HW7 Tianbo Zhang 1938501

24.1. 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. (This comes from the German "Eigenwert." The corresponding abbreviation for eigenvector is "ev," from "Eigenvektor.")

(a) If λ is an ew of A and $\mu \in \mathbb{C}$, then $\lambda - \mu$ is an ew of $A - \mu I$.
(b) If A is real and λ is an ew of A, then so is $-\lambda$.
(c) If A is real and λ is an ew of A, then so is $\bar{\lambda}$.
(d) If λ is an ew of A and A is nonsingular, then λ^{-1} is an ew of A^{-1}.
(e) If all the ew's of A are zero, then $A = 0$.
(f) If A is hermitian and λ is an ew of A, then $|\lambda|$ is a singular value of A.
(g) If A is diagonalizable and all its ew's are equal, then A is diagonal.

a) True

Suppose λ is an ew of A then we have:

$Av = \lambda v$

Then note:

$(A-\mu I)v = Av - \mu v = \lambda v - \mu v = (\lambda - \mu)v$

Hence $\lambda - \mu$ is a ew of $A - \mu I$. $\square$

b) True

Let λ be ew of A.

Then we have:

$Av = \lambda v \rightarrow A(-v) = \lambda(-v)$

$\rightarrow -Av = -\lambda v$ $\square$

c) True

Because characteristic polynomial of A will have complex roots in conjugate pairs.

d) True

Let λ be ew of A.

Then we have:

$Av = \lambda v \rightarrow A^{-1}Av = A^{-1}\lambda v$

$\Rightarrow v = \lambda A^{-1} v$

$\Rightarrow \lambda^{-1} v = A^{-1} v$ $\square$

e) True

Note by definition eigenvectors are non zero

$\Rightarrow Av = \lambda v = 0 \cdot v$

$\Rightarrow A = 0$

f) True

Let λ be ew of A.

Then we have:

$Av = \lambda v \rightarrow v^* A v = \lambda v^* v$

$\rightarrow \lambda$ is real

Note singular value of A is sqrt(ew of A^2)

$= |\lambda|$

g) True

Let $A = PDP^{-1}$

Since ews are all equal we have:

$A = P(\lambda I)P^{-1} = \lambda PP^{-1} = \lambda I$

Thus A is diagonal.

24.2. Here is *Gerschgorin's theorem*, which holds for any $m \times m$ matrix A, symmetric or nonsymmetric. *Every eigenvalue of A lies in at least one of the m circular disks in the complex plane with centers a_{ii} and radii $\sum_{j \neq i} |a_{ij}|$. Moreover, if n of these disks form a connected domain that is disjoint from the other $m - n$ disks, then there are precisely n eigenvalues of A within this domain.*

(a) Prove the first part of Gerschgorin's theorem. (Hint: Let λ be any eigenvalue of A, and x a corresponding eigenvector with largest entry 1.)

(b) Prove the second part. (Hint: Deform A to a diagonal matrix and use the fact that the eigenvalues of a matrix are continuous functions of its entries.)

(c) Give estimates based on Gerschgorin's theorem for the eigenvalues of

$$A = \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix}, \qquad |\epsilon| < 1.$$

(d) Find a way to establish the tighter bound $|\lambda_3 - 1| \leq \epsilon^2$ on the smallest eigenvalue of A. (Hint: Consider diagonal similarity transformations.)

c) i) The first row: $a_{11} = 8$ is the center, radius $= 1$

ii) The second row: $a_{22} = 4$ is the center, radius $= 1 + |\varepsilon|$

iii) The third row: $a_{33} = 1$ is the center, radius $= |\varepsilon|$

λ_1 in circle 1: $|\lambda_1 - 8| \leq 1$

λ_2 in circle 2: $|\lambda_2 - 4| \leq 1 + |\varepsilon|$

λ_3 in circle 3: $|\lambda_3 - 1| \leq |\varepsilon|$

d) First note by Theorem 24.3 we kno :

If $B = X^{-1}AX$ and X is non singular, then B has the same eigenvalues as A.

