MATH-6600, CLA Problem Set No. 8, 11-19-15

Name: Michael Hennessey

1. Prove true or false. Throughout, $A \in \mathbb{C}^{m\times m}$ unless otherwise indicated.

   (a) If λ is an eigenvalue of A and $\mu \in \mathbb{C}$, then $\lambda - \mu$ is an eigenvalue of $A - \mu I$.

   True.

$$Ax = \lambda x$$
$$Ax - \mu x = \lambda x - \mu x$$
$$(A - \mu I)x = (\lambda - \mu)x$$

   Then $\lambda - \mu$ is an eigenvalue of $A - \mu I$.

   (b) If A is real and λ is an eigenvalue of A, then so is $-\lambda$.

   False.

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

   A is real with eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 1$ and characteristic polynomial

$$p(\lambda) = (3 - \lambda)(1 - \lambda).$$

   Clearly $\lambda = -3$ or $\lambda = -1$ do not satisfy the characteristic polynomial.

   (c) If A is real and λ is an eigenvalue of a, then so is $\bar{\lambda}$.

   True.

$$Ax = \lambda x$$
$$\bar{A}\bar{x} = \bar{\lambda}\bar{x} \implies A\bar{x} = \bar{\lambda}\bar{x}$$

   (d) If λ is an eigenvalue of A and A is nonsingular, then λ^{-1} is an eigenvalue of A^{-1}.
   True.
   If A is nonsingular, it is diagonalizable. Then

$$A^{-1} = (X^{-1}\Lambda X)^{-1} = X^{-1}\Lambda^{-1}X.$$

   As Λ is the diagonal matrix of eigenvalues, inverting A amounts to inverting the elements of Λ, the eigenvalues of A. Then since the above equation is an eigenvalue decomposition of A^{-1} and λ^{-1} is a diagonal entry in Λ^{-1}, λ^{-1} is an eigenvalue of A^{-1}.

(e) If all the eigenvalues of A are zero, then $A = 0$.

False.

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

$$p(\lambda) = \lambda^2 \implies \lambda_{1,2} = 0$$

(f) If A is hermitian and λ is an eigenvalue of A, then $|\lambda|$ is a singular value of A.

True.
$A^*A = V^*\Sigma^2 V$ has eigenvalues σ_i^2. $A^*A = A^2$ has eigenvalue decomposition $X^{-1}\Lambda^2 X = V^*\Sigma^2 V$. Then $\lambda^2 = \sigma_i^2$. Therefore, the singular values of A are

$$\sigma_i = |\lambda|.$$

(g) If A is diagonalizable and all its eigenvalues are equal, than A is diagonal.

True.

$$A = X^{-1} \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix} \left[\begin{array}{c|c|c} x_1 & \dots & x_m \end{array} \right]$$

$$= X^{-1} \left[\begin{array}{c|c|c} \lambda x_1 & \dots & \lambda x_m \end{array} \right] = \lambda X^{-1} X = \lambda I$$

2. Let

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

with $\epsilon = 10^{-10}$.

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

Solution:

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

$$D_1 : \{z| \ |z-3| \leq 3\}$$
$$D_2 : \{z| \ |z-2| \leq 3\}$$
$$D_3 : \{z| \ |z-1| \leq \epsilon\}$$

Where D_i corresponds to the ith row of $A+B$. Therefore $\lambda_1 \in D_1$, $\lambda_2 \in D_2$, and $\lambda_3 \in D_3$ by Gershgorin's Circle theorem.

(b) Improve the estimates from (a) by judicious choices of diagonal similarity transformations of the form

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d^2 \end{bmatrix}.$$

Solution:

We begin by doing a diagonal similarity transform of $A+B$ to get

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

Thus we have

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

This gives the three Gershgorin disks:

$$D_1 : \{z| \ |z-3| \leq 3d\}$$
$$D_2 : \{z| \ |z-2| \leq 3d\}$$
$$D_3 : \{z| \ |z-1| \leq \frac{\epsilon}{d^2}\}$$

where D_i corresponds to the ith row of $D^{-1}(A+B)D$. Thus $\lambda_1 \in D_1$, $\lambda_2 \in D_2$, and $\lambda_3 \in D_3$ by Gershgorin's Circle theorem. To "zoom" in on each eigenvalue we can choose appropriate values of d such that the radius is as small as possible around each eigenvalue and the disk in question has no intersection with the other two disks.

