Math 584 Homework 6

Due Soon

By Marvyn Bailly

Problem 1 *Exercise 23.1*

Solution.
Let A be a nonsingular square matrix and let $A = QR$ and $A^*A = U^*U$ be QR and Cholesky factorizations. Since A is nonsingular, it has full rank and thus $A = QR$ is unique where $r_{jj} > 0$ and Q is unitary. Observe that

$$A^*A = (QR)^*(QR) = R^*Q^*QR = R^*IR = R^*R.$$

Next let's show that A has a unique Cholesky factorization which is true when A^*A is Hermitian positive definite. We know that A^*A is Hermitian and since A is nonsingular, $Ax \neq 0$ for all $x \neq 0$. Then

$$x^*A^*Ax = (Ax)^*(Ax) = \|Ax\|_2^2 > 0.$$

Thus A^*A is positive definite. Therefore A^*A is Hermitian positive definite. Thus we know that A has a unique Cholesky factorization which implies that $A^*A = R^*R = U^*U$ and thus $R = U$. $\square$

Problem 2 *Exercise 24.1*

Solution.
Consider $A \in \mathbb{C}^{m\times m}$.

(a) This is true. Let λ be an eigenvalue of A and let $\mu \in \mathbb{C}$. Then

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

Thus we have that $\lambda - \mu$ is an eigenvalue of $A - \mu I$.

(b) This is false. Consider the following counterexample. Let A be a diagonal matrix where the absolute value of the elements along the main diagonal are different for each element. Then we know that all the eigenvalues of A are the elements along the main diagonal of A but their negative is not a eigenvalue.

(c) This is true. Let A be real and λ be an eigenvalue of A. Then we know that

$$Ax = \lambda x \implies \overline{Ax} = \overline{\lambda x}.$$

Since A is real

$$\overline{Ax} = \overline{\lambda x} \implies A\overline{x} = \overline{\lambda}\overline{x}.$$

Thus we have that $\overline{\lambda}$ is an eigenvalue of A.

(d) This is true. Let λ be an eigenvalue of A and suppose A is nonsingular. Then we have

$$Ax = \lambda x \implies x = A^{-1}\lambda x = \lambda A^{-1}x \implies \lambda^{-1}x = A^{-1}x.$$

Thus λ^{-1} is an eigenvalue of A^{-1}.

(e) This is false. Consider the following counterexample. Let A be a strictly upper triangular matrix. Then we know that the elements along the main diagonal are the eigenvalues of A and are thus all zero. But A has at least one nonzero element above the main diagonal and thus $A \neq 0$.

(f) This is true. Let A be Hermitian and let λ be an eigenvalue of A. As A is hermitian, we know that A is a normal matrix. Thus we have that

$$A = Q\Lambda Q^* = Q|\Lambda|\mathrm{sgn}(\Lambda)Q^*,$$

is the singular value decomposition of A with $\Sigma = |\Lambda|$. Thus $|\lambda|$ is an singular value of A.

(g) This is true. Let A be such that it is diagonalizable and all its eigenvalues are equal. Since A is diagonalizable $X^{-1}AX = \Lambda$ where Λ is a diagonal matrix with A's eigenvalues along the main diagonal. Since the eigenvalues of A are all equal, $\Lambda = \lambda I$ for some λ. Then

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

Thus A is a diagonal matrix.

□

Problem 3 *Exercise 24.2 (c) and (d)*

Solution.

(c) Consider the symmetric matrix

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

From Gerschgorin's theorem, every eigenvalue of A lies in at least one eigenvalue in each of the following disks

$$D_1 = \{z \in \mathbb{C} : |z - 8| \leq 1\}, D_2 = \{z \in \mathbb{C} : |z - 4| \leq 1 + \epsilon\}, D_3 = \{z \in \mathbb{C} : |z - 1| \leq \epsilon\}.$$

From homework 1 we know that symmetric matrices have real eigenvalues and thus we have that

$$\lambda_1 \in [7, 9], \quad \lambda_2 \in [3 - \epsilon, 3 + \epsilon], \quad \text{and} \quad \lambda_3 \in [1 - \epsilon, 1 + \epsilon].$$

