目录

- 引言
- 幂法
- 反幂法
- Householder方法
- QR算法

研究动机

- 前面几节讨论的是求解矩阵最大（小）特征值及其对应特征向量的方法
- 若要求求出所有特征值及其特征向量，应该用什么方法呢？
- 下面将讨论的以正交相似变换为基础的一类方法即是解决这类问题的方法
- 首先，讨论对于一般实矩阵$\boldsymbol{A} \in \mathbf{R}^{n\times n}$利用正交相似变换约化到什么程度的问题

正交相似变换

- 定理9.10 设$A \in \mathbf{R}^{n\times n}$，则存在一个正交阵$\boldsymbol{R}$，使

$$\boldsymbol{R}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{R} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1s} \\ & T_{22} & \cdots & T_{2s} \\ & & \ddots & \vdots \\ & & & T_{ss} \end{bmatrix}$$

其中对角块为一阶或二阶矩阵

- 每一个一阶对角块即为$\boldsymbol{A}$的实特征值
- 每一个二阶对角块的两个特征值是$\boldsymbol{A}$的一对共轭复特征值
- 参考Golub和Loan的《Matrix Computations》

上Hessenberg（海森堡）阵

- 定义9.2 一方阵$\boldsymbol{B}$，如果当$i > j+1$时有$b_{ij}=0$，则称$\boldsymbol{B}$为上Hessenberg阵，即

$$\boldsymbol{B}=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & \ddots & \ddots & \vdots \\ & & b_{n,n-1} & b_{nn} \end{bmatrix}$$

- 本节讨论如下两个问题
  1. 用正交相似变换约化一般实矩阵为上Hessenberg阵
  2. 用正交相似变换约化对称阵为对称三对角阵
  - 这样，就变成求转换后矩阵的特征值问题

初等反射阵

- **定义9.3** 设向量$\boldsymbol{w}$满足$\|\boldsymbol{w}\|_2 = 1$，矩阵$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}$称为初等反射阵，记作$\boldsymbol{H}(\boldsymbol{w})$，即

$$\boldsymbol{H}(\boldsymbol{w}) = \begin{bmatrix} 1-2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1-2w_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -2w_{n-1}w_n \\ -2w_nw_1 & \cdots & -2w_nw_{n-1} & 1-2w_n^2 \end{bmatrix}$$

其中，$\boldsymbol{w} = (w_1, w_2, \cdots, w_n)^{\mathrm{T}}$

- **定理9.11** 初等反射阵$\boldsymbol{H}$是对称阵($\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{H}$)、正交阵($\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} = \boldsymbol{I}$)和对合阵($\boldsymbol{H}^2 = \boldsymbol{I}$)

初等反射阵（续）

- 只证$\boldsymbol{H}$的正交性，其他显然

$$\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} = \boldsymbol{H}^2 = \left(\boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)\left(\boldsymbol{I} - 2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}\right)$$

$$= \boldsymbol{I} - 4\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}} + 4\boldsymbol{w}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}\right)\boldsymbol{w}^{\mathrm{T}} = \boldsymbol{I}$$

- 设向量$\boldsymbol{u} \neq \boldsymbol{0}$，则显然$\boldsymbol{H} = \boldsymbol{I} - 2\dfrac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|_2^2}$是一个初等反射阵

Householder方法

- 初等反射阵在计算上的意义是它能用来约化矩阵，例如设向量$\boldsymbol{a} \neq \boldsymbol{0}$，可选择一初等反射阵$\boldsymbol{H}$使$\boldsymbol{H}\boldsymbol{a} = \sigma \boldsymbol{e}_1$，这种约化矩阵的方法称为Householder方法

- **定理9.12** 设$\boldsymbol{x}, \boldsymbol{y}$为两个不相等的$n$维向量，$\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2$，则存在一个初等反射阵$\boldsymbol{H}$，使$\boldsymbol{H}\boldsymbol{x} = \boldsymbol{y}$
  - 令$\boldsymbol{w} = \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2}$，则得到一个初等反射阵

定理9.12的证明

$$\boldsymbol{H}=\boldsymbol{I}-2\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}=\boldsymbol{I}-2\frac{(\boldsymbol{x}-\boldsymbol{y})}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}\left(\boldsymbol{x}^{\mathrm{T}}-\boldsymbol{y}^{\mathrm{T}}\right)$$

