1. Eigenvalues and eigenvectors

Definition 1 设 A 是复数域 $\mathbb{C}$ 上的 n 阶矩阵，若存在数 $\lambda \in \mathbb{C}$ 和非零的 n 维向量 x 使得：

$$Ax = \lambda x$$

就称 λ 为矩阵 A 的特征值，x 是 A 对应 λ 的特征向量

Definition 2 设 n 阶矩阵 $A = (a_{ij})$ 则

$$f(\lambda) = \det(\lambda I - A)$$

$$= \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为 A 的特征多项式，$\lambda I - A$ 称为 A 的特征矩阵，

$$\det(\lambda I - A) = 0$$

称为 A 的特征方程

Theorem 1 (Rank-Nullity Theorem 秩-零化度定理) 设 $A \in \mathbb{F}^{m\times n}$ 定义零空间 $N(A) = \{x \in \mathbb{F}^n \mid Ax = 0\}$，则有：

$$\operatorname{rank}(A) + \dim(N(A)) = n$$

Theorem 2 下面两式等价：

(a). λ 是 A 的特征值，α 是 A 关于 λ 的特征向量

(b). λ 满足 $|\lambda I - A| = 0$，α 满足 $\alpha \neq 0$ $(\lambda I - A)\alpha = 0$

Proposition 1 (性质) λ 是 A 的特征值，$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A 对应 λ 的 m 个特征向量：

(a) $\forall a, b \in \mathbb{F}$，$b\lambda + a$ 是 $bA + aI$ 的特征值
$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 $bA+aI$ 关于 $b\lambda+a$ 的特征向量

(b) $\forall k \in \mathbb{N}_+$，$\lambda^k$ 是 A^k 的特征值，$\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A^k 关于 λ^k 的特征向量

(c) 如果 A 可逆，$\lambda \neq 0$. 则 λ^{-1} 是 A^{-1} 的特征值 $\{\alpha_i \mid i=1,2,\cdots,m\}$ 是 A^{-1} 关于 λ^{-1} 的特征向量

(d). λ 也是 A^T 的特征值

(e). 任意 α_i 的非零线性组合都是 A 关于 λ 的特征向量

性质

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

$\operatorname{rank}(A) \geq$ A 的非零特征值个数

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$$

Theorem 3 (Linear Independence of eigenvectors) $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是矩阵 A 的 s 个不同的特征值，$\{\alpha_{ij} \mid j=1,2,\cdots,r_i\}$ 是 A 关于 λ_i 的 r_i 个线性无关的特征向量。则整体集 $\{\alpha_{ij} \mid j=1,2,\cdots,r_i,\ i=1,2,\cdots,s\}$ 也是线性无关的。

$\Rightarrow$ 对每一个特征值，它的特征向量是线性无关的

$\Rightarrow$ 对 A 的不同的特征值 λ_i，取 λ_i 的相互独立的特征向量

它们整体仍然线性独立

2. Diagonalization of Square Matrices

Definition 3 (Diagonalization 对角化) 如果 $A = PDP^{-1}$ 其中 P 可逆，D 为对角阵，则称 A 对角化

Theorem 4 A可以被对角化 当且仅当 $A \in \mathbb{F}^{n\times n}$，且 A 有 n 个线性无关的特征向量

Corollary 1 已知 $A \in \mathbb{F}^{n\times n}$，若 A 有 n 个互不相同的特征值，则 A 可以被对角化 （不可反推） $\Rightarrow$ $\cancel{\Leftarrow}$

Remark： 考虑重数

Theorem 5 已知 $\{\lambda_i : i=1,2,\cdots,s\}$ 是 A 的 s 个不同的特征值，用

$$V_i = \{ x \in \mathbb{F}^n \mid (\lambda_i I - A)x = 0 \}$$ （可以理解为解空间）

表示 λ_i 的特征空间. $\dim(V_i)$ 称为 λ_i $(i=1,2,\cdots,s)$ 的几何重数 (geometric multiplicity). A 可对角化，当且仅当 $\sum_{i=1}^{s}\dim(V_i) = n$

$$\boxed{A\text{可对角化} \iff \sum_{i=1}^{s}\dim(V_i) = n}$$

Remark: $(\lambda_i I - A)x = 0$ 的 x 有无穷，将它们集合得到特征空间，
目标，找到 V_i 的基，从而表示整个空间
($\dim(V_i)$ 即基个数)

