线性代数课堂讲义

内积空间中的相关变换、谱分解、正定性与 SVD

玖伍

2026 年 5 月 31 日

摘要

本讲义依据 101 教材《代数学 2》相关内容整理。定义、定理、命题、推论、证明和例题尽量保持教材中的数学表述，并适当修改为课堂讲义形式；少量串联性文字均以课堂讲义方式标注。

目录

1 记号与总览

本讲义中的内积空间均默认定义在 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$ 上；若涉及特征多项式分裂条件，会在结论中明示。线性变换主要记为 ϕ,ψ,φ，矩阵主要记为 A,B,U,P。若 α 是 V 的一组基，则 $[\phi]_\alpha$ 表示 ϕ 在 α 下的矩阵。

对矩阵 A，A^T 表示转置，$\overline{A}$ 表示共轭，$A^H=\overline{A}^T$ 表示共轭转置。标准内积空间 $\mathbb{F}^n$ 上由矩阵 A 定义的线性变换记为 $l_A: x\mapsto Ax$。

课堂提示

本讲义的主线是：先用伴随变换把共轭转置矩阵抽象到内积空间中，再用正规变换刻画“可由标准正交特征向量基对角化”的条件；自伴变换给出实谱和谱分解，正定性把二次型、自伴变换、矩阵特征值联系起来，最后由半正定平方根推出极分解和奇异值分解。

2 伴随变换

本节目标

- 掌握伴随变换的定义及有限维情形的存在唯一性。
- 理解伴随变换与共轭转置矩阵 A^H 的对应关系。
- 熟悉伴随运算的基本性质及简单计算。

2.1 定义与基本结论

定义 2.1. 设 ϕ 是内积空间 V 上的线性变换。若存在 V 上的线性变换 ψ 满足

$$(\phi(u),v)=(u,\psi(v)),\qquad \forall u,v\in V,$$

则称 ψ 是 ϕ 的一个伴随变换，表为 $\psi=\phi^*$。

定理 2.2. *设 ϕ 是有限维内积空间 V 上的线性变换，α 是 V 的一组标准正交基，ϕ 在基 α 下的矩阵是 A。那么*

1. *ϕ 存在唯一的伴随变换，即 V 上有唯一的线性变换 ϕ^* 满足*

$$(\phi(u),v)=(u,\phi^*(v)),\qquad \forall u,v\in V;$$

2. *此变换 ϕ^* 在基 α 下的矩阵是 A^H，即 $[\phi^*]_\alpha=[\phi]_\alpha^H$。*

证明. 先证明 ϕ^* 是一个映射，即证明对任意向量 $v\in V$，存在唯一的向量 $w\in V$ 使得

$$(\phi(u),v)=(u,w),\qquad \forall u\in V.$$

那么即可定义 $\phi^*(v)=w$。若令 $x=[u]_\alpha, y=[v]_\alpha, z=[w]_\alpha$，则 $\phi(u)$ 在 α 下的坐标为 Ax，因此

$$\begin{aligned}(\phi(u),v)=(u,w)\,,\ \forall u\in V &\iff (Ax)^T\overline{y}=x^TA^T\overline{y}=x^T\overline{z},\ \forall x\in\mathbb{F}^n\\ &\iff \overline{x}^TA^Hy=\overline{x}^Tz,\ \forall x\in\mathbb{F}^n\\ &\iff A^Hy=z.\end{aligned}$$

由此即知结论成立。由定义可直接验证映射 ϕ^* 是线性的，留作练习。由上面的计算可知 $\phi^*(v)$ 在 α 下的坐标为 $A^H[v]_\alpha$，因此 ϕ^* 在 α 下的矩阵为 A^H。 □

推论 2.3. 设 ϕ 是有限维内积空间 V 上的线性变换，则

$$(u,\phi(v))=(\phi^*(u),v)\,,\qquad \forall u,v\in V.$$

证明. 由内积的对称性有

$$(u,\phi(v))=\overline{(\phi(v),u)}=\overline{(v,\phi^*(u))}=(\phi^*(u),v)\,.$$

□

推论 2.4. 设 $\mathbb{F}^n$ 是 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$ 上的标准内积空间，$A\in\mathbb{F}^{n\times n}$。那么 $(l_A)^*=l_{A^H}$。

例 2.5. 设 e_1,e_2 是 $\mathbb{C}^2$ 的标准基。定义线性变换

$$\phi:\mathbb{C}^2\to\mathbb{C}^2,\qquad \begin{pmatrix}a_1\\a_2\end{pmatrix}\mapsto\begin{pmatrix}2a_1i+3a_2\\a_1-a_2\end{pmatrix}.$$

那么

$$[\phi]_{(e_1,e_2)}=\begin{pmatrix}2i&3\\1&-1\end{pmatrix},\qquad [\phi^*]_{(e_1,e_2)}=\begin{pmatrix}-2i&1\\3&-1\end{pmatrix}.$$

从而

$$\phi^*:\mathbb{C}^2\to\mathbb{C}^2,\qquad \begin{pmatrix}a_1\\a_2\end{pmatrix}\mapsto\begin{pmatrix}-2a_1i+a_2\\3a_1-a_2\end{pmatrix}.$$

2.2 伴随运算的性质

对 $\mathbb{F}=\mathbb{C}$ 或 $\mathbb{R}$ 上的有限维内积空间 V，定义映射

$$*:\mathrm{End}(V)\to\mathrm{End}(V),\qquad \phi\mapsto\phi^*.$$

由伴随变换的定义，可以证明对任意 $\phi,\psi\in\mathrm{End}(V)$ 有

$$(\phi+\psi)^*=\phi^*+\psi^*,\qquad (c\phi)^*=\overline{c}\,\phi^*,\ \forall c\in\mathbb{F},$$
$$(\phi\psi)^*=\psi^*\phi^*,\qquad \phi^{**}=\phi,\qquad \mathrm{id}_V^*=\mathrm{id}_V.$$

命题 2.6. 令 ϕ 为 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 上的线性变换。若 ϕ 有一个特征值 $\lambda\in\mathbb{F}$，则 ϕ^* 有一个特征值 $\overline{\lambda}$。

证明. 设 $u\neq 0$ 是 ϕ 关于特征值 λ 的特征向量。那么对任意 $v\in V$ 有

$$0=(0,v)=((\phi-\lambda\mathrm{id}_V)(u),v)=(u,(\phi-\lambda\mathrm{id}_V)^*(v))=\big(u,(\phi^*-\overline{\lambda}\mathrm{id}_V)(v)\big)\,.$$

即 $u\perp\mathrm{Im}(\phi^*-\overline{\lambda}\mathrm{id}_V)$，这表明 $u\notin\mathrm{Im}(\phi^*-\overline{\lambda}\mathrm{id}_V)$。因此 $\phi^*-\overline{\lambda}\mathrm{id}_V$ 不是满射，从而也不是单射，即 $\mathrm{Ker}(\phi^*-\overline{\lambda}\mathrm{id}_V)\neq\{0\}$。这证明了 $\overline{\lambda}$ 是 ϕ^* 的一个特征值。 □

本节小结

- 有限维内积空间上线性变换的伴随唯一存在。
- 在标准正交基下，伴随变换对应于矩阵的共轭转置。
- 伴随运算反转复合顺序：$(\phi\psi)^* = \psi^*\phi^*$。
- 若 λ 是 ϕ 的特征值，则 $\overline{\lambda}$ 是 ϕ^* 的特征值。

