第九章 矩阵特征值特征向量计算

- 幂法
- 幂法的加速与降阶
- 反幂法
- 实对称矩阵Jacobi方法

引言

- 问题
  - 一个或几个绝对值最大(最小)的特征值(特征向量)
  - 全部特征值 (特征向量)
- 来源
  - 振动问题、稳定性问题和许多科学工程实际问题
- 两类方法
  - 求特征多项式的根
    - 高次多项式求根问题有特殊的困难
    - 重根的计算精度较低
    - 对舍入误差敏感
  - 迭代法

幂法

- 原理
  - 构造迭代序列$x^{(0)}, x^{(1)} = Ax^{(0)}, \ldots, x^{(k+1)} = Ax^{(k)}, \ldots$
  - 序列中大特征值的分量比重越来越大.
    - 设A有完全的特征系统$\lambda_j, v_j, j = 1,2, \ldots, n. |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_p| > |\lambda_{p+1}| \geq \cdots$
    - $x^{(0)} = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$
    - $x^{(k)} = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n$
      $\approx c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_p \lambda_p^k v_p$
    - 从而可得到一个或几个最大特征值的信息. 并将它们和对应的特征向量计算出来

例

- 例1求$A=\begin{pmatrix}\frac{1}{4} & \frac{1}{5}\\ \frac{1}{5} & \frac{1}{6}\end{pmatrix}$特征值特征向量

$$x^{(1)}=Ax^{(0)},\ x^{(2)}=Ax^{(1)},\cdots,x^{(k)}=Ax^{(k-1)},\cdots$$

k	x_1^k	x_2^k
0	1	0
1	0.25	0.2
2	0.10250	0.08333
3	0.042292	0.034389
4	0.017451	0.014190
…	…	…

$$\frac{x_1^{(2)}}{x_1^{(1)}}=0.41,\qquad \frac{x_1^{(3)}}{x_1^{(2)}}=0.412\ 6,\qquad \frac{x_1^{(4)}}{x_1^{(3)}}=0.412\ 63$$

$$\frac{x_2^{(2)}}{x_2^{(1)}}=0.416\ 65,\qquad \frac{x_2^{(3)}}{x_2^{(2)}}=0.412\ 68,\qquad \frac{x_2^{(4)}}{x_2^{(3)}}=0.412\ 63$$

求模最大特征值特征向量

- 情况一： $p = 1$
  - $x^{(k)} \approx c_1\lambda_1^k v_1, x^{(k+1)} \approx c_1\lambda_1^{k+1} v_1$(不妨设$c_1 \neq 0$)
  - $x^{(k+1)} \approx \lambda_1 x^{(k)}$由此得$\lambda_1$及对应特征向量
  - 算法（向量规一化以免溢出）
    - 任取$x = (x_1, x_2, \dots, x_n)^T, \mu = 0$
    - for $k = 1,2, \dots K$

$$\lambda = x_r : |x_r| = \max|x_r|$$

$$y = \frac{x}{\lambda}, x = Ay$$

if $|\lambda - \mu| \leq \varepsilon$, break end

$$\mu = \lambda$$

end

算例

- 例2.矩阵 $A=\begin{pmatrix}1 & -3 & 2\\ 4 & 4 & -1\\ 6 & 3 & 5\end{pmatrix}$ $v_1 \approx (0.2999, 0.0666, 1)^T$
  λ_1 =6.9992 (特征值特征向量)

k	$y^{(k)}$			$x^{(k+1)}$		
0	1	1	1	0	7	14
1	0.000 0	0.500 0	1	0.300 0	1	6.500 0
2	0.076 9	0.153 8	1	1.615 4	0.077 0	5.923 0
3	0.272 7	–0.013 0	1	2.311 8	0.038 8	6.597 2
4	0.350 4	0.005 9	1	2.332 7	0.425 2	6.419 3
5	0.363 4	0.066 2	1	2.164 7	0.718 5	7.379 1
6	0.293 7	0.097 5	1	2.001 2	0.564 6	7.051 5
7	0.283 7	0.080 0	1	2.043 6	0.456 8	6.942 2
8	0.294 4	0.065 5	1	2.097 8	0.439 5	6.962 8
9	0.301 3	0.063 1	1	2.111 9	0.457 6	6.997 1
10	0.301 8	0.065 4	1	2.105 6	0.468 9	7.007 1
11	0.300 5	0.066 9	1	2.099 8	0.469 6	7.003 7
12	0.299 8	0.067 1	1	2.098 7	0.467 4	7.000 0
13	0.299 8	0.066 8	1	2.099 5	0.466 2	6.999 1
14	0.299 9	0.066 6	1	2.100 1	0.466 0	6.999 2

情况二

- $p = 2, \lambda_1 = -\lambda_2$
- $x^{(k)} \approx c_1\lambda_1^k v_1 + c_2(-\lambda_1)^k v_2$，（不妨设$c_1 c_2 \neq 0$）
- $x^{(k+2)} \approx \lambda_1^2 x^{(k)}$，由此得$\lambda_1, \lambda_2$是比例常数的正负两平方根对应特征向量
- $x^{(k+1)} + \lambda_1 x^{(k)} \approx 2c_1\lambda_1^{k+1} v_1$
- $x^{(k+1)} + \lambda_2 x^{(k)} \approx 2c_2\lambda_2^{(k+1)} v_2$

