UNIT I

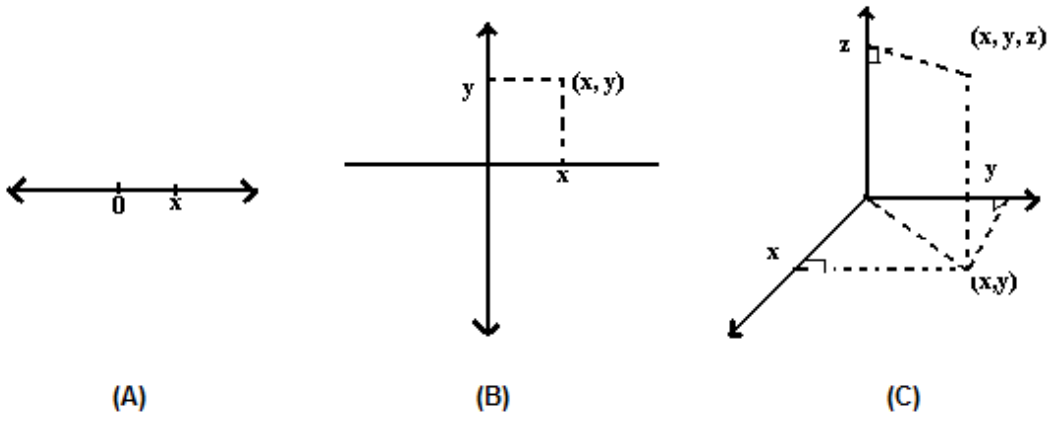
VECTORS

Certain Physical quantities such as mass, area, density, volume, etc., that possess only magnitude are called *scalars*. On the other hand, there are physical quantities such as force, displacement, velocity, acceleration, etc that has both magnitude and direction. Such quantities are called *vectors*.

The concept of a vector is essential for the whole course. It provides the foundation and geometric motivation for everything that follows. Hence the properties of vectors, both algebraic and geometric, will be discussed in this unit.

1.1 Definition of Points in n-space

We know that, once a unit length is selected, a number x can be used to represent a point on a line. A pair of numbers (i.e. a couple of numbers) (x, y) can be used to represent a point in the plane. A triple of numbers (x, y, z) can be used to represent a point in space. The following pictures illustrate these representations:



(A) (B) (C)

We can say that a single number represents a point in 1-space (A), a couple represents a point in 2-space (B) and a triple represents a point in 3-space (C).

Although we cannot draw a picture to go further, we can say that a quadruple of numbers (x, y z, w) or (x_1, x_2, x_3, x_4) represent a point in 4-sapce.

Activity 1.1.1: Define a point in n-space, where n is a positive integer.

The set of all points in n-space is represented by $\Re^n$. For the point X in $\Re^n$, represented by n-tuple of real numbers (x_1, x_2, …, x_n), the numbers x_1, $x_{2;}$…, x_n are called the **coordinates** of X.

Example 1.1.1 The space we live in can be considered as a 3 space. After selecting an origin and a coordinate system, we can describe the position of a point (body, particle, etc.) by 3 coordinates. We can extend this space to a 4 dimensional space, with a fourth coordinate, for example, time. If you select the origin of the time axis as the birth of Christ, how do we describe a body with negative time co-ordinate? What if the birth of the earth is taken as the origin of time?

If A = (a_1, a_2, ., a_n) and B = (b_1, b_2, …, b_n) are points in the same space $\Re^n$, and if c is a real number then

i. A and B are equal (or represent the same point) if $a_1 = b_1$, $a_2 = b_2$, … and $a_n = b_n$.

ii. A + B, A – B and cA are defined to the points whose coordinates are ($a_1 + b_1$, $a_2 + b_2$, …, $a_n + b_n$), (a_1-b_1, a_2-b_2, …,$a_n - b_n$) and (ca_1, ca_2, …, ca_n), respectively.

Example 1.1.2 1) Let A = (1,2), B = (-3,4) , then A+B=(-2,6), A-B = (4,-2), -3A=(-3,-6)

2) Let X = (1, 0, π, 4), Y = (2, 4,-2π,-6), then 2X+Y = (4, 4, 0, 2) and X-(1/2) Y = (0,-2, 2π, 7).

1.2. Vectors in n-space, Geometric interpretation in 2 and 3- spaces

Every pair of distinct points A and B in $\Re^n$ determines a directed line segment with initial point at A and terminal point at B. We call such a directed line segment a **vector** and denote it by $\overrightarrow{AB}$. The length of the line segment is the magnitude of the vector. Although $\overrightarrow{AA}$ has zero length, and strictly speaking, no direction, it is convenient to view it as a vector. It is called a **zero** or a **null vector**. It is often denoted by $\vec{O}$.

Two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ will be considered to be equal (or equivalence), $\overrightarrow{AB} = \overrightarrow{CD}$, if they have the same magnitude and direction.

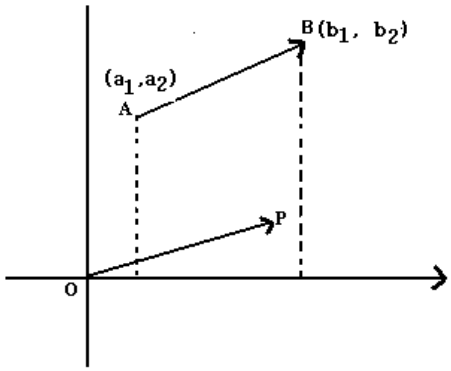
Notice that the definition of equality of two vectors does not require that the vectors have the same initial and terminal points. Rather it suggests that we can move vectors freely provided we make no change in magnitude and direction.

Activity 1.2.1

Let (a_1, a_2) be the coordinate representation of A and let (b_1, b_2) be that of B.Let **P** be the point $(b_1-a_1, b_2 -a_2)$. If O is the origin, is $\overrightarrow{\mathbf{OP}}$ in the direction of $\overrightarrow{AB}$?

Is length of $\overrightarrow{\mathbf{OP}}$ equal to the length of $\overrightarrow{AB}$?

Is $\overrightarrow{\mathbf{OP}} = \overrightarrow{AB}$?



If your answer for the above questions is yes, then we can conclude that any vector V= $\overrightarrow{AB}$ in the plane is a vector $\overrightarrow{\mathbf{OP}}$ with initial point at the origin. This is the only vector whose initial point is the origin and P = B – A, which is equal to $\overrightarrow{AB}$. Moreover, V = $\overrightarrow{\mathbf{OP}}$ is uniquely determined by its terminal point P. If **P** = (x, y), then we shall write **V** = (x, y) and refer to it as the coordinate representation of V relative to the chosen coordinate system. In view of this, we shall call (x, y) either a point or a vector, depending on the interpretation which we have in mind. So if V = $\overrightarrow{AB}$, then we can write V = B – A. In view of this two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal (or equivalence) if B – A = D – C.

Example 1.2.1: If P = (1, 3), Q = (-1,0), R = (0, -1) and S = (-2, -4) then $\overrightarrow{QP} = \overrightarrow{RS}$

As numbers can be added subtracted and multiplies, vectors can be combined in the following ways.

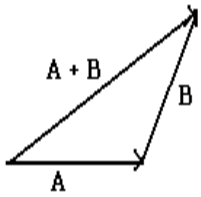
Let A = (a_1, a_2) and B = (b_1, b_2) vectors and t be a real number.

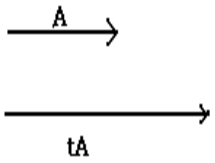
The sum A + B = $(a_1 + b_1, a_2 + b_2)$

The difference A – B = $(a_1 - b_1, a_2 – b_2)$

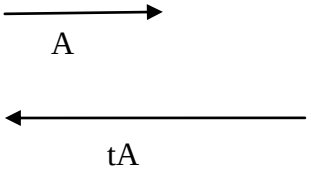
The scalar multiple tA = $(t\, a_1, t\, a_2)$

The geometric interpretation of the above vector operations is that A + B is a vector obtained by placing the initial point of B on the terminal point of A.If t > 0,then tA is a vector in the direction of A.What about if t<0? A and tA are said to have opposite direction.(see figure a and b)





for t > 0 Fig. a



for t < 0 Fig. b

We can extend the above notions to vectors in $\mathfrak{R}^n$ but the geometric interpretations for n > 3 are difficult. Hence we focus on algebraic aspects of vectors.

If A = $(a_1, a_2, \ldots, a_n)$ and B = $(b_1, b_2, \ldots, b_n)$ are vectors in $\mathfrak{R}^n$ and if t is any real number, then

$$\boldsymbol{A+B=(a_1+b_1, a_2+b_2, .., a_n+b_n)}$$

$$tA=(ta_1, ta_2, \ldots, ta_n)$$

A and B are said to be **parallel** vectors iff either A = tB for some real number t or B = tA. In other words, A and B are parallel iff one is a scalar multiple of the other.

Using the arrow notation $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ are parallel if and only if

Q – P = t(S – R) or S – R = t(Q – P) for some real number t.

Activity 1.2.2: 1) Let A = (6, -2, 4). Find two vectors C and D which are parallel to A. Are C and D also parallel to each other?

2) Let P = (3,7), Q = (-4,2), R = (5,1), S = (-16,-14).

Is $\overrightarrow{PQ}$ parallel to $\overrightarrow{RS}$?

Using the above definitions and applying the associative and commutative properties of real numbers, one can prove the following theorem.

Theorem 1.2.1 Let A, B and C be any members of $\mathfrak{R}^n$, and let **m** and **n** be any real numbers.Then

a) $m(nA) = (mn)A$

b) $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$ } Associativity

c) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ Commutativity

d) $(m+n)A=mA+nA$

e) $m(A+B)=mA+mB$ } Distributive property

f) $0.\mathbf{A}=\mathbf{O}$

g) $\mathbf{A}+\mathbf{O}=\mathbf{A}$

Proof c) $\mathbf{A}+\mathbf{B} = (\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_n)+(\mathbf{b}_1,\mathbf{b}_2,...,\mathbf{b}_n)$

$= (\mathbf{a}_1+\mathbf{b}_1,\mathbf{a}_2+\mathbf{b}_2,...,\mathbf{a}_n+\mathbf{b}_n)$

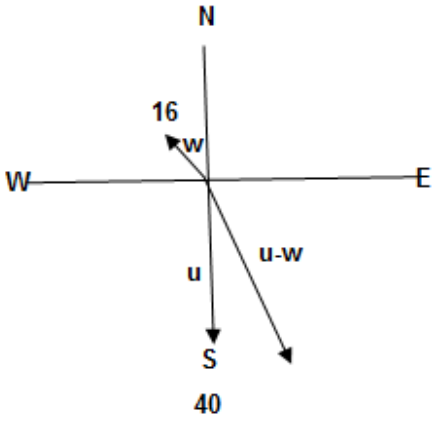
$= (\mathbf{b}_1+\mathbf{a}_1,\mathbf{b}_2+\mathbf{a}_2,...,\mathbf{b}_n+\mathbf{a}_n)$

$= (\mathbf{b}_1,\mathbf{b}_2,...,\mathbf{b}_n) + (\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_n)$

$= \mathbf{B}+\mathbf{A}$

Example 1.2.3

A boat captain wants to travel due south at 40 knots. If the current is moving northwest at 16 knots, in what direction and magnitude should he works the engine?



Solution: we have **u = v + w** where u corresponds to the engine's vector and w corresponds to the velocity of the current. We have **u = -40j** and **w =** $-8\sqrt{2}\,\mathbf{i} + 8\sqrt{2}\,\mathbf{j}$

Hence v = u – w = -40j – ($-8\sqrt{2}$ i + $8\sqrt{2}$ j) = $8\sqrt{2}$ i – (40 + $8\sqrt{2}$)j. The magnitude is $\sqrt{(8\sqrt{2})^2+(40+8\sqrt{2})^2}$ = 52.5.

The direction is $\arctan\left(-\dfrac{40+8\sqrt{2}}{8\sqrt{2}}\right)$ = -1.35 radians

Exercise 1.2.1:

1. Given three vectors A = (1, 1, 1), B = (-1, 2, 3) and C = (0, 3, 4), find
   a. A+B
   b. 2A – B
   c. A+B – C
   d. A – 3B + 10C

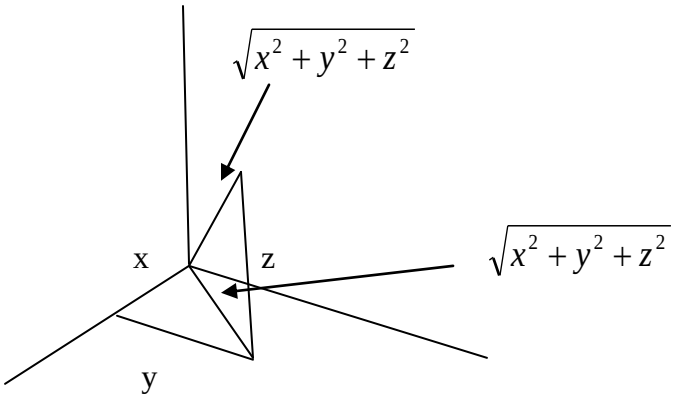
2. Determine whether α and β can be found to satisfy the vector equations

   a. $(2, 1, 0) = \alpha(-2, 0, 2) + \beta(1, 1, 1)$

   b. $(-3, 1, 2) = \alpha(-2, 0, 2) + \beta(1,1,1)$

The Distance Formula

The distance formula is derived from the three dimensional version of the Pythagorean theorem, which is displayed below,



The distance between two points (x_1, y_1, z_1) and (x_2, y_2, z_2) is given by:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

1.3 Scalar product and norm of vector, orthogonal projection, and direction cosines

Let $A = (a_1, a_2,\ldots,a_n)$ and $B = (b_1, b_2, \ldots, b_n)$ be two vectors. The **scalar product** of A and B is the number A.B defined by

$$\mathbf{A.B} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \ldots + \mathbf{a}_\mathbf{n}\mathbf{b}_\mathbf{n}$$

Note: The scalar product is also called a **dot product** or **inner product.**

Example 1.3.1

1) Given A = (3, -1) and B = (2, 3), then A.B = 3

2) $\mathbf{i}.\mathbf{j} = \mathbf{O}$, where i = (1,0,0) and j = (0,1,0).

The scalar product satisfies many of the laws that hold for real numbers. The basic ones are:

a) $\mathbf{A}.\mathbf{B} = \mathbf{B}.\mathbf{A}$

b) $\mathbf{t}(\mathbf{A}.\mathbf{B}) = (\mathbf{tA}).\mathbf{B} = \mathbf{A}.(\mathbf{tB})$

c) $(\mathbf{A}+\mathbf{B}).\mathbf{C} = \mathbf{A}.\mathbf{C} + \mathbf{B}.\mathbf{C}$

d) If $\mathbf{A} = \mathbf{O}$ is the zero vector, then $\mathbf{A}.\mathbf{A} = 0$, and otherwise $\mathbf{A}.\mathbf{A} > 0$

Activity 1.3.1: Proof the above properties.

Example 1.3.2: Given A = (3, 2,-1) and B = (2, 0, 3), and C = (1,-1, 1), then

a. A.B = B.A = 3
b. 2(A.B) = 6
c. (A+B).C = A.C + B.C = 5

Activity 1.3.2: Find A.A, B.B, and C.C. Are all positive values?

Exercise 1.3.1: Given the vectors A = (1,-1,2), B = (-2,0,2), C = (3,2,1). Evaluate

a. A.C b. B.C c. (A + B).C d. A.(2B + 3C) e. (A.B)C

The **length**, or **norm** or **magnitude** of vector A = (a_1, a_2, …, a_n), denoted by **||A||**, can be expressed in terms of the scalar product.By definition

$$\|\mathbf{A}\|^2 = \mathbf{a}_1^2 + \mathbf{a}_2^2 + ... + \mathbf{a}_\mathbf{n}^2$$

and $\mathbf{A}.\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_\mathbf{n}).(\mathbf{a}_1, \mathbf{a}_2 ..., \mathbf{a}_\mathbf{n}) = \mathbf{a}_1^2 + \mathbf{a}_2^2 + ... + \mathbf{a}_\mathbf{n}^2$

Thus, the norm of the vector A is the number:

$$\|\mathbf{A}\| = \sqrt{\mathbf{a}_1^2 + \mathbf{a}_2^2 + ... + \mathbf{a}_\mathbf{n}^2}$$

Example 1.3.3: If A = (-1, 2, 3) then $\|\mathbf{A}\| = \sqrt{1+4+9} = \sqrt{14}$

Note: 1. A **unit vector** is a vector having norm or length 1.

2. $\|A\| = 0 \Leftrightarrow A = 0$

Example 1.3.4: If $A=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$, then $\|A\| = \sqrt{(1/3)+(1/3)+(1/3)} = 1$

Any non-zero vector can be fully represented by providing its magnitude and a unit vector along its direction. Let $\hat{A}$ be a **unit vector in the direction of** A. Then

$$\hat{A}=\frac{A}{\|A\|}$$

Example 1.3.5: Given a vector A = (1, 1, 1). Find a unit vector in the direction of A.
Solution:

$\|A\| = \sqrt{1+1+1} = \sqrt{3}$, then the unit vector in the direction of A is:

$$\hat{A}=\frac{A}{\|A\|} = \frac{(1,1,1)}{\sqrt{3}} = \left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)=\frac{1}{\sqrt{3}}(1,1,1)$$

Activity 1.3.3: 1. Given three vectors A = (1, 1, 1), B = (-1, 2, 3) and C = (0, 3, 4), find the unit vector in the direction of A+B – C.

2. The vectors i (1, 0, 0), j = (0, 1, 0) and k = (0, 0, 1) are unit vectors in the direction of positive x, y and z axis, respectively. Find a unit vector in the direction of A = (-1, 2, 3).

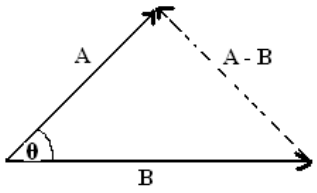
Let A, B be two n-tuples of vectors. We define the distance between A and B to be

$$\|A-B\|=\sqrt{(A-B).(A-B)}$$

Note: Let A be any vector and $x \in \Re$, then $\|A\|=\|-A\|$ and $\|xA\|=|x|\|A\|$

The following theorem gives us a geometric interpretation for the scalar product.

Theorem 1.3.1: Let $\mathbf{A}=(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3)$ and $\mathbf{B}=(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)$ be non-zero vectors and let θ be the angle between $\mathbf{A}$ and $\mathbf{B}$ $(0 \le \theta \le \pi)$.



Then $\mathbf{A}.\mathbf{B} = \|\mathbf{A}\|\,\|\mathbf{B}\|\cos\theta$

Proof: Consider a triangle formed by $\mathbf{A},\mathbf{B}$ and $\mathbf{A}-\mathbf{B}$(Why?)

$\Rightarrow (a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2 = a_1^2+a_2^2+a_3^2+b_1^2+b_2^2+b_3^2-2\|A\|\,\|B\|\cos\theta$

After cancellation, we get, $\mathbf{a}_1\mathbf{b}_1+\mathbf{a}_2\mathbf{b}_2+\mathbf{a}_3\mathbf{b}_3 = \|\mathbf{A}\|\,\|\mathbf{B}\|\cos\theta$

$\Rightarrow \quad \mathbf{A}.\mathbf{B} = \|\mathbf{A}\|\,\|\mathbf{B}\|\cos\theta$

Activity 1.3.4: Given two non-zero vectors A and B, how do you find the angle between them? Take, for example, A = (2, -1, 2), B = (1, -1, 0) and find the angle between them.

Two non-zero vectors are said to be **orthogonal (Perpendicular)** if the angle between them is $\frac{\pi}{2}$.

Note: Two non-zero vectors A and B are said to be orthogonal (Perpendicular) if **A.B= 0**.

Activity 1.3.5: 1. Which of the following pairs of vectors are perpendicular?

a. (1,-1,1), and (2,1,5)
b. (-5,2,7), and (3,-1,2)
c. (1,-1,1), and (2,3,1)
d. (π,2,1), and (2,- π,0)

2. Suppose A.B = A.C, what can you deduce about A, B and C?

Exercise 1.3.2: **1.** Show that $\|A-B\| = \|A+B\|$ if and only if A.B=0.

2. Let $A_1, A_2, \ldots, A_r$ be non-zero vectors such that $A_i.A_j = o$ if $i \neq j$.
Let $c_1, c_2, \ldots, c_r$ be numbers such that $c_1A_1+ c_2A_2+\ldots..+ c_rA_r = 0$.
Show that $c_i = 0$ for all i = 1,2, 3, …, r.

The following are two of the important inequalities

Theorem 1.3.2 Let A and B be vectors. Then

a) $|\mathbf{A}.\mathbf{B}| \le \|\mathbf{A}\| \ \|\mathbf{B}\|$ (Schwarz inequality)

b) $\|\mathbf{A}+\mathbf{B}\| \le \|\mathbf{A}\| + \|\mathbf{B}\|$ (Triangle inequality)

Proof a) If one of A or B is a zero vector then both sides of the inequality are equal to 0. Suppose both A and B are non-zero.

From $\mathbf{A}.\mathbf{B} = \|\mathbf{A}\| \|\mathbf{B}\| \cos\theta$,

$|\boldsymbol{A}.\boldsymbol{B}| = \|\boldsymbol{A}\| \ \|\boldsymbol{B}\| \ |\cos\theta|$ which implies $|\mathbf{A}.\mathbf{B}| \le \|\mathbf{A}\| \ \|\mathbf{B}\|$ (Why?)

b) From the fact that $\|\mathbf{V}\|^2 = \mathbf{V}.\mathbf{V}$,

$$\begin{aligned}
\|\mathbf{A}+\mathbf{B}\|^2 &= (\mathbf{A}+\mathbf{B}) \,.\, (\mathbf{A}+\mathbf{B}) \\
&= \mathbf{A}.\mathbf{A} + 2\mathbf{A}.\mathbf{B} + \mathbf{B}.\mathbf{B} \\
&= \|\mathbf{A}\|^2 + 2\mathbf{A}.\mathbf{B} + \|\mathbf{B}\|^2 \\
&\le \|\mathbf{A}\|^2 + 2|\mathbf{A}.\mathbf{B}| + \|\mathbf{B}\|^2 \qquad (\textbf{why ?}) \\
&\le \|\mathbf{A}\|^2 + 2\|\mathbf{A}\| \|\mathbf{B}\| + \|\mathbf{B}\|^2 \qquad (\textbf{why ?}) \\
&= \left(\|\mathbf{A}\| + \|\mathbf{B}\|\right)^2
\end{aligned}$$

By taking the square roots,

$$\|\mathbf{A}+\mathbf{B}\| \le \|\mathbf{A}\| + \|\mathbf{B}\|$$

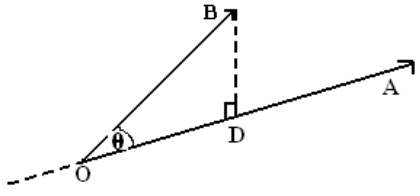
Remark: The inequalities of Theorem 1.3.2 hold true also for any vectors A and B in $\mathbf{R^n}$.

We now define the component of a vector in the direction of another vector.

Let A and B be two non-zero vectors.

Let θ be the angle between them.

From the terminal point of B drop a perpendicular to the



line containing A.

The vector $\overrightarrow{\mathbf{OD}}$ has magnitude $\left\|\overrightarrow{\mathbf{OD}}\right\| = \|\mathbf{B}\|\ |\cos\theta|$, $0 \le \theta \le \pi$, and its direction is either the same as that of A or opposite to it depending on whether θ is acute or obtuse.

Since $\frac{1}{\|\mathrm{A}\|}\mathrm{A}$ is a unit vector in the direction of A and since $\overrightarrow{\mathbf{OD}}$ has magnitude $\|\mathrm{B}\|\ |\cos\theta|$ and is in the direction of A or opposite to A, we can write

$$\overrightarrow{\mathrm{OD}} = \left(\|\mathrm{B}\|\cos\theta\right)\frac{1}{\|\mathrm{A}\|}\mathrm{A}$$

$$\text{or} \quad \overrightarrow{\mathrm{OD}} = \left(\frac{\mathrm{A.B}}{\|\mathrm{A}\|^2}\right)\mathrm{A} \qquad \text{(why?)}$$

This vector is called the **projection of B along A** and is denoted by $\mathbf{Proj}_{\mathbf{A}}\mathbf{B}$.

That is, $\mathrm{Pr}oj_A^B = \left(\frac{A.B}{\|A\|^2}\right)A$

We call $t = \frac{A.B}{\|A\|^2}$ the component of B along A.

Note: 1. $\mathrm{Pr}oj_A^B$ is **parallel** to A. That is $\mathrm{Pr}oj_A^B = tA$, for some $t \in \Re$

2. B - $\mathrm{Pr}oj_A^B$ is **orthogonal (perpendicular) to** A.

Example 1.3.6: Let A = (3, -1, -2) and B = $\left(2, -3, \frac{1}{2}\right)$, then

$$\mathbf{Proj}_{\mathbf{A}}\mathbf{B} = \left(\frac{\mathbf{A.B}}{\|\mathbf{A}\|^2}\right)\mathbf{A} = \left(\frac{6+3-1}{9+1+4}\right)\mathbf{A} = \left(\frac{12}{7}, \frac{-4}{7}, \frac{-8}{7}\right)$$

$$\mathbf{Proj_B A} = \left(\frac{\boldsymbol{A.B}}{\|\boldsymbol{B}\|^2}\right)\boldsymbol{B} = \left(\frac{6+3-1}{4+9+\frac{1}{4}}\right)\boldsymbol{B} = \left(\frac{64}{53}, \frac{-96}{53}, \frac{16}{53}\right)$$

Activity 1.3.4: One application of projections of vector arises in the definition of the work done by a force on a moving body. Find another application.

Given a non-zero vector $u = (u_1, u_2, u_3)$ in $\Re^3$. The **direction cosines** of the vector u are:

$$Cos\alpha = \frac{u_1}{\|u\|},\ Cos\beta = \frac{u_2}{\|u\|},\ Cos\gamma = \frac{u_3}{\|u\|}$$

Where the direction angles, α, β, and γ are the angles that the vector makes with the positive x, y, and z-axes respectively.

Remark: $Cos^2\alpha + Cos^2\beta + Cos^2\gamma = 1$

Example 1.3.7: Let u = (1, -2, 3). Find the direction cosines of u.

Solution: Since $\|u\| = \sqrt{u.u} = \sqrt{(1,-2,3).(1,-2,3)} = \sqrt{1+4+9} = \sqrt{14}$

Thus the direction cosines are: $Cos\alpha = \frac{u_1}{\|u\|} = \frac{1}{\sqrt{14}}$, $Cos\alpha = \frac{u_2}{\|u\|} = \frac{-2}{\sqrt{14}}$, and $Cos\gamma = \frac{u_3}{\|u\|} = \frac{3}{\sqrt{14}}$

Activity 1.3.5: Is $Cos^2\alpha + Cos^2\beta + Cos^2\gamma = 1$?

1.4 The Vector product

The second type of product of two vectors is the cross product. Unlike the dot product, the cross product of two vectors is a vector.

Definition 1.4.1 The cross product (or vector product) A x B of two vectors A = (a_1, a_2, a_3) and B = (b_1, b_2, b_3) is defined by

A x B = ($a_2 b_3 - a_3 b_2$, $a_3 b_1 - a_1 b_3$, $a_1 b_2 - a_2 b_1$)

Note that the cross product is defined in $\Re^3$

Example 1.4.1 Let A = (2, -1, 3) and B = (-1, -2, 4)

A x B = (-4 + 6, -3 – 8, -4 –1) = (2, -11, -5)

Activity 1.4.1: Find B x A.

The following are some of the basic properties of cross product

Theorem 1.4.1 For vectors A, B and C,

1) A x B = -(B x A)
2) A x A = O
3) tA x B = t(A x B) = A x (tB), $t \in \Re$
4) $\|\mathbf{AxB}\|^2 = \|\mathbf{A}\|^2\|\mathbf{B}\|^2 - (\mathbf{A}.\mathbf{B})^2$
5) C. (A x B) = B. (C x A) = A. (B x C)
6) (A + B) x C = (A x C) + (B x C)
7) C x (A + B) = (C x A) + (C x B)
8) A . (A x B) = 0 and B. (A x B) = 0 (that is, AxB is perpendicular to both A and B.)
9) (AxB) x C = (A. C) B – (B.C)A

Proof : The following is the proof for 1, 2 and 8. The rest are left as an exercise

1) From the definition of cross product,

   A x B = $(a_2 b_3 – a_3 b_2,\ a_3 b_1 – a_1 b_3,\ a_1 b_2 – a_2 b_1)$

   For B x A, interchange A and B to obtain

   B x A = $(b_2 a_3 – b_3 a_2,\ b_3 a_1 - b_1 a_3,\ b_1 a_2 - b_2 a_1)$

   = $(a_2 b_3 – a_3 b_2,\ a_3 b_1 - a_1 b_3,\ a_1 b_2 - a_2 b_1)$

   = - (A x B)

2) A x A = $(a_2 a_3 – a_3 a_2,\ a_3 a_1 - a_1 a_3,\ a_1 a_2 - a_2 a_1)$

   = (0, 0, 0)

8) Setting C = A in 5) yields

A . (A x B) = B . (A x A)

= B . 0 (why?)

= 0

By setting C = B in 5),

B .(A x B) = A . (B x B)

= A . 0 = 0

This shows that for non zero vectors A and B, the cross product A x B is orthogonal to both A and B.

Activity 1.4.2: Are the usual commutative and associative laws valid? i.e. for any vectors A, B and C in $\Re^3$, is A x B = B x A?

Is A x (B x C) = (A x B) x C?

Exercise 1.4.1: Let A = (2,1,0), B = (2,-1,1) and C = (0,1,1). Find

a. AxB b. (AxB)xC c. (A.C)B – (B.C)A d. BxC e. Ax(BxC)

From 4) of theorem 1.4.1, we derive an important formula for the norm of the cross product.

$$\|\mathbf{AxB}\|^2 = \|\mathbf{A}\|^2 \|\mathbf{B}\|^2 - (\mathbf{A}.\mathbf{B})^2$$

$$= \|\boldsymbol{A}\|^2 \|\boldsymbol{B}\|^2 - \|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2 \boldsymbol{\cos^2\theta} \quad (\theta \text{ is the angle between A and B})$$

$$= \|\mathbf{A}\|^2 \|\mathbf{B}\|^2(1-\cos^2\theta)$$

$$= \|\mathbf{A}\|^2 \|\mathbf{B}\|^2 \sin^2\theta$$

$$\Rightarrow \|\mathbf{AxB}\| = \|\mathbf{A}\| \|\mathbf{B}\| \sin\theta \quad (\text{For } 0 \le \theta \le \pi, \sin\theta \text{ is non- negative})$$

Activity 1.4.3:

- For the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, find $\mathbf{i} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{k}$ and $\mathbf{k} \times \mathbf{i}$. What is $\mathbf{j} \times \mathbf{i}$?
- If A and B are parallel, what is A ×B?
- If A and B are orthogonal, What is $\|\mathbf{A} \times \mathbf{B}\|$?

Exercise 1.4.1:

1. Find a unit vector perpendicular to both A = (2,-3,1) and B = (1,2,-4).
2. Prove that (A – B)x(A + B) = 2(AxB).

1.5. Application on area and volume

Let **u** and **v** be vectors and consider the parallelogram that the two vectors make. Then

$$\|u \times v\| = \textbf{Area of the Parallelogram}$$

The direction of **uxv** is a right angle to the parallelogram that follows the right hand rule.

To find the volume of the parallelepiped spanned by three vectors **u, v,** and **w**, we find the triple product:

$$|u.(vxw)| = \textbf{Volume}$$

This can be found by computing the determinant of the three vectors:

$$\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = u_1(v_2w_3 - v_3w_2) - v_1(u_2w_3 - u_3w_2) + w_1(u_2v_3 - u_3v_2)$$

Example 1.5.1: 1. Find the area of the parallelogram which is formed by the two vectors u= (1, 3, 2) and v= (-2, 1, 3).

2. Find the volume of the parallelepiped spanned by the vectors

u = (3, -2, -1), v = (1, 3, 2), and w = (-2, 1, 3).

Solution: 1. The area of the parallelogram is given by:

$$\|u \times v\| = \|(1,3,2) \times (-2,1,3)\| = \|(7,-7,7)\| = \sqrt{147}$$

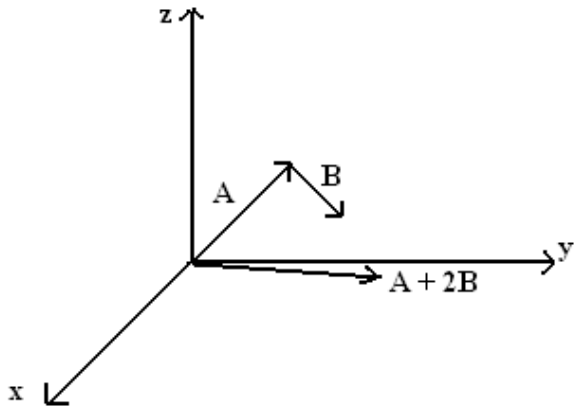
2. The volume of the parallelepiped spanned by the three vectors is:

$$|u.(vxw)| = |(3,-2,-1).((1,3,2) \times (-2,1,3))| = |(3,-2,-1).(7,-7,7)| = 28$$

Exercise: Find the area of the triangle having vertices at u = (3, -2, -1), v = (1, 3, 2), and w = (-2, 1, 3).