- 易得

$$\boldsymbol{H}\boldsymbol{x}=\boldsymbol{x}-2\frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}\left(\boldsymbol{x}^{\mathrm{T}}-\boldsymbol{y}^{\mathrm{T}}\right)\boldsymbol{x}$$

$$=\boldsymbol{x}-2\frac{(\boldsymbol{x}-\boldsymbol{y})\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}\right)}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}$$

- 根据$\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2$，可得

$$\|\boldsymbol{x}-\boldsymbol{y}\|_2^2=(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y})=2\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}\right)$$

- 所以

$$\boldsymbol{H}\boldsymbol{x}=\boldsymbol{x}-(\boldsymbol{x}-\boldsymbol{y})=\boldsymbol{y}$$

定理9.12的应用

- 易知，$\boldsymbol{w}=\frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}$是使$\boldsymbol{H}\boldsymbol{x}=\boldsymbol{y}$成立的唯一长度等于1的向量（不计符号）

- **推论** 设向量$\boldsymbol{x}\in\mathbf{R}^n\ (\boldsymbol{x}\neq\mathbf{0})$，$\sigma=\pm\|\boldsymbol{x}\|_2$，且$\boldsymbol{x}\neq-\sigma\boldsymbol{e}_1$，则存在一个初等反射阵

$$\boldsymbol{H}=\boldsymbol{I}-2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|_2^2}\equiv\boldsymbol{I}-\rho^{-1}\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$$

使$\boldsymbol{H}\boldsymbol{x}=-\sigma\boldsymbol{e}_1,\boldsymbol{u}=\boldsymbol{x}+\sigma\boldsymbol{e}_1,\rho=\|\boldsymbol{u}\|_2^2/2$

  - 应用定理9.12，令$\boldsymbol{y}=-\sigma\boldsymbol{e}_1$，已知$\|\boldsymbol{y}\|_2=\|\boldsymbol{x}\|_2$，$\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{x}+\sigma\boldsymbol{e}_1$

定理9.12的应用（续）

- 如何选择$\sigma = \pm\|\boldsymbol{x}\|_2$？
  - 设$\boldsymbol{x} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^{\mathrm{T}} \neq \boldsymbol{0}$，则

$$\boldsymbol{u} = \boldsymbol{x} + \sigma \boldsymbol{e}_1 = (\alpha_1 + \sigma, \alpha_2, \cdots, \alpha_n)^{\mathrm{T}}$$

  - 因此

$$\rho = \frac{\|\boldsymbol{u}\|_2^2}{2} = \frac{(\alpha_1 + \sigma)^2 + \alpha_2^2 + \cdots + \alpha_n^2}{2}$$

$$= \frac{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 + \sigma^2}{2} + \alpha_1 \sigma$$

$$= \sigma^2 + \alpha_1 \sigma \quad = \sigma(\sigma + \alpha_1) \qquad (\sigma = \pm\|\boldsymbol{x}\|_2)$$

  - 如果σ和α_1异号，那么计算$(\sigma + \alpha_1)$时有效数字可能损失，取σ和α_1有相同的符号，即取

$$\sigma = \operatorname{sgn}(\alpha_1)\,\|\boldsymbol{x}\|_2$$

用正交相似变换约化矩阵

- 设

$$A=\left[\begin{array}{c|cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array}\right] \equiv \begin{bmatrix} a_{11} & \boldsymbol{A}_{12}^{(1)} \\ \boldsymbol{a}_{21}^{(1)} & \boldsymbol{A}_{22}^{(1)} \end{bmatrix}$$

- 步1：不妨设$\boldsymbol{a}_{21}^{(1)} \neq \mathbf{0}$，否则这一步不需约化，选择初等反射阵$\boldsymbol{R}_1$使$\boldsymbol{R}_1\boldsymbol{a}_{21}^{(1)} = -\sigma_1\boldsymbol{e}_1$，其中

$$\begin{cases} \sigma_1 = \operatorname{sgn}(a_{21})\left(\sum_{i=2}^{n} a_{i1}^2\right)^{1/2} \\ \boldsymbol{u}_1 = \boldsymbol{a}_{21}^{(1)} + \sigma_1\boldsymbol{e}_1 \\ \rho_1 = \dfrac{1}{2}\|\boldsymbol{u}_1\|_2^2 = \sigma_1(\sigma_1 + a_{21}) \\ \boldsymbol{R}_1 = \boldsymbol{I} - \rho_1^{-1}\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}} \end{cases}$$

