数值分析

主讲教师：贺慧霞

北京航空航天大学数学科学学院

第三章 矩阵特征值与特征向量的算法

一、特征值问题及其性质

定义1　已知 $\mathrm{A}=(a_{ij})_{n\times n}$，则称

$$\phi(\lambda)=\det(\lambda\mathrm{I}-\mathrm{A})=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^{n}-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^{n}|A|$$

为 A 的**特征多项式**.

定义2 特征多项式 $\phi(\lambda)=\det(\lambda I - A)=0$的根称为A的**特征值**，$\lambda(A)$表示$A$的所有特征值的集合.

设λ为A的特征值,相应的齐次方程组 $(\lambda I - A)x = 0$ 的非零解x称为A的对应于λ的**特征向量**.

定理1 设λ是矩阵$A\in R^{n\times n}$的特征值，x是对应的非零特征向量，则

(1) $c\lambda$是cA的特征值(常数$c\neq 0$);

(2) $\lambda - p$为$A-pI$的特征值,即$(A-pI)x=(\lambda-p)x$;

(3) λ^k是A^k的特征值,即$A^k x=\lambda^k x$;

(4) 设A非奇异，则$\lambda\neq 0$且$\frac{1}{\lambda}$为A^{-1}的特征值,即$A^{-1}x=\frac{1}{\lambda}x$.

定理2 若$\lambda_i(i=1,\cdots,n)$是矩阵A的特征值，则

(1) $\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{n}a_{ii}=\mathrm{tr}(A)$, (2) $\det(A)=\lambda_1\cdots\lambda_n$.

定理3 设$A\in R^{n\times n}$，则 $\lambda(A^T)=\lambda(A)$.

定理2 (1) 设矩阵$A\in R^{n\times n}$可对角化，即存在非奇异矩阵P使

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

的充分必要条件是A具有n个线性无关的特征向量.

(2) 如果矩阵$A \in R^{n\times n}$有m个$(m \le n)$不同的特征值$\lambda_1, \lambda_2, \cdots, \lambda_m$，则对应的特征向量 $x_1, x_2, \cdots, x_m$ 线性无关.

定义 设$A \in R^{n\times n}$为对称矩阵，对于任一非零向量$x \neq 0$，称

$$R(x) = \frac{(Ax, x)}{(x, x)},$$

为对应于向量x的**瑞利(Rayleigh)商**.

定理4 设A为分块上三角阵，

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ & A_{22} & \cdots & A_{2m} \\ & & \ddots & \vdots \\ & & & A_{mm} \end{bmatrix}$$

其中每个对角块 A_{ii} 均为方阵，则 $\lambda(A)=\bigcup_{i=1}^{m}\lambda(A_{ii})$.

定理5 若A与B为相似矩阵，即∃非奇异P使$P^{-1}AP=B$,则

(1) A与B有相同的特征值；

(2) 若y是B的特征向量，则Py是A的特征向量.

定理 6(对称矩阵的正交约化) 设 $A \in R^{n \times n}$ 为对称矩阵，则

(1) A的特征值均为实数;

(2) A有 n 个线性无关的特征向量;

(3) 存在正交矩阵 P 使得

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

且 $\lambda_i (i = 1,2,\cdots,n)$ 为A的特征值，而 $P = (u_1, u_2, \cdots, u_n)$ 的列向量 u_j 为对应于 λ_j 的特征向量.

定理7 设$A\in R^{n\times n}$为对称矩阵(其特征值依次记为$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$)，则

(1). $$\lambda_n \leq \frac{(Ax,x)}{(x,x)} \leq \lambda_1$$ (对任何非零$x\in R^n$)；

(2). $$\lambda_1 = \max_{\substack{x\in R^n \\ x\neq 0}} \frac{(Ax,x)}{(x,x)};\qquad \lambda_n = \min_{\substack{x\in R^n \\ x\neq 0}} \frac{(Ax,x)}{(x,x)}. \qquad (1.3)$$

证明 只证(1)，关于(2)自己作练习.

由于A为实对称矩阵，可将 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 对应的特征向量 $x_1, x_2, \cdots, x_n$ 正交规范化，则有 $(x_i, x_j)=\delta_{ij}$，设$x\neq 0$为R^n中任一向量，则有

$$x=\sum_{i=1}^{n}\alpha_i x_i, \quad \|x\|_2=\sqrt{\sum_{i=1}^{n}\alpha_i^2}\neq 0.$$

于是

$$\lambda_n \le \frac{(Ax,x)}{(x,x)}=\frac{\sum_{i=1}^{n}\alpha_i^2\lambda_i}{\sum_{i=1}^{n}\alpha_i^2}\le \lambda_1.$$

从而(1)成立. 结论(1)说明瑞利商必位于λ_n和λ_1之间.

特征值的应用背景

- 求系统的固有振动频率
- 求机械系统的振动频谱（模态）
- 系统的稳定性分析
- 物理学中的某些临界值的确定

常用的数值求特征值的方法

- **幂法与反幂法**

  幂法是一种求实矩阵A的按模最大的特征值λ_1及其对应的特征向量x_1的方法。特别适合于大型稀疏矩阵。

- **Jacobi法**

  Jacobi是一种求实对称矩阵A的全部特征值及其对应的特征向量的方法。特别适合于中小型对称矩阵。

- **QR分解法**

  QR方法是一种变换方法，是计算一般(中小型)矩阵全部特征值问题的最有效方法之一.

- ……

3.1 幂法与反幂法

幂法是一种计算矩阵**主特征值**(矩阵按模最大的特征值)及对应特征向量的迭代方法.

二、幂法

1、主特征值是单根.

