第四章　矩阵特征值与特征向量的计算

第四章 矩阵特征值与特征向量的计算

- 引言
- 幂法
- 反幂法

4.1 引言

定义 设A是n阶矩阵，如果存在数λ和n维非零向量α，使

$$A\alpha=\lambda\alpha$$

则称λ为方阵A的一个特征值，α为方阵A对应于特征值λ的一个特征向量.

1° A的特征值 λ 由它的特征方程 $\varphi(\lambda)=\det(\lambda I-A)=0$ 的根确定.

2° 设λ为A的特征值,求齐次线性方程组 $(\lambda I-A)x=0$ 的非零解，便得到A 的属于λ 的特征向量.

设 $A\alpha=\lambda\alpha\ (\alpha\neq 0)$, 则 $(\lambda I-A)\alpha=0$.

α是 $(\lambda I-A)X=0$ 的非零解.

求A的特征值与特征向量的步骤如下：

(1) 求$|\lambda I-A|=0$的根：$\lambda_1,\lambda_2,\cdots,\lambda_k$;

(2) 求$(\lambda_i I-A)X=0$的基础解系：

$$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_{r_i}},$$

则A对应于λ_i的特征向量为：

$$k_1\alpha_{i_1}+k_2\alpha_{i_2}+\cdots+k_{r_i}\alpha_{r_i}\qquad (k_1,k_2,\cdots,k_{r_i}\ \text{不全为零}).$$

回顾几个基本结论:

定理1 若$\lambda_i(i=1,2,\ldots,n)$为$A$的特征值，则有

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

（2） $$\det(A)=\lambda_1\lambda_2\cdots\cdots\lambda_n$$

定理2 设A与B为相似矩阵，即存在可逆矩阵P，使得$P^{-1}AP=B$，则

(1) A与B有相同的特征值;

(2) 若x为B的一个特征向量，则Px为A的特征向量.

定理3 (Gerschgorin's定理)设$A=(a_{ij})_{n\times n}$，则A的每一个特征值，必属于下述某个圆盘之中：

$$|\lambda - a_{ii}| \le \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|, \quad (i = 1\ldots\ldots n)$$

且如果一个特征向量的第i个分量绝对值最大，则对应的特征值一定属于第i个圆盘中.

例1 估计方阵特征值的范围 $A=\begin{bmatrix} 1 & 0.1 & 0.2 & 0.3 \\ 0.5 & 3 & 0.1 & 0.2 \\ 1 & 0.3 & -1 & 0.5 \\ 0.2 & -0.3 & -0.1 & -4 \end{bmatrix}$

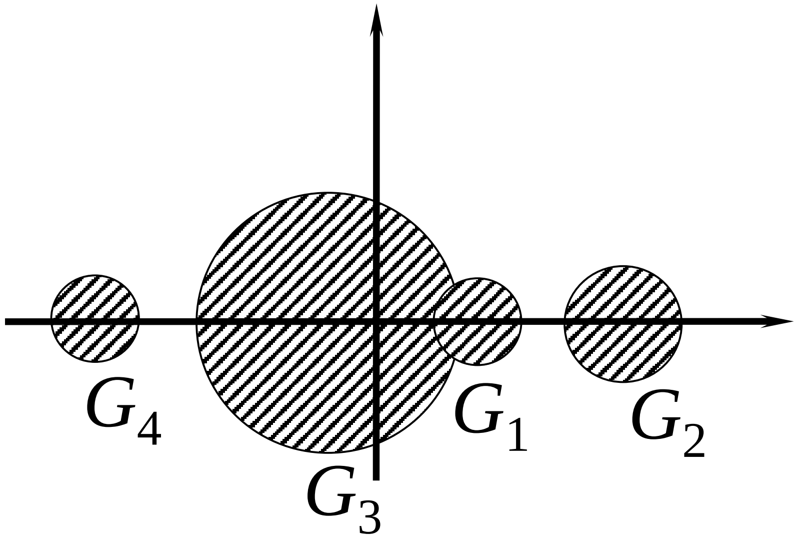
解:

$G_1=\{z: |z-1|\le 0.6\};$

$G2=\{z: |z-3|\le 0.8\};$

$G_3=\{z: |z+1|\le 1.8\};$

$G4=\{z: |z+4|\le 0.6\}.$



注：定理称A的n个特征值全落在n个盖氏圆上，但未说明每个圆盘内都有一个特征值.

