你永远不知道一个强迫症能干出什么事情

倪兴程[1]

2024 年 10 月 18 日

[1]Email: 19975022383@163.com

Todo list

目录

Chapter 1

线性空间

1.1 线性空间

1.1.1 线性空间的概念与基本性质

Definition 1.1. 设 S 是一个非空集合，$S \times S$ 是 S 与自身的一个 *Cartesian product*（定义可见**??**），则 $f : S \times S \to S$ 称为 S 上一个二元代数运算 *(binary algebraic operation)*，简称为 S 上的一个运算。

线性空间的定义

Definition 1.2. 设 X 是一个非空集合，F 是一个域。如果 X 上有一个运算，即 $f : (\alpha, \beta) \to \gamma(\alpha, \beta, \gamma \in X)$，将该运算称为加法 *(addition)*，把 γ 称为 α 与 β 的和 *(sum)*，记作 $\alpha + \beta = \gamma$；同时 F 与 X 有一个运算，即 $g : (k, \alpha) \to \delta(k \in F,\ \alpha, \delta \in X)$，将该运算称为纯量乘法 *(scalar multiplication)*，把 δ 称为 k 与 α 的纯量乘积 *(scalar multiple)*，记作 $k\alpha = \delta$。若上述两个运算还满足以下 8 条运算法则：

域的概念

1. $\forall\, \alpha, \beta \in X,\ \alpha + \beta = \beta + \alpha$；
2. $\forall\, \alpha, \beta, \gamma \in X,\ (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$；
3. X 中有一个元素，记作 $\mathbf{0}$，称为 X 的零元 *(zero vector)*，它使得：
$$\forall\, \alpha \in X,\ \alpha + \mathbf{0} = \alpha$$
4. 对于任意的 $\alpha \in X$，存在与之对应的 $\beta \in X$，称为 α 的负元 *(additive inverse)*，记作 $-\alpha$，它使得：
$$\alpha + \beta = \mathbf{0}$$
5. F 中的单位元 1 满足 $\forall\, \alpha \in X,\ 1\alpha = \alpha$；
6. $\forall\, \alpha \in X,\ k, l \in F,\ (kl)\alpha = k(l\alpha)$；

7. $\forall\, \alpha \in X,\ k, l \in F,\ (k+l)\alpha = k\alpha + l\alpha$；

8. $\forall\, \alpha, \beta \in X,\ k \in F,\ k(\alpha+\beta) = k\alpha + k\beta$。

那么称 X 是域 F 上的一个线性空间 *(linear space)*，称 X 中的元素为向量 *(vector)*。

线性空间的基本性质

Property 1.1.1. 域 F 上的线性空间 X 具有如下性质：

1. X 中的零元是唯一的；

2. X 中每个元素的负元是唯一的；

3. $\forall\, \alpha \in X,\ 0\alpha = \mathbf{0}$；

4. $\forall\, k \in F,\ k\mathbf{0} = \mathbf{0}$；

5. 设 $k \in F,\ \alpha \in X$。如果 $k\alpha = \mathbf{0}$，那么 $k = 0$ 或 $\alpha = \mathbf{0}$。

6. $\forall\, \alpha \in X,\ (-1)\alpha = -\alpha$；

Proof. (1) 假设 X 中有两个零元 $\mathbf{0}_1, \mathbf{0}_2$ 且 $\mathbf{0}_1 \neq \mathbf{0}_2$，由线性空间运算法则 (3) 可得：

$$\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1,\ \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2$$

而由线性空间运算法则 (1) 可得：

$$\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1$$

于是 $\mathbf{0}_1 = \mathbf{0}_2$，产生矛盾，所以 X 中的零元是唯一的。

(2) 任取 X 中的一个元素 α，假设它有两个负元 β_1, β_2。由线性空间运算法则 (4)(3)(2) 可得：

$$(\beta_1 + \alpha) + \beta_2 = \mathbf{0} + \beta_2 = \beta_2$$
$$(\beta_1 + \alpha) + \beta_2 = \beta_1 + (\alpha + \beta_2) = \beta_1 + \mathbf{0} = \beta_1$$

所以 $\beta_1 = \beta_2$，产生矛盾。由 α 的任意性，X 中每个元素的负元都是唯一的。

(3) 由线性空间运算法则 (7) 可得：

$$0\alpha + 0\alpha = (0+0)\alpha = 0\alpha$$

两边同时加上 -0α 可得：

$$0\alpha + 0\alpha + (-0\alpha) = 0\alpha + (-0\alpha)$$

由线性空间运算法则 (2)(4) 和 (3) 可得：

$$0\alpha = \mathbf{0}$$

(4) 由线性空间运算法则 (8) 和 (3) 可得：

$$k\mathbf{0} + k\mathbf{0} = k(\mathbf{0} + \mathbf{0}) = k\mathbf{0}$$

两边加上 $-k\mathbf{0}$ 再由线性空间运算法则 (2)(4) 和 (3) 可得：

$$k\mathbf{0} = \mathbf{0}$$

(5) 如果 $k \neq 0$，依次由线性空间运算法则 (5)、(6) 和线性空间基本性质 (4) 可得：

$$\alpha = 1\alpha = (k^{-1}k)\alpha = k^{-1}(k\alpha) = k^{-1}\mathbf{0} = \mathbf{0}$$

(6) 由线性空间运算法则 (5) 与 (7) 以及线性空间基本性质 (3) 可得：

$$\alpha + (-1)\alpha = 1\alpha + (-1)\alpha = (1-1)\alpha = 0\alpha = \mathbf{0}$$

再由负元的定义，$(-1)\alpha = -\alpha$。 □

Definition 1.3. 设 X 是域 F 上的线性空间。由性质 *1.1.1(2)*，定义 $f : (\alpha, \beta) \to \alpha + (-\beta) \in X (\alpha, \beta \in X)$，将该运算称为减法 *(subtraction)*，把 $\alpha + (-\beta)$ 称为 α 与 β 的差 *(difference)*，记作 $\alpha - \beta = \alpha + (-\beta)$。

1.1.2 线性相关与线性无关

Definition 1.4. X 是域 F 上的线性空间。按照一定顺序写出的有限多个向量（允许有相同的向量）称为 X 的一个向量组 *(set of vectors)*，如 $\alpha_1, \alpha_2, \ldots, \alpha_n$。

Definition 1.5. X 是域 F 上的线性空间。对于 X 中的一组向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 和 F 中的一组元素 $k_1, k_2, \ldots, k_n$，作纯量乘法和加法得到：

$$\sum_{i=1}^{n} k_i \alpha_i \in X$$

称该向量为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的一个线性组合 *(linear combination)*。

Definition 1.6. X 是域 F 上的线性空间。若 $\beta \in X$ 可以表示成向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的一个线性组合，则称 β 可以由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ **线性表出**。

Definition 1.7. X 是域 F 上的线性空间。按照如下方式定义 X 中对象的与：

Theorem 1.1. 在域 F 上的线性空间 X 中，如果向量组的一个部分组线性相关，那么这个向量组线性相关。

Proof. 取 X 中的向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$，其部分组 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_m}$，$i_m, m \leqslant n$ 线性相关，即 F 中存在不全为 0 的一组元素 $k_{i_1}, k_{i_2}, \ldots, k_{i_m}$ 使得：

$$k_{i_1}\alpha_{i_1} + k_{i_2}\alpha_{i_2} + \cdots + k_{i_m}a_{i_m} = \mathbf{0}$$

研究对象	线性相关	线性无关
X 中的向量组 $\alpha_1,\alpha_2,\dots,\alpha_n$	F 中有不全为 0 的元素 $k_1,k_2,\dots,k_n$ 使得 $\sum\limits_{i=1}^{n}k_i\alpha_i=\mathbf{0}$	从 $\sum\limits_{i=1}^{n}k_i\alpha_i=\mathbf{0}$ 可以推出 $k_1=k_2=\cdots=k_n=0$
空集		定义空集是线性无关的
X 的非空有限子集	给这个子集的元素一种编号所得的向量组线性相关	给这个子集的元素一种编号所得的向量组线性无关
X 的无限子集 W	W 有一个有限子集线性相关	W 的任一有限子集都线性无关

在：

$$l_1\alpha_1+l_2\alpha_2+\cdots+l_n\alpha_n$$

中取 $l_{i_j}=k_{i_j}$，$j=1,2,\dots,m$，其余系数为 0，则 $l_1,l_2,\dots,l_n$ 不全为 0 同时上式值为 $\mathbf{0}$，即向量组 $\alpha_1,\alpha_2,\dots,\alpha_n$ 线性相关。 □

Corollary 1.1. 在域 F 上的线性空间 X 中，包含 $\mathbf{0}$ 的向量组是线性相关的。

Proof. $\mathbf{0}$ 作为向量组的部分组是线性相关的。 □

Theorem 1.2. 在域 F 上的线性空间 X 中，元素个数大于 1 的向量集 W 线性相关当且仅当 W 中至少有一个向量可以由其余向量中的有限多个线性表出，从而 W 线性无关当且仅当 W 中的每一个向量都不能由其余向量中的有限多个线性表出。

Proof. 由线性相关的定义立即得出。 □

Definition 1.8. X 是域 F 上的线性空间。设向量 $\beta\in X$ 可以由向量集 $W\subseteq X$ 中有限多个向量线性表出，则称 β 可以由向量集 W 线性表出。

Theorem 1.3. X 是域 F 上的线性空间，向量 $\beta\in X$ 可以由向量集 $W\subseteq X$ 线性表出，则表示方法唯一的充分必要条件是 W 线性无关。

Proof. **(1) 充分性：**因为 β 可以由向量集 W 线性表出，假设此时 β 有以下两种表出方式：

$$\beta=k_1\alpha_1+\cdots+k_r\alpha_r+k_{r+1}u_1+\cdots+k_{r+s}u_s$$
$$\beta=l_1\alpha_1+\cdots+l_r\alpha_r+l_{r+1}v_1+\cdots+l_{r+t}v_t$$

其中 $\alpha_1,\alpha_2,\dots,\alpha_r,u_1,u_2,\dots,u_s,v_1,v_2,\dots,v_t\in W,\ k_1,\dots,k_{r+s},l_1,\dots.l_{r+t}\in F,\ r,s,t\geqslant 0$。二式作差可得：

$$\mathbf{0}=(k_1-l_1)\alpha_1+\cdots+(k_r-l_r)\alpha_r+k_{r+1}u_1+\cdots+k_{r+s}u_s-l_{r+1}v_1-\cdots-l_{r+t}v_t$$

因为 W 线性无关，所以向量组 $\alpha_1,\alpha_2,\dots,\alpha_r,u_1,u_2,\dots,u_s,v_1,v_2,\dots,v_t$ 线性无关，于是：

$$k_1-l_1=0,\cdots,k_r-l_r=0,k_{r+1}=0,\cdots,k_{r+s}=0,l_{r+1}=0,\dots,l_{r+t}=0$$

所以两个表出方式完全相同，β 由 W 线性表出的表示方法唯一。

(2) 必要性：如果 W 线性相关，则 W 有一个有限子集 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关，于是 F 中有不全为 0 的元素 $k_1, k_2, \dots, k_n$ 使得：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = \mathbf{0}$$

由于 β 可以由 W 线性表出，所以：

$$\beta = l_1\alpha_1 + \cdots + l_n\alpha_n + l_{n+1}v_1 + \cdots + l_{n+s}v_s$$

其中 $l_i \in F,\ i = 1, 2, \dots, n;\ v_j \in X,\ j = 1, 2, \dots, s$。将上两式相加可得：

$$\beta = (l_1 + k_1)\alpha_1 + \cdots + (l_n + k_n)\alpha_n + l_{n+1}v_1 + \cdots + l_{n+s}v_s$$

因为 $k_1, k_2, \dots, k_n$ 不全为 0，所以有序元素组：

$$(l_1, \dots, l_n, l_{n+1}, \dots, l_{n+s}) \neq (l_1 + k_1, \dots, l_n + k_n, l_{n+1}, \dots, l_{n+s})$$

于是 β 由 W 线性表出的方式不唯一，矛盾，所以 W 线性无关。 □

Theorem 1.4. 在域 F 上的线性空间 X 中，若向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关，则向量 β 可以由向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表出的充分必要条件为 $\alpha_1, \alpha_2, \dots, \alpha_n, \beta$ 线性相关。

Proof. **(1) 充分性**：因为 $\alpha_1, \alpha_2, \dots, \alpha_n, \beta$ 线性相关，所以域 F 中存在不全为 0 的元素 $k_1, k_2, \dots, k_n, l$ 使得：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n + l\beta = \mathbf{0}$$

若 $l = 0$，则域 F 中存在不全为 0 的元素 $k_1, k_2, \dots, k_n$ 使得：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = \mathbf{0}$$

这与向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关矛盾，所以 $l \neq 0$。于是：

$$\beta = -\frac{k_1}{l}\alpha_1 - \frac{k_2}{l}\alpha_2 - \cdots - \frac{k_n}{l}\alpha_n$$

即向量 β 可以由向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表出。

(2) 必要性：因为 β 可以由向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表出，所以域 F 中存在不全为 0 的元素 $k_1, k_2, \dots, k_n, l$ 使得：

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

移项即可得到：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n - \beta = \mathbf{0}$$

因为 $-1 \neq 0$，所以 $\alpha_1, \alpha_2, \dots, \alpha_n, \beta$ 线性相关。 □

极大线性无关组与秩

Definition 1.9. 设 X 是域 F 上的一个线性空间。向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 中的一个部分组若满足下述条件：

1. 本身线性无关；
2. 若该部分组不等于向量组，则从向量组的其余向量中任取一个向量添加进该部分组都将使部分组线性相关。若该部分组等于向量组，则跳过此条件。

则称该部分组是向量组的一个极大线性无关组 *(maximal linearly independent system)*。

Definition 1.10. 如果向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的每一个向量都可以由向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性表出，那么称向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性表出。如果向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 和向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 可以互相线性表出，则称两个向量组等价 *(equivalent)*，记作：

$$\{\alpha_1, \alpha_2, \ldots, \alpha_s\} \cong \{\beta_1, \beta_2, \ldots, \beta_r\}$$

Theorem 1.5. X 是数域 K 上的一个线性空间。X 中向量组的等价是 X 中向量组的一个等价关系。

Proof. (1) 反身性与 (2) 对称性显然成立。

(3) 传递性：若向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性表出，且向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 可以由向量组 $\gamma_1, \gamma_2, \ldots, \gamma_t$ 线性表出，于是有：

$$\alpha_i = \sum_{j=1}^{r} k_{j_i} \beta_j, \ i = 1, 2, \ldots, s \quad \beta_j = \sum_{l=1}^{t} q_{l_j} \gamma_l, \ j = 1, 2, \ldots, r$$

于是：

$$\alpha_i = \sum_{j=1}^{r} k_{j_i} \left(\sum_{l=1}^{t} q_{l_j} \gamma_l \right) = \sum_{j=1}^{r} \sum_{l=1}^{t} k_{j_i} q_{l_j} \gamma_l = \sum_{l=1}^{t} \left(\sum_{j=1}^{r} k_{j_i} q_{l_j} \right) \gamma_l, \ i = 1, 2, \ldots, s$$

即向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由向量组 $\gamma_1, \gamma_2, \ldots, \gamma_t$ 线性表出。传递性得证。 □

Theorem 1.6. 一个向量组与它的任意一个极大线性无关组等价。

Proof. 设 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 是一个向量组，任取它的一个极大线性无关组 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$。显然 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$ 可以由 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性表出。由极大线性无关组的定义与定理 1.4可直接得到 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$ 线性表出。 □

Corollary 1.2. 一个向量组的任意两个极大线性无关组等价。

Proof. 由上一个定理以及向量组等价的对称性与传递性可直接推出。 □

Theorem 1.7. 在域 F 上的线性空间 X 中，若向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 可以由向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性表出，同时 $r > s$，那么向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性相关。

Proof. 考虑方程：

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_r\beta_r = \mathbf{0}$$

因为向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 可以由向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性表出，所以：

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1s}\alpha_s \\ \beta_2 = a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2s}\alpha_s \\ \vdots \\ \beta_r = a_{r1}\alpha_1 + a_{r2}\alpha_2 + \cdots + a_{rs}\alpha_s \end{cases}$$

则有：

$$\begin{aligned} k_1\beta_1 + k_2\beta_2 + \cdots + k_r\beta_r &= k_1(a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1s}\alpha_s) \\ &\quad + k_2(a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2s}\alpha_s) + \cdots \\ &\quad + k_r(a_{r1}\alpha_1 + a_{r2}\alpha_2 + \cdots + a_{rs}\alpha_s) \\ &= (k_1a_{11} + k_2a_{21} + \cdots + k_ra_{r1})\alpha_1 \\ &\quad + (k_1a_{12} + k_2a_{22} + \cdots + k_ra_{r2})\alpha_2 + \cdots \\ &\quad + (k_1a_{1s} + k_2a_{2s} + \cdots + k_ra_{rs})\alpha_s \\ &= \mathbf{0} \end{aligned}$$

将上式看作 $k_1, k_2, \ldots, k_r$ 的线性方程，考虑如下齐次线性方程组：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rs} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

因为 $s < r$，所以上述齐次线性方程组必有非零解。取它的一个非零解 $(k_1, k_2, \ldots, k_r)$，由性质 1.1.1(3) 和线性空间运算法则 (3) 即可得：

齐次方程组的解

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_r\beta_r = 0\alpha_1 + 0\alpha_2 + \cdots + 0\alpha_s = \mathbf{0}$$

于是向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性相关。 □

Corollary 1.3. 上述定理可以得到如下推论：

1. 在域 F 上的线性空间 X 中，若向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 可以由向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性表出，同时向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性无关，就有 $r \leqslant s$。

2. 等价的线性无关的向量组所含向量的个数相同。

3. 一个向量组的任意两个极大线性无关组所含向量的个数相同。

Proof. (1) 是上述定理的逆否命题，(2) 可由 (1) 直接得到，(3) 可由 (2) 直接得到。 □

Definition 1.11. 向量组的极大线性无关组所含向量的个数称为这个向量组的秩 *(rank)*。把向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的秩记作 $\operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_s\}$。全由零向量组成的向量组的秩规定为 0。

Theorem 1.8. 从向量组秩的定义可推出下述定理：

1. 在域 F 上的线性空间 X 中，向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性无关的充分必要条件是它的秩等于它所含向量的个数。
2. 如果向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性表出，那么：
$$\operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_s\} \leqslant \operatorname{rank}\{\beta_1, \beta_2, \ldots, \beta_r\}$$
3. 等价的向量组具有相等的秩。

Proof. (1) $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性无关 ⇔ 极大线性无关组就是自身 ⇔ 秩等于所含向量的个数。

(2) 任取向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的一个极大线性无关组 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$，再取向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 的一个极大线性无关组 $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_m}$。因为向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可以由向量组 $\beta_1, \beta_2, \ldots, \beta_r$ 线性表出，所以 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$ 可以由 $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_m}$ 线性表出，因为 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$ 线性无关，由推论 1.3(1) 可得，$n \leqslant m$，即 $\operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_s\} \leqslant \operatorname{rank}\{\beta_1, \beta_2, \ldots, \beta_r\}$。

(3) 由 (2) 可直接推得。 □

基与维数

Definition 1.12. 设 X 是域 F 上的一个线性空间。X 中的向量集 S 若满足下述条件：

1. 本身线性无关；
2. X 中的每一个向量都可以由 S 中有限多个向量线性表出。

则称 S 是向量组的一个基 *(basis)*。

Theorem 1.9. 任一域 F 上的任一线性空间 X 都有一个基。

该定理的证明不提供，涉及 Zorn 引理。

Definition 1.13. X 是域 F 上的线性空间。如果 X 的一个基是由有限多个向量组成的，那么称 X 是有限维的 *(finite-dimensional)*；如果 X 有一个基含有无穷多个向量，则称 X 是无限维的 *(infinite-dimensional)*。

Theorem 1.10. 如果域 F 上的线性空间 X 是有限维的，那么 X 的任意两个基所含向量的个数相同。

Proof. 由有限维线性空间的定义，X 存在一个基只有有限多个向量，记为 $\alpha_1, \alpha_2, \ldots, \alpha_n$。再取 X 的一个基 S，从中取出 $n+1$ 个向量 $\beta_1, \beta_2, \ldots, \beta_{n+1}$（考虑 S 可能含有无数个向量的情况）。因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的基，所以 $\beta_1, \beta_2, \ldots, \beta_{n+1}$ 可以由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出。因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关，所以 $\beta_1, \beta_2, \ldots, \beta_{n+1}$ 线性相关，而此时 S 也应线性相关，矛盾，所以 S 中的元素小于 $n+1$ 个。设 $S = \{\beta_1, \beta_2, \ldots, \beta_m\}$，则 S 与 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 都线性无关且二者等价，由推论 1.3(2)，它们具有相同的秩，即所含向量个数相同。由 S 的任意性与向量组等价的传递性、对称性，X 的任意两个基所含向量的个数相同。 □

Corollary 1.4. 如果域 F 上的线性空间 X 是无限维的，那么 X 的任意一个基都含有无穷多个向量。

Proof. 如果 X 有一个基由有限多个向量组成，那么 X 也是一个有限维线性空间，而有限维线性空间所有基所含向量的个数都相同，那么 X 就不可能有一个基含有无穷多个向量，与无穷维线性空间的定义矛盾。 □

Definition 1.14. X 是域 F 上的线性空间。如果 X 是有限维的，那么把 X 的基所含向量的个数称为 X 的，记作 $\dim V$；如果 X 是无限维的，那么记 $\dim V = +\infty$。

Theorem 1.11. n 维线性空间中任意 $n+1$ 个向量都线性相关。

Proof. 任意 $n+1$ 个向量都可以由一组基线性表出，而每一组基所含向量个数都是 n，由定理 1.7，这 $n+1$ 个向量线性相关。。 □

Theorem 1.12. 设 X 是域 F 上的 n 维线性空间，则：

1. X 中任意 n 个线性无关的向量都是 X 的一组基；
2. 如果 X 中任一向量都可以由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出，则 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基。

Proof. (1) 任取 X 中 n 个线性无关的向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$，对任意的 $\beta \in X$，由上一个定理，向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta$ 线性相关。由定理 1.4，β 可以由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出。由基的定义，$\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基。由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的任意性，命题成立。

(2) 取 X 的一组基 $\beta_1, \beta_2, \ldots, \beta_n$，则 $\beta_1, \beta_2, \ldots, \beta_n$ 可以由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出。由定理 1.8(2)，$n = \operatorname{rank}\{\beta_1, \beta_2, \ldots, \beta_n\} \leqslant \operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \leqslant n$，所以 $\operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_n\} = n$，$\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关。由 (1)，$\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基。 □

Theorem 1.13. X 是域 F 上的 n 维线性空间，则 X 中任意一个线性无关的向量组都可以扩充成 X 的一组基。

Proof. 任取 X 中一个线性无关的向量组 $\alpha_1, \alpha_2, \ldots, \alpha_r$，若 $r = n$，由定理 1.12(1) 可知，$\alpha_1, \alpha_2, \ldots, \alpha_r$ 是 X 的一组基；若 $r < n$，则 X 中必定存在一个元素不能由 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 线性表出，将其记为 α_{r+1}，否则的话 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 就是 X 的一组基，进而 X 的维数应是 r，矛盾。不断重复上述过程即可得到一个向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$，这就是 X 的一组基。 □

坐标与坐标变换

Definition 1.15. X 是域 F 上的 n 维线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一个基，由定理 *1.3* 可知 X 中任一向量 α 由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出的方式唯一：

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$$

把系数构成的 n 元有序数组写成列向量的形式，得到 $(a_1, a_2, \ldots, a_n)^T$，该列向量被称为 α 在基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 下的坐标 *(coordinate)*。

接下来我们要讨论的是 X 中某一向量在不同基下的坐标之间有什么关系，首先我们需要定义什么叫不同的基。

Definition 1.16. X 是域 F 上的 n 维线性空间。X 中的两个向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与 $\beta_1, \beta_2, \ldots, \beta_n$ 如果满足 $\alpha_i = \beta_i$，$i = 1, 2, \ldots, n$，那么称这两个向量组**相等**。

Definition 1.17. X 是域 F 上的 n 维线性空间。给定 V 的两个基：

$$\alpha_1, \alpha_2, \ldots, \alpha_n \quad \beta_1, \beta_2, \ldots, \beta_n$$

因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 V 的一个基，所以有：

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n \\ \beta_2 = a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n \\ \cdots\cdots\cdots \\ \beta_n = a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n \end{cases}$$

模仿矩阵乘法的定义将上式写作：

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

将上式右端的矩阵记作 A,称它是基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 到基 $\beta_1, \beta_2, \ldots, \beta_n$ 的过渡矩阵 *(transition matrix)*。于是上式可以写作：

$$(\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

由于这种写法是模仿矩阵乘法的定义，所以矩阵乘法所满足的运算法则对于这种写法也成立。

Theorem 1.14. X 是域 F 上的 n 维线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一个基，且向量组 $\beta_1, \beta_2, \ldots, \beta_n$ 满足：

$$(\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

则 $\beta_1, \beta_2, \ldots, \beta_n$ 是 X 的一个基当且仅当 A 是可逆矩阵。

齐次方程组的解与逆矩阵

Proof. 由可得：

$$
\begin{aligned}
\beta_1,\beta_2,\ldots,\beta_n\text{是 } X \text{ 的一组基} &\Leftrightarrow \text{由}(\beta_1,\beta_2,\ldots,\beta_n)\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=\mathbf{0}\text{可推出}k_1=k_2=\cdots=k_n=0\\
&\Leftrightarrow \text{由}(\alpha_1,\alpha_2,\ldots,\alpha_n)A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=\mathbf{0}\text{可推出}k_1=k_2=\cdots=k_n=0\\
&\Leftrightarrow \text{由}A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=\mathbf{0}\text{可推出}k_1=k_2=\cdots=k_n=0\\
&\Leftrightarrow \text{齐次线性方程组}Ax=\mathbf{0}\text{只有零解}\\
&\Leftrightarrow A\text{可逆}
\end{aligned}
$$

