BITS Pilani
Pilani Campus

Mathematical Foundations for Data Science

MFDS Team

innovate achieve lead

BITS Pilani
Pilani Campus

DSECL ZC416, MFDS

Webinar#2

Agenda

1. Homework 3.4 Problems
2. Rank Nullity Theorem for Linear Transformations.
3. Problem on Eigenvalues and Eigenvectors.
4. Power method exemplification
5. Problem on Rayleigh’s method.
6. Illustration of Gerschgorin’s result for a few cases.

3.4 Homework Problems

Q1. Let B = (b_1, b_2,..., b_{r-1}, b_r, b_{r+1},..., b_n) be a non-singular matrix. If column b_r is replace by 'a' and that the resulting matrix is called B_a along with $a = \sum_{i=1}^{i=n} y_i\, b_i$, then state the necessary and sufficient condition for B_a to be non-singular.

Solution :

Given matrix **B=(b_1,b_2,...,b_{r-1},b_r,b_{r+1},...,b_n)** is non-singular

$\rightarrow$ det(B)≠0

Therefore, the vectors are linearly independent.

Therefore $\alpha_1\mathbf{b}_1+\alpha_2\mathbf{b}_2+\ldots+\alpha_{r-1}\mathbf{b}_{r-1}+\alpha_r\mathbf{b}_r+\alpha_{r+1}\mathbf{b}_{r+1}+\ldots+\alpha_n\mathbf{b}_n=0$

Implies α_i=0 for i=1,2,3,…,r,…n -------- (I)

The column vector $\mathbf{b}_r$ is replaced by vector '**a**', to get a matrix

$\mathbf{B}_a=[\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_{r-1},\mathbf{a},\mathbf{b}_{r+1},\ldots,\mathbf{b}_n]$ Where, $\mathbf{b}_r = \boldsymbol{a} = \sum_{i=1}^{i=n} y_i\, \boldsymbol{b_i}$.

$\mathbf{B}_a = [\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_{r-1}, \sum_{i=1}^{i=n} y_i\, \boldsymbol{b_i}\ ,\ \mathbf{b}_{r+1},\ldots,\mathbf{b}_n]$

Necessary Condition

Assume B_a is non-singular then $\det(B_a)\neq 0$

i.e

$\alpha_1\mathbf{b}_1+\alpha_2\mathbf{b}_2+\ldots\ \alpha_{r-1}\mathbf{b}_{r-1}+\alpha_r\, [\, \sum_{i=1}^{i=n} y_i\, \boldsymbol{b_i}] + \alpha_{r+1}\mathbf{b}_{r+1}+\ldots+\alpha_n\mathbf{b}_n=0$

$(\alpha_1+\alpha_r\, y_1)\, \mathbf{b}_1 + (\alpha_2+\alpha_r\, y_2)\, \mathbf{b}_2 + \ldots + \alpha_r\, y_r\, \mathbf{b}_r + \ldots + (\alpha_n+\alpha_r\, y_n)\, \mathbf{b}_n = 0$

Implies

$\alpha_i+\alpha_r\, y_i = 0$ for all i=1,2,3,…r-1,r+1, ….n and also $\alpha_r\, y_r = 0$ for i=r

$\alpha_r = 0$, from (I)

Therefore, the required condition for B_a is to be non-singular is $y_r \neq 0$.

Therefore , y_r =0 or $y_r \neq 0$

If y_r =0 then in the system $\boldsymbol{a} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{i=n}} y_{\boldsymbol{i}}\, \boldsymbol{b_i}$) ,i.e in augmented matrix **[B:a])** the r^th column will be zero and it leads to det(B)=0. This contradicts that B is non-singular.

Therefore, $y_r \neq 0$.

Sufficient Condition

Suppose $y_r \neq 0$ then in matrix $\mathbf{B_a}$ the vector ‘ $\boldsymbol{a} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{i=n}} y_i\, \boldsymbol{b_i}$ ’ will become a new independent column.

Therefore, the columns of $\mathbf{B_a}$ are linearly independent .

Hence B_a in non-singular.

Rank Nullity Theorem for Linear Transformation

Theorem: Let $T : V \to W$ be a linear transformation, with V be finite dimensional vector space. Then

dim(V)=dim(R(T))+dim(N(T)) =Rank of T+ Nullity of T

Where R(T) is Range of T and N(T) is Kernel of T.

