DIRECTORATE OF DISTANCE EDUCATION

UNIVERSITY OF NORTH BENGAL

MASTER OF SCIENCES- MATHEMATICS

SEMESTER -III

LINEAR ALGEBRA

DEMATH3OLEC1

BLOCK-1

UNIVERSITY OF NORTH BENGAL

Postal Address:

The Registrar,

University of North Bengal,

Raja Rammohunpur,

P.O.-N.B.U., Dist-Darjeeling,

West Bengal, Pin-734013,

India.

Phone: (O) +91 0353-2776331/2699008

Fax: (0353) 2776313, 2699001

Email: regnbu@sancharnet.in ; regnbu@nbu.ac.in

Wesbsite: www.nbu.ac.in

First Published in 2019

FOREWORD

The Self Learning Material (SLM) is written with the aim of providing simple and organized study content to all the learners. The SLMs are prepared on the framework of being mutually cohesive, internally consistent and structured as per the university's syllabi. It is a humble attempt to give glimpses of the various approaches and dimensions to the topic of study and to kindle the learner's interest to the subject

We have tried to put together information from various sources into this book that has been written in an engaging style with interesting and relevant examples. It introduces you to the insights of subject concepts and theories and presents them in a way that is easy to understand and comprehend.

We always believe in continuous improvement and would periodically update the content in the very interest of the learners. It may be added that despite enormous efforts and coordination, there is every possibility for some omission or inadequacy in few areas or topics, which would definitely be rectified in future.

We hope you enjoy learning from this book and the experience truly enrich your learning and help you to advance in your career and future endeavours.

LINEAR ALGEBRA

BLOCK-1

Unit 1: Matrix Representation Of Linear Transformations. 6

Unit 2: System Of Linear Equations 24

Unit-3 Solution Set Of Linear Equation...44

Unit-4 Linear Transformation And Determinant 61

Unit-5 Vector Spaces In Linear Transformation 90

Unit-6 Vector Space In Linear Transformation Ii 107

Unit-7 Linear Transformation 125

BLOCK-2

Unit 8: Dual Space And Linear Transformation

Unit 9: Inner Product Space

Unit 10: Quadratic Forms

Unit 11: Jordan Cannonical Form

Unit 12: Bilinear Forms

Unit 13: Annihilating Polynomials & Decomposition - I

Unit 14: Annihilating Polynomials & Decomposition - Ii

BLOCK-1 LINEAR ALGEBRA

The study of linear algebra includes the topics of vector algebra, **matrix** algebra, and the theory of vector spaces. Linear algebra originated as the study of linear equations, including the solution of simultaneous linear equations. An equation is linear if no **variable** in it is multiplied by itself or any other variable.

Two important concepts emerge in linear algebra to help facilitate the expression and solution of systems of simultaneous linear equations. They are the vector and the matrix. Vectors correspond to directed line segments. They have both magnitude (length) and direction. Matrices are rectangular arrays of numbers. They are used in dealing with the coefficients of simultaneous equations. Using vector and matrix notation, a system of linear equations can be written, in the form of a single equation, as a matrix times a vector.

Linear algebra has a wide variety of applications. It is useful in solving network problems, such as calculating current flow in various branches of complicated electronic circuits, or analyzing traffic flow patterns on city streets and interstate highways. Linear algebra is also the basis of a process called linear programming, widely used in business to solve a variety of problems that often contain a very large number of variables.

UNIT 1: MATRIX REPRESENTATION OF LINEAR TRANSFORMATIONS.

