Polynomials of Matrices

Consider a polynomial $f(t) = a_n t^n + a_{n-1}t^{n-1} + ... + a_1 t + a_0$ over a field $\mathbb{F}$. If A is any square matrix, then we define

$$f(A) = a_n A^n + a_{n-1}A^{n-1} + ... + a_1 A + a_0 I$$

where I is the identity matrix. In particular, we say that A is a root of $f(t)$ if $f(A) = 0$, the zero matrix.

Example : Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then $A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$. Let $f(t) = 2t^2 - 3t + 5$ and $g(t) = t^2 - 5t - 2$

Then

$$f(A) = 2A^2 - 3A + 5I = \begin{bmatrix} 14 & 20 \\ 30 & 44 \end{bmatrix} + \begin{bmatrix} -3 & -6 \\ -9 & -12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 16 & 14 \\ 21 & 37 \end{bmatrix}$$

and

$$g(A) = A^2 - 5A - 2I = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} + \begin{bmatrix} -5 & -10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Thus, A is a zero of $g(t)$.

Theorem: (Cayley Hamilton) Every matrix A is a root of its characteristic polynomial.

Remark: Suppose $A = [a_{ij}]$ is a triangular matrix. Then $tI - A$ is a triangular matrix with diagonal entries $t - a_{ii}$; hence,

$$\Delta(t) = det(tI - A) = (t - a_{11})(t - a_{22}) \dots (t - a_{nn})$$

Observe that the roots of $\Delta(t)$ are the diagonal elements of A.

Example: Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$. Its characteristic polynomial is

$\Delta(t) = |tI - A| = \begin{bmatrix} t-1 & -3 \\ -4 & t-5 \end{bmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$

As expected from the Cayley Hamilton theorem, A is a root of $\Delta(t)$; that is,

$$\Delta(t) = A^2 - 6A - 7I = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix} + \begin{bmatrix} -6 & -18 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} -7 & 0 \\ 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Now suppose A and B are similar matrices, say $B = P^{-1}AP$, where P is invertible. We show that A and B have the same characteristic polynomial. Using $B = P^{-1}AP$, we have

$$\begin{aligned} \Delta_B(t) &= det(tI - B) = det(tI - P^{-1}AP) = det(P^{-1}tIP - P^{-1}AP) \\ &= det(P^{-1}(tI - A)P) = det(P^{-1})det(tI - A)det(P) \end{aligned}$$

Using the fact that determinants are scalars and commute and that $det(P^{-1})det(P) = 1$, we finally obtain

$$\Delta_B(t) = det(tI - A) = \Delta_A(t)$$

Thus, we have proved the following theorem.

Theorem : Similar matrices have the same characteristic polynomial.

Theorem: A square matrix $A_{n\times n}$ is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors. In this case, the diagonal elements of D are the corresponding eigenvalues and $D = P^{-1}AP$, where P is the matrix whose columns are the eigenvectors.

Eigenvalues and Eigenvectors of a Matrix Let A be an $n \times n$ matrix over a field $\mathbb{F}$. A scalar $\lambda \in \mathbb{F}$ is said to be an eigenvalue of A if

$$Ax = \lambda x$$

for some nonzero column vector $x \in \mathbb{F}$. Such a vector x is referred to as an eigenvector corresponding to the eigenvalue (λ).

From above definition its is clear that $Ax = \lambda x$ is equivalent to $Ax - \lambda x = 0 \implies (A - \lambda I)x = 0$ (where I is a $n \times n$ identity matrix), which has a non-zero solution if and only if $A - \lambda I$ is singular. So finding λ is equivalent in solving

$$det(A - \lambda I) = 0$$

Above equation represents the polynomial of degree n (with λ as the variable). In order to get λ all we have to do is to find the roots of the n degree polynomial $p(\lambda) = det(A - \lambda I)$. The coefficient of the $p(\lambda)$ can in general be complex number. So we have

$$p(\lambda) = det(A - \lambda I) = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \dots (\lambda - \lambda_n)$$

where $\lambda_1, \lambda_2 \dots \lambda_n$ are the roots of the polynomial $p(\lambda)$. These scalars are the eigen value of the A. In general, $\lambda_1, \lambda_2 \dots \lambda_n$ need not be necessarily be distinct.

Putting $\lambda = 0$ in $p(\lambda)$ we have $det(A) = \lambda_1\lambda_2 \dots \lambda_n$ and comparing the coefficient of λ we get $a_{11} + a_{22} + \ldots + a_{nn} = trace(A) = (\lambda_1 + \lambda_1 + \ldots + \lambda_n)$.

The eigenvectors x corresponding to the each eigenvalue of A are the solutions to the linear equation system $(A-\lambda I)x = 0$ that is, the null space of $(A - \lambda I)$. We call this space the eigenspace of A corresponding to λ.

Example : Find the eigenvalue of the following matrix.

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

$$\begin{aligned} det(A-\lambda I) &= det\left(\begin{bmatrix} 1-\lambda & 2 & -1 \\ 1 & 0-\lambda & 1 \\ 4 & -4 & 5-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(-\lambda(5-\lambda)+4) - 2(5-\lambda-4) - 1(-4+4\lambda) \\ &= (1-\lambda)(\lambda^2-5\lambda+4) - 2(1-\lambda) + 4(1-\lambda) \\ &= (1-\lambda)(\lambda^2-5\lambda+6) \\ &= (1-\lambda)(\lambda-2)(\lambda-3) \end{aligned}$$

Clearly, $det(A - \lambda I) = 0 \implies \lambda = \{1, 2, 3\}$ are the eigenvalues of matrix A (Distinct eigenvalues). Suppose a matrix A can be diagonalized as above, say $P^{-1}AP = D$, where D is diagonal. Then A has the extremely useful diagonal factorization $A = PDP^{-1}$.

THEOREM : Prove that an $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigen vectors.

PROOF : (i) If A is diagonalizable then A has n linearly independent eigen vector.

Consider $\lambda_1, \lambda_2, \dots \lambda_n$ are the n eigen values (not necessarily distinct) of the matrix A. If A is diagonalizable then we know by similarity transformation argument that there exist an $n \times n$ invertible matrix P such the

$$D = P^{-1}AP$$

where, D is a diagonal matrix with $\lambda_1, \lambda_2, \dots \lambda_n$ as diagonal entries. Above equation can be written as

$$\begin{gathered} D = P^{-1}AP \\ PD = AP \end{gathered}$$

Now, the fact that P is invertible implies that P consists of n linearly independent columns. Let $x_1, x_2, \dots, x_n$ are n linearly independent columns of P.

$$P = \begin{bmatrix} | & | & \cdots & | \\ | & | & \cdots & | \\ x_1 & x_2 & \dots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}$$

From above collected information about P and D matrices we can write

$$\begin{aligned} PD &= AP \\ \begin{bmatrix} | & | & \cdots & | \\ | & | & \cdots & | \\ x_1 & x_2 & \dots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} &= A \begin{bmatrix} | & | & \cdots & | \\ | & | & \cdots & | \\ x_1 & x_2 & \dots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} \\ \begin{bmatrix} | & | & \cdots & | \\ | & | & \cdots & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \dots & \lambda_n x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} &= \begin{bmatrix} | & | & \cdots & | \\ | & | & \cdots & | \\ Ax_1 & Ax_2 & \dots & Ax_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix} \end{aligned}$$

From above equation it is quite clear that $Ax_i = \lambda_i \quad \forall i = 1 \dots n$ which from the definition of eigen vectors implies that $x_i's$ are eigen vectors of A corresponding to each $\lambda_i's$. But from the fact that P is an invertible matrix and P is formed by stacking the $x_i's$ column-wise, $\{x_i's\}$ are n eigen vectors of A.