Let $X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix} \rightarrow X^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{x} & 0 \\ 0 & 0 & \frac{1}{x^2} \end{bmatrix}$

Then we have:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{x} & 0 \\ 0 & 0 & \frac{1}{x^2} \end{bmatrix} \begin{bmatrix} 8 & 1 & 0 \\ 1 & 4 & \varepsilon \\ 0 & \varepsilon & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 1 & 0 \\ \frac{1}{x} & \frac{4}{x} & \frac{\varepsilon}{x} \\ 0 & \frac{\varepsilon}{x^2} & \frac{1}{x^2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix} = \begin{bmatrix} 8 & x & 0 \\ \frac{1}{x} & 4 & \varepsilon x \\ 0 & \frac{\varepsilon}{x} & 1 \end{bmatrix}$$

Note: ① For circle 1, radius $= |x|$

② For circle 2, radius $= |\frac{1}{x}| + |\varepsilon x|$

③ For circle 3, radius $= |\frac{\varepsilon}{x}|$

Set $|x| = |\varepsilon|^{-1}$

Then for the third circle we have:

$|\frac{\varepsilon}{x}| = \frac{|\varepsilon|}{|x|} = |\varepsilon|^2 = \varepsilon^2$

$\Rightarrow |\lambda_3 - 1| \leq \varepsilon^2$

28.1. What happens if you apply the unshifted QR algorithm to an orthogonal matrix? Figure out the answer, and then explain how it relates to Theorem 28.4.

Theorem 28.4. *Let the pure QR algorithm (Algorithm 28.1) be applied to a real symmetric matrix A whose eigenvalues satisfy $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_m|$ and whose corresponding eigenvector matrix Q has all nonsingular leading principal submatrices. Then as $k \to \infty$, $A^{(k)}$ converges linearly with constant $\max_j |\lambda_{j+1}|/|\lambda_j|$ to $\operatorname{diag}(\lambda_1, \ldots, \lambda_m)$, and $\underline{Q}^{(k)}$ (with the signs of its columns adjusted as necessary) converges at the same rate to Q.*

First note the Unshifted QR algorithm:

$$A^{(0)} = A$$

$$A^{(k-1)} = Q^{(k)} R^{(k)}$$

$$A^{(k)} = R^{(k)} Q^{(k)}$$

$$\underline{Q}^{(k)} = Q^{(1)} Q^{(2)} \cdots Q^{(k)}$$

Then note if A is an orthogonal matrix,

For $A = QR$, R is the identity

Then if multiply the factors in reverse order,

we'll get Q.

Then note that this will make no progress.

Then note all e.w will have magnitude 1.

Thus do not satisfy the hypothesis in 28.4.

4. (*from A. Greenbaum*) For $A \in \mathbb{R}^{m \times m}$, define the matrix exponential:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Suppose A is diagonalizable: $A = X \Lambda X^{-1}$. Show that $e^{tA} = X e^{t\Lambda} X^{-1}$, where

$$e^{t\Lambda} = \begin{bmatrix} e^{t\lambda_1} & & \\ & \ddots & \\ & & e^{t\lambda_m} \end{bmatrix}.$$

First note by properties of tylor series expansion we have:

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!}$$

Then note for $e^{t\Lambda}$ we have:

$$e^{t\Lambda} = \sum_{k=0}^{\infty} \frac{(t\Lambda)^k}{k!}$$

Then note since t is scalar we have:

$$(tA)^k = t^k A^k = t^k \overbrace{A \cdot A \cdot A \cdots A}^{k}$$

Then since we know A is diagonalizable and $A = X \Lambda X^{-1}$

We can plug $A = X \Lambda X^{-1}$ back in $(tA)^k$ we have:

$$(tA)^k = t^k \underbrace{(X\Lambda X^{-1})(X\Lambda X^{-1}) \cdots (X\Lambda X^{-1})}_{k}$$

$$= t^k X \Lambda^k X^{-1}$$

Then plug this back in e^{tA} we have:

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k X \Lambda^k X^{-1}}{k!} = X\left(\sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!}\right) X^{-1} = X e^{t\Lambda} X^{-1}$$

$\square$

5. (*From A. Greenbaum*) Consider the matrix

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

Taking A to be a 10 by 10 matrix, try the following:

(a) What information does Gerschgorin's theorem give you about this matrix?

(b) Implement the power method to compute an approximation to the eigenvalue of largest absolute value and its corresponding eigenvector. Include in your writeup a printout your of code together with the eigenvalue/eigenvector pair that you computed. Once you have a good approximate eigenvalue, look at the error in previous approximations. Create a plot showing the error at each iteration, and comment on the rate of convergence. Does it match the theory?