4.2 幂法

用于求矩阵的按模最大的特征值与相应的特征向量的近似值.

设A为n阶实矩阵,

$\lambda_i, u_i \quad (i=1,2,\cdots,n)$为$A$的特征值和相应的特征向量,

且满足: $|\lambda_1|>|\lambda_2|\geq|\lambda_3|\geq\cdots\geq|\lambda_n|$

$u_1, u_2, \cdots, u_n$,线性无关.

幂法:求λ_1及其对应的特征向量.

λ_1通常称为主特征值.

幂法基本思想：

给定初始非零向量$x^{(0)}$，由矩阵A构造一向量序列

$$\begin{cases} x^{(1)} = Ax^{(0)} \\ x^{(2)} = Ax^{(1)} = A^2 x^{(0)} \\ \cdots \\ x^{(k+1)} = Ax^{(k)} = A^{k+1} x^{(0)} \\ \cdots \end{cases}$$

在一定条件下，当k充分大时： $\lambda_1 \approx \dfrac{x_i^{(k+1)}}{x_i^{(k)}}$

相应的特征向量为： $x^{(k+1)}$

幂法的理论依据：

对任意向量$x^{(0)}$，有 $x^{(0)}=\sum_{i=1}^{n} t_i u_i$， 设$t_1$不为零.

$$x^{(k+1)}=Ax^{(k)}=A^{k+1}x^{(0)}$$

$$=\sum_{i=1}^{n} A^{k+1} t_i u_i=\sum_{i=1}^{n} t_i \lambda_i^{k+1} u_i$$

$$=\lambda_1^{k+1}\left[t_1 u_1+\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1} t_2 u_2+\cdots+\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} t_n u_n\right]$$

$$\approx \lambda_1^{k+1} t_1 u_1$$

故 $\lambda_1 \approx x_i^{(k+1)} \big/ x_i^{(k)}$

$x^{(k+1)}$为λ_1的特征向量的近似向量(除一个因子外).

- 如果$x^{(0)}$的选取恰恰使得$t_1=0$，幂法仍能进行. 因为计算过程中会有舍入误差, 迭代若干次后, 必然会产生一个向量$x^{(k)}$, 它在u_1方向上的分量不为零, 这样以后的计算就满足所设条件.

- 因为 $x^{(k)} \approx \lambda_1^k t_1 u_1$， 计算过程中可能会出现上溢($|\lambda_1|>1$)或下溢成为0 ($|\lambda_1|<1$). 为避免出现这一情形, 实际计算时每次迭代所求的向量都要归一化.

归一化过程：

设有一向量$x\neq 0$, 将其归一化得到向量

$$y=\frac{x}{\max(x)}$$

其中$\max(x)$表示向量x的绝对值最大的分量，即如果有

$$|x_{i_0}|=\max_{1\le i\le n}|x_i|$$

则$\max(x)=x_{i_0}$，且i_0为所有绝对值最大的分量中的最小下标.

例　$x=(1,-8,7)^T$，则 $\max(x)=-8$，归一化向量为

$$y=\frac{x}{\max(x)}=(-0.125,1,-0.875)^T$$

幂法的计算公式：

任取初始向量$x^{(0)}=y^{(0)}\neq 0$，对$k=1, 2, \ldots$，构造向量序列 $\{x^{(k)}\}$，$\{y^{(k)}\}$

$$\begin{cases} x^{(k)} = Ay^{(k-1)} \\ \alpha_k = \max\left(x^{(k)}\right) \\ y^{(k)} = \dfrac{x^{(k)}}{\alpha_k} \end{cases}$$

当k充分大时， $\alpha_k \approx \lambda_1$， $y^{(k)} \approx \dfrac{u_1}{\max(u_1)}$