第一章　空间与形式

5　对偶空间简介

在本节中 V 是域 F 上的 n 维线性空间.

线性空间 $\mathrm{Hom}(V,F)$ 称为 V 的对偶空间, 记为 V^*. 换言之, V^* 是 V 上所有线性函数的集合, 其中的加法和数乘由 $\mathrm{Map}(V,F)$ 给出. 设 $f\in V^*$, $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基. $\mathbf{x}=x_1\mathbf{e}_1+\ldots x_n\mathbf{e}_n\in V$. 则

$$f(\mathbf{x})=x_1f(\mathbf{e}_1)+\cdots+x_nf(\mathbf{e}_n)=(f(\mathbf{e}_1),\ldots,f(\mathbf{e}_n))\begin{pmatrix}x_1\\ \vdots\\ c_n\end{pmatrix}.$$

反之, 设 $p=\alpha_1x_1+\cdots+\alpha_nx_n\in F[x_1,\ldots,x_n]$. 则 p 可以解释为由 $p(\mathbf{e}_1)=\alpha_1,\ldots,p(\mathbf{e}_n)=\alpha_n$ 在 F^n 上定义的线性函数, 其中 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 F^n 的标准基.

定理 5.1 设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基. 则在 V^* 中存在唯一的一组基 $\mathbf{e}_1^*,\ldots,\mathbf{e}_n^*$ 满足 $\mathbf{e}_i^*(\mathbf{e}_j)=\delta_{i,j}, i,j\in\{1,2,\ldots,n\}$. 特别地, $\dim(V^*)$=n. ($\mathbf{e}_1^*,\ldots,\mathbf{e}_n^*$ 称为 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 的对偶基.)

证明. 对 $i\in\{1,2,\ldots,n\}$, 设 $\mathbf{e}_i^*\in V^*$ 满足 $\mathbf{e}_i^*(\mathbf{e}_j)=\delta_{i,j}$, $j=1,2,\ldots,n$. 由线性映射基本定理 II 可知这样的 $\mathbf{e}_i^*$ 存在且唯一. 我们只要证明 $\mathbf{e}_1^*,\ldots,\mathbf{e}_n^*$ 是 V^* 的基即可.

设 $\alpha_1, \ldots, \alpha_n \in F$ 使得 $\alpha_1 \mathbf{e}_1^* + \cdots + \alpha_n \mathbf{e}_n^* = \mathbf{0}^*$, 其中 $\mathbf{0}^*$ 代表 V^* 中的零元, 即零函数. 设 $j \in \{1, 2, \ldots, n\}$. 则

$$0 = \mathbf{0}^*(\mathbf{e}_j) = \left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i^*\right)(\mathbf{e}_j) = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{i=1}^{n} \alpha_i \delta_{i,j} = \alpha_j.$$

于是 $\alpha_1 = \cdots = \alpha_n = 0$, 即 $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ 线性无关.

再设 $f \in V^*$ 且 $f(\mathbf{e}_j) = \beta_j, j = 1, 2, \ldots, n$. 令

$$g = \beta_1 \mathbf{e}_1^* + \cdots + \beta_n \mathbf{e}_n^*.$$

则对于任意 $j \in \{1.2. \ldots, n\}$, 我们有

$$g(\mathbf{e}_j) = \left(\sum_{i=1}^{n} \beta_i \mathbf{e}_i^*\right)(\mathbf{e}_j) = \sum_{i=1}^{n} \beta_i \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{i=1}^{n} \beta_i \delta_{i,j} = \beta_j.$$

再由线性映射基本定理 II 中的唯一性可知, $f = g$. □

例 5.2 对偶基为取坐标提供方便. 设 $\mathbf{x} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$. 证明: 对任意 $i \in \{1, 2, \ldots, n\}$, $x_i = \mathbf{e}_i^*(\mathbf{x})$.

证明. 我们计算

$$\mathbf{e}_i^*(\mathbf{x}) = \mathbf{e}_i^*\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} x_j \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \delta_{i,j} = x_i. \quad \square$$

由此我们可以得出 $\mathbf{x} = \mathbf{0}$ 当且仅当 $\mathbf{e}_i^*(\mathbf{x}) = 0$ 对任意 $i \in \{1, 2, \ldots, n\}$ 成立.

引理 5.3 设 $f_1, \ldots, f_n$ 是 V^* 的一组基, $\mathbf{x}, \mathbf{y} \in V$. 则 $\mathbf{x} = \mathbf{y}$ 当且仅当 $f_i(\mathbf{x}) = f_i(\mathbf{y})$, $i = 1, 2, \ldots, n$.

证明. 设 $\mathbf{z} \in V$. 只要证明:

$$\mathbf{z} = \mathbf{0} \iff f_1(\mathbf{z}) = \cdots = f_n(\mathbf{z}) = \mathbf{0}.$$

"$\Longrightarrow$" 是显然的.

