高代笔记

所用教材为蓝以中——《高等代数简明教程》

1 线性空间和线性变换

线性空间的基本概念

定义 1 设 V 是一个非空集合，K 是一个数域，又设

(i) 在 V 中定义了一种运算，称为**加法**，即对 V 中的任意两个元素 α, β，都按某一法则对应于 V 内唯一确定的一个元素，记为 $\alpha + \beta$；

(ii) 在 K 中的数与 V 中的元素间定义了一种运算，称为**数乘**，即对 V 中的任意元素 α 和数域 K 中的任意数 k，都按某一法则对应于 V 内唯一确定的一个元素，记之为 $k\alpha$.

如果加法与数乘满足下面列出的八条运算法则，那么称 V 是数域 K 上的一个**线性空间**:

1. 对任意 $\alpha, \beta, \gamma \in V$，有 $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$;
2. 对任意 $\alpha, \beta \in V$，有 $\alpha + \beta = \beta + \alpha$;
3. 存在一个元素 $0 \in V$，使对一切 $\alpha \in V$，有 $\alpha + 0 = \alpha$，此元素 0 称为 V 的**零元素**;
4. 对任一 $\alpha \in V$ 都存在 $\beta \in V$，使 $\alpha + \beta = 0$，β 称为 α 的一个**负元素**;
5. 对数域中的数 1，有 $1 \cdot \alpha = \alpha$;
6. 对任意 $k, l \in K, \alpha \in V$，有 $(kl)\alpha = k(l\alpha)$;
7. 对任意 $k, l \in K, \alpha \in V$，有 $(k + l)\alpha = k\alpha + l\alpha$;
8. 对任意 $k \in K, \alpha, \beta \in V$，有 $k(\alpha + \beta) = k\alpha + k\beta$.

因为线性空间是从向量空间抽象出来的，所以我们把线性空间的元素也称为**向量**；显然，它已不再具有三维几何空间中向量的几何的直观意义，也不再具有 m 维向量空间中向量的 m 元有序数组的具体形式. 线性空间中的零元素称为**零向量**，一个向量 α 的负元素则称为 α 的**负向量**.

命题 1 设 V 是数域 K 上的一个线性空间，如果在 V 中存在 n 个线性无关的向量

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n \tag{I}$$

使 V 中的任一向量均能被 (I) 线性表示，那么，有:

- 任给 V 内的 n 个线性无关向量
$$\eta_1, \eta_2, \cdots, \eta_n \tag{II}$$
则 V 中的任一向量都能被 (II) 线性表示；
- 如果 $\eta_1, \eta_2, \cdots, \eta_s$ 是 V 的一个线性无关向量组，且 V 内任一向量均能被它们线性表示，则 $s = n$.

定义 1 设 V 是数域 K 上的一个线性空间。如果 V 中存在 n 个线性无关的向量

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n,$$

使 V 中任一向量均能被此向量组线性表示，则 V 称为 n **维线性空间**，记做 $\dim V = n$。上述向量组称为 V 的一组**基**。单由零向量组成的线性空间称为**零空间**。零空间的维数定义为零。

从命题 1 可知：在 n 维线性空间中，任意 n 个线性无关向量都是它的一组基，反之，它的任意一组基也一定恰好含有 n 个向量。

显然，在一个 n 维线性空间中，任意 $n+1$ 个向量都是线性相关的，所以，在一个 n 维线性空间中最多只有 n 个线性无关的向量。

如果一个线性空间有 n 个线性无关向量，而任意 $n+1$ 个向量都线性相关，那么它即为 n 维线性空间。

如果在一个线性空间中存在任意多个线性无关的向量，则称之为**无限维线性空间**。而 n 维线性空间则统称为**有限维线性空间**

命题 1 有一个明显的推论:

推论 设 V 是数域 K 上的 n 维线性空间。$a_1, a_2, \cdots, a_n$ 是 V 中一个向量组，使 V 中任一向量可被它线性表示，则此向量组为 V 的一组基。

向量的坐标 设 V 是数域 K 上的 n 维线性空间，而 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是它的一组基。任给 $a \in V$，有表示式

$$a = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n.$$

由于 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 线性无关，由此前命题，这个表达式的系数是唯一确定的。我们称 $a_1, a_2, \ldots, a_n$ 为 a 在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 下的**坐标**。

对数域 K 上任意多项式 $f(x) = a_0x^m + a_1x^{m-1} + a_2x^{m-2} + \cdots + a_{m-1}x + a_m$, 定义 $f'(x) = a_0mx^{m-1} + a_1(m-1)x^{m-2} + a_2(m-2)x^{m-3} + \cdots + a_{m-1}$, 称 $f'(x)$ 为 $f(x)$ 的**形式微商**，它仍为 K 上的一个多项式，但次数比 $f(x)$ 的次数减 1 或变为零多项式。

现在设 $f(x)$ 是 $K[x]_n$ 内任意多项式，a 为 K 内任意数，那么有 $f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_{n-1}(x-a)^{n-1}$. 对 $f(x)$ 逐次求形式微商，再令 $x = a$ 代入，立即得

$$\begin{aligned}
a_0 &= f(a), \\
a_1 &= f'(a), \\
a_2 &= \frac{1}{2!}f''(a), \\
&\vdots \\
a_{n-1} &= \frac{1}{(n-1)!}f^{(n-1)}(a).
\end{aligned}$$

于是 $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}$. 这称为数域 K 上多项式的 Taylor 展开式。由此即可知 $f(x)$ 在基 $1, (x-a), (x-a)^2, \cdots, (x-a)^{n-1}$ 下的坐标表示式为

$$f(x) = (1, x-a, (x-a)^2, \cdots, (x-a)^{n-1}) \begin{bmatrix} f(a) \\ f'(a) \\ \frac{f''(a)}{2!} \\ \vdots \\ \frac{f^{(n-1)}(a)}{(n-1)!} \end{bmatrix}.$$

命题 2 设 V 是数域 K 上的 n 维线性空间。在 V 内取定一组基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. 设 V 内向量组

$a_1, a_2, \cdots, a_s$ 中各向量在此组基下的坐标表示式为 $a_i = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A_i \quad (i = 1, 2, \cdots, s)$, 则 $a_1, a_2, \cdots, a_s$ 线性相关（无关）的充分必要条件是 $A_1, A_2, \cdots, A_s$ 在 K^n 内线性相关（无关）。

定义 2 设在 n 线性空间 V 内给定两组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n,$$

$$\eta_1, \eta_2, \cdots, \eta_n,$$

每个 η_i 能被 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性表示。设

$$\begin{gathered}
\eta_1 = t_{11}\varepsilon_1 + t_{21}\varepsilon_2 + \cdots + t_{n1}\varepsilon_n, \\
\eta_2 = t_{12}\varepsilon_1 + t_{22}\varepsilon_2 + \cdots + t_{n2}\varepsilon_n, \\
\vdots \\
\eta_n = t_{1n}\varepsilon_1 + t_{2n}\varepsilon_2 + \cdots + t_{nn}\varepsilon_n.
\end{gathered}$$

再度借助矩阵乘法法则，把上面的公式形式地写成

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix}.$$

称 T 为从基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到基 $\eta_1, \eta_2, \cdots, \eta_n$ 的**过渡矩阵**。

命题 3 在数域 K 上的 n 维线性空间 V 内给定一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$. T 是 K 上一个 n 方阵。命 $(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)T$. 则有

(i) 如果 $\eta_1, \eta_2, \cdots, \eta_n$ 是 V 的一组基，则 T 可逆；

(ii) 如果 T 可逆，则 $\eta_1, \eta_2, \cdots, \eta_n$ 是 V 的一组基。

坐标变换公式

设 V 中一个向量 α 在第一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标为 $x_1, x_2, \cdots, x_n$，即

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

又设 α 在第二组基 $\eta_1, \eta_2, \cdots, \eta_n$ 下的坐标为 $y_1, y_2, \cdots, y_n$，即

$$\alpha = y_1\eta_1 + y_2\eta_2 + \cdots + y_n\eta_n = (\eta_1, \eta_2, \cdots, \eta_n) \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

现设两组基间的过渡矩阵为 T，即

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

令

$$X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad Y=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}.$$

那么

$$\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X=(\eta_1,\eta_2,\cdots,\eta_n)Y.$$

以关系式

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T$$

代入，得

$$\begin{aligned}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X&=[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T]Y\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(TY).\end{aligned}$$

由于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组基，线性无关，它们的两个线性组合相等时，对应系数相等，故有 $X=TY$.

子空间和商空间

定义 3 设 V 是数域 K 上的一个线性空间，M 是 V 的一个非空子集。如果 M 关于 V 内的加法与数乘运算也组成数域 K 上的一个线性空间，则 M 称为 V 的一个**子空间**。

命题 4 线性空间 V 的一个非空子集 M 是一个子空间的充分必要条件是，它满足以下两个条件：

1. 它对加法封闭，即对 M 内任意两个向量 α,β，有 $\alpha+\beta\in M$;
2. 它对数乘运算封闭，即对任一 $\alpha\in M$ 和任一 $k\in K$，有 $k\alpha\in M$.

证明. 条件的必要性是显然的，我们证明充分性。设条件 (i) 和 (ii) 满足。我们只要证明：零向量属于 M; M 中任一向量 α 的负向量 $-\alpha$ 属于 M。那么，线性空间定义中的八条性质就都具备，于是 M 就是子空间了。因为 M 非空，必有某个 $\alpha\in M$。由条件 (ii) 知，$0\cdot\alpha=0\in M$。另一方面，对任意 $\alpha\in M$，$(-1)\cdot\alpha=-\alpha\in M$。至此命题得证。 □

现在设 $a_1,a_2,\cdots,a_s$ 是 V 内一个向量组，作它们的所有可能的线性组合，由此所得的 V 的子集记做 $L(a_1,a_2,\cdots,a_s)$. 如果采用集合论中惯用的记号，它可写做

$$L(a_1,a_2,\cdots,a_s)=\{k_1a_1+k_2a_2+\cdots+k_sa_s\mid k_i\in K,i=1,2,\cdots,s\}.$$

这个记号的意思是：集合 $L(a_1,a_2,\cdots,a_s)$ 的元素是由花括号中所表达出来的向量组成的，而其中线性组合的各个系数 $k_i(i=1,2,\cdots,s)$ 各自遍历数域 K 中的一切元素。显然，子集 $L(a_1,a_2,\cdots,a_s)$ 满足命题 2.1 中的条件 (i) 和 (ii)，故它是 V 的一个子空间，这个子空间称为**由向量组** $a_1,a_2,\cdots,a_s$ **生**

成的子空间。容易看出，向量组 $a_1, a_2, \cdots, a_s$ 的任一极大线性无关部分组都是子空间 $L(a_1, a_2, \cdots, a_s)$ 的一组基，而由此向量组的秩即为 $L(a_1, a_2, \cdots, a_s)$ 的维数。

不难发现，对于一个线性空间 V，单由它的零向量组成的子集也满足命题 4 中的条件，所以它也是 V 的一个子空间，被称为**零子空间**，记作 $\{0\}$，此外，V 也被认为是自身的一个子空间，这两个子空间被称为**平凡子空间**，其余的空间都被称为**非平凡子空间**。

命题 5 设 V 是数域 K 上的 n 维线性空间，M 是 V 的一个非零子空间，则 M 内任一组基都可以扩充成 V 的一组基。

证明. n 维线性空间中至多有 n 个线性无关的向量，于是 M 内线性无关向量的个数也不能大于 n，故 M 也是有限维线性空间。设 $\dim M = r$，且 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 为 M 的一组基。于是 $M = L(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r)$。

若 $M = V$，则 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 已经是 V 的一组基。如 $M \neq V$，则存在 $\varepsilon_{r+1} \in V$，而 $\varepsilon_{r+1} \notin M$。此时 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}$ 必线性无关。因若 $\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{r+1}$ 线性相关，则 ε_{r+1} 可被 $\varepsilon_1, \cdots, \varepsilon_r$ 线性表示，于是 $\varepsilon_{r+1} \in L(\varepsilon_1, \cdots, \varepsilon_r) = M$，与假设矛盾。

如果 $L(\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{r+1}) \neq V$，再继续上述步骤。经过 $n - r$ 次这种步骤后，得到 V 内 n 个线性无关的向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，故知它们组成 V 的一组基。 □

推论 一个有限维线性空间 V 内任意 r 个线性无关的向量 $a_1, a_2, \cdots, a_r$ 都可以扩充为 V 的一组基。

子空间的交与和

定义 4 设 M_1, M_2 为线性空间 V 的两个子空间，称 $M_1 \cap M_2$ 为它们的**交**。又命 $M = \{\alpha_1 + \alpha_2 \mid \alpha_1 \in M_1, \alpha_2 \in M_2\}$, 称为 M_1 与 M_2 的**和**，记做 $M_1 + M_2$。

命题 6 设 M_1, M_2 是线性空间 V 的两个子空间，则它们的交 $M_1 \bigcap M_2$ 与和 $M_1 + M_2$ 仍是 V 的子空间. 一般的，可以递归定义更多子空间的交与和.

命题 7（维数公式） 设 M_1, M_2 是线性空间 V 的两个有限维子空间，则有

$$dim(M_1 + M_2) = dimM_1 + dimM_2 - dim(M_1 \bigcap M_2)$$

推论 设 $M_1, M_2, \cdots, M_k$ 是线性空间 V 的 k 个有限维子空间，则

$$dim(M_1 + M_2 + \cdots + M_k) = dimM_1 + dimM_2 + \cdots + dimM_k$$

子空间的直和

定义 4 设 M_1, M_2 是线性空间 V 的两个子空间，$M_1 + M_2 = M$。如果对 M 内任一向量 α，其表示式 $\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2)$ 是唯一的，则称 M 是 M_1 与 M_2 的**直和**（亦称 $M_1 + M_2$ 是直和），记做 $M = M_1 \oplus M_2$.

命题 8 设 M_1, M_2 是数域 K 上线性空间 V 的两个有限维子空间，则下面几条互相等价：

1. $M_1 + M_2$ 是直和；
2. 0 向量表法唯一，即由 $0 = \alpha_1 + \alpha_2 \quad (\alpha_1 \in M_1, \alpha_2 \in M_2)$, 必定有 $\alpha_1 = \alpha_2 = 0$;
3. $M_1 \cap M_2 = \{0\}$;
4. $\dim M_1 + \dim M_2 = \dim(M_1 + M_2)$.

