

二次型的矩阵表示

定义 7.1

在实数域 $\mathbb{R}$ 上，一个含 n 个变元 $x_1,\dots,x_n$ 的**二次型** $Q(x_1,\dots,x_n)$ 是一个齐次的二次多项式

$$Q(x_1,\dots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

其中 $a_{ji}=a_{ij}$ $(i,j=1,2,\dots,n)$ 为实系数。

（若 $a_{ij}\neq a_{ji}$，$\Rightarrow a_{ij}'=a_{ji}'=\frac{a_{ij}+a_{ji}}{2}$）

利用矩阵的语言，记 $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{T}}$，$A=(a_{ij})_{n\times n}$，则

$$Q(x_1,\dots,x_n)=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x},$$

称实对称矩阵 A 为二次型 $Q(x_1,\dots,x_n)$ 的**矩阵**，A 的秩称为**二次型的秩**。

多元二次多项式：二次项的全体组成一个二次齐次多项式，称之为二次型。
每一个二次型和一个n阶实对称矩阵一一对应。

同一二次型在不同基下面的矩阵有什么关系？

设线性空间 V 的一组基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n$，$V$ 上的二次型 $Q(x_1, \dots, x_n)$ 在这组基下的表示为：

$$Q(x_1, \dots, x_n) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}.$$

V 有另一组基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n$ 且有

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n) P,$$

则

$$Q(y_1, \dots, y_n) = (P\boldsymbol{y})^{\mathrm{T}} A (P\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}} P^{\mathrm{T}} A P \boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}} B \boldsymbol{y},$$

其中对称矩阵 $B = P^{\mathrm{T}} A P$。

定义 7.2

对于实数域 $\mathbb{R}$ 上两个 n 阶矩阵 A, B，如果存在一个可逆的实矩阵 P 使得

$B = P^{\mathrm{T}} A P$, 则称 A 和 B 是**相合的**，或者说 B **相合于**A，矩阵 P 称为**相合变换矩阵**。

相合关系是**等价关系**

二次型的标准型

定义 7.3

如果二次型 $Q(x_1,\ldots,x_n)=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 经过可逆线性变换 $\boldsymbol{x}=P\boldsymbol{y}$ 化为下列平方和的简单形式：

$$Q(x_1,\ldots,x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}}=\mu_1y_1^2+\cdots+\mu_ny_n^2=\boldsymbol{y}^{\mathrm{T}}B\boldsymbol{y},$$

其中 $B=\mathrm{diag}(\mu_1,\ldots,\mu_n)$ 是实对角阵，称这种平方和形式为二次型的标准形。

二次型 $Q(x_1,\ldots,x_n)=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，其中 A 为实对称矩阵

$$\Rightarrow \text{存在正交矩阵 } P \text{ 使得 } P^{-1}AP=P^{\mathrm{T}}AP=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

$$\Rightarrow Q(x_1,\ldots,x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2$$

定理 7.1

实数域 $\mathbb{R}$ 上的任何一个二次型 $Q(x_,\ldots,x_n)$ 都可以经过正交变换 $\boldsymbol{x}=P\boldsymbol{y}$ 化为标准形

$$Q(x_1,\ldots,x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2.$$

参考定理 6.12

定理 7.2

对于任何一个实二次型

$$Q(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

均可通过配平方法找到可逆变换 $\boldsymbol{x}=P\boldsymbol{y}$，将二次型化为标准形

$$Q(x_1,\ldots,x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}}=\mu_1y_1^2+\cdots+\mu_ny_n^2.$$

这里的$\mu_1,\mu_2,\ldots,\mu_n$不要求是矩阵的特征值。

证明 对变量的个数作归纳法. 对于 $n=1$, $Q(x_1)=a_{11}x_1^2$, 结论是显然的. 假设定理对 $n-1$ 元的二次型成立, 则对于 n 元的情形, 分三种情况讨论.

