特征值与特征向量 BY 范中平

I 基本概念

1. 定义

设$A \in \mathbf{C}^{n\times n}$，若存在$\lambda \in \mathbf{C}$及非零列向量$X \in \mathbf{C}^n$，使得

$$AX = \lambda X,$$

则称λ为A的一个**特征值**，称X为A的关于λ的一个**特征向量**．（记为$\lambda \sim A$，$X \sim \lambda$）

举例，单位阵，$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$，照镜子，照相，

Google Page Rank:

Google Page Rank 算法的目的是对多个网站的重要性进行排序．以下图所示的网站链接为例．

根据各网站被链接的情况建立排序矩阵A，然后求排序向量X，使得

$$\mathrm{AX} = \mathrm{X},$$

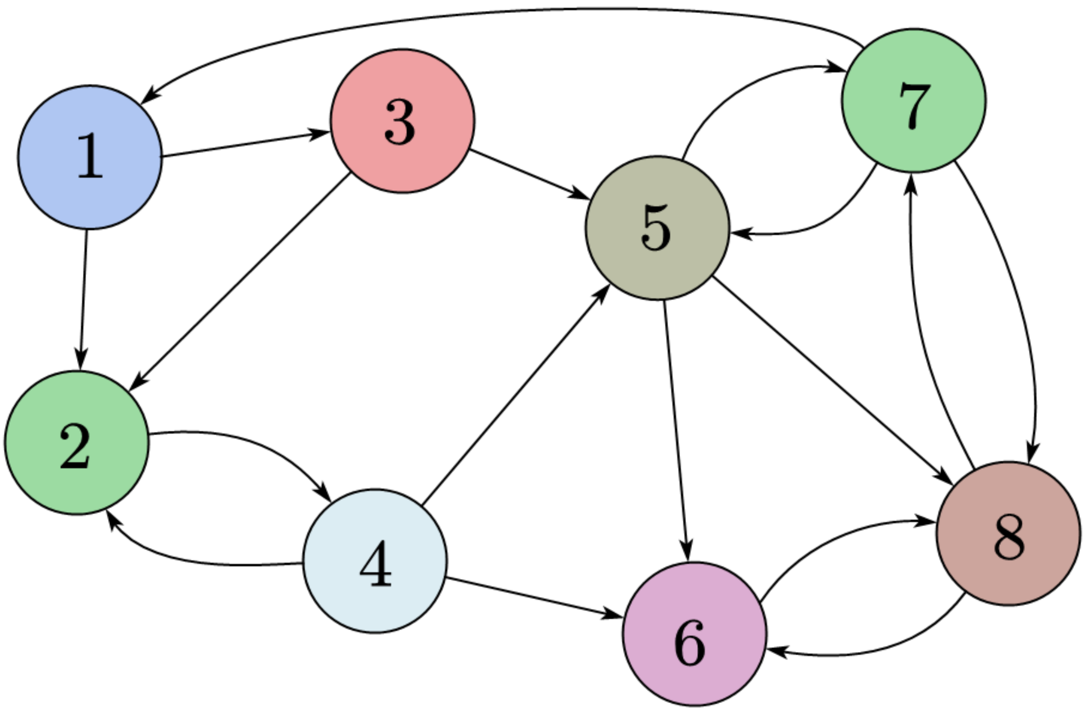
有

$$\mathrm{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \\ 1/2 & 0 & 1/2 & 1/3 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/3 & 1 & 1/3 & 0 \end{pmatrix}, \mathrm{X} = \begin{pmatrix} 0.0600 \\ 0.0675 \\ 0.0300 \\ 0.0675 \\ 0.0975 \\ 0.2025 \\ 0.1800 \\ 0.2950 \end{pmatrix}$$

从而得到，重要性排序

$$\text{网站 } 8 > \text{网站 } 6 > \text{网站 } 7 > \text{网站 } 5 > \text{网站 } 4 = \text{网站 } 2 > \text{网站 } 1 > \text{网站 } 3.$$

不难看出，Google Page Rank 算法的实质就是求排序矩阵A关于特征值1的特征向量X.



