Chapter 7

Matrix Eigenfunctions

One of the most important topic in linear algebra is determination of eigenvalues and eigenvectors. For a square matrix, eigenfunctions (eigenvalue and eigenvector) plays a significant role in the field of Applied Mathematics, Applied Physics, Economics, Astronomy, Engineering and Statistics. The analysis of electrical circuit, small oscillations, frequency analysis in digital system, etc. can be done with the help of eigenvalues and eigenvectors. These are useful in the study of canonical forms of a matrix under similarity transformations and in the study of quadratic forms, especially the extrema of quadratic form.

7.1 Matrix Polynomial

If the elements of a matrix A be polynomial in x with degree n atmost, then,

$$A = x^n A_0 + x^{n-1} A_1 + \cdots + x A_{n-1} + A_n, \qquad (7.1)$$

where A_i are the square matrices of the same order as that of A. Such a polynomial (7.1) is called *matrix polynomial* of degree n, provided the leading co-efficient $A_0 \neq \underline{0}$. The symbol x is called *indeterminate*. For example,

$$\begin{bmatrix} 1+x & x^2-1 & 1 \\ 2 & x^2+x+2 & 2 \\ x^2+3 & x & x^2+5 \end{bmatrix} = x^2 \begin{bmatrix} 0&1&0 \\ 0&1&0 \\ 1&0&1 \end{bmatrix} + x \begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&1&0 \end{bmatrix} + \begin{bmatrix} 1&-1&1 \\ 2&2&2 \\ 3&0&5 \end{bmatrix}$$

$$= A_0 x^2 + A_1 x + A_2,$$

where the coefficients A_0, A_1, A_2 are real matrices of order 3×3 as of A is a matrix polynomial of degree 2. We say that the metric polynomial is r−rowed, if the order of each of the matrix coefficients $A_i; i = 1, 2, \cdots, n$ be r. Two matrix polynomials are said to be equal, if and only if the coefficients of the like powers of the indeterminate x are same.

7.1.1 Polynomials of Matrices

Let us consider a polynomial $f(x) = c_0 x^n + c_1 x^{n-1} + \cdots + c_{n-1} x + c_n$ over a field F. Let A be a given square matrix, then we define,

$$f(A) = c_0 A^n + c_1 A^{n-1} + \cdots + c_{n-1} A + c_n I, \qquad (7.2)$$

where I is an unit matrix of the same order as A, is a *polynomial of matrix* A. A polynomial $f(x)$ is said to annihilate A if $f(A) = \underline{0}$, the zero matrix. Let f and g be polynomials. For any square matrix A and scalar k,

(i) $(f+g)(A) = f(A) + g(A)$

(ii) $(fg)(A) = f(A)g(A)$

(iii) $(kf)(A) = kf(A)$

(iv) $f(A)g(A) = g(A)f(A)$

(iv) tells us that any two polynomials in A commute.

7.1.2 Matrices and Linear Operator

Let $T : V \to V$ be a linear operator on a vector space V. Powers of T are defined by the composition operation, i.e.

$$T^2 = T.T, \quad T^3 = T^2.T, \ldots$$

Also, for a polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$, we define $f(T)$ like matrices as

$$f(T) = a_n T^n + \vdots + a_1 T + a_0 I$$

where I is the identity mapping. We say that T is zero or root of $f(x)$ if $f(T) = \underline{0}$, the zero mapping. Suppose A is a matrix representation of a linear operator T. Then $f(A)$ is the matrix representation of $f(T)$, and, in particular, $f(T) = \underline{0}$ if and only if $f(A) = \underline{0}$.

Ex 7.1.1 *Let* $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$. *Find* $f(A)$, *where* $f(x) = x^2 - 3x + 7$ *and* $f(x) = x^2 - 6x + 13$.

Solution: Given that, $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$, therefore,

$$A^2 = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix}.$$

For $f(x) = x^2 - 3x + 7$ the value of $f(A) = A^2 - 3A + 7I$ becomes,

$$f(A) = \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix} - 3\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 12 & 9 \end{pmatrix}.$$

For $f(x) = x^2 - 6x + 13$ the value of $f(A) = A^2 - 6A + 13I$ becomes,

$$f(A) = \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix} - 6\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 13\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In the second case, A is the root of $f(x)$.

7.2 Characteristic Polynomial

Characteristic polynomial of a matrix

If $A = [a_{ij}]_{n\times n}$ be a given square matrix of order n over the field F, then, an ordinary polynomial in λ of the n^{th} degree with scalar coefficients as

$$\chi_A(\lambda) = |A - \lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}, \quad (7.3)$$

where, I is the unit matrix of order n, is defined as *characteristic polynomial or characteristic matrix* of A. The equation

$$\chi_A(\lambda) = |A - \lambda I| = 0 \quad (7.4)$$

is defined as the characteristic equation of the matrix A. The degree of the characteristic equation is the same as the order of the square matrix A. Let us write it as $\chi_A(\lambda) = C_0\lambda^n + C_1\lambda^{n-1} + C_2\lambda^{n-2} + \cdots + C_{n-1}\lambda + C_n = 0$

where the coefficients C_i are functions of the elements of A. It can be shown that

$$C_0 = (-1)^n; \quad C_1 = (-1)^{n-1}\sum_{i=1}^{n} a_{ii},$$

$$C_2 = (-1)^{n-2} \times \text{sum of the principle minors of order 2},$$

and so on. Lastly, $C_n = det\ A$. Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, be a matrix of order 2, then the characteristic polynomial of A is,

$$\begin{aligned}\chi_A(\lambda) = |A - \lambda I| &= \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) \\ &= \lambda^2 - tr(A)\ \lambda + |A|,\end{aligned}$$

where, $tr(A)$ denotes the trace of A, i.e., the sum of the diagonal elements of A. Similarly, for a matrix of order 3, the characteristic polynomial is

$$\begin{aligned}\chi_A(\lambda) = |A - \lambda I| &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} \\ &= \lambda^3 - tr(A)\ \lambda^2 + (A_{11}+A_{22}+A_{33})\lambda - |A|,\end{aligned}$$

where, A_{11}, A_{22}, A_{33} are the cofactors of a_{11}, a_{22}, a_{33} respectively. In general, if A be a square matrix of order n, then the characteristic polynomial is

$$\begin{aligned}\chi_A(\lambda) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \\ &= \lambda^n - S_1\lambda^{n-1} + S_2\lambda^{n-2} - \cdots + (-1)^n S_n\end{aligned}$$

where S_k is the sum of the principal minors of A of order k.

Ex 7.2.1 *Find the characteristic polynomial of* $A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix}$.

Solution: The characteristic polynomial of A is given by,

$$\begin{aligned}\chi_A(\lambda) = |A - \lambda I| &= \begin{vmatrix} 1-\lambda & 3 \\ 4 & 5-\lambda \end{vmatrix} \\ &= (\lambda - 1)(\lambda - 5) - 12 = \lambda^2 - 6\lambda - 7.\end{aligned}$$

Also, $tr(A) = 1 + 5 = 6$ and $|A| = -7$, so, $\chi_A(\lambda) = \lambda^2 - 6\lambda - 7$.

7.2.1 Eigen Value

If the matrix A be of order n, then the characteristic equation of A is an n^{th} degree equation in λ. The roots of

$$\chi_A(\lambda) = |A - \lambda I| = 0 \tag{7.5}$$

are defined as characteristic roots or latent roots or eigen values of the square matrix A. The spectrum of A is the set of distinct characteristic roots of A. Thus, if $A = [a_{ij}]_{n\times n}$, then the eigen values of the matrix A is obtained from the characteristic equation

$$|A - \lambda I| = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \cdots + p_n = 0,$$

where $p_1, p_2, \cdots, p_n$ can be expressed in terms of elements a_{ij}'s of the matrix $A = [a_{ij}]_{n\times n}$ over F. Clearly $\chi_A(\lambda)$ is a monic (i.e., the leading coefficients is 1) polynomial of degree n, since the heighest power of λ occurs in the term $\prod_{i=1}^{n}(\lambda - a_{ii})$ in $\chi_A(\lambda)$. So by fundamental theorem of algebra $\chi_A(\lambda)$ has exactly n (not necessarily distinct) roots. We usually denote the characteristic roots of A by $\lambda_1, \lambda_2, \cdots \lambda_n$, so that

$$\chi_A(\lambda) = |A - \lambda I| = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

Ex 7.2.2 *The characteristic roots of a 3×3 matrix are known to be in arithmetic progression. Determine them, given* $tr(A) = 15$ *and* $|A| = 80$.

Solution: Let the characteristic roots of the matrix A be $a - d, a, a + d$. Then,

$$(a - d) + a + (a + d) = 15 \Rightarrow a = 5.$$

$$\text{Also,}\quad (a - d)a(a + d) = 80 \Rightarrow (a^2 - d^2)a = 80 \Rightarrow d = 3.$$

Therefore, the characteristic roots of the matrix are $2, 5, 8$.

7.2.2 Eigen Vector

Let us consider, the matrix equation

$$AX = \lambda X,\ \text{i.e.,}(A - \lambda I)X = \underline{0}, \tag{7.6}$$

where $A = [a_{ij}]_{n\times n}$ is a given $n \times n$ matrix and $X = [x_1, x_2, \cdots, x_n]^T$ is a $n \times 1$ column non null matrix and λ is scalar as

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Thus corresponding to each eigen value λ_i of the matrix A, there is a non null solution $(A - \lambda_i I)X = \underline{0}$. If $X = X_i$ be the corresponding non null solution, then the column vector X_i is defined as eigen or inverient or characteristic or latent vector or pole. Determination of scalar λ and the non-null vector X, satisfying $AX = \lambda X$, is known as the eigen value problem.

Ex 7.2.3 *Determine the eigen values and eigen vector of the matrix* $\begin{pmatrix} 2&2&1 \\ 1&3&1 \\ 1&2&2 \end{pmatrix}$.

Solution: The characteristic equation of the given matrix A is $|A - \lambda I| = 0$, i.e.,

$$\begin{vmatrix} 2-\lambda & 2 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$\text{or,}\quad (2-\lambda)(\lambda^2 - 5\lambda + 4) + 3(\lambda - 1) = 0$$

$$\text{or,}\quad (\lambda - 1)^2(\lambda - 5) = 0 \Rightarrow \lambda = 1, 1, 5.$$

Thus the eigen values of the given matrix are $1, 1, 5$ and 1 is an 2 fold eigen value of the matrix A. The spectrum of A is $1, 1, 5$. Corresponding to $\lambda = 1$, consider the equation $(A - I)X = \underline{0}$, where A is the given matrix and $X = [x_1, x_2, x_3]^T$. The coefficient matrix is given by

$$\begin{pmatrix} 1&2&1 \\ 1&2&1 \\ 1&2&1 \end{pmatrix} \sim \begin{pmatrix} 1&2&1 \\ 0&0&0 \\ 0&0&0 \end{pmatrix}.$$

The system of equation is equivalent to $x_1 + 2x_2 + x_3 = 0$. We see that $[1, 0, -1]^T$ is one of the non null column solution, which is a eigen vector corresponding to the eigen value $\lambda = 1$. For $\lambda = 5$, the coefficient matrix is given by

$$\begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{pmatrix} \sim \begin{pmatrix} 0 & 8 & -8 \\ 0 & -4 & 4 \\ 1 & 2 & -3 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

The system of equation is equivalent to $-x_2 + x_3 = 0, x_1 - x_3 = 0$ so that $x_3 = 1$ gives $x_1 = x_2 = 1$. Hence $[1, 1, 1]^T$ is a eigen vector corresponding to the eigen value $\lambda = 5$.

Cayley-Hamilton theorem

Theorem 7.2.1 *Every square matrix satisfies its characteristic equation.*

Proof: Let A be a matrix of order n and I be the unit matrix of the same order. Its characteristic equation is $|A - \lambda I| = 0$, i.e.,

$$\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n = 0;\quad a_i = \text{ scalars}$$

$$\Rightarrow |A - \lambda I| = (-1)^n\{\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n\}.$$

We are to show that $A^n + a_1A^{n-1} + a_2A^{n-2} + \cdots + a_{n-1}A + a_nI = \underline{0}$. Now, cofactors of the elements of the matrix $A - \lambda I$ are polynomial in λ of degree at most $(n-1)$, so $adj(A - \lambda I)$ is a matrix polynomial in λ of degree $(n-1)$ as,

$$adj(A - \lambda I) = \lambda^{n-1}A_1 + \lambda^{n-2}A_2 + \cdots + \lambda A_{n-1} + A_n,$$

where A_i are suitable matrices of order n, each of which will contain terms with same powers of λ. Now, using the relation, $(A - \lambda I)adj(A - \lambda I) = |A - \lambda I|I$, we get,

$$(A - \lambda I)[\lambda^{n-1}A_1 + \lambda^{n-2}A_2 + \cdots + \lambda A_{n-1} + A_n]$$

$$= (-1)^n\{\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n\}I.$$

This relation is true for all values of λ, so, equating coefficients of like powers $\lambda^n, \lambda^{n-1}, \cdots$ of λ from both sides, we get,

$$-A_1 = (-1)^n I$$
$$AA_1 - IA_2 = (-1)^n a_1 I$$
$$AA_2 - A_3 = (-1)^n a_2 I$$
$$\vdots$$
$$AA_n = (-1)^n a_n I.$$

Multiplying these relations successively by $A^n, A^{n-1}, \cdots, A, I$ respectively and adding we get,

$$\underline{0} = (-1)^n[A^n + a_1A^{n-1} + a_2A^{n-2} + \cdots + a_{n-1}A + a_nI]$$
$$\Rightarrow A^n + a_1A^{n-1} + a_2A^{n-2} + \cdots + a_{n-1}A + a_nI = \underline{0}. \qquad (7.7)$$

Thus every matrix A is a root of its characteristic polynomial.

