第四章 求矩阵特征值和特征向量的方法

讲授：

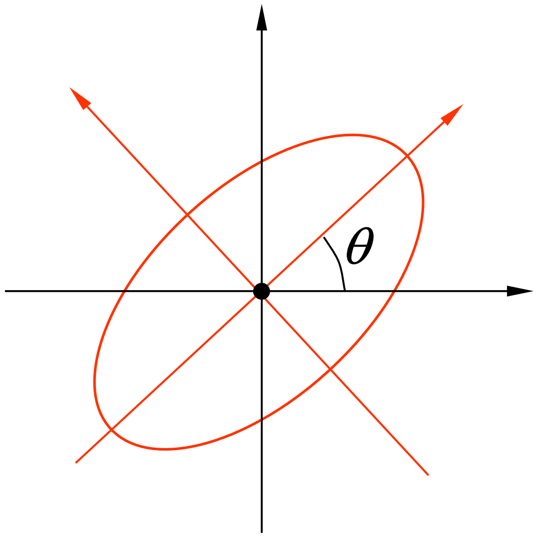
求矩阵特征值及特征向量的常用数值方法的构造和原理。

重点论述：

重点论述幂法的构造内容。

§4.2 基本概念

幂法？反幂法？



1、矩阵A的特征多项式：

$$f_A(\lambda)=\det(A-\lambda I)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}, A\in R^{n\times n}.$$

2、矩阵A的特征多项式方程：

$$f_A(\lambda)=\det(A-\lambda I)=0$$

3、矩阵A的特征值：

$$f_A(\lambda)=\det(A-\lambda I)=0\text{的根}$$

4、特征值与特征向量的关系：

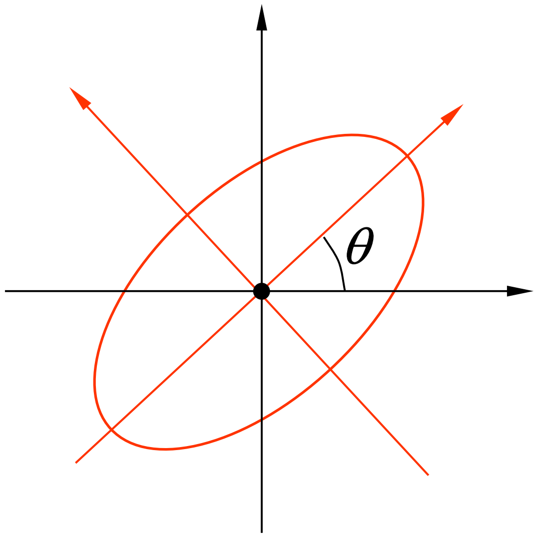
$$A\vec{x}=\lambda\vec{x}\,,\vec{x}\neq 0$$

问题

特征方程是一个n次代数方程，求根很复杂且特征方程对舍入误差很敏感，特别当n较大时，这些问题更突出。

由于这些原因，实用中在求解代数特征值问题时一般不用如上的线性代数的方法，而采用本章介绍的**迭代加变换的计算机求解**方法，这些方法具有编程简单，对舍入误差不敏感等优点。

