MA/CSC 580, Homework #5
Due Tuesday, November 16, 2018

Lindsay Eddy

1. Let $A \in R^{n,n}$ be a square matrix.

   (a) Let $(\lambda, \mathbf{x})$ be an eigenpair of A. Show that $(\lambda^k, \mathbf{x})$ is an eigenpair of A^k for any integer k. Can k be negative or zero?

   (b) If $A = A^H$ (the conjugate transpose of A and so on) be a symmetric matrix, use the conclusion above to show the following

   $$\|A\|_2 = \max_{1\leq i\leq n} |\lambda_i(A)|, \qquad \|A^{-1}\|_2 = \frac{1}{\min_{1\leq i\leq n} |\lambda_i(A)|}, \qquad cond_2(A) = \frac{\max_{1\leq i\leq n} |\lambda_i(A)|}{\min_{1\leq i\leq n} |\lambda_i(A)|}.$$

   (c) Also show that all eigenvalues of A are real and eigenvectors corresponding to different eigenvalues are orthogonal. **Hint:** Consider $Ax = \lambda_i x$, and $y^H A^H = \bar{\lambda}_i y$, $y^h Ax$ and $x^h Ay$.

   (d) If A is a lower/upper triangular matrix, show that $\lambda_i(A) = a_{ii}$, that is, the eigenvalues of A are the diagonal entries. Is such a matrix always diagonalizable?

   (a) Proof for positive k:

   $$A^k x = A^{k-1}(Ax) = \lambda A^{k-1}x = \lambda^2 A^{k-2}x = \cdots = \lambda^k x$$

   Proof for $k = 0$:

   $$A^0 x = Ix = x = 1x = \lambda^0 x$$

   Proof for negative k:
   Note that $k = -j$ for some $j > 0$. Then

   $$A^k x = A^{-j}x = (A^{-1})^j x = (A^{-1})^j \lambda^{-j}(\lambda^j x) = (A^{-1})^j \lambda^{-j} A^j x = \lambda^{-j} x = \lambda^k x$$

   (b)

   $$\|A\|_2 = \max_i \sqrt{\lambda_i(A^H A)} = \max_i \sqrt{\lambda_i(A^2)}$$

   By part (a), for every eigenvalue λ_i of A, λ_i^2 is an eigenvalue of A^2. Also, for every eigenvalue λ_j^2 of A^2, λ_j is an eigenvalue of A, ($A^2 x = \lambda_j^2 x \Rightarrow A\lambda_j x = \lambda_j^2 x \Rightarrow Ax = \lambda_j x$). Since $f(x) = x^2$ is a monotonically increasing function, $\max_i \sqrt{\lambda_i(A^2)} = \max_i \sqrt{\lambda_i(A)^2}$. Thus

   $$\|A\|_2 = \max_i \sqrt{\lambda_i(A)^2} = \max_i |\lambda_i(A)|.$$

   $$\|A^{-1}\|_2 = \max_i \sqrt{\lambda_i((A^{-1})^H A^{-1})} = \max_i \sqrt{\lambda_i((A^{-1})^2)} = \max_i \sqrt{\lambda_i(A^{-2})}$$

   As we did previously, we can use part (a) and the fact that $f(x) = x^{-2}$ is a monotonically decreasing function to show that $\max_i \sqrt{\lambda_i(A^{-2})} = 1/\big(\min_i \sqrt{\lambda_i(A)^2}\big)$. Thus

   $$\|A^{-1}\|_2 = \frac{1}{\min_i \sqrt{\lambda_i(A)^2}} = \frac{1}{\min_i |\lambda_i(A)|}.$$

   $$cond_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}.$$

(c) Let (λ_i, x) be an eigenvalue of A where $\|x\|_2 = 1$. Then

$$\lambda_i = \lambda_i x^H x = x^H(\lambda_i x) = x^H(Ax) = (x^H A^H)x = \overline{\lambda}_i x^H x = \overline{\lambda_i}.$$

That is, all eigenvalues of A are real.