3 正规变换与标准正交对角化

本节目标

- 理解正交相似、酉相似与 Schur 定理。
- 掌握正规变换可由标准正交特征向量基对角化的充要条件。
- 认识自伴、反自伴、保距变换作为正规变换的典型特例。

3.1 正交相似、酉相似与 Schur 定理

定义 3.1. 两个 n 阶实（复）矩阵 A, B 称为正交相似的（酉相似的），是指存在正交矩阵（酉矩阵）U 使得

$$U^TAU = B \qquad (U^HAU = B).$$

定理 3.2. *设 ϕ 为 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 上的线性变换，且特征多项式 $f_\phi(x)$ 在 $\mathbb{F}$ 上分裂。那么存在 V 的标准正交基 ξ 使得 $[\phi]_\xi$ 是上三角形矩阵。*

证明. 对 V 的维数 n 归纳。当 $n = 1$ 时结论显然。假设结论对 $n-1$ 维内积空间上特征多项式分裂的线性变换成立。由于 $f_\phi(x)$ 分裂，ϕ 有特征值 λ。由命题 3.1.1，ϕ^* 有特征值 $\overline{\lambda}$。设 w 是 ϕ^* 关于特征值 $\overline{\lambda}$ 的特征向量，并不妨假定 w 是单位向量。令 $W = \mathrm{Span}\{w\}$。

首先可以证明 $W^\perp$ 是 ϕ-不变的。事实上，对任意 $v \in W^\perp$ 和 $u \in W$ 有

$$(\phi(v), u) = (v, \phi^*(u)) = \left(v, \overline{\lambda}u\right) = \lambda\,(v, u) = 0.$$

因此 $\phi(v) \in W^\perp$。容易证明 $\phi|_{W^\perp}$ 的特征多项式整除 ϕ 的特征多项式，因此也在 $\mathbb{F}$ 上分裂。由归纳假设，$W^\perp$ 有标准正交基 η 使得 $[\phi|_{W^\perp}]_\eta$ 是上三角形矩阵。那么由 $V = W^\perp \oplus W$，$\xi = \eta \cup \{w\}$ 是 V 的标准正交基，并且 $[\phi]_\xi$ 是上三角形矩阵。 □

推论 3.3. *任意复方阵 A 均酉相似于一个上三角形复矩阵。若 n 阶实方阵有 n 个实特征值（计重数），那么 A 正交相似于一个上三角形实矩阵。*

3.2 正规变换

定义 3.4. 1. 设 ϕ 是有限维内积空间 V 上的线性变换。若

$$\phi^*\phi = \phi\phi^*,$$

则称 ϕ 是正规变换。

2. 如果实（复）方阵 A 满足 $AA^T = A^TA$（$AA^H = A^HA$），则称 A 是正规矩阵。

性质 3.5. 设 ϕ 是 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 上的线性变换，ϕ 在 V 的一组标准正交基下的矩阵为 A。那么，ϕ 是正规变换当且仅当 A 是正规矩阵。

命题 3.6. 设 ϕ 是有限维内积空间 V 上的正规变换。那么

1. 对任意 $v \in V$ 有 $\|\phi(v)\| = \|\phi^*(v)\|$；
2. 对任意 $c \in \mathbb{F}$，线性变换 $\phi - c\mathrm{id}_V$ 是正规的；
3. 若 v 是 ϕ 关于特征值 $\lambda \in \mathbb{F}$ 的特征向量，则 v 也是 ϕ^* 关于特征值 $\overline{\lambda}$ 的特征向量；
4. 若 λ_1 和 λ_2 是 ϕ 不同的特征值，则 $V_{\lambda_1} \perp V_{\lambda_2}$。

证明. 对任意 $v \in V$ 有

$$\|\phi(v)\|^2 = (\phi(v), \phi(v)) = (\phi^*\phi(v), v) = (\phi\phi^*(v), v) = (\phi^*(v), \phi^*(v)) = \|\phi^*(v)\|^2.$$

对 $c \in \mathbb{F}$，直接计算可得

$$(\phi - c\mathrm{id}_V)(\phi - c\mathrm{id}_V)^* = (\phi - c\mathrm{id}_V)^*(\phi - c\mathrm{id}_V).$$

若 $(\phi - \lambda\mathrm{id}_V)(v) = 0$，由上一结论及第一条得

$$\left\|(\phi^* - \overline{\lambda}\mathrm{id}_V)(v)\right\| = \|(\phi - \lambda\mathrm{id}_V)(v)\| = 0,$$

故 $\phi^*(v) = \overline{\lambda}v$。若 $v_i \in V_{\lambda_i}$，则

$$\lambda_1 (v_1, v_2) = (\phi(v_1), v_2) = (v_1, \phi^*(v_2)) = \left(v_1, \overline{\lambda_2}v_2\right) = \lambda_2 (v_1, v_2).$$

由 $\lambda_1 \neq \lambda_2$ 即得 $(v_1, v_2) = 0$。 □

定理 3.7. 设 ϕ 是 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 上的线性变换。那么 V 有一组由 ϕ 的特征向量构成的标准正交基（即 ϕ 关于一组标准正交基可对角化）当且仅当 ϕ 是一个特征多项式在 $\mathbb{F}$ 上分裂的正规变换。

证明. 必要性。设 ξ 是由 ϕ 的特征向量构成的 V 的标准正交基。那么 $[\phi]_\xi$ 为对角矩阵，从而 ϕ 的特征多项式分裂，并且 $[\phi^*]_\xi = [\phi]_\xi^H$ 也是对角矩阵。因此

$$[\phi\phi^*]_\xi = [\phi]_\xi[\phi^*]_\xi = [\phi^*]_\xi[\phi]_\xi = [\phi^*\phi]_\xi,$$

即 ϕ 正规。

充分性。设 ϕ 的特征多项式分裂且是正规变换。根据 Schur 定理，存在 V 的标准正交基 $\xi = \{\xi_1, \xi_2, \ldots, \xi_n\}$ 使得 $A = [\phi]_\xi$ 为上三角形矩阵。特别地，ξ_1 是 ϕ 的特征向量。若 $\xi_1, \ldots, \xi_{k-1}$ 是 ϕ 的特征向量，设 $\phi(\xi_j) = \lambda_j\xi_j$。由命题 3.2.1 有 $\phi^*(\xi_j) = \overline{\lambda_j}\xi_j$。由于

$$\phi(\xi_k) = a_{1k}\xi_1 + a_{2k}\xi_2 + \cdots + a_{kk}\xi_k,$$

两边与 ξ_j 作内积可得

$$a_{jk} = (\phi(\xi_k), \xi_j) = (\xi_k, \phi^*(\xi_j)) = \left(\xi_k, \overline{\lambda_j}\xi_j\right) = 0.$$

