W.D. Henshaw **Math 6800: Solutions for Problem Set 8**

1. NLA 24.1 *For each of the following statements, prove that ...*

Solution:

(a) TRUE: if $\lambda \in \lambda(A)$, then $Ax = \lambda x$, $x \neq 0$, then $(A - \mu I)x = (\lambda - \mu)x$ and $\lambda - \mu \in \lambda(A - \mu I)$.

(b) FALSE: If $A = I$ then A has eigenvalues 1 but -1 is not an eigenvalue.

(c) TRUE: if $Ax = \lambda x$, $x \neq 0$, then taking conjugates of both sides gives $A\bar{x} = \bar{\lambda}\bar{x}$.

(d) TRUE: if A is nonsingular then $\lambda \neq 0$. If $Ax = \lambda x$, $x \neq 0$, then multiplying by A^{-1} gives $x = \lambda A^{-1}x$, i.e. $A^{-1}x = \lambda^{-1}x$

(e) FALSE: The matrix

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

has all eigenvalues 0 but $A \neq 0$.

(f) TRUE: If $A = A^*$ and $Ax = \lambda x$, $x \neq 0$, then $A^*Ax = \lambda A^*x = \lambda\bar{\lambda}x = |\lambda|^2 x$ But the eigenvalues of A^*A are the squares of the singular values σ_i of A.

(g) TRUE: If $A = X\mathrm{diag}(\lambda)X^{-1}$ then $A = \lambda XIX^{-1} = \lambda I$

2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & \epsilon & \epsilon \\ \epsilon & 0 & \epsilon \\ \epsilon & \epsilon & 0 \end{bmatrix},$$

with ϵ a small positive perturbation, with $\epsilon \leq 10^{-3}$.

(a) Estimate the locations of the eigenvalues of $A + B$ by using Gershgorin's theorem.

(b) Improve the estimate for $\lambda_1 \approx 1$ by judicious choice of diagonal similiarity transformation of the form

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix},$$

for some $d > 0$.

Solution:

(a) Gershgorin's theorem implies that the eigenvalues lie in the union of the three overlapping disks:

$$\begin{aligned} \mathcal{D}_1 &: |z - 1| \leq 2\epsilon, \\ \mathcal{D}_2 &: |z - 2| \leq 2\epsilon, \\ \mathcal{D}_3 &: |z - 3| \leq 2\epsilon \end{aligned}$$

and thus since the disks are distinct we can say that

$$\begin{aligned} |\lambda_1 - 1| &\leq 2\epsilon, \\ |\lambda_2 - 2| &\leq 2\epsilon, \\ |\lambda_3 - 3| &\leq 2\epsilon. \end{aligned}$$

(a) Under a diagonal scaling

$$D^{-1}(A+B)D = \begin{bmatrix} 1 & d\epsilon & d\epsilon \\ d^{-1}\epsilon & 2 & \epsilon \\ d^{-1}\epsilon & \epsilon & 3 \end{bmatrix}$$

Now the Gershgorin disks are

$$\begin{aligned} \mathcal{D}_1 &: |z-1| \leq 2d\epsilon, \\ \mathcal{D}_2 &: |z-2| \leq \frac{\epsilon}{d} + \epsilon, \\ \mathcal{D}_3 &: |z-3| \leq \frac{\epsilon}{d} + \epsilon \end{aligned}$$

To obtain a better estimate for λ_1 we can choose d to be $\mathcal{O}(\epsilon)$ but not so small that $\mathcal{D}_1$ overlaps $\mathcal{D}_2$ or $\mathcal{D}_3$. To avoid $\mathcal{D}_1$ overlapping $\mathcal{D}_2$ we should choose

$$\begin{aligned} & 1 + 2d\epsilon < 2 - \frac{\epsilon}{d} - \epsilon, \\ \implies \quad & \frac{\epsilon}{d} < 1 - \epsilon - 2d\epsilon \end{aligned}$$

This implies we can choose $d = \epsilon + \mathcal{O}(\epsilon^2)$ and we obtain the estimate

$$|\lambda_1 - 1| \leq 2\epsilon^2 + \mathcal{O}(\epsilon^3).$$

3. (NLA 26.3) One of the best known results of eigenvalue perturbation theory is the *Bauer-Fike theorem. Suppose $A \in \mathbb{C}^{m\times m}$ is diagonalizable wih $A = V\Lambda V^{-1}$, and let $\delta A \in \mathbb{C}^{m\times m}$ be arbitrary. The every eigenvalue of $A + \delta A$ lies in at least one of the m circular disks in the complex plane of radius $\kappa(V)\|\delta A\|_2$ centred at the eigenvalues of A, where κ is the 2-norm condition number.*

(a) Prove the Bauer-Fike thereom by using the equivalence of conditions (i) and (iv) in Exercise 26.1.

