HW5

1(a) Since A is symmetric, $A^T = A$

If A has comple eigenvalues, they must appear in conjugate pairs

Suppose for the sake of contradiction that $\lambda \neq \bar{\lambda}$ are eigenvalues of A

$A\vec{x} = \lambda\vec{x}$, $A\vec{x} = \bar{\lambda}\vec{x}$

$\vec{x}^T\lambda\vec{x} = \vec{x}^T A\vec{x} = (\vec{x}^T A^T \vec{x})^T = (\vec{x}^T A \vec{x})^T = (\vec{x}^T \bar{\lambda}\vec{x})^T = \vec{x}^T\bar{\lambda}\vec{x}$

$\Rightarrow \lambda = \bar{\lambda}$ contradict

So all eigenvalues are real □

(b) $R_1 = 2 + 0.3 + 0.7 = 3$ $D_1 = \{|z+6| \leq 3\} = [-9, -6]$

$R_2 = 2 + 0.1 + 0.05 = 2.15$ $D_2 = \{|z+4| \leq 2.15\} = [-6.15, -1.85]$

$R_3 = 0.3 + 0.1 + 0.1 + 0.1 = 0.6$ $D_3 = \{|z-2| \leq 0.6\} = [1.4, 2.6]$

$R_4 = 0.05 + 0.1 = 0.15$ $D_4 = \{|z-4| \leq 0.15\} = [3.85, 4.15]$

$R_5 = 0.7 + 0.1 = 0.8$ $D_5 = \{|z-6| \leq 0.8\} = [5.2, 6.8]$

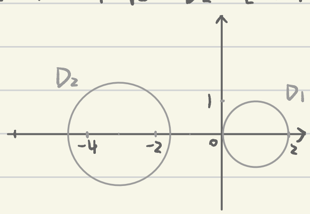
So $D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 = [-9, -1.85] \cup [1.4, 2.6] \cup [3.85, 4.15] \cup [5.2, 6.8]$

(c) No. Since D_1 and D_2 are not disjoint, it is possible that λ with greatest absolute value lies in the intersection of D_1 and D_2, and thus is a root of multiplicity 2 of $\det(A - \lambda I) = 0$, i.e. λ is not simple

(d) Since $0 \notin D$, A is invertible.

2(a) $R_1 = 1$ $D_1 = [0, 2]$

$R_2 = |-1-i| = \sqrt{2}$ $D_2 = [-3-\sqrt{2}, -3+\sqrt{2}]$



(b) Since D_1 and D_2 are disjoint, D_3 must intersect with one of the two disks

$R_3 = |2i| = 2$ $([z-2, z+2] \cap D_1) \cup ([z-2, z+2] \cap D_2) \neq \phi$

So $z \in [-5-\sqrt{2}, 4]$

(c) $D_1 \cap D_2 \cap D_3 \neq \phi$

So $z \in [-2, -1+\sqrt{2}]$

3(a) See the code

(b) See the code

When $x_0 = [1\ 2\ -1]^T$, x_k converge to around $[0.7, 0, -0.7]$, λ_1 converge to -5.97

When $x_0 = [1\ 2\ 1]^T$, x_k converge to around $[0.58, 0.58, 0.58]$, λ_1 converge to 3.00

Compared to result of Matlab inbuilt eig function X_k and λ_k converges to the largest $|\lambda_1|$ and its corresponding eigenvector $\vec{v}_1$

Note that $[1\ 2\ 1]^T \vec{v}_1 = 0$ violates the prerequisite of power method, the algorithm converge to the second largest $|\lambda_2|$.

(c) See the code

(d) See the code. As I picked $\alpha_1, \alpha_2, \alpha_3$ each close to $\lambda_1, \lambda_2, \lambda_3$, and with appropriate choice of x_0, λ each converge to $\lambda_1 \approx -6$, $\lambda_2 \approx 0$, $\lambda_3 \approx 3$, $\vec{x}_k$ each converge to $\vec{v}_1 = [0.7, 0, -0.7]$, $\vec{v}_2 = [-0.4, 0.8, -0.4]$, $\vec{v}_3 = [-0.58, -0.58, -0.58]$, the corresponding eigenvectors.