§4.3 幂法



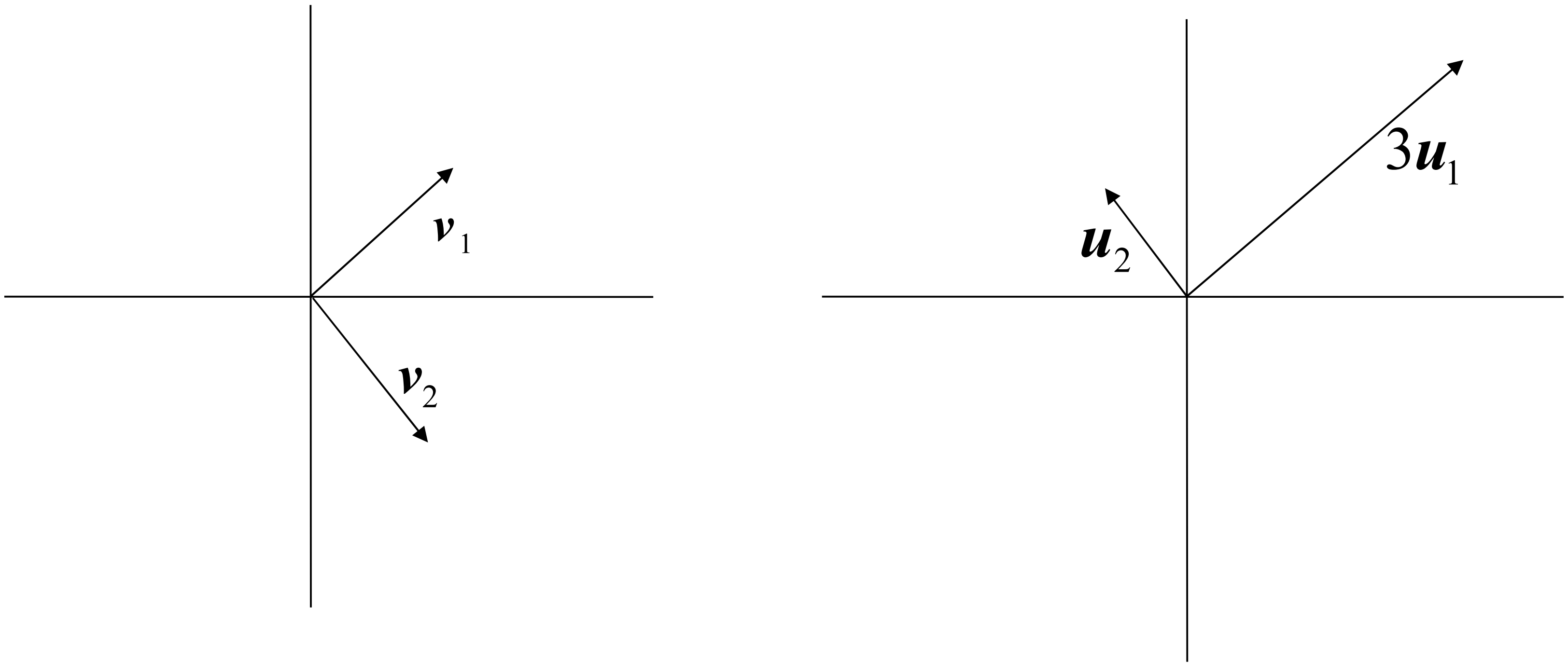
一、基本思想

利用矩阵的特征值与特征向量的关系构造迭代向量序列来求矩阵按模最大的特征值及其相应特征向量。

幂法作用

求矩阵按模最大的特征值及其相应特征向量。

矩阵作用效果



二、构造原理

设 $x^{(1)},x^{(2)},\dots,x^{(n)}$是A的$n$个线性无关的特征向量，$\lambda_1,\lambda_2,\dots,\lambda_n$是对应的特征值。

$$\forall\ V^{(0)} \neq 0 \in R^n, \Rightarrow V^{(0)} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \dots + \alpha_n x^{(n)}$$

用A左乘，并利用$Ax^{(k)} = \lambda_k x^{(k)}$，有

$$AV^{(0)} = \alpha_1 Ax^{(1)} + \alpha_2 Ax^{(2)} + \dots + \alpha_n Ax^{(n)}$$

$$= \alpha_1\lambda_1 x^{(1)} + \alpha_2\lambda_2 x^{(2)} + \dots + \alpha_n\lambda_n x^{(n)}$$

记$V^{(k)} = AV^{(k-1)}$，有

$$V^{(k)} = A^k V^{(0)} = \alpha_1\lambda_1^k x^{(1)} + \alpha_2\lambda_2^k x^{(2)} + \cdots + \alpha_n\lambda_n^k x^{(n)}$$

$$= \lambda_1^k \left(\alpha_1 x^{(1)} + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x^{(2)} + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k x^{(n)} \right)$$

假设$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|, \alpha_1 \neq 0$

$$\because \left| \lambda_k / \lambda_1 \right| < 1 \Rightarrow V^{(k)} \Rightarrow \lambda_1^k \alpha_1 x^{(1)}, k \to \infty$$

考虑向量V的分量比，对于第i个分量，有

$$\frac{V_i^{(k)}}{V_i^{(k-1)}} = \lambda_1 \frac{\alpha_1 x_i^{(1)} + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k x_i^{(2)} + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k x_i^{(n)}}{\alpha_1 x_{i-1}^{(1)} + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k x_{i-1}^{(2)} + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k x_{i-1}^{(n)}} \xrightarrow{k \to \infty} \lambda_1$$

$$\boxed{\Rightarrow \frac{V_i^{(k)}}{V_i^{(k-1)}} \approx \lambda_1}$$

$\because V^{(k)} \approx \alpha_1 \lambda_1^k x^{(1)}$，故$V^{(k)}$是$\lambda_1$对应的近似特征向量。

