第三章　内积空间

1.3　单位正交基

设 $\dim(V)=n$, $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 中两两正交的单位向量. 称为 V 的一组单位正交基. 根据第三章第一讲引理 1.15 (ii), $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基.

例 1.17 在标准欧式空间 $\mathbb{R}^n$ 中, 标准基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是一组标准正交基. 在 $\mathbb{R}^2$ 中,

$$\mathbf{u}=(\cos(\theta),\sin(\theta))^t,\quad \mathbf{v}=(-\sin(\theta),\cos(\theta))$$

是一组标准正交基.

定理 1.18 *(Gram-Schmidt 正交化)* 设 $\mathbf{v}_1,\dots,\mathbf{v}_k\in V$ 线性无关. 则存在两两正交的单位向量 $\epsilon_1,\dots,\epsilon_k$ 使得

$$\langle \mathbf{v}_1,\dots,\mathbf{v}_i\rangle=\langle \epsilon_1,\dots,\epsilon_i\rangle,$$

$i=1,2,\dots,k$. 特别地, V 有单位正交基.

证明. 对 k 归纳. 当 $k=1$ 时, 取 ϵ_1 为 $\mathbf{v}_1$ 的单位化向量即可. 设存在两两正交的单位向量 $\epsilon_1,\dots,\epsilon_{i-1}$ 使得

$$\langle \mathbf{v}_1,\dots,\mathbf{v}_{i-1}\rangle=\langle \epsilon_1,\dots,\epsilon_{i-1}\rangle.$$

令

$$\epsilon_i'=\mathbf{v}_i-(\mathbf{v}_i|\epsilon_1)\epsilon_1-\cdots-(\mathbf{v}_i|\epsilon_{i-1})\epsilon_{i-1}. \tag{1}$$

我们先来验证

$$\underbrace{\langle \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i \rangle}_{V_i} = \underbrace{\langle \epsilon_1, \ldots, \epsilon_{i-1}, \epsilon_i' \rangle}_{W_i'}. \tag{2}$$

根据 (1), $\mathbf{v}_i \in W_i'$. 而归纳假设蕴含 $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1} \in W_i'$. 故 $V_i \subset W_i'$. 反之，归纳假设蕴含 $\langle \epsilon_1, \ldots, \epsilon_{i-1} \rangle \subset V_i$, 而 (1) 蕴含 $\epsilon_i' \in \langle \epsilon_1, \ldots, \epsilon_{i-1}, \mathbf{v}_i \rangle$. 故 $\epsilon_i' \in V_i$. 由此得出 $W_i' \subset V_i$. 等式 (2) 成立. 特别地，我们有 $\dim(W_i') = i$. 故 $\epsilon_i' \neq \mathbf{0}$.

我们利用 (1) 计算得：

$$(\epsilon_i' | \epsilon_j) = (\mathbf{v}_i | \epsilon_j) - \sum_{\ell=1}^{k-1} (\mathbf{v}_k | \epsilon_\ell)(\epsilon_\ell | \epsilon_j) = (\mathbf{v}_i | \epsilon_j) - (\mathbf{v}_i | \epsilon_j) = 0.$$

故 ϵ_i' 与 ϵ_j 正交. 令 ϵ_i 是 ϵ_i' 的单位化向量. 则 $\epsilon_1, \ldots, \epsilon_{i-1}, \epsilon_i$ 是两两正交的单位向量. 根据 (2), $\langle \mathbf{v}_1, \ldots, \mathbf{v}_i \rangle = \langle \epsilon_1, \ldots, \epsilon_i \rangle$.

□

例 1.19 设 $\mathbb{R}^4$ 是标准欧式空间,

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

计算子空间 $U = \langle \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \rangle$ 的一组单位正交基.

解. 由 *Gram-Schmidt* 正交化得

$$\epsilon_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{2}} \mathbf{u}_1.$$

$$\epsilon_2' = \mathbf{u}_2 - (\mathbf{u}_2|\epsilon_1)\epsilon_1 = \mathbf{u}_2 - \|\mathbf{u}_1\|^{-2}(\mathbf{u}_2|\mathbf{u}_1)\mathbf{u}_1 = \mathbf{u}_2 - \frac{1}{2}\mathbf{u}_1 = \begin{pmatrix} -\frac{1}{2} \\ -1 \\ \frac{1}{2} \\ -1 \end{pmatrix}.$$

$$\epsilon_2 = \frac{\epsilon_2'}{\|\epsilon_2\|} = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 \\ -2 \\ 1 \\ -2 \end{pmatrix}.$$

$$\epsilon_3' = \mathbf{u}_3 - (\mathbf{u}_3|\epsilon_1)\epsilon_1 - (\mathbf{u}_3|\epsilon_2)\epsilon_2 = \frac{1}{5} \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \end{pmatrix}.$$

$$\epsilon_3 = \frac{1}{\sqrt{10}} \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \end{pmatrix}.$$