"$\Longleftarrow$". 假设 $\mathbf{z} \neq \mathbf{0}$. 由线性映射基本定理 II 可知, 存在 $f \in V^*$ 使得 $f(\mathbf{z}) = 1$. 因为 $f_1, \ldots, f_n$ 是 V^* 的基底, 所以存在 $\alpha_1, \ldots, \alpha_n \in F$ 使得 $f = \sum_{i=1}^n \alpha_i f_i$. 于是, $f(\mathbf{z}) = \left(\sum_{i=1}^n \alpha_i f_i\right)(\mathbf{z}) = \sum_{i=1}^n \alpha_i f_i(\mathbf{z}) = \mathbf{0}$. 矛盾. $\square$

6 双线性型

本节中 V 是 F 上的 n 维线性空间, $n > 0$.

6.1 定义和矩阵表示

设

$$\begin{array}{rccc} f: & V \times V & \longrightarrow & F \\ & (\mathbf{x}, \mathbf{y}) & \mapsto & f(\mathbf{x}, \mathbf{y}). \end{array}$$

如果对任意的 $\alpha, \beta \in F$ 和 $\mathbf{z} \in V$ 满足

$$f(\alpha\mathbf{x} + \beta\mathbf{z}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y}) + \beta f(\mathbf{z}, \mathbf{y})$$

和

$$f(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \alpha f(\mathbf{x}, \mathbf{y}) + \beta f(\mathbf{x}, \mathbf{z}),$$

则称 f 是 V 上的双线性型.

例 6.1 设 f 是 V 上的双线性型. 则对任意的 $\mathbf{u},\mathbf{v},\mathbf{x},\mathbf{y}\in V$ 和 $\alpha,\beta\in F$, 我们有:

$$\begin{aligned} f(\mathbf{u}+\mathbf{v},\mathbf{x}+\mathbf{y}) &= f(\mathbf{u},\mathbf{x}+\mathbf{y})+f(\mathbf{v},\mathbf{x}+\mathbf{y}) \\ &= f(\mathbf{u},\mathbf{x})+f(\mathbf{u},\mathbf{y})+f(\mathbf{v},\mathbf{x})+f(\mathbf{v}+\mathbf{y}) \end{aligned}$$

且

$$f(\alpha\mathbf{x},\beta\mathbf{y})=\alpha f(\mathbf{x},\beta\mathbf{y})=\alpha\beta f(\mathbf{x},\mathbf{y}).$$

此外

$$f(\mathbf{0},\mathbf{y})=f(\mathbf{0}+\mathbf{0},\mathbf{y})=f(\mathbf{0},\mathbf{y})+f(\mathbf{0},\mathbf{y}) \implies f(\mathbf{0},\mathbf{y})=0.$$

同理, $f(\mathbf{x},\mathbf{0})=0$.

定理 6.2 设 V 的一组基是 $\mathbf{e}_1,\dots,\mathbf{e}_n$, f 是 V 上的双线性型. 则存在唯一的矩阵 $A\in \mathrm{M}_n(F)$ 使得,

$$\forall\ \mathbf{x}=(\mathbf{e}_1,\dots,\mathbf{e}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\ \mathbf{y}=(\mathbf{e}_1,\dots,\mathbf{e}_n)\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

我们有

$$f(\mathbf{x},\mathbf{y})=(x_1,\dots,x_n)A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

事实上, $A=(f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}$. 称 A 是 f 在 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵表示.

证明. 我们计算

$$\begin{aligned}
f(\mathbf{x},\mathbf{y}) &= f\left(\sum_{i=1}^{n} x_i\mathbf{e}_i,\ \sum_{j=1}^{n} y_j\mathbf{e}_j\right)\\
&= \sum_{i=1}^{n} x_i f\left(\mathbf{e}_i,\ \sum_{j=1}^{n} y_j\mathbf{e}_j\right)\\
&= \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{n} y_j f(\mathbf{e}_i,\ \mathbf{e}_j)\right)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} f(\mathbf{e}_i,\mathbf{e}_j)x_i y_j.
\end{aligned}$$

令 $A = (f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}$. 由上式直接验证得

$$f(\mathbf{x},\mathbf{y}) = (x_1,\ldots,x_n)A\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}.$$

再设 $B = (b_{i,j}) \in \mathrm{M}_n(F)$ 使得

$$f(\mathbf{x},\mathbf{y}) = (x_1,\ldots,x_n)B\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}.$$

对 $i, j \in \{1, 2, \ldots, n\}$, 取 $\mathbf{x} = \mathbf{e}_i, \mathbf{y} = \mathbf{e}_j$, 则

$$f(\mathbf{e}_i, \mathbf{e}_j) = \begin{pmatrix}0, \ldots, \underset{\substack{\uparrow\\ i}}{1}, \ldots, 0\end{pmatrix} \; B \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} \begin{matrix}\\ \\ \to j\\ \\ \\ \end{matrix} = \vec{B}_i \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} \begin{matrix}\\ \\ \to j\\ \\ \\ \end{matrix} = b_{i,j}.$$

于是 $A = B$. $\square$

例 6.3 设 $V = \mathbb{R}^2$. 对任意的

$$\mathbf{x} = \begin{pmatrix}x_1\\ x_2\end{pmatrix}, \mathbf{y} = \begin{pmatrix}y_1\\ y_2\end{pmatrix}, \quad f(\mathbf{x}, \mathbf{y}) := \det\begin{pmatrix}x_1 & y_1\\ x_2 & y_2\end{pmatrix}.$$

验证 f 是 V 上的双线性型, 并求它在标准基下的矩阵.