□

Theorem 1.15. X 是域 F 上的 n 维线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_n$ 和 $\beta_1,\beta_2,\ldots,\beta_n$ 是 X 上的两个基，A 是由基 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 到基 $\beta_1,\beta_2,\ldots,\beta_n$ 的过渡矩阵。若 X 中向量 α 在这两个基下的坐标分别为：

$$x=(x_1,x_2,\ldots,x_n)^T,\ y=(y_1,y_2,\ldots,y_n)^T$$

则有：

$$y=A^{-1}x$$

Proof. 因为：

$$
\begin{aligned}
\alpha&=(\alpha_1,\alpha_2,\ldots,\alpha_n)(x_1,x_2,\ldots,x_n)^T\\
&=(\beta_1,\beta_2,\ldots,\beta_n)(y_1,y_2,\ldots,y_n)^T\\
&=(\alpha_1,\alpha_2,\ldots,\alpha_n)A(y_1,y_2,\ldots,y_n)^T
\end{aligned}
$$

所以：

$$(x_1,x_2,\ldots,x_n)^T=A(y_1,y_2,\ldots,y_n)^T$$

即 $x=Ay$。因为 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 和 $\beta_1,\beta_2,\ldots,\beta_n$ 是 X 上的两个基，所以 A 可逆，于是 $y=A^{-1}x$。 □

1.1.3 子空间

子空间的定义及其判别

Definition 1.18. X 是域 F 上的一个线性空间，E 是 X 的一个非空子集。若 E 对于 X 上的加法和纯量乘法也构成域 F 上的一个线性空间，则称 E 是 X 的一个线性子空间 *(linear subspace)*，简称为子空间 *(subspace)*。

Theorem 1.16. X 是域 F 上的一个线性空间，E 是 X 的一个非空子集。E 是 X 的一个子空间的充分必要条件是 E 对 X 中的加法和纯量乘法封闭，即：

$$\alpha,\beta \in E \Rightarrow \alpha+\beta \in E$$
$$k \in F, \alpha \in E \Rightarrow k\alpha \in E$$

Proof. **(1) 必要性：**因为 E 是 X 的子空间，由子空间的定义，E 中的加法和纯量乘法就是 X 中的加法和纯量乘法，由加法和纯量乘法的定义，E 对 X 中的加法和纯量乘法封闭。

(2) 充分性：由条件可知此时已经对 E 定义了加法和纯量乘法，且就是 X 中的加法和纯量乘法，还需要证明的是 E 对于它自身的加法和纯量乘法满足线性空间的八条运算法则。对 $\forall\, \alpha,\beta,\gamma \in E$, $\forall\, k,l \in F$：

1. 因为 $\alpha,\beta,\gamma \in E$，所以 $\alpha,\beta,\gamma \in X$，对于 X 上的加法，有 $\alpha+\beta=\beta+\alpha$, $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$，而 X 上的加法就是 E 上的加法，所以 E 上的加法满足线性空间运算法则 (1)(2)；同理，E 上的纯量乘法满足线性空间运算法则 (5)(6)(7)(8)；

2. 因为 E 不是空集，所以存在 $\delta \in E$。因为 $\delta \in X$，所以由 X 中的纯量乘法可得 $0\delta = \mathbf{0}_x \in E$。对任意的 $\alpha \in E$，有 $\alpha \in X$，根据 X 中的加法有 $\alpha + \mathbf{0}_x = \alpha$，于是 $\mathbf{0}_x$ 是 E 中的零元，即 E 满足线性空间运算法则 (3)；

3. 因为 E 对纯量乘法封闭，所以 $(-1)\alpha \in E$。因为 $\alpha+(-1)\alpha = [1+(-1)]\alpha = 0\alpha = \mathbf{0}$（第一步到第二步由 1，第三步到第四步因为 $\alpha \in X$），所以 $(-1)\alpha$ 是 α 的负元，即 E 满足线性空间运算法则 (4)。

□

子空间的性质

Theorem 1.17. 设 X 是域 F 上的一个线性空间，E 是 X 的任意一个子空间，则有 $\dim E \leqslant \dim X$。X 是有限维时等号成立当且仅当 $E = X$。

Proof. 由定理 1.13可得，E 的一组基可以扩充成 X 的一组基，所以 $\dim E \leqslant \dim X$。当 $E = X$ 时，显然 $\dim E = \dim X$。当 $\dim E = \dim X$ 时，由定理 1.12可知 E 的一组基就是 X 的一组基，所以 X 中的任一向量可以由 E 的基线性表出，于是 $X \subset E$，从而 $E = X$。 □

张成的子空间

Definition 1.19. 设 X 是域 F 上的一个线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_n \in X$，称：

$$W = \left\{ \sum_{i=1}^{n} k_i \alpha_i : k_i \in F \right\}$$

为向量组 $\alpha_1, \alpha_2, \ldots, \alpha_n$ **张成的线性子空间**，记作 $< \alpha_1, \alpha_2, \ldots, \alpha_n >$。

Property 1.1.2. 设 X 是域 F 上的一个线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_n \in X$。$< \alpha_1, \alpha_2, \ldots, \alpha_n >$ 具有如下性质：

1. $< \alpha_1, \alpha_2, \ldots, \alpha_n >$ 是 X 的一个子空间；
2. $< \alpha_1, \alpha_2, \ldots, \alpha_n >$ 是 X 中包含 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的最小的子空间；
3. $< \alpha_1, \alpha_2, \ldots, \alpha_n >$ 的极大线性无关组是 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的基；
4. $\dim < \alpha_1, \alpha_2, \ldots, \alpha_n > = \operatorname{rank}\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$；
5. 若 $\beta_1, \beta_2, \ldots, \beta_m \in X$，则有：
$$< \alpha_1, \alpha_2, \ldots, \alpha_n > = < \beta_1, \beta_2, \ldots, \beta_m > \Leftrightarrow \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \cong \{\beta_1, \beta_2, \ldots, \beta_m\}$$
6. 若 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基，则 $X = < \alpha_1, \alpha_2, \ldots, \alpha_n >$。

Proof. (1) 显然 $< \alpha_1, \alpha_2, \ldots, \alpha_n >$ 中的元素对 X 中的加法与纯量乘法封闭。

(2) 由子空间对加法与纯量乘法的封闭性，包含 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的子空间必然包含 $< \alpha_1, \alpha_2, \ldots, \alpha_n >$。

(3) 显然。

(4) 由 (3) 直接得到。

(5) 充分性显然，必要性由反证法可得。

(6) 显然。 □

子空间的交与和

Theorem 1.18. 设 X 是域 F 上的一个线性空间，I 是一个指标集，对任意的 $i \in I$ 有 X_i 是 X 的子空间，则

$$\bigcap_{i \in I} X_i = \{\alpha : \alpha \in X_i, \ \forall\, i \in I\}$$

也是 X 的子空间。

Proof. 任取 $\alpha, \beta \in \bigcap_{i \in I} X_i$ 和 $k_1, k_2 \in F$。对任意的 $i \in I$，因为 X_i 是 X 的子空间，$\alpha, \beta \in X_i$，所以 $k_1\alpha + k_2\beta \in X_i$，于是 $k_1\alpha + k_2\beta \in \bigcap_{i \in I} X_i$。由定理 1.16可知 $\bigcap_{i \in I} X_i$ 是 X 的子空间。 □

Definition 1.20. 设 X 是域 F 上的一个线性空间，$X_1, X_2, \ldots, X_n$ 是 X 的子空间。定义：

$$\sum_{i=1}^{n} X_i = \left\{ \sum_{i=1}^{n} \alpha_i : \alpha_i \in X_i \right\}$$

Theorem 1.19. 设 X 是域 F 上的一个线性空间，$X_1, X_2, \ldots, X_n$ 是 X 的子空间，$n \in \mathbb{N}^+$，则 $X_1 + X_2 + \cdots + X_n$ 也是 X 的子空间。

Proof. 任取 $\alpha, \beta \in X_1 + X_2 + \cdots + X_n$ 和 $k_1, k_2 \in F$，则：

$$\alpha = \gamma_1 + \gamma_2 + \cdots + \gamma_n,\ \beta = \delta_1 + \delta_2 + \cdots + \delta_n$$

其中 $\gamma_i, \delta_i \in X_i$，$i = 1, 2, \ldots, n$，于是：

$$k_1\alpha + k_2\beta = \sum_{i=1}^{n} (k_i\gamma_i + k_2\delta_i)$$

因为 X_i 是 X 的子空间，所以 $k_i\gamma_i + k_2\delta_i \in X_i$，$i = 1, 2, \ldots, n$，于是 $k_1\alpha + k_2\beta \in X_1 + X_2 + \cdots + X_n$。由定理 1.16可得 $X_1 + X_2 + \cdots + X_n$ 是 X 的子空间。 □

Lemma 1.1. 设 X 是域 F 上的一个线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n \in X$，则：

$$< \alpha_1, \alpha_2, \ldots, \alpha_m > + < \beta_1, \beta_2, \ldots, \beta_n > = < \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n >$$

Proof. 由子空间与子空间和的定义可直接得到。 □

Theorem 1.20. 设 X 是域 F 上的一个线性空间，X_1, X_2 为 X 的有限维子空间，则 $X_1 \cap X_2$，$X_1 + X_2$ 也是有限维的，并且有：

$$\dim X_1 + \dim X_2 = \dim(X_1 + X_2) + \dim(X_1 \cap X_2)$$

Proof. 显然 $X_1 \cap X_2$ 是有限维的，设 $X_1, X_2, X_1 \cap X_2$ 的维数分别为 n_1, n_2, m。取 $X_1 \cap X_2$ 的一个基 $\alpha_1, \alpha_2, \ldots, \alpha_m$，把它分别扩充为 X_1 和 X_2 的一个基 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n_1-m}$ 与 $\alpha_1, \alpha_2, \ldots, \alpha_m, \gamma_1, \gamma_2, \ldots, \gamma_{n_2-m}$。由引理 1.1和性质 1.1.2(6) 可得：

$$\begin{aligned} X_1 + X_2 &= < \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n_1-m} > + < \alpha_1, \alpha_2, \ldots, \alpha_m, \gamma_1, \gamma_2, \ldots, \gamma_{n_2-m} > \\ &= < \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n_1-m}, \gamma_1, \gamma_2, \ldots, \gamma_{n_2-m} > \end{aligned}$$

所以 $\dim(X_1 + X_2) < n_1 + n_2 - m$，即 $X_1 + X_2$ 是有限维的。

设：

$$k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_{n_1-m}\beta_{n_1-m} + q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m} = \mathbf{0}$$

于是：

$$k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_{n_1-m}\beta_{n_1-m} = -(q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m})$$

左边属于 X_1,右边属于 X_2,所以它们属于 $X_1 \cap X_2$。于是右边的向量也可以由 $\alpha_1, \alpha_2, \ldots, \alpha_m$ 线性表出，即：

$$(q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m}) = p_1\alpha_1 + \cdots + p_m\alpha_m$$

因为 $\alpha_1, \alpha_2, \ldots, \alpha_m, \gamma_1, \gamma_2, \ldots, \gamma_{n_2-m}$ 是 X_2 的一个基，所以：

$$q_1 = q_2 = \cdots = q_{n_2-m} = p_1 = p_2 = \cdots = p_m = 0$$

又因为 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n_1-m}$ 是 X_1 的一个基，所以：

$$k_1 = k_2 = \cdots = k_m = l_1 = l_2 = \cdots = l_{n_1-m} = 0$$

于是 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n_1-m}, \gamma_1, \gamma_2, \ldots, \gamma_{n_2-m}$ 线性无关。由基的定义，它们是 X 的一组基，于是有：

$$\begin{aligned}
\dim(X_1 + X_2) &= m + n_1 - m + n_2 - m \\
&= \dim(X_1 \cap X_2) + \dim(X_1) - \dim(X_1 \cap X_2) + \dim(X_2) - \dim(X_1 \cap X_2) \\
&= \dim(X_1) + \dim(X_2) - \dim(X_1 \cap X_2)
\end{aligned}$$

即：

$$\dim X_1 + \dim X_2 = \dim(X_1 + X_2) + \dim(X_1 \cap X_2)$$

□

子空间的直和

Definition 1.21. 设 X 是域 F 上的线性空间，$X_1, X_2, \ldots, X_n$ 是 X 的子空间。如果 $X_1 + X_2 + \cdots + X_n$ 中的每个向量 α 都能唯一地表示为：

$$\alpha = \sum_{i=1}^{n} \alpha_i,\ \alpha_i \in X_i$$

则称和 $X_1 + X_2 + \cdots + X_n$ 为直和 *(direct sum)*，记为 $X_1 \oplus X_2 \oplus \cdots \oplus X_n$，也可以写作 $\oplus_{i=1}^{n} X_i$。若 $X = X_1 \oplus X_2$，则称 X_2 为 X_1 的补空间 *(complement)*。

Theorem 1.21. 设 X 是域 F 上的线性空间，$X_1, X_2, \ldots, X_n$ 是 X 的有限维子空间，则下列命题等价：

1. 和 $X_1 + X_2 + \cdots + X_n$ 是直和；
2. 和 $X_1 + X_2 + \cdots + X_n$ 中零向量的表示方法唯一；
3. $X_i \cap \left(\sum_{j \neq i} X_j \right) = \mathbf{0},\ i, j = 1, 2, \ldots, n$；
4. $\dim(X_1 + X_2 + \cdots + X_n) = \dim X_1 + \dim X_2 + \cdots + \dim X_n$；
5. $X_i,\ i = 1, 2, \ldots, n$ 的基合起来是和 $X_1 + X_2 + \cdots + X_n$ 的一个基。

前三条在 $X_1, X_2, \ldots, X_n$ 是无限维子空间时也成立。

Proof. $1 \Rightarrow 2$：由直和的定义是显然的。

$2 \Rightarrow 3$：任取 $\alpha \in X_i \cap \left(\sum\limits_{j\neq i} X_j\right)$。因为 $\alpha \in X_i$，X_i 是一个子空间，所以 $-\alpha \in X_i$。因为 $\alpha \in \left(\sum\limits_{j\neq i} X_j\right)$，所以 α 可表示为：

$$\alpha = \sum_{j\neq i} \alpha_j,\ j = 1, 2, \ldots, i-1, i+1, \ldots n$$

于是：

$$\mathbf{0} = \alpha + (-\alpha) = -\alpha + \sum_{j\neq i} \alpha_j$$

因为 $\mathbf{0} = \mathbf{0} + \mathbf{0} + \cdots + \mathbf{0}$，所以 $-\alpha = \mathbf{0}$，$\alpha = \mathbf{0}$。由 α 的任意性，$X_i \cap \left(\sum\limits_{j\neq i} X_j\right) = \mathbf{0}$，$i, j = 1, 2, \ldots, n$。

$3 \Rightarrow 1$：假设和 $X_1 + X_2 + \cdots + X_n$ 不是直和，则存在 $\alpha \in X_1 + X_2 + \cdots + X_n$ 有两种表示方式，设：

$$\alpha = \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i,\ \alpha_i, \beta_i \in X_i,\ \alpha_i \neq \beta_i$$

则：

$$\alpha_i - \beta_i = \sum_{j\neq i} (\beta_j - \alpha_j),\ i, j = 1, 2, \ldots, n$$

注意到左边属于 X_i，右边属于 $\sum\limits_{j\neq i} X_j$，所以：

$$\alpha_i - \beta_i \in X_i \cap \left(\sum_{j\neq i} X_j\right) = \mathbf{0}$$

于是 $\alpha_i = \beta_i$，矛盾，因此和 $X_1 + X_2 + \cdots + X_n$ 是直和。

$1 \Rightarrow 4$：因为和 $X_1 + X_2 + \cdots + X_n$ 是直和，所以 $X_i \cap \left(\sum\limits_{j\neq i} X_j\right) = \mathbf{0}$，$i, j = 1, 2, \ldots, n$。由定理 1.20可得：

$$\begin{aligned}\dim\left(\sum_{i=1}^{n} X_i\right) &= \dim\left(X_1 + \sum_{i=2}^{n} X_i\right) = \dim X_1 + \dim\left(\sum_{i=2}^{n} X_i\right) - \dim\left(X_1 \cap \sum_{i=2}^{n} X_i\right) \\ &= \dim X_1 + \dim\left(\sum_{i=2}^{n} X_i\right)\end{aligned}$$

注意到：

$$X_2 \cap \sum_{i=3}^{n} X_i \subset X_2 \cap \left(X_1 + \sum_{i=3}^{n} X_i\right) = \mathbf{0}$$

所以：

$$\dim\left(\sum_{i=2}^{n} X_i\right) = \dim X_2 + \dim\left(\sum_{i=3}^{n} X_i\right)$$

由数学归纳法可得：

$$\dim(X_1 + X_2 + \cdots + X_n) = \dim X_1 + \dim X_2 + \cdots + \dim X_n$$

$4 \Rightarrow 5$：在 X_i 中取一个基 $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{ir_i}$，$i = 1, 2, \ldots, n$，由引理 1.1和性质 1.1.2(6) 可得：

$$\begin{aligned}\sum_{i=1}^{n} X_i &=< \alpha_{11}, \alpha_{12}, \ldots, \alpha_{1r_1} > + < \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2r_2} > + \cdots + < \alpha_{n1}, \alpha_{n2}, \ldots, \alpha_{nr_n} > \\ &=< \alpha_{11}, \alpha_{12}, \ldots, \alpha_{1r_1}, \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2r_2}, \ldots, \alpha_{n1}, \alpha_{n2}, \ldots, \alpha_{nr_n} >\end{aligned}$$

因为：

$$\dim(X_1 + X_2 + \cdots + X_n) = \dim X_1 + \dim X_2 + \cdots + \dim X_n = \sum_{i=1}^{n} r_i$$

所以 α_{ij}，$i = 1, 2, \ldots, n$，$j = 1, 2, \ldots, r_i$ 线性无关，否则的话，$X_1 + X_2 + \cdots + X_n$ 中线性无关的向量数目就会小于 $\sum_{i=1}^{n} r_i$，这与它的维数产生了矛盾。因此 α_{ij}，$i = 1, 2, \ldots, n$，$j = 1, 2, \ldots, r_i$ 是 $X_1 + X_2 + \cdots + X_n$ 的一个基，即 X_i，$i = 1, 2, \ldots, n$ 的基合起来是和 $X_1 + X_2 + \cdots + X_n$ 的一个基。

$5 \Rightarrow 1$：在 X_i 中取一个基 $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{ir_i}$，$i = 1, 2, \ldots, n$，则 α_{ij}，$i = 1, 2, \ldots, n$，$j = 1, 2, \ldots, r_i$ 是 $X_1 + X_2 + \cdots + X_n$ 的一个基。设:

$$\mathbf{0} = \alpha_1 + \alpha_2 + \cdots + \alpha_n,\ \alpha_i \in X_i$$

则有：

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = \sum_{i=1}^{n}\sum_{j=1}^{r_i} k_{ij}\alpha_{ij} = \mathbf{0}$$

$k_{ij} \in F$。因为 α_{ij}，$i = 1, 2, \ldots, n$，$j = 1, 2, \ldots, r_i$ 是 $X_1 + X_2 + \cdots + X_n$ 的一个基，所以 $k_{ij} = 0$，$i = 1, 2, \ldots, n$，$j = 1, 2, \ldots, r_i$，于是 $\alpha_i = \mathbf{0}$，即和 $X_1 + X_2 + \cdots + X_n$ 中零向量的表示方法唯一，所以和 $X_1 + X_2 + \cdots + X_n$ 是直和。 □

Theorem 1.22. 设 X 是域 F 上的线性空间，则 X 的任一子空间 E 都有补空间。

Proof. **(1)** $\dim X = n < +\infty$：取 E 的一个基 $\alpha_1, \alpha_2, \ldots, \alpha_m$，把它扩充为 X 的一个基 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n-m}$，由引理 1.1可知：

$$\begin{aligned}X &=< \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n-m} > \\ &=< \alpha_1, \alpha_2, \ldots, \alpha_m > + < \beta_1, \beta_2, \ldots, \beta_{n-m} >= E + W\end{aligned}$$

其中 $W =< \beta_1, \beta_2, \ldots, \beta_{n-m} >$。因为 $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{n-m}$ 线性无关，由定理 1.1可知 $\beta_1, \beta_2, \ldots, \beta_{n-m}$ 线性无关，所以 $\beta_1, \beta_2, \ldots, \beta_{n-m}$ 是 W 的一组基。于是 E 的一

组基和 W 的一组基合起来就是 X 的一组基。由定理 1.21可知 $X = E \oplus W$，于是 W 是 E 的补空间。

(2) $\dim X = +\infty$：考虑商空间 X/E，设其一个基为 $\alpha_i + E,\ i \in I$，其中 I 是一个指标集。由推论 1.5可知 $\alpha_i,\ i \in I$ 线性无关。令：

$$W = \left\{ \sum_{i=1}^{n} k_i \alpha_{r_i} : k_i \in F,\ n \in \mathbb{N}^+ \right\}$$

显然 $\alpha_i,\ i \in I$ 是 W 的一个基。下面证明 W 是 E 的一个补空间。

任取 $\alpha \in X$，因为 $\alpha_i + E,\ i \in I$ 是 X/E 的一个基，所以：

$$\alpha + E = \sum_{i=1}^{m} l_i \alpha_{r_i} + E,\ l_j \in F$$

即：

$$\alpha - \sum_{i=1}^{m} l_i \alpha_{r_i} \in E$$

于是 α 可以表示为 E 和 W 中两个元素的和。由 α 的任意性，$X = E + W$。

任取 $\beta \in W \cap E$，因为 $\beta \in W$，所以：

$$\beta = \sum_{i=1}^{t} p_i a_{r_i},\ p_i \in F$$

又因为 $\beta \in E$，所以：

$$E = \beta + E = \sum_{i=1}^{t} p_i a_{r_i} + E = \sum_{i=1}^{t} p_i (a_{r_i} + E)$$

因为 $\alpha_i + E,\ i \in I$ 线性无关，所以 $p_i = 0,\ i = 1, 2, \ldots, t$（$E$ 是 X/E 的零元），于是 $\beta = \mathbf{0}$。由定理 1.21(3) 可知 $X = E \oplus W$，W 是 E 的补空间。

综上，X 的任一子空间 E 都有补空间。 □

1.1.4 线性空间的同构

Definition 1.22. 设 X, Y 为域 F 上的线性空间。如果存在 X 到 Y 的一个双射 σ，使得对于任意的 $\alpha, \beta \in X$，$k_1, k_2 \in F$，有：

$$\sigma(k_1\alpha + k_2\beta) = k_1\sigma(\alpha) + k_2\sigma(\beta)$$

则称 σ 是 X 到 Y 的一个同构映射 *(isomorphism)*，此时称 X 与 Y 同构 *(isomorphic)*，记作 $X \cong Y$。

Property 1.1.3. 设 X, Y 为域 F 上的线性空间，且 X, Y 同构，σ 是 X 到 Y 的同构映射，则：

1. $\sigma(\mathbf{0}_X) = \mathbf{0}_Y$；
2. 对于任意的 $\alpha \in X$，有 $\sigma(-\alpha) = -\sigma(\alpha)$；

3. 对于任意的 $\alpha_1, \alpha_2, \ldots, \alpha_n \in X,\ k_1, k_2, \ldots, k_n \in F$，有：

$$\sigma\left(\sum_{i=1}^{n} k_i\alpha_i\right) = \sum_{i=1}^{n} k_i\sigma(\alpha_i)$$

4. $\alpha_1, \alpha_2, \ldots, \alpha_n \in X$ 线性相关当且仅当 $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$ 线性相关；
5. 如果 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基，则 $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$ 是 Y 的一组基；
6. 若 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基，则 $\alpha \in X$ 在基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 下的坐标和 $\sigma(\alpha) \in Y$ 在基 $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$ 下的坐标相同。
7. 若 E 是 X 的一个子空间，则 $\sigma(E)$ 是 Y 的一个子空间。若 $\dim E = n < +\infty$，则 $\dim \sigma(E) = n$；
8. 线性空间的同构是一个等价关系，其等价类被称为同构类。

Proof. (1) 任取 $\alpha \in X$，由线性空间的运算法则和同构映射的定义可得 $\sigma(\mathbf{0}_X) = \sigma(0\alpha) = 0\sigma(\alpha) = \mathbf{0}_Y$。

(2) 显然 $\sigma(-\alpha) = \sigma[(-1)\alpha] = (-1)\sigma(\alpha) = -\sigma(\alpha)$。

(3) 由同构映射的定义直接可得。

(4) 设：

$$\sum_{i=1}^{n} k_i\alpha_i = \mathbf{0}_X,\ k_i \in F$$

由 (3) 和 (1) 可得：

$$\sum_{i=1}^{n} k_i\sigma(\alpha_i) = \sigma(\mathbf{0}_X) = \mathbf{0}_Y$$

若 k_i 不全为 0，即 $\alpha_1, \alpha_2, \ldots, \alpha_n \in X$ 或 $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$ 线性相关，显然此时另一个也线性相关。

(5) 由 (4) 可得 $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n)$ 线性无关。任取 $\beta \in Y$，因为 σ 是一个满射，则存在 $\alpha \in X$ 使得 $\sigma(\alpha) = \beta$。因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是 X 的一组基，所以存在 $k_1, k_2, \ldots, k_n \in F$ 使得：

$$\alpha = \sum_{i=1}^{n} k_i\alpha_i$$

由 (3) 可得：

$$\beta = \sigma(\alpha) = \sigma\left(\sum_{i=1}^{n} k_i\alpha_i\right) = \sum_{i=1}^{n} k_i\sigma(\alpha_i)$$

