7.2 幂法和反幂法

7.2.1 幂法和加速方法

7.2.2 反幂法和原点位移

7.2.1 幂法和加速方法

设矩阵 $A \in R^{n*n}$ 的 n 个特征值满足

$$|\lambda_1| > |\lambda_2| \geq \ldots\ldots \geq |\lambda_n|, \quad (7.2.1)$$

对应的 n 个特征向量 $x_1, x_2, \ldots\ldots, x_n$ 线性无关 .称模最大的特征值 λ_1 为主特征值 ，称对应的特征向量 x_1 为主特征向量 .

幂法用于主特征值和特征向量.它的基本思想是任取一个非零的初始向量 v_0,由矩阵A构造一向量序列

$$v_k = Av_{k-1} = A^k v_0, k = 1,2\ldots..$$

由假设 v_0 可表示为

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \ldots. + \alpha_n x_n. \quad (7.2.2)$$

若记 $(v_k)_i$ 为 v_k 的第 i 个分量，则有

$$v_k = A^k v_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k x_i$$

$$= \lambda_1^k [\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^k x_i] = \lambda_1^k (\alpha_1 x_1 + \varepsilon_K),$$

$$\frac{(v_{k+1})_i}{(v_k)_i} = \frac{\lambda_1 (\alpha_1 x_1 + \varepsilon_{k+1})_i}{(\alpha_1 x_1 + \varepsilon_k)_i},$$

其中 $\varepsilon_k = \sum_{i=2}^{n} \alpha_i (\lambda_i / \lambda_1)^k x_i$。若 $\alpha_1 \neq 0, (x_1)_i \neq 0$，

则由 $\lim_{k \to \infty} \varepsilon_k = 0$ 知

$$\lim_{k \to \infty} \frac{v_k}{\lambda_1^k} = \alpha_1 x_1, \lim_{k \to \infty} \frac{(v_{k+1})_i}{(v_k)_i} = \lambda_1。$$

可见，当 k 充分大时，v_k 近似于主特征值，v_{k+1} 与 v_k 的对应非零分量的比值近似于主特征值。

在实际计算中,需要对计算结果进行规范化。因为当$|\lambda_1|<1$时，v_k趋于零，当$|\lambda_1|>1$时,v_k的非零分量趋于无穷。 从而计算时会出现下溢 或上溢。为此,对$Z=(z_1,z_2,\ldots\ldots,z_n)^T\in R^n$,记$\max(Z)=z_i$,其中$|z_i|=\|Z\|_\infty$.这样,我们有

如下幂法的实用的计算公式:

$$\begin{cases} v_0 = u_0 \neq 0, \\ v_k = Au_{k-1}, \\ u_k = v_k / \max(v_k), k=1,2,\ldots。 \end{cases} \qquad \textbf{(7.2.3)}$$

定理 7.4设$A\in R^{n*n}$的特征值 $\lambda_i\,(i=1,2,\ldots,n)$满足(7.2.1)并且有对应的$n$个线性无关的特征向量 $x_i\,(i=1,2,\ldots n)$。给定初值向量 $v_0=\sum_{i=1}^{n}\alpha_i x_i$, $\alpha_i\neq 0$，则由（7.2.3)生成的向量序列有

$$\lim_{k\to\infty} u_k = \frac{x_1}{\max(x_1)}, \lim_{k\to\infty}\max(v_k)=\lambda_1。$$

证：由(7.2.3)有

由(7.2.2)有 $v_k = \dfrac{A^k v_0}{\max(A^{k-1}v_0)}, u_k = \dfrac{A^k v_0}{\max(A^k v_0)}$。

$$A^k v_0 = \lambda_1^k[\alpha_1 x_1 + \sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k x_i] = \lambda_1^k(\alpha_1 x_1 + \varepsilon_k),$$

$$u_k = \frac{A^k v_0}{\max(A^k v_0)} = \frac{\lambda_1^k(\alpha_1 x_1 + \varepsilon_k)}{\max[\lambda_1^k(\alpha_1 x_1 + \varepsilon_k)]}$$

$$= \frac{\alpha_1 x_1 + \varepsilon_k}{\max(\alpha_1 x_1 + \varepsilon_k)} \to \frac{x_1}{\max(x_1)}(k \to \infty)。$$

