第八章
矩阵特征值和特征向量的计算

§1 引言

§2 乘幂法与反幂法

§3* Jacobi方法

§1 引言

设**A**为n阶方阵，若数λ满足

$$\mathbf{Ax} = \lambda\mathbf{x} \quad \mathbf{x} \neq \mathbf{0}$$

则λ称为**A**的一个特征值。非零向量 $\mathbf{x}$ 称为与特征值λ对应的特征向量。

求**A**的特征值，可通过求$\det(\mathbf{A}-\lambda\mathbf{I})=0$的$n$个根得到；对应的特征向量可通过求$(\mathbf{A}-\lambda_i\mathbf{I})\mathbf{x}=\mathbf{0}$ $i=1,2,\cdots n$ 的非零解向量得到。

将$\det(\mathbf{A}-\lambda\mathbf{I})=0$展开为$\lambda$的多项式，未必得到精确的特征方程；且$n$增加，计算量迅速增加。

计算矩阵特征值及特征向量的数值方法：**迭代法和变换法**

§2 乘幂法与反幂法

一 乘幂法

求按模最大的特征值(主特征值)和相应的特征向量。

基本思想：采用迭代法求实矩阵的主特征值对应的特征向量，进而再求出主特征值。

设$\mathbf{A}\in\mathbf{R}^{n\times n}$的特征值为$|\lambda_1|\geq|\lambda_2|\geq\cdots\geq|\lambda_n|$，对应的特征向量$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$线性无关（构成$\mathbf{R}^n$上的一组基）。

对$\forall\mathbf{V}^{(0)}\in\mathbf{R}^n$ $(\mathbf{V}^{(0)}\neq\mathbf{0})$，设$\mathbf{V}^{(0)}=c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_n\mathbf{x}_n=\sum_{j=1}^{n}c_j\mathbf{x}_j$

构造向量序列

$$\mathbf{V}^{(1)}=\mathbf{A}\mathbf{V}^{(0)}=c_1\lambda_1\mathbf{x}_1+c_2\lambda_2\mathbf{x}_2+\cdots+c_n\lambda_n\mathbf{x}_n=\sum_{j=1}^{n}c_j\lambda_j\mathbf{x}_j$$

$$\mathbf{V}^{(2)}=\mathbf{A}\mathbf{V}^{(1)}=c_1\lambda_1^2\mathbf{x}_1+c_2\lambda_2^2\mathbf{x}_2+\cdots+c_n\lambda_n^2\mathbf{x}_n=\sum_{j=1}^{n}c_j\lambda_j^2\mathbf{x}_j$$

$$\vdots$$

$$\mathbf{V}^{(k)}=\mathbf{A}\mathbf{V}^{(k-1)}=c_1\lambda_1^k\mathbf{x}_1+c_2\lambda_2^k\mathbf{x}_2+\cdots+c_n\lambda_n^k\mathbf{x}_n=\sum_{j=1}^{n}c_j\lambda_j^k\mathbf{x}_j$$

即 $\mathbf{V}^{(k)}=\mathbf{A}\mathbf{V}^{(k-1)}=\mathbf{A}(\mathbf{A}\mathbf{V}^{(k-2)})=\mathbf{A}^2\mathbf{V}^{(k-2)}=\cdots=\mathbf{A}^k\mathbf{V}^{(0)}$

$\mathbf{V}^{(k)}$是用$\mathbf{A}$的k次幂左乘$\mathbf{V}^{(0)}$得到，故该方法称为**乘幂法**。

分下列情况讨论：

① $|\lambda_1|>|\lambda_j|\ (j=2,3,\cdots,n)$ （主特征值是特征方程的单根）

$$\mathbf{V}^{(k)}=\lambda_1^k(c_1\mathbf{x}_1+c_2(\frac{\lambda_2}{\lambda_1})^k\mathbf{x}_2+\cdots+c_n(\frac{\lambda_n}{\lambda_1})^k\mathbf{x}_n)$$

