NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: MTH 212

COURSE TITLE: LINEAR ALGEBRA

UNIT 1 LINEAR TRANSFORMATIONSI

CONTENTS

1.0 INTRODUCTION

You have already learnt about a vector space and several concepts related to it. In this unit we initiate the study of certain mappings between two vector spaces, called linear transformations. The importance of these mappings can be realized from the fact that, in the calculus of several variables, every continuously differentiable function can be replaced, to a first approximation, by a linear one. This fact is a reflection of a general principle that every problem on the change of some quantity under the action of several factors can be regarded, to a first approximation, as a linear problem. It often turns out that this gives an adequate result. Also, in physics it is important to know how vectors behave under a change of the coordinate system. This requires a study of linear transformations.

In this unit we study linear transformations and their properties, as well as two spaces associated with a linear transformations and their properties, as well as two spaces associated with a linear transformation, and their dimensions. Then, we prove the existence of linear transformations with some specific properties, as discuss the notion of an isomorphism between two vector spaces, which allows us to say that all finite-dimensional vector spaces of the same dimension are the “Same”, in a certain sense.

Finally, we state and prove the Fundamental Theorem of Homomorphism and some of its corollaries, and apply them to various situations.

2.0 OBJECTIVES

After reading this unit, you should be able to:

- Verify the linearity of certain mappings between vector spaces;
- Construct linear transformations with certain specified properties;
- Calculate the rank and nullity of a linear operator;
- Prove and apply the Rank Nullity Theorem;
- Define an isomorphism between two vector spaces;
- Show that two vector spaces are isomorphic if and only if they have the same dimension;
- Prove and use the Fundamental Theorem of homormorphism.

3.0 MAIN CONTENT

3.1 Linear Transformations

By now you are familiar with vector spaces IR^2 and IR^3. Now consider the mapping $f:R^2 \rightarrow R^3$:f(x,y) = (x,y,0) (see fig. 1).

F is a well defined function. Also notice that

(i) f((a,b) + (c, d)) = f((a + c, b + d)) = (a + c, B + d,0) = (a, b, 0) + (c,d,0)

= f((a, b)) + f((c,d)), for (a,b), (c,d)€ R^2 and

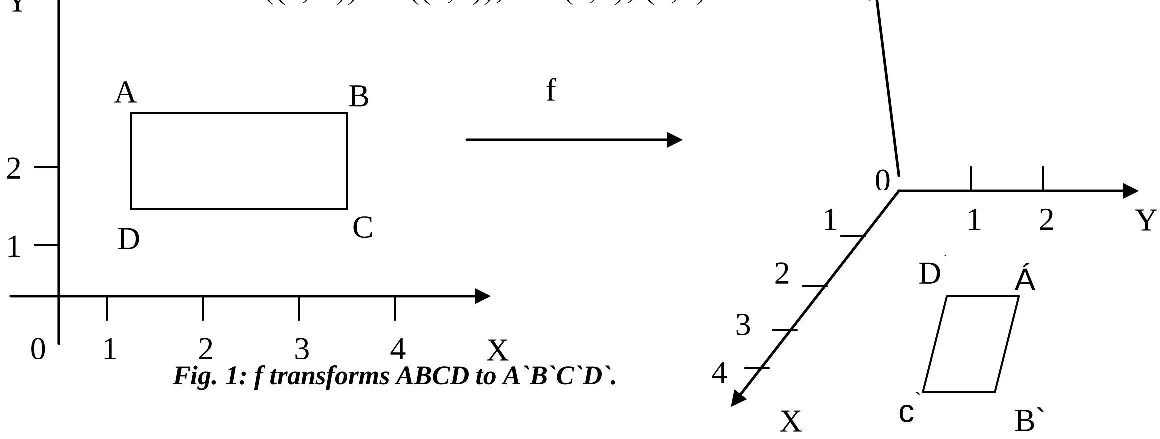

Fig. 1: f transforms ABCD to A`B`C`D`.

(ii) for any α € R and (a,b) € R2, f((αa, αb)) = (αa, αb, 0) = α (a,b,0) = αf((a,b)).
So we have a function f between two vector spaces such that (i) and (ii) above hold true.

(i) Says that the sum of two plane vectors is mapped under f to the sum to sum of their images under f.

(ii) Says that a line in the plane R^2 is mapped under f to a line in R^2.

The properties (i) and (ii) together say that f is linear, a term that we now define.

Definition: Let U and V be vector spaces over a field F. A linear transformation (or linear operator) from U to V is a function T: U $\rightarrow$ V, such that LT1) $T(u_1 + u_2) = T(u_1) + T(u_1)$, for $U_1, U_2 \in U$, and LT2) $T(\alpha u) = \alpha T(u)$ for $\alpha \in F$ and $U \in U$.

The conditions LT1 and LT2 can be combined to give the following equivalent condition. LT3) $T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2)$, for $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$.

What we are saying is that [LT1 and LT2] $\Leftrightarrow$LT3. This can be easily shown as follows:

We will show that LT3 $\Rightarrow$LT1 and LT3 $\Rightarrow$ LT2. Now, LT3 is true $\forall$ $\alpha_1\alpha_2 \in F$. Therefore, it is certainly true for $\alpha_{1 = 1}\, \alpha_{2\,,}$ that is, LT1 holds.

Now, to show that LT2 is true, consider $T(\alpha u)$ for any $\alpha \in F$ and $u \in U$. We have $T(\alpha u) = T(\alpha u + 0.\ u) = \alpha T(u) + 0.T(u) = \alpha T(u)$, thus proving that LT2 holds.

You can try and prove the converse now. That is what the following exercise is all about!

E E1) Show that the conditions LT1 and LT2 together imply LT3.

Before going further, let us note two properties of any linear transformation T:U$\rightarrow$ V, which follow from LT1 (or LT2, or LT3).

LT4) T(0) = 0. Let's see why this is true. Since T(0) = T(0 + 0) = T(0) +T(0) (by LT1, we subtract T(0) from both sides to get T(0) = 0.

LT5) $T(-u) = - T(u)$ $\forall$ $u \in U$. why is this so? Well, since $0 = T(0) = T(u - u) = T(u) + T(-u)$, we get $T(-u) = - T(u)$.

E E2) Can you show how LT4 and LT5 will follow from LT2?

Now let us look at some common linear transformations.

Example 1: Consider the vector space U over a field F, and the function $T:U \rightarrow U$ defined by $T(u) = u$ for all $u \in U$.

Show that T is a linear transformation. (This transformation is called the **identity transformation,** and is denoted by I_u, or Just I, if the underlying vector space is understood).

Solution: For any $\alpha, \beta \in F$ and $u_1, u_2 \in U$, we have

$T(\alpha u_1 + \beta u_2) = \alpha u_1 + \beta u_2 = \alpha T(u_1) + \beta T(u_2)$

Hence, LT3 holds, and T is a linear transformation.

Example 2: Let $T: U \rightarrow V$ be defined by $T(u) = 0$ for all $u \in U$.

Check that T is a linear transformation. (It is called the null or **Zero Transformation,** and is denoted by **0**).

Solution: For any $\alpha, \beta, \in F$ and $u_1, u_2 \in U$, we have $T(\alpha u_1 + \beta u_2) = 0 = \alpha.0 + \beta.0 = \alpha T(u_1) + \beta T(u_2)$.

Therefore, T is linear transformation.

Example 3: Consider the function $pr_1:R^n \rightarrow R$, defined by $pr_1[(x_1, \ldots\ldots\ldots x_n)] = x_1$. Shows that this is a linear transformation. (This is called the projection on the first coordinate. Similarly, we can define $pr_1:R_n \rightarrow R$ by $pr_1 [(x_1 \ldots. X_i\text{-}1, X_i, \ldots\ldots\ldots, x_n)] = x_1$ to be the **projection** on the i^{th}**Coordinate** for i = 2, ...,n. For instance, $pr_2:R^3 \rightarrow R$: $pr_2(x,y,z) = y$.)

Solution: We will use LT3 to show that projection is a linear operator. For $\alpha,\beta \in R$ and $(x_1 \ldots\ldots\ldots x_n)$, $(y_1 \ldots\ldots\ldots\ldots y_n)$ in R^n, we have

$Pr_1 [\alpha(x_1 \ldots\ldots\ldots\ldots x_n) + \beta(y_1 \ldots\ldots\ldots\ldots y_n)]$

$= pr_1 (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2 \ldots\ldots\ldots\ldots \alpha x_n + \beta y_n) = \alpha x_1 + \beta y_1$
$= \alpha pr_1,[(x_1,\ldots..,x_n)] + \beta pr_1 [(y_1,\ldots.,y_n)]$.

Thus pr1 (and similarly pr1) is a linear transformation.

Before going to the next example, we make a remake about projections.

Remark: Consider the function $p:R^3 \rightarrow R^2$: $p(x,y,z) = (x,y)$. this is a projection from R^3 on to the xy-plane. Similarly, the functions f and g, from $R^3 \rightarrow R^2$, defined by $f(x,y,z) = (x,z)$ and $g(x,y,z) = (y,z)$ are projections from R^3 onto the xz-plane and the yz-plane, respectively.

In general, any function $\theta : R^n \rightarrow R^m$ (n> m), which is defined by dropping any (n – m) coordinates, is a projection map.

Now let us see another example of a linear transformation that is very geometric in nature.

Example 4: Let $T:R2 \rightarrow R^2$ be defined by $T(x,y) = (x, - y)\ \forall\ x,y \in R$. Show that T is a linear transformation. (This is the **reflection** in the x-axis that we show in fig 2).

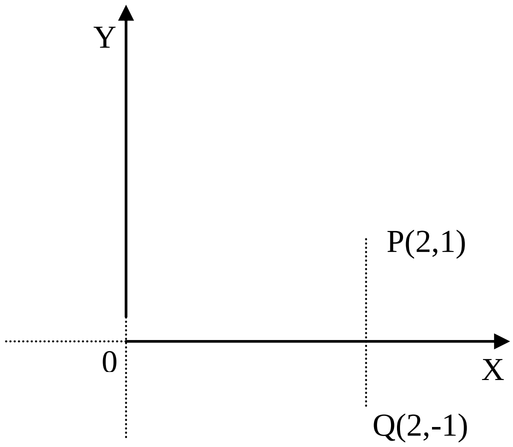

Fig 2: Q is the reflection of P in the X-axis.

Solution: For $\alpha, \beta \in R$ and $(x_1y_1), (x_2, y_2) \in R^2$, we have

$T[\alpha (x_1,y_1) + \beta (x_2,y_2) = T (\alpha x_1, + \beta x_2, \alpha y_1 + \beta y_2) = (\alpha x_1 + \beta x_2, - ay_1 - \beta y_2)$

$= \quad \alpha (x1, - y1) + \beta (x2, - y2)$
$= \quad \alpha T (x_1, y_1) + \beta T(x_2, y_2)$.

Therefore, T is a linear transformation.

So, far we've given examples of linear transformations. Now we give an example of a very important function which is not linear. This example's importance lies in its geometric applications.

Example 5: Let u_0 be a fixed non-zero vector in U. Define $T : U \rightarrow U$ by $T(u) = u + u_0 \forall u \in U$. Show that T is not a linear transformation. (T is called the translation by u_0. See Fig 3 for a geometrical view).

Solution: T is not a linear transformation since LT4 does not hold. This is because $T(0) = u_0 \neq 0$

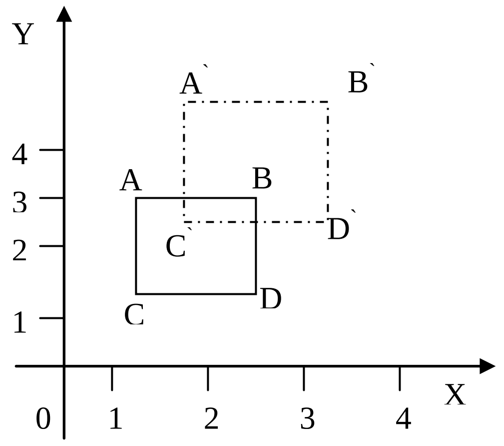

Fig. 3: A`B`C`D` is the translation of ABCD by (1,1).

Now, try the following Exercises.

E E3) Let $T: R^2 \rightarrow R^2$ be the reflection in the y-axis. Find an expression for T as in Example 4. Is T a linear operator?

E E4): For a fixed vector (a_1, a_2, a_3) in R_3, define the mapping $T:R^3 \rightarrow R$ by $T(x_1,x_2,x_3) = a_1x_1 + a_2x_2 + a_3x_3$. Show that T is a linear transformation. Note that $T(x_1,x_2,x_3)$ is the dot product of (x_1,x_2,x_3) and (a_1,a_2,a_3) (ref. sec. 2.4).

E E5) Show that the map $T:R^3 \rightarrow R^3$ defined by $T(x_1,x_2,x_3) = (x_1 + x_2 - x_3, 2x_1 - x_2, x_2 + 2x_3)$ is a linear operator.

You came across the real vector space P_n, of all polynomials of degree less than or equal to n, in Unit 4. The next exercise concerns it.

E E6) Let $f \in P_n$ be given by

$F(x) = \alpha_0 + \alpha_1 x + \ldots\ldots\ldots\ldots + \alpha_n x^n$, $\alpha_1 \in R \ \forall\, i$.

We define $(Df)(x) = \alpha + 2\alpha_2 x + \ldots\ldots\ldots + n\,\alpha_n x^{n-1}$.

Show that D:Pn is a linear transformation. (Observe that Df is nothing but the derivative of f. D is called the **differentiation operator.**)

In Unit 3 we introduced you to the concept of a quotient space. We now define a very useful linear transformation, using this concept.

Example 6: Let W be a subspace of a vector space U over a field F. W gives rise to the quotient space U/W. Consider the map $T:U \rightarrow U/W$ defined by $T(u) = u + W$. Show that T is a linear transformation.

Solution: for $\alpha\ \beta \in$ **F** and $u_1, u_2 \in U$ we have

$$T(\alpha u_1 + \beta u_2) = (\alpha u_1 + \beta u_2) + W = (\alpha u_1 + W) + (\beta u_2 + W).$$
$$= \alpha (u_1 + W) + \beta(u_2 + W)$$
$$= \alpha T(u_1) + \beta T(u_2)$$

Thus, T is a linear transformation.

Now solve the following exercise. Which is about plane vectors.

E E7) Let $u_1 = (1 - 1)$, $u_2 = (2, - 1)$, $u_3 = (4 - 3)$, $v_1 = (1,0)$ $v_2 = (0, 1)$ and $v_3 = (1, 1)$ be 6 vectors in R^2,. Can you define a linear transformation $T:R^2 \rightarrow R_2$ such that $(T(u_1) = V_1$, i = 1,2,3?

(Hint: Note that $2u_1 + u_2 = u_3$ and $v_1 + v_2 = v_3$).

You have already seen that a linear transformation $T:U \rightarrow V$ must satisfy $T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2)$, for $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$. More generally, we can show that,

LT6: $T(\alpha_1 u_1 + \ldots + \alpha_n u_n) = \alpha_1 T(u_1) + \ldots + \alpha_n T(u_n)$,

Where $\alpha_1 \in$ F and u_1, U.

Let us show this by induction, that is, we assume the above relation for n = m, and prove it for m + 1, Now.

$T(\alpha u_1 + \ldots + \alpha_m u_m + \alpha_{m+1} u_{m+1})$

= $T(u + \alpha_{m+1} u_{m+1})$, where $u = \alpha_1 u_1 + \ldots\ldots + \alpha_m u_m$
= $T(u) + \alpha_{m+1} T(u_{m+1})$, since the result holds for n = 2
= $T(\alpha_1 u_1 + \ldots\ldots + \alpha_m u_m) + \alpha_{m+1} T(u_{m+1})$
= $\alpha_1 T(u_1) + \ldots T\, \alpha_m T(u_m) + \alpha_{m+1} T(u_{m+1})$, since we have assumed the result for n = m.

Thus, the result is true for n = m+1. Hence, by induction, it holds true for all n.

Let us now come to a very important property of any linear transformation $T:U \rightarrow V$. In Unit 4 we mentioned that every vector space has a basis. Thus, U has a basis. We will now show that T is completely determined by its values on a basis of U. More precisely, we have

Theorem 1: Let S and T be two linear transformation from U to V, where $\dim_1 U = n$. Let $(e_1, \ldots\ldots\ldots e_n)$ be a basis of U. Suppose $S(e_1)$ for i = 1, ….., n. Then $S(u) = T(u)$ for all $u \in U$.

Proof: Let u $\in$ U. Since $(e_1, ….., e_n)$ is a basis of U, u can be uniquely written as u = $a_1e_1 + ….. + \alpha_n e_n$, where the α_1 are scalars.

Then, $S(u) = S(\alpha_1 e_1 + …….. + {}_n e_n)$

= $\alpha_1 S(e_1) + …... + \alpha_n S(e_n)$, by LT6
= $\alpha_1 T(e_1) + ……. + \alpha_n T(e_n)$
= $\alpha_1(a_1e_1 + …….. + \alpha_n e_n)$, by LT6
= T(u).

What we have just proved is that once we know the values of T on a basis of U, then we can find T(u) for any u $\in$ U.

Note: Theorem 1 is true even when U is not finite – dimensional. The proof, in this case, is on the same lines as above.

Let us see how the idea of Theorem 1 helps us to prove the following useful result.

Theorem 2: Let V be a real vector space and **T:R**$\rightarrow$ V be a linear transformation. Then there exists v $\in$ V such that $T(\alpha) = \alpha v \ \forall \alpha \in$ **R**.

Proof: A basis for R is (1). Let T(1) = V $\in$ V. then, for any $\alpha \in$ R, $T(\alpha) = \alpha T(1) = \alpha$ v.

Once you have read Sec. 5, 3 you will realize that this theorem says that T® is a vector space of dimension one, whose basis is [T(1)].

Now try the following exercise, for which you will need Theorem 1.

E E8) We define a linear operator $T:R^2 \rightarrow R^2$: T(1,0) = (0,1) and T(0,5) = 1,0). What is T(3,5)? What is T(5,3)?

Now we shall prove a very useful theorem about linear transformations, which is linked to Theorem 1

Theorem 3: Let) $e_1 …., e_n$) be a basis of U and let $v_1, …….v_n$ be any n vectors in V. Then there exists one and only one linear transformation $T:U \rightarrow V$ such that $T(e_1) = v_1$, i = 1, …….n.

Proof: Let u $\in$ U. Then u can be uniquely written as u $= a_1 e_1 + ………, T \alpha_n e_n$ (see Unit 4, Theorem 9).

Define T(u) = $\alpha_1 v_1 + ….+ \alpha_n v_n$. The T defines a mapping from U to V such that $T(e_1) = v_1 \forall$ i = 1, …….n,. Let us now show that T is linear. Let a, b be scalars and u, u` $\in$ U. The $\exists$ scalar $\alpha_1, …..,\alpha_n, \beta_1,………, \beta_n$ such that u = $\alpha_1 e_1 + ……, + \alpha_n e_n$ and u` = $\beta_1 e_1 + …… + \beta_n e_n$.

Then au + bu` = $(a\alpha_1 + b\beta_1) e_1 + ….., + (a \alpha_{\mathbf{n}} + b\beta_{\mathbf{n}}) e_{\mathbf{n}}$.

Hence, T(au + bu`) = $(a\alpha_1 + b\beta_1) v_1 + …., + (a\alpha_n + b\beta_{\mathbf{n}}) v_{\mathbf{n}} = a(\alpha_1 v_1 + ….. + \alpha_{\mathbf{n}} v_{\mathbf{n}}) + b(\beta_1 v_1 + ….+ \beta_{\mathbf{n}} v_{\mathbf{n}})$ = aT(u) + bT(u`)

Therefore, T is a linear transformation with the property that $T(e_1) = v_1 \forall$ i. Theorem 1 now implies that T is the only linear transformation with the above properties.

Let's see how Theorem 3 can be used.

Example 7: e_1 = (1,0,0). e_2 = (0, 1, 0) and e_3 = 0, 0, 1) form the standard basis of R^3. Let (1,2), (2,3) and (3,4) be three vectors in R^2. Obtain the linear transformation T: $R^3 \rightarrow R^2$ such that T(e1) = (1,2), $T(e_2)$ = (2,3) and $T(e_3)$ = (3,4).

Solution: By Theorem 3 we know that $\exists$ T:$R^3 \rightarrow R^2$ such that $T(e_1)$ = (1,2), $T(e_2)$ = (2,3), and $T(e_3)$ = (3,4). We want to know what T(x) is, for any x = $(x_1, x_2, x_3) \in R^3$, Now, X = $x_1 e_1 + x_2 e_2 + x_3 e_3$.

Hence, $T(x) = x_1 T(e_1) + x_2 T(e_2) + x_3 T(e_3)$

$= \quad x_1 (1,2) + x_2 (2,3) + x_3 (3,4)$

$= \quad (x_1 + 2x_2 + 3x_3, \; 2x_1 + 3x_2 + 4x_3)$

Therefore, $T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 2x_1 + 3x_2 + 4x_3)$ is the definition of the linear transformation T.

E E9) Consider the complex field **C**. It is a vector space over **R**

a) What is its dimension over **R**? Give a basis of **C** over **R**.

b) Let $\alpha, \beta \in$ R. Give the linear transformation which maps the basis elements of **C**. obtained in (a), onto α and β, respectively.

Let us now look at some vector spaces that are related to a linear operator.

3.2 Spaces Associated with a Linear Transformation

In Unit 1 you found that given any function, there is a set associated with it, namely, its range. We will now consider two sets which are associated with any linear transformation, T. These are the range and the kernel of T.

3.3 The Range Space and the Kernel

Let U and V be vector spaces over a field F. Let T:U → V be a linear transformation. We will define the range of T as well as the Kernel of T. At first, you will see them as sets. We will prove that these sets are also vector spaces over F.

Definition: The range of T, denoted by R(T), is the set {T(x)] x ∈ U}/ The **kernel** (or null space) of **T**, denoted by **Ker T,** is the set {x ∈ U] T(x) = 0 } Note that R(T) ⊆ V and Ker T ⊆ U.

To clarify these concepts consider the following examples.

Example 8: Let I: V → V be the identity transformation (see Example 1). Find R(I) and Ker I.

Solution: R(I) = {I(v) | v ∈ V} = {v| v ∈ V} = V. Also, Ker I = {v ∈ V| I(v) = 0} = {v ∈ V| v = 0} = {0}.

Example 9: Let T: R^3→ R be defined by $T(x_1, x_2, x_3) = 3x_1 + x_2 + 2x_3$ Find R(T) and Ker T.

Solution: R(T) = {x ∈ R|∃ x_1, x_2, x_3, ∈ R with $3x_1 + x_2 + 2x_3$. = x}. For example, 0 ∈ R(T). Since 0 = 3.0 + 0 + 2.0 = T(0, 0, 0,)

Also, I E R (T), since I = 3.1/3 + 0 + 2.0 = T(1/3, 0, 0), or I = 3.0 + 1 + 2.0 = T(0, 1, 0), or I = T(0, 0, ½) or I = T (1/6, ½, 0).

Now can you see that R(T) is the whole real line R? This is because, for any α ∈ R, α = α 1 = α T(1/3,0,0) = T($^{\alpha}/_3$, 0, 0) ∈ R (T).

Ker T = { $(x_1, x_2, x_3) \in R^3$| $3x_1 + x_2 + 2x_3 = 0$}.

For example, (0,0,0) ∈ Ker T. But (1, 0,0) ∉ Ker T . :. Ker T ≠R3. In fact, Ker T is the plane $3x_1 + x_2 +2x_3 = 0$ in R^3.

Example 10: Let T:R^3→ R^3 be defined by $T(x_1,x_2, x_3) = (x_1 – x_2 + 2x_3, 2x_1 + x_2, - x_1 – 2x_2 + 2x_3)$. Find R(T) and Ker T.

Solution: To find R(T), we must find conditions on y_1, y_2, $y_3 \in$ R so that (y_1, y_2, y_3) $\in$R(T), i.e. ., we must find some $(x_1, x_2, x_3) \in R^3$ so that $(y_1, y_2, y_3) = T(x_1, x_2, x_3) = (x_1 – x_2 + 2x_3, 2x_1 + x_1 – 2x_2 + 2x_3)$.

This means
$x_1 – x_2 + 2x_3 = y_1$ ………………. (1)
$2x + x_2 = y_2$ …………………..(2)
$-x_1 = 2x_2 + 2x_3 = y_3$ ……………. (3)

Subtracting 2 times Equation (1) from Equation (2) and adding Equations (1) and (3) we get.

$3x_2 – 4x_3 = y_2 + - 2y_1$ ………………… (4)
and $- 3x_2 + 4x_3 = y_1 + y_3$ ……………… (5)

Adding Equations (4) and (5) we get

$y_2 – 2y_1 + y_1 + y_3 = 0$, that is , $y_2 + y_3 = y_1$,

Thus, $(y_1, y_2, y_3 \in R (T) \Rightarrow y_2 + y_3 = y_1$.

On the other hand, if y2 + y3 = y1. We can choose

$$X_3 = 0,\ x_2 = \frac{y_2 - 2y_1}{3} \text{ and } x_1 = y_1 + \frac{y_2 – 2y_1}{3} = \frac{y_1 + y_2}{3}$$

Then, we see that $T (x_1, x_2, x_3) = (y_1, y_2 , y_3)$.
Thus, $y_2 + y_3 = y_1 \Rightarrow (y_1, y_2, y_3) \in R(T)$.
Hence, $R(T) = \{(y_1,y_2,y_3) \in R^3 | y_2 + y_3 = y_1\}$

Now $(x_1,x_2,x_3) \in$ Ker T if and only if the following equations are true:
$x_1 – x_2 + 2x_3 = 0$
$2x_1 + x_2 = 0$
$-x_2 – 2x_2 + 2x_3 = 0$

Of course $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ is a solution. Are there other solutions? To answer this we proceed as in the first part of this example. We see that $3x_2 = 4x_3 = 0$. Hence, $x_3 = (3/4)\ x_2$.

Also, $2x_1 + x_2 = 0 \Rightarrow x_1 = - x^2/_2$.

Thus, we can give arbitrary values to x_2 and calculate x_1 and x_3 in terms of x_2. Therefore, Ker T = $\{(- {}^{\alpha}/_2, \alpha, (3/4)\alpha): \alpha \in R\}$.

In this example, we se that finding R(T) and Ker T amounts to solving a system of equations. In Unit 9, you will learn a systematic way of solving a system of linear equations by the use of matrices and determinants.

The following exercises will help you in getting used to R(T) and Ker T.

E E10) Let T be the zero transformation given in Example 2. Find Ker T and R(T). Does $I \in R(T)$?

E E11) Find R(T) and Ker T for each of the following operators.

a) $T: R^3 \rightarrow R^2 : T(x, y, z) = (x, y)$

b) $T: R^3 \rightarrow R : T(x, y, z) = z$

c) $T: R^3 \rightarrow R^3 : T(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_1 + x_2 + x_3, x_1 + x_2 + x_3)$.

(Note that the operators in (a) and (b) are projections onto the xy-plane and the z-axis, respectively).

Now that you are familiar with the sets R9T) and Ker T, we will prove that they are vector spaces.

Theorem 4: Let U and V be vector spaces over a field F. Let $T: U \rightarrow V$ be a linear transformation. Then Ker T is a subspace of U and R(T) is a subspace of V.

Proof: Let $x_1, x_2, \in$ Ker $T \subseteq U$ and $\alpha_1, \alpha_2 \in \mathbf{F}$. Now, by definition, $T(x_1) = T(x_2) = 0$.

Therefore, $\alpha_1 T(x_1) + \alpha_2 T(x_2) = 0$

But $\alpha_1 T(x_1) + \alpha_2 T(x_2) = T(\alpha_1 x_1 + \alpha_2 x_2)$.

Hence, T $(\alpha_1 x_1 + \alpha_2 x_2) = 0$
This means that $\alpha_1 x_1 + \alpha_2 x_2 \in$ Ker T.

Thus, by Theorem 4 of Unit 3, Ker T is a subspace of U.

Let $y_1, y_2 \in R(T) \subseteq V$, and $\alpha_1, \alpha_2 \in F$. then, by definition of R(T), there exist $x_1, x_2 \in U$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$.

So, $\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 T(x_1) + \alpha_2 T(x_2)$

$= \quad T(\alpha_1 x_1 + \alpha_2 x_2)$.

Therefore, $\alpha_1 y_1 + \alpha_2 y_2 \in$ R (T), which proves that R(T) is a subspace of V.

Now that we have proved that R(T) and Ker T are vector spaces, you know, from Unit 4, that they must have a dimension. We will study these dimensions now.

3.4 Rank and Nullity

Consider any linear transformation T:U → V, assuming that dim U is finite. Then Ker T, being a subspace of U, has finite dimension and dim (Ker T) $\leq$ dim U. Also note that R(T) = T(U), the image of U under T, a fact you will need to use in solving the following exercise.

E E12) Let $\{e_1 \ldots\ldots e_n\}$ be a basis of U Show that R(T) is generated by $\{T(e_1), \ldots\ldots\ldots\ldots, T(en)\}$.

From E12 it is clear that, if dim U = N, then dim R(T) $\leq$ n. Thus, dim R(T) is finite, and the following definition is meaningful.

Definition: The **rank** of T is defined to be the dimension of R(T), the range space of T. The **nullity** of T is defined to be the dimension of Ker T, the kernel (or the null space) of T.

Thus, **rank (T)** = dim R(T) and nullity (T) = dim Ker T.

We have already seen that rank (T) $\leq$ dim U and nullity (T) $\leq$ dim U.

Example 11: Let T:U → V be the zero transformation given in example 2. What are the rank and nullity of T?

Solution: In E10 you saw that R(T) = (0) and Ker T = U, Therefore, rank (T) = 0 and nullity (T) = dim U.

Note that rank (T) + nullity (T) = dim U, in this case.

E E13) If T is the identity operator on V, find rank (T) and nullity (T).

E E14) Let D be the differentiation operator in E6. Give a basis for the range space of D and for Ker D. What are rank (D) and nullity (D)?

In the above example and exercises you will find that for T:U → V, rank (T) + nullity (T) = dim U. In fact, this is the most important result about rank and nullity of a linear operator. We will now state and prove this result.

Theorem 5: Let U and V be vector spaces over a field F and dim U = n. Let T:U → V be a linear operator. Then rank (T) + nullity (T) = n.

Proof: Let nullity (T) = m, that is, dim Ker T = m, Let (e_1 ….., e_m) be a basis of Ker T. We know that Ker T is a subspace of U. thus, by theorem 11 of Unit 4, we can extend this basis to obtain a basis (e_1, ….., e_m, e_{m+1}, ……, e_n) of U. We shall show that {T(e_{m+1}), ………, T(e_n)} is a basis of R(T). Then, our result will follow because dim R(T) will be n – m = n – nullity (T).

Let us first prove that {T(e_{m+1}), ……, T(e_n)} spans, or generates, R(T). Let y ∈ R (T). Then, by definition of R(T), there exists x ∈ U such that T(x) = y.

Let $x = c_1e_1 + \ldots + c_me_m + c_{m+1}e_{m+1} + \ldots + c_ne_n$, $c_1 \in F\ \forall\ i$.
Then

$y = T(x) = c_1T(e_1) + \ldots. + c_mT(e_m) + c_{m+1}) + \ldots..+ c_nT(e_n)$

$= \quad c_{m+1} T(e_{m+1}) + \ldots.. + c_nT(e_n)$,

Because $T(e_1) = .. = T(e_m) = 0$, since $T(e_1) \in$ Ker T $\forall$ i = 1, …., m. :. any $y \in R(T)$ is a linear combination of $\{T(e_{m+1}), \ldots., T(e_n)\}$. Hence, R(T) is spanned by $\{T(e_{m+1}), \ldots., T(e_n)\}$. It remains to show that the set $\{T(e_{m+1}), \ldots., T(e_n)\}$ is linearly independent. For this, suppose there exist $a_{m+1}, \ldots., a_n \in F$ with $a_{m+1}T(e_{m+1}) + \ldots + a_n T(e_n) = 0$.

Then, $T(a_{m+1}e_{m+1} + \ldots + a_ne_n) = 0$
Hence, $a_{m+1} e_{m+1} + \ldots + a_ne_n \in$ Ker T, which is generated by $\{e_1, \ldots\ldots, e_m\}$.

Therefore, there exist $a_1, \ldots\ldots, a_m \in F$ such that $a_{m+1} e_{m+1} + \ldots. + a_ne_n = a_1e_1 + \ldots + a_m e_m \Rightarrow (- a_1) e_1 + \ldots. + (- a_m) e_m + a_{m+1} e_{m+1} + \ldots + a_n e_n = 0$.

Since $\{e_1,\ldots\ldots, e_n\}$ is a basis of U, it follows that this set is linearly independent. Hence, $- a_1 = 0$, ---, $-a_m = 0$, $a_{m+1} = 0$, …., $a_n = 0$. In particular, $a_{m+1} = \ldots. = a_n = 0$, which we wanted to prove.

Therefore, dim R(T) = n – m = n – nullity (T), that is, rank (T + nullity (T) = n. let us see how this theorem can be useful.

Example 12: Let $L:R^3 \rightarrow R$ be the map given by $L(x,y,z) = x + y + z$. What is nullity (L)?

Solution: In this case it is easier to obtain R(L), rather than Ker L. Since L (1,0,0) = 1 $\neq$ 0, R(L) $\neq$ {0}, and hence dim R(L) $\neq$ =. Also, R(L) is a subspace of R. Thus, dim R(L) $\leq$ dim R = 1. therefore, the only possibility for dim R(L) is dim R(L) = 1,. By Theorem 5, dim Ker L + dim R(L) = 3.

Hence, dim Ker L = 3 – 1 = 2. That is, nullity (L) = 2.

E E15) Give the rank and nullity of each of the linear transformations in E11.

E E16) Let U and V be real vector spaces and T:U → V be a linear transformation, where dim U = 1. Show that R(T) is either a point or a line.

Before ending this section we will prove a result that links the rank (or nullity) of the composite of two linear operators with the rank (or nullity) of each of them.

Theorem 6: Let V be a vector space over a field F. Let S and T be linear operators from V to V. Then

a) rank (ST) min (rank (S), rank (T))
b) nullity (ST) $\geq$ max (nullity (S), nullity (T))

Proof: We shall prove (a). Note that (ST) (v) = S(T(v)) for any v $\in$ V

Now, for any y $\in$ R (ST), $\exists$ v $\in$ V such that,

Y = (ST) (v) = S (T(v)) …………………….1

Now, (1) $\Rightarrow$ y $\in$ R(S).

Therefore, R(ST) $\subseteq$ R(S). This implies that rank (ST) $\leq$ rank (S).

Again, (1) $\Rightarrow$ y $\in$ S (R(T)), since T(v) $\in$ R(T).

:. R (ST) $\subseteq$ S (R(T)), so that dim R(ST) $\leq$ dim S(R(T)) $\leq$ dim R(T) (since dim L(U) $\leq$ U, for any linear operator (0).

Therefore, rank (ST) $\leq$ rank (T).
Thus, rank (ST) $\leq$ min (rank (S), rank (R)).
The proof of this theorem will be complete, once you solve the following exercise.

E E17) Prove (b) of Theorem 6 using the Rank Nullity Theorem.

We would now like to discuss some linear operators that have special properties.

3.5 Some types of Linear Transformations

Let us recall, from Unit 1, that there can be different types of functions, some of which are one-one, onto or invertible. We can also define such types of linear transformations as follows:

Definition: Let $T: U \rightarrow V$ be a linear transformation.

a) T is called **one-one** (or injective) if, for $u_1, u_2 \in U$ with $u_1, \notin u_2$, we have $T(u_1) \notin T(u_2)$. If T is injective, we also say T is 1 – 1.
Note that T is **1 – 1 if T** $(\mathbf{u_1})$ $= T(u_2) \Rightarrow u_1 = u_2$.

b) T is called onto (or **surjective)** if, for each $v \in V$, $\exists u \in U$ such that $T(u) = v$, that is $R(T) = V$.

Can you think of examples of such functions? The identity operator is both one-one and onto. Why is this so? Well, $I:V \rightarrow V$ is an operator such that, if $v_1, v_2 \in V$ with $v_1 \neq v_2$ then $I(v_1) \neq I(v_2)$ Also, $R(I) = V$, so that I is onto.

E E18) Show that the zero operator $0: R \rightarrow R$ is not one – one.

Theorem7: $T:U \rightarrow V$ is one-one if and only if Ker T = (0).

Proof: First assume T is one – one. Let $u \in$ Ker T. Then $T(u) = 0 = T(0)$. This means that $u = 0$. thus, Ker T = (0). Conversely, let Ker T = (0). Suppose $u_1, u_2 \in U$ with $T(u_1) = T(u_2) \Rightarrow T(u_2 - u_2) = 0 \Rightarrow u_1 - u_2 \in \text{Ker} \Rightarrow u_1 - u_2 = 0 \Rightarrow u_1 = u_2$. therefore T is 1 - 1

Suppose now that T is a one – one and onto linear transformation from a vector space U to a vector space V. Then, from Unit 1 (Theorem 4), we know that T^{-1} exists. But is T^{-1} linear? The answer to this question is 'yes', as is shown in the following theorem.

Theorem 8: Let U and V be vector spaces ovet a field F. Let $T:U \rightarrow V$ be a none-one and onto linear transformation. Then $T^{-1}: V \rightarrow U$ is a linear transformation.

In fact, T^{-1} is also 1 – 1 and onto.

Proof: Let $y_1, y_2 \in V$ and $\alpha_1, \alpha_2 \in F$. Suppose $T^{-1}(y_1) = x_1$ and $T^{-1}(y_2) = x_2$. then, by definition, $y_1 = T(x_1)$ and $y_2 = T(x_2)$.

Now, $\alpha_1 y_1 + \alpha_1 y_2 = \alpha_1 T(x_1) + \alpha_2 T(x_2) = T(\alpha_1 x_1 + \alpha_1 x_2)$

Hence, $T^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1 + \alpha_2 x_2$

$= \alpha_1 T^{-1}(y_1) + \alpha_2 T^{-1}(y_2)$ **t-1 (y) = x ⇔ T(x) = y**

This shows that T^{-1} is a linear transformation.

We will now show that T^{-1} is 1 -1, for this, suppose $y_1, y_2 \in V$ such that $T^{-1}(y_1) = T^{-1}(y_2)$ Let $x_1 = T^{-1}(y_1)$ and $x_2 = T^{-1}(y_2)$.

Then $T(x_1) = y_1$ and $T(x_2) = y_2$. We know that $x_1 = x_2$. Therefore, $T(x_1) = T(x_2)$, that is, $y_1 = y_2$. thus, we have shown that $T^{-1}(y_1) = T^{-1}(y_2) \Rightarrow y_1 = y_2$, proving that T^{-1} is 1 – 1. T-1 is also surjective because, for any $u \in U$, $\exists\, T(u) = v \in V$ such that $T^{-1}(v) = u$.

Theorem 8 says that a one-one and onto linear transformation is **invertible**, and the inverse is also a one-one and onto linear transformation.

This theorem immediately leads us to the following definition.

Definition: Let U and V be vector spaces over a field F, and let $T:U \rightarrow V$ be a one-one and onto linear transformation. The T is called an **Isomorphism** between U and V. In this case we say that U and V are **isomorphic vector spaces.** This is denoted by $U \approx V$.

An obvious example f an isomorphism is the identity operator. Can you think of any other? The following exercise may help.

E E19) Let $T:R^3 \to R^3$: $T(x, y, z) = (x + y, y, z)$. Is T an isomorphism? Why? Define T^{-1}, if it exist

E E20) Let $T : R^3 \rightarrow R^2$: $T(x, y, z) = (x + y, y + z)$. Is T an isomorphism?

In all these exercises and examples, have you noticed that if T is an isomorphism between U and V then T^{-1} is an isomorphism between V and U?

Using these properties of an isomorphism we can get some useful results, like the following:

Theorem 9: Let $T: U \rightarrow V$ be an isomorphism. Suppose $\{e_1, \ldots, e_n)$ is a basis of U. then $\{T(e_1), .., T(e_n)\}$ is a basis of V.

Proof: First we show that the set $\{T(e_1), \ldots, T(e_n)\}$ spans V. Since T is onto, R(T) = V. thus, from E12 you know that $\{T(e_1),.., T(e_n)\}$ spans V.

Let us now show that $\{T(e_1),.., T(e_n)\}$ is linearly independent. Suppose there exist scalars $c_1,\ldots, c_n$, such that $c_1 T(e_1) + \ldots + c_n T(e_n) = 0$ ………1

We must show that $c_1 = \ldots = c_n = 0$

Now, (1) implies that
$T(c_1 e_1 + \ldots + c_n e_n) = 0$
Since T is one-one and $T(0) = 0$, we conclude that $c_1e_1 + \ldots + c_n e_n = 0$.
But $\{e_1, \ldots, e_n\}$ is linearly independent. Therefore, $c_1 = \ldots = c_n = 0$.

Thus, we have shown that {T (e1),…., T(en)} is a basis of V.

Remark: The argument showing the linear independence of $\{T(e_1), \ldots, T(e_n)\}$ in the above theorem can be used to prove that any one-one linear transformation $T:U \rightarrow V$ maps any linearly independent subset of U onto a linearly independent subset of V (see E22).

We now give an important result equating 'isomorphism' with '1 -1 ' and with 'onto' in the finite-dimensional case.

Theorem 10: Let $T: U \rightarrow V$ be a linear transformation where U. V are of the same finite dimension. Then the following statements are equivalent.

a) T is 1 – 1

b) T is onto.
c) T is an isomorphism.

Proof: To prove the result we will prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a). Let dim U = dim V = n.

Now (a) implies that Ker T = (0+ (from Theorem 7), Hence, nullity (T) = (0). Therefore by Theorem 5, rank (T) = n that is dim R(T) = n = dim V. But R (T) is a subspace of V. thus, by the remark following Theorem 12 of Unit 4, we get R (T) = V, i.e., T is onto, i.e., (b) is true. So (a) $\Rightarrow$ (b).

Similarly, if (b) holds then rank (T) = n, and hence, nullity (T) = 0. consequently, Ker T = {0}, and T is one-one. Hence, T is one-one and onto, i.e. . t is an isomorphism. Therefore, (b) implies (c).

That (a) follows from 9c) is immediate from the definition of an isomorphism.

Hence, our result is proved.

Caution: Theorem 10 is true for **finite-dimensional spaces** U and V, of the same **dimension.** It is not true, otherwise. Consider the following counter-example.

Example 13: (To show that the spaces have to be finite-dimensional): Let V be the real vector space of all polynomials. Let D:V $\rightarrow$ V be defined by D $(a_0 + a_1 x + .. + a_r x^r) = a_1 + 2a_2x + \ldots + ra_r x^{r-1}$. then show that D is onto but not 1 -1.

Solution: Note that V has infinite dimension, a basis being $\{1, x, x^2, \ldots\}$. D is onto because any element of V is of the form $a_0 + a_1 x + \ldots + a_nx_n$ = D

$$a_0 x + \left(\frac{a_1}{2} x_2 + \ldots + \frac{a_n}{n+1} x^{n+1} \right)$$

D is not 1 – 1 because, for example, $1 \neq 0$ but D (1) = D (0) = 0.

The following exercise shows that the statement of Theorem 10 is false if dim U $\neq$ dim V.

E E12) Define a linear operator T: $R^3 \rightarrow R^2$ such that T is onto but T is not 1 - 1. Note that dim $R^3 \neq$ dim R^2.

Let us use Theorems 9 and 10 to prove our next result.

Theorem 11: Let T:V → V be a linear transformation and let $\{e_1, \ldots, e_n\}$ be a basis of V. Then T is one-one and onto if and only if $\{T(e_1), \ldots, T(e_n)\}$ is linearly independent.

Proof: Suppose T is one-one and onto. The t is an isomorphism. Hence, by Theorem 9, $\{T(e_1), \ldots., T(e_n)\}$ is a basis. Therefore, $\{T(e_1), \ldots, T(e_n)\}$ is linearly independent.

Conversely, suppose $\{T(e_1), \ldots., T(e_n)\}$ is linearly independent. Since $(e_1, \ldots, e_n\}$ is a basis of V, dim V = n. therefore, any linearly independent subset of an vectors is a basis of V (by Unit 4, Theorem 5, Cor. 1). Hence, $\{T(e_1), \ldots, T(e_n)\}$ is a basis of V. Then, any element v of V is of the form

$$v = \sum_{i=1}^{n} c_i \; T(e_1) = T\left(\sum_{i=1}^{n} c_i e_i\right),$$ where $c_1, \ldots., c_n$ are scalars. Thus, T is onto, and we can use Theorem 10 to say that T is an isomorphism.

Here are some exercise now.

E E22) a) Let T : U → V be a one-one linear transformation and let $\{u_1, \ldots, U_k\}$ be a linearly independent subset of U. show that the set $\{T(u_1), \ldots., T(u_k)\}$ is linearly independent.

b) Is it true that every linear transformation maps every linearly independent set of vectors into a linearly independent set?

d) Show that every linear transformation maps a linearly dependent set of vectors onto a linearly dependent set of vectors.

E E23) Let $T: R^3 \rightarrow R^3$ be defined by $T(x_1, x_2, x_3) = (x_1 + x_3, x_2 + x_3, x_1 + x_2)$. Is T invertible? If yes, find a rule for T^{-1} like the one which defines T.

We have seen, in Theorem 9, that if $T: U \rightarrow V$ is an isomorphism, then T maps a basis of U onto a basis of V. Therefore, dim U = dim V. In other words, if U and V are isomorphic then dim U = dim V. the natural question arises whether the converse is also true. That is, if dim U = dim V, both being finite, can we say that U and V are isomorphic? The following theorem shows that this is indeed the case.

Theorem 12: Let U and V be finite-dimensional vector spaces over F. the U and V are isomorphic if and only if dim U = dim V.

Proof: We have already seen that if U and V are isomorphic then dim U= dim V. Conversely, suppose dim U = dim V = n. We shall show that U and V are isomorphic. Let $\{e_1,\ldots, e_n\}$ be a basis of U and $\{f_1,\ldots, f_n\}$ be a basis of V. By Theorem 3, there exists a linear transformation $T:U \rightarrow V$ such that $T(e_1) = f_1$, i = 1, ….., n.

We shall show that T is 1 -1 .
Let $u = c_1e_1 + \ldots + c_n e_n$ be such that $T(u) = 0$
Then $0 = T(u) = c_1 T(e_1) + \ldots + c_n T(e_n)$
$= c_1f_1 + \ldots + c_nf_n$.

Since $\{f_1,\ldots, f_n\}$ is a basis of V. we conclude that $c_1 = c_2 = \ldots = c_n = 0$. Hence, u = 0. Thus, Ker T = (0) and, by Theorem 7, T is one – one.

Therefore, by Theorem 10, T is an isomorphism, and U = V. An immediate consequence of this theorem follows:

Corollary: Let V be a real (or complex) vector space of dimension n. Then V is isomorphic to R" (or C"). respectively.

Proof: Since dim R_n = n = dim $_R$V, we get $V \approx R^n$. Similarly, if dim $_c$V = n, then $V \approx c_n$.