1.6 Lines and planes

Activity 1.6.1: Given two vectors A and B with B ≠ 0, consider vectors of the form A + tB, where t varies over all real numbers. The vector A + 2B, for example is as shown in the figure. Show the vectors A + tB for t = 0, 1, 3, $\frac{1}{2}$, -1, -2, $\frac{-3}{2}$, what kind of figure is generated by these vectors? i.e. what is the collection of all points which are terminal points of A + tB?

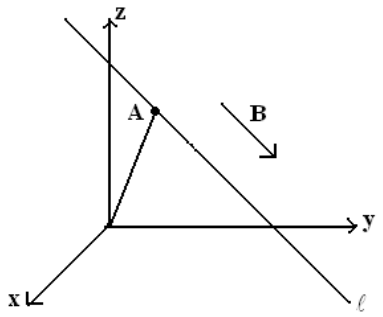


Definition 1.6.1: A **line** ℓ is any set of the form $\{\mathbf{p} : \mathbf{p} = \mathbf{A} + \mathbf{tB},\ \mathbf{t} \in \Re\}$ where B is assumed to be a non-zero vector and A is a fixed point on the line.

Note that if (x, y, z) is on line ℓ and if A = (a_1, a_2, a_3) and B = (b_1, b_2, b_3) then (x, y, z) = (a_1, a_2, a_3) + t(b_1, b_2, b_3) for some real number t.

Activity 1.6.2: Is point A on ℓ? Is B parallel to a vector formed by any two points of ℓ?

P = A + tB is a **vector equation** of a line through A



Example 1.6.1 Find equation of a line through P_1 = (0, 1, 2) and P_2 = (-1, 1, 1).

Solution: We need a point A on the line and a vector B parallel to the vector formed by two point of the line.

Take A = P_1 and B = $P_2 - P_1$. Then

A + t B = (0, 1, 2) + t (-1, 0, -1)

(x, y, z) = (0, 1, 2) + t (-1, 0, -1) is equation of the line. By giving distinct values for t we will obtain distinct points on the line. Find some of the points.

Note: The equation of a line passing through points A and B is given by:

P = A + t(B – A) or P = (1 – t)A + B, $t \in \Re$

Exercise 1.6.1: Let the line L_1 passes through the points (5,1,7) and (6,0,8) and the line L_2 passes through the points (3,1,3) and $(-1,3,\alpha)$. Find the value of α for which the two lines intersect.

Suppose P = (x, y, z) is a point on line ℓ through A = $(a_1, a_2\ a_3)$ in the direction of B = (b_1, b_2, b_3). Then p = A + tB $\Rightarrow$ (x, y, z) = (a_1, a_2, a_3) + t(b_1, b_2, b_3) or equivalently

$$x = a_1 + b_1t$$
$$y = a_2 + b_2t$$
$$z = a_3 + b_3t$$

These equations are **parametric equation** of a line and t is called a **parameter**.

Activity 1.6.3: 1) Find the parametric equation of a line that contains (2, -1, 1) and is parallel to the vector (3, $\frac{1}{2}$, 0).

2) From the parametric equation of a line in $\Re^3$, derive the equation

$$\frac{\mathbf{y}-\mathbf{a}_1}{\mathbf{b}_1} = \frac{\mathbf{y}-\mathbf{a}_2}{\mathbf{b}_2} = \frac{\mathbf{z}-\mathbf{a}_3}{\mathbf{b}_3}$$

It is called **standard form** of equation of a line.

If the line is on a plane show that the standard form reduces to an equation of the form y = mx + c.

Two lines ℓ and m given by $A_1 + tB_1$ and $A_2 + tB_2$ are said to be **parallel** if B_1 and B_2 are parallel. That is the vectors $P_1 - Q_1$ and $P_2 - Q_2$ are parallel for any two points P_1, Q_1 of ℓ and P_2, Q_2 of m.

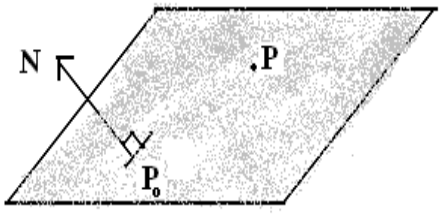
Let ℓ be a line through A in the direction of B ($B \neq 0$). Consider the distance between ℓ and the origin. This distance is the minimum of the lengths of all vectors with initial point the origin and terminal point on ℓ. That is, minimum of $\|A+tB\|$ for any real number t.

Now put $f(t) = \|A+tB\|^2$

This is a quadratic function whose graph opens upward: $f(t) = \|A\|^2 + 2t(A.B) + t^2\|B\|^2$

So it has minimum at: $t = \dfrac{-2A.B}{2\|B\|^2} = \dfrac{-A.B}{\|B\|^2}$

Let P_o be a point and N be a non zero vector. We define the **plane** passing through P_o to perpendicular to N to be the collection of all points P such that the vector $\overrightarrow{\mathbf{p_o p}}$ is perpendicular N.



According to our definition, if **P** is any point on the plane through P_0 and perpendicular to N, then $\mathbf{N} \cdot \overrightarrow{\mathbf{p_o p}} = 0$ or $N \cdot (p - p_0) = 0$

Activity 1.6.4: Starting from the equation $\mathbf{N} \cdot \overrightarrow{\mathbf{p_o p}} = 0$, show that equation of a plane through point $P_o = (x_o, y_o, z_o)$ perpendicular to N = (a, b, c) is

ax + by + cz = d where $\mathbf{d = ax_o + by_o + cz_o}$.

This equation can be written as $N.P = d$. The vector N is said to be **normal** to the plane. Hence a plane is any set of the form **{P: N.P = d}.** Where N is a given non-zero vector and d is a given number.

Example 1.6.2: Find an Equation of the plane that contains point (-2, 4, 5) and that is normal to (7, 0, -6).

Solution: The equation of the plane is given by 7(x+2)+0(y-4)-6(z-5) = 0 or 7x- 6z = -44

Activity 1.6.5: Does this plane intersect the y-axis?

Two planes in 3 spaces are said to be **parallel** if their normal vectors are parallel. They are said to be **perpendicular** if their normal vectors are perpendicular. The **angle** between two planes is defined to be the angle between their normal vectors.

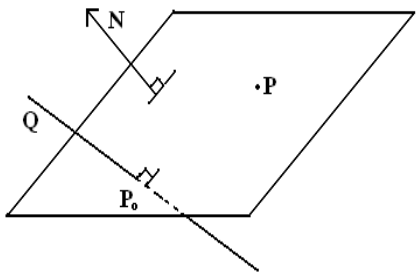
Activity 1.6.6:

1. A plane passes through (-1, 2, 3) and is perpendicular to the y-axis. What is the equation?
2. Consider the planes x + 2y - 3z = 2 and 15x - 9y – z = 2. Are they parallel or perpendicular planes? Or neither parallel nor perpendicular?

Exercise 1.6.2: Find the equation of the plane passing through the three points

$P_1 = (2,1,1)$, $P_2 = (3,-1,1)$, $P_3 = (4,1,-1)$.

Let Q be a point outside a plane normal to N. We define the distance from point Q to the plane as follows. Let P_o be the point of intersection of the line through Q, in the direction of N, and the plane through P. The distance d from Q to the plane is the distance between Q and P_o.



Now we find a formula for this distance. Clearly $d = \left\| \text{Pr}oj_{\overrightarrow{QP}} \overrightarrow{QP} \right\| = \left\| \text{Pr}\mathbf{oj}_{\mathbf{N}} \overrightarrow{\mathbf{QP}} \right\|$

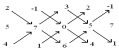
However, $\text{Pr}\mathbf{oj}_{\mathbf{N}} \overrightarrow{\mathbf{QP}} = \left(\frac{\mathbf{N}.\overrightarrow{\mathbf{QP}}}{\|\mathbf{N}\|^2} \right) \mathbf{N}$

Hence $\mathbf{d} = \left\| \left(\frac{\mathbf{N}.\overrightarrow{\mathbf{QP}}}{\|\mathbf{N}\|^2} \right) \mathbf{N} \right\| = \frac{\left| \mathbf{N}.\overrightarrow{\mathbf{QP}} \right|}{\|\mathbf{N}\|^2} \|\mathbf{N}\| = \frac{\left| \mathbf{N}.\overrightarrow{\mathbf{QP}} \right|}{\|\mathbf{N}\|}$

Therefore the **distance d** of a point Q from a plane through **P** which is normal to **N** is given by:

$$\mathbf{d} = \frac{\left| \mathbf{N}.\overrightarrow{\mathbf{QP}} \right|}{\|\mathbf{N}\|}$$

Exercises 1.1

1. Let A = (0, 1, 5) ,B =  . Find the angle between A and B.
2. Which of he following vectors are parallel or perpendicular to (1, 1, -1)?
   a) (2, 2, -2) d) (1, 0, 1)
   b) (2, -2, 0) e) $\left(\frac{-1}{2}, \frac{-1}{2}, \frac{1}{2} \right)$
   c) (-2, 2, 2) f) 
3. a) Find all vectors that are orthogonal to $E_1 = (1, 0, 0)$
   b) Find all vectors that are orthogonal to both E_1 and $E_3 = (0, 0, 1)$
   c) Find all vectors that are orthogonal to E_1, E_2 and $E_3 = (0, 0, 1)$
4. Find a non-zero vector orthogonal to (1, 2, -1)
5. Find a unit vector in the direction of (3, -1, 2, 4)
6. Let $\mathbf{U}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}} \right)$, $\mathbf{U}_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)$, $\mathbf{U}_3 = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0 \right)$

a) Show that each u_1, u_2, u_3 is orthogonal to the other two and that each is a unit vector

b) Find the projection of E_1 on each of u_1, u_2, u_3

c) Find the projection of $A = (a_1, a_2, a_3)$ on u_1.

7. In the following cases compute (A x B).C

   a) A = (1, 2, 0) B = (-3, 1, 0), C = (4, 9, -3)

   b) A = (-3, 1, -2) B = (2, 0, 4), C = (1, 1, 1)

8. Prove that two non-zero vectors A and B are perpendicular if and only if $|\mathbf{A}| \leq |\mathbf{A} + \mathbf{tB}|$ for every number t.

9. If A + B + C = 0. Show that A x B = B x C = C x A

10. Find a formula for the area of a parallelogram whose vertices, in order, are P, Q, R & S.

11. Show that $|\mathbf{A}+\mathbf{B}|\ |\mathbf{A}-\mathbf{B}| \leq |\mathbf{A}|^2 + |\mathbf{B}|^2$

12. Find parametric equations of lines through

    a) (-5, -6, 8) and (1, 3, 7) b. (10, 3, 1) and (6, -2, -3)

13. Find equation of a plane through points (0, 1, 0), (0, -1, -1) and (1, 2, 1)

14. Find a point of intersection of the lines {p:p = (1, -5, 2) + t(-1, 1, 0)} and {p:p = (3, -3, 1) + t(4, 0, -1)}

15. Prove that the line {p:p = (1, 3, -1) + t(0, 3, 5)} lies entirely in the plane {(x, y, z): 2x – 5y + 3z = -16}

16. Find the intersection point of the lines {p:p = (-2, 4, 6) + t(1, 2, 3)} and {p:p = (-2, 1, 6) + t(3, 2, 1)}. Find equation of the plane containing the two lines.

17. Find all points of intersection of the line {p:p = t(1, -3, 6)} and the plane {p:x + 3y + z = 2}

18. Prove that if B.N = 0 and A is on the plane $\{p:(p - p_o).N = 0\}$ then the entire line {p:p = A + tB} lies in the plane.

19. Find a line through the point $p_o = (-5, 2, 1)$ and normal to the plane

{(x, y, z): x = y}

20. Find a line through (x_o, y_o, z_o) and normal to the plane {(x, y, z): ax + by + cz = d}

21. Let ℓ be a line given by A + tB and let P be a point on ℓ different from A. For point Q not on ℓ, show that the distance d of Q form the line ℓ is given by: $$\mathbf{d} = \frac{\left\|\mathbf{B} \times \overrightarrow{\mathbf{PQ}}\right\|}{\|\mathbf{B}\|}$$

22. Let ℓ be the line x = 1 + 2t, y = -1 + 3t, z = -5 + 7t. Find the two points on ℓ at a distance 3 units from the plane 2(x-1) + 2(y+3) –z = 0

23. The set of all points equidistant from (0, 1, 5) and 5, -1, 3) is a plane. Find the equation.

UNIT II

VECTOR SPACES

2.1 The axioms of a vector space

Definition 2.1.1: Let **K** be a set of numbers. We shall say that **K** is **field** if it satisfies the following conditions:

a) If x, y are elements of K then x + y and xy are also elements of K.

b) If x is an element of K, then –x is also an element of K. Furthermore, if $x \neq 0$, then x^{-1} is also an element of K.

c) 0 and 1 are elements of K.

Example 2.1.1: The set of all real numbers $\Re$ and the set of all complex numbers $\mathbb{C}$ are fields.

Activity 2.1.1: Are $\mathbb{Z}$ (The set of all integers) and Q (the set of all rational numbers fields?

Remark: The essential thing about a field is that its elements can be added and multiplied and the results are also elements of the field. Moreover, every element can be divided by a non-zero element.

Definition 2.1.2: A vector space V over a field K is a set of objects which can be added and can be multiplied by elements of K. It satisfies the following properties.

V_1) For any $u, v \in V$ and $a \in K$, we have

$$u + v \in V \quad \text{and} \quad a\,u \in V$$

V_2) For any $u, v, w \in V$,

$$(u + v) + w = u + (v + w)$$

V_3) There is an element of V, denoted by O (called the zero element), such that

$0 + u = u + 0 = u$ for all elements u of V.

V_4) For $u \in V$, there exists $-u \in V$ such that

$u + (-u) = 0$

V_5) For $u, v \in V$, we have

$u + v = v + u$

V_6) For $u, v \in V$ and $a \in k$,

$a(u + v) = au + av$

V_7) For $u \in V$ and $a, b \in k$, $(a + b) u = au + bu$ and $(ab) u = a(bu)$

V_8) For $u \in v$,

$1u = u$

Activity 2.1.2: What is the name given for each of the above properties?

Other properties of a vector space can be deduced from the above eight properties. For example, the property $0u = O$ can be proved as :

$0u + u = 0u + 1.u$ (by V_8)

$= (0 + 1) u$ (by V_7)

$= 1. u$

$= u$

By adding $-u$ to both sides of $ou + u = u$, we have $0u = O$

Examples of different models of a vector space

1) Consider sets $\Re^2$ and $\Re$

   Clearly, for $V = \Re^2$ and $k = \Re$, properties V_1 and V_2 hold.

   The element 0 of $\Re^2$ that satisfied v_3 is $0 = (0, 0)$.

   The other 5 properties can be easily verified. Hence $\Re^2$ is a vector space over$\Re$.

2) Let $V = \Re^2$ and $K = \mathbb{C}$

   For any $u, v \in \Re^2$, we have $u + v \in \Re^2$.

   But for $a \in \mathbb{C}$, au is not always in $\Re^2$.

For example for a = 3i and u = (1, -2), au = (3i, -2i) $\notin \Re^2$.

Hence $\Re^2$ is not vector space over $\mathbb{C}$.

Thus when dealing with vector spaces, we shall always specify the field over which we take the vector space.

3) Let F be the set of all functions form $\Re$ to $\Re$, for any f and g in F, f + g is a function from $\Re$ to $\Re$ defined by (f + g) (x) = f(x) + g(x).

For $a \in \Re$, af = af(x) is in F.

The zero element O of F is the zero function f(x) = 0 for all $x \in \Re$.

By verifying the other properties, we can see that F is a vector space over $\Re$.

The algebraic properties of elements of an arbitrary vector space are very similar to those of elements of $\Re^2$, $\Re^3$, or $\Re^n$. Consequently, we call elements of a vector space as **vectors**

Activity 2.2.1: Which of the following are vector spaces?

a) C on $\Re^2$
b) C^n over $\mathbb{C}$
c) Q^n over Q
d) $\Re^n$ over $\mathbb{C}$

2.3 Subspaces, Linear Combinations and generators

Definition2.3.1: Suppose V is a vector space over k and W is a subset of V. If, under the addition and scalar multiplication that is defined on V, W is also a vector space then we call W a **subspace** of V.

Using this definition and the axioms of a vector space, we can easily prove the following: A subset W of a vector space V is called a **subspace** of V if:

i) **W** is closed under addition. That is, if u, w ∈ W, then u + w ∈ W
ii) **W** is closed under scalar multiplication. That is, if u∈W and a ∈ k, then au∈W.
iii) **W** contains the additive identity 0.

Then as W $\subseteq$ V, properties $V_1 - V_8$ are satisfied for the elements of W.

Hence W itself is a vector space over k. We call W a **subspace** of V.

Example 2.3.1: Consider H = {(x, y): x, y $\in \Re$ and x + 4y = 0} .

H is a subset of the vector space $\Re^2$ over $\Re$. To show that H is a subspace of V, it is enough to show the above three properties hold in H.

Let u = (x_1, y_1) and w = (x_2, y_2) be in H. Then $x_1 + 4y_1 = 0$ and $x_2 + 4y_2 = 0$

u + w = $(x_1 + x_2, y_1 + y_2)$ and $(x_1 + x_2) + 4(y_1 + y_2) = x_1 + 4y + x_2 + 4y_2 = 0 + 0 = 0$

Which shows $\boldsymbol{u + w \in H}$.

For $\boldsymbol{a \in \Re}$, $\boldsymbol{au = (ax_1, ax_2)}$ and $\boldsymbol{(ax_1) + 4(ax_2) = a(x_1 + 4x_2) = 0a} = 0$.

Hence au $\in$ H. Now, the element O of $\Re^2$ is (0, 0). 0 + 4(0) = 0. Hence O = (0, 0) is in H.

$\therefore$ H is a subspace of $\Re^2$

Activity 2.3.1: Take any vector A in $\Re^3$. Let W be the set of all vectors B in $\Re^3$ where B.A = 0. Discuss whether W is a subspace of $\Re^3$ or not.

Definition 2.3.2: Let $v_1, v_2, \ldots, v_n$ be elements of a vector space V over k. Let $x_1, x_2, \ldots, x_n$ be elements of k. Then an expression of the form $x_1v_1 + x_2v_2 + \ldots + x_n v_n$ is called a **linear combination** of $v_1, v_2, \ldots, v_n$..

Example 2.3.2: The sum 2(3, 1) + 4(-1, 2) +(1, 0) is a linear combination of (3, 1), (-1, 2) and (1, 0). As this sum is equal to (3, 10), we say that (3, 10) is a linear combination of the three ordered pairs.

Activity 2.3.2:

i) Take two elements v_1 and v_2 of $\Re^3$. Let **W** be the set of all linear combinations of v_1 & v_2. Show that W is a subspace of $\Re^3$. W is called the subspace **generated** by v_1 and v_2

ii) Generalize i) by showing that a set w generated by elements $v_1, v_2, \ldots, v_n$ of a vector space **V** is a subspace.

2.4. Linear dependence and independence of vectors

Definition 2.4.1: Let **V** be a vector space over k. Elements $v_1, v_2, \ldots, v_n$ of **V** are said to be **linearly independent** if and only if the following condition is satisfied:
whenever $a_1, a_2, \ldots, a_n$ are in k such that $a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0$, then $a_i = 0$ for all $i = 1, 2, \ldots, n$.
If the above condition does not hold, the vectors are called **linearly dependent**. In other words $v_1, v_2, \ldots, v_n$ are linearly dependent if and only if there are numbers $a_1, a_2, \ldots, a_n$ where $a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0$ for at least one non-zero a_i.

Example 2.4.1: Consider $v_1 = (1, -1,1)$, $v_2 = (2, 0, -1)$ and $v_3 = (2, -2, 2)$

i) $a_1v_1 + a_2v_2 = a_1 (1, -1, 1) + a_2 (2, 0, -1) = (a_1 + 2a_2, -a_1, a_1 – a_2)$

$a_1v_1 + a_2v_2 = 0 \Rightarrow a_1 + 2a_2 = 0, -a_1 = 0$ and $a_1 – a_2 = 0$

$\Rightarrow a_1 = 0$ and $a_2 = 0$

Hence v_1 & v_2 are linearly independent.

ii) $a_1v_1 + a_2v_3 = a_1 (1, -1, 1) + a_2 (2, -2, 2)$

$= (a_1 + 2a_2, -a_1 – 2a_2 , a_1 +2 a_2)$

$a_1v_1 + a_2v_3 = 0 \Rightarrow a_1 + 2a_2 = 0, -a_1 – 2a_2 = 0$ and $a_1 +2 a_2 = 0$

$\Rightarrow a_1 = -2a_2$

Take $a_1 = 2$ and $a_2 = -1$, we get $2(1, -1, 1) + (-1) (2, -2, 2) = 0$.

As the constants are not all equal to zero, v_1 and v_3 are linearly dependent.

Activity 2.4.1: Show that v_1, v_2 and v_3 are also linearly dependent.

Remark: If vectors are linearly dependent, at least one of them can be written as a linear combination of the others.

Activity 2.4.2: Show that (1, 0, 0, …,0), (0, 1,0,…)…, (0,0,0, …, 1) are linearly independent vectors in $\Re^n$.

2.5 Bases and dimension of a vector space

Definition 2.5.1: If elements $e_1, e_2, \ldots, e_n$ of a vector space **V** are linearly independent and generate **V**, then the set $B = \{e_1, e_2, \ldots, e_n\}$ is called a **basis** of **V**. we shall also say that the elements $e_1, e_2, \ldots, e_n$ **constitute** or **form** a basis of **V**.

Example 2.5.1:

1) Show that $e_1 = (0, -1)$ and $e_2 = (2, 1)$ form a basis of $\Re^2$.

Solution: we have to show that

i) e_1 and e_2 are linearly independent

ii) They generate $\Re^2$ i.e every element (x,y) of $\Re^2$ can be written as a linear combination of $e_1$and e_2.

i) $a_1 e_1 + a_2 e_2 = O \Rightarrow a_1(0, -1) + a_2(2,1) = (0, 0)$

$\Rightarrow 2a_2 = 0$ and $-a_1 + a_2 = 0$

$\Rightarrow a_2 = 0$ and $a_1 = 0$

$\therefore e_1$ and e_2 are linearly independent

ii) $(x, y) = a_1e_2 + a_2 e_2 \Rightarrow (x, y) = (0, -a_1) + (2a_2, a_2)$

$\Rightarrow x = 2a_2$ and $y = -a_1 + a_2$

$\Rightarrow a_2 = \frac{x}{2}$ and $a_1 = a_2 - y$ … (*)

$= \frac{x-2y}{2}$

Therefore, given any (x, y), we can find a_1 and a_2 given by (*) and (x, y) can be written as a linear combination of e_1 and e_2 as

$$(x, y) = \left(\frac{x-2y}{2}\right)(0,-1) + \left(\frac{x}{2}\right)(2,1)$$

For example, (4, 3) $= \left(\frac{4-6}{2}\right)(0,-1) + \left(\frac{4}{2}\right)(2,1)$

Or (4, 3) = -(0, -1) + 2(2, 1)

Note that {(1, 0), (0, 1)} is also a basis of $\Re^2$. Hence a vector space can have two or more basis. Find other bases of $\Re^2$.

2) Show that $e_1 = (2, 1, 0)$ and $e_2 = (1, 1, 0)$ form a basis of $\Re^3$.

Solution: $e_1 = (2, 1, 0)$ and $e_2 = (1, 1, 0)$ are linearly independent but they do not generate $\Re^3$. There are no numbers a_1 and a_2 for which

$$(3, 4, 2) = a_1 (2, 1, 0) + a_2 (1, 1, 0).$$

Hence {(2, 1, 0), (1,1,0)} is not a basis of $\Re^3$.

The vectors $E_1 = (1, 0, 0)$, $E_2 = (0, 1, 0)$, $E_3 = (0, 0, 1)$ are linearly independent and every element (x, y, z) of $\Re^3$ can be written as

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)$$
$$= xE_1 + yE_2 + zE_3$$

Hence $\{E_1, E_2, E_3\}$ is a basis of $\Re^3$.

Note that the set of elements $E_1 = (1, 0, 0,\ldots,0)$, $E_2 = (0, 1, 0, \ldots 0),\ldots,E_n = (0, 0, 0, \ldots,1)$ is a basis of $\Re^n$. It is called a **standard basis**.

Let $B = \{e_1, e_2, \ldots, e_n\}$ be a basis of **V**. since B generates **V**, any u in **V** can be represented as $u = a_1e_1 + a_2 e_2 + \ldots + a_n e_n$. Since the e_i are linearly independent, such a representation is unique. We call $(a_1, a_2, \ldots, a_n)$ the **coordinate vector** of u with respect to the basis B, and we call a_i the **i – th coordinate.**

Example 2.5.2

1) In 1) of example 3.3.1 The coordinate vector of (4,3) with respect to the basis {(0, -1), (2,1)} is (-1, 2). But with respect to the standard basis it is (4, 3).
Find coordinates of (4,3) in some other basis of $\Re^2$.

2) Consider the set **V** of all polynomial functions f: $\Re \to \Re$ which are of degree less than or equal to 2.

Every element of **V** has the form $f(x) = bx^2 + cx + d$, where $b, c, d \in \Re$

V is a vector space over $\Re$ (show).

Clearly, $e_1 = x^2$, $e_2 = x$ and $e_3 = 1$ are in **V** and $a_1e_1 + a_2 e_2 + a_3e_3 = O$

(0 is the zero function)

$\Rightarrow a_1x^2 + a_2e_2 + a_3 e_3 = 0$ for all x

$\Rightarrow a_1 = a_2 = a_3 = 0.$

Which shows e_1, e_2 and e_3 are linearly independent

$bx^2 + cx + d = a_1e_1 + a_2e_2 + a_3e_3$ for all x

$\Rightarrow bx^2 + cx + d = a_1x^2 + a_2x + a_3$

$\Rightarrow b = a_1$, $c = a_2$ and $d = a_3$

Thus e_1, e_2 and e_3 generate **V**.

$\therefore \{x^2, x_1\ 1\}$ is a basis of **V** and the coordinate vector of an element

$f(x) = bx^2 + cx + d$ is (b, c, d)

The coordinate vector of $x^2 - 3x + 5$ is (1, -3, 5)

Activity 2.5.1: Show that the polynomials

$E_1 = (x - 1)^2 = x^2 - 2x + 1$

$E_2 = x - 1$

and $E_3 = 1$

form a basis of a vector space **V** defined in **2)** of **example 2.5.2**. What is the coordinate of $\mathbf{f(x) = 2x^2 - 5x + 6}$ with respect to the basis $\mathbf{\{E_1, E_2, E_3\}}$?

E = {(1, 0, 0), (0,1,0), (0,0,1)} and B = {(-1,1,0), (-2, 0, 2), (1, 1, 1)} are bases of $\Re^3$ and each has three elements. Can you find a basis of $\Re^3$ having two elements? four elements? The main result of this section is that any two bases of a vector space have the same number of elements. To prove this, we use the following theorem.

Theorem 2.5.1: Let V be a vector space over the field K. Let $\{v_1, v_2,\ldots,v_n\}$ be a basis of V. If $w_1, w_2,\ldots,w_m$ are elements of V, where m > n, then $w_1, w_2, \ldots, w_m$ are linearly dependent.

Proof (reading assignment)

Theorem 2.5.2: Let V be a vector space and suppose that one basis B has n elements, and another basis W has m elements. Them m = n.

Proof: As B is a basis, m > n is impossible. Otherwise by theorem 3.4.1, W will be a linearly dependent set. Which contradicts the fact that W is a basis. Similarly, as W is a basis, n > m is also impossible. Hence n = m.

Definition 2.5.2: Let V be a vector space having a basis consisting of n elements. We shall say that n is the **dimension** of V. It is denoted by **dim V**.

Remarks : 1. If V = {0}, then V doesn't have a basis, and we shall say that dim v is zero.

2. The zero vector space or a vector space which has a basis consisting of a finite number of elements, is called **finite dimensional**. Other vector spaces are called **infinite dimensional**.

Example 2.5.3:

1) $\Re^3$ over $\Re$ has dimension 3. In general $\Re^n$ over $\Re$ has dimension n.
2) $\Re$ over $\Re$ has dimension 1. In fact, {1} is a basis of $\Re$, because
   $a.1 = 0 \Rightarrow a = 0$ and
   any number $x \in \Re$ has a unique expression $x = x.1$.

Definition 2.5.3: The set of elements $\{v_1, v_2, \ldots,v_n\}$of a vector space V is said to be a **maximal set of linearly independent elements** if $v_1, v_2, \ldots,v_n$ are linearly independent and if given any element w of V, the elements $w,v_1, v_2, \ldots, v_n$ are linearly dependent.

Example 2.5.4: In $\Re^3$ {(1, 0, 0), (0, 1, 1), (0, 2, 1)} is a maximal set of linearly independent elements.

We now give criteria which allow us to tell when elements of a vector space constitute a basis.

Theorem 2.5.3: Let **V** be a vector space and $\{v_1, v_2, \ldots, v_n\}$be a maximal set of linearly independent elements of **V**. Then $\{v_1, v_2, \ldots, v_n\}$is a basis of **V**.

Proof: It suffices to show that $v_1, v_2, \ldots, v_n$ generate **V**. (Why?)

Let $\boldsymbol{w} \in \boldsymbol{v}$.

Then w, $v_1, v_2, \ldots, v_n$ are linearly dependent (why?).

Hence there exist numbers $a_o, a_1, a_2, \ldots, a_n$ not all 0 such that

$a_o w + a_1v_1 + a_2v_2 + \ldots + a_nv_n = O$

In particular $\boldsymbol{a_o} \neq 0$ (why?

Therefore, by solving for w,

$$w = \frac{-a_1}{a_o}V_1 - \frac{a_2}{a_o}V_2 \ldots \frac{-a_n}{a_o}V_n$$

This proves that w is a linear combination of $v_1, v_2, \ldots, v_n$.

Theorem 2.5.4: Let dim **V** = n, and let $v_1, v_2, \ldots, v_n$ be linearly independent elements of v. Then

$\{v_1, v_2, \ldots, v_n\}$ is a basis of v.

Proof: According to theorem 3.4.1, $\{v_1, v_2, \ldots, v_n\}$ is a maximum set of linearly independent elements of **V**.

Hence it is a basis by theorem 2.5.3

Corollary 2.5.1: Let W be a subspace of V. If dim W = dimV, then V = W

Proof: Exercise

2.6. Direct sum and direct product of subspaces

Let **V** be a vector space over the field K. Let **U, W** be subspaces of **V**. We define the **sum** of **U** and **W** to be the subset of **V** consisting of all sums **u + w** with $u \in U$ and

$w \in W$. We denote this sum by **U +W** and it is a subspace of **V**. Indeed, if $u_1, u_2 \in U$ and $w_1, w_2 \in W$ then

$$(u_1 + w_1) + (u_2 + w_2) = u_1 + u_2 + w_1 + w_2 \in U + W$$

If $c \in K$, then

$$c(u_1 + w_1) = cu_1 + cw_1 \in U + W$$

Finally, $0 + 0 \in U + W$. This proves that U + W is a subspace.

Definition 2.6.1: A vector space V is a **direct sum** of U and W if for every element v in V there exist unique elements $u \in U$ and $w \in W$ such that $v = u + w$.

Theorem 2.6.1: Let V be a vector space over the field K, and let U, W be subspaces. If U + W = V, and if $U \cap W = \{0\}$, then V is the direct sum of U and W.

Proof: Exercise

Note: When V is the direct sum of subspaces U, W we write:

$$V = U \oplus W$$

Theorem 2.6.2: Let V be a finite dimensional vector space over the field K. Let W be a subspace. Then there exists a subspace U such that V is the direct sum of W and U.

Proof: Exercise

Theorem 2.6.3: If **V** is a finite dimensional vector space over the field K, and is the direct sum of subspaces U, W then

dim V = dim U + dim W

Proof: Exercise

Remark: We can also define V as a direct sum of more than two subspaces. Let W_1, W_2, …., W_r be subspaces of V. We shall say that V is their **direct sum** if every element of can be expressed in a unique way as a sum

$$v = w_1 + w_2 + ... + w_r$$

With w_i in W_i.

Suppose now that U, W are arbitrarily vector spaces over the field K(i.e. not necessarily subspaces of some vector space). We let UXW be the set of all pairs (u, w) whose first component is an element u of U and whose second component is an element w of W. We define the addition of such pairs component wise, namely, if $(u_1, w_1) \in UXW$ and $(u_2, w_2) \in UXW$ we define

$$(u_1, w_1) + (u_2, w_2) = u_1 + u_2, w_1 + w_2)$$

If $c \in K$, we define the product $c(u_1, w_1)$ by

$$c(u_1, w_1) = (cu_1, cw_1)$$

It is then immediately verified that UXW is a vector space, called the **direct product** of U and W.

Note: If n is a positive integer, written as a sum of two positive integers, $n = r + s$, then we see that K^n is the direct product $K^r XK^s$ and $\dim(UXW) = \dim U + \dim W$.

Example 2.6.1: Let, $V = R^3$, $U = \{(0,0, x_3), x_3 \in \Re\}$, and

$W = \{(x_1, x_2, 0), x_1, x_2 \in \Re\}$. Show that V is the direct sum of W and U.

Solution: Since V, U and W are vector spaces, and in addition to that U and W are subspaces of V.

The sum of U and W is:

U + W = $\{(x_1, x_2, x_3), x_1, x_2, x_3 \in \Re\} = R^3 = V$

Thus; V = U + W

The intersection of U and W is: $U \cap W = \{0\}$

Therefore, V is the direct sum of W and U.

