

第 02 章 线性代数基础

2.1 课程内容

本章位于绪论之后、矩阵计算之前。第 4--8 章会反复使用本章结论：直接法使用正定性、正交变换和矩阵范数；迭代法使用谱半径与非负矩阵；最小二乘使用投影和 SVD；Krylov 与特征值算法使用不变子空间、Schur 分解和扰动界。

学习这些定义时，要注意它们之间的关系。例如

$$\text{null}(A) = \text{range}(A^H)^\perp, \quad \|A\|_2 = \sigma_1(A), \quad P_{\text{range}(A)} = AA^\dagger. \quad (2.1)$$

这些关系把几何、谱和数值计算连接起来。

Tip

本页要点： 后续算法会反复用到子空间、矩阵分解和扰动界，要把这些概念联系起来理解。

2.2 本章主要内容

本章分为四个部分。

1. 基本概念与常用变换：线性空间、基、线性变换、秩、特征值、正定性、Householder 与 Givens。
2. 距离与范数：向量/矩阵范数、谱半径、非负矩阵与 Perron--Frobenius 理论。
3. 矩阵分解：Schur分解、SVD、低秩近似和 Cos--Sin 分解。
4. 扰动分析：线性系统条件数、逆矩阵扰动、特征值与奇异值扰动界。

后续证明默认有限维空间。实数情形使用转置 $(\cdot)^T$ ，复数情形使用共轭转置 $(\cdot)^H$ 。

Tip

本页要点： 先统一向量、矩阵、范数和谱的记号，后续算法与误差分析都要用到它们。

第2.1节 基础知识与常用变换

2.4 Linear Spaces

线性空间由向量集合 $\mathcal{V}$ 和标量域 $\mathbb{F}$ 组成。向量加法与标量乘法必须满足八条公理。对任意 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ 和 $a, b \in \mathbb{F}$:

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \mathbf{v} + \mathbf{u}, \\ (\mathbf{u} + \mathbf{v}) + \mathbf{w} &= \mathbf{u} + (\mathbf{v} + \mathbf{w}), \\ \mathbf{u} + \mathbf{0} &= \mathbf{u}, \\ \mathbf{u} + (-\mathbf{u}) &= \mathbf{0}, \end{aligned} \quad (2.2)$$

以及

$$\begin{aligned}
a(\mathbf{u} + \mathbf{v}) &= a\mathbf{u} + a\mathbf{v}, \\
(a + b)\mathbf{u} &= a\mathbf{u} + b\mathbf{u}, \\
(ab)\mathbf{u} &= a(b\mathbf{u}), \\
1\mathbf{u} &= \mathbf{u}.
\end{aligned} \tag{2.3}$$

这些公理保证线性组合 $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ 在空间内有明确定义。

Tip

本页要点： 线性空间的闭合性和基的表示是理解投影、分解与子空间算法的基础。

2.5 Basis of A Linear Space

向量组 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset \mathcal{V}$ **线性无关**，是指

$$\sum_{i=1}^n a_i \mathbf{v}_i = \mathbf{0} \implies a_1 = \cdots = a_n = 0. \tag{2.4}$$

它张成 $\mathcal{V}$ ，是指每个 $\mathbf{v} \in \mathcal{V}$ 都可表示为

$$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{v}_i. \tag{2.5}$$

同时线性无关且张成 $\mathcal{V}$ 的向量组称为一组基。若 $\sum_i a_i \mathbf{v}_i = \sum_i b_i \mathbf{v}_i$ ，相减并由线性无关性得 $a_i = b_i$ ，所以基下的表示唯一。

有限维空间中任意两组基含相同数量的向量，这个数量定义为 $\dim \mathcal{V}$ ，也是空间内线性无关向量的最大个数。空间维数不随基改变，坐标分量则取决于所选基。

Note

为什么两组基所含的向量个数相同？ 先证明一个更一般的结论：若 $\mathbf{v}_1, \dots, \mathbf{v}_r$ 线性无关，而 $\mathbf{w}_1, \dots, \mathbf{w}_s$ 张成 $\mathcal{V}$ ，则 $r \leq s$ 。

用指标集记录尚未被换出的 $\mathbf{w}_j$ ：令 $I_0 = \{1, \dots, s\}$ ，并要求

$$\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \cup \{\mathbf{w}_j : j \in I_k\} \text{ 张成 } \mathcal{V}, \quad |I_k| = s - k. \tag{2.6}$$

当 $k=0$ 时，这是因为 I_0 对应的 $\mathbf{w}_1, \dots, \mathbf{w}_s$ 已张成 $\mathcal{V}$ 。假设已经构造出 I_{k-1} ，且 $1 \leq k \leq \min(r, s)$ 。由于 $\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\} \cup \{\mathbf{w}_j : j \in I_{k-1}\}$ 张成 $\mathcal{V}$ ，存在系数使

$$\mathbf{v}_k = \sum_{i=1}^{k-1} a_i \mathbf{v}_i + \sum_{j \in I_{k-1}} b_j \mathbf{w}_j. \tag{2.7}$$

如果所有 $b_j = 0$ ，则 $\mathbf{v}_k \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ ，这与 $\mathbf{v}_1, \dots, \mathbf{v}_r$ 线性无关矛盾。因此存在某个 $j_* \in I_{k-1}$ ，使 $b_{j_*} \neq 0$ 。从上式解出

$$\mathbf{w}_{j_*} = \frac{1}{b_{j_*}} \left(\mathbf{v}_k - \sum_{i=1}^{k-1} a_i \mathbf{v}_i - \sum_{j \in I_{k-1} \setminus \{j_*\}} b_j \mathbf{w}_j \right). \tag{2.8}$$

令 $I_k = I_{k-1} \setminus \{j_*\}$ 。上式说明被换出的 w_{j_*} 属于新集合的张成空间，而原来的其他向量显然也属于新集合的张成空间，所以新集合仍张成 $\mathcal{V}$ ，且 $|I_k| = s - k$ 。

如果 $r > s$ ，递推进行到 $k = s$ 后， $I_s = \emptyset$ ，因此 $v_1, \dots, v_s$ 张成 $\mathcal{V}$ 。于是 $v_{s+1} \in \mathcal{V}$ 必属于 $\text{span}\{v_1, \dots, v_s\}$ ，这与线性无关性矛盾。因此 $r \leq s$ 。当两组向量都为基时，交换两组基的角色可得 $s \leq r$ ，从而 $r = s$ 。这里的“个数”是基中向量的数目，也就是空间的维数 $\dim \mathcal{V}$ 。

Tip

本页要点： 选定基以后，抽象向量才有了可计算的坐标表示。

2.6 Linear Transformation and Matrix

映射 $T: \mathcal{V} \rightarrow \mathcal{W}$ 称为线性变换，如果

$$T(u + v) = Tu + Tv, \quad T(av) = aTv. \quad (2.9)$$

由两式可得 $T(\sum_i a_i v_i) = \sum_i a_i T v_i$ ，所以只要知道基向量的像，就完全确定 T 。

设 $B_1 = (v_1, \dots, v_n)$ 是定义域基， $B_2 = (w_1, \dots, w_m)$ 是值域基。把

$$T v_j = \sum_{i=1}^m a_{ij} w_i \quad (2.10)$$

的系数组成矩阵 $[T]_{B_2}^{B_1} = A = (a_{ij})$ ，则

$$[Tv]_{B_2} = A[v]_{B_1}. \quad (2.11)$$

在该表示下，变换复合对应矩阵乘法，核对应矩阵零空间，像对应矩阵列空间。改变基只改变矩阵表示，不改变变换的核、像维数等内在性质。

Note

矩阵表示是怎样得到的？ 式 (2.10) 中， A 的第 j 列就是 $T v_j$ 在基 B_2 下的坐标向量。

对任意 $v = \sum_j x_j v_j$ ，由线性性及式 (2.10)，有 $T v = \sum_j x_j T v_j = \sum_j x_j \sum_i a_{ij} w_i = \sum_i (\sum_j a_{ij} x_j) w_i$ 。因此 $T v$ 在 B_2 下的第 i 个坐标是 $\sum_j a_{ij} x_j$ ，这正是 $A[v]_{B_1}$ 的第 i 个分量，从而得到式 (2.11)。

Note

为什么复合映射对应矩阵乘法？ 另设线性映射 $U: \mathcal{W} \rightarrow \mathcal{Z}$ ，选定 $\mathcal{Z}$ 的一组基 B_3 ，记 $[U]_{B_3}^{B_2} = B$ 。对任意 $v \in \mathcal{V}$ ，先应用 T ，再应用 U ，其坐标满足 $[U(Tv)]_{B_3} = B[Tv]_{B_2} = BA[v]_{B_1}$ 。因此 $[U \circ T]_{B_3}^{B_1} = BA$ 。

这里两个矩阵对中间空间 $\mathcal{W}$ 必须使用同一组基 B_2 ；先做 T 、后做 U ，对应的乘积顺序是 BA 。

Tip

本页要点： 线性变换在选定基下由矩阵表示，换基只改变坐标而不改变变换本身。

2.7 Subspaces and Direct Sum

非空集合 $\mathcal{S} \subseteq \mathcal{V}$ 是子空间，当且仅当它对线性组合封闭：

$$\mathbf{x}, \mathbf{y} \in \mathcal{S}, a, b \in \mathbb{F} \implies a\mathbf{x} + b\mathbf{y} \in \mathcal{S}. \quad (2.12)$$

两个子空间之和定义为

$$\mathcal{S}_1 + \mathcal{S}_2 = \{\mathbf{x}_1 + \mathbf{x}_2 : \mathbf{x}_i \in \mathcal{S}_i\}. \quad (2.13)$$

若每个和分解唯一，则记为直和 $\mathcal{S}_1 \oplus \mathcal{S}_2$ 。以下条件等价：

$$\mathcal{S}_1 + \mathcal{S}_2 = \mathcal{S}_1 \oplus \mathcal{S}_2, \quad \mathcal{S}_1 \cap \mathcal{S}_2 = \{\mathbf{0}\}, \quad (2.14)$$

以及

$$\dim(\mathcal{S}_1 + \mathcal{S}_2) = \dim \mathcal{S}_1 + \dim \mathcal{S}_2. \quad (2.15)$$

这些条件也等价于：若 $\mathbf{x}_i \in \mathcal{S}_i$ 且 $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{0}$ ，则 $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{0}$ 。若 $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{y}_1 + \mathbf{y}_2$ ，则 $\mathbf{x}_1 - \mathbf{y}_1 = \mathbf{y}_2 - \mathbf{x}_2 \in \mathcal{S}_1 \cap \mathcal{S}_2$ ，交只有零元时便有 $\mathbf{x}_i = \mathbf{y}_i$ ，即分解唯一。任意子空间在有限维空间中都可补成某个直和分解 $\mathcal{S} \oplus \mathcal{S}_c = \mathcal{V}$ ，但补空间一般不唯一。

④ Note

维数公式的推导。 记 $r = \dim \mathcal{S}_1$ 、 $s = \dim \mathcal{S}_2$ 。取 $\mathcal{S}_1 \cap \mathcal{S}_2$ 的基 $\mathbf{c}_1, \dots, \mathbf{c}_q$ ，分别添上 $\mathbf{u}_1, \dots, \mathbf{u}_{r-q}$ 和 $\mathbf{v}_1, \dots, \mathbf{v}_{s-q}$ ，补成两个子空间的基。这三组向量合起来张成 $\mathcal{S}_1 + \mathcal{S}_2$ 。

若 $\sum_i a_i \mathbf{c}_i + \sum_j b_j \mathbf{u}_j + \sum_k d_k \mathbf{v}_k = \mathbf{0}$ ，移项得到 $\sum_k d_k \mathbf{v}_k = -\sum_i a_i \mathbf{c}_i - \sum_j b_j \mathbf{u}_j \in \mathcal{S}_1 \cap \mathcal{S}_2$ 。因此存在系数 e_i 使 $\sum_k d_k \mathbf{v}_k = \sum_i e_i \mathbf{c}_i$ 。两边都是 $\mathcal{S}_2$ 的基 $\{\mathbf{c}_i, \mathbf{v}_k\}$ 的线性组合，线性无关性给出 $d_k = 0$ 和 $e_i = 0$ 。代回原式，再由 $\mathcal{S}_1$ 的基 $\{\mathbf{c}_i, \mathbf{u}_j\}$ 线性无关，得到 $a_i = b_j = 0$ 。所以合并组是 $\mathcal{S}_1 + \mathcal{S}_2$ 的基， $\dim(\mathcal{S}_1 + \mathcal{S}_2) = r + s - q$ 。交集为 $\{\mathbf{0}\}$ 时即得式 (2.15)；反之，式 (2.15) 也迫使 $q = 0$ 。

💡 Tip

本页要点： 直和要求分解存在且唯一，交集只有零向量是唯一性的条件。

2.8 Fundamental Subspaces

向量组的张成空间为

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\} = \left\{ \sum_{i=1}^m a_i \mathbf{v}_i : a_i \in \mathbb{F} \right\}. \quad (2.16)$$

对 $A \in \mathbb{R}^{m \times n}$ ，四个基本子空间是

$$\begin{aligned} \text{range}(A) &= \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m, \\ \text{null}(A) &= \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n, \\ \text{range}(A^T) &\subseteq \mathbb{R}^n, \\ \text{null}(A^T) &\subseteq \mathbb{R}^m. \end{aligned} \quad (2.17)$$

四个基本子空间满足正交关系

$$\text{range}(A)^\perp = \text{null}(A^T), \quad \text{range}(A^T)^\perp = \text{null}(A). \quad (2.18)$$

秩定义为 $\text{rank}(A) = \dim \text{range}(A)$ 。秩-零化度定理给出

$$\text{rank}(A) + \dim \text{null}(A) = n. \quad (2.19)$$

Note

正交关系和秩-零化度从哪里来? 对任意 $\mathbf{z}$, $\mathbf{z} \in \text{range}(A)^\perp$ 当且仅当 $\mathbf{z}^T A\mathbf{x} = 0$ 对所有 $\mathbf{x}$ 成立, 也就等价于 $A^T \mathbf{z} = \mathbf{0}$ 。因此 $\text{range}(A)^\perp = \text{null}(A^T)$, 另一条关系同理。

取 $\text{null}(A)$ 的基 $\mathbf{z}_1, \dots, \mathbf{z}_k$, 再添 $\mathbf{w}_1, \dots, \mathbf{w}_{n-k}$ 补成定义域基。任意 $\mathbf{x}$ 都能写成 $\sum_i \alpha_i \mathbf{z}_i + \sum_j \beta_j \mathbf{w}_j$, 因此 $A\mathbf{x} = \sum_j \beta_j A\mathbf{w}_j$, 说明这些像向量张成 $\text{range}(A)$ 。若 $\sum_j c_j A\mathbf{w}_j = \mathbf{0}$, 则 $\sum_j c_j \mathbf{w}_j \in \text{null}(A)$, 可写成 $\sum_i d_i \mathbf{z}_i$ 。于是 $\sum_j c_j \mathbf{w}_j - \sum_i d_i \mathbf{z}_i = \mathbf{0}$ 是补全基的线性关系, 故所有 c_j, d_i 均为零。于是 $A\mathbf{w}_j$ 线性无关并构成值域的基, $\text{rank}(A) = n - k$, 得到式 (2.19)。

Tip

本页要点: 四个基本子空间把矩阵的行、列、零空间及其正交补联系起来。

2.9 Range and Null Spaces

首先证明

$$\text{null}(A) = \text{null}(A^T A). \quad (2.20)$$

若 $A\mathbf{x} = \mathbf{0}$, 左乘 A^T 即得 $A^T A\mathbf{x} = \mathbf{0}$ 。反之, 若 $A^T A\mathbf{x} = \mathbf{0}$, 则

$$0 = \mathbf{x}^T A^T A\mathbf{x} = \|A\mathbf{x}\|_2^2, \quad (2.21)$$

所以 $A\mathbf{x} = \mathbf{0}$ 。复数情形把 T 换成 H 。

再证明

$$\text{range}(A) = \text{range}(AA^T). \quad (2.22)$$

由第 2.8 页式 (2.18)、本页式 (2.20) (将 A 换为 A^T) 及 AA^T 的对称性, 得到

$$\text{range}(AA^T)^\perp = \text{range}(A)^\perp. \quad (2.23)$$

有限维内积空间中 $(\mathcal{S}^\perp)^\perp = \mathcal{S}$, 所以正交补相同的两个子空间也相同, 即得式 (2.22)。

在式 (2.22) 中将 A 换为 A^T , 便有

$$\text{range}(A^T) = \text{range}(A^T A). \quad (2.24)$$

这些关系也可由后续的 SVD 直接读出。

Tip

本页要点: 处理矩阵方程时, 要同时看右端是否属于值域, 以及零空间中有哪些自由度。

2.10 Rank of A Matrix

行秩等于列秩，因此可统一记为 $\text{rank}(A)$ 。基本性质包括

$$\text{rank}(A) \leq \min(m, n), \quad \text{rank}(A^T) = \text{rank}(A), \quad (2.25)$$

$$\text{rank}(A^T A) = \text{rank}(A A^T) = \text{rank}(A), \quad \text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B). \quad (2.26)$$

第一条来自列向量至多有 m 个线性无关、总数又为 n ；第三条来自第 2.9 页的零空间、列空间关系及秩-零化度定理；最后一条来自

$$\text{range}(A + B) \subseteq \text{range}(A) + \text{range}(B). \quad (2.27)$$

若 X, Y 非奇异，则左乘或右乘只改变基，不改变秩：

$$\text{rank}(XAY) = \text{rank}(A). \quad (2.28)$$

若 $r = \text{rank}(A)$ ，从 A 的 n 列中选出 r 个线性无关的列，组成 $X \in \mathbb{R}^{m \times r}$ 。它们构成 $\text{range}(A)$ 的一组基，所以未选入 X 的其余列也都能唯一写成这些列的线性组合。对 A 的每一列写 $\mathbf{a}_j = X\mathbf{y}_j$ ，按原列顺序组成 $Y = [\mathbf{y}_1, \dots, \mathbf{y}_n] \in \mathbb{R}^{r \times n}$ 。若 $\mathbf{a}_j$ 是所选的第 ℓ 列，其坐标 $\mathbf{y}_j$ 就是第 ℓ 个标准基向量； Y 因而包含所有列的坐标。由此得到满秩分解

$$A = XY, \quad \text{rank}(X) = \text{rank}(Y) = r. \quad (2.29)$$

其中 $\text{rank}(X) = r$ 来自所选列的线性无关性； Y 在所选列对应的位置包含 r 个标准基向量，因此 $\text{rank}(Y) = r$ 。

秩还决定矩阵在左右换基下的标准形：存在非奇异方阵 $S \in \mathbb{R}^{m \times m}$ 、 $T \in \mathbb{R}^{n \times n}$ ，使

$$A = SJT^{-1}, \quad J = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}. \quad (2.30)$$

先取值域的一组基 $\mathbf{u}_1, \dots, \mathbf{u}_r$ ，分别选原像 $\mathbf{v}_i$ 使 $A\mathbf{v}_i = \mathbf{u}_i$ ，再用零空间的一组基 $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ 补齐。这些向量线性无关：对它们的零线性组合左乘 A ，由 $\mathbf{u}_i$ 独立得前 r 个系数为零，再由零空间基独立得其余系数为零。因此 $T = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ 非奇异。把 $\mathbf{u}_1, \dots, \mathbf{u}_r$ 补成 $\mathbb{R}^m$ 的基，组成 S ，便有 $AT = SJ$ ，即上式。 $r = 0$ 时 $A = 0$ ，标准形也成立。

同尺寸矩阵若可通过左右乘非奇异矩阵相互转化，就称为等价矩阵。等价变换保持秩；反过来，同秩矩阵都可化为同一个 J ，所以彼此等价。这里左右换基可以独立选择，不同于要求同一方阵换基的相似变换。

Note

行秩与列秩、可逆变换与秩。 由第 2.8 页的 $\text{range}(A^T)^\perp = \text{null}(A)$ ，有 $\dim \text{range}(A^T) = n - \dim \text{null}(A)$ ；再用式 (2.19) 得 $\text{rank}(A^T) = \text{rank}(A)$ 。若 X, Y 可逆，右乘 Y 不改变值域，左乘 X 作可逆线性变换，均不改变值域维数，因此式 (2.28) 成立。

Tip

本页要点： 秩同时刻画列空间维数、行空间维数和线性方程组中独立信息的数量。

2.11 Eigenvalues and Eigenvectors

对 $A \in \mathbb{C}^{n \times n}$, 右特征对满足

$$Au = \lambda u, \quad u \neq 0; \quad (2.31)$$

左特征向量满足 $v^H A = \lambda v^H$ 、 $v \neq 0$ 。谱和谱半径定义为

$$\lambda(A) = \{\lambda : \det(\lambda I - A) = 0\}, \quad \rho(A) = \max_{\lambda \in \lambda(A)} |\lambda|. \quad (2.32)$$

因为 $\det(\lambda I - A) = \det((\lambda I - A)^T) = \det(\lambda I - A^T)$, A 与 A^T 有相同的特征多项式, 因此特征值及其代数重数相同。

特征多项式

$$p_A(t) = \det(tI - A) = \prod_{i=1}^r (t - \lambda_i)^{n(\lambda_i)}, \quad \sum_{i=1}^r n(\lambda_i) = n, \quad (2.33)$$

其中 $\lambda_1, \dots, \lambda_r$ 是互异特征值, 根 λ 在多项式中的重数 $n(\lambda)$ 称为代数重数。特征空间

$$\mathcal{E}_\lambda = \text{null}(A - \lambda I) \quad (2.34)$$

的维数是几何重数 $m(\lambda)$, 并且

$$1 \leq m(\lambda) \leq n(\lambda) \leq n. \quad (2.35)$$

④ Note

为什么几何重数不超过代数重数? 设 $g = m(\lambda)$, 取特征空间的一组基并补成 $\mathbb{C}^n$ 的基, 组成矩阵 S 。前 g 个基向量均满足 $Av_i = \lambda v_i$, 所以 $S^{-1}AS = \begin{bmatrix} \lambda I_g & C \\ 0 & B \end{bmatrix}$ 。分块三角矩阵的行列式是对角块行列式的乘积, 故 $p_A(t) = (t - \lambda)^g \det(tI - B)$, 于是 $n(\lambda) \geq g = m(\lambda)$ 。每个特征值都有非零特征向量, 且 p_A 为 n 次多项式, 故 $1 \leq m(\lambda) \leq n(\lambda) \leq n$ 。秩-零化度定理还给出 $m(\lambda) = n - \text{rank}(A - \lambda I)$ 。

④ Note

从 Jordan 块看两种重数。 对同一特征值 λ , 若它对应的 Jordan 块共有 k 个, 大小分别为 $s_1, \dots, s_k$, 则代数重数 $n(\lambda) = s_1 + \dots + s_k$, 几何重数 $m(\lambda) = k$ 。前者是这些块的总阶数, 后者是块数, 因为每块只贡献一个独立特征方向。例如, 一个 3 阶块和一个 2 阶块对应同一特征值时, 代数重数为 5, 几何重数为 2。每块至少有一阶, 所以 $m(\lambda) \leq n(\lambda)$; 等号成立当且仅当这些块都是一阶。

💡 Tip

本页要点: 特征值方程寻找在线性变换下只改变尺度的特殊方向。

2.12 Eigenvalues of Some Matrices

常见矩阵的谱可由定义直接判断。

- 零矩阵只有特征值 0; 单位矩阵只有特征值 1。
- 标量矩阵 zI 的特征值全为 z , 因为 $zIx = zx$ 对任意非零向量成立。
- 投影矩阵满足 $P^2 = P$ 。对特征向量作用两次得 $\lambda^2 x = \lambda x$, 故特征值只能为 0 或 1。