解. 设 $\alpha, \beta \in F, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$,

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{z}, \mathbf{y}) &= \det(\alpha\mathbf{x} + \beta\mathbf{z}, \mathbf{y}) = \det(\alpha\mathbf{x}, \mathbf{y}) + \det(\beta\mathbf{z}, \mathbf{y}) \\ &= \alpha \det(\mathbf{x}, \mathbf{y}) + \beta \det(\mathbf{z}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y}) + \beta f(\mathbf{z}, \mathbf{y}). \end{aligned}$$

类似 f 对第二个变元线性. 于是 f 是双线性型.

注意到 $f(\mathbf{e}_1, \mathbf{e}_1) = f(\mathbf{e}_2, \mathbf{e}_2) = 0$, $f(\mathbf{e}_1, \mathbf{e}_2) = 1$ 而 $f(\mathbf{e}_2, \mathbf{e}_1) = -1$. 于是

$$f(\mathbf{x}, \mathbf{y}) = (x_1, x_2) \underbrace{\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}}_{A} \begin{pmatrix}y_1\\ y_2\end{pmatrix}.$$

例 6.4 设 $A \in \mathrm{M}_n(F)$. 则

$$\begin{array}{rccc} f: & F^n \times F^n & \longrightarrow & F \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} & \mapsto & (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \end{array}$$

是 F^n 上的双线性型, f 在标准基下的矩阵是 A.

证明. 设 $\mathbf{x} = (x_1, \ldots, x_n)^t, \mathbf{y} = (y_1, \ldots, y_n)^t, \mathbf{z} = (z_1, \ldots, z_n)^t$, $\alpha, \beta \in F$. 则

$$\begin{aligned} f(\alpha\mathbf{x} + \beta z, y) &= (\alpha\mathbf{x} + \beta\mathbf{z})^t A\mathbf{y} \\ &= \alpha(\mathbf{x}^t A\mathbf{y}) + \beta(\mathbf{z}^t A\mathbf{y}) \\ &= \alpha f(\mathbf{x}, \mathbf{y}) + \beta f(\mathbf{z}, \mathbf{y}). \end{aligned}$$

于是, f 对第一个变元线性. 类似可验证 f 对第二个变元也线性. 从而 f 是双线性型. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 F^n 的标准基, $A = (a_{i,j})_{n\times n}$. 则 $f(\mathbf{e}_i, \mathbf{e}_j) = a_{i,j}$, $i, j \in \{1, 2, \ldots, n\}$. 由定理 6.2, f 在标准基下的矩阵是 A. $\square$

设 f 是 V 上的双线性型, f 在 V 的两组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\epsilon_1, \ldots, \epsilon_n$ 下的矩阵分别是 A 和 B. 再设

$$(\epsilon_1, \ldots, \epsilon_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)P, \quad P \in \mathrm{GL}_n(F).$$

设

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n = u_1\epsilon_1 + \cdots + u_n\epsilon_n,$$

和

$$\mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_n\mathbf{e}_n = v_1\epsilon_1 + \cdots + v_n\epsilon_n.$$

则由坐标变换公式可得

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= (x_1,\dots,x_n)A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= (u_1,\dots,u_n)\underbrace{P^tAP}_{B}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. \end{aligned}$$

根据定理 6.2, $B = P^tAP$. 反之, 给定 F^n 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $P \in \mathrm{GL}_n(F)$, f 在给定基底下的矩阵是 A, 则 $(\epsilon_1,\dots,\epsilon_n) = (\mathbf{e}_1,\dots,\mathbf{e}_n)P$ 是 F^n 的一组基. 由上述计算可知 f 在新的基底下的矩阵是 P^tAP.

反推上述过程可知, 如果双线性型 $f(\mathbf{x},\mathbf{y})$ 在基底 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵是 A, 且 $B = P^tAP$, 其中 $P \in \mathrm{GL}_n(F)$. 则 B 是该双线性型在基底 $(\mathbf{e}_1,\dots,\mathbf{e}_n)P$ 下的矩阵.

