Advanced Numerical Analysis Homework 8

Michael Nelson

Throughout this homework, $\|\cdot\|$ denotes the ℓ_2-norm. We also let $\langle \cdot, \cdot \rangle$ denote the standard Euclidean inner-product on $\mathbb{C}^m$ (thus

$$\langle x, y \rangle = \sum_{i=1}^{m} x_i \overline{y}_i$$

for all $x, y \in \mathbb{C}^m$). Finally we set $\varepsilon = \varepsilon_{\text{machine}}$ to be the machine coefficient.

1 Problem 1

Exercise 1. 1. Let H be the initial input matrix for the shifted QR iteration. Show that

$$(H - \mu^{(k)} I) \cdots (H - \mu^{(2)} I)(H - \mu^{(1)} I) = \underline{Q}^{(k)} \underline{R}^{(k)}.$$

In practice, H is upper Hessenberg, but we do not need this assumption here.

2. Other than the Wilkinson shift, we may also let $\mu^{(k)} = h_{nn}^{(k-1)}$ if $h_{n,n-1}^{(k-1)}$ is small. Assume, for example, that

$$H^{(k-1)} = \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \delta & h_{n,n}^{(k-1)} \end{pmatrix}.$$

After the application of $n-2$ Givens rotations to $H^{(k-1)} - h_{n,n}^{(k-1)} I$, we have the intermediate matrix

$$H_{\text{tmp}}^{(k-1)} = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & \delta & 0 \end{pmatrix}$$

(make sure you understand why it is of this form), and the last Givens rotation is needed on the left before we compute $H^{(k)}$ by transposed Givens rotations. Show that the new matrix $H^{(k)} = R^{(k)} Q^{(k)} + h_{n,n}^{(k-1)} I$ satisfies $h_{n,n-1}^{(k)} = -b\delta^2/(a^2 + \delta^2)$. What does this observation suggest, if $|h_{n,n-1}^{(k-1)}| = |\delta| \ll 1$, and either $|b| < 2|a|$ (δ can be arbitrary) or if $|\delta| < a^2/|b|$?

3. What can we say about $h_{n,n-1}^{(k)}$ if A is real symmetric, such that $H^{(k-1)}$ is also real symmetric (hence tridiagonal)? In particular, does this entry decrease more slowly or more rapidly in the symmetric case than in the nonsymmetric case?

Solution 1. 1. First let us recall that for all k, the QR algorithms gives us the following relations:

$$H^{(k-1)} - \mu^{(k)} = Q^{(k)} R^{(k)}$$
$$H^{(k)} - \mu^{(k)} = R^{(k)} Q^{(k)}.$$

Using this, we wish to show that for all k, we have

$$\prod_{i=1}^{k} (H - \mu^{(i)}) = (H - \mu^{(k)}) \cdots (H - \mu^{(2)})(H - \mu^{(1)}) = \underline{Q}^{(k)} \underline{R}^{(k)}, \tag{1.1}$$

where $\underline{Q}^{(k)} = Q^{(1)}Q^{(2)}\cdots Q^{(k)}$ and $\underline{R}^{(k)} = R^{(k)}R^{(k-1)}\cdots R^{(1)}$. We prove (1.1) by induction on k. The base $k = 1$ is clear, so assume we have shown (1.1) for all $j < k$ where $k > 1$, and we wish to show it holds for k. We have

$$\begin{aligned}
\underline{Q}^{(k)}\underline{R}^{(k)} &= \underline{Q}^{(k-1)}Q^{(k)}R^{(k)}\underline{R}^{(k-1)} \\
&= \underline{Q}^{(k-1)}(H^{(k-1)} - \mu^{(k)})\underline{R}^{(k-1)} \\
&= \underline{Q}^{(k-1)}H^{(k-1)}\underline{R}^{(k-1)} - \mu^{(k)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= \underline{Q}^{(k-1)}(R^{(k-1)}Q^{(k-1)} + \mu^{(k-1)})\underline{R}^{(k-1)} - \mu^{(k)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= \underline{Q}^{(k-1)}R^{(k-1)}Q^{(k-1)}\underline{R}^{(k-1)} + \mu^{(k-1)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} - \mu^{(k)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= \underline{Q}^{(k-1)}R^{(k-1)}Q^{(k-1)}\underline{R}^{(k-1)} + \mu^{(k-1)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} - \mu^{(k)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= (H - \mu^{(k-1)})\underline{Q}^{(k-1)}\underline{R}^{(k-1)} + \mu^{(k-1)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} - \mu^{(k)}\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= (H - \mu^{(k)})\underline{Q}^{(k-1)}\underline{R}^{(k-1)} \\
&= (H - \mu^{(k)})\prod_{i=1}^{k-1}(H - \mu^{(i)}) \\
&= \prod_{i=1}^{k}(H - \mu^{(i)}).
\end{aligned}$$

2. For this part of the problem we simplify notation by setting $H := H^{(k-1)}$. Thus $H' := H - h_{nn}$ has the form

$$H' = \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \delta & 0 \end{pmatrix}.$$

If G_1 is the first givens Gotations, then G_1H' has the form

$$G_1H' = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \delta & 0 \end{pmatrix}.$$

Similarly, if G_2 and G_3 are the second and third Givens rotations respectively, then $G_3G_2G_1H'$ has the form

$$G_3G_2G_1H' = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & \delta & 0 \end{pmatrix}.$$