因此 $\phi(\xi_k) = a_{kk}\xi_k$。由归纳法即知所有 ξ_i 都是 ϕ 的特征向量。 □

推论 3.8. 设 ϕ 是有限维酉空间 V 上的线性变换。那么如下陈述等价：

1. V 有一组由 ϕ 的特征向量构成的标准正交基；

2. ϕ 在 V 的任意标准正交基下对应的矩阵都酉相似于一个对角矩阵；

3. ϕ 是正规变换。

推论 3.9. $A \in \mathbb{C}^{n\times n}$ 酉相似于对角矩阵当且仅当 A 是正规矩阵。

3.3 自伴、反自伴与保距变换

定义 3.10. 1. 设 ϕ 是有限维内积空间 V 上的线性变换。若 $\phi^* = \phi$，则称 ϕ 是自伴变换。

2. 若复（或实）方阵 A 满足 $A = A^H$（或 $A = A^T$），即是 Hermite 矩阵或实对称矩阵，则统称 A 是自伴矩阵。

引理 3.11. 设 ϕ 是有限维内积空间 V 上的自伴变换。那么 ϕ 的特征多项式 $f_\phi(x)$ 总是分裂的，并且 ϕ 的特征值均为实数。

定理 3.12. 设 ϕ 是有限维内积空间 V 上的线性变换。那么下列陈述等价：

1. V 有一组由 ϕ 的特征向量构成的标准正交基，且对应特征值均为实数；

2. ϕ 是自伴变换。

特别地，若 V 是有限维欧氏空间，则 V 有一组由 ϕ 的特征向量构成的标准正交基当且仅当 ϕ 是一个自伴变换。

推论 3.13. *1.* $A \in \mathbb{C}^{n\times n}$ 酉相似于实对角矩阵当且仅当 A 是 *Hermite* 矩阵；

2. $A \in \mathbb{R}^{n\times n}$ 正交相似于对角矩阵当且仅当 A 是实对称矩阵。

例 3.14. 设

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

由于 $A = A^T$，它对应的线性变换是自伴变换，因此可由标准正交特征向量基对角化。直接计算得

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

取

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad Q = (u_1, u_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

则 Q 是正交矩阵，并且

$$Q^T A Q = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

这个例子展示了自伴矩阵的特征子空间天然正交，归一化后即可组成标准正交基。

例 3.15. 设 ϕ 是内积空间 V 上的正规变换，W 是 V 的子空间。证明：若 W 是 ϕ-不变的，则 W 也是 ϕ^*-不变的，从而 ϕ_W 仍是正规变换。

证明. 记 $\phi_W = \phi|_W$。取 W 的一组标准正交基，并扩充为 V 的一组标准正交基。由于 W 是 ϕ-不变的，ϕ 在这组基下的矩阵可写成分块形式

$$A = \begin{pmatrix} B & C \\ O & D \end{pmatrix},$$

其中 B 是 ϕ_W 在 W 的上述标准正交基下的矩阵。

由 ϕ 正规可知 $A^HA = AA^H$。直接计算得

$$A^HA = \begin{pmatrix} B^HB & B^HC \\ C^HB & C^HC + D^HD \end{pmatrix}, \qquad AA^H = \begin{pmatrix} BB^H + CC^H & CD^H \\ DC^H & DD^H \end{pmatrix}.$$

比较左上角分块并取迹，得到

$$\operatorname{tr}(B^HB) = \operatorname{tr}(BB^H) + \operatorname{tr}(CC^H).$$

由于 $\operatorname{tr}(B^HB) = \operatorname{tr}(BB^H)$，故 $\operatorname{tr}(CC^H) = 0$，从而 $C = O$。于是

$$A = \begin{pmatrix} B & O \\ O & D \end{pmatrix}, \qquad A^H = \begin{pmatrix} B^H & O \\ O & D^H \end{pmatrix}.$$

这说明 W 也是 ϕ^*-不变的。再由 $A^HA = AA^H$ 比较左上角分块可得

$$B^HB = BB^H,$$

所以 ϕ_W 是正规变换。 □

定义 3.16. 若域 $\mathbb{F}$ 上的有限维内积空间 V 上的线性变换 ϕ 满足 $\phi^* = -\phi$，则称 ϕ 是反自伴变换。若方阵 $A \in \mathbb{F}^{n\times n}$ 满足 $A^H = -A$，则称 A 是反自伴矩阵。特别地，当 $\mathbb{F} = \mathbb{C}$ 时，称 A 是反 Hermite 矩阵；当 $\mathbb{F} = \mathbb{R}$ 时，A 就是反对称实矩阵。

定理 3.17. *设 ϕ 是有限维酉空间 V 上的线性变换，那么如下陈述等价：*

1. *ϕ 是反自伴变换，即在 V 的任意一组标准正交基下的矩阵为反 Hermite 矩阵；*
2. *ϕ 在 V 的某组标准正交基下的矩阵是对角元均为纯虚数的对角矩阵；*
3. *ϕ 的特征值均为纯虚数且 V 有一组由 ϕ 的特征向量构成的标准正交基。*

命题 3.18. *设 ϕ 是 n 维内积空间 V 上的线性变换，那么下面的陈述等价：*

1. *ϕ 是保距变换；*
2. *对 V 的任意标准正交基 α，$\phi(\alpha)$ 仍是标准正交基；*
3. *存在 V 的一组标准正交基 α，使得 $\phi(\alpha)$ 是标准正交基；*
4. *在 V 的任意标准正交基下，ϕ 的矩阵是正交矩阵或酉矩阵；*
5. *在 V 的某组标准正交基下，ϕ 的矩阵是正交矩阵或酉矩阵。*

定理 3.19. 设 ϕ 是 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 的线性变换。那么如下陈述等价：

1. V 有一组由 ϕ 的特征向量构成的标准正交基，且所有特征值的模均为 1；
2. ϕ 是一个特征多项式在 $\mathbb{F}$ 上分裂的保距变换。

本节小结

- Schur 定理把特征多项式分裂的线性变换化为标准正交基下的上三角形矩阵。
- 正规变换正是“可在标准正交特征向量基下对角化”的核心条件。
- 酉空间中复多项式总分裂，因此正规变换与酉对角化等价。
- 自伴、反自伴、保距变换都可作为正规变换的重要特例来理解。

4 正交投影与谱分解

本节目标

- 将矩阵对角化中的分块投影解释为线性变换的投影。
- 掌握正交投影的自伴幂等刻画。
- 理解正规变换的谱分解及其矩阵版本。

4.1 投影与正交投影

定义 4.1. 对线性空间 V 及其子空间 W_1, W_2，若 $V = W_1 \oplus W_2$，则可定义线性变换

$$\phi : V \to V, \qquad v_1 + v_2 \mapsto v_1, \qquad \forall v_1 \in W_1,\ v_2 \in W_2.$$

此线性变换 ϕ 称为 V 在子空间 W_1 上关于子空间 W_2 的投影，简称 V 的投影。

事实 4.2. 设 ϕ 是 V 在子空间 W_1 上关于子空间 W_2 的投影。那么

1. W_1, W_2 是 ϕ-不变子空间且 $\phi|_{W_1} = \mathrm{id}_{W_1}$，$\phi|_{W_2} = 0$；
2. $\operatorname{Im}\phi = W_1$，$\operatorname{Ker}\phi = W_2$。

命题 4.3. 线性空间 V 上的线性变换 ϕ 是 V 的一个投影当且仅当 ϕ 是幂等变换，即 $\phi^2 = \phi$。

证明. 必要性由投影的定义直接验证即得。下面证明充分性。假设 $\phi^2 = \phi$。对任意 $v \in V$ 有

$$v = \phi(v) + v - \phi(v).$$

显然 $\phi(v) \in \operatorname{Im}\phi$，并且 $\phi(v - \phi(v)) = \phi(v) - \phi^2(v) = 0$，从而 $v - \phi(v) \in \operatorname{Ker}\phi$。这表明 $V = \operatorname{Im}\phi + \operatorname{Ker}\phi$。若 $v \in \operatorname{Im}\phi \cap \operatorname{Ker}\phi$，则存在 $w \in V$ 使得 $v = \phi(w)$，从而

$$v = \phi(w) = \phi^2(w) = \phi(v) = 0.$$

因此 $V = \operatorname{Im}\phi \oplus \operatorname{Ker}\phi$，再由分解及定义即知 ϕ 是 V 在 $\operatorname{Im}\phi$ 上关于 $\operatorname{Ker}\phi$ 的投影。 □

