EIGENVALUE PROBLEM

When multiplied with A, almost all vectors change directions. Certain exceptional vectors $\vec{x}$ are in the same direction as $A\vec{x}$; those are the eigenvectors.
If A is multiplied by an eigenvector $\vec{x}$, the vector $A\vec{x}$ is λ times the original $\vec{x}$.

$$A\vec{x} = \lambda\vec{x} \qquad \text{—— (1)}$$

The number λ is called the eigenvalue. It tells whether the special vector $\vec{x}$ is streched or shrunk or reversed or left unchanged - when it is multiplied by A. The eigenvalue could also be zero! Then $A\vec{x} = 0\vec{x}$ means that this eigenvector is in the nullspace. $\Rightarrow$ A is singular.

* ~~Z~~ Zero vector is never an eigenvector.

An application of eigenvalues & eigenvectors is to compute the powers of a matrix because e.g.,

$$A^2\vec{x} = A(A\vec{x}) = A(\lambda\vec{x}) = \lambda(A\vec{x}) = \lambda(\lambda\vec{x}) = \lambda^2\vec{x}$$

when matrix is squared, its eigenvalue is squared while the eigenvector remains the same as that for A.
This fact can be applied to ~~obtain~~ determine powers of a matrix, as we shall see.

Ex. Permutation matrix: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ symmetric matrix

guess the eigenvector and eigenvalues!

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\lambda = 1$

$\begin{bmatrix} -1 \\ +1 \end{bmatrix}$ $\lambda = -1$

Real eigenvalues; orthogonal eigenvectors.

To compute eigenvalues and eigenvectors, we start with

$A\vec{x} = \lambda \vec{x}$ (called the eigenvalue problem)

$\Rightarrow (A - \lambda I)\vec{x} = 0$

For nonzero sol^n (i.e. for (nonzero) eigenvector) $\vec{x}$,
$\vec{x}$ <u>should</u> be in the null space of $(A - \lambda I)$
or $A - \lambda I$ <u>should</u> be singular.
or $\det(A - \lambda I) = 0$ → root of characteristic equations are the eigenvalues
for each eigenvalue λ, solve $(A-\lambda I)\vec{x} = 0$ to find an eigenvector.

Characteristic equation.

Later
* <u>Caley-Hamilton theorem</u>:- Every square matrix satisfies its characteristic polynomial.

<u>Bad news</u>: Elimination (row exchange, row additions) does not preserve the eigenvalues. The upper triangular matrix U has its eigenvalues sitting along the diagonal - they are pivots. But they are not the eigenvalues of A!

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \lambda_A = 0, 2 \quad \xrightarrow{R_2 \to R_2 - R_1} \quad \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}}_{U} \quad \lambda_U = 0, 1.$$

<u>Good news</u>:-
* The product of the n eigenvalues equals the determinant of the matrix.
* The sum of the n eigenvalues equals sum of the n diagonal entries of A. This sum along the main diagonal is called the trace of A:

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}.$$

A matrix may have repeated eigenvalues but it may or may not have the full set of linearly **independent** eigenvectors.

E.g.

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

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 1-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (-1-\lambda)\left[(1-\lambda)^2 - 4\right] = 0 \quad \Rightarrow \lambda = -1 \text{ or } 1-\lambda = \pm 2$$

$$\Rightarrow \lambda = -1, -1, 3$$

e.vector for $\lambda = -1$

$$(A-\lambda I)\vec{x} = \begin{bmatrix} +2 & 2 & 0 \\ 2 & +2 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

two linearly independent e.vectors.

e.vector for $\lambda = 3$

$$(A-\lambda I)\vec{x} = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} +1 \\ 1 \\ 0 \end{bmatrix}$$

is an eigenvector

The above are three independent eigenvector.

Diagonalization:- Suppose the $n \times n$ matrix A has n linearly independent eigenvectors $x_1, x_2, \dots, x_n$. Put them into ~~the~~ the columns of an eigenvector matrix S. Then $S^{-1}AS$ is the eigenvalue matrix Λ.

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Proof:

$$AS = A\begin{bmatrix} \vec{x}_1 & \cdots & \vec{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{x}_1 & \cdots & \lambda_n\vec{x}_n \end{bmatrix} = \begin{bmatrix} \vec{x}_1 & \cdots & \vec{x}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & O \\ & \lambda_2 & \\ O & & \ddots \ \lambda_n \end{bmatrix}.$$

$$\boxed{AS = S\Lambda} \Rightarrow \boxed{S^{-1}AS = \Lambda} \text{ or } \boxed{A = S\Lambda S^{-1}}$$

↑ another facorization.

* Without n independent eigenvectors, ~~A~~ cannot ~~be~~ digonalized.

$$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2 S^{-1}$$

$$\vdots$$

$$A^k = S\Lambda^k S^{-1}.$$

What will be A^{100} of the matrix

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

e. values are $-1, 0$.

$$(A-(-1)I)\vec{x} = \begin{bmatrix} 0 & 0 \\ 1 & +1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 \\ 1 \end{bmatrix} = X_1$$

$$(A - 0I)\vec{x} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \quad X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$S = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} \qquad S^{-1} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$A^{100} = S\Lambda^{100}S^{-1} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (-1)^{100} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}. \quad \square$$

→ A $n \times n$ matrix B is called Similar to a $n \times n$ matrix A if there is a non-singular $n \times n$ matrix T such that

$$B = T^{-1} A T.$$

→ Similar matrices have the same
(a) Characteristic equation and eigenvalues
(b) determinant and invertibility
(c) trace
(d) rank and nullity.

Note that two matrices may have equal eigenvalues but not similar. E.g.,
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Clearly, $\quad P_A(x) = P_B(x) = 1 - 2x + x^2 = (x-1)^2,$

So A and B have equal characteristic polynomials.
Thus, A and B have equal eigenvalues $\lambda_1 = 1$, $\lambda_2 = 1$.
Conversly, Suppose there is a non-singular matrix S so that $A = S^{-1} B S$. Then

$$A = S^{-1} B S = S^{-1} I_2 S = S^{-1} S = I_2$$

Clearly, $A \neq I_2$, and contradiction tells us that A and B are not similar matrices but have equal eigenvalues.

POWER METHOD:

Dominant Eigenvalue and Dominant Eigenvector :-

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix A. λ_1 is called the dominant eigenvalue of A if
$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \geq 0.$$

The eigenvector corresponding to the dominant eigenvalue λ_1 is called the dominant eigenvector of A.

* Not every matrix has a dominant eigenvalue!

E.g.,
$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (\text{with eigenvalues } \lambda_1 = 1, \lambda_2 = -1)$$
has no dominant eigenvalue.

or
$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (\text{with eigenvalue } \lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 1)$$
has no dominant eigenvalue.

Assumptions for Power Method :-

1. A has a dominant eigenvalue with a corresponding eigenvector.
2. A has n linearly independent eigenvectors.

Note that If A does not have n linearly independent eigenvectors, the power method may still be successful but it is not guaranteed.

Procedure :-

→ Choose an initial approximation $\vec{x}_0 \neq 0$ for a dominant eigen vector.