Proof: First, Let us prove that N(T) is subspace of V

Let $v_1, v_2 \in N(T) \Rightarrow T(v_1) = 0, T(v_2) = 0$

Consider $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2) = 0 \Rightarrow \alpha v_1 + \beta v_2 \in N(T)$

Hence N(T) is subspace of V.

Let $\{v_1, v_2, v_3 ......v_n\}$ be the basis of N(T). Hence Nullity of T is n i.e.

dim(N(T))=n

Let us pick the basis for V be $S = \{v_1, v_2, v_3 ......v_n, v_{n+1}, v_{n+2}, ...v_{n+m}\}$ implies (I)

dim(V)=n + m

Let us consider $w \in R(T) \Rightarrow w = T(v)$ for v is in V

Since v is in V

$$v = \alpha_1 v_1 + \alpha_2 v_2 ..... + \alpha_n v_n + \alpha_{n+1} v_{n+1} + \alpha_{n+2} v_{n+2} ...... + \alpha_{n+m} v_{n+m}$$

Operating T on both sides

$$w = T\left(\alpha_1 v_1 + \alpha_2 v_2 ..... + \alpha_n v_n + \alpha_{n+1} v_{n+1} + \alpha_{n+2} v_{n+2} ...... + \alpha_{n+m} v_{n+m}\right)$$

$$w = \alpha_1 T(v_1) + \alpha_2 T(v_2) ..... + \alpha_n T(v_n) + \alpha_{n+1} T(v_{n+1}) + \alpha_{n+2} T(v_{n+2}) ..... + \alpha_{n+m} T(v_{n+m})$$

$$w = \alpha_{n+1} T(v_{n+1}) + \alpha_{n+2} T(v_{n+2}) ..... + \alpha_{n+m} T(v_{n+m}), \text{ since } \sum_{j=1}^{n} \alpha_j T(v_j) = 0$$

$$\Rightarrow R(T) = span\{T(v_{n+1}), T(v_{n+2}), ... T(v_{n+m})\}$$

Now, let us show that $\{T(v_{n+1}), T(v_{n+2}), ... T(v_{n+m})\}$ is linearly independent. Let us consider

$$\beta_{n+1} T(v_{n+1}) + \beta_{n+2} T(v_{n+2}) ..... + \beta_{n+m} T(v_{n+m}) = 0$$

$$T\left(\beta_{n+1} v_{n+1} + \beta_{n+2} v_{n+2} ..... + \beta_{n+m} v_{n+m}\right) = 0$$

$$\sum_{i=1}^{m} T\left(\beta_{i+n} v_{i+n}\right) = \sum_{i=1}^{m} \beta_{i+n} T\left(v_{i+n}\right) = 0$$

This implies $\beta_{i+n} = 0$ for each i. From (I)

Hence $\{T(v_{n+1}), T(v_{n+2}), \ldots . T(v_{n+m})\}$ is linearly independent and forms

basis of R(T) and hence dim(R(T))=m.

Implies

dim(V)=dim(R(T))+dim(N(T)) =Rank of T+ Nullity of T

MID SEM-2021

Example

Let V and W be vector spaces of dimensions n and m respectively over $\mathbb{R}$ and $T : V \to W$ be a linear transformation. Aravind starts with the basis $\{u_1, u_2, \ldots, u_r\}$, $(r < n)$ of $\text{Ker}(T)$ and adds a few more elements of V namely, $\{u_{r+1}, u_{r+2}, \ldots, u_n\}$ such that the combined set $\{u_1, u_2, \ldots, u_n\}$ becomes a basis for V. How should Aravind proceed to prove that $\text{Range}(T)$ has dimension $n - r$, in order that the rank-nullity theorem is satisfied? Assumptions, if any, are to be stated clearly.

Solution:

Let $b \in Range(T)$and $v \in V$ and let $T(v) = b$. Since $\{u_1, u_2, \ldots, u_r, u_{r+1} \ldots \ldots, u_n\}$ is a basis for V, we can write

$$v = \sum_{i=1}^{n} \alpha_i u_i = \sum_{i=1}^{r} \alpha_i u_i + \sum_{i=r+1}^{n} \alpha_i u_i .$$

Hence,

$$b = T(v) = T(\textstyle\sum_{i=r+1}^{n} \alpha_i u_i)$$

since $T(u_i)_{i=1\ldots r} = 0$.

