

UNIT-ONE: VECTORS

UNIT OBJECTIVES:

At the end of this unit each student should be able to:

- Know representation of vectors geometrically.
- Understand operation of vectors algebraically and geometrically.
- Realize properties of operations on vectors.
- Learn about projection of a vector along another vector.
- Understand the dot product and cross product of two vectors.
- Know the norm of a vector.
- Develop parametric and non-parametric equation of a given line.
- Develop the parametric form and normal form of equation of a plane.

Introduction

Physical quantities which are completely described with magnitude alone is called **scalars**. On the other hand physical quantities which are determined with both magnitude and direction are known as **vectors**. In this unit our goal is not to define vectors at all, but simply will insist that they can be added or multiplied by a number in a familiar algebraic rules. Thus we will review some of the basic theory of vectors in two and three dimensions. This helps us to visualize the geometric concepts.

1.1-Definition of Points In n-Space

Objectives:

At the end of this section, students should be able to

- Add any vectors if they have the same components.
- Multiply all vectors by scalars.
- Represent vectors geometrically in two and three spaces in order to visualize.
- Use properties of vector operations in **n-space** to calculate vector component.

Definition: If n is a positive integer, then an **ordered n -tuple** is a sequence of n real numbers $(a_1, a_2, a_3, \dots, a_n)$. The set of all ordered n -tuples is called **n -space** and is denoted by $\mathfrak{R}^n$.

When $n = 2$ or 3 , it is customary to use the terms **ordered pair** and **ordered triple**, respectively, rather than *ordered 2-tuple* and *ordered 3-tuple*. When $n = 1$, each ordered n -tuple consists of one real number, so $\mathfrak{R}^1$ may be viewed as the set of real numbers. It is usual to write $\mathfrak{R}$ rather than $\mathfrak{R}^1$ for this set. It might have occurred to you in the study of 3-space that the symbol (a_1, a_2, a_3) has two different geometric interpretations: it can be interpreted as a point, in which case, a_1, a_2 and a_3 are the coordinates (Figure 1.1.1a), or it can be interpreted as a vector, in which case a_1, a_2 and a_3 are the components (Figure 1.1.1b). It follows, therefore, that an ordered n -tuple $(a_1, a_2, a_3, \dots, a_n)$ can be viewed either as a “generalized point” or as a “generalized vector”—the distinction is mathematically unimportant. Thus we can describe the 5-tuple $(-2, 4, 0, 1, 6)$ either as a point in $\mathfrak{R}^5$ or as a vector in $\mathfrak{R}^5$.

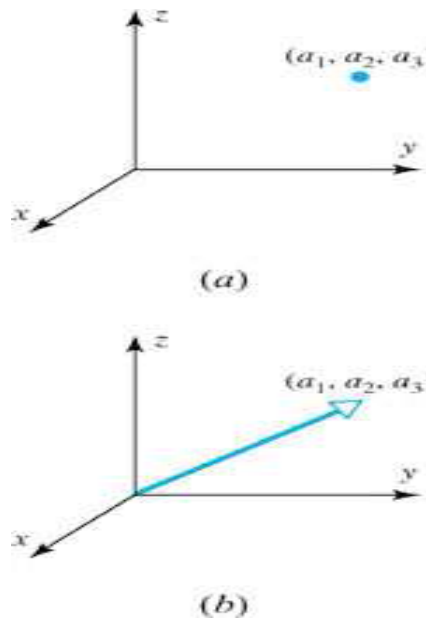


Figure 1.1.1 Geometric representation of (a_1, a_2, a_3) as a point and a vector.

Definition: Two vectors $\mathbf{u} = (u_1, u_2, u_3, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, v_3, \dots, v_n)$ in $\mathfrak{R}^n$ are called **equal** ($\mathbf{u} = \mathbf{v}$) if $u_1 = v_1, u_2 = v_2, u_3 = v_3, \dots, u_n = v_n$. The **sum** $\mathbf{u} + \mathbf{v}$ is defined by $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \dots, u_n + v_n)$ and if k is any scalar, the **scalar multiple**

$k\mathbf{u}$ is defined by $k\mathbf{u} = (ku_1, ku_2, ku_3, \dots, ku_n)$. The operations of addition and scalar multiplication in this definition are called the **standard operations** on $\mathfrak{R}^n$. The **zero vector** in $\mathfrak{R}^n$ is denoted by $\mathbf{0}$ and is defined to be the vector $\mathbf{0} = (0, 0, 0, \dots, 0)$. If $\mathbf{u} = (u_1, u_2, u_3, \dots, u_n)$ is any vector in $\mathfrak{R}^n$, then the **negative** (or **additive inverse**) of $\mathbf{u}$ is denoted by $-\mathbf{u}$ and is defined by $-\mathbf{u} = (-u_1, -u_2, -u_3, \dots, -u_n)$. The **difference** of vectors in $\mathfrak{R}^n$ is defined by $\mathbf{v} - \mathbf{u} = \mathbf{v} + (-\mathbf{u})$ or, in terms of components $\mathbf{v} - \mathbf{u} = (v_1 - u_1, v_2 - u_2, v_3 - u_3, \dots, v_n - u_n)$.

Some Examples of Vectors in Higher-Dimensional Spaces

Experimental Data. A scientist performs an experiment and makes n numerical measurements each time the experiment is performed. The result of each experiment can be regarded as a vector $\mathbf{y} = (y_1, y_2, y_3, \dots, y_n)$ in $\mathfrak{R}^n$ in which $y_1, y_2, y_3, \dots, y_n$ are the measured values.

Storage and Warehousing. A national trucking company has 15 depots for storing and servicing its trucks. At each point in time the distribution of trucks in the service depots can be described by a 15 tuple $\mathbf{x} = (x_1, x_2, x_3, \dots, x_{15})$ in which x_1 is the number of trucks in the first depot, x_2 is the number in the second depot, and so forth.

Economics. Our approach to economic analysis is to divide an economy into sectors (manufacturing, services, utilities, and so forth) and to measure the output of each sector by a dollar value. Thus, in an economy with 10 sectors the economic output of the entire economy can be represented by a 10-tuple $\mathbf{s} = (s_1, s_2, s_3, \dots, s_{10})$ in which the numbers $s_1, s_2, s_3, \dots, s_{10}$ are the outputs of the individual sectors.

Properties of Vector Operations in n -Space

Properties of addition and scalar multiplication of vectors in $\mathfrak{R}^n$ are listed in the following theorem. The proofs are all easy and are left as exercises.

Theorem: Properties of Vectors in $\mathfrak{R}^n$

If $\mathbf{u} = (u_1, u_2, u_3, \dots, u_n)$, $\mathbf{v} = (v_1, v_2, v_3, \dots, v_n)$, and $\mathbf{w} = (w_1, w_2, w_3, \dots, w_n)$ are vectors in $\mathfrak{R}^n$ and k and m are scalars, then:

- (a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- (b) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
- (c) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
- (d) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$; that is, $\mathbf{u} - \mathbf{u} = \mathbf{0}$
- (e) $k(m\mathbf{u}) = (km)\mathbf{u}$
- (f) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
- (g) $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$
- (h) $1\mathbf{u} = \mathbf{u}$

This theorem enables us to manipulate vectors in $\Re^n$ without expressing the vectors in terms of components. For example, to solve the vector equation $\mathbf{x} + \mathbf{u} = \mathbf{v}$ for $\mathbf{x}$, we can add $-\mathbf{u}$ to both sides and proceed as follows:

$$\begin{aligned}\mathbf{x} + \mathbf{u} + (-\mathbf{u}) &= \mathbf{v} + (-\mathbf{u}) \\ \Rightarrow \mathbf{x} + (\mathbf{u} - \mathbf{u}) &= \mathbf{v} - \mathbf{u} \\ \Rightarrow \mathbf{x} + \mathbf{0} &= \mathbf{v} - \mathbf{u} \\ \Rightarrow \mathbf{x} &= \mathbf{v} - \mathbf{u}\end{aligned}$$

The reader will find it instructive to name the parts of the Theorem that justify the last three steps in this computation.

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

In this section, vectors in 2-space and 3-space will be introduced geometrically, arithmetic operations on vectors will be defined, and some basic properties of these arithmetic operations will be established.

Geometric Vectors

Vectors can be represented geometrically as directed line segments or arrows in 2-space or 3-space. The direction of the arrow specifies the direction of the vector, and the length of the arrow describes its magnitude. The tail of the arrow is called the **initial point** of the vector, and the tip of the arrow the **terminal point**. Symbolically, we shall denote vectors in lowercase boldface type (for instance, $\mathbf{a}$, $\mathbf{k}$, $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{x}$). When

discussing vectors, we shall refer to numbers as **scalars**. As shown in Figure 1.2a, the initial point of a vector $\mathbf{v}$ is A and the terminal point is B, we write $\mathbf{v} = \overrightarrow{AB}$.

Vectors with the same length and same direction, such as those in Figure 1.2 b, are called **equivalent**. Since we want a vector to be determined solely by its length and direction, equivalent vectors are regarded as **equal** even though they may be located in different positions. If $\mathbf{v}$ and $\mathbf{w}$ are equivalent, we write $\mathbf{v} = \mathbf{w}$.

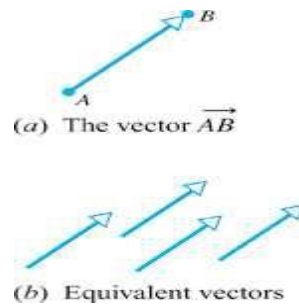


Figure 1.2

Definition: If $\mathbf{v}$ and $\mathbf{w}$ are any two vectors, then the **sum** $\mathbf{v} + \mathbf{w}$ is the vector determined as follows: Position the vector $\mathbf{w}$ so that its initial point coincides with the terminal point of $\mathbf{v}$. The vector $\mathbf{v} + \mathbf{w}$ is represented by the arrow from the initial point of $\mathbf{v}$ to the terminal point of $\mathbf{w}$ (Figure 1.3 a).

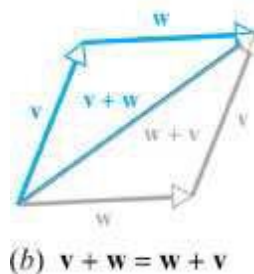
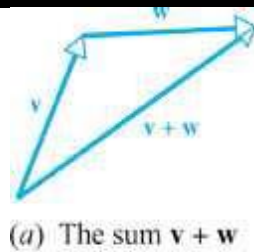


Figure 1.3

In Figure 1.3b we have constructed two sums, $\mathbf{v} + \mathbf{w}$ (diagonal of upper triangle) and $\mathbf{w} + \mathbf{v}$ (diagonal of lower triangle). It is evident that $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ and that the sum

coincides with the diagonal of the parallelogram determined by $\mathbf{v}$ and $\mathbf{w}$ when these vectors are positioned so that they have the same initial point.

The vector of length zero is called the **zero vector** and is denoted by $\mathbf{0}$. We define

$$\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0}.$$

for every vector $\mathbf{v}$. Since there is no natural direction for the zero vector, we shall agree that it can be assigned any direction that is convenient for the problem being considered. If $\mathbf{v}$ is any nonzero vector, then $-\mathbf{v}$, the **negative** of $\mathbf{v}$, is defined to be the vector that has the same magnitude as $\mathbf{v}$ but is oppositely directed (Figure 1.2.3). This vector has the property $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ (Why?) In addition, we define $-\mathbf{0} = \mathbf{0}$. Subtraction of vectors is defined as follows:

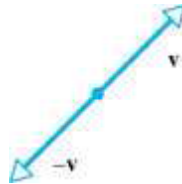


Figure 1.4

The negative of $\mathbf{v}$ has the same length as $\mathbf{v}$ but is oppositely directed.

Definition: If $\mathbf{v}$ and $\mathbf{w}$ are any two vectors, then the **difference** of $\mathbf{w}$ from $\mathbf{v}$ is defined by

$$\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w}) \text{ (Figure 1.5a).}$$

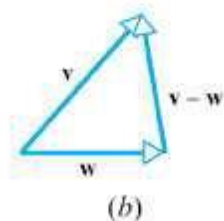
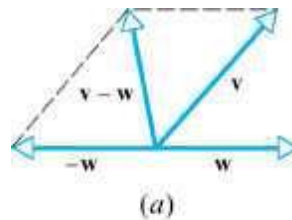


Figure 1.5

To obtain the difference $\mathbf{v} - \mathbf{w}$ without constructing $-\mathbf{w}$, position $\mathbf{v}$ and $\mathbf{w}$ so that their initial points coincide; the vector from the terminal point of $\mathbf{w}$ to the terminal point of $\mathbf{v}$ is then the vector $\mathbf{v} - \mathbf{w}$ (Figure 1.5b).

Definition: If $\mathbf{v}$ is a nonzero vector and k is a nonzero real number (scalar), then the **product** $k\mathbf{v}$ is defined to be the vector whose length is $|k|$ times the length of $\mathbf{v}$ and whose direction is the same as that of $\mathbf{v}$ if $k > 0$ and opposite to that of $\mathbf{v}$ if $k < 0$. We define $k\mathbf{v} = \mathbf{0}$ if $k = 0$ or $\mathbf{v} = \mathbf{0}$.

Figure 1.6 illustrates the relation between a vector $\mathbf{v}$ and the vector $\frac{1}{2}\mathbf{v}$, $(-1)\mathbf{v}$, $2\mathbf{v}$, and $(-3)\mathbf{v}$. Note that the vector $(-1)\mathbf{v}$ has the same length as $\mathbf{v}$ but is oppositely directed. Thus $(-1)\mathbf{v}$ is just the negative of $\mathbf{v}$; that is, $(-1)\mathbf{v} = -\mathbf{v}$.

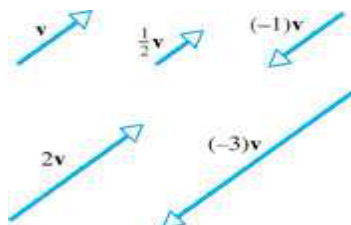


Figure 1.6

A vector of the form $k\mathbf{v}$ is called a **scalar multiple** of $\mathbf{v}$. As evidenced by Figure 1.6, vectors that are scalar multiples of each other are parallel. Conversely, it can be shown that nonzero parallel vectors are scalar multiples of each other. We omit the proof.

Vectors in Coordinate Systems

Problems involving vectors can often be simplified by introducing a rectangular coordinate system. For the moment we shall restrict the discussion to vectors in 2-space (the plane). Let $\mathbf{v}$ be any vector in the plane, and assume, as in Figure 1.7, that $\mathbf{v}$ has been positioned so that its initial point is at the origin of a rectangular coordinate system. The coordinates (v_1, v_2) of the terminal point of $\mathbf{v}$ are called the **components of** $\mathbf{v}$, and we write $\mathbf{v} = (v_1, v_2)$

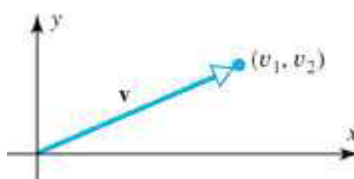


Figure 1.7: v_1 and v_2 are the components of $\mathbf{v}$.

If equivalent vectors, $\mathbf{v}$ and $\mathbf{w}$, are located so that their initial points fall at the origin, then it is obvious that their terminal points must coincide (since the vectors have the same length and direction); thus the vectors have the same components. Conversely, vectors with the same components are equivalent since they have the same length and the same direction. In summary, two vectors $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ are equivalent if and only if $v_1 = w_1$ and $v_2 = w_2$.

The operations of vector addition and multiplication by scalars are easy to carry out in terms of components. As illustrated in Figure 1.2.7, if $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ then $\mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2)$.

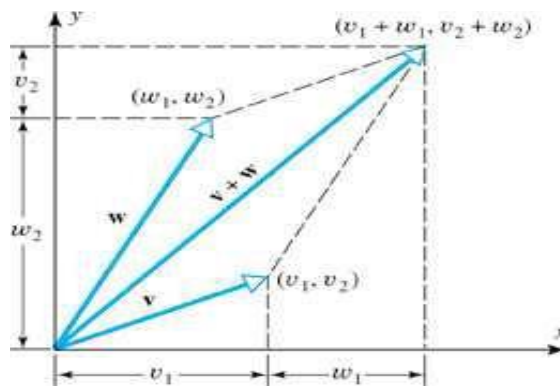


Figure 1.8

(1) If $\mathbf{v} = (v_1, v_2)$ and k is any scalar, then by using a geometric argument involving similar triangles, it can be shown that $k\mathbf{v} = (kv_1, kv_2)$

(2) Thus, for example, if $\mathbf{v} = (1, -2)$ and $\mathbf{w} = (7, 6)$, then

$$\mathbf{v} + \mathbf{w} = (1, -2) + (7, 6) = (1+7, -2+6) = (8, 4) \text{ and}$$

$$4\mathbf{v} = 4(1, -2) = (4(1), 4(-2)) = (4, -8)$$

Since, $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1)\mathbf{w}$, it follows from Formulas 1 and 2 that $\mathbf{v} - \mathbf{w} = (v_1 - w_1, v_2 - w_2)$ (Verify.)

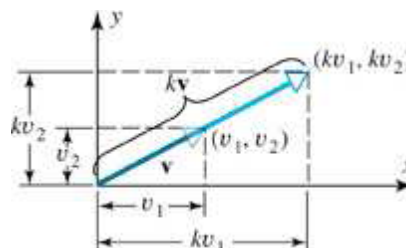


Figure 1.9

Vectors in 3-Space

Just as vectors in the plane can be described by pairs of real numbers, vectors in 3-space can be described by triples of real numbers by introducing a **rectangular coordinate** system. To construct such a coordinate system, select a point O , called the **origin**, and choose three mutually perpendicular lines, called **coordinate axes**, passing through the origin. Label these axes x , y , and z , and select a positive direction for each coordinate axis as well as a unit of length for measuring distances (Figure 1.10a). Each pair of coordinate axes determines a plane called a **coordinate plane**. These are referred to as the **xy -plane**, the **xz -plane**, and the **yz -plane**. To each point P in 3-space we assign a triple of numbers (x, y, z) , called the **coordinates of P** , as follows: Pass three planes through P parallel to the coordinate planes, and denote the points of intersection of these planes with the three coordinate axes by X , Y , and Z (Figure 1.10b). The coordinates of P are defined to be the signed lengths $x = OX$, $y = OY$, $z = OZ$.

In Figure 1.11 *a* we have constructed the point whose coordinates are $(4, 5, 6)$ and in Figure 1.11 *b* the point whose coordinates are $(-3, 2, -4)$.

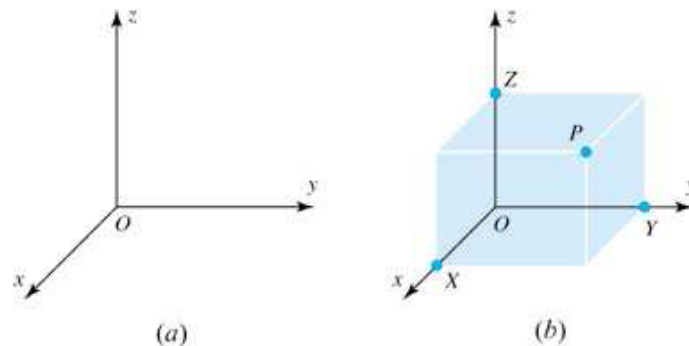


Figure 1.10

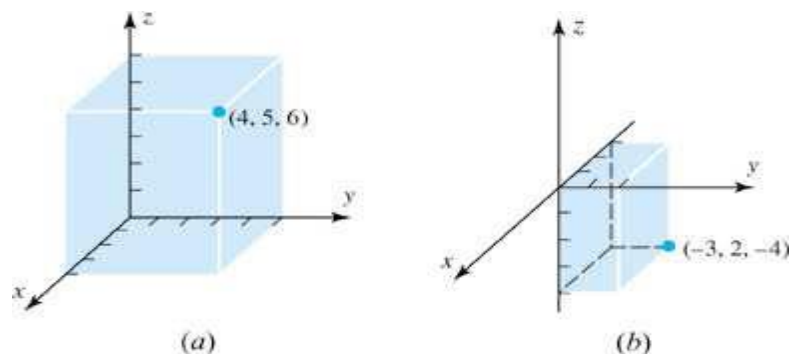


Figure 1.11

Rectangular coordinate systems in 3-space fall into two categories: left-**handed** and **right-handed**. A right-handed system has the property that an ordinary screw pointed in the positive direction on the z -axis would be advanced if the positive x -axis were rotated 90° toward the positive y -axis (Figure 1.12a); the system is left-handed if the screw would be retracted (Figure 1.12b).

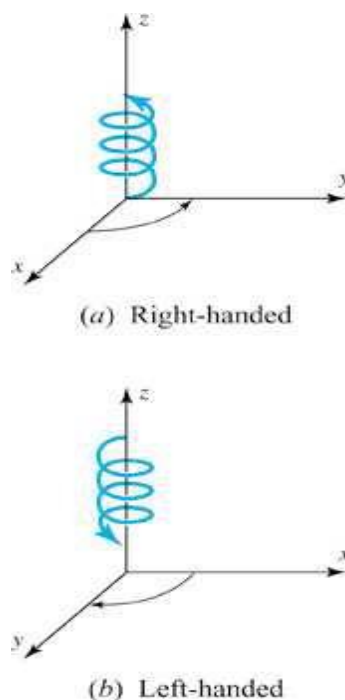


Figure 1.12



Remark: In this module we shall use only right-handed coordinate systems.

If, as in Figure 1.12, a vector $\mathbf{v}$ in 3-space is positioned so its initial point is at the origin of a rectangular coordinate system, then the coordinates of the terminal point are called the **components** of $\mathbf{v}$, and we write $\mathbf{v} = (v_1, v_2, v_3)$.

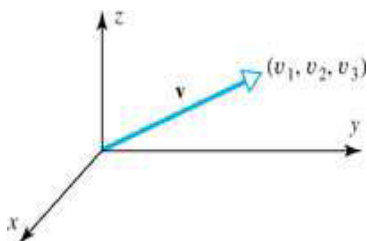


Figure 1.13

If $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ are two vectors in 3-space, then arguments similar to those used for vectors in a plane can be used to establish the following results

- $\mathbf{v}$ and $\mathbf{w}$ are equivalent if and only if

$$v_1 = w_1, v_2 = w_2, \text{ and } v_3 = w_3$$

$$\text{➤ } \mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$$

$$\text{➤ } k\mathbf{v} = (kv_1, kv_2, kv_3), \text{ where } k \text{ is any scalar.}$$

Example 1 : Vector Computations with Components

If $\mathbf{v} = (1, -3, 2)$ and $\mathbf{w} = (4, 2, 1)$, then $\mathbf{v} + \mathbf{w} = (5, -1, 3)$, $2\mathbf{v} = (2, -6, 4)$,

$$-\mathbf{w} = (-4, -2, -1), \mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w}) = (-3, -5, 1).$$

Sometimes a vector is positioned so that its initial point is not at the origin. If the vector

$\vec{P_1P_2}$ has initial point $P_1(x_1, y_1, z_1)$ and terminal point

$$P_2(x_2, y_2, z_2) \text{ then } \vec{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

That is, the components of $\vec{P_1P_2}$ are obtained by subtracting the coordinates of the initial point from the coordinates of the terminal point. This may be seen using Figure 1.14:

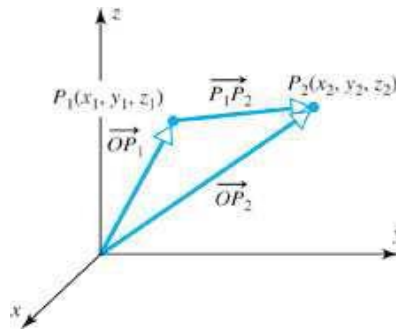


Figure 1.14

The vector $\vec{P_1P_2}$ is the difference of vectors $\vec{OP_2}$ and $\vec{OP_1}$, so $\vec{P_1P_2} = \vec{OP_2} - \vec{OP_1}$

$$\Rightarrow (x_2, y_2, z_2) - (x_1, y_1, z_1)$$

$$\Leftrightarrow (x_2 - x_1, y_2 - y_1, z_2 - z_1)$$

Example 2: Finding the Components of a Vector

The components of the vector $\mathbf{v} = \vec{P_1P_2}$ with initial point $P_1(2, -1, 4)$ and terminal point $P_2(7, 5, -8)$ are $\mathbf{v} = (7 - 2, 5 - (-1), (-8) - 4) = (5, 6, -12)$.

In 2-space the vector with initial point $P_1(x_1, y_1)$ terminal point $P_2(x_2, y_2)$ is

$$\vec{P_1P_2} = (x_2 - x_1, y_2 - y_1)$$

☺ **Test yourself!**

- Let $\mathbf{u} = (2, -7, 1)$, $\mathbf{v} = (-3, 0, 4)$, $\mathbf{w} = (0, 5, -8)$. Find a. $3\mathbf{u} - 4\mathbf{v}$ b. $2\mathbf{u} + 3\mathbf{v} - 5\mathbf{w}$

2. Find x and y if $(4, y) = x(2, 3)$.

3. Compute a. $\begin{pmatrix} 7 \\ -4 \\ 2 \end{pmatrix} + \begin{pmatrix} -3 \\ -1 \\ 5 \end{pmatrix}$ b. $\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} -2 \\ 4 \end{pmatrix}$

4. Prove that $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

5. Find the vector of $A = (1, 1, 3)$ and $B = (-1, 1, 5)$ and sketch the graph.

1.3- Scalar Product and the Norm of a Vector, Orthogonal Projection, Directional Cosines

Objectives:

At the end of this section, the students should be able to

- Calculate the norm of a vector.
- Define the unit vector.
- Determine the scalar (dot) product of nonzero vectors.
- Find the orthogonal projection of vectors in two and three spaces.
- Compute directional angle

Norm of a vector

The **length** of a vector $\mathbf{u}$ is often called the **norm** of $\mathbf{u}$ and is denoted by $\|\mathbf{u}\|$. It follows from the Theorem of Pythagoras that the norm of a vector $\mathbf{u} = (u_1, u_2)$ in 2-space is $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2}$ (Figure 1.16a). Let $\mathbf{u} = (u_1, u_2, u_3)$ be a vector in 3-space. Using Figure 1.16 b and two applications of the Theorem of Pythagoras, we obtain

$$\|\mathbf{u}\|^2 = (OR)^2 + (RP)^2 = (OQ)^2 + (OS)^2 + (RP)^2 = u_1^2 + u_2^2 + u_3^2$$

Thus $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$. A vector of norm 1 is called a **unit vector**.

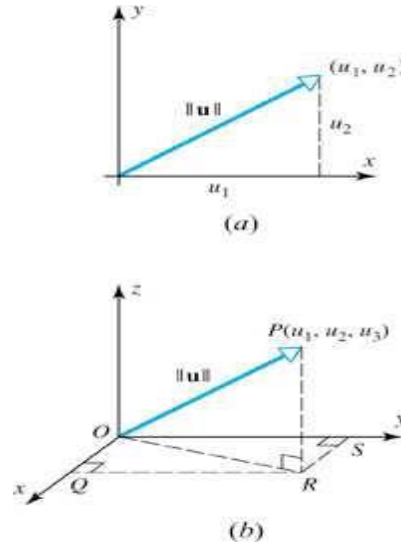


Figure 1.16

If $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ are two points in 3-space, then the **distance** d between them is the norm of the vector $\vec{P_1P_2}$ (Figure 1.3.2). Since

$$\vec{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$$

it follows that

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

Similarly, if $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ are points in 2-space, then the distance between them is given by

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

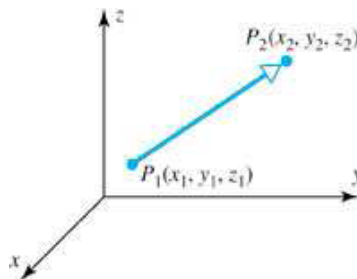


Figure 1.17

The distance between P_1 and P_2 is the norm of the vector $|\vec{P_1P_2}|$.

Example 1: Finding Norm and Distance

The norm of the vector $\mathbf{u} = (-3, 2, 1)$ is $\|\mathbf{u}\| = \sqrt{(-3)^2 + (2)^2 + (1)^2} = \sqrt{14}$.

The distance d between the points $P_1(2, -1, -5)$ and $P_2(4, -3, 1)$ is

$$d = \sqrt{(4-2)^2 + (-3+1)^2 + (1+5)^2} = \sqrt{44} = 2\sqrt{11}.$$

From the definition of the product $k\mathbf{u}$, the length of the vector $k\mathbf{u}$ is $|k|$ times the length of $\mathbf{u}$. Expressed as an equation, this statement says that

$$\|k\mathbf{u}\| = |k| \|\mathbf{u}\|$$

This useful formula is applicable in both 2-space and 3-space.

☺ Test yourself!

1. Find the norm of $\mathbf{v}$.

(a) $\mathbf{v} = (4, -3)$

(b) $\mathbf{v} = (-7, 2, -1)$

(c) $\mathbf{v} = (0, 6, 9)$

2. Find the distance between P_1 and P_2 .

(a) $P_1(3, 4), P_2(5, 7)$

(b) $P_1(7, -5, 1), P_2(-7, -2, -1)$.

3. Let $\mathbf{v} = (-1, 2, 5)$. Find all scalars k such that $\|k\mathbf{v}\| = 4$.

4. (a) Show that if $\mathbf{v}$ is any nonzero vector, then $\frac{1}{\|\mathbf{v}\|} \mathbf{v}$ is a unit vector.

(b) Use the result in part (a) to find a unit vector that has the same direction as the vector $\mathbf{v} = (3, 4)$.

(c) Use the result in part (a) to find a unit vector that is oppositely directed to the vector $\mathbf{v} = (-2, 3, -6)$.

Scalar Product (Dot Product); Orthogonal Projections

Scalar Product of Vectors: It is also called dot product.

Let $\mathbf{u}$ and $\mathbf{v}$ be two nonzero vectors in 2-space or 3-space, and assume these vectors have been positioned so that their initial points coincide. By the **angle between $\mathbf{u}$ and $\mathbf{v}$** , we shall mean the angle θ determined by $\mathbf{u}$ and $\mathbf{v}$ that satisfies $0 \leq \theta \leq \pi$ (Figure 1.3.3).

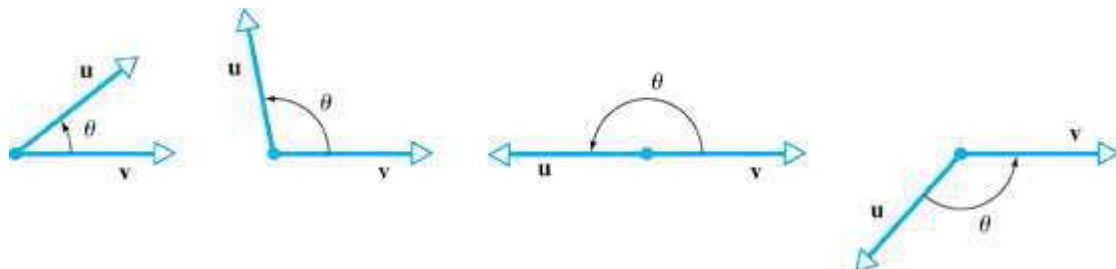


Figure 1.17: The angle θ between $\mathbf{u}$ and $\mathbf{v}$ satisfies $0 \leq \theta \leq \pi$.

Definition: If $\mathbf{u}$ and $\mathbf{v}$ are vectors in 2-space or 3-space and θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, then the **dot product** or **Euclidean inner product** $\mathbf{u} \cdot \mathbf{v}$ is defined by

$$\mathbf{u} \cdot \mathbf{v} = \begin{cases} \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta, & \text{if } \mathbf{u} \neq \mathbf{0} \text{ and } \mathbf{v} \neq \mathbf{0} \\ 0, & \text{if } \mathbf{u} = \mathbf{0} \text{ and } \mathbf{v} = \mathbf{0} \end{cases} \quad (1)$$

Example 1: Dot Product

As shown in Figure 1.19, the angle between the vectors $\mathbf{u} = (0, 0, 1)$ and $\mathbf{v} = (0, 2, 2)$ is 45° . Thus

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = (\sqrt{0^2 + 0^2 + 1^2})(\sqrt{0^2 + 2^2 + 2^2}) \left(\frac{1}{\sqrt{2}} \right) = 2$$

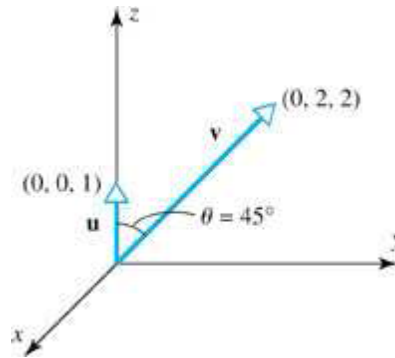


Figure 1.19

Component Form of the Dot Product

For purposes of computation, it is desirable to have a formula that expresses the dot product of two vectors in terms of the components of the vectors. We will derive such a formula for vectors in 3-space; the derivation for vectors in 2-space is similar.

Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be two nonzero vectors. If, as shown in Fig. 1.20, θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, then the law of cosines yields

$$\|\vec{P_1 P_2}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2 \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta. \quad (2)$$

Since $\vec{PQ} = \mathbf{v} - \mathbf{u}$, we can rewrite 2 as $\|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{v} - \mathbf{u}\|^2)$

or $\mathbf{u} \cdot \mathbf{v} = \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{v} - \mathbf{u}\|^2)$

Substituting

$$\begin{aligned}\|\mathbf{u}\|^2 &= u_1^2 + u_2^2 + u_3^2, \\ \|\mathbf{v}\|^2 &= v_1^2 + v_2^2 + v_3^2 \quad \text{and} \\ \|\mathbf{v} - \mathbf{u}\|^2 &= (v_1 - u_1)^2 + (v_2 - u_2)^2 + (v_3 - u_3)^2\end{aligned}$$

we obtain, after simplifying,

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 \quad (3)$$

Although we derived this formula under the assumption that $\mathbf{u}$ and $\mathbf{v}$ are nonzero, the formula is also valid if $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$ (verify).

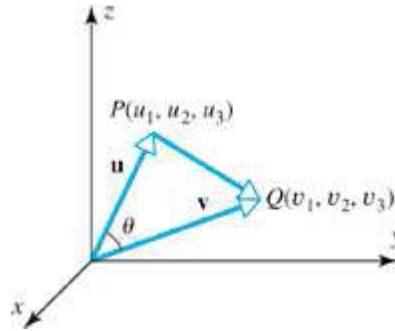


Figure 1.20

If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are two vectors in 2-space, then the formula corresponding to 3 is $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2$ (4)

Finding the Angle Between Vectors

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, then Formula 1 can be written as

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \quad (5)$$

Example 2: Dot Product Using (3)

Consider the vectors $\mathbf{u} = (2, -1, 1)$ and $\mathbf{v} = (1, 1, 2)$. Find $\mathbf{u} \cdot \mathbf{v}$ and determine the angle θ between $\mathbf{u}$ and $\mathbf{v}$.

Solution: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 = (2)(1) + (-1)(1) + (1)(1) = 3$.

For the given vectors we have $\|\mathbf{u}\| = \|\mathbf{v}\| = \sqrt{6}$, so from (5), $\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{3}{\sqrt{6}\sqrt{6}}$

$= \frac{1}{2}$. Thus, $\theta = 60^\circ$

The following theorem shows how the dot product can be used to obtain information about the angle between two vectors; it also establishes an important relationship between the norm and the dot product.

Theorem: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in 2- or 3-space.

(a) $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$; that is, $\|\mathbf{v}\| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$

(b) If the vectors $\mathbf{u}$ and $\mathbf{v}$ are nonzero and θ is the angle between them, then

$$\theta \text{ is acute} \Leftrightarrow \mathbf{u} \cdot \mathbf{v} > 0$$

$$\theta \text{ is obtuse} \Leftrightarrow \mathbf{u} \cdot \mathbf{v} < 0$$

$$\theta = \pi/2 \Leftrightarrow \mathbf{u} \cdot \mathbf{v} = 0$$

Proof (a): Since the angle θ between $\mathbf{v}$ and $\mathbf{v}$ is 0, we have $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\| \|\mathbf{v}\| \cos \theta = \|\mathbf{v}\|^2 \cos 0 = \|\mathbf{v}\|^2$.

Proof (b): since θ satisfies $0 \leq \theta \leq \pi$, it follows that θ is acute if and only if $\cos \theta > 0$, that θ is obtuse if and only if $\cos \theta < 0$, and that $\theta = \frac{\pi}{2}$ if and only if $\cos \theta = 0$. But $\cos \theta$

has the same sign as $\mathbf{u} \cdot \mathbf{v}$ since $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$, $\|\mathbf{u}\| > 0$, and $\|\mathbf{v}\| > 0$.

Thus, the result follows.

Example 3: Finding Dot Products from Components

If $\mathbf{u} = (1, -2, 3)$, $\mathbf{v} = (-3, 4, 2)$, and $\mathbf{w} = (-3, 4, 2)$, then

$$\mathbf{u} \cdot \mathbf{v} = (1)(-3) + (-2)(4) + (3)(2) = -5 \qquad \mathbf{u} \cdot \mathbf{v} = (1)(-3) + (-2)(4) + (3)(2) = 0$$

$$\mathbf{v} \cdot \mathbf{w} = (-3)(-3) + (4)(4) + (2)(2) = 21$$

Therefore, $\mathbf{u}$ and $\mathbf{v}$ make an obtuse angle, $\mathbf{v}$ and $\mathbf{w}$ make an acute angle, and $\mathbf{u}$ and $\mathbf{w}$ are perpendicular.

Orthogonal Vectors

Perpendicular vectors are also called **orthogonal** vectors. In light of the Theorem above, two *nonzero* vectors are orthogonal if and only if their dot product is zero. If we agree to consider $\mathbf{u}$ and $\mathbf{v}$ to be perpendicular when either or both of these vectors is $\mathbf{0}$, then we can state without exception that *two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (perpendicular) if and only if $\mathbf{u} \cdot \mathbf{v} = 0$* . To indicate that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors, we write $\mathbf{u} \perp \mathbf{v}$.

Example 4: A Vector Perpendicular to a Line

Show that in 2-space the nonzero vector $\mathbf{n} = (a, b)$ is perpendicular to the line $ax + by + c = 0$.

Solution: Let $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ be distinct points on the line, so that

$$\begin{aligned} ax_1 + by_1 + c &= 0 \\ ax_2 + by_2 + c &= 0 \end{aligned} \tag{6}$$

Since the vector $\vec{P_1P_2} = (x_2 - x_1, y_2 - y_1)$ runs along the line (Figure 1.21), we need only show that $\mathbf{n}$ and $\vec{P_1P_2}$ are perpendicular. But on subtracting the equations in (6), we obtain $a(x_2 - x_1) + b(y_2 - y_1) = 0$ which can be expressed in the form $(a, b) \cdot (x_2 - x_1, y_2 - y_1) = 0$ or $\mathbf{n} \cdot \vec{P_1P_2} = 0$

Thus $\mathbf{n}$ and $\vec{P_1P_2}$ are perpendicular.

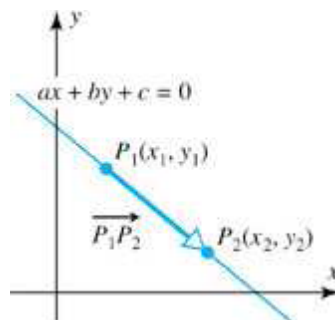


Figure 1.21

The following theorem lists the most important properties of the dot product. They are useful in calculations involving vectors.

Theorem: Properties of the Dot Product

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in 2-or 3-space and k is a scalar, then

- (a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- (b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
- (c) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (k\mathbf{v})$
- (d) $\mathbf{v} \cdot \mathbf{v} > 0$ if $\mathbf{v} \neq 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ if $\mathbf{v} = 0$

Proof: We shall prove (c) for vectors in 3-space and leave the remaining proofs as exercises. Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$; then

$$\begin{aligned}
 k(\mathbf{u} \cdot \mathbf{v}) &= k(u_1v_1 + u_2v_2 + u_3v_3) \\
 &= (ku_1)v_1 + (ku_2)v_2 + (ku_3)v_3 \\
 &= (k\mathbf{u}) \cdot \mathbf{v}
 \end{aligned}$$

Similarly, $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$

An Orthogonal Projection

In many applications it is of interest to “decompose” a vector $\mathbf{u}$ into a sum of two terms, one parallel to a specified nonzero vector $\mathbf{a}$ and the other perpendicular to $\mathbf{a}$. If $\mathbf{u}$ and $\mathbf{a}$ are positioned so that their initial points coincide at a point Q , we can decompose the vector $\mathbf{u}$ as follows (Figure 1.3.7): Drop a perpendicular from the tip of $\mathbf{u}$ to the line through $\mathbf{a}$, and construct the vector $\mathbf{w}_1$ from Q to the foot of this perpendicular. Next form the difference $\overline{\mathbf{w}_2} = \overline{\mathbf{u} - \mathbf{w}_1}$.

As indicated in Figure 1.3.7, the vector $\mathbf{w}_1$ is parallel to $\mathbf{a}$, the vector $\mathbf{w}_2$ is perpendicular to $\mathbf{a}$, and $\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u}$

The vector $\mathbf{w}_1$ is called the **orthogonal projection of $\mathbf{u}$ on $\mathbf{a}$** or sometimes the **vector component of $\mathbf{u}$ along $\mathbf{a}$** . It is denoted by $\text{proj}_{\mathbf{a}} \mathbf{u}$ (7)

The vector $\mathbf{w}_2$ is called the **vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$** . Since we have $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1$, this vector can be written in notation 7 as $\mathbf{w}_2 = \mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u}$

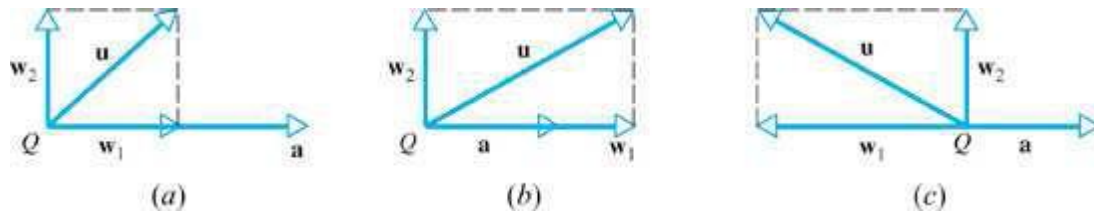


Figure 1.22

The vector $\mathbf{u}$ is the sum of $\mathbf{w}_1$ and $\mathbf{w}_2$, where $\mathbf{w}_1$ is parallel to $\mathbf{a}$ and $\mathbf{w}_2$ is perpendicular to $\mathbf{a}$.

The following theorem gives formulas for calculating $\text{proj}_{\mathbf{a}} \mathbf{u}$ and $\mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u}$.

Theorem:

If $\mathbf{u}$ and $\mathbf{a}$ are vectors in 2-space or 3-space and if $\mathbf{a} \neq \mathbf{0}$, then

$$\begin{aligned}
 \text{proj}_{\mathbf{a}} \mathbf{u} &= \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} && \text{(vector component of } \mathbf{u} \text{ along } \mathbf{a}) \\
 \mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u} &= \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} && \text{(vector component } \mathbf{u} \text{ orthogonal to } \mathbf{a})
 \end{aligned}$$

Proof: Let $\mathbf{w}_1 = \text{proj}_{\mathbf{a}} \mathbf{u}$ and $\mathbf{w}_2 = \mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u}$. Since $\mathbf{w}_1$ is parallel to $\mathbf{a}$, it must be a scalar multiple of $\mathbf{a}$, so it can be written in the form $\mathbf{w}_1 = k\mathbf{a}$. Thus

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 = k\mathbf{a} + \mathbf{w}_2 \quad (8)$$

Taking the dot product of both sides of 8 with $\mathbf{a}$ and using the above Theorems yields

$$\mathbf{u} \cdot \mathbf{a} = (k\mathbf{a} + \mathbf{w}_2) \cdot \mathbf{a} = k \|\mathbf{a}\|^2 + \mathbf{w}_2 \cdot \mathbf{a} \quad (9)$$

But $\mathbf{w}_2 \cdot \mathbf{a} = 0$ since $\mathbf{w}_2$ is perpendicular to $\mathbf{a}$; so 9 yields $k = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}$.

Since $\text{proj}_{\mathbf{a}} \mathbf{u} = \mathbf{w}_1 = k\mathbf{a}$ we obtain $\text{proj}_{\mathbf{a}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a}$

Example 5: Vector Component of $\mathbf{u}$ Along $\mathbf{a}$

Let $\mathbf{u} = (2, -1, 3)$ and $\mathbf{a} = (4, -1, 2)$ Find the vector component of $\mathbf{u}$ along $\mathbf{a}$ and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.

Solution:

$$\mathbf{u} \cdot \mathbf{a} = (2)(4) + (-1)(-1) + (3)(2) = 15$$

$$\|\mathbf{a}\|^2 = 4^2 + (-1)^2 + 2^2 = 21$$

Thus the vector component of $\mathbf{u}$ along $\mathbf{a}$ is

$$\text{proj}_{\mathbf{a}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} = \frac{15}{21} (4, -1, 2) = \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right)$$

and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$ is

$$\mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u} = (2, -1, 3) - \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right) = \left(-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7}\right)$$

As a check, the reader may wish to verify that the vectors $\mathbf{u} - \text{proj}_{\mathbf{a}} \mathbf{u}$ and $\mathbf{a}$ are perpendicular by showing that their dot product is zero.

A formula for the length of the vector component of $\mathbf{u}$ along $\mathbf{a}$ can be obtained by writing

$$\begin{aligned} \|\text{proj}_{\mathbf{a}} \mathbf{u}\| &= \left\| \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \right\| \\ &= \left| \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \right| \|\mathbf{a}\| \\ &= \frac{|\mathbf{u} \cdot \mathbf{a}|}{\|\mathbf{a}\|^2} \|\mathbf{a}\| \end{aligned}$$

$$\text{which yields } \|\text{proj}_{\mathbf{a}} \mathbf{u}\| = \frac{|\mathbf{u} \cdot \mathbf{a}|}{\|\mathbf{a}\|} \quad (10)$$

If θ denotes the angle between $\mathbf{u}$ and $\mathbf{a}$, then $\mathbf{u} \cdot \mathbf{a} = \|\mathbf{u}\| \|\mathbf{a}\| \cos \theta$, so 10 can also be

written as $\|\text{proj}_{\mathbf{a}} \mathbf{u}\| = \|\mathbf{u}\| |\cos \theta|$ (11)

(Verify.) A geometric interpretation of this result is given in Figure 1.23.

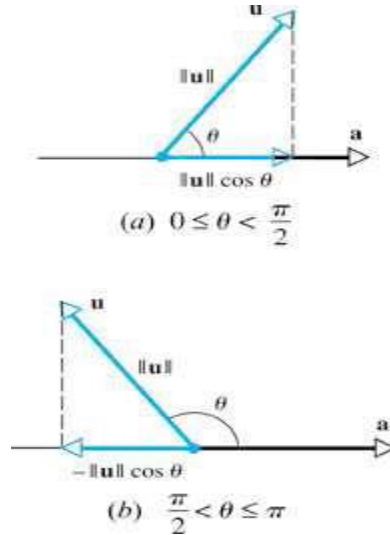


Figure 1.23

As an example, we will use vector methods to derive a formula for the distance from a point in the plane to a line.

Example 6: Distance Between a Point and a Line

Find a formula for the distance D between point $P_0(x_0, y_0)$ and the line $ax + by + c = 0$.

Solution:

Let $Q(x_1, y_1)$ be any point on the line, and position the vector $\mathbf{n} = (a, b)$ so that its initial point is at Q .

By virtue of Example 4, the vector $\mathbf{n}$ is perpendicular to the line (Figure 1.3.9). As indicated in the figure, the distance D is equal to the length of the orthogonal projection of $\overrightarrow{QP_0}$ on $\mathbf{n}$; thus, from 10,

$$D = \|\text{proj}_{\mathbf{n}} \overrightarrow{QP_0}\| = \frac{|\overrightarrow{QP_0} \cdot \mathbf{n}|}{\|\mathbf{n}\|}$$

But

$$\begin{aligned} \overrightarrow{QP_0} &= (x_0 - x_1, y_0 - y_1) \\ \overrightarrow{QP_0} \cdot \mathbf{n} &= a(x_0 - x_1) + b(y_0 - y_1) \\ \|\mathbf{n}\| &= \sqrt{a^2 + b^2} \end{aligned}$$

So

$$D = \frac{|a(x_0 - x_1) + b(y_0 - y_1)|}{\sqrt{a^2 + b^2}} \quad (12)$$

Since the point $Q(x_1, y_1)$ lies on the line, its coordinates satisfy the equation of the line, so $ax_1 + by_1 + c = 0$ or $c = -ax_1 - by_1$

Substituting this expression in 12 yields the formula

$$D = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} \quad (13)$$

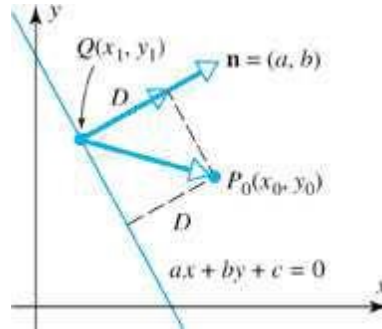


Figure 1.3.9

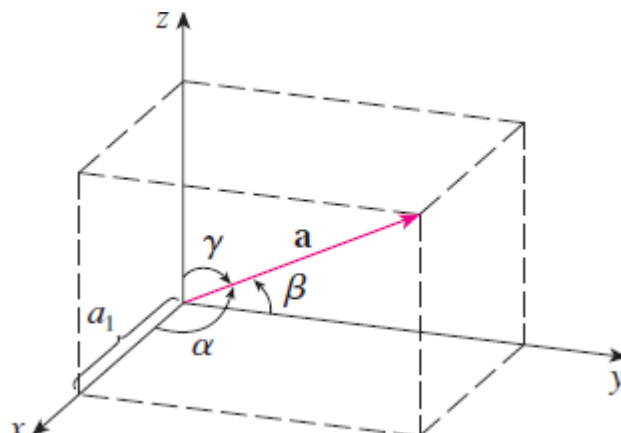
Example 8: Using the Distance Formula

It follows from Formula 13 that the distance D from the point $(1, -2)$ to the line $3x + 4y - 6 = 0$ is

$$D = \frac{|(3)(1) + 4(-2) - 6|}{\sqrt{3^2 + 4^2}} = \frac{|-11|}{\sqrt{25}} = \frac{11}{5}$$

Direction Angles and Direction Cosines

The **direction angles** of a nonzero vector $\mathbf{a}$ are the angles α, β , and γ (in the interval $[0, \pi]$) that $\mathbf{a}$ makes with the positive x, y , and z -axes. (See Figure below.)



The cosines of these direction angles $\cos\alpha, \cos\beta$, and $\cos\gamma$, are called the **direction**

cosines of the vector **a**. Using the formula of angles between two vectors , we obtain

$$\cos\alpha = \frac{\mathbf{a} \cdot \mathbf{i}}{\|\mathbf{a}\|\|\mathbf{i}\|} = \frac{a_1}{\|\mathbf{a}\|}, \cos\beta = \frac{a_2}{\|\mathbf{a}\|}, \text{ and } \cos\gamma = \frac{a_3}{\|\mathbf{a}\|}$$

Example : Find the direction angles of the vector $\mathbf{a} = (1, 2, 3)$

Solution : Since $\|\mathbf{a}\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$ then

$$\begin{aligned} \cos\alpha &= \frac{1}{\sqrt{14}}, \cos\beta = \frac{2}{\sqrt{14}}, \text{ and } \cos\gamma = \frac{3}{\sqrt{14}} \\ \Rightarrow \alpha &= \cos^{-1}\left(\frac{1}{\sqrt{14}}\right) \approx 74^\circ, \beta = \cos^{-1}\left(\frac{2}{\sqrt{14}}\right) \approx 58^\circ \text{ and } \gamma = \cos^{-1}\left(\frac{3}{\sqrt{14}}\right) \approx 37^\circ \end{aligned}$$

☺ Test yourself!

1. Find $\mathbf{u} \cdot \mathbf{v}$.
 - (a) $\mathbf{u} = (2, 3), \mathbf{v} = (5, -7)$
 - (b) $\mathbf{u} = (-2, 2, 2), \mathbf{v} = (1, 7, -4)$
 2. In each part of Exercise 1, find the cosine of the angle θ between $\mathbf{u}$ and $\mathbf{v}$.
 3. Determine whether $\mathbf{u}$ and $\mathbf{v}$ make an acute angle, make an obtuse angle, or are orthogonal.
 - (a) $\mathbf{u} = (6, 1, 4), \mathbf{v} = (2, 0, -3)$
 - (b) $\mathbf{u} = (0, 0, -1), \mathbf{v} = (1, 1, 1)$
 4. Find the orthogonal projection of $\mathbf{u}$ on $\mathbf{a}$.
 - (a) $\mathbf{u} = (-1, -2), \mathbf{a} = (-2, 3)$
 - (b) $\mathbf{u} = (1, 0, 0), \mathbf{a} = (4, 3, 8)$
 5. In each part of Exercise 4, find the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.
-

1.4- The vector product

In many applications of vectors to problems in geometry, physics, and engineering, it is of interest to construct a vector in 3-space that is perpendicular to two given vectors. In this section we shall show how to do this.

Objectives:

- Find the cross product of vectors.
- Compute determinant for two and three-order vectors.
- Apply cross product to find area and volum of figures.

Definition: The **cross product** $\mathbf{a} \times \mathbf{b}$ of two vectors $\mathbf{a}$ and $\mathbf{b}$, unlike the dot product, is a vector. For this reason it is also called the **vector product**. Note that $\mathbf{a} \times \mathbf{b}$ is defined only when $\mathbf{a}$ and $\mathbf{b}$ are *three-dimensional* vectors.

1 DEFINITION If $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, then the **cross product** of $\mathbf{a}$ and $\mathbf{b}$ is the vector

$$\mathbf{a} \times \mathbf{b} = \langle a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1 \rangle$$

This may seem like a strange way of defining a product. The reason for the particular form of Definition 1 is that the cross product defined in this way has many useful properties, as we will soon see. In particular, we will show that the vector is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. In order to make Definition 1 easier to remember, we use the notation of determinants.

A **determinant of order 2** is defined by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

For example, $\begin{vmatrix} 2 & 1 \\ -6 & 4 \end{vmatrix} = 2(4) - 1(-6) = 14$

A **determinant of order 3** can be defined in terms of second-order determinants as follows :

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

Observe that each term on the right side of Equation 2 involves a number a_i in the first row of the determinant, and a_i is multiplied by the second-order determinant obtained from the left side by deleting the row and column in which a_i appears. Notice also the minus sign in the second term. For example,

$$\begin{vmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \\ -5 & 4 & 2 \end{vmatrix} = 1 \begin{vmatrix} 0 & 1 \\ 4 & 2 \end{vmatrix} - 2 \begin{vmatrix} 3 & 1 \\ -5 & 2 \end{vmatrix} + (-1) \begin{vmatrix} 3 & 0 \\ -5 & 4 \end{vmatrix} \\ = 1(0 - 4) - 2(6 + 5) + (-1)(12 - 0) = -38$$

If we now rewrite Definition 1 using second-order determinants and the standard basis

Vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, we see that the cross product of the vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$

$$\boxed{3} \quad \mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}$$

In view of the similarity between Equations 2 and 3, we often write

$$\boxed{4} \quad \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Although the first row of the symbolic determinant in Equation 4 consists of vectors, if we expand it as if it were an ordinary determinant using the rule in Equation 2, we obtain Equation 3. The symbolic formula in Equation 4 is probably the easiest way of remembering and computing cross products.

EXAMPLE 1 If $\mathbf{a} = \langle 1, 3, 4 \rangle$ and $\mathbf{b} = \langle 2, 7, -5 \rangle$, then

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 3 & 4 \\ 2 & 7 & -5 \end{vmatrix} \\ &= \begin{vmatrix} 3 & 4 \\ 7 & -5 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 4 \\ 2 & -5 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 3 \\ 2 & 7 \end{vmatrix} \mathbf{k} \\ &= (-15 - 28)\mathbf{i} - (-5 - 8)\mathbf{j} + (7 - 6)\mathbf{k} = -43\mathbf{i} + 13\mathbf{j} + \mathbf{k} \end{aligned}$$

EXAMPLE 2 Show that $\mathbf{a} \times \mathbf{a} = \mathbf{0}$ for any vector $\mathbf{a}$ in V_3 .

SOLUTION If $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$, then

$$\begin{aligned} \mathbf{a} \times \mathbf{a} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \end{vmatrix} \\ &= (a_2a_3 - a_3a_2)\mathbf{i} - (a_1a_3 - a_3a_1)\mathbf{j} + (a_1a_2 - a_2a_1)\mathbf{k} \\ &= 0\mathbf{i} - 0\mathbf{j} + 0\mathbf{k} = \mathbf{0} \end{aligned}$$

One of the most important properties of the cross product is given by the following theorem.

THEOREM The vector $\mathbf{a} \times \mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.