4(a)

$$\vec{x}_k = c_1\lambda_1^k \vec{v}_1 + \sum_{i=2}^{n} c_i\lambda_i^k \vec{v}_i = c_1\lambda_1^k \left(\vec{v}_1 + \sum_{i=2}^{n} \frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)$$

$$r_k = \frac{\left(c_1\lambda_1^k\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)\right)^T A\, c_1\lambda_1^k\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}{\left(c_1\lambda_1^k\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)\right)^T c_1\lambda_1^k\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}$$

$$= \frac{c_1^2\lambda_1^{2k}\left(\vec{v}_1^T + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\right) A\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}{c_1^2\lambda_1^{2k}\left(\vec{v}_1^T + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\right)\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}$$

$$= \frac{\left(\vec{v}_1^T A + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T A\right)\left(\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}{\vec{v}_1^T\vec{v}_1 + \vec{v}_1^T\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\vec{v}_1 + \left(\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\right)\left(\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}$$

$\vec{v}_i^T\vec{v}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

$$= \frac{\vec{v}_1^T\lambda_1\vec{v}_1 + \vec{v}_1^T\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \lambda_i\vec{v}_i + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\lambda_1\vec{v}_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T \cdot \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \lambda_i\vec{v}_i}{\vec{v}_1^T\vec{v}_1 + \vec{v}_1^T\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i + \sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\vec{v}_1 + \left(\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i^T\right)\left(\sum_{i=2}^{n}\frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v}_i\right)}$$

$$= \frac{\lambda_1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}\lambda_i}{1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}} = \frac{\lambda_1\left(1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k+1}\right)}{1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}}$$

$$= \frac{\lambda_1\left(1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k} + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k+1} - \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}\right)}{1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}}$$

$$= \lambda_1\left(1 + \underbrace{\frac{\sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}\left(\frac{\lambda_i}{\lambda_1} - 1\right)}{1 + \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}}}_{a_k}\right)$$

$$a_k\left(\frac{\lambda_2}{\lambda_1}\right)^{-2k} = \frac{\sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_2}\right)^{2k}\left(\frac{\lambda_i}{\lambda_1}-1\right)}{1+\sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\underbrace{\left(\frac{\lambda_i}{\lambda_1}\right)^{2k}}_{\to 0}}$$

Since for $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$, $\lim_{k\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^{2k} = 0$

$$a_k\left(\frac{\lambda_2}{\lambda_1}\right)^{-2k} = \sum_{i=2}^{n}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_2}\right)^{2k}\left(\frac{\lambda_i}{\lambda_1}-1\right)$$

$|\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_j| > \cdots \geq |\lambda_n|$, consider $\{\lambda_2, \lambda_3, \cdots \lambda_j\}$ such that $\lambda_3 = \lambda_4 = \cdots \lambda_j = \lambda_2$

$$a_k\left(\frac{\lambda_2}{\lambda_1}\right)^{-2k} = \sum_{i=2}^{j}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}-1\right) + \sum_{i=j+1}^{n}\left(\frac{c_i}{c_1}\right)^2\underbrace{\left(\frac{\lambda_i}{\lambda_2}\right)^{2k}}_{\to 0}\left(\frac{\lambda_i}{\lambda_1}-1\right) = \sum_{i=2}^{j}\left(\frac{c_i}{c_1}\right)^2\left(\frac{\lambda_i}{\lambda_1}-1\right) = C$$

Thus $r_k = \lambda_1(1+a_k)$ where $a_k\left(\frac{\lambda_2}{\lambda_1}\right)^{-2k} \to C$ as $k\to\infty$. C does not depend on k. □