证明. 采用轮转方式证明这些命题等价。

(i) ⇒ (ii) 按定义，M_1+M_2 内任一向量表法唯一，因而零向量的表法当然是唯一的。

(ii) ⇒ (iii) 用反证法。若 $M_1\cap M_2\neq\{0\}$，则有一 $\beta\in M_1\cap M_2$，$\beta\neq 0$。于是 $\beta\in M_1$，$-\beta\in M_2$，而 $0=\beta+(-\beta)$. 这与零向量表法唯一的假设矛盾。

(iii) ⇒ (iv) 利用命题 7 即得。

(iv) ⇒ (i) 利用命题 7 知 $\dim(M_1\cap M_2)=0$, 即 $M_1\cap M_2=\{0\}$。对任一 $\alpha\in M_1+M_2$, 如果 $\alpha=\alpha_1+\alpha_2=\alpha_1'+\alpha_2'\quad(\alpha_1,\alpha_1'\in M_1;\alpha_2,\alpha_2'\in M_2)$, 则有 $\alpha_1-\alpha_1'=\alpha_2'-\alpha_2$. 于是 $\alpha_1-\alpha_1'\in M_1\cap M_2=\{0\}$, 即 $\alpha_1-\alpha_1'=0$，$\alpha_2-\alpha_2'=0$。这说明 $\alpha_1=\alpha_1'$，$\alpha_2=\alpha_2'$，因而 α 表法唯一。

□

定义 5 设 $M_1,M_2,\cdots,M_k$ 为线性空间 V 的子空间，$M_1+M_2+\cdots+M_k=M$。如果对 M 中任一向量 α, 表达式 $\alpha=\alpha_1+\alpha_2+\cdots+\alpha_k\quad(\alpha_i\in M_i,i=1,2,\cdots,k)$ 是唯一的, 则称 M 是 $M_1,M_2,\cdots,M_k$ 的直和（亦称 $M_1+M_2+\cdots+M_k$ 是直和），记做

$$M=M_1\oplus M_2\oplus\cdots\oplus M_k=\bigoplus_{i=1}^{k}M_i.$$

命题 9 设 $M_1,M_2,\cdots,M_k$ 是数域 K 上线性空间 V 的有限维子空间，则下列命题互相等价：

1. $M=M_1+M_2+\cdots+M_k$ 是直和；
2. 零向量的表法唯一，即若 $0=\alpha_1+\alpha_2+\cdots+\alpha_k\quad(\alpha_i\in M_i,i=1,2,\cdots,k)$, 则 $\alpha_1=\alpha_2=\cdots=\alpha_k=0$;
3. $M_i\cap\left(\sum_{\substack{j=1\\j\neq i}}^{k}M_j\right)=\{0\}\quad(i=1,2,\cdots,k)$;
4. $\dim\sum_{i=1}^{k}M_i=\sum_{i=1}^{k}\dim M_i$.

证明. 我们只证明 $(iv)\Rightarrow(i)$，$(iii)\Rightarrow(iv)$ 可以使用数学归纳法.

(iv) ⇒ (i) 先证 $M_i\cap\left(\sum_{\substack{j=1\\j\neq i}}^{k}M_j\right)=\{0\}\quad(i=1,2,\cdots,k)$.

对任一 i，按维数公式及其推论，有

$$0\leqslant\dim\left(M_i\cap\left(\sum_{\substack{j=1\\j\neq i}}^{k}M_j\right)\right)=\dim M_i+\dim\sum_{\substack{j=1\\j\neq i}}^{k}M_j-\dim\sum_{j=1}^{k}M_j. \tag{1}$$

$$\leqslant\dim M_i+\sum_{\substack{j=1\\j\neq i}}^{k}\dim M_j-\dim\sum_{j=1}^{k}M_j. \tag{2}$$

$$=\sum_{j=1}^{k}\dim M_j-\dim\sum_{j=1}^{k}M_j. \tag{3}$$

$$\tag{4}$$

由假设，$\dim \sum_{j=1}^{k} M_j = \sum_{j=1}^{k} \dim M_j$，故

$$\dim\left(M_i \cap \left(\sum_{\substack{j=1\\ j\neq i}}^{k} M_j\right)\right) = 0,$$

即

$$M_i \cap \left(\sum_{\substack{j=1\\ j\neq i}}^{k} M_j\right) = \{0\}.$$

现设 $\alpha \in \sum_{i=1}^{k} M_i$，证明 α 的表法唯一。命 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k = \alpha_1' + \alpha_2' + \cdots + \alpha_k'$, 其中 $\alpha_i, \alpha_i' \in M_i (i = 1, 2, \cdots, k)$。对任一 i，有 $\alpha_i - \alpha_i' = (\alpha_1' - \alpha_1) + \cdots + (\alpha_{i-1}' - \alpha_{i-1})$ $+(\alpha_{i+1}' - \alpha_{i+1}) + \cdots + (\alpha_k' - \alpha_k) \in \sum_{\substack{j=1\\ j\neq i}}^{k} M_j$. 于是 $\alpha_i - \alpha_i' \in M_i \cap \left(\sum_{\substack{j=1\\ j\neq i}}^{k} M_j\right) = \{0\}$, 因此 $\alpha_i = \alpha_i'$，α 的表法唯一，从而 M 是直和. □

命题 10 设 M 是数域 K 上有限维线性空间 V 的一个子空间，则必存在 V 的子空间 N，使

$$V = M \oplus N.$$

证明. 若 $M = \{0\}$，则取 $N = V$。下面设 $M \neq \{0\}$。在 M 内取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$，则 $M = L(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r)$。另一方面，按命题 8，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 可扩充成 V 的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \cdots, \varepsilon_n$。令 $N = L(\varepsilon_{r+1}, \cdots, \varepsilon_n)$，则显见有 $V = M + N$。又因为 $\dim V = n = r + (n-r) = \dim M + \dim N$，故：$V = M \oplus N$。 □

定义 6 上述命题指出的子空间 N 称为子空间 M 的一个**补空间**。一个子空间的补空间不是唯一的，因为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 可以有多种方式（实际上有无穷多种方式）扩充成空间的一组基。

商空间

设 α 是 V 中任意一个向量，定义 V 的子集 $\alpha + M = \{\alpha + m \mid m \in M\}$. 我们称 $\alpha + M$ 为一个模 M 的同余类或称为模 M 的剩余类，而 α 称为这个同余类的一个代表。

关于模 M 的同余类，有如下简单的性质：

1. $\alpha' \equiv \alpha (\mod M) \Longleftrightarrow \alpha' - \alpha \in M \Longleftrightarrow \alpha' \in \alpha + M$;
2. $\alpha' \in \alpha + M \Longleftrightarrow \alpha' + M = \alpha + M$;
3. $\alpha + M = 0 + M \Longleftrightarrow \alpha \in M$;
4. 若 $\alpha' + M \neq \alpha + M$，则 $(\alpha' + M) \cap (\alpha + M) = \emptyset$.

命集合 $\overline{V}$ 表示 V 内向量模 M 的同余类的全体构成的集合，定义：

- $(\alpha + M) + (\beta + M) = (\alpha + \beta) + M$
- 对任意 $k \in K$, 定义：

$$k(\alpha + M) = k\alpha + M$$

不难验证这个定义关于 K 上的加法和数乘确实构成了 K 上的一个线性空间，这个线性空间被称为 V 对子空间 M 的**商空间**，记作 V/M.

命题 11 设 V 是数域 K 上的 n 维线性空间，M 是 V 的一个 m 维子空间，则 $\dim V/M = n-m$。

证明. 在 M 内取一组基 $\varepsilon_1,\cdots,\varepsilon_m$，扩充为 V 的一组基:$\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$。我们来证明 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 是 V/M 的一组基。

(i) 证 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 线性无关。设有 $k_{m+1}\overline{\varepsilon}_{m+1}+\cdots+k_n\overline{\varepsilon}_n=\overline{0}$, 那么 $k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n\in M$.

由此推知 $k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n=k_1\varepsilon_1+\cdots+k_m\varepsilon_m$.

但 $\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$ 线性无关，由上式立即推出 $k_{m+1}=\cdots=k_n=0$。这表明 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 在 V/M 内线性无关。

(ii) 对任意 $\overline{\alpha}\in V/M$，证明 $\overline{\alpha}$ 可被 $\overline{\varepsilon}_{m+1},\overline{\varepsilon}_{m+2},\cdots,\overline{\varepsilon}_n$ 线性表示。设 $\overline{\alpha}=a+M$，有

$$\alpha=k_1\varepsilon_1+\cdots+k_m\varepsilon_m+k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n.$$

于是

$$\begin{aligned}\overline{\alpha}&=\overline{k_1\varepsilon_1+\cdots+k_m\varepsilon_m+k_{m+1}\varepsilon_{m+1}+\cdots+k_n\varepsilon_n} &(5)\\&=k_1\overline{\varepsilon}_1+\cdots+k_m\overline{\varepsilon}_m+k_{m+1}\overline{\varepsilon}_{m+1}+\cdots+k_n\overline{\varepsilon}_n &(6)\\&=k_{m+1}\overline{\varepsilon}_{m+1}+\cdots+k_n\overline{\varepsilon}_n, &(7)\end{aligned}$$

上面用到了 $\varepsilon_i\in M(i=1,2,\cdots,m)$，故 $\overline{\varepsilon}_i=\overline{0}$。

上面推理指明 $\overline{\varepsilon}_{m+1},\cdots,\overline{\varepsilon}_n$ 是 V/M 的一组基，命题得证。 □

线性映射和线性变换

定义 7 设 U,V 为数域 K 上的两个线性空间，f 为 U 到 V 的一个映射，且满足如下条件:

1. 对任意 $\alpha,\beta\in U$，有
$$f(\alpha+\beta)=f(\alpha)+f(\beta);$$
2. 对任意 $\alpha\in U,k\in K$，有
$$f(k\alpha)=kf(\alpha),$$

则称 f 为 U 到 V 的一个**线性映射**。U 到 V 的全体线性映射所成的集合记做 $\mathrm{Hom}(U,V)$。

根据上述不难推出，对任意 $\alpha_1,\alpha_2,\cdots,\alpha_l\in U,k_1,k_2,\cdots,k_l\in K$，有:

$$f(k_1\alpha_1+k_2\alpha_2+\cdots+k_l\alpha_l)=k_1f(\alpha_1)+k_2f(\alpha_2)+\cdots+k_lf(\alpha_l)$$

以及: $f(0)=0,f(-\alpha)=-f(\alpha)$

【例】 设 V 是数域 K 上的线性空间，M 是它的一个子空间。定义映射:

$$\varphi:V\longrightarrow V/M,$$

$$\alpha\longmapsto\alpha+M,$$

则有

$$\varphi(\alpha+\beta)=(\alpha+\beta)+M=(\alpha+M)+(\beta+M)=\varphi(\alpha)+\varphi(\beta);$$

$$\varphi(k\alpha)=k\alpha+M=k(\alpha+M)=k\varphi(\alpha).$$

于是 φ 是 V 到商空间 V/M 的线性映射。这个映射称为**自然映射**。显然自然映射是满射，但不是单射（当 $M\neq\{0\}$ 时）。因为取 $\alpha,\beta\in M$, $\alpha\neq\beta$, 此时

$$\varphi(\alpha)=\alpha+M=0+M=\beta+M=\varphi(\beta).$$

另外，从 φ 的定义可知：$\varphi(\alpha)=\overline{0}$ 的充分必要条件是 $\alpha\in M$。

命题 12 设 U,V 是数域 K 上的线性空间，f 是 U 到 V 的线性映射，且为单射。则 U 内向量组 $\alpha_1,\alpha_2,\cdots,\alpha_l$ 线性无关的充分必要条件是 $f(\alpha_1),f(\alpha_2),\cdots,f(\alpha_l)$ 在 V 内线性无关。

命题 13 设 U,V 是数域 K 上的线性空间，$\dim U=n$. f 是 U 到 V 的线性映射。如果 f 是双射，则对 U 的任一组基 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$, $f(\epsilon_1),f(\epsilon_2),\cdots,f(\epsilon_n)$ 为 V 的一组基，从而 $\dim V=n$。

定义 8 设 U 与 V 是数域 K 上的线性空间，如果存在 U 到 V 的线性映射 f 同时又是双射，则称 U 与 V **同构**，而映射 f 称为 U 到 V 的**同构映射**.

命题 14 设 U，V 是数域 K 上的线性空间，f 是 U 到 V 的同构映射，则 f^{-1} 是 V 到 U 的同构映射（只需证明 f^{-1} 是线性映射即可）.

不难看出，对于数域 K 上的 n 维线性空间 U 和 V，它们都与 K^n 同构，从而由同构的传递性知 U 和 V 同构.

定义 9 设 U,V 是数域 K 上的线性空间，f 是 U 到 V 的线性映射。定义 $\operatorname{Ker} f=\{a\in U\mid f(a)=0\}$, 称为线性映射 f 的**核**。又定义 $\operatorname{Im} f=\{f(a)\mid a\in U\}$, 称为线性映射 f 的**像集**。

命题 15 设 U,V 是数域 K 上的线性空间，$f\in\operatorname{Hom}(U,V)$。则有：

1. $\operatorname{Ker} f$ 是 U 的子空间，f 是单射的充分必要条件是 $\operatorname{Ker} f=\{0\}$；
2. $\operatorname{Im} f$ 是 V 的子空间，定义 $\operatorname{Coker} f=V/\operatorname{Im} f$。$f$ 是满射的充分必要条件是

$$\operatorname{Coker} f=\{0\}.$$

   $\operatorname{Coker} f$ 称为线性映射 f 的**余核**。

命题 16 自然映射 $\varphi(V)=V/M$，有 $Ker\,\varphi=M, Im\,\varphi=V/M$.

命题 17 设 U,V 是数域 K 上的线性空间，$f\in\operatorname{Hom}(U,V)$，则 $U/\operatorname{Ker} f$ 与 $\operatorname{Im} f$ 同构。

证明. 定义 $U/\operatorname{Ker} f$ 到 $\operatorname{Im} f$ 的映射 $\tau(a+\operatorname{Ker} f)=f(a)\in\operatorname{Im} f\quad(\forall a\in U)$.（或者写成 $\tau(\overline{a})=f(a)$。）首先要验证上面的定义在逻辑上无矛盾（因为 $U/\operatorname{Ker} f$ 中元素的表示法不唯一）。设有 $a+\operatorname{Ker} f=\beta+\operatorname{Ker} f$, 则 $\beta=a+m(m\in\operatorname{Ker} f)$. 于是 $f(\beta)=f(a)+f(m)=f(a)$, 这表明 $\tau(a+\operatorname{Ker} f)=f(a)=f(\beta)=\tau(\beta+\operatorname{Ker} f)$.