三、分析

$$\because |\lambda_1| > 1 \Rightarrow \lambda_1^k \to \infty \Rightarrow V^{(k)}$$出现上溢错误。

定理：假设矩阵A的n个特征值满足$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$，任取非零向量$V^{(0)} \in R^n$，有计算格式

$$u^{(0)} = V^{(0)}, \quad \begin{cases} V^{(k)} = Au^{(k-1)} \\ m_k = \max\left(V^{(k)}\right) \quad k = 1, 2, \cdots \\ u^{(k)} = V^{(k)} / m_k \end{cases}$$

规范化的思想

$$\Rightarrow \underline{\lim_{k \to \infty} u^{(k)} = \frac{x^{(1)}}{\max\left(x^{(1)}\right)}}, \qquad \underline{\lim_{k \to \infty} m_k = \lambda_1}$$

式中$\max\{V^{(k)}\}$表示$V^{(k)}$绝对值最大的分量，$\{u^{(k)}\}$是规范化向量。

证明 $\because V^{(1)} = Au^{(0)} = AV^{(0)},$

$$u^{(1)} = \frac{V^{(1)}}{\max V^{(1)}} = \frac{Au^{(0)}}{\max\left(Au^{(0)}\right)} = \frac{AV^{(0)}}{\max\left(AV^{(0)}\right)}$$

$$V^{(2)} = Au^{(1)} = A\frac{AV^{(0)}}{\max\left(AV^{(0)}\right)} = \frac{A^2V^{(0)}}{\max\left(AV^{(0)}\right)}$$

$$u^{(2)} = \frac{V^{(2)}}{\max\left(V^{(2)}\right)} = \frac{A^2V^{(0)}}{\max\left(V^{(2)}\right)\cdot\max\left(AV^{(0)}\right)} = \frac{A^2V^{(0)}}{\max\left(A^2V^{(0)}\right)}$$

一般的有

$$\boxed{V^{(k)} = \frac{A^kV^{(0)}}{\max\left(A^{k-1}V^{(0)}\right)}, u^{(k)} = \frac{A^kV^{(0)}}{\max\left(A^kV^{(0)}\right)}}$$

记$\varepsilon^{(k)} = \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k x^{(i)} \Rightarrow \lim_{k\to\infty} \varepsilon^{(k)} = 0, \because \left|\lambda_k / \lambda_1\right| < 1$

再由$A^kV^{(0)}=\lambda_1^k\left(\alpha_1 x^{(1)}+\varepsilon^{(k)}\right)$ **有**

$$\lim_{k\to\infty} u^{(k)}=\lim_{k\to\infty}\frac{\lambda_1^k\left(\alpha_1 x^{(1)}+\varepsilon^{(k)}\right)}{\max\left(\lambda_1^k\left(\alpha_1 x^{(1)}+\varepsilon^{(k)}\right)\right)}=\frac{x^{(1)}}{\max\left(x^{(1)}\right)};$$

$$\lim_{k\to\infty} m_k=\lim_{k\to\infty}\max\left(\frac{A^kV^{(0)}}{\max\left(A^{k-1}V^{(0)}\right)}\right)$$

$$=\lim_{k\to\infty}\frac{\lambda_1^k\left(\alpha_1 x^{(1)}+\varepsilon^{(k)}\right)}{\max\left(\lambda_1^{k-1}\left(\alpha_1 x^{(1)}+\varepsilon^{(k-1)}\right)\right)}=\lambda_1$$

规范化幂法算法

1) 输入矩阵$A, V^{(0)}$和精度ε, 使用中取$V^{(0)}=\{1,1,\cdots,1\}$

2）$k \Leftarrow 1$

3）$V^{(k)} \Leftarrow Au^{(k-1)}$

4）$m_k \Leftarrow \max\left(V^{(k)}\right), m_{k-1} \Leftarrow \max\left(V^{(k-1)}\right)$

5）$u^{(k)} \Leftarrow V^{(k)} / \mathrm{m}_k$

6）如果$\left|\mathrm{m}_{\mathrm{k}}-\mathrm{m}_{\mathrm{k}-1}\right| < \varepsilon$, 则输出$u^{(k)}, m_k$, 停止

7）$k \Leftarrow k+1, goto \quad 3)$

【例 4-1】 用幂法求 $\boldsymbol{A}=\begin{bmatrix}133 & 6 & 135\\ 44 & 5 & 46\\ -88 & -6 & -90\end{bmatrix}$ 按模最大的特征值及其特征向量，要求误差 $<10^{-4}$.

解 取 $\boldsymbol{V}^{(0)}=\{1,1,1\}^{\mathrm{T}}$，按规范化幂法计算得下表结果.