(b) By Gerschgorin Thm, $\lambda_{max} \in [20, 22]$

$$\vec{x}_1 = A\vec{x}_0 = \begin{bmatrix} * \\ * \\ * \\ * \\ 21 \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 21 \end{bmatrix} = \lambda_{max}\vec{x}_0 \qquad \vec{y}_1 = \frac{\vec{x}_1}{\|\vec{x}_1\|} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 21 \end{bmatrix}$$

Note that $\lambda_{max}\vec{x}_0 = A\vec{x}_0$, $\vec{x}_0$ is the corresponding eigenvector of λ_{max}

So $\vec{y}_k$ generated by power method converge to the eigenvector belong to λ_{max}.

(c) $\varepsilon_k = |r_k - \lambda_1| = |\lambda_1(1+a_k) - \lambda_1| = |\lambda_1 a_k|$

$\varepsilon_{k+5} = |\lambda_1 a_{k+5}|$

Since $\lambda_1 \approx 21$ $\lambda_2 \approx -9$, $a_k\left(\frac{\lambda_2}{\lambda_1}\right)^{-2k} \to c$ as $k\to\infty$

$$\lim_{k\to\infty}\frac{|\varepsilon_{k+5}|}{|\varepsilon_k|} = \lim_{k\to\infty}\frac{|a_{k+5}|}{|a_k|} = \lim_{k\to\infty}\frac{\left|c\left(\frac{\lambda_2}{\lambda_1}\right)^{2(k+5)}\right|}{\left|c\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right|} = \left(\frac{\lambda_2}{\lambda_1}\right)^{10}$$

$\rho = -\log_{10}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{10}\right) \approx 3.68$

So we could gain 3 to 4 more correct decimals

5 (a) Prove by induction

Base Case: - $A_p^{(3)} = \begin{bmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{bmatrix}$ A of odd size

$$\det(A_p^{(3)} - xI) = -x(-x(-a_2 - x) + a_1) - a_0 = -x^3 - a_2x^2 - a_1x - a_0 = -p^{(3)}(x)$$

$$\det(A_p^{(3)} - xI) = 0 \iff p^{(3)}(x) = 0$$

- $A_p^{(4)} = \begin{bmatrix} 0 & 0 & 0 & -a_0 \\ 1 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & -a_2 \\ 0 & 0 & 1 & -a_3 \end{bmatrix}$ A of even size

$$\det(A_p^{(4)} - xI) = -x(-x(-x(-a_3 - x) + a_2) - a_1) + a_0$$
$$= x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = p^{(4)}(x)$$

$$\det(A_p^{(4)} - xI) = 0 \iff p^{(4)}(x) = 0$$

Induction: $A_p^{(n)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}$ $A_p^{(n+1)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 & -a_n \end{bmatrix}$

$$\det(A_p^{(n+1)} - xI) = -x \cdot \det\begin{bmatrix} -x & 0 & 0 & \cdots & 0 & -a_1 \\ 1 & -x & 0 & \cdots & 0 & -a_2 \\ & & \cdots & & & \\ 0 & 0 & 0 & & 1 & -a_n - x \end{bmatrix} + (-1)^{n+2}(-a_0) \cdot \det\underbrace{\begin{bmatrix} 1 & -x & 0 & \cdots & 0 \\ 0 & 1 & -x & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & & 1 \end{bmatrix}}_{\text{triangular}}$$

- When n is even, $n+1$ is odd

$$\det(A_p^{(n+1)} - xI) = -x(x^n + a_nx^{n-1} + \cdots + a_1) - a_0 = -x^{n+1} - a_nx^n - \cdots - a_1x - a_0 = -p^{(n+1)}(x)$$

- When n is odd, $n+1$ is even

$$\det(A_p^{(n+1)} - xI) = -x(-x^n - a_nx^{n-1} - \cdots - a_1) + a_0 = x^{n+1} + a_nx^n + \cdots + a_1x + a_0 = p^{(n+1)}(x)$$

