矩阵的特征值与特征向量

定义：设 A 是数域 F 上的 n 阶矩阵，矩阵 $\lambda E - A$ 称为 A 的特征矩阵。

行列式

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

称为 A 的特征多项式。

n 次代数方程 $\left|\lambda E_n - A\right| = 0$ 称为 A 的特征方程.

它的根称为 A 的特征根（或特征值）.

矩阵 A 的所有特征根的全体称为 A 的谱，记为 $\sigma(A)$.

$(\lambda E_n - A)X = 0$ 称为矩阵 A 的特征方程组.

特征值和特征向量的性质

设 $A=\left[a_{ij}\right]_{n\times n}$，易见，它的特征多项式是关于$\lambda$ 的n次多项式，不妨设之为

$$f(\lambda)=\left|\lambda I-A\right|=C_0\lambda^n+C_1\lambda^{n-1}+\cdots+C_{n-1}\lambda+C_n$$

即

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

$$=C_0\lambda^n+C_1\lambda^{n-1}+\cdots+C_{n-1}\lambda+C_n,\ (\Delta 1)$$

考虑上式左端行列式的展开式，它除了

$$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn}) \qquad (\Delta 2)$$

这一项含有 n个形如$(\lambda-a_{ii})$的因式外，其余各项最多含有 $n-2$ 个这样的因式。于是$\lambda^{n},\lambda^{n-1}$只能由$(\Delta 2)$产生。比较$(\Delta 1)$两端的系数，得

$$C_0=1$$

$$C_1=-(a_{11}+a_{22}+\cdots+a_{nn})$$

在式（5.1.5）中，令 $\lambda=0$，得

$$C_n=(-1)^n|A|$$

另外,根据多项式理论，n 次多项式

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

在复数域上有 n个根，不妨设为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，又由于$f(\lambda)$的首项系数 $C_0=1$，于是有

$$f\ \ \lambda\ \ = |\lambda I - A|$$

$$= \ \lambda - \lambda_1 \ \ \lambda - \lambda_2 \ \cdots \ \lambda - \lambda_n$$

$$= \lambda^n - \ \lambda_1 + \cdots + \lambda_n \ \ \lambda^{n-1} + \cdots$$

$$+ \ -1 \ ^n \lambda_1 \lambda_2 \cdots \lambda_n$$

比较 (5.1.5) 和 (5.1.9)， 得

$$C_1 = -\left(\lambda_1 + \cdots + \lambda_n\right)$$

$$C_n = \left(-1\right)^n \lambda_1 \lambda_2 \cdots \lambda_n$$

于是可得特征值的重要性质:

$$(1)\quad |A| = \lambda_1 \lambda_2 \cdots \lambda_n$$

$$(2)\quad \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$$

由 (1) 易见，矩阵A 可逆的充要条件是它的所有特征值都不为零。

性质1：

若 A 的特征值是 λ，X 是 A 的对应于 λ 的特征向量，则

(1) kA 的特征值是 $k\lambda$.（k 是任意常数)

(2) A^m 的特征值是 λ^m.（m 是正整数)

(3) 若 A 可逆，则 A^{-1} 的特征值是 λ^{-1}.

且 X 仍然是矩阵 kA, A^m, A^{-1}

分别对应于 $k\lambda, \lambda^m, \lambda^{-1}$ 的特征向量。

(4) $f(x)$ 为x的多项式，则 $f(A)$ 的特征值为 $f(\lambda)$.

性质2： 矩阵 A 和 A^T 的特征值相同。

几个概念：特征子空间，特征根的代数重复度、几何重复度。

（1） n 阶矩阵 A 的属于特征值 λ_0 的全部特征向量再添上零向量，可以组成 R^n 的一个子空间，称之为矩阵 A 的属于特征值 λ_0 的特征子空间，记为 $\boldsymbol{V}_{\lambda_0}$，不难看出 $\boldsymbol{V}_{\lambda_0}$ 正是特征方程组

$$(\lambda_0 I - A)X = 0$$

的解空间。

(2)设 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 是 A 的 r 个互不相同的特征值，对应的重数分别为 $p_1, p_2,\cdots, p_r$ ，则称 p_i 为 λ_i 的代数重复度。

$$\left|\lambda I - A\right| = (\lambda-\lambda_1)^{p_1}(\lambda-\lambda_2)^{p_2}\cdots(\lambda-\lambda_r)^{p_r}$$