迭代次数 k	m_k	$u^{(k)}$	$\lvert m_k-m_{k-1}\rvert$
0		(1, 1, 1)	
1	44.42335766	{1., 0.3467153285, −0.6715328467}	229.5766423
2	44.92343082	{1., 0.3341275058, −0.6672691423}	0.5000731606
3	44.99546459	{1., 0.3333729572, −0.6667020234}	0.07203376236
4	44.99977337	{1., 0.3333351894, −0.6666684279}	0.004308781874
5	44.99998937	{1., 0.3333334179, −0.6666667492}	0.0002160020115
6	44.99999952	{1., 0.3333333371, −0.6666666704}	0.0000101441501

此结果说明迭代 6 次，求得误差为 0.0000101441501 的按模最大的特征值为 44.99999952 及其对应的一个特征向量为 {1,0.3333333371,−0.6666666704}.

本题矩阵 $\boldsymbol{A}$ 的三个特征值为 {45,2,1}，可见所求结果很好.

思考：如何求矩阵按模最小的特征值及特征向量

反幂法

用幂法求矩阵按模最小的特征值的方法

如果有$|\lambda_1| \geq \ldots \geq |\lambda_2| > |\lambda_n| > 0 \Rightarrow A^{-1}$存在

由$Ax^{(k)} = \lambda_k x^{(k)} \Rightarrow A^{-1}x^{(k)} = \lambda_k^{-1}x^{(k)}$

说明λ_k^{-1}是A^{-1}的特征值，$x^{(k)}$是其对应的特征向量。

$$|\lambda_1| \geq \ldots \geq |\lambda_2| > |\lambda_n| > 0 \Rightarrow |\lambda_n^{-1}| > |\lambda_{n-1}^{-1}| \geq \ldots \geq |\lambda_1^{-1}|$$

对A^{-1}用幂法，可以求出λ_n^{-1}，

继而求出安模最小的特征值λ_n。

避免矩阵求逆

反幂法迭代公式 $\mathbf{V}^{(k)} = \mathbf{A}^{-1}\ \boldsymbol{u}^{(k-1)}$

用解线性方程组 $\mathbf{AV}^{(k)} = \boldsymbol{u}^{(k-1)}$ 的方法求得 $\mathbf{V}^{(k)}$

思考：如何求矩阵所有的特征值及特征向量

$$V^{(k)} = A^k V^{(0)} = \alpha_1 \lambda_1^k x^{(1)} + \alpha_2 \lambda_2^k x^{(2)} + \cdots + \alpha_n \lambda_n^k x^{(n)}$$

