# 第一章　空间与形式

## 9　实二次型

在本节中, $V$ 是 $\mathbb{R}$ 上的 $n$ 维线性空间.

### 9.1　惯性定理

**定理 9.1** *(Sylvester)* 设 $q$ 是 $V$ 上的二次型. 则存在 $q$ 的一组规范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得在该基下 $q$ 的矩阵为

$$\begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix}$$

且 $k + \ell = \operatorname{rank}(q)$. 进而, 如果 $q$ 在另一组规范基下的矩阵是

$$\begin{pmatrix} E_s & O & O \\ O & -E_t & O \\ O & O & O \end{pmatrix}$$

则 $k = s$ 和 $t = \ell$.

**证明.** 设 $r = \operatorname{rank}(q)$. 由上一节推论 7.17 及其证明可知, 存在 $q$ 的一组基使得 $q$ 在该基下的矩阵是

$$A = \operatorname{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0),$$

其中 $r=\operatorname{rank}(q)$, $\lambda_1,\ldots,\lambda_r\in\mathbb{R}\setminus\{0\}$. 适当调整基底中元素的顺序, 我们可进一步设

$$\lambda_1,\ldots,\lambda_k\in\mathbb{R}^+,\quad \lambda_{k+1},\ldots,\lambda_{k+\ell}\in\mathbb{R}^-,\quad \text{且}\quad k+\ell=r.$$

令 $P$ 为 $n$ 阶可逆矩阵

$$\operatorname{diag}\left((\sqrt{\lambda_1})^{-1},\ldots,(\sqrt{\lambda_k})^{-1},(\sqrt{-\lambda_{k+1}})^{-1},\ldots,(\sqrt{-\lambda_{k+\ell}})^{-1},1,\ldots,1\right).$$

则

$$P^tAP=\begin{pmatrix}E_k & O & O\\ O & -E_\ell & O\\ O & O & O\end{pmatrix}.$$

设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\epsilon_1,\ldots,\epsilon_n$ 是两组 $q$ 的规范基, 所对应的矩阵分别是

$$B=\begin{pmatrix}E_k & O & O\\ O & -E_\ell & O\\ O & O & O\end{pmatrix}\quad \text{和}\quad C=\begin{pmatrix}E_s & O & O\\ O & -E_t & O\\ O & O & O\end{pmatrix}.$$

设 $\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n=y_1\epsilon_1+\cdots+y_n\epsilon$. 则

$$q(\mathbf{x})=x_1^2+\cdots+x_k^2-x_{k+1}^2-\cdots-x_r^2=y_1^2+\cdots+y_s^2-y_{s+1}^2-\cdots-y_r^2.$$

假设 $k>s$. 则 $\ell<t$. 这是因为 $k+\ell=s+t=r$. 令 $U=\langle\mathbf{e}_1,\ldots,\mathbf{e}_k\rangle$, $W=\langle\epsilon_{s+1},\ldots,\epsilon_n\rangle$. 则

$$\dim(U\cap W)=\dim(U)+\dim(W)-\dim(U+W)\geq k+n-s-n=k-s>0.$$

于是由非零向量 $\mathbf{v} \in U \cap W$. 由 $U$ 和 $W$ 的生成元可知, 存在 $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, 不全为零, 和 $\beta_{s+1}, \ldots, \beta_n \in \mathbb{R}$ 使得

$$\mathbf{v} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_k \mathbf{e}_k = \beta_{s+1} \epsilon_{s+1} + \cdots + \beta_n \epsilon_n.$$

于是

$$q(\mathbf{v}) = \alpha_1^2 + \cdots + \alpha_k^2 > 0 \quad \text{且} \quad q(\mathbf{v}) = -\beta_{s+1}^2 - \cdots - \beta_r^2 \leq 0.$$

矛盾. □

利用上述定理中的记号, 我们有如下定义.

