第三章　内积空间

1.2　长度、距离、角度和正交

定义 1.9 设 V 是欧式空间, $\mathbf{x} \in V$. 称 $\sqrt{(\mathbf{x}|\mathbf{x})}$ 是 $\mathbf{x}$ 的长度, 记为 $\|\mathbf{x}\|$. 再设 $\mathbf{y} \in V$. 则 $\|\mathbf{x}-\mathbf{y}\|$ 称为 $\mathbf{x}$ 到 $\mathbf{y}$ 之间的距离.

由内积的正定性可知, $\|\mathbf{x}\|$ 是良定义的且 $\|\mathbf{x}\| = 0$ 当且仅当 $\mathbf{x} = \mathbf{0}$. 从而, $\mathbf{x} = \mathbf{y}$ 等且仅当 $\|\mathbf{x}-\mathbf{y}\| = \mathbf{0}$. 由双线性可知 $\|\mathbf{x}\| = \|-\mathbf{x}\|$. 从而, $\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{y}-\mathbf{x}\|$.

定理 1.10 设 $\mathbf{x}, \mathbf{y} \in V$. 则 $|(\mathbf{x}|\mathbf{y})| \le \|\mathbf{x}\|\|\mathbf{y}\|$ *(Cauchy-Bunyakovsky* 不等式*)*. 特别地, $|(\mathbf{x}|\mathbf{y})| = \|\mathbf{x}\|\|\mathbf{y}\|$ 当且仅当 $\mathbf{x}, \mathbf{y}$ 线性相关.

证明. 当 $\mathbf{x} = \mathbf{0}$ 或 $\mathbf{y} = \mathbf{0}$ 时定理中的结论显然成立. 设 $\mathbf{y} \neq \mathbf{0}$ 和 λ 是任意实数. 则 $(\mathbf{x}+\lambda\mathbf{y}|\mathbf{x}+\lambda\mathbf{y}) \ge 0$. 利用双线性性和对称性得

$$(\mathbf{x}+\lambda\mathbf{y}|\mathbf{x}+\lambda\mathbf{y}) = (\mathbf{y}|\mathbf{y})\lambda^2 + 2(\mathbf{x}|\mathbf{y})\lambda + (\mathbf{x}|\mathbf{x}) \ge 0.$$

于是, $\Delta := 4(\mathbf{x}|\mathbf{y})^2 - 4(\mathbf{y}|\mathbf{y})(\mathbf{x}|\mathbf{y}) \le 0$. 由此得出

$$(\mathbf{x}|\mathbf{y})^2 \le (\mathbf{x}|\mathbf{x})(\mathbf{y}|\mathbf{y}) \implies |(\mathbf{x}|\mathbf{y})| \le \|\mathbf{x}\|\|\mathbf{y}\|.$$

注意到 $|(\mathbf{x}|\mathbf{y})| = \|\mathbf{x}\|\|\mathbf{y}\|$ 当且仅当 $\Delta = 0$ 当且仅当存在 $\lambda \in \mathbb{R}$ 使得

$$(\mathbf{x}+\lambda\mathbf{y}|\mathbf{x}+\lambda\mathbf{y}) = 0.$$

这结论等价于 $\mathbf{x}+\lambda\mathbf{y} = \mathbf{0}$ (内积正定性). 在 $\mathbf{y} \neq 0$ 得条件下, 上述结论等价于 $\mathbf{x}$ 和 $\mathbf{y}$ 线性相关. □

例 1.11 *在标准欧式空间中, Cauchy-Bunyakovsky 不等式是对任意 $x_1,\ldots,x_n,y_1,\ldots,y_n\in\mathbb{R}$*

$$|x_1y_1+\cdots+x_ny_n| \leq \sqrt{x_1^2+\cdots+x_n^2}\sqrt{y_1^2+\cdots+y_n^2}.$$

在上一讲例 1.2 定义的矩阵欧式空间中, Cauchy-Bunyakovsky 不等式是对任意 $A,B\in \mathrm{M}_n(\mathbb{R})$,

$$|\mathrm{tr}(A^tB)| \leq \sqrt{\mathrm{tr}(A^tA)}\sqrt{\mathrm{tr}(B^tB)}.$$

在上一讲例 1.3 定义的多项式欧式空间中, Cauchy-Bunyakovsky 不等式是对任意 $f,g\in\mathbb{R}[x]^{(n)}$,

$$|\int_a^b f(x)g(x)dx| \leq \sqrt{\int_a^b f(x)^2dx}\sqrt{\int_a^b g(x)^2dx}.$$

设 $\mathbf{x},\mathbf{y}\in V$, 且存在 $\alpha\in\mathbb{R}$ 使得 $\mathbf{x}=\alpha\mathbf{y}$ 或 $\mathbf{y}=\alpha\mathbf{x}$. 则称 $\mathbf{x},\mathbf{y}$ 平行. 如果 $\alpha\geq 0$, 则称 $\mathbf{x},\mathbf{y}$ 同向. 如果 $\alpha\leq 0$, 则称 $\mathbf{x},\mathbf{y}$ 反向. 有时也称 $\mathbf{0}$ 是迷向的.