(ii) If A has n linearly independent eigen vectors than A is digonalizable.

Let $\{x_i's\}_{i=1\dots n}$ are the n independent vectors corresponding to n (not necessarily distinct) eigen values $\{\lambda_i\}_{i=1\dots n}$ of A respectively. Clearly, from definition of eigen vectors, we have

$$Ax_i = \lambda_i \qquad \forall i = 1 \dots n$$

Further, let P be an matrix with n linearly independent columns which are eigen vectors of A.

$$
\begin{aligned}
AP &= A\begin{bmatrix} | & | & \dots & | \\ | & | & \dots & | \\ x_1 & x_2 & \dots & x_n \\ | & | & \dots & | \\ | & | & \dots & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & \dots & | \\ | & | & \dots & | \\ Ax_1 & Ax_2 & \dots & Ax_n \\ | & | & \dots & | \\ | & | & \dots & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & \dots & | \\ | & | & \dots & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \dots & \lambda_n x_n \\ | & | & \dots & | \\ | & | & \dots & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & \dots & | \\ | & | & \dots & | \\ x_1 & x_2 & \dots & x_n \\ | & | & \dots & | \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}
\end{aligned}
$$

Hence, $AP = PD$, where D is the diagonal matrix with eigen values as its diagonals. Now, since P is invertible, P^{-1} exists. Hence $D = P^{-1}AP$ which says that A is diagonalizable.

Remark: Using this factorization, the algebra of A reduces to the algebra of the diagonal matrix D, which can be easily calculated. Specifically, suppose $D = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$. Then $A^m = (PDP^{-1})^m = PD^mP^{-1} = diag(\lambda_1, \lambda_2, \dots, \lambda_n)P^{-1}$. More generally, for any polynomial $f(t)$, $f(A) = f(PDP^{-1}) = Pf(D)P^{-1} = Pdiag(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n))P^{-1}$ Furthermore, if the diagonal entries of D are nonnegative, let

$$B = P\ diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})P^{-1}$$

Then B is a nonnegative square root of A; that is, $B^2 = A$ and the eigenvalues of B are nonnegative.

Example : Let $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$. Find A^4 and $f(t) = t^3 - 5t^2 + 3t + 6$ then find $f(A)$.

Example : It can be easily verified that $\lambda_1 = 1$ and $\lambda_1 = 4$ are the eigenvalues of A and $v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are corresponding eigenvectors. Observe that v_1 and v_2 are linearly independent and hence form a basis of $\mathbb{R}^2$. Accordingly, A is diagonalizable. Furthermore, let P be the matrix whose columns are the eigenvectors v_1 and v_2. That is, let

$$P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}, P^{-1} = \tfrac{1}{3}\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$$

Then A is similar to the diagonal matrix

$$D = P^{-1}AP = \tfrac{1}{3}\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

As expected, the diagonal elements 1 and 4 in D are the eigenvalues corresponding, respectively, to the eigenvectors v_1 and v_2, which are the columns of P. In particular, A has the factorization

$$A = PDP^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Accordingly,

$$A^4 = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 256 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 171 & 85 \\ 170 & 86 \end{bmatrix}$$

Moreover, suppose $f(t) = t^3 - 5t^2 + 3t + 6$; hence $f(1) = 5$ and $f(4) = 2$. Then

$$f(A) = Pf(D)P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}$$

Lastly, we obtain a positive square root of A. Specifically, using $\sqrt{1} = 1$ and $\sqrt{4} = 2$, we obtain the matrix

$$B = P\sqrt{D}P^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{4}{3} \end{bmatrix}$$

where $B^2 = A$ and where B has positive eigenvalues 1 and 2.

Theorem: Suppose $v_1, v_2, \dots, v_n$, are nonzero eigenvectors of a matrix A belonging to distinct eigenvalues $\lambda_1, \lambda_2, \dots \lambda_n$. Then $v_1, v_2, \dots, v_n$, are linearly independent.

Theorem: Suppose the characteristic polynomial $\Delta_A(\lambda)$ of an matrix $A_{n\times n}$ is a product of n distinct factors, say, $\Delta_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n)$. Then A is similar to a diagonal matrix $D = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Algebraic and Geometric Multiplicity If λ is an eigenvalue of a matrix A, then the algebraic multiplicity of A is defined to be the multiplicity of λ as a root of the characteristic polynomial of A, while the geometric multiplicity of A is defined to be the dimension of its **eigenspace**, dim E_λ.

Theorem: The geometric multiplicity of an eigenvalue λ of a matrix A does not exceed its algebraic multiplicity.
Diagonalization Algorithm The input is an square matrix $A_{n\times n}$.

1. Find the characteristic polynomial $\Delta_A(\lambda)$ of A.
2. Find the roots of $\Delta_A(\lambda)$ to obtain the eigenvalues of λ
3. Repeat (a) and (b) for each eigenvalue λ of A
   (a) Form the matrix $A - \lambda I$ by subtracting λ down the diagonal of A.
   (b) Find a basis for the solution space of the homogeneous system $(A - \lambda I)x = 0_{n\times 1}$. (These basis vectors are linearly independent eigenvectors of A belonging to λ.)
4. Consider the collection $S = \{v_1, v_2, \dots, v_m\}$ of all linearly independent eigenvectors (corresponding to all eigenvalues) obtained in Step 3.
   (a) If $m \neq n$, then A is not diagonalizable.
   (b) If in $m = n$, then A is diagonalizable. Specifically, let P be the matrix whose columns are the eigenvectors $v_1, v_2, \dots, v_n$. Then $D = P^{-1}AP = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$ where A is the eigenvalue corresponding to the eigenvector v.

DIAGONALIZING REAL SYMMETRIC MATRICES
There are many real matrices A that are not diagonalizable. In fact, some real matrices may not have any (real) eigenvalues. However, if A is a real symmetric matrix, then these problems do not exist. Namely, we have the following theorems.

Theorem : Let A be a real symmetric matrix. Then each root λ of its characteristic polynomial is real.

Theorem : Let A be a real symmetric matrix. Suppose u and v are eigenvectors of A belonging to distinct eigenvalues λ_1 and λ_1. Then u and v are orthogonal, that is, $u^T v = 0$.

Theorem : Let A be a real symmetric matrix. Then there exists an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal.
The orthogonal matrix P is obtained by normalizing a basis of orthogonal eigenvectors of A as illustrated below. In such a case, we say that A is orthogonally diagonalizable.

Orthogonal Diagonalization Algorithm The input is a real symmetric matrix A.