(1) $a_{11}\neq 0$. 如果 $a_{11}=0$, 但存在某个 $a_{ii}\neq 0$, 只需将 x_1 和 x_i 对换, 使二次型成为上述形式即可. 这时

$$\begin{aligned}
Q(x_1,\cdots,x_n)&=a_{11}x_1^2+2\sum_{j=2}^{n}a_{1j}x_1x_j+\sum_{i,j=2}^{n}a_{ij}x_ix_j\\
&=a_{11}\left(x_1+\sum_{j=2}^{n}a_{1j}a_{11}^{-1}x_j\right)^2-a_{11}^{-1}\left(\sum_{j=2}^{n}a_{1j}x_j\right)^2+\sum_{i,j=2}^{n}a_{ij}x_ix_j\\
&=a_{11}\left(x_1+\sum_{j=2}^{n}a_{1j}a_{11}^{-1}x_j\right)^2+\sum_{i,j=2}^{n}b_{ij}x_ix_j.
\end{aligned}$$

上式中

$$\sum_{i,j=2}^{n}b_{ij}x_ix_j=-a_{11}^{-1}\left(\sum_{j=2}^{n}a_{1j}x_j\right)^2+\sum_{i,j=2}^{n}a_{ij}x_ix_j$$

是一个含 $n-1$ 个变元 $x_2,x_3,\cdots,x_n$ 的实二次型. 令

$$\begin{cases}
y_1=x_1+\sum\limits_{j=2}^{n}a_{1j}a_{11}^{-1}x_j,\\
y_2=x_2,\\
\cdots\cdots\cdots\cdots\\
y_n=x_n
\end{cases}
\quad 或 \quad
\begin{cases}
x_1=y_1-\sum\limits_{j=2}^{n}a_{1j}a_{11}^{-1}y_j,\\
x_2=y_2,\\
\cdots\cdots\cdots\cdots\\
x_n=y_n,
\end{cases}$$

其对应的变换矩阵是

$$\boldsymbol{S}=\begin{pmatrix}
1 & -a_{12}a_{11}^{-1} & \cdots & -a_{1n}a_{11}^{-1}\\
0 & 1 & \cdots & 0\\
\vdots & \vdots & & \vdots\\
0 & 0 & \cdots & 1
\end{pmatrix}.$$

经过变换 $\boldsymbol{x}=\boldsymbol{S}\boldsymbol{y}$, 原二次型化为

$$\widetilde{Q}(y_1,\cdots,y_n)=a_{11}y_1^2+\sum_{i,j=2}^{n}b_{ij}y_iy_j.$$

由归纳假设, 对于 $n-1$ 元的二次型 $\sum\limits_{i,j=2}^{n}b_{ij}y_iy_j$, 存在一个可逆的线性变换

$\boldsymbol{y}=\boldsymbol{T}\boldsymbol{z}$, 使得

$$\sum_{i,j=2}^{n}b_{ij}y_iy_j=\sum_{i=2}^{n}\mu_iz_i^2.$$

令 $z_1=y_1,\mu_1=a_{11}$, 则原二次型化为标准形 $\sum\limits_{i=1}^{n}\mu_iz_i^2$. 相应的变换为 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{z}$, 这里

$$\boldsymbol{P}=\boldsymbol{S}\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{T}\end{pmatrix}.$$

(2) 如果所有的 $a_{ii}=0$, $i=1,2,\cdots,n$, 但至少有一个 $a_{1j}\neq 0$, $j>1$. 不妨设 $a_{12}\neq 0$, 令

$$\begin{cases} x_1 = y_1 - y_2, \\ x_2 = y_1 + y_2, \\ x_3 = y_3, \\ \cdots\cdots\cdots\cdots \\ x_n = y_n, \end{cases}$$

它是可逆的线性变换, 对应的变换矩阵为

$$\boldsymbol{P} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

于是

$$\begin{aligned} Q(x_1, \cdots, x_n) &= 2a_{12}x_1x_2 + \cdots = 2a_{12}(y_1^2 - y_2^2) + \cdots \\ &= 2a_{12}y_1^2 - 2a_{12}y_2^2 + \cdots \end{aligned}$$

化为关于变元 $y_1, y_2, \cdots, y_n$ 的二次型, 且平方项 y_1^2 的系数不为零, 从而化成情形 (1).