- 三角矩阵的特征值是对角元，因为 $\det(\lambda I - A)$ 是对角因子的乘积。
- 幂零矩阵满足 $A^k = 0$ ，若 $Ax = \lambda x$ ，则 $0 = A^k x = \lambda^k x$ ，故 $\lambda = 0$ 。
- 对合矩阵满足 $A^2 = I$ ，故特征值满足 $\lambda^2 = 1$ ，即 $\lambda = \pm 1$ 。
- Hermitian 矩阵特征值为实数；实反对称矩阵的非零特征值为纯虚共轭对。
- 对称正定矩阵的特征值严格为正：由 $Ax = \lambda x$ 得 $\lambda = x^T Ax / \|x\|_2^2 > 0$ 。
- 正交/酉矩阵满足 $\|Ax\|_2 = \|x\|_2$ ，故特征值模为 1。
- 对列随机 Markov 矩阵， $A \geq 0$ 且 $\mathbf{1}^T A = \mathbf{1}^T$ ，所以 1 是左特征值，也是 A 的特征值。各列和为 1 给出 $\|A\|_1 = 1$ ；由第 2.26 页的 $\rho(A) \leq \|A\|_1$ 可知，其余特征值的模也不超过 1。
- 奇异矩阵含特征值 0；可逆矩阵不含 0。

可以用这些结论快速检查计算结果。例如，若计算出的 Hermitian 特征值的虚部超出预期舍入误差范围，应检查输入、算法或实现是否符合假设。

④ Note

几个结论的短推导。 若 $Ax = \lambda x$ ，则 Hermitian 情形中 $\lambda \|x\|_2^2 = x^H Ax \in \mathbb{R}$ ，故 λ 为实数；实反对称情形 $A^H = -A$ ，同一个标量同时等于自己的相反共轭，因此 $\lambda + \bar{\lambda} = 0$ 。由于实矩阵的特征多项式系数为实数，任何非实特征值都与其共轭成对出现。若 A 正交或酉，则 $\|Ax\|_2 = \|x\|_2$ ，而左边等于 $|\lambda| \|x\|_2$ ，所以 $|\lambda| = 1$ 。若 A 对称正定且 $y \neq 0$ ，令 $x = A^{-1}y$ ，则 $y^T A^{-1}y = x^T Ax > 0$ ，故 A^{-1} 也正定。

🔗 Tip

本页要点： 矩阵结构会限制特征值的范围，可用这些限制快速检查计算结果。

2.13 Geometric Meaning

对称矩阵例子

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad (2.36)$$

直接相乘可得 $A(1, 1)^T = 3(1, 1)^T$ 、 $A(1, -1)^T = (1, -1)^T$ 。对应特征向量 $v_1 = (1, 1)^T$ 、 $v_2 = (1, -1)^T$ 的长度均为 $\sqrt{2}$ ，除以该长度得到单位长度的特征向量：

$$\lambda_1 = 3, \quad u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \lambda_2 = 1, \quad u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (2.37)$$

任意单位向量写为 $x = c_1 u_1 + c_2 u_2$ ，其中 $c_1^2 + c_2^2 = 1$ 。于是

$$Ax = 3c_1 u_1 + c_2 u_2. \quad (2.38)$$

因此单位圆经 A 映射为沿 u_1, u_2 方向半轴长分别为 3 和 1 的椭圆。特征方向不发生转向，只按特征值缩放。

Note

这个例子为什么能画成“沿特征方向纯缩放”？这里的 A 是实对称矩阵，可以写成 $A = Q\Lambda Q^T$ ，其特征向量构成标准正交基。因此，单位圆沿特征向量方向的两个半轴分别只被缩放 $|\lambda_i|$ 倍；这种描述依赖于实对称性。

对一般矩阵，特征向量即使存在，也可能不正交，甚至不能组成一组基；这时特征对仍表示某些不变方向，但不能把整个变换理解为沿正交主轴的纯缩放。要描述一般实矩阵或复矩阵如何把单位球变成椭球，应使用 SVD：右奇异向量给出输入主方向，左奇异向量给出输出主方向，奇异值给出相应的伸缩倍数。

对称矩阵和更一般的正规矩阵是特殊情形：左右奇异向量可以取自同一组正交特征向量（至多相差符号或复相位），并且 $\sigma_i = |\lambda_i|$ ；只有半正定时才可同时取 $U = V$ 。

Tip

本页要点： 单位圆经对称矩阵映射为椭圆，其主轴方向为特征方向，半轴长为特征值的绝对值。

2.14 Geometric Meaning: Alternative Perspective

若 A 对称正定，存在标准正交特征向量 $\mathbf{u}_i$ ，且 $\lambda_i > 0$ 。写

$$\mathbf{x} = \sum_{i=1}^n c_i \mathbf{u}_i, \quad (2.39)$$

则正交性给出

$$\mathbf{x}^T A \mathbf{x} = \sum_{i=1}^n \lambda_i c_i^2. \quad (2.40)$$

因此二次曲面 $\mathbf{x}^T A \mathbf{x} = 1$ 在特征坐标中为

$$\sum_{i=1}^n \frac{c_i^2}{(1/\sqrt{\lambda_i})^2} = 1, \quad (2.41)$$

其第 i 个半轴长度为 $1/\sqrt{\lambda_i}$ 。大特征值对应短半轴，小特征值对应长半轴。

这里的 $\mathbf{u}_i$ 是欧氏单位特征向量；椭球的半轴向量则按 A -范数归一化，对应 $\mathbf{u}_i/\sqrt{\lambda_i}$ 。它的 A -范数平方为 $\mathbf{u}_i^T A \mathbf{u}_i/\lambda_i = 1$ ，欧氏长度为 $1/\sqrt{\lambda_i}$ ，恰好等于该半轴长。

若令 $\mathbf{z} = A^{1/2} \mathbf{x}$ ，则 $\|\mathbf{z}\|_2^2 = \mathbf{x}^T A \mathbf{x}$ 。反之，令 $\mathbf{x} = A^{-1/2} \mathbf{z}$ ，单位圆被映射为上述椭圆。这给出正定二次型与加权内积的几何意义。

本页仍用 $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ ，所以等值线 $\mathbf{x}^T A \mathbf{x} = 1$ 是半轴长为 $1/\sqrt{3}$ 和 1 的椭圆，与上一页单位圆经 A 映射所得的椭圆不同。由 $A^T = A$ 得 $\nabla(\mathbf{x}^T A \mathbf{x} - 1) = 2A\mathbf{x}$ ，因此 $\mathbf{n} = A\mathbf{x}$ 是未归一化的法向量。若 $\mathbf{x} = \alpha \mathbf{n}$ ，则 $A\mathbf{x} = \alpha^{-1} \mathbf{x}$ ，径向方向也就是特征方向；对本例的椭圆，这恰好发生在主轴端点。

Tip

本页要点： 正定二次型等值面的半轴长为特征值平方根的倒数，不要与单位圆经矩阵映射的像混淆。

2.15 Eigenspace

特征值 λ 的特征空间为

$$\mathcal{E}_\lambda = \ker(A - \lambda I). \quad (2.42)$$

它是子空间，其维数就是几何重数 $m(\lambda)$ 。更一般地，子空间 $\mathcal{S} \subseteq \mathbb{C}^n$ 对 A 不变，是指

$$A\mathcal{S} \subseteq \mathcal{S}. \quad (2.43)$$

特征空间和 $\text{range}(A)$ 都是不变子空间；后一结论由 $A(A\mathbf{x}) \in \text{range}(A)$ 直接得到。

若 $X \in \mathbb{C}^{n \times k}$ 的列构成 $\mathcal{S}$ 的一组基，则不变性等价于存在 $B \in \mathbb{C}^{k \times k}$ 使

$$AX = XB. \quad (2.44)$$

原因是每一列 $A\mathbf{x}_j$ 都必须能由 X 的列线性表示，其系数正好组成 B 。若 X 方阵且非奇异，则

$$A = XBX^{-1}, \quad (2.45)$$

即 A 与 B 相似，因而 $\lambda(A) = \lambda(B)$ 。若 X 仅列满秩，则 B 是 A 限制在不变子空间上的矩阵表示。

此时 B 的特征值也是 A 的特征值：若 $B\mathbf{y} = \lambda\mathbf{y}$ 、 $\mathbf{y} \neq 0$ ，则 $A(X\mathbf{y}) = X(B\mathbf{y}) = \lambda X\mathbf{y}$ 。由于 X 列满秩， $X\mathbf{y} \neq 0$ ，故它就是 A 对应的特征向量。

Tip

本页要点： 特征空间包含同一特征值对应的全部方向，讨论重根和不变子空间时要注意它的维数。

2.16 Commonly Used Matrices

常见结构及其定义如下。

- 方阵：行列数相同；可逆矩阵：存在 A^{-1} 。
- 奇异矩阵： $\det A = 0$ ，等价于秩不足。
- 对角矩阵、三角矩阵：非零位置受限，便于求解和读出特征值。
- 可对角化矩阵：存在非奇异矩阵 X ，使 $X^{-1}AX$ 为对角矩阵，即可以选出一组特征向量作为全空间的基。
- 正规矩阵 (normal matrix)： $A^H A = A A^H$ ；Hermitian/实对称矩阵是其子类。
- 正交矩阵：实方阵满足 $Q^T Q = I$ ；酉矩阵：复方阵满足 $Q^H Q = I$ 。两者都保持二范数。
- 正半定/正定矩阵：Hermitian，且对所有非零 $\mathbf{x}$ ，分别满足 $\mathbf{x}^H A \mathbf{x} \geq 0$ 或 > 0 。
- 投影矩阵： $P^2 = P$ ；若再有 $P^H = P$ ，则为正交投影。
- 置换矩阵：每行每列恰有一个 1，是正交矩阵。

矩阵可能同时属于多个集合，例如实对称正定矩阵也是 Hermitian 矩阵、正规矩阵和可逆矩阵。利用这些结构，可以选择计算量更小、稳定性更好的算法。

💡 Tip

本页要点： 先识别矩阵的结构及其包含关系，才能选择合适的分解和求解方法。

2.17 Example 1: Rotation Matrices

逆时针旋转角 θ 的矩阵为

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}. \quad (2.46)$$

若 $\mathbf{x} = (x, y)^T = r(\cos \phi, \sin \phi)^T$ ，利用和角公式可得

$$R_\theta \mathbf{x} = r \begin{bmatrix} \cos(\phi + \theta) \\ \sin(\phi + \theta) \end{bmatrix}, \quad (2.47)$$

所以该矩阵确实把方向角增加 θ 。两列标准正交，因此

$$R_\theta^T R_\theta = I, \quad \det R_\theta = 1. \quad (2.48)$$

由正交性， $R_\theta^{-1} = R_\theta^T = R_{-\theta}$ ，并保持距离与夹角。直接相乘并使用和角公式还得到

$$R_\theta R_\phi = R_{\theta+\phi}. \quad (2.49)$$

💡 Tip

本页要点： 旋转矩阵把几何旋转写成线性变换，正交性保证长度不变。

2.18 Example 2: Pauli Matrices

Pauli 矩阵由 Wolfgang Pauli 于 1927 年引入，用于描述自旋 1/2 态的旋转及沿三个空间轴的自旋测量。三个矩阵为

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (2.50)$$

它们都是 Hermitian 且满足 $\sigma_j^2 = I$ ，故特征值只能是 ± 1 。以

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (2.51)$$

为例， $\sigma_x |\rightarrow\rangle = |\rightarrow\rangle$ ，所以它是 σ_x 的 +1 特征态；但

$$\sigma_z |\rightarrow\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |\leftarrow\rangle \quad (2.52)$$

不与 $|\rightarrow\rangle$ 共线，说明原态 $|\rightarrow\rangle$ 不是 σ_z 的特征态。这个矩阵乘积不表示测量后的电子态。沿 z 方向测量时，原态在 $|+z\rangle = (1, 0)^T$ 、 $|-z\rangle = (0, 1)^T$ 两个本征态上的系数模平方均为 1/2，所以测得 ± 1 并得到对应本征态的概率各为 1/2。

这个例子把特征向量的代数定义与可观测量、基变换联系起来。

💡 Tip

本页要点： Pauli 矩阵展示了复数矩阵、特征态和量子计算中的基本结构。

2.19 Symmetric Positive Definiteness

实对称矩阵 $A \in \mathbb{R}^{n \times n}$ 称为正定矩阵，如果

$$\mathbf{x}^T A \mathbf{x} > 0, \quad \forall \mathbf{x} \neq \mathbf{0}. \quad (2.53)$$

正定矩阵必非奇异：若 $A\mathbf{x} = \mathbf{0}$ 且 $\mathbf{x} \neq \mathbf{0}$ ，则 $\mathbf{x}^T A \mathbf{x} = 0$ ，矛盾。若 A 正定，则 A^{-1} 也正定；若 X 列满秩，则对任意非零 $\mathbf{y}$ ， $X\mathbf{y} \neq \mathbf{0}$ ，从而

$$\mathbf{y}^T X^T A X \mathbf{y} = (X\mathbf{y})^T A (X\mathbf{y}) > 0, \quad (2.54)$$

所以 $X^T A X$ 正定。

📌 Note

A^{-1} 与 $X^T A X$ 的对称性由转置规则直接验证。令 $\mathbf{x} = A^{-1}\mathbf{y} \neq \mathbf{0}$ ，则 $\mathbf{y}^T A^{-1}\mathbf{y} = \mathbf{x}^T A \mathbf{x} > 0$ ，故逆矩阵也正定。另外，取 $\mathbf{x} = \mathbf{e}_i$ 得 $a_{ii} > 0$ ，所以正定矩阵的所有对角元都为正。

Sylvester 判据

判据。 设 $A_k = A(1:k, 1:k)$ 为 A 的前 $k \times k$ 主子矩阵。实对称矩阵 A 正定当且仅当

$$D_k := \det(A_k) > 0, \quad k = 1, \dots, n. \quad (2.55)$$

相应地， A 负定当且仅当

$$(-1)^k D_k > 0, \quad k = 1, \dots, n. \quad (2.56)$$

证明。 若 A 正定，则任意非零 $\mathbf{y} \in \mathbb{R}^k$ 都满足

$$\mathbf{y}^T A_k \mathbf{y} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}^T A \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} > 0. \quad (2.57)$$

因此每个 A_k 都正定，其特征值全为正，故 $D_k > 0$ 。

反之，设所有 $D_k > 0$ ，并令 $D_0 = 1$ 。对阶数归纳构造无主元 LDL^T 分解。 $n = 1$ 时由 $A = [D_1]$ 、 $D_1 > 0$ 立即成立。若 $n \geq 2$ ，分块写成

$$A = \begin{bmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{w} & B \end{bmatrix}. \quad (2.58)$$

因为 $a_{11} = D_1 \neq 0$ ，有对称消元恒等式

$$A = \begin{bmatrix} 1 & 0 \\ \mathbf{w}/a_{11} & I \end{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} 1 & \mathbf{w}^T/a_{11} \\ 0 & I \end{bmatrix}, \quad S = B - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}}. \quad (2.59)$$

若 S_j 是 S 的前 $j \times j$ 主子矩阵, 则同一分块消元给出

$$\det S_j = \frac{D_{j+1}}{D_1} > 0, \quad j = 1, \dots, n-1. \quad (2.60)$$

归纳假设应用于 S , 并嵌入上式, 得到

$$A = LDL^T, \quad D = \text{diag}(d_1, \dots, d_n), \quad d_k = \frac{D_k}{D_{k-1}} > 0. \quad (2.61)$$

这里 D 是对角矩阵, D_k 是顺序主子式; L 为非奇异单位下三角矩阵, 而 $D_k = d_1 \cdots d_k$ 给出最后一个比值公式。于是对任意 $\mathbf{x} \neq 0$,

$$\mathbf{x}^T A \mathbf{x} = (L^T \mathbf{x})^T D (L^T \mathbf{x}) > 0. \quad (2.62)$$

故 A 正定。最后, A 负定等价于 $-A$ 正定, 而 $\det(-A_k) = (-1)^k D_k$; 对 $-A$ 应用已证结论即得负定判据。证毕。

Tip

本页要点: 正定性表示二次型具有正能量, 也是许多数值算法保证稳定性所需的条件。

2.20 Theorem 2.1: Gershgorin Circle Theorem

定理 2.1. 设 $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, 定义行 Gershgorin 圆盘

$$G_i(A) = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}. \quad (2.63)$$

则

$$\lambda(A) \subseteq \bigcup_{i=1}^n G_i(A). \quad (2.64)$$

若其中 m 个圆盘组成一个与其余圆盘不相交的连通区域 G , 则 G 中恰含 m 个特征值, 按代数重数计。

证明. 令 $(\lambda, \mathbf{x})$ 为特征对, 取 k 使 $|x_k| = \|\mathbf{x}\|_\infty$ 。由于 $\mathbf{x} \neq 0$, 有 $x_k \neq 0$ 。第 k 行方程为

$$(\lambda - a_{kk})x_k = \sum_{j \neq k} a_{kj}x_j. \quad (2.65)$$

两边取绝对值并除以 $|x_k|$:

$$|\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j \neq k} |a_{kj}|. \quad (2.66)$$

所以 $\lambda \in G_k(A)$, 第一部分成立。

为证明计数结论, 令

$$A(t) = D + t(A - D), \quad D = \text{diag}(a_{11}, \dots, a_{nn}), \quad t \in [0, 1]. \quad (2.67)$$

第 i 个圆盘为

$$G_i(t) = \left\{ z : |z - a_{ii}| \leq t \sum_{j \neq i} |a_{ij}| \right\} \subseteq G_i(1). \quad (2.68)$$

设 I 是组成 G 的那 m 个圆盘的指标集。两个圆盘并集是互不相交的紧集，故可选区域的有向边界 Γ ，使 $i \in I$ 的圆盘位于区域内，其余圆盘位于区域外。 Γ 可由多条闭曲线组成，外边界取正向，孔洞的内边界取负向。对每个 t ，特征值都在相应圆盘内，所以没有特征值落在 Γ 上。

利用复分析中的辐角原理，

$$N(t) = \frac{1}{2\pi i} \int_{\Gamma} \frac{\partial_z \det(zI - A(t))}{\det(zI - A(t))} dz \quad (2.69)$$

等于区域内的特征值个数，按代数重数计。被积函数连续依赖于 t ，而 $N(t)$ 取整数值，故它在连通区间 $[0, 1]$ 上恒定。 $t=0$ 时 $A(0) = D$ ，区域内恰有 $i \in I$ 的 m 个对角元，所以 $N(1) = N(0) = m$ 。证毕。

若 A 严格行对角占优，即

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad (2.70)$$

则所有圆盘都不含原点，因此 $0 \notin \lambda(A)$ ，矩阵非奇异。

由于 A^T 与 A 的谱相同，对 A^T 应用定理还可得到列圆盘 $\{z : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ji}|\}$ 。这些定位结果也引出后续问题：矩阵元素的小扰动会使特征值移动多少？

Tip

本页要点： Gershgorin 圆盘给出特征值的包围区域，而不是每个特征值的精确位置。

2.21 Sherman - Morrison - Woodbury Formula

设 A, B 为同阶非奇异矩阵，**预解式恒等式 (resolvent identity)** 为

$$B^{-1} = A^{-1} - B^{-1}(B - A)A^{-1}. \quad (2.71)$$

证明。 展开 $B^{-1}(B - A)A^{-1} = A^{-1} - B^{-1}$ ，移项即得。它把逆矩阵之差与原矩阵之差联系起来；乘法次序须保持不变。

Sherman 与 Morrison 于 1950 年给出秩一更新公式，Woodbury 同年给出低秩更新的推广。

设 $A \in \mathbb{R}^{n \times n}$ 非奇异， $U, V \in \mathbb{R}^{n \times k}$ ，并假设 $I + V^T A^{-1} U$ 非奇异。低秩更新的逆为

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1}U)^{-1}V^T A^{-1}. \quad (2.72)$$

令右端为 X ，记 $M = I + V^T A^{-1}U$ 。直接相乘：

$$\begin{aligned} (A + UV^T)X &= I - UM^{-1}V^T A^{-1} + UV^T A^{-1} - UV^T A^{-1}UM^{-1}V^T A^{-1} \\ &= I + U[I - M^{-1} - (M - I)M^{-1}]V^T A^{-1} \\ &= I. \end{aligned} \quad (2.73)$$

方阵存在右逆即非奇异，且该右逆就是逆矩阵，公式得证。**意义在于：只需已有的 A^{-1} 作用和一个 $k \times k$ 小矩阵求逆，而不必重新分解 $n \times n$ 矩阵。**

当 $k = 1$ 、 $U = \mathbf{u}$ 、 $V = \mathbf{v}$ 时，得到 Sherman—Morrison 公式

$$(A + \mathbf{u}\mathbf{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^T A^{-1}}{1 + \mathbf{v}^T A^{-1}\mathbf{u}}. \quad (2.74)$$

分母非零是必要条件；若它为零，则矩阵行列式引理给出 $\det(A + \mathbf{u}\mathbf{v}^T) = 0$ 。

💡 Tip

本页要点： 低秩更新可归结为求解小规模线性系统，关键前提是其系数矩阵可逆。

2.22 Theorem 2.2: Cayley – Hamilton Theorem

定理 2.2. 设

$$p(\lambda) = \det(\lambda I - A) \quad (2.75)$$

是 $A \in \mathbb{C}^{n \times n}$ 的特征多项式，则

$$p(A) = 0. \quad (2.76)$$

若 A 可逆，则 A^{-1} 可表示为 A 的次数不超过 $n - 1$ 的多项式。

证明. 写

$$p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 \quad (2.77)$$

以及

$$\text{adj}(\lambda I - A) = \lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + B_{n-1}. \quad (2.78)$$

伴随矩阵恒等式为

$$(\lambda I - A) \text{adj}(\lambda I - A) = p(\lambda)I. \quad (2.79)$$

比较 $\lambda^n, \lambda^{n-1}, \dots, 1$ 的系数，依次得到

$$\begin{aligned} B_0 &= I, \\ B_j - AB_{j-1} &= c_{n-j}I, \quad j = 1, \dots, n-1, \\ -AB_{n-1} &= c_0I. \end{aligned} \quad (2.80)$$

递推关系给出

$$B_j = A^j + c_{n-1}A^{j-1} + \cdots + c_{n-j}I. \quad (2.81)$$

取 $j = n - 1$ 并代入最后一个系数关系，得到

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I = 0, \quad (2.82)$$

即 $p(A) = 0$ 。这里通过比较标量多项式系数完成证明，避免了把矩阵直接代入含非交换因子的恒等式。

若 A 可逆, 则 $c_0 = (-1)^n \det(A) \neq 0$ 。将上式右乘 A^{-1} 并解出逆矩阵:

$$A^{-1} = -\frac{1}{c_0} (A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1I). \quad (2.83)$$

证毕。

Tip

本页要点: Cayley–Hamilton 定理把高次矩阵幂化为次数不超过 $n-1$ 的矩阵多项式。

第2.2节 距离与范数

2.24 Vector Norms

向量范数 $\|\cdot\|$ 满足正定性、三角不等式和绝对齐次性:

$$\|\mathbf{x}\| \geq 0, \quad \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}, \quad (2.84)$$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \quad \|a\mathbf{x}\| = |a| \|\mathbf{x}\|. \quad (2.85)$$

常用 p -范数为

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty, \quad \|\mathbf{x}\|_\infty = \max_i |x_i|. \quad (2.86)$$

有限维空间中所有范数等价。例如

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_2 \leq n \|\mathbf{x}\|_\infty. \quad (2.87)$$

前两项分别由 $|x_i|^2 \leq \sum_j |x_j|^2$ 和 $\sum_i |x_i|^2 \leq (\sum_i |x_i|)^2$ 得到; 第三项用 Cauchy–Schwarz, 最后一项用 $|x_i| \leq \|\mathbf{x}\|_\infty$ 。

