第十一章　矩阵特征值和特征向量的计算

《高等工程数学》(第三版)　姚仰新, 王福昌, 罗家洪, 庄楚强

华南理工大学出版社出版

2016 年 10 月 20 日

目录

序言

科学与工程技术领域的许多问题常常要求计算矩阵的特征值和特征向量，即对 n 阶方阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，求解 n 维非零向量 $\boldsymbol{x} \in \mathbb{R}^n$ 和常数 λ 使得 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ 成立。由方程组 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ 易知，特征值 λ 是特征多项式 $|\lambda\boldsymbol{E} - \boldsymbol{A}|$ 的零点，而特征向量 $\boldsymbol{x}$ 是齐次线性方程组 $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}_{n\times 1}$ 非零解。对于高阶方阵，由 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$ 求特征值往往十分困难。本章介绍几种计算矩阵特征值和特征向量的数值方法，虽然它们都是近似方法，但是能有效地控制误差大小。

对于给定的矩阵，事先估计特征值的分布往往是必要的。著名的圆盘定理是估计特征值的有效方法。

序言

定理 (Gerschgorin 圆盘定理)

设矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$，记复平面上以 a_{ii} 为中心，$r_i=\sum\limits_{\substack{j=1\\j\neq i}}^{n}|\,a_{ij}\,|$ 为半径的 n 个圆盘为

$$R_i=\{\lambda\mid|\lambda-a_{ii}|\leqslant r_i\},(i=1,2,\cdots,n),$$

则：

(1) 矩阵 $\boldsymbol{A}$ 的任一个特征值至少位于其中一个圆盘内;

(2) 在 m 个圆盘相互连通 (而与其余 $n-m$ 个圆盘互不联通) 的区域内，含有 $\boldsymbol{A}$ 的 m 个特征值.

注：(1) 结论 (2) 中, $m=1$ 时表明在每一个孤立的圆盘中仅有一个特征值。(2) r_i 是矩阵 $\boldsymbol{A}$ 的第 i 行去除第 i 个的元素后元素的模的和。

序言

例 (11.1)

试讨论 $\boldsymbol{A}=\begin{bmatrix}4 & 1 & 0\\1 & 0 & -1\\1 & 1 & -4\end{bmatrix}$ 的特征值的分布。

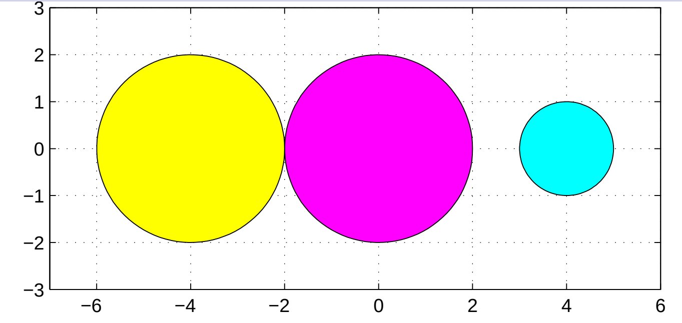

图: 圆盘

序言

解

解：由矩阵 $\boldsymbol{A}$ 确定的 3 个圆盘为

$$R_1=\{\lambda \mid |\lambda-4|\leqslant 1\}, R_2=\{\lambda \mid |\lambda|\leqslant 2\}, R_3=\{\lambda \mid |\lambda+4|\leqslant 2\}.$$

由于第一个圆盘与其余两个圆盘互不相通，故其中有且仅有一个特征值，而复特征值必成对共轭出现，即在区间 $[3,5]$ 内有一个实特征值，其余两个圆盘的交点 $\lambda=-2$ 不是特征值，因此 $\boldsymbol{A}$ 特征值分别位于区间 $[-6,-2],[-2,2],[3,5]$ 内。实际上，用 *MATLAB* 中的 eig 可求得特征值分别为 $-3.7601,-0.4429,4.2030$.

```
>> A = [4 1 0; 1 0 -1;1 1 -4];
>> eig(A)
ans =
4.2030
-3.7601
-0.4429
```

序言

在主成分分析中，需要计算协方差矩阵的特征值和特征向量，而协方差矩阵，是一个实对称矩阵。这里给出 n 阶实对称矩阵的特征值估计：如果 $\boldsymbol{A}$ 是 n 阶实对称矩阵，则有

(1) $\lambda_n \leqslant \dfrac{(\boldsymbol{Ax},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})} \leqslant \lambda_1, \forall\, \boldsymbol{0} \neq \boldsymbol{x} \in \mathbb{R}^n$;

(2) $\lambda_n = \min\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{(\boldsymbol{Ax},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}$;

(3) $\lambda_1 = \max\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{(\boldsymbol{Ax},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}$.

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ 为 $\boldsymbol{A}$ 的特征值。

证明.

因为 $\boldsymbol{0} \neq \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{A}^T = \boldsymbol{A}, \boldsymbol{Ax} = \lambda\boldsymbol{x}$，故 $(\boldsymbol{Ax},\boldsymbol{x}) = (\boldsymbol{Ax})^T\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{Ax} = \lambda\boldsymbol{x}^T\boldsymbol{x} = \lambda(\boldsymbol{x},\boldsymbol{x})$, 整理得 $\dfrac{(\boldsymbol{Ax},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})} = \lambda$，上面三个结论显然成立。 $\square$

11.1 乘幂法与反幂法

一、乘幂法

实际问题中，矩阵 $\boldsymbol{A}$ 按模最大的特征值起着重要的作用，并且希望有较高的计算精度。因此，求矩阵按模最大特征值的问题十分重要。

设 n 阶实矩阵 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$，对应的特征值满足

$$|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|,$$

下面求按模最大的特征值 λ_1 和对应特征向量 $\boldsymbol{x}_1$.

任取一 n 维非零向量 $\boldsymbol{u}_0$，则 $\boldsymbol{u}_0$ 一定可以表示为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 的线性组合： $\boldsymbol{u}_0 = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_n\boldsymbol{x}_n$，两边乘以 $\boldsymbol{A}^k$，并利用 $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i, (i = 1, 2, \cdots, n)$，记 $\boldsymbol{u}_k = \boldsymbol{A}^k\boldsymbol{u}_0$，故有

$$\begin{aligned}\boldsymbol{u}_k = \boldsymbol{A}^k\boldsymbol{u}_0 &= \alpha_1\lambda_1^k\boldsymbol{x}_1 + \alpha_2\lambda_2^k\boldsymbol{x}_2 + \cdots + \alpha_n\lambda_n^k\boldsymbol{x}_n \\ &= \lambda_1^k[\alpha_1\boldsymbol{x}_1 + \alpha_2(\frac{\lambda_2}{\lambda_1})^k\boldsymbol{x}_2 + \cdots + \alpha_n(\frac{\lambda_n}{\lambda_1})^k\boldsymbol{x}_n].\end{aligned}$$

设 $\alpha_1 \neq 0$，由于 $\left|\dfrac{\lambda_i}{\lambda_1}\right| < 1 (i = 2, 3, \cdots, n)$ ，故对充分大的 k ，有 $\boldsymbol{u}_k \approx \alpha_1 \lambda_1^k \boldsymbol{x}_1 \approx \lambda_1 \boldsymbol{u}_{k-1}$，$\boldsymbol{u}_k$ 与 $\boldsymbol{u}_{k-1}$ 几乎仅差一个常数因子，故

$$\lambda_1 \approx \frac{(\boldsymbol{u}_k)_i}{(\boldsymbol{u}_{k-1})_i}, (i = 1, 2, \cdots, n.)$$

由 $\boldsymbol{u}_k = \boldsymbol{A}^k \boldsymbol{u}_0 = \boldsymbol{A} \boldsymbol{u}_{k-1}$.