Remark: Let V be a vector space over F and let $B = \{e_1,\ldots, e_n\}$ be a basis
n

of V. Each v € V can be uniquely expressed as $v = \sum_{i=1} \alpha_1 e_1$. Recall that $\alpha_1, \ldots, \alpha_n$ are called the coordinates of v with respect to B (refer to sec. 4.4.1).

Define $\theta : V \rightarrow F^n : \theta(v) = (\alpha_1, \ldots, \alpha_n)$. Then θ is an isomorphism from V to F^n. This is because θ is 1 -1, since the coordinates of v with respect to B are uniquely determined **Thus, V** $\approx F^n$.

We end this section with an exercise.

E E24) Let T: U → V be a one- one linear mapping. Show that T is onto if and only if dim U = dim V. (of course, you must assume that U and V are finite dimensional spaces).

Now let us look at isomorphism between quotient spaces.

3.6 Homomorphism Theorems

Linear transformation are also called **vector space homomorphisms.** There is a basic theorem which uses the properties of homomorphisms to show the isomorphism of certain quotient spaces (ref. Unit 3). It is simple to prove. But is very important because it is always being used to prove more advanced theorems on vector spaces. (in the Abstract Algebra course we will prove this theorem in the setting of groups and rings)

Theorem 13: Let V and W be vector spaces over a firld F and T:V → W be a linear transformation. Then **V/Ker T ≈ R (T)**.

Proof: You know that Ker T is a subspace of V, so that V/Ker T is a well defined vector space over F. Also $R(T) = \{T(v) \mid V \in V\}$. To proof the theorem let us define θ V/Ker T → R (T) by $\theta(v + \text{Ker } T) = T(v)$.

Firstly, we must show that θ is a well defined function, that is, if v +Ker T = v\` + Ker T then $\theta(v + \text{Ker } T) = \theta(v` + \text{Ker } T)$, i.e. T(v) = T (v\`).

Now, v + Ker T = v\` + Ker T$\Rightarrow$ (v –v\`) $\in$ Ker T (see Unit 3, E23)

$\Rightarrow$ T (v-v\`) = 0 $\Rightarrow$ T(v) = T(v\`). and hence, θ is well defined.

Next, we check that θ is a linear transformation. For this, let a , b $\in$ F and v. v` $\in$V. then $\theta$\{a (v + Ker T) + b(v` + Ker T)\}
= θ (av + bv` + Ker T) (ref. Unit 3)
= T (av + bv`)
= aT(v) + bT (v`), since T is linear.
= a θ (v + Ker T) + b θ (v` + Ker T).
Thus, θ is a linear transformation.

We end the proof by showing that θ is an isomorphism. θ is 1 -1 (because θ (v + Ker T) = $\theta \Rightarrow$ T(v) = 0 $\Rightarrow$ v $\in$ Ker T $\Rightarrow$ v + Ker T = 0 (in V/Ker T).

Thus, Ker θ =\{0\})

θ is onto (because any element of R (T) is T (v) = θ (v) = θ (v + Ker T)).
So have prove that θ is an isomorphism. This proves that V/Ker T = R (T).

Let us consider an immediate useful application of Theorem 13.

Example 14: Let V be a finite-dimensional space and let S and t be linear transformations from V to V. show that. Rank (ST) = rank (T) – dim (R (T) $\cap$ Ker S).

Solution: We have $V \xrightarrow{T} V \xrightarrow{S} V$. ST is the composition of the operators S and T. which you have studied in Unit 1, and will also study in Unit 6. Now, we apply Theorem 13 to the homomorphism θ : T(v) → ST(V): θ (T(v)) = (ST) (v).

Now, Ker θ = $\{ x \in T(V) \mid S(x) = 0 \}$ = Ker S $\cap$ T (V) = Ker S $\cap$ R (T). Also R(θ) = ST (V), since any element of ST (V) is (ST) (v) = θ (T(v)). Thus.

$$\frac{T(V)}{\text{Ker } S \cap T(V)} \approx ST(V)$$

Therefore,

$$\dim \frac{T(V)}{\text{Ker } S \cap T(V)} = \dim ST(V)$$

That is, dim T(V) – dim (Ker S $\cap$ T (V)) = dim ST (V), which is what we had to show.

E E25) Using Example 14 and the Rank Nullity Theorem, show that nullity (ST) = nullity (T) + dim (R(T) $\cap$ Ker S).

Now let us see another application of Theorem 13.

Example 15: Show that $R^3/R \approx R^2$.

Solution: Note that we can consider R as a subspace of R^3 for the following reason: any element a of R is equated with the element (α 0, 0) of R^3. Now, we define a function f: $R^3 \rightarrow R^2$: $f(\alpha,\beta,\gamma) = (\beta,\gamma)$. then f is a linear transformation and Ker f = $\{(\alpha, 0, 0) \mid \alpha \in R\} \approx R$. Also f is onto, since any element (α, β) of R^2 is $f(0, \alpha, \beta)$. Thus, by Theorem 13, $R^3/R \approx R^2$.

Note: In general, for any $n \geq m$, $R^n/R^m \approx R^{n-m}$.. Similarly, $C^{n-m} \approx C^n/C^m$ for $n \geq m$. The next result is a corollary to the Fundamental Theorem of Homomorphism. But, before studying it, read unit 3 for definition of the sum of spaces.

Corollary 1: Let A and Be be subspaces of a vector space V. then $A + B/B \approx A/A \cap B$.

Proof: we define a linear function $T: A \rightarrow \frac{A+B}{B}$ by $T(a) = a + B$

T is well defined because a + B is an element of $\frac{A+B}{B}$ (since $a = a + 0 \in A+B$).

T is a linear transformation because, for α_1, α_2 in F and α_1, α_2 in A, we have
$T(\alpha_1 a_1 + \alpha_2 a_2) = \alpha_1 a_1 + \alpha_2 a_2 + B = \alpha_1 (a_1 + B) + \alpha_2 (a_2 + B)$
$= \alpha_1 T(a_1) + \alpha_2 T(a_2)$

Now we will show that T is surjective. Any element of $\frac{A+B}{B}$ is of the form a + b + B, where $a \in A$ and $b \in B$.

Now $a + b + B = a + B + b + B = a + B + B$, since $b \in B$.

A + B

= a + B, since B is the zero element of $\dfrac{\quad}{B}$

= T (a), proving that T is surjective.

$\therefore R(T) = \dfrac{A + B}{B}$

We will now prove that Ker T = A ∩ B.

If a ∈ Ker T, then a ∈ A and T(a) = 0. This means that a + B = B, the zero element of $\dfrac{A + B}{B}$ Hence, a ∈ B (by Unit 3, E23). Therefore, a ∈A ∩ B.

Thus, Ker T ⊆ A ∩ B. On the other hand, a ∈ A ∩ B⇒ a ∈ A and a ∈ B⇒ a ∈ A and a + B = B ⇒ a ∈ A and T(a) = T(0) = 0

⇒ a ∈ Ker T.

This proves that A ∩ B = Ker T.

Now using Theorem 13, we get

A/Ker T ≈ R (T)

That is, A/(A ∩ B) ≈ (A + B)/B

E E26) Using the corollary above, show that A ⊕ B/B ≈ A (⊕) denotes the direct sum of defined in sec. 3.6).

There is yet another interesting corollary to the Fundamental Theorem of Homorphism.

Corollary 2: Let W be a subspace of a vector space V. Then, for any subspace U of V containing W.

$$\frac{V/W}{U/W} \approx V/U$$

Proof: This time we shall prove the theorem with you. To start with let use define a function T: V/W → V/U : t (v +) = v + U. Now try E 27

E E27 a) Check that T is well defined.
b) Prove that T is a linear transformation.
c) What are the spaces Ker T and R(T)?

So, is the theorem proved? Yes; apply theorem 13 to T. we end the unit by summarizing what we have done in it.

4.0 CONCLUSION

In this unit we have covered the following points.

(1) A linear transformation from a vector space U over F to a vector space V over F is a function T: U → V such that,

LT1) $T(u_1 + u_2) = T(u_1) + T(u_2) \forall u_1, u_2 \in U$, and

LT 2) $T(\alpha u) = \alpha T(u)$, for $\alpha \in F$ and $u \in U$.

These conditions are equivalent to the single condition LT3) $T(\alpha u_1 + \beta u_2) = \alpha T(u_1) + \beta T(u_2)$ for $\alpha, \beta \in F$ and $u_1, u_2 \in U$.

(2) Given a linear transformation T: U → V.

i) The Kernel of T is the vector space $\{u \in U \mid T(u) = 0\}$, denoted by Ker T.
ii) The range of T is the vector space $\{T(u) \mid u \in U\}$, denoted by r (T),
iii) The rank of T = dim $_1$R(T)
iv) The nullity of T = dim $_1$Ker T.

(3) Let U and V be finite-dimensional vector spaces over F and T: U → V be a linear transformation. Then rank (T) + nullity (T) = dim U.

(4) Let T: U → V be a linear transformation then
T is one-one if T (u_1) = T (u_2) ⇒ u_1 = u2 ~~∀~~ u_1, u_2 ∈ U
(i) T is onto if, for any v ∈ V ∃ u ∈ U such that T(u) = v.
(ii) T is an isomorphism (or is invertible) if it is one-one and onto, and then U and V are called isomorphic spaces. This is denoted by U ≈ V.

(5) T : U → V is

(i) one-one if and only if Ker T = (0)
(ii) onto if and only if R (T) = V

(6) Let U and V be **finite**-dimensional vector spaces with the **same dimension.** The T: U → V is 1 -1 iff T is onto iff t is an isomorphism

(7) Two finite – dimensional vector spaces U and V are isomorphic if and only if dim U = dim V.

(8) Let V and W be vector spaces over a field F, and T:V → W be a linear transformation. Then V/Ker T ≈ R(T).

5.0 SUMMARY

E1) For any a1, a2 ∈ F and u1, u2 ∈ U, we know that a1u1 ∈ U and a2u2 ∈ U. therefore, by LT1.

T $(\alpha_1 u_1 + \alpha_1 u_2) = T(\alpha_1 u_1) + T(\alpha_2 u_2)$
$= \alpha_1 T(u_1) + \alpha_1 T(u_2)$, by LT2
Thus, LT3 is true.

E 2) By LT2, T (0, u) = 0.T (u) for any u ∈ U. Thus, T (0) = 0. Similarly, for any u ∈ U, T (-u) = T ((- 1)u) = (-1) T(u) = - T (u).

E3) T (x,y) = (- x, y) ~~∀~~ (x, y) ∈ R^2. (See the geometric view in Fig.4) T is a linear operator. This can be proved the same way as we did in Example 4.

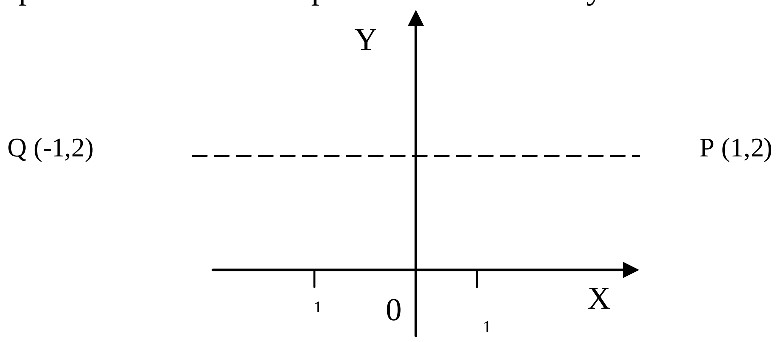

Fig 4 Q is the reflection of 1 in the y-axis

E4) $T((x_1, x_2, x_3) + (y_1, y_2, y_3)) = T(x_1 + y_2, x_3 + y_3)$

$= \quad a_1 (x_1 = y_1) + a_2 (x_2 + y_2) + a_3 (x_3 + y_3)$

$= \quad (a_1x_1 = a_2 x_2 + a_3 x_3) + (a_1, y_1 + a_2 y_2 + a_3y_3)$

$= \quad T(x_1, x_2, x_3) + T(y_1, y_2, y_3)$

Also, for any $\alpha \in R$,

$T(\alpha(x_1, x_2, x_3)) = a_1\alpha x_1 + a_2\alpha x_2 + a_3\alpha x_3$

$= \quad \alpha (a_1 x_1 + a_2 x_2 + a_3 x_3) = \alpha T(x_1, x_2, x_3)$.

Thus, LT1 and LT2 hold for T.

(5) We will check if LT1 and LT2 hold firstly.

$T((x_1 x_2 x_3) + (y_1, y_2, y_3)) = T(x_1 + y_1, x_2 + y_2, x_3 + y_3)$

$= (x_1 + y_1 + x_2 + y_2 – x_3 – y_3, 2x_1 + 2y_1 – x_2 - y_2, x_2 + y_2 + 2x_3 + 2y_3)$
$= (x_1 + x_2 – x_3, 2x_1 – x_2, x_2 + 2x_3) + (y_1 + y_2 – y_3, 2y_1 – y_2 + 2y_3)$
$= T(x_1, x_2, x_3) + T(y_1, y_2, y_3)$ showing that LT1 holds.

Also, for any $\alpha \in R$.

$T(\alpha (x_1, x_2, x_3)) = T(\alpha x_1, \alpha x_2, \alpha x_3)$

$= \quad (\alpha x_1 + \alpha x_2 – \alpha x_3, 2\alpha x_1 - \alpha x_2, \alpha x_2 + 2\alpha x_3)$
$= \quad \alpha (x_1, + x_2 – x_3, 2x_1 – x_2, x_2 + 2x_3) = \alpha T(x_1, x_2, x_3)$, showing that LT2 holds.

E6) We want to show that $D(\alpha f + \beta g) = \alpha D(f) + \beta D(g)$, for any $\alpha, \beta \in R$ and f, g $\in P_n$. Now, let $f(x) = a_0 + a_1 x + a_2x^2 + \ldots + a_nx^n$ and $g(x) = b_0 + b_1 x + \ldots + b_nx^n$.

Then $(\alpha f + \beta g)(x) = \alpha a_0 + \beta b_0) + (\alpha a_1 + \beta b_1) x + \ldots + (\alpha a_n + \beta b_n)$
$:.\ [D(\alpha f + \beta g)](x) = (\alpha a_0 + \beta b_1) + 2(\alpha a_2 + \beta b_2) x + \ldots + n(\alpha a_n + \beta b_n) x^{n-1}$
$= \alpha (a_1 + 2a_2x + \ldots + na_nx^{n-1}) + \beta (b_1 + 2b_2x + \ldots + nb_nx^{n-1})$
$= \alpha\ (Df)(x) + \beta (Dg)(x) = (\alpha Df + \beta Dg)(x)$
Thus, $D(\alpha f + \beta g) = \alpha Df + \beta Dg$, showing that D is a linear map.

E7 No. Because, it T exists, then

$T(2u_1 + u_2) = 2T(u_1) + T(u_2)$.
But $2u_1 + u_2 = u_3$. :. $T(2u_1 + u_2) = T(u_3) = v_3 = (1, 1)$.
On the other hand, $2T(u_1) + T(u_2) = 2v_1 + v_2 = (2,0) + (0, 1)$

= (2, 1) ≠ v3. T
Therefore, LT3 is violated. Therefore, no such T exists.

E8) Note that {(1,0), (0,5)} is a basis for R^2.

Now (3,5) = 3(1,0) + (0, 5).
Therefore, T(3,5) = 3T (1,0) + T (0,5) = 3 (0,1) + (1,0) = (1,3).
Similarly, (5,3) = 5 (1,0) + 3/5 (0, 5).
Therefore, T(5,3) = 5 (1,0) + 3/5 (1,0) = (3/5, 5).
Note that T (5,3) ≠ T (3, 5)

E9) a) $\dim_R C = 2$, a basis being {1,i}, $i = \sqrt{-1}$.
b) Let T:C → R be such that T (1) = α, T (i) = β.

Then, for any element x + iy ∈ C (x, y ∈ R), we have T (x + iy) = xT (1) + yT(i) = xα+ yβ. Thus, T is defined by T(x + iy) = xα + yβ ∀ x + i y ∈ C.

E10) T: U → V : T (U) = 0 ∀ u ∈ U.

:. Ket T = {u E U | T (u) = 0} = U
R(T) = { T(u) |u ∈, U} = {0}. :. 1 ∉ R (T).

E11) a) $R(T) = \{T(x, y, z) | (x, y, z) \in R^3\} = \{(x,y) | (x,y,z) \in R^3 = R^2$.
Ker T = {(x, y, z)|T (x,y,z) = 0} = {(x, y,z)| (x,y) = (0,0)}

= {0,0,z)| z ∈ R}
:. Ker T is the z-axis.

b) R(T) = {z ∈| (x,y,z) ∈ R3} = R
Ker T = {(x,y, 0) |x, y, ∈ R} = xy – plane in R^3.

c) $R(T) = \{(x, y,z) \in R^3 \in | \exists\ x_1, x_2, x_3 \in R$ such that $x = x_1 + x_2 + x_3 = y = z\}$
$= \{x, x, x) \in R^3 | x = x_1 + x_2 + x_3$ for some $x_1, x_2, x_3 \in R\}$
$= \{(x, x, x) \in R^3 | x \in R\}$

Because, for any x ∈ R, (x, x, x) = T (x, 0, 0)
:. R(T) is generated by {(1,1,1)}.
Ker T = $\{(x_1, x_2, x_3) | x_1 + x_2 + x_3 = 0\}$, which is the plane $x_1 + x_2 + x_3 = 0$, in R^3.

E12) Any element of R(T) is of the form T(u), u ∈ U. Since $\{e_1,….e_n)$ generates U, ∃ scalars α_1,…., an such that $u = \alpha_1 e_1 + … + \alpha_n e_n$.

Then $T(u) = \alpha_1 T(e_1) + … + \alpha_n T(e_n)$, that is, T (u) is in the linear span of $\{T(e_1), …., T(e_n)\}$.

:. $\{T(e_1) \ldots\ldots, T(e_n)\}$ generates R(T).

E13) T: V → V: T (v_0 = v. Since R(T) = V and Ker T = (0), we see that rank (T) = dim V, nullity (T) = 0.

E14) $R(D) = \{a_1 + 2a_2x + \ldots + na_nx^{n-1} \mid a_1,\ldots,a_n \in R\}$
Thus, $R(D) \subseteq P^{n-1}$. But any element $b_0 + b_1 x +,\ldots + b_{n-1} x^{n-1}$, in

$$\text{Pn-1 is D} \quad b_0\, x \left(+ \frac{b_1}{2} x_2 + \ldots + \frac{b_{n-1}}{n} x_n \right) \in R\,(D)$$

Therefore, $R(D) = P_{n-1}$.

:. a basis for R(D) is $\{1, x,\ldots, x^{n-1}\}$, and rank (D) = n.

$\text{Ker } D = \{a_0 + a_1x + \ldots + a_nx^n \mid a_1 + 2a_2x +,\ldots + na_n x^{n-1} = 0, a_i \in R\ \forall_i\}$
$= \{a_0 + a_1x + \ldots + a_nx^n \mid a_1 = 0, a_2 = 0, \ldots, a_n = 0, a_i \in R\ \forall_i\}$
$= \{a_0 \mid a_0 \in R\} = R$

:. A basis for Ker D is {1}.

⇒ nullity (D) = 1.

E15 a) we have show that $R(t) = R^2$. :. rank (t) = 2
Therefore, nullity (T) = dim $R^3 - 2 = 1$.

b) rand (T) = 1, nullity (T) = 2
(c) R (T) is generated by {(1,1,1)}. :. Rank (T) = 1.
:. Nullity (T) = 2.

E 16 Now rank (T) + nullity (T) = dim U = 1.
Also rank (T) ≥ 0.
:. The only values rank (T) can take are 0 and 1. If rank (T) = 0, then
dim R(T) = 0.
Thus, R(T) = {0}, that is, R(T) is a point.

If rank (T) = 1, then dim R(T) = 1, That is, R(T) is a vector space over R generated by a single element, v, say. Then R(T) is the line $R_v = \{\alpha v \mid \alpha \in R\}$.

E17) By Theorem 5, nullity (ST) = dim V – rank (ST). By (a) of Theorem 6, we know that – rank (ST) ≥ - rank (S) and – rank (ST) ≥ – rank (T).
:. Nullity (ST) ≥ dim V – rank (S) and nullity (ST) ≥ dim V – rank (T).

Thus, nullity (ST) $\geq$ nullity (S) and nullity (ST) $\geq$ nullity (T). That is, nullity (ST) $\geq$ max {nullity (S), nullity (T)}.

E18) Since $1 \not\in 2$, but 0 (1) = 0 (2) = 0, we find that 0 is not 1– 1

E19) Firstly note that T is a linear transformation. Secondly, T is 1-1 because T (x, y, z) = T (x`, y`, z`) $\Rightarrow$ (x, y z) = (x`, y`, z`)
Thirdly, T is onto because any (x, y, z) $\in R^3$ can be written as T(x, - y, y, z)
:., T is an isomorphism. :. $T^{-1} : R^3 \rightarrow R^3$ exists and is defined by T^{-1} (x, y, z) = (x – y, y, z).

E20) T is not an isomorphism because T is not 1-1, since (1-1, 1) $\in$ Ker T.

E21) The linear operator in E11) (a) suffices.

E22) a) Let $\alpha_1,\ldots.\alpha_k \in$ F such that α_1 $T(u_1) +\ldots+ \alpha_k$ $T(u_k) = 0$
$\Rightarrow T(\alpha_1 u_1 + \ldots + a_k u_k) = 0 = T(0)$
$\Rightarrow \alpha_1 u_1 + \ldots + a_k u_k = 0$, since T is 1 -1
$\Rightarrow \alpha_1 = 0,\ldots, a_k = 0$, since $\{u_1, \ldots.u_k\}$ is linearly independent
:. $\{T(u_1),\ldots, T(u_k)\}$ is linearly independent.

b) No. For example, the zero operator maps every linearly independent set to {0}, which is not linearly independent.

c) Let $T:U \rightarrow V$ be a linear operator, and $\{u_1,\ldots., u_n\}$ be a linearly dependent set of vectors in U. We have to show that $\{T(u_1),\ldots,T(u_n)\}$ is linearly dependent. Since $\{u_1,\ldots..,u_n\}$ is linearly dependent, $\exists$ scalars $a_1,\ldots.,a_n$, not all zero, such that $a_1u_1 + \ldots.+ a_nu_n = 0$.

Then a_1 $T(u_1) +\ldots+ a_n$ $T(u_n)$ = T (0) = 0, so that $\{T(u_1),\ldots.T(u_n)\}$ is linearly dependent.

E23) T is a linear transformation now, if (x, y, z) $\in$ Ker T, then T(x, y, z) = (0, 0 , 0).

:., x + y = 0 = y + z = x + z $\Rightarrow$ x = 0 = y = z
$\Rightarrow$ Ker T = { (0,0,0)}

$\Rightarrow$ T is 1 – 1
:. by Theorem 10, T is invertible.

To define T^{-1} $R^3 \rightarrow R^3$ suppose T^{-1} (x, y, z) = (a, b, c).
Then T (a, b, c) = (x, y, z)

$\Rightarrow$ (a + b, b+ c, a+ c) = (x, y, z)
$\Rightarrow$ a + b = x, b + c = y, a + c = z

x + z – y   x + y – z   y + z - x

$\Rightarrow$ a = $\frac{}{2}$, b = $\frac{}{2}$ a = $\frac{}{2}$

$$:.\ T^{-1}(x, y, z) = \left(\frac{x + z - y}{2}\ \ \frac{x + y - z}{2}\ \ \frac{y + z - x}{2}\right) \text{ for any } (x, y, z) \in R^3.$$

E24) T: U → V is 1 -1. **Suppose T is onto. Then T is an isomorphism and dim U = dim V,** by Theorem 12. Conversely suppose dim U = dim V. Then T is onto by theorem 10.

E25) The Rank Nullity Theorem and Example 14 give
Dim V – nullity (ST) = dim V – nullity (T) – dim (R(T) $\cap$ Ker S)
$\Rightarrow$ nullity (ST) = nullity (T) + dim (R(T) $\cap$ Ker S)

E26) In the case of the direct sum A $\oplus$ B, we have A $\cap$ B = {0}

$$:.\ \frac{.A \oplus B}{B} \approx$$

E27) a) v + W = v\` + W $\Rightarrow$ v-v\` $\in$ W $\subseteq$ U $\Rightarrow$ v-v\` $\in$ U $\Rightarrow$ v + U = v\` + U
$\Rightarrow$ T(v+W) = T (v\` + W)
:. T is well defined.

b) For any v + W, v\` + W in V/W and scalar a, b, we have
T(a(v + W) + b (v\` + W)) = T (av + bv\`+W) = av + bv\` + U
= a(v + U) + b (v\` + U) = aT (v + W) + bT (v\` + W).
:. T is a linear operator.

c) ker T = { v + W | v + U = U}, since U is the “zero” for V/U.
= {v + W | v $\in$U} = U/W.
R(T) = {v + U | v $\in$ V} = V/U.

6.0 TOTUR MARKED – ASSEMENT

7.0 REFERENCE/FURTHER READING

UNIT 2 LINEAR TRANSFORMATION – II

CONTENTS

1.0 INTRODUCTION

In the last unit we introduced you to linear transformations and their properties. We will now show that the set of all linear transformations form a vector space U to a vector space V forms a vector space itself, and its dimension is dim U (dim V). In particular, we define and discuss the dual space of a vector space.

In Unit 1 we defined the composition of two functions. Over here we will discuss the composition of two linear transformations and show that it is again a linear operator. Note that we use the terms 'linear transformation' interchangeably.

Finally, we study polynomials with coefficients from a field F, in a linear operator T:V → V. You will see that every such T satisfies a polynomial equation g(x) = 0. that is, if we substitute T for x in g(x) we get the zero transformation. We will, then, define the minimal polynomial of an operator and discuss some of its properties. These ideas will crop up again in Unit 11.

You must revise Units 1 and 5 before going further.

2.0 OBJECTIVES

After reading this unit, you should be able to

- Prove and use the fact that L (U, V) is a vector space of dimension (dim U) (dim V);
- Use dual bases, whenever convenient;

- Obtain the composition of two linear operators, whenever possible;
- Obtain the minimal polynomial of a linear transformation T:V → V in some simple cases;
- Obtain the inverse of an isomorphism T: V → V if its minimal polynomial is known.

3.0 MAIN CONTENT

3.1 Introduction

3.3 The Vector Space L (U, V)

By now you must be quite familiar with linear operators, as well as vector spaces. In this section we consider the set of al linear operators from one vector space to another, and show that it forms a vector space.

Let U, V be vector spaces over a field F. consider the set of all linear transformations from U to V. We denote this set by L(U, V).

We will now define addition and scalar multiplication in L(U, V) so that L (U,V) becomes a vector space.

Suppose S, T $\in$ L (U, V) (that is, S and T are linear operators from U to V). we define (S + T: U → V by

$(S + T) (u) = S (u) + T (u) \forall u \in U$.

Now, for $a_1, a_2 \in F$ and $u_1, u_2 \in U$, we have

(S + T) (a1 u1 + a2 u2)

$$
\begin{aligned}
&= S(a_1 u_1 + a_2 u_2) + T (a_1, u_1 + a_2 u_2) \\
&= a_1 S (u_1) + a_2 S (u_2) + a_1 T(u_1) + a_2 T(u_2) \\
&= a_1(S(u_1) + T(u_1)) + a_2 (S(u_2) + T(u_2)) \\
&= a_1 (S + T) (u_1) + a_2 (S + T) (u_2)
\end{aligned}
$$

Hence, S + T $\in$ L (U, V).

Next, suppose S $\in$ L (U, V) and $\alpha \in$ F. We define α S: U → V as follows:
$(\alpha S) (u) = \alpha S(u) \forall u \in U$.

Is αS a linear operator? To answer this take $\beta_1, \beta_2\beta \in F$ and $u_1, u_2 \in U$. Then, $(\alpha S) (\beta_1, u_1 + b_2 u_2) = \alpha S (\beta_1 u_1 + \beta_2 u_2) = \alpha[\beta_1 S(u_1) + \beta_2 S(u_2)]$

$= \beta 1 (\alpha S) (u_1) + \beta_2 (\alpha S) (u_2)$

Hence, $\alpha S \in L(U, V)$.

So we have successfully defined addition and scalar multiplication on L(U, V).

E E1 Show that the set L (U, V) is a vector space over F with respect to the operations of addition and multiplication by scalars defined above. (Hint: The zero vector in this space is the zero transformation).

Notation: For any vector space V we denote L (V, V) by A(V).

Let U and V be vector spaces over F of dimensions m and n, respectively. We have already observed that L(U, V) is a vector space over F. therefore, it must have a dimension. We now show that the dimension of L(U, V) is mn.

Theorem 1: Let U, V be vector spaces over a field F of dimensions m and n, respectively then L(U, V) is a vector space of dimension mn.

Proof: Let $\{e_1,\ldots e_m\}$ be a basis of U and $\{f_1,\ldots f_n\}$ be a basis of V. By Theorem 3 of Unit 5, there exists a unique linear transformation $E_{11} \in L(U, V)$, such that

$E_{11}(e_1) = f_1, E_{11}(e_2) = 0, \ldots, E_{11}(e_m) = 0$
Similarly, $E_{12} \in L(U, V)$ such that

$E_{12}(e_1) = 0, E_{12}(e_2) = f_1, E_{12}(e_3) = 0, \ldots, E_{12}(e_m) = 0$.
In general, there exist $E_{ij} \in L(U, V)$ for I = 1, …., n, j = 1, …., m, such that $E_{ij}(e_j) = f_i$ and $E_{ij}(e_k) = 0$ for $j \neq k$.
To get used to these Eij try the following exercise before continuing the proof.

E E2) Clearly define E_{2m}, E_{32} and E_{mn},

Now, let us go on with the proof of Theorem 1.

If $u = c_1 e_1 + \ldots + c_m e_m$, where $c_1 \in F \ \forall$ i, then $E_{ij}(u) = c_j f_i$.

We complete the proof by showing that $\{E_{ij} | i = 1, \ldots m\}$ is a basis of L(U, V).

Let us first show that set is linearly independent over F. for this, suppose

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij} E_{ij} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

where $c_{ij} \in$ F. we must show that $c_{ij} = 0$ for all i, j.

(1) Implies that

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij} E_{ij}(ek) = 0 \quad \forall\, k = 1, \ldots, m.$$

Thus, by definition of E_{ij}'s, we get

$$\sum_{i=1}^{n} c_{ik} f_i = 0,$$

But, $\{f_1, \ldots, f_n\}$ is a basis for V thus, $c_{ik} = 0$ for all I = 1, …., n.

But this is true for all k = 1, …., m.

Hence, we conclude that $c_{ij} = 0 \ \forall$ i, j. therefore, the set of E_{ij}'s is linearly independent.

Next, we show that the set $\{E_{ij} \mid i = 1, \ldots, n, j = 1, \ldots, m\}$ spans L(U, V). Suppose $T \in L(U, V)$.

Now, for each j such that $1 \leq j \leq m$, $T(e_j) \in V$. since $\{f_1, \ldots. f_n\}$ is a basis of V, there exist scalars $c_{ij}, \ldots, c_{in}$ such that

$$T(e_j) = \sum_{i=1}^{n} c_{ij} f_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

we shall prove that

$$T = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij} E_{ij} \qquad \ldots\ldots\ldots\ldots\ldots\ldots .(3)$$

By Theorem 1 of Unit 5 it is enough to show that, for each k with $1 \leq k \leq m$,

$$T(e_k) = \sum_{i}\sum_{j} c_{ij} E_{ij}(e_k).$$

Now, $\sum_{i}^{n}\sum_{j}^{m} c_{ij} E_{ij}(e_k). = \sum_{i=1}^{n} c_{ik} f_i = T(e_k)$, by (2). This implies (3).

Thus, we have proved that the set of mn elements $\{E_{ij} \mid i = 1, \ldots, n, j = 1, \ldots m\}$ is a basis for L(U,V).

Let us see some ways of using this theorem.

Example 1: Show that $L(R^2, R)$ is a plane.

Solution: $L(R^2, R)$ is a real vector space of dimension 2 x 1 = 2.

Thus, by Theorem 12 of Unit 5 $L(R^2, R) \approx R^2$, the real plane.

Example 2: Let U, V be vector spaces of dimensions m and n, respectively. Suppose W is a subspace of V of dimension p ($\leq$ n). Let

$X = \{ T \in L(U, V): T(u) \in W \text{ for all } u \in U\}$

Is X a subspace of L(U, V)? if yes, find its dimension.

Solution: $X = \{T \in L(U, V) \mid T(U) \subseteq W\} = L(U, W)$. thus, X is also a vector space. Since it is a subset of L(U, V), it is a subspace of L(U, V). By Theorem 1, dim X = mp.

E E3 What can be a basis for $L(R^2, R)$, and for $L(R, R^2)$? Notice that both these spaces have the same dimension over R.

After having looked at L(U, V), we now discuss this vector space for the particular case when V = F.

3.4 The Dual Space

The vector space L(U, V), discussed in sec. 2.2, has a particular name when V = F.

Definition: Let U be a vector space over F. Then the space L(U, F) is called the dual space of U*, and is denoted by U.

In this section we shall study some basic properties of U*. the elements of U have a specific name, which we now give.

Definition: A linear transformation T:U → F is called a **linear functional**. Thus, a linear functional on U is a function T: U → F such that $T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_1 T(u_2)$, for $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$.
For example, the map $f:R^3 \to R$:f $(x_1, x_2, x_3) = a_1 x_1 + a_2 x_2 + a_3 + a_3$, where $a_1, a_2, a_3 \in R$ are fixed, is a linear functional on R^3. You have already seen this in Unit 5 (E4).

We now come to a very important aspect of the dual space.

We know that the space V*, of linear functional on V, is a vector space. Also, if dim V = m, then dim V* = m, by Theorem 1. (Remember, dim F = 1).

Hence, we see that dim V = dim V*. From Theorem 12 of unit 5, it follows that the vector spaces V and, V* are isomorphic.

We now construct a special basis for V*. Let $\{e_1,\ldots.e_m\}$ be a basis for V. by Theorem 3 of Unit 5, for each i = 1, …, m, there exists a unique linear functional f_i on V such that

$$f_i(e_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

$$= \delta_{ij}$$

The Kronecker delta function is

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

We will prove that the linear functional $f_1, \ldots, f_m$, constructed above, form a basis of V*.

Since dim V = dim V* = m, it is enough to show that the set $\{f_1,\ldots.f_m\}$ is linearly independent. For this we suppose $e_1,\ldots,c_m \in F$ such that $c_1f_1 + \ldots + c_m f_m = 0$.

We must show that $C_1 = 0$, for all i.

Now $\sum_{j=1}^{n} c_j f_j = 0$

$$\Rightarrow \sum_{j=1}^{n} \left(c_j f_j \quad (e_i) = \right) 0, \text{ for each i.}$$

$$\Rightarrow \sum_{j=1}^{n} c_j \; (f_j(e_i)) = 0 \; \forall \; i.$$

$$\Rightarrow \sum_{j=1}^{n} c_j \; \delta_{ij} = 0 \; \forall \; i \Rightarrow ci = 0 \; \forall \; i$$

Thus, the set $\{f_1,\ldots.f_m\}$ is a set of m linearly independent elements of a vector space V* of dimension m, Thus, from Unit 4 (Theorem 5, Cor. 1), it forms a basis of V*.

Definition: The basis $\{f_1,\ldots.,f_m\}$ of V* is called the dual basis of the basis $\{e_1,\ldots., e_m\}$ of V.

We now come to the result that shows the convenience of using a dual basis.

Theorem 2: Let V be a vector space over F of dimension n, $\{e_1,\ldots e_n\}$ be a basis of V and $\{f_1,\ldots f_n\}$ be the dual basis of $\{e_1,\ldots,e_n\}$.Then, for each $f \in V^*$.

$$f = \sum^{n} f(e_j) f_i.$$

$$i=1$$

and, for each $v \in V$,

$$v = \sum_{i=1}^{n} f_i.(v)e_i.$$

Proof: Since $\{f_1, \ldots, f_n\}$ is a basis of V^*, for $f \in V^*$ there exist scalars $c_1, \ldots c_n$ such that

$$f = \sum_{i=1}^{n} c_i f_i.$$

Therefore,

$$f(e_j) = \sum_{i=1}^{n} c_i f_i.(e_j)$$

$$= \sum_{i=1}^{n} c_i \delta_{ij}, \text{ by definition of dual basis.}$$

$$= c_j.$$

This implies that $c_1 = f(e_i)$ ∀ $i = 1, \ldots, n$. therefore, $f = \Sigma f_1$, Similarly, for $v \Sigma\in V$, there exist scalars $a_1, \ldots, a_n$ such that

$$v = \sum_{i=1}^{n} a_i e_i.$$

Hence,

$$f_j(v) = \sum_{i=1}^{n} a_i f_i.(e_j)$$

$$= \sum_{i=1}^{n} a_i \delta_{ij}.$$

$$= a_j,$$

and we obtain

$$v = \sum_{i=1}^{n} f_i (v) e_j$$

Let us see an example of how this theorem works.

Example 3: Consider the basis $e_1 = (1, 0 – 1)$, $e_2 = (1, 1, 0)$ of C^3 over C. Find the dual basis of $\{e_1, e_2, e_3\}$.

Solution: Any element of C^3 is $v = (z_1, z_2, z_3)$, $z_1 \in C$. Since $\{e_1, e_2, e_3\}$ is a basis, we have $\alpha_1, \alpha_2, \alpha_3 \in$ C. Since That

$$V = \{ z_1, z_2, z_3\} = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$
$$= (\alpha_1 + \alpha_2 + \alpha_3, \alpha_2 + \alpha_3, - \alpha_1 + \alpha_2)$$

Thus, $\alpha\alpha + \alpha_2 + \alpha_3 = z_1$
$\alpha_2 + \alpha_3 = z_2$
$-\alpha_1 + \alpha_2 = z_3$.

These equations can be solved to get
$\alpha_1 = z_1 – z_2$, $\alpha_2 = z_1 – z_2 + z_3$, $\alpha_3 = 2z_2 – z_3 – z_3$

Now, by Theorem 2,

$v = f_1, (v) e_1 + f_2 (v) e_2 + f_3 (v) e_3$, where $\{f_1, f_2, f_3\}$ is the dual basis. Also $v = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$.

Hence, $f_1, (v) \alpha_1, f_2(v) = \alpha_2$, $f_3(v) = \alpha_3$ ~~V~~ $v \in C^3$.

Thus, the dual basis of $\{e_1, e_2, e_3\}$ is $\{f_1, f_2, f_3\}$, where f_1, f_2, f_3 will be defined as follows:

$f_1 (z_1, z_2, z_3) = \alpha_1 = z_1 – z_2$.
$f_1(z_1, z_2, z_3) = \alpha_2 = z_1 –z_2 + z_3$
$f_1 (z_1, z_2, z_3) = \alpha_3 = 2z_2 – z_1 – z_3$.

E E5 What is the dual basis for the basis $\{1, x, x^2\}$ of the space

$P_2 = (a_0 + a_1x + a_2x^2 | a_1 \in R\}$?

Now let us look at the dual of the dual space. If you like, you may skip this portion and go straight to sec. 6. 4.

Let V be an n-dimensional vector space. We have already seen that V and V* are isomorphic because dim V = dim V*. The dual of V* is called the **second dual of V** and is denoted by V**. We will show that $V \approx V^{**}$.

Now any element of V** is a linear transformation from V* to F. Also, for any $v \in V$ and $f \in V^*$, $f(v) \in F$. So we define a mapping $\phi : V \rightarrow V^{**}: v \rightarrow \phi\ v$, where $(\phi\ v)(f) = f(v)$ for all $f \in V^*$ and $v \in V$. (Over here we will us $\phi\ (v)$ and ϕv interchangeably).

Note that, for any $v \in V$, $\phi\ v$ is a well defined mapping from $V^* \rightarrow F$. We have to check that it is a linear mapping.

Now, for $c_1, c_2 \in F$ and $f_1, f_2 \in V^*$.

$(\phi\ v)\ (c_1, f_1 + c_2\ f_2) = c_1\ f_1 + c_2\ f_2)\ (v)$

$= c_1\ f_1\ (v) + c_2\ f_2\ (v)$

$= c_1\ (\phi\ v)\ (f_1) + c_2\ (\phi\ v)\ (f_2)$

$\therefore \phi\ v \in L(V^*, F) = V^{**}, \forall v.$

Furthermore, the map $\varnothing : V \rightarrow V^{**}$ is linear. This can be seen as follows: for $c_1, c_2 \in F$ and $v_1, v_2, \in V$.

$\varnothing\ (c1\ v1 + c2\ v2)\ (f) = f(c_1\ v_1 + c_2\ v_2)$

$= c_1 f(v_1) + c_2\ f(v_2)$

$= c_1\ (\varnothing\ v_1)\ (f) + c_2\ (\varnothing\ v_2\ (f)$

$= (c_1 \varnothing\ v_1 + c_2 \varnothing\ v_2)\ (f).$

This is true $\forall$ f $\in$ V*. Thus, 0 $(c_1 v_1 + c_2 v_2) = c_1\varnothing (v_1) + c_2\varnothing (v_2)$.

Now that we have shown that 0 is linear, we want to show that it is actually an isomorphism. We will show that 0 is 1 -1. For this. By Theorem 7 of Unit 5, it suffices to show that $\varnothing$ (v) = 0 implies v = 0. Let $\{f_1,\ldots,f_n\}$ be the dual basis of a basis $\{e_1,\ldots,e_n\}$ of V.

By Theorem 2. we have $v = \sum_{i=1}^{n} f_1(v) e_i$.

Now $\varnothing$ (v) = 0 $\Rightarrow$ ($\varnothing$v) (f1) = 0 $\forall$ i = 1, …., n.

$\Rightarrow f_1(v) = 0$ $\forall$ i = 1, ….,n

$\Rightarrow v = \Sigma f_1 (v) e_1 = 0$

Hence, it follows that $\varnothing$ is 1 -1. thus, $\varnothing$ is an isomorphism (Unit 5, Theorem 10).

What we have just proved is the following theorem.

Theorem 3: The map $\varnothing$: V → V**, defined by ($\varnothing$ v) (f) = f(v) $\forall$ v $\in$ V and f $\in$ V*, is an isomorphism.

We now give an important corollary to this theorem.

Corollary: Let ψ be a linear functional on V* (i.e., ψ E V**).

Then there exists a unique v $\in$ V such that ψ (f) = f(v) for all f $\in$V*.

Proof: By Theorem 3, since $\varnothing$is an isomorphism, it is onto and 1 -1. thus, there exists a unique v $\in$ V such that $\varnothing$ (v) = ψ. This by definition, implies that ψ (f) = ($\varnothing$v) (f) = f(v) for all f $\in$V*.

Using the second dual try to prove the following exercise.

E E6) Show that each basis of V* is the dual of some basis of V.

In the following section we look at the composition of linear operators, and the vector space A(v), where V is a vector space over F.

3.5 Composition of Linear Transformations

Do you remember the definition f the composition of functions, which you studied in Unit 1? Let us now consider the particular case of the composition of two linear transformations. Suppose T: U → V W, defined by $S_0T(u) = S(T(u))$ ∀ $u \in U$.

This is diagrammatically represented in Fig. 1.

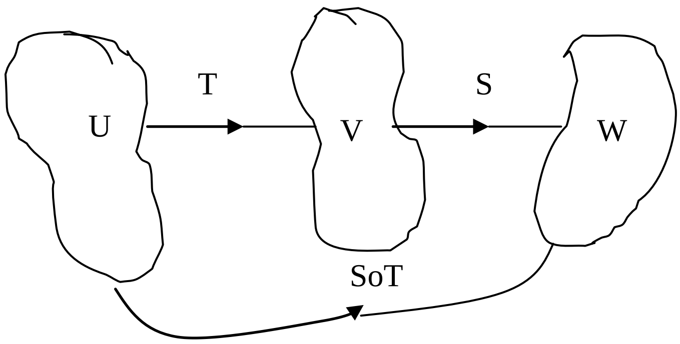

***Fig 1:** SoT is the composition of S and T.*

The first question which comes to our mind is whether S0T is linear. The affirmative answer is given by the following result.

Theorem 4: Let U, V, W be vector spaces over F. suppose $S \in L(V, W)$ and $T \in L(U, V)$. Then $SoT \in L(U, W)$.

Proof: All we need to prove is the linearity of the map SoT. Let $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$. Then

$$
\begin{aligned}
SoT(\alpha_1 u_1 + a_2 u_2) &= S(T(\alpha_1 u_1 + \alpha_2 u_2)) \\
&= S(\alpha_1 T(u_1) + \alpha_2 T(u_2)), \text{ since T is linear} \\
&= \alpha_1 S(T(u_1)) + \alpha_2 S(T(u_2)), \text{ since S is linear} \\
&= a_1 SoT(u_1) + \alpha_2 SoT(u_2)
\end{aligned}
$$

This shows that $SoT \in L(U, W)$

Try the following exercises now

E E7) Let I be the identity operator on V. show that SoI = IoS = S for all $S \in A(V)$.

E E8) Prove that So0 = 0oS = 0 for all S ∈ A (V). where 0 is the null operator.

We now make an observation.

Remark: Let S:V → V be an invertible linear transformation (ref. Sec 1.4), that is an isomorphism. Then, by Unit 5, Theorem 8, $S^{-1} \in$ L (V, V) = A (V).

Since S^{-1} oS (v) = v and SoS^{-1} (v) = v for all v ∈ V.

SoS^{-1} = S^{-1} 0S = I_v, where I, denotes the identity transformation on V. this remark leads us to the following interesting result.

Theorem 5: Let V be a vector space over a field F. A linear transformation S ∈ A(V) is an isomorphism if an only if ∃ T ∈ A (V) such that SoT = I = ToS.

Proof: Let us first assume that S is an isomorphism. Then, the remark above tells us that ∃ $S^{-1} \in$ A(V) such that SoS^{-1} = I = S^{-1}oS. Thus, we have T (= S-1) such that SoT = ToS = I.

Conversely, suppose T exists in A(V), such that SoT = I = ToS. We want to show that S is 1 -1 and onto

We first show that S is 1 -1, that is, Ker S = {0}. Now, x ∈ Ker S ⇒ S(x) = 0 ⇒ ToS (x) = T{0} = 0 ⇒ I(x) = 0 ⇒ x = 0. thus, Ker S = {0}.

Next, we show that S is onto, that is, for any v ∈ V, ∃ u ∈ V such that S(u) = v. Now, for any v ∈ V,

v = I (v) = SoT (v) = S (T(v)) = S(u), where u = T(v) ∈ V. thus, S is onto.

Hence, S is 1 -1 and onto, that is, S is an isomorphism.

Use Theorem 5 to solve the following exercise.

E E9) Let S $(x_1, x_2) = (x_2 = x_1)$ and T $(x_1, x_2) = (- x_2, x_1)$. Find SoT and ToS. Is S (or T) invertible?

Now, let us look at some examples involving the composite of linear operators.

