§4 酉空间的分解与投影

程光辉

2020 年 3 月 13 日

1 欧式 (酉) 空间

定义 1 若 $\boldsymbol{V}_n(\mathbf{P})$ 上的映射 $(\boldsymbol{x},\boldsymbol{y})$: $\boldsymbol{V}_n(\mathbf{P})\times\boldsymbol{V}_n(\mathbf{P})\to\mathbf{P}$ 满足:

(1) $(\boldsymbol{x},\boldsymbol{x})\geq\mathbf{0}$; $(\boldsymbol{x},\boldsymbol{x})=\mathbf{0}$ 当且仅当 $\boldsymbol{x}=\mathbf{0}$; (正定性)

(2) $(\boldsymbol{x},\boldsymbol{y})=\overline{(\boldsymbol{y},\boldsymbol{x})}$, $\forall\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{V}_n(\mathbf{P})$;((共轭) 对称性)

(3) $(\boldsymbol{x},\boldsymbol{\lambda}\boldsymbol{y})=\boldsymbol{\lambda}(\boldsymbol{x},\boldsymbol{y})$, $\forall\boldsymbol{\lambda}\in\boldsymbol{P},\forall\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{V}_n(\mathbf{P})$; (齐次性)

(4) $(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{x},\boldsymbol{z})$, $\forall\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\boldsymbol{V}_n(\mathbf{P})$, (可加性)

则称映射 $(\boldsymbol{x},\boldsymbol{y})$ 是 $\boldsymbol{V}_n(\mathbf{P})$ 上的内积; 若在 $\boldsymbol{n}$ 维线性空间 $\boldsymbol{V}_n(\mathbf{P})$ 中定义了内积, 则称该空间为内积空间.

例 1 *(1)* $\mathbf{C}^n$ 中的标准内积, $\forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$ 定义 $(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{y}$;

(2) $\mathbf{C}^{m\times n}$ 中的标准内积, $\forall\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{m\times n}$, 定义

$$(\boldsymbol{A},\boldsymbol{B})=tr(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{B});$$

(3) 闭区间 $[\boldsymbol{a},\boldsymbol{b}]$ 全体连续函数构成的空间 $\boldsymbol{C}[\boldsymbol{a},\boldsymbol{b}]$ 中的内积, $\forall\boldsymbol{f},\boldsymbol{g}\in\boldsymbol{C}[\boldsymbol{a},\boldsymbol{b}]$, 定义

$$(\boldsymbol{f},\boldsymbol{g})=\int_a^b\overline{\boldsymbol{f}}\boldsymbol{g}d\boldsymbol{t};$$

(4) 若 $\boldsymbol{A}$ 为实对称正定矩阵, 则 $\mathbf{R}^n$ 可定义内积

$$(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{A}}=\sqrt{\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{y}},$$

称为 $\boldsymbol{A}$-内积.

定义 2 当 $\mathbf{P}=\mathbf{R}$ 时, $\boldsymbol{V}_n(\mathbf{R})$ 称为欧式空间; 当 $\mathbf{P}=\mathbf{C}$ 时, $\boldsymbol{V}_n(\mathbf{C})$ 称为酉空间.

定义 3 对任意 $\boldsymbol{x} \in \boldsymbol{V}_n(\mathbf{C})$, 称

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$$

为向量 $\boldsymbol{x}$ 的长度.

定理 1 向量长度的性质:

(1) $\|\boldsymbol{\lambda x}\| = |\boldsymbol{\lambda}|\|\boldsymbol{x}\|$;

(2) $\|\boldsymbol{x}-\boldsymbol{y}\|^2 + \|\boldsymbol{x}+\boldsymbol{y}\|^2 = 2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2$ *(平行四边形法则)*;

(3) $|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$ *(Cauchy 不等式)*;

(4) $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ *(三角不等式)*.