Deduction 7.2.1 Now, if $|A| \neq 0$, then A^{-1} exists. In this case, multiplying (7.7) by A^{-1}, we get,

$$A^{n-1} + a_1A^{n-2} + a_2A^{n-3} + \cdots + a_{n-1}I + a_nA^{-1} = \underline{0}$$
$$\Rightarrow A^{-1} = -\frac{1}{a_n}A^{n-1} - \frac{a_1}{a_n}A^{n-2} - \frac{a_2}{a_n}A^{n-3} - \cdots - \frac{a_{n-1}}{a_n}I.$$

Therefore, Cayley-Hamilton theorem can be applied to find the inverse of a matrix.

Result 7.2.1 Suppose $A = [a_{ij}]$ be a triangular matrix. Then $A - \lambda I$ is a triangular matrix with diagonal entries $a_{ii} - \lambda$, and hence, the characteristic polynomial is

$$|A - \lambda I| = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).$$

Result 7.2.2 Suppose the characteristic polynomial of an n square matrix A is a product of n distinct factors, then A is similar to the diagonal matrix $D = diag(a_{11}, a_{22}, \cdots, a_{nn})$.

Ex 7.2.4 *What are the possible eigen values of a square matrix A (over the field $\Re$) satisfying $A^3 = A$?*

Solution: According to Cayley-Hamilton theorem, every square matrix satisfies its own characteristic equation, so $A^3 = A$ becomes

$$\lambda^3 = \lambda \Rightarrow \lambda(\lambda^2 - 1) = 0 \Rightarrow \lambda = -1, 0, 1.$$

Ex 7.2.5 *Verify Cayley-Hamilton theorem for $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ and hence find A^{-1} and A^6.*

Solution: The characteristic equation of the given matrix A is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 3-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 4\lambda + 4 = 0.$$

By Cayley-Hamilton theorem, we have $A^2 - 4A + 4I = \underline{0}$. Now,

$$\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} - 4\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 4 \\ -4 & 8 \end{pmatrix} - \begin{pmatrix} 4 & 4 \\ -4 & 12 \end{pmatrix} + \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \underline{0}.$$

so the Cayley-Hamilton theorem is verified. Therefore,

$$A^{-1} = I - \frac{1}{4}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{4}\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$
$$= \frac{1}{4}\left[\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}\right] = \frac{1}{4}\begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}.$$

Now divide λ^6 by $\lambda^2 - 4\lambda + 4$, we get,

$$\lambda^6 = (\lambda^2 - 4\lambda + 4)(\lambda^4 + 4\lambda^3 + 12\lambda^2 + 32\lambda + 80) + 192\lambda - 320$$
$$= 192\lambda - 320 \Rightarrow A^6 = 192A - 320I = \begin{pmatrix} -128 & 192 \\ -192 & 256 \end{pmatrix}.$$

Ex 7.2.6 *If $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$, express $2A^5 - 3A^4 + A^2 - 5I$ as a linear polynomial in A.*

Solution: The characteristic equation of the given matrix A is $|A - \lambda I| = 0$, i.e.,

$$\begin{vmatrix} 3-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 5\lambda + 5 = 0.$$

By Cayley-Hamilton theorem, we have $A^2 - 5A + 5I = \underline{0}$. Now divide $2\lambda^5 - 3\lambda^4 + \lambda^2 - 5$ by $\lambda^2 - 5\lambda + 5$, we get,

$$2\lambda^5 - 3\lambda^4 + \lambda^2 - 5 = (\lambda^2 - 5\lambda + 5)(2\lambda^3 + 7\lambda^2 + 25\lambda + 91) + 330\lambda - 460$$
$$= 330\lambda - 460$$
$$\Rightarrow 2A^5 - 3A^4 + A^2 - 5I = 330A - 460I.$$

Ex 7.2.7 *Verify Cayley-Hamilton theorem for $A = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix}$ and hence find A^{-1}.*

Solution: The characteristic equation of the given matrix A is

$$|A - \lambda I| = 0 \Rightarrow \begin{vmatrix} 0-\lambda & 0 & 1 \\ 3 & 1-\lambda & 0 \\ -2 & 1 & 4-\lambda \end{vmatrix} = 0$$
$$\text{or,} \quad \lambda^3 - 5\lambda^2 + 6\lambda - 5 = 0.$$

By Cayley-Hamilton theorem, we have $A^3 - 5A^2 + 6A - 5I = \underline{0}$. For verification, we have,

$$A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix}$$
$$A^3 = \begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} -5 & 5 & 14 \\ -3 & 4 & 15 \\ -13 & 19 & 51 \end{pmatrix}.$$

Therefore, the expression $A^3 - 5A^2 + 6A - 5I$ becomes,

$$\begin{pmatrix} -5 & 5 & 14 \\ -3 & 4 & 15 \\ -13 & 19 & 51 \end{pmatrix} - 5\begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix} + 6\begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \underline{0}.$$

Thus, Cayley-Hamilton theorem is verified. To find A^{-1}, we get,

$$A^3 - 5A^2 + 6A - 5I = \underline{0}$$
$$\Rightarrow A^{-1} = \frac{1}{5}(A^2 - 5A + 6I) = \frac{1}{5}\begin{pmatrix} 4 & 1 & -1 \\ -12 & 2 & 3 \\ 5 & 0 & 0 \end{pmatrix}.$$

Ex 7.2.8 *Verify Cayley-Hamilton theorem for* $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ *and hence find* A^{-1} *and* A^{50}.

Solution: The characteristic equation of the given matrix A is

$$|A - \lambda I| = 0 \Rightarrow \begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & 0-\lambda & 1 \\ 0 & 1 & 0-\lambda \end{vmatrix} = 0$$
$$\text{or,} \quad \lambda^3 - \lambda^2 - \lambda + 1 = 0.$$

By Cayley-Hamilton theorem, we have $A^3 - A^2 - A + I = \underline{0}$. For verification, using A, the expression $A^3 - A^2 - A + I$ becomes,

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \underline{0}.$$

Hence the Cayley-Hamilton theorem is verified. Now, using the relation $A^3 - A^2 - A + I = \underline{0}$, we have, $A^{-1} = -A^2 + A + I$, i.e.,

$$A^{-1} = -\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

From the relation $A^3 - A^2 - A + I = \underline{0}$, we see that

$$A^3 = A^2 + A - I$$
$$\Rightarrow A^4 = A^3 + A^2 - A = A^2 + A^2 - I$$
$$\Rightarrow A^5 = A^3 + A^3 - A = A^3 + A^2 - I.$$

Thus we get, for every integer $n \geq 3$, we have $A^n = A^{n-2} + A^2 - I$. Using this recurrence relations, we have, $A^4 = A^2 + A^2 - I, A^6 = A^4 + A^2 - I$, i.e.,

$$A^4 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}; \quad A^6 = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

From symmetry, we see that $A^{50} = \begin{pmatrix} 1 & 0 & 0 \\ 25 & 1 & 0 \\ 25 & 0 & 1 \end{pmatrix}$.

Ex 7.2.9 *A matrix A has eigen values* 1 *and* 4 *with corresponding eigenvectors* $(1, -1)^T$ *and* $(2, 1)^T$ *respectively. Find the matrix A.* *[Gate'97]*

Solution: Let the required matrix be $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Then from the equation $AX = \lambda X$, we have,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ -1 \end{pmatrix} \Rightarrow \begin{matrix} a_{11} - a_{12} = 1 \\ a_{21} - a_{22} = -1 \end{matrix}.$$
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = 4\begin{pmatrix} 2 \\ -1 \end{pmatrix} \Rightarrow \begin{matrix} 2a_{11} + a_{12} = 8 \\ 2a_{21} + a_{22} = 4 \end{matrix}.$$

Solving these equations, we get $a_{11} = 3, a_{12} = 2, a_{21} = 1$ and $a_{22} = 2$. Therefore, the matrix is $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$.

Theorem 7.2.2 *If the eigen values of A are distinct, then the eigen vectors are linearly independent.*

Proof: Let x_k be the eigen vector of an $n \times n$ square matrix A corresponding to the eigen value λ_k, where $\lambda_k; k = 1, 2, \cdots, n$ are distinct. Let $x_k = [x_{k1}, x_{k2}, \cdots, x_{kn}]^T$ for $k = 1, 2, \cdots, n$. Thus we have, $Ax_k = \lambda_k x_k$; $k = 1, 2, \cdots, n$. Therefore,

$$A^2 x_k = A(Ax_k) = \lambda_k(Ax_k) = \lambda_k(\lambda_k x_k) = \lambda^2 x_k.$$

By the principle of mathematical induction, we conclude $A^p x_k = \lambda_k^p x_k$, for any positive integer p. Let us consider the relation, $X = c_1x_1 + c_2x_2 + \cdots + c_nx_n = \theta$; where c_i's are scalars. Equating i^{th} component to 0, we get,

$$c_1x_{1i} + c_2x_{2i} + \cdots + c_nx_{ni} = 0,$$

it is true for $i = 1, 2, \cdots, n$. Since $x = \theta$, so $Ax = \theta$. Therefore,

$$A(c_1x_1 + c_2x_2 + \cdots + c_nx_n) = \theta$$
$$\text{or,} \quad c_1(Ax_1) + c_2(Ax_2) + \cdots + c_n(Ax_n) = \theta$$
$$\text{or,} \quad c_1\lambda_1x_1 + c_2\lambda_2x_2 + \cdots + c_n\lambda_nx_n = \theta.$$

Equating i^{th} component to 0, we get,

$$c_1\lambda_1x_{1i} + c_2\lambda_2x_{2i} + \cdots + c_n\lambda_nx_{ni} = 0.$$

Again, equating the i^{th} component of $A^2x = \theta$ to zero, gives

$$c_1\lambda_1^2x_{1i} + c_2\lambda_2^2x_{2i} + \cdots + c_n\lambda_n^2x_{ni} = 0.$$

Continuing this process and lastly we get,

$$c_1\lambda_1^{n-1}x_{1i} + c_2\lambda_2^{n-1}x_{2i} + \cdots + c_n\lambda_n^{n-1}x_{ni} = 0.$$

The n equations in n unknowns have a non null solution, if and only if,

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & & \vdots & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{vmatrix} = 0,$$

which is true if and only if some two λ's are equal, which is contradictory to the hypothesis. Thus the system has no non null solution. Therefore, we must have

$$c_1x_{1i} = c_2x_{2i} = \cdots = c_nx_{ni} = 0; \quad i = 1, 2, \cdots, n.$$

Hence, $c_1x_1 = c_2x_2 = \cdots = c_nx_n = \theta$. But $x_1, x_2, \cdots, x_n$ are non null vectors, since they are eigen vectors, hence $c_1 = c_2 = \cdots = c_n = 0$. This shows that $\{x_1, x_2, \cdots, x_n\}$ is linearly ind

Hence $a_2 = \frac{1}{2} \times [1-2-9] = -5$ and $B_2 = A_2 - a_2 I = \begin{pmatrix} 6 & -18 & -2 \\ -3 & 3 & 1 \\ 0 & 0 & -4 \end{pmatrix}$.

$(iii)\, A_3 = AB_2 = \begin{pmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 6 & -18 & -2 \\ -3 & 3 & 1 \\ 0 & 0 & -4 \end{pmatrix} = \begin{pmatrix} -12 & 0 & 0 \\ 0 & -12 & 0 \\ 0 & 0 & -12 \end{pmatrix}$.

Thus $a_3 = \frac{1}{3}$ Tr $A_3 = -12$ and $B_3 = A_3 - a_3 I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Hence the characteristic polynomial is $\lambda^3 - 6\lambda^2 + 5\lambda + 12 = 0$. The eigenvalues of A are $-1, 3, 4$. To find the eigenvector corresponding to the eigenvalue $\lambda = -1$, let us take $e_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $e_1 = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}$, $e_2 = \begin{pmatrix} -2 \\ 1 \\ -4 \end{pmatrix}$.

From the formula (7.10) we get the results of calculation in the following table

λ_i	I	II	III	$X^{(i)}$
	0	−1	−2	−3
$\lambda_1 = -1$	0	0	1	1
	1	3	−4	0
	0	3	−2	1
$\lambda_2 = 3$	0	0	1	1
	9	−9	−4	−4
	0	4	−2	2
$\lambda_3 = 4$	0	0	1	1
	16	−12	−4	0

7.2.3 Eigen Space

If the eigen values of A are real then the eigen vectors $X_1, X_2, \cdots, X_n \in \Re^n$. The subspace generated by the non null vectors is known as eigen or characteristic space of the matrix A and is denoted by E_λ. If λ is an eigen value of A, then the *algebraic multiplicity* of λ is defined to be the multiplicity of λ as a root of the characteristic polynomial of A, while the *geometric multiplicity* of λ is defined to be the dimension of its eigen space, i.e., $dimE_\lambda$. The geometric multiplicity of an eigen value $\lambda \leq$ the algebraic multiplicity of the eigen value λ. If the geometric multiplicity of an eigen value $\lambda =$ the algebraic multiplicity of the eigen value λ, then λ is said to be regular.