若 $1 \leq p, q \leq \infty$ 且 $1/p + 1/q = 1$, Hölder 不等式为

$$|\mathbf{x}^H \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q. \quad (2.88)$$

当 $1 < p, q < \infty$ 且两范数非零时, 令 $a_i = |x_i|/\|\mathbf{x}\|_p$ 、 $b_i = |y_i|/\|\mathbf{y}\|_q$ 。Young 不等式 $ab \leq a^p/p + b^q/q$ 给出

$$\sum_i a_i b_i \leq \frac{1}{p} \sum_i a_i^p + \frac{1}{q} \sum_i b_i^q = \frac{1}{p} + \frac{1}{q} = 1, \quad (2.89)$$

再乘回两范数, 并用 $|\mathbf{x}^H \mathbf{y}| \leq \sum_i |x_i y_i|$, 即得 Hölder。端点 $p=1, q=\infty$ 直接由 $\sum_i |x_i y_i| \leq \|\mathbf{y}\|_\infty \sum_i |x_i|$ 得到; 另一端交换两者即可。任一向量为零时, 两端均为零。

Tip

本页要点: 向量范数给出长度和误差的尺度, Hölder 不等式统一了常用的 p 范数关系。

2.25 Matrix Norms

对实矩阵 $A \in \mathbb{R}^{m \times n}$ ，由向量 p -范数诱导的矩阵范数定义为

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}. \quad (2.90)$$

它直接给出作用界

$$\|Ax\|_p \leq \|A\|_p \|x\|_p. \quad (2.91)$$

对任意相容维数的 A, B ,

$$\|AB\|_p = \sup_{x \neq 0} \frac{\|A(Bx)\|_p}{\|x\|_p} \leq \|A\|_p \|B\|_p, \quad (2.92)$$

这个性质称为次乘性。若 $A \in \mathbb{R}^{m \times n}$ 、 $B \in \mathbb{R}^{n \times k}$ ，则三个范数分别作用于 $m \times n$ 、 $n \times k$ 、 $m \times k$ 矩阵，虽然都记为 $\|\cdot\|_p$ ，定义域和值域并不相同。

最常用的三种诱导范数有显式表达式：

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad \|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|. \quad (2.93)$$

第一式是最大绝对列和。三角不等式给出 $\|Ax\|_1 \leq \sum_j |x_j| \sum_i |a_{ij}| \leq (\max_j \sum_i |a_{ij}|) \|x\|_1$ ；取达到最大列和的第 j 个坐标向量 e_j ，等号成立。第二式是最大绝对行和：逐行有 $|(Ax)_i| \leq \sum_j |a_{ij}| |x_j| \leq \sum_j |a_{ij}| \|x\|_\infty$ ；选最大行和所在的行 i ，令 $x_j = \text{sign}(a_{ij})$ ，零元素处取 $x_j = 1$ ，便有 $\|x\|_\infty = 1$ 且该行达到上界。

二范数则为

$$\|A\|_2^2 = \max_{\|x\|_2=1} x^H A^H A x = \lambda_{\max}(A^H A), \quad \|A\|_2 = \sqrt{\lambda_{\max}(A^H A)}. \quad (2.94)$$

理由是 $A^H A$ 为 Hermitian 半正定矩阵。把单位向量展开到它的标准正交特征基中，上面的二次型就是各特征值的加权平均，权重非负且和为 1；取最大特征值对应的单位特征向量时达到最大值。实数情形 $A^H = A^T$ ，这个公式同样适用于复矩阵。

Frobenius 范数定义为

$$\|A\|_F = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2} = \sqrt{\text{tr}(A^H A)}. \quad (2.95)$$

更一般的相容不等式

$$\|Ax\|_\alpha \leq \|A\|_{\alpha \leftarrow \beta} \|x\|_\beta \quad (2.96)$$

实际涉及值域范数、定义域范数和算子范数三种范数；它们不必属于同一类型。

Tip

本页要点： 矩阵范数既可衡量矩阵本身，也可衡量它对向量的放缩作用。

2.26 Spectral Radius

谱半径定义为

$$\rho(A) = \max_{\lambda \in \lambda(A)} |\lambda|. \quad (2.97)$$

谱半径描述矩阵幂的渐近行为。例如在线性迭代 $\mathbf{x}_k = A^k \mathbf{x}_0$ 中，任意初值都趋于零当且仅当 $\rho(A) < 1$ ，证明见性质 3；性质 6 则给出矩阵幂的渐近增长率。

1. 对任意满足次乘性的矩阵范数， $\rho(A) \leq \|A\|$ 。诱导范数情形可直接由特征对 $A\mathbf{x} = \lambda\mathbf{x}$ 得到

$$|\lambda|\|\mathbf{x}\| = \|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|. \quad (2.98)$$

2. 对任意 $\varepsilon > 0$ ，存在一个向量范数诱导的矩阵范数，使 $\|A\|_* \leq \rho(A) + \varepsilon$ 。

④ Note

性质 2 的证明。 先在复数域上将 A 相似变换为上三角矩阵。对阶数归纳：一阶时结论成立；当 $n > 1$ 时，取一个特征向量并扩充为基，组成可逆矩阵 S_1 ，则

$$S_1^{-1}AS_1 = \begin{bmatrix} \lambda_1 & \mathbf{r}^T \\ 0 & A_1 \end{bmatrix}. \quad (2.99)$$

对 A_1 使用归纳假设，取可逆矩阵 S_2 使 $S_2^{-1}A_1S_2$ 上三角。令 $S = S_1 \text{diag}(1, S_2)$ ，便有 $A = STS^{-1}$ ，其中 $T = (t_{ij})$ 上三角。由 $\det(tI - A) = \prod_i (t - t_{ii})$ ，可知其对角元就是 A 的特征值，故 $\max_i |t_{ii}| = \rho(A)$ 。

令 $M = \max_i \sum_{j>i} |t_{ij}|$ ，取 $\delta = \min\{1, \varepsilon/(1+M)\} > 0$ ，并令 $D_\delta = \text{diag}(1, \delta, \dots, \delta^{n-1})$ 。左乘 D_δ^{-1} 缩放各行，右乘 D_δ 缩放各列，所以

$$\begin{aligned} (D_\delta^{-1}TD_\delta)_{ij} &= \delta^{j-i}t_{ij} \quad (i \leq j), \\ \|D_\delta^{-1}TD_\delta\|_\infty &= \max_i \left(|t_{ii}| + \sum_{j>i} \delta^{j-i}|t_{ij}| \right) \\ &\leq \rho(A) + \delta M \leq \rho(A) + \varepsilon. \end{aligned} \quad (2.100)$$

这里用了 $0 < \delta \leq 1$ ，因此 $j > i$ 时 $\delta^{j-i} \leq \delta$ 。定义向量范数 $\|\mathbf{x}\|_* = \|D_\delta^{-1}S^{-1}\mathbf{x}\|_\infty$ 。两个矩阵均可逆，故非零向量的范数严格为正；齐次性和三角不等式继承自无穷范数。由

$$\begin{aligned} \|A\mathbf{x}\|_* &= \|(D_\delta^{-1}TD_\delta)(D_\delta^{-1}S^{-1}\mathbf{x})\|_\infty \\ &\leq (\rho(A) + \varepsilon)\|\mathbf{x}\|_*, \end{aligned} \quad (2.101)$$

对所有非零向量取比值的上确界，即得 $\|A\|_* \leq \rho(A) + \varepsilon$ 。

3. $A^k \rightarrow 0$ 当且仅当 $\rho(A) < 1$ 。若 $\rho(A) < 1$ ，在性质 2 中取 $\varepsilon = (1 - \rho(A))/2$ ，记 $q = \rho(A) + \varepsilon < 1$ 。由诱导范数的次乘性，

$$\|A^k\|_* \leq \|A\|_*^k \leq q^k \rightarrow 0. \quad (2.102)$$

由第 2.24 页的有限维范数等价性，这也给出 $A^k \rightarrow 0$ 。反之，若 $A^k \rightarrow 0$ ，对任意特征对 $A\mathbf{x} = \lambda\mathbf{x}$ 、 $\mathbf{x} \neq \mathbf{0}$ ，有 $A^k\mathbf{x} = \lambda^k\mathbf{x} \rightarrow \mathbf{0}$ ，所以 $|\lambda| < 1$ ，即 $\rho(A) < 1$ 。

4. 级数 $\sum_{k=0}^{\infty} A^k$ 收敛当且仅当 $\rho(A) < 1$ 。若 $\rho(A) < 1$ ，沿用性质 3 的范数和 $q < 1$ ，有 $\sum_{k=0}^{\infty} \|A^k\|_*$

$\leq \sum_{k=0}^{\infty} q^k = 1/(1-q) < \infty$, 故矩阵级数绝对收敛。反之, 级数收敛必有 $A^k \rightarrow 0$, 再由性质 3 得 $\rho(A) < 1$ 。

5. 收敛时, 有限几何和恒等式

$$(I - A) \sum_{k=0}^N A^k = I - A^{N+1} \quad (2.103)$$

中 $A^{N+1} \rightarrow 0$, 且 $\rho(A) < 1$ 保证 1 不是特征值, 故 $I - A$ 可逆。左乘 $(I - A)^{-1}$, 再令 $N \rightarrow \infty$, 得到

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1}. \quad (2.104)$$

6. 对任意固定的矩阵范数, Gelfand 公式为

$$\lim_{k \rightarrow \infty} \|A^k\|^{1/k} = \rho(A). \quad (2.105)$$

将第 2.24 页的范数等价性用于有限维矩阵空间, 存在 $c > 0$, 使 $\|M\| \geq c\|M\|_2$ 。取模最大的特征值及其特征向量, 便有 $\|A^k\|_2 \geq \rho(A)^k$, 因而 $\|A^k\| \geq c\rho(A)^k$ 。

另一方面, 对任意 $\varepsilon > 0$, 取性质 2 构造的诱导范数。范数等价性给出与 k 无关的常数 $C_\varepsilon > 0$, 使 $\|A^k\| \leq C_\varepsilon \|A^k\|_* \leq C_\varepsilon (\rho(A) + \varepsilon)^k$ 。将上下界开 k 次方, 再令 $k \rightarrow \infty$, 得到

$$\rho(A) \leq \liminf_{k \rightarrow \infty} \|A^k\|^{1/k} \leq \limsup_{k \rightarrow \infty} \|A^k\|^{1/k} \leq \rho(A) + \varepsilon. \quad (2.106)$$

因为 $\varepsilon > 0$ 任意, 极限等于 $\rho(A)$, 也包含 $\rho(A) = 0$ 的情形。若矩阵范数满足次乘性, 则 $\|A^k\|^{1/k} \leq \|A\|$, 取极限便证明性质 1 的一般情形。

④ Note

实矩阵的情形。 即使上三角化使用了复矩阵 S , 性质 2 中的 $\|x\|_*$ 限制在实向量上仍是向量范数, 并满足同一个作用界。若性质 1 中使用的是实向量范数, 而特征值为复数, 可先取与 k 无关的常数 $C > 0$, 使 $\|M\|_2 \leq C\|M\|$ 。于是 $|\lambda|^k \leq \|A^k\|_2 \leq C\|A^k\| \leq C\|A\|^k$; 开 k 次方并令 $k \rightarrow \infty$, 仍得 $|\lambda| \leq \|A\|$, 无须把复特征向量代入实向量范数。

💡 Tip

本页要点: 谱半径控制渐近增长, 但非正规矩阵仍可能出现显著瞬态放大。

2.27 Non-negative Matrices

对实矩阵, $A \geq 0$ 表示逐元素 $a_{ij} \geq 0$, $A > 0$ 表示逐元素严格为正。逐元素偏序定义为

$$A \geq B \iff A - B \geq 0. \quad (2.107)$$

若 $A, B \geq 0$, 则加法和乘法保持非负性。若复矩阵 C 满足 $|C| \leq B$ 且 $B \geq 0$, 则

$$|C^k| \leq B^k \quad (2.108)$$

可由归纳和三角不等式得到。取最大绝对行和范数, 则

$$\|C^k\|_{\infty} \leq \|B^k\|_{\infty}. \quad (2.109)$$

两边开 k 次方并用第 2.26 页的 Gelfand 公式, 得到谱半径的单调性

$$|C| \leq B \implies \rho(C) \leq \rho(B). \quad (2.110)$$

非负矩阵可描述人口、产出、概率和网页链接等非负量, 其主特征向量往往也有相应的应用含义。

Tip

本页要点: 非负矩阵保留了概率和数量的非负性, 主特征向量因此常有直接解释。

2.28 Theorem 2.3: Irreducible Matrices

矩阵 $A \in \mathbb{R}^{n \times n}$ ($n \geq 2$) 可约, 是指存在置换矩阵 P , 使

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}. \quad (2.111)$$

其中两个对角块都是非空方阵。左乘 P 、右乘 P^T 是用同一置换重排行和列。若不存在这样的置换, 则称 A 不可约。

分块形式有什么用? 特征多项式可以分解为

$$\det(tI - A) = \det(tI - A_{11}) \det(tI - A_{22}), \quad \lambda(A) = \lambda(A_{11}) \cup \lambda(A_{22}). \quad (2.112)$$

这是因为置换相似不改变行列式, 而分块上三角矩阵的行列式等于对角块行列式之积。因此只需分别求两个小矩阵的特征值, 代数重数按重根次数相加。

对线性方程 $Ax = b$, 令 $y = Px = (y_1^T, y_2^T)^T$, 并将置换后的右端写为 $Pb = (b_1^T, b_2^T)^T$, 则

$$A_{22}y_2 = b_2, \quad A_{11}y_1 = b_1 - A_{12}y_2. \quad (2.113)$$

若两个对角块可逆, 就可以先求 y_2 , 再求 y_1 , 最后由 $x = P^T y$ 恢复原变量。第二组方程不依赖 y_1 , 第一组却可能依赖 y_2 , 这就是两个子系统之间的单向依赖。

对非负矩阵, 以 $a_{ij} > 0$ 表示从顶点 i 到顶点 j 有边, **不可约性等价于这个有向图强连通**。可约时, 式 (2.111) 的左下零块使第二组顶点无法走到第一组, 故图不强连通。反之, 若图不强连通, 取一个不能到达全部顶点的顶点, 令 $\mathcal{R}$ 为从它出发能够到达的顶点集。从 $\mathcal{R}$ 不可能有边通向其补集, 否则该终点也应属于 $\mathcal{R}$ 。将补集排在前、 $\mathcal{R}$ 排在后, 就得到左下块为零的形式, 故矩阵可约。

定理 2.3. 设 $n \geq 2$ 、 $A \geq 0$ 。则 A 不可约当且仅当

$$(I + A)^{n-1} > 0. \quad (2.114)$$

证明. 沿用上面的有向图约定。矩阵幂的展开

$$(A^k)_{ij} = \sum_{i_1, \dots, i_{k-1}} a_{ii_1} a_{i_1 i_2} \cdots a_{i_{k-1} j} \quad (2.115)$$

表明 $(A^k)_{ij} > 0$ 当且仅当存在从 i 到 j 的长度为 k 的有向游走（允许重复经过顶点和边）。加入 I 等价于每个顶点加入自环，因此 $((I + A)^{n-1})_{ij} > 0$ 当且仅当存在长度不超过 $n - 1$ 的路，较短路径可用自环补齐。

由上面的等价关系，不可约时图强连通，即任意 i, j 之间都有有向路。对 $i \neq j$ ，删去游走中重复顶点之间的回路，可得到长度至多 $n - 1$ 的简单路径；对 $i = j$ ，使用自环即可。因此 $(I + A)^{n-1}$ 严格为正。反之，若该矩阵严格正，则任意 i, j 可达，图强连通，故 A 不可约。证毕。

Tip
本页要点： 不可约性等价于有向图强连通，是 Perron–Frobenius 理论的结构前提。

2.29 Theorem 2.4: Perron – Frobenius Theorem

定理 2.4. 若 $n \geq 2$ ，且 $A \in \mathbb{R}^{n \times n}$ 非负、不可约，则 $\rho(A) > 0$ 是 A 的代数单特征值，并存在严格正向量 $\mathbf{x}_* > 0$ 使

$$A\mathbf{x}_* = \rho(A)\mathbf{x}_*. \tag{2.116}$$

证明.

(1) 正矩阵情形。 设 $B > 0$ ，令

$$\Delta = \{\mathbf{x} \geq 0 : \|\mathbf{x}\|_1 = 1\}. \tag{2.117}$$

映射

$$\Phi(\mathbf{x}) = \frac{B\mathbf{x}}{\|B\mathbf{x}\|_1} \tag{2.118}$$

将紧凸集 Δ 连续映到自身。由 Brouwer 不动点定理，存在 $\mathbf{x}_* \in \Delta$ ，使 $\Phi(\mathbf{x}_*) = \mathbf{x}_*$ 。因为 $B > 0$ ，有 $B\mathbf{x}_* > 0$ ，所以 $\mathbf{x}_* > 0$ 。记其第 i 个分量为 $x_{*,i}$ ，并记

$$B\mathbf{x}_* = \alpha\mathbf{x}_*, \quad \alpha = \|B\mathbf{x}_*\|_1 > 0. \tag{2.119}$$

任取特征对 $B\mathbf{z} = \lambda\mathbf{z}$ ，选 i 使

$$\frac{|z_i|}{x_{*,i}} = \max_j \frac{|z_j|}{x_{*,j}}. \tag{2.120}$$

则

$$|\lambda||z_i| \leq \sum_j b_{ij}|z_j| \leq \frac{|z_i|}{x_{*,i}} \sum_j b_{ij}x_{*,j} = \alpha|z_i|. \tag{2.121}$$

故 $|\lambda| \leq \alpha$ 。由于 α 本身是特征值， $\alpha = \rho(B)$ 。

再证特征空间一维。若 $B\mathbf{y} = \alpha\mathbf{y}$ 且 $\mathbf{y}$ 为实向量，令

$$c = \max_j \frac{y_j}{x_{*,j}}, \quad \mathbf{w} = c\mathbf{x}_* - \mathbf{y} \geq 0. \tag{2.122}$$

取达到最大值的下标 i_0 , 则 $w_{i_0} = cx_{*,i_0} - y_{i_0} = 0$ 。又由 $Bx_* = \alpha x_*$ 和 $By = \alpha y$, 有 $Bw = cBx_* - By = \alpha w$ 。

下面证明 $w = 0$, 也就是 $y = cx_*$ 。反设 $w \neq 0$ 。由于 $w \geq 0$, 必有某个分量 $w_{j_0} > 0$ 。再由 $B > 0$, 得到 $(Bw)_{i_0} = \sum_j b_{i_0j} w_j \geq b_{i_0j_0} w_{j_0} > 0$ 。但 $Bw = \alpha w$ 和 $w_{i_0} = 0$ 又给出 $(Bw)_{i_0} = \alpha w_{i_0} = 0$, 矛盾。因此反设不成立, $w = 0$, 从而 $y = cx_*$ 。

所以实特征向量均为 x_* 的倍数。复特征向量的实部和虚部均属于这个实特征空间, 因此复特征向量也为 x_* 的复数倍, 故 $\ker(B - \alpha I)$ 一维。

对 B^T 同样得到左特征向量 $y > 0$, 满足

$$y^T B = \alpha y^T. \quad (2.123)$$

下面证明 α 的代数重数也为 1。令 $\mathcal{S} = \text{null}(y^T)$ 。因为 $y \neq 0$, 由第 2.8 页的秩 - 零化度定理, $\dim \mathcal{S} = n - 1$; 又有 $y^T x_* > 0$, 故 $x_* \notin \mathcal{S}$ 。取 $\mathcal{S}$ 的一组基组成矩阵 X , 则 $Q = [x_*, X]$ 可逆。

对任意 $v \in \mathcal{S}$, 有 $y^T Bv = \alpha y^T v = 0$, 所以 $Bv \in \mathcal{S}$ 。因此存在矩阵 C 使 $BX = XC$, 结合 $Bx_* = \alpha x_*$, 得到

$$Q^{-1}BQ = \begin{bmatrix} \alpha & 0 \\ 0 & C \end{bmatrix}, \quad \det(tI - B) = (t - \alpha) \det(tI - C). \quad (2.124)$$

若 $\det(\alpha I - C) = 0$, 则存在 $z \neq 0$ 使 $Cz = \alpha z$ 。由于 X 列满秩, $Xz \neq 0$, 并且 $B(Xz) = X(Cz) = \alpha Xz$ 。特征空间一维, 故 Xz 必为 x_* 的非零倍数; 但 $Xz \in \mathcal{S}$, 而 $x_* \notin \mathcal{S}$, 矛盾。因此 $\det(\alpha I - C) \neq 0$, 特征多项式中的因子 $t - \alpha$ 恰好出现一次, α 的代数重数为 1。

(2) 不可约非负矩阵情形。 由第 2.28 页的定理 2.3,

$$B = (I + A)^{n-1} > 0. \quad (2.125)$$

设 $Bx_* = \beta x_*$ 是 (1) 中得到的正 Perron 特征对。因为 A 与 B 可交换,

$$B(Ax_*) = A(Bx_*) = \beta Ax_*. \quad (2.126)$$

β 的特征空间一维, 故 $Ax_* = \alpha x_*$ 。由 $A \geq 0$ 、 $x_* > 0$ 得 $\alpha \geq 0$; 不可约矩阵不是零矩阵, 所以至少一行满足 $(Ax_*)_i > 0$, 从而 $\alpha > 0$ 。

对任意 $Az = \lambda z$, 重复最大分量比值估计得到

$$|\lambda| |z_i| \leq \sum_j a_{ij} |z_j| \leq \frac{|z_i|}{x_{*,i}} \sum_j a_{ij} x_{*,j} = \alpha |z_i|. \quad (2.127)$$

因此 $|\lambda| \leq \alpha$, 即 $\alpha = \rho(A)$ 。先证明对应的特征空间一维。任何属于 $\rho(A)$ 的特征向量 z 都满足

$$Bz = (I + A)^{n-1} z = (1 + \rho(A))^{n-1} z. \quad (2.128)$$

而 $Bx_* = (1 + \rho(A))^{n-1} x_*$ 。因此 $(1 + \rho(A))^{n-1} = \beta$, 由 B 的 Perron 特征空间一维, z 必为 x_* 的倍数。

矩阵 A^T 仍不可约非负。将上面的正特征向量存在性论证用于 A^T , 并由第 2.11 页的转置矩阵谱相同结论, 得到 $y > 0$ 满足

$$y^T A = \rho(A) y^T. \quad (2.129)$$

现在 $\mathbf{y}^T \mathbf{x}_* > 0$ ，且 $\rho(A)$ 的特征空间一维。将 (1) 中从 $\mathcal{S} = \text{null}(\mathbf{y}^T)$ 出发的论证用于 A ，便知 $\rho(A)$ 在特征多项式中也只出现一次，故其代数重数为 1。证毕。

邻接矩阵、游走与渐近行为。 Perron 在 1907 年研究了正矩阵，Frobenius 在 1912 年将结论推广到不可约非负矩阵。图的邻接矩阵给出了这一定理的一个直观应用。

若 $a_{ij} \in \{0, 1\}$ 表示从顶点 i 到顶点 j 是否有边，则第 2.28 页式 (2.115) 中，每个乘积恰好对应一条长度为 k 的有向游走：沿途所有边都存在时乘积为 1，否则为 0。因此 $(A^k)_{ij}$ 就是从 i 到 j 的长度为 k 的游走数。这里的 walk 允许重复顶点和边；若矩阵元素是一般非负权重， $(A^k)_{ij}$ 则是这些游走的边权乘积之和。

若存在正整数 m 使 $A^m > 0$ ，称 A 为**本原矩阵 (primitive matrix)**。它一定不可约，否则其所有幂经同一置换后仍有左下零块。本原性比不可约性更强，例如 $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 不可约，但其幂在自身与 I 之间交替，始终不严格正。