Also, since the set $\{u_i\}_{r+1}^{n}$ is linearly independent,

$\{T\{u_i\}_{i=r+1}^{n}\}$ *are also linearly independent*.

Therefore, the dimension of $Range(T) = n - r$.

Eigen Values and Eigen Vectors

Example-1:

Let γ be the real root of a polynomial equation of degree 9 with integer coefficients. Construct the matrix $A = \begin{bmatrix} 2 & \gamma & 0 \\ \gamma & 2 & \gamma \\ 0 & \gamma & 2 \end{bmatrix}$

With this information, is it possible to

a) **derive all the possible values of γ so that A has all non-zero eigen-values?**

b) **Calculate the necessary condition on γ so that all the eigenvalues of A are positive?**

Solution:

If matrix has only nonzero eigenvalues , then determinant is not zero this means $2-\gamma^2 \neq 0$. So $\gamma \neq \sqrt{2}$

Similarly, one necessary condition for all positive eigenvalues is $2-\gamma^2 > 0$. In summary $-\sqrt{2} \leq \gamma \leq \sqrt{2}$

Example-2:

A math professor teaching about eigenvalues and eigenvectors introduces a matrix $B_{n \times n}$ having an interesting property that $B^k = 0$ for some $k < n$. He then ask the students to work out the following questions.

(i) estimate the trace and determinant of B

(ii) derive all the eigenvalues of B.

Are the above tasks possible? If so, derive the results and if not, give reasons.

Solution:

Let λ be an eigenvalue of B for some eigenvector x.

Since $B^k = 0$.

It means that λ should satisfy $\lambda^k = 0$.

This means $\lambda = 0$

Hence B has only zero eigenvalues.

Hence Trace=0 and Determinant=0

Homework 4.4

Example-3:

Q .a) Let P be a real square matrix satisfying $P=P^T$ and $P^2=P$.

i) Can the matrix P have complex eigenvalues?

If so, construct an example, else, justify your answer.

ii) What are the eigenvalues of P?

b) Given the following matrix $A = \begin{bmatrix} 1 & 2 & r \\ c & 1 & 7 \\ c & 1 & 7 \end{bmatrix}$

where c and r are arbitrary real numbers and $5.5 < r \leq 6.5$, and the fact that $\lambda_1 = 3$ is one of the eigenvalues, is it possible to determine the other two eigenvalues? If so, compute them and give reasons for your answer.

Solution :

a) i) The matrix P cannot have complex eigenvalues.

Justification :

Given $P=P^2$ and $P^T=P$

Implies P is a real symmetric matrix.

By the property of eigenvalues real symmetric matrices will have only real eigenvalues.

ii) Let eigenvalue of P be λ then the eigenvalue of P^2 will be λ^2.

Hence, $\lambda^2=\lambda$ (Since $P=P^2$) $\rightarrow$ $\lambda^2-\lambda=0$, $\lambda(\lambda-1)=0$

That implies, $\lambda=0$ or 1.

b)

Given $A = \begin{bmatrix} 1 & 2 & r \\ c & 1 & 7 \\ c & 1 & 7 \end{bmatrix}$ and one of the eigenvalue is λ_1=3.

Since matrix A has 2 identical rows. det(A)=0.

Implies, one of the eigenvalue must be zero say λ_2=0.

W.K.T ,Trace of A=sum of the eigenvalues

So, $\lambda_1+\lambda_2+\lambda_3 = 9$

Implies λ_3=6.

Since matrix A has 2 identical rows.

The value of 'r' will not play any role in computing eigenvalues.

Power Method – Example

Find the dominant eigenvalue & corresponding eigenvector of the matrix

$$A = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix}$$

Solution:

Step 1 : Choose column vector $u_o = [\,1,-1,-1,1]^T$

Step 2 : Multiply the matrix by the matrix [A] by $u_o = y_1$

$$A = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{Bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ -3 \\ 5 \\ 7 \end{Bmatrix} \Rightarrow \begin{Bmatrix} 1 \\ -3 \\ 5 \\ 7 \end{Bmatrix} = 7 \begin{Bmatrix} 0.1429 \\ -0.4286 \\ 0.7143 \\ 1 \end{Bmatrix}$$