综上讨论，求主特征值 λ_1 及相应特征向量的步骤如下：

1 任给 n 维初始向量 $\boldsymbol{u}_0 \neq \boldsymbol{0}$;

2 按 $\boldsymbol{u}_k = \boldsymbol{A} \boldsymbol{u}_{k-1}, k = 1, 2, \cdots,$ 计算 $\boldsymbol{u}_k$;

3 如果从某个常数后，$\frac{(\boldsymbol{u}_k)_i}{(\boldsymbol{u}_{k-1})_i} \approx c$常数, $(i = 1, 2, \cdots, n.)$，则取 $\lambda_1 = c$ ，而 $\boldsymbol{u}_k$ 就是与 λ_1 对应的一个近似特征向量。

注：算法收敛速度主要取决于 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 的大小。

例 (11.2)

求矩阵 $\boldsymbol{A}=\begin{bmatrix}\frac{1}{4} & \frac{1}{5}\\ \frac{1}{5} & \frac{1}{6}\end{bmatrix}$ 的按模最大特征值及其特征向量。

```
A = [1/4 1/5;1/5 1/6]; u0 = [1;0]; uk = u0;for k = 1:4 u1 = A*u0; u0=u1; uk = [uk u0]; end; uk
uk =
1.0000 0.2500 0.1025 0.0423 0.0175
0 0.2000 0.0833 0.0344 0.0142
>> uk(:,3:end)./uk(:,2:end-1)
ans =
0.4100 0.4126 0.4126
0.4167 0.4127 0.4126
```

11.1 乘幂法与反幂法

由计算结果可知，比值 $\dfrac{(\boldsymbol{u}_k)_1}{(\boldsymbol{u}_{k-1})_1}$ 与 $\dfrac{(\boldsymbol{u}_k)_2}{(\boldsymbol{u}_{k-1})_2}$ 很快趋于同一常数，取 $\lambda_1 \approx 0.4126$, 相应的特征向量取为 $\boldsymbol{x} = (0.0175, 0.0142)^T$.

值得注意的是，当 $|\lambda_1| > 1$(或者 $|\lambda_1| < 1$) 时，k 充分大，将使 u_k 的分量的绝对值过大 (或者过小)，以至于运算无法继续进行，为此，对乘幂法需做规范化运算。对非零向量 $\boldsymbol{u}$，规范化向量取 $\boldsymbol{y} = \boldsymbol{u}/\max\{\boldsymbol{u}\}$, 其中 $\max\{\boldsymbol{u}\}$ 是指 $\boldsymbol{u}$ 的绝对值最大的分量，例如设 $\boldsymbol{u} = (1, 0, -2, 5, -6)^T$，则 $\max\{\boldsymbol{u}\} = -6$，其规范化向量 $\boldsymbol{v} = (-1/6, 0, 1/3, -5/6, 1)^T$，且 $\|\boldsymbol{v}\|_\infty = 1$.

乘幂法的规范化计算公式如下，

任取初始向量 $\boldsymbol{v}_0 = \boldsymbol{u}_0 \neq 0$，对 $k = 1, 2, \cdots$，依次计算 $\boldsymbol{u}_k = \boldsymbol{A}\boldsymbol{v}_{k-1}$，$\mu_k = \max(\boldsymbol{u}_k)$，$\boldsymbol{v}_k = \dfrac{\boldsymbol{u}_k}{\mu_k}$，$(k = 1, 2, \cdots)$

比如上式得

$$\boldsymbol{v}_k = \frac{\boldsymbol{A}^k \boldsymbol{u}_0}{\max(\boldsymbol{A}^k \boldsymbol{u}_0)} = \frac{\alpha_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\dfrac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i}{\max\left[\alpha_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\dfrac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i\right]},$$

且有

$$\lim_{k\to\infty} \boldsymbol{v}_k = \frac{\boldsymbol{x}_1}{\max(\boldsymbol{x}_1)},$$

其收敛速度由比值 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 来确定。另一方面，

$$\mu_k = \max(\boldsymbol{u}_k) = \lambda_1 \frac{\max\left[\alpha_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\dfrac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i\right]}{\max\left[\alpha_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\dfrac{\lambda_i}{\lambda_1}\right)^{k-1} \boldsymbol{x}_i\right]}, \qquad (11.1)$$

则有

$$\lim_{k\to\infty} \mu_k = \lambda_1,$$

因此当 k 足够大时，可取 $\boldsymbol{A}$ 按模最大特征值和相应的特征向量为，

$$\lambda_1 \approx \mu_k = \max(\boldsymbol{u}_k), \boldsymbol{x} \approx \boldsymbol{v}_k = \frac{\boldsymbol{u}_k}{\mu_k}.$$

下面给出算法对应的 Matlab 程序，和一个例子。

第十一章 矩阵特征值和特征向量的计算

《高等工程数学》(第三版) 姚仰新等

目录

序言

11.1 乘幂法与反幂法

11.2 雅可比 (Jacobi) 方法

11.3 QR 方法

11.4 广义特征值与特征向量

参考文献

例 (11.3)

求矩阵

$$\boldsymbol{A}=\begin{bmatrix}-4 & 14 & 0\\ -5 & 13 & 0\\ -1 & 0 & 2\end{bmatrix}$$

的按模最大特征值及其特征向量。

11.1 乘幂法与反幂法

例 (11.3)

求矩阵

$$A = \begin{bmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

的按模最大特征值及其特征向量。

幂法的程序为:

```
function [m, u] = powermethod(A, ep, it_max)
% 求矩阵最大模特征值的幂法，调用格式为
%   [m, u] = powermethod(A, ep, it_max)
% 其中
% A 为矩阵，ep 为精度要求，默认为1e-5,
% it_max 为最大迭代次数，默认为100
% m 为模最大的特征值，u 为 m 对应的特征向量
```

```
if nargin < 3 it_max = 100; end
if nargin <2 ep = 1e-5; end;
n = length(A); u = ones(n,1); k = 0; m1 = 0;
while k <= it_max
      v = A*u; [vmax, i ] = max(abs(v));
      m  = v(i); u = v/m;
      if abs(m-m1)<ep break; end
      m1 = m; k = k+1;
end
```

11.1 乘幂法与反幂法

上例中结果为:

```
>> A = [-4   14    0 ;-5    13    0 ; -1   0   2];
 [m, u] = powermethod(A)
m =

 6.0000
 u =
 1.0000
 0.7143
 -0.2500
```

二、加速方法

计算矩阵 $\boldsymbol{A}$ 按模最大特征值的乘幂法的收敛速度，取决于比值 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 的大小，比值越小，数量速度越快。当比值接近于 1 时，速度就很慢，为了提高收敛速度，减少计算量，这里介绍两种常用的加速方法。

二、加速方法

计算矩阵 $\boldsymbol{A}$ 按模最大特征值的乘幂法的收敛速度，取决于比值 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 的大小，比值越小，数量速度越快。当比值接近于 1 时，速度就很慢，为了提高收敛速度，减少计算量，这里介绍两种常用的加速方法。