(3) 如果对所有的 $a_{ii} = 0$, $i = 1, 2, \cdots, n$ 及 $a_{1j} = 0$, $j = 1, 2, \cdots, n$, 则由对称性, 得 $a_{j1} = 0$, $j = 1, 2, \cdots, n$. 因此二次型简化为

$$Q(x_1, \cdots, x_n) = \sum_{i,j=2}^{n} a_{ij}x_ix_j.$$

这是一个含有 $n-1$ 个变元的二次型, 由归纳假设, 它经过线性变换可以化为标准形.

总之, 对于任一种形式的实二次型都存在可逆的线性变换将其化为标准形. $\square$

定理 7.3

对每一个实二次型都可以通过初等变换使之化为标准形 $\Leftrightarrow$ 对每一个实对称矩阵 A，存在一系列初等矩阵 $P_1, \ldots, P_r$，使得 A 相合于实对角矩阵

$$P_r^{\mathrm{T}}P_{r-1}^{\mathrm{T}}\cdots P_1^{\mathrm{T}}AP_1P_2\cdots P_r = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}.$$

根据定理 7.2，这个结论是非常显然的，下面我们使用矩阵的语言重新叙述：

对矩阵 $\boldsymbol{A}$ 的阶数 n 进行归纳. 当 $n = 1$ 时, 结论是显然的. 假设对于 $n-1$ 阶对称矩阵, 定理已成立, 则对 n 阶对称矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$, 分如下情形讨论:

(1) 如果 $a_{11} \neq 0$, 作“把第 1 列乘 $-a_{i1}a_{11}^{-1}$ 加到第 i 列”的初等变换, 同时作“把第 1 行乘 $-a_{i1}a_{11}^{-1}$ 加到第 i 行”的初等变换, $i = 2, 3, \cdots, n$, 由于 $\boldsymbol{A}$ 是对称矩阵, 经过上述变换, 矩阵 $\boldsymbol{A}$ 化为 $\mathrm{diag}(a_{11}, \boldsymbol{A}_{n-1})$, 即存在可逆矩阵 $\boldsymbol{P}_1$, 使得

$$P_1^{\mathrm{T}} A P_1 = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & A_{n-1} \end{pmatrix}.$$

这里 A_{n-1} 是一个 $n-1$ 阶的实对称矩阵. 根据归纳假设, A_{n-1} 相合于对角矩阵, 即存在可逆方阵 P_2, 使得 $P_2^{\mathrm{T}} A_{n-1} P_2 = \Lambda$, Λ 为对角矩阵. 令 $P = P_1 \mathrm{diag}(1, P_2)$, 则

$$P^{\mathrm{T}} A P = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix}.$$

(2) 当 $a_{11} = 0$ 时, 如果有一个 $a_{ii} \neq 0, i = 2, 3, \cdots, n$, 则只要把 A 的第一行与第 i 行互换, 再把第 1 列和第 i 列互换, 就成为第一种情形. 这个过程就是对 A 作相合变换 $S_{1i} A S_{1i}$.

(3) 当 $a_{ii} = 0$, $i = 1, 2, \cdots, n$ 时, 则第一行的其他元素至少有一个不等于零. 否则 A 的第一行、第一列均为 0, 从而可归纳为 $n-1$ 的情形. 不妨设 $a_{1i} \neq 0$, 则作初等相合变换

$$T_{1i}(1) A T_{1i}(1)^{\mathrm{T}} = \begin{pmatrix} 2a_{1i} & C \\ C^{\mathrm{T}} & A_{n-1} \end{pmatrix},$$

于是问题又归结到第一种情形. 这样就完成了定理的证明. □

我们如何计算得到矩阵P使得$P^T AP = diag(\mu_1, \mu_2, \dots, \mu_n)$呢？

$$\left(A \vdots I\right) \xrightarrow[\text{对 } I \text{ 只作初等行变换}]{\text{对 } A \text{ 作成对的初等行列变换}} \left(P^T AP \vdots P^T\right)$$

$$\begin{pmatrix} A \\ I \end{pmatrix} \xrightarrow[\text{对 } I \text{ 只作初等列变换}]{\text{对 } A \text{ 作成对的初等行列变换}} \begin{pmatrix} P^T AP \\ P \end{pmatrix}$$