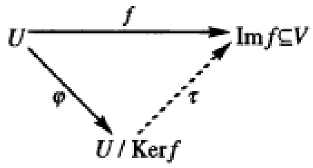
(i) 证明 τ 是线性映射。对 $\overline{\alpha},\overline{\beta}\in U/\operatorname{Ker}f$, 有

$$\begin{aligned}\tau(k\overline{\alpha}+l\overline{\beta})&=\tau(\overline{k\alpha+l\beta})=f(k\alpha+l\beta)\\&=kf(\alpha)+lf(\beta)\\&=k\tau(\alpha+\operatorname{Ker}f)+l\tau(\beta+\operatorname{Ker}f)\\&=k\tau(\overline{\alpha})+l\tau(\overline{\beta}).\end{aligned}$$

(ii) 证明 τ 是单射。设 $\tau(a+\operatorname{Ker}f)=\tau(\beta+\operatorname{Ker}f)$, 于是 $f(a)=f(\beta)$, 即 $f(a-\beta)=f(a)-f(\beta)=0$, 由此知 $a-\beta\in\operatorname{Ker}f$, 于是 $a+\operatorname{Ker}f=\beta+\operatorname{Ker}f$.

(iii) τ 显然为满射，因为 $\operatorname{Im}f=\{f(a)\mid a\in U\}$.

综上所述知 τ 为 $U/\operatorname{Ker}f$ 到 $\operatorname{Im}f$ 的同构映射。 □



根据上述命题可以得到以下推论：

推论 1 设 U,V 是数域 K 上的两个线性空间，$dimU=n,f\in Hom(U,V),$, 则 $dimKerf+dimImf=dimU$.（利用命题 17 同构和商空间的基的构造证明）

推论 2 给定数域 K 上的 $m\times n$ 矩阵 A，设 $r(A)=r$，则齐次线性方程组 $AX=0$ 的基础解系中含 $n-r$ 个向量.

线性映射的运算

设 U,V 是数域 K 上的线性空间.

1) 给定 $f,g\in\mathrm{Hom}(U,V)$，定义 $(f+g)(a)=f(a)+g(a)\quad(\forall a\in U)$,

则 $f+g\in\mathrm{Hom}(U,V)$. 这是因为

$$\begin{aligned}(f+g)(ka+l\beta)&=f(ka+l\beta)+g(ka+l\beta)\\&=kf(\alpha)+lf(\beta)+kg(\alpha)+lg(\beta)\\&=k(f(\alpha)+g(\alpha))+l(f(\beta)+g(\beta))\\&=k(f+g)(\alpha)+l(f+g)(\beta).\end{aligned}$$

$f+g$ 称为 f 与 g 的**加法**.

2) 给定 $f\in\mathrm{Hom}(U,V)$ 及 $k\in K$，定义 $(kf)(a)=kf(a)\quad(\forall a\in U)$,

则 $kf\in\mathrm{Hom}(U,V)$. 这是因为

$$\begin{aligned}(kf)(k_1a+l\beta)&=kf(k_1a+l\beta)\\&=k(k_1f(a)+lf(\beta))\\&=k_1kf(a)+lkf(\beta)\\&=k_1(kf)(a)+l(kf)(\beta).\end{aligned}$$

kf 称为 k 与 f 的**数乘**.

容易验证，在 $\mathrm{Hom}(U,V)$ 内定义的上述加法、数乘满足线性空间定义中的八条运算法则，其中零元素为如下 U 到 V 的**零映射**:

$$\mathbf{0}(a)=0 \quad (\forall a\in U).$$

显然 $\mathbf{0}\in\mathrm{Hom}(U,V)$，且对任意 $f\in\mathrm{Hom}(U,V)$，有 $f+\mathbf{0}=f$. 又对任意 $f\in\mathrm{Hom}(U,V)$，定义 $(-f)(a)=-f(a) \quad (\forall a\in U)$. 显然 $-f\in\mathrm{Hom}(U,V)$，且 $f+(-f)=\mathbf{0}$.

因此，$\mathrm{Hom}(U,V)$ 关于上述加法、数乘也组成 K 上的线性空间.

3) 设 U,V,W 都是数域 K 上的线性空间. 如果 $f\in\mathrm{Hom}(U,V)$, $g\in\mathrm{Hom}(V,W)$，定义 $(gf)(a)=g(f(a)) \quad (\forall a\in U)$, 则显然 $gf\in\mathrm{Hom}(U,W)$. 这是因为

$$\begin{aligned}(gf)(ka+l\beta)&=g(f(ka+l\beta))\\&=g(kf(a)+lf(\beta))\\&=kg(f(a))+lg(f(\beta))\\&=k(gf)(a)+l(gf)(\beta).\end{aligned}$$

gf 称为 g 与 f 的乘法. g 与 f 只有满足上述条件时才能相乘.

线性映射的乘法满足如下运算法则（在下面总假定出现的乘法是有意义的）.

1. 乘法结合律: $(fg)h=f(gh)$.
2. 加法与乘法分配律:
$$f(g+h)=fg+fh;$$
$$(f+g)h=fh+gh.$$
3. 对任意 $k\in K$, $k(fg)=(kf)g=f(kg)$.

这里同样要注意两点:

1. 线性映射的乘法一般不可交换次序;
2. 线性映射的乘法没有消去律，即由等式 $fg=fh, f\neq 0$ 并不能推出 $g=h$；同样，由等式 $gf=hf, f\neq 0$ 也不能推出 $g=h$。

线性映射的矩阵

现在设 U, V 是数域 K 上的线性空间，且设 $\dim U=n$, $\dim V=m$. 设 $f\in\mathrm{Hom}(U,V)$. 为了把 f 具体化，我们在 U 内取定一组基 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$, 在 V 内也取定一组基 $\eta_1,\eta_2,\cdots,\eta_m$. 我们先证明一个基本命题.

命题 18 记号如上述。我们有如下结论:

1. U 到 V 的任一线性映射 f 由它在 U 的基 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$ 处的作用唯一决定，就是说，如果又有 $g\in\mathrm{Hom}(U,V)$ 使 $g(\epsilon_i)=f(\epsilon_i)(i=1,2,\cdots,n)$，则 $g(a)=f(a) \quad (\forall a\in U)$;
2. 任给 V 内 n 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 必存在唯一的 $f\in\mathrm{Hom}(U,V)$, 使 $f(\epsilon_i)=\alpha_i(i=1,2,\cdots,n)$。

证明. (i) 对 U 内任一向量 a, 设 $a=x_1\epsilon_1+x_2\epsilon_2+\cdots+x_n\epsilon_n$, 则 $f(a)=x_1f(\epsilon_1)+x_2f(\epsilon_2)+\cdots+x_nf(\epsilon_n)$ $=x_1g(\epsilon_1)+x_2g(\epsilon_2)+\cdots+x_ng(\epsilon_n)=g(a)$.

(ii) 定义 U 到 V 的映射 f 如下：若 $a = x_1\epsilon_1 + x_2\epsilon_2 + \cdots + x_n\epsilon_n$，则令 $f(a) = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$。现在来证 $f \in \mathrm{Hom}(U,V)$。若又有 $\beta = y_1\epsilon_1 + y_2\epsilon_2 + \cdots + y_n\epsilon_n$，则 $f(ka+l\beta) = f((kx_1+ly_1)\epsilon_1+(kx_2+ly_2)\epsilon_2+\cdots+(kx_n+ly_n)\epsilon_n) = (kx_1+ly_1)\alpha_1+(kx_2+ly_2)\alpha_2+\cdots+(kx_n+ly_n)\alpha_n$ $= k(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n)+l(y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n) = kf(a)+lf(\beta)$. 现在显然有 $f(\epsilon_i) = \alpha_i$。满足此条件的线性映射的唯一性由 (1) 立即推出。 □

现在定义 $\mathrm{Hom}(U,V)$ 到 $M_{m,n}(K)$ 的一个映射 σ 如下：对 $f \in \mathrm{Hom}(U,V)$，如果 f 在 U, V 所取定的基下的矩阵为 A，则令 $\sigma(f) = A$. 那么上面命题 19 的 (i) 说明 σ 为单射（不同线性映射的矩阵是不同的），而 (ii) 则说明 σ 为满射（任一 $m \times n$ 矩阵都是某线性映射的矩阵）。于是集合 $\mathrm{Hom}(U,V)$ 和集合 $M_{m,n}(K)$ 之间存在一一对应。不但如此，这个双射 σ 还保持 $\mathrm{Hom}(U,V)$ 的运算和 $M_{m,n}(K)$ 内运算的对应关系。为了证明这一点，我们先来引进一个记号。

设 $f \in \mathrm{Hom}(U,V)$,对 U 内任意向量组 $a_1,a_2,\cdots,a_k$,我们约定 $f(a_1,a_2,\cdots,a_k) = (f(a_1),f(a_2),\cdots,f(a_k))$.

使用这个记号将使我们下面的推理简化。这是因为这一记号具有如下一个性质：令 $a \in U$,则 $a = x_1\epsilon_1+x_2\epsilon_2+\cdots+x_n\epsilon_n = (\epsilon_1,\epsilon_2,\cdots,\epsilon_n)X$, $f(a) = x_1f(\epsilon_1)+x_2f(\epsilon_2)+\cdots+x_nf(\epsilon_n) = (f(\epsilon_1),f(\epsilon_2),\cdots,f(\epsilon_n))X$ $= [f(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)]X$.

把上面两个式子合并，可以写成 $f(a) = f[(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)X] = [f(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)]X$.

上面式子右边的等式形式上与结合律一致。

一般地,设 $a_1 = (\epsilon_1,\cdots,\epsilon_n)X_1,\cdots,a_s = (\epsilon_1,\cdots,\epsilon_n)X_s$, 我们有 $f(a_1) = (f(\epsilon_1),\cdots,f(\epsilon_n))X_1,\cdots,f(a_s) = (f(\epsilon_1),\cdots,f(\epsilon_n))X_s$.

以 $X_1,\cdots,X_s$ 为列向量排成 $n \times s$ 矩阵 X，则有 $(a_1,\cdots,a_s) = (\epsilon_1,\cdots,\epsilon_n)X$.

而 $(f(a_1),\cdots,f(a_s)) = (f(\epsilon_1),\cdots,f(\epsilon_n))X$.

按上面的约定，有 $f(a_1,\cdots,a_s) = (f(a_1),\cdots,f(a_s))$.

把上面式子合并，即写成 $f[(\epsilon_1,\cdots,\epsilon_n)X] = (f(\epsilon_1),\cdots,f(\epsilon_n))X = [f(\epsilon_1,\cdots,\epsilon_n)]X$.

这是更一般的形式结合律。下面的推理中，我们将经常使用上面约定的记号及其形式结合律。

命题 20 记号如上所述。我们有如下结论：

1. 对 $f,g \in \mathrm{Hom}(U,V)$, $k,l \in K$, 有

$$\sigma(kf+lg) = k\sigma(f)+l\sigma(g);$$

   因此，σ 是 $\mathrm{Hom}(U,V)$ 到 $M_{m,n}(K)$ 的线性空间同构。于是我们有

$$\dim \mathrm{Hom}(U,V) = \dim M_{m,n}(K) = mn.$$

2. 若 $f \in \mathrm{Hom}(U,V)$, $g \in \mathrm{Hom}(V,W)$, 则 $\sigma(gf) = \sigma(g)\sigma(f)$.

证明. (i) 设 $\sigma(f)=A$, $\sigma(g)=B$, 则

$$
\begin{aligned}
&(kf+lg)(\epsilon_1),(kf+lg)(\epsilon_2),\cdots,(kf+lg)(\epsilon_n)\\
&=k(f(\epsilon_1),f(\epsilon_2),\cdots,f(\epsilon_n))+l(g(\epsilon_1),g(\epsilon_2),\cdots,g(\epsilon_n))\\
&=k(\eta_1,\eta_2,\cdots,\eta_m)A+l(\eta_1,\eta_2,\cdots,\eta_m)B\\
&=(\eta_1,\eta_2,\cdots,\eta_m)(kA+lB).
\end{aligned}
$$

上式表明线性映射 $kf+lg$ 在所取定基下的矩阵为 $kA+lB=k\sigma(f)+l\sigma(g)$，亦即 $\sigma(kf+lg)=k\sigma(f)+l\sigma(g)$.

(ii) 设在 U,V,W 内分别取一组基

$$\epsilon_1,\cdots,\epsilon_n;\quad \eta_1,\cdots,\eta_m;\quad \delta_1,\cdots,\delta_s.$$

又设

$$(f(\epsilon_1),f(\epsilon_2),\cdots,f(\epsilon_n))=(\eta_1,\eta_2,\cdots,\eta_m)A,$$

$$(g(\eta_1),g(\eta_2),\cdots,g(\eta_m))=(\delta_1,\delta_2,\cdots,\delta_s)B.$$

那么，我们有（使用上面约定记号）

$$
\begin{aligned}
&(gf(\epsilon_1),gf(\epsilon_2),\cdots,gf(\epsilon_n))\\
&=g(f(\epsilon_1),f(\epsilon_2),\cdots,f(\epsilon_n))\\
&=g[(\eta_1,\eta_2,\cdots,\eta_m)A]\\
&=[g(\eta_1,\eta_2,\cdots,\eta_m)]A\\
&=[(\delta_1,\delta_2,\cdots,\delta_s)B]A\\
&=(\delta_1,\delta_2,\cdots,\delta_s)(BA).
\end{aligned}
$$

上面的结果表示 $gf\in\mathrm{Hom}(U,W)$ 在所取定的基下的矩阵为 $BA=\sigma(g)\sigma(f)$，于是 $\sigma(gf)=\sigma(g)\sigma(f)$. □

线性变换的基本概念

定义 10 设 V 是数域 K 上的线性空间，A 是 V 到自身的一个线性映射，则称 A 为 V 内的一个**线性变换**（注意不一定是满射）,V 内全体线性变换所成的集合记为 $End(V)$。以下是几种特殊的线性变换：

1) 零变换 $\mathbf{0}$. 对任意 $a\in V$, $\mathbf{0}a=0$. 这是 V 内的一个线性变换，称为零变换. 对任意 $A\in\mathrm{End}(V)$, 有 $\mathbf{0}A=A\mathbf{0}=\mathbf{0}$.