→ Proceed as follows:

$$A\vec{x}_0 = c_1 \vec{x}_1 \quad (\text{say})$$
$$A\vec{x}_1 = c_2 \vec{x}_2$$
$$A\vec{x}_2 = c_3 \vec{x}_3$$
$$\vdots \qquad \vdots$$
$$A\vec{x}_K = c_{K+1} \vec{x}_K$$

For large value of K and by properly scaling this sequence, we shall see that we obtain a good approximation of the dominant eigenvector of A.

Ex: Complete five iterations of the power method to approximate a dominant eigenvector of $A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$.

Solⁿ We begin with a non-zero approximation

$$\vec{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$A\vec{x}_0 = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -10 \\ -4 \end{bmatrix} = \underbrace{-10}_{c_1} \underbrace{\begin{bmatrix} 1 \\ 0.4 \end{bmatrix}}_{\vec{x}_1}$$

$$A\vec{x}_1 = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0.4 \end{bmatrix} = \begin{bmatrix} -2.8 \\ -1 \end{bmatrix} = \underbrace{-2.8}_{c_2} \underbrace{\begin{bmatrix} 1 \\ 0.357 \end{bmatrix}}_{\vec{x}_2}$$

$$A\vec{x}_2 = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0.357 \end{bmatrix} = \begin{bmatrix} -2.28 \\ -0.785 \end{bmatrix} = -2.28 \underbrace{\begin{bmatrix} 1 \\ 0.344 \end{bmatrix}}_{\vec{x}_3}$$

$$A\vec{x}_3 = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0.344 \end{bmatrix} = \begin{bmatrix} -2.13 \\ -0.72 \end{bmatrix} = -2.13 \underbrace{\begin{bmatrix} 1 \\ 0.338 \end{bmatrix}}_{\vec{x}_4}$$

$$A\vec{x}_4 = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0.338 \end{bmatrix} = \begin{bmatrix} -2.06 \\ -0.69 \end{bmatrix} = -2.06 \underbrace{\begin{bmatrix} 1 \\ 0.\cancel{335} \to 0.335 \end{bmatrix}}_{\vec{x}_5}$$

The approximations seems to approach to $\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$, which is actually a dominant eigenvector of A corresponding to eigenvalue $\lambda = -2$. #

Suppose we do not know the dominant eigenvalue of A. The following theorem provides a formula for determining the eigenvalue corresponding to a given eigenvector.

Theorem:- If $\vec{x}$ is an eigenvector of A, then its corresponding eigenvalue is given by

$$\lambda = \frac{A\vec{x}\cdot\vec{x}}{\vec{x}\cdot\vec{x}} = \frac{(A\vec{x})^T\vec{x}}{\vec{x}^T\vec{x}}.$$

The quotient is called the Rayleigh quotient.

Proof. Since $\vec{x}$ is an eigenvector of A, we know that

$$A\vec{x} = \lambda\vec{x} \Rightarrow A\vec{x}\cdot\vec{x} = \lambda\vec{x}\cdot\vec{x}$$

$$\Rightarrow \lambda = \frac{A\vec{x}\cdot\vec{x}}{\vec{x}\cdot\vec{x}}. \quad \#$$

Dominant eigenvalue in the previous example:

$$\vec{x}_5 = \begin{bmatrix} 1 \\ 0.335 \end{bmatrix}$$

$$\lambda = \frac{A\vec{x}\cdot\vec{x}}{\vec{x}\cdot\vec{x}} \simeq \frac{A\vec{x}_5\cdot\vec{x}_5}{\vec{x}_5\cdot\vec{x}_5} = \frac{-2.25}{1.11} = -2.03$$

which is a good approximation of the dominant eigenvalue $\lambda = -2$.

Why does the power method work?

Let A has n <u>independent</u> eigenvectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ and corresponding ~~to~~ the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and let λ_1 be the dominant eigenvalue, i.e.,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|.$$

The intial approximation $\vec{x}_0$ can be written as the linear combination of the eigenvectors:

$$\vec{x}_0 = \beta_1 \vec{x}_1 + \beta_2 \vec{x}_2 + \cdots + \beta_n \vec{x}_n$$

(If $\beta_1 = 0$, the power method may not converge and one would need to choose a different initial guess.)

$$A\vec{x}_0 = \beta_1 (A\vec{x}_1) + \beta_2 (A\vec{x}_2) + \cdots + \beta_n (A\vec{x}_n)$$
$$= \beta_1 \lambda_1 \vec{x}_1 + \beta_2 \lambda_2 \vec{x}_2 + \cdots + \beta_n \lambda_n \vec{x}_n$$

$$A^2\vec{x}_0 = \beta_1 \lambda_1 (A\vec{x}_1) + \beta_2 \lambda_2 (A\vec{x}_2) + \cdots + \beta_n \lambda_n (A\vec{x}_n)$$
$$= \beta_1 \lambda_1^2 \vec{x}_1 + \beta_2 \lambda_2^2 \vec{x}_2 + \cdots + \beta_n \lambda_n^2 \vec{x}_n$$

$\vdots$

$$A^k \vec{x}_0 = \beta_1 \lambda_1^k \vec{x}_1 + \beta_2 \lambda_2^k \vec{x}_2 + \cdots + \beta_n \lambda_n^k \vec{x}_n$$
$$= \lambda_1^k \left[\beta_1 \vec{x}_1 + \beta_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{x}_2 + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{x}_k \right]$$

each of $\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_1}, \ldots, \frac{\lambda_n}{\lambda_1}$ is less than 1 is absolute value.

$\cancel{\Rightarrow \lim_{k\to\infty} A^k \vec{x}_0 = \lim_{k\to\infty} \lambda_1^k \beta \vec{x}}$ These fraction approach to zero as k approaches infinity.

This implies that

$$A^k \vec{x}_0 \approx \lambda_1^k \beta_1 \vec{x}_1$$

improves as k increases.

Since $\vec{x}_1$ is the dominant eigenvector, it follows that any scalar multiple of $\vec{x}_1$ is also a dominat eigenvector.

Thus we have shown that $A^k \vec{x}_0$ approaches a multiple of the dominant eigenvector of A. □

* The above proof provides some insight into the rate of convergence of the power method. That is if the eigenvalues of A are ordered so that

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

then the power method will converge quickly if $|\lambda_2|/|\lambda_1|$ is small, and slowly if $|\lambda_2|/|\lambda_1|$ is close to 1.

<u>Ex.</u> $A = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}$ has eigenvalues $\lambda_1 = 10$ and $\lambda_2 = -1$

$|\lambda_2|/|\lambda_1| = 0.1$. For this matrix only 4 iterations are required to obtain successive approximations that agree when rounded to three significant digits.