相合不变量与分类

二次型的标准型并不唯一，所以它不可以作为相合等价类的代表元。

定理 7.4

设 A 是一个 n 阶实对称矩阵，则存在可逆矩阵 P，使得

$$P^{\mathrm{T}}AP = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{pmatrix}, \quad \mathrm{rank}(A) = r + s \leq n$$

其中 r 是标准形中正项的项数，s 是负项的项数。上式右边的对角矩阵称为矩阵 A 的规范形。

$\Longleftrightarrow$ 给定实二次型 $Q(x_1, \ldots, x_n)$，存在可逆方阵 P 使得

$Q(x_1, \ldots, x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}} = y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2$，上式称为二次型 $Q(x_1, \ldots, x_n)$ 的规范形。

定理 7.5

惯性定理 实二次型 $Q(x_1, \ldots, x_n) = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的规范形中正项数 r 和负项数 s 是由二次型 $Q(x_1, \ldots, x_n)$ 唯一确定的，或者说是由实对称矩阵 A 唯一确定的。

r为正惯性指数，s为负惯性指数，它们是相合等价类的**全系不变量**。

$r+s$就是二次型的秩，即$rank(A) = r + s$
$r - s = 2r - rank(A)$称为二次型（或者矩阵A）的符号差。

二次曲线和二次曲面的分类

设欧氏空间R^2中取定一个直角坐标系：$[O;\boldsymbol{e}_1,\boldsymbol{e}_2]$，平面二次曲线的一般方程为：

$$F(x,y)=a_{11}x^2+2a_{12}xy+a_{22}y^2+2b_1x+2b_2y+c=0,$$

其中a_{11},a_{12},a_{22}不全为 0.

通过分类我们想要得到：

(1) 确定$F(x,y)$所对应的是哪一类曲线（直接通过系数来判断）；

(2) 确定$F(x,y)$的标准方程。

若采用齐次坐标，则

$$F(x,y)=\begin{pmatrix}x & y & 1\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\\ b_1 & b_2 & c\end{pmatrix}\begin{pmatrix}x\\ y\\ 1\end{pmatrix}$$

$$=\begin{pmatrix}X^{\mathrm{T}} & 1\end{pmatrix}\begin{pmatrix}A & B\\ B^{\mathrm{T}} & c\end{pmatrix}\begin{pmatrix}X\\ 1\end{pmatrix}$$

其中 $a_{21}=a_{12}$，$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$ 是 2 阶实对称方阵，$B=\begin{pmatrix}b_1\\ b_2\end{pmatrix}$，$X=\begin{pmatrix}x\\ y\end{pmatrix}$。

如果取一个新的直角坐标系：$[\widetilde{O};\widetilde{\boldsymbol{e}}_1,\widetilde{\boldsymbol{e}}_2]$，则

$$X=T\widetilde{X}+X_0\Longleftrightarrow\begin{pmatrix}X\\ 1\end{pmatrix}=\begin{pmatrix}T & X_0\\ \mathbf{0} & 1\end{pmatrix}\begin{pmatrix}\widetilde{X}'\\ 1\end{pmatrix},$$

其中 T 为正交矩阵（若保持定向，要求 $\det(T)=1$）。

$$\Rightarrow \widetilde{F}(\widetilde{x},\widetilde{y})=\begin{pmatrix}\widetilde{X} & 1\end{pmatrix}\begin{pmatrix}T^{\mathrm{T}} & \mathbf{0}\\ X_0^{\mathrm{T}} & 1\end{pmatrix}\begin{pmatrix}A & B\\ B^{\mathrm{T}} & c\end{pmatrix}\begin{pmatrix}T & X_0\\ \mathbf{0} & 1\end{pmatrix}\begin{pmatrix}\widetilde{X}\\ 1\end{pmatrix}$$

$$=\begin{pmatrix}\widetilde{X} & 1\end{pmatrix}\begin{pmatrix}T^{\mathrm{T}}AT & T^{\mathrm{T}}AX_0+T^{\mathrm{T}}B\\ X_0^{T}AT+B^{\mathrm{T}}T & X_0^{\mathrm{T}}AX_0+2B^{\mathrm{T}}X_0+c\end{pmatrix}\begin{pmatrix}\widetilde{X}\\ 1\end{pmatrix}$$

