3.2 SIMILAR MATRICES

We shall now introduce the idea of similar matrices and study the properties of similar matrices.

DEFINITION

An nxn matrix A is said to be similar to a nxn matrix B if there exists a nonsingular nxn matrix P such that,

$P^{-1} A P = B$

We then write,

$A \sim B$

Properties of Similar Matrices

(1) Since $I^{-1} A I = A$ it follows that $A \sim A$

(2) $A \sim B \Rightarrow \exists P$, nonsingular such that., $P^{-1} A P = B$

$\Rightarrow A = P B P^{-1}$

$\Rightarrow A = Q^{-1} B P$, where $Q = P^{-1}$ is nonsingular

$\Rightarrow \exists$ nonsingular Q show that $Q^{-1} B Q = A$
$\Rightarrow B \sim A$

Thus

$A \sim B \Rightarrow B \sim A$

(3) Similarly, we can show that

$A \sim B, B \sim C \Rightarrow A \sim C.$

(4) Properties (1), (2) and (3) above show that similarity is an equivalence relation on the set of all nxn matrices.

(5) Let A and B be similar matrices. Then there exists a nonsingular matrix P such that

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

Now, let $C_A(\lambda)$ and C_B (λ) be the characteristic polynomials of A and B respectively. We have,

$$C_A(\lambda) = |\lambda I - A| = |\lambda I - P^{-1}BP|$$

$$= |\lambda P^{-1}P - P^{-1}BP|$$

$$= |P^{-1}(\lambda I - B)P|$$

$$= |P^{-1}||\lambda I - B||P|$$

$$= |\lambda I - B| \text{ since } |P^{-1}||P| = 1$$

$$= C_B(\lambda)$$

Thus " SIMILAR MATRICES HAVE THE SAME CHARACTERISTIC POLYNOMIALS ".

(6) Let A and B be similar matrices. Then there exists a nonsingular matrix P such that

$$A = P^{-1} B P$$

Now for any positive integer k, we have

$$A^k = \underbrace{(P^{-1}BP)(P^{-1}BP).....(P^{-1}BP)}_{\text{k times}}$$

$$= P^{-1} B^k P \text{ (since } PP^{-1}=I)$$

Therefore,

$$A^k = O_n \Leftrightarrow P^{-1} B^k P = O_n$$

$$\Leftrightarrow B^k = O_n$$

" Thus if A and B are similar matrices then $A^k = O_n \Leftrightarrow B^k = O_n$ ".

7) Let A and B be any two square matrices of the same order, and let $p(\lambda) = a_0 + a_1\lambda + ... + a_k\lambda$ be any polynomial.

Then

$$p(A) = a_0 I + a_1 A + ... + a_k A^k$$

$$= a_0 I + a_1 P^{-1} BP + a_2 P^{-1} B^2 P + ... + a_k P^{-1} B^k P$$

$$= P^{-1}\left[a_0 I + a_1 B + a_2 B^2 + ... + a_k B^k\right]P$$

$$= P^{-1} p(B) P$$

Thus

$$p(A) = O_n \Leftrightarrow P^{-1} p(B) P = O_n$$

$$\Leftrightarrow p(B) = O_n$$

Thus " IF A and B ARE SIMILAR MATRICES THEN FOR ANY POLYNOMIAL $p(\lambda)$; p (A) = $O_n \Leftrightarrow$ p (B) = O_n ".

(8) Let A be any matrix. By A(A) we denote the set of all polynomials $p(\lambda)$ such that

p(A) = O_n, i.e.

$\mathscr{A}$ (A) = $\{p(\lambda) : p(A) = O_n\}$

Now from (6) it follows that,

" IF A AND B ARE SIMILAR MATRICES THEN A(A) = A (B) ".

The set $\mathscr{A}$ (A) is called the set of " ANNIHILATING POLYNOMIALS OF A ". Thus similar matrices have the same set of annihilating polynomials.