定义 4 两个非零向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的夹角定义为

$$\cos\boldsymbol{\theta} = \frac{|(\boldsymbol{x}, \boldsymbol{y})|}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}\sqrt{(\boldsymbol{y}, \boldsymbol{y})}} = \frac{|\boldsymbol{x}^H\boldsymbol{y}|}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}.$$

显然，当 $\boldsymbol{x}^H\boldsymbol{y} = 0$ 时, $\boldsymbol{\theta} = \frac{\pi}{2}$, 此时称向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交. 因此，两个常数向量正交定义如下.

定义 5 若两个常数向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积为零, 即 $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^H\boldsymbol{y} = 0$, 则称它们是正交的, 并记为 $\boldsymbol{x} \perp \boldsymbol{y}$.

定义 6 若 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 是 $\boldsymbol{V}_n(\mathbf{C})$ 的两个子空间, 若 $\forall \boldsymbol{v}_1 \in \boldsymbol{V}_1$ 和 $\forall \boldsymbol{v}_2 \in \boldsymbol{V}_2$, 有 $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$, 则称它们是正交的, 并记为 $\boldsymbol{V}_1 \perp \boldsymbol{V}_2$.

2 正交补子空间

定义 7 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，称 $\mathrm{N}(\boldsymbol{A}) = \{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x} = 0,\ \boldsymbol{x} \in \mathbf{C}^n\}$ 为 $\boldsymbol{A}$ 的核 *(或零空间 **Null**)*, $\mathrm{R}(\boldsymbol{A}) = \{\boldsymbol{y}|\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},\ \forall \boldsymbol{x} \in \mathbf{C}^n\}$ 为 $\boldsymbol{A}$ 的值域 *(**Range**)*.

定理 2 设 $\boldsymbol{A} \in \mathbf{C}^{n\times m}$，$\boldsymbol{B} \in \mathbf{C}^{n\times s}$，则 $\mathrm{R}(\boldsymbol{A}) \perp \mathrm{R}(\boldsymbol{B})$ 的充要条件是 $\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$.

证明: 设 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m)$，$\boldsymbol{B} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_s)$.

(必要性) 易知 $\boldsymbol{\alpha}_i \in \mathrm{R}(\boldsymbol{A})$，$i = 1, \cdots, m$ 和 $\boldsymbol{\beta}_j \in \mathrm{R}(\boldsymbol{B})$，$j = 1, \cdots, s$，因 $\mathrm{R}(\boldsymbol{A}) \perp R(\boldsymbol{B})$，有 $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j) = \boldsymbol{\alpha}_i^H\boldsymbol{\beta}_j = 0$，即 $\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$.

(充分性) 对 $\forall \boldsymbol{y}_A \in \mathrm{R}(\boldsymbol{A})$ 和 $\forall \boldsymbol{y}_B \in \mathrm{R}(\boldsymbol{B})$，则存在向量 $\boldsymbol{x}_A \in \mathbf{C}^m$ 和 $\boldsymbol{x}_B \in \mathbf{C}^s$ 使得 $\boldsymbol{y}_A = \boldsymbol{A}\boldsymbol{x}_A$ 和 $\boldsymbol{y}_B = \boldsymbol{B}\boldsymbol{x}_B$.

$\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$, 于是有

$$(\boldsymbol{y}_A, \boldsymbol{y}_B) = \boldsymbol{y}_A^H\boldsymbol{y}_B = \boldsymbol{x}_A^H\boldsymbol{A}^H\boldsymbol{B}\boldsymbol{x}_B = 0.$$

由于 $\boldsymbol{y}_A$ 和 $\boldsymbol{y}_B$ 的任意性，即 $\mathrm{R}(\boldsymbol{A}) \perp \mathrm{R}(\boldsymbol{B})$.