定义 6.5 *设 $A, B \in \mathrm{M}_n(F)$. 如果存在 $P \in \mathrm{GL}_n(F)$ 使得 $B = P^tAP$, 则称 B 合同于 A, 记为 $B \sim_c A$.*

我们来验证 $\sim_c$ 是等价关系. 对任意 $A \in \mathrm{M}_n(F)$, $A = E^tAE \implies A \sim_c A$. 自反性成立. 设 $B \sim_c A$. 则存在

$P \in \mathrm{GL}_n(F)$ 使得 $B = P^t AP$. 于是

$$A = (P^t)^{-1} BP^{-1} = (P^{-1})^t BP^{-1} \implies A \sim_c B.$$

对称性成立. 设 $A \sim_c B$, $B \sim_c C$. 则存在 $P, Q \in \mathrm{GL}_n(F)$ 使得

$$\begin{aligned} & A = P^t BP,\ B = Q^t CQ \\ & \implies A = P^t Q^t CQP = (QP)^t C(QP) \\ & \implies A \sim_c C. \end{aligned}$$

传递性成立.

从以上论述我们看出, 一个双线性型在不同基底下的矩阵是合同的. 而两个彼此合同的矩阵一定是一个双线性型在不同基底下的矩阵. 于是, 研究双线性型等价于研究方阵在合同意义下的等价类. 利用矩阵的语言, 我们所要研究的问题是: $\mathrm{M}_n(F)/\sim_c$ 含有多少不同的等价类？在每个等价类中可否找出一个"标准" 的代表元? 这个代表矩阵中应该含有尽可能多个 0, 而非零元素出现的位置应该尽可能有规律.

命题 6.6 设 $A, B \in \mathrm{M}_n(F)$. 若 $A \sim_c B$, 则 $\mathrm{rank}(A) = \mathrm{rank}(B)$.

证明. 设 $P \in \mathrm{GL}_n(F)$ 使得 $A = P^t BP$. 因为 P 满秩, 所以 $\mathrm{rank}(A) = \mathrm{rank}(B)$. (见上学期讲义第二章推论 4.3) $\square$.

定义 6.7 设 f 是 V 上的双线性型, A 是 f 在 V 的某组基下的矩阵. 则 f 的秩定义为 $\mathrm{rank}(A)$, 记为 $\mathrm{rank}(f)$.

由上述命题可知, $\mathrm{rank}(f)$ 是良定义的. 下例说明双线性型可以通过矩阵给出.

例 6.8 合同关系保持对称和斜对称性. 设 $A \in \mathrm{M}_n(F)$ (斜)对称, 且 $A \sim_c B$. 证明 B 也(斜)对称.

证明. 设 A 斜对称. 因为 $A \sim_c B$, 所以存在 $P \in \mathrm{GL}_n(F)$ 使得 $B = P^tAP$. 则

$$B^t = (P^tAP)^t = P^tA^tP = -P^tAP = -B.$$

对称情形类似. □

6.2 对称双线性型

定义 6.9 设 f 是 V 上的双线性型. 如果对任意 $\mathbf{x},\mathbf{y}\in V$, $f(\mathbf{x},\mathbf{y}) = f(\mathbf{y},\mathbf{x})$, 则称 f 是对称双线性型.

命题 6.10 设 f 是 V 上的双线性型, $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基, A 是 f 在 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下的矩阵. 则 f 是对称的当且仅当 A 是对称的.

证明. 设 f 对称. 则对任意 $i,j \in \{1,\ldots,n\}$, 我们有 $f(\mathbf{e}_i,\mathbf{e}_j) = f(\mathbf{e}_j,\mathbf{e}_i)$. 故 $A = (f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}$ 对称. 反之, 设 A

对称, $\mathbf{x}$ 和 $\mathbf{y}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标分别是 $(x_1, \ldots, x_n)^t$ 和 $(y_1, \ldots, y_n)^t$. 则

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= (x_1, \ldots, x_n) A (y_1, \ldots, y_n)^t \\ &= (y_1, \ldots, y_n) A^t (x_1, \ldots, x_n)^t \quad (\because F \text{ 中的元素转置不变}) \\ &= (y_1, \ldots, y_n) A (x_1, \ldots, x_n)^t = f(\mathbf{x}, \mathbf{y}). \quad (\because A^t = A) \end{aligned}$$

故 f 对称. □

空间 V 上的所有对称双线性型记为 $\mathcal{L}_2^+(V)$. 本节的主要结果是

定理 6.11 *设 F 的特征不等于 2 且 $f \in \mathcal{L}_2^+(V)$. 则 V 中有一组基使得 f 在该基下的矩阵是对角阵.*

证明该定理需要对称双线性型的极化公式. 设 F 的特征不等于 2 且 $f \in \mathcal{L}_2^+(V)$. 则对于任意 $\mathbf{x}, \mathbf{y} \in V$,

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \left(f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - f(\mathbf{x}, \mathbf{x}) - f(\mathbf{y}, \mathbf{y}) \right). \tag{1}$$

验证如下:

$$\begin{aligned} &\frac{1}{2} \left(f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - f(\mathbf{x}, \mathbf{x}) - f(\mathbf{y}, \mathbf{y}) \right) \\ &= \frac{1}{2} \left(f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y}) - f(\mathbf{x}, \mathbf{x}) - f(\mathbf{y}, \mathbf{y}) \right) \quad (\text{双线性}) \\ &= \frac{1}{2} \left(f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) \right) \\ &= f(\mathbf{x}, \mathbf{y}). \quad (\text{对称性}) \end{aligned}$$

定理 6.11 的证明. 如果 f 是零映射, 则 f 在 V 的任意基底下的矩阵都是零矩阵. 定理显然成立. 设 f 不是零映射.