To show that eigenvectors corresponding to different eigenvalues are orthogonal, we will consider the eigenpairs (x, λ_I) and (y, λ_j) where $\lambda_i \neq \lambda_j$.

First, note that

$$(y^H Ax)^H = x^H A^H y = x^H Ay = x^H \lambda_j y = \lambda_j x^H y$$

Thus, $(\lambda_j x^H y)^H = \left((y^H Ax)^H\right)^H = y^H Ax$. Also, note that $\lambda_k = \overline{\lambda}_k$ for all k, since all eigenvalues of A are real. So,

$$\lambda_i y^H x = y^H \lambda_i x = y^H Ax = (\lambda_j x^H y)^H = y^H x \overline{\lambda}_j = \lambda_j y^H x$$

We have shown that $\lambda_i y^H x = \lambda_j y^H x$. Since $\lambda_i \neq \lambda_j$, $y^H x = 0$. That is, x and y are orthogonal.

(d) Prove that the eigenvalues of an upper/lower triangular matrix are the diagonal entries:

The eigenvalues of a matrix are the solutions to the characteristic polynomial $\det(A - \lambda I)$. If A is a triangular matrix, the zeros cause the characteristic polynomial to be

$$\prod_{i=1}^{n} (\lambda - a_{ii}) = 0.$$

The diagonal elements, a_{ii}, are the soultions to the characteristic polynomial and are thus the eigenvalues.

Triangular matrices are not always diagonalizable. A matrix is diagonalizable if and only if it has n linearly independent eigenvectors. Consider the upper triangular matrix

$$U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix}.$$

Using Matlab's `eig` function, I found that the eigenvectors of U are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$$

and of course, scalar multiples of these eigenvectors. Clearly, v_1 and v_2 are not linearly independent, so U is not diagonalizable. Since U is upper triangular, this is a counterexample.

2. Let $A \in R^{n,n}$ be the following matrix

$$A = \begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_n \end{pmatrix},$$

Assume that $\lambda_i \neq \lambda_j$ if $i \neq j$. Find a similarity transform so that the Gershgorin circles of the transformed matrix are separated. **Hint:** See the proof of the convergence of stationery iterative methods.

Consider

$$\epsilon = \frac{1}{4} \min_{i,j(i \neq j)} \left|\lambda_i - \lambda_j\right|.$$

Since $\lambda_i \neq \lambda_j$ if $i \neq j$, $\epsilon > 0$.

We define D as

$$D = \begin{bmatrix} 1 & & & & \\ & \epsilon & & & \\ & & \epsilon^2 & & \\ & & & \ddots & \\ & & & & \epsilon^{n-1} \end{bmatrix}.$$

Then we use a similarity transform to get B:

$$\begin{aligned} B &= D^{-1}AD \\ &= \begin{bmatrix} 1 & & & & \\ & \epsilon^{-1} & & & \\ & & \epsilon^{-2} & & \\ & & & \ddots & \\ & & & & \epsilon^{-(n-1)} \end{bmatrix} \begin{bmatrix} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_n \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \epsilon & & & \\ & & \epsilon^2 & & \\ & & & \ddots & \\ & & & & \epsilon^{n-1} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \epsilon & & & \\ & \lambda_2 & \epsilon & & \\ & & \ddots & \ddots & \\ & & & \ddots & \epsilon \\ & & & & \lambda_n \end{bmatrix} \end{aligned}$$

The Gershogorin circles of B are:

$$\left|z - \lambda_1\right| \leq \epsilon, \quad \left|z - \lambda_2\right| \leq \epsilon, \quad \ldots \quad \left|z - \lambda_n\right| \leq \epsilon$$

I will now prove that the Gershogorin circles of B are separated. For the sake of contradiction, assume x is in two distinct Gershogorin circles. That is, $\left|x - \lambda_k\right| \leq \epsilon$ and $\left|x - \lambda_\ell\right| \leq \epsilon$. Then, by the triangle inequality,

$$\left|\lambda_k - \lambda_\ell\right| \leq \left|\lambda_k - x\right| + \left|x - \lambda_\ell\right| \leq 2\epsilon = \frac{1}{2} \min_{i,j(i \neq j)} \left|\lambda_i - \lambda_j\right| \leq \frac{1}{2}\left|\lambda_k - \lambda_\ell\right| < \left|\lambda_k - \lambda_\ell\right|.$$

This is a contradiction. Thus, x cannot be in two distinct Gershogorin circles of B. That is, the Gershogorin circles of B are separated.