若 $c_1\neq 0$， 由于$\left|\frac{\lambda_j}{\lambda_1}\right|<1\,(j\geq 2)$， 当$k$充分大时，

$$\mathbf{V}^{(k)}\approx\lambda_1^k c_1\mathbf{x}_1$$

即$\mathbf{V}^{(k)}$是相应于λ_1的近似特征向量。

设 $V_i^{(k)}$表示$\mathbf{V}^{(k)}$的第i个分量，则

$$\frac{V_i^{(k+1)}}{V_i^{(k)}}\approx\frac{\lambda_1^{k+1}c_1(\mathbf{x}_1)_i}{\lambda_1^k c_1(\mathbf{x}_1)_i}=\lambda_1$$

$\mathbf{V}^{(k)}=c_1\lambda_1^k\mathbf{x}_1+c_2\lambda_2^k\mathbf{x}_2+\cdots+c_n\lambda_n^k\mathbf{x}_n$ $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}\Rightarrow\mathbf{A}(k\mathbf{x})=\lambda(k\mathbf{x})$

② $\lambda_1=\lambda_2=\cdots=\lambda_r, |\lambda_1|>|\lambda_{r+1}|\geq\cdots\geq|\lambda_n|$（主特征值是特征方程的重根）

$$\mathbf{V}^{(k)}=c_1\lambda_1^k\mathbf{x}_1+c_2\lambda_2^k\mathbf{x}_2+\cdots+c_n\lambda_n^k\mathbf{x}_n=\lambda_1^k[\sum_{j=1}^{r}c_j\mathbf{x}_j+\sum_{j=r+1}^{n}c_j(\frac{\lambda_j}{\lambda_1})^k\mathbf{x}_j]$$

若 $c_1\neq 0, c_2\neq 0,\cdots,c_r\neq 0$ 由于 $\left|\frac{\lambda_j}{\lambda_1}\right|<1\ (j\geq r+1)$，当$k$充分大时，

$$\mathbf{V}^{(k)}\approx\lambda_1^k\sum_{j=1}^{r}c_j\mathbf{x}_j$$

即$\mathbf{V}^{(k)}$仍是相应于 λ_1 的近似特征向量。

且

$$\frac{V_i^{(k+1)}}{V_i^{(k)}}\approx\frac{\lambda_1^{k+1}(\sum_{j=1}^{r}c_j\mathbf{x}_j)_i}{\lambda_1^{k}(\sum_{j=1}^{r}c_j\mathbf{x}_j)_i}=\lambda_1$$

Remark 此时λ_1的特征向量子空间不是一维，故$\mathbf{V}^{(k)}$只是相应于λ_1的特征子空间中的一个向量，且从不同的$\mathbf{V}^{(0)}$出发，得到的$\mathbf{V}^{(k)}$可能线性无关。

$$\mathbf{A}\mathbf{x}_1=\lambda\mathbf{x}_1, \mathbf{A}\mathbf{x}_2=\lambda\mathbf{x}_2\Rightarrow\mathbf{A}(k_1\mathbf{x}_1+k_2\mathbf{x}_2)=\lambda(k_1\mathbf{x}_1+k_2\mathbf{x}_2)$$

③ $\lambda_1 = -\lambda_2, |\lambda_1| = |\lambda_2| > |\lambda_j| \ (j = 3,4,\cdots,n)$（主特征值互为相反实数）

④ $\lambda_1 = \overline{\lambda_2}, |\lambda_1| > |\lambda_3| \geq \cdots \geq |\lambda_n|$（主特征值为特征方程的共轭复根）

（见其他书籍）

具体做法 取$\forall \mathbf{V}^{(0)} \in \mathbf{R}^n \ (\mathbf{V}^{(0)} \neq \mathbf{0})$。

$$\mathbf{V}^{(1)} = \mathbf{A}\mathbf{V}^{(0)} \qquad \lambda_1^{(1)} = \frac{V_i^{(1)}}{V_i^{(0)}}$$

$$\mathbf{V}^{(2)} = \mathbf{A}\mathbf{V}^{(1)} \qquad \lambda_1^{(2)} = \frac{V_i^{(2)}}{V_i^{(1)}} \qquad \text{若 } \left|\lambda_1^{(2)} - \lambda_1^{(1)}\right| \leq \varepsilon$$