2) 单位变换 $\boldsymbol{E}$. 对任意 $a\in V$, $\boldsymbol{E}a=a$. 这显然也是 V 内的一个线性变换，称为单位变换或恒等变换. 对任意 $A\in\mathrm{End}(V)$, $\boldsymbol{E}A=A\boldsymbol{E}=A$.

3) 数乘变换 $\boldsymbol{k}$. 设 k 是数域 K 内一个固定的数. 对任意 $a\in V$, 定义: $\boldsymbol{k}a=ka$. 不难验证，这也是 V 内的一个线性变换，称为数乘变换.

4) 投影变换 $\boldsymbol{P}$. 设 M 是 V 的一个子空间. 按命题 10，存在 V 的子空间 N, 使 $V=M\oplus N$. 于是，对任意 $a\in V$, 有唯一分解式 $a=a_1+a_2\quad(a_1\in M,\ a_2\in N)$. 我们定义 V 内一个变换 $\boldsymbol{P}$ 如下: $\boldsymbol{P}a=a_1$.

我们证明 P 是一个线性变换.

(i) 设 $\alpha,\beta\in V$, 又 $\alpha=\alpha_1+\alpha_2,\quad\beta=\beta_1+\beta_2\quad(\alpha_1,\beta_1\in M;\alpha_2,\beta_2\in N)$. 则 $\alpha+\beta=(\alpha_1+\beta_1)+(\alpha_2+\beta_2)\quad(\alpha_1+\beta_1\in M;\alpha_2+\beta_2\in N)$. 按 P 的定义，有 $P(\alpha+\beta)=\alpha_1+\beta_1=P\alpha+P\beta$.

(ii) 对任意 $k\in K$, 有 $k\alpha=k\alpha_1+k\alpha_2\quad(k\alpha_1\in M;k\alpha_2\in N)$, $P(k\alpha)=k\alpha_1=kP\alpha$.

线性变换 P 称为 V 对子空间 M（关于直和分解式 $V=M\oplus N$）的投影变换。注意投影变换依赖于直和分解式的具体形式。

线性变化 P 称为 V 对子空间 M（关于直和分解式 $V=M\oplus N$）的**投影变换**，投影变换依赖于直和分解式的具体形式.

定义 $End(V)$ 到 $M_n(K)$ 的映射 σ：对 $A\in End(V)$，$\sigma(A)$ 为 A 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵，于是 σ 是一个双射（一一对应），而且有如下性质：

1) $\sigma(kA+lB)=k\sigma(A)+l\sigma(B)$**;**

2) $\sigma(AB)=\sigma(A)\sigma(B)$**.**

因为单位变换 E 在 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$ 下的矩阵为 E, 我们有 $E=AB=BA\iff E=\sigma(A)\sigma(B)=\sigma(B)\sigma(A)$.

所以, A 为可逆线性变换的充分必要条件是 $\sigma(A)$ 为 K 上可逆 n 阶方阵, 而且 $\sigma(A^{-1})=\sigma(A)^{-1}$ (其中用到了 σ 为单、满映射这一条件).(并且注意上述的 A，B 等指的是线性变换并非其对应的矩阵，后文中 A,B 也会用相同的字母表示这两种不同的意义，一般地，写成 $\vec{A}(\varepsilon_1\cdots,\varepsilon_n)=(\varepsilon_1,\cdots,\varepsilon_n)A$，前者表示线性变换而后者表示矩阵)

命题 21 设线性变换 A 在一组基下的矩阵为 A, 又设向量 α 在这组基下的坐标为

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

则 $A\alpha$ 在这组基下的坐标为 AX.

证明. 设这组基为 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$. 采用形式的写法, 有 $A\alpha=A(x_1\epsilon_1+x_2\epsilon_2+\cdots+x_n\epsilon_n)=[A(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)]X$ $=[(A(\epsilon_1,\epsilon_2,\cdots,\epsilon_n))A]X=(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)(AX)$. □

命题 22 设

$$\epsilon_1,\epsilon_2,\cdots,\epsilon_n;$$

$$\eta_1,\eta_2,\cdots,\eta_n$$

是线性空间 V 的两组基，其过渡矩阵是 $T=(t_{ij})$，即

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\epsilon_1,\epsilon_2,\cdots,\epsilon_n)T.\tag{1}$$

又设线性变换 $\mathcal{A}$ 在这两组基下的矩阵分别是 A 和 B，则 $B = T^{-1}AT$.

证明. 由线性变换的矩阵的定义，有

$$\mathcal{A}(\epsilon_1, \epsilon_2, \cdots, \epsilon_n) = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)A,$$

$$\mathcal{A}(\eta_1, \eta_2, \cdots, \eta_n) = (\eta_1, \eta_2, \cdots, \eta_n)B.$$

把 (1) 式代入上面的第二个等式，得

$$\mathcal{A}[(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)T] = [(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)T]B.$$

利用形式运算的结合律，有

$$\begin{aligned} \mathcal{A}[(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)T] &= [(\mathcal{A}(\epsilon_1, \epsilon_2, \cdots, \epsilon_n))T] && (8)\\ &= [(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)A]T && (9)\\ &= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)(AT) && (10)\\ &= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)(TB). && (11) \end{aligned}$$

因为 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 线性无关，故有 $AT = TB$.

而 T 可逆，因而有 $B = T^{-1}AT$. □

定义 11 对数域 K 上的两个 n 阶方阵 A 与 B，如果存在 K 上一个 n 阶可逆的方阵 T，使 $B = T^{-1}AT$，则称 B 与 A 在 K 内 **相似**，记做 $B \sim A$。

矩阵的相似关系具有如下性质：

1. 反身性：$A \sim A$。这是因为 $A = E^{-1}AE$;
2. 对称性：若 $B \sim A$，则 $A \sim B$。这是因为当 $B = T^{-1}AT$ 时，有 $A = (T^{-1})^{-1}BT^{-1}$;
3. 传递性：若 $A \sim B, B \sim C$，则 $A \sim C$。这是因为当 $A = T_1^{-1}BT_1$，$B = T_2^{-1}CT_2$ 时，有

$$A = T_1^{-1}(T_2^{-1}CT_2)T_1 = (T_2T_1)^{-1}C(T_2T_1).$$

这说明矩阵的相似关系是一个等价关系。我们把数域 K 上全体 n 阶方阵的集合 $M_n(K)$ 在相似关系下的等价类叫做 **相似类**。于是 $M_n(K)$ 可分解为互不相交的相似类的并。

由于上述性质，我们可以把 K 上 n 阶方阵的集合 $M_n(K)$ 中的元素按相似关系进行分类，凡是相互之间存在相似关系的矩阵属于同一类，不同的相似类之间没有公共元素（交是空集）。下面一个命题阐明了相似的实际意义。

命题 23 数域 K 上两个 n 阶方阵 A, B 相似的充分必要条件是，它们是 V 内某一线性变换在两组基下的矩阵。

线性变换的特征值和特征向量

定义 12 设 V 是数域 K 上的一个线性空间，$\mathcal{A}$ 是 V 内一个线性变换。如果对 K 内一个数 λ，存在 V 的一个向量 $\xi \neq 0$，使 $\mathcal{A}\xi = \lambda\xi$. 则称 λ 为 $\mathcal{A}$ 的一个 **特征值**，而 ξ 称为属于特征值 λ 的 **特征向量**。

这里要注意两点：

1. 特征向量 ξ 一定要是非零向量；
2. λ 必须属于数域 K，否则数乘 $\lambda\xi$ 没有意义。

对于数域 K 内任一数 λ, 我们定义 $V_\lambda = \{a \in V \mid Aa = \lambda a\}$.

容易看出，由于 $A0 = \lambda \cdot 0$, 故 $0 \in V_\lambda$, 即 V_λ 非空。如果 $a, \beta \in V_\lambda, k, l \in K$, 则 $Aa = \lambda a, A\beta = \lambda\beta$, 故 $A(ka + l\beta) = kAa + lA\beta = k\lambda a + l\lambda\beta = \lambda(ka + l\beta)$, 即 $ka + l\beta \in V_\lambda$. 于是 V_λ 关于加法、数乘封闭，即 V_λ 为 V 的子空间。

我们有下列两个明显的事实：

1. $\lambda \in K$ 是线性变换 A 的特征值的充分必要条件是 $V_\lambda \neq \{0\}$. 此时 V_λ 称为 A 的属于特征值 λ 的**特征子空间**. V_λ 中任何非零向量都是 A 的属于特征值 λ 的特征向量；
2. 要找出 A 的属于特征值 λ 的全部特征向量，只要决定出特征子空间 V_λ, 特别地，当 V_λ 是有限维子空间时，只要找出它的一组基，就等于找出 V_λ 中的所有向量。

以上两个问题都需要通过计算来实现。我们知道，在线性空间和线性变换的问题中，凡是需要具体计算时，就必须在 V 内取定一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，求出每个向量在此组基下的坐标：$a = x_1\epsilon_1 + x_2\epsilon_2 + \cdots + x_n\epsilon_n = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)X$, 再求出线性变换在此组基下的矩阵 $(A\epsilon_1, A\epsilon_2, \cdots, A\epsilon_n) = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)A$.

按前述命题有 $Aa = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)(AX)$, 而 $\lambda a = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)(\lambda X)$.

于是 V 内的定义式 $Aa = \lambda a$ 现在等价于 K^n 内的关系式 $AX = \lambda X$，而 $a \neq 0$ 等价于 $X \neq 0$。

现在我们有如下基本关系：

1. $V_\lambda \neq \{0\} \iff$ 有 $0 \neq a \in V$ 使 $Aa = \lambda a \iff$ 有 $0 \neq X \in K^n$ 使 $AX = \lambda X$;
2. $0 \neq a \in V$，满足 $Aa = \lambda a \iff 0 \neq X \in K^n$，满足等式 $AX = \lambda X$.

现设

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

由 $AX = \lambda X$ 推知 $\lambda X - AX = 0$，即 $(\lambda E - A)X = 0$。具体写出来，就是

$$\begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

上式是数域 K 内 n 个未知量 n 个方程的齐次线性方程组，它有非零解的充分必要条件是

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

于是，对上面提出的两个基本问题，答案如下：

1. $\lambda \in K$ 为 A 的特征值，即 $V_\lambda \neq \{0\}$ 的充分必要条件是它满足

$$|\lambda E - A| = 0;$$

2. $a \in V$ 满足 $Aa = \lambda a$，即 $a \in V_\lambda$ 的充分必要条件是 a 在 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的坐标 X 满足齐次线性方程组 $(\lambda E - A)X = 0$，这也说明了可以通过 K^n 到 V 的同构映射，通过 X 的解空间的基得到特征向量空间 V_λ 的一组基.

给定数域 K 上的 n 阶方阵 $A = (a_{ij})$，令

$$f(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix},$$

从行列式的完全展开式易知 $f(\lambda)$ 为 λ 的多项式，其系数属于数域 K，$f(\lambda)$ 称为方阵 A 的 **特征多项式**。$f(\lambda)$ 属于数域 K 的根称为方阵 A 的 **特征根**或 **特征值**。因为数域 K 上的方阵都可看做复数域上的方阵，在这样看的时候，$f(\lambda)$ 在复数域内的全部根都认为是 A 的特征根或特征值。所以，说到矩阵 A 的特征值时，不能忘记究竟是把该矩阵看做哪个数域上的方阵。当 λ 是 A 的特征值时，齐次线性方程组 $(\lambda E - A)X = 0$ 的每个非零解 X_0 就称为 A 的属于特征值 λ 的 **特征向量**。注意此时 λ 与 X_0 属于同一个数域。

特征值和特征向量的计算法

现在我们把计算线性变换 A 的特征值和特征向量的步骤归纳如下：

1. 在 V 中给定一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，求 A 在这组基下的矩阵 A.
2. 计算特征多项式 $f(\lambda) = |\lambda E - A|$.
3. 求 $f(\lambda) = 0$ 的属于数域 K 的那些根 $\lambda_1, \lambda_2, \cdots, \lambda_s$.
4. 对每个 $\lambda_i (i = 1, 2, \cdots, s)$ 求齐次线性方程组 $(\lambda E - A)X = 0$ 的一个基础解系。这个齐次线性方程组具体写出来就是

$$\begin{bmatrix} \lambda_i - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_i - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_i - a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

   注意其中的 λ_i 是在步骤 3 中求出的，是已知数，不是未知量。
5. 以步骤 4 中求出的基础解系为坐标写出 V 中一个向量组，它就是 V_{λ_i} 的一组基。

命题 24 相似的矩阵有相同的特征多项式。

证明. 设 $B = T^{-1}AT$，则

$$|\lambda E - B| = |\lambda E - T^{-1}AT| = |T^{-1}(\lambda E - A)T|$$

$$= |T^{-1}| \cdot |\lambda E - A| \cdot |T| = |\lambda E - A| \cdot |T^{-1}| \cdot |T|$$

$$= |\lambda E - A| \cdot |E| = |\lambda E - A|.$$

□

注意此命题的逆命题一般不成立。即如果两个 n 阶方阵特征多项式相同，它们未必相似。

因为一个线性变换 $\mathcal{A}$ 在不同基下的矩阵是相似的，根据上述命题，它们的特征多项式相同。因而，我们把 $\mathcal{A}$ 在任一组基下的矩阵的特征多项式称为 $\mathcal{A}$ 的 **特征多项式**。

命题 23 设 A 是数域 K 上的 n 阶方阵，则

$$f(\lambda) = |\lambda E - A| = \lambda^n - \operatorname{Tr}(A)\lambda^{n-1} + \cdots + (-1)^n |A|,$$