1、艾特金 (Aitken) 方法

由 (11.1) 式知道

$$\mu_k = \max(\boldsymbol{u}_k) = \lambda_1 \left[1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^k\right)\right],$$

因而有

$$\mu_k - \lambda_1 \approx c\left(\frac{\lambda_2}{\lambda_1}\right)^k,$$

其中 c 是某个与 k 无关的常数。当比值 $\left|\dfrac{\lambda_2}{\lambda_1}\right|$ 接近于 1 时，有

$$\frac{\mu_{k+1} - \lambda_1}{\mu_k - \lambda_1} \approx \frac{\mu_{k+2} - \lambda_1}{\mu_{k+1} - \lambda_1}$$

整理得

$$\lambda_1 \approx \mu_k - \frac{(\mu_{k+1} - \mu_k)^2}{(\mu_{k+2} - 2\mu_{k+1} + \mu_k)},$$

我们可以用

$$\tilde{\mu}_k = \mu_k - \frac{(\mu_{k+1} - \mu_k)^2}{(\mu_{k+2} - 2\mu_{k+1} + \mu_k)}$$

作为 μ_k 的校正值。

11.1 乘幂法与反幂法

整理得

$$\lambda_1 \approx \mu_k - \frac{(\mu_{k+1}-\mu_k)^2}{(\mu_{k+2}-2\mu_{k+1}+\mu_k)},$$

我们可以用

$$\tilde{\mu}_k = \mu_k - \frac{(\mu_{k+1}-\mu_k)^2}{(\mu_{k+2}-2\mu_{k+1}+\mu_k)}$$

作为 μ_k 的校正值。

例 (11.3)

采用乘幂法求方阵 $\boldsymbol{A}$ 的按模最大特征值，并用艾特金法加速，其中

$$\boldsymbol{A} = \begin{bmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{bmatrix}.$$

```
%幂法及其加速方法
A = [-4 14 0 ; -5 13 0 ; -1 0 2];
u0 = [1;1;1]; v0 = u0;
n = 12; mu = zeros(n,1); T112 = zeros(n,6);
for k = 1:n
    u1 = A*v0; [maxu,i] = max(abs(u1)); mu(k) =
        u1(i); v1 = u1/mu(k);
    if k<3
        T112(k,1:5) = [k mu1 v1'];
    else
        muka = mu(k-2)-(mu(k-1)-mu(k-2))^2/(mu(k
            )-2*mu(k-1)+mu(k-2));
        T112(k,:) = [k mu(k) v1' muka];
    end
    u0 = u1;v0 = v1;
end
```

11.1 乘幂法与反幂法

```
>> T112(:,2:end)
ans =
  6.0008    1.0000    0.8000    0.1000         0
 6.0008    1.0000    0.7500   -0.1111         0
 6.5000    1.0000    0.7308   -0.1880    6.2667
 6.2308    1.0000    0.7222   -0.2209    6.0625
 6.1111    1.0000    0.7182   -0.2359    6.0154
 6.0545    1.0000    0.7162   -0.2431    6.0038
 6.0270    1.0000    0.7152   -0.2466    6.0010
 6.0135    1.0000    0.7148   -0.2483    6.0002
 6.0067    1.0000    0.7145   -0.2492    6.0001
 6.0034    1.0000    0.7144   -0.2496    6.0000
 6.0017    1.0000    0.7143   -0.2498    6.0000
 6.0008    1.0000    0.7143   -0.2499    6.0000
```

本题特征值为 $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 2.$

2、原点平移法

假设 $\boldsymbol{A}$ 的特征值按模由大至小次序排列的特征值为 $\lambda_1^{\boldsymbol{A}}, \lambda_2^{\boldsymbol{A}}, \cdots, \lambda$ 令矩阵 $\boldsymbol{B} = \boldsymbol{A} - p\boldsymbol{I}$,

则由特征值求法可知 $\lambda_i^{\boldsymbol{B}} = \lambda_i^{\boldsymbol{A}} - p, (i = 1, 2, \cdots, n)$ 为矩阵 $\boldsymbol{B}$ 的矩阵特征值，且与 $\lambda_i^{\boldsymbol{A}}$ 对应的特征向量相同。适当选取 p, 使得 $\lambda_1^{\boldsymbol{A}} - p$ 仍为 $\boldsymbol{B}$ 的按模最大的特征值，且有

$$\max_{2\leqslant k\leqslant n}\left|\frac{\lambda_k^{\boldsymbol{A}} - p}{\lambda_1^{\boldsymbol{A}} - p}\right| < \left|\frac{\lambda_2^{\boldsymbol{A}} - p}{\lambda_1^{\boldsymbol{A}} - p}\right|,$$

可用乘幂法求出 $\boldsymbol{B}$ 的按模最大特征值 $\lambda_1^{\boldsymbol{B}}$, 则对应 $\boldsymbol{A}$ 的最大特征值 $\lambda_1^{\boldsymbol{A}} = \lambda_1^{\boldsymbol{B}} + p$.

11.1 乘幂法与反幂法

2、原点平移法

假设 $\boldsymbol{A}$ 的特征值按模由大至小次序排列的特征值为 $\lambda_1^{\boldsymbol{A}}, \lambda_2^{\boldsymbol{A}}, \cdots, \lambda$

令矩阵 $\boldsymbol{B} = \boldsymbol{A} - p\boldsymbol{I}$,

则由特征值求法可知 $\lambda_i^{\boldsymbol{B}} = \lambda_i^{\boldsymbol{A}} - p, (i = 1, 2, \cdots, n)$ 为矩阵 $\boldsymbol{B}$ 的矩阵特征值，且与 $\lambda_i^{\boldsymbol{A}}$ 对应的特征向量相同。适当选取 p, 使得 $\lambda_1^{\boldsymbol{A}} - p$ 仍为 $\boldsymbol{B}$ 的按模最大的特征值，且有

$$\max_{2\leqslant k\leqslant n}\left|\frac{\lambda_k^{\boldsymbol{A}} - p}{\lambda_1^{\boldsymbol{A}} - p}\right| < \left|\frac{\lambda_2^{\boldsymbol{A}} - p}{\lambda_1^{\boldsymbol{A}} - p}\right|,$$

可用乘幂法求出 $\boldsymbol{B}$ 的按模最大特征值 $\lambda_1^{\boldsymbol{B}}$, 则对应 $\boldsymbol{A}$ 的最大特征值 $\lambda_1^{\boldsymbol{A}} = \lambda_1^{\boldsymbol{B}} + p$.

例 (11.3')

采用乘幂法求方阵 $\boldsymbol{A} = \begin{bmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{bmatrix}$ 的按模最大特征值，并用原点平移法加速.