Step 3 : Normalize the resulting vector obtained in step 2 by dividing each component by the largest in magnitude

$u_1 = y_1 / 7 = [0.1429, -0.4286 , 0.7143,1]^T$

Normalizing factor $m_1 = 7$

Power Method - Example

Step 4 :

$$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{Bmatrix} 0.1429 \\ -0.4286 \\ 0.7143 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0.1429 \\ -5.2857 \\ 5.5714 \\ 5.8571 \end{Bmatrix} = 5.8571 \begin{Bmatrix} 0.0244 \\ -0.9024 \\ 0.9512 \\ 1.0000 \end{Bmatrix}$$

$u_2 = y_2 / 5.8571$ (normalizing factor m_2)

Now Repeating steps 2 and 3

$$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{Bmatrix} 0.0244 \\ -0.9024 \\ 0.9512 \\ 1.0000 \end{Bmatrix} = \begin{Bmatrix} -1.6342 \\ -11.6830 \\ 8.4147 \\ 5.1464 \end{Bmatrix} = 11.6830 \begin{Bmatrix} -0.1399 \\ -1.0000 \\ 0.7203 \\ 0.4405 \end{Bmatrix}$$

$u_3 = y_3 / 11.6830$ (normalizing factor m_3)

Power Method – Example

Continue the process

$$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{Bmatrix} -0.1399 \\ -1.0000 \\ 0.7203 \\ 0.4405 \end{Bmatrix} = \begin{Bmatrix} -2.9791 \\ -13.2798 \\ 7.3215 \\ 1.3633 \end{Bmatrix} = 13.2798 \begin{Bmatrix} -0.2243 \\ -1.0000 \\ 0.5513 \\ 0.1027 \end{Bmatrix}$$

$u_4 = y_4 / 13.2798$
(normalizing factor m_4)

Continue the process

$$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{Bmatrix} -0.2243 \\ -1.0000 \\ 0.5513 \\ 0.1027 \end{Bmatrix} = \begin{Bmatrix} -3.5703 \\ -13.4486 \\ 6.3083 \\ -0.8327 \end{Bmatrix} = 13.4486 \begin{Bmatrix} -0.2655 \\ -1.0000 \\ 0.4690 \\ -0.0619 \end{Bmatrix}$$

$u_5 = y_5 / 13.4486$
(normalizing factor m_5)

m_1	m_2	m_3	m_4	m_5	m_6
7	5.8571	11.6830	13.2798	13.4486	13.5310

Change in normalizing factor m_i is now negligible
LARGEST Eigenvalue is $m_4 = 13.4486$ and
Corresponding Eigen vector $u_4 = [-0.2655, -1.000, 0.469, -0.0619]^T$

Problem on Rayleigh's method

We have Rayleigh's coefficient $\lambda_{\max} = \dfrac{X^T A X}{X^T X}$

$$= \frac{\begin{bmatrix} -0.2655 & -1.0000 & 0.4696 & -0.0619 \end{bmatrix} \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix} \begin{bmatrix} -0.2655 \\ -1.0000 \\ 0.4696 \\ -0.0619 \end{bmatrix}}{\begin{bmatrix} -0.2655 & -1.0000 & 0.4696 & -0.0619 \end{bmatrix} \begin{bmatrix} -0.2655 \\ -1.0000 \\ 0.4696 \\ -0.0619 \end{bmatrix}}$$

$$\lambda_{\max} = \frac{16.98979478}{1.29484602} = 13.12109279217617$$

Illustration Gerschgorin's result for a few cases

Theorem : Every eigenvalue of matrix A satisfies

$$|\lambda - A_{ii}| \le \sum_{j \ne i} |A_{ij}|, \; i, j \in \{1,2,3..n\}$$

Example: For illustration $A = \begin{bmatrix} 5 & 0 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ -1.5 & 1 & -2 & 1 \\ -1 & 1 & 3 & -3 \end{bmatrix}$