1. Find the characteristic polynomial $\Delta_A(\lambda)$ of A.
2. Find the eigenvalues of A, which are the roots of $\Delta_A(\lambda)$.
3. For each eigenvalue λ of A in Step 2, find an orthogonal basis of its eigenspace. (Here in the case when some eigenvalue yields more than one independent eigenvectors, one may use Gram-Schmidt process to find the orthogonal eigenvectors corresponding to the same eigenvalue. Note the fact that the eigenvectors corresponding to distinct eigenvalues are already orthogonal as A is symmetric).
4. Normalize all eigenvectors in Step 3, which then forms an orthonormal basis of $\mathbb{R}^n$.
5. Now find matrix P, whose columns are the normalized eigenvectors in Step 4.
   Now it can be easily checked that columns of P are normalized and P is a orthogonal matrix i.e. $P^{-1} = P^T$ and $|P| = 1$.

Application to Quadratic Forms
Let q be a real polynomial in variables $x_1, x_2, \dots, x_n$ such that every term in q has degree two; that is

$$q(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} c_i x_i^2 + \sum_{i=1, i<j}^{n} d_{ij} x_i x_j$$

Then q is called a quadratic form. If there are no cross-product terms that is, all $d_{ij} = 0$, then q is said to be diagonal. The above quadratic form q determines a real symmetric matrix $A = [a_{ij}]_{n\times n}$, where $a_{ii} = c_i$ and $a_{ij} = a_{ji} = \frac{1}{2}d_{ij}$. Namely, q can be written in the matrix form

$$q(X) = X^T AX$$

where $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the column vector of the variables. Furthermore, suppose $X = PY$ is a linear substitution of the variables. Then substitution in the quadratic form yields $q(Y) = (PY)^T A(PY) = Y^T(P^T AP)Y$.
Thus $P^T AP$ is the matrix representation of q in the new variables. We seek an orthogonal matrix P such that the orthogonal substitution $X = PY$ yields a diagonal quadratic form, that is, for which $P^T AP$ is diagonal. Since P is orthogonal, $P^T = P^{-1}$, and hence $P^T AP = P^{-1}AP$. The above theory yields such an orthogonal matrix P.

Example: Consider the quadratic form

$$q(x, y) = 2x^2 - 4xy + 5y^2 = X^T AX, \text{ where } A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \text{ and } X = \begin{bmatrix} x \\ y \end{bmatrix}$$

One can easily verify that $\lambda_1 = 6$ and $\lambda_2 = 1$ and $x_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ are the corresponding eigenvectors. Clearly the eigenvectors are orthogonal but not normalize. We first normalize then by dividing them with their respective norm

$$\hat{x}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ and } \hat{x}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Now the matrix P is formed by taking normalized eigen vectors $\hat{x}_1$ and $\hat{x}_2$ as the column vectors. (This matrix is orthogonal matrix as $P^{-1} = P^T$)

$$P = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}, \quad P^{-1}AP = P^T AP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

The matrix P corresponds to the following linear orthogonal substitution of the variables x and y in terms of the variables x' and y'

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = PY = \begin{bmatrix} x \\ y \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y' \\ \frac{-2}{\sqrt{5}}x' + \frac{1}{\sqrt{5}}y' \end{bmatrix} \Longrightarrow \begin{array}{l} x = \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y' \\ y = \frac{-2}{\sqrt{5}}x' + \frac{1}{\sqrt{5}}y' \end{array}$$

This substitution in $q(x, y)$ yields the diagonal quadratic form $q(x', y') = 6x'^2 + y'^2$.

MINIMAL POLYNOMIAL

Let A be any square matrix. Let $J(A)$ denote the collection of all polynomials $f(T)$ for which A is a root, i.e., for which $f(A) = 0_{n\times n}$. The set $J(A)$ is not empty; since the CayleyHamilton theorem tells us that the characteristic polynomial $\Delta_A(\lambda)$ of A belongs to $J(A)$. Let $m_A(t)$ denote the monic polynomial of lowest degree in $J(A)$. (Such a polynomial $m(t)$ exists and is unique). We call $m(t)$ the minimal polynomial of the matrix A.

Remark: A polynomial $f(t) = 0$ is monic if its leading coefficient (coefficient of highest power term) equals one.

Theorem : The minimal polynomial $m_A(t)$ of a matrix A divides every polynomial that has A as a zero. In particular, $m_A(t)$ divides the characteristic polynomial $\Delta_A(\lambda)$ of A.

There is an even stronger relationship between $m_A(t)$ and $\Delta_A(\lambda)$.

Theorem : The characteristic polynomial $\Delta_A(\lambda)$ and the minimal polynomial $m_A(t)$ of a matrix A have the same irreducible factors.

Remark : Its is important to note that $m_A(t)$ need not be equal to $\Delta_A(\lambda)$, only that any irreducible factor of one must divide the other. In particular, since a linear factor is irreducible,$m_A(t)$ and $\Delta_A(\lambda)$ have the same linear factors. Hence they have the same roots.

Theorem : A scalar λ is an eigenvalue of the matrix A if and only if λ is a root of the minimal polynomial of A.

Jordan Block

Consider the following two r-square matrices, where $a \neq 0$

$$J(\theta, r) = \begin{bmatrix} \theta & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \theta & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \theta & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \theta & 1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & \theta \end{bmatrix}, \quad A = \begin{bmatrix} \theta & a & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \theta & a & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \theta & a & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \theta & a \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & \theta \end{bmatrix}$$

The first matrix, called a Jordan Block, has θ on the diagonal, ls on the superdiagonal (consisting of the entries above the diagonal entries), and 0s elsewhere. The second matrix A has θs on the diagonal, as on the superdiagonal, and 0s elsewhere. Thus A is a generalization of $J(A, r)$. One can show that

$$f(\lambda) = (\lambda - \theta)^r$$

is both the characteristic and minimal polynomial of both $J(\theta, r)$ and A.

Companion Matrix

Consider an arbitrary monic polynomial

$$f(\lambda) = \lambda^{\,n} + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots + a_1\lambda + a_0$$

Let $C(f)$ be the n-square matrix with ls on the subdiagonal (consisting of the entries below the diagonal entries), the negatives of the coefficients in the last column, and 0s elsewhere as follows:

$$C(f) = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 & -a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 0 & 0 & -a_{n-3} \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & \ldots & 0 & 0 & 1 & -a_{n-1} \end{bmatrix}$$

Then $C(f)$ is called the companion matrix of the polynomial $f(t)$. Moreover, the minimal $m_{C(f)}(\lambda)$ polynomial and the characteristic polynomial $\Delta_{C(f)}(t)$ of the companion matrix $C(f)$ are both equal to the original polynomial $f(t)$.

Example: Find a matrix whose minimal polynomial is $f(\lambda) = \lambda^3 - 8\lambda^2 + 5\lambda + 7$. **Solution:** One such matrix is the companion matrix corresponding to the given polynomial.

$$m_{C(f)}(\lambda) = \begin{bmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -7 \\ 1 & 0 & -5 \\ 0 & 1 & 8 \end{bmatrix}$$

CHARACTERISTIC AND MINIMAL POLYNOMIALS OF BLOCK MATRICES

Characteristic Polynomial and Block Triangular Matrices

Suppose M is a block triangular matrix, say $\begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$ where A_1 and A_2 are square matrices. Then $(M - \lambda I)$ is also a block triangular matrix whose diagonal entries are $A_1 - \lambda I$ and $A_2 - \lambda I$. Thus

$$|M - \lambda I| = \begin{vmatrix} A_1 - \lambda I & B \\ 0 & A_2 - \lambda I \end{vmatrix} = |A_1 - \lambda I| \times |A_2 - \lambda I|$$

That is, the characteristic polynomial of M is the product of the characteristic polynomials of the diagonal blocks A_1 and A_2. By induction, we obtain the following useful result.