推论 1.12 *设 $\mathbf{x},\mathbf{y}\in V$. 则 $\|\mathbf{x}+\mathbf{y}\|\leq\|\mathbf{x}\|+\|\mathbf{y}\|$. 等式成立等且仅当 $\mathbf{x}$ 和 $\mathbf{y}$ 同向.*

证明. 我们计算

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}) && \text{(长度的定义)}\\
&= (\mathbf{x}|\mathbf{x}) + 2(\mathbf{x}|\mathbf{y}) + (\mathbf{y}|\mathbf{y}) && \text{(双线性和对称性)}\\
&= \|\mathbf{x}\|^2 + 2(\mathbf{x}|\mathbf{y}) + \|\mathbf{y}\|^2 && \text{(长度的定义)}\\
&\le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 && \text{(Cauchy-Bunyakovsky)}\\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.
\end{aligned}$$

于是, $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$.

下面验证等式成立的充要条件. 不妨设 $\mathbf{y} \neq 0$. 由上面计算可知等式成立当且仅当 $(\mathbf{x}|\mathbf{y}) = \|\mathbf{x}\|\|\mathbf{y}\|$. 根据定理 1.10, 此时存在 $\alpha \in \mathbb{R}$ 使得 $\mathbf{x} = \alpha\mathbf{y}$. 于是, $(\mathbf{x}|\mathbf{y}) = \|\mathbf{x}\|\|\mathbf{y}\|$ 等价于 $(\alpha\mathbf{y}|\mathbf{y}) = \|\alpha\mathbf{y}\|\|\mathbf{y}\|$, 即 $\alpha\|\mathbf{y}\|^2 = |\alpha|\|\mathbf{y}\|^2$. 换言之, $\alpha = |\alpha|$. 即 $\mathbf{x}, \mathbf{y}$ 同向. □

设 $\mathbf{x} \in V \setminus \{\mathbf{0}\}$. 如果 $\|\mathbf{x}\| = 1$, 则称 $\mathbf{x}$ 是单位向量. 设 $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 则 $\mathbf{v}/\|\mathbf{v}\|$ 是与 $\mathbf{v}$ 同向的单位向量, 称为 $\mathbf{v}$ 的单位化向量.

定义 1.13 设 $\mathbf{x}, \mathbf{y} \in V \setminus \{\mathbf{0}\}$. 称

$$\arccos\left(\frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)$$

是 $\mathbf{x}, \mathbf{y}$ 的夹角.

根据 Cauchy-Bunyakovsky 不等式, 夹角是良定义的. 它的通常取值范围是 $[0, \pi]$.

定义 1.14 设 $\mathbf{x}, \mathbf{y} \in V$. 如果 $(\mathbf{x}|\mathbf{y}) = 0$, 则称 $\mathbf{x}$ 和 $\mathbf{y}$ 正交, 记为 $\mathbf{x} \perp \mathbf{y}$.

零向量与任何向量都正交.

引理 1.15 设 $\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_k \in V$, 其中 $\mathbf{x}_1, \ldots, \mathbf{x}_k$ 非零.

(i) $\mathbf{x} \perp \mathbf{x} \iff \mathbf{x} = \mathbf{0}$.

(ii) 如果 $\mathbf{x}_1, \ldots, \mathbf{x}_k$ 两两正交, 则它们线性无关.

证明. (i) 注意到

$$(\mathbf{x}|\mathbf{x}) = 0 \iff \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}.$$

(ii) 因为对任意 $i, j \in \{1, 2, \ldots, k\}$ 满足 $i \neq j$, 我们有 $(\mathbf{x}_i, \mathbf{x}_j) = 0$, 所以 $\mathrm{G}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \mathrm{diag}((\mathbf{x}_1|\mathbf{x}_1), \ldots, (\mathbf{x}_k|\mathbf{x}_k))$. 因为对任意 $i \in \{1, 2, \ldots, k\}$, $(\mathbf{x}_i|\mathbf{x}_i) \neq 0$, 所以

$$\mathrm{rank}(\mathrm{G}(\mathbf{x}_1, \ldots, \mathbf{x}_k)) = k.$$

根据上一讲命题 1.7, $\mathbf{x}_1, \ldots, \mathbf{x}_k$ 线性无关. □

例 1.16 *(勾股定理)* 设 $\mathbf{x}, \mathbf{y} \in V$. 证明 $\mathbf{x} \perp \mathbf{y}$ 当且仅当

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

证明. 因为 $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2(\mathbf{x}|\mathbf{y}) + \|\mathbf{y}\|^2$, 所以

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

当且仅当 $(\mathbf{x}|\mathbf{y}) = 0$, 即 $\mathbf{x} \perp \mathbf{y}$. □

1.3 单位正交基

设 $\dim(V) = n$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 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), $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基.