Activity 2.6.1: 1. Let, $V = R^3, U = \{(x_1, 0, x_3), x_1, x_3 \in \Re\}$, and

$W = \{(0, x_2, 0), x_2 \in \Re\}$. Show that V is the direct sum of W and U.

2. Let, $V = R^3, U = \{(x_1, x_2, 0), x_1, x_2 \in \Re\}$, and

$W = \{(0,0, x_3), x_3 \in \Re\}$. Show that V is the direct sum of W and U.

Exercise:

1. Let, $V = R^2, U = \left\{(\frac{a+b}{2}, \frac{a+b}{2}), a, b \in \Re\right\}$, and $W = \left\{(\frac{a-b}{2}, \frac{b-a}{2}), a, b \in \Re\right\}$.

Show whether V is the direct sum of W and U or not.

2. Let, $V = R^3, U = \{(x_1, x_2, 0), x_1, x_2 \in \Re\}$, and $W = \{(0, x_2, x_3),\ x_2, x_3 \in \Re\}$

Show whether V is the direct sum of W and U or not.

Exercise 2.1

1. Let k be the set of all numbers which can be written in the form $\mathbf{a} + \mathbf{b}\sqrt{2}$, where a, b are rational numbers. Show that k is a field.
2. Show that the following sets form subspaces
   a. The set of all (x, y) in $\Re^2$ such that x = y
   b. The set of all (x, y) in $\Re^2$ such that x – y = 0
   c. The set of all (x, y, z) in $\Re^3$ such that x + y = 3z
   d. The set of all (x, y, z) in $\Re^3$ such that x = y and z = 2y
3. If U and W are subspaces of a vector space V, show that $\mathbf{U} \cap \mathbf{W}$ and $\mathbf{U} + \mathbf{W}$ are subspaces.
4. Decide whether the following vectors are linearly independent or not (on $\Re$)
   a) (π, 0) and (0, 1)
   b) (-1, 1, 0) and (0, 1, 2)
   c) (0, 1, 1), (0, 2, 1), and (1, 5, 3)
5. Find the coordinates of X with respect to the vectors A, B and C
   a. X = (1, 0, 0), A = (1, 1, 1), B = (-1, 1, 0), C = (1, 0, -1)
   b. X = (1, 1, 1) , A = (0, 1, -1), B = (1, 1, 0), C = (1, 0, 2)
6. Prove: The vectors (a, b) and (c, d) in the plane are linearly dependent if and only if ad – bc = 0

7. Find a basis and the dimension of the subspace of $\Re^4$ generated by
   {(1, -4, -2, 1), (1, -3, -1, 2), (3, -8, -2, 7)}.
8. Let W be the space generated by the polynomials $x^3 + 3x^2 – x + 4$, and
   $2x^3 + x^2 – 7x – 7$. Find a basis and the dimension of W.
9. Let $V = \{(a, b, c, d) \in \Re^4: b – 2c + d = 0\}$
   $W = \{(a, b, c, d) \in \Re^4: a = d, \ b = 2c\}$
   Find a basis and dimension of
   a) V b) W c) $\mathbf{V \cap W}$
10. What is the dimension of the space of 2 x 2 matrices? Give a basis for this space. Answer the same question for the space of n x m matrices.
11. Find the dimensions of the following
    a) The space of n x n matrices all of whose elements are 0 except possibly the diagonal elements.
    b) The space of n x n upper triangular matrices
    c) The space of n x n symmetric matrices
    d) The space of n x n diagonal matrices
12. Let V be a subspace of $\Re^3$. What are the possible dimensions for V? Show that if $V \neq \Re^3$, then either V = {0}, or V is a straight line passing through the origin, or V is a plane passing through the origin.

UNIT III

Matrices

The concept of matrices has had its origin in various types of linear problems, the most important of which concerns the nature of solutions of any given system of linear equations. Matrices are also useful in organizing and manipulating large amounts of data. Today, the subject of matrices is one of the most important and powerful tools in Mathematics which has found applications to a very large number of disciplines such as Engineering, Business and Economics, statistics etc.

3.1 Definition of a matrix

Definition 3.1.1: A rectangular arrangement of mn numbers (real or complex) in to m horizontal rows and n vertical columns enclosed by a pair of brackets [ ], such as

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m1} & \dots & a_{mn} \end{bmatrix}$$

is called an $\boldsymbol{m \times n}$ (read "m by n") **matrix** or a matrix of **order** $\boldsymbol{m \times n}$.

Parentheses ( ) are also commonly used to enclose numbers constituting matrices.

Before we go any further, we need to familiarize ourselves with some terms that are associated with matrices. The numbers in a matrix are called the **entries** or the **elements** of the matrix. For the entry $\boldsymbol{a_{ij}}$, the first subscript $\boldsymbol{i}$ specify the row and the second subscript $\boldsymbol{j}$ the column in which the entry appears. That is, $\boldsymbol{a_{ij}}$ is an element of matrix A which is located in the i^{th} row and j^{th} column of the matrix A. Whenever we talk about a matrix, we need to know the **order** of the matrix.

The **order** of a matrix is the number of rows and columns it has. When we say a matrix is a 3 by 4 matrix, we are saying that it has 3 rows and 4 columns. The rows are always mentioned first and the columns second. This means that a 3×4 matrix does not have the same order as a 4×3 matrix. It must be noted that even though an $m \times n$ matrix contains

mn elements, the entire matrix should be considered as a single entity. In keeping with this point of view, matrices are denoted by single capital letters such as **A, B, C** and so on.

Remark: By the size of a matrix or the dimension of a matrix we mean the order of the matrix.

Example 3.1.1: Let $\boldsymbol{A} = \begin{bmatrix} 1 & 5 & 2 \\ 0 & 3 & 6 \end{bmatrix}$.

Solution: Since A has 2 rows and 3 columns, we say A has order 2×3, where the number of rows is specified first. The element 6 is in the position a_{23} (read *a* two three) because it is in row 2 and column 3.

Example 3.1.2: What is the value of $\boldsymbol{a_{23}}$ and $\boldsymbol{a_{32}}$ in $\boldsymbol{A} = \begin{bmatrix} -1 & 4 & 7 \\ 2 & 3 & 1 \\ 5 & 7 & 8 \end{bmatrix}$ ***?***

Solution: $\boldsymbol{a_{23}}$, the element in the second row and third column, is 1 and $\boldsymbol{a_{32}}$, the element in the third row and second column, is 7. What is the size of this matrix?

Activity 3.1.1: 1. Suppose A is a 5x7 matrix, then

a. A has 7 rows. (True/False)

b. a_{ij} is an element of A for i = 6 and j = 4.(True/False)

c. For what values of i and j, a_{ij} is an element of A?

2. Suppose $A = \begin{bmatrix} 4 & -7 & 5 \\ 8 & 1 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 8 \\ -7 & 1 \\ 5 & 6 \end{bmatrix}$

a. What is the order of A and B?

b. Find a_{22}, a_{13}, b_{13} and b_{31}.

It is customary to abbreviate the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

by the symbol $(a_{ij})_{m\times n}$ or more simply (a_{ij}). This notation merely indicates what type of symbols we are using to denote the general entry.

Example 3.1.3: Form a 4 by 5 matrix, *B*, such that $b_{ij} = i + j$.

Solution: Since the number of rows is specified first, this matrix has four rows and five columns.

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \end{bmatrix}.$$

Activity 3.1.2: Form a 4 by 3 matrix, *B*, such that

a) $b_{ij} = i \times j$ b) $b_{ij} = (-1)^{i+j}$

Definition 3.1.2: Two matrices A and B are said to be **equal**, written A = B, if they are of the same order and if all corresponding entries are equal.

For example, $\begin{bmatrix} 5 & 1 & 0 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 2+3 & 1 & 0 \\ 2 & 3 & 2\times 2 \end{bmatrix}$ but $\begin{bmatrix} 9 & 2 \end{bmatrix} \neq \begin{bmatrix} 9 \\ 2 \end{bmatrix}$. Why?

Example 3.1.4: Given the matrix equation $\begin{bmatrix} x+y & 6 \\ x-y & 8 \end{bmatrix} = \begin{bmatrix} 1 & 6 \\ 3 & 8 \end{bmatrix}$. Find x and y.

Solution: By the definition of equality of matrices, $\begin{matrix} \boldsymbol{x}+\boldsymbol{y}=1 \\ \boldsymbol{x}-\boldsymbol{y}=3 \end{matrix}$

solving gives x = 2 and y = -1.

Activity 3.1.3: Find the values of x, y, z and w which satisfy the matrix equation

a. $\begin{bmatrix} \boldsymbol{x}-\boldsymbol{y} & 2\boldsymbol{x}+\boldsymbol{z} \\ 2\boldsymbol{x}-\boldsymbol{y} & 3\boldsymbol{z}+\boldsymbol{w} \end{bmatrix} = \begin{bmatrix} -1 & 5 \\ 0 & 13 \end{bmatrix}$

b. $\begin{bmatrix} x+3 & 2y+x \\ z-1 & 4w-6 \end{bmatrix} = \begin{bmatrix} 0 & -7 \\ 3 & 2w \end{bmatrix}$

3.2. Types of matrices: Square, identity, scalar, diagonal, triangular, symmetric, and skew symmetric matrices

Certain types of matrices, which play important roles in matrix theory, are now considered.

Row Matrix: A matrix that has exactly one row is called **a row matrix**. For example, the matrix $\boldsymbol{A}=\begin{bmatrix} 5 & 2 & -1 & 4 \end{bmatrix}$ is a row matrix of order 1×4.

Column Matrix: A matrix consisting of a single column is called **a column matrix**. For example, the matrix $\boldsymbol{B}=\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$ is a 3×1 column matrix.

Zero or Null Matrix: A matrix whose entries are all 0 is called a **zero** or **null matrix**. It is usually denoted by $0_{\boldsymbol{m}\times\boldsymbol{n}}$ or more simply by **0**. For example, $0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is a 2×4 zero matrix.

Square Matrix: An $\boldsymbol{m}\times\boldsymbol{n}$ matrix is said to be a **square matrix** of **order n** if **m = n**. That is, if it has the same number of columns as rows.

For example, $\begin{bmatrix} -3 & 4 & 6 \\ 2 & 1 & 3 \\ 5 & 2 & -1 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1 \\ 5 & 6 \end{bmatrix}$ are square matrices of order 3 and 2 respectively.

In a square matrix $\boldsymbol{A = (a_{ij})}$ of order n, the entries $\boldsymbol{a_{11}, a_{22}, ..., a_{nn}}$ which lie on the diagonal extending from the left upper corner to the lower right corner are called the **main diagonal entries**, or more simply the **main diagonal**. Thus, in the matrix C = $\begin{bmatrix} 3 & 2 & 4 \\ 1 & 6 & 0 \\ 5 & 1 & 8 \end{bmatrix}$ the entries $c_{11} = 3,\ c_{22} = 6$ and $c_{33} = 8$ constitute the main diagonal.

Note: The sum of the entries on the main diagonal of a square matrix A of order n is called the **trace of A**. That is, Trace of A = $\sum_{i=1}^{n} a_{ii}$.

Activity 3.2.1: Find the trace of C in the above example.

Triangular Matrix: A square matrix is said to be an upper (lower) triangular matrix if all entries below (above) the main diagonal are zeros.

For example, $\begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 2 \\ 0 & 0 & -3 \end{bmatrix}$ and $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 6 & 1 & 2 & 0 \\ -2 & -4 & 8 & 6 \end{bmatrix}$ are upper and lower triangular matrices, respectively.

Diagonal Matrix: A square matrix is said to be diagonal if each of the entries not falling on the main diagonal is zero. Thus a square matrix $\boldsymbol{A = (a_{ij})}$ is diagonal if $\boldsymbol{a_{ij} = 0}$ for $\boldsymbol{i \neq j}$.

Activity 3.2.2: What about for i = j?

For example, $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix}$ is a diagonal matrix.

Notation: A diagonal matrix A of order n with diagonal elements $a_{11}, a_{22},...,a_{nn}$ is denoted by $A = diag\,(a_{11}, a_{22},...,a_{nn})$

Scalar matrix: A diagonal matrix whose all the diagonal elements are equal is called a scalar matrix.

For example, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is a scalar matrix.

Note: Let $A = (a_{ij})$ be a square matrix. A is a scalar matrix if and only if $a_{ij} = \begin{cases} k, if i = j \\ 0, if i \neq j \end{cases}$.

Identity Matrix or Unit Matrix: A square matrix is said to be identity matrix or unit matrix if all its main diagonal entries are 1's and all other entries are 0's. In other words, a diagonal matrix whose all main diagonal elements are equal to 1 is called an identity or unit matrix. An identity matrix of order n is denoted by I_n or more simply by **I**.

For example, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is identity matrix of order 3. $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is identity matrix of order 2.

Note: Let $A = (a_{ij})$ be a square matrix. A is an identity matrix if and only if

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

3.3 Algebra of matrices

Activity 3.3.1: $\begin{bmatrix} 30 & 18 \\ 20 & 12 \\ 16 & 10 \end{bmatrix} + \begin{bmatrix} 36 & 20 \\ 24 & 18 \\ 20 & 12 \end{bmatrix} = \begin{bmatrix} 66 & ? \\ ? & 30 \\ ? & ? \end{bmatrix}$. Can you guess what number should appear in the entries marked by question mark?

Addition of matrices: Let A and B be two matrices of the same order. Then the addition of A and B, denoted by A + B, is the matrix obtained by adding corresponding entries of A and B. Thus, if

$\boldsymbol{A} = (\boldsymbol{a_{ij}})_{\boldsymbol{m}\times\boldsymbol{n}}$ and $\boldsymbol{B} = (\boldsymbol{b_{ij}})_{\boldsymbol{m}\times\boldsymbol{n}}$, then $\boldsymbol{A} + \boldsymbol{B} = (\boldsymbol{a_{ij}} + \boldsymbol{b_{ij}})_{\boldsymbol{m}\times\boldsymbol{n}}$.

Remark: Notice that we can add two matrices if and only if they are of the same order. If they are, we say they are conformable for addition. Also, the order of the sum of two matrices is same as that of the two original matrices.

Activity 3.3.2: Given the matrices A, B, and C below

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 5 & 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & -1 & 3 \\ 2 & 4 & 2 \\ 3 & 6 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}$$

Find, if possible. **a)** A + B **b)** B + C

If **A** is any matrix, the negative of **A**, denoted by **–A**, is the matrix obtained by replacing each entry in **A** by its negative. For example, if

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 \\ 5 & 4 \\ -6 & 0 \end{bmatrix}, \text{ then } -\boldsymbol{A} = \begin{bmatrix} -2 & 1 \\ -5 & -4 \\ 6 & 0 \end{bmatrix}$$

Properties of Addition of Matrices

1. Matrix addition is commutative. That is, if **A** and **B** are two matrices of the same order, then **A + B = B + A**.
2. Matrix addition is associative. That is, if **A, B** and **C** are three matrices of the same order, then **(A + B) + C = A + (B + C).**
3. Existence of additive identity. That is, if 0 is the zero matrix of the same order as that of the matrix **A**, then **A + 0 = A = 0 + A**.
4. Existence of additive inverse. That is, if A is any matrix, then

$$\mathbf{A + (-A) = 0 = (-A) + A}$$

Note: The zero matrix plays the same role in matrix addition as the number zero does in addition of numbers.

Subtraction of Matrices: Let A and B be two matrices of the same order. Then by **A – B**, we mean **A + (-B).** In other words, to find **A – B** we subtract each entry of **B** from the corresponding entry of **A**.

Example 3.3.1: Let $\boldsymbol{A} = \begin{bmatrix} 4 & -1 \\ 2 & 3 \\ 5 & -7 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ 5 & -2 \\ 6 & 1 \end{bmatrix}$

Then $\boldsymbol{A-B} = \begin{bmatrix} 4-0 & -1-2 \\ 2-5 & 3-(-2) \\ 5-6 & -7-1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -3 & 5 \\ -1 & -8 \end{bmatrix}$

Multiplication of a Matrix by a Scalar

Let **A** be an $\boldsymbol{m \times n}$ matrix and *k* be a real number (called **a scalar**). Then the multiplication of **A** by *k*, denoted by *k* **A**, is the $\boldsymbol{m \times n}$ matrix obtained by multiplying each entry of **A** by *k*. This operation is called **scalar multiplication**.

Example 3.3.2: If $\boldsymbol{A} = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$ ***and*** $\boldsymbol{B} = \begin{bmatrix} 7 & 6 & 3 \\ 1 & 4 & 5 \end{bmatrix}$. Find 2A + 3B.

Solution: $2\boldsymbol{A} = 2\begin{bmatrix} 0 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 6 \\ 4 & 2 & 8 \end{bmatrix}$ and $3B = 3\begin{bmatrix} 7 & 6 & 3 \\ 1 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 21 & 18 & 9 \\ 3 & 12 & 15 \end{bmatrix}$

$$2A + 3B = \begin{bmatrix} 0 & 4 & 6 \\ 4 & 2 & 8 \end{bmatrix} + \begin{bmatrix} 21 & 18 & 9 \\ 3 & 12 & 15 \end{bmatrix} = \begin{bmatrix} 21 & 22 & 15 \\ 7 & 14 & 23 \end{bmatrix}$$

Example 3.3.3: Express the matrix equation $x\begin{bmatrix} 2 \\ 1 \end{bmatrix} - y\begin{bmatrix} 3 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 8 \\ 11 \end{bmatrix}$ as a system of equations and solve.

Solution: The given matrix equation gives

$$\begin{bmatrix} 2\boldsymbol{x} \\ \boldsymbol{x} \end{bmatrix} - \begin{bmatrix} 3\boldsymbol{y} \\ 5\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} 16 \\ 22 \end{bmatrix} \Rightarrow \begin{bmatrix} 2\boldsymbol{x} - 3\boldsymbol{y} \\ \boldsymbol{x} - 5\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} 16 \\ 22 \end{bmatrix}$$

By equality of matrices we have

$$2\boldsymbol{x} - 3\boldsymbol{y} = 16$$
$$\boldsymbol{x} - 5\boldsymbol{y} = 22$$

Solving gives x = 2, y = -4

Properties of scalar multiplications

1. If **A** and **B** are two matrices of the same order and if *k* is a scalar, then
   $$k(A + B) = kA + kB$$
2. If $\boldsymbol{k_1}$ and $\boldsymbol{k_2}$ are two scalars and if **A** is a matrix, then
   $$(k_1 + k_2)A = k_1A + k_2A$$
3. If $\boldsymbol{k_1}$ and $\boldsymbol{k_2}$ are two scalars and if **A** is a matrix, then
   $$(k_1k_2)A = k_1(k_2A) = k_2(k_1A)$$

Multiplication of Matrices

While the operations of matrix addition and scalar multiplication are fairly straightforward, the product **AB** of matrices **A** and **B** can be defined under the condition that *the number of columns of **A** must be equal to the number of rows of **B***. If the number of columns in the matrix **A** equals the number of rows in the matrix **B**, we say that the matrices are conformable for the product **AB**.

Because of wide use of matrix multiplication in application problems, it is important that we learn it well. Therefore, we will try to learn the process in a step by step manner. We first begin by finding a product of a row matrix and a column matrix.

Example 3.3.4: Given A = $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$ and B = $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, find the product AB.

Solution: The product is a 1×1 matrix whose entry is obtained by multiplying the corresponding entries and then forming the sum.

$$AB = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$= [\,(2a + 3b + 4c)]$$

Note that AB is a 1×1 matrix, and its only entry is 2a + 3b + 4c.

Example 3.3.4: Given A = $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$ and B = $\begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$, find the product AB.

Solution: AB = $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = [10 + 18 + 28] = [56]$

Note: In order for a product of a row matrix and a column matrix to exist, the number of entries in the row matrix must be the same as the number of entries in the column matrix.

Example 3.3.5: Here is an application: Suppose you sell 3 T-shirts at \$10 each, 4 hats at \$15 each, and 1 pair of shorts at \$20. Then your total revenue is

$$\underbrace{\begin{pmatrix} 10 & 15 & 20 \end{pmatrix}}_{Price} \underbrace{\begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix}}_{Quantity} = \underbrace{\left((10 \times 3) + (15 \times 4) + (20 \times 1)\right)}_{Revenue} = (110)$$

Example 3.3.6: Given A = $\begin{bmatrix} 2 & 3 & 4 \end{bmatrix}$ and B = $\begin{bmatrix} 5 & 3 \\ 6 & 4 \\ 7 & 5 \end{bmatrix}$, find the product AB.

Solution: We already know how to multiply a row matrix by a column matrix. To find the product AB, in this example, we will be multiplying the row matrix A to both the first and second columns of matrix B, resulting in a 1×2 matrix.

$$AB = [\, 2 \cdot 5 + 3 \cdot 6 + 4 \cdot 7 \qquad 2 \cdot 3 + 3 \cdot 4 + 4 \cdot 5 \,]$$
$$= [\, 56 \quad 38 \,]$$

We have just multiplied a 1×3 matrix by a matrix whose order is 3×2. So unlike addition and subtraction, it is possible to multiply two matrices with different dimensions as long as the number of entries in the rows of the first matrix is the same as the number of entries in columns of the second matrix.

Activity 3.3.3: **1.** Given the matrices E, F, G and H, below

$$E = \begin{bmatrix} 1 & 2 \\ 4 & 2 \\ 3 & 1 \end{bmatrix} \quad F = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} \quad G = \begin{bmatrix} 4 & 1 \end{bmatrix} \quad H = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$$

Find, if possible. a) GH b) FH c) EF d) FE

2. Given the matrices R, S, and T below.

$$R = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \\ 2 & 3 & 1 \end{bmatrix} \quad S = \begin{bmatrix} 0 & -1 & 2 \\ 3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \quad T = \begin{bmatrix} -2 & 3 & 0 \\ -3 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}$$

Find 2RS – 3ST.

We summarize matrix multiplication as follows:

In order for product AB to exist, the number of columns of A, must equal the number of rows of B. If matrix A is of dimension $m \times n$ and B of dimension $n \times p$, the product AB will have the dimension $m \times p$. Let $\boldsymbol{A = (a_{ij})}$ be an $\boldsymbol{m \times n}$ matrix and $\boldsymbol{B = (b_{jk})}$ be an $\boldsymbol{n \times p}$ matrix. Then the product **AB** is the $\boldsymbol{m \times p}$ matrix defined by $\boldsymbol{AB = (c_{ik}),}$ where

$$\boldsymbol{c_{ik} = a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk} = \sum_{j=1}^{n} a_{ij} b_{jk}}\text{ , I = 1,2,…,m and k = 1,2,…,p}$$

Thus, the product **AB** is the $m \times p$ matrix, where each entry c_{ik} of **AB** is obtained by multiplying corresponding entries of the i^{th} row of **A** by those of the k^{th} column of **B** and then finding the sum of the results.

Activity 3.3.4: 1. If a matrix A is 3x5 and the product AB is 3x7, then what is the order of B?

2. How many rows does X have if XY is a 2x6 matrix?

Remark: The definition refers to the product **AB**, in that order, **A** is the left factor called **pre factor** and **B** is the right factor called **post factor.**

Example 3.3.7: Find the product **AB** if $A = \begin{bmatrix} 1 & -4 \\ 5 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 4 & 1 & 6 \\ 2 & 7 & 3 & 8 \end{bmatrix}$

Solution: Since the number of columns of **A** is equal to the number of rows of B, the product **AB=C** is defined. Since **A** is 3×2 and B is 2×4, the product AB will be 3×4

$$AB = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix}$$

The entry c_{11} is obtained by summing the products of each entry in row 1 of A by the corresponding entry in column 1 of B, that is.

$c_{11} = (1)(-2) + (-4)(2) = -10$. Similarly, for C_{21} we use the entries in row 2 of A and those in column 1 of B, that is C_{21} = (5) (-2) + (3) (2) = -4.

Also, C_{12} = (1) (4) + (-4) (7) = -24

C_{13} = (1) (1) + (-4) (3) = -11

C_{14} = (1) (6) + (-4) (8) = -26

C_{22} = (5) (4) + (3) (7) = 41

C_{23} = (5) (1) + (3) (3) = 14

C_{24} = (5) (6) + (3) (8) = 54

C_{31} = (0) (-2) + (2) (2) = 4

$C_{32} = (0)(4) + (2)(7) = 14$

$C_{33} = (0)(1) + (2)(3) = 6$

$C_{34} = (0)(6) + (2)(8) = 16$

Thus $\boldsymbol{AB} = \begin{bmatrix} -10 & -24 & -11 & -26 \\ -4 & 41 & 14 & 54 \\ 4 & 14 & 6 & 16 \end{bmatrix}$

Observe that the product **BA** is not defined since the number of columns of **B** is not equal to the number of rows of **A**. This shows that matrix multiplication is not commutative. That is, for any two matrices **A** and **B**, it is usually the case that $\boldsymbol{AB} \neq \boldsymbol{BA}$ (even if both products are defined).

Example 3.3.8: Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$, then

$AB = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$, $BA = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. Thus, $AB \neq BA$.

Activity 3.3.4: 1. Which of the following are defined?

i) $\begin{pmatrix} 5 & 2 & 10 \end{pmatrix}\begin{pmatrix} 3 \\ -4 \end{pmatrix}$ iii) $\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix}\begin{pmatrix} \boldsymbol{c} & \boldsymbol{d} \end{pmatrix}$

ii) $\begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \end{pmatrix}\begin{pmatrix} \boldsymbol{c} & \boldsymbol{d} \end{pmatrix}$ iv) $\begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \end{pmatrix}\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix}$

2. If $\boldsymbol{A} = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} \\ 3 & 5 \end{bmatrix}$, find a and b such that AB = BA.

Note: **1.** AB = 0 does not necessarily imply A = 0 or B = 0.

2. AB = AC does not necessarily imply B = C.

Example 3.3.9: 1) Let $A = \begin{bmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix}$, then $AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

2) Let $A = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 1 & -1 & -2 \\ 3 & -2 & -1 & -1 \\ 2 & -5 & -1 & 0 \end{bmatrix}$

then $AB = \begin{bmatrix} -3 & -3 & 0 & 1 \\ 1 & 15 & 0 & -5 \\ -3 & 15 & 0 & -5 \end{bmatrix} = AC$. But $B \neq C$.

In this chapter, we will be using matrices to solve linear systems. Later, we will be asked to express linear systems as the **matrix equation AX = B**, where A, X, and B are matrices. The matrix A is called the **coefficient matrix.**

Example 3.3.10: Verify that the system of two linear equations with two unknowns:

$ax + by = h$
$cx + dy = k$ can be written as AX = B, where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \end{bmatrix} \text{ and } B = \begin{bmatrix} h \\ k \end{bmatrix}$$

Solution: If we multiply the matrices A and X, we get

$$AX = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

If AX = B, then

$$\begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} = \begin{bmatrix} h \\ k \end{bmatrix}$$

If two matrices are equal, then their corresponding entries are equal.

Therefore, it follows that

$ax + by = h$
$cx + dy = k$

Activity 3.3.5: Express the following system as AX = B.

$$\begin{aligned} 2x + 3y - 4z &= 5 \\ 3x + 4y - 5z &= 6 \\ 5x \qquad\quad - 6z &= 7 \end{aligned}$$

Properties of Matrix multiplication

If A, B and C are any matrices, and if I is an identity matrix, then the following hold, whenever the dimensions of the matrices are such that the products are defined.

A(BC) = (AB)C	Associative Law
A·I = I·A =A (The order of I and A is the same)	Multiplicative Identity Law
A(B + C) = AB + AC	Left Distributive Law
(A + B)C =AC+BC	Right Distributive Law
A·**0** = **0**·A = **0**	Multiplication by Zero

Remark: For real numbers, a multiplied by itself n times can be written as a^n.

Similarly, a square matrix A multiplied by itself n times can be written as A^n. Therefore, A^2 means AA, A^3 means AAA and so on.

Exercise:

1. If $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \\ 1 & -3 & -1 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 4 & 5 & 6 \\ -1 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$ and $\boldsymbol{C} = \begin{bmatrix} -1 & -2 & 1 \\ -1 & 2 & 3 \\ -1 & -2 & 2 \end{bmatrix}$

Find each of the following

(i) A + B ii) 2B – 3C

iii) A + B – C iv) A – 2B + 3C v) 2A – C

2. Let $\boldsymbol{A}=\begin{bmatrix} 5 & 2 \\ -1 & 3 \end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix} 2 & 4 \\ 6 & 1 \end{bmatrix}$, $\boldsymbol{C}=\begin{bmatrix} 1 & 3 \\ 7 & 2 \end{bmatrix}$

Find the following: i) AB ii) BC iii) (AB)C iv) A(BC)

3. If $\boldsymbol{A}=\begin{bmatrix} 4 & -1 & -4 \\ 4 & 0 & -4 \\ 3 & -1 & -3 \end{bmatrix}$, compute $\mathbf{A^2}$**.** Is it equal to I_3, where I_3 is the identity matrix of order 3?

Transpose of a matrix

Definition 3.3.1: Let **A** be an $\boldsymbol{m\times n}$ matrix. The transpose of **A**, denoted by $\boldsymbol{A'}$ or $\boldsymbol{A^t}$, is the $\boldsymbol{n\times m}$ matrix obtained from **A** by interchanging the rows and columns of **A**. Thus the first row of **A** is the first column of $\mathbf{A^t}$, the second row of **A** is the second column of $\mathbf{A^t}$ and so on.

Example 3.3.11: If $\boldsymbol{A}=\begin{pmatrix} 2 & -4 & 6 \\ 3 & 1 & 4 \end{pmatrix}$, then $\boldsymbol{A^t}=\begin{pmatrix} 2 & 3 \\ -4 & 1 \\ 6 & 4 \end{pmatrix}$

Activity 3.3.6: Find a 3x3 matrix A for which $\mathbf{A = A^t}$.

Properties: a) If **A** and **B** have the same order, $\boldsymbol{(A\pm B)^t = A^t \pm B^t}$.

b) For a scalar $\boldsymbol{k}$, $\boldsymbol{(kA)^t = kA^t}$.

c) If A is $\boldsymbol{m\times n}$ and **B** is $\boldsymbol{n\times p}$, then $(AB)^t = B^t A^t$.

d) $\boldsymbol{\left(A^t\right)^t = A}$

Definition 3.3.2: A square matrix **A** is said to be **orthogonal** if $\boldsymbol{AA^t = A^tA = I}$

Example 3.3.12: $\boldsymbol{A}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ is orthogonal (verify)

Definition 3.3.3: A square matrix $A = (a_{ij})$ is said to be **symmetric** if $A^t = A$, or equivalently, if $a_{ij} = a_{ji}$ for each i and j.

Example 3.3.13: $A = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 0 & -3 \\ 5 & -3 & 6 \end{pmatrix}$ is symmetric.

Activity 3.3.7: 1. For $A = \begin{pmatrix} a & 3 & 4 & 8 \\ b & c & -3 & 9 \\ d & e & f & 10 \\ g & h & i & j \end{pmatrix}$ is to be a symmetric matrix, what numbers should the letters **a** to j represent?

2. a) Does a symmetric matrix have to be square?

b) Are all square matrices symmetric?

Definition 3.3.4: A square matrix $A = (a_{ij})$ is said to be **skew symmetric** if $A^t = -A$, or equivalently, if $a_{ij} = -a_{ji}$ for each i and j

Remark: $a_{ii} = -a_{ii} \Rightarrow 2a_{ii} = 0$ **or** $a_{ii} = 0$**.** Hence elements of main diagonal of a skew-symmetric matrix are all zero.

Example 3.3.14: For $A = \begin{pmatrix} 0 & 5 & 7 \\ -5 & 0 & 3 \\ -7 & -3 & 0 \end{pmatrix}$, $A^t = \begin{pmatrix} 0 & -5 & -7 \\ 5 & 0 & -3 \\ 7 & 3 & 0 \end{pmatrix} = -A$.

So A is skew-symmetric.

Properties of symmetric and skew-symmetric matrices

1. For any square matrix A, $A + A^t$ is symmetric and $A - A^t$ is skew- symmetric
2. If A and B are two symmetric (or skew symmetric) matrices of the same order, then so is A + B. Proof (exercise)
3. If A is symmetric or skew symmetric, then so is kA. Proof (exercise)

4. Let A and B be symmetric matrices of the same order. The product AB is symmetric if and only if AB = BA.

   Proof: AB is symmetric $\Rightarrow (\boldsymbol{AB})^{\boldsymbol{t}} = \boldsymbol{AB} \Rightarrow \boldsymbol{B^t A^t} = \boldsymbol{AB}$

   $\Rightarrow \boldsymbol{BA} = \boldsymbol{AB}$

   Suppose **AB = BA.** Then $\mathbf{(AB)^t = B^tA^t = BA = AB}$

Exercise

1. a) Form a 4 by 5 matrix, *B*, such that $b_{ij} = i*j$, where * represents multiplication.

b) What is B^T? **c)** Is *B* symmetric? Why or why not?