问题1： 假设实矩阵A具有n个线性无关的特征向量$x_1, x_2, \ldots, x_n$；相应的特征值$\lambda_1, \lambda_2, \ldots, \lambda_n$，已知矩阵的主特征值$\lambda_1$为实数,且满足

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|,$$

讨论求λ_1及其对应特征向量的方法.

则任意非零向量v可表示为 $v = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$,

取非零向量v_0: $v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$,

假设$\alpha_1 \neq 0$.

幂法的基本思想是:

任取**非零**的初始向量v_0，由矩阵A构造一向量序列$\{v_k\}$

$$\begin{cases} v_1 = Av_0, \\ v_2 = Av_1 = A^2 v_0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ v_{k+1} = Av_k = A^{k+1} v_0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases} \quad (2.1)$$

称为**迭代向量**，由假设，v_0可唯一表示为

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \ (\text{设} a_1 \neq 0), \quad (2.2)$$

定义： 由已知非零向量v_0及矩阵A的乘幂A^k构造向量序列$\{v_k\}$以计算矩阵A的**主特征值及相应**特征向量的方法称为**幂法**.

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \quad (\text{设} a_1 \neq 0), \qquad (2.2)$$

$$v_k = Av_{k-1} = A^2 v_{k-2} = \cdots = A^k v_0 = \alpha_1 A^k x_1 + \alpha_2 A^k x_2 + \cdots + \alpha_n A^k x_n$$

$$= \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \cdots + \alpha_n \lambda_n^k x_n$$

$$= \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2^k}{\lambda_1^k} \right) x_2 + \cdots + \alpha_n \left(\frac{\lambda_n^k}{\lambda_1^k} \right) x_n \right]$$

$$\varepsilon_k \xrightarrow{k \to \infty} 0$$

当$\alpha_1 \neq 0$时， $v_k \approx \lambda_1^k \alpha_1 x_1$， $Av_k = v_{k+1} \approx \lambda_1^{k+1} \alpha_1 x_1 = \lambda_1 v_k$

即v_k可以近似的作为属于λ_1的特征向量.

$$\boxed{v_k = \lambda_1^k \left[\alpha_1 x_1 + \varepsilon_k \right]}$$

$$v_k = \lambda_1^k \left[\alpha_1 x_1 + \varepsilon_k \right]$$

用 $(v_k)_i$ 表示 v_k 的第 i 个分量，则当 k 充分大时，有

$$\frac{(v_{k+1})_i}{(v_k)_i} = \frac{\lambda_1^{k+1}\left[\alpha_1(\vec{x}_1)_i + (\vec{\varepsilon}_{k+1})_i\right]}{\lambda_1^{k}\left[\alpha_1(\vec{x}_1)_i + (\vec{\varepsilon}_{k})_i\right]} = \lambda_1 \left\{ \frac{\alpha_1(\vec{x}_1)_i + (\vec{\varepsilon}_{k+1})_i}{\alpha_1(\vec{x}_1)_i + (\vec{\varepsilon}_{k})_i} \right\} \xrightarrow{k\to\infty} \lambda_1$$

即为 A 的主特征值 λ_1 的近似值.

此时两相邻迭代向量的对应非零分量的比值一定收敛到主特征值！

定理 1： 1.假设实矩阵 A 具有 n 个线性无关的特征向量 x_1，x_2，…，x_n；2.相应的特征值 λ_1, λ_2，…，λ_n 满足 $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$，于是对任何非零初始向量 v_0，

$$\lambda_1 \approx \frac{(v_{k+1})_i}{(v_k)_i}, \quad 及\ x_1 \approx v_k .$$

迭代公式实质上是由矩阵A的乘幂 A^k与非零向量v_0相乘来构造向量序列$\{v_k\}=\{A^k v_0\}$，从而计算主特征值λ_1及其对应的特征向量，这就是幂法的思想.

$$\frac{(v_{k+1})_i}{(v_k)_i} \to \lambda_1 \quad (k \to \infty).$$

的收敛速度由比值

$$r = \left|\frac{\lambda_2}{\lambda_1}\right|,$$

来确定，r 越小收敛越快，但当$r\approx 1$时收敛可能很慢.

2、主特征值是重根.

问题2： 假设实矩阵A具有n个线性无关的特征向量x_1，x_2，…，x_n；相应的特征值λ_1,λ_2，…，λ_n，已知矩阵的**主特征值λ_1**为实数，且满足

$$|\lambda_1|=|\lambda_2|=\cdots=|\lambda_r|>|\lambda_{r+1}|\geq\ldots\geq|\lambda_n|,$$

讨论求λ_1及其对应特征向量的方法.

$\forall\vec{u}_0\neq\vec{0}$，$\vec{u}_0=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n$，假设$\alpha_1,\cdots,\alpha_r$不全为零.

$$u_k=A^ku_0=\lambda_1^k\left[\sum_{i=1}^{r}\alpha_ix_i+\sum_{j=r+1}^{n}\alpha_j\left(\frac{\lambda_j^k}{\lambda_1^k}\right)x_i\right]=\lambda_1^k\left[\sum_{i=1}^{r}\alpha_ix_i+\varepsilon_k\right]$$

$\vec{u}_k\approx\lambda_1^k\sum_{i=1}^{r}\alpha_i\vec{x}_i$，即$u_k$可以近似的作为属于$\lambda_1$的特征向量.

A的最大特征值是r重时，幂方法仍然有效

注记：

1.该方法的本质是$u_k = A^k u_0$，因此得名**幂方法**；

2. A有n个线性无关的特征向量时，幂方法有效

3. 当矩阵没有n个线性无关的特征向量时，幂法仍然有效，只是收敛速度很慢，应换其他方法.