Theorem 6 设不同的特征值 λ_i ($i=1,2,\cdots,s$)，则**特征多项式** (characteristic polynomial) $|\lambda I - A|$ 可表示为：

$$|\lambda I - A| = \prod_{i=1}^{s} (\lambda - \lambda_i)^{r_i}$$

我们称 r_i 为 λ_i 的**代数重数** (algebraic multiplicity)

(a) 每个特征值对应的几何重数 ≤ 代数重数

$$\dim(V_i) \le r_i \quad (i=1,2,\cdots,s)$$

(b) A 可对角化，当且仅当 几何重数 = 代数重数

A is diagonalizable

$$\Leftrightarrow \forall i \quad \dim(V_i) = r_i$$

(c) 如果 A 是可对角化的，取 A 的每个 λ_i 的特征空间的基，它们共同构成 $\mathbb{F}^n$ 的一组基

Corollary 2 A 可对角化

$$\Leftrightarrow \quad r_i + \operatorname{rank}(\lambda_i I - A) = n \quad i=1,2,\cdots,s$$

3. Eigenvalues and eigenvectors of real symmetric matrices

实对称矩阵

①首先考虑，复数域上的矩阵：

Definition 3 如果矩阵 $\overline{A}$ 满足 $\overline{A}[i,j] = \overline{A[i,j]}$，则称 $\overline{A}$ 为 A 的共轭矩阵 (complex conjugate)

也即 $A = (x_{ij})$ $\overline{A} = (\overline{x_{ij}})$

Proposition 1 (propositions of complex conjugate matrices)

(a). $\overline{A+B} = \overline{A} + \overline{B}$

$\overline{kA} = \overline{k} \cdot \overline{A}$

$\overline{AB} = \overline{A}\,\overline{B}$

$\overline{A^T} = (\overline{A})^T$

$|\overline{A}| = \overline{|A|}$

(b) 如果 A 可逆，则 $\overline{A^{-1}} = (\overline{A})^{-1}$

(c) A 是实矩阵 $\Leftrightarrow$ $\overline{A} = A$

②下面，我们仅考虑实矩阵 $A \in \mathbb{R}^{n\times n}$

Lemma 1 已知 λ 是 A 的特征值，x 是 A 关于 λ 的特征向量

则 $\overline{\lambda}$ 也是 A 的特征值，$\overline{x}$ 是 A 关于 $\overline{\lambda}$ 的特征向量

Theorem 7 已知 A 是实对称矩阵（RSM: real symmetric matrix）则 A 的所有特征值均为实数.

$$A \in \mathbb{R}^{n\times n} \text{ and } A^T = A \quad \Rightarrow \quad \lambda_i \in \mathbb{R}$$

Theorem 8 已知 A 是实对称矩阵，则 A 的不同特征根的特征向量是相互正交的（orthogonal）

4. Orthogonal Diagonalization of RSM

① 正交对角化

Definition 4 如果方阵 Q 满足 $Q^TQ = QQ^T = I$，则称方阵 Q 是正交阵（orthogonal matrix）

☆ **Theorem 9** A 是实对称矩阵，等价于存在正交阵 Q，使得 Q^TAQ 是一个对角阵。我们称 A 可正交对角化：

A 是实对称阵

$$\Leftrightarrow \quad A = QDQ^T$$

其中 Q 为正交阵（$QQ^T = I$），D 为对角阵

Theorem 10 (Spectral Theorem 谱定理)

如果 A 是实对称矩阵，则有：

(a). A 有 n 个特征根（包含重根）

(b). A 特征根的几何重数等于代数重数

(c). A 的不同特征根对应的特征向量相互正交

(d). A 可以正交对角化 Orthogonally Diagonalizable

② 构造正交阵

Theorem 11 (Gram - Schmidt process) 假设 V 是 $\mathbb{R}^n$ 的子空间，$\{\alpha_i \mid i=1,2,\cdots s\}$ 是 V 的一组基。记 $\eta_1 = \alpha_1$，从及：

$$\eta_i = \alpha_i - \sum_{j=1}^{i-1} (\eta_j^T \eta_j)^{-1} (\alpha_i^T \eta_j)\, \eta_j \quad (i=1,2,\cdots,n)$$