Proof: In order to show that $\mathbf{a} \times \mathbf{b}$ is orthogonal to $\mathbf{a}$, we compute their dot product as follows:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} a_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} a_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} a_3 \\ &= a_1(a_2b_3 - a_3b_2) - a_2(a_1b_3 - a_3b_1) + a_3(a_1b_2 - a_2b_1) \\ &= a_1a_2b_3 - a_1a_3b_2 - a_1a_2b_3 + a_1a_3b_2 + a_1a_2b_3 - a_1a_2b_3 \\ &= 0 \end{aligned}$$

A similar computation shows that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0$. Therefore $\mathbf{a} \times \mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. If $\mathbf{a}$ and $\mathbf{b}$ are represented by directed line segments with the same initial point (as in Figure 1), then Theorem 5 says that the cross product $\mathbf{a} \times \mathbf{b}$ points in a direction perpendicular to the plane through $\mathbf{a}$ and $\mathbf{b}$. It turns out that the direction of $\mathbf{a} \times \mathbf{b}$ is given by the *right-hand rule*: If the fingers of your right hand curl in the direction of a rotation (through an angle less than 180°) from $\mathbf{a}$ to $\mathbf{b}$, then your thumb points in the direction of $\mathbf{a} \times \mathbf{b}$. Now that we know the direction of the vector $\mathbf{a} \times \mathbf{b}$, the remaining thing we need to complete its geometric description is its length $\|\mathbf{a} \times \mathbf{b}\|$. This is given by the following theorem.

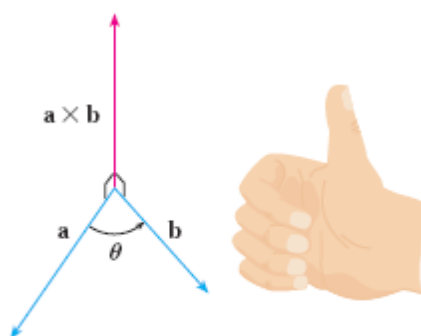


FIGURE 1

6 THEOREM If θ is the angle between $\mathbf{a}$ and $\mathbf{b}$ (so $0 \leq \theta \leq \pi$), then

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta$$

Proof: From the definitions of the cross product and length of a vector, we have

$$\|\mathbf{a} \times \mathbf{b}\|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2$$

$$\begin{aligned}
 &= a_2^2 b_3^2 - 2a_2 a_3 b_2 b_3 + a_3^2 b_2^2 + a_3^2 b_1^2 - 2a_1 a_3 b_1 b_3 \\
 &\quad + a_1^2 b_3^2 + a_1^2 b_2^2 - 2a_1 a_2 b_1 b_2 + a_2^2 b_1^2 \\
 &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2 \\
 &= |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 \\
 &= |\mathbf{a}|^2 |\mathbf{b}|^2 - |\mathbf{a}|^2 |\mathbf{b}|^2 \cos^2 \theta \\
 &= |\mathbf{a}|^2 |\mathbf{b}|^2 (1 - \cos^2 \theta) \\
 &= |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \theta
 \end{aligned}$$

Taking square roots and observing that $\sqrt{\sin^2 \theta} = \sin \theta$ because $\sin \theta \geq 0$ when $0 \leq \theta \leq \pi$, we have

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta$$

Since a vector is completely determined by its magnitude and direction, we can now say that $\mathbf{a} \times \mathbf{b}$ is the vector that is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, whose orientation is determined by the right hand rule, whose length is $|\mathbf{a}| |\mathbf{b}| \sin \theta$.

In fact, that is exactly how physicists *define* $\mathbf{a} \times \mathbf{b}$.

7 COROLLARY Two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel if and only if

$$\mathbf{a} \times \mathbf{b} = \mathbf{0}$$

Proof: Two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel if and only if $\theta = 0$ or π . In either case

$$\sin \theta = 0, \text{ so } |\mathbf{a} \times \mathbf{b}| = 0 \text{ and therefore } \mathbf{a} \times \mathbf{b} = \mathbf{0}.$$

The geometric interpretation of Theorem 6 can be seen by looking at Figure 2. If $\mathbf{a}$ and $\mathbf{b}$ are represented by directed line segments with the same initial point, then they determine a parallelogram with base $|\mathbf{a}|$, altitude $|\mathbf{b}| \sin \theta$, and area

$$A = |\mathbf{a}| (|\mathbf{b}| \sin \theta) = |\mathbf{a} \times \mathbf{b}|$$

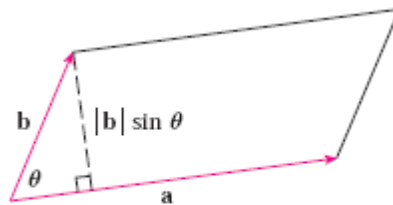


FIGURE 2

Thus we have the following way of interpreting the magnitude of a cross product.

The length of the cross product $\mathbf{a} \times \mathbf{b}$ is equal to the area of the parallelogram determined by $\mathbf{a}$ and $\mathbf{b}$.

Example 3: Find a vector perpendicular to the plane that passes through the points $P(1, 4, 6)$, $Q(-2, 5, -1)$, and $R(1, -1, 1)$.

SOLUTION The vector $\vec{PQ} \times \vec{PR}$ is perpendicular to both $\vec{PQ}$ and $\vec{PR}$ and is therefore perpendicular to the plane through P , Q , and R . We know from (12.2.1) that

$$\vec{PQ} = (-2 - 1)\mathbf{i} + (5 - 4)\mathbf{j} + (-1 - 6)\mathbf{k} = -3\mathbf{i} + \mathbf{j} - 7\mathbf{k}$$

$$\vec{PR} = (1 - 1)\mathbf{i} + (-1 - 4)\mathbf{j} + (1 - 6)\mathbf{k} = -5\mathbf{j} - 5\mathbf{k}$$

We compute the cross product of these vectors:

$$\begin{aligned} \vec{PQ} \times \vec{PR} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 1 & -7 \\ 0 & -5 & -5 \end{vmatrix} \\ &= (-5 - 35)\mathbf{i} - (15 - 0)\mathbf{j} + (15 - 0)\mathbf{k} = -40\mathbf{i} - 15\mathbf{j} + 15\mathbf{k} \end{aligned}$$

So the vector $\langle -40, -15, 15 \rangle$ is perpendicular to the given plane. Any nonzero scalar multiple of this vector, such as $\langle -8, -3, 3 \rangle$, is also perpendicular to the plane.

Example 4: Find the area of the triangle with vertices

$P(1, 4, 6)$, $Q(-2, 5, -1)$, and $R(1, -1, 1)$.

Solution: In Example 3 we computed that $\vec{PQ} \times \vec{PR} = \langle -40, -15, 15 \rangle$

The area of the parallelogram with adjacent sides $\vec{PQ}$ and $\vec{PR}$ is the length of this cross product:

$$|\vec{PQ} \times \vec{PR}| = \sqrt{(-40)^2 + (-15)^2 + 15^2} = 5\sqrt{82}$$

If we apply Theorems 5 and 6 to the standard basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ using $\theta = \pi/2$, we obtain

$$\begin{array}{lll} \mathbf{i} \times \mathbf{j} = \mathbf{k} & \mathbf{j} \times \mathbf{k} = \mathbf{i} & \mathbf{k} \times \mathbf{i} = \mathbf{j} \\ \mathbf{j} \times \mathbf{i} = -\mathbf{k} & \mathbf{k} \times \mathbf{j} = -\mathbf{i} & \mathbf{i} \times \mathbf{k} = -\mathbf{j} \end{array}$$

Observe that

$$\mathbf{i} \times \mathbf{j} \neq \mathbf{j} \times \mathbf{i}$$

Thus the cross product is not commutative. Also

$$\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$$

- ❑ So the associative law for multiplication does not usually hold; that is, in general, $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$

However, some of the usual laws of algebra *do* hold for cross products. The following theorem summarizes the properties of vector products.

8 THEOREM If $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are vectors and c is a scalar, then

1. $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
2. $(c\mathbf{a}) \times \mathbf{b} = c(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (c\mathbf{b})$
3. $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$
4. $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
5. $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$
6. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$

These properties can be proved by writing the vectors in terms of their components and using the definition of a cross product. We give the proof of Property 5 and leave the remaining proofs as exercises.

Proof Of Property 5

If $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$, $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, and $\mathbf{c} = \langle c_1, c_2, c_3 \rangle$, then

$$\begin{aligned} \text{9} \quad \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) &= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1) \\ &= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1 \\ &= (a_2b_3 - a_3b_2)c_1 + (a_3b_1 - a_1b_3)c_2 + (a_1b_2 - a_2b_1)c_3 \\ &= (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \end{aligned}$$

There is an important difference between the dot product and cross product of two vectors—the dot product is a scalar and the cross product is a vector. The following theorem gives some important relationships between the dot product and cross product and also shows that $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

Theorem:- Relationships Involving Cross Product and Dot Product

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in 3-space, then

(a) $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ is orthogonal to $\mathbf{u}$)

- (b) $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ is orthogonal to $\mathbf{v}$)
- (c) $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (Lagrange's identity)
- (d) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ (relationship between cross and dot products)
- (e) $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u}$ (relationship between cross and dot products)

Proof (a) Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$. Then

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) &= (u_1, u_2, u_3) \cdot (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1) \\ &= u_1(u_2v_3 - u_3v_2) + u_2(u_3v_1 - u_1v_3) + u_3(u_1v_2 - u_2v_1) = 0 \end{aligned}$$

Proof (b) Similar to (a).

Proof (c) Since

$$\|\mathbf{u} \times \mathbf{v}\|^2 = (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 \quad (1)$$

and

$$\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 = (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \quad (2)$$

the proof can be completed by “multiplying out” the right sides of 1 and 2 and verifying their equality.

Proof (d) and (e) As an Exercise.

Example 5: $\mathbf{u} \times \mathbf{v}$ Is Perpendicular to $\mathbf{u}$ and to $\mathbf{v}$

Consider the vectors $\mathbf{u} = (1, 2, -2)$ and $\mathbf{v} = (3, 0, 1)$

Then we can show that $\mathbf{u} \times \mathbf{v} = (2, -7, -6)$

Since $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = (1)(2) + (2)(-7) + (-2)(-6) = 0$

and $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = (3)(2) + (0)(-7) + (1)(-6) = 0$

$\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, as guaranteed by the Theorem above.

$\mathbf{u} \times \mathbf{v}$

1.5- Applications on area and volume

Geometric Interpretation of Cross Product

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in 3-space, then the norm of $\mathbf{u} \times \mathbf{v}$ has a useful geometric interpretation. Lagrange's identity, given in the above Theorem, states that

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \quad (1)$$

If θ denotes the angle between $\mathbf{u}$ and $\mathbf{v}$, then $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$ so 1 can be rewritten as

$$\begin{aligned}\|\mathbf{u} \times \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \cos^2 \theta \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 (1 - \cos^2 \theta) \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2 \theta\end{aligned}$$

Since $0 \leq \theta \leq \pi$, it follows that $\sin \theta \geq 0$, so this can be rewritten as

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin \theta \quad (2)$$

But $\|\mathbf{v}\| \sin \theta$ is the altitude of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ (Figure 1.5.1).

Thus, from 2, the area A of this parallelogram is given by

$$A = (\text{base})(\text{altitude}) = \|\mathbf{u}\| \|\mathbf{v}\| \sin \theta = \|\mathbf{u} \times \mathbf{v}\|$$

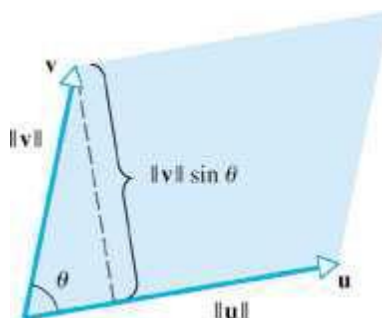


Figure 1.5.1

This result is even correct if $\mathbf{u}$ and $\mathbf{v}$ are collinear, since the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ has zero area and from 2 we have $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ because $\theta = 0$ in this case. Thus we have the following theorem.

Theorem 1.5.1- Area of a Parallelogram

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in 3-space, then $\|\mathbf{u} \times \mathbf{v}\|$ is equal to the area of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$.

Example 1: Area of a Triangle

Find the area of the triangle determined by the points

$$P_1 = (2, 2, 0), P_2 = (-1, 0, 2), \text{ and } P_3 = (0, 4, 3).$$

Solution

The area A of the triangle is $\frac{1}{2}$ the area of the parallelogram determined by the vectors

$\vec{P_1 P_2}$ and $\vec{P_1 P_3}$ (Figure 1.5.2). Using the method discussed previously $\vec{P_1 P_2} = (-3, -2, 2)$,

and $\vec{P_1 P_3} = (-2, 2, 3)$. It follows that $\vec{P_1 P_2} \times \vec{P_1 P_3} = (-10, 5, -10)$ and consequently,

$$A = \frac{1}{2} \left\| \vec{P_1 P_2} \times \vec{P_1 P_3} \right\| = \frac{1}{2} (15) = \frac{15}{2}$$

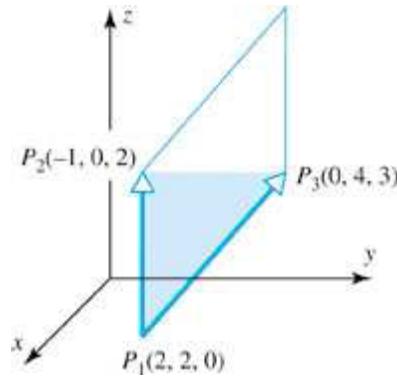


Figure 1.5.2

Definition –Scalar triple product

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in 3-space, then $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is called the **scalar triple product** of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. The scalar triple product of $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, and $\mathbf{w} = (w_1, w_2, w_3)$ can be calculated from the formula

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \quad (3)$$

Since,

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \mathbf{u} \cdot \left(\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \mathbf{k} \right) \\ &= \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} u_1 - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} u_2 + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} u_3 \\ &= \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \end{aligned}$$

Example 2: Calculating a Scalar Triple Product

Calculate the scalar triple product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ of the vectors

$$\mathbf{u} = 3\mathbf{i} - 2\mathbf{j} - 5\mathbf{k}, \quad \mathbf{v} = \mathbf{i} + 4\mathbf{j} - 4\mathbf{k}, \quad \mathbf{w} = 3\mathbf{j} + 2\mathbf{k}$$

Solution:

From 2,

$$\begin{aligned}
 \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \begin{vmatrix} 3 & -2 & -5 \\ 1 & 4 & -4 \\ 0 & 3 & 2 \end{vmatrix} \\
 &= 3 \begin{vmatrix} 4 & -4 \\ 3 & 2 \end{vmatrix} - (-2) \begin{vmatrix} 1 & -4 \\ 0 & 2 \end{vmatrix} + (-5) \begin{vmatrix} 1 & 4 \\ 0 & 3 \end{vmatrix} \\
 &= 60 + 4 - 15 = 49
 \end{aligned}$$

 **Remark:** The symbol $(\mathbf{u} \cdot \mathbf{v}) \times \mathbf{w}$ makes no sense because we cannot form the cross product of a scalar and a vector. Thus no ambiguity arises if we write $\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}$ rather than $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$. However, for clarity we shall usually keep the parentheses. It follows from 3 that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u})$, since the 3×3 determinants that represent these products can be obtained from one another by two row interchanges. (Verify.) These relationships can be remembered by moving the vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ clockwise around the vertices of the triangle in Figure 1.4.3.

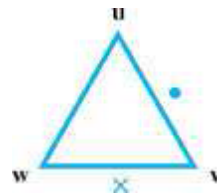


Figure 1.5.3

Geometric Interpretation of Determinants: The next theorem provides a useful geometric interpretation of 2×2 and 3×3 determinants.

Theorem 1.5.2-

The absolute value of the determinant $\det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix}$ is equal to the area of the parallelogram in 2-space determined by the vectors $\mathbf{u}$ and $\mathbf{v}$. ($\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$). (See Figure 1.4.4a.)

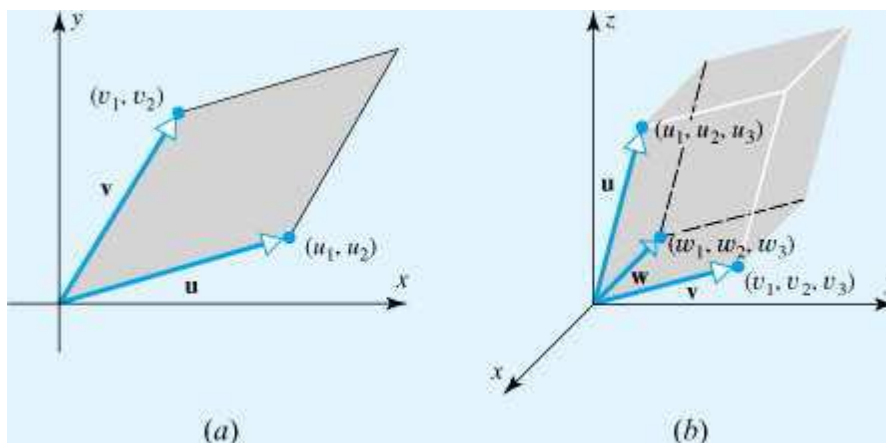


Figure 1.5.4

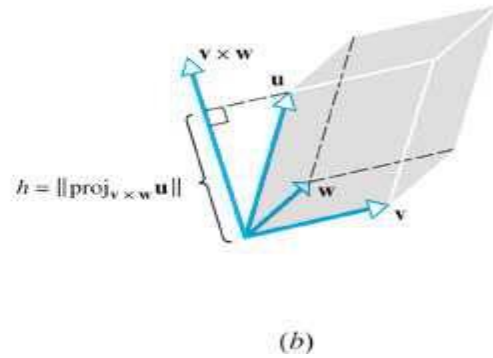
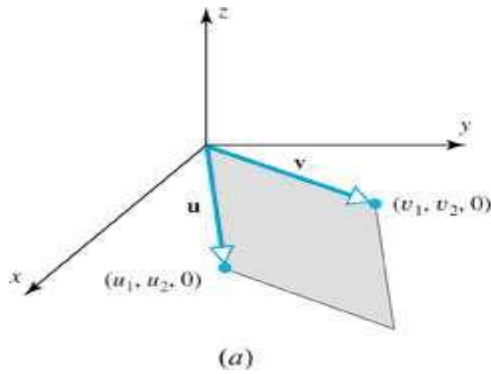
a) The absolute value of the determinant

$$\det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}$$

is equal to the volume of the parallelepiped in 3-space determined by the vectors $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, and $\mathbf{w} = (w_1, w_2, w_3)$. (See Figure 1.4.4b.)

Proof (a) The key to the proof is to use Theorem 1.4.1. However, that theorem applies to vectors in 3-space, whereas $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are vectors in 2-space. To circumvent this “dimension problem,” we shall view $\mathbf{u}$ and $\mathbf{v}$ as vectors in the xy -plane of an xyz -coordinate system (Figure 1.5.5a), in which case these vectors are expressed as $\mathbf{u} = (u_1, u_2, 0)$ and $\mathbf{v} = (v_1, v_2, 0)$. Thus

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & 0 \\ v_1 & v_2 & 0 \end{vmatrix} = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{k} = \det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \mathbf{k}$$


Figure 1.5.5

It now follows from Theorem 1.5.1 and the fact that that $\|\mathbf{k}\| = 1$ the area A of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ is

$$A = \|\mathbf{u} \times \mathbf{v}\| = \left\| \det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \mathbf{k} \right\| = \left| \det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \right| \|\mathbf{k}\| = \left| \det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \right|$$

which completes the proof.

Proof (b) As shown in Figure 1.5.5b, take the base of the parallelepiped determined by $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ to be the parallelogram

determined by $\mathbf{v}$ and $\mathbf{w}$. It follows from Theorem 1.5.1 that the area of the base is $\|\mathbf{v} \times \mathbf{w}\|$ and, as illustrated in Figure 1.5.5b, the height h of the parallelepiped is the length of the orthogonal projection of $\mathbf{u}$ on $\|\mathbf{v} \times \mathbf{w}\|$. Therefore,

$$h = \|\text{proj}_{\mathbf{v} \times \mathbf{w}} \mathbf{u}\| = \frac{|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|}{\|\mathbf{v} \times \mathbf{w}\|}$$

It follows that the volume V of the parallelepiped is

$$V = (\text{area of base}) \cdot \text{height} = \|\mathbf{v} \times \mathbf{w}\| \frac{|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|}{\|\mathbf{v} \times \mathbf{w}\|} = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$$

so from 3,

$$V = \left| \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \right|$$

which completes the proof.

 **Remark:** If V denotes the volume of the parallelepiped determined by vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, then it follows

$$V = \left[\begin{array}{l} \text{volume of parallelepiped} \\ \text{determined by } \mathbf{u}, \mathbf{v}, \text{ and } \mathbf{w} \end{array} \right] = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| \quad (4)$$

From this, we can conclude that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \pm V$ where the $+$ or $-$ results depending on whether $\mathbf{u}$ makes an acute or an obtuse angle with $\mathbf{v} \times \mathbf{w}$. Formula 4 leads to a useful test for ascertaining whether three given vectors lie in the same plane. Since three vectors not in the same plane determine a parallelepiped of positive volume, it follows from 4 $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = 0$ that if and only if the vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ lie in the same plane. Thus we have the following result.

Theorem 1.5.3

If the vectors $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, and $\mathbf{w} = (w_1, w_2, w_3)$ have the same initial point, then they lie in the same

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \left| \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \right| = 0$$

The product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is called the **vector triple product of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$** . This is used to derive Kepler's First Law of planetary motion

☺ Test yourself!

1. Let $\mathbf{u} = (3, 2, -1)$, $\mathbf{v} = (0, 2, -3)$, and $\mathbf{w} = (2, 6, 7)$. Compute

- (a) $\mathbf{v} \times \mathbf{w}$
- (b) $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$
- (c) $\mathbf{u} \times (\mathbf{v} - 2\mathbf{w})$

2. Find a vector that is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

- (a) $\mathbf{u} = (-6, 4, 2)$, $\mathbf{v} = (3, 1, 5)$
- (b) $\mathbf{u} = (-2, 1, 5)$, $\mathbf{v} = (3, 0, -3)$

3. Find the area of the parallelogram determined by $\mathbf{u}$ and

- (a) $\mathbf{u} = (1, -1, 2)$, $\mathbf{v} = (0, 3, 1)$
- (b) $\mathbf{u} = (3, -1, 4)$, $\mathbf{v} = (6, -2, 8)$

4. Find the area of the triangle having vertices P, Q, and R.

- (a) $P(2, 6, -1)$, $Q(1, 1, 1)$, $R(4, 6, 2)$
- (b) $P(4, 6, 2)$, $Q(0, 3, 4)$, $R(6, 1, 8)$

5. Find a vector $\mathbf{v}$ that is orthogonal to the vector $\mathbf{u} = (2, -3, 5)$.

6. Find the scalar triple product.

- (a) $\mathbf{u} = (-1, 2, 4)$, $\mathbf{v} = (3, 4, -2)$, $\mathbf{w} = (-1, 2, 5)$
- (b) $\mathbf{u} = (3, -1, 6)$, $\mathbf{v} = (2, 4, 3)$, $\mathbf{w} = (5, -1, 2)$

7. Suppose that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 3$. Find

- (a) $\mathbf{u} \cdot (\mathbf{w} \times \mathbf{v})$
- (b) $(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}$
- (c) $\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$

8. Find the volume of the parallelepiped with sides $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

- (a) $\mathbf{u} = (2, -6, 2)$, $\mathbf{v} = (0, 4, -2)$, $\mathbf{w} = (2, 2, -4)$
 - (b) $\mathbf{u} = (3, 1, 2)$, $\mathbf{v} = (4, 5, 1)$, $\mathbf{w} = (1, 2, 4)$
-

1.6- Lines and Planes

Objectives

- Find parametric equation of a line.
- Calculate the normal vector of a plane.
- Compute the distance between two planes.

A line in the xy -plane is determined when a point on the line and the direction of the line (its slope or angle of inclination) are given. The equation of the line can then be written using the point-slope form. Likewise, a line L in three-dimensional space is determined when we know a point $P_0(x_0, y_0, z_0)$ on L and the direction of L . In three dimensions the direction of a line is conveniently described by a vector, so we let $\mathbf{v}$ be a vector parallel to L . Let $P(x, y, z)$ be an arbitrary point on L and let $\mathbf{r}_0$ and $\mathbf{r}$ be the position vectors of P_0 and P (that is, they have representations $\vec{OP_0}$ and $\vec{OP}$). If $\mathbf{a}$ is the vector with representation $\vec{P_0P}$, as in Figure 1.6.1, then the Triangle Law for vector addition gives $\mathbf{r} = \mathbf{r}_0 + \mathbf{a}$.

But, since $\mathbf{a}$ and $\mathbf{v}$ are parallel vectors, there is a scalar t such that $\mathbf{a} = t\mathbf{v}$. Thus

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} \quad (1)$$

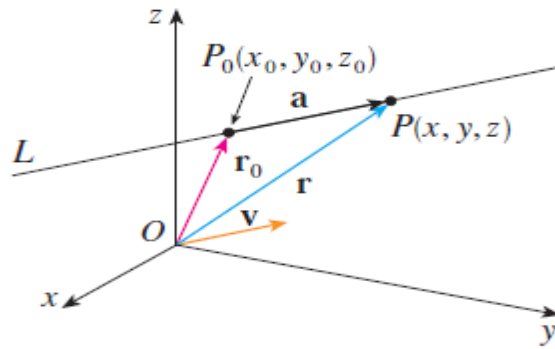


Figure 1.6.1

which is a **vector equation** of L . Each value of the **parameter** t gives the position vector $\mathbf{r}$ of a point on L . In other words, as t varies, the line is traced out by the tip of the vector $\mathbf{r}$. As Figure 1.6.2 indicates, positive values of t correspond to points on L that lie on one side of P_0 , whereas negative values of t correspond to points that lie on the other side of P_0 . If the vector $\mathbf{v}$ that gives the direction of the line L is written in component form as $\mathbf{v} = (a, b, c)$, then we have $t\mathbf{v} = (ta, tb, tc)$. We can also write $\mathbf{r} = (x, y, z)$ and $\mathbf{r}_0 = (x_0, y_0, z_0)$, so the vector equation (1) becomes $(x, y, z) = (x_0 + ta, y_0 + tb, z_0 + tc)$

Two vectors are equal if and only if corresponding components are equal. Therefore we have the three scalar equations: $x = x_0 + at$, $y = y_0 + bt$, $z = z_0 + ct$ (2)

Where $t \in \mathbb{R}$. These equations are called **parametric equations** of the line L through the point $P_0(x_0, y_0, z_0)$ and parallel to the vector $\mathbf{v} = (a, b, c)$. Each value of the parameter t gives a point (x, y, z) on L .

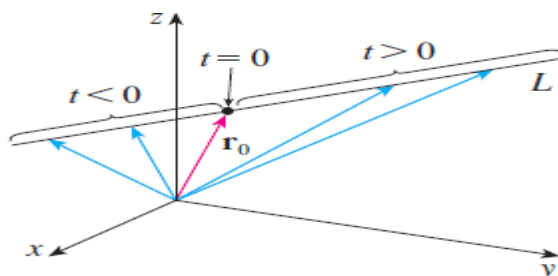


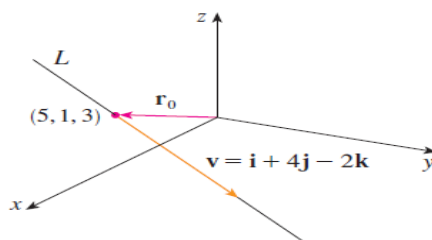
Figure 1.6.2

Example 1:

(a) Find a vector equation and parametric equations for the line that passes through the point $(5, 1, 3)$ and is parallel to the vector $\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}$.

(b) Find two other points on the line.

Solution:



(a) Here $\mathbf{r}_0 = (5, 1, 3) = 5\mathbf{i} + \mathbf{j} + 3\mathbf{k}$ and $\mathbf{v} = \mathbf{i} + 4\mathbf{j} - 2\mathbf{k}$, so the vector equation (1) becomes $\mathbf{r} = (5\mathbf{i} + \mathbf{j} + 3\mathbf{k}) + t(\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}) \Rightarrow \mathbf{r} = (5+t)\mathbf{i} + (1+4t)\mathbf{j} + (3-2t)\mathbf{k}$

Or Parametric equations are $x = 5 + t, y = 1 + 4t, z = 3 - 2t$

(b) Choosing the parameter value $t = 1$ gives $x = 6, y = 5$, and $z = 1$ so $(6, 5, 1)$ is a point on the line. Similarly $t = -1$, gives the point $(4, -3, 5)$.

The vector equation and parametric equations of a line are not unique. If we change the point or the parameter or choose a different parallel vector, then the equations change. For instance, if, instead of $(5, 1, 3)$, we choose the point $(6, 5, 1)$ in Example 1, then the parametric equations of the line become $x = 6 + t, y = 5 + 4t, z = 1 - 2t$. Or, if we stay with the point $(5, 1, 3)$ but choose the parallel vector $2\mathbf{i} + 8\mathbf{j} - 4\mathbf{k}$, we arrive at the

equations $x = 5 + 2t, y = 1 + 8t, z = 3 - 4t$. In general, if a vector $\mathbf{v} = (a, b, c)$ is used to describe the direction of a line L , then the numbers a, b , and c are called **direction numbers** of L . Since any vector parallel to $\mathbf{v}$ could also be used, we see that any three numbers proportional to a, b , and c could also be used as a set of direction numbers for L . Another way of describing a line L is to eliminate the parameter from Equations 2. If none of a, b , or c is 0, we can solve each of these equations for t , equate the results, and obtain

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \quad (3)$$

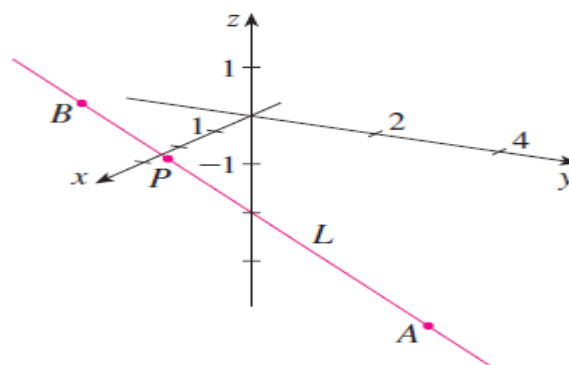
These equations are called **symmetric equations** of L . Notice that the numbers a, b , and c that appear in the denominators of Equations 3 are direction numbers of L , that is, components of a vector parallel to L . If one of a, b , or c is 0, we can still eliminate t . For instance, if $a = 0$, we could write the equations of L as

$$x = x_0, \frac{y - y_0}{b} = \frac{z - z_0}{c} \quad \text{This means that } L \text{ lies in the vertical plane } x = x_0.$$

Example 2:

- (a) Find parametric equations and symmetric equations of the line that passes through the points $A(2, 4, -3)$ and $B(3, -1, 1)$.
- (b) At what point does this line intersect the xy -plane?

Solution:



- (a) We are not explicitly given a vector parallel to the line, but observe that the vector $\mathbf{v}$ with representation line AB is parallel to the line and $\mathbf{v} = (3 - 2, -1 - 4, 1 - (-3)) = (1, -5, 4)$

Thus direction numbers are $a=1$, $b=-5$, and $c=4$. Taking the point $(2,4,-3)$ as P_0 , we see that parametric equations (2) are $x = 2 + t, y = 4 - 5t, z = -3 + 4t$ and symmetric equations (3) are $\frac{x-2}{1} = \frac{y-4}{-5} = \frac{z+3}{4}$

(b) The line intersects the xy -plane when $z = 0$, so we put $z = 0$ in the symmetric equations and obtain $\frac{x-2}{1} = \frac{y-4}{-5} = \frac{3}{4}$. This gives $x = \frac{11}{4}$ and $y = \frac{1}{4}$, so the line intersects the xy -plane at the point $\left(\frac{11}{4}, \frac{1}{4}, 0\right)$.

In general, the procedure of Example 2 shows that direction numbers of the line L through the points $P_0(x_0, y_0, z_0)$ and $P_1(x_1, y_1, z_1)$ are $x_1 - x_0, y_1 - y_0$, and $z_1 - z_0$ and so symmetric equations of L are

$$\frac{x-x_0}{x_1-x_0} = \frac{y-y_0}{y_1-y_0} = \frac{z-z_0}{z_1-z_0}$$

The vector equation of a line through the (tip of the) vector $\mathbf{r}_0$ in the direction of a vector $\mathbf{v}$ is $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$. If the line also passes through (the tip of) $\mathbf{r}_1$, then we can take $\mathbf{v} = \mathbf{r}_1 - \mathbf{r}_0$ and so its vector equation is $\mathbf{r} = \mathbf{r}_0 + t(\mathbf{r}_1 - \mathbf{r}_0) = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$

The line segment from $\mathbf{r}_0$ to $\mathbf{r}_1$ is given by the parameter interval $0 \leq t \leq 1$.

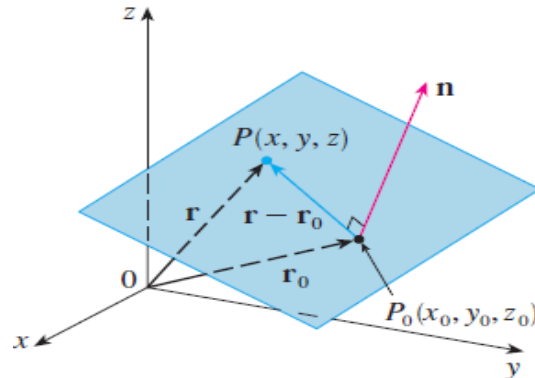
4 The line segment from $\mathbf{r}_0$ to $\mathbf{r}_1$ is given by the vector equation

$$\mathbf{r}(t) = (1-t)\mathbf{r}_0 + t\mathbf{r}_1 \quad 0 \leq t \leq 1$$

PLANES

Although a line in space is determined by a point and a direction, a plane in space is more difficult to describe. A single vector parallel to a plane is not enough to convey the “direction” of the plane, but a vector perpendicular to the plane does completely specify its direction. Thus a plane in space is determined by a point $P_0(x_0, y_0, z_0)$ in the plane and a vector $\mathbf{n}$ that is orthogonal to the plane. This orthogonal vector $\mathbf{n}$ is called a **normal vector**. Let $P(x, y, z)$ be an arbitrary point in the plane, and let $\mathbf{r}_0$ and $\mathbf{r}$ be the position vectors of P_0 and P . Then the vector $\mathbf{r} - \mathbf{r}_0$ is represented by a line P_0P . (See Figure below.) The normal vector $\mathbf{n}$ is orthogonal to every vector in the given plane. In

particular, $\mathbf{n}$ is orthogonal to $\mathbf{r}-\mathbf{r}_0$ and so we have $\mathbf{n} \cdot (\mathbf{r}-\mathbf{r}_0) = 0$ which can be rewritten as $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0$. And this Equation is called a **vector equation of the plane**. To obtain a scalar equation for the plane, we write $\mathbf{n} = (a, b, c)$, $\mathbf{r} = (x, y, z)$, and $\mathbf{r}_0 = (x_0, y_0, z_0)$. Then the above vector equation becomes $(a, b, c) \cdot (x - x_0, y - y_0, z - z_0) = 0$ or $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$. And this Equation is the **scalar equation of the plane through with normal vector $\mathbf{n} = (a, b, c)$** .



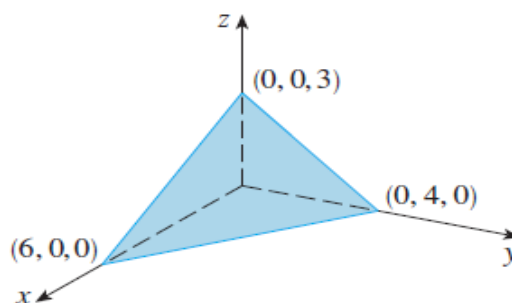
Example-1:

Find an equation of the plane through the point $(2, 4, -1)$ with normal vector $\mathbf{n} = (2, 3, 4)$.

Find the intercepts and sketch the plane.

Solution: Putting $a = 2, b = 3, c = 4, x_0 = 2, y_0 = 4, z_0 = -1$ in the scalar equation of the plane we see that an equation of the plane is $2(x - 2) + 3(y - 4) + 4(z + 1) = 0$ or $2x + 3y + 4z = 12$

To find the x-intercept we set $y = z = 0$ in this equation and obtain $x = 6$. Similarly, the y-intercept is 4 and the z-intercept is 3. This enables us to sketch the portion of the plane that lies in the first octant (see Figure below).



Example-2: Find an equation of the plane that passes through the points $P(1, 3, 2), Q(3, -1, 6)$ and $R(5, 2, 0)$

Solution: The vectors **a** and **b** corresponding to line PQ and line PR are $\mathbf{a} = (2, -4, 4)$, $\mathbf{b} = (4, -1, -2)$. Since both **a** and **b** lie in the plane, their cross product $\mathbf{a} \times \mathbf{b}$ is orthogonal to the plane and can be taken as the normal vector.

$$\text{Thus, } \mathbf{n} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} i & j & k \\ 2 & -4 & 4 \\ 4 & -1 & -2 \end{vmatrix} = 12i + 20j + 14k$$

With the point $P(1, 3, 2)$ and the normal

vector **n**, an equation of the plane is

$$12(x-1) + 20(y-3) + 14(z-2) = 0 \Rightarrow 6x + 10y + 7z = 50$$

Note : Two planes are **parallel** if their normal vectors are parallel. For instance, the planes $x + 2y - 3z = 4$ and $2x + 4y - 6z = 3$ are parallel because their normal vectors are $\mathbf{n}_1 = (1, 2, -3)$ and $\mathbf{n}_2 = (2, 4, -6)$ and $\mathbf{n}_2 = 2\mathbf{n}_1$. If two planes are not parallel, then they intersect in a straight line and the angle between the two planes is defined as the acute angle between their normal vectors.

SUMMARY OF UNIT-ONE

❖ vectors are physical quantities that have both magnitude and direction, and thus can not be described or represented by a single real number. Geometrically vectors are represented by a directed line segment or an arrow. We define a vector in n-space to be an n-tuple of numbers: $(x_1, x_2, \dots, x_n)$

Thus, $\mathfrak{R}^n = \{(x_1, x_2, \dots, x_n : x_i \in \mathfrak{R} \text{ for all } i \text{ such that } 1 \leq i \leq n)\}$

❖ For any vectors **u**, **v** and **w** and any scalars **r** and **s** we have:

- a) $u + v = v + u$
- b) $u + (v + w) = (u + v) + w$
- c) $v + 0 = v = 0 + v$
- d) $v + (-v) = 0$
- e) $(r + s)v = rv + sv$
- f) $r(u + v) = ru + rv$

$$g) (rs)v = r(sv) = s(rv)$$

$$h) 1v = v$$

$$i) 0v = 0 = r0$$

The above properties are properties of vector addition and scalar multiplication.

❖ In the xy -plane a vector v whose initial point is the origin and terminal point is

$$A(x, y) \text{ is represented by } v = (x, y) \text{ or } v = x_i + y_j$$

where $i = (1, 0)$ and $j = (0, 1)$ and we have

$$a) \|v\| = \sqrt{x^2 + y^2} \text{ where } x \text{ and } y \text{ are called the first and second component of } v.$$

$$b) \text{ If } p = (x_1, y_1) \text{ then vector } AP = A - P = (x, y) - (x_1, y_1) = (x - x_1, y - y_1)$$

$$c) kv = k(x, y) = (kx, ky)$$

$$d) \text{ If } u = (a, b) \text{ then } v + u = (x + a, y + b) \text{ and } v - u = (x - a, y - b)$$

$$e) x = \|v\| \cos \theta, y = \|v\| \sin \theta \text{ and } \tan \theta = \frac{y}{x}$$

❖ If θ is the angle between non zero vectors u and v the $u \cdot v$

$$u \cdot v = \|u\| \|v\| \cos \theta \text{ and } -1 \leq \frac{u \cdot v}{\|u\| \|v\|} \leq 1$$

❖ Given two non zero vectors u, v then

$$a) u \text{ and } v \text{ are perpendicular to each other, denoted by } u \perp v, \text{ if } u \cdot v = 0$$

$$b) u \text{ and } v \text{ are parallel, denoted by } u // v \text{ if there exists a scalar } k \text{ such that } u = kv.$$

❖ Let u, v and w be vectors $\mathbb{R}^n$ and $c, d \in \mathbb{R}$ then

$$a) u + v \text{ is a vector in } \mathbb{R}^n$$

$$b) u + v = v + u$$

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

$$d) u + 0 = u$$

$$e) u + (-u) = 0$$

$$f) cu \text{ is a vector in } \mathbb{R}^n$$

$$g) c(u + v) = cu + cv$$

$$h) (c + d)u = cu + du$$

$$i) c(du) = (cd)u = d(cu)$$

$$j) 1u = u$$

The above properties are properties of vector addition and scalar multiplication

❖ Let $u = (x_1, x_2, \dots, x_n)$ and $v = (y_1, y_2, \dots, y_n)$ be vectors in $\mathfrak{R}^n$ then

a) $\|u\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$

b) $u \cdot v = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$

c) $d(u, v) = \|u - v\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$

d) $|u \cdot v| \leq \|u\| \|v\|$, the Cauchy Schwarz Inequality

e) $\|u + v\| \leq \|u\| + \|v\|$, the triangle inequality

❖ Let L be a line which Passes through the point $p_0(x_0, y_0)$ and $A = (a_1, a_2)$ be a vector parallel to L then

I. $P = P_0 + tA$ where P is any point on L, is a parametric vector equation of L.

II. $ax + by + c = 0$, where $a = \frac{1}{a_1}$, $b = \frac{-1}{a_2}$, $c = \frac{-1}{a_1}x_0 + \frac{1}{a_2}y_0$

is a rectangular form of equation of L.

❖ Let $v_0 = (x_0, y_0, z_0)$ be a point in a plane Π and $N = (a, b, c)$ be a non zero vector to the plane Π . If $v = (x, y, z)$ is any point of the plane Π , then $ax + by + cz = ax_0 + by_0 + cz_0$ is the normal form of equation of the plane Π .

❖ Let $u = u_1i + u_2j + u_3k$ and $v = v_1i + v_2j + v_3k$ be vectors in $\mathfrak{R}^3$ then

$$u \times v = (u_2 v_3 - u_3 v_2)i - (u_1 v_3 - u_3 v_1)j + (u_1 v_2 - u_2 v_1)k$$

$$= \det \begin{bmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

❖ If u, v and w are vectors in $\mathfrak{R}^n$ and $c \in \mathfrak{R}$, then

I. $u \times v = -(v \times u)$

II. $c(u \times v) = (cu) \times v = u \times (cv)$

III. $u \times (v + w) = (u \times v) + (u \times w)$

IV. $u \times 0 = 0 = 0 \times u$

V. $u \times u = 0$

VI. $u \cdot (v \times w) = (u \times v) \cdot w$

The above properties are algebraic properties of the cross product

❖ If u and v are non zero vectors in $\mathbb{R}^n$, then

- I. $u \times v$ is orthogonal to both u and v
- II. The angle θ between u and v is given by $\|u \times v\| = \|u\| \|v\| \sin \theta$.
- III. u and v are parallel iff $u \times v = 0$
- IV. The parallelogram having u and v as adjacent sides has an area of $\|u \times v\|$

The above properties are geometric properties of the cross product

😊 Test yourself!

1. Determine whether each statement is true or false.

- (a) Two lines parallel to a third line are parallel.
- (b) Two lines perpendicular to a third line are parallel.
- (c) Two planes parallel to a third plane are parallel.
- (d) Two planes perpendicular to a third plane are parallel.
- (e) Two lines parallel to a plane are parallel.
- (f) Two lines perpendicular to a plane are parallel.
- (g) Two planes parallel to a line are parallel.
- (h) Two planes perpendicular to a line are parallel.
- (i) Two planes either intersect or are parallel.

2. Show that the lines L_1 and L_2 with parametric equations

$x = 1 + t, y = -2 + 3t, z = 4 - t$ and $x = 2s, y = 3 + s, z = -3 + 4s$ are skew lines; that is, they do not intersect and are not parallel (and therefore do not lie in the same plane).

3. Find the point at which the line with parametric equations $x = 2 + 3t, y = -4t, z = 5 + t$ intersects the plane $4x + 5y - 2z = 18$

4. Find the distance between the parallel planes $10x + 2y - 2z = 5$ and $5x + y - z = 1$

Miscellaneous Exercises-1

1. (a) Find the components of the vector w if $\|w\| = 5$ and the direction angle for w is $\theta = 300^\circ$
 - (b) Let $u = (3, 1, -5)$, $v = (\sqrt{2}, 1, -\sqrt{3})$, then find the components of $u + v$, $3u + 2v$, $-5u$, $u - v$.
3. (a) Find the parametric vector equation of the line that passes through $(1, 4)$ and
 - I. Parallel to the line $y = 3x - 1$
 - II. Perpendicular to the line $x - 2y = 5$
 - (b) A triangle has $(-1, -2)$, $(5, -3)$, and $(1, 2)$ as the coordinates of its vertices, then find the vector equations of its medians
2. (a) Let x and y be unit vectors and the angle between them is $\frac{\pi}{6}$ then find $\|x + y\|$ and $\|x - y\|$
 - (b) Let u and v be vectors with $\frac{\pi}{3}$ an angle between them. If $\|u\| = 2$ and $\|v\| = 3$ Then find
 - (i) $u \cdot v$
 - (ii) $u \cdot u$
 - (iii) $(u + v)^2$
 - (iv) $(3u - 2v)$
 - (v) $(u - 2v)$
 - (vi) $(3u - 2v) \cdot (3v + 2u)$
 - (c) Find a unit vector whose direction is opposite to $3i - 6j + k$.
6. (a) Find a unit vector that is orthogonal to both $u = i - 4j + k$ and $v = 2i + 3j$
 - (b) Find the area of a parallelogram that has $u = -3i + 4j + k$ and $v = -2j + 6k$ as adjacent sides.
7. Find an equation of the plane passing through the given points
 - a) $(1, -2, 1)$, $(-1, -1, 7)$ and $(2, -1, 3)$
 - b) $(0, -1, 0)$, $(1, 1, 0)$, $(2, 1, 2)$
8. Prove that
 - a) If $\|u + v\| = \|u - v\|$ then $\|u\| = \|v\|$
 - b) If $u \cdot (v + w) = v \cdot (u - w)$, then $w \perp u + v$
 - c) If the cosine of the angle between the vectors $2i + kj$ and $ki + j$ is $\frac{3}{\sqrt{10}}$ then $k = 1$ or $k = 2$

UNIT TWO: VECTOR SPACES

UNIT OBJECTIVES:

At the end of this unit each student should be able to:

- Know axioms of vector spaces
- Understand the difference between linearly dependence and linearly independence
- Identify base and dimension of vector space.
- Know subspace of vector space.
- Understand direct sum and direct product of subspaces
- Learn linear combinations of vectors
- Know generators of vector space

Introduction

In the Previous Chapter, we listed ten Special Properties of Vector addition and Scalar multiplication in $\mathbb{R}^n$. Suitable definitions of addition and scalar multiplication reveal that many other mathematical quantities such as Polynomials functions and Matrices also share these ten properties. We call any non empty set that satisfies these properties (or axioms) a vector space, and the objects in the set are called vectors. It is important to realize that the following definition of a vector space is precisely that a definition. We do not need to prove any thing, since we are simply listing the axioms required of vector spaces. We call this type of definition an abstraction because we are abstracting a collection of properties for a more general setting.

2.1 The Axioms of Vector Space

Objectives

At the end of this section, the learners expected to:

- Define vector space.
- List all properties of vector space.
- Use axioms of vector space to show whether a function is a vector space over the given domain or not.
- State and prove Cancellation law for vector addition.

Definition 2.1: Definition of a Vector Space

Let V be a set on which two operations (vector addition and scalar multiplication) are defined. If the following axioms are satisfied for every u, v and w in V and for every scalar (an element of a field F) c and d , then V is called vector space over a field F .

Addition:

- A-1 $u + v$ is in V ----- Closure under addition
- A-2 $u + v = v + u$ ----- Commutative Property
- A-3 $u + (v + w) = (u + v) + w$ ----- Associative property
- A-4 V has a zero vector 0 such that ---- Existence of Additive Identity for every u in V , $u + 0 = u$
- A-5 For every u in V , there is ----- Existence of additive inverse a vector in V denoted by $-u$ Such that $u + (-u) = 0$ where 0 is the identity element

Scalar Multiplication:

- A-6 $c u$ is in V ----- Closure under scalar multiplication
- A-7- $c(u + v) = cu + cv$ ----- Distributive property
- A-8- $(c+d)u = cu + du$ ----- Distributive Property
- A-9- $c(du) = (cd)u = d(cu)$ ----- Associative Property
- A-10- $1(u) = u$, where $1 \in F$ --- -- Scalar Identity



Remark 2.1

1. It is important to realize that a vector space consists of four entities: a set of vectors, a set of scalars, and the two operations. When you refer to a vector space V , be sure that all four entities are clearly stated or understood,
- 2- A vector space together with the defined operation addition is Abelian group.

Class Activity- 2.1

- 1 - Is a vector space V can be empty set? What ever your answer comes with justification
- 2 - Identify the difference between vector space axioms 7 and 8 (A-7 & A-8)
- 3 - Is the zero vectors invertible under vector addition? If so, find its additive inverse

2.2 Examples of different models of a vector space

2.2.1 The Axioms of vector space

Throughout mathematics one encounters many examples of mathematical objects that can be added to each other and multiplied by numbers. Geometric vectors and vectors in n -space studied in the preceding chapter are examples of such objects. Other examples are real-valued functions, complex numbers, and polynomials over the real. In this chapter we discuss a general mathematical concept which includes these examples and many more.

Throughout this chapter, by a “field” we shall mean either the field of rational numbers, we shall, however, use a neutral letter k , since it is necessary to deal with each of these fields. The elements of k will also be called numbers or scalars.

Definition 2.2.1

A vector space V over a field is a set of objects which can be added and multiplied by elements of such that the following axioms are satisfied

Axiom 1: (closure under addition)

For all $v_1, v_2 \in V$, there exists a unique element of V called the
Sum of v_1 and v_2 , denoted $v_1 + v_2$

Axiom 2: (closure under multiplication by elements of k)

For every $v \in V$ and all $a \in k$,
there exists an element of V called the product of a and v , denoted av .

Axiom 3: (commutative law for addition)

For all $v_1, v_2 \in V$, we have $v_1 + v_2 = v_2 + v_1$

➤ Examples of Vectorspaces

Example-1 Let $V = \mathbb{R}^+$ and $b \in \mathbb{R}$. Define: $x_1 + x_2 = x_1 x_2 = \forall x_1, x_2 \in \mathbb{R}^+$, usual product in $\mathbb{R}$ and $bx = x^b \forall x \in \mathbb{R}^+$, then show that V is a vector space over $\mathbb{R}$ with 1 as the zero element.

Solution:

1. Let $x_1, x_2 \in V = \mathbb{R}^+$ then $x_1 + x_2 = x_1 x_2 \in \mathbb{R}^+$

$\Rightarrow x_1 + x_2 \in V$, Since $\mathbb{R}^+$ is closed under usual multiplication

$\Rightarrow V$ is closed under vector addition

2. Let $x_1, x_2 \in V = \mathbb{R}^+$ then $x_1 + x_2 = x_1 x_2 = x_2 x_1$ (Why?)

$$= x_2 + x_1$$

$$\Rightarrow x_1 + x_2 = x_2 + x_1 \quad \forall x_1, x_2 \in \mathbb{R}^+$$

$\Rightarrow$ Vector addition is commutative

3. Let $x_1, x_2, x_3 \in V$ then

$$(x_1 + x_2) + x_3 = (x_1 x_2) + x_3 \quad \text{by definition of addition in } V$$

$$= (x_1 x_2) x_3, \quad \text{by definition of addition in } V$$

$$= x_1 (x_2 x_3), \quad (\text{Why?})$$

$$= x_1 (x_2 + x_3), \quad \text{by definition of '+' in } V$$

$$= x_1 + (x_2 + x_3), \quad \text{by definition of '+' in } V$$

$$\text{Thus, } (x_1 + x_2) + x_3 = x_1 + (x_2 + x_3) \quad \forall x_1, x_2, x_3 \in V$$

$\Rightarrow$ Vector addition is associative.

4. Suppose e is a zero vector in V then

$$e + x = x = x + e, \quad \forall x \in V$$

$$\text{Now, } e + x = x$$

$$\Rightarrow e.x = x, \quad \text{By definition of '+' in } V$$

$$\Rightarrow e = 1, \quad (\text{How?})$$

$$\Rightarrow \text{The zero vector in } V = \mathbb{R}^+ \text{ is } 1$$

$$\text{Therefore, } \exists 1 \in V \text{ such that } 1 + x = x = x + 1 \quad \forall x \in V$$

$\Rightarrow$ Additive identity exists in V

5. Let $x \in V = \mathbb{R}^+$ suppose $y \in V$ is

An additive inverse of x then

$$x + y = 1 = y + x$$

But, c

$$\Rightarrow x.y = 1$$

$$\Rightarrow y = 1/x \quad \text{--- (Why?)}$$

$$\text{Hence, } \forall x \in V \quad \exists y = 1/x \in V \text{ such that } x + y = 1$$

$\Rightarrow$ Every element in V has an additive inverse

6. Let $x \in V$ and $a \in \mathbb{R}$ then $ax = x^a \in V$ (why?)

$$\Rightarrow ax \in V \quad \forall a \in \mathbb{R} \text{ And } \forall x \in V$$

$\Rightarrow V$ is closed under scalar multiplication.

7. Let $a \in \mathfrak{R}$ and $x, y \in V$ then

$$\begin{aligned}
 a(x+y) &= a(x.y) && \text{By definition of '+' in } V \\
 &= (x.y)^a && \text{by definition of '.' in } V \\
 &= x^a.y^a && \text{(Why?)} \\
 &= x^a + y^a && \text{by definition of '+' in } V = \mathfrak{R}^+ \\
 &= ax + ay && \text{by definition of '.' in } V = \mathfrak{R}^+
 \end{aligned}$$

Therefore, $a(x+y) = ax + ay \quad \forall a \in \mathfrak{R} \text{ and } \forall x, y \in V$

8. Let $a, b \in \mathfrak{R}$ and $x \in V = \mathfrak{R}^+$ then

$$\begin{aligned}
 (a+b)x &= x^{a+b}, && \text{by definition of scalar multiplication in } V \\
 &= x^a x^b, && \text{(Why?)} \\
 &= x^a + x^b, && \text{by definition of '+' in } V \\
 &= ax + bx, && \text{by definition of '.' In } V
 \end{aligned}$$

$$\Rightarrow (a+b)x = ax + bx \quad \forall a, b \in \mathfrak{R} \text{ And } \forall x \in V$$

9. Let $a, b \in \mathfrak{R}$ and $x \in V = \mathfrak{R}^+$ then

$$\begin{aligned}
 a(bx) &= a(x^b), && \text{by definition of '.' in } V \\
 &= (x^b)^a, && \text{by definition of '.' in } V \\
 &= x^{ba}, && \text{Why?} \\
 &= x^{ab} \\
 &= (ab)x
 \end{aligned}$$

$$\Rightarrow a(bx) = (ab)x \quad \forall a, b \in \mathfrak{R} \text{ And } \forall x \in V$$

10. Let $x \in V = \mathfrak{R}^+$ then

$$1. x = x^1 = x$$

$$\Rightarrow 1.x = x \quad \forall x \in V$$

Hence, by 1-10 we have $V = \mathfrak{R}^+$ is a vector space over $\mathfrak{R}$ with the given addition and scalar multiplication defined on it.

Example -2 Let P_2 be the set of all polynomials over $\mathfrak{R}$ of degree less than or equal to 2 together with zero polynomial then show that P_2 is a vector space over $\mathfrak{R}$ with the usual addition and scalar multiplication of polynomials.

That is, let $P_2 = \{ax^2 + bx + c : a, b, c \in \mathbb{R}\}$ and $P(x) = a_2x^2 + a_1x + a_0$, and $q(x) = b_2x^2 + b_1x + b_0 \in P_2$

Define

$P(x) + q(x) = (a_2x^2 + a_1x + a_0) + (b_2x^2 + b_1x + b_0) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$ and $c \cdot P(x) = c(a_2x^2 + a_1x + a_0) = (ca_2)x^2 + (ca_1)x + (ca_0)$ Then P_2 is a vector space over $\mathbb{R}$

Solution:

Verification of each of the ten vector space axioms is a straight forward application of the properties of real numbers.