4.为防止u_k过大或过小，我们在每一步迭代中，对迭代向量**规范化**,使得其范数等于1.

设有一非零向量v,将其规范化得到向量 $y = \dfrac{u}{\| u \|}$,

$y_0 = u_0,$

$$\begin{cases} u_1 = Ay_0 = Au_0, \longrightarrow & y_1 = \dfrac{u_1}{\|u_1\|} = \dfrac{Au_0}{\|Au_0\|}, \\ u_2 = Ay_1 = \dfrac{A^2u_0}{\|Au_0\|}, \longrightarrow & y_2 = \dfrac{u_2}{\|u_2\|} = \dfrac{A^2u_0}{\|A^2u_0\|}, \\ \cdots\cdots & \cdots\cdots \\ u_k = Ay_{k-1} = \dfrac{A^ku_0}{\|A^{k-1}u_0\|}, \longrightarrow & y_k = \dfrac{u_k}{\|u_k\|} = \dfrac{A^ku_0}{\|A^ku_0\|}. \end{cases}$$

$$\text{一般情况下迭代公式}\begin{cases}\text{选取非零向量}u_0, \\ u_k = Au_{k-1}, (k=1,2,\cdots)\end{cases}$$

$$\text{实际使用的迭代公式为：}\begin{cases}\text{选取非零向量}u_0 \\ y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ u_k = Ay_{k-1}, (k \ge 1)\end{cases}$$

$$u_k = \frac{Au_{k-1}}{\| u_{k-1} \|} = \frac{A^2 u_{k-2}}{\| Au_{k-2} \|} = \cdots = \frac{A^k u_0}{\| A^{k-1} u_0 \|}.$$

$\forall \vec{u}_0 \ne \vec{0}$， $\vec{u}_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$， 假设 $\alpha_1 \ne 0$.

$$y_k = \frac{u_k}{\| u_k \|} = \frac{A^k u_0}{\| A^k u_0 \|}$$

证明： 设$u_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$，

$$u_k = Au_{k-1} = A^2 u_{k-2} = \cdots = A^k u_0 = \alpha_1 A^k x_1 + \alpha_2 A^k x_2 + \cdots + \alpha_n A^k x_n$$

$$= \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \cdots + \alpha_n \lambda_n^k x_n$$

$$= \lambda_1^k \left[\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2^k}{\lambda_1^k} \right) x_2 + \cdots + \alpha_n \left(\frac{\lambda_n^k}{\lambda_1^k} \right) x_n \right]$$

$$\| A^k u_0 \| = | \lambda_1^k | \left[\left\| \alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2^k}{\lambda_1^k} \right) x_2 + \cdots + \alpha_n \left(\frac{\lambda_n^k}{\lambda_1^k} \right) x_n \right\| \right]$$

$$\text{一般情况下迭代公式}\begin{cases}\text{选取非零向量}u_0, \\ u_k = Au_{k-1}, (k=1,2,\cdots)\end{cases}$$

实际使用的迭代公式为：

$$\begin{cases}\text{选取非零向量}u_0 \\ y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ u_k = Ay_{k-1}, (k \geq 1)\end{cases}$$

$$u_k = \frac{Au_{k-1}}{\|u_{k-1}\|} = \frac{A^2u_{k-2}}{\|Au_{k-1}\|} = \cdots = \frac{A^k u_0}{\|A^{k-1}u_0\|}.$$

$\forall \vec{u}_0 \neq \vec{0}$， $\vec{u}_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$， 假设 $\alpha_1 \neq 0$.

$$y_k = \frac{u_k}{\|u_k\|} = \frac{A^k u_0}{\|A^k u_0\|} = \left(\frac{\lambda_1}{|\lambda_1|}\right)^k \frac{\alpha_1 x_1 + \alpha_2\left(\dfrac{\lambda_2^k}{\lambda_1^k}\right)x_2 + \cdots + \alpha_n\left(\dfrac{\lambda_n^k}{\lambda_1^k}\right)x_n}{\left\|\alpha_1 x_1 + \alpha_2\left(\dfrac{\lambda_2^k}{\lambda_1^k}\right)x_2 + \cdots + \alpha_n\left(\dfrac{\lambda_n^k}{\lambda_1^k}\right)x_n\right\|}$$

$$y_k = \left(\frac{\lambda_1}{|\lambda_1|}\right)^k \frac{\alpha_1 x_1 + \alpha_2\left(\dfrac{\lambda_2^k}{\lambda_1^k}\right)x_2 + \cdots + \alpha_n\left(\dfrac{\lambda_n^k}{\lambda_1^k}\right)x_n}{\left\|\alpha_1 x_1 + \alpha_2\left(\dfrac{\lambda_2^k}{\lambda_1^k}\right)x_2 + \cdots + \alpha_n\left(\dfrac{\lambda_n^k}{\lambda_1^k}\right)x_n\right\|}$$

如果$\lambda_1 > 0, \Rightarrow y_k \xrightarrow{k\to\infty} \dfrac{\alpha_1 x_1}{\|\alpha_1 x_1\|}$，如果$\lambda_1 < 0, \Rightarrow y_k \approx (-1)^k \dfrac{\alpha_1 x_1}{\|\alpha_1 x_1\|}$，

当$k \to +\infty$时，

$$Ay_k \sim A\left[(\pm 1)^k \frac{\alpha_1 x_1}{\|\alpha_1 x_1\|}\right] = (\pm 1)^k \frac{\alpha_1 A x_1}{\|\alpha_1 x_1\|} = (-1)^k \frac{\lambda_1 \alpha_1 x_1}{\|\alpha_1 x_1\|} = \lambda_1 y_k.$$

这说明y_k可以近似地作为A的属于特征值λ_1对应的特征向量.