2. Given $\boldsymbol{A} = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 4 & 5 \\ 1 & 3 & 6 \end{bmatrix}$ and B= $\begin{bmatrix} -2 & 4 & 3 \\ 5 & 1 & 7 \\ 2 & 3 & 8 \end{bmatrix}$. Verify that

i) $(A+B)^t = A^t + B^t$, ii) $(AB)^t = B^t A^t$ iii) $(2A)^t = 2A^t$

3. Let $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$, is $A^t A$ is symmetric?

3.4 Elementary row and column operations

***Elementary row operations*:**

1. (Replacement) Replace one row (say R_i) by the sum of itself and a multiple of another row (say R_j). This is abbreviated as $R_i \to k R_j + R_i$.
2. (Interchange) Interchange two rows (say R_i *and* R_j). This is abbreviated as $R_i \leftrightarrow R_j$.
3. (Scaling) Multiply all entries in a row (say R_i) by a nonzero constant (scalar) k. This is abbreviated as $R_i \to k R_i$

For elementary column operations "row" by "column" in (1), (2) and (3) above.

We say that two matrices are **row equivalent** if one is obtained from the other by a finite sequence of elementary row operations.

It is important to note that row operations are reversible. If two rows are interchanged, they can be returned to their original positions by another interchange. If a row is scaled by a nonzero constant c, then multiplying the new row by $\frac{1}{c}$ produces the original row. Finally, consider a replacement operation involving two rows, say rows i and j, and suppose c times row i is added to row j to produce a new row j. To "reverse" this operation, add – c times row i to the new row j and obtain the original row j.

Example 3.4.1: Find the elementary row operation that transforms the first matrix in to the second, and then find the reverse row operation that transforms the second matrix in to the first.

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -4 \\ 0 & -3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -2 \\ 0 & -3 & 4 \end{bmatrix}$$

Solution: $\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -4 \\ 0 & -3 & 4 \end{bmatrix} \xrightarrow{R_2 \longrightarrow \frac{1}{2} R_2} \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -2 \\ 0 & -3 & 4 \end{bmatrix}$

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -4 \\ 0 & -3 & 4 \end{bmatrix} \xleftarrow{2R_2 \longleftarrow R_2} \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -2 \\ 0 & -3 & 4 \end{bmatrix}$$

Activity 3.4.1: Find the elementary row operation that transforms the first matrix in to the second, and then find the reverse row operation that transforms the second matrix in to the first.

a) $\begin{bmatrix} 0 & 5 & -3 \\ 1 & 5 & -2 \\ 2 & 1 & 8 \end{bmatrix}, \quad \begin{bmatrix} 1 & 5 & -2 \\ 0 & 5 & -3 \\ 2 & 1 & 8 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 2 & -5 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 3 & -5 \end{bmatrix}$

3.5 Row reduced echelon form of a matrix

In the definition that follows, a **nonzero row (or column)** in a matrix means a row (or column) that contains at least one non-zero entry; a **leading entry** of a row refers to the left most nonzero entry (in a non zero row).

Definition 3.5.1: A matrix is in **echelon** form (or **row echelon** form) if it has the following three properties:

1) All non zero rows are above any rows of all zeros.
2) Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3) All entries in a column below a leading entry are zero.

If a matrix in echelon form satisfies the following additional condition then it is in **reduced echelon** form (or **row reduced echelon** form)

4) The leading entry in each non zero row is 1
5) Each leading 1 is the only nonzero entry in its column.

Example 3.5.1: The following matrices are in row echelon form, in fact the second matrix is in row reduced echelon form

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & \frac{5}{2} \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Definition 3.5.2: (i) A matrix which is in row echelon form is called an **echelon matrix**.

(ii) A matrix which is in row reduced echelon form is called a **reduced echelon matrix**.

Note: 1) Each matrix is row equivalent to one and only one row reduced echelon matrix. But a matrix can be row equivalent to more than one echelon matrices.

2) If matrix A is row equivalent to an echelon matrix U, we call U an echelon form of A. If U is in reduced echelon form, we call U the reduced echelon form of A.

Activity 3.5.1: Determine which of the following matrices are in row reduced echelon form and which others are in row echelon form (but not in reduced echelon form)

a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 0 & -5 & 0 & -8 & 3 \\ 0 & 1 & 4 & -1 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3.6. Rank of a matrix using elementary row/column operations

Let A be an mxn matrix and U be an echelon or the reduced echelon form of A. The rank of A is denoted by Rank (A) and is defined as the number of non zero rows of U.

Example 3.6.1: Find the rank of each of the matrices given in the above activity.

Solution: a) has rank 2; b) has rank 1; c, d, and e have rank 3.

Activity 3.6.1: Find the row reduced echelon form of each of the following matrices and determine the rank.

a) $\begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \\ 3 & 5 & 7 & 9 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & 3 \end{bmatrix}$

3.7. System of Linear equations

In this section we will present certain systematic methods for solving system of linear equations.

Definition 3.7.1: A **linear equation** in the variables $x_1, x_2, ..., x_n$ over the real field $\Re$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \qquad (1)$$

where b and the coefficients $a_1, a_2, ..., a_n$ are given real numbers.

Definition 3.7.2: A **system of linear equations** (or **a linear system**) is a collection of one or more linear equations involving the same variables, say $x_1, x_2, ..., x_n$.

Now consider a system of m linear equations in n-unknowns $x_1, x_2, ..., x_n$:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ &. \\ &. \\ &. \\ &a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{aligned} \qquad (2)$$

If $b_1 = b_2 = ... = b_m = 0$ then we say that the system is **homogeneous**. If $b_i \neq 0$ for some i∈{1,2,3, . . ., m} then the system is called **non homogeneous**. In matrix notation, the linear system (2) can be written as $AX = B$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & & & \\ . & & & \\ . & & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} \text{ and } \quad B = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{bmatrix}$$

We call **A** the **coefficient matrix** of the system (2).

Observe that entries of the k-th column of **A** are the coefficients of the variable $\mathbf{x_k}$ in (2).

The $\mathbf{m}\times\mathbf{(n+1)}$ matrix whose first **n** columns are the columns of A (the coefficient matrix) and whose last column is B is called the **augmented matrix** of the system. We denote it by **[A|B]**. The augmented matrix determines the system (2) completely because it contains all the coefficients and the constants to the right side of each equation in the system. For example for the non homogeneous linear system

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 2 \\ x_2 - 2x_3 &= 4 \\ -2x_1 - 3x_2 - 3x_3 &= 5 \end{aligned} \qquad (3)$$

The matrix $A = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -2 \\ -2 & -3 & -3 \end{bmatrix}$ is the coefficient matrix and $\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & -2 & 4 \\ -2 & -3 & -3 & 5 \end{bmatrix}$ is the augmented matrix.

Are the coefficient matrix and the augmented matrix of a homogeneous linear system equal? Why?

A **solution** of a linear system in n-unknowns $x_1, x_2, ..., x_n$ is an n-tuple $(s_1, s_2, ..., s_n)$ of real numbers that makes each of the equations in the system a true statement when s_i is substituted for x_i, i = 1,2, . . ., n. The set of all possible solutions is called the **solution set** of the linear system. We say that two linear systems are **equivalent** if they have the same solution set.

Activity 3.7.1: 1) Give the coefficient matrix and the augmented matrix of the linear system

$$\begin{aligned} x_1 + 3x_2 &= 2x_3 \\ x_1 &= 11x_3 \\ 2x_1 - x_2 + 4x_3 &= 0 \end{aligned}$$

2) Does the linear system $\begin{aligned} -x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0 \end{aligned}$ have a solution?

3) Find the solution set of the linear system $\begin{aligned} u + v &= 2 \\ u - v &= 1 \end{aligned}$

How many solutions does it have?

4) Given the system $\begin{cases} 2x_1 - x_2 + \frac{3}{2}x_3 = 8 \\ x_1 - 4x_3 = -7 \end{cases}$. Is $\left(5, \frac{13}{2}, 3\right)$ a solution of the linear system? What about (-7, -22, 0) and $\left(3, \frac{-25}{2}, 1\right)$?

The activity given above illustrates the following general fact about linear systems.

A system of linear equations has either

1. no solution, or
2. exactly one solution, or
3. Infinitely many solutions.

We say that a linear system is **consistent** if it has either one solution or infinitely many solutions; a system is **inconsistent** if it has no solution.

Activity 3.7.2:

1. The homogeneous linear system $AX = O$ is consistent for any $m \times n$ matrix **A**. Explain, why?
2. Consider a linear system of two equations in two unknowns, give geometric interpretation if the system has

   i) no solution ii) exactly one solution iii) many solutions

   Do the same for a linear system of three equations in three unknowns.

Solving a linear system

This is the process of finding the solutions of a linear system. We first see the technique of elimination (Gaussian elimination method) and then we add two more techniques, matrix inversion method and Cramer's rule.

Gaussian Elimination Method

The Gaussian elimination method is a standard method for solving linear systems. It applies to any system, no matter whether $m < n$, $m = n$ or $m > n$ (where m and n are number of equations and variables respectively). We know that equivalent linear systems

have the same solutions. Thus the basic strategy in this method is to replace a given system with an equivalent system, which is easier to solve.

The basic operations that are used to produce an equivalent system of linear equations are the following:

1. Replace one equation by the sum of itself and a multiple of another equation.
2. Interchange two equations
3. Multiply all the terms in an equation by a non zero constant.

Activity 3.7.3: Why these three operations do not change the solution set of the system?

We illustrate this technique by using the following example.

Example 3.7.1: Solve the system

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 2 \\ x_2 - 2x_3 &= 4 \\ -2x_1 - 3x_2 - 3x_3 &= 5 \end{aligned}$$

Solution: We perform the elimination procedure with and without matrix notation of the system. For each step we put the resulting system and its augmented matrix side by side for comparison:

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

We keep x_1 in the first equation and eliminate it from the other equations. For this replace the third equation by the sum of itself and two times equation 1.

$$\begin{aligned} 2.[eq.1]: &\quad 2x_1 + 6x_2 - 2x_3 = 4 \\ +[eq.3]: &\quad -2x_1 - 3x_2 - 3x_3 = 5 \\ \hline [New\ eq.3] &\quad 3x_2 - 5x_3 = 9 \end{aligned}$$

We write the new equation in place of the original third equation:

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 2 \\ x_2 - 2x_3 &= 4 \\ 3x_2 - 5x_3 &= 9 \end{aligned} \qquad \xrightarrow{R_3 \to R_3 + 2R_1} \qquad \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & -2 & 4 \\ 0 & 3 & -5 & 9 \end{bmatrix}$$

Next use the x_2 in equation 2 to eliminate $3x_2$ in equation 3.

$$\begin{array}{ll} -3.[eq.2]: & -3x_2 + 6x_3 = -12 \\ +[eq.3]: & \underline{3x_2 - 5x_3 = 9} \\ [New\,eq.3] & x_3 = -3 \end{array}$$

The resulting equivalent system is:

$$\begin{array}{r} x_1 + 3x_2 - x_3 = 2 \\ x_2 - 2x_3 = 4 \\ x_3 = -3 \end{array} \xrightarrow{R_3 \to R_3 - 3R_2} \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

Now we eliminate the $-2x_3$ term form equation 2. For this we use x_3 in equation 3.

$$\begin{array}{ll} 2.[\boldsymbol{eq}.3]: & 2x_3 = -6 \\ +[\boldsymbol{eq}.2]: & \underline{x_2 - 2x_3 = 4} \\ [\boldsymbol{New\,eq}.3] & x_3 = -3 \end{array}$$

From this we get

$$\begin{array}{r} x_1 + 3x_2 - x_3 = 2 \\ x_2 = -2 \\ x_3 = -3 \end{array} \xrightarrow{R_2 \to R_2 + 2R_3} \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

Again by using the x_3 term in equation 3, we eliminate the $-x_3$ term in equation 1.

$$\begin{array}{ll} 1.[\boldsymbol{eq}.3]: & x_3 = -3 \\ +[\boldsymbol{eq}.1]: & \underline{x_1 + 3x_2 - x_3 = 2} \\ [\boldsymbol{New\,eq}.1] & x_1 + 3x_2 = -1 \end{array}$$

Thus we get the system

$$\begin{array}{r} x_1 + 3x_2 = -1 \\ x_2 = -2 \\ x_3 = -3 \end{array} \xrightarrow{R_1 \to R_1 + R_3} \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

Finally, we eliminate the $3x_2$ term in equation 1. We use the x_2 term in equation 2 to eliminate the $3x_2$ term above it.

$$\begin{array}{ll} -3.[\boldsymbol{eq}.2]: & -3x_2 = 6 \\ +[\boldsymbol{eq}.1]: & \underline{x_1 + 3x_2 = 2} \\ [\boldsymbol{New\,eq}.1] & x_1 = 5 \end{array}$$

So we have an equivalent system (to the original system) that is easier to solve.

$$\begin{aligned} x_1 &= 5 \\ x_2 &= -2 \\ x_3 &= -3 \end{aligned} \qquad \xrightarrow{R_1 \rightarrow R_1 - 3R_2} \qquad \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

Thus the system has only one solution, namely (5, -2, -3) or $\mathbf{x}_1 = 5, \mathbf{x}_2 = -2, \mathbf{x}_3 = -3$. To verify that (5, -2, -3) is a solution, substitute these values in to the left side of the original system, and compute:

$$5 + 3(-2) - (-3) = 5 - 6 + 3 = 2$$

$$-2 - 2(-3) = -2 + 6 = 4$$

$$-2(5) - 3(-2) - 3(-3) = -10 + 6 + 9 = 5$$

It is a solution, as it satisfies all the equation in the given system (3).

Example 3.7.1, illustrates how operations in a linear system correspond to operations on the appropriate rows of the augmented matrix. The three basic operations listed earlier correspond to the three elementary row operations on the augmented matrix.

Let us see how elementary row operations on the augmented matrix of a given linear system can be used to determine a solution of the system. Suppose a system of linear equations is changed to a new one via row operations on its augmented matrix. By considering each type of row operation it is easy to see that any solution of the original system remains a solution of the new system.

Conversely, since the original system can be produced via row operations on the new system, each solution of the new system is also a solution of the original system. From this we have the following important property.

- If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Thus to solve a linear system by elimination we first perform appropriate row operations on the augmented matrix of the system to obtain the augmented matrix of an equivalent

linear system which is easier to solve and use back substitution on the resulting new system. This method can also be used to answer questions about existence and uniqueness of a solution whenever there is no need to solve the system completely.

In Gaussian elimination method we either transform the augmented matrix to an echelon matrix or a reduced echelon matrix. That is we either find an echelon form or the reduced echelon form of the augmented matrix of the system. An echelon form of the augmented matrix enables us to answer the following two fundamental questions about solutions of a linear system. These are:

1. Is the system consistent; that is, does at least one solution exists?
2. If a solution exists, is it the only one; that is, is the solution unique?

Example 3.7.2: Determine if the following system is consistent. If so how many solutions does it have?

$$\begin{aligned} \boldsymbol{x}_1 - \boldsymbol{x}_2 + \boldsymbol{x}_3 &= 3 \\ \boldsymbol{x}_1 + 5\boldsymbol{x}_2 - 5\boldsymbol{x}_3 &= 2 \\ 2\boldsymbol{x}_1 + \boldsymbol{x}_2 - \boldsymbol{x}_3 &= 1 \end{aligned}$$

Solution: The augmented matrix is $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 1 & 3 \\ 1 & 5 & -5 & 2 \\ 2 & 1 & -1 & 1 \end{bmatrix}$

Let us perform a finite sequence of elementary row operations on the augmented matrix.

$$[A|B] = \begin{bmatrix} 1 & -1 & 1 & 3 \\ 1 & 5 & -5 & 2 \\ 2 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{R_2 \to -R_1 + R_2} \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & 6 & -6 & -1 \\ 2 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{R_3 \to -2R_1 + R_3}$$

$$\begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & 6 & -6 & -1 \\ 0 & 3 & -3 & -5 \end{bmatrix} \xrightarrow{R_3 \to \frac{-1}{2}R_2 + R_3} \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & 6 & -6 & -1 \\ 0 & 0 & 0 & -\frac{9}{2} \end{bmatrix}$$

The corresponding linear system of the last matrix is

$$\begin{aligned} x_1 - x_2 + x_3 &= 3 \\ 6x_2 - 6x_3 &= -1 \qquad (*) \\ 0 &= \frac{-9}{2} \end{aligned}$$

But the last equation $0.x_1 + 0.x_2 + 0.x_3 = \frac{-9}{2}$ is never true. That is there are no values x_1, x_2, x_3 that satisfy the new system (*). Since (*) and the original linear system have the same solution set, the original system is inconsistent (has no solution).

Example 3.7.3: Use Gaussian elimination to solve the linear system

$$\begin{aligned} 2\mathbf{x} - \mathbf{y} + \mathbf{z} &= 2 \\ -2\mathbf{x} + \mathbf{y} + \mathbf{z} &= 4 \\ 6\mathbf{x} - 3\mathbf{y} - 2\mathbf{z} &= -9 \end{aligned}$$

Solution: The augmented matrix of the given system is

$$[\mathbf{A}|\mathbf{B}] = \begin{bmatrix} 2 & -1 & 1 & 2 \\ -2 & 1 & 1 & 4 \\ 6 & -3 & -2 & -9 \end{bmatrix}$$

Let us find an echelon form of the augmented matrix first. From this we can determine whether the system is consistent or not . If it is consistent we go a head to obtain the reduced echelon form of **[A|B] ,** which enable us to describe explicitly all the solutions.

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ -2 & 1 & 1 & 4 \\ 6 & -3 & -2 & -9 \end{bmatrix} \xrightarrow{R_2 \to R_1 + R_2} \begin{bmatrix} 2 & -1 & 1 & 2 \\ 0 & 0 & 2 & 6 \\ 6 & -3 & -2 & -9 \end{bmatrix} \xrightarrow{R_3 \to -3R_1 + R_3}$$

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & -5 & -15 \end{bmatrix} \xrightarrow{R_3 \to \frac{5}{2}R_2 + R_3} \begin{bmatrix} 2 & -1 & 1 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \to \frac{1}{2}R_2}$$

$$\begin{bmatrix} 2 & -1 & 1 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to -R_2 + R_1} \begin{bmatrix} 2 & -1 & 0 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to \frac{1}{2}R_1}$$

$$\begin{bmatrix} 1 & \frac{-1}{2} & 0 & \frac{-1}{2} \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The associated linear system to the reduced echelon form of **[A|B]** is

$$\begin{aligned} \mathbf{x} - \tfrac{1}{2}\mathbf{y} &= \tfrac{-1}{2} \\ \mathbf{z} &= 3 \\ 0 &= 0 \end{aligned}$$

The third equation is $0\mathbf{x} + 0\mathbf{y} + 0\mathbf{z} = 0$. It is not an inconsistency, it is always true what ever values we take for ***x***, ***y***, ***z***. The system is consistent and if we assign any value λ for ***y*** in the first equation, we get $\mathbf{x} = \frac{-1}{2} + \frac{1}{2}\lambda$. From the second we have z = 3. Thus $\mathbf{x} = \frac{-1}{2} + \frac{1}{2}\lambda$, $\boldsymbol{y} = \lambda$ and $\boldsymbol{z}$ = 3 is the solution of the given system, where λ is any real number. There are an infinite number of solutions, for example,

$\boldsymbol{x} = \frac{-1}{2}$, $\boldsymbol{y}$ = 0, $\boldsymbol{z}$ = 3

$\boldsymbol{x}$ = 0, $\boldsymbol{y}$ = 1, $\boldsymbol{z}$ = 3 and so on.

In vector form the general solution of the given system is of the form $\left(\frac{-1}{2} + \frac{1}{2}\lambda, \lambda, 3\right)$ where $\lambda \in \Re$. What dose this represents in $\Re^3$?

Remark: A system of linear equation $\boldsymbol{AX} = \boldsymbol{B}$ is **consistent** iff the ranks of the coefficient matrix and the augmented matrix are equal.

Activity 3.7.4: Find the solution set of the system: $\begin{aligned} x_1 + x_2 + x_4 &= 1 \\ x_1 - x_2 + x_3 &= -2 \\ x_2 + x_3 - x_4 &= 0 \end{aligned}$.

Exercise 3.1:

1. Find the solution set of the following system:

a. $\begin{aligned} 2x_1 - x_2 - 3x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= -2 \\ 4x_1 - 3x_2 + x_3 &= -3 \\ x_1 - 5x_3 &= 3 \end{aligned}$

c. $\begin{aligned} x_1 + 3x_2 + 5x_3 - 2x_4 &= 11 \\ 3x_1 - 2x_2 - 7x_3 + 5x_4 &= 0 \\ 2x_1 + x_2 + x_4 &= 7 \end{aligned}$

b. $\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 0 \\ 3x_1 - x_2 + 5x_3 - x_4 &= 0 \\ 2x_1 + x_2 + x_4 &= 0 \end{aligned}$

d. $\begin{aligned} x_1 + 3x_2 + 2x_3 + 5x_4 &= 10 \\ 3x_1 - 2x_2 - 5x_3 + 4x_4 &= -5 \\ 2x_1 + x_2 - x_3 + 5x_4 &= 5 \end{aligned}$

2. For what values of λ and μ the system: $\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 10 \\ x + 2y + \lambda z &= \mu \end{aligned}$, has

   i. No solution ii. Unique solution iii. Infinitely many solution

3. Let M_{mxn} = the set of all mxn matrices. Is M_{mxn} a vector space under matrix addition and scalar multiplication?

4. Let $W = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & x \end{pmatrix} \middle| x \in \Re \right\}$. Is W a subspace of M_{2x2}, where

   $M_{2x2} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a,b,c,d \in \Re \right\}$?

5. Let $W = \left\{ \begin{pmatrix} 0 & x \\ 0 & y \end{pmatrix} \middle| x,y \in \Re \right\}$. Is W a subspace of M_{2x2}, where

   $M_{2x2} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a,b,c,d \in \Re \right\}$?

UNIT IV

Determinants

4.1 Definition of a Determinant

We interrupt our discussion of matrices to introduce the concept of the *determinant function*. Remember that a matrix is simply an ordered arrangement of elements; it is meaningless to assign a single numerical value to a matrix.

However, if A is a square matrix, then the determinant function associates with A exactly one numerical value called the **determinant of A,** that gives us valuable information about the matrix. By denoting the determinant of A by $|\boldsymbol{A}|$ or **det A** we can think of the *determinant function* as correspondence:

$$\underset{square\ matrix}{A} \quad \rightarrow \quad \underset{determinant\ of\ A}{|A|}$$

In this case, the straight bars do NOT mean absolute value; they represent the determinant of the matrix. We will see some of the uses of the determinant in the subsequent sections. For now, let's find out how to compute the determinant of a matrix so that we can use it later.

Definition 4.1.1: (Determinant of order 1): Let $A = [a_{11}]$ be a square matrix of order 1. Then determinant of A is defined as the number $\boldsymbol{a_{11}}$ itself. That is, $|\boldsymbol{a_{11}}| = \boldsymbol{a_{11}}$.

Example $|3| = 3$, $|-5| = -5$ and $|0| = 0$

Definition 4.1.2: (Determinant of order 2): Let $\begin{bmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} \end{bmatrix}$ be a 2×2 matrix, then

$$|\boldsymbol{A}| = \boldsymbol{a_{11}a_{22}} - \boldsymbol{a_{12}a_{21}}.$$

That is, the determinant of a 2×2 matrix is obtained by taking the product of the entries in the main diagonal and subtracting from it the product of the entries in the other diagonal.

To define the determinant of a square matrix A of order n(n > 2), we need the concepts of the minor and the cofactor of an element.

Let $|\boldsymbol{A}| = |\boldsymbol{a_{ij}}|$ be a determinant of order n. The **minor of a_{ij},** is the determinant that is left by deleting the *i*th row and the *j*th column. It is denoted by $\boldsymbol{M_{ij}}$**.**

For example, given the 3 x 3 determinant $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$. The minor of a_{11} is

$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$, the minor of a_{12} is $M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$, and so on.

Let $|\boldsymbol{A}| = |\boldsymbol{a_{ij}}|$ be a determinant of order n. The **cofactor of a_{ij}** denoted $\mathbf{C}_{ij}$ **or** $\mathbf{A}_{ij}$**,** is

defined as $(-1)^{i+j} M_{ij}$, where $i + j$ is the sum of the row number i and column number j in which the entry lies. Thus $C_{ij} = \begin{cases} M_{ij}, \text{ if } i+j \text{ is even} \\ -M_{ij}, \text{ if } i+j \text{ is odd} \end{cases}$. For example, the cofactor of a_{12} in the 3 x 3 determinant $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ is

$$C_{12} = (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

Example 4.1.1: Evaluate the cofactor of each of the entries of the matrix: $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{pmatrix}$

Solution: $C_{11} = -1$, $C_{21} = 1$, $C_{31} = -1$, $C_{12} = 8$, $C_{13} = -5$, $C_{22} = -6$, $C_{32} = 2$, $C_{23} = 3$, $C_{33} = -1$

Activity 4.1.1: Evaluate the cofactor of each of the entries of the given matrices:

a. $\begin{pmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$ **b.** $\begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$

Definition 4.1.3 :(Determinant of order n): If A is a square matrix of order n (n >2), then its determinant may be calculated by multiplying the entries of any row (or column) by their cofactors and summing the resulting products. That is,

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in} \quad \textbf{Or} \quad \det A = a_{1j}C_{1j} + a_{2j}C_{2j} + ... + a_{nj}C_{nj}$$

Remark: It is a fact that determinant of a matrix is unique and does not depend on the row or column chosen for its evaluation.

Example 4.1.2: Find the value of $\begin{vmatrix} 1 & -3 & 4 \\ 0 & 2 & 5 \\ -2 & 6 & 3 \end{vmatrix}$

Solution: Choose a given row or column. Let us arbitrarily select the first row. Then

$$\begin{vmatrix} 1 & -3 & 4 \\ 0 & 2 & 5 \\ -2 & 6 & 3 \end{vmatrix} = (1)\begin{vmatrix} 2 & 5 \\ 6 & 3 \end{vmatrix} + (-3)(-1)\begin{vmatrix} 0 & 5 \\ -2 & 3 \end{vmatrix} + 4\begin{vmatrix} 0 & 2 \\ -2 & 6 \end{vmatrix} = 1(6-30) + 3(0+10) + 4(0+4)$$

$$= 22$$

If we had expanded along the first column, then

$$\begin{vmatrix} 1 & -3 & 4 \\ 0 & 2 & 5 \\ -2 & 6 & 3 \end{vmatrix} = (1)\begin{vmatrix} 2 & 5 \\ 6 & 3 \end{vmatrix} + 0 + (-2)\begin{vmatrix} -3 & 4 \\ 2 & 5 \end{vmatrix} = 1(6-30) - 2(-15-8) = 22, \text{ as before}$$

Example 4.1.3: Find the value of $|A| = \begin{vmatrix} 1 & 2 & 0 & -1 \\ 3 & -1 & 4 & 1 \\ -2 & 0 & -3 & 3 \\ 4 & 3 & 1 & 2 \end{vmatrix}$

Solution: Expanding along first row, we have

$$|A| = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + a_{14}C_{14} = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} - a_{14}M_{14}$$

$$= (1)\begin{vmatrix} -1 & 4 & 1 \\ 0 & -3 & 3 \\ 3 & 1 & 2 \end{vmatrix} - 2\begin{vmatrix} 3 & 4 & 1 \\ -2 & -3 & 3 \\ 4 & 1 & 2 \end{vmatrix} + 0 + (1)\begin{vmatrix} 3 & -1 & 4 \\ -2 & 0 & -3 \\ 4 & 3 & 1 \end{vmatrix} = \quad 54 - 94 + 13 = -27$$

Activity 4.1.2: Compute the determinant of A if:

a. $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{pmatrix}$

c. $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$

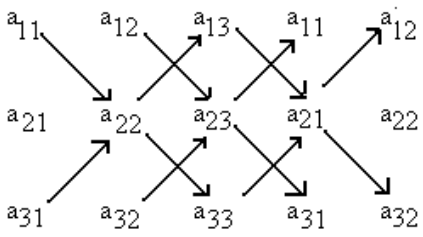
b. $A = \begin{pmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{pmatrix}$

d. $A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 5 & -2 & 0 & 0 \\ -8 & 3 & 1 & 0 \\ 4 & -7 & 5 & 2 \end{pmatrix}$

Note: 1. $\det I_n = 1$**,** where I_n is an identity matrix of order n.

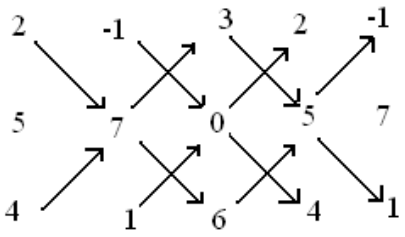
2. $\det A =$ The product of the diagonal elements, if A is a diagonal matrix or lower triangular matrix or upper triangular matrix.

The following diagram called Sarrus' diagram, enables us to write the value of the determinant of order 3 very conveniently. This technique does not hold for determinant of higher order.



Working Rule: Make the Sarrus' diagram by repeating the first two columns of the determinant as shown below. Then multiply the elements joined by arrows. Assign the positive sign to an expression if it is formed by a downward arrow and negative sign to an expression if it is formed by an upward arrow.

Value: $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$



Example 4.1.4: Find the value of $|\boldsymbol{A}| = \begin{vmatrix} 2 & -1 & 3 \\ 5 & 7 & 0 \\ 4 & 1 & 6 \end{vmatrix}$ with the help of Sarrus' diagram

Solution: The Sarrus' diagram for the given determinant is to the right. Thus the value of the determinant is

$$|A| = (2)(7)(6) + (-1)(0)(4) + (3)(5)(1) - (4)(7)(3) - (1)(0)(2) - (6)(5)(-1) = 45$$

Exercise:

1) Evaluate the following determinants: a) $\begin{vmatrix} 3 & 2 \\ -5 & -4 \end{vmatrix}$ b) $\begin{vmatrix} -2 & -\boldsymbol{a} \\ -\boldsymbol{a} & 2 \end{vmatrix}$ c) $\begin{vmatrix} 2 & 1 & 5 \\ -3 & 4 & -1 \\ 0 & 6 & -1 \end{vmatrix}$

2) Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$. Determine each of the following

a) the minor of a_{21} b) the minor of a_{22} c) the cofactor of a_{22}

d) the cofactor of a_{23} e) the cofactor of a_{32}.

4.2. Properties of Determinants

We now state some useful properties of determinants. These properties help a good deal in the evaluation of determinants. We use the notations R_i and C_j to denote respectively the *i-th* row and the j-*th* column of a determinant.

Property 1: The value of a determinant remains unchanged if rows are changed into columns and columns into rows. That is,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \quad \text{or} \quad \det A = \det A^t$$

Remark: In terms of matrices, if A is a square matrix, then $|\boldsymbol{A}| = |\boldsymbol{A}'|$.

Property 2: If any two rows (or columns) of a matrix are interchanged, the value of the determinant so obtained is the negative of the value of the original determinant. That is,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}.$$

Remark: The notation $\boldsymbol{R_i \leftrightarrow R_j (C_i \leftrightarrow C_j)}$ is used to represent interchange of *ith* and *jth* row (column).

Property 3: If any two rows (or columns) of a matrix are identical, the value of the determinant is zero. That is, $\begin{vmatrix} \boldsymbol{a_1} & \boldsymbol{b_1} & \boldsymbol{c_1} \\ \boldsymbol{a_1} & \boldsymbol{b_1} & \boldsymbol{c_1} \\ \boldsymbol{a_2} & \boldsymbol{b_2} & \boldsymbol{c_2} \end{vmatrix} = 0$. R_1 and R_2 are identical

Property 4: If each element of a row (or column) of a matrix is multiplied by a constant k, the value of the determinant so obtained is k times the value of the original determinant. That is, $\begin{vmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \boldsymbol{a_{13}} \\ \boldsymbol{ka_{21}} & \boldsymbol{ka_{22}} & \boldsymbol{ka_{23}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{32}} & \boldsymbol{a_{33}} \end{vmatrix} = \boldsymbol{k} \begin{vmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \boldsymbol{a_{13}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{32}} & \boldsymbol{a_{33}} \end{vmatrix}$

Remark: 1. The notation $\boldsymbol{R_i \to k\ R_i (C_i \to k C_i)}$ is used to represent multiplication of each element of ith row (column) by the constant k.