于是, U 得一组单位正交基是 $\epsilon_1, \epsilon_2, \epsilon_3$. □

注解 1.20 定理 *1.18* 的证明称为 *Gram-Schmidt* 正交化. 该定理说明

$$\epsilon_i \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_i \rangle \quad 和 \quad \mathbf{v}_i = \langle \epsilon_1, \ldots, \epsilon_i \rangle.$$

故存在上三角的矩阵 $S, T \in \mathrm{M}_k(\mathbb{R})$ 使得

$(\epsilon_1, \ldots, \epsilon_k) = (\mathbf{v}_1, \ldots, \mathbf{v}_k)S$ 和 $(\mathbf{v}_1, \ldots, \mathbf{v}_k) = (\epsilon_1, \ldots, \epsilon_k)T.$

因为 $\mathbf{v}_1, \ldots, \mathbf{v}_k$ 和 $\epsilon_1, \ldots, \epsilon_k$ 是两组线性无关的向量, 所以 S, T 都是可逆的. 下面证明:

$$T = \begin{pmatrix} (\mathbf{v}_1|\epsilon_1) & (\mathbf{v}_2|\epsilon_1) & \cdots & (\mathbf{v}_k|\epsilon_1) \\ & (\mathbf{v}_2|\epsilon_2) & \cdots & (\mathbf{v}_k|\epsilon_2) \\ & & \ddots & \vdots \\ & & & (\mathbf{v}_k|\epsilon_k) \end{pmatrix}. \tag{3}$$

设 $T = (t_{i,j})_{k\times k}$. 因为 T 是上三角形矩阵, 所以

$$\mathbf{v}_j = t_{1,j}\epsilon_1 + \cdots + t_{j,j}\epsilon_j.$$

设 $i \in \{1, 2, \ldots, j\}$. 我们计算

$$(\mathbf{v}_j|\epsilon_i) = \left(\sum_{\ell=1}^{j} t_{\ell,j}(\epsilon_\ell|\epsilon_i)\right) = \sum_{\ell=1}^{j} t_{\ell,j}(\epsilon_\ell|\epsilon_i) = t_{i,j}.$$

于是, (3) 成立. □

命题 1.21 设 V 的一组单位正交基是 $\mathbf{e}_1, \ldots, \mathbf{e}_n$, $\mathbf{x}, \mathbf{y} \in V$ 在这组基下的坐标分别是 $(x_1, \ldots, x_n)^t$ 和 $(y_1, \ldots, y_n)^t$. 则

$$(\mathbf{x}|\mathbf{y}) = x_1y_1 + \cdots + x_ny_n.$$

证明. 因为 $\mathrm{G}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = E_n$, 所以第三章第一讲命题 1.7 蕴含结论. □

命题 1.22 设 V 的一组单位正交基是 $\mathbf{e}_1, \ldots, \mathbf{e}_n$, $\mathbf{x} \in V$. 则 $\mathbf{x}$ 在该基下的第 i 个坐标分量是 $(\mathbf{x}|\mathbf{e}_i)$, $i = 1, 2, \ldots, n$.

证明. 设 $\mathbf{x}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标是 $(x_1, \ldots, x_n)^t$. 则

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

定理 1.23 设 V 和 W 是两个 n-维欧氏空间, 其中的内积分别记为 $(\,|\,)_V$ 和 $(\,|\,)_W$. 则存在线性同构 $\phi : V \longrightarrow W$ 满足对任意 $\mathbf{x}, \mathbf{y} \in V$,

$$(\mathbf{x}|\mathbf{y})_V = (\phi(\mathbf{y})|\phi(\mathbf{y}))_W.$$

证明. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组单位正交基, 而 $\epsilon_1, \ldots, \epsilon_n$ 是 W 的一组单位正交基. 则存在线性映射 ϕ 使得 $\phi(\mathbf{e}_i) = \epsilon_i$, $i = 1, 2, \ldots, n$. (第一章第二讲定理 4.11). 进而, ϕ 是线性同构(第一章第二讲定理 4.12 的证明). 设 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ 和 $\mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_n\mathbf{e}_n$. 根据命题 1.21,

$$(\mathbf{x}|\mathbf{y})_V = x_1y_1 + \cdots + x_ny_n.$$

因为 $\phi(\mathbf{x}) = x_1\epsilon_1 + \cdots + x_n\epsilon_n$ 和 $\phi(\mathbf{y}) = y_1\epsilon_1 + \cdots + y_n\epsilon_n$, 所以命题 1.21 蕴含

$$(\phi(\mathbf{x})|\phi(\mathbf{y}))_W = x_1y_1 + \cdots + x_ny_n.$$

我们有 $(\mathbf{x}|\mathbf{y})_V = (\phi(\mathbf{x})|\phi(\mathbf{y}))_W$. $\square$