再设 $n=\dim(V)$. 我们对 n 归纳. 当 $n=1$ 时, 定理显然成立. 设 $n>1$ 且定理对 $n-1$ 成立.

由极化公式 (1), 存在 $\mathbf{e}_1\in V$ 使得 $f(\mathbf{e}_1,\mathbf{e}_1)\neq 0$. 令 $W=\{\mathbf{v}\in V\mid f(\mathbf{v},\mathbf{e}_1)=0\}$. 可直接验证 W 是子空间. 我们来证明

$$V=\langle \mathbf{e}_1\rangle\oplus W. \tag{2}$$

首先, 设 $\mathbf{x}\in\langle\mathbf{e}_1\rangle\cap W$. 则 $\mathbf{x}=\lambda\mathbf{e}_1$, 其中 $\lambda\in F$, 且 $f(\lambda\mathbf{e}_1,\mathbf{e}_1)=0$. 于是 $\lambda f(\mathbf{e}_1,\mathbf{e}_1)=0$. 因为 $f(\mathbf{e}_1,\mathbf{e}_1)\neq 0$, 所以 $\lambda=0\implies \mathbf{x}=\mathbf{0}$. 由第一章第一讲定理 1.12 (iii), $\langle\mathbf{e}_1\rangle+W$ 是直和. 由第一章第二讲命题 4.15, 只要证明 $\dim(W)=n-1$ 即可. 考虑线性映射

$$\begin{aligned}\phi:\ V &\longrightarrow F\\ \mathbf{x} &\mapsto f(\mathbf{x},\mathbf{e}_1).\end{aligned}$$

则 $W=\ker(\phi)$. 因为 $\phi(\mathbf{e}_1)\neq 0$, 所以 $\dim(\mathrm{im}(\phi))\geq 1$. 但 $\mathrm{im}(\phi)\subset F$ 且 $\dim F=1$. 于是 $\mathrm{im}(\phi)=F$. 特别地 $\dim(\mathrm{im}(\phi))=1$. 由对偶公式, $\dim(W)=n-1$. 直和分解 (2) 成立.

设 $g\in\mathcal{L}_2(W)$ 满足对任意 $\mathbf{x},\mathbf{y}\in W$, $g(\mathbf{x},\mathbf{y})=f(\mathbf{x},\mathbf{y})$. 由归纳假设存在 W 的一组基 $\mathbf{e}_2,\ldots,\mathbf{e}_n$ 使得 g 在该基下的矩阵是对角的, 即对任意的 $i,j\in\{2,3,\ldots,n\}, i\neq j$, 我们有 $g(\mathbf{e}_i,\mathbf{e}_j)=f(\mathbf{e}_i,\mathbf{e}_j)=0$. 由 (2), $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ 线性无关 (见第一章第一讲定理 1.12 (ii))且 $\dim(V)=n$. 于是

$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ 是 V 的一组基. 由 W 的定义可知

$$f(\mathbf{e}_i, \mathbf{e}_1) = 0, \quad i = 2, 3, \ldots, n.$$

在由 f 的对称性可知

$$f(\mathbf{e}_1, \mathbf{e}_i) = 0, \quad i = 2, 3, \ldots, n.$$

综上所述 $f(\mathbf{e}_i, \mathbf{e}_j) = 0$, $i, j \in \{1, 2, \ldots, n\}$ 且 $i \neq j$. 于是 f 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵是对角阵. □

定义 6.12 设 $f \in \mathcal{L}_2^+(V)$, f 在 V 的基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵是对角阵. 则称 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 f 的一组规范基. 设双线性型 f 在一组规范基下的矩阵为 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 则

$$f(\mathbf{x}, \mathbf{y}) = \lambda_1 x_1 y_1 + \cdots + \lambda_n x_n y_n$$

称为与规范基对应的规范型, 其中 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, $\mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_n\mathbf{e}_n$.

推论 6.13 设 F 的特征不等于 2, $A \in \mathrm{SM}_n(F)$. 则 A 合同于一个对角阵.

证明. 考虑双线性型

$$\begin{array}{cccc} f: & F^n \times F^n & \longrightarrow & F \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} & \mapsto & (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \end{array}$$

因为 A 对称, 所以 f 对称. 由上述定理存在 F^n 的一组基使得 f 在该基下的矩阵是对角阵 B. 则 $A \sim_c B$. $\square$

例 6.14 求 $P \in \mathrm{GL}_3(\mathbb{R})$ 使得

$$P^t \underbrace{\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{A} P$$

是对角矩阵.

