

# 第八章

## 矩阵特征值与特征向量的计算

**§8.0 问题描述**

**§8.1 乘幂法与反幂法**

## §4.0 问题描述

设 $A$ 为 $n \times n$ 矩阵, 所谓 $A$ 的特征问题是求数 $\lambda$ 和非零向量 $x$ , 使

$$Ax = \lambda x$$

成立。数 $\lambda$ 叫做 $A$ 的一个特征值, 非零向量 $x$ 叫做与特征值 $\lambda$ 对应的特征向量。这个问题等价于求使方程组 $(A - \lambda I)x = 0$ 有非零解的数 $\lambda$ 和相应的非零向量 $x$ 。

线性代数理论中是通过求解特征多项式 $\det(A - \lambda I) = 0$ 的零点而得到 $\lambda$ , 然后通过求解退化的方程组 $(A - \lambda I)x = 0$ 而得到非零向量 $x$ 。当矩阵阶数很高时, 这种方法极为困难。目前用数值方法计算矩阵的特征值以及特征向量比较有效的方法是迭代法和变换法。

## §4.1 乘幂法与反幂法

### 一、乘幂法

通过求矩阵特征向量求出特征值的一种迭代方法，它用以求按模最大的特征值和相应的特征向量。

设实矩阵A的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$ ，相应的特征向量 $x_1, x_2, \dots, x_n$ 线性无关。设A的特征值按模排序为：

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

则对任一非零向量  $U^{(0)} \in R^n$ ，可以得到：

$$U^{(0)} = a_1 x_1 + a_2 x_2 + \dots + a_n x_n = \sum_{j=1}^n a_j x_j$$

根据特征值的定义  $Ax_i = \lambda_i x_i$

令  $U^{(k+1)} = AU^{(k)}, k = 0, 1, 2, \dots$ ，可以构造一个向量序列，

$$U^{(1)} = AU^{(0)} = \lambda_1 a_1 x_1 + \lambda_2 a_2 x_2 + \dots + \lambda_n a_n x_n = \sum_{j=1}^n \lambda_j a_j x_j$$

$$U^{(2)} = AU^{(1)} = \lambda_1^2 a_1 x_1 + \lambda_2^2 a_2 x_2 + \cdots + \lambda_n^2 a_n x_n = \sum_{j=1}^n \lambda_j^2 a_j x_j$$

$$\vdots$$

$$U^{(k)} = AU^{(k-1)} = \lambda_1^k a_1 x_1 + \lambda_2^k a_2 x_2 + \cdots + \lambda_n^k a_n x_n = \sum_{j=1}^n \lambda_j^k a_j x_j$$

$$U^{(k)} = A^k U^{(0)} = \lambda_1^k (a_1 x_1 + \sum_{j=2}^n (\frac{\lambda_j}{\lambda_1})^k a_j x_j)$$

若  $a_1 \neq 0$ , 由于  $\left| \frac{\lambda_i}{\lambda_1} \right| < 1 (i \geq 2)$ , 故  $k$  充分大时,

$$U^{(k)} = \lambda_1^k (a_1 x_1 + \varepsilon_k) \approx \lambda_1^k a_1 x_1$$

**$U^{(k)}$  是相应于  $\lambda_1$  的近似特征向量**

**设  $U_l^{(k)}$  表示  $U^{(k)}$  的第  $l$  个分量。**

$$\frac{U_l^{(k)}}{U_l^{(k-1)}} = \frac{\lambda_1^k a_1 (x_1)_l + (\varepsilon_k)_l}{\lambda_1^{k-1} a_1 (x_1)_l + (\varepsilon_{k-1})_l} \approx \lambda_1, l = 1, 2, \cdots, n$$

综上所述，求矩阵主特征值及相应的特征向量的**计算步骤**如下：

Step1:任给  $n$  维初始向量  $U^{(0)} \neq 0$ ;

Step2:按  $U^{(k)} = AU^{(k-1)}$  ( $k=1,2,\dots$ ) 计算  $U^{(k)}$ ;

Step3:如果  $k$  从某个数后分量比

$$\frac{U_l^{(k)}}{U_l^{(k-1)}} \approx c(\text{常数}), l = 1, 2, \dots, n$$

则取  $\lambda_1 \approx c$ ，而  $U^{(k)}$  就是与  $\lambda_1$  对应的一个近似特征向量。

上述方法即**乘幂法**。

**Remark1:**具体计算时,  $U^{(0)}$ 的选取很难保证一定有  $\alpha_1 \neq 0$ 。但是, 由于舍入误差的影响, 只要迭代次数足够多, 如  $U^{(m)} = a'_1 x_1 + a'_2 x_2 + \cdots + a'_n x_n$ , 就会有  $a'_1 \neq 0$ , 因而最后结论是成立的。对于  $U_l^{(k)} = 0$  的情形, 由于对任意  $l$  均有上面的结论, 故只要取另外的使  $U_l^{(k)} \neq 0$  即可。