$x_0 = \begin{bmatrix} 1.000 \\ 1.000 \end{bmatrix}$, $x_1 = \begin{bmatrix} 0.818 \\ 1.000 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0.835 \\ 1.000 \end{bmatrix}$, $x_3 = \begin{bmatrix} 0.833 \\ 1.000 \end{bmatrix}$, $x_4 = \begin{bmatrix} 0.8333 \\ 1.000 \end{bmatrix}$

②

$A = \begin{bmatrix} -4 & 10 \\ 7 & 5 \end{bmatrix}$ has eigenvalues $\lambda_1 = 10$, $\lambda_2 = -9$,

the ratio $|\lambda_2|/|\lambda_1| = 0.9$. The power method does not produce successive approximations that agree to three significant digits until 68 iterations have been performed.

x_0	x_1	x_2		x_{66}	x_{67}	x_{68}
$\begin{bmatrix} 1.000 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.941 \\ 1.000 \end{bmatrix}$	$\cdots$	$\begin{bmatrix} 0.715 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.714 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} 0.714 \\ 1.000 \end{bmatrix}$.

Inverse power method

Lemma If $(\lambda_i, \vec{x}_i)$ are the eigenvalue/eigenvector pairs of A, $1 \le i \le n$, then $(A - \sigma I)^{-1}$ has eigenvalue/eigenvector pairs $\left(\frac{1}{\lambda_i - \sigma}, \vec{x}_i\right)$.

Let σ is close (but not equal to) to some eigenvalue λ_1. Define

$$\mu_1 = \frac{1}{\lambda_1 - \sigma}, \quad \mu_2 = \frac{1}{\lambda_2 - \sigma}, \quad \cdots, \quad \mu_n = \frac{1}{\lambda_n - \sigma}$$

the eigenvalues of $(A - \sigma I)^{-1}$ and ordered in such a way that $|\mu_1| > |\mu_2| \geq |\mu_3| \geq \cdots \geq |\mu_n|$.

The inverse power method is simply the power method applied to $(A - \sigma I)^{-1}$.

(Do we really need to compute inverses?)

If $\vec{x}_0$ is the initial guess vector, chosen appropriately, then the sequences $\vec{Y}_k = (A - \sigma I)^{-1} \vec{X}_k$,

$$\vec{X}_{k+1} = \frac{\vec{Y}_k}{\| \vec{Y}_k \|}$$

and $c_{k+1} = \dfrac{\vec{Y}_k^T \vec{X}_k}{\vec{X}_k^T \vec{X}_k}$ (Rayleigh quotient) will converge to the dominant

Ex. $A = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}$ has eigenvalues $\lambda_1 = 10$ and $\lambda_2 = -1$

Power method for the dominant eigenvalue.

$$\vec{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vec{x}_1 = A\vec{x}_0 = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 11 \end{bmatrix} \longrightarrow \vec{x}_1 = \frac{\vec{x}_1}{\|\vec{x}_1\|_\infty} = \begin{bmatrix} 0.8182 \\ 1 \end{bmatrix}$$

$$\vec{x}_2 = A\vec{x}_1 = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}\begin{bmatrix} 0.8182 \\ 1 \end{bmatrix} = \begin{bmatrix} 8.2727 \\ 9.9091 \end{bmatrix} \longrightarrow \vec{x}_2 = \frac{\vec{x}_2}{\|\vec{x}_2\|_\infty} = \begin{bmatrix} 0.8349 \\ 1 \end{bmatrix}$$

$$\vec{x}_3 = A\vec{x}_2 = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}\begin{bmatrix} 0.8349 \\ 1 \end{bmatrix} = \begin{bmatrix} 8.33945 \\ 10.0092 \end{bmatrix} \longrightarrow \vec{x}_3 = \frac{\vec{x}_3}{\|\vec{x}_3\|_\infty} = \begin{bmatrix} 0.8332 \\ 1 \end{bmatrix}$$

$$\vec{x}_4 = A\vec{x}_3 = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}\begin{bmatrix} 0.8332 \\ 1 \end{bmatrix} = \begin{bmatrix} 8.3327 \\ 9.9991 \end{bmatrix} \longrightarrow \vec{x}_4 = \frac{\vec{x}_4}{\|\vec{x}_4\|_\infty} = \begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}$$

$$\vec{x}_5 = A\vec{x}_4 = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix}\begin{bmatrix} 0.8333 \\ 1 \end{bmatrix} = \begin{bmatrix} 8.3334 \\ 10.0001 \end{bmatrix} \longrightarrow \vec{x}_5 = \frac{\vec{x}_5}{\|\vec{x}_5\|_\infty} = \begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}$$

eigenvector is $\vec{x} = \begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}$

and associated eigenvalue is $\dfrac{(A\vec{x}) \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{\begin{bmatrix} 8.3333 \\ 10.0000 \end{bmatrix} \cdot \begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}\begin{bmatrix} 0.8333 \\ 1 \end{bmatrix}}$

$$= 10.0000$$

The other e.value is -1.

Let $\sigma = -0.5$

$$\mu_1 = \frac{1}{\lambda_1 - (-0.5)}, \quad \mu_2 = \frac{1}{\lambda_2 - (-0.5)} \text{ are the eigenvalues}$$

of $(A - \sigma I)^{-1}$.

$$B = A - \sigma I = \begin{bmatrix} 4 & 5 \\ 6 & 5 \end{bmatrix} - \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix} = \begin{bmatrix} 4.5 & 5 \\ 6 & 5.5 \end{bmatrix}$$

Let $\vec{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\vec{Y}_0 = (A - \sigma I)^{-1} \vec{x}_0 = B^{-1} \vec{x}_0$$

$$\Rightarrow (A - \sigma I)\vec{Y}_0 = \vec{X}_0 \Rightarrow B \vec{Y}_0 = \vec{X}_0$$

$$\begin{bmatrix} 4.5 & 5 \\ 6 & 5.5 \end{bmatrix} \vec{Y}_0 = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\vec{X}_0} \xrightarrow[\text{system}]{\text{Solve linear}} \cancel{\vec{Y}_0 = \begin{bmatrix} 71.8182 \\ 17.2727 \end{bmatrix} \longrightarrow \vec{X}_1 = \frac{\vec{Y}_0}{\|\vec{Y}_0\|_\infty} = \begin{bmatrix} 1 \\ 0.2405 \end{bmatrix}}$$

$$\vec{Y}_0 = \begin{bmatrix} -0.0952 \\ 0.2857 \end{bmatrix} \longrightarrow \vec{x}_1 = \frac{\vec{Y}_0}{\|\vec{Y}_0\|_\infty} = \begin{bmatrix} -0.3333 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 4.5 & 5 \\ 6 & 5.5 \end{bmatrix} \vec{Y}_1 = \underbrace{\begin{bmatrix} -0.3333 \\ 1 \end{bmatrix}}_{\vec{x}_1} \longrightarrow \vec{Y}_1 = \begin{bmatrix} 1.3016 \\ -1.2381 \end{bmatrix} \longrightarrow \vec{x}_2 = \frac{\vec{Y}_1}{\|\vec{Y}_1\|_\infty} = \begin{bmatrix} 1 \\ -0.9512 \end{bmatrix}$$

$$\vec{Y}_2 = \begin{bmatrix} -1.95354 \\ 1.9582 \end{bmatrix} \longrightarrow \vec{x}_3 = \frac{\vec{Y}_2}{\|\vec{Y}_2\|_\infty} = \begin{bmatrix} -0.9976 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 4.5 & 5 \\ 6 & 5.5 \end{bmatrix} \vec{Y}_2 = \underbrace{\begin{bmatrix} +1.0000 \\ -0.9512 \end{bmatrix}}_{\vec{X}_2} \longrightarrow \vec{Y}_2$$

$$\vec{Y}_3 = \begin{bmatrix} 1.99751 \\ -1.99729 \end{bmatrix} \longrightarrow \vec{x}_4 = \frac{\vec{Y}_3}{\|\vec{Y}_3\|_\infty} = \begin{bmatrix} 1 \\ -0.9999 \end{bmatrix}$$

$$\begin{bmatrix} 4.5 & 5 \\ 6 & 5.5 \end{bmatrix} \vec{Y}_3 = \underbrace{\begin{bmatrix} -0.9976 \\ 1 \end{bmatrix}}_{\vec{x}_3} \longrightarrow \vec{Y}_3$$