(c) Using $s = 1$ as a shift in inverse iteration, find the eigenvalue that is closest to 1 and its corresponding eigenvector. Include in your writeup a printout of your code together with the eigenvalue/eigenvector pair that you computed. Comment on the rate of convergence of inverse iteration with $s = 1$ as a shift.

a) By Gerschgorin's theorem we know:

For λ_1, λ_9 we have:

$|\lambda_1 - 2| \leq 1$, $|\lambda_9 - 2| \leq 1$

For each eigenvalue $\lambda_2, \cdots, \lambda_9$ of A we have:

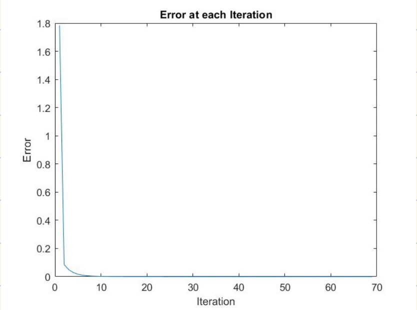
$|\lambda_i - 2| \leq 2$

Thus all the eigenvalues lies in $[0, 4]$

b)

```
% AMATH 584 HW7 Problem 4 Part b
% Tianbo Zhang 1938501
% Implement the Power Iteration to compute an approximation to the
% eigenvalue of largest absolute value and its corresponding eigenvector

% Matrix setup
n = 10;
A = 2 * eye(n,n);
for i = 1:n-1
    A(i+1,i) = -1;
    A(i,i+1) = -1;
end

% Precondition setup
w = rand(n,1);
v = w / norm(w);
error = 1.e-6;
cur_error = 10;
error_list = [];

% Start the Power Iteration
while cur_error > error
    w = A*v;
    v = w / norm(w);
    ew = v' * A * v;
    cur_error = norm(A*v - ew * v);
    error_list(end+1) = cur_error;
end

plot([1:length(error_list)], error_list);
title("Error at each Iteration");
xlabel("Iteration");
ylabel("Error");
```

```
ew =

    3.9190


v =

    0.1216
   -0.2331
    0.3250
   -0.3899
    0.4228
   -0.4213
    0.3858
   -0.3195
    0.2280
   -0.1187
```



The approximate lambda converges at a relatively fast.

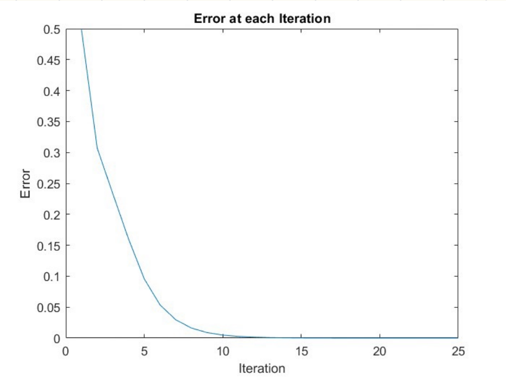
The ratio of errors is multiplied by about 0.88 at each iteration.

It does match the theory.

c)

```
% AMATH 584 HW7 Problem 4 Part b
% Tianbo Zhang 1938501
% Implement the Inverse Iteration with s = 1 shift to compute an approximation
% to the eigenvalue that is closest to 1 and its corresponding eigenvector

% Matrix setup
n = 10;
A = 2 * eye(n,n);
for i = 1:n-1
    A(i+1,i) = -1;
    A(i,i+1) = -1;
end

% Precondition setup
s = 1;
w = rand(n,1);
v = w / norm(w);
error = 1.e-6;
shift_A = A - s*eye(n);
cur_error = 10;
error_list = [];

% Start the iteration
while cur_error > error
    w = shift_A \ v;
    v = w / norm(w);
    ew = v' * A * v;
    cur_error = norm(A * v - ew * v);
    error_list(end+1) = cur_error;
end

plot([1:length(error_list)], error_list);
title("Error at each Iteration");
xlabel("Iteration");
ylabel("Error");
```

```
ew =

    1.1692


v =

    0.3879
    0.3223
   -0.1201
   -0.4221
   -0.2305
    0.2305
    0.4221
    0.1201
   -0.3223
   -0.3879
```



This method converge faster than power method.

The error decreases rapidly than hit the balance state.