作业8

(数值算法与案例分析)

李维杰

2024 年 11 月 7 日

题目1. 设$A \in \mathbb{C}^{n\times n}, x \in \mathbb{C}^n$.假设$X = [x, Ax, ..., A^{n-1}x]$是非奇异的.
证明: $X^{-1}AX$ 是一个上Hessenberg阵.

解答. 设$X^{-1}AX = H$,其中$H = (h_{i,j})$.则有$AX = XH$,即

$$\begin{bmatrix} Ax & A^2x & ... & A^nx \end{bmatrix} = \begin{bmatrix} x & Ax & ... & A^{n-1}x \end{bmatrix} H.$$

进一步展开,即得

$$A^kx = \sum_{i=1}^{n} h_{i,k}A^{i-1}x.$$

由于X是非奇异的,故$x, Ax, ..., A^{n-1}x$线性无关.则由上式可以直接得出

$$\begin{cases} h_{i,k} = 1, i = k+1 \\ h_{i,k} = 0, i \neq k+1 \end{cases}, k = 1, 2, ..., n-1.$$

故

$$H = \begin{bmatrix} 0 & 0 & \cdots & 0 & h_{1,n} \\ 1 & 0 & \cdots & 0 & h_{2,n} \\ 0 & 1 & \cdots & 0 & h_{3,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & h_{n,n} \end{bmatrix},$$

这是一个上Hessenberg阵.

题目2. 设$A \in \mathbb{C}^{n\times n}$是一个未缩减的上Hessenberg阵.其中未缩减的意思是对于$1 \leqslant i \leqslant n-1$,均有$A_{i+1,i} \neq 0$.假设$A$是奇异的.
证明: 0 特征值将在一步QR迭代后出现在矩阵的右下角.
另外考虑:保持A是一个奇异的上Hessenberg阵.当某些$A_{i+1,i} = 0$时会出现什么状况?

解答. 由于A是奇异的,则对A做QR分解后,可得分块矩阵$Q = \begin{bmatrix} \tilde{Q} & 0 \end{bmatrix}$和$R = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix}$.
故一步迭代后,即得

$$A' = RQ = \begin{bmatrix} \tilde{Q}\tilde{R} & 0 \\ 0 & 0 \end{bmatrix}.$$

也即0特征值将出现在一步迭代后矩阵的右下角.
当存在某些$A_{i+1,i} = 0$时,若采用Givens变换不断消去bulge的算法进行QR分解,则算法会在$G(i, i+1)$处中断导致迭代无法完成. 故此时必须采用其它QR分解的算法(如MGS等)对A进行QR分解,同上述证明可得0特征值会出现在一步迭代后矩阵的右下角.

题目3. 设

$$A = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & 1 & 0 & \\ & & & 1 & 0 \end{bmatrix}.$$

试评价下面两种算法作用于该矩阵时的收敛性情况:
(1) 朴素QR算法,
(2) Francis' double-shift QR算法.

解答. (1) 注意到A的一个QR分解就是$A = AI$,则一步迭代后变为$A^{(1)} = IA = A$, 故该矩阵在朴素QR算法下不会收敛.
(2) 由于矩阵A上Hessenberg化后仍为矩阵A,其右下角2×2子矩阵$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$的特征值为0,则对该矩阵使用Francis' double-shift QR算法的第一步迭代等价于朴素QR分解,由(1)知朴素QR迭代过程不会改变矩阵A,故在Francis' double-shift算法下该矩阵同样不会收敛.

题目4. 设

$$A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

设计一个算法来生成一个正交矩阵Q,使得

$$Q^T A Q = \begin{bmatrix} b & c \\ 0 & a \end{bmatrix}.$$

解答. 设$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$,则

$$Q^T A Q = \begin{bmatrix} a\cos^2\theta + b\sin^2\theta + c\sin\theta\cos\theta & -a\sin\theta\cos\theta + b\sin\theta\cos\theta + c\cos^2\theta \\ -a\sin\theta\cos\theta + b\sin\theta\cos\theta - c\sin^2\theta & a\sin^2\theta + b\cos^2\theta - c\sin\theta\cos\theta \end{bmatrix}.$$