11.1 乘幂法与反幂法

```
>>  % 幂法及原点平移加速方法
A = [-4   14    0 ;  -5    13    0 ;   -1   0   2];
p = 2; B = A - p*eye(size(A));
u0 = [1;1;1]; v0 = u0;
n = 12; mu = zeros(n,1); T112 = zeros(n,5);
for k = 1:n
    u1 = B*v0; [maxu,i] = max(abs(u1)); mu(k) = u1(i); v1 = u1/mu(k);
    T112(k,:) = [k mu(k) v1'];
    u0 = u1;v0 = v1;
end
T112
```

11.1 乘幂法与反幂法


```
T112 =
1.0000     8.0000     1.0000     0.7500    -0.1250
2.0000     4.5000     1.0000     0.7222    -0.2222
3.0000     4.1111     1.0000     0.7162    -0.2432
4.0000     4.0270     1.0000     0.7148    -0.2483
5.0000     4.0067     1.0000     0.7144    -0.2496
6.0000     4.0017     1.0000     0.7143    -0.2499
7.0000     4.0004     1.0000     0.7143    -0.2500
8.0000     4.0001     1.0000     0.7143    -0.2500
9.0000     4.0000     1.0000     0.7143    -0.2500
10.0000     4.0000     1.0000     0.7143    -0.2500
11.0000     4.0000     1.0000     0.7143    -0.2500
12.0000     4.0000     1.0000     0.7143    -0.2500
```

可见 $\boldsymbol{B}$ 的最大特征值为 $\lambda_1^{\boldsymbol{B}} = 4$，从而 $\lambda_1^{\boldsymbol{A}} = \lambda_1^{\boldsymbol{B}} + p = 4 + 2 = 6.$

11.1 乘幂法与反幂法

三、反幂法

反幂法用来求非奇异矩阵按模最小的特征值和相应特征向量。约定同上。

步骤如下：

(1) 任给 n 维初始向量 $\boldsymbol{u}_0 \neq \boldsymbol{0}$;

(2) 按 $\boldsymbol{u}_k = \boldsymbol{A}^{-1}\boldsymbol{u}_{k-1}, k = 1, 2, \cdots .$;

(3) 如果从某个常数后，$\frac{(\boldsymbol{u}_k)_i}{(\boldsymbol{u}_{k-1})_i} \approx c$常数, $(i = 1, 2, \cdots, n.)$，则取 $\lambda_n = \dfrac{1}{c}$，而 $\boldsymbol{u}_k$ 就是与 λ_n 对应的一个近似特征向量。

实际计算 $\boldsymbol{A}^{-1}$ 困难，故上述迭代公式 $\boldsymbol{u}_k = \boldsymbol{A}^{-1}\boldsymbol{u}_{k-1}, k = 1, 2, \cdots .$ 常写为 $\boldsymbol{A}\boldsymbol{u}_k = \boldsymbol{u}_{k-1}, k = 1, 2, \cdots .$。反幂法计算量大。

```
function [m, u] = powerinvmethod(A, ep, it_max)
% 求矩阵最小模特征值的反幂法, 调用格式为
%  [m, u] = powerinvmethod(A, ep, it_max)
% 其中
% A 为矩阵, ep 为精度要求, 默认为1e-5,
% it_max 为最大迭代次数, 默认为100
% m 为模最小的特征值, u 为 m 对应的特征向量
```

```
if nargin < 3 it_max = 100; end
if nargin <2 ep = 1e-5; end;
n = length(A); u = ones(n,1); k = 0; m1 = 0;
while k <= it_max
      v = A\u; [vmax, i ] = max(abs(v));
      m  = v(i); u = v/m;
      if abs(m-m1)<ep break; end
          m1 = m; k = k+1;
      end
      m = 1/m;
end
```

例

求矩阵

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

的按模最小特征值及其特征向量。

11.1 乘幂法与反幂法

```
>> A= [2 -1 0;-1 2 -1;0 -1 2];
[m, u] = powerinvmethod(A)

m =
 0.5858
u =
 0.7071
 1.0000
 0.7071
```

11.1 乘幂法与反幂法

```
>> A= [2 -1 0;-1 2 -1;0 -1 2];
[m, u] = powerinvmethod(A)

m =
 0.5858
u =
 0.7071
 1.0000
 0.7071
```

例 (11.3”)

采用乘幂法求方阵 $\boldsymbol{A}=\begin{bmatrix}-4 & 14 & 0\\ -5 & 13 & 0\\ -1 & 0 & 2\end{bmatrix}$ 的按模最小特征值，并用原点平移法加速.

11.1 乘幂法与反幂法


```
%反幂法求最小特征值
A = [-4  14  0 ; -5  13  0 ;  -1  0  2];
u0 = [1;1;1]; v0 = u0;
n = 40; mu = zeros(n,1); T113 = zeros(n,5);
for k = 1:n
    u1 = A\v0; [maxu,i] = max(abs(u1)); mu(k) = u1(i);
    v1 = u1/mu(k);
    T113(k,:) = [k 1/mu(k) v1'];
    u0 = u1;v0 = v1;
end
T113
```

```
T113 =

1.0000      2.1176     -0.1176      0.1176      1.0000
2.0000      2.4286     -0.4286     -0.1429      1.0000
3.0000      2.4950     -0.4950     -0.2178      1.0000
4.0000      2.4634     -0.4634     -0.2195      1.0000
5.0000      2.3922     -0.3922     -0.1912
    ··· ··· ··· ··· ··· ··· ··· ··· ··· ··· ··· ···
28.0000      2.0001     -0.0001     -0.0000      1.0000
29.0000      2.0000     -0.0000     -0.0000      1.0000
```

可以看出，反幂法收敛的速度很慢，直到第 29 步才开始收敛到最小特征值 2。

```
%原点平移
p = 1.5;B = A-p*eye(size(A));
u0 = [1;1;1]; v0 = u0;
n = 20; mu = zeros(n,1); T114 = zeros(n,5);
for k = 1:n
    u1 = B\v0; [maxu,i] = max(abs(u1)); mu(k) =
        u1(i);
    v1 = u1/mu(k);
    T114(k,:) = [k 1/mu(k)+p v1'];
    u0 = u1;v0 = v1;
end
T114
```

11.1 乘幂法与反幂法

```
T114 =
1.0000     2.2941    -0.2941    -0.0588     1.0000
2.0000     2.3053    -0.3053    -0.1368     1.0000
3.0000     2.1547    -0.1547    -0.0750     1.0000
4.0000     2.0605    -0.0605    -0.0299     1.0000
5.0000     2.0213    -0.0213    -0.0106     1.0000
6.0000     2.0072    -0.0072    -0.0036     1.0000
7.0000     2.0024    -0.0024    -0.0012     1.0000
8.0000     2.0008    -0.0008    -0.0004     1.0000
9.0000     2.0003    -0.0003    -0.0001     1.0000
10.0000     2.0001    -0.0001    -0.0000     1.0000
11.0000     2.0000    -0.0000    -0.0000     1.0000
```

经过原点平移后，只要经过 11 步就可以收敛到最小特征值 2 了，但是如何选择平移参数 p 是个问题。

一、雅可比 (Jacobi) 方法的基本思想

雅可比 (Jacobi) 方法是求实对称阵全部特征值和特征向量的一种迭代方法。

由矩阵理论可知，实对称矩阵 $\boldsymbol{A}$ 的特征值均为实数，且存在正交阵 $\boldsymbol{R}$, 使得 $\boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ ，而 $\boldsymbol{R}$ 的第 i 列恰为 λ_i 的特征向量，亦即 $\boldsymbol{A}$ 有 n 个两两正交的特征向量。若记 $\boldsymbol{A}_1 = \boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R}$, 显然 $\boldsymbol{A}_1$ 也为对称阵。