2. $\det(kA) = k^n \det A$**,** where k is any real number and A is an nxn matrix**.**

Property 5: If the elements of a row (or column) of a determinant are added k times the elements of another row (or column), the value of the determinant so obtained is equal to the value of the original determinant. That is,

$$\begin{vmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \boldsymbol{a_{13}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{32}} & \boldsymbol{a_{33}} \end{vmatrix} = \begin{vmatrix} \boldsymbol{a_{11} + ka_{31}} & \boldsymbol{a_{12} + ka_{32}} & \boldsymbol{a_{13} + ka_{33}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{32}} & \boldsymbol{a_{33}} \end{vmatrix}$$

Property 6: If each element of a row (or column) of a determinant is the sum of two elements, the determinant can be expressed as the sum of two determinants. That is,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31}+b_1 & a_{32}+b_2 & a_{33}+b_3 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Example 4.2.1: Find the value of the determinant $|A| = \begin{vmatrix} 1 & 18 & 72 \\ 2 & 40 & 148 \\ 3 & 45 & 150 \end{vmatrix}$

Solution: By applying various properties of determinants, we make maximum number of zeros in a row or a column. We shall make maximum number of zeros in C_1. Performing the operations $R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - 2R_1$ (property 5),

$$|A| = \begin{vmatrix} 1 & 18 & 72 \\ 0 & 4 & 4 \\ 0 & 9 & 6 \end{vmatrix} = \begin{vmatrix} 4 & 4 \\ 9 & 6 \end{vmatrix} = 24 - 36 = -12$$

Example 4.2.2: Show that $\begin{vmatrix} y+z & x & y \\ z+x & z & x \\ x+y & y & z \end{vmatrix} = (x+y+z)(x-z)^2$

Solution: Let $\Delta = \begin{vmatrix} y+z & x & y \\ z+x & z & x \\ x+y & y & z \end{vmatrix}$. Performing $R_1 \to R_1 + R_2$, $R_1 \to R_1 + R_3$, we get

$$\Delta = \begin{vmatrix} 2(x+y+z) & x+y+z & x+y+z \\ z+x & z & x \\ x+y & y & z \end{vmatrix} = (x+y+z) \begin{vmatrix} 2 & 1 & 1 \\ z+x & z & x \\ x+y & y & z \end{vmatrix}$$

Performing $C_1 \to C_1 - 2C_2$ and $C_2 \to C_2 - C_3$, we get

$$= \Delta = (x+y+z) \begin{vmatrix} 0 & 0 & 1 \\ x-z & z-x & x \\ x-y & y-z & z \end{vmatrix} \qquad \text{(Property 5)}$$

$$= (x+y+z)(x-z)^2$$

Activity 4.2.1: Evaluate the following determinants by using the properties listed above:

a) $\begin{vmatrix} 3 & 1 & 43 \\ 2 & 7 & 35 \\ 1 & 3 & 17 \end{vmatrix}$ b) $\begin{vmatrix} 2 & 4 & 6 \\ 7 & 9 & 11 \\ 8 & 10 & 12 \end{vmatrix}$ c) $\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix}$

Product of two determinants

Theorem 4.2.1: The determinant of the product of two matrices of order n is the product of their determinants. That is, $|\boldsymbol{AB}| = |\boldsymbol{A}|\,|\boldsymbol{B}|$.

Example 4.2.3: Let $\boldsymbol{A} = \begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix}$, then

$$|AB| = |A||B| = \begin{vmatrix} 3 & 0 \\ 4 & 1 \end{vmatrix} \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} = (3)(3) = 9$$

Example 4.2.4: Let A and B be 3x3 matrix with det A = 2 and det B = -3.

Find det ($2AB^t$).

Solution: $\det(2AB^t) = 2^3 \det A \det B^t = 8(2)(-3) = -48$, since det B = det B^t.

4.3 Adjoint and Inverse of a matrix

Definition 4.3.1: Let A = (a_{ij}) be a square matrix of order n and let C_{ij} be the cofactor of a_{ij}. Then the **adjoint of A**, denoted by **adj A**, is defined as the transpose of the cofactor matrix (C_{ij}).

Example 4.3.1: Find adj A, if $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 4 & 3 & 2 \end{bmatrix}$

Solution: We have C_{11}=-3, C_{12}=6, C_{13}= -3, C_{21}=5, C_{22}=-10, C_{23}=5, C_{31}=2,C_{32}=-4, C_{33}=2.

Thus, $adj\ A = \begin{bmatrix} -3 & 5 & 2 \\ 6 & -10 & -4 \\ -3 & 5 & 2 \end{bmatrix}$.

Activity 4.3.1: Find adj A if $A=\begin{bmatrix}1 & 5 & 0\\ 2 & 4 & -1\\ 0 & -2 & 0\end{bmatrix}$

Properties of the Adjoint of a matrix

1. If A is a square matrix of order n, then
   A(adj A) = |A| $\mathbf{I_n}$ = (adj A)A, where $\mathbf{I_n}$ is an identity matrix of order n.
2. If A is a square matrix of order n, then $\boldsymbol{adj\ (A')=(adjA)'}$
3. If A and B are two square matrices of the same order, then
   adj(AB) = adj(B) adj (A).

Example 4.3.2: If $\boldsymbol{A}=\begin{bmatrix}2 & 1 & 3\\ 2 & 0 & 1\\ -4 & 5 & 6\end{bmatrix}$, verify that A(adjA) = |A| I_3 = (adjA)A

Solution: We have |A| = 2(-5)-1(12+4)+3(10) = -10 – 16 + 30 = 4

Now C_{11}=-5, C_{12}=-16, C_{13}=10, C_{21} = 9, C_{22}=24, C_{23}=-14, C_{31}=1, C_{32}=4, C_{33}=-2.

Therefore, $adj\ A=\begin{bmatrix}-5 & 9 & 1\\ -16 & 24 & 4\\ 10 & -14 & -2\end{bmatrix}$

Hence $\boldsymbol{A(adjA)}=\begin{bmatrix}2 & 1 & 3\\ 2 & 0 & 1\\ -4 & 5 & 6\end{bmatrix}\begin{bmatrix}-5 & 9 & 1\\ -16 & 24 & 4\\ 10 & -14 & -2\end{bmatrix}=\begin{bmatrix}4 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 4\end{bmatrix}=4\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=|\boldsymbol{A}|\boldsymbol{I}_3$

Similarly, it can be proved that $\boldsymbol{(adjA)A}=|\boldsymbol{A}|\boldsymbol{I}_3$

Definition 4.3.2: Let A be a square matrix of order n. Then a square matrices B of order n, if it exists, is called an **inverse of A** if **AB = BA = $\mathbf{I_n}$**. A matrix A having an inverse is called **an invertible matrix**. It may easily be seen that if a matrix A is invertible, its inverse is **unique**. The inverse of an invertible matrix A is denoted by $\mathbf{A^{-1}}$.

Does every square matrix possess an inverse? To answer this let us consider the matrix $\boldsymbol{A}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$. If B is any square matrix of order 2, we find that **AB = BA = 0.**

We thus see that there cannot be any matrix **B** for which **AB** and **BA** both are equal to $\mathbf{I_2}$**.** Therefore A is not invertible. Hence, we conclude that a square matrix may fail to have an inverse. However, if A is a square matrix such that $|A| \neq 0$, then A is invertible and

$A^{-1} = \frac{1}{|A|} adj\ A$. For, we know that $A(adjA) = (adiA)A = |A|\ I_n$

$A\left(\frac{1}{|A|} adjA\right) = \left(\frac{1}{|A|} adjA\right)A = I_n$. Thus A is invertible and $A^{-1} = \frac{1}{|A|} adjA$.

A square matrix A is said to be **singular (not invertible)** if $|A| = 0$, and it is called **non-singular (invertible)** if $|A| \neq 0$.

Example 4.3.3: Find λ if the matrix $A = \begin{bmatrix} 6 & 7 & -1 \\ 3 & \lambda & 5 \\ 9 & 11 & \lambda \end{bmatrix}$ has no inverse.

Solution: $6(\lambda^2 - 55) - 7(3\lambda - 45) - (33 - 9\lambda) = 0$

- $\lambda^2 - 2\lambda - 8 = 0$
- $(\lambda - 2)(\lambda - 4) = 0$
- $\lambda = 2$ or $\lambda = 4$

Example 4.3.4: If $A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{bmatrix}$, then $|A| = 3(-3+2) - 1(2+1) + 2(4+3) = 8$

Since $|A| \neq 0, A$ is non-singular or invertible.

Activity 4.3.2: Find $\mathbf{A^{-1}}$.

Further, if $B = \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 4 \\ 5 & -2 & 8 \end{bmatrix}$ then $|B| = 0$ and it is singular.

Note: If A is an invertible nxn matrix, then $AA^{-1} = I_n$ and $\det A^{-1} = \frac{1}{\det A}$, where $\det A \neq 0$.

Properties of the inverse of a matrix

1. A square matrix is invertible if and only if it is non-singular.
2. The inverse of the inverse is the original matrix itself, i.e. $\left(\boldsymbol{A^{-1}}\right)^{-1} = \boldsymbol{A}$
3. The inverse of the transpose of a matrix is the transpose of its inverse, i.e.,
$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$$
4. If A and B are two invertible matrices of the same order, then AB is also invertible and moreover, $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

Example 4.3.5: Find the inverse of the matrix $A=\begin{bmatrix} 4 & -2 & 1 \\ 7 & 3 & 3 \\ 2 & 0 & 1 \end{bmatrix}$.

Solution: $|A| = (4)(3) - (-2)(1) + 1(-6) = 8 \neq 0$. Thus $\boldsymbol{A^{-1}}$ exists and is given by $\boldsymbol{A^{-1}} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{adjA}$. To find *adjA*, let C_{ij} denote the cofactor of a_{ij}, the element in the *ith* row and *jth* column of |A|. Thus C_{11}=3, C_{12} = -1, C_{13}=-6, C_{21} = 2, C_{22}=2, C_{23}=-4, C_{31}=-9 C_{32}=-5 and C_{33} =26.

$adj\ A = \begin{bmatrix} 3 & 2 & -9 \\ -1 & 2 & -5 \\ -6 & -4 & 16 \end{bmatrix}$. Hence $\boldsymbol{A^{-1}} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{adj}\ \boldsymbol{A} = \frac{1}{8}\begin{bmatrix} 3 & 2 & -9 \\ -1 & 2 & -5 \\ -6 & -4 & 16 \end{bmatrix}$

Activity 4.3.3: 1. Find the inverse of A, if i) $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ ii) $\boldsymbol{A} = \begin{bmatrix} 1 & 4 & 0 \\ -1 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

2. Find matrix A such that $\boldsymbol{A}\begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 9 & 4 \end{bmatrix}$.

3. If $AX = b$ then $X = A^{-1}b$. (True/False)

4.4. Cramer's rule for solving system of linear equations (homogeneous and non homogeneous)

Suppose we have to solve a system of n linear equations in n unknowns **Ax** = **b.** Let $\mathbf{A_i(b)}$ be the matrix obtained from **A** by replacing column **i** by the vector **b** and $\mathbf{A^k}$ be the **k**-th column vector of matrix **A**.

$$A_i(b) = [\mathbf{A^1}\ \mathbf{A^2}\ \ldots\ \mathbf{b}\ \ldots\ \mathbf{A^n}]$$

column i

Now let $\mathbf{e_1}, \mathbf{e_2}, \ldots \mathbf{e_n}$ be columns of the n × n identity matrix **I** and $\mathbf{I_i(x)}$ be the matrix obtained from **I** by replacing column **i** by **x.**

If **Ax** = **b** then by using matrix multiplication we have

$$\mathbf{AI_i(x) = A[e_1\ \ldots\ x\ \ldots\ e_n] = [Ae_1\ \ldots\ Ax\ \ldots\ Ae_n]}$$
$$\mathbf{= [A^1\ \ldots\ b\ \ldots\ A^n] = A_i(b)}$$

By the multiplicative property of determinants, (det**A**)(det$\mathbf{I_i(x))}$ = det $\mathbf{A_i(b)}$

The second determinant on the left is x_i. (Make a cofactor expansion along the ith row.)

Hence (det **A).** x_i = det $\mathbf{A_i(b).}$ Therefore if det$\mathbf{A \neq 0}$ then we have $\mathbf{x_i} = \dfrac{\mathbf{det\, A_i(b)}}{\mathbf{det\, A}}$.

This method for finding the solutions of n linear equations in n unknowns is known as Cramer's Rule.

Example 4.4.1: Solve the following system of linear equations by Cramer's Rule.

$$2\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 = 6$$
$$\mathbf{x}_1 + 4\mathbf{x}_2 - 2\mathbf{x}_3 = -4$$
$$3\mathbf{x}_1 + \mathbf{x}_3 = 7$$

Solution: Matrix form of the given system is **Ax** = **b**

where $\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 4 & -2 \\ 3 & 0 & 1 \end{pmatrix}$ $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{pmatrix}$ **and** $\mathbf{b} = \begin{pmatrix} 6 \\ -4 \\ 7 \end{pmatrix}$

By Cramer's Rule, $\mathbf{x_i} = \dfrac{\mathbf{\det A_i(b)}}{\mathbf{\det A}}$ **(i** $= 1, 2, 3$**)**

$$\det A = \begin{vmatrix} 2 & -1 & 1 \\ 1 & 4 & -2 \\ 3 & 0 & 1 \end{vmatrix} = 3$$

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{\begin{vmatrix} 6 & -1 & 1 \\ -4 & 4 & -2 \\ 7 & 0 & 1 \end{vmatrix}}{2} = \frac{6}{3} = 2 \quad , \; x_2 = \frac{\det A_2(b)}{\det A} = \frac{\begin{vmatrix} 2 & 6 & 1 \\ 1 & -4 & -2 \\ 3 & 7 & 1 \end{vmatrix}}{2} = \frac{-3}{3} = -1$$

and $x_3 = \dfrac{\det A_3(b)}{\det A} = \dfrac{\begin{vmatrix} 2 & -1 & 6 \\ 1 & 4 & -4 \\ 3 & 0 & 7 \end{vmatrix}}{2} = \dfrac{3}{3} = 1.$

Example 4.4.2: Solve the following system of linear equations by Cramer's Rule.

$$\begin{aligned} 2x_1 + x_2 &= 7 \\ -3x_1 + 2x_3 &= -8 \\ x_2 + 2x_3 &= -3 \end{aligned}$$

Solution: Matrix form of the given system is $\mathbf{Ax} = \mathbf{b}$

where $A = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ *and* $b = \begin{pmatrix} 7 \\ -8 \\ -3 \end{pmatrix}$

By Cramer's Rule, $\mathbf{x_i} = \dfrac{\mathbf{\det A_i(b)}}{\mathbf{\det A}}$ **(i** $= 1, 2, 3$**)**

$$\det A = \begin{vmatrix} 2 & 1 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 4$$

$$x_1 = \frac{\det A_1(b)}{\det A} = \frac{\begin{vmatrix} 7 & 1 & 0 \\ -8 & 0 & 1 \\ -3 & 1 & 2 \end{vmatrix}}{4} = \frac{6}{4} = \frac{3}{2} \quad , \quad x_2 = \frac{\det A_2(b)}{\det A} = \frac{\begin{vmatrix} 2 & 7 & 0 \\ -3 & -8 & 1 \\ 0 & -3 & 2 \end{vmatrix}}{4} = \frac{16}{4} = 4 \quad \text{and}$$

$$x_3 = \frac{\det A_3(b)}{\det A} = \frac{\begin{vmatrix} 2 & 1 & 7 \\ -3 & 0 & -8 \\ 0 & 1 & -3 \end{vmatrix}}{4} = \frac{-14}{4} = \frac{-7}{2}.$$

Activity 4.4.1: 1. Use Cramer's rule to solve each of the following

a) $2z + 3 = y + 3x$
$x - 3z = 2y + 1$
$3y + z = 2-2x$

b) $-a + 3b - 2c = 7$
$3a + 3c = -3$
$2a + b + 2c = -1$

2. Consider the system of homogeneous n linear equations in n unknowns **Ax** = **0.** Discuss about the cases that we can use Cramer's Rule.

For the non homogeneous system $\mathbf{Ax} = \mathbf{b}$, if $\det A = 0$, then the Cramer's rule does not give any information whether or not the system has a solution. However, in the case of homogeneous system we have the following useful theorem.

Theorem 4.4.1: A system of n homogenous linear equations in n unknowns **Ax = 0**, has a non trivial solution if $\det A = 0$. If $\det A \neq 0$, it has only the trivial solution $x_1 = x_2 = \ldots = x_n = 0$.

Example 4.4.3: Let $A = \begin{pmatrix} 6 & 7 & -1 \\ 3 & \lambda & 5 \\ 9 & 11 & \lambda \end{pmatrix}$. Find the value(s) of λ if $Ax = 0$ has non-zero solution.

Solution: det A = 0

- $\begin{vmatrix} 6 & 7 & -1 \\ 3 & \lambda & 5 \\ 9 & 11 & \lambda \end{vmatrix} = 0$
- $6(\lambda^2 - 55) - 7(3\lambda - 45) - (33 - 9\lambda) = 0$
- $\lambda^2 - 2\lambda - 8 = 0$
- $(\lambda - 2)(\lambda - 4) = 0$ or $\lambda = 2, \lambda = 4$

4.5 The rank of a matrix by subdeterminants

Let us consider a square matrix A of order n (i.e n rows and n columns). If there are at least k rows and k columns which must be deleted in order to obtain a non vanishing determinant, then the order of the highest ordered non vanishing determinant in A is given by r = n – k and this number is defined as the rank of A.

We are now giving the definition of rank, in terms of determinant.

Definition 4.5.1: The rank of a matrix A, r(A), is equal to the order of the highest ordered non vanishing determinant in A.

Remark: The rank of a matrix is the order of the largest sub square matrix whose determinant is different from zero.

Note: If A is a non singular matrix of order n, then the rank of A is equal to n.

1. Rank of A is unique
2. Every matrix has a rank
3. If A is a matrix of order m × n, then r(A) ≤ minimum of $\{m, n\}$
4. If r(A) = k then every minor of order k + 1, k + 2 etc is zero
5. A is a matrix of order n × n. A is non singular ⇔ r (A) = n
6. Rank of I_n = n
7. A is a matrix of order m × n. If every k^{th} order minor (k < m, k < n) is zero then r(A) < k.

8. A is matrix of order m × n. If there is a minor of order k (k < m, k < n) which is not zero, then r(A) ≥ k.
9. If A is a null matrix, then the rank of A is defined as zero. That is, $r(A) \geq 1$ if A is non-zero matrix.

Example 4.5.1: Obtain the rank of the matrix $A=\begin{pmatrix} -1 & 2 & 0 \\ 3 & 7 & 1 \\ 5 & 9 & 3 \end{pmatrix}_{3\times 3}$

Solution: det A = -1 (21 – 9) - 2 (9 – 5) + 0 = -20 ≠ 0

∴ A is non singular and hence r (A) = 3

Example 4.5.2: Obtain the rank of the matrix $A=\begin{pmatrix} -1 & -2 & -3 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix}$

Solution: | A | = -1(24 – 25) + 2(18 – 20) – 3(15 – 16) = 0. So R (A) < 3 … (1)

A minor of order 2 of A is $\begin{vmatrix} -1 & -2 \\ 3 & 4 \end{vmatrix} = 2 \neq 0$. So R(A) ≥ 2 ……(2)

∴ From (1) and (2); rank (A) = 2.

Example 4.5.3: Find the rank of the matrix $A = \begin{pmatrix} -1 & 2 & 3 \\ -2 & 4 & 6 \end{pmatrix}_{2\times 3}$

Solution: Since A is a 2x3 matrix, rank of A ≤ 2. But every determinant of order 2 in A is zero. So rank of A < 2 …..(1)

But A is a non zero matrix ⇒ r(A) ≥ 1 …. (2)

∴ from (1) and (2) r(A) = 1

Example 4.5.4: Find the rank of the matrix $A=\begin{pmatrix} 1 & 0 & -4 & 5 \\ 2 & -1 & 3 & 0 \\ 8 & 1 & 0 & -7 \end{pmatrix}$.

Solution: A is a 3x4 matrix. So rank of A ≤ 3 ……(1).

A non vanishing determinant of order 3 in A is

$\begin{vmatrix} 1 & 0 & -4 \\ 2 & -1 & 3 \\ 8 & 1 & 0 \end{vmatrix}$ = -43. So rank of A ≥ 3 ……(2).

Combining (1) and (2) we have rank of A = 3.

Example 4.5.5: The 3×4 matrix A = $\begin{pmatrix} 1 & -1 & 1 & 2 \\ 2 & -2 & 2 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$ has a row that is a constant multiple of another row (i.e $R_2 = 2R_1$). This matrix possesses four square sub matrices order 3: $\begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ 1 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 1 & 3 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 1 & 2 \\ -2 & 2 & 4 \\ 3 & 3 & 4 \end{pmatrix}$.

The determinant of each of these matrices is zero. Because, in each case the second row is a constant multiple of the first row. Thus the rank of the matrix A cannot be equal to 3. That is rank of A < 3. However, it is easy to find a 2 x 2 sub matrix of A whose determinant is different from zero. Take $\begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix}$ its determinant is equal to -4 ≠ 0.

This indicates that rank(A) = 2.

4.6. Determinant and volume

This is related to the idea of triple dot product.

If A is a 2x2 matrix, the **area of the parallelogram** determined by the columns of A is **det A**.

If A is a 3x3 matrix, the **volume of the parallelepiped** determined by the columns of A is **det A**.

4.7. Eigenvalues and Eigenvectors of a matrix

Definition 4.7.1: Let A be an n x n matrix over a field K. A scalar λ in K is said to be an **eigenvalue** of A iff there is a non-zero vector X in K^n such that

$$AX = \lambda X \text{ …… } (*)$$

If λ is an eigenvalue of A then any vector v in K^n satisfying (*) is called an **eigenvector** of A corresponding to λ. (In our case K is the set of all real numbers)

Note: Intuitively, an eigenvector of A is a vector that remains unchanged or reversed in direction when A operates on the vector space. If the corresponding value is 1, then it is unchanged in magnitude also and it then said to be **invarian**t under A.

Activity 4.7.1: Let $A=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}, u=\begin{bmatrix}6\\ -5\end{bmatrix}, and\ v=\begin{bmatrix}3\\ -2\end{bmatrix}$.

i) Are u and v eigenvectors of A?

ii) Show that 7 is an eigenvalue of A.

How to determine the Eigenvalues and corresponding Eigenvectors of a Matrix?

X is an eigenvector with eigenvalue $\lambda \Leftrightarrow AX = \lambda X \Leftrightarrow (A - \lambda I_n) X = 0$ …(**)

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

X = 0 is the trivial solution of (**). Further solutions will exist iff $\left|A-\lambda I_n\right| = 0$.

Hence, solving the equation $\left|A-\lambda I_n\right| = 0$ gives the eigenvalue(s) of A.

For each eigenvalue λ, the corresponding eigenvector is found by substituting λ back into the equation $\left|A-\lambda I_n\right| X = 0$.

Note: i) The polynomial $\left|A-\lambda I_n\right|$ is called the **characteristic polynomial of A.**

ii) The equation $\left|A-\lambda I_n\right| = 0$ is called the **characteristic equation**.

Example 4.7.1: Let $A=\begin{bmatrix}1 & 6\\ 5 & 2\end{bmatrix}$. Find the eigenvalues and the corresponding eignvectors of A.

Solution: $$|A - \lambda I_2| = 0 \Leftrightarrow \left| \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$$\Leftrightarrow \begin{vmatrix} 1-\lambda & 6 \\ 5 & 2-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^2 - 3\lambda - 28 = 0$$

$$\Leftrightarrow \lambda = 7 \text{ or } \lambda = -4$$

The corresponding eigenvectors can now be found as follow:

For $\lambda = 7$: $(A - 7I_2)X = 0 \Leftrightarrow \left[\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 7 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Leftrightarrow \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow y = x$$

Hence, any vector of the type $\beta \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, where β is any real number, is an eigenvector corresponding to the eigenvalue 7.

For $\lambda = -4$: $(A + 4I_2)X = 0 \Leftrightarrow \left[\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} + 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Leftrightarrow \begin{bmatrix} 5 & 6 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow y = \frac{-5}{6}x$$

Hence, any vector of the type $\beta \begin{bmatrix} 1 \\ \frac{-5}{6} \end{bmatrix}$, where β is any real number, is an eigenvector corresponding to the eigenvalue –4.

Note: If X is an eigenvector with eigenvalue λ, αX is also an eigenvector with the same eigenvalue, where α is a non-zero scalar.

The following theorem summarizes our results so far.

Theorem 4.7.1: If A is an n x n matrix, then the following statements are equivalent:

i) λ is an eigenvalue of A.

ii) There is a non-zero vector $X \in K^n$ such that $AX = \lambda X$.

iii) The system of equations $(A - \lambda I)X = 0$ has non-trivial solutions.

iv) λ is a solution of the characteristic equation $\det(A - \lambda I)$ in K.

Example 4.7.2: Let $A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. Find the eigenvalues and the corresponding eignvectors of A.

Solution: Characteristic equation of A: $\begin{vmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} = 0$

$\Leftrightarrow \quad (3 - \lambda)\ (3 - \lambda)\ (5 - \lambda) - 4(5 - \lambda) = 0$

$\Leftrightarrow \quad [(3 - \lambda)^2 - 4]\ (5 - \lambda) = 0$

$\Leftrightarrow \quad (\lambda^2 - 6\lambda + 5)\ (\lambda - 5) = 0$

$\Leftrightarrow \quad (\lambda - 1)\ (\lambda - 5)^2 = 0$

So, eigenvalues of A are: $\lambda = 1$ and $\lambda = 5$.

To find the corresponding eigen vectors, we substitute the values of λ in the equation

$(A - \lambda I)\, X = 0$. That is, $\begin{bmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ ….(*)

For $\lambda = 1$, (*) becomes: $\begin{bmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow x = y, \quad z = 0$

Thus, the eigenvectors corresponding to eigenvalue 1 are vectors of the form:

$$X = \begin{bmatrix} x \\ x \\ 0 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$ for any x in the set of real numbers.

For $\lambda = 5$, (*) becomes: $\begin{bmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow x = -y$

Thus, eigenvectors of A corresponding to eigenvalues 5 are vectors of the form.

$$X = \begin{bmatrix} x \\ -x \\ z \end{bmatrix} = \begin{bmatrix} x \\ -x \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Exercise:

Find the eigenvalues and the corresponding eigenvectors of the matrices:

a) $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$ b) $A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix}$ d) $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

4.8. Diagonalization of symmetric matrix

Theorem 4.8.1: If A is an $n \times n$ symmetric matrix, then the eigenvectors of A associated with distinct eigenvalues are orthogonal.

Proof:

Let $x_1 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $x_2 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ be eigenvectors of ***A*** associated with distinct eigenvalues λ_1 and λ_2, respectively, i.e.,

$$Ax_1 = \lambda_1 x_1 \text{ and } Ax_2 = \lambda_2 x_2 .$$

Thus,

$$x_1^t Ax_2 = x_1^t \left(Ax_2\right) = x_1^t \lambda_2 x_2 = \lambda_2 x_1^t x_2$$

and

$$x_1^t Ax_2 = x_1^t A^t x_2 = \left(x_1^t A^t\right)x_2 = \left(Ax_1\right)^t x_2 = \left(\lambda_1 x_1\right)^t x_2 = \lambda_1 x_1^t x_2 .$$

Therefore,

$$x_1^t Ax_2 = \lambda_2 x_1^t x_2 = \lambda_1 x_1^t x_2 .$$

Since $\lambda_1 \neq \lambda_2$, $\boldsymbol{x_1^t x_2 = O}$.

Example 4.8.1: Let $A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix}$.

A is a symmetric matrix. The characteristic equation is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & 2 \\ 0 & \lambda + 2 & 0 \\ 2 & 0 & \lambda - 3 \end{vmatrix} = (\lambda + 2)(\lambda - 4)(\lambda + 1) = 0 .$$

The eigenvalues of ***A*** are $-2, 4, -1$. The eigenvectors associated with these eigenvalues are

$$x_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} (\lambda = 2),\ x_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} (\lambda = 4),\ x_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} (\lambda = -1).$$

Thus, x_1, x_2, x_3 **are orthogonal.**

Note: If A is an $n \times n$ symmetric matrix, then there exists an orthogonal matrix P such that

$$\boldsymbol{D = P^{-1}AP = P^t AP},$$

Where $col_1(P), col_2(P), \ldots, col_n(P)$ are n linearly independent eigenvectors of A and the diagonal elements of D are the eigenvalues of A associated with these eigenvectors.

Example 4.8.2: Let $A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$.

Please find an orthogonal matrix ***P*** and a diagonal matrix ***D*** such that $D = P^t AP$.

Solution: We need to find the orthonormal eigenvectors of ***A*** and the associated eigenvalues first. The characteristic equation is

$$f(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda & -2 & -2 \\ -2 & \lambda & -2 \\ -2 & -2 & \lambda \end{vmatrix} = (\lambda + 2)^2 (\lambda - 4) = 0.$$

Thus, $\lambda = -2, -2, 4.$

1. As $\lambda = -2,$ solve for the homogeneous system

$$(-2I - A)x = 0.$$

The eigenvectors are

$$t\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, t, s \in R,\ t \neq 0 \text{ or } s \neq 0.$$

$\Rightarrow v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ are two eigenvectors of ***A***. However, the two eigenvectors are not orthogonal. We can obtain two orthonormal eigenvectors via ***Gram-Schmidt process***. The orthogonal eigenvectors are

$$v_1^* = v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$v_2^* = v_2 - \frac{v_2 \cdot v_1^*}{v_1^* \cdot v_1^*} v_1^* = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}.$$

Standardizing these two eigenvectors results in

$$w_1 = \frac{v_1^*}{\|v_1^*\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$w_2 = \frac{v_2^*}{\|v_2^*\|} = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}.$$

2. As $\lambda = 4,$ solve for the homogeneous system

$$(4I - A)x = 0.$$

The eigenvectors are : $r\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, r \in R,\ r \neq 0$.

$\Rightarrow v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvectors of $\boldsymbol{A}$. Standardizing the eigenvector results in

$$w_3 = \frac{v_3}{\|v_3\|} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}.$$

Thus,

$$P = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix},$$

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$

and $D = P^t AP$.

Note: For a set of vectors $v_1, v_2, \ldots, v_n$, we can find a set of orthogonal vectors $v_1^*, v_2^*, \ldots, v_n^*$ via **Gram-Schmidt process:**

$$v_1^* = v_1$$

$$v_2^* = v_2 - \frac{v_2 \cdot v_1^*}{v_1^* \cdot v_1^*} v_1^*$$

$$\vdots$$

$$v_i^* = v_i - \frac{v_i \cdot v_{i-1}^*}{v_{i-1}^* \cdot v_{i-1}^*} v_{i-1}^* - \frac{v_i \cdot v_{i-2}^*}{v_{i-2}^* \cdot v_{i-2}^*} v_{i-2}^* - \cdots - \frac{v_i \cdot v_2^*}{v_2^* \cdot v_2^*} v_2^* - \frac{v_i \cdot v_1^*}{v_1^* \cdot v_1^*} v_1^*$$

$$\vdots$$

$$v_n^* = v_n - \frac{v_n \cdot v_{n-1}^*}{v_{n-1}^* \cdot v_{n-1}^*} v_{n-1}^* - \frac{v_n \cdot v_{n-2}^*}{v_{n-2}^* \cdot v_{n-2}^*} v_{n-2}^* - \cdots - \frac{v_n \cdot v_2^*}{v_2^* \cdot v_2^*} v_2^* - \frac{v_n \cdot v_1^*}{v_1^* \cdot v_1^*} v_1^*$$

UNIT V

LINEAR TRANSFORMATIONS

5.1. Definition of linear Transformations and examples

A linear transformation is a special type of function. Hence before the discussion about linear maps, it is helpful to revise the concept of functions.

Activity 5.1.1: Recall about the meaning and properties of a function. Also try to recall about related concepts like domain, range, one to one, on to, composition, and inverse.

Recall that a function (mapping) consists of the following:

i) a set X, each of whose element is mapped

ii) a set Y, to which each element of **x** is mapped

iii) a rule (correspondence) *f*, which associates with each element **x** of X a single element $f(\mathbf{x})$ of Y.

A function (mapping) *f* from X to Y is written as $f: X \to Y$.

If *f* maps an element **a** of X in to element **b** of Y then we write $f(\mathbf{a}) = \mathbf{b}$. Here **b** is called the image of **a** under *f* and **a** is called pre-image of **b**. Moreover, for any subset A of X, the set of all images of elements of A is denoted by $f[A]$ and we call it the image of A under *f*. On the other hand if B is a subset of Y the set of all elements of X whose images are in B is denoted by $f^{-1}[B]$ and we call it the pre-image of B under *f*. Using set notation the image of A and the pre-image of B under *f* can be defined respectively as

$$f[A] = \{\mathbf{b} \in Y \mid \mathbf{b} = f(\mathbf{a}) \text{ for some } \mathbf{a} \in A\} \text{ and } f^{-1}[B] = \{\mathbf{a} \in X \mid f(\mathbf{a}) \in B\}$$

Note: For any function $f: X \to Y$,

1. Domain of f = X and range of f = f[X] the image of set X under *f*.
2. *f* is one-to-one (injective) if and only if for any $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

3. f is on to (surjective) if and only if for every **b** in Y there exists an element **a** in X such that $f(\mathbf{a}) = \mathbf{b}$, i.e $f[X] = Y$ or range of $f = Y$.
4. We say that f is invertible (has an inverse) if and only if there is a function g from Y into X such that

   i) gof is the identity function on X, i.e $(gof)(x) = x$ for all $x \in X$.

   ii) fog is the identity function on Y i.e $(fog)(y) = y$ for all $y \in Y$.

   This function $g: Y \to X$ is unique whenever it exists and we call it the inverse of f. It is denoted by f^{-1}. Moreover from (i) and (ii) it follows that $f(x) = y$ if and only if $\mathbf{f}^{-1}(\mathbf{y}) = \mathbf{x}$. Observe that if there is a function $g: Y \to X$ satisfying (i) then f is one-to-one. Similarly if there is a function $g: Y \to X$ satisfying (ii) then f is on to. What do you conclude from this?

   Further one can easily verify that if $f: Y \to X$ is one-to one and on to then there exists a unique function $g: Y \to X$ which satisfies (i) and (ii). Thus f is invertible if and only if it is one-to-one and on to.

Activity 5.1.2:

1. Let $A = [2, \infty)$ and $B = [-4, \infty)$. Define a function

   $f: A \to B$ by $f(x) = x^2 - 4x$.

   a) show that f is one-to-one

   b) Show that f is on to

   c) Find the inverse of f
2. Let $g: \Re^2 \to \Re^2$ be given by $g(a, b) = (a + b, b)$

   i) Is g one-to-one? Verify!

   ii) Is g onto? Verify!

   iii) Does g have inverse? If so find its inverse.