For instance, let $P(x) = a_2x^2 + a_1x + a_0, q(x) = b_2x^2 + b_1x + b_0 \in P_2$

Because of set of real numbers is closed under addition, it follows that $a_2 + b_2, a_1 + b_1, a_0 + b_0$ are real numbers, and

$$P(x) + q(x) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$$

$$= c_2x^2 + c_1x + c_0, \text{ Where}$$

$$c_2 = a_2 + b_2, c_1 = a_1 + b_1, c_0 = a_0 + b_0 \in \mathbb{R}$$

$$\Rightarrow P(x) + q(x) = c_2x^2 + c_1x + c_0 \in P_2, \text{ Since it is a polynomial of degree two or less}$$

$$\Rightarrow P_2 \text{ is closed under addition.}$$

Similarly, we can use the fact that the set of real numbers is closed under scalar multiplication to show that P_2 is closed under scalar multiplication. To verify the commutative axiom of addition, we write

$$P(x) + q(x) = (a_2x^2 + a_1x + a_0) + (b_2x^2 + b_1x + b_0)$$

$$= (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0), (\text{Why?})$$

$$= (b_2 + a_2)x^2 + (b_1 + a_1)x + (b_0 + a_0)$$

$$= (b_2x^2 + b_1x + b_0) + (a_2x^2 + a_1x + a_0)$$

$$= q(x) + P(x)$$

$$\Rightarrow P(x) + q(x) = q(x) + P(x) \forall P, q \in P_2$$

$$\Rightarrow \text{addition in } P_2 \text{ is commutative}$$

Class activity 2-2

In the above example -2

1. Can you see where we used the commutative property of addition of real number?
2. The zero vectors in this space is the zero polynomial given by $0(x) = 0 \quad \forall x$, why?
3. Try to verify the other vector space axioms. You then may conclude that P_2 is a vector space over $\mathfrak{R}$
4. Let P_n be the set of all polynomials over $\mathfrak{R}$ of degree less than or equal to n together with the zero polynomial then show that P_n with the usual operations of polynomial addition and scalar multiplication is a vector space over $\mathfrak{R}$

Example-3 Let $V = \{(x, 2x) : x \in \mathfrak{R}\}$ then show that V is a vector space over $\mathfrak{R}$ with Standard operations (componentwise addition and multiplication)

Solution:

1. Let $u, v \in V$ then $u = (x, 2x)$ for some $x \in \mathfrak{R}$ and $v = (y, 2y)$ for some $y \in \mathfrak{R}$.

$$\begin{aligned}
 \text{Now, } u+v &= (x, 2x) + (y, 2y) \\
 &= (x+y, 2x+2y), \text{ By definition '+' in } V \\
 &= (x+y, 2(x+y)) \\
 &= (z, 2z), \text{ Where } z = x+y \in \mathfrak{R}
 \end{aligned}$$

$$\Rightarrow u+v = (z, 2z) \text{ For some } z = x+y \in \mathfrak{R}$$

$$\Rightarrow u+v \in V, \text{ by definition of the set } V$$

$$\Rightarrow V \text{ is closed under vector addition.}$$

2. Let $u = (x, 2x), v = (y, 2y) \in V$ then

$$\begin{aligned}
 u+u &= (x, 2x) + (y, 2y) \\
 &= (x+y, 2x+2y) \\
 &= (y+x, 2y+2x); \text{ Since addition is commutative in } \mathfrak{R} \\
 &= (y, 2y) + (x, 2x), \text{ by definition of '+' in } V \\
 &= v+u
 \end{aligned}$$

$$\Rightarrow u+v = v+u \quad \forall u, v \in V$$

$$\Rightarrow \text{Vector addition is commutative in } V$$

3. Let $u = (x, 2x), v = (y, 2y), w = (z, 2z) \in V$ then

$$\begin{aligned}
 (u + v) + w &= [(x, 2x) + (y, 2y)] + (z, 2z) \\
 &= (x + y, 2x + 2y) + (z, 2z), \text{ By definition of '+' in } V \\
 &= ((x + y) + z, (2x + 2y) + 2z), \text{ by definition of '+' in } V \\
 &= (x + (y + z), 2x + (2y + 2z)) ; \text{ Since the usual addition is} \\
 &\quad \text{associative in } \mathfrak{R} \\
 &= (x, 2x) + (y + z, 2y + 2z); \text{ By definition '+' in } V \\
 &= (x, 2x) + [(y, 2y) + (z, 2z)] ; \text{ By definition of '+' in } V \\
 &= u + (v + w)
 \end{aligned}$$

$$\Rightarrow (u + v) + w = u + (v + w) \quad \forall u, v, w \in V$$

$\Rightarrow$ The defined addition is associative in V

4. Since $0 \in \mathfrak{R}$ we have

$$0 = (0, 0) \in V \text{ And } \forall u = (x, 2x) \in V,$$

$$0 + u = (0, 0) + (x, 2x) = (0 + x, 0 + 2x)$$

$$= (x, 2x)$$

$$= u$$

$$\Rightarrow 0 + u = u \quad \forall u \in V$$

$$\Rightarrow 0 = (0, 0) \text{ is the additive identity element in } V$$

$$\Rightarrow \text{Additive identity element exists in } V$$

5. Let $u = (x, 2x) \in V$ then $-u = (-x, -2x) \in V$ and

$$u + -u = (x, 2x) + (-x, -2x)$$

$$= (x + -x, 2x + -2x)$$

$$= (0, 0)$$

$$\Rightarrow u + -u = 0$$

$$\Rightarrow \forall u \in V \quad \exists -u \in V \text{ Such that } u + -u = 0$$

$$\Rightarrow \text{Every element in } V \text{ has an additive inverse.}$$

6. Let $a \in \mathfrak{R}$ and $u = (x, 2x) \in V$ then

$$a \cdot u = a(x, 2x)$$

$$= (ax, 2ax), \text{ by definition of scalar multiplication}$$

$$= (y, 2y); \text{ Where } y = ax \in \mathfrak{R}$$

$$\Rightarrow a \cdot u = (y, 2y) \text{ For some } y = ax \in \mathfrak{R}$$

$\Rightarrow au \in V$, why?

$\Rightarrow V$ is closed under scalar multiplication.

$$\begin{aligned}
 7. \text{ Let } a \in \mathfrak{R} \text{ and } u = (x, 2x), v = (y, 2y) \in V \text{ then } a(u + v) &= a((x, 2x) + (y, 2y)) \\
 &= a(x + y, 2x + 2y), \text{ By definition of '+' in } V \\
 &= (a(x + y), a(2x + 2y)) \text{ by definition of scalar multiplication in } V \\
 &= (ax + ay, 2ax + 2ay), \text{ Why?} \\
 &= (ax, 2ax) + (ay, 2ay), \text{ by definition of scalar multiplication in } V \\
 &= a(x, 2x) + a(y, 2y) \\
 &= au + av \\
 \Rightarrow a(u + v) &= au + av \quad \forall a \in \mathfrak{R} \forall u, v \in V
 \end{aligned}$$

$$\begin{aligned}
 8. \text{ Let } a, b \in \mathfrak{R} \text{ and } u = (x, 2x) \in V \text{ then } (a + b)u &= (a + b)(x, 2x) \\
 &= ((a + b)x, (a + b)2x) ; \text{ By definition of scalar multiplication in } V \\
 &= (ax + bx, 2ax + 2bx) ; \text{ Why?} \\
 &= (ax, 2ax) + (bx, 2bx) ; \text{ By definition of '+' in } V \\
 &= a(x, 2x) + b(x, 2x) ; \text{ Why?} \\
 &= au + bu \\
 \Rightarrow (a + b)u &= au + bu \quad \forall a, b \in \mathfrak{R} \text{ And } \forall u \in V
 \end{aligned}$$

$$\begin{aligned}
 9. \text{ Let } a, b \in \mathfrak{R} \text{ and } u = (x, 2x) \in V \text{ then} \\
 a(bu) &= a(b(x, 2x)) \\
 &= a(bx, 2bx) ; \text{ By definition of scalar multiplication in } V \\
 &= (a(bx), a(2bx)) ; \text{ Why?} \\
 &= ((ab)x, (ab)2x) ; \text{ Why?} \\
 &= ab(x, 2x) ; \text{ Why?} \\
 &= (ab)u \\
 \Rightarrow a(bu) &= (ab)u \quad \forall a, b \in \mathfrak{R} \text{ And } \forall u \in V
 \end{aligned}$$

$$\begin{aligned}
 10. \text{ Let } u = (x, 2x) \in V \text{ then} \\
 1.u &= 1(x, 2x) = (1.x, 1.2x) \\
 &= (x, 2x) \\
 &= u
 \end{aligned}$$

$$\Rightarrow 1.u = u \forall u \in V$$

Therefore, by 1-10 we've V is a vector space over $\mathfrak{R}$

➤ **Examples of sets, which are, not vector space**

Example-1 : Show that the set of rational numbers Q is not a vector space with standard operations over $\mathfrak{R}$

Solution Since $\sqrt{2} \in \mathfrak{R}$ and $1 \in Q$ such that

$$\sqrt{2} \cdot 1 = \sqrt{2} \notin Q \text{ We've } Q \text{ is not closed under scalar multiplication.}$$

Example-2: Let $V = \mathfrak{R}^3$, the set of ordered triple of real numbers, then show that V is not a vector space over $\mathfrak{R}$ with the standard operation of addition and the following non-standard definition of scalar multiplication:

$$a(x_1, x_2, x_3) = (ax_1, ax_2, 0)$$

Solution Since $1 \in \mathfrak{R}$ we've $u = (1,1,1) \in V$ and $1 \cdot (1,1,1) = (1,1,1,0) = (1,1,0)$

$$\Rightarrow 1 \cdot (1,1,1) = (1,1,0) \neq (1,1,1) = u$$

$$\Rightarrow 1 \cdot (1,1,1) \neq (1,1,1) = u$$

$$\Rightarrow 1 \cdot u \neq u$$

$$\Rightarrow V \text{ Doesn't satisfy one of the vector space axioms}$$

$$\Rightarrow V \text{ is not a vector space over } \mathfrak{R}$$

Infact, show that V satisfies the first nine axioms.



SRemark 2.2

A single failure of one of the ten vector space axioms suffices to show that a set is not a vector space. That is to show that a set V is not a vector space we need only find one thing wrong regardless of how many of the ten axioms are satisfied.

Theorem 2.2.1 (Cancellation law for vector addition)

Let V be a vector space over a field F and $u, v, w \in V$.

If $u + w = v + w$ then $u = v$

Proof Suppose $u + w = v + w$ then $(u + w) + -w = (v + w) + -w$ but,

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

$$\Rightarrow u + (w + -w) = v + (w + -w); \text{ Since vector addition is associative}$$

$$\Rightarrow u + o = v + o; \text{ Since } w + -w = 0$$

$\Rightarrow u = v$, Since 0 is additive identity element in V

Therefore, $u + w = v + w \Rightarrow u = v$ and hence the proof

Theorem 2.2.2 (Properties of scalar multiplication)

Let V be a vector space over a field F then

1- $a0 = 0 \quad \forall a \in F$

2- $0u = 0 \quad \forall u \in V$

3- $(-a)u = -(au) \quad \forall a \in F$ And $\forall u \in V$ In particular, $(-1)u = -u \quad \forall u \in V$

4. $au = 0 \Rightarrow a = 0$ Or $u = 0$, Where $a \in F$ and $u \in V$

Proof:

1. Let $a \in F$

$$a \cdot 0 = a(0 + 0); \text{ since } 0 + 0 = 0$$

$$= a0 + a0; \text{ By A-7}$$

$$\Rightarrow a0 = a0 + a0 \text{-----} (*)$$

$$\text{Also, we know that } 0 + a0 = a0 + 0 = a0 \text{-----} (**)$$

Now, combining $(*)$ and $(**)$ we get

$$0 + a0 = a0 = a0 + a0$$

$$\Rightarrow 0 + a0 = a0 + a0$$

$$\Rightarrow 0 = a0; \text{ by theorem 2.1}$$

$$\text{Thus, } a0 = 0 \quad \forall a \in F$$

2. Let $u \in V$ then $0u = (0 + 0)u$; since 0 is additive identity element in F .

$$\text{Now, } 0u = (0 + 0)u$$

$$= 0u + 0u; \text{ by A-8}$$

$$\Rightarrow 0u = 0u + 0u$$

$$\Rightarrow 0 + 0u = 0u = 0u + 0u; \text{ Since } 0 + 0u = 0u$$

$$\Rightarrow 0 + 0u = 0u + 0u$$

$$\Rightarrow 0 = 0u; \text{ By theorem 2.1}$$

$$\text{Hence, } 0u = 0 \quad \forall u \in V$$

3. Let $a \in F$ and $u \in V$ then $-a \in F, -(a)u \in V$

$$\text{And } 0 = 0u; \text{ by 2-above}$$

$$= (-a + a)u; \text{ since } -a + a = 0$$

$$= (-a)u + au; \text{ By A-8}$$

$$\Rightarrow (-a)u + au = 0$$

$$\Rightarrow (-a)u + au = 0 = au + (-a)u; \text{ Since '+' is commutative in } V$$

$$\Rightarrow \text{The additive inverse of } au \text{ is } (-a)u \text{ in } V$$

$$\Rightarrow -(au) = (-a)u$$

$$\text{Therefore } -(au) = (-a)u \quad \forall a \in F, \quad \forall u \in V$$

$$\text{If } a = 1 \text{ then } -(1u) = (-1)u \Rightarrow -u = (-1)u$$

4. Let $a \in F, u \in V$ and suppose that $au = 0$

Claim: either $a = 0$ or $u = 0$

If $a = 0$, we are done!

If $a \neq 0$ then a^{-1} exists in F and

$$au = 0$$

$$\Rightarrow a^{-1}(au) = a^{-1}0$$

$$\Rightarrow (a^{-1}a)u = a^{-1}0, \text{ By A-9}$$

$$\Rightarrow (a^{-1}a)u = 0, \text{ By 1- above}$$

$$\Rightarrow 1u = 0, \text{ Since } a^{-1}a = 1$$

$$\Rightarrow u = 0; \text{ By A-10}$$

Thus, $\forall a \in F$ and $u \in V$ such that $au = 0$ implies either $a = 0$ or $u = 0$

☺ Test yourself!

1. Find the zero vectors of the following vector spaces with standard operations.

A. p_3 = the set of all polynomials over $\mathbb{R}$ of degree ≤ 3 together with zero Vector.

B. $C(-\infty, \infty)$ = The set of all real valued continuous function defined on the real line.

C. $\mathbb{R}^n = \{(x_1, x_2, x_3, \dots, x_n) : x_i \in \mathbb{R} \forall i \text{ such that } 1 \leq i \leq n\}$

= The set of all ordered n-tuples of real numbers

2. Describe the additive inverse of a vector u in the given vector space with standard operations.

A. p_3 over $\mathbb{R}$

B. $C(-\infty, \infty)$ over $\mathbb{R}$

C. $\mathbb{R}^n$ over $\mathfrak{R}$

3. Determine whether the given set, together with the indicated operations and field F is a vector space over a field F . If it is not, identify at least one of the ten vector space axioms that fail.

A. $C[0, 1]$ = the set of all real valued continuous functions defined on $[0, 1]$

With the standard operations and $F = \mathfrak{R}$.

B. $\mathfrak{R}$ = the set of real numbers with standard operations and F = the set of complex numbers $\mathbb{C}$

C. The set of all fourth degree polynomials over $\mathbb{R}$ with standard operations and $F = \mathfrak{R}$

D. The set $\{(a, b); a \geq 0 \text{ and } b \in \mathbb{R}\}$ with standard operation in $\mathbb{R}^2$ and $F = \mathfrak{R}$

4. Rather than using the standard definitions of addition and scalar multiplication in $\mathbb{R}^2$ suppose that we define the following two operations

A. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$

$$a(x, y) = (ax, y)$$

B. $(x_1, y_1) + (x_2, y_2) = (x_1, 0)$

$$a(x, y) = (ax, ay)$$

C. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$

$$a(x, y) = (\sqrt{a}x, \sqrt{a}y)$$

With these new definitions, is $\mathbb{R}^2$ vector space over $\mathfrak{R}$? Justify your answer!

5. Rather than using the standard definitions of addition and scalar multiplication in $\mathbb{R}^3$, suppose that we define the following two operations:

A. $(a, b, c) + (d, e, f) = (a + d, b + e, c + f)$

$$\beta(a, b, c) = (\beta a, \beta b, 0)$$

B. $(a, b, c) + (d, e, f) = (0, 0, 0)$

$$\beta(a, b, c) = (\beta a, \beta b, \beta c)$$

C. $(a, b, c) + (d, e, f) = (a + d + 1, b + e + 1, c + f + 1)$

$$\beta(a, b, c) = (\beta a, \beta b, \beta c)$$

With these new definitions, is $\mathbb{R}^3$ is a vector space over $\mathfrak{R}$? Justify your answer!

6. Let V be a vector space over a field F then prove that.

- A. The zero vector is unique
- B. The additive inverse of a vector is unique
- C. Cancellation law for scalar multiplication in V

That is, if $\beta u = \beta v$ and $\beta \neq 0$ then $u = v$ where

$$u, v \in V \text{ And } \beta \in F.$$

7. Let F be a field then show that $F^n = \{(x_1, x_2, \dots, x_n); x_i \in F \text{ with } 1 \leq i \leq n\}$ with Component wise addition and scalar multiplication is a vector space over F .

2.3 Subspaces, linear combinations and generators

In this section, we are going to discuss definition of subspaces, examples of subspaces, sum, union, intersection of two subspaces and test for subspace.

In most important applications in linear algebra, vector space occurs as subspaces of larger subspaces. We say that a subset of a vector space is a subspace if it is itself a vector space (with the same operations).

Objectives

At the end of this section learners should be able to

- Define a subspaces of a vector space.
- For any vector space, describe the “smallest” and “largest” subspaces of vector space.
- State and proof the theorem of **test for subspace**.
- Show the intersection of two subspaces is also a subspace.

2.3.1. Subspaces

Definition 2.2 (definition of subspace of a vector space)

Let V be a vector space over a field F and W be a subset of V , then we say that W is a subspaces of V if W is itself a vector space over F under the operations of addition and scalar multiplication in V .

Class Activity 2-3.1

1. Let $V =$ the set of all real valued continuous functions defined on $[0,1]$ and W be the set of real valued differentiable functions on $[0,1]$ then show that
 - A) V is a vector space over $\mathbb{R}$ with standard operations
 - B) W is a subspace of V
2. Let $V = \mathbb{R}^4$ and $W = \{(0, x, y, z); x, y, z \in \mathbb{R}\}$ then show that
 - A) V with the standard operations is a vector space over $\mathbb{Q}$.
 - B) W is a subspace of V .

To establish that a set W is a vector space we must verify all the ten vector space axioms. However if W is a subset of a larger vector space V (and the operations defined on W are the same as those defined on V), then most of the ten properties are inherited from the larger space and need no verification. The following theorem tells us that it is sufficient to test for closure in order to establish a subset of a vector space is a subspace.

Theorem 2.3.1.1- Test for a subspace

Let V be a vector space over a field F and W be an empty subset of V , then W is a subspace of V over F if and only if the following closure conditions hold:

- I) if u and v are in W , then $u+v$ is in W
- II) If u is in W and $\beta \in F$, then βu is in W

Proof: Let V be a vector space over a field F and $\emptyset \neq W \subseteq V$

($\Rightarrow$) Suppose W is a subspace of V over F then by definition of subspace we have W is a vector space over F under the induced operations and hence W is closed under addition and scalar multiplication.

$$\Rightarrow u + v \in W \quad \forall u, v \in W \quad \text{And} \quad \beta u \in W \quad \forall \beta \in F, \forall u \in W$$

$\Rightarrow$ I) and II) hold.

($\Leftarrow$) Suppose I) and II) hold then W is closed under addition and scalar multiplication.

$\Rightarrow$ A-1 and A-6 are satisfied ----- (Δ)

Also, let $u, v, w \in W$ and $a, b \in F$ then

$$u, v, w \in V \quad \text{and} \quad a, b \in F \quad ; \text{Since } W \subseteq V$$

$\Rightarrow$ A-2, A-3, A-7, A-8, A-9, and A-10 are satisfied automatically, by definition of V ----- $(\Delta\Delta)$

Furthermore, since $w \neq \phi$ we've $\exists u_1 \in W$ and hence $\beta u_1 \in W \quad \forall \beta \in F$; by II

In particular, for $\beta = 0$ we get $0u_1 \in W$

$\Rightarrow 0 \in W$; Since $0u_1 = 0$ by theorem 2-2

$\Rightarrow W$ Has additive identity element; why?

$\Rightarrow$ A-4 is satisfied ----- $(\Delta\Delta\Delta)$

In addition, because W is closed under scalar multiplication, it follows that

$\forall u \in W \quad (-1)u \in W$

$\Rightarrow -u \in W$; by theorem 2-2

$\Rightarrow$ every element in W is invertible with respect to the induced additive operation

$\Rightarrow$ A-5 is satisfied ----- $(\Delta\Delta\Delta\Delta)$

Thus, by (Δ) , $(\Delta\Delta)$, $(\Delta\Delta\Delta)$ and $(\Delta\Delta\Delta\Delta)$ we have W is a vector space over F .

$\Rightarrow W$ is a subspace of V over F .



Remark 2-3:

1. If W is a subspace of a vector space V over a field F then both W and V must have the same zero vector O . Justify!
2. Since a subspace of a vector space is itself a vector space, it must contain the zero vector

In fact, the simplest subspace of a vector space is the one consisting of only the zero

$$\text{Vector} = \{0\} = W = \{0\}$$

Definition 2-3.1.1 Let V be a vector space over a field F then

A) V and $\{0\}$ are called improper (trivial) subspaces of V .

B) Subspaces of V other than V and $\{0\}$ are called proper (non trivial)

Subspaces of V

Observe that Improper subspaces of a vector space always exist while the proper Subspace of vector spaces may not exist.

Example (1) Let V be the set of all real valued functions defined on $\mathfrak{R}$ and

$W = \{f : \mathfrak{R} \rightarrow \mathfrak{R} / f \text{ is even function}\}$ then show that

- A) With the standard operations, is a vector space over $\mathfrak{R}$
- B) W is a subspace of V over $\mathfrak{R}$

Solution A) Left as an exercise to you.

B) Let $V = \{f / f : \mathfrak{R} \rightarrow \mathfrak{R} \text{ is a function}\}$ and

$W = \{f : \mathfrak{R} \rightarrow \mathfrak{R} / \text{is even function}\}$ then

$$W = \{f : \mathfrak{R} \rightarrow \mathfrak{R} / f(x) = f(-x) \forall x \in \mathfrak{R}\}$$

I) Clearly, $W \subseteq V$; why?

Also, consider the function f_0 that has a value of zero for all x .

That is, $f_0(x) = 0, x$ is any real number. Then $f_0 : \mathfrak{R} \rightarrow \mathfrak{R}$ and $f_0(x) = 0 = f_0(-x)$

$$\Rightarrow f_0 \in W$$

$$\Rightarrow W \neq \emptyset. \text{ Thus, } \emptyset \neq W \subseteq V$$

II) Let $F, G \in W$ then $F, G : \mathfrak{R} \rightarrow \mathfrak{R}$ Such that both are even functions and hence $F + G : \mathfrak{R} \rightarrow \mathfrak{R}$ such that $F+G$ is even function,

Since $(F+G)(x) = F(x) + G(x) = F(-x) + G(-x)$; why?

$$= (F + G)(-x) \forall x \in \mathfrak{R}$$

Hence, W is closed under the induced vector addition.

III) Let $a \in \mathfrak{R}$ and $f \in W$ then

$af : \mathfrak{R} \rightarrow \mathfrak{R}$ And $(af)(x) = af(x)$; Why?

$$= af(-x); \text{ Why?}$$

$$= (af)(-x) \forall x \in \mathfrak{R}$$

$$\Rightarrow af \in W$$

$\Rightarrow W$ is closed under the induced scalar multiplication.

Therefore, by I), II), III) and test for subspace theorem we have W is a subspace of V over $\mathfrak{R}$.

Example 2: Show that the set of first quadrant vector $W = \{(x, y) : x \geq 0 \text{ and } y \geq 0\}$, with the standard operations, is not a subspace of $\mathfrak{R}^2$ over $\mathfrak{R}$.

Solution: To see this, we note that $(1, 2)$ is in W , but the scalar multiple $(-1)(1, 2) = (-1, -2)$ is not in W . Therefore, W is not closed under scalar multiplication and hence W is not a subspace of $\mathfrak{R}^2$ over $\mathfrak{R}$. Infact, one can show that W is nonempty and closed under addition

Class Activity 2-3.2

1. Let k small letter be a whole number and P_k be the set of all polynomials over $\mathfrak{R}$ of degree less than or equal to k then show that

a) P_k With the standard operations is a vector space over $\mathfrak{R}$.

b) P_j a subspace of $P_k \quad \forall j \leq k$. Thus we can write $P_0 \subset P_1 \subset P_2 \subset P_3 \subset \dots \subset P_n$ and this is sequence of subspaces that are nested in each other

2. **(Subspaces of functions)**

Let V_5 be the vector space of all real valued functions defined on $[0, 2]$ and let

V_1, V_2, V_3 and V_4 be defined as:

V_1 = The set of all polynomial functions over $\mathfrak{R}$

V_2 = The set of all real valued functions that is differentiable on $[0, 2]$

V_3 = The set of all real valued function that is continuous on $[0, 2]$

V_4 = The set of all real valued functions that are integrable on $[0, 2]$ then show that V_i is a subspace of $V_j \quad \forall i \leq j$ and $V_1 \subset V_2 \subset V_3 \subset V_4 \subset V_5$ (nesting of subspaces)

3. **(Subspace of $\mathfrak{R}^2$)**

Prove that if W is subset of $\mathfrak{R}^2$, then it is a subspace of $\mathfrak{R}^2$ over $\mathfrak{R}$ if and only if one of the following is true

a) W consists of the single point $(0, 0)$ That is, $W = \{(0, 0)\}$

b) W consists of all points on a line that passes through the origin. That is,

$$W = \{(x, y) : y = 2x\}$$

d) W consists of all of $\mathbb{R}^2$. That is $W = \mathbb{R}^2$

4. (Subspace of $\mathbb{R}^3$)

Show that a subset W of $\mathbb{R}^3$ is subspace of $\mathbb{R}^3$ over $\mathbb{R}$ (with the standard operations) if and only if it has one of the following forms.

a) W consists of the single point $(0,0,0)$; $W = \{(0,0,0)\}$

b) W Consists of all points on a line that passes through the origin.

c) W consists of all points on a plane that passes through the origin

d) W Consists of $\mathbb{R}^3$

Theorem 2-3.1.2: The intersection of two subspaces is also a subspace.

That is, U and W are both subspace of a vector space V over a field F then the intersection of U and W , denoted by $U \cap W$, is also a subspace of V over F .

Proof: I) Since U and W are both subspace of V , we know that both contain the zero vector. Thus, $0 \in U \cap W$ and hence $U \cap W$ is nonempty.

Also, since $U \cap W \subseteq U$ and $U \subseteq V$ we have $U \cap W \subseteq V$

$$\Rightarrow U \cap W \subseteq V$$

Therefore, $\emptyset \neq U \cap W \subseteq V$

II) Let $v_1, v_2 \in U \cap W$ then $v_1, v_2 \in U$ and $v_1, v_2 \in W$

$$\Rightarrow v_1 + v_2 \in U \text{ And } v_1 + v_2 \in W; \text{ since both } U \text{ and } W \text{ are subspaces of } V$$

$$\Rightarrow v_1 + v_2 \in U \cap W$$

$$\Rightarrow U \cap W \text{ is closed under addition.}$$

III) Let $b \in F$ and $v \in U \cap W$ then $v \in U$ and $v \in W$ thus, $bv \in U$ and $bv \in W$; since both U and W are subspaces of V over F .

$$\Rightarrow bv \in U \cap W$$

Therefore, $bv \in U \cap W \quad \forall b \in F \text{ and } \forall v \in U \cap W$

$$\Rightarrow U \cap W \text{ is closed under scalar multiplication.}$$

Hence the proof



Remark 2-4:

1. Theorem 2.4 tells us that the intersection of two subspaces of a vector space V is itself a subspace of V but one can show that the union of two subspaces of a vector space V is not (in general) a subspace of V .
2. Let U, W and V be vector spaces over a field F such that U is a subspace of W and W is a subspace of V , then one can show that U is also a subspace of V .

😊 **Test yourself!**

1. Determine whether the set W is a subspace of $\mathbb{R}^3$ over $\mathbb{R}$ with standard operations. Justify your answer
 - a) $W = \{(a, b, 2a - 3b) : a, b \in \mathbb{R}\}$
 - b) $W = \{(r, r - t, t) : r, t \in \mathbb{R}\}$
 - c) $W = \{(x, y, xy) : x, y \in \mathbb{R}\}$
 - d) $W = \{(x, y, 3) : x, y \in \mathbb{R}\}$
2. Which of the following subset of $C(-\infty, \infty)$ are subspaces of $C(-\infty, \infty)$ over $\mathbb{R}$?
 - a) The set of all real valued nonnegative functions: $f(x) \geq 0 \forall x \in \mathbb{R}$
 - b) The set of all real valued even functions: $f(x) = f(-x) \forall x \in \mathbb{R}$
 - c) The set of all real valued odd functions: $f(-x) = -f(x) \forall x \in \mathbb{R}$
 - d) The set of all real valued constant functions: $f(x) = c \forall x \in \mathbb{R}$
 - e) The set of all functions such that: $f(0) = 0; 0 \in \mathbb{R}$
 - f) The set of all functions such that: $f(0) = 1; 0, 1 \in \mathbb{R}$
3. Prove that a non empty set W is a subspace of a vector space V over field F if and only if $ax + by$ is an element of $W \forall x, y \in W$ and $\forall a, b \in F$.
4. Give an example showing that the union of two subspaces of a vector space V over a field F is not necessarily a subspace of V over F .
5. Let W be a subspace of a vector space V over a field F then prove that zero vector O in V is also the zero vector in W .
6. Let V be a vector space over a field F and $x, y, z \in V$ then show that the set of all linear combinations of x, y and z . $W = \{ax + by + cz : a, b, c \in F\}$ is a subspace of V over F . This subspace is called the span of $\{x, y, z\}$
7. Let $W_1, W_2, \dots, W_n$ be subspaces of a vector space V over a field F then show that

a) $\cap_{i=1}^n w_i$ is also a subspace of V over F .

b) $\sum_{i=1}^n w_i = w_1 + w_2 + \dots + w_n = \{x_1 + x_2 + \dots + x_n : x_i \in w_i\}$ is also a subspace of V over F .

8. Let V be any vector space. Describe the “smallest” and “largest” subspaces of

2.3.2. Spanning sets and Linear Independence

In this section we are going to develop procedures for representing each vector in vector space as a linear combination of a selected number of vectors in the space and determine a given subset of a vector space is linearly independent or not.

Objectives:

At the end of this section each student should be able to:

- Write vector space as a linear combinations of another vector space.
- Define a spanning set or generators of a vector space V .
- Illustrate whether a given subset of a vector space is linearly independent or not.
- State and prove the spanning set theorem.

Spanning Sets

Definition 2.3.2.1: Definition of linear combination of vectors

Let V be a vector space over a field F and $v, x_1, x_2, \dots, x_n$ be vectors in V then we say that v is a linear combination of $x_1, x_2, \dots$, and x_n if v can be expressed as

$$v = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n; \text{ Where } \beta_i \in F, 1 \leq i \leq n$$

Example 1: Example of linear combinations

Let $v_1 = (1,3,1), v_2 = (0,1,2)$ and $v_3 = (1,0,-5)$ be vectors in $\mathbb{R}^3$ then since

$v_1 = 3v_2 + v_3 = 3(0,1,2) + (1,0,-5) = (1,3,1)$. We have v_1 is a linear combination of v_2 and v_3 .

Example 2: Finding a linear combination

Let $v = (1,1,1), x_1 = (1,2,3), x_2 = (0,1,2)$ and $x_3 = (-1,0,1)$ be vectors in $\mathbb{R}^3$ then write a linear combination of x_1, x_2 and x_3 .

Solution: We need to find scalars β_1, β_2 and β_3 such that $v = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

$$\text{But } v = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$\begin{aligned} \Rightarrow (1,1,1) &= \beta_1(1,2,3) + \beta_2(0,1,2) + \beta_3(-1,0,1) \\ &= (\beta_1 - \beta_3, 2\beta_1 + \beta_2, 3\beta_1 + 2\beta_2 + \beta_3) \end{aligned}$$

Then, by equating the corresponding components, we arrive at the following system of linear equations.

$$\begin{cases} \beta_1 - \beta_3 = 1 \\ 2\beta_1 + \beta_2 = 1 \\ 3\beta_1 + 2\beta_2 + \beta_3 = 1 \end{cases}$$

$$\Rightarrow \beta_1 = 1 + r, \beta_2 = -1 - 2r, \beta_3 = r \text{ This is an infinite number of solutions, by}$$

Gauss-Jordan elimination method

For instance, to obtain one solution, we could let $r = 1$, then $\beta_1 = 2, \beta_2 = -3$ and $\beta_3 = 1$

Hence; $v = 2x_1 - 3x_2 + x_3$. other choices for r would yield other ways to write v as a linear combination of x_1, x_2 and x_3 .

Definition 2.3.2.2: Definition of spanning set of a vector space

Let $S = \{x_1, x_2, \dots, x_n\}$ be a subset of a vector space V over a field F . The set S is called a spanning set (or a generator set) of V if every vector in V can be expressed as a linear combination of vectors in S . That is, if $\forall v \in V \quad \exists \beta_1, \beta_2, \dots, \beta_n \in F$ such that $v = \sum_{i=1}^n \beta_i x_i$. in such cases we say that S spans (or generates) V over F .

Example 3: Examples of spanning sets

a) Let $x_1 = (1,0,0), x_2 = (0,1,0)$ and $x_3 = (0,0,1)$ be vectors in $\mathbb{R}^3$ then since

$$\forall v = (x, y, z) \in \mathbb{R}^3, \quad v = (x, y, z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$$

$$= xx_1 + yx_2 + zx_3 \text{ We have } S = \{x_1, x_2, x_3\} \text{ spans or generates } \mathbb{R}^3$$

b) The set $S = \{1, x, x^2\}$ generates of P_2 , since any polynomial $p(x) = a + bx + cx^2$

$$\text{in } P_2 \text{ can be written as } p(x) = a(1) + b(x) + c(x^2) = a + bx + cx^2$$

☞ **Note:** The spanning sets given in example -3 are called the standard spanning sets of $\mathbb{R}^3$ and p_2 , respectively

Definition 2.3.2.3 Let V be a vector space over a field F and S be non empty subset of V then the set of all finite linear combination of elements of S is called span of S and span of S is denoted by $\text{span}(S)$. That is span

$$(s) = \left\{ \sum_{i=1}^n \beta_i x_i : \beta_i \in F \text{ and } x_i \in S \quad \forall_i \text{ such that } 1 \leq i \leq n \right\}$$

Theorem 2.3.2.1: $\text{Span}(S)$ is a subspace of V

Let V be a vector space over a field F and S is nonempty subset of V then $\text{span}(S)$ is a subspace of V .

Moreover, $\text{span}(s)$ is the smallest subspace of V that contains S , in the sense that every other subspace of V that contains S must contain $\text{span}(s)$

Claim-1 $\text{span}(S)$ is a subspace of V over F

Proof:

I. Let S be nonempty subset of V then $\exists x_1 \in s$ such that $0x_1 = 0 \in \text{span}(s)$

$$= \left\{ \sum_{i=1}^n \beta_i x_i : \beta_i \in F \text{ and } x_i \in s; 1 \leq i \leq n \right\}$$

But $0 \in \text{span}(S)$ implies $\text{span}(S) \neq \emptyset \quad \text{---} (\nabla)$

Also, let $x \in \text{span}(s)$ then $x = \sum_{i=1}^n \beta_i x_i$ for some $\beta_i \in F$ and $x_i \in S$ where $1 \leq i \leq n$

Now, $x = \sum_{i=1}^n \beta_i x_i$ with $\beta_i \in F$ and $x_i \in s$ for $1 \leq i \leq n$

$\Rightarrow x \in V$; Since $S \subseteq V$ and V is closed under addition and scalar multiplication

Therefore, $\forall x (x \in \text{span}(s) \Rightarrow x \in V)$

$$\Rightarrow \text{span}(s) \subseteq V \quad \text{---} (\nabla \nabla)$$

Thus, by (∇) and $(\nabla \nabla)$ we get $\emptyset \neq \text{span}(s) \subseteq V$.

II. Let $u, v \in \text{span}(s)$ then

$$u = \sum_{i=1}^m a_i x_i \quad \text{and, } v = \sum_{i=1}^n \beta_i y_i \quad \text{for some}$$

$$a_1, a_2, \dots, a_m, \beta_1, \beta_2, \dots, \beta_n \in F \text{ and}$$

$$x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n \in S$$

$$\text{Now, } u + v = \sum_{i=1}^m a_i x_i + \sum_{i=1}^n \beta_i y_i$$

$$= \sum_{j=1}^{m+n} x_j z_j \text{ with } x_j = \begin{cases} a_j, & 1 \leq j \leq m \\ \beta_j, & m+1 \leq j \leq n \end{cases} \text{ and } z_j = \begin{cases} x_j; & 1 \leq j \leq m \\ y_j; & m+1 \leq j \leq n \end{cases}$$

$$\Rightarrow u + v = \sum_{j=1}^{m+n} \ell_j z_j \text{ with } \ell_j \in F \text{ and } z_j \in S \forall_j (1 \leq j \leq m+n) \text{ (why?)}$$

$\Rightarrow u + v$ is a finite linear combination of elements of S

$$\Rightarrow u + v \in \text{span}(S)$$

$\Rightarrow \text{Span}(S)$ is closed under addition

III. Let $a \in F$ and $u \in \text{span}(S)$ then $u = \sum_{i=1}^n \beta_i x_i$ for some $\beta_i \in F$ and $x_i \in S (1 \leq i \leq n)$

$$\text{and hence } au = a \sum_{i=1}^n \beta_i x_i$$

$$= \sum_{i=1}^n (\beta_i x_i), \quad \text{Why?}$$

$$= \sum_{i=1}^n (a\beta_i) x_i, \quad \text{Why?}$$

$$= \sum_{i=1}^n (\alpha_i x_i), \text{ Where } \alpha_i = a\beta_i \in F \forall_i (1 \leq i \leq n) \text{ and } x_i \in S$$

$$\Rightarrow au = \sum \alpha_i x_i \text{ with } \alpha_i \in F \text{ and } x_i \in S \forall_i$$

$\Rightarrow au$ is a finite linear combination of elements of S

$$\Rightarrow au \in \text{span}(S) \forall a \in F \text{ and } \forall u \in \text{span}(S)$$

$\Rightarrow \text{Span}(S)$ is closed under scalar multiplication. Therefore, by i), ii) and iii) we get

$\text{span}(S)$ is a subspace of V over F

Claim -2 $\text{Span}(S)$ is the smallest subspace containing S

i) Let then spans (S) and hence $\text{span}(S)$.

ii) By claim 1 we have shown that $\text{span}(S)$ is a subspace of V .

iii) Let W be a subspace of V containing S then W is closed under addition and scalar multiplication --- (*)

Now, let $x \in \text{span}(S)$ then $x = \sum_{i=1}^n \beta_i x_i$ with $\beta_i \in F$ and

$$x_i \in S \forall_i (1 \leq i \leq n)$$

$$\Rightarrow x = \sum_{i=1}^n \beta_i x_i \text{ Where } \beta_i \in F \text{ and } x_i \in W \forall_i; \text{ as } S \subseteq W$$

$$\Rightarrow x = \sum_{i=1}^n \beta_i x_i \in W; \text{ by } (*)$$

$$\Rightarrow x \in W$$

Thus every element in $\text{span}(S)$ is also in W and hence $\text{span}(S) \subseteq W$

Therefore, by i, ii, and iii we have $\text{span}(S)$ is the smallest subspace of V containing S .

2.4. Linear Dependence and Independence of vectors

Definition 2.4.1: Let V be a vector space over a field F and $S = \{x_1, x_2, \dots, x_n\}$ be subset of V then S is called linearly independent if the vector equation $\sum_{i=1}^n \beta_i x_i = 0$ has only the trivial solution. If there are also non-trivial solutions, then S is called linearly dependent. That is,

a) If $\sum_{i=1}^n \beta_i x_i = 0 \Rightarrow \beta_i = 0 \forall_i (1 \leq i \leq n)$ then S is a linearly independent set.

b) If $\exists \beta_1, \beta_2, \dots, \beta_n \in F$ not all zero such that $\sum_{i=1}^n \beta_i x_i = 0$ then S is linearly dependent set

Class Activity 2-4:

Testing for Linear Independence

1. Which of the following sets are linearly independent in $\mathbb{R}^3$ and which are not?

a) $\{(3,1,-4), (2,2,-3), (0,-4,1)\}$

b) $\{(0,2,-4), (1,-2,-1), (1,-4,3)\}$

c) $\{(3,1,-4), (2,2,-3), (0,-4,1)\}$

d) $\{(1,2,3), (0,1,2), (-2,0,1)\}$

e) $\{(1,0,0), (0,4,0), (0,0,-6), (1,5,-3)\}$

2. Determine which of the sets in P_2 are linearly independent

a) $\{2 - x, 2x - x^2, 6 - 5x + x^2\}$

b) $\{x^2 - 1, 2x + 5\}$

c) $\{x^2 + 3x + 1, 2x^2 + x - 1, 4x\}$

d) $\{1 + x - 2x^2, 2 + 5x - x^2, x + x^2\}$

e) $\{1, x, x^2\}$



Remark 2-5: Testing for Linear Independence and Linear Dependence

A. Let $S = \{u_1, u_2, \dots, u_n\}$ be a set of vectors in a vector space V over a field F . To determine whether S is linearly independent or linearly dependent, perform the following steps.

➤ Step-1: From the vector equation $\sum_{i=1}^n \beta_i u_i = O$, write a homogenous system

of linear equations in the variables $\beta_1, \beta_2, \dots, \beta_n$.

➤ Step-2: Use Gauss ian elimination method to solve the system for

$\beta_1, \beta_2, \dots$, and β_n .

➤ Step-3: If the system has only the trivial solution,

$\beta_1 = 0 = \beta_2 = \dots = \beta_n$, then the set S is linearly independent. If the system also has non trivial solutions, then S is linearly dependent.

B. Every set of vectors in a vector space is either linearly independent or linearly dependent. Thus a single test suffices.

C. We say that an infinite set S is linearly independent if every finite subset of S is a linearly independent set.

Theorem 2-4: A Property of Linearly Dependent Sets

Let $S = \{x_1, x_2, \dots, x_n\}, n \geq 2$ be a subset of a vector space V over a field F then S is Linearly dependent if and only if at least one of the vectors x_i can be written as a linear combination of the other vectors in S .

Proof: ($\Rightarrow$) Suppose $S = \{x_1, x_2, \dots, x_n\}$ is linearly dependent set. Then there exist

scalars (elements of F) $\beta_1, \beta_2, \dots, \beta_n$ not all zero such that $\sum_{i=1}^n \beta_i x_i = 0$

Because at least one of the coefficients must be non zero, no generality is lost

by assuming that $\beta_j \neq 0$ ($1 \leq j \leq n$) Now, $\sum_{i=1}^n \beta_i x_i = 0$

$$\Rightarrow \beta_1 x_1 + \dots + \beta_j x_j + \dots + \beta_n x_n = 0$$

$$\Rightarrow \beta_j x_j = - \sum_{\substack{i=1 \\ i \neq j}}^n \beta_i x_i$$

$$\Rightarrow x_j = -\beta_j^{-1} \sum_{i \neq j}^n \beta_i x_i = 0, \text{ Why?}$$

$$\Rightarrow x_j = - \sum_{i \neq j}^n (\beta_j^{-1} \beta_i) x_i, \text{ where } \partial_i = \beta_j^{-1} \beta_i \in F$$

$$\Rightarrow x_j = \partial_1 x_1 + \partial_2 x_2 + \dots + \partial_{j-1} x_{j-1} + \partial_{j+1} x_{j+1} + \dots + \partial_n x_n$$

$\Rightarrow x_j$ is a linear combination of the other vectors.

($\Leftarrow$) Suppose at least one of the vectors x_j is a linear combination of the other vectors.

Now, without loss of generality assume x_1 is a linear combination of the other vectors.

$$\text{Then } x_1 = \sum_{i=2}^n \beta_i x_i$$

$$\Rightarrow \sum_{i=2}^n \beta_i x_i - x_1 = 0$$

$$\Rightarrow (-1) x_1 + \beta_2 x_2 + \dots + \beta_n x_n = 0$$

$$\Rightarrow S = \{x_1, x_2, \dots, x_n\} \text{ is linearly dependent, since at least one coefficient, } -1$$

is non-zero. Hence the result.

Example-4

Let $S = \{1+x-2x^2, 2+5x-x^2, x+x^2\}$ then show that S is linearly dependent subset of P_2

Solution:

Let $f_1 = 1+x-2x^2$, $f_2 = 2+5x-x^2$ and $f_3 = x+x^2$ then one can observe that

$$f_2 = 2f_1 + 3f_3 \text{ And hence } f_2 \text{ is a linear combination of } f_1 \text{ and } f_3$$

Thus, by theorem 2.6 we have S is linearly dependent.

Example 5:

Let V be a vector space over a field F and $u, v \in V$ then show that u and v are linearly dependent if and only if one is a scalar multiple of the other.

Solution:

It is an immediate consequence of theorem 2.6. It is a simple test for determining whether two vectors are linearly dependent.

☺ **Test yourself!**

1. Determine which vectors u, v , and w can be written as linear combinations of the vectors in S .
 - a) $S = \{(1, 2, -2), (2, -1, 1)\}$, $u = (1, 17, -17)$, $v = (3, -\frac{2}{3}, \frac{2}{3})$ and $w = (8, -4, 3)$
 - b) $S = \{(2, 0, 7), (2, 4, 5), (2, -12, 13)\}$, $u = (4, -20, 24)$, $v = (-1, 0, 0)$ and $w = (6, 24, 9)$
2. Determine whether the given set S spans $\mathbb{R}^3$. If the set does not span $\mathbb{R}^3$, describe the subspace that it does span.
 - a) $S = \{(1, 0, 1), (1, 1, 0), (0, 1, 1)\}$
 - b) $S = \{(1, 0, 3), (2, 0, -1), (4, 0, 5), (2, 0, 6)\}$
 - c) $S = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$
 - d) $S = \{(1, 2, 3), (0, 1, 2), (-1, 0, 1)\}$
3. Determine whether the set S is linearly independent or dependent
 - a) $S = \{(0, 0), (1, -1)\}$
 - b) $S = \{(4, -4, -2), (-2, 2, 1)\}$
 - c) $S = \{(1, 0, 0), (0, 4, 0), (0, 0, -6), (1, 5, -3)\}$
 - d) $S = \{(1, 2, 3, 4), (1, 0, 1, 2), (1, 4, 5, 6)\}$
4. For which values of x are the following sets linearly independent?
 - a) $S = \{(x, 1, 1), (1, x, 1), (1, 1, x)\}$
 - b) $S = \{(x, 1, 1), (1, 0, 1), (1, 1, 3x)\}$
5. For which values of r are the following sets linearly independent?
 - a) $S = \{(r, 0, 0), (0, 1, 0), (0, 0, 1)\}$
 - b) $S = \{(r, r, r), (r, 1, 0), (r, 0, 1)\}$

6. Determine whether the set $S = \{y^2 - 2y, y^3 + 8, y^3 - y^2, y^3 - y^2, y^2 - 4\}$ Spans or generates p_3
7. Prove the following
 - a) If S_1 is a sub set of S_2 and S_1 is linearly dependent then so is S_2 .
 - b) Any set of vectors containing the zero vector is linearly dependent.
 - c) A non empty sub set of linearly independent vectors is linearly independent.
 - d) Given that $\{x_1, x_2, \dots, x_n\}$ is linearly independent set $\{x_1, x_2, \dots, x_n, x\}$ of vectors but the set is linearly dependent, then x is a linear combination of the x_i 's.
 - e) If $S = \{u, v\}$ is a linearly independent set, then the set $\{u + v, u - v\}$ is linearly independent.
8. The set $\{(1, 2, 3), (1, 0, -2), (-1, 0, 2)\}$ is linearly dependent, but $(1, 2, 3)$ can not be written as a linear combination of $(1, 0, -2)$ & $(-1, 0, 2)$ and why does this not contradict theorem 2.6.
9. Under what conditions will a set consisting of a single vector be linearly independent?

2.5. Bases and Dimension of a Vector space

Objectives

At the end of this section students are expected

- Define a basis of a vector space.
- Define the dimension of a vector space.
- Find a base of a subspace and state its dimension.

2.5.1. Basis of a vector space

Definition 2.5.1: Definition of Basis

A set of vectors $S = \{x_1, x_2, \dots, x_n\}$ in a vector space V over a field F is called basis for V if i- S spans V and ii- S is linearly independent.



Remark 2.5:

1. The above definition tells us that a basis has two features. A basis S must have enough vectors to span V , but not so many vectors that one of them could be written as a linear combination of the other vectors in S .
2. An infinite set S is said to be a basis for V if
 - i- S spans V and ii- S is linearly independent. That is, if every element of V is a finite linear combination of elements of S and every finite subset of S is linearly independent. Thus a basis of a vector space V may consists of infinite number of vectors.

Class Activity 2.5.1:

1. Show that $s\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is a basis for R^3 . S is called the standard basis for R^3 .
2. Let $e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = \{(0, 0, \dots, 1)\}$ then show that $S = \{e_1, e_2, \dots, e_n\}$ is a basis for R^n S is called a standard basis for R^n .
3. Show that
 - a) $S = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$ is a non standard basis for R^3
 - b) $S = \{(1, 1), (1, -1)\}$ is a non standard basis for R^2
 - c) $S = \{1, x, x^2, \dots, x^n\}$ is a standard basis for P^n

Theorem 2.5.1: Uniqueness of Basis Representation

Let V be a vector space over field F and $s = \{x_1, x_2, \dots, x_n\}$ be a basis for V then every vector in V can be written in one and only one way as a linear combination of vectors in S .

Proof

i) **Existence:**

Let $v \in V$ then $v = \sum_{i=1}^n a_i x_i$ where $a_i \in F$ and $x_i \in F$ because of S generates V

$\forall_i = 1, 2, \dots, n$. This

ii) **Uniqueness:**

To prove that a given vector can be represented in only one way, we assume that V has another representation

$$v = \sum_{i=1}^n b_i x_i \text{ where } b_i \in F \text{ and } x_i \in S \quad \forall i = 1, 2, \dots, n$$

$$\text{Then we get } V = \sum_{i=1}^n a_i x_i = \sum_{i=1}^n b_i x_i$$

$$\Rightarrow \sum_{i=1}^n a_i x_i - \sum_{i=1}^n b_i x_i = 0$$

$$\Rightarrow \sum_{i=1}^n (a_i - b_i) x_i = 0$$

$$\Rightarrow a_i - b_i = 0 \quad \forall i = 1, 2, \dots, n. \text{ This is because of the fact that } S \text{ is}$$

Linearly independent

Thus, $a_i = b_i = 0 \quad \forall i = 1, 2, \dots, n$ and hence V has only one representation for the given basis S .

Example 1: Uniqueness of Basis Representation

Let $U = (u_1, u_2, u_3)$ be any vector in R^3 then show that the equation

$u = a_1 v_1 + a_2 v_2 + a_3 v_3$ has a unique solution for the set $S = \{v_1, v_2, v_3\}$ where

$$v_1 = (1, 2, 3), v_2 = (0, 1, 2), \& v_3 = (-2, 0, 1)$$

Solution:

Claim 1 S is a basis for R^3

$$\text{i) Suppose } \sum_{i=1}^3 a_i v_i = 0$$

$$\Rightarrow a_1 v_1 + a_2 v_2 + a_3 v_3 = 0$$

$$\Rightarrow a_1(1, 2, 3) + a_2(0, 1, 2) + a_3(-2, 0, 1) = (0, 0, 0)$$

$$\Rightarrow \begin{cases} a_1 - 2a_3 = 0 \\ 2a_1 + a_2 = 0 \\ 3a_1 + 2a_2 + a_3 = 0 \end{cases}$$

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

$$\Rightarrow S \text{ is linearly independent set}$$

$$\text{ii) Let } v = (x, y, z) \in R^3 \text{ and suppose that}$$

$$\begin{aligned}
 a_1 v_1 + a_2 v_2 + a_3 v_3 &= \sum_{i=1}^n a_i v_i \\
 \Rightarrow (x, y, z) &= a_1(1, 2, 3) + a_2(0, 1, 2) + a_3(-2, 0, 1) \\
 \Rightarrow (x, y, z) &= (a_1, -2a_3, 2a_1 + a_2, 3a_1 + 2a_2 + a_3) \\
 \Rightarrow \begin{cases} a_1 - 2a_3 = x \\ 2a_1 + a_2 = y \\ 3a_1 + 2a_2 + a_3 = z \end{cases} \\
 \Rightarrow a_1 &= -x + 4y - 2z \\
 a_2 &= 2x - 7y + 4z \\
 a_3 &= -x + 2y - z
 \end{aligned}$$

Thus, $\forall v = (x, y, z) \in \mathbb{R}^3$

$$\exists a_1 = -x + 4y - 2z, a_2 = 2x - 7y + 4z \text{ and } a_3 = -x + 2y - z \in \mathbb{R}$$

Such that $v = a_1 v_1 + a_2 v_2 + a_3 v_3$ and hence s generates $\mathbb{R}^3$

There fore, by I) and II) we get S is a basis for $\mathbb{R}^3$

Claim- 2

The vector equation $u = a_1 v_1 + a_2 v_2 + a_3 v_3$ has a unique solution.

Now, Since S is a basis for $\mathbb{R}^3$ we've by theorem 2.7, uniqueness of basis representation, and the equation $u = a_1 v_1 + a_2 v_2 + a_3 v_3$ has unique solution. Infact, by the result of claim-1 we get

$$\begin{aligned}
 a_1 &= -u_1 + 4u_2 - 2u_3 \\
 a_2 &= 2u_1 - 7u_2 + 4u_3 \text{ and } a_3 = -u_1 + 2u_2 - u_3
 \end{aligned}$$

For instance, the vector $u = (1, 0, 0)$ can be represented uniquely as a linear combination of v_1, v_2 and v_3 as $(1, 0, 0) = -v_1 + 2v_2 - v_3$,

Definition 2.5.2: Coordinate Representation relative to a Basis

Let V be a vector space over a field F with basis $\beta = \{x_1, x_2, \dots, x_n\}$ and $x \in V$. If $x = \sum_{i=1}^n a_i x_i$ where $a_i \in F \forall i = 1, 2, \dots, n$ then the scalars $a_1, a_2, \dots, a_n$ are called

the coordinates of x with respect to the basis β . The coordinate vector of x relative to β is the vector in $\mathbb{R}^n$ denoted by $(x)_\beta = (a_1, a_2, \dots, a_n)$

Example 2:

In the example-1 above we have

$u = (1, 0, 0) = -v_1 + 2v_3 - v_3$ and hence $(u)_\beta = (-1, 2, -1)$ is the coordinate of U with respect to the basis $\beta = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$.

 **Note that:** The order of the vectors in the basis is important. In the context of coordinate representation, $\beta = \{x_1, x_2, \dots, x_n\}$ is sometimes called an ordered basis.

Theorem 2.5.2 Basis and Linear Dependence

If $S = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space V over a field F , then every set containing more than n vectors in V is linearly dependent. That is, the maximum number of linearly independent vectors in V is n .

Proof:

Let $S_1 = \{u_1, u_2, \dots, u_m\}$ be any set of m vectors in V , where $m > n$. To show that S_1 is linearly dependent, we need to find scalars $k_1, k_2, \dots, k_m$ not all zero such that $k_1 u_1 + k_2 u_2 + \dots + k_m u_m = 0 \dots (1)$

Since S is a basis for V , it follows that each u_i is a linear combination of vectors in S ,

and hence $u_j = \sum_{i=1}^n c_{ij} v_i$ where $c_{ij} \in F \forall i = 1, 2, \dots, n$ & $\forall j = 1, 2, \dots, m$.

Substituting each of these representations of u_i in to (1) and regrouping terms produces

$$\sum_{i=1}^n d_i v_i = 0 \text{ where } d_i = \sum_{j=1}^m c_{ij} k_j \quad \text{However, since the } v_i \text{'s form a linearly}$$

independent set, we conclude that each $d_i = 0$. Thus the following system of equations

$$\text{is obtained. } c_{11}k_1 + c_{12}k_2 + \dots + c_{1m}k_m = 0$$

$$c_{21}k_1 + c_{22}k_2 + \dots + c_{2m}k_m = 0$$

$$C_{21}k_1 + c_{22}k_2 + \dots + c_{2m}k_m = 0$$

$$C_{n1}k_1 + c_{n2}k_2 + \dots + c_{nm}k_m = 0$$

But this homogeneous system has fewer equations than variables $k_1, k_2, \dots, k_m$ and hence it must have nontrivial solutions. Consequently, S_1 is linearly dependent. Hence the proof

Example-3: Linearly Dependent sets in $\mathbb{R}^3$, P_3 & $\mathbb{R}^n$

a) Since $\mathbb{R}^3$ has a basis consisting of three vectors, the set

$\{(1,2,-1), (1,1,0), (2,3,0), (5,9,-1)\}$ must be linearly dependent.

b) Since P_3 has a basis consisting of four vectors, the set

$\{1, 1+x, 1-x, 1+x+x^2, 1-x+x^2\}$ must be linearly dependent.

c) Since, $\mathbb{R}^n$ has the standard basis

Containing of n vectors, it follows from theorem 2-8 that every set of vectors in $\mathbb{R}^n$ containing more than n vectors must be linearly dependent.