$\vdots$

$$\vec{Y}_4 = \begin{bmatrix} -1.9999 \\ 1.9999 \end{bmatrix} \longrightarrow \vec{x}_5 = \frac{\vec{Y}_4}{\|\vec{Y}_4\|} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\vec{Y}_5 = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \longrightarrow \vec{x}_6 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mu = \frac{\vec{Y}_5 \cdot \vec{x}_5}{\vec{x}_5 \cdot \vec{x}_5} = \frac{\begin{bmatrix} 2 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}}{\begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \frac{-4}{2} = -2$$

$$\lambda = \frac{1}{\mu} + \sigma = \frac{1}{(-2)} + (-0.5) = -1.$$

Geršgorin Discs :-

Let A be a $n \times n$ matrix and R_i denotes the discs in the complex plane with center a_{ii} and radius $\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$, i.e.,

$$D_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right\}, \quad i = 1, 2, \ldots, n$$

Geršgorin discs

Geršgorin Circle Theorem (1931):

Every eigenvalue of a $n \times n$ (real or complex) matrix A lies within at least one of the Geršgorin discs D_i.

Proof:- Suppose λ is an eigenvalue of the matrix A and its associated eigenvector is $\vec{x}$.

We have $A\vec{x} = \lambda \vec{x}$, —— ① $\vec{x} \neq 0$.

Let x_i be the component of $\vec{x}$ having the largest absolute value. Then comparison of i^{th} row on both sides of ① gives

$$\sum_{j=1}^{n} a_{ij} x_j = \lambda x_i$$

$$\Rightarrow (\lambda - a_{ii}) x_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} x_j$$

$$\Rightarrow \lambda - a_{ii} = \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} \frac{x_j}{x_i}$$

$$\Rightarrow |\lambda - a_{ii}| = \left| \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} \left| a_{ij} \frac{x_j}{x_i} \right| = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \left| \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

since $|x_j| \leq |x_i|$

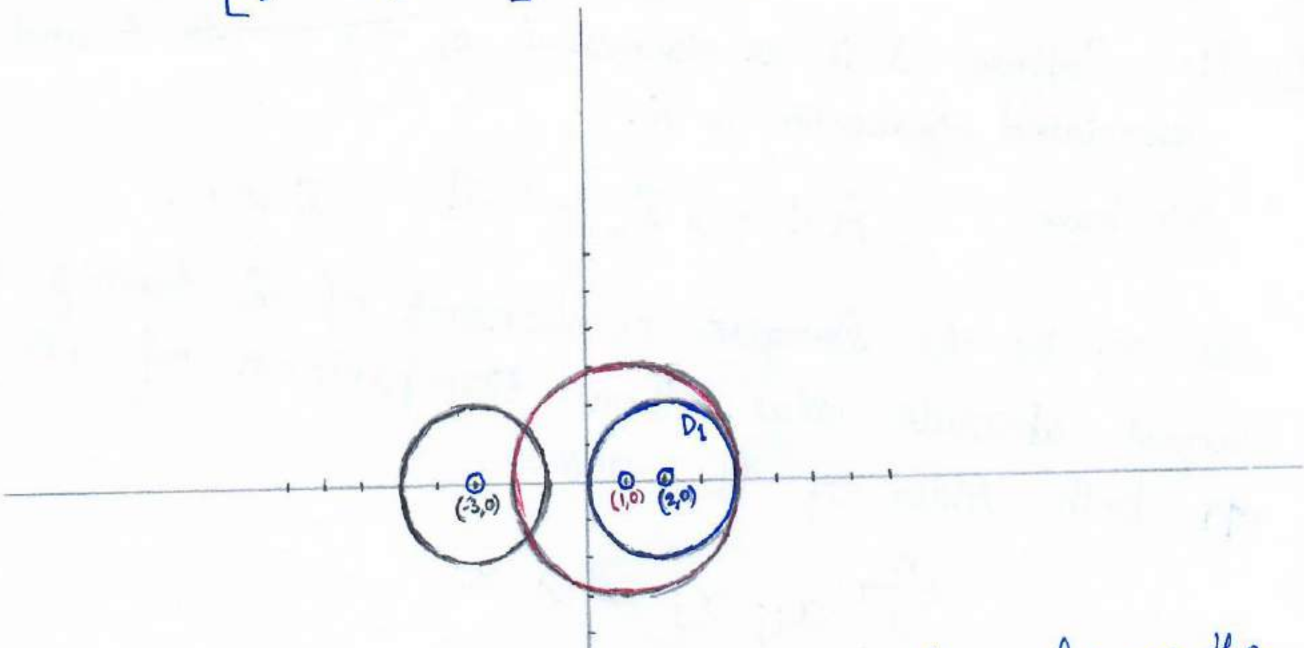
$$\Rightarrow \lambda \in D_i.$$

□

Example:-

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 1 & 1 & 2 \\ 1 & -1 & -3 \end{bmatrix} \quad \begin{aligned} r_1 &= 0+2=2 \\ r_2 &= 1+2=3 \\ r_3 &= 1+|-1|=2 \end{aligned}$$

Actual eigenvalues
$-3, 2, 1$



* It is not necessary that the eigenvalues lie on the centers of the discs.