用正交相似变换约化矩阵（续）

- 令$\boldsymbol{U}_1 = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R}_1 \end{bmatrix}$，则

$$\boldsymbol{A}_2 = \boldsymbol{U}_1\boldsymbol{A}_1\boldsymbol{U}_1 = \begin{bmatrix} a_{11} & \boldsymbol{A}_{12}^{(1)}\boldsymbol{R}_1 \\ \boldsymbol{R}_1\boldsymbol{a}_{21}^{(1)} & \boldsymbol{R}_1\boldsymbol{A}_{22}^{(1)}\boldsymbol{R}_1 \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{A}_{11}^{(2)} & \boldsymbol{a}_{21}^{(2)} & \boldsymbol{A}_{13}^{(2)} \\ \boldsymbol{O} & \boldsymbol{a}_{22}^{(2)} & \boldsymbol{A}_{23}^{(2)} \end{bmatrix}$$

其中

$\boldsymbol{A}_{11}^{(2)} \in \mathbf{R}^{2\times 1}$，　$\boldsymbol{a}_{22}^{(2)} \in \mathbf{R}^{n-2}$，　$\boldsymbol{A}_{23}^{(2)} \in \mathbf{R}^{(n-2)\times(n-2)}$

用正交相似变换约化矩阵（续）

- 步k：设对$\boldsymbol{A}$已进行了第$k-1$步正交相似约化，即$\boldsymbol{A}_k$有形式

$$\boldsymbol{A}_k=\boldsymbol{U}_{k-1}\boldsymbol{A}_{k-1}\boldsymbol{U}_{k-1}=\left[\begin{array}{cccc:c:ccc}
a_{11}^{(1)} & a_{12}^{(2)} & \cdots & & a_{1k}^{(k)} & a_{1,k+1}^{(k)} & \cdots & a_{1n}^{(k)} \\
-\sigma_1 & a_{22}^{(2)} & \cdots & & a_{2k}^{(k)} & a_{2,k+1}^{(k)} & \cdots & a_{2n}^{(k)} \\
 & \ddots & \ddots & & \vdots & \vdots & & \vdots \\
 & & -\sigma_{k-1} & & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{k,n}^{(k)} \\
\hdashline
 & & & & a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\
 & & & & \vdots & \vdots & & \vdots \\
 & & & & a_{nk}^{(k)} & a_{n,k+1}^{(k)} & \cdots & a_{nn}^{(k)}
\end{array}\right]$$

用正交相似变换约化矩阵（续）

- 步k：设对$\boldsymbol{A}$已进行了第$k-1$步正交相似约化，即$\boldsymbol{A}_k$有形式

$$\boldsymbol{A}_k=\boldsymbol{U}_{k-1}\boldsymbol{A}_{k-1}\boldsymbol{U}_{k-1}=\begin{bmatrix}\boldsymbol{A}_{11}^{(k)} & \boldsymbol{a}_{12}^{(k)} & \boldsymbol{A}_{13}^{(k)}\\ \boldsymbol{O} & \boldsymbol{a}_{22}^{(k)} & \boldsymbol{A}_{23}^{(k)}\end{bmatrix}$$

其中

$$\boldsymbol{A}_{11}^{(k)}\in\mathbf{R}^{k\times(k-1)},\quad \boldsymbol{a}_{22}^{(2)}\in\mathbf{R}^{n-k},\quad \boldsymbol{A}_{23}^{(k)}\in\mathbf{R}^{(n-k)\times(n-k)}$$

- 设$\boldsymbol{a}_{22}^{(k)}\neq\mathbf{0}$，选择初等反射阵$\boldsymbol{R}_k$，使$\boldsymbol{R}_k\boldsymbol{a}_{22}^{(k)}=-\sigma_k\boldsymbol{e}_1$，其中

用正交相似变换约化矩阵（续）

$$\begin{cases} \sigma_k = \operatorname{sgn}\left(a_{k+1,k}^{(k)}\right)\left(\sum_{i=k+1}^{n} a_{ik}^2\right)^{1/2} \\ \boldsymbol{u}_k = \boldsymbol{a}_{22}^{(k)} + \sigma_k \boldsymbol{e}_1 \\ \rho_k = \dfrac{1}{2}\|\boldsymbol{u}_k\|_2^2 = \sigma_k(\sigma_k + a_{k+1,n}^{(k)}) \\ \boldsymbol{R}_k = \boldsymbol{I} - \rho_k^{-1}\boldsymbol{u}_k\boldsymbol{u}_k{}^{\mathrm{T}} \end{cases}$$