则 $\lambda_1 \approx \lambda_1^{(2)}$，$\mathbf{V}^{(2)}$是与$\lambda_1$相应的近似特征向量。

否则 $\mathbf{V}^{(3)} = \mathbf{A}\mathbf{V}^{(2)} \qquad \lambda_1^{(3)} = \dfrac{V_i^{(3)}}{V_i^{(2)}} \qquad$ 若 $\left|\lambda_1^{(3)} - \lambda_1^{(2)}\right| \leq \varepsilon$

则$\lambda_1 \approx \lambda_1^{(3)}$，$\mathbf{V}^{(3)}$是与$\lambda_1$相应的近似特征向量。……

否则 $\mathbf{V}^{(k+1)} = \mathbf{A}\mathbf{V}^{(k)} \quad \lambda_1^{(k+1)} = \dfrac{V_i^{(k+1)}}{V_i^{(k)}}, k = 1,2,\cdots \quad$ 若$\left|\lambda_1^{(k+1)} - \lambda_1^{(k)}\right| \leq \varepsilon$，

即 $\dfrac{V_i^{(k+1)}}{V_i^{(k)}}$趋于定值，则$\lambda_1 \approx \lambda_1^{(k+1)}, \mathbf{V}^{(k+1)}$是相应于$\lambda_1$的近似特征向量。

$\mathbf{V}^{(k)} = \mathbf{A}\mathbf{V}^{(k-1)}$，$\mathbf{V}^{(k)} \approx \lambda_1^k c_1 \mathbf{x}_1, \mathbf{V}^{(k)} \approx \lambda_1^k \sum_{j=1}^{r} c_j \mathbf{x}_j,$

$\lambda_1 \approx V_i^{(k+1)} / V_i^{(k)} \quad V_i^{(k)}$ 为$\mathbf{V}^{(k)}$的第i个分量

Remarks

(1) 若计算相邻迭代向量的分量比时，其中某个分量为零，则计算其它分量的比值。

(2) 乘幂法中，迭代向量$\mathbf{V}^{(k)}=\mathbf{A}\mathbf{V}^{(k-1)}$的分量可能会出现绝对值非常大(当$|\lambda_1|>1$) 的现象，从而造成计算上溢。特征向量允许相差一个常数因子，实用中常每进行m步对迭代向量$\mathbf{V}^{(k)}$做一次规范化。

设$\max\mathbf{V}^{(k)}$表示向量$\mathbf{V}^{(k)}$按模最大的分量。则用$\tilde{\mathbf{V}}^{(k)}=\dfrac{\mathbf{V}^{(k)}}{\max\mathbf{V}^{(k)}}$代替$\mathbf{V}^{(k)}$继续迭代。$m$取值视情况而定。

$\mathbf{V}^{(k)}$分量绝对值非常小(当$|\lambda_1|<1$)时，会造成计算下溢，则用$\tilde{\mathbf{V}}^{(k)}=\dfrac{\mathbf{V}^{(k)}}{\min\mathbf{V}^{(k)}}$代替 $\mathbf{V}^{(k)}$继续迭代。

$$\mathbf{A}\mathbf{x}=\lambda\mathbf{x}\Rightarrow\mathbf{A}(k\mathbf{x})=\lambda(k\mathbf{x})$$

二 反(逆)幂法

求A按模最小的特征值和相应的特征向量。

设非奇异矩阵 $\mathbf{A}\in\mathbf{R}^{n\times n}$， 且A无零特征值。

当A的特征值满足$|\lambda_1|\ge|\lambda_2|\ge\cdots>|\lambda_n|>0$，对应的特征向量为$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$ 。$\mathbf{A}^{-1}$的特征值满足

$$|1/\lambda_n|>|1/\lambda_{n-1}|\ge\cdots\ge|1/\lambda_1|$$

并且A对应于λ_n的特征向量与$\mathbf{A}^{-1}$对应于$1/\lambda_n$的特征向量相同。

对 $\mathbf{A}^{-1}$用乘幂法求解的主特征值是 $1/\lambda_n$，特征向量是$\mathbf{x}_n$。从而可得A按模最小的特征值λ_n及相应的特征向量 $\mathbf{x}_n$ 。

