第九章　特征值和特征向量的计算.

特征值的求解方法主要分为迭代法和分解法.

9.1 幂法

幂法的基本思想：令 $x^{(k)} = A^k x^{(0)}$

当 k 充分大时，归一化的 $\vec{x}^{(k)}$ 趋近于特征向量 $\vec{V}_1$

1° 证明：假设 $A \in \mathbb{R}^{n\times n}$，A 有 n 个线性无关的特征向量 $\vec{V}_1, \vec{V}_2, \cdots, \vec{V}_n$，

且相应特征值按模递减，且 λ_1 是单实根.

$$令\ \vec{x}^{(0)} = \sum_{i=1}^{n} a_i \vec{V}_i$$

$$则\ \vec{x}^{(k)} = A^k \sum_{i=1}^{n} a_i \vec{V}_i = \sum_{i=1}^{n} a_i \lambda_i^k \vec{V}_i = \lambda_1^k \left(\sum_{i=1}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{V}_i \right)$$

$$当\ k\to\infty\ 时，\vec{x}^{(k)} \approx \lambda_1^k a_1 \vec{V}_1 \quad 且\ \frac{x_i^{(k+1)}}{x_i^{(k)}} \to \lambda_1 .$$

注：当 $\left|\frac{\lambda_2}{\lambda_1}\right| \approx 1$ 时，收敛速度慢；当 $\left|\frac{\lambda_2}{\lambda_1}\right|$ 小时，收敛速度快。

归一化操作：

$$\begin{cases} \vec{y}^{(k)} = \dfrac{\vec{x}^{(k)}}{\|\vec{x}^{(k)}\|_\infty} \\ \vec{x}^{(k+1)} = A\vec{y}^{(k)} \end{cases}$$

类似地，可证当 $k\to\infty$ 时，$|\vec{y}^{(k)}| \to \left|\dfrac{a_1\vec{V}_1}{\|a_1\vec{V}_1\|_\infty}\right|$ 且 $\|\vec{x}^{(k)}\|_\infty \to |\lambda_1|$

2° 若 λ_1 是重根，即 $\lambda_1 = \lambda_2 = \cdots = \lambda_r$，$(r \le n)$，

$$则\ \vec{y}^{(k)} = \frac{\vec{x}^{(k)}}{\|\vec{x}^{(k)}\|_\infty} = \left(\frac{\lambda_1}{|\lambda_1|}\right)^k \frac{a_1\vec{V}_1 + a_2\vec{V}_2 + \cdots + a_r\vec{V}_r + a_{r+1}\left(\frac{\lambda_{r+1}}{\lambda_1}\right)^k \vec{V}_{r+1} + \cdots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{V}_n}{\left\|a_1\vec{V}_1 + a_2\vec{V}_2 + \cdots + a_r\vec{V}_r + a_{r+1}\left(\frac{\lambda_{r+1}}{\lambda_1}\right)^k \vec{V}_{r+1} + \cdots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{V}_n\right\|_\infty}$$

$$\therefore |\vec{y}^{(k)}| \to \frac{a_1\vec{V}_1 + a_2\vec{V}_2 + \cdots + a_r\vec{V}_r}{\|a_1\vec{V}_1 + a_2\vec{V}_2 + \cdots + a_r\vec{V}_r\|_\infty}，仍是\ \lambda_1\ 的特征向量.$$

且 $\|\vec{x}^{(k)}\|_\infty \to |\lambda_1|$

注：1) 若计算结果不收敛，则应改用它法.

2) 幂法的特点：优点：算法简单，能同时求特征值和特征向量，适合稀疏矩阵.（因为每步迭代计算量小.）

缺点：收敛速度慢.

9.2 幂法的加速与降阶

1° 加速（原点位移法）

用 $A-\lambda_0 I$ 代替 A 进行迭代时，新的特征值为 $\lambda_i-\lambda_0$，如果能使

$\max\left\{\left|\dfrac{\lambda_2-\lambda_0}{\lambda_1-\lambda_0}\right|, \cdots, \left|\dfrac{\lambda_n-\lambda_0}{\lambda_1-\lambda_0}\right|\right\}$ 有效地减小，则实现了幂法的加速。

例如，对正定矩阵，可令 $\lambda_0=\frac{1}{2}(\lambda_2+\lambda_n)$，$(\lambda_1\neq\lambda_2)$.

注：由于矩阵特征值的分布并不清楚，所以常常需预估一个 λ_0.