推论 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

$$\mathbf{R}(\boldsymbol{A}) \perp \mathbf{N}(\boldsymbol{A}^H); \quad \mathbf{N}(\boldsymbol{A}) \perp \mathbf{R}(\boldsymbol{A}^H).$$

推论 2 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

(1) $\dim\mathbf{R}(\boldsymbol{A}) + \dim\mathbf{N}(\boldsymbol{A}^H) = \boldsymbol{m}$;

(2) $\dim\mathbf{R}(\boldsymbol{A}^H) + \dim\mathbf{N}(\boldsymbol{A}) = \boldsymbol{n}$;

(3) $\mathbf{C}^m = \mathbf{R}(\boldsymbol{A}) \oplus \mathbf{N}(\boldsymbol{A}^H)$;

(4) $\mathbf{C}^n = \mathbf{R}(\boldsymbol{A}^H) \oplus \mathbf{N}(\boldsymbol{A})$.

定义 8 设酉空间 $\boldsymbol{V}_n(\mathrm{C})$ 的两个正交子空间 $\boldsymbol{V}_1$，$\boldsymbol{V}_2$，有 $\boldsymbol{V}_1 \perp \boldsymbol{V}_2$，且 $\boldsymbol{V}_1 + \boldsymbol{V}_2 = \boldsymbol{V}_n(\mathrm{C})$，则称 $\boldsymbol{V}_2$ 为 $\boldsymbol{V}_1$ 的正交补空间，记为 $\boldsymbol{V}_2 = \boldsymbol{V}_1^{\perp}$.

定理 3 酉空间 $\boldsymbol{V}_n(\mathrm{C})$ 的任意子空间 $\boldsymbol{V}_1$ 都有唯一的正交补.

证明： 若 $\boldsymbol{V}_1$ 是平凡子空间 $\{\mathbf{0}\}$ 或 $\boldsymbol{V}_n(\mathbf{C})$，则显然成立.

若 $\boldsymbol{V}_1$ 不是平凡子空间，取 $\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_m$ 为标准正交基，它可以扩充为 $\boldsymbol{V}_n(\mathbf{C})$ 的一组标准正交基

$$\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_m, \boldsymbol{\epsilon}_{m+1}, \cdots, \boldsymbol{\epsilon}_n.$$

即有 $\boldsymbol{V}_2 = \boldsymbol{L}(\boldsymbol{\epsilon}_{m+1}, \cdots, \boldsymbol{\epsilon}_n)$ 为 $\boldsymbol{V}_1$ 的正交补.

(唯一性) 设 $\boldsymbol{V}_2$，$\boldsymbol{V}_3$ 都是 $\boldsymbol{V}_1$ 的正交补，则

$$\boldsymbol{V}_n(\mathrm{C}) = \boldsymbol{V}_1 + \boldsymbol{V}_2, \quad \boldsymbol{V}_n(\mathrm{C}) = \boldsymbol{V}_1 + \boldsymbol{V}_3.$$

对 $\forall \boldsymbol{\alpha}_2 \in \boldsymbol{V}_2$，即有 $\boldsymbol{\alpha}_2 \in \boldsymbol{V}_n(\mathbf{C})$, 于是 $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3$, 其中 $\boldsymbol{\alpha}_1 \in \boldsymbol{V}_1$，$\boldsymbol{\alpha}_3 \in \boldsymbol{V}_3$. 又因为 $\boldsymbol{\alpha}_2 \perp \boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_3 \perp \boldsymbol{\alpha}_1$，所以

$$\mathbf{0} = (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) + (\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1),$$

即 $\boldsymbol{\alpha}_1 = \mathbf{0}$. 因此，$\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 \in \boldsymbol{V}_3$，即 $\boldsymbol{V}_2 \subseteq \boldsymbol{V}_3$.

同理可证 $\boldsymbol{V}_3 \subseteq \boldsymbol{V}_2$，即 $\boldsymbol{V}_2 = \boldsymbol{V}_3$. 唯一性得证.

3 投影与幂等矩阵

定义 9 设 $\boldsymbol{V}_n(\mathrm{C})$ 是线性空间，如果线性变换 $\boldsymbol{T}: \boldsymbol{V}_n(\mathrm{C}) \to \boldsymbol{V}_n(\mathrm{C})$ 具有 $\boldsymbol{T}^2 = \boldsymbol{T}$ 的性质，则称 $\boldsymbol{T}$ 是 $\boldsymbol{V}_n(\mathrm{C})$ 上的投影 *(也称投影算子或幂等算子)*.