Example 4: Let T : $R^2 \rightarrow R^2$ and S:$R^3 \rightarrow$ be defined by

$T(x_1, x_2) = (x_1, x_2, x_1 + x_2)$ and $S(x_1, x_2, x_3) = (x_1, x_2)$. Find SoT and ToS.

Solution: First, note that T $\in$ L (R^2, R^3) and S $\in$ L (R^3, R^2). :. SoT and ToS are both well defined linear operators. Now,

SoT (x_1, x_2) = S $(T(x_1, x_2))$ = S $(x_1, x_2, x_1 + x_2) = (x_1, x_2)$.

Hence, SoT = the identity transformation of $R^2 = I_{R2}$.

Now, ToS $(x_1, x_2, x_3) = T(S(x_1, x_2, x_3))$ = T $(x_1, x_2) = x_1, x_2, x_1 + x_2)$.

In this case SoT $\in$ A (R^2), while ToS $\in$ A (R^3). Clearly, SoT, $\neq$ ToS. Also, note that SoT = I, but ToS $\neq$ I.

Remark: Even if SoT and ToS both being to A(V), SoT may not be equal to ToS. We give such an example below.

Example 5: Let S, T $\in$ A (R^2) be defined by T $(x_1, x_2) = (x_1 \_ x_2, x_1 - x_2)$ and S $(x_1, x_2) = 0, x_2)$. Show that SoT $\neq$ ToS.

Solution: You can check that SoT $(x_1, x_2) = (0, x_1 - x_2)$ and ToS $(x_1, x_2) = (x_2 - x_2)$. Thus, $\exists$ $(x_1, x_2) \in R^2$ such that SoT $(x_1, x_2) \neq$ ToS (x_1, x_2) (for instance, SoT (1, 1) $\neq$ ToS (1, 1)). That is, SoT $\neq$ ToS.

Note: Before checking whether SoT is a well defined linear operator. You must be sure that both S and T are well defined linear operators.

Now try to solve the following exercise.

E E10) Let T $(x_1, x_2) = (0, x_1, x_2)$ and S $(x_1, x_2, x_3) = (x_1, + x_2, x_2 + x_3)$ Find SoT and ToS. When is SoT = ToS?

E E11) Let T $(x_1, x_2) = 2x_1, x_1 + 2x_2)$ for $(x_1, x_2) \in R^2$, and S $(x_1, x_2, x_3) = (x_1 + 2x_2, 3x_1 - x_2, x_3)$ for $(x_1, x_2, x_3) \in R^3$. Are SoT and ToS defined? If yes, find them.

E E12) Let U, V, W, Z be vector spaces over F. suppose T $\in$ L(U, V), S $\in$ L (V, W) and R $\in$ L(W, Z). show that (RoS) oT = Ro (SoT).

E E13) Let S, T $\in$ A (V) and S be invertible. Show that rank (ST) = rank (TS) = rank (T). (ST).

So far we have discussed the composition of linear transformation. We have seen that if S, T $\in$ A (V), then SoT $\in$ A (V), where V is a vector space of dimension n. Thus, we have introduced another binary operation (see 1. 5.2) in A (V), namely, the composition of operators, denoted by o. Remember, we already have the binary operations given in Sec. 6.2 In the following theorem we sate some simple properties that involve all these operations.

Theorem 6: Let R. S. T $\in$ A (V) and let $\alpha \in$ F. Then

(a) Ro (S + T) = RoS + RoT, and (S + T) oR = SoR + ToR.
(b) α(SoT) = αSoT = SoαT.

Proof: a) for any V $\in$ V,

Ro (S + T) (v) = R ((S + T) (v)) = R (S(v) + T (v))
= R(S(v) + R (T(v))
= (RoS) (v) + (RoT) (v)
= (RoS + RoT) (v)

Hence, Ro (S + T) = RoS + RoT.

Similarly, we can prove that (S + T) oR = SoR + ToR

b) For any v E V, α (SoT) (v) = α (S(T(v))
= (αS) (T(v))
= (αSoT) (v)

Therefore, α (SoT) = α SoT.

Similarly, we can show that α (SoT) = SoαT.

Notification: In future we shall be writing ST in place of SoT. Thus, ST(u) = S(T(u)) = (SoT) u. Also, if T $\in$ A (V), we write T^0 = I, T^1 = T, T^2 = ToT and, in general, T^n = T^{n-1} oT = ToT^{n-1}.

The properties of A(V), stated in theorems 1 and 6 are very important and will be used implicitly again and again. To get used to A(V) and the operations in it try the following exercises.

E E14) Consider S, T:$R^2 \rightarrow R^2$ defined by S(x_1, x_2) = x_1, - x_2) and T (x_1, x_2) = (x_1 + x_2, x_2 – x_3). What are S + T, ST, TS, So (S-T) and (S – T) oS?

E E15) Let S E A (V), dim V = n and rank (S) = r, Let
M = {T $\in$ A (V) | ST = 0},
N = {T $\in$ A (V) | TS = 0}

a) Show that M and N are subspaces of A (V).
b) Show that M = L (V, Ker S). What is dim M?

By now you must have got used to handling the elements of A(V). the next section deals with polynomials that are related to these elements.

3.6 Minimal Polynomial

Recall that a polynomial in one variable x over F is of the form p(x) = $a_0 + a_1x +,\ldots.+ a_n x_n$, where $a_0, a_1,\ldots.., a_n \in$ F.

If an $\neq$ o, then p(x) is said to be of degree n. If an = 1, then p (x) is called a **monic polynomial** of degree n. For example, $x^2 + 5x + 6$ is a monic polymial of degree 2. The set of all polynomials in x with coefficients in F is denoted by F [x].

Definition: For a polynomial p, as above, and an operator T $\in$ A (V), we define p (T) = $a_0 I + a_1 T +\ldots.+ a_nT^n$.

Since each of I, T,…., $T^n \in$ A (V), we find P (T) $\in$ A (V). We say P (T) $\in$ F [T]. If q is another polynomial in x over F, then P (T) q (T) = q (T) = P(T), that is, P(T) and q (T) commute with each other. This can be seen as follows:

Let q(T) = $b_0 I + b_1 T + \ldots + b_m T^m$

Then p (T) q (T) $= (a_0 I + a_1 T +\ldots.+ a_n T_n) (b_0I + b_1 T +\ldots.+ b_mT^m)$
$= a_0 b_0 I + (a_0 b_1 + a_1 b_0) T +\ldots + a_n b_m T_m T^{n+m}$
$= (b_0 I + b_1 T +\ldots.+ b_mT^m) (a_0 I + a_1 T +\ldots.+ a_n T^n)$
= q (T) p (T).

E E16) Let p, q E F [x] such that p (T) = 0, q (T) = 0. Show that (p + q) (T) = 0. ((p + q) (x) means p (x) + q (x)).

E E17) Check that $(2I + 3S + S^3)$ commutes with $(S + 2S^4)$, for $S \in A(R^n)$.

We now go on to prove that given any $T \in A(V)$ we can find a polynomial $g \in F[x]$ such that

$g(T) = 0$, that is, $g(T)(v) = 0 \ \forall \ v \in V$.

Theorem 7: Let V be a vector space over F of dimension n and $T \in A(V)$. Then there exists a non-zero polynomial g over F such that $g(T) = 0$ and the degree of g is at most n^2.

Proof: We have already seen that A(V) is a vector space of dimension n^2. Hence, the set $\{I, T, T^2, \ldots, T^{n2}\}$ of $n^2 + 1$ vectors of A (V), must be linearly dependent (ref. Unit 4, Theorem 7). Therefore, there must exist $a_0, a_1, \ldots, a_n{}^2 \in F$ (not all zero) such that $a_0 I + a_1 T + \ldots + a_n{}^2 T^{n2} = 0$.

Let g be the polynomial of degree at most n^2, such that $g(T) = 0$.

The following exercise will help you in getting used to polynomials in x and T.

E E18) Give an example of polynomials g (x) and h (x) in r [x], for which g (I) = 0 and h (0) = 0, where I and 0 are the identity and zero transformations in A (R^3).

E E19) Let T E A (V). then we have a map $\varnothing$ from F [x] to A (V) given by $\varnothing(p) = p(T)$ show that, for a $b \in F$ and $p, q \in F[x]$,

a) $\varnothing(ap + bq) = a\varnothing(p) + b\varnothing(q)$.
b) $\varnothing(pq) = \varnothing(p)\varnothing(q)$.

In Theorem 7: we have proved that there exists some $g \in F[x]$ with $g(T) = 0$. But, if $g(T) = 0$, then $(\alpha g)(T) = 0$, for any $\alpha \in F$. Also, if $\deg g \leq n^2$. Thus, there are infinitely many polynomials that satisfy the conditions in theorem 7. But if we insist on some more conditions on the polynomial g, then we end up with one and only one polynomial which will satisfy these conditions and the conditions in Theorem 7. Let us see what the conditions are.

Theorem 8: Let $T \in A(V)$. then there exists a unique **monic polynomial p of smallest degree** such that $p(T) = 0$.

Proof: Consider the set $S = \{g \in F[x] \mid g(T) = 0\}$. This set is non-empty since, by Theorem 7, there exists a non-zero polynomial g, of degree at most n^2, such that $g(T) = 0$. Now consider the set $D = \{\deg f \mid f \in S\}$. Then D is a subset of $N \cup \{0\}$, and therefore, it must have a minimum element, m. Let $h \in S$ such that $\deg h = m$. then $h(T) = 0$ and $\deg h \leq \deg g \ \forall\ g \in S$.

If $h = a_0 + a_1 x + \ldots + a_m x^m$, $a_m \neq 0$, then $p = a_m^{-1} h$ is a monic polynomial such that $p(T) = 0$, Also $\deg p = \deg h \leq \deg g \ \forall\ g \in S$. Thus, we have shown that there exists a monic polynomial p, of least degree, such that $p(T) = 0$.

We now show that p is unique, that is, if q is any monic polynomial of smallest degree such that $q(T) = 0$, then $p = q$. But this is easy. Firstly, since $\deg p \leq \deg g \ \forall\ g \in S$, $\deg p \leq \deg q$.

Similarly, $\deg q \leq \deg p$. :. $\text{Deg } p = \deg q$.

Now suppose $p(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + x^n$ and $q(x) = b_0 + b_1 x + \ldots + b_{n-1} x^{n-1} + x^n$.

Since $p(T) = 0$ and $q(T) = 0$, we get $(p - q)(T) = 0$. But $p - q = (a_0 - b_0) + \ldots + (a_{n-1} - b_{n-1}) x^{n-1}$. Hence, $(p - q)$ is a polynomial of degree strictly less than the degree of p, such that $(p - q)(T) = 0$. That is, $p - q \in S$ with $\deg(p - q) < \deg p$. This is a contradiction to the way we chose p, unless $p - q = 0$, that is, $p = q$. :. P is the unique polynomial satisfying the conditions of Theorem 8.

This theorem immediately leads us to the following definition.

Definition: For T E A (V), the unique monic polynomial p of smallest degree such that $p(T) = 0$ is called the **minimal polynomial of T.**

Note that the minimal polynomial p, of T, is uniquely determined by the following three properties.

1) p is a monic polynomial over F
2) $p(T) = 0$

3) if $g \in F(x)$ with $g(T) = 0$, then $\deg p \leq \deg g$.

Consider the following example and exercises.

Example 6: For any vector space V, find the minimal polynomials for I, the identity transformation, and 0, the zero transformation.

Solution: Let $p(x) = x - 1$ and $q(x) = x$. Then p and q are monic such that $p(I) = 0$ and $q(0) = 0$. Clearly no non-zero polynomials of smaller degree have the above properties. Thus $x - 1$ and x are the required polynomials.

E E20) Define $T: R^3 \rightarrow R^3$: $T(x_1, x_2, x_3) = (0, x_1, x_2)$. Show that the minimal polynomial of T is x^3.

E E21) Define $T: R^n \rightarrow R^n$: $T(x_1, \ldots\ldots, x_{n-1})$. What is the minimal polynomial of T? (Does E 20 help you?

E E22) Let $T: R^3 \rightarrow R^3$ be defined by

$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3)$. Show that $(T^2 - I)(T - 3I)$

= 0. what is the minimal polynomial of T?

We will now state and prove a criterion by which we can obtain the minimal polynomial of linear operator T, once we know any polynomial $f \in F[x]$ with f (T) = 0. It says that the minimal polynomial must be a factor of any such f.

Theorem 9: Let $T \in A(V)$ and let p (x) be the minimal polynomial of T. Let f (x) be any polynomial such that f (T) = 0. Then there exists a polynomial g (x) such that f (x) = p (x) g (x).

Proof: The division algorithm states that given f (x) and p (x), there exist polynomials g (x) and h (x) such that f (x) = p (x) g (x) + h (x), where h (x) = 0 or deg h (x) < deg p (x). Now,

0 = f (T) = p (T) g (T) + h (T) = h (T), since p (T^) = 0
Therefore, if h (x) $\neq$ 0, then h (T) = 0, and deg h (x) < deg p (x).

This contradicts the fact that p (x) is the minimal polynomial of T. hence, h (x) = 0 and we get f (x) = p (x) g (x).

Using this theorem, can you obtain the minimal polynomial of T in E22 more easily? Now we only need to check if T-I, T + I or T – 3I are 0.

Remark: if dim V = n and $T \in A(V)$, we have seen that the degree of the minimal polynomial p of $T \leq n^2$. We will study a systematic method of finding the minimal polynomial of T, and some applications of this polynomial. But now we will only illustrate one application of the concept of the minimal polynomial by proving the following theorem.

Theorem 10: Let $T \in A(V)$. Then T is invertible if and only if the constant term in the minimal polynomial of T is not zero.

Proof: Let $p(x) = a_0 + a_1 x + \ldots + a_m x^{m-1} + x^m$ be the minimal polynomial of T. Then $a_0 I + a_1 T + \ldots + a_{m-1} T^{m-1} + T^m – 0$.
$\Rightarrow T(a_1 I + \ldots + a_{m-1} T^{m-2} + T^{m-1}) = -a_0 I$ …………………………..(1)

Firstly, we will show that if T^{-1} exists, then $a_0 \neq 0$. On the contrary, suppose $a_0 = 0$ Then (1) implies that $T(a_1 I + \ldots. + T^{m-1}) = 0$.. Multiplying both sides by T^{-1} on the left, we get $a_1 I + \ldots + T^{m-1} = 0$.

This equation gives us a monic polynomial $q(x) = a_1 + \ldots. + x^{m-1}$ such that $q(T) = 0$ and deg q < deg p. this contradicts the fact that p is the minimal polynomial of T. Therefore, if T-1 exists the constant term in the minimal polynomial of T cannot be zero.

Conversely; suppose the constant term in the minimal polynomial of T is not zero, that is, $a_0 \neq 0$. Then dividing Equation (1) on both sides by $(-a_0)$, we get

$T((-a_1/a_0) I + \ldots + (-I/a_0) T^{m-1}) = I$
Let $S = (-a_1/a_0) I + \ldots + (-I/a_0) T^{m-1}$.

Then we have ST = I and TS = I. This shows, by Theorem 5, that T^{-1} exists and $T^{-1} = S$.

E E23) Let Pn be the space of all polynomials of degree ≤ n. Consider the linear operator D: $P_2 \rightarrow P_2$ given by $D(a_0 + a_1 x + a_2 x^2) = a_1 + 2a_2 x$. (Note that D is just the differentiation operator.) Show that $D^4 = 0$. What is the minimal polynomial of D? is D invertible?

E E24) Consider the reflection transformation given in Unit 5, Example 4, Find its minimal polynomial. Is T invertible? If so, find its inverse.

E E25) Let the minimal polynomial of $S \in A(V)$ be x^n, $n \geq I$. Show that there exists $v_0 \in V$ such that the set $\{v_0, S(v_0), \ldots, S^{n-1}(V_0)\}$ is linearly independent.

We will now end the unit by summarizing what we have covered in it.

2.6 Summary

In this unit we covered the following points.

i. L(U, V, the vector space of all linear transformations from U to V is of dimension (dim U) (dim V).

ii. The dual space of a vector space V is L (V, F) = V*, and is isomorphic to V.

iii. If $\{e_1, \ldots, e_n\}$ is a basis of V and $\{f_1, \ldots, f_n\}$ is its dual basis, then

$$f = \sum_{i=I}^{n} f(e_i) f_i \; \forall f \in V^* \text{ and } v = \sum_{i=I}^{n} f_i(v) e_i \; \forall v \in V.$$

iv. Every vector space is isomorphic to its second dual.

v. Suppose S E L (V, W) and $T \in L(U, V)$. Then their composition $SoT \in L(U, W)$.

vi. $S \in A(V) = L(V, V)$ is an isomorphism if and only if there exists $T \in A(V)$ such that SoT = I = ToS.

vii. For $T \in A(V)$ there exists a non-zero polynomial $g \in F[x]$, of degree at most n^2, such that g (T) = 0, where dim V = n.

viii. The minimal polynomial of T and f is a polynomial p, of smallest degree such that p(T) = 0.

ix. If p is the minimal polynomial of T and f is a polynomial such that f(T) = 0, then there exists a polynomial g (x) such that f(x) g (x).

x. Let $T \in A(V)$. Then T^{-1} exists if and only if the constant term in the minimal polynomial of T is not zero.

2.7 Solutions/Answers

E1) We have to check that VS1 – VS10 are satisfied by L(U,V). we have already shown that VS1 and VS6 are true.

VS_2: For any L, M, N $\in$ L (U, V), we have $\forall$ u $\in$ U, [(L + M) + N] (u)

= (L + M) (u) + N (u) = [L(u) + M(u)] + N (u)

= L(u) + [M(u) + N (u)], since addition is associative in V.

= [L + (M + N)] (u)

:. (L + M) + N = L + (M + N).

VS3: 0 : U $\rightarrow$ V : 0 (u) = 0 $\forall$ u $\in$ U is the zero element of L (U, V).

VS4: For any S $\in$ L(U,V), (-1) S = - S, is the additive inverse of S.

VS5: Since addition is commutative in V, S + T = T + S $\forall$ S, T in L (U, V).

VS7: $\forall \alpha \in$ F and S, T E L (U, V),

α (S + T) = (αS + αT) (u) $\forall$ u $\in$ U.

:. α (S + T) = αS + αT.

VS8: $\forall$, $\alpha\beta \in$ F and S $\in$ L (U, V), ($\alpha + \beta$) S = αS + βS.

VS9: $\forall \alpha\beta \in$ F and S $\in$ L (U, V), ($\alpha\beta$) S = α (βS).

VS10: VS $\in$ L (U, V), 1. S = S.

E2) E_{2m} (e_m) = f_2 and E_{2m} (ei) = 0 for i $\neq$ m.

E_{32} (ei) f_3 and E_{32} (e_i) = 0 for i = 2.

$$E_{mn}(e_i) = \begin{cases} f_m, \text{ if } i = n \\ 0 \text{ otherwise} \end{cases}$$

E3) Both spaces have dimension 2 over R. A. basis for L (R^2, R) is {E_{11}, E_{12}}, where E_{11} (1, 0) = 1, E_{11} (0, 1) = 0, E_{12} (1, 0) = 0, E_{12}, (0, 1) = 1. A basis for L (R, R^2) is {E_{11}, E_{21}}, where E_{11} (1) = (1, 0), E_{21} (1) = (0, 1).

E4 Let f: $R^3 \rightarrow$ R be any linear functional. Let f (1, 0, 0) = a_1, f (0, 1, 0) = a_2, f(0, 0, 1) = a_3. Then, for any x ⇋ (x_1, x_2, x_3), we have x = x_1 (1, 0,0) + x_2 (0, 1, 0) + x_3 (0, 0, 1).

:. f (x) = x_1 f (1, 0,0) + x_2 f (0, 1, 0) + x_3 f (0, 0, 1)

= $a_1 x_1 + a_2 x_2 + a_3 x_3$.

E5 Let the dual basis be {f_1, f_2, f_3}. Then, for any v $\in$ P_2, v = f_1 (v). 1 + f_2 (v). x + f_3 (v). x^2

:. If v = $a_0 + a_1x + a_2x^2$, then f1 (v) = a_0, f_2 (v) = a_1, f_3 (v) = a_2.

That is, f_1 ($a_0 + a_1x + a_2x^2$) = a_0, f_2 ($a_0 + a_1x + a_2x^2$) = a_1, f_3 ($a_0 + a_1x + a_2x^2$) = a_3, for any $a_0 + a_1 x + a_2 x^2 \in P_2$.

E6) Let $\{f_1,\ldots, f_n\}$ be a basis of V* Let its dual basis be $\{\theta_1,\ldots, \theta_n\}$, $\theta_i \in V^{**}$. Let $e_i \in V$ such that $\varnothing (e_i) = \theta_i$ (ref. Theorem 3) for i = 1,…., n.
Then $\{e_1,\ldots,e_n\}$ is a basis of V, since $\varnothing^{-1}$ is an isomorphism and maps a basis to $\{e_1,\ldots, e_n\}$. Now $f_i (e_j) = \varnothing(e_j) (f_i) = \theta_j (f_i) = \delta_{ji}$, by definition of a dual basis.
:. $\{f_1,\ldots, f_n\}$ is the dual of $\{e_1,\ldots e_n\}$.

E7) For any $S \in A (V)$ and for any $v \in V$,
SoI (v) = S (v) and IoS (v) = I (S(v)) = S (v).
:. SoI = S = IoS.

E8) V $S \in A(V)$ and $v \in V$,
So0 (v) = S (0) = 0, and
0oS (v) = 0 (S (v)) = 0.
:. So0 = 0oS = 0.

E9) $S \in A$ (R2), $T \in A (R^2)$.
SoT $(x_1, x_2) = S (-x_2, x_1) = (x_1, x_2)$
ToS $(x_1, x_2) = T (x_1, - x_1) = x_1, x_2)$
$\forall (x_1, x_2) \in R^2$.
:. SoT = ToS = I, and hence, both S and T are invertible.

E10) $T \in L (R^2, R^3)$, $S \in L (R^3, R^2)$. :. SoT $\in A (R^2)$, ToS $\in A (R^3)$.
:. SoT and ToS can never be equal.
Now SoT $(x_1, x_2) = S (0, x_1, x_2) = (x_1, x_1 + x_2) \forall (x_1, x_2) \in R^2$
Also, ToS $(x_1, x_2, x_3) = T (x_1 + x_2, x_2 + x_3) = (0, x_1 + x_2, x_2 + x_3) \forall (x_1, x_2, x_3) \in R^3$.

E11) Since $T \in A (R^2)$ and $S \in A (R^3)$, SoT and ToS are not defined.

E12) Both (RoS) oT and Ro (SoT) are in L (U, Z). For any $u \in U$,
[(RoS) oT] (u) = (RoS) [T(u)] = R[S(T(u))] = R [SoT) (u)] = [Ro(SoT)] (u).
:. (RoS) oT = Ro (SoT).

E13) By Unit 5, Theorem 6, rank (SoT) ≤ rank (T).
Also, rank (T) = rank (IoT) = rank $((S^{-1}$ oS) oT)
= rank $(S^{-1}$o (SoT)) ≤ rank (SoT) (by Unit 5, Theorem 6).
Thus, rank (SoT) ≤ rank (T) ≤ rank (SoT).
:. Rank (SoT) = rank (T).
Similarly, you can show that rank (ToS) = rank (T).

E14) (S + T) (x, y) = (x, - y) + (x + y, y – x) = (2x + y – x)
ST (x, y) = S (x + y, y- x) = (x + y, x – y)
TS (x, y) = T (x, - y) = (x – y, - (x + y))
[So (S-T)] (x, y) = S (-y, x – 2y) = (- y, 2y – x)

$[(S - T)oS] (x, y) = (S - T) (x, - y) = (x, y) - (x - y, - (x + y)) = (y, 2y + x) \forall (x, y) \in R^2$.

E15) a) W first show that if A, B $\in$ M and α $\beta \in$ F, then $\alpha A + \beta B \in$ M. Now. So $(\alpha A + \beta B)$ = SoαA + SoβB, by Theorem 6.

= α(SoA) + β(SoB), again by Theorem 6

= $\alpha_0 + \beta_0$, since A B $\in$ M.

= 0.

:. $\alpha A + \beta B \in$ M. and M is a subspace of A (V).

Similarly, you can show that N is a subspace of A (V).

b) For any T $\in$ M, ST (v) = 0 $\forall v \in V$. :. T(v) $\in$ Ker S $\forall v \in V$.

:. R (T), the range of T, is a subspace of Ker S.

:. T $\in$ L (V, Ker S). :. M $\subseteq$ L (V, Ker S).

Conversely, for any T $\in$ L (V, Ker S), T $\in$ A (V) such that S (T(v)) = 0 $\forall$ v $\in$ V. :. ST = 0. :. T $\in$ M.

:. L (V, Ker S) $\subseteq$ M.

:. We have proved that M = L (V, Ker S).

:. dim M = (dim V) (nullity S, by Theorem 1.

= n (n – r), by the Rank Nullity Theorem.

E16) (p + q) (T) = p (T) + q (T) = 0 + 0 = 0.

E17) $(2I + 3S + S^3) (S + 2S^4) = 2I + 3S + S^3) S + (2I + 3S + S^3) (2S^4)$

$= 2S + 3S^2 + S^4 + 4S^4 + 6S^5 + 2S^7$

$= 2S + 3S^2 + 5S^4 + 6S^5 + 2S^7$

Also, $(S + 2S^4) (2I + 3S + S^3) = 2S + 3S^2 + 5S^4 + 6S^5 + 2S^7$

:. $(S + 2S^4) (2I + 3S + S^3) = (2I + 3S + S^3) (S + 2S^4)$.

E18) Consider g (x) = x – 1 $\in$ R [x]. Then g (I) = I -1 I = 0.

Also, if h (x) = x, then h (0) = 0.

Notice that the degrees of g and h are both I $\leq$ dim R^3.

E19) Let $p = a_0 + a_1x + \ldots + a_n x^n$, $q = b_0 + b_1x + \ldots + b_m x^m$.

a) Then $ap + bq = aa_0 + aa_1x + \ldots + aa_nx^n + bb_0 + bb_1x + \ldots. + bb_m x^m$.

:.$\phi (ap + bq) = aa_0I + aa_1T + \ldots. + aa_n T^n + bb_0I + bb_1 T + \ldots. + bb_mT^{m.}$

= $ap (T) + bq (T) = a \phi (p) + b \phi (q)$

b) $pq = (a_0 + a_1 x + \ldots + a_n x^n) (b_0 + b_1 x + \ldots + b_m x_m)$

= $a_0b_0 + (a_1 b_0 + a_0 b_1) x + \ldots + a_n b_m x^{n+m}$

:. $\phi (pq) = a_0 b_0I + (a_1 b_0 + a_0 b_1) T + \ldots + a_n b_m T^{n+m}$

= $(a_0I + a_1T + \ldots + a_n T^n) (b_0I + b_1T + \ldots. + b_m T^m)$

= $\phi (p) \phi (q)$.

E20) $T \in A(R^3)$. Let $p(x) = x^3$. Then p is a monic polynomial. Also, $p(T)(x_1, x_2, x_3) = T^3(x_1, x_2, x_3) = T^2(0, x_1, x_2) = T(0, 0, x_1) = (0, 0, 0) \forall (x_1, x_2, x_3) \in R^3$.
:. p(T) = **0**.

We must also show that no monic polynomial q of smaller degree exists such that q (T) = 0.

Suppose $q = a + bx + x^{-}$ and q (T) = 0
Then $(aI + bT + T^2)(x_1, x_2, x_3) = (0, 0, 0)$
$\Leftrightarrow a(x_1, x_2, x_3) + b(0, x_1, x_2) + (0, 0, x_1) = (0, 0, 0)$
$\Leftrightarrow ax_1 = 0, ax_2 + bx_1 = 0, ax_3 + bx_2 + x_1 = 0 \forall (x_1, x_2, x_3) \in R^3$.
$\Leftrightarrow$ a = 0, b = 0 and x1 = 0. But x1 can be non-zero.
:. q does not exist.
:. p is a minimal polynomial of T.

E21) Consider $p(x) = x^n$. Then p (T) = 0 and no non- zero polynomial q of lesser degree exists such that q (T) = 0. this ca be checked on the lines of the solution of E20.

E22) $(T^2 - I)(T - 3I)(x_1, x_2, x_3)$
$= (T^2 - I)((3x_1, x_1 - x_2, 2x_1 + x_2 + x_3) - (3x_1, 3x_2, 3x_3))$
$= (T^2 - I)(0, x_1 - 4x_2, 2x_1 + x_2 - 2x_3)$
$= T(0, -x_1 + 4x_2, 3x_1 - 3x_2 - 2x_3) - (0, x_1 - 4x_2, 2x_1 + x_2 - 2x_3)$
$= (0, x_1 - 4x_2, 2x_1 + x_2 - 2x_3) - (0, x_1 - 4x_2, 2x_1 + x_2 - 2x_3)$
$= (0, 0, 0) \forall (x_1, x_2, x_3) \in R^3$.
:. $(T^2 - I)(T - 3I) = 0$

Suppose $\exists$ $q = a + bx + x^2$ such that q (T) = 0. Then $q(T)(x_1, x_2, x_3) = (0, 0, 0) \forall (x_1, x_2, x_3) \in R^3$. This means that a + 3b + 9 = 0, $(b + 2)x_1 + (a = b + 1)x_2 = 0$, $(2b + 9)x_1 + bx_2 + (a + b + 1)x_3 = 0$. Eliminating a and b, we find that these equations can be solved provided $5x_1 - 2x_2 - 4x_3 = 0$. But they should be true for any $(x_1, x_2, x_3) \in R^3$.
:. The equations can't be solved, and q does not exist. :., the minimal polynomial of T is $(x_2 - I)(x - 3)$.

E23) $D^4(a_0 + a_1x + a_2x^2) D^3(a_1, + 2a_2x) = D^2(2a_2) = D(0) = 0 \forall a_0 + a_1x + a_2x^2 \in P^2$.

:. D4 = 0.

The minimal polynomial of D can be D, D^2, D^3 or D^4. Check that $D^3 = 0$, but $D^2 \neq 0$.
:. The minimal polynomial of D is $p(x) = x^3$. Since p has no non-zero constant term,

D is not an isomorphism.

E24) $T: R^2 \rightarrow$: T(x,y) = (x, - y).

Check that T^2-I = 0

:. The minimal polynomial p must divide x^2 – I.

:. P(x) can be x – 1, x + 1 or x^2 – 1. Since T – 1 ≠ 0 and T + I ≠ 0, we see that p (x_0 = x^2 – 1.

By Theorem 10, T is invertible. Now T^2 – 1 = 0

:. T (-T) = 1, :. T-1 = -T.

E25) Since the minimal polynomial of S is x^n, S^n = 0 and $S^{n-1} \neq 0$. : ∃ $v_0 \in$ V such that

S^{n-1} (v_0) ≠ 0. Let a_1, a_2,….., $a_n \in$ F such that

a_1 v_0 + a_2 S(v_0) + …+ a_n S^{n-1} (0) = 0 …………………….. (1)

Then, applying S^{n-1} to both sides of this equation, we get a_1S^{n-1} (v_0) + …+ a_nS^{2n-1}(v_0) = 0

⇒ a_1S^{n-1} (v_0) = 0, since S^n = 0 S^{n+1} = …. = S^{2n-1}

⇒ a_1 = 0.

Now (1) reduces to a_2S (v_0) + …+ a_nS^{n-1} (v_0) = 0.

Applying S^{n-2} to both sides we get a_2 = 0. In this way we get a_1 = 0 ∀ i = 1, …., n.

:. The set {v_0, S (v_0),…., S^{n-1} (v_0)} is linearly independent.

UNIT 2 MATRICES I

1.0 INTRODUCTION

You have studied linear transformations in Units 1 and 2 we will now study a simple means of representing them, namely, by matrices the plural form of ‘matrix’). We will show that, given a linear transformation, we can obtain a matrix associated to it, and vice versa. Then, as you will see, certain properties of a linear transformation can be studied more easily if we study the associated matrix instead. For example, you will see in Block 3, that it is often easier to obtain the characteristic roots of a matrix than of a linear transformation.

Matrices were introduced by the English Mathematician, Arthur Cayley, in 1858. He came upon this notion in connection with linear substitutions. Matrix theory now occupies an important position in pure as well as applied mathematics. In physics one comes across such terms as matrix mechanics, scattering matrix, spin matrix, annihilation and creation matrices. In economics we have the input-output matrix and the pay off matrix; in statistics we have the transition matrix; and in engineering, the stress matrix, strain matrix, and many other matrices.

Matrices are intimately connected with linear transformations. In this unit we will bring out this link. We will first define matrices and derive algebraic operations on matrices from the corresponding operations on linear transformations. We will also discuss some special types of matrices. One type, a triangular matrix, will be used often in Unit 6. You will also study invertible matrices in some detail, and their connection with change of bases. In Block 2 we will often refer to the material on change of bases so do spend some time on sec 3.6

To realize the deep connection between matrices and linear transformations, you should go back to the exact sport in Units 1 and 2 to which frequent references are made.

This unit may take you a little longer to study, than previous ones, but don't let that worry you. The material in it is actually very simple.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- Define and give examples of various types of matrices;
- Obtain a matrix associated to a given linear transformation
- Define a linear transformation, if you know its associated matrix;
- Evaluate the sum, difference, product and scalar multiples of matrices;
- Obtain the transpose and conjugate of a matrix;
- Determine if a given matrix is invertible;
- Obtain the inverse of a matrix;
- Discuss the effect that the change of basis has on the matrix of a linear transformation.

3.1 Vector Space of Matrices

Consider the following system of three simultaneous equations in four unknowns:

$x - 2y + 4^2 + t = 0$
$x + {}^1/_2 y + 11t = 0$
$3y - 5z = 0$

The coefficients of the unknowns, x, y, z and t can be arranged in rows and columns to form a rectangular array as follows:

1 -2 4 2 (Coefficients of the first equation)

$$\begin{pmatrix} 1 & \frac{1}{2} & 0 & 11 \\ 0 & 3 & -5 & 0 \end{pmatrix} \begin{matrix} \text{(Coefficients of the second equation)} \\ \text{(Coefficients of the third equation)} \end{matrix}$$

Such a rectangular array (or arrangement) of numbers is called a matrix. A matrix is usually enclosed within square brackets [ ] or round brackets ( )

$$\begin{bmatrix} 1 & -2 & 4 & 1 \\ 1 & \frac{1}{2} & 0 & 11 \\ 0 & 3 & -5 & 0 \end{bmatrix} \qquad \begin{pmatrix} 1 & -2 & 4 & 1 \\ 1 & \frac{1}{2} & 0 & 11 \\ 0 & 3 & -5 & 0 \end{pmatrix}$$

The numbers appearing in the various positions of a matrix are called the **entries** (or **elements)** of the matrix. Note that the same number may appear at two or more different positions of a matrix. For example, 1 appears in 3 different positions in the matrix given above.

In the matrix above, the three horizontal rows of entries have 4 elements each. These are called the **rows** of this matrix. The four vertical rows of entries in the matrix, having 3 elements each, are called its **columns.** Thus, this matrix has three rows and four columns. We describe this by saying that this is a matrix of size 3 x 4 (“3 by 4” or “3 cross 4”), or that this is a 3 x 4 matrix. The rows are counted from top to bottom and the columns are counted from left to right. Thus, the first row is (1, -2, 4, 1), the second row is (1, ½, 0, 11), and so on. Similarly,

The first column is $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, the second column is $\begin{pmatrix} -2 \\ \frac{1}{2} \\ 3 \end{pmatrix}$, and so on.

Note that each row is a 1 x 4 matrix and each column is a 3 x 1 matrix,

We will now define a matrix of any size.

3.1.1 Definition of a Matrix

Let us see what we mean by a matrix of size m x n, where m and n are any two natural numbers.

Let F be a field.
A rectangular array

A_{11} a_{12}………… A_{1n}

A_{21} a_{22} ………… a_{2n}
……………………………
……………………………
A_{m1} a_{m2}………… a_{mn}

Of mn elements of F arranged in m rows and n columns is called a **matrix of size m x n;** or an **m x n matrix,** over **F.** You must remember that the mn entries need not be distinct.

The element at the intersection of the ith row and the jth column is called the (i,j) th elements. For example, in the m x n matrix above, the (2, n) the element is a2n, which is the intersection of the 2^{nd} row and the nth column.

A brief notation for this matrix is $[a_{ij}]$ mn, or simply $[a_{ij}]$, if m and n need not be stressed. We also denote matrices by capital letters A, B, C, … etc. The set of all m x n matrices over **F** is denoted by $\mathbf{M}_{mxn}$ (**F**).

Thus, $[1,\sqrt{2}\,] \in \mathbf{M}_{1x2}(\mathbf{R})$.

If m = n, then the matrix is called a square matrix. The set of all n x n matrices over F is denoted by $\mathbf{M}_n$ (**F**).

In an m x n matrix each row is a 1 x n matrix and is also called a row vector. Similarly, each column is an m x 1 matrix and is also called a **column vector**.

Let us look at a situation in which a matrix can arise.

Example 1: There are 20 male and 5 female students in the B.Sc. (Math. Hons) 1 year class in a certain college, 15 male and 10 female student's in B.Sc. (Math. Hon's) II year and 12 male and 10 female students in B.Sc. (Math. Hon's) III year. How does this information give rise to a matrix?

Solution: One of the ways in which we can arrange this information in the form of a matrix is as follows:

B.Sc. I B.Sc. II B.Sc. III

$$\begin{array}{l} \text{Male} \\ \text{Female} \end{array} \begin{pmatrix} 20 & 15 & 12 \\ 5 & 10 & 10 \end{pmatrix}$$

This is a 2 x 3 matrix.

Another way could be the 3 x 2 matrix.

Female Male

B.Sc. I
B.Sc. II
B.Sc. II

$$\begin{pmatrix} 5 & 20 \\ 10 & 15 \\ 10 & 12 \end{pmatrix}$$

Either of these matrix representations immediately shows us how many male/female students there are in any class.

To get used to matrices and their elements, you can try the following exercises.

E Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 5 & 3 & 2 \\ 5 & 4 & 1 & 5 \\ 0 & 3 & 2 & 0 \end{pmatrix}$

Give the

a) (1, 2)th elements of A and B.
b) Third row of A.
c) Second column of A and the first column of B.
d) Forth row of B.

E E2) Write two different 4 x 2 matrices.

How did you solve E 2? Did the (I, j) th entry of one differ from the (I, j) th entry of the other for some I and j? if not, then they were equal. For example, the two 1 x 1 matrices [2] and [2] are equal. But [2] $\neq$ [3], since their entries at the (1, 1) position differ.

Definition: Two matrices are said to be **equal** if

i. They have the same size, that is, they have the same number of rows as well as the same number of columns, and
ii. Their elements, at all the corresponding positions, are the same.

The following example will clarify what we mean by equal matrices.

Example 2: if $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} x & y \\ z & 3 \end{pmatrix}$, then what are x, y and z?

Solution: Firstly both matrices are of the same size, namely, 2 x 2, Now, for these matrices to be equal the (I j) th elements of both must be equal $\forall$ I, j. therefore, we must have x = I, y = 0, z = 2.

E E3) Are [1] and $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ equal? Why?

Now that you are familiar with the concept of a matrix, we will link it up with linear transformations.

3.1.2 Matrix of a Linear Transformation

We will now obtain a matrix that corresponds to a given linear transformation. You will see how easy it is to go from matrices to linear transformations, and back. Let U and V be vector spaces over a field **F,** of dimensions n and m, respectively. Let

B_1 = $\{e_1,\ldots..e_n\}$ be an ordered basis of U, and
B_2 = $\{f_1,\ldots.., f_m\}$ be an ordered basis of V, (By an ordered basis

We mean that the order in which the elements of the basis are written is fixed. Thus, an ordered basis $\{e_1, e_2\}$ is not equal to an ordered basis $\{e_2, e_1\}$).

Given a linear transformation T:U → V, we will associate a matrix to it. For this, we consider T (e_1), ……, T (e_n), which are all elements of V and hence, they are linear combinations of $f_1,\ldots.., f_m$. Thus, there exist mn scalars α_{ij}, such that

$T(e_1) = \alpha_{11} f_1 + \alpha_{21} f_2 + \ldots\ldots + \alpha_{m1} f_m$

$T(e_1) = \alpha_{ij} f_1 + \alpha_{21} f_2 + \ldots\ldots\ldots + a_{mj} f_m$

$T(e_n) = \alpha_{ij} f_1 + \alpha_{2n} f_2 + \ldots\ldots\ldots + \alpha_{mn} f_m$.

From these n equations we form an m x n matrix whose first column consists of the coefficients of the first equation, second column consists of the coefficients of the second equation, and so on. This matrix.

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & .. & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & .. & \alpha_{2n} \\ .. & .. & .. & .. \\ \alpha_{m1} & \alpha_{m2} & .. & \alpha_{mn} \end{pmatrix}$$

is called the **matrix of T with respect to the bases B_1 and B_2. Notice that the coordinate vector of T (e_j) is the jth column of A.**

We use the notation $[T]_{B1}$, B_2 for this matrix. Thus, to obtain $[T]B_1$ B_2 we consider T (e_i) $\forall$ $e_i \in B_1$, and write them as linear combinations of the elements of B_2.

If T $\in$ L (V, V), B is a basis of V and we take $B_1 = B_2 = B$, then $[T]_{B,\ B}$ is called the matrix of T with respect to the basis B, and can also be written as $[T]_B$.

Remark: Why do we insist on order bases? What happens if we interchange the order of the elements in B= to $\{e_n, e_1, \ldots, e_{n-1}\}$? The matrix $[T]B_1$, B_2 also changes, the last column becoming the first column now. Similarly, if we change the positions of the f_1's in B_2, the rows of $[T]_{B1,\ B2}$ will get interchanged.

Thus, to obtain a unique matrix corresponding to T, we must insist on B_1 and B_2 being ordered bases. Henceforth, while discussing **the matrix of a linear mapping, we will always assume that our bases are ordered bases**.

We will now give an example, followed by some exercises.

Example 3: Consider the linear operator
T: $R^3 \rightarrow$: T (x, y, z) = (x, y). Choose bases B_1 and B_2 of R^3 and R^2, respectively. Then obtain $[T]B_1$. B_2.

Solution: Let $B_1 = \{e_1, e_2, e_3\}$, where $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Let $B_2 = \{f_1, f_2\}$, where $f_1 = \{1, 0)$, $f_2 = \{0, 1)$, Note that B_1 and B_2 are the standard bases of R^3 and R^2, respectively.

$T(e_1) = (1,0) = f_1 = 1.\ f_1 + 0.f_2$
$T(e_2) = (0, 1) = f_2 = 0.f_1 + 1.f_2$
$T(e_3) = (0, 0) = 0f_1 + 0f_2.$

Thus, $[T]B_1.B_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

E E4) Choose two other bases B1 and B`2 of R3 and R2, respectively. (In Unit 4 you came across a lot of bases of both these vector spaces) For T in the example above, give the matrix $[T]B^{`}_1B^{`}_2$

What E4 shows us is that the matrix of a transformation depends on the bases that we use for obtaining it. The next two exercises also being out the same fact.

E E 5) Write the matrix of the linear transformation $T: R^3 \rightarrow R^2: T(x, y, z) = (x + 2y + 2z, 2x + 3y + 4z)$ with respect to the standard bases of R^3 and R^2.

E E 6) What is the matrix of T, in E5, with respect to the bases
B'_1 = {(1, 0, 0), (0, 1, 0), (1, -2, 1)} and
B'_2 = {(1, 2), (2, 3)}?

The next exercise is about an operator that you have come across often

E E7) Let V be the vector space of polynomials over R of degree < 3, in the variable t. let D:V → V be the differential operator given in Unit 5 (E6, when n = 3). Show that the matrix of D with respect to the basis $\{1, t, t^2, t^3\}$ is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

So far, given a linear transformation, we have obtained a matrix from it. This works the other way also. That is given a matrix we can define a linear transformation corresponding to it.

Example 4: Describe T: $R^3 \rightarrow R^3$ such that

$$[T]_B = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix},$$ where B is the standard basis of R^3.

Solution: Let B = $\{e_1, e_2, e_3)$, Now, we are given that

$T(e1) = 1.e_1 + 2.e_2 + 3.e_3$
$T(e2) = 2.e_1 + 3.e_2 + 1.e_3$
$T(e3) = 4.e_1 + 1.\ e_2 + 2.e_3$

You know that any element of R^3 is $(x, y, z) = xe_1 + ye_2 + ze_3)$

Therefore, $T(x, y, z) = T(xe_1 + ye_2 + ze_3)$

$= xT(e_1) + yT(e_2) + zT(e_3)$, since T is linear.
$= x(e_1 + 2e_2 + 3e_3) + y(2e_1 + 3e_2 + e_3) + z(4e_1 + e_2 + 2e_3)$
$= (x + 2y + 4z)e_1 + (2x + 3y + z)e_2 + (3x + y + 2z)e_3$.
$= (x + 2y + 4z, 2x + 3y + z, 3x + y + 2z)$

:. T: $R^3 \rightarrow R^3$ is defined by T (x, y, z) = (x + 2y + 4z, 2x + 3y + z, 3x + y + 2z)

Try the following exercises now.

E E8) Describe T: $R^3 \rightarrow R^2$ such that

$$[T]B_1.\ B_2 = \left(\begin{array}{ccc} & 1 & 1 \\ 0 & 1 & 1 \end{array}\right)\begin{array}{c} 0 \\ \ \end{array},$$ where B_1 and B_2 are the standard

bases of R^3 and R^2, respectively.

E E9) Find the linear operator T: C $\rightarrow$ C whose matrix, with respect to the basis {1, i} is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

(Note that C, the field of complex numbers, is a vector space over R, of dimension 3)

Now we are in a position to define the sum of matrices and multiplication of a matrix by a scalar.

3.1.3 Sum and Multiplication by Scalars

In Unit 5 you studied about the sum and scalar multiples of linear transformations. In the following theorem we will see what happens to the matrices associated with the linear transformations that are sums or scalar multiples of given linear transformations.

Theorem 1: Let U and V be vector spaces over F, of dimensions n and m, respectively. Let B_1 and B_2 be arbitrary bases of u and V, respectively. (Let us abbreviate $[T]B_1. B_2$ to [T] during this theorem.) Let S, T $\in$ L (U, V) and $\alpha \in F$. Suppose $[S] = [a_{ij}]$, $[T] = [b_{ij}]$. Then

$$[S + T] = [a_{ij} + b_{ij}], \text{ and}$$
$$[\alpha S] = [\alpha a_{ij}]$$

Proof: Suppose $B_1 = \{e_1, e_2, \ldots\ldots, e_n\}$ and $B_2 = \{f_1, f_2, \ldots.., f_m\}$. Then all the matrices to be considered here will be of size m x n.

Now, by our hypothesis,

$$S(e_j) = \sum_{i=1}^{m} a_{ij} f_i \; \forall \; j = 1,\ldots, n \text{ and}$$

$$T(e_j) = \sum_{i=1}^{m} b_{ij} f_i \; \forall \; j = 1,\ldots, n \text{ and}$$

:., $(S + T)(e_i) = S(e_j) + T(e_i)$ (by definition of S + T)

$$= \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} b_{ij} f_i$$

$$= \sum_{i=1}^{m} (a_{ij} + b_{ij}) f_i.$$

Thus, by definition of the matrix with respect to B_1 and B_2, we get $[S + T] = a_{ij} + b_{ij}]$.