定义 4.4. 若内积空间 V 关于子空间 W 有正交补分解 $V = W \oplus W^\perp$，则 V 在 W 上关于 $W^\perp$ 的投影

$$\phi : V \to V, \qquad v_1 + v_2 \mapsto v_1, \qquad \forall v_1 \in W,\ v_2 \in W^\perp$$

称为 V 在子空间 W 上的正交投影，简称 V 的正交投影。

命题 4.5. 内积空间 V 上的线性变换 ϕ 是 V 的一个正交投影当且仅当 ϕ 是自伴幂等变换，即

$$\phi^2 = \phi = \phi^*.$$

证明. 必要性。设 ϕ 是在子空间 W 上的正交投影，那么 $V = W \oplus W^\perp$。由命题 3.5.1 有 $\phi^2 = \phi$。任取 $u, v \in V$ 并作分解 $u = u_1 + u_2$，$v = v_1 + v_2$，其中 $u_1, v_1 \in W$，$u_2, v_2 \in W^\perp$。那么

$$(\phi(u), v) = (u_1, v_1 + v_2) = (u_1, v_1), \qquad (u, \phi(v)) = (u_1 + u_2, v_1) = (u_1, v_1).$$

故 $\phi = \phi^*$。

充分性。设 $\phi^2 = \phi = \phi^*$。由命题 3.5.1，$V = \operatorname{Im}\phi \oplus \operatorname{Ker}\phi$。若 $u \in \operatorname{Im}\phi$，$v \in \operatorname{Ker}\phi$，则 $u = \phi(u)$，从而

$$(u, v) = (\phi(u), v) = (u, \phi(v)) = 0.$$

这表明 $\operatorname{Ker}\phi \subseteq (\operatorname{Im}\phi)^\perp$。反向包含由直和分解和维数比较得到，因此 $\operatorname{Ker}\phi = (\operatorname{Im}\phi)^\perp$，故 ϕ 是正交投影。 □

4.2 谱分解定理

定理 4.6. 设 ϕ 是域 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上有限维内积空间 V 上的正规变换，特征多项式 $f_\phi(x)$ 在 $\mathbb{F}$ 上分裂，其所有不同的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_k$。对 $i = 1, 2, \ldots, k$，令 ϕ_i 是 V 到 V_{λ_i} 的正交投影。那么下列结论成立：

1. $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$；
2. $V_{\lambda_i}^\perp = \bigoplus_{j \neq i} V_{\lambda_j}$，$i = 1, 2, \ldots, k$；
3. $\phi_i \phi_j = \delta_{ij} \phi_i$，$i, j = 1, 2, \ldots, k$；
4. $\mathrm{id}_V = \phi_1 + \phi_2 + \cdots + \phi_k$；
5. $\phi = \lambda_1 \phi_1 + \lambda_2 \phi_2 + \cdots + \lambda_k \phi_k$。

特别地，上述结论对有限维欧氏空间上的自伴变换成立。

证明. 由定理 3.2.2 知 ϕ 可对角化，故第一条成立。令 $W_i = \bigoplus_{j \neq i} V_{\lambda_j}$。由命题 3.2.1 有 $V_{\lambda_i} \perp V_{\lambda_j}$，$\forall i \neq j$，因此 $W_i \subseteq V_{\lambda_i}^\perp$。由第一条可得二者维数相等，于是 $W_i = V_{\lambda_i}^\perp$。

由第一、二条，对 $i = 1, 2, \ldots, k$ 有正交投影

$$\phi_i : V \to V_{\lambda_i}, \qquad v = v_1 + v_2 + \cdots + v_k \mapsto v_i, \qquad v_l \in V_{\lambda_l}.$$

于是 $(\phi_i \phi_j)(v) = \phi_i(v_j) = \delta_{ij} v_i = \delta_{ij} \phi_i(v)$，第三条成立。第四条由上式直接得到。最后，

$$\phi(v) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = (\lambda_1 \phi_1 + \lambda_2 \phi_2 + \cdots + \lambda_k \phi_k)(v).$$

□

上述定理中 ϕ 的互异特征值之集 $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ 称为 ϕ 的谱，分解

$$\mathrm{id}_V = \phi_1 + \phi_2 + \cdots + \phi_k$$

称为由 ϕ 诱导的恒等算子解消，分解

$$\phi = \lambda_1\phi_1 + \lambda_2\phi_2 + \cdots + \lambda_k\phi_k$$

称为 ϕ 的谱分解。若不计特征值的顺序，则 ϕ 的谱分解是唯一的。

下面给出矩阵谱分解的具体算法。对 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上的实对称矩阵或复正规矩阵 A，存在正交矩阵或酉矩阵 U 使得

$$U^H AU = \mathrm{diag}\{\lambda_1 I_{n_1}, \ldots, \lambda_k I_{n_k}\},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$ 是 A 的所有不同特征值。若

$$U = (\xi_{11}, \ldots, \xi_{1n_1}, \ldots, \xi_{k1}, \ldots, \xi_{kn_k})$$

的列向量是 $\mathbb{F}^n$ 的一组标准正交基，并且

$$\xi_{i1}, \xi_{i2}, \ldots, \xi_{in_i}$$

是特征子空间 V_{λ_i} 的一组标准正交基，$i = 1, 2, \ldots, k$，那么有谱分解

$$A = U\,\mathrm{diag}\{\lambda_1 I_{n_1}, \ldots, \lambda_k I_{n_k}\}U^H = \sum_{i=1}^{k} \lambda_i \sum_{j=1}^{n_i} \xi_{ij}\xi_{ij}^H = \sum_{i=1}^{k} \lambda_i A_i,$$

其中

$$A_i = U\,\mathrm{diag}\{O_{n_1\times n_1}, \ldots, I_{n_i}, \ldots, O_{n_k\times n_k}\}U^H = \sum_{j=1}^{n_i} \xi_{ij}\xi_{ij}^H, \qquad i = 1, 2, \ldots, k.$$

容易直接验证 A_i 满足矩阵谱分解推论中的各条性质。

例 4.7. 设

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

取

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

直接计算可知 u_1, u_2, u_3 分别是 A 关于特征值 $3, 1, 4$ 的单位特征向量。到对应特征子空间

的正交投影为

$$P_3 = u_1 u_1^T = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$P_1 = u_2 u_2^T = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$P_4 = u_3 u_3^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

于是

$$P_i P_j = O \quad (i \neq j), \qquad P_1 + P_3 + P_4 = I_3,$$

并且谱分解写成

$$A = 1P_1 + 3P_3 + 4P_4.$$

例如

$$A^2 = 1^2 P_1 + 3^2 P_3 + 4^2 P_4 = \begin{pmatrix} 5 & 4 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 16 \end{pmatrix}.$$

例 4.8. 设正规变换 ϕ 的互异特征值为 $\lambda_1, \ldots, \lambda_k$，对应谱投影为 $\phi_1, \ldots, \phi_k$。证明：

$$\phi_i = \prod_{j \neq i} \frac{\phi - \lambda_j \mathrm{id}_V}{\lambda_i - \lambda_j}, \qquad i = 1, 2, \ldots, k.$$