解. 设 f 是 $\mathbb{R}^3$ 上对称双线性型, 它在标准基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的矩阵是 A.

步骤 1. 选取 ϵ_1 使得 $f(\epsilon_1, \epsilon_1) \neq 0$. 令 $\epsilon_1 = \mathbf{e}_1 + \mathbf{e}_2$. 则

$$\begin{aligned} f(\epsilon_1, \epsilon_1) &= f(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2) \\ &= f(\mathbf{e}_1, \mathbf{e}_1) + f(\mathbf{e}_1, \mathbf{e}_2) + f(\mathbf{e}_2, \mathbf{e}_1) + f(\mathbf{e}_2, \mathbf{e}_2) = 2. \end{aligned}$$

步骤 2. 确定 $W = \ker(f(\mathbf{x}, \epsilon_1))$ 的一组基. 我们计算

$$f(\mathbf{x}, \epsilon_1) = (x_1, x_2, x_3)A(\mathbf{e}_1 + \mathbf{e}_2) = x_1 + x_2 + 2x_3.$$

解方程 $x_1 + x_2 + 2x_3 = 0$ 得到 W 的一组基

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

步骤 3. 求 $g := f|_{W\times W}$ 在 $\mathbf{w}_1, \mathbf{w}_2$ 下的矩阵

$$B = \begin{pmatrix} f(\mathbf{w}_1, \mathbf{w}_1) & f(\mathbf{w}_1, \mathbf{w}_2) \\ f(\mathbf{w}_2, \mathbf{w}_1) & f(\mathbf{w}_2, \mathbf{w}_2) \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ -2 & -4 \end{pmatrix}.$$

到此降维到 W 上的对称双线性型 g.

步骤 1. 选取 ϵ_2 使得 $g(\epsilon_2, \epsilon_2) \neq 0$. 令 $\epsilon_2 = \mathbf{w}_1$.

步骤 2. 确定 $Z = \ker(g(\mathbf{x}, \epsilon_2))$ 的一组基. 我们计算

$$g(\mathbf{y}, \epsilon_2) = (y_1, y_2)B \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -2y_1 - 2y_2.$$

解方程 $-2y_1 - 2y_2 = 0$ 得到解空间的一组基

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \Longrightarrow Z \text{ 的基是 } (\mathbf{w}_1, \mathbf{w}_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

于是, f 在 $\mathbb{R}^3$ 中的一组规范基是

$$\epsilon_1 = \mathbf{e}_1 + \mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon_2 = \mathbf{w}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \epsilon_3 = \mathbf{w}_1 - \mathbf{w}_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

从 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 到 $\epsilon_1, \epsilon_2, \epsilon_3$ 的矩阵是

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

计算得

$$P^t A P = \operatorname{diag}(2, -2, -2).$$

例 6.15 设

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SM}_2(\mathbb{Z}_2).$$

证明 A 不合同于对角方阵.

证明. 设

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}_2)$$

使得 $P^t AP = \mathrm{diag}_2(u, v)$. 则

$$P^t AP = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} c & d \\ a & b \end{pmatrix} = \begin{pmatrix} 2ac & ad+bc \\ ad+bc & 2bd \end{pmatrix} = \mathrm{diag}_2(u, v).$$

于是 $u = v = 0$ $(\because\ 2 = 0)$. 由此可知 $\mathrm{rank}(A) = 0$. 矛盾.

推论 6.16 设 $f \in \mathcal{L}_2^+(V)$ 且 $r = \mathrm{rank}(f)$. 则存在 V 的一组规范基使得 f 在该基下的矩阵是

$$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \cdots, 0),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_r \in F \setminus \{0\}$.

证明. 由定理 6.11, 存在 f 规范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$. 于是

$$f(\mathbf{e}_i, \mathbf{e}_j) = 0, \quad i, j \in \{1, 2, \ldots, n\} \text{ 且 } i \neq j.$$

因为 $r = \mathrm{rank}(A)$, 所以在 $f(\mathbf{e}_1, \mathbf{e}_1), \ldots, f(\mathbf{e}_n, \mathbf{e}_n)$ 中恰好有 r 个非零. 适当调整下标后, 我们可以得到一组新的规范

基 $\epsilon_1,\ldots,\epsilon_n$ 满足 $\{\epsilon_1,\ldots,\epsilon_n\}=\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$,

$f(\epsilon_i,\epsilon_i)\neq 0,\ i=1,2,\ldots,r,$ 且 $f(\epsilon_j,\epsilon_j)=0,\ j=r+1,r+2,\ldots,n.$

令 $\lambda_i=f(\epsilon_i,\epsilon_i)$, $i=1,2,\ldots,n$, 即可. □

推论 6.17 设 $A\in \mathrm{SM}_n(F)$ 且 $\mathrm{rank}(A)=r$. 则存在 F 中非零元素 $\lambda_1,\ldots,\lambda_r$ 使得 $A\sim_c \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_r,0,\cdots,0)$.