定理 4 设 $\boldsymbol{T}$ 是 $\boldsymbol{V_n}(\mathbf{C})$ 上的投影，则

$$\boldsymbol{V_n}(\mathbf{C})=\mathbf{R}(\boldsymbol{T})\oplus\mathbf{N}(\boldsymbol{T}).$$

证明: 对 $\forall\boldsymbol{\alpha}\in\boldsymbol{V_n}(\mathbf{C})$，则有 $\boldsymbol{\alpha_1}=\boldsymbol{T}(\boldsymbol{\alpha})$，记 $\boldsymbol{\alpha_2}=\boldsymbol{\alpha}-\boldsymbol{\alpha_1}$. 因为 $\boldsymbol{T}=\boldsymbol{T^2}$，有

$$\boldsymbol{T}(\boldsymbol{\alpha_2})=\boldsymbol{T}(\boldsymbol{\alpha}-\boldsymbol{\alpha_1})=\boldsymbol{T}(\boldsymbol{\alpha})-\boldsymbol{T}(\boldsymbol{\alpha_1})=\boldsymbol{T}(\boldsymbol{\alpha})-\boldsymbol{T^2}(\boldsymbol{\alpha})=\mathbf{0}.$$

因此，$\boldsymbol{\alpha_2}\in\mathbf{N}(\boldsymbol{T})$. 所以 $\boldsymbol{V_n}(\mathbf{C})=\mathbf{R}(\boldsymbol{T})+\mathbf{N}(\boldsymbol{T})$.

$\forall\boldsymbol{\beta}\in\mathbf{R}(\boldsymbol{T})\bigcap\mathbf{N}(\boldsymbol{T})$，则 $\boldsymbol{\beta}\in\mathbf{R}(\boldsymbol{T})$ 且 $\boldsymbol{\beta}\in\mathbf{N}(\boldsymbol{T})$，即有 $\exists\boldsymbol{\gamma}\in\boldsymbol{V_n}(\mathbf{C})$，使得 $\boldsymbol{\beta}=\boldsymbol{T}(\boldsymbol{\gamma})$ 和 $\boldsymbol{T}(\boldsymbol{\beta})=\mathbf{0}$. 进一步可得

$$\boldsymbol{\beta}=\boldsymbol{T}(\boldsymbol{\gamma})=\boldsymbol{T^2}(\boldsymbol{\gamma})=\boldsymbol{T}(\boldsymbol{\beta})=\mathbf{0}.$$

综上，有 $\boldsymbol{V_n}(\mathbf{C})=\mathbf{R}(\boldsymbol{T})\oplus\mathbf{N}(\boldsymbol{T})$.

定理 5 设 $\boldsymbol{V_n}(\mathbf{C})=\boldsymbol{V_1}\oplus\boldsymbol{V_2}$，则存在投影 $\boldsymbol{T}$，使得

$$\mathbf{R}(\boldsymbol{T})=\boldsymbol{V_1},\quad \mathbf{N}(\boldsymbol{T})=\boldsymbol{V_2}.$$

证明: 因为 $\boldsymbol{V_n}(\mathbf{C})=\boldsymbol{V_1}\oplus\boldsymbol{V_2}$，对 $\forall\boldsymbol{\alpha}\in\boldsymbol{V_n}(\mathbf{C})$，则 $\boldsymbol{\alpha}$ 可以唯一的分解为 $\boldsymbol{\alpha}=\boldsymbol{\alpha_1}+\boldsymbol{\alpha_2}$，其中 $\boldsymbol{\alpha_1}\in\boldsymbol{V_1}$，$\boldsymbol{\alpha_2}\in\boldsymbol{V_2}$. 定义线性映射 $\boldsymbol{T}$ 满足 $\boldsymbol{T}(\boldsymbol{\alpha})=\boldsymbol{\alpha_1}$，即

$$\boldsymbol{T}(\boldsymbol{\alpha})=\boldsymbol{T}(\boldsymbol{\alpha_1}+\boldsymbol{\alpha_2})=\boldsymbol{T}(\boldsymbol{\alpha_1})+\boldsymbol{T}(\boldsymbol{\alpha_2})=\boldsymbol{\alpha_1},$$