于是 β 可由 $\sigma(\alpha_i),\ i = 1, 2, \ldots, n$ 线性表出。由 β 的任意性，$\sigma(\alpha_i),\ i = 1, 2, \ldots, n$ 是 Y 的一组基。

(6) 由 (3)(5) 直接得到。

(7) 任取 $\alpha,\beta\in\sigma(E)$, $k_1,k_2\in F$，考虑 $k_1\alpha+k_2\beta$。因为 σ 是 X 到 Y 的一个双射，所以对于 α,β，存在 $a,b\in X$ 满足 $\sigma(a)=\alpha$, $\sigma(b)=\beta$。因为 E 是 X 的一个子空间，所以 $k_1a+k_2b\in E$，于是 $\sigma(k_1a+k_2b)=k_1\alpha+k_2\beta\in\sigma(E)$，所以 $\sigma(E)$ 是 Y 的一个子空间。由 (5) 可直接得到有限维情况下 E 与 $\sigma(E)$ 之间的维数关系。

(8) 反身性由恒等映射保证，对称性由双射保证，传递性由复合映射可直接得到。 □

Theorem 1.23. 设 X,Y 为域 F 上的有限维线性空间，则 X 与 Y 同构的充分必要条件为它们的维数相同，于是维数是有限维线性空间同构类的完全不变量。

Proof. **(1) 必要性：** 由性质 1.1.3(5) 直接得到。

(2) 充分性： 设二者维数都是 n,取 X 的一组基 $\alpha_1,\alpha_2,\dots,\alpha_n$ 和 Y 的一组基 $\beta_1,\beta_2,\dots,\beta_n$。令：

$$\sigma:\sum_{i=1}^n k_i\alpha_i\longrightarrow\sum_{i=1}^n k_i\beta_i$$

显然它是一个线性映射并且是一个双射，于是 X 与 Y 同构。 □

1.1.5 商空间

商空间的定义

Theorem 1.24. 设 X 是域 F 上的一个线性空间，L 是 E 的子空间。对于 $\forall\, x,y\in X$，若 $x-y\in L$，则称 x,y 是等价的，记为 $x\sim y$。该关系是一个等价关系。对于该关系的任意等价类 α，任取 $x\in\alpha$，α 可由 $\{x+y:y\in L\}$ 来表示，简记为 $x+L$。

Proof. 任取 $x,y,z\in X$。

(1) 因为 $x-x=\mathbf{0}\in L$，所以该关系满足自反性。

(2) 若 $x\sim y$，即 $x-y\in L$，因为 L 是一个线性空间，所以 $y-x=(-1)(x-y)\in L$，于是 $y\sim x$，该关系满足对称性。

(3) 若 $x\sim y$, $y\sim z$，则有 $x-y\in L$, $y-z\in L$，因为 L 是一个线性空间，所以 $x-z=(x-y)+(y-z)\in L$，即 $x\sim z$。该关系满足传递性。

综上，该关系是一个等价关系。

下证明表示方法的正确性。

对任意的 $a\in\alpha$，因为 $x\in\alpha$，所以 $x-a\in L$，于是存在 $z\in L$ 使得 $x-a=-z$，即 $a=x+z$。 □

Theorem 1.25. 设 X 是域 F 上的一个线性空间，L 是 X 的子空间，$\hat{X}$ 为 X 中所有等价类构成的集合，确定等价类的关系为定理 *1.24*中的关系。对任意的 $\alpha,\beta\in\hat{X}$, $k\in F$，在 $\hat{X}$ 中定义线性运算如下：

$$\alpha+\beta=x+y+L,\ x\in\alpha,\ y\in\beta$$

$$k\alpha=kx+L,\ x\in\alpha$$

则 $\hat{X}$ 成为一个线性空间，称其为 X 关于 L 的商空间 *(quotient space)*，记作 X/L。

Proof. 先证明上述线性运算与 x, y 的选择无关。

(1) 任取 $x' \in \alpha,\ y' \in \beta$，满足 $x \neq x',\ y \neq y'$。于是有：

$$x' + y' + L = x + y + (x' - x) + (y' - y) + L$$

因为 $x, x' \in \alpha,\ y, y' \in \beta$，所以 $x' - x, y' - y \in L$，于是 $x' + y' + L = x + y + L$。

(2) 任取 $x' \in \alpha$，满足 $x \neq x'$。于是有：

$$kx' + L = kx + k(x' - x) + L$$

因为 $x, x' \in \alpha$,所以 $x' - x \in L$。因为 L 是线性空间,所以 $k(x' - x) \in L$,于是 $kx' + L = kx + L$。

下证明 $\hat{X}$ 是一个线性空间。

1. 任取 $\hat{X}$ 中的两个元素 $\alpha = x + L,\ \beta = y + L$。因为 $x, y \in X$，X 是一个线性空间，所以 $x + y = y + x$，于是 $\alpha + \beta = x + y + L = y + x + L = \beta + \alpha$。由 α, β 的任意性，$\hat{X}$ 上的加法满足线性空间运算法则 (1)；

2. 任取 $\hat{X}$ 中的三个元素 $\alpha = x + L,\ \beta = y + L,\ \gamma = z + L$。因为 $x, y, z \in X$，X 是一个线性空间，所以 $(x + y) + z = x + (y + z)$，于是 $(\alpha + \beta) + \gamma = (x + y) + z + L = x + (y + z) + L = \alpha + (\beta + \gamma)$。由 α, β, γ 的任意性，$\hat{X}$ 上的加法满足线性空间运算法则 (2)；

3. 任取 $\alpha = x + L \in \hat{X}$，则 $\alpha + \mathbf{0} + L = x + \mathbf{0} + L = x + L = \alpha$，于是 $\mathbf{0} + L$ 是 $\hat{X}$ 中的零元。$\hat{X}$ 上的加法满足线性空间运算法则 (3)；

4. 任取 $\alpha = x + L \in \hat{X}$，则 $\alpha + -x + L = x + (-x) + L = \mathbf{0} + L$，于是 $-x + L$ 是 α 的负元。由 α 的任意性，$\hat{X}$ 上的加法满足线性空间运算法则 (4)；

5. 任取 $\alpha = x + L \in \hat{X}$。因为 $x \in X$，X 是一个线性空间，所以 $1x = x$，于是 $1\alpha = 1x + L = x + L = \alpha$。由 α 的任意性，$\hat{X}$ 上的纯量乘法满足线性空间运算法则 (5)；

6. 任取 $\alpha = x + L \in \hat{X},\ k, l \in F$。因为 $x \in X$，X 是一个线性空间，所以 $(kl)x = k(lx)$，于是 $(kl)\alpha = (kl)x + L = k(lx) + L = k(l\alpha)$。由 α, k, l 的任意性，$\hat{X}$ 上的纯量乘法满足线性空间运算法则 (6)；

7. 任取 $\alpha = x + L \in \hat{X},\ k, l \in F$。因为 $x \in X$,X 是一个线性空间,所以 $(k + l)x = kx + lx$，于是 $(k + l)\alpha = (k + l)x + L = kx + lx + L = k\alpha + l\alpha$。由 α, k, l 的任意性，$\hat{X}$ 上的纯量乘法满足线性空间运算法则 (7)；

8. 任取 $\hat{X}$ 中的两个元素 $\alpha = x + L,\ \beta = y + L$。因为 $x, y \in X$，X 是一个线性空间，所以 $k(x + y) = ky + kx$，于是 $k(\alpha + \beta) = k(x + y) + L = kx + ky + L = k\beta + k\alpha$。由 α, β, k 的任意性，$\hat{X}$ 上的加法满足线性空间运算法则 (8)。

综上，$\hat{X}$ 是一个线性空间。 □

商空间的性质

Theorem 1.26. 设 X 是域 F 上的一个有限维线性空间，E 是 X 的一个子空间，则：

$$\dim(X/E) = \dim(X) - \dim(E)$$

Proof. 设 $\dim X = n$, $\dim E = m$，取 E 的一组基 $\alpha_1, \alpha_2, \ldots, \alpha_m$，把它扩充为 X 的一组基 $\alpha_1, \alpha_2, \ldots, \alpha_n$。任取 $\beta + W \in X/E$，因为 $\beta \in X$，于是：

$$\beta + E = \sum_{i=1}^{n} k_i\alpha_i + E = \sum_{i=1}^{m} k_i\alpha_i + \sum_{i=m+1}^{n} k_i\alpha_i + E = \sum_{i=m+1}^{n} k_i\alpha_i + E$$

其中 $k_i \in F$, $i = 1, 2, \ldots, n$。上式表明，对任意的 $\beta + W \in X/E$，都可以用 $\alpha_i + W$, $i = m+1, m+2, \ldots, n$ 的线性组合表示。下证明它们线性无关。

设：

$$\sum_{i=m+1}^{n} k_i\alpha_i + E = E$$

则 $\sum\limits_{i=m+1}^{n} k_i\alpha_i \in E$，于是存在 $l_j \in F$, $j = 1, 2, \ldots, m$ 使得：

$$\sum_{j=1}^{m} l_j\alpha_j = \sum_{i=m+1}^{n} k_i\alpha_i$$

因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关，所以 $k_i = l_j = 0$, $i = m+1, m+2, \ldots, n$, $j = 1, 2, \ldots, m$，因此 $\alpha_i + W$, $i = m+1, m+2, \ldots, n$ 线性无关，即它们是 X/E 的一组基，$\dim(X/E) = n - m = \dim X - \dim E$。 □

Definition 1.23. 设 X 是域 F 上的一个线性空间，E 是 X 的一个子空间。若 X/E 是有限维的，则称 $\dim(X/E)$ 是 E 在 X 中的余维数 *(codimension)*，记作 $\operatorname{codim} E$。

标准映射

Definition 1.24. 设 X 是域 F 上的一个线性空间，E 是 X 的一个子空间，定义映射：

$$\pi : \alpha \longrightarrow \alpha + E, \ \forall\, \alpha \in X$$

称之为标准映射 *(canonical mapping)*。

Property 1.1.4. 设 X 是域 F 上的一个线性空间，E 是 X 的一个子空间，π 是 X 到 X/E 的标准映射，则 π 具有如下性质：

1. X/E 中的一个元素 $\alpha + E$ 在 π 下的原像是：

$$\{\alpha + \beta : \beta \in E\}$$

2. π 是一个满射。当且仅当 E 是零空间时，π 是一个双射；

3. π 是一个线性映射；

4. 若 $\alpha_1, \alpha_2, \ldots, \alpha_n \in X$ 线性相关，则 $\pi(\alpha_1), \pi(\alpha_2), \ldots, \pi(\alpha_n)$ 线性相关。

Proof. (1) 显然：

$$\begin{aligned}\gamma \in \pi^{-1}(\alpha + E) \Leftrightarrow \gamma + E = \alpha + E \Leftrightarrow \gamma - \alpha \in E \\ \Leftrightarrow \exists\, \beta \in E,\ \gamma = \alpha + \beta \Leftrightarrow \gamma \in \{\alpha + \beta : \beta \in E\}\end{aligned}$$

(2) 满射是显然的结论。当 E 是零空间时，π 是 X 上的恒等变换，恒等变换显然是双射。当 π 是双射时，π 是一个单射，即对任何的 $\alpha + E \in X/E$，有且仅有 $\alpha \in X$ 使得 $\pi(\alpha) = \alpha + E$，也即不存在 $\beta \in X,\ \beta \neq \alpha$，使得 $\beta - \alpha \in E$，此时 E 只能为零空间。

(3) $\forall\, \alpha, \beta \in X,\ k_1, k_2 \in F$，有：

$$\pi(k_1\alpha + k_2\beta) = k_1\alpha + k_2\beta + E = k_1\alpha + E + k_2\alpha + E = k_1\pi(\alpha) + k_2\pi(\beta)$$

于是 π 是一个线性映射。

(4) 因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性相关，所以存在不全为零的 $k_1, k_2, \ldots, k_n \in F$ 使得：

$$\sum_{i=1}^{n} k_i\alpha_i = \mathbf{0}$$

由 (3) 可得：

$$\pi\left(\sum_{i=1}^{n} k_i\alpha_i\right) = \sum_{i=1}^{n} k_i\pi(\alpha_i) = \mathbf{0}$$

于是 $\pi(\alpha_i),\ i = 1, 2, \ldots, n$ 线性相关。 □

Corollary 1.5. 设 X 是域 F 上的一个线性空间，E 是 X 的一个子空间。若 $\{\alpha_i + W : i \in I\}$ 是 X/E 的一组基，I 是一个指标集，则 $\alpha_i,\ i \in I$ 线性无关。

Proof. 就 X/E 是无限维的给出证明，有限维情况类似。

若此时 $\alpha_i,\ i \in I$ 线性相关，则存在 $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n}$ 线性相关，其中 $n \in \mathbb{N}^+$。由性质 1.1.4(4) 可得 $\alpha_{i_j} + W,\ j = 1, 2, \ldots, n$ 线性相关，于是 $\{\alpha_i + W : i \in I\}$ 线性相关，矛盾。 □

1.2 线性变换

1.2.1 线性变换的定义与基本性质

Definition 1.25. 设 X, Y 是域 F 上的线性空间，X 到 Y 上的一个映射 $\mathcal{T}$ 如果对任意的 $\alpha, \beta \in X$ 和任意的 $k_1, k_2 \in F$，有：

$$\mathcal{T}(k_1\alpha + k_2\beta) = k_1\mathcal{T}\alpha + k_2\mathcal{T}\beta$$

则称 $\mathcal{T}$ 是 X 到 Y 的一个线性映射 *(linear mapping)*。若 $Y = X$，则称 $\mathcal{T}$ 为 X 上的线性变换 *(linear transformation)*；若 $Y = F$，则称 $\mathcal{T}$ 为 X 上的线性函数 *(linear function)*。

线性映射空间与线性映射的运算

Definition 1.26. 设 X, Y 是域 F 上的线性空间，将 X 到 Y 的所有线性映射组成的集合记为 $\mathrm{Hom}(X, Y)$。设 $\mathcal{T}_1, \mathcal{T}_2 \in \mathrm{Hom}(X, Y)$，$k \in F$，定义线性映射的加法与纯量乘法如下：

$$(\mathcal{T}_1 + \mathcal{T}_2)\alpha = \mathcal{T}_1\alpha + \mathcal{T}_2\alpha,\ \forall\, \alpha \in X$$
$$(k\mathcal{T}_1) = k\mathcal{T}_1\alpha,\ \forall\, \alpha \in X$$

容易验证 $\mathrm{Hom}(X, Y)$ 成为域 F 上的一个线性空间。$X = Y$ 时将 $\mathrm{Hom}(X, Y)$ 简记为 $\mathrm{Hom}(X)$。

Definition 1.27. 设 X, Y 是域 F 上的线性空间，$\mathcal{T}_1, \mathcal{T}_2 \in \mathrm{Hom}(X, Y)$，定义线性映射的减法如下：

$$\mathcal{T}_1 - \mathcal{T}_2 = \mathcal{T}_1 + (-\mathcal{T}_2)$$

Definition 1.28. 设 X, Y, Z 是域 F 上的线性空间，$\mathcal{T}_1 \in \mathrm{Hom}(X, Y), \mathcal{T}_2 \in \mathrm{Hom}(Z, Y)$，定义线性映射乘法如下：

$$\forall\, \alpha \in X,\ (\mathcal{T}_2\mathcal{T}_1)\alpha = \mathcal{T}_2(\mathcal{T}_1\alpha)$$

Theorem 1.27. 设 X, Y, Z 是域 F 上的线性空间，$\mathcal{T}_1 \in \mathrm{Hom}(X, Y), \mathcal{T}_2 \in \mathrm{Hom}(Z, Y)$，则 $\mathcal{T}_2\mathcal{T}_1 \in \mathrm{Hom}(X, Z)$。

Proof. 只需注意到对任意的 $\alpha, \beta \in X$，$k_1, k_2 \in F$，有：

$$\begin{aligned}(\mathcal{T}_2\mathcal{T}_1)(k_1\alpha + k_2\beta) &= \mathcal{T}_2[\mathcal{T}_1(k_1\alpha + k_2\beta)] = \mathcal{T}_2(k_1\mathcal{T}_1\alpha + k_2\mathcal{T}_1\beta)\\ &= k_1\mathcal{T}_2(\mathcal{T}_1\alpha) + k_2\mathcal{T}_2(\mathcal{T}_1\beta) = k_1(\mathcal{T}_2\mathcal{T}_1)\alpha + k_2(\mathcal{T}_2\mathcal{T}_1)\beta\end{aligned}$$ □

Definition 1.29. 设 X 是域 F 上的线性空间，$\mathcal{T} \in \mathrm{Hom}(X)$，定义 $\mathcal{T}$ 的正整数指数幂如下：

$$\mathcal{T}^n = \underbrace{\mathcal{T} \cdot \mathcal{T} \cdots \mathcal{T}}_{n\text{个}\mathcal{T}},\ n \in \mathbb{N}^+$$

若 $\mathcal{T}$ 可逆，还可以定义 $\mathcal{T}$ 的负整数指数幂如下：

$$\mathcal{T}^{-n} = (\mathcal{T}^{-1})^n,\ n \in \mathbb{N}^+$$

Definition 1.30. 设 X, Y 是域 F 上的线性空间，$\mathcal{T}$ 是 X 到 Y 上的一个线性映射，分别称：

$$\{\alpha \in X : \mathcal{T}\alpha = \mathbf{0}_Y\},\quad \{\mathcal{T}\alpha : \alpha \in X\}$$

为 $\mathcal{T}$ 的核 *(kernel)* 与象 *(image)*，将它们分别记作 $\mathrm{Ker}\,\mathcal{T}$ 和 $\mathrm{Im}\,\mathcal{T}$。

Property 1.2.1. 设 X, Y 是域 F 上的线性空间，$\mathcal{T}$ 是 X 到 Y 上的线性映射，则：

1. 若 $\mathcal{T}$ 可逆，则 $\mathcal{T}$ 是 X 到 Y 上的同构映射；

2. $\mathcal{T}\mathbf{0}_X = \mathbf{0}_Y$；
3. 对于任意的 $\alpha \in X$，有 $\mathcal{T}(-\alpha) = -\mathcal{T}\alpha$；
4. 对于任意的 $\alpha_1, \alpha_2, \ldots, \alpha_n \in X$, $k_1, k_2, \ldots, k_n \in F$，有：
$$\mathcal{T}\left(\sum_{i=1}^{n} k_i\alpha_i\right) = \sum_{i=1}^{n} k_i\mathcal{T}\alpha_i$$
这表明，如果 X 是有限维的，那么只要知道 X 的一个基在 $\mathcal{T}$ 下的象，那么 X 中所有向量在 $\mathcal{T}$ 下的象就都确定了。
5. 若 $\alpha_1, \alpha_2, \ldots, \alpha_n \in X$ 线性相关，则 $\mathcal{T}\alpha_1, \mathcal{T}\alpha_2, \ldots, \mathcal{T}\alpha_n$ 线性相关。
6. $\operatorname{Ker}\mathcal{T}$ 和 $\operatorname{Im}\mathcal{T}$ 分别是 X 和 Y 的子空间；
7. $\mathcal{T}$ 是单射当且仅当 $\operatorname{Ker}\mathcal{T} = \mathbf{0}_X$；
8. $\mathcal{T}$ 是满射当且仅当 $\operatorname{Im}\mathcal{T} = Y$。
9. $X/\operatorname{Ker}\mathcal{T}$ 与 $\operatorname{Im}\mathcal{T}$ 在映射：
$$\sigma : \alpha + \operatorname{Ker}\mathcal{T} \longrightarrow \mathcal{T}\alpha$$
下同构；
10. 若 X 是有限维的，则 $\operatorname{Ker}\mathcal{T}$ 和 $\operatorname{Im}\mathcal{T}$ 都是有限维的，且有：
$$\dim X = \dim(\operatorname{Ker}\mathcal{T}) + \dim(\operatorname{Im}\mathcal{T})$$
11. 若 $\dim X = \dim Y = n < +\infty$，则 $\mathcal{T}$ 是单射当且仅当 $\mathcal{T}$ 是满射。

Proof. (1)(2)(3)(4)(5)(8) 证明都是显然的，只需参考性质 1.1.3即可，这是因为线性映射只比同构映射少了双射这一条件，所以同构映射不涉及双射条件的性质对于线性映射也成立。

(6) 任取 $\alpha, \beta \in \operatorname{Ker}\mathcal{T}$ 和 $k_1, k_2 \in F$，则有：
$$\mathcal{T}(k_1\alpha + k_2\beta) = k_1\mathcal{T}\alpha + k_2\mathcal{T}\beta = \mathbf{0}$$
于是 $k_1\alpha + k_2\beta \in \operatorname{Ker}\mathcal{T}$，所以 $\operatorname{Ker}\mathcal{T}$ 是 X 的子空间。

任取 $\mathcal{T}\alpha, \mathcal{T}\beta \in \operatorname{Im}\mathcal{T}$ 和 $k_1, k_2 \in F$，则有：
$$k_1\mathcal{T}\alpha + k_2\mathcal{T}\beta = \mathcal{T}(k_1\alpha + k_2\beta)$$
因为 X 是一个线性空间，所以 $k_1\alpha + k_2\beta \in X$，于是 $\mathcal{T}(k_1\alpha + k_2\beta) \in \operatorname{Im}\mathcal{T}$，因此 $\operatorname{Im}\mathcal{T}$ 是 Y 的子空间。

(7) **充分性**：假设此时 $\mathcal{T}$ 不是单射，则存在 $\mathcal{T}\alpha, \mathcal{T}\beta \in \mathcal{T}$ 使得 $\mathcal{T}\alpha = \mathcal{T}\beta$ 且 $\alpha \neq \beta$，而此时 $\mathcal{T}\alpha - \mathcal{T}\beta = \mathcal{T}(\alpha - \beta) = \mathbf{0}_Y$，由已知条件可得 $\alpha - \beta = \mathbf{0}_X$，即 $\alpha = \beta$，矛盾。

必要性：由 (2) 可知 $\mathcal{T}\mathbf{0}_X=\mathbf{0}_Y$，因为 $\mathcal{T}$ 是一个单射，所以 $\operatorname{Ker}\mathcal{T}=\mathbf{0}_X$。

(8) 由满射的定义立即可得。

(9) 先证明 σ 是一个映射。若 $\alpha+\operatorname{Ker}\mathcal{T}=\beta+\operatorname{Ker}\mathcal{T}$，则 $\alpha-\beta\in\operatorname{Ker}\mathcal{T}$，即 $\mathcal{T}(\alpha-\beta)=\mathcal{T}\alpha-\mathcal{T}\beta=\mathbf{0}_Y$，于是 $\mathcal{T}\alpha=\mathcal{T}\beta$，所以 σ 是一个映射。

任取 $\alpha+\operatorname{Ker}\mathcal{T},\beta+\operatorname{Ker}\mathcal{T}\in X/\operatorname{Ker}\mathcal{T}$ 和 $k_1,k_2\in F$，则有：

$$\begin{aligned}\sigma[k_1(\alpha+\operatorname{Ker}\mathcal{T})+k_2(\beta+\operatorname{Ker}\mathcal{T})]&=\sigma(k_1\alpha+k_2\beta+\operatorname{Ker}\mathcal{T})=\mathcal{T}(k_1\alpha+k_2\beta)\\&=k_1\mathcal{T}\alpha+k_2\mathcal{T}\beta=k_1\sigma(\alpha+\operatorname{Ker}\mathcal{T})+k_2(\beta+\operatorname{Ker}\mathcal{T})\end{aligned}$$

所以 σ 是一个线性映射。

显然 σ 是一个满射。

若存在 $\alpha+\operatorname{Ker}\mathcal{T},\beta+\operatorname{Ker}\mathcal{T}\in X$ 满足 $\alpha+\operatorname{Ker}\mathcal{T}\neq\beta+\operatorname{Ker}\mathcal{T}$ 且 $\mathcal{T}\alpha=\mathcal{T}\beta$，则此时有 $\mathcal{T}(\alpha-\beta)=\mathcal{T}\alpha-\mathcal{T}\beta=\mathbf{0}_Y$，所以 $\alpha-\beta\in\operatorname{Ker}\mathcal{T}$，即 $\alpha+\operatorname{Ker}\mathcal{T}=\beta+\operatorname{Ker}\mathcal{T}$，矛盾，因此 σ 是个单射。

综上，σ 是一个双射且是一个线性映射，于是 $X/\operatorname{Ker}\mathcal{T}$ 与 $\operatorname{Im}\mathcal{T}$ 在 σ 下同构。

(10) 因为 X 是有限维的，所以 $X/\operatorname{Ker}\mathcal{T}$ 和 $\operatorname{Ker}\mathcal{T}$ 都是有限维的。由定理 1.23和 (9) 可知 $\dim(X/\operatorname{Ker}\mathcal{T})=\dim(\operatorname{Im}\mathcal{T})$，于是 $\operatorname{Im}\mathcal{T}$ 也是有限维的。由定理 1.26可得：

$$\dim(\operatorname{Im}\mathcal{T})=\dim(X/\operatorname{Ker}\mathcal{T})=\dim X-\dim(\operatorname{Ker}\mathcal{T})$$

(11) 由 (7) 和 (10) 可得：

$$\begin{aligned}\mathcal{T}\text{是单射}&\Leftrightarrow\operatorname{Ker}\mathcal{T}=\mathbf{0}_X\Leftrightarrow\dim(\operatorname{Ker}\mathcal{T})=0\\&\Leftrightarrow\dim Y=\dim X=\dim(\operatorname{Im}\mathcal{T})\Leftrightarrow\mathcal{T}\text{是满射}\end{aligned}$$ □

Definition 1.31. 设 X,Y 是域 F 上的线性空间，$\mathcal{T}\in\operatorname{Hom}(X,Y)$，称 $Y/\operatorname{Im}\mathcal{T}$ 为 $\mathcal{T}$ 的余核 *(cokernel)*。