3. (a): Derive the Power method using $\|\mathbf{x}\|_1$ or $\|\mathbf{x}\|_\infty$ scaling and show its convergence under appropriate conditions. (b): Compare with different method for general and symmetric matrices. You can write a Matlab code to test different methods; and then carry out some analysis.

   (a) The Power method using the infinity-norm is:

   for $k = 1$ until convergence

   $$y^{(k+1)} = Ax^{(k)}$$
   $$x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_\infty}$$
   $$\mu^{(k+1)} = \frac{y_p^{(k+1)}}{x_p^{(k)}}$$

   end

   Examples of stopping criteria which can be used are $|\mu^{(k+1)} - \mu^{(k)}| <$tol, or $\|y^{(k+1)} - y^{(k)}\|_\infty <$ tol.

   Proof of convergence:

   Let $v_1, \dots, v_n$ be the eigenvectors of A. Assume $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ and that $x_0 = \sum_{i=1}^n \alpha_i v_i$ with $\alpha_1 \neq 0$.

   Note that

   $$y^{(k)} = Ax^{(k-1)} = A\frac{y^{(k-1)}}{\|y^{(k-1)}\|_\infty} = \gamma_k Ay^{(k-1)} = \gamma_k\gamma_{k-1}A^2y^{(k-2)} = \gamma_k\gamma_{k-1}\dots\gamma_2A^{k-1}y^{(1)} = C_kA^ky^{(0)}.$$

   Also, since $x^{(k)}$ is defined as $x^{(k)} = y^{(k)}/\|y^{(k)}\|_\infty$, $\|x^{(k)}\|_\infty = 1$.

   Furthermore,

   $$A^kx^{(0)} = \lambda_1^k\left(\alpha_1v_1 + \alpha_2v_2\left(\frac{\lambda_2}{\lambda_1}\right)^k + \cdots + \alpha_nv_n\left(\frac{\lambda_n}{\lambda_2}\right)^k\right).$$

   Therefore,

   $$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty}\frac{x^{(k)}}{\|x^{(k)}\|_\infty} = \frac{\alpha_1v_1}{\|\alpha_1v_1\|_\infty} = \pm v_1.$$

   Note that for the infinity-norm is it not the case that

   $$\lim_{k\to\infty}\left(x^{(k)}\right)^TAx^{(k)} = \lambda_1$$

   because $(x^{(k)})^Tx^{(k)} = \|x^{(k)}\|_2^2$, not $\|x^{(k)}\|_\infty^2$. However, when using the Power method with infinity-norm scaling, it is not that case that $\|x^{(k)}\|_2 = 1$; rather, $\|x^{(k)}\|_\infty = 1$. Thus we must compute μ in a different way.

   Note that if we define μ as

   $$\mu^{(k+1)} = \frac{y_p^{(k+1)}}{x_p^{(k)}}$$

   then

   $$\lim_{k\to\infty}\mu^{(k+1)} = \lim_{k\to\infty}\frac{y_p^{(k+1)}}{x_p^{(k)}} = \lim_{k\to\infty}\frac{(Ax^{(k)})_p}{x_p^{(k)}} = \frac{(Av_1)_p}{(v_1)_p} = \frac{\lambda_1(v_1)_p}{(v_1)_p} = \lambda_1.$$

   (b) I am comparing the Power method using the 2-norm and the infinity-norm. Matrices A and C from question 6 are used to test the different power methods. We choose ($n =$ 10,20,40,80,160). The computed dominant eigenvalues are compared to the "true" dominant eigenvalues. For matrix A, the true dominant eigenvector is found using the lemma about the eigenvalues of tridiagonal matrices from class. Thus, the dominant eigenvalues of A is $\lambda_1 = 2 - 2\cos\left(\frac{n\pi}{n+1}\right)$.