思考：已知在实数s附近有特征值，如何求此特征值

Inverse Power Iteration

Given initial vector x_0 and shift s

for $j = 1, 2, 3, \ldots$

$u_{j-1} = x_{j-1}/||x_{j-1}||_2$

Solve $(A - sI)x_j = u_{j-1}$

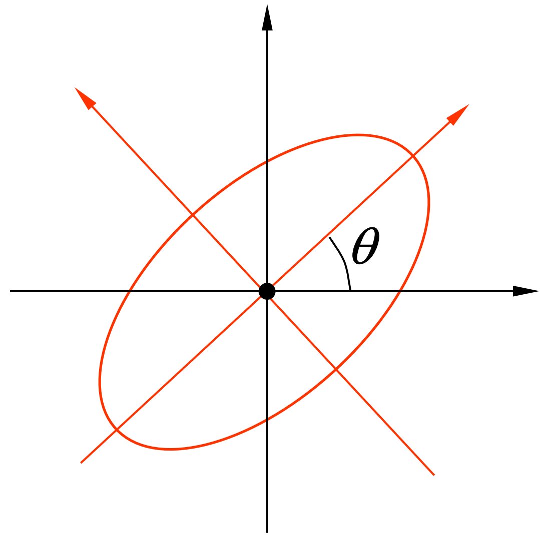
$\lambda_j = u_{j-1}^T x_j$

end

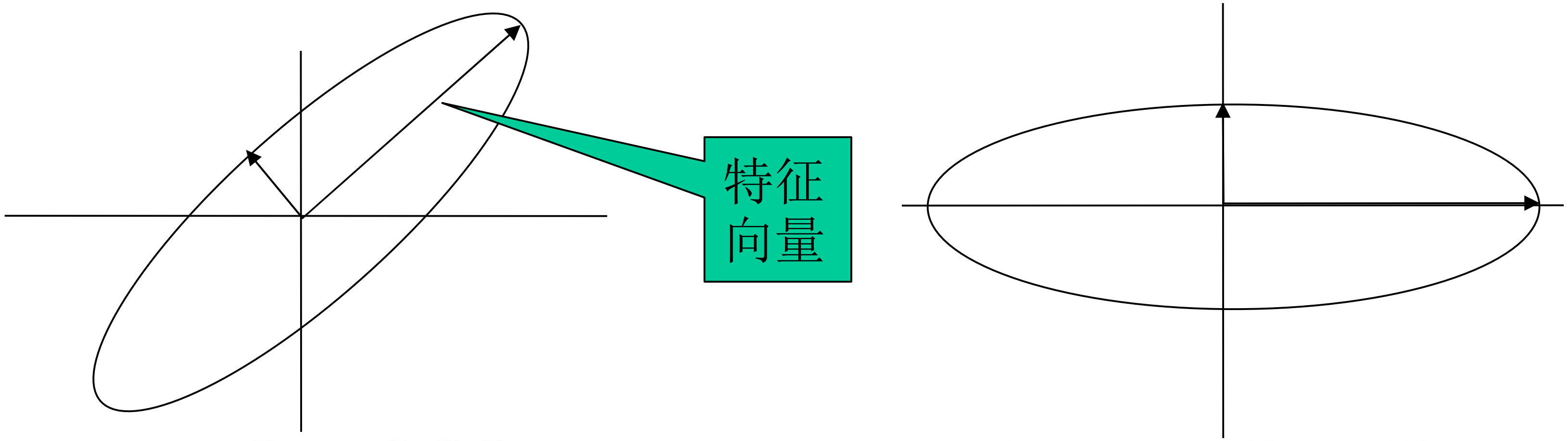
$u_j = x_j/||x_j||_2$

思考：算法的收敛速度

§4.4 Jacobi方法



Jacobi方法又称旋转法，是求实对称矩阵特征值和特征向量的方法



$$f(x_1,x_2)=[x_1,x_2]\begin{bmatrix}a_{11} & a_{12}\\ a_{12} & a_{22}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

二阶旋转变换

$$\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix},|\varphi|\leqslant\frac{\pi}{4}$$

$$f(x_1,x_2)=[y_1,y_2]\begin{bmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{bmatrix}\begin{bmatrix}a_{11} & a_{12}\\ a_{12} & a_{22}\end{bmatrix}\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix}$$

$$=[y_1,y_2]\begin{bmatrix}b_{11} & 0\\ 0 & b_{22}\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix}$$

二次型 ⇕ **实对称矩阵**

特殊二次型 ⇕ **实对角矩阵**

N阶旋转矩阵，
也叫Givens矩阵

$$J(i,j,\varphi)=\begin{array}{cc} \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & \cos\varphi & & & & -\sin\varphi & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & \sin\varphi & & & & \cos\varphi & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ i\text{行} \\ \\ \\ \\ j\text{行} \\ \\ \\ \\ \end{array} \\ \begin{array}{ccccccccccc} & & & \quad i\text{列} & & & & \quad j\text{列} & & & \end{array} & \end{array}$$

单位正交矩阵

N阶实对称矩阵，对应N元二项式；也可通过一系列的旋转将N维空间的椭球面的轴对齐到坐标轴。

$$\dots J_n^T \dots J_2^T J_1^T A J_1 J_2 \dots J_n \dots = \Lambda$$

定义$J_1 J_2 \dots J_n \dots = Q$,则有$Q^T A Q = \Lambda$,

同时$Q^T Q = I$, $Q^T = Q^{-1}$,故$Q^{-1} A Q = \Lambda$

$$AQ = Q\Lambda$$

$$AQ_j = \lambda_{ii} Q_j$$

实对称矩阵一定正交相似于对角阵；对角元素为特征值；正交矩阵列向量为对应的特征向量；

一、基本思想

将实对称矩阵进行一系列的相似正交变换使其约化成一个近似对角矩阵，然后利用相似正交变换的关系来求全部特征值和特征向量。

Jacobi方法的作用

求**实对称矩阵**的**全部特征值**和特征向量。

由于使用的正交相似变换主要采用旋转变换，故称Jacobi方法为**旋转法**。

1、$J(i,j,\varphi)$的特点**是正交矩阵**

2、选择合适的φ对 A做相似变换

$$J^T(i,j,\varphi)AJ(i,j,\varphi)=A^{(1)}=\left(a_{ij}^{(1)}\right),$$ 可使$a_{ij}^{(1)}=a_{ji}^{(1)}=0$

$$\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}=\begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix}$$

$$b_{11}=a_{11}\cos^2\varphi+a_{22}\sin^2\varphi+a_{12}\sin 2\varphi$$
$$b_{12}=0.5(a_{22}-a_{11})\sin 2\varphi+a_{12}\cos 2\varphi$$
$$b_{22}=a_{11}\sin^2\varphi+a_{22}\cos^2\varphi-a_{12}\sin 2\varphi$$

