第二章　线性算子

定义 11.11 设 $A \in \mathrm{M}_n(F)$. 把 A 看成从 F^n 上的线性算子所对应的初等因子组称为矩阵 A 的初等因子组.

定理 11.12 设 $A \in \mathrm{M}_n(\mathbb{C})$. 在不计 *Jordan* 块出现顺序的前提下, A 的 *Jordan* 标准型由 A 的初等因子组唯一确定.

证明. 设 $\mathcal{A} : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ 由公式 $\mathcal{A}(\mathbf{x}) = A\mathbf{x}$ 确定. 设

$$V = V_1 \oplus \cdots \oplus V_k$$

是 V 的一个 $\mathcal{A}$-不可分子空间分解. 令 $\mathcal{A}_i = \mathcal{A}_{V_i}$ 和 $\mu_i = \mu_{\mathcal{A}_i}, i = 1, 2, \ldots, k$. 根据代数学基本定理, 存在 $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ (不必两两不同), $d_1, \ldots, d_k \in \mathbb{Z}^+$ 使得

$$\mu_1 = (t - \alpha_1)^{d_1}, \ldots, \mu_k = (t - \alpha_k)^{d_k}.$$

于是, $\mathcal{A}$ 的初等因子组是

$$\{(t - \alpha_1)^{d_1}, \ldots, (t - \alpha_k)^{d_k}\}.$$

根据第二章第五讲引理 11.1, $\mathcal{A}$ 在 V 的某组基下的矩阵是

$$J_A = \begin{pmatrix} J_{d_1}(\alpha_1) & & \\ & \ddots & \\ & & J_{d_k}(\alpha_k) \end{pmatrix}.$$

于是, $A \sim_s J_A$ 且 J_A 由$\mathcal{A}$ 的初等因子组唯一确定. □

由上述定理可知, 记号 J_A 以及把 J_A 称为 A 的 Jordan 标准型都是合理的. 进而, 计算复数域上方阵的 Jordan 标准型等价于计算该矩阵的初等因子组.

注解 11.13 设 $A \in \mathrm{M}_n(\mathbb{C})$ 且 $\mathrm{spec}_{\mathbb{C}}(A) = \{\lambda_1, \ldots, \lambda_s\}$. 则 $\chi_{\mathcal{A}}$ 的不可约因子是 $t-\lambda_1, \ldots, t-\lambda_s$. 我们可以把上一讲定理 *11.10* 中的 $R(i,\ell)$ 和 $N(i,\ell)$ 分别记为 $R(\lambda_i,\ell)$ 和 $N(\lambda_i,\ell)$. 此时的重数公式是

$$N(\lambda_i,\ell) = R(\lambda_i,\ell-1) + R(\lambda_i,\ell+1) - 2R(\lambda_i,\ell),$$

其中 $\lambda_i \in \mathrm{spec}_{\mathbb{C}}(A), \ell \in \mathbb{Z}^+$.

例 11.14 设:

$$A = \begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 2 & -2 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}_7(\mathbb{C}).$$

计算 J_A.

解. 由计算机计算得:

$$\chi_{\mathcal{A}} = t^7 - 9\,t^6 + 34\,t^5 - 70\,t^4 + 85\,t^3 - 61\,t^2 + 24\,t - 4 = (t-2)^2(t-1)^5.$$

设 $\lambda_1 = 2$ 和 $\lambda_2 = 1$. 显然 $R(\lambda_1, 0) = 7$. 由计算机得 $R(\lambda_1, 1) = 6, R(\lambda_1, 2) = 5$. 于是,

$$N(\lambda_1, 1) = R(\lambda_1, 0) + R(\lambda_1, 2) - 2R(\lambda_1, 1) = 7 + 5 - 2 \times 6 = 0.$$

由计算机得 $R(\lambda_1, 3) = 5.$, 于是,

$$N(\lambda_1, 2) = R(\lambda_1, 1) + R(\lambda_1, 3) - 2R(\lambda_1, 2) = 6 + 5 - 2 \times 5 = 1.$$

因为 λ_1 的代数重数等于 2, 所以当 $\ell > 2$ 时, $N(\lambda_1, \ell) = 0$.
显然 $R(\lambda_2, 0) = 7$. 由计算机得 $R(\lambda_2, 1) = 4, R(\lambda_2, 2) = 2$. 于是,

$$N(\lambda_2, 1) = R(\lambda_2, 0) + R(\lambda_2, 2) - 2R(\lambda_2, 1) = 7 + 2 - 2 \times 4 = 1.$$

由计算机计算得 $R(\lambda_2, 3) = 2$. 于是,

$$N(\lambda_2, 2) = R(\lambda_2, 1) + R(\lambda_2, 3) - 2R(\lambda_2, 2) = 4 + 2 - 2 \times 2 = 2.$$

因为 λ_2 的代数重数等于 5, 所以当 $\ell > 2$ 时, $N(\lambda_2, \ell) = 0$.