Theorem : Suppose M is a block triangular matrix with diagonal blocks $A_1, A_2, \ldots, A_n$. Then the characteristic polynomial of M is the product of the characteristic polynomials of the diagonal blocks $A_1, A_2, \ldots, A_n$ that is,

$$\Delta_M(\lambda) = \Delta_{A_1}(\lambda)\Delta_{A_2}(\lambda)\ldots\Delta_{A_n}(\lambda)$$

Example: Find the characteristic polynomial of the matrix $M = \begin{bmatrix} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{bmatrix}$.

Solution: The above matrix can be portioned as $M = \left[\begin{array}{cc|cc} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ -- & -- & -- & -- \\ 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{array}\right]. = \begin{bmatrix} A_1 & B \\ 0_{2\times 2} & A_2 \end{bmatrix}$

Where $A_1 = \begin{bmatrix} 9 & -1 \\ 8 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 3 & 6 \\ -1 & 8 \end{bmatrix}, B = \begin{bmatrix} 5 & 7 \\ 2 & -4 \end{bmatrix}, 0_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Clearly M is a block triangular matrix with diagonal blocks A_1 and A_2. The characteristic polynomial of M would be equal to the product of characteristic polynomial of the diagonal blocks A_1 and A_2.

The characteristic polynomial of A_1 is given by

$$\Delta_{A_1}(\lambda) = |A_1 - \lambda I| = \begin{vmatrix} 9-\lambda & -1 \\ 8 & 3-\lambda \end{vmatrix} = (9-\lambda)(3-\lambda) - (-1)\times 8 = 27 - 9\lambda - 3\lambda + \lambda^2 + 8 = \lambda^2 - 12\lambda + 35 = (\lambda - 5)(\lambda - 7)$$

The characteristic polynomial of A_1 is given by

$$\Delta_{A_2}(\lambda) = |A_2 - \lambda I| = \begin{vmatrix} 3-\lambda & 6 \\ -1 & 8-\lambda \end{vmatrix} = (3-\lambda)(8-\lambda) - (-1)\times 6 = 24 - 8\lambda - 3\lambda + \lambda^2 + 6 = \lambda^2 - 11\lambda + 30 = (\lambda - 5)(\lambda - 6)$$

Hence the characteristic polynomial of M is given by

$$\Delta_M(\lambda) = \Delta_{A_1}(\lambda) \times \Delta_{A_2}(\lambda) = \{(\lambda-5)(\lambda-7)\} \times \{(\lambda-5)(\lambda-6)\} = (\lambda-5)^2(\lambda-6)(\lambda-7)$$

Minimal Polynomial and Block Diagonal Matrices

Theorem : Suppose M is a block triangular matrix with diagonal blocks $A_1, A_2, \ldots, A_n$. Then the minimal polynomial of M is equal to the least common multiple (LCM) of the minimal polynomials of the diagonal blocks $A_1, A_2, \ldots, A_n$.

Example: Find the characteristic polynomial of the block diagonal matrix $M = \left[\begin{array}{cc|cc|c} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ \hline 0 & 0 & 0 & 0 & 7 \end{array}\right]$.

Solution: The given matrix can be seen as

$$M = \begin{bmatrix} A_1 & 0_{2\times2} & 0_{2\times1} \\ 0_{2\times2} & A_2 & 0_{2\times1} \\ 0_{1\times2} & 0_{1\times2} & A_3 \end{bmatrix}$$

Where $A_1 = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} 4 & 2 \\ 3 & 5 \end{bmatrix}, A_3 = \begin{bmatrix} 7 \end{bmatrix}, 0_{2\times2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, 0_{1\times2} = \begin{bmatrix} 0 & 0 \end{bmatrix} 0_{2\times1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Delta_{A_1}(\lambda) = |A_1 - \lambda I| = \begin{vmatrix} 2-\lambda & 5 \\ 0 & 2-\lambda \end{vmatrix} = (2-\lambda)(2-\lambda) - 0 \times 5 = (\lambda-2)^2$$

The characteristic polynomial of A_1 is given by

$$\Delta_{A_2}(\lambda) = |A_2 - \lambda I| = \begin{vmatrix} 4-\lambda & 2 \\ 3 & 5-\lambda \end{vmatrix} = (4-\lambda)(5-\lambda) - 3 \times 2 = 20 - 4\lambda - 5\lambda + \lambda^2 - 6 = \lambda^2 - 9\lambda + 14 = (\lambda-2)(\lambda-7)$$

$$\Delta_{A_3}(\lambda) = |A_3 - \lambda I| = (\lambda - 7)$$

Hence the characteristic polynomial of M is given by

$$\Delta_M(\lambda) = \Delta_{A_1}(\lambda) \times \Delta_{A_2}(\lambda) \times \Delta_{A_3}(\lambda) = \{(\lambda-2)^2\} \times \{(\lambda-2)(\lambda-7)\} \times \{(\lambda-7)\} = (\lambda-2)^3(\lambda-7)^2$$

It can be easily verified that the minimal polynomials $m_{A_1}(t), m_{A_2}(t)$ and $m_{A_3}(t)$, of the diagonal blocks A_1, A_2, A_3 are $(\lambda-2)^2, (\lambda-2)(\lambda-7)$ and $(\lambda-7)$ respectively .i.e.
$m_{A_1}(t) = (\lambda-2)^2, m_{A_2}(t) = (\lambda-2)(\lambda-7)$ and $m_{A_3}(t) = (\lambda-7)$
The minimal polynomial of M, $m(t)$ is equal to the least common multiple of $m_{A_1}(t), m_{A_2}(t)$ and $m_{A_3}(t)$ Thus $m(t) = (\lambda-2)^2(\lambda-7)$.

Inner Product Spaces

Let $(V, +, .)$ be areal vector space. Suppose to each pair of elements $u, v \in V$ there is assigned a real number, denoted by $< u, v >$. This function is called a real inner product on V if it satisfies the following axioms

1. **Linear Property** $\langle a\, u + b\, v, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$
2. **Symmetric Property** $\langle u, v\rangle = \langle v, u\rangle$
3. **Positive Definite Property** $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0$ iff $u = 0$

ORTHOGONALITY

Let V by an inner product space. The vectors $u, v \in V$ are said to be orthogonal and u is said to be orthogonal to v if

$$\langle u, v\rangle = 0$$

The relation is clearly symmetric, that is, if u is orthogonal to v, then $\langle v, u\rangle = 0$, and so v is orthogonal to u. Also $0 \in V$ is orthogonal to every $v \in V$ since

$$\langle 0, v\rangle = \langle 0v, v\rangle = 0\langle v, v\rangle = 0$$

Example : Consider the vectors $u = (1, 1, 1), v = (1, 2, -3), w = (1, -4, 3)$ in $\mathbb{R}^3$. Then $\langle u, v\rangle = 1 + 2 - 3 = 0$, $\langle v, w\rangle = 1 - 4 + 3 = 0$, $\langle v, w\rangle = 1 - 8 - 9 = -16$.
Thus u is orthogonal to v and w, but v and w are not orthogonal.