取φ使$b_{12}=0$，得φ要满足 $\tan 2\varphi=\dfrac{2a_{12}}{a_{11}-a_{22}}$

选取矩阵A的非对角元素的非零元，计算旋转角度，构造旋转矩阵J；A左乘J的转置，右乘J,更新A;
矩阵的第i,j行和第i,j列元素值改变;A(i,j)=A(j,i)=0

$$\varphi = \frac{1}{2}\tan^{-1}\frac{2a_{ij}}{a_{ii}-a_{jj}}$$

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

$A^{(1)}$的计算公式:

$$a_{ii}^{(1)} = a_{ii}\cos^2\varphi + a_{jj}\sin^2\varphi + a_{ij}\sin 2\varphi$$

$$a_{ij}^{(1)} = a_{ji}^{(1)} = 0.5\left(a_{jj} - a_{ii}\right)\sin 2\varphi + a_{ij}\cos 2\varphi$$

$$a_{jj}^{(1)} = a_{ii}\sin^2\varphi + a_{jj}\cos^2\varphi - a_{ij}\sin 2\varphi$$

$$a_{ip}^{(1)} = a_{pi}^{(1)} = a_{ip}\cos\varphi + a_{jp}\sin\varphi \qquad p \ne i, j$$

$$a_{jp}^{(1)} = a_{pj}^{(1)} = a_{jp}\cos\varphi + a_{ip}\sin\varphi \qquad p \ne i, j$$

$$a_{pq}^{(1)} = a_{qp}^{(1)} = a_{pq} = a_{qp} \qquad p, q \ne i, j$$

反复进行上述操作时，之前变成0的元素，可能变成非零元。

定理保证算法可行

定理：设实对称矩阵$\mathrm{A}^{(0)}=\mathrm{A}\in R^{n\times n}, J\left(i_k, j_k, \varphi_k\right)=J_k, k=1,2,\cdots$

是n阶旋转矩阵序列，若记 $A^{(k)}=J_k^T A^{(k-1)} J_k$

$$\Rightarrow \lim_{k\to\infty} A^{(k)}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

算法终止性判断：

$$E\left(A^{(k)}\right)=\sum_{i\neq j}\left(a_{ij}^{(k)}\right)^2<\varepsilon$$

Jacobi算法：

（1） 给定精度 ε，准备存放特征向量的矩阵 $\boldsymbol{Q} \Leftarrow \boldsymbol{I}, \boldsymbol{I}$ 为单位矩阵；

（2） 对 $k=1,2,\cdots$，

①选矩阵 $\boldsymbol{A}^{(k)}$ 非对角元素绝对值最大者 $|a_{pq}^{(k)}| = \max\limits_{i \neq j} |a_{ij}^{(k)}|$ 得到行标 p、列标 q；

②计算旋转角 $\varphi_k = \dfrac{1}{2} \arctan \dfrac{2a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}}$ 和旋转矩阵 $\boldsymbol{J}_k = \boldsymbol{J}(p,q,\varphi_k)$；

③ $\boldsymbol{Q} \Leftarrow \boldsymbol{Q}\boldsymbol{J}_k, \boldsymbol{A}^{(k)} \Leftarrow \boldsymbol{J}_k^{\mathrm{T}} \boldsymbol{A}^{(k-1)} \boldsymbol{J}_k$；

④计算 $E(\boldsymbol{A}^{(k)})$，并判别 $E(\boldsymbol{A}^{(k)}) < \varepsilon$，若成立，输出 $\boldsymbol{A}^{(k)}$（其主对角线元素为所求的全部特征值）和 $\boldsymbol{Q}$（其 n 个列为相应的 n 个特征向量），停止；否则做 $k \Leftarrow k+1$，转①.