由此得出

$$J_A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad \square$$

例 11.15 设 $A \in \mathrm{M}_5(\mathbb{C})$. 设

$$\mathrm{rank}(A) = 3,\ \mathrm{rank}(A^2) = 2,\ \mathrm{rank}(A+E) = 4,\ \mathrm{rank}((A+E)^2) = 3.$$

求 J_A

解. 由秩的条件可知 $\lambda_1 = 0$ 和 $\lambda_2 = -1$ 是 A 的两个特征根. 根据上一讲定理 *11.10*,

$$N(\lambda_1, 1) = R(\lambda_1, 0) + R(\lambda_1, 2) - 2R(\lambda_1, 1) = 5 + 2 - 2 \times 3 = 1$$

和

$$N(\lambda_2, 1) = R(\lambda_2, 0) + R(\lambda_2, 2) - 2R(\lambda_2, 1) = 5 + 3 - 2 \times 4 = 0.$$

注意到 $\mathrm{rank}(A) = 3$ 和 $\mathrm{rank}(A + E) = 4$ 分别蕴含 λ_1 的几何重数是 2 和 λ_2 的几何重数是 1. 由此得出 $N(\lambda_1, 2) = 1$

和 $N(\lambda_2,2)=1$. 我们有

$$J_A=\begin{pmatrix}0&&&&\\&0&1&&\\&0&0&&\\&&&-1&1\\&&&0&-1\end{pmatrix}.\quad\square$$

12 矩阵相似的判定

引理 12.1 设 $A,B\in\mathrm{M}_n(F)$ 且 $A\sim_s B$. 则对任意 $f\in F[t]$, $f(A)\sim_s f(B)$. 特别地, $\mathrm{rank}(f(A))=\mathrm{rank}(f(B))$.

证明. 设 $A=P^{-1}BP$, 其中 $P\in\mathrm{GL}_n(F)$. 因为对任意 $k\in\mathbb{N}$, $A^k=P^{-1}B^kP$, 所以 $f(A)=P^{-1}f(B)P$. $\square$

定理 12.2 *(相似判别法I)* 设 $A,B\in\mathrm{M}_n(F)$. 则 $A\sim_s B$ 当且仅当 A 和 B 由共同的初等因子组.

证明. 设 $A\sim_s B$. 则 $\mu_A=\mu_B$ (第二章第二讲命题 4.9). 设 $p_1,\ldots p_s$ 是 μ_A 的两两互素的首一的不可约因子. 则它们也是 μ_B 的两两互素的首一的不可约因子. 根据引理 12.1, 对任意 $\ell\in\mathbb{N}$, $i\in\{1,2,\ldots,s\}$, $\mathrm{rank}(p_i(\mathcal{A})^\ell)=\mathrm{rank}(p_i(\mathcal{B})^\ell)$. 由上一讲定理 11.10, A 和 B 由共同的初等因子组.

反之, 设 A 和 B 由共同的初等因子组 $\{p_1,\dots,p_k\}$. 把 A 和 B 看成 F^n 的算子分别记为 $\mathcal{A}$ 和 $\mathcal{B}$. 令

$$F^n = V_1 \oplus \cdots \oplus V_k = W_1 \oplus \cdots \oplus W_k,$$

其中 V_i 是 $\mathcal{A}$-不可分的, W_i 是 $\mathcal{B}$-不可分的, $i{=}1,2,\dots,k$. 调整下标后可再设 $\mathcal{A}_{V_i}$ 和 $\mathcal{B}_{W_i}$ 的极小多项式都是 p_i. 令

$$p_i = t^{d_i} + \alpha_{i,d_i-1}t^{d_i-1} + \cdots + \alpha_{i,1}t + \alpha_{i,0},$$