Now, $(\alpha S)(e_j) = \alpha(S(e_j))$ (by definition of αS)

$$= \alpha \left(\sum_{i=1}^{m} a_{ij} f_i \right)$$

Two matrices can be added if and only if they are of the same size

$$= \sum_{i=1}^{m} (\alpha a_{ij}) f_i$$

Thus, $[\alpha S] = [\alpha a_{ij}]$

Theorem 1 motivates us to define the sum of 2 matrices in the following way.

Definition: Let A and B be the following two m x n matrices.

$$A = \begin{pmatrix} a_{11} & a_{12} & . \; . & a_{1n} \\ a_{21} & a_{22} & . \; . & a_{2n} \\ : & . & . \; . & . \\ : & . & . \; . & . \\ a_{m1} & a_{m2} & . \; . & a_{mn} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & b_{12} & . \; . & b_{1n} \\ b_{21} & b_{12} & . \; . & b_{2n} \\ . & . & . \; . & . \\ . & . & . \; . & . \\ b_{m1} & b_{m2} & . \; . & b_{mn} \end{pmatrix}.$$

Then the sum of A and B is defined to be the matrix

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & . \; . & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & . \; . & a_{2n} + b_{2n} \\ . \; . & . \; . & . \; . & . \; . \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & . \; . & a_{mn} + b_{nm} \end{pmatrix}.$$

In other words, A + B is the m x n matrix whose (i, j)th element is the sum of the (i, j)th element of A and the (i, j) th element of B.

Let us see an example of how two matrices are added.

$$1 \quad \begin{pmatrix} 4 & 5 \end{pmatrix} 0 \quad \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Example 5: What is the sum of 0 1 0 and 1 4 5 ?

Solution: Firstly, notice that both the matrices are of the same size (otherwise, we can't add them). Their sum is

$$\begin{pmatrix} 1+0 & 4+1 & 5+0 \\ 0+1 & 1+4 & 0+5 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 5 \\ 1 & 5 & 5 \end{pmatrix}$$

E E10) What is the sum of

a) $\begin{pmatrix} 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$?

b) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$?

Now, let us define the scalar multiple of a matrix, again motivated by theorem 1.

Definition: let a be a scalar, i.e., $\alpha \in F$, and Let $A = [a_{ij}]_{mxn}$, Then we define the **scalar multiple of the matrix A by the scalar α to be the matrix.**

$$\begin{pmatrix} \alpha a_{11}. & \alpha a_{12} & . & . & \alpha a_{1n} \\ \alpha a_{21} & aa_{22} & . & . & \alpha a_{2n} \\ . & . & . & . & . \end{pmatrix}$$

$\alpha a_{m1} \quad \alpha a_{m2} \quad . \quad . \quad \alpha a_{mn}$

In other words, αA is the m x n matrix whose (i,i) th element is a times the (i,j) th element of A.

Example 6: What is 2A, where $\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ 0 & 0 & 0 \end{pmatrix}$?

Solution: We must multiply each entry of A by 2 to get 2A.

Thus,

$$2A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{2}{3} \\ 0 & 0 & 0 \end{pmatrix}$$

E E11) Calculate $3\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $3\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1. \end{bmatrix}\right)$

Remark: The way we have defined the sum and scalar multiple of matrices allows us to write Theorem 1 as follows:

$[S + T]B_1.B_2 \quad = \quad [S]B_1.B_2 + [T]B_1.B_2.$

$[\alpha S]B_1. B_2 \quad = \quad \alpha[S]B_1.B_2.$

The following exercise will help you in checking if you have understood the contents of sections 7.2.2 and 7.2.3.

E E12) Define $S: R^2 \rightarrow R^3$: S (x, y) = (x, 0,y) and $T: R^2 \rightarrow R^3$: T(x,y). Let
B^1 and B^2 be the standard bases for R^2 and R^3 respectively.

Then what are $[S]B_1.B_2$, $[T]B_1.B_2$, $[S + T]B_1. B_2$, for any $\alpha \in R$.

We now want to show that the set of all m x n matrices over F is actually a vector space over F.

3.1.4 M_{mxn} (F) is a Vector Space

After having defined the sum and scalar multiplication of matrices, we enumerate the properties of these operations. This will ultimately lead us to prove that the set of all m x n matrices over F is a vector space over F. Do keep the properties VS1 – VS10 (of Unit 3) in mind. For any A = [aij], B = $[b_{ij}]$, C = $[c_{ij}] \in M_{mxn}$ (F) and $\alpha\beta \in F$, we have

i) Matrix addition is associative:

(A + B) + C = A + (B + C), since

$(a_{ij} + b_{ij} + c_{ij} = a_{ij} + (b_{ij} + c_{ij}) \forall$ i, j, as they are element of a field.

ii) Additive identity: the matrix of the zero transformation (see unit 5), with respect to any basis, will have 0 as all its entries. This is called the zero matrix. Consider the zero matrix 0, of size m x n. then, for any $A \in M_{mxn}$ (F).

A + 0 = 0 + A = A,

Since $a_{ij} + 0 = 0 + a_{ij} = a_{ij} \forall$ i, j.

Thus, 0 is the additive identity for M_{mxn} (F).

iii) Additive inverse: Given $A \in M_{mxn}$ (F) we consider the matrix (-1) A. then
A + (-1) A = (-1) A + A = 0

This is because the (i,j)th element of (-1) A is $-a_{ij}$, and $a_{ij} + (-a_{ij}) = 0 = (-a_{ij}) + a_{ij} \forall$ i.j.
Thus, (-1) A is the additive inverse of A. We denote (-1) A by – A.

iv) Matrix addition is commutative:

$A + B = B + A$
This is true because $a_{ij} + b_{ij} + a_{ij} \forall$ i. j.

v) $\alpha(A + B) = \alpha A + \alpha B$.

vi) $(\alpha + \beta) A = \alpha A + \beta A$

vii) $(\alpha\beta) A = \alpha(\beta A)$

viii) 1. A = A

E E13) Write out the formal profs of the properties (v) – (viii) given above.

These eight properties imply that M_{mxn} (F) is a vector space over F

Now that we have shown that M_{mxn} (F) is a vector space over F, we know it must have a dimension.

3.1.5 Dimension of M_{mxn} (F) over F

What is the dimension of Mmxn (F) over F? to answer this question we prove the following theorem. But, before you go further, check whether you remember the definition of a vector space isomorphism (Unit 5).

Theorem 2: Let U and V be vector spaces over F of dimensions n and m., respectively. Let B_1 and B_2 be a pair of bases of u and V, respectively. The mapping ϕ: L(U, V) → M_{mxn} (F), given by ϕ (T) = $[T]B_1$. B_2 is a vector space isomorphism.

Proof: The fact that ϕ is a linear transformation follows from Theorem 1. We proceed to show that the map is also 1-1 and onto. For the rest of the proof we shall denote [S] B_1. B_2 by [S] only. And take $B_1 = \{e_1,\ldots,e_n\}$, $B_2 = \{f_1, f_2,\ldots, f_m\}$

ϕ is 1 -1 : Suppose S, T $\in$ L (U, V) be such that ϕ (S) = ϕ (T).

Then [S] = [T]. Therefore, $S(e_j) = T(e_j) \forall e_j \in B_1$.

Thus, by Unit 5 (Theorem, 1), we have S = t.

ϕ is on 0 : if $A \in M_{mxn}$ (F) we want to construct $T \in L$ (U, V)

Such that ϕ (T) = A. Suppose $A = [a_{ij}]$. Let $v_1, \ldots. v_n \in V$ such that

$$v_1 = \sum_{i=1}^{m} a_{ij} f_i \text{ for } j = 1, \ldots, n.$$

Then, by Theorem 3 of Unit 5, there exists a linear transformation $T \in L$ (U, V) such that

$$T(ej) = vj = \sum_{i=1}^{m} a_{ij} f_i.$$

Thus, by definition, ϕ (T) = A.

Therefore, ϕ is a vector space isomorphism.

A corollary to this theorem gives us the dimension of M_{mxn} (F).

Corollary: Dimension of M_{mxn} (F) = mn.

Proof: Theorem 2 tells us the M_{mxn} (F) is isomorphic to L (U, V). Therefore, $\dim_F M_{mxn}$ (F) = $\dim_F$ (L(U, V) (by Theorem 12 of Unit 5) = mn, from Unit 6 (Theorem 1).

Why do you think we chose such a roundabout way for obtaining dim M_{mxn} (F)? We could as well have tried to obtain mn linearly independent m x n matrices and show that they generate M_{mxn} (F). But that would be quite tedious (see E16). Also, we have done so much work on L(U, V) so why not use that! And, doesn't the way we have used seem neat?

Now for some exercises related to Theorem 2.

E E 14) At most, how many matrices can there be in any linearly independent subject of M_{2x3} (F)?

E E15) Are the matrices [1, 0] and [1, -1] linearly independent over R?

E E16) Let E_{ij} be an m x n matrix whose (i, j)th element is 1 and the other elements are 0. Show that $\{E_{ij}: \leq i \leq m. 1 \leq j \leq h\}$is a basis of M_{mxn} (F) over F. Conclude that $\dim_F M_{mxn}$ (F) = mn.

Now we move on to the next section, where we see some ways of getting new matrices from given ones.

3.3 New Matrices From Old

Given any matrix we can obtain new matrices from them in different ways. Let us see three of these ways.

3.2.1 Transpose

Suppose $A = \begin{pmatrix} 1 & 0 & 9 \\ & 2 & 5 \end{pmatrix} \quad 9$

From this we for a matrix whose first and second columns are the first and second rows of A, respectively. That is, we obtain.

$$B = \begin{pmatrix} 1 & 2 \\ 0 & 5 \\ 9 & 9 \end{pmatrix}$$

Then B is called the transpose of A. Note that A is also the transpose of B, since the rows of B are the columns of A. here A is a 2 x 3 matrix and B is a 3 x 2 matrix.

In general, if A = $[a_{ij}]$ is an m x n matrix. Then the n x m matrix whose ith column is the ith row of A, is called the transpose of A. the transpose of a A is denoted by A^t (the notation and A' is also widely used.)

Note that, if A = $[a_{ij}]_{mxn}$, then At = $[b_{ij}]_{mxn}$ where b_{ij} is the intersection of the ith row and the jth column of At. :. b_{ij} is the intersection of the jth row and ith column of A, i.e.., a_{ji}. :. $b_{ij} = a_{ji}$.

1

E E17) Find At, where a = $\begin{pmatrix} 2 & \\ 2 & 0 \end{pmatrix}$.

We now given theorem that lists some properties of the transpose.

Theorem 3: Let A, B $\in M_{mxn}$ (F) and $\alpha \in$ F. Then,

a) $(A + B)^t = A^t + B^t$

b) $(\alpha A)^t = \alpha A^t$.

c) $(A^t)^t = A$

Proof a) Let A = $[a_{ij}]$ and B = $[b_{ij}]$. Then A + B = $[a_{ij} + b_{ij}]$.
Therefore, (A + B)t = $[c_{ij}]$, where

c_{ij} = the (j, i)th element of A + B = $a_{ij} + b_{ji}$.
= sum of the (j, i)th elements of A and B
= sum of the (i, j)th elements of A^t and B^t.
= (i, j)th element of $A^t + B^t$.

Thus, $\mathbf{(A + B)^t - A^t + B^t}$.

We leave you to complete the proof of this theorem. In fact that is what E 18 says!

E E18) Prove (b) and (c) of Theorem 3.

E E19) Show that, if $A = A^t$, then A must be a square matrix.

E 19 leads us to some definitions.

Definitions: A square matrix A such that $A^t = A$ is called a **symmetric matrix**. A square matrix A such that $A^t = -A$, is called **a skew–symmetric matrix.** For example, the matrix in E17, and

$\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}$ 2
are both symmetric matrices.

0 $\begin{pmatrix} 2 & \\ & -2 \end{pmatrix}$ 0 is an example of a skew-symmetric matrix since

$$\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}^t = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$$

E E20) Take a 2 x 2 matrix A. Calculate $A + A^t$ and $A – A^t$. Which of these is symmetric and which is skew-symmetric?

What you have shown in E20 is true for a square matrix of any size, namely, for any A $\in M_n$ (F), $A + A^t$ is symmetric and $A – A^t$ is skew-symmetric.

We now given another ways of getting a new matrix from a given matrix over the complex field.

3.2.2 Conjugate

If A is a matrix over C, then the matrix obtained by replacing each entry of A by its complex conjugate is called the conjugate of A, and is denoted by Ā.

Three properties of conjugates, which are similar to those of the transpose, are

a) A + B = Ā + □, for A, B $\in M_{mxn\ ά}$(C).
b) αA = ά Ā, for a $\in$ C and A $\in M_{mxn}$ (C)
c) $\tilde{\bar{A}}$ A = A, for A $\in M_{mxn}$ (C)

Let us see an example of obtaining the conjugate of a matrix.

Example 7: Find the conjugate of
$$\begin{pmatrix} 1 & : \\ 2+i & -3-2i \end{pmatrix}$$

Solution: By definition, the required matrix will be

$$\begin{pmatrix} 1 & -1 \\ 2-i & -3+2i \end{pmatrix}$$

Example 8: What is the conjugate of $\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$?

Solution: Note that this matrix has only real entries. Thus, the complex conjugate of each entry is itself. This means that the conjugate of this matrix is itself.

This example leads us to make the following observation.

Remark: Ā = Ā if and only if A is a real matrix.

Try the following exercise now.

E E21) Calculate the conjugate of $\begin{pmatrix} i & 2 & \\ & 3 & i \end{pmatrix}$.

We combine what we have learnt in the previous two sub-section now.

3.2.3 Conjugate Transpose

Given a matrix A ∈ M_{mxn}(F) we form a matrix B by taking the conjugate of At. Then B = $\bar{A}^t$, is called the **Conjugate transpose of A.**

Example 9: Find $\bar{A}^t$ where A $\begin{pmatrix} 1 & i \\ 2 + i & -3 - 2i. \end{pmatrix}$

Solution: Firstly, $A^t = \begin{pmatrix} 1 & 2 + i \\ i & -3 - 2i, \end{pmatrix}$ Then

$$\bar{A}^t = \begin{pmatrix} 1 & 2 - i \\ -1 & -3 + 2i. \end{pmatrix}$$

Now, note a peculiar occurrence. If we first calculate Ā and then take its transpose, we get the same matrix, namely, $\bar{A}^t$. That is, $(\bar{A})^t = \bar{A}^t$. In general, $(\bar{A})^t$- $\bar{A}^t$ ∀ A ∈ M_{mxn} (C),

E E22) Show that $A = \bar{A}^t \Rightarrow A$ is a square matrix.

E22 leads us to the following definitions.

Definitions: A square matrix A for which $\bar{A}^t = A$ is called a **Hermitian matrix.** A square matrix A is called a **Skew-Hermitian matrix** if $\bar{A}^t = -A$.

For example, the matrix $\begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}$ is Hermitian, whereas the

Matrix $\begin{pmatrix} i & 1+i \\ -1+I & 0 \end{pmatrix}$ is a skew-Harmitian matrix.

Note: If $A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$ Then $A = A^t = \bar{A}^t$ (Since the entries are all real).

:. A is symmetric as well as Hermitian. In fact, for a real matrix A, A is Hermitian if A is symmetric. Similarly, A is skew-Harmitian if A is skew-symmetric.

We will now discuss two important and often-used, types of square matrices.

3.3 Some Types of Matrices

In this section we will define a diagonal matrix and a triangular matrix.

3.3.1 Diagonal Matrix

Let U and V be vector spaces over F of dimension n. Let $B_1 = \{e_1, \ldots .. e_n\}$ and $B_2 = \{f_i, \ldots \ldots ., f_n\}$ be based of U and V, respectively. Let $d_1, \ldots \ldots, d_n \in F$. Consider the transformation $T:U \rightarrow V:T(a_1 e_1 + \ldots + a_n e_n) = a_1 d_1 f_1 + \ldots + a_n d_n f_n$

Then $T(e_1) = d_1 f_1, T(e_2) = d_2 f_2, \ldots ., T(e_n) = d_n f_n$.

$$:.[T]B_1.B_2 = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ & . & . & \ldots \end{pmatrix}$$

$0 \quad 0 \quad \ldots \quad d_n.$

Such a matrix is called a diagonal matrix. Let us se what this means.

Let A $[a_{ij}]$ be a square matrix. The entries $a_{11}, a_{22}, \ldots, a_{mn}$ are called the **diagonal entries** of A. This is because they lie along the diagonal, from left to right, of the matrix. All the other entries of A are called the **off-diagonal entries** of A.

A square matrix whose off-diagonal entries are zero (i.e., $a_{ij} = 0 \ \forall \ i \neq j$) is called a **diagonal matrix**. The diagonal matrix.

$$\begin{pmatrix} d_1 & 0 & 0 & .\ . & 0 \\ 0 & d_2 & 0 & .\ . & 0 \\ : & : & : & :\ : & : \\ 0 & 0 & 0 & .\ . & d_n. \end{pmatrix}$$

Is denoted by **diag** $(d_1, d_2, \ldots\ldots\ldots, d_n)$.

Note: The d_i's may or may not be zero. What happens if all the d_i's are zero? Well, we get the n x n zero matrix, which corresponds to the zero operator.
If $d_i = 1 \ \forall \ I = 1, \ldots\ldots\ldots\ldots, n$, we get the identity matrix, I_n (or I, when the size is understood).

E E23) Show that I, is the matrix associated to the identity operator from R^n to R^n.

if $\alpha \in F$, the linear operator $\alpha I: R^n \rightarrow R^n$: $\alpha I(v) = \alpha v$, for all $v \in R^n$, is called a **Scalar operator.** Its matrix with respect to any basis is $\alpha I = \text{diag}\ (\alpha, \alpha, \ldots\ldots\ldots, \alpha)$. Such a matrix is called a **Scalar matrix.** It is a diagonal matrix whose diagonal entries are all equal. With this much discussion on diagonal matrices, we move onto describe triangular matrices.

3.3.2 Triangular Matrix

Let $B = \{e_1, e_2, \ldots\ldots, e_n\}$ be a basis of a vector space V. let $S \in L$ (V, V) be an operator such that

$S(e_1) = a_{11}e_1.$

$S(e_2) = a_{12}e_1 + a_{22}e_2$

$\vdots \quad \vdots \quad \vdots$

$S(e_n) = a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n,$

Then, the matrix of S with respect to B is

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdot\ \cdot & a_{1n} \\ 0 & a_{22} & \ldots & \cdot\ \cdot & a_{2n} \\ \vdots & \vdots & \ldots & \cdot\ \cdot & \vdots \\ 0 & 0 & 0\ . & \cdot\ \cdot & a_{nn} \end{pmatrix}.$$

Note that $a_{ij} = 0 \ \forall\ i > j$.

A square matrix A such that $a_{ij} = 0 \ \forall\ i > j$ is called an **Upper Triangular Matrix.** If $a_{ij} = 0 \ \forall\ i \geq j$, then A is called **Strictly Upper Triangular.**

For example $\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are all upper triangular,

while $\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$ is strictly upper triangular.

Note that every strictly upper triangular matrix is an upper triangular matrix.

Now let $T : V \rightarrow V$ be an operator such that $T(e_j)$ is a linear combination of e_j, $e_{j+1}, \ldots, e_n \mid \forall_j$ The matrix of T with respect to B is

$$[T]B = \begin{pmatrix} b_{11} & 0 & 0 & \cdot\ \cdot & 0 \\ b_{21} & b_{22} & \vdots & \cdot\ \cdot & 0 \\ \vdots & \vdots & \vdots & \vdots\ \vdots & \vdots \end{pmatrix}$$

$b_{n1} \quad b_{n2} \quad b_{n3} \quad . \; . \quad b_{nn}.$

Note that $b_{ij} = 0 \; \forall \; i < j$

Such a matrix is called a **Lower Triangular Matrix.** If $b_{ij} = 0$ for all $I \leq j$, then B is said to be a strictly **Lower Triangular Matrix**.

The matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 5 & 0 \end{pmatrix}$$

is a strictly lower triangular matrix. Of course, it is also lower triangular!

Remark: If A is an upper triangular 3 x 3 matrix, say

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \text{ and } A^t = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{pmatrix} \text{ a lower triangular matrix}$$

In fact, for any n x n upper triangular matrix A, its transpose is lower triangular, and vice versa.

E E24) If an upper triangular matrix A is symmetric, then show that it must be a diagonal matrix.

E E25) Show that the diagonal entries of a skew-symmetric matrix are all zero, but the converse is not true.

Let us now se how to define the product of two or more matrices.

3.4 Matrix Multiplication

We have already discussed scalar multiplication. Now we see how to multiply two matrices. Again, the motivation for this operation comes from linear transformations.

3.4.1 Matrix of the Composition of Linear Transformations

Let U, V and W be vector spaces over F, of dimension p, n and m, respectively, Let B_1, B_2 and B_3 be bases of these respective spaces. Let $T \in L(U, V)$ and $S \in L(V, W)$. Then ST (= S T) $\in$ L (U, W) (see sec 6.4).

Suppose $[T]B_1.B_2 = B = [b_{jk}]_{n \times p}$

and $[S]B_2,B_3 = A = [a_{ij}]_{mxn}$

We ask: What is the matrix $[ST]B_1.B_3$?

To answer this we suppose

$B_1 = \{e_1, e_2 \ldots\ldots\ldots, e_n\}$
$B_2 = \{f_1, f_2, \ldots\ldots\ldots, f_n\}$
$B_3 = \{g_1, g_2, \ldots\ldots\ldots, g_m\}$

Then, we know that $T(e_k) = \sum_{j=1}^{n} b_{jk} f_j \forall k = 1, 2, \ldots\ldots, p,$

and $S(f_j) \quad \sum_{i=1}^{m} a_{ij} g_i \forall j = 1, 2, \ldots.., n.$

n

Therefore, $SoT(e_k) = S(T(e_k)) = S \sum_{i=1} b_{jk} f_j = b_{jk} S(f_1) + b_{2k} S(f_2) + \ldots + b_{nk} S(f_n)$

$$= b_{1k} \sum_{i=1}^{m} a_{ij} g_i + b_{2k} \sum_{i=1}^{m} a_{12} g_i + \ldots + b_{nk} \sum_{i=1}^{m} a_{in} g_i$$

$$= \sum_{i=1}^{m} (a_{ij} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk}) g_i,$$ on collecting the

coefficients of g_i.

Thus, $[ST]B_1. B_3 = [c_{ik}]_{mxp}$, where $c_{ik} = \sum_{i=1}^{m} a_{ij} b_{jk}$

We define the matrix $[c_{ik}]$ to be the product AB.

So, let us see how we obtain AB from A and B

Let $a = [a_{ij}]_{mxn}$ $B = [b_{ik}]_{mxn}$ be two matrices over F of sizes m x n and n x p, respectively. We define AB to be the m x p matrix C whose (i,k) th entry is

$$c_{ik} = \sum_{i=1}^{n} a_{ij} b_{jk} = a_{ij} b_{1k} + a_{12} b_{2k} + \ldots + a_{in} b_{nk}.$$

In order to obtain the (I,k) th element of AB, take the ith row. Of A and the kth column of B are both n-tuples. Multiply their corresponding elements and add up all these products.

For example, if the 2nd row of A = [1 2 3], and the 3rd column of

$$B = \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix}$$, then the (2,3) entry of AB = 1x4 + 2x5 + 3x 6 = 32

Note that two matrices A and B can only be multiplied if the number of columns of A = the number of rows of B. the following illustration may help in explaining what we do to obtain the product of two matrices.

$$
A \qquad\qquad\qquad\qquad\qquad B
$$

$$
\begin{pmatrix}
a_{11} & a_{12} & .. & a_{1n} \\
a_{21} & a_{22} & .. & a_{2n} \\
. & . & .. & .. \\
a_{11} & a_{12} & .. & a_{in} \\
. & . & .. & .. \\
a_{m1} & a_{m2} & .. & a_{mn}
\end{pmatrix}
\begin{pmatrix}
b_{11} & b_{12} & .. & b_{ik} & .. & b_{1p} \\
b_{21} & b_{22} & .. & b_{2k} & .. & b_{2p} \\
. & . & .. & . & .. & . \\
. & . & .. & . & .. & . \\
. & . & .. & . & .. & . \\
b_{n1} & b_{n2} & .. & b_{nk} & .. & b_{np}
\end{pmatrix}
=
$$

$$
AB
$$

$$
\begin{pmatrix}
c_{11} & c_{12} & .. & c_{ik} & .. & c_{ip} \\
c_{21} & c_{22} & .. & c_{2k} & .. & c_{2p} \\
. & . & .. & . & .. & . \\
c_{11} & c_{12} & .. & c_{jk} & .. & c_{ip} \\
. & . & .. & . & .. & . \\
c_{m1} & c_{m2} & .. & c_{mk} & .. & c_{mp}
\end{pmatrix}.
$$

Where $c_{ik} = \sum_{i=1}^{m} a_{ij} b_{jk}$

Note: This is a very new kind of operation so take your time in trying to understand it. To get you used to matrix multiplication we consider the product of a row and a column matrix.

Let A = $[a_1, a_2, \ldots, a_n]$ be a |xn matrix and $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. be an n x 1 matrix. Then AB is the [x] matrix

$[a_1 b_1 + a_2 b_2 + \text{- - - - - -} + a_n b_n]$

E E26) What is $[1 \quad 0 \quad 0] \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$?

Now for another example.

Example 10: Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 7 & 0 & 8 \\ 0 & 0 & 9 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ 3 & 5 \\ 4 & 0 \end{bmatrix}$

Find AB, if it is defined.

Solution: A B is defined because the number of columns of A = 3 = number of rows of B.

$$AB = \begin{bmatrix} 1.2 + 0.3 + 0.4 & 1.1 + 0.5 + 0.0 \\ 7.2 + 0.3 + 8.4 & 7.1 + 0.5 + 8.0 \\ 0.2 + 0.3 + 9.4 & 0.1 + 0.5 + 9.0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 46 & 7 \\ 36 & 0 \end{bmatrix}$$

Notice that BA is not defined because the number of columns of B = 2 ≠ number of rows of A. thus, if AB is defined then BA may not be defined.

In fact, even if AB and BA are both defined it is possible that AB ≠ BA. Consider the following example.

Example 11: Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$

Is AB = BA?

Solution: AB is a 2 x 2 matrix. BA is a 3 x 3 matrix. So AB and BA are both defined. But they are of different sizes. Thus, AB ≠ BA.

Another point of difference between multiplication of numbers and matrix multiplication is that A ≠ 0. B ≠0, but AB can be zero.

1 1 1 0
For example. If A = 1 1 , B = -1 0

Then AB =
1 x 1 + 1(-1) 1 x 0 + 1 x 0
= 0 0
1 x 1 + 1 (-1) 1 x 0 + 1 x 0 0 0

So, you see, the product of two non-zero matrices can be zero.

The following exercises will given you some practice in matrix multiplication.

E E27) Let A =
1 1 1 0
, B =
0 1 1 1

Write AB and BA, if defined.

E E28) Let C =
1 1 0 0 1
= 0 1 0 , D = 1 1 1
0 0

Write C + D, CD and DC, if defined. Is CD = DC?

E E29) With A, B as in E 27, calculate $(A + B)^2$ and $A^2 + 2AB + B^2$. Are they equal? (Here A^2 means A. A.).

E E30 Let A = $\begin{pmatrix} -bd & d \\ -d2b & db \end{pmatrix}$, b, d ∈ F. Find A^2.

E E31) Calculate $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and ([x y z] $\begin{pmatrix} 1 & 0 & 0 \\ & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ 0

E E32) Take a 3 x 2 matrix A whose end row consists of zeros only. Multiply it by any 2 x 4 matrix B. Show that the 2^{nd} row of AB consists of zeros only. (In fact, for any two matrices A and B such that AB is defined. If the ith row of A is the zero vector, then the ith row of AB is also the zero vector. Similarly, if the jth column of B is the zero vector, then the jth column of AB is the zero vector)

We now make an observation.

Remark: If T ∈ L (U, V) and S ∈ L (V, W), then

$[ST]B_1.\ B_3 = [S]B_2.B_3\ [T]\ B_1.B_2$, where B_1, B_2, B_3 are the bases of U, V, W, respectively.

Let us illustrate this remark.

Example 12: Let $T:R^2 \rightarrow R^3$ be a linear transformation such that T (x, y) = (2x + y, x + 2y, x + t). let S: $R^3 \rightarrow R^2$ be defined by S (x, y, z) = (-y + 2z, y – z). Obtain the matrices $[T]\ B_1.B_2$, $[S]B_2.\ B_1$, and $[SoT]B_1$, and verify that $[SoT]B_1 = [S]B_2.B_1\ [T]B_1.B_2$, where B_1 and B_2 are the standard bases in R^2 and R^3, respectively.

Solution: Let B_1 =[e_1, e_2}, B_2 = {f_1, f_2, f_3}.

Then $T(e_1)$ = T(1,0) = (2, 1, 1) = $2f_1 + f_2 + f_3$

$T(e_2)$ = T(0,1, - (1, 2, 1) = $f_1 + 2f_2 + f_3$.

Thus, $[T]B_1.B_2 = \begin{matrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{matrix}$

Also,

$S(f_1) = S(1,0,0) = (0,0) = 0.e_1 + 0.e_2$
$S(f_2) = S(0,1,0) = (-1, 1) = -e_1 + e_2$
$S(f_3) = S(0, 0, 1) = (2, -1) = 2e_1 - e_2.$

Thus,

$$[S]B_2.B_1 = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 1 & -1 \end{pmatrix}$$

So,

$$[S]B_2.B_1\ [T]B_1.B_2 = \begin{pmatrix} 0 & -1 & 2 & 2 \\ & & & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

Also, SoT (x, y) = S (2x = y, x = 2y, x + y)
= (-x – 2y + 2x = 2y, x + 2y – x – y)
= (x, y)

Thus, SoT = I, the identify map.

This means $[SoT]B_1 = I_2$

Hence, $[SoT]B_1 = [S]B_2.B_1\ [T]B_1.B_2$.

E E33) Let S: $R^3 \rightarrow R^3$: S(x, y, z) = (0, x, y), and T:$R^3 \rightarrow R^3$: T(x, y, z) = (x, 0, y) show that $[SoT]_n = [S]_B\ [T]_B$' where B is the standard basis of R^3.

We will now look a little closer at matrix multiplication.

3.4.2 Properties of a Matrix Product

We will now state 5 properties concerning matrix multiplication. (Their proofs could get a little technical, and we prefer not to give them here).

(i) **Associative Law:** if A, B, C are m x n, n x p and p x q matrices, respectively, over F, then (AB) C = A (BC), i.e.., matrix multiplication is associative.

(ii) **Distributive Law:** If A is an m x n matrix and B, C are n x p matrices, then A (B + C) = AB + AC.
Similarly, if, A and B are m x n matrices, and C is an n x p matrix, then (A + B) C = AC + BC.

(iii) **Multiplicative identity:** In Sec. 7.4.1, we defined the identity matrix I_n. This acts as the multiplicative identity for matrix multiplication. We have $AI_n = A$, $I_m A = A$, for every m x n matrix A.

(iv) If a ∈ F, and A, B are m x n and n x p matrices over F, respectively then $\alpha(AB) = (\alpha A) B = A (\alpha B)$.

(v) If A, B are m x n, n x p matrices over F, respectively, then $(AB)^t = B^t A^t$. (This says that the operation of taking the transpose of a matrix is anti-commutative).
These properties can help you in solving the following exercises.

E E34) Show that $(A + B)^2 = A^2 + AB + BA + B^2$, for any two n x n matrices A and B.

E E3) For $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$,

Show that 2 (AB) = (2A) B.

E E36) Let $A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -4 & 0 \\ 2 & -1 & 3 \\ 0 & -2 & 4 \end{pmatrix}$

Find (AB) and B^1 A^1. are they equal?

E E37) Let A, B be two symmetric n x n matrices over F. Show that AB is symmetric if and only if AB = BA.

The following exercise is a nice property of the product of diagonal matrices.

E E38) Let A, B be two diagonal n x n matrices over F. Show that AB is also a diagonal matrix.

Now we shall go on to introduce you to the concept of an invertible matrix.

3.5 Invertible Matrices

In this section we will first explain what invertible matrices are. Then we will see what we mean by the matrix of a change of basis. Finally, we will show you that such matrix must be invertible.

3.5.1 Inverse of a Matrix

Just as we defined the operations on matrices by considering them on linear operators first, we give a definition of invertibility for matrices based on considerations of invertibility of linear operators.

It may help you to recall what we mean by an invertible linear transformation. A linear transformation T: U → V is invertible if

(a) T is 1 – 1 and onto, or , equivalently,
(b) There exists a linear transformation S: V → U such that SoT = I_u, ToS = I_v.

In particular, T ∈ L (V, V) is said to be invertible if ∃ S ∈ L (V, V) such that ST = TS = I.

We have the following theorem involving the matrix of an invertible linear operator.

Theorem 4: Let V be an n-dimensional vector space over a field F, and B be a basis of V. Let T ∈ L (V, V). T is invertible if there exists A ∈ M_n (F) such that $[T]_B A = I_n = A [T]_B$.

Proof: Suppose T is invertible Then ∃ S ∈ L(V, V) such that TS = ST = I. Then, by Theorem 2, $[TS]_B = [ST]_B = I$. That is, $[T]_B [S]_B = [S]_B [T]_B = I$. Take $A = [S]_B$. Then $[T]_B A = I = A [T]_B$.

Conversely, suppose ∃ a matrix A such that $[T]_B A = A [T]_B = I$.

Let S ∈ L (V, V) be such that $[S]_B = A$. (S exists because of Theorem 2) Then $[T]_B [S]_B = [S]_B [T]_B = I = [I]_B$. Thus, $[TS]_B = [ST]_B = [I]_B$.

So, by Theorem 2, ST = ST = I. that is, T is invertible.

Theorem 4 motivates us to give the following definition.

Definition: A matrix ∈ M_n(F) is said to be invertible if ∃ B ∈ M_n(F) such that AB = BA = I_n.

Remember, only a square matrix can be invertible.

I_n is an example of an invertible matrix, since $I_n \cdot I_n = I_n$. on the other hand the n x n zero matrix **0** is not invertible, since $0A = 0 \neq In$, for any A.
Note that Theorem 4 says that T is invertible iff $[T]_B$ is invertible. We give another example of an invertible matrix now.

Example 13: Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ invertible?

Solution: Suppose A were invertible. Then $\exists\, B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that AB = I = BA. Now.

$$AB = I \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a + c & b + d \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow c = 0, d = 1, a = 1, b = -1$$

$:.B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, Now you can also check that BA = I.

Therefore, A is invertible.
We now show that if an inverse of a matrix exists, it must be unique.

Theorem 5: Suppose $A \in M_n(F)$ is invertible. There exists a unique matrix $B \in M_n(F)$ such that AB = BA = I.

Proof: Suppose $B, C \in M_n(F)$ are two matrices such that AB = BA = I, and AC = CA = I. then B = BI = B (AC) = (BA) C = IC = C.

Because of Theorem 5 we can make the following definition.

Definition: Let A be an invertible matrix. The unique matrix B such that AB = BA is I is called the inverse of A and is denoted by A^{-1}.

Let us take an example.

Example 14: Calculate the product AB, where

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$$

Use this to calculate A^{-1}.

Solution: Now $AB = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a + b \\ 0 & 1 \end{pmatrix}$.

Now, how can we use this to obtain A^{-1}? Well, if AB = I, then a + b = 0. So, if we take

$$B = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix},$$

We get AB = BA = I. Thus, $A^{-1} = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$

E E39) is the matrix $\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$ invertible? If so, find its inverse.

We will now make a few observations about the matrix inverse, in the form of a theorem.

Theorem 6: a) If A is invertible, then

(i) A-1 is invertible and (A^{-1}) -1 = A,

(iii) At is invertible and $(A^t)^{-1} = (A\text{-}1)^t$.

(b) If A, B $\in M_n$ (F) are invertible, then AB is invertible and $(AB)^{-1} = B^{-1} A^{-1}$

Proof: (a) By definition.,

$A A^{-1} = A^{-1} A = I$ ………………………………. (1)

(i) Equation (I) shows that A^{-1} is invertible and $(A^{-1})^t = A$.

(ii) If we take transposes in Equation (I) and use the property that $(AB)^t = B^t A^t$, we get $(A^{-1})^t A^t = A^t (A^{-1})^t = I^t = I$.
So A^t is invertible and $(A_t)^{-1} = (A^{-1})^t$.

(b) To prove this we will use the associativity of matrix multiplication.
Now $(AB)(B^{-1} A^{-1}) = [A (BB^{-1})] A^{-1} = A A^{-1} = I$.
$(B^{-1} A^{-1})(AB) = B^{-1} [(A^{-1} A) B] = B^{-1} B = I$.

So AB is invertible and $(AB)^{-1} = B^{-1} A^{-1}$.

We now relate matrix invertibility with the linear independence of its rows or columns. When we say that the m rows of $A = [a_{ij}] \in M_{mxn}(F)$ are linearly independent, what do we mean? Let $R_1,\ldots,R_m$ be the m row vectors $[a_{11}, a_{12},\ldots, a_{1n}]$, $[a_{21},\ldots,a_{2n}],\ldots,[a_{m1},\ldots,a_{mn}]$, respectively. We say that they are linearly independent if, whenever $\exists\ a_1,\ldots, a_m \in F$ such that $a_1R_1 + \ldots + a_m R_m = 0$,

Then, $a_1 = 0, \ldots, a_m = 0$.

Similarly, the n columns $C_1, \ldots, C_n$ of A are linearly independent if $b_1C_1 + \ldots b_nC_n = 0 \Rightarrow b_1 = 0, b_2 = 0,\ldots, b_n = 0$, where $b_1,\ldots,b_n \in F$.

We have the following result.

Theorem 7: Let $A \in M_n(F)$. Then the following conditions are equivalent

(a) A is invertible
(b) The columns of A are linearly independent.
(c) The rows of A are linearly independent.

Proof: We first prove (a) $\Leftrightarrow$ (b), using Theorem 4, Let V be an n-dimensional vector space. Over F and $B = \{e_1,\ldots,e_n\}$ be a basis of V. Let $T \in L(V, V)$ be such that $[T]_B = A$. then A is invertible iff T is invertible iff T(e1), $T(e_2),\ldots, T(e_n)$ are linearly independent (see Unit 5 Theorem 9). Now we define the map

$$\theta : V \rightarrow M_{nx1}(F): (a_1e_1 + \ldots + a_n e_n) = \begin{pmatrix} a_1 \\ : \\ . \\ a_n \end{pmatrix}$$

Before continuing the proof we give an exercise.

E E40) Show that θ is a well-defined isomorphism.

Now let us go on with proving. Theorem 7.

Let $C_1, C_2, \ldots, C_n$ be the columns of A. Then θ (T(ei)) = C_i for all i = 1,... n. Since θ is an isomorphism, $T(e_1), \ldots, T(e_n)$ are linearly independent iff $C_1, C_2, \ldots, C_n$ are linearly independent. Thus, A is invertible iff $C_1, \ldots, C_n$ are linearly independent. Thus, we have proved (a) $\Leftrightarrow$ (b).

Now, the equivalence of (a) and (c) follows because A is invertible $\Leftrightarrow$ A^t is invertible $\Leftrightarrow$ the columns of A^t are linearly independent (as we have just shown) $\Leftrightarrow$ the rows of A are linearly independent (since the columns of A^t are the rows of A).

So we have shown that (a) $\Leftrightarrow$ (c).

Thus, the theorem is proved.

From the following example you can see how Theorem 7 can be useful.

Example 15:

Let $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in M_3(R)$.

Determine whether or not A is invertible.

Solution: let R_1, R_2, R_3 be the rows of A. We will show that they are linearly independent.

Suppose $x R_1 + yR_2 + zR_3 = 0$, where x, y, z, $\in$ R. Then,

X (1,0,1) + y (0, 1,1) + z (1,1,1) = (0, 0, 0). This gives us the following equations.

$$\begin{aligned} x + z &= 0 \\ y + z &= 0 \\ x + y + z &= 0 \end{aligned}$$

On solving these we get x = 0, y = 0 z = 0.

Thus, by Theorem 7, A is invertible.

E E41) Check if

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{pmatrix} \in M_3 (Q) \text{ is invertible.}$$

We will now see how we associate a matrix to a change of basis. This association will be made use of very often in the next block.

3.5.2 Matrix of Change of Basis

Let V be an n-dimensional vector space over F. Let B = $\{e_1, e_2, \ldots\ldots\ldots, e_n\}$ and B' = $\{e'_1, e'_2 \ldots\ldots\ldots e'_n\}$ be two bases of V. Since $e'_j \in$. V, for every j, it is a linear combination of the elements of B. Suppose,

$$e'_j = \sum_{i=1}^{n} a_{ij} e_i \forall j = 1, \ldots\ldots\ldots, n$$

The n x n matrix A = $[a_{ij}]$ is called the matrix of the change of basis from B to B' It is denoted by MB^B.

Note that A is the matrix of the transformation T $\in$ L (V, V) such that T $(e_j) = e'_j \forall$ j = 1, ……, n, with respect to the basis B. Since $\{e'_j, \ldots\ldots, e'_n\}$ is a basis of V, from Unite 5 we see that T is 1 – 1 and onto. Thus T is invertible. So A is invertible. Thus, the matrix of the change of basis from B to B' s invertible.

Note: a) $M^B{}_B$ = In, This is because, in this case $e'_j = e_j \forall i = 1, 2, \ldots\ldots, n$.

b) $M^{B'}{}_B$ = [I] B', B. This is because

$$I(e'_j) = e'_j = \sum_{i=1} a_{ij} e_i \forall j = 1, 2, \ldots\ldots, n$$

Now suppose A is any invertible matrix. By Theorem 2, $\exists$ T $\in$ L (V, V) such that $[T]_R$ = A. Since A is invertible, T is invertible. Thus, T is 1 – 1 and onto. Let f_i = T

(e_i) ∀ i = 1, 2 . . . n, Then B' = {f_1, f_2,…..,f_n} is also a basis of V, and the matrix of change of basis from B to B' is A.

In the above discussion, we have just proved the following theorem.

Theorem 8: Let B = {e_1, e_2, ...,e_n} be a fixed basis of V. the mapping B' → $M^{B'}_B$ is a 1 -1 and onto correspondence between the set of all bases of V and the set of invertible n x n matrices over F.

Let us se an example of how to obtain $M^{B'}_B$.

Example 16: In R^2, B = {e_1, e_2} is the standard basis. Let B' be the basis obtained by rotating B through a angle θ in the anti-clockwise direction (see Fig. 1). Then B' = (e'_1, e'_2) where e'_1 = (cos θ, sin θ), e'_2 = (-sin θ, cos θ). Fin $M^{B'}_B$.

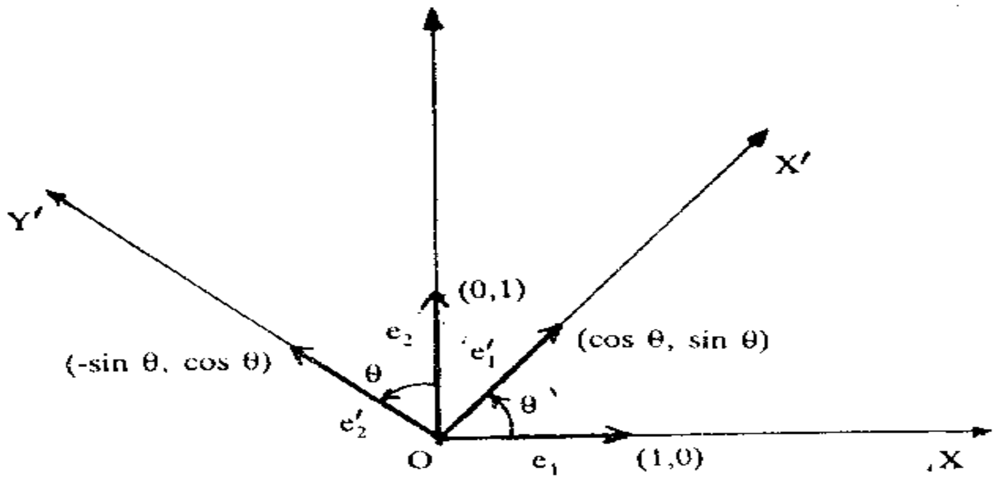

Fig. 1: Change of basis.

Solution: e'_1 = Cos θ (1.0) + sin θ (0.1), and
= -sin θ (1,0) + cos θ (0.1)

Thus, $M^{B'}_B = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Try the following exercise.

E E42) Let B be the standard basis of R^3 and B' be another basis such that

$M^{B'}_B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ & 0 & 0 \end{pmatrix}_3$

What are the elements of B'?

What happens if we change the basis more than once? The following theorem tells us something about the corresponding matrices.

Theorem 9: Let B, B', B" be three bases of V. Then $M^{B'}_{B}$ $M^{B'}_{B}$ = $M^{B'}_{B}$.

Proof: Now, $M^{B'}_{B}$ $M^{B'}_{B}$ = [I]B'. B [I]B'.B
= [IoI]B'.B' = $M^{B'}_{B}$.

An immediate useful consequence is

Corollary: Let B.B' be two bases of V. then $M^{B'}_{B}$ $M^{B'}_{B}$ = I = $M^{B'}_{B}$ $M^{B'}_{B}$

That is, $(M^{B'-1}_{B})$ = $M^{B'}_{B}$.

Proof: By Theorem 9,
$M^{B'}_{B}$ $M^{B'}_{B}$ = $M^{B'}_{B}$ = I

Similarly, $M^{B'}_{B}$ $M^{B'}_{B}$ = $M^{B'}_{B}$ = I.

But, how does the change of basis affect the matrix associated to a given linear transformation? In sec. 7.2 we remarked that the matrix of a linear transformation depends upon the pair of bases chosen. The relation between the matrices of a transformation with respect to two pairs of bases can be described as follows.

Theorem 10: Let T ∈ L (U, V), Let $B_1 = \{e_1, \ldots\ldots, e_n\}$ and $B_2 = \{f_1, \ldots\ldots, f_m\}$ be a pair of bases of U and V, respectively.

Let $B'_1 = \{e'_1, \ldots\ldots, e'_n\}$, $B'_2 = \{f'_1, \ldots\ldots f'_m\}$ be another pair of bases of U and V, respectively. Then.

$[T]B'_1.B'_2$ = $M^{B'}_{B}$ $[T]B_1.B_2$ $M^{B'1}_{B1}$.

Proof: $[T]B'_1.B'_2$ = [I Vo To Iu]$B'_1.B'_2$ = [IV]$B_1.B_2$ [Iu]$B'_1.B_1$
(where I_u = identity map on U and Iv = identity map on V)

= $M^{B'2}_{B2}$ $[T]B_1.B_2$ $M^{B'1}_{B1}$.

Now, a corollary to Theorem 10, which will come in handy in the next block.