Theorem 1.28. 设 X,Y 是域 F 上的线性空间，$\mathcal{T}\in\operatorname{Hom}(X,Y)$，则 $\mathcal{T}$ 是满射当且仅当 $\operatorname{Coker}\mathcal{T}=\mathbf{0}$。

Proof. $\mathcal{T}$ 是满射 $\Leftrightarrow\operatorname{Im}\mathcal{T}=Y\Leftrightarrow Y/\operatorname{Im}\mathcal{T}=\mathbf{0}$。这里的 $\mathbf{0}$ 实际上是商空间的零元，也即 Y。 □

1.2.2 线性映射的矩阵表示

Definition 1.32. 设 X,Y 分别为域 F 上的 m 维、n 维线性空间，$\mathcal{T}$ 是 X 到 Y 的一个线性映射。由性质 *1.2.1(4)* 可知

1.2.3 常见线性映射

Definition 1.33. 设 X,Y 是域 F 上的线性空间，$\mathcal{T}_1\in\operatorname{Hom}(X,Y)$，$\mathcal{T}_2\in\operatorname{Hom}(X)$，则：

1. 若对任意的 $\alpha \in X$，有 $\mathcal{T}_1\alpha = \mathbf{0}_Y$，则称 $\mathcal{T}_1$ 为 X 到 Y 的零映射 *(zero mapping)*，记作 $\mathcal{O}$；
2. 若对任意的 $\alpha \in X$，有 $\mathcal{T}_2\alpha = \alpha$，则称 $\mathcal{T}_2$ 为 X 到 Y 的恒等变换 *(identity transformation)*，记作 $\mathcal{I}$；
3. 给定 $k \in F$，若对任意的 $\alpha \in X$，有 $\mathcal{T}_2\alpha = k\alpha$，则称 $\mathcal{T}_2$ 为 X 到 Y 的数乘变换 *(scalar transformation)*，记作 $\mathcal{K}$；
4. 若 $\mathcal{T}_2^2 = \mathcal{I}$，则称 $\mathcal{T}_2$ 为对合变换 *(involution)*；
5. 若 $\mathcal{T}_2^2 = \mathcal{T}_2$，则称 $\mathcal{T}_2$ 为幂等变换 *(idempotent transformation)*；
6. 若存在 $n \in \mathbb{N}^+$ 使得 $\mathcal{T}_2^n = \mathcal{O}$，则称 $\mathcal{T}_2$ 为幂零变换 *(nilpotent transformation)*，使得 $\mathcal{T}_2^n = \mathcal{O}$ 成立的最小正整数 n 被称为是 $\mathcal{T}_2$ 的幂零指数 *(nilpotent index)*；
7. 若 E 和 W 是 X 的子空间，且有 $X = E \oplus W$，若对任意的 $\alpha = \alpha_1 + \alpha_2 \in X$，$\alpha_1 \in E$，$\alpha_2 \in W$，有：
$$\mathcal{T}_2\alpha = \alpha_1$$
则称 $\mathcal{T}_2$ 为平行于 W 在 E 上的投影变换 *(projection transformation)*，记为 $\mathcal{P}_E$。

Definition 1.34. 设 X 是域 F 上的线性空间，$\mathcal{T}_1, \mathcal{T}_2 \in \mathrm{Hom}(X)$。若 $\mathcal{T}_1\mathcal{T}_2 = \mathcal{T}_2\mathcal{T}_1 = \mathcal{O}$，则称 $\mathcal{T}_1$ 和 $\mathcal{T}_2$ 正交。

投影变换

Property 1.2.2. 设 X 是域 F 上的线性空间，E 和 W 是 X 的子空间，且有 $X = E \oplus W$，$\mathcal{P}_E$ 是平行于 W 在 E 上的投影变换，$\mathcal{P}_W$ 是平行于 E 在 W 上的投影变换，则：

1. $\mathcal{P}_E$ 是线性变换；
2. $\mathcal{P}_E$ 是幂等变换；
3. 幂等变换 $\mathcal{T} \in \mathrm{Hom}(X)$ 是平行于 $\mathrm{Ker}\,\mathcal{T}$ 在 $\mathrm{Im}\,\mathcal{T}$ 上的投影变换，此时 $X = \mathrm{Im}\,\mathcal{T} \oplus \mathrm{Ker}\,\mathcal{T}$；
4. $\mathcal{P}_E$ 与 $\mathcal{P}_W$ 正交；
5. $\mathcal{P}_E + \mathcal{P}_W = \mathcal{I}$；
6. 若 $\mathcal{T}_1, \mathcal{T}_2 \in \mathrm{Hom}(X)$，$\mathcal{T}_1$ 和 $\mathcal{T}_2$ 是正交的幂等变换，且 $\mathcal{T}_1 + \mathcal{T}_2 = \mathcal{I}$，则 $X = \mathrm{Im}\,\mathcal{T}_1 \oplus \mathrm{Im}\,\mathcal{T}_2$，并且有 $\mathcal{T}_1$ 是平行于 $\mathrm{Im}\,\mathcal{T}_2$ 在 $\mathrm{Im}\,\mathcal{T}_1$ 上的投影，$\mathcal{T}_2$ 是平行于 $\mathrm{Im}\,\mathcal{T}_1$ 在 $\mathrm{Im}\,\mathcal{T}_2$ 上的投影；

Proof. (1) 任取 $\alpha, \beta \in X$ 和 $k_1, k_2 \in F$，其中 $\alpha = \alpha_1 + \alpha_2$，$\beta = \beta_1 + \beta_2$，$\alpha_1, \beta_1 \in E$，$\alpha_2, \beta_2 \in W$，于是有：

$$\mathcal{P}_E(k_1\alpha + k_2\beta) = \mathcal{P}_E[(k_1\alpha_1 + k_2\beta_1) + (k_1\alpha_2 + k_2\beta_2)] = k_1\alpha_1 + k_2\beta_1 = k_1\mathcal{P}_E\alpha + k_2\mathcal{P}_E\beta$$

所以 $\mathcal{P}_E$ 是线性变换。

(2) 任取 $\alpha = \alpha_1 + \alpha_2 \in X$, $\alpha_1 \in E$, $\alpha_2 \in W$，则：

$$\mathcal{P}_E(\mathcal{P}_E\alpha) = \mathcal{P}_E(\alpha_1) = \alpha_1 = \mathcal{P}_E\alpha$$

所以 $\mathcal{P}_E$ 是幂等变换。

(3) 任取 $\alpha \in X$，则 $\mathcal{T}\alpha \in \operatorname{Im}\mathcal{T}$。因为：

$$\mathcal{T}(\alpha - \mathcal{T}\alpha) = \mathcal{T}\alpha - \mathcal{T}^2\alpha = \mathcal{T}\alpha - \mathcal{T}\alpha = \mathbf{0}$$

所以 $\alpha - \mathcal{T}\alpha \in \operatorname{Ker}\mathcal{T}$。因为 $\alpha = \mathcal{T}\alpha + \alpha - \mathcal{T}\alpha$，所以 $X = \operatorname{Im}\mathcal{T} + \operatorname{Ker}\mathcal{T}$。

任取 $\beta \in \operatorname{Im}\mathcal{T} \cap \operatorname{Ker}\mathcal{T}$，则存在 $\gamma \in X$ 使得 $\mathcal{T}\gamma = \beta$，且有 $\mathcal{T}\beta = \mathbf{0}$，于是：

$$\mathbf{0} = \mathcal{T}\beta = \mathcal{T}(\mathcal{T}\gamma) = \mathcal{T}^2\gamma = \mathcal{T}\gamma = \beta$$

所以 $\operatorname{Im}\mathcal{T} \cap \operatorname{Ker}\mathcal{T} = \mathbf{0}$，由定理 1.21(3) 可知 $X = \operatorname{Im}\mathcal{T} \oplus \operatorname{Ker}\mathcal{T}$。

记 $\operatorname{Im}\mathcal{T} = E$, $\operatorname{Ker}\mathcal{T} = W$，对上述 $\alpha = \mathcal{T}\alpha + \alpha - \mathcal{T}\alpha$，有 $\mathcal{P}_E\alpha = \mathcal{T}\alpha$。由 α 的任意性，$\mathcal{T} = \mathcal{P}_E$。

(4) 任取 $\alpha = \alpha_1 + \alpha_2 \in X$, $\alpha_1 \in E$, $\alpha_2 \in W$，则：

$$\mathcal{P}_E(\mathcal{P}_W\alpha) = \mathcal{P}_E\alpha_2 = \mathbf{0},\ \mathcal{P}_W(\mathcal{P}_E\alpha) = \mathcal{P}_W\alpha_2 = \mathbf{0}$$

所以 $\mathcal{P}_E\mathcal{P}_W = \mathcal{P}_W\mathcal{P}_E = \mathcal{O}$，即 $\mathcal{P}_E$ 与 $\mathcal{P}_W$ 正交。

(5) 任取 $\alpha = \alpha_1 + \alpha_2 \in X$, $\alpha_1 \in E$, $\alpha_2 \in W$，则：

$$(\mathcal{P}_E + \mathcal{P}_W)\alpha = \mathcal{P}_E\alpha + \mathcal{P}_W\alpha = \alpha_1 + \alpha_2 = \alpha$$

于是 $\mathcal{P}_E + \mathcal{P}_W = \mathcal{I}$。

(6) 任取 $\alpha \in X$，则 $\alpha = (\mathcal{T}_1 + \mathcal{T}_2)\alpha = \mathcal{T}_1\alpha + \mathcal{T}_2\alpha$，所以 $X = \operatorname{Im}\mathcal{T}_1 + \operatorname{Im}\mathcal{T}_2$。

任取 $\beta \in \operatorname{Im}\mathcal{T}_1 \cap \operatorname{Im}\mathcal{T}_2$，则存在 $\gamma, \delta \in X$ 使得 $\mathcal{T}_1\gamma = \beta$, $\mathcal{T}_2\delta = \beta$，于是有 $\beta = \mathcal{T}_1\gamma = \mathcal{T}_1^2\gamma = \mathcal{T}_1(\mathcal{T}_1\gamma) = \mathcal{T}_1\beta$。因为 $\mathcal{T}_1$ 与 $\mathcal{T}_2$ 正交，所以：

$$\mathcal{T}_1\beta = \mathcal{T}_1(\mathcal{T}_2\delta) = (\mathcal{T}_1\mathcal{T}_2)\delta = \mathbf{0}$$

于是 $\beta = \mathbf{0}$，即 $\operatorname{Im}\mathcal{T}_1 \cap \operatorname{Im}\mathcal{T}_2 = \mathbf{0}$。由定理 1.21可得 $X = \operatorname{Im}\mathcal{T}_1 \oplus \operatorname{Im}\mathcal{T}_2$。

任取 $\varepsilon = \mathcal{T}_1\varepsilon_1 + \mathcal{T}_2\varepsilon_2 \in X$，其中 $\varepsilon_1, \varepsilon_2 \in X$。因为 $\mathcal{T}_1$ 与 $\mathcal{T}_2$ 正交、$\mathcal{T}_1$ 是幂等变换，所以显然有：

$$\mathcal{T}_1\varepsilon = \mathcal{T}_1(\mathcal{T}_1\varepsilon_1 + \mathcal{T}_2\varepsilon_2) = \mathcal{T}_1^2\varepsilon_1 + (\mathcal{T}_1\mathcal{T}_2)\varepsilon_2 = \mathcal{T}_1\varepsilon_1$$

于是 $\mathcal{T}_1$ 是平行于 $\operatorname{Im}\mathcal{T}_2$ 在 $\operatorname{Im}\mathcal{T}_1$ 上的投影。$\mathcal{T}_2$ 同理。 □

Corollary 1.6. *设 X 是域 F 上的线性空间，由性质 1.2.2(3) 可得到如下推论：*

1. *若 $X = E \oplus W$，$\mathcal{P}_E$ 为平行于 W 在 E 上的投影变换，则：*

$$E = \operatorname{Im}\mathcal{P}_E,\ W = \operatorname{Ker}\mathcal{P}_E$$

2. X 的任一子空间 E 是平行于 E 的一个补空间在 E 上的投影变换的象；

3. X 的任一子空间 E 是平行于 E 在 E 的一个补空间上的投影变换的核。

Proof. (1) 任取 $\alpha = \alpha_1 + \alpha_2 \in X$, $\alpha_1 \in E$, $\alpha_2 \in W$，则 $\mathcal{P}_E\alpha = \alpha_1 \in E$，所以 $\operatorname{Im}\mathcal{P}_E \subset E$。任取 $\beta \in E$，有 $\mathcal{P}_E\beta = \beta \in \operatorname{Im}\mathcal{P}_E$，所以 $E \subset \operatorname{Im}\mathcal{P}_E$。因此 $E = \operatorname{Im}\mathcal{P}_E$。

任取 $\gamma \in W$，有 $\mathcal{P}_E\gamma = \mathbf{0}$，所以 $\gamma \in \operatorname{Ker}\mathcal{P}_E$，于是 $W \subset \operatorname{Ker}\mathcal{P}_E$。任取 $\delta = \delta_1 + \delta_2 \in \operatorname{Ker}\mathcal{P}_E$, $\delta_1 \in E$, $\delta_2 \in W$，则 $\mathcal{P}_E\delta = \delta_1 = \mathbf{0}$，所以 $\delta = \delta_2 \in W$，于是 $\operatorname{Ker}\mathcal{P}_E \subset W$。因此 $W = \operatorname{Ker}\mathcal{P}_E$。

(2) 由定理 1.22可知 E 必定存在一个补空间 W，于是 $X = E \oplus W$，由 (1) 即可得到 $E = \operatorname{Im}\mathcal{P}_E$。

(3) 与 (2) 类似可得。

□

1.3 内积空间

内积空间的定义

Definition 1.35. 设 X 为实（复）数域 K 上的线性空间。若 X 中任意一对元素 x, y 都对应于 K 中的一个数，记为 (x, y)，满足：

1. 线性性：$(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$，这里 $z \in X$。

2. 对称性：当 K 为实数域时，$(x, y) = (y, x)$；当 K 为复数域时，$(x, y) = \overline{(y, x)}$。

3. 非负性：$(x, x) \geqslant 0$，等号成立当且仅当 $x = \mathbf{0}$。

那么就称 X 为实（复）内积空间 *(inner product space)*，称 (x, y) 为元素 x, y 的内积 *(inner product)*。

Chapter 2

矩阵

2.1 矩阵空间

Definition 2.1. 由 $s \cdot m$ 个数排成 s 行、m 列的一张表称为一个 $s \times m$矩阵 *(matrix)*，通常用大写英文字母表示，其中的每一个数称为这个矩阵的一个元素，第 i 行与第 j 列交叉位置的元素称为矩阵的 (i, j) 元，记作 $A(i; j)$。一个 $s \times m$ 矩阵可以简单地记作 $A_{s \times m}$。如果矩阵 A 的 (i, j) 元是 a_{ij}，那么可以记作 $A = (a_{ij})$。如果一个矩阵的行数和列数相同，则称它为方阵，n 行 n 列的方阵也成为 n 阶矩阵。对于两个矩阵 A 和 B，如果它们的行数都等于 s 且列数都等于 m，同时还有 $A(i; j) = B(i; j), i = 1, 2, \ldots, s,\ j = 1, 2, \ldots, m$，那么称 A 和 B 相等，记作 $A = B$。

2.1.1 矩阵的运算

加减法与数量乘法

Definition 2.2. 将数域 K 上所有 $s \times m$ 矩阵组成的集合记作 $M_{s \times m}(K)$，当 $s = m$ 时，$M_{s \times s}(K)$ 可以简记作 $M_s(K)$。在 $M_{s \times m}(K)$ 中定义如下运算：

1. 加法：

$$\forall\ A = (a_{ij}), B = (b_{ij}) \in M_{s \times m}(K),\ A + B = (a_{ij} + b_{ij})$$

2. 纯量乘法：

$$\forall\ k \in K,\ \forall\ A = (a_{ij}),\ kA = (ka_{ij})$$

那么 $M_{s \times m}(K)$ 构成一个线性空间。

Proof. 首先证明如上定义的加法和纯量乘法对 $M_{s \times m}(K)$ 是封闭的。由数域中加法和乘法的封闭性，$a_{ij} + b_{ij} \in K$，$ka_{ij} \in K\ i = 1, 2, \ldots, s,\ j = 1, 2, \ldots, m$，所以如上定义的加法与纯量乘法对 $M_{s \times m}(K)$ 是封闭的。

接下来证明如上定义的加法和纯量乘法满足线性空间中的 8 条运算法则：

1. 因为数域内的数满足加法交换律与加法结合律，所以 $M_{s\times m}(K)$ 上的加法满足线性空间运算法则 (1)(2)；

2. 取一个元素全为 0 的 $s\times m$ 矩阵，将其记作 $\mathbf{0}$，显然对 $\forall\, A \in M_{s\times m}(K)$，有 $A+\mathbf{0}=A$，因此 $M_{s\times m}(K)$ 中存在零元且它就是元素全为 0 的 $s\times m$ 矩阵，称其为零矩阵 (zero matrix)，就记作 $\mathbf{0}$。因此，$M_{s\times m}(K)$ 上的加法满足线性空间运算法则 (3)；

3. 对 $\forall\, A \in M_{s\times m}(K)$，取 $-A=(-a_{ij})$，则有 $A+(-A)=(a_{ij}-a_{ij})=\mathbf{0}$。由 A 的任意性，$M_{s\times m}(K)$ 中的每个元素都具有负元，将 $\forall\, A \in M_{s\times m}(K)$ 的负元就记作 $-A$。因此，$M_{s\times m}(K)$ 上的加法满足线性空间运算法则 (4)；

4. 因为数域内的数满足乘法结合律和乘法分配律，同时它们乘 1 的积是自身，所以 $M_{s\times n}$ 上的纯量乘法满足线性空间运算法则 (5)(6)(7)(8)。

证明完毕。 □

Definition 2.3. 定义 $M_{s\times m}(K)$ 上矩阵的减法如下：设 $A,B\in M_{s\times m}(K)$，则：

$$A-B \overset{def}{=} A+(-B)$$

乘法

Definition 2.4. 设 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times m}$，令 $C=(c_{ij})_{s\times m}$，其中：

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj},\ i=1,2,\ldots,s,\ j=1,2,\ldots,m$$

则矩阵 C 称作矩阵 A 与 B 的乘积，记作 $C=AB$。

初等变换

Definition 2.5. 称以下变换为矩阵的初等行变换 *(elementary row operation)*：

1. 把一行的倍数加到另一行上；
2. 互换两行的位置；
3. 用一个非零数乘某一行。

称以下变换为矩阵的初等列变换 *(elementary column operation)*：

1. 把一列的倍数加到另一列上；
2. 互换两列的位置；
3. 用一个非零数乘某一列。

2.1.2 矩阵的行列式

2.2 矩阵的向量空间

Definition 2.6. 设 $A=(\alpha_1,\alpha_2,\dots,\alpha_n)\in M_{m\times n}(K)$，将：

$$\left\{\sum_{i=1}^{n} k_i\alpha_i : k_i\in K\right\} \overset{def}{=} \mathcal{M}(A)$$

Theorem 2.1. 设 $A\in M_{m\times n}(K)$，则：

$$\mathcal{M}(A)=\mathcal{M}(AA^T)$$

Proof. 由定义，显然 $\mathcal{M}(AA^T)\subset\mathcal{M}(A)$。对于任意的 $x\perp\mathcal{M}(AA^T)$，有 $x^TAA^T=\mathbf{0}$，于是 $||A^Tx||^2=x^TAA^Tx=0$，即 $A^Tx=\mathbf{0}$，于是 $x\perp\mathcal{M}(A)$。 □

回头改证明，同时注意数域问题

Theorem 2.2. 设 $A\in M_{m\times n}(\mathbb{C})$，则有：

$$\operatorname{rank}(AA^H)=\operatorname{rank}(A^HA)=\operatorname{rank}(A)$$

线性方程组解链接

Proof. 只需证明方程 $A^HAx=\mathbf{0}$ 与 $Ax=\mathbf{0}$ 同解。注意到 $Ax=\mathbf{0}$ 则必然有 $A^HAx=\mathbf{0}$，而若 $A^HAx=\mathbf{0}$，则必有 $x^HA^HAx=||Ax||=0$，所以 $Ax=\mathbf{0}$。

链接方程组秩与维数的公式

$$n-\operatorname{rank}(A^HA)=n-\operatorname{rank}(A)$$

所以：

$$\operatorname{rank}(A^HA)=\operatorname{rank}(A)$$

同理可得：

$$\operatorname{rank}(AA^H)=\operatorname{rank}(A^H)=\operatorname{rank}(A)$$

于是有：

$$\operatorname{rank}(AA^H)=\operatorname{rank}(A^HA)=\operatorname{rank}(A)$$ □

2.3 线性方程组

Definition 2.7. 设 $x_1,x_2,\dots,x_n$ 为 n 个未知数，若一个方程具有如下形式：

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b$$

其中，$a_1,a_2,\dots,a_n$ 为系数 *(coefficient)*，b 为常数项 *(constant term)*，则称该方程为线性方程 *(linear equation)*。由 m 个形如上式的方程组成的方程组：

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

被称为 n 元线性方程组 *(system of linear equations, SLE)*。由矩阵乘法的定义，该方程组也可以写作矩阵形式：

$$Ax = b$$

其中，

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Definition 2.8. 给定线性方程组 $Ax = b$，称如下矩阵

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

为该线性方程组的增广矩阵 *(augmented matrix)*，记为 $[A|b]$。

Definition 2.9. 一个矩阵被称为行阶梯形矩阵 *(row echelon form, REF)*，如果它满足以下条件：

1. 所有零行（全为零的行）位于非零行的下方；
2. 若某一行非零，则该行的首个非零元素（称为主元 *(pivot)*）位于该行之前所有行的主元右侧。

一个矩阵被称为简化行阶梯形矩阵 *(reduced row echelon form, RREF)*，如果满足以下条件：

1. 它是阶梯形矩阵；
2. 每个非零行的主元都是 1；
3. 每个主元所在列的其他元素均为 0。

Theorem 2.3. 任意一个矩阵都可以经过一系列初等行变换化成行阶梯形矩阵，进而可以经过一系列初等行变换化成简化行阶梯形矩阵。

Definition 2.10. 设增广矩阵化简后变为阶梯形矩阵，称每一行主元所在列所对应的未知数为主变量 *(pivot variable)*，同时称非主元所在列对应的未知数为自由未知量 *(free variable)*。

Theorem 2.4. 数域 K 上线性方程组 $Ax = b$ 有解的充分必要条件为 $\mathrm{rank}(A) = \mathrm{rank}([A|b])$。

Proof. □

2.4 矩阵的等价关系

2.4.1 相抵

Definition 2.11. $A, B \in M_{s\times m}(K)$，如果满足下述条件中的任意一个：

1. A 能够通过初等行变换和初等列变换变成 B；

2. 存在数域 K 上的 s 阶初等矩阵 $P_1, P_2, \ldots, P_t$ 与 m 阶初等矩阵 $Q_1, Q_2, \ldots, Q_n$ 使得：

$$P_t \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_n = B$$

3. 存在数域 K 上的 s 阶可逆矩阵 P 与 m 阶可逆矩阵 Q 使得：

$$PAQ = B$$

则称 A 与 B 相抵 *(equivalent)*。

上述三个条件显然是等价的。

Theorem 2.5. 相抵是 $M_{s\times m}(K)$ 上的一个等价关系。在相抵关系下，矩阵 A 的等价类称为 A 的**相抵类**。

Proof. 证明是显然的。 □

Theorem 2.6. 设 $A \in M_{s\times m}(K)$，且 $\operatorname{rank}(A) = r$。如果 $r > 0$，那么 A 相抵于如下形式的矩阵：

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

称该矩阵为 A 的**相抵标准形**。如果 $r = 0$，则 A 相抵于零矩阵，此时称零矩阵为 A 的**相抵标准形**。

Proof. 一个矩阵通过初等行变换一定可以变成一个简化行阶梯型矩阵，再由初等列变换即可得到上述矩阵。 □

Theorem 2.7 (相抵的完全不变量)**.** $A, B \in M_{s\times m}(K)$，$A$ 与 B 相抵当且仅当它们的秩相同。

Proof. **(1) 必要性**：初等行变换和初等列变换不改变矩阵的秩。

(2) 充分性：若 A, B 的秩相同，则它们的相抵标准形相同。因为相抵是一个等价关系，由等价关系的对称性与传递性即可得到 A 与 B 相抵。 □

2.4.2 相似

Definition 2.12. $A, B \in M_n(K)$。如果存在可逆矩阵 $P \in M_n(K)$，使得：

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

则称 A 与 B**相似** *(similar)*。

Theorem 2.8. 相似是 $M_n(K)$ 上的一个等价关系。在相似关系下，矩阵 A 的等价类称为 A 的**相似类**。

Proof. 证明是显然的。 □

Property 2.4.1 (相似的不变量)**.** 相似的矩阵具有相同的行列式值、秩、迹、特征多项式、特征值（包括重数相同）。

Proof. 设 $A, B \in M_n(K)$ 且 A 与 B 相似，于是存在可逆矩阵 $P \in M_n(K)$ 使得 $P^{-1}AP = B$。

(1)$|A| = |P^{-1}AP| = |P^{-1}|\,|B|\,|P| = |P^{-1}|\,|P|\,|B| = |B|$。

(2) 初等行变换与初等列变换不改变矩阵的秩。

(3) 由**??**(3) 可得 $\mathrm{tr}(A) = \mathrm{tr}(P^{-1}BP) = \mathrm{tr}(BPP^{-1}) = \mathrm{tr}(B)$。