- 设$\boldsymbol{U}_k = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R}_k \end{bmatrix}$，则

$$\boldsymbol{A}_{k+1} = \boldsymbol{U}_k\boldsymbol{A}_k\boldsymbol{U}_k = \begin{bmatrix} \boldsymbol{A}_{11}^{(k)} & \boldsymbol{a}_{12}^{(k)} & \boldsymbol{A}_{13}^{(k)}\boldsymbol{R}_k \\ \boldsymbol{O} & \boldsymbol{R}_k\boldsymbol{a}_{22}^{(k)} & \boldsymbol{R}_k\boldsymbol{A}_{23}^{(k)}\boldsymbol{R}_k \end{bmatrix}$$

用正交相似变换约化矩阵（续）

$$\boldsymbol{A}_{k+1}=\boldsymbol{U}_k\boldsymbol{A}_k\boldsymbol{U}_k=\begin{bmatrix}\boldsymbol{A}_{11}^{(k)} & \boldsymbol{a}_{12}^{(k)} & \boldsymbol{A}_{13}^{(k)}\boldsymbol{R}_k \\ \boldsymbol{O} & -\sigma_k\boldsymbol{e}_1 & \boldsymbol{R}_k\boldsymbol{A}_{23}^{(k)}\boldsymbol{R}_k\end{bmatrix}$$

- 由上式知，$\boldsymbol{A}_{k+1}$的左上角$k+1$阶子阵为上Hessenberg阵，从而约化又进了一步
- 重复这过程，直到

$$\boldsymbol{A}_{n-1}=\boldsymbol{U}_{n-2}\cdots\boldsymbol{U}_2\boldsymbol{U}_1\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_2\cdots\boldsymbol{U}_{n-2}$$

$$=\begin{bmatrix}a_{11} & \times & \times & \cdots & \times \\ -\sigma_1 & a_{22}^{(2)} & \times & \cdots & \times \\ & -\sigma_2 & a_{33}^{(3)} & \ddots & \vdots \\ & & \ddots & \ddots & \times \\ & & & -\sigma_{n-1} & a_{nn}^{(n-1)}\end{bmatrix}$$

用正交相似变换约化矩阵（续）

- **定理9.13** 设$\boldsymbol{A} \in \mathrm{R}^{n\times n}$，则存在初等反射阵$\boldsymbol{U}_1, \boldsymbol{U}_2, \cdots, \boldsymbol{U}_{n-2}$使

  $\boldsymbol{U}_{n-2} \cdots \boldsymbol{U}_2\boldsymbol{U}_1\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_2 \cdots \boldsymbol{U}_{n-2} = \boldsymbol{C}$（上Hessenberg阵）

  在$\boldsymbol{A}_k \rightarrow \boldsymbol{A}_{k+1} = \boldsymbol{U}_k\boldsymbol{A}_k\boldsymbol{U}_k$的进一步约化中，需要计算$\boldsymbol{R}_k$和$\boldsymbol{A}_{13}^{(k)}\boldsymbol{R}_k$，$\boldsymbol{R}_k\boldsymbol{A}_{23}^{(k)}\boldsymbol{R}_k$

$$\boldsymbol{A}_{k+1} = \boldsymbol{U}_k\boldsymbol{A}_k\boldsymbol{U}_k = \begin{bmatrix} \boldsymbol{A}_{11}^{(k)} & \boldsymbol{a}_{12}^{(k)} & \boldsymbol{A}_{13}^{(k)}\boldsymbol{R}_k \\ \boldsymbol{O} & -\sigma_k\boldsymbol{e}_1 & \boldsymbol{R}_k\boldsymbol{A}_{23}^{(k)}\boldsymbol{R}_k \end{bmatrix}$$

- 用初等反射阵正交相似约化$\boldsymbol{A}$为上Hessenberg阵，大约需要$\frac{5}{3}n^3$次乘法运算

特征值求解

- 由于$\boldsymbol{U}_k$都是正交阵，所以$\boldsymbol{A}_1 \sim \boldsymbol{A}_2 \sim \cdots \sim \boldsymbol{A}_{n-1}$，求$\boldsymbol{A}$的特征值问题，就转化为求上Hessenberg阵$\boldsymbol{C}$的特征值问题
  - 由定理9.13，记$\boldsymbol{P} = \boldsymbol{U}_{n-2} \cdots \boldsymbol{U}_2 \boldsymbol{U}_1$则