We shall discuss more about annihilating polynomials later.

We now investigate the following question? Given an nxn matrix A, when is it similar to a "simple matrix"? What are simple matrices? The simplest matrix we know is the zero matrix O_n. Now A ~ $O_n \Leftrightarrow$ There is a nonsingular matrix P such that A = $P^{-1} O_n$ P = O_n.

$\therefore$ " THE ONLY MATRIX SIMILAR TO O_n IS O_n ITSELF ".

The next simple matrix we know is the identity matrix I_n. Now A ~ $I_n \Leftrightarrow$ there is a nonsingular P such that A = $P^{-1} I_n$ P $\Leftrightarrow$ A = I_n.

Thus "THE ONLY MATRIX SIMILAR TO I_n IS ITSELF ".
Similarly the only matrix similar to a scalar matrix kI_n , (where k is a scalar),is kI_n itself.

The next class of simple matrices are the DIAGONAL MATRICES. So we now ask the question " Which type of nxn matrices are similar to diagonal matrices"?

Suppose now A is an nxn matrix; and A is similar to a diagonal matrix,

$$D=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & \bigcirc & \\ & & \ddots & \\ \bigcirc & & & \lambda_n \end{pmatrix}$$

(λ_I not necessarily distinct).

Then there exists a nonsingular matrix P such that

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

$\Rightarrow$

$$AP = PD \qquad \text{..........(1)}$$

$$Let P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1i} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2i} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{ni} & \cdots & p_{nn} \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$Let P_i = \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$$ denote the i^{th} column of P.

Now the i^{th} column of AP is

$$\begin{pmatrix} a_{11}p_{1i} + a_{12}p_{2i} + \ldots\ldots + a_{1n}p_{ni} \\ a_{21}p_{1i} + a_{22}p_{2i} + \ldots\ldots + a_{2n}p_{ni} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}p_{1i} + a_{n2}p_{2i} + \ldots\ldots + a_{nn}p_{ni} \end{pmatrix}$$

which is equal to $AP_{i.}$
Thus the i^{th} column of AP, the l.h.s. of (1), is AP_i.

Now the i^{th} column of P D is

$$\begin{pmatrix} p_{1i}\lambda_i \\ p_{2i}\lambda_i \\ \vdots \\ p_{ni}\lambda_i \end{pmatrix} = \lambda_i \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix} = \lambda_i P_i$$

Thus the i^{th} column of P D, the r.h.s. of (1), is $\lambda_i P_i$. Since l.h.s. = r.h.s. by (1) we have

$AP_i = \lambda_i P_i$; i = 1, 2, …., n ……………..(2)

Note that since P is nonsingular no column of P can be zero vector. Thus none of the column vectors P_i are zero. Thus we conclude that,

"IF A IS SIMILAR TO A DIAGONAL MATRIX D THEN THE DIAGONAL ENTRIES OF D MUST BE THE EIGENVALUES OF A AND IF $P^{-1}AP = D$ THEN THE i^{th} COLUMN VECTOR OF P MUST BE AN EIGENVECTOR CORRESPON DING TO THE EIGENVALUE WHICH IS THE i^{th} DIAGONAL ENTRY OF D".

<u>Note:</u>

The n columns of P must be linearly independent since P is nonsingular and thus these n columns give us n linearly independent eigenvectors of A

Thus the above result can be restated as follows: A is similar to a diagonal matrix D and P^{-1} A P = D $\Rightarrow$ A has n linearly independent eigenvectors; taking these as the columns of P we get P^{-1} A P we get D where the i^{th} diagonal entry of D is the eigenvalue corresponding to the i^{th} eigenvector, namely the I^{th} column vector of P. .

Conversely, it is now obvious that if A has n linearly independent eigenvectors then A is similar to a diagonal matrix D and if P is the matrix whose i^{th} column is the eigenvector, then D is P^{-1} A P and i^{th} diagonal entry of D is the eigenvalue corresponding to the i^{th} eigenvector.