(b) Suppose that A is normal. Show that for each eigenvalue $\tilde{\lambda}_j$ of $A + \delta A$, there is an eigenvalue λ_j of A such that

$$|\tilde{\lambda}_j - \lambda_j| \leq \|\delta A\|_2. \tag{1}$$

Solution:

For this problem let $\|\cdot\|$ denote the 2-norm.

(a) Let λ_j denote an eigenvalue of A and let $z = \tilde{\lambda}$ denote an eigenvalue of $A + \delta A$. Futhermore define

$$\epsilon = \|\delta A\|. \tag{2}$$

Then statement (i) in exercise 24.2 is satfisfied: *z is an eigenvalue of $A + \delta A$ for some δA with* $\|\delta A\| \leq \epsilon$.

But statement (i) implies statement (ii) and thus

$$\|(\tilde{\lambda} I - A)^{-1}\| \geq \frac{1}{\epsilon} = \frac{1}{\|\delta A\|}, \tag{3}$$

Whence

$$\begin{aligned}
\frac{1}{\|\delta A\|} &\leq \|(\tilde{\lambda} I - A)^{-1}\| = \|(\tilde{\lambda} I - V\Lambda V^{-1})^{-1}\| = \| \left(V(\tilde{\lambda} I - \Lambda)V^{-1} \right)^{-1} \| && (4)\\
&= \|V(\tilde{\lambda} I - \Lambda)^{-1} V^{-1}\| && (5)\\
&\leq \|V\| \|V^{-1}\| \, \|(\tilde{\lambda} I - \Lambda)^{-1}\| && (6)\\
&= \kappa(V) \max_j |\tilde{\lambda} - \lambda_j|^{-1} = \kappa(V) \frac{1}{\min_j |\tilde{\lambda} - \lambda_j|}. && (7)
\end{aligned}$$

Whence

$$\min_j |\tilde{\lambda} - \lambda_j| \leq \kappa(V) \|\delta A\| \tag{8}$$

and thus there exists one $\lambda_j \in \lambda(A)$ such that $\tilde{\lambda}$ lies in the disk with centre λ_j and radius $\kappa(V)\|\delta A\|$.

(b) If A is normal then A is unitarily diagonalizable,

$$A = U\Lambda U^*, \qquad U^* U = I, \tag{9}$$

and $\kappa(U) = 1$. Thus, from part (a)

$$\min_j |\tilde{\lambda} - \lambda_j| \leq \|\delta A\|, \tag{10}$$

and therefore there is an eigenvalue λ_j such that

$$|\tilde{\lambda} - \lambda_j| \leq \|\delta A\|. \tag{11}$$

4. Write a Matlab code `[W,H] = hessenberg(A)` to transform an $m \times m$ matrix A to upper Hessenberg form, H, by similarity transformations using Householder reflectors,

$$A = QHQ^*.$$

Here Q is represented implicitly in terms of the Householder vectors v_k stored in W. Also write a Matlab function `[Q] = formQh(A)` that takes W and generates the matrix Q.

Test your routine on the $m \times m$ matrix $A = [a_{ij}]$ with entries

$$\begin{aligned}
a_{ij} &= 9, && \text{for } i = j,\\
a_{ij} &= \frac{1}{(i+j)} && \text{for } i \neq j
\end{aligned}$$

and $m = 5$. Check that your routines are correct by confirming that H is upper Hessenberg, Q is unitary and $A = QHQ^*$.
Output A, H, W, Q, $\|Q^*Q - I\|_2$, and $\|A - QHQ^*\|_2$.