Example : Consider the functions $\sin t$ and $\cos t$ in the vector space $\mathbb{C}[\pi, \pi]$ of continuous functions on the closed interval $[\pi, \pi]$. Then $\langle \sin t, \cos t\rangle = \sum_{\pi}^{-\pi} \sin t \cos t dt = \frac{1}{2}[\sin^2 t]_{-\pi}^{\pi} = 0 - 0 = 0$

Thus $\sin t$ and $\cos t$ are orthogonal functions in the vector space $\mathbb{C}[\pi, \pi]$.

GRAM-SCHMIDT ORTHOGONALLZATION PROCESS

Suppose $u_1, u_2, \ldots, u_n$ is a basis of an inner product space V One can use this basis to construct an orthogonal basis $v_1, v_2, \ldots v_n$ of V as follows. Set

$$
\begin{array}{lcl}
v_1 & = & u_1 \\
v_2 & = & u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \\
v_3 & = & u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2 \\
\ldots & \ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\ldots & \ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
v_n & = & u_n - \frac{\langle u_n, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_n, v_2\rangle}{\langle v_2, v_2\rangle} v_2 \ldots \frac{\langle u_n, v_{n-1}\rangle}{\langle v_{n-1}, v_{n-1}\rangle} v_{n-1}
\end{array}
$$

Each v_k is orthogonal to the preceeding v's. Thus $v_1, v_2, \ldots v_n$ form an orthogonal basis for V as claimed. Normalizing each v will then yield an orthonormal basis for V. The above construction is known as the GramSchmidt orthogonalization process.

Remark: Suppose $\{u_1, u_2, \ldots, u_n\}$ are linearly independent, and they form a basis for U. Applying the GramSchmidt orthogonalization process to the $u_1, u_2, \ldots, u_n$ yields an orthogonal basis for $v_1, v_2, \ldots v_n$.

Example: Apply the GramSchmidt orthogonalization process to find an orthogonal basis and then an orthonormal basis for the subspace U of $\mathbb{R}^4$ spanned by $u_1 = (1,1,1,1), u_2 = (1,2,4,5), u_3 = (1,-3,-4,-2)$. **Solution:** GramSchmidt orthogonalization is applied as follows

1. First set $v_1 = u_1 = (1,1,1,1)$.
2. Compute $v_2 = u_2 - \dfrac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (1,2,4,5) - \dfrac{12}{4}(1,1,1,1) = (-2,-1,1,2)$.
3. $v_3 = u_3 - \dfrac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \dfrac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2 = (1,-3,-4,-2) - \dfrac{-8}{4}(1,1,1,1) - \dfrac{-7}{10}(-2,-1,1,2) = (\dfrac{8}{5}, \dfrac{-17}{10}, \dfrac{-13}{10}, \dfrac{7}{5}) = \dfrac{1}{10}(16,-17,-13,14)$.

Thus v_1, v_2, v_3 form an orthogonal basis for U. Normalize these vectors to obtain an orthonormal basis $\{\hat{v}_1, \hat{v}_2, \hat{v}_3\}$ of U. We have $\|v_1\|^2 = 4$, $\|v_2\|^2 = 10$, $\|v_3\|^2 = \frac{910}{100}$, so $\hat{v}_1 = \frac{1}{2}(1,1,1,1), \hat{v}_2 = \frac{1}{\sqrt{10}}(-2,-1,1,2), \hat{v}_3 = \frac{1}{\sqrt{910}}(16,-17,-13,14)$

Example: Let $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$.

1. Find all eigenvalues of A.
2. Find maximal set S of all linearly independent eigenvalues.
3. Find algebraic and geometric multiplicity of all eigenvalues of A.
4. Is A diagonalizable? If yes, find P such $D = P^{-1}AP$ is diagonalizable.
5. Find the minimal polynomial of A.

Solution: First we find the characteristic polynomial $\Delta(t)$ of A.

1. $\Delta(t) = \begin{vmatrix} 4-\lambda & 1 & -1 \\ 2 & 5-\lambda & -2 \\ 1 & 1 & 2-\lambda \end{vmatrix}$
   $= (4-\lambda)\begin{vmatrix} 5-\lambda & -2 \\ 1 & 2-\lambda \end{vmatrix} - 1\begin{vmatrix} 2 & -2 \\ 1 & 2-\lambda \end{vmatrix} - 1\begin{vmatrix} 2 & 5-\lambda \\ 1 & 1 \end{vmatrix}$
   $= (4-\lambda)\{(5-\lambda)(2-\lambda) - 1(-2)\} - (2 \times (2-\lambda) - (-2) \times 1) - (2 \times 1 - (5-\lambda) \times 1)$
   $= (4-\lambda)(10 - 5\lambda - 2\lambda + \lambda^2 + 2) - (4 - 2\lambda + 2) - (2 - 5 + \lambda)$
   $= (4-\lambda)(12 - 7\lambda + \lambda^2) - 6 + 2\lambda + 3 - \lambda$
   $= 48 - 28\lambda + 4\lambda^2 - 12\lambda + 7\lambda^2 - \lambda^3 - 3 + \lambda$
   $= -\lambda^3 + 11\lambda^2 - 39\lambda + 45$

   Now, $\Delta(t) = 0 \Longrightarrow -\lambda^3 + 11\lambda^2 - 39\lambda + 45 = 0 \Longrightarrow (\lambda-3)^2(\lambda-5) = 0 \Longrightarrow \lambda_1 = 3, \lambda_2 = 3, \lambda_3 = 5$ are the eigenvalues of matrix A.
2. To find eigenvector corresponding to the eigenvalue $\lambda_1 = \lambda_2 = 3$ we solve the system $(A - \lambda_1)x = 0$ i.e.

$$
\begin{aligned}
&\begin{bmatrix} 4-\lambda_1 & 1 & -1 \\ 2 & 5-\lambda_1 & -2 \\ 1 & 1 & 2-\lambda_1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
\Longrightarrow &\begin{bmatrix} 4-3 & 1 & -1 \\ 2 & 5-3 & -2 \\ 1 & 1 & 2-3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
\Longrightarrow &\begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}
$$

$$\begin{array}{rcrcrcl} & x_1 & + & x_2 & - & x_3 & = & 0 \\ \Longrightarrow & 2x_1 & + & 2x_2 & - & 2x_3 & = & 0 \\ & x_1 & + & x_2 & - & x_3 & = & 0 \end{array}$$

Since the last two equations are scalar multiple of the first equation.

$$\Longrightarrow x_1 + x_2 - x_3 = 0$$

Since above is one equation in three unknowns, to solve it we choose $(3-1) = 2$ variable, say x_1 and x_2 arbitrarily.

$$\Longrightarrow x_3 = x_1 + x_2$$

$$\Longrightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Hence $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ are two linearly independent eigenvectors corresponding to the eigenvalue $\lambda_1 = \lambda_2 = 3$

Now, to find eigenvector corresponding to the eigenvalue $\lambda_3 = 5$ we solve the system $(A - \lambda_3)x = 0$ i.e.