Theorem 7.2.3 *The eigen vector of an $n \times n$ matrix A over a field F corresponding to an eigen value λ of A together with the zero column vector is a subspace of $V_n(F)$.*

Proof: Let E_λ be the set of all eigen vectors of A corresponding to the eigen value λ. Obviously, each vector of E_λ is $n \times 1$ column vector. Let $X_1, X_2 \in E_\lambda$ and $c_1, c_2 \in F$. Then, $AX_1 = \lambda X_1$ and $AX_2 = \lambda X_2$. Now,

$$\begin{aligned} A(c_1X_1 + c_2X_2) &= A(c_1X_1) + A(c_2X_2) = c_1(AX_1) + c_2(AX_2) \\ &= c_1(\lambda X_1) + c_2(\lambda X_2) = \kappa(c_1X_1 + c_2X_2). \end{aligned}$$

It shows that $c_1X_1 + c_2X_2 \in E_\lambda$, if $X_1, X_2 \in E_\lambda$. Hence $E_\lambda \cup \{\theta\}$, where θ is the zero column vector in $V_n(F)$, is a subspace of $V_n(F)$. $E_\lambda \cup \{\theta\}$ is known as the characteristic subspace corresponding to the eigen value λ or eigen space of λ.

Characteristic polynomial of block diagonal matrices

Let A be a block triangular matrix, say, $A = \begin{pmatrix} A_1 & B \\ \underline{0} & A_2 \end{pmatrix}$ where A_1 and A_2 are square matrices. Then $A - \lambda I$ is also a block triangular matrix, with diagonal blocks $A_1 - \lambda I$ and $A_2 - \lambda I$. Thus,

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

Thus the characteristic polynomial of A is the product of the characteristic polynomials of the diagonal blocks A_1 and A_2. In general, let A is a block triangular matrix with diagonal blocks $A_1, A_2, \cdots, A_r$; then the characteristic polynomial of A is

$$|A - \lambda I| = |A_1 - \lambda I|\ |A_2 - \lambda I| \cdots |A_r - \lambda I|.$$

Ex 7.2.11 *Find the characteristic polynomial of the block triangular matrix*

$$A = \left[\begin{array}{cc:cc} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ \hdashline 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{array}\right].$$

Solution: The given block triangular matrix A can be written in the form $A = \begin{bmatrix} A_1 & B \\ \underline{0} & A_2 \end{bmatrix}$. Now, the characteristic polynomial of A_1 and A_2 are

$$|A_1 - \lambda I| = \lambda^2 - 12\lambda + 35 = (\lambda - 5)(\lambda - 7),$$
$$|A_2 - \lambda I| = \lambda^2 - 11\lambda + 30 = (\lambda - 5)(\lambda - 6).$$

Accordingly, the characteristic polynomial of A is

$$|A - \lambda I| = (\lambda - 5)(\lambda - 7)(\lambda - 5)(\lambda - 6) = (\lambda - 5)^2(\lambda - 6)(\lambda - 7).$$

Characteristic polynomial of linear operator

Let $T: V \to V$ be a linear operator on a vector space $V(F)$ with finite dimension. For any polynomial $f(t) = c_0 + c_1t + \cdots + c_nt^n$, let us define

$$f(T) = c_0I + c_1T + \cdots + c_nT^n,$$

where I is the identity mapping and powers of T are defined by the composition operation. The characteristic polynomial of the linear operator T is defined to be the characteristic polynomial of the matrix representation of T. Cayley-Hamilton states that

"A linear operator T is a zero of the characteristic polynomial".

Eigen function : Let $T: V \to V$ be a linear operator on a vector space with finite dimension. A scalar λ is called an eigenvalue of T if $\exists$ a non null vector α such that,

$$T(\alpha) = \lambda\alpha.$$

Every vector satisfying this relation is called an eigen vector of T corresponding to the eigen value λ. If λ is an eigen value of T if $T - \lambda I$ is non singular. The set E_λ, which is the kernal of $T - \lambda I$, of all eigen vectors belonging to an eigen value λ is a subspace of V, called the eigen space of λ.

Note that if A and B are matrix representation of T, then $B = P^{-1}AP$, where P is a change of basis matrix. Thus A and B are similar and they have the same characteristic polynomial. Accordingly, the characteristic polynomial of T is independent of the particular basis in which the matrix representation of T is computed.

Ex 7.2.12 *For the linear operator $T : V \to V$, find all eigen values and a basis for eigenspace $T(x, y) = (3x + 3y, x + 5y)$.*

Solution: The matrix A that represents the linear operator $T : V \to V$ relative to the standard basis of $\Re^2$ as

$$A = [T] = \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix}.$$

The characteristic polynomial of a linear operator is equal to the characteristic polynomial of any matrix A that represents the linear operator. Therefore, the characteristic polynomial for the linear operator $T : V \to V$ is given by

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & 3 \\ 1 & 5-\lambda \end{vmatrix} = \lambda^2 - 8\lambda + 15.$$

Ex 7.2.13 *Let $T : \Re^3 \to \Re^3$ be defined by*

$$T(x, y, z) = (2x + y - 2z, 2x + 3y - 4z, x + y - z).$$

Find all eigen values of T and find a basis of each eigen space.

Solution: The matrix A that represents the linear operator $T : V \to V$ relative to the standard basis of $\Re^2$ is

$$A = [T] = \begin{pmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{pmatrix}.$$

The characteristic polynomial of a linear operator is equal to the characteristic polynomial of any matrix A that represents the linear operator. The characteristic polynomial for the linear operator $T : V \to V$ is given by

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 & -2 \\ 2 & 3-\lambda & -4 \\ 1 & 1 & -1-\lambda \end{vmatrix}$$

$$= \lambda^3 - 4\lambda^2 + 5\lambda - 2 = (\lambda - 1)^2(t - 2).$$

Thus the eigen values of A are 1, 2. Now we find the linearly independent eigenvectors for each eigenvalue of A. Corresponding to $\lambda = 1$, consider the equation $(A - I)X = \underline{0}$, where $X = [x_1, x_2, x_3]^T$. The coefficient matrix is given by

$$A - I = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 2 & -4 \\ 1 & 1 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\Rightarrow x_1 + x_2 - 2x_3 = 0.$$

We see that $[1, -1, 0]^T$ and $[2, 0, 1]^T$ are two linearly independent eigen vector corresponding to the eigen value $\lambda = 1$. Similarly, for $\lambda = 2$, we obtain,

$$A - 2I = \begin{pmatrix} 0 & 1 & -2 \\ 2 & 1 & -4 \\ 1 & 1 & -3 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 0 \\ 1 & 1 & -3 \end{pmatrix}.$$

$$x_1 + x_2 - 3x_3 = 0 \text{ and } x_2 - 2x_3 = 0.$$

We see that $[1, 2, 1]^T$ is a solution and so it is the eigen vector corresponding to the eigen value $\lambda = 2$.

Ex 7.2.14 *For the linear operator $D : V \to V$ defined by $D(f) = \frac{df}{dt}$, where, V is the space of functions with basis $S = \{\sin t, \cos t\}$, find the characteristic polynomial.*

Solution: First we are to find the matrix A representing the differential operator D relative to the basis S. Now,

$$D(\sin t) = \cos t = 0.\sin t + 1.\cos t$$

$$D(\cos t) = -\sin t = (-1).\sin t + 0.\cos t$$

$$\Rightarrow A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Therefore, the characteristic polynomial for the linear operator $D : V \to V$ is given by

$$|A - \lambda I| = \begin{vmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 + 1.$$

7.3 Diagonalization

Diagonalization of a matrix

A given n square matrix A with eigen values $\lambda_1, \lambda_2, \cdots, \lambda_n$ is said to be diagonalisable, if $\exists$ a non singular matrix P such that

$$D = P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \tag{7.12}$$

is diagonal. Thus the $n \times n$ matrix A is said to be diagonalisable, if A is similar to an $n \times n$ diagonal matrix. Below we are to derive a necessary and sufficient for diagonalizability of a matrix A.

Theorem 7.3.1 *An $n \times n$ matrix A over the field F is diagonalisable if and only if A has n linearly independent eigen vectors.*

Proof: First let A is diagonalisable. Then by definition A is similar to a diagonal matrix $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, where the eigen values of A are $\lambda_1, \lambda_2, \cdots, \lambda_n$ and there exists a non-singular matrix P of order n such that

$$A = PDP^{-1}, \text{ i.e., } AP = PD.$$

Let $X_1, X_2, \cdots, X_n$ be n column vectors of P, then,

$$i^{th} \text{ column vector of } AP = i^{th} \text{ column vector of } DP.$$

Therefore, $AX_i = \lambda_i X_i$. Since λ_i is an eigen value of A, this relation $AX_i = \lambda_i X_i$ shows that X_i is the eigen vector of A corresponding to the eigen value λ_i. Therefore, eigen vectors of A are n column vector of P. P is non-singular, so these vectors are LI in $V_n(F)$. Consequently, A has n linearly independent eigen vectors.

Conversely, let $X_1, X_2, \cdots, X_n$ be n linearly independent eigen vectors of A corresponding to the eigen values $\lambda_1, \lambda_2, \cdots, \lambda_n$ respectively. Then, $AX_i = \lambda_i X_i$. Let P be an $n \times n$ matrix whose i^{th} column vector is X_i. Since X_i's are linearly independent in $V_n(F)$, P is non-singular. If D is a diagonal matrix, than,

$$D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n), \text{ so that } AP = PD, \text{ i.e., } A = PDP^{-1}.$$

Consequently, A is similar to D and A is diagonalisable.

Ex 7.3.1 *Diagonalise the matrix* $A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$, *if possible.*

Solution: The characteristic equation of the given matrix A is

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{vmatrix} = 0$$

$$\text{or,} \quad (2+\lambda)^2(4-\lambda) = 0 \Rightarrow \lambda = -2, -2, 4.$$

Thus the eigen values of the given matrix are $-2, -2, 4$ and -2 is an 2-fold eigen value of the matrix A. Corresponding to $\lambda = -2$, consider the equation $(A+2I)X = \underline{0}$, where $X = [x_1, x_2, x_3]^T$. The coefficient matrix is given by

$$A + 2I = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The system of equation is equivalent to $x_1 - x_2 + x_3 = 0$. We see that $[1,1,0]^T$ and $[1,0,-1]^T$ generate the eigen space of the eigen value -2 and they form a basis of the eigen space E_{-2} of -2. For $\lambda = 4$, the coefficient matrix is given by

$$A + 4I = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\Rightarrow x_1 + x_2 - x_3 = 0, 2x_2 - x_3 = 0$$

so that $x_3 = 2$ gives $x_1 = x_2 = 1$. Hence $[1,1,2]^T$ is a eigen vector corresponding to the eigen value $\lambda = 5$. Thus $[1,1,2]^T$ generates the eigen space of the eigen value 4 and they form a basis of the eigen space E_4 of 4. These three vectors $[1,1,0]^T, [1,0,-1]^T$ and $[1,1,2]^T$ are LI, so the given matrix A is diagonalisable and the diagonalising matrix is

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} \text{ so that } P^{-1} = -\frac{1}{2}\begin{pmatrix} 1 & -3 & 1 \\ -2 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}$$

$$\Rightarrow P^{-1}AP = -\frac{1}{2}\begin{pmatrix} 1 & -3 & 1 \\ -2 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

where the diagonal elements are eigen values of A.

Ex 7.3.2 *Show that the matrix* $A = \begin{pmatrix} 3 & -5 \\ 2 & -3 \end{pmatrix}$ *is diagonalizable over the complex field* C.

Solution: The characteristic polynomial of A is

$$|A - \lambda I| = \lambda^2 - (3-3)\lambda + (-9+10) = \lambda^2 + 1.$$

Now, we consider the following two subcases:

(i) A is a matrix over the real field $\Re$, then the characteristic polynomial has no real roots. Thus A has no eigen values and no eigen vectors, and so A is not diagonalizable.

(ii) A is a matrix over the complex field C. Then it has two distinct eigen values i and $-i$.

Therefore, $X_1 = (5, 3-i)^T$ and $X_2 = (5, 3+i)^T$ are the linearly independent eigen vectors of A corresponding to the eigen values i and $-i$ respectively. Thus,

$$P = \begin{pmatrix} 5 & 5 \\ 3-i & 3+i \end{pmatrix} \text{ and } P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

As expected, the diagonal entries in D are the eigen values of A. Therefore, the matrix A is diagonalizable over the complex field C.

Definition 7.3.1 For an r-fold eigenvalue λ of the matrix A, r is called the *algebraic multiplicity* of λ. If k be the number of linearly independent eigenvectors corresponding to an eigenvalue λ then k is the *geometric multiplicity* of λ. The geometric multiplicity of an eigenvalue is less than or equal to its algebraic multiplicity. If the geometric multiplicity of λ is equal to its algebraic multiplicity, then λ is said to be *regular*.

Ex 7.3.3 *Let* $A = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}$. *Find the algebraic and geometric multiplicities of the eigenvalues. Also, diagonalise A, if possible.*

Solution: The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} = 0$$

$$\text{or,} \quad \lambda^2 - 4\lambda + 4 = 0 \Rightarrow (\lambda-2)^2 = 0.$$

Therefore, the eigenvalues are $\lambda = 2, 2$. Hence the algebraic multiplicity of 2 is 2. Let $[x_1, x_2]^T$ be the eigenvector corresponding to 2. Then

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

$$\text{or,} \quad \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\text{or, } x_1 + x_2 = 0. \text{ Let } x_2 = k \text{ then } x_1 = -k.$$

Thus the eigenvectors are $\begin{bmatrix} -k \\ k \end{bmatrix} = k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, i.e., there is only one independent eigenvector corresponding to $\lambda = 2$. So, the geometric multiplicity of the eigenvalue 2 is 1. Since the number of independent eigenvectors is 1, of the matrix A of order 2×2, so A is not diagonasable.