2 正交补

定义 2.1 设 $U_1, U_2 \subset V$ 是子空间. 如果对于任意的 $\mathbf{u}_1 \in U_1$ 和 $\mathbf{u}_2 \in U_2$ 我们有 $\mathbf{u}_1 \perp \mathbf{u}_2$, 则称 U_1 和 U_2 正交, 记为 $U_1 \perp U_2$.

记号. 设 $\mathbf{x} \in V$ 且 $S \subset V$. 如果 $\mathbf{x}$ 与 S 中元素都正交, 则记为 $\mathbf{x} \perp S$.

定理 2.2 设 $U \subset V$ 是子空间. 令 $U^{\perp} := \{\mathbf{x} \in V \mid \mathbf{x} \perp U\}$. 则

(i) $U^{\perp}$ 是子空间且 $U \perp U^{\perp}$;

(ii) $V = U \oplus U^{\perp}$ (称 $U^{\perp}$ 是 U 的正交补).

(iii) $(U^{\perp})^{\perp} = U$.

证明. (i) 设 $\mathbf{x}, \mathbf{y} \in U^{\perp}$ 和 $\alpha, \beta \in \mathbb{R}$. 则对任意 $\mathbf{u} \in U$,

$$(\alpha\mathbf{x} + \beta\mathbf{y}|\mathbf{u}) = \alpha(\mathbf{x}|\mathbf{u}) + \beta(\mathbf{y}|\mathbf{u}) = 0.$$

于是, $(\alpha\mathbf{x} + \beta\mathbf{y}) \perp \mathbf{u}$. 我们得到 $\alpha\mathbf{x} + \beta\mathbf{y} \in U^{\perp}$. 由 $U^{\perp}$ 的定义可知, $U \perp U^{\perp}$.

(ii) 设 $\mathbf{x} \in U \cap U^{\perp}$. 则 $\mathbf{x} \perp \mathbf{x} = 0$. 于是, $\mathbf{x} = \mathbf{0}$ (第三章第一讲引理 1.15 (i)). 我们只要证明 $V = U + U^{\perp}$ 即可. 如果 $U = \{\mathbf{0}\}$, 则 $U^{\perp} = V$. 结论显然成立. 设 $U \neq \{\mathbf{0}\}$. 根据定理 1.18, U 有单位正交基. 设其为 $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$. 对任意向量 $\mathbf{x} \in V$, 令

$$\mathbf{y} = (\mathbf{x}|\mathbf{e}_1)\mathbf{e}_1 + \cdots + (\mathbf{x}|\mathbf{e}_d)\mathbf{e}_d$$

和 $\mathbf{z} = \mathbf{x} - \mathbf{y}$. 因为 $\mathbf{y} \in U$, 所以我们只要证明 $\mathbf{z} \in U^{\perp}$ 即可. 设 $\mathbf{u}$ 是 U 中任意向量. 则存在 $\alpha_1, \ldots, \alpha_d \in \mathbb{R}$, 使得 $\mathbf{u} = \alpha_1\mathbf{e}_1 + \cdots + \alpha_d\mathbf{e}_d$. 于是,

$$(\mathbf{z}|\mathbf{u}) = (\mathbf{x}-\mathbf{y}|\mathbf{u}) = (\mathbf{x}|\mathbf{u}) - (\mathbf{y}|\mathbf{u}) = \sum_{i=1}^{d} \alpha_i(\mathbf{x}|\mathbf{e}_i) - \sum_{i=1}^{d} \alpha_i(\mathbf{x}|\mathbf{e}_i) = 0.$$

(iii) 由正交补得定义可知 $U \subset (U^{\perp})^{\perp}$. 根据 (ii),

$$V = U \oplus U^{\perp} = U^{\perp} \oplus (U^{\perp})^{\perp}.$$

于是, $\dim(U) = \dim((U^{\perp})^{\perp})$ (第一章第二讲命题 4.16). 根据第一章第二讲命题 4.15 (i), 我们得到 $U = (U^{\perp})^{\perp}$. □

推论 2.3 设 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 是 V 中的单位正交向量. 则 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 可扩充为 V 的一组单位正交基.