于是有方程组:

$$\begin{cases} a\cos^2\theta + b\sin^2\theta + c\sin\theta\cos\theta = b \\ -a\sin\theta\cos\theta + b\sin\theta\cos\theta + c\cos^2\theta = c \\ -a\sin\theta\cos\theta + b\sin\theta\cos\theta - c\sin^2\theta = 0 \\ a\sin^2\theta + b\cos^2\theta - c\sin\theta\cos\theta = a \end{cases},$$

解此方程组,得$\tan\theta = \frac{b-a}{c}$,也即

$$\begin{cases} \sin\theta = \frac{b-a}{\sqrt{(b-a)^2+c^2}} \\ \cos\theta = \frac{c}{\sqrt{(b-a)^2+c^2}} \end{cases}.$$

由此即可生成符合题意的Q.随机生成10个矩阵A,并设$\tilde{A} = \begin{bmatrix} b & c \\ 0 & a \end{bmatrix}$,由上述算法确定$Q$,计算相应的$\left\| Q^T A Q - \tilde{A} \right\|_{\mathsf{F}}$如下图 (代码见Problem4.m)

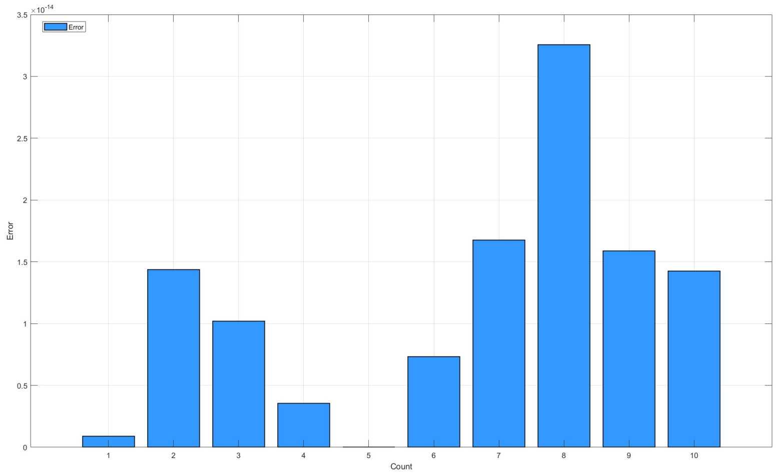

图 1: 对A进行正交相似变换后与目标矩阵$\tilde{A}$之间的误差

题目5. 执行下列用于矩阵上Hessenberg化的算法:

(a) 使用Householder reflections;

(b) 使用基于MGS正交化的Arnoldi变换.

随机生成一些矩阵并计算相应的Hessenberg分解 $A = QHQ^*$.通过计算$\|Q^*AQ - H\|_\mathsf{F}$ 和 $\|Q^*Q - I\|_\mathsf{F}$来检查程序的准确性.并说明有什么发现.

解答. (代码见Problem5.m)

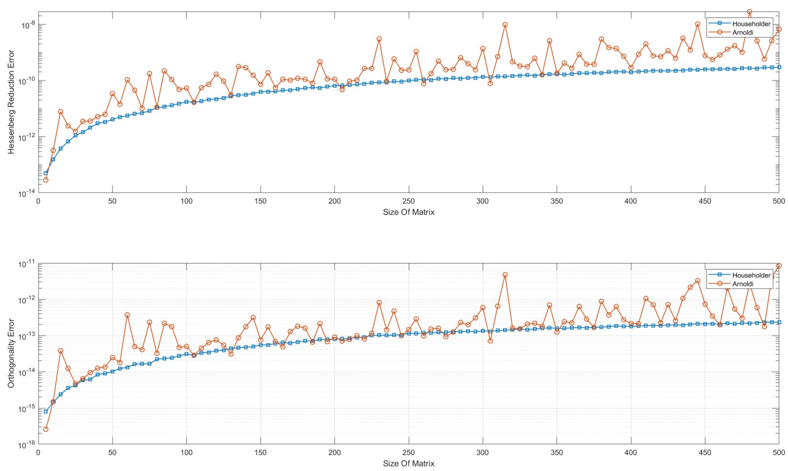

图 2: 两种算法的误差表现

由图知基于Householder变换的矩阵上Hessenberg化算法具有更小的误差和更优越的稳定性.