证明. 由上述推论直接可得. □

例 6.18 设 $A\in \mathrm{SM}_n(\mathbb{C})$ 且 $r=\mathrm{rank}(A)$. 则

$$A\sim_c \begin{pmatrix} E_r & O\\ O & O\end{pmatrix}.$$

证明. 由推论 *6.17*, $A\sim_c \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_r,0,\cdots,0)$, 其中 $\lambda_1,\ldots,\lambda_r$ 是非零复数. 由代数学基本定理 $\sqrt{\lambda_1},\ldots,\sqrt{\lambda_r}$ 是复数. 令

$$P=\mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}},\ldots,\frac{1}{\sqrt{\lambda_r}},\underbrace{1,\ldots,1}_{n-r}\right).$$

则 $P\in \mathrm{GL}_n(\mathbb{C})$ 且对称. 直接计算

$$\begin{aligned} A &\sim_c P^t\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_r,0,\cdots,0)P\\ &=P^t\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_r,0,\cdots,0)P\\ &=\begin{pmatrix} E_r & O\\ O & O\end{pmatrix}.\end{aligned}$$

注解 6.19 设 $f \in \mathcal{L}_2^+(\mathbb{C}^n)$ 且 $r = \operatorname{rank}(f)$. 则存在 $\mathbb{C}^n$ 的一组基 $\epsilon_1, \ldots, \epsilon_n$ 使得对任意 $\mathbb{C}^n$ 中向量 $\mathbf{x} = \sum_{i=1}^n x_i\epsilon_i$ 和 $\mathbf{y} = \sum_{i=1}^n y_i\epsilon_i$, 我们有

$$f(\mathbf{x}, \mathbf{y}) = x_1y_1 + \cdots + x_ry_r.$$

6.3 行列相伴消元

设 $F_{i,j}$ 是 n 阶第一类初等矩阵, $i, j \in \{1, \ldots, n\}$, $F_{i,j}(\lambda)$ 是第二类初等矩阵, 其中 $i, j \in \{1, \ldots, n\}$, $i \neq j$, $\lambda \in F$.

引理 6.20 设 $A = (a_{k,\ell}) \in \mathrm{SM}_n(F)$, $B = F_{i,j}^t A F_{i,j}$ 是对称矩阵且

$$B = \begin{array}{cc} \begin{array}{ccccccc} \phantom{a_{1,1}} & & \downarrow^{i} & & \downarrow^{j} & & \phantom{a_{1,n}} \end{array} & \\ \begin{pmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,i} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1} & \cdots & \underline{a_{j,j}} & \cdots & a_{j,i} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & \underline{a_{i,i}} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,i} & \cdots & a_{n,n} \end{pmatrix} & \begin{array}{l} \\ \\ \to i \\ \\ \to j \\ \\ \\ \end{array} \end{array}.$$

证明. 因为 $F_{i,j}$ 对称, 所以 $B \sim_c A$. 由 A 对称得出 B 对称. 由初等行变换可知 $F_{i,j}A$ 仅仅置换 A 的第 i 和第 j 行. 再由初等列变换可知 $(F_{i,j}A)F_{i,j}$ 仅仅置换 $F_{i,j}A$ 的第 i 和第 j 列. 从而, $B = F_{i,j}AF_{i,j}$. $\square$

引理 6.21 设 $A=(a_{k,\ell})\in \mathrm{SM}_n(F)$, $B=F_{i,j}(\lambda)^t AF_{i,j}(\lambda)$ 是对称矩阵且

$$B=\begin{array}{c}\begin{array}{ccccccc} & & \downarrow^i & & \downarrow^j & & \end{array}\\ \begin{pmatrix} a_{1,1} & \cdots & a_{1,i} & \cdots & a_{1,j}+\lambda a_{1,i} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,i} & \cdots & \underline{a_{i,j}+\lambda a_{i,i}} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j,1}+\lambda a_{i,1} & \cdots & \underline{a_{j,i}+\lambda a_{i,i}} & \cdots & \boxed{a_{j,j}+2\lambda a_{i,j}+\lambda^2 a_{i,i}} & \cdots & a_{j,n}+\lambda a_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i} & \cdots & a_{n,j}+\lambda a_{n,i} & \cdots & a_{n,n} \end{pmatrix}\begin{array}{l} \\ \\ \to i \\ \\ \to j \\ \\ \\ \end{array}\end{array}.$$

证明. 矩阵 B 显然对称. 由初等行变换可知 $F_{i,j}^t(\lambda)A$ 仅仅把 A 的第 i 行通乘 λ 后加到第 j 行上. 再由初等列变换可知 $(F_{i,j}(\lambda)^t A)F_{i,j}(\lambda)$ 仅仅把 A 的第 i 列通乘 λ 后加到第 j 列上. 从而, $B=F_{i,j}(\lambda)^t AF_{i,j}(\lambda)$. □

对对称矩阵做有限次上述两个引理中的操作得到的矩阵称为通过(初等)行列相伴变换得到的矩阵.