STRUCTURE

1.0 Objective

1.1 Introduction

1.2 Matrices

1.3 Special Matrices

1.4 Operations On Matrices

1.5 Multiplication Of Matrices

1.6 Inverse Of Matrix

1.7 Some Special Matrix

1.8 Summary

1.9 Keywords

1.10 Questions for review

1.11 Suggested Readings

1.12 Answers To Check Your Progress

1.0 OBJECTIVE

Understand the basic concept of matrices

Learn different operation on matrices

Understand how to multiply the matrices

Understand the inverse and special matrices

1.1 INTRODUCTION

The matrix has a long history of application in solving linear equations. They were known as arrays until the 1800's. The term "matrix" (Latin for "womb", derived from *mater*—mother) was coined by James Joseph Sylvester in 1850, who understood a matrix as an object giving rise to a number of determinants today called minors, that is to say, determinants of smaller matrices that are derived from the original one by removing columns and rows. An English mathematician named Cullis was the first to use modern bracket notation for matrices in 1913 and he simultaneously demonstrated the first significant use of the notation $A=a_{i,j}$ to represent a matrix where $a_{i,j}$ refers to the element found in the *i*th row and the *j*th column. Matrices can be used to compactly write and work with multiple linear equations, referred to as a system of linear equations, simultaneously. Matrices and matrix multiplication reveal their essential features when related to linear transformations, also known as linear maps.

1.2 MATRICES

Definition of a Matrix

Definition 1.1.1. A rectangular array of numbers is called a **matrix**. The horizontal arrays of a matrix are called its **rows** and the vertical arrays are called its **columns**. Let *A* be a matrix having *m* rows and *n* columns. Then, *A* is said to have **order** $m \times n$ or is called a matrix of **size** $m \times n$ and can be represented in either of the following forms:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \text{ or } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

where *aij* is the entry at the intersection of the *i*th row and *j*th column.

One writes $A \in \mathbb{M}_{m,n}(\mathbb{C})$

Notes

to say that A is an $m \times n$ matrix with complex entries, $A \in \mathbb{M}_{m,n}(\mathbb{R})$ to say that A is an $m \times n$

matrix with real entries and $A = [aij]$, when the order of the matrix is understood from the context. We will also use $A[i, :]$ to denote the i-th row of A, $A[:, j]$ to denote the j-th column of A and aij or $(A)ij$, for the (i, j)-th entry of A

For example consider

$$A = \begin{bmatrix} 1 & 3+\mathbf{i} & 7 \\ 4 & 5 & 6-5\mathbf{i} \end{bmatrix}$$

then $A[1, :] = [1\ 3 + \mathbf{i}\ 7]$,

$$A[:, 3] = \begin{bmatrix} 7 \\ 6-5\mathbf{i} \end{bmatrix}$$

and $a_{22} = 5$. In general, in row vector commas are inserted to differentiate between entries. Thus, $A[1, :] = [1, 3 + \mathbf{i}, 7]$. A matrix having only one column is called a **column vector** and a matrix with only one row is called a **row vector**. All our vectors will be column vectors and will be represented by bold letters. Thus, $A[1, :]$ is a row vector and $A[:, 3]$ is a column vector.

Example: The system of linear equations $2x + 3y = 5$ and $3x + 2y = 6$ can be identified with the matrix

$$A = \left[\begin{array}{cc|c} 2 & 3 & 5 \\ 3 & 2 & 6 \end{array}\right]$$

Note that x and y are variables with the understanding that x is associated with $A[:, 1]$ and y is associated with $A[:, 2]$.

Definition 1.1.3. Two matrices $A = [aij]$, $B = [bij] \in \mathbb{M}_{m,n}(\mathbb{C}))$ are said to be **equal** if $a_{ij} = b_{ij}$, for each $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. In other words, two matrices are said to be equal if they have the same order and their corresponding entries are equal.

1.3 SPECIAL MATRICES

Definition 1.2.1. 1. A matrix in which each entry is zero is called a **zero-matrix**, denoted **0**.

For example,

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \mathbf{0}_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2. A matrix that has the same number of rows as the number of columns, is called a **square matrix**. A square matrix is said to have order n if it's order is $n \times n$ and is denoted either by writing $A \in \mathbb{M}_{m,n}(\mathbb{R})$ or $A \in \mathbb{M}_{m,n}(\mathbb{C})$, depending on whether the entries are real or complex numbers, respectively.

3. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$,

(a) Then, the entries $a_{11}, a_{22}, \ldots, ann$ are called the diagonal entries and they constitute the **principal diagonal** of A.

(b) Then, A is said to be a **diagonal** matrix if $a_{ij} = 0$ for $i \neq j$, denoted diag$(a_{11}, \ldots, ann)$.

For example, the zero matrix $\mathbf{0}_n$ and $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ are two diagonal matrices.

(c) If $A = \text{diag}(a_{11}, \ldots, a_{nn})$ and $a_{ii} = d$ for all $i = 1, \ldots, n$ then the diagonal matrix A is called a **scalar matrix**.

(d) Then, $A = \text{diag}(1, \ldots, 1)$ is called the **identity matrix**, denoted In, or in short I.

For example, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

For $1 \leq i \leq n$, define $\mathbf{e}i = I_n[:, i]$, a matrix of order $n \times 1$. Then, the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, where $\boldsymbol{e_i} \in \mathbb{M}_{m,1}(\mathbb{C})$, for $1 \leq i \leq n$, is called the **standard basis** of $\mathbb{C}^n$. Note that even though the order of the column vectors $\mathbf{e}_i$'s depend on n, we don't mention it as the size is understood from the context. For example, if $\mathbf{e}_1 \in \mathbb{C}^2$ then, e_1^T= [1, 0]. If $\mathbf{e}1 \in \mathbb{C}^3$ then, e_1^T= 1 = [1, 0, 0] and so on.

5. Let $A = [a_{ij}]$ be a square matrix.

(a) Then, A is said to be an **upper triangular** matrix if $a_{ij} = 0$ for $i > j$.
(b) Then, A is said to be a **lower triangular** matrix if $a_{ij} = 0$ for $i < j$.
(c) Then, A is said to be **triangular** if it is an upper or a lower triangular matrix.

For example, $\begin{bmatrix} 0 & 1 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix}$ is upper triangular, $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is lower triangular and the

6. An $m \times n$ matrix $A = [aij]$ is said to have an **upper triangular form** if $aij = 0$ for all $i > j$. For example, the matrices have upper triangular forms.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

1.4 OPERATIONS ON MATRICES

Definition 1.3.1. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$,

1. Then, the **transpose** of A, denoted $A^T = [b_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$,and $b_{ij} = a_{ji}$, for all i, j.
2. Then, the **conjugate transpose** of A, denoted $A^* = [c_{ij}] \in \mathbb{M}_{n,m}(\mathbb{C})$,and $c_{ij} = \overline{a_{ji}}$, for all i, j, where for $a \in \mathbb{C}$, $\overline{a}$ denotes the complex-conjugate of a.

Thus, if **x** is a column vector then $\mathbf{x}^T$ and $\mathbf{x}^*$ are row vectors and vice-versa. For example, if

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } A^* = A^T = \begin{bmatrix} 1 & 0 \\ 4 & 1 \\ 5 & 2 \end{bmatrix}, \text{ whereas if } A = \begin{bmatrix} 1 & 4+\mathrm{i} \\ 0 & 1-\mathrm{i} \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 1 & 0 \\ 4+\mathrm{i} & 1-\mathrm{i} \end{bmatrix}$$

$$\text{and } A^* = \begin{bmatrix} 1 & 0 \\ 4-\mathbf{i} & 1+\mathbf{i} \end{bmatrix}.$$

Note that $A^* \neq A^T$.

Theorem 1.3.2. *For any matrix A,* $(A^*)^* = A$. *Thus,* $(A^T)^T = A$.

Proof. Let $A = [a_{ij}]$, $A^* = [b_{ij}]$ and $(A^*)^* = [c_{ij}]$. Clearly, the order of A and $(A^*)^*$ is the same. Also, by definition $c_{ij} = \overline{b_{ji}} = \overline{\overline{a_{ij}}} = a_{ij}$ for all i, j and hence the result follows.