其中 $\alpha_{i,n-1},\dots,\alpha_{i,1},\alpha_{i,0} \in F$. 因为 V_i 是 $\mathcal{A}$-不可分的, 所以它是 $\mathcal{A}$-循环的. 于是存在 $\mathbf{v}_i \in V$ 使得 $V_i = F[\mathcal{A}_{V_i}] \cdot \mathbf{v}_i$. 由此得出 $\mathcal{A}_{U_i}$ 在基底 $\mathbf{v}_i, \mathcal{A}(\mathbf{v}_i), \dots \mathcal{A}^{d_i-1}(\mathbf{v}_i)$ 下的矩阵是

$$M_i = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_{i,0} \\ 1 & 0 & \cdots & 0 & -\alpha_{i,1} \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{i,d_i-1} \end{pmatrix} \in \mathrm{M}_{d_i}(F).$$

同理存在 $\mathbf{w}_i \in V$ 使得 $W_i = F[\mathcal{B}_{V_i}]\mathbf{w}_i$, 且 $\mathcal{B}_{W_i}$ 在基底 $\mathbf{w}_i, \mathcal{B}(\mathbf{w}_i), \dots \mathcal{B}^{d_i-1}(\mathbf{w}_i)$ 下的矩阵也是 M_i. 由第二章第三讲定理 5.9,

$$A \sim_s \begin{pmatrix} M_1 & O & \cdots & O \\ O & M_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & M_k \end{pmatrix} \quad \text{和} \quad B \sim_s \begin{pmatrix} M_1 & O & \cdots & O \\ O & M_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & M_k \end{pmatrix}.$$

于是, $A \sim_s B$. □

定理 12.3 *(相似判别法II) 设 $A, B \in \mathrm{M}_n(F)$. 则 $A \sim_s B$ 当且仅当下述两点同时成立:*

(i) $\chi_A = \chi_B$, 或 $\mu_A = \mu_B$;

(ii) 设 $p_1, \ldots, p_s$ 是 χ_A 或 μ_A 在 $F[t]$ 中的两两互素的(首一的)不可约因子, 且

$$\forall\, i \in \{0, 1, \ldots, n+1\}, j \in \{1, 2, \ldots, s\},\ \mathrm{rank}(p_j(A)^i) = \mathrm{rank}(p_j(B)^i).$$

证明. 设 $A \sim_s B$. 则 $\chi_A = \chi_B$ 和 $\mu_A = \mu_B$ (第二章第三讲定义 7.6 后的讨论和第二章第二讲命题 4.9). 由引理 12.1, 对任意 $i \in \{0, 1, \ldots, n\}$, $j \in \{1, 2, \ldots, s\}$,

$$p_j(A)^i \sim_s p_j(B)^i \implies \mathrm{rank}(p_j(A)^i) = \mathrm{rank}(p_j(B)^i).$$

反之, 设 $\chi_A = \chi_B$ 和 (ii), 或 $\mu_A = \mu_B$ 和 (ii) 成立. 则 A 和 B 由共同的初等因子组(上一讲定理 11.10). 于是, $A \sim_s B$ (定理 12.2). □

定理 12.4 *(相似判别法III) 设 $A, B \in \mathrm{M}_n(F)$. 则 $A \sim_s B$ 当且仅当对任意 $f \in F[t]$, $\mathrm{rank}(f(A)) = \mathrm{rank}((f(B))$.*

证明. 设 $A \sim_s B$. 引理 12.1 蕴含, 对任意 $f \in F[t]$, $\mathrm{rank}(f(A)) = \mathrm{rank}((f(B))$. 反之, 因为 $\mathrm{rank}(\mu_A(A)){=}0$, 所

以 $\operatorname{rank}(\mu_A(B))$=0. 于是 $\mu_A(B)$=O. 由此可知 $\mu_B|\mu_A$(第二章第二讲引理 4.2). 同理 $\mu_A|\mu_B$. 于是, $\mu_A = \mu_B$. 由定理 12.3 可知, $A \sim_s B$. □.

例 12.5 设 $A \in \mathrm{M}_n(\mathbb{C})$. 证明: $A \sim_s A^t$.

证明. 注意到对任意 $f \in F[t]$, $f(A)^t$=$f(A^t)$. 故

$$\operatorname{rank}(f(A)) = \operatorname{rank}(f(A)^t) = \operatorname{rank}(f(A^t)).$$

由定理 *12.4* 可知, $A \sim_s A^t$. □

例 12.6 设 $A, B \in \mathrm{M}_n(\mathbb{R})$. 证明: 如果存在 $P \in \mathrm{GL}_n(\mathbb{C})$ 使得 $A = P^{-1}BP$, 则存在 $Q \in \mathrm{GL}_n(\mathbb{R})$ 使得 $A = Q^{-1}BQ$.