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{C}$$

  - $\boldsymbol{y}$是$\boldsymbol{C}$的对应特征值λ的特征向量，则$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{y}$为$\boldsymbol{A}$的对应特征值$\lambda$的特征向量，且

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{U}_1 \boldsymbol{U}_2 \cdots \boldsymbol{U}_{n-2} \boldsymbol{y}$$

$$= \left(\boldsymbol{I} - \lambda_1^{-1} \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathrm{T}}\right) \cdots \left(\boldsymbol{I} - \lambda_{n-2}^{-1} \boldsymbol{u}_{n-2} \boldsymbol{u}_{n-2}^{\mathrm{T}}\right) \boldsymbol{y}$$

约化对称矩阵

- **定理9.14** 设$A \in \mathbb{R}^{n\times n}$为对称阵，则存在初等反射阵$U_1, U_2, \cdots, U_{n-2}$使

$$A_{n-1} = U_{n-2}\cdots U_1 A U_1 \cdots U_{n-2}$$

$$= \begin{bmatrix} c_1 & b_1 & & & \\ b_1 & c_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & c_{n-1} & b_{n-1} \\ & & & b_{n-1} & c_n \end{bmatrix} \equiv C$$

- 由定理9.13，存在初等反射阵$U_1, U_2, \cdots, U_{n-2}$，使$A_{n-1}$为上Hessenberg阵，但$A_{n-1}$又为对称阵，因此$A_{n-1}$为对称三对角阵

约化对称矩阵（续）

$$A_{k+1}=U_kA_kU_k=\begin{bmatrix} A_{11}^{(k)} & a_{12}^{(k)} & A_{13}^{(k)}R_k \\ O & -\sigma_k e_1 & R_kA_{23}^{(k)}R_k \end{bmatrix}$$

- 由于对称性，在由$A_k \to A_{k+1}=U_kA_kU_k$的计算过程中，只需计算$R_k$和$R_kA_{23}^{(k)}R_k$
- 更进一步，只计算$R_kA_{23}^{(k)}R_k$对角线下面的元素
- 将对称阵A用初等反射阵正交相似约化为对称三对角阵约需做$\frac{2}{3}n^3$次乘法运算

$$A_2 = U_1 A_1 U_1 = \begin{bmatrix} a_{11} & A_{12}^{(1)} R_1 \\ R_1 a_{21}^{(1)} & R_1 A_{22}^{(1)} R_1 \end{bmatrix}$$

例9.5

- 用Householder方法将下述矩阵化为Hessenberg阵

$$A_1 = A = \begin{bmatrix} -4 & -3 & -7 \\ 2 & 3 & 2 \\ 4 & 2 & 7 \end{bmatrix}$$

1. 对于$k = 1$，确定变换

$$U_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & & R_1 \end{bmatrix}, \qquad a_{21}^{(1)} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

其中R_1为初等反射阵且使

$$R_1 a_{21}^{(1)} = -\sigma_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \sigma_1 = \left\| a_{21}^{(1)} \right\|_2 = \sqrt{20} \approx 4.472136$$

例9.5（续）

$$\boldsymbol{A}_2 = \boldsymbol{U}_1\boldsymbol{A}_1\boldsymbol{U}_1 = \begin{bmatrix} a_{11} & \boldsymbol{A}_{12}^{(1)}\boldsymbol{R}_1 \\ \boldsymbol{R}_1\boldsymbol{a}_{21}^{(1)} & \boldsymbol{R}_1\boldsymbol{A}_{22}^{(1)}\boldsymbol{R}_1 \end{bmatrix}$$

$$\boldsymbol{u}_1 = \boldsymbol{a}_{21}^{(1)} + \sigma_1\rho_1 = \begin{pmatrix} 2+\sqrt{20} \\ 4 \end{pmatrix} \approx \begin{pmatrix} 6.472136 \\ 4 \end{pmatrix}$$

$$\rho_1 = \sigma_1(\sigma_1 + a_{21}) = \sqrt{20}\left(\sqrt{20}+2\right) \approx 28.94427$$

$$\boldsymbol{R}_1 = \boldsymbol{I} - \rho_1^{-1}\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}$$