证明. 设 $U = \langle \mathbf{e}_1, \ldots, \mathbf{e}_d \rangle$. 则 $U^{\perp}$ 有一组单位正交基 $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$. 根据定理 2.2, $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 是 V 的一组单位正交基. □

注解 2.4 上述推论也可以由 *Gram-Schmidt* 正交化直接得出. 这是因为从 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 出发可以得到 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{v}_{d+1}, \ldots, \mathbf{v}_n$. 对这组基做 *Gram-Schmidt* 正交化得到的单位正交基的前 d 个元素仍然是 $\mathbf{e}_1, \ldots, \mathbf{e}_d$.

例 2.5 设标准欧式空间 $\mathbb{R}^3$ 的标准基是 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. 则 $\langle \mathbf{e}_1 \rangle^{\perp} = \langle \mathbf{e}_2, \mathbf{e}_3 \rangle$.

$$\langle \mathbf{e}_1 \rangle^{\perp\perp} = \langle \mathbf{e}_2, \mathbf{e}_3 \rangle^{\perp} = \langle \mathbf{e}_1 \rangle. \quad \square$$

例 2.6 设 $\mathbb{R}^4$ 是标准欧式空间,

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

计算子空间 $\langle\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\rangle^\perp$ 的一组基.

解. 注意到 $\mathbf{x}=(x_1,x_2,x_3)^t\in\langle\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\rangle^\perp$ 当且仅当 $(\mathbf{x}|\mathbf{u}_1)=(\mathbf{x}|\mathbf{u}_2)=(\mathbf{x}|\mathbf{u}_3)=\mathbf{0}$. 即

$$(\mathbf{u}_1^t,\mathbf{u}_2^t,\mathbf{u}_3^t)\mathbf{x}=\mathbf{0}\iff\begin{pmatrix}1&0&1&0\\0&-1&1&-1\\1&1&1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

换言之, $\langle\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\rangle^\perp$ 是上述方程组的解空间. 直接计算该解空间的基是 $(0,1,0,-1)^t$. □

例 2.7 设标准欧式空间 $\mathbb{R}^3$ 中子空间 U 是方程组

$$\begin{cases}x_1+x_2-x_3=0\\2x_1+3x_2=0\end{cases}$$

的解空间. 求 $U^\perp$ 的一组基.

解. 上述方程组的系数矩阵

$$A=\begin{pmatrix}1&1&-1\\2&3&0\end{pmatrix}.$$

由方程组可知 $\vec{A}_1^t, \vec{A}_2^t \in U^\perp$. 因为 $\mathrm{rank}(A) + \dim(U) = 3$, 所以 $U^\perp = \langle \vec{A}_1^t, \vec{A}_2^t \rangle$. *(定理 2.2)*. 因为 $\mathrm{rank}(A) = 2$, 所以 $U^\perp$ 的一组基是 $\vec{A}_1^t, \vec{A}_2^t$. $\square$

3 正交投影

设 $\mathbf{x} \in V$, W 是 V 的子空间, π_W 是关于直和分解 $V = W \oplus W^\perp$ 从 V 到 W 的投影. 则 $\pi_W(\mathbf{x})$ 称为 $\mathbf{x}$ 在 W 中的正交投影.

命题 3.1 *利用上面的记号,*

(i) 设 $\mathbf{y} \in W$. 则 $\mathbf{y}$ 是 $\mathbf{x}$ 在 W 中的正交投影当且仅当 $(\mathbf{x} - \mathbf{y}) \perp W$.

(ii) 设 $\mathbf{y}$ 是 $\mathbf{x}$ 在 W 中的投影. 则对任意 $\mathbf{w} \in W$,

$$\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{w}\|.$$

证明. (i) 设 $\pi_{W^\perp}$ 是从 V 到 $W^\perp$ 关于直和 $V = W \oplus W^\perp$ 的投影. 如果 $\mathbf{y} = \pi_W(\mathbf{x})$. 则

$$\mathbf{x} = \mathbf{y} + \pi_{W^\perp}(\mathbf{x}) \implies \mathbf{x} - \mathbf{y} = \pi_{W^\perp}(\mathbf{x}) \in W^\perp \implies (\mathbf{x} - \mathbf{y}) \perp W.$$

反之, 设 $\mathbf{z} := \mathbf{x} - \mathbf{y} \in W^\perp$. 则 $\mathbf{x} = \mathbf{y} + \mathbf{z} = \pi_W(\mathbf{x}) + \pi_{W^\perp}(\mathbf{x})$. 根据 $\mathbf{x}$ 在直和分解 $V = W \oplus W^\perp$ 分解中的唯一性, 我们

有 $\mathbf{y}=\pi_W(\mathbf{x})$ 和 $\mathbf{z}=\pi_{W^\perp}(\mathbf{x})$. 特别地, $\mathbf{y}$ 是 $\mathbf{x}$ 在 W 中的正交投影.