We begin with D_1. Here we want to zoom in on $|z-3|$ and stop D_2 from intersecting D_1. We must also be careful to not make d so large that D_3 begins intersecting D_1. To choose d such that these conditions are satisfied we expand the absolute value inequalities that define the disks:

$$D_1 : 3-3d \leq z \leq 3+3d$$
$$D_2 : 2-3d \leq z \leq 2+3d$$
$$D_3 : 1-\frac{\epsilon}{d^2} \leq z \leq 1+\frac{\epsilon}{d^2}$$

Since we want d to be as small as possible, and since D_2 will not be an issue, the only constraint we must keep in mind is:

$$1+\frac{\epsilon}{d^2} < 3+3d$$

To simplify the analysis we let $d = \frac{\sqrt{\epsilon}}{a}$. We then have

$$1+a^2 < 3+3\frac{\sqrt{\epsilon}}{a} \implies a^2 \approx 2$$

For clarity, we choose $a = \sqrt{1.8} = \frac{3}{\sqrt{5}}$, which makes $d = \frac{\sqrt{5\epsilon}}{3}$. Then we know λ_1 is in the Gershgorin disk

$$D_1' : \{z : |z-3| \leq \sqrt{5\epsilon}\}.$$

Moving on to D_2 we use the same inequalities as above, except our constraints are that we need

$$2+3d < 3-3d \text{ and } 1+\frac{\epsilon}{d^2} < 2-3d.$$

To satisfy the first constraint all we need is $d < \frac{1}{6}$. To satisfy the second, however, we can make a similar substitution as before, with $d = \frac{\sqrt{\epsilon}}{b}$. The second constraint then becomes:

$$1+b^2 < 2-\frac{3}{b}\sqrt{\epsilon} \implies b^2 < 1.$$

Again, for clarity, we choose $b = \sqrt{.9} = \dfrac{3}{\sqrt{10}}$, which makes $d = \dfrac{\sqrt{10\epsilon}}{3}$. Then we know that λ_2 lies in the Gershgorin disk

$$D_2' : \{z : |z - 2| \leq \sqrt{10\epsilon}\}.$$

Lastly, we look at D_3. Here we are only concerned with removing the overlap from D_2 and D_1. Using the same equalities as before, we note that if $D_2 \cap D_3 = \emptyset$ then $D_1 \cap D_3 = \emptyset$. Therefore, we are only worried about the constraint

$$1 + \frac{\epsilon}{d^2} < 2 - 3d.$$

Except here we are only interested in making d small enough so that D_2 does not intersect D_3. We can easily see that we want $d < \dfrac{1}{3}$. For a very good estimate, we can let $d = .3 = \dfrac{3}{10}$. Then we know λ_3 is in the Gershgorin disk

$$D_3' = \{z : |z - 1| \leq \frac{100}{9}\epsilon\}.$$

3. (a) Let $A \in \mathbb{C}^{m\times m}$ be tridiagonal and hermitian, with all its sub- and superdiagonal entries nonzero. Prove that the eigenvalues of A are distinct.

   Proof: Let λ be an eigenvalue of A. Then $B = A - \lambda I$ is singular, and therefore has rank$(A - \lambda I) \leq m - 1$. We then note the $m - 1 \times m$ submatrix $B_{2:m,1:m}$ is upper triangular whose diagonal entries are non-zero by our assumptions on A. Hence $B_{2:m,1:m}$ has $m - 1$ linearly independent columns and is therefore of full rank. Therefore, $A - \lambda I$ must have rank $m - 1$. Therefore, the null space of B is spanned by one unique eigenvector of A corresponding to λ. Since A is Hermitian, it is nonsingular and therefore has m linearly independent eigenvectors, by the finite-dimensional spectral theorem. Thus all λ must be distinct.

   (b) On the other hand, let A be upper-Hessenberg, with all its subdiagonal entries nonzero. Give an example that shows that the eigenvalues of A are not necessarily distinct.

   Solution:

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

   has eigenvalues all 0.

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,

$$A = QHQ^*.$$

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

$$a_{ij} = 9, \quad \text{for } i = j,$$

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

and $m = 5$. Output $A, H, W, Q, \|Q^*Q - I\|_2$, and $\|A - QHQ^*\|_2$.

Code:

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

[~,m]=size(A);
for k=1:m-2
    I=(k+1):m;
    x=A(k+1:m,k);
    e=zeros(m-k,1);
    e(1)=1;
    if x(1)==0
        V(I,k)=norm(x,2)*e+x;
    else
        V(I,k)=sign(x(1))*norm(x,2)*e+x;
    end
    V(I,k)=V(I,k)/norm(V(I,k),2);
    A(k+1:m,k:m)=A(k+1:m,k:m)-2*V(I,k)*(V(I,k)'*A(k+1:m,k:m));
    A(1:m,k+1:m)=A(1:m,k+1:m)-2*(A(1:m,k+1:m)*V(I,k))*V(I,k)';
end
W=V;
H=A(1:m,1:m);


function Q=formQh(W)

[m,~]=size(W);

for i=1:m
    x=zeros(m,1);
    x(i)=1;
    for k=m-2:-1:1
        x(k:m)=x(k:m)-2*W(k:m,k)*(W(k:m,k)'*x(k:m));

    end
    Q(1:m,i)=x;
end
```

```
%hessenberg matrix precursor
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
```

Results:

```
>> hessenbergprecursor
>> [W,H]=hessenberg(A);
>> Q=formQh(W);
>> A

A =

    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

>> H

H =

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

>> W

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

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

>> norm(Q'*Q-eye(5),2)

ans =

   9.7450e-16

>> norm(A-Q*H*Q',2)

ans =

   1.8432e-14
```