Row-1 $|\lambda - 5| \le 1$

Row-2 $|\lambda - 0| \le 3$

Row-3 $|\lambda + 2| \le 3.5$

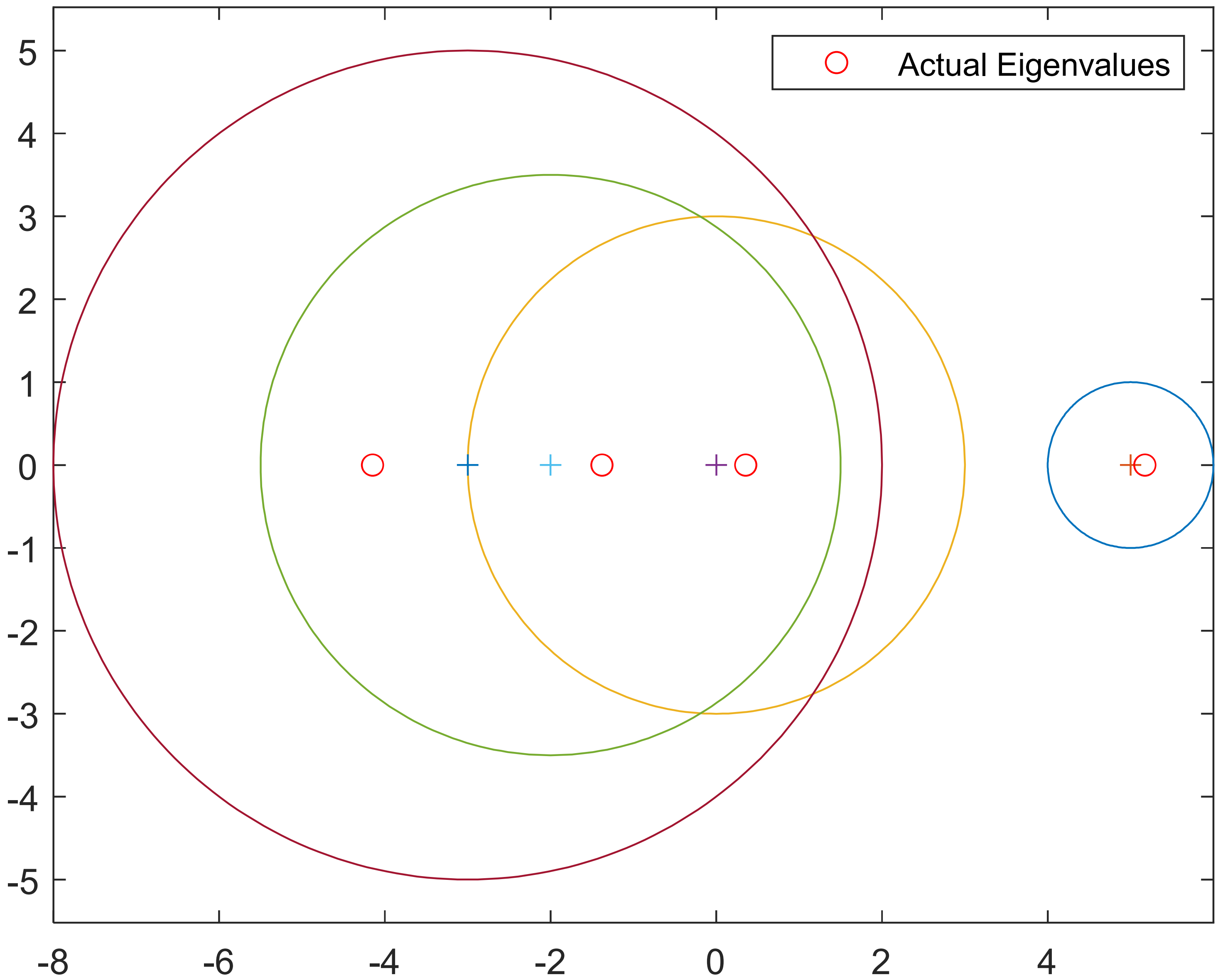
Row-3 $|\lambda + 3| \le 5$

Col-1 $|\lambda - 5| \le 3.5$

Col-2 $|\lambda - 0| \le 2$

Col-3 $|\lambda + 2| \le 4$

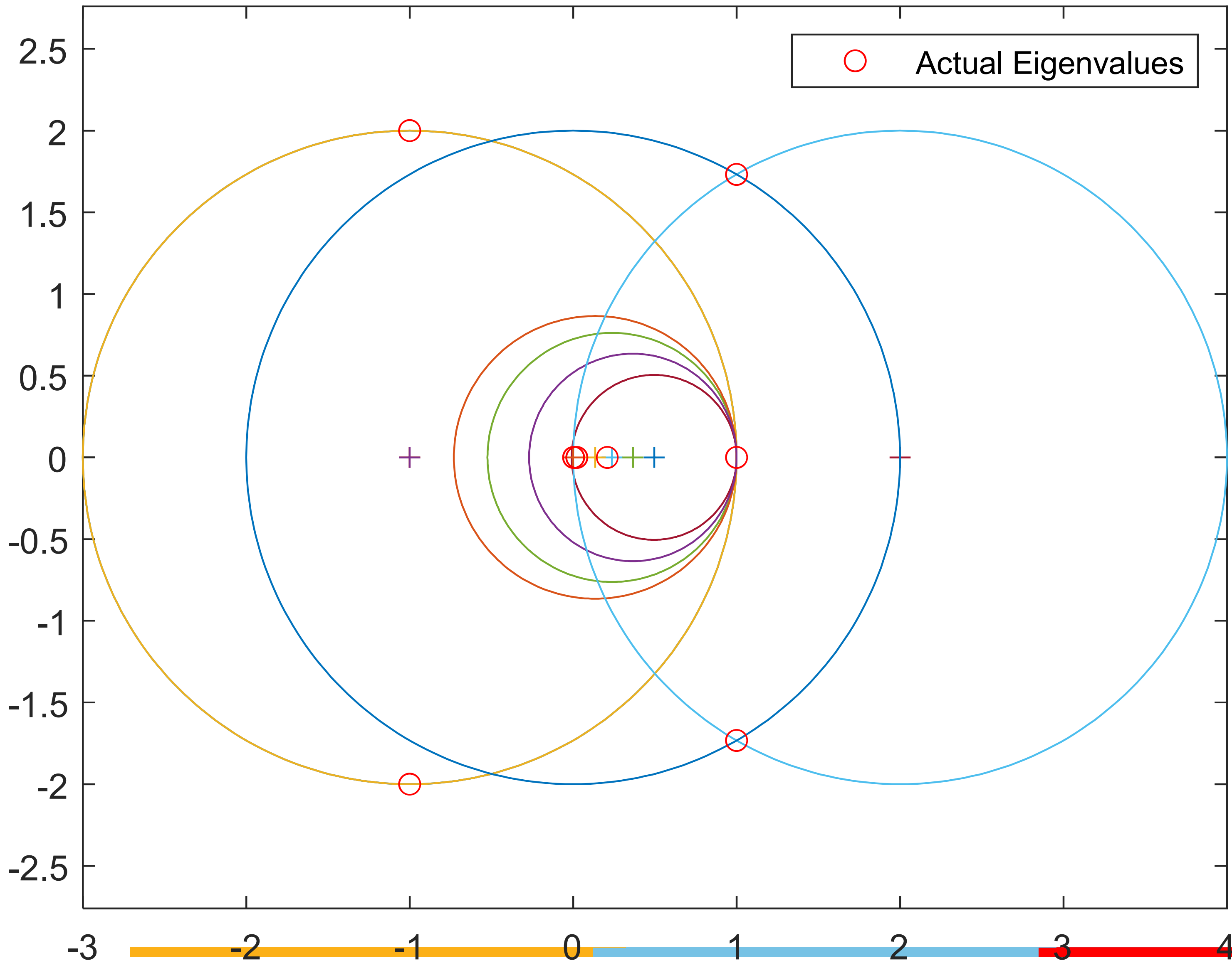
Col-4 $|\lambda + 3| \le 3$



Actual Eigenvalues
5
4
3
2
1
0
-1
-2
-3
-4
-5
-8
-6
-4
-2
0
2
4

Example-2

$$\begin{bmatrix} -1.0000 & 2.0000 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.0000 & -1.0000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2379 & 0.5145 & 0.1201 & 0.1275 & 0 & 0 \\ 0 & 0 & 0.1943 & 0.4954 & 0.1230 & 0.1873 & 0 & 0 \\ 0 & 0 & 0.1827 & 0.4955 & 0.1350 & 0.1868 & 0 & 0 \\ 0 & 0 & 0.1084 & 0.4218 & 0.1045 & 0.3653 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.0000 & 2.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2.0000 & 0 \end{bmatrix}$$



Actual Eigenvalues
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-2
-1
0
1
2
3
4

Thank You