由此可见，寻求实对称矩阵 $\boldsymbol{A}$ 的特征值与特征向量归结为求正交矩阵 $\boldsymbol{R}$ ，但直接求 $\boldsymbol{R}$ 是困难的，雅克比提出用一系列平面旋转矩阵逐次将 $\boldsymbol{A}$ 约化为对角阵。

11.2 雅可比 (Jacobi) 方法

一、雅可比 (Jacobi) 方法的基本思想

雅可比 (Jacobi) 方法是求实对称阵全部特征值和特征向量 的一种迭代方法。

由矩阵理论可知，实对称矩阵 $\boldsymbol{A}$ 的特征值均为实数，且存在正交阵 $\boldsymbol{R}$, 使得 $\boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ ，而 $\boldsymbol{R}$ 的第 i 列恰为 λ_i 的特征向量，亦即 $\boldsymbol{A}$ 有 n 个两两正交的特征向量。若记 $\boldsymbol{A}_1 = \boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R}$，显然 $\boldsymbol{A}_1$ 也为对称阵。

由此可见，寻求实对称矩阵 $\boldsymbol{A}$ 的特征值与特征向量归结为求正交矩阵 $\boldsymbol{R}$ ，但直接求 $\boldsymbol{R}$ 是困难的，雅克比提出用一系列平面旋转矩阵逐次将 $\boldsymbol{A}$ 约化为对角阵。

二、平面旋转矩阵

现以二阶实对称矩阵 $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 为例，其中 $a_{12} = a_{21} \neq 0$. 对任意 $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ ，矩阵 $\boldsymbol{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ 显然是正交阵，

11.2 雅可比 (Jacobi) 方法

作正交相似变换

$$\boldsymbol{A}_1 = \boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R} = \boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R},$$

$\boldsymbol{A}_1$ 的元素 $a_{ij}^{(1)}$ 容易求得为

$$\begin{cases} a_{11}^{(1)} = a_{11}\cos^2\theta + 2a_{12}\cos\theta\sin\theta + a_{22}\sin^2\theta, \\ a_{22}^{(1)} = a_{11}\sin^2\theta - 2a_{12}\sin\theta\cos\theta + a_{22}\cos^2\theta, \\ a_{12}^{(1)} = \dfrac{1}{2}(a_{22} - a_{11})\sin 2\theta + a_{12}\cos 2\theta = a_{21}^{(1)}, \end{cases}$$

令 $a_{12}^{(1)} = a_{21}^{(1)} = 0$，得

$$\tan 2\theta = \frac{2a_{12}}{a_{11} - a_{22}}, \tag{11.2}$$

所以，如果取 θ 满足 (11.2) 式，则必有

$$\boldsymbol{A}_1 = \boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R} = \begin{bmatrix} a_{11}^{(1)} & 0 \\ 0 & a_{(22)}^{(1)} \end{bmatrix}$$

11.2 雅可比 (Jacobi) 方法

从而，原矩阵 $\boldsymbol{A}$ 的两个特征值就是 $a_{11}^{(1)}$ 或 $a_{22}^{(1)}$，而 $\boldsymbol{A}$ 对应的特征向量为 $\boldsymbol{R}_1=(\cos\theta,\sin\theta)^T$, $\boldsymbol{R}_2=(-\sin\theta,\cos\theta)^T$.

比如 $\boldsymbol{A}=\begin{bmatrix}1 & 1/3\\ 1/3 & 1/2\end{bmatrix}$,

由 (11.2), $\tan 2\theta=\dfrac{2a_{12}}{a_{11}-a_{22}}=\dfrac{2\times 1/3}{1-1/2}=4/3$,

从而 $\tan 2\theta=\dfrac{\sin 2\theta}{\cos 2\theta}=\dfrac{2\tan\theta}{1-\tan^2\theta}$,

可以解得 $\tan\theta=\dfrac{1}{2}, \tan\theta=-2$. 故取 $\tan\theta=\dfrac{1}{2}$ 对应 θ, $R_1=[2/\sqrt{5},1/\sqrt{5}]$, $R_2=[-1/\sqrt{5},2/\sqrt{5}]$,

令矩阵 $\boldsymbol{R}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$，计算可得

$$\boldsymbol{R}^T\boldsymbol{A}\boldsymbol{R}=\begin{bmatrix}7/6 & 0\\ 0 & 1/3\end{bmatrix},$$

特征值 $\lambda_1=\dfrac{7}{6},\lambda_2=\dfrac{1}{3}$.

11.2 雅可比 (Jacobi) 方法

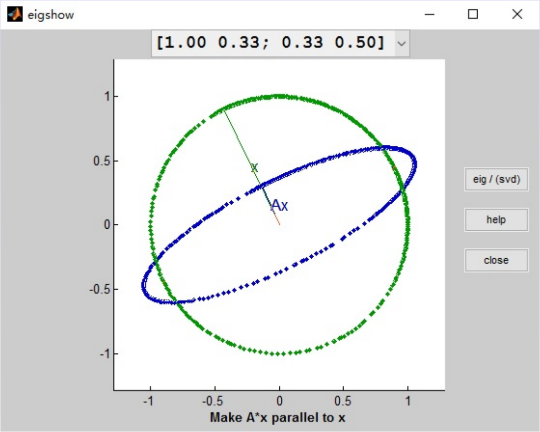

图: 用于显示特征值及特征向量的 eigshow 结果

11.2 雅可比 (Jacobi) 方法

下面推广到 n 维情况，引进 $\mathbb{R}^n$ 中平面旋转变换矩阵

$$\boldsymbol{R}_{pq}(\theta) = \begin{bmatrix}
1 & & & & & & & & \\
 & \ddots & & & & & & & \\
 & & 1 & & & & & & \\
 & & & \cos\theta & \cdots & -\sin\theta & & & \\
 & & & 1 & & & & & \\
 & & & \vdots & \ddots & \vdots & & & \\
 & & & & & 1 & & & \\
 & & & \sin\theta & \cdots & \cos\theta & & & \\
 & & & & & & 1 & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & 1
\end{bmatrix}
\begin{array}{l}
\\ \\ \\ \cdots 第\ p\ 行 \\ \\ \\ \\ \cdots 第\ q\ 行 \\ \\ \\ \\
\end{array}$$

称 $\boldsymbol{R}_{pq}(\theta)$ 平面旋转矩阵 (Givens 变换)。 $\boldsymbol{R}_{pq}(\theta)$ 具有如下性质:

(1) $\boldsymbol{R}_{pq}(\theta)$ 为正交矩阵，即 $\boldsymbol{R}_{pq}^{-1}(\theta) = \boldsymbol{R}_{pq}^{T}(\theta)$;

(2) 如果 $\boldsymbol{A}$ 为对称矩阵，则 $\boldsymbol{B} = \boldsymbol{R}_{pq}^{T}(\theta)\boldsymbol{A}\boldsymbol{R}_{pq}(\theta)$ 也为对称阵，且与 $\boldsymbol{A}$ 有相同的特征值。

(3) $\boldsymbol{R}_{pq}(\theta)\boldsymbol{A}$ 仅改变 $\boldsymbol{A}$ 的第 p 行和第 q 行元素， $\boldsymbol{A}\boldsymbol{R}_{pq}(\theta)$ 仅改变 $\boldsymbol{A}$ 的第 p 列和第 q 列元素。