同理,可得到 $v_k = \dfrac{\lambda_1^k(\alpha_1 x_1 + \varepsilon_k)}{\max[\lambda_1^{k-1}(\alpha_1 x_1 + \varepsilon_{k-1})]} = \dfrac{\lambda_1(\alpha_1 x_1 + \varepsilon_k)}{\max(\alpha_1 x_1 + \varepsilon_{k-1})},$

$$\max(v_k) = \lambda_1 \frac{\max(\alpha_1 x_1 + \varepsilon_k)}{\max(\alpha_1 x_1 + \varepsilon_{k-1})} \to \lambda_1(k \to \infty)。$$

定理得证。

由定理的证明可见，幂法的收敛速度由 $|\lambda_2/\lambda_1|$的大小确定。若 A的特征值不满足（7.2.1)，将有不同的情况。如果$\lambda_1=\lambda_2=\ldots=\lambda_r$，且$|\lambda_r|>|\lambda_{r+1}|$，可以作类似的分析，对初始向量(7.2.2)和计算公式(7.2.3)有

$$\lim_{k\to\infty} u_k=\frac{\sum_{i=1}^{r}\alpha_i x_i}{\max(\sum_{i=1}^{r}\alpha_i x_i)},\ \lim_{k\to\infty}\max(v_k)=\lambda_1。$$

可见，u_k仍收敛于一个主特征向量。对特征值的其他情况，讨论较为复杂。完整的幂法程序要加上各种情况的判断。

例7.1 用幂法求矩阵

$$A=\begin{pmatrix}1 & 1 & 0.5\\ 1 & 1 & 0.25\\ 0.5 & 0.25 & 2\end{pmatrix}$$

的主特征值和主特征向量.

解：取初始向量 $u_0=(1,1,1)^T$，按(7.2.3)的计算结果如表 7－1。

表7-1

K	u_k^T	$\max(v_k)$
0	(1.0000,1.0000,1)	
1	(0.9091,0.8182,1)	2.7500000
5	(0.7651,0.6674,1)	2.5887918
10	(0.7494,0.6508,1)	2.5380029
15	(0.7483,0.6497,1)	2.5366256
20	(0.7482,0.6497,1)	2.5365323

矩阵 A 的主特征值和特征向量 的准确值(8位数字)分别为 $\lambda_1=2.5365258$，$x_1^*=(0.74822116,0.64966116,1)^T$。可见迭代20次后，所得的主特征值 有5位有效数字。

从定理7.4的证明中易见，当k充分大时，有$\left|\max(v_k)-\lambda_1\right| \approx c\left|\lambda_2/\lambda_1\right|^k$。因此，幂法是线性收敛的方法。当$|\lambda_2|$接近于$|\lambda_1|$时，收敛很慢。这时，一个补救的办法是采用加速收敛的方法。下面简要的介绍两种加速方法。

1.*Aitken*外推法

记$m_k=\max(v_k)$。对幂法的计算结果进行外推加速：

$$\tilde{m}_k = m_m - \frac{(m_k - m_{k-1})^2}{m_k - 2m_{k-1} + m_{k-2}}, \quad k \geq 3,$$

$$(\tilde{u}_k)_j = (u_k)_j - \frac{((u_k)_j - (u_{k-1})_j)^2}{(u_k)_j - 2(u_{k-1})_j + (u_{k-2})_j}, (u_k)_j \neq 1。$$

2. Rayleigh商加速

若$A \in R^{n*n}$为对称矩阵，则幂法所得的规范化向量u_k的*Rayleigh*商给出特征值λ_1较好的近似值，

$$\frac{(Au_k, u_k)}{(u_k, u_k)} = \lambda_1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right)。$$

7.2.2 反幂法和原点位移

反幂法用来计算矩阵按模最小的特征值及其特征向量。设$A \in R^{n*n}$为非奇异矩阵，它的特征值满足

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_{n-1}| > |\lambda_n| > 0, \quad \textbf{(7.2.4)}$$

则A^{-1}的特征值 λ_1^{-1}，λ_2^{-1},....., λ_n^{-1}满足

$$|\lambda_n^{-1}| > |\lambda_{n-1}^{-1}| \geq \ldots\ldots \geq |\lambda_1^{-1}|,$$

即λ_n^{-1}是A^{-1}的主特征值。因此，对A^{-1}应用幂法可得矩阵A的按模最小的特征值及其特征向量，称为反幂法，计算公式为

$$\begin{cases} v_0 = u_0 \neq 0, \\ v_k = A^{-1} u_{k-1}, \\ u_k = v_k / \max(v_k), k = 1,2\ldots. \end{cases} \quad \textbf{(7.2.5)}$$