Theorem 2-5.3: Number of Vectors in a Basis

If a vector space V over a field F has one basis with n vectors, then every basis for V has n vectors.

Proof:

Let $S_1 = \{x_1, x_2, \dots, x_n\}$ and $S_2 = \{y_1, y_2, \dots, y_m\}$ be basis for V over F then since

S_1 is a basis and S_2 is linearly independent we get $m \leq n$ --- (*) Also, since S_2

is a basis for V and S_1 is linearly independent we get $n \leq m$ --- (**)

Now, Combining (*) and (**) we have $m \leq n$ and $n \leq m$ which intern implies $m=n$.

Example-4

a) Since the standard basis for $\mathbb{R}^3$, $\beta = \{(1,0,0), (0,1,0), (0,0,1)\} = \{i, j, k\}$ has three

vectors we have the set $\{(3,2,1), (7,-1,4)\}$ is not a basis for $\mathbb{R}^3$.

b) Since the standard basis for P_3 , $\beta = \{1, x, x^2, x^3\}$ has four elements we have the set

$\{x+2, x^2, x^3-1, 3x+1, x^2-2x+3\}$ is not a basis for P_3 .

2.5.2 The Dimension of a Vector Space

Definition 2.5.3: Definition of dimension of a vector space

If a vector space V over a field F has a basis consisting of n vectors, then the number n is called the dimension of V , denoted by $\dim(V) = n$. If V consists of the zero vector alone, then the dimension of V is defined as zero. In general, let V be a vector space, over a field F . Suppose there is some whole number n such that V contains a set of n vectors that are linearly independent, while every set of $(n+1)$ vectors in V is linearly dependent then V is called finite dimensional and n is called the dimension of V .

Example -5

- a) Since $\beta = \{e_1, e_2, e_3, \dots, e_n\}$ form a basis for R^n over R we have $\dim(R^n) = n$
- b) Since $\beta = \{1, x, x^2, \dots, x^n\}$ form a basis for P_n over R we have $\dim(P_n) = n+1$

Class Activity 2.5.2

Finding the dimension of a subspace

1. Let W be a subspace of a vector space over a field F with dimension n then show that $\dim W$ is less than or equal to n
2. Determine the dimensions of the following subspaces of R^3 over R
 - a) $W = \{(d, c-d, c) : c, d \in R\}$
 - b) $W = \{(2b, b, 0) : b \in R\}$
3. Find the dimension of the subspace W of R^4 over R spanned by $S = \{(-1, 2, 5, 0), (3, 0, 1, -2), (-5, 4, 9, 2)\}$

Theorem 2.10: Basis Tests in an n-Dimensional space

Let V be a vector space over a field F with dimension n

- a) If $\beta = \{x_1, x_2, \dots, x_n\}$ is a linearly independent set of vectors in V , then S is a basis for V over F .
- b) If $\beta = \{v_1, v_2, \dots, v_n\}$ spans V , then S is a basis for V over F .

Proof: Left as an exercise

☺ **Test yourself!**

1. Determine whether β is a basis for the indicated vector space
 - a) $\beta = \{(-1, 2, 0, 0), (2, 0, -1, 0), (3, 0, 0, 4), (0, 0, 5, 0)\}$ for $\mathbb{R}^4$ over $\mathbb{R}$
 - b) $\beta = \{(1, i)\}$ for $\mathbb{C}$ over $\mathbb{R}$
 - c) $\beta = \{(1, 5, 3), (0, 1, 2), (0, 0, 6)\}$ for $\mathbb{R}^3$ over $\mathbb{R}$
2. Find all subsets of the set
 - a) $S = \{(1, 0), (0, 1), (1, 1)\}$ that form a basis for $\mathbb{R}^2$ over $\mathbb{R}$
 - b) $S = \{(1, 3, -2), (-4, 1, 1), (-2, 7, -3), (2, 1, 1)\}$ that form a basis for $\mathbb{R}^3$ over $\mathbb{R}$
3. Find a basis for
 - a) $\mathbb{R}^2$ over $\mathbb{R}$ that includes the vector $(1, 1)$
 - b) $\mathbb{R}^3$ over $\mathbb{R}$ that includes the set $S = \{(1, 0, 2), (0, 1, 1)\}$
4. Let $w_1 = \{(2t, t) : t \in \mathbb{R}\}$

$$w_2 = \{(0, t) : t \in \mathbb{R}\}$$

$$w_3 = \{(2r, r, -r) : r \in \mathbb{R}\}$$

$$w_4 = \{(2s - t, s, t) : s, t \in \mathbb{R}\}$$
 - a) Show that w_1 and w_2 are subspaces of $\mathbb{R}^2$ over $\mathbb{R}$
 - b) Show that w_3 and w_4 are subspaces of $\mathbb{R}^3$ over $\mathbb{R}$
 - c) Give a geometric description for each W_i ($1 \leq i \leq 4$)
 - d) Find a basis and dimension of W_i ($1 \leq i \leq 4$)
5. Determine whether the given statement is true or false. Give a reason for your answer.
 - a) If $\dim(V) = n$ then there exists a set of $n-1$ vectors in V that will span V
 - b) If $\dim(V) = n$, then there exists a set of $n+1$ vectors in V that will span V
 - c) If $\dim(V) = n$, then any set of $(n+1)$ vectors in V must be dependent.
 - d) If $\dim(V) = n$, then any set of $(n-1)$ vectors in V must be independent.
6. Prove that the vector space of all polynomials over $\mathbb{R}$ is infinite dimensional
7. Let $\{v_1, v_2, \dots, v_n\}$ be a basis for a vector space V over a field F

- a. Prove that $\{v_1, v_1 + v_2, v_1 + v_2 + v_3, \dots, v_1 + v_2 + \dots + v_n\}$ is also a basis for V over F
- b. Let $\{a_1, a_2, \dots, a_n\}$ be scalars, none of which is zero. Prove that $\{a_1 v_1, a_2 v_2, \dots, a_n v_n\}$ is also a basis for V over F .
- c. Suppose that a vector V has coordinate set $\{\beta_1, \beta_2, \dots, \beta_n\}$ relative to the original basis
8. Determine its coordinate set relative to
- I) the basis of part (a)
- II) The basis of part (b)
-

2.6 Direct sum and Direct Product of Subspaces

Objectives

At the end of this section students should be able to

- Define sums of subspaces of a vector space.
- Determine the smallest subspace of a vector space over a field.
- Show vector spaces over a field can be expressed as scalar multiplication defined component wise over a field.

Definition 2.6.1: Sums of subspaces of a vector space

Let $W_1, W_2, \dots, W_n$ be subspace of a vector space V over a field F then the set of all vectors of the form $v_1 + \dots + v_n$ where $v_i \in W_i$ is called the sum of the subspaces or their span, and is denoted by

$$W_1 + W_2 + \dots + W_n = \{v_1 + v_2 + \dots + v_n : v_i \in W_i \ \forall i = 1, 2, \dots, n\}$$

Theorem 2.6.1:

Let $W_1, W_2, \dots, W_n$ be subspaces of a vector space over a field F then

$W_1 + W_2 + W_3 + \dots + W_n$ is the smallest subspace of V containing

$W_1, W_2, \dots$ and W_n .

Proof:

I) Let $x \in W_i$ where i is a fixed element of $\{1, 2, \dots, n\}$ then

$$x = 0 + 0 + \dots + 0 + x + 0 + \dots + 0 \in W_1 + \dots + W_i + \dots + W_n$$

$$\Rightarrow x \in W_1 + \dots + W_i + \dots + W_n$$

$$\Rightarrow \forall x (x \in W_i \Rightarrow x \in W_1 + \dots + W_i + \dots + W_n)$$

$$\Rightarrow W_i \subseteq W_1 + W_2 + \dots + W_i + \dots + W_n \text{ for each } i \in \{1, 2, \dots, n\}$$

$$\Rightarrow \text{for each } i \in \{1, 2, \dots, n\} \text{ we have } W_i \subseteq W_1 + \dots + W_2 + \dots + W_n$$

$$\Rightarrow W_1 + W_2 + \dots + W_n \text{ contains } W_1, W_2, \dots, W_n \dots$$

II)

➤ Since $W_1, W_2, \dots, W_n$ are subspaces of V over F . We have let

$$0 \in W_i \subseteq V \quad \forall_i = 1, 2, \dots, n \quad \text{and hence}$$

$$0 = 0 + 0 + \dots + 0 \in W_1 + W_2 + \dots + W_n \subseteq V$$

$$\Rightarrow \phi \neq W_1 + W_2 + \dots + W_n \subseteq V$$

➤ $x, y \in W_1 + W_2 + \dots + W_n$ then $x = x_1 + x_2 + \dots + x_n$ and

$$y = y_1 + y_2 + \dots + y_n \text{ where } x_i, y_i \in W_i \quad \forall_i = 1, 2, \dots, n$$

$$\text{Thus } x + y = (x_1 + \dots + x_n) + (y_1 + y_2 + \dots + y_n)$$

$$= (x_1 + y_1) + (x_2 + y_2) + \dots + (x_n + y_n); \text{ why?}$$

$$\in W_1 + W_2 + \dots + W_n,$$

$$\text{since } x_i, y_i \in W_i \text{ and } W_i \text{ is a subspace of } V \quad \forall_i = 1, 2, \dots, n$$

$$\Rightarrow x + y \in W_1 + W_2 + \dots + W_n.$$

➤ Let $x \in W_1 + W_2 + \dots + W_n$ and $\beta \in F$ then

$$x = x_1 + \dots + x_n \text{ with } x_i \in W_i \quad \forall_i = 1, 2, \dots, n \text{ and hence } \beta x = \beta(x_1 + \dots + x_n)$$

$$= \beta x_1 + \beta x_2 + \dots + \beta x_n$$

$$\in W_1 + W_2 + \dots + W_n, \text{ Since } x_i \in W_i \text{ and } W_i \text{ is subspace of } V \text{ over } F$$

$$\forall_i = 1, 2, \dots, n.$$

$$\Rightarrow \beta x \in W_1 + W_2 + \dots + W_n$$

Thus, we have $W_1 + W_2 + \dots + W_n$ is a subspace of V over F .

III) Let W be a subspace of V containing $W_1, W_2, \dots, W_n$. Let $x \in W_1 + W_2 + W_n$

then $x = x_1 + \dots + x_n$ with $x_i \in W_i \quad \forall_i = 1, 2, \dots, n$

$\Rightarrow x = x_1 + \dots + x_n$ with $x_i \in W$; since $W_i \subseteq W \quad \forall_i = 1, 2, \dots, n$

$\Rightarrow x \in W$ as W is a subspace of V over F

Thus, $\forall x (x \in W_1 + W_2 + \dots + W_n \Rightarrow x \in W)$

$\Rightarrow W_1 + W_2 + \dots + W_n \subseteq W$

Hence, $W_1 + W_2 + \dots + W_n$ is the smallest subspace of V containing

$W_1, W_2, \dots, \text{and } W_n$, by I) II) III).

Definition 2.6.2: Direct sums of subspaces

Let $W_1, W_2, \dots, W_n$ be subspaces of a vector spaces V over a field F , such that

$\bigcap_{i=1}^n W_i = \{0\}$ then the direct sum of $W_1, W_2, \dots, W_n$, denoted by

$W_1 \oplus W_2 \oplus W_3 \oplus \dots \oplus W_n$ is the subspace $W_1 + W_2 + \dots + W_n$. That is, if

$W_i \cap \sum_{j \neq i} W_j = \{0\}$ then the subspace $W_1 + W_2 + \dots + W_n$ is called the direct sum of

$W_1, W_2, \dots, W_n$

and we write $W_1 \oplus W_2 \oplus W_3 \oplus \dots \oplus W_n$ instead of $W_1 + W_2 + \dots + W_n$.

Example-1

Let $V = \mathbb{R}^2, F = \mathbb{R}$

$M = \{(x, 0) : x \in \mathbb{R}\}$ and $N = \{(0, x) : x \in \mathbb{R}\}$ then find $M \oplus N$.

Solution:

One can observe that M and N are subspaces of V over $F = \mathbb{R}$.

Also, for $\forall v \in M \cap N$ we have $v \in M$ and $v \in N$

$\Rightarrow v = (x, 0)$ and $v = (0, y)$ for some $x, y \in \mathbb{R}$.

$\Rightarrow v = (x, 0) = (0, y)$

$\Rightarrow (x, 0) = (0, y)$

$\Rightarrow x = 0 = y$

$\Rightarrow v = (0, 0)$; the zero vector in $\mathbb{R}^2$.

Thus $M \cap N = \{(0,0)\} = \{0\}$

$$\begin{aligned} \text{Moreover, } M \oplus N &= M + N = \{m + n : m \in M \text{ and } n \in N\} \\ &= \{(x,0) + (0,y) : x, y \in R\} \\ &= \{(x+0, 0+y) : x, y \in R\} \\ &= \{(x,y) : x, y \in R\} \\ &= R^2 ; \text{why?} \end{aligned}$$

$$\Rightarrow M \oplus N = M + N = R^2$$

$$\Rightarrow M \oplus N = R^2$$

Thus the direct sum of M and N is $R^2 = V$

Theorem 2.6.2: Let $W_1, W_2, \dots, W_n$ be subspace of a vector space V over a field F then $V = W_1 \oplus W_2 \oplus \dots \oplus W_n$ if and only if every element of V is uniquely expressed as a sum of elements of $W_1, W_2, \dots, W_n$.

Proof: ($\Rightarrow$) Suppose $V = W_1 \oplus W_2 \oplus W_3 \oplus \dots \oplus W_n$ then

$V = W_1 + W_2 + \dots + W_n$ and $W_i \cap W_j = \{0\} \quad \forall_i = 1, 2, \dots, n$. and hence for

$x \in V$ we get $x = x_1 + x_2 + \dots + x_n$ where $x_i \in W_i \quad \forall_i = 1, 2, \dots, n$

Thus, every element of V is expressed as sum of elements of $W_1, W_2, \dots, W_n$.

Now, to show that the expression is unique we also suppose that

$x = y_1 + y_2 + \dots + y_n$ with $y_i \in W_i \quad \forall_i = 1, 2, \dots, n$

Then, $x = x_1 + x_2 + \dots + x_n = y_1 + y_2 + \dots + y_n$ with $x_i, y_i \in W_i$

$\Rightarrow x_1 + x_2 + \dots + x_i + \dots + x_n = y_1 + y_2 + \dots + y_i + \dots + y_n$ where $x_i, y_i \in W_i$

$\Rightarrow (x_i - y_i) = (y_1 - x_1) + (y_2 - x_2) + \dots + (y_{i-1} - x_{i-1}) + (y_{i+1} - x_{i+1}) + \dots + y_n - x_n$

Where $x_i, y_i \in W_i, \forall_i = 1, 2, \dots, n$,

$\Rightarrow (x_i - y_i) \in W_1 + W_2 + \dots + W_{i-1} + W_{i+1} + \dots + W_n = \sum_{j \neq i} W_j$

$\Rightarrow x_i - y_i \in W_i \cap \sum_{j \neq i} W_j$

$\forall_i = 1, 2, \dots, n$, since W_i is subspace of V over F we have

$$x_i - y_i \in W_i$$

$$\Rightarrow x_i - y_i \in W_i \quad n \sum_{j \neq i} w_j = \{0\} \quad \forall_i = 1, 2, \dots, n$$

$$\Rightarrow x_i - y_i \in \{0\} \quad \forall_i = 1, 2, \dots, n$$

$$\Rightarrow x_i - y_i = 0 \quad \forall_i = 1, 2, \dots, n$$

$$\Rightarrow x_i = y_i \quad \forall_i = 1, 2, \dots, n.$$

Hence, every element x of V is uniquely expressed as

$$x = x_1 + x_2 + \dots + x_n \text{ with } x_i \in W_i \quad \forall_i = 1, 2, \dots, n.$$

 Remark 2.7

1. Observe that $W_i \cap \sum_{j \neq i} W_j = \{0\}$ is the same as

$$x_1 + x_2 + \dots + x_n = 0 \text{ and } x_i \in W_i \text{ implies } x_i = 0 \quad \forall_i = 1, 2, \dots, n$$

2. Let V be a vector space over a field F and M, N be subspace of V then

$$V = M \oplus N \text{ if and only if } V = M + N \text{ and } M \cap N = \{0\}$$

Theorem 2.6.3

Let M and N be subspaces of vector space V (over a field F) of finite dimension.

If $V = M \oplus N$ then $\dim V = \dim M + \dim N$

Proof

Let $\dim M = m$ and $\dim N = n$ then there exist sets

$\beta_1 = \{x_1, x_2, \dots, x_m\}$ and $\beta = \{y_1, y_2, \dots, y_n\}$ Which are basis for M and N respectively.

I) Since $V = M \oplus N$, then $\forall v \in V$ we have

$$v = m_1 + n_1, \text{ where } m_1 \in M \text{ and } n_1 \in N$$

$$= \sum_{i=1}^m a_i x_i + \sum_{i=1}^n b_i y_i \text{ where } a, b \in F, \text{ since } \beta \text{ and } \beta \text{ are basis for } M \text{ and } N$$

Respectively

$\Rightarrow$ Every element v in V is a linear combination of. $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$

$\Rightarrow \beta = \{x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n\}$ Generates V

II) Suppose for scalars $a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n$ we have

$$\Rightarrow \sum_{i=1}^m a_i x_i - \sum_{i=1}^n b_i y_i = 0 \Rightarrow \sum_{i=1}^m a_i x_i = \sum_{i=1}^n b_i y_i$$

$$\Rightarrow \sum_{i=1}^m a_i x_i, \sum_{i=1}^n b_i y_i \in \cap nN; \text{ since } \sum_{i=1}^m a_i x_i \in \cap \text{ and } \sum_{i=1}^n b_i y_i \in N$$

$$\Rightarrow \sum_{i=1}^m a_i x_i, \sum_{i=1}^n b_i y_i \in \{0\}; \text{ since } \cap nN = \{0\}$$

$$\Rightarrow \sum_{i=1}^m a_i x_i = 0 = \sum_{i=1}^n b_i y_i$$

$$\Rightarrow a_i = 0 \forall_i = 1, 2, \dots, m \text{ \& } b_i = 0 \forall_i = 1, 2, \dots, n$$

This is because of the fact that β_1 and β_2 are linearly independent over F .

Thus, $\sum_{i=1}^m a_i x_i + \sum_{i=1}^n b_i y_i = 0 \Rightarrow a_1 = 0 = \dots = a_m = b_1 = b_2 = \dots = b_n$

$\Rightarrow \beta = \{x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n\}$ is a linearly independent set

Therefore, by I) and II) we get β is a basis for V over F

$\Rightarrow \dim V = \text{the number of elements in } \beta$

$$= m + n$$

$$= \dim M + \dim N$$

Thus $\dim V = \dim M + \dim N$

Class Activity 2.6

1. Let V be a finite dimensional vector space, then show that any linearly independent set can be extended by adding elements, to get a basis
2. Let W be a subspace of a finite dimensional vector space V , then show that there is another subspace N such that $v = W \oplus N$
3. Let M, N be subspace of a finite dimensional vector space V over a field F , then show that $\dim M + \dim N = \dim(M \cap N) + \dim(M + N)$.

Definition 2.6.3 (Direct products)

Let $V_1, V_2, \dots, V_n$ be vector spaces over a field F then the collection of n -tuples $(x_1, x_2, \dots, x_n)$ where $x_i \in V_i$ with addition and scalar multiplication defined component wise is called the direct product of $V_1, V_2, \dots, \text{and } V_n$ and it is denoted by $V_1 \times V_2 \times V_3 \times \dots \times V_n$.

That is, $V_1 \times V_2 \times \dots \times V_n = \{(x_1, x_2, \dots, x_n) : x_i \in V_i \ \forall_i = 1, 2, \dots, n\}$

Theorem 2.6.4: Let $V_1, V_2, \dots, V_n$ be vector spaces over a field F then show that $V_1 \times V_2 \times \dots \times V_n = \{(x_1, x_2, \dots, x_n) : x_i \in V_i \ \forall_i = 1, 2, \dots, n\}$ With addition and scalar multiplication defined component wise is a gain a vector space over a field F .

Proof: Exercise

 **Note that** The direct product of $V_1, V_2, \dots, V_n$ is also referred to as the (external) direct sum of $V_1, V_2, \dots, V_n$ and denoted by $V_1 \oplus V_2 \oplus V_3 \oplus \dots \oplus V_n$.

Definition 2.6.4 (Internal direct Sum)

Let $N_1, N_2, \dots, N_k$ be subspaces of a vector space V over a field F such that $V = N_1 \oplus N_2 \oplus \dots \oplus N_k$ then V is said to be the (internal) direct sum of $N_1, N_2, \dots, N_k$.

 **Test yourself!**

1. Let $V = \mathbb{R}^3$, W be the subspace generated by $(2, 1)$ U be the subspace generated by $(0, 1)$ and M be the subspace generated by $(1, 1)$ over $\mathbb{R}$ then show that

a) $V = W \oplus U$

b) $V = W \oplus M$

- c) Find dimension of W , U and M
2. Let M and N be subspaces of a finite vector space V over a field F such that $\dim M = 2$, $\dim N = 3$ and $\dim (M \cap N) = 1$, then find $\dim (M + N)$
3. Show that $R^x \not\subset xR$ with component wise addition and scalar multiplication is a vector space over R .
4. Let $W_1, W_2, \dots, W_n$ be subspaces of a vector space V over a field F then show that the following are equivalent.
- a) The map $\Pi: W_1 \times W_2 \times \dots \times W_n \rightarrow W_1 + W_2 + \dots + W_n$ defined by $\Pi(a_1, a_2, \dots, a_n) = a_1 + a_2 + \dots + a_n$ is an isomorphism.
- b) $W_j \cap (W_1 + W_2 + \dots + W_{j-1} + W_{j+1} + \dots + W_n) = \{0\} \forall j = 1, 2, \dots, n$.
- c) Every $x \in W_1 + W_2 + \dots + W_n$ can be written uniquely in the form $a_1 + a_2 + \dots + a_n$ With $a_i \in W_i$.
5. Let $W_1, W_2, \dots, W_n$ be subspaces of a finite dimensional vector space V , and let β_i be a basis for W_i then show that
- a) The ordered set β obtained by listing the bases $\beta_1, \beta_2, \dots, \beta_n$ in order is a basis of V if and only if V is the direct sum $W_1 \oplus W_2 \oplus \dots \oplus W_n$
- b) $\dim (W_1 + W_2 + \dots + W_n) \leq (\dim W_1) + \dots + (\dim W_n)$ with equality if and only if $W_j \cap \sum_{i \neq j} W_i = \{0\} \forall j = 1, 2, \dots, n$.

SUMMARY OF UNIT TWO

- ❖ Let V be an arbitrary nonempty set of objects on which two operations are defined: addition, and multiplication by scalars (numbers). By **addition** we mean a rule for associating with each pair of objects u and v in V an object $u + v$, called the **sum** of u and v ; by **scalar multiplication** we mean a rule for associating with each scalar k and each object u in V an object ku , called the

scalar multiple of $\mathbf{u}$ by k . If the following axioms are satisfied by all objects $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and all scalars k and m , then we call V a **vector space** and we call the objects in V **vectors**.

1. If $\mathbf{u}$ and $\mathbf{v}$ are objects in V , then $\mathbf{u} + \mathbf{v}$ is in V
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
4. There is an object $\mathbf{0}$ in V , called a **zero vector** for V , such that $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u}$ in V .
5. For each $\mathbf{u}$ in V , there is an object in V , called a **negative** of $\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$.
6. If k is any scalar and $\mathbf{u}$ is any object in V , then $k\mathbf{u}$ is in V .
7. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
8. $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$
9. $k(m\mathbf{u}) = (km)(\mathbf{u})$
10. $1.\mathbf{u} = \mathbf{u}$

❖ Let V be a vector space, $\mathbf{u}$ a vector in V , and k a scalar; then:

- (a) $\mathbf{0}.\mathbf{u} = \mathbf{0}$
- (b) $k.\mathbf{0} = \mathbf{0}$
- (c) $(-1)\mathbf{u} = -\mathbf{u}$
- (d) If $k.\mathbf{u} = \mathbf{0}$, then $k = \mathbf{0}$ or $\mathbf{u} = \mathbf{0}$.

- ❖ The additive identity element in V is unique.
- ❖ The additive inverse of every element of V is unique.
- ❖ A subset W of a vector space V is called a **subspace** of V if W is itself a vector space under the addition and scalar multiplication defined on V .
- ❖ If W is a set of one or more vectors from a vector space V , then W is a subspace of V if and only if the following conditions hold.
 - (a) If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W , then $\mathbf{u} + \mathbf{v}$ is in W .
 - (b) If k is any scalar and $\mathbf{u}$ is any vector in W , then $k\mathbf{u}$ is in W .
- ❖ A vector $\mathbf{w}$ is called a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_r$ if it can be expressed in the form $\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \dots + k_r\mathbf{v}_r$

Where $k_1, k_2, \dots, k_r$ are scalars.

- ❖ If $S = \{v_1, v_2, \dots, v_r\}$ is a nonempty set of vectors, then the vector equation $k_1v_1 + k_2v_2 + \dots + k_rv_r = \mathbf{0}$ has at least one solution, namely $k_1 = 0, k_2 = 0, \dots, k_r = 0$. If this is the only solution, then S is called a **linearly independent** set. If there are other solutions, then S is called a **linearly dependent** set.
 - ❖ A set S with two or more vectors is
 - (a) Linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S .
 - (b) Linearly independent if and only if no vector in S is expressible as a linear combination of the other vectors in S .
 - ❖ A finite set of vectors that contains the zero vector is linearly dependent.
 - ❖ A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.
 - ❖ If V is any vector space and $S = \{v_1, v_2, \dots, v_n\}$ is a set of vectors in V , then S is called a **basis** for V if the following two conditions hold:
 - (a) S is linearly independent.
 - (b) S spans V .
 - ❖ If $S = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space V , then every vector v in V can be expressed in the form $v = c_1v_1 + c_2v_2 + \dots + c_nv_n$ in exactly one way.
 - ❖ A nonzero vector space V is called **finite-dimensional** if it contains a finite set of vectors $\{v_1, v_2, v_3, \dots, v_n\}$ that forms a basis. If no such set exists, V is called **infinite-dimensional**.
- In addition, we shall regard the zero vector space to be finite dimensional.
- ❖ The **dimension** of a finite-dimensional vector space V , denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V . In addition, we define the zero vector space to have dimension zero.
 - ❖ Let $W_1, W_2, \dots, W_n$ be subspace of a vector space V over a field F then the set of all vectors of the form $v_1 + \dots + v_n$ where $v_i \in W_i$ is called the sum of the

subspaces or their span, and is denoted by

$$W_1 + W_2 + \dots + W_n = \{v_1 + v_2 + \dots + v_n : v_i \in W_i \quad \forall i = 1, 2, \dots, n\}$$

❖ Let $W_1, W_2, \dots, W_n$ be subspaces of a vector spaces V over a field F , such

that $\bigcap_{i=1}^n W_i = \{0\}$ then the direct sum of $W_1, W_2, \dots, W_n$, denoted by

$$W_1 \oplus W_2 \oplus W_3 \oplus \dots \oplus W_n \text{ is the subspace } W_1 + W_2 + \dots + W_n.$$

❖ Let $V_1, V_2, \dots, V_n$ be vector spaces over a field F then the collection of n -tuples

$(x_1, x_2, \dots, x_n)$ where $x_i \in V_i$ with addition and scalar multiplication defined

component wise is called the direct product of $V_1, V_2, \dots$, and V_n , denoted by

$$V_1 \times V_2 \times V_3 \times \dots \times V_n \text{ is a vector space over a field } F$$

❖ Let $N_1, N_2, \dots, N_k$ be subspaces of a vector space V over a field F such that

$$V = N_1 \oplus N_2 \oplus \dots \oplus N_k \text{ then } V \text{ is said to be the (internal) direct sum of}$$

$$N_1, N_2, \dots, N_k.$$

UNIT-THREE: MATRICES

UNIT OBJECTIVES:

At the end of this unit each student should able to:

- Learn a matrix and know properties of matrices
- Understand how to operate with matrices
- Know properties of matrix operations
- Realize inverses of a matrices
- Know types of matrix
- Work with elementary row/column operations

- Represent system of linear equations by a matrix.

Introduction

In working with a system of linear equations such as

$$x_1 - 2x_2 + 3x_3 + x_4 = -2$$

$$2x_1 + 3x_2 - 7x_3 - 5x_4 = 1$$

$$-x_1 + 10x_2 + x_3 - 2x_4 = 1$$

$$8x_1 + 4x_2 - 11x_3 + 8x_4 = 0$$

Only the coefficients and their respective positions are important. Thus these coefficients can be efficiently arranged in a rectangular array called a "matrix".

Moreover, certain abstract objects in higher mathematics can also be represented by matrices. Thus, matrices can be used to solve systems of linear equations. They have also numerous applications.

3.1: Definitions of Matrices

1- The concept of a Matrix

Definition 3.1: Definition of a Matrix

Let m, n be natural numbers. An $m \times n$ (read "m by n") matrix is a collection of mn numbers arranged in a rectangular array.

$$\begin{array}{c}
 \text{n columns} \\
 \downarrow \\
 \begin{array}{c}
 \begin{array}{c} a_{11} \ a_{12} \ \cdots \ a_{1n} \\ a_{21} \ a_{22} \ \cdots \ a_{2n} \\ ' \quad ' \quad ' \\ ' \quad ' \quad ' \\ ' \quad ' \quad ' \\ a_{m1} \ a_{m2} \ a_{mn} \end{array} \\
 \text{m rows} \rightarrow
 \end{array}
 \end{array}$$

in which each entry, a_{ij} , of the matrix is a number. An $m \times n$ matrix has m rows (horizontal lines) and n columns (vertical lines).

 Remarks 3.1

1. The numbers in the above matrix are called the matrix entries and are denoted by a_{ij} , Where i, j are indices (which are natural numbers) with $1 \leq i \leq m$ and $1 \leq j \leq n$.

The index i is called the row index or row subscript because it gives the position in the horizontal lines, and j is the column index or column subscript because it gives the position in the vertical lines. So a_{ij} is the entry which appears in the i^{th} row and j^{th} column of the matrix.

2. The plural of matrix is matrices. If each entry of a matrix is a real number / complex number, then the matrix is called a real/ complex matrix.
3. An $m \times n$ matrix is said to be of order $m \times n$. If $m=n$, then the matrix is square of order n . For a square matrix, the entries $a_{11}, a_{22}, a_{33}, \dots$ are called the main diagonal entries.
4. We can abbreviate the above matrix by writing $A = (a_{ij})_{m \times n}$.

Example-1: Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ e & \sqrt{2} & -1 & 5 \\ \frac{1}{2} & -3 & 0 & 10 \end{pmatrix}$$

Since A has three rows and four columns, we say that A is a 3×4 matrix or A is a matrix of order 3×4 . Moreover, $a_{11} = 0, a_{12} = 1, a_{13} = 2, a_{14} = 3$

$$a_{21} = e, a_{22} = \sqrt{2}, a_{23} = -1, a_{24} = 5, a_{31} = \frac{1}{2}, a_{32} = -3, a_{33} = 0, a_{34} = 10$$

Example-2: The following matrices have the indicated orders.

(a) Order: 1×1

$$[2]$$

(c) Order: 3×3

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(b) Order: 1×4

(d) Order: 3×2

$$\begin{bmatrix} 0 & 2 & -1 & 3 \end{bmatrix} \quad \begin{bmatrix} e & 1 \\ 2 & -1 \\ -8 & 1 \end{bmatrix}$$

☺ Test yourself!

1. What is the order of each of the following matrices?

$$(a) \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 5 & 6 \end{bmatrix} \quad (b) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (c) [\sqrt{3}] \quad (d) \begin{bmatrix} 4 & 5 & 6 \\ 1 & -2 & 4 \\ 0 & 1 & 3 \\ -2 & 0 & _ \end{bmatrix}$$

2. What are the entries a_{21} , a_{12} , and a_{23} of the matrix

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 0 & 5 \\ 0 & 7 & -1 \end{bmatrix} ? \quad B = \begin{bmatrix} \cos A & -\sin A & \tan A \\ 0 & 1 & 2 \\ \sin A & -\cos A & 0 \end{bmatrix}$$

3. What are the main diagonal entries of the following matrices?

$$(a) \begin{pmatrix} 1 & 1+i \\ 0 & i \end{pmatrix} \quad (b) \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & 0 & 7 \\ 0 & 4 & 1 \\ 3 & 5 & 8+i \end{bmatrix}$$

4. Construct a 4x3 matrix $B = (b_{ij})$ where

a) $b_{ij} = i + j$ c) $b_{ij} = i \cdot j$
 b) $b_{ij} = i - j$ d) $b_{ij} = i/j$

3.2. The Algebra of Matrices

In this section we concentrated on the mechanics of the three basic matrix operations: Matrix addition, scalar multiplication, and matrix multiplication and develop an algebra of matrices. You will see that this algebra shares many (but not all) of the properties of the algebra of scalars.

Objectives

At the end of this section students should be able to

- Manipulate matrix addition and subtraction.
- Identify equal matrix using definition of equal matrix.
- Perform scalar multiplication and addition simultaneously.
- Compute the product of matrices

1. Equality of matrices

Definition 3.2 Definition of Equality of Matrices

Two Matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are equal, denoted by $A = B$ if and only if

I) they have the same order ($m \times n$) and

II) $a_{ij} = b_{ij} \quad \forall i, j$ Such that $1 \leq i \leq m$ and $1 \leq j \leq n$.

That is, two matrices of the same order are said to be equal if their corresponding entries are equal.

Example-3 Equality of Matrices

Solve for x, y, z and w in the following matrix equation:

$$\begin{bmatrix} x+1 & 0 \\ 5 & z-1 \end{bmatrix} = \begin{bmatrix} 2 & y \\ w & -2 \end{bmatrix}$$

Solution:

Because two matrices are equal only if the corresponding entries are equal, we conclude that: $x+1=2, y=0, w=5$ and $z-1=-2$

$$\Rightarrow x=1, y=0, w=5, z=-1.$$

2. Addition and subtraction of matrices

Definition 3.3: Definition of matrix Addition

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be matrices of order $m \times n$, then their sum is an $m \times n$ matrix given by $A+B = [a_{ij} + b_{ij}]$ That is, we add the two matrices of the same order by adding their corresponding entries and the sum of two matrices of different orders is undefined.

Example-4: Addition of Matrices

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

Then find a) $A+B$ b) $A+C$.

Solution:

a) Since A and B have different orders, we have $A + B$ is undefined.

$$\begin{aligned} \text{b) } A+C &= \begin{bmatrix} 2 & 1 & 0 \\ 4 & 0 & -1 \\ 3 & -2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2+1 & 1+0 & 0+0 \\ 4+0 & 0+-1 & 0 \\ 3+0 & -2+0 & 2+1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 4 & -1 & 0 \\ 3 & -2 & 3 \end{bmatrix} \end{aligned}$$

Definition 3.4: Definition of Matrix Subtraction

Let A and B be two Matrices of the same order then the difference of A and B , denoted by $A - B$, is defined to be the matrix obtained by subtracting each element of B from the corresponding elements of A . That is if

$$A = (a_{ij})_{m \times n} \text{ \& } B = (b_{ij})_{m \times n} \text{ then } A - B = (a_{ij} - b_{ij})_{m \times n}$$

Example-5: Subtraction of Matrices

$$\text{Let } A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 3 & 0 \\ -1 & 2 & 5 \end{pmatrix} \text{ then}$$

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

 **Note that:** Two matrices of the same order are said to be conformable for addition and subtraction. Only conformable matrices can be added or subtracted.

3. Multiplication of a matrix by scalar

Definition 3.5: Definition of scalar multiplication

Let $A = (a_{ij})_{m \times n}$ be a matrix and b be scalar, and then a scalar multiple of A by b is an $m \times n$ matrix given by: $bA = (ba_{ij})$. That is, bA is an $m \times n$ matrix obtained from A by multiplying each of its elements by b .

Example-6: Scalar multiplication

Let $A = \begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix}$ then find

- a) $3A$
- b) $-B$
- c) $3A - B$

Solution:

$$\text{a) } 3A = 3 \begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3(1) & 3(2) & 3(4) \\ 3(-3) & 3(0) & 3(-1) \\ 3(2) & 3(1) & 3(2) \end{bmatrix} = \begin{bmatrix} 3 & 6 & 12 \\ -9 & 0 & -3 \\ 6 & 3 & 6 \end{bmatrix}$$

$$\text{b) } -B = (-1)B = (-1) \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} (-1)2 & (-1)0 & (-1)0 \\ (-1)1 & (-1)(-4) & (-1)3 \\ (-1)(-1) & (-1)3 & (-1)2 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ -1 & 4 & -3 \\ 1 & -3 & -2 \end{bmatrix}$$

$$\text{c) } 3A - B = \begin{bmatrix} 3 & 6 & 12 \\ -9 & 0 & -3 \\ 6 & 3 & 6 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 12 \\ -10 & 4 & -6 \\ 7 & 0 & 4 \end{bmatrix}$$

 Note:

- (1) If A is any matrix and b is any scalar then $bA = Ab$
- (2) $-A$ represent the scalar product $(-1)A$.

Matrix Multiplication

It is a basic matrix operation. At first glance the following definition may seem unusual. You will see later, however, that this definition of the product of two matrices has many practical applications.

Definition 3.6: Definition of Matrix multiplication

Let $A = [a_{ij}]$ be an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the product AB is

$$\text{an } m \times p \text{ matrix given by } AB = [c_{ij}] \text{ where } c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \\ = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} + \dots + a_{in} b_{nj}$$

That is, the ij^{th} element of AB is obtained by multiplying each element of the i^{th} row of A (denoted by A_i) in to the corresponding element of the j^{th} column of B (denoted by B_j) and adding the products or equivalently, the ij^{th} element of AB is the dot product of the i^{th} row vector of A by the j^{th} column vector of B and hence

$$AB = [c_{ij}] \text{ where } c_{ij} = A_i \cdot B_j$$

Example-7: Finding the Product of two Matrices

$$\text{Let } A = \begin{bmatrix} -1 & 3 \\ 4 & 2 \\ 5 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} \text{ then find the product } AB.$$

Solution:

First note that the product AB is defined because A has order 3×2 , and will take the form.

$$\begin{bmatrix} -1 & 3 \\ 4 & 2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

To find c_{11} (the entry in the first row and first column of the product) we take the dot product of the first row of A and the first column of B That is, $c_{11} = A_1 \cdot B^1$

$$= (-1 \ 3) \cdot \begin{pmatrix} -3 \\ -4 \end{pmatrix} = (-1)(-3) + 3(-4) = 3 - 12 = -9$$

Similarly, to find c_{12} , we take the dot product of the first row of A and the second column of B and hence $c_{12} = A_1 \cdot B^2$

$$= (-1 \ 3) \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ = (-1)(2) + (3)(1) = -2 + 3 = 1$$

Continuing this pattern produces the following results.

$$c_{21} = A_2 B^1 = (4 - 2) = \begin{pmatrix} -3 \\ -4 \end{pmatrix} \cdot (4) (-3) + (-2)(-4) = -4$$

$$c_{22} = A_2 B^2 = (4 - 2) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot (4) (2) + (-2)(1) = 6$$

$$c_{31} = A_3 B^1 = (5 \ 0) \cdot \begin{pmatrix} -3 \\ -4 \end{pmatrix} = (5) (-3) + (0)(-4) = -15$$

$$c_{32} = A_3 B^2 = (5 \ 0) \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = (5) (2) + (0)(1) = 10$$

Thus the product is

$$AB = \begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} -9 & 1 \\ -4 & 6 \\ -15 & 10 \end{bmatrix}$$



Remarks 3.2

1. The product of two matrices A and B is defined if the number of columns of A is equal to the number of rows of B , say if A is an $m \times n$ matrix and B is an $n \times p$ matrix. In this case, the product AB is an $m \times p$ matrix.
2. Let $A = (a_{ij}) m \times n$ and $B = (b_{ij}) n \times p$ then

$$AB = \begin{bmatrix} A_1 B^1 & A_1 B^2 & A_1 B^3 & \dots & A_1 B^p \\ A_2 B^1 & A_2 B^2 & A_2 B^3 & \dots & A_2 B^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_m B^1 & A_m B^2 & A_m B^3 & \dots & A_m B^p \end{bmatrix}$$

Where $A_i B_j$ is the dot product of the i^{th} row of A and j^{th} column of B

$\forall_i = 1, 2, \dots, m$ and $j = 1, 2, \dots, p$

3. If the product that AB is defined, then it is not necessary that BA must also be defined. For instance, if A is of order 3×3 and B is of order 3×2 , then clearly AB is defined but BA is not defined, as the number of columns of B is not equal to the number of rows of A .

Class Activity 3.1

1. Compute the products AB and BA for the following values of A and B

$$a) \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix}, B = \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix}$$

$$b) \quad A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$c) \quad A = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

2. Find

$$(a) \quad x, y, z \text{ and } w \text{ if } \begin{pmatrix} 2 & 1 \\ x & y \end{pmatrix} \begin{pmatrix} 8 & -3 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} z & w \\ -16 & 8 \end{pmatrix}$$

$$(b) \quad a, b, c, \& d \text{ if } \begin{pmatrix} a & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 4 & 7 \\ -3 & 6 \end{pmatrix} = \begin{pmatrix} 13 & 1 \\ c & d \end{pmatrix}$$

$$3. \text{ Let } A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}. \text{ Show that } AB = BA$$

4. Construct two matrices A and B such that the product AB , BA are defined and $AB \neq BA$.

5. Properties of Matrix Operations

We begin by listing several properties of matrix addition and scalar multiplication

Theorem 3.1: Properties of Matrix Addition and scalar multiplication.

Let A, B, C be an $m \times n$ matrices and c and d be scalars, then the following properties are true.

1. $A + B = B + A$ ----- commutative property of Addition
2. $A + (B + C) = (A + B) + C$ ----- Associative property of Addition
3. $(cd)A = c(dA)$
4. $1 A = A$
5. $c(A + B) = cA + cB$ -----Distributive property
6. $(c + d)A = cA + dA$ -----Distributive property A

Proof:

Proof of these six properties follows directly from the definitions of matrix addition, scalar multiplication, and the corresponding properties of scalars. For example, to

prove the commutative property of matrix addition, we let $A = [a_{ij}]$ and $B = [b_{ij}]$. Then, using the commutative property of scalars, we have

$$A + B = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = B + A$$

similarly, to prove property 5, we use the distributive property for scalars of multiplication over addition and hence

$$c(A + B) = [c(a_{ij} + b_{ij})] = [ca_{ij} + cb_{ij}] = cA + cB.$$

The proofs of the remaining four properties are left as an exercise.

 **Note:** If A is an $m \times n$ matrix and O_{mn} is the $m \times n$ matrix consisting entirely of zeros then one can observe that $A + O_{mn} = A$. We call O_{mn} a zero matrix or null matrix and it serves as the additive identity for the set of all $m \times n$ matrices. For example, the matrix $O_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ serves as the additive identity for the set of all 2×3 matrices.

Theorem 3.2: Properties of zero Matrices

Let A be an $m \times n$ matrix and c be a scalar then

a) $A + O_{mn} = A$

b) $A + (-A) = O_{mn}$

c) $cA = O_{mn} \Rightarrow c = 0 \text{ or } A = O_{mn}$

Proof: Left as an exercise.

Remarks 3.3:

- Property b) above can be described by saying that the matrix $-A$ is the additive inverse of A .
- The algebra of real numbers (scalars) and the algebra of matrices have many similarities. For example, compare the following solutions.

Real numbers

(Solve for x)

$$x + a = b$$

$$\Rightarrow x + a + (-a) = b + (-a)$$

$$\Rightarrow x + 0 = b - a$$

$$\Rightarrow x = b - a$$

mxn Matrices

(Solve for x)

$$x + A = B$$

$$\Rightarrow x + A + (-A) = B + (-A)$$

$$\Rightarrow x + 0 = B - A$$

$$\Rightarrow x = B - A$$

Theorem: 3.3: Properties of Matrix Multiplication

Let A , B and C be matrices with orders such that the given matrix products are defined and b is a scalar, then the following properties are true.

- a) $A(BC) = (AB)C$ -----Associative property of multiplication
- b) $A(B + C) = AB + AC$ -----Left distributive property
- c) $(A + B)C = AC + BC$ -----Right distributive property
- d) $b(AB) = (bA)B = A(bB)$

Proof:

To prove property a) We compute the entry in the i^{th} row and j^{th} column of both $A(BC)$ and $(AB)C$. To begin, we assume that A has order $m \times n$, B has order $n \times p$,

and C has order $p \times q$. Then the j^{th} column of BC is

$$\begin{bmatrix} b_{11} c_{1j} + b_{12} c_{2j} + \dots + b_{1p} c_{pj} \\ b_{21} c_{1j} + b_{22} c_{2j} + \dots + b_{2p} c_{pj} \\ \vdots \\ b_{n1} c_{1j} + b_{n2} c_{2j} + \dots + b_{np} c_{pj} \end{bmatrix}$$

This implies that the i^{th} entry of the j^{th} column of $A(BC)$ is

$a_{i1}(b_{11} c_{1j} + \dots + b_{1p} c_{pj}) + \dots + a_{in}(b_{n1} c_{1j} + b_{np} c_{pj})$. On the other hand, the i^{th} row of AB is $(a_{i1} b_{11} + \dots + a_{in} b_{n1}), \dots, (a_{i1} b_{1p} + \dots + a_{in} b_{np})$ which implies that the i^{th} entry of the j^{th} column of $(AB)C$ is

$$(a_{i1} b_{11} + \dots + a_{in} b_{n1})c_{1j} + \dots + (a_{i1} b_{1p} + \dots + a_{in} b_{np})c_{pj}.$$

By distributing and regrouping, we can see these two $(ij)^{th}$ entries are equal and we conclude that $A(BC) = (AB)C$.

To prove b) we find the entry in the i^{th} row and j^{th} column of $A(B+C)$ to be

$$a_{i1}(b_{1j} + c_{1j}) + \dots + a_{in}(b_{nj} + c_{nj})$$

Moreover, the entry in the i^{th} row and j^{th} column of $AB + AC$ is

$$(a_{i1} b_{1j} + \dots + a_{in} b_{nj}) + (a_{i1} c_{1j} + \dots + a_{in} c_{nj})$$

Again, by distributing and regrouping, we can see that these two $(ij)^{th}$ entries are equal and we conclude that $A(B + C) = AB + AC$.

The proofs of c) and d) left as an exercises.

Class Activity 3.2

1. Let $A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix}$ then solve for X in the equation $3X + A = B$

2. Let $A = \begin{bmatrix} -1 & 0 & 5 \\ 7 & -2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 7 & 0 \\ -3 & -1 & 0 \\ 1 & 0 & 5 \end{bmatrix}$ and $C = \begin{bmatrix} -1 & -1 \\ 2 & 0 \\ 0 & 4 \end{bmatrix}$ then show that

$$A(BC) = (AB)C$$

3. Consider $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \\ 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & -3 \\ 3 & -4 & 5 \end{bmatrix}$ and $C = \begin{bmatrix} -3 & -4 & 5 \\ -5 & 6 & -7 \end{bmatrix}$ then show that

a) $A(B + C) = AB + AC$

b) $(B + C)A = BA + CA$

4. Let $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$ then show that $AB \neq BA$ and conclude that matrix multiplication is not commutative.

5. Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$ then show that $AC = BC$, Yet $A \neq B$ and conclude that a general cancellation property for matrix multiplication does not hold.

Definition 3.7: Identity matrix

A special type of square matrix of order n that has 1's on the main diagonal and 0's elsewhere is called the identity matrix of order n and it is denoted by I_n .

That is,
$$I_n = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

It serves as the identity for matrix multiplication.

For instance, if $n = 1, 2$, or 3 we have

$$I_1 = [1], \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem 3.4 Properties of the identity Matrix

If A is a matrix of order $m \times n$, then

- a) $A I_n = A$
- b) $I_m A = A$
- c) If $m=n$ then $A I_n = I_n A = A$

Proof: Left as an exercise

Example-8:

Let $A = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}$ then

$$a) A I_2 = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} = A = I_2 A$$

$$b) I_3 B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}$$

 **Note:** For repeated multiplication of square matrices, we use the same exponential notation used with real numbers.

That is, $A^1 = A$

$$A^2 = AA \text{ and for } k \in \mathbb{N}$$

We define $A^k = \underbrace{AA \cdots A}_{K \text{ factors}}$

It is convenient also to define $A^0 = I_n$, where A is a square matrix of order n

As a result one can observe that the following properties.

- a) $A^j A^k = A^{j+k}$
- b) $(A^j)^k = A^{jk}$

😊 Test yourself!

1. Perform the indicated operations, given that $a=3, b=-4, c=-2$ and

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- (a) $B(CA)$
- (d) $(a-b)(B-C)$

- (b) $B + C)A$ (e) $B(C + 0)$
 (c) $(a + b)B$ $B(CA)$ (f) $B(aA)$

2. Perform the indicated operations, given that $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$

- (a) A^2 (c) $A + I_2 A$
 (b) $A(I_2 + A)$ (d) A^4

3. Let $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ be a polynomial function over $\mathbb{R}$ and A be an $n \times n$ matrix, then using the definition

$f(A) = a_0I_n + a_1A + a_2A^2 + \dots + a_nA^n$, find $f(A)$ if

(a) $f(x) = x^2 - 10x + 24$ and $A = \begin{bmatrix} 8 & -4 \\ 2 & 2 \end{bmatrix}$

(b) $f(x) = x^3 - 10x^2 + 31x - 30$ and $A = \begin{bmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{bmatrix}$

4. Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ then compute (a) A^{19} (b) A^{20}

5. Find a matrix A that satisfies the given equation

(a) $A^2 = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$ (b) $A^3 = \begin{bmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 27 \end{bmatrix}$

6. Let A, B and C be matrices with orders such that the given matrix operations are defined. Determine whether the given statement is true or false. Give a reason for your answer.

- (a) $(A + B)^2 = A^2 + 2AB + B^2$
 (b) $(A + B)(A - B) = A^2 - B^2$
 (c) $AB = O \Rightarrow A = O$ or $B = O$
 (d) $AC = BC \Rightarrow A = B$

7. Let $M_{m \times n}$ be the set of all $m \times n$ matrices over $\mathbb{R}$ then

- a) Show that $M_{m \times n}$ with the standard operations is a vector space over $\mathbb{R}$
 b) Find a basis for $M_{m \times n}$ and its dimension.

8. Let M_n be the set of all $n \times n$ matrices over R then show that M_n with the standard operations is a ring with unity which is not commutative.

9. Let $A = \left\{ \begin{bmatrix} a & b \\ c & o \end{bmatrix} : a, b, c, \in R \right\}$ and

$B = \left\{ \begin{bmatrix} a & b \\ c & 1 \end{bmatrix} : a, b, c \in R \right\}$ then with standard operation determine whether

(a) A is a vector space over R .

(b) B is a commutative ring.

3.3 Types of Matrices: Square, identity, scalar, diagonal, triangular, symmetric, an Skew symmetric matrices

Objectives

At the end of this section students should be able to

- Identify types of matrices.
- Find trace of matrices.
- Compute the inverse of matrix.
- Proof properties of inverse and transpose of matrices.
- Determine the product of special matrices.

1) **Row Matrix:** A matrix which has only one row. It is a $1 \times n$ matrix, where n is arbitrary natural number. It is also called a row vector

Example-1: $(-1 \ 0 \ 0 \ 4)$ is a row matrix with order 1×4 .

$(1 \ -2 \ -3)$ is a 1×3 row matrix.

2) **Column Matrix:** A matrix which has one column. It is also called a column vector.

Example-2: $\begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$ is a 3 by 1 column matrix.

3) **Zero matrix:** is a matrix of arbitrary size in which all the entries are zero. It is also called null matrix and denoted by $\mathbf{0}$.

Example-3: The following are null matrices

$$(0 \ 0) \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

↑ order 1x2 ↑ order 2x2 ↑ order 3x2

- 4) **A square matrix**: Is a matrix where the number of rows is equal to the number of columns. It is an $n \times n$ matrix; where $n \in \mathbb{N}$.

Example 4: $\begin{bmatrix} 2 & 1 \\ 3 & 8 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & -3 \\ 4 & 1 & 3 \\ 6 & 7 & \frac{1}{2} \end{bmatrix}$

Square matrix of
Order 2x2

square matrix of
order 3x3

Let $A = (a_{ij})$ be a square matrix of order n , then

- a) The entries $a_{ii}; 1 \leq i \leq n$ are called main diagonal entries of A .
b) The sum of the main diagonal entries of A is called the trace of A and it is

denoted by $Tr(A)$ That is, $Tr(A) = a_{11} + a_{22} + a_{33} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$

Example-5: Find the trace of the following matrix

(a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & 4 \\ 3 & 1 & 3 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Solution: (a) $Tr(A) = \sum_{i=1}^3 a_{ii} = 1 + (-2) + 3 = 2$

(b) $Tr(B) = \sum_{i=1}^3 b_{ii} = b_{11} + b_{22} + b_{33} = 1 + 1 + 1 = 3$

- 5) **Diagonal Matrix**: Is a square matrix in which all entries that are not on the main diagonal are zero.

Example-6: The following are Diagonal matrices

(a) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$(c) \begin{bmatrix} a_{11} & 0 & 0 \dots & 0 \\ 0 & a_{22} & 0 \dots & 0 \\ 0 & 0 & a_{33} \dots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \dots & a_{nn} \end{bmatrix}$$

Note: If $A = (a_{ij})_{n \times n}$ is a diagonal matrix then $a_{ij} = 0 \forall i \neq j$ while a_{ii} can be zero or different from zero.

6) **Scalar Matrix:** Is a diagonal matrix whose main diagonal elements are equal. That is, if $A = (a_{ij})$ is a square matrix such that

$$a_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ c, & \text{if } i = j \end{cases} \quad \text{Where } c \text{ is a scalar, then } A \text{ is called scalar matrix.}$$

Example-7: $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

7) **Identity Matrix:** Is a diagonal matrix in which all the main diagonal elements are equal to 1. It is also called a unit matrix.

Example-8: $[1]$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are unit matrices of order 1, 2 and 3

respectively

 **Note**

1. The $n \times n$ identity matrix is denoted by I_n

2. The identity matrix behaves like 1 in multiplication of numbers: If A is a $m \times n$ matrix, then $I_m A = A$ and $A I_n = A$

3- $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$ Are some short hand ways of drawing the matrix I_n .

8) **Upper triangular Matrix:** Is a square matrix in which all the elements below the main diagonal are zero. That is, if $A = (a_{ij})_{n \times n}$ is a square matrix such that

$$a_{ij} = 0 \text{ for all } i > j \quad \text{then } A \text{ is said to be upper triangular matrix.}$$

Example-9:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ are upper triangular matrices of order 3 and 4}$$

respectively.

9) **Lower triangular matrix:** Is a square matrix whose entries above the main diagonal are 0.

$$\text{Example-10: } \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ are lower triangular matrices of order 3}$$

and 4 respectively.

 **Note :**

- 1) We often indicate that a whole region in a matrix consists of zeros by leaving it blank or by putting in a single O.
- 2) We use * to indicate an arbitrary undetermined entry of a matrix.

Definition 3.8: The Transpose of a Matrix

The transpose of an $m \times n$ matrix $A = (a_{ij})$ is the $n \times m$ matrix A^t obtained by reflecting about the diagonal: $A^t = [b_{ij}]$ where $b_{ij} = a_{ji}$. It is a matrix obtained by interchanging the corresponding rows and columns of the given matrix.

Example-11: The Transpose of a matrix

$$(a) A = \begin{bmatrix} 3 \\ -4 \end{bmatrix} \quad (b) B = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(c) C = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

Solution: (a) $A^t = [3 - 4]$ (b) $B^t = \begin{bmatrix} 3 & 4 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(c) $C^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \dots & a_{m3} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{mn} \end{bmatrix}$

Theorem 3.5: Properties of Transposes

Let A and B be matrices with orders such that the given matrix operations are defined and c is a scalar, then the following properties are true

- (a) $(A^t)^t = A$ -----Transpose of transpose
- (b) $(A + B)^t = A^t + B^t$ -----Transpose of a sum
- (c) $(cA)^t = c(A^t)$ -----Transpose of a scalar multiple
- (d) $(AB)^t = B^t A^t$ -----Transpose of a product
- (e) $(I_n)^t = I_n$

Proof: Exercise

Definition 3.9: A matrix A is said to be symmetric if it is equal to its transpose. That is, if $A^t = A$ then A is called symmetric matrix.

 **Note:** A symmetric matrix is necessarily a square matrix

Example-12: Let $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 3 \\ 2 & 3 & 7 \end{bmatrix}$ then since $A = A^t$ we have A is symmetric

Definition 3.10: A square matrix A is said to be skew-symmetric if $A^t = -A$.

Example-13: Let $A = \begin{bmatrix} 2 & 0 \\ -2 & 0 \end{bmatrix}$, then $A^t = -A$ and hence A is skew symmetric.