n	10	20	40	80	160
k	47	128	297	745	1407
Rel. Error	6.0341×10^{-2}	1.6726×10^{-2}	4.4133×10^{-3}	1.1834×10^{-3}	6.0293×10^{-4}
Residual	2.0146×10^{-5}	2.7372×10^{-4}	3.3184×10^{-3}	7.8869×10^{-3}	1.3970×10^{-2}

(a) Matrix A with infinity-norm scaling.

n	10	20	40	80	160
k	529	559	331	624	839
Rel. Error	6.7187×10^{-7}	7.0403×10^{-7}	2.4093×10^{-7}	6.1276×10^{-7}	1.0498×10^{-6}
Residual	1.3027×10^{-4}	2.4304×10^{-4}	6.0696×10^{-6}	1.4570×10^{-5}	6.9975×10^{-5}

(b) Matrix C with infinity-norm scaling.

n	10	20	40	80	160
k	29	88	261	717	1558
Rel. Error	6.0342×10^{-2}	1.6724×10^{-2}	4.4060×10^{-3}	1.1555×10^{-3}	4.0397×10^{-4}
Residual	9.0740×10^{-4}	1.2844×10^{-3}	1.3826×10^{-3}	1.4044×10^{-3}	1.4123×10^{-3}

(c) Matrix A with 2-norm scaling.

n	10	20	40	80	160
k	99	388	508	351	541
Rel. Error	7.2823×10^{-8}	3.4935×10^{-7}	1.4660×10^{-4}	2.9879×10^{-7}	4.9043×10^{-7}
Residual	5.1527×10^{-6}	4.3499×10^{-6}	9.9481×10^{-2}	6.9722×10^{-6}	7.5195×10^{-5}

(d) Matrix C with 2-norm scaling.

Figure 1: Results of testing the Power Method with different scalings. The number of iterations until convergence, k, relative error, $\|\lambda_1(\text{calc}) - \lambda_1(\text{true})\|_2 / \|\lambda_1(\text{true})\|_2$, and residual, $\|Ax - \lambda x\|_2$, for various n for each matrix. Initial guess $[1\ 1 \ldots 1]^T$ was used for all matrices.

The Matlab `eig` function was used to compute the dominant eigenvalues for matrix C. Note that A is a symmetric matrix and C is a general matrix.

For each n, each matrix, and each method, the number of iterations, k, the relative error, and residual was computed. This data is displayed in Figure 1. The power method sucessfully finds the dominant eigenvalues with both scalings.

4.

(a) Use the Gershgorin's theorem to locate the intervals that contain the eigenvalues of A

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -6 & 0 \\ 0 & 1 & 9 \end{bmatrix}.$$

(b) Can A have complex eigenvalues? Why?

(c) Is A diagonalizable? Why?

(d) Apply one step Power method using the 2-norm, and the x_p notation.

(e) Assume that eigenvalues A satisfy $|\lambda_1| > |\lambda_2| > |\lambda_3|$, apply one step ***shifted inverse Power method*** to approximate the eigenvalue λ_2 and its eigenvector with initial guess $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$.

Hint: You can use Matlb command [l u p]= lu(A) to find $PA = LU$ decomposition.

(a) The intervals that contain the eigenvalues of A are

$$|z| \leq 2 \qquad |z+6| \leq 1 \qquad |z-9| \leq 1.$$

(b) No, A cannot have complex eigenvalues because the Gershogorin circles do not intersect. Consequently, there is exactly one eigenvalue in each circle, but since complex eigenvalues occur in pairs which have the same complex modulus (in real-valued matrices), A cannot have complex eigenvalues.

(c) A has 3 distinct eigenvalues, this A is diagonalizable.

(d) When applying the Power method, we will use inital guess $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$.