In either case, $\det(A_p^{(n+1)} - xI) = 0 \iff p^{(n+1)}(x) = 0$

Thus the roots of $p(x)$ could be computed as the eigenvalues of A_p. □

(b) See the code

6. $\vec{x} = \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ $\vec{v}' = \vec{x} + c\vec{e}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + 3\vec{e}_1 = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$ $\vec{v} = \begin{bmatrix} 0 \\ 4 \\ 2 \\ 2 \end{bmatrix}$

$$H_1 = I - \frac{2}{\vec{v}^T\vec{v}}\vec{v}\vec{v}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ 0 & -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \qquad H_1AH_1 = \begin{bmatrix} 2 & -3 & 0 & 0 \\ -3 & 1 & 3 & 4 \\ 0 & 3 & -3 & -9 \\ 0 & 4 & -9 & -2 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \vec{v}' = \vec{x} + c\vec{e_1} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} + 5\vec{e_1} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \quad \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 4 \end{bmatrix}$$

$$H_2 = I - \frac{2}{\vec{v}^T\vec{v}} \vec{v}\vec{v}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \quad H_2H_1AH_1H_2 = \begin{bmatrix} 2 & -3 & 0 & 0 \\ -3 & 1 & -5 & 0 \\ 0 & -5 & -11 & -3 \\ 0 & 0 & -3 & 6 \end{bmatrix}$$

$$Q = H_2H_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ 0 & \frac{14}{15} & -\frac{2}{15} & -\frac{1}{3} \\ 0 & \frac{2}{15} & -\frac{11}{15} & \frac{2}{3} \end{bmatrix}$$

7(a)(b) See the code

(c) From the graph, we can see that larger k leads to a wider spread in the value of eigenvalues.

Note that powe method iterations $\vec{x_k} = c_1\lambda_1^k(\vec{v_1} + \sum_{i=2}^{n} \frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \vec{v_i})$

$\vec{x_k} - c_1\lambda^k\vec{v_1} \propto C\left(\frac{\lambda_2}{\lambda_1}\right)^k$, so the sequence converges at the speed $\frac{|\lambda_2|}{|\lambda_1|}$

When $k<1$, $\frac{|\lambda_2|}{|\lambda_1|}$ is smaller, with λ_1 fixed at one.

So $k<1$ improves the speed of convergence.

8(a) $L_k(x_i) = \begin{cases} 1 & i=k \\ 0 & i \neq k \end{cases}$ could be constructed as a linear combination of basis of P_n.

So $L_k(x_i) \in P_n$

$\forall i \in \{0, \dots n\}$, $\sum_{k=0}^{n} \alpha_k L_k(x_i) = 0$

$$\alpha_0 L_0(x_i) + \alpha_1 L_1(x_i) + \dots + \alpha_n L_n(x_i) = \alpha_i L_i(x_i) = \alpha_i = 0$$

So $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$

(b) $\forall i \in \{0, \dots n\}$, $\alpha_0 L_0(x_i) + \alpha_1 L_1(x_i) + \dots + \alpha_n L_n(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_n x_i^n$

$$\alpha_i = \alpha_i L_i(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_n x_i^n = (1 \ \ x_i \ \ x_i^2 \ \dots \ x_i^n)\vec{\beta}$$

$$\vec{\alpha} = \begin{bmatrix} \beta_0 + x_0\beta_1 + x_0^2\beta_2 + \dots + x_0^{n-1}\beta_{n-1} + x_0^n\beta_n \\ \beta_0 + x_1\beta_1 + x_1^2\beta_2 + \dots + x_1^{n-1}\beta_{n-1} + x_0^n\beta_n \\ \dots \\ \beta_0 + x_n\beta_1 + x_n^2\beta_2 + \dots + x_n^{n-1}\beta_{n-1} + x_n^n\beta_n \end{bmatrix} = V\vec{\beta}$$

(c) $\mathrm{cond}(V) \gg 1$, this basis transform cannot be performed accurately.