Class Activity 3.2

1. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$, $B = \begin{bmatrix} -3 & -1 \\ 2 & 1 \end{bmatrix}$ and $c = 4$ then verify that

(a) $(A^t)^t = A$ (b) $(A+B)^t = A^t + B^t$

(c) $(cB)^t = cB^t$ (d) $(AB)^t = B^t A^t$

2. Determine whether the given matrix is symmetric, skew symmetric, or neither

(a) $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$ (b) $B = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}$ (c) $C = \begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & -3 \\ 1 & 3 & 0 \end{bmatrix}$ (d) $D = \begin{bmatrix} 2 & 5 & 3 \\ -3 & 6 & 0 \\ 4 & 1 & 1 \end{bmatrix}$

3. Give an example of two 2×2 matrices A and B such that $(AB)^t \neq A^t B^t$

Definition 3.11: Definition of an inverse of a matrix

An $n \times n$ matrix A is invertible (or nonsingular) if there exists an $n \times n$ matrix B such that $AB = I_n = BA$; where I_n is the identity matrix of order n . The matrix B is called the (multiplicative) inverse of A . A matrix that doesn't have an inverse is called noninvertible (or singular). The inverse of a matrix A is denoted by A^{-1} .

Example-14:

Find the inverse of the following matrices

(a) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (b) $B = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$

Solution:

(a) To find the inverse of A . We try to solve the matrix equation

$$AX = I \text{ for } X \text{ But, } AX = I$$

$$\Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_{11} + x_{21} & x_{12} + x_{22} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Now, by equation the corresponding entries, we obtain the following two systems of linear equations.

$$\left\{ \begin{array}{l} x_{11} + x_{21} = 1 \\ x_{21} = 0 \end{array} \right. \text{ and } \left\{ \begin{array}{l} x_{12} + x_{22} = 0 \\ x_{22} = 1 \end{array} \right.$$

$$\Rightarrow x_{11} = 1, x_{21} = 0, \quad x_{12} = -1 \text{ and } x_{22} = 1$$

$$\Rightarrow x = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Therefore, the inverse of A is $A^{-1} = x = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

(b) Suppose $B^{-1} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ be the inverse of B then $B^{-1}B = I_2 = B B^{-1}$ But,

$$B B^{-1} = I_2$$

$$\Rightarrow \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 2b_{11} + b_{21} & 2b_{12} + b_{22} \\ 5b_{11} + 3b_{21} & 5b_{12} + 3b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \left\{ \begin{array}{l} 2b_{11} + b_{21} = 1 \\ 5b_{11} + 3b_{21} = 0 \end{array} \right. \text{ and } \left\{ \begin{array}{l} 2b_{12} + b_{22} = 0 \\ 5b_{12} + 3b_{22} = 1 \end{array} \right.$$

$$\Rightarrow b_{11} = 3, \quad b_{21} = -5, \quad b_{12} = -1 \text{ and } b_{22} = 2$$

$$\Rightarrow B^{-1} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

Infact, one can check that $B B^{-1} = I_2 = B^{-1}B$

Remarks

1. Non square matrices do not have inverses. To see this, note that if A is of order $m \times n$ and B is of order $n \times m$ (where $m \neq n$), then the products AB and BA are of different orders and therefore could not be equal to each other.

2. Not all square matrices possess inverse. For example, the matrix $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ doesn't have inverse.
3. We will see later that square matrix A is invertible if there is a square matrix B such that either one of the two relations $AB = I$ or $BA = I$ holds, or then B is the inverse of A .

Theorem 3.6: Uniqueness of an Inverses Matrix

If A is an invertible matrix, then its inverse is unique

Proof:

Since A is invertible, we know that it has at least one inverse B such that $AB = I = BA$. Suppose that A has another inverse C such that $AC = I = CA$, then

$$\begin{aligned} B &= I B = (CA)B \\ &= C(AB); \text{ Why?} \\ &= CI \quad j \text{ why?} \\ &= C \end{aligned}$$

$\Rightarrow B = C$ And it follows that the inverse of a matrix is unique

Theorem 3.7: Properties of Inverse Matrices

Let A be an invertible matrix, $K \in \mathbb{N}$ and C is a scalar then the following are true

- (a) $(A^{-1})^{-1} = A$
- (b) $(A^K)^{-1} = A^{-1}A^{-1} \dots A^{-1} \} K \text{ factors}$
- (c) $(cA)^{-1} = c^{-1}A^{-1}; c \neq 0$
- (d) $(A^t)^{-1} = (A^{-1})^t$

Proof:

- (a) Suppose A^{-1} is the inverse of A

$$\Rightarrow A^{-1}A = I = A A^{-1}$$

$\Rightarrow$ the inverse of A^{-1} is A ; by definition of an inverse

$$\Rightarrow (A^{-1})^{-1} = A$$

- (c) Since $(cA) (c^{-1}A^{-1}) = (cc^{-1}) AA^{-1}$; by property of scalar multiplication.

$$= (1) I$$

$$= I$$

And $(c^{-1}A^{-1})(cA) = (c^{-1}c)A^{-1}A = (1)I = I$ we have $(cA)(c^{-1}A^{-1}) = I = (c^{-1}A^{-1})(cA)$

And it follows that $c^{-1}A^{-1}$ is the inverse of CA .

$$\Rightarrow (cA)^{-1} = c^{-1}A^{-1}.$$

As an exercise show property b & d .

 **Note:** For nonsingular matrices, the exponential notion used for repeated multiplication of square matrices can be extended to include exponents that are negative integers. This may be done by defining A^{-k} to be

$$A^{-k} = A^{-1} A^{-1} \cdots A^{-1} \quad (k \text{ factors})$$

With this convention we can show that the following properties.

$$(a) A^j A^k = A^{j+k}$$

$$(b) (A^j)^k = A^{jk} \quad \text{For any integers } j \text{ and } K.$$

Theorem 3.8: The inverse of a product

Let A and B be invertible matrices of order n , then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}$$

Proof:

$$\text{Since } (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}$$

$$= A(I)A^{-1}$$

$$= (AI)A^{-1}$$

$$= AA^{-1}$$

$$= I$$

And in a similar way we can show that

$$(B^{-1}A^{-1})(AB) = I \quad \text{we have } (AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB) \text{ and hence we get } B^{-1}A^{-1}$$

is the inverse of AB .

$$\Rightarrow B^{-1}A^{-1} = (AB)^{-1} \text{ and } AB \text{ is invertible}$$

Remarks

1. **Theorem 3.8** says that the inverse of a product of two invertible matrices is the product of their inverses taken in the reverse order. This can be generalized to include the product of several invertible matrices.

$$(A_1 A_2 \cdots A_n)^{-1} = A_n^{-1} A_{n-1}^{-1} \cdots A_3^{-1} A_2^{-1} A_1^{-1}$$

2. The inverse of AB is usually not equal to $A^{-1}B^{-1}$.

Theorem 3.9: Cancellation Properties

Let C be an invertible matrix, then the following properties hold.

- (a) If $AC = BC$, then $A = B$ ----- Right cancellation property
 (b) If $CA = CB$, then $A = B$ ----- Left cancellation property

Proof:

- (a) Suppose $AC = BC$, then since C is invertible we have $(AC)C^{-1} = (BC)C^{-1}$

$$\Rightarrow A(CC^{-1}) = B(CC^{-1})$$

$$\Rightarrow AI = BI$$

$$\Rightarrow A = B$$

- (b) Similar to (a)

 **Note:** If C is not invertible then cancellation is not usually valid and hence theorem 3.9 can be applied only if C is an invertible matrix.

 Test yourself!

1. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & 4 \\ 3 & 1 & 2 \end{bmatrix}$, then find a matrix B such that

- (a) $A + B$ is a diagonal matrix.
 (b) $A + B$ is a scalar matrix.
 (c) $A + B$ is a unit matrix.

2. Which of the following matrices are diagonal Matrices?

(a) $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

3. Let A and B be square matrices of order n and c is a scalar then prove that

- (a) $Tr(A + B) = Tr(A) + Tr(B)$
 (b) $Tr(AC) = cTr(A)$
 (c) $Tr(AB) = Tr(BA)$
 (d) If B is invertible, then $Tr(A) = Tr(BAB^{-1})$

4. Prove that

- (a) If the product AB is a square matrix, then the product BA is defined
- (b) If both products AB and BA are defined, then AB and BA are square matrices
- (c) If A and B are diagonal matrices of the same order, then $AB = BA$.
- (d) the product of two upper triangular matrices is upper triangular.

5. Let $D = \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \\ & & & d_n \end{bmatrix}$ be a diagonal matrix and

$A = [a_{ij}]$ be an $n \times n$ Matrix then

- (a) Compute the products DA and AD
- (b) Compute the product of two diagonal matrices
- (c) When is a diagonal matrix D invertible? If D is invertible find its inverse.

6. Show that the following matrix equation has no solution

(a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} A = I_2$; where A is a 2×2 matrix with real entries.

(b) $AB - BA = In$; where A and B are $n \times n$ matrices with real entries.

7. Prove the following

- (a) If A is an $n \times m$ matrix, AA^t and $A^t A$ are symmetries
- (b) If A is an $n \times n$ matrix, then $A - A^t$ is skew symmetric
- (c) If A and B are $n \times n$ skew symmetric matrices, then $A + B$ is skew symmetric
- (d) The main diagonal of a skew symmetric matrix consists entirely of zeros.
- (e) Let A and B be symmetric $n \times n$ matrix, then AB is symmetric if and only if.

$$AB = BA.$$

8. Find the inverse of the following matrices if it exists

(a) $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a, b, c, d \in R$.

(b) $B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 4 \\ 3 & 6 & 5 \end{bmatrix}$ (c) $C = \begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}$

(d) $D = \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$; where $a, b, c \in R$.

9. Find x such that

(a) $A = \begin{bmatrix} 3 & x \\ -2 & -3 \end{bmatrix}$ is equal to its own inverse

(b) $A = \begin{pmatrix} 4 & x \\ -2 & -3 \end{pmatrix}$ is singular.

(c) $(2x)^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

10. Prove that followings

(a) If $A^2 = A$ then either $A = I$ or A is singular

(b) If A is invertible and $AB = 0$, then $B = 0$

(c) If A, B and C are square matrices and $ABC = I$, then B is invertible and $B^{-1} = CA$.

(d) If A is nilpotent then $I + A$ is invertible. (Note: A square matrix A is called nilpotent if $A^k = 0$ for some $k > 0$)

(e) Let A, B be $m \times n$ and $n \times m$ matrices. Prove that $I_n - AB$ is invertible if and only if $I_m - BA$ is invertible.

(f) A matrix which has a row of zeros is not invertible.

3.4. Elementary row and column operations

Objectives:

At the end of this section students should able to:

- State the elementary row /column operations.
- Define Elementary matrix and identity elementary matrices.
- Use elementary matrices to find inverses.

1. Elementary Row Operations

Consider the matrices $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & -2 \\ 3 & 1 & -1 \end{bmatrix}$,

$$D = \begin{bmatrix} -2 & 4 & -2 \\ 2 & 1 & -2 \\ 3 & 1 & -1 \end{bmatrix} \quad E = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 7 & -4 \\ 2 & 1 & -2 \end{bmatrix}$$

Observe that:

I) B is obtained from A by interchanging the first and second rows of A .

II) D is obtained from C by multiplying the first row of C by -2 .

III) F is obtained from matrix E by replacing the second row of E by a new row.

This is made up by multiplying the first row by -3 and adding to the second row of E .

Such operations on rows of a matrix as described above in (I), (II) and (III) are called elementary row operations. These operations will be very useful in finding inverse of a matrix and solving systems of linear equations.

Similarly, elementary column operations can be defined. Hence we have the following definition.

Definition 3.12: An elementary operation on a matrix is either elementary row operation or elementary column operation and is of the following three types.

Type I: The interchange of any two rows or (columns)

Type II: The multiplication of any row or (column) by a non zero number

Type III: The addition of a multiple of one row (or column) to another row (or column)

 **Notations:** We shall use the following notations for the three types of elementary operations.

1. The interchange of i^{th} and j^{th} rows (columns) is denoted by $R_i \leftrightarrow R_j$ ($C_i \leftrightarrow C_j$)
2. The multiplication of the i^{th} row (column) by a none zero number k is denoted by $R_i \rightarrow kR_i$ ($C_i \rightarrow kC_i$)
3. The addition of k times the j^{th} row (column) to i^{th} row (column) is denoted by $R_i \rightarrow R_i + kR_j$ ($C_i \rightarrow C_i + kC_j$)

 **Note:**

- a) $R_i \rightarrow kR_i$ means replace the i^{th} row by k times the i^{th} row
- b) $R_i \rightarrow R_i + kR_j$ Means replace the i^{th} row by the sum of itself and k times the j^{th} row.
- c) Although elementary operations are simple to perform, they involve a lot of arithmetic. Because it is easy to make a mistake, we suggest that you get in

the habit of noting the elementary operation performed in each step so that you can go back to check your work

Definition 3.13: Row Equivalent and column Equivalent Matrices

Two matrices are said to be row-equivalent or (column equivalent) if one can be obtained from the other by a finite sequence of elementary row or (column) operations. That is,

a) $A = (a_{ij})_{m \times n}$ is row equivalent to $B = (b_{ij})_{m \times n}$, denoted by $A \overset{\text{row}}{\equiv} B$, if B is attainable from A by successive operations of finitely many elementary row operations on A.

b) $A = (a_{ij})_{m \times n}$ is column equivalent to $B = (b_{ij})_{m \times n}$, denoted by $A \overset{\text{column}}{\equiv} B$, if B is attainable from A by successive operations of finitely many elementary column operations on A.

Example-1

Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ -2 & 3 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$ then show that

(a) $A \overset{\text{row}}{\equiv} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ (b) $B \overset{\text{row}}{\equiv} I_3$

Solution:

(a)
$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ -2 & 3 & -2 \end{bmatrix} \xrightarrow{R_2 \rightarrow -3R_1 + R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ -2 & 3 & -2 \end{bmatrix}$$

$$\xrightarrow{R_3 \rightarrow 2R_1 + R_3} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ 0 & 7 & -2 \end{bmatrix} \xrightarrow{R_3 \rightarrow R_2 + R_3} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -7 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned}
 \text{(b) } B = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} &\xrightarrow{R_2 \rightarrow R_2 + R_3} \begin{bmatrix} 1 & 3 & 3 \\ 0 & 1 & 0 \\ 1 & 3 & 4 \end{bmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_1} \begin{bmatrix} 1 & 3 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
 &\xrightarrow{R_1 \rightarrow R_1 - 3R_2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \rightarrow R_1 - 3R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

Class Activity:

1) Let $M_{m \times n}$ be the set of all $m \times n$ matrices then show that the relation row equivalent on $M_{m \times n}$ is an equivalence relation.

2) Let $A \equiv B$ then show that $A^{\text{row}} \equiv B^{\text{column}}$

3) Let $A = \begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ then show that $A \equiv B$.

2. Elementary Matrices

Definition 3.14: Definition of Elementary Matrix

Let A be an $n \times n$ matrix such that A can be Obtained from I_n by a single elementary row /column operation then A is called an elementary matrix.

Example-2: Elementary Matrices and Non elementary Matrices

Which of the following matrices are elementary? For those that are describe the corresponding elementary row operation.

$$\begin{aligned}
 \text{(a) } A = I_n \quad \text{(b) } B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{(c) } C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
 \text{(d) } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{(e) } E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{(f) } F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{(g) } G = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}
 \end{aligned}$$

Solution:

(a) Since $A = I_n$ can be obtained from itself by multiplying any one of its rows by 1.

(i.e. $I_n \xrightarrow{R_1 \leftarrow 1R_1} A$) we've I_n is elementary matrix.

(b) Since B can be obtained from I_3 by multiplying the second row of

I_3 by -4 we have $I_3 \xrightarrow{R_2 \rightarrow -4R_2} B$ and hence B is an elementary matrix.

(c) Since C is obtained by multiplying the fourth row of I_4 by 0 we have C is not elementary matrix. Note that, row multiplication must be a non zero constant.

(d) D is not elementary because it is not square matrix.

(e) Since $I_3 \xrightarrow{R_2 \leftrightarrow R_3} E$ we've E is elementary.

(f) This matrix is not elementary because two elementary row operations are required to obtain it from I_3 .

(g) Since $I_2 \xrightarrow{R_2 \rightarrow 4R_1 + R_2} G$ we have G is elementary.

 **Note** Elementary Matrices are useful because they enable us to use matrix multiplication to perform elementary row/ column operations.

Example-3: Elementary Matrices and Elementary Row Operations

$$\text{Let } E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Let } A_1 = \begin{bmatrix} 0 & 2 & 1 \\ 1 & -3 & 6 \\ 3 & 2 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 2 & 6 & -4 \\ 0 & 1 & 3 & 1 \end{bmatrix} \text{ and } A_3 = \begin{bmatrix} 1 & 0 & -1 \\ -2 & -2 & 3 \\ 0 & 4 & 5 \end{bmatrix}$$

Then

$$(a) E_1 A_1 = \begin{bmatrix} 1 & -3 & 6 \\ 0 & 2 & 1 \\ 3 & 2 & -1 \end{bmatrix}, \text{ this product matrix is a matrix obtained by interchanging the}$$

first two rows of A. And E_1 is an elementary matrix in which the first two rows of I_3 has been interchanged.

$$(b) E_2 A_2 = \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & 3 & -2 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

Here, E_2 is an elementary matrix in which the second row of I_3 has been multiplied by

$$\frac{1}{2} \text{ and } A \xrightarrow{R_2 \rightarrow \frac{1}{2} R_2} E_2 A_2$$

$$(c) E_3 A_3 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 1 \\ 0 & 4 & 5 \end{bmatrix} \quad \text{Here, } \begin{matrix} I_3 \xrightarrow{R_2 \rightarrow 2R_1 + R_2} E_3 \text{ and} \\ A_3 \xrightarrow{R_2 \rightarrow 2R_1 + R_2} E_3 A_3 \end{matrix}$$

Theorem 3.10: Representing Elementary Operations

Let A be an $m \times n$ matrix

(a) Let E be the elementary matrix obtained by performing an elementary row operation on I_m such that the same elementary row operation is performed on A , then the resulting matrix is given by the product EA .

(b) Let E be the elementary matrix obtained by performing an elementary column operation on I_n such that the same elementary column operation is performed on A , then the resulting matrix is given by AE .

Proof:

To prove the above theorem we use the following fact:

Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$, A_i be i^{th} row of A and A^j be the j^{th} column of B

then one can show that

I) $(AB)_i = A_i \cdot B =$ which is the product of the i^{th} row of A and B .

II) $(AB)^j = A \cdot B^j =$ which is the product of A and the j^{th} column of B .

(a) 1- Let E be an elementary matrix obtained from I_m by applying elementary

operation $R_i \leftrightarrow R_j$ then $(EA)_i = E_i \cdot A = (I_m)_j \cdot A = (I_m A)_j = A_j = j^{\text{th}}$ row of A ,

$(EA)_j = E_j \cdot A = (I_m)_i \cdot A = (I_m A)_i = A_i = i^{\text{th}}$ row of A and for $k \neq ij$ we have

$(EA)_k = E_k A = (I_m)_k A = (I_m A)_k = A_k$ Thus,

$(EA)_i = A_j, (EA)_j = A_i$ and $(EA)_k = A_k \quad \forall k \neq i, j$

$\Rightarrow EA$ is a matrix obtained from A by applying the elementary row operation

$$R_i \leftrightarrow R_j$$

2- Let E be an elementary matrix obtained from I_m by applying elementary operation $R_i \rightarrow cR_j$ where c is a non zero scalar and suppose that we apply

$R_i \rightarrow cR_j$ on A then

$$(EA)_i = E_i \cdot A = [C(I_m)_i] A = C[(I_m)_i A] = C(I_m \cdot A)_i = CA_i \text{ and for } j \neq i$$

$$(EA)_j = E_j \cdot A = A = (I_m)_j \cdot A = (I_m \cdot A)_j = A_j$$

$$\Rightarrow (EA)_i = cA_i \text{ and } (EA)_j = A_j \quad \forall_i \neq j$$

$\Rightarrow EA$ is a matrix obtained from A by applying elementary row operation

$$R_i \rightarrow cR_i$$

3-Let E be an elementary matrix obtained from I_m by applying the elementary operation $R_i \rightarrow cR_j + R_i$ where c is a scalar and suppose that we apply

$R_i \rightarrow cR_j + R_i$ on A then

$$\begin{aligned} (EA)_i &= E_i \cdot A = [c(I_m)_j + (I_m)_i] A \\ &= [c(I_m)_j] \cdot A + (I_m)_i \cdot A \\ &= C[(I_m)_j \cdot A] + (I_m \cdot A)_i \\ &= C(I_m \cdot A)_j + A_i \\ &= C A_j + A_i \end{aligned}$$

And $\forall_j \neq i$ we have $(EA)_j = E_j \cdot A = (I_m)_j \cdot A = (I_m \cdot A)_j = A_j$

Thus, $(EA)_i = CA_j + A_i$ and $(EA)_j = A_j \quad \forall_j \neq i$.

Therefore, EA is a matrix obtained from A by applying the elementary row operation

$$R_i \rightarrow CR_j + R_i.$$

Hence, by (1) (2) and (3) we have the proof.

(b) Similar to (a).

Remarks

Be sure to remember that,

- 1) In theorem 3.10 (a), A is multiplied on the left by the elementary matrix E and
- 2) In theorem 3.10 (b), A is multiplied on the right by the elementary matrix E , which involves column operations.

Corollary 3.1

Let A and B be an $m \times n$ matrices

(a) If $A \overset{\text{row}}{\equiv} B$, Then $B = PA$, where P is a product of elementary matrices, which involves elementary row operations.

(b) If $A \overset{\text{column}}{\equiv} B$, then $B = AQ$, where Q is a product of elementary column operations.

Proof: Apply theorem 3.10 repeatedly.

Example-4: Let

$$A = \begin{bmatrix} 0 & 1 & 3 & 5 \\ 1 & -3 & 0 & 2 \\ 2 & -6 & 2 & 0 \end{bmatrix} \text{ And } B = \begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & -2 \end{bmatrix},$$

Then find a sequence of elementary matrices such that

B is the product of those elementary matrices and A ,

Solution:

$$A = \begin{bmatrix} 0 & 1 & 3 & 5 \\ 1 & -3 & 0 & 2 \\ 2 & -6 & 2 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 2 & -6 & 2 & 0 \end{bmatrix} = A_1$$

$$\xrightarrow{R_3 \rightarrow -2R_1 + R_3} \begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & -4 \end{bmatrix} = A_2 \xrightarrow{R_3 \rightarrow \frac{1}{2}R_3} \begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & -2 \end{bmatrix} = B$$

Observe that

$A_1 = E_1 A$; Where E_1 is an elementary matrix obtained from I_3 by applying $R_1 \leftrightarrow R_2$

$A_2 = E_2 A_1$; where E_2 is An elementary matrix obtained from I_3 by applying

$$R_3 \rightarrow -2R_1 + R_3$$

$B = E_3 A_2$; whhere E_3 is An elementary matrix obtained from I_3 by applying

$$R_3 \rightarrow \frac{1}{2}R_3$$

Thus $B = E_3 A_2$

$$= E_3 E_2 A_1; \text{ Since } A_2 = E_2 A_1$$

$$= E_3 E_2 E_1 A; \text{ Since } A_1 = E_1 A$$

$$\text{Where } E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \text{ and } E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} \end{bmatrix}$$

Definition 3.15: Let A, B be an $m \times n$ matrices. We say that A is equivalent to B , denoted by $A \equiv B$, if B can be attainable from A by successive application of finitely many elementary row or column operations.

Theorem 3.11

Let A, B be a $m \times n$ matrices then $A = B$ if and only if $B = PAQ$; where P is product of an $m \times m$ elementary matrices which involves row operations and Q is product of an $n \times n$ elementary matrices which involves column operations.

Proof Exercise

Theorem 3.12: (Elementary matrices are Invertible.) If E is an elementary matrix, then E^{-1} exists and is an elementary matrix.

Proof: Since elementary operations are reversible then to find the inverse of an elementary matrix E , we simply reverse the elementary row operation used to obtain E . If E is an elementary matrix such that

- (a) $I_n \xrightarrow{R_i \leftrightarrow R_j} E$ then put $E^{-1} = E$
- (b) $I_n \xrightarrow{R_i \rightarrow cR_i} E$ then put $E^{-1} = E_1$ where E_1 is an elementary matrix such that $I_n \xrightarrow{R_i \rightarrow \frac{1}{c}R_i} E_1$
- (c) $I_n \xrightarrow{R_i \rightarrow R_i + cR_j} E$ then put $E^{-1} = E_2$ where E_2 is an elementary matrix such that $I_n \xrightarrow{R_i \rightarrow R_i - cR_j} E_2$

And to check that $EE^{-1} = I_n = E^{-1}E$, apply theorem 3.10

Example-5:

$$\text{Let } E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \text{ and } E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \text{ then}$$

$$E_1^{-1} = E_1, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \quad \text{and} \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Test yourself!

1. Let $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -6 & 2 & 3 \end{bmatrix}$ then show that $A \equiv I_3^{\text{row}}$

2. Let $B = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{bmatrix}$ then show that $B \equiv I_4^{\text{row}}$.

3. Determine whether the matrix is elementary. If it is, state the elementary row operation that was used to produce it.

(a) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -3 & 1 \end{bmatrix}$

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

- Find an elementary matrix E such that $EA = B$
 - Find an elementary matrix E such that $EA = C$
 - Find an elementary matrix E such that $EB = A$
 - Find an elementary matrix E such that $EC = A$
5. Let E be an elementary matrix obtained by interchanging two rows in I_n and A be an $n \times n$ matrix then
- How will EA compare with A
 - Find E^2
6. Let E be an elementary matrix obtained by multiplying a row in I_n by a non zero constant c and A be an $n \times n$ matrix then
- How will EA compare with A ?
 - Find E^2

3.5. Row reduced echelon form of a Matrix

Objectives

At the end of this section learners are expected to

- To define row-echelon form of a matrix.
- Identify a row-echelon form of matrix.
- Solve system of linear equation using row-echelon form of a matrix.
- Compute the inverse of a square matrix.

1. Row-Echelon Matrix

Definition 3.16: Definition of Row-Echelon Form of a Matrix

A matrix is said to be in row-echelon form if

1. All the non-zero rows, if any, precede all zero rows, if any. That is, all rows Consisting entirely of zeros occur at the bottom of the matrix.
2. In any non-zero row after the first row, the number of zeros preceding the first non-zero elements is greater than the number of such zeros in the preceding row. That is, the first non-zero entry of row $i+1$ is to the right of the first non-zero entry of row i .
3. The first non-zero entry in each non-zero row is 1. This entry is called a pivot or a leading 1.

Example-1: Determine whether the given matrix is in row-echelon form or not.

$$(a) \begin{bmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

$$(b) \begin{bmatrix} 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

$$(d) \begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -4 \end{bmatrix}$$

Solution: The matrices shown in parts (a) and (b) are in row-echelon form, while the matrices shown in parts (c) and (d) are not in row-echelon form.

☞ **Note:** One can show that every matrix is row equivalent to a matrix in row-echelon form.

Definition 3.17: Definition of Reduced Row-Echelon Form of a matrix.

A matrix M in row-echelon form is in reduced row-echelon form if the first non-zero element in each non-zero row is the only non-zero element in its column.

That is, M is in reduced row-echelon form if M is in row-echelon form and if every column that has a leading 1 has zeros in every position above and below its leading 1.

☞ **Note that:** It can be shown that the reduced row echelon matrix obtained from a given matrix A by elementary operations is unique. That is, that it does not depend on the particular sequence of operations used.

Example-2: The following matrices are in reduced row-echelon

$$(a) \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem 3.13: Let A be an $m \times n$ matrix, then A is row equivalent to an $m \times n$ reduced row-echelon form of a matrix.

Proof: Exercise

Example- 3

Let $A = \begin{bmatrix} 0 & 2 & -1 & 4 \\ 3 & 2 & 0 & 2 \\ 3 & 3 & 3 & 4 \end{bmatrix}$ then find the unique reduced row-echelon matrix that is row

equivalent to the matrix A .

$$\begin{aligned} A &\xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 3 & 2 & 0 & 2 \\ 0 & 2 & -1 & 4 \\ 3 & 3 & 3 & 4 \end{bmatrix} \xrightarrow{R_1 \rightarrow \frac{1}{3}R_1} \begin{bmatrix} 1 & \frac{2}{3} & 0 & \frac{2}{3} \\ 0 & 2 & -1 & 4 \\ 3 & 3 & 3 & 4 \end{bmatrix} \\ &\xrightarrow{R_3 \rightarrow -3R_1 + R_3} \begin{bmatrix} 1 & \frac{2}{3} & 0 & \frac{2}{3} \\ 0 & 2 & -1 & 4 \\ 0 & 1 & 3 & 2 \end{bmatrix} \xrightarrow{R_2 \rightarrow \frac{1}{2}R_2} \begin{bmatrix} 1 & \frac{2}{3} & 0 & \frac{2}{3} \\ 0 & 1 & -\frac{1}{2} & 2 \\ 0 & 1 & 3 & 2 \end{bmatrix} \end{aligned}$$

$$\xrightarrow{R_1 \rightarrow \frac{-2}{3}R_2 + R_1} \begin{bmatrix} 1 & 0 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & \frac{-1}{2} & 2 \\ 0 & 0 & \frac{7}{2} & 0 \end{bmatrix} \xrightarrow{R_3 \rightarrow \frac{2}{7}R_3} \begin{bmatrix} 1 & 0 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & \frac{-1}{2} & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$R_3 \rightarrow -R_2 + R_3$$

$$\xrightarrow{R_1 \rightarrow \frac{-1}{3}R_3 + R_1} \begin{bmatrix} 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$R_2 \rightarrow \frac{1}{2}R_3 + R_2$$

Thus the reduced row-echelon matrix which is equivalent to A is $\begin{bmatrix} 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

Class Activity:

Reduce the following Matrices in to reduced row-echelon matrix.

$$A = \begin{bmatrix} 4 & 7 & -14 & 10 \\ 2 & 3 & -4 & -4 \\ 1 & 1 & 1 & 6 \end{bmatrix} \quad (b) \quad B = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ -2 & 4 & 5 \end{bmatrix}$$

2. Elementary Matrices and Non singularity

Recall that an elementary matrix is nonsingular as the inverse of the elementary operation

$$R_i \leftrightarrow R_j, R_i \rightarrow CR_i \text{ and } R_i \rightarrow CR_j + R_i \text{ are } R_i \leftrightarrow R_j, R_i \rightarrow \frac{1}{c}R_i \text{ and}$$

$$R_i \rightarrow R_i - CR_j \text{ respectively.}$$

Theorem 3.14: Let A be an $n \times n$ matrix then A is row-equivalent to I_n if and only if A is nonsingular

Proof ($\Rightarrow$) suppose $A \stackrel{\text{row}}{\equiv} I_n$ then by corollary 3.1 we get $I_n = PA$; where P is a product of finite number of elementary matrices.

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

$$\Rightarrow A \text{ is invertible or non singular.}$$

($\Leftarrow$) Suppose A is an invertible then by theorem 3.13 we have $A \equiv B$; where B is a reduced row-echelon matrix.

Claim $B = I_n$

Let $B = (b_{ij})_{n \times n}$

- If $b_{ii} \neq 0 \quad \forall_i = 1, 2, \dots, n$ then the structure of a matrix of reduced echelon form guarantees that $B = I_n$.
- If $b_{ii} = 0$ for some i and i is the first such element on the diagonal of B

Case-1 $i = n$ then the n^{th} row is a row of zeros

Case-2 $i < n$

Consider the $(i+1)^{\text{th}}$ row B

It has at least $i+1$ zeros before the first non-zero entry then one of the rows after the i^{th} row must be a row of zeros, Hence, in both cases, we have at least one row of zeros in B . But B is invertible and hence B^{-1} exists. Thus $BB^{-1} = I_n$ has at least one row of zeros; which is contradiction.

Therefore, $B = I_n$.

Corollary 3.2: A property of Invertible Matrices

A square matrix A of order n is invertible if and only if it can be written as a product of elementary matrices.

Proof ($\Leftarrow$) Suppose A is a product of elementary matrices, Let

$A = E_1 E_2 \dots E_k$, where E_i is an elementary matrix $\forall_i = 1, 2, \dots, k$ then since each elementary matrix is invertible we get

$A^{-1} = (E_1 E_2 \dots E_k)^{-1} = E_k^{-1} E_{k-1}^{-1} \dots E_2^{-1} E_1^{-1}$; As the product of invertible matrices is invertible

$\Rightarrow A^{-1}$ Exists and hence A is invertible.

($\Rightarrow$) Suppose A is invertible then by theorem 3.14 we get A is row equivalent to I_n .

But $A \equiv I_n \Rightarrow I_n = E_k E_{k-1} \dots E_3 E_2 E_1 A$ where E_i is an elementary matrix $\forall_i = 1, 2, \dots, k$

$$\Rightarrow A = (E_1^{-1} E_2^{-1} E_3^{-1} \dots E_k^{-1}) I_n$$

$$\Rightarrow A = E_1^{-1} E_2^{-1} E_3^{-1} \dots E_k^{-1}$$

$\Rightarrow A$ Can be expressed as the product of elementary matrices, as each E_i^{-1} is elementary matrix

Hence, the proof is completed.

Example-4: Writing a matrix as the Product of elementary matrices

Find a sequence of elementary matrices whose product is $A = \begin{bmatrix} -1 & -2 \\ 3 & 8 \end{bmatrix}$

Solution:

We begin by finding a sequence of elementary row operations that can be used to rewrite A in reduced row-echelon form.

<u>Matrix</u>	<u>Elementary Row Operations</u>	<u>Elementary Matrix</u>
$A = \begin{bmatrix} -1 & -2 \\ 3 & 8 \end{bmatrix}$		
$A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}$	$\xleftarrow{R_1 \rightarrow -R_1}$	$E_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
$A_2 = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix}$	$\xleftarrow{R_2 \rightarrow R_2 + -3R_1}$	$E_2 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$
$A_3 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$	$\xleftarrow{R_2 \rightarrow \frac{1}{2}R_2}$	$E_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
$I_2 = A_4 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$	$\xleftarrow{R_1 \rightarrow -2R_1 + R_1}$	$E_4 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

Now from the matrix product we get

$$\Rightarrow I_2 = E_4 A_3 = E_4 E_3 A_2 = E_4 E_3 E_2 A_1 = E_4 E_3 E_2 E_1 A$$

$$\Rightarrow A = E_1^{-1} E_2^{-1} E_3^{-1} \dots E_4^{-1}$$

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

$$= \begin{bmatrix} -1 & -2 \\ 3 & 8 \end{bmatrix}$$

Corollary 3.3: Let A be an $n \times n$ matrix. If A can be reduced to I_n by a sequence of elementary row operations, then the same sequence of elementary row operations performed on I_n produces A^{-1} .

Proof:

Suppose $A \xrightarrow{\text{row}} I_n$ then A^{-1} exists and $I_n = E_k \cdots E_3 E_2 E_1 A$ where E_i is an elementary matrix $\forall_i = 1, 2, \dots, K$.

But $I_n = E_k \cdots E_3 E_2 E_1 A$

$$\Rightarrow I_n A^{-1} = (E_k \cdots E_3 E_2 E_1 A) A^{-1}$$

$$\Rightarrow A^{-1} = (E_k \cdots E_3 E_2 E_1) A A^{-1}$$

$$\Rightarrow A^{-1} = (E_k \cdots E_3 E_2 E_1) I_n$$

$$\Rightarrow A^{-1} = E_k \cdots E_3 E_2 E_1$$

In other words, a sequence of elementary matrices that reduces A to the identity also can be used to reduce the identity I to A^{-1} . Applying the corresponding sequence of elementary row operations to the matrices A and I simultaneously, we have

$E_k \cdots E_3 E_2 E_1 [A|I] = [I|A^{-1}]$. Of course, if A is singular, then no such sequence can be found.

 **Remark: Finding the Inverse of a Matrix by Gauss-Jordan Elimination**

Let A be a square matrix of order n

1. Write the $n \times 2n$ matrix that consists of the given matrix A on the left and the $n \times n$ identity matrix I on the right to obtain $[A|I]$. Note that we separate the matrices A and I by a line. We call this process adjoining the matrices A and I .
2. If possible, row reduce A to I using elementary row operations on the entire matrix $[A|I]$. The result will be the matrix $[I|A^{-1}]$. If this is not possible, then A is not invertible.
3. Check your work by multiplying to see that $AA^{-1} = I = A^{-1}A$.

Example-5: If

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -3 \\ -6 & 2 & 3 \end{bmatrix} \text{ then find its inverse}$$

Solution

We begin by adjoining the identity matrix to A to form the matrix

$$\begin{aligned} [A|I] &= \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ -6 & 2 & 3 & 0 & 0 & 1 \end{array} \right] \\ \xrightarrow{R \rightarrow -R_1 + R_2} & \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & -4 & 3 & 6 & 0 & 1 \end{array} \right] \\ \xrightarrow{R_3 \rightarrow 6R_1 + R_3} & \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & -4 & 3 & 6 & 0 & 1 \end{array} \right] \\ \xrightarrow{R_3 \rightarrow 4R_2 + R_3} & \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 2 & 4 & 1 \end{array} \right] \\ \xrightarrow{R_3 \rightarrow -R_3} & \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & -4 & -1 \end{array} \right] \\ \xrightarrow{R_3 \rightarrow R_3 + R_2} & \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & -3 & -1 \\ 0 & 0 & 1 & -2 & -4 & -1 \end{array} \right] \\ \xrightarrow{R_1 \rightarrow R_2 + R_1} & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -3 & -1 \\ 0 & 1 & 0 & -3 & -3 & -1 \\ 0 & 0 & 1 & -2 & -4 & -1 \end{array} \right] = [I|A^{-1}] \end{aligned}$$

Therefore, A is invertible and its inverse is $A^{-1} = \begin{bmatrix} -2 & -3 & -1 \\ -3 & -3 & -1 \\ -2 & -4 & -1 \end{bmatrix}$

Class Activity

1. Find the inverse of the following matrix by using Gauss-Jordan elimination

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} \quad (b) \begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & 1 & 2 & 1 \\ 2 & 3 & 4 & 1 \\ 3 & 3 & 3 & 1 \\ 1 & 2 & 3 & 1 \end{bmatrix}$$

2. Show that the following matrices have no inverses

$$(a) \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ -2 & 3 & -2 \end{bmatrix} \quad (b) \begin{bmatrix} 3 & -1 \\ -6 & 2 \end{bmatrix}$$

 **Note:** - Using Gauss-Jordan elimination to find the inverse of a matrix works well (even as a computer technique) for matrices of order 3x3 or greater. For 2x2 matrices, however, many people prefer to use a formula for the inverse, rather than find the inverse by Gauss-Jordan elimination. The formula is given by:

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \text{ where } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ with } ad-bc \neq 0.$$

Thus A is invertible if and only if $ad-bc \neq 0$. Why?

Test yourself!

1. Determine whether the given matrix is in row-echelon form. If it is, determine whether it is also in reduced row echelon form.

$$(a) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix} \quad (b) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(d) \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (e) \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (f) \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

2. Find a unique reduced row-echelon matrix that is row equivalent to the given matrix

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

3. Describe all possible

(a) 2×2 Reduced row-echelon matrices

(b) 3×3 Reduced row-echelon matrices

4. Determine conditions on a, b, c , and d such that the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ will be row equivalent to the given matrix

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

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

5. Factor the matrix A into a product of elementary matrices

(a) $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} 4 & -1 \\ 3 & -1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 2 & 0 & -8 \\ 2 & 2 & 2 \\ 0 & 2 & 7 \end{bmatrix}$

6. Use elementary matrices to find the inverse of

(a) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{bmatrix}, c \neq 0$ (b) $A = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 \\ b & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}, c \neq 0$

7. Find the inverse of the following matrices if it exists.

(a) $\begin{bmatrix} \frac{1}{a} & 0 \\ a & a \end{bmatrix}; a \neq 0$

(b) $\begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 0 \\ -2 & 0 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 5 & 5 \end{bmatrix}$

7. Show that $A = \begin{bmatrix} \sin Q & \cos Q \\ -\sin Q & \sin Q \end{bmatrix}$ is invertible

3.6. Rank of a Matrix Elementary Row/Column Operation

Objectives

- Find the rank of matrix.

If A is an $m \times n$ matrix, then the subspace of $\mathfrak{R}^n$ spanned by the row vectors of A is called the **row space** of A , and the subspace of $\mathfrak{R}^m$ spanned by the column vectors of A is called the **column space** of A . The solution space of the homogeneous system of equations $Ax = 0$, which is a subspace of $\mathfrak{R}^n$, is called the **nullspace** of A . If s of the n columns of an $m \times n$ matrix A are linearly independent, the number s is called the **column rank** of matrix A . When the s nonzero columns $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s$ are regarded as m element column vectors belonging to a vector space R^m , these vectors will span a subspace of R^m .

 **Note**-Let A be any matrix. Then the row rank and column rank of A are equal.

Definition: The **rank** of matrix A , denoted by $\text{rank}(A)$, is the value common to the row and column ranks of A .

 **Note**- Let A be any matrix. Then $\text{rank}(A) = \text{rank}(A^t)$.

Example:- Let

$$A = \begin{bmatrix} 1 & 0 & 3 & 0 & 4 & 0 \\ 2 & 1 & 7 & 0 & 10 & 1 \\ 1 & 0 & 3 & 2 & 6 & 4 \\ 1 & 0 & 3 & 0 & 4 & 0 \end{bmatrix}$$

Then the row-reduced echelon form of A is

$$B = \begin{bmatrix} 1 & 0 & 3 & 0 & 4 & 0 \\ 0 & 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

showing that the number of leading columns is 3, so the row rank of A is 3, and hence its *rank* is 3.

 **Test yourself!**

Find the row-reduced echelon form of the given matrix, its rank, a basis for its row space, and a basis for its column space.

$$1. \begin{bmatrix} 1 & 3 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 4 & 1 & 3 \end{bmatrix}, \quad 2. \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 1 & 0 & 4 & 5 \\ 0 & 1 & 2 & 4 \end{bmatrix}, \quad 3. \begin{bmatrix} 3 & 0 & 2 & 6 & 0 \\ 4 & 1 & 0 & 11 & 3 \\ 2 & 0 & 2 & 4 & 0 \\ 3 & 0 & 0 & 6 & 3 \end{bmatrix}$$

3.7. System of linear Equations.

Objectives

At the end of this section learners should be able to

- Determine the coefficient matrix and the augmented matrix of linear system.
- Use Gaussian elimination with back-substitution to solve linear system
- Compute linear system using Gauss-Jordan elimination method.

Recall that

1. A linear equation in n variables $x_1, x_2, \dots, x_n$ has the form

$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$ Where the coefficients $a_1, a_2, \dots, a_n$ are real numbers and the constant term b is a real number.

The number a_1 is the leading coefficient and x_1 is the leading variable.

2. A solution of a linear equation in n variables $a_1x_1 + a_2x_2 + \dots + a_nx_n$, is a sequence of n -real numbers $t_1, t_2, \dots, t_n$ arranged so that the equation is satisfied when the values $x_1 = t_1, x_2 = t_2, x_3 = t_3, \dots, x_n = t_n$ are substituted in to the equation.

Definition 3.18: systems of Linear equations

A system of m linear equations in n variables is a set of m -equations, each of which is linear in the same n -variables:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n = b_3$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$a_{m1} x_1 + a_{m2} x_2 + a_{m3} x_3 + \dots + a_{mn} x_n = b_m$$

 Remark:

The double subscript notation indicates that a_{ij} is the coefficient of x_j in the i^{th} equation.

Definition 3.19:

A solution of the above system of linear equations is a sequence of numbers $t_1, t_2, t_3, \dots, t_n$ that is a solution of each of the linear equations in the system.

It can happen that a system of linear equations has exactly one solution, an infinite number of solutions, or no solution. A system of linear equations is called consistent if it has at least one solution and inconsistent if it has no solution.

 Note: Number of solutions of a system of linear equations

For a system of linear equations in n-variables, precisely one of the following is true.

- (a) The system has exactly one solution (consistent system)
- (b) The system has an infinite number of solutions (consistent system)
- (c) The system has no solution (inconsistent system)

 Notation: Matrix Notation of system of linear equations

One very common use of matrices is to represent a system of linear equations. Matrix notation was introduced in the nineteenth century to provide a short-hand way of writing linear equations. The system of m linear equations in n variables

$$a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1$$

$$a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = b_2$$

$$\vdots \quad \quad \quad \vdots$$

$$a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n = b_m$$

Can be written in matrix notation as $AX = B$

$$\text{Where } A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \text{ And } AX \text{ is the matrix product}$$

Here, in matrix notation $AX = B$, we have the following definitions.

- (1) $A = (a_{ij})_{m \times n}$ is called the coefficient matrix of the system.
- (2) A and B are column vectors in R^n and R^m respectively. x is unknown and B is a given column vector.
- (3) The matrix derived from the coefficients and constant terms of a system of linear m equations is called the augmented matrix of the system. That is, the matrix $[A|B]$ is called the augmented matrix of the system.

Example-1:

Use matrix notation to denote the following system of linear equations. Also find the coefficient matrix and the augmented matrix of the system.

$$\begin{aligned} -x_2 + 2x_3 &= 2 \\ 3x_1 + 4x_2 - 6x_3 &= 1 \end{aligned}$$

Solution:

$$\begin{cases} -x_2 + 2x_3 = 2 \\ 3x_1 + 4x_2 - 6x_3 = 1 \end{cases}$$

$$\Rightarrow \begin{cases} 0x_1 - 1x_2 + 2x_3 = 2 \\ 3x_1 + 4x_2 - 6x_3 = 1 \end{cases}$$

$$\Rightarrow \begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\Rightarrow (i) \ Ax = B \text{ where } A = \begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix}, \ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ is the matrix}$$

notation for the given system.

$\Rightarrow$ (ii) $A = \begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix}$ Is the coefficient matrix of the given system and

$\Rightarrow$ (iii) $[A|B] = \begin{bmatrix} 0 & -1 & 2 & 2 \\ 3 & 4 & -6 & 1 \end{bmatrix}$ is the associated augmented matrix of the system.

2. Gassing Elimination and Gauss-Jordan Elimination

(I) Gaussian Elimination

It is a procedure for solving a system of linear equations.

Rewriting a system of linear equations in row-echelon form usually involves a chain of equivalent systems, each of which is obtained by using one of the three basic operations. This process is called Gaussian elimination, after the German mathematician Carl Friedrich Gauss (1777-1855)

Theorem 3.15: Let $AX = B$ be represent the system of linear equations and $M = [A|B]$ be the block matrix.

If $m' = [A'|B']$ is row equivalent to

m then the solutions of $A'x = B'$ are the same as those of $Ax = B$.

Proof

Since M' is obtained by a sequence of elementary row operations, we have

$m' = E_r \dots E_2 E_1 M$ Let $P = E_r \dots E_2 E_1$. This matrix is invertible, why?

Also, $M' = [A'|B'] = [PA|PB]$, How?

If X is a solution of the original system $AX = B$, so X also solves the new system

Conversely, if $A'X = B'$, then

$AX = p^{-1}A'X = p^{-1}B = B$, so x solves the system $AX = B$ Too, Hence the proof

Gaussian Elimination with Back Substitution

The general procedure for using Gaussian elimination with back substitution to solve a system of linear equations is summarized as follows:

Step-1 Write the augmented matrix of the system

Step- 2 Use elementary row operations to rewrite the augmented matrix in

Row-echelon form

Step-3 Write the system of linear equations corresponding to the matrix in Row-echelon form, and use back-substitution to find the solution.

Example-2

Use Gauss-Jordan elimination with back-substitution to solve the following system.

$$x_2 + x_3 - 2x_4 = -3$$

$$x_1 + 2x_2 - x_3 = 2$$

$$2x_1 + 4x_2 + x_3 - 3x_4 = -2$$

$$x_1 - 4x_2 - 7x_3 - x_4 = -19$$

Solution:

Step-1: The augment matrix for this system is $M = [A|B]$ where, $M = [A|B] =$

$$\left[\begin{array}{cccc|c} 0 & 1 & 1 & -2 & -3 \\ 1 & 2 & -1 & 0 & 2 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{array} \right]$$

Step-2: using the elementary operations

$R_1 \leftrightarrow R_2, R_3 \rightarrow -2R_1 + R_3, R_4 \rightarrow -R_1 + R_4, R_4 \rightarrow 6R_2 + R_4, R_3 \rightarrow \frac{1}{3}R_3$ and $R_4 \rightarrow \frac{-1}{13}R_4$ respectively we get the row-echelon form of the matrix M,

$$\text{Which is } M' = [A'|B'] = \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right]$$

Step-3: The corresponding system of linear equations to the matrix M' is

$$x_1 + 2x_2 - x_3 = 2$$

$$x_2 + x_3 - 2x_4 = -1$$

$$x_3 - x_4 = -2$$

$$x_4 = 3$$

and using back substitution, we can determine that the solution is

$$x_1 = -1, x_2 = 2, x_3 = 1, x_4 = 3$$

(II) Gauss-Jordan Elimination

With Gaussian elimination, we apply elementary row operations to a matrix to obtain a row equivalent row-echelon form. A second method of elimination called Gauss Jordan elimination after Carl Gauss and Wilhelm Jordan (1842-1899), Continues the reduction process until a reduced row-echelon form is obtained.

Example-3: use Gauss-Jordan elimination method to solve the system

$$\begin{aligned}x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17\end{aligned}$$

Solution

Step-1: The associated augmented matrix of the system is

$$M = [A|B] = \left[\begin{array}{ccc|c} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{array} \right]$$

Step-2: You can show that the reduced row-echelon form of the matrix M is given by

$$M' = [A'|B'] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

Step-3: The corresponding system of linear equations to the matrix M' is

$$x = 1, y = -1 \text{ and } z = 2$$

Note:

- (1) If in the elimination process, you obtain a row with zeros except for the last entry, it is unnecessary to continue the elimination process. You can simply conclude that the system is inconsistent.
- (2) If in the row-echelon form system the number of equations and the number of unknowns are equal then the system will have a unique solution.
- (3) Observe that if in the row-echelon form system the number of equations is less than the number of unknown, then the system will have infinitely many solutions.

Class Activity

1. Use Gaussian elimination with back substitution to solve.

$$2x + y - 2z = 10$$

$$(a) \quad 3x + 2y + 2z = 1$$

$$5x + 4y + 3z = 4$$

$$x + 2y - 3z = 4$$

$$(c) \quad x + 3y + z = 11$$

$$2x + 5y - 4z = 13$$

$$2x + 6y + 2z = 22$$

$$x + 2y - 3z = 6$$

$$(b) \quad 2x - y + 4z = 2$$

$$4x + 3y - 2z = 14$$

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

$$(d) \quad x_1 - x_3 = 6$$

$$2x_1 - 3x_2 + 5x_3 = 4$$

$$3x_1 + 2x_2 - x_3 = 1$$

2. Use Gauss-Jordan elimination method to solve

$$(a) \quad x_1 - x_2 + 3x_3 = 0$$

$$(b) \quad 2x_1 + 4x_2 - 2x_3 = 0$$

$$2x_1 - x_2 + 3x_3 = 0$$

$$3x_1 + 5x_2 = 1$$

Definition 3.20: Homogeneous system of linear equations

A system of linear equations in which each linear equation is equal to zero is called a homogeneous system and it is denoted by $Ax = 0$.

It is easy to see that a homogeneous system must have at least one solution. Specifically, if all variables in a homogeneous system have the value zero, then each of the equations must be satisfied. Such a solution is called trivial or obvious.



Test yourself!

1. In exercise a-d, the augmented matrix of a system of equations is given. Express the system in ordinary equation notation.

$$(a) \quad \left[\begin{array}{cc|c} 2 & -3 & 1 \\ 4 & 7 & 2 \end{array} \right]$$

$$(b) \quad \left[\begin{array}{ccc|c} 2 & 3 & 5 & 2 \\ 1 & 6 & 9 & 0 \end{array} \right]$$

$$(c) \quad \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 1 & -1 & 4 & -2 & 3 \\ 4 & 2 & 5 & 0 & 2 \end{array} \right]$$

$$(d) \quad \left[\begin{array}{ccc|c} 1 & 7 & 0 & 4 \\ 2 & 3 & 1 & 6 \\ -1 & 0 & -2 & 3 \end{array} \right]$$

2. Find the solution set of the system of linear equations represented by the given augmented

matrix. (a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 2 & \end{array} \right]$

$$(b) \quad \left[\begin{array}{cccc|c} 1 & 2 & 1 & 0 & \\ 0 & 0 & 1 & -1 & \\ 0 & 0 & 0 & 0 & \end{array} \right]$$

$$(c) \quad \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 4 \\ 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right]$$

3. Using either Gaussian elimination with back-substitution or Gauss-Jordan elimination solve the given system of linear equations

$$\begin{array}{ll}
 (a) \quad \begin{array}{l} x + 3y = 0 \\ 2x - y = 0 \end{array} & (b) \quad \begin{array}{l} x + 3y - 2z = 0 \\ x - 8y + 8z = 0 \\ 3x - 2y + 4z = 0 \end{array} \\
 \\
 \begin{array}{l} 2x + y - z + 2w = -6 \\ 3x + 4y + w = 1 \end{array} & \begin{array}{l} x_1 + x_2 - 2x_3 + 3x_4 + 2x_5 = 9 \\ 3x_1 + 3x_2 - x_3 + x_4 + x_5 = 5 \end{array} \\
 (c) \quad \begin{array}{l} x + 5y + 2z + 6w = -3 \\ 5x + 2y - z - w = 3 \end{array} & (d) \quad \begin{array}{l} 2x_1 + 2x_2 - x_3 + x_4 - 2x_5 = 1 \\ 4x_1 + 4x_2 + x_3 - 3x_5 = 4 \\ 8x_1 + 5x_2 - 2x_3 - x_4 + 2x_5 = 3 \end{array}
 \end{array}$$

4. Find the values of a,b and c if possible such that the given system of linear equations has

- (I) a unique solution (II) no solution and
(III) An infinite number of solutions

$$\begin{array}{ll}
 (a) \quad \begin{array}{l} x + y = 2 \\ y + z = 2 \\ x + z = 2 \\ ax + by + cz = 0 \end{array} & (b) \quad \begin{array}{l} x + y = 0 \\ y + z = 0 \\ x + z = 0 \\ ax + by + cz = 0 \end{array}
 \end{array}$$

5. Is it possible that a system of linear equations with fewer equations than variables may have no solution? If so, give an example.
6. If A is an invertible matrix, then show that the system of linear equations represented by $Ax = B$ has a unique solution given by $x = A^{-1}B$
7. Prove that every homogeneous system of linear equations is consistent. Moreover, if the system has fewer equations than variables, then it must have an infinite number of solutions.
8. Let A be a $n \times n$ matrix, then prove that the following statement are equivalent.
- A is invertible
 - $Ax = B$ has a unique solution for every $n \times 1$ matrix B
 - $Ax = 0$ has only the trivial solution
 - A is row equivalent to I_n .
 - A can be written as the product of elementary matrices.
9. Let $m' = [A' | B']$ be a reduced row-echelon matrix, then prove that the system of equations $A'X = B'$ has solution if and only if there is no pivot in the last column B^1 .

SUMMARY OF UNIT THREE

- ❖ A **matrix** is a rectangular array of numbers. The numbers in the array are called the **entries** in the matrix.
- ❖ Two matrices are defined to be **equal** if they have the same size and their corresponding entries are equal. In matrix notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then $A = B$ if and only if $(A)_{ij} = (B)_{ij}$, or, equivalently, $a_{ij} = b_{ij}$ for all i and j .
- ❖ If A and B are matrices of the same size, then the **sum** $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A , and the **difference** $A - B$ is the matrix obtained by subtracting the entries of B from the corresponding entries of A . Matrices of different sizes cannot be added or subtracted.
- ❖ If A is any matrix and c is any scalar, then the **product** cA is the matrix obtained by multiplying each entry of the matrix A by c . The matrix cA is said to be a **scalar multiple** of A .
- ❖ If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the **product** AB is the $m \times n$ matrix whose entries are determined as follows. To find the entry in row i and column j of AB , single out row i from the matrix A and column j from the matrix B . Multiply the corresponding entries from the row and column together, and then add up the resulting products.
- ❖ If A is any $m \times n$ matrix, then the **transpose of A** , denoted by A^t , is defined to be the $n \times m$ matrix that results from interchanging the rows and columns of A ; that is, the first column of A^t is the first row of A , the second column of A^t is the second row of A , and so forth.
- ❖ If A is a square matrix, then the **trace of A** , denoted by $\text{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A . The trace of A is undefined if A is not a square matrix.
- ❖ Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.

$$(a) \ A + B = B + A \quad (\text{Commutative law for addition})$$

$$(b) \ A + (B + C) = (A + B) + C \quad (\text{Associative law for addition})$$

$$(c) \ A(BC) = (AB)C \quad (\text{Associative law for multiplication})$$

$$(d) \ A(B + C) = AB + AC \quad (\text{Left distributive law})$$

$$(e) (B + C)A = BA + CA \quad (\text{Right distributive law})$$

$$(f) A(B - C) = AB - AC$$

$$(g) (B - C)A = BA - CA$$

$$(h) a(B + C) = aB + aC$$

$$(i) a(B - C) = aB - aC$$

$$(j) (a + b)C = aC + bC$$

$$(k) (a - b)C = aC - bC$$

$$(l) a(bC) = (ab)C$$

$$(m) a(BC) = (aB)C = B(aC)$$

$$(n) A + 0 = 0 + A = A$$

❖ If R is the reduced row-echelon form of an $n \times n$ matrix A, then either R has a row of zeros or R is the identity matrix I_n .