Corollary: Let T ∈ L (V, V) and B, B' be two bases of V. Then $[T]B' = P^{-1}\ [T]_B^P$, where P = $M^{B'}_{B}$.

Proof: [T]B = $M^{B'}_{B}$ [T]B $M^{B'}_{B}$ = P-1 [T]BP, by the corollary to Theorem 9. Let us now recapitulate all that we have covered in this unit.

4.0 SUMMARY

We briefly sum up what has been done in this unit.

1) We defined matrices and explained the method of associating matrices with linear transformations.
2) We showed what we mean by sums of matrices and multiplication of matrices by scalars.
3) We proved that M_{mxn} (F) is a vector space of dimension mn over F.
4) We defined the transpose of a matrix, the conjugate of a complex matrix, the conjugate transpose of a complex matrix, a diagonal matrix, identity matrix, scalar matrix and lower and upper triangular matrices.
5) We defined the multiplication of matrices and showed its connection with the composition of linear transformations. Some properties of the matrix product were also listed and used.
6) The concept of an invertible matrix was explained.
7) We defined the matrix of a change of basis, and discussed the effect of change of bases on the matrix of a linear transformation.

3.6 Solutions/Answers

E1) a) You want the elements in the 1^{st} row and the 2^{nd} column. They are 2 and 5, respectively.

b) [0 0 7]

c) The second column of A is $\begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}$

The first column of B is also $\begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix}$

d) B only has 3 rows. Therefore, there is no 4^{th} row of B.

E2) They are infinitely many answers. We give.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 0 \\ 4 & 0 \end{pmatrix}$$

E3) No. Because they are of different sizes.

E4) Suppose B'_1 = {(1, 0, 1), (0, 2, - 1)} and B'_2 = {0, 1), (1, 0)}

Then T (1, 0, 1) = (1,) = 0. (0,1) + 1. (1, 0)
T (0, 2, - 1) = (0,2) = 2. (0, 1) + 0, (1, 0)
T (1, 0, 0) = (1, 0) = 0. (0, 1) + 1. (1, 0).

$$:. [T]B'_1.B'_2 = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

E5) $B_1 = \{e_1, e_2, e_3\}$ $B_2 = \{f_1, f_2\}$ are the standard bases (given in Example 3).
T (e_1) = T (1, 0, 0) = (1, 2) = $f_1 + 2f_2$
T (e_2) = T (0, 1, 0) = (2, 3) = $2f_1 + 3f_2$
T (e_3) = T (0, 0, 1) = (2, 4) = $2f_1 + 4f_2$.

$$:. [T]B_1.B_2 = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \end{pmatrix}$$

E6) T (1, 0, 0) = (1, 2) = 1. 91, 2) = 0. (2, 3)
T (0, 1, 0) = (2, 3) = 0. (1, 2) + 1. (2, 3)
T (1, - 2, 1) = (-1, 0) = 3. (1, 2) -2 (2, 3)

$$:. [T]\ B'_1.B'_2 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}$$

E7) Let B = $\{1, t, t^2, t^3\}$. Then
D (1) = $0.1 + 0.t + 0.t^2 + 0.t^3$
D(t) = $1 = 1.1 + 0.t + 0.t^2 + 0.t^3$
$D(t^2)$ = $2t = 0.1 + 2.t + 0.t^2 + 0.t^3$
$D(t^3)$ = $3t^2 = 0.1 + 2.t + 3.t^2 + 0.t^3$.
Therefore, $[D]_B$ is the given matrix.

E8) We know that
T (e_1) = f_1
T (e_2) = $f_1 + f_2$
T (e_3) = f_2

Therefore, for any (x, y, z) $\in R^3$.

T (x, y, z) = T $(xe_1 + ye_2 + ze_3)$ = $xT(e_1) + yT(e_2) + zT(e_3)$
= $xf_1 + y(f_1 + f_2) + zf_2 = (x + y)\ f_1 + (y + z)f_2$
= (x + y, y + z)
That is, $R:R^3 \to R^2$: T (x, y, z) = (x + y, y + z)

E9) We are given that
T(1) = 0. 1 + 1.i = i
T(i) = (-1).1 + 0.1 = -1
:., for any a + ib E C, we have

T (a + ib) = a T (1) + bT (i) = ai – b

E10) a) Since [1 2] is of size 1 x 2 and $\begin{bmatrix}3\\1\end{bmatrix}$ is of size 2 x 1,
The sum of these matrices is not defined.

c) Both matrices are of the same size, namely, 2 x 2. their sum is the matrix.

$$\begin{pmatrix} 1 + (-1) & 0 + 0 \\ 0 + 0 & 1 + (-1) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

E11) $3\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}3\\6\end{bmatrix}, \quad 3\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\3\end{bmatrix}$

and $3\left(\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}\right) = 3\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}3\\9\end{bmatrix}$

Notice that $3\left(\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}\right) = 3\begin{bmatrix}1\\2\end{bmatrix} + 3\begin{bmatrix}0\\1\end{bmatrix}$

E12) B1 = {(1,0), (0, 1), B2 = {(1,0,0), (0,1,0), (0,01)}
Now S(1,0) = (1,0,0)
S(0,1) = (0, 0, 1)

$$:.[S]_{B_1.B_2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$, a 3 x 2 matrix

Again, T (1, 0) = (0, 1, 0)
T(0, 1) = (0,0 1)

$$:.[T]_{B_1.B_2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$, a 3 x 2 matrix

$\therefore [S + T]_{B_1 . B_2} = [S]_{B_1.B_2} + [T]_{B_1. B_2} =$

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

And $[\alpha S]_{B_1.B_2} = \alpha[S]_{B_1.B_2} = \alpha \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 0 & \\ 0 & 1 & \end{pmatrix} = \begin{pmatrix} 0 & \\ 0 & 0 \\ 0 & \alpha \end{pmatrix}$, for any $\alpha \in R$

E13) We will prove (v) and (vi) here. You can prove (vii) and (viii) in a similar way .

v) $\alpha(A + B) = \alpha([a_{ij}]) = \alpha[a_{ij} + b_{ij}] = [\alpha a_{ij} + \alpha b_{ij}]$
$= [\alpha a_{ij}] + [\alpha b_{ij}] = \alpha A + \alpha B.$

vi) prove it using the fact that $(\alpha + \beta)a_{ij} = \alpha a_{ij} + \beta a_{ij}$.

E14) Since dim M_{2x3} (R) is 6, any linearly independent subset can have 6 elements, at most.

E15) Let $\alpha, \beta \in R$ such that $\alpha[1,0] + \beta[1, - 1] = [0, 0]$

Then $[\alpha + \beta, - \beta] = [0, 0]$. Thus, $\beta = 0, \alpha = 0$.
$\therefore$ The matrices are linearly independent.

E16) $E_{11} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ and so on

Now any m x n matrix $A = [a_{ij}] = a_{11} E_{11} + a_{12}E_{12} + \dots + a_{mn} E_{mn}$ (for example, in the 2 x 2 situation.

$$\begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} = 2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 3 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Thus, $[E_{ij}]$ {i = 1,…, m, j = 1,…,n} generates M_{mxn} (F). Also, if α_{ij}, I = 1,…., m, J = 1,…..,n, be scalars such that $\alpha_{11} E_{11} + \alpha_{12} E_{12} + \dots + \alpha_{mn} E_{mn} = 0$.

Thus

We get
$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix} = \begin{pmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

Therefore, $a_{ij} = 0 \ \forall$ I j.

Hence, the given set is linearly independent. :. It is a basis of M_{mxn} (F). The number of elements in this basis is mn.

:. dim M_{mxn} (R) = mn.

E17) $A^t = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$, In this case $A^t = A$.

E18) b) $\alpha A = [\alpha a_{ij}]$. :. $(\alpha A)^t = [b_{ij}]$, where

b_{ij} = (j ,i)th element of $\alpha A = \alpha a_{ji}$
= a times the (j, i)th element of A
= a times the (i, j)th element of A^t
= (i, j)th element of αA^t.

:. $(\alpha A)^t = A$.

E19) Let A be an m x n matrix. Then A^t is an n x m matrix.
:., for $A = A^t$, their sizes must be the same, that is, m = n.
:. A must be a square matrix.

E20) Let $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, be a square matrix over a field F.

Then $A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$

:. $A + A^t = \begin{pmatrix} a + a & b + c \\ c + b & d + d \end{pmatrix} = \begin{pmatrix} 2a & b + c \\ b + c & 2d \end{pmatrix}$, and

$A - A^t = \begin{pmatrix} a - b & b - c \\ c - b & d - d \end{pmatrix} = \begin{pmatrix} 0 & b - c \\ -(b - c) & 0 \end{pmatrix}$

You can check that $(A + A^t) = A + A^t$ and $(A - A^t)^t = -(A - A^t)$.
:. $A + A^t$ is symmetric and $A - A^t$ is skew-symmetric.

E21) $\begin{pmatrix} -i & 2 \\ 3 & -i \end{pmatrix}$

E22) The size of Āt is the same as the size of At. :. A = Āt. Implies that the sizes of A and At are the same. :. A is a square matrix.

E23) $I: R^n \rightarrow R^n : I(x_1,\ldots,x_n) = (x_1,\ldots,x_n)$.
Then, for any basis $B = \{e_1,\ldots,e_n\}$ of R^n. $I(e_i) = e_i$.

$$:.[I]_B = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I_n.$$

E24) Since A is upper triangular, all its elements below the diagonal are zero. Again, since A = At, a lower triangular matrix, all the entries of A above the diagonal are zero:., all the off-diagonal entries of A are zero. :. A is a diagonal matrix.

E25) Let A be a skew-symmetric matrix. Then $A = -A^t$. Therefore,

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} -a_{11} & -a_{21} & \ldots & -a_{n1} \\ -a_{12} & -a_{22} & \ldots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \ldots & -a_{nn} \end{pmatrix}.$$

:., for any $i = 1,\ldots, n$, $a_{ij} = -a_{ij} = 0 \Rightarrow a_{ij} = 0$.

The converse is not true. For example, the diagonal entries of $\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ are zero, but this matrix is not skew-symmetric.

E26) $[1 \times 1 + 0 \times 2 + 0 \times 3] = [1]$

E27) $AB = \begin{pmatrix} 1 \times 1 + 1 \times 1 & 1 \times 0 + 1 \times 1 \\ 0 \times 1 + 1 \times 1 & 0 \times 0 + 1 \times 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$

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

E28) C + D is not defined.

CD is a 2 x 2 matrix and DC is a 3 x 3 matrix: CD ≠ DC.

$$CD = \begin{pmatrix} 1 \times 0 + 1 \times 1 + 0 \times 0 & 1 \times 1 + 1 \times 1 + 0 \times 0 \\ 0 \times 0 + 1 \times 1 + 0 \times 0 & 0 \times 1 + 1 \times 1 + 0 \times 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

$$DC = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \times 1 + 1 \times 0 & 0 \times 1 + 1 \times 1 & 0 \times 0 + 1 \times 0 \\ 1 \times 1 + 1 \times 0 & 1 \times 1 + 1 \times 1 & 1 \times 0 + 1 \times 0 \\ 0 \times 1 + 0 \times 0 & 0 \times 1 + 0 \times 1 & 0 \times 0 + 0 \times 0 \end{pmatrix}$$

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

E29) $A + B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ∴ $(A + B)^2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$

Also $A^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

$B2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

2. $AB = 2 \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$

∴ $A^2 + 2AB + B^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}$

∴ $(A + B)^2 \neq A^2 + 2AB + B^2$.

E30) $A^2 = \begin{pmatrix} -bd & b \\ d^2 b & db \end{pmatrix} \begin{pmatrix} -bd & b \\ d2\,b & db \end{pmatrix} = \begin{pmatrix} b^2d^2 + d^2b^2 & -b^2d + db^2 \\ -b^2 d^3 + d^3 b^2 & d^2 b^2 + d^2 b^2 \end{pmatrix} \begin{pmatrix} 2d^2 & b^2 \\ 0 & \end{pmatrix}$

E31) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 2y \\ 3z \end{pmatrix}$

$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$

$$[x\ y\ x] \begin{matrix} 0 & 2 & 0 \\ 0 & 0 & 3 \end{matrix} = [x\ \ 2y\ \ 3z]$$

E32) We take $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 1 & 1 \end{pmatrix}$, Then

$AB = \begin{pmatrix} 9 & 12 & 5 & 2 \\ 0 & 0 & 0 & 0 \\ 7 & 11 & 10 & 1 \end{pmatrix}$, You can see that the 2^{nd} row of AB is zero

E33) $[S]_B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, $[T]_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

$$\therefore [S]_B\ [T]_B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{Also, } [SoT]_B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = [S]_B\ \ [T]_B$$

E34) $(A + B)^2 = (A + B) = A(A + B) + B(A + B)$ (by distributivity)

$= A^2 + AB + BA + B^2$ (by distributiveity)

E35) $AB = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ 9 & 22 & 15 \end{pmatrix}$ $\therefore 2(AB) = \begin{pmatrix} -2 & -16 & -5 \\ 2 & -4 & -10 \\ 18 & 44 & 30 \end{pmatrix}$

On the other hand, $(2A)\,B = \begin{pmatrix} 4 & -2 \\ 2 & 0 \\ -6 & 8 \end{pmatrix} \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$

$$= \begin{pmatrix} -2 & -16 & -20 \\ 2 & -4 & -10 \\ 18 & 44 & 30 \end{pmatrix}$$

:. 2(AB) = (2A)B

E36) AB = $\begin{pmatrix} 0 & -7 & -3 \\ -11 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}$. :. $(AB)^t = \begin{pmatrix} 0 & -11 & 0 \\ -7 & -4 & 0 \\ -3 & 6 & 0 \end{pmatrix}$

Also, $B^t A^t = \begin{pmatrix} 1 & 2 & 4 \\ & -4 & -1 \\ 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -11 & 0 \\ -7 & -4 & 0 \\ -3 & 6 & 0 \end{pmatrix}$

$= (AB)^t$.

E37) First, suppose AB is symmetric. Then $AB = (AB)^t = B^t A^t = BA$, since A and B are symmetric.

Conversely, suppose AB = BA. Then

$(AB)^t = B^tA^t = BA = AB$, so that AB is symmetric.

E38) let $A = \text{diag}(d_1, \ldots, d_n)$, $B = \text{diag}(e_1, \ldots, e_n)$. Then

$$AB = \begin{pmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ . & . & . & \ldots & . \\ : & : & : & \ldots & : \\ 0 & 0 & : & \ldots & d_n \end{pmatrix} \begin{pmatrix} e_1 & 0 & 0 & \ldots & 0 \\ 0 & e_2 & 0 & \ldots & 0 \\ . & . & . & ::: & : \\ : & : & : & \ldots & : \\ 0 & 0 & : & \ldots & e_n \end{pmatrix}$$

$$= \begin{pmatrix} d_1 e_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 e_2 & 0 & \ldots & 0 \\ 0 & 0 & d_3 e_3 & \ldots & 0 \\ : & : & : & & : \\ 0 & 0 & 0 & \ldots & d_n e_n \end{pmatrix}.$$

$= \text{diag}(d_1 e_1, d_2 e_2, \ldots, d_n e_n)$.

E39) Suppose it is invertible. Then E A = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that

$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}$

2 -1 c d = 0 1 = c d 2 -1

This gives us $A = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$ which is the same as the given matrix.

This shows that the given matrix is invertible and, in fact, $\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$

E40) Firstly, θ is a well defined map. Secondly, check that $\theta(v_1 + v_2) = \theta(v_1) + \theta(v_2)$, and $\theta(\alpha v) = \alpha\theta(v)$ for $v, v_1, v_2, \in V$ and $\alpha \in F$. Thirdly, show that $\theta(v) = 0 \Rightarrow v = 0$, that is θ is 1 – 1. Then; by Unit 5 (Theorem 10), you have shown that θ is an isomorphism.

E41) We will show that its columns are linearly independent over Q. Now, if $x, y, z \in Q$ such that

$$x \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + z \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$, we get the equations

$$2x + z = 0$$

$$z = 0$$

$$3y = 0$$

on solving them we get $x = 0, y = 0, z = 0$.

:. the given matrix is linearly independent.

E42) Let $B = \{e_1, e_2, e_3\}$ $B' = \{f_1, f_2, f_3\}$. Then

$f_1 = 0e_1 + 1e_2 + 0e_3 = e_2$

$f_2 = e_1 + e_2$

$f_3 = e_1 + 3e_3$

:. $B' = \{e_2, e_1 + e_2, e_1 + 3e_3\}$.

UNITS 4 MATRICES – II

1.0 INTRODUCTION

In Unit 3 we introduced you to a matrix and showed you how a system of linear equations can give us a matrix. An important reason for which linear algebra arose is the theory of simultaneous linear equations. A system of simultaneous linear equations can be translated into a matrix equation, and solved by using matrices.

The study of the rank of a matrix is a natural forerunner to the theory of simultaneous linear equations. Because, it is in terms of rank that we can find out whether a simultaneous system of equations has a solution or not. In this unit we start by studying the rank and inverse of a matrix. Then we discuss row operations on a matrix and use them for obtaining the rank and inverse of a matrix. Finally, we apply this knowledge to determine the nature of solutions of a system of linear equations. The method of solving a system of linear equations tat we give here is by "successive elimination of variable". It is also called the Gaussian elimination process.

With this unit we finish Block 2. In the next block we will discuss concepts that are intimately related to matrices.

2.0 OBJECTIVES

After reading this unit, you should be able to

- Obtain the rank of a matrix;
- Reduce a matrix to the echelon form;
- Obtain the inverse of a matrix by row-reduction;
- Solve a system of simultaneous linear equations by the method of successive elimination of variables.

3.0 MAIN COURSE

3.1 Rank Of A Matrix

Consider any mxn matrix A, over a field F. We can associate two vector spaces with it, in a very natural way. Let us see what they are. Let A = $[a_{ij}]$. A has m rows, say, $R_1, R_2, \ldots, R_m$, where $R_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $R_2 = (a_{21}, a_{22}, \ldots, a_{2n}) \ldots$, $R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$.

Thus, $R_1 \in$ i , and ∀ A = $\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix}$

The subspace of Fn generated by the row vectors $R_1, \ldots R_m$ of A, is called the **row space** of A, and is denoted by **RS (A)**.

Example 1: If A = $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, does (0,0,1) $\in$ RS (A)?

Solution: The row space of a is the subspace of R3 generated by (1, 0, 0) and (0,1,0). Therefore, RS (A) = {a, b, 0) | a, b $\in$ R). Therefore (0, 0, 1) $\notin$ RS (A).

The dimension of the row space of A is called the row rank of A, and is denoted by p_r (A).

Thus, p_r (A) = maximum number of linear independent rows of A.

In Example 1, $p_r(A) = 2$ = number of rows of A. But consider the next example.

Example 2: If A = $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}$, find $p_r(A)$

Solution: The row space of A is the subspace of R^2 generated by (1,0), (0,1) and (2,0). But (2,0) already lies in the vector space generated by (1,0) and 0,1), since (2,0) = 2 (1,0). Therefore, the row of A is generated by the linear independent vectors (1,0) and (0,1). Thus, p_r (A) = 2.

So, in Example 2, $p_r(A)$ < number of rows of A.

In general, for m x n matrix A, RS (A) is generated by m vectors. Therefore, $p_r(A) \leq m$. Also, RS (A) is a subspace of F^n and $\dim_F F^n = n$. Therefore, $p_r(A) \leq n$.

Thus, for any mxn matrix A, $0 \leq p_r(A) \leq \min(m,n)$.

E E1) show that $A = 0 \Leftrightarrow p_r(A) = 0$.

Just as we have defined the row space of A, we can define the column space of A. each column of A is an m-tuple, and hence belongs to Fm. We denote the column of A by $C_1, \ldots, C_n$. The subspace of Fm generated by $\{C_1, \ldots C_n\}$ is called the column space of A and is denoted by CS (A). The dimension of CS (A) is called the column rank of A, and is denoted of p_c (A). Again, since CS (A) is generated by n vectors and is a subspace of Fm, we get $0 \leq pc(A) \leq \min(m, n)$.

E E2) Obtain the column rank and row rank of A = $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

In E2 you may have noticed that the row and column ranks of A equal. In fact, in Theorem 1, we prove that $p_r(A) = p_c(A)$, for any matrix A. But first, we prove a lemma.

Lemma 1: Let A, B be two matrices over F such that AB is defined. Then

a) $CS(AB) \subseteq CS(A)$,
b) $RS(AB) \subseteq RS(B)$.

Thus, $p_c(AB) \leq p_c(A)$, $p_r(AB) \leq p_r(B)$.

Proof: (a) Suppose A = $[a_{ij}]$ is an n x p matrix. Then, from sec. 3.5, you know that the jth column of C = AB will be

$$\begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{1k} b_{kj} \\ \sum_{k=1}^{n} a_{2k} b_{kj} \\ \vdots \\ \sum_{k=1}^{n} a_{mk} b_{kj} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{m} \end{bmatrix} b_{1j} + \ldots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{mn} \end{bmatrix} b_{nj}$$

$$= C_1 b_{1j} + \ldots + C_n b_{nj}.$$

Where $C_1,\ldots C_n$ are the columns of A.

Thus, the columns of AB are linear combinations of the columns of A. thus, the columns of AB ∈ CS(A). So, CS(AB) ⊆ CS(A).

Hence, p_c (AB) ≤ p_c(A).

b) By a similar argument as above, we get RS (AB) ⊆ RS(B), and so, p_r(AB) ≤ p_r(B).

E E3) Prove (b) of Lemma 1.

We will now use Lemma 1 for proving the following theorem.

Theorem 1: pc(A) = pr(A), for any matrix **A** over **F.**

Proof: Let A ∈ Mmxn (F) Suppose p_r(A) = r and p_c(A) = t.

Now, RS(A) = [{R_1, R_2,…., m where R_1, R_2, ….,R_m are the rows of A. let {e1,e2, …er,} be a basis of RS(A). Then R_1 is a linear combination of $e_1 \ldots e_r$, for each i =1…,m. Let

$$R_i = \sum_{j=1}^{r} b_{ij} e_j, \; i\ 1, 2, \ldots m,$$ where $b_{ij} \in F$ for $1 \le i \le m$, $1 \le j \le r$.

We can write these equations in matrix form as

$$\begin{bmatrix} R_1 \\ \vdots \\ \; \end{bmatrix} \qquad \begin{bmatrix} b_{11} \ldots\ldots b_{1r} \\ \; \\ \; \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ \; \end{bmatrix}$$

$$\begin{matrix} & & \ldots\ldots\ldots\ldots \\ & = & \ldots\ldots\ldots\ldots \\ & & \ldots\ldots\ldots\ldots \\ R_m & & b_{m1}\ldots\ldots b_{mr} & e_r \end{matrix}$$

So, A = BE, where B = $[b_{ij}]$ is an mxr matrix and E is the rxn matrix with rows e_1, $e_2,\ldots e_r$. (remember, e1 $\in F^n$, for each i=1,…., r.)

So, t = pc (A) = (BE) ≤ pc (B), by Lemma 1.

≤ min (m,r,)

≤ r

Thus, t ≤ r.

Just as we got A = BE above, we get A = $[f_1,\ldots f_t]$ D, where { $f_1,\ldots f_t$} is a basis of the column space of a and D is a t x n matrix. Thus, r = $p_r(A) \le p_r(D) \le t$, by Lemma 1. So we get r ≤ and t ≤ r. This gives us r = t.

Theorem 1 allows us to make the following definition.

Definition: the integer $p_c(A)$ (=$p_r(A)$ is called the rank of A, and is denoted by p(A).

You will see that theorem 1 is very helpful if we want to prove any fact about p(A). If it is easier to deal with the rows of A we can prove the fact for $p_r(A)$. similarly, if it is easier to deal with the columns of A, we can prove the fact for $p_c(A)$. While proving Theorem 3 we have used this facility that theorem 1 gives us.

Use theorem 1 to solve the following exercises.

E E4) If A,B are two matrices such that AB is defined then show that p(AB) ≤ min (p(A), p(B)).

E E5) Suppose C ≠ 0 $\in M_{mx1}(F)$, and R ≠ 0 $\in M_{1xn}(F)$, then show that the rank of the mxn matrix CR is 1. (Hint: use E4).

Does the term 'rank' seem familiar to you? Do you remember studying about the rank of a linear transformation in Unit 2? We now see if the rank of a linear transformation is related to the rank of its matrix. The following theorem brings forth the precise relationship. (God through sec. 2.3 before further.)

Theorem 2:Let U.V be vector spaces over F of dimensions n and m, respectively. Let B_1 be a basis of U and B_2 be a basis of V. Let $T \in L(U,V)$.

Then $R(T) \sim CS([T]_{B1,\ B2})$.

Proof: Let $B_1 = \{e_1, e_2\}$ and $B_2 = \{f_1, f_2, \ldots f_m\}$. As in the proof of Theorem 7 of unit 3, $\theta: V \rightarrow M_{m,1}(F)$: $\theta(v)$ = coordinate vector of v with respect to the basis B_2, is an isomorphism.

Now. $R(T) = |\{ T(e_1), T(e_2) \ldots T(e_n)\}|$. Let $A = [T]_{B_1,\ B_2,}$ have $C_1, C_2, \ldots C_n$ as its columns.

Then $CS(A) = |\{C_1, C_2 \ldots C_n\}|$. Also, $\theta(T(e1)) = Ci \quad i = 1, \ldots, n$.

Thus, $\theta: v \rightarrow CS(A)$ is an isomorphism. $\therefore R(T) \cong CS(A)$

In particular. Dim R(T) = dim CS(A) = p(A).

That is, rank (T) = p(A).

Theorem 2 leads us to the following corollary. It says that pre-multiplying or post-multiplying a matrix by invertible matrices does not alter its rank.

Corollary 1: Let A be an m x n matrix. Let P,Q be mxm and nxn invertible matrices, respectively.

Then p (PAQ) = p(A).

Proof: Let $T \in L(U,V)$ be such that $[T]_{B_1,\ B_2,} = A$. We are given matrices Q and P^{-1}. Therefore, by theorem 8 of Unit 3, $\exists$ bases B_1 and B_2 of U and V, respectively, such that $Q = M^{B_1}_{B_1,}$ and $P^{-1} = M^{B_2}_{B_2}$

Then, by theorem 10 of Unit 7,

$$[T]_{B_1,\ B_2,} = M^{B_2}_{B_2} [T]_{B_1,\ B_2,} M^{B_1}_{B_1,} = PAQ$$

In other words, we can change the bases suitable so that the matrix of T with respect to the new bases is PAQ.

So, by theorem 2, p(PAQ)= rank (T) = Thus, p (PAQ) = p(A).

E E6) Take $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{bmatrix}$, $P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ $Q = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Obtain PAQ and show that p(PAQ) = p(A).

Now we state and prove another corollary to Theorem 2. This corollary is useful because it transforms any matrix into a very simple matrix, a matrix whose entries are 1 and 0 only.

Corollary 2: Let A be an mxn matrix with rant r. Then ∃ invertible matrices P and Q such that PAQ = $\begin{bmatrix} 1_r & 0 \\ 0 & 0 \end{bmatrix}$

Proof: Let T ∈ L((,V) be such that $[T]_{B_1, B_2}$ = A. since p(A) = r, rank (T) = r. ∴ nullity (T) = n-r (Unit 2, Theorem 5).

Let $\{u_1, u_2, \ldots u_{n-r}\}$ b a basis of Ker T. We extend this to form the basis $B_1 = u_1, u_2, \ldots u_{n-r}, u_{n-r+1}, \ldots, u_n\}$ of U. then $\{T(u_{n-r+1}), \ldots, T(u_n)$ is a basis of R(T) (see Unit 5, proof of Theorem 5). Extend this set to form a basis B_2 of V, say B_2 = $\{T(u_{n-r+1}, \ldots, T(u_n), v_1, \ldots, v_{m-r}\}$. Let us reorder the elements of B_1 and write it as B1 = $\{ u_{n-r+1}, u_n, u1, \ldots . u_{n-r}\}$.

Then, by definition, $[T]_{B_1, B_2} = \begin{bmatrix} 1r & 0_{1x(n-r)} \\ 0_{(m-r)\ xr} & 0_{(m-r)x(n-r)} \end{bmatrix}$

Where 0_{sxt} denotes the zero matrix of size sxt. (Remember that $u_1, \ldots, u_{n-r} \in$ Ker T.)
Hence, PAQ = $\begin{bmatrix} 1_r & 0 \\ 0 & 0 \end{bmatrix}$

Where Q = $M^{B_1}_{B_1}$ and P = $M^{B_2}_{B_2}$, by Theorem 10 of Unit 3.

$\begin{bmatrix} 1_r & 0 \\ 0 & 0 \end{bmatrix}$

Note is called the normal form of the matrix A.

Consider the following example, which is the converse of E5.

Example 3: A is an mxn matrix of rank 1, show that ∃ C ≠ 0 M_{mx1}(F) and R ≠ 0 in M_{1xn} (F) such that a = CR.

Solution: By corollary 2 (above), ∃ P,Q such that

$$PAQ = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}, \text{ since } p(A) = 1.$$

$$= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [1 \;\; 0 \;\; \ldots \;\; 0]$$

$$\therefore A = p^{-1}(PAQ)Q^{-1} = p^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [1 \;\; 0 \;\; \ldots \;\; 0]\, A^{-1} = CR.$$

Where $C = p^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \neq 0$. $R = [1 \;\; 0 \;\; \ldots \;\; 0]\, Q^{-1} \neq 0$.

E E7) What is the normal form of diag (1,2,3)?

The solution of E7 is a particular case of the general phenomenon: the normal form of an nxn invertible matrix is 1_n.

Let us now look at some ways of transforming a matrix by playing around with its rows. The idea is to get more and more entries of the matrix to be zero. This will help us in solving systems of linear equations.

3.2 Elementary Operations

Consider the following set of 2 equations in 3 unknowns x, y and z:

$$x + y + z = 1$$
$$2x + 3z = 0$$

How can you express this system of equations in matrix form?

One way is

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

In general, if a system of m linear equations in n variable, $x_1, \ldots, x_n$, is

$$a_{11}x_1 + a_{12}\,x_2 + .. + a_{1n}x_n = b_1$$

$$\vdots \qquad \vdots \qquad \qquad \vdots$$

$$a_{m1}x_1 + a_{m2}\,x_2 + .. + a_{mn}x_n = b_m$$

where $a_{ij}, b_i \in F \forall \ i = 1, \ldots, m$ and $j = 1, \ldots. n$, then this can be expressed as AX = B,

where $A = [a_{ij}]_{mxn}$, $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

In this section we will study methods of changing the matrix A to a very simple form so that we can obtain an immediate solution to the system of linear equations AX = B. For this purpose, we will always be multiplying A on the left or the right by a suitable matrix. In effect, we will be applying elementary row or column operations on a.

3.3 Elementary Operations on a matrix

Let A be an mxn matrix. As usual, we denote its rows by $R_1, \ldots, R_m$, and columns by $C_1, \ldots C_n$. We call the following operations elementary row operations:

1) Interchanging R1 and Rj for $i \neq j$.
2) Multiplying R1 by some $a \in F$, $a \neq 0$.
3) Adding aR1 to R1, where $i \neq j$ and $a \in F$.

We denote the operation (1) by R_{ij}, (2) by $R_i(a)$, (3) by $R_{ij}(a)$.

For example. if $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}$

Then $R_{12}(A) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ (interchanging the two rows).

Also $R_2(3)(A) = \begin{bmatrix} 1 & 2 & 3 \\ 0x3 & 1x3 & 2x3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \end{bmatrix}$

and $R_{12}(2)(A) = \begin{bmatrix} 1 + 0 \text{ x } 2 & 2 + 1 \text{ x } 2 & 3+2x2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 7 \\ 0 & 1 & 2 \end{bmatrix}$

E E8) If A = $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, what is

a) R_{21} (A) b) R_{32} o R_{21} (A) c) $R_{13}(-1)$ (A)?

Just as we defined the row operations, we can define the three column operations as follows:

1) Interchanging C_1 and C_1 for $i \neq j$ denoted by C_{ij}.
2) Multiplying C_1 by $a \in F$, $a \neq 0$, denoted by C_i,(a).
3) Adding a C_1 to C_1, where $a \in F$, denoted by C_{ij} (a).

For example, if A = $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$

Then C21(10) (A) = $\begin{bmatrix} 1 & 13 \\ 2 & 24 \end{bmatrix}$

and C12(10) (A) = $\begin{bmatrix} 31 & 3 \\ 42 & 4 \end{bmatrix}$

We will now prove a theorem which we will use in sec. 8.3.2 for obtaining the rank of a matrix easily.

Theorem 3: elementary operations on a matrix do not alter its rank.

Proof: the way we will prove the statement is to show that the row space remains unchanged under row operations and the column space remains unchanged under column operations. This means that the row rank and the column rank remain unchanged. This immediately shows, by Theorem 1, that the rank of the matrix remains unchanged.

Now, let us show that the row space remains unaltered. Let $R_1,\ldots R_m$ be the rows of a matrix A. Then the row space of A is generated by $\{R_1 \ldots . R_i \ldots R_j \ldots . R_m\}$. On applying R_{ij} to A, the rows of A remain the same. Only their order gets changed. Therefore, the row space of R_{ij}(A) is the same as the rows space of A.

If we apply $R_1(a)$, for $a \in F$, $a \neq 0$, then any linear combination of $R_1,\ldots,R_m$ is $a_1R_1 + \ldots + a_mR_m = a_1 + \ldots + \frac{a_i}{a} aR_i + \ldots + a_mR_m$, which is a linear combination of $R_1,\ldots,aR_i,\ldots,R_m$.

Thus, $|\{R_1\ldots R_i\ldots R_m\}| = [\{R_1\ldots aR_i\ldots R_m\}]$. That is, the row space of A is the same as the row space of $R_i(a)(A)$.

If we apply $R_{ij}(a)$, for $a \in F$, then any linear combination
$B_1R_1 + \ldots + b_iR_j + \ldots + b_iR_j .. + b_mR_m = b_1R_1 + .. + b_i(R_i + aR_j) + .. + (b_j - b_ia)R_j + \ldots + b_mR_m$.
Thus, $[\{R_1,\ldots,R_m\}] = [\{R_1,\ldots,R_i + aR_j,\ldots, R_j,\ldots,R_m\}]$.

Hence, the row space of A remains unaltered under any elementary row operations.

We can similarly show that the column space remains unaltered under elementary column operations.

Elementary operations lead to the following definition.

Definition: A matrix obtained by subjecting 1_n to an elementary row or column operation is called an elementary matrix.

For example, $C_{12}(I_3) = C_{12}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is an elementary matrix.

Since there are six types of elementary operations, we get six types of elementary matrices, but not all them are different.

E E9) Check that $R_{23}(I_4) = C_{23}(I_4), R_2(2)(1_4) = C_2(2)(I_4)$ and $R_{12}(3)(I_4) = C_{21}(3)(I_4)$

In general, $R_{ij}(I_n) = C_{ij}(I_n)$, $R_i(a)(I_n) = C_i(a)(I_n)$ for $a \neq 0$, and $R_{ij}(a)(I_n) = C_{ij}(a)(I_n)$ for $i \neq j$ and $a \in F$.

Thus, there are only three types of elementary matrices. We denote
$R_{ij}(I) = C_{ij}(I)$ by E_{ij},

$R_i(a)(I) = C_i(a)(I)$, (if $a \neq 0$) by $E_i(a)$ and
$R_{ij}(a)(I) = C_{ij}(a)(I)$ by $E_{ij}(a)$ for $i \neq j$, $a \in F$.

E_{ij}, $E_i(a)$ and $E_{ij}(a)$ are called the elementary matrices corresponding to the pairs R_{ij} and C_{ij}.
$R_i(a)$ and $C_i(a)$, $R_{ij}(a)$ and $C_{ij}(a)$, respectively.

Caution: $E_{ij}(a)$ corresponds to $C_{ij}(a)$, and not $C_{ij}(a)$.

Now, see what happens to the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$ if we multiply it on the left by

$$E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ We get

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix} = R_{12}(A)$$

Similarly, $AE_{12} = C_{12}(A)$.

Again, consider $E_3(2)A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$

$$= \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 4 & 2 & 0 \end{bmatrix} = R_3\,(2)(A)$$

Similarly $AE_3(2) = C_3(2)(A)$

Finally, $E_{13}\,(5)A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 15 \\ 2 & 1 & 10 \end{bmatrix}$

$= R_{13}\,(5)(A)$

But, $AE_{13}\,(5) = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 15 \\ 2 & 1 & 10 \end{bmatrix}$

$= C_{31}(5)(A)$.

What you have just seen are example of a general phenomenon. We will not state this general result formally. (Its proof is slightly technical, and so, we skip it.)

Theorem 4: for any matrix A

a) $R_{ij}(A). = E_{ij}\ A$
b) $R_i(a)(A) = E_i(a)A$, for $a \neq 0$.
c) $R_{ij}(a)(A). = E_{ij}\ (a)A$
d) $C_{ij}(A) = AE_{ij}$
e) $C_i(a)(A) = AE_i(a)$, for $a \neq\ )$
f) $C_{ij}(a)(A) = Ae_{ij}(a)$

In (f) note the change of indices i and j.

An immediate corollary to this theorem shows that all the elementary matrices are invertible (see Sec. 7.6).

Corollary: An elementary matrix is invertible. In fact,

a) $E_{ij}E_{ij} = I$,
b) $E_i(a\text{-}1)E_i(a) = I$, for $a \neq 0$.
c) $E_{ij}\ (-a)\ E_{ij}(a) = I$.

Proof: we prove (a) only and leave the rest to you (see E10).
Now, from Theorem 4,

$E_{ij}E_{ij} = R_{ij}) = R_{ij}(R_{ij}(I)) = I$, by definition of R_{ij}.

E E10) Prove (b) and (c) of the corollary above.

The corollary tells us that the elementary matrices are invertible and the inverse of an elementary matrix is also an element matrix of the same type.

E E11) Actually multiply the two 4 x 4 matrices E_{13} (-2) and $E_{13}(2)$ to get I_3.

And now we will introduce you to a very nice type of matrix, which any matrix can be transformed to by applying elementary operations.

3.4 Row – reduced Echelon Matrices

Consider the matrix

$$\begin{bmatrix} 1 & 0 & 9 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

In this matrix the three non-zero rows come before the zero row, and the first non-zero entry in each non-zero row is 1. Also, below this 1, are only zero. This type of matrix has a special name, which we now give.

Definition: An m x n matrix A is called a row-reduced echelon matrix if

a) The non-zero rows come before the rows,
b) In each non-zero row, the first non-zero entry is 1, and
c) The first non-zero entry in every non-zero row (after the first row) is to the right of the first non-zero entry in the preceding row.

Is $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix}$ a row-reduced echelon matrix? Yes. It satisfies all the conditions of the definition. On the other hand are not row-reduce echelon matrices, since they violate conditions (a), 9b) and (c), respectively.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

The matrix

$$\begin{bmatrix} 0 & 1 & 3 & 4 & 9 & 7 & 8 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 5 & 6 & 10 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 7 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is a 6 x 11 row-reduced echelon matrix. The dotted line in it s to indicate the step-like structure of the non-zero rows.

But, why bring in this type of a matrix? Well the following theorem gives us one good reason.

Theorem 5: The rank of a row-reduced echelon matrix is equal to the number of its non-zero rows.

Proof: Let $R_1, R_2, \ldots R_r$ be the non-zero rows of an m x n row-reduced echelon matrix, E. Then RS(E) is generated by $R_1, \ldots R_r$. We want to show that $R_1, \ldots, R_r$ are linearly independent. Suppose R_1 has its first non-zero entry in column k_1, R_2 in column k_2, and so on. Then, for any r scalars $c_1, \ldots c_r$ such that c_1 R_1 that c_1, $R_1 + C_2R_2 + \ldots + c_rR_r = 0$ we immediately get

$$
\begin{array}{llllll}
 & k_1 & k_2 & k_r & & \\
 & \downarrow & \downarrow & \downarrow & & \\
c_1 & [0, \ldots\ldots, 0. & 1, \ldots * \ldots, & \ldots * \ldots\ldots, & , * \ldots\ldots &]* \\
+ c_2 & [0, \ldots\ldots\ldots\ldots\ldots\ldots & \ldots, 0, 1, & \ldots * \ldots\ldots, & , * \ldots\ldots &]* \\
 & \vdots\ \vdots & \vdots & & \vdots & \\
+ c_r & [0, \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots & \ldots, 0, 1, & \overset{*}{\ldots\ldots\ldots} &]^{*} \\
= & . [0, \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots & 0]
\end{array}
$$

where * denotes various entries that aren't bothering to calculation.

This equation gives us the following equations (when we equate the k_1th entries, the k_2th entries,…., the k_rth entries on both sides of the equation):

$c_1 = 0 = c_1(*) + c_2 = 0 \ldots, c1(*) + c_2(*) + \ldots + _{r-1}(*) + c_r = 0$.

On solving these equations we get

$c_1 = 0 = c_2 \ldots = c_r. \therefore R_1 \ldots, R_r$ are linearly independent $\therefore p(E) = r$.

Not only is it easy to obtain the rank of an echelon matrix, one can also solve linear equations of the type AX =B more easily if A is in echelon form.

Now here is some good news!

Every matrix can be transformed to the row echelon form by a series of elementary row operations. We say that the matrix is reduced to the echelon form. Consider the following example.

Example 4: let A = $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 2 & 4 & 1 & 10 & 2 \end{bmatrix}$

Reduce a to the row echelon form.

Solution: the first column of a is zero. The second is non-zero. The (1,2)th element is 0. we want 1 at this position. We apply R12 to A and get.

$$A_1 = \begin{pmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 2 & 4 & 1 & 10 & 2 \end{pmatrix}$$

The (1,2)th, entry has become 1. Now we subtract multiples of the first row from other rows so that the (2,2)th, (3,2)th (4,2)th and (5,2)th entries become zero. So e apply R_{ij} (-1) and R_{51})-2) and get

$$A_2 = \begin{pmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 12 & 2 \end{pmatrix}$$

Now, beneath the entries of the first row we have zeros in the first 3 columns, and in the fourth column we find non-zero entries. We want 1 at the (2,4)the position, so we interchange the 2nd 3rd rows. We get

$$A_3 = \begin{pmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 12 & 0 \end{pmatrix}$$

We now subtract suitable multiples if the 2nd row from the 3rd, 4th and 5th rows so that the (3,4)th (4,4(th (5,4)th entries all become zero. ∴

$$A_3 \overset{R_{42}^{(-1)},\ R_{42}^{(-3)}}{\sim} \begin{pmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$A \overset{R}{\sim} B$ means that on apply the operation R to A we get matrix B.

Now we have zero below the entries of the 2nd row, except for the 6th column. The (3,6)th element is 1. We subtract suitable multiples of the 3rd row from the 4th and 5th rows so that the (4,6)th elements become zero. ∴

$$A_4 \overset{R_{43}^{(-1)}}{\sim} \begin{pmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Linear transformations and Matrices

And now we have achieved a row echelon matrix. Notice that we applied 7 elementary operations to a to obtain this matrix.

In general, we have the following theorem.

Theorem 6: Every matrix can be reduced to a row-reduced echelon matrix by a finite sequence of elementary row operations.

The proof of this result is just a repetition of the process that you went through in Example 4 for practice, we give you the following exercise.

E E12) Reduce the matrix $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix}$ to echelon form.

Theorem 6 leads us to the following definition.

Definition: If a matrix A is reduced to a row-reduce echelon matrix E by a finite sequence of elementary row operations then E is called a **row-reduced echelonform** (or, the row echelon form) of A. We now give a useful result that immediately follows from Theorem 3 and 5.

Theorem 7: Let E be a row-reduced echelon form of A. Then the rank of A = number of non-zero rows of E.

Proof: We obtain E from A by applying elementary operations. Therefore by Theorem 3. p(A) = p(E). Also. P(E) = the number of non-zero rows of E, by Theorem 5.

Thus, we have proved the theorem.

Let us look at some examples to actually see how the echelon form of a matrix simplifies matters.

Example 5: Find p(A), where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 6 \end{bmatrix}$$

by reducing it to its row-reduced echelon form.

Solution: $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 6 \end{bmatrix} \overset{R_{21}(-1)}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 3 \end{bmatrix} \overset{R_2(1/3)}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix}$

which is the desired row-reduced echelon form. This has 2 non-zero rows. Hence, p(A) =

E E13) Obtain the row-reduced echelon form of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & 5 \\ 2 & 1 & 7 & 6 \\ 4 & 5 & 7 & 10 \end{bmatrix}$$

Hence determine the rank of the matrix.

By now must have got used to obtaining row echelon forms. Let us discuss some ways of applying this reduction.

3.6 Applications Of Row-Reduction

In this section we shall see how to utilize row-reduction for obtaining the inverse of a matrix, and for solving a system of leaner equations.

3.7 Inverse of a Matrix

In Theorem 4 you discovered that applying a row transformation to a matrix a is the same as multiplying it on the left by a suitable elementary matrix. Thus applying a series of row transformations to A is the same as pre-multiplying A by a series of elementary matrices. This means that after the nth row transformation we obtain the matrix $E_nE_{n-1}\ldots E_2 E_1A$. Where $E_1E_2,\ldots.E_n$, are elementary matrices.

Now, how do we use this knowledge for obtaining the inverse of an invertible matrix? Suppose we have an n x n invertible matrix A. We know that A = 1A. Where $1 = 1_n$. Now, we apply a series of elementary row operations $E_1,\ldots E$ to A so that A gets transformation to 1_n

Thus,

$1 = E_sE_{s-1}\ldots E_2E_1A = E\ E_{s-1}\ldots\ E_2E_1\ (1A)$

$= (E\ E_{s-1}\ldots E_2E_1 1)A = BA$

where $B = E\ \ldots.\ E_1I$. Then B is the inverse of A!

Note that we are reducing A to 1 and not only to the echelon form.

We illustrate this below.

Example 6: Determine if the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

is invertible. If it is invertible, find its inverse.

Solution: Can we transform A to 1? If so, then A will be invertible.

$$\text{Now, } A = IA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

To transform A we will be pre-multiplying it by elementary matrices. We will also be pre-multiplying 1A by these matrices. Therefore, as a is transformed to 1, the same transformations are done to 1 on the right hand side of the matrix equation given above. Now

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} A \text{ (applying } R_{21}(-2) \text{ and } R_{31}(-3) \text{ to A)}$$

$$\Rightarrow \quad = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 5 & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{bmatrix} A \text{ (applying } R_2(-1) \text{ and R3(-1))}$$

$$\Rightarrow \quad = \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & -18 \end{bmatrix} \begin{bmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -7 & 5 & -1 \end{bmatrix} A \text{ (applying } R_{12}(-2) \text{ and } R_{32}(-5))$$

$$\Rightarrow \quad = \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ 7/18 & -5/18 & 1/19 \end{bmatrix} A \text{ (applying } R_3(-1/18))$$

$$\Rightarrow \quad = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -5/18 & 1/18 & 7/18 \\ 1/18 & 7/18 & -5/18 \\ 7/18 & -5/18 & 1/18 \end{bmatrix} A \text{ (applying } R_{13}(7) \text{ and } R_{23}(-5))$$

$$\text{Hence, A is invertible and its inverse is } B = 1/18 \begin{bmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{bmatrix}$$

To make sure that we haven't made a careless mistake at any stage, check the answer by multiplying B with A. your answer should be 1.