用$\mathbf{A}^{-1}$代替A作乘幂，求A按模最小的特征值及相应特征向量的方法称为**反幂法**。

若$\mathbf{Ax}=\lambda\mathbf{x}$，则$\mathbf{A}^{-1}\mathbf{x}=\lambda^{-1}\mathbf{x}$

反幂法的**计算过程**

① 任取初始向量 $\mathbf{V}^{(0)} \in \mathbf{R}^n \ (\mathbf{V}^{(0)} \neq \mathbf{0})$

② 计算 $\mathbf{V}^{(k+1)} = \mathbf{A}^{-1}\mathbf{V}^{(k)} \quad k = 0,1,2,\cdots$

③ 根据 $\mathbf{V}^{(k)}$ 分量的变化趋势，求出$\mathbf{A}^{-1}$按模最大的特征值$1/\lambda_n$及对应特征向量$\mathbf{x}_n$的近似值，从而得到λ_n及对应特征向量$\mathbf{x}_n$的近似值。

为避免求逆，上面迭代公式**改写**为

$$\mathbf{A}\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)}$$

它是反幂法的主要工作量。

每步迭代中，矩阵**A**不变。故做三角分解**A=LU**，使得每次迭代只需解二个三角形方程组

$$\begin{cases} \mathbf{Ly} = \mathbf{V^{(k)}} \\ \mathbf{UV^{(k+1)}} = \mathbf{y} \end{cases}$$

反幂法的应用

若**A**的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，对应的特征向量为$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$。则 $\mathbf{B}=\mathbf{A}-p\mathbf{I}$ 的特征值为λ_1-p，$\lambda_2-p,\cdots,\lambda_n-p$，对应的特征向量也为$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$。

已知矩阵**A**之特征值λ_s的近似值p，求特征值λ_s及相应的特征向量。

一般有 $0<|\lambda_s-p|<|\lambda_j-p| \quad (j\neq s)$

$\mathbf{A}-p\mathbf{I}$（**I**为单位矩阵）按模最小的特征值是λ_s-p，对应的特征向量与**A**相应于λ_s的特征向量相同。

对$\mathbf{A}-p\mathbf{I}$用反幂法（即对$(\mathbf{A}-p\mathbf{I})^{-1}$用乘幂法)。

$$\mathbf{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i \Rightarrow (\mathbf{A}-p\mathbf{I})\mathbf{x}_i=(\lambda_i-p)\mathbf{x}_i$$

对$\mathbf{A}-p\mathbf{I}$用反幂法（即对$(\mathbf{A}-p\mathbf{I})^{-1}$用乘幂法），有

① 任取初始向量 $\mathbf{V}^{(0)} \in \mathbf{R}^n \; (\mathbf{V}^{(0)} \neq \mathbf{0})$

② 计算 $(\mathbf{A}-p\mathbf{I})\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)} \quad k=0,1,2,\cdots$ 对$(\mathbf{A}-p\mathbf{I})^{-1}$实施乘幂法

③ 根据 $\mathbf{V}^{(k)}$ 分量的变化趋势，求出$(\mathbf{A}-p\mathbf{I})^{-1}$按模最大特征值$1/(\lambda_s-p)$及对应特征向量 $\mathbf{x}_s$ 的近似值。从而得到λ_s 及对应特征向量 $\mathbf{x}_s$的近似值。

$$(\mathbf{A}-p\mathbf{I})^{-1}\mathbf{x}_s = \frac{1}{\lambda_s - p}\mathbf{x}_s \Rightarrow (\mathbf{A}-p\mathbf{I})\mathbf{x}_s = (\lambda_s - p)\mathbf{x}_s \Rightarrow \mathbf{A}\mathbf{x}_s = \lambda_s \mathbf{x}_s$$

总结

特征值、特征向量计算的方法

- 迭代法

  乘幂法–按模最大的特征值和相应的特征向量

  反幂法-按模最小的特征值和相应的特征向量

  反幂法的应用-已知特征值的近似值，求该特征值和相应的特征向量