特征子空间 V_{λ_0} 的维数 q_i 为 λ_i 的几何重复度。

$$q_i = n - \mathrm{rank}(\lambda_i I - A)$$

(3)设 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 是 A 的 r 个互不相同的特征值，q_i 是 λ_i 的几何重复度，$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{iq_i}$ 是对应于 λ_i 的 q_i 个线性无关的特征向量，则 A 的所有这些特征向量 $\alpha_{11},\alpha_{12},\cdots,\alpha_{1q_1};\alpha_{21},\alpha_{22},\cdots,\alpha_{2q_2};\cdots;\alpha_{r1},\alpha_{r2},\cdots,\alpha_{rq_r}$ 线性无关。

$\Downarrow$

r 个特征子空间的和是直和 $V_{\lambda_1}\oplus V_{\lambda_2}\cdots\oplus V_{\lambda_r}$

（4）一个特征向量不能属于不同的特征值。

(6) 矩阵 A 的任一特征值 λ_i 的几何重复度 q_i 不大于它的代数重复度 p_i.

证: 特征向量: $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{iq_i}$ 对应于特征值 λ_i

$$\boxed{\alpha_{i1},\alpha_{i2},\cdots,\alpha_{iq_i},\beta_1,\beta_2,\cdots,\beta_{n-q_i}} \longrightarrow R^n \text{ 的一组基}$$

令 $P=[\alpha_{i1},\alpha_{i2},\cdots,\alpha_{iq_i},\beta_1,\beta_2,\cdots,\beta_{n-q_i}]$ 则

$$AP=[A\alpha_{i1},A\alpha_{i2},\cdots,A\alpha_{iq_i},A\beta_1,A\beta_2,\cdots,A\beta_{n-q_i}]$$

$$=[\boxed{\lambda_i\alpha_{i1},\lambda_i\alpha_{i2},\cdots,\lambda_i\alpha_{iq_i},A\beta_1,A\beta_2,\cdots,A\beta_{n-q_i}}]$$

每个列向量都可以由 P 的列向量组线性表出

$$= \underbrace{[\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{iq_i}, \beta_1, \beta_2, \cdots, \beta_{n-q_i}]}_{P} \left[\begin{array}{ccc:c} \lambda_i & & & \\ & \ddots & & * \\ & & \lambda_i & \\ \hdashline & 0 & & A_1 \end{array}\right]$$

$$\Rightarrow \quad P^{-1}AP = \left[\begin{array}{ccc:c} \lambda_i & & & \\ & \ddots & & * \\ & & \lambda_i & \\ \hdashline & 0 & & A_1 \end{array}\right]$$

$$|\lambda E_n - A| = |\lambda E_n - P^{-1}AP| = (\lambda - \lambda_i)^{q_i} \, |\lambda E_{n-q_i} - A_1|$$

特征值 λ_i 的几何重复度 q_i 不大于它的代数重复度 p_i.

矩阵的可对角化条件

n 维线性空间 V上的线性变换 f，是否存在 V 的一个基使得 f 在这个基下的矩阵为对角矩阵？

f 在给定的一个基下的矩阵表示为 A，那么 A 是否相似于一个对角矩阵？

矩阵 A 相似于一个对角矩阵的条件是什么？

定义：若 n 阶矩阵 A与对角形矩阵相似，则称 A可对角化，也称 $\boldsymbol{A}$ 是单纯矩阵。

定理1.10.2 n 阶矩阵 A 可对角化的充要条件是有 n个线性无关的特征向量。

不妨假设 n 阶方阵 A 可相似于对角阵，即存在可逆矩阵 P ，使得

$$P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

或 $$AP = Pdiag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

令 $$P = [X_1, X_2, \cdots, X_n]$$

并将之代入上式，得

$$A[X_1, X_2, \cdots, X_n]$$
$$= [X_1, X_2, \cdots, X_n] \; diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

即 $[AX_1, AX_2, \cdots, AX_n] = [\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n]$

从而有

$$AX_i = \lambda_i X_i \qquad i = 1,2,\cdots,n$$

由 P 可逆知，$X_i \neq 0 (i = 1,2,\cdots,n)$ 且

$$X_1, X_2, \cdots, X_n$$

线性无关从而 $X_1, X_2, \cdots, X_n$ 是 A 的 n 个线性无关的特征向量， $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的 n 个特征值。

反之，若 n 阶方阵 A 有 n 个线性无关的特征向量，不妨设为 $X_1, X_2, \cdots, X_n$，则存在相应的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ ，使得

$$AX_i = \lambda_i X_i \qquad i = 1,2,\cdots,n$$

此时，令

$$P = \left[X_1, X_2, \cdots, X_n\right]$$

显然 P 可逆，且有

$$P^{-1}AP = diag\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$$

即矩阵 A 相似于对角阵.