其中 $\operatorname{Tr}(A)$ 为 A 的主对角线元素之和，即为 A 的迹。

证明. 已知 $f(\lambda)$ 为 λ 的多项式，设 $f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_m$. 显然 $a_m = f(0) = |-A| = (-1)^n|A|$. 故只需证

$$m = n, \quad a_0 = 1, \quad a_1 = -\operatorname{Tr}(A).$$

对 n 作数学归纳法。当 $n = 1$ 时 $f(\lambda) = |\lambda - a_{11}| = \lambda - a_{11}$，显然成立。下面设对 K 上 $n-1$ 阶方阵命题成立。当 $A = (a_{ij}) \in M_n(K)$ 时，我们对 $f(\lambda)$ 求形式微商。一方面，

$$f'(\lambda) = a_0 m\lambda^{m-1} + a_1(m-1)\lambda^{m-2} + \cdots + a_{m-1}.$$

另一方面，按行列式微分公式

$$\frac{d}{d\lambda}f(\lambda) = \frac{d}{d\lambda}|\lambda E - A| = \sum_{i=1}^{n} |(\lambda E - A)_i|,$$

其中 $(\lambda E - A)_i$ 表示把 $\lambda E - A$ 的第 i 行对 λ 求形式微商。我们把 $|(\lambda E - A)_i|$ 对第 i 行展开，再利用归纳假设，有 $|(\lambda E - A)_i| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & \vdots & \cdots & -a_{1n} \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ -a_{n1} & -a_{n2} & \cdots & \vdots & \cdots & \lambda - a_{nn} \end{vmatrix} = |\lambda E_{n-1} - A\binom{i}{i}| = \lambda^{n-1} - \operatorname{Tr}(A\binom{i}{i})\lambda^{n-2} + \cdots = \lambda^{n-1} - (\operatorname{Tr}(A) - a_{ii})\lambda^{n-2} + \cdots$. 代回原式得 $\frac{df(\lambda)}{d\lambda} = \sum_{i=1}^{n}\left[\lambda^{n-1} - (\operatorname{Tr}(A) - a_{ii})\lambda^{n-2} + \cdots\right]$ $= n\lambda^{n-1} - (n-1)\operatorname{Tr}(A)\lambda^{n-2} + \cdots = f'(\lambda) = a_0 m\lambda^{m-1} + a_1(m-1)\lambda^{m-2} + \cdots$. 我们有 $m = n$ 且 $a_0 m = n, \quad -(n-1)\operatorname{Tr}(A) = a_1(m-1).$，于是 $a_0 = 1, a_1 = -\operatorname{Tr}(A)$. □

对于特征多项式在复数域的情况，假设它在复数域的 n 个根分别是 $\lambda_1, \lambda_2, \cdots, \lambda_n$，那么

$$\sum_{i=1}^{n} \lambda_i = Tr(A), \qquad \Pi_{i=1}^{n}\lambda_i = |A|$$

下面来讨论方阵特征多项式的一个有趣性质。我们知道，两个 n 阶方阵 A, B 乘积一般不可交换：$AB \neq BA$，但我们可以证明它们的特征多项式却是一样的。实际上，借助分块矩阵运算的技巧，我们可以得到更一般的结果。

例 设 A 是数域 K 上的 $n \times m$ 矩阵，B 是 K 上的 $m \times n$ 矩阵，则 $\lambda^m|\lambda E_n - AB| = \lambda^n|\lambda E_m - BA|$. 特别地，当 $m = n$ 时 $|\lambda E - AB| = |\lambda E - BA|$.

解 我们有 $\begin{bmatrix} E_m & 0 \\ -A & E_n \end{bmatrix} \begin{bmatrix} \lambda E_m & B \\ \lambda A & \lambda E_n \end{bmatrix} = \begin{bmatrix} \lambda E_m & B \\ 0 & \lambda E_n - AB \end{bmatrix}$.

两边取行列式，有 $\begin{vmatrix} \lambda E_m & B \\ \lambda A & \lambda E_n \end{vmatrix} = \begin{vmatrix} \lambda E_m & B \\ 0 & \lambda E_n - AB \end{vmatrix} = \lambda^m |\lambda E_n - AB|$.

另一方面，我们又有 $\begin{bmatrix} \lambda E_m & B \\ \lambda A & \lambda E_n \end{bmatrix} \begin{bmatrix} E_m & 0 \\ -A & E_n \end{bmatrix} = \begin{bmatrix} \lambda E_m - BA & B \\ 0 & \lambda E_n \end{bmatrix}$,

两边取行列式，得 $\begin{vmatrix} \lambda E_m & B \\ \lambda A & \lambda E_n \end{vmatrix} = \begin{vmatrix} \lambda E_m - BA & B \\ 0 & \lambda E_n \end{vmatrix} = \lambda^n |\lambda E_m - BA|$.

比较上、下两式即得所要的公式。

设 A 是 n 维线性空间 V 内的一个线性变换。如果在 V 内存在一组基 $\eta_1, \eta_2, \cdots, \eta_n$，使 A 在这组基下的矩阵成对角形，我们就说 A 的矩阵可对角化。现在我们来考查一下，什么样的线性变换其矩阵可对角化。

命题 24 数域 K 上 n 维线性空间 V 内一个线性变换 A 的矩阵可对角化的充分必要条件是，A 有 n 个线性无关的特征向量。

证明. 必要性若 A 在基 $\eta_1, \eta_2, \cdots, \eta_n$ 下的矩阵成对角形，即

$$(A\eta_1, A\eta_2, \cdots, A\eta_n) = (\eta_1, \eta_2, \cdots, \eta_n) \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

则有

$$A\eta_1 = \lambda_1 \eta_1, \quad A\eta_2 = \lambda_2 \eta_2, \quad \cdots, \quad A\eta_n = \lambda_n \eta_n.$$

于是 $\eta_1, \eta_2, \cdots, \eta_n$ 是 A 的 n 个线性无关特征向量。

充分性如果 A 有 n 个线性无关特征向量 $\eta_1, \eta_2, \cdots, \eta_n$，把它们取作 V 的一组基，显然，A 在这组基下的矩阵成对角形。 □

究竟什么样的线性变换才具有 n 个线性无关的特征向量呢？我们首先给出一个充分条件。

命题 25 线性变换 A 的属于不同特征值的特征向量线性无关。

证明. 取 A 的 k 个不同特征值 $\lambda_1, \lambda_2, \cdots, \lambda_k$，它们分别对应于特征向量 $\xi_1, \xi_2, \cdots, \xi_k$。对 k 作数学归纳法。

当 $k = 1$ 时，$\xi_1 \neq 0$，当然线性无关。设命题在 $k-1$ 个不同特征值的情况下已经成立，证明 k 个不同特征值的情况下命题也成立。

考查向量等式

$$l_1\xi_1 + l_2\xi_2 + \cdots + l_k\xi_k = 0. \tag{1}$$

两边用 A 作用，利用 A 的线性性质，得

$$l_1A\xi_1 + l_2A\xi_2 + \cdots + l_kA\xi_k = 0.$$

因为 $A\xi_i = \lambda_i\xi_i$，故有

$$l_1\lambda_1\xi_1 + l_2\lambda_2\xi_2 + \cdots + l_k\lambda_k\xi_k = 0. \tag{2}$$

以 λ_1 乘 (1) 式再与 (2) 式相减，得 $l_2(\lambda_1 - \lambda_2)\xi_2 + \cdots + l_k(\lambda_1 - \lambda_k)\xi_k = 0.$

按归纳假设，$\xi_2, \cdots, \xi_k$ 线性无关，故 $l_2(\lambda_1 - \lambda_2) = \cdots = l_k(\lambda_1 - \lambda_k) = 0.$

因为 k 个特征值互不相同，由上式即得 $l_2 = \cdots = l_k = 0.$

代入 (1) 式，因 $\xi_1 \neq 0$，就有 $l_1 = 0$。这就证明 $\xi_1, \xi_2, \cdots, \xi_k$ 是线性无关的。 □

推论 如果 n 维线性空间 V 内的线性变换 A 有 n 个不同的特征值，那么它的矩阵可对角化。

命题 26 设 A 是数域 K 上 n 维线性空间 V 内的线性变换，$\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 K 内 k 个不同的数，令 $V_{\lambda_i} = \{a \in V \mid Aa = \lambda_i a\}$，则子空间的和 $\sum_{i=1}^k V_{\lambda_i}$ 是直和。

证明. 按本章定理，只要证 0 向量表法唯一。设 $0 = a_1 + a_2 + \cdots + a_k \quad (a_i \in V_{\lambda_i})$. 因 $a_i \in V_{\lambda_i}$，当 $a_i \neq 0$ 时，它是属于 A 的特征值 λ_i 的特征向量。上式表示 $a_1, a_2, \cdots, a_k$ 中不为 0 的向量线性相关（因为它们的和为 0），但它们属于 A 的不同特征值，这与命题 4.3 矛盾。故 $a_1 = a_2 = \cdots = a_k = 0$，即 $\sum_{i=1}^k V_{\lambda_i}$ 为直和。 □

命题 27 设 A 是数域 K 上 n 维线性空间 V 的线性变换，$\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 A 的全部互不相同的特征值。则 A 的矩阵可对角化的充分必要条件是

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} = \bigoplus_{i=1}^k V_{\lambda_i}.$$

在 A 的矩阵可对角化的情况下，在每个 V_{λ_i} 中任取一组基，合并后即为 V 的一组基，在该组基下 A 的矩阵为对角矩阵。

证明. 必要性　　设 $M = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} \subseteq V.$

如果 A 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的矩阵成对角形，则每个 ϵ_i 均为 A 的特征向量，必属于某个特征子空间 V_{λ_i}，从而属于 M。故对任意 $a \in V$，有 $a = a_1\epsilon_1 + a_2\epsilon_2 + \cdots + a_n\epsilon_n \in M.$ 这表明 $V \subseteq M$，于是 $V = M = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$。

充分性　　设 $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$. 在每个 V_{λ_i} 中取一组基，合并得 V 的一个向量组 (I)，则 (I) 是 V 的一组基，此组基全由特征向量组成。这表明 A 在基 (I) 下的矩阵成对角矩阵。 □

在 V 中取定一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，设 $(A\epsilon_1, A\epsilon_2, \cdots, A\epsilon_n) = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)A.$

对 A 的每个特征值 λ_i 求解齐次线性方程组 $(\lambda_i E - A)X = 0$ 得一个基础解系 $X_{i1}, X_{i2}, \cdots, X_{it_i}$，这里 $t_i = n - r(\lambda_i E - A)$。令

$$\begin{aligned}\eta_{i1} &= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)X_{i1},\\ \eta_{i2} &= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)X_{i2},\\ &\vdots\\ \eta_{it_i} &= (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)X_{it_i}.\end{aligned}$$

前面已指出，$\eta_{i1}, \eta_{i2}, \cdots, \eta_{it_i}$ 即为 V_{λ_i} 的一组基，把它们合并得 V 的一组基 (I)，A 在基 (I) 下的矩阵为对角矩阵

$$D = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_2 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \lambda_k \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} t_1 \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} t_2 \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} \vdots \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} t_k \end{matrix}$$

在 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 到 (I) 的过渡矩阵的列向量为 (I) 中每个向量 η_{ij} 在 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的坐标，即为 X_{ij}。故只要把所求出的基础解系作为列向量依次排列，即得过渡矩阵：

$$T = (X_{11}X_{12}\cdots X_{1t_1}X_{21}X_{22}\cdots X_{2t_2}\cdots X_{k1}X_{k2}\cdots X_{kt_k}).$$

此时有 $T^{-1}AT = D$。现在 A 的特征多项式 $f(\lambda) = |\lambda E - D| = (\lambda - \lambda_1)^{t_1}(\lambda - \lambda_2)^{t_2}\cdots(\lambda - \lambda_k)^{t_k}$，特征值 λ_i 的重数 $t_i = \dim V_{\lambda_i}$。

对数域 K 上每个 n 阶方阵 A，我们总可以把它看做 K 上 n 维线性空间 V 内一个线性变换 A 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的矩阵。同时上述指出了相似矩阵的 T 矩阵的求法.

NOTE: 利用矩阵相似于某个对角矩阵可以轻松计算出该矩阵的幂.

定义 13 设 A 是线性空间 V 内的一个线性变换。如果 M 是 V 的一个子空间，且对任意 $a \in M$，有 $Aa \in M$，则称 M 为 A 的一个**不变子空间**。这时 A 可以看做 M 内的一个线性变换，称为 A 在 M 内的**限制**，记做 $A|_M$。

显然，零子空间 $\{0\}$ 和 V 本身都是 A 的不变子空间，称它们为 A 的**平凡不变子空间**。除此之外的不变子空间称为 A 的**非平凡不变子空间**。例如，当 λ 是 A 的一个特征值时，V_λ 是 A 的一个不变子空间，且 $V_\lambda \neq \{0\}$。如果 A 不是数乘变换，则 $V_\lambda \neq V$，此时 V_λ 就是 A 的一个非平凡不变子空间。

命题 28 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换。在 V 内存在一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，使 A 在这组基下的矩阵成准对角形的充分必要条件是，V 可以分解为 A 的不变子空间 $M_1, M_2, \cdots, M_s$ 的直和

$$V = M_1 \oplus M_2 \oplus \cdots \oplus M_s = \bigoplus_{i=1}^{s} M_i.$$

证明. 必要性　若 $\mathcal{A}$ 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下成下面准对角形

$$A = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s \end{bmatrix},$$

把这组基相应分成 s 段

$$\epsilon_{11}, \epsilon_{12}, \cdots, \epsilon_{1n_1}, \epsilon_{21}, \epsilon_{22}, \cdots, \epsilon_{2n_2}, \cdots, \epsilon_{s1}, \epsilon_{s2}, \cdots, \epsilon_{sn_s},$$

其中 $n_i(i = 1, 2, \cdots, s)$ 为 A_i 的阶，这时应有

$$(\mathcal{A}\epsilon_{i1}, \mathcal{A}\epsilon_{i2}, \cdots, \mathcal{A}\epsilon_{in_i}) = (\epsilon_{i1}, \epsilon_{i2}, \cdots, \epsilon_{in_i})A_i.$$

令 $M_i = L(\epsilon_{i1}, \epsilon_{i2}, \cdots, \epsilon_{in_i})$，则 M_i 为 $\mathcal{A}$ 的不变子空间，且

$$\dim M_i = n_i, \quad \sum_{i=1}^{s} M_i = V.$$

而

$$\dim \sum_{i=1}^{s} M_i = \dim V = n = \sum_{i=1}^{s} n_i = \sum_{i=1}^{s} \dim M_i,$$