(ii) 注意到 $\mathbf{x}-\mathbf{w}=(\mathbf{x}-\mathbf{y})+(\mathbf{y}-\mathbf{w})$. 由 (i) 和 $\mathbf{y}-\mathbf{w}\in W$ 可知, $(\mathbf{x}-\mathbf{y})\perp(\mathbf{y}-\mathbf{w})$. 再利用勾股定理得到

$$\|\mathbf{x}-\mathbf{w}\|^2=\|\mathbf{x}-\mathbf{y}\|^2+\|\mathbf{y}-\mathbf{w}\|^2\geq\|\mathbf{x}-\mathbf{y}\|^2.$$

故 $\|\mathbf{x}-\mathbf{y}\|\leq\|\mathbf{x}-\mathbf{w}\|$. $\square$

根据上述命题 (ii), 我们把 $\|\mathbf{x}-\pi_W(\mathbf{x})\|$ 称为 $\mathbf{x}$ 到 W 的距离, 记为 $\mathrm{d}(\mathbf{x},W)$.

下面我们来推导计算 $\mathbf{x}\in V$ 到 $W=\langle\mathbf{w}_1,\ldots,\mathbf{w}_d\rangle$ 的正交投影和距离的公式, 其中 $\mathbf{w}_1,\ldots,\mathbf{w}_d$ 是 W 的一组基.

设 $A=\mathrm{G}(\mathbf{w}_1,\ldots,\mathbf{w}_d)$. 根据第三章第一讲命题 1.6, A 可逆. 设 $\mathbf{y}=\alpha_1\mathbf{w}_1+\cdots+\alpha_d\mathbf{w}_d$, 其中 $\alpha_1,\ldots,\alpha_d\in\mathbb{R}$ 待定. 根据命题 3.1 (i), $\mathbf{y}$ 是 $\mathbf{x}$ 在 W 中的正交投影当且仅当

$$(\mathbf{x}-\mathbf{y})\perp\mathbf{w}_i,\quad i=1,2,\ldots,d.$$

换言之, $(\mathbf{x}-\mathbf{y}|\mathbf{w}_i)=0$, $i=1,2,\ldots,d$. 利用 Gram 矩阵直接计算得到

$$A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_d\end{pmatrix}=\begin{pmatrix}(\mathbf{v}|\mathbf{w}_1)\\ \vdots\\ (\mathbf{v}|\mathbf{w}_d)\end{pmatrix}. \tag{4}$$

再利用 Cramer 法则得出

$$\alpha_j=\frac{\det(A_j)}{\det(A)}, \tag{5}$$

其中 A_j 是把 A 中第 j 列换成 $((\mathbf{v}|\mathbf{w}_1),\ldots,(\mathbf{v}|\mathbf{w}_d))^t$ 得到的矩阵, $j=1,\ldots,d$.

进而, 根据勾股定理, 我们有

$$\mathrm{d}(\mathbf{x},W)^2=\|\mathbf{x}\|^2-\|\alpha_1\mathbf{w}_1+\cdots+\alpha_d\mathbf{w}_d\|^2=(\mathbf{x}|\mathbf{x})-(\alpha_1,\ldots,\alpha_d)A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_d\end{pmatrix}. \tag{6}$$

定理 3.2 设 $\mathbf{x}\in V$, W 是 V 的子空间, $\mathbf{w}_1,\ldots,\mathbf{w}_d$ 是 W 的一组基. 则

$$\mathrm{d}(\mathbf{x},W)^2=\frac{\det(\mathrm{G}(\mathbf{x},\mathbf{w}_1,\ldots,\mathbf{w}_d))}{\det(\mathrm{G}(\mathbf{w}_1,\ldots,\mathbf{w}_d))}.$$

证明. 如前文所述, 令 $A=\mathrm{G}(\mathbf{w}_1,\ldots,\mathbf{w}_d)$. 把 $\det(\mathrm{G}(\mathbf{x},\mathbf{w}_1,\ldots,\mathbf{w}_d))$