证明. 令

$$p_i(x) = \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j}.$$

则 $p_i(\lambda_i) = 1$，且当 $j \neq i$ 时 $p_i(\lambda_j) = 0$。由谱分解 $\phi = \sum_{j=1}^k \lambda_j \phi_j$ 以及 $\phi_i \phi_j = \delta_{ij} \phi_i$ 可知，对任意多项式 $g(x)$ 有

$$g(\phi) = \sum_{j=1}^k g(\lambda_j) \phi_j.$$

取 $g = p_i$，得到

$$p_i(\phi) = \sum_{j=1}^k p_i(\lambda_j) \phi_j = \phi_i,$$

即得所证公式。 □

推论 4.9. 设 ϕ 是 $\mathbb{F}$ 上有限维内积空间 V 上的线性变换且特征多项式 $f_\phi(x)$ 在 $\mathbb{F}$ 上分裂。那么 ϕ 是保距变换当且仅当 ϕ 正规且其特征值的模均为 1。

推论 4.10. 设 ϕ 是 $\mathbb{F}$ 上有限维内积空间 V 上的正规变换且特征多项式 $f_\phi(x)$ 在 $\mathbb{F}$ 上分裂。那么 ϕ 是自伴的当且仅当 ϕ 的特征值均为实数。

推论 4.11. 设 A 是 n 阶实对称矩阵或复正规矩阵，其不同的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_k$，其中每个 λ_i 的代数重数是 n_i。那么存在 $\mathbb{R}$ 或 $\mathbb{C}$ 上的 n 阶矩阵 $A_1, A_2, \ldots, A_k$ 使得如下陈述成立：

1. $A = \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k$；
2. $A_i A_j = O$，$\forall i, j = 1, 2, \ldots, k,\ i \neq j$；
3. A_i 是秩为 n_i 的幂等 Hermite 矩阵，$i = 1, 2, \ldots, k$；
4. $I_n = A_1 + A_2 + \cdots + A_k$。

课堂提示

谱分解把"对角矩阵的每个对角块"解释为到相应特征子空间的正交投影。它不仅是对角化的等价表达，也便于计算多项式 $g(\phi)$、判断保距性和自伴性。

本节小结

- 投影等价于幂等变换，正交投影等价于自伴幂等变换。
- 正规变换的不同特征子空间两两正交。
- 谱分解为 $\phi = \sum_i \lambda_i \phi_i$，其中 ϕ_i 是到特征子空间的正交投影。
- 实对称矩阵和复正规矩阵都有相应的矩阵谱分解。

5 二次型与自伴变换的正定性

本节目标

- 统一实二次型、自伴变换、实对称矩阵的正定性定义。
- 掌握正定与半正定的等价刻画。
- 理解 Hermite 矩阵和酉二次型的对应版本。

5.1 欧氏空间与实二次型

设 q 是 n 维欧氏空间 V 上的二次型，在 V 的一组标准正交基 $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的矩阵为 A，从而 A 是实对称矩阵，那么

$$q(v) = x^T A x, \qquad x = [v]_\alpha \in \mathbb{R}^n.$$

定义 V 上线性变换 ϕ 使得 $[\phi]_\alpha = A$。那么 ϕ 是一个自伴变换，且

$$(\phi(v), v) = x^T A x = q(v).$$

定义 5.1. 设 q, A 和 ϕ 如上所述。若对任意非零向量 $v \in V$，有

$$q(v) = (\phi(v), v) \geq 0 \quad (\leq 0),$$

也等价于对任意非零向量 $x \in \mathbb{R}^n$，有

$$x^T A x \geq 0 \quad (\leq 0),$$

则称实二次型 q、自伴变换 ϕ 和实对称矩阵 A 是半正定的（半负定的）；若上面两式中不等号严格成立，则称实二次型 q、自伴变换 ϕ 和实对称矩阵 A 是正定的（负定的）；若它们既非半正定的也非半负定的，则称其是不定的。

设 A 为 n 阶方阵。对 $k = 1, 2, \ldots, n$，将 A 的前 k 行与前 k 列交叉位置元素构成的矩阵称作 A 的 k 阶顺序主子阵并记为 A_k，称行列式 $|A_k|$ 为 A 的 k 阶顺序主子式。

定理 5.2. *设 A 为 n 阶实对称矩阵，那么下面的陈述等价：*

1. *$\mathbb{R}^n$ 上的实二次型 $f(x) = x^T A x$ 正定，即实对称矩阵 A 正定；*
2. *上述 f 的正惯性指数等于 n；*
3. *A 的所有特征值均为正数；*
4. *A 与单位矩阵合同；*
5. *存在 n 阶可逆实矩阵 B，使得 $A = B^T B$；*
6. *A 的 k 阶顺序主子式为正数，$k = 1, 2, \ldots, n$。*

证明. (1) $\Rightarrow$ (2)。设 M 为可逆矩阵，使得 $M^T A M = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}$。若存在 $d_i < 0$，则对 $x = M e_i$ 有

$$x^T A x = e_i^T M^T A M e_i = d_i < 0,$$

与 A 正定矛盾。因此 $d_i > 0$。

(2) $\Rightarrow$ (3)。设 $\lambda_1, \ldots, \lambda_n$ 是 A 的所有特征值，那么存在正交矩阵 U 使得

$$U^T A U = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}.$$

由正惯性指数为 n 可得 $\lambda_i > 0$。

(3) $\Rightarrow$ (4)。令 $D = \mathrm{diag}\{1/\sqrt{\lambda_1}, 1/\sqrt{\lambda_2}, \ldots, 1/\sqrt{\lambda_n}\}$，则

$$(UD)^T A U D = D^T U^T A U D = I_n.$$

(4) $\Rightarrow$ (5)。若 $M^T A M = I_n$，则 $A = B^T B$，其中 $B = M^{-1}$。

(5) $\Rightarrow$ (1)。若 $A = B^T B$ 且 B 可逆，则对任意非零向量 $x \in \mathbb{R}^n$，有 $Bx \neq 0$，从而

$$f(x) = x^T A x = x^T B^T B x = (Bx)^T B x > 0.$$

(1) $\Rightarrow$ (6)。对任意非零向量 $x \in \mathbb{R}^k$，由分块形式可得 $x^T A_k x > 0$，因此 A_k 正定。由已证的 (1) $\Rightarrow$ (3)，A_k 的特征值均为正数，因此 $|A_k| > 0$。

(6) $\Rightarrow$ (1)。对 n 作归纳。设结论对 $n-1$ 成立，并设 A 的顺序主子式均为正数。记

$$A = \begin{pmatrix} A_{n-1} & b \\ b^T & d \end{pmatrix}.$$

由归纳假设，A_{n-1} 正定且可逆。令 $d' = d - b^T A_{n-1}^{-1} b$，则 $|A| = |A_{n-1}| d'$，故 $d' > 0$。于是 A 合同于 $\mathrm{diag}(A_{n-1}, d')$，从而 A 正定。 □