证明. 因为存在 $P \in \mathrm{GL}_n(\mathbb{C})$ 使得 $A = P^{-1}BP$, 所以对任意 $f \in \mathbb{C}[t]$, $\operatorname{rank}(f(A)) = \operatorname{rank}(f(B))$ (定理 *12.4*. 特别地, 任意 $g \in \mathbb{R}[t]$, $\operatorname{rank}(g(A)) = \operatorname{rank}(g(B))$. 再由定理 *12.4*, 存在 $Q \in \mathrm{GL}_n(\mathbb{R})$ 使得 $A = Q^{-1}BQ$. □

例 12.7 设 $A \in \mathrm{GL}_n(\mathbb{C})$. 证明:

$$J_A = \begin{pmatrix} J_{d_1}(\lambda_1) & O & \cdots & O \\ O & J_{d_2}(\lambda_2) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{d_k}(\lambda_k) \end{pmatrix}$$

蕴含

$$J_{A^{-1}}=\begin{pmatrix} J_{d_1}(\lambda_1^{-1}) & O & \cdots & O \\ O & J_{d_2}(\lambda_2^{-1}) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{d_k}(\lambda_k^{-1}) \end{pmatrix}.$$

证明. 首先, $A \in \mathrm{GL}_n(\mathbb{C})$ 蕴含 A 的特征根非零. 于是, λ_i^{-1} 有意义. 设 $\lambda \in \mathrm{spec}_{\mathbb{C}}(A)$. 则

$$|\lambda E - A| = 0 \iff |A||\lambda A^{-1} - E| = 0 \iff |\lambda^{-1}E - A^{-1}| = 0.$$

即 $\lambda \in \mathrm{spec}_{\mathbb{C}}(A) \iff \lambda^{-1} \in \mathrm{spec}_{\mathbb{C}}(A^{-1})$. 于是, $p_\lambda = t - \lambda$ 是 χ_A 的因子当且仅当 $q_\lambda = t - \lambda^{-1}$ 是 $\chi_{A^{-1}}$ 的因子. 设 $\ell \in \{0, 1, \ldots, n\}$. 则

$$\mathrm{rank}((A - \lambda E)^\ell) = \mathrm{rank}(A^\ell (E - \lambda A^{-1})^\ell) = \mathrm{rank}((\lambda^{-1}E - A^{-1})^\ell).$$

于是, 对任意 $\ell \in \{0, 1, \ldots, n\}$, $\mathrm{rank}(p_\lambda(A)^\ell) = \mathrm{rank}(q_\lambda(A)^\ell)$. 根据上一讲定理 *11.10*, 对任意 $\lambda \in \mathrm{spec}_{\mathbb{C}}(A)$, p_λ 在 A 的初等因子组中的重数等于 q_λ 在 A^{-1} 的初等因子组中的重数. 故对任意 $m \in \mathbb{Z}^+$, $J_m(\lambda)$ 出现在 J_A 中的重数等于 $J_m(\lambda^{-1})$ 出现在 $J_{A^{-1}}$ 中的重数. $\square$

注解 12.8 上述例子说明: 如果 $\mathrm{spec}_{\mathbb{C}}(A)=\{1,-1\}$, 则 $A \sim_s A^{-1}$.

第三章　内积空间

1　欧式空间

约定: 在本节中 V 是实数域 $\mathbb{R}$ 上的有限维线性空间.

1.1　V 上的内积

定义 1.1 设 $f(\mathbf{x},\mathbf{y})$ 是 V 上的对称双线性型满足 $f(\mathbf{x},\mathbf{x})$ 是正定的. 则称 (V,f) 是一个欧式空间, f 是 V 上的内积.

例 1.2 *(标准欧式空间)* 设 $V=\mathbb{R}^n$,

$$\begin{aligned} f:\ V\times V &\longrightarrow \mathbb{R} \\ (\mathbf{x},\mathbf{y}) &\mapsto \mathbf{x}^t\mathbf{y}. \end{aligned}$$

注意到 $f(\mathbf{x},\mathbf{y})=\mathbf{x}^tE\mathbf{y}$. 于是, f 是 $\mathbb{R}^n$ 对称双线性型, 且 $f(\mathbf{x},\mathbf{x})=\mathbf{x}^t\mathbf{x}$ 是正定的. 于是, (V,f) 是欧式空间.