2. 意义

在接下来学习的绝大部分时候，讨论的A, λ, X都是实的．

已知任何一个n阶实方阵A都一一对应于一个$\mathbf{R}^n$到$\mathbf{R}^n$的线性变换F_A，也就是对任意的$\alpha \in \mathbf{R}^n$，

$$F_A(\alpha) = A\alpha.$$

若存在非零实向量β，使得β与$A\beta$是平行的(规定零向量与任意向量平行)，也就是存在实数λ，使得

$$A\beta=\lambda\beta,$$

则当$\lambda>0$时，β与$A\beta$同向；当$\lambda<0$时，β与$A\beta$反向．按照定义，λ为A的一个特征值，而β为A的关于λ的一个特征向量．

复数的情形可以看作是实数情形的自然推广，并且在接下来的学习中将会看到，在复数域下讨论保证了特征值的存在性．通过研究特征值与特征向量，可以更加清楚的研究方阵的性质与结构，同时有许多其他方面的数学和物理问题可以归结为研究在某个线性变换下保持方向平行而大小变化的向量，也就是研究某线性变换下的特征值与特征向量．

特征值与特征向量的定义中考虑复矩阵有更广泛的意义．在讨论特征值与特征向量时，可以只考虑实矩阵．但作为实矩阵的推广，复矩阵并不陌生，绝大部分之前学习的在实数情形下成立的结果，都可以推广到复数的情形下．

3．分析

根据特征值与特征向量的定义和线性方程组系数矩阵与解的关系知，

$$\lambda\text{为}A\text{的一个特征值},$$

$$<=>\ \exists X\neq 0,\ \text{使得}AX=\lambda X,$$

$$<=>\ \exists X\neq 0,\ \text{使得}(\lambda I-A)X=0,$$

$$<=>\ (\lambda I-A)\text{不可逆},$$

$$<=>\ (\lambda I-A)\text{的行列式等于}0.$$

由此，

定理 1: (1) λ为A的一个特征值等价于

$$|\lambda I-A|=0;$$

(2) X为A的关于特征值λ的一个特征向量等价于

$$(\lambda I-A)X=0,\qquad X\neq 0.$$

因为$(\lambda I-A)X=0$的几个解向量的线性组合仍为$(\lambda I-A)X=0$的解，由定理 1，就知道，关于同一特征值λ的几个特征向量$X_1,X_2\cdots X_m$，以不全为零的系数$k_1,k_2\cdots k_m$作线性组合，得到的向量仍为特征值λ的特征向量，也就是说：

定理 2: 对$\lambda\sim A$且$X_1,X_2\cdots X_m\sim\lambda$，若系数$k_1,k_2\cdots k_m$不全为0，则

$$k_1X_1+k_2X_2+\cdots+k_mX_2\sim\lambda\sim A.$$

还要注意定义中的几点：

1) $\mathbf{C}^{n\times n}$表示n阶复方阵全体，包含n阶实方阵全体$\mathbf{R}^{n\times n}$；
2) A为n阶复方阵，包含了A为n阶实方阵的情形；
3) A的特征值λ的取值范围为复数域，实方阵的特征值不一定是实数；举例$\begin{pmatrix}0&1\\-1&0\end{pmatrix}$
4) A的关于λ的特征向量X总是非零的复向量，包含了X为非零的实向量的情形；举例
5) A的特征值不一定唯一；举例
6) 特征值λ的特征向量一定不唯一：若$AX=\lambda X$，则$A(kX)=\lambda(kX)$.

II 求法

1．准备

对n阶方阵A，

- 称多项式

$$|\lambda I-A|,$$

为A的**特征多项式**，显然$|\lambda I-A|$为关于λ的n次多项式；一般记$f(\lambda)=|\lambda I-A|$.

- 称方阵

$$(\lambda \mathrm{I}-A)$$

为A的**特征矩阵**;

- 称关于λ的方程

$$|\lambda \mathrm{I}-A|=0$$

为A的**特征方程**，显然$|\lambda \mathrm{I}-A|=0$为关于λ的n次方程**.**

由定理 1 知，λ为A的一个特征值等价于λ为$|\lambda \mathrm{I}-A|=0$的一个根.