推论 5.3. n 阶实对称矩阵 A 负定当且仅当 A 的 k 阶顺序主子式 $|A_k|$ 的符号为 $(-1)^k$，$k=1,2,\ldots,n$。

例 5.4. 设 $a_1,a_2,\ldots,a_n$ 为互不相同的非负整数。证明：n 阶实对称矩阵

$$A=\left(\frac{1}{a_i+a_j+1}\right)_{n\times n}$$

正定。

证明. 注意到

$$\frac{1}{a_i+a_j+1}=\int_0^1 t^{a_i+a_j}\,\mathrm{d}t.$$

对任意非零向量 $x=(x_1,\ldots,x_n)^T\in\mathbb{R}^n$，有

$$\begin{aligned} x^TAx &= \sum_{i=1}^n\sum_{j=1}^n x_ix_j\int_0^1 t^{a_i+a_j}\,\mathrm{d}t \\ &= \int_0^1\left(\sum_{i=1}^n x_it^{a_i}\right)^2\mathrm{d}t. \end{aligned}$$

由于 $a_1,a_2,\ldots,a_n$ 互不相同，函数 $t^{a_1},t^{a_2},\ldots,t^{a_n}$ 线性无关。若 $x\neq 0$，则

$$\sum_{i=1}^n x_it^{a_i}$$

不是零函数，因此其平方在 $[0,1]$ 上连续、非负且不恒为零，从而

$$\int_0^1\left(\sum_{i=1}^n x_it^{a_i}\right)^2\mathrm{d}t>0.$$

故对任意非零 $x\in\mathbb{R}^n$ 都有 $x^TAx>0$，所以 A 正定。 □

注 5.5. 也可以尝试证明这个矩阵的行列式为正，然后用归纳法说明它是正定矩阵。

例 5.6. 判断

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

的正定性。

对任意 $x=(x_1,x_2,x_3)^T\in\mathbb{R}^3$，有

$$x^TAx=x_1^2+2x_1x_2+x_2^2+2x_3^2=(x_1+x_2)^2+2x_3^2\geq 0,$$

所以 A 半正定。取非零向量 $x=(1,-1,0)^T$，则 $x^TAx=0$，因此 A 不是正定矩阵。

从特征值角度也可验证：前两个坐标上的矩阵 $\begin{pmatrix}1&1\\1&1\end{pmatrix}$ 的特征值为 $2,0$，第三个方向的特征值为 2，故 A 的特征值为 $2,2,0$，全为非负但不全为正。

定理 5.7. 设 A 为 n 阶实对称矩阵，$r=\mathrm{rank}(A)$，那么下面的陈述等价：

1. $\mathbb{R}^n$ 上的实二次型 $f(x)=x^TAx$ 半正定，即 A 半正定；
2. 上述 f 的正惯性指数为 r，负惯性指数为 0；
3. A 的所有特征值均非负；
4. A 与 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{n\times n}$ 合同；
5. 存在秩为 r 的 n 阶实矩阵 B，使得 $A=B^TB$。

5.2 酉空间与酉二次型

定义 5.8. 1. 设 ϕ 是 n 维酉空间 V 上的自伴变换，在一组标准正交基下的矩阵为 A，称 V 上实值函数 $q(v)=(\phi(v),v)$ 为一个酉二次型。

2. 若对任意非零向量 $v\in V$，有

$$q(v)=(\phi(v),v)\geq 0 \quad (\leq 0),$$

也等价于对任意非零向量 $x\in\mathbb{C}^n$，有

$$x^HAx\geq 0 \quad (\leq 0),$$

则称酉二次型 q、自伴变换 ϕ 和 Hermite 矩阵 A 是半正定的（半负定的）；若上面两式中不等号严格成立，则称它们是正定的（负定的）；若它们既非半正定的也非半负定的，则称其是不定的。

定理 5.9. 设 A 为 n 阶 *Hermite* 矩阵，那么下面的陈述等价：

1. $\mathbb{C}^n$ 上的酉二次型 $f(x)=x^HAx$ 正定，即 *Hermite* 矩阵 A 正定；
2. A 的所有特征值均为正数；
3. A 与单位矩阵 *Hermite* 合同，即存在可逆复矩阵 M，使得 $M^HAM=I_n$；
4. 存在 n 阶可逆复矩阵 B，使得 $A=B^HB$；
5. A 的所有 k 阶顺序主子式均为正数，$k=1,2,\dots,n$。

推论 5.10. n 阶 *Hermite* 矩阵 A 负定当且仅当 A 的 k 阶顺序主子式 $|A_k|$ 的符号为 $(-1)^k$，$k=1,2,\dots,n$。

定理 5.11. 设 A 为 n 阶 *Hermite* 矩阵，$r=\mathrm{rank}(A)$，那么下面的陈述等价：

1. $\mathbb{C}^n$ 上的酉二次型 $f(x)=x^HAx$ 半正定，即 A 半正定；
2. A 的所有特征值均非负；
3. A 与 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{n\times n}$ 是 *Hermite* 合同的；
4. 存在秩为 r 的 n 阶复矩阵 B，使得 $A=B^HB$。

推论 5.12. *对有限维内积空间 V 上的自伴变换 ϕ，下面的陈述等价：*

1. *ϕ 是正定（半正定）变换；*
2. *ϕ 的所有特征值均是正的（非负的）；*
3. *存在 V 上（可逆）线性变换 ψ，使得 $\phi = \psi^*\psi$。*

命题 5.13. *一个 n 阶实对称（Hermite）矩阵 A 是正定的当且仅当 A 是某个 n 维欧氏空间（酉空间）的内积在一组标准正交基下的度量矩阵。*

例 5.14. 设 A 是 n 阶正定矩阵，B 是 $m \times n$ 实矩阵。证明：B^TAB 正定当且仅当 $\mathrm{rank}(B) = n$。

证明. 对任意 $x \in \mathbb{R}^n$，有

$$x^TB^TABx = (Bx)^TA(Bx).$$

若 $\mathrm{rank}(B) = n$，则任意非零 x 都满足 $Bx \neq 0$。由于 A 正定，故

$$(Bx)^TA(Bx) > 0,$$

所以 B^TAB 正定。

反过来，若 $\mathrm{rank}(B) < n$，则存在非零 $x \in \mathbb{R}^n$ 使得 $Bx = 0$，于是

$$x^TB^TABx = 0,$$

这与 B^TAB 正定矛盾。因此必须有 $\mathrm{rank}(B) = n$。 □

课堂提示

正定性可以从多个等价角度判断：二次型取值、惯性指数、特征值、合同标准形、B^HB 或 B^TB 分解、顺序主子式。课堂计算中顺序主子式判别最直接，理论推导中特征值和分解更灵活。

本节小结

- 实二次型、自伴变换、实对称矩阵的正定性是一组等价语言。
- 正定实对称矩阵等价于全部特征值为正，也等价于所有顺序主子式为正。
- 半正定对应非负特征值和秩为 r 的平方分解。
- Hermite 矩阵和酉二次型有平行的正定与半正定判别。

6 极分解与奇异值分解

本节目标

- 理解半正定变换平方根的存在性和唯一性。
- 掌握线性变换和矩阵的极分解。
- 掌握线性变换和矩阵的奇异值分解。

6.1 半正定平方根与极分解

命题 6.1. *对有限维内积空间 V 上的自伴变换 ϕ，下面的陈述等价：*

1. *ϕ 是半正定（正定）变换；*
2. *ϕ 有半正定（正定）的平方根；*
3. *ϕ 有自伴（自伴且可逆）的平方根。*

证明. 由第 3 章中结论，存在 V 的标准正交基 α，使得 ϕ 在 α 下的矩阵是一个实对角矩阵

$$\operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}.$$

若 ϕ 半正定，则 $\lambda_i \geq 0$。令 φ 在 α 下的矩阵为

$$\operatorname{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\},$$

则 φ 半正定且 $\varphi^2 = \phi$；若 ϕ 正定，则 φ 亦然。第二条推出第三条显然。若 φ 自伴且 $\varphi^2 = \phi$，那么 $\phi = \varphi^2 = \varphi^*\varphi$ 是半正定的，并且若 φ 可逆，则 ϕ 正定。 □