不同范数选取下的特征值的计算

$$y_k \approx (\pm 1)^k \frac{\alpha_1 x_1}{\|\alpha_1 x_1\|_2},$$

1. 取范数为2－范数时

令$\beta_k = y_{k-1}^T u_k$， 由于$u_k = A y_{k-1} \sim \lambda_1 y_{k-1}$

$$\beta_k = y_{k-1}{}^T A y_{k-1} \sim y_{k-1}^T \lambda_1 y_{k-1}, \quad k \to \infty.$$

$$\lim_{k\to\infty} \beta_k = \lim_{k\to\infty} y_{k-1}^T \lambda_1 y_{k-1} = \lambda_1 \| y_{k-1} \|_2^2 = \lambda_1$$

对应的迭代公式
$$\begin{cases} \forall\ 0 \neq u_0 \in R^n \\ \eta_{k-1} = \sqrt{u_{k-1}^T u_{k-1}} \\ y_{k-1} = u_{k-1} / \eta_{k-1} \\ u_k = A y_{k-1} \\ \beta_k = y_{k-1}^T u_k, \qquad (k = 1, 2, \ldots) \end{cases}$$

$\dfrac{|\beta_k - \beta_{k-1}|}{|\beta_k|} \le \varepsilon$时终止程序，

当前的β_k与y_{k-1}分别为λ_1和与它对应的特征向量.

编制程序容易，迭代一次所需时间较短

2. 取范数为∞－范数时

假设u_{k-1}的第r个分量为模最大的分量,令$\beta_k=\dfrac{\vec{e}_r^{\,T}u_k}{\vec{e}_r^{\,T}y_{k-1}}$,$= sign(h_r^{(k-1)})h_r^{(k)}$

由于$u_k=Ay_{k-1}$ $\sim\lambda_1 y_{k-1}$ $\beta_k=\dfrac{\vec{e}_r^{\,T}u_k}{\vec{e}_r^{\,T}y_{k-1}}\sim\dfrac{\vec{e}_r^{\,T}\lambda_1 y_{k-1}}{\vec{e}_r^{\,T}y_{k-1}}=\lambda_1,$ $\Leftrightarrow\lim\limits_{k\to\infty}\beta_k=\lambda_1.$

对应的迭代公式

$$\begin{cases}\text{选取非零向量}u_0=(h_0^{(1)},h_0^{(2)},\cdots,h_0^{(n)})^T\\ y_{k-1}=\dfrac{u_{k-1}}{\|u_{k-1}\|_\infty},k=1,2,\cdots\\ u_k=Ay_{k-1}=(h_1^{(k)},\cdots,h_n^{(k)})^T \qquad |h_r^{(k-1)}|=\|u_{k-1}\|_\infty\\ \beta_k=sign(h_r^{(k-1)})h_r^{(k)},\ \ k=1,2,\ldots\end{cases}$$

$\dfrac{|\beta_k-\beta_{k-1}|}{|\beta_k|}\le\varepsilon$时终止程序，当前的$\beta_k$与$y_{k-1}$分别为$\lambda_1$和与它对应的特征向量.

选主元，所需时间较长，但舍入误差小.

理解：$$\beta_k=\frac{\vec{e}_r^{\,T}u_k}{\vec{e}_r^{\,T}y_{k-1}}=sign(h_r^{(k-1)})h_r^{(k)}$$

$$u_{k-1}=(1,-3,2,6)^T,|h_4^{(k-1)}|=\|u_{k-1}\|_\infty,h_4^{(k-1)}=6,$$

$$y_{k-1}=\frac{u_{k-1}}{\|u_{k-1}\|_\infty}=(\frac{1}{6},-\frac{1}{2},\frac{1}{3},1)^T,\quad \vec{e}_4=(0,0,0,1),$$

设$u_k=(0,-7,1,3)=Ay_{k-1}=(h_1^{(k)},h_2^{(k)},h_3^{(k)},h_4^{(k)})$,

$$\beta_k=\frac{\vec{e}_4^{\,T}u_k}{\vec{e}_4^{\,T}y_{k-1}}=h_4^{(k)}=3=sign(h_4^{(k-1)})h_4^{(k)}.$$

例2: 用幂法求矩阵$A=\begin{bmatrix}2 & -1 & 0\\ 0 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}$的按模最大的特征值和相应的特征向量。取$u_0=(0,0,1)$, $\dfrac{|\beta_k-\beta_{k-1}|}{|\beta_k|}<\dfrac{1}{2}\times 10^{-3}$.

解: $u_0=(0,0,1)^T$, $|h_3^{(0)}|=\|u_0\|_\infty=1$, $y_0=(0,0,1)^T$, $\mathrm{sign}(h_3^{(0)})=1$,

$u_1=Ay_0=(0,-1,2)^T, h_3^{(1)}=2$, $y_1=\dfrac{u_1}{|h_3^{(1)}|}=(0,-\dfrac{1}{2},1)^T$,

$\beta_1=\mathrm{sgn}(h_3^{(0)})h_3^{(1)}=2$, $\mathrm{sign}(h_3^{(1)})=1$,

$u_2=Ay_1=(0.5,-2,2.5)^T, h_3^{(2)}=2.5$, $y_2=\dfrac{u_2}{|h_3^{(2)}|}=(0.2,-0.8,1)^T$,