* Does every disk have its own eigenvalue? NO

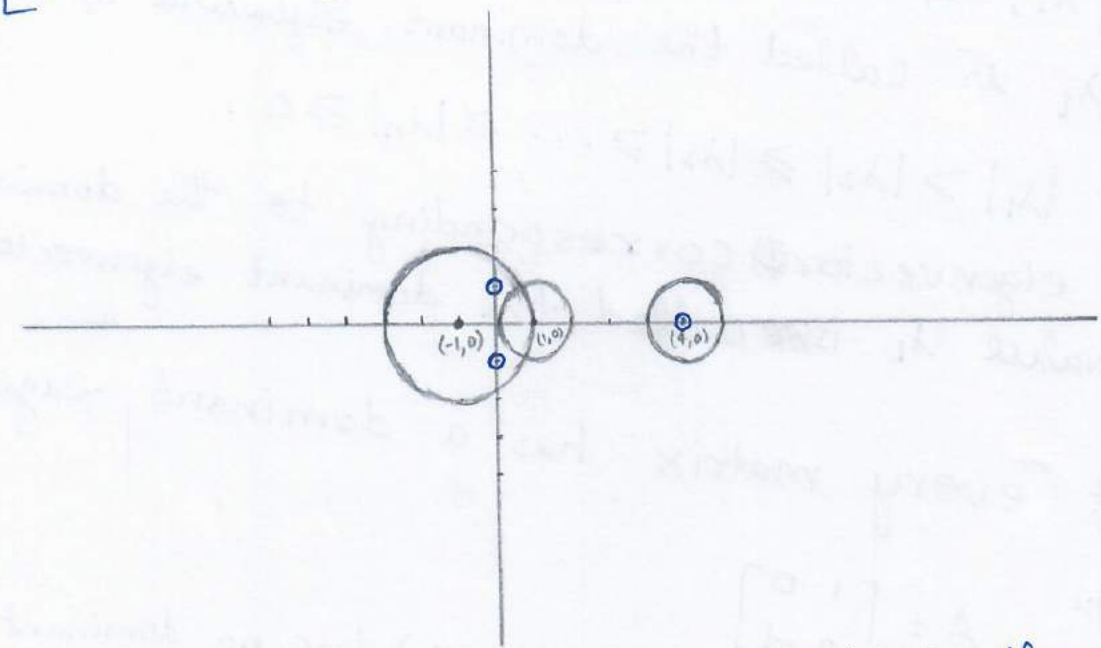
Theorem (Extension of the Geršgorin theorem)
Let A be a $n \times n$ (real or complex) matrix. If $\cancel{D_1, D_2, \ldots, D_k}$ $D_{i_1}, D_{i_2}, \ldots, D_{i_k}$ are Geršgorin discs of A that are disjoint from the remaining $(n-k)$ discs, then their union contains exactly k eigenvalues of A (counting multiplicities).

(Proof skipped)

Example:-

$$\begin{bmatrix} 1 & -1 & 0 \\ 2 & -1 & 0 \\ 0 & 1 & 4 \end{bmatrix} \quad \begin{aligned} r_1 &= |-1| + |0| = 1 \\ r_2 &= |+2| + |0| = 2 \\ r_3 &= |0| + |1| = 1 \end{aligned}$$

Actual eigenvalues $\{4, i, -i\}$



* In the case, where all discs are disjoint, the matrix is diagonalizable. Why?
  ⇒ Eigenvalues are distinct. We have full set of independent eigenvectors.

Remarks ① A and A^T have same eigenvalues, therefore we could have used the columns of A rather than the rows to make our Geršgorin discs.

② Since similar matrices have ~~the~~ same eigenvalues, we could have used the discs from SAS^{-1} to approximate the eigenvalues of A.

Example: Determine the Gersgorin circles for the matrix

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix},$$

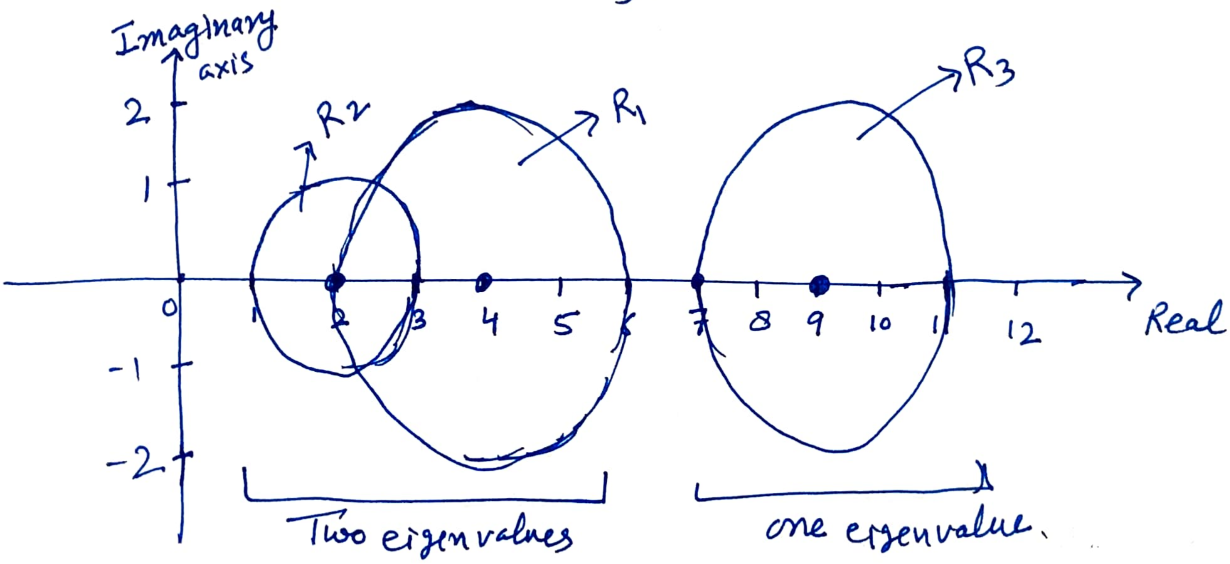
and use these to find the bounds for the spectral radius of A.

Solⁿ The circles in the Gersgorin theorem are

$$R_1 = \{z \in \mathbb{C} : |z-4| \le 2\},$$

$$R_2 = \{z \in \mathbb{C} : |z-2| \le 1\},$$

and $R_3 = \{z \in \mathbb{C} : |z-9| \le 2\}$.



Because R_1 and R_2 are disjoint from R_3, there are precisely two eigenvalues within $R_1 \cup R_2$ and one within R_3. Moreover, $\rho(A) = \max_{1 \le i \le 3} |\lambda_i|$, so $7 \le \rho(A) \le 11$.

Tridiagonal Matrices:

A tridiagonal matrix has non-zero elements on the main diagonal, the first diagonal below this, and the first diagonal above the main diagonal only.

For example, the following matrices are ~~tridg~~ tridiagonal

$$\begin{pmatrix} 1 & 4 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 5 & 3 & 0 \\ 0 & 3 & 7 & 2 \\ 0 & 0 & 2 & 5 \end{pmatrix}, \quad \text{etc.}$$

Householder's Tridiagonalization Method:

To reduce a given real symmetric $n \times n$ matrix A by $(n-2)$ successive similarity transformations to tridiagonal form.

Start with $A_0 = A$

$A_1 = P_1 A_0 P_1 \rightarrow$ creates necessary zeros in the first row and first column

$A_2 = P_2 A_1 P_2 \rightarrow$ " in the second row and 2nd column.

$\vdots$

$A_{n-2} = P_{n-2} A_{n-3} P_{n-2}$

All P_i $(i = 1, 2, \cdots n-2)$ are orthogonal and symmetric matrices (i.e., $P_i^{-1} = P_i$ and $P_i^T = P_i$). This implies that A_0 is similar to all other A_i's. Hence, A_{n-2} and A_0 have the same eigenvalues.