E E14) show that $\begin{bmatrix} 0 & 1 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 7 \end{bmatrix}$ is invertible. Find its inverse.

Let us now look at another application of row-reduction.

3.7 Solving a System of Linear Equations

Any system of m linear equations, in n unknowns x1….x_n, is

$$\begin{array}{ccc} a_n x_1 + \ldots + & a1_n x_n & = b_1 \\ \vdots & \vdots & \vdots \\ a_{m1} x_1 + \ldots + & a1_{mn} x_n & = b_m \end{array}$$

where all the a_{ij} and bi are scalars

This can be written in matrix form as

$$AX = B., \text{ where } a = [a_{ij}],\ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},\ B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

If B = 0, the system is called homogenous. In this situation we are in a position to say how many linearly independent solutions the system of equations has.

Theorem 8: The number of linearly impendent solutions of the matrix equation AX = 0 is n-r, where A is an mxn matrix and r = p(A).

Proof: In Unit 7 you studied that given the matrix A, we can obtain a linear transformation T: $F^n \rightarrow F^m$ such that $[T]_{B.B} = A$, where B and B′ are bases of F^n and F^m, respectively.

Now, X = $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a solution of AX=0 if and only if it lies in ker T(since T(X)=AX)

Thus, the number of linearly independent solutions is dim ker T nullity (T) = n-rank (T) (Unit 5. Theorem 5.)

Also, rank (T) = p(A) (Theorem 2)

Thus, the number of linearly independent solutions is n-p(A).

This theorem is very useful for finding out whether a homogeneous system has any non-trivial solution or not.

Example 7 consider the system of 3 equations in 3 unknowns:

$$3x - 2y + z = 0$$
$$x + y = 0$$
$$x - 3z = 0$$

How many solutions does it have which are linearly independent over R?

Solution: Here our coefficient matrix, A = $\begin{bmatrix} 3 & -2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

Thus, n = 3. We have to find r. For this, we apply the row-reduction method.

We Obtain A ~ $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$, which is in echelon form and has rank 3.

Thus, p(A) = 3.

Thus, the number of linearly independent solutions is 3 – 3 = 0. This means that this system of equation has no non-zero solution.

In Example 7 the number of unknowns was equal to the number of equations, that is, n = m. What happens if n>m?

A system of m homogeneous equations in n unknowns has a non-zero solution if n>m, why? Well in n>m, then the rank of the coefficient matrix is less than or equal to m, and hence, less than. So n-r> 0 Therefore, at least one non-zero solution exists.

Note: If a system AX = 0 has one solution, X_0 then it has an infinite number of solutions of the form cX_0, $c \in F$. This is because $AX_0 = 0 \Rightarrow A(cX_0) = 0 \ \forall c \in F$.

E E15) Give a set of linearly independent solutions for the system of equations

$$x + 2y + 3z = 0$$
$$2x + 4y + z = 0$$

Now consider the general equation AX = B, where A is an mxn matrix. We form the augmented matrix |AB|. This is an m x (n+1) matrix whose last column is the matrix B. Here, we also include the case B = 0.

Interchanging equations, multiplying an equation by a non-zero scalar, and adding to any equation a scalar times some other equation does not alter the set of solutions of the system of equations. In other words, if we apply elementary row operations on |AB| then the solution set does nit change.

The following result tells us under what conditions the system AX = B has a solution.

Theorem 9: The system of linear equations given by the matrix equation AX = B has a solution if P(A) = p([AB]).

Proof: AX = B represents the system.

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \ldots + a1_n x_n = b_1 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a1_{mn}x_n = b_m \end{matrix}$$

This is the same as

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n - b_1 = 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a1_{mn}x_n - b_m = 0 \end{matrix}$$

which is represented by $|AB| \begin{bmatrix} X \\ -1 \end{bmatrix} = 0$. Therefore, any solution of AX = B is also a solution of $|AB| \begin{bmatrix} X \\ -1 \end{bmatrix} = 0$ and vice versa. By Theorem 8; this system has a solution if and only if n + 1> p([AB]).

Now, if the

Equation $|AB| \begin{bmatrix} X \\ -1 \end{bmatrix} = 0$ has a solution, say $\begin{bmatrix} c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}$, then $c_1C_1 + c_2C_2 + \ldots + c_nC_n = B$, where

$C_1 \ldots C_n$ are the columns of A. That is, B is a linear combination of the C_i's, ∴, RS ([AB]) = RS(A), ∴ p([AB]).

Conversely, if p([AB]), then the number of linearly independent columns of A and |AB| are the same. Therefore, B must be a combination of the columns $C_1, \ldots C_n$ of A.

Let $B = a1C1 + \ldots + a_nC_n$ $a1 \in F$ i. ∀

Then a solution of AX = B is $X = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

Thus, AX = B has a solution if and only if p(A) = p(|AB|).

Remark: If A is invertible then the system AX=B has the unique solution X= A^{-1} B.

Now, once we know that the system given by AX = B is consistent, how do we find a solution? We utilize the method of successive (or Gaussian) elimination. This method is attributed to the famous German mathematician, Carl Friedrich Gauss (1777-1855) (see Fig. 1). Gauss was called the "prince of mathematicians" by his contemporaries. He did a great amount of work in pure mathematics as well as probability theory of errors, geodesy, mechanics, electro-magnetism and optics.

To apply the method of Gaussian elimination, we first reduce |AB| to its row echelon form.

E. Then, we write out the equations E $\begin{bmatrix} X \\ -1 \end{bmatrix} = 0$ and solve them, which is simple.
Let us illustrate the method.

Example 8: Solve the following system by using the Gaussian elimination process.
x + 2y + 3z = 1
2x + 4y + z = z

Solution: the given system is the same as

$$\begin{bmatrix} x \\ y \\ z \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Fig. 1: Carl Friedrich Gaus

We first reduce the coefficient matrix to echelon form.

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 1 & 2 \end{bmatrix} \underset{\sim}{R1(-2)} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & -5 & 0 \end{bmatrix} \underset{\sim}{} R(-1/5) \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

This gives us an equivalent to x = 2y and z = 0.
x +2y + 3z = 1 and z = 0.

These are again equivalent to x = 1 – 2y and z = 0.
We get the solution in terms of a parameter. Put y = α. Then x = 1 - 2α, y = α, z = 0 is a solution, for any scalar α, Thus, the solution set is {(1 - 2$\alpha\alpha$ 0) | $\alpha \in$R}.

Now let us look at an example where B =), that is, the system is homogeneous

Example 9: Obtain a solution a solution set of the simultaneous equations.
x + 2y + 5t = 0
2x + y + 7z +61 = 0
4x +5y + 7z + 16t = 0

Solution: The matrix of coefficients is

$$A = \begin{pmatrix} 1 & 2 & 0 & 5 \\ 2 & 1 & 7 & 6 \\ 4 & 5 & 7 & 16 \end{pmatrix}$$

The given system is equivalent to AX = 0. A row-reduced echelon form of this matrix is

$$\begin{pmatrix} 1 & 2 & 0 & 5 \\ 0 & 1 & -7/3 & 4/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then the given system is equivalent to

$$\left.\begin{array}{l} x + y + 5t = 0 \\ y - (7/3)\,z + (4/3)t = 0 \end{array}\right\} \Rightarrow \begin{array}{l} x = (-\,14/3)z - (7/3)t \\ y = (7?3)z - (4/3)t \end{array}$$

which is the solution in terms of z and t. Thus, the solution set of the given system of equations, in terms of two parameters α and β, is

$\{((\,-\,14/3)\alpha - (7/3)\,\beta, (7/3)\,\alpha - (4/3)\,\beta, \alpha, \beta) \mid \alpha, \beta \in \mathbf{R}\}$

This is a two-dimensional vector subspace of $\mathbf{R}^4$ with basis

$\{(1\text{-}14/3, 7/3, 1, 0), (-7/3, -4/3, 0, 1)\}$

For practice we give you the following exercise.

E E16) Use the Gaussian method to obtain solution sets of the following system of equations.

$4x_1 - 3x_2 + x_3 - 7 = 0$

$x_1 - 2x_2 - 2x_3 - = 0$

$3x_1 - x_2 = 2x_3 + 1 = 0$

And now we are near the end of this unit.

4.0 CONCLUSION

5.0 SUMMARY

In this unit we covered the following points.

- We defined the row rank, column rank and rank of a mtrix, and showed that the are equal.
- We proved that the rank of a linear transformation is equal to the rank of its matrix.
- We defined the six elementary row and column operations.
- We have shown you how to reduce a matrix to the row-reduced echelon form.
- We have used the echelon form to obtain the inverse of a matrix.
- We proved that the number of linearly independent solutions of a homogenous system of equations given by the matrix equation AX = 0 is n-r, where r = rank of A n = number of columns of A.
- We proved that the system of linear equations given by the matrix equation AX = B is consistent if and only if p(A) = p([AB]).
- We have shown you how to solve a system of linear equations by the process of successive elimination of variables, that is, the Gaussian method.

Solutions/Answers

E1) A is the mxn zero matrix $\Leftrightarrow$ RS (A) = {0}. $\Leftrightarrow p_r(A) = 0$.

E2) The column space of A is the subspace of R^2 generated by (1,0), (0,2), (1,1). Now $\dim_R CS(A) \leq \dim_R R^2 = 2$. Also (1,0) and (0,2), are linearly independent.

$\therefore$ {(1,0),(0,2)} is a basis of CS(A), and pc(A) = 2.

The row space of A is the subspace of R^3 generated by (1,0,1) and (0,2,1). These vectors are linearly independent and hence, form a basis of RS(A). $\therefore p_r(A) = 2$.

E3) The ith row of C = AB is

$[c_{i1}\ c_{i2} \ldots c_{ip}]$

$$= \left[\sum_{k=1}^{n} a_{ik}\, b_{k1} \quad \sum_{k=1}^{n} a_{ik} b_{k2} \ \ldots\ldots \sum_{k=1}^{n} a_{ik} b_{kp}\right]$$

= ai1 [b11 b12 ..n1p] + ai1 [b21 b22…b2p] +…. + ain [bn1 bn2…bnp], a linear combination of the rows of B. $\therefore$ Rs.(AB)$\subseteq$RS(B) $\therefore$ pr (AB) $\leq$ pr(B).

E4) By Lemma 1, p(AB) $\leq$ pc(A) = p(A)

Also p(AB) $\leq$ pr(B) = p(B).

$\therefore$ p(AB) $\leq$ min (p(A), p(B)).

E5) $p(CR) \leq \min(p(R))$

But $p(C) \leq \min(m,1) = 1$. Also $C \neq 0$. $\therefore p(C) = 1. \therefore p(CR) \leq 1$.

Now, if c = $\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$, r = $[B_1, \ldots\ldots, B_n]$, then

$$\begin{matrix} a_1b_1 & a_1b_2+\ldots+ & a_1b_n \\ a_2b_1 & a_2b_2+\ldots+ & a_2b_n \\ \vdots & \vdots & \vdots \\ a_mb_1 +\ldots+ & a_mb_2 = & a_mb_m \end{matrix}$$

Since $C \neq 0$, $a_i \neq 0$, for some i. Similarly $b_j \neq 0$ for some j. $\therefore a_i b_j \neq 0$.

$\therefore CR \neq 0$.

$\therefore p(CR) \neq 0. \therefore p(CR) = 1$.

E6) PAQ = $\begin{bmatrix} 0 & -2 & -2 \\ -3 & -4 & -3 \end{bmatrix}$ The rows of PAQ are linearly **independent**. $\therefore p(PAQ) = 2$. Also the rows of a are linearly independent. $\therefore p(PAQ) = p(A)$.

E7)Let A = $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ Then $p(A) = 3. \therefore$ A's normal form is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

b) $R_{32} \text{ o } R_{21}(A) = R_{32} \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

c) $\begin{matrix} 0 + 0 \text{ x })- 1) & 0 + 1 \text{ x } (- 1) & 1 + 0 \text{ x } (- 1) \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

E9) R23 (14)= $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C_{23}(1_4)$

$R_2(2)(1_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C_2(2)(1_4)$

$$R_{12}(3)\,(1_4) = \begin{bmatrix}1&3&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = C_{21}\,(3)\,(1_4)$$

E10) $E_i\,(a^{-1})\,E_i(a) = R_i(a^{-1})\,(E_i(a)) = R_i(a^{-1})\,R_i(a)\,(1) = 1.$
This proves (b)

$E_{ij}\,(-a)\,(E_{ij}\,(a) = R_{ij}\,(-a)\,(E_{ij}(a)) = R_{ij}\,(-a)\,(1)) = 1$, providing (c).

E11 E13 (-2) E13 (2) = $\begin{bmatrix}1&0&-2&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&2&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$

E12) $\begin{bmatrix}1&2&0\\0&1&0\\3&1&0\end{bmatrix} \; R_{31}(-3) \sim \begin{bmatrix}1&2&0\\0&1&0\\0&-5&0\end{bmatrix} \; R_{32}(5) \sim \begin{bmatrix}1&2&0\\0&1&0\\0&0&0\end{bmatrix}$

E13) $\begin{bmatrix}1&2&0&5\\2&1&7&6\\4&5&7&10\end{bmatrix} \; R_{21}(-2),\, R_{31}(-4) \sim \begin{bmatrix}1&2&0&5\\0&-3&7&-4\\0&-3&7&10\end{bmatrix} \; R_{32}(5)$

R2(-1/3) $\begin{bmatrix}1&2&0&5\\0&1&-7/13&4/3\\0&-3&7&-10\end{bmatrix}$

$R_{32}\,(3) \sim \begin{bmatrix}1&2&0&5\\0&1&-7/3&4/3\\0&0&0&-6\end{bmatrix} \; R_3(-1/6) \sim \begin{bmatrix}1&2&0&5\\0&1&-7/3&4/3\\0&0&0&1\end{bmatrix}$

$\therefore \rho(A) = 3$

E14) A = $\begin{bmatrix}0&1&3\\2&3&5\\3&5&7\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} A$

$\Rightarrow \begin{bmatrix}0&1&3\\2&3&5\\3&5&7\end{bmatrix} = \begin{bmatrix}0&1&3\\2&3&5\\3&5&7\end{bmatrix}$ A (applying R_{12}

$\Rightarrow \begin{bmatrix}1&3/2&5/2\\0&1&3\\0&1/2&-1/2\end{bmatrix} = \begin{bmatrix}1&1/2&0\\1&0&0\\0&-3/2&1\end{bmatrix}$ A (applying $R_1(1/2)$, $R_{31}\,(-3)$)

$$\Rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} -3/2 & 1/2 & 0 \\ 1 & 0 & 0 \\ -1/2 & -3/2 & 1 \end{bmatrix} A \text{ (applying } R_{12}(-3/2), R_{32}(-1/2))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ 1/4 & -9/4 & 3/2 \\ 1/4 & 3/4 & -1/2 \end{bmatrix} A \text{ (applying } R_3(-1/2), R_{23}(-3) \text{ and } R_{13}(2))$$

$\therefore$ A is invertible, and $A^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 1/4 & -9/4 & 3/2 \\ 1/4 & 3/4 & -1/2 \end{bmatrix}$

E15) The given system is equivalent to

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Now, the rank of $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$ is 2. $\therefore$, the number of linear independent solutions is 3 – 2 = 1. $\therefore$ any non-zero solution will be a linearly independent solution. Now, the given equations are equivalent to

x + 2y = - 3z ….. (1)
2x + 4y = - z …. (2)

(- 3) times Equation (2) added to Equation (1) gives -5x – 10y = 0.

$\therefore$ x = -2. Then (1) gives z = 0. Thus, a solution is (-2, 1, 0) $\therefore$, a set of linearly independent solutions is {(-2,1,0)}.

Note that you can get several answers to this exercise. But any solution will be
α (-2, 1,0), for some $\alpha \in$ R.

E16) The augmented matrix is [AB].

$$= \begin{bmatrix} 4 & -3 & 1 & 7 \\ 1 & -2 & -2 & 3 \\ 3 & -1 & 2 & -1 \end{bmatrix}$$ Its row-reduced echelon form is

$$\begin{bmatrix} 1 & -2 & -21 & 3 \\ 0 & 1 & 9/5 & -1 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

Thus, the given system of equations is equivalent to

$x1 - 2x_2 - 2x_3 = 3$
$x_2 + (9/5)\, x_3 = -1$
$x_3 = 5.$

We can solve this system to get the unique solution $x_1 = -7$, $x_2 = -10_0$ $x_3 = 5$.

Eigenvalues And Eigenvectors

This section consist of three units in which we first introduce you to the theory of determinants, and give its applications in solving systems of linear equations.

The theory of determinants was originated by Leibniz in 1693 while studying systems of simultaneous linear equations. The mathematician Jacobi was perhaps the most prolific contributor to the theory of determinants. In fact, there is a particular kind of determinant that is named Jacobian, after him. The mathematicians Cramer and Bezout used determinants extensively for solving systems of linear equations.

In Unit 5 we have given a self-contained treatment of determinants, including the standard properties of determinants. We have also given formula for obtaining the inverse of a matrix, and have explained Cramer's Rule. We end this unit by discussing the determinant rank.

In unit 6 we discuss eigenvalues and eigenvectors. Their use first appeared in the study of quadratic forms. The concepts that you will study in this unit were developed by Arthur Cayley and others during the 1840s. What you will discover in the unit is the algebraic eigenvalue problem and methods of finding eigenvalues and linearly independent eigenvectors.

In unit 7 we introduce you to the characteristic polynomial. We give a proof of the Cayley-Hamilton theorem and give its applications. We also discuss the minimal polynomial of a matrix ad of a linear transformation.

If you are interested in knowing more about the material covered in this block, you can refer to the books listed in the course introduction. These books will be available at your study centre.

Notations And Symbols

Mn(F)	set of all nxn matrices over F
Vn(F)	Mnx1 (F)
Det (A), |A|	determinant of the matrix A
πp a1	the product of ais such that I satisfies property P
det (T)	determinant of the linear operator T
Adj (A)	adjoint of the matrix A
Tr (A)	trace of the mnatrix A
Wλ	eigenspace corresponding to the eigenvalue λ

UNIT 5 DETERMINANTS

Introduction

In Unit 4 we discussed the successive elimination method for solving a system of linear equations. In this unit we introduce you to another method, which depends on the concept of a determinant function. Determinants were used by the German mathematician Leibniz (1646 – 716) and the Swiss mathematician Vandermonde (1735-1796) gave the first systematic presentation of the theory of determinants.

There are several ways of developing the theory of determinants. In section 5.2 we approach it in one way. In section 5.3 you will study the properties of determinants and certain other basic facts about them. We go on to give applications in solving a system of linear equations (Cramer's Rule) and obtaining the inverse of a matrix. We also define the determinant of a linear transformation. We end with discussing a method of obtaining the rank of a matrix.

Throughout this unit F will denote field of characteristic zero (M_n (F) will denote the set of n x n matrices over F and V_n (F) will denote the space of all n x 1 matrices over F, that is,

$$V_n(F) = \left\{ X = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \middle| \; a_i \in F \right\}$$

The concept of a determinant must be understood properly because you will be using it again and again. Do spend more time on section 5.2, if necessary. We also advise you to revise unit 1-4 before starting this unit.

2.0 OBJECTIVES

After completing this unit, you should be able to

- Evaluate the determinant of a square matrix; using various properties of determents;
- Obtain the adjoint of a square matrix;
- Compute the inverse of an invertible matrix, using its dajoint;
- Apply Cramer's rule to solve a system o f linear equations;
- Evaluate the determinant of a linear transformation;
- Evaluate the rank of a matrix by using the concept of the determinant rank.

3.0 MAIN CONTENT

3.1 Defining Determinants

There are many ways of introducing and defining the determinant function from M_n (F) to F. In this section we give one of them, the classical approach. This was given by the French mathematician Laplace (1749- 1827), and still very much in use.

We will define the determinant function det: M_n (F) $\to$ F by induction on n. That is, we will define it for n = 1,2,3, and then define it for any n, assuming the definition for n-1.

When n = 1, for any A $\in M_1(F)$ we have a = [a], for some a $\in$F. In this case we define det (A) = det ([a]) = a.

For example, det ([5]) = -5.

When n = 2, for any A = $\begin{matrix} a_{11} & a12 \\ a_{21} & a_{22} \end{matrix}$ $\in M_2(F)$, we define.

For example det $\left(\begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}\right)$ a13 = 0 x 3 – 1 x (- 2) = 2.

When n = 2, for any a = $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in M_3$ (F), we define

det (A) using the definition for the case n = 2 as followings:

det (A) = a det $\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right)$ $-a_{12}$ det $\left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}\right)$ $+a_{13}$ det $\left(\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}\right)$

That is, det (a) = $(-1)^{1+1}$ a_{11} (det of the matrix left after deleting the row and column containing a_{11}) + $(-1)^{1+2}$ a_{12} (det of the matrix left after deleting left row and column containing a_{12}) + $(-1)^{1+2}$ (det of the matrix left after deleting the row and column containing a_{13}

Note that the power of (-1) that is attached to a_{11}, is 1+ j for j = 1+ j for j = 1,2,3.

So, det (A)=a_{11} (a_{22} a_{33} – a_{23} a_{32}) – a_{12} (a_{21} a_{33} – a_{23} a_{31}) + a_{13} (a_{21} a_{32}– a_{22} a_{31}).

In fact, we could have calculated | A | from the second row also as follows:

$$|A| = (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a^{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{2+2} a_{22} \begin{vmatrix} a^{11} & a^{13} \\ a31 & a33 \end{vmatrix} + (-1)^{2+3} a1_{23} \begin{vmatrix} a^{11} & a^{12} \\ a31 & a32 \end{vmatrix}$$

Similarly, expanding by the third row, we get

$$|A| = (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a^{13} \\ a_{22} & a_{23} \end{vmatrix} + (-1)^{3+2} a_{32} \begin{vmatrix} a^{11} & a^{13} \\ a_{21} & a_{23} \end{vmatrix} + (-1)^{3+3} a_{33} \begin{vmatrix} a^{11} & a^{12} \\ a21 & a22 \end{vmatrix}$$

All 3 ways of obtaining | A | lead to the same value.

Consider the following example.

Example 1: Let

$$A = \begin{bmatrix} 1 & 2 & 6 \\ 5 & 4 & 1 \\ 7 & 3 & 2 \end{bmatrix} \quad \text{Calculate } |A|$$

Solution: we want to obtain

$$|A| = \begin{vmatrix} 1 & 2 & 6 \\ 5 & 4 & 1 \\ 7 & 3 & 2 \end{vmatrix}$$

let A_{ij} denote the matrix obtained by deleting the ith row and jth column of A let us expand by the first row. Observe that

$$A_{11} = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix},\ A12 = \begin{bmatrix} 5 & 1 \\ 7 & 2 \end{bmatrix},\ A13 = \begin{bmatrix} 5 & 4 \\ 7 & 3 \end{bmatrix}$$

Thus,

$$|A_{11}| = \begin{vmatrix} 4 & 1 \\ 3 & 2 \end{vmatrix} = 4 \text{ x } 2 - 1\text{x } 3 = 5,\ |A_{12}| = \begin{vmatrix} 5 & 1 \\ 7 & 2 \end{vmatrix} = 5 \text{ x } 2 - 1 \text{ x } 7 = 3,\ |A_{13}| = \begin{vmatrix} 5 & 4 \\ 7 & 3 \end{vmatrix}$$

$$= 5 \text{ x } 3 - 4 \text{ x } 7 = -13.$$

Thus,

$$|A| =(-1)^{1+1} X\ 1 X\ |A_{11}| + (-1)1+2\ X 2 X\ |A_{12}| + (-1)^{1+3} X 6\ X\ |A_{13}| = 5 – 6 – 8 = - 79.$$

E E1) Now obtain A of Example 1, by expanding by the second row, and the third row does the value of A depend upon the row used for calculating it?

Now, let us see how this definition is extended to define det (A) for any n x n matrix A, n ≠ 1.

$$\text{When } A = \begin{array}{cccc} A_{11} & a_{12} & \ldots & a1n \\ a_{21} & a_{22} & \ldots & a_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{11} & a_{12} & \ldots & a_{in} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{array}$$

The ith row as follows:

$\det (A) = (-1)^{1+1} a_{11} \det (A_{11}) + (- 1)^{1+2} a_{12} \det (A_{12}) + \ldots + (- 1)^{1+n} a_{in} \det (A_{in})$, where A_{ij} is the (n – 1) x (n – 1) matrix obtained from A by deleting the ith row and the jth column, and I is a fixed integer with $1 \leq i \leq n$.

We, thus, wee that $\det (A) = \sum_{j=1}^{n} (- 1)^{i+j} a_{ji} \det (A_{ji})$,

define the determinant of an n x n matrix A in terms of the determinants of the (n-1) x (n-1) matrices a_{ij}, j = 1,2, …………, n.

Note: while calculating |A| , we prefer to expand along a row that has the maximum number of zeros, This cuts downs the number of terms to be calculated.

The following example will help you to get used to calculating determinants.

Example 2: Let

$$A = \begin{vmatrix} -3 & -2 & 0 & 2 \\ 2 \quad 1 & 0 & -1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & -3 & 1 \end{vmatrix}$$ Calculate |A |

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

The first three rows have one zero each. Let us expand along third row. Observe that $a_{32} = 0$. So we don't need to calculate A_{32}. Now,

$$A_{31} = \begin{bmatrix} -2 & 0 & 2 \\ 1 & 0 & 1 \\ 1 & -3 & 1 \end{bmatrix},\ A_{33} = \begin{bmatrix} -3 & -2 & 2 \\ 2 & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix},\ A_{34} = \begin{bmatrix} -3 & -2 & 0 \\ 2 & 1 & 0 \\ 2 & 1 & -3 \end{bmatrix}$$

We will obtain $|A_{31}|$, $|A_{33}|$, and $|A_{34}|$ by expanding along the second, third and second row, respectively.

$$\therefore,\ |A31| = \begin{vmatrix} -2 & 0 & 2 \\ 1 & 0 & -1 \\ 1 & -3 & 1 \end{vmatrix}$$

$$= (-1)^{2+1}.1.\begin{vmatrix} 0 & 2 \\ -3 & 1 \end{vmatrix} + (-1)^{2+2}.0.\begin{vmatrix} -2 & 2 \\ 1 & 1 \end{vmatrix} + (-1)^{2+3}.(-1).\begin{vmatrix} -2 & 0 \\ 1 & -3 \end{vmatrix}$$

(expansion along the second row)

$$= (-1).6 + 0 + (-1)(-1).6$$
$$= -6 + 6 = 0.$$

$$|A_{33}| = \begin{vmatrix} -3 & -2 & 2 \\ 2 & 1 & -1 \\ 2 & 1 & 1 \end{vmatrix} = (-1)^{3+1}.2.\begin{vmatrix} -2 & 2 \\ 1 & -1 \end{vmatrix} + (-1)^{2+3}.1.\begin{vmatrix} -3 & -2 \\ 2 & 1 \end{vmatrix}$$

$$+ 9-1)^{3+3}.1.\begin{vmatrix} -3 & -2 \\ 2 & 1 \end{vmatrix}$$ (expansion along the third row)

$$= 1.2.0 + (-1).1.(-1) + 1.1.1$$

$= 1 + 1 = 2.$

$$|A_{34}| = \begin{vmatrix} -3 & -2 & 0 \\ 2 & 1 & 0 \\ 2 & 1 & -3 \end{vmatrix} = (-1)^{2+1}.2. \begin{vmatrix} -2 & 0 \\ 1 & -3 \end{vmatrix} + (-1)^{2+2}\ .1. \begin{vmatrix} -3 & 0 \\ 2 & -3 \end{vmatrix}$$

$$+ (-1)\ 2+3\ .0. \begin{matrix} -3 & -2 \\ 2 & 1 \end{matrix} \text{ (expansion along the second row)}$$

$= (-1) .2. 6 + 1.1.9 + 0$
$= -12 + 9 = -3.$

Thus, the required determinant is given by

$|A| = a_{31} |A_{31}| -a_{32} |A_{32}| + a_{33} |A_{33}| - a_{34} |A_{34}|$
$= 1.0 – 0 + 1.2 -2.)-3) = 8.$

E E2) Calculate |At|, where A is the matrix in
a) Example 1,
b) Example 2.

At this point we mention that there are two other methods of obtaining determinants – via permutations and via multilinear forms. We will not be doing these methods here. Fr purposes of actual calculation of determinants the method that we have given is normally used. The other methods are used to prove various properties of determinants.

So far we have looked at determinant algebraically only. But there is a geometrical interpretation of determinants also, which we now give.

Determinant as area and volume: Let $u =(a_1, a_2)$ and $v = (b_1,b_2)$ be two vectors in R^2. Then, the magnituted of the area of the parallelogram spanned by u and v (see fig. 1) is the absolute

value of det (u,v) = $\begin{matrix} a_1 & b_1 \\ a_2 & b_2. \end{matrix}$

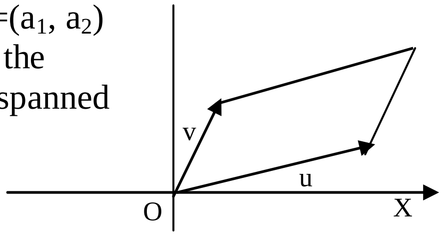

Fig. 1: the shaded area is det (u, v)

In fact, what we have just said is true for any n > 0. Thus, if u_1, u_2, ….u_n are n vectors in

R^n, then the absolute of det ($u_1 u_2$…., un) is the magnitude of the volume of the n-dimensional box spanned by u_1, u_2, …. u_n.

Try this exercise now.

E E3 What is the magnitude of the volume of the box in R^3 spanned by I, j and k?

Let us, now study some properties of the determinant function.

3.5 Porperties Of Determinants

In this section we will state some properties of determinants, mostly proof. We will take examples and check that these properties hold for them.

Now, for any A $\in M_n$ (F) we shall denote its columns by C_1, C_2, ….C_n Then we have the following 7 properties, P1 – P7.

P1: If C_i is an n x 1 vector over F, then
det (C_1……,C_{i-1}, C_i+ C_1+1,…..,C_n)
= det (C1…., C_i, C_i + 1,…., C_n) + det (C_1, …., C_i -1, $C_i + {}_1$,…., C_n).
P2: IF C_i = C_j, for any i$\neq$ j, then det ((C_1,C_2,…,C_n) = 0.
P3 If C_i and C_j are internachanged (i $\neq$ j) to form a new matrix B, then det B = -det (C_1, C_2……, C_n).
P4: For $\propto \in$F.
det ($C_{1..}$…$C_{i-1} \propto C_{i+1}$, …., C_n) = $\propto$ det (C_1, C_2,…, C_n).
Thus, det ($\propto C_1$, $\propto C_2$,…, $\propto C_n$) = $\propto^n$ det (C_1,…,C_n).
Now, using P1, P2 and P4, we find that for I $\neq$ j and $\propto \in$F,
det ($C_{1..}$…C_i + $\propto C_j$,…, C_j …., C_n) = det (C_1,…., C_i,…, C_j …., C_n) + $\propto$ det (C_1,…., C_i,…, C_j …., C_n). = det (C_1, C_2, …., C_n).

Thus, we have
P5: for any $\alpha \in$ F and i $\neq$ j, det (C_1,…., C_i,…, C_j ,C_{i+1},…., C_n).

Another property that we give is

P6: det (A) = det (A^t) ∀A ∈ M_n (F). (In E2 you saw that this property was true for Examples 1 and 2. its proof uses the permutation approach to determinants.)

Using P6, and the fact that det (A) can be obtained by expanding along any row, we get P7" For A ∈ Mn (F), we can obtain det (A) by expanding along any column. That is, for a fixed k,

$|A| = (-1)^{1+k} a_{1k} |A_{1k}| + (-1)^{2+k} a_{2k} |A_{2k}| + \ldots + (-1)^{n+k} a_{nk} |A_{nk}|$
An important remark now.

Remark: Using P6, we can immediately say that P1 – P5 are valid when columns are replaced by rows.

Using the notation of Unit 8, P3 say that
det (R_{ij} (A)) = - det (A) = det (C_{ij} (A)).

P4 says that
det (R_i (α) (A)) = α^n det (A) = det (C_i (α) (A)), ∀α∈F, and P5 says that
det (R_{ij} (α) (A)) = det (A) = det (Cij) (α) (A), ∀α∈ F.

Wewill now illustrate how useful the properties P1 – P7 are

Example 3: Obtain et (A), where A is

a) $\begin{vmatrix} 1 & 6 & 0 \\ 2 & 7 & 2 \\ 1 & 6 & 0 \end{vmatrix}$ b) $\begin{vmatrix} 1 & 2 & -1 & -3 \\ 2 & 4 & 5 & 0 \\ 0 & 2 & -1 & -2 \\ -1 & 0 & 0 & 1 \end{vmatrix}$

Solution: a) Since the first and third rows or A (R_1 and R_3) coincide, |A| = 0, by P2 and P6.

b) $|A| = \begin{vmatrix} 1 & 2 & -1 & -3 \\ 2 & 4 & 5 & 0 \\ 0 & 2 & -1 & -2 \\ -1 & 0 & 0 & 1 \end{vmatrix}$

$= \begin{vmatrix} 1 & 2 & -1 & -3 \\ 2 & 4 & 5 & 0 \\ 0 & 2 & -1 & -2 \\ 0 & 2 & -1 & -2 \end{vmatrix}$, by adding R1 to R4.

= 0, since R3 = R4.

Try the following exercise now.

E E4) Calculation $\begin{vmatrix} 1 & 3 & 0 \end{vmatrix}$ $\begin{vmatrix} 2 & 3 & 5 \end{vmatrix}$

$$\begin{matrix} 2 & 1 & 2 \\ 1 & 3 & 0 \end{matrix} \quad \text{and} \quad \begin{matrix} 1 & 0 & 1 \\ 4 & 6 & 10 \end{matrix}$$

Now we give some examples of determinants that you may come arose often.

Example 4: Let

$$A = \begin{matrix} A & b & b & b \\ b & a & b & b, \\ b & b & a & b \\ b & b & b & a \end{matrix} \quad \text{where } a, b \in R.$$

Calculate |A|

Solution:

$$A = \begin{vmatrix} a & b & b & b \\ b & a & b & b, \\ b & b & a & b \\ b & b & b & a \end{vmatrix}$$

$$= \begin{vmatrix} a + 3b & a + 3b & a + 3b & a + 3b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{vmatrix}$$ (by adding the second, third and fourth rows to the first row, and applying P5)

$$= \begin{vmatrix} a + 3b & 0 & 0 & 0 \\ b & a-b & 0 & 0 \\ b & 0 & a - b & 0 \\ b & 0 & 0 & -b \end{vmatrix}$$ (by subtracting the first column from every other column, and using P5)

$$= (a + 3b \begin{vmatrix} a - b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & a - b \end{vmatrix}$$ (expanding along the first row)

$$= (a + 3b)\ (a - b)^3.$$

In Example 4 we have used an important, and easily proved fact, namely, det (diag $(a_1, a_2\ldots, a_n)) = a1\ a_2\ldots a_n,\ a_1 \in F$ i.

This is true because,

$$\begin{vmatrix} a_1 & 0 & 0 & \ldots & 0 \\ 0 & a2 & 0 & \ldots & 0 \\ : & : & : & \ldots & : \\ 0 & 0 & 0 & \ldots & \alpha_n \end{vmatrix} = \alpha_n\,\alpha_{n\ldots}\,\alpha_n \begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ . & . & \ldots & . \\ : & : & \ldots & : \\ 0 & 0 & & 1 \end{vmatrix}, \text{ by P4}$$

$$= \alpha_1\,\alpha_2\,\ldots\ldots\,\alpha_n\,|1_n|$$

$$= \alpha_1\,\alpha_2\,\ldots\ldots\,\alpha_{n,}\ \text{since}|1_n| = 1.$$

Example 5: show that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix} = \underset{i<j}{\pi}\ (xj - xi),\ 1 \le\ I \le j \le\ 4$$

(This is know as the vandermonde's determinant of order 4)
Solution: the given determinant

$$= \begin{vmatrix} 1 & 0 & 0 & 0 \\ x_1 & x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ x_1^2 & x_2^2 - x_1^2 & x_3^2 - x_1^2 & x_4^2 - x_1^2 \\ x_1^3 & x_2^3 - x_1^3 & x_3^3 - x_1^3 & x_4^3 - x_1^3 \end{vmatrix}$$ (by subtracting the first column from every other column)

$$\begin{vmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ (x_2 - x_1)(x_2 + x_1) & (x_3 - x_1)(x_3 + x_1) & (x_4 - x_1)(x_4 + x_1) \\ (x_2 - x_1)(x_2^2 + x_1^2 \_ x_2x_1) & (x_3 - x_1)(x_3^2 + x_1^2 + x_3x_1) & (x_4 - x_1)(x_4^2 + x_1^2 + x_4x_1) \end{vmatrix}$$

(by expanding along the first row and factorizing the entries)

$$= (x_2 - x_1)(x_3 + x_1)(x_4 + x_1) \begin{vmatrix} 1 & 1 & 1 \\ x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ x_2^2 + x_1^2 + x_2x_1 & x_3^2 + x_1^2 + x_3x_1 & x_4^2 + x_1^2 + x_4x_1 \end{vmatrix}$$

(by taking out $(x_2 - x_1)$ $(x_3 + x_1)$, and $(x_4 - x_1)$from column 1,2 and 3 respectively).

$$= (x_2 - x_1)\ (x_3 - x_1) \begin{vmatrix} 1 & 0 & 0 \\ x_2 + x_1 & x_3 - x_1 & x_4 - x_2 \\ x_2^2 + x_1^2 + x_2x_1 & x_3^2 - x_2^2 + (x_3 - x_2)x_1 & x_4^2 - x_2^2 + (x_4 - x_2)x_1 \end{vmatrix}$$

(by subtracting the first column from the second and third columns)

$$= (x_2 - x_1)(x_3 + x_1)(x_4 + x_1) \begin{vmatrix} x_3 & x_4 - x_2 \\ (x_3 - x_2)(x_3 + x_2 + x_1) & (x_4 - x_2)(x_4 + x_2 + x_1) \end{vmatrix}$$

(expanding by the first row and factorizing the entries)

$$= (x_2 - x_1)(x_3 + x_1)(x_4 + x_1)(x_3 - x_2)(x_4 - x_2) \begin{vmatrix} 1 & 1 \\ x_3 + x_2 + x_1 & x_4 + x_2 + x_1 \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_2)(x_4 - x_3)$$

$$= \underset{i<j}{\pi} (xj - xi),\ 1 \le\ i, \le j \le\ 4$$

try the following exercise now

E E5) What are $\begin{vmatrix} a & 0 & 0 \\ \alpha & b & 0 \\ \beta & \tau & c \end{vmatrix}$ and $\begin{vmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{vmatrix}$?

The answer of E4 is part of a general phenomenon, namely, the determinant of an upper or lower triangular matrix is the product of its diagonal elements.

The proof of this is immediate because,

$$\begin{vmatrix} a_{11} & * & \ldots & * \\ 0 & a_{22} & \ldots & * \\ 0 & 0 & . & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{22} & * & \ldots & * \\ 0 & a_{33} & \ldots & * \\ 0 & . & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{vmatrix} \text{ (expanding along } C_1\text{)}$$

$= \ldots = a_{11}\, a_{22} \ldots a_{nn}$, each time expanding along the first column.

In the Calculus course you must have come across df/dt =f'(t), where f is a function of it. The next exercise involves this.

E E6) Let us define the function θ(t) by

$$\theta(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix}$$

Show that $\theta'(t) = \begin{vmatrix} f(t) & g(t) \\ f''(t) & g''(t) \end{vmatrix}$

And now, let us study a method for obtaining the invertible matrix.

5.1 INVERSE PF A MATRIX

In this section we first obtain the determinant of the product of two matrices and then define an adjoint of a matrix. Finally, we see the conditions under which a matrix is invertible, and, when it is invertible, we give its inverse in terms of its adjoint.

5.1.1 Product Formula

In unit 7 you studied matrix multiplication, let us see what happens to the determinant of a product of matrices.

Theorem 1: Let A and B be n X n matrices over F. Then det (AB) = (A) det (B).

We will not do proof here since it is slightly complicated. But let us verify theorem 1 for some cases.

Example 6: Calculate |A|, |B| and |AB| when

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 10 & 9 \\ 0 & 3 & 8 \\ 0 & 0 & 5 \end{bmatrix}.$$

Solution: We want to verify theorem ! for our pair pf matrices. Now, on expanding by the third row, we get |A| = 1.

Also, $|B| = 30$, which can be immediately seen since B is a triangular matrix

Since $AB = \begin{bmatrix} 2 & 10 & 19 \\ 6 & 33 & 35 \\ 0 & 0 & 5 \end{bmatrix}$, $|AB| = 5\begin{vmatrix} 2 & 10 \\ 6 & 33 \end{vmatrix} = 30$

$= |A| \; |B|$.

You can verify theorem 1 for the following situation.

E E7) show that $|AB| = |A| \; |B|$, where

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \\ 3 & -3 & 5 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 2 & 0 \\ 5 & -3 & 3 \end{bmatrix}$$

Theorem 1 can be extended to a product of m n x n matrices,
$A_1, A_2, \ldots, A_m$. That is,
$\det(A_1 A_2 \ldots A_m) = \det(A_1) \det(A_2) \ldots\ldots \det(A_m)$

Now let us look at an example in which Theorem 1 simplifies calculations.

Example 7: for a, b, c, $\in R$, calculate

$$\begin{vmatrix} a^2 + 2bc & c^2 + 2ab & b^2 + 2ac \\ b^2 + 2ac & a^2 + 2bc & c^2 + 2ab \\ c^2 + 2ab & b^2 + 2ac & a^2 + 2bc \end{vmatrix}$$

Solution: The solution is very simple. The given matrix is equal to

$$\begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}^2 .$$ Therefore,

We get the required determinant to be

$$\left|\begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}\right|^2 = \begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}^2 \quad \text{(by theorem 1)}$$

$$= (a^3 + b^3 + c^3 - 3abc)^2,$$

because $\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix} = a \begin{vmatrix} a & b \\ c & a \end{vmatrix} - b \begin{vmatrix} c & b \\ b & a \end{vmatrix} + c \begin{vmatrix} c & a \\ b & c \end{vmatrix}$

$$= a(a^2 - bc) - b(ac - b^2) + c(c^2 - ab)$$
$$= a^3 + b^3 + c^3 - 3abc.$$

Now, you know that $AB \neq BA$, in general. But, det (AB) = det (BA), since both are equal to the scalar det (A) det (B).

On the other hand, det $(A + B) \neq$ det (a) + det (b), in general. The following exercise is an example.

E e8) Let $A = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$ $B = \begin{matrix} -1 & 0 \\ 0 & -1 \end{matrix}$ show that det $(A + B) \neq$ det (A) + det (B).

What we have just said is that det. Is not a linear function.

We now give an immediate corollary to theorem 1.

Corollary 1: If $A \in M_n$ (F) is invertible, then det (A^{-1}) = 1/det (A).

Proof: Let $B \in M_n$ (F) such that AB = 1. Then det(AB = det(A) det (B) = det(1) =1

Thus, det(A) $\neq$ 0 and det (B) = 1/det(A). In particular,

Det (A^{-1}) = 1/det (A).

A matrix B us similar to a matrix A if there exists a non-singular matrix P such that $P^{-1} AP = B$

Another corollary to Theorem 1 is

Corollary 2: Similar matrics have the same determinant.

Proof: if B is similar to A, then B = P-1 AP for some invertible matrix P. Thus, by
Theorem 1, det(B) = det(P-1 AP)

= det (P-1) det(A) et (P) = 1/det (P). det(P), by cor. 1

= det(A).

we use this corollary to introduce you to the determinant of a linear transformation. At each stage you have seen the very close relationship between linear transformations and matrices. Here too, you will see this closeness.

Definifion: Let T:V → V be a linear transformation on a finite-dimensional on-zero vector space V. Let A = $[T]_B$ be the matrix of T with respect to a given basis B of V.

Then we define the determinant of T by det(T) = det(A).

This definition is independent of the basis of V that is chosen because, if we choose another basis B′ of V we obtain the matrix A′ = $[T]_{B'}$,, which is similar to A (see Unit 7, Cor. To Theorem 10). Thus, det (A′) = det (A).

We have the following example and exercises.

Example 8: Find det (T) where we define T: R3 → R3 by
T(x1,x2, x3) = ($3x_1 + x_3$, $-2x_1 + x_2$,$-x+2x2+4x_3$)
Solution: Let B = {(1,0,0), (0,1,0), (0,0,1)} be the standard ordered basis of R3. Now,
T(1,0,0) = (3,-2,-1) = 3(1,0,0) -2(0,1,0) -1(0,0,1)
T(0,1,0) = (0,1,2) = 0(1,0,0) + 1(0,1,0) + 2(0,0,1)
T(0,0,1) = (1,0,4) = 191,0,0) + 0(0,1,0) + 4(0,01)

$$\therefore, A = [T]_B = \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}$$

So, by definition,

$$\det(T) = \det(A) = \begin{vmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{vmatrix}$$

$$= 3\begin{vmatrix} 1 & 0 \\ 2 & 4 \end{vmatrix} + 1\begin{vmatrix} -2 & 1 \\ -1 & 2 \end{vmatrix} = 12 - 3 = 9.$$

E E9) Find the determinant of the zero operator and the identity operator from
R3 → R3.

E E10) Consider the differential operator

$D: P2 \rightarrow P2 : D(a_0 + a_1x + a_2x^2) = a_1 + 2a_1x.$

What is det(D)?

Let us now see what the adjoint of a square matrix is, and how it will help us in obtaining the inverse of an invertible matrix.

5.1.2 Adjoint of a Matrix.

In section 9.2 we used the notation A_{ij} for the matrix obtained from a square matrix A by deleting its ith row and jth column. Related to this we define the (i,j)th cofactor of A (or the cofactor of a_{ij}) to be $(-1)^{i+j}|A_{ij}|$. It is denoted by C_{ij}. That is $Cij = (-1)^{i+j}|A|$.

Consider the following example.

Example 9: Obtain the cofactors C12 and C23 of the matrix $A = \begin{bmatrix} 0 & 2 & -1 \\ 3 & 4 & 1 \\ 2 & 1 & 6 \end{bmatrix}$

Solution: $C_{12} = (-1)^{2+3}|A_{12}| = -\begin{vmatrix} 3 & 1 \\ 2 & 6 \end{vmatrix} = -16$

$C_{23} = (-1)^{2+3}|A_{23}| = -\begin{vmatrix} 3 & 1 \\ 2 & 6 \end{vmatrix} = 4.$

In the following result we give a relationship between the elements of a matrix and their cofactors.

Theorem 2: Let $A = [aij]_n x_n$. Then,

a) $a_{i1}C_{i1} + a_{i2}C_{i2} \ldots + a_{in}C_{in} = \det(A) = a_{1i}C_{1i} + a_{2i}C_{2i} + .. + a_{ni}C_{ni}.$
b) $a_{i1}C_{j1} + a_{i1}C_{j2} \ldots + a_{in}C_{jn} = 0 = a_{i1}C_{j1} + a_{2i}C_{2j} + .. + a_{ni}C_{nj}$ if $i \neq j$.