设实对称矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$，记 $\boldsymbol{B} = \boldsymbol{R}_{pq}^{T}(\theta)\boldsymbol{A}\boldsymbol{R}_{pq}(\theta) = (b_{ij})_{n\times n}$，则它们之间有如下关系:

$$\begin{cases} b_{pp} = a_{pp}\cos^2\theta + a_{pq}\sin 2\theta + a_{qq}\sin^2\theta, \\ b_{qq} = a_{pp}\sin^2\theta - a_{pq}\sin 2\theta + a_{qq}\cos^2\theta, \\ b_{pq} = b_{qp} = \frac{1}{2}(a_{qq} - a_{pp})\sin 2\theta + a_{pq}\cos 2\theta, \\ b_{ip} = b_{pi} = a_{ip}\cos\theta + a_{iq}\sin\theta (i \neq p, q), \\ b_{iq} = b_{qi} = -a_{jp}\sin\theta + a_{jq}\cos\theta (j \neq p, q), \\ b_{ij} = a_{ij} (i, j \neq p, q). \end{cases}$$

由上述关系式可得

$$\begin{cases} b_{pi}^2 + b_{qi}^2 = a_{pi}^2 + a_{qi}^2, (i \neq p, q) \\ b_{jp}^2 + b_{iq}^2 = a_{jp}^2 + a_{jq}^2, (j \neq p, q) \\ b_{pp}^2 + b_{qq}^2 + 2b_{pq}^2 = a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2. \end{cases} \tag{11.3}$$

11.2 雅可比 (Jacobi) 方法

又注意到

$$\|\boldsymbol{B}\|_F^2 = \|\boldsymbol{A}\|_F^2, \text{即} \sum_{i,j=1}^{n} b_{ij}^2 = \sum_{i,j=1}^{n} a_{ij}^2, \tag{11.4}$$

由 (11.3) 式和 (11.4) 式可以得到， $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的对角线和非对角线之间的关系。

$$\sum_{i\neq j} b_{ij}^2 - 2b_{pq}^2 = \sum_{i\neq j} a_{ij}^2 - 2a_{pq}^2,$$

$$\sum_{i=1}^{n} b_{ii}^2 + 2b_{pq}^2 = \sum_{i=1}^{n} a_{ii}^2 + 2a_{pq}^2.$$

如果 $a_{pq} \neq 0$, 适当选取角 θ，使得 $b_{pq} = b_{qp} = \frac{1}{2}(a_{qq} - a_{pp})\sin 2\theta + a_{pq}\cos 2\theta = 0$， 只需取角 θ 满足

$$\cot 2\theta = \frac{a_{pp} - a_{qq}}{2a_{pq}} = \tau, |\ \theta\ | \leqslant \frac{\pi}{4}. \tag{11.5}$$

从而

$$\sum_{i\neq j} b_{ij}^2 = \sum_{i\neq j} a_{ij}^2 - 2a_{pq}^2 < \sum_{i\neq j} a_{ij}^2,$$

$$\sum_{i=1}^{n} b_{ii}^2 = \sum_{i=1}^{n} a_{ii}^2 + 2a_{pq}^2 > \sum_{i=1}^{n} a_{ii}^2.$$

结果表明，这样确定的旋转矩阵 $\boldsymbol{R}_{pq}(\theta)$ ，对 $\boldsymbol{A}$ 作正交变换后，非对角线元素的平方和减少了 $2a_{pq}^2$ ，而对角线元素平方和增加了 $2a_{pq}^2$ 。

根据 (11.5) 式，利用三角公式可知 $t=\tan\theta$ 满足

$$t^2+2\tau t-1=0,$$

此方程有两个根，为保证 $|\theta|\leqslant \frac{\pi}{4}$，取绝对值较小者为

$$t=\begin{cases}\operatorname{sgn}(\tau)/(|\tau|+\sqrt{1+\tau^2}), & \tau\neq 0,\\ 1, & \tau=0.\end{cases} \tag{11.7}$$

11.2 雅可比 (Jacobi) 方法

于是

$$\cos\theta = (1+t^2)^{-\frac{1}{2}}, \sin\theta = t\cos\theta. \tag{11.8}$$

例 (11.4)

用雅克比方法求对称阵

$$\boldsymbol{A} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 5 & 1 \\ 2 & 1 & 6 \end{bmatrix}$$

的全部特征值。

解：取 $a_{12} = 2$ 为非对角主元素，即 $p = 1, q = 2$，于是

$$\tau = \frac{a_{11} - a_{22}}{2a_{12}} = -0.25,$$

$$t = \operatorname{sgn}(\tau)/(|\ \tau\ | + \sqrt{1+\tau^2}) = -0.780776,$$

11.2 雅可比 (Jacobi) 方法

$$\cos\theta = (1+t^2)^{-\frac{1}{2}} = 0.788206, \sin\theta = t\cos\theta = -0.61542,$$

从而

$$\boldsymbol{R}_1 = \boldsymbol{R}_{pq}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.788205 & 0.615412 & 0 \\ -0.615412 & 0.788205 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

所以

$$\boldsymbol{A}_1 = \boldsymbol{R}_1^T \boldsymbol{A} \boldsymbol{R}_1 = \begin{bmatrix} 2.438447 & 0 & 0.960999 \\ 0 & 6.561553 & 2.019030 \\ 0.960999 & 2.019030 & 6 \end{bmatrix}$$

再取 $a_{23}^{(1)} = 2.019030$ 作为非对角主元，即 $p=2, q=3$，可得

$$\boldsymbol{A}_2 = \boldsymbol{R}_2^T \boldsymbol{A} \boldsymbol{R}_2 = \begin{bmatrix} 2.438447 & 0.630997 & 0.724818 \\ 0.630997 & 8.319236 & 0 \\ 0.724818 & 0 & 4.242317 \end{bmatrix}$$

11.2 雅可比 (Jacobi) 方法

$$\boldsymbol{A}_3 = \begin{bmatrix} 2.183296 & 0.595195 & 0 \\ 0.595195 & 8.319236 & 0.209521 \\ 0 & 0.209521 & 4.497468 \end{bmatrix}$$

$$\boldsymbol{A}_4 = \begin{bmatrix} 2.126094 & 0 & -0.020044 \\ 0 & 8.376437 & 0.208560 \\ -0.020044 & 0.208560 & 4.497468 \end{bmatrix}$$

$$\boldsymbol{A}_5 = \begin{bmatrix} 2.126094 & -0.001073 & -0.020015 \\ -0.001073 & 8.387619 & 0 \\ -0.020015 & 0 & 4.486287 \end{bmatrix}$$

$$\cdots,$$

$$\boldsymbol{A}_{10} = \begin{bmatrix} 2.125924 & 0 & 0 \\ 0 & 8.387619 & 0 \\ 0 & 0 & 4.486456 \end{bmatrix}$$

11.2 雅可比 (Jacobi) 方法

雅可比方法

```
function [D, R] = EigJacobi(A, ep)
%% 求对称矩阵全部特征值的Jacobi方法, 调用方法为
%%      [D, R] = EigJacobi(A, ep)
%% 其中A 为矩阵, ep 为精度要求, 缺省为1e-5
%%      it_max 为最大迭代次数, 缺省为100
%%      D 的对角线元素为特征值, R 的列为对应的特征向量
if nargin < 2  ep = 1e-5;     end
n = length(A); R = eye(n);
while 1
      % 求非对角线上的最大元素
      Amax = 0;
      for l = 1 : n-1
            for k = l+1 : n
                  if abs(A(l, k)) > Amax
```