At this point, we set

$$G_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & a/\lambda & \delta/\lambda \\ 0 & 0 & 0 & -\delta/\lambda & a/\lambda \end{pmatrix}$$

to be the last Givens rotation where $\lambda = \sqrt{a^2 + \delta^2}$ and we set $Q^* = G_4G_3G_2G_1$ (thus $Q = G_1^*G_2^*G_3^*G_4^*$ since each G_i is orthonormal). Then Q^*H' has the form

$$Q^*H' = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \lambda & ab/\lambda \\ 0 & 0 & 0 & 0 & -b\delta/\lambda \end{pmatrix} := R.$$

Finally we construct a new matrix $H'' := RQ + h_{nn}$ and the goal is to show that

$$h''_{n,n-1} = -b\delta^2/(a^2+\delta^2) = -b(\delta/\lambda)^2.$$

This is clear though since

$$\begin{pmatrix} \lambda & ab/\lambda \\ 0 & -\delta b/\lambda \end{pmatrix} \begin{pmatrix} a/\lambda & -\delta/\lambda \\ \delta/\lambda & a/\lambda \end{pmatrix} = \begin{pmatrix} \times & \times \\ -b(\delta/\lambda)^2 & \times \end{pmatrix}.$$

where $\left(\begin{smallmatrix} \lambda & ab/\lambda \\ 0 & -\delta b/\lambda \end{smallmatrix}\right)$ is the bottom right 2×2 submatrix of R and where $\left(\begin{smallmatrix} a/\lambda & -\delta/\lambda \\ \delta/\lambda & a/\lambda \end{smallmatrix}\right)$ is the bottom right 2×2 submatrix of G_4^*. In particular, if $|\delta| \ll 1$ and either $|b| < 2|a|$ (δ can be arbitrary) or if $|\delta| < a^2/|b|$, then $|h''_{n,n-1}| \ll 1$.

3. In the symmetric case, we have $b = \delta$. Thus $h''_{n,n-1} = -\delta^3/(a^2+\delta^2)$. It decreases more rapidly in the symmetric case.

2 Problem 2

Exercise 2. Implement the single-shift QR step in MATLAB; that is, given an upper Hessenberg $H^{(k-1)}$ and shift $\mu^{(k)}$, we compute

$$Q^{(k)}R^{(k)} = H^{(k-1)} - \mu^{(k)} I$$

by Givens rotations and then use these Givens rotations to compute

$$H^{(k)} = R^{(k)}Q^{(k)} + \mu^{(k)} I.$$

Make simple changes in my code to enforce the use of single (Wilkinson) shift only, even if complex arithmetic is needed. Assemble your single shift code with the up-loaded subroutines. Test it with the matrix obtained by

```
load west0479;
A = full(west0479);
```

Compare the eigenvalues of your final $H^{(k)}$ (use ordeig) with those of A. Be aware that the ordering of eigenvalues must be consistent to make a meaningful comparison.

Solution 2.

3 Problem 3

Exercise 3. Implement the Arnoldi's method without and with reorthogonalization, and test the orthogonality of the column vectors in U_{50} for the matrix A generated by

```
u = cos((0:2048)/2048*pi);
A = vander(u);
```

Is the reorthogonalization effective for generating an orthonormal basis? Use Arnoldi with reorthogonalization to compute the 11 dominant eigenvalues and eigenvectors of aerofoil new, using $m = 30, 60, 100, 150$ dimensional Krylov subspaces. For each m, plot all eigenvalues $\{\lambda_i\}_{i=1}^n$ of A together with the eigenvalues $\{\mu_i\}_{i=1}^m$ of H_m on the complex plane. Intuitively, how do $\{\mu_i\}_{i=1}^m$ approximate $\{\lambda_i\}_{i=1}^n$ as m increases? Give the relative eigenresidual norm

$$\frac{\|AU_m w_i - \mu_i U_m w_i\|}{\|AU_m w_i\|}$$

for $1 \le i \le 11$ of the desired eigenpairs for each m in a table.

Solution 3. Table:

$\frac{\|AU_m w_i - \mu_i U_m w_i\|}{\|AU_m w_i\|}$	$m = 30$	$m = 60$	$m = 100$	$m = 150$
$i = 1$				
$i = 2$				
$i = 3$				
$i = 4$				
$i = 5$				
$i = 6$				
$i = 7$				
$i = 8$				
$i = 9$				
$i = 10$				
$i = 11$				

4 Problem 4

Exercise 4. Read the implicit double-shifted QR step for real nonsymmetric matrices and the overall QR iteration. Then read my codes to see how the described algorithms are implemented. Debug my code to compare numerically if the double shifted QR step gives a new upper Hessenberg matrix that is numerically the same as the upper Hessenberg matrix obtained by using the pair of complex conjugate shifts successively in two single-shifted QR steps. Use your own words to summarize (not to repeat) the overall QR iteration.

Solution 4.

Appendix

QR Post Process

```
function [Q,R] = QRpostprocess(Q,R)

[m,n] = size(Q);

for i = 1:n
   if R(i,i) < 0
      R(i,:) = -R(i,:);
      Q(:,i) = -Q(:,i);
   end;
end;
```

QR Algorithm

```
function [Qk,Rk,Ak] = QRalgorithm(A,k)

[m,n] = size(A); Ak = A; Rk = eye(n); Qk = eye(n);

for i = 1:k
   [Q,R] = qr(Ak);
   [Q,R] = QRpostprocess(Q,R);
   Ak = R*Q;
   Qk = Qk*Q;
   Rk = R*Rk;
end;
```

Simultaneous Iteration

```
function [Qk,Rk,Ak] = SimultaneousIteration(A,k)

[m,n] = size(A); Qk = eye(n); Rk = eye(n);

for i = 1:k
   Z = A*Qk;
   [Qk,R] = qr(Z);
   [Qk,R] = QRpostprocess(Qk,R);
   Ak = Qk'*A*Qk;
   Rk = R*Rk;
end;
```