例 1.17 在标准欧式空间 $\mathbb{R}^n$ 中, 标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 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 $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ 线性无关. 则存在两两正交的单位向量 $\epsilon_1, \ldots, \epsilon_k$ 使得

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

$i = 1, 2, \ldots, k$. 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 $\epsilon_1, \ldots, \epsilon_{i-1}$ 使得

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

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,

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

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 $\epsilon_1, \epsilon_2, \epsilon_3$. □

命题 1.20 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

$$(\mathbf{x}|\mathbf{y}) = x_1 y_1 + \cdots + x_n y_n.$$

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 1.21 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

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

例 1.22 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 *1.20*,

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

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 *1.20* 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. □

例 2.4 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 2.5 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 2.6 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 $\operatorname{rank}(A) + \dim(U) = 3$, 所以 $U^\perp = \langle \vec{A}_1^t, \vec{A}_2^t \rangle$. *(定理 2.2)*. 因为 $\operatorname{rank}(A) = 2$, 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 $\|\mathbf{x}-\mathbf{y}\|\leq\|\mathbf{x}-\mathbf{w}\|$. □

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 $\mathrm{d}(\mathbf{x},W)$.

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. 因为 $\mathbf{x}=\pi_W(\mathbf{x})+\pi_{W^\perp}(\mathbf{x})$, 所以

$$\det(\mathrm{G}(\mathbf{x},\mathbf{w}_1,\ldots,\mathbf{w}_d))=||\pi_{W^\perp}(\mathbf{x})||^2\det(\mathrm{G}(\mathbf{w}_1,\ldots,\mathbf{w}_d)). \tag{3}$$

验证过程详见注 3.3. 故

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

注解 3.3 *上述证明中 (3) 的推导过程如下:*

设 $\mathbf{y}=\pi_W(\mathbf{x})$ 和 $\mathbf{z}=\pi_{W^\perp}(\mathbf{x})$. 则

$$\det(\mathrm{G}(\mathbf{x},\mathbf{w}_1,\ldots,\mathbf{w}_d))=\det(\mathrm{G}(\mathbf{y}+\mathbf{z},\mathbf{w}_1,\ldots,\mathbf{w}_d))$$

$$=\det\begin{pmatrix}(\mathbf{y}+\mathbf{z}|\mathbf{y}+\mathbf{z}) & (\mathbf{y}+\mathbf{z}|\mathbf{w}_1) & \ldots & (\mathbf{y}+\mathbf{z}|\mathbf{w}_d)\\ (\mathbf{w}_1|\mathbf{y}+\mathbf{z}) & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \vdots & \vdots & \ddots & \vdots\\ (\mathbf{w}_d|\mathbf{y}+\mathbf{z}) & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}$$

$$=\det\begin{pmatrix}(\mathbf{y}+\mathbf{z}|\mathbf{y}) & (\mathbf{y}+\mathbf{z}|\mathbf{w}_1) & \ldots & (\mathbf{y}+\mathbf{z}|\mathbf{w}_d)\\ (\mathbf{w}_1|\mathbf{y}) & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \vdots & \vdots & \ddots & \vdots\\ (\mathbf{w}_d|\mathbf{y}) & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}$$

$$+\det\begin{pmatrix}(\mathbf{y}+\mathbf{z}|\mathbf{z}) & (\mathbf{y}+\mathbf{z}|\mathbf{w}_1) & \ldots & (\mathbf{y}+\mathbf{z}|\mathbf{w}_d)\\ (\mathbf{w}_1|\mathbf{z}) & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \vdots & \vdots & \ddots & \vdots\\ (\mathbf{w}_d|\mathbf{z}) & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}$$

$$=\underbrace{\det\begin{pmatrix}(\mathbf{y}|\mathbf{y}) & (\mathbf{y}|\mathbf{w}_1) & \ldots & (\mathbf{y}|\mathbf{w}_d)\\ (\mathbf{w}_1|\mathbf{y}) & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \vdots & \vdots & \ddots & \vdots\\ (\mathbf{w}_d|\mathbf{y}) & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}}_{D_1}+\underbrace{\det\begin{pmatrix}\color{blue}{(\mathbf{z}|\mathbf{y})} & \color{blue}{(\mathbf{z}|\mathbf{w}_1)} & \color{blue}{\ldots} & \color{blue}{(\mathbf{z}|\mathbf{w}_d)}\\ (\mathbf{w}_1|\mathbf{y}) & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \vdots & \vdots & \ddots & \vdots\\ (\mathbf{w}_d|\mathbf{y}) & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}}_{D_2}$$