（式中利用 $X_0^{\mathrm{T}}B=B^{\mathrm{T}}X_0$）

命题 7.6

对于 $F(x,y)$ 以下三个数：

$$I_1 = \mathrm{tr}(A) = a_{11} + a_{22},\ I_2 = \det(A) = a_{11}a_{22} - a_{12}^2,\ I_3 = \begin{vmatrix} A & B \\ B^{\mathrm{T}} & c \end{vmatrix}$$

在直角坐标变换下保持不变，称为二次多项式 $F(x,y)$ 的正交不变量，简称不变量。

利用I_1, I_2, I_3可以对曲线$F(x,y)$进行分类：

情形 (I) 当 $I_2 = |A| = \lambda_1\lambda_2 \neq 0$ 时 $\Rightarrow A$ 可逆

依据 I_1, I_2, I_3 不同取值，二次曲线 C 可分为 5 类：

(1) $I_2 > 0$，I_3 与 I_1 异号 $\Rightarrow C$ 是椭圆，$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$；
(2) $I_2 > 0$，I_3 与 I_1 同号 $\Rightarrow C$ 是虚椭圆，$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$；
(3) $I_2 > 0$，$I_3 = 0 \Rightarrow C$ 退化成点，$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$；
(4) $I_2 < 0$，$I_3 \neq 0 \Rightarrow C$ 是双曲线，$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$；
(5) $I_2 < 0$，$I_3 = 0 \Rightarrow C$ 是一对相交直线，$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$。

几何量：二次曲线的中心、对称轴、渐近线等

情形 (II) 当 $I_2 = |A| = \lambda_1\lambda_2 = 0$ 时 $\Rightarrow \lambda_2 = I_1$(设 $\lambda_1 = 0, \lambda_2 \neq 0$)

依据 I_1, I_2, I_3 不同取值，二次曲线 C 可分为 4 类：

(1) $I_2 = 0$，$I_3 \neq 0 \Rightarrow C$ 是抛物线，$y^2 - 2px = 0$；
(2) $I_2 = I_3 = 0$，$K_1 < 0 \Rightarrow C$ 是一对平行直线，$y^2 - a^2 = 0$；
(3) $I_2 = I_3 = 0$，$K_1 > 0 \Rightarrow C$ 是一对虚平行直线，$y^2 + a^2 = 0$；
(4) $I_2 = I_3 = 0$，$K_1 = 0 \Rightarrow C$ 是一对重合直线，$y^2 = 0$，

其中 $K_1 = \begin{vmatrix} a_{11} & b_1 \\ b_1 & c \end{vmatrix} + \begin{vmatrix} a_{22} & b_2 \\ b_2 & c \end{vmatrix}$。

设 n 维欧氏空间 V 取定一个直角坐标系：$[O;\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n]$，二次超曲面的一般方程为

$$F(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j+2\sum_{i=1}^{n}b_ix_i+c=0,$$

其中 $a_{ij}=a_{ji}\ (i,j=1,2,\ldots,n)$。

采用齐次坐标，则

$$F(x_1,\ldots,x_n)=\begin{pmatrix}x_1 & \cdots & x_n & 1\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1n} & b_1\\ \vdots & & \vdots & \vdots\\ a_{n1} & \cdots & a_{nn} & b_n\\ b_1 & \cdots & b_n & c\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\\ 1\end{pmatrix}$$

$$=\begin{pmatrix}X^{\mathrm{T}} & 1\end{pmatrix}\begin{pmatrix}A & B\\ B^{\mathrm{T}} & c\end{pmatrix}\begin{pmatrix}X\\ 1\end{pmatrix}$$

如果取一个新的直角坐标系：$[\widetilde{O};\widetilde{\boldsymbol{e}}_1,\widetilde{\boldsymbol{e}}_2,\ldots,\widetilde{\boldsymbol{e}}_n]$，则

$$X=T\widetilde{X}+X_0\Longleftrightarrow\begin{pmatrix}X\\ 1\end{pmatrix}=\begin{pmatrix}T & X_0\\ \mathbf{0} & 1\end{pmatrix}\begin{pmatrix}\widetilde{X}'\\ 1\end{pmatrix},$$