Deduction 7.3.1 Suppose a matrix A can be diagonalized as $P^{-1}AP = D$, where, D is diagonal. Then A has the extremely useful diagonal factorization $A = PDP^{-1}$. Using the factorization, the algebra of A reduces to the algebra of the diagonal matrix D, which can be easily evaluated. Suppose $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then,

$$A^m = \left(PDP^{-1}\right)^m = PD^mP^{-1} = P\ diag(\lambda_1^m, \lambda_2^m, \cdots, \lambda_n^m)\ P^{-1}.$$

More generally, for a polynomial $f(t)$,

$$f(A) = f\left(PDP^{-1}\right) = Pf(D)P^{-1} = Pdiag\left(f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)\right)P^{-1}.$$

Furthermore, if the diagonal entries of D are nonnegative, let .

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

Then B is nonnegative square root of A, i.e., $B^2 = A$ and the eigen values of A are nonnegative.

Ex 7.3.4 *Let* $A = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$, *find* $f(A)$, *where* $f(t) = t^3 - 5t^2 + 3t + 6$ *and* A^{-1}.

Solution: The characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = (\lambda-1)(\lambda-4).$$

Thus the eigen values of A are 1,4. We see that $X_1 = (1,-2)^T$ and $X_2 = (1,1)^T$ are linearly independent eigen vectors corresponding to $\lambda_1 = 1$ and $\lambda_2 = 4$ respectively, and hence form a basis of $\Re^2$. Therefore, A is diagonalisable. Let,

$$P = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \text{ so, } P^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$$

$$\Rightarrow P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = D.$$

Thus the diagonal elements are eigen values of A. Using the diagonal factorization $A = PDP^{-1}$, and $1^4 = 1$ and $4^4 = 256$, we get,

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

Also, $f(1) = 5$ and $f(4) = 2$, hence,

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

Using $\sqrt{1} = 1$ and $\sqrt{4} = 2$, we obtain,

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

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

Ex 7.3.5 *Let* $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$. *(a) Find all eigen values and corresponding eigenvectors.*

(b) Find a nonsingular matrix P such that $D = P{-}1AP$ is diagonal,
(c) Find A^6 and $f(A)$, where $f(t) = t^4 - 3t^3 - 6t^2 + 7t + 3$.
(d) Find a matrix B such that $B^3 = A$ and B has real eigenvectors.

Solution: (a) The characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = (\lambda-1)(\lambda-4).$$

Thus the eigen values of A are 1,4. The eigen vectors corresponding to $\lambda_1 = 1$ and $\lambda_2 = 4$ are $X_1 = (2,-1)^T$ and $X_2 = (1,1)^T$ respectively.
(b) Since the vectors X_1 and X_2 are linearly independent, so A is diagonalisable. Finally, let P be the matrix whose columns are the unit vectors X_1 and X_2 respectively, then,

$$P = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \text{ so, } P^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$$

$$\Rightarrow D = P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix},$$

where the diagonal elements are eigen values of A.
(c) Using the diagonal factorization $A = PDP^{-1}$, and $1^6 = 1$ and $4^6 = 4096$, we get,

$$A^6 = PD^6P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4096 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1366 & 2230 \\ 1365 & 2731 \end{pmatrix}.$$

Also, $f(1) = 2$ and $f(4) = -1$, hence,

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

(d) Here $\begin{pmatrix} 1 & 0 \\ 0 & \sqrt[3]{4} \end{pmatrix}$ is the real cube root of D. Hence the real cube root of A is

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

$$= \frac{1}{3}\begin{pmatrix} 2+\sqrt[3]{4} & -2+2\sqrt[3]{4} \\ -1+\sqrt[3]{4} & 1+2\sqrt[3]{4} \end{pmatrix}.$$

7.3.1 Orthogonal Diagonalisation

A square matrix A is said to be orthogonally diagonalisable, if there exists an orthogonal non-singular matrix P such that

$$P^{-1}AP = \text{ a diagonal matrix.}$$

In this case, P is said to diagonalise A orthogonally.

Theorem 7.3.2 *A square matrix is orthogonally diagonalisable, if and only if it is real symmetric.*

Proof: First let A be orthogonally diagonalisable, then there exists an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix, say $P^{-1}AP = D$, where D is an diagonal matrix. Since P is an orthogonal matrix, we have $P^T = P^{-1}$ and so,

$$A = P^{-1}DP = P^TDP$$
$$\Rightarrow A^T = [P^TDP]^T = P^TD^TP$$
$$= P^TDP = A, \text{ as } D \text{ is diagonal so, } D^T = D,$$

shows that A is an symmetric matrix. Conversely, let A be a real symmetric matrix of order n, the A has n linearly independent eigen vectors. Using Gram-Schmidt process, these n eigen vectors can be converted to a set of linearly indendent orthogonal vectors. This orthogonal set can be normalized to get a set of n orthogonal eigen vectors. Let P be the $n \times n$ matrix whose column vectors are these n orthonormal eigen vectors. Clearly, P is an orthogonal matrix and $P^{-1}AP$ is a diagonal matrix. Thus, A is orthogonally diagonalisable, if it be symmetric.

Ex 7.3.6 *Let* $A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$, *find an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal.*

Solution: The characteristic polynomial of A is

$$|A - \lambda I| = \lambda^2 - (7-1)\lambda + (-7-9) = (\lambda-8)(\lambda+2).$$

Thus the eigenvalues of A are $-2, 8$. The eigen vectors corresponding to $\lambda_1 = -2$ and $\lambda_2 = 8$ are $X_1 = (1, -3)^T$ and $X_2 = (3, 1)^T$ respectively. Since A is symmetric, the eigen vectors X_1 and X_2 are orthogonal. Normalize X_1 and X_2 to obtain, respectively, the unit vectors

$$\hat{X}_1 = \left(\frac{1}{\sqrt{10}}, \frac{-3}{\sqrt{10}}\right) \text{ and } \hat{X}_2 = \left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right).$$

Finally, let P be the matrix whose columns are the unit vectors $\hat{X}_1$ and $\hat{X}_2$ respectively, then,

$$P = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix} \text{ and } P^{-1} = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix}$$

$$\Rightarrow D = P^{-1}AP = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & -2 \end{pmatrix}.$$

As expected, the diagonal entries in D are the eigen values of A.

Ex 7.3.7 *Diagonalise the matrix* $A = \begin{pmatrix} 6 & 4 & -2 \\ 4 & 12 & -4 \\ -2 & -4 & 13 \end{pmatrix}$, *if possible.*

Solution: Here the given matrix A is a real symmetric matrix. The characteristic equation of the given matrix A is

$$|A - \lambda I| = \begin{vmatrix} 6-\lambda & 4 & -2 \\ 4 & 12-\lambda & -4 \\ -2 & -4 & 13-\lambda \end{vmatrix} = 0$$

$$\text{or, } (4-\lambda)(\lambda^2 - 27\lambda + 162) = 0$$

$$\text{or, } (\lambda - 9)(\lambda - 18)(4 - \lambda) = 0 \Rightarrow \lambda = 4, 9, 18.$$

Thus the eigen values of the given matrix are $4, 9, 18$. Corresponding to $\lambda = 4$, consider the equation $(A - 4I)X = \underline{0}$, where and $X = [x_1, x_2, x_3]^T$. The coefficient matrix is given by

$$A - 4I = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 8 & -4 \\ -2 & -4 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

$$\Rightarrow x_1 + 2x_2 - x_3 = 0, 7x_3 = 0.$$

We see that, $[-2, 1, 0]^T$ generates the eigen space of the eigen value 4 and forms a basis of the eigen space E_4 of 4. For $\lambda = 9$, the coefficient matrix is given by

$$A - 9I = \begin{pmatrix} -3 & 4 & -2 \\ 4 & 3 & -4 \\ -2 & -4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 7 & -6 \\ 0 & 5 & -4 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\Rightarrow x_1 + 7x_2 - 6x_3 = 0, 5x_2 - 4x_3 = 0$$

so that $x_3 = 5$ gives $x_1 = 2, x_2 = 4$. Hence $[2, 4, 5]^T$ is a eigen vector corresponding to the eigen value $\lambda = 9$. Thus $[2, 4, 5]^T$ generates the eigen space of the eigen value 9 and they form a basis of the eigen space E_9 of 9. For $\lambda = 18$, the coefficient matrix is given by

$$A - 18I = \begin{pmatrix} -12 & 4 & -2 \\ 4 & -6 & -4 \\ -2 & -4 & -5 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\Rightarrow 2x_1 + x_3 = 0, x_2 + x_3 = 0$$

so that $x_3 = -2$ gives $x_1 = 1, x_2 = 2$. Hence $[1, 2, -2]^T$ is a eigen vector corresponding to the eigen value $\lambda = 18$. Thus $[1, 2, -2]^T$ generates the eigen space of the eigen value 18 and they form a basis of the eigen space E_{18} of 18. These three vectors $[1, 1, 0]^T, [1, 0, -1]^T$ and $[1, 1, 2]^T$ are linearly independent and orthogonal, so the given matrix A is diagonalisable and the diagonalising orthogonal matrix is

$$P = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & \frac{2}{3} \\ 0 & \frac{5}{3\sqrt{5}} & -\frac{2}{3} \end{pmatrix} \text{ so that } P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{pmatrix},$$

where the diagonal elements are eigen values of A.

Diagonalization of linear operator

A linear operator $T : V \to V$ is said to be diagonalisable, if it can be represented by a diagonal matrix D. According to the definition, the linear operator $T : V \to V$ is diagonalisable if and only if there exists a basis $S = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ of V for which,

$$T(\alpha_1) = \lambda_1\alpha_1, T(\alpha_2) = \lambda_2\alpha_2, \cdots, T(\alpha_n) = \lambda_n\alpha_n. \tag{7.13}$$

In such a case, T is represented by the diagonal matrix $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ relative to the basis $S = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$.

Ex 7.3.8 *Each of the following real matrices defines a linear transformation on* $\Re^2$:

$$(a)\ A = \begin{pmatrix} 5 & 6 \\ 3 & -2 \end{pmatrix}; \quad (b)\ B = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}; \quad (c)\ C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}.$$

Find for each matrix, all eigen values and maximum set S of linearly indeperndent eigen vectors. Which of these linear operators are diagonalisable?

Solution: (a) The characteristic polynomial of A is

$$\begin{vmatrix} 5-\lambda & 6 \\ 3 & -2-\lambda \end{vmatrix} = \lambda^2 - 3\lambda - 28 = (\lambda + 4)(\lambda - 7).$$

Therefore, the eigen values of A are $-4, 7$. For $\lambda_1 = -4$, if $X_1 = (x_1, x_2)^T$ be the non null eigen vector, then

$$AX_1 = -4X_1 \Rightarrow \left.\begin{matrix} 9x_1 + 6x_2 = 0 \\ 3x_1 + 2x_2 = 0 \end{matrix}\right\} \Rightarrow 3x_1 + 2x_2 = 0.$$

Thus, $X_1 = (2, -3)^T$ is a eigen vector corresponding to $\lambda_1 = -4$. Similarly, $X_2 = (3, 1)^T$ is a eigen vector corresponding to $\lambda_2 = 7$.

So, $S = \{(2, -3), (3, 1)\}$ is a maximal set of linearly independent eigen vectors. Since S is a basis of $\Re^2$, A is diagonalisable. Using the basis S, A can be represented by the diagonal matrix

$$D = \begin{pmatrix} 2 & 3 \\ -3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 5 & 6 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 0 \\ 0 & 7 \end{pmatrix}.$$

(b) The characteristic polynomial of B is

$$\begin{vmatrix} 1-\lambda & -1 \\ 2 & -1-\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda + i)(\lambda - i).$$

There is no real characteristic root of B. Thus, B, a real matrix representing a linear transformation on $\Re^2$, has no eigen values and no eigen vectors. Hence in particular, B is not diagonalisable in $\Re^2$.

As a polynomial over C, the eigen values of B are $-i, i$. Therefore, $X_1 = (1, 1+i)^T$ and $X_2 = (1, 1-i)^T$ are the linearly independent eigen vectors of A corresponding to the eigen values i and $-i$ respectively. Now $S = \{(1, 1+i), (1, 1-i)\}$ is a basis of C^2 consisting of eigen vectors of B. Using this basis B, B can be represented by the diagonal matrix

$$D = \begin{pmatrix} 1 & 1 \\ 1+i & 1-i \end{pmatrix}^T \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1+i & 1-i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

As expected, the diagonal entries in D are the eigen values of A. Therefore, the matrix A is diagonalizable over the complex field C.

(c) The characteristic polynomial of C is

$$\begin{vmatrix} 5-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 8\lambda + 16 = (\lambda - 4)^2.$$

Therefore, the eigen values of C are 4, 4. For $\lambda_1 = 4$, if $X_1 = (x_1, x_2)^T$ be the non null eigen vector, then

$$CX_1 = 4X_1 \Rightarrow x_1 - x_2 = 0.$$

The homogeneous system has only one independent solution, say $(1,1)^T$, so, $(1,1)^T$ is an eigen vector of C. Furthermore, since there are no other eigen values, the solution set $S = \{(1,1)\}$ is a maximal set of linearly independent eigen vectors of C. Since S is not a basis of $\Re^2$, C is not diagonalisable.

* 7.4 Minimal Polynomial

It turns out that in the case of some matrices having eigen values with multiplicity greater than unity, there may exist polynomials of degree less than n, which equal to the zero matrix.