We will not be proving this theorem here. We only mention that (a) follows immediately from the definintion of det (A), since det (A) = $(-1)^{i+1}$ a_{i1} $|A_{i1}|$ + … + $(-1)^{i+n}$ a_{in} $|A_{in}|$.

E E11)Verify (b) of theorem 2 for the matrix in example 9 and i=1, j=2 or 3.

Now, we can define the adjoint of a matrix.

Definition: Let A = [aij] be any n x n matrix. Then the adjoint of A is the n x n matrix, denoted by Adj(A), and defined by

$$\begin{bmatrix} C_{11} & C_{12} \dots & C_{in} \\ C_{21} & C_{22} \dots & C_{2n} \\ \vdots & \vdots & \vdots \\ C_{n1} & C_{n2} \dots & C_{nn} \end{bmatrix}^t \qquad \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} \dots & & C_{nn} \end{bmatrix}$$

Where Cij denotes the (i,j)th cofactor of A.

Thus, Adj (A) is the n x n matrix which is the transpose of the matrix of corresponding cofactors of A.

Let us look at an example.

Example 10: Obtain the adjoint of the matrix A = $\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$

Solution: $C_{11} = (-1)^{1+1}\begin{vmatrix} 1 & 0 \\ 0 & \cos\theta \end{vmatrix} = \cos\theta$

$C_{12} = (-1)\ 1+2 \begin{vmatrix} 0 & 0 \\ \sin\theta & \cos\theta \end{vmatrix} = 0$

$C_{13} = \begin{matrix} 0 & 1 \\ \sin\theta & 0 \end{matrix} = -\sin\theta$

$C_{21} = 0$, $C_{22} = \cos^2\theta + \sin^2\theta = 1$, $C_{23} = 0$.

$C_{31} = \sin\theta$, $C_{32} = 0$, $C_{23} = \cos\theta$.

$$\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ & & \\ & & \end{bmatrix}^t \qquad \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ & & \\ & & \end{bmatrix}$$

$\therefore$, Adj (A) = $\begin{matrix} 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{matrix}$ = $\begin{matrix} 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{matrix}$

Now you can try the following exercise.

E E12) find Adj(A), where A = $\begin{bmatrix} 2 & 3 & -1 \\ 0 & 0 & 6 \\ 0 & 0 & 5 \end{bmatrix}$

In Unit 7 you came across one method of finding out id a matrix is invertible. The following theorem uses the adjoint to give another way of finding out if a matrix A is invertible. It also gives us A^{-1}, if A is invertible.

Theorem 3: Let A be an n x n matrix over F, then
A. (Adj(A)) = (Adj(A)). A = det(A) 1.

Proof: Recall matrix multiplication from Unit 7. Now

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{matrix} \quad \begin{matrix} C_{11} & C_{12} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n1} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{matrix}$$

by Theorem 2 we know that ai1Ci1 + …. + ainCin = et(A), and
ai1Cj1 + ai2Ci2 + … + ainCjn = 0 if i $\neq$ j. Therefore,

$$A\,(\text{Adj}\,(A)) = \begin{matrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ : & : & \cdots & : \\ 0 & 0 & \cdots & \text{Det}\,(A) \end{matrix}$$

$$\begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ & & \cdots & \\ & & \cdots & \\ & & \cdots & \\ & & \cdots & \end{matrix}$$

$$= \det(A) \begin{matrix} 0 & 0 & & 0 \\ : & : & & : \\ 0 & 0 & & 1 \end{matrix} = \det(A)\, 1.$$

Similarly, (Adj (A)) .A = det (A) 1.

An immediate corollary shows us how to calculate the inverse of a matrix, if it exists.

Corollary: let A be an n x matrix over F. Then a is invertible if an only if det (A) ≠ 0, then

A-1 = (1/det (A)) Adj (A).

Proof: If a is invertible, then A-1 exists and A-1 A= 1. so, by theorem 1, Det(A-1) det (A) = det (1) = 1.∴, det (A) ≠ 0.

Conversely, if det (A) ≠ 0, then Theorem 3 says that

$$A\left(\frac{1}{|A|}\text{Adj }(A)\right) = 1 = \left(\frac{1}{|A|}\text{ Adj }(A)\right) A$$

$$\therefore A^{-1} = \frac{1}{|A|}\text{ Adj }(A).$$

We will use the result in the following example.

Example 11: Let

$$A = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \qquad \text{Find } A^{-1}$$

Solution:

$$\det(A) = (-1)^{2+2} .1. \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}$$ (by expansion along the second row)

$$= \cos^2\theta + \sin^2\theta = 1$$

also, from Example 10 we know that

$$\text{Adj}(A) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

Therefore, A^{-1} = (1/det (A)) Adj(A) = Adj(A).

You should also verify that Adj(A) is A^{-1} by calculating A.Adj(A) and Adj(A).A.

You can use theorem 3 for solving the following exercises.

E E13) Can you find A^{-1} for the matrix in E 12?

E E14) find the adjoint and inverse of the matrix A in E7

E E15) If A^{-1} exists, does $[Adj(A)]^{-1}$?

Now we go to the next section, in which we apply our knowledge of determinants to obtain solutions of systems of linears equations.

5.2 Systems Of Linear Equations

Consider the system of n linear equations in n unknowns, given by

$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$
$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$
. . . .
. . . .
$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n$
which is the same as

$$AX = B, \text{ where } A = [aij],\ X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix},\ B = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{bmatrix}.$$

In Section 8.4 we discussed the Gaussian elimination method for obtaining a solution of this system. In this section we give a rule due to the mathematician Cramer, for solving a system of linear equations when the number of equations equals the number of variables.

Theorem 4: Let the matrix equation of a system of linear equations be

$$AX = B, \text{ where } A = [a_{ij}]_{nxn},\ X = \begin{bmatrix} x_1 \\ . \\ . \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ . \\ . \\ b_n \end{bmatrix}$$

Let the columns of A be $C_1, C_2, \ldots. C_n$. If $\det(A) \neq 0$, the given system has a unique solution, namely,
$X_1 = D_1/D, \ldots.\ x_n = D_n/D$, where
$D_i = \det(C_1, \ldots.C_{i-1}, B, C_{i-1}, \ldots., C_n)$
= determinant of the matrix obtained from A by replacing the ith column by B, and D = det (A).

Proof: Since $|A| \neq 0$, the corollary to Theorem 3 says that A-1 exists
Now AX = B $\Rightarrow$ A-1 AX = A-1B
$\Rightarrow$ IX = (1/D) Adji(A) B

$$X = (1/D)\begin{bmatrix} C_{11} & C_{21} \dots & C_{n1} \\ C_{12} & C_{22} \dots & C_{n2} \\ \cdot & \cdot \cdots & \cdot \\ \vdots & \vdots & \vdots \\ C_{1n} & C_{n2} \dots & C_{nn} \end{bmatrix} \begin{matrix} b_1 \\ b_2 \\ \vdots \\ \\ b_n \end{matrix}$$

Thus,

$$\begin{matrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{matrix} = (1/D) \begin{bmatrix} C_{11}b_1 + C_{21}b_2 + \dots + C_{n1}b_n \\ C_{12}b_1 + C_{22}b_2 + \dots + C_{n2}b_n \\ . \quad . \quad . \\ . \quad . \quad . \\ C_{1n}b_1 + C_{2n}b_2 + \dots + C_{nn}b_n \end{bmatrix}$$

Now, D_i = det (C_1, ….., C_{i-1}, B, C_{i+1}, …., C_n). Expanding along the ith column, we get $D_i = C_{1i}b_1 + C2_ib_2 + \dots + C_{ni}b_n$.

Thus,

$$\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix} = 1/D \begin{bmatrix} D_1 \\ D_1 \\ . \\ . \\ D_n \end{bmatrix}$$

which gives us Cramer's rule, namely,
$x_1 = D_1/D, x_2 = D_2/D, \dots, x_n = D_n/D$.
The following example and exercise may help you to practice using Cramer's rule.

Example 12: Solve the following system using Cramer's rule:

$2x + 3y - z = 2$
$x + 2y + z = -1$
$2x + y - 6z = 4$

Solution: The give system is equivalent to AX = B, where

$$A = \begin{matrix} 2 & 3 & -1 \\ 1 & 2 & 1 \\ 2 & 1 & -6 \end{matrix}, \quad X = \begin{matrix} x \\ y \\ z \end{matrix}, \quad B = \begin{matrix} 2 \\ -1 \\ 4 \end{matrix}$$ Therefore, apply the rule,

$$\begin{vmatrix} 2 & 3 & -1 \\ -1 & 2 & 1 \end{vmatrix} \quad \begin{vmatrix} 2 & 3 & -1 \\ 1 & -1 & \end{vmatrix} 1 \quad \begin{vmatrix} 2 & 3 & 2 \\ 1 & 2 & -1 \end{vmatrix}$$

$$x = \frac{4 \quad 1 \quad -6}{\begin{vmatrix} 2 & 3 & -1 \\ 1 & 2 & 1 \\ 4 & 1 & -6 \end{vmatrix}}, \quad y = \frac{2 \quad 4 \quad -6}{\begin{vmatrix} 2 & 3 & -1 \\ 1 & -1 & 1 \\ 2 & 4 & -6 \end{vmatrix}}, \quad z = \frac{2 \quad 1 \quad 4}{\begin{vmatrix} 2 & 3 & 2 \\ 1 & 2 & -1 \\ 2 & 1 & 4 \end{vmatrix}}$$

After calculating, we get
X = 23, y = 14, z = -6.

Substitute these values in the given equations to check that we haven't made a mistake in our calculations.

E E16) Solve, by Cramer's rule, the following system of equations.

$$\begin{aligned} X + 2y \quad & +4z = 1 \\ 2x + 3y \quad -z & = 3 \\ x \quad -3z & = 2 \end{aligned}$$

Now let us see what happens if B = 0. Remember, in Unit 8 you saw that AX = 0 has n-r linearly independent solutions, where r = rank a. The following theorem tells us this condition in terms of det(A).

Theorem 5: The homogeneous system AX = 0 has a non-trivial solution if and only if det(A) = 0

Proof: first assume that AX = 0has a non-trivial solution. Suppose, if possible, that det(A) ≠ 0. Then Cramer's Ruler's says that AX = 0 has only the trivial solution X = 0 (because each Di=0 in Theorem 4). This is a contradiction to our assumption. Therefore, det(A) = 0.

Conversely, if det (A) = 0, then A is not invertible. ∴, the linear mapping $A : V_n(F) \to V_n(F)$: A(X) = AX is not invertible .∴, this mapping is not one-one Therefore, Ker A ≠ 0, that is AX = 0 for some non-zero X∈Vn(F). Thus, AX = 0 has a non-trivial solution.

You can use Theorem 5 to solve the following exercise.

E E17) Does the sytem

$$\begin{aligned} 2x + 3y \quad + z &= 0 \\ x - y \quad -z &= 0 \\ 4x + 6y \quad + 2z &= 0 \end{aligned}$$

have a non-zero solution?

And now we introduce you to the determinant rank of a matrix, which leads us to another method of obtaining the rank of a matrix.

5.3 The Determinant Rank

In Unit 5 and 8 you were introduced to the rank of a linear transformation and the rank of a matrix, respectively. Then we related the two ranks. In this section we will discuss the determinant rank and show that it is the rank of the concerned matrix. First we give a necessary and sufficient condition for n vectors in $V_n(F)$ to be linearly dependent.

Theorem 6: Let $X_1, X_2, \ldots, X_n \in V_n$ (F). Then $X_1, X_2, \ldots, X_n$ are linearly dependent over the field F if and only if det $(X_1X_2, \ldots X_n) = 0$.

Proof: Let U = $(X_1X_2, \ldots X_n)$ be the n x n matrix whose column vectors are $X_1X_2, \ldots X_n$. Then $X_1X_2, \ldots X_n$ are linearly dependent over F if and only if there exist scalars $a_1, a_2, \ldots, a_n \in F$, not all zero, such that a1 $X_1 + a_2X_2 + \ldots + a_nX_n = 0$.

Now, $U \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ \alpha_n \end{bmatrix} = (X_1X_2, \ldots X_n) \begin{matrix} a_1 \\ a_2 \\ . \\ . \\ \alpha n \end{matrix}$

$= X_1 a_1 + X_2a_2 + \ldots + X_na_n$
$= a_1X_1 + a_2X_2 + \ldots + \alpha_nX_n$.

Thus, $X_1X_2,\ldots X_n$ are linearly dependent over F if and only if UX = 0 for some non-

$$\text{zero } X = \begin{matrix} a_1 \\ a_2 \\ . \\ \alpha_n \end{matrix} \in V_n\ (F).$$

But this happens if and only if det (U) = 0, by Theorem 5. Thus, Theorem 6 is proved.

Theorem 6 is equivalent to the statement $X_1X_2,\ldots X_n \in Vn$ (F) are linearly independent if and only if det $(X_1X_2,\ldots\ldots X_n) \neq 0$.

You can use Theorem 6 for solving the following exercises.

E E18) Check if the vectors $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$ are linearly independent over R.

.

Now, consider the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 2 & 3 \end{bmatrix}$

A submatrix of is a matrix that can be obtained from A by deleting some rows and columns.

Since two rows of A are equal we know that |A| = 0. But consider its 2 x 2 submatrix

$A_{13} = \begin{matrix} 0 & 4 \\ 1 & 2 \end{matrix}$ Its determinant is – 4 ≠ 0. In this case we say that the determinant rank of A is 2.

In general, we have the following definition.

Definition: Let A be an m x n matrix. If A ≠ 0, then the determinant rank of A is the largest positive integer r such that

i) there exists an r x r submatrix of A whose determinant is non-zero, and
ii) for s > r, the determinant of any s x s submatrix of A is 0.

Note: The determinant rank r of any m x n matrix is defined, not only of a square matrix. Also $r \leq \min(m, n)$.

Example 13: Obtain the determinant rank of $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

Solution: Since A is a 3 x 2 matrix, the largest possible value of its determinant rank can be 2. Also the submatrix $\begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix}$ of A has determinant (- 3) ≠ 0.

∴, the determinant rank of A is 2.

Try the following exercise now.

E E19) Calculate the determinant rank of A, where A =

a) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}$, b) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

And now we come to the reason for introducing the determinant rank – it gives us another method for obtaining the rank of a matrix.

Theorem 7: The determinant rank of an m x n matrix A is equal to the rank of A.

Proof: Let the determinant rank of A be r. Then there exists an r x r submatrix of A whose determinant is non-zero. By Theorem 6, its column vectors are linearly independent. It follows by the definition of linear independence, that these column vectors, when extended to the column vectors of A, remain linearly independent. Thus, A has at least r linearly independent column vectors. Therefore, by definition of the rank of a matrix,

$r \leq \text{rank}(A) = p(A)$ …… (1)

also, by definition of p (A), we know that the number of linearly independent rows that A has is p (A). These rows form a p (A) x n matrix p (A). Thus, B will have p(A) linearly independent columns. Retaining these linearly independent columns of B we get a p(A) x p(A) submatrix C of B. so, C is a submatrix of A whose determinant will be non-zero, by theorem 6, since its columns are linearly independent. Thus, by the definition of the determinant rank of A, we get

$p(A) \leq r$ ……..(2)

(1) and (2) give us us p(A) = r.

we will use Theorem 7 in the following example.

Example 14: Find the rank of

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

Solution: det (A) = 0. but det $\left(\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}\right) = -7 \neq 0.$

Thus, by theorem 7, p(A) = 2.

Remark: This example show Theorem 7 can simplify the calculation of the rank of a matrix in some cases. We don't have to reduce a matrix to echelon form each time. But, in (a) of the following exercise, we see a situation where using this method seems to be as tedious as the row-reduction method.

E E20) Use Theorem 7 to find the rank of A, where A =

a) $\begin{bmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \\ 4 & 3 & 1 & 7 \end{bmatrix}$, b) $\begin{bmatrix} 2 & 3 & 5 & 1 \\ 1 & -1 & 2 & 1 \end{bmatrix}$

E20 (a) shows how much time can be taken by using this method. On the other hand, E20 (b) shows how little time it takes to obtain p(A), using the determinant rank. Thus, the method to be used for obtaining p(A) varies from case to case.

We end this unit by briefly mentioning what we have cover in it.

5.4 Summary

In this unit we have covered the following points.

1) The definition of the determinant of a square matrix.

2) The properties P1-P7, of determinants.
3) The statement and use of the fact that det(AB) = det(A) det(B).

4) The definintion of the determinant of a linear transformation from U to V, where dim U = dim V.

5) The definition of the adjoints of a square matrix.

6) The use of adjoints to obtain the inverse of an invertible matrix.

7) The proof and use of Cramer's rule for solving a system of linear equations.

8) The proof of the fact that the homogeneous system of linear equations AX = 0 has a non-zero solution if and only if det(A) = 0.

9) The definition of the determinant rank, and the proof of the fact that rank of A = determinant rank of A.

5.5 Solutions/Answers

E1) On expanding by the 2nd row we get

$$|A| = -5\,|A_{21}| + 4|A_{22}| - |A_{23}|.$$

Now, $|A_{21}| = \begin{vmatrix} 2 & 6 \\ 3 & 2 \end{vmatrix} = 4 - 18 = -14$

$$|A_{22}| = \begin{vmatrix} 1 & 6 \\ 7 & 2 \end{vmatrix} = 2 - 42 = -40.$$

$$|A_{23}| = \begin{vmatrix} 1 & 2 \\ 7 & 3 \end{vmatrix} = 3 - 14 = -11.$$

$\therefore$ $|A| = (-5)(-14) + (-40) - (-11) = -79.$

Expanding by the 3rd row, we get

$$|a| = 7|a31| - 3|A32| + 2|A33| = 7\begin{vmatrix} -2 & 6 \\ 4 & 1 \end{vmatrix} - 3\begin{vmatrix} 1 & 6 \\ 5 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 2 \\ 5 & 4 \end{vmatrix}$$

$= 7(-22) - 3(-29) + 2(-6) = -79.$

Thus, $|A| = -79$, irrespective of the row that we use to obtain it.

E2)a) $A^t = \begin{vmatrix} 1 & 5 & 7 \\ 2 & 4 & 3 \\ 6 & 1 & 2 \end{vmatrix}$, $\therefore$, on expanding by the first row, we get

$$|A^t| = 1\begin{vmatrix} 4 & 3 \\ 1 & 2 \end{vmatrix} - 5\begin{vmatrix} 2 & 3 \\ 6 & 2 \end{vmatrix} + 7\begin{vmatrix} 2 & 4 \\ 6 & 1 \end{vmatrix} = 5 + 70 + /(-22) = -79.$$

b) $A^t = \begin{vmatrix} -3 & 2 & 1 & 2 \\ -2 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \\ 2 & -1 & 2 & 1 \end{vmatrix}$ Since the 3rd row has the maximum

Number of zeros, we expand along it. Then

$$|A| = 1\begin{vmatrix} -3 & 2 & 2 \\ -2 & 1 & 1 \\ 2 & -1 & 1 \end{vmatrix} - (-3)\begin{vmatrix} -3 & 2 & 1 \\ -2 & 1 & 0 \\ 2 & -1 & 2 \end{vmatrix} = 2 + 3(2) = 8.$$

E3) The magnitude of the required volume is the modulus of

$$\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

we draw the box in fig. 2.

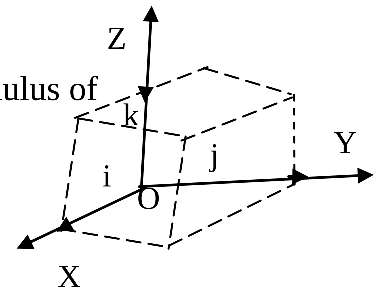

Fig. 2

E4) the first determinant is zero, using the row equivalent of P2.
The second determinant is zero, using the row equivalent of P5, since $R_3 = 2R_1$.

E5 $\begin{vmatrix} a & 0 & 0 \\ \alpha & b & 0 \\ & & \end{vmatrix} = a\begin{vmatrix} b & 0 \\ r & c \end{vmatrix} = a\,b\,c.$

β_T c

$$\begin{vmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{vmatrix} = a \begin{vmatrix} b & f \\ 0 & c \end{vmatrix} = a\ b\ c.$$

E6 $\theta(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix} = f(t') - f'(t)\, g(t).$

$\therefore \theta'(t) = f'(t)\, g'(t) + f(t')\, g'(t) - \{f'(t)\, g(t) + f'(t)\, g'(t)\},$

since $\frac{d}{dt}(fg) = \frac{df}{dt}\, g + f\frac{dg}{dt}$

$= f(t)\, g''(t) - f''(t)\, g(t) = \begin{vmatrix} f(t) & g(t) \\ f''(t) & g''(t) \end{vmatrix}$

E7) note that B is obtained from A by interchanging C_1 and C_3.

$\therefore |B| = - |A|$

Now $|A| = \begin{vmatrix} 2 & -2 \\ -3 & -5 \end{vmatrix} \begin{vmatrix} - & 0 \\ 3 & -3 \end{vmatrix} 3 = 4 + 6 = 10.$ $\therefore$, $|B| = - 10.$

Also $|AB| = \begin{vmatrix} -6 & 3 & -2 \\ -14 & 10 & -6 \\ 28 & -21 & 18 \end{vmatrix}$

$= \begin{vmatrix} -6 & 3 & -2 \\ -14 & 10 & -6 \\ 0 & -1 & 6 \end{vmatrix}$ adding $2R_2$ to R_3.

$= \begin{vmatrix} -6 & -2 \\ -14 & -6 \end{vmatrix} \begin{vmatrix} -6 & 3 \\ -14 & 10 \end{vmatrix}$, expanding along R_3.

$= 8 - 108 = - 100 = |A|\ |B|.$

E8) $|A| = 1 = |B|. \therefore |A| + |B| = 2.$

But $a + B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ …. $|A+B| = 0$

E9) let B be the standard basis of R3. The zero operators is

$0{:}R^3 : 0\ (x) = 0 \forall\ x \in R^3$. Now, $[0]_B = 0$.

$\therefore$ det (0) = 0.

1: $R^3 - R^3 : (x) = x \forall\ x \in R^3$, is the identify operator or R3. Now, $[I]_B = I_3$

$\therefore$ det(I) = det (I_3) = 1.

E10) The standard basis for P2, is $\{1,x,x^2\}$

Now D(1) = 0, D(x) = 1, D(x) = 1, $D(x^2) = 2x$,

$$\therefore |D|_B = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{vmatrix}$$

$$\therefore \det(D) = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

E11) $a_{11}C_{21} + a_{12}C_{22} + a_{13}C_{23} = 2(-1)^{2+2} \begin{vmatrix} 0 & -1 \\ 2 & 6 \end{vmatrix} + (-1)(-1)^{2+3} \begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} = 0$

similarly, check that $a_{11}C_{21} + a_{12}C_{22} + a_{31}C_{32} = 0$,

$a_{11}C_{31} + a_{12}C_{32} + a_{13}C_{33} = 0 = a_{11}C_{13} + a_{21}C_{23} + a_{31}C_{33}$

E12) $C_{11} = 0$, $C_{12} = 0$, $C_{13} = 0$, $C_{21} = 15$, $C_{22} = 10$, $C_{23} = 0$, $C_{31} = 18$, $C_{32} = -12$, $C_{33} = 0$.

$$\therefore \text{Adj}(A) = \begin{vmatrix} 0 & -15 & 18 \\ 0 & 10 & -12 \\ 0 & 0 & 0 \end{vmatrix}$$

E13) Since $|A| = 0$ A^{-1} does not exist.

E14) From E7 we know that $|A| = 10$.

Now, $C_{11} = 4$, $C_{12} = 6$, $C_{13} = -6$,
$C_{21} = 3$, $C_{22} = 8$, $C_{23} = 3$,
$C_{31} = 2$, $C_{32} = 2$, $C_{33} = 2$.

$$\therefore \text{Adj}(A) = \begin{vmatrix} 4 & 3 & 2 \\ 6 & 8 & 2 \\ -6 & 3 & 2 \end{vmatrix}$$

$$\therefore A^{-1} = \frac{1}{A} \text{ Adj }(A) = \frac{1}{10} \begin{vmatrix} 4 & 3 & 2 \\ 6 & 8 & 2 \\ -6 & 3 & 2 \end{vmatrix}.$$

verify that the matrix we have obtained is right, by multiplying it by A.

E15) Since A. Adj (A) = |A| I = Adj (A). A, and $|A| \neq 0$, we find that

$[\text{Adj}(A)]^{-1}$ exists, and is $\frac{1}{|A|}$ A.

E16) This is of the form AX = B, where

$$A = \begin{vmatrix} 1 & 2 & 4 \\ 2 & 3 & -1 \\ 1 & 0 & -3 \end{vmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

$$\therefore D_1 = \begin{vmatrix} 1 & 2 & 4 \\ 3 & 3 & -1 \\ 2 & 0 & -3 \end{vmatrix} = -19$$

$$D_1 = \begin{vmatrix} 1 & 1 & 4 \\ 2 & 3 & -1 \\ 1 & 2 & -3 \end{vmatrix} = 2$$

$$D_1 = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ 1 & 0 & 2 \end{vmatrix} = 1$$

$$D = |A| = -11$$

$$\therefore x = \frac{D_1}{D} = \frac{19}{11}\ y = \frac{D_2}{D} = \frac{-2}{11}\ z = \frac{D_3}{D} = \frac{-1}{11}$$

E17) The given system is equivalent to AX = 0, where

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -1 & -1 \\ 4 & 6 & 2 \end{bmatrix}$$

now, the third row of A is twice the first row of A.

$\therefore$, by P2 and P4 of Section 9.3, | A | = 0.

$\therefore$, Theorem 5, the given system has a non-zero solution.

E18) $\begin{vmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \\ 1 & 1 & 0 \end{vmatrix} = -3 + 2 = -1 \neq 0$. $\therefore$ the given vectors are linearly independent.

E19) a) Since $|A| \neq 0$, the determinant rank of A is 3.

b) As in Example 13, the determinant rank of A is 2.

E20) a) The determinant rank of A $\leq$ 3.

Now the determinant of the 3 x 3 submatric $\begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & -1 \\ 4 & 3 & 1 \end{bmatrix}$ is zero.

Also, the determinant of the 3 x 3 submatrix $\begin{bmatrix} 3 & 2 & 5 \\ 1 & -1 & 2 \end{bmatrix}$ is zero.

4 1 7

In fact, you can check that the determinant of any of the 3x3 submatrices is zero. Now let us look at the 2 x 2 submatrices of A. Since $\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5 \neq 0$

we find that p(A) = 2.

b) The determinant rank of A ≤ 2.

Now $\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -5 \neq 0$. $\therefore$, p(A) = 2.

UNIT 6

EIGENVALUES AND EIGENVECTORS

Structure

6.1 Introduction

In Unit 5 you have studied about the matrix of a linear transformation. You have had several opportunities, in earlier units, to observe that the matrix of a linear transformation depends on the choice of the bases of the concerned vector spaces.

Let V be an n-dimentional vector space over F, and let $T : V \to V$ be a linear transformation. In this unit we will consider the problem of finding a suitable basis B, of the vector space V, such that the n x n matrix [T]B is a diagonal matrix. This problem can also be see as: given an n xx n matrix A, find a suitable n x n non-singular matrix P such that P-1 AP is a diagonal matrix (see Unit 7, Cor. To Theorem 10). It is in this context that the study of eigenvalues and eigenvectors plays a central role. This will be seen in Section 10.4.

The eigenvalue problem involves the evaluation of all the eigenvalues and eigenvectors of a linear transformation or a matrix. The solution of this problem has basic applications in almost all branches of the sciences, technology and the social science besides its fundamental role in various branches of pure and applied mathematics. The emergence of computers and the availability of modern computing facilities has further strengthened this study, since they can handle very large systems of equations.

In Section 6.2 we define eigenvalues and eigenvectors. We go on to discuss a method of obtaining them, in Section 6.3. In this section we will also define the characteristic polynomial, of which you will study more in the next unit.

Objectives

After studying this unit, you should be able to

- Obtain the characteristic polynomial of a linear transformation or a matrix;
- Obtain the eigenvalues, eigenvectors and eigenspaces of a linear transformation or a matrix;
- Obtain a basis of a vector space V with respect to which the matrix of a linear transformation T : V $\rightarrow$ V is in diagonal form;
- obtain a non-singular matrix P which diagonalises a given diagonalizable matrix A.

6.2 THE ALGEBRAIC EIGENVALUE PROBLEM

Consider the linear mapping T : $R^2 \rightarrow R^2$: T(x,y) = (2x, y). Then, T(1,0) = (2,0) = 2(1,0). $\therefore$ T(x,y) = 2(x,y) = (1,0) $\neq$ (0,0). In this situation we say that 2 is an eigenvalue of T. But what is an eigenvalue?

Definitions: An eigenvalue of a linear transformation T: V $\rightarrow$ V is a scalar such that there exists a non-zero x $\in$ V is called an eigenvector of T with respect t the eigenvalue . (In our example above, (1,0) is an eigenvector of T with respect to the eigenvalue 2)

Thus, a vector x $\in$ V is an eigenvector of the linear transformation T if
i) x is non-zero, and
ii) Tx = λx for some scalar $\lambda \in$F.

The fundamental algebraic eigenvalue problem deals with the determination of all the eigenvalues of a linear transformation. Let us look at some examples of how we can find eigenvales.

Example 1: Obtain an eignevalue and a corresponding eigenvector for the linear operator T :$R^3 \rightarrow R^3$: T(x,y,z) = (2x, 2y, 2z).

Solution: Clearly, T(x,y,z) = 2(x,y,z) $\forall$ (x,y,z) $\in R^3$. Thus, 2 is an eigenvalue of T. Any non-zero element of R^3 will be an eigenvector of T corresponding to 2.

Example 2: Obtain an eigenvalue and a corresponding eigenvector of T: C3 $\rightarrow$ C3: T(x,y,z) = (ix,iy,z).

Solution: Firstly note that T is a linear operator. Now, if $\lambda \in$ C is an eigencalue, then $\exists$ (x,y,z) $\neq$ (0,0,0) such that T(x,y,z) =λ (x,y,z) $\Rightarrow$ (ix,iy,z) = (λzλy, λz).
$\Rightarrow$ ix = λx, -iy = λy, z = λz

These equations are satisfied if λ = i, y = 0, z = 0
$\therefore$, λ = I is an eigenvalues with a corresponding eigenvector being (1,0,0) (or (x,0,0) for any x $\neq$0)

(1) is also satisfied if λ = -i,x = 0, z = 0 or f λ = 1, x = 0, y = 0. There, -i and 1 are also eigenvalues with corresponding eigenvectors (0,y,0) and 0,0z) respectively for any $\neq$ 0, z $\neq$0.

Do try the following exercise now.

E E1) Let T : $R^2 \rightarrow R^2$ be defined by T(x,y) = Obtain an eigenvalue and a corresponding eigenvector of T.

Warning: The zero vector can never be an eigenvector. But,$\in$ F can be an eigevalue. For example, 0 is an eigenvalue of the linear operator in E 1, a corresponding eigenvector being (0,1).

Now we define a vector space corresponding to an eigenvalue of T:V $\rightarrow$ V. Suppose $\lambda \in$ F is an eigenvalue of the linear transformation T. Define the set W_A = {x $\in$ V| T(x) = λx}.

= {0} $\cup$ {eigenvectors of T corresponding to λ}.

So, a vector v $\in$W, if and only if v = 0 is an eigenvector of T corresponding to λ.
Now, x$\in W_\lambda \Leftrightarrow$ Tx = λIx, I being the identity operator.

$\Leftrightarrow$ (T - λI)x = 0

$\Leftrightarrow$ x$\in$ Ker (T - λI)

$\therefore W_\lambda$ = Ker (T - λI), and hence, $W_\lambda$is a subspace of V (ref. unit 5, Theorem 4).
Since $\lambda$is an eigenvalue of T, it has an eigenvector, which must be non-zero. Thus, $W_\lambda$is non-zero.

Definition: For an eigenvalue λ of T, the non-zero subspace W_λ is called the eigenspace of t associated with the eigenvalue.

Example 3: Obtain W_2 for the linear operator given in Example 1.

Solution: W_2 = {(x,y,z) $\in R^3$ | T(x,y,z) = 2(x,y,z)}

= {(x,y,z) $\in R^3$ | (2x,2y,2z) = 2(x,y,z)} = R^3.

Now, try the following exercise.

E E2) For T in Example 2, obtain the complex vector spaces W_i, W_{-i} and W_1.

As with every other concept related to linear transformations, we can define eigenvalues and eigenvectors for matrices also. Let us do so.

Let A be any n x n matrix over the field F.

As we have said in Unit 2 (Theorem 5), the matrix A becomes a linear transformation from $V_n(F)$ to $V_n(F)$, if we define
$A: V_n(F) \quad V_n(F) : A(X) = AX$.
Also , you can see that $[A]Bo = A$, where.

$$Bo = \left\{ e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ . \\ . \\ . \\ . \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ . \\ . \\ . \\ . \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ 1 \end{bmatrix} \right\}$$

is the standard ordered basis of $V_n(F)$ to $V_n(F)$, with respect to the standard basis B_o, is A itself.
This is why we denote the linear transformation A by A itself.

Looking at matrices as linear transformations in the above manner will help you in the understanding of eigenvalues and eigenvectors for matrices.

Definition: A scalar λ is an eigenvalue of an n x n matrix a over F if there exists $X \in V_n(F)$, $X \neq 0$, such that $AX = \lambda X$ are eigenvectors of the matrix A corresponding to the eigenvalue λ.

Let us look at a few examples.

Example 4: Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ Obtain an eigenvalue and a corresponding Eigenvector of A

Solution: Now $A\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\end{bmatrix}$ This shows that 1 is an

eigenvalue and $\begin{bmatrix}1\\0\\0\end{bmatrix}$ is an eigenvector corresponding to it.

Infact, $A\begin{bmatrix}0\\1\\0\end{bmatrix} = 2\begin{bmatrix}0\\1\\0\end{bmatrix}$ and $A\begin{bmatrix}0\\0\\1\end{bmatrix} = 3\begin{bmatrix}0\\0\\1\end{bmatrix}$.

Thus, 2 and a3 are eigenvalues of A, with corresponding

eigenvectors $\begin{bmatrix}0\\1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\\0\end{bmatrix}$, respectively.

Example 5: Obtain an eigenvalue and a corresponding eigenvector

of $A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \in M_2(R)$.

Solution: Suppose $\lambda \in R$ is an eigenvalue of A. then

$\exists\ x = \begin{bmatrix}x\\y\end{bmatrix} \neq \begin{bmatrix}0\\0\end{bmatrix}$ such that $AX = \lambda X$, that is, $\begin{bmatrix}-y\\x+2y\end{bmatrix} = \begin{bmatrix}\lambda x\\\lambda y\end{bmatrix}$

So for what values of λ, x and y are the equations $-y = \lambda x$ and $x + 2y = \lambda y$ satisfied? Note that $x \neq 0$ and $y \neq 0$, because if either is zero then the other will have to be zero. Now, solving our equations we get $\lambda = 1$.

The an eigenvector corresponding to it is $\begin{bmatrix}1\\-1\end{bmatrix}$

Now you can solve an eigenvalue problem yourself!

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

E E3 Show that 3 is an eigenvalue of .Find 2 corresponding eigenvectors.

Just as we defined an eigenspace associated with a linear transformation we define the eigenspace W_λ, corresponding to an eigenvalue λ of an n x n matrix A, as follows:
$W\lambda = \{X \in Vn(F),| AX = \lambda X\} = \{X \in Vn (F) \mid (A - \lambda I) X = 0\}$

For example., the eigenspace W1, in the situation of Example 4, is

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in V_3(R) \;\middle|\; A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in V_3(R) \;\middle|\; \begin{bmatrix} x \\ 2y \\ 3z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\}$$

$$= \left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \;\middle|\; x \in R \right\}$$ which is the same as $\{(x,0,0) \mid x \in R\}$.

E E4) Find W3 for the matrix in E3.

The algebraic eigenvalue problem for matrices is to determine all the eigenvalues and eigenvectors of a given matrix. In fact, the eigenvalues and eigenvectors of an n x n matrix A are precisely the eigenvalues ad eigenvector of a regarded as a linear transformation from $V_n(F)$ to $V_n(F)$.

We end this section with the following remark:

A scalar λ is an eigenvalue of the matrix a if and only if $(A - \lambda I)X = 0$ has a non-zero solution, i.e., if and only if det $(A- \lambda I) = 0$

Similarly, λ is an eigenvalue of the linear transformation T if and only if det $(T - \lambda I) = 0$

So far we have been obtaining eigenvalues by observation, or by some calculations that may not give us all the eigenvalues of a given matrix or linear transformation. The remark above suggests where to look for all the eigenvalues In the next section we determine eigenvalues and eigenvectors explicitly.

6.3 OBTAINING EIGENVALUES AND EIGENVECTORS

In the previous section we have seen that a scalar λ is an eigenvalue of a matrix A if and only if det $(A) - \lambda I) = 0$. In this section we shall see how this equation helps us to solve the eigenvalue problem.

6.3.1 Characteristic Polynomial

Once we know that λ is an eigenvalue of a matrix A, the eigenvectors can easily be obtained by finding non-zero solutions of the system of equations given by $AX = \lambda X$.

Now, if $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ and $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

the equation $AX = \lambda X$ becomes

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

write it out, we get the following system of equations.

$a_{11}x_1 + a_{12}x_2 + \ldots.. + a_{1n}x_n = \lambda x_1$
$a_{21}x_1 + a_{22}x_2 + \ldots.. + a_{2n}x_n = \lambda x_2$
$. \quad . \quad \ldots.. \quad . \quad . \quad .$
$a_{n1}x_1 + a_{n2}x_2 + \ldots.. + a_{nn}x_n = \lambda x_1$
This equivalent to the following system.

$a_{11} -\lambda)x_1 + a_{12}x_2 + \ldots.. + a_{1n}x_n = 0$
$a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots.. + a_{2n}x_n = 0$
$\ldots.. \quad ..$
$a_{n1}x_1 + a_{n2}x_2 + \ldots.. + a_{nn} -\lambda)x_n = \lambda x_1$

This homogeneous system of linear equations has a non-trivial solution if and only if the determinant of the coefficient matrix is equal to 0 (by Unit 9, theorem 5). Thus, λ is an eigenvalue of A if and only if

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots.. & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots.. & a_{2n} \\ . & . & \ldots.. & . \\ . & . & \ldots.. & . \\ a_{n1} & a_{n1} & \ldots.. & a_{nn} - \lambda \end{vmatrix} = 0$$

Now, $\det(\lambda I – A) = (- 1)n \det(A - \lambda I)$ (multiplying each row by (- 1)). Hence, $\det(\lambda I – A) = 0$ if and only if $\det (A - \lambda\lambda= 0$.

This leads us to define the concept of the characteristic polynomial.

Definition: Let $A = [a_{ij}]$ be any n x n matrix. Then the characteristic polynomial of the matrix A is defined by

$$f_A(t) = \det(tI – A)$$

$$= \begin{vmatrix} t – a_{11} & - a_{12} & \ldots.. & - a_{1n} \\ - a_{21} & t – a_{22} & \ldots.. & - a_{2n} \\ : & : & \ldots.. & : \\ - a_{n1} & -a_{n2} & \ldots.. & t – a_{nn} \end{vmatrix}$$

$$= t'' + c_1t''^{\,1} + c_2t''^{\,2} + \ldots. + c_{n1}t + c_n$$

where the coefficients $c_1, c_2, \ldots, c_n$ depend on the entries a_{ij} of the matrix A.

The equation $f_A(t) = 0$ is the characteristic equation of A.

When no confusion arises, we shall simply write f(t) in place of fA(t).

Consider the following example.

Example 6: Obtain the characteristic polynomial of the matrix

1 2

0 -1 .

Solution: The required polynomial is $\begin{vmatrix} t-1 & -2 \\ 0 & t+1 \end{vmatrix}$

$= (t-1)(t+1) = t^2 - 1.$

Note try this exercise.

E E5) Obtain the characteristic polynomial of the matrix

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

Note that λ is an eigenvalue of A iff $\det(\lambda I - A) = f_A(\lambda) = 0$, that is, iff$\lambda$ is a root of the characteristic polynomial $f_A(t)$, defined above. Due to this fact, eigenvalues are also called characteristic root, and eigenvectors are called chrematistic vectors.

For example, the eigenvalues of the matrix in Example 6 are the roots of the polynomial t2 – 1, namely, 1 and 9 – 1).

E E6) Find the eigenvalues of the matrix in E5.

Now, the characteristic polynomial $f_A(t)$ is a polynomial of degreen. Hence, it can have n roots at the most. Thus, an n x n matrix has two eigenvalues, at the most. for example, the matrix in Example 6 has two eigenvalues, 1 and – 1 , and the matrix in E5 has 3 eigenvalues.

Now we will prove a theorem that will help us in section 10.4.

Theorem 1: similar matrice have the same eigenvalues.

Proof: Let an n x n matrix B be similar to an n x n matrix A.

Then, by definition, B= P-1 AP, for some invertible matrix P.

Now, the characteristic polynomial of B,

fB(t) = det(tI – B)
= det(tI –P^{-1} AP)
= det (P^{-1}(tI- A), since P-1tIP = tP-1P = tI
= det(P^{-1}) det(tI- A) det(P) (see sec. 9.4)
= det(tI-A) det(P^{-1}) det(P)
= f_A(t) det(P^{-1}P)
= f_A(t), since det(P-1P) = det(I) = 1.

Thus, the roots of f_B (t) and f_A(t) coincide. Therefore, the eigenvalues of A and B are the same.

Let us consider some more examples so that the concepts mentioned in this section become absolutely clear to you.

Example 7: find the eigenvalues and eigenvectors of the matrix

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

Solution: In solving E6 you found that the eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = -2$. Now we obtain the eigenvectors of A.

The eigenvectors of A with respect to the eigenvalue λ_1 =1 are the non-trivail solutions of

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

which gives the equations

$$\left.\begin{aligned} 2x_3 &= x_1 \\ x_1 + x_3 &= x_2 \\ x_2 - 2x_3 &= x_3 \end{aligned}\right\} \Rightarrow \quad \begin{aligned} x_1 &= 2x_3 \\ x_2 &= x_1 + x_3 = 3x_3 \\ x_3 &= x_3 \end{aligned}$$

The eigenvectors corresponding to $\lambda_1 = 0$ are given by

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

which gives $\quad x_1 + x_2 = 0$

$$\begin{aligned} -x_1 - x_2 &= 0 \\ -2x_1 - 2x_2 + 2x_3 + x_4 &= 0 \\ x_1 + x_2 - x_3 &= 0 \end{aligned}$$

The first and last equations give $x_3 = 0$. Then, the third equation gives $x_4 = 0$. The first equation gives $x_1 = -x_2$.

Thus, the eigenvectors are

$$\begin{bmatrix} -x2 \\ x2 \\ 0 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} x_2 \neq 0, x_2 \in R.$$

The eigenvectors corresponding to $\lambda_2 = 1$ are given by

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

which gives $\quad x_1 + x_2 = x_1$

$$\begin{aligned} -x_1 - x_2 &= x_2 \\ -2x_1 - 2x_2 + 2x_3 + x_4 &= x_3 \\ x_1 + x_2 - x_3 &= x_4 \end{aligned}$$

The first two equations give $x_2 = 0$ and $x_1 = 0$. Then the last equation gives $x_4 = -x_3$. Thus, the eigenvectors are

$$\begin{bmatrix} 0 \\ 0 \\ x_3 \\ -x_3 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \quad x_3 \neq 0, x_3 \in R.$$

Example 9: Obtain the eigenvalues and eigenvectors of

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

Solution: the characteristic polynomial of A = $f_A(t)$ = det(tI-A)

$$= \begin{vmatrix} t & -1 & 0 \\ -1 & t & 0 \\ 0 & 0 & t-1 \end{vmatrix} = (t+1)\,(t-1)^2$$

Therefore, the eigenvalues are λ_1= and $\lambda_2 = 1$.

The eigenvectors corresponding to $\lambda_1 = -1$ are given by

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (-1) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

which is equivalent to

$$\begin{aligned} x_2 &= x_1 \\ x_1 &= x_2 \\ x_3 &= x_3 \end{aligned}$$

The last equation gives $x_3 = 0$. Thus, the eigenvectors are

$$\begin{bmatrix} x_1 \\ -x_1 \\ 0 \end{bmatrix} = x_1 \begin{matrix} 1 \\ -1 \\ 0 \end{matrix}, \quad x1 \neq 0,\ x1 \in R.$$

The eigenvectors corresponding to $\lambda 2 = 1$ are given by

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (-1) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

which gives

$$\begin{aligned} x_2 &= x_1 \\ x_1 &= x_2 \\ x_3 &= x_3 \end{aligned}$$

Thus, the eigenvectors are

$$\begin{bmatrix} x_1 \\ x_1 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

where x_1, x_3 are real numbers, not simultaneously 0.

Note that, corresponding to $\lambda_2 = 1$, there exist two linearly independent eigenvectors.

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$ which form a basis of the eigenspace W_1.

Thus, W-1 is 1 – dimensional, while dim R W1 = 2.
Try the following exercises now.

E E7) find the eigenvalues and bases for the eigenspaces of the matrix.

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

E E8) Find the eigenvectors of the diagonal matrix

$$D = \begin{bmatrix} a_1 & 0 & 0 & . & . & 0 \\ 0 & a_2 & 0 & . & . & 0 \\ 0 & 0 & a_3 & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & a_n \end{bmatrix}$$, where $a_i \neq a_j$ for $i \neq j$.

We now turn to the eigenvalues and eigenvectors of linear transformations.

6.3.2 Eigenvalues of Linear Transformations

As in section 10.2, let $T:V \to V$ be a linear transformation on a finite-dimensional vector space V over the field F. we have seen that

$\lambda \in F$ is an eigenvalue of T

$\Leftrightarrow \det(T - \lambda I) = 0$

$\Leftrightarrow \det(\lambda I - T) = 0$

$\Leftrightarrow \det(\lambda I - A) = 0$, where $A = [T]_B$ is the matrix of T with respect to a basis B of V. Note that $[\lambda I - T]B = \lambda I - [T]_B$.

This shows that λ is an eigenvalue of T if and only if λ is an eigenvalue of the matrix $A = [T]_B$, where B is a basis of V. We define the characteristic polynomial of the linear transformation T to be same as the characteristic polynomial of the matrix $A = [T]_B$. where B is basis V.

This definition does not depend on the basis B chosen, since similar matrice have the same characteristic polynomial (Theorem 1), and the matrices of the same linear transformation T with respect to two different ordered bases of V are similar.

Just as for matrices, the eigenvalues of T are precisely the roots of the characteristic polynomial of T.

Example 10: Let $T : R^2 \to R^2$ be the linear transformation which maps $e_1 = (1,0)$ to $e_2 = (0,1)$ and e_2 to $- e_1$. Obtain the eigenvalues of T.