❖ If A is a square matrix, and if a matrix B of the same size can be found such that $AB = BA = I$, then A is said to be **invertible** and B is called an **inverse** of A. If no such matrix B can be found, then A is said to be **singular**.

❖ The matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if $ad - bc \neq 0$, in which case the inverse is given by the formula

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \frac{d}{ad - bc} & -\frac{b}{ad - bc} \\ -\frac{c}{ad - bc} & \frac{a}{ad - bc} \end{bmatrix}$$

❖ If A and B are invertible matrices of the same size, then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}$$

❖ A product of any number of invertible matrices is invertible, and the inverse of the product is the product of the inverses in the reverse order.

❖ If the sizes of the matrices are such that the stated operations can be performed, then

$$(a) ((A)^T)^T = A$$

$$(b) (A + B)^T = A^T + B^T \text{ and } (A - B)^T = A^T - B^T$$

$$(c) (kA)^T = kA^T, \text{ where } k \text{ is any scalar}$$

$$(d) (AB)^T = B^T A^T$$

- ❖ The transpose of a product of any number of matrices is equal to the product of their transposes in the reverse order.
- ❖ If A is an invertible matrix, then A^t is also invertible and

$$(A^T)^{-1} = (A^{-1})^T$$
- ❖ An $n \times n$ matrix is called an **elementary matrix** if it can be obtained from the $n \times n$ identity matrix I_n by performing a single elementary row operation.
- ❖ Every elementary matrix is invertible, and the inverse is also an elementary matrix.
- ❖ If A is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false.
 - (a) A is invertible.
 - (b) $Ax = 0$ has only the trivial solution.
 - (c) The reduced row-echelon form of A is I_n .
 - (d) A is expressible as a product of elementary matrices.
- ❖ To find the inverse of an invertible matrix A , we must find a sequence of elementary row operations that reduces A to the identity and then perform this same sequence of operations on I_n to obtain A^{-1} .
- ❖ A square matrix A is called **symmetric** if $A = A^t$.
- ❖ An elementary operation on a matrix is either elementary row operation or elementary column operation and is of the following three types
 - Type I: The interchange of any two rows or (columns)
 - Type II: The multiplication of any row or (column) by a non zero number
 - Type III: The addition of a multiple of one row (or column) to another row (or column)
- ❖ A matrix is said to be in row-echelon form if
 1. All the non-zero rows, if any, precede all zero rows, if any. That is, all rows consisting entirely of zeros occur at the bottom of the matrix.
 2. In any non-zero row after the first row, the number of zeros preceding the first non-zero element is greater than the number of such zeros in the preceding

row. That is, the first non-zero entry of row $i+1$ is to the right of the first non-zero entry of row i .

3. The first non-zero entry in each non-zero row is 1. This entry is called a pivot or a leading 1.

❖ A matrix M in row-echelon form is in reduced row-echelon form if the first non-zero element in each non-zero row is the only non-zero element in its column. That is, M is in reduced row-echelon form if M is in row-echelon form and if every column that has a leading 1 has zeros in every position above and below its leading 1.

UNIT FOUR: DETERMINANT

UNIT-OBJECTIVES:

At the end of this unit each student should know;

- The definition of determinant and its properties.
- How to find the inverse of a matrix by using the concept of determinant.
- How to solve system of linear equations by using Cramer's rule.
- How to find the rank of a matrix using determinant
- How to find Eigen value and eigenvector of a matrix.
- How to Diagonalize a symmetric matrix

4.1 Definition of Determinant

Objectives

At the end of this section learners should be able to

- Find all the minors and cofactors of a square matrix of order two and three.
- Compute the determinant of triangular matrix.
- Discuss the concept of determinants of $n \times n$ matrix.

Consider a system of linear equations

$$ax + by = e$$

$$cx + dy = f$$

You recall that, one of the methods available to us to solve such a system was the elimination method. In this unit, using determinants, we shall exhibit a very efficient computational method to solve system of linear equations. We first define determinant and study its properties.

Definition 4.1: To every square matrix A with elements from the set of real numbers there is assigned a specific real number called the determinant of A . It is usually denoted by $\det(A)$ or $|A|$.

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ then determinant of } A \text{ is denoted by } \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Remarks:

- Observe that determinant is a function from the set of all square matrices to the set of real numbers, and it is defined only for square matrix. In fact, if $A = (a_{ij})$ is a square matrix over a field F ($a_{ij} \in F$) then $|A| \in F$. In particular for $F = \mathbb{C}$ = The set of complex numbers then $\det A \in \mathbb{C}$. From now on wards unless we have stated we consider matrices over R .
- The determinant of an $n \times n$ matrix is called determinant of order n . Before we go to the definition of determinant of order n ($n \in \mathbb{N}$) we shall consider determinants of order 1, 2 and 3.

You recall that a single number a is considered as a 1×1 matrix. Hence if

$A = [a_{11}]$ then the determinant of A is defined to be the number itself. That is,

$$\det(A) = |a_{11}| = a_{11}.$$

1) Determinant of order

Definition 4.2: The determinant of the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is defined to be the number $ad - bc$. That is, $\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

A convenient method for remembering the formula is to note that $ad - bc$ is just the difference of the products of diagonally opposite entries.

That is, $|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

Example-1: Find the determinants of the following matrices

(a) $A = \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix}$ (b) $B = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$ (c) $C = \begin{bmatrix} 1 & 2 \\ 6 & 4 \end{bmatrix}$

Solution: (a) $|A| = \begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix} = (2)(2) - (-3)(1) = 4 + 3 = 7$

(b) $|B| = \begin{vmatrix} 3 & 1 \\ 6 & 2 \end{vmatrix} = (3)(2) - (6)(1) = 6 - 6 = 0$

(c) $|C| = \begin{vmatrix} 1 & 2 \\ 6 & 4 \end{vmatrix} = (1)(4) - (6)(2) = 4 - 12 = -8$

 **Remark:** Note that the determinant of a matrix A can be positive, zero, or negative.

Definition 4.3: Definition of Minors and cofactors of a matrix

Let $A = (a_{ij})$ be a square matrix then

1- The determinant of the matrix obtained by deleting or removing the i^{th} row and j^{th} column of A is called the minor of the element a_{ij} and it is denoted by M_{ij} .

That is, $M_{ij} = |A_{ij}|$; where A_{ij} is a matrix in which the i^{th} row and j^{th} column of A are removed.

2-The product of the minor of the element a_{ij} and the number $(-1)^{i+j}$ is called the cofactor of a_{ij} and it is denoted by C_{ij} , That is, $C_{ij} = (-1)^{i+j} M_{ij} = (-1)^{i+j} |A_{ij}|$

Example-2: Find all the minors and cofactors of $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 4 \\ -3 & 2 & 5 \end{bmatrix}$

Solution: To find the minor M_{ij} of a_{ij} , we delete the i^{th} row j^{th} column of A and evaluate the determinant of the resulting matrix

$$\text{Minor of } a_{11} = M_{11} = |A_{11}| = \begin{vmatrix} 1 & 4 \\ 2 & 5 \end{vmatrix} = (1)(5) - (2)(4) = -3$$

$$\text{Minor of } a_{12} = M_{12} = |A_{12}| = \begin{vmatrix} 1 & 4 \\ -3 & 5 \end{vmatrix} = (1)(5) - (4)(-3) = 17,$$

$$\text{Minor of } a_{13} = M_{13} = |A_{13}| = \begin{vmatrix} 1 & 1 \\ -3 & 2 \end{vmatrix} = (1)(5) - (1)(2) - (1)(-3) = 5,$$

$$\text{Minor of } a_{21} = M_{21} = |A_{21}| = \begin{vmatrix} 1 & 0 \\ 2 & 5 \end{vmatrix} = (1)(5) - 0 = 5,$$

$$\text{Minor of } a_{22} = M_{22} = |A_{22}| = \begin{vmatrix} 2 & 0 \\ -3 & 5 \end{vmatrix} = (2)(5) - 0(-3) = 10,$$

$$\text{Minor of } a_{23} = M_{23} = |A_{23}| = \begin{vmatrix} 2 & 1 \\ -3 & 2 \end{vmatrix} = (2)(2) - (1)(-3) = 7,$$

$$\text{Minor of } a_{31} = M_{31} = |A_{31}| = \begin{vmatrix} 1 & 0 \\ 1 & 4 \end{vmatrix} = (1)(4) - (0)(1) = 4,$$

$$\text{Minor of } a_{32} = M_{32} = |A_{32}| = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} = (2)(4) - (0)(1) = 8,$$

$$\text{Minor of } a_{33} = M_{33} = |A_{33}| = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = (2)(1) - (1)(1) = 1,$$

Now, to find the cofactor C_{ij} of a_{ij} , we multiply M_{ij} with $(-1)^{i+j}$

$$C_{11} = (-1)^{1+1} M_{11} = -3 \quad C_{21} = (-1)^{2+1} M_{21} = -5 \quad C_{31} = (-1)^{3+1} M_{31} = 4$$

$$C_{12} = (-1)^{1+2} M_{12} = -17 \quad C_{22} = (-1)^{2+2} M_{22} = 10 \quad C_{32} = (-1)^{3+2} M_{32} = -8$$

$$C_{13} = (-1)^{1+3} M_{13} = 5 \quad C_{23} = (-1)^{2+3} M_{23} = -7 \quad C_{33} = (-1)^{3+3} M_{33} = 1$$

Remark

- (i). The minors and cofactors of a matrix differ at most in sign. To obtain the cofactors of a matrix, first find the minors and then apply the following checker board pattern of +’s and -’s.

Sign Pattern for Cofactors

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

3x3 matrix

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

4x4 matrix

$$\begin{bmatrix} + & - & + & - & + & - & - \\ - & + & - & + & - & - & - \\ + & - & + & - & + & - & - \\ - & + & - & + & - & - & - \\ + & - & + & - & + & - & - \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$n \times n$ matrix

Note that odd positions (where $i + j$ is odd) have negative signs, and even positions (where $i + j$ is even) have positive signs.

ii) If A is an $n \times n$ matrix, then the order of A_{ij} is $(n-1) \times (n-1)$

3) Determinant of Order Three

Definition 4.4: The determinant of the 3×3 matrix

$$\begin{aligned} A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} & \text{ is defined to be the number } |A| = a_{11} C_{11} + a_{12} C_{12} + a_{13} C_{13} \\ &= a_{11}(-1)^{1+1} M_{11} + a_{12}(-1)^{1+2} M_{12} + a_{13}(-1)^{1+3} M_{13} \\ &= a_{11} |A_{11}| + a_{12} |A_{12}| + a_{13} |A_{13}| \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31} \end{aligned}$$

Remarks: The formula used to compute determinants is known as expansion by a row (column). Thus, if we consider the first row, then the determinant of A will be equal to the sum of the products obtained by multiplying the elements of the first row by their respective cofactors.

Example-3 Let $A = \begin{bmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & 0 & 1 \end{bmatrix}$, then determine $|A|$

- (a) By expanding along the first row
- (b) By expanding along the second row
- (c) By expanding along the first column

Solution:

$$(a) C_{11} = + \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1, C_{12} = - \begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = -(-5) = 5, \text{ and } C_{13} = + \begin{vmatrix} 3 & -1 \\ 4 & 0 \end{vmatrix} = 4$$

$$\text{Thus, } |A| = a_{11} C_{11} + a_{12} C_{12} + a_{13} C_{13}, \text{ First row expansion}$$

$$= 0(-1) + 2(5) + 1(4) = 14$$

$$(b) \text{ Since } C_{21} = - \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = -2, C_{22} = + \begin{vmatrix} 0 & 1 \\ 4 & 1 \end{vmatrix} = -4,$$

$$\text{And } C_{23} = - \begin{vmatrix} 0 & 2 \\ 4 & 0 \end{vmatrix} = -(-8) = 8 \text{ we have}$$

$$|A| = a_{21} C_{21} + a_{22} C_{22} + a_{23} C_{23}, \text{ second row expansion}$$

$$= 3(-2) + (-1)(-4) + 2(8)$$

$$= 14$$

(c) Since

$$C_{11} = + \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1, C_{21} = - \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = -2 \text{ and } C_{31} = + \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} = 5 \text{ we have}$$

$$|A| = a_{11} C_{11} + a_{21} C_{21} + a_{31} C_{31}, \text{ First column expansion}$$

$$= 0(-1) + 3(-2) + 4(5) = 14$$

As an exercise try some other possibilities to see that the determinant of A can be evaluated by expanding by any row or any column. This result is stated formally in the following theorem 4.1, called Laplace's Expansion of a determinant, after the French Mathematician Pierre-Simon Laplace (1749-1827).

3-Determinant of order n

Definition 4.5: Definition of the Determinant of a Matrix

Let A be a square Matrix of order n where $n > 2$, then the determinant of A is the sum of the entries in the first row of A multiplied by their cofactors, That is,

$$|A| = \sum_{j=1}^n a_{1j} C_{1j} = a_{11} C_{11} + a_{12} C_{12} + \dots + a_{1n} C_{1n}.$$

When this definition is used to evaluate a determinant, we say that we are expanding by cofactors.

Theorem 4.1: Expansion by cofactor

Let A be a square matrix of order n , then the determinant of A is given

$$|A| = \sum_{j=1}^n a_{ij} c_{ij} = a_{i1} c_{i1} + a_{i2} c_{i2} + \dots + a_{in} c_{in} \quad (i^{\text{th}} \text{ row expansion}) \text{ or}$$

$$|A| = \sum_{i=1}^n a_{ij} c_{ij} = a_{1j} c_{1j} + a_{2j} c_{2j} + \dots + a_{nj} c_{nj} \quad (j^{\text{th}} \text{ column expansion})$$

Proof: Exercise

☞ **Note:** when expanding by cofactors we do not need to evaluate the cofactors of zero entries, because a zero entry times its cofactor is zero. That is, $a_{ij} c_{ij} = (0) c_{ij} = 0$. Thus the row (or column) containing the most zeros is usually the best choice for expansion by cofactors.

Example-4 Find the determinant of

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

Solution:

Inspecting this matrix, we see that in third column of A we have more number of zeros than the other rows and columns. Thus we can eliminate some of the work in the expansion by using the third column. Therefore,

$$\begin{aligned} |A| &= a_{13} C_{13} + a_{23} C_{23} + a_{33} C_{33} + a_{43} C_{43} \\ &= 3 C_{13} + 0(C_{23}) + 0(C_{33}) + 0(C_{43}) \\ &= 3 C_{13} \end{aligned}$$

$$\text{But } C_{13} = (-1)^{1+3} \begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 4 & -2 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 4 & -2 \end{vmatrix} \quad \text{Again, expanding by cofactors in the second}$$

$$\text{row we get } C_{13} = (0) (-1)^3 \begin{vmatrix} 1 & 2 \\ 4 & -2 \end{vmatrix} + (2) (-1)^4 \begin{vmatrix} -1 & 2 \\ 3 & -2 \end{vmatrix} + (3) (-1)^5 \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix}$$

$$= 0 + 2(1)(-4) + 3(-1)(-7) = 13. \text{ Thus we obtain } |A| = 3(13) = 39$$

Theorem 4.2: Determinant of a Triangular matrix

Let $A = (a_{ij})$ be a triangular matrix of order n , and then its determinant is the product of the entries on the main diagonal.

Proof: We use mathematical induction to prove this theorem for the case in which A is an upper triangular matrix. The case in which A is lower triangular can be proven similarly. If A has order 1, then $A = [a_{11}]$ and the determinant is given by $|A| = a_{11}$.

Assuming that the theorem is true for any upper triangular matrix of order $k - 1$, we now consider an upper triangular matrix A of order k . Expanding by the k^{th} row, we obtain

$$|A| = a_{k1} C_{k1} + a_{k2} C_{k2} + \dots + a_{kk} C_{kk}$$

$$= a_{kk} C_{kk}$$

$$= a_{kk} (-1)^{k+k} M_{kk} = a_{kk} (-1)^{2k} M_{kk}$$

$$= a_{kk} M_{kk}$$

$$= a_{kk} |A_{kk}|$$

$$= a_{kk} (a_{11} a_{22} a_{33} \dots a_{k-1,k-1}), \text{ Since } A_{kk} \text{ is}$$

order $k - 1$ we can apply induction assumption to it

$$= a_{11} a_{22} a_{33} \dots a_{kk}$$

This completes the proof

Example-5 (a) Let $A = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}$ and $B = I_n$

Then since a diagonal matrix is both upper and lower triangular we have by theorem

$$4.2. |A| = (-1)(3)(2)(4)(-2) = 48 \text{ and } |B| = (1)(1) \dots (1) = 1 = |I_n|$$

☺ **Test yourself!**

1. Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ -1 & 4 & 5 \end{bmatrix}$, then find the determinant of A

- (a) By expanding along the first row
 (b) By expanding along the first column
 (c) By expanding along the third row.

What can you conclude from the result (a), (b) and (c).

(ii). Find the determinant of the given matrix by using expansion by cofactors.

(a) $\begin{bmatrix} 1 & 4 & -2 \\ 3 & 2 & 0 \\ -1 & 4 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} x & y & 1 \\ 2 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} -0.4 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.2 \\ 0.3 & 0.2 & 0.2 \end{bmatrix}$

(d) $\begin{bmatrix} 5 & 2 & 0 & 0 & -2 \\ 0 & 1 & 4 & 3 & 2 \\ 0 & 0 & 2 & 6 & 3 \\ 0 & 0 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ (e) $\begin{bmatrix} 4 & 3 & -2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & -7 & 13 & 12 \\ 6 & -2 & 5 & 6 & 7 \\ -1 & 4 & 2 & 0 & 9 \end{bmatrix}$

(iii). Verify the following equation

(a) $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (a-b)(b-c)(c-a)$

(b) $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} = (a-b)(b-c)(c-a)(a+b+c)$

(c) $\begin{vmatrix} x & 0 & c \\ -1 & x & b \\ 0 & -1 & a \end{vmatrix} = ax^2 + bx + c$

(iv) Solve the following equations

(a) $\begin{vmatrix} 2 & x-4 \\ 1 & 4 \end{vmatrix} = 0$

(b) $\begin{vmatrix} 1 & 4 \\ 3x & x+22 \end{vmatrix} = 0$

(c) $\begin{vmatrix} \lambda-1 & -4 \\ -2 & \lambda+1 \end{vmatrix} = 0$

(d) $\begin{vmatrix} y+1 & 0 & 0 \\ 4 & y & 3 \\ 2 & 8 & y+5 \end{vmatrix} = 0$

- (v) The determinant of a 2×2 matrix involves 2 products. The determinant of a 3×3 matrix involves 6 triple products. Show that the determinant of a 4×4 matrix involves 24 quadruple products. In general, the determinant of an $n \times n$ matrix involves $n!$ n -fold products.
-

4.2: Properties of Determinant

Objectives

At the end of this section learners are expected

- Discuss the properties of determinants.
- Evaluate a determinant using elementary Row operations.

It should be clear that calculating the determinant of a large matrix from the definition can be quite long and cumbersome. For instance, the determinant of a 6×6 matrix involves six 5×5 cofactors. Each of these in turn involves five 4×4 cofactors each of which involves four 3×3 cofactors for a total of $120 \times 3 \times 3$ cofactors to be calculated. Consequently other methods are often used to compute determinants. These methods depend on the properties of determinants which are stated as follows.

Property-1: Elementary Row operations and Determinants

Let A and B be square matrices

- (a) If B is obtained from A by interchanging two rows (or columns) of A , then

$$|B| = -|A|$$

- (c) If B is obtained from A by adding a multiple of a row (or a column) of A to another row (or c column) A then the determinant is unchanged. That is, $|B| = |A|$.

- (c) If B is obtained from A by multiplying a row (or a column) of A by a non zero c constant c , then $|B| = c|A|$.

Proof: (a) we use induction to prove this.

- (i). Assume that A and B are 2×2 matrices such that

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}$$

Then, we have

$$|A| = a_{11} a_{22} - a_{12} a_{21} \quad \text{and} \quad |B| = a_{21} a_{12} - a_{11} a_{22} \\ \Rightarrow |B| = -(a_{11} a_{22} - a_{12} a_{21}) = -|A|$$

(ii). Now assume that the property is true for matrices of order

$(n-1) \times (n-1)$. That is, if $A \xrightarrow{R_i \leftrightarrow R_j} B$ then $|B| = -|A|$
where A and B are any $(n-1) \times (n-1)$ matrices.

(iii) Let A be an $n \times n$ matrix such that B is obtained from A by interchanging two rows of A . Then to find $|A|$ and $|B|$, we expand along a row other than the two interchanged rows. Without loss of generality let us expand along the first row of B ; then

$$|B| = \sum_{j=1}^n a_{1j} c_{1j}.$$

But by our induction assumption, the cofactors of B will be the negatives of the cofactors of A because the corresponding $(n-1) \times (n-1)$ matrices have two rows interchanged. Thus,

$$|B| = \sum_{j=1}^n a_{1j} C'_{1j} \quad \text{where } C'_{1j} \text{ is the cofactors of } A \quad \forall j = 1, 2, \dots, n \\ = - \sum_{j=1}^n a_{1j} C_{1j} \\ = -|A|$$

$$\Rightarrow |B| = -|A| \quad \text{and the Proof is complete.}$$

The proofs of (b) and (c) are left as exercises.

Example-1: Show that the determinant of an elementary matrix E of

(i) First kind (add a multiple of one row to another) is 1

(ii) Second kind (row interchange) is -1

(iii) Third kind (multiply a row by a non zero constant c) is c

Solution: We know that $|I_n| = 1$. Thus, we have

$$(i) |E| = |I_n|, \text{ by property 1 (b)} \\ = 1$$

$$\begin{aligned} \text{(ii) } |E| &= -|I_n|, \text{ by property 1 (a)} \\ &= -1 \end{aligned}$$

$$\begin{aligned} \text{(iii) } |E| &= c |I_n|, \text{ by property 1 (c)} \\ &= c \cdot 1 = c \end{aligned}$$

Example-2: Evaluating a determinant using elementary Row operations.

Let $A = \begin{bmatrix} 2 & -3 & 10 \\ 1 & 2 & -2 \\ 0 & 1 & -3 \end{bmatrix}$ then find determinant of A .

Solution:

$$\begin{aligned} &= -(-7) \begin{vmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 1 & -3 \end{vmatrix}, \text{ factor } -7 \text{ out of the second row} \\ &= 7 \begin{vmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 1 & -3 \end{vmatrix} = 7 \begin{vmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{vmatrix}, \text{ by } R_3 \rightarrow -R_2 + R_3 \\ &= 7(1)(1)(-1) = -7, \text{ Since } \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{bmatrix} \text{ is triangular matrix} \\ |A| &= \begin{vmatrix} 2 & -3 & 10 \\ 1 & 2 & -2 \\ 0 & 1 & -3 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & -2 \\ 2 & -3 & 10 \\ 0 & 1 & -3 \end{vmatrix}, \text{ by } R_1 \leftrightarrow R_2 \\ &= - \begin{vmatrix} 1 & 2 & -2 \\ 0 & -7 & 14 \\ 0 & 1 & -3 \end{vmatrix}, \text{ by } R_2 \rightarrow -2R_1 + R_2 \end{aligned}$$

 **Note:** In the above solution we have changed the given matrix in to triangular matrix.

Property-2: Determinant of a transpose

If A is a square matrix, then $|A| = |A^t|$

Proof: Exercise(Hint: use mathematical induction and Laplace's expansion of determinant)

Example-3: Show that $|A| = |A^t|$ for the matrix $A = \begin{bmatrix} 3 & 1 & -2 \\ 2 & 0 & 0 \\ -4 & -1 & 5 \end{bmatrix}$

Solution:

We expand by cofactors along the second row to find $|A|$

To find the determinant of $A^t = \begin{bmatrix} 3 & 2 & -4 \\ 1 & 0 & -1 \\ -2 & 0 & 5 \end{bmatrix}$ We expand by cofactors down the

second column to get $|A^t| = \begin{vmatrix} + & - & + \\ 3 & 2 & -4 \\ 1 & 0 & -1 \\ -2 & 0 & 5 \end{vmatrix} = -2 \begin{vmatrix} 1 & -1 \\ 2 & -5 \end{vmatrix} = (-2)(3) = -6$

Thus, $|A| = -6 = |A^t| \Rightarrow |A| = |A^t|$



Notation

Let A is a square matrix of order n with columns $A^1, A^2, A^3 \dots A^j \dots$ and A^n

And rows $A_1, A_2, A_3 \dots A_j \dots$ and A_n then we can also denote $\det(A)$ by

$\det(A^1, A^2, A^3, \dots, A^j, \dots, A^n)$ or $\det(A_1, A_2, A_3 \dots A_j \dots, A_n)$

Remarks – property - 1 can be expressed as

(1) **a.** $\det(A^1, A^2, A^3, \dots, cA^j, \dots, A^n) = c \det(A^1, A^2, A^3, \dots, A^j, \dots, A^n)$

b. $\det(A_1, A_2, A_3 \dots cA_j \dots, A_n) = c \det(A_1, A_2, A_3 \dots A_j \dots, A_n)$ where c is a non zero constant

(2) **a.** $\det(A^1, A^2, A^3, \dots, A^i, \dots, A_j, \dots, A^n) = - \det(A^1, A^2, A^3, \dots, A^j, \dots, A^i, \dots, A^n)$

b. $\det(A_1, A_2, A_3 \dots A_i, \dots, A_j \dots, A_n) = - \det(A_1, A_2, A_3 \dots A_j, \dots, A_i, \dots, A_n)$

(3) **a.** $\det(A^1, A^2, A^3, \dots, cA^j + A^i, \dots, A^j, \dots, A^n)$

$$= \det(A^1, A^2, A^3, \dots, A^i, \dots, A^j, \dots, A^n)$$

b. $\det(A_1, A_2, A_3 \dots cA_j + A_i, \dots, A_j \dots, A_n) = \det(A_1, A_2, A_3 \dots A_i, \dots, A_j \dots, A_n)$ where c is constant number.

Property-3: Linearity of determinant

Let A is a square matrix of order n then

a. the function $|A|$ is linear in the rows of the matrix

$$\text{i.e., } \det(A_1, A_2, A_3 \dots cA_i + A_i, \dots, A_j, \dots, A_n) = c \det(A_1, A_2, A_3 \dots A_i, \dots, A_j, \dots, A_n) \\ + \det(A_1, A_2, A_3 \dots A_i, \dots, A_j, \dots, A_n)$$

where c is a constant number.

b. the function $|A|$ is linear in the columns of the matrix

$$\text{That is, } \det(A^1, A^2, A^3 \dots cA^i + A^{i'}, \dots, A^j, \dots, A^n) \\ = c \det(A^1, A^2, A^3 \dots A^i, \dots, A^j, \dots, A^n) + \det(A^1, A^2, A^3 \dots A^{i'}, \dots, A^j, \dots, A^n)$$

Where c is a constant number.

Proof: Exercise. Hint (use mathematical induction)

Class activity

1. Show that

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

$$(b.) \begin{vmatrix} 1 & 2 & 4 \\ 2(5) & 2(6) & 2(3) \\ 2 & -1 & 0 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 3 \\ 2 & -1 & 0 \end{vmatrix}$$

2. Show that

$$(a.) \begin{vmatrix} a_1 & a_2 + x & a_3 \\ b_1 & b_2 + y & b_3 \\ c_1 & c_2 + z & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & x & a_3 \\ b_1 & y & b_3 \\ c_1 & z & c_3 \end{vmatrix}$$

$$(b.) \begin{vmatrix} a_1 & a_2 & xa_3 \\ b_1 & b_2 & xb_3 \\ c_1 & c_2 & xc_3 \end{vmatrix} = x \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

3. Show that

$$(a.) \begin{vmatrix} b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = - \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$(b) \begin{vmatrix} a_1 & a_2 & a_3 \\ kb_1 & kb_2 & kb_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = k \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$(c) \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 + kc_1 & b_2 + kc_2 & b_3 + kc_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Property-4: Conditions that yield a zero determinant

Let A is a square matrix of order n and any of the following conditions is true, then

$$|A| = 0$$

- (a) An entire row(or an entire column) consists of zero
- (b) Two rows (or columns) are equal.
- (c) One row(or two columns) is a multiple of other row (column)

Proof: Let A be an $n \times n$ matrix with rows $A_1, A_2, A_3, \dots, A_i, \dots, A_j, \dots, A_n$

- (a) Without loss of generality suppose that the first row of A is a row of zeros then

$$\begin{aligned} \det(A) &= \det(0, A_2, A_3, \dots, A_j, \dots, A_n) = \det(0, 0, A_2, A_3, \dots, A_j, \dots, A_n) \\ &= 0 \det(0, A_2, A_3, \dots, A_j, \dots, A_n) = 0 \end{aligned}$$

Hence, the proof.

- (b) Without loss of generality assume that $A_1 = A_2$ then

$$\begin{aligned} \det(A) &= \det(A_1, A_2, A_3, \dots, A_j, \dots, A_n) = \det(A_1 - A_2, A_2, A_3, \dots, A_j, \dots, A_n) \\ &= \det(0, A_2, A_3, \dots, A_j, \dots, A_n), \text{ as } A_1 = A_2 \\ &= 0, \text{ by (a)} \end{aligned}$$

- (c) Without loss of generality assume that $A_1 = cA_2$ then

$$\begin{aligned} \det(A) &= \det(A_1, A_2, A_3, \dots, A_j, \dots, A_n) = \det(cA_2, A_2, A_3, \dots, A_j, \dots, A_n) \\ &= c \det(A_2, A_2, A_3, \dots, A_j, \dots, A_n) = c \cdot 0 = 0, \text{ by (b)} \end{aligned}$$

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

(a) Since A_2 is a row of zeros we have $|A| = 0$

(b) Since $B_1 = B_3$ we have $|B| = 0$

(c) Since $C^3 = -3C^1$ we have $|C| = 0$

Further Properties of determinant

Theorem 4.3: Determinant of a scalar multiple of a matrix

If A is an $n \times n$ matrix and c is a scalar, then the determinant of cA is given by $|cA| = c^n |A|$

Proof: $|cA| = \det(cA) = \det(cA^1, cA^2, cA^3, \dots, cA^n)$
 $= c \det(A^1, cA^2, cA^3, \dots, cA^n)$
 $= c^2 \det(A^1, A^2, cA^3, \dots, cA^n)$
 $= c^3 \det(A^1, A^2, A^3, \dots, cA^n)$
 $\vdots$
 $= c^n \det(A^1, A^2, A^3, \dots, A^n)$
 $= c^n \det(A) = c^n |A|$

Example- 5: If A is a matrix of order 3 and $|A^t| = 5$ then find determinant of $10A$

Solution: $|10A| = 10^3 |A| = 10^3 |A^t| = 10^3 (5) = 5000$

Theorem 4.4: Let E be an elementary matrix and A be an arbitrary square matrix then $|EA| = |E||A|$

In general, if $E_1, E_2, E_3 \dots E_k$ are elementary matrix then

$$|E_k \dots E_3 E_2 E_1 A| = |E_k| \dots |E_3| |E_2| |E_1| |A|$$

Proof: If E is an elementary matrix obtained from identity matrix I by

I- interchanging two rows then

$$\begin{aligned} |E| &= -|I| = -1 \text{ and hence } |EA| = -|A|, \text{ by property (1)} \\ &= -1|A| = |E||A|, \text{ since } |E| = -1 \\ \Rightarrow |EA| &= |E||A| \end{aligned}$$

II- by multiplying a row of I by a non zero content c then

$$\begin{aligned} |E| &= c \text{ and hence} \\ |EA| &= C |A|, \text{ by property (1)} \\ &= |E| |A|, \text{ since } |E| = C \end{aligned}$$

III- by adding a multiple of one row of I to another row of I , then $|E| = 1$ and

$$\begin{aligned} \text{hence } |EA| &= |A|, \text{ by property (1)} \\ &= 1 |A| \\ &= |E| |A|, \text{ since } |E| = 1 \end{aligned}$$

Theorem 4.5: Determinant of a matrix product

If A and B are square matrices of order n , then $|AB| = |A| |B|$. In general,

$$|A_1 A_2 \dots A_k| = |A_1| |A_2| \dots |A_k|.$$

Proof: We note that this is theorem 4.4 if A is an elementary matrix.

Case 1: If A is invertible, then A can be written as a product of elementary matrices

$$\begin{aligned} A &= E_k \dots E_2 E_1 \text{ and hence } |AB| = |E_k \dots E_3 E_2 E_1 B| \\ &= |E_k| \dots |E_3| |E_2| |E_1| |B| \\ &= |E_k \dots E_3 E_2 E_1| |B|, \text{ by induction} \\ &\quad \text{and theorem 4.4} \\ &= |A| |B|. \end{aligned}$$

Case 2: If A is singular, then A is row equivalent to a matrix with an entire row of zeros. Thus, $|A| = 0$. Moreover, because A is singular we have AB is also singular. (For if AB is non singular then $A [B (AB)^{-1}] = I$ and hence A is nonsingular, which is contradiction). Thus, $|AB| = 0 = 0|B| = |A| |B|$.

Hence the result.

Theorem 4.6: Determinant of an Invertible Matrix

A square matrix A is invertible (non singular) if and only if $|A| \neq 0$

Proof: ($\Rightarrow$) Suppose A is invertible, then

$$\begin{aligned} AA^{-1} &= I \Rightarrow |AA^{-1}| = |I| \\ &\Rightarrow |A| |A^{-1}| = 1 \\ &\Rightarrow |A| \neq 0 \text{ and } |A^{-1}| \neq 0. \end{aligned}$$

Thus, $|A| \neq 0$.

($\Leftarrow$) Suppose $|A| \neq 0$, then using Gauss-Jordan elimination, we find a matrix B , in reduced row echelon form, that is row equivalent to A . Because B is in reduced row-echelon form, it must be the identity matrix I or it must have at least one row that consists entirely of zeros. However, if B has a row of all zeros, then $|B| = 0$, which would imply that $|A| = 0$, which is contradiction. Thus, $B = I$ and hence A is row equivalent to $I \Rightarrow A$ is invertible.

Theorem 4.7: Determinant of an Inverse Matrix

If A is invertible, then $|A^{-1}| = \frac{1}{|A|}$.

Proof: Suppose A is invertible then $AA^{-1} = I$ and

$$\begin{aligned} |A| \neq 0 &\Rightarrow |AA^{-1}| = |I| \\ &\Rightarrow |A| |A^{-1}| = 1 \\ &\Rightarrow |A^{-1}| = \frac{1}{|A|} \end{aligned}$$

Example-6: Use a determinant to decide whether A is singular or non-singular. If A is invertible, then find $|A^{-1}|$.

$$\begin{aligned} \text{(a) } A &= \begin{bmatrix} 0 & 2 & -1 \\ 3 & -2 & 1 \\ 3 & 2 & -1 \end{bmatrix} & \text{(b) } A &= \begin{bmatrix} 1 & 0 & 3 \\ 3 & -2 & 1 \\ 3 & 2 & 1 \end{bmatrix} \end{aligned}$$

Solution:

(a) Because $|A| = 0$ we conclude that A has no inverse

(b) Since $|A| = 4 \neq 0$ we have A is invertible and $|A^{-1}| = \frac{1}{|A|} = \frac{1}{4}$

😊 **Test yourself!**

1. Which property of determinants is illustrated by the given equation?

$$(a) \begin{vmatrix} -4 & 5 \\ 12 & -15 \end{vmatrix} = 0 \quad (b) \begin{vmatrix} 5 & 10 \\ 2 & -7 \end{vmatrix} = 5 \begin{vmatrix} 1 & 2 \\ 2 & -7 \end{vmatrix}$$

$$(c) \begin{vmatrix} -4 & 3 & 2 \\ 8 & 0 & 0 \\ -4 & 3 & 2 \end{vmatrix} = 0 \quad (d) \begin{vmatrix} 1 & 3 & 4 \\ -7 & 2 & -5 \\ 6 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 4 & 3 \\ -7 & -5 & 2 \\ 6 & 2 & 1 \end{vmatrix}$$

$$(e) \begin{vmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{vmatrix} = 6^4 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (g) \begin{vmatrix} 1 & -3 & 2 \\ 5 & 2 & 1 \\ -1 & 0 & 6 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 2 \\ 0 & 17 & -11 \\ -1 & 0 & 6 \end{vmatrix}$$

$$(f) \begin{vmatrix} 5 & 0 & 10 \\ 25 & -30 & 40 \\ -15 & 5 & 20 \end{vmatrix} = 5^3 \begin{vmatrix} 1 & 0 & 2 \\ 5 & -6 & 8 \\ -3 & 1 & 4 \end{vmatrix} \quad (h) \begin{vmatrix} 3 & 2 & 4 \\ -2 & 1 & 5 \\ 5 & -7 & -20 \end{vmatrix} = \begin{vmatrix} 3 & 2 & -6 \\ -2 & 1 & 0 \\ 5 & -7 & 15 \end{vmatrix}$$

2. Use elementary row or column operations to evaluate the following determinant

$$(a) \begin{vmatrix} 1 & -1 & 8 & 4 & 2 \\ 2 & 6 & 0 & -4 & 3 \\ 2 & 0 & 2 & 6 & 2 \\ 0 & 2 & 8 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 \end{vmatrix} \quad (b) \begin{vmatrix} 3 & -2 & 4 & 3 & 1 \\ -1 & 0 & 2 & 1 & 0 \\ 5 & -1 & 0 & 3 & 2 \\ 4 & 7 & -8 & 0 & 0 \\ 1 & 2 & 3 & 0 & 2 \end{vmatrix}$$

$$(c) \begin{vmatrix} -2 & 0 & 1 & 3 & -2 \\ -2 & 1 & 3 & 2 & -1 \\ 1 & 0 & -1 & 2 & 3 \\ 3 & -1 & 2 & 4 & -3 \\ 1 & 1 & 3 & 2 & 0 \end{vmatrix} \quad (d) \begin{vmatrix} 1 & -2 & 7 & 9 \\ 3 & -4 & 5 & 5 \\ 3 & 6 & 1 & -1 \\ 4 & 5 & 3 & 2 \end{vmatrix}$$

3. Find the determinant of the given elementary matrix where $b \neq 0$.

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

4. Verify the following equation

$$(a) \begin{vmatrix} a+b & a & a \\ a & a+b & a \\ a & a & a+b \end{vmatrix} = b^2(3a+b)$$

$$(b) \begin{vmatrix} 1+a & 1 & 1 \\ 1 & 1+b & 1 \\ 1 & 1 & 1+c \end{vmatrix} = abc(1 + \frac{1}{a} + \frac{1}{b} + \frac{1}{c}), a \neq 0, b \neq 0, c \neq 0.$$

5. Let A and B square matrices of order 4 such that $|A| = -5$, $|B| = 3$ then find

$$(a) |AB| \quad (b) |A^3| \quad (c) |3B| \quad (d) |(AB)^t| \quad (e) |A^{-1}|$$

6. Find the value (s) of b such that A is singular

$$(a) A = \begin{bmatrix} b-1 & 3 \\ 2 & b-2 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 0 & 3 \\ 2-1 & 0 \\ 4 & 2 & b \end{bmatrix}$$

7. Prove that

$$(a) \text{ If } A \text{ is an } n \times n \text{ skew symmetric matrix, then } |A| = (-1)^n |A|.$$

$$(b) \text{ If } A \text{ is a skew symmetric matrix of odd order, then } |A| = 0$$

(c) If A and B are $n \times n$ matrices such that

$$I. AB = I \text{ then } |A| \neq 0 \text{ and } |B| \neq 0$$

$$II. AB \text{ is singular, then either } A \text{ or } B \text{ is singular}$$

8. Find two 2×2 matrices such that

$$I. |A| + |B| \neq |A+B|$$

$$II. |A| + |B| = |A+B|$$

4.3 Adjoint and inverse of a Matrix

In this section we develop an explicit formula for the inverse of a non singular matrix.

In the next section we use this formula to derive a theorem known as Cramer's Rule.

Objectives

At the end of this section students are expected

- Define adjoint of a matrix.
- Find inverse of a matrix using adjoint.

Definition 4.6: Adjoint of a Matrix

Let A be a square matrix of order n with cofactors $C_{ij} = (-1)^{i+j} |A_{ij}|$ then the transpose of the matrix of cofactors of A is called the ad joint of A and is denoted by $adj(A)$.

$$\text{That is, } adj(A) = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & & & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}^t = \begin{bmatrix} C_{11} & C_{21} & C_{31} & \dots & C_{n1} \\ C_{12} & C_{22} & C_{32} & \dots & C_{n2} \\ \vdots & & & & \vdots \\ C_{1n} & C_{2n} & C_{3n} & \dots & C_{nn} \end{bmatrix}$$

Where $C_{ij} = (-1)^{i+j} |A_{ij}|$ and A_{ij} is the matrix obtained by deleting the i^{th} row and j^{th} column of A .

Example 1: Let $A = \begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{bmatrix}$ then find $adj(A)$.

Solution: Let C_{ij} be the cofactor of A then

$$C_{11} = (-1)^{1+1} |A_{11}| = + |A_{11}| = \begin{vmatrix} -2 & 1 \\ 0 & -2 \end{vmatrix} = 4,$$

$$C_{12} = (-1)^{1+2} |A_{12}| = - |A_{12}| = - \begin{vmatrix} 0 & 1 \\ 1 & -2 \end{vmatrix} = -(-1) = 1,$$

$$C_{13} = (-1)^{1+3} |A_{13}| = + |A_{13}| = \begin{vmatrix} 0 & -2 \\ 1 & 0 \end{vmatrix} = 2$$

$$C_{21} = (-1)^{2+1} |A_{21}| = - \begin{vmatrix} 3 & 2 \\ 0 & -2 \end{vmatrix} = -(-6) = 6,$$

$$C_{22} = (-1)^{2+2} |A_{22}| = + \begin{vmatrix} -1 & 2 \\ 1 & -2 \end{vmatrix} = 0,$$

$$C_{23} = (-1)^{2+3} |A_{23}| = - \begin{vmatrix} -1 & 3 \\ 1 & 0 \end{vmatrix} = -(-3) = 3,$$

$$C_{31} = (-1)^{3+1} |A_{31}| = + \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} = 7,$$

$$C_{32} = (-1)^{3+2} |A_{32}| = - \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -(-1) = 1,$$

$$C_{33} = (-1)^{3+3} |A_{33}| = \begin{vmatrix} -1 & 3 \\ 0 & -1 \end{vmatrix} = 2$$

Thus, the matrix of cofactors of A is

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 6 & 0 & 3 \\ 7 & 1 & 2 \end{bmatrix} \text{ and hence } adj(A)$$

$$= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}^t = \begin{bmatrix} 4 & 1 & 2 \\ 6 & 0 & 3 \\ 7 & 1 & 2 \end{bmatrix}^t = \begin{bmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{bmatrix}.$$

Theorem 4.8: The Inverse of a Matrix Given by its Adjoint

If A is an $n \times n$ invertible matrix, then $A^{-1} = \frac{1}{|A|} adj(A)$.

Proof: To prove this first consider the product $A adj(A)$.

$$\text{Now, } A adj(A) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & \dots & c_{j1} & \dots & c_{n1} \\ c_{12} & c_{22} & \dots & c_{j2} & \dots & c_{n2} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{1n} & c_{2n} & \dots & c_{jn} & \dots & c_{nn} \end{bmatrix}$$

Then the entry in the i th row and j th column of the product $A adj(A)$ is given by

$$a_{i1} c_{j1} + a_{i2} c_{j2} + \dots + a_{in} c_{jn}$$

If $i = j$, then the sum $a_{i1} c_{j1} + a_{i2} c_{j2} + \dots + a_{in} c_{jn}$

$$= \sum_{j=1}^n a_{ij} c_{ij}$$

$= |A|$ —, (*) Since this sum is the cofactor expansion of along its i th row.

On the other hand, if $i \neq j$, then the sum $a_{i1} c_{j1} + a_{i2} c_{j2} + \dots + a_{in} c_{jn} = 0$. To see this, consider the following matrix B in which the j th row of A has been replaced with the i th row of A.

$$\begin{bmatrix}
 a_{11} & a_{12} & \cdots & a_{1n} \\
 a_{21} & a_{22} & \cdots & a_{2n} \\
 \vdots & & & \vdots \\
 a_{i1} & a_{i2} & \cdots & a_{in} \\
 \vdots & \vdots & & \vdots \\
 a_{j1} & a_{j2} & \cdots & a_{jn} \\
 \vdots & \vdots & & \vdots \\
 a_{n1} & a_{n2} & \cdots & a_{nn}
 \end{bmatrix}
 \begin{matrix}
 \\
 \\
 \\
 \leftarrow i\text{th row} \\
 \\
 \\
 \leftarrow j\text{th row}
 \end{matrix}$$

Cofactor expansion along the j th row of B produces

$$|B| = a_{j1} C_{j1} + a_{j2} C_{j2} + \cdots + a_{jn} C_{jn} = 0, \text{ because } B \text{ has two identical rows}$$

$$\Rightarrow a_{j1} C_{j1} + a_{j2} C_{j2} + \cdots + a_{jn} C_{jn} = |B| = 0 \quad \text{--- (**)}$$

Therefore, by (*) and (**) we have

$$A \operatorname{adj}(A) = (r_{ij})_{n \times n} \text{ with } r_{ij} = |A|, \text{ if } i = j \\ 0, \text{ if } i \neq j$$

$$\Rightarrow \operatorname{adj}(A) = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix} = |A| \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = |A| I$$

$$\Rightarrow A \operatorname{adj}(A) = |A| I$$

$$\Rightarrow A \left[\frac{1}{|A|} \operatorname{adj}(A) \right] = I, \text{ since } A \text{ is invertible we have } |A| \neq 0$$

$$\Rightarrow A^{-1} = \frac{1}{|A|} \operatorname{adj}(A)$$

Example-2: use the adjoint of $A = \begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{bmatrix}$ to find A^{-1} .

Solution:

$$\begin{aligned}
 \text{I. (i) } |A| &= \begin{vmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{vmatrix} = \begin{vmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 0 & 3 & 0 \end{vmatrix}, \text{ by } R_3 \rightarrow R_1 + R_3 \\
 &= -1 \begin{vmatrix} -2 & 1 \\ 3 & 0 \end{vmatrix} = -1(-3) = 3, \text{ by cofactor expansion along the first column}
 \end{aligned}$$

$$\Rightarrow |A| = 3$$

II.

$$C_{11} = + |A_{11}| = + \begin{vmatrix} -2 & 1 \\ 0 & -2 \end{vmatrix} = 4,$$

$$C_{12} = - |A_{12}| = - \begin{vmatrix} 0 & 1 \\ 1 & -2 \end{vmatrix} = -(-1) = 1,$$

$$C_{13} = + |A_{13}| = + \begin{vmatrix} 0 & -2 \\ 1 & 0 \end{vmatrix} = 2,$$

$$C_{21} = - |A_{21}| = - \begin{vmatrix} 3 & 2 \\ 0 & -2 \end{vmatrix} = -(-6) = 6,$$

$$C_{22} = + |A_{22}| = + \begin{vmatrix} -1 & 2 \\ 1 & -2 \end{vmatrix} = 0,$$

$$C_{23} = - |A_{23}| = - \begin{vmatrix} -1 & 3 \\ 1 & 0 \end{vmatrix} = -(-3) = 3,$$

$$C_{31} = + |A_{31}| = + \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} = 7,$$

$$C_{32} = - |A_{32}| = - \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -(-1) = 1,$$

$$C_{33} = + |A_{33}| = + \begin{vmatrix} -1 & 3 \\ 0 & -2 \end{vmatrix} = 2,$$

$$\Rightarrow \text{The matrix of cofactors of } A = \begin{bmatrix} 4 & 1 & 2 \\ 6 & 0 & 3 \\ 7 & 1 & 2 \end{bmatrix}$$

$$\Rightarrow \text{adj}(A) = \begin{bmatrix} 4 & 1 & 2 \\ 6 & 0 & 3 \\ 7 & 1 & 2 \end{bmatrix}^t = \begin{bmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{bmatrix}$$

$$\text{III. } A^{-1} = \frac{1}{|A|} \text{adj}(A) = \frac{1}{3} \begin{bmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & 2 & \frac{7}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & 1 & \frac{2}{3} \end{bmatrix}$$

$$\text{In fact you can check that } AA^{-1} = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Test yourself!

1. Find the adjoint of the matrix A . Then use the adjoint to find the inverse of A , if possible

$$(a) A = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (d) A = \begin{bmatrix} -1 & 2 & 0 & 1 \\ 3 & -1 & 4 & 1 \\ 0 & 0 & 1 & 2 \\ -1 & 1 & 1 & 2 \end{bmatrix}$$

2. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then show that

$$a) \operatorname{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$b) A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \text{ provided that } |A| \neq 0.$$

3. Prove that

a) If A is an $n \times n$ invertible matrix, then $\operatorname{adj}(A^{-1}) = [\operatorname{adj}(A)]^{-1}$

b) If A is an $n \times n$ singular matrix, then $A[\operatorname{adj}(A)]$ is the zero matrix.

4. Let A be an $n \times n$ non singular matrix then prove that

$$a) |\operatorname{adj}(A)| = |A|^{n-1}$$

$$b) \operatorname{adj}[\operatorname{adj}(A)] = |A|^{n-2} A$$

4.4 Cramer's Rule for solving system of linear equations (Homogenous and non homogenous)

Objectives

At the end of this section students should be able to

- Use Cramer's Rule to solve the given system of linear equations.
- Find volume of the tetrahedron with given vertices.

Cramer's Rule, named after Gabriel Cramer (1704 - 1752), is a formula that uses determinants to solve a system of n linear equations in n -variables. This rule can be applied only to systems of linear equations that have unique solutions.

Theorem 4.9: Cramer's Rule

If a system of n linear equations in n variables has a coefficient matrix A with a non-zero determinant $|A|$, then the solution to the system

$$Ax = B \text{ is given by } x_j = \frac{\det(A^1, A^2, \dots, A^{j-1}, B, A^{j+1}, \dots, A^n)}{|A|} \quad \forall j = 1, 2, \dots, n.$$

Proof: since $|A| \neq 0$ we have A is invertible and hence the system $Ax = B$ has a unique solution given by $x = A^{-1}B$.

$$\text{Let } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ then } x = A^{-1} B$$

$$\Rightarrow x = \frac{1}{|A|} \text{adj}(A) B =$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{|A|} \text{adj}(A) B = \frac{1}{|A|} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} C_{11} b_1 + C_{21} b_2 + \dots + C_{n1} b_n \\ C_{12} b_1 + C_{22} b_2 + \dots + C_{n2} b_n \\ \vdots \\ C_{1j} b_1 + C_{2j} b_2 + \dots + C_{nj} b_n \\ \vdots \\ C_{1n} b_1 + C_{2n} b_2 + \dots + C_{nn} b_n \end{bmatrix}$$

$$\Rightarrow x_j = \frac{1}{|A|} (C_{1j} b_1 + C_{2j} b_2 + \dots + C_{nj} b_n) \\ = \frac{1}{|A|} \det(A^1, A^2, \dots, A^{j-1}, B, A^{j+1}, \dots, A^n), \text{how?}$$

Therefore $\forall_j = 1, 2, \dots, n$ we have

$x_j = \frac{1}{|A|} \det (A^1, A^2, \dots, A^{j-1}, B, A^{j+1}, \dots, A^n)$ and the proof is complete.

Example: Use Cramer's Rule to solve the following system of linear equations for x, y & z

$$-x + 2y - 3z = 1$$

$$2x + \quad \quad z = 0$$

$$3x - 4y + 4z = 2$$

Solution: Let $A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 0 & 1 \\ 3 & -4 & 4 \end{bmatrix}$, $x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and

$B = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, then the system can be represented by $Ax = B$.

and since $|A| = \begin{vmatrix} -1 & 2 & -3 \\ 2 & 0 & 1 \\ 3 & -4 & 4 \end{vmatrix} = 10 \neq 0$

Then by Cramer's rule the system $Ax = B$ has a unique solution given by

$$x = \det \frac{(B, A^2, A^3)}{|A|}, \quad y = \det \frac{(A^1, B, A^3)}{|A|} \quad \text{and} \quad z = \det \frac{(A^1, A^2, B)}{|A|}$$

$$\Rightarrow x = \frac{\begin{vmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 2 & -4 & 4 \end{vmatrix}}{10}, \quad y = \frac{\begin{vmatrix} -1 & 1 & -3 \\ 2 & 0 & 1 \\ 3 & 2 & 4 \end{vmatrix}}{10}, \quad \text{and} \quad z = \frac{\begin{vmatrix} -1 & 2 & 1 \\ 2 & 0 & 0 \\ 3 & -4 & 2 \end{vmatrix}}{10}$$

$$\Rightarrow x = \frac{4}{5}, \quad y = -\frac{3}{2} \quad \text{and} \quad z = -\frac{8}{5} \quad \text{Thus the solution set for}$$

$Ax = B$ is given by S.S. = $\{(x, y, z) : x = \frac{4}{5}, y = -\frac{3}{2} \text{ and } z = -\frac{8}{5}\}$

☺ **Test yourself!**

1. Use Cramer's Rule to solve the given system of linear equations

(a) $3x - 5y = 13$

$2x + 7y = 81$

(c) $4x_1 - x_2 - x_3 = 1$

$2x_1 + 2x_2 + 3x_3 = 10$

$5x_1 - 2x_2 - 2x_3 = -1$

(b) $x + y + z = 3$

$2x - y + 4z = 5$

$x - 3y - 9z = -11$

(d) $4x_1 - 2x_2 + 3x_3 = -2$

$2x_1 + 2x_2 + 5x_3 = 16$

$8x_1 - 5x_2 - 2x_3 = 4$

2. Use Cramer's Rule to solve the following system of linear equations for x and y

$$ax + (1-a)y = 1$$

$$(1-a)x + ay = 3$$

For what value (s) of a will the system be inconsistent?

3. Determine the values of a and b for the following given system

I. Has a unique solution

II. Has no solution

III. Has infinitely many solutions

(a) $3x - ay = 1$

$6x + 4y = b$

(b) $3x - 2y + z = b$

$5x - 8y + 9z = 3$

$2x + y + az = -1$

4. Consider the system of linear equations in $\cos A$, $\cos B$ and $\cos C$ for the triangle

$$c \cos B + b \cos C = a$$

$$c \cos A + a \cos C = b$$

$$b \cos A + a \cos B = c$$

then use Cramer's Rule to solve for $\cos C$ and use the result to verify the law of cosines,

$$c^2 = a^2 + b^2 - 2ab \cos c$$

5. Show that

(a) The area of a triangle whose vertices are

(x_1, y_1) , (x_2, y_2) and (x_3, y_3) is given by

$$\text{Area} = \text{the absolute value of } \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

(b) The volume of the tetrahedron whose vertices are

(x_1, y_1, z_1) , (x_2, y_2, z_2) , (x_3, y_3, z_3) , (x_4, y_4, z_4) is given by

$$v = \text{the absolute value of } \frac{1}{6} \begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}$$

4.5. The Rank of a matrix by sub determinants

Objectives

At the end of this section learners should be able to

- Determine the rank of non square matrix.
- Explain the similarity and difference of rank, trace and determinant.

The rank of a general $M \times N$ matrix is an important concept, particularly in the solution of sets of simultaneous linear equations. Like the trace and determinant, the rank of matrix A is a single number (or algebraic expression) that depends on the elements of A . Unlike the trace and determinant, however, the rank of a matrix can be defined even when A is not square. As we shall see, there are two equivalent definitions of the rank of a general matrix. Firstly, the rank of a matrix may be defined in terms of the linear independence of vectors. Suppose that the columns of an $M \times N$ matrix are interpreted as the components in a given basis of N (M -component) vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ as follows:

$$A = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_N \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix}.$$

Then the rank of A , denoted by $\text{rank } A$ or by $R(A)$, is defined as the number of linearly independent vectors in the set $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$, and equals the dimension of the vector space spanned by those vectors. Alternatively, we may consider the rows of A to contain the components in a given basis of the M (N -component) vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M$ as follows:

$$A = \begin{pmatrix} \leftarrow & \mathbf{w}_1 & \rightarrow \\ \leftarrow & \mathbf{w}_2 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{w}_M & \rightarrow \end{pmatrix}.$$

It may then be shown that the rank of A is also equal to the number of linearly independent vectors in the set $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M$. From this definition it should be clear that the rank of A is unaffected by the exchange of two rows (or two columns) or by the multiplication of a row (or column) by a constant.

Furthermore, suppose that a constant multiple of one row (column) is added to another row (column): for example, we might replace the row $\mathbf{w}_i$ by $\mathbf{w}_i + c\mathbf{w}_j$. This also has no effect on the number of linearly independent rows and so leaves the rank of A unchanged. We may use these properties to evaluate the rank of a given matrix. A second (equivalent) definition may be given of the rank of a matrix and uses the concept of sub matrices. A sub matrix of A is any matrix that can be formed from the elements of A by ignoring one or more than one, row or column. It may be shown that the rank of a general $M \times N$ matrix is equal to the size of the largest square sub matrix of A whose determinant is non-zero. Therefore, if a matrix A has an $r \times r$ sub matrix S with $|\mathbf{S}| \neq 0$, but no $(r+1) \times (r+1)$ sub matrix with non-zero determinant then the rank of the matrix is r . From either definition it is clear that the rank of A is less than or equal to the smaller of M and N .