$\beta_2=\mathrm{sgn}(h_3^{(1)})h_3^{(2)}=2.5$, $\mathrm{sign}(h_3^{(2)})=1$, ……

$$u_8 = Ay_7 = (2.7650948, -2.9981848, 2.9990924)^T, h_3^{(8)} = 2.9990924,$$

$$y_8 = \frac{u_8}{\left|h_3^{(8)}\right|} = (0.9219772, -0.9996973, 1)^T, \ \beta_8 = \mathbf{sign}(h_3^{(7)})h_3^{(8)} = 2.9990924,$$

$$\mathbf{sign}(h_3^{(7)}) = 1,$$

$$u_9 = Ay_8 = (2.8436517, -2.9993946, 2.9996973)^T, h_3^{(9)} = 2.9996973,$$

$$\mathbf{sign}(h_3^{(8)}) = 1, \quad \beta_9 = \mathbf{sign}(h_3^{(8)})h_3^{(9)} = 2.9996973,$$

k	v_k			u_k			β_k
0	0	0	1	0	0	1	
1	0	-1	2	0	0.5	1	2
2	0.5	-2	2.5	0.2	-0.8	1	2.5
……	……	……	……	……	……	……	……
8	2.7650948	-2.9981848	2.9990924	0.9219772	-0.9996973	1	2.9990924
9	2.8436517	-2.9993946	2.9996973				2.9996973

由于$\dfrac{|\beta_9-\beta_8|}{|\beta_9|}=\dfrac{2.9996973-2.9990924}{2.9996973}=\dfrac{0.0006049}{2.9996973}\approx 0.0002017<\dfrac{1}{2}\times 10^{-3}$,

所以 $\lambda_1 \approx \beta_9 = 2.9996973,$

相应的特征向量为$\boldsymbol{x}_1 \approx \boldsymbol{u}_8=\dfrac{\boldsymbol{v}_8}{|h_3^{(8)}|}=(0.9219772,-0.9996973,1)^T$.

例3：用幂法求矩阵

$$A=\begin{bmatrix} 6 & -12 & 6 \\ -21 & -3 & 24 \\ -12 & -12 & 51 \end{bmatrix}$$

按模最大的特征值 λ_1 和属于 λ_1 的特征向量，要求 $|\beta_k-\beta_{k-1}|/|\beta_k|\leqslant 0.000\,1$

书本3.1节例1，自学看懂.

数值分析

主讲教师：贺慧霞

北京航空航天大学数学科学学院

第三章 矩阵特征值与特征向量的算法

3.1 幂法与反幂法

幂法是一种计算矩阵按模最大的特征值与其对应的特征向量的迭代方法.

适合于大型稀疏矩阵

反幂法

1.可以计算矩阵按模最小的特征值与其对应的特征向量;

2.求一个给定近似特征值对应的特征向量.

二、反幂法

问题1： 假设实矩阵A具有n个线性无关的特征向量$x_1, x_2, \ldots, x_n$;相应的特征值$\lambda_1, \lambda_2, \ldots, \lambda_n$，已知矩阵的**按模最小的特征值$\lambda_n$**为实数,且满足

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_{n-1}| > |\lambda_n|$$

讨论求λ_n及其对应特征向量的方法.

设A非奇异，则$\lambda_i \neq 0$，且$\dfrac{1}{\lambda_i}$是A^{-1}的特征值，即 $A^{-1}x_i = \dfrac{1}{\lambda_i}x_i$

A^{-1}的特征值满足$\left|\dfrac{1}{\lambda_n}\right| > \left|\dfrac{1}{\lambda_{n-1}}\right| \geq \cdots \geq \left|\dfrac{1}{\lambda_2}\right| \geq \left|\dfrac{1}{\lambda_1}\right|$

A^{-1}的特征值满足$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \geq \cdots \geq \left|\frac{1}{\lambda_2}\right| \geq \left|\frac{1}{\lambda_1}\right|$

对于 A^{-1} 应用幂方法 $\begin{cases} y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ u_k = A^{-1} y_{k-1} \end{cases} \Longrightarrow \begin{cases} y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ A u_k = y_{k-1} \end{cases}$

可计算出A^{-1}的按模最大的特征值,其倒数恰好为 A 的按模最小的特征值

定理1 假设实矩阵A具有n个线性无关的特征向量$x_1, x_2, \ldots, x_n$,相应的特征值$\lambda_1, \lambda_2, \ldots, \lambda_n$满足

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n| > 0,$$

则对任意的非零初始向量$u_0 = \sum_{i=1}^{n} \alpha_i x_i$(至少有一个$\alpha_i \neq 0$),由反幂法迭代公式构造的向量序列$\{y_k\},\{u_k\}$满足

$$(1)\lim_{k\to\infty} y_k = \frac{x_n}{\|x_n\|}, \quad (2)\lim_{k\to\infty}\|u_k\| = \frac{1}{\lambda_n}.$$

$$\begin{cases} u_0 \neq 0 \\ y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ Au_k = y_{k-1} \end{cases}$$

1. 取范数为∞－范数时

对应的迭代公式 $\begin{cases} \text{选取非零向量} \boldsymbol{u}_0 \\ \boldsymbol{y}_{k-1} = \dfrac{\boldsymbol{u}_{k-1}}{\| \boldsymbol{u}_{k-1} \|_\infty}, k = 1,2,\cdots \\ A\boldsymbol{u}_k = \boldsymbol{y}_{k-1} \\ \beta_k = sign(h_r^{(k-1)}) h_r^{(k)}, \quad k = 1,2,\ldots \end{cases}$ $\quad |h_r^{(k-1)}| = \| \boldsymbol{u}_{k-1} \|_\infty$

$\dfrac{|\beta_k - \beta_{k-1}|}{|\beta_k|} \le \varepsilon$时终止程序.

$\lambda_n \approx \dfrac{1}{\beta_k}, \boldsymbol{y}_{k-1}$可近似作为矩阵A的属于$\lambda_n$的特征向量.