**定义 9.2** 称 $k$ 是 $q$ 的正惯性指数, $\ell$ 是 $q$ 的负惯性指数, $(k, \ell)$ 是 $q$ 的签名.

上述定理的矩阵版如下.

**推论 9.3** 设 $A \in \mathrm{SM}_n(\mathbb{R})$. 则存在 $k, \ell \in \mathbb{N}$ 使得

$$A \sim_c \begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix}$$

且 $k + \ell = \mathrm{rank}(q)$. 进而, 如果

$$A \sim_c \begin{pmatrix} E_s & O & O \\ O & -E_t & O \\ O & O & O \end{pmatrix}$$

则 $k = s$ 和 $t = \ell$.

利用上述推论中的记号, 我们由如下定义.

**定义 9.4** 称 $k$ 是 $A$ 的正惯性指数, $\ell$ 是 $A$ 的负惯性指数, $(k,\ell)$ 是 $A$ 的签名.

**推论 9.5** 设 $A, B \in \mathrm{SM}_n(\mathbb{R})$. 则 $A \sim_c B$ 当且仅当 $A$ 和 $B$ 有共同的签名.

**证明.** 设 $A$ 的签名是 $(k,\ell)$, $B$ 的签名是 $(s,t)$.

如果 $A \sim_c B$, 则由 $\sim_c$ 的传递律和对称律得出

$$A \sim_c \begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix} \Longrightarrow B \sim_c \begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix}.$$

根据推论 9.3, $(k,\ell)$ 也是 $B$ 的签名.

反之，设 $(k,\ell) = (s,t)$. 则

$$A \sim_c \begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix} \quad \text{和} \quad B \sim_c \begin{pmatrix} E_k & O & O \\ O & -E_\ell & O \\ O & O & O \end{pmatrix}.$$

因为 $\sim_c$ 是等价关系, 所以 $A \sim_c B$. $\square$

**例 9.6** 设

$$\begin{array}{rccc} q: & \mathrm{M}_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & A & \mapsto & \mathrm{tr}(A^2). \end{array}$$

证明 $q$ 是二次型并求其签名.

**证明.** 设 $A=(a_{i,j})$. 则

$$q(A)=\mathrm{tr}(A^2)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{i,j}a_{j,i}=\sum_{i=1}^{n}a_{i,i}^2+\sum_{1\le i<j\le n}2a_{i,j}a_{j,i}.$$

于是, $q$ 是二次型. 考虑可逆坐标变换

$$z_{i,i}=a_{i,i},\ i=1,2,\ldots,n,\quad \text{和}\quad a_{i,j}=z_{i,j}+z_{j,i},\ a_{j,i}=z_{i,j}-z_{j,i},$$

其中 $i,j\in\{1,2,\ldots,n\}, i\neq j$. 则

$$q=\sum_{i=1}^{n}z_{i,i}^2+\sum_{1\le i<j\le n}2(z_{i,j}^2-z_{j,i}^2).$$

于是, $q$ 的正惯性指数是

$$\frac{n(n+1)}{2}$$

而负惯性指数是

$$\frac{n(n-1)}{2}.$$

从而 $q$ 的签名为

$$\left(\frac{n(n+1)}{2},\ \frac{n(n-1)}{2}\right)$$

## 9.2 （半）正定二次型

**定义 9.7** 设 $q$ 是 $V$ 上的二次型.

*(i)* 如果对于任意 $\mathbf{x} \in V$, $q(\mathbf{x}) \geq 0$, 则称 $q$ 是半正定的 *(semi-positive definite);*

*(ii)* 如果对于任意 $\mathbf{x} \in V \setminus \{\mathbf{0}\}$, $q(\mathbf{x}) > 0$, 则称 $q$ 是正定的 *(positive definite);*

*(iii)* 如果对于任意 $\mathbf{x} \in V$, $q(\mathbf{x}) \leq 0$, 则称 $q$ 是半负定的 *(semi-negative definite);*