(4)(5) 参考定理 2.23。 □

2.4.3 合同

Definition 2.13. $A, B \in M_n(K)$。如果存在可逆矩阵 $C \in M_n(K)$，使得：

$$C^TAC = B$$

则称 A 与 B**合同** *(congruent)*，记作 $A \cong B$。如果对称矩阵 A 合同于一个对角矩阵，那么称这个对角矩阵为 A 的一个**合同标准形**。

Theorem 2.9. 合同是 $M_n(K)$ 上的一个等价关系。在合同关系下，矩阵 A 的等价类称为 A 的**合同类**。

Proof. 证明是显然的。 □

Definition 2.14. 对 n 阶矩阵的行作初等行变换，再对该矩阵的同样标号的列作相同的初等列变换，这种变换被称为**成对初等行、列变换**。

Lemma 2.1. $A, B \in M_n(K)$，则 A 合同于 B 当且仅当 A 经过一系列成对初等行、列变换可以变成 B，此时对 I 作其中的初等列变换即可得到可逆矩阵 C，使得 $C^TAC = B$。

Proof. 由可逆矩阵的初等矩阵分解，可得：

$$\begin{aligned} A \cong B &\Leftrightarrow \text{存在数域 } K \text{ 上的可逆矩阵 } C\text{，使得} C^TAC = B \\ &\Leftrightarrow \text{存在数域 } K \text{ 上的初等矩阵 } P_1, P_2, \ldots, P_t \text{ 使得} \\ &\qquad C = P_1P_2\cdots P_t \\ &\qquad P_t^T \cdots P_2^T P_1^T A P_1 P_2 \cdots P_t = B \end{aligned}$$

□

Theorem 2.10. 数域 K 上的任一对称矩阵都合同于一个对角矩阵。

Proof. 对数域 K 上对称矩阵的阶数 n 作数学归纳法，。

当 $n=1$ 时，因为矩阵合同于自身，同时一阶矩阵都是对角矩阵，所以结论成立。

假设 $n-1$ 阶对称矩阵都合同于对角矩阵，考虑 n 阶矩阵 $A=(a_{ij})$。

情形一： $a_{11}\neq 0$

把 A 写成分块矩阵的形式，然后对 A 作初等行变换与初等列变换可得：

$$\begin{pmatrix} a_{11} & A_1 \\ A_1^T & A_2 \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & A_1 \\ \mathbf{0} & A_2-a_{11}^{-1}A_1^TA_1 \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & A_2-a_{11}^{-1}A_1^TA_1 \end{pmatrix}$$

于是有：

$$\begin{pmatrix} 1 & \mathbf{0} \\ -a_{11}^{-1}A_1^T & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & A_1 \\ A_1^T & A_2 \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}A_1 \\ \mathbf{0} & I_{n-1} \end{pmatrix} = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & A_2-a_{11}^{-1}A_1^TA_1 \end{pmatrix}$$

因为 A 是一个对称矩阵，所以 A_2 是一个对称矩阵，于是：

$$(A_2-a_{11}^{-1}A_1^TA_1)^T = A_2^T-a_{11}^{-1}A_1^T(A_1^T)^T = A_2-a_{11}^{-1}A_1^TA_1$$

所以 $A_2-a_{11}^{-1}A_1'A_1$ 是 $n-1$ 阶对称矩阵。由归纳假设可知存在可逆矩阵 $C\in M_{n-1}(K)$ 使得 $C^T(A_2-a_{11}^{-1}A_1'A_1)C=D$，其中 D 是一个对角矩阵，即：

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C^T \end{pmatrix} \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & A_2-a_{11}^{-1}A_1^TA_1 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix} = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$$

于是有：

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ -a_{11}^{-1}A_1 & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & A_1 \\ A_1^T & A_2 \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}A_1 \\ \mathbf{0} & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix} = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$$

因为：

$$\left[\begin{pmatrix} 1 & -a_{11}^{-1}A_1 \\ \mathbf{0} & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix}\right]^T = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix}^T \begin{pmatrix} 1 & -a_{11}^{-1}A_1 \\ \mathbf{0} & I_{n-1} \end{pmatrix}^T = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ -a_{11}^{-1}A_1 & I_{n-1} \end{pmatrix}$$

并且：

$$\begin{pmatrix} 1 & -a_{11}^{-1}A_1 \\ \mathbf{0} & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C \end{pmatrix}$$

是一个可逆矩阵，所以 A 合同于对角矩阵：

$$\begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$$

情形二： $a_{11}=0$, **存在** $i\neq 1$ **使得** $a_{ii}\neq 0$

把 A 的第 $1, i$ 行呼唤，再把所得矩阵的第 $1, i$ 列呼唤，得到的矩阵 B 的 $(1,1)$ 元即为 $a_{ii} \neq 0$。根据情形一的讨论，B 合同于一个对角矩阵。因为 B 是由 A 作成对初等行、列变换得到的，由引理 2.1可得 $A \cong B$。由合同的传递性，A 也合同于一个对角矩阵。

情形三：$a_{ii} = 0,\ \forall\, i = 1, 2, \ldots, n$，**存在** $a_{ij} \neq 0,\ i \neq j$

把 A 的第 j 行加到第 i 行上，再把所得矩阵的第 j 列加到第 i 列上，得到的矩阵 E 的 (i,i) 元即为 $2a_{ij} \neq 0$。由情形二的讨论，E 合同于一个对角矩阵。因为 E 是由 A 作成对初等行、列变换得到的，由引理 2.1可得 $A \cong E$。由合同的传递性，A 也合同于一个对角矩阵。

情形四：$A = \mathbf{0}$

因为 $\mathbf{0}$ 是一个对角矩阵，所以结论显然成立。 □

Theorem 2.11. 设对角矩阵 B 是对称矩阵 A 的合同标准形，则 B 对角线上不为 0 的元素的个数等于 A 的秩。

Proof. 因为 $A \cong B$，所以存在可逆矩阵 C 使得 $C^T AC = B$，于是 $\operatorname{rank}(A) = \operatorname{rank}(B)$。 □

实对称矩阵的合同规范形

Theorem 2.12. 对于任意的对称矩阵 $A \in M_n(\mathbb{R})$，A 都合同于对角矩阵 $\operatorname{diag}\{1, 1, \ldots, 1, -1, -1, \ldots, -1, 0, 0, \ldots, 0\}$，系数为 1 的平方项个数称为 A 的正惯性指数 *(positive inertia index)*，系数为 -1 的平方项个数称为 A 的负惯性指数 *(negative inertia index)*，这个对角矩阵称为 A 的**合同规范形**。

Proof. 任取矩阵 $A \in M_n(\mathbb{R})$，由定理 2.10可得 A 合同一个对角矩阵 B。对 B 作成对初等行、列变换可将 B 对角线上的元素重新排列，使得正值在前，负值在中间，零值在最后，如此得到对角矩阵 C，C 可写作：

$$C = \begin{pmatrix} c_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_p & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -c_{p+1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & -c_{p+2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -c_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

其中 $c_1, c_2, \ldots, c_r > 0$。再对 C 作成对初等行、列变换，即先对第 i 行除 $\sqrt{c_i}$，再对第 i 列

除 $\sqrt{c_i}$，$i=1,2,\dots,n$，即可得到对角矩阵 D：

$$D=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

由引理 2.1可得，$D\cong C$，$C\cong B$，又因为 $A\cong B$，由合同的传递性与对称性即可得 $A\cong D$。由 A 的任意性结论得证。 □

复对称矩阵的合同规范形

Theorem 2.13. 对于任意的 $A\in M_n(\mathbb{C})$，A 都合同于对角矩阵 $\mathrm{diag}\{1,1,\dots,1,0,0,\dots,0\}$，这个对角矩阵称为 A 的合同规范形。

Proof. 任取矩阵 $A\in M_n(\mathbb{C})$，由定理 2.10可得 $A\cong B=\mathrm{diag}\{b_1,b_2,\dots,b_r,0,0,\dots,0\}$，其中 r 是矩阵 B 的秩，$b_1,b_2,\dots,b_r\neq 0$。设 $b_j=r_j\cos\theta_j+ir_j\sin\theta_j$，$\theta_j\in[0,2\pi)$，$j=1,2,\dots,r$。因为：

$$\left[\sqrt{r_j}\left(\cos\frac{\theta_j}{2}+i\sin\frac{\theta_j}{2}\right)\right]^2=b_j$$

将 $\sqrt{r_j}\left(\cos\frac{\theta_j}{2}+i\sin\frac{\theta_j}{2}\right)$ 记作 $\sqrt{b_j}$，作成对初等行、列变换，即先对第 j 行除 $\sqrt{b_j}$，再对第 j 列除 $\sqrt{b_j}$，则可得到矩阵 $C=\mathrm{diag}\{1,1,\dots,1,0,0\dots,0\}$，其中 1 的个数为 r。由引理 2.1可得，$B\cong C$。因为 $A\cong B$，由合同的传递性，$A\cong C$。由 A 的任意性，结论成立。 □

2.5 相抵的应用

2.5.1 广义逆

Definition 2.15. 设 $A\in M_{m\times n}(K)$，一切满足方程组：

$$AXA=A$$

的矩阵 X 都被称为是 A 的*广义逆 (generalized inverse)*，记为 A^-。

Theorem 2.14. 设非零矩阵 $A \in M_{m\times n}(K)$，$\operatorname{rank}(A) = r$ 且：

$$A = P \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q$$

其中 P, Q 分别为数域 K 上的 m 阶可逆矩阵和 n 阶可逆矩阵，则矩阵方程：

$$AXA = A$$

一定有解，且其通解可表示为：

$$X = Q^{-1} \begin{pmatrix} I_r & B \\ C & D \end{pmatrix} P^{-1}$$

其中 B, C, D 分别为数域 K 上任意的 $r\times(m-r)$, $(n-r)\times r$, $(n-r)\times(m-r)$ 矩阵。

Proof. 若 X 是上述矩阵方程的一个解，则：

$$P \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QXP \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q = P \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q$$
$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QXP \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

将 QXP 写作如下分块矩阵的形式：

$$QXP = \begin{pmatrix} H & B \\ C & D \end{pmatrix}$$

其中 H, B, C, D 分别为数域 K 上任意的 $r\times r$, $r\times(m-r)$, $(n-r)\times r$, $(n-r)\times(m-r)$ 矩阵。于是：

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} H & B \\ C & D \end{pmatrix} \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$
$$\begin{pmatrix} H & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

所以 $H = I_r$，因此：

$$X = Q^{-1} \begin{pmatrix} I_r & B \\ C & D \end{pmatrix} P^{-1} \qquad \square$$

Property 2.5.1. 设 $A, \in M_{m\times n}(K)$, $B \in M_{m\times q}(K)$, $C \in M_{p\times n}(K)$，则广义逆 A^- 具有如下性质：

1. A^- 唯一的充分必要条件为 A 可逆，此时 $A^- = A^{-1}$；
2. $\operatorname{rank}(A^-) \geqslant \operatorname{rank}(A) = \operatorname{rank}(AA^-) = \operatorname{rank}(A^-A)$；

3. 若 $\mathcal{M}(B) \subset \mathcal{M}(A), \mathcal{M}(C) \subset \mathcal{M}(A^T)$，则 $C^T A^- B$ 与 A^- 的选择无关；

4. $A(A^T A)^- A^T$ 与 $(A^T A)^-$ 的选择无关；

数域问题

5. $A(A^T A)^- A^T A = A,\ A^T A(A^T A)^- A^T = A^T$；

6. 若 A 对称，则 $[(A)^-]^T = (A)^-$。

Proof. (1) **充分性**：若 A 可逆，则 $r=n$，由 A^- 的通解公式，显然此时 A^- 唯一。

必要性：若 A^- 唯一，则 $r=n$，显然此时 A 可逆。

(2) 由 A^- 的通解公式，$\mathrm{rank}(A^-) \geqslant r = \mathrm{rank}(A)$。因为：

$$AA^- = P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QQ^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix} P^{-1} = P\begin{pmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{pmatrix} p^{-1}$$

$$A^- A = Q^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix} P^{-1} P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q = Q^{-1}\begin{pmatrix} I_r & \mathbf{0} \\ C & \mathbf{0} \end{pmatrix} Q$$

显然，$\mathrm{rank}(AA^-) = \mathrm{rank}(A^- A) = \mathrm{rank}(A) = r$。

(3) 由已知条件，存在矩阵 D_1, D_2 使得 $B = AD_1,\ C = A^T D_2$，于是：

$$C^T A^- B = D_2^T AA^- AD_1 = D_2^T AD_1$$

(4) 由定理 2.1可知 $\mathcal{M}(A^T) = \mathcal{M}(A^T A)$，于是存在矩阵 B 使得 $A^T = A^T AB$，所以有：

$$A(A^T A)^- A^T = B^T A^T A(A^T A)^- A^T AB = B^T A^T AB$$

与 $(A^T A)^-$ 无关。

(5) 设 $B = A(A^T A)^- A^T A - A$，则：

$$\begin{aligned}
B^T B &= \{A^T A[(A^T A)^-]^T A^T - A^T\}[A(A^T A)^- A^T A - A] \\
&= A^T A[(A^T A)^-]^T A^T A(A^T A)^- A^T A - A^T A[(A^T A)^-]^T A^T A \\
&\quad - A^T A(A^T A)^- A^T A + A^T A \\
&= A^T A[(A^T A)^-]^T A^T A - A^T A[(A^T A)^-]^T A^T A - A^T A + A^T A = \mathbf{0}
\end{aligned}$$

所以 $B = \mathbf{0}$（考虑 $B^T B$ 主对角线上的元素），于是 $A(A^T A)^- A^T A = A$。

设 $C = A^T A(A^T A)^- A^T - A^T$，则：

$$\begin{aligned}
CC^T &= [A^T A(A^T A)^- A^T - A^T]\{A[(A^T A)^-]^T A^T A - A\} \\
&= A^T A(A^T A)^- A^T A[(A^T A)^-]^T A^T A - A^T A(A^T A)^- A^T A \\
&\quad - A^T A[(A^T A)^-]^T A^T A + A^T A \\
&= A^T A[(A^T A)^-]^T A^T A - A^T A - A^T A[(A^T A)^-]^T A^T A + A^T A = \mathbf{0}
\end{aligned}$$

所以 $C = \mathbf{0}$，于是 $A^T A(A^T A)^- A^T = A^T$。

(6) 此时有：

$$AXA = A \Leftrightarrow A^T X^T A^T = A^T \Leftrightarrow AX^T A = A$$

□

2.5.2 Moore-Penrose 广义逆

Definition 2.16. 设 $A \in M_{m\times n}(\mathbb{C})$。若 $X \in M_{n\times m}(\mathbb{C})$ 满足：

$$\begin{cases} AXA = A \\ XAX = X \\ (AX)^H = AX \\ (XA)^H = XA \end{cases}$$

则称 X 为 A 的 *Moore-Penrose* 广义逆，记作 A^+，上述方程组被称为 A 的 *Penrose* 方程组。

满秩分解导出的广义逆

Theorem 2.15. 设 $A \in M_{m\times n}(\mathbb{C})$，则 A 的 *Penrose* 方程组一定有唯一解。对 A 进行满秩分解，设 $A = BC$，其中 B, C 分别为列满秩矩阵与行满秩矩阵，则 A 的 *Penrose* 方程组的解可表示为：

$$X = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$$

Proof. 由定理 2.2可知 $(B^HB)^{-1}, (CC^H)^{-1}$ 存在，将上述 X 代入 A 的 Penrose 方程组可得：

$$\begin{aligned}
XAX &= C^H(CC^H)^{-1}(B^HB)^{-1}B^HBCC^H(CC^H)^{-1}(B^HB)^{-1}B^H \\
&= C^H(CC^H)^{-1}(B^HB)^{-1}B^H = X \\
AXA &= BCC^H(CC^H)^{-1}(B^HB)^{-1}B^HBC = BC = A \\
(AX)^H &= X^HA^H = B[(B^HB)^{-1}]^H[(CC^H)^{-1}]^HCC^HB^H \\
&= B[(B^HB)^{-1}]^H[(CC^H)^{-1}]^HCC^HB^H \\
&= B[(B^HB)^H]^{-1}[(CC^H)^H]^{-1}CC^HB^H \\
&= B(B^HB)^{-1}(CC^H)^{-1}CC^HB^H \\
&= B(B^HB)^{-1}B^H \\
&= B(CC^H)(CC^H)^{-1}(B^HB)^{-1}B^H = AX \\
(XA)^H &= A^HX^H = C^HB^HB[(B^HB)^{-1}]^H[(CC^H)^{-1}]^HC \\
&= C^H(CC^H)^{-1}C = C^H(CC^H)^{-1}(B^HB)^{-1}(B^HB)C = XA
\end{aligned}$$

于是 X 与 A 的 Penrose 方程组相容，所以 X 是解。 □

奇异值分解导出的广义逆

Theorem 2.16. 设 $A \in M_{m\times n}(\mathbb{C})$，则有：

$$A^+ = Q\begin{pmatrix} \Lambda^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}P^H$$

其中 P, Q, Λ 为 A 的奇异值分解中相关矩阵。

Proof. 将之代入到 A 的 Penrose 方程组中可得：

$$AQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HA = P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H$$

$$= P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H = A$$

$$Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HAQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H = Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$$

$$= Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$$

$$AQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H = Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HA = I$$

因为 I 是 Hermitian 矩阵，于是 $Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$ 与 A 的 Penrose 方程组相容，所以它是解。 □

Moore-Penrose 广义逆的性质

Property 2.5.2. *设 $A \in M_{m\times n}(\mathbb{C})$，则 A 的 Moore-Penrose 广义逆 A^+ 具有如下性质：*

1. A^+ 是唯一的；
2. $(A^+)^+ = A$；
3. $(A^+)^H = (A^H)^+$；
4. $\operatorname{rank}(A^+) = \operatorname{rank}(A)$；
5. 若 A 是一个 *Hermitian* 矩阵，则：

$$A^+ = P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$$

   其中 Λ 为 A 的非零特征值构成的对角矩阵，P 是一个正交矩阵；

6. 若 α 是一个非零向量，则 $\alpha^+ = \frac{\alpha^H}{||\alpha||^2}$；
7. $I - A^+A \geqslant \mathbf{0}$；
8. $(A^HA)^+ = A^+(A^H)^+$；
9. $A^+ = (A^HA)^+A^H = A^H(AA^H)^+$。

Proof. (1) 设 X_1, X_2 都是 A 的 Penrose 方程组的解，则：

$$\begin{aligned}X_1 &= X_1AX_1 = X_1(AX_2A)X_1 = X_1(AX_2)(AX_1)\\&= X_1(AX_2)^H(AX_1)^H = X_1(AX_1AX_2)^H = X_1X_2^H(AX_1A)^H\\&= X_1X_2^HA^H = X_1(AX_2)^H = X_1AX_2 = X_1(AX_2A)X_2\\&= (X_1A)(X_2A)X_2 = (X_1A)^H(X_2A)^HX_2 = (X_2AX_1A)^HX_2\\&= (X_2A)^HX_2 = X_2AX_2 = X_2\end{aligned}$$

所以 Penrose 方程组的解是唯一的。

(2) 由 Penrose 方程的对称性可直接得到。

(3) 由 A^+ 的奇异值分解表示（定理 2.16）可得：

$$\begin{aligned}(A^+)^H &= \left[Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\right]^H = P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}^H Q^H\\&= P\begin{pmatrix}(\Lambda^{-1})^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H = P\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\end{aligned}$$

将其代入 A^H 的 Penrose 方程组可得：

$$\begin{aligned}A^H(A^+)^HA^H &= Q\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\\&= Q\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H = A^H\\(A^+)^HA^H(A^+)^H &= P\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\\&= P\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H = (A^+)^H\\[A^H(A^+)^H]^H &= [(A^+)^HA^H]^H = A^+A = I\end{aligned}$$

因为 I 是 Hermitian 矩阵，于是 $(A^+)^H$ 与 A^H 的 Penrose 方程组相容，所以 $(A^+)^H = (A^H)^+$。

(4) 由 A^+ 的奇异值分解表示（定理 2.16）显然可得 $\mathrm{rank}(A^+) = \mathrm{rank}(\Lambda)$，而 $\mathrm{rank}(\Lambda) = \mathrm{rank}(A)$，所以有 $\mathrm{rank}(A^+) = \mathrm{rank}(A)$。

(5) 因为 A 是一个 Hermitian 矩阵，由性质 2.6.2(3) 可知存在正交矩阵 P 使得：

$$A = P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$$

将 $P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$ 代入 A 的 Penrose 方程组可得：

$$\begin{aligned}AP\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HA &= P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\\ &= P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H = A\end{aligned}$$

$$\begin{aligned}P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HAP\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H &= P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\\ &= P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\end{aligned}$$

$$\left[AP\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\right]^H = \left[P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HA\right]^H = P\begin{pmatrix}I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$$

因为 $P\begin{pmatrix}I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$ 是 Hermitian 矩阵，于是 $P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H$ 与 A 的 Penrose 方程组相容，所以 $P\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H = A^+$。

(6) 将 $\frac{\alpha^H}{||\alpha||^2}$ 代入 α 的 Penrose 方程组可得：

$$\begin{aligned}\alpha\frac{\alpha^H}{||\alpha||^2}\alpha &= \alpha\\ \frac{\alpha^H}{||\alpha||^2}\alpha\frac{\alpha^H}{||\alpha||^2} &= \frac{\alpha^H}{||\alpha||^2}\\ \left(\alpha\frac{\alpha^H}{||\alpha||^2}\right)^H = \left(\frac{\alpha^H}{||\alpha||^2}\alpha\right)^H &= 1\end{aligned}$$

显然 $\frac{\alpha^H}{||\alpha||^2} = \alpha^+$。

(7) 由 A^+ 的奇异值分解表示（定理 2.16）可得：

$$\begin{aligned}I - A^+A = I - Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H &= I - Q\begin{pmatrix}I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\\ &= I - \begin{pmatrix}I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & I\end{pmatrix} \cong \begin{pmatrix}I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\end{aligned}$$

由定理 2.43(3) 的第三条可知 $I - A^+A \geqslant \mathbf{0}$。

(8) 由 (3) 可得：

$$\begin{aligned}A^+(A^H)^+ = A^+(A^+)^H &= Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ^H\\ &= Q\begin{pmatrix}\Lambda^{-1}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H = \begin{pmatrix}\Lambda^{-1}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\end{aligned}$$

由 A 的奇异值分解（定理 2.46）可得：

$$\begin{aligned} A^HA=&\left[P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\right]^H P\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H\\ =&Q\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^H=\begin{pmatrix}\Lambda^H\Lambda & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\end{aligned}$$

将 $A^+(A^H)^+$ 代入 A^HA 的 Penrose 方程组中即可验证得到 $(A^HA)^+=A^+(A^H)^+$。

(9) 由 (8)、(3) 和 A^+ 的奇异值分解表示（定理 2.16）可得：

$$\begin{aligned} (A^HA)^+A^H&=A^+(A^H)^+A^H=A^+(A^+)^HA^H\\ &=Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\\ &=Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H=A^+\\ A^H(AA^H)^+&=A^H(A^H)^+A^+=A^H(A^+)^HA^+\\ &=Q\begin{pmatrix}\Lambda^H & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^HP\begin{pmatrix}(\Lambda^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}Q^HQ\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H\\ &=Q\begin{pmatrix}\Lambda^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}P^H=A^+\end{aligned}$$

□

2.5.3 线性方程组的解

Theorem 2.17. 数域 K 上 n 元非齐次线性方程组 $Ax=\beta$ 有解的充分必要条件为对 A 的任一广义逆 A^- 都有：

$$\beta=AA^-\beta$$

Proof. **(1) 必要性**：若 $Ax=\beta$ 有解，取其一个解 α，于是对 A 的任一广义逆有：

$$\beta=A\alpha=AA^-A\alpha=AA^-\beta$$

(2) 充分性：若此时对 A 的任一广义逆 A^- 有 $\beta=AA^-\beta$，则方程组可化为：

$$Ax=AA^-\beta$$

容易看出 $A^-\beta$ 就是 $Ax=\beta$ 的一个解。 □

齐次方程组解的结构

Theorem 2.18. 若数域 K 上 n 元齐次线性方程组 $Ax=\mathbf{0}$ 有解，则它的通解为：

$$x=(I_n-A^-A)y$$

其中 A^- 是 A 的任意一个给定的广义逆，y 取遍 K^n 中的列向量。

Proof. 任取 $y \in K^n$，有：

$$A(I_n - A^- A)y = Ay - AA^- Ay = Ay - Ay = \mathbf{0}$$

所以对任意的 $y \in K^n$，$(I_n - A^- A)y$ 都是 $Ax = \mathbf{0}$ 的解。

若 η 是 $Ax = \mathbf{0}$ 的一个解，则：

$$(I_n - A^- A)\eta = \eta - A^- A\eta = \eta - A^- \mathbf{0} = \eta$$

所以 $Ax = \mathbf{0}$ 的任意一个解 x 都可以表示为 $(I_n - A^- A)x$ 的形式。

综上，$Ax = \mathbf{0}$ 的通解为 $x = (I_n - A^- A)y$。 □

非齐次方程组解的结构

Theorem 2.19 (结构 1). 若数域 K 上 n 元非齐次线性方程组 $Ax = \beta$ 有解，则它的通解为：

$$x = A^- \beta + (I_n - A^- A)y$$

其中 A^- 是 A 的任意一个给定的广义逆，y 取遍 K^n 中的列向量。

Proof. 由定理 2.17的充分性可知对于给定的这一 A^-，$A^- \beta$ 为 $Ax = \beta$ 的一个特解，而由定理 2.18可知齐次线性方程组 $Ax = \mathbf{0}$ 的通解为 $(I_n - A^- A)y$，由可得 $Ax = \beta$ 的通解为 $x = A^- \beta + (I_n - A^- A)y$。 □