由**代数基本定理**知,

在复数域中，一元n次方程一定有n个根.

注意这n个根可以相同，可以是复数．由此λ一定有n个取值满足$|\lambda \mathrm{I}-A|=0$，从而保证了特征值的存在性，也就是

命题 3: n阶方阵A在复数域下一定有n个特征值.

2. 步骤

在方程中，设$x_1, x_2 \cdots, x_t$为n次方程$f(\mathrm{x})=0$的全部的不同根，有

$$f(\mathrm{x})=(x-x_1)^{n_1}(x-x_2)^{n_2}\cdots(x-x_t)^{n_t}$$

且

$$\sum_{i=1}^{t} n_i=n,$$

则对$i=1\cdots t$，称x_i为$f(\mathrm{x})=0$的n_i**重根**.

例如，对方程$\mathrm{x}^2-2x+1=0$，有$\mathrm{x}=1$为方程的2重根.

我们接下来给出求特征值与特征向量的一般方法:

(1) 求特征多项式

$$f(\lambda)=|\lambda \mathrm{I}-A|;$$

(2) 解$f(\lambda)=0$，对$f(\lambda)$作复数域下的因式分解,

$$f(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_t)^{n_t}$$

其中$t\le n$，求出$f(\lambda)=0$的全部的t个不同的根

$$\lambda_1, \lambda_2, \cdots, \lambda_t.$$

其中对$i=1\cdots t$，第i个根λ_i的重数为n_i，称n_j为λ_j的**代数重数**.

(3) 对$i=1\cdots t$，由齐次方程组

$$(\lambda_i \mathrm{I}-A)X=0,$$

求出一个基础解系

$$X_1^{(i)}, \cdots X_{m_i}^{(i)}.$$

由定理 1 知，λ_i的任意特征向量都可以表示为

$$k_1 X_1^{(i)}+\cdots+k_{m_i} X_{m_i}^{(i)},$$

其中$k_1, \cdots k_{m_i}$为任意常数.

对$i=1\cdots t$，称$(\lambda_i \mathrm{I}-A)X=0$的基础解系向量个数$m_i$为特征值$\lambda_i$的**几何重数**.

举例, P224 例 1.

III 性质

1. 重数

对n阶方阵A，设$\lambda_1,\cdots,\lambda_t$为$A$的全部不同的特征值.

已知

(1) λ_i的代数重数n_i为λ_i作为特征方程 $f(\lambda)=0$的根的重数;

(2) λ_i的几何重数m_i为方程$(\lambda_i \mathrm{I}-A)X=0$的基础解系的向量个数.

命题 4: 对n阶方阵A，A的全部不同的特征值$\lambda_1,\cdots,\lambda_t$的代数重数的和为

$$\sum_{i=1}^{t} n_i = n.$$

定理 5: 对n阶方阵A的一个特征值λ_i，有

λ_i的代数重数n_i大于等于λ_i的几何重数m_i. (P235 定理 5.8)

证明：设$(\lambda_i \mathrm{I}-A)X=0$的基础解系为

$$X_1^{(i)},\cdots X_{m_i}^{(i)},$$

将其扩充为$\mathbf{C}^n$的一组基

$$X_1^{(i)},\cdots X_{m_i}^{(i)},Y_{m_i+1},\cdots,Y_n,$$

令方阵

$$P=\left(X_1^{(i)},\cdots X_{m_i}^{(i)},Y_{m_i+1},\cdots,Y_n\right),$$

则P可逆且有

$$AP=\left(A\left(X_1^{(i)},\cdots X_{m_i}^{(i)}\right),A\left(Y_{m_i+1},\cdots,Y_n\right)\right)=\left(\lambda_i\left(X_1^{(i)},\cdots X_{m_i}^{(i)}\right),A\left(Y_{m_i+1},\cdots,Y_n\right)\right)$$

$$=\left(X_1^{(i)},\cdots X_{m_i}^{(i)},Y_{m_i+1},\cdots,Y_n\right)\begin{pmatrix}\lambda_i I_{m_i} & C\\ 0 & D\end{pmatrix},$$