Definition 1.3.3. Let $A = [a_{ij}]$, $B = [b_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$,M$m,n$(C). Then, the **sum** of A and B, denoted $A + B$, is defined to be the matrix $C = [c_{ij}] \in \mathbb{M}_{m,n}(\mathbb{C})$,with $c_{ij} = a_{ij} + b_{ij}$.

$$\text{For example, if } A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix} \text{ then } 5A = \begin{bmatrix} 5 & 20 & 25 \\ 0 & 5 & 10 \end{bmatrix} \text{ and } (2+\mathrm{i})A = \begin{bmatrix} 2+\mathrm{i} & 8+4\mathrm{i} & 10+5\mathrm{i} \\ 0 & 2+\mathrm{i} & 4+2\mathrm{i} \end{bmatrix}.$$

Definition 1.3.4. Let $A = [aij] \in \mathbb{M}_{m,n}(\mathbb{C})$,Then, the **product** of $k \in$ C with A, denoted kA,
is defined as $kA = [ka_{ij}] = [a_{ij}k] = Ak$.

Theorem 1.3.5. *Let A, B, C* $\in \mathbb{M}_{m,n}(\mathbb{C})$ *and let k, l* $\in \mathbb{C}$. *Then,*

1. $A + B = B + A$ *(commutativity).*

2. $(A + B) + C = A + (B + C)$ *(associativity).*

3. $k(lA) = (kl)A$.

4. $(k + l)A = kA + lA$.

Proof. Part 1.

Let $A = [a_{ij}]$ and $B = [b_{ij}]$. Then, by definition
$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$

as complex numbers commute.

Definition 1.2.6. Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$
1. Then, the matrix $\mathbf{0}m\times n$ is called the **additive identity** as $A + \mathbf{0} = \mathbf{0} + A = A$.

2. Then, there exists a matrix B with $A + B = \mathbf{0}$. This matrix B is called the **additive inverse** of A, and is denoted by $-A = (-1)A$.

Check your progress

1. Define the following

a. Square matrix

b. Upper Triangular matrix

c. Lower Triangular matrix [HINT: Provide the definition with example]

2. Define the product of matrices and state its properties [HINT: Provide the definition and statement of its properties]

1.5 MULTIPLICATION OF MATRICES

Definition 1.4.1. Let A = [aij] ∈ $\mathbb{M}_{m,n}(\mathbb{C})$ and B = [bij] ∈ $\mathbb{M}_{n,r}(\mathbb{C})$. Then, the **product** of A and B, denoted AB, is a matrix C = [cij] ∈ $\mathbb{M}_{m,r}(\mathbb{C})$ with

For example,

$$\text{if } A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \text{ and } B = \begin{bmatrix} \alpha & \beta & \gamma & \delta \\ x & y & z & t \\ u & v & w & s \end{bmatrix} \text{ then} \quad (1)$$

$$AB = \begin{bmatrix} a\alpha + bx + cu & a\beta + by + cv & a\gamma + bz + cw & a\delta + bt + cs \\ d\alpha + ex + fu & d\beta + ey + fv & d\gamma + ez + fw & d\delta + et + fs \end{bmatrix}.$$

Note that the rows of the matrix AB can be written directly as

(AB)[1, :] = a [α, β, γ, δ] + b [x, y, z, t] + c [u, v, w, s] = aB[1, :] + bB[2, :] + cB[3, :]
(AB)[2, :] = dB[1, :] + eB[2, :] + fB[3, :]
(2)

and similarly, the columns of the matrix AB can be written directly as

$$(AB)[:,1] = \begin{bmatrix} a\alpha + bx + cu \\ d\alpha + ex + fu \end{bmatrix} = \alpha\ A[:,1] + x\ A[:,2] + u\ A[:,3], \quad (3)$$

(AB)[:, 2] = β A[:, 1] + y A[:, 2] + v A[:, 3], · · · , (AB)[:, 4] = δ A[:, 1] + t A[:, 2] + s A[:, 3].