<u>2-norm:</u>

$$y^{(1)} = Ax^{(0)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -6 & 0 \\ 0 & 1 & 9 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -6 \\ 1 \end{bmatrix}$$

$$x^{(1)} = \frac{y^{(1)}}{\|y^{(1)}\|_2} = \frac{1}{\sqrt{38}} \begin{bmatrix} 1 \\ -6 \\ 1 \end{bmatrix}$$

$$\mu^{(1)} = (x_{(1)})^T A x_{(1)} = \frac{1}{38} \begin{bmatrix} 1 & -6 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & -6 & 0 \\ 0 & 1 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ -6 \\ 1 \end{bmatrix}$$

$$= \frac{1}{38} \begin{bmatrix} 1 & -6 & 1 \end{bmatrix} \begin{bmatrix} -5 \\ 37 \\ 3 \end{bmatrix} = -224/38.$$

<u>x_p notation:</u>

$$y^{(1)} = Ax^{(0)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -6 & 0 \\ 0 & 1 & 9 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -6 \\ 1 \end{bmatrix}$$

$$x^{(1)} = \frac{y^{(1)}}{\|y^{(1)}\|_\infty} = \begin{bmatrix} 1/6 \\ -1 \\ 1/6 \end{bmatrix}$$

$$\mu^{(1)} = \frac{y_p^{(1)}}{x_p^{(0)}} = \frac{-6}{1} = -6$$

(e) Note that λ_2 is in the Gershogorin circle centered at -6, and that λ_2 is closer to -6 than any other eigenvalue is. Thus we will use $\sigma = -6$ for our approximation to λ_2. The first step of the shifted inverse power method follows:

First we must solve the following system of equations for $y^{(1)}$.

$$(A + 6I)y^{(1)} = x^{(0)}$$

$$\begin{bmatrix} 6 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 15 \end{bmatrix} y^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

The $PA = LU$ decomposition of $A + 6I$ is

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/6 & -1/6 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 6 & 1 & 1 \\ 0 & 1 & 15 \\ 0 & 0 & 7/3 \end{bmatrix} \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Therefore we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/6 & -1/6 & 1 \end{bmatrix} \begin{bmatrix} 6 & 1 & 1 \\ 0 & 1 & 15 \\ 0 & 0 & 7/3 \end{bmatrix} y^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

First we solve (using forward substitution):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/6 & -1/6 & 1 \end{bmatrix} z^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

to get

$$z^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

then we solve (using backward substitution)

$$\begin{bmatrix} 6 & 1 & 1 \\ 0 & 1 & 15 \\ 0 & 0 & 7/3 \end{bmatrix} y^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

to get

$$y^{(1)} = \begin{bmatrix} 1 \\ -45/7 \\ 3/7 \end{bmatrix}.$$

Now we have

$$x^{(1)} = y^{(1)}/\|y^{(1)}\|_2 = \frac{7}{\sqrt{2083}} \begin{bmatrix} 1 \\ -45/7 \\ 3/7 \end{bmatrix}.$$

Finally, we compute

$$\mu^{(1)} = (x^{(1)})^T A x^{(1)} = \frac{49}{2083} \begin{bmatrix} 1 & -45/7 & 3/7 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & -6 & 0 \\ 0 & 1 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ -45/7 \\ 3/7 \end{bmatrix} \approx -6.1512.$$

5. Let $x = \begin{bmatrix} 3 & 0 & -1 & 2 \end{bmatrix}^T$, and $y = \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}^T$.

(a) Is there a Householder matrix P such that $Px = y$? Explain.

(b) Let $\tilde{y} = \alpha y$, find the scalar α and a Householder matrix P such that $Px = \tilde{y}$.

(c) Find the QR decomposition of the following matrix.

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 5 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 2 \end{bmatrix}.$$

(a) There is no Householder matrix P such that $Px = y$. Householder matricies are orthogonal, therefore $\|Px\|_2 = \|x\|_2$. Since $Px = y$ implies $\|Px\|_2 = \|y\|_2$, it must be the case that $\|x\|_2 = \|y\|_2$. However, $\|x\|_2 = \sqrt{14} \neq 1 = \|y\|_2$.