令 $\lambda_1 = \rho(A)$ ，取右 Perron 向量 $\mathbf{v} > 0$ 和左 Perron 向量 $\mathbf{u} > 0$ ，满足

$$A\mathbf{v} = \lambda_1\mathbf{v}, \quad \mathbf{u}^T A = \lambda_1\mathbf{u}^T, \quad \mathbf{u}^T \mathbf{v} = 1. \quad (2.130)$$

其中 $\mathbf{v}$ 就是上面 $\mathbf{x}_*$ 的正倍数。由于左右向量均严格正，可以固定其中一个，再缩放另一个，使内积为 1。若 A 本原，则

$$\frac{A^k}{\lambda_1^k} \rightarrow \mathbf{v}\mathbf{u}^T, \quad (A^k)_{ij} \sim \lambda_1^k v_i u_j. \quad (2.131)$$

渐近式的证明。 先说明正矩阵 B 除 Perron 根以外的特征值都严格位于谱圆内部。在 (1) 的式 (2.121) 中，若 $|\lambda| = \alpha$ ，两处不等式都必须取等号。因为 $b_{ij} > 0$ ，第二处取等号迫使所有 $|z_j|/x_{*,j}$ 相同；第一处是三角不等式，其等号又迫使所有非零 z_j 具有同一复相位。因此 $\mathbf{z}$ 是 $\mathbf{x}_*$ 的复数倍，代入特征方程便得 $\lambda = \alpha$ 。

现在取 $B = A^m > 0$ 。由 $B\mathbf{v} = \lambda_1^m \mathbf{v}$ 和 (1) 的最大分量比值估计， $\rho(B) = \lambda_1^m$ 。若 $A\mathbf{z} = \lambda\mathbf{z}$ 且 $|\lambda| = \lambda_1$ ，则 $B\mathbf{z} = \lambda^m \mathbf{z}$ ，而 $|\lambda^m| = \rho(B)$ 。由刚证明的结论及 Perron 特征空间一维， $\mathbf{z}$ 必为 $\mathbf{v}$ 的倍数，因此 $\lambda = \lambda_1$ 。所以 A 的其余特征值都满足 $|\lambda| < \lambda_1$ 。

沿用 (1) 的不变补空间构造，在 $\mathcal{S} = \text{null}(\mathbf{u}^T)$ 中选基组成 X ，令 $Q = [\mathbf{v}, X]$ ，便有 $Q^{-1}AQ = \text{diag}(\lambda_1, C)$ ，且 $\rho(C) < \lambda_1$ 。又因 $\mathbf{u}^T \mathbf{v} = 1$ 、 $\mathbf{u}^T X = 0$ ，矩阵 $\mathbf{v}\mathbf{u}^T$ 在这组基下为 $\text{diag}(1, 0)$ 。因此

$$\frac{A^k}{\lambda_1^k} - \mathbf{v}\mathbf{u}^T = Q \begin{bmatrix} 0 & 0 \\ 0 & (C/\lambda_1)^k \end{bmatrix} Q^{-1} \rightarrow 0. \quad (2.132)$$

最后一步使用第 2.26 页的性质 3，因为 $\rho(C/\lambda_1) < 1$ 。再逐项取极限，并用 $v_i u_j > 0$ ，即得式 (2.131)。当 $\lambda_1 > 1$ 时，游走数以 λ_1^k 的量级增长，因子 $v_i u_j$ 则反映起点与终点的影响； $\lambda_1 = 1$ 时没有这种指数增长。

④ Note

本原性与归一化。 页末渐近式要求本原性；仅有不可约性时，可能存在模同为 $\rho(A)$ 的其他特征值，矩阵幂会振荡。左右 Perron 向量还须满足 $\mathbf{u}^T \mathbf{v} = 1$ ；若分别归一化为单位长度，极限应写成 $\mathbf{v}\mathbf{u}^T / (\mathbf{u}^T \mathbf{v})$ 。

📌 Tip

本页要点： Perron–Frobenius 定理保证不可约非负矩阵有唯一的正主特征方向。

2.30 Google Matrix and PageRank

第 1.12 页从网页链接引出 PageRank，本页用 Perron - Frobenius 定理说明排名向量为什么存在且唯一。设共有 $n \geq 2$ 个网页， M_{ij} 表示从网页 j 沿链接跳到网页 i 的概率。没有出链的网页按第 1.12 页的方法补成均匀跳转列，因而

$$M \geq 0, \quad \sum_{i=1}^n M_{ij} = 1 \quad (j = 1, \dots, n), \quad \mathbf{1}^T M = \mathbf{1}^T. \quad (2.133)$$

这里 $\mathbf{1} \in \mathbb{R}^n$ 是全 1 向量。不等式 $M \geq 0$ 指每个元素非负，列和为 1 表示从任一网页出发的转移概率总和为 1。

以概率 d 沿链接行走、以概率 $1 - d$ 均匀跳到任一网页，得到 Google 矩阵

$$G = dM + \frac{1-d}{n} \mathbf{1} \mathbf{1}^T, \quad 0 < d < 1. \quad (2.134)$$

矩阵 $\mathbf{1} \mathbf{1}^T / n$ 的每列都为均匀概率向量。由 $\mathbf{1}^T \mathbf{1} = n$,

$$\mathbf{1}^T G = d \mathbf{1}^T M + (1-d) \mathbf{1}^T = \mathbf{1}^T. \quad (2.135)$$

所以 G 仍是列随机矩阵，且 1 是它的左特征值。矩阵与转置有相同特征多项式，因此 1 也是 G 的特征值，给出 $\rho(G) \geq 1$ 。另一方面，第 2.25 页的最大绝对列和公式与第 2.26 页的谱半径界给出

$$\|G\|_1 = \max_j \sum_i |G_{ij}| = 1, \quad \rho(G) \leq \|G\|_1 = 1. \quad (2.136)$$

两边合起来得到 $\rho(G) = 1$ 。均匀跳转还使每个元素严格为正：

$$G_{ij} = dM_{ij} + \frac{1-d}{n} \geq \frac{1-d}{n} > 0 \quad (1 \leq i, j \leq n). \quad (2.137)$$

因此任意两个网页之间都有正概率的一步转移， G 不可约。由第 2.29 页的 Theorem 2.4, $\rho(G) = 1$ 是代数单特征值，并有严格正的右特征向量。把该向量除以其分量和，便得到唯一的 PageRank 向量

$$G\mathbf{r} = \mathbf{r}, \quad \mathbf{r} > 0, \quad \mathbf{1}^T \mathbf{r} = 1. \quad (2.138)$$

唯一性来自特征空间只有一维，而分量和为 1 又固定了向量的尺度。 r_i 表示稳定分布中位于网页 i 的概率，概率越大，排名越靠前；本页的 $\mathbf{r}$ 与第 1.12 页的 $\mathbf{p}$ 表示同一对象。

④ Note

备注：左特征向量与右特征向量。 列和为 1 给出的是 $\mathbf{1}^T G = \mathbf{1}^T$ ，即全 1 向量是左特征向量；网页排名由右特征向量 $G\mathbf{r} = \mathbf{r}$ 给出。列随机并不意味着 $G\mathbf{1} = \mathbf{1}$ ，排名一般也不均匀。

💡 Tip

本页要点： 均匀跳转使 Google 矩阵严格为正，Perron - Frobenius 定理保证有唯一的归一化 PageRank 向量。

2.31 Orthogonality

实矩阵 $Q = [q_1, \dots, q_n] \in \mathbb{R}^{n \times n}$ 正交，是指

$$Q^T Q = I, \quad (2.139)$$

等价于列向量满足 $q_i^T q_j = \delta_{ij}$ 。复数情形用酉矩阵

$$Q^H Q = I. \quad (2.140)$$

由于 $Q^{-1} = Q^H$ ，有

$$\|Qx\|_2^2 = x^H Q^H Q x = \|x\|_2^2. \quad (2.141)$$

更一般地，二范数和 Frobenius 范数酉不变：

$$\|QAZ\|_2 = \|A\|_2, \quad \|QAZ\|_F = \|A\|_F \quad (2.142)$$

对任意相容酉矩阵 Q, Z 成立。二范数结论来自单位球在酉变换下不变；Frobenius 结论来自迹的循环不变性。**正交变换不放大向量的二范数误差，是构造稳定算法的重要基础。**

💡 Tip

本页要点： 正交或酉变换保持二范数，是构造稳定数值算法的基础。

2.32 Projections

投影 $P: \mathcal{V} \rightarrow \mathcal{V}$ 是满足

$$P^2 = P \quad (2.143)$$

的线性变换。再次投影不改变结果，因而 $P^k = P$ 对任意正整数 k 成立。由 $(I - P)^2 = I - P$ 和 $(P^T)^2 = (P^2)^T = P^T$ ， $I - P$ 与 P^T 也都是投影。

任意 x 可分解为

$$x = Px + (I - P)x. \quad (2.144)$$

其中 $Px \in \text{range}(P)$ ，由 $P(I - P) = 0$ 可知 $(I - P)x \in \text{null}(P)$ 。值域与零空间的交只有零向量，因此

$$\mathcal{V} = \text{range}(P) \oplus \text{null}(P), \quad \text{range}(I - P) = \text{null}(P). \quad (2.145)$$

若 $Px = \lambda x$ ，则 $\lambda^2 = \lambda$ ，故投影矩阵的特征值只能是 0 或 1。

📌 Note

备注： 在实数空间的标准内积下， $y \in \text{null}(P^T)$ 等价于 $y^T Px = 0$ 对所有 x 成立，故 $\text{null}(P^T) = \text{range}(P)^\perp$ 。两空间交为 $\{0\}$ ，维数之和为 n ，因此 $\mathcal{V} = \text{range}(P) \oplus \text{null}(P^T)$ 是正交直和。直和只要求分解唯一，并不要求正交；同一个子空间可以有不同的补空间。

本页的分解 $x = Px + (I - P)x$ 沿 $\text{null}(P)$ 投影，第二部分一般不属于 $\text{null}(P^T)$ 。若 $P = P^T$ ，两个零空间相同，两种分解才重合，这就是下一页的正交投影。复数情形将 P^T 换成 P^H 。

💡 Tip

本页要点： 投影矩阵满足幂等性，重复投影不会继续改变已经得到的子空间分量。

2.33 Orthogonal Projections

投影 P 若满足 $P^H = P$ ，则称为正交投影。此时

$$\text{null}(P) = \text{range}(P)^\perp. \quad (2.146)$$

若 $\mathbf{x} = P\mathbf{x} + (I - P)\mathbf{x}$ ，两部分内积为

$$(P\mathbf{x})^H(I - P)\mathbf{x} = \mathbf{x}^H(P - P^2)\mathbf{x} = 0. \quad (2.147)$$

因此 $P\mathbf{x}$ 是 $\mathbf{x}$ 在目标子空间中的最佳二范数近似。对任意 $\mathbf{s} \in \text{range}(P)$ ，Pythagoras 恒等式给出

$$\|\mathbf{x} - \mathbf{s}\|_2^2 = \|(I - P)\mathbf{x}\|_2^2 + \|P\mathbf{x} - \mathbf{s}\|_2^2, \quad (2.148)$$

故唯一最小点是 $\mathbf{s} = P\mathbf{x}$ 。

对 $\mathbf{s} \in \text{range}(P)$ ，幂等性给出 $P\mathbf{s} = \mathbf{s}$ 。到同一子空间 $\mathcal{S}$ 的正交投影也唯一：若 P_1, P_2 都满足条件，则 $(P_1 - P_2)\mathbf{x}$ 同时属于 $\mathcal{S}$ 和 $\mathcal{S}^\perp$ ，只能为零。

若 $A \in \mathbb{C}^{m \times n}$ 列满秩，则到 $\text{range}(A)$ 的正交投影为

$$P_A = A(A^H A)^{-1} A^H. \quad (2.149)$$

它满足 $P_A^H = P_A$ 、 $P_A^2 = P_A$ ，且 $P_A A = A$ ，所以其像正好是 $\text{range}(A)$ 。

💡 Tip

本页要点： 正交投影把向量分解为子空间分量与其正交补分量，且投影矩阵满足幂等性。

2.34 Matrix Form of Projection

设 $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{R}^n$ 都是 k 维子空间， $V, W \in \mathbb{R}^{n \times k}$ 的列分别构成它们的标准正交基，并且

$$\mathbb{R}^n = \mathcal{S}_1 \oplus \mathcal{S}_2^\perp. \quad (2.150)$$

本页的投影目标是 $\mathcal{S}_1$ ，投影方向是 $\mathcal{S}_2^\perp$ 。先说明 $W^T V$ 可逆：若 $W^T V \mathbf{c} = \mathbf{0}$ ，则 $V\mathbf{c}$ 与 W 的每一列正交，故 $V\mathbf{c} \in \mathcal{S}_2^\perp$ ；又因 $V\mathbf{c} \in \mathcal{S}_1$ ，直和条件给出 $V\mathbf{c} = \mathbf{0}$ 。由 V 列满秩得 $\mathbf{c} = \mathbf{0}$ ，所以方阵 $W^T V$ 非奇异。所求投影矩阵为

$$P = V(W^T V)^{-1} W^T. \quad (2.151)$$

对任意 $\mathbf{x}$ ，令 $\mathbf{c} = (W^T V)^{-1} W^T \mathbf{x}$ ，则 $P\mathbf{x} = V\mathbf{c} \in \mathcal{S}_1$ 。同时

$$W^T(\mathbf{x} - P\mathbf{x}) = \mathbf{0}, \quad (2.152)$$

故余量属于 $\text{null}(W^T) = \mathcal{S}_2^\perp$ 。这正是沿 $\mathcal{S}_2^\perp$ 到 $\mathcal{S}_1$ 的分解，由直和条件可知分解唯一。

此外, $PV = V$, 所以 $\text{range}(P) = \mathcal{S}_1$, 并有 $P^2 = PV(W^TV)^{-1}W^T = P$ 。由于 V 列满秩且 W^TV 可逆, $Px = 0$ 当且仅当 $W^Tx = 0$, 故 $\text{null}(P) = \mathcal{S}_2^\perp$ 。

一般情形称为斜投影。若 $\mathcal{S}_1 = \mathcal{S}_2$, 可取 $W = V$, 公式化为正交投影 $P = VV^T$ 。两个极端例子是: $\mathcal{S}_1 = \mathcal{S}_2 = \mathbb{R}^n$ 时 $P = I$; $\mathcal{S}_1 = \mathcal{S}_2 = \{0\}$ 时 $P = 0$ 。

Tip

本页要点: V 决定投影到哪里, $W^T(x - Px) = 0$ 决定沿哪个方向投影。

2.35 Derivation of Orthogonal Projection Matrix

设 $V = [v_1, \dots, v_k] \in \mathbb{R}^{n \times k}$ 的列是子空间 $\mathcal{S}$ 的标准正交基, 即

$$V^TV = I_k. \quad (2.153)$$

先看任意非零向量 v : 它所张成的直线上的投影可写成 av 。残差正交条件 $v^T(x - av) = 0$ 给出 $a = (v^Tx)/(v^Tv)$ 。当 $v = v_i$ 是单位长度基向量时, 分母为 1, 因而

$$\text{proj}_{v_i}(x) = (v_i^Tx)v_i. \quad (2.154)$$

把各正交分量相加:

$$Px = \sum_{i=1}^k v_i v_i^T x = VV^T x, \quad P = VV^T. \quad (2.155)$$

由 $P^T = P$ 和 $P^2 = V(V^TV)V^T = P$ 可知, 这确实是正交投影。

若基不正交但 V 列满秩, 则有

$$P = V(V^TV)^{-1}V^T. \quad (2.156)$$

Tip

本页要点: 推导正交投影公式时, 抓住残差与目标子空间正交这一条件。

2.36 Theorem 2.5: Householder Transformation

Householder 矩阵定义为

$$H = I - 2 \frac{uu^T}{u^Tu}, \quad u \neq 0. \quad (2.157)$$

它关于过原点、法向为 u 的超平面作反射: 若 $y \perp u$, 则 $H(y + au) = y - au$ 。因此 -1 是沿 u 方向的特征值, 其余 $n-1$ 个特征值为 1, 且 $\det H = -1$ 。将 u 单位化后, 便得到 $H = I - 2uu^T$ 。

定理 2.5. 给定非零 $x \in \mathbb{R}^n$, 可构造 Householder 矩阵使

$$Hx = \alpha e_1, \quad |\alpha| = \|x\|_2. \quad (2.158)$$

证明. 取

$$\alpha = -\text{sign}(x_1)\|\mathbf{x}\|_2, \quad \mathbf{u} = \mathbf{x} - \alpha\mathbf{e}_1, \quad (2.159)$$

并约定 $\text{sign}(0) = 1$ 。此时 $|u_1| = |x_1| + \|\mathbf{x}\|_2 > 0$ ，既保证 $\mathbf{u} \neq 0$ ，又避免首分量相减时的消去。

记

$$P = \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}. \quad (2.160)$$

因为 $P^T = P$ 、 $P^2 = P$,

$$H^T = H, \quad H^T H = (I - 2P)^2 = I - 4P + 4P^2 = I. \quad (2.161)$$

所以 H 对称且正交。由 $\mathbf{u} = \mathbf{x} - \alpha\mathbf{e}_1$ 和 $\alpha^2 = \|\mathbf{x}\|_2^2$,

$$\mathbf{u}^T\mathbf{u} = \|\mathbf{x}\|_2^2 - 2\alpha x_1 + \alpha^2 = 2(\|\mathbf{x}\|_2^2 - \alpha x_1) = 2\mathbf{u}^T\mathbf{x}. \quad (2.162)$$

代回得到

$$H\mathbf{x} = \mathbf{x} - 2\mathbf{u}\frac{\mathbf{u}^T\mathbf{x}}{\mathbf{u}^T\mathbf{u}} = \mathbf{x} - \mathbf{u} = \alpha\mathbf{e}_1. \quad (2.163)$$

证毕。

💡 Tip

本页要点： Householder 反射用一次正交变换把向量对齐到坐标轴。

2.37 Givens Transformation

实 Givens 旋转作用于两个坐标：

$$G(\Phi) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \cos \Phi, \quad s = \sin \Phi. \quad (2.164)$$

因为 $c^2 + s^2 = 1$ ，有 $G^T G = I$ 。它将 $(x_1, x_2)^T$ 变为 $(cx_1 + sx_2, -sx_1 + cx_2)^T$ 。在固定坐标系中，这对应向量顺时针旋转 Φ ；也可解释为保持向量不动、坐标轴逆时针旋转后的坐标变换。LAPACK 的实 Givens 变换采用上述符号约定。

把平面向量对应到复数 $z = x_1 + ix_2$ ，同一旋转就是乘以单位复数 $u = \cos \Phi - i \sin \Phi = e^{-i\Phi}$ ，因为

$$u(x_1 + ix_2) = (cx_1 + sx_2) + i(-sx_1 + cx_2). \quad (2.165)$$

给定 $(x_1, x_2)^T \neq 0$ ，令

$$r = \sqrt{x_1^2 + x_2^2}, \quad c = x_1/r, \quad s = x_2/r, \quad (2.166)$$

则

$$G \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}. \quad (2.167)$$

④ Note

备注: Givens 与 Householder 的适用场景。

嵌入 $n \times n$ 单位矩阵后, Givens 只改变两行或两列, 适合逐个消去稀疏矩阵中的元素。Householder 一次作用于一个完整向量, 适合稠密批量消元; 二者都是正交变换。

复数情形常用

$$G = \begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix}, \quad c \in \mathbb{R}, \quad c^2 + |s|^2 = 1, \quad (2.168)$$

选择 c, s 使第二分量为零。数值库会通过缩放计算 r , 避免平方和上溢/下溢。

💡 Tip

本页要点: Givens 旋转只改两个坐标, 适合逐元素消去和稀疏更新。

第2.3节 奇异值分解

2.39 Theorem 2.6: Schur Decomposition

定理 2.6. 对任意 $A \in \mathbb{C}^{n \times n}$, 存在酉矩阵 U , 使

$$U^H A U = T, \quad (2.169)$$

其中 T 为上三角矩阵, 且其对角元正是 A 的特征值。

Schur 分解对每个复方阵都存在, 不要求矩阵可对角化。实矩阵也存在实 Schur 形式

$$Q^T A Q = T, \quad (2.170)$$

其中 Q 正交, T 为由 1×1 实特征值块和 2×2 共轭复特征值块组成的拟上三角矩阵。这样即使 A 有复特征值, 也能在实数范围内完成分解。上述复 Schur 分解的证明见第 2.40 页。

💡 Tip

本页要点: 每个复方阵都存在 Schur 分解, 不要求有完整的特征向量基。

2.40 Proof of Schur Decomposition

Theorem 2.6 的证明。 利用 Householder 变换, 对阶数 n 归纳。 $n = 1$ 时, 取 $U = [1]$ 、 $T = A$ 即可。设结论对 $n - 1$ 阶矩阵成立。复数域上的特征多项式有根, 因此可取 A 的特征

$$A x = \lambda x, \quad x \neq 0. \quad (2.171)$$

先把特征向量变到第一个坐标轴。第 2.36 页的 Householder 构造可推广到复数情形: 当 $x_1 \neq 0$ 时令 $\alpha = x_1/|x_1|$, 当 $x_1 = 0$ 时令 $\alpha = 1$, 再取

$$p = -\alpha\|\mathbf{x}\|_2, \quad \mathbf{v} = \mathbf{x} - p\mathbf{e}_1, \quad H = I - 2\frac{\mathbf{v}\mathbf{v}^H}{\mathbf{v}^H\mathbf{v}}. \quad (2.172)$$

这里 $p \neq 0$, 且 $v_1 = \alpha(|x_1| + \|\mathbf{x}\|_2) \neq 0$. 与实数情形相同, $\mathbf{v}\mathbf{v}^H/(\mathbf{v}^H\mathbf{v})$ 是正交投影, 故 $H^H = H$ 、 $H^H H = I$. 又因 $\bar{p}x_1 = -\|\mathbf{x}\|_2|x_1|$ 为实数,

$$\begin{aligned} \mathbf{v}^H \mathbf{x} &= \|\mathbf{x}\|_2^2 - \bar{p}x_1, \\ \mathbf{v}^H \mathbf{v} &= 2(\|\mathbf{x}\|_2^2 - \bar{p}x_1) = 2\mathbf{v}^H \mathbf{x}, \\ H\mathbf{x} &= \mathbf{x} - \mathbf{v} = p\mathbf{e}_1. \end{aligned} \quad (2.173)$$

由 $H^2 = I$ 得 $H\mathbf{e}_1 = \mathbf{x}/p$. 因此 $H A H \mathbf{e}_1 = H A \mathbf{x}/p = \lambda H \mathbf{x}/p = \lambda \mathbf{e}_1$, 说明 $H A H$ 的第一列只有首元可能非零. 写成分块形式,

$$H A H = \begin{bmatrix} \lambda & \mathbf{w}^H \\ 0 & A_1 \end{bmatrix}. \quad (2.174)$$

由归纳假设, 存在 $n-1$ 阶酉矩阵 U_1 , 使 $U_1^H A_1 U_1 = T_1$ 为上三角矩阵. 令

$$U = H \operatorname{diag}(1, U_1), \quad (2.175)$$

则 U 是酉矩阵, 且

$$U^H A U = \begin{bmatrix} \lambda & \mathbf{w}^H U_1 \\ 0 & T_1 \end{bmatrix} \quad (2.176)$$

上三角. 上三角矩阵的特征多项式为对角因子的乘积, 因此其对角元正是 A 的全部特征值. 证毕.

Tip

本页要点: 先用 Householder 变换把一个特征向量对齐到坐标轴, 再对剩余子块归纳.