11.2 雅可比 (Jacobi) 方法


```
% 终止准则
    if Amax < ep  break;  end
    % 计算三角函数
    d=(A(i,i)-A(j,j)) / (2*A(i,j));
    if abs(d) < 1e-10
        t = 1;
    else
        t = sign(d)/(abs(d)+sqrt(d^2+1));
    end
    c = 1/sqrt(t^2+1); s = c * t; %cos, sin
    % 旋转计算
    for l = 1 : n
        if l == i
          Aii=A(i,i)*c^2+A(j,j)*s^2+2*A(i,j)*s*c;
          Ajj=A(i,i)*s^2+A(j,j)*c^2-2*A(i,j)*s*c ;
          A(i,j)=(A(i,i)-A(j,j))*s*c+A(i,j)*(c^2-s
```

11.2 雅可比 (Jacobi) 方法

```
        Rli = R(l,i)*c + R(l,j)*s;
        Rlj = -R(l,i)*s + R(l,j)*c;
       R(l,i) = Rli; R(l,j) = Rlj;
        end
end
D = diag(diag(A));
```

对于例 11.4 中矩阵，输入

```
>> A= [4 2 2;2 5 1;2 1 6]; [D,R] = EigJacobi(A)
```

11.2 雅可比 (Jacobi) 方法

得到

```
D =
2.1259          0          0
0          8.3876          0
0          0          4.4865
R =
0.8280     0.5387     -0.1555
-0.4697     0.5149     -0.7172
-0.3062     0.6669      0.6793
```

11.3 QR 方法

QR 方法是目前求一般矩阵全部特征值最有效并最广泛应用的方法之一。在实际应用中，经常用 Householder 变换将实矩阵 $\boldsymbol{A}$ 分解为正交矩阵 $\boldsymbol{Q}$ 和上三角矩阵 $\boldsymbol{R}$.

先介绍矩阵的 Householder 变换。

设 $\boldsymbol{u}$ 为 n 维实向量且 $\boldsymbol{u}^T\boldsymbol{u}=1$(单位向量), 称矩阵

$$\boldsymbol{H}=\boldsymbol{I}-2\boldsymbol{u}\boldsymbol{u}^T,$$

为 Householder 矩阵。

容易验证

$$\boldsymbol{H}^T\boldsymbol{H}=I,\boldsymbol{H}^T=\boldsymbol{H},$$

即 $\boldsymbol{H}$ 为对称正交阵。

对任意具有相同长度的向量 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$，即 $\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2$，令

$$\boldsymbol{u}=\frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2},$$

即 $\boldsymbol{u}^T\boldsymbol{u}=1$，$\boldsymbol{H}=\boldsymbol{I}-2\boldsymbol{u}\boldsymbol{u}^T=\boldsymbol{I}-2\dfrac{(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^T}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}$,

$$\begin{aligned}\boldsymbol{H}\boldsymbol{x}&=\boldsymbol{x}-2\frac{(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^T}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}\boldsymbol{x}\\&=\boldsymbol{x}-2\frac{(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}^T\boldsymbol{x}-\boldsymbol{y}^T\boldsymbol{x})}{\boldsymbol{x}^T\boldsymbol{x}-2\boldsymbol{y}^T\boldsymbol{x}+\boldsymbol{y}^T\boldsymbol{y}}\\&\overset{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2}{\underset{\boldsymbol{x}^T\boldsymbol{x}=\boldsymbol{y}^T\boldsymbol{y}}{=\!=\!=\!=}}\boldsymbol{x}-(\boldsymbol{x}-\boldsymbol{y})=\boldsymbol{y}.\end{aligned}$$

称 $\boldsymbol{H}$ 是把向量 $\boldsymbol{x}$ 变换到 $\boldsymbol{y}$ 的 Householder 变换矩阵。有的书中叫镜像矩阵。

一、QR 分解

设 $\boldsymbol{A}$ 的第 j 个列向量为 $\boldsymbol{a}_j$, 即

$$\boldsymbol{A}=[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n],\ \boldsymbol{a}_j=(a_{1j},a_{2j},\cdots,a_{nj})^T,$$

$(j=1,2,\cdots,n.)$

取 $\boldsymbol{x}=\boldsymbol{a}_1, \boldsymbol{y}=-\mathrm{sign}(a_{11})\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1$, $\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2$, 把向量 $\boldsymbol{x}$ 变换到 $\boldsymbol{y}$ 的 Householder 变换矩阵为

$$\boldsymbol{H}_1=\boldsymbol{I}-\frac{2(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^T}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}, \boldsymbol{H}_1\boldsymbol{a}_1=-\mathrm{sign}(a_{11})\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1.$$

记 $\alpha_1=-\mathrm{sign}(a_{11})\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1$, 从而

$$\boldsymbol{H}_1\boldsymbol{A}=\boldsymbol{H}_1\cdot(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)=\begin{bmatrix}\alpha_1 & a_{12}^{(2)} & \cdots & a_{1n}^{(2)}\\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)}\end{bmatrix}$$

取 $\overline{\boldsymbol{a}}_2^2=(a_{22}^{(2)},a_{32}^{(2)},\cdots,a_{n2}^{(2})$, $\overline{\boldsymbol{e}}_1=(1,0,\cdots,0)$ 为 $(n-1)$ 为单位向量，类似于上述方法，可以构造 $(n-1)$ 维空间上的 Householder 矩阵

$$\overline{\boldsymbol{H}}_2=\overline{\boldsymbol{I}}-2\overline{\boldsymbol{u}}\overline{\boldsymbol{u}}^T,$$

使得

$$\overline{\boldsymbol{H}}_2\overline{\boldsymbol{a}}_2^{(2)} = \alpha_2\overline{\boldsymbol{e}}_1.$$

又

$$\boldsymbol{H}_2 = \begin{bmatrix} 1 & 0 \\ 0 & \overline{\boldsymbol{H}}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \overline{\boldsymbol{I}}_2 - 2\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \overline{\boldsymbol{I}}_2 \end{bmatrix} - 2\begin{bmatrix} 0 \\ \overline{\boldsymbol{u}} \end{bmatrix}\begin{bmatrix} 0 \\ \overline{\boldsymbol{u}} \end{bmatrix}^T$$

是 n 维空间上的 Householder 矩阵，并且

$$\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & \overline{\boldsymbol{H}}_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & \alpha_2 & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}$$

一直进行下去，直到 $(n-1)$ 步，得

$$\boldsymbol{H}_{n-1}\boldsymbol{H}_{n-2}\cdots\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A} \triangleq \boldsymbol{R}.$$

$\boldsymbol{R}$ 为上三角矩阵， $\boldsymbol{H}_{n-1}\boldsymbol{H}_{n-2}\cdots\boldsymbol{H}_2\boldsymbol{H}_1$ 是正交矩阵，从而 $\boldsymbol{Q}=\boldsymbol{H}_1^T\boldsymbol{H}_2^T\cdots\boldsymbol{H}_{n-2}^T\boldsymbol{H}_{n-1}^T$ 也是正交矩阵，进而得到矩阵 $\boldsymbol{A}$ 的 QR 分解： $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$.