When does then a matrix have n linearly independent eigenvectors? It can be shown that a matrix A has n linearly independent eigenvectors $\Leftrightarrow$ the algebraic multiplicity of each eigenvalue of A is equal to its geometric multiplicity. Thus

A IS SIMILAR TO A DIAGONAL MATRIX $\Leftrightarrow$
FOR EVERY EIGENVALUE OF A, ITS ALGEBRAIC MULTIPLICITY IS EQUAL TO ITS GEOMETRIC MULTPLICITY".

RECALL; if $C(\lambda)=(\lambda-\lambda_1)^{a_1}(\lambda-\lambda_2)^{a_2}.....(\lambda-\lambda_k)^{a_k}$ where λ_1, λ_2, ….., λ_k are the distinct eigenvalues of A, then a_i is called the algebraic multiplicity of the eigenvalue λ_i. Further, let

$$\omega_i=\{x : Ax=\lambda_i x\}$$

be the eigensubspace corresponding to λ_i. Then g_i = dim ω_i is called the geometric multiplicity of λ_i.

Therefore, we have,

" If A is an nxn matrix with $C(\lambda)=(\lambda-\lambda_1)^{a_1}\cdots(\lambda-\lambda_k)^{a_k}$ where λ_1, ….., λ_k are the district eigenvalues of A, then A is similar to a diagonal matrix $\Leftrightarrow$ $a_i = g_i$ (=dimω_i) ; $1\le i \le k$".

<u>Example:</u>

Let us now consider

$$A=\begin{pmatrix}-9 & 4 & 4\\ -8 & 3 & 4\\ -16 & 8 & 7\end{pmatrix}$$

On page 87, we found the characteristic polynomial of A as

$C(\lambda) = (\lambda+1)^2 (\lambda - 3)$

Thus $\lambda_1 = -1$; $a_1 = 2$
$\lambda_2 = 3$; $a_2 = 1$

On pages 83 and 84 we found,

W_1 = eigensubspace corresponding to λ = -1

$$= \left\{ x : x = A_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + A_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

W_2 = eigensubspace corresponding to λ = 3

$$= \left\{ x : x = k \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$$

Thus dim $W_1 = 2$ $\quad \therefore g_1 = 2$
dim $W_2 = 1$ $\quad \therefore g_2 = 1$

$$\text{Thus } \left.\begin{array}{l} a_1 = 2 = g_1 \\ a_2 = 1 = g_2 \end{array}\right\} \text{ and hence A must be similar.}$$

to a diagonal matrix. How do we get P such that $P^{-1}AP$ is a diagonal matrix? Recall the columns of P must be linearly independent eigenvectors. From ω_1 we get two linearly eigenvectors, namely, $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$; and from ω_2 we get third as $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Thus if we take these as columns and write,

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

Then $P^{-1} = \begin{pmatrix} 1 & 0 & -1/2 \\ 2 & -1 & -1/2 \\ -2 & 1 & 1 \end{pmatrix}$; and it can be verified that

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & -\tfrac{1}{2} \\ 2 & -1 & -\tfrac{1}{2} \\ -2 & 1 & 1 \end{pmatrix}\begin{pmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ a diagonal matrix., whose diagonal entries are the eigenvalues of the matrix A.

Thus we can conclude that A is similar to a diagonal matrix, i.e., $P^{-1}AP = D$
$\Rightarrow$ A has n linearly independent eigenvectors namely the n columns of P.

Conversely, A has n linearly independent eigenvectors.
$\Rightarrow$ $P^{-1}AP$ is a diagonal matrix where the columns of P are taken to be the n linearly eigenvectors.

We shall now see a class of matrices for which it is easy to decide whether they are similar to a diagonal matrix; and in which case the P^{-1} is easy to compute. But we shall first introduce some preliminaries.