2.41 Remarks on Schur Decomposition

酉相似变换保持特征值、二范数、Frobenius 范数和角度, 因此 Schur 三角化比一般相似变换更适合数值计算. 上三角矩阵 T 的对角元可按指定顺序排列, 但 Schur 向量和非对角元一般不唯一.

正规矩阵满足 $A^H A = A A^H$. 若 $A = U T U^H$ 是 Schur 分解, 则 T 也是正规矩阵. 上三角正规矩阵必为对角矩阵: 比较 $T T^H$ 与 $T^H T$ 的 $(1,1)$ 元得到第一行非对角元平方和为零; 删去第一行列并归纳即可. 因此

$$A \text{ 为正规矩阵} \iff A = U \Lambda U^H. \quad (2.177)$$

这里 U 为酉矩阵, Λ 为对角矩阵. 反过来, 具有这种分解的矩阵满足

$$A^H A = U \Lambda^H \Lambda U^H, \quad A A^H = U \Lambda \Lambda^H U^H. \quad (2.178)$$

对角矩阵满足 $\Lambda^H \Lambda = \Lambda \Lambda^H$, 故 $A^H A = A A^H$, 即 A 为正规矩阵. 若 A 是 Hermitian 矩阵, 则其特征值为实数: 对任意特征对 $A\mathbf{x} = \lambda\mathbf{x}$,

$$\lambda \mathbf{x}^H \mathbf{x} = \mathbf{x}^H A \mathbf{x} = \overline{\mathbf{x}^H A \mathbf{x}}, \quad (2.179)$$

而 $\mathbf{x}^H \mathbf{x} > 0$, 故 $\lambda = \bar{\lambda}$ 。反过来, 若 $A = U\Lambda U^H$ 且 Λ 为实对角矩阵, 则 $A^H = U\Lambda^H U^H = U\Lambda U^H = A$ 。这就证明了本页两组等价关系。

Tip

本页要点: 每个复方阵都可酉三角化, 正规矩阵还可酉对角化, Hermitian 矩阵的特征值进一步为实数。

2.42 Theorem 2.7: Singular Value Decomposition

定理 2.7. 任意 $A \in \mathbb{C}^{m \times n}$ 都可写为

$$A = U\Sigma V^H, \quad (2.180)$$

其中 $U \in \mathbb{C}^{m \times m}$ 、 $V \in \mathbb{C}^{n \times n}$ 是酉矩阵, $\Sigma \in \mathbb{R}^{m \times n}$ 只有主对角线上可能非零, 且

$$\sigma_1 \geq \cdots \geq \sigma_p \geq 0, \quad p = \min(m, n). \quad (2.181)$$

这称为完整 SVD (full/regular SVD)。 σ_i 是奇异值, U 、 V 的列分别是左、右奇异向量。由 U, V 的酉性,

$$\begin{aligned} AV &= U\Sigma, & A^H U &= V\Sigma^T, \\ Av_i &= \sigma_i u_i, & A^H u_i &= \sigma_i v_i, \quad 1 \leq i \leq p. \end{aligned} \quad (2.182)$$

其中 $\Sigma^H = \Sigma^T$, 逐列等式也适用于零奇异值。若 A 为实矩阵, 可取实正交矩阵 U, V , 将 H 换为 T 。完整构造见第 2.43 页。

Tip

本页要点: SVD 对任意长方矩阵都存在, 并把尺度与正交方向分离。

2.43 Proof of Singular Value Decomposition

证明一: 归纳法。 以下用共轭转置统一处理复数情形, 实数时 H 换为 T 。对 $p = \min(m, n)$ 归纳, 同时证明最大的奇异值等于 $\|A\|_2$ 。 $A = 0$ 时结论成立; 设 $A \neq 0$ 。由单位球面的紧性, 存在单位向量 $\mathbf{x}$, 使

$$\sigma := \|A\|_2 = \|A\mathbf{x}\|_2 > 0, \quad \mathbf{y} := \frac{A\mathbf{x}}{\sigma}, \quad \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1. \quad (2.183)$$

把 $\mathbf{x}$ 和 $\mathbf{y}$ 分别扩充为标准正交基, 得到酉矩阵

$$V_1 = [\mathbf{x}, V_2] \in \mathbb{C}^{n \times n}, \quad U_1 = [\mathbf{y}, U_2] \in \mathbb{C}^{m \times m}. \quad (2.184)$$

由 $A\mathbf{x} = \sigma\mathbf{y}$ 和 $U_1^H \mathbf{y} = \mathbf{e}_1$, 矩阵 $U_1^H A V_1$ 的第一列为 $\sigma\mathbf{e}_1$, 于是

$$A_1 := U_1^H A V_1 = \begin{bmatrix} \sigma & \mathbf{w}^H \\ 0 & B \end{bmatrix}, \quad \mathbf{w} \in \mathbb{C}^{n-1}, \quad B \in \mathbb{C}^{(m-1) \times (n-1)}. \quad (2.185)$$

由第 2.31 页式 (2.142) 的酉不变性, $\|A_1\|_2 = \|A\|_2 = \sigma$ 。令 $\mathbf{e}_1 \in \mathbb{C}^m$ 为第一个坐标向量, 并取 $\mathbf{z} = A_1^H \mathbf{e}_1$, 则

$$\|A_1\|_2 = \sigma, \quad \mathbf{z} = A_1^H \mathbf{e}_1 = \begin{bmatrix} \sigma \\ \mathbf{w} \end{bmatrix}. \quad (2.186)$$

由 $\sigma > 0$ 可知 $\mathbf{z} \neq \mathbf{0}$ 。利用 $\|\mathbf{e}_1\|_2 = 1$ 和第 2.25 页式 (2.90)–(2.91) 的算子范数作用界,

$$\|\mathbf{z}\|_2^2 = |\mathbf{e}_1^H \mathbf{A}_1 \mathbf{z}| \leq \|\mathbf{A}_1 \mathbf{z}\|_2 \leq \|\mathbf{A}_1\|_2 \|\mathbf{z}\|_2 = \sigma \|\mathbf{z}\|_2. \quad (2.187)$$

约去 $\|\mathbf{z}\|_2$, 得到 $\|\mathbf{z}\|_2 \leq \sigma$ 。结合式 (2.186),

$$\begin{aligned} \sigma^2 + \|\mathbf{w}\|_2^2 &= \|\mathbf{z}\|_2^2 \leq \sigma^2, \\ \mathbf{w} &= \mathbf{0}, \quad \mathbf{A}_1 = \begin{bmatrix} \sigma & 0 \\ 0 & B \end{bmatrix}. \end{aligned} \quad (2.188)$$

当 $p = 1$ 时, B 没有行或没有列, 式 (2.188) 已是所需的矩形对角矩阵, 唯一的奇异值为 $\sigma = \|\mathbf{A}\|_2$, 归纳起点成立。

以下设 $p > 1$ 。对任意 $\mathbf{t} \in \mathbb{C}^{n-1}$, 将首分量补零后代入 $\mathbf{A}_1$ 的作用界, 得到

$$\|\mathbf{B}\mathbf{t}\|_2 = \left\| \mathbf{A}_1 \begin{bmatrix} 0 \\ \mathbf{t} \end{bmatrix} \right\|_2 \leq \sigma \|\mathbf{t}\|_2, \quad \|\mathbf{B}\|_2 \leq \sigma. \quad (2.189)$$

因为 $\min(m-1, n-1) = p-1$, 可对 B 应用归纳假设, 写成

$$B = U_B \Sigma_B V_B^H, \quad \tau_1 \geq \cdots \geq \tau_{p-1} \geq 0, \quad \tau_1 = \|\mathbf{B}\|_2. \quad (2.190)$$

定义

$$\tilde{U} := U_1 \text{diag}(1, U_B), \quad \tilde{V} := V_1 \text{diag}(1, V_B). \quad (2.191)$$

$\tilde{U}$ 和 $\tilde{V}$ 仍是酉矩阵, 并且由 (2.188)–(2.190)

$$\mathbf{A} = \tilde{U} \begin{bmatrix} \sigma & 0 \\ 0 & \Sigma_B \end{bmatrix} \tilde{V}^H =: \tilde{U} \Sigma \tilde{V}^H. \quad (2.192)$$

由 (2.189) 及归纳假设, $\sigma = \|\mathbf{A}\|_2 \geq \|\mathbf{B}\|_2 = \tau_1$, 所以

$$\sigma_1 = \sigma \geq \sigma_2 = \tau_1 \geq \cdots \geq \sigma_p = \tau_{p-1} \geq 0. \quad (2.193)$$

式 (2.192) 中的 Σ 恰为 $m \times n$ 的矩形对角矩阵, 完成归纳。左右奇异向量关系随即由式 (2.182) 得到。证毕。

证明二：对 $\mathbf{A}^H \mathbf{A}$ 作谱分解。 矩阵 $\mathbf{A}^H \mathbf{A}$ 是 Hermitian 半正定, 因为

$$\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} = \|\mathbf{A} \mathbf{x}\|_2^2 \geq 0. \quad (2.194)$$

由第 2.41 页从 Schur 分解得到的结论 (式 (2.177)), Hermitian 矩阵 $\mathbf{A}^H \mathbf{A}$ 可酉对角化。半正定性保证其特征值非负, 记为 σ_i^2 。因此存在酉矩阵 V , 使

$$V^H \mathbf{A}^H \mathbf{A} V = \text{diag}(\sigma_1^2, \dots, \sigma_n^2), \quad \sigma_1 \geq \cdots \geq \sigma_n \geq 0. \quad (2.195)$$

对每个正特征值定义

$$\mathbf{u}_i = \frac{\mathbf{A} \mathbf{v}_i}{\sigma_i}. \quad (2.196)$$

这些向量标准正交, 因为

$$\mathbf{u}_i^H \mathbf{u}_j = \frac{\mathbf{v}_i^H A^H A \mathbf{v}_j}{\sigma_i \sigma_j} = \delta_{ij}. \quad (2.197)$$

若 $\sigma_i = 0$, 则

$$\|A \mathbf{v}_i\|_2^2 = \mathbf{v}_i^H A^H A \mathbf{v}_i = 0, \quad A \mathbf{v}_i = 0. \quad (2.198)$$

正奇异值对应的 $\mathbf{u}_i$ 标准正交, 因此其个数 $r \leq m$ 。将它们补成 $\mathbb{C}^m$ 的标准正交基, 组成酉矩阵 U 。令 $p = \min(m, n)$, 取 $\Sigma \in \mathbb{R}^{m \times n}$ 的前 p 个主对角元为 $\sigma_1, \dots, \sigma_p$, 其余元素为零; 当 $n > m$ 时, $i > m$ 的 σ_i 本来就为零。逐列比较得到

$$AV = U\Sigma, \quad A = U\Sigma V^H. \quad (2.199)$$

证毕。这种证明利用前面的 Schur 推论, 直接构造左右奇异向量。

Tip

本页要点: 最大伸长方向使右上块必须为零, 剩下的子块再递归分解, 就得到 SVD。

2.44 Reduced and Condensed SVD

本页讨论 $m \geq n$ 的情形。完整 SVD 为

$$A = U\Sigma V^H, \quad U \in \mathbb{C}^{m \times m}, \quad \Sigma \in \mathbb{R}^{m \times n}, \quad V \in \mathbb{C}^{n \times n}. \quad (2.200)$$

删除 U 中只与 Σ 的零行相乘的列, 得到 thin/reduced SVD

$$A = U_n \Sigma_n V_n^H, \quad U_n \in \mathbb{C}^{m \times n}, \quad U_n^H U_n = I_n, \quad \Sigma_n = \text{diag}(\sigma_1, \dots, \sigma_n). \quad (2.201)$$

若 $r = \text{rank}(A)$, 再删除所有零奇异值对应部分, 得到 condensed SVD

$$A = U_r \Sigma_r V_r^H, \quad (2.202)$$

其中

$$U_r \in \mathbb{C}^{m \times r}, \quad V_r \in \mathbb{C}^{n \times r}, \quad \Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) > 0. \quad (2.203)$$

这里 $U_r^H U_r = V_r^H V_r = I_r$ 。由 $A = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^H$ 可见, 删去的项都乘着零奇异值, 因此三种形式表示同一个矩阵, 并未损失 A 的信息。区别在于保留了多少与零空间相关的正交基; 薄矩阵的列标准正交, 但它本身未必是方阵。

Tip

本页要点: 薄 SVD 去掉矩形分解中多余的列; 紧致 SVD 进一步只保留非零奇异值对应的部分。

2.45 Geometric Meaning of SVD

SVD 的存在性已在第 2.43 页证明, 这里看它怎样描述线性变换的几何作用。对实矩阵 $A \in \mathbb{R}^{m \times n}$, 写成 $A = U\Sigma V^T$, 其中 U 和 V 分别是 m 阶、 n 阶正交矩阵, $r = \text{rank}(A)$, 正奇异值为 $\sigma_1 \geq \dots \geq \sigma_r > 0$ 。矩阵乘积作用于向量时从右向左计算, 因此三步顺序是 V^T 、 Σ 、 U 。

第一步: V^T 改变输入方向。 令 $y = V^T x$, 则 $y_i = v_i^T x$ 是输入向量在右奇异向量 v_i 方向上的坐标。因为 $V^T V = I$, 有 $\|y\|_2 = \|x\|_2$, 所以单位球的形状和大小不变, 只改变方向。本页将这一步称为 Rotation/Reflection: 正交矩阵的行列式可以为 -1 , 所以也可能包含反射。

第二步: Σ 沿坐标轴拉伸或压缩。 前 r 个坐标分别乘以 σ_i , 其余输入坐标被消去; 输出中剩余的坐标为零。 $\sigma_i > 1$ 表示该方向被拉长, $0 < \sigma_i < 1$ 表示被缩短, 零奇异值对应的方向则被完全压扁。矩形矩阵的输入、输出维数不同, 也在这一步体现。

第三步: U 改变输出方向。 U 把第 i 个输出坐标轴转到左奇异向量 u_i 的方向, 同样只作正交变换, 不再改变长度。三步合起来就是

$$\begin{aligned} x &= \sum_{i=1}^n y_i v_i, & y_i &= v_i^T x, \\ Ax &= U\Sigma V^T x = \sum_{i=1}^r \sigma_i y_i u_i, \\ Av_i &= \sigma_i u_i, & 1 \leq i \leq r, \\ \|x\|_2^2 &= \sum_{i=1}^n y_i^2, & \|Ax\|_2^2 &= \sum_{i=1}^r \sigma_i^2 y_i^2. \end{aligned} \tag{2.204}$$

这也解释了左右奇异向量的区别: v_i 是输入的特殊方向, u_i 是它经映射后的输出方向, σ_i 是长度的放大倍数。一般并不是同一个方向, 因此奇异向量不应直接当作特征向量。

为什么奇异值是椭球的半轴长? 先取单位球体 $\|x\|_2 \leq 1$ 。记输出 $z = Ax$, 其在 u_i 方向上的坐标为 $z_i = u_i^T z = \sigma_i y_i$ 。于是 $\sum_{i=1}^r (z_i / \sigma_i)^2 = \sum_{i=1}^r y_i^2 \leq 1$ 。反过来, 满足这个不等式的任意 r 个坐标, 都可以取 $y_i = z_i / \sigma_i$ 、其余 $y_i = 0$ 得到。因此单位球体的像恰是位于 $\text{span}\{u_1, \dots, u_r\}$ 中的椭球体, 半轴方向为 u_i , 半轴长为 σ_i 。特别地, $x = \pm v_i$ 被映到半轴端点 $\pm \sigma_i u_i$ 。

再考虑单位球面的像 $\{Ax : \|x\|_2 = 1\}$ 。若 A 列满秩, 即 $r = n$, 其像是上述椭球体的边界; 若 $r < n$, 则可以在零空间方向补足输入长度, 使 $\sum_{i=r+1}^n y_i^2 = 1 - \sum_{i=1}^r y_i^2$, 而不改变输出, 球面的像因而包含整个低维椭球体, 不只是边界。 $m > r$ 时, 像位于输出空间的一个 r 维子空间中; $r = 0$ 时, 整个像退化为原点。

SVD 给出的方向和尺度既能描述几何变形, 也能用于判断秩、分析条件数、求最小二乘解和构造低秩近似。

Tip

本页要点: 右奇异向量给出输入方向, 左奇异向量给出输出方向, 奇异值给出对应的伸缩倍数。

2.46 Geometric Meaning of SVD and EVD

沿用上一页的 SVD 几何解释。对 $A = U\Sigma V^T$, 令 $y = V^T x$, 则

$$Ax = U\Sigma y. \tag{2.205}$$

V^T 与 U 改变方向, Σ 决定椭球半轴长度, 零奇异值对应被压扁的方向。EVD 则描述不变方向: $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ 使特征向量留在原直线上, 负特征值会反转方向。

④ Note

和 EVD 的几何区别。 若 $A = X\Lambda X^{-1}$, EVD 的特征向量列 X 一般不是正交矩阵, 因此沿特征向量方向展开后, 长度和角度不会像 SVD 那样保持不变。实对称矩阵是特殊情形: $A = Q\Lambda Q^T$, $Q^T Q = I$, 这时特征方向正交, 几何图像才可以直接看成沿正交主轴的缩放。

🔗 Tip

本页要点: SVD 把任意实矩阵分成定义域旋转、按奇异值缩放和值域旋转三步。

2.47 Comparison of SVD and EVD

两种分解的适用范围不同。SVD $A = U\Sigma V^H$ 对每个实或复长方形矩阵都存在, 左、右奇异向量分别给出值域和定义域的正交方向, σ_i 是方向增益。EVD $AX = X\Lambda$ 需要方阵, 并且只有在拥有特征向量基时才能写成对角形式; 特征向量方向可能不正交且对扰动敏感。

求秩、计算条件数、构造最佳低秩近似或求解最小二乘问题时, 应优先考察奇异值; 分析动力系统模态、稳定性或不变子空间时, 则还需分析特征值、Schur 形式和非正规耦合。两种分解不能互相替代: SVD 描述输入到输出的方向增益, EVD 描述保持方向不变的谱结构。

④ Note

两种分解何时相互联系? 由 $A = U\Sigma V^H$ 得 $A^H A = V(\Sigma^T \Sigma)V^H$, 所以 σ_i^2 (以及必要的补零) 是 $A^H A$ 的特征值。实数情形把 H 换为 T ; 完整 SVD 中 Σ 可能为矩形, 此时应写 $\Sigma^T \Sigma$, 不能直接写 Σ^2 。

若 A 为实对称矩阵, 由第 2.41 页有 $A = Q\Lambda Q^T$, 从而 $A^T A = Q\Lambda^2 Q^T$, 所以奇异值是特征值的绝对值, 按绝对值重新排序后有 $\sigma_i = |\lambda_i|$ 。负特征值还表示方向反转, 这个符号要吸收到左奇异向量中, 因此一般不能取 $U = V = Q$ 。若 A 进一步正定, 则所有特征值都为正, 可以取 $U = V = Q$ 、 $\Sigma = \Lambda$ 。

🔗 Tip

本页要点: SVD 适合秩和低秩近似, EVD 适合模态和不变子空间; 两者揭示不同结构。

2.48 Comments on SVD and EVD

SVD、EVD 与 Schur 分解之间有如下关系。若 $Q^H A Q = D + N$ 是 Schur 形式, 其中 $D = \text{diag}(\lambda_i)$ 、 N 严格上三角, 则

$$\begin{aligned}\sigma(A) &= \sigma(Q^H A Q) = \sigma(D + N), \\ \sigma_{\min}(A) &\leq \min_i |\lambda_i| \leq \max_i |\lambda_i| \leq \sigma_{\max}(A), \\ \max_{i,j} \left| \frac{\lambda_i}{\lambda_j} \right| &\leq \kappa_2(A), \quad A \text{ 非奇异}.\end{aligned}\tag{2.206}$$

第一行来自奇异值的酉不变性。由于 $(Q^H A Q)^H (Q^H A Q) = Q^H A^H A Q$, 它与 $A^H A$ 酉相似, 因此二者特征值相同, 开平方后得到同一组奇异值。

第二行来自奇异值对伸缩倍数的界：对单位特征向量 $\mathbf{x}$ ，有 $\sigma_{\min}(A) \leq \|A\mathbf{x}\|_2 = |\lambda| \leq \sigma_{\max}(A)$ 。对全部特征值取最小值和最大值，再在非奇异条件下相除，即得第三行的条件数界。

当 A 为正规矩阵时 $N = 0$ ，于是 $A = QDQ^H$ ，并且 $A^H A = Q \operatorname{diag}(|\lambda_i|^2) Q^H$ ，所以奇异值作为多重集恰为 $|\lambda_i|$ 。一般非正规情形中， N 描述偏离正交特征向量基的耦合；特征值比值可能远小于 $\kappa_2(A)$ ，因而 EVD 的特征值信息不能替代 SVD 的方向放大因子。

对每个奇异三元组， $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ 把输入方向、输出方向和放大倍数区分开来，适用于任意长方矩阵；另有 $A^H \mathbf{u}_i = \sigma_i \mathbf{v}_i$ 。

Tip

本页要点： SVD 的奇异方向总是正交，而 EVD 的特征向量未必正交。

2.49 SVD and Norms

SVD 中正奇异值的个数就是 $\operatorname{rank}(A)$ 。由酉不变性， $\|A\|_2 = \|\Sigma\|_2$ ；对任意单位向量 $\mathbf{y}$ ，

$$\|\Sigma \mathbf{y}\|_2^2 = \sum_i \sigma_i^2 |y_i|^2 \leq \sigma_1^2, \quad (2.207)$$

而取 $\mathbf{y} = \mathbf{e}_1$ 达到等号，因此

$$\|A\|_2 = \sigma_1. \quad (2.208)$$

Frobenius 范数满足

$$\|A\|_F^2 = \operatorname{tr}(A^H A) = \operatorname{tr}(V \Sigma^H \Sigma V^H) = \sum_i \sigma_i^2. \quad (2.209)$$

核范数定义为

$$\|A\|_* = \sum_i \sigma_i. \quad (2.210)$$

核范数是凸且连续的，常在机器学习中用作促进低秩结构的正则项。**三种范数分别对应奇异值向量的 ∞ 、2 和 1 范数。**

Note

最小增益。 当 $m \geq n$ 时，SVD 还给出 $\min_{\mathbf{x} \neq 0} \|A\mathbf{x}\|_2 / \|\mathbf{x}\|_2 = \sigma_n$ 。令 $\mathbf{y} = V^H \mathbf{x}$ ，则比值等于 $(\sum_i \sigma_i^2 |y_i|^2 / \sum_i |y_i|^2)^{1/2}$ ，是奇异值平方的加权平均的平方根，因而不小于 σ_n ；取 $\mathbf{x} = \mathbf{v}_n$ 达到等号。若 A 列不满秩，则 $\sigma_n = 0$ 。

将 $A^H A$ 的特征值从大到小排列，有 $\sigma_i(A) = \sqrt{\lambda_i(A^H A)}$ ， $1 \leq i \leq \min(m, n)$ 。若 A 是正规矩阵，则 $A = Q\Lambda Q^H$ ，从而奇异值按重数等于特征值的模；实对称矩阵是其中的特例。

Tip

本页要点： SVD 把谱范数、Frobenius 范数和核范数都归结为奇异值的范数。

2.50 Theorem 2.8: Low-Rank Approximation

定理 2.8. 设紧致 SVD 为

$$A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H, \quad \sigma_1 \geq \cdots \geq \sigma_r > 0. \quad (2.211)$$