引理 6.22 设域 F 的特征不等于 2, $A\in \mathrm{SM}_n(F)$. 如果 A 中对角线上元素都等于零但 $A\neq O$, 则我们可以通过行列相伴变换把 A 变成对称矩阵 $B=(b_{i,j})$ 使得 $b_{1,1}\neq 0$. 特别地, $A\sim_c B$.

证明. 设 $A=(a_{k,\ell})_{n\times n}$, 其中某个 $a_{i,j}\neq 0$, 且 $i\neq j$. 则 $F_{i,j}(1)^t AF_{i,j}(1)$ 在第 j 行 j 列处的元素是

$$a_{j,j}+2a_{i,j}+a_{i,i}=2a_{i,j}\neq 0.$$

这是因为引理 6.21 和 $2\neq 0$. 根据引理 6.20,

$$B=F_{1,j}(F_{i,j}(1)^t AF_{i,j}(1))F_{1,j}. \quad \square$$

定理 6.23 设域 F 的特征不等于 2, $A \in \mathrm{SM}_n(F)$. 则我们可以通过初等行伴列变换得到对角矩阵.

证明. 我们对 n 归纳. 当 $n=1$ 时, A 是对角阵. 定理显然成立. 设 $n>1$ 且定理对 $n-1$ 成立. 我们考虑 n 阶对称矩阵 A. 如果 $A=O$, 则 A 已经是对角矩阵. 下面设 $A=(a_{i,j}) \neq O$.

由引理 6.22, 我们可以进一步假设 $a_{1,1} \neq 0$. 由引理 6.21,

$$F_{1,n}\left(-\frac{a_{n,1}}{a_{1,1}}\right)^t \cdots F_{1,2}\left(-\frac{a_{2,1}}{a_{1,1}}\right)^t A \underbrace{F_{1,2}\left(-\frac{a_{2,1}}{a_{1,1}}\right) \cdots F_{1,n}\left(-\frac{a_{n,1}}{a_{1,1}}\right)}_{M} = \begin{pmatrix} a_{1,1} & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & B \end{pmatrix}.$$

其中 $B \in \mathrm{SM}_{n-1}(F)$. 由归纳假设存在 $Q \in \mathrm{GL}_n(F)$ 是第一类和第二类初等矩阵之积使得 Q^tBQ 是对角矩阵. 令

$$P = \begin{pmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & Q \end{pmatrix}.$$

则 $(MP)^tA(MP)$ 是对角阵. □

例 6.24 设

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathrm{SM}_3(\mathbb{R}).$$

利用行列相伴变换把 A 化成对角阵 B, 并计算 $P \in \mathrm{GL}_3(\mathbb{R})$ 使得 $B = P^tAP$.

解.

$$(A|E)=\begin{pmatrix}0&1&1&|&1&0&0\\1&0&1&|&0&1&0\\1&1&0&|&0&0&1\end{pmatrix}$$

$$\xrightarrow{\text{把第}2\text{行加到第}1\text{行}}\begin{pmatrix}1&1&2&|&1&1&0\\1&0&1&|&0&1&0\\1&1&0&|&0&0&1\end{pmatrix}$$

$$\xrightarrow{\text{对称操作}}\begin{pmatrix}2&1&2&|&1&1&0\\1&0&1&|&0&1&0\\2&1&0&|&0&0&1\end{pmatrix}$$

$$\xrightarrow{\text{第}1\text{行通乘 }-\frac{1}{2}\text{ 加到第}2\text{行}}\begin{pmatrix}2&1&2&|&1&1&0\\0&\frac{-1}{2}&0&|&\frac{-1}{2}&\frac{1}{2}&0\\2&1&0&|&0&0&1\end{pmatrix}$$

$$\xrightarrow{\text{对称操作}}\begin{pmatrix}2&0&2&|&1&1&0\\0&\frac{-1}{2}&0&|&\frac{-1}{2}&\frac{1}{2}&0\\2&0&0&|&0&0&1\end{pmatrix}$$

$$\xrightarrow{\text{第}1\text{行通乘 }-1\text{ 加到第}3\text{行}}\begin{pmatrix}2&0&2&|&1&1&0\\0&\frac{-1}{2}&0&|&\frac{-1}{2}&\frac{1}{2}&0\\0&0&-2&|&-1&-1&1\end{pmatrix}$$

$$\xrightarrow{\text{对称操作}}\begin{pmatrix}2&0&0&|&1&1&0\\0&\frac{-1}{2}&0&|&\frac{-1}{2}&\frac{1}{2}&0\\0&0&-2&|&-1&-1&1\end{pmatrix}$$

设

$$P=\begin{pmatrix}1&\frac{-1}{2}&-1\\1&\frac{1}{2}&-1\\0&0&1\end{pmatrix}\implies P^tAP=\operatorname{diag}\left(2,-\frac{1}{2},-2\right).$$