2. 取范数为2－范数时的迭代公式

$$\begin{cases} y_{k-1} = \dfrac{u_{k-1}}{\left\| u_{k-1} \right\|} \\ Au_k = y_{k-1} \end{cases}$$

对应的迭代公式
$$\begin{cases} \forall\ 0 \neq u_0 \in R^n \\ \eta_{k-1} = \sqrt{u_{k-1}^T u_{k-1}} \\ y_{k-1} = u_{k-1} / \eta_{k-1} \\ Au_k = y_{k-1} \\ \beta_k = y_{k-1}^T u_k \\ \qquad (k = 1, 2, \ldots) \end{cases}$$

$\dfrac{|\beta_k - \beta_{k-1}|}{|\beta_k|} \le \varepsilon$时终止程序.

$r = \dfrac{|\lambda_n|}{|\lambda_{n-1}|}$越小，收敛速度越快.

$r \to 1$,收敛可能很慢.

$\lambda_n \approx \dfrac{1}{\beta_k}$, y_{k-1}可近似作为矩阵A的属于λ_n的特征向量.

反幂法的解题步骤:

1.对A进行三角分解$A=LU$,(或者$PA=LU$,P是置换矩阵），

2.取非零向量v_0,计算$\|v_0\|$，得$u_0=\dfrac{v_0}{\|v_0\|}$,

3. $LUv_1=u_0,\Leftrightarrow\begin{cases}Ly_1=u_0\\Uv_1=y_1\end{cases}\Rightarrow v_1\Rightarrow u_1=\dfrac{v_1}{\|v_1\|}$,

4. $LUv_2=u_1,\Leftrightarrow\begin{cases}Ly_2=u_1\\Uv_1=y_2\end{cases}\Rightarrow v_2\Rightarrow u_2=\dfrac{v_2}{\|v_2\|},\cdots\cdots$

5. $LUv_k=u_{k-1},\Leftrightarrow\begin{cases}Ly_k=u_{k-1}\\Uv_k=y_k\end{cases}\Rightarrow v_k\Rightarrow u_k=\dfrac{v_k}{\|v_k\|}$,

【例】用反幂法求$A=\begin{pmatrix}2&1&0\\1&3&1\\0&1&4\end{pmatrix}$的按模最小的特征值及对应的特征向量.

LU分解计算u_k, y_k, β_k

【解】把A做三角分解，得

$$L=\begin{pmatrix}1&0&0\\0.5&1&0\\0&0.4&1\end{pmatrix}, U=\begin{pmatrix}2&1&0\\0&2.5&1\\0&0&3.6\end{pmatrix},$$

迭代公式 $\begin{cases}\forall\ 0\neq u_0\in R^n\\ \eta_{k-1}=\sqrt{u_{k-1}^T u_{k-1}}\\ y_{k-1}=u_{k-1}/\eta_{k-1}\\ Au_k=y_{k-1}\\ \beta_k=y_{k-1}^T u_k,(k=1,2,\ldots)\end{cases}$

$u_0=(0,0,1)^T$ 则$y_0=(0,0,1)^T$

$Au_1=LUu_1=y_0=(0,0,1)^T$ 记$w_1=Uu_1=(w_{11},w_{12},w_{13})^T$,

$$Lw_1=\begin{pmatrix}1&0&0\\0.5&1&0\\0&0.4&1\end{pmatrix}\begin{pmatrix}w_{11}\\w_{12}\\w_{13}\end{pmatrix}=\begin{pmatrix}0\\0\\1\end{pmatrix}, \Rightarrow \begin{pmatrix}w_{11}\\w_{12}\\w_{13}\end{pmatrix}=\begin{pmatrix}0\\0\\1\end{pmatrix},$$

$$Uu_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2.5 & 1 \\ 0 & 0 & 3.6 \end{pmatrix}\begin{pmatrix} u_{11} \\ u_{12} \\ u_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \Rightarrow \begin{pmatrix} u_{11} \\ u_{12} \\ u_{13} \end{pmatrix} = \begin{pmatrix} 1/18 \\ -1/9 \\ 5/18 \end{pmatrix}, \quad \beta_1 = y_0^{\ T} u_1 = \frac{5}{18}.$$

$$y_1 = \frac{u_1}{\| u_1 \|_2} = \frac{(1,-2,5)^T}{\sqrt{30}}$$

迭代公式
$$\begin{cases} \forall\ 0 \neq u_0 \in R^n \\ \eta_{k-1} = \sqrt{u_{k-1}^T u_{k-1}} \\ y_{k-1} = u_{k-1} / \eta_{k-1} \\ Au_k = y_{k-1} \\ \beta_k = y_{k-1}^T u_k, (k = 1,2,\ldots) \end{cases}$$

$$Au_2 = LUu_2 = y_1,$$

记$w_2 = Uu_2 = (w_{21}, w_{22}, w_{23})^T$,

$$Lw_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0.5 & 1 \end{pmatrix}\begin{pmatrix} w_{21} \\ w_{22} \\ w_{23} \end{pmatrix} = \frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}, \Rightarrow \begin{pmatrix} w_{21} \\ w_{22} \\ w_{23} \end{pmatrix} = \ldots\ldots$$

例 2 用反幂法求矩阵

$$A = \begin{bmatrix} 6 & -12 & 6 \\ -21 & -3 & 24 \\ -12 & -12 & 51 \end{bmatrix}$$

的按模最小的特征值和相应的特征向量，要求 $|\beta_k^{-1}-\beta_{k-1}^{-1}|/|\beta_k^{-1}|\leqslant 0.005$。