2. 计算$\boldsymbol{R}_1\boldsymbol{A}_{22}^{(1)}$，记

$$\boldsymbol{A}_{22}^{(1)} = \begin{bmatrix} 3 & 2 \\ 2 & 7 \end{bmatrix} \equiv (\boldsymbol{a}_1, \boldsymbol{a}_2)$$

$$\boldsymbol{R}_1\boldsymbol{A}_{22}^{(1)} = (\boldsymbol{R}_1\boldsymbol{a}_1, \boldsymbol{R}_1\boldsymbol{a}_2) = \begin{bmatrix} -3.130496 & -7.155419 \\ -1.788855 & 1.341640 \end{bmatrix}$$

$$\boldsymbol{R}_1\boldsymbol{a}_i = \left(\boldsymbol{I} - \rho_1^{-1}\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}\right)\boldsymbol{a}_i = \boldsymbol{a}_i - \left(\rho_1^{-1}\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{a}_i\right)\boldsymbol{u}_1 \quad (i = 1,2)$$

例9.5（续）

$$A_2 = U_1 A_1 U_1 = \begin{bmatrix} a_{11} & A_{12}^{(1)} R_1 \\ R_1 a_{21}^{(1)} & R_1 A_{22}^{(1)} R_1 \end{bmatrix}$$

3. 计算$A_{12}^{(1)} R_1$及$\left(R_1 A_{22}^{(1)}\right) R_1$，即计算

$$\begin{bmatrix} A_{12}^{(1)} \\ \left(R_1 A_{22}^{(1)}\right) \end{bmatrix} R_1 \equiv \begin{bmatrix} b_1^{\mathrm{T}} \\ b_2^{\mathrm{T}} \\ b_3^{\mathrm{T}} \end{bmatrix} R_1 = \begin{bmatrix} b_1^{\mathrm{T}} R_1 \\ b_2^{\mathrm{T}} R_1 \\ b_3^{\mathrm{T}} R_1 \end{bmatrix}$$

$$= \begin{bmatrix} 7.602634 & -0.447212 \\ 7.800003 & -0.399999 \\ -0.399999 & 2.200000 \end{bmatrix}$$

其中

$$b_i^{\mathrm{T}} R_1 = b_i^{\mathrm{T}} - \left(\rho_1^{-1} b_i^{\mathrm{T}} u_1\right) u_1^{\mathrm{T}} \quad (i = 1,2,3)$$

例9.5（续）

$$A_2 = U_1 A_1 U_1 = \begin{bmatrix} a_{11} & A_{12}^{(1)} R_1 \\ R_1 a_{21}^{(1)} & R_1 A_{22}^{(1)} R_1 \end{bmatrix}$$

4. 得到$A_2 = U_1 A_1 U_1$

$$A_2 = \begin{bmatrix} -4 & A_{12}^{(1)} R_1 \\ -\sigma_1 & \\ 0 & R_1 A_{22}^{(1)} R_1 \end{bmatrix}$$

$$= \begin{bmatrix} -4 & 7.602634 & -0.447212 \\ -4.472136 & 7.800003 & -0.399999 \\ 0 & -0.399999 & 2.200000 \end{bmatrix}$$

为上Hessenberg阵

目录

- 引言
- 幂法
- 反幂法
- Householder方法
- QR算法

QR算法

- Francis（在1961年、1962年）利用矩阵的QR分解建立了计算矩阵特征值的QR方法
- QR方法是一种变换方法，是计算一般矩阵（中小型矩阵）全部特征值问题的最有效的方法之一
- 目前，QR方法主要用来计算
  1. 上Hessenberg阵的全部特征值问题，
  2. 对称三对角阵的全部特征值问题
- QR方法具有收敛快、算法稳定等特点

QR算法（续）

- 对于一般矩阵$\boldsymbol{A} \in \mathbf{R}^{n\times n}$(或对称阵)
  - 首先用Householder方法将$\boldsymbol{A}$化为上Hessenberg阵$\boldsymbol{B}$(或对称三对角阵)
  - 然后再用QR方法计算$\boldsymbol{B}$的全部特征值问题
  - 注意： QR方法可以直接处理任意矩阵$\boldsymbol{A}$

- 平面旋转矩阵来约化矩阵
  - 同样是一种用正交相似变换约化矩阵的方法
  - 是QR方法的基础

平面旋转矩阵

$$
\boldsymbol{P}_{i,j} = \begin{bmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & c & & & & s & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & -s & & & & c & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{bmatrix}
$$