编程时可注意节约内存空间

特征向量的正交性得到保证

过关Jacobi算法:

先取阈值 $\alpha_1=\sqrt{E\ (\mathbf{A})}/n$，然后按行的顺序依次把非对角元素 a_{ij} 与 α_1 相比，若 $|a_{ij}|<\alpha_1$，则不做变换，否则做变换将 a_{ij} 化为零，如此多次直至所有非对角元素的绝对值都小于 α_1 后，再选 $\alpha_2=\alpha_1/2$，重复上述过程，继续下去选阈值 $\alpha_1>\alpha_2>\alpha_3>\cdots>\alpha_m>0$，直至 $\alpha_m<\varepsilon$ 则终止计算，这里 $\alpha_k=\alpha_{k-1}/k$，当然也可以选其他的阈值，只要 α_k 单调减小即可．

过关 **Jacobi** 方法使确定旋转矩阵的过程大大缩短，因此可以加快求解过程．

编程时可注意节约内存空间
特征向量的正交性得到保证

例题：用Jacobi方法求矩阵

$$A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

的所有特征值及其相应特征向量，要求误差 $\varepsilon<10^{-5}$

解：**按Jacobi方法的算法，**

$$A^{(1)}=\begin{pmatrix} 1 & 0 & -0.707107 \\ 0 & 3 & -0.707107 \\ -0.707107 & -0.707107 & 2 \end{pmatrix},$$

$$J_1=\begin{pmatrix} 0.707107 & -0.707107 & 0 \\ 0.707107 & 0.707107 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E\left(A^{(1)}\right)=2$$

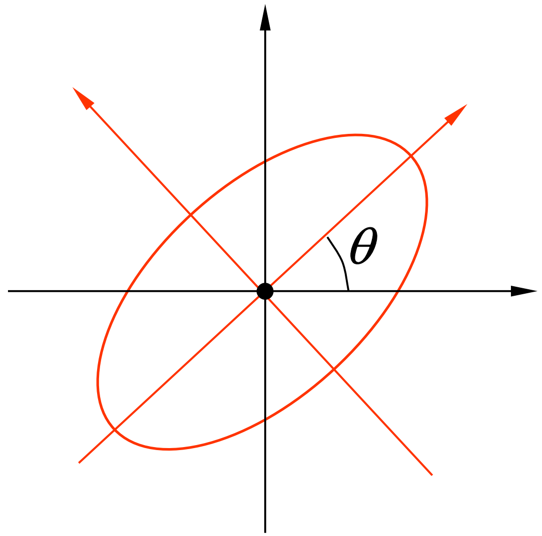
$$A^{(2)}=\begin{pmatrix} 0.633975 & -0.325058 & 0 \\ -0.325058 & 3 & -0.627963 \\ 0 & -0.627963 & 2.36603 \end{pmatrix},$$

$$J_2=\begin{pmatrix} 0.627963 & -0.707107 & -0.325058 \\ 0.627963 & 0.707107 & -0.325058 \\ 0.459701 & 0 & 0.888074 \end{pmatrix}, E\left(A^{(2)}\right)=1$$

$$A^{(6)}=\begin{pmatrix} 0.585788 & 0.00203811 & -0.24\times10^{-4} \\ 0.00203811 & 3.41421 & 0 \\ -0.24\times10^{-4} & 0 & 2 \end{pmatrix}, E\left(A^{(6)}\right)=0.83\times10^{-5}$$

得 λ_1=0.585788，λ_2=3.41421，λ_3=2

§4.5 QR方法



一、基本思想

利用矩阵的QR分解，通过逆序相乘产生对原矩阵的一系列正交相似变换，使其变化为一个近似的上三角矩阵来求全部特征值。

QR方法的作用

求**一般矩阵**的**全部特征值**。

QR分解： 将矩阵化为一个正交矩阵Q和一个上三角矩阵R相乘的形式:A=QR。

二、构造与分析

定理(实Schur分解定理）：

$$\exists 正交矩阵 Q \in R^{n\times n}, \Rightarrow Q^T AQ = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ & B_{22} & \cdots & B_{2n} \\ & & \ddots & \\ & & & B_{nn} \end{pmatrix}$$

其中B_{ii}是1×1或2×2的小矩阵。

当B_{ii}是1×1时，它就是A的特征值，

当为2×2时，其特征值就是A的一对共轭特征值。

定理：若$A \in R^{n\times n}$可逆，且有n个不同的特征值，

记$A = A^{(1)}$，则矩阵序列$A^{(k)}$本质上收敛于上三角块矩阵，

这里$A^{(k)}$的构造为：若$A^{(k)} = Q_k R_k \Rightarrow A^{(k+1)} = R_k Q_k$.