其中 T 为正交矩阵（若保持定向，要求 $\det(T)=1$）。

$$\Rightarrow\widetilde{F}(\widetilde{x}_1,\widetilde{x}_2,\ldots,\widetilde{x}_n)=\begin{pmatrix}\widetilde{X} & 1\end{pmatrix}\begin{pmatrix}T^{\mathrm{T}} & \mathbf{0}\\ X_0^{\mathrm{T}} & 1\end{pmatrix}\begin{pmatrix}A & B\\ B^{\mathrm{T}} & c\end{pmatrix}\begin{pmatrix}T & X_0\\ \mathbf{0} & 1\end{pmatrix}\begin{pmatrix}\widetilde{X}\\ 1\end{pmatrix}$$

$$=\begin{pmatrix}\widetilde{X} & 1\end{pmatrix}\begin{pmatrix}T^{\mathrm{T}}AT & T^{\mathrm{T}}AX_0+T^{\mathrm{T}}B\\ X_0^{\mathrm{T}}AT+B^{\mathrm{T}}T & X_0^{\mathrm{T}}AX_0+2B^{\mathrm{T}}X_0+c\end{pmatrix}\begin{pmatrix}\widetilde{X}\\ 1\end{pmatrix}$$

（式中利用 $X_0^{\mathrm{T}}B=B^{\mathrm{T}}X_0$）

定理 7.7

对于实对称矩阵 $A \in \mathbb{R}^{n\times n}$，列矩阵 $B \in \mathbb{R}^n$ 以及实数 c，一定存在正交矩阵 T 以及列矩阵 $X_0 \in \mathbb{R}^n$ 使得矩阵

$$\begin{pmatrix} T^{\mathrm{T}} & \mathbf{0} \\ X_0^{\mathrm{T}} & 1 \end{pmatrix} \begin{pmatrix} A & B \\ B^{\mathrm{T}} & c \end{pmatrix} \begin{pmatrix} T & X_0 \\ \mathbf{0} & 1 \end{pmatrix}$$

具有以下简化形式：

(1) 当 $\mathrm{rank}(A \vdots B) = \mathrm{rank}(A)$ 时：

$$\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_n & \\ & & & c' \end{pmatrix}$$

其中 $\lambda_1, \ldots, \lambda_n$ 是矩阵 A 的特征值，c' 是一个实数；

定理 7.7

(2) 当 $\mathrm{rank}(A \vdots B) < \mathrm{rank}(A)$ 时：

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_k & & & & \\ & & & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & c' \\ & & & & & c' & 0 \end{pmatrix}$$

其中 $\lambda_1, \ldots, \lambda_k$ $(k < n)$ 是矩阵 A 的非零特征值，c' 是一个非零实数。

命题 7.8

n 维欧氏空间中的二次超曲面在适当选取直角坐标系后可以简化成以下两种方程之一：

$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 = c,\ \lambda_1, \ldots, \lambda_n, c \in \mathbb{R},$$

或

$$\lambda_1 x_1^2 + \cdots + \lambda_k x_k^2 = 2cx_n,\ k < n, \lambda_1, \ldots, \lambda_k, c \text{ 均为非零实数。}$$

根据定理 7.7 这个结论是显然的。
所以当$n=3$时，一共有17类标准方程。

正定二次型

定义 7.4

n 元实二次型

$$Q(x_1,\dots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$$

称为**正定二次型**，如果对任意非零向量 $\boldsymbol{x}\in\mathbb{R}^n$ 均有

$$Q(x_1,\dots,x_n)>0.$$

正定的二次型对应的矩阵 A 称为**正定矩阵**，简记为 $A>0$。

内积的度量矩阵是正定矩阵。

定理 7.9

n 元实二次型

$$Q(x_1,\dots,x_n)=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$$

正定的充要条件是 Q 的正惯性指数为 n。

$\Longleftrightarrow$ 实对称方阵 A 正定的充要条件是 A 相合于单位阵。

根据定理 7.4：

$\Rightarrow$ $\exists$ 可逆阵 P 使得

$$Q(x_1,\dots,x_n)\Big|_{\boldsymbol{x}=P\boldsymbol{y}}=y_1^2+\cdots+y_r^2-y_{r+1}^2\cdots-y_{r+s}^2,$$