<u>To determine P_i's:</u> we use the Householder transformation as

$$P_i = I - 2\, \underline{v}_i\, \underline{v}_i^T, \quad (i = 1, 2, \cdots, n-2),$$

where $\underline{v}_i$ is a unit vector with its first 'i' components zero; thus

$$\underline{v}_1 = \begin{bmatrix} 0 \\ - \\ - \\ - \\ - \end{bmatrix}, \quad \underline{v}_2 = \begin{bmatrix} 0 \\ 0 \\ - \\ - \\ - \end{bmatrix}, \quad \cdots \quad \underline{v}_{n-2} = \begin{bmatrix} \left.\begin{matrix} 0 \\ 0 \\ \vdots \\ 0 \end{matrix}\right\} n-2 \\ - \\ - \end{bmatrix}$$

<u>For 5×5 matrix:</u>

$$\begin{bmatrix} - & - & & & \\ - & - & - & - & - \\ & - & - & - & - \\ & - & - & - & - \\ & - & - & - & - \end{bmatrix} \qquad \begin{bmatrix} - & - & & & \\ - & - & - & & \\ & - & - & - & - \\ & & - & - & - \\ & & - & - & - \end{bmatrix} \qquad \begin{bmatrix} - & - & & & \\ - & - & - & & \\ & - & - & - & \\ & & - & - & - \\ & & & - & - \end{bmatrix}$$

1st step $\qquad$ 2nd step $\qquad$ 3rd step.

$$A_1 = P_1 A_0 P_1 \qquad A_2 = P_2 A_1 P_2 \qquad A_3 = P_3 A_2 P_3$$

Step-1: Computing $\underline{v_1} = [v_{11}, v_{21}, v_{31}, \ldots v_{n1}]^T$.

$$v_{11} = 0,$$

$$v_{21} = \sqrt{\frac{1}{2}\left(1 + \frac{|a_{21}|}{S_1}\right)},$$

$$v_{j1} = \frac{a_{j1}\, \mathrm{sgn}(a_{21})}{2 v_{21} S_1}, \quad j = 3, 4, 5, \ldots, n,$$

where

$$S_1 = \sqrt{a_{21}^2 + a_{31}^2 + \cdots + a_{n1}^2} > 0$$

and
$$\mathrm{sgn}(a_{21}) = \begin{cases} +1 & \text{if } a_{21} \geq 0 \\ -1 & \text{if } a_{21} < 0. \end{cases}$$

With this, we compute $P_1 = I - 2\underline{v_1}\,\underline{v_1}^T$, and then

$$A_1 = P_1 A P_1 = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & & & \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}.$$

Step-2: Computing $\underline{v_2} = [v_{12}, v_{22}, v_{32}, \ldots, v_{n2}]^T$.

$$v_{12} = v_{22} = 0,$$

$$v_{32} = \sqrt{\frac{1}{2}\left(1 + \frac{|a_{32}^{(1)}|}{S_2}\right)},$$

$$V_{j2} = \frac{a_{j2}^{(1)} \, \operatorname{sgn}(a_{32}^{(1)})}{2 V_{32} S_2}, \quad j = 4, 5, \ldots, n,$$

where

$$S_2 = \sqrt{(a_{32}^{(1)})^2 + (a_{42}^{(1)})^2 + \cdots + (a_{n2}^{(1)})^2}.$$

With this, we compute $P_2 = I - 2 \underline{V_2}\, \underline{V_2}^T$, and then

$$A_2 = P_2 A_1 P_2 = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & & & \\ a_{n1}^{(2)} & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}.$$

<u>Step-3:</u> Computing $\underline{V_3} = [V_{13}, V_{23}, V_{33}, \ldots, V_{n3}]^T$.

$$V_{13} = V_{23} = V_{33} = 0$$

$$V_{43} = \sqrt{\frac{1}{2}\left(1 + \frac{|a_{43}^{(2)}|}{S_3}\right)}$$

$$V_{j3} = \frac{a_{j3}^{(2)} \, \operatorname{sgn}(a_{j3}^{(2)})}{2 V_{43} S_3}, \quad j = 5, 6, \ldots, n,$$

and so on.

Example: To tridiagonalize the real symmetric matrix

$$A = \begin{bmatrix} 6 & 4 & 1 & 1 \\ 4 & 6 & 1 & 1 \\ 1 & 1 & 5 & 2 \\ 1 & 1 & 2 & 5 \end{bmatrix}.$$

Step-1:

$$S_1 = \sqrt{4^2+1^2+1^2} = \sqrt{18}$$

$$sgn(a_{21}) = sgn(4) = 1.$$

we compute $\underline{v_1} = [v_{11}, v_{21}, v_{31}, v_{41}]^T$,

$$v_{11} = 0$$

$$v_{21} = \sqrt{\frac{1}{2}\left(1+\frac{|a_{21}|}{S_1}\right)} = \sqrt{\frac{1}{2}\left(1+\frac{4}{\sqrt{18}}\right)} = 0.985599$$

$$v_{31} = \frac{a_{31}\, sgn(a_{21})}{2\, v_{21}\, S_1} = \frac{1\times 1}{2\times 0.985599\times\sqrt{18}} = 0.119573$$

$$v_{41} = \frac{a_{41}\, sgn(a_{21})}{2\, v_{21}\, S_1} = 0.119573$$

Thus,

$$\underline{v_1} = \begin{bmatrix} 0 \\ 0.985599 \\ 0.119573 \\ 0.119573 \end{bmatrix}.$$

$$P_1 = I - 2\underline{V_1}\,\underline{V_1}^T$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.942809 & -0.235702 & -0.235702 \\ 0 & -0.235702 & 0.971405 & -0.028595 \\ 0 & -0.235702 & -0.028595 & 0.971405 \end{bmatrix}$$

$$A_1 = P_1 A P_1 = \begin{bmatrix} 6 & -4.24264 & 0 & 0 \\ -4.24264 & 7 & -1 & -1 \\ 0 & -1 & 4.5 & 1.5 \\ 0 & -1 & 1.5 & 4.5 \end{bmatrix}.$$

Step-2:

$$S_2 = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}$$

$$sgn(a_{32}^{(1)}) = sgn(-1) = -1$$

$$\underline{V_2} = \begin{bmatrix} 0 \\ 0 \\ 0.923880 \\ 0.382683 \end{bmatrix} = \begin{bmatrix} V_{12} \\ V_{22} \\ V_{32} \\ V_{42} \end{bmatrix}.$$

$$P_2 = I - 2\underline{V_2}V_2^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -0.707107 & -0.707107 \\ 0 & 0 & -0.707107 & 0.707107 \end{bmatrix}$$

$$A_2 = P_2 A_1 P_2 = \begin{bmatrix} 6 & -4.24264 & 0 & 0 \\ -4.24264 & 7 & 1.41421 & 0 \\ 0 & 1.41421 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

This is tridiagonal. Since our given matrix has order $n=4$, we needed $n-2=2$ steps to accomplish this reduction. Note that the matrix A_2 is similar to the given matrix A.