*(iv)* 如果对于任意 $\mathbf{x} \in V \setminus \{\mathbf{0}\}$, $q(\mathbf{x}) < 0$, 则称 $q$ 是负定的 *(negative definite);*

*(v)* 如果 $q$ 既不是半正定也不是半负定的, 则称 $q$ 是不定的 *(indefinite).*

设 $A \in \mathrm{SM}_n(\mathbb{R})$, $q_A$ 是 $\mathbb{R}^n$ 上在标准基下矩阵为 $A$ 的二次型. 如果 $q_A$ 是半正定(正定的, 半负定的, 负定的, 不定的), 则称 $A$ 是半正定(正定的, 半负定的, 负定的, 不定的).

**命题 9.8** 设 $\dim(V) = n$ 且 $q$ 是 $V$ 上的二次型. 它的签名是 $(k, \ell)$.

*(i)* $q$ 是半正定的当且仅当 $\ell = 0$;

*(ii)* $q$ 是正定的当且仅当 $k = n$;

*(iii)* $q$ 是半负定的当且仅当 $k = 0$;

*(iv)* $q$ 是负定的当且仅当 $\ell = n$;

*(v)* $q$ 是不定的当且仅当 $k > 0$ 且 $\ell > 0$.

**证明.** 设 $q$ 在规范基 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的规范型是

$$q(\mathbf{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+\ell}^2,$$

其中 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ 是 $V$ 中的任意向量.

(i) 若 $\ell = 0$, 则 $q(\mathbf{x}) = x_1^2 + \cdots + x_k^2 \geq 0$. 即 $q$ 半正定. 反之, 假设 $\ell > 0$. 令 $\mathbf{x} = \mathbf{e}_{k+1}$. 则 $q(\mathbf{x}) = -1 < 0$. 矛盾. 故 $\ell = 0$.

(ii) 若 $k = n$, 则 $q(\mathbf{x}) = x_1^2 + \cdots + x_n^2$. 于是, 当 $\mathbf{x} \neq \mathbf{0}$ 时, $q(\mathbf{x}) > 0$, 即 $q$ 正定. 反之, 假设 $k < n$. 令 $\mathbf{x} = \mathbf{e}_n$. 则 $q(\mathbf{e}_n) = 0$. 于是 $q$ 不是正定的. 矛盾.

(iii) 与 (i) 类似.

(iv) 与 (ii) 类似.

(v) 排除 (i), (iii) 情形即可. □

**注解 9.9** 半正定, 正定, 半负定, 负定和不定二次型分别有下列规范型

$$x_1^2 + \cdots + x_k^2, \quad x_1^2 + \cdots + x_n^2, \quad -x_1^2 - \cdots - x_\ell^2, \quad -x_1^2 - \cdots - x_n^2,$$

和

$$x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+\ell}^2,$$

在最后的规范型中 $k > 0, \ell > 0$.

类似地, 我们有

**命题 9.10** 设 $A \in \mathrm{SM}_n(\mathbb{R})$. 它的签名是 $(k, \ell)$.

*(i)* $A$ *是半正定的当且仅当* $\ell = 0$*;*

*(ii)* $A$ *是正定的当且仅当* $k = n$*;*

*(iii)* $A$ *是半负定的当且仅当* $k = 0$*;*

*(iv)* $A$ *是负定的当且仅当* $\ell = n$*;*

*(v)* $A$ *是不定的当且仅当* $k > 0$ *且* $\ell > 0$*.*

**例 9.11** 证明 $(a) \in \mathrm{M}_1(\mathbb{R})$ 是正定的当且仅当 $a > 0$. 判定

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

是不是正定的.

**证明.** 矩阵 $(a)$ 对应 $\mathbb{R}$ 上的二次型 $q(x) = ax^2$. 而 $q$ 正定当且仅当 $a > 0$.

利用行列相伴变换或 *Jacobi* 公式可得

$$A \sim_c \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \Longrightarrow A \text{ 的签名是 } (1, 1).$$

于是 $A$ 是不定的.