Remark 1.4.2. Observe the following:

1. In this example, while AB is defined, the product BA is not defined. However, for square matrices A and B of the same order, both the product AB and BA are defined.

2. The product AB corresponds to operating (adding or subtracting multiples of different rows) on the rows of the matrix B(see Equation (1.2.2)). This is **row method** for calculating the matrix product.

3. The product AB also corresponds to operating (adding or subtracting multiples of different columns) on the columns of the matrix A (see Equation (1.2.3)). This is **column method** for calculating the matrix product.

4. Let A and B be two matrices such that the product AB is defined. Then, verify that

(a) Then, verify that (AB)[i, :] = A[i, :]B. That is, the i-th row of AB is obtained by multiplying the i-th row of A with B.

(b) Then, verify that (AB)[:, j] = AB[:, j]. That is, the j-th column of AB is obtained by multiplying A with the j-th column of B.

Hence,

$$AB = \begin{bmatrix} A[1,:]B \\ A[2,:]B \\ \vdots \\ A[n,:]B \end{bmatrix} = [A\ B[:,1],\ A\ B[:,2],\ \ldots,\ A\ B[:,p]]. \qquad (4)$$

Example: Let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Use the row/column method of matrix multiplication to

1. find the second row of the matrix *AB*.

Solution: By Remark 1.4.2.4, $(AB)[2, :] = A[2, :]B$ and hence

$(AB)[2, :] = 1 \cdot [1, 0, -1] + 0 \cdot [0, 0, 1] + 1 \cdot [0, -1, 1] = [1, -1, 0]$.

2. find the third column of the matrix *AB*.

Solution: Again, by Remark 1.2.9.4, $(AB)[:, 3] = A\, B[:, 3]$ and hence

$$(AB)[:, 3] = -1 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Definition 1.4.3. Two square matrices A and B are said to **commute** if AB = BA.

Remark 1.4.4. Note that if A is a square matrix of order n and if B is a scalar matrix of order n then AB = BA. In general, the matrix product is not commutative. For example, consider $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ Then, verify that $AB = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = BA$.

Theorem 1.4.5. Suppose that the matrices A, B and C are so chosen that the matrix multiplications are defined.

1. Then, (AB)C = A(BC). That is, the matrix multiplication is associative.
2. For any $k \in \mathbb{R}$, $(kA)B = k(AB) = A(kB)$.
3. Then, A(B + C) = AB + AC. That is, multiplication distributes over addition.
4. If $A \in \mathbb{M}_n(\mathbb{C})$ then $AI_n = I_nA = A$.

Proof. Part 1. Let A = [aij] ∈ $\mathbb{M}_{n,r}(\mathbb{C})$), B = [bij] ∈ $\mathbb{M}_{n,p}(\mathbb{C})$) and C = [cij] ∈ $\mathbb{M}_{p,q}(\mathbb{C})$. Then,

$$(BC)_{kj} = \sum_{\ell=1}^{p} b_{k\ell}c_{\ell j} \quad \text{and} \quad (AB)_{i\ell} = \sum_{k=1}^{n} a_{ik}b_{k\ell}.$$

Therefore,

$$\begin{aligned}\left(A(BC)\right)_{ij} &= \sum_{k=1}^{n} a_{ik}\left(BC\right)_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{\ell=1}^{p} b_{k\ell}c_{\ell j}\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{ik}\left(b_{k\ell}c_{\ell j}\right) \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{p}\left(a_{ik}b_{k\ell}\right)c_{\ell j} = \sum_{\ell=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{k\ell}\right)c_{\ell j} = \sum_{\ell=1}^{T}\left(AB\right)_{i\ell}c_{\ell j} = \left((AB)C\right)_{ij}.\end{aligned}$$

Using a similar argument, the next part follows.

1.6 INVERSE OF A MATRIX

Definition 1.5.1. Let $A \in \mathbb{M}_n$ ($\mathbb{C}$)

1. Then, a square matrix B is said to be a **left inverse** of A, if BA = I_n.
2. Then, a square matrix C is called a **right inverse** of A, if AC = I_n.
3. Then, A is said to be **invertible** (or is said to have an **inverse**) if there exists a matrix

B such that AB = BA = I_n.