命题 6.2. *若 φ, φ' 是有限维内积空间 V 上的半正定变换且 $\varphi^2 = \varphi'^2$，则 $\varphi = \varphi'$。*

证明. 记 $\phi = \varphi^2 = \varphi'^2$。由 φ 半正定知 φ 可对角化且有非负特征值，那么 φ 关于特征值 $\lambda \geq 0$ 的特征子空间即为 ϕ 关于特征值 λ^2 的特征子空间。对 φ' 作类似讨论可以看出，φ 和 φ' 可对角化且有相同的特征值和特征子空间，因此 $\varphi = \varphi'$。 □

定理 6.3. *设 V 是有限维内积空间，ϕ 是 V 上的线性变换，那么有*

1. *存在 V 上的保距线性变换 ψ 和唯一的半正定变换 φ，使得 $\phi = \psi\varphi$。此时 φ 即为 $\phi^*\phi$ 唯一的半正定平方根；*
2. *若 ϕ 可逆，则第一条中分解唯一。*

证明. 由推论 6.2.3 或由定义直接验证可知 $\phi^*\phi$ 是半正定变换。若 $\phi = \psi\varphi$，其中 ψ 是保距线性变换，φ 是半正定变换，则

$$\phi^*\phi = \varphi^*\psi^*\psi\varphi = \varphi^*\varphi = \varphi^2,$$

故 φ 是 $\phi^*\phi$ 唯一的半正定平方根。

下面证明分解存在。由 $\varphi^2 = \phi^*\phi$，对任意 $v \in V$ 有

$$\|\varphi(v)\|^2 = (\varphi(v), \varphi(v)) = \big(v, \varphi^2(v)\big) = (v, \phi^*\phi(v)) = (\phi(v), \phi(v)) = \|\phi(v)\|^2.$$

定义映射

$$\psi_1 : \operatorname{Im}\varphi \to \operatorname{Im}\phi, \qquad \varphi(v) \mapsto \phi(v).$$

若 $\varphi(u) = \varphi(v)$，则

$$\|\phi(u) - \phi(v)\| = \|\phi(u-v)\| = \|\varphi(u-v)\| = 0,$$

从而 $\phi(u) = \phi(v)$，故 ψ_1 合理定义。由上式可知 ψ_1 保长度从而是单射；又 $\operatorname{Ker}\varphi = \operatorname{Ker}\phi$，由维数定理有 $\dim\operatorname{Im}\varphi = \dim\operatorname{Im}\phi$，因此 ψ_1 是保距同构。

任取保距同构 $\psi_2 : (\operatorname{Im}\varphi)^\perp \to (\operatorname{Im}\phi)^\perp$，并定义

$$\psi = \psi_1 \oplus \psi_2 : V = \operatorname{Im}\varphi \oplus (\operatorname{Im}\varphi)^\perp \to V = \operatorname{Im}\phi \oplus (\operatorname{Im}\phi)^\perp.$$

那么 ψ 是保距线性变换，且 $\psi\varphi(v) = \psi_1(\varphi(v)) = \phi(v)$。若 ϕ 可逆，则 φ 亦可逆，从而 $\psi = \phi\varphi^{-1}$。 □

注 6.4. 需要指出，对酉空间上线性变换 ϕ 的极分解 $\phi = \psi\varphi$，其中 ψ 是保距线性变换，φ 是半正定变换，ψ 和 φ 都可酉相似对角化，但未必有同时使其对角化的标准正交基。

推论 6.5. *设 A 为 n 阶实（复）方阵，那么有*

1. *存在正交（酉）矩阵 U 和唯一的半正定实对称（Hermite）矩阵 B，使得 $A = UB$。此时 B 即为 A^HA 唯一的半正定平方根；*
2. *若 A 可逆，则第一条中分解唯一。*

例 6.6. 设

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

则

$$A^TA = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \qquad B = (A^TA)^{1/2} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

由于 A 可逆，极分解唯一，且

$$U = AB^{-1} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

于是

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

其中 U 是正交矩阵，B 是半正定实对称矩阵。

例 6.7. 再看一个不可逆矩阵的极分解。设

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

显然 A 不可逆。先计算

$$A^TA = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

因此

$$B = (A^TA)^{1/2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

此时 B 的对角元中有且仅有一个 0。由于 B 不可逆，不能用 $U = AB^{-1}$ 来求正交矩阵 U。观察到

$$\operatorname{Im} B = \operatorname{Span}\{e_2, e_3\}, \qquad \operatorname{Im} A = \operatorname{Span}\{e_1, e_3\},$$

并且

$$Be_2 = 2e_2, \qquad Ae_2 = 2e_1, \qquad Be_3 = 3e_3, \qquad Ae_3 = 3e_3.$$

所以只需选择一个正交矩阵 U，使得

$$Ue_2 = e_1, \qquad Ue_3 = e_3.$$

例如取

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

则 U 是正交矩阵，并且

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

于是

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

就是 A 的一个极分解。

6.2 奇异值分解

设 n 维内积空间 V 上有线性变换 ϕ。半正定变换 $\phi^*\phi$ 有非负实特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$。令

$$s_i = \sqrt{\lambda_i}, \qquad i = 1, 2, \ldots, n,$$

称为 ϕ 的奇异值。若 $\lambda = \lambda_i$ 是 $\phi^*\phi$ 的一个特征值，则其特征子空间的维数称为奇异值 $s = s_i$ 的重数。

定理 6.8. 设 n 维内积空间 V 上的线性变换 ϕ 有奇异值 $s_1, s_2, \ldots, s_n$，那么存在 V 的标准正交基 $\xi_1, \xi_2, \ldots, \xi_n$ 和 $\eta_1, \eta_2, \ldots, \eta_n$ 使得

$$\phi(\xi_i) = s_i \eta_i, \qquad i = 1, 2, \ldots, n.$$

此时有

$$\phi(v) = (v, \xi_1)\, s_1\eta_1 + (v, \xi_2)\, s_2\eta_2 + \cdots + (v, \xi_n)\, s_n\eta_n, \qquad \forall v \in V.$$

证明. 作极分解 $\phi = \psi\varphi$，其中 ψ 为保距线性变换，φ 为 $\phi^*\phi$ 的半正定平方根。那么 φ 的特征值即为 $s_1, s_2, \ldots, s_n$，且存在 V 的标准正交基 $\xi_1, \xi_2, \ldots, \xi_n$，使得 ξ_i 是 φ 关于 s_i 的特征向量。