**例 9.12** 设 $p,q\in\mathcal{Q}(V)$是半正定的. 证明 $p+q$ 也是半正定的; 若上述 $p,q$ 中还有一个是正定的, 则 $p+q$ 也是正定的.

**证明.** 设 $p,q$ 是半正定的. 则对任意的 $\mathbf{x}\in V$, $p(\mathbf{x})\geq 0$, $q(\mathbf{x})\geq 0$. 于是

$$(p+q)(\mathbf{x})=p(\mathbf{x})+q(\mathbf{x})\geq 0.$$

再设 $p$ 是正定的. 则对于任意的 $\mathbf{x}\in V\setminus\{\mathbf{0}\}$, $p(\mathbf{x})>0$. 于是

$$(p+q)(\mathbf{x})=p(\mathbf{x})+q(\mathbf{x})>0. \quad \square$$

类似地可证明下列结论.

**注解 9.13** 设 $A,B\in\mathrm{SM}_n(\mathbb{R})$ 是半正定的, 则 $A+B$ 也是半正定的; 进一步设 $A,B$ 中还有一个是正定的, 则 $A+B$ 也是正定的.

设 $q\in\mathcal{Q}(V)$. 定义

$$C_q=\{\mathbf{v}\in V\mid q(\mathbf{v})=0\}.$$

称 $C_q$ 为 $q$ 确定的锥面(cone).

**例 9.14** 设 $q\in\mathcal{Q}(V)$. 证明 $C_q$ 是 $V$ 的子空间当且仅当 $q$ 是半正定或半负定的.

**证明.** 设 $q$ 的签名是 $(k,\ell)$.

设 $C_q$ 是 $V$ 中的子空间. 如果 $k>0$ 且 $\ell>0$. 根据惯性定理 $q$ 在 $V$ 的某组基 $\mathbf{e}_1,\ldots,\mathbf{e}_k,\mathbf{e}_{k+1},\ldots,\mathbf{e}_n$ 下的规范型是 $q(\mathbf{x})=x_1^2+\cdots+x_k^2-x_{k+1}^2+\cdots+x_{k+\ell}^2$, 其中 $\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_k\mathbf{e}_k+x_{k+1}\mathbf{e}_{k+1}+\cdots+x_n\mathbf{e}_n$. 于是 $q(\mathbf{e}_k+\mathbf{e}_{k+1})=1-1=0$ 且 $q(\mathbf{e}_k-\mathbf{e}_{k+1})=0$. 由此可知

$$\mathbf{e}_k+\mathbf{e}_{k+1},\mathbf{e}_k-\mathbf{e}_{k+1}\in C_q \implies 2\mathbf{e}_k\in C_q \implies q(2\mathbf{e}_k)=0 \implies 4=0.$$

矛盾. 于是 $k=0$ 或 $\ell=0$. 即 $q$ 是半负定或半正定的.

反之, 不妨设 $q$ 是半正定的. 根据惯性定理 $q$ 在 $V$ 的某组基 $\mathbf{e}_1,\ldots,\mathbf{e}_k,\mathbf{e}_{k+1},\ldots,\mathbf{e}_n$ 下的规范型是

$$q(\mathbf{x})=x_1^2+\cdots+x_k^2.$$

设 $U=\langle\mathbf{e}_{k+1},\ldots,\mathbf{e}_n\rangle$. 则 $U\subset C_q$. 设 $\mathbf{v}=v_1\mathbf{e}_1+\cdots+v_n\mathbf{e}_n\in C_q$. 则

$$q(\mathbf{v})=v_1^2+\cdots+v_k^2=0.$$

因为 $v_1,\ldots,v_k\in\mathbb{R}$, 所以 $v_1=\cdots=v_k=0$. 于是

$$\mathbf{v}=v_{k+1}\mathbf{e}_{k+1}+\cdots+v_n\mathbf{e}_n\in U,$$

即 $C_q\subset U$. 综上所述, $U=C_q$. $\square$

## 9.3 （半）正定矩阵的等价条件

实数域的一个基本性质是：设 $x_1,\ldots,x_n\in\mathbb{R}$. 则

$$x_1^2+\cdots+x_n^2\geq 0;\quad x_1^2+\cdots+x_n^2=0 \implies x_1=\cdots=x_n=0.$$