(d) Next we wish to establish the tighter bound $|\lambda_3 - 1| \leq \epsilon^2$ on the smallest eigenvalue of A. In the case that $\epsilon = 0$, the disk D_3 becomes

$$D_3 = \{z \in \mathbb{C} : |z - 1| \leq 0\} \implies \lambda_3 = 1$$

and we are done since there is no tighter bound. Assume that $\epsilon \neq= 0$. Since we showed in the last part that A has distinct eigenvalues, we know that A is diagonalizable. Let $X^{-1}AX = B$ be the diagonalization of the A. From theorem 24.3, B has the same real eigenvalues as A. We want the matrix B to be in the form

$$B = \begin{pmatrix} 8 & \star & \star \\ \star & 4 & \star \\ 0 & \epsilon^2 & 1 \end{pmatrix},$$

as this will bound $|\lambda_3 - 1| \leq \epsilon^2$ by Gerschgorin's theorem. We can achieve this form of B by considering

$$\begin{aligned} \begin{pmatrix} 8 & \star & \star \\ \star & 4 & \star \\ 0 & \epsilon^2 & 1 \end{pmatrix} &= X^{-1} \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} X \\ &= \begin{pmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1/c \end{pmatrix} \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \\ &= \begin{pmatrix} 8 & b/a & 0 \\ a/b & 4 & \epsilon c/b \\ 0 & \epsilon b/c & 1 \end{pmatrix}. \end{aligned}$$

Since we want $\epsilon^2 = \epsilon b/c$, let $a = b = \epsilon$ and $c = 1$. Thus we have that A and B are similar with

$$X = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and thus

$$B = \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & \epsilon^2 & 1 \end{pmatrix} = \begin{pmatrix} 1/\epsilon & 0 & 0 \\ 0 & 1/\epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 8 & 1 & 0 \\ 1 & 4 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} = X^{-1}AX.$$

By Gerschgorin's Theorem we have that

$$D_1 = \{z \in \mathbb{C} : |z - 8| \leq 1\}, D_2 = \{z \in \mathbb{C} : |z - 4| \leq 2\}, D_3 = \{z \in \mathbb{C} : |z - 1| \leq \epsilon^2\}.$$

Since D_3 is disjoint from D_1 and D_2, D_3 contains one eigenvalue. We have that D_1 and D_2 correspond to larger magnitude eigenvalues of A and B than D_3. And since A and B must have real eigenvalues, we have that

$$|\lambda_3 - 1| \leq \epsilon^2.$$

□

Problem 4 *For $A \in \mathbb{R}^{m\times m}$, define the matrix exponential:*

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

(a) Show that the matrix-valued function $Y(t) := e^{tA}$ satisfies $Y'(t) = AY(t), t \geq 0, Y(0) = I$

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

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

Solution.

(a) Let $A \in \mathbb{R}^{m\times m}$. Define $Y(t) = e^{tA}$. Using the definition of matrix exponential,

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

Then we see that

$$Y'(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}A^k}{(k-1)!} = A\sum_{k=1}^{\infty} \frac{t^{k-1}A^{k-1}}{(k-1)!} = A\sum_{j=0}^{\infty} \frac{t^j A^j}{j!} = AY(t) \quad \forall t \geq 0.$$

And we have that

$$Y(0) = 0^0 \cdot I + 0 \cdot A + 0 \cdot \frac{A^2}{2!} + \ldots = I + 0 = I.$$

(b) Now assume that A is diagonalizable of the form $A = X\Lambda X^{-1}$. Note that

$$A^n = \left(X\Lambda X^{-1}\right)^n = \left(X\Lambda X^{-1}\right)\cdot\left(X\Lambda X^{-1}\right)\cdots\left(X\Lambda X^{-1}\right) = X(\Lambda I \cdot \Lambda I \cdots I\Lambda)X^{-1} = X\Lambda^n X^{-1}.$$

Using this, observe that

$$\begin{aligned}
e^{tA} &= I + tA + \frac{t^2A^2}{2!} + \frac{t^3A^3}{3!} + \ldots \\
&= I + tX\Lambda X^{-1} + \frac{t^2\left(X\Lambda X^{-1}\right)^2}{2!} + \frac{t^3\left(X\Lambda X^{-1}\right)^3}{3!} + \ldots \\
&= I + tX\Lambda X^{-1} + \frac{t^2X\Lambda^2X^{-1}}{2!} + \frac{t^3X\Lambda^3X^{-1}}{3!} + \ldots \\
&= \sum_{k=0}^{\infty} \frac{t^kX\Lambda^kX^{-1}}{k!}
\end{aligned}$$

$$
\begin{aligned}
&= X \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} X^{-1} \\
&= X \begin{pmatrix} \sum_{k=0}^{\infty} \frac{t^k \lambda_1^k}{k!} & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty} \frac{t^k \lambda_m^k}{k!} \end{pmatrix} X^{-1} \\
&= X \begin{pmatrix} e^{t\lambda_1} & & \\ & \ddots & \\ & & e^{t\lambda_m} \end{pmatrix} X^{-1} \\
&= X e^{t\Lambda} X^{-1}
\end{aligned}
$$

$\square$

Problem 5 *By hand, find a Householder reflector Q and an upper Hessenberg matrix H such that $Q^*AQ = H$, where*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Solution.
Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Following the algorithm outlined in lecture 26, let

$$x = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

Then,

$$\begin{aligned} v &= \mathrm{sign}(x_1)\|x\|_2 e_1 + x \\ &= \|x\|_2 + x \\ &= \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{2}+1 \\ 1 \end{pmatrix}. \end{aligned}$$