故有

$$V = M_1 \oplus M_2 \oplus \cdots \oplus M_s.$$

充分性　设

$$V = M_1 \oplus M_2 \oplus \cdots \oplus M_s,$$

其中 M_i 为 $\mathcal{A}$ 的 n_i 维不变子空间。在每个 M_i 内取一组基，合并成 V 的一个向量组 (I)，由以前的推论知 (I) 是 V 的一组基，且 $\mathcal{A}$ 在此组基下矩阵为准对角形，主对角线上由 s 个小块矩阵 $A_1, A_2, \cdots, A_s$ 组成，A_i 为 $\mathcal{A}|_{M_i}$ 在 M_i 内取定基下的矩阵。 □

命题 29 设 $\mathcal{A}$ 是数域 K 上 n 维线性空间 V 内的一个线性变换。如果 $\mathcal{A}$ 的矩阵可对角化，则对 $\mathcal{A}$ 的任意不变子空间 M，$\mathcal{A}|_M$ 的矩阵也可对角化。

证明. 设 $\mathcal{A}$ 的全部互不相同特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$，由定理 4.2 知现在

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}.$$

令 $N_i = M \cap V_{\lambda_i}(i = 1, 2, \cdots, k)$。我们来证明

$$M = N_1 \oplus N_2 \oplus \cdots \oplus N_k.$$

(i) 证明和 $\sum_{i=1}^{k} N_i$ 为直和。只要证零向量表法唯一。设 $0 = a_1 + a_2 + \cdots + a_k$, 现在 $a_i \in N_i \subseteq V_{\lambda_i}$，由于 $\sum_{i=1}^{k} V_{\lambda_i}$ 为直和，故必 $a_i = 0$。于是 $\sum_{i=1}^{k} N_i$ 为直和。

(ii) 证明 $\sum_{i=1}^k N_i = M$。显然 $\sum_{i=1}^k N_i \subseteq M$。反之，设 α 为 M 内任意向量，则因 $\alpha \in V$，有

$$\alpha = a_1 + a_2 + \cdots + a_k \quad (a_i \in V_{\lambda_i}),$$

$$A\alpha = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k,$$

$$A^{k-1}\alpha = \lambda_1^{k-1} a_1 + \lambda_2^{k-1} a_2 + \cdots + \lambda_k^{k-1} a_k,$$

即

$$\begin{bmatrix} \alpha \\ A\alpha \\ \vdots \\ A^{k-1}\alpha \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix}.$$

上式右端 k 阶方阵的行列式为范德蒙德行列式，$\lambda_1, \lambda_2, \cdots, \lambda_k$ 互不相同，故该行列式不为 0，即此 k 阶方阵可逆，设其逆矩阵为 $T \in M_k(K)$，则

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = T \begin{bmatrix} \alpha \\ A\alpha \\ \vdots \\ A^{k-1}\alpha \end{bmatrix}.$$

因 M 为 A 的不变子空间, 故 $\alpha, A\alpha, \cdots, A^{k-1}\alpha \in M$, 于是由上式得出 $a_i \in M$, 即 $a_i \in V_{\lambda_i} \cap M = N_i$。由此推知 $M \subseteq \sum_{i=1}^k N_i$，即 $M = \sum_{i=1}^k N_i$。

综合上述两方面的结果知

$$M = N_1 \oplus N_2 \oplus \cdots \oplus N_k.$$

在每个 N_i 中取一组基，因 $N_i \subseteq V_{\lambda_i}$，此组基全由 A 的特征向量组成，它们合并成 M 的一组基，在此组基下 $A|_M$ 的矩阵即为对角形。 $\square$

商空间中的诱导变换

设 V 是数域 K 上的线性空间，A 是 V 内一个线性变换。现设 M 是 A 的一个不变子空间。我们在商空间 V/M 定义一个变换如下：$A(\alpha + M) = A\alpha + M$. (左式可写成 $A\overline{\alpha} = \overline{A\alpha}$)。

首先要指出这一定义在逻辑上无矛盾。设有 $\beta + M = \alpha + M$，我们需要证明 $A(\beta + M) = A\beta + M = A\alpha + M = A(\alpha + M)$。现在 $\beta = \alpha + m, m \in M$，因 M 为 A 的不变子空间，故 $A\beta = A\alpha + Am$ 推出：$A\beta + M = A\alpha + M$。这就是我们需要的结论。

考查此前定义的自然映射 $\varphi : V \to V/M$，其中 $\varphi(\alpha) = \alpha + M = \overline{\alpha}$。把这记号用到上面的定义式，得 $A(\alpha + M) = A\varphi(\alpha) = A\alpha + M = \varphi(A\alpha)$. 这表明 φ 和上面定义的 V/M 内变换 A 可交换。

现在来证 A 是 V/M 内线性变换。我们有 $A(k\overline{\alpha} + l\overline{\beta}) = A(\overline{k\alpha + l\beta}) = A\varphi(k\alpha + l\beta) = \varphi A(k\alpha + l\beta) = \varphi(kA\alpha + lA\beta) = k\varphi(A\alpha) + l\varphi(A\beta)) = kA\varphi(\alpha) + lA\varphi(\beta) = kA\overline{\alpha} + lA\overline{\beta}$.

上面定义的 V/M 内的线性变换 A 称为 V 内线性变换 A 在商空间 V/M 内的**诱导变换**。我们用同一个记号 A 代表 V 内线性变换 A 及商空间 V/M 内的诱导变换 A，其具体含意从上下文中立即看出，

不致混淆（如果 A 作用在 V 的向量 α 上，它代表的是 V 内的线性变换；如果它作用在 $\overline{\alpha}=\alpha+M$ 上，则代表的是 V/M 内的诱导变换）。

现在设 $\dim V=n$。在 M 内取一组基 $\epsilon_1,\epsilon_2,\cdots,\epsilon_r$，扩充为 V 的一组基 $\epsilon_1,\cdots,\epsilon_r,\epsilon_{r+1},\cdots,\epsilon_n$。在上一段开头部分已指出（这里继续使用该处记号，不再重复）A 在此组基下的矩阵为如下分块形式：

$$A=\begin{bmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{bmatrix},$$

其中 $A_{22}=\begin{bmatrix}a_{r+1\,r+1} & a_{r+1\,r+2} & \cdots & a_{r+1\,n}\\ a_{r+2\,r+1} & a_{r+2\,r+2} & \cdots & a_{r+2\,n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n\,r+1} & a_{n\,r+2} & \cdots & a_{nn}\end{bmatrix}$.

在以前的证明中已指出 $\overline{\epsilon}_{r+1},\cdots,\overline{\epsilon}_n$ 为 V/M 的一组基，现在显见有（注意 $\overline{\epsilon}_1=\overline{\epsilon}_2=\cdots=\overline{\epsilon}_r=\overline{0}$）
$A\overline{\epsilon}_{r+i}=\overline{A\epsilon_{r+i}}=a_{1\,r+i}\overline{\epsilon}_1+\cdots+a_{r\,r+i}\overline{\epsilon}_r+a_{r+1\,r+i}\overline{\epsilon}_{r+1}+\cdots+a_{n\,r+i}\overline{\epsilon}_n$,

故

$$A\overline{\epsilon}_{r+1}=a_{r+1\,r+1}\overline{\epsilon}_{r+1}+a_{r+2\,r+1}\overline{\epsilon}_{r+2}+\cdots+a_{n\,r+1}\overline{\epsilon}_n,$$

$$A\overline{\epsilon}_{r+2}=a_{r+1\,r+2}\overline{\epsilon}_{r+1}+a_{r+2\,r+2}\overline{\epsilon}_{r+2}+\cdots+a_{n\,r+2}\overline{\epsilon}_n,$$

$$\vdots$$

$$A\overline{\epsilon}_n=a_{r+1\,n}\overline{\epsilon}_{r+1}+a_{r+2\,n}\overline{\epsilon}_{r+2}+\cdots+a_{nn}\overline{\epsilon}_n.$$

于是诱导变换 A 在 V/M 的基 $\overline{\epsilon}_{r+1},\cdots,\overline{\epsilon}_n$ 下的矩阵为 $(A\overline{\epsilon}_{r+1},A\overline{\epsilon}_{r+2},\cdots,A\overline{\epsilon}_n)=(\overline{\epsilon}_{r+1},\overline{\epsilon}_{r+2},\cdots,\overline{\epsilon}_n)A_{22}$.

因为 $|\lambda E-A|=\begin{vmatrix}\lambda E-A_{11} & -A_{12}\\ 0 & \lambda E-A_{22}\end{vmatrix}=|\lambda E-A_{11}|\cdot|\lambda E-A_{22}|$,

上面的关系式给出 V 内线性变换 A 及 $A|_M$ 和商空间 V/M 内诱导变换 A 的特征多项式之间的关系，同时也给出了它们的特征值之间的关系。我们把这关系严格阐述如下。

命题 30 设 A 是数域 K 上 n 维线性空间 V 内的线性变换，M 是 A 的不变子空间。若 A 在 V 内特征多项式为 $f(\lambda)$，$A|_M$ 特征多项式为 $g(\lambda)$，A 在 V/M 的诱导变换特征多项式为 $h(\lambda)$，则 $f(\lambda)=g(\lambda)h(\lambda)$.

命题 31 设 V 是数域 K 上的 n 维线性空间，A 是 V 内的线性变换。如果 A 的特征多项式的根都属于 K，则在 V 内存在一组基，在该组基下 A 的矩阵为上三角矩阵。

证明. 对 n 作数学归纳法。$n=1$ 时命题显然成立。设对 $n-1$ 维线性空间命题成立。当 $\dim V=n$ 时，由假没设 A 必有一特征值 λ_0，设 $A\epsilon_1=\lambda_0\epsilon_1$，其中 $\epsilon_1\neq 0$。令 $M=L(\epsilon_1)$，则 M 为 A 的一维不变子空间，于是 V/M 为 K 上 $n-1$ 维线性空间。根据命题 4.7，A 在 V/M 内的诱导变换 A 的特征多项式的根都是 V 内线性变换 A 的特征多项式的根，从而都属于 K，按归纳假设，在 V/M 内存在一组基 $\overline{\epsilon}_2=\epsilon_2+M,\quad \overline{\epsilon}_3=\epsilon_3+M,\quad\cdots,\quad \overline{\epsilon}_n=\epsilon_n+M$. 使 A 在此组基下矩阵成上三角形，即 $A\overline{\epsilon}_i=a_{2i}\overline{\epsilon}_2+a_{3i}\overline{\epsilon}_3+\cdots+a_{ii}\overline{\epsilon}_i\quad(i=2,3,\cdots,n)$.

(i) 先证 $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$ 为 V 的一组基。这只要证它们线性无关即可。设 $k_1\epsilon_1+k_2\epsilon_2+\cdots+k_n\epsilon_n=0$. 两边用自然映射 φ 作用（注意 $\varphi(\epsilon_1)=\overline{\epsilon}_1=\overline{0}$）: $\varphi(k_1\epsilon_1+k_2\epsilon_2+\cdots+k_n\epsilon_n)=k_1\varphi(\epsilon_1)+k_2\varphi(\epsilon_2)+\cdots+k_n\varphi(\epsilon_n)$

$= k_2\bar{\epsilon}_2 + \cdots + k_n\bar{\epsilon}_n = \bar{0}$. 因 $\bar{\epsilon}_2, \cdots, \bar{\epsilon}_n$ 为 V/M 的一组基，故 $k_2 = \cdots = k_n = 0$。代回原式，因 $\epsilon_1 \neq 0$，推得 $k_1 = 0$。

(ii) 再证在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下 A 的矩阵成上三角形。注意如下事实：因 $M = L(\epsilon_1)$，而 $A\overline{\alpha} = \overline{\beta} \Leftrightarrow A(\alpha+M) = A\alpha+M = \beta+M$, 于是 $A\alpha = \beta+k\epsilon_1 = k\epsilon_1+\beta(k \in K)$。现在 $A\bar{\epsilon}_i = a_{2i}\bar{\epsilon}_2+a_{3i}\bar{\epsilon}_3+\cdots+a_{ii}\bar{\epsilon}_i = a_{2i}\epsilon_2 + a_{3i}\epsilon_3 + \cdots + a_{ii}\epsilon_i$.

根据上面的一般关系式，我们有 $A\epsilon_i = k_i\epsilon_1 + a_{2i}\epsilon_2 + a_{3i}\epsilon_3 + \cdots + a_{ii}\epsilon_i \quad (i = 2, 3, \cdots, n)$, $A\epsilon_1 = \lambda_0\epsilon_1$.

于是 A 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下的矩阵为如下上三角形：

$$A = \begin{bmatrix} \lambda_0 & k_2 & k_3 & \cdots & \cdots & k_n \\ 0 & a_{22} & a_{23} & \cdots & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & \cdots & a_{3n} \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

□

2 双线性函数与二次型

双线性函数

定义 1 设 V 是数域 K 上的线性空间。如果对 V 中任一向量 α，都按某个给定的法则 f 对应于 K 内一个唯一确定的数，记做 $f(\alpha)$，并且满足如下条件：

(i) $f(\alpha + \beta) = f(\alpha) + f(\beta) \quad (\forall\, \alpha, \beta \in V)$;

(ii) $f(k\alpha) = kf(\alpha) \quad (\forall\, \alpha \in V, k \in K)$,

则称 f 为 V 内一个**线性函数**。

定义 2 设 V 是数域 K 上的线性空间。如果 V 中任意一对有序向量 (α, β) 都按照某一法则 f 对应于 K 内唯一确定的一个数，记做 $f(\alpha, \beta)$，且

(i) 对任意 $k_1, k_2 \in K, \alpha_1, \alpha_2, \beta \in V$，有

$$f(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1 f(\alpha_1, \beta) + k_2 f(\alpha_2, \beta);$$

(ii) 对任意 $l_1, l_2 \in K, \alpha, \beta_1, \beta_2 \in V$，有

$$f(\alpha, l_1\beta_1 + l_2\beta_2) = l_1 f(\alpha, \beta_1) + l_2 f(\alpha, \beta_2),$$

则称 $f(\alpha, \beta)$ 是 V 上的一个**双线性函数**。

现设 V 是 n 维线性空间。在 V 内取定一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n.$$

又设

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n;$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

那么，按定义，有

$$\begin{aligned} f(\alpha,\beta) &= f\left(\sum_{i=1}^n x_i\varepsilon_i, \sum_{j=1}^n y_j\varepsilon_j\right) \\ &= \sum_{i=1}^n x_i f\left(\varepsilon_i, \sum_{j=1}^n y_j\varepsilon_j\right) \\ &= \sum_{i=1}^n\sum_{j=1}^n x_i y_j f(\varepsilon_i,\varepsilon_j). \end{aligned}$$

由此可见，$f(\alpha,\beta)$ 由它在一组基处的函数值唯一确定。显然，现在有类似于线性变换的两条结论。

命题 1 设 V 是数域 K 上的 n 维线性空间。在 V 内取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$。则有

(i) V 上一个双线性函数 $f(\alpha,\beta)$ 由它在此组基处的函数值 $f(\varepsilon_i,\varepsilon_j)(i,j=1,2,\cdots,n)$ 唯一确定。换句话说，如果有一个双线性函数 $g(\alpha,\beta)$ 满足 $g(\varepsilon_i,\varepsilon_j)=f(\varepsilon_i,\varepsilon_j)$，则 $g(\alpha,\beta)\equiv f(\alpha,\beta)$。

(ii) 任给数域 K 上一个 n 阶矩阵 $A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ $(a_{ij}\in K)$, 必存在 V 上唯一的一个双线性函数 $f(\alpha,\beta)$，使 $f(\varepsilon_i,\varepsilon_j)=a_{ij}\quad(i,j=1,2,\cdots,n)$.