**Remark2:**以上讨论只是说明了乘幂法的基本原理。当  $|\lambda_1|$  太小或太大时, 将会使  $U^{(k)}$  分量的绝对值过小或过大, 以致运算无法继续进行。因此, 实际计算时, 常常是每进行  $m$  步迭代进行一次规范化, 如用

$$V^{(m)} = U^{(m)} / \max (U^{(m)})$$

其中,  $\max(U^{(m)})$  表示向量  $U^{(m)}$  的绝对值最大的分量。

代替  $U^k$  继续迭代。由于特征向量允许差一个非零常数因子，因而从  $U^k$  往后继续迭代与从  $U^k$  往后继续迭代的收敛速度是相同的，但规范化的做法有效防止了溢出现象。至于  $m$  的选取，可以自由掌握，如取  $m=1, 5$  等等。

**Remark3:**若主特征值是重特征值，如

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

则有

$$U^{(k)} = \lambda_1^k \left( \sum_{i=1}^r a_i x_i + \sum_{i=r+1}^n \left( \frac{\lambda_i}{\lambda_1} \right)^k a_i x_i \right)$$

从而

$$\frac{U_l^{(k)}}{U_l^{(k-1)}} = \lambda_1 \left[ \frac{\sum_{i=1}^r a_i (x_i)_l + \sum_{i=r+1}^n a_i \left( \frac{\lambda_i}{\lambda_1} \right)^k (x_i)_l}{\sum_{i=1}^r a_i (x_i)_l + \sum_{i=r+1}^n a_i \left( \frac{\lambda_i}{\lambda_1} \right)^{k-1} (x_i)_l} \right] \xrightarrow[k \rightarrow \infty]{} \lambda_1, l = 1, 2, \cdots, n$$

由此可得乘幂法的算法。但是应该注意到，在重特征值的情形下，从不同的非零初始向量出发迭代，可能得到主特征值的几个线性无关的特征向量。

**Remark4:**由上述推导可知，乘幂法收敛的快慢取决于比值  $|\lambda_2|/|\lambda_1|$  的大小，该比值越小收敛越快。

由此便提出了乘幂法的加速收敛方法，如 **Rayleigh** 商加速法、原点平移法等。

**Remark5:**对于  $\lambda_1 = -\lambda_2$ ，或  $\lambda_1$  与  $\lambda_2$  共轭等情形，也可类似进行计算，具体可参阅相关教材。



## 二、反幂法

计算矩阵按模最小的特征值及相应的特征向量。

设矩阵A非奇异，用  $A^{-1}$  代替A作幂的方法就成为反幂法。当A的特征值满足  $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$  时， $A^{-1}$  的特征值满足  $\left| \frac{1}{\lambda_1} \right| \leq \left| \frac{1}{\lambda_2} \right| \leq \dots \leq \left| \frac{1}{\lambda_n} \right|$ ，并且A对应于 $A^{-1}$ 的相应的特征向量是相同的。

对  $A^{-1}$  用反幂法求解按模最大的特征值是  $\frac{1}{\lambda_n}$ ，特征向量是  $x_n$ ，即是A的按模最小的特征值和特征向量。

反幂法的**计算步骤**如下:

Step1: 任取  $U^{(0)} \neq 0$ ;

Step2: 计算  $U^{(k)} = A^{-1} U^{(k-1)} (k=1, 2, \dots)$ ;

Step3: 如果  $k$  从某个数后分量比

$$\frac{U_l^{(k)}}{U_l^{(k-1)}} \approx c (\text{常数}), l = 1, 2, \dots, n$$

则取  $\lambda_n \approx \frac{1}{c}$ , 而  $U^{(k)}$  就是与  $\lambda_n$  对应的一个近似特征向量。

**Remark1:** 实际计算时一般并不求  $A^{-1}$ ，而是将算法中的迭代公式  $U^{(k)} = A^{-1} U^{(k-1)}$  改为解方程组  $A U^{(k)} = U^{(k-1)}$ 。由于每步所解方程组具有相同的系数矩阵  $A$ ，故常常是先将  $A$  进行三角分解，然后转化为每步只需用回代公式求解两个三角方程组。这样可以减少计算工作量。

**Remark2:** 若已知矩阵  $A$  的某个特征值  $\lambda_i$  的相对分离较好的近似值  $p$ 。不要求  $p$  的近似程度有多好，只要求  $j \neq i$  时， $|\lambda_i - p| < |\lambda_j - p|$ ，则  $\frac{1}{\lambda_i - p}$  便是  $(A - pI)^{-1}$  的主特征值。

这样一来，就可以使用反幂法求解矩阵的在某点附近的特征值及其特征向量。

**本课程到此全部结束，**

**祝大家考试取得好成绩！**