令 $\eta_i = \psi(\xi_i)$。由于 ψ 是保距线性变换，因此 $\eta_1, \eta_2, \ldots, \eta_n$ 亦是 V 的标准正交基，且

$$\phi(\xi_i) = \psi(\varphi(\xi_i)) = \psi(s_i\xi_i) = s_i\eta_i.$$

又

$$v = (v, \xi_1)\,\xi_1 + (v, \xi_2)\,\xi_2 + \cdots + (v, \xi_n)\,\xi_n,$$

以 ϕ 作用即得定理最后一个等式。 □

推论 6.9. 对 n 阶实（复）矩阵 A，存在正交（酉）矩阵 P, Q，使得

$$A = Q\,\mathrm{diag}\{s_1, s_2, \ldots, s_n\}P^H,$$

其中 $s_1, s_2, \ldots, s_n$ 是半正定矩阵 A^HA 特征值的非负平方根。

例 6.10. 继续设

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

由 $A^TA = \mathrm{diag}\{1, 4\}$ 可知 A 的奇异值为 $1, 2$。取

$$\xi_1 = e_1, \quad \xi_2 = e_2, \qquad \eta_1 = e_2, \quad \eta_2 = e_1,$$

则

$$A\xi_1 = e_2 = 1 \cdot \eta_1, \qquad A\xi_2 = 2e_1 = 2 \cdot \eta_2.$$

因此可取

$$P = I_2, \qquad Q = (\eta_1, \eta_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \Sigma = \mathrm{diag}\{1, 2\},$$

并得到

$$A = Q\Sigma P^T.$$

例 6.11. 设 ϕ 是有限维内积空间 V 上的线性变换。证明：ϕ 的零奇异值的重数等于 $\dim\operatorname{Ker}\phi$。

证明. 奇异值来自半正定变换 $\phi^*\phi$ 的特征值的非负平方根。因此零奇异值的重数就是 $\phi^*\phi$ 关于特征值 0 的特征子空间维数，即 $\dim\operatorname{Ker}(\phi^*\phi)$。

对任意 $v \in V$，

$$\phi^*\phi(v) = 0 \quad\Longrightarrow\quad \|\phi(v)\|^2 = (\phi(v), \phi(v)) = (\phi^*\phi(v), v) = 0 \quad\Longrightarrow\quad \phi(v) = 0.$$

反向包含 $\operatorname{Ker}\phi \subseteq \operatorname{Ker}(\phi^*\phi)$ 显然成立，所以

$$\operatorname{Ker}(\phi^*\phi) = \operatorname{Ker}\phi.$$

故零奇异值的重数等于 $\dim\operatorname{Ker}\phi$。 □

课堂提示

极分解可类比非零复数的极坐标：$\phi = \psi\varphi$ 中 ψ 负责“保长度的旋转或反射”，φ 是由 $\phi^*\phi$ 决定的半正定“伸缩”。SVD 则进一步把这种伸缩放到两组标准正交基之间。

本节小结

- 半正定自伴变换存在唯一半正定平方根。
- 极分解为 $\phi = \psi\varphi$，其中 ψ 保距，$\varphi = (\phi^*\phi)^{1/2}$。
- 奇异值是 $\phi^*\phi$ 特征值的非负平方根。
- SVD 给出 $\phi(\xi_i) = s_i\eta_i$，矩阵形式为 $A = Q\Sigma P^H$。

7 思考与练习

A 组：基础核对题

A1. 对有限维内积空间 V 和 $\phi, \psi \in \mathrm{End}(V)$，证明：

$$(\phi+\psi)^* = \phi^* + \psi^*, \quad (c\phi)^* = \bar{c}\,\phi^*, \quad (\phi\psi)^* = \psi^*\phi^*, \quad \phi^{**} = \phi, \quad \mathrm{id}_V^* = \mathrm{id}_V.$$

A2. 设

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

求 l_A 的伴随，并判断 A 是否自伴、正规、保距。

A3. 设 ϕ 是有限维内积空间 V 上的线性变换，$\psi_1 = \phi + \phi^*$，$\psi_2 = \phi\phi^*$。证明：$\psi_1 = \psi_1^*$，$\psi_2 = \psi_2^*$。

A4. 判断矩阵

$$A = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 1 & -2 \\ 3 & -2 & 8 \end{pmatrix}$$

是否正定，并说明理由。

A5. 设 ϕ 是有限维内积空间上的自伴变换。证明：若 ϕ 正定，则 ϕ 可逆。

B 组：证明提高题

B1. 设 ϕ 是有限维内积空间 V 上的线性变换。证明：若 ϕ 可逆，则 ϕ^* 也可逆且 $(\phi^*)^{-1} = (\phi^{-1})^*$。

B2. 设 ϕ 是内积空间 V 上的线性变换。证明：

$$(\mathrm{Im}\,\phi^*)^\perp = \mathrm{Ker}\,\phi.$$

若 V 维数有限，进一步证明 $\mathrm{Im}\,\phi^* = (\mathrm{Ker}\,\phi)^\perp$。

B3. 设 ϕ 是正规变换。证明：若 $\lambda_1 \neq \lambda_2$ 是 ϕ 的特征值，则 $V_{\lambda_1} \perp V_{\lambda_2}$。

B4. 设 $A=(a_{ij})_{n\times n}$ 是一个实矩阵，且对任意 $1\le i\le n$，有

$$2a_{ii}>\sum_{j=1}^{n}|a_{ij}|.$$

这样的矩阵称为绝对对角占优矩阵。证明：

(i) $|A|>0$；

(ii) 如果 A 对称，则 A 正定。

B5. 设 ϕ 是内积空间 V 上的线性变换。证明：ϕ 是保距同构当且仅当 ϕ 的所有奇异值都等于 1。

B6. 设 A 是一个正定矩阵，B 是一个半正定矩阵。证明：如果 $A-B$ 半正定，则 $|A|\ge|B|$。

B7. 设 A,B 均为实对称矩阵或复正规矩阵。证明：$AB=BA$ 当且仅当存在正交矩阵或酉矩阵 U，使得 U^HAU 和 U^HBU 均为对角矩阵。

B8. 设 A,B 是 $m\times n$ 阶实矩阵。证明：$A^TA=B^TB$ 当且仅当存在 m 阶正交矩阵 Q，使得 $A=QB$。

C 组：综合应用题

C1. 对 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$，定义标准内积空间 $\mathbb{F}^3$ 上的线性变换

$$\phi:\mathbb{F}^3\to\mathbb{F}^3,\qquad (z_1,z_2,z_3)\mapsto(z_3,2z_1,3z_2).$$

令 φ 是 $\phi^*\phi$ 的半正定平方根。求一个 $\mathbb{F}^3$ 上的保距线性变换 ψ，使得 $\phi=\psi\varphi$。

C2. 对矩阵

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix},$$

求 A^HA 的特征值和 A 的奇异值，并写出一组奇异值分解。

C3. 证明：对任意 n 阶复矩阵 A，存在 n 阶对角矩阵

$$J=\operatorname{diag}\{\pm1,\pm1,\dots,\pm1\}$$

使得 $|A+J|\ne0$。

C4. 设 A 是正定矩阵，$\alpha_1,\alpha_2,\dots,\alpha_n$ 均为非零的 n 元实列向量，且当 $i\ne j$ 时 $\alpha_i^TA\alpha_j=0$。证明：$\alpha_1,\alpha_2,\dots,\alpha_n$ 线性无关。

C5. 设正规变换 ϕ 的谱分解为

$$\phi=\lambda_1\phi_1+\lambda_2\phi_2+\cdots+\lambda_k\phi_k.$$

证明对任意多项式 $g(x)$ 有

$$g(\phi)=g(\lambda_1)\phi_1+g(\lambda_2)\phi_2+\cdots+g(\lambda_k)\phi_k.$$

C6. 设 A 是一个 n 阶 Hermite 矩阵，B 是一个 Hermite 半正定矩阵。证明：AB 的特征值全为实数。

C7. 设 $A, B \in \mathbb{C}^{n\times n}$，$A$ 的特征值全为 1，且存在正整数 k 使得

$$A^k B = BA^k.$$

证明：$AB = BA$。