二、QR 方法

记 $\boldsymbol{A}_1$ 为 $\boldsymbol{A}$ ，对 $\boldsymbol{A}_1$ 进行 QR 分解： $\boldsymbol{A}_1=\boldsymbol{Q}_1\boldsymbol{R}_1$ ，然后把矩阵 $\boldsymbol{Q}_1,\boldsymbol{R}_1$ 逆序相乘，得 $\boldsymbol{A}_2=\boldsymbol{R}_1,\boldsymbol{Q}_1$, 再以 $\boldsymbol{A}_2$ 代替 $\boldsymbol{A}_1$。重复以上步骤，依次类推，因此 QR 方法的迭代公式为

$$\begin{cases}\boldsymbol{A}_k=\boldsymbol{Q}_k\boldsymbol{R}_k,\\ \boldsymbol{A}_{k+1}=\boldsymbol{R}_k\boldsymbol{Q}_k,\end{cases}$$

11.3 QR 方法

这样产生的矩阵序列有两个基本性质：

(1) 每个 $\boldsymbol{A}_k$ 都与 $\boldsymbol{A}$ 相似。其实，$\boldsymbol{A}_{k+1}=\boldsymbol{R}_k\boldsymbol{Q}_k=\boldsymbol{Q}_k^{-1}\boldsymbol{A}\boldsymbol{Q}_k=\cdots=\boldsymbol{Q}_k^{-1}\boldsymbol{Q}_{k-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k$，若令 $\boldsymbol{P}_k=\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k$, 则上式可以写成

$$\boldsymbol{A}_{k+1}=\boldsymbol{P}_k^{-1}\boldsymbol{A}_1\boldsymbol{P}_k=\boldsymbol{P}_k^T\boldsymbol{A}_1\boldsymbol{P}_k,$$

故 $\boldsymbol{A}_{k+1}$ 与 $\boldsymbol{A}_1$ 相似，从而有相同的特征值。

(2) 若令 $\boldsymbol{H}_k=\boldsymbol{R}_k\boldsymbol{R}_{k-1}\cdots\boldsymbol{R}_1$，则有

$$\begin{aligned}\boldsymbol{P}_k\boldsymbol{H}_k&=\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k\boldsymbol{R}_k\boldsymbol{R}_{k-1}\cdots\boldsymbol{R}_1\\&=\boldsymbol{P}_{k-1}\boldsymbol{A}_k\boldsymbol{H}_{k-1}=\cdots=\boldsymbol{A}_k,\end{aligned}$$

即 $\boldsymbol{A}_k$ 的 QR 分解是 $\boldsymbol{A}_k=\boldsymbol{P}_k\boldsymbol{H}_k$.

显然，为了求得 $\boldsymbol{A}$ 的特征值，只需序列 $\boldsymbol{A}_k$ 能趋于分块三角阵即可，从而可求出 $\boldsymbol{A}$ 的特征值。

例 (11.5)

用 QR 分解方法求矩阵 $\boldsymbol{A}$ [a]的特征值，其中

$$\boldsymbol{A}=\begin{bmatrix}5 & -2 & -5 & -1\\ 1 & 0 & -3 & 2\\ 0 & 2 & 2 & -3\\ 0 & 0 & 1 & -2\end{bmatrix}$$

解： $\boldsymbol{A}$ 的特征值为 $-1,4,1\pm 2\mathrm{i}$，现对 $\boldsymbol{A}$ 进行 QR 分解，

$$\boldsymbol{A}_1=\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{R}_1$$

$$=\begin{bmatrix}0.980581 & -0.037743 & 0.176604 & 0.076472\\ 0.196116 & 0.188713 & -0.883022 & -0.382360\\ 0 & 0.981307 & 0.176604 & 0.076472\\ 0 & 0 & 0.397360 & -0.917663\end{bmatrix}$$

$$\times\begin{bmatrix}5.099020 & -1.961161 & -5.491252 & -0.588348\\ 0 & 2.038099 & 1.585188 & -2.528752\end{bmatrix}$$

$$\boldsymbol{A}_2 = \boldsymbol{R}_1 \boldsymbol{Q}_1$$

$$= \begin{bmatrix} 4.615385 & -5.951149 & 1.428692 & 1.259780 \\ 0.399704 & 1.940171 & -2.524558 & 1.662478 \\ 0 & 2.469568 & -0.853801 & 3.190620 \\ 0 & 0 & 0.303869 & -0.701754 \end{bmatrix}.$$

继续算下去，得

$$\boldsymbol{A}_{12} = \begin{bmatrix} 4.001004 & -5.444217 & -3.035584 & -0.844712 \\ 0.000440 & 1.297699 & -1.059398 & 0.324239 \\ 0 & 3.860187 & 0.701345 & -4.014928 \\ 0 & 0 & 0.000037 & -1.000049 \end{bmatrix}$$

定义

对于矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 若存在标量 λ 和向量 $\boldsymbol{x}$, 使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{B}\boldsymbol{x}$$

则称 λ 是矩阵 $\boldsymbol{A}$ 关于矩阵 $\boldsymbol{B}$ 的广义特征值, 称 $\boldsymbol{x}$ 为对应特征值 λ 的特征向量。

11.4 广义特征值与特征向量

定义

对于矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 若存在标量 λ 和向量 $\boldsymbol{x}$, 使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{B}\boldsymbol{x}$$

则称 λ 是矩阵 $\boldsymbol{A}$ 关于矩阵 $\boldsymbol{B}$ 的广义特征值, 称 $\boldsymbol{x}$ 为对应特征值 λ 的特征向量。

在 MATLAB 中，eig() 函数也可以求矩阵的广义特征值与特征向量，其使用格式为

[V,D] = eig(A,B)

[V,D] = eig(A,B,flag)

返回 D 的对角线上的元素为矩阵 $\boldsymbol{A}$ 的特征值，矩阵 V 的每一列是对应的特征向量，即 AV=BVD。参数 flag 详细说明了计算特征值和特征向量的方法。

11.4 广义特征值与特征向量

例

求广义特征值. 设

$$\boldsymbol{A} = \begin{bmatrix} 1.746 & 0.940 \\ 1.246 & 1.898 \end{bmatrix}$$

,

$$\boldsymbol{B} = \begin{bmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{bmatrix}$$

(1) 求矩阵 $\boldsymbol{A}$ 关于矩阵 $\boldsymbol{B}$ 的广义特征值;
(2) 用矩阵 $\boldsymbol{A}\boldsymbol{B}^{-1}$ 和 $\boldsymbol{A}^{-1}\boldsymbol{B}$ 的的特征值的倒数求求矩阵 $\boldsymbol{A}$ 关于矩阵 $\boldsymbol{B}$ 的广义特征值;

解：(1) 特征值：2, 1071333.99986002；(2) 数值上看不出差异。

参考文献

姚仰新, 罗家洪, 庄楚强.
高等工程数学 (第三版).
华南理工大学出版社, 广州, 2016-08.

薛毅.
数值分析与科学计算.
科学出版社, 北京, 2011-06.

邱启荣.
矩阵论与数值分析——理论及其工程应用.
清华大学出版社, 北京, 2013-01.

薛定宇, 陈阳泉.
高等应用数学问题的 MATLAB 求解 (第三版).
清华大学出版社，北京，2013-10.

欢迎使用!

PRODUCED BY: WANG Fu-Chang

ADDRESS: Institute of Disaster Prevention of CEA Sanhe, 065201, China

QQ : 1626194111

EMAIL: cidpmath@126.com