这个性质的矩阵版如下: 设

$$\mathbf{x}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in\mathbb{R}^n,\ \mathbf{x}^t\mathbf{x}\geq 0;\quad \mathbf{x}^t\mathbf{x}=0\Longrightarrow\mathbf{x}=\mathbf{0}.$$

**引理 9.15** 设 $A\in\mathbb{R}^{m\times n}$. 则 $A^tA\in\mathrm{SM}_n(\mathbb{R})$, 半正定且 $\mathrm{rank}(A^tA)=\mathrm{rank}(A)$.

**证明.** 计算 $(A^tA)^t=A^t(A^t)^t=A^tA$ 得出 $A^tA\in\mathrm{SM}_n(\mathbb{R})$.

设 $\mathbf{x}=(x_1,\ldots,x_n)^t\in\mathbb{R}^n$, $\mathbf{y}=A\mathbf{x}\in\mathbb{R}^m$. 则

$$\mathbf{x}^t(A^tA)\mathbf{x}=\mathbf{y}^t\mathbf{y}\geq 0.$$

故 $A^tA$ 半正定.

设 $U,V$ 是分别是以 $A$ 和 $A^tA$ 为系数的齐次线性方程组的解空间. 则 $U\subset V$. 反之, 设 $\mathbf{x}\in V$. 则

$$A^tA\mathbf{x}=\mathbf{0}_n\in\mathbb{R}^n.$$

于是, $\mathbf{y}^t\mathbf{y}=0$. 由此得出 $\mathbf{y}=\mathbf{0}_m\in\mathbb{R}^m$, 即 $\mathbf{x}\in U$. 由此得出 $U=V$. 特别地, $\dim(U)=\dim(V)$. 由方程组版的对偶定理, $\mathrm{rank}(A)=\mathrm{rank}(A^tA)$. $\square$

**定理 9.16** 设 $A\in\mathrm{SM}_n(\mathbb{R})$. 则

*(i)* $A$ 半正定当且仅当存在 $B\in\mathrm{M}_n(\mathbb{R})$ 使得 $A=B^tB$.

*(ii)* $A$ 正定当且仅当存在 $B\in\mathrm{GL}_n(\mathbb{R})$ 使得 $A=B^tB$.

**证明.** (i) 设 $A$ 半正定. 则由矩阵版的惯性定理存在 $P \in \mathrm{GL}_n(\mathbb{R})$ 使得

$$P^t A P = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix},$$

其中 $r = \mathrm{rank}(A)$, 也是 $A$ 的正惯性指数. 于是

$$A = \left(P^{-1}\right)^t \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} P^{-1} = \left(P^{-1}\right)^t \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} \underbrace{\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} P^{-1}}_{B} = B^t B.$$

反之, 根据引理 9.15, $A = B^t B$ 是半正定的.