链接线性方程组的关系

Theorem 2.20 (结构 2). 若数域 K 上 n 元非齐次线性方程组 $Ax = \beta$ 有解，则它的通解为：

$$x = A^- \beta$$

A^- 取遍 A 的所有广义逆。

Proof. 由定理 2.17的充分性可知对于任意的 A^-，$A^- \beta$ 都是 $Ax = \beta$ 的解。

对于 $Ax = \beta$ 的任意一个解 y，由定理 2.19可知存在 A 的一个广义逆 G 和 K^n 上的一个列向量 z，使得：

$$y = G\beta + (I_n - GA)z$$

因为 $\beta \neq \mathbf{0}$，所以 $\beta^H \beta \neq 0$，于是存在数域 K 上的矩阵 $B = z(\beta^H \beta)^{-1} \beta^H$ 使得 $B\beta = z$，于是：

$$y = G\beta + (I_n - GA)B\beta = [G + (I_n - GA)B]\beta$$

因为：

$$\begin{aligned} A[G + (I_n - GA)B]A &= AGA + A(I_n - GA)BA \\ &= A + ABA - AGABA \\ &= A + ABA - ABA = A \end{aligned}$$

所以 $G + (I_n - GA)B$ 是 A 的一个广义逆，即 $Ax = \beta$ 的任一解可以表示为 $A^- \beta$。 □

Theorem 2.21. 在数域 K 上相容线性方程组 $Ax=\beta$ 的解集中，$x_0=A^+\beta$ 为长度最小者。

Proof. 由定理 2.19可知，$Ax=\beta$ 的通解可以表示为：

$$x=A^+\beta+(I-A^+A)y$$

于是：

$$\begin{aligned}||x|| &= [A^+\beta+(I-A^+A)y]^H[A^+\beta+(I-A^+A)y]\\ &= ||x_0||+\beta^H(A^+)^H(I-A^+A)y\\ &\quad +y^H(I-A^+A)^HA^+\beta+y^H(I-A^+A)^H(I-A^+A)y\\ &= ||x_0||+2\beta^H(A^+)^H(I-A^+A)y+||(I-A^+A)y||\end{aligned}$$

由性质 2.5.2(9) 可得：

$$\begin{aligned}(A^+)^H(I-A^+A) &= (A^+)^H-(A^+)^HA^+A=(A^H)^+-(A^H)^+A^+A\\ &= (A^H)^+-[A(A^H)]^+A=\mathbf{0}\end{aligned}$$

于是有 $2\beta^H(A^+)^H(I-A^+A)y=0$。因为 $||(I-A^+A)y||\geqslant 0$，等号成立当且仅当 $(I-A^+A)y=\mathbf{0}$，所以 $x=A^+\beta=x_0$ 时长度最小。 □

2.6 相似的应用

2.6.1 特征值与特征向量

Definition 2.17. $A\in M_n(K)$。如果 K^n 中存在非零列向量 α，使得：

$$A\alpha=\lambda\alpha,\ \lambda\in K$$

则称 λ 是 A 的一个特征值 *(eigenvalue)*，α 是 A 属于特征值 λ 的一个特征向量 *(eigenvector)*。

求解特征值与特征向量

Definition 2.18. $A\in M_n(K)$，称 $|\lambda I-A|$ 为 A 的特征多项式 *(characteristic polynomial)*。

Theorem 2.22. $A\in M_n(K)$，则：

1. λ 是 A 的一个特征值当且仅当 λ 是 A 的特征多项式在数域 K 中的一个根；
2. α 是 A 属于特征值 λ 的一个特征向量当且仅当 α 是齐次线性方程组 $(\lambda I-A)x=\mathbf{0}$ 的一个非零解。

Proof. 显然：

$$
\begin{aligned}
&\lambda\text{是}A\text{的一个特征值，}\alpha\text{是}A\text{属于}\lambda\text{的一个特征向量}\\
\Leftrightarrow & A\alpha=\lambda\alpha,\ \alpha\neq\mathbf{0},\ \lambda\in K\\
\Leftrightarrow & (\lambda I-A)\alpha=\mathbf{0},\ a\neq\mathbf{0},\ \lambda\in K\\
\Leftrightarrow & \alpha\text{是齐次线性方程组}(\lambda I-A)x=\mathbf{0}\text{的一个非零解，}\lambda\in K\\
\Leftrightarrow & |\lambda I-A|=0,\ \alpha\text{是齐次线性方程组}(\lambda I-A)x=\mathbf{0}\text{的一个非零解，}\lambda\in K\\
\Leftrightarrow & \lambda\text{是多项式}|\lambda I-A|\text{在}K\text{中的一个根，}\\
& \alpha\text{是齐次线性方程组}(\lambda I-A)x=\mathbf{0}\text{的一个非零解，}\lambda\in K
\end{aligned}
$$

□

特征向量的性质

Property 2.6.1. $A\in M_n(K)$，其特征向量具有如下性质：

1. 设 λ 是 A 的一个特征值，则 A 属于 λ 的所有特征向量构成 K^n 的一个子空间。因此，把齐次线性方程组 $(\lambda I-A)x=\mathbf{0}$ 的解空间称为 A 属于 λ 的特征子空间 *(eigenspace)*；

2. A 的属于不同特征值的特征向量是线性无关的。

Proof. (1) 任取 $k_1,k_2\in K$ 和 A 属于特征值 λ 的两个特征向量 α,β，则

$$A(k_1\alpha+k_2\beta)=k_1A\alpha+k_2A\beta=k_1\lambda\alpha+k_2\lambda\beta=\lambda(k_1\alpha+k_2\beta)$$

于是 $k_1\alpha+k_2\beta$ 也是 A 属于特征值 λ 的特征向量。由定理 1.16可知 A 属于 λ 的所有特征向量构成 K^n 的一个子空间。

(2) 我们来证明：设 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 是 $A\in M_n(K)$ 的不同的特征值，$a_{j1},a_{j2},\ldots,a_{jr_j}$ 是 A 属于 λ_j 的线性无关的特征向量，$j=1,2,\ldots,m$，则向量组：

$$a_{11},a_{12},\ldots,a_{1r_1},a_{21},a_{22},\ldots,a_{2r_2},a_{m1},a_{m2},\ldots,a_{mr_m}$$

线性无关。

1. 证明对 $n=2$ 成立：对于 λ_1 和 λ_2 的线性无关的特征向量 $a_{11},a_{12},\ldots,a_{1r_1}$ 和 $a_{21},a_{22},\ldots,a_{2r_2}$，设：

$$k_1a_{11}+k_2a_{12}+\cdots+k_{r_1}a_{1r_1}+l_1a_{21}+l_2a_{22}+\cdots+l_{r_2}a_{2r_2}=\mathbf{0}$$

两边同乘 A 可得：

$$
\begin{aligned}
k_1Aa_{11}+k_2Aa_{12}+\cdots+k_{r_1}Aa_{1r_1}+l_1Aa_{21}+l_2Aa_{22}+\cdots+l_{r_2}Aa_{2r_2}=\mathbf{0}\\
k_1\lambda_1a_{11}+k_2\lambda_1a_{12}+\cdots+k_{r_1}\lambda_1a_{1r_1}+l_1\lambda_2a_{21}+l_2\lambda_2a_{22}+\cdots+l_{r_2}\lambda_2a_{2r_2}=\mathbf{0}
\end{aligned}
$$

因为 $\lambda_1 \neq \lambda_2$，所以 λ_1, λ_2 不全为 0。设 $\lambda_2 \neq 0$，在上上上个式子两端乘以 λ_2（若 $\lambda_2 = 0$，则同乘 λ_1）得：

$$k_1\lambda_2 a_{11} + k_2\lambda_2 a_{12} + \cdots + k_{r_1}\lambda_2 a_{1r_1} + l_1\lambda_2 a_{21} + l_2\lambda_2 a_{22} + \cdots + l_{r_2}\lambda_2 a_{2r_2} = \mathbf{0}$$

于是：

$$k_1(\lambda_1 - \lambda_2)a_{11} + k_2(\lambda_1 - \lambda_2)a_{12} + \cdots + k_{r_1}(\lambda_1 - \lambda_2)a_{1r_1} = \mathbf{0}$$

因为 $\lambda_1 \neq \lambda_2$，所以：

$$k_1 a_{11} + k_2 a_{12} + \cdots + k_{r_1} a_{1r_1} = \mathbf{0}$$

因为 $a_{11}, a_{12}, \ldots, a_{1r_1}$ 线性无关，所以 $k_1 = k_2 = \cdots = k_{r_1} = 0$，从而：

$$l_1 a_{21} + l_2 a_{22} + \cdots + l_{r_2} a_{2r_2} = \mathbf{0}$$

因为 $a_{21}, a_{22}, \ldots, a_{2r_2}$ 线性无关，所以 $l_1 = l_2 = \cdots = l_{r_2} = 0$。

综上，向量组 $a_{11}, a_{12}, \ldots, a_{1r_1}, a_{21}, a_{22}, \ldots, a_{2r_2}$ 线性无关。

2. 归纳假设：假设对 n 个不同的特征值都有上述结论（即 n 个不同特征值的线性无关的特征向量构成的向量组线性无关），下面来证明对 $n+1$ 个不同的特征值也成立。

设：

$$k_{11}a_{11} + k_{12}a_{12} + \cdots k_{1r_1}a_{1r_1} + \cdots + k_{nr_n}a_{nr_n} + l_1 a_{(n+1)1} + l_2 a_{(n+1)2} + \cdots + l_{r_{n+1}} a_{(n+1)r_{n+1}} = \mathbf{0}$$

两边同乘 A 可得：

$$\begin{aligned}
&k_{11}Aa_{11} + k_{12}Aa_{12} + \cdots k_{1r_1}Aa_{1r_1} + \cdots + k_{nr_n}Aa_{nr_n}\\
&+l_1 Aa_{(n+1)1} + l_2 Aa_{(n+1)2} + \cdots + l_{r_{n+1}}Aa_{(n+1)r_{n+1}} = \mathbf{0}\\
&k_{11}\lambda_1 a_{11} + k_{12}\lambda_1 a_{12} + \cdots k_{1r_1}\lambda_1 a_{1r_1} + \cdots + k_{nr_n}\lambda_n a_{nr_n}\\
&+l_1\lambda_{n+1}a_{(n+1)1} + l_2\lambda_{n+1}a_{(n+1)2} + \cdots + l_{r_{n+1}}\lambda_{n+1}a_{(n+1)r_{n+1}} = \mathbf{0}
\end{aligned}$$

2.1.$\lambda_{n+1} \neq 0$：若 $\lambda_{n+1} \neq 0$，则在上上上式两边同乘 λ_{n+1} 可得：

$$\begin{aligned}
&k_{11}\lambda_{n+1}a_{11} + k_{12}\lambda_{n+1}a_{12} + \cdots k_{1r_1}\lambda_{n+1}a_{1r_1} + \cdots + k_{nr_n}\lambda_{n+1}a_{nr_n}\\
&+l_1\lambda_{n+1}a_{(n+1)1} + l_2\lambda_{n+1}a_{(n+1)2} + \cdots + l_{r_{n+1}}\lambda_{n+1}a_{(n+1)r_{n+1}} = \mathbf{0}
\end{aligned}$$

于是有：

$$k_{11}(\lambda_{n+1}-\lambda_1)a_{11} + k_{12}(\lambda_{n+1}-\lambda_1)a_{12} + \cdots k_{1r_1}(\lambda_{n+1}-\lambda_1)a_{1r_1} + \cdots + k_{nr_n}(\lambda_{n+1}-\lambda_n)a_{nr_n} = \mathbf{0}$$

由归纳假定 $a_{11}, a_{12}, \ldots, a_{1r_1}, \ldots, a_{nr_n}$ 线性无关，所以

$$k_{11}(\lambda_{n+1} - \lambda_1) = k_{12}(\lambda_{n+1} - \lambda_1) = \cdots = k_{1r_1}(\lambda_{n+1} - \lambda_1) = \cdots = k_{nr_n}(\lambda_{n+1} - \lambda_n) = 0$$

因为 λ_i，$i = 1, 2, \ldots, n$ 之间互不相同，所以 $\lambda_{n+1} - \lambda_1, \lambda_{n+1} - \lambda_2, \ldots, \lambda_{n+1} - \lambda_n$ 不为 0，于是 $k_{11} = k_{12} = \cdots = k_{1r_1} = \cdots = k_{nr_n} = 0$，所以：

$$l_1 a_{(n+1)1} + l_2 a_{(n+1)2} + \cdots + l_{r_{n+1}} a_{(n+1)r_{n+1}} = \mathbf{0}$$

因为 $a_{(n+1)1}, a_{(n+1)2}, \dots, a_{(n+1)r_{n+1}}$ 线性无关，所以有 $l_1 = l_2 = \dots = l_{r_{n+1}} = 0$。

综上 $a_{11}, a_{12}, \dots, a_{1r_1}, \dots, a_{nr_n}, a_{(n+1)1}, a_{(n+1)2}, \dots, a_{(n+1)r_{n+1}}$ 线性无关。

2.2.$\lambda_{n+1} = 0$：若 $\lambda_{n+1} = 0$，则此时有：

$$\begin{aligned}
&k_{11}\lambda_1 a_{11} + k_{12}\lambda_1 a_{12} + \cdots k_{1r_1}\lambda_1 a_{1r_1} + \cdots + k_{nr_n}\lambda_n a_{nr_n} \\
&+ l_1\lambda_{n+1}a_{(n+1)1} + l_2\lambda_{n+1}a_{(n+1)2} + \cdots + l_{r_{n+1}}\lambda_{n+1}a_{(n+1)r_{n+1}} \\
&= k_{11}\lambda_1 a_{11} + k_{12}\lambda_1 a_{12} + \cdots k_{1r_1}\lambda_1 a_{1r_1} + \cdots + k_{nr_n}\lambda_n a_{nr_n} = \mathbf{0}
\end{aligned}$$

由归纳假定 $a_{11}, a_{12}, \dots, a_{1r_1}, \dots, a_{nr_n}$ 线性无关，所以 $k_{11}\lambda_1 = k_{12}\lambda_1 = \cdots = k_{1r_1}\lambda_1 = \cdots = k_{nr_n}\lambda_n = 0$。因为 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都不是 0（$\lambda_i,\ i = 1, 2, \dots, n+1$ 互不相同，已经有 $\lambda_{n+1} = 0$ 了），所以 $k_{11} = k_{12} = \cdots = k_{1r_1} = \cdots = k_{nr_n} = 0$，于是有：

$$l_1 a_{(n+1)1} + l_2 a_{(n+1)2} + \cdots + l_{r_{n+1}} a_{(n+1)r_{n+1}} = \mathbf{0}$$

因为 $a_{(n+1)1}, a_{(n+1)2}, \dots, a_{(n+1)r_{n+1}}$ 线性无关，所以有 $l_1 = l_2 = \dots = l_{r_{n+1}} = 0$。

综上，$a_{11}, a_{12}, \dots, a_{1r_1}, \dots, a_{nr_n}, a_{(n+1)1}, a_{(n+1)2}, \dots, a_{(n+1)r_{n+1}}$ 线性无关。

假设存在属于不同特征值的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关，则有：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = \mathbf{0}$$

其中 $k_1, k_2, \dots, k_m$ 不全为 0。注意到 $\alpha_i,\ i = 1, 2, \dots, m$ 可由其对应特征值的特征子空间中的一组基线性表出，于是有：

$$\alpha_i = \sum_{n=1}^{r_i} l_n \beta_{in}$$

其中 $\beta_{in},\ n = 1, 2, \dots, r_i$ 为 α_i 对应特征值的特征子空间的一组基，所以：

$$\sum_{i=1}^{m} k_i \sum_{n=1}^{r_i} l_n \beta_{in} = \sum_{i=1}^{m}\sum_{n=1}^{r_i} k_i l_n \beta_{in} = \mathbf{0}$$

而 $\beta_{in},\ i = 1, 2, \dots, m,\ n = 1, 2, \dots, r_i$ 是线性无关的，所以：

$$k_i l_n = 0,\ \forall\, i = 1, 2, \dots, m,\ n = 1, 2, \dots, r_i$$

因为 $k_1, k_2, \dots, k_m$ 不全为 0，所以存在一组 l_n 全为 0，于是 α_i 中存在零向量，而特征向量不是零向量，矛盾。 □

Theorem 2.23. 相似的矩阵有相同的特征多项式，进而有相同的特征值（包括重数相同）。

Proof. 设 $A, B \in M_n(K)$ 且 A 与 B 相似，于是存在可逆矩阵 $P \in M_n(K)$ 使得 $P^{-1}AP = B$，就有：

$$\begin{aligned}
|\lambda I - B| &= |\lambda I - P^{-1}AP| = |P^{-1}\lambda I P - P^{-1}AP| \\
&= |P^{-1}(\lambda I - A)P| = |P^{-1}|\,|\lambda I - A|\,|P| = |\lambda I - A|
\end{aligned}$$ □

几何重数与代数重数

Definition 2.19. $A \in M_n(K)$，λ 是 A 的一个特征值。把 A 属于 λ 的特征子空间的维数叫作 λ 的几何重数 *(geometric multiplicity)*，把 λ 作为 A 的特征多项式的根的重数叫作 λ 的代数重数 *(algebraic multiplicity)*。

Theorem 2.24. $A \in M_n(K)$，λ_1 是 A 的一个特征值，则 λ_1 的几何重数不超过它的代数重数。

Proof. 设 A 属于特征值 λ_1 的特征子空间 W_1 的维数为 r。在 W_1 中取一个基 $\alpha_1, \alpha_2, \ldots, \alpha_r$，把它扩充为 K^n 的一组基 $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_{n-r}$。令：

$$P = (\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_{n-r})$$

则 P 是数域 K 上的 n 阶可逆矩阵，并且有：

$$\begin{aligned}
P^{-1}AP &= P^{-1}(A\alpha_1, A\alpha_2, \ldots, A\alpha_r, A\beta_1, A\beta_2, \ldots, A\beta_{n-r}) \\
&= P^{-1}(\lambda_1\alpha_1, \lambda_1\alpha_2, \ldots, \lambda_1\alpha_r, A\beta_1, A\beta_2, \ldots, A\beta_{n-r}) \\
&= (\lambda_1\varepsilon_1, \lambda_1\varepsilon_2, \ldots, \lambda_1\varepsilon_r, P^{-1}A\beta_1, P^{-1}A\beta_2, \ldots, P^{-1}A\beta_{n-r}) \\
&= \begin{pmatrix} \lambda_1 I_r & B \\ \mathbf{0} & C \end{pmatrix}
\end{aligned}$$

由定理 2.23可得：

$$\begin{aligned}
|\lambda I - A| &= \begin{vmatrix} \lambda I_r - \lambda_1 I_r & -B \\ \mathbf{0} & \lambda I_{n-r} - C \end{vmatrix} \\
&= |\lambda I_r - \lambda_1 I_r|\, |\lambda I_{n-r} - C| \\
&= (\lambda - \lambda_1)^r |\lambda I_{n-r} - C|
\end{aligned}$$

即 λ_1 的几何重数小于或等于 r，也即 λ_1 的几何重数小于或等于它的代数重数。 □

2.6.2 矩阵的对角化

Definition 2.20. 如果 n 阶矩阵 A 能够相似于一个对角矩阵,那么称 A 可对角化 *(diagonalizable)*。

研究矩阵是否可对角化是为了计算矩阵的幂，因为对角矩阵的幂是很好计算的。

Theorem 2.25 (矩阵可对角化的第一个充分必要条件)**.** $A \in M_n(K)$ 可对角化的充分必要条件为：A 有 n 个线性无关的特征向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$，此时令 $P = (\alpha_1, \alpha_2, \ldots, \alpha_n)$，则：

$$P^{-1}AP = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$

其中 λ_i 是 α_i 所属的特征值，$i = 1, 2, \ldots, n$。上述对角矩阵称为 A 的**相似标准形**，除了主对角线上元素的排列次序外，A 的相似标准形是唯一的。

Proof. 显然：

$$
\begin{aligned}
& A\text{与对角矩阵}D = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}\text{相似，其中}\lambda_i \in K,\ i = 1, 2, \ldots, n \\
\Leftrightarrow & \text{如果存在可逆矩阵 } P \in M_n(K)\text{，使得}P^{-1}AP = D \\
& \text{即}AP = PD \\
& \text{即}A(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)D \\
& \text{即}(A\alpha_1, A\alpha_2, \ldots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \ldots, \lambda_n\alpha_n) \\
\Leftrightarrow & K^n\text{中有 } n \text{ 个线性无关的列向量}\alpha_1, \alpha_2, \ldots, \alpha_n\text{使得}A\alpha_i = \lambda_i\alpha_i,\ i = 1, 2, \ldots, n
\end{aligned}
$$

□

Theorem 2.26 (矩阵可对角化的第二个充分必要条件)**.** *$A \in M_n(K)$ 可对角化的充分必要条件是：A 的属于不同特征值的特征子空间的维数之和等于 n。*

Proof. **(1) 充分性：** 由性质 2.6.1(2) 和定理 2.25的充分性可直接得出。

(2) 必要性： 设 A 的所有不同的特征值是 $\lambda_1, \lambda_2, \ldots, \lambda_m$，它们的几何重数分别为 $r_1, r_2, \ldots, r_m$。若此时 A 的属于不同特征值的特征子空间的维数之和不等于 n，由定理 2.24可知此时 $r_1 + r_2 + \cdots + r_m < n$，那么 A 没有 n 个线性无关的特征向量，由定理 2.25的必要性可得 A 不可以对角化。 □

Corollary 2.1. *$A \in M_n(K)$ 如果有 n 个不同的特征值，那么 A 可对角化。*

Theorem 2.27 (矩阵可对角化的第三个充分必要条件)**.** *$A \in M_n(K)$ 可对角化的充分必要条件是：A 的特征多项式的全部复根都属于 K，且 A 的每个特征值的几何重数等于它的代数重数。*

Proof. **(1) 充分性：** 由定理 2.26的充分性可直接得到。

(2) 必要性： 因为 A 可对角化，由可对角化的定义可知 A 相似于：

$$\mathrm{diag}(\lambda_1, \cdots, \lambda_1, \ldots, \lambda_m, \ldots, \lambda_m) \in M_n(K)$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_m$ 是 A 的全部不同的特征值，每个特征值重复的次数为对应特征子空间的维数，λ_i 对应特征子空间的维数记为 r_i，$i = 1, 2, \ldots, m$。因为相似的矩阵具有相同的特征多项式，所以：

$$|\lambda I - A| = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_m)^{r_m}$$

于是 A 的特征多项式的根为 $\lambda_1, \lambda_2, \ldots, \lambda_m$。因为 $\mathrm{diag}(\lambda_1, \cdots, \lambda_1, \ldots, \lambda_m, \ldots, \lambda_m) \in M_n(K)$，所以 $\lambda_1, \lambda_2, \ldots, \lambda_m \in K$，于是 A 的特征多项式的全部根都属于 K 且每一个特征值的代数重数等于它的几何重数。 □

2.6.3 Hermitian 矩阵的对角化

Definition 2.21. *若对于 $A, B \in M_n(\mathbb{C})$，存在一个 n 阶正交矩阵 Q，使得 $Q^{-1}AQ = B$，则称 A* **正交相似**于 *B。*

Theorem 2.28. 正交相似是 $M_n(\mathbb{C})$ 上的一个等价关系。

Theorem 2.29. $A \in M_n(\mathbb{C})$。若 A 正交相似与一个对角矩阵 D，则 A 一定是 *Hermitian* 矩阵。

Proof. 因为 A 正交相似于 D，所以存在正交矩阵 Q 使得 $Q^{-1}AQ = D$，即 $A = QDQ^{-1}$，于是有：

$$A^H = (QDQ^{-1})^H = (Q^{-1})^H D^H Q^H = (Q^H)^H DQ^{-1} = QDQ^{-1} = A$$

所以 A 是一个 Hermitian 矩阵。 □

Corollary 2.2. 正交相似一定相似，相似不一定正交相似。

Proof. 设非 Hermitian 矩阵 A 相似于一个对角矩阵 D，若 A 正交相似于 D，则 A 得是一个 Hermitian 矩阵，而 A 不是一个 Hermitian 矩阵。 □

Property 2.6.2. 设 *Hermitian* 矩阵 $A, B \in M_n(\mathbb{C})$，则：

1. A 的特征多项式的每一个根都是实数，从而都是 A 的特征值；
2. A 属于不同特征值的特征向量是正交的；
3. A 一定正交相似于由它的特征值构成的对角矩阵；
4. A 与 B 正交相似的充分必要条件为 A 与 B 相似。

Proof. (1) 设 λ 是 A 的特征多项式的任意一个根，将 A 看作是复数域 $\mathbb{C}$ 上的矩阵，取 A 属于特征值 λ 的一个特征向量 α，考虑 $\mathbb{C}^n$ 中的内积，有：

$$(A\alpha, \alpha) = (\lambda\alpha, \alpha) = \lambda(\alpha, \alpha)$$
$$(\alpha, A\alpha) = (\alpha, \lambda\alpha) = \overline{\lambda}(\alpha, \alpha)$$
$$(A\alpha, \alpha) = (A\alpha)^H\alpha = \alpha^H A^H \alpha = \alpha^H A\alpha = (\alpha, A\alpha)$$

所以 $\lambda(\alpha, \alpha) = \overline{\lambda}(\alpha, \alpha)$。因为 α 是特征向量，所以 $\alpha \neq \mathbf{0}$，于是 $\lambda = \overline{\lambda}$，因此 λ 是一个实数。由 λ 的任意性，结论成立。

(2) 设 λ_1, λ_2 是 A 的不同的特征值（由 (1) 得它们都是实数），α_1, α_2 分别是 A 属于 λ_1, λ_2 的一个特征向量，考虑 $\mathbb{C}^n$ 上的标准内积：

$$\begin{aligned}\lambda_1(\alpha_1, \alpha_2) &= (\lambda_1\alpha_1, \alpha_2) = (A\alpha_1, \alpha_2) = A(\alpha_1, \alpha_2) = (\alpha_1, A^H\alpha_2) \\ &= (\alpha_1, A\alpha_2) = (\alpha_1, \lambda_2\alpha_2) = \overline{\lambda_2}(\alpha_1, \alpha_2) = \lambda_2(\alpha_1, \alpha_2)\end{aligned}$$