按第一行展开得

$$
\begin{aligned}
&\det(\mathrm{G}(\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_d)) \\
&= \det(A)(\mathbf{x}|\mathbf{x}) - \sum_{j=1}^{d} (\mathbf{x}|\mathbf{w}_j) \det(A_j) \quad (A_j \text{ 由 (5) 给出}) \\
&= \det(A) \left((\mathbf{x}|\mathbf{x}) - \sum_{j=1}^{d} (\mathbf{x}|\mathbf{w}_j) \frac{\det(A_j)}{\det(A)} \right) \\
&= \det(A) \left((\mathbf{x}|\mathbf{x}) - \sum_{j=1}^{d} (\mathbf{x}|\mathbf{w}_j) \alpha_j \right) \quad (\text{根据 (5)}) \\
&= \det(A)\Big((\mathbf{x}|\mathbf{x}) - (\alpha_1, \ldots, \alpha_d) \begin{pmatrix} (\mathbf{x}|\mathbf{w}_1) \\ \vdots \\ (\mathbf{x}|\mathbf{w}_d) \end{pmatrix} \Big) \\
&= \det(A)\Big((\mathbf{x}|\mathbf{x}) - (\alpha_1, \ldots, \alpha_d) A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \Big) \quad (\text{根据 (4)}) \\
&= \det(A) \mathrm{d}(\mathbf{x}, W) \quad (\text{根据 (6)}) \quad \square
\end{aligned}
$$

注解 3.3 设 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组基. 我们说明

$$P_n = \sqrt{\det(\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_n))}$$

代表 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 构成的平行多面体的体积.

当 $n = 1$ 时, $P_1 = \sqrt{(\mathbf{v}_1|\mathbf{v}_1)} = \|\mathbf{v}_1\|$.

当 $n = 2$ 时, 由定理 *3.2* 可知, $P_2 = \mathrm{d}(\mathbf{v}_1, \langle \mathbf{v}_2 \rangle)^2 \|\mathbf{v}_1\|^2$. 于是, $\mathrm{G}(\mathbf{v}_1, \mathbf{v}_2)$ 是 $\mathbf{v}_1, \mathbf{v}_2$ 构成的平行四边形的面积的平方.

对于 $n > 2$ 的情形, 定理 *3.2* 蕴含,

$$P_n^2 = \mathrm{d}(\mathbf{v}_1, \langle \mathbf{v}_2, \ldots, \mathbf{v}_n \rangle)^2 \mathrm{G}(\mathbf{v}_2, \ldots, \mathbf{v}_n)^2.$$

注意到 $\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = (\mathbf{v}_1, \ldots, \mathbf{v}_n)^t (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, 其中 $\mathbf{v}_i$ 理解为该向量在 V 的一组单位正交基下的坐标. 于是,

$$P_n = |\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)|.$$

故 $\mathbb{R}^n$ 中 n 个线性无关列向量构成的行列式可以理解为这些向量组成的平行多面体的"有向体积".

在本节的最后, 我们简单地介绍最小二乘法 (the least square method). 我们把 $\mathbb{R}^n$ 看作标准欧式空间. 设 $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ 和 $\mathbf{x} = (x_1, \ldots, x_n)$. 考虑方程组,

$$A\mathbf{x} = \mathbf{b}. \tag{7}$$

设 $\mathbf{b}$ 在 $V_c(A)$ 中的正交投影是 $\mathbf{v}$=$\alpha_1 \vec{A}^{(1)} + \cdots + \alpha_n \vec{A}^{(n)}$. 则

$$(\alpha_1, \ldots, \alpha_n)^t.$$

称为 (7) 的一个最小平方解.

注意到 (7) 有解当且仅当 $\mathbf{b} \in V_c(A)$. 此时 $\mathbf{v} = \mathbf{b}$. 故每个 (7) 的最小平方解都是 (7) 的解, 反之亦然. 当 A 列满秩时, (7) 的最小平方解是唯一的.

例 3.4 设 x 代表某种杂质, y 代表产品的成品率. 已知

$$y = a\,x\% + b.$$

根据下列实验数据

$x\%$	*3.6*	*3.7*	*3.8*	*4.0*	*4.1*	*4.2*
y	*1.0*	*0.9*	*0.9*	*0.6*	*0.56*	*0.35*

求 a, b.

解. a, b 满足的方程组是

$$\begin{cases} 3.6a + b = 1.0 \\ 3.7a + b = 0.9 \\ 3.8a + b = 0.9 \\ 4.0a + b = 0.6 \\ 4.1a + b = 0.56 \\ 4.2a + b = 0.35 \end{cases}$$

该方程组无解. 计算其最小平方解得到 $a{=}-1.05, b{=}4.81$. 故 $y = -1.05\,x\% + 4.81$.