(ii) 设 $A$ 正定. 则由矩阵版的惯性定理存在 $P \in \mathrm{GL}_n(\mathbb{R})$ 使得 $P^t A P = E_n$. 于是, $A = \left(P^{-1}\right)^t P^{-1}$. 反之, 根据引理 9.15, $A = B^t B$ 是半正定的. 因为 $B$ 可逆, 所以 $A$ 满秩. 于是, $A$ 的正惯性指数等于 $n$. □

**例 9.17** 设 $A$ 是正定矩阵. 证明 $\det(A) > 0$ 且 $A^{-1}$ 也正定.

**证明.** 由上述定理 *(ii)*, $A = P^t P$, 其中 $P \in \mathrm{GL}_n(\mathbb{R})$. 则 $\det(A) = \det(P)^2 > 0$, 且

$$A^{-1} = (P^t P)^{-1} = P^{-1} \left(P^t\right)^{-1} = P^{-1} \left(P^{-1}\right)^t.$$

再由上述定理 *(ii)*, $A^{-1}$ 正定. □

**定理 9.18** 设 $A \in \mathrm{SM}_n(\mathbb{R})$. 设 $\Delta_k$ 是 $A$ 的 $k$ 阶顺序主子式, $k = 1, 2, \ldots, n$. 则下列命题等价.

*(i)* $A$ *正定;*

*(ii)* $A$ *的任何* $k$ *阶主子式都大于零;*

*(iii)* $\Delta_1 > 0,\ \ldots,\ \Delta_n > 0.$

**证明.** (i) $\Longrightarrow$ (ii) 设 $B$ 是由 $A$ 中第 $i_1, \ldots, i_k$ 行和 $i_1, \ldots, i_k$ 列构成的子矩阵. 其中 $1 \le i_1 < \cdots < i_k \le n$. 令 $\mathbf{x} \in \mathbb{R}^n$, 对任意 $j \in \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\}$, 第 $j$ 个坐标等于零, 第 $i_\ell$ 个坐标等于 $x_{i_\ell}$, $\ell = 1, 2, \ldots, k$. 再令 $\mathbf{y} = (x_{i_1}, \ldots, x_{i_k})^t$. 假设 $\mathbf{y} \ne \mathbf{0}_k$. 则 $\mathbf{x} \ne \mathbf{0}_n$. 于是

$$0 < \mathbf{x}^t A \mathbf{x} = \mathbf{y}^t B \mathbf{y}.$$

由 $\mathbf{y}$ 的任意性可知, $B$ 正定. 根据上例, $\det(B) > 0$.

(ii) $\Longrightarrow$ (iii) 显然.

(iii) $\Longrightarrow$ (i) 由 Jacobi 公式,

$$A \sim_c \operatorname{diag}\left(\frac{\Delta_1}{\Delta_0}, \frac{\Delta_2}{\Delta_1}, \ldots, \frac{\Delta_n}{\Delta_{n-1}}\right),$$

其中 $\Delta_0 = 1$. 于是 $A$ 合同于一个对角矩阵, 其对角线上的元素都是正实数. 于是 $A$ 的正惯性指数等于 $n$. $\square$

**例 9.19** 设

$$A = \begin{pmatrix} \lambda & -1 & 1 \\ -1 & \lambda & -1 \\ 1 & -1 & \lambda \end{pmatrix}$$

问 $\lambda$ 为何值时 $A$ 是正定的, $A$ 是负定的?

**解.** $A$ 的三个顺序主子式分别是

$$\Delta_1 = \lambda,\ \Delta_2 = \det\begin{pmatrix}\lambda & -1\\ -1 & \lambda\end{pmatrix} = \lambda^2 - 1,\ \Delta_3 = \det(A) = (\lambda-1)^2(\lambda+2).$$

$A$ 正定当且仅当 $\Delta_1 > 0$, $\Delta_2 > 0$ 且 $\Delta_3 > 0$. 即 $\lambda > 1$.

$A$ 负定当且仅当 $-A$ 正定. 而 $-A$ 的三个主子式是

$$\Omega_1 = -\lambda, \quad \Omega_2 = \lambda^2 - 1, \quad \Omega_3 = \det(A) = -(\lambda-1)^2(\lambda+2).$$

$A$ 负定当且仅当 $\Omega_1 > 0$, $\Omega_2 > 0$ 且 $\Omega_3 > 0$. 即 $\lambda < -2$.