Solution: Let $A = [T]B = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$, where $B = \{e_1,e_2\}$.

The characteristic polynomial of T = the characteristic polynomial of A

$$= \begin{vmatrix} T & 1 \\ -1 & t \end{vmatrix} = t^2 + 1, \text{ which has no real roots.}$$

Hence, the linear transformation T has n real eigenvalues. But, it has two complex eigenvalues i and - i

Try the following exercise now.

E E9) Obtain the eigenvalues and eigenvectors of the differential operator
$D: P2 \to P2$:

$D(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$, for $a_0, a_1, a_2 \in R$.

E E10) show that the eigenvalues of a square matrix A coincide with those of A^t.

E E11) Let a be an invertible matrix. If λ is an eigenvalue of A, show that $\lambda \neq 0$ and that λ^{-1} is an eigenvalue of A^{-1}.

Now that we have discussed a method of obtaining the eigenvalues and eigenvectors of a matrix, let us see how they help in transforming any square matrix into a diagonal matrix.

6.4 Diagonalisation

In this section we start with proving a theorem that discusses the linear independence

Theoroem 2: Let T: V $\rightarrow$ V be a linear transformation on a finite-dimensional vector space V over the field F. $\lambda_1, \lambda_2, \ldots.\lambda_m$ be the distinct eigenvalues of T and $v_1, v_2, \ldots.,$ v_m be eigenvector of T corresponding to $\lambda_1, \lambda_2, \ldots. \lambda_m$, respectively. Then $v_1, v_2, \ldots., v_m$ are linearly independent over F.

Proof: We know that

$Tv_i = \lambda_i v_i, \lambda_i \in F$, $0 \neq v_i \in V$ for i = 1,2,…, m, and $\lambda_i \neq \lambda_j$ for $i \neq j$.

Suppose, if possible, that $\{v_1 v_2, \ldots., v_m\}$ is a linearly dependent set. Now, the single non-zero vector v_1 is linearly independent. We choose $r(\leq m)$ such that

$\{v_1, v_2, \ldots, v_{r-1}\}$is linearly independent and $\{v_1, v_2, \ldots, v_{r-1}\ v_r\}$is linearly dependent. Then $v_r, a_1v_1 + a_2v_2 + \ldots + a_{r-1}v_{r-1}$ ……. (1)
for some $a_1, a_2, \ldots, a_{r-1}\ v_{r-1}$ in F.
Applying T, we get
$Tv_r = a_1 Tv_1 + \ldots + a_{r-1} Tv_{r-1}$. This gives
$\lambda_r v_r = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \ldots + a_{r-1}\lambda_{r-1}v_{r-1}$ ....(2)

Now, we multiply (1) by λ_r and subtract it from (2), to get
$0 = a_1 (\lambda_1 - \lambda_r) v_1 + a_2 (\lambda_2 - \lambda_r)v_2 + \ldots + a_{r-1} (\lambda_{r-1} - \lambda_r)v_{r-1}$
Since the set $\{v_1, v_2, \ldots, v_{r-1}\}$if linearly independent, each coefficients in the above equation must be 0. thus, we have $a_i ((\lambda_1 - \lambda_r) = 0$ - 1,2 …….r-1.

But $\lambda_2 \neq \lambda_r$ for i = 1,2,……. R-1. hence $(\lambda_i - \lambda_r) \neq 0$ for i = 1,2….., r-1, and we must have $a_i = 0$ for i= 1,2,……, r-1. However, this is not possible since (1) would imply that $v_r = 0$, and, being an eigenvector, v_r can never be 0. Thus, we reach a contradiction.

Hence, the assumption we started with must be wrong. Thus, $\{v_1, v_2, \ldots, v_m\}$must be linearly independent, and the theorem is proved.

We will use theorem 2 to choose a basis for a vector space V so that the matrix $[T]_B$ is a diagonal matrix.

Definition: A linear transformation T : V $\to$ V on a finite-dimensional vector space V is said to be diagonalizable if there exists a basis $B = \{v_1, v_2, \ldots, v_n\}$of V such that the matrix of t with respect to the basis B is diagonal. That is,

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

where $\lambda_1, \lambda_2, \ldots \lambda_n$ are scalars which need not be distinct.

The next theorem tells us under what conditions a linear transformation is diagonalizable.

Theorem 3: A linear transformation T, on a finite-dimensional vector space V, is diagonalizable if and only if there exists a basis of V consisting of eigenvector of T.

Proof: suppose that T is diagonalizable. By definition, there exists a basis B = $\{v_1,v_2,\ldots, v_n\}$ of V, such that

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & . \\ .. & .. & .. & \ldots & . \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

By definition of $[T]_B$, we must have
$Tv_1 = \lambda_1 v_1$, $Tv_2 = \lambda_2 v_2$, ………, $Tv_n = \lambda_n v_n$.

Since basis vectors are always non-zero, $v_1,v_2,\ldots, v_n$ are non-zero. Thus, we find that $v_1,v_2,\ldots, v_n$ are eigenvecyors of T.

Conversely, let b = $\{ v_1,v_2,\ldots, v_n\}$ be a basis of V consisting of eigenvectors of T. Then, are exist scalars $a_1,a_2,\ldots, a_n$, not necessarily distinct, such that $Tv_1 = a_1$, $Tv_2 = a_2v_2,\ldots., Tv_n = a_nv_n$.

But then we have

$$[T]_B = \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \alpha_n \end{pmatrix}$$ which means that T is diagonalizable.

the next theorem combines theorem 2 and 3

Theorem 4: Let $T:V \rightarrow V$ be a linear transformation, where V is an n-dimensional vector space. Assume that T has n distinct eigenvalues. Then T is diagonalizable.

Proof: Let $\lambda_1\lambda_2,\ldots\ldots\lambda_n$ be the n distinct eigenvalues of T. Then there exist eigenvectors $\lambda_1\lambda_2,\ldots\ldots\lambda_n$ corresponding to the eigenvalues $\lambda_1\lambda_2,\ldots\ldots\lambda_n$, respectively. By theorem 2, the set, $v_1v_2,\ldots\ldots v_n$, is linearly independent and has n vectors, where n = dim V. Thus, from Unit 5 (corollary to Theorem 5), B = $\{v_1v_2,\ldots\ldots v_n\}$is a basis of V consisting of eigenvectors of T. Thus, by theorem 3, T is diagonalizable.

Just as we have reached the conclusion of Theorem 4 for linear transformations, we define diagonalisability of a matrix, and reach a similar conclusion for matrices.

Definition: An n x n matrix A is said to be diagonalizable if A is similar t a diagonal matrix, that is, $P^{-1}AP$ is diagonal for some non-singular nxn matrix P.

Note that the matrix A is diagonalizable if and only if the matrix A, regarded as a linear transformation $A:V_n(F)\rightarrow V_n(F) = AX$, is diagonalizable.

Thus, Theorem 2,3, and 4 are true for the matrix A regarded as a linear transformation from $V_n(F)$ to $V_n(F)$. Therefore, given an nxn matrix A, we know that it is diagonalizable if it has n distinct eigenvalues.

We now give a practical method of diagonalaising a matrix.

Theorem 5: Let a be an nxn matrix having n distinct eigenvalues $\lambda_1\lambda_2,\ldots\ldots\lambda_n$. let $X_1,X_2,\ldots\ldots X_n\in V_n(F)$ be eigenvectors of A corresponding to $\lambda_1\lambda_2,\ldots\ldots\lambda_n$, respectively. Let $P = (X_1,X_2,\ldots\ldots X_n)$ be the nxn matrix having $X_1,X_2,\ldots\ldots X_n$ as its column vectors. Then.

$P\text{-}1AP = \text{diag}\ (\lambda_1\lambda_2,\ldots\ldots\lambda_n)$.

Proof: By actual multiplication, you can see that

$$
\begin{aligned}
AP &= A(X_1,X_2,\ldots\ldots X_n) \\
&= (AX_1,AX_2,\ldots\ldots AX_n) \\
&= (\lambda_1X_1,\lambda_2X_2,\ldots\ldots\lambda_nX_n \\
&= (X_1,X_2,\ldots\ldots X_n) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ . & . & \ldots & . \\ . & . & \ldots & . \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \\
&= P\ \text{diag}\ (\lambda_1\lambda_2,\ldots\ldots\lambda_n)
\end{aligned}
$$

now, by Theorem 2, the column vectors of P are linearly independent. This means that P is invertible (Unit 9, Theorem 6). Therefore, we can pre-multiply both sides of the matrix equation $AP = P\ \text{diag}\ (\lambda_1\lambda_2,\ldots\ldots\lambda_n)$.

Let us see how this theorem works in practice.

Example 11: Diagonalise the matrix

$$
A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & -6 \\ 2 & -2 & 3 \end{bmatrix}
$$

Solution: The characteristic polynomial of A = f(t) =

$$
\begin{vmatrix} t-1 & -2 & 0 \\ -2 & t-1 & 6 \end{vmatrix} = (t-5)\ (t-3)\ (t+3).
$$

-2 2 t-3

Thus, the eigenvalues of A are $\lambda_1 = 5$, $\lambda_2 = 3$, $\lambda_3 = -3$ since they are all distinct, A is diagonalizable (by theorem 4). You can find the eigenvectors by the method already explained to you. Tight now you can directly verify that.

$$A\begin{pmatrix}1\\2\\-1\end{pmatrix} = 5\begin{pmatrix}1\\2\\-1\end{pmatrix}, \quad A\begin{pmatrix}1\\1\\0\end{pmatrix} = 3\begin{pmatrix}1\\1\\0\end{pmatrix} \text{ and } A\begin{pmatrix}-1\\2\\1\end{pmatrix} = -3\begin{pmatrix}-1\\2\\1\end{pmatrix}$$

Thus, $\begin{pmatrix}1\\2\\-1\end{pmatrix}$, $\begin{pmatrix}1\\1\\0\end{pmatrix}$ and $\begin{pmatrix}-1\\1\\0\end{pmatrix}$

are eigenvectors corresponding to the distinct eigenvalues 5,3 and -3, respectively. By Theorem 5, the matrix which diagonalises A is given by

$$P = \begin{pmatrix}1 & 1 & -1\\2 & 1 & 2\\-1 & 0 & 1\end{pmatrix}.$$ Check, by actual multiplication, that

$$P = \begin{pmatrix}5 & 0 & 0\\0 & 1 & 0\\0 & 0 & -3\end{pmatrix},$$ which is in diagonal form.

The following exercise will give you some practice in diagonalising matrices.

E E12) Are the matrices in Examples 7,8 and 9 diagonalisable? If so, diagonalise them.

We end this unit by summarizing what has been done in it.

6.5 Summary

As in previous unit, in this unit also we have treated linear transformations along with the analogous matrix version. We have covered the following point here.

1) The definition of eigenvalues, eigenvectors and eigenspaces of linear transformations and matrices.

2) The definition of the characteristic polynomial and characteristic equation of a linear transformation (or matrix).

3) A scalar λ is an eigenvalue of a linear transformation T(or matrix A) if and only if it is a root of the characteristic polynomial of T (or A).

4) A method of containing all the eigenvalues and eigenvectors of a linear transformation (or matrix).

5) Eigenvectors of a linear transformation (or matrix) corresponding to distinct eigenvalues are linearly independent.

6) A linear transformation T: V $\rightarrow$ V is diagonalizable if and only if V has a basis consisting of eigenvectors of t.

7) A linear transformation (or matrix) is diagonalizable if its eigenvaules are distinct.

6.6 Solutions/Answers

E1) Supose $\lambda \in R$ is an eigenvalue. Then $\exists$ $(x,y) \neq (0,0)$ such that $T(x,y) = \lambda (x,y)$ $\Rightarrow (x,0) = (\lambda x, \lambda y) \Rightarrow \lambda x = x, \lambda y = 0$. These equations are satisfied if $\lambda = 1, y = 0$ $\therefore$, 1 is an eigenvalue. A corresponding eigenvector of (1,0). Note that there are infinitely many eigenvectors corresponding to 1, namely, $(x,0) \quad 0 \neq \forall x \in R$.

E2) $W_1 = \{(x,y,z) \in C3 \mid T(x,y,z) = i(x,y,z)\}$

$= \{(x,y,z) \in C3 \mid (ix, -iy, z) = (ix,iy,iz)\}$

$= \{(x,0,0) \mid x \in C\}$.

Similarly, you can show that $w_{-1} = \{(0,x,0) \mid x \in C\}$ and $W_1 = \{(0,0,x) \mid x \in C\}$.

E3) if is an eigencalve , then $\exists \begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ such that

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 3 \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow x + 2y = 3x \text{ and } 3y = 3y.$$

These equations are satisfied by x = 1, y = 1 and x = 2, y = 2.

$\therefore$ 3 is an eigenvalue, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ are eigevectors corresponding to 3.

E4) $W_3 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in V_2(R) \;\middle|\; \begin{bmatrix} x+2y \\ 3y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix} \right\}$

$= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in V_2(R) \;\middle|\; x = y \right\} = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} \;\middle|\; x \in R \right\}$

This is the 1-dimensional real subspace of $V_2(R)$ whose basis is $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

W5) It is $\begin{vmatrix} t & 0 & -2 \\ -1 & t & -1 \\ 0 & -1 & t+2 \end{vmatrix} = t \begin{vmatrix} t & -1 \\ -1 & t+2 \end{vmatrix} + \begin{vmatrix} 0 & -2 \\ -1 & t+2 \end{vmatrix}$

$= \{t^2(t+2)-t\} - 2 = t^3 + 2t^2 - t-2.$

E6 The eigenvalues are the roots of the polynomial $t^3+2t^2 - t-2 = (t-1)(t+1)(t+2)$

$\therefore$ they are 1, -1, -2.

E7 $fA(t) = \begin{bmatrix} t-2 & 1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{bmatrix} = (t-2)^2 (t-3)$

$\therefore$ the eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 3$.

The eigenvectors corresponding to λ_1 are given by

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

This leads us to the equations.

$$\begin{aligned} 2x + y &= 2x \\ y - z &= 2y \\ 2y + 4z &= 2z \end{aligned} \quad \Rightarrow \quad \begin{aligned} x &= x \\ y &= 0 \\ z &= 0 \end{aligned}$$

$\therefore W_2 = \left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \right\}$ $x \in R$ $\therefore$, a basis for W_2 is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

The eigenvectors corresponding to λ_2 are given by

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$ This gives us the equations.

$$\begin{matrix} 2x + y = 3x \\ y - z = 3y \\ 2y + 4z = 3z \end{matrix} \Rightarrow \begin{matrix} x = x \\ y = x \\ z = 2x \end{matrix}$$

$\therefore$ W3 = $\left\{ \begin{pmatrix} x \\ x \\ -2x \end{pmatrix} \right\}$ $x \in R$ $\therefore$, a basis for W_2 is $\left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$

E8) $f_D(t) = \begin{vmatrix} t-a_1 & 0 & \dots & 0 \\ 0 & t-a_2 & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & t-a_n \end{vmatrix} = $ (t-a1) (t-a2) …. (t-an)

$\therefore$, its eigenvalues are $a_1, a_2 \dots, a_n$.

The eigenvectors corresponding to a1 are given by

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ . & . & \dots & : \\ 0 & 0 & \dots & a_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ : \\ x_n \end{pmatrix} = a_1\begin{pmatrix} x_1 \\ x_2 \\ : \\ x_n \end{pmatrix}$$

This gives us the equations

$$\left\{\begin{matrix} a_1x_1 = a_1x_1 \\ a_2x_2 = a_2x_2 \\ : \quad : \end{matrix}\right\} \qquad \begin{matrix} x_1 = x_1 \\ x_2 = 0 \\ : \quad : \end{matrix}$$

(since $a_n \neq a_j = a_j$ for $i \neq j$

$\therefore$ The eigenvector corresponding to a1 are $\begin{pmatrix} x_1 \\ 0 \\ : \\ 0 \end{pmatrix}$, $x_1 \neq 0$, x_1, $\in R$

Similarly, the eigenvectors corresponding to a_1 are $\begin{pmatrix} 0 \\ : \\ 0 \\ , \\ x_1 \\ 0 \\ : \end{pmatrix}$ $x_2 \neq 0$, x_1, $\in R$.

0

E9) $B = \{1, x, x^2\}$ is a basis of P_2

Then $[D]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

∴, the characteristic polynomial of D is $\begin{vmatrix} t & -1 & 0 \\ 0 & t & -2 \\ 0 & 0 & t \end{vmatrix} = t^3$

∴, the only eigenvvalue of D is $\lambda = 0$.

The eigenvectors corresponding to $\lambda = 0$ are $a_0 + a_1x + 4_2x^2$, where $D(a_0 + a_1x + a_2x^2)$

$= 0$, that is $a_1 + 2a_2x = 0$.

This gives a1 = 0, a2 = 0 ∴, the set of eigenvectors corresponding to $\lambda = 0$ are $\{a0 | a0 \in R, a0 \neq 0\} = R \setminus \{0\}$.

E10) $|tI - A| = |(tI - A)^t|$, since $|A^t| = |A|$.

$= |tI - A^t|$, since $I^t = I$ and $(B -)1 = B^t - C^t$.

∴, the eigenvalues of A are the same as those of A^t

E11) Let X be an eigenvector corresponding to λ. Then $X \neq 0$ and $AX = \lambda X$.

∴ A-1 (AX) = A-1 (λX).

$\Leftrightarrow (A^{-1}A)\, x = \lambda(A^{-1}X)$

$\Leftrightarrow X = \lambda(A^{-1}X)$

$\Leftrightarrow \lambda \neq 0$, since $X \neq 0$.

Also, $X = \lambda(A^{-1}X) \Rightarrow \lambda^{-1}X \Leftrightarrow \lambda^{-1}$ is an eigenvalue of A^{-1}.

E12) Since the matrix in Example 7 has distinct eigenvalues 1, -1 and -2, it is diagonalizable. Eigenvectors corresponding to.

These eigenvalue are $\begin{bmatrix} 1 \\ \\ 1 \end{bmatrix}$ 3 $\begin{bmatrix} -2 \\ \\ 1 \end{bmatrix}$ 1 $\begin{bmatrix} 1 \\ \text{and} \\ 1 \end{bmatrix}$ 0 , respectively

∴, if P $\begin{bmatrix} 2 & -2 & -1 \\ & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ 0, then $\begin{bmatrix} 0 & 0 & 2 \\ P\text{-}1 & & 1 \\ 0 & 1 & -2 \end{bmatrix}$ 0 $\begin{bmatrix} 1 & 0 & 0 \\ 1 & P = & 0 \\ 0 & 0 & -2 \end{bmatrix}$ -1 0.

The matrix in Example 8 is not diagonalizable. This is because it only has two distinct eigenvalues and, corresponding to each , it has only linearly independent eigenvector. ∴, we cannot find a basis of $V_4(F)$ consisting of eigenvectors. And now apply Theorem 3.

The matrix in Example 9 is diagonalizable though it only has two distinct eigenvalue. This is because corresponding to $\lambda_1 = -1$ there is one linear independent eigenvector, but corresponding to $\lambda_2 = 1$ there exist two linearly independent eigenvectors. Therefore, we can form a basis V3(R) consisting of the eigenvectors.

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

The matrix $P = \begin{matrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$ is invertible , and

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

UNIT 7

CHARACTERISTIC AND MINIMAL POLYNOMIAL

7.1 Introduction

This unit is basically a continuation of the previous unit, but the emphasis is on a different aspect of the problem discussed in the previous unit.

Let $T: V \rightarrow V$ be a linear transformation on a n-dimensional vector space V over the field F. The two most important polynomials that are associated with T are the characteristic polynomial of T and the minimal polynomial of T. we defined the former in unit 10 and the latter in unit 6.

In this unit we first show that every square matrix (of linear transformation $T: V \rightarrow V$) satisfies its characteristic equation, and use this to compute the inverse of the concerned matrix (or linear transformation), if it exists.

Then we define the minimal polynomial of a square matrix, and discuss the relationship between the characteristic and minimal polynomials. This leads us to a simple way of obtaining the minimal polynomial of a matrix (or linear transformation).

We advise you to study Unit 2.4 and 6 before starting this unit.

Objectives

After studying this unit, you should be able to

- State and prove the Cayley-Hamilton theorem;
- Find the inverse of an invertible matrix using this theorem;
- Prove that a scalar λ is an eigenvalue if and only it is a root of the minimal polynomial;
- Obtain the minimal polynomial of a matrix (or linear transformation) if the characteristic polynomial is known.

7.2 Cayley-Hamilton Theorem

In this section we present the Cayley-Hamilton theorem, which is related to the characteristic equation of a matrix. It is named after the British mathematicians Arthur Cayley (1821-1895) and William Hamilton (1805 – 1865, who were responsible for a lot of work done in the theorem of determinants.

Let us consider the 3 x 3 matrix $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix}$

Then $tI - A = \begin{pmatrix} t & -1 & -2 \\ 1 & t-2 & -1 \\ 0 & -3 & t-2 \end{pmatrix}$

Let Δji be the (i,j)th cofactor of (tI – a.

Then $\Delta_{11} = (t-2)^2 - 3 = t^2\ 4t + 1$, $\Delta_{12} = t - 2$, $\Delta_{13} = -3$, $\Delta_{21} = t + 4$, $\Delta_{22} = t^2 - 2t$,

$\Delta_{23} = 3t$, $\Delta_{31} = 2t – 3$, $\Delta_{32} = t – 2$, $\Delta_{33} = t^2 – 2t + 1$.

$$\therefore \text{Adj}\,(tI – A) = \begin{bmatrix} t^2 – 4t + 1 & t + 4 & 2t – 3 \\ t – 2 & t^2 – 2t & t – 2 \\ - 3 & 3t & t^2 – 2t + I \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} t2 + \begin{bmatrix} -4 & 1 & 2 \\ 1 & -2 & 1 \\ 0 & 3 & -2 \end{bmatrix} t + \begin{bmatrix} 1 & 4 & -3 \\ -2 & 0 & -2 \\ -3 & 0 & 1 \end{bmatrix}$$

This is a polynomial in t of degree 2, with matrix coefficients.

Similarly, if we consider the n x n Matrix A = $[a_{ij}]$, then adj(tI – A) is a polynomial of degree ≤ n -1, with matrix coefficients. Let

Adj (tI – A) = $B_1t^{n-1} + B_2t^{n-2} + \ldots + B_{n-1} + B_n$ …(1)
Where B1,….., B_n are n x n matrices over F.
Now, the characteristic polynomial of A is given by
F(t) = $f_A(t)$ = det(tI – A) = |tI – A|

$$= \begin{vmatrix} t – a_{11} & - a_{12} & . & . & . & - a_{1n} \\ - a_{21} & t – a_{22} & . & . & . & - a_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ - a_{n1} & -a_{n2} & & \ldots\ldots & & t – a_{nn} \end{vmatrix}, \text{ where } a = [a_{ij}]$$

$= t^n + c_1t^{n-1} + c_2t^{n-2} + \ldots\ldots + c_{n-1}t + c_n$, …(2)
where the coefficients (1) and (2) to prove the Cayley-Hamilton theorem.

Theorem 1 (Cayley-Hamilton): Let $f(t) = t^n + c_1t^{n-1} + \ldots\ldots + c_{n-1}t + c_n$ be the characteristic Polynomial of an n x n matrix A. then,
$F(A) = A^n + c_1 A^{n-1} + c_2A^{n-2} + \ldots. + c_{n-1}A + c_n1 = 0$
(Note that over here 0 denotes the n x n zero matrix, and $I = I_n$.)
(tI – A) adj(tI – A) = Adj(tI – A)
= det(tI – A)I
= f(t)I.

now equation (1) above says that
Adj(tI) = $B_1t^{n-1} + B_2t^{n-2} + \ldots\ldots + B_n$, where B_k is an n x n matrix for k = 1,2,….,n.
Thus, we have
$(tI – A) (B_1t^{n-1} + B_2t^{n-2} + B_3t^{n-3} + \ldots\ldots + B_{n-2}t^2 + B_{n-1}t + B_n)$

$= f(t) I$
$= It^n + c_1t^{n-1} + c_2It^{n-2} + \ldots\ldots + c_{n-1}It^2 + c_{n-1}It + c_nI$, substituting the value of f(t).
$= f(t)$

Now, comparing constant term and the coefficients of t, t^2,….., t^n on both sides we get.

$$
\begin{aligned}
&- AB_n = c_nI \\
B_n &- AB_{n-1} = c_n\, tI \\
B_{n-1} &- Ab_{n-2} = c_n\text{-}2I \\
&\ \vdots \\
B, &- AB_2 = c_2I \\
B_2 &- AB_1 = c_1I \\
&B_1 = I
\end{aligned}
$$

Pre-multiplying the first equation by I, the second A, third by a_2, ….., that last by An, and adding all these equations, we get.$0 = c_nI + c_{n-1}A + c_{n-2}A^2 + \ldots + c_2A^{n-2} + c_1A^{n-1} + A^n = f(A)$.
Thus, $f(A) = A^n + c_1a^{n-1} + c_2A^{n-2} \ldots + c_{n-1}A + c_{nI} = 0$, and the Cayley-Hamilton theorem is proved.

This theorem can also stated as
"Every square matrix satisfies its characteristic polynomial".
Remark 1: You may be templed to give the following 'quick' proof Theorem 1:

$f(t) = \det (tI – A)$
$\Rightarrow$ $f(A) = \det (AI – A) = \det (A- A) = \det (0) = 0$

This proof is false. Why? Well , the left hand side of the above equation, f(A), is an n x n matrix while the right hand side is the scalar 0, being the value of det(0).

Now, as usual, we give the analogue of Theorem 1 for linear transformations.

Theorem 2 (Calyley-Hamilton): Let T be a linear transformation on a finite-dimensional vector space V. if f(t) is the characteristic polynomial of T, then f(T) = 0.

Proof: Let dim V = n, and let B = {v1, v2, ……, vn} be a basis of V. In Unit 10 we have observed that
f(t) = the characteristic polynomial of T
= the characteristic polynomial of the matrix [T]B.

Let [T]B = A.

If $f(t)\ t^n + c_1t^{n-1} + c_2t^{n-2} + \ldots + c_{n-1}t + c_n$, then, by Theorem 1,

$F(A) = A^n + c_1A^{n-1} + c_2A^{n-2} + \ldots..+ c_{n-1}A + c_nI = 0.$

Now, in Theorem 2 of Unit 7 we proved that $[\quad]_B$ is a vector space isomorphism. Thus,

$$[f(T)]_B = [T^n + c_1T^{n-1} + c_2T^{n-2} + \ldots + cn\text{-}1T + c_nI]_B$$
$$= [T]^n_B + c_1[T]^{n}_B + c_2[T]^{n}_B + \ldots \_ c_{n-1}[T]_B + c_n[I]_B$$
$$= [A^n + c_1A^{n-1} + c_2A^{n-2} + \ldots + cn\text{-}1A + c_nI$$
$$= f(A) = 0$$

Again, using the one-one property of [ ]B, this implies that f(T0 = 0.
Thus, Theorem 2 is true.

Let us look at some examples now.

Example 1:Verify the Cayley-Hamilton theorem for $A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$

Solution: The characteristic polynomial of A is

$$\begin{vmatrix} t-2 & -1 \\ 1 & t \end{vmatrix} - t^2 - 3t + 2$$

$\therefore$, we want to verify that $A^2 -3A + 21 = 0$.

Now, $A^2 = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ -3 & -2 \end{bmatrix}$

$\therefore$ $A2 - 3A + 21 = \begin{bmatrix} 7 & 6 \\ -3 & -2 \end{bmatrix} - \begin{bmatrix} 9 & 6 \\ -3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$\therefore$, the Cayley-Hamilton theorem is true in this case.

E verify the Cayley-Hamilton theorem for A, where A =

a) $\begin{bmatrix} 7 & 6 & 0 \\ 2 & 3 & 0 \\ -2 & -2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -2 & -1 \end{bmatrix}$, c) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 3 & 3 & 4 \end{bmatrix}$

We will now use Theorem 1 to prove a result that gives us a method for obtaining the inverse of an invertible matrix.

Theorem 3: Let $f(t) = t + c_1t^{n-1} + \ldots.. + c^{n-1}t + c_n$ be the characteristic polynomial of an nxn matrix A. Then exists if $c_n \neq 0$ and, in this case,

$A^{-1} = -c_n^{-1} (A^{n-1} + \ldots.. + c_{n-1}I)$.

Proof: By Theorem 1,

$f(A) = A^n + c_1A^{n-1} + \ldots + c_{n-1}A + c_nI = 0$

$\Rightarrow A(A^{n-1} + c_1A^{n-2} + \ldots + c_{n-1}I) = - c_nI$

and $(A^{n-1} + c_1A^{n-2} + \ldots + c_{n-1}I)A = - c_nI$

$\Rightarrow A [-c_n^{-1} (A^{n-1} + c_1A^{n-2} + \ldots + c_{n-1}I)] = I$

$= [- c_n^{-1} (A^{n-1} + c_1A^{n-2} + \ldots + c_{n-1}I)] = A$.

Thus, A is invertible, and

$A^{n-1} = - c_1A^{n-2} + \ldots + c_{n-1}I)$.

Let us see how Theorem 3 works in practice.

Example 2: Is $A = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & 1 & 3 \end{pmatrix}$ invertible? If so, find A^{-1}.

Solution: The characteristic polynomial of A, f(t)

$$= \begin{vmatrix} t-2 & -1 & -1 \\ 1 & t-2 & 1 \\ 1 & -1 & t-3 \end{vmatrix} = t^3 - 7t^2 + 19t - 19.$$

Since the constant term of $f(t) = -19 \neq 0$, a is invertible.

Now, by Theorem 1, $f(A) = A^3 - 7A^2 + 19A - 19I = 0$

$\Rightarrow (1/19)\ A\ (A^2 - 7A + 19I) = I$

Therefore, $A^{-1} = (1/19)\ (A^2 - 7A + 19I)$

Now, $A^2 = \begin{pmatrix} 2 & 5 & 4 \\ -3 & 2 & -6 \\ -6 & 4 & 7 \end{pmatrix}$

$$\begin{pmatrix} 7 & -2 & -3 \\ 4 & 7 & 1 \\ 1 & -3 & 5 \end{pmatrix}$$

$\therefore A^{-1} = 1/19$

To make sure that there has been no error in calculation, multiply this matrix by a. you should get I!

Now try the following exercise.

E E2) For the matrices in E1, obtain A-1, wherever possible.

Now let us look closely at the minimal polynomial.

7.3 Minimal Polynomial

In Unit 6 we define the minimal polynomial of a linear transformation $T:V \to V$. We said that it is the **monic Polynomial of least degree** with coefficients in F, which is satisfied by T. But, we weren't able to give a method of obtaining the minimal polynomial of T.

In this section we will show that the minimal polynomial divides the characteristic polynomial. Moreover, the roots of the minimal polynomial are the same as those of the characteristic polynomial. Since it is easy to obtain the characteristic polynomial of T, these facts will give us a simply way of finding the minimal polynomial of T.

Let us first recall some properties of the minimal polynomial of T that we gave in Unit 6. Let p9t) be the minimal polynomial of T, then
MP1) p(t) is a monic polynomial with coefficients in F.

MP2 If q(t) is a non-zero polynomial over F such that deg q< deg p, then q(T) $\neq$ 0.
MP3 If, for some polynomial g(over F, g(T) = 0, then p(t)| g(t). That is, there Exists a polynomial h(t) over F such that g(t) = p(t)h(t).

We will now obtain the first link in the relationship between the minimal polynomial and the characteristic polynomial of a linear transformation divides its characteristic polynomial.
Proof: Let the characteristic polynomial and the minimal polynomial of T be f(t) and p(t), respectively. By Theorem 2, f(T) = 0. Therefore, by MP4, p(t) divides f(t), as desired.

Before going on to show the full relationship between the minimal and characteristic polynomials, we state (but don't prove!) two theorems that will be used again and again, in this course as well as other courses.

Theorem 5 (Division algorithan for polynomials): Let f and g be two polynomials in t with coefficients in a field F such that f $\neq$ 0. then
a) there exist polynomials and r with coefficients in F such that g = fq + r, where r = 0 or deg r < deg f, and
b) if we also have g + fq_1 + r_1,with r_1 = 0 or deg r_1< deg f, then q = q_1 and r = r_1

An immediate corollary follows.
Corollary: if g is a polynomial over F with $\neq\in$ as a root then g(t) = (t-λ) q(t), for some polynomial q over F.

Proof: By the division algorithm, taking f = (t-λ) we get
g(t) = (t-λ) q(t) + r(t), …….(1)
with r = 0 or deg r < deg (t-λ) = 1.
If deg r < 1, then r is a constant.
Putting t = λ in (1) gives us
g(λ)= r(λ) = r, since r is a constant. But g(λ) = 0, since λ is a root of g. $\therefore$, r = 0. Thus, the only possibility is r = 0. Hence, g(t) = (t - λ) q(t).

And now we come to a very important result that you may be using often, without realizing coefficients has at least one root in C.

In other words, this theorem says that any polynomial f(t) = $\alpha_{n-1}t^{n-1}$ + … + α1t + α_0 (where α_0, $\alpha_n \in$ C $\alpha_n \neq 0$, n $\geq$ 1) has at least one root in C.

Remark 2: In Theorem 6, if $\lambda_1 \in$ C is a root of f(t) = 0, then by theorem 5, f(t) = (t-λ_1)f_1(t). Here deg f1 = n -1. If f1 (t) is not constant, then the equation f1(t) = 0 has a root $\lambda_2 \in$ C, and f1(t) = (t-λ_2)f_2(t). Consequently, f(t) = (t-λ1) (t-λ_2) f_2(t). Here deg f_2 = n-2. Using the fundamental theorem repeatedly, we get
F(t) = αn(t-λ_1) (t-λ_2) …. (t-λn) for some λ_1,λ_2, …..λ_n in C, which are not necessarily distinct. (This process has to stop after n steps since deg f = n.) thus, all the roots of

f(t) belong to C and these are n in number. They may not all be distinct. Suppose $\lambda_1, \lambda_2, \ldots. \lambda_k$ are the distinct roots, and they are repeated m_1, m_2, …., m_k times, respectively. Then $m_1 + m_2 + \ldots, + m_k = n$, and $f(t) = \alpha_n (t- \lambda_1)^{m1} (t - \lambda_2)m2 \ldots (t- \lambda k)^{mk}$.

For example, the polymial equation t3 – it2 + t – i = 0 has no real roots, but it has two distinct complex roots, namely, t(= $\sqrt{-1}$) and - i . And we write $t^3 - it^2 + t - i = (t-i)^2 (t + i)$. Here i is repeated twice and – i only occurs once.

We can similarly show that any polynomial f(t) over r can be written as a product of linear polynomials and quadratic polynomials. For example the real polynomial $t^3 - 1 = (t-1) (t^2 + t +1)$.

Now we go to show the second and final link that relates the minimal and characteristic polynomials of T:V → V, where V is a vector space over F. let p(t) be the minimal polynomial of T. We will show that a scalar λ is a root of p(t). The proof will utilize the following remark.

Remark 3: if λ is an eigenvalue of T, then Tx - λx for some x ∈ V, x ≠ 0. But tx = λx ⇒ $T^2x = T(Tx) = T(\lambda x) = \lambda^2 x$. By induction it is easy to see that $T^k x = \lambda^k x$ for all k. Now, if $g(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots.. + a_1 t + a_0$ is a polynomial over F, then $g(T)= a_n t^n + a_{n-1} T^{n-1} T^{n-1} + \ldots\ldots + a_1 T + a_0 I$.

This means that
$g(T)x = a_n T^n + a_{n-1} T^{n-1} x + \ldots + a_1 Tx + a_0 x$
$= a_n \lambda^n + a_{n-1} \lambda^{n-1} x + \ldots + a_1 \lambda x + a_0 x$
$= g(\lambda) x$
Thus, λ is an eigenvalue of T ⇒ g(λ) is an eigenvalue of g(T).

Now for the theorem.

Teorem 7: Let T be a linear transformation on a finite-dimensional vector V over the field f. Then λ∈ F is an eigenvalue of T if and only if λ is a root of the minimal **polynomial of T have the same roots.**

Proof: Let p be the minimal polynomial of T and let λ∈ F. Suppose λ is an eigenvalue of T. Then Tx = λx for some 0 ≠ x ∈V. Also, by Remark 3, P9T)x = But p(T) = 0. Thus, 0 = p(λ) = 0, that is, λ is a root of p(t).

Conversely, let λ be a root of p(λ) = 0 and , by Theorem 5, p(t) = (t-λ)q(t), deg q < deg p, q ≠ 0. By the property MP3, ∃ v ∈ V such that q(T) v ≠ 0.
Let x = q(T)v ≠ 0. Then,
(T - λI)x = (t - λI) q(T)v = p(T)v =0

⇒ Tx - λx = 0 ⇒ Tx = λx. Hence, λ is an eigenvalue of T.

so, λ is an eigenvalue of T iff λ is a root of the minimal polymial of t.

In Unit 10 we have already observed that λ is an eigenvalue of T if and only if λ is a root of the characteristic polynomial of T. Hence, we have shown that both the minimal and characteristic polynomials of T have the same roots, namely, the eigenvalues of T.

Caution: Though the roots of the characteristic polynomial ad the minimal polynomial coincide, the two polynomials are not the same, in general.

For example, if the characteristic polynomial T: $R^4 \rightarrow R^4$ is $(t+1)^2 (t-2)^2$, then the minimal polynomial could be (t = 1) (t -2) or $(t+1)^2 (t-2)$, or $(t+1)(t-2)^2$, or even $(t+1)^2(t-2)^2$, depending on which of these polynomials is satisfied by T.

In general, let f(t) = (t -λ1)n1 (t -λ2)n2 … (t - λ)nr be the characteristic polynomial of a linear transformation T, where deg f = n ($\therefore$, $n_1 + n_2 + \ldots + n_r = n$) and $\lambda_1, \ldots\lambda_r \in C$ are distinct. Then the minimal polynomial p(t) is given by
$P(t) = (t - \lambda_1)^{m1} (t - \lambda 2)^{m2} (t - \lambda_r)^{mr}$, where $1 \leq m_1 \leq n_i$ for i = 1,2, ….., r.
In case T has n distinct eigenvalues, then
$f(t) = (t-\lambda_1)(t-\lambda_2) \ldots\ldots (t-\lambda_n)$
and therefore,
$p(t) = (t - \lambda_1)(t-\lambda_2) \ldots\ldots (t-\lambda_n) = f(t)$.

E E3 what can the minimal polynomial of T: $R^3 \rightarrow R^3$ be if its characteristic polynomial if
a) t^3, b) t(t -1) (t + 2)?

Analogues to the definition of the minimal polynomial of a linear transformation, we define the minimal polynomial of a matrix.

Definition: The minimal polynomial of a matrix A over f is the monic polynomial p(t) such that

i) p(A) = 0, and
ii) if q(t) is a non-zero polynomial over F such that deg q < deg p, then $q(A) \neq 0$.

We state two theorems which are analogues to Theorem 4 and 7. Their proofs also similar to those of Theorems 4 and 7

Theorem 8: The minimal polynomial and the characteristic polynomial.

Theorem 9: The roots of the minimal polynomial and characteristic polynomial of a matrix are the same, and are the eigenvalues of the matrix.

let us use these theorems now.

Example 3: Obtain the minimal polynomial of A = $\begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$

Solution: The characteristic polynomial of A =

$$f(t) = \begin{vmatrix} t-5 & 6 & 6 \\ 1 & t-4 & -2 \\ -3 & 6 & t+4 \end{vmatrix} = (t-1)(t-2)^2.$$

Therefore, the minimal polynomial p(t) is either (t -1) (t -2) or (t-1) (t-2)2

Since (A – 1) (A – 21)

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

the minimal polynomial of

Example 4: find the minimal polynomial of

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$$

Solution: The characteristic polynomial of A =

$$f(t) = \begin{vmatrix} t-3 & 6 & 6 \\ -2 & t-2 & 1 \\ -2 & -2 & t \end{vmatrix} = (t-1)(t-2)^2.$$

Again, as before, the minimal polynomial p(t) of A is either (t – 1) (t-2) or (t-1) $(t-2)^2$. But, in this case,

$$(A-I)(A-21) = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{pmatrix} \neq 0.$$

Hence, $p(t) \neq (t-1)(t-2)$. Thus, $p(t) = (t-1)(t-2)^2$.

Now, let T be a linear transformation for V to v, and B be a basis of V. Let A = $[T]_B$. If g(t) is any polynomial with coefficients in f, then g(T) = 0 if and only it g(A) = 0. Thus, the minimal polynomial of T is the same as the minimal of A. so, For example, if t : $R^3 \rightarrow R^3$ is a linear operator which is represented with respect to the standard basis, by the matrix in Example 3., then its minimal polynomial is (t-1) (t-2).

Example 5: what can the minimal polynomial of T: $R^4 \rightarrow R^4$ be if the characteristic polynomial of $[T]_B$ is

a) $(t-1)(t^3 + 1)$, b) $(t^2 + 1)^2$?

Here, B is the standard basis of R^4.

Solution: a) now $(t-1)(t^3 + 1) = (t + 1)(t-1)(t^2-t+1)$. This has 4 distinct complex roots, of which only 1 and -1 are real. Since all the roots are distinct this polynomial is also the minimal polynomial of T.

b) $(t2 +1)^2$ has no real roots. It has 2 repeated complex roots, i –i . Now, the minimal polynomial must be a real polynomial that divides the characteristic polynomial. ∴, it can be $(t^2 + 1)$ or $(t^2 + 1)^2$.

This example shows you that if the minimal polynomial is a real polynomial, then it need not be a product of linear polynomial only. Of course, over C it will always be a product of linear polynomials.

Try the following exercises now.

E E4) Find the minimal polynomial of

a) $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$

b) T: $R \rightarrow R^3$: T(x,y,z) = (x + y, y +z, z +x)

The next exercise involves the concept of the trace of a matrix. If A = [a_{ij}] $\in M_n(F)$, then the trace of A, denoted by Tr(A) is – (coefficient of tn-1 in $f_A(t)$).

E5) Let A = [aij] $\in M_n(F)$. For the matrix A given in E4, show that

Tr(A) = (sum of its eigenvalues)

= (sum of its diagonal elements)

We end the unit by recapitulating what we have done in it.

7.4 Summary

In this unit we have covered the following points.

1) The proof of the Cayley-Hamilton theorem, which says that every square matrix (or linear transformation T:V $\to$ V) satisfies its characteristic equation.

2) The use of the Cayley-Hamilton theorem to find the inverse of a matrix.

3) The definition of the minimal polynomial of a matrix.

4) The proof of the fact that the minimal polynomial and the characteristic polynomial of a linear transformation (or matrix) have the same roots. These roots are precisely the eigenvalues of the concerned linear transformation (or matrix).

5) A method for obtaining the minimal polynomial of a linear transformation (or matrix).

7.5 Solutions/Answers

E1) a) $f_A(t) = \begin{vmatrix} t-1 & 0 & 0 \\ -2 & t-3 & 0 \\ 2 & 2 & t-1 \end{vmatrix} = (t-1)^2 (t-2)$

Now, $(A-1)^2 (A-31) = 0$. $\therefore$ A satisfies $f_A(t)$.

b) $f_A(t) = \begin{vmatrix} t & -1 & 0 \\ -3 & t & -1 \\ -1 & 2 & t+1 \end{vmatrix} = t^3 + t^2 - t - 4.$

Now, $A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 1 \\ 1 & 1 & -1 \\ -7 & 3 & -1 \end{pmatrix}$

$\therefore A^3 = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ 1 & 1 & -1 \\ -7 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 2 \\ 8 & -5 & 4 \end{pmatrix}$

Now, $A^3 + A^2 - A - 41 = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 2 \\ 8 & -5 & 4 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ 1 & 1 & -1 \\ -7 & 3 & -1 \end{pmatrix}$

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

c) $f_A(t) = \begin{pmatrix} t-1 & 0 & -1 \\ 0 & t-3 & -1 \\ -3 & -3 & t-1 \end{pmatrix} = t^3 - 8t^2 + 13t$

$\begin{matrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 3 & 3 & 4 \end{matrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 3 & 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 3 & 5 \\ 3 & 12 & 7 \\ 15 & 21 & 22 \end{pmatrix}$

Now, A2 = $\begin{pmatrix} & & \end{pmatrix}$

$\therefore$ A1 = $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 3 & 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 3 & 5 \\ 3 & 12 & 7 \\ 15 & 21 & 22 \end{pmatrix} \begin{pmatrix} 19 & 24 & 27 \\ 24 & 57 & 43 \\ 81 & 168 & 176 \end{pmatrix}$

$$+ \begin{pmatrix} 13 & 0 & 13 \\ 0 & 39 & 13 \\ 39 & 39 & 52 \end{pmatrix} = 0.$$

$\therefore$ A satisfies its characteristic polynomial.

E2) a) The constant term of $f_A(t)$ is $-3 \neq 0$. $\therefore$ A is invertible.
Now, $A^3 - 5A^2 + 7A - 31 = 0$.

$\therefore A^{-1} = \frac{1}{3}(A^2 - 5A + 71)$

$$= \frac{1}{3}\left(\begin{bmatrix} 1 & 0 & 0 \\ 8 & 9 & 0 \\ -8 & -8 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 10 & 15 & 0 \\ -10 & -10 & 5 \end{bmatrix} + \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix} \right)$$

$$= \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & 2 & 3 \end{bmatrix}$$

Pre-multiply by A to check that our calculations are right.

b) A is invertible and $A^{-1} = \frac{1}{4} (A^2 + A-1)$

$$\frac{1}{4} \begin{bmatrix} 2 & 1 & 1 \\ 4 & 0 & 0 \\ -6 & 1 & -3 \end{bmatrix}$$

c) A is not invertible, by Theorem 3.

E 3) a) the minimal polynomial can be t, t^2 or t^3.

c) The minimal polynomial can only be t(t-1) (t +2).

E 4) a) $f_A(t) = \begin{vmatrix} t & -1 & 0 & -1 \\ -1 & t & -1 & 0 \\ 0 & -1 & t & -1 \\ -1 & 0 & -1 & t \end{vmatrix} = t^2(t-2)(t+2)$

$\therefore$, the minimal polynomial can be t(t -2) (t + 2) or $t^2(t-2)(t+2)$.

Now A(A -21) (A + 21) = 0. $\therefore$, t(t -2) 9t + 2) is the minimal polynomial of A

.

b) The matrix of T with respect to the standard basis is

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

Then $fA(t) = \begin{vmatrix} t-1 & -1 & 0 \\ 0 & t-1 & -1 \\ -1 & 0 & t-1 \end{vmatrix} = t^3 - t^2 - t.$

This has 3 distinct roots: $0, \frac{1+i\sqrt{5}}{2}, \frac{1-i\sqrt{5}}{2}$

$\therefore$ the minimal polynomial is the same as $f_A(t)$.

E5) Sum of diagonal elements = 0

Sum of eigenvalues = 0 -2 + 2 = 0 and Tr(A) = - (coeff. of t^3 in $f_A(t)$) = 0.

$\therefore$ Tr(A) = sum of diagonal elements of A.

= sum of eigenvalues of A.