其中

$$\mathrm{P}\begin{pmatrix}C\\ D\end{pmatrix}=A\left(Y_{m_i+1},\cdots,Y_n\right).$$

记分块矩阵

$$B=\begin{pmatrix}\lambda_i I_{m_i} & C\\ 0 & D\end{pmatrix}$$

则有

$$\mathrm{AP}=\mathrm{PB},\mathrm{B}=\mathrm{P}^{-1}\mathrm{AP}$$

计算A的特征多项式有,

$$|\lambda\mathrm{I}-A|=|\lambda\mathrm{I}-PBP^{-1}|=|\lambda\mathrm{I}-B|=(\lambda-\lambda_i)^{m_i}|\lambda\mathrm{I}-D|$$

所以

$$m_i\le n_i.$$

举例，$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$，$\lambda=1$.

2. 运算律

考虑特征值与其他方阵性质和运算间的联系.

设 $\mathrm{X}\sim\lambda\sim\mathrm{A}$和$\mathrm{Y}\sim\mu\sim\mathrm{B}$,

- 与加法，若$\mathrm{X}=\mathrm{Y}$，则$\mathrm{X}\sim(\lambda+\mu)\sim(\mathrm{A}+\mathrm{B})$；否则，无;
- 与数乘,

$$\mathrm{X}\sim \mathrm{k}\lambda\sim \mathrm{kA};$$

- 与幂,

$$X\sim\lambda^m\sim A^m;$$

- 与逆，若A可逆，则$\lambda\neq 0$，且

$$X\sim\lambda^{-1}\sim A^{-1};$$

- 与多项式，对多项式$F(x)$，

$$X\sim F(\lambda)\sim F(A);$$

- 与转置，$\lambda\sim A^T$，但X不一定为A^T的关于λ的特征向量；

证明：设$A^TX=\rho X$，有$|\rho I-A^T|=|\rho I-A|$，即A^T与A的特征值相同.

举例，$\begin{pmatrix}1\\0\end{pmatrix}\sim 0\sim\begin{pmatrix}0&1\\0&0\end{pmatrix}$，虽然$0\sim\begin{pmatrix}0&0\\1&0\end{pmatrix}$，但$\begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$；

设$\lambda_1,\cdots,\lambda_n$为$A=(a_{ij})$的全部n个特征值，(区分于t个不同的特征值)

- 与迹，称

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

为A的**迹**，有

$$\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i,$$

(P226 定理 5.2) 举例

证明：已知

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

考虑等号左右两边$(n-1)$次项的系数，对左边有

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11}&-a_{12}&\cdots&-a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&-a_{2n}\\\vdots&\vdots&\ddots&\vdots\\-a_{m1}&-a_{m2}&\cdots&\lambda-a_{mn}\end{vmatrix},$$

不难验证，其中$(n-1)$次项的系数只能由$\prod_{i=1}^{n}(\lambda-a_{ii})$产生，为$-\sum_{i=1}^{n}a_{ii}$；对右边$\prod_{i=1}^{n}(\lambda-\lambda_i)$，其$(n-1)$次项的系数为$-\sum_{i=1}^{n},\lambda_i$，所以有$|\lambda I-A|=\prod_{i=1}^{n}(\lambda-\lambda_i)$.

- 与行列式，

$$|A|=\prod_{i=1}^{n}\lambda_i.$$

(P226 定理 5.2) 举例

证明：由$|\lambda I-A|=\prod_{i=1}^{n}(\lambda-\lambda_i)$，令$\lambda=0$，则有

$$|-A|=\prod_{i=1}^{n}(-\lambda_i).$$

3. 特征向量

设A为任意n阶方阵，

命题 6: A的相同特征值下特征向量全体和零向量的并集构成一个向量空间.

证明：由定理 2，此集合对向量加法和数乘保持封闭.

命题 7: A的不同特征值下的的特征向量无关.