对 $0 \leq k \leq r$, 令

$$A_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^H, \quad (2.212)$$

约定 $\sigma_{r+1} = 0$ 。则

$$\min_{\text{rank}(B) \leq k} \|A - B\|_F = \left(\sum_{i=k+1}^r \sigma_i^2 \right)^{1/2}, \quad (2.213)$$

$$\min_{\text{rank}(B) \leq k} \|A - B\|_2 = \sigma_{k+1}, \quad (2.214)$$

且 A_k 同时达到两种最小值。

证明. 先设 $k < r$ 。对任意 $\text{rank}(B) \leq k$, 有 $\dim \ker(B) \geq n - k$, 可在其中取标准正交向量 $\mathbf{x}_1, \dots, \mathbf{x}_{n-k}$ 。令 $\mathbf{v}_i$ 是 $A^H A$ 的标准正交特征向量, 并置

$$\alpha_i = \sum_{j=1}^{n-k} |\mathbf{v}_i^H \mathbf{x}_j|^2. \quad (2.215)$$

由 Bessel 不等式, $0 \leq \alpha_i \leq 1$; 交换求和并用 $\mathbf{x}_j$ 单位正交, 得到 $\sum_i \alpha_i = n - k$ 。把奇异值补零到 n 个, 令 $a_i = \sigma_i^2$ 且 $a_1 \geq \cdots \geq a_n \geq 0$ 。置 $\beta_i = 1 - \alpha_i$, 则

$$0 \leq \beta_i \leq 1, \quad \sum_i \beta_i = k. \quad (2.216)$$

在这些约束下, $\sum_i a_i \beta_i$ 的最大值为 $\sum_{i=1}^k a_i$ 。确实, 若 $i < j$, $\beta_i < 1$, $\beta_j > 0$, 把 $\min(1 - \beta_i, \beta_j)$ 的权重从 j 移到 i 不会减小该和; 有限次移动得到 $\beta_1 = \cdots = \beta_k = 1$ 。因此

$$\sum_i a_i \alpha_i = \sum_i a_i - \sum_i a_i \beta_i \geq \sum_{i=k+1}^n a_i = \sum_{i=k+1}^r \sigma_i^2. \quad (2.217)$$

将 $\mathbf{x}_1, \dots, \mathbf{x}_{n-k}$ 补成标准正交基。对 $M = A - B$, Frobenius 范数在任意标准正交基上满足

$$\|M\|_F^2 = \sum_{j=1}^n \|M \mathbf{x}_j\|_2^2. \quad (2.218)$$

所以

$$\begin{aligned}
\|A - B\|_F^2 &\geq \sum_{j=1}^{n-k} \|(A - B)\mathbf{x}_j\|_2^2 \\
&= \sum_{j=1}^{n-k} \|A\mathbf{x}_j\|_2^2 = \sum_i \sigma_i^2 \alpha_i \\
&\geq \sum_{i=k+1}^r \sigma_i^2,
\end{aligned} \tag{2.219}$$

其中第二行使用 $B\mathbf{x}_j = 0$ 。

对二范数，由维数公式

$$\dim \ker(B) + \dim \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\} \geq n + 1, \tag{2.220}$$

两个子空间有非零交。取交中的单位向量 $\mathbf{x} = \sum_{i=1}^{k+1} c_i \mathbf{v}_i$ ，则

$$\begin{aligned}
\|A - B\|_2^2 &\geq \|(A - B)\mathbf{x}\|_2^2 = \|A\mathbf{x}\|_2^2 \\
&= \sum_{i=1}^{k+1} \sigma_i^2 |c_i|^2 \geq \sigma_{k+1}^2.
\end{aligned} \tag{2.221}$$

取 $B = A_k$ 时，

$$A - A_k = \sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H, \tag{2.222}$$

故其 Frobenius 范数和二范数分别等于上述尾部平方和开方与 σ_{k+1} ，所有下界同时达到。当 $k = r$ 时 $A_r = A$ ，两种误差均为零。证毕。

💡 Tip

本页要点： Eckart--Young--Mirsky 定理说明截断 SVD 是最佳低秩近似。

2.51 SVD and Numerical Rank

精确秩对任意非零奇异值都计数，但浮点数据含噪时，需要阈值 $\varepsilon \geq 0$ 。数值秩可定义为：

$$r_\varepsilon(A) = \min\{\text{rank}(B) : \|A - B\|_2 \leq \varepsilon\}. \tag{2.223}$$

由第 2.50 页的 Eckart--Young--Mirsky 定理（定理 2.8），

$$r_\varepsilon(A) = \#\{i : \sigma_i > \varepsilon\}. \tag{2.224}$$

也就是说，保留严格大于阈值的奇异值；在信号处理中，其余部分常作为噪声舍去。仅考虑浮点舍入误差时，常用的基准阈值是

$$\varepsilon = \max(m, n)u\sigma_1. \tag{2.225}$$

其中 u 是单位舍入误差；若数据本身含噪，还应考虑噪声尺度。只看绝对阈值会依赖数据缩放，奇异值曲线中的明显转折或间隙也可帮助选取阈值。

Tip

本页要点： 数值秩必须结合阈值和噪声尺度判断，不能只看精确秩。

2.52 SVD for Invertible Matrices

若 $A \in \mathbb{C}^{n \times n}$ 可逆，则所有奇异值均为正。由 $A = U\Sigma V^H$ 得

$$A^{-1} = V\Sigma^{-1}U^H. \quad (2.226)$$

所以逆矩阵的奇异值为 $1/\sigma_i(A)$ ，特别地

$$\|A^{-1}\|_2 = \frac{1}{\sigma_n(A)}, \quad \kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1(A)}{\sigma_n(A)}. \quad (2.227)$$

行列式满足

$$|\det A| = |\det U| |\det \Sigma| |\det V^H| = \prod_{i=1}^n \sigma_i(A), \quad (2.228)$$

因为酉矩阵行列式的模为 1。

Tip

本页要点： 可逆矩阵的所有奇异值均为正，因此 SVD 可以直接给出逆矩阵。

2.53 SVD and Moore – Penrose Pseudoinverse

对任意 $A \in \mathbb{C}^{m \times n}$ ，Moore–Penrose 伪逆 $A^\dagger$ 由四个条件唯一确定：

$$AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (2.229)$$

$$(AA^\dagger)^H = AA^\dagger, \quad (A^\dagger A)^H = A^\dagger A. \quad (2.230)$$

若紧致 SVD 为

$$A = U_r \Sigma_r V_r^H, \quad (2.231)$$

则

$$A^\dagger = V_r \Sigma_r^{-1} U_r^H. \quad (2.232)$$

代入可得

$$AA^\dagger = U_r U_r^H, \quad A^\dagger A = V_r V_r^H, \quad (2.233)$$

二者分别是到 $\text{range}(A)$ 和 $\text{range}(A^H)$ 的正交投影，满足式 (2.230)。再利用 $U_r^H U_r = V_r^H V_r = I$ ，有 $AA^\dagger A = U_r \Sigma_r V_r^H = A$ 和 $A^\dagger AA^\dagger = V_r \Sigma_r^{-1} U_r^H = A^\dagger$ ，式 (2.229) 也成立。伪逆只倒置非零奇异值，零奇异值保持为零。

💡 Tip

本页要点： 伪逆只倒置非零奇异值，并给出最小范数最小二乘解。

2.54 SVD and Fundamental Subspaces

紧致 SVD 可写为秩一和

$$A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H. \quad (2.234)$$

于是每一列都属于 $\text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ ，且 $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ 表明这些 $\mathbf{u}_i$ 都在列空间中。因此

$$\text{range}(A) = \text{range}(AA^H) = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_r\}. \quad (2.235)$$

其中 $AA^H = U_r \Sigma_r^2 U_r^H$ 给出相同的值域。对 A^H 同理，

$$\text{range}(A^H) = \text{range}(A^H A) = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}. \quad (2.236)$$

完整 SVD 中其余右奇异向量满足 $A\mathbf{v}_i = 0$ ，故

$$\text{null}(A) = \text{span}\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}, \quad (2.237)$$

$$\text{null}(A^H) = \text{span}\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}. \quad (2.238)$$

这同时证明 $\text{rank}(A) = \text{rank}(A^H) = r$ 。由 $AA^H = U_r \Sigma_r^2 U_r^H$ 和 $A^H A = V_r \Sigma_r^2 V_r^H$ ，还得到 $\text{rank}(AA^H) = \text{rank}(A^H A) = r$ 。

💡 Tip

本页要点： SVD 同时给出矩阵的秩、值域和零空间，是后续推导的统一工具。

2.55 SVD and Orthogonal Projections

把完整 SVD 的奇异向量按秩分块：

$$U = [U_r, U_0], \quad V = [V_r, V_0]. \quad (2.239)$$

由第 2.54 页的子空间关系，四个正交投影为

$$\begin{aligned} P_{\text{range}(A)} &= U_r U_r^H = AA^\dagger, \\ P_{\text{null}(A^H)} &= U_0 U_0^H = I - AA^\dagger, \\ P_{\text{range}(A^H)} &= V_r V_r^H = A^\dagger A, \\ P_{\text{null}(A)} &= V_0 V_0^H = I - A^\dagger A. \end{aligned} \quad (2.240)$$

这些等式把四个基本子空间的正交直和关系写成矩阵恒等式。例如

$$\mathbb{C}^m = \text{range}(A) \oplus \text{null}(A^H) \quad (2.241)$$

对应 $U_r U_r^H + U_0 U_0^H = I$ 。

💡 Tip

本页要点： SVD 中的左右奇异向量给出四个基本子空间的正交投影。

2.56 Projection to Column Space

由 $A = U_r \Sigma_r V_r^H$,

$$AA^H = U_r \Sigma_r^2 U_r^H \quad (2.242)$$

由 U_r 列满秩且 Σ_r^2 可逆, $\text{range}(AA^H) = \text{range}(U_r \Sigma_r^2) = \text{range}(U_r) = \text{range}(A)$, 所以

$$\text{range}(A) = \text{range}(AA^H). \quad (2.243)$$

此外, $A^\dagger = V_r \Sigma_r^{-1} U_r^H$ 、 $A^\dagger A = V_r V_r^H$, 所以 $\text{range}(A^\dagger) = \text{range}(A^\dagger A) = \text{range}(V_r) = \text{range}(A^H)$ 。这说明 $A^\dagger A$ 是定义域中行空间的正交投影, $AA^\dagger$ 是值域中列空间的正交投影。

正交投影可统一写成

$$P_{\text{range}(A)} = AA^\dagger. \quad (2.244)$$

若 A 列满秩, $A^\dagger = (A^H A)^{-1} A^H$, 故

$$P_A = A(A^H A)^{-1} A^H. \quad (2.245)$$

④ Note

高亮等式为什么成立？ 这不是伪逆定义直接给出的, 而是由紧致 SVD 推出。若 A 列满秩, 设 $A = U_r \Sigma_r V_r^H$, 则 $r = n$, 且 Σ_r 可逆, 所以 $A^H A$ 可逆。于是 $A^H A = V_r \Sigma_r^2 V_r^H$, 从而 $(A^H A)^{-1} = V_r \Sigma_r^{-2} V_r^H$ 。代入可得 $(A^H A)^{-1} A^H = V_r \Sigma_r^{-2} V_r^H V_r \Sigma_r U_r^H = V_r \Sigma_r^{-1} U_r^H = A^\dagger$ 。

行满秩的公式可用同样的计算得到。因此, 两条显式公式分别对应列满秩和行满秩, 不能把其中一条的适用条件省略。若 A 不是列满秩, $A^H A$ 必然不可逆, 应回到 SVD 的伪逆公式。

若 A 行满秩, 则 $A^\dagger = A^H (AA^H)^{-1}$ 。伪逆公式不要求满秩, 因此比正规方程形式更一般。

💡 Tip

本页要点： 投影到列空间把最小二乘的几何条件写成了可计算的矩阵公式。

2.57 Distance Between Subspaces

设 $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{R}^n$ 维数相同, P_1, P_2 是相应正交投影。定义

$$\text{dist}(\mathcal{S}_1, \mathcal{S}_2) = \|P_1 - P_2\|_2. \quad (2.246)$$

它取值于 $[0, 1]$: 两个子空间相同时为 0, 存在互相正交的主方向时达到 1。若最大主角为 $\theta_{\max}$, 则

$$\|P_1 - P_2\|_2 = \sin \theta_{\max}. \quad (2.247)$$

令两子空间的公共维数为 k ，把各自的一组标准正交基扩充为 $\mathbb{R}^n$ 的标准正交基，组成正交矩阵

$$W = [W_1, W_2], \quad Z = [Z_1, Z_2], \quad (2.248)$$

其中 W_1, Z_1 各有 k 列，分别张成 $\mathcal{S}_1, \mathcal{S}_2$ ，所以 $P_1 = W_1 W_1^T$ 、 $P_2 = Z_1 Z_1^T$ 。下面用二范数的正交不变性和定义计算投影差的范数；复数情形将 T 换成 H ，正交矩阵换成酉矩阵，证明相同。

④ Note

分块范数为何取两个交叉块的最大值？ 先设 $0 < k < n$ 。由 $W^T W_1 = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$ 和 $Z_1^T Z = [I_k \ 0]$ ，得到

$$\begin{aligned} M := W^T(P_1 - P_2)Z &= \begin{bmatrix} W_1^T Z_1 & W_1^T Z_2 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} W_1^T Z_1 & 0 \\ W_2^T Z_1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & W_1^T Z_2 \\ -W_2^T Z_1 & 0 \end{bmatrix}. \end{aligned} \quad (2.249)$$

左右乘正交矩阵不改变二范数，故 $\|P_1 - P_2\|_2 = \|M\|_2$ 。记 $B = W_1^T Z_2$ 、 $C = W_2^T Z_1$ ，并令 $\alpha = \max\{\|B\|_2, \|C\|_2\}$ 。

对分块向量 $\mathbf{x} = (\mathbf{a}^T, \mathbf{b}^T)^T$ ，有 $M\mathbf{x} = ((B\mathbf{b})^T, -(C\mathbf{a})^T)^T$ ，所以

$$\begin{aligned} \|M\mathbf{x}\|_2^2 &= \|B\mathbf{b}\|_2^2 + \|C\mathbf{a}\|_2^2 \\ &\leq \alpha^2(\|\mathbf{a}\|_2^2 + \|\mathbf{b}\|_2^2) = \alpha^2\|\mathbf{x}\|_2^2. \end{aligned} \quad (2.250)$$

因而 $\|M\|_2 \leq \alpha$ 。反过来，分别取 $\mathbf{a} = \mathbf{0}$ 或 $\mathbf{b} = \mathbf{0}$ ，对另一个单位长度分块取上确界，可得 $\|M\|_2 \geq \|B\|_2$ 和 $\|M\|_2 \geq \|C\|_2$ 。结合上下界，

$$\|P_1 - P_2\|_2 = \|M\|_2 = \max\{\|W_1^T Z_2\|_2, \|W_2^T Z_1\|_2\}. \quad (2.251)$$

若 $k = 0$ 或 $k = n$ ，则两个投影同为零矩阵或单位矩阵，距离为零；把空块视为范数为零的零映射，上式仍成立。下一页再证明两个交叉块的二范数相等。

💡 Tip

本页要点： 等维子空间之间的距离可由相应正交投影之差的范数衡量。

2.58 Distance Between Subspaces, Continued

沿用第 2.57 页的分块，设 $0 < k < n$ ，令

$$Q = W^T Z = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \quad (2.252)$$

为正交矩阵。

④ Note

两个交叉块的范数为什么相等? 对任意 $\|x\|_2 = 1$, 由 $Q \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} Q_{11}x \\ Q_{21}x \end{bmatrix}$ 和 $Q^T Q = I$, 有 $\|Q_{11}x\|_2^2 + \|Q_{21}x\|_2^2 = 1$. 对单位向量取最大值, 得到 $\|Q_{21}\|_2^2 = 1 - \min_{\|x\|_2=1} \|Q_{11}x\|_2^2 = 1 - \sigma_{\min}(Q_{11})^2$. 对 Q^T 同理得到 $\|Q_{12}^T\|_2^2 = 1 - \sigma_{\min}(Q_{11}^T)^2$. 方阵 Q_{11} 与其转置的奇异值相同, 且 $\|Q_{12}^T\|_2 = \|Q_{12}\|_2$, 故两个交叉块的二范数相等。

因此

$$\|Q_{12}\|_2 = \|Q_{21}\|_2 = \sqrt{1 - \sigma_{\min}(Q_{11})^2}. \quad (2.253)$$

结合第 2.57 页的分块范数结论, 以及 $Q_{12} = W_1^T Z_2$ 、 $Q_{21} = W_2^T Z_1$, 得到

$$\text{dist}(\mathcal{S}_1, \mathcal{S}_2) = \|W_1^T Z_2\|_2 = \|W_2^T Z_1\|_2. \quad (2.254)$$

💡 Tip

本页要点: 子空间距离可以通过投影矩阵之差测量, 奇异值给出对应的夹角信息。

2.59 Cos-Sin Decomposition

Cos-Sin 分解描述正交矩阵相对于两个分块子空间的结构。设

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \in \mathbb{R}^{n \times n} \quad (2.255)$$

正交。分块正交条件给出

$$Q_{11}^T Q_{11} + Q_{21}^T Q_{21} = I, \quad (2.256)$$

所以两个块的非平凡奇异值可配对写为

$$c_i = \cos \theta_i, \quad s_i = \sin \theta_i, \quad c_i^2 + s_i^2 = 1. \quad (2.257)$$

存在块对角正交矩阵 $U = \text{diag}(U_1, U_2)$ 、 $V = \text{diag}(V_1, V_2)$, 使

$$U^T Q V = \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & C & 0 & 0 & S & 0 \\ 0 & 0 & 0 & 0 & 0 & I \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & -S & 0 & 0 & C & 0 \\ 0 & 0 & I & 0 & 0 & 0 \end{bmatrix}, \quad C = \text{diag}(c_i), \quad S = \text{diag}(s_i). \quad (2.258)$$

这里 $0 < c_i, s_i < 1$; 取值为 0 或 1 的情形已单独列为零块和单位块, 各块尺寸由原分块决定, 允许某些块为空。角 θ_i 刻画相应子空间的主角, 非平凡部分由二维旋转块组成。

💡 Tip

本页要点: Cos-Sin 分解把两个子空间的相对位置化为主角和二维旋转块。

2.60 Theorem 2.9: Cos-Sin Decomposition for Unitary Matrices

定理 2.9. 设 $n = k + j$, $k \leq j$, 并令

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad (2.259)$$

其中 $Q_{11} \in \mathbb{C}^{k \times k}$ 、 $Q_{22} \in \mathbb{C}^{j \times j}$, 且 Q 为酉矩阵。则存在酉矩阵 $U_1, V_1 \in \mathbb{C}^{k \times k}$ 、 $U_2, V_2 \in \mathbb{C}^{j \times j}$ 和非负对角矩阵

$$C = \text{diag}(c_1, \dots, c_k), \quad S = \text{diag}(s_1, \dots, s_k), \quad (2.260)$$

满足 $C^2 + S^2 = I$, 并且

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}^H Q \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} = \begin{bmatrix} C & S & 0 \\ -S & C & 0 \\ 0 & 0 & I_{j-k} \end{bmatrix}. \quad (2.261)$$

其中 $c_i = \cos \theta_i$ 、 $s_i = \sin \theta_i$, 可取 $0 \leq \theta_1 \leq \dots \leq \theta_k \leq \pi/2$ 。完整证明见 [02线性代数基础.pdf](#) 第 2.62 页。

Tip

本页要点: Cos-Sin 分解用主角和二维旋转块刻画两个子空间的相对位置。

2.61 Theorem 2.10: Thin Cos-Sin Decomposition

定理 2.10. 设

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \in \mathbb{R}^{(m_1+m_2) \times n}, \quad Q^T Q = I_n, \quad (2.262)$$

其中 $m_1, m_2 \geq n$ 。则存在正交矩阵 $U_1 \in \mathbb{R}^{m_1 \times m_1}$ 、 $U_2 \in \mathbb{R}^{m_2 \times m_2}$ 、 $V_1 \in \mathbb{R}^{n \times n}$, 以及非负对角矩阵 $C, S \in \mathbb{R}^{n \times n}$, 满足 $C^2 + S^2 = I$, 使

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}^T \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} V_1 = \begin{bmatrix} C \\ 0 \\ S \\ 0 \end{bmatrix}. \quad (2.263)$$

两个零块的行数分别为 $m_1 - n$ 和 $m_2 - n$ 。可取 $C = \text{diag}(\cos \theta_i)$ 、 $S = \text{diag}(\sin \theta_i)$, 其中 $0 \leq \theta_1 \leq \dots \leq \theta_n \leq \pi/2$ 。完整证明见第 2.62 页。

Tip

本页要点: 薄 Cos-Sin 分解利用列正交条件, 得到配对的余弦、正弦因子。

2.62 Proof of Cos-Sin Decomposition

Theorem 2.9 的证明

由 SVD Theorem 2.7 ([02线性代数基础.pdf](#) 第 2.42--2.44 页), 对 Q_{11} 作分解:

$$Q_{11} = U_1 C V_1^H, \quad C = \text{diag}(c_i), \quad 0 \leq c_i \leq 1. \quad (2.264)$$

由 $Q^H Q = I$ 的左上块等式 $Q_{11}^H Q_{11} + Q_{21}^H Q_{21} = I$, 得到

$$(Q_{21} V_1)^H (Q_{21} V_1) = I - C^2 =: S^2. \quad (2.265)$$

对每个 $s_i > 0$, 令

$$\mathbf{u}_{2,i} = -\frac{Q_{21} \mathbf{v}_{1,i}}{s_i}. \quad (2.266)$$

上述 Gram 矩阵等式表明这些向量标准正交。若 $s_i = 0$, 则 $Q_{21} \mathbf{v}_{1,i} = 0$; 在已有向量的正交补中任选单位向量作为 $\mathbf{u}_{2,i}$ 。将所得 k 个向量补成 $\mathbb{C}^j$ 的标准正交基, 组成 U_2 。于是

$$U_2^H Q_{21} V_1 = \begin{bmatrix} -S \\ 0 \end{bmatrix}. \quad (2.267)$$

记

$$\tilde{Q} = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}^H Q \begin{bmatrix} V_1 & 0 \\ 0 & I_j \end{bmatrix}. \quad (2.268)$$

$\tilde{Q}$ 仍为酉矩阵。其前 k 列为 M , 构造 N 如下:

$$M = \begin{bmatrix} C \\ -S \\ 0 \end{bmatrix}, \quad N = \begin{bmatrix} S & 0 \\ C & 0 \\ 0 & I_{j-k} \end{bmatrix}. \quad (2.269)$$

由 $C^2 + S^2 = I$ 和 $CS = SC$, 有 $N^H N = I$ 、 $M^H N = 0$, 故 N 的列构成 $\text{range}(M)^\perp$ 的标准正交基。 $\tilde{Q}$ 的后 j 列也是该空间的标准正交基, 所以存在酉矩阵 V_2 , 使这些后列右乘 V_2 后等于 N , 得到定理中的块矩阵。上述补基过程也覆盖 $s_i = 0$ 或 $c_i = 0$ 的情形。证毕。

Theorem 2.10 的证明

由 SVD Theorem 2.7 ([02线性代数基础.pdf](#) 第 2.42--2.44 页), 对 Q_1 作分解, 取正交矩阵 U_1, V_1 使

$$U_1^T Q_1 V_1 = \begin{bmatrix} C \\ 0 \end{bmatrix}, \quad C = \text{diag}(c_1, \dots, c_n), \quad 1 \geq c_1 \geq \dots \geq c_n \geq 0. \quad (2.270)$$