If $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$; $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ are any two vectors in C^n, we define the INNER PRODUCT OF x with y (which is denoted by (x,y)) as,

$$(x, y) = x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n$$

$$= \sum_{i=1}^{n} x_i\overline{y}_i$$

Example 1:

If $x = \begin{pmatrix} i \\ 2+i \\ -1 \end{pmatrix}$; $y = \begin{pmatrix} 1 \\ 1-i \\ i \end{pmatrix}$; then,

$$(x, y) = i.\overline{1} + (2+i)\overline{(1-i)} + (-1)(\overline{i})$$

$$= i + (2+i)(1+i) + (-1)(-i) = 1 + 5i$$

Whereas $(y,x) = 1(\bar{i}) + (1-i)\overline{(2+i)} + (i)\overline{(-1)} = 1 - 5i$

We now observe some of the properties of the inner product, below:

(1) For any vector x in C^n, we have

$$(x,x) = \sum_{i=1}^{n} x_i \overline{x}_i = \sum_{i=1}^{n} |x_i|^2 ,$$

which is real ≥ 0. Further,

$$(x,x) = 0 \Leftrightarrow \sum_{i=1}^{n} |x_i|^2 = 0$$

$$\Leftrightarrow x_i = 0; 1 \leq i \leq n$$

$$\Leftrightarrow x = \theta_n$$

Thus,

(x,x) is real and ≥ 0 and (x,x)= 0 $\Leftrightarrow$ x = θ_n

(2) $$(x,y) = \sum_{i=1}^{n} x_i \overline{y}_i = \left(\overline{\sum_{i=1}^{n} y_i \overline{x}_i} \right)$$

$$= \overline{(y,x)}$$

Thus,

$$(x,y) = \overline{(y,x)}$$

(3) For any complex number α, we have,

$$(\alpha x, y) = \sum_{i=1}^{n} (\alpha x_i) \overline{y}_i = \alpha \sum_{i=1}^{n} x_i \overline{y}_i$$

$$= \alpha (x,y)$$

Thus

$(\alpha x,y) = \alpha\,(x,y)$ for any complex number α.

We note,

$$(x, \alpha y) = \overline{(\alpha y, x)} \quad \text{by (2)}$$

$$= \overline{\alpha (y,x)} = \overline{\alpha}\,\overline{(y,x)} = \overline{\alpha}(x,y)$$

(4) $$(x+y, z) = \sum_{i=1}^{n} (x_i + y_i)\overline{z}_i$$

$$= \sum_{i=1}^{n} x_i \overline{z_i} + \sum_{i=1}^{n} y_i \overline{z_i}$$

$$= (x, y) + (x, z)$$

Thus

$$\begin{cases} (x + y, z) = (x,z) + (y,z) \text{ and similarly} \\ (x, y + z) = (x, y) + (x, z) \end{cases}$$

We say that two vectors x and y are <u>ORTHOGONAL</u> if (x, y) = 0.

<u>Example</u> (1) If x = $\begin{pmatrix} 1 \\ i \\ -i \end{pmatrix}; y = \begin{pmatrix} -1 \\ i \\ 0 \end{pmatrix},$

then,

$$(x, y) = 1\overline{(-1)} + i\overline{(i)} + (-i)\overline{(0)}$$

= -1 + 1 = 0

Thus x and y are orthogonal.

(2) If $x = \begin{pmatrix} 1 \\ i \\ -i \end{pmatrix}, y = \begin{pmatrix} -1 \\ a \\ 1 \end{pmatrix}$

then

$$(x, y) = -1 + \overline{a}i - i$$

$\therefore$ x, y orthogonal

$$\Leftrightarrow -(1 + i) + \overline{a}i = 0$$

$$\Leftrightarrow \overline{a} = \left(\frac{1 + i}{i}\right) = -i(1 + i) = 1 - i$$

$$\Leftrightarrow a = 1 + i$$