► Determine the rank of the matrix

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

The largest possible square sub matrices of A must be of dimension 3×3 . Clearly, A possesses four such sub matrices, the determinants of which are given by

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

$$\begin{vmatrix} 1 & 0 & -2 \\ 2 & 2 & 2 \\ 4 & 3 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & -2 \\ 0 & 2 & 2 \\ 1 & 3 & 1 \end{vmatrix} = 0.$$

The next largest square sub matrices of A are of dimension 2×2 . Consider, for example, the 2×2 sub matrix formed by ignoring the third row and the third and fourth columns of A ; this has determinant

$$\begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} = 1 \times 0 - 2 \times 1 = -2.$$

Since its determinant is non-zero, A is of rank 2 and we need not consider any other 2×2 sub matrices.

In the special case in which the matrix A is a square $N \times N$ matrix, by comparing either of the above definitions of rank with our discussion of determinants, we see that $|A| = 0$ unless the rank of A is N . In other words, A is singular unless $R(A) = N$

4.6. Determinant and volume

Objectives

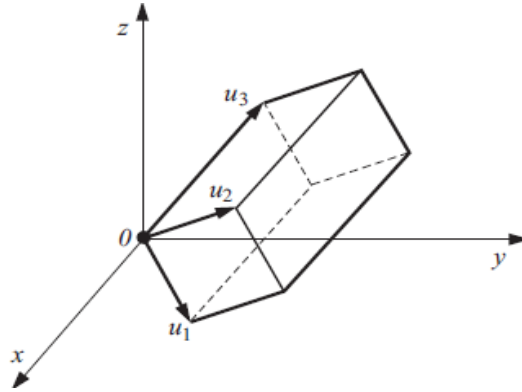
At the end of this section learners should be able to

- Find the volume of parallelepiped using determinant.

Determinants are related to the notions of area and volume as follows. Let $u_1, u_2, u_3, \dots, u_n$ be vectors in $\mathbb{R}^n$. Let S be the (solid) parallelepiped determined by the vectors, that is $S = \{a_1 u_1 + a_2 u_2 + a_3 u_3 + \dots + a_n u_n : 0 \leq a_i \leq 1 \text{ for } i = 1, 2, 3, \dots, n\}$ (When $n = 2$, S is a parallelogram.) Let $V(S)$ denote the volume of S (or area of S when $n = 2$) then $V(S) = \text{absolute value of det}(A)$, where A is the matrix with rows $u_1, u_2, u_3, \dots, u_n$. In general, $V(S) = 0$ iff the vectors $u_1, u_2, u_3, \dots, u_n$ do not form a coordinate system for $\mathbb{R}^n$. That is iff the vectors are linearly dependent.

Example: Let $u_1 = (1, 1, 0)$, $u_2 = (1, 1, 1)$, $u_3 = (0, 2, 3)$ then find the volume of $V(S)$ the parallelepiped S in $\mathbb{R}^3$ determined by the three vectors.

Solution: Diagrammatically the parallelepiped S in $\mathbb{R}^3$ determined by the three vectors is given by the figure below.



And the volume $V(S)$ is evaluated as the absolute value of the determinant of the matrix whose rows are the vectors u_1, u_2, u_3 . That is

$$V(S) = \text{the absolute value of } \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 3 \end{vmatrix} = |3 + 0 + 0 - 0 - 2 - 3| = |-2| = 2$$

Hence, $V(S) = 2$.

☺ **Test yourself!**

1. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at $(1, 0, -2)$, $(1, 2, 4)$, $(7, 1, 0)$.
2. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at $(1, 4, 0)$, $(-2, -5, 2)$, $(-1, 2, -1)$.
3. What is the volume of the parallelepiped with four of its vertices at $(0, 0, 0)$, $(1, 2, 2)$, $(2, 1, 2)$, and $(2, 2, 1)$? Where are the other four vertices?
4. Let $P(1, 0, 1)$, $Q(1, 1, 1)$, and $R(2, 2, 1)$. Choose S so that $PQRS$ is a parallelogram, and compute its area. Choose T, U, V so that $OPQRSTU$ is a tilted box, and compute its volume.

4.7. Eigen Value and Eigenvector of a Matrix

Objectives

At the end of this section learners should be able to

- Define Eigen value and Eigen vector.
- Find the Eigen values and corresponding Eigen spaces of a matrix.
- Show the relationship between Eigen values and its main diagonal of triangular matrix.

Definition of Eigen value and Eigen vector

Let A be an $n \times n$ matrix over a field F , then the scalar λ is called an Eigen value of A if there is a non zero vector $v \in F^n$ such that $Av = \lambda v$. The vector v is called an Eigen vector of A corresponding to λ .

✎ Remarks:

1. The term Eigen value is derived from the German word eigenwerte, meaning "proper value"
2. An eigenvector v cannot be zero. Allowing v to be the zero vector would render the definition meaningless because $A0 = \lambda 0$ is true for all values of λ . An Eigen value of $\lambda = 0$, however, is possible.
3. A matrix can have more than one Eigen value.

Example-1:

Let $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ then verify that $v_1 = (1, 0)$ and $v_2 = (0, 1)$ are eigenvectors of A corresponding to the Eigen values $\lambda_1 = 2$ and $\lambda_2 = -1$ respectively.

Solution:

Since $Av_1 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2v_1$, we have $v_1 = (1, 0)$ is an eigenvector of A corresponding to the Eigen values $\lambda_1 = 2$.

Also, since $Av_2 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1v_2$

We have $v_2 = (0, 1)$ is an eigenvector of A corresponding to the Eigen value $\lambda_2 = -1$.

Class Activity

1. If an Eigen value exists then show that it is unique.

2. Let $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ then verify that $v_1 = (-3, -1, 1)$ and $v_2 = (1, 0, 0)$ are eigenvectors of A and find their corresponding Eigen values.

Theorem: Eigenvectors of λ Form a subspace

Let A be an $n \times n$ matrix over field F with an Eigen value λ , then the set of all Eigen vectors of λ together with the zero vector,

$$\{0\} \cup \{v : v \text{ is an eigen vector of } A \text{ with eigen value } \lambda\}$$

is a subspace of F^n . we call this subspace the Eigen space of A belonging to λ

Proof:

Let A_λ be the set $\{0\} \cup \{v : v \text{ is an eigen vector of } A \text{ with eigen value } \lambda\}$

i. Since $0 \in A_\lambda$ we have $\phi \neq A_\lambda \subseteq F^n$.

ii. Let $v_1, v_2, \in A_\lambda$ then $Av_1 = \lambda v_1, Av_2 = \lambda v_2$

And hence $A(v_1 + v_2) = Av_1 + Av_2$ Why?

$$= \lambda v_1 + \lambda v_2$$

$$= \lambda(v_1 + v_2); \text{ why?}$$

$$\Rightarrow A(v_1 + v_2) = \lambda(v_1 + v_2)$$

$\Rightarrow v_1 + v_2$ is also an Eigen vector of A corresponding to λ

iii. Let $v \in A_\lambda$ and c be a non zero scalars then

$$Av = \lambda v \text{ and hence}$$

$$A(cv) = c(Av) = c(\lambda v) = \lambda(cv), \text{ Why?}$$

$$\Rightarrow A(cv) = \lambda(cv)$$

$\Rightarrow cv$ is also an eigenvector of A with Eigen value λ

$$\Rightarrow cv \in A_\lambda$$

In fact, if $c=0$ then $cv=0 \Rightarrow cv \in A_\lambda$

Hence, by (i), (ii) and (iii) we have A_λ is a subspace of F^n .

 Note

1. The sum of two Eigenvectors with the same Eigen value λ is also an Eigenvector with Eigen value λ .
2. A nonzero multiple of an Eigenvector with Eigen value λ is also an Eigenvector with Eigen value λ

Definition

- 1- A_λ is called the Eigen space or spectrum of A belonging to λ or Null space of $\lambda I - A = N(\lambda I - A)$.
- 2- The set of all Eigen values of A is called the spectrum of A and it is denoted by $\text{spec}(A)$. That is, $\text{Spec}(A) = \{\lambda: \lambda \text{ is an eigen value of } A\}$

Note: Intuitively, an Eigen vector 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 invariant under A .

Example:-2

Find the Eigen values and corresponding Eigen spaces of $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

Solution

i. Let $v_1 = (x, 0)$ then $Av_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} -x \\ 0 \end{bmatrix} = -1 \begin{bmatrix} x \\ 0 \end{bmatrix} = -1 v_1$

$\Rightarrow$ The eigenvectors corresponding to $\lambda_1 = -1$ are the nonzero vectors on the x-axis and hence

$A_{-1} = \{(x, 0) : x \in \mathbb{R} \setminus \{0\}\}$ is eigenspace of A belonging to eigen value $\lambda_1 = -1$.

ii. Let $v_2 = (0, y)$ then $Av_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} = 1 \begin{bmatrix} 0 \\ y \end{bmatrix} = 1 v_2$

$\Rightarrow$ The eigenvectors on the corresponding to $\lambda_2 = 1$ are the non zero vectors on the y-axis and hence

$A_1 = \{ (0, y) : y \in \mathbb{R} \setminus \{0\} \}$ is eigenspace of A belonging to eigenvalue $\lambda_2 = 1$.

iii. For $v = (x, y) \in \mathbb{R}^2$, we have

$$Av = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}, \text{ which is a reflection in the } y - \text{axis}.$$

Thus, the eigenspace corresponding to $\lambda_1 = -1$ is the x-axis and the eigenspace corresponding to $\lambda_2 = 1$ is the $y - \text{axis}$.

Question: How to determine the eigenvectors and corresponding eigenvalues of a matrix?

Theorem: Eigenvalues and eigenvectors of a matrix

Let A be an $n \times n$ matrix over a field F

- 1 - An eigen value of A is a scalar λ such that $\det(\lambda I_n - A) = 0$
- 2 - The eigen vectors of A corresponding to λ are the non zero solutions x of $(\lambda I_n - A)x = 0$.

Proof:

Let $x \in F^n$ be an eigen vector of A with eigen value λ then

$$\begin{aligned} Ax &= \lambda x \\ \Leftrightarrow Ax &= \lambda I_n x, \text{ since } I_n x = x \quad \forall x \in F^n \\ \Leftrightarrow (A - \lambda I_n)x &= 0 \\ \Leftrightarrow A - \lambda I_n &\text{ is singular, why?} \\ \Leftrightarrow |A - \lambda I_n| &= \det(A - \lambda I_n) = 0 \end{aligned}$$

This completes the proof.

 Note:

1. The equation $|A - \lambda I_n| = 0$ is called the characteristic equation of A .

And the matrix $A - \lambda I_n$ is called characteristic matrix of A

2. The determinantal polynomial $|A - \lambda I_n| = \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0$ is called the characteristic polynomial of A and it is denoted χ_A . This definition tells us that the eigenvalues of an $n \times n$ matrix A corresponds to the roots of the characteristic polynomial of A . Because the characteristic polynomial of A is of degree n , A can have at most n -distinct eigenvalues.
3. For each eigenvalue λ , the corresponding eigenvector is found by substituting λ back in to the characteristic equation $|A - \lambda I_n| = 0$ of A .

Example 3: Find the eigenvalues and corresponding eigenvectors of

$$(a) \ A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \quad (b) \ B = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Solution:

(a) The characteristic equation of A is

$$\begin{aligned} |A - \lambda I_2| &= 0 \\ \Leftrightarrow \left| \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| &= 0 \\ \Leftrightarrow \begin{vmatrix} 1-\lambda & 6 \\ 5 & 2-\lambda \end{vmatrix} &= 0 \\ \Leftrightarrow \lambda^2 - 3\lambda - 28 &= 0 \\ \Leftrightarrow (\lambda - 7)(\lambda + 4) &= 0 \\ \Leftrightarrow \lambda = 7 \text{ or } \lambda = -4. \end{aligned}$$

The corresponding eigenvectors are now found

Let $\lambda = 7$ then $(A - 7I_2)x = 0$

$$\begin{aligned} \Leftrightarrow \left[\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 7 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Leftrightarrow \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} -6x + 6y = 0 \\ 5x - 5y = 0 \end{cases} \\ \Leftrightarrow y &= x \end{aligned}$$

Hence, any vector of type $\beta (1, 1)$, where β is a non zero scalar, is an eigenvector corresponding to eigenvalue $\lambda = 7$.

Let $\lambda = -4$ then $(A + 4I_2)x = 0$

$$\begin{aligned} &\Leftrightarrow \left[\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} + 4 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right] \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Leftrightarrow \begin{bmatrix} 5 & 6 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} 5x + 6y = 0 \\ 5x + 6y = 0 \end{cases} \\ &\Leftrightarrow y = -\frac{5}{6}x \end{aligned}$$

Hence, any vector of type $\beta(1, -\frac{5}{6})$; where β is a non zero scalar, is an eigenvector corresponding to eigenvalue $\lambda = -4$

Try checking that $Ax = \lambda_i x$ for the eigenvalues and eigenvectors found in this example.

(b) Exercise

 **Remark:**

Steps in finding Eigen values and Eigen vectors

Let A be an $n \times n$ matrix over a field $\mathfrak{R}$

Step-1: Form the characteristic equation $|A - \lambda I_n| = 0$. It will be a polynomial equation of degree n in the variable λ .

Step-2: Find the real roots of the characteristic equation. These are the eigen values of A.

Step-3: For each eigen value λ_i , find the eigen vectors corresponding to λ_i by solving the homogeneous system

$(A - \lambda_i I_n) x = 0$. This may requires row-reducing an $n \times n$ matrix. The resulting reduced row-echelon form must have at least one row of zeros.

Class Activity:

- Find the eigen values and corresponding eigen vectors for the following matrix.
What is the dimension of the eigen space of each eigen value?

$$(a) A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (b) B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & -10 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

- For the matrix given in Q – 1 (a) and (b) find a basis for each of the corresponding eigen spaces.
- If A is an $n \times n$ triangular matrix, then show that its eigen values are the entries on its main diagonal.

4.8. Diagonalization of a symmetric matrix

Objectives

At the end of this section students are expected to:

- Recognize steps for diagonalizing of matrix.
- Compute an orthogonal matrix that diagonalizes a symmetric matrix.

A square matrix A is called **diagonalizable** if there is an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix; the matrix P is said to **diagonalize** A .

 **Remark:**

If A is an $n \times n$ matrix, then the following are equivalent.

- A is diagonalizable.
- A has n linearly independent eigenvectors.

Procedure for Diagonalizing a Matrix

The preceding remark guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and the proof provides the following method for Diagonalizing A .

Step 1: Find n linearly independent eigenvectors of A , say $P_1, P_2, P_3, \dots, P_n$.

Step 2: Form the matrix P having $P_1, P_2, P_3, \dots, P_n$ as its column vectors.

Step 3: The matrix $P^{-1}AP$ will then be diagonal with $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ as its successive diagonal entries, where λ_j is the eigenvalue corresponding to P_j for $j = 1, 2, 3, \dots, n$

 **Note:**

- If $v_1, v_2, v_3, \dots, v_k$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_k$ then $\{v_1, v_2, v_3, \dots, v_k\}$ is a linearly independent set.
- If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Diagonalization of Symmetric Matrices

If A is a symmetric matrix, then

- The eigenvalues of A are all real numbers.

(b) Eigenvectors from different eigenspaces are orthogonal

Steps in diagonalizing symmetric matrix

Step -1: Find a basis for each eigenspace of A .

Step -2: Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step -3: Form the matrix P whose columns are the basis vectors constructed in Step 2. This matrix orthogonally diagonalizes A .

Example: Find an orthogonal matrix P that diagonalizes

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Solution:

The characteristic equation of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 4 & -2 & -2 \\ -2 & \lambda - 4 & -2 \\ -2 & -2 & \lambda - 4 \end{bmatrix} = (\lambda - 2)^2 (\lambda - 8) = 0$$

Thus the eigenvalues of A are $\lambda_{1,2} = 2$ and $\lambda_3 = 8$

It can be shown that

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{form a basis for the}$$

eigenspace corresponding to $\lambda = 2$. Applying the Gram–Schmidt process to $\{\mathbf{u}_1, \mathbf{u}_2\}$

yields the following orthonormal eigenvectors (verify): $\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$

The eigenspace corresponding to $\lambda = 8$ has $u_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as a basis. Applying the Gram–

Schmidt process to $\{u_3\}$ yields $\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$. Finally, using $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ as column

vectors, we obtain $P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

Which orthogonally diagonalizes A . (As a check, the reader may wish to verify that $P^T A P$ is a diagonal matrix)

☺ Test yourself!

Orthogonally diagonalize the following matrices

1. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

3. $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$

4. $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

6. $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

7. $\begin{bmatrix} 3 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

9. $\begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$

Summary of Unit Four

1. Let $F = \mathbb{R}$ or $\mathbb{C}$ (the set of complex numbers) and $M_{n \times n}(F)$ be the space of $n \times n$ matrices over F then the determinant is a function which assigns to each square matrix A a number called the determinant of A and denoted by $\det A$ or $|A|$. Thus, $\det : M_{n \times n}(F) \rightarrow F$

Note

- I. $\det A$ can also be viewed as a function of the columns of A . In that case it is written as $\det (A_1, A_2, \dots, A_n)$
- II. Similarly, $\det A$ can be viewed as a function of the rows of A and written as $\det (A^1, A^2, \dots, A^n)$

2. Let $A = (a_{ij})$ be a square matrix of order n ($n \geq 2$) then the determinant of A can be expressed as:

$$\text{I. } |A| = \sum_{j=1}^n (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{j=1}^n a_{ij} c_{ij} \dots \dots i^{\text{th}} \text{ row expansion or}$$

$$\text{II. } |A| = \sum_{i=1}^n (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{i=1}^n a_{ij} c_{ij} \dots \dots j^{\text{th}} \text{ column expansion or}$$

Where A_{ij} is a matrix obtained from A by removing the i^{th} row and j^{th} column of A and $C_{ij} = (-1)^{i+j} |A_{ij}|$, which is the cofactor of a_{ij} .

3. Properties of $n \times n$ determinants

Let $A = (a_{ij})$ be a square matrix of order n over a field F then

a) Additivity in columns:

Suppose the k^{th} column A^k of A is the sum of two columns B^k and C^k then

$$\begin{aligned} \det (A^1, A^2, \dots, A^k, \dots, A^n) &= \det (A^1, A^2, \dots, B^k + C^k, \dots, A^n) \\ &= \det(A^1, A^2, \dots, B^k, \dots, A^n) + \det(A^1, A^2, \dots, C^k, \dots, A^n) \end{aligned}$$

b) Homogeneity in columns

$$\det(A^1, A^2, \dots, \lambda A^k, \dots, A^n) = \lambda \det(A^1, A^2, \dots, A^k, \dots, A^n) \forall \lambda \in F$$

c) If two adjacent columns of A are equal, say

$$A^k = A^{k+1} \text{ for some } 1 \leq k \leq n-1, \text{ then } \det A = 0$$

d) If two adjacent columns of A are interchanged then the determinant changes by sign. That is,

$$\det(A^1, A^2, \dots, A^k, A^{k+1}, \dots, A^n) = -\det(A^1, A^2, \dots, A^{k+1}, A^k, \dots, A^n) \text{ where } 1 \leq k \leq n-1$$

e) If two columns of A are equal, say $A^j = A^k$ for $j \neq k$, then $\det A = 0$

f) If a scalar multiple of one column is added to another column, then the value of the determinant does not change. That is,

$$= \det(A^1, A^2, \dots, A^j, \dots, A^k + \lambda A^j, \dots, A^n) = \det(A^1, A^2, \dots, A^j, \dots, A^k, \dots, A^n)$$

g) If any two columns of A are interchanged, then the determinant changes by sign.

h) If one of the columns of A is zero, then $\det A = 0$

i) If $A^1, A^2, \dots, A^n$ are linearly dependent, then $\det A = 0$

j) If A is an upper (lower) triangular matrix then $\det A = a_{11} a_{22} \dots a_{nn}$

4. Let $A = (a_{ij})_{n \times n}$ and $c_{ij} = (-1)^{i+j} |A_{ij}|$ then

a) the adjoint of A is given by:

$$\text{adj}(A) = [c_{ij}]_{n \times n}^t$$

b) $A [\text{adj}(A)] = [\text{adj}(A)] A = (\det A) I_n$

c) If $\det A \neq 0$, then $A^{-1} = \left[\frac{1}{\det A} \right] \text{adj}(A)$

5. Equivalent Conditions

Let A be an $n \times n$ matrix then the following statements are equivalent

- A is invertible
- $Ax = B$ has a unique solution for every $n \times 1$ matrix B
- $Ax = 0$ has only the trivial solution
- A is row equivalent to I_n
- A can be written as the product of elementary matrices.
- $|A| = \det(A) \neq 0$

6. Cramer's Rule states that a system of n-linear equations in n unknowns, $Ax = B$ with $|A| \neq 0$ has a unique solution given by:

$$x_j = \frac{\det(A^1, A^2, \dots, A^{j+1}, B, A^{j+1}, \dots, A^n)}{\det(A)}$$

Where $A = (a_{ij})_{n \times n}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

7. Let $A = (a_{ij})_{n \times n}$

- The determinant of any sub matrix of order k is called k^{th} order minor of A
- The rank of A is the order of the highest order non zero minor of A

8. Let $A = (a_{ij})_{m \times n}$

- Column rank of A is the dimension of the sub space of $\mathbb{R}^n$ generated by $A^1, A^2, \dots, A^n$ = The maximum number of linearly independent columns of A
- Row rank of A is the dimension of the subspace of $\mathbb{R}^m$ generated by $A_1, A_2, \dots, A_m$
= The maximum number of linearly independent rows of A.
- Row ranks of A = column rank of A = rank of A
- rank of A = n-dimension of solution space of $Ax = 0$

Miscellaneous Exercise-4

- Find the determinant of the following matrices:

$$\begin{aligned}
 (a) & \begin{bmatrix} 2 & 1 & 1 \\ 4 & 3 & 5 \\ 2 & 1 & 2 \end{bmatrix} & (b) & \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} & (c) & \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ -2 & -2 & 1 & 3 \\ 1 & 2 & -2 & 3 \end{bmatrix} \\
 (d) & \begin{bmatrix} -1 & 5 & 20 \\ 0 & 4 & 8 \\ 0 & 0 & 6 \end{bmatrix} & (e) & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 7 & 3 \end{bmatrix}
 \end{aligned}$$

3. Let A and B be 3×3 matrices such that $\det(A) = 4$ and $\det B = 5$ then find the value of

$$\begin{aligned}
 (a) & |AB| & (c) & |2AB| \\
 (b) & |3A| & (d) & |A^{-1}B|
 \end{aligned}$$

3. For each of the following compute A^{-1}

$$\begin{aligned}
 (a) & A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} & (b) & \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ -2 & 2 & -1 \end{bmatrix} & (c) & A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

4. Use Cramer's Rule to solve for each of the following systems

$$\begin{aligned}
 (a) & \begin{cases} x + 2y + 3z = 8 \\ 2x - y + z = 7 \\ -y + z = 1 \end{cases} & (b) & \begin{cases} x + y + 2z = 0 \\ 3x - y - z = 3 \\ 2x + 5y - 3z = 4 \end{cases}
 \end{aligned}$$

5. Evaluate

$$\begin{vmatrix} a & 1 & 1 & 1 & 1 \\ 1 & a & 1 & 1 & 1 \\ 1 & 1 & a & 1 & 1 \\ 1 & 1 & 1 & a & 1 \\ 1 & 1 & 1 & 1 & a \end{vmatrix}$$

6. a) Find the adjoint of

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

- b) If A is a triangular matrix, then show that $\text{adj}(A)$ is also triangular

7. (a) Let $A = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ e & f & g & h \\ x & y & z & w \end{bmatrix}$, then find $\text{rank}(A)$

(b) Let A be a square matrix of order n , then show that A is invertible if $\text{rank}(A) = n$

8. Let $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ be two vectors in $\mathbb{R}^3$, then show that their cross

Product $A \times B$ can be restated using determinants $A \times B = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$

Where $i = (1, 0, 0)$, $j = (0, 1, 0)$, $k = (0, 0, 1)$ are the standard unit vectors in $\mathbb{R}^3$

9. Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & b & c & d \\ e & f & g & h \end{bmatrix}$ $B = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Then prove that a) $\det A = \begin{vmatrix} c & d \\ g & h \end{vmatrix}$ b) $\det(B) = \begin{vmatrix} a & b & d \\ e & f & h \end{vmatrix}$

10. Let A and B square matrices of order n then prove or disprove the following

a) $\det(A + B) = \det(A) + \det(B)$

b) $\det[(A + B)^2] = [\det(A + B)]^2$

UNIT-FIVE: LINEAR TRANSFORMATIONS

Unit Objectives:

After studying this unit, you should be able to:

- Know linear transformation and state its properties
- Understand the kernel and Range of a linear transformation and find the kernel of a given linear transformation.
- Develop the composition of linear transformations
- Know the inverse of a given linear transformation.
- Know the standard matrix for the given linear transformation

5.1: Linear transformations and Examples

In this unit we discuss functions that map a vector space V in to a vector space W . This type of functions is denoted by $T: V \rightarrow W$. We use the standard function terminology for such functions. For instance, V is called the domain of T . If v is in V and w is in W such that $T(v) = w$, then w is called the image of v under T . The set of all images of vectors in V is called the range of T , and the set of all v in V such that $T(v) = w$ is called the preimage of w . For a vector $x = (x_1, x_2, \dots, x_n)$ in R^n , it would be technically correct to use double parentheses to denote $T(x)$ as $T(x) = T((x_1, x_2, \dots, x_n))$. For convenience, however, one set of parentheses is dropped, producing $T(x) = T(x_1, x_2, \dots, x_n)$

In this unit our interest centers on functions from vector space to another that preserve the operations of vector addition and scalar multiplication. We call such functions linear transformations.

Objectives

At the end of this section students are expected

- Determine whether the given function is a linear transformation or not.
- Show that a translation in the plane is not linear transformation.
- State and proof properties of linear transformation.
- Compute the linear transformation of functions.

Definition-5.1: Definition of a Linear Transformation

Let V and W be vector spaces over a field F (usually $F = R$), then the function

$T: V \rightarrow W$ is called a linear transformation (or a linear map) of V in to W if

(I) $T(u + v) = T(u) + T(v) \quad \forall u, v \in V$ and

(II) $T(au) = aT(u) \quad \forall u \in V$ and for all scalars $a \in F$

 Remarks

- (1) A linear transformation $T: V \rightarrow V$ from a vector space in to itself is called a linear operator.
- (2) we say that a linear transformation is operation preserving, because the same resul occurs whether the operations of addition and scalar multiplication are performed before or after the linear transformation is applied.
- (3) Although we use the same symbols to denote the vector operations in both V and W , you should note that the operations may be different, as indicated below.

$$\begin{array}{ccccccc}
 T(u + v) & = & T(u) + T(v) & & T(au) & = & cT(u) \\
 \uparrow & & \uparrow & & \uparrow & & \uparrow \\
 \text{Addition in } V & & \text{Addition in } W & & \text{Scalar} & & \text{Scalar} \\
 & & & & \text{multiplication in } V & & \text{multiplication in } W
 \end{array}$$

Example-1: Show that the following functions are linear operators

(a) $T: R^2 \rightarrow R^2$ defined by $T(x, y) = (y, x)$

(b) $T: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ given by $T(x, y) = (-iy, ix)$

Solution:

- (a) To show that the function T is a linear transformation, we must show that it preserves addition and scalar multiplication. To do this, we let

$u = (a, b)$ and $v = (x, y)$ be vectors in R^2 and c be any real number then

(I) Since $(u + v) = (a, b) + (x, y) = (a + x, b + y)$, we have $T(u + v) = T(a + x, b + y)$

$= (b + y, a + x), \text{ by definition of } T$

$= (a, b) + (y, x), \text{ by definition of '+' in } R^2$

$= T(a, b) + T(x, y), \text{ by definition of } T$

$= T(u) + T(v), \text{ by definition of } u \text{ \& } v$

$\Rightarrow T(u + v) = T(u) + T(v) \quad \forall u, v \in R^2 \text{ and}$

(II) Since $cu = c(a, b) = (ca, cb)$, we have $T(cu) = T(ca, cb) = (cb, ca)$, by definition of T

$$\begin{aligned}
 &= c(b, a), && \text{by definition of scalar multiplication in } R^2 \\
 &= cT(a, b), && \text{by definition of } T \\
 &= cT(u), && \text{by definition of } u
 \end{aligned}$$

$$\Rightarrow T(cu) = cT(u) \quad \forall c \in R \text{ and } \forall u \in R^2$$

Therefore, by (I) and (II) T is a linear transformation and hence T is a linear operator as $V = W = R^2$.

(b) Let $u = (z_1, z_2), v = (z_3, z_4) \in \mathbb{C}^2$ and $\beta \in \mathbb{C}$ then

(I) Since $u + v = (z_1, z_2) + (z_3, z_4) = (z_1 + z_3, z_2 + z_4)$,

we have $T(u + v) = T(z_1 + z_3, z_2 + z_4)$

$$\begin{aligned}
 &= (-i(z_2 + z_4), i(z_1 + z_3)) \\
 &= (-i z_2 - i z_4, i z_1 + i z_3) \\
 &= (-i z_2, i z_1) + (-i z_4, i z_3) \\
 &= T(z_1, z_2) + T(z_3, z_4) \\
 &= T(u) + T(v)
 \end{aligned}$$

$$\Rightarrow T(u + v) = T(u) + T(v) \quad \forall u, v \in \mathbb{C}^2 \text{ And}$$

(II)

Since $\beta u = \beta(z_1, z_2) = (\beta z_1, \beta z_2)$ we have $T(\beta u) = T(\beta z_1, \beta z_2) = (-i\beta z_2, i\beta z_1)$

$$\begin{aligned}
 &= \beta(-i z_2, i z_1) \\
 &= \beta T(z_1, z_2) = \beta T(u)
 \end{aligned}$$

$$\Rightarrow T(\beta u) = \beta T(u) \quad \forall \beta \in \mathbb{C} \text{ and } \forall u \in \mathbb{C}^2$$

Thus, by (I) and (II) we have T is a linear operator.

Class Activity 5.1

1. Let $T : R^2 \rightarrow R^2$ be defined by $T(x, y) = (x - y, x + 2y) \quad \forall (x, y) \in R^2$ then

- Show that T is a linear transformation
- Find the image of $v = (-1, 2)$
- Find the pre image of $w = (-1, 11)$

2. Let $T : R^2 \rightarrow R^3$ be defined by $T(x) = Ax$; where $Ax = \begin{bmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \forall (x, y) \in R^2$

(a) Show that T is a linear transformation from R^2 in to R^3

(b) If $x = (2, -1)$ then find $T(x)$

3. Construct functions that are not linear transformations.

Theorem 5-1: Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a function then T is a linear transformation if and only if

$$T(au + v) = aT(u) + T(v) \quad \forall a \in F \text{ and } \forall u, v \in V$$

Proof Exercise

Theorem 5.2: The Linear Transformation Given by a Matrix

Let A be an $m \times n$ matrix, then the function $T : R^n \rightarrow R^m$ defined by $T(x) = Ax$ is a linear transformation from R^n in to R^m . In order to conform to matrix multiplication with an $m \times n$ matrix, the vectors in R^n are represented by $n \times 1$ matrices and the vectors in R^m are represented by $m \times 1$ matrices.

Proof

(I) Let $A = (a_{ij})_{m \times n}$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be a vector in R^n then $T(x) = Ax$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}, \text{ which is a vector in } R^m.$$

$$\Rightarrow T(x) = Ax \in R^m$$

$$\Rightarrow T \text{ is well defined}$$

(II) Let x, y be vectors in R^n then $T(x + y) = A(x + y) = Ax + Ay$, by distributive

property of matrix multiplication

$$\begin{aligned}\Rightarrow T(x + y) &= Ax + Ay \\ &= T(x) + T(y) \quad \forall x, y \in R^n\end{aligned}$$

(III) Let x be a vector in R^n and $c \in R$ then $T(cx) = A(cx)$

$= c(Ax);$ by commutative property of scalar multiplication with matrix multiplication

$$= cT(x)$$

$$\Rightarrow Tc(x) = cT(x) \quad \forall c \in R \text{ and } \forall x \in R^n$$

Therefore, by (I), (II) and (III) we have T is a linear transformation from R^n in to R^m

Example-2:

$$A = \begin{bmatrix} 2 & -3 \\ -5 & 0 \\ 0 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 3 & 0 \\ 4 & 2 & 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 3 & 1 & 0 & 0 \end{bmatrix} \text{ then}$$

(a) The function defined by $T(x) = Ax$ is a linear transformation from R^2 in to R^3 .

Because the order of the matrix A is 3×2

(b) Since the order of the matrix B is 3×3 , we have the function $T : R^3 \rightarrow R^3$ defined by $T(x) = Bx$ is a linear transformation from R^3 in to R^3 .

(b) Since the order of the matrix C is 2×4 , we have the function

$T : R^4 \rightarrow R^2$ defined by $T(x) = Cx$ is a linear transformation from R^4 in to R^2 .

Example-3: Rotation in the Plane

Show that the linear transformation

$$T : R^2 \rightarrow R^2 \text{ given by the matrix } A = \begin{bmatrix} \cos Q & -\sin Q \\ \sin Q & \cos Q \end{bmatrix}$$

has the property that it rotates every vector in R^2 counterclockwise about the origin through the angle Q .

Solution:

From theorem 5.2, we know that T is a linear transformation. To show that it rotates every vector in R^2 counterclockwise through the angle Q , we let $v = (x, y)$ be a vector in R^2 , then using polar coordinates, we can write v as $v = (x, y) = (r \cos \beta, r \sin \beta)$,

where r is the length of v and β is the angle from the positive x -axis counterclockwise to the vector v .

Now, applying the linear transformation T to v produces

$$\begin{aligned} T(v) = Av &= \begin{bmatrix} \cos Q & -\sin Q \\ \sin Q & \cos Q \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos Q & -\sin Q \\ \sin Q & \cos Q \end{bmatrix} \begin{bmatrix} r \cos \beta \\ r \sin \beta \end{bmatrix} \\ &= \begin{bmatrix} r \cos Q \cos \beta - r \sin Q \sin \beta \\ r \sin Q \cos \beta + r \cos Q \sin \beta \end{bmatrix} = \begin{bmatrix} r \cos(Q + \beta) \\ r \sin(Q + \beta) \end{bmatrix} \end{aligned}$$

$\Rightarrow$ the vector $T(v)$ has the same length as v , which is r

Furthermore, since the angle from the positive x -axis to $T(v)$ is $Q + \beta$, $T(v)$ is the vector that results from rotating the vector v counterclockwise through the angle Q .

 Note:

- (1) The linear transformation in example 3 is called a rotation in $\mathbb{R}^2$. Rotation in $\mathbb{R}^2$ preserve both vector length and the angle between two vectors. That is, the angle between u and v is equal to the angle between $T(u)$ and $T(v)$.
- (2) The $m \times n$ zero matrix corresponds to the zero transformation from $\mathbb{R}^n$ in to $\mathbb{R}^m$ and the $n \times n$ identity matrix. I_n corresponds to the identity transformation from $\mathbb{R}^n$ in to $\mathbb{R}^n$.

 Test yourself!

1. Determine whether the given function is a linear transformation or not

- (a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x, y) = (2x + y, x - 3y)$
- (b) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $T(x, y, z) = (x + y, x - y, z)$
- (c) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $T(x, y) = (\sqrt{x}, xy, \sqrt{y})$
- (d) $T : P_2 \rightarrow P_2$ defined by $T(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$
- (e) $T : M_{2 \times 2} \rightarrow M_{2 \times 4}$ defined by $T(A) = AB$ where $B \in M_{2 \times 4}$

$$(f) \quad T : M_{3 \times 3} \rightarrow M_{3 \times 3} \text{ defined by } T(A) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A$$

2. (a) Let $T : M_{m \times n} \rightarrow M_{n \times m}$ defined by $T(A) = A^t$ then show that T is a linear transformation where $M_{m \times n}$ is the set of all m by n matrices

- (b) Show that $T : C[a, b] \rightarrow \mathbb{R}$ defined by $T(f) = \int_a^b f(x) dx$; where $C[a, b]$ is the set of all real valued continuous functions on $[a, b]$, is a linear transformation.
3. (a) Let $C'[a, b]$ be the set of all real valued functions whose derivatives are continuous on $[a, b]$ then show that the differential operator $D_x : C'[a, b] \rightarrow C[a, b]$ defined by $D_x(f) = \frac{d}{dx}(f) = f'(x) \forall f \in C'[a, b]$, defines a linear transformation.
- (b) For the linear transformation D_x defined in (a), find the pre image of the following functions (i) $f(x) = 2x + 1$ (ii) $f(x) = e^x$ (iii) $f(x) = \sin x$
4. Let T be a linear transformation from $C[0, 1]$ into $\mathbb{R}$ defined by $T(f) = \int_0^1 f(x) dx$ then find
- (a) $T(3x^2 - 2)$ (b) $T(x^3 - x^5)$ (c) $T(4x - 6)$
5. Let T be a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$ given by $T(x, y) = (x \cos Q - y \sin Q, x \sin Q + y \cos Q)$ then find
- (a) $T(4, 4)$ for $Q = 45^\circ$
 (b) $T(4, 4)$ for $Q = 30^\circ$
 (c) $T(5, 0)$ for $Q = 120^\circ$
6. Using the definition:
- A vector u is a fixed point of a linear transformation $T : V \rightarrow V$ if $T(u) = u$ then
- (a) Prove that 0 is a fixed point of any linear transformation $T : V \rightarrow V$
 (b) Prove that the set of all fixed points of a linear transformation $T : V \rightarrow V$ is a subspace of V .
 (c) Determine all fixed points of the linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by
 (i) $T(x, y) = (x, 2y)$ (ii) $T(x, y) = (y, x)$
7. Recall that a translation is a function of the form $T(x, y) = (x - h, y - k)$, where at least one of the constants h and k is non-zero then
- (a) show that a translation in the plane is not linear transformation
 (b) For the translation $T(x, y) = (x - 2, y + 1)$, determine the images of $(0, 0), (2, -1)$ and $(5, 4)$
 (c) Show that a translation in the plane has no fixed points.

8. For the linear transformation $T : R^3 \rightarrow R^3$ given by $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, called a projection in

R^3 find (a) $T(-1, 2, 2)$

(b) The pre image of $(1, 2, 0)$

(c) Explain why the vector $(1, 1, 1)$ has no pre image under this transformation.

Properties of Linear Transformations

Two simple linear transformations are the zero transformation and the identity transformation, which are defined as

(a) $T : V \rightarrow W$ defined by $T(v) = 0 \forall v \in V$ ----- zero transformation

(b) $T : V \rightarrow W$ defined by $T(v) = v \forall v \in V$ -----Identity transformation

Theorem 5.3: Properties of Linear Transformations

Let V and W be vector spaces over a field F and T be a linear transformation from V in to W then

1. $T(0) = 0$
2. $T(-v) = -T(v) \quad \forall v \in V$
3. $T(u - v) = T(u) - T(v) \quad \forall u, v \in V$
4. If $v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$, where $c_i \in F, v_i \in V$ then

$$T(v) = T\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i T(v_i)$$

Proof:

1. Since $0v = 0 \forall v \in V$, we have

$$\begin{aligned} T(0) &= T(0v) = 0T(v), \text{ since } T \text{ is linear} \\ &= 0, \text{ why?} \end{aligned}$$

$$\begin{aligned} \Rightarrow T(0) &= 0 \leftarrow \text{the zero vector in } W \\ &\quad \uparrow \text{zero vector in } V \end{aligned}$$

2. Since $-v = (-1)v \quad \forall v \in V$, we have

$$T(-v) = T[(-1)v] = (-1)T(v) = -T(v); \text{ why?}$$

3. Since $u - v = u + (-v) \quad \forall u, v \in V$, we have $T(u - v) = T(u + (-v)) = T(u) + T(-v)$,

since T is linear.

$$\begin{aligned} &= T(u) + -T(v), \text{ by 2 above} \\ &= T(u) - T(v) \end{aligned}$$

4. We apply induction on n .

$$\begin{aligned} \text{(i) If } n = 2, \text{ then } v &= c_1 v_1 + c_2 v_2 \text{ and } T(v) = T(c_1 v_1 + c_2 v_2) = T(c_1 v_1) + T(c_2 v_2) \\ &= c_1 T(v_1) + c_2 T(v_2) \end{aligned}$$

$$\Rightarrow T(v) = T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2)$$

(ii) Assume the assertion is true for n . That is, if $v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$ then

$$T(v) = c_1 T(v_1) + c_2 T(v_2) + \dots + c_n T(v_n) = \sum_{i=1}^n c_i T(v_i)$$

$$\text{(iii) Now, if } v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n + c_{n+1} v_{n+1} = \sum_{i=1}^{n+1} c_i v_i \text{ then}$$

$$\begin{aligned} T(v) &= T\left(\sum_{i=1}^{n+1} c_i v_i\right) \\ &= T\left(\sum_{i=1}^n c_i v_i + c_{n+1} v_{n+1}\right) \\ &= T\left(\sum_{i=1}^n c_i v_i\right) + c_{n+1} T(v_{n+1}), \text{ since } T \text{ is linear} \\ &= \sum_{i=1}^n c_i T(v_i) + c_{n+1} T(v_{n+1}), \text{ by induction assumption} \\ &= \sum_{i=1}^{n+1} c_i T(v_i) \end{aligned}$$

$$\Rightarrow T(v) = T\left(\sum_{i=1}^{n+1} c_i v_i\right) = \sum_{i=1}^{n+1} c_i T(v_i)$$

$\Rightarrow$ the assertion is also true for $n+1$.

Thus, by i), ii), and iii) we have the assertion is true for all $n \in \mathbb{N}$. That is, if $v =$

$$\sum_{i=1}^n c_i v_i \text{ Then } T\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i T(v_i) \quad \forall n \in \mathbb{N}.$$

Example-1: Show that the function $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x, y) = (x, y + 3)$ is not a linear transformation.

Solution: For if T is a linear transformation then by property-1 above, we get

$$T(0) = 0. \text{ But, } T(0) = 0$$

$$\Rightarrow T(0,0) = (0,0)$$

$$\Rightarrow (0,0+3) = (0,0), \text{ by definition of } T.$$

$$\Rightarrow (0,3) = (0,0)$$

$\Rightarrow 3 = 0$; This is contradiction. Thus, our assumption is wrong and hence T is not a linear transformation.

 **Note:** If any one of the properties in theorem 5.3 is not satisfied for a function $T: V \rightarrow W$, then the function T is not a linear transformation. Thus theorem 4.3 provides us a quick way to spot functions that are not linear transformations.

Example-2: Let $T: R^3 \rightarrow R^3$ be a linear transformation such that

$$T(1,0,0) = (2, -1, 4), T(0,1,0) = (1, 5, -2) \text{ and } T(0,0,1) = (0, 3, 1) \text{ then find } T(2,3,4)$$

Solution: Since we know that $\beta = \{(1,0,0), (0,1,0), (0,0,1)\}$ is a standard basis for

$$R^3 \text{ we have } \forall v \in R^3, v = (x, y, z) = x(1,0,0) + y(0,1,0) + z(0,0,1).$$

$$\text{In particular for, } v_1 = (2,3,4) = 2(1,0,0) + 3(0,1,0) + 4(0,0,1)$$

$$\Rightarrow v_1 = 2e_1 + 3e_2 + 4e_3 \text{ And hence by property-4 above we have}$$

$$\begin{aligned} T(v_1) &= T(2e_1 + 3e_2 + 4e_3) = 2T(e_1) + 3T(e_2) + 4T(e_3) \\ &= 2T(1,0,0) + 3T(0,1,0) + 4T(0,0,1) \end{aligned}$$

$$= 2(2, -1, 4) + 3(1, 5, -2) + 4(0, 3, 1)$$

$$= (4, -2, 8) + (3, 15, -6) + (0, 12, 4)$$

$$= (4 + 3 + 0, -2 + 15 + 12, 8 + -6 + 4)$$

$$= (7, 25, 6)$$

$$\Rightarrow T(2,3,4) = (7, 25, 6)$$

Theorem 5.4: Let V be a vector space over a field F , $S = \{v_1, v_2, v_3, \dots, v_n\}$ be a linearly dependent subset of V and $T: V \rightarrow V$ be a linear transformation then the set $\{T(v_1), T(v_2), \dots, T(v_n)\}$ is linearly dependent subset of V

Proof: Suppose S is linearly dependent subset of V then there exist scalars

$$\beta_1, \beta_2, \dots, \beta_n \in F, \text{ not all zero such that } \sum_{i=1}^n \beta_i v_i = 0,$$

$$\text{But } \sum_{i=1}^n \beta_i v_i = 0,$$

$$\Rightarrow T\left(\sum_{i=1}^n \beta_i v_i\right) = T(0)$$

$$\Rightarrow \sum_{i=1}^n \beta_i T(v_i) = 0, \text{ why?}$$

$$\Rightarrow \text{there exist scalars } \beta_1, \beta_2, \dots, \beta_n \text{ Not all zero such that } \sum_{i=1}^n \beta_i T(v_i) = 0$$

$\Rightarrow \{T(v_1), T(v_2), \dots, T(v_n)\}$ is a linearly dependent set. And hence the proof.

Theorem 5.5: let V and W be vector spaces over a field F such that

$\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for V and $w_1, w_2, \dots, w_n$ are arbitrary elements of W , then there exists a unique linear transformation $T : V \rightarrow W$ such that

$$T(v_i) = w_i \quad \forall_i = 1, 2, \dots, n.$$

Proof:

(i) Existence

Let $v \in V$ then $v = \sum_{i=1}^n \beta_i v_i$ with unique $\beta_i \in F \quad \forall_i = 1, 2, \dots, n$. Now, define

$$T : V \rightarrow W \text{ by } T(v) = T\left(\sum_{i=1}^n \beta_i v_i\right) = \sum_{i=1}^n \beta_i T(v_i) = \sum_{i=1}^n \beta_i w_i, \text{ Where } w_i = T(v_i) \forall_i \text{ then}$$

one can show that

(a) T is well defined

(b) $T(v_i) = w_i$

(c) T is linear transformation Thus, there exists a linear transformation

$$T : V \rightarrow W \text{ such that } T(v_i) = w_i \quad \forall_i = 1, 2, \dots, n$$

(ii) Uniqueness

Suppose $f : V \rightarrow W$ is also a linear transformation such that

$f(v_i) = w_i, \forall i = 1, 2, \dots, n$. Now, let $v \in V$ then $v = \sum_{i=1}^n \beta_i v_i$ and hence

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n \beta_i v_i\right) = \sum_{i=1}^n \beta_i f(v_i). \text{ Since } f \text{ is linear} \\ &= \sum_{i=1}^n \beta_i w_i, \text{ Since } f(v_i) = w_i \forall i = 1, 2, \dots, n \\ &= \sum_{i=1}^n \beta_i T(v_i), \text{ since } T(v_i) = w_i \forall i = 1, 2, \dots, n \\ &= T\left(\sum_{i=1}^n \beta_i v_i\right), \text{ since } T \text{ is linear} \\ &= T(v), \text{ since } v = \sum_{i=1}^n \beta_i v_i \end{aligned}$$

$$\Rightarrow f(v) = T(v) \forall v \in V.$$

$$\Rightarrow f = T$$

Therefore, T is unique and hence the proof.

 **Note:** Theorem 5.5. tells us that a linear transformation $T : V \rightarrow W$ is completely determined by its action on a basis of V . In other words, if

$$T(v_1), T(v_2), \dots, \text{ and } T(v_n)$$

are defined for a basis $\{v_1, v_2, \dots, v_n\}$ then T is defined for any v in V .

Test yourself!

1. Let T be a linear transformation from R^2 to R^2 such that

$$T(1,1) = (1,0) \text{ and } T(1,-1) = (0,1), \text{ then find}$$

$$(a) T(1,0) \quad (b) T(0,2)$$

2. Let $T : R^2 \rightarrow R^2$ be a linear transformation such that

$$T(1,0) = (1,1) \text{ and } T(0,1) = (-1,1) \text{ then find}$$

$$(a) T(1,4) \quad (b) T(-2, 1)$$

3. Let $T : P_2 \rightarrow P_2$ be a linear transformation such that

$$T(1) = x, T(x) = x+1 \text{ and } T(x^2) = 1+x+x^2, \text{ then find } T(2-6x+x^2)$$

Where P_2 is the set of all polynomials of degree less than or equal to 2.

4. Suppose $T : R^2 \rightarrow R^2$ such that $T(1,0) = (1,0)$ and $T(0,1) = (0,0)$ then

(a) Determine $T(x, y)$ for $(x, y) \in R^2$

(b) Give a geometric description of T

5. Suppose $f : R^2 \rightarrow R^2$ such that $f(1,0) = (0,1)$ and $f(0,1) = (1,0)$ then

(a) Determine $f(x, y)$ for $(x, y) \in R^2$

(b) Give a geometric description of f .

6. Let $S = \{v_1, v_2, v_3\}$ be a set of linearly independent vectors in R^3 . Find a linear transformation $T : R^3 \rightarrow R^3$ such that $\{T(v_1), T(v_2), T(v_3)\}$ is linearly dependent.

7. Let $e_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $e_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $e_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $e_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ then

(a) Show that $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ is a basis for $M_{2 \times 2}$.

(b) Let $T : M_{2 \times 2} \rightarrow M_{2 \times 2}$ be a linear transformation such that

$$T(e_{11}) = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, T(e_{12}) = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}, T(e_{21}) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \text{ and } T(e_{22}) = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix} \text{ then}$$

$$\text{Find (i) } T \left(\begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix} \right) \quad \text{(ii) } T \left(\begin{bmatrix} x & y \\ z & w \end{bmatrix} \right)$$

5.2. The Rank and Nullity of a Linear Transformation and example

Objectives

At the end of this section students should be able to

- Define Kernel and Image of a linear transformation.
- Define Rank and Nullity of a Linear Transformation.
- Define isomorphic of a linear transformation.
- Compute Kernel and image of linear transformations.
- Find basis and dimension of kernel in various subspaces.
- State and proof theorem of isomorphic.

Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation then we know that $T(0) = 0$. Now, before you go to the next section try to answer the following questions.

Q1. Do you think that there exist a non zero vector v such that $T(v) = 0$?

Q2. When do you say that $\{v \in V : T(v) = 0\} = \{0\}$?

Q3. Is $\{v \in V : T(v) = 0\}$ a subspace of V ?

Q4. Is $\{T(v) : v \in V\}$ a subspace of V ?

Definition 5.2:- Definition of kernel and Image of a linear transformation

Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation then

1). The set of all vectors v in V such that $T(v) = 0$ is called the kernel (or null space) of

T and is denoted by $\ker(T)$ or $N(T)$. That is, $\ker(T) = N(T) = \{v \in V : T(v) = 0\}$

2). The set of all vectors w in W such that $T(v) = w$ for some v in V is called the image

(or range) of T and is denoted by

$$\text{Im}(T) \text{ or } R(T) \text{ That is, } \text{Im}(T) = R(T) = \{w \in W : w = T(v) \text{ for some } v \in V\} \\ = \{T(v) : v \in V\}$$

 **Remark:** Since $T(0) = 0$ we have always $0 \in \ker(T)$ and $0 \in \text{Im}(T)$

Example-1

Find the Kernel of the following linear transformations.

(a) $T: R^2 \rightarrow R^3$ defined by $T(x, y) = (x - 2y, 0, -x)$

(b) $T_1: R^3 \rightarrow R^2$ defined by $T(x) = Ax$, where $A = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix}$

Solution:

(a) $\text{Ker}(T) = N(T) = \{v \in R^2 : T(v) = 0\}$

$$= \{(x, y) \in R^2 : T(x, y) = (0, 0, 0)\}$$

$$= \{(x, y) \in R^2 : (x - 2y, 0, -x) = (0, 0, 0)\}$$

$$= \{(x, y) \in R^2 : x - 2y = 0, -x = 0\}$$

$$= \{(x, y) \in R^2 : x = 0 = y\}$$

$$= \{(0, 0)\}, \text{ which is set containing the zero vector} = \{0\}$$

(b) $\ker(T_1) = N(T_1) = \{x \in R^3 : T_1(x) = 0\}$

$$\begin{aligned}
 &= \{x \in R^3 : Ax = 0\} \\
 &= \left\{ (x, y, z) \in R^3 : \begin{bmatrix} 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\
 &= \left\{ (x, y, z) \in R^3 : \begin{bmatrix} x - y - 2z \\ -x + 2y + 3z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\
 &= \{(x, y, z) \in R^3 : x - y - 2z = 0 \text{ and } -x + 2y + 3z = 0\} \\
 &= \{(x, y, z) \in R^3 : x = z, y = -z\} \\
 &= \{(z, -z, z) : z \in R\} = \{z(1, -1, 1) : z \in R\} = \text{span}\{(1, -1, 1)\}
 \end{aligned}$$

Observe that

- (1) $(0, 0, 0)$ $(1, -1, 1) \in \ker(T_1)$ and hence $\ker(T_1)$ contains non-zero vector and the zero vector.
- (2) $\ker(T_1)$ contains an infinite number of vectors.

Example-2: Find the range of the following linear transformations

- (a) $T : R^3 \rightarrow R^3$ defined by $T(x, y, z) = (0, 0, 0)$
- (b) $T : P_2 \rightarrow P_1$ given by $T(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$

Solution:

(a) $\text{Im}(T) = R(T) = \{w \in R^3 : w = T(v) \text{ for some } v \in R^3\}$

Now, if $w = (a, b, c) \in R(T)$ then $\exists v = (x, y, z) \in R^3$ such that $w = T(v)$.

But let $w = T(v) \Rightarrow (a, b, c) = T(x, y, z)$

$$\Rightarrow (a, b, c) = (0, 0, 0), \text{ by definition of } T$$

$$\Rightarrow a = 0 = b = c$$

$$\Rightarrow w = (a, b, c) = (0, 0, 0) \in \{0\}$$

Thus, $R(T) \subseteq \{0\} \dots R(*)$ also since T is a linear transformation we have $T(0) = 0$ and hence $\{0\} \subseteq R(T) \dots (**)$. Therefore, by $(*)$ and $(**)$ we get $R(T) = \{0\} = \{(0, 0, 0)\}$

(c) Claim: $R(T) = P_1$

Let $w \in P_1 \Rightarrow w = ax + b$; where $a, b \in \mathfrak{R}$ and $T(0 + bx + \frac{a}{2}x^2) = b + 2\frac{a}{2}x = b + ax = w$

$$\Rightarrow \forall w = ax + b \in P_1 \text{ there exist } v = 0 + bx + \frac{a}{2}x^2 \in P_2 \text{ such that } T(v) = w$$

$$\Rightarrow R(T) = P_1.$$

Class Activity 5.2: Find the kernel and image of the following linear transformations.

- (a) The projection $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ given by $T(x, y, z) = (x, y, 0)$
- (b) The identity transformation $T : V \rightarrow V$ given by $T(v) = v \quad \forall v \in V$
- (c) The zero transformation $T : V \rightarrow W$ given by $T(v) = 0 \quad \forall v \in V$.
- (d) $T : M_{m \times n} \rightarrow M_{n \times m}$ defined by $T(A) = A^t$
- (e) $T : P_2 \rightarrow \mathbb{R}$ be given by $T(p) = \int_0^1 P(x)dx$.

Theorem 5.6: Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation then

- (a) $\ker T$ is a subspace of the domain V
- (b) $\text{Im}(T) = R(T)$ is a subspace of W

Proof:

- (a) We know that $\ker(T) = \{v \in V : T(v) = 0\}$.
 - i) Since $T(0) = 0$ we have $0 \in \ker(T)$ and hence $\ker(T)$ is a nonempty subset of V .
 - ii) Let $u, v \in \ker(T)$ then $T(u) = 0, T(v) = 0$ and hence

$$T(u + v) = T(u) + T(v) = 0 + 0 = 0$$

$$\Rightarrow u + v \in \ker(T)$$

$$\Rightarrow \ker(T) \text{ is closed under addition.}$$
 - iii) Let $a \in F$ and $u \in \ker(T)$ then $T(u) = 0$ and hence

$$T(au) = aT(u) = a \cdot 0 = 0, \text{ why?}$$

$$\Rightarrow au \in \ker(T)$$

$$\Rightarrow \ker(T) \text{ is closed under scalar multiplication.}$$

Therefore, by i), ii), iii) we get $\ker(T)$ is a subspace of V .

(b) Exercise

Definition 5.3 Let A be an $m \times n$ matrix then

1. The row space of A is the subspace of $\mathbb{R}^n$ spanned by the row vectors of A .
2. The column space of A is the subspace of $\mathbb{R}^m$ spanned by the column vectors of A .
3. The dimension of the row (or column) space of a matrix A is called the rank of A and is denoted by $\text{rank}(A)$.