$$\begin{bmatrix} 4-\lambda_3 & 1 & -1 \\ 2 & 5-\lambda_3 & -2 \\ 1 & 1 & 2-\lambda_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 4-5 & 1 & -1 \\ 2 & 5-5 & -2 \\ 1 & 1 & 2-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{array}{rcrcrcl} & -x_1 & + & x_2 & - & x_3 & = & 0 \\ \Longrightarrow & 2x_1 & + & 0x_2 & - & 2x_3 & = & 0 \\ & x_1 & + & x_2 & - & 3x_3 & = & 0 \end{array}$$

Since the third equation is sum of the first and second equations.

$$\Longrightarrow \begin{array}{rcrcrcl} -x_1 & + & x_2 & - & x_3 & = & 0 \\ 2x_1 & + & 0x_2 & - & 2x_3 & = & 0 \end{array}$$

Since above is a system of two equations in three unknowns, to solve it we choose $(3-2) = 1$ variable, say x_3 arbitrarily. $\Longrightarrow x_3 = x_1 \Longrightarrow x_2 = x_1 + x_3 = 2x_1$

$$\Longrightarrow x = \begin{bmatrix} x_1 \\ 2x_1 \\ x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Hence $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_3 = 5$. Hence $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$

3. Since $\lambda_1 = \lambda_2 = 3$ i.e. 3 is twice repeated root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_1 = \lambda_2 = 3$ is two. There are two linearly independent eigenvectors corresponding to the eigenvalue $\lambda_1 = \lambda_2 = 3 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_1 = \lambda_2 = 3$ is two.
   Also $\lambda_3 = 5$ is a root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_3 = 5$ is one. There is one eigenvector corresponding to the eigenvalue $\lambda_3 = 5 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_3 = 5$ is one.

4. Since A has three linearly independent eigenvectors $\Longrightarrow A$ is diagonalizable. The matrix P is formed by choosing the eigenvectors of A as column vectors. i.e.

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

It can be easily verified that

$$P^{-1}AP = D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

5. Since minimal polynomial and characteristic polynomials have same set of irreducible factors, we start with considering the product of $(A-3I)(A-5I)$ as $\lambda_1 = 3$ and $\lambda_3 = 5$ are the distinct roots. It can be easily seen that

$$(A-3I)(A-5I) = \begin{bmatrix} 4-3 & 1 & -1 \\ 2 & 5-3 & -2 \\ 1 & 1 & 2-3 \end{bmatrix} \begin{bmatrix} 4-5 & 1 & -1 \\ 2 & 5-5 & -2 \\ 1 & 1 & 2-5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence the minimal polynomial of A is $m_A(\lambda) = (\lambda-3)(\lambda-5) = \lambda^2 - 8\lambda + 15$

Example: Let $A = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$.

1. Find all eigenvalues of A.
2. Find maximal set S of all linearly independent eigenvalues.
3. Find algebraic and geometric multiplicity of all eigenvalues of A.
4. Is A diagonalizable? If yes, find P such $D = P^{-1}AP$ is diagonalizable.
5. Find the minimal polynomial of A.

Solution: First we find the characteristic polynomial $\Delta(t)$ of A.

1. $\Delta(t) = \begin{vmatrix} 3-\lambda & -1 & 1 \\ 7 & -5-\lambda & 1 \\ 6 & -6 & 2-\lambda \end{vmatrix}$
$= (3-\lambda)\begin{vmatrix} -5-\lambda & 1 \\ -6 & 2-\lambda \end{vmatrix} + 1\begin{vmatrix} 7 & 1 \\ 6 & 2-\lambda \end{vmatrix} + 1\begin{vmatrix} 7 & -5-\lambda \\ 6 & -6 \end{vmatrix}$
$= (3-\lambda)\{(-5-\lambda)(2-\lambda) - 1(-6)\} + (7\times(2-\lambda) - (1)\times 6) + (7\times(-6) - (-5-\lambda)\times 6)$
$= (3-\lambda)(-10+5\lambda-2\lambda+\lambda^2+6) + (14-7\lambda-6) + (-42+30+6\lambda)$
$= (3-\lambda)(-4+3\lambda+\lambda^2) + 8 - 7\lambda - 12 + 6\lambda$
$= -12 + 9\lambda + 3\lambda^2 + 4\lambda - 3\lambda^2 - \lambda^3 - 4 - \lambda$
$= -\lambda^3 + 12\lambda - 16$

Now, $\Delta(t) = 0 \Longrightarrow -\lambda^3 + 11\lambda - 16 = 0 \Longrightarrow (\lambda-2)^2(\lambda+4) = 0 \Longrightarrow \lambda_1 = 2, \lambda_2 = 2, \lambda_3 = -4$ are the eigenvalues of matrix A.

2. To find eigenvector corresponding to the eigenvalue $\lambda_1 = 2$ we solve the system $(A-\lambda_1)x = 0$ i.e.

$$\begin{bmatrix} 3-\lambda_1 & -1 & 1 \\ 7 & -5-\lambda_1 & 1 \\ 6 & -6 & 2-\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Longrightarrow \begin{bmatrix} 3-2 & -1 & 1 \\ 7 & -5-2 & 1 \\ 6 & -6 & 2-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Longrightarrow \begin{array}{ccccccc} x_1 & - & x_2 & + & x_3 & = & 0 \\ 7x_1 & - & 7x_2 & + & x_3 & = & 0 \\ 6x_1 & - & 6x_2 & & & = & 0 \end{array}$$
$$\Longrightarrow \begin{array}{ccccccc} x_1 & - & x_2 & + & x_3 & = & 0 \\ & & & & x_3 & = & 0 \end{array}$$

The system has only one independent solution; for example, $x = 1, y = 1, z = 0$. Thus $u = (1,1,0)$ form a basis for the eigenspace of $\lambda_1 = 2$.

Now, to find eigenvector corresponding to the eigenvalue $\lambda_3 = -4$ we solve the system $(A-\lambda_3)x = 0$ i.e.

$$\begin{bmatrix} 3-\lambda_2 & -1 & 1 \\ 7 & -5-\lambda_2 & 1 \\ 6 & -6 & 2-\lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 3+4 & -1 & 1 \\ 7 & -5+4 & 1 \\ 6 & -6 & 2+4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{array}{rcrcrcl} 7x_1 & - & x_2 & + & x_3 & = & 0 \\ 7x_1 & - & x_2 & + & x_3 & = & 0 \\ 6x_1 & - & 6x_2 & + & 6x_3 & = & 0 \end{array}$$

$$\Longrightarrow \begin{array}{rcrcrcl} x_1 & - & x_2 & + & x_3 & = & 0 \\ 6x_2 & - & 6x_3 & = & 0 \end{array}$$

The system has only one independent solution, for example $x_1 = 0, x_2 = 1, x_3 = 1$. Thus $v = (0, 1, 1)$ forms a basis for the eigenspace of $\lambda_2 = -4$

3. Since $\lambda_1 = \lambda_2 = 2$ is a twice repeated root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_1 = 2$ is one. There is one eigenvector corresponding to the eigenvalue $\lambda_1 = 2 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_1 = 2$ is one.
   Also $\lambda_2 = -4$ is a root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_2 = -4$ is one. There is one eigenvector corresponding to the eigenvalue $\lambda_2 = -4 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_2 = -4$ is one.

4. Since A has at most two linearly independent eigenvectors. A is not similar to a diagonal matrix. $\Longrightarrow A$ is not diagonalizable.