Minimal polynomial of a matrix

Let the characteristic polynomial of a matrix A be

$$\chi(\lambda) = (\lambda_1 - \lambda)^{d_1} (\lambda_2 - \lambda)^{d_2} \cdots (\lambda_l - \lambda)^{d_l}; \quad \sum_{i=1}^{l} d_i = n.$$

The Cayley-Hamilton theorem states that

$$\chi(A) = (\lambda_1 I - A)^{d_1} (\lambda_2 I - A)^{d_2} \cdots (\lambda_l I - A)^{d_l} = \underline{0}.$$

If $r_1, r_2, \cdots, r_l$ be the smallest positive integers for which,

$$J(A) \equiv (\lambda_1 I - A)^{r_1} (\lambda_2 I - A)^{r_2} \cdots (\lambda_l I - A)^{r_l} = \underline{0}$$

where, $r_i \le d_i$; $(1 \le i \le l)$ then,

$$J(\lambda) = (\lambda_1 - \lambda)^{r_1} (\lambda_2 - \lambda)^{r_2} \cdots (\lambda_l - \lambda)^{r_l} \tag{7.14}$$

is called the minimal polynomial of the matrix A. The degree of the minimal polynomial of an $n \times n$ matrix A is atmost n. It follows at once that if all the eigen values of a matrix are distinct, its minimal polynomial equal to characteristic polynomial.

Theorem 7.4.1 *The minimal polynomial $m(t)$ of a matrix A divides every polynomial which has A as zero. In particular, $m(t)$ divides the characteristic polynomial of A.*

Proof: Suppose $f(t)$ is a polynomial for which $f(A) = \underline{0}$. By the division algorithm, $\exists$ polynomials $q(t)$ and $s(t)$ for which

$$f(t) = m(t)q(t) + r(t) \tag{7.15}$$

where either $r(t) = 0$ or $degr(t) < degm(t)$. Substituting $t = A$ in (7.15) and using the fact that $f(A) = \underline{0}$ and $m(A) = \underline{0}$, we get $r(A) = \underline{0}$. If $r(t) \neq 0$, then by the division algorithm $degr(t) < degm(t)$. This means that there is a polynomial $r(t)$ of degree less than that of $m(t)$ such that $r(A) = \underline{0}$, which is a contradiction, as by definition of minimal polynomial, $m(t)$ is a polynomial of least degree such that $m(A) = \underline{0}$. Hence $r(t) = 0$ and so

$$f(t) = m(t)q(t), \quad i.e., m(t) \text{ divides } f(t).$$

As a particular case, since A satisfies its own characteristic equation, by Cayley-Hailton's theorem $m(t)$ divides the characteristic polynomial.

Theorem 7.4.2 *Let $m(t)$ be the minimal polynomial of an n square matrix A. Then the characteristic polynomial of A divides $(m(t))^n$.*

Proof: Let the minimal polynomial of an n square matrix A be

$$m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1} t + c_r.$$

Define, the matrices B_j as follows

$$\begin{aligned} B_0 &= I & &\text{so } I = B_0 \\ B_1 &= A + c_1 I & &\text{so } c_1 I = B_1 - AB_0 \\ B_2 &= A^2 + c_1 A + c_2 I & &\text{so } c_2 I = B_2 - AB_1 \\ &\vdots & &\vdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I & &\text{so } c_{r-1} I = B_{r-1} - AB_{r-2} \end{aligned}$$

Also, we have,

$$\begin{aligned} -AB_{r-1} &= c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1} A + c_r I) \\ &= c_r I - m(A) = c_r I. \end{aligned}$$

Set, $B(t) = t^{r-1} B_0 + t^{r-2} B_1 + \cdots + t B_{r-2} + B_{r-1}$, we get,

$$\begin{aligned} (tI - A)B(t) &= (t^r B_0 + t^{r-1} B_1 + \cdots + t B_{r-1}) \\ &\quad - (t^{r-1} AB_0 + t^{r-2} AB_1 + \cdots + AB_{r-1}) \\ &= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} t I + c_r I = m(t) I \end{aligned}$$

$|tI - A||B(t)| = |m(t)I| = (m(t))^n$; taking determinant.

Since $|B(t)|$ is a polynomial, $|tI - A|$ divides $(m(t))^n$. Hence the characteristic polynomial divides $(m(t))^n$.

Theorem 7.4.3 *The characteristic polynomial and the minimal polynomial of a matrix A have the same irreducible factors.*

Proof: Let $f(t)$ is an irreducible polynomial. If $f(t)$ divides the minimal polynomial $m(t)$, then as $m(t)$ divides the characteristic polynomial, $f(t)$ must divide the characteristic polynomial. On the other hand, if $f(t)$ divides the characteristic polynomial of A, by the above theorem, $f(t)$ must divide $(m(t))^n$. But $f(t)$ is irreducible, hence $f(t)$ also divides $m(t)$. Thus $m(t)$ and the characteristic polynomial have the same irreducible factors.

Result 7.4.1 This theorem does not say that $m(t)$ =characteristic polynomial, only that any irreducible factor of one must divide the other. In particular, since a linear factor is irreducible, $m(t)$ and characteristic polynomial have the same linear factors, so that they have the same roots. Thus we conclude, a scalar λ is an eigen value of the matrix A if and only if λ is a root of the minimal polynomial of A.

Ex 7.4.1 *Find the characteristic and minimal polynomials of each of the following matrices.*

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 4 & -2 \\ -1 & -1 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 2 & -1 \\ 3 & 8 & -3 \\ 3 & 6 & -1 \end{pmatrix}.$$

Solution: The characteristic polynomial $\chi_A(\lambda)$ of A is

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ 2 & 4-\lambda & -2 \\ -1 & -1 & 3-\lambda \end{vmatrix}$$

$$= 24 - 28\lambda + 10\lambda^2 - \lambda^3 = (\lambda - 2)^2(6 - \lambda).$$

The characteristic polynomial $\chi_B(\lambda)$ of B is

$$|B - \lambda I| = \begin{vmatrix} 3-\lambda & 2 & -1 \\ 3 & 8-\lambda & -3 \\ 3 & 6 & -1-\lambda \end{vmatrix}$$

$$= 24 - 28\lambda + 10\lambda^2 - \lambda^3 = (\lambda - 2)^2(6 - \lambda).$$

Thus the characteristic polynomial of both matrices is same. Since the characteristic polynomial and the minimal polynomial have the same irreducible factors, it follows that both $(t-2)$ and $(6-t)$ must be factors of $m(t)$. Also, $m(t)$ must divide the characteristic polynomial. Hence it follows that $m(t)$ must be one of the following:

$$f(t) = (t-2)(6-t) \text{ or } g(t) = (t-2)^2(6-t).$$

(i) By Cayley-Hailton's theorem, $g(A) = \Delta(A) = \underline{0}$, so we need only test $f(t)$. Now,

$$(A - 2I)(6I - A) = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 3 & -1 & 1 \\ -2 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix} = \underline{0}.$$

Therefore, $m(t) = (t-2)(6-t)$ is the minimal polynomial of A.

(ii) By Cayley-Hailton's theorem, $g(B) = \Delta(B) = \underline{0}$, so we need only test $f(t)$. Now,

$$(B - 2I)(6I - B) = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \\ 3 & 6 & -3 \end{pmatrix}\begin{pmatrix} 3 & -2 & 1 \\ -3 & -2 & 3 \\ -3 & -6 & 7 \end{pmatrix} = \underline{0}.$$

Therefore, $m(t) = (t-2)(6-t)$ is the minimal polynomial of B.

Deduction 7.4.1 Consider the following two n square matrices as

$$J(\lambda, n) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \text{ and } A = \begin{pmatrix} \lambda & a & 0 & \cdots & 0 & 0 \\ 0 & \lambda & a & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda & a \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

where $a \neq 0$. The matrix $J(\lambda, n)$, called *Jordan Block* has λ's on the diagonal, 1's on the superdiagonal and 0's elsewhere. The matrix A, which is the generalization of $J(\lambda, n)$, has λ's on the diagonal, a's on the super diagonal and 0's elsewhere. Now we see that $f(t) = (t - \lambda)^n$ is the characteristic and minimal polynomial of both $J(\lambda, n)$ and A.

Ex 7.4.2 *Find the minimal polynomial of the matrix* $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$.

Solution: The characteristic polynomial $\chi(t)$ of A is given by,

$$|A - tI| = \begin{vmatrix} 2-t & 1 & 0 & 0 \\ 0 & 2-t & 0 & 0 \\ 0 & 0 & 2-t & 0 \\ 0 & 0 & 0 & 5-t \end{vmatrix} = (t-2)^3(t-5).$$

Since the characteristic polynomial and the minimal polynomial have the same irreducible factors, it follows that both $t-2$ and $t-5$ must have factors of $m(t)$. Also, $m(t)$ must divide the characteristic polynomial. Hence it follows that $m(t)$ must be one of the following three polynomials:

$$(i)\ m(t) = (t-2)(t-5),\ (ii)\ m(t) = (t-2)^2(t-5),\ (iii)\ m(t) = (t-2)^3(t-5).$$

For the type (i), we have,

$$(A - 2I)(A - 5I) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} -3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq \underline{0}.$$

For the type (ii), we have,

$$(A - 2I)(A - 2I)(A - 5I) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}\begin{pmatrix} -3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \underline{0}.$$

For the type (iii), we have obviously, $(A - 2I)^3(A - 5I) = \underline{0}$ follows from Cayley-Hailton's theorem. Since $m(t)$ is minimal polynomial, we have $m(t) = (t-2)^2(t-5)$.

Deduction 7.4.2 Let us consider an arbitrary monic polynomial

$$f(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0.$$

Let us consider an n^{th} order square matrix A with 1's on the subdiagonal, last column $[-c_0, -c_1, \cdots, -c_{n-1}]^T$ and 0's elsewhere as follows

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix},$$

then A is called the *companion matrix* of the polynomial $f(t)$. Moreover, the characteristic and minimal polynomial of the comparison matrix A are both equal to the original polynomial $f(t)$.

Ex 7.4.3 *Find a matrix whose minimal polynomial is* $t^3 - 5t^2 + 6t + 8$.

Solution: Here the given monic polynomial is $f(t) = t^3 - 5t^2 + 6t + 8$. Let A be the comparison matrix of the polynomial $f(t)$, then by definition

$$A = \begin{pmatrix} 0 & 0 & -8 \\ 1 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}.$$

Also the characteristic and minimal polynomial of the comparison matrix A are both equal to the original given polynomial $f(t)$.

Minimal polynomial of linear operator

The minimal polynomial of the operator T is defined independently of the theory of matrices, as the monic polynomial of the lowest degree with leading coefficient 1, which has T as zero. However, for any polynomial $f(t)$,

$$f(T) = \underline{0} \text{ if and only if } f(A) = \underline{0},$$

where A is any matrix representation of T. Accordingly, T and A have the same minimal polynomials.

(i) The minimal polynomial $m(t)$ of a linear operator T divides every polynomial that has T as a zero. In particular, the minimal polynomial $m(t)$ divides the characteristic polynomial of T.

(ii) The characteristic and minimal polynomials of a linear operator T have the same irreducible factors.

(iii) A scalar λ is an eigen value of a linear operator if and only if λ is a root of the minimal polynomial $m(T)$ of T.

Minimal polynomial of block diagonal matrices

Let A be a block diagonal matrix with diagonal blocks $A_1, A_2, \cdots, A_r$. Then the minimal polynomial of A is equal to the least common multiple of the minimal polynomials of the diagonal blocks A_i.

Ex 7.4.4 *Find the characteristic and minimal polynomial of the block diagonal matrix*

$$A = \left[\begin{array}{cc:cc:c} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 7 \end{array}\right]$$

Solution: The given block diagonal matrix can be written in the form

$$A = diag(A_1, A_2, A_3); \quad \text{where, } A_1 = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} 4 & 2 \\ 3 & 5 \end{bmatrix}, A_3 = [7].$$

The characteristic polynomials of A_1, A_2, A_3 are

$$|A_1 - \lambda I| = (\lambda - 2)^2; \quad |A_2 - \lambda I| = (\lambda - 2)(\lambda - 7); \quad |A_3 - \lambda I| = \lambda - 7.$$

Thus the characteristic polynomial of A is

$$|A - \lambda I| = |A_1 - \lambda I| \; |A_2 - \lambda I| \; |A_3 - \lambda I|$$
$$= (\lambda - 2)^2(\lambda - 2)(\lambda - 7)(\lambda - 7) = (\lambda - 2)^3(\lambda - 7)^2.$$

The minimal polynomials $m_1(t), m_2(t), m_3(t)$ of the diagonal blocks A_1, A_2, A_3 respectively, are equal to the characteristic polynomials, i.e.,

$$m_1(t) = (t-2)^2; \quad m_2(t) = (t-2)(t-7); \quad m_3(t) = t - 7.$$

But $m(t)$, the minimal polynomial of A is equal to the least common multiple of $m_1(t), m_2(t)$ and $m_3(t)$, i.e., $m(t) = (t-2)^2(t-7)$.

7.5 Bilinear Forms

Let V be an n dimensional Euclidean space, then $B(\alpha, \beta) : V \times V \to \Re$ is said to be a bilinear form if it is linear, homogeneous, with respect to both the arguments α, β. If B is a symmetric Bilinear form then we can write

$$B(\alpha, \alpha) = (B\alpha, \alpha) = \alpha^T B\alpha \tag{7.16}$$

where $B = [b_{ij}]$ real symmetric $n \times n$ matrix, known as the matrix of the quadratic form, $B(\alpha, \alpha)$ is known as a *quadratic form.* For example, the expression

$$5x_1y_1 + 2x_1y_2 - 3x_1y_3 + 7x_2y_1 - 5x_2y_2 + 3x_3y_3$$

is a bilinear form in the variables x_1, x_2 and y_1, y_2, y_3. If we change the base vector such that $\alpha = P\alpha'$ then,

$$\alpha'^T P^T B P\alpha' = B(P\alpha', P\alpha'). \tag{7.17}$$

The matrices of the two quadratic forms (7.16) and (7.17) connected by the transformation, i.e., if we change co-ordinates in a quadratic form, its matrix is change to a matrix which is congruent to the matrix of the original quadratic form. In other words if we have two quadratic form whose matrices are congruent to each other then they represent the same quadratic form only with respect to two coordinate system connected by a non singular transformation.