且 $\operatorname{rank}(A)=r+s$，$r$ 为正惯性指数，s 为负惯性指数。

$\Rightarrow$ 矩阵 A 正定 $\Leftrightarrow r=n,s=0$

定理 7.10

设方阵 $A=(a_{ij})_{n\times n}$ 是实对称方阵，则下述命题等价：

(1) 方阵 A 是正定的；
(2) 方阵 A 的每个特征值都是正的；
(3) 存在正定对称方阵 A_1，使得 $A=A_1^2$；
(4) 存在可逆方阵 P，使得 $A=P^{\mathrm{T}}P$；
(5) 方阵 A 的每个主子式都是正的；
(6) 方阵 A 的每个顺序主子式都是正的。

Pf: (1)=>(2):

A为实对称阵 => $\exists$ 正交阵 P s.t. $P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$

$$
\begin{aligned}
Q(x_1,x_2,\cdots,x_n) &= X^TAX\\
&=(PY)^TAPY\\
&=Y^TP^TAPY\\
&=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2
\end{aligned}
$$

取 $\vec{x_j}=P(0,\cdots,\underset{\uparrow j}{1},\cdots,0)^T \Rightarrow Q(x_1,x_2,\cdots,x_n)=\lambda_j>0,\ j=1,2,3,\cdots,n$

(2)=>(3) $P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$

$$
\begin{aligned}
A &= P\,\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)P^{-1}\\
&=P\,\mathrm{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})P^{-1}P\,\mathrm{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})P^{-1}\\
&=A_1^2
\end{aligned}
$$

(3)=>(4) $A_1^T=A_1$ 显然成立 故: $A=A_1^TA_1$

(4)=>(5)
$$
\left|A\begin{pmatrix} i_1 & i_2 & i_3\cdots i_k\\ i_1 & i_2 & i_3\cdots i_k\end{pmatrix}\right| \overset{\text{Binet-Cauchy公式}}{=} \sum_{1\le j_1<j_2<\cdots<j_k\le n}\left|A_1^T\begin{pmatrix} i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|\left|A_1\begin{pmatrix} i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|
$$
$$
=\sum_{1\le j_1<\cdots<j_k\le n}\left|A_1\begin{pmatrix} i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|^2\ge 0
$$

若 $\left|A\begin{pmatrix} i_1 i_2\cdots i_k\\ j_1 j_2\cdots j_k\end{pmatrix}\right|=0$ 则 $\left|A_1\begin{pmatrix} i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|=0$ 对 $\forall\ 1\le j_1<\cdots<j_k\le n$

则 $\det(A_1)=0$，A_1不可逆 和题意不符

∴ A的每个主子式都是正的

(5) ⇒ (6) 显然

(6) ⇒ (1) 数学归纳法

1° $n=1$ 时显然是成立的

2° 假设 $n-1$ 时成立，A_1 为 $n-1$ 阶实对称矩阵

令 $A_{n\times n} = \begin{pmatrix} A_1 & \vec{y}_1^T \\ \vec{y}_1 & a_{nn} \end{pmatrix}$

$$\begin{pmatrix} I_{n-1} & \\ -\vec{y}_1 A_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_1 & \vec{y}_1^T \\ \vec{y}_1 & a_{nn} \end{pmatrix} \begin{pmatrix} I_{n-1} & -A_1^{-1}\vec{y}_1^T \\ & 1 \end{pmatrix}$$

$$= \begin{pmatrix} A_1 & \\ & a_{nn} - \vec{y}_1 A_1^{-1} \vec{y}_1^T \end{pmatrix}$$

$\det(A) = \det(A_1)\left(a_{nn} - \vec{y}_1 A_1^{-1}\vec{y}_1^T\right) > 0 \Rightarrow a_{nn} - \vec{y}_1 A_1^{-1}\vec{y}_1^T > 0$