例 1.3 设 $V=\mathrm{M}_n(\mathbb{R})$,

$$\begin{aligned} f:\ V\times V &\longrightarrow \mathbb{R} \\ (X,Y) &\mapsto \mathrm{tr}(X^tY). \end{aligned}$$

下面我们来验证 f 是 V 上的内积.

首先设 $\alpha, \beta \in \mathbb{R}$, $A, B \in V$,

$$\begin{aligned} f(\alpha A+\beta B, Y) &= \operatorname{tr}((\alpha A+\beta B)^t Y) && (f \text{ 的定义}) \\ &= \operatorname{tr}((\alpha A^t+\beta B^t) Y) && (\text{转置的性质}) \\ &= \operatorname{tr}(\alpha(A^t Y)+\beta(B^t Y)) && (\text{矩阵乘法分配律}) \\ &= \alpha \operatorname{tr}(A^t Y)+\beta \operatorname{tr}(B^t Y) && (\operatorname{tr} \text{ 是线性函数}) \\ &= \alpha f(A, Y)+\beta f(B, Y) && (f \text{ 的定义}). \end{aligned}$$

于是, f 关于第一个变元是线性的. 类似地可验证 f 关于第二个变元也是线性的. 故 f 是双线性型. 注意到

$$f(Y, X)=\operatorname{tr}(Y^t X)=\operatorname{tr}((Y^t X)^t)=\operatorname{tr}(X^t Y)=f(X, Y).$$

于是, f 是对称的. 设 $X=(x_{i,j}) \neq O$. 则

$$f(X, X)=X^t X=\sum_{i=1}^n \sum_{j=1}^n x_{i,j}^2>0.$$

于是, $f(X, X)$ 是正定的. 由此得出 (V, f) 是欧式空间.

例 1.4 设 $V=\mathbb{R}[x]^{(n)}$, $a, b \in \mathbb{R}$ 且 $a<b$.

$$\begin{array}{rccc} \phi: & V \times V & \longrightarrow & \mathbb{R} \\ & (f, g) & \mapsto & \int_a^b f(x) g(x)\, dx. \end{array}$$

下面我们来验证 ϕ 是 V 上的内积. 首先设 $\alpha, \beta \in \mathbb{R}$, $p, q \in V$.

则

$$
\begin{aligned}
\phi(\alpha p+\beta q, g) &= \int_a^b (\alpha p(x)+\beta q(x))g(x))dx \quad (\phi \text{ 的定义}) \\
&= \alpha\int_a^b p(x)g(x)dx+\beta\int_a^b q(x)g(x)dx \quad (\int_a^b \text{ 线性}) \\
&= \alpha f(p,g)+\beta f(q,g) \quad (f \text{ 的定义}).
\end{aligned}
$$

于是, ϕ 关于第一个变元是线性的. 类似地可验证 ϕ 关于第二个变元也是线性的. 故 ϕ 是双线性型. ϕ 显然是对称的. 设 $f \in R[x]^{(n)} \setminus \{0\}$. 则 $\phi(f,f) = \int_a^b f(x)^2dx > 0$. 于是, $\phi(f,f)$ 是正定的. 由此得出 (V,ϕ) 是欧式空间.

我们把欧式空间 V 上的内积记为 $(\,|\,)$. 即

$$
\begin{aligned}
(\,|\,):\ V\times V &\longrightarrow \quad \mathbb{R} \\
(\mathbf{x},\mathbf{y}) &\mapsto \quad (\mathbf{x}|\mathbf{y}).
\end{aligned}
$$

利用上述符号, 内积的基本性质如下:

(i) (双线性)对任意 $\mathbf{x},\mathbf{y},\mathbf{z}\in V$, $\alpha,\beta\in\mathbb{R}$,

$$(\alpha\mathbf{x}+\beta y|\mathbf{z}) = \alpha(\mathbf{x}|\mathbf{z})+\beta(\mathbf{y}|\mathbf{z}), \quad (\mathbf{x}|\alpha\mathbf{y}+\beta\mathbf{z}) = \alpha(\mathbf{x}|\mathbf{y})+\beta(\mathbf{x}|\mathbf{z}).$$

(ii) (对称性)对任意 $\mathbf{x},\mathbf{y}\in V$, $(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})$.