7.5.1 Real Quadratic Forms

A homogeneous expression of the second degree in any number of variables of the form

$$Q(\alpha, \alpha) = \alpha^T B\alpha = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j; \quad a_{ij} = a_{ji}$$
$$= (x_1, x_2, \cdots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \tag{7.18}$$

where a_{ij} are constants belonging to a field of numbers and $x_1, x_2, \cdots, x_n$ are variables, belonging to a field of numbers (not necessarily same) is defined as a quadratic form. If the variables assumes real variables only, the form is said to be quadratic form in real variables. When the constants a_{ij} and the variables x_i's are all real, then the expression is said to be a real quadratic form with B as associated matrix. For example,

(i) $x_1^2+2x_1x_2+3x_2^2$ is real quadratic forms in 2 variables with associated matrix

$$B_1=\begin{pmatrix}1&1\\1&3\end{pmatrix}.$$

(ii) $x_1^2+3x_2^2+3x_3^2-4x_2x_3+4x_3x_1-2x_1x_2$ is real quadratic forms in 3 variables with associated matrix

$$B_2=\begin{pmatrix}1&-1&2\\-1&-3&-2\\2&-2&3\end{pmatrix}..$$

Now, A real quadratic form $Q(\alpha,\alpha)$ is said to be

(i) positive definite, if $Q>0$, for all $\alpha\neq\theta$.

(ii) positive semi-definite, if $Q\geq 0$, for all α and $Q=0$ for some $\alpha\neq\theta$.

(iii) negative definite, if $Q<0$, for all $\alpha\neq\theta$.

(iv) negative semi-definite, if $Q\leq 0$, for all α and $Q=0$ for some $\alpha\neq\theta$.

(v) indefinite, if $Q\geq 0$, for some $\alpha\neq\theta$ and $Q\leq 0$ for some $\alpha\neq\theta$.

These five classes of quadratic forms are called *value classes*.

Ex 7.5.1 *Find the quadratic form that corresponds to a symmetric matrix* $A=\begin{pmatrix}5&-3\\-3&8\end{pmatrix}$.

Solution: The quadratic form $Q(\alpha)$ that corresponds to a symmetric matrix A is $Q(\alpha)=\alpha^TA\alpha$, where $\alpha=(x_1,x_2)^T$ is the column vectors of unknowns. Thus,

$$Q(\alpha)=(x_1,x_2)\begin{pmatrix}5&-3\\-3&8\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=5x_1^2-6x_1x_2+8x_2^2.$$

Ex 7.5.2 *Examine whether the quadratic form* $5x^2+y^2+5z^2+4xy-8xz-4yz$ *is positive definite or not.*

Solution: The given quadratic form can be written as

$$Q(x,y,z)=5x^2+y^2+5z^2+4xy-8xz-4yz$$
$$=(2x+y-2z)^2+x^2+z^2.$$

Since $Q>0$, for all (x,y,z) and $Q=0$, only when $x=y=z=0$, i.e., $\alpha=\theta$. Hence Q is positive definite. Alternatively, if $Q(\alpha,\alpha)=\alpha^TB\alpha$, then the associated matrix B is given by, $B=\begin{pmatrix}5&2&-4\\2&1&-2\\4&-2&5\end{pmatrix}$. The principal minors of B are

$$5,\quad \begin{vmatrix}1&-2\\-2&5\end{vmatrix}=9,\quad |B|=17$$

are all positive, hence the given quadratic form Q is positive definite.

Ex 7.5.3 *Prove that the real quadratic* $Q=ax^2+bxy+cy^2$ *is positive definite, if* $a>0$ *and* $b^2<4ac(a,b,c\neq 0)$.

Solution: The given quadratic form can be written as

$$Q(x,y,z)=ax^2+bxy+cy^2$$
$$=a\left[(x+\frac{b}{2a}y)^2+\frac{4ca-b^2}{4a^2}y^2\right].$$

Since the expression Q is positive definite, so, $Q\geq 0$, i.e., if $a>0$ and $b^2<4ac(a,b,c\neq 0)$. Also, if $Q(\alpha,\alpha)=\alpha^TB\alpha$, then the associated matrix B is given by, $B=\begin{pmatrix}a&\frac{b}{2}\\\frac{b}{2}&c\end{pmatrix}$. The principal minors of B are a, $\begin{vmatrix}a&\frac{b}{2}\\\frac{b}{2}&c\end{vmatrix}=a^2-\frac{b^2}{4}$. For $a,b,c\neq 0$, and if $a>0$ and $b^2<4ac$, then, all principal minors are positive, hence the given quadratic form Q is positive definite.

Ex 7.5.4 *Examine whether the quadratic form* $x^2+2y^2+2z^2-2xy+2xz-4yz$ *is positive definite or not.*

Solution: The given quadratic form can be written as

$$Q(x,y,z)=x^2+2y^2+2z^2-2xy+2xz-4yz$$
$$=(y-z-x)^2+(y-z)^2.$$

Since $Q\geq 0$, for all (x,y,z) and if we take $x=0,y=z=1$, then $Q=0$. Hence $Q\geq 0$, for all $\alpha\neq\theta$ and so Q is positive semi-definite.

Ex 7.5.5 *Examine whether the quadratic form* $x^2+y^2-2z^2+2xy-2yz-2xz$ *is positive definite or not.*

Solution: The given quadratic form can be written as

$$Q(x,y,z)=x^2+y^2-2z^2+2xy-2yz-2xz$$
$$=(x+y-z)^2-3z^2.$$

We see that, $Q\geq 0$, for some $\alpha\neq\theta$ and $Q\leq 0$ for some $\alpha\neq\theta$. For example, if $(x,y,z)=(1,0,0)$, then $Q>0$, $(x,y,z)=(0,0,1)$, then $Q<0$, $(x,y,z)=(1,-1,0)$, then $Q=0$, so, the given expression Q is indefinite.

7.6 Canonical Form

Let us consider the real quadratic form $Q(\alpha,\alpha)=\alpha^TA\alpha$, where A is a real symmetric matrix of order n. From spectral theorem, we know that eigen vector of A forms an orthonormal basis of V. Let P be the n square matrix whose columns are orthogonal eigen vector of A then, $|P|\neq 0$ and $P^T=P^{-1}$ and hence the non-singular linear transformation $\alpha'=P\alpha$ will transform $\alpha^TA\alpha$ to

$$\alpha^TP^TAP\alpha=\alpha'^TP^{-1}AP\alpha'=\alpha'^TD\alpha'=Q'(\alpha',\alpha'), \tag{7.19}$$

where $D=P^{-1}AP=P^TAP$ is the symmetric diagonal matrix whose element in the diagonal are the eigen values of the matrix A, i.e., $D=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)=P^{-1}AP$. Now, $Q'(\alpha',\alpha')$ is real quadratic form and it is called a linear transformation of Q. The matrix D of Q' is congruent to A and Q' is said to be congruent to Q. When expressed in terms of coordinates, the equation (7.19) has the form

$$Q'(\alpha',\alpha')=\alpha'^TD\alpha'=\lambda_1x_1'^2+\lambda_2x_2'^2+\cdots+\lambda_nx_n'^2. \tag{7.20}$$

Let $\lambda_1, \lambda_2, \cdots, \lambda_p$ be the positive eigen values of A, $\lambda_{p+1}, \lambda_{p+2}, \cdots, \lambda_r$ the negative eigen values of A and $\lambda_r, \lambda_{r+1}, \cdots, \lambda_n$ be the zero eigen values of A, where r is the rank of A. If A be an $n \times n$ real symmetric matrix of rank $r (\leq n)$, then $\exists$ a non singular matrix P such that $P^T AP$, i.e., D becomes diagonal with the form

$$\begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & 0 \end{bmatrix}; \quad 0 \leq p \leq r.$$

Thus, if $p, r-p, n-r$ are defined to be *the positive, the negative and the zero indices of inertia*, and it is expressed by writing

$$I_n(A) = (p, r-p, n-r). \tag{7.21}$$

The quantity, $p-(r-p) = s$ is defined as the *signature*. We can reduce the equation (7.20) to the further signature form applying the following transformation

$$\left.\begin{array}{ll} x_i' = \frac{1}{\sqrt{|\alpha_i|}} x_i''; & i = 1, 2, \cdots, r \\ x_i' = x_i''; & i = r+1, \cdots, n \end{array}\right\} \tag{7.22}$$

The equation (7.22) transforms to (7.20) into the quadratic form

$$x_1''^2 + \cdots + x_p''^2 - x_{p+1}''^2 - \cdots - x_r''^2. \tag{7.23}$$

We have reduce the quadratic form (7.20) to the quadratic form (7.23) which is the sum of the square terms with coefficients as $+1$ and -1 respectively. The quadratic form (7.20) is called the canonical or normal form of Q. The number of positive terms in the normal form is called index.

Deduction 7.6.1 Sylvesters law of inertia: Sylvestes law of inertia states that when a quadratic form is reduced to a normal form similar to (7.23), the rank and signature of the form remains invariant, i.e., $I_n(A)$ is independent of the method of reducing (7.20) to the canonical form (7.23).

Deduction 7.6.2 Classification of quadratic forms : A quadratic form $Q(\alpha, \alpha) = \alpha^T A\alpha$ is said to be a *positive definite* if $Q(\alpha, \alpha) > 0; \forall \alpha \neq \theta$, negative definite if $Q(\alpha, \alpha) < 0; \forall \alpha \neq \theta$, positive semi definite if $Q(\alpha, \alpha) \geq 0; \forall \alpha$, negative semi definite if $Q(\alpha, \alpha) \leq 0; \forall \alpha$ and is said to be indefinite if $Q(\alpha, \alpha)$ can take positive value for some $\alpha \neq \theta$ as well as negative value for some other $\alpha \neq \theta$. Thus

(i) the quadratic form $Q(\alpha, \alpha) = \alpha^T A\alpha$ is positive definite if $(r = n)$, all the eigen values are positive, i.e., $I_n(A) = (n, 0, 0)$. In this case, the canonical form becomes $x_1''^2 + x_2''^2 + \cdots + x_n''^2$.

(ii) the quadratic form $Q(\alpha, \alpha) = \alpha^T A\alpha$ is negative definite if all the eigen values are negative, i.e., $I_n(A) = (0, n, 0)$. In this case, the canonical form becomes $-x_1''^2 - x_2''^2 - \cdots - x_n''^2$.

(iii) the quadratic form $Q(\alpha, \alpha) = \alpha^T A\alpha$ is positive semi definite if $r < n, r-p = 0$, i.e., $I_n(A) = (p, 0, n-p)$ and positive semi definite if $I_n(A) = (0, r, n-r)$.

(iv) the quadratic form $Q(\alpha, \alpha) = \alpha^T A\alpha$ is indefinite if $I_n(A) = (p, r-p, n-r)$, where $p > 0$ and $r-p > 0$.

Each quadratic form must be of one of these five types. Sylvester's criterion states that the real symmetric matrix A is positive definite if and only if all its principal minors of A are positive. This remains valid if we replace the word 'positive' everywhere by the word 'non-negative'.

Ex 7.6.1 *Reduce the quadratic form $5x_1^2 + x_2^2 + 10x_3^2 - 4x_2x_3 - 10x_3x_1$ to the normal form.*

Solution: The given quadratic form can be written as

$$Q(\alpha, \alpha') = (x_1 \ x_2 \ x_3) \begin{pmatrix} 5 & 0 & -5 \\ 0 & 1 & -2 \\ -5 & -2 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

where, the associated symmetric matrix is given by $A = \begin{pmatrix} 5 & 0 & -5 \\ 0 & 1 & -2 \\ -5 & -2 & 10 \end{pmatrix}$. Let us apply congruence operations on A to reduce it to the normal form

$$A \xrightarrow{R_3 + R_2} \begin{pmatrix} 5 & 0 & -5 \\ 0 & 1 & -2 \\ 0 & -2 & 5 \end{pmatrix} \xrightarrow{C_3 + C_2} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 5 \end{pmatrix}$$

$$\xrightarrow{R_3 + 2R_2} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_3 + 2C_2} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{\sqrt{5}} R_1} \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\frac{1}{\sqrt{5}} C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The rank of the quadratic form is $r = 3$ and the number of positive indices of inertia is $p = 3$, which is the index. Therefore, the signature of the quadratic form is $2p - r = 3$. Here, $n = r = p = 3$, so, the quadratic form is positive definite.