对应的迭代公式

$$\begin{cases} \forall\ 0 \neq u_0 \in R^n \\ \eta_{k-1} = \sqrt{u_{k-1}^T u_{k-1}} \\ y_{k-1} = u_{k-1}/\eta_{k-1} \\ Au_k = y_{k-1} \\ \beta_k = y_{k-1}^T u_k \\ \qquad (k = 1,2,\ldots) \end{cases}$$

$$u_0 = (1,1,1)^T$$

$$y_0 = \frac{(1,1,1)^T}{\sqrt{3}}$$

LU分解计算u_k, y_k, β_k

表 3-2　例 2 计算结果

k	$\boldsymbol{u}_k^{\mathrm{T}}$			$\boldsymbol{y}_k^{\mathrm{T}}$			β_k
0	1.000 0	1.000 0	1.000 0	0.577 4	0.577 4	0.577 4	
1	−0.032 1	−0.070 6	−0.012 8	−0.408 2	−0.898 1	−0.163 3	−0.066 7
2	0.068 0	0.084 4	0.008 5	0.627 8	0.778 4	0.078 1	−0.104 9
3	−0.066 4	−0.104 9	−0.038 8	−0.510 5	−0.806 5	−0.298 1	−0.126 4
4	0.058 2	0.085 6	0.028 0	0.542 4	0.798 5	0.261 0	−0.107 1
5	−0.059 5	−0.090 0	−0.030 1	−0.531 4	−0.803 5	−0.268 4	−0.112 0
6	0.059 4	0.088 8	0.029 6	0.535 9	0.800 9	0.267 1	−0.110 9
7	−0.059 4	−0.089 2	−0.029 7				−0.111 2

β_7 已满足精度要求。所以，$\boldsymbol{A}$ 的按模最小的特征值为 $\lambda_3 \approx \frac{1}{\beta_7} = -8.992\,8$，相应的特征向量为 $\boldsymbol{x}_3 \approx (0.535\,9, 0.800\,9, 0.267\,1)^{\mathrm{T}}$。

二、带原点平移的反幂法----对$(A-pI)^{-1}$用幂法

定理1 设λ是矩阵$A\in R^{n\times n}$的特征值，x是对应的非零特征向量，则

(1) $c\lambda$是cA的特征值 (常数$c\neq 0$);

(2) $\lambda - p$为$A-pI$的特征值，即$(A-pI)x=(\lambda-p)x$;

(3) λ^k是A^k的特征值，即$A^k x=\lambda^k x$;

(4) 设A非奇异，则$\lambda\neq 0$且$\dfrac{1}{\lambda}$为A^{-1}的特征值，即$A^{-1}x=\dfrac{1}{\lambda}x$.

$$\begin{cases} \forall 0\neq u_0 \\ y_{k-1}=\dfrac{u_{k-1}}{\|u_{k-1}\|} \\ u_k=(A-pI)^{-1}y_{k-1} \end{cases} \Longrightarrow \begin{cases} \forall 0\neq u \\ y_{k-1}=\dfrac{u_{k-1}}{\|u_{k-1}\|} \\ (A-pI)u_k=y_{k-1} \end{cases}$$

定理2 假设实矩阵A具有n个线性无关的特征向量$x_1, x_2, \dots, x_n$,.相应的特征值$\lambda_1, \lambda_2, \dots, \lambda_n$，而p为$\lambda_j$的近似值,$(A-pI)^{-1}$存在,且

$$|\lambda_j - p| \ll |\lambda_i - p|, \quad i \neq j,$$

对任意的非零初始向量$u_0 = \sum_{i=1}^{n} \alpha_i x_i$(至少存在一个$\alpha_i \neq 0$),由反幂法迭代公式构造的向量序列$\{u_k\},\{y_k\}$

$$\begin{cases} \forall\ u_0 \in R^n \\ \eta_{k-1} = \sqrt{u_{k-1}^T u_{k-1}} \\ y_{k-1} = u_{k-1}/\eta_{k-1} \\ (A-pI)u_k = y_{k-1} \\ \beta_k = y_{k-1}^T u_k \\ \qquad (k=1,2,\dots) \end{cases}$$

$$\begin{cases} y_{k-1} = \dfrac{u_{k-1}}{\|u_{k-1}\|} \\ Au_k = y_{k-1} \end{cases}$$

2－范数时的迭代公式

取范数为∞－范数时

对应的迭代公式

$$\begin{cases} \text{选取非零向量} u_0 \\ y_{k-1} = \dfrac{u_{k-1}}{\| u_{k-1} \|_\infty}, k = 1, 2, \cdots \\ (A - pI)u_k = y_{k-1} \\ \beta_k = sign(h_r^{(k-1)})h_r^{(k)}, \quad k = 1, 2, \ldots \quad | h_r^{(k-1)} | = \| u_{k-1} \|_\infty \end{cases}$$

收敛速度由 $r = \dfrac{|\lambda_j - p|}{\min\limits_{i \neq j} |\lambda_i - p|}$ 确定. $\dfrac{|\beta_k - \beta_{k-1}|}{|\beta_k|} \le \varepsilon$ 时终止程序.

$\lambda_n \approx p + \dfrac{1}{\|u_k\|}$, y_{k-1} 可近似作为矩阵A的属于 λ_n 的特征向量.