令 $\vec{x} = \boxed{\begin{pmatrix} I_{n-1} & 0 \\ -A_1^{-1}\vec{y} & 1 \end{pmatrix}} \begin{pmatrix} \vec{z}_{n-1} \\ z_n \end{pmatrix}$

$$\therefore \vec{x}^T A \vec{x} = \underline{\vec{z}_{n-1}^T A_1 \vec{z}_{n-1}} + \left(a_{nn} - \vec{y}_1 A^{-1} \vec{y}_1^T\right) z_n^2 \geq 0$$

故A是正定的

定理 7.11

设 A, B 是两个 n 阶实对称矩阵，其中 A 是正定的，则存在 n 阶可逆矩阵 P 使得 A 和 B 同时相合于对角阵，即

$$P^T A P = I,\ P^T B P = D,$$

其中 D 为对角阵。

证明：

显然存在 P_1 使得：$P_1^T A P_1 = I$，令 $B_1 = P_1^T B P_1$

此时 B_1 仍然是实对称矩阵，故存在正交矩阵 P_2 使得 $P_2^T B_1 P_2 = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ D

令 $P = P_1 P_2 \Rightarrow PAP = P_2^T P_1^T A P_1 P_2 = P_2^T I P_2 = I$

$PBP = P_2^T P_1^T A P_1 P_2 = P_2^T B_1 P_2 = D$

n 个变元的二次型 $Q(x_1, \ldots, x_n) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ 的分类：

(1) 正定二次型：$r = n, s = 0 \Leftrightarrow$ 正定矩阵；
(2) 半正定二次型：$r < n, s = 0 \Leftrightarrow$ 半正定矩阵，简记为：$A \geqslant 0$；
(3) 负定二次型：$r = 0, s = n \Leftrightarrow$ 负定矩阵，简记为：$A < 0$；
(4) 半负定二次型：$r = n, s < n \Leftrightarrow$ 半负定矩阵，简记为：$A \leqslant 0$；
(5) 不定二次型：其他情形 $r = n, s = 0 \Leftrightarrow$ 不定矩阵；

思考：负定矩阵 A 的顺序主子式满足什么条件?

例 7.5

设 A 为 3 阶实对称矩阵，且满足条件：$A^2 + 2A = \mathbf{0}$。已知 $\mathrm{rank}(A) = 2$，求

(1) 矩阵 A 的全部特征值；
(2) 当 k 为何值时，矩阵 $A + kI$ 为正定矩阵。

例 7.6

设二次型 $Q(x_1, x_2, x_3) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = ax_1^2 + 2x_2^2 - 2x_3^2 + 2bx_1x_3\ (b > 0)$。已知矩阵 A 的全部特征值之和为 1，全部特征值之积为 -12，

(1) 求 a，b 的值；
(2) 利用正交变换将二次型 $Q(x_1, x_2, x_3)$ 化为标准形，并写出所用的正交变换和对应的正交矩阵。

例 7.7

已知二次型 $Q(x_1, x_2, x_3) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ 在正交变换 $\boldsymbol{x} = P\boldsymbol{y}$ 下的标准形为 $y_1^2 + y_2^2$ 且 P 的第三列为 $\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}^{\mathrm{T}}$，

(1) 求矩阵 A；
(2) 证明 $A + I$ 为正定矩阵。

Conclusion

1. 线性代数的 3 种等价关系：

（1） 相抵：

定义 3.14

设 A, B 是 $m \times n$ 矩阵，如果存在可逆方阵 P 和 Q 使得

$$B = PAQ,$$

则称 A 和 B **相抵**。

（2） 相似 ：

定义 5.3

设 A, B 为数域 F 上的两个 n 阶方阵，如果存在数域 F 上的 n 阶可逆方阵 T，使得 $B = T^{-1}AT$，则称 A 与 B（在数域 F 上）**相似**，记为 $A \sim B$。

（3） 相合

定义 6.4

设 $A, B \in F^{n \times n}$，若存在数域 F 上的 n 阶可逆方阵 P，使得 $B = P^{\mathrm{T}}AP$，则称 A 与 B（在数域 F 上）**相合**，记作 $A \sim B$。

2．实对称矩阵的性质

3．$A - I$模型：设A为n阶实方阵，且$A^2 = A$，证明：A可相似对角化。

（$A^2 = A$可以替换为$A^2 - I = 0, A^2 - 3A + 2I = 0$）