2° 降阶

假设 A 为实对称矩阵，则 A 有 n 个相互正交的特征向量。

记 $A^{(1)}=A$，构造

$$A^{(2)}=A^{(1)}-\lambda_1\frac{\vec{v}_1\vec{v}_1^T}{(\vec{v}_1^T\vec{v}_1)}$$

则
$$A^{(2)}\vec{v}_1=A^{(1)}v_1-\lambda_1\frac{\vec{v}_1\vec{v}_1^T\vec{v}_1}{(\vec{v}_1^T\vec{v}_1)}=\vec{0}$$

$$A^{(2)}\vec{v}_i=A^{(1)}v_i-\lambda_1\vec{v}_1\frac{\vec{v}_1^T\vec{v}_i}{\vec{v}_1^T\vec{v}_1}=\lambda_i\vec{v}_i\quad(i=2,\cdots,n)$$

∴ 此时，$A^{(2)}$ 的按模最大特征值为 λ_2，按幂法可得 λ_2 和 $\vec{v}_2$

一般地，可以定义

$$A^{(k+1)}=A^{(k)}-\lambda_k\frac{\vec{v}_k\vec{v}_k^T}{(\vec{v}_k^T\vec{v}_k)}$$，从而求得 λ_k 和 $\vec{v}_k$.

注：此法求出的特征值精度会下降，一般只能求前几个特征值和特征向量.

9.3 反幂法

用途：求矩阵的按模最小特征值和特征向量.

假设矩阵 A 是实非奇异阵，且有 n 个线性无关的特征向量 $\vec{v}_1,\cdots,\vec{v}_n$.
且对应特征值按模递减。

∵ $A\vec{v}_i=\lambda_i\vec{v}_i$

∴ $A^{-1}\vec{v}_i=\frac{1}{\lambda_i}\vec{v}_i$，即 A^{-1} 的特征值为 $\frac{1}{\lambda_i}$，且按模最大特征值为 $\frac{1}{\lambda_n}$.

① 解线性方程组法

因为 A^{-1} 很难求，一般解法为：

$$\begin{cases}\vec{y}^{(k)}=\dfrac{\vec{x}^{(k)}}{\|\vec{x}^{(k)}\|_\infty}\\ A\vec{x}^{(k+1)}=\vec{y}^{(k)}\end{cases}\quad(k=0,1,2,\cdots)$$

作 LU 分解得
$$\begin{cases}\vec{y}^{(k)}=\dfrac{\vec{x}^{(k)}}{\|\vec{x}^{(k)}\|_\infty}\\ L\vec{z}^{(k)}=\vec{y}^{(k)}\\ U\vec{x}^{(k+1)}=\vec{z}^{(k)}\end{cases}$$

② 原点加速法

$(A-\lambda_0 I)^{-1}$ 的特征值为 $\frac{1}{\lambda_1-\lambda_0}, \cdots, \frac{1}{\lambda_n-\lambda_0}$.

1° 若取 $\lambda_0 \approx \lambda_l$，则 $\left|\frac{\frac{1}{\lambda_l-\lambda_0}}{\frac{1}{\lambda_k-\lambda_0}}\right|$ $(k\neq l)$ 可以很大，

能很快收敛. 因为 $A-\lambda_0 I$ 几乎奇异，所以用列主元 LU分解求解。

2° 若已知近似的 $\vec{x}_l$，则 $\lambda_l \approx \frac{\vec{x}_l^T A \vec{x}_l}{\vec{x}_l^T \vec{x}_l}$，可令 $\lambda_0=\lambda_l$.

③ 特征值修正.

若已知近似的特征值 λ^*，可令 $\lambda_0=\lambda^*$，求出 $\frac{1}{\lambda-\lambda^*}$，

收敛速度一般很快。

9.4 平行迭代法

基本算法 ⊂ 幂法 → 正交归一化

用途：同时计算前几个较大的特征值及对应特征向量。

对于高阶对称稀疏矩阵比较有效。

9.5 QR算法

用途：求实非奇异阵的全部特征值。

QR分解定理：若 $A \in R^{n\times n}$ 且非奇异，则 ∃ Q, R, st.

$$A=QR.$$

且除矩阵的行(列)符号外，分解是唯一的。

① 基本QR算法

取 $A_1=A$，做迭代

$$\begin{cases} A_k=Q_k R_k \\ A_{k+1}=R_k Q_k. \end{cases} \quad (k=0,1,\cdots)$$

$\because A_{k+1}=R_k Q_k = Q_k^{-1} A_k Q_k$

$\therefore A_{k+1} \sim A_k$，有相同的特征值.

② QR算法的收敛性.

本质收敛：序列 $\{A_k\}$ 收敛于上三角（或分块上三角）矩阵，

其对角线元素（或子块）有确定的极限.（简单情形收敛到对角阵.）

定理：若 A 可对角化，且 $A=XDX^{-1}$，D的各对角元按模递减排列，且 $X^{-1}=L_y U_y$，

则 $\{A_k\}$ 本质收敛。

③ QR分解法

1° Schmidt 正交化

2° Household 变换

3° Givens 变换

④ 用反幂法求特征向量

已知近似的λ^*，令$\lambda_0=\lambda^*$，用反幂法可找出更精确的λ和对应特征向量$\vec{v}$

⑤ 带原点位移的QR方法

取$A_1=A$，作迭代

$$\begin{cases} A_k-\mu_k I = Q_k R_k \\ A_{k+1}=R_k Q_k+\mu_k I \end{cases} \quad (k=1,2,\cdots)$$

其中，μ_k称为第k次迭代的原点位移量，可取$a_{nn}^{(k+1)}$.