Remark:

(1) The matrix P_r is symmetric.

$$\begin{aligned} P_r^T &= (I - 2\underline{V}_r \underline{V}_r^T)^T \\ &= I^T - 2(\underline{V}_r \underline{V}_r^T)^T \\ &= I - 2(\underline{V}_r^T)^T \underline{V}_r^T \\ &= I - 2\underline{V}_r \underline{V}_r^T \\ &= P_r. \end{aligned}$$

(2) The matrix P_r is also orthogonal because $\underline{V}_r$ is a unit vector, so that $\underline{V}_r^T \underline{V}_r = 1$, and thus

$$P_r P_r^T = P_r^2 \qquad (\text{Since } P_r^T = P_r \text{ as } (1)).$$

$$= (I - 2\underline{v_r}\,\underline{v_r}^T)^2$$

$$= I + 4(\underline{v_r}\,\underline{v_r}^T)(\underline{v_r}\,\underline{v_r}^T) - 4\underline{v_r}\,\underline{v_r}^T$$

$$= I + 4\underline{v_r}\underbrace{(\underline{v_r}^T\underline{v_r})}_{1}\underline{v_r}^T - 4\underline{v_r}\,\underline{v_r}^T$$

$$= I + 4\underline{v_r}\,\underline{v_r}^T - 4\underline{v_r}\,\underline{v_r}^T$$

$$= I$$

$$\Rightarrow \quad P_r^{-1} = P_r^T = P_r.$$

(3) Consider

$$B = P_{n-2} A_{n-3} P_{n-2}$$

$$= P_{n-2} \cdot P_{n-3} A_{n-4} P_{n-3} P_{n-2}$$

$$\vdots$$

$$= P_{n-2} P_{n-3} \cdots P_1 A_0 P_1 P_2, \cdots P_{n-3} P_{n-2}$$

$$= (P_{n-2}^{-1} P_{n-3}^{-1} \cdots P_1^{-1}) A_0 (P_1 P_2 \cdots P_{n-3} P_{n-2})$$

$$= Q^{-1} A Q$$

where $Q = P_1 P_2 \cdots P_{n-3} P_{n-2}$.

Thus, B and A are similar.

The QR method :- We begin with a symmetric matrix in tridiagonal form. If this is not the form, the first step is to apply Householder's method to compute a symmetric, tridiagonal matrix similar to the given matrix.

$$A = \begin{bmatrix} a_1 & b_2 & 0 & \cdots & 0 \\ b_2 & a_2 & b_3 & \ddots & \vdots \\ 0 & b_3 & a_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & 0 & b_n & a_n \end{bmatrix}$$

If $b_2 = 0$ or $b_n = 0$ then we immediately a_1 or a_n as an eigenvalue of A. The QR method takes advantage of this observation by successively decreasing the values of the entries below the main diagonal until $b_2 \approx 0$ or $b_n \approx 0$.

When $b_j = 0$ for some j. where $2 < j < n$, the problem can be reduced to the smaller matrices, instead of A,

$$\begin{bmatrix} a_1 & b_2 & 0 & \cdots & 0 \\ b_2 & a_2 & b_3 & \ddots & \vdots \\ 0 & b_3 & a_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_{j-1} \\ 0 & \cdots & 0 & b_{j-1} & a_{j-1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_j & b_{j+1} & 0 & \cdots & 0 \\ b_{j+1} & a_{j+1} & b_{j+2} & \ddots & 0 \\ 0 & b_{j+2} & a_{j+2} & \ddots & b_n \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_n & a_n \end{bmatrix}$$

If none of the b_j are zero, the QR method proceeds by forming a sequence of matrices $A_1, A_2, A_3, \ldots$, as follows.

$A_1 = Q_1 R_1$, Q_1 is orthogonal and R_1 is upper triangular.

$A_2 = R_1 Q_1 = Q_2 R_2$, Q is " " R_2 " " .

and so on.

In other words, we write $A_s = Q_s R_s$, $A_{s+1} = R_s . Q_s$.

Eigenvalues of A will be the eigenvalues of A_s, since they are similar.

The tridiagonal matrix B has $n-1$ generally nonzero entries below the main diagonal. These are $b_{21}, b_{32}, \ldots, b_{n,n-1}$. We multiply B from the left by a matrix C_2 such that $C_2 B = [b_{jk}^{(2)}]$ has $b_{21}^{(2)} = 0$.

We multiply this by a matrix C_3 such that

$$C_3 C_2 B = [b_{jk}^{(3)}] \text{ has } b_{32}^{(2)} = 0, \text{ etc.}$$

After $(n-1)$ such multiplications we are left with an upper triangular matrix R_0, namely

$$C_n C_{n-1} \cdots C_3 C_2 B_0 = R_0.$$

The C_j are very simple. C_j has 2×2 submatrix

$$\begin{bmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{bmatrix} \quad (\theta_j \text{ Suitable}) \leftarrow \text{Rotation matrix}$$

in rows $j-1$ and j and columns $j-1$ and j, entries 1 everywhere else on the main diagonal and all other entries 0. For instance if $n=4$, ~~writing~~

$$C_2 = \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & 0 & 0 \\ -\sin\theta_2 & \cos\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad C_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_4 & \sin\theta_4 \\ 0 & 0 & -\sin\theta_4 & \cos\theta_4 \end{bmatrix}$$

$$C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_3 & \sin\theta_3 & 0 \\ 0 & -\sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

<u>Note:-</u> C_j's are orthogonal.

Hence there product is orthogonal.

Why?

$$C_i C_i^T = I \quad \text{for } i = 2, \dots, n$$

there product:

$$(C_2 C_3 \dots C_n)(C_2 C_3 \dots C_n)^T$$

$$= (C_2 C_3 \dots C_n)(C_n^T C_{n-1}^T \dots C_3^T C_2^T)$$

$$= C_2 C_3 \dots C_{n-1} I C_{n-1}^T \dots C_3^T C_2^T$$

$$= I. \qquad \square$$

We have $(C_n C_{n-1} \dots C_3 C_2) B_0 = R_0$

$$\Rightarrow \quad B_0 = \underbrace{(C_n \dots C_3 C_2)^{-1}}_{Q_0} R_0$$

where

$$Q_0 = (C_n \dots C_3 C_2)^{-1} = C_2^{-1} C_3^{-1} \dots C_n^{-1}$$

$$= C_2^T C_3^T \dots C_{n-1}^T C_n^T$$

This is our QR-factorization of B_0.

$$B_1 = R_0 Q_0 = R_0 C_2^T C_3^T \dots C_{n-1}^T C_n^T \quad \text{——} \circledast$$

We do not need Q_0 expicitly. But to get B_1 from $\circledast$ we first compute $(R_0 C_2^T)$, then $(R_0 C_2^T) C_3^T$ etc.

Similarly in further steps that produce $B_2, B_3, \dots$

How to find $\cos\theta_j$ and $\sin\theta_j$?