5. Similar to previous example, considering the product of irreducible factors $(A - \lambda_1 I)(A - \lambda_3 I) = (A - 2I)(A - 4I)$ does not yield the null matrix of order three. Hence we need to consider the $(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) = (A - 2I)(A - 2I)(A + 4I)$ which would surely yield the yield the null matrix of order three (being all factors used to form the characteristic polynomial). Hence in this case both the characteristic polynomial and minimal polynomial are same and $m_A(\lambda) = \Delta_A(\lambda) = (\lambda - 2)(\lambda - 2)(\lambda + 4) = \lambda^3 - 12\lambda + 16$

Example: Let $A = \begin{bmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{bmatrix}$.

1. Find all eigenvalues of A.
2. Find maximal set S of all linearly independent eigenvalues.
3. Find algebraic and geometric multiplicity of all eigenvalues of A.
4. Is A diagonalizable? If yes, find P such $D = P^{-1}AP$ is diagonalizable.
5. Find the minimal polynomial of A.

Solution: First we find the characteristic polynomial $\Delta(t)$ of A.

1. $\Delta(t) = \begin{vmatrix} 11-\lambda & -8 & 4 \\ -8 & -1-\lambda & -2 \\ 4 & -2 & -4-\lambda \end{vmatrix}$
   $= (11-\lambda)\begin{vmatrix} -1-\lambda & -2 \\ -2 & -4-\lambda \end{vmatrix} + 8\begin{vmatrix} -8 & -2 \\ 4 & -4-\lambda \end{vmatrix} + 4\begin{vmatrix} -8 & -1-\lambda \\ 4 & -2 \end{vmatrix}$
   $= (11-\lambda)\{(-1-\lambda)(-4-\lambda) + 2(-2)\} + 8(-8 \times (-4-\lambda) - (-2) \times 4) + 4(-8 \times (-2) - 4(-1-\lambda))$
   $= (11-\lambda)(4 + \lambda + 4\lambda + \lambda^2 - 4) + 8(32 + 8\lambda + 8) + 4(20 + 4\lambda)$
   $= (11-\lambda)(5\lambda + \lambda^2) + 8(40 + 8\lambda) + 4(1 + \lambda)$
   $= (55\lambda + 11\lambda^2 - 5\lambda^2 - \lambda^3) + (320 + 64\lambda) + (80 + 16\lambda)$
   $= -\lambda^3 + 6\lambda^2 + 135\lambda + 400$
   Now, $\Delta(t) = 0 \Longrightarrow -\lambda^3 + 6\lambda^2 + 135\lambda + 400 = 0 \Longrightarrow (\lambda+5)^2(\lambda - 16) = 0 \Longrightarrow \lambda_1 = -5, \lambda_2 = -5, \lambda_3 = 16$ are the eigenvalues of matrix A.

2. To find eigenvector corresponding to the eigenvalue $\lambda_1 = \lambda_2 = -5$ we solve the system $(A - \lambda_1)x = 0$ i.e.

$$\begin{bmatrix} 11-\lambda_1 & -8 & 4 \\ -8 & -1-\lambda_1 & -2 \\ 4 & -2 & -4-\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 11+5 & -8 & 4 \\ -8 & -1+5 & -2 \\ 4 & -2 & -4+5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 16 & -8 & 4 \\ -8 & 4 & -2 \\ 4 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{array}{rcrcrcl} 16x_1 & - & 8x_2 & + & 4x_3 & = & 0 \\ -8x_1 & + & 4x_2 & - & 2x_3 & = & 0 \\ 4x_1 & - & 2x_2 & + & x_3 & = & 0 \end{array}$$

Since the last two equations are scalar multiple of the first equation.

$$\Longrightarrow 16x_1 - 8x_2 + 4x_3 = 0$$

Since above is one equation in three unknowns, to solve it we choose $(3-1) = 2$ variable, say x_1 and x_2 arbitrarily.

$$\Longrightarrow x_3 = -4x_1 + 2x_2$$

$$\Longrightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ -4x_1+2x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ -4x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 2x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Hence $\begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ are two linearly independent eigenvectors corresponding to the eigenvalue $\lambda_1 = \lambda_2 = -5$

Now, to find eigenvector corresponding to the eigenvalue $\lambda_3 = 16$ we solve the system $(A - \lambda_1)x = 0$ i.e.

$$\begin{bmatrix} 11-\lambda_3 & -8 & 4 \\ -8 & -1-\lambda_3 & -2 \\ 4 & -2 & -4-\lambda_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 11-16 & -8 & 4 \\ -8 & -1-16 & -2 \\ 4 & -2 & -4-16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} -5 & -8 & 4 \\ -8 & -17 & -2 \\ 4 & -2 & -20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \begin{array}{rcrcrcl} -5x_1 & - & 8x_2 & + & 4x_3 & = & 0 \\ -8x_1 & - & 17x_2 & - & 2x_3 & = & 0 \\ 4x_1 & - & 2x_2 & - & 20x_3 & = & 0 \end{array}$$

Above system of equation is not linearly independent (determinant of coefficient matrix is zero), since the third equation is a linear combination of the first and second equations (two times second equation minus 4 times first equation).

$$\Longrightarrow \begin{array}{rcrcrcl} -5x_1 & - & 8x_2 & + & 4x_3 & = & 0 \\ -8x_1 & - & 17x_2 & - & 2x_3 & = & 0 \end{array}$$

Since above is a system of two equations in three unknowns, to solve it we choose $(3-2) = 1$ variable, say x_3 arbitrarily.

$$\Longrightarrow \begin{array}{rcrcr} -5x_1 & - & 8x_2 & = & -4x_3 \\ -8x_1 & - & 17x_2 & = & 2x_3 \end{array} \Longrightarrow x_1 = 4x_3 \Longrightarrow x_2 = -2x_3$$

$$\Longrightarrow x = \begin{bmatrix} 4x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$$

Hence $\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_3 = 16$. Hence $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} \right\}$

3. Since $\lambda_1 = \lambda_2 = -5$ i.e. -5 is twice repeated root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_1 = \lambda_2 = -5$ is two. There are two linearly independent eigenvectors corresponding to the eigenvalue $\lambda_1 = \lambda_2 = -5 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_1 = \lambda_2 = -5$ is two.
   Also $\lambda_3 = 16$ is a root of the characteristic equation $\Longrightarrow$ algebraic multiplicity of eigenvalue $\lambda_3 = 16$ is one. There is one eigenvector corresponding to the eigenvalue $\lambda_3 = 16 \Longrightarrow$ geometric multiplicity of eigenvalue $\lambda_3 = 16$ is one.

4. Since A has three linearly independent eigenvectors $\Longrightarrow A$ is diagonalizable. The matrix P is formed by choosing the eigenvectors of A as column vectors. i.e.
$$P = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -2 \\ -4 & 2 & 1 \end{bmatrix}.$$
   It can be easily verified that
$$P^{-1}AP = D = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 16 \end{bmatrix}.$$

5. Like last example, in this case also the minimal polynomial and characteristic polynomial are same. $m_A(\lambda) = \Delta_A(\lambda) = \lambda^3 - 6\lambda^2 - 135\lambda - 400$.