$\because A_{k+1}=R_kQ_k+\mu_kI$

$= Q_k^{-1}(A_k-\mu_kI)Q_k+\mu_kI$

$= Q_k^{-1}A_kQ_k$

$\therefore A_{k+1}\sim A_k$，有相同的特征值.

9.6 Jacobi 迭代法

用途：求实对称矩阵的特征值和特征向量。适用于阶数不高的满对称阵.

① 基本思想

取$A_0=A$，做迭代：$A_k=R_k^TA_{k-1}R_k$，其中R_k为正交阵，

逐步将A_k变换为对角阵D.

记$V_k=R_1\cdots R_k$，则V_k也是正交阵，且$\begin{cases} D=V_k^TAV_k. \\ AV_k=V_kD \end{cases}$

即 D的对角元为A的特征值，V_k的列向量为A的特征向量。

② 平面旋转矩阵.

平面旋转矩阵：$R(i,j,\varphi)=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & \cos\varphi & \cdots & -\sin\varphi & \\ & & \vdots & \ddots & \vdots & \\ & & \sin\varphi & \cdots & \cos\varphi & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$

对对称矩阵A作正交相似变换：$A_1 = R_{ij}^T A R_{ij}$

则
$$
\begin{cases}
a_{ii}^{(1)} = a_{ii}\cos^2\varphi + a_{jj}\sin^2\varphi + 2a_{ij}\cos\varphi\sin\varphi \\
a_{jj}^{(1)} = a_{ii}\sin^2\varphi + a_{jj}\cos^2\varphi - 2a_{ij}\cos\varphi\sin\varphi \\
a_{il}^{(1)} = a_{li}^{(1)} = a_{il}\cos\varphi + a_{jl}\sin\varphi \qquad (l\ne i,j) \\
a_{jl}^{(1)} = a_{lj}^{(1)} = -a_{il}\sin\varphi + a_{jl}\cos\varphi \qquad (l\ne i,j) \\
a_{ij}^{(1)} = a_{ji}^{(1)} = \frac{1}{2}(a_{jj}-a_{ii})\sin 2\varphi + a_{ij}\cos 2\varphi \\
a_{lm}^{(1)} = a_{ml}^{(1)} = a_{ml} = a_{lm}
\end{cases}
$$

若取φ，使得 $\cot 2\varphi = \frac{a_{ii}-a_{jj}}{2a_{ij}}$，则旋转得到 $a_{ij}^{(1)} = a_{ji}^{(1)} = 0$.

每次取按模最大的非对角元所在的行和列为i,j进行如上旋转，经过无穷次平面旋转可以使矩阵对角化。

③收敛性

设A的非对角线元素平方和为 $E(A) = \sum_{i\ne j} a_{ij}^2$

$$\because (a_{il}^{(k+1)})^2 + (a_{jl}^{(k+1)})^2 = (a_{il}^{(k)}\cos\varphi + a_{jl}^{(k)}\sin\varphi)^2 + (-a_{il}^{(k)}\sin\varphi + a_{jl}^{(k)}\cos\varphi)^2$$
$$= (a_{il}^{(k)})^2 + (a_{jl}^{(k)})^2$$

同理 $(a_{li}^{(k+1)})^2 + (a_{lj}^{(k+1)})^2 = (a_{li}^{(k)})^2 + (a_{lj}^{(k)})^2$

又 $(a_{lm}^{(k+1)})^2 = (a_{ml}^{(k+1)})^2 = (a_{lm}^{(k)})^2 = (a_{ml}^{(k)})^2$

$a_{ij}^{(k+1)} = a_{ji}^{(k+1)} = 0$

$$\therefore E(A_{k+1}) = E(A_k) - 2(a_{ij}^{(k)})^2$$

对角线元素之和

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

∴ 正交相似变换矩阵元素的平方和不变，因上述变换对角线元素平方和增大。

按②中的i,j取法时，$(a_{ij}^{(k)})^2 \ge \frac{E(A_k)}{n^2-n}$

$$\therefore E(A_{k+1}) = E(A_k) - 2(a_{ij}^{(k)})^2 \le E(A_k) - 2\frac{E(A_k)}{n^2-n} = \left(1-\frac{2}{n^2-n}\right)E(A_k)$$

$\therefore E(A_k) \le \left(1-\frac{2}{n^2-n}\right)^k E(A)$，即 $\lim_{k\to\infty} E(A_k) = 0$.

∴ Jacobi迭代法收敛

注：Jacobi迭代法得到的系数矩阵不一定稀疏。