"本质上收敛"指 的主对角线上的元素或子块有确定的极限，其它元素或子块不管是否有极限。

逆序相乘的本质：

$$\because A^{(k)} = Q_k R_k \Rightarrow R_k = Q_k^{-1} A^{(k)} = Q_k^T A^{(k)}$$

$$\therefore A^{(k+1)} = R_k Q_k = Q_k^T A^{(k)} Q_k$$

说明矩阵序列是相似变换得到，有相同特征值。

QR法算法

1) 输入矩阵$A, k \Leftarrow 1$

2）对A做QR分解得到矩阵Q和R

3）$A \Leftarrow RQ$

4）如果A是对角线为1×1或2×2的上三角块矩阵，求出1×1或2×2的特征值，停止

5）$k \Leftarrow k+1, goto \quad 2)$

三、QR分解的方法

1、Hessenberg矩阵

$$\begin{pmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ & * & * & \cdots & * \\ & & \ddots & \ddots & \vdots \\ & & & * & * \end{pmatrix}$$

2、镜面反射矩阵:

$$P = I - \beta^{-1} V V^T, \quad V \neq 0 \in R^n, \beta = 0.5 \|V\|_2^2$$

镜面反射矩阵P也称为Householder满足：$P^T=P, P^{-1}=P$

3、定理：$\forall$非零$x=\{x_1,x_2,\cdots,x_n\}^T\in R^n$，$\exists \mathrm{P}\in \mathrm{R}^{n\times n}$，满足

$\underline{Px=-\sigma e_1\quad}$, $\sigma=\pm\|x\|_2$, $e_1=\{1,0,\cdots,0\}^T$，P为镜面反射矩阵。

证明：令 $u=x+\sigma e_1$

做一个镜面反射矩阵：$P=I-\beta^{-1}uu^T,\beta=0.5\|u\|_2^2$

$$\because x^Tx=\|x\|_2^2=\sigma^2, e_1^Te_1=1, x^Te_1=e_1^Tx=x_1,$$

$$(x+\sigma e_1)^T(x+\sigma e_1)=2(\sigma^2+\sigma x_1)$$

$$\therefore \beta=0.5u^T\cdot u=\sigma^2+\sigma x_1=\sigma(\sigma+x_1)$$

$$\therefore Px=(I-\beta^{-1}uu^T)x=x-(x+\sigma e_1)=-\sigma e_1$$

为避免$\beta=\sigma(\sigma+x_1)$有相近数相减，选$\sigma=(\operatorname{sgn} x_1)\|x\|_2$

4、对 $x \neq 0 \in R^n$，找方阵P，满足$Px = -\sigma e_1$ 的方法

1) 计算 $\sigma = (\operatorname{sgn} x_1)\|x\|_2$

2) 计算 $u = x + \sigma e_1$

3) 计算 $\beta = \sigma(\sigma + x_1)$

4) 取 $P = I - \beta^{-1} u u^T$

例题：对$x=\{-1,1,-1,1\}^{\mathrm{T}}$，试构造正交矩阵P，

满足$Px=-\sigma e_1$ ，这里$\sigma=\pm\|x\|_2$，$e_1=\{1,0,0,0\}^T$

解： $\sigma=(\operatorname{sgn} x_1)\|x\|_2$=-$\sqrt{4}=-2$

$$u=x+\sigma e_1=(-3,1,-1,1)^T$$

$$\beta=\sigma(\sigma+x_1)=-2(-2-1)6$$

$$P=I-\beta^{-1}uu^T=I-\frac{1}{6}\begin{pmatrix}-3\\1\\-1\\1\end{pmatrix}\begin{pmatrix}-3 & 1 & -1 & 1\end{pmatrix}=\frac{1}{6}\begin{pmatrix}-3 & 3 & -3 & 3\\3 & 5 & 1 & -1\\-3 & 1 & 5 & 1\\3 & -1 & 1 & 5\end{pmatrix}$$

本章介绍的幂法、Jacobi 旋转法及 QR 方法是在计算机上常用的求矩阵特征值和特征向量的方法．这些方法都是用矩阵迭代来构造算法的．

幂法计算简单，特别适用于高阶稀疏矩阵情况，但其收敛速度有时不能令人满意，需要采用位移技术来对幂法进行加速，这方面的内容可以参考文献［4］．

Jacobi 方法是古典方法，它收敛快、精度高，便于并行计算且算法稳定．用 Jacobi 方法求出的特征向量有较好的正交性，但当矩阵阶数较大时，收敛速度减慢，因此 Jacobi 方法适用于求低阶对称矩阵的全部矩阵特征值和特征向量．

QR 方法是 20 世纪 60 年代发展起来的方法，被称为计算数学最值得注意的算法之一，也是目前求任意矩阵全部特征值的最有效方法．