由列正交条件 $Q_1^T Q_1 + Q_2^T Q_2 = I$, 得到

$$(Q_2 V_1)^T (Q_2 V_1) = I - C^2 =: S^2. \quad (2.271)$$

对 $s_i > 0$, 令

$$\mathbf{u}_{2,i} = Q_2 \mathbf{v}_{1,i} / s_i. \quad (2.272)$$

Gram 矩阵等式保证这些向量标准正交。对 $s_i = 0$ 的指标, 有 $Q_2 \mathbf{v}_{1,i} = 0$, 在已有向量的正交补中补选单位向量。继续补成 $\mathbb{R}^{m_2}$ 的标准正交基, 组成 U_2 , 便有

$$U_2^T Q_2 V_1 = \begin{bmatrix} S \\ 0 \end{bmatrix}. \quad (2.273)$$

与 Q_1 的分解合并即得结论； $C^2 + S^2 = I$ 已由列正交性给出。证毕。

Tip

本页要点： 先对一个块作 SVD，再由列正交性构造另一个块的配对因子，并补全正交基。

2.63 Example: Information Geometry

模型有很多参数，为什么预测有时只对少数参数组合敏感？可以用 SVD 分析当前参数附近的情形。

设参数为 $\theta \in \mathbb{R}^n$ ，预测观测为 $f(\theta) \in \mathbb{R}^m$ 。灵敏度矩阵 $J = \partial f / \partial \theta \in \mathbb{R}^{m \times n}$ 的第 j 列描述第 j 个参数对各观测量的二阶影响，即 $\delta f \approx J \delta \theta$ 。

观测模型为 $y = f(\theta) + \varepsilon$ ，其中 $\varepsilon \sim N(0, C)$ ，协方差矩阵 C 对称正定且不依赖参数。高斯对数似然中与参数有关的部分是 $\ell = -\frac{1}{2}(y - f)^T C^{-1}(y - f)$ ，对参数求导得 $g = \nabla_{\theta} \ell = J^T C^{-1} \varepsilon$ 。Fisher 信息矩阵定义为这一梯度的二阶矩，因此

$$F = \mathbb{E}[gg^T] = J^T C^{-1} \mathbb{E}[\varepsilon \varepsilon^T] C^{-1} J = J^T C^{-1} J. \quad (2.274)$$

左乘 $C^{-1/2}$ 后，噪声的协方差变为 $C^{-1/2} C C^{-1/2} = I$ ，称为“白化”。它把观测变化放到统一的噪声尺度下比较；若 C 是对角矩阵，就是将各观测分量除以各自的噪声标准差。对加权灵敏度矩阵作完整 SVD，利用 $U^T U = I$ ，得到

$$C^{-1/2} J = U \Sigma V^T, \quad F = (C^{-1/2} J)^T (C^{-1/2} J) = V \Sigma^T \Sigma V^T. \quad (2.275)$$

当 Σ 是方形对角矩阵时， $\Sigma^T \Sigma = \Sigma^2$ ；完整 SVD 中的 Σ 若为矩形矩阵，则只能写成 $\Sigma^T \Sigma$ ，与第 2.47 页的说明一致。令 $p = \min(m, n)$ ，参数扰动对应的噪声加权变化量满足

$$\|C^{-1/2} J \delta \theta\|_2^2 = \delta \theta^T F \delta \theta = \sum_{i=1}^p \sigma_i^2 (v_i^T \delta \theta)^2. \quad (2.276)$$

沿 v_i 方向取 $\delta \theta = \alpha v_i$ ，上式就是 $\alpha^2 \sigma_i^2$ 。对相同大小的参数变化，大奇异值方向更容易从噪声中辨识，小奇异值方向则较难辨识；零奇异值方向不产生一阶观测变化。若只有前 k 个奇异值显著， $\mathcal{S}_k = \text{span}\{v_1, \dots, v_k\}$ 就给出局部有效参数子空间。这里保留的是参数的线性组合，不一定是某几个原始参数。

Note

适用范围。 这一判断针对当前参数附近，并要求先固定参数的单位或缩放；改变参数尺度会改变奇异值和所选方向。它不意味着所有复杂模型都能全局降到低维。

页中论文：Machta、Chachra、Transtrum 与 Sethna, *Parameter Space Compression Underlies Emergent Theories and Predictive Models*, Science 342, 604 - 607 (2013), DOI: [10.1126/science.1238723](https://doi.org/10.1126/science.1238723)。

Tip

本页要点： 对噪声加权的灵敏度矩阵做 SVD，可找出局部最容易辨识的参数组合。

2.64 Example: From Text to Latent Semantics

潜在语义分析把“文档包含哪些词”组织为矩阵，再用低秩 SVD 建立词和文档的共享低维空间。若 $A \in \mathbb{R}^{N_d \times N_t}$ 的行对应文档、列对应词项，则

$$A \approx A_k = U_k \Sigma_k V_k^T. \tag{2.277}$$

文档坐标可取 $U_k \Sigma_k$ ，词项坐标可取 $V_k \Sigma_k$ 。原空间中不同但在相似文档中出现的词，可能在低秩空间中获得相近坐标。

低秩截断过滤小奇异值方向，可能去除噪声和偶然共现，也可能丢失稀有但重要的语义。秩 k 必须根据奇异值能量、下游任务和验证数据选择。

 **Tip**

本页要点： 潜在语义分析先构造词项矩阵，再用低秩 SVD 提取主题方向。现代主流大语言模型通常不通过 SVD 构造语义表示；潜在语义分析（LSA）是较早的经典方法，两者的表示构造方式不同。

2.65 Data Analysis Step 1: Preprocessing

示例包含 5 篇短文档，主题覆盖数值代数、机器学习、篮球和足球。预处理依次执行：转为小写、分词、去标点，并删除停用词。结果是每篇文档的 token 列表。

词形还原、复合词处理和停用词表都会改变词汇表及后续矩阵。本例采用简单规则，便于逐项复算。

 **Tip**

本页要点： 预处理规则决定矩阵的列，改变词形或停用词规则也会改变后续 SVD 的结果。

2.66 Data Analysis Step 1: Preprocessing (continued)

把所有保留词去重并排序，得到 22 个词项。对 5 篇文档，原始计数矩阵为

$$N = (n_{dt}) \in \mathbb{R}^{5 \times 22}, \tag{2.278}$$

其中 n_{dt} 是词 t 在文档 d 中的出现次数。页中的每一行是一个长度为 22 的稀疏向量。

行索引对应文档，列索引对应排序后的词汇表；改变词汇表顺序时必须同步重排矩阵列。

 **Tip**

本页要点： 文本预处理决定词项矩阵的行列含义，停用词和词形规则必须先固定。

2.67 Data Analysis Step 2: Computing Weights

词频定义为

$$\text{TF}(t, d) = \frac{n_{dt}}{\sum_s n_{ds}}. \tag{2.279}$$

分母是文档 d 中保留 token 的总数，因此 TF 表示该词在该文档中的相对频率。示例中文档长度为 5 或 6，所以出现一次的词权重分别为 $1/5 = 0.20$ 或 $1/6 \approx 0.17$ 。

只用 TF 会让在所有文档中普遍出现的词占据较大权重。逆文档频率 IDF 用文档层面的稀有程度对其重新缩放。

Tip
本页要点： TF 用词频除以文档长度，先消除不同文档长度对词频的影响。

2.68 Data Analysis Step 2: Computing Weights (IDF)

设文档总数为 N ，包含词 t 的文档数为 $\text{df}(t)$ 。本例使用

$$\text{IDF}(t) = \ln \frac{N}{\text{df}(t)} + 1. \tag{2.280}$$

加 1 使出现在全部文档中的词也有正权重。对 $N = 5$ ，三种文档频率对应的权重为

$\text{df}(t)$	1	2	3
$\text{IDF}(t)$	$\ln 5 + 1$	$\ln(5/2) + 1$	$\ln(5/3) + 1$
近似值	2.61	1.92	1.51

(2.281)

文档频率越低，IDF 越大；这有助于突出主题专用词，但稀有词不一定具有更强的语义区分度。

Tip
本页要点： IDF 根据词语跨文档出现的普遍程度赋予权重，越稀有的词权重越高。

2.69 Data Analysis Step 2: Computing Weights (multiplication)

最终权重为

$$A_{dt} = \text{TF}(t, d) \text{IDF}(t). \tag{2.282}$$

例如，文档 1 长度为 6。若 `algebra` 出现在 3 篇文档中，则

$$A_{1,\text{algebra}} = \frac{1}{6} \left(\ln \frac{5}{3} + 1 \right) \approx 0.17 \times 1.51 \approx 0.25. \tag{2.283}$$

若 `important` 只出现在一篇文档中，则

$$A_{1,\text{important}} = \frac{1}{6} (\ln 5 + 1) \approx 0.17 \times 2.61 \approx 0.43. \tag{2.284}$$

页中小数只为显示而舍入；SVD 应使用未舍入的全精度权重，否则多个舍入误差会改变相近奇异值和向量。

Tip
本页要点： TF-IDF 的最终权重是词频与逆文档频率的乘积。

2.70 Data Analysis Step 2: Computing Weights (TF – IDF matrix)

把每篇文档的 22 个权重按固定词汇表顺序排列，得到

$$A \in \mathbb{R}^{5 \times 22}. \tag{2.285}$$

矩阵大部分元素为零，因为单篇短文档只含少数词。行向量描述文档，列向量描述词在不同文档中的分布。

④ Note

备注：大规模语料的存储与计算。

在更大语料上，稀疏的 TF–IDF 矩阵可使用 CSR/CSC 存储。完整 SVD 的因子通常是稠密的，存储和计算成本较高，因此通常只计算前 k 个奇异三元组。

💡 Tip

本页要点： TF – IDF 同时考虑词在单篇文档中的频率和跨文档的稀有程度。

2.71 Applying SVD to TF–IDF Matrix

示例的前五个奇异值约为

$$1.176, \quad 1.071, \quad 1.044, \quad 0.873, \quad 0.753. \tag{2.286}$$

前 k 个分量保留的平方 Frobenius 能量比例为

$$\eta_k = \frac{\sum_{i=1}^k \sigma_i^2}{\sum_i \sigma_i^2}. \tag{2.287}$$

页中 $k = 2$ 时约为 51.13%， $k = 3$ 时约为 73.15%。由于 σ_2 与 σ_3 接近，在第二个分量后截断会分开两个权重相近的方向。 σ_3 后的间隔最大，因此本例保留三个分量，但这里并没有很突出的奇异值截断点。

文档嵌入为

$$Z_d = U_3 \Sigma_3 \in \mathbb{R}^{5 \times 3}. \tag{2.288}$$

奇异向量整体变号不改变重构和相似性，所以坐标符号本身没有绝对语义，只有相对几何关系有意义。

💡 Tip

本页要点： 在 TF – IDF 数据上，截断 SVD 用少量方向保留文档之间的主要相似性。

2.72 Dimensionality Reduction to 3D Space

词项嵌入取

$$Z_t = V_3 \Sigma_3 \in \mathbb{R}^{22 \times 3}. \tag{2.289}$$

等价地，取 V^T 的前三行，并把第 i 行乘以 σ_i ，再按词项读取列。页中的坐标显示：第一坐标区分体育词与其他词；第二坐标聚集 numerical/algebra/AI；第三坐标进一步区分篮球和足球相关词。

Note

备注：低维坐标轴不一定唯一。

解释坐标轴时应注意旋转不唯一性：若一组奇异值相等，相应奇异子空间内可任意正交旋转。稳定对象是子空间及其内积，不一定是单根坐标轴。

Tip

本页要点： 低维坐标由截断 SVD 的右奇异向量给出，可用于观察词项之间的主题结构。

2.73 Document Similarity Calculation

两个非零向量 u, v 的余弦相似度为

$$\cos \theta = \frac{u^T v}{\|u\|_2 \|v\|_2}. \quad (2.290)$$

它只比较方向，忽略整体尺度。对归一化行向量，全部两两相似度可用 Gram 矩阵一次计算。

在原 TF-IDF 空间中，文档 1、2、5 的两两相似度约为 0.15, 0.15, 0.29；截断至三维后，它们都沿第二坐标轴，相似度变为 1。文档 3、4 分别讨论篮球和足球，相似度在两个空间中均约为 0.12；它们与文档 1、2、5 的相似度均为零。

低秩投影会改变相似度，并可能合并原有差异，因此仍需结合任务选择 k 。

Tip

本页要点： 余弦相似度只比较方向不比较长度，适合衡量文档主题而非篇幅。

第2.4节 矩阵的扰动分析

2.75 Condition of A Linear System

设 A 非奇异、 $b \neq 0$ ，并考虑

$$(A + \varepsilon F)x(\varepsilon) = b + \varepsilon f, \quad x(0) = x_*. \quad (2.291)$$

当 ε 足够小时， $\det(A + \varepsilon F) \neq 0$ ，所以 $x(\varepsilon) = (A + \varepsilon F)^{-1}(b + \varepsilon f)$ 在零点附近光滑。将

$$x(\varepsilon) = x_* + \varepsilon \dot{x}(0) + O(\varepsilon^2) \quad (2.292)$$

代入并比较 ε 一阶项：

$$A\dot{x}(0) + Fx_* = f, \quad \dot{x}(0) = A^{-1}(f - Fx_*). \quad (2.293)$$

于是

$$\|x(\varepsilon) - x_*\| \leq |\varepsilon| \|A^{-1}\| (\|f\| + \|F\| \|x_*\|) + O(\varepsilon^2). \quad (2.294)$$

由 $\|b\| = \|Ax_*\| \leq \|A\| \|x_*\|$ ，得到相对界

$$\frac{\|x(\varepsilon) - x_*\|}{\|x_*\|} \leq |\varepsilon| \kappa(A) \left(\frac{\|F\|}{\|A\|} + \frac{\|f\|}{\|b\|} \right) + O(\varepsilon^2), \quad (2.295)$$

其中

$$\kappa(A) = \|A\| \|A^{-1}\|. \quad (2.296)$$

条件数大表示某些相对输入扰动可能被明显放大，并不表示每个方向的扰动都会造成很大的解误差。

④ Note

相对界中的条件数如何出现？ 矩阵扰动项可写为 $\|A^{-1}\| \|F\| = \kappa(A) \|F\| / \|A\|$ 。右端扰动项再用 $1/\|x_*\| \leq \|A\|/\|b\|$ ，便得到式 (2.295)。

💡 Tip

本页要点： 条件数衡量数据扰动被问题放大的程度，解释解误差时必须考虑它。

2.76 Examples of Ill-Conditioned Systems

对非奇异矩阵 A ，特征值模的最大比值只给出二范数条件数的下界：

$$\max_{i,j} \frac{|\lambda_i|}{|\lambda_j|} = \rho(A) \rho(A^{-1}) \leq \|A\|_2 \|A^{-1}\|_2 = \kappa_2(A). \quad (2.297)$$

这里 A^{-1} 的特征值为 $1/\lambda_i$ ，再用第 2.26 页的谱半径界即可。正规矩阵取等号；非正规矩阵的两个量可能相差很大。

上双对角矩阵。 第一个矩阵的主对角元均为 1，上副对角元均为 2。令 N 的上副对角元为 1，其余元素为零，则 $A = I + 2N$ 。由于 $N^n = 0$ ，有限几何和给出

$$A^{-1} = \sum_{k=0}^{n-1} (-2N)^k, \quad (A^{-1})_{ij} = \begin{cases} (-2)^{j-i}, & j \geq i, \\ 0, & j < i. \end{cases} \quad (2.298)$$

当 $n \geq 2$ 时， $\|A\|_\infty = 3$ ，而 A^{-1} 的最大绝对行和为 $\sum_{k=0}^{n-1} 2^k = 2^n - 1$ 。所以

$$\kappa_\infty(A) = 3(2^n - 1). \quad (2.299)$$

二范数下也有指数增长。由 $|(A^{-1})_{1n}| = 2^{n-1}$ 得 $\|A^{-1}\|_2 \geq 2^{n-1}$ ，而 $\|A\|_2 \geq \|Ae_1\|_2 = 1$ 。另一方面， $\|N\|_2 = 1$ 给出 $\|A\|_2 \leq 3$ 及 $\|A^{-1}\|_2 \leq \sum_{k=0}^{n-1} 2^k = 2^n - 1$ 。因此

$$2^{n-1} \leq \kappa_2(A) \leq 3(2^n - 1), \quad n \geq 2. \quad (2.300)$$

虽然 A 的特征值全为 1，条件数仍按 2^n 的量级增长；特征值远离零并不能保证非正规矩阵良态。

📌 Important

讲课时需强调： 对这个上副对角元为 2 的矩阵，条件数是指数增长，不能写成 $O(n^2)$ 。特征值全为 1 并不意味着奇异值也都接近 1。

Hilbert 矩阵。 第二个矩阵满足 $H_{ij} = 1/(i+j-1)$ 。它是单项式 $1, t, \dots, t^{n-1}$ 在区间 $[0, 1]$ 上的 Gram 矩阵，因为

$$H_{ij} = \int_0^1 t^{i-1} t^{j-1} dt, \quad \mathbf{c}^T H \mathbf{c} = \int_0^1 \left(\sum_{i=1}^n c_i t^{i-1} \right)^2 dt > 0 \quad (\mathbf{c} \neq \mathbf{0}). \quad (2.301)$$

最后一个严格不等式来自非零多项式不可能在整个区间上恒为零，因此 H 正定。随着阶数增大，单项式之间会出现严重的近似线性相关，使 H 的最小特征值很小。这说明正定性也不能保证矩阵良态。

更精细的渐近分析给出 $\kappa_2(H) = O((1 + \sqrt{2})^{4n}/\sqrt{n})$ 。不推导这一精细估计，也可以用一个多项式直接说明其病态程度至少呈指数增长。取 $(1 - t)^{n-1}$ 的系数向量 $\mathbf{c}$ ，则

$$\begin{aligned} \mathbf{c}^T H \mathbf{c} &= \int_0^1 (1 - t)^{2n-2} dt = \frac{1}{2n-1}, \\ \|\mathbf{c}\|_2^2 &= \sum_{j=0}^{n-1} \binom{n-1}{j}^2 = \binom{2n-2}{n-1}. \end{aligned} \quad (2.302)$$

最后一个等式可比较 $(1 + t)^{n-1}(1 + t)^{n-1}$ 中 t^{n-1} 的系数得到。由第 2.41 页的正交对角化，将 $\mathbf{c}$ 展开在特征向量基上，其 Rayleigh 商是特征值的加权平均，因此 $\lambda_{\min}(H) \leq (\mathbf{c}^T H \mathbf{c})/\|\mathbf{c}\|_2^2$ 。同理， $\lambda_{\max}(H) \geq \mathbf{e}_1^T H \mathbf{e}_1 = 1$ ，所以

$$\kappa_2(H) \geq (2n-1) \binom{2n-2}{n-1} \geq 4^{n-1}. \quad (2.303)$$

第二个不等式利用了中央二项式系数不小于同一行 $2n-1$ 个二项式系数的平均值，而这些系数之和为 2^{2n-2} 。这个下界已足以说明：即使矩阵对称正定，其条件数也可能随阶数迅速增长。

Tip

本页要点： 特征值远离零或矩阵对称正定，都不能单独保证问题良态；二范数条件数取决于最大、最小奇异值之比。

2.77 Theorem 2.11: Inverse of Perturbed Matrix

定理 2.11. 若 $F \in \mathbb{R}^{n \times n}$ 、 $\|F\|_p < 1$ ，则

$$(I - F)^{-1} = \sum_{k=0}^{\infty} F^k, \quad \|(I - F)^{-1}\|_p \leq \frac{1}{1 - \|F\|_p}. \quad (2.304)$$

证明. 若 $I - F$ 奇异，则存在 $\mathbf{x} \neq \mathbf{0}$ 使 $F\mathbf{x} = \mathbf{x}$ ，从而 $\|\mathbf{x}\|_p \leq \|F\|_p \|\mathbf{x}\|_p$ ，与 $\|F\|_p < 1$ 矛盾。

有限几何和满足

$$\left(\sum_{k=0}^N F^k \right) (I - F) = (I - F) \left(\sum_{k=0}^N F^k \right) = I - F^{N+1}. \quad (2.305)$$

由于 $\|F^{N+1}\|_p \leq \|F\|_p^{N+1} \rightarrow 0$ ，令 $N \rightarrow \infty$ 得 Neumann 展开。三角不等式给出

$$\|(I - F)^{-1}\|_p \leq \sum_{k=0}^{\infty} \|F\|_p^k = \frac{1}{1 - \|F\|_p}. \quad (2.306)$$

证毕。

应用：扰动矩阵的逆。 若 A 非奇异且

$$\gamma = \|A^{-1}E\|_p < 1, \quad (2.307)$$

则 $A + E$ 非奇异，并且

$$\|(A + E)^{-1} - A^{-1}\|_p \leq \frac{\|E\|_p \|A^{-1}\|_p^2}{1 - \gamma}. \quad (2.308)$$

推导。 先分解

$$A + E = A(I + A^{-1}E). \quad (2.309)$$

令 $F = -A^{-1}E$ ，便知第二个因子可逆，因此 $A + E$ 也可逆。由第 2.21 页的预解式恒等式（式 (2.71)），取 $B = A + E$ ，再代入 $(A + E)^{-1} = (I + A^{-1}E)^{-1}A^{-1}$ ，得到

$$\begin{aligned} (A + E)^{-1} - A^{-1} &= -(A + E)^{-1}EA^{-1} \\ &= -(I + A^{-1}E)^{-1}A^{-1}EA^{-1}. \end{aligned} \quad (2.310)$$

取范数，再用 $\|(I + A^{-1}E)^{-1}\|_p \leq (1 - \gamma)^{-1}$ ，即得结论。证毕。

 **Tip**

本页要点： Neumann 级数把小扰动下的逆矩阵写成可控的几何级数。

2.78 Sensitivity of Matrix Inversion

由第 2.50 页定理 2.8 的最佳低秩近似结论，非奇异方阵 A 满足

$$\min_{\text{rank}(B) < n} \|A - B\|_2 = \sigma_n(A). \quad (2.311)$$

这里取截断 SVD 的秩为 $n - 1$ 即可达到等号。因此最小奇异值正是 A 到奇异矩阵集合的二范数距离。 σ_n 很小意味着只需很小的绝对扰动就可能使矩阵不可逆。

再看这种敏感性如何反映在线性方程组的解中。若 $A = U\Sigma V^H$ 非奇异，则

$$A^{-1}\mathbf{b} = V\Sigma^{-1}U^H\mathbf{b} = \sum_{i=1}^n \frac{\mathbf{u}_i^H \mathbf{b}}{\sigma_i} \mathbf{v}_i. \quad (2.312)$$

小奇异值会放大 $\mathbf{b}$ 在相应左奇异向量方向的分量。最坏相对放大由

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_n} \quad (2.313)$$

刻画。因而绝对敏感性取决于 $1/\sigma_n$ ，相对敏感性则要结合最大奇异值一起判断。

 **Tip**

本页要点： 逆矩阵对小奇异值方向最敏感，误差会在这些方向被显著放大。

2.79 Theorem 2.12: Bauer – Fike Theorem

定理 2.12 (Bauer – Fike) . 设 $A \in \mathbb{C}^{n \times n}$ 可对角化, 写成

$$A = X^{-1} \Lambda X. \quad (2.314)$$

其中 X 非奇异、 Λ 为对角矩阵。对任意 $B \in \mathbb{C}^{n \times n}$ 和 $\mu \in \lambda(B)$, 存在 $\lambda \in \lambda(A)$ 使

$$|\mu - \lambda| \leq \kappa_2(X) \|A - B\|_2. \quad (2.315)$$

证明. 若 $\mu \in \lambda(A)$, 取 $\lambda = \mu$ 即可。以下设 $\mu \notin \lambda(A)$, 则 $\mu I - A$ 非奇异。取单位特征向量 x 使 $Bx = \mu x$, 由 $(\mu I - A)x = (B - A)x$ 得

$$1 = \|x\|_2 \leq \|(\mu I - A)^{-1}\|_2 \|B - A\|_2. \quad (2.316)$$

而

$$(\mu I - A)^{-1} = X^{-1}(\mu I - \Lambda)^{-1}X, \quad (2.317)$$

所以

$$\|(\mu I - A)^{-1}\|_2 \leq \|X^{-1}\|_2 \|X\|_2 \max_{\lambda \in \lambda(A)} |\mu - \lambda|^{-1}. \quad (2.318)$$