如此，$\{\eta_i : i=1,2,\cdots,n\}$ 是 V 的一组正交基。进一步，我们有：

$$\left\{\sum_{i=1}^{t} k_i\alpha_i \;\middle|\; k_i \in \mathbb{R},\ i=1,2,\cdots,t\right\} = \left\{\sum_{i=1}^{t} k_i\eta_i \;\middle|\; k_i \in \mathbb{R},\ i=1,2,\cdots,t\right\}$$

根据 Theorem 11 构造 $A = QDQ^T$ 的 Q 和 D：

(a). 解 $|\lambda I - A| = 0$ 得 $\{\lambda_i^* : i=1,2,\cdots,s\}$，以及对应的代数重数 $\{r_i : i=1,2,\cdots,s\}$

(b). 对 $\forall\, i=1,2,\cdots,s$ 找一组 $V_i = \{x \in \mathbb{F}^n \mid (\lambda_i^* I - A)x = 0\}$ 解空间的基 $\{\alpha_{ij}^* \mid j=1,2,\cdots,r_i\}$

(c). 将 $\{\alpha_{ij}^*\}$ 转为正交基 $\{\eta_{ij}\}$

(d). 令 $Q = (\eta_{11}, \cdots \eta_{1r_1}, \cdots, \eta_{s1}, \cdots, \eta_{sr_s})$

$D = \mathrm{diag}(\lambda_1^* I_{r_1}, \lambda_2^* I_{r_2}, \cdots, \lambda_s^* I_{r_s})$

如此 $A = QDQ^T$

Definition 5 (Spectral Decomposition) 记 $Q = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ $D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

基于正交对角化，我们有 spectral decomposition formula of A:

$$A = QDQ^T = \sum_{i=1}^{n} \lambda_i \alpha_i \alpha_i^T$$

Proposition 3 A 是实对称阵. 记 λ_{min} 和 λ_{max} 为 A 的最小和最大的特征值。对任意，模长为 1 的向量 $x \in \mathbb{R}^n$ ($\|x\| = 1$)，有:

$$\lambda_{min} \leq x^T A x \leq \lambda_{max}$$

5. Definiteness of Matrices

正定. 负定. 半正定. 半负定

Definition 6 如果对 $\forall x \in \mathbb{R}^n$ $x \neq 0$ 均有 $x^T A x > 0$，则称 A 是正定的 (positive definite, PD)， $-A$ 是负定的 (negative definite)

如果对 $\forall x \in \mathbb{R}^n$ $x \neq 0$ 均有 $x^T A x \geqslant 0$ 则称 A 是半正定的 (positive semi-definite, PSD)，$-A$ 是半负定的

Theorem 12

实对称阵是正定的，当且仅当 它的特征值均正

实对称阵是半正定的，当且仅当 它的特征值均非负

6. Square roots of positive semi-definite matrices

半正定阵的平方根

Definition 7 对任意，实矩阵 A，若存在矩阵 C，使 $A = CC$ 则称 C 是 A 的平方根（Square Root）

半正定阵存在平方根

$$A^{\frac{1}{2}} = QD^{\frac{1}{2}}Q^T \qquad D = \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})$$

正定阵存在平方根（和其倒数）

$$A^{-\frac{1}{2}} = QD^{-\frac{1}{2}}Q^T \qquad D = \operatorname{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \cdots, \frac{1}{\sqrt{\lambda_n}}\right)$$

Remark: 半正定阵的平方根不唯一

半正定阵的非负平方根唯一

Theorem 13 Any Positive Semi-Definite Matrix A has a unique PSD square root $A^{\frac{1}{2}} = QD^{\frac{1}{2}}Q^T$

Remark: 以后讨论的均是唯一的 PSD square root