MATH-6800
Vignesh Ramakrishnan
RIN: 662028006
Due: Thursday December 1, 2022

MATH 6800: Problem Set 8

1. For each of the following statements, prove that it is true or give an example to show it is false. Throughout, $A \in \mathbb{C}^{m\times m}$ unless otherwise indicated, and "ew" stands for eigenvalue.

   (a) If λ is an ew of A and $\mu \in \mathbb{C}$, then $\lambda - \mu$ is an ew of $A - \mu I$
   Let $x \in \mathbb{C}^m$ be an eigen vector of A. Then,

   $$\begin{aligned} Ax &= \lambda x \\ Ax - \mu x &= \lambda x - \mu x \\ (A - \mu I)\, x &= (\lambda - \mu)\, x \end{aligned}$$

   Hence this claim is true.

   (b) If A is real and λ is an ew of A, then so is $-\lambda$
   This is a false claim. Example,

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

   The eigenvalues are $\lambda = 1, 2$ and $-1, -2$ are not eigenvalues of A.

   (c) If A is real and λ is an ew of A, then so is $\bar{\lambda}$
   Let $x \in \mathbb{C}^m$ be an eigen vector of A. Then,

   $$\begin{aligned} Ax &= \lambda x \\ (Ax)^* &= (\lambda x)^* \\ Ax^* &= \bar{\lambda} x^* \end{aligned}$$

   Since, $A = A^*, A \in \mathbb{R}^{m\times m}$. Therefore this claim is true.

   (d) If λ is an ew of A and A is nonsingular, then $(\lambda)^{-1}$ is an eigenvalue of $(A)^{-1}$.
   Let $x \in \mathbb{C}^m$ be an eigen vector of A.

   $$\begin{aligned} Ax &= \lambda x \\ (A)^{-1} Ax &= \lambda (A)^{-1} x \\ \frac{1}{\lambda} x &= (A)^{-1} x \end{aligned}$$

   Therefore this claim is true.

   (e) If ew's of A are 0, then $A = 0$.
   Let,

   $$A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}$$

   The eigenvalues of this matrix are 0, but A is non-zero. Hence, that claim is false.

(f) If A is hermitian and λ is an ew of A, then $|\lambda|$ is a singular value of A.
The square-root of the eigen values of A^*A are the singular values of A. Let $x \in \mathbb{C}^m$ be an eigen vector of A.

$$\begin{aligned} Ax &= \lambda x \\ A^*Ax &= \lambda A^*x \\ (A^*A)\, x &= \lambda^2 x \end{aligned}$$

Singular values of A, $\sigma = \sqrt{\lambda^2} = |\lambda|$. Hence this is true.

(g) If A is diagonalizable and all its ew's are equal, then A is diagonal.
$A = U\Lambda U^{-1} \implies A = \lambda I$. hence, A is a diagonal matrix. This is true.

2. Let,

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

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

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

$$A + B = \begin{bmatrix} 1 & \varepsilon & \varepsilon \\ \varepsilon & 2 & \varepsilon \\ \varepsilon & \varepsilon & 3 \end{bmatrix}$$

From Gershgorin's theorem the three eigenvalues of this Matrix will lie within these three discs:

Disk	Center	Radius
D_1	1	$\lvert 2\varepsilon \rvert$
D_2	2	$\lvert 2\varepsilon \rvert$
D_3	3	$\lvert 2\varepsilon \rvert$

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

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

for some $d > 0$.
Through similarity transformation, $A + B$ has the same eigenvalues as $D^{-1}(A + B)\, D$. I choose D to be

$$D = \begin{bmatrix} d/\varepsilon & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & d/\varepsilon \end{bmatrix}$$

$$
\begin{aligned}
D^{-1}(A+B)D &= \begin{bmatrix} \varepsilon/d & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & \varepsilon/d \end{bmatrix} \begin{bmatrix} 1 & \varepsilon & \varepsilon \\ \varepsilon & 2 & \varepsilon \\ \varepsilon & \varepsilon & 3 \end{bmatrix} \begin{bmatrix} d/\varepsilon & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & d/\varepsilon \end{bmatrix} \\
&= \begin{bmatrix} \varepsilon/d & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & \varepsilon/d \end{bmatrix} \begin{bmatrix} d/\varepsilon & 2\varepsilon & d \\ d & 4 & d \\ d & 2\varepsilon & 3d/\varepsilon \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2\varepsilon^2/d & \varepsilon \\ d/2 & 2 & d/2 \\ \varepsilon & 2\varepsilon^2/d & 3 \end{bmatrix}
\end{aligned}
$$

For $\lambda \approx 1$, we have now a better estimate of what the eigenvalue would be as it is:

$$
\mathcal{D}_1 : |1 - z| \leq \left|2\varepsilon^2/d + \varepsilon\right| = \mathcal{O}(\varepsilon^2)
$$

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 with $A = V\Lambda V^{-1}$, 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 $\kappa(V)\,||\delta A||_2$ centered 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.