合并两式并整理, 得到

$$\min_{\lambda \in \lambda(A)} |\mu - \lambda| \leq \kappa_2(X) \|A - B\|_2. \quad (2.319)$$

证毕。这里 X^{-1} 的列组成特征向量基, 且 $\kappa_2(X^{-1}) = \kappa_2(X)$ 。该基接近线性相关时, 特征值可能高度敏感。

💡 Tip

本页要点: Bauer–Fike 界中的特征向量矩阵条件数刻画了谱扰动上界中的放大因子。

2.80 Theorem 2.13: Eigenvalues of A Normal Matrix

定理 2.13 (Krylov – Bogoljubov – Weinstein) . 设 $A \in \mathbb{C}^{n \times n}$ 为正规矩阵, $\|x\|_2 = 1$, 并令 Rayleigh 商

$$\mu(x) = x^H A x. \quad (2.320)$$

则

$$\min_{\lambda \in \lambda(A)} |\lambda - \mu(x)| \leq \|(A - \mu(x)I)x\|_2. \quad (2.321)$$

📌 Note

这个定理说明了什么? 把 $(\mu(x), x)$ 看作一个近似特征对, 残量为 $r = Ax - \mu(x)x$ 。残量为零时, 它就是精确特征对; 对于正规矩阵, 小残量保证 $\mu(x)$ 接近至少一个真实特征值, 距离不超过 $\|r\|_2$ 。几何上, 以 $\mu(x)$ 为中心、以残量范数为半径的闭圆盘至少包含一个特征值。这个界无需预先知道真实特征值即可计算, 但不确定圆盘中是哪一个特征值, 也不直接保证特征向量的方向准确。

谱间隔为什么影响特征向量？ 对所关注的单特征值 λ_k ，谱间隔是它到其余特征值的最小距离 $g = \min_{j \neq k} |\lambda_j - \lambda_k|$ 。两个特征值很接近时，把它们的特征向量混合起来，也可能得到很小的残量，而混合后的方向未必接近其中任何一个。

例如，取 $A = \text{diag}(1, 1 + \varepsilon)$ 、 $\mathbf{x} = (1, 1)^T / \sqrt{2}$ ，其中 $\varepsilon > 0$ 。直接计算得 $\mu = 1 + \varepsilon/2$ ，故 $\mathbf{r} = \varepsilon(-1, 1)^T / (2\sqrt{2})$ ， $\|\mathbf{r}\|_2 = \varepsilon/2$ 。虽然残量随 $\varepsilon \rightarrow 0$ 趋于零， $\mathbf{x}$ 与两个特征向量方向却始终都成 45° 角。原因是谱间隔也等于 ε ，残量与间隔之比始终为 $1/2$ ，并没有变小。

因此，要由残量判断某个单特征值对应的特征向量是否准确，除了确认 μ 接近该特征值，还要看残量相对于谱间隔是否足够小。若目标是重特征值，则应讨论与其特征子空间的接近程度，而不是与其中某个指定特征向量的接近程度。

证明。 由第 2.41 页式 (2.177) 的 Schur 分解推论，正规矩阵 A 可酉对角化，故可写为

$$A = U\Lambda U^H. \quad (2.322)$$

令 $\mathbf{y} = U^H \mathbf{x}$ ，则 $\sum_i |y_i|^2 = 1$ 。记 $\mu = \mu(\mathbf{x})$ ，由酉不变性，

$$\|(A - \mu I)\mathbf{x}\|_2^2 = \sum_i |y_i|^2 |\lambda_i - \mu|^2 \geq \min_i |\lambda_i - \mu|^2. \quad (2.323)$$

两边开平方即得结论。证毕。

取 $\mathbf{x} = \mathbf{e}_i$ 时， $\mu = a_{ii}$ ，残量范数为

$$\|(A - a_{ii}I)\mathbf{e}_i\|_2 = \left(\sum_{j \neq i} |a_{ji}|^2 \right)^{1/2} = \left(\sum_{j \neq i} |a_{ij}|^2 \right)^{1/2}. \quad (2.324)$$

最后一个等号来自 $A^H A = A A^H$ 的第 i 个对角元相等，即第 i 行、列的平方范数相等。因此以 a_{ii} 为中心、上述范数为半径的每个圆盘至少含一个特征值，但这些圆盘的并集不一定包含全部特征值。

① Note

与盖氏圆盘定理有什么区别？ 第 2.20 页定理 2.1 适用于任意方阵，圆心也是 a_{ii} ，半径却是 $R_i = \sum_{j \neq i} |a_{ij}|$ 。它的基本结论是：**每个特征值至少落在一个盖氏圆盘内**，因此圆盘并集包含全部特征值；这个包含结论本身不保证每个圆盘都含有特征值。

本页要求 A 为正规矩阵，半径为 $\rho_i = (\sum_{j \neq i} |a_{ij}|^2)^{1/2}$ ，满足 $\rho_i \leq R_i$ 。它保证的是：**每个这样的较小圆盘至少含有一个特征值**。不同圆盘可能包含同一个特征值，所以“每个圆盘都含有特征值”不能推出“所有特征值都在这些圆盘中”。

例如，取实对称矩阵 $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ 。向量 $(1, 1, 1)^T$ 对应特征值 2；对满足 $x_1 + x_2 + x_3 = 0$ 的向量，有

$A\mathbf{x} = -\mathbf{x}$ ，这个二维子空间对应特征值 -1 。因此三个特征值为 2, -1 , -1 。本页的三个圆盘完全重合，都是 $|z| \leq \sqrt{2}$ ，其中含有 -1 ，却漏掉了 2；盖氏圆盘则都是 $|z| \leq 2$ ，包含全部特征值。

💡 Tip

本页要点： 对正规矩阵，单位向量的残量范数给出其 Rayleigh 商到最近特征值的距离上界。

2.81 Theorem 2.14: Hoffman – Wielandt Theorem

定理 2.14. 若 $A, B \in \mathbb{C}^{n \times n}$ 都是正规矩阵, 谱分别为 $\{\lambda_i\}$ 和 $\{\mu_i\}$, 则存在排列 π 使

$$\sum_{i=1}^n |\lambda_i - \mu_{\pi(i)}|^2 \leq \|A - B\|_F^2. \quad (2.325)$$

证明. 由第 2.41 页式 (2.177) 的 Schur 分解推论, 正规矩阵 A, B 均可酉对角化, 故可写为

$$A = U\Lambda U^H, \quad B = VMV^H. \quad (2.326)$$

令 $W = U^H V$. Frobenius 范数酉不变, 所以

$$\|A - B\|_F = \|\Lambda - WMW^H\|_F. \quad (2.327)$$

置 $s_{ij} = |w_{ij}|^2$. W 的每行每列都是单位向量, 故

$$s_{ij} \geq 0, \quad \sum_j s_{ij} = 1, \quad \sum_i s_{ij} = 1. \quad (2.328)$$

即 $S = (s_{ij})$ 双随机。展开平方并使用行列和:

$$\begin{aligned} \|A - B\|_F^2 &= \sum_i |\lambda_i|^2 + \sum_j |\mu_j|^2 - 2 \operatorname{Re} \sum_{i,j} \bar{\lambda}_i \mu_j s_{ij} \\ &= \sum_{i,j} s_{ij} |\lambda_i - \mu_j|^2. \end{aligned} \quad (2.329)$$

接下来只需一个配对引理: 对任意双随机矩阵 $S = (s_{ij})$ 和实数 c_{ij} , 存在排列 π , 使

$$\sum_{i=1}^n c_{i, \pi(i)} \leq \sum_{i,j} s_{ij} c_{ij}. \quad (2.330)$$

引理的初等证明见下面的备注。取 $c_{ij} = |\lambda_i - \mu_j|^2$, 结合式 (2.329), 立即得到

$$\sum_{i=1}^n |\lambda_i - \mu_{\pi(i)}|^2 \leq \sum_{i,j} s_{ij} |\lambda_i - \mu_j|^2 = \|A - B\|_F^2. \quad (2.331)$$

证毕。

④ Note

配对引理的证明。 思路是逐步调整 S , 保持行列和为 1, 使加权和 $L(S) = \sum_{i,j} c_{ij} s_{ij}$ 不增加, 同时减少正元素, 直到得到排列矩阵。

(1) 找到可调整的闭合回路。 把每一行和每一列各画成一个点; 若 $s_{ij} > 0$, 就连接第 i 行与第 j 列。一个点若只有一条连线, 该元素必为 1, 其另一端也没有其它连线。去掉这些已经配好的行列; 若还有剩余, 每个点至少有两点连线。沿连线前进且不立即原路返回, 由于点数有限, 必会遇到重复的点, 从而找到一个闭合回路。回路在行、列之间交替, 故连线数为偶数。

(2) 沿回路交替加减。 在回路对应的矩阵位置交替标 $+1, -1$ ，其它位置标 0 ，组成矩阵 $D = (d_{ij})$ 。每个涉及的行、列恰有一个正号和一个负号，所以 $S + tD$ 的行列和不变。必要时把全部符号反转，便可使 $L(D) \leq 0$ 。取 $t = \min_{d_{ij}=-1} s_{ij} > 0$ ，则 $S + tD$ 仍非负，且 $L(S + tD) = L(S) + tL(D) \leq L(S)$ 。至少一个被减的正元素变成零；被加的位置原本也为正，因此正元素总数严格减少。

(3) 有限步完成配对。 只要矩阵还不是排列矩阵，就能重复上述调整。正元素数每次减少，故有限步后得到某个排列矩阵 P ，且 $L(P) \leq L(S)$ 。若 $P_{i,\pi(i)} = 1$ ，就有 $L(P) = \sum_i c_{i,\pi(i)}$ ，引理得证。

若 A, B Hermitian 且特征值降序排列，重排不等式说明最优匹配保持次序，故

$$\sum_i (\lambda_i - \mu_i)^2 \leq \|A - B\|_F^2. \quad (2.332)$$

矩形矩阵的奇异值版本。 设 $A, B \in \mathbb{C}^{m \times n}$ ，记 $p = \min(m, n)$ ，两者的奇异值均按降序排列。分别记 $\sigma_i = \sigma_i(A)$ 、 $\tau_i = \sigma_i(B)$ ，下述推导对 $m \geq n$ 和 $m < n$ 均成立。把矩形矩阵嵌入 Hermite 矩阵，即可将奇异值问题转化为特征值问题。

(1) 构造 Hermite 扩张。 定义

$$H(A) = \begin{bmatrix} 0_{m \times m} & A \\ A^H & 0_{n \times n} \end{bmatrix}, \quad H(B) = \begin{bmatrix} 0_{m \times m} & B \\ B^H & 0_{n \times n} \end{bmatrix}. \quad (2.333)$$

它们都是 $(m+n) \times (m+n)$ Hermite 矩阵，可应用式 (2.332)。这里 $H(A)$ 表示新构造的矩阵，上标 H 仍表示共轭转置。

(2) 说明扩张矩阵的特征值。 由第 2.42 页的完整 SVD 及式 (2.182)，左右奇异向量满足 $Av_i = \sigma_i(A)u_i$ 、 $A^H u_i = \sigma_i(A)v_i$ 。因此，对 $1 \leq i \leq p$ ，

$$\begin{aligned} H(A) \frac{1}{\sqrt{2}} \begin{bmatrix} u_i \\ \pm v_i \end{bmatrix} &= \frac{1}{\sqrt{2}} \begin{bmatrix} \pm Av_i \\ A^H u_i \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} \pm \sigma_i(A) u_i \\ \sigma_i(A) v_i \end{bmatrix} \\ &= \pm \sigma_i(A) \frac{1}{\sqrt{2}} \begin{bmatrix} u_i \\ \pm v_i \end{bmatrix}. \end{aligned} \quad (2.334)$$

这说明每个奇异值 $\sigma_i(A)$ 都给出一对特征值 $+\sigma_i(A)$ 和 $-\sigma_i(A)$ ，对应的 $2p$ 个向量标准正交。若 $m > n$ ，完整 SVD 中剩余的 u_i 满足 $A^H u_i = 0$ ，故 $\begin{pmatrix} u_i \\ 0 \end{pmatrix}$ 是零特征向量；若 $n > m$ ，则剩余的 v_i 满足 $Av_i = 0$ ，故 $\begin{pmatrix} 0 \\ v_i \end{pmatrix}$ 是零特征向量。这样共得到 $m+n$ 个标准正交特征向量。

所以 $H(A)$ 的特征值按降序排列为

$$\sigma_1(A), \dots, \sigma_p(A), \underbrace{0, \dots, 0}_{|m-n| \text{ 个}}, -\sigma_p(A), \dots, -\sigma_1(A). \quad (2.335)$$

$H(B)$ 的特征值同理。这个列表也包含秩亏情形：若某个奇异值为零，对应的正、负两项都为零；中间的 $|m-n|$ 个零只是长方形尺寸额外带来的零。

(3) 代入特征值扰动界。 两组特征值按上述次序配对，正、负两部分给出相同的平方差，中间的零互相抵消，因而

$$\begin{aligned}
& \sum_{j=1}^{m+n} \left(\lambda_j(H(A)) - \lambda_j(H(B)) \right)^2 \\
&= \sum_{i=1}^p \left(\sigma_i(A) - \sigma_i(B) \right)^2 + \sum_{i=1}^p \left(-\sigma_i(A) + \sigma_i(B) \right)^2 \\
&= 2 \sum_{i=1}^p \left(\sigma_i(A) - \sigma_i(B) \right)^2.
\end{aligned} \tag{2.336}$$

扩张矩阵的两个非零块分别为 $A - B$ 及其共轭转置，故

$$\begin{aligned}
\|H(A) - H(B)\|_F^2 &= \|A - B\|_F^2 + \|(A - B)^H\|_F^2 \\
&= 2\|A - B\|_F^2.
\end{aligned} \tag{2.337}$$

把式 (2.336) 和 (2.337) 代入式 (2.332)，再约去两边的因子 2，得到

$$\sum_{i=1}^p (\sigma_i(A) - \sigma_i(B))^2 \leq \|A - B\|_F^2. \tag{2.338}$$

当 $m \geq n$ 时，上式即为 $\sum_{i=1}^n (\sigma_i - \tau_i)^2 \leq \|A - B\|_F^2$ 。原矩阵不必是方阵，也不需要满足正规性。

💡 Tip

本页要点： Hoffman–Wielandt 定理用 Frobenius 范数控制正规矩阵特征值的整体移动。

2.82 Theorem 2.15: Courant – Fischer Theorem

设 A Hermitian，特征值按

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \tag{2.339}$$

排列，Rayleigh quotient 为

$$R_A(x) = \frac{x^H A x}{x^H x}. \tag{2.340}$$

定理 2.15. 对 $k = 1, \dots, n$,

$$\lambda_k = \max_{\dim \mathcal{S}=k} \min_{x \in \mathcal{S} \setminus \{0\}} R_A(x) \tag{2.341}$$

且

$$\lambda_k = \min_{\dim \mathcal{T}=n-k+1} \max_{x \in \mathcal{T} \setminus \{0\}} R_A(x). \tag{2.342}$$

特别地，

$$\lambda_1 = \max_{x \neq 0} R_A(x), \quad \lambda_n = \min_{x \neq 0} R_A(x). \tag{2.343}$$

完整证明见 [02线性代数基础.pdf](#) 第 2.83 页。

💡 Tip

本页要点： Courant–Fischer 原理把特征值写成 Rayleigh 商的极值。

2.83 Proof of Courant – Fischer Theorem

Theorem 2.15 的证明. Hermitian 矩阵 A 也是正规矩阵。由第 2.41 页式 (2.177) 的 Schur 分解推论，可取一组标准正交特征向量 $\mathbf{v}_i$ ，满足

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i. \quad (2.344)$$

先取

$$\mathcal{S}_k = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}. \quad (2.345)$$

任意非零 $\mathbf{x} = \sum_{i=1}^k c_i \mathbf{v}_i$ 满足

$$R_A(\mathbf{x}) = \frac{\sum_{i=1}^k \lambda_i |c_i|^2}{\sum_{i=1}^k |c_i|^2} \geq \lambda_k. \quad (2.346)$$

取 $\mathbf{x} = \mathbf{v}_k$ 达到等号，所以

$$\max_{\substack{\dim \mathcal{S}=k \\ \mathbf{x} \in \mathcal{S} \\ \mathbf{x} \neq 0}} \min R_A(\mathbf{x}) \geq \lambda_k. \quad (2.347)$$

反向地，任意 k 维子空间 $\mathcal{S}$ 与

$$\mathcal{T}_k = \text{span}\{\mathbf{v}_k, \dots, \mathbf{v}_n\} \quad (2.348)$$

必有非零交，因为两者维数和为 $k + (n - k + 1) = n + 1$ 。交中向量只含 $\mathbf{v}_k, \dots, \mathbf{v}_n$ 分量，故其 Rayleigh 商不超过 λ_k 。因此每个 $\mathcal{S}$ 都有

$$\min_{\substack{\mathbf{x} \in \mathcal{S} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) \leq \lambda_k, \quad (2.349)$$

第一式得证。

对第二式，取上述 $(n - k + 1)$ 维空间 $\mathcal{T}_k$ 。其内任意非零 $\mathbf{x}$ 有 $R_A(\mathbf{x}) \leq \lambda_k$ ，且 $\mathbf{v}_k$ 达到等号，所以

$$\min_{\dim \mathcal{T}=n-k+1} \max_{\substack{\mathbf{x} \in \mathcal{T} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) \leq \lambda_k. \quad (2.350)$$

反之，任意 $(n - k + 1)$ 维子空间 $\mathcal{T}$ 与 $\mathcal{S}_k$ 的维数和为 $n + 1$ ，故有非零交向量。该向量的 Rayleigh 商不小于 λ_k ，所以任意 $\mathcal{T}$ 上的最大值都不小于 λ_k 。第二式得证。证毕。

💡 Tip

本页要点： Courant – Fischer 证明把特征值极值转化为子空间上的 Rayleigh 商极值。

2.84 Theorem 2.16: Weyl Theorem

先给出子空间压缩的特征值分离。若 A Hermitian, $U \in \mathbb{C}^{n \times (n-1)}$ 满足 $U^H U = I$, 并令 $C = U^H A U$, 则

$$\lambda_1(A) \geq \lambda_1(C) \geq \lambda_2(A) \geq \cdots \geq \lambda_{n-1}(C) \geq \lambda_n(A). \quad (2.351)$$

令 $\mathcal{U} = \text{range}(U)$ 。由 $U^H U = I$, 映射 $\mathbf{y} \mapsto U\mathbf{y}$ 保持长度, 且 $R_C(\mathbf{y}) = R_A(U\mathbf{y})$ 。因此 C 的 Rayleigh 商就是 A 限制在 $\mathcal{U}$ 上的 Rayleigh 商。将 Theorem 2.15 (第 2.82—2.83 页) 的最大值范围限制在 $\mathcal{U}$ 内, 得到

$$\lambda_k(C) = \max_{\substack{\mathcal{S} \subseteq \mathcal{U} \\ \dim \mathcal{S} = k}} \min_{\substack{\mathbf{x} \in \mathcal{S} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) \leq \max_{\substack{\mathcal{S} \subseteq \mathbb{C}^n \\ \dim \mathcal{S} = k}} \min_{\substack{\mathbf{x} \in \mathcal{S} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) = \lambda_k(A). \quad (2.352)$$

再将第二种 min--max 形式的最小值范围限制在 $\mathcal{U}$ 内, 得到

$$\lambda_k(C) = \min_{\substack{\mathcal{T} \subseteq \mathcal{U} \\ \dim \mathcal{T} = n-k}} \max_{\substack{\mathbf{x} \in \mathcal{T} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) \geq \min_{\substack{\mathcal{T} \subseteq \mathbb{C}^n \\ \dim \mathcal{T} = n-k}} \max_{\substack{\mathbf{x} \in \mathcal{T} \\ \mathbf{x} \neq 0}} R_A(\mathbf{x}) = \lambda_{k+1}(A). \quad (2.353)$$

合并即有 $\lambda_k(A) \geq \lambda_k(C) \geq \lambda_{k+1}(A)$, 对 $k = 1, \dots, n-1$ 串联便得到上述分离关系。

定理 2.16. 若 A, B Hermitian, 特征值分别降序排列为 λ_i 和 μ_i , 则

$$\lambda_i + \lambda_{\min}(B - A) \leq \mu_i \leq \lambda_i + \lambda_{\max}(B - A), \quad (2.354)$$

从而

$$|\mu_i - \lambda_i| \leq \|B - A\|_2. \quad (2.355)$$

证明. 令 $E = B - A$ 。由 $R_B = R_A + R_E$ 和 $\lambda_{\min}(E) \leq R_E(\mathbf{x}) \leq \lambda_{\max}(E)$, 对任意非零 $\mathbf{x}$,

$$R_A(\mathbf{x}) + \lambda_{\min}(E) \leq R_B(\mathbf{x}) \leq R_A(\mathbf{x}) + \lambda_{\max}(E). \quad (2.356)$$

对上侧不等式使用 Theorem 2.15 的第一种 min--max 形式 (02线性代数基础.pdf 第 2.82—2.83 页):

$$\begin{aligned} \mu_i &= \max_{\dim \mathcal{S} = i} \min_{\mathbf{x} \in \mathcal{S} \setminus \{0\}} R_B(\mathbf{x}) \\ &\leq \max_{\dim \mathcal{S} = i} \min_{\mathbf{x} \in \mathcal{S} \setminus \{0\}} (R_A(\mathbf{x}) + \lambda_{\max}(E)) \\ &= \lambda_i + \lambda_{\max}(E). \end{aligned} \quad (2.357)$$

同一公式用于下侧不等式得到

$$\mu_i \geq \lambda_i + \lambda_{\min}(E). \quad (2.358)$$

最后, Hermitian 矩阵满足 $\|E\|_2 = \max\{|\lambda_{\min}(E)|, |\lambda_{\max}(E)|\}$, 故 $|\mu_i - \lambda_i| \leq \|E\|_2$ 。证毕。

Tip

本页要点: Weyl 不等式把 Hermitian 矩阵的特征值移动限制在扰动范数内。

本章小结

Important

第 2 章回顾

1. **从子空间理解方程。** 值域决定右端是否可达, 零空间决定解有哪些自由度; 四个基本子空间的正交关

系把存在性、唯一性与最小二乘联系起来。正交投影就是把向量分成可表示的部分与无法表示的正交残差。

2. **区分长度、放大与渐近行为。** 范数给出误差的尺度，诱导矩阵范数衡量一次作用的最大放大，谱半径控制矩阵幂的渐近增长。Gershgorin 圆盘与 Perron - Frobenius 理论则利用矩阵结构提供谱信息。
3. **正交变换保持二范数。** Householder 与 Givens 在消去元素时保持二范数，Schur 分解用酉相似把方阵化为上三角。对角 EVD 需要特征向量基，SVD 对任意长方矩阵都成立；特征值描述不变方向，奇异值描述输入到输出的伸缩。
4. **SVD 可用于多类问题。** 奇异向量给出值域、零空间及其投影，伪逆给出最小范数最小二乘解，截断 SVD 给出最佳低秩近似。数值秩还取决于阈值与噪声尺度。
5. **把求解精度与问题敏感性分开。** 对非奇异矩阵， $\kappa_2(A) = \sigma_{\max}(A)/\sigma_{\min}(A)$ 。很小的奇异值意味着某些输入方向被强烈压缩，求逆时便会放大扰动；条件数描述这种固有敏感性。
6. **使用谱扰动界时要核对前提。** 对可对角化矩阵，Bauer - Fike 界中的特征向量矩阵条件数解释了非正规谱为何敏感；正规矩阵有更好的谱距离控制。对 Hermitian 矩阵，可由 Courant - Fischer 的子空间极值刻画推导 Weyl 不等式，用扰动范数控制特征值的移动。