4 正交矩阵与正交等价

设欧式空间 V 由两组单位正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\epsilon_1, \ldots, \epsilon_n$, 矩阵 $P \in \mathrm{GL}_n(\mathbb{R})$ 满足 $(\epsilon_1, \ldots, \epsilon_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)P$. 则对任

意 $i, j \in \{1, 2, \ldots, n\}$, 我们有

$$\delta_{i,j} = (\epsilon_i|\epsilon_j) = ((\mathbf{e}_1, \ldots, \mathbf{e}_n)\vec{P}^{(i)}|(\mathbf{e}_1, \ldots, \mathbf{e}_n)\vec{P}^{(j)}) = (\vec{P}^{(i)})^t\vec{P}^{(j)}.$$

由此得出 $P^tP = E$, 进而 $PP^t = E$.

定义 4.1 设 $P \in \mathrm{GL}_n(\mathbb{R})$. 如果 $P^t = P^{-1}$, 则称 P 是正交矩阵. 所有 n 阶正交矩阵的集合记为 $\mathrm{O}_n(\mathbb{R})$.

显然, E 是正交矩阵.

命题 4.2 集合 $\mathrm{O}_n(\mathbb{R})$ 是 $\mathrm{GL}_n(\mathbb{R})$ 的子群.

证明. 根据第一学期第四章第一讲命题 2.24, 只要证明对任意 $P, Q \in \mathrm{O}_n(\mathbb{R})$, $PQ^{-1} \in \mathrm{O}_n(\mathbb{R})$. 我们计算

$$(PQ^{-1})^t(PQ^{-1}) = (Q^{-1})^tP^tPQ^{-1} = (Q^t)^tQ^{-1} = QQ^{-1} = E.$$

于是, $(PQ^{-1})^t = (PQ^{-1})^{-1}$. □

命题 4.3 *(i)* 如果 $P \in \mathrm{O}_n(\mathbb{R})$, 则 $\det(P) = \pm 1$.

(ii) $P \in \mathrm{O}_n(\mathbb{R})$ 当且仅当 P 的列向量是标准欧式空间 $\mathbb{R}^n$ 中的一组单位正交基.

(iii) $P \in \mathrm{O}_n(\mathbb{R})$ 当且仅当 P 的行向量是标准欧式空间 $\mathbb{R}^{1\times n}$ 中的一组单位正交基.

证明. (i) 因为 $P^tP = 1$, 所以 $\det(P^tP) = 1$. 于是, $\det(P^t)\det(P) = \det(P)^2 = 1$. 故 $\det(P) = \pm 1$.

(ii) 对任意 $i, j \in \{1, \ldots, n\}$,

$$(\vec{P}^i|\vec{P}^j) = \delta_{i,j} \iff (\vec{P}^i)^t \vec{P}^j = \delta_{i,j} \iff P^tP = 1.$$

(iii) 考虑矩阵 P^t 即可. □

例 4.4 证明: $P \in \mathrm{O}_2(\mathbb{R})$ 当且仅当存在 θ 使得

$$P = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad 或 \quad P = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

证明. 设

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

是正交矩阵. 由命题 *4.3 (ii)* 可知, $a^2 + c^2 = 1$ 可知. 我们不妨设 $a = \cos(\theta)$. 则 $c = \pm\sin(\theta)$. 由命题 *4.3 (iii)* 可知, $a^2 + b^2 = 1$. 于是, $b = \pm\sin(\theta)$. 同理 $c^2 + d^2 = 1$ 得出 $d = \pm\cos(\theta)$.

情形 1. $c = \sin(\theta)$, $b = -\sin(\theta)$. 由 $ab + cd = 0$ 得出 $d = \cos(\theta)$. 此时

$$P = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

情形 2. $c=\sin(\theta)$, $b=\sin(\theta)$. 由 $ab+cd=0$ 得出 $d=-\cos(\theta)$. 此时

$$P=\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{pmatrix}.$$

其它情形可在上述情形中把 θ 换为 $-\theta$ 得到. □

命题 4.5 设欧式空间 V 由基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\epsilon_1,\ldots,\epsilon_n$, 矩阵 $P\in\mathrm{GL}_n(\mathbb{R})$ 满足 $(\epsilon_1,\ldots,\epsilon_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)P$. 再设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是单位正交基. 则 $\epsilon_1,\ldots,\epsilon_n$ 是单位正交基当且仅当 $P\in\mathrm{O}_n(\mathbb{R})$.