   Using (i) and (ii) conditions of 26.1, $||\delta A||_2 \leq \varepsilon$ (i), $\left|\left|\left(\tilde{\lambda}_j I - A\right)^{-1}\right|\right|_2 \geq \varepsilon^{-1}$ (ii). From (ii) we have $\varepsilon \left|\left|\left(\tilde{\lambda}_j I - A\right)\right|\right|_2 \geq 1$.

$$
\begin{aligned}
&||\delta A||_2 \left|\left|\left(\tilde{\lambda}_j I - A\right)^{-1}\right|\right|_2 \leq \varepsilon \left|\left|\left(\tilde{\lambda}_j I - A\right)^{-1}\right|\right|_2 \\
\implies &||\delta A||_2 \left|\left|\left(\tilde{\lambda}_j I - A\right)^{-1}\right|\right|_2 \geq 1
\end{aligned}
$$

   Now, $A = V\Lambda V^{-1}$,

$$
\begin{aligned}
||\delta A||_2 \left|\left|\left(\tilde{\lambda}_j I - V\Lambda V^{-1}\right)^{-1}\right|\right|_2 &\geq 1 \\
||\delta A||_2 \left|\left|\left(V\left(\tilde{\lambda}_j I - \Lambda\right)V^{-1}\right)^{-1}\right|\right|_2 &\geq 1 \\
||\delta A||_2 \, ||V^{-1}||_2 \left|\left|\left(\tilde{\lambda}_j I - \Lambda\right)^{-1}\right|\right|_2 ||V||_2 &\geq 1 \\
||\delta A||_2 \, \kappa(V) \left|\left|\left(\tilde{\lambda}_j I - \Lambda\right)^{-1}\right|\right|_2 &\geq 1
\end{aligned}
$$

   Here, $\left|\left|\left(\tilde{\lambda}_j I - \Lambda\right)^{-1}\right|\right|_2 = \max_{||x||_2 \neq 0} \frac{\left|\left|\left(\tilde{\lambda}_j I - \Lambda\right)^{-1} x\right|\right|_2}{||x||_2} = \frac{1}{\min_{\tilde{\lambda}_j \in \Lambda(A)} \left|\tilde{\lambda}_j - \lambda\right|}$.

$$
\left|\tilde{\lambda}_j - \lambda_j\right| \leq ||\delta A||_2 \, \kappa(V)
$$

(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

$$\left|\tilde{\lambda}_j - \lambda_j\right| \leq ||\delta A||_2$$

In this case, if A is normal, V is unitary in which case $\kappa(V) = 1$. Therefore,

$$\left|\tilde{\lambda}_j - \lambda_j\right| \leq ||\delta A||_2$$

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(W)** that takes W and generates the matrix Q. Test your routine on the $m \times m$ matrix $A = [a_{ij}]$ with entries

$$a_{ij} = 9, \text{ for } i = j,$$
$$a_{ij} = \frac{1}{i+j}, \text{ for } i \neq j$$

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 $||A - QHQ^*||_2$.

The functions are in Listings 1 and 2. The script to generate A is in Listing 3 and the solutions are also attached.

```
function [W,H] = hessenberg(A)

m = size(A,1);
H = A;

for k=1:m-2
    x  = H(k+1:m,k);
    vk = sign(x(1))*norm(x,2)*eye(m-k,1) + x;
    vk = vk/norm(vk,2);

    H(k+1:m,k:m) = H(k+1:m,k:m) - 2*vk*(vk'*H(k+1:m,k:m));
    H(1:m,k+1:m) = H(1:m,k+1:m) - 2*(H(1:m,k+1:m)*vk)*vk';
    W(k+1:m,k) = vk;
end

end
```

Listing 1: Hessenberg function

```
function Q = formQ(W)

[m,n] = size(W);
Q = eye(m);

if m >=n
    % Algorithm 10.3 - implicit calculation of Q using ei
    for i=1:m          % perform on each column of I
```

```
        for k=n:-1:1  % perform Q*ei to get back column of Q
            vk = W(k:m,k);
            Q(k:m,i) = Q(k:m,i) - 2*vk*(vk'*Q(k:m,i));
        end
    end
else
    disp("Algorithm works only for m >= n");
end

end
```

Listing 2: formQ function

```
clc
clear
%% generate A
m = 5;
A = zeros(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
%%
[W,H] = hessenberg(A);
Q = formQ(W);

disp('A formed through QHQ^* = ');
disp(Q*H*Q');

disp('Hessenberg matrix = ');
disp(H);

disp('W = ');
disp(W);

disp('Q = ');
disp(Q);

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

Listing 3: script for Q4

```
A formed through QHQ^* =
    9.0000    0.3333    0.2500    0.2000    0.1667
    0.3333    9.0000    0.2000    0.1667    0.1429
    0.2500    0.2000    9.0000    0.1429    0.1250
    0.2000    0.1667    0.1429    9.0000    0.1111
    0.1667    0.1429    0.1250    0.1111    9.0000

Hessenberg matrix =
    9.0000   -0.4913         0         0         0
   -0.4913    9.4289    0.1080   -0.0000    0.0000
         0    0.1080    8.8463    0.0400         0
         0         0    0.0400    8.8507   -0.0208
         0         0         0   -0.0208    8.8741

W =
         0         0         0
    0.9161         0         0
    0.2777   -0.8281         0
    0.2222   -0.3721   -0.7992
    0.1851   -0.4194   -0.6010

Q =
    1.0000         0         0         0         0
         0   -0.6785    0.6754   -0.2850   -0.0479
         0   -0.5088   -0.1666    0.7519    0.3847
         0   -0.4071   -0.4524    0.0300   -0.7929
         0   -0.3392   -0.5580   -0.5938    0.4701

|| Q^*Q - I || = 6.31e-16
|| A - QHQ^* || = 7.87e-15
>>
```