Ex 7.6.2 *Let $Q = x^2 + 6xy - 7y^2$. Find the orthogonal substitution that diagonalizes Q.*

Solution: The symmetric matrix A that represents Q is $A = \begin{pmatrix} 1 & 3 \\ 3 & -7 \end{pmatrix}$. The characteristic polynomial of A is

$$|A - \lambda I| = \lambda^2 - (1-7)\lambda + (7-6) = (\lambda + 8)(\lambda - 2).$$

The eigenvalues of A are $-8, 2$. Thus using x_1 and x_2 as new variables, a diagonal form of Q is

$$Q(x_1, x_2) = 2x_1^2 - 8x_2^2.$$

The corresponding orthogonal substitution is obtained by finding an orthogonal set of eigen vectors of A. The eigen vector corresponding to $\lambda_1 = -8$ and $\lambda_2 = 2$ are $X_1 = (-1, 3)^T$ and $X_2 = (3, 1)^T$ respectively. Since A is symmetric, the eigen vectors X_1 and X_2 are orthogonal. Now, we normalize X_1 and X_2 to obtain respectively the unit vectors

$$\hat{X}_1 = \left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)^T \text{ and } \hat{X}_2 = \left(-\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right)^T.$$

Finally, let P be the matrix whose columns are the unit vectors X_1 and X_2 respectively, and then $(x, y)^T = P(x_1, x_2)^T$ is the required orthogonal change of coordinates, i.e.,

$$P = \begin{pmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \text{ and } x = \frac{1}{\sqrt{10}}(3x_1 - x_2), \ y = \frac{1}{\sqrt{10}}(x_1 + 3x_2).$$

One can also express x_1 and x_2 interms of x and y by using $P^{-1} = P^T$ as

$$x_1 = \frac{1}{\sqrt{10}}(3x+y); \quad x_2 = \frac{1}{\sqrt{10}}(-x+3y).$$

Classification of conics

(*i*) The general equation of a quadratic conic in two variables x and y can be written in the form

$$ax^2 + 2hxy + by^2 + gx + fy + c = 0$$

$$(x \;\; y)\begin{pmatrix} a & h \\ h & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (g \;\; f)\begin{pmatrix} x \\ y \end{pmatrix} + c = 0$$

$$\text{or,} \quad \alpha^T A\alpha + K^T\alpha + c = 0, \tag{7.24}$$

where $A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$ is a real symmetric matrix and hence it is orthogonally diagonalizable, $K^T = (g \;\; f)$ and $\alpha^T = (x, y)$. Let λ_1, λ_2 be the eigen values of the real symmetric matrix A, the corresponding eigen vectors be α_1, α_2 respectively, so that for $P = [\alpha_1, \alpha_2]$, $P^{-1} = P^T$ i.e., P is a orthogonal matrix and

$$AP = PD, \text{ where, } D = diag(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

If we apply the rotation $\alpha = P\alpha'$, then equation (7.24) reduces to,

$$\alpha'^T P^T AP\alpha' + K^T P\alpha' + c = 0$$

$$\text{or,} \quad \lambda_1 x'^2 + \lambda_2 y'^2 + g'x' + f'y' + c = 0 \tag{7.25}$$

where the rotation $\alpha = P\alpha'$ transform the principal axes into coordinate axes. Let us now apply the translation $x' = x'' + \delta, y' = y'' + \mu$.

If $\lambda_1 \neq 0$, coefficient of x'' may be made to be zero for a suitable choice of δ and if $\lambda_2 \neq 0$, coefficient of y'' may be made to be zero for a suitable choice of μ. Therefore, the conic (7.25) can be transformed to one of the three general forms

1. Let $I_n(A) = (2, 0, 0)$, then the standard form becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \quad \text{ellipse}$$
$$= 0; \quad \text{a single point.}$$

2. Let $I_n(A) = (1, 1, 0)$, i.e., rank of $A = 1$. In this case, one of λ_1 and λ_2 is zero, so, the standard form becomes

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1; \quad \text{hyperbola}$$
$$= 0; \quad \text{pair of intersecting straight lines}$$

3. Let $I_n(A) = (1, 0, 1)$, then the standard form becomes

$$x^2 - 4y = 0; \quad \text{parabola}$$
$$= 1; \quad \text{pair of straight lines}$$
$$= 0; \quad \text{a single straight line}$$

(*ii*) The general equation of a quadratic conic in the variables x, y, z can be written in the form

$$ax^2 + by^2 + cz^2 + 2hxy + 2gyz + 2fzx + ux + vy + wz + d = 0$$

$$(x \;\; y \;\; z)\begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + (u \;\; v \;\; w)\begin{pmatrix} x \\ y \\ z \end{pmatrix} + d = 0$$

$$\text{or,} \quad \alpha^T A\alpha + K^T\alpha + d = 0, \tag{7.26}$$

where $A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$ is a real symmetric matrix and hence it is orthogonally diagonalizable, $K^T = (u \;\; v \;\; w)$ and $\alpha^T = (x, y, z)$. Rotating the coordinate axes to coincide with orthogonal eigen axes or principal axes and translating the origin suitably the quadratic form (7.26) can be reduced to one of the following six general forms, assuming that $\lambda_1 > 0$ and in the final expression constant on the right hand side, if any, is positive. Therefore, the conic (7.26) can be transformed to one of the six general forms

1. Let $I_n(A) = (3, 0, 0)$, i.e., rank of $A = 3$. In this case, none of $\lambda_1, \lambda_2, \lambda_3$ is zero. Then the standard form becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1; \quad \text{ellipsoid}$$
$$= 0; \quad \text{a single point.}$$

2. Let $I_n(A) = (2, 1, 0)$, then the standard form becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1; \quad \text{elliptic hyperboloid of one sheet}$$
$$= 0; \quad \text{elliptic cone}$$

3. Let $I_n(A) = (1, 2, 0)$, then the standard form becomes

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1; \quad \text{elliptic hyperboloid of two sheets}$$
$$= 0; \quad \text{elliptic cone}$$

4. Let $I_n(A) = (2, 0, 1)$, then the standard form becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z; \quad \text{elliptic paraboloid}$$
$$= 1; \quad \text{elliptic cylinder}$$
$$= 0; \quad \text{a single point}$$

5. Let $I_n(A) = (1, 1, 0)$, then the standard form becomes

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z; \quad \text{hyperbolic paraboloid}$$
$$= 1; \quad \text{hyperbolic cylinder}$$
$$= 0; \quad \text{pair of intersecting planes}$$

6. Let $I_n(A) = (1, 0, 2)$, then the standard form becomes

$$\frac{x^2}{a^2} = \delta y + \mu z; \quad \text{parabolic cylinder}$$
$$= 1; \quad \text{a pair of planes}$$
$$= 0; \quad \text{a simple plane}$$

Ex 7.6.3 *Reduce $2y^2 - 2xy - 2yz + 2zx - x - 2y + 3z - 2 = 0$ into canonical form.*

Solution: The quadratic equation $2y^2 - 2xy - 2yz + 2zx - x - 2y + 3z - 2 = 0$ in x, y, z, can be written as

$$(x\ y\ z)\begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + (-1\ \ -2\ \ 3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} - 2 = 0,$$
$$X^T AX + BX - 2 = 0.$$

The characteristic equation of A is

$$|A - \lambda I| = 0 \Rightarrow \begin{vmatrix} -\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & -\lambda \end{vmatrix} = 0$$
$$\lambda^3 - 2\lambda^2 - 3\lambda = 0 \Rightarrow \lambda = -1, 0, 3.$$

The eigen vectors corresponding to the eigen values $3, -1, 0$ are $k_1(1, -2, 1), k_2(1, 0, -1)$ and $k_3(1, 1, 1)$. The orthogonal eigen vectors are

$$(\tfrac{1}{\sqrt{6}}, -\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}), (\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}) \text{ and } (\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}})$$

respectively. Let $P = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & \sqrt{3} & \sqrt{2} \\ -2 & 0 & \sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{pmatrix}$, then, $P^TAP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and

$$BP = (\sqrt{2}\ \ -2\sqrt{2}\ \ 0).$$

By the orthogonal transformation, $X = PX'$, where $X'^T = (x'\ y'\ z')$, the equation reduces to

$$3x'^2 - y'^2 + \sqrt{6}x' - 2\sqrt{2}y' - 2 = 0$$
$$\text{or, } 3(x' + \frac{1}{\sqrt{6}})^2 - (y' + \sqrt{2})^2 = \frac{1}{2}.$$

Let us applying the transformation

$$x'' = x' + \tfrac{1}{\sqrt{6}}, y'' = y' + \sqrt{2}, z'' = z',$$

the equation finally reduces to $3x''^2 - y''^2 = \frac{1}{2}$, which is canonical form and it represent a hyperbolic cylinder.

Ex 7.6.4 *Reduce the equation $2x^2 + 5y^2 + 10z^2 + 4xy + 6xz + 12yz$ into canonical form.*

Solution: The given quadratic form in x, y, z, can be written in the form

$$Q(\alpha, \alpha') = (x\ y\ z)\begin{pmatrix} 2 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = X^TAX,$$

where, the associated symmetric matrix is A. Let us apply congruence operations on A to reduce it to the normal form

$$A \xrightarrow{R_2 - R_1, R_3 - \frac{3}{2}R_1} \begin{pmatrix} 2 & 2 & 3 \\ 0 & 3 & 3 \\ 0 & 3 & \frac{11}{2} \end{pmatrix} \xrightarrow{C_2 - C_1, C_3 - \frac{3}{2}C_1} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 3 & \frac{11}{2} \end{pmatrix}$$
$$\xrightarrow{R_3 - R_2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} \xrightarrow{C_3 - C_2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix}$$
$$\xrightarrow{\frac{1}{\sqrt{2}}R_1, \frac{1}{\sqrt{2}}R_2, \sqrt{\frac{2}{5}}R_3} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{\frac{5}{2}} \end{pmatrix} \xrightarrow{\frac{1}{\sqrt{2}}C_1, \frac{1}{\sqrt{2}}C_2, \sqrt{\frac{2}{5}}C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The rank of the quadratic form is $r = 3$ and the number of positive indices of inertia is $p = 3$, which is the index. Therefore, the signature of the quadratic form is $2p - r = 3$. Here, $n = r = p = 3$ so, the quadratic form is positive definite. The corresponding normal form is $x^2 + y^2 + z^2$.

Ex 7.6.5 *Obtain a non-singular transformation that will reduce the quadratic form $x^2 + 2y^2 + 3z^2 - 2xy + 4yz$ to the normal form.*

Solution: The given quadratic form in x, y, z, can be written as

$$Q(\alpha, \alpha') = (x\ y\ z)\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 2 \\ 0 & 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = X^TAX,$$

where, the associated symmetric matrix is A. Let us apply congruence operations on A to reduce it to the normal form

$$A \xrightarrow{R_2 + R_1} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix} \xrightarrow{C_2 + C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$
$$\xrightarrow{R_3 - 2R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow{C_3 - 2C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The rank of the quadratic form is $r = 3$ and the number of positive indices of inertia is $p = 2$, which is the index. Therefore, the signature of the quadratic form is $2p - r = 1$. Here, $n = r = 3$ and $p = 1 < r$, so, the quadratic form is indefinite. The corresponding normal form is $x^2 + y^2 - z^2$. Let $X = PX'$, where, $X'^T = (x'\ y'\ z')$ and P is non-singular, transforms the form into the normal form X'^TDX', then $D(= P^TAP)$ is a diagonal matrix. By the property of elementary matrices, we get,

$$E_{32}(-2)E_{21}(1)A\{E_{21}\}^T\{E_{32}(-2)\}^T = D$$
$$\Rightarrow P^T = E_{32}(-2)E_{21}(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & -2 & 1 \end{pmatrix}.$$

Thus the transformation $X = PX'$ becomes,

$$x = x' + y' - 2z', y = y' - 2z', z = z'.$$

Ex 7.6.6 *Show that the quadratic form $x_1x_2 + x_2x_3 + x_3x_1$ can be reduced to the canonical form $y_1^2 - y_2^2 - y_3^2$, by means of the transformation*

$$x_1 = y_1 - y_2 - y_3, x_2 = y_1 + y_2 - y_3, x_3 = y_3.$$

Solution: The quadratic form $x_1x_2 + x_2x_3 + x_3x_1$ in x_1, x_2, x_3, can be written as

$$Q(\alpha,\alpha') = (x_1 \ \ x_2 \ \ x_3)\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = X^TAX,$$

where, the associated symmetric matrix is A. Let us apply congruence operations on A to reduce it to the normal form

$$A \xrightarrow{R_1+R_2} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \xrightarrow{C_1+C_2} \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 \end{pmatrix}$$

$$\xrightarrow{R_2 - \frac{1}{2}R_1, R_3 - R_1} \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow{C_2 - \frac{1}{2}C_1, C_3 - C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$\xrightarrow{2R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow{2C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The rank of the quadratic form is $r = 3$ and the number of positive indices of inertia is $p = 1$, which is the index. The corresponding normal form is $y_1^2 - y_2^2 - y_3^2$. By the property of elementary matrices, we get,

$$E_2(2)E_{31}(-1)E_{12}(-\tfrac{1}{2})^TE_{21}(1)A[E_{21}(1)]^T[E_{12}(-\tfrac{1}{2})]^T[E_{31}(-1)]^T[E_2(2)]^T = D$$

$$\Rightarrow P^T = E_2(2)E_{31}(-1)E_{12}(-\tfrac{1}{2})^TE_{21}(1)$$

$$= \begin{pmatrix} 1&0&0 \\ 0&2&0 \\ 0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0 \\ 0&1&0 \\ -1&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0 \\ -\frac{1}{2}&1&0 \\ 0&0&1 \end{pmatrix}\begin{pmatrix} 1&1&0 \\ 0&2&0 \\ 0&0&1 \end{pmatrix} = \begin{pmatrix} 1&1&0 \\ -1&1&0 \\ -1&-1&1 \end{pmatrix}.$$

Let $X = PY'$, where, $Y^T = (y_1 \ \ y_2 \ \ y_3)$ and P is non-singular, transforms the quadratic form into the normal form $= Y^TDY$, then $D(= P^TAP)$ is a diagonal matrix. Thus the transformation $X = PY$ becomes,

$$x_1 = y_1 - y_2 - y_3, x_2 = y_1 + y_2 - y_3, x_3 = y_3.$$

Ex 7.6.7 *Reduce the quadratic form $2x_1x_3 + x_2x_3$ to diagonal form.*

Solution: Since the diagonal terms are absent and the coefficient of x_1x_3 is non zero we make the change of variables $x_1 = y_1, x_2 = y_2$ and $x_3 = y_3 + y_1$. The quadratic form is transform to

$$\begin{aligned} & 2y_1^2 + y_1y_2 + 2y_1y_3 + y_2y_3 \\ &= 2(y_1 + \frac{1}{4}y_2 + \frac{1}{2}y_3)^2 - \frac{1}{8}y_2^2 - \frac{1}{2}y_3^2 + \frac{1}{2}y_2y_3. \end{aligned}$$

With $z_1 = y_1 + \frac{1}{4}y_2 + \frac{1}{2}y_3, z_2 = y_2$ and $z_3 = y_3$, the quadratic form becomes,

$$2z_1^2 - \frac{1}{8}z_2^2 - \frac{1}{2}z_3^3 + \frac{1}{2}z_2z_3 = 2z_1^2 - \frac{1}{8}(z_2 - 2z_3)^2.$$

Finally, making the transformation $z_1 = u_1, z_2 - 2z_3 = u_2$ and $z_3 = u_3$ the quadratic form becomes $2u_1^2 - \frac{1}{8}u_2^2$.