(b) Since $Px = \tilde{y}$, $\sqrt{14} = \|Px\|_2 = \|\tilde{y}\|_2 = \|\alpha y\|_2 = |\alpha|\|y\|_2 = |\alpha|$. Thus, $\alpha = \pm\sqrt{14}$. To avoid cancellation error, we choose $\alpha = \sqrt{14}$. This implies $\tilde{y} = \begin{bmatrix} -\sqrt{14} & 0 & 0 & 0 \end{bmatrix}^T$. To find the Householder matrix such that $Px = \tilde{y}$, we note that $P = I - 2\omega\omega^T$ and $\omega = (x - \tilde{y})/\|x - \tilde{y}\|_2$. Thus

$$\omega = \begin{bmatrix} 3 + \sqrt{14} \\ 0 \\ -1 \\ 2 \end{bmatrix} \Big/ \sqrt{28 + 6\sqrt{14}}$$

and

$$\begin{aligned} P &= I - 2\omega\omega^T \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \frac{2}{28 + 6\sqrt{14}} \begin{bmatrix} 3 + \sqrt{14} \\ 0 \\ -1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 + \sqrt{14} & 0 & -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \frac{2}{28 + 6\sqrt{14}} \begin{bmatrix} 25 + 6\sqrt{14} & 0 & -3 - \sqrt{14} & 6 + 2\sqrt{14} \\ 0 & 1 & 0 & 0 \\ -3 - \sqrt{14} & 0 & 1 & -2 \\ 6 + 2\sqrt{14} & 0 & -2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} -0.8018 & 0 & 0.2673 & -0.5345 \\ 0 & 1 & 0 & 0 \\ 0.2673 & 0 & 0.9604 & 0.0793 \\ -0.5345 & 0 & 0.0793 & 0.8414 \end{bmatrix} \end{aligned}$$

(c) The first two columns of A are already in the correct form, so we will focus only on the third column of A. The goal is to find a P' such that $P'x = y$ where

$$x = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix}.$$

We require $\|x\|_2 = \|y\|_2$, so $\|y\|_2 = \|x\|_2 = \sqrt{5}$. Therefore $\alpha = \pm\sqrt{5}$. Since the first component of x is zero, we arbitrarily choose $\alpha = +\sqrt{5}$. Then

$$\omega = \frac{x-y}{\|x-y\|_2} = \begin{bmatrix} -\sqrt{5} \\ -1 \\ 2 \end{bmatrix} / \sqrt{10}$$

and

$$P' = I - 2\omega\omega^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{5}\begin{bmatrix} 5 & \sqrt{5} & -2\sqrt{5} \\ \sqrt{5} & 1 & -2 \\ -2\sqrt{5} & -2 & 4 \end{bmatrix} = \begin{bmatrix} 0 & -1/\sqrt{5} & 2/\sqrt{5} \\ -1/\sqrt{5} & 4/5 & 2/5 \\ 2/\sqrt{5} & 2/5 & 1/5 \end{bmatrix}$$

. Now

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/\sqrt{5} & 2/\sqrt{5} \\ 0 & 0 & -1/\sqrt{5} & 4/5 & 2/5 \\ 0 & 0 & 2/\sqrt{5} & 2/5 & 1/5 \end{bmatrix}$$

and

$$PA = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/\sqrt{5} & 2/\sqrt{5} \\ 0 & 0 & -1/\sqrt{5} & 4/5 & 2/5 \\ 0 & 0 & 2/\sqrt{5} & 2/5 & 1/5 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 0 & 5 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 5 & 3 \\ 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = R.$$

To get $A = QR$, we note $PA = R$, and $P^T = P^{-1}$, therefore $A = P^T R$. Note that P^T is an orthogonal matrix, so $Q = P^T$. Thus $A = QR$ where $Q = P^T$ and R is above.

6. Write a computer program to find the least dominant eigenvalue of the matrix A and corresponding the unit eigenvector. Test your code for the following matrices:

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 2 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 2 & -1 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & 1 & 2 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & 2 \end{bmatrix};$$

C is a n by n random matrix generated by (i), generate a diagonal matrix, $n = 10; D = diag(rand(n, 1) * 100); S = rand(n, n); C = inv(S) * D * S$. In this way, we can generate a matrix with known eigenvalues.