于是有 $(\lambda_1 - \lambda_2)(\alpha_1, \alpha_2) = 0$。因为 $\lambda_1 \neq \lambda_2$，所以 $(\alpha_1, \alpha_2) = 0$。

(3) 对 n 作数学归纳法。

当 $n = 1$ 时，$(1)^{-1}A(1) = A$，结论成立。

假设对于 $n-1$ 阶的实对称矩阵都成立，考虑 n 阶实对称矩阵 A。

由 (2) 可知 A 必有特征值，取 A 的一个特征值 λ_1 和 A 属于 λ_1 的一个特征向量 η_1，满足 $||\eta_1||=1$。把 η_1 扩充为 $\mathbb{C}^n$ 的一个基并进行 Schimidt 正交化和单位化，可得到 $\mathbb{C}^n$ 的一个标准正交基 $\eta_1,\eta_2,\dots,\eta_n$。令：

$$Q_1=(\eta_1,\eta_2,\dots,\eta_n)$$

显然 Q_1 是一个正交矩阵，于是有 $Q_1^{-1}Q_1=(Q_1^{-1}\eta_1,Q_1^{-1}\eta_2,\dots,Q_1^{-1}\eta_n)=(e_1,e_2,\dots,e_n)$。注意到：

$$Q_1^{-1}AQ_1=Q_1^{-1}(A\eta_1,A\eta_2,\dots,A\eta_n)=(Q_1^{-1}\lambda\eta_1,Q_1^{-1}A\eta_2,\dots,Q_1^{-1}A\eta_n)=\begin{pmatrix}\lambda_1 & \alpha\\ \mathbf{0} & B\end{pmatrix}$$

因为 $(Q_1^{-1}AQ_1)^H=Q_1^HA^H(Q_1^{-1})^H=Q_1^{-1}A(Q_1^H)^H=Q_1^{-1}AQ_1$，所以 $Q_1^{-1}AQ_1$ 是一个对称阵，于是 $\alpha=\mathbf{0}$，B 是一个 $n-1$ 阶 Hermitian 阵。由归纳假设，存在 $n-1$ 阶正交矩阵 Q_2 使得 $Q_2^{-1}BQ_2=\mathrm{diag}\{\lambda_2,\lambda_3,\dots,\lambda_n\}$。令：

$$Q=Q_1\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2\end{pmatrix}$$

则：

$$Q^HQ=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2^H\end{pmatrix}Q_1^HQ_1\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2\end{pmatrix}=I$$

即 Q 是一个正交矩阵。同时：

$$\begin{aligned}Q^{-1}AQ&=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2^H\end{pmatrix}Q_1^HAQ_1\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2\end{pmatrix}=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2^H\end{pmatrix}\begin{pmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & B\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_2\end{pmatrix}\\&=\begin{pmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & Q_2^HBQ_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & \mathrm{diag}\{\lambda_2,\lambda_3,\dots,\lambda_n\}\end{pmatrix}=\mathrm{diag}\{\lambda_1,\lambda_2,\dots,\lambda_n\}\end{aligned}$$

所以 A 正交相似于对角矩阵 $\mathrm{diag}\{\lambda_1,\lambda_2,\dots,\lambda_n\}$。由 $AQ=Q\,\mathrm{diag}\{\lambda_1,\lambda_2,\dots,\lambda_n\}$ 可以得到 $\lambda_2,\lambda_3,\dots,\lambda_n$ 是 A 的特征值。

综上，结论成立。

(4) **必要性**：正交相似也是相似。

充分性：因为 A 与 B 相似，由定理 2.23可知 A 与 B 有相同的特征值（包括重数）$\lambda_1,\lambda_2,\dots,\lambda_n$。由 (3) 可得 A 与 B 都正交相似于 $\mathrm{diag}\{\lambda_1,\lambda_2,\dots,\lambda_n\}$（考虑 λ_i 的顺序的话只需要更改 Q 中列向量的顺序）。因为正交相似具有对称性与传递性，所以 A 正交相似于 B。 □

求解正交矩阵 Q

Theorem 2.30. 对于 *Hermitian* 阵 $A\in M_n(\mathbb{C})$，求正交矩阵 Q 使得 $Q^{-1}AQ$ 为对角阵的步骤如下：

1. 求出 A 的所有特征值 $\lambda_1, \lambda_2, \ldots, \lambda_m$；
2. 对于每一个特征值 λ_i，求得其特征子空间的一组基 $\alpha_{1i}, \alpha_{2i}, \ldots, \alpha_{r_i i}$，并对它们进行 *Schimidt* 正交化与单位化，得到 $\eta_{1i}, \eta_{2i}, \ldots, \eta_{r_i i}$；
3. 令 $Q = (\eta_{11}, \eta_{21}, \ldots, \eta_{r_1 1}, \ldots, \eta_{r_m m})$，$Q$ 即为所求。

Proof. 由可知 η_{ij}，$i = 1, 2, \ldots, r_j$，$j = 1, 2, \ldots, m$ 是 A 属于 λ_j 的特征值。根据性质 2.6.2(2) 可知：

Schimidt 正交化链接

$$
\begin{aligned}
Q^{-1}AQ &= Q^H(A\eta_{11}, A\eta_{21}, \ldots, A\eta_{r_1 1}, \ldots, A\eta_{r_m m}) \\
&= \begin{pmatrix} \eta_{11}^H \\ \eta_{21}^H \\ \vdots \\ \eta_{r_1 1}^H \\ \vdots \\ \eta_{r_m m}^H \end{pmatrix} (\lambda_1\eta_{11}, \lambda_1\eta_{21}, \ldots, \lambda_1\eta_{r_1 1}, \ldots, \lambda_m\eta_{r_m m}) \\
&= \operatorname{diag}\{\lambda_1\eta_{11}^H\eta_{11}, \lambda_1\eta_{21}^H\eta_{21}, \ldots, \lambda_1\eta_{r_1 1}^H\eta_{r_1 1}, \ldots, \lambda_m\eta_{r_m m}^H\eta_{r_m m}\} \\
&= \operatorname{diag}\{\lambda_1, \ldots, \lambda_1, \ldots, \lambda_m, \ldots, \lambda_m\}
\end{aligned}
$$

□

实对称矩阵特征值的极值性质

Theorem 2.31. 设 $A \in M_n(\mathbb{R})$，A 的特征值从大到小记作 $\lambda_1, \lambda_2, \ldots, \lambda_n$，$\varphi_1, \varphi_2, \ldots, \varphi_n$ 为对应的标准正交化特征向量，则：

$$
\max_{x \neq \mathbf{0}} \frac{x^TAx}{x^Tx} = \lambda_1 = \varphi_1^TA\varphi_1 \quad \min_{x \neq \mathbf{0}} \frac{x^TAx}{x^Tx} = \lambda_n = \varphi_n^TA\varphi_n
$$

Proof. 由性质 2.6.2(3) 可知存在一个正交矩阵 Q 使得 $Q^{-1}AQ = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \Lambda$。对任意的 $x \in \mathbb{R}^n$，因为 Q 为正交矩阵，Q 可逆，所以关于 y 的非齐次线性方程组 $Qy = x$ 有唯一解，于是对于这个存在且唯一的 y，有：

$$
\begin{aligned}
\frac{x^TAx}{x^Tx} = \frac{y^TQ^TAQy}{y^TQ^TQy} = \frac{y^T\Lambda y}{y^Ty} = \frac{\sum\limits_{i=1}^n \lambda_i y_i^2}{\sum\limits_{i=1}^n y_i^2} \leqslant \lambda_1 \frac{\sum\limits_{i=1}^n y_i^2}{\sum\limits_{i=1}^n y_i^2} = \lambda_1 \\
\frac{x^TAx}{x^Tx} = \frac{y^TQ^TAQy}{y^TQ^TQy} = \frac{y^T\Lambda y}{y^Ty} = \frac{\sum\limits_{i=1}^n \lambda_i y_i^2}{\sum\limits_{i=1}^n y_i^2} \geqslant \lambda_n \frac{\sum\limits_{i=1}^n y_i^2}{\sum\limits_{i=1}^n y_i^2} = \lambda_n
\end{aligned}
$$

当 y 为 $(1, 0, 0, \ldots, 0)^T$ 时第一式取等号，当 y 为 $(0, 0, \ldots, 0, 1)^T$ 时第二式取等号，此时 x 分别为 φ_1 和 φ_n。 □

2.7 合同的应用——二次型

Definition 2.22. 数域 K 上的一个 n 元二次型 *(quadratic form)*是系数在 K 中的 n 个变量的二元齐次多项式，它的一般形式为：

$$f(x_1,x_2,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

其中 $a_{ij}=a_{ji}$，$1\leqslant i,j\leqslant n$。矩阵：

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{12} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{nn}\end{pmatrix}$$

被称为二次型 $f(x_1,x_2,\ldots,x_n)$ 的矩阵，它是一个对称矩阵，主对角元依次是 $x_1^2,x_2^2,\ldots,x_n^2$ 的系数，(i,j) 元是 x_ix_j 系数的一半，其中 $i\neq j$。令：

$$x=(x_1,x_2,\ldots,x_n)^T$$

则二次型 $f(x_1,x_2,\ldots,x_n)$ 可写作 x^TAx。

Definition 2.23. 令 $x=(x_1,x_2,\ldots,x_n)^T$，$y=(y_1,y_2,\ldots,y_n)^T$，可逆矩阵 $C\in M_n(K)$，则关系式 $x=Cy$ 称为变量 $x_1,x_2,\ldots,x_n$ 到变量 $y_1,y_2,\ldots,y_n$ 的一个非退化线性变换 *(invertible linear transformation)*。如果 C 是一个正交矩阵，则称变量变换 $x=Cy$ 为一个正交变换 *(orthogonal transformation)*。

Definition 2.24. 对于数域 K 上的两个 n 元二次型 x^TAx 与 y^TAy，如果存在一个非退化线性变换 $x=Cy$，把 x^TAx 变成 y^TBy，那么称二次型 x^TAx 与 y^TBy **等价**，记作 $x^TAx\cong y^TBy$。如果二次型 x^TAx 等价于一个只含平方项的二次型，那么称这个只含平方项的二次型是 x^TAx 的一个**标准形**。

Theorem 2.32. 数域 K 上两个 n 元二次型 x^TAx 与 y^TBy 等价当且仅当 n 阶对称矩阵 A 与 B 合同，于是二次型的等价也是一个等价关系。

Proof. **(1) 充分性**：因为 $A\cong B$，所以存在可逆矩阵 C 使得 $C^TAC=B$。作非退化线性变换 $x=Cy$，可得到 $(Cy)^TA(Cy)=y^TC^TACy=y^TBy$，所以 $x^TAx\cong y^TBy$。

(2) 必要性：因为 $x^TAx\cong y^TBy$，所以存在非退化线性变换 $x=Cy$，C 是一个可逆矩阵，把 x^TAx 变为 y^TBy，即 $(Cy)^TA(Cy)=y^TC^TACy=y^TBy$，所以 $C^TAC=B$，即 $A\cong B$。

因为合同是一个等价关系，显然可得二次型的等价也是一个等价关系。 □

Theorem 2.33. 数域 K 上任一 n 元二次型都等价于一个只含平方项的二次型。

Proof. 当二次型的矩阵是对角矩阵时该二次型只含平方项，由定理 2.10与定理 2.32可立即得出结论。 □

Theorem 2.34. 设 n 元二次型 x^TAx 的矩阵 A 合同于对角矩阵 $D = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}$，即存在可逆矩阵 C 使得 $C^TAC = D$。令 $x = Cy$，则可以得到 x^TAx 的一个标准形：

$$d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2$$

Proof. 将 $x = Cy$ 代入可得：

$$x^TAx = (Cy)^TA(Cy) = y^TC^TACy = y^TDy = \sum_{i=1}^{n} d_iy_i^2$$ □

Theorem 2.35. 数域 K 上 n 元二次型 x^TAx 的任一标准形中，系数不为 0 的平方项个数等于它的矩阵 A 的秩。

Proof. 设 n 元二次型 x^TAx 经过非退化线性变换 $x = Cy$ 化成标准形 $d_1y_1^2 + d_2y_2^2 + \cdots + d_ry_r^2$，其中 $d_1, d_2, \ldots, d_r$ 都不为 0，则：

$$C^TAC = \mathrm{diag}\{d_1, d_2, \ldots, d_r, 0, \ldots, 0\}$$

于是 $\mathrm{diag}\{d_1, d_2, \ldots, d_r, 0, \ldots, 0\}$ 是 A 的一个合同标准形。由定理 2.11可得 $\mathrm{rank}(A) = r$。 □

Definition 2.25. 称二次型 x^TAx 的矩阵 A 的秩为二次型 x^TAx 的秩。

2.7.1 二次型的规范形

实二次型的规范形

Definition 2.26. 实数域上的二次型称为**实二次型**。由定理 *2.12*可知 n 元实二次型 x^TAx 的矩阵 A 合同于一个对角矩阵 $\mathrm{diag}\{1, 1, \ldots, 1, -1, -1, \ldots, -1, 0, 0, \ldots, 0\}$，再由定理 *2.32*可知经过一个适当的非退化线性变换可以将 x^TAx 化作：

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

称此形式为二次型 x^TAx 的**规范形**，其特征为：只含平方项且平方项系数为 $1, -1, 0$，系数为 1 的平方项在最前面，系数为 -1 的平方项在中间，系数为 0 的平方项在最后。实二次型 x^TAx 的规范形被两个自然数 p 和 r 决定。

Theorem 2.36 (Sylvester's Law of Inertia)**.** n 元实二次型 x^TAx 的规范形是唯一的。

Proof. 设 n 元实二次型 x^TAx 的秩为 r，假设 x^TAx 分别经过非退化线性变换 $x = Cy$ 和 $x = Bz$ 变成两个规范形：

$$x^TAx = y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - y_{p+2}^2 - \cdots - y_r^2$$
$$x^TAx = z_1^2 + z_2^2 + \cdots + z_q^2 - z_{q+1}^2 - z_{q+2}^2 - \cdots - z_r^2$$

要证规范形唯一，即证 $p=q$。

由 $x=Cy$ 和 $x=Bz$ 可知，经过非退化线性变换 $z=(B^{-1}C)y$ 后有：

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

记 $D=B^{-1}C=(d_{ij})$。假设 $p>q$，我们想找到变量 $y_1,y_2,\ldots,y_n$ 的一组取值，使得上式右端大于 0，而左端小于或等于 0，从而产生矛盾。令：

$$y=(y_1,y_2,\ldots,y_p,0,0,\ldots,0)^T$$

其中 $y_1,y_2,\ldots,y_p$ 是待定的实数，使得变量 $z_1,z_2,\ldots,z_q$ 的值全为 0。因为 $z=Dy$，所以：

$$\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1p} \\ d_{21} & d_{22} & \cdots & d_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ d_{q1} & d_{q2} & \cdots & d_{qp} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_q \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

因为 $p>q$,所以上述齐次线性方程组有非零解,即存在非零向量 $y=(y_1,y_2,\ldots,y_p,0,0,\ldots,0)^T$ 使得 $z_1=z_2=\cdots=z_q=0$。此时有：

$$\begin{aligned} z_1^2+z_2^2+\cdots+z_q^2-z_{q+1}^2-z_{q+2}^2-\cdots-z_r^2 &\leqslant 0 \\ y_1^2+y_2^2+\cdots+y_p^2-y_{p+1}^2-y_{p+2}^2-\cdots-y_r^2 &> 0 \end{aligned}$$

矛盾。因此 $p\leqslant q$。同理可得 $q\leqslant p$，于是 $p=q$，规范形唯一。 □

Definition 2.27. 在实二次型 x^TAx 的规范形中，系数为 1 的平方项个数 p 称为 x^TAx 的正惯性指数，系数为 -1 的平方项个数 $r-p$ 称为 x^TAx 的负惯性指数，正惯性指数减去负惯性指数所得的差 $2p-r$ 称为 x^TAx 称为 x^TAx 的符号差 *(signature)*。

Theorem 2.37. 两个 n 元实二次型等价

$\Leftrightarrow$ 它们的规范形相同

$\Leftrightarrow$ 它们的秩相等，并且正惯性指数也相等。

Proof. 第一条由定理 2.36以及二次型等价的传递性、对称性可直接得到（必要性的证明中需要考虑规范形的定义，然后使用定理 2.36），第二条是显然的。 □

显然矩阵 A 的正惯性指数与负惯性指数就等于二次型 x^TAx 的正惯性指数与负惯性指数，也等于 A 的合同标准形主对角线上大于 0 的元素的个数与小于 0 的个数。

Theorem 2.38. 两个 n 阶实对称矩阵合同 $\Leftrightarrow$ 它们的秩相等，并且正惯性指数也相等。

Proof. 由定理 2.32可得矩阵合同等价于各自对应的二次型等价，再由定理 2.37可得两个二次型的秩与正惯性指数都相等。因为矩阵的秩与正惯性指数等于对应的二次型的秩与正惯性指数，所以结论成立。 □

复二次型的规范形

Definition 2.28. 复数域上的二次型称为**复二次型**。由定理 2.13可知 n 元复二次型 x^TAx 的矩阵 A 合同于一个对角矩阵 $\mathrm{diag}\{1,1,\ldots,1,0,0,\ldots,0\}$，再由定理 2.32可知经过一个适当的非退化线性变换可以将 x^TAx 化作：

$$z_1^2+z_2^2+\cdots+z_r^2$$

称此形式为二次型 x^TAx 的**规范形**，其特征为：只含平方项且平方项系数为 $1,0$，系数为 1 的平方项在前面，系数为 0 的平方项在后面。

Theorem 2.39. 复二次型 x^TAx 的规范形是唯一的。

Proof. 复二次型 x^TAx 的规范形完全由它的秩 r 所决定。 □

Theorem 2.40. 两个 n 元复二次型等价

$\Leftrightarrow$ 它们的规范形相同

$\Leftrightarrow$ 它们的秩相等。

Proof. 第一条由定理 2.39以及二次型的传递性、对称性可直接得到（必要性的证明中需要考虑规范形的定义，然后使用定理 2.39），第二条是显然的。 □

2.7.2 正定二次型与正定矩阵

Definition 2.29. 如果对 $\mathbb{R}^n$ 中任意非零列向量 α，都有 $\alpha^TA\alpha>0$，则称 n 元实二次型 x^TAx 是正定 *(positive definite)*的。

Definition 2.30. 若实二次型 x^TAx 是正定的，则称实对称矩阵 A 是正定的，并称 A 为正定矩阵 *(positive definite matrix)*，记为 $A>0$。

Theorem 2.41. n 元实二次型 x^TAx 是正定的当且仅当它的正惯性指数等于 n。

Proof. **(1) 必要性**：设 x^TAx 是正定的，作非退化线性变换 $x=Cy$ 化成规范形：

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

如果 $p<n$，则 y_n^2 的系数为 0 或 -1，取 $y=(0,0,\ldots,1)^T$，则有 $y^TC^TACy=-y_n^2$ 为 0 或 -1，取 $\alpha=Cy$ 即有 $\alpha^TA\alpha$ 为 0 或 -1，与二次型 x^TAx 的正定性矛盾，所以 $p=n$。

(2) 充分性：设 x^TAx 的正惯性指数等于 n，则可以作一个非退化线性变换 $x=Cy$ 将该二次型化作规范形：

$$y^TC^TACy=y_1^2+y_2^2+\cdots+y_n^2$$

因为矩阵 C 可逆，所以关于 y 的齐次线性方程组 $C^{-1}x=\mathbf{0}$ 只有零解。任取非零向量 $\alpha\in\mathbb{R}^n$，则 $C^{-1}\alpha$ 不是零向量，令 $y=C^{-1}\alpha$，于是 $\alpha^T(C^{-1})^TC^TACC^{-1}\alpha>0$，即 $\alpha^TA\alpha>0$。由 α 的任意性，x^TAx 是正定的。 □

Theorem 2.42. 由上述定理可得到如下推论：

1. 对于 n 元实二次型 x^TAx，下述说法等价：
   - x^TAx 是正定的；
   - x^TAx 的规范形为 $y_1^2+y_2^2+\cdots+y_n^2$；
   - x^TAx 的标准形中的 n 个系数都大于 0；
2. 与正定二次型等价的实二次型也是正定的；
3. 对于 n 阶实对称矩阵 A，下述说法等价：
   - A 是正定的；
   - A 的正惯性指数为 n；
   - $A \cong I$；
   - A 的合同标准形中主对角元都大于 0；
   - A 的特征值都大于 0；
   - A 的顺序主子式都大于 0。
4. 与正定矩阵合同的实对称矩阵也是正定矩阵。
5. 正定矩阵的行列式大于 0；

Proof. (1)1 $\Leftrightarrow$ 2：由上一定理，x^TAx 正定当且仅当它的正惯性指数为 n，而 x^TAx 的正惯性指数为 n 当且仅当它的规范形为 $y_1^2+y_2^2+\cdots+y_n^2$。

2 $\Rightarrow$ 3：由标准形化规范形的步骤，若 x^TAx 的规范形为 $y_1^2+y_2^2+\cdots+y_n^2$，则其标准形中的 n 个系数必然都大于 0;

3 $\Rightarrow$ 2：当 x^TAx 的标准形中的 n 个系数都大于 0 时，也必然可以将其化为 $y_1^2+y_2^2+\cdots+y_n^2$。

(2) 由 (4)、定理 2.32和正定矩阵的定义可直接得到。

(3)1 $\Rightarrow$ 2：因为 A 是正定的，所以 n 元二次型 x^TAx 是正定的，由上一定理可得 x^TAx 的正惯性指数为 n。因为 A 的正惯性指数等于 x^TAx 的正惯性指数，所以 A 的正惯性指数为 n。

2 $\Rightarrow$ 3：因为 A 的正惯性指数为 n，由矩阵正惯性指数的定义，A 合同于 I。

3 $\Rightarrow$ 4：因为 A 合同于 I，由合同规范形的定义，I 是 A 的合同规范形，由合同标准型化合同规范形的步骤，A 的合同标准型中主对角元都大于 0。

4 $\Rightarrow$ 5：由性质 2.6.2(3) 可知 $A \cong \mathrm{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$，其中 λ_i，$i=1,2,\ldots,n$ 是 A 的特征值。显然 $\mathrm{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$ 是 A 的一个合同标准型，因为 A 的合同标准型中主对角元都大于 0，所以 A 的特征值都大于 0。

5 $\Rightarrow$ 2：显然。

2 $\Rightarrow$ 1：由定理 2.32、上一定理和矩阵正定的定义可直接得到。

$1 \Rightarrow 6$：设 n 阶实对称矩阵 A 是正定的，则对于 $k = 1, 2, \ldots, n-1$，把 A 写成分块矩阵：

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

其中 $|A_k|$ 是 A 的 k 阶顺序主子式。在 $\mathbb{R}^k$ 中任取一个非零向量 δ，因为 A 是正定矩阵，所以：

$$\begin{pmatrix} \delta \\ \mathbf{0} \end{pmatrix}^T A \begin{pmatrix} \delta \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \delta^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} A_k & B_1 \\ B_1^T & B_2 \end{pmatrix} \begin{pmatrix} \delta \\ \mathbf{0} \end{pmatrix} = \delta^T A_k \delta > 0$$

由 δ 的任意性，A_k 是正定矩阵。由 (5)，$|A_k| > 0,\ k = 1, 2, \ldots, n-1,\ |A| > 0$。

$6 \Rightarrow 1$：对实对称矩阵 A 的阶数 n 作数学归纳法。

当 $n = 1$ 时，因为 A 的顺序主子式都大于 0，所以 A 的唯一一个元素大于 0，显然此时 A 是正定矩阵。

假设对于 $n-1$ 阶实对称矩阵命题为真，考虑 n 阶实对称矩阵 $A = (a_{ij})$，将其写作分块矩阵的形式：

$$A = \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{pmatrix}$$

其中 A_{n-1} 是 $n-1$ 阶实对称矩阵，因为 A_{n-1} 的所有顺序主子式是 A 的 1 到 $n-1$ 阶顺序主子式，它们都大于 0，由归纳假设可得 A_{n-1} 是正定的。根据 (5) 可知 A_{n-1} 可逆。由 (3) 的第三条可知存在可逆矩阵 $C \in M_{n-1}(\mathbb{R})$ 使得 $C^T A_{n-1} C = I$。因为：

可逆矩阵行列式链接

$$\begin{pmatrix} I & \mathbf{0} \\ -\alpha^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{pmatrix} \begin{pmatrix} I & -A_{n-1}^{-1}\alpha \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn} - \alpha^T A_{n-1}^{-1}\alpha \end{pmatrix}$$

注意到：

$$\begin{pmatrix} I & \mathbf{0} \\ -\alpha^T A_{n-1}^{-1} & 1 \end{pmatrix}^T = \begin{pmatrix} I & (-\alpha^T A_{n-1}^{-1})^T \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} I & -A_{n-1}^{-1}\alpha \\ \mathbf{0} & 1 \end{pmatrix}$$

且：

$$\begin{pmatrix} I & \mathbf{0} \\ -\alpha^T A_{n-1}^{-1} & 1 \end{pmatrix}$$

可逆，所以 A 合同于矩阵：

$$\begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn} - \alpha^T A_{n-1}^{-1}\alpha \end{pmatrix}$$

因为：

$$\begin{aligned} \begin{vmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn} - \alpha^T A_{n-1}^{-1}\alpha \end{vmatrix} &= \begin{vmatrix} I & \mathbf{0} \\ -\alpha^T A_{n-1}^{-1} & 1 \end{vmatrix} \begin{vmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{vmatrix} \begin{vmatrix} I & -A_{n-1}^{-1}\alpha \\ \mathbf{0} & 1 \end{vmatrix} \\ &= \begin{vmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{vmatrix} = |A| \end{aligned}$$