It can be checked that the transformation from x's to u's is $x_1 = u_1 - \frac{1}{4}u_2 - u_3, x_2 = u_2 + 2u_3$ and $x_3 = u_1 - \frac{1}{4}u_2$. The u's can be expressed in terms of x's as $u_1 = \frac{1}{2}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3, u_2 = 2x_1 + x_2 - 2x_3$ and $u_3 = x_3 - x_1$. Thus the given quadratic form can be written as

$$2\left(\frac{1}{2}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3\right)^2 - \frac{1}{8}(2x_1 + x_2 - 2x_3)^2$$

$$\text{or,}\ \left(\frac{1}{\sqrt{2}}x_1 + \frac{1}{2\sqrt{2}}x_2 + \frac{1}{\sqrt{2}}x_3\right)^2 - \left(\frac{1}{\sqrt{2}}x_1 + \frac{1}{2\sqrt{2}}x_2 - \frac{1}{\sqrt{2}}x_3\right)^2,$$

with coefficients $1, -1$ and 0.

7.6.1 Jordan Canonical Form

We have shown that every complex matrix is similar to an upper triangular matrix. Also it is similar to a diagonal matrix if and only if its minimal polynomial has distinct roots. When the minimal polynomial has repeated roots then the Jordan canonical form theorem implies that it is similar to $D + N$, where D is a diagonal matrix with eigen values as the diagonal elements of D and N is a nilpotant matrix with a suitable simplified form namely the first super diagonal elements of N are either 1 or 0, with at least 1. In other words the matrix is similar to a block diagonal matrix, where each diagonal block is of the form

$$\begin{pmatrix} \lambda' & & \\ & \lambda' & \\ & & \lambda' \end{pmatrix}$$

whose diagonal elements are all equal to an eigen value λ and first super diagonal element are all 1. This block diagonal form is known as *Jordan canonical form.*

Result 7.6.1 Let $T : V \to V$ be a linear operator, whose characteristic and minimal polynomial are

$$P(\lambda) = (\lambda - \lambda_1)^{n_1}\cdots(\lambda - \lambda_r)^{n_r}; m(\lambda) = (\lambda - \lambda_1)^{m_1}\cdots(\lambda - \lambda_r)^{m_r}$$

where λ_i are the distinct eigen values with algebraic multiplicity n_i, m_i respectively with $m_i \le n_i$, then T has a block diagonal matrix representation J whose diagonal entries are of the form J_{ij}, where

$$J_{ij} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

For each λ_i the corresponding blocks J_{ij} has the following properties

(i) There is at least one J_{ij} of order m_i and all other J_{ij} with λ_i as diagonal element are all of order $\le m_i$.

(ii) The sum of orders of J_{ij} in n_i.

(iii) The number of J_{ij} having diagonal element λ_i = the geometric multiplicity of λ_i.

(iv) The J_{ij} of each possible order is uniquely determined.

A k^{th} order Jordan submatrix referring to the number λ_0 is a matrix of order $k, 1 \le k \le n$, of the form

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \\ 0 & 0 & 0 & \cdots & \lambda_0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_0 \end{pmatrix}.$$

In other words, one and same number λ_0 form the field F occupies the principal diagonal, with unity along the diagonal immediately above and zero elsewhere. Thus

$$[\lambda_0], \begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}, \begin{bmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{bmatrix}$$

are respectively Jordan submatrices of first, second and third order. A *Jordan matrix of order n* is a matrix of order n having the form

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & J_n \end{pmatrix}.$$

The elements along the principal diagonal are *Jordan submatrices or Jordan blocks* of certain orders, not necessarily distinct, referring to certain numbers (not necessarily distinct either) lying in the field F. Thus, a matrix is a Jordan matrix if and only if it has form

$$\begin{pmatrix} \lambda_1 & \varepsilon_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \varepsilon_2 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & \varepsilon_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

where $\lambda_i; i = 1, 2, \cdots, n$ are arbitrary numbers in F and every $\varepsilon_j; j = 1, 2, \cdots, n-1$ is equal to unity or zero. Note that if $\varepsilon_j = 1$, then $\lambda_j = \lambda_{j+1}$. Diagonal matrices are a special case of Jordan matrices. These are Jordan matrices whose submatrices are of order 1.

Theorem 7.6.1 *Let J be a Jordan block of order k. Then J has exactly one eigenvalue, which is equal to the scalar on the main diagonal. The corresponding eigenvectors are the non zero scalar multiples of the k dimensional unit coordinate vector* $[1, 0, \cdots, 0]$.

Proof: Suppose that the diagonal entries of J are equal to λ. A column vector $X = [x_1, x_2, \cdots, x_k]^T$ satisfies the equation $JX = \lambda X$ if and only if its components satisfy the following k scalar equations:

$$\begin{aligned} \lambda x_1 + x_2 &= \lambda x_1 \\ \lambda x_2 + x_3 &= \lambda x_2 \\ &\vdots \\ \lambda x_{k-1} + x_k &= \lambda x_{k-1} \\ \lambda x_k &= \lambda x_k. \end{aligned}$$

From the first $(k-1)$ equations, we obtain $x_2 = x_3 = \cdots = x_k = 0$, so λ is an eigenvalue for J and all eigenvectors have the same form $x_1[1, 0, \cdots, 0]$ with $x_1 \neq 0$. To show that λ is the only eigenvalue for J, assume that $JX = \mu X$ for some scalar $\mu \neq \lambda$. Then the components satisfy the following k scalar equations

$$\begin{aligned} \lambda x_1 + x_2 &= \mu x_1 \\ \lambda x_2 + x_3 &= \mu x_2 \\ &\vdots \\ \lambda x_{k-1} + x_k &= \mu x_{k-1} \\ \lambda x_k &= \mu x_k. \end{aligned}$$

Because $\lambda \neq \mu$, the last relation gives $x_k = 0$ and from the other equations we get $x_{k-1} = x_{k-2} = \cdots = x_2 = x_1 = 0$. Hence only zero vector satisfies $JX = \mu X$, so no scalar different from λ can be an eigen value for J.
This theorem describes all the eigenvalues and eigenvectors of a Jordan block.

Ex 7.6.8 *Find all possible Jordan canonical forms for those matrices whose characteristic polynomial $\Delta(t)$ and the minimal polynomial $m(t)$ are as follows*

1. $\Delta(t) = (t-2)^5;\ m(t) = (t-2)^2$.
2. $\Delta(t) = (t-7)^5;\ m(t) = (t-7)^2$.
3. $\Delta(t) = (t-2)^7;\ m(t) = (t-2)^3$.
4. $\Delta(t) = (t-2)^4(t-5)^3;\ m(t) = (t-2)^2(t-5)^3$.
5. $\Delta(t) = (t-2)^4(t-3)^2;\ m(t) = (t-2)^2(t-3)^2$.

Solution: (1) Since $\Delta(t)$ has degree 5, J must be a 5×5 matrix, and all diagonal elements must be 2, since 2 is the only eigenvalue. Moreover, since the exponent of $t-2$ in $m(t)$ is 2, J must have one Jordan block of order 2, and the other must be of order 2 or 1. Thus there are only two possibilities

(i) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2]\right)$.

(ii) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2], [2], [2]\right)$.

(2) In the similar ways, there are only two possibilities

(i) $J = diag\left(\begin{bmatrix} 7 & 1 \\ 0 & 7 \end{bmatrix}, \begin{bmatrix} 7 & 1 \\ 0 & 7 \end{bmatrix}, [7]\right)$.

(ii) $J = diag\left(\begin{bmatrix} 7 & 1 \\ 0 & 7 \end{bmatrix}, [7], [7], [7]\right)$.

(3) Let M_k denote a Jordan block with $t = 2$ of order k. Then, in the similar ways, there are only four possibilities

(i) $diag(M_3, M_3, M_1)$.

(ii) $diag(M_3, M_2, M_2)$.

Scanned by CamScanner

(iii) $diag(M_3, M_2, M_1, M_1)$.

(iv) $diag(M_3, M_1, M_1, M_1, M_1)$.

(4) The Jordan canonical form is one of the following block diagonal matrices:

(i) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}\right)$.

(ii) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2], [2], \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}\right)$.

The first matrix occurs if the T has two independent eigenvectors belonging to the eigenvalue 2 and the second matrix occurs if the linear operator T has three independent eigenvectors belonging to 2.

(5) The Jordan canonical form is one of the following block diagonal matrices:

(i) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}\right)$.

(ii) $J = diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2], [2], \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}\right)$.

Result 7.6.2 Let V be an n dimensional linear space with complex scalars and let $T : V \to V$ be a linear transformation of V into itself. Then there is a basis for V relative to which T has a block diagonal matrix representation $diag(J_1, J_2, \cdots, J_m)$, with each J_k being a Jordan block.

Ex 7.6.9 *Find all possible Jordan canonical forms for a linear operator $T : V \to V$ whose characteristic polynomial $\Delta(t) = (t-2)^3(t-5)^2$. In each case, find the minimal polynomial $m(t)$.*

Solution: Since $t-2$ has exponent 3 in $\Delta(t)$, 2 must appear three times on the diagonal. Similarly, 5 must appear twice. Thus there are six possibilities:

$$(i)diag\left(\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}\right), \quad (ii)diag\left(\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, [5], [5]\right)$$

$$(iii)diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, [2], \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}\right), \quad (iv)diag\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}; [2], [5], [5]\right)$$

$$(v)diag\left([2], [2], [2] \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}\right), \quad (vi)diag\left([2], [2], [2], [5], [5]\right).$$

The exponent in the minimal polynomial $m(t)$ is equal to the size of the largest block. Thus

$(i)m(t) = (t-2)^3(t-5)^2$, $(ii)m(t) = (t-2)^3(t-5)$, $(iii)m(t) = (t-2)^2(t-5)^2$
$(iv)m(t) = (t-2)^2(t-5)$, $(v)m(t) = (t-2)(t-5)^2$, $(vi)m(t) = (t-2)(t-5)$.

Ex 7.6.10 *Verify that the matrix $A = \begin{pmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{pmatrix}$ has eigen values $2, -1, -1$. Find a non singular matrix C with initial entry $C_{11} = 1$ that transforms A to the Jordan canonical form* $C^{-1}AC = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$.

Solution: The characteristic equation of the matrix A is $|A - \lambda I| = 0$, i.e.,

$$\begin{vmatrix} -1-\lambda & 3 & 0 \\ 0 & 2-\lambda & 0 \\ 2 & 1 & -1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (2-\lambda)(1+\lambda)^2 = 0 \Rightarrow \lambda = -1, -1, 2.$$

The corresponding eigenvector corresponding to $\lambda = 2$ is obtained by solving equations

$$-3x + 3y = 0; 2x + y - 3z = 0 \Rightarrow x = y = z.$$

The eigenvector corresponding to $\lambda = 2$ is $k(1,1,1)$, where k is nonzero constant. The corresponding eigenvector corresponding to $\lambda = -1$ is obtained by solving equations

$$3y = 0; 2x + y = 0 \Rightarrow x = y = 0.$$

The eigenvector corresponding to $\lambda = -1$ is $(0,0,a)$, where a is arbitrary nonzero number. We construct the matrix C whose first two columns are the eigenvectors corresponding to $\lambda = 2$ and $\lambda = -1$. Since $C_{11} = 1$, we must have $k = 1$. The third column is chosen in such a way that $AC = CB$, where B is the Jordan canonical form, say $C = \begin{pmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & d \end{pmatrix}$. Therefore,

$$\begin{pmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & d \end{pmatrix} = \begin{pmatrix} 1 & 0 & b \\ 1 & 0 & c \\ 1 & a & d \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

$$\Rightarrow -b + 3c = -b, 2c = -c, 2b + c - d = a - d \Rightarrow c = 0, a = 2b.$$

Hence $C = \begin{pmatrix} 1 & 0 & b \\ 1 & 0 & 0 \\ 1 & 2b & d \end{pmatrix}$, where $b \neq 0$ and d is arbitrary.

7.7 Functions of Matrix

As we define and study various functions of a variable in algebra, it is possible to define and evaluate functions of a matrix. We shall study the following functions of a matrix in this chapter: integral powers(positive and negative), fractional powers(roots), exponential, logarithmatic, trigonometric and hyperbolic functions.

There are two methods by which a function of a matrix can be evaluated. The first is a rather straightforward method based on the diagonalization of a matrix and is therefore applicable to diagonizable matrices only. The second method is based on the existence of a minimal polynomial and can be used to evaluate functions of any matrix.

Functions of diagonizable matrix

Let A be a diagonizable matrix and let P be a diagonalizing matrix for A, so that

$$P^{-1}AP = \Lambda, \quad A = P\Lambda P^{-1}, \tag{7.27}$$

where Λ is a diagonal matrix containing eigenvalues of A. Now, if f is any function of a matrix, then we have

$$f(A) = Pf(\Lambda)P^{-1}. \tag{7.28}$$

Thus, if we can define a function of a diagonal matrix, we can define and evaluate the function of any diagonalizable matrix. The discussion of this chapter evidently applies to square matrices only.