$$+\underbrace{\det\begin{pmatrix}\color{blue}{(\mathbf{y}|\mathbf{z})} & (\mathbf{y}|\mathbf{w}_1) & \ldots & (\mathbf{y}|\mathbf{w}_d)\\ \color{blue}{(\mathbf{w}_1|\mathbf{z})} & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \color{blue}{\vdots} & \vdots & \ddots & \vdots\\ \color{blue}{(\mathbf{w}_d|\mathbf{z})} & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}}_{D_3}+\underbrace{\det\begin{pmatrix}(\mathbf{z}|\mathbf{z}) & \color{red}{(\mathbf{z}|\mathbf{w}_1)} & \color{red}{\ldots} & \color{red}{(\mathbf{z}|\mathbf{w}_d)}\\ \color{red}{(\mathbf{w}_1|\mathbf{z})} & (\mathbf{w}_1|\mathbf{w}_1) & \ldots & (\mathbf{w}_1|\mathbf{w}_d)\\ \color{red}{\vdots} & \vdots & \ddots & \vdots\\ \color{red}{(\mathbf{w}_d|\mathbf{z})} & (\mathbf{w}_d|\mathbf{w}_1) & \ldots & (\mathbf{w}_d|\mathbf{w}_d)\end{pmatrix}}_{D_4}$$

因为 $\mathbf{y}\in W$, 所以 $\mathbf{y},\mathbf{w}_1,\ldots,\mathbf{w}_d$ 线性相关. 由上一讲命题 *1.7*, $D_1=0$. 因为 $\mathbf{z}\perp W$ 且 $\mathbf{y}\in W$, 所以 D_2 的第一行中的元素*(*蓝色部分*)*都是零. 于是, $D_2=0$. 同理, D_3 的

第一列中的元素(蓝色部分)都是零. 故 $D_3=0$. 我们得到

$$\det(\mathrm{G}(\mathbf{x},\mathbf{w}_1,\ldots,\mathbf{w}_d))=D_4.$$

在根据 $\mathbf{z}\perp W$, D_4 中的红色部分都是零. 即

$$D_4=(\mathbf{z}|\mathbf{z})\mathrm{G}(\mathbf{w}_1,\mathbf{w}_d)=\mathrm{d}(\mathbf{x},W)^2\mathrm{G}(\mathbf{w}_1,\mathbf{w}_d).$$

注解 3.4 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, $P_2^2=\mathrm{d}(\mathbf{v}_1,\langle\mathbf{v}_2\rangle)^2\|\mathbf{v}_1\|^2$. 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 $\mathbf{x} = (x_1, \ldots, x_n)$. 考虑线性方程组,

$$A\mathbf{x} = \mathbf{b}. \tag{4}$$

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 (4) 的一个最小平方解.

注意到 (4) 有解当且仅当 $\mathbf{b} \in V_c(A)$. 此时 $\mathbf{v} = \mathbf{b}$. 故每个 (4) 的最小平方解都是 (4) 的解, 反之亦然. 当 A 列满秩时, (4) 的最小平方解是唯一的.

注解 3.5 *计算最小二乘解的一种方法如下: 设 $\mathbf{b}$ 在 $V_c(A)$ 上的正交投影是*

$$\mathbf{v} = \alpha_1\vec{A}^{(1)} + \cdots + \alpha_n\vec{A}^{(n)},$$

其中 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ 待定. 则

$$(\mathbf{b} - \mathbf{v}) \perp \vec{A}^{(j)}, \quad j = 1, 2, \ldots, n.$$

即

$$\left(\mathbf{b} - \mathbf{v} | \vec{A}^{(j)}\right) = 0, \quad j = 1, 2, \ldots, n.$$

转换方程组得

$$G\left(\vec{A}^{(1)},\dots,\vec{A}^{(n)}\right)\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}=\begin{pmatrix}\left(\vec{A}^{(1)}|\mathbf{b}\right)\\ \vdots\\ \left(\vec{A}^{(n)}|\mathbf{b}\right)\end{pmatrix}.$$

故最小平凡解 $(\alpha_1,\dots,\alpha_n)^t$ 唯一当且仅当 $\operatorname{rank}(A)=n$.

例 3.6 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 $\det(P)=\pm1$.

(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

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

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. □