- 其中$c = \cos\theta$，$s = \sin\theta$

平面旋转矩阵（续）

- 引理1 设$\boldsymbol{x}=(\alpha_1,\alpha_2,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n)^{\mathrm{T}}$，其中$\alpha_i,\alpha_j$不全为零，则可选一平面旋转矩阵$\boldsymbol{P}_{ij}$使$\boldsymbol{y}\equiv\boldsymbol{P}_{ij}\boldsymbol{x}=\left(\alpha_1,\alpha_2,\cdots,\alpha_i^{(1)},\cdots,\alpha_j^{(1)},\cdots,\alpha_n\right)^{\mathrm{T}}$，其中

$$\alpha_i^{(1)}=\sqrt{\alpha_i^2+\alpha_j^2} \tag{9.4.2}$$

$$\alpha_j^{(1)}=0 \tag{9.4.3}$$

$$\begin{cases} c=\alpha_i/\sqrt{\alpha_i^2+\alpha_j^2} \\ s=\alpha_j/\sqrt{\alpha_i^2+\alpha_j^2} \end{cases} \tag{9.4.4}$$

平面旋转算法

- 注意到，$\boldsymbol{P}_{ij}\boldsymbol{x}$只改变$\boldsymbol{x}$的第$i$个及第$j$个元素
- 用平面旋转阵进行左变换可产生一算法，设给定$\boldsymbol{x}=\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$，计算$c=\cos\theta, s=\sin\theta$，$v=\sqrt{\alpha^2+\beta^2}$，使$\boldsymbol{P}_{ij}\boldsymbol{x}=\begin{bmatrix}c & s\\ -s & c\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}=\begin{bmatrix}v\\ 0\end{bmatrix}$
- 为了防止溢出，将x规范化，有

$$\eta\equiv\|\boldsymbol{x}\|_\infty=\max\{|\alpha|,|\beta|\}\neq 0,\quad \boldsymbol{x}'=\boldsymbol{x}/\eta=\begin{bmatrix}\alpha/\eta\\ \beta/\eta\end{bmatrix}$$

  - 于是

$$\begin{cases}c'=c, \quad s'=s\\ v'=v/\eta\end{cases}$$

用平面旋转矩阵约化矩阵

- **定理9.15** 如果$\boldsymbol{A}$为非奇异矩阵，则存在正交矩阵$\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_{n-1}$（即一系列平面旋转矩阵）使

$$\boldsymbol{P}_{n-1}, \cdots, \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} \equiv \boldsymbol{R} \tag{9.4.5}$$

且$r_{ii} > 0 \quad (i = 1,2,\cdots,n-1)$

- 由于$\boldsymbol{A}$的第1列一定存在$a_{j1} \neq 0$，于是，如果$a_{j1} \neq 0\ (j = 2,3,\cdots,n)$，应用算法1，即存在平面旋转矩阵$\boldsymbol{P}_{12}, \boldsymbol{P}_{13}, \cdots, \boldsymbol{P}_{1n}$，使

定理9.15的证明

$$\boldsymbol{P}_{1n}\cdots\boldsymbol{P}_{13}\boldsymbol{P}_{12}\boldsymbol{A}=\begin{bmatrix} r_{11} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}\equiv\boldsymbol{A}^{(2)}$$

- 记$\boldsymbol{P}_{1n}\cdots\boldsymbol{P}_{12}=\boldsymbol{P}_1$，则

$$\boldsymbol{P}_1\boldsymbol{A}=\boldsymbol{A}^{(2)}$$

- 同理如果$a_{2j}^{(2)}\neq 0\ (j=3,\cdots,n)$，应用算法1，存在平面旋转矩阵$\boldsymbol{P}_{23},\cdots,\boldsymbol{P}_{2n}$(记$\boldsymbol{P}_2=\boldsymbol{P}_{2n}\cdots\boldsymbol{P}_{23}$)，使

定理9.15的证明（续）

$$P_2P_1A=\begin{bmatrix} r_{11} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} \\ & r_{22} & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ & & \vdots & & \vdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}$$

- 注意到，构造$\boldsymbol{P}_2$的矩阵数量减少了1个

- 重复上述过程，最后得到：存在正交阵$\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_{n-1}$，使式(9.4.5)成立