Tabulate the number of iterations, the relative error of the eigenvalue, and the residue vector $\|Ax - \lambda x\|_2$ for $n = 10, 20, 40, 80, 160$. For the second matrix, also try $n = 11, 21, 41, 81, 161$. Explain your results. **Hint:** Use the Lemma learned in class to find the exact eigenvalues of A for the first matrix; and the Matlab function $eig(A)$ for the second matrix.

To find the least dominant eigenvalue, I used the inverse power method. The stopping criteria was reaching a error of less than the tolerance of 10^{-16}. Error was here defined as $\|x_{k+1} - x_k\|_2 / \|x_k\|_2$. An initial guess of $\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T$ was used for each matrix. The theoretical ("true") least dominant eigenvalue for matrix A was calculated using $\lambda_n = 2 - 2\cos(\pi/(n+1))$. This comes from the lemma stating that the eigenvalues for tridiagonal matricies are $\lambda_i = \alpha + 2\beta\cos(k\pi/(n+1))$.
When I tried to find the least dominant eigenvalue for matrix B using even n, the tolerance was never met. To find out why this was the case, I used Matlab's `eig` function to find the eigenvalues of B for even n (n = 6,10,18, and 34). In each case, the eigenvalues were all complex and occured in pairs. As a result, there were always two least dominant eigenvalues. In this case, the inverse power mthod does not converge. This is most likely the reason that the tolerance we never met when I ran my code.

Matrix A:

n	10	20	40	80	160
k	7	7	7	7	7
Rel. Error	6.9548×10^{-14}	4.3954×10^{-14}	3.3403×10^{-14}	3.1717×10^{-15}	2.3905×10^{-13}
Residual	5.8835×10^{-8}	1.3393×10^{-8}	3.3349×10^{-9}	8.4269×10^{-10}	2.1254×10^{-10}

Matrix B:

n	11	21	41	81	161
k	381	1173	3992	14,254	51,961
Rel. Error	4.4409×10^{-16}	1.8874×10^{-15}	1.9984×10^{-15}	2.2204×10^{-16}	5.1070×10^{-15}
Residual	1.9679×10^{-6}	1.9880×10^{-6}	1.9979×10^{-6}	1.9988×10^{-6}	2.0000×10^{-6}

Matrix C:

n	10	20	49	80	160
k	15	183	12	11	98
Rel. Error	2.3842×10^{-5}	1.4716×10^{-7}	2.9484×10^{-6}	6.7998×10^{-7}	1.3785×10^{-6}
Residual	2.1076×10^{-6}	6.7305×10^{-6}	8.6643×10^{-7}	1.4012×10^{-7}	6.5135×10^{-7}

Figure 2: The number of iterations until convergence, k, relative error, $\|\lambda_n(\text{calc}) - \lambda_n(\text{true})\|_2 / \|\lambda_n(\text{true})\|_2$, and residual, $\|Ax - \lambda x\|_2$, for various n for each matrix. Initial guess $[1\ 1 \dots 1]^T$ was used for all matrices.

Note that the number of interations for matrix B is consistenly high, even for even n. It seemed likely that this was because the inital guess vector had a very low v_n component (where λ_n denotes the least dominant eigenvalue and v_n denotes the corresponding eigenvector with $\|v_n\|_2 = 1$). To test this, I ran my code for matrix B with starting vector $\begin{bmatrix} 1 & 0 & 1 & 0 & \dots & 1 \end{bmatrix}^T$ and all other parameters unchanged. With this starting vector, only 2 iterations were needed for each value of n ($n = 21, 41, 81, 161$). This quick convergence strongly supports the hypothesis.

n	11	21	41	81	161
k	2	2	2	2	2
Rel. Error	6.6613×10^{-16}	1.8874×10^{-15}	1.5543×10^{-15}	3.3307×10^{-16}	5.1070×10^{-15}
Residual	5.5236×10^{-16}	9.7087×10^{-17}	5.1544×10^{-16}	4.0110×10^{-17}	3.2469×10^{-17}

Figure 3: Data for matrix B using intial guess vector $[1\ 0\ 1\ 0 \dots 1]^T$.