证明. (i) 已在前面证明。现证结论 (ii)。在 V 内定义二元函数如下：若 $\alpha = x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$, $\beta = y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)Y$, 则令 $f(\alpha,\beta)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_iy_j = X^tAY$.

若 $k_1\alpha_1+k_2\alpha_2 = k_1(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X_1+k_2(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X_2 = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(k_1X_1+k_2X_2)$，则 $f(k_1\alpha_1+k_2\alpha_2,\beta)=(k_1X_1+k_2X_2)^tAY=k_1X_1^tAY+k_2X_2^tAY=k_1f(\alpha_1,\beta)+k_2f(\alpha_2,\beta)$.

类似地，有 $f(\alpha,l_1\beta_1+l_2\beta_2)=l_1f(\alpha,\beta_1)+l_2f(\alpha,\beta_2)$.

故 $f(\alpha,\beta)$ 为 V 内双线性函数。又因 $\varepsilon_i=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}$ (第 i 个是 1), $\quad\varepsilon_j=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}$ (第 j 个是 1)

代入 $f(\alpha,\beta)$ 的定义式，立知 $f(\varepsilon_i,\varepsilon_j)=a_{ij}$.（并且唯一性已知） □

定义 2 我们称

$$A=\begin{bmatrix} f(\varepsilon_1,\varepsilon_1) & f(\varepsilon_1,\varepsilon_2) & \cdots & f(\varepsilon_1,\varepsilon_n) \\ f(\varepsilon_2,\varepsilon_1) & f(\varepsilon_2,\varepsilon_2) & \cdots & f(\varepsilon_2,\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(\varepsilon_n,\varepsilon_1) & f(\varepsilon_n,\varepsilon_2) & \cdots & f(\varepsilon_n,\varepsilon_n) \end{bmatrix}$$

为双线性函数 $f(\alpha,\beta)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵。

推论 设 A,B 是数域 K 上的两个 n 方阵。如果对 K 上任意 $n\times 1$ 阵 X,Y，都有 $X^tAY=X^tBY$，则有 $A=B$。

设 $f(\alpha,\beta)$ 为 V 内一个双线性函数。在 V 内取定两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n;$$

$$\eta_1,\eta_2,\cdots,\eta_n.$$

设 $f(\varepsilon_i,\varepsilon_j)=a_{ij},\quad f(\eta_i,\eta_j)=b_{ij}\quad (i,j=1,2,\cdots,n).$

记 $A=(a_{ij})$, $B=(b_{ij})$，则 A,B 分别是 $f(\alpha,\beta)$ 在这两组基下的矩阵。

令

$$\begin{cases}\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X,\\ \beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)Y,\end{cases}$$

$$\begin{cases}\alpha=\bar{x}_1\eta_1+\bar{x}_2\eta_2+\cdots+\bar{x}_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)\overline{X},\\ \beta=\bar{y}_1\eta_1+\bar{y}_2\eta_2+\cdots+\bar{y}_n\eta_n=(\eta_1,\eta_2,\cdots,\eta_n)\overline{Y}.\end{cases}$$

又设两组基之间的过渡矩阵为 T

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)T.$$

于是有坐标变换公式 $X=T\overline{X};\quad Y=T\overline{Y}.$

将上面的式子代入 $f(\alpha,\beta)$ 的矩阵表达式中，有

$$f(\alpha,\beta)=X^tAY=(T\overline{X})^tA(T\overline{Y})$$

$$=\overline{X}^t(T^tAT)\overline{Y}=\overline{X}^tB\overline{Y}.$$

根据命题 1 的推论，有 $B=T^tAT.$

这就是同一个双线性函数 $f(\alpha,\beta)$ 在两组不同的基下的矩阵之间的关系。

定义 3 给定数域 K 上两个 n 阶方阵 A,B。如果存在 K 上一个可逆的 n 阶方阵 T，使 $B=T^tAT$，则称 B 与 A 在 K 内**合同**。

命题 2 数域 K 上两个 n 阶方阵 A,B 合同的充分必要条件是它们是 K 上 n 维线性空间 V 内一个双线性函数 $f(\alpha,\beta)$ 在两组基下的矩阵。

显然，矩阵之间的合同关系也具有如下几条基本性质

1. 反身性: $A = E^t AE$;
2. 对称性: 若 $B = T^t AT$，则 $A = (T^{-1})^t BT^{-1}$;
3. 传递性: 若 $A = T_1^t BT_1$，$B = T_2^t CT_2$，则

$A = T_1^t BT_1 = T_1^t (T_2^t CT_2)T_1 = (T_2T_1)^t C(T_2T_1)$.

因此，矩阵的合同是一个等价关系。在此等价关系下的等价类称做**合同类**。我们自然也会想从每个合同类中挑选出一个最简单的矩阵（最好是对角矩阵）来作为该合同类的代表。使用双线性函数的语言，就是对每一个双线性函数 $f(\alpha,\beta)$，要设法在 V 内找出一组基，使 $f(\alpha,\beta)$ 在这组基下的矩阵具有最简单的形式。

定义 3 设 V 是数域 K 上的线性空间，$f(\alpha,\beta)$ 是 V 内的一个双线性函数。如果对任意 $\alpha,\beta\in V$，有 $f(\alpha,\beta) = f(\beta,\alpha)$，则称 $f(\alpha,\beta)$ 是一个**对称双线性函数**。(类似地根据反对称矩阵的性质可以定义反对称双线性函数)

现设 $f(\alpha,\beta)$ 是 V 内一个对称双线性函数. 我们定义 $Q_f(\alpha) = f(\alpha,\alpha)$，称为 $f(\alpha,\beta)$ 决定的**二次型函数**. 如在 V 取定基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 又令 $f(\varepsilon_i,\varepsilon_j) = a_{ij}$，$A = (a_{ij})$. 那么 $f(\alpha,\beta) = X^t AY$ （X,Y 为 α,β 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标)，于是

$$Q_f(\alpha) = f(\alpha,\alpha) = X^t AX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad (a_{ij} = a_{ji}).$$

上式称为二次型函数在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的**解析表达式**. 若已知对称双线性函数，则其二次型函数 $Q_f(\alpha)$ 被唯一决定. 反之，因为

$$Q_f(\alpha+\beta) = f(\alpha+\beta,\alpha+\beta) = f(\alpha,\alpha) + 2f(\alpha,\beta) + f(\beta,\beta),$$

故有

$$f(\alpha,\beta) = \frac{1}{2}[Q_f(\alpha+\beta) - Q_f(\alpha) - Q_f(\beta)].$$

这表示反过来二次型函数也唯一决定对称双线性函数 $f(\alpha,\beta)$. 就是说，如果另有 V 对称双线性函数 $g(\alpha,\beta)$，使 $Q_g(\alpha) = Q_f(\alpha)$，则 $g(\alpha,\beta) = f(\alpha,\beta)$. 别地，若 $Q_f(\alpha) = 0$，则 $f(\alpha,\beta) = 0$.

下面是关于对称双线性函数的基本定理.

命题 3 设 V 是数域 K 上的 n 线性空间，$f(\alpha,\beta)$ 是 V 内的一个对称双线性函数，则在 V 存在一组基，使 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形.

证明. 对 V 的维数 n 作数学归纳法.

当 $n = 1$ 时定理是显然的. 设对 $n-1$ 维线性空间定理成立，证明对 n 维线性空间定理也成立.

首先，若 $f(\alpha,\beta) = 0$，定理自然成立. 如果不是这种情况，则 $Q_f(\alpha) = f(\alpha,\alpha) \neq 0$.

于是，可在 V 内取定 ε_1，使 $f(\varepsilon_1,\varepsilon_1) = d_1 \neq 0$. 把 ε_1 扩充成 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$.

令 $\begin{cases} \eta_1 = \varepsilon_1, \\ \eta_i' = \frac{f(\varepsilon_1,\varepsilon_i)}{d_1}\varepsilon_1 - \varepsilon_i \quad (i = 2,3,\cdots,n). \end{cases}$

显然，$\eta_1,\eta_2',\cdots,\eta_n'$ 是 V 的一组基，且 $f(\eta_1,\eta_i') = f\left(\eta_1,\frac{f(\varepsilon_1,\varepsilon_i)}{d_1}\varepsilon_1 - \varepsilon_i\right) = \frac{f(\varepsilon_1,\varepsilon_i)}{d_1}f(\varepsilon_1,\varepsilon_1) - f(\varepsilon_1,\varepsilon_i) = 0 \quad (i = 2,3,\cdots,n)$.

命 $M=L(\eta_2',\cdots,\eta_n')$. 这是一个 $n-1$ 维线性空间，$f(\alpha,\beta)$ 可以看做 M 内的对称双线性函数. 对任意 $\alpha\in M$，有 $\alpha=k_2\eta_2'+\cdots+k_n\eta_n'$，于是 $f(\alpha,\eta_1)=f(\eta_1,\alpha)=k_2f(\eta_1,\eta_2')+\cdots+k_nf(\eta_1,\eta_n')=0$.

按归纳假设, 在 M 内存在一组基 $\eta_2,\cdots,\eta_n$, 使 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形, 即有 $f(\eta_i,\eta_j)=d_1\delta_{ij}\quad(i,j=2,3,\cdots,n)$.

因为 $\eta_1,\eta_2,\cdots,\eta_n$ 与 $\eta_1,\eta_2',\cdots,\eta_n'$ 等价,故它也是 V 的一组基,且 $f(\eta_i,\eta_j)=d_1\delta_{ij}\quad(i,j=1,2,\cdots,n)$.

于是 $f(\alpha,\beta)$ 在这组基下的矩阵成对角形. ■ □

推论 设 A 是数域 K 上的一个 n 阶对称方阵，则存在 K 上的一个可逆方阵 T，使 $T'AT=D$ 为对角形。

定义 4 设 $f(\alpha,\beta)$ 是 n 维线性空间 V 内的一个双线性函数，它在某一组基下的矩阵 A 的秩 $r(A)$ 称为 $f(\alpha,\beta)$ 的**秩**. 如果 A 是满秩的，即 $r(A)=n$，则称 $f(\alpha,\beta)$ 是**满秩双线性函数**（或称**非退化双线性函数**）.（显然给定的双线性函数的秩（双线性函数在某一组基下的矩阵的秩）是确定的）

定义 5 以数域 K 的元素作系数的 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次函数

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\quad(a_{ij}=a_{ji})\tag{1}$$

称为数域 K 上的一个**二次型**，其系数所成的矩阵

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

是数域 K 上的 n 阶对称矩阵，称为此**二次型的矩阵**. A 的秩 $r(A)$ 称为此**二次型的秩**.

定义 6 考查系数属于数域 K 的如下的 n 个变量的线性变数替换

$$\begin{cases}x_1=t_{11}y_1+t_{12}y_2+\cdots+t_{1n}y_n,\\x_2=t_{21}y_1+t_{22}y_2+\cdots+t_{2n}y_n,\\\quad\vdots\\x_n=t_{n1}y_1+t_{n2}y_2+\cdots+t_{nn}y_n.\end{cases}\tag{2}$$

如果其系数矩阵

$$T=\begin{bmatrix}t_{11}&t_{12}&\cdots&t_{1n}\\t_{21}&t_{22}&\cdots&t_{2n}\\\vdots&\vdots&\ddots&\vdots\\t_{n1}&t_{n2}&\cdots&t_{nn}\end{bmatrix}\quad(t_{ij}\in K)$$

可逆，则 (2) 式称为数域 K 上的**可逆线性变数替换**。

令 $X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad Y=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix},$

则 (2) 式可表示成矩阵形式 $X = TY$.

设 $S = (s_{ij}) = T^{-1}$，则有 $Y = SX$. 具体写出就是

$$\begin{cases} y_1 = s_{11}x_1 + s_{12}x_2 + \cdots + s_{1n}x_n, \\ y_2 = s_{21}x_1 + s_{22}x_2 + \cdots + s_{2n}x_n, \\ \quad \cdots \\ y_n = s_{n1}x_1 + s_{n2}x_2 + \cdots + s_{nn}x_n. \end{cases} \tag{3}$$

(3) 式称为 (2) 式的**逆变换**.

如果对二次型 (1) 做上述可逆线性变数替换，我们有

$f = X^tAX \xrightarrow{X=TY} (TY)^tA(TY) = Y^t(T^tAT)Y = g.$

现在因 $(T^tAT)^t = T^tA^tT = T^tAT$，即 T^tAT 仍为 K 上对称矩阵，故上式的 g 为变量 $y_1, y_2, \cdots, y_n$ 的二次型，其矩阵为 T^tAT.

命题 4 给定 K 上两个二次型 $f = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j \quad (a_{ij} = a_{ji})$, $g = \sum_{i=1}^n \sum_{j=1}^n b_{ij}y_iy_j \quad (b_{ij} = b_{ji})$, 它们的矩阵分别为 $A = (a_{ij})$, $B = (b_{ij})$. 则存在 K 上可逆线性变数替换 $X = TY$, 使 f 变成 g 的充分必要条件是 B 与 A 在 K 内合同，即 $B = T^tAT$.