As it is already mentioned a linear transformation is a special type of function. What is/are the condition(s) that a function must satisfy so as to be a linear transformation?

The following definition will enable us to give a complete answer for this question.

Definition 5.1.1: Let V and W be vector spaces over the same field K. A function

T: V $\rightarrow$ W is called a linear transformation (or a linear mapping) of V in to W if it satisfies the following conditions:

i) $T(u + v) = T(u) + T(v) \quad \forall\, u,v \in V$

ii) $T(\alpha u) = \alpha\, T(u) \quad \forall \alpha \in K$ and $\forall u \in V$

Note: 1. Using condition (ii) of the definition, one can show that $T(O_v) = O_w$ where O_v and O_w are zero vectors in V and W respectively.

$T(O_v) = T(\mathbf{0}.u)$ (because $\mathbf{0}.u = O_v$ for any $u \in V$)

$= \mathbf{0}T(u)$ (by (ii),$\mathbf{0} \in$ K (The zero element in the field K))

$= O_W$ (Why?)

This proves that a linear mapping maps a zero vector in to zero vector.

2) The two conditions in the definition are equivalent to

$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) \quad \forall \alpha, \beta \in K$ and $\forall u, v \in V$

Let us prove that a function T from vector space V in to W over the same field K is a linear transformation iff $T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2)$ for any $\alpha_1, \alpha_2 \in K$ and for any $v_1, v_2 \in V$.

Proof: First suppose T: V $\rightarrow$ W is a linear transformation.

Since $\alpha_1 v_1, \alpha_2 v_2 \in V\ \forall \alpha_1, \alpha_2 \in K$ and $\forall v_1, v_2 \in V$,

$T(\alpha_1 v_1 + \alpha_2 v_2) = T(\alpha_1 v_1) + T(\alpha_2 v_2)$ by condition (i) of definition 4.1.1

$= \alpha T(v_1) + \alpha_2\, T(v_2)$ by condition (ii) of definition 4.1.1

Thus $T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2) \quad \forall \alpha_1, \alpha_2 \in K$ and $\forall v_1, v_2 \in V$.

Next suppose that

$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2\, T(v_2) \quad \forall \alpha_1, \alpha_2 \in K$ and $\forall v_1, v_2 \in V$ *

We have to prove that T is a linear transformation.

$T(v_1 + v_2) = T(1.v_1) + 1.v_2)$, where 1 is the multiplicative identity in K.

$= 1.T(v_1) + 1.T(v_2)$ From (*)

$= T(v_1) + T(v_2)$.

$T(\alpha_1 v_1) = T(\alpha_1 v_1 + \mathbf{0}.v_2)$, where $\mathbf{0}$ is the zero element of K.

$= \alpha_1 T(v_1) + \mathbf{0}.T(v_2)$, from (*)

$= \alpha_1 T(v_1) + O_{w,}$ ($\mathbf{0}.T(v_2) = O_w$ – zero vector in W)

$= \alpha_1 T(v_1)$

Therefore $T(v_1 + v_2) = T(v_1) + T(v_2)$and $T(\alpha_1 v_1) = \alpha_1 T(v_1)$

$\forall\ \alpha_1 \in k, \forall v_1, v_2 \in V$.

Consequently T: V $\rightarrow$ W is a linear transformation if and only if

$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2)\ \forall \alpha_1, \alpha_2 \in K$ and $\forall v_1, v_2 \in V$.

Observe that by induction we also have a more general relation

$T(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \ldots + \alpha T(v_n)$ or

$$T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i T(v_i)$$ for any n vectors $v_1, v_2, \ldots, v_n$ in V

and for any n scalars $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ in K whenever T:V$\rightarrow$W is a linear transformation.

3) A linear mapping T: V$\rightarrow$ V is often called a linear operator on V.

Now let us see some examples of a linear transformation (or a linear mapping).

Example 5.1.1 Let V be a vector space over the field K. Then the mapping I: V $\rightarrow$ V given by I(v) = v $\forall v \in V$ is a linear transformation. To prove this let $u,v \in V$ and $\alpha \in K$. Then $u + v \in V$ and $\alpha u \in V$ as V is a vector space. Since I(x) = x $\forall x \in V$, we have

i) I (u + v) = u + v and I(u) + I(v) = u + v

Thus I (u + v) = I(u) + I(v)

ii) I (αu) = αu and αI(u) = αu

Thus I (αu) = αI(u)

Therefore I is a linear transformation. We call I identity transformation.

Example 5.1.2: Let T be a mapping from a vector space V over a field K into it self given by $T(v) = O_v \ \forall v \in V$ where O_v is a zero vector in V.

Then T is a linear transformation. (Verify!). We call this linear transformation the zero transformation.

Example 5.1.3: Let V be the vector space of all differentiable real valued functions of real variables on an open interval (a, b). Then the mapping $D: V \to V$ given by $\mathbf{D(f) = f'}$ (where $\mathbf{f'}$ is the derivative of **f**) is a linear transformation. This can be easily verified by using the properties of derivative.

Example 5.1.4: Let $T : M_n \to M_n$(where M_n is the vector space of n x n matrices over $\Re$) be given by $T(A) = A^t$. Show that T is a linear transformation.

To show this let $A, B \in M_n$ and $\alpha \in \Re$.Then using the properties of transpose of matrices we have,

$(A + B)^t = A^t + B^t$ and $(\alpha A)^t = \alpha(A^t) = \alpha A^t$

Thus $T(A + B) = (A + B)^t$ by definition of T

$= A^t + B^t$

$= T(A) + T(B)$

Also $T(\alpha A) = (\alpha A)^t = \alpha A^t = \alpha T(A)$

Therefore T is a linear transformation.

Example 5.1.5: Let a function $L: \Re^2 \to \Re^3$ be given by $L(x, y) = (x + y, y, x)$

Is L a linear transformation?

Solution: Let $(a, b), (c, d) \in \Re^2$ and $\alpha \in \Re$.

Then (i) $T((a, b) + (c, d)) = T(a + c, b + d)$

$= (a + c + b + d, b + d, a + c)$

$= (a + b, b, a) + (c + d, d, c))$

$= T(a, b) + T(c, d)$

ii) $T(\alpha(a, b)) = T(\alpha a, \alpha b)$

$= (\alpha a + \alpha b, \alpha b, \alpha a)$

$= \alpha(a + b, b, a)$

$= \alpha\ T(a, b)$

Thus T is a linear transformation.

Example 5.1.6: Is the mapping N: $\Re^3 \to \Re$ defined by $N(x, y, z) = \|(x, y, z)\|$ a linear transformation?

Solution: No, take $(0, 1, 2), (3, 0, 1) \in \Re^3$

$N((0, 1,2) + (3, 0, 1)) = N(3, 1, 3)$

$= \|(3,1,3)\| = \sqrt{3^2 + 1^2 + 3^2}$

$= \sqrt{19}$

$N(0, 1, 2) + N(3, 0, 1) = \|(0,1,2)\| + \|3, 0, 1\|$

$= \sqrt{0^2 + 1^2 + 2^2} + \sqrt{3^2 + 0^2 + 1^2}$

$= \sqrt{5} + \sqrt{10}$

Since $\sqrt{19} \neq \sqrt{5} + \sqrt{10}$, $N((0,1,2) + (3, 0, 1)) \neq N(0,1,2) + (3,0,1)$

Thus N is not a linear transformation as $N(u + v) = N(u) + N(v)$ must hold true for all vectors u, v in $\Re^3$.

Remark: To show that a mapping T from a vector space V in to W over the same field K is not a linear transformation it suffices to show that there exists two vectors $v_1, v_2 \in V$ such that $T(v_1 + v_2) \neq T(v_1) + T(v_2)$ or there exists scalar $\alpha \in V$ such that $\mathbf{T(\alpha v) \neq \alpha T(v)}$.

Activity 5.1.3:

1. Show that the mapping $T : \Re^3 \to \Re^2$ defined by $T(x, y, z) = (x-y, x - z)$ is a linear transformation.

2. Is the mapping L: $\Re^3 \to \Re^2$ defined by L (a, b, c) = (| a |, 0) a linear transformation? Justify your answer!

Let us add one more example. Recall that vectors $v_1, v_2, v_3, \ldots v_m$ in a vector space V over a field K are linearly independent iff $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = O_v$ (where $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$) implies $\alpha_1 = \alpha_2 = \ldots. = \alpha_n = 0$

Example 5.1.7: Let T be a linear transformation from a vector space V in to W over the same field K. Prove that the vectors $v_1, v_2, v_3, \ldots, v_n \in V$ are linearly independent if $T(v_1), T(v_2), T(v_3), \ldots, T(v_n)$ are linearly independent vectors in W.

Solution: Suppose $T(v_1), T(v_2), T(v_3), \ldots, T(v_n)$ are linearly independent vectors in W where $v_1, v_2, v_3, \ldots, v_n$ are vectors in V and $T:V \to W$ is a linear transformation. To prove that $v_1, v_2, v_3, \ldots, v_n$ are linearly independent.

Let $\alpha_1 \alpha_2, \alpha_3, \ldots, \alpha_n \in K$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \ldots + \alpha_n v_n = O_v$$

Then $T(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \ldots + \alpha_n v_n) = T(O_v)$

So $\alpha_1 T(v_1) + \alpha_2 T(v_2) + \alpha_3 T(v_3) + \ldots + \alpha_n T(v_n) = O_w$

But $T(v_1), T(v_2), T(v_3), \ldots, T(v_n)$ are linearly independent.

Hence $\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_n = 0$. Thus we have shown that

$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \ldots + \alpha_n v_n = O_v$ implies

$$\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_n = 0.$$

Consequently $v_1, v_2, v_3, \ldots, v_n$ are linearly independent.

Exercise

1. Determine whether or not each of the following mappings is linear transformation.
   a) T: $\Re^2 \to \Re^2$ given by T (x, y) = (x + y, x)
   b) T: $\Re^2 \to \Re$ given by T (x, y) = xy
   c) T: $\Re^3 \to \Re^2$ given by T (x_1, x_2, x_3) = ($1 + x_1, x_2$)

d) L: $\Re^3 \to \Re^2$ given by L (x, y, z) = (z + x, y)

e) L: $\Re^2 \to \Re^2$ given by L (p, q) = (p^3, q^3)

2. Let M_2 denote the vector space of 2×2 matrices over $\Re$.

Let T: $M_2 \to M_2$ be given by $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+b & 2c \\ -3a & c+d \end{pmatrix}$

Is T a linear mapping? Justify your answer!

3. Let V be the vector space of $m \times n$ matrices over $\Re$. Let P be a fixed $m \times m$ matrix and Q a fixed $n \times n$ matrix over $\Re$. Show that the mapping L: $V \to V$ defined by T(A) = PAQ is a linear transformation.

4. Show that the mapping F: $\Re^2 \to \Re$ defined by F(a, b) = |a – b| is not a linear transformation.

5. Let G : $V \to W$ be a linear transformation where V and W are vector spaces over the same field K.

Prove that if $u_1, u_2, u_3, \ldots, u_n$ are linearly dependent vectors in V then $G(u_1), G(u_2), G(u_3), \ldots, G(u_n)$ are linearly dependent vectors in W.

6. Let U, V, and W be vector spaces over the same field K.

If g: $U \to V$ and f : $V \to W$ are linear transformations show that fog is also a linear transformation form U in to W.

7) i) Let A = (a, b, c) be a fixed given vector in $\Re^3$. Define T: $\Re^3 \to \Re$ by

T(X) = A•X $\forall X \in \Re^3$.

(A•X is the scalar (dot) product of A and X). Show that T is a linear transformation.

ii) Let A be as in (i) define T: $\Re^3 \to \Re$ by T(X) = A•X + 4. Show that T is not a linear transformation.

8) Let V be the space of n x 1 matrices over $\Re$ and let W be the space of $m \times 1$ matrices over $\Re$. Let A be a fixed $m \times n$ matrix over $\Re$. Define T: $V \to W$ by T(X) = AX, $\forall X \in V$.

Prove that i) T is a linear transformation.

ii) T is a zero transformation if and only if A is the zero matrix.

9) Let V be the vector space of all $n \times n$ matrices over $\Re$ and let B a fixed $n \times n$ matrix.

If T: $V \to V$ is defined by $T(A) = AB - BA$ then verify that T is a linear transformation.

10) Let V be a vector space over $\Re$, and $f: V \to \Re$, $g: V \to \Re$ be two linear transformations. Let $F: V \to \Re^2$ be the mapping defined by $F(v) = (f(v), g(v))$. Show that F is a linear transformation.

11) Let V, W be two vector spaces over the same field K and let $F: V \to W$ be a linear transformation. Let U be the subset of V consisting of all elements u such that $F(u) = O_w$. Prove that U is a subspace of V.

12) Let $F: \Re^3 \to \Re^4$ be a linear transformation. Let P be a point of $\Re^3$ and A is a non zero element of $\Re^3$. Describe F[S], where $S = \{X \in \Re^3 | X = P + t A, t \in \Re\}$. (Distinguish the cases when $F(A) = 0$ and $F(A) \neq 0$).

Properties of Linear transformation

Recall that in the previous section, we have seen that if T: $V \to W$ is a linear transformation from vector space V in to W over the same field K then

i) $T(O_v) = O_w$

ii) $\mathbf{T}\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) = \sum_{i=1}^{n} \alpha \mathbf{T}(\mathbf{v}_i)$ $\forall \alpha_i \in K$ and $\forall v_i \in V$, $i = 1, 2, \ldots, n$.

We now state two other basic properties in the following theorem. The proof is left for you as an exercise.

Theorem 5.1.1: Let T be a linear transformation from a vector space V in to W over the same field K. Then i) $T(-v) = -T(v)$ $\forall v \in V$

ii) $T(v_1 - v_2) = T(v_1) - T(v_2)$ $\forall v_1, v_2 \in V$

Our next theorem asserts that a linear transformation from a given finite dimensional vector space V in to any vector space W is completely determined by its values on the elements of a given basis of V.

Theorem 5.1.2: Let V and W be vector spaces over the field K. Let $\{v_1, v_2, v_3, \ldots, v_n\}$be a basis of V. If $\{w_1, w_2, w_3, \ldots, w_n\}$is a set of arbitrary vectors in W, then there exists a unique linear transformation $F: V \to W$ such that $F(v_j) = w_j$ for $j = 1, 2, \ldots n$.

To prove the theorem, we need to

a) define a function F from V into W such that $F(v_i) = w_i$ for all $i = 1, 2, 3, \ldots n$.

b) show F is a linear transformation

c) show that F is unique.

Proof: Let V and W be vector spaces over the field K. Let $\{v_1, v_2, v_3, \ldots, v_n\}$ be a basis of V and $\{w_1, w_2, w_3, \ldots, w_n\}$ be any set of n-vectors in W.

Since $\{v_1, v_2, v_3, \ldots, v_n\}$ is a basis of V, for any $v \in V$ there exist unique scalars $a_1, a_2, a_3, \ldots, a_n \in K$ such that

$$\mathbf{v} = \sum_{i=1}^{n} \mathbf{a_i v_i}$$

a) Define $F : V \to W$ by $F(v) = a_1w_1 + a_2 w_2 + a_3w_3 + \ldots + a_nw_n$

i.e $F(v) = \sum_{i=1}^{n} a_i w_i$ where $\mathbf{v} = \sum_{i=1}^{n} \mathbf{a_i v_i}$

Every element v of V is mapped to only one element of W as the scalars a_i 's are unique. As W is a vector space, every linear combination of vectors in W is also in W. So $\sum_{i=1}^{n} a_i w_i \in W$. Thus F is a function from V in to W

Moreover, since $v_i = 0.v_1 + 0.v_2 + \ldots + Ov_{i-1} + 1.v_i + 0.v_{i+1} + \ldots + 0.v_n$

for any $i = 1, 2, 3, \ldots, n$, we have

$F(v_i) = 0.w_1 + 0.w_2 + \ldots + 0w_{i-1} + 1.w_i + 0.w_{i+1} + \ldots + 0.w_n = w_i$

So $F(v_i) = w_i$ for each $i = 1, 2, 3, \ldots, n$

b) We show that F is a linear transformation,

i.e $F(u + v) = F(u) + F(v)$ and $F(\alpha v) = \alpha F(v)$ $\forall \alpha \in K$ and $\forall u, v \in V$.

To do this, let $x, y \in V$ and $\alpha \in K$

Then $x = \sum_{i=1}^{n} x_i v_i$ and $y = \sum_{i=1}^{n} y_i v_i$ for some unique scalars $x_1, x_2, x_3, \ldots, x_n$, $y_1, y_2, y_3, \ldots, y_n$ in K.

(i) $F(x + y) = F\left(\sum_{i=1}^{n} x_i v_i + \sum_{i=1}^{n} y_i v_i\right)$

$= F\left(\sum_{i=1}^{n} (x_i + y_i) v_i\right)$

$= \sum_{i=1}^{n} (x_i + y_i) w_i$ by definition of F

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

$= \mathbf{F}\left(\sum_{\mathbf{i}=1}^{\mathbf{n}} \mathbf{x_i}\, \mathbf{v_i}\right) + \mathbf{F}\left(\sum_{\mathbf{i}=1}^{\mathbf{n}} \mathbf{y_i}\, \mathbf{v_i}\right)$ by definition of F

$= F(x) + F(y)$

$F(x + y) = F(x) + F(y) \quad \forall x, y \in V$

ii) $F(\alpha x) = F\left(\alpha \sum_{i=1}^{n} x_i v_i\right)$

$= F\left(\sum_{i=1}^{n} (\alpha x_i) v_i\right)$

$= \sum_{i=1}^{n} (\alpha x_i) w_i$ by definition of F

$= \alpha \sum_{i=1}^{n} x_i w_i$

$= \alpha\, F\left(\sum_{i=1}^{n} x_i v_i\right)$ by definition of F

$= \alpha\, F(x)$

So $F(\alpha x) = \alpha F(x) \quad \forall \alpha \in K$ and $\forall x \in V$

Therefore from (i) and (ii) we conclude that F is a linear transformation from V in to W.

c) In (a) and (b) we have shown the existence of a linear transformation $F : V \to W$ such that $F(v_i) = w_i$ for all $i = 1, 2, 3, \ldots, n$.

To prove that F is unique, suppose that $G: V \to W$ is a linear transformation such that $G(v_i) = w_i \ \forall_i \in \{1, 2, \ldots, n\}$

Let x be any vector in V then $x = \sum_{i=1}^{n} x_i v_i$ for some unique scalars $x_1, x_2, x_3, \ldots, x_n$ in K.

Thus $$G(x) = G\left(\sum_{i=1}^{n} x_i v_i\right)$$

$$= \sum_{i=1}^{n} x_i G(v_i) \text{ as G is a linear transformation}$$

$$= \sum_{i=1}^{n} x_i w_i \quad \text{as } G(v_i) = w_i \text{ for each } i = 1, 2, 3, \ldots, n.$$

$$= F\left(\sum_{i=1}^{n} x_i v_i\right) \quad \text{by definition of F.}$$

$$= F(x)$$

Since $G(x) = F(x)$ for any $x \in V$, we conclude that $G = F$.

This proves that F is unique. With this we complete the proof of the theorem.

Remark: **1)** The vectors $w_1, w_2, w_3, \ldots, w_n$ in theorem 4.2.2 are completely arbitrary; they may be linearly dependent, independent or they may even be equal to each other. But the number of these vectors in W must be equal with that of the number of basis vectors of V.

2) In determining the linear transformation from V in to W the assumption that $\{v_1, v_2, \ldots, v_n\}$ is a basis of V is essential.

Example 5.1.8:

a) Is there a linear transformation T from $\Re^2$ in to $\Re^2$ such that T(2, 3) = (4, 5) and T(1, 0) = (0, 0)?

b) How many linear transformations satisfying the given conditions do we have?

Solution: a) Yes. The two vectors (2, 3) and (1, 0) are linearly independent and hence they form a basis for $\Re^2$. Thus according to theorem 4.2.2, there is unique linear transformation from $\Re^2$ in to $\Re^2$ such that T(2,3) = (4, 5) and T(1, 0) = (0, 0).

b) As it is verified above, we have only one linear transformation satisfying the given conditions**.**

Example 5.1.9: Describe explicitly the linear transformation T: $\Re^2 \to \Re^2$ such that

T(2, 3) = (4, 5) and T(1, 0) = (0, 0).

Solution: Let (x, y) $\in \Re^2$. Then (x, y) can be expressed as a linear combination of (2, 3) and (1, 0)as {(2,3), (1,0)} form a basis of $\Re^2$.

So (x, y) = a(2, 3) + b(1, 0) for some a, b $\in \Re$ or (x, y) = (2a + b, 3a)

Hence we have x = 2a + b and y = 3a, which in turn implies $a = \frac{y}{3}$ and

$$b = x - \frac{2y}{3}$$

Thus (x, y) = $\frac{y}{3}(2, 3) + \left(x - \frac{2y}{3}\right)(1, 0)$

$$T(x, y) = T\left(\frac{y}{3}(2,3) + \left(x - \frac{2}{3}y\right)(1, 0)\right)$$

$$= \frac{y}{3}T(2,3) + \left(x - \frac{2}{3}y\right)T(1, 0)$$

$$= \frac{y}{3}(4,5) + \left(x - \frac{2}{3}y\right)(0, 0)$$

$$= \left(\frac{4}{3}y \quad , \quad \frac{5y}{3}\right) + (0, 0)$$

Therefore $T(x, y) = \left(\frac{4}{3}y, \frac{5y}{3}\right) \forall$ (x, y) $\in \Re^2$

Observe that the image of any vector (a, b) $\in \Re^2$ under the linear transformation of

example 4.2.2 is $\mathbf{b}\left(\frac{4}{3},\frac{5}{3}\right)$. So the image of $\Re^2$ under T is the line through (0, 0) with direction vector $\left(\frac{4}{3},\frac{5}{3}\right)$.

Activity 5.1.4:

Let T: $\Re^2 \rightarrow \Re^2$ be as in example 4.2.2 above, i.e $\mathbf{T(a,b)} = \left(\frac{4}{3}\mathbf{b}, \frac{5}{3}\mathbf{b}\right)$

i) Let A = {(x, y)| $x^2 + y^2 = 1$}. Find the image of A under T i.e T[A].

ii) Describe the set containing all elements in $\Re^2$ whose images is (0,0)

Example 5.1.10: Find a linear transformation T : $\Re^2 \rightarrow \Re^3$ such that

T(1, 2) = (3, -1, 5) and T(0, 1) = (2, 1, -1)

Solution: β = {(1, 2), (0, 1)} is a basis of $\Re^2$. (verify)

Thus there exist a unique linear transformation (by theorem 4.2.2) T: $\Re^2 \rightarrow \Re^3$ such that T(1,2) = (3, -1, 5) and T(0, 1) = (2, 1, -1).

To describe this unique linear transformation explicitly, suppose (x, y) $\in \Re^2$

Then (x, y) = t_1 (1, 2) + t_2 (0, 1) for some $t_1, t_2 \in \Re$ as β is a basis of $\Re^2$.

= $(t_1, 2t_1 + t_2)$

Thus we have $\begin{cases} \mathbf{x} = \mathbf{t}_1 \\ \mathbf{y} = 2\mathbf{t}_1 + \mathbf{t}_2 \end{cases}$

Which in turn implies $t_1 = x$ and $t_2 = y - 2x$. So (x, y) = x(1, 2) + (y – 2x) (0, 1)

and T(x, y) = x T(1, 2) + (y - 2x) T(0, 1)

= x(3, -1, 5) + (y – 2x) (2, 1, -1)

= (3x, -x, 5x) + (2y – 4x, y-2x, 2x – y)

= (2y – x, y – 3x, 7x – y)

Therefore the required linear transformation is given by

T(x, y) = (2y – x, y – 3x, -y + 7x).

Activity 5.1.5:

1 a) Find a linear transformation T: $\Re^2 \to \Re^2$ such that T(1, 0) = (1, 1) and T(0, 1) = (-1, 2)

b) Prove that T maps the square with vertices (0, 0), (1, 0), (1, 1) and (0, 1) in to a parallelogram.

2. a) Is there a linear transformation T:$\Re^3 \to \Re^3$ such that T(0, 1, 2) = (3, 1, 2) and T(1, 1, 1) = (2, 2, 2)?

b) If your answer in (a) is yes,

(i) find T.

(ii) is it unique? Why?

Example 5.1.11: Is there a linear transformation L from $\Re^2$ in to $\Re^2$ such that

L(1, -1, 1) = (1, 0) and L(1, 1, 1) = (0, 1)?

Solution: Even though (1, -1, 1) = (1, 1,1) are linearly independent vectors in $\Re^3$,

They do not from a basis of $\Re^3$. This is because a basis of $\Re^3$ must contain three linearly independent vectors as $\Re^3$ a 3-dimensional vector space. So theorem 4.2.2 does not guarantee us the existence of a linear transformation L : $\Re^3 \to \Re^2$ such that L(1, -1, 1) = (1, 0) and L(1, 1, 1) = (0, 1)? But to make use of the theorem we need to have a basis of $\Re^3$ containing the two given vectors v_1 = (1, -1,1) and v_2 = (1, 1, 1) and a vector v_3 different from v_1 and v_2. That is we need to find a vector v_3 so that $\{v_1, v_2, v_3\}$ is a basis of $\Re^3$. For this put v_3 = (0, 0, 1).

Since α (1, -1, 1) + β(1, 1, 1) + γ(0, 0, 1) = (0,0,0) for some $\alpha,\beta,\gamma \in \Re$ implies $\alpha = \beta = \gamma = 0$, (1, -1, 1), (1,1,1) and (0, 0, 1) are three linearly independent vectors in $\Re^3$. In addition to this let us take one more vector in $\Re^2$, say (0, 0). Now we have a basis {(1, -1, 1), (1, 1, 1), (0, 0, 1)} of $\Re^3$ and three vectors (1, 0), (0, 1), (0,0) in $\Re^2$. Thus by theorem 4.2.2 there exists a unique linear transformation L: $\Re^3 \to \Re^2$ such that

L(1, -1, 1) = (1, 0), L(1, 1, 1) = (0, 1) and L(0, 0, 1) = (0, 0).

To find this linear transformation, first let us write any (x, y, z) $\in \Re^3$ as a linear combination of (1, -1, 1), (1,1,1) and (0, 0, 1). This is possible as they form basis of $\Re^3$.

Set (x, y, z) = a(1, -1, 1) + b(1, 1,1) + c(0, 0, 1); a, b, c $\in \Re$ and solve for a, b, c interms of x, y, and z.

$$(x, y, z) = (a, -a, a) + (b, b, b) + (0, 0, c)$$

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

$$\Rightarrow \begin{cases} x = a+b \\ y = -a+b \\ z = a+b+c \end{cases}$$

$$\Rightarrow \quad b = \frac{x+y}{2}, \quad a = \frac{x-y}{2}, \quad c = z - x$$

Thus $(x, y, z) = \frac{(x-y)}{2}(1,-1,1) + \frac{(x+y)}{2}(1,1,1) + (z - x)(0,0,1)$

$$L(x, y, z) = \frac{(x-y)}{2} L(1,-1,1) + \frac{(x+y)}{2} L(1,1,1) + (z-x) L(0,0,1)$$

$$= \frac{x-y}{2}(1, 0) + \frac{(x+y)}{2}(0, 1) + (z-x)(0, 0)$$

$$= \left(\frac{x-y}{2}, \ \frac{x+y}{2} \right)$$

Therefore we have obtained a linear transformation L: $\Re^3 \to \Re^2$ given by

$$L(x, y, z) = \left(\frac{x-y}{2}, \ \frac{x+y}{2} \right)$$

Moreover $\quad L(1,-1,1) = \left(\frac{1-(-1)}{2}, \ \frac{1+(-1)}{2} \right) = (1,0)$ and

$$L(1, 1,1) = \left(\frac{1-1}{2}, \ \frac{1+1)}{2} \right) = (0, 1).$$

Consequently we can say that there is a linear transformation L: $\Re^3 \to \Re^2$ satisfying the two given conditions L(1, -1, 1) = (1, 0) and L(1,1,1) = (0, 1).

Note: The linear transformation T: V $\rightarrow$ W whose existence and uniqueness is guaranteed by theorem 4.2.2 depends on the given basis vectors of V and the given vectors of W whose number equals the number of basis vectors in V.

Is L: $\Re^3 \rightarrow \Re^2$ given by L(x, y, z) = $\left(\frac{x-y}{2}, \frac{x+y}{2}\right)$ the only linear mapping that can satisfy the requirements of the question in the above example? Replace (0, 0, 1) by (1, 0, 0) in the solution of the above example and find a linear transformation
L: $\Re^3 \rightarrow \Re^2$ such that L(1, -1, 1) = (1, 0) and L(1, 1, 1) = (0, 1). Do the same by replacing (0, 0) by (1, 1) and (0, 0,1) by (1, 0, -1).

Exercise:

1. Find a linear transformation

   a) T: $R^2 \rightarrow \Re^2$ such that T(1, 0) = (2, -1) and T(0,1) = (-2, 1)
   b) T: $\Re^2 \rightarrow \Re^2$ such that T(1, 2) = (3, 0) and T(2, 1) = (1, 2)
   c) L: $\Re^2 \rightarrow \Re^3$ such that L(2, -5) = (-1, 2, 3) and L(3, 4) = (0, 1, 5)
   d) L: $\Re^3 \rightarrow \Re^2$ such that L(3, 1,1) = (1, 0), L(2, -1, 5) = (0, 1) and L(4, 0, -3) = (-1, 1)
   e) F: $\Re^2 \rightarrow \Re^3$ such that T(2, 1) = (3, -1, 5) and T(1, 0) = (2, 1, -1)
   f) G: $\Re^3 \rightarrow \Re^2$ such that G(2, 0, 1) = 1, 1) and G(0, 3, 0) = (-1, 1)

2. If $v_1 = (1, -1)$ $w_1 = (1, 0)$
   $v_2 = (2, -1)$ $w_2 = (0, 1)$
   $v_3 = (-3, 2)$ $w_3 = (1, 1)$
   Is there a linear transformation G from $\Re^2$ into $\Re^2$ such that $G(v_i) = w_i$ for i = 1, 2, 3 ? If so find G.
3. Let T: V $\rightarrow$ W be a linear mapping . Let v be an element of V. Show that T(-v) = -T(v)
4. Let L: $\Re^2 \rightarrow \Re^2$ be a linear mapping, such that
   L(3, 1) = (1, 2) and L(-1, 0) = (1, 1) compute L(1, 0)

5.2. The Rank and Nullity of a linear transformation and examples

In this section we will discuss in detail about two important sets related to a linear transformation T: V $\rightarrow$ W where V and W are vector spaces over the same field K. One of them is a subset of V and the other is a subset of W. Is there an element v in V such that T(v) = O_w -the zero vector in W? We know that T(O_v) = O_w, so there is at least one element v in V such that T(v) = O_w. Hence we have a non-empty subset

U = {u $\in$ V| T(u) = O_w} of V. It is a subspace of V because

i) $O_v \in U$ as T(O_v) = O_w

ii) If $u_1, u_2 \in U$ then T(u_1) = T(u_2) = O_w

So T($u_1 + u_2$) = T(u_1) + T(u_2) as T is a linear transformation

= $O_w + O_w$

= O_w

Thus $u_1 + u_2 \in U$ for any $u_1, u_2 \in U$

iii) if u $\in$ U and $\alpha \in K$, then T(u) = O_w

T(αu) = αT(u) as T is a linear transformation

= $\alpha.O_w$

= O_w

So α u $\in$ U for any $\alpha \in K$ and for any u $\in$ U

From (i), (ii) and (iii) it follows that U is a subspace of V.

On the other hand we know that, the range of T is a subset of W. It is also a subspace of W(verify!).

This section is concerned with these two special subspaces. Let us begin with the following definition.

Definition 5.2.1: Let V, W be vector spaces over K and T: V $\rightarrow$ W be a linear transformation.

i) The set of elements v $\in$ V such that T(v) = O_w is called the **Kernel of T** or the **null space of T**.

ii) The set of elements w in W such that there exists an element v of V such that T(v) = w is called the **image** or the **range of T**.

Notation: We shall denote the Kernel and image of a linear transformation T: V $\rightarrow$ W by **KerT** and **ImT** respectively. So KerT = $\{u \in V | T(u) = O_w\}$ and ImT = $\{w \in W | w = T(v)$ for some $v \in V\}$

At the beginning of this section we have proved that kerT is a subspace of V. In the theorem below we prove that ImT is a subspace of W.

Theorem 5.2.1: Let T be a linear transformation from a vector space V in to W over the some field K. Then (a) Ker T is a subspace of V

(b) ImT is a subspace of W.

Proof : a) It is already proved

b) Clearly ImT is a subset of W

i) Since $T(O_v) = O_w$, $O_w \in$ ImT.

ii) Let $w_1, w_2 \in$ ImT. Then there exists $u_1, u_2 \in V$ such that $T(u_1) = w_1$ and $T(u_2) = w_2$. Since V is a vector space, $u_1 + u_2 \in V$. Moreover $T(u_1 + u_2) = T(u_1) + T(u_2) = w_1 + w_2$. Thus $w_1 + w_2 \in$ ImT as there exists a vector $v \in V$ such that $T(v) = w_1 + w_2$ $(v = u_1 + u_2)$.

So we have $w_1 + w_2 \in$ ImT $\forall w_1, w_2 \in W$.

iii) Let $\alpha \in K$ and $w \in$ ImT. Then there exists $v \in V$ such that T(v) = w as $w \in$ ImT. Since V is a vector space over K, $\alpha v \in V$. Moreover $T(\alpha v) = \alpha T(v) = \alpha w$. Hence $\alpha w \in$ ImT. From (i), (ii) and (iii) it follows that ImT is a subspace of W.