给定A的特征值λ_j的一个近似值p,求p对应的特征向量x_j的步骤:

1.对A进行三角分解$A-pI=LU$,(或者$P(A-pI)=LU$,P是置换矩阵),

2.取非零向量$v_0=(1,1,\cdots,1)^T$,

3. $LUv_1=v_0=(1,1,\cdots,1)^T \Leftrightarrow \begin{cases} Ly_1=v_0 \\ Uv_1=y_1 \end{cases} \Rightarrow v_1 \Rightarrow u_1=\dfrac{v_1}{\|v_1\|}$,

4. $LUv_2=u_1, \Leftrightarrow \begin{cases} Ly_2=u_1 \\ Uv_2=y_2 \end{cases} \Rightarrow v_2 \Rightarrow u_2=\dfrac{v_2}{\|v_2\|}, \cdots\cdots$

5. $\lambda_k \approx p+\dfrac{1}{\|v_k\|}$,对应的特征向量为$u_k=\dfrac{v_k}{\|v_k\|}$.

【例】用反幂法求$A=\begin{pmatrix}2&1&0\\1&3&1\\0&1&4\end{pmatrix}$的对应于特征值$p=1.2679$

(精确特征值为$\lambda_3=3-\sqrt{3}$)的特征向量(计算两步即可).

【解】由选主元的三角分解将A-pI分解为 $(A-pI)=LU$,其中

$$L=\begin{pmatrix}1&0&0\\0&1&0\\0.7321&-0.26807&1\end{pmatrix},U=\begin{pmatrix}1&1.7321&1\\0&1&2.7321\\0&0&0.29405\times10^{-3}\end{pmatrix},P=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix},$$

(1) $v_0=(1,1,1)^T$, $(A-pI)v_1=LUv_1=v_0$, 令$\begin{cases}Ly_1=v_0\\Uv_1=y_1\end{cases}$,

$$A - pI = \begin{pmatrix} 0.7321 & 1 & 0 \\ 1 & 1.7321 & 1 \\ 0 & 1 & 2.7321 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 1.7321 & 1 \\ 0.7321 & 1 & 0 \\ 0 & 1 & 2.7321 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1.7321 & 1 \\ 0.7321 & 1 - 0.7321 * 1.7321 & 0 \\ 0 & 1 & 2.7321 \end{pmatrix} = \begin{pmatrix} 1 & 1.7321 & 1 \\ 0.7321 & -0.26807 & 0 \\ 0 & 1 & 2.7321 \end{pmatrix}$$

$$\xrightarrow{r_2 \leftrightarrow r_3} \begin{pmatrix} 1 & 1.7321 & 1 \\ 0 & 1 & 2.7321 \\ 0.7321 & -0.26807 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1.7321 & 1 \\ 0 & 1 & 2.7321 \\ 0.7321 & -0.26807 & \begin{matrix} 0 - 0.7321 * 1 \\ +0.26807 * 2.7321 \end{matrix} \end{pmatrix}$$

【例】用反幂法求$A=\begin{pmatrix}2 & 1 & 0\\ 1 & 3 & 1\\ 0 & 1 & 4\end{pmatrix}$的对应于特征值$p=1.2679$

(精确特征值为$\lambda_3=3-\sqrt{3}$)的特征向量(计算两步即可).

【解】由选主元的三角分解将A-pI分解为 $(A-pI)=LU$,其中

$$L=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0.7321 & -0.26807 & 1\end{pmatrix},U=\begin{pmatrix}1 & 1.7321 & 1\\ 0 & 1 & 2.7321\\ 0 & 0 & 0.29405\times10^{-3}\end{pmatrix},P=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix},$$

(1) $v_0=(1,1,1)^T$, $(A-pI)v_1=LUv_1=v_0$, 令$\begin{cases}Uv_1=y_1\\ Ly_1=v_0\end{cases}$,

(1) $\quad v_0 = (1,1,1)^T$, $\quad$ 解方程组$(A - pI)v_1 = LUv_1 = v_0$,

$$y_1 = (1,1,0.53597)^T,$$

$$v_1 = (6802.1,-4978.8,1822.7)^T,$$

$$u_1 = \frac{v_1}{\| v_1 \|_\infty} = (1,-0.73195,0.26796)^T,$$

(2)解方程组$(A - pI)v_2 = LUv_2 = Pu_1$, 令$\begin{cases} Ly_2 = Pu_1 = (-0.73195,0.26796,1)^T \\ Uv_2 = y_2 \end{cases}$,

$$y_2 = (-0.73195,0.26796,1.60769)^T, \quad v_2 = (20404.6,-14937.2,5467.4)^T,$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.7321 & -0.26807 & 1 \end{pmatrix}, U = \begin{pmatrix} 1 & 1.7321 & 1 \\ 0 & 1 & 2.7321 \\ 0 & 0 & 0.29405\times 10^{-3} \end{pmatrix}, P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$v_2=(20404.6,-14937.2,5467.4)^T,$$

$$u_2=\frac{v_2}{\|v_2\|_\infty}=(1,-0.73205,0.26795)^T,$$

λ_3=1.2679对应的特征向量是

$$x_3=(1,1-\sqrt{3},2-\sqrt{3})^T\approx(1,-0.73206,0.26795)^T,$$

由此看出u_2是x_3的相当好的近似.

特征值$\lambda_3\approx 1.2679+\dfrac{1}{\|v_2\|}\approx 1.26794901$,

λ_3的真实值$\lambda_3=3-\sqrt{3}=1.26794912\cdots$.

总　结

- 幂法可以用来求矩阵模最大的特征值和特征向量；

A $\to$ A^{-1} 幂方法

- 反幂法可以用来求矩阵模最小的特征值和特征向量；

- 当$|\lambda_1|=|\lambda_2|$，但$\lambda_1= -\lambda_2$ 时，直接幂法失败；

  当$|\lambda_{n-1}|=|\lambda_n|$，但$\lambda_{n-1}= -\lambda_n$ 时，直接反幂法失败；

- 当是多重特征值时，幂法和反幂法仍有效。

作　业

- 教材第65页：习题1、3，6，7
- 课外阅读：《C数值算法》第11章