Remark: Since A is symmetric, instead of finding P one may find an orthogonal matrix Q also such that $Q^{-1}AQ = Q^TAQ = D$. To do so Q is formed by choosing mutually **orthonormal** eigenvectors of A. Since A is symmetric, eigenvectors corresponding to distinct eigenvalues would always be orthogonal. To find mutually orthogonal eigenvectors (corresponding to same eigenvalue $\lambda_1 = \lambda_2 = -5$) from two linearly independent eigenvectors $u_1 = \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ one may use Gram-Schmidt process as follows.
we know that $u_3 = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$.

Choosing $v_1 = u_1$. Then v_2 is given by
$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \frac{-8}{17} \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix} = \begin{bmatrix} \frac{8}{17} \\ 1 \\ \frac{2}{17} \end{bmatrix}$$
Its important to notice that v_2 is not only orthogonal to u_1 but u_3 also.
Hence $v_3 = u_3$ and the matrix Q is formed by taking normalized vectors v_1, v_2 and v_3 as the columns of Q.
$\hat{v}_1 = \frac{v_1}{||v_1||} = \frac{1}{\sqrt{17}} \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix}$
$\hat{v}_2 = \frac{v_3}{||v_3||} = \frac{1}{\sqrt{357}} \begin{bmatrix} 8 \\ 17 \\ 2 \end{bmatrix}$
$\hat{v}_3 = \frac{v_3}{||v_3||} = \frac{1}{\sqrt{21}} \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$
$Q = \begin{bmatrix} \frac{1}{\sqrt{17}} & \frac{8}{\sqrt{357}} & \frac{4}{\sqrt{21}} \\ 0 & \frac{17}{\sqrt{357}} & -\frac{2}{\sqrt{21}} \\ -\frac{4}{\sqrt{17}} & \frac{2}{\sqrt{357}} & \frac{1}{\sqrt{21}} \end{bmatrix}$ **DEFINITION**

Hermitian Matrix: An n-square complex matrix $A = (a_{ij})$ is said to be a Hermitian matrix if $A^* = A$; that is, $a_{ij} = \overline{a_{ij}}$ for all i and j. In case of real matrices, we say that A is real symmetric if $A^T = A$.

Skew Hermitian Matrix A square matrix A is skew-Hermitian if $A^* = -A$, equivalently $a_{ij} = -a_{ij}$ It is immediate that the entries on the main diagonal of a Hermitian matrix are necessarily real. A matrix A is skew-Hermitian if and only if iA is Hermitian.

Unitary Matrices A square matrix $Q \in \mathbb{C}^{n \times n}$ is unitary (in real case we also call it orthogonal) if $Q^* = Q^{-1}$ i.e $Q^*Q = I$. Let $q_1, q_2, \ldots q_n$ are the n columns of unitary matrix then we have $q_i^* q_j = 0$ for all i, j such that $i \neq j$ and $q_i^* q_i = 1$ for all i.

THEOREM: All the eigen values of an hermitian matrix are real.

PROOF: Let A be an hermitian matrix, hence $A^* = A$. Let λ be an eigen value corresponding to eigen vector x of A. Then we have
$$Ax = \lambda x \tag{1}$$

$$(Ax)^* = (\lambda x)^*$$

$$x^* A^* = x^* \overline{\lambda} \tag{2}$$

where, since λ is a scalar, $\overline{\lambda}$ denotes its complex conjugate. Multiplying x on both sides of equation (2) from right, we have

$$x^* A^* x = x^* x \overline{\lambda} \tag{3}$$

Now multiplying x^* on both sides of equation (1) from left, we have

$$x^* A x = \lambda x^* x \tag{4}$$

Now since $A^* = A$, noting this fact and then Comparing equations 3 and 4 we get

$$\lambda x^* x = x^* x \overline{\lambda}$$

Now, $x^* x$ is a non-zero scalar. Therefore we have proved that

$$\lambda = \overline{\lambda}$$

Hence, all the eigen values of an hermitian matrix are real.

THEOREM: Prove that the eigen vectors corresponding to distinct eigen values of Hermitian matrix are orthogonal.

PROOF: Let A be an hermitian matrix, hence $A^* = A$. Let λ_i be an eigen value corresponding to eigen vector x_i of A for all $i = 1 \ldots n$. Consider any i, j such that $i \neq j$, then we have

$$Ax_i = \lambda_i x_i \tag{5}$$

$$Ax_j = \lambda_j x_j \tag{6}$$

Multiplying x_j^* on both sides of eq. 5 we get

$$\begin{aligned} x_j^*(Ax_i) &= x_j^*(\lambda_i x_i) \\ &= \lambda_i (x_j^* x_i) \end{aligned} \tag{7}$$

Multiplying x_i^* on both sides of eq. 6 we get

$$\begin{aligned} x_i^* A^* x_j &= x_i^* A x_j \qquad \text{Using the fact that } A = A^* \\ &= x_i^*(\lambda_j x_j) \\ &= \lambda_j (x_i^* x_j) \end{aligned} \tag{8}$$

Now, $x_i^* A^* x_j = x_i^* A x_j = x_j^* A x_i$, so we have,

$$\begin{gathered} \lambda_i (x_i^* x_j) = \lambda_j (x_i^* x_j) \\ (\lambda_i - \lambda_j)(x_i^* x_j) = 0 \end{gathered}$$

But from the fact that eigen values of A are distinct and nonzero, difference $\lambda_i - \lambda_j$ can never be zero. Therefore, $x_i^* x_j = 0 \quad \forall i, j$ such that $i \neq j$ which proves that eigen vectors are orthogonal.

THEOREM: Eigen value of skew Hermitian matrix is either purely imaginary, or zero.

PROOF: Let A be a skew hermitian matrix, hence $A^* = -A$. Let λ be an eigen value corresponding to eigen vector x of A. Then we have

$$Ax = \lambda x \tag{9}$$

$$(Ax)^* = (\lambda x)^*$$

$$x^* A^* = x^* \overline{\lambda} \tag{10}$$

where, since λ is a scalar, $\overline{\lambda}$ denotes its complex conjugate. Multiplying x on both sides of equation (2) from right, we have

$$x^* A^* x = x^* x \overline{\lambda} \tag{11}$$

Now multiplying x^* on both sides of equation (1) from left, we have

$$x^* A x = \lambda x^* x \tag{12}$$

Now since $A^* = -A$, noting this fact and then Comparing equation 11 and 12 we get

$$\lambda x^* x = -x^* x \overline{\lambda}$$

Now, x^*x is a non-zero scalar. Therefore we have proved that

$$\lambda = -\overline{\lambda}$$

Clearly, λ is zero or purely imaginary.

THEOREM: The eigen values of unitary matrix all have absolute values 1.

PROOF: Let λ be the eigen value corresponding to normalized eigen vector x ($\parallel x \parallel = 1$) of an unitary matrix Q.

$$\begin{aligned}
|\lambda| &= |\lambda| \parallel x \parallel && \text{Since } \parallel x \parallel = 1 \\
&= \parallel \lambda x \parallel && \\
&= \parallel Qx \parallel && \text{Since } Qx = \lambda x \\
&= \parallel x \parallel && \text{since } \parallel x \parallel = \parallel Qx \parallel \\
&= 1 && \text{As assumed } \parallel x \parallel = 1
\end{aligned}$$