由于 $\boldsymbol{\alpha}$ 的任意性, 有 $\boldsymbol{T}(\boldsymbol{\alpha_1})=\boldsymbol{\alpha_1}$, 进而可得 $\boldsymbol{T}(\boldsymbol{\alpha_2})=\mathbf{0}$. 因此，$\boldsymbol{\alpha_2}\in\mathbf{N}(\boldsymbol{T})$. 由于 $\boldsymbol{\alpha}$ 的任意性和直和关系，有 $\mathbf{R}(\boldsymbol{T})=\boldsymbol{V_1}$，$\mathbf{N}(\boldsymbol{T})=\boldsymbol{V_2}$.

又因为

$$\boldsymbol{T^2}(\boldsymbol{\alpha})=\boldsymbol{T}(\boldsymbol{\alpha_1})=\boldsymbol{\alpha_1}=\boldsymbol{T}(\boldsymbol{\alpha}),$$

有 $\boldsymbol{T}$ 为投影.

4 正交投影

定义 10 设 $\boldsymbol{T}$ 是 $\boldsymbol{V_n}(\mathbf{C})$ 上的投影，$\boldsymbol{V_n}(\mathbf{C})=\mathbf{R}(\boldsymbol{T})\oplus\mathbf{N}(\boldsymbol{T})$. 如果 $\mathbf{R}^{\perp}(\boldsymbol{T})=\mathbf{N}(\boldsymbol{T})$，则称 $\boldsymbol{T}$ 是正交投影.

定理 6 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{A^2}=\boldsymbol{A}$，则 $\boldsymbol{A}$ 是正交投影的充分必要条件是 $\boldsymbol{A^H}=\boldsymbol{A}$.

证明: (充分性) 因为 $\boldsymbol{A^2}=\boldsymbol{A}$，知 $\mathbf{C}^n=\mathbf{R}(\boldsymbol{A})\oplus\mathbf{N}(\boldsymbol{A})$. $\forall\boldsymbol{y}\in\mathbf{R}(\boldsymbol{A})$，$\forall\boldsymbol{x}\in\mathbf{N}(\boldsymbol{A})$，则 $\exists\boldsymbol{z}\in\mathbf{C}^n$ 使得 $\boldsymbol{y}=\boldsymbol{Az}$，$\boldsymbol{Ax}=\mathbf{0}$. 于是有

$$(\boldsymbol{y},\boldsymbol{x})=\boldsymbol{y^H x}=\boldsymbol{z^H A^H x}=\boldsymbol{z^H A x}=\mathbf{0},$$

即 $\boldsymbol{y} \perp \boldsymbol{x}$. 由于 $\boldsymbol{y}$, $\boldsymbol{x}$ 的任意性，得 $\mathbf{R}^{\perp}(\boldsymbol{A}) = \mathbf{N}(\boldsymbol{A})$.

(必要性) 因为 $\boldsymbol{A}^2 = \boldsymbol{A}$，$\forall \boldsymbol{x} \in \mathbf{C}^n$，则 $\boldsymbol{A}\boldsymbol{x} \in \mathbf{R}(\boldsymbol{A})$，$\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} \in \mathbf{N}(\boldsymbol{A})$. 因为 $\mathbf{R}^{\perp}(\boldsymbol{A}) = \mathbf{N}(\boldsymbol{A})$，有

$$\mathbf{0} = (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x},\ \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\boldsymbol{H}}(\boldsymbol{A} - \boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})\boldsymbol{x},$$

由 $\boldsymbol{x}$ 的任意性，知 $\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}$ 为 Hermitian 矩阵，即 $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{H}}$.