(iii) (正定性)对任意 $\mathbf{x}\in V$,

$$(\mathbf{x}|\mathbf{x})\geq 0 \quad \text{且} \quad (\mathbf{x}|\mathbf{x}) = 0 \iff \mathbf{x}=\mathbf{0}.$$

定义 1.5 设 V 是欧式空间, $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. 定义

$$\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m) = ((\mathbf{v}_i|\mathbf{v}_j))_{m\times m}.$$

称之为 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 的 Gram 矩阵.

由内积的对称性可知 $\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ 是对称的.

命题 1.6 设 V 是欧式空间, $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$, $\alpha_1, \ldots, \alpha_m$, $\beta_1, \ldots, \beta_m \in \mathbb{R}$. 令

$$\mathbf{x} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m \text{ 和 } \mathbf{y} = \beta_1\mathbf{v}_1 + \cdots + \beta_m\mathbf{v}_m).$$

则

$$(\mathbf{x}|\mathbf{y}) = (\alpha_1, \ldots, \alpha_m)\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}.$$

证明. 利用内积的双线性性得出

$$\begin{aligned}(\mathbf{x}|\mathbf{y}) &= (\sum_{i=1}^m \alpha_i\mathbf{v}_i|\sum_{j=1}^m \alpha_j\mathbf{v}_j)\\ &= \sum_{i=1}^m\sum_{j=1}^m \alpha_i\beta_j(\mathbf{v}_i|\mathbf{v}_j) \quad (\text{双线性})\\ &= (\alpha_1, \ldots, \alpha_m)\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}. \quad \square\end{aligned}$$

命题 1.7 设 V 是欧式空间, $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. 则 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 线性相关当且仅当 $\operatorname{rank}(\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m)) < m$.

证明. 设 $\operatorname{rank}(\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m)) < m$. 则存在 $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ 不全为零, 使得

$$\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

令 $\mathbf{x} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m$. 由命题 1.6 和上式可知,

$$(\mathbf{x}|\mathbf{x}) = (\alpha_1, \ldots, \alpha_m)\mathrm{G}(\mathbf{v}_1, \ldots, \mathbf{v}_m) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = 0.$$

根据内积的正定性推出 $\mathbf{x} = \mathbf{0}$. 故 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 线性相关.

反之, 设 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 线性相关. 则存在 $\beta_1, \ldots, \beta_m \in \mathbb{R}$, 不全为零, 使得

$$\beta_1 \mathbf{v}_1 + \cdots + \beta_m \mathbf{v}_m = \mathbf{0}.$$

令

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = G(\mathbf{v}_1, \ldots, \mathbf{v}_m) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$$

则对 $i=1,2,\dots,m$,

$$
\begin{aligned}
0 &= (\mathbf{v}_i|\mathbf{0}) = (\mathbf{v}_i|\sum_{j=1}^{m}\beta_j\mathbf{v}_j) \\
&= (0,\dots,0,\underset{\substack{\uparrow \\ i}}{1},0,\dots,0)G(\mathbf{v}_1,\dots,\mathbf{v}_m)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix} \\
&= \gamma_i.
\end{aligned}
$$

于是, 以 $\mathrm{G}(\mathbf{v}_1,\dots,\mathbf{v}_m)$ 为系数矩阵的齐次线性方程组有非平凡解. 故 $\mathrm{rank}(\mathrm{G}(\mathbf{v}_1,\dots,\mathbf{v}_m)<m$. □

注解 1.8 由上述两个命题可知, $\mathrm{G}(\mathbf{v}_1,\dots,\mathbf{v}_m)$ 是半正定的. 它是正定的当且仅当 $\mathbf{v}_1,\dots,\mathbf{v}_m$ 线性无关.

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 $|(\mathbf{x}|\mathbf{y})| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ *(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 $(\mathbf{x}+\lambda\mathbf{y}|\mathbf{x}+\lambda\mathbf{y}) \geq 0$. 利用双线性性和对称性得

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

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

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

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 $f, g \in \mathbb{R}[x]^{(n)}$,

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

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 $\alpha \ge 0$, 则称 $\mathbf{x}, \mathbf{y}$ 同向. 如果 $\alpha \le 0$, 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* $\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$. *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), $\langle \mathbf{v}_1,\ldots,\mathbf{v}_i\rangle = \langle \epsilon_1,\ldots,\epsilon_i\rangle$.

$\square$

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 的一组单位正交基是 $\epsilon_1, \epsilon_2, \epsilon_3$. □