推论 如果 K 上二次型 $f = X^tAX$ 在 K 上的可逆线性变数替换 $X = TY$ 下变为 $g = Y^tBY$, 则 f, g 为 V 内同一个二次型函数 $Q_f(\alpha)$ 在两组基下的解析表达式.

证明. 现在 B 与 A 合同. 按命题 1.2, 它们是 V 内对称双线性函数 $f(\alpha, \beta)$ 在两组基下的矩阵，在此两组基下 $Q_f(\alpha)$ 解析表达式分别为 X^tAX, Y^tBY. □

给定数域 K 上两个二次型 f, g, 若 f 可经 K 上可逆线性变数替换化为 g, 则称 f 与 g **等价**, 记做 $f \sim g$. 命题 4 说明, f 与 g 等价的充分必要条件是它们的矩阵合同. 合同关系是数域 K 上 n 阶对称矩阵集合中的一个等价关系. 因此, 二次型的合同关系也是数域 K 上全体二次型所成的集合内的一个等价关系, 于是二次型按此关系划分为互不相交的等价类. 二次型理论的一个基本问题, 就是要从每个等价类中挑选出一个最简单的二次型来作为该等价类的代表.

形如 $d_1z_1^2 + d_2z_2^2 + \cdots + d_nz_n^2 = Z^tDZ$ 的二次型称为**标准形**, 其矩阵为对角形

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}.$$

显然, 这种二次型是较简单的.

命题 5 给定数域 K 上的一个二次型 $f = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j \quad (a_{ij} = a_{ji})$, 则存在 K 上一个可逆方阵 T, 使在线性变数替换 $X = TZ$ 下此二次型变为如下标准形

$$d_1z_1^2 + d_2z_2^2 + \cdots + d_nz_n^2.$$

这等价于：

1) 用双线性函数的语言：数域 K 上 n 维线性空间 V 内任一对称双线性函数的矩阵都可对角化（即 V 内存在一组基，使该对称双线性函数在此组基下的矩阵成对角形）。

2) 用矩阵论的语言：数域 K 上任一对称的 n 阶方阵都合同于一个对角矩阵。

3) 用二次型的语言：数域 K 上任一二次型都可经一个可逆线性变数替换化为标准形。

二次型标准形的计算方法

1）二次型 (1) 中有某个变量平方项的系数不为零，例如 $a_{11} \neq 0$。此时把二次型对 x_1 配方得 $f = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j = a_{11}\left[x_1^2 + 2x_1 \cdot \frac{a_{12}}{a_{11}}x_2 + \cdots + 2x_1 \cdot \frac{a_{1n}}{a_{11}}x_n\right] + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j = a_{11}\left[x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right]^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}x_ix_j$,

作变数替换 $\begin{cases} y_1 = x_1 + \dfrac{a_{12}}{a_{11}}x_2 + \cdots + \dfrac{a_{1n}}{a_{11}}x_n, \\ y_2 = x_2, \\ \quad\vdots \\ y_n = x_n. \end{cases}$

反解为 $\begin{cases} x_1 = y_1 - \dfrac{a_{12}}{a_{11}}y_2 - \cdots - \dfrac{a_{1n}}{a_{11}}y_n, \\ x_2 = y_2, \\ \quad\vdots \\ x_n = y_n. \end{cases}$

写成矩阵形式 $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & & 0 \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

不难看出,这实际上相当于基变换。经过上述变数替换后,二次型化作 $a_{11}y_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}y_iy_j \quad (b_{ij} = b_{ji})$. 然后再对上式右边的 $n-1$ 个变量 $y_2, y_3, \cdots, y_n$ 的二次型继续进行计算。

如果 $a_{11} = 0$，而某个 $a_{ii} \neq 0$，则对 x_i 配方。

2）所有 $a_{ii} = 0 \ (i = 1, 2, \cdots, n)$，而有一个 $a_{ij} \neq 0 \ (i < j)$，则做变数替换 $\begin{cases} x_i = y_i + y_j, \\ x_j = y_i - y_j, \\ x_k = y_k \quad (k \neq i, j). \end{cases}$ 这就可以把二次型化为第一种情况。

在变数替换的过程中，需要记录坐标变换矩阵，最终得到整个过程的坐标变换矩阵.(按顺序连乘)

复二次型的分类

通过整理合同类的对角矩阵（调整新基的顺序使得不为 0 的对角元素集中在前面），并且对调整后的对角阵元素归一化处理（通过开方换元的可逆现行变数替换），可以得到更特殊的复二次型：

命题 6 数域 $\mathbb{C}$ 上任一个二次型 f 都等价于如下一个二次型:

$$u_1^2 + u_2^2 + \cdots + u_r^2,$$

其中 r 等于二次型 f 的秩，这称为该二次型的**规范形**。规范形是唯一的。

规范形由二次型的秩唯一决定，所以它是唯一的。因为在证明中用到对某些数做开方运算，所以这个结论对一般数域不成立。

秩不同的二次型显然不等价，秩相同的二次型因为有相同的规范形，因而互相等价。二次型的秩可为 $0, 1, 2, \cdots, n$，共 $n+1$ 种可能，所以复数域 $\mathbb{C}$ 上的二次型一共有 $n+1$ 个不同的等价类。

同样的思路用于处理实二次型:

实二次型的分类

命题 7 实数域 $\mathbb{R}$ 上任一个二次型 f 都等价于如下一个实二次型:

$$u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_r^2.$$

它称为 f 的**规范形**。规范形是唯一的。

证明. 现在只需证明规范形的唯一性。规范形中的 r 等于 f 的秩，是唯一确定的，我们只需证明正平方项的个数 p 也是唯一确定的就可以了。

设 f 有两个规范形

$$u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_r^2$$

$$v_1^2 + \cdots + v_q^2 - v_{q+1}^2 - \cdots - v_r^2$$

这表明在 $\mathbb{R}$ 上 n 维线性空间 V 内存在一组基 $\eta_1, \eta_2, \cdots, \eta_n$，使当 $\alpha = u_1\eta_1 + u_2\eta_2 + \cdots + u_n\eta_n$ 时，$Q_f(\alpha) = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_r^2$.

在 V 内又存在一组基 $\omega_1, \omega_2, \cdots, \omega_n$，使当 $\alpha = v_1\omega_1 + v_2\omega_2 + \cdots + v_n\omega_n$ 时，$Q_f(\alpha) = v_1^2 + \cdots + v_q^2 - v_{q+1}^2 - \cdots - v_r^2$.

现令 $M = L(\eta_1, \cdots, \eta_p)$，则当 $\alpha \in M, \alpha \neq 0$ 时，$\alpha = u_1\eta_1 + \cdots + u_p\eta_p \quad (u_i$ 不全为0).

于是 $Q_f(\alpha) = u_1^2 + \cdots + u_p^2 > 0$. 又令 $N = L(\omega_{q+1}, \cdots, \omega_n)$,则当 $\alpha \in N$ 时,有 $\alpha = v_{q+1}\omega_{q+1} + \cdots + v_r\omega_r$.

于是 $Q_f(\alpha) = -v_{q+1}^2 - \cdots - v_r^2 \leq 0$. 这表明 $M \cap N = \{0\}$. 按维数公式，我们有

$$\begin{aligned} n = \dim V &\geq \dim(M+N) = \dim M + \dim N \\ &= p + (n-q). \end{aligned}$$

这表明 $p - q \leq 0$，即 $p \leq q$。由于 p, q 地位对称，同理应有 $q \leq p$，于是 $p = q$。 □

上述命题通常称为**惯性定律**。p 称为实二次型 f 的**正惯性指数**，$r-p$ 称为 f 的**负惯性指数**，它们的差 $p-(r-p) = 2p-r$ 称为 f 的**符号差**。

注意，二次型函数的取值和选择的基无关（选基只是为了量化出计算过程）

定义 6 实数域 $\mathbb{R}$ 上的一个二次型

$$f = X^tAX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad (a_{ij} = a_{ji}), \tag{1}$$

如果它的秩 r 和正惯性指数都等于变量个数 n，则称为一个**正定二次型**。正定二次型的矩阵称为**正定矩阵**。与 f 对应的二次型函数 $Q_f(\alpha)$ 称为**正定二次型函数**。

如果实二次型 (1) 正定，则它的规范形为

$$y_1^2 + y_2^2 + \cdots + y_n^2 = Y^tEY.$$

命题 8 对于实二次型 (1)，下列命题等价：

(i) f 正定；

(ii) A 在 $\mathbb{R}$ 内合同于单位矩阵 E，亦即存在实数域上 n 阶可逆矩阵 T，使 $A = T^tET = T^tT$；

(iii) f 对应的二次型函数 $Q_f(\alpha) > 0$ $(\forall \alpha \in V, \alpha \neq 0)$.

推论 若 A 正定，则 $|A| > 0$.

下面来给出利用实对称矩阵 A 的子式来判断其是否正定的法则．设 $A = (a_{ij})$ 为 n 阶实对称矩阵，$1 \leq i_1 < i_2 < \cdots < i_r \leq n$. 称 A 的 r 子式

$$A\begin{Bmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{Bmatrix}$$

为 A 的一个 **r 阶主子式**.

$$A\begin{Bmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{Bmatrix} \quad (1 \leq k \leq n)$$

则称为 A 的 **k 阶顺序主子式**.

命题 9 给定 n 元实二次型

$$f = X^tAX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad (a_{ij} = a_{ji}).$$

则 f 正定的充分必要条件是其矩阵 A 的各阶顺序主子式都大于零，即

$$A\begin{Bmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{Bmatrix} > 0 \quad (k = 1, 2, \cdots, n).$$

证明. 必要性 $\mathbb{R}$ 上 n 维线性空间 V 内对称双线性函数 $f(\alpha, \beta)$ 在基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 下矩阵为 A. 令 $M = L(\epsilon_1, \epsilon_2, \cdots, \epsilon_k)$. 把 $f(\alpha, \beta)$ 限制在 M 内，在 M 的基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_k$ 下它的矩阵为

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}.$$

因 $\forall\alpha\in M,\alpha\neq 0,Q_f(\alpha)>0$. 由此前的知识知

$$A\begin{Bmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{Bmatrix}=|A_k|>0.$$

充分性　对 n 作数学归纳法. 当 $n=1$ 时, $f=a_{11}x_1^2,a_{11}>0$, 显然 $a_{11}x_1^2>0(x_1\neq 0)$, 故 f 正定. 现设对 $n-1$ 个变元的实二次型命题成立. 考查 V 的子空间 $M=L(\epsilon_1,\epsilon_2,\cdots,\epsilon_{n-1})$. $f(\alpha,\beta)$ 限制在 M 内，在基 $\epsilon_1,\epsilon_2,\cdots,\epsilon_{n-1}$ 下的矩阵为

$$A_{n-1}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{bmatrix},$$

其各阶顺序主子式 >0. 按归纳假设, $\forall\alpha\in M,\alpha\neq 0,Q_f(\alpha)>0$. 由于 A_{n-1} 合同于 E_{n-1}. 于是 M 内存在一组基 $\eta_1,\eta_2,\cdots,\eta_{n-1}$, 使 $f(\alpha,\beta)$ 在该组基下矩阵为 E_{n-1}. 即

$$f(\eta_i,\eta_j)=\delta_{ij}\quad(i,j=1,2,\cdots,n-1). \tag{2}$$

将 $\eta_1,\eta_2,\cdots,\eta_{n-1}$ 添加 ξ 使成 V 的一组基。令 $\zeta=\xi-f(\eta_1,\xi)\eta_1-f(\eta_2,\xi)\eta_2-\cdots-f(\eta_{n-1},\xi)\eta_{n-1}$, 则 $\eta_1,\eta_2,\cdots,\eta_{n-1},\zeta$ 与 $\eta_1,\eta_2,\cdots,\eta_{n-1},\xi$ 等价, 也是 V 的一组基。因（利用 (2) 式）$f(\eta_i,\zeta)=f(\eta_i,\xi)-f(\eta_i,\xi)f(\eta_i,\eta_i)=f(\eta_i,\xi)-f(\eta_i,\xi)=0$. 故 $f(\alpha,\beta)$ 在基 $\eta_1,\eta_2,\cdots,\eta_{n-1},\zeta$ 下的矩阵为

$$B=\begin{bmatrix} E_{n-1} & 0 \\ 0 & f(\zeta,\zeta) \end{bmatrix}.$$

B 与 A 是 $f(\alpha,\beta)$ 在不同基下的矩阵，互相合同，即有 $T\in M_n(\mathbb{R})$，$|T|\neq 0$，使 $T^tAT=B$，于是（注意 $|A|>0$）

$$d=f(\zeta,\zeta)=|B|=|T^tAT|=|T|^2\cdot|A|>0.$$

令 $\eta_n=\frac{1}{\sqrt{d}}\zeta$，则

$$f(\eta_n,\eta_i)=\frac{1}{\sqrt{d}}f(\zeta,\eta_i)=0\quad(i=1,2,\cdots,n-1),$$

$$f(\eta_n,\eta_n)=\frac{1}{d}f(\zeta,\zeta)=1.$$

而 $\eta_1,\eta_2,\cdots,\eta_{n-1},\eta_n$ 为 V 的一组基，在此基下 $f(\alpha,\beta)$ 的矩阵为 E_n。即 A 合同于 E_n，从而 f 正定。 □

n 元实二次型 $f=X^tAX$ 可以划分为以下几个大类:

1) 正定二次型;

2) **半正定二次型**: 其规范形为 $y_1^2+y_2^2+\cdots+y_r^2$, 即 f 的正惯性指数 $p=$ 秩r. 显然，f 半正定的充分必要条件是对一切 $\alpha\in V,Q_f(\alpha)=X^tAX\geq 0$, 半正定型的矩阵称为**半正定矩阵**;

3) **负定二次型**: 其规范形为 $-y_1^2-y_2^2-\cdots-y_n^2$, 即 f 的负惯性指数 $q=$ 秩$r=n$. 显然，f 负定的充分必要条件是对一切 $\alpha\neq 0,Q_f(\alpha)=X^tAX<0$, 负定二次型的矩阵称**负定矩阵**;

4) **半负定二次型**：其规范形为 $-y_1^2-y_2^2-\cdots-y_r^2$, 即 f 的负惯性指数 $q=$ 秩r. 显然，f 半负定的充分必要条件是对一切 $\alpha, Q_f(\alpha)=X^tAX\leq 0$, 半负定二次型的矩阵称**半负定矩阵**；

5) 除上述四类之外，其他实二次型都称为**不定型**。