所以 $|A_{n-1}|(a_{nn}-\alpha^T A_{n-1}^{-1}\alpha)=|A|>0$，而 $|A_{n-1}|>0$，所以 $a_{nn}-\alpha^T A_{n-1}^{-1}\alpha>0$。因为：

$$\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}^T \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{pmatrix} \begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$$
$$= \begin{pmatrix} C^T A_{n-1} C & \mathbf{0} \\ \mathbf{0} & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{pmatrix} = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{pmatrix}$$

而：

$$B=\begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{pmatrix}$$

主对角线上的元素都大于 0，由 (3) 的第四条可知 B 是一个正定矩阵。因为 $|C|1=|C|\neq 0$，所以：

可逆矩阵行列式链接

$$\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$$

可逆。于是：

$$\begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{pmatrix}$$

合同于 B。根据合同的传递性，A 合同于正定矩阵 B。由 (4)，A 是一个正定矩阵。

(4) 设 A 是一个正定矩阵，B 是一个实对称矩阵且合同于 A。由 (3) 的第三条可知 A 合同于 I，根据合同的传递性，B 也合同于 I。再由 (3) 的第三条可得 B 也是一个正定矩阵。

(5) 设 A 是一个正定矩阵，由 (3) 的第三条可得 $A\cong I$，即存在可逆矩阵 C，使得 $C^TAC=I$，于是：

$$|C^TAC|=|C^T|\,|A|\,|C|=|A|\,|C|^2=1$$

因为 $|C|^2>0$，所以 $|A|>0$。 □

半正定二次型与半正定矩阵

Definition 2.31. 如果对 $\mathbb{R}^n$ 中任意非零列向量 α，都有 $\alpha^T A\alpha\geqslant 0$，则称 n 元实二次型 x^TAx 是半正定 *(positive semidefinite)* 的。

Definition 2.32. 若实二次型 x^TAx 是半正定的，则称实对称矩阵 A 是半正定的，并称 A 为半正定矩阵 *(positive semidefinite matrix)*，记为 $A\geqslant 0$。

Theorem 2.43. 由上述定理可得到如下推论：

1. 对于 n 元实二次型 x^TAx，$\mathrm{rank}(A)=r$，下述说法等价：

   - x^TAx 是半正定的；
   - x^TAx 的正惯性指数等于 r；
   - x^TAx 的规范形为 $y_1^2+y_2^2+\cdots+y_r^2$；
   - x^TAx 的标准形中的 n 个系数都非负；

2. 与半正定二次型等价的实二次型也是半正定的；

3. 对于 n 阶实对称矩阵 A，$\text{rank}(A) = r$，下述说法等价：

   - A 是半正定的；
   - A 的正惯性指数为 r；
   - $A \cong \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$；
   - A 的合同标准形中主对角元都非负；
   - A 的特征值都非负；
   - A 的主子式都非负。

4. 与半正定矩阵合同的实对称矩阵也是半正定矩阵。

5. 半正定矩阵的行列式为 0；

Proof. (1)$1 \Rightarrow 3$：作非退化线性变换 $x = Cy$ 把 x^TAx 化作规范形：

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

若 $p < r$，取 $\alpha = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$，其中只有第 r 位为 1，则 $(C\alpha)^TA(C\alpha) = \alpha C^TAC\alpha = -1$，与 x^TAx 的非负定性矛盾，所以 $p = r$。

$3 \Rightarrow 2$：显然。

$2 \Rightarrow 4$：显然。

$4 \Rightarrow 1$：作非退化线性变换 $x = Cy$ 把 x^TAx 化作一个标准形 $d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2$，其中 $d_i \geqslant 0$, $i = 1, 2, \ldots, n$。任取 $\alpha \in \mathbb{R}^n$ 且 $\alpha \neq \mathbf{0}$。因为 C 可逆，所以 $C^{-1}x = \mathbf{0}$ 只有零解，于是 $C^{-1}\alpha = (b_1, b_2, \ldots, b_n) \neq \mathbf{0}$，所以：

$$(C^{-1}\alpha)^TC^TACC^{-1}\alpha = \sum_{i=1}^{n} d_ib_i^2 \geqslant 0$$

而：

$$(C^{-1}\alpha)^TC^TACC^{-1}\alpha = \alpha^T(C^{-1})^TC^TACC^{-1}\alpha = \alpha^T(C^T)^{-1}C^TACC^{-1}\alpha = \alpha^TA\alpha$$

所以 $\alpha^TA\alpha \geqslant 0$。由 α 的任意性，x^TAx 半正定。

(2) 由 (4)、定理 2.32和半正定矩阵的定义可直接得到。

(3)$1 \Rightarrow 2$：因为 A 是半正定的，所以 x^TAx 是半正定的。由 (1) 的第二条，x^TAx 的正惯性指数等于 r，而 A 的正惯性指数等于 x^TAx 的正惯性指数，所以 A 的正惯性指数为 r。

$2 \Rightarrow 3$：因为 A 的正惯性指数为 r，由矩阵正惯性指数的定义，$A \cong \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$。

$3 \Rightarrow 4$：因为 $A \cong C = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$，所以 C 是 A 的合同规范形。由合同标准形化合同规范形的步骤，A 的合同标准形中主对角元都大于 0。

$4 \Rightarrow 5$：由性质 2.6.2(3) 可知 $A \cong \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$，其中 λ_i，$i = 1, 2, \ldots, n$ 是 A 的特征值。显然 $\mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ 是 A 的一个合同标准型，因为 A 的合同标准型中主对角元都非负，所以 A 的特征值都非负。

$5 \Rightarrow 2$：因为 $\mathrm{rank} = r$，所以 A 的相似标准形主对角线上的元素有且只有 r 个非零，由条件它们也非负，于是它们为正数，显然此时 A 的正惯性指数为 r。

$2 \Rightarrow 1$：由定理 2.32、(1) 的第二条和矩阵半正定的定义可直接得到。

$1 \Rightarrow 6$：

有空证明

$6 \Rightarrow 5$：

(4) 设 A 是一个半正定矩阵，B 是一个实对称矩阵且合同于 A。由 (3) 的第三条可知 $A \cong C = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$，根据合同的传递性，$B \cong C$。再由 (3) 的第三条可得 B 也是一个半正定矩阵。

(5) 设 A 是一个 n 阶半正定矩阵，由 (3) 的第三条，存在可逆矩阵 C 使得：

$$C^T A C = B = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

而 $\mathrm{rank}(B) = r$，因为可逆变换不改变矩阵的秩，所以 $\mathrm{rank}(A) = r < n$，于是 $|A| = 0$。 □

负定矩阵

Definition 2.33. 如果对 $\mathbb{R}^n$ 中任意非零列向量 α，都有 $\alpha^T A \alpha < 0$，则称 n 元实二次型 $x^T A x$ 是负定 *(negative definite)* 的。

Definition 2.34. 若实二次型 $x^T A x$ 是负定的，则称实对称矩阵 A 是负定的，并称 A 为负定矩阵 *(negative definite matrix)*，记为 $A < 0$。

Theorem 2.44. 对称矩阵 $A \in M_n(\mathbb{R})$ 负定的充分必要条件为：它的奇数阶顺序主子式都小于 0，偶数阶顺序主子式都大于 0。

Proof. 设 $|A_k|$ 为 A 的 k 阶顺序主子式，由定理 2.42(3) 的第六条：

$$\begin{aligned} & A\text{是负定矩阵} \\ \Leftrightarrow & (-A)\text{是正定矩阵} \\ \Leftrightarrow & (-1)^k |A_k| > 0 \\ \Leftrightarrow & \begin{cases} |A_k| > 0, & k\text{为偶数} \\ |A_k| < 0, & k\text{为奇数} \end{cases} \end{aligned}$$

□

2.8 特殊矩阵

2.8.1 幂等阵

Definition 2.35. 若矩阵 $A \in M_n(K)$ 满足 $A^2 = I_n$，则称 A 为幂等矩阵 *(idempotent matrix)*。

Property 2.8.1. 设 $A \in M_n(K)$ 是一个幂等阵，$\mathrm{rank}(A) = r$，则：

1. A 的特征值只能是 1 或 0；
2. $\mathrm{tr}(A) = \mathrm{rank}(A)$；
3. A幂等 $\Leftrightarrow \mathrm{rank}(A) + \mathrm{rank}(I_n - A) = n$；
4. 存在秩为 r 的 $B \in M_n(K)$ 使得 $A = B(B^T B)^- B^T$；

Proof. (1) 设 λ 为 A 的一个特征值，φ 为对应的特征向量，因为 A 是一个幂等阵，所以 $A^2\varphi = A\varphi = \lambda\varphi$，又因为：

$$A^2\varphi = AA\varphi = A\lambda\varphi = \lambda A\varphi = \lambda^2\varphi$$

所以 $(\lambda^2 - \lambda)\varphi = \mathbf{0}$。因为 φ 是特征向量，所以 $\varphi \neq \mathbf{0}$，于是 $\lambda^2 - \lambda = 0$，即 $\lambda = 1$ 或 $\lambda = 0$。由 λ 的任意性，结论成立。

(2) 因为 $\mathrm{rank}(A) = r$，所以存在可逆矩阵 P, Q 使得：

$$A = P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q = \begin{pmatrix} P_1 & P_2 \end{pmatrix} \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$$

其中 P_1 为 $n \times r$ 矩阵，Q_1 为 $r \times n$ 矩阵，于是 $A = P_1 Q_1$。因为 A 是一个幂等阵，所以：

$$P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QP \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q = P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q$$

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QP \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} P_1 & P_2 \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$\begin{pmatrix} Q_1 P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

即 $Q_1 P_1 = I_r$。由**??**(3) 可得：

$$\mathrm{tr}(A) = \mathrm{tr}(P_1 Q_1) = \mathrm{tr}(Q_1 P_1) = \mathrm{tr}(I_r) = r = \mathrm{rank}(A)$$

(3) □

Property 2.8.2. 设 $A \in M_n(\mathbb{C})$ 是一个 *Hermitian* 幂等阵，则：

1.
2. $\mathrm{rank}(A) = \mathrm{tr}(A)$。

Proof. (1) (2) 由性质 2.6.2(3) 和 (1) 可知存在一个正交矩阵 Q 使得：

$$A = Q^{-1}\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q$$

由性质 2.4.1可得：

$$\operatorname{rank}(A) = \operatorname{rank}(I_r) = r$$

根据**??**(3) 可得：

$$\operatorname{tr}(A) = \operatorname{tr}\left[Q^{-1}\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q\right] = \operatorname{tr}\left[\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} QQ^{-1}\right] = \operatorname{tr}\left[\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\right] = r$$

所以有 $\operatorname{rank}(A) = \operatorname{tr}(A)$。 □

2.9 矩阵的分解

2.9.1 SVD 分解

Theorem 2.45. 设 $A \in M_{m\times n}(\mathbb{C})$，则 $AA^H, A^H A$ 是半正定矩阵。

Proof. 设 λ_i，$i = 1, 2, \ldots, n$ 是矩阵 $A^H A$ 的特征值，ξ_i 是对应的特征向量，则：

$$A^H A\xi_i = \lambda_i \xi_i \to \xi_i^H A^H A\xi_i = \lambda_i \xi_i^H \xi_i \to ||A\xi_i||^2 = \lambda_i ||\xi_i||^2$$

由于左式非负，所以右式非负，而 $||\xi_i||^2$ 非负，因此 λ_i 非负，由定理 2.43(3) 的第五条可知 AA^T 是半正定矩阵。 □

Theorem 2.46. 设 $A \in M_{m\times n}(\mathbb{C})$，$\operatorname{rank}(A) = r$，则存在两个正交矩阵 $P \in M_m(\mathbb{C})$, $Q \in M_n(\mathbb{C})$ 使得：

$$A = P\begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q^H$$

其中 $\Lambda = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_r\}$，$\lambda_i > 0$，$\lambda_i^2$ 为 $A^H A$ 的正特征值。

Proof. 由定理 2.2可知 $\operatorname{rank}(A^H A) = \operatorname{rank}(A)$。于是 $A^H A$ 确实有 r 个正特征值。因为 $A^H A$ 是一个 Hermitian 矩阵，由性质 2.6.2可知存在正交矩阵 $Q \in M_n(\mathbb{C})$ 使得：

$$Q^H A^H A Q = \begin{pmatrix} \Lambda^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

记 $B = AQ$，则：

$$B^H B = \begin{pmatrix} \Lambda^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

这表明 B 的列向量相互正交，且前 r 个列向量的长度分别为 $\lambda_1, \lambda_2, \ldots, \lambda_r$，后 $n-r$ 个列向量为零向量，于是存在一个正交矩阵 $P \in M_m(\mathbb{C})$ 使得：

$$B = P \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

因为 $B = AQ$，所以：

$$A = P \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q^{-1} = P \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} Q^H$$ □

Definition 2.36. 设 $A \in M_{m\times n}(\mathbb{C})$，$\operatorname{rank}(A) = r$，$A^H A$ 的正特征值为 λ_i，$i = 1, 2, \ldots, r$，称 $\delta_i = \sqrt{\lambda_i}$ 为矩阵 A 的奇异值 *(singular value)*。

Chapter 3

线性模型

Definition 3.1.

$$\begin{cases} y = X\beta + \varepsilon \\ \mathrm{E}(\varepsilon) = \mathbf{0} \\ \mathrm{Cov}(\varepsilon) = \sigma^2 I_n \end{cases}$$

其中 y 为 $n \times 1$ 观测向量，X 为 $n \times p$ 的设计矩阵，β 为 $p \times 1$ 的未知参数向量，ε 为随机误差，σ^2 为误差方差。

Definition 3.2. 称方程 $X^T X\beta = X^T y$ 为**正则方程**。

Theorem 3.1. 对于定义 3.1，$\hat{\beta} = (X^T X)^- X^T y$ 是其唯一的最小二乘解。

Proof. 注意到：

$$\begin{aligned} Q(\beta) &= ||y - X\beta||^2 = (y - X\beta)^T (y - X\beta) \\ &= y^T y - y^T X\beta - \beta^T X^T y - \beta^T X^T X\beta \\ &= y^T y - 2y^T X\beta - \beta^T X^T X\beta \end{aligned}$$

$$\frac{\partial y^T X\beta}{\beta} = X^T y, \quad \frac{\partial \beta^T X^T X\beta}{\beta} = 2X^T X\beta$$

$$\frac{\partial Q(\beta)}{\partial \beta} = 2X^T y - 2X^T X\beta = 0$$

$$X^T X\beta = X^T y$$

由定理 2.1可知方程 $X^T X\beta = X^T y$ 是相容的，根据定理 2.20可知其通解为：

$$\hat{\beta} = (X^T X)^- X^T y$$

其中 $(X^T X)^-$ 是 $X^T X$ 的任意一个广义逆矩阵。

对任意的 β，有：

$$\begin{aligned} Q(\beta) &= ||y - X\beta||^2 = ||y - X\hat{\beta} + X\hat{\beta} - X\beta||^2 = ||y - X\hat{\beta} + X(\hat{\beta} - \beta)||^2 \\ &= ||y - X\hat{\beta}||^2 + ||X(\hat{\beta} - \beta)||^2 + 2(y - X\hat{\beta})^T X(\hat{\beta} - \beta) \end{aligned}$$

注意到正则方程即为：

$$X^T(y - X\beta) = \mathbf{0}$$

于是：

$$2(y - X\hat{\beta})^T X(\hat{\beta} - \beta) = 2[X^T(y - X\hat{\beta})]^T(\hat{\beta} - \beta) = 0$$

所以：

$$Q(\beta) = ||y - X\hat{\beta}||^2 + ||X(\hat{\beta} - \beta)||^2$$

上第二项总是非负的，由范数的性质其为 0 当且仅当 $X\hat{\beta} = X\beta$，即当且仅当 $X^TX\beta = X^TX\hat{\beta} = X^Ty$，所以使 $Q(\beta)$ 达到最小值的 β 必为正则方程的解 $\hat{\beta} = (X^TX)^-X^Ty$。 □

推导 3.1. 若 $\mathrm{rank}(X) = p$，则 X 的列向量组线性无关。考虑二次型 y^TX^TXy，$y^TX^TXy = 0 \Leftrightarrow ||Xy|| = 0 \Leftrightarrow Xy = \mathbf{0}$，而 X 的列向量是线性无关的，所以不存在非零向量的 y 使得 $Xy = \mathbf{0}$，于是 y^TX^TXy 是一个正定二次型，X^TX 是一个正定矩阵。由定理 2.42(3) 的第五点和可得 X^TX 可逆。此时 $\hat{\beta} = (X^TX)^{-1}X^Ty$，称 $\hat{\beta}$ 为 β 的最小二乘估计 *(least squares estimate)*。

行列式等于特征值的积，行列式大于 0 矩阵可逆

Definition 3.3. 若存在 $n \times 1$ 向量 α 使得 $\mathrm{E}(\alpha^Ty) = c^T\beta$ 对一切的 β 成立，则称 $c^T\beta$ 为可估函数 *(estimable function)*。

Property 3.0.1. 对于定义 3.1，$c^T\beta$ 和 $d^T\beta$ 是可估函数，$\hat{\beta}$ 是正则方程的解，则：

1. 使 $c^T\beta$ 成为可估函数的全体向量 c 构成 $\mathcal{M}(X^T)$；
2. 若 $c_1^T\beta$ 和 $c_2^T\beta$ 都是可估函数，则对任意常数 a_1, a_2，$a_1c_1^T\beta + a_2c_2^T\beta$ 也是可估函数；
3. 线性无关的可估函数组最多有 $\mathrm{rank}(X)$ 个可估函数；
4. $c^T\hat{\beta}$ 与 $(X^TX)^-$ 的选择无关；
5. $c^T\hat{\beta}$ 为 $c^T\beta$ 的无偏估计；
6. $\mathrm{Var}(c^T\hat{\beta}) = \sigma^2c^T(X^TX)^-c$，$\mathrm{Cov}(c^T\hat{\beta}, d^T\hat{\beta}) = \sigma^2c^T(X^TX)^-d$，且与 $(X^TX)^-$ 的选择无关；
7. $c^T\hat{\beta}$ 是 $c^T\beta$ 唯一的 *BLUE*；
8. 设 $\varphi_i = c_i^T\beta$，$i = 1, 2, \ldots, k$ 都是可估函数，$\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$，则 $\varphi = \sum\limits_{i=1}^{k} \alpha_i\varphi_i$ 也是可估的，且 $\hat{\varphi} = \sum\limits_{i=1}^{k} \alpha_ic_i^T\hat{\beta}$ 是 φ 的 *BLU* 估计。

Proof. (1) $c^T\beta$ 是可估函数 $\Leftrightarrow$ 存在 $n \times 1$ 向量 α 使得 $\mathrm{E}(\alpha^Ty) = \alpha^T\mathrm{E}(y) = \alpha^TX\beta = c^T\beta$ 对一切的 β 成立 $\Leftrightarrow c = X^T\alpha$。

(2) 由 (1) 直接可得。

(3) 由 (1) 和直接可得。

转置秩不变

(4) 因为 $c^T\beta$ 可估，由 (1) 可知存在 $n\times 1$ 向量 α 使得 $c = X^T\alpha$，于是：

$$c^T\hat{\beta} = \alpha^T X(X^TX)^- X^T y$$

由性质 2.5.1(4) 即可得出结论。

(5) 因为 $c^T\beta$ 可估，由 (1) 可知存在 $n\times 1$ 向量 α 使得 $c = X^T\alpha$，根据性质 2.5.1(5) 可得：

$$\mathrm{E}(c^T\hat{\beta}) = \mathrm{E}[\alpha^T X(X^TX)^- X^T y] = \alpha^T X(X^TX)^- X^T X\beta = \alpha^T X\beta = c^T\beta$$

(6) 因为 $c^T\beta, d^T\beta$ 是可估函数，所以存在 α,γ 使得 $c = X^T\alpha, d = X^T\gamma$。由**??**(3) 和性质 2.5.1(6)(5) 可知：

$$\begin{aligned}\mathrm{Cov}(c^T\hat{\beta}, d^T\hat{\beta}) &= \mathrm{Cov}[\alpha^T X(X^TX)^- X^T y, \gamma^T X(X^TX)^- X^T y]\\ &= \alpha^T X(X^TX)^- X^T\,\mathrm{Cov}(y) X[(X^TX)^-]^T X^T\gamma\\ &= \alpha^T X(X^TX)^- X^T\sigma^2 I_n X(X^TX)^- X^T\gamma\\ &= \sigma^2\alpha^T X(X^TX)^- d\\ &= \sigma^2 c^T(X^TX)^- d\end{aligned}$$

由性质 2.5.1(4) 及上第三行可知 $\mathrm{Cov}(c^T\hat{\beta}, d^T\hat{\beta})$ 与 $(X^TX)^-$ 的选择无关。

(7) 无偏性由 (5) 可得，线性性由正则方程可知，下证方差最小。设 a^Ty 为 $c^T\beta$ 的任一无偏估计，由 (1) 的过程可知 $c = X^Ta$。根据性质 2.5.2(3) 和 (6) 可得：

$$\begin{aligned}\mathrm{Var}(a^Ty) - \mathrm{Var}(c^T\hat{\beta}) &= \sigma^2[a^Ta - c^T(X^TX)^- c]\\ &= \sigma^2[a^T - c^T(X^TX)^- X^T][a - X(X^TX)^- c]\\ &= \sigma^2\|a - X(X^TX)^- c\|^2 \geqslant 0\end{aligned}$$

上式第一行到第二行是由于性质 2.5.2(3)：

$$\begin{aligned}&[a^T - c^T(X^TX)^- X^T][a - X(X^TX)^- c]\\ =&a^Ta - a^TX(X^TX)^- c - c^T(X^TX)^- X^Ta + c^T(X^TX)^- X^TX(X^TX)^- c\\ =&a^Ta - c^T(X^TX)^- c - a^TX(X^TX)^- c + a^TX(X^TX)^- X^TX(X^TX)^- c\\ =&a^Ta - c^T(X^TX)^- c - a^TX(X^TX)^- c + a^TX(X^TX)^- c\\ =&a^Ta - c^T(X^TX)^- c\end{aligned}$$

由范数的性质可知 $\mathrm{Var}(a^Ty) = \mathrm{Var}(c^T\hat{\beta})$ 当且仅当 $a = X(X^TX)^- c$，由性质 2.5.2(3) 可知 $a = X(X^TX)^- c \Leftrightarrow a^T = c^T(X^TX)^- X^T \Leftrightarrow a^Ty = c^T(X^TX)^- X^Ty = c^T\hat{\beta}$。

(8) 因为 $\varphi_1, \varphi_2, \ldots, \varphi_k$ 都是可估函数，所以存在 $b_1, b_2, \ldots, b_k$ 使得 $\mathrm{E}(b_i^Ty) = c_i^T\beta$，于是：

$$\mathrm{E}\left(\sum_{i=1}^k \alpha_i b_i^T y\right) = \sum_{i=1}^k \alpha_i\,\mathrm{E}(b_i^Ty) = \sum_{i=1}^k \alpha_i c_i^T\beta = \sum_{i=1}^k a_i\varphi_i = \varphi$$

所以取 $\alpha = \sum_{i=1}^{k} \alpha_i b_i$ 即可得到 $\mathrm{E}(\alpha^T y) = \varphi$，$\varphi$ 是可估的。

由 (5) 可得 $c_i^T \hat{\beta}$ 是 $c_i^T \beta$ 的无偏估计，所以：

$$\mathrm{E}(\hat{\varphi}) = \mathrm{E}\left(\sum_{i=1}^{k} \alpha_i c_i^T \hat{\beta}\right) = \sum_{i=1}^{k} \alpha_i \,\mathrm{E}(c_i^T \hat{\beta}) = \sum_{i=1}^{k} \alpha_i c_i^T \beta = \varphi$$

即 $\hat{\varphi}$ 是一个无偏估计。

令 $c = \sum_{i=1}^{k} \alpha_i c_i$，则 $\varphi = c^T \beta$。设 $\gamma^T y$ 是 φ 的一个无偏估计，于是由 (7) 可得：

$$\mathrm{Var}(\gamma^T y) - \mathrm{Var}(c^T \hat{\beta}) = \sigma^2 ||\gamma - X(X^T X)^{-} c||^2$$

上式等于 $0 \Leftrightarrow \gamma^T y = c^T \hat{\beta} = \hat{\varphi}$，即 $\hat{\varphi}$ 是唯一的 BLUE。 $\square$

Definition 3.4. 对于定义 3.1，若 $c^T \beta$ 是可估函数，称 $c^T \hat{\beta}$ 为 $c^T \beta$ 的 *LS* 估计，其中 $\hat{\beta}$ 为正则方程的解。

Definition 3.5. 称 $\hat{e} = y - X\hat{\beta}$ 为**残差向量**。

Property 3.0.2. 对于定义 3.1，$\hat{\beta}$ 为正则方程的解，则残差向量 $\hat{e}$ 满足 $\mathrm{E}(\hat{e}) = 0$, $\mathrm{Cov}(\hat{e}) = \sigma^2(I - P_X)$。

Proof. 由可知： 对称幂等阵

$$\begin{aligned} \mathrm{E}(\hat{e}) &= \mathrm{E}(y - X\hat{\beta}) = \mathrm{E}[I_n y - X(X^T X)^{-} X^T y] = (I_n - P_X)\,\mathrm{E}(y) \\ &= (I_n - P_X) X \beta = (X - X)\beta = 0 \\ \mathrm{Cov}(\hat{e}) &= \mathrm{Cov}[(I_n - P_X) y] = (I_n - P_X)\,\mathrm{Cov}(y)(I_n - P_X) = \sigma^2 (I_n - P_X) \end{aligned}$$

$\square$