情况三

- $p=2, \lambda_2 = \vec{\lambda}_1$
- $x^{(k)} \approx c_1\lambda_1^k v_1 + \vec{c}_1 \, \vec{\lambda}_1^{\,k} \, \vec{v}_1$
- $x^{(k+1)} \approx c_1\lambda_1^{k+1} v_1 + \vec{c}_1 \, \vec{\lambda}_1^{\,k+1} \, \vec{v}_1$
- $x^{(k+2)} \approx c_1\lambda_1^{k+2} v_1 + \vec{c}_1 \, \vec{\lambda}_1^{\,k+2} \, \vec{v}_1$
- 令$p = -\lambda_1 - \lambda_2, q = \lambda_1\lambda_2$
- $x^{(k+2)} + px^{(k+1)} + qx^{(k)} \approx 0$
- 可用两用两个分量或最小二乘法求出p, q进而解得λ_1，λ_2及
- $x^{(k+2)} - \lambda_2 x^{(k+1)} = c_1(\lambda_1 - \lambda_2)\lambda_1^{k+1} v_1$
- $x^{(k+2)} - \lambda_1 x^{(k+1)} = c_2(\lambda_2 - \lambda_1)\lambda_2^{k+1} v_2$

幂法加速

- 计算λ_1的收敛速度
  - $p = 1$ 近似略去了比$|\lambda_1|^k$小得多的$|\lambda_2|^k$, 计算出的特征值和规一的特征向量误差为$O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$
- 位移加速
  - 矩阵$\boldsymbol{A} - s\boldsymbol{I}$的特征值$\mu_j = \lambda_j - s$
  - $\boldsymbol{A}$对称正定, 取$s = \frac{(\lambda_n+\lambda_2)}{2}, \frac{|\lambda_2-s|}{|\lambda_1-s|} < \left|\frac{\lambda_2}{\lambda_1}\right|$,可用幂法求$\mu_1$再求$\lambda_1$
  - 其它一些情况也可位移加速

反幂法

- 计算模最小的特征值和特征向量
  - A^{-1}的特征值$\mu_j = \frac{1}{\lambda_j}$, 特征向量与$A$同
  - 用幂法求A^{-1}模最大特征值特征向量
    - $x^{(k+1)} = A^{-1}x^{(k)}$: LU分解解方程组$Ax^{(k+1)} = x^{(k)}$
- 位移加速
  - 对$A - sI$用反幂法可得最接近s的特征值
  - 某特征值近似λ^*可取为s用反幂法改善并求特征向量. 通常一二次迭代即可.
  - 结合二分法可求某区间所有特征值特征向量

反幂法改善特征值

- 算法取特征值近似λ^*为s用反幂法改善并求特征向量.

1. 输入矩阵A，近似值λ^*，初始向量x，误差限ε，最大容许迭代次数N.
2. 置$k = 1, \mu = 1$.
3. 作三角分解$A - \lambda^* I = LU$
4. 求$x_r \Rightarrow \alpha, |x_r| = \max\limits_{1\le i\le n}|x_i|$.
5. 计算x的新值

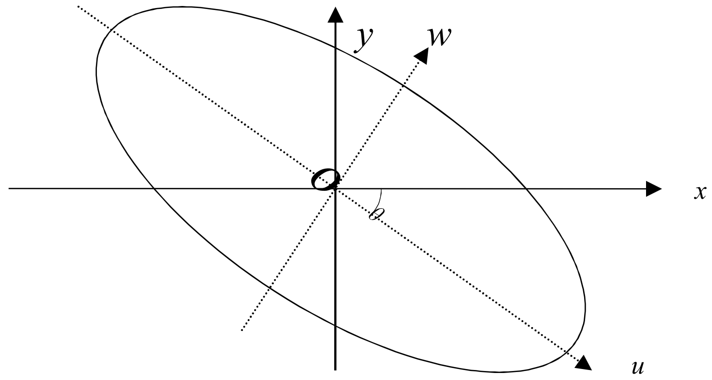
$$y = \frac{x}{\alpha}$$
$$Lz = y$$
$$Ux = z$$

6. 若$\left|\frac{1}{\alpha} - \frac{1}{\mu}\right| < \varepsilon$,则置$\lambda = \lambda^* + \frac{1}{\alpha}$,输出$\lambda, x$,停机；
7. 若$k < N$,则$k + 1 \Rightarrow k, \alpha \Rightarrow \mu$,转步骤4；否则，输出失败信息，停机

降阶

- 利用λ_1, v_1构造新矩阵计算其它特征值
  - 实对称矩阵特征向量组标准正交λ_j, v_j
  - 作$A^{(2)} = A - \lambda_1 v_1 v_1^T$
    - $A^{(2)}$仍以v_j为特征向量，特征值$0, \lambda_2, \dots, \lambda_n$
    - 可用幂法求λ_2, v_2
- 降阶精度将越来越差，实际一般采用的方法是QR结合解耦的算法