Example 5.2.1: Let $G : \Re^3 \rightarrow \Re^2$ be defined by $G(x, y, z) = (x + y, y - z)$

a) Show that G is a linear transformation

b) Find the ker G and ImG.

Solution: a) Let (x, y, z), $(u, v, w) \in \Re^3$ and $\alpha \in \Re$

Then $G((x, y, z) + (u, v, w)) = G(x + u, y + v, z + w)$

$$= (x+u+y+v, y+v-z-w)$$
$$= (x+y, y-z)+(u+v, v-w)$$
$$= G(x,y,z)+G(u,v,w)$$

and $G(\alpha(x,y,z)) = G(\alpha x, \alpha y, \alpha z)$
$$= (\alpha x+\alpha y, \alpha y-\alpha z)$$
$$= \alpha(x+y,\ y-z)$$
$$= \alpha G(x,y,z)$$

Therefore G is a linear transformation.

b) $\ker G = \{(p,q,r)\in\Re^3 \mid G(p,q,r)=(0,0)\}$
$$= \{(p,q,r)\in\Re^3 \mid (p+q,q-r)=(0,0)\}$$
$$= \{(p,q,r)\in\Re^3 \mid p+q=0 \text{ and } q-r=0\}$$
$$= \{(p,q,r)\in\Re^3 \mid -p=q=r\}$$
$$= \{(p,-p,-p) \mid P\in\Re\}$$
$$= \{p(1,-1,-1) \mid p\in\Re\}$$

So kerG is the subspace of $\Re^3$ generated by (1,-1,-1). List at least four elements of ker G.

$$\text{Im}G = \{(d,e)\in\Re^2 \mid G(x,y,z)=(d,e) \text{ for some } (x,y,z)\in\Re^3\}$$
$$= \{(d,e)\in\Re^2 \mid (x+y,y-z)=(d,e) \text{ for some } (x,y,z)\in\Re^3\}$$
$$= \{(d,e)\in\Re^2 \mid x+y=d \text{ and } y-z=e \text{ for some } x,y,z\in\Re\}$$
$$= \{(x+y,y-z) \mid x,y,z\in\Re\}$$
$$= \{(x+y,0)+(0,y-z) \mid x,y,z\in\Re\}$$
$$= \{(\mathbf{x}+\mathbf{y})(1,0)+(\mathbf{y}-\mathbf{z})(0,1) \mid \mathbf{x},\mathbf{y},\mathbf{z}\in\Re\}$$
$$= \{t_1(1,0)+t_2(0,1) \mid t_1=x+y\in\Re \text{ and } t_2=y-z\in\Re\}$$

Thus ImG is the subspace of $\Re^2$ generated by (1, 0) and (0, 1).

That is $\text{Im}\mathbf{G}=\Re^2$ (why?) observe that dim (ker G) = 1 and dim (ImG) = 2 .

In the following theorem we state an equivalent condition that help us to determine whether a given linear transformation T is one-to-one (injective) or not, using kerT.

Theorem 5.2.2: Let T: $V \to W$ be a linear transformation.

Then T is one-to-one if and only if ker T = $\{O_v\}$.

Proof: Suppose T is one-to-one. We need to show that kerT contains only the zero vector. Let $x \in$ kerT. Then $T(x) = O_w$.

But $T(O_v) = O_w$. So we have $T(x) = T(O_v)$, which implies $x = O_v$ as T is one-to-one. Therefore kerT = $\{O_v\}$

Conversely suppose kerT = $\{O_v\}$. We wish to show that T is one-to-one.

Now let $T(x_1) = T(x_2)$; $x_1, x_2 \in V$. Then $T(x_1) + -T(x_2) = O_w$ or

$T(x_1) + T(-x_2) = O_w$. Thus we have $T(x_1 + -x_2) = O_w$, which implies

$x_1 + (-x_2) \in$ ker T. Since ker T contains only the zero vector by assumption,

$x_1 + -x_2 = O_v$ and hence $x_1 = x_2$. Consequently T is one-to-one.

Let us illustrate how theorem 4.3.2 can be used to check whether a given linear transformation is one-to-one or not.

Example 5.2.2: Let T: $\Re^2 \to \Re^3$ be a linear transformation given by
T(a, b) = (a + b, a – b, b).

i) Find ker T and ImT.

ii) Is T one-to-one? Why?

Solution: (i) Ker T = $\{ (x, y) \in \Re^2 \mid T(x, y) = (0,0,0)\}$

$= \{(x, y) \in \Re^2 \mid T(x + y, x - y, y) = (0, 0, 0)\}$

$= \{(x, y) \in \Re^2 \mid x + y = 0, x - y = 0 \text{ and } y = 0\}$

$= \{(x, y) \in \Re^2 \mid x = 0 \text{ and } y = 0\} = \{(0, 0)\}$

From the definition of ImT, it follows that

$$(u,v,w) \in \mathrm{Im}T \Leftrightarrow (u,v,w) = T(a,b) \text{ for some } (a,b) \in \Re^2$$

$$\Leftrightarrow (u,v,w) = (a+b, a-b, b) \text{ for some } a,b \in \Re$$

$$\Leftrightarrow u = a+b, v = a-b \text{ and } w = b \text{ for some } a,b \in \Re$$

$$\Leftrightarrow u - v = 2b \text{ and } w = b \text{ for some } a, b \in \Re$$

$$\Leftrightarrow u - v = 2w \text{ where } u, v, w \in \Re$$

$$\Leftrightarrow u = v + 2w \text{ where } u, v, w \in \Re$$

$$\text{Thus } \mathrm{Im}T = \{(u, v, w) \mid u, v, w \in \Re \text{ and } u = v + 2w\}$$

$$= \{(v + 2w, v, w) \mid v, w \in \Re\}$$

$$= \{(v, v, 0) + (2w, 0, w) \mid v, w \in \Re\}$$

$$= \{v(1,1,0) + w(2,0,1) \mid v, w \in \Re\}$$

Therefore ImT is a subspace of $\Re^3$ generated by (1, 1, 0) and (2, 0,1).

What is dim(kerT)? What is dim(ImT)?

ii) Since ker T = {(0, 0)} (contains only the zero vector of $\Re^2$), T is one – to – one by theorem 4.3.2. But it is not on to why?

Notice that whenever kerT $\neq \{O_v\}$, we can conclude that T is not one-to-one.

Activity 5.2.1:

Let T: $\Re^3 \to \Re^3$ be given by T(x, y, z) = (x + y, 2y, 2y – x)

a) Show that T is a linear transformation

b) i) Find ker T

ii) Is T one-to-one? Verify

iii) Find ImT.

Does a linear transformation maps linearly independent vectors in to linearly independent vectors?

Consider the linear transformations L and T from $\Re^3$ in to $\Re^3$ given by

L(a, b, c) = (a, c, 0) and T(x, y, z) = (x, y + x, z + y). The vectors (1, 0, 0) and (2, 1, 0) are linearly independent in $\Re^3$. But L(1, 0, 0) = (1, 0, 0) and

L(2, 1, 0) = (2, 0, 0) are linearly dependent vectors in $\Re^3$. On the other hand

T(1, 0, 0) = (1, 1, 0) and T(2, 1,0) = (2, 3, 1) linearly independent. Thus from this particular instance we conclude that a linear transformation may or may not map linearly independent vectors in to linearly independent vectors.

Under what condition does it map linearly independent vectors in to linearly independent vectors?

The following theorem provides a sufficient condition for this.

Theorem 5.2.3: Let T: V $\rightarrow$ W be a linear transformation and $v_1, v_2, v_3 \ldots, v_n$ be linearly independent vectors in V. If kerT = $\{0_v\}$ then $T(v_1)$, $T(v_2)$, $T(v_3),\ldots, T(v_n)$ are linearly independent vectors in W.

Proof: Let $x_1, x_2, x_3 \ldots, x_n$ be scalars such that

$x_1T(v_1) + x_2T(v_2) + x_3 T(v_3) + \ldots + x_nT(v_n) = O_w$

Then by linearity of T we get, $T(x_1v_1 + x_2v_2 + x_3v_3 + \ldots + x_nv_n) = O_w$.

Hence $x_1v_1 + x_2 v_2 + x_3v_3 + \ldots + x_nv_n \in$ ker T. But ker T = $\{O_v\}$. Thus

$x_1v_1 + x_2 v_2 + x_3v_3 + \ldots x_nv_n = O_v$. Since $v_1, v_2, v_3, \ldots, v_n$ are linearly independent vectors in V, it follows that $x_1 = x_2 = x_3 = \ldots = x_n = 0$

This completes the proof.

In the theorem we have proved that if the kernel of a linear transformation T contains only the zero vector i.e T is 1 – 1 then T maps linearly independent vectors in to linearly independent vectors.

The next theorem relates the dimension of the kernel and image of a linear transformation L: V $\rightarrow$ W with the dimension of V. Before going to it let us have the following definition.

Definition 5.2.2: Let L be a linear transformation from a vector space V in to W over the field K.

(a) The dimension of the Kernel (the null space) of L is called the **nullity** of L.

(b) The dimension of the Image (the range) of L is called the **rank** of L.

Example 5.2.3: The linear transformation G: V $\rightarrow$ W with V = $\Re^3$ and $W = \Re^2$ of example 4.3.1 is given by $G(x, y, z) = (x + y, y - z)$.

We have seen that its kernel is a subspace of $\Re^3$ generated by (1, -1, 1). So dim (ker G) =1 . That is Nullity of G = 1. Its image is a subspace of $\Re^2$ with basis {(1, 0), (0, 1)}. So dim (ImG) = 2 = Rank of G. Since dim V= 3 and dim(ker G) + dim (ImG) = 3, we have dim V = Nullity of G + Rank of G.

Example 5.2.4: Consider the linear transformation $\mathbf{T}: \Re^2 \rightarrow \Re^3$ given in example 4.3.2. We have seen that its kernel contains only the zero vector, so dim (ker T) = 0. The set {(1, 1, 0), (2, 0, 1)} is a basis of its image, so dim (ImT) = 2. Again we have dim $\left(\Re^2\right)$ = Nullity of T + Rank of T

Theorem 5.2.4: (Rank-nullity theorem) Let V and W be vector spaces over the same field K. Let L: V $\rightarrow$ W be a linear transformation. If V is finite dimensional vector space then dim V = nullity of L + rank of L

i.e dim V = dim (ker L) + dim (ImL)

Proof: Since V is finite dimensional vector space, it is obvious that Ker L and ImL =L (V) are finite dimensional. Moreover dim (Ker L), dim (ImL) $\leq$ dim V (Verify)

Let $\{u_1, u_2, \ldots, u_p\}$ and $\{w_1, w_2, \ldots, w_q\}$ be basis of kerL and ImL respectively ($p, q \leq$ dim V) Then there exist $v_1, v_2, \ldots, v_q \in V$ such that $L(v_i) = w_i$ for i = 1, 2, 3,…q as $w_i \in$ ImL.

Claim: $\beta = \{u_1, u_2, \ldots, u_p, v_1, v_2, \ldots, v_q\}$ is a bais of V.

Now we show that

i) β generates V

ii) β is linearly independent.

i) Let $v \in V$. Then $L(v) \in$ ImL and hence there exist unique scalars $b_1, b_2, \ldots, b_q$ in K such that $L(v) = b_1w_1 + b_2w_2 + \ldots + b_qw_q$.

So $L(v) = b_1L(v_1) + b_2L(v_2) + \ldots + b_qL(v_q)$ as $L(v_i) = w_i$ for i = 1, 2, ..q

$= L(b_1v_1 + b_2v_2 + \ldots + b_qv_q)$

Thus $L(v-b_1v_1 - b_2v_2 - \ldots - b_qv_q) = O_w$. Which implies $v - b_1v_1 - b_2v_2 - \ldots - b_qv_q \in \ker L$.

Hence $v - b_1v_1 - b_2v_2 - \ldots - b_qv_q = a_1u_1 + a_2u_2 + \ldots + a_pu_p$ for some unique scalars $a_1, a_2, \ldots, a_p$ in K. Therefore $v = a_1u_1 + a_2u_2 + \ldots + a_pu_p + b_1v_1 + b_2v_2 + \ldots + b_qv_q$

From the this we conclude that β generates V.

ii) Suppose $\alpha_1u_1 + \alpha_2u_2 + \ldots + \alpha_pu_p + r_1v_1 + r_2v_2 + \ldots + r_qv_q = O_v$ …… (1)

where $\alpha_1, \alpha_2 \ldots, \alpha_p r_1, r_2, \ldots, r_q \in K$

Then $L(\alpha_1u_1 + \alpha_2u_2 + \ldots + \alpha_pu_p + r_1v_1 + r_2v_2 + \ldots + r_qv_q) = L(O_v) = O_w$

So we have $\alpha_1L(u_1) + \alpha_2L(u_2) + \ldots + \alpha_pL(u_p) + r_1L(v_1) + r_2L(v_2) + \ldots + r_qL(v_q) = O_w$

$\Rightarrow$ $r_1w_1 + r_2w_2 + \ldots + r_qw_q = O_w$ since $L(u_j) = O_w$ and

$L(v_i) = w_i \; \forall j = 1, 2, \ldots, p$ and $\forall_i = 1, 2, \ldots, q$

$\Rightarrow$ $r_1 = r_2 = \ldots = r_q = 0$, since $\{w_1, w_2, \ldots, w_q\}$ is basis of ImL.

$\Rightarrow$ $\alpha_1u_1 + \alpha_2u_2 + \ldots + \alpha_pu_p = O_v$ (replace $r_1, r_2 \ldots r_q$ by 0 in (1)

$\Rightarrow$ $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$, since $\{u_1, u_2, \ldots, u_p\}$ is a basis of Ker L.

Thus we have shown that

$\alpha_1u_1 + \alpha_2u_2 + \ldots + \alpha_pu_p + r_1v_1 + r_2v_2 + \ldots + r_qv_q = O_v$

$\Rightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_p = r_1 = v_2 = \ldots = r_q = 0$

That is β is a linearly independent set in V. From (i) and (ii) it follows that

$\beta = \{u_1, u_2, \ldots, u_p, v_1, v_2, \ldots, v_q\}$ is a basis of V.

Hence $\dim V = p + q = \dim(\ker L) + \dim(\text{Im}L)$

Therefore dim V = Nullity of L + Rank of L.

Example 5.2.5 : 1) Let $T: \Re^4 \rightarrow \Re^3$ be a linear transformation defined by

$T(a, b, c, d) = (a - b + c + d, a + 2c - d, a + b + 3c - 3d) \; \forall a, b, c, d \in \Re$

i) Find the rank and nullity of T.

ii) Is T one-to-one? Is T on to?

Solution: i) First let us find nullity of T.

Let $(x, y, z, w) \in \ker T$

Then T(x, y, z, w) = (0, 0, 0)

$\Rightarrow$ (x – y + z + w, x + 2z – w, x + y + 3z – 3w) = (0,0,0)

$$\Rightarrow \begin{cases} x - y + z + w = 0 \ \ldots (1) \\ x + 2z - w = 0 \quad \ldots (2) \\ x + y + 3z - 3w = 0 \ \ldots (3) \end{cases}$$

Adding (1) and (3) we get 2x + 4z – 2w = 0

By dividing both sides of this equation by 2, we get equation (2)

i.e. x + 2z – w = 0. So w = x + 2z, and y = x + z + w = 2x + 3z

Thus (x, y, z, w) = (x, 2x + 3z, z, x + 2z)

= (x, 2x, 0, x) + (0, 3z, z, 2z)

= x(1, 2, 0, 1) + z(0, 3, 1, 2)

Hence the vectors (1, 2, 0, 1) and (0, 3, 1, 2) generate ker T as every (x, y, z, w) in ker T can be written as a linear combination of (1, 2, 0, 1) and (0, 3, 1, 2). Moreover they are linearly independent vectors in ker T. Therefore {(1, 2, 0, 1),(0, 3, 1, 2)} is a basis of ker T and hence dim ker T = 2. That is nullity of T is 2. Now by using rank-nullity theorem, we can easily determine the rank of T.

$\dim(\Re^4)$ = Nullity of T + rank of T

4 = 2 + Rank of T

Rank of T= 2

ii) T is not one-to-one because kerT $\neq \{0\}$. Since Rank of T = 2 which is different from $\dim(\Re^3)$, ImT $\neq \Re^3$. So T is not on to.

Example 5.2.6: Find a linear transformation T: $\Re^3 \to \Re^4$ whose range (image) is spanned by (1, 2, 0, -4), (2, 0, -1, -3).

Solution: Given ImT is spanned by {(1, 2, 0, -4), (2, 0, -1, -3)}

Let us include a vector (0,0,0,0) in this set which will not affect the given spanning property. Now by using, the standard basis of $\Re^3$ and theorem 4. 2.2, one can get a linear transformation T: $\Re^3 \to \Re^4$ such that

T(1, 0, 0) = (1, 2, 0, -4), T(0, 1, 0) = (2, 0, -1, -3) and T(0,0,1) = (0,0,0,0)

If $(x, y, z) \in \Re^3$, $T(x, y, z) = T(x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1))$

$= x\,T(1, 0, 0) + yT(0, 1, 0) + zT(0, 0, 1)$

$= (x, 2x, 0, -4x) + (2y, 0, -y, -3y) + z(0, 0, 0, 0)$

Therefore $T(x, y, z) = (x + 2y, 2x, -y, -4x-3y)$ is the required linear transformation.

/

Example 5.2.7: Let $L: \Re^3 \to \Re^3$ be a linear transformation defined by

$L(x, y, z) = (x + y+ z, 2x + y – z, 3x + 2y)$.

i) Find a basis for ker L

ii) Find a basis for ImL iii) Verify rank nullity theorem for L

Solution: i) From the definition of kerL we have,

$(a, b, c) \in \ker L \Leftrightarrow L(a, b, c) = (0, 0, 0)$

$\Leftrightarrow (a + b + c, 2a + b – c, 3a + 2b) = (0, 0, 0)$

$$\Leftrightarrow \begin{cases} a + b + c = 0 \\ 2a + b - c = 0 \\ 3a + 2b = 0 \end{cases}$$

From $3a + 2b = 0$, we get $b = \frac{-3}{2}a$

Substituting this in $a + b + c = 0$ and $2a + b – c = 0$ we get $c = \frac{a}{2}$.

Thus any element (a, b, c) of ker L, can be expressed as

$$(a, b, c) = \left(a, \frac{-3}{2}a, \frac{a}{2}\right) = a\left(1, \frac{-3}{2}, \frac{1}{2}\right)$$

That is $\left(1, \frac{-3}{2}, \frac{1}{2}\right)$ generates ker L. So $\left\{\left(1, \frac{-3}{2}, \frac{1}{2}\right)\right\}$ is a basis of ker L

ii) Again by using the definition for ImL, we have

$(u, v, w) \in \text{ImL} \Leftrightarrow (u, v, w) = L(x, y, z)$ for some $(x, y, z) \in \Re^3$

$\Rightarrow$ $(u, v, w) = (x + y + z, 2x + y – z, 3x + 2y)$, where $x, y, z \in \Re$

$$\Rightarrow \quad \begin{cases} x + y + z = u \\ 2x + y - z = v \\ 3x + 2y = w \end{cases}$$

By adding the first two equations we get, $3x + 2y = u + v$. But $3x + 2y = w$. So $w = u + v$

Thus any vector (u, v, w) in ImL can be written as

$$\begin{aligned} (u, v, w) &= (u, v, u + v) \; ; \; u, v \in \Re \\ &= (u, o, u) + (o, v, v) \; ; \; u, v \in \Re \\ &= u(1, 0, 1) + v(0, 1, 1) \; ; \; u, v \in \Re \end{aligned}$$

Hence every vector in ImL is a linear combination of (1, 0, 1) and (0, 1, 1).

So {(1, 0, 1), (0, 1, 1)} generates ImL. Moreover (1, 0, 1) and (0, 1, 1) are linearly independent (verify). Consequently {(1, 0, 1), (0, 1, 1)} is a basis of ImL.

iii) dim(kerL) = 1 as its basis contains only one non-zero vector, dim(ImL) = 2 as its basis contains two vectors. dim (kerL) + dim (ImL) = 1 + 2 = 3

i.e. nullity of L + rank of L = dim($\Re^3$) as stated in the theorem.

Recall that a linear transformation T: $V \to W$

i) is one-to-one iff ker T = $\{0_v\}$

ii) is on to iff ImT = W

Now suppose V and W are respectively n and m dimensional vector spaces over the same field K. If n > m, is there a one-to-one linear transformation from V in to W?

Suppose there is, then ker T = $\{0_v\}$ and hence dim kerT = 0. So dimV = dim (ImT) by using rank-nullity theorem. Thus dim (ImT) > dim W. But this is not possible as ImT is a subspace of W. Therefore there is no one-to-one linear transformation T from V in to W if dim V > dimW.

If n < m, does these exist an on to linear transformation T: $V \to W$?

If it exists ImT = W and dimW = dim (TmT). Thus form rank-nullity theorem it follows that n = dim ker T + m, which implies dim ker T = n – m < 0 which is not possible. Therefore there does not exist a linear transformation T : $V \to W$ which is on to if dim V is less than dim W.

Is there a one – to – one linear mapping from $\Re^3$ to $\Re^2$?

Is there an on to linear mapping from $\Re^3$ to $\Re^4$?

Exercise:

1 For each of the following linear transformation, find a basis and dimension of its image and kernel.

i) $T : \Re^3 \rightarrow \Re^3$ defined by $T(x, y, z) = (x, x + 2y, y)$

ii) $F : \Re^3 \rightarrow \Re^2$ defined by $F(a, b, c) = (a, c)$

iii) $L : \Re^4 \rightarrow \Re^3$ defined by $L(x, y, z, t) = (x – y + z + t, x + 2z – t, x + y + 3z- 3t)$

iv) $T_A : \Re^4 \rightarrow \Re^3$ defined by $T_A(X) = A.X$ where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ -2 & -1 & 2 & 1 \\ 1 & -4 & -4 & 2 \end{pmatrix}$$ and X is a column vector in $\Re^4$

2) Find a linear transformation $T : \Re^3 \rightarrow \Re^3$ such that $\{(x, y, z): 4x – 3y + z = 0\}$ is the i) Kernel of T ii) image of T

3) Find a linear transformation $F : \Re^4 \rightarrow \Re^3$ whose kernel is generated by $(1, 2, 3, 4)$ and $(1, 1, 0, 1)$.

4) Find a linear transformation $T : \Re^3 \rightarrow \Re^4$ whose image is generated by $(1, -1, 2, 3)$ and $(2, 3, 0, 1)$.

5) Let $L: V \rightarrow W$ be a liner transformation. Let $w \in W$ and $v_o \in V$ such that $L(v_o) = w$. Show that $L(x) = w$ iff $x = v_o + u$ where u is an element of the kernel of L.

6) Let the linear transformation $T : \Re^3 \rightarrow \Re^3$ be defined by $T(u, v, w) = (u + v – 2w, u + 2v – w, 2u + v)$ Find the rank and nullity of T.

7. Show that the linear transformation $T : \Re^3 \rightarrow \Re^3$ for which $T(e_1) = e_1 + e_2$, $T(e_2) = 2e_2 + e_3$ and $T(e_3) = e_1 + e_2 + e_3$ is both one-to-one and on to.

8. Let the linear transformation from $\Re^3$ in to $\Re^3$ be defined by

   $T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3)$

   i) If (a, b, c) is a vector in $\Re^3$, what are the conditions on a, b, and c so that the vector be in the range of T? What is the rank of T ?

   ii) What are the conditions on a, b, and c so that (a, b, c) be in the null space of T? What is the nullity of T?

   iii) Is T one-to-one? Why?
   Is T on-to? Why?

5.3. Algebra of Linear transformations

In the study of linear transformation from vector space V in to W, it is of fundamental importance that the set of these transformations possesses a natural vector space structure.

The set of linear transformations from a vector space V into itself has even more algebraic structure because of ordinary composition of function, provides a multiplication of such transformations. We shall explore these ideas in this section.

Notation: We denote the set of all linear transformations from a vector space V in to W over the field K by L(V, W). While using the notation L(V, W), it should be noticed that V and W are vector spaces over the same field . Let $T, S \in L(V, W)$ and $\alpha \in K$. Then T and S are linear transformations from V in to W and hence they are functions from V in to W. Thus

i) the sum of T and S, T + S is a function from V in to W defined by $(T + S)(v) = T(v) + S(v)\ \forall v \in V$

ii) the scalar multiple of T by α, αT is a function from V in to W defined by $(\alpha T)(v) = \alpha(T(v))\ \forall v \in V$.

We now state a theorem that asserts $T + S$ and αT are linear transformations form V in to W for any $T, S \in L(V, W)$ and $\alpha \in K$; and L (V, W) is a vector space with addition and multiplication by scalars defined as above.

We shall prove the first two assertions and leave the last one as an exercise.

Theorem 5.3.1: Let V and W be vector spaces over the field K. Let T and S be linear transformations from V in to W and $\alpha \in K$. Then

i) the function $T + S$ is a linear transformation from V in to W.
   i.e $T + S \in L(V, W)$.

ii) the function αT is a linear transformation from V in to W.
   i.e $\alpha T \in L(V, W)$.

iii) L (V, W) the set of all linear transformations from V in to W with respect to the operation of vector addition and scalar multiplication defined as $(T + S)(v) = T(v) + S(v)$ and $(\alpha T(v) = \alpha T(v)$ for all $v \in V$, is a vector space over K.

Proof: i) We need to show that $T + S$ is a linear transformation from V in to W for any linear transformation T and S from V in to W. For this let $v_1, v_2 \in V$ and $\beta \in K$.

Then $(T + S)(v_1 + v_2) = T(v_1 + v_2) + S(v_1 + v_2)$ why?

$= T(v_1) + T(v_2) + S(v_1) + Sv_2)$ why?

$= T(v_1) + S(v_1) + T(v_2) + S(v_2)$ why?

$= (T + S)(v_1) + (T + S)(v_2)$ why?

Moreover $(T + S)(\beta v_1) = T(\beta v_1) + S(\beta v_1)$

$= \beta T(v_1) + \beta S(v_1)$ why?

$= \beta(T(v_1) + S(v_1))$ why?

$= \beta(T + S)(v_1)$

Therefore $T + S$ is a linear transformation.

ii) Let T be a linear transformation form V into W and $\alpha \in K$

Then for any $v_1, v_2 \in V$ and $\lambda \in K$.

We have $(\alpha T)(v_1 + v_2) = \alpha(T(v_1 + v_2))$

$= \alpha(T(v_1) + T(v_2))$ why?

$= \alpha(T(v_1)) + \alpha(T(v_2))$ why?

$= (\alpha T)(v_1) + (\alpha T)(v_2)$

and $(\alpha T)(\lambda v_1) = \alpha(T(\lambda v_1))$

$= \alpha(\lambda T(v_1))$ why?

$= (\alpha\lambda)(T(v_1))$

$= (\lambda\alpha)(T(v_1))$

$= \lambda(\alpha(T(v_1)))$

$= \lambda(\alpha T)(v_1)$

Therefore αT is a linear transformation.

iii) Here we need to verify that all the conditions in the definition of a vector space are satisfied. We leave this to the students.

Note: In the above theorem, we have asserted that L(V,W) is a vector space over K. With this we can consider every linear transformation in L(V, W) as a vector. The zero vector in this vector space will be the zero transformation that sends every vector of V in to the zero vector in W.

Example 5.3.1: Let $T : \Re^3 \to \Re^2$ and $H : \Re^3 \to \Re^2$ be linear transformations given by

$T(x, y, z) = (3x, y + z)$ and $H(x, y, z) = (2x – z, y)$

a. Are T + H and 3T – 4H linear transformations form $\Re^3$ in to $\Re^2$

b. Find (i) T + H (c) 3T – 4H

Solution: a) Since T and H are linear transformation from $\Re^3$ in to $\Re^2$, we have $T, H \in L(\Re^3, \Re^2)$ and hence $T + H, 3T – 4H \in L(\Re^3, \Re^2)$ as $L(\Re^3, \Re^2)$ is a vector space over $\Re$. So T + H and 3T – 4H are linear transformations from $\Re^3$ in to $\Re^2$.

b) i) $(T + H)(x, y, z) = T(x, y, z) + H(x, y, z)$

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

$= (5x – z, 2y + z)$

ii) $(3T - 4H)(x, y, z) = (3T + (-4H))(x, y, z)$

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

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

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

$= (9x, 3y + 3z) + (-8x + 4z, -4y)$

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

Theorem 5.3.2: Let V, W and Z be vector spaces over the field K. Let T and S be linear transformations from V in to W and from W into Z respectively. Then the compose function SoT defined by $(SoT)(v) = S(T(v))$ for all $v \in V$ is a linear transformation.

Proof: Let T and S be as in the hypothesis of the theorem Let $v_1, v_2 \in V$ and $r \in K$

Then $(SoT)(v_1 + v_2) = S(T(v_1 + v_2))$

$= S(T(v_1) + T(v_2))$ why?

$= S(T(v_1)) + S(T(v_2))$ why?

$= (SoT)(v_1) + (SoT)(v_2)$

and $(SoT)(rv_1) = S(T(rv_1))$

$= S(rT(v_1))$ because T is a linear transformation.

$= rS(T(v_1))$ why?

$= r(SoT)(v_1)$

Therefore SoT is a linear transformation.

Notation: For the sake of brevity we shall simply denote the composition SoT of S and T by ST.

Activity 5.3.1:

1) Give two linear transformation S and T from $\Re^3$ in to it self such that $ST \neq TS$. From this you may conclude that composition of linear transformations is not commutative.
2) Is composition of linear transformations associative? Justify your answer!
3) Let U, V and W be vector spaces over the field K.
   Let $S, S' \in L(U,V)$ and $T, T' \in L(V, W)$

Verify each of the following.

i) $T(S + S') = TS + TS'$

ii) $(T + T')S = TS + T'S$

iii) $\alpha(T S) = (\alpha T)S = T(\alpha S)$ for all $\alpha \in K$

Remember that by a linear operator on vector space V we mean a linear transformation from a vector space V in to itself.

If T is a linear operator on vector space V, we can compose T with T to get TT which is again a linear operator on V. We shall use the notation $T^2 = TT$, $T^3 = TTT$ and in general $T^n = T.T \ldots T$ (the composite of T with itself n times) for n = 1, 2, 3, …. We define $T^o = I$ (identity mapping) if $T \neq 0$ (the zero mapping).

Example 5.3.2: Let T be a linear operator on a vector space V over the field F.

(i) If $T^2 = O$ the zero mapping then what can you say about the relation of the range of T to the kernel of T?

ii) Give an example of a linear operator on $\Re^2$ such that $T^2 = O$ but $T \neq O$.

Solution: i) Method 1: $T^2 = O \Rightarrow T^2(v) = O(v), \ \forall v \in V$

$\Rightarrow T(T(v)) = O_v$ – zero vector in V.

$\Rightarrow T(v) \in \ker T \ \forall v \in V$

But T(v) is an arbitrary element of Range of T for any $v \in V$.

Therefore Range of $T \subseteq \ker T$.

Method 2: Let $w \in$ Range of T. Then $T(v) = w$ for some $v \in V$.

So $T(w) = T(T(v)) = T^2(v) = O_v$ as

$T^2(v) = O_v \ \forall v \in V$. Hence $T(v) = w \in \ker T$

Therefore, Range of T is a subset of kerT

ii) Consider the linear map $T: \Re^2 \rightarrow \Re^2$ such that

$T(a, b) = (0, a) \ \forall (a, b) \in \Re^2$

Then $T \neq O$ as $T(1, 1) = (0, 1) \neq (0, 0)$

Now $T^2 (a, b) = (T.T) (a, b)$

$= T(T(a, b))$

$= T(0, a)$

$= (0, 0)$

So, $T^2(a, b) = (0, 0) \ \forall (a, b) \in \Re^2$

Hence, $T^2 = O$

Example 5.3.3: Let $T:\Re^2 \to \Re^2$ be a linear operator on $\Re^2$ given by $T(a, b) = (a + b, a + b)$. Show that $T(T - 2I) = O$, where I is the identity mapping on $\Re^2$ and O is the zero mapping on $\Re^2$.

Solution: $(T - 2I)(x, y) = T(x, y) - 2I(x, y)$

$= (x + y, x + y) - 2(x, y) = (y - x, x - y)$

So, $[T(T - 2I)](x, y) = T((T - 2I)(x, y))$

$= T(y - x, x - y)$

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

$= (0, 0)$

Thus $T(T - 2I) = O$

Activity 5.3.2:

1. Let T and S be linear operators on $\Re^2$. Does TS = O imply either T = O or S = O? Explain! Give a counter example if your answer is No.
2. Let V be finite dimensional vector space over the field F and T be a linear operator on V. Suppose that rank (T^2) = rank T. Prove that the range and null space of T have only the zero vector in common.

Now let us discuss about inverse of a linear transformation. Recall that a function T from V in to W is called invertible if there exists a functions S from W in to V such that ST is the identity function on V and TS is the identity function on W. If T is invertible the function S is unique and is denoted by $T^{-1}$and is called the inverse of T.

$T^{-1}(w) = v \Leftrightarrow T(v) = w$ whenever T^{-1} exists. Furthermore we know that T is invertible iff T is one- to- one and on to.

Theorem 5.3.3: Let V and W be vector spaces over the field F and Let T be a linear transformation from V in to W. If T is invertible then the inverse function T^{-1} is a linear transformation.