在(7.2.5)中，向量v_k可以通过解方程组 $Av_k = u_{k-1}$得到。

定理7.5: 设非奇异矩阵$A \in R^{n*n}$的特征值$\lambda_i (i=1,2,\ldots\ldots,n)$满足(7.2.4)，并且有对应的$n$个线性无关的特征向量$x_i (i=1,2,\ldots,n)$。给定初始向量$v_0=\sum_{i=1}^{n}\alpha_i x_i, \alpha_n \neq 0$, 则由(7.2.5)生成的向量序列有

$$\lim_{k\to\infty} u_k = \frac{x_n}{\max(x_n)}, \lim_{k\to\infty}\max(v_k)=\frac{1}{\lambda_n}。$$

反幂法的一个重要应用 是利用“原点位移”， 求指定点附近的某个特 征值和对应的特征向量。

如果矩阵$(A-PI)^{-1}$存在，显然其特征值为$(\lambda_i - p)^{-1}, i=1,2,\ldots,n$, 对应的特征向量仍然是$x_i (i=1,2,\ldots,n)$。如果$p$是$A$的特征值$\lambda_j$的一个近似值，且

$$\left|\lambda_j - p\right| < \left|\lambda_i - p\right|, i \neq j, \qquad (7.2.6)$$

即$(\lambda_j - p)^{-1}$是$(A-PI)^{-1}$的主特征值，可用反幂法计算相应的特征值和特征向量，计算公式为

$$\begin{cases} u_0 = v_0 \neq 0, \\ v_k = (A - PI)^{-1} u_{k-1}, \\ u_k = v_k / \max(v_k), k = 1,2,\ldots。\end{cases} \qquad (7.2.7)$$

定理7.6设$A \in R^{n*n}$的特征值$\lambda_i (i = 1,2,\ldots,n)$对应的特征向量$x_i$ $(i = 1,2,\ldots,n)$线性无关，p为λ_j的近似值，满足(7.2.6),$(A - PI)^{-1}$存在。

给定初始向量$v_0 = \sum_{i=1}^{n} \alpha_i x_i, \alpha_j \neq 0$,则由(7.2.7)生成的向量序列有

$$\lim_{k \to \infty} u_k = \frac{x_j}{\max(x_j)}, \lim_{k \to \infty} \max(v_k) = \frac{1}{\lambda_j - p}。$$

由该定理可知，$p + [\max(v_k)]^{-1}$是特征值λ_j的近似值，对应的近似 特征向量为 u_k。迭代收敛速度由比值 $\sigma = \max_{i \neq j} \left| (\lambda_j - p)/(\lambda_i - p) \right|$来确定。

反幂法迭代公式 (7.2.7)中的 v_k是通过解方程组

$$(A - pI) v_k = u_{k-1}$$

求得的，为了节省计算 工作量，可以先将 $(A-PI)$进行三角分解

$$P(A-pI)=LU,$$

其中 P 为排列阵。

只要选择的P是λ_i的一个较好的近似且特征值分离情况较好，一般σ很小，收敛将是较快的。 实验表明，按下述方法 选择$v_0=u_0$是较好的：选u_0使

$$Uv_1=L^{-1}Pu_0=(1,1,\dots,1)^T,$$

用回代求解可得 v_1。

例7.2 用反幂法求下列矩阵的接近于**P=1.2679**的特征值(精确特征值$\lambda_3=3-\sqrt{3}$)及其特征向量(用5位浮点数进行计算),

$$A=\begin{pmatrix}2&1&0\\1&3&1\\0&1&4\end{pmatrix}。$$

解：用列选主远元的三角分 解将$A-pI$分解为

$$P(A-pI)=LU,$$

其中

$$P=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix},L=\begin{pmatrix}1&0&0\\0&1&0\\0.7321&-0.26807&1\end{pmatrix},$$

$$U=\begin{pmatrix}1&1.7321&1\\0&1&2.7321\\0&0&0.29405\times10^{-3}\end{pmatrix}。$$

由$Uv_1=(1,1,1)^T$得

$$v_1=(12692,-9290.3,3400.8)^T,$$

$$u_1=(1,-0.73198,0.26795)^T。$$

由$LUv_2=Pu_1$得

$$v_2=(20404,-14937,5467.4)^T,$$

$$u_2=(1,-0.73206,0.26796)^T。$$

由此可得特征值 $\lambda_3 (=1.2679492)$的近似值

$$1.2679+\frac{1}{20404}=1.267949$$

λ_3对应的特征向量是

$$x_3=(1,1-\sqrt{3},2-\sqrt{3})^T\approx(1,-0.73205,0.26795)^T。$$

由此可见，u_2是x_3的相当好的近似。