Solution:

The codes are given below. Here are the results,

```
A =
   9.000000000000000   0.333333333333333   0.250000000000000   0.200000000000000   0.166666666666667
   0.333333333333333   9.000000000000000   0.200000000000000   0.166666666666667   0.142857142857143
   0.250000000000000   0.200000000000000   9.000000000000000   0.142857142857143   0.125000000000000
   0.200000000000000   0.166666666666667   0.142857142857143   9.000000000000000   0.111111111111111
   0.166666666666667   0.142857142857143   0.125000000000000   0.111111111111111   9.000000000000000
H =
   9.000000000000000  -0.491313432432789                   0                   0                   0
  -0.491313432432789   9.428927612471918   0.107982271026731  -0.000000000000001   0.000000000000000
                   0   0.107982271026730   8.846327621703727   0.039957947387906                   0
                   0                   0   0.039957947387906   8.850664789996072  -0.020796583661351
                   0                   0                   0  -0.020796583661355   8.874079975828286
W =
                   0                   0                   0                   0                   0
   0.916093208017819                   0                   0                   0                   0
   0.277722913024081  -0.828062814556595                   0                   0                   0
   0.222178330419265  -0.372096036007008  -0.799238799047971                   0                   0
   0.185148608682721  -0.419352495087941  -0.601013595600264                   0                   0
Q =
   1.000000000000000                   0                   0                   0                   0
                   0  -0.678453531552758   0.675417393438013  -0.284994216277152  -0.047858613159621
                   0  -0.508840148664569  -0.166616441869197   0.751858607324624   0.384745758908068
                   0  -0.407072118931655  -0.452430095455764   0.029992667631721  -0.792905882562814
                   0  -0.339226765776379  -0.557994009525314  -0.593790679590698   0.470085647032515
Hessenberg: || Q^* Q - I || = 6.31e-16,  || A - Q H Q^* || = 7.87e-15
```

Listing 1: ps8.m

```
% Problem set 8

clear; % clear variables
format long; % show more digts on the output
% format short; % show more digts on the output

% --- Construct the matrix:
m=5;
A=zeros(m,m);
for i=1:m
  for j=1:m
    if i==j
      A(i,j) = 9;
    else
      A(i,j) = 1/( i+j );
    end;
  end;
end
A


[H,W]=hessenberg(A);
H
W

Q = formQh(W);

```

```
Q

fprintf('Hessenberg: || Q^* Q - I || = %8.2e,  || A - Q H Q^* || = %8.2e\n',norm(Q'*Q-eye(m)),norm(A-Q*H*Q',2));
```

Listing 2: hessenberg.m

```
function [H,W] = hessenberg( A )
%
%   Compute an implicit representation of the similarity transform to an
%   upper Hessenberg matrix H,
%         A = Q H Q^*
%
%   A (input) : m x m matrix
%   W (output) : m x m lower triangular matrix with columns the Housholder vectors v_k
%   H (output) : m x m upper Hessenberg matrix
%

[m,n]=size(A);

W=zeros(m,n);

for k=1:m-2

  vk = A(k+1:m,k); % partial column k of A
  % Householder vector vk = x + sign(x1) ||x|| e_k :

  % Note: sign(0)=0 may cause failure
  % vk(1) = vk(1) + sign(vk(1))*norm(vk);
  if( vk(1)>0 ) s = 1; else s=-1; end
  vk(1) = vk(1) + s*norm(vk);
  vk = vk/norm(vk,2);

  A(k+1:m,k:n) = A(k+1:m,k:n) - (2*vk)*(vk'*A(k+1:m,k:n));

  A(1:m,k+1:m) = A(1:m,k+1:m) - (A(1:m,k+1:m)*vk)*(2*vk');

  W(k+1:m,k)=vk;  % save vk in lower triangular part of W

end
H=A;
```

Listing 3: formQh.m

```
function [Q] = formQh( W )
%
%   Compute the unitary matrix Q given the output W from the function hessenberg
%
%   W (input) : m x m lower triangular matrix with columns the Housholder vectors v_k
%   Q (output) : mxm unitary matrix
%

[m,n]=size(W);

```

```
Q=eye(m);

% Q = Q1 * Q2 * ...
% Qk = I - 2 vk vk^*
for k=m-2:-1:1
  Q(k+1:m,:) = Q(k+1:m,:) - 2*W(k+1:m,k)*(W(k+1:m,k)'*Q(k+1:m,:)); % this could be optimized
end
```