证明. 设 $\epsilon_1,\ldots,\epsilon_n$ 是单位正交基. 由引进正交矩阵的概念的推导过程可知, $P\in\mathrm{O}_n(\mathbb{R})$. 反之, 设 $P\in\mathrm{O}_n(\mathbb{R})$. 对任意 $i,j\in\{1,2,\ldots,n\}$, 命题 4.3 (ii) 蕴含

$$(\epsilon_i|\epsilon_j)=((\mathbf{e}_1,\ldots,\mathbf{e}_n)\vec{P}^{(i)}|(\mathbf{e}_1,\ldots,\mathbf{e}_n)\vec{P}^{(j)})=(\vec{P}^{(i)})^t\vec{P}^{(j)}=\delta_{i,j}.$$

故 $\epsilon_1,\ldots,\epsilon_n$ 是单位正交基. □

设 V 有两组单位正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\epsilon_1,\ldots,\epsilon_n$. 则存在唯一的 $P\in\mathrm{O}_n(\mathbb{R})$ 使得 $(\mathbf{e}_1,\ldots,\mathbf{e}_n)=(\epsilon_1,\ldots,\epsilon_n)P$. 设 $\mathcal{A}\in\mathcal{L}(V)$ 在 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\epsilon_1,\ldots,\epsilon_n$ 下的矩阵分别是 A 和 B. 根据第二章第一讲第 2.2 节第一段,

$$B=P^{-1}AP=P^tAP\quad(\because\ P\in\mathrm{O}_n(\mathbb{R})).$$

定义 4.6 设 $A, B \in \mathrm{M}_n(\mathbb{R})$. 如果存在 $P \in \mathrm{O}_n(\mathbb{R})$ 使得 $B = P^{-1}AP$, 则称 A 与 B 正交等价(正交相似), 记为 $A \sim_o B$.

我们来验证 $\sim_o$ 是等价关系. 因为 $E \in \mathrm{O}_n(\mathbb{R})$, 所以对任意 $A \in \mathrm{M}_n(\mathbb{R})$, $A = E^{-1}AE$. 故 $A \sim_o A$. 自反性成立.

设 $A, B \in \mathrm{M}_n(\mathbb{R})$ 满足 $A \sim_o B$. 则存在 $P \in \mathrm{O}_n(\mathbb{R})$ 使得 $B = P^{-1}AP$. 于是, $A = PBP^{-1}$. 根据命题 4.2, $P^{-1} \in \mathrm{O}_n(\mathbb{R})$, 我们得到 $B \sim_o A$. 对称性成立.

再设 $C \in \mathrm{M}_n(\mathbb{R})$ 且 $A \sim_o B$ 和 $B \sim_o C$. 则存在 $P, Q \in \mathrm{O}_n(\mathbb{R})$ 使得 $B = P^{-1}AP$ 和 $C = Q^{-1}BQ$. 于是, $C = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$. 根据命题 4.2, $PQ \in \mathrm{O}_n(\mathbb{R})$. 故 $A \sim_o C$. 传递律成立. 验证完毕.

注解 4.7 符号如定义 *4.6*, 如果 $A \sim_o B$, 则 $A \sim_s B$ 且 $A \sim_c B$. 这是因为正交矩阵的逆和转置相等. 由此可得, 矩阵的相似不变量和合同不变量都是正交等价的不变量.

例 4.8 设

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

证明*:* $A \sim_s B$, $A \sim_c B$ 但 $A \not\sim_o B$.

证明. 显然 $\chi_A = \chi_B = t^2$. 因为 $\mathrm{rank}(A) = \mathrm{rank}(B)$ 且 $\mathrm{rank}(A^2) = \mathrm{rank}(B^2)$. 根据相似判定法则 I, $A \sim_s B$. 设

$P = \mathrm{diag}(1/\sqrt{2}, 1/\sqrt{2})$. 则 $P^t A P = B$. 于是, $A \sim_c B$.

假设存在 $Q \in \mathrm{O}_2(\mathbb{R})$ 使得 $Q^t A Q = B$. 根据例 *4.4*, 我们有

$$Q = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad 或 \quad Q = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

如果 Q 为前者, 则 $AQ = QB$ 蕴含

$$\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

于是, $\sin(\theta) = 0$ 且 $\cos(\theta) = 0$. 矛盾. 类似地可证明 Q 也不可能等于后者. $\square$

问题. 给定 $A \in \mathrm{M}_n(\mathbb{R})$,

1. 求它在正交等价下的标准型.

2. 给定 $A, B \in \mathrm{M}_n(\mathbb{R})$, 判定它们是否正交等价.