平面旋转



- 二次曲线化主轴
  - $f(x,y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$
  - $f(u,w) = b_{11}u^2 + b_{22}w^2$

- 二阶实对称矩阵A对角化

$$A = A^T = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, V = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = V^T A V = B^T$$

  - $$b_{11} = c^2 a_{11} + s^2 a_{22} + 2cs a_{12}$$

$$b_{22} = s^2 a_{11} + c^2 a_{22} - 2cs a_{12}$$

$$b_{12} = -cs(a_{11} - a_{22}) + (c^2 - s^2) a_{12}$$

$$= -(a_{11} - a_{22})/2 \sin 2\theta + a_{12} \cos 2\theta$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix}$$

$$c = cos\theta, s = sin\theta$$

$$\cot 2\theta = \frac{a_{11} - a_{22}}{2a_{12}}$$

$$|\theta| \le \frac{\pi}{4}$$

平面旋转:一般情况

- 一般实对称矩阵对角化-Jacobi方法
  - 一般情况下实对称矩阵平面旋转可化某非对角元为零，之后再化另一非对角元为零时己得零元可能又非零. 但可证明依一定方式化零, 则愈来愈接近对角阵.
  - 典型步:$p-q$平面旋转化pq元为零

$$A^T=A, B=V^TAV=B^T$$
$$b_{pp}=c^2a_{pp}+s^2a_{qq}+2csa_{pq}$$
$$b_{qq}=s^2a_{pp}+c^2a_{qq}-2csa_{pq}$$
$$b_{pq}=-cs(a_{pp}-a_{qq})+(c^2-s^2)a_{pq}$$
$$=-(a_{pp}-a_{qq})\frac{\sin 2\theta}{2}+a_{pq}\cos 2\theta$$
$$b_{pl}=b_{lp}=ca_{pl}+sa_{ql}, l\neq p,q$$
$$b_{ql}=b_{lq}=-sa_{pl}+ca_{ql}$$

$$V=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & -s & & \\ & & & \ddots & & & \\ & & s & & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

$b_{pq}=0$, 当

$$c=\cos\theta, s=\sin\theta$$
$$\cot 2\theta=\frac{a_{11}-a_{22}}{2a_{12}}$$

实对称矩阵Jacobi方法

- 古典Jacobi方法
  - 每次旋转将模最大非对角元化零直到所有非对角元小于误差要求的ε.
    - A变换成B非对角元平方和$E(B) \le cE(A)$，变换k次将下降c^k倍$c = \left(1 - \frac{2}{n(n-1)}\right)$，$c^k \to 0$，当$k \to \infty$
    - 从而证明矩阵序列$AA_1, \dots, A_k, \dots$收敛于对角阵.
- 其它次序
  - 循环Jacobi方法. 逐行逐列将对角线上部非对角元化零, 反复进行 直到所有非对角元小于误差要求的ε
- Jacobi方法适用于小的“满矩阵”全部特征值问题

算例

- 例4. 用Jacobi算法计算$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$的特征值.

解　首先取$i=1, j=2$，得$2\varphi=\frac{\pi}{2}$（或$-\frac{\pi}{2}$），所以$cos\varphi=sin\varphi=\frac{1}{\sqrt{2}}$

$$A_1=V_{12}^T A V_{12}$$

$$=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} 1 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 3 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 \end{bmatrix}$$

算例

再取$i=1$，$j=3$，$\tan 2\varphi=\sqrt{2}$，则$\sin\varphi\approx 0.45969$, $\cos\varphi\approx 0.88808$

$A_2=V_{13}^{T}A_1V_{13}$

$$=\begin{bmatrix}0.88808 & 0 & 0.45969\\ 0 & 1 & 0\\ -0.45969 & 0 & 0.888.8\end{bmatrix}*\begin{bmatrix}1 & 0 & -\frac{1}{\sqrt{2}}\\ 0 & 3 & -\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2\end{bmatrix}*\begin{bmatrix}0.88808 & 0 & -0.45969\\ 0 & 1 & 0\\ 0.45969 & 0 & 0.888.8\end{bmatrix}$$

$$=\begin{bmatrix}0.63398 & -0.32505 & 0\\ -0.32505 & 3 & -0.62797\\ 0 & -0.62797 & 2.36603\end{bmatrix}$$

继续做下去，直到非对角线元素趋向于零，进行9次旋转变换后，得

$$A_9=\begin{bmatrix}0.58578 & 0.00000 & 0.00000\\ 0.00000 & 2.00000 & 0.00000\\ 0.00000 & 0.00000 & 3.41421\end{bmatrix}$$

A_9的对角线元素即是A的特征值，相应的特征向量为$\begin{bmatrix}0.50000 & 0.70710 & 0.50000\\ 0.70710 & 2.00000 & -0.70710\\ 0.50000 & -0.70710 & 0.50000\end{bmatrix}$