 **Remarks :** One can show that

- (a) If $A = (a_{ij})_{m \times n}$ is row equivalent to $B = (b_{ij})_{m \times n}$ then, the row space of A is equal to the row space of B .
- (b) If $A \overset{\text{row}}{\equiv} B$ where B is a matrix in row-echelon form, then the non-zero row vectors of B form a basis for the row space of A .
- (c) If A is an $m \times n$ matrix, then the row space and column space of A have the same dimension. That is, $\text{row rank}(A) = \text{column rank}(A)$.
- (d) Let A be an $m \times n$ matrix, then the set of all solutions of the homogeneous system of linear equations $Ax = 0$, $\{x \in R^n : Ax = 0\}$, is a subspace of R^n . We call this subspace the solution space of the system. It is also called the null space of the matrix A . Moreover, the dimension of the solution space of $Ax = 0$ is called the nullity of A .
- (e) Column space of A is equal to the row space of A^t and hence to find a basis for the column space of a matrix. We take transpose of A and then using elementary row operations to rewrite A^t in row-echelon form.

Corollary of Theorem 5.6: Let $T : R^n \rightarrow R^m$ be a linear transformation given by $T(x) = Ax$, where A is an $m \times n$ matrix then

- (a) $\ker(T)$ is equal to the solution space of $Ax = 0$
- (b) $R(T)$ is equal to the column space of A .

Proof: You can observe that it is an immediate consequence of definition 5.3, the above remark and theorem 5.6.

Example-3:

Let $T : R^5 \rightarrow R^4$ be defined by

$$T(x) = Ax, \text{ where } x \text{ is in } R^5 \text{ and } A = \begin{bmatrix} 1 & 2 & 0 & 1 & -1 \\ 2 & 1 & 3 & 1 & 0 \\ -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 8 \end{bmatrix}$$

then find

(a) Basis and dimension for $\ker(T)$ as a subspace of R^5

(b) A basis and dimension of $R(T)$ as a subspace R^4 .

Solution:

$$(a) \ker(T) = \{x \in R^5 : T(x) = 0\}$$

$$= \{(x_1, x_2, x_3, x_4, x_5) = (0, 0, 0, 0, 0)\}$$

But the equation $T = (x_1, x_2, x_3, x_4, x_5) = (0, 0, 0, 0, 0)$

$$\Rightarrow \begin{bmatrix} 1 & 2 & 0 & 1 & -1 \\ 2 & 1 & 3 & 1 & 0 \\ -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix};$$

which is a homogeneous system $Ax = 0$

Now, the associated augmented matrix of this system is

$$\Rightarrow \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & -1 & 0 \\ 2 & 1 & 3 & 1 & 0 & 0 \\ -1 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 8 & 0 \end{array} \right]$$

$\Rightarrow x_1 = -2x_3 + x_5, x_2 = x_3 + 2x_5$ and $x_4 = 4x_5$ letting $x_3 = s$ and $x_5 = t$ we have

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s + t \\ s + 2t \\ s \\ 4s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Thus, One basis for the kernel of T is given by $\beta = \{(-2, 1, 1, 0, 0), (1, 2, 0, -4, 1)\}$ and

$\dim [\ker(T)] = 2$

(b) We begin by taking A^t , since $R(T) = \{Ax : x \in R^5\}$ = column space of A
 = row space of A_t .

Now, $A_t = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 3 & -2 & 0 \\ 1 & 1 & 0 & 2 \\ -1 & 0 & 1 & 8 \end{bmatrix}$ then using Gauss-Jordan elimination A_t is reduced to the

following reduced row-echelon form

$\begin{bmatrix} 1 & 0 & 0 & -2 \leftarrow w_1 \\ 0 & 1 & 0 & 4 \leftarrow w_2 \\ 0 & 0 & 1 & 6 \leftarrow w_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ Then one basis for the range of T is given by the non-zero rows of

the reduced row-echelon form of A_t .

$\Rightarrow \beta = \{w_1, w_2, w_3\} = \{(1, 0, 0, -2), (0, 1, 0, 4), (0, 0, 1, 6)\}$ is a basis for $R(T)$ and

$$\dim(R(T)) = 3.$$

Definition 4.4: Definition of Rank and Nullity of a Linear Transformation

Let $T : V \rightarrow W$ is a linear transformation

- (a) The dimension of the kernel of T is called the nullity of T .
- (b) The dimension of the range of T is called the rank of T .

 **Note:** If T is given by a matrix A , then the rank of T is equal to the rank of A .

Theorem 5.7: Sum of Rank and Nullity

Let $T : V \rightarrow W$ be a linear transformation from an n -dimensional vector space V in to a vector space W , and then the sum of the dimensional of the range and kernel is equal to the dimension of the domain V . That is, $rank + nullity = n$ or $\dim(range) + \dim(kernel) = \dim(domain) = \dim(V)$

Proof: Exercise

Example-4: Let $T : R^5 \rightarrow R^7$ be a linear transformation then find

- (a) The Nullity of T if the rank of T is 2.
- (b) The Rank of T if $\ker(T) = \{0\}$
- (c) The Rank of T if the Nullity of T is 4.

Solution:

(a) Since by Theorem 5.7, $\text{Nullity}(T) + \text{rank}(T) = \dim(\text{domain of } T)$ we have

$$\begin{aligned}\text{Nullity}(T) + 2 &= \dim(R^5) \\ \Rightarrow \text{Nullity}(T) + 2 &= 5 \\ \Rightarrow \text{Nullity}(T) &= 5 - 2 = 3\end{aligned}$$

(b) Since $\text{rank}(T) + \text{Nullity}(T) = \dim(\text{domain } T)$. We have

$$\begin{aligned}\text{rank}(T) + \dim(\ker(T)) &= \dim(R^5) \\ \Rightarrow \text{rank}(T) &= \dim(R^5) - \dim(\ker(T)) \\ &= 5 - \dim(\{0\}) \\ &= 5 - 0 = 5.\end{aligned}$$

$$\begin{aligned}\text{(c) } \text{rank}(T) &= \dim(\text{domain}(T) - \text{Nullity}(T)) \\ &= 5 - 4 = 1\end{aligned}$$

One-to-One and On to Linear transformations

Recall that a function $T : V \rightarrow W$ is called

1. One to-one if the pre image of every w in the range consists of a single vector.

That is, T is one to one if $\forall u, v \in V$ such that $T(u) = T(v) \Rightarrow u = v$

2. Onto if every element in W has a pre image in V . In other words, T is onto

when $W = R(T)$. That is, T is onto if $\forall w \in W \exists v \in V$ such that $T(v) = w$.

Theorem 5.8: Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation, then

- (a) T is one-to-one if and only if $\ker(T) = \{0\}$.
- (b) If W is finite dimensional then T is onto if and only if $\text{rank}(T) = \dim(W)$
- (c) If $\dim(V) = \dim W = n$, then T is one-to-one if and only if it is onto.

Proof:

- (a) ($\Rightarrow$) Suppose T is one-to-one. We claim that $\ker(T) = \{0\}$. To see this, let

$u \in \ker(T)$, then $T(u) = 0$ and also we know that $T(0) = 0$.

Thus, $T(u) = 0 = T(0) \Rightarrow T(u) = T(0)$

$$\Rightarrow u = 0$$

$\Rightarrow \ker(T)$ only contains the zero vector and hence $\ker(T) = \{0\}$.

($\Leftarrow$) Suppose $\ker(T) = \{0\}$. We want to show that T is one-to-one.

To see this, let $u, v \in V$ such that $T(u) = T(v)$

$$\Rightarrow T(u) - T(v) = 0$$

$$\Rightarrow T(u - v) = 0$$

$$\Rightarrow u - v \in \ker(T)$$

$$\Rightarrow u - v \in \{0\}, \text{ since } \ker(T) = \{0\}$$

$$\Rightarrow u - v = 0$$

$$\Rightarrow u = v$$

Therefore, $T(u) = T(v) \Rightarrow u = v$ and hence T is one-to-one. Hence, the proof.

(b) Since T is on to if and only if $W = R(T)$ we have $\dim w = \dim[R(T)]$

$$\Leftrightarrow \dim W = \text{rank}(T)$$

$$W = R(T) \text{ we have } \dim w = \dim[R(T)] \Leftrightarrow \dim W = \text{rank}(T)$$

(c) ($\Rightarrow$) Suppose T is one to one then $\ker(T) = \{0\}$ and hence $\text{nullity}(T) = 0$.

$$\begin{aligned} \text{Thus, } \dim(\text{rang of } T) &= \dim(V) - \dim(\ker(T)) \\ &= n - 0 \\ &= n \\ &= \dim(W) \end{aligned}$$

$$\Rightarrow \dim[R(T)] = \dim(W)$$

$$\Rightarrow T \text{ is onto, by (b)}$$

$$\Rightarrow \dim[R(T)] = \dim(W)$$

$$\Rightarrow T \text{ is onto, by (b)}$$

($\Leftarrow$) Suppose T is onto, then

$$W = R(T) \text{ and hence } \dim[R(T)] = \dim(W) = n = \dim(V)$$

$$\text{Thus, } \dim[\ker(T)] = \dim(V) - \dim[R(T)]$$

$$= n - n = 0$$

$$\Rightarrow \dim[\ker(T)] = 0$$

$\Rightarrow \ker(T) = \{0\}$, as $\ker(T)$ is subspace of V .

$\Rightarrow T$ is one-to-one, by (a)

Class Activity 5.3:

1. Let $T : R^n \rightarrow R^m$ be a linear transformation defined by $T(x) = Ax$ where A is given below, then find the nullity and rank of T and determine whether T is one to one, onto, or neither

$$(a) A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (c) A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad (d) A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

2. Let T be a linear transformation given by $T(x) = Ax$ where A is given below then find (i) $\ker(T)$ (ii) $\text{Nullity}(T)$ (iii) $\text{rang}(T)$, and (iv) $\text{rank}(T)$

$$(a) A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (b) A = \begin{bmatrix} 2 & 2 & -3 & 1 & 13 \\ 1 & 1 & 1 & 1 & -1 \\ 3 & 3 & -5 & 0 & 14 \\ 6 & 6 & -2 & 4 & 16 \end{bmatrix} \quad (c) A = \begin{bmatrix} 3 & -2 & 6 & -1 & 15 \\ 4 & 3 & 8 & 10 & -14 \\ 2 & -3 & 4 & -4 & 20 \end{bmatrix}$$

$$(d) A = \begin{bmatrix} \frac{4}{9} & -\frac{4}{9} & \frac{2}{9} \\ -\frac{4}{9} & \frac{4}{9} & -\frac{2}{9} \\ \frac{2}{9} & -\frac{2}{9} & \frac{1}{9} \end{bmatrix}$$

Definition 5.5 Definition of Isomorphism

A linear transformation $T : V \rightarrow W$ which is both one-to-one and onto is called an isomorphism. Moreover, if there exists an isomorphism from V to W , then V and W are said to be isomorphic to each other.

Theorem 5.9: Isomorphic spaces and dimension

Two finite dimensional vector spaces V and W are isomorphic if and only if they are of the same dimension.

Proof: ($\Rightarrow$) Suppose V is isomorphic to W , where $\dim(V) = n$. then there exists a linear transformation $T : V \rightarrow W$ that is one-to-one and onto.

Now, T is one-to-one $\Rightarrow \ker(T) = \{0\}$

$$\Rightarrow \dim[\ker(T)] = 0$$

$$\Rightarrow \dim[R(T)] = \dim(V) - \dim[\ker(T)]$$

$$= n - 0 = n$$

$$\Rightarrow \dim[R(T)] = n \text{ --- (1)}$$

In addition, because T is onto, we can conclude that $R(T) = W$

$$\Rightarrow \dim(W) = \dim[R(T)]$$

$$= n, \text{ by (1)}$$

$$\Rightarrow \dim(W) = n = \dim V$$

$\Rightarrow W$ and V Have the same dimension.

($\Leftarrow$) Suppose $\dim V = \dim W = n$, finite, then there exist

$\beta = \{v_1, v_2, \dots, v_n\}$ and $\beta' = \{w_1, w_2, \dots, w_n\}$ which are basis of V and W

respectively. Now, let $v \in V$ then $v = \sum_{i=1}^n c_i v_i$ and define $T : V \rightarrow W$ by

$$T(v) = T\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i w_i \quad \text{then one can show that } T \text{ is a well defined}$$

linear transformation which is both one to one and onto. Thus V and W are isomorphic.

Example-5: the following vector spaces are isomorphic to each other.

(a) $R^4 = 4\text{-space}$

(b) $M_{4 \times 1} = \text{space of all } 4 \times 1 \text{ matrices}$

(c) $M_{2 \times 2} = \text{space of all } 2 \times 2 \text{ matrices}$

(d) $P_3 = \text{space of all Polynomials of degree 3 or less.}$

(e) $V = \{(x_1, x_2, x_3, x_4, 0) : x \in R\}$ Which is a subspace of R^5 .

Exercises 5.2

1. Find the kernel and image of the following linear transformations

(a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T(x, y) = (x - y, y - x)$

(b) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4, T(x, y, z, w) = (0, 0, 0, 0)$

(c) $T : P_3 \rightarrow \mathbb{R}, T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0$

(d) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3, T(x, y, z) = (z, x, y)$

2. Let T be a linear transformation given by $T(x) = AX$ where A is given below then find a basis and dimension for (i) $\ker(T)$ (ii) $R(T)$

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

(c) $A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 3 & 1 & 2 & -1 \\ -4 & -3 & -1 & -3 \\ -1 & -2 & 1 & 1 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 2 \\ -1 & -2 \\ 1 & 1 \end{bmatrix}$

3. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation. Use the given information to find

$N(T)$ and give a geometric description of the kernel and range of T .

(a) $\text{rank}(T) = 0$

(b) $\text{rank}(T) = 3$

(c) T is the counterclockwise rotation of 45° about the Z axis;

$T(x, y, z) = \left(\frac{\sqrt{2}}{2}x, \frac{-\sqrt{2}}{2}y, \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y, z \right)$

(d) T is the reflection through the yz -coordinate plane; $T(x, y, z) = (-x, y, z)$

(e) T is the projection on to the vector $v = (1, 2, 2)$; $T(x, y, z) = \frac{x+2y+2z}{9}(1, 2, 2)$

4. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation that projects u onto $v = (2, -1, 1)$

(a) Find the rank and nullity of T

(b) Find a basis for the kernel of T

(c) Repeat (a) & (b) for $v = (3, 0, 4)$

5. Which of the following vector spaces are isomorphic to $\mathbb{R}^6$?

(a) $M_{2 \times 3}$ (b) P_6 (c) $C[0, 6]$ (d) $M_{6 \times 1}$

(e) P_5 (f) $\{(x_1, x_2, x_3, 0, x_5, x_6, x_7) : x_i \in \mathbb{R}\}$

6. (a) Let $T : M_{n \times n} \rightarrow M_{n \times n}$ be given by $T(A) = A - A^t$, then show that $\ker(T)$ is the set of $n \times n$ symmetric matrices.

(b) Let $T : P_4 \rightarrow P_3$ be given by $T(f) = f'$. What is kernel of T ?

7. (a) Determine a relationship between m, n, j and k such that $M_{m \times n}$ is isomorphic to $M_{j \times k}$.

(b) Let B be an invertible $n \times n$ matrix. Prove that the linear transformation

$T : M_{n \times n} \rightarrow M_{n \times n}$ Given by $T(A) = AB$, is an isomorphism.

8. Let $A : V \rightarrow W$ be a linear transformation then prove that

(a) A^{-1} exists if and only if $N(A) = \{0\}$

(b) If A is one-to-one then $\dim(V) = \dim[R(A)]$

(c) If $\dim V = \dim W = n$, then $R(A) = W$ if and only if A^{-1} exists.

(d) If M is a subspace of V such that $V = M \oplus N(A)$ then the restriction of A to M is an isomorphism of M on to $R(A)$. That is, M and $R(A)$ are isomorphic.

5.3 Algebra of Linear transformations

Before you go to the next discussion, what do you say about the following questions?

1. Is the sum of two linear transformations also a linear transformation?
2. Is the difference of two linear transformations also a linear transformation?
3. Is the product of two linear transformations also a linear transformation?
4. What about the composition of two linear transformations?

Definition 5.6: Let T_1, T_2 be linear transformations from a vector space V over a field F into a vector space W over F and c be a scalar, then we define $T_1 + T_2$ and cT_1 as

$$(a) (T_1 + T_2)(v) = T_1(v) + T_2(v) \quad \forall v \in V$$

$$(b) (cT_1)(v) = cT_1(v) \quad \forall v \in V$$

Theorem 5.10

(1) $T_1 + T_2$ and cT_1 are linear transformations from V into W .

(2) $L(V, W)$ is a vector space over F where $L(V, W)$ is the set of all linear transformations from V in to W . That is,

$$L(V, W) = \{T : V \rightarrow W \mid T \text{ is linear transformation}\}$$

Proof:

(1) (i) Let $u, v \in V$ and $a \in F$ then

$$\begin{aligned} (T_1 + T_2)(au + v) &= T_1(au + v) + T_2(au + v), \text{by definition(a)} \\ &= [aT_1(u) + T_1(v)] + [aT_2(u) + T_2(v)], \text{since } T_1 \text{ and } T_2 \text{ are linear transformations} \\ &= [aT_1(u) + aT_2(u)] + [T_1(v) + T_2(v)] \\ &= a(T_1 + T_2)(u) + (T_1 + T_2)(v) \end{aligned}$$

$\Rightarrow T_1 + T_2$ is a linear transformation V in to W

(ii) Let $u, v \in V$ and $a \in F$ then $(cT_1)(au + v)$

$$\begin{aligned} &= (cT_1)(au + v), \text{by definition(b)} \\ &= c[aT_1(u) + T_1(v)], \text{since } T \text{ is linear} \\ &= ca T_1(u) + cT_1(v) \\ &= a(cT_1)(u) + (cT_1)(v) \end{aligned}$$

$\Rightarrow cT_1$ is also a linear transformation.

(2) Exercise Hint Try to show that $L(V, W)$ is a subspace of $\{f : V \rightarrow W \mid f \text{ is a function}\}$

Definition 5.7: Composition of linear Transformations

Let U, V and W be vector spaces over a field F and $T_1 : U \rightarrow V, T_2 : V \rightarrow W$ be linear transformations then the composition of T_1 with T_2 , denoted by $T_2 \circ T_1$, is defined by $(T_2 \circ T_1)(u) = T_2(T_1(u))$ for all u in U .

Remark:

(1) $T_2 \circ T_1 : U \rightarrow W$ And the composition is not defined unless the range of T_1 lies within the domain of T_2 .

- (2) Composition of linear transformation is not commutative.
- (3) Composition of linear transformation is associative.

Theorem 5.11: Let U, V and W be vector spaces over a field F and $T_1 : U \rightarrow V$ and $T_2 : V \rightarrow W$ be linear transformations then

- (1) $T_2 \circ T_1 : U \rightarrow W$ is also a linear transformation.
- (2) Let $L(V, V)$ be the set of all Linear transformation from V in to V then with the usual addition of functions and multiplication defined as function composition then $L(V, V)$ is a ring with unity.

Proof:

(1) (i) Let $u, v \in U$ then

$$\begin{aligned}
 (T_2 \circ T_1)(u + v) &= T_2(T_1(u + v)), \text{ By definition of composition} \\
 &= T_2(T_1(u) + T_1(v)), \text{ Since } T_1 \text{ linear} \\
 &= T_2(T_1(u)) + T_2(T_1(v)), \text{ Since } T_2 \text{ is linear} \\
 &= (T_2 \circ T_1)(u) + (T_2 \circ T_1)(v), \text{ By definition of composition}
 \end{aligned}$$

$$\Rightarrow (T_2 \circ T_1)(u + v) = (T_2 \circ T_1)(u) + (T_2 \circ T_1)(v) \quad \forall u, v \in U$$

(ii) Let $u \in U$ and $b \in F$ then

$$\begin{aligned}
 (T_2 \circ T_1)(bu) &= T_2(T_1(bu)), \text{ By definition of composition} \\
 &= T_2(b T_1(u)), \text{ Since } T_1 \text{ linear} \\
 &= b T_2(T_1(u)), \text{ Since } T_2 \text{ is linear} \\
 &= b(T_2 \circ T_1)(u), \text{ by definition of composition}
 \end{aligned}$$

$$\Rightarrow (T_2 \circ T_1)(bu) = b(T_2 \circ T_1)(u) \quad \forall b \in F \text{ and } \forall u \in U$$

Therefore, by (i) and (ii) we get $T_2 \circ T_1$ is a linear transformation.

(2) Exercise!

Example-1: Find $T_1 \circ T_2$ and $T_2 \circ T_1$ where the linear transformations T_1 and T_2 are given below.

(a) $T_1 : R^2 \rightarrow R^3, T_1(x, y) = (x, x + y, y)$

$T_2 : R^3 \rightarrow R^2, T_2(x, y, z) = (0, y)$

$$(b) T_1 : R \rightarrow R^2, T(x) = (x, 3x)$$

$$T_2 : R^2 \rightarrow R, T(x, y) = y + 2x$$

Solution:

$$\begin{aligned} (a) (i) (T_1 \circ T_2)(x, y, z) &= T_1(T_2(x, y, z)) = \\ &= T_1(0, y) = (0, 0 + y, y) = (0, y, y) \\ &\Rightarrow (T_1 \circ T_2)(x, y, z) = (0, y, y) \end{aligned}$$

$$\begin{aligned} (ii) (T_2 \circ T_1)(x, y) &= T_2(T_1(x, y)) = T_2(x, x + y, y) = (0, x + y) \\ &\Rightarrow (T_2 \circ T_1)(x, y) = (0, x + y) \end{aligned}$$

(b) Exercise!

Definition 5.8: Definition of Inverse Linear Transformation

Let V and W be vector spaces over a field F and $T_1 : V \rightarrow W$ and $T_2 : W \rightarrow V$ be linear transformations such that $(T_2 \circ T_1)(v) = v \forall v \in V$ and $(T_1 \circ T_2)(w) = w \forall w \in W$ then T_2 is called the inverse of T_1 and we write $T_2 = T_1^{-1}$.

 **Note** Not every linear transformation is invertible.

Theorem 5.12 Let $T : V \rightarrow W$ be a linear transformation then

- (a) T is invertible if and only if T is an isomorphism
- (b) If T^{-1} exists, then it is also a linear transformation

Proof Exercise!

Example-2: Determine whether the transformation T has an inverse or not. If it does, find T^{-1} .

$$(a) T : R^2 \rightarrow R^2 \text{ defined by } T(x, y) = (x, y + kx)$$

$$(b) T : R^3 \rightarrow R^3 \text{ defined by } T(x, y, z) = (x, y, 0)$$

Solution:

$$\begin{aligned} (a) \text{ since } \ker(T) &= \{(x, y) \in R^2 : T(x, y) = (0, 0)\} \\ &= \{(x, y) \in R^2 : (x, y + kx) = (0, 0)\} \\ &= \{(x, y) \in R^2 : x = 0 \text{ and } y + kx = 0\} \end{aligned}$$

$= \{(0,0)\} = \{0\}$ We have T is one-to-one and also

$\forall (a,b) \in R^2 \exists (a,b-ka) \in R^2$ such that

$$T(a,b-ka) = (a,b-ka+ka) = (a,b)$$

$\Rightarrow T$ is onto. Therefore, T is one-to-one and onto and hence T is invertible and its inverse is given $T^{-1}(x,y) = (x,y-kx)$, Infact you can check that $T^{-1} \circ T = T \circ T^{-1} = I$

(b) Since $(1, 2, 3) \in R^3$ and there doesn't exist $(x,y,z) \in R^3$ such that $T(x,y,z) = (1, 2, 3)$ we have T is not onto, and hence T is not isomorphism $\Rightarrow T$ is not invertible.

 **Note:** To see that there does not exist $(x,y,z) \in R^3$ such that $T(x,y,z) = (1, 2, 3)$. suppose $\exists (x,y,z) \in R^3$ such that $T(x,y,z) = (1, 2, 3)$.

$$\Rightarrow (x,y,0) = (1, 2, 3)$$

$$\Rightarrow 0 = 3; \text{ Which is contradiction and hence the fact}$$

Exercise 5.3

1. Let $T : V \rightarrow U$ and $S : U \rightarrow W$ be linear transformations then prove that

(a) If S and T are both one-to-one, then so is SoT .

(b) $\ker(T) \subseteq \ker(SoT)$

(c) If SoT is onto, then so is S

2. Let $S : V \rightarrow W$ and $T : V \rightarrow W$ be linear transformations then

$$\text{Prove that } \text{rank}(S+T) \leq \text{rank}(S) + \text{rank}(T)$$

3. Give an example of two linear transformations T and S such that

(a) $\text{Rank}(S+T) = \text{rank}(S) + \text{rank}(T)$

(b) $\text{Rank}(S+T) < \text{rank}(S) + \text{rank}(T)$

4. Determine whether the linear transformations represented by the matrix A is

(i) One-to-one

(ii) Onto

(iii) Invertible and if it does find its inverse

(a)

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & 1 \end{bmatrix}$$

$$(d) A = \begin{bmatrix} 4 & 0 & 7 \\ 5 & 5 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

5.4. Matrix Representation of a linear transformation

Brainstorming Question

Which of the following representations of $T : R^4 \rightarrow R^3$ is better?

$T(x, y, z, w) = (2x + y - z + 2w, x + 5y + 2z + 6w, 3x + 4y + w)$ or

$$T(v) = Av = \begin{bmatrix} 2 & 1 & -1 & 2 \\ 1 & 5 & 2 & 6 \\ 3 & 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

Answer: The second representation is better than the first for at least three reasons;

- (i) It is simpler to write
- (ii) It is simpler to read and
- (iii) It is more easily adapted for computer use.

Later we will see that matrix representation of a linear transformation also has some theoretical advantages. In this section we will see that for linear transformations involving finite dimensional vector spaces, matrix representation is always possible.

The key to representing a linear transformation $T : V \rightarrow W$ by a matrix is to determine how it acts on a basis of V . Once we know the image of every vector in the basis, we can use the properties of linear transformations to determine $T(v)$ for any v in V .

For convenience, first we will see that linear transformations from R^n into R^m , relative to the standard bases in R^n and R^m . Then at the end of this section we see linear transformations from R^n into R^m , relative to the non standard bases and linear transformations defined on general vector spaces.

Recall that the standard basis for R^n is given by $\beta = \{(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 1)\} = \{e_1, e_2, \dots, e_n\}$

Theorem 5.13: Standard Matrix for a Linear Transformation

Let $T : R^n \rightarrow R^m$ be a linear transformation such that

$$T(e_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \quad T(e_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Then the $m \times n$ matrix whose columns correspond to $T(e_i)$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \text{ is such that } T(v) = Av \quad \forall v \in R^n. \text{ } A \text{ is called the standard}$$

matrix for T.

Proof:

Claim- $Av = T(v) \quad \forall v \in R^n$,

to see this, let $V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in R^n$ then $V = v_1 e_1 + v_2 e_2 + \dots + v_n e_n$

$$\Rightarrow T(v) = T(v_1 e_1 + v_2 e_2 + \dots + v_n e_n)$$

$$= T(v_1 e_1) + T(v_2 e_2) + \dots + T(v_n e_n), \quad \text{Since } T \text{ is linear}$$

$$= v_1 T(e_1) + v_2 T(e_2) + \dots + v_n T(e_n), \text{ since } T \text{ is linear ----- (*)}$$

Also, the matrix product Av is given by

$$\begin{aligned} Av &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11} v_1 + a_{12} v_2 + \dots + a_{1n} v_n \\ a_{21} v_1 + a_{22} v_2 + \dots + a_{2n} v_n \\ \vdots \\ a_{m1} v_1 + a_{m2} v_2 + \dots + a_{mn} v_n \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
 &= v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\
 &= v_1 T(e_1) + v_2 T(e_2) + \cdots + v_n T(e_n)
 \end{aligned}$$

$\Rightarrow Av = v_1 T(e_1) + v_2 T(e_2) + \cdots + v_n T(e_n) \dots (**)$ Now, combining (*) and (**) we get

$$= T(v) = v_1 T(e_1) + v_2 T(e_2) + \cdots + v_n T(e_n) = Av$$

$$\Rightarrow T(v) = Av \quad \forall v \in R^n.$$

Example-1: Find the standard matrix for the linear transformation $T : R^3 \rightarrow R^2$ defined by

$$(a) T(x, y, z) = (x - 2y, 2x + y) \quad (b) T(x, y, z) = (x - 3y, z + 3x)$$

Solution:

(a) We begin by finding the images of e_1, e_2 and e_3 .

$$T(e_1) = T(1, 0, 0) = (1 - 2 \cdot 0, 2 \cdot 1 + 0) = (1, 2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$T(e_2) = T(0, 1, 0) = (0 - 2 \cdot 1, 2 \cdot 0 + 1) = (-2, 1) = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ and}$$

$$T(e_3) = T(0, 0, 1) = (0 - 2 \cdot 0, 2 \cdot 0 + 0) = (0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and}$$

Thus, by Theorem 5.13, the columns of A consists of

$T(e_1), T(e_2)$ and $T(e_3)$ and we have

$$A = [T(e_1) \mid T(e_2) \mid T(e_3)] = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \text{ and hence } T(v) = Av$$

$$\text{As a check we note that } Av = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 2y \\ 2x + y \end{bmatrix} \text{ which is equivalent}$$

to $T(x, y, z) = (x - 2y, 2x + y)$,

$$(b) T(e_1) = T(1, 0, 0) = (1, 3) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$T(e_2) = T(0, 1, 0) = (-3, 0) = \begin{bmatrix} -3 \\ 0 \end{bmatrix} \text{ and}$$

$$T(e_3) = T(0, 0, 1) = (0, 1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow A = [T(e_1) \mid T(e_2) \mid T(e_3)] = \begin{bmatrix} 1 & -3 & 0 \\ 3 & 0 & 1 \end{bmatrix} \text{ and hence } T(v) = Av.$$

Remarks:

1. A little practice will enable you to determine the standard matrix for a linear transformation by inspection. For instance, the standard matrix for the linear transformation given by $T(x, y, z) = (x - 2y + z, 5x + 3y, 4x + y - 2z)$ is found by using the coefficients of x, y and z to form the rows of A

That is, $T(x, y, z) = (x - 2y + z, 5x + 3y, 4x + y - 2z)$

$$= (1x - 2y + 1z, 5x + 3y + 0z, 4x + 1y - 2z)$$

$$\Rightarrow A = \begin{bmatrix} 1 & -2 & 1 \\ 5 & 3 & 0 \\ 4 & 1 & -2 \end{bmatrix}$$

2. The standard matrix for the zero transformation from R^n into R^m is the $m \times n$ zero matrix, and the standard matrix for the identity transformation from R^n into R^n is I_n .

Theorem 5.14

(a) Let $T_1 : R^n \rightarrow R^m$ and $T_2 : R^m \rightarrow R^p$ be linear transformations with standard matrices A_1 and A_2 then the standard matrix for the composition of T_1 with T_2 , $T_2 \circ T_1$ is given by the matrix product $A_2 A_1$. In general, if $T_1, T_2, \dots, T_n$ are linear transformations with standard matrices $A_1, A_2, \dots, A_n$ then the standard matrix of the composition $T_n \circ T_{n-1} \circ \dots \circ T_2 \circ T_1$ is given by the product $A_n, A_{n-1}, A_{n-2}, \dots, A_2, A_1$.

(b) Let $T : R^n \rightarrow R^n$ be a linear transformation with standard matrix A , then T is invertible if and only if A is invertible. Moreover, if T is invertible then the standard matrix for T^{-1} is A^{-1} .

Proof:

(a) Let $v \in R^n$ then

$$\begin{aligned}(T_2 \circ T_1)(v) &= T_2(T_1(v)) = T_2(A_1 v), \text{ Since } A_1 \text{ is standard matrix for } T_1 \\ &= A_2 (A_1 v), \text{ since } A_2 \text{ is standard matrix for } T_2 \\ &= (A_2 A_1)v, \text{ since Matrix multiplication is associative}\end{aligned}$$

$$\Rightarrow (T_2 \circ T_1)(v) = (A_2 A_1)v \quad \forall v \in R^n$$

$\Rightarrow$ the standard matrix for $T_2 \circ T_1$ is the matrix product $A_2 A_1$.

(b) $(\Rightarrow)$ Suppose T is invertible and A' is the standard matrix for T^{-1}

$$(T^{-1} \circ T)(v) = v = (T \circ T^{-1})(v)$$

$$\Rightarrow (A' A)v = I(v) = (A A')v$$

$$\Rightarrow (A' A)v = I_n v = (A A')v$$

$$\Rightarrow A' A = I_n = A A'$$

$\Rightarrow A'$ is the inverse of A

$$\Rightarrow A' = A^{-1}.$$

$\Rightarrow A$ is invertible

$(\Leftarrow)$ Suppose A is invertible with inverse A^{-1} then $A A^{-1} = I_n = A^{-1} A$

$$\Rightarrow (A A^{-1})v = I_n v = (A^{-1} A)v \quad \forall v \in R^n$$

$$\Rightarrow A(A^{-1}v) = I(v) = A^{-1}(A(v)) \quad \forall v \in R^n$$

$$\Rightarrow A(A^{-1}v) = I_n v = A^{-1}(A)v \quad \forall v \in R^n$$

$$\Rightarrow T(A^{-1}v) = I(v) = A^{-1}T(v), \text{ set } T = A^{-1}v$$

$$\Rightarrow T(T^{-1}v) = I(v) = T(T(v))$$

$$\Rightarrow (T \circ T^{-1})(v) = I(v) = (T \circ T^{-1})(v) \quad \forall v \in R^n$$

$$\Rightarrow T \circ T^{-1} = I = T^{-1} \circ T$$

$\Rightarrow T$ is invertible with inverse T^{-1}

$$\Rightarrow T^{-1} = T^{-1} \text{ and } T^{-1}(v) = T^{-1}(v) = A^{-1}v$$

$\Rightarrow$ the standard matrix for T^{-1} is A^{-1} . And hence the proof

(B) $(\Rightarrow)$ Suppose T is invertible and A^{-1} is the standard matrix for T^{-1} then $(T^{-1} \circ T)(v) = v = (T \circ T^{-1})(v)$

$$(v) = v = (T \circ T^{-1})(v)$$

$$\Rightarrow (A^1 \ A) V = I(V) = (A \ A^1) V, O$$

$$\Rightarrow (A^1 \ A) V = I_n V = (A \ A^1) V$$

$$\Rightarrow A^1 A = I_n = A A^1$$

$$\Rightarrow A^1 \text{ is the inverse of } A$$

$$\Rightarrow A^1 = A^{-1}.$$

$$\Rightarrow A \text{ is invertible}$$

($\Leftarrow$) Suppose A is invertible with inverse A^{-1} then $AA^{-1} = I_n = A^{-1}A$

$$\Rightarrow (A \ A^{-1}) V = I_n V = (A^{-1} \ A) V \quad \forall v \in R_n$$

$$\Rightarrow A(A^{-1} V) = I(V) = A^{-1}(AV) \quad \forall v \in R_n$$

$$\Rightarrow T(A^{-1} V) = I(V) = A^{-1}(T(v)), \text{ set } T^{-1} = A^{-1}V$$

$$\Rightarrow T(T^{-1} V) = I(V) = T^{-1}(T(v))$$

$$\Rightarrow (T \circ T^{-1})(v) = I(V) = (T^{-1} \circ T)(v) \quad \forall v \in R^n$$

$$\Rightarrow T \circ T^{-1} = I = T^{-1} \circ T$$

$$\Rightarrow T \text{ is invertible with inverse } T^{-1}$$

$$\Rightarrow T^{-1} = T^1 \text{ and}$$

$$T^{-1}(v) = T^1(v) = A^{-1}v$$

$$\Rightarrow \text{the standard matrix for } T^{-1} \text{ is } A^{-1}.$$

Example-2:

1- Let T_1 and T_2 be linear transformations from R^3 in to R^3 such that $T_1(x, y, z) = (2x + y, 0, x + z)$ and $T_2(x, y, z) = (x - y, z, y)$ then find the standard matrices for the compositions.

(a) $T_2 \circ T_1$

(b) $T_1 \circ T_2$

(2) Let $T : R^3 \rightarrow R^3$ be a linear transformations defined by $T(x, y, z) = (2x + 3y + z, 3x + 3y + z, 2x + 4y + z)$ then show that T is invertible, and find its inverse.

Solution:

1) Let A_1 and A_2 be the standard matrices for T_1 & T_2 respectively then.

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{then the standard matrix for}$$

$$(a) T_2 \circ T_1 \text{ is } A_2 A_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ and}$$

$$(b) T_1 \circ T_2 \text{ is } A_1 A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

2) Let A be a standard matrix for T then

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 3 & 1 \\ 2 & 4 & 1 \end{bmatrix} \quad \text{and you can show that}$$

$$A^{-1} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix}, \text{ by using elementary row operations}$$

$\Rightarrow A$ is invertible

$\Rightarrow T$ is invertible and the standard matrix for T^{-1} is A^{-1} More over, $T^{-1}(v) = A^{-1}v$

$$\Rightarrow T^{-1}(x, y, z) = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + y \\ -x + z \\ 6x - 2y - 3z \end{bmatrix}$$

$$\Rightarrow T^{-1}(x, y, z) = (-x + y, -x + z, 6x - 2y - 3z)$$

Transformation Matrix for Non- standard Bases

Let V and W be finite-dimensional vector spaces over a field F with bases β and β' , respectively, where $\beta = \{v_1, v_2, \dots, v_n\}$. If $T: V \rightarrow W$ is a linear transformation such that

$$[T(v_1)]\beta' = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, [T(v_2)]\beta' = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, [T(v_n)]\beta' = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Where $[T(v_i)] \beta'$ is the coordinate matrix of $T(v_i)$ relative to the basis β' , then the $m \times n$ Matrix whose n columns correspond to $[T(v_i)] \beta'$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \text{ is such that } T(v) = A[v] \beta \text{ for every } v \text{ in } V \text{ where}$$

$[v] \beta$ is the coordinate matrix of v relative to β . A is called the matrix of T relative to the bases β and β'

 **Note** $T(v) = A[v] \beta = [T(v)] \beta'$.

Example-3: Let $T: R^2 \rightarrow R^2$ be a linear transformation defined by $T(x, y) = (x + y, 2x - y)$ then find the matrix of T relative to the bases $\beta = \{(1, 2), (-1, 1)\} = \{v_1, v_2\}$ and $\beta' = \{(1, 0), (0, 1)\} = \{w_1, w_2\}$

(b) $T(v) =$ where $v = (2, 1)$ by using the matrix of T relative to the bases β and β' .

Solution

$$(a) \quad T(v_1) = T(1, 2) = (1 + 2, 2 \cdot 1 - 2) = (3, 0) = 3w_1 + 0w_2$$

$$\Rightarrow [T(v_1)] \beta' = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \text{ and}$$

$$T(v_2) = T(-1, 1) = (-1 + 1, 2(-1) - 1) = (0, -3) = 0w_1 - 3w_2$$

$$\Rightarrow [T(v_2)] \beta' = \begin{bmatrix} 0 \\ -3 \end{bmatrix}, \text{ which is the coordinate matrix of } T(v_2) \text{ relative to}$$

$$\beta' = \{(1, 0), (0, 1)\}.$$

$$\text{Thus, the matrix for } T \text{ relative to } \beta \text{ and } \beta' \text{ is } A = [[T(v_1)] \beta' \mid [T(v_2)] \beta'] = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$$

(b) Using the basis $\beta = \{(1, 2), (-1, 1)\}$, we find that $v = (2, 1) = 1(1, 2) - 1(-1, 1)$

$$\Rightarrow [v]_{\beta} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\text{Therefore, } T(v) = A[v]_{\beta} = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = [T(v)]_{\beta'}$$

Finally, Since $\beta' = \{(1, 0), (0, 1)\}$, it follows that $T(v) = 3(0, 1) = (3, 3)$

Note If $V=W$ and $\beta = \beta'$, then the matrix A is called the matrix of T relative to the basis β in such cases the matrix of the identity transformations simple I

To see this, let $\beta = \{v_1, v_2, \dots, v_n\}$. Since $I(v_i) = v_i$, we have

$$[T(v_1)]_{\beta} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, [T(v_2)]_{\beta} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, [T(v_n)]_{\beta} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

$$\Rightarrow A = [T(v_1)]_{\beta} | [T(v_2)]_{\beta} | \dots | [T(v_n)]_{\beta}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = I_n$$

Example-4:

Let $D_x : P_2 \rightarrow P_1$ be the differential operator that a quadratic polynomial on to its derivative

P' . Find the matrix for D_x used the bases $\beta = \{1, x, x^2\}$ and $\beta' = \{1, x\}$

Solution: The derivatives of the basis vectors are

$$D_x(1) = 0 = 0(1) + 0(x)$$

$$D_x(x) = 1 = 1(1) + 0(x)$$

$$D_x(x^2) = 2x = 0(1) + 2(x)$$

$\Rightarrow$ The coordinate matrices relative to

$$\beta' \text{ are } [D_x(1)]_{\beta'} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, [D_x(x)]_{\beta'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, [D_x(x^2)]_{\beta'} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\Rightarrow \text{The matrix for } D_x \text{ is } A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Moreover, for $P(x) = a + bx + cx^2 \in P_2$ we have $[P]_\beta = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and

$$A[P]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \end{bmatrix} = b + 2cx = D_x[a + bx + cx^2]$$

Exercises 5.4

1. Find the standard matrix for the linear transformation T

(a) $T(x, y) = (3x + 2y, 2y - x)$

(b) $T(x, y) = (4x + y, 0, 2x - 3y)$

(c) $T(x, y, z) = (5x - 3y, 2y + z, 2z + 4y, 5x + 3y)$ (d) $T(x, y, z, w) = (2x - z, 3y - 4w, 4z - x, y + w)$

2. Find the standard matrices for $T_2 \circ T_1$ & $T_1 \circ T_2$

(a) $T_1 : \mathbb{R}^3 \rightarrow \mathbb{R}^3, T_1(x, y, z) = (x + 2y, y - z, -2x + y + 2z)$
 $T_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^3, T_2(x, y, z) = (y + z, x + z, 2y - 2z)$

(b) $T_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^3, T_1(x, y) = (-x + 2y, x + y, x - y)$
 $T_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^2, T_2(x, y, z) = (x - 3y, z + 3x)$

3. Determine whether the given linear transformation is invertible? If it is, find its inverse.

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

(b) $T(x, y) = (x + y, 3x + 3y)$

(c) $T(x, y) = (2x, 0)$

(d) $T(x, y, z, w) = (x - 2y, y, y + Z, z)$

4. Find the matrix of T relative to the given bases β and β' .

(a) $T : P_2 \rightarrow P_3$ defined by $T(P) = xP$ and

$\beta = \{1, x, x^2\}, \beta' = \{1, x, x^2, x^3\}$

(b) $T : P_2 \rightarrow P_4$ defined by $T(P) = x^2P$

$\beta = \{1, x, x^2\}, \beta' = \{1, x, x^2, x^3, x^4\}$

5. Let $\beta = \{1, x, x^2, x^3\}$ be a basis for P_3 and let $T : P_3 \rightarrow P_4$ be the linear transformation given by $T(x^k) = \int_0^x t^k dt$.

(a) Find the matrix A for T with respect to β and the standard basis for P_4 .

(b) Use (a) to integrate $P(x) = 6 - 2x + 3x^3$.

6. (a) Find the standard matrix for T where $T(v) = kv$ for $v \in \mathbb{R}^n$.
- (b) Find the matrix for T relative to the standard bases for $M_{2 \times 3}$ and $M_{3 \times 2}$ where $T : M_{2 \times 3} \rightarrow M_{3 \times 2}$ defined by $T(A) = A^t$
- (c) Show that T defined in (b) is an isomorphism and find the matrix for the inverse of T .

5.5: Eigenvalues and Eigenvectors of a Linear Transformation

Eigenvectors and eigenvalues can be defined for linear operators as well as matrices. A scalar λ is called an *eigenvalue* of a linear operator $T : V \rightarrow V$ if there is a nonzero vector v in V such that $T(v) = \lambda v$. The vector v is called an *eigenvector* of T corresponding to λ . Equivalently, the eigenvectors of T corresponding to λ are the nonzero vectors in the kernel of $\lambda I - T$. This kernel is called the *eigenspace* of T corresponding to λ .

Note:

1. Eigen values are also called characteristic values or latent roots.
2. If $v \neq 0$ then λ is uniquely determined, because

$$\begin{aligned} \lambda_1 v = \lambda_2 v &\Rightarrow (\lambda_1 - \lambda_2)v = 0 \\ &\Rightarrow \lambda_1 = \lambda_2, \text{ as } v \neq 0 \end{aligned}$$
3. If v and w are eigen vectors with eigen value λ then
 - i. $v + w$ is also an eigen vector with eigen value λ because

$$T(v + w) = T(v) + T(w) = \lambda v + \lambda w = \lambda(v + w)$$
 - ii. Each scalar multiple cv is also an eigen vector with eigen value λ , because: $T(cv) = cT(v) = c(\lambda v) = \lambda(cv)$ where c is a constant.
4. Let $T : V \rightarrow V$ be a linear operator with

$$\ker(T) \neq \{0\}, \text{ If } 0 \neq v \in \ker(T) \text{ then}$$

$$T(v) = 0 = 0.v \text{ and hence } v \text{ is an eigen vector of } T \text{ with eigen}$$

value 0.

Example-

- a) Let $id : V \rightarrow V$ be the identity linear operator then every non zero vector in V is an eigen vector of id with an eigen value 1, since, $id(v) = v = 1.v$
- b) Let $D : V \rightarrow V$ be the differential operator on the vector space of differentiable functions. Since $D(e^{5t}) = 5 e^{5t}$ we have e^{5t} is an eigenvector of (D) with eigenvalue 5.
- c) Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear operator which rotates each $v \in \mathbb{R}^2$ by an angle of 90° then one can see that $\forall \lambda \in \mathbb{R}, T(v) \neq \lambda v$. Therefore, T has no eigenvectors and hence no eigenvalues.

Exercises 5.5

1. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation, whose matrix relative to the non standard

basis $\beta = \{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$ is $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ then find

- a) the characteristic polynomial of T
 - b) the eigen values of T
 - c) the corresponding eigen space of T
2. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by $T(v) = \text{Proj}_u$, where u is a fixed vector in $\mathbb{R}^2$ then
 - a) find the characteristic polynomial of T
 - b) Show that the eigen values of A (the standard matrix of T) are 0 and 1.
3. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear operator defined by $T(x, y, z) = (x + 3y, 3x + y, -2z)$ then find the characteristic polynomial of T
4. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a rotation of $\mathbb{R}^2$ counter clock wise about the origin through an angle Q then find
 - a) The characteristic polynomial P of T
 - b) The eigen values of T

5.6: Eigenspace of a linear transformation

The Eigen space of a linear operator

Let $T : V \rightarrow V$ be a linear operator over a field F and $\lambda \in F$ then

V_λ = The set of all eigenvectors of T with eigenvalue λ is a subspace of V

Proof: Exercise

☞ **Note** V_λ is called the eigenspace or spectrum of T belonging to λ .

Theorem 1: Let V be a vector space over a field F and $T : V \rightarrow V$ be a linear operator then

- a) $\lambda \in F$ is an eigen value of T if and only if the operator $T - \lambda I$ is singular.
- b) If $v_1, v_2, \dots, v_n$ are eigen vectors of T belonging to distinct eigen values $\lambda_1, \lambda_2, \dots, \lambda_n$, then $v_1, v_2, \dots, v_n$ are linearly independent

Proof:

- a) Let λ is an eigen value of T then $\exists$ a non zero vector $v \in V$ such that
$$T(v) = \lambda v$$

$$\begin{aligned}
 &\Leftrightarrow T(v) = (\lambda I)v \\
 &\Leftrightarrow (T - \lambda I)(v) = 0 \\
 &\Leftrightarrow v \in \ker(T - \lambda I) \\
 &\Leftrightarrow \ker(T - \lambda I) \neq \{0\} \Leftrightarrow T - \lambda I \text{ is not one to one} \\
 &\Leftrightarrow T - \lambda I \text{ is singular (not invertible)}
 \end{aligned}$$

b) Exercise.

Example:

Let $T: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ be defined by $T(x, y, z) = (2x + y - 2z, 2x + 3y - 4z, x + y - z)$. Find all eigenvalues of T , and find a basis of each eigenspace. Is T diagonalizable? If so, find the basis S of $\mathbf{R}^3$ that diagonalizes T , and find its diagonal representation D .

First find the matrix A that represents T relative to the usual basis of $\mathbf{R}^3$ by writing down the coefficients of x, y, z as rows, and then find the characteristic polynomial of A (and T). We have

$$A = [T] = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{aligned} \operatorname{tr}(A) &= 4, & |A| &= 2 \\ A_{11} &= 1, & A_{22} &= 0, & A_{33} &= 4 \\ & \sum_i A_{ii} &= 5 \end{aligned}$$

Therefore $\Delta(t) = t^3 - 4t^2 + 5t - 2 = (t - 1)^2(t - 2)$, and so $\lambda = 1$ and $\lambda = 2$ are the eigenvalues of A (and T). We next find linearly independent eigenvectors for each eigenvalue of A .

(i) Subtract $\lambda = 1$ down the diagonal of A to obtain the matrix

$$M = \begin{bmatrix} 1 & 1 & -2 \\ 2 & 2 & -4 \\ 1 & 1 & -2 \end{bmatrix}, \quad \text{corresponding to} \quad x + y - 2z = 0$$

Here y and z are free variables, and so there are two linearly independent eigenvectors belonging to $\lambda = 1$. For example, $u = (1, -1, 0)$ and $v = (2, 0, 1)$ are two such eigenvectors.

(ii) Subtract $\lambda = 2$ down the diagonal of A to obtain

$$M = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & -4 \\ 1 & 1 & -3 \end{bmatrix}, \quad \text{corresponding to} \quad \begin{aligned} y - 2z &= 0 \\ 2x + y - 4z &= 0 \\ x + y - 3z &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x + y - 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

Only z is a free variable. Here $w = (1, 2, 1)$ is a solution.

Thus T is diagonalizable, since it has three independent eigenvectors. Specifically, choosing

$$S = \{u, v, w\} = \{(1, -1, 0), (2, 0, 1), (1, 2, 1)\}$$

as a basis, T is represented by the diagonal matrix $D = \operatorname{diag}(1, 1, 2)$.

Exercise 5.6

1-

Prove the following for a linear operator (matrix) T :

- (a) The scalar 0 is an eigenvalue of T if and only if T is singular.
- (b) If λ is an eigenvalue of T , where T is invertible, then λ^{-1} is an eigenvalue of T^{-1} .

2-

- (a) We have that 0 is an eigenvalue of T if and only if there is a vector $v \neq 0$ such that $T(v) = 0v$, that is, if and only if T is singular.
- (b) Since T is invertible, it is nonsingular; hence, by (a), $\lambda \neq 0$. By definition of an eigenvalue, there exists $v \neq 0$ such that $T(v) = \lambda v$. Applying T^{-1} to both sides, we obtain

$$v = T^{-1}(\lambda v) = \lambda T^{-1}(v), \quad \text{and so} \quad T^{-1}(v) = \lambda^{-1}v$$

Therefore λ^{-1} is an eigenvalue of T^{-1} .

3- Let $T : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ be a linear operator represented by the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ with

respect to the standard basis of $\mathbb{C}^2$ then find

- a) The characteristic polynomial of T
- b) The minimal polynomial of T
- c) The eigen values of T
- d) Is T diagonalizable?

4- Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear operator represented by the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ with

respect to the standard basis of $\mathbb{R}^2$ then find

- a) The characteristic polynomial of T
- b) The minimal polynomial M of T
- c) The eigen values of T

5-Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by

$T(x, y, z) = (0, x + y, y + z)$ be a linear transformation then find

- a) The characteristic polynomial of T
- b) Eigenvalues and Eigenvectors

Summary of Unit Five

- Let V and W be vector spaces over a field F (usually $F=\mathbb{R}$), then the function $T:V \rightarrow W$ is called a linear transformation (or a linear map) of V in to W if
 $T(u + v) = T(u) + T(v) \quad \forall u, v \in V$ and
 $T(au) = aT(u) \quad \forall u \in V$ and for all scalars $a \in F$

- The above two properties can be combined in the formula given by
 $T(au + bv) = aT(u) + bT(v) \quad \forall a, b \in F$ and $\forall u, v \in V$

- By induction we also have a more generalization property for a linear transformation T for n set of vectors v_i as :

$$T\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i T(v_i) \text{ for scalars } c_i$$

- Let V and W be vector spaces over a field F and $T : V \rightarrow W$ be a linear transformation then The set of all vectors v in V such that $T(v) = 0$ is called the kernel (or null space) of T and is denoted by $\ker(T)$ or $N(T)$. The set of all vectors w in W such that $T(v) = w$ for some v in V is called the image (or range) of T and is denoted by $m(T)$ or $R(T)$.
- Let $T : V \rightarrow W$ is a linear transformation then the dimension of the kernel of T is called the nullity of T and the dimension of the range of T is called the rank of T .
- A function $T : V \rightarrow W$ is called one to-one if the pre image of every w in the range consists of a single vector and onto if every element in W has a pre image in V .

- Let T_1, T_2 be linear transformations from a vector space V over a field F in to a vector space W over F and c be a scalar then $T_1 + T_2$ and cT_1 are linear transformations from V in to W and $L(V, W)$ is a vector space over F where $L(V, W)$ is the set of all linear transformations from V in to W
- Let V and W be vector spaces over a field F and $T_1 : V \rightarrow W$ and $T_2 : W \rightarrow V$ be linear transformations such that $(T_2 \circ T_1)(v) = v \forall v \in V$ and $(T_1 \circ T_2)(w) = w \forall w \in W$ then T_2 is called the inverse of T_1 and we write $T_2 = T_1^{-1}$.
- Let $T_1 : R^n \rightarrow R^m$ and $T_2 : R^m \rightarrow R^p$ be linear transformations with standard matrices A_1 and A_2 then the standard matrix for the composition of T_1 with T_2 , $T_2 \circ T_1$ is given by the matrix product $A_2 A_1$.
In general, if $T_1, T_2, \dots, T_n$ are linear transformations with standard matrices $A_1, A_2, \dots, A_n$ then the standard matrix of the composition $T_n \circ T_{n-1} \circ \dots \circ T_2 \circ T_1$ is given by the product $A_n A_{n-1} A_{n-2} \dots A_2 A_1$.
- Let $T : R^n \rightarrow R^n$ be a linear transformation with standard matrix A , then T is invertible if and only if A is invertible. Moreover, if T is invertible then the standard matrix for T^{-1} is A^{-1} .
- Let A be an $n \times n$ matrix over a field F ($F = \mathbb{R}$ or $\mathbb{C}$) then a scalar λ is called an eigenvalue of A if there is a non zero vector $x \in F^n$ such that $Ax = \lambda x$. The vector x is called the eigenvector of A corresponding to λ .
- Let V be a vector space over a field F ($F = \mathbb{R}$ or $\mathbb{C}$) and $T : V \rightarrow V$ be linear transformation then $v \in V$ is said to be an eigenvector of T if $\exists \lambda \in F$ such that $T(v) = \lambda v$. λ is called an eigenvalue of T belonging to the eigenvector v . We also say that v is an eigenvector with eigenvalue λ .
- Let V be a vector space over a field F and $T : V \rightarrow V$ be a linear transformation then $V_\lambda =$ the set of all eigenvector of T with eigenvalue λ . is a subspace of V .

- Let A be an $n \times n$ matrix then the characteristic polynomial of A is given by $x_A(t) = |tI - A| = t^n + c_{n-1}t^{n-1} + \dots + c_1t + c_0$ and the characteristic equation of A is given by $x_A(t) = 0$. That is, $|tI - A| = 0$
- λ is an eigenvalue of A if $x_A(\lambda) = 0$
- The eigenvectors of A corresponding to λ are the non zero solutions of $(\lambda I - A)x = 0$
- If A is triangular matrix then its eigenvalues are the elements on its main diagonal

References

1. Bernard Kolman & David R. Hill, Elementary linear algebra, 8th ed., Prentice Hall, 2004
2. Larson/Edwards, Elementary Linear Algebra, D.C. Heath and Company, Lexington, 1988
3. H. Anton and C. Rorres, Elementary linear algebra 8th ed, John Wiley & Sons, INC., 1994
4. K. Hoffman & R. Kunze, Linear Algebra, 2nd ed., Prentice Hall INC. 1971
5. S. Lipschutz, Theory and problems of linear algebra, 2nd ed. McGraw-Hill 1991
6. D. C. Lay, Linear algebra and its applications, Pearson Addison Wesley, 2006
7. Demissu Gameda, An Introduction to Linear Algebra, Department of Mathematics, AAU, 2000
8. Serge Lang; Linear Algebra
9. Schaum's Outline in Linear Algebra
10. Piage and swift: Linear Algebra
11. Beaumont: Linear Algebra
12. Halms: Finite Dimensional Vector space
13. Nomizu: Fundamentals of Linear Algebra
14. Isaak and Manougian, Basic Concepts of Linear Algebra, 1st ed., George J. McLead Limited, 1976
15. J.N. Sharma and et al, Linear Algebra, Krishna Prakashan Media(P) Ltd., 2003
16. Otto Bretscher, Linear algebra with application, 3rd ed., Prentice Hall, 2005