Then we can compute F to be

$$\begin{aligned} F &= I - \frac{2vv^*}{v^*v} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{(\sqrt{2}+1)(\sqrt{2}+1)+1} \begin{pmatrix} \sqrt{2}+1 \\ 1 \end{pmatrix} \begin{pmatrix} \sqrt{2}+1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2+\sqrt{2}} \begin{pmatrix} 3+2\sqrt{2} & \sqrt{2}+1 \\ \sqrt{2}+1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \sqrt{2}+1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}. \end{aligned}$$

Thus we have found the Householder reflector Q to be,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Thus we have that the upper Hessenberg matrix H is given by

$$
\begin{aligned}
Q^*AQ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}^* \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & -5/\sqrt{2} & 1/\sqrt{2} \\ 1 & -3\sqrt{2} & 1/\sqrt{2} \\ 1 & -\sqrt{2} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -5\sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2} & 5/2 & -1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.
\end{aligned}
$$

□

Problem 6 *Consider the matrix*

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

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. Turn in a listing of your code together with the eigenvalue/eigenvector pair that you computed. Once you have a good approximate eigenvalue, look at the error in previous approximations and comment on the rate of convergence of the power method.

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

Solution.

(a) From Gerschgorin's Theorem, we see that all the eigenvalue of A are within a disk of radius 2 centered at 2 (note that the $a_{1,1}$ and $a_{10,10}$ have radius 1 but are within the radius 2 neighborhoods). Since A is symmetric the eigenvalues of A must be real. Thus the eigenvalues of A are within $[0, 4]$.

(b) I created the following code in MatLab that runs the power method on A

```
%form A
A = 2*eye(10) - diag(ones(9,1),1) - diag(ones(9,1),-1);
%set v to be some vector with norm 1
v = eye(10,1);
%define lambdas
lambdaOld = -Inf;
lambdaNew = Inf;
%define v
vOld = v;

while lambdaOld~=lambdaNew
    w = A*vOld;
    v = w/norm(w);
    lambdaOld = lambdaNew;
```

```
        lambdaNew = transpose(v)*A*vOld;
        vOld = v;
    end
    disp(lambdaOld)
    disp(v)
```

which outputs that the eigenvalue is

```
    3.918985947228988
```

and the corresponding eigenvector is

```
    0.120131202514620
   -0.230530080785626
    0.322252768349909
   -0.387868437271821
    0.422061300037665
   -0.422061261854957
    0.387868334846436
   -0.322252634201185
    0.230529957504833
   -0.120131129242539
```

To study the convergence of the power method, I added the following lines of code to the while loop to check the error every steps.

```
if mod(counter,10) == 0
    if oldError == -1
        newError = abs(lambdaExact - lambdaNew);
        disp([newError,newError/oldError])
        oldError = newError;
    else
        oldError = abs(lambdaExact - lambdaNew);
        disp(abs(lambdaExact - lambdaNew))
    end
end
```

Studying the output of the new code shows that every ten steps our error reduces by about a factor of 0.3. To compute the rate of convergence we can take the tenth root of this (since it is every ten steps) to get that the rate of convergence is around 0.88. There seems to be some small errors due to rounding and not having the exact eigenvalue.

(c) I created the following code following the outlined script in chapter 23.

```
A = 2*eye(10) - diag(ones(9,1),1) - diag(ones(9,1),-1);

%set v to be some vector with norm 1
v = eye(10,1);

% Initialize lambda variables
lambdaOld = Inf;
lambdaNew = 0;

%set v
vOld = v;

%counter
counter = 0;

while lambdaOld~=lambdaNew
    w = (A - eye(10))\vOld;
    v = w/norm(w);
    lambdaOld = lambdaNew;
    lambdaNew = transpose(v)*A*v;
    vOld = v;
    counter = counter + 1;
end

disp(lambdaOld)
disp(v)
disp(counter)
```

This outputs

```
 1.169169973996227

 0.387868389593145
 0.322252705904122
-0.120131163350454
-0.422061282263745
```

```
-0.230530023398824
 0.230530014891641
 0.422061279628888
 0.120131168406708
-0.322252696646980
-0.387868382525122

30
```

which shows that the eigenvalue is 1.169169973996227. To studying the convergence, we can add the same code as before

```
if mod(counter,5) == 0
    if oldError == -1
        newError = abs(lambdaExact - lambdaNew);
        disp([newError,newError/oldError])
        oldError = newError;
    else
        oldError = abs(lambdaExact - lambdaNew);
        disp(abs(lambdaExact - lambdaNew))
    end
end
```

which shows every five steps the error reduces by a factor of around 0.002. Taking the fifth root of this we get that the rate of convergence is around 0.29. We once again have some small errors going on so we do not see this exact behavior in the code but an approximation.

□