Proof: Suppose $T : V \to W$ is invertible then there exists a unique function $T^{-1}: W \to V$ such that $T^{-1}(w) = v \Leftrightarrow T(v) = w$ for all $w \in W$. Moreover T is 1-1 and on to. We need to show that T^{-1} is linear. Let $w_1, w_2 \in W$ and $\alpha \in F$ Then there exist unique vectors $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$ as T is 1-1 and on to. So $T^{-1}(w_1) = v_1$ and $T^{-1}(w_2) = v_2$.

$T(v_1 + v_2) = T(v_1) + T(v_2)$ because T is linear

$= w_1 + w_2$

Thus $T^{-1}(w_1 + w_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2)$ for any $w_1, w_2 \in W$.

Since $T(\alpha v_1) = \alpha w_1$, $T^{-1}(\alpha w_1) = \alpha v_1 = \alpha T^{-1}(w_1)$. Therefore T^{-1} is a linear transformation.

Remark: Let $T : V \to W$ and $S : W \to Z$ be invertible linear transformations. Then

i) $T^{-1} : W \to V$ is also invertible and $\left(T^{-1}\right)^{-1} = T$

ii) $ST : V \to Z$ is also invertible and $(ST)^{-1} = T^{-1}S^{-1}$

Example 5.3.4: Let $T : \Re^3 \to \Re^3$ be a linear transformation given by $T(x, y, z) = (x, 3y, z)$. Is T invertible? If so, find a rule for T^{-1} like the one which defines T.

Solution: We know that if kerT = {O} then T is one to-one

$$\begin{aligned} \text{KerT} &= \{(a,b,c) \in \Re^3 \mid T(a,b,c) = (0,0,0)\} \\ &= \{(a,b,c) \in \Re^3 \mid (a,3b,c) = (0,0,0)\} \\ &= \{(a,b,c) \in \Re^3 \mid a = 0,\ 3b = 0 \ \text{and}\ c = 0\} \\ &= \{(0,0,0)\}. \end{aligned}$$

Therefore T is one to one.

Moreover from rank-nullity theorem, $\dim(\Re^3) = \dim(\ker T) + \dim(\text{Im}T)$.

$3 = 0 + \dim(\text{Im}T)$. So $\dim(\text{Im}T) = 3$. Since ImT is a subspace of $\Re^3$

and dim(ImT) = 3, we have ImT = $\Re^3$. Hence T is on to as ImT = $\Re^3$.

Therefore T is invertible as it one to- one and on to.

To find T^{-1}, let T (x, y, z) = (u, v, w).

Then (x, 3y, z) = (u, v, w). So x = u, $y = \frac{v}{3}$, $z = w$.

But T(x, y, z) = (u, v, w) $\Leftrightarrow$ T^{-1} (u, v, w) = (x, y, z).

Thus T^{-1} (u, v, w) = $\left(\mathbf{u}, \frac{\mathbf{v}}{3}, \mathbf{w}\right)$.

Whenever V is a finite dimensional vector space, one – to – oneness can be related to linear independence and rank as stated in the theorem below.

Theorem 5.3.4: Let T : V $\rightarrow$ W be a linear transformation and V be finite dimensional vector space with dimV = n. Then the following statements are equivalent.

i) T is 1 – 1

ii) If $v_1, v_2, v_3, \ldots, v_k$ are linearly independent vectors in V then $T(v_1), T(v_2), T(v_3), \ldots, T(v_k)$ are linearly independent vectors in ImT

iii) If $\{v_1, v_2, v_3, \ldots, v_n\}$ is a basis of V then $\{T(v_1), T(v_2), T(v_3), \ldots, T(v_n)\}$ is a basis of ImT.

iv) Dim (ImT) = n

Proof: Left for students

Activity 5.3.3:

Prove theorem 4.4.4. Follow the following steps

1. Show that (iv) $\Rightarrow$ (i)
2. Show that (i) $\Rightarrow$ (ii)
3. Show that (ii) $\Rightarrow$ (iii)
4. Show that (iii) $\Rightarrow$ (iv)
5. Form 1, 2, 3 and 4, can we conclude that (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv)? How?

Remark: For finite dimensional vector spaces V and W over the field K with dimV = dimW, we have the following results about any linear transformation $T : V \to W$.

a) T is $1 - 1 \Rightarrow$ T is invertible

b) T is on to $\Rightarrow$ T is invertible.

Example 5.3.5: Let $L : V \to V$ be a linear operator such that $L^2 = O$. Show that $I - L$ is invertible. (I is the identity mapping on V.)

Solution: Clearly $I - L$ is a linear operator on V.

Let $x \in \ker(I-L)$. Then $(I - L)(x) = O_v$

That $I(x) - L(x) = O_v$. So we have $L(x) = x$ and $L^2(x) = L(x)$.

But $L^2 = O$ and hence $L^2(x) = O_v$. Thus $x = L(x) = O_v$.

Therefore $\ker (I - L) = \{O_v\}$. Consequently $I - L$ is one to one. In this case it is not sufficient to have $I - L$ is $1 - 1$ to conclude that it is invertible as V can be infinite dimensional. (Nothing is assumed about V whether it is finite or infinite dimensional). So we need to show that $I - L$ is on to. For this let $v \in V$. Then $L(v) \in V$ and $L(v) + v \in V$. Moreover

$$\begin{aligned}(I - L)(L(v) + v) &= I(L(v)+v) - L(L(v) + v) \\ &= L(v) + v - L^2(v) - L(v) \\ &= v \qquad\qquad (L^2(v) = O_v)\end{aligned}$$

Thus $L(v) + v$ is the pre-image of v under $I - L$. So for any vector $v \in V$ there exists a vector $x = L(v) + v$ in V such that $(I - L)(x) = v$. From this it follows that $(I - L): V \to V$ is on to. Therefore $I - L$ is invertible as it is $1 - 1$ and on to.

Exercise:

1. Let T and S be linear operators on $\Re^2$ defined by $T(x, y) = (y, x)$ and $S(a, b) = (a, 0)$
   i) How do you describe T and S geometrically?
   ii) Give rules like the one defining T and S for each of the linear transformations $S - 2T$, ST, TS, T^2, S^2.

2. Let T be the linear operator on $\Re^3$ defined by
   $T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3)$. Is T invertible? If so, find a rule for T^{-1}.
3. For the linear operator T of exercise 2, Show that $(T^2 - I)(T - 3I) = O$
   (I – identitymapping and O zero mapping)
4. Let T be a linear transformation from $\Re^3$ in to $\Re^2$ and Let U be a linear transformation from $\Re^2$ in to $\Re^3$. Prove that the linear transformation UT is not invertible.
5. $L : \Re^3 \to \Re^3$ be a linear transformation. Show that L is invertible and find L^{-1} for:
   a) $L(x, y, z) = (x - y, \quad x + z, \quad y + 2z)$
   b) $L(x, y, z) = (2x - y + z, \; x + y, \quad 3x + y + z)$

6. a) Let $S : V \to V$ be a linear operator such that $S^2 - S + I = O$ (where I is the identity mapping on V and O is the zero mapping on V) Show that S^{-1} exists and is equal to I – S.

   b) Let T be a linear operator on a vector space V, and assume that $L^3(v) = O_v$ for all $v \in V$. Show that I – L is invertible.

7. Let F and G be linear operators on a vector space V over the set of real numbers.. Assume that FG = GF. Show that

   i) $(F + G)^2 = F^2 + 2FG + G^2$

   ii) $(F + G)(F - G) = F^2 - G^2$

5.4.Matrix representation of a linear transformation

In this section we shall investigate the strong relationship that exists between linear transformation and matrices.

Linear transformation associated with a given matrix

Let us consider a matrix $\boldsymbol{A = (a_{ij})_{m \times n}}$ over a field $\boldsymbol{K}$ (*i.e* $\boldsymbol{a_{ij} \in K}$ for each i = 1,2,…,m and for each j = 1, 2, …,n). We can associate a mapping $\boldsymbol{T_A : K^n \to K^m}$ by defining $\boldsymbol{T_A(X) = AX}$ for any column vector **X** in $\mathbf{K^n}$.

Activity 5.4.1: Show that T_A is a linear transformation

Definition 5.4.1: For any matrix $A=(a_{ij})_{m\times n}$ over a field K. The mapping $T_A : K^n \to K^m$ defined by $T_A(X)=AX \quad \forall X \in K^n$ is called the linear transformation associated with the matrix A.

Note: In this definition, each element **X** of $\mathbf{K^n}$ should be considered as column vector otherwise A**X** is not defined.

Example 5.4.1: a) Let $A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. Then $T_A : \Re^2 \to \Re^2$ is an identity linear transformation.

b) Let $B=\begin{pmatrix}1 & 0\\ 2 & 1\\ 3 & 0\end{pmatrix}$. Then $T_B\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}1 & 0\\ 2 & 1\\ 3 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\\ 2x+y\\ 3x\end{pmatrix}$

Thus $T_B:\Re^2 \to \Re^3$ given by $T_B\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\\ 2x+y\\ 3x\end{pmatrix}$ is the linear transformation associated with matrix B.

c) Let $A=\begin{bmatrix}1 & 3\\ 0 & 1\end{bmatrix}$, $u=\begin{pmatrix}0\\ 2\end{pmatrix}$, $v=\begin{pmatrix}2\\ 0\end{pmatrix}$, $w=\begin{pmatrix}2\\ 2\end{pmatrix}$ and T_A be the linear transformation associated with matrix A.

Find i) $T_A(u), T_A(v)$ and $T_A(w)$

ii) The image of the square with vertices $\begin{pmatrix}0\\ 0\end{pmatrix},\begin{pmatrix}2\\ 0\end{pmatrix},\begin{pmatrix}2\\ 2\end{pmatrix}$ and $\begin{pmatrix}0\\ 2\end{pmatrix}$.

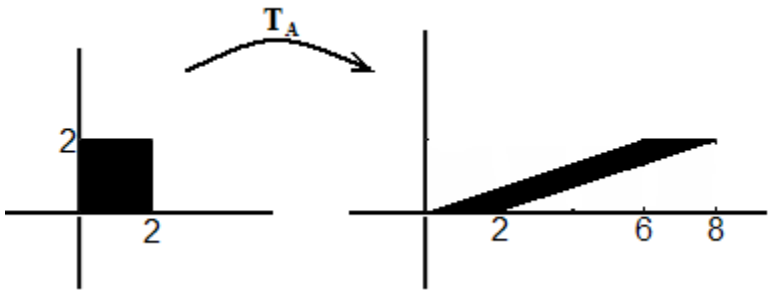
Solution:

i) $T_A(u)=Au=\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix}\begin{pmatrix}0\\ 2\end{pmatrix}=\begin{pmatrix}6\\ 2\end{pmatrix}$

$$T_A(v) = AV = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$T_A(w) = Aw = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$$

ii) T_A deforms the given square as if the top of the square were pushed to the right while the base is held fixed (see the figure below).



We call such transformation shear transformation.

Activity 5.4.2:

1. Let A be a 7×5 matrix. What must a and b be inorder to define $T : \Re^a \to \Re^b$ by $T(x) = AX$
2. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Give a geometrical description of the linear transformation $\mathbf{T_A}$ associated with A.
3. Let $B = \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$
   i) Find the linear transformation **T** associated with **B**.
   ii) Find **X** in $\Re^2$ whose image under $\mathbf{T_B}$ is **b**
   iii) Determine whether the system $BX = b$ has no solution, unique solution or many solutions

In view of definition 4.5.1 we can study system of linear equations with the help of linear transformation associated with the coefficient matrix of the system. Consider the system $\boldsymbol{AX = b}$ where $\boldsymbol{A}$ is an $\boldsymbol{m \times n}$ real matrix. Then $\boldsymbol{AX = b}$ iff $\boldsymbol{T_A(X) = b}$ where $\boldsymbol{T_A}$ is the linear transformation associated with matrix A. Thus the system AX = **b** has a solution iff **b** is in the range of $\mathbf{T_A}$. If there is exactly one element $\boldsymbol{X \in \Re^n}$ whose image is **b** under $\mathbf{T_A}$ then the system $\boldsymbol{AX = b}$ has exactly one solution. But if **b** has more than one pre-images under $\boldsymbol{T_A}$ then the system has more than one solution. If there is no $\boldsymbol{X \in \Re^n}$ such that $\boldsymbol{T_A(X) = b}$, (i.e b is not in the range of $\boldsymbol{T_A}$), then the system has no solution.

Activity 5.4.3:

Prove that if ker $T_A = \{O\}$, then the system $\boldsymbol{AX = b}$ has at most one solution.

Further if $b = O$ the zero column vector in $\Re^m$ then the homogeneous system $AX = O$ has at least one solution. What is this solution?

The solution set of $AX = O$ is the kernel of $\boldsymbol{T_A}$. So the solution set of $AX = O$ is a subspace of $\Re^n$. Suppose $\dim(\boldsymbol{ker\, T_A}) = \boldsymbol{k}$ and $\{\boldsymbol{v_1, v_2, v_3, ..., v_k}\}$ is a basis of ker $\boldsymbol{T_A}$. Then any solution v of $\boldsymbol{AX = 0}$ can be expressed as

$$\boldsymbol{v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + ... + \alpha_k v_k} \text{ where } \alpha_1, \alpha_2, \alpha_3, ..., \alpha_k \text{ are scalars.}$$

Now let X_o be one particular solution of the non-homogeneous system $\boldsymbol{AX = b}$ and w be any solution of $\boldsymbol{AX = b}$.

Then $\boldsymbol{AX_o = b = Aw}$ so we have, $\boldsymbol{T_A(X_o) = T_A(w)}$. This in turn implies $\boldsymbol{T_A(X_o - w) = 0}$ and hence $\boldsymbol{w - X_o \in ker\, T_A}$. Thus $\boldsymbol{w - x_o = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k}$ or $w = x_o + \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k$ where $\alpha_1, \alpha_2, ..., \alpha_k$ are scalars.

Therefore if X_o is one particular solution of the non-homogeneous system $\boldsymbol{AX = b}$ then every solution w of $\boldsymbol{AX = b}$ is given by $w = x_o + \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 ... + \alpha_k v_k$ where $\{\boldsymbol{v_1, v_2, v_3, ..., v_k}\}$ is a basis ker $\boldsymbol{T_A}$ and $\boldsymbol{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_k}$ are scalars.

This called the **general solution** of the non homogeneous system $\boldsymbol{AX = b}$.

Example 5.4.2

Let $\boldsymbol{A}=\begin{pmatrix} 1 & 0 & 3 \\ -2 & 1 & -3 \end{pmatrix}$, $\boldsymbol{b}=\begin{pmatrix} -4 \\ 9 \end{pmatrix}$ and $\boldsymbol{T_A}$ be the linear transformation associated with matrix A.

i) Find ker $\boldsymbol{T_A}$?

ii) Is $b \in \mathrm{Im}T_A$?

iii) Is there more than one $\boldsymbol{X}$ whose image under $\boldsymbol{T_A}$ is **b**?

iv) Describe the solution set of $\boldsymbol{AX} = \boldsymbol{b}$

Solution: $\boldsymbol{T_A}: \Re^3 \to \Re^2$ is given by

$$\boldsymbol{T_A}\begin{pmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \boldsymbol{x_3} \end{pmatrix}=\begin{pmatrix} 1 & 0 & 3 \\ -2 & 1 & -3 \end{pmatrix}\begin{pmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \boldsymbol{x_3} \end{pmatrix}=\begin{pmatrix} \boldsymbol{x_1}+3\boldsymbol{x_3} \\ -2\boldsymbol{x_1}+\boldsymbol{x_2}-3\boldsymbol{x_3} \end{pmatrix}$$

i) $\ker T_A = \{X \in \Re^3 | T_A(X) = 0\}$

$$\text{So } \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \\ \boldsymbol{c} \end{pmatrix} \in \boldsymbol{ker\, T_A} \Leftrightarrow \begin{pmatrix} \boldsymbol{a}+3\boldsymbol{c} \\ -2\boldsymbol{a}+\boldsymbol{b}-3\boldsymbol{c} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} a+3c=0 \\ -2a+b-3c=0 \end{matrix}\Big\}$$

$$\Leftrightarrow a=-3c=b$$

$$\text{Thus } \boldsymbol{ker\, T_A}=\left\{\begin{pmatrix} -3\boldsymbol{c} \\ -3\boldsymbol{c} \\ \boldsymbol{c} \end{pmatrix} \middle| \boldsymbol{c} \in \Re\right\}=\left\{\boldsymbol{c}\begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix} \middle| \boldsymbol{c} \in \Re\right\}$$

ii) Since dim(ker $\boldsymbol{T_A}$) = 1 and 3 = dim (ker $\boldsymbol{T_A}$) + dim(Im$\boldsymbol{T_A}$),

dim(Im$\boldsymbol{T_A}$) = 2. Hence T_A is on to. So every element in $\Re^2$ has a pre-image under $\boldsymbol{T_A}$. Consequently $\boldsymbol{b} \in \boldsymbol{Im\, T_A}$ as $\boldsymbol{b} \in \Re^2$.

iii) Since **b** has a pre-image under T_A , there exists $x_o \in \Re^3$ such that $T_A(x_o) = b$. Moreover for any $\alpha \in \Re$, $w = x_o + \alpha \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix}$ is a pre-image of **b**, as

$$T_A(w) = T_A(x_o + \alpha \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix} = T_A(x_o) + \alpha T_A \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix} \quad = \quad b + \alpha.\mathrm{O} = \mathbf{b}$$

Thus we have more than one element in $\Re^3$ whose image under T_A is **b**.

iv) $\begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix}$ is a particular solution of $AX = b$

as $\begin{pmatrix} 1 & 0 & 3 \\ -2 & 1 & -3 \end{pmatrix} \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \end{pmatrix}$.

Therefore, $\left\{ \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix} \;\middle|\; q \in \Re \right\}$ is the solution set of $AX = b$

Exercise:

1) Let $A = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & -5 \\ -4 & 2 & 1 \end{pmatrix}$ and T_A be the linear transformation associated with matrix A. Find X such that $T_A(X) = \mathrm{O}$

2) Let $A = \begin{pmatrix} 1 & 3 & 4 & -3 \\ 0 & 1 & 3 & -2 \\ 3 & 7 & 6 & -5 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ -1 \\ 7 \end{pmatrix}$ and T_A be the linear transformation associated with matrix A.

i) Find ker T_A ii) Is **b** in the range of T_A

iii) Describe the solution set of $AX = b$

3) Suppose $T: \Re^5 \to \Re^2$ and $T(X) = AX$ for some matrix A and each X in $\Re^5$. How many rows and columns does A have?

4) Use a rectangular coordinate system to plot $\begin{pmatrix} 4 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ 2 \end{pmatrix}$, and their images under the linear transformation T_A associated with matrix A. (Make a separate sketch for each exercise). Give a geometric description of what T_A does to a vector $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ in $\Re^2$ where

i) $A = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$ ii) $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

iii) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ iv) $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

The matrix of a linear transformation

In the previous subsection we have seen that associated to any given $m \times n$ matrix A there is a linear transformation $T_A: \Re^n \to \Re^m$ defined by $T_A(X) = AX$. In this subsection we shall see the reverse process, that is to find a matrix associated to a given linear transformation from a finite dimensional vector space **V** in to finite dimensional vector space **W** over the same field **K**.

Before going further let us recall about the coordinates of an element **V** of a finite dimensional vector space V with respect to a given ordered basis β of V. What do we mean by ordered basis? Suppose $\beta = \{b_1, b_2, ..., b_n\}$ is an ordered basis for a finite dimensional vector space V over a field K and **v** is in V. The **coordinates of v relative to the basis** β (or the β **coordinates of v**) are the scalars $c_1, c_2, ..., c_n$ in K such that $v = c_1 b_1 + c_2 b_2 + ... + c_n b_n$.

If $c_1, c_2, \ldots, c_n$ are the β -coordinates of **v** then the vector in K^n,

$$[v]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix}$$

is the **coordinate vector of v relative to** β or the β **-coordinate vector of v.**

Example 5.4.3: (i) The coordinates of (x, y, z) relative the standard basis

$\{(1,0,0),(0,1,0),(0,0,1)\}$ of $\Re^3$ are simply x, y and z, since

(x, y, z) = x(1, 0, 0)+ y(0, 1, 0+ z(0, 0, 1).

ii) Consider a basis β = {$\mathbf{b_1}$, $\mathbf{b_2}$} of $\Re^2$, where $b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The coordinate vector of $X = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ relative to β is $[X]_\beta = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$,

since X = (-2)$\mathbf{b_1}$ + 3$\mathbf{b_2}$.

Activity 5.4.4:

1) Find the vector X determined by the coordinate vector $[X]_\beta = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ where

$$\beta = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}$$

2) Find the coordinate vector $[X]_\beta$ of $X = \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}$ relative to the basis

$$\beta = \left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\} \text{ of } \Re^3$$

3) Let P_2 be the space of polynomial functions from $\Re$ in to $\Re$ of degree two or less and P be an element of P_2 defined by $P(t) = 1 + 4t + 7t^2$

   a) Find the coordinate vector of P relative to the standard basis $\beta = \{f_0, f_1, f_2\}$ where $f_0(t) = 1$, $f_1(t) = t$ and $f_2(t) = t^2$.

   b) Find the coordinate vector of P relative to the basis $\beta' = \{g_0, g_1, g_2\}$ where $g_0(t) = 1 + t^2$, $g_1(t) = t + t^2$ and $g_2(t) = 1 + 2t + t^2$

Now let us deal with the matrix of a given linear transformation.

Let V be an n-dimensional vector space over the field K and let W be an m-dimensional vector space over K. Let $\beta = \{v_1, v_2, ..., v_n\}$ and $\beta' = \{w_1, w_2, ..., w_m\}$ be ordered bases of V and W respectively. Suppose $T : V \to W$ is a linear transformation.

Then for any x in V, $T(x) \in W$ and $T(x)$ can be expressed as a linear combination of elements of the basis β'. So we have,

$$\begin{aligned}
T(v_1) &= a_{11}w_1 + a_{21}w_2 + \dots + a_{m1}w_m \\
T(v_2) &= a_{12}w_1 + a_{22}w_2 + \dots + a_{m2}w_m \\
&\vdots \\
T(v_j) &= a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m \\
&\vdots \\
T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \dots + a_{mn}w_m
\end{aligned} \qquad (1)$$

where a_{ij}'s are scalars in K.

Writing the coordinates of $T(v_1), T(v_2), \dots, T(v_n)$ successively as columns of a matrix we get

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ . & . & & . & & . \\ . & . & & . & & . \\ . & . & & . & & . \\ a_{i1} & a_{i2} & & a_{ij} & & a_{in} \\ . & . & & . & & . \\ . & . & & . & & . \\ . & . & & . & & . \\ a_{m1} & a_{m2} & & a_{mj} & & a_{mn} \end{bmatrix}_{m\times n} \quad \dots. \qquad (2)$$

That is $M = \left[\left[T(v_1)\right]_{\beta'}\left[T(v_2)\right]_{\beta'} \dots\left[T(v_j)\right]_{\beta'}\dots\left[T(v_n)_{\beta'}\right]\right]$

The matrix M in (2) is called a **matrix representation of T** or **the matrix for T relative to the bases β and β'.** If β and β' are standard bases of V and W respectively, we call the matrix M in (2) the **standard matrix** for the linear transformation T.

Example 5.4.4: Define $T : \Re^3 \to \Re^2$ by

$$T(x_1, x_2, x_3) = (3x_1 - 2x_2 + x_3, - x_1 + x_2 + 5x_3)$$

a) Find the standard matrix for T.

b) Find the matrix of T relative to the ordered bases

β = {(1,0,1), (0,1,1), (0,0,1)} and β' = {(1,0), (1,1)} of $\Re^3$ and $\Re^2$ respectively.

Solution: a) Here we use the standard basis of $\Re^3$ and $\Re^2$.

T(1,0,0) = (3, -1) = 3(1,0) + -1(0,1)

T(0,1,0) = (-2,1) = -2(1,0) + 1(0,1)

T(0,0,1) = (1,5) = 1(1,0) + 5(0,1)

Writing the coordinates of T(1,0,0), T(0,1,0), T(0,0,1) as first, second and third columns of a matrix we get the standard matrix M for T as

$$M = \begin{bmatrix} 3 & -2 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$

b) T(1,0,1) = (4,4) = 0(1,0) + 4(1,1)

T(0,1,1) = (-1,6) = -7(1,0) + 6(1,1)

T(0,0,1) = (1,5) = -4(1,0) + 5(1,1)

The matrix of T relative to β and β' is $M = \begin{bmatrix} 0 & -7 & -4 \\ 4 & 6 & 5 \end{bmatrix}$

Our next task is to examine how the matrix M in (2) determines the linear transformation T. If $x = x_1v_1 + x_2v_2 + \ldots + x_nv_n$ is a vector in V, then the coordinate vector of x relative to

$$\beta = [X]_\beta = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

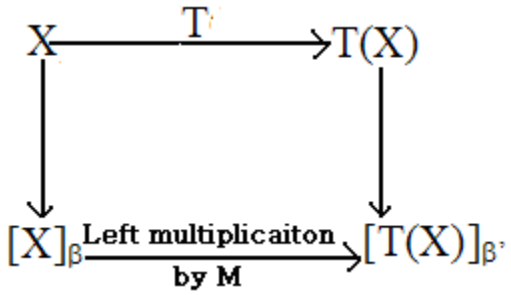
and $T(x) = T(x_1v_1 + x_2v_2 + \ldots + x_nv_n) = x_1T(v_1) + x_2T(v_2) + \ldots + x_nT(v_n)$ ….. (3)

Using the basis β' in W, we can rewrite (3) in terms of coordinate vectors relative to β' as $[T(x)]_{\beta'} = x_1[T(v_1)]_{\beta'} + x_2[T(v_2)]_{\beta'} + \ldots + x_n[T(v_n)]_{\beta'}$ … (4)

Further the vector equation (4) can be written as a matrix equation?

$$[T(X)]_{\beta'} = M[X]_\beta \quad \ldots. \qquad (5)$$

Thus if $[X]_\beta$ is the coordinate vector of x relative to β, then the equation in (5) shows that $M[X]_\beta$ is the coordinate vector of the vector **T(X)** relative to β'.



Note: In case when W is the same as V and the basis β' is the same as β, the matrix M in (2) is called the **matrix for T relative to** β and is denoted by $[T]_\beta$.

Activity 5.4.5: Using equation (3), verify equations (4) and (5).

Example 5.4.5: Let $\boldsymbol{T}: \Re^3 \to \Re^2$ be the linear transformation defined by T(x, y, z) = (3x + 2y - 4z, x -5y + 3z). Find the matrix of T relative to the bases $B_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)$ of $\Re^3$ and $B_2 = \{(1, 3), (2, 5)\}$ of $\Re^2$.

Solution: T(1, 1, 1) = (1, -1) = -7(1, 3) + 4(2, 5)

T(1, 1, 0) = (5, -4) = -33(1, 3) + 19(2, 5)

T(1, 0, 0) = (3, 1) = -13(1, 3) + 8(2, 5)

The matrix M of T relative to the bases B_1 and B_2 is:

$$\boldsymbol{M} = \begin{bmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{bmatrix}$$

Example 5.4.6: Let $\beta = \{b_1, b_2, b_3\}$ be a basis for a vector space V over the set of real numbers. Find $T(3b_1 - 4b_2)$, where T is a linear transformation form V in to V whose matrix relative to $\boldsymbol{\beta}$ is

$$[\boldsymbol{T}]_{\boldsymbol{\beta}} = \begin{bmatrix} 0 & -6 & 1 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

Solution: Let $x = 3b_1 - 4b_2$. Then the coordinate vector of x relative to $\boldsymbol{\beta}$ $[X]_{\beta} = \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}$

and the coordinate vector of ***T(X)*** relative to $\boldsymbol{\beta}$ is

$$[\boldsymbol{T(X)}]_{\boldsymbol{\beta}} = [\boldsymbol{T}]_{\boldsymbol{\beta}}[\boldsymbol{X}]_{\boldsymbol{\beta}} = \begin{bmatrix} 0 & -6 & 1 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 24 \\ -20 \\ 11 \end{bmatrix}$$

Hence $T(x) = 24b_1 - 20b_2 + 11b_3$.

Exercise:

1. Let $\mathbf{F}:\Re^3 \to \Re^2$ be defined by F(x, y, z) = (z - x, x + y). Find the matrix associated with F with respect to the standard bases of $\Re^2$ and $\Re^3$.
2. Let $\mathbf{T}:\Re^2 \to \Re^3$ defined by T(a, b) = (a, b, a+2b). Find the matrix of T relative to the bases B_1= {(1,1), (2,0)} and B_2 = {(1,1,1), (1,1,0), (0,1,1)}.
3. Let $\mathbf{A}=\begin{pmatrix}1 & 2 & -1\\ 3 & 4 & 0\end{pmatrix}$ and T_A be a linear mapping from $\Re^3$ to $\Re^2$ defined by $\mathbf{T_A(v) = Av}$ where v is a column vector in $\Re^3$. Find the matrix of T_A relative to the bases $\mathbf{B}_1 = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ and $\mathbf{B}_2 = \left\{\begin{pmatrix}1\\3\end{pmatrix}, \begin{pmatrix}2\\5\end{pmatrix}\right\}$ of $\Re^3$ and $\Re^2$ respectively.
4. Suppose that $\boldsymbol{\beta} = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \boldsymbol{b_3}\}$ and $\boldsymbol{\beta'} = \{\boldsymbol{d_1}, \boldsymbol{d_2}\}$ be a basis for real vector spaces V and W, respectively. Let $\boldsymbol{T:V \to W}$ be a linear transformation with the property that $T(b_1) = 3d_1 - 5d_2$, $T(b_2) = -d_1 + 6d_2$, $T(b_3) = 4d2$. Find the matrix M for T relative to $\boldsymbol{\beta}$ and $\boldsymbol{\beta'}$.

5.5. Eigenvalues and Eigenvectors of a linear Transformation

In this section we look for vectors that are sent by a linear operator into scalar multiple of themselves. Recall that a linear operator is a linear transformation from a vector space to it self.

Definition 5.5.1:

Let T: V $\to$ V be a linear operator on a vector space V over a field K. An **eigenvalue** of T is a scalar λ in K such that there is a non-zero vector v in V with T(v) = λv. If λ is an eigenvalue of T, then

a) any vector v in V such that T(v) = λv is called an **eigenvector** of T associated with the eigenvalue λ;

b) the collection of all vectors v of V such that $T(v) = \lambda v$ is called the **eigenspace** associated with λ.

Note:

1. One of the meanings of the word "eigen" in German is "Proper". Thus eigen values are also called proper values or characteristic values or latent roots.
2. If υ and w are eigenvectors associated with eigenvalue λ. Then:
   - i) $\upsilon + w$ is also an eigenvector with eigenvalue λ, because:
     $T(\upsilon + w) = T(\upsilon) + T(w) = \lambda\upsilon + \lambda w = \lambda(\upsilon + w)$
   - ii) each scalar multiple $k\upsilon$, $k \neq 0$, is also an eigenvector with eigenvalue λ, because: $T(k\upsilon) = kT(\upsilon) = k(\lambda\upsilon) = \lambda(k\upsilon)$.

Activity 5.5.1: Let $T: V \to V$ be a linear operator with $\ker T \neq \{0\}$. Prove that every non zero vector in ker T is an eigenvector of T with eigenvalue 0.

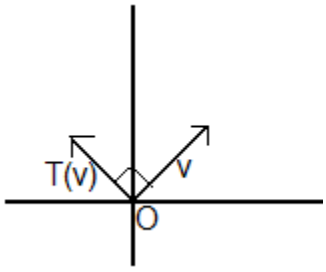
Example 5.5.1:

a) Let id: $V \to V$ be the identity operator.

Every non-zero vector in V is an eigenvector of id with an eigenvalue 1, since:

$Id(\upsilon) = \upsilon = 1.\upsilon$

b) Let $T: \Re^2 \to \Re^2$ be a linear operator which rotates each $\upsilon \in \Re^2$ by an angle of 90^o.



T has no eigenvalues and hence no eigenvectors.

c) Let D: $V \to V$ be the differential operator on the vector space of differentiable functions. We have $D(e^{5t}) = 5e^{5t}$.

Hence, e^{5t} is an eigenvector of D corresponding eigenvalue 5.

5.6. Eigenspace of a linear transformation

Let T: V $\rightarrow$ V a linear operator over a field K.

Let $\lambda \in K$. Let V_λ = The set of all eigenvectors of T with eigenvalue λ.

Claim V_λ is a subspace of V.

Proof:

i) $T(0) = 0 = \lambda \cdot 0$, Showing that $0 \in V_\lambda$.

ii) $\upsilon_1, \upsilon_2 \in V_\lambda \Rightarrow T(\upsilon_1 + \upsilon_2) = T(\upsilon_1) + T(\upsilon_2)$

$= \lambda\upsilon_1 + \lambda\upsilon_2$

$= \lambda(\upsilon_1 + \upsilon_2)$

$\Rightarrow \upsilon_1 + \upsilon_2 \in V_\lambda$

iii) $\upsilon \in V_\lambda, \alpha \in K \Rightarrow T(\alpha\upsilon) = \alpha T(\upsilon) = \alpha(\lambda\upsilon) = \lambda(\alpha\upsilon)$

$\Rightarrow \alpha\upsilon \in V_\lambda$

Thus we have proved that the eigenspace associated with λ is a subspace of V.