定理1.10.3 n阶矩阵A可对角化的充要条件是A的每一个特征值的几何重复度等于代数重复度。

$\boldsymbol{q}_i$ 为 $(\lambda_i \boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 基础解系中向量个数，所以

$$\boldsymbol{q}_i=\boldsymbol{n}-\operatorname{rank}(\lambda_i \boldsymbol{I}-\boldsymbol{A}), \quad \boldsymbol{q}_i \leq \boldsymbol{p}_i$$

推论 若$\boldsymbol{n}$阶方阵$\boldsymbol{A}$有$\boldsymbol{n}$个互异的特征值（即特征多项式无重根），则$\boldsymbol{A}$可相似对角化。

同时对角化

引理： 设 $\boldsymbol{A} \in \boldsymbol{C}^{n\times n}, \boldsymbol{B} \in \boldsymbol{C}^{m\times m}$，且 $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}$，

则 C 可以对角化的充要条件是 A, B 都可以对角化。

定理 1.10.7 设 $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{C}^{n\times n}$ 都可以对角化，则 $\boldsymbol{A}, \boldsymbol{B}$ 同时对角化的充要条件是 $\boldsymbol{AB} = \boldsymbol{BA}$.

证明：必要性 若存在 $P\in C_n^{n\times n}$ 使得

$$P^{-1}AP=\text{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

$$P^{-1}BP=\text{diag}(\mu_1,\mu_2,\cdots,\mu_n)$$

$$\longrightarrow \left(P^{-1}AP\right)\left(P^{-1}BP\right)=\text{diag}(\lambda_1,\cdots,\lambda_n)\text{diag}(\mu_1,\cdots,\mu_n)$$

$$=\text{diag}(\mu_1,\cdots,\mu_n)\text{diag}(\lambda_1,\cdots,\lambda_n)=\left(P^{-1}BP\right)\left(P^{-1}AP\right)$$

$$\longrightarrow P^{-1}ABP=P^{-1}BAP \longrightarrow AB=BA.$$

充分性：A, B 都可以对角化，所以存在 $S \in C_n^{n\times n}$ 使得

$$S^{-1}AS = \begin{bmatrix} \lambda_1 E_1 & & & \\ & \lambda_2 E_2 & & \\ & & \ddots & \\ & & & \lambda_h E_h \end{bmatrix} = A$$

其中 $\lambda_1, \lambda_2, \cdots \lambda_h$ 互不相同。

令 $S^{-1}BS = B$, 则由 $AB = BA$, 可得

$$AB = BA$$

将 B 分块，B_{ij} 的行数与 E_i 相同，列数与 E_j 相同.

$$B=\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1h} \\ B_{21} & B_{22} & \cdots & B_{2h} \\ \vdots & \vdots & & \vdots \\ B_{h1} & B_{h2} & \cdots & B_{hh} \end{bmatrix},$$

$$A=\begin{bmatrix} \lambda_1 E_1 & & & \\ & \lambda_2 E_2 & & \\ & & \ddots & \\ & & & \lambda_h E_h \end{bmatrix},$$

$$AB=BA$$

$$\downarrow$$

$$(\lambda_i E_i) B_{ij} = B_{ij} (\lambda_j E_j)$$

$$\because \lambda_i \neq \lambda_j, \quad (i \neq j)$$

$$\downarrow$$

$$B_{ij}=0, \quad (i \neq j)$$

即

$$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_{11} & & & \\ & \boldsymbol{B}_{22} & & \\ & & \ddots & \\ & & & \boldsymbol{B}_{hh}\end{bmatrix}, \qquad \left(=\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S}\right)$$

因为 $\boldsymbol{B}$ 可对角化，从而 $\boldsymbol{B}$ 可对角化，由前面引理知，$\boldsymbol{B}_{ii}, i=1,2,\cdots,h$ 可对角化，即存在

可逆矩阵 $\boldsymbol{T}_i$，使得 $\boldsymbol{T}_i^{-1}\boldsymbol{B}_{ii}\boldsymbol{T}$ 是对角形，$i=1,2,\cdots,h$

令

$$T=\begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_h \end{bmatrix},$$

则 $T^{-1}BT,\ T^{-1}AT$ ，即 $T^{-1}S^{-1}BST,\ T^{-1}S^{-1}AST$ 均为对角形矩阵.

从而，令 $P=ST$，则

$P^{-1}BP,\ P^{-1}AP$ 均为对角形矩阵。