Lemma 1.5.2. Let A ∈ $\mathbb{M}_n$ ($\mathbb{C}$)If that there exist B, C ∈ $\mathbb{M}_n$ ($\mathbb{C}$)such that AB = I_n and CA = I

n then B = C.

Proof. Note that C = CI_n = C(AB) = (CA)B = I_nB = B.

Remark 1.5.3. Lemma **1.5.2.** implies that whenever A is invertible, the inverse is unique.

Thus, we denote the inverse of A by A^{-1}. That is,

$$AA^{-1} = A^{-1}A = I.$$

Example: Prove that the matrices

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Solution: Suppose there exists C such that CA = AC = I. Then, using matrix product

A[1, :]C = (AC) [1, :] = I[1, :] = [1, 0, 0] and A[2, :]C = (AC)[2, :] = I[2, :] = [0, 1, 0].

But A[1, :] = A[2, :] and thus [1, 0, 0] = [0, 1, 0], a contradiction. Similarly, if there exists D such that BD = DB = I then

DB[:, 1] = (DB)[:, 1] = I[:, 1], DB[:, 2] = (DB)[:, 2] = I[:, 2] and DB[:, 3] = I[:, 3].

But B[:, 3] = B[:, 1] + B[:, 2] and hence I[:, 3] = I[:, 1] + I[:, 2], a contradiction.

Theorem 1.5.4. Let A and B be two invertible matrices. Then,

1. $(A^{-1})^{-1} = A$.
2. $(AB)^{-1} = B^{-1} A^{-1}$.
3. $(A^*)^{-1} = (A^{-1})^*$.

Proof.

Part 1. Let B = A−1 be the inverse of A. Then, AB = BA = I. Thus, by definition,
B is invertible and B−1 = A. Or equivalently, $(A^{-1})^{-1} = A$.

Part 2. By associativity $(AB)(B^{-1} A^{-1}) = A(BB^{-1})A^{-1} = I = (B^{-1} A^{-1})(AB)$.

Part 3 As $AA^{-1} = A^{-1}A = I$, we get $(AA^{-1})^* = (A^{-1}A)^* = I^*$. Or equivalently,

$(A^{-1})^*. A^* = A^*(A^{-1})^* = I$. Thus, by definition $(A^*)^{-1} = (A^{-1})^*$.

1.7 SOME MORE SPECIAL MATRICES

Definition 1.6.1. 1. For $1 \leq k \leq m$ and $1 \leq l \leq n$, define a matrix $k` \in \mathbb{M}_n$ $(\mathbb{C})$ by $(\mathcal{E}_{kl})ij$ = (1 0,, if (otherwise $k,l) = (i, j)$. Then, the matrices

$$(\mathcal{E}_{k\ell})_{ij} = \begin{cases} 1, & \text{if } (k,\ell) = (i,j) \\ 0, & \text{otherwise.} \end{cases}$$

are called the **standard basis** elements for $\mathbb{M}_{m,n}$ $(\mathbb{C})$

So, if $\mathcal{E}_{kl} \in \mathbb{M}_{2,3}$ $(\mathbb{C})$ then

$$\mathcal{E}_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \mathcal{E}_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

and

$$\mathcal{E}_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}.$$

<u>**1.6.2**</u> Let $A \in \mathbb{M}_{m,n}$ $(\mathbb{R})$

(a) Then, A is called **symmetric** if $A^T = A$. For example,

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$$

(b) Then, A is called **skew-symmetric** if $A^T = -A$. For example,

$$A = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}.$$

(c) Then, A is called **orthogonal** if $AA^T = A^T A = I$. For example,

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

(d) Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is said to be a **permutation matrix** if A has exactly one non-zero entry, namely 1, in each row and column. For example, I_n, for each positive integer