基本QR方法

- □ **定理9.16** 如果$\boldsymbol{A} \in \mathbf{R}^{n\times n}$为非奇异矩阵，则$\boldsymbol{A}$可分解为一正交阵$\boldsymbol{Q}$与上三角阵$\boldsymbol{R}$的乘积，即$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$，且当$\boldsymbol{R}$对角元素都为正数时分解唯一
- □ 设$\boldsymbol{A} = (a_{ij}) \in \boldsymbol{R}^{n\times n}$且对$\boldsymbol{A}$进行QR分解，即$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$，于是可得到一新矩阵

$$\boldsymbol{B} = \boldsymbol{R}\boldsymbol{Q} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} \quad (\boldsymbol{R} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A})$$

  - ■ 显然，$\boldsymbol{B}$是由$\boldsymbol{A}$经过正交相似变换得到，因此$\boldsymbol{B}$与$\boldsymbol{A}$特征值相同
  - ■ 再对$\boldsymbol{B}$进行QR分解，又可得一新的矩阵，重复这过程可得到矩阵序列

基本QR方法（续）

- 按上述递推法则构造矩阵序列$\{\boldsymbol{A}_k\}$的过程
- **定理9.17**（基本QR方法） 设$\boldsymbol{A}_1=\boldsymbol{A}\in\mathbf{R}^{n\times n}$， QR算法为

$$\begin{cases}\boldsymbol{A}_k=\boldsymbol{Q}_k\boldsymbol{R}_k & (\boldsymbol{Q}_k^{\mathrm{T}}\boldsymbol{Q}_k=\boldsymbol{I},\boldsymbol{R}_k\text{为上三角阵})\\ \boldsymbol{A}_{k+1}=\boldsymbol{R}_k\boldsymbol{Q}_k & (k=1,2,\dots)\end{cases}$$

记$\widetilde{\boldsymbol{Q}}_k\equiv\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k,\widetilde{\boldsymbol{R}}_k\equiv\boldsymbol{R}_k\cdots\boldsymbol{R}_2\boldsymbol{R}_1$则有

1. $\boldsymbol{A}_{k+1}$相似于$\boldsymbol{A}_k$， 即$\boldsymbol{A}_{k+1}=\boldsymbol{Q}_k^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{Q}_k$
2. $\boldsymbol{A}_{k+1}=(\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k)^{\mathrm{T}}\boldsymbol{A}_1(\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k)=\widetilde{\boldsymbol{Q}}_k^{\mathrm{T}}\boldsymbol{A}_1\widetilde{\boldsymbol{Q}}_k$
3. $\boldsymbol{A}^k$ 的QR分解式为$\boldsymbol{A}^k=\widetilde{\boldsymbol{Q}}_k\widetilde{\boldsymbol{R}}_k$

QR方法的收敛性

- 定理9.18 设$\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$
  1. 如果$\boldsymbol{A}$的特征值满足：
  $$|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|>0$$
  2. $\boldsymbol{A}$有标准形$\boldsymbol{A}=\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$，其中$\boldsymbol{D}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$

  则由QR算法产生的$\{\boldsymbol{A}_k\}$本质上收敛于上三角阵，即
  $$\boldsymbol{A}_k\xrightarrow{\text{本质上}}\boldsymbol{R}=\begin{bmatrix}\lambda_1 & \times & \cdots & \times\\ & \lambda_2 & \cdots & \vdots\\ & & \ddots & \times\\ & & & \lambda_n\end{bmatrix}\quad(k\to\infty)$$

QR方法的收敛性（续）

换言之

$$a_{ii}^{(k)} \to \lambda_i \quad (k \to \infty)$$

当$i > j$时，$a_{ij}^{(k)} \to 0 \quad (k \to \infty)$

当$i < j$时，$a_{ij}^{(k)}$极限不一定存在

- **定理9.19** 如果对称阵$\boldsymbol{A}$满足定理9.18条件，则由QR算法产生的$\{\boldsymbol{A}_k\}$收敛于对角阵

总结

- 引言
  - 特征多项式、特征值、特征向量
  - 特征值的性质
- 幂法
  - 幂法的流程、幂法的收敛性
  - 规范化幂法
  - 原点平移法、Rayleigh商加速法
- 反幂法
  - 反法的流程、反幂法的收敛性
  - 反幂法的加速

总结

- Householder方法
  - 正交相似变换、上Hessenberg阵
  - 初等反射阵、Householder方法
  - 用正交相似变换约化矩阵、约化对称矩阵

- QR算法
  - 平面旋转矩阵、用平面旋转矩阵约化矩阵
  - 矩阵的QR分解
  - 基本QR方法、QR方法的收敛性