$\cos\theta_2$ and $\sin\theta_2$ in C_2 must be such that $b_{21}^{(2)} = 0$ in the product

$$C_2 B = \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & 0 & \cdots & 0 \\ -\sin\theta_2 & \cos\theta_2 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & & \cdots & 1 \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & & \bigcirc \\ \boxed{b_{21}} & b_{22} & b_{32} & \\ & b_{32} & b_{33} & \ddots \\ \bigcirc & & \ddots & b_{nn} \end{bmatrix}$$

$$b_{21}^{(2)} = -\sin\theta_2 b_{11} + \cos\theta_2 b_{21} = 0$$

$$\Rightarrow \tan\theta_2 = \frac{b_{21}}{b_{11}}$$

$$\Rightarrow \sin\theta_2 = \frac{b_{21}}{\sqrt{b_{11}^2 + b_{21}^2}} = \frac{b_{21}/b_{11}}{\sqrt{1 + (b_{21}/b_{11})^2}},$$

$$\cos\theta_2 = \frac{b_{11}}{\sqrt{b_{11}^2 + b_{21}^2}} = \frac{1}{\sqrt{1 + (b_{21}/b_{11})^2}}.$$

Similarly,

For $b_{32}^{(3)} = -\sin\theta_3 b_{22} + \cos\theta_3 b_{32}$

$$\Rightarrow \tan\theta_3 = \frac{b_{32}}{b_{22}}$$

$$\sin\theta_3 = \frac{b_{32}}{\sqrt{b_{22}^2 + b_{32}^2}} = \frac{b_{32}/b_{22}}{\sqrt{1 + (b_{32}/b_{22})^2}}$$

$$\cos\theta_3 = \frac{b_{22}}{\sqrt{b_{22}^2 + b_{32}^2}} = \frac{1}{\sqrt{1 + (b_{32}/b_{22})^2}}$$

Ex. To calculate the eigenvalues of

$$A = \begin{bmatrix} 6 & 4 & 1 & 1 \\ 4 & 6 & 1 & 1 \\ 1 & 1 & 5 & 2 \\ 1 & 1 & 2 & 5 \end{bmatrix}$$ by QR method.

$\downarrow$ Householder's method
Tridiagonal form

$$B^* = \begin{bmatrix} 6 & -\sqrt{18} & 0 & 0 \\ -\sqrt{18} & 7 & \sqrt{2} & 0 \\ 0 & \sqrt{2} & 6 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Clearly, 3 is an eigenvalue of B^*. So we can take B as

$$B = \begin{bmatrix} 6 & -\sqrt{18} & 0 \\ -\sqrt{18} & 7 & \sqrt{2} \\ 0 & \sqrt{2} & 6 \end{bmatrix} = \begin{bmatrix} 6 & -4.24264 & 0 \\ -4.24264 & 7 & 1.41421 \\ 0 & 1.41421 & 6 \end{bmatrix}$$

First step :-

$$C_2 = \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & 0 \\ -\sin\theta_2 & \cos\theta_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \cancel{C_2 B =}$$

$b_{21}^{(1)}$ should be zero. $\quad -6\sin\theta_2 - \sqrt{18}\cos\theta_2 = 0$

$$\Rightarrow \quad \tan\theta_2 = -\frac{\sqrt{18}}{6} = -\frac{3\sqrt{2}}{6} = -\frac{1}{\sqrt{2}}$$

$$\sin\theta_2 = -\frac{1}{\sqrt{3}} \qquad \cos\theta_2 = \sqrt{\frac{2}{3}}$$

$$C_2 B = \begin{bmatrix} \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -3\sqrt{2} & 0 \\ -3\sqrt{2} & 7 & \sqrt{2} \\ 0 & \sqrt{2} & 6 \end{bmatrix} = \begin{bmatrix} 3\sqrt{6} & -\frac{13}{\sqrt{3}} & -\frac{\sqrt{2}}{3} \\ 0 & \frac{4\sqrt{2}}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 0 & \sqrt{2} & 6 \end{bmatrix}$$

$$C_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_3 & \sin\theta_3 \\ 0 & -\sin\theta_3 & \cos\theta_3 \end{bmatrix} \qquad C_2B = \begin{bmatrix} 3\sqrt{6} & -13/\sqrt{3} & -\frac{\sqrt{2}}{3} \\ 0 & 4\sqrt{2}/\sqrt{3} & 2/\sqrt{3} \\ 0 & \sqrt{2} & \sqrt{6} \end{bmatrix}$$

$$-\frac{4\sqrt{2}}{\sqrt{3}}\sin\theta_3 + \sqrt{2}\cos\theta_3 = 0$$

$$\Rightarrow \tan\theta_3 = \sqrt{3}/4 \Rightarrow \sin\theta_3 = \frac{\sqrt{3}}{\sqrt{19}}, \quad \cos\theta_3 = \frac{4}{\sqrt{19}}.$$

$$C_3C_2B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{\sqrt{19}} & \frac{\sqrt{3}}{\sqrt{19}} \\ 0 & -\frac{\sqrt{3}}{19} & \frac{4}{\sqrt{19}} \end{bmatrix} \begin{bmatrix} 3\sqrt{6} & -13/\sqrt{3} & -\sqrt{2}/3 \\ 0 & 4\sqrt{2}/\sqrt{3} & 2/\sqrt{3} \\ 0 & \sqrt{2} & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 3\sqrt{6} & -13/\sqrt{3} & -\sqrt{2}/3 \\ 0 & \sqrt{38/3} & 26/\sqrt{57} \\ 0 & 0 & 22/\sqrt{19} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 7.34847 & -7.50555 & -0.816497 \\ 0 & 3.55903 & 3.44378 \\ 0 & 0 & 5.04715 \end{bmatrix}}_{R_0}$$

$$\Rightarrow B = \underbrace{C_2^T C_3^T}_{Q_0} R_0$$

$$B_1 = R_0Q_0 = R_0 C_2^T C_3^T$$

$$= \begin{bmatrix} 10.3333 & -2.0548 & 0 \\ -2.0548 & 4.03509 & 2.00553 \\ 0 & 2.00553 & 4.63158 \end{bmatrix}$$

Second step:-

$$\cos\theta_2 = \frac{1}{\sqrt{1+(b_{21}/b_{11})^2}} \qquad \sin\theta_2 = \frac{b_{21}/b_{11}}{\sqrt{1+(b_{21}/b_{11})^2}}$$

$$= 0.980797 \qquad = -0.195033$$

$$C_2 B_1 = \begin{bmatrix} 10.5357 & -2.80232 & -0.391146 \\ 0 & 3.55684 & 1.96702 \\ 0 & 2.00553 & 4.63158 \end{bmatrix}$$

$$\cos\theta_3 = \frac{1}{\sqrt{1+(b_{32}^{(1)}/b_{22}^{(1)})^2}} \qquad \sin\theta_3 = \frac{b_{32}^{(1)}/b_{22}^{(1)}}{\sqrt{1+(b_{32}^{(1)}/b_{22}^{(1)})^2}}$$

$$= 0.871072 \qquad = 0.491155$$

$$R_1 = C_3 C_2 B_1 = \begin{bmatrix} 10.5357 & -2.80232 & -0.391146 \\ 0 & 4.0833 & 3.98824 \\ 0 & 0 & 3.06833 \end{bmatrix}$$

$$B_1 = \underbrace{C_2^T C_3^T}_{Q_1} R_1$$

$$B_2 = R_1 Q_1 = R_1 C_2^T C_3^T = \begin{bmatrix} 10.8799 & -0.796379 & 0 \\ -0.796379 & 5.44739 & 1.50703 \\ 0 & 1.50703 & 2.67273 \end{bmatrix}$$

The off-diagonal entries are somewhat smaller in absolute value than those of B_1, but too large for the diagonal entries to be good approximation for eigenvalues of B.

⋮

After 9 steps, eigenvalues are 11, 6, 2.

Eigenvalues of A are 11, 6, 3, 2.