证明:

用归纳法证明A的t个不同特征值下的的特征向量无关. 设$\lambda_1, \cdots, \lambda_t$为$A$的$t$个不同的特征值，且

$$X_i \sim \lambda_i \sim A, \qquad i = 1, \cdots, t.$$

当$t = 2$时，已知

$$AX_1 = \lambda_1 X_1, \qquad AX_2 = \lambda_2 X_2.$$

设

$$k_1 X_1 + k_2 X_2 = 0,$$

有

$$A(k_1 X_1 + k_2 X_2) = k_1 \lambda_1 X_1 + k_2 \lambda_2 X_2 = A0 = 0,$$

将$k_1 X_1 = -k_2 X_2$代入，得

$$-k_2 \lambda_1 X_2 + k_2 \lambda_2 X_2 = k_2(-\lambda_1 + \lambda_2) X_2 = 0.$$

因为$X_2 \neq 0, -\lambda_1 + \lambda_2 \neq 0$ 所以

$$k_2 = 0,$$

同理，有

$$k_1 = 0.$$

因此，X_1, X_2无关，命题成立.

假设$t = N - 1$时，命题成立；当$t = N$时，设

$$\sum_{i=1}^{N} k_i X_i = 0 \tag{1}$$

将A左乘到式(1)的等号左右两边得,

$$\sum_{i=1}^{N} k_i \lambda_i X_i = 0 \tag{2}$$

将λ_N左乘到式(1)的等号左右两边得,

$$\sum_{i=1}^{N} k_i \lambda_N X_i = 0 \tag{3}$$

式(3)减式(2)得,

$$\sum_{i=1}^{N-1} k_i (\lambda_N - \lambda_i) X_i = 0.$$

由假设知，$X_1, \cdots, X_{N-1}$无关，则对$i = 1, \cdots N - 1$,

$$k_i (\lambda_N - \lambda_i) = 0.$$

因为$\lambda_N - \lambda_i \neq 0$，所以

$$k_i = 0, \qquad i = 1, \cdots N - 1,$$

代入式(1)有

$$k_N = 0,$$

从而有$t = N$时，$X_1, \cdots, X_N$无关. 由归纳法知，命题对任意的正整数t都成立.

定理 8: 对n阶方阵A，设$\lambda_1, \cdots, \lambda_t$为$A$的不同的特征值，且$X_1^{(i)}, \cdots X_{l_i}^{(i)}$为无关的$\lambda_i$的特征向量，$i = 1, \cdots t$. 则向量组

$$\bigcup_{i=1}^{t} \left\{ X_1^{(i)}, \cdots X_{l_i}^{(i)} \right\}$$

是无关的.

证明：设

$$\sum_{i=1}^{t} (k_1^{(i)} X_1^{(i)} + \cdots + k_{l_i}^{(i)} X_{l_i}^{(i)}) = 0.$$

假设$\exists i$，使得

$$k_1^{(i)} X_1^{(i)} + \cdots + k_{l_i}^{(i)} X_{l_i}^{(i)} \neq 0.$$

因为$k_1^{(i)}X_1^{(i)}+\cdots+k_{l_i}^{(i)}X_{l_i}^{(i)}$为关于$\lambda_i$的特征向量，所以

$\sum_{i=1}^{t}(k_1^{(i)}X_1^{(i)}+\cdots+k_{l_i}^{(i)}X_{l_i}^{(i)})$是$A$的不同特征值的特征向量以系数$1$进行的线性组合. 由命题 7 知，向量组

$$\left\{k_1^{(i)}X_1^{(i)}+\cdots+k_{l_i}^{(i)}X_{l_i}^{(i)}\,\middle|\, i=1,\cdots t\right\}$$

是无关的，矛盾.
因此

$$k_1^{(i)}X_1^{(i)}+\cdots+k_{l_i}^{(i)}X_{l_i}^{(i)}=0.$$

由$X_1^{(i)},\cdots X_{l_i}^{(i)}$无关知，

$$k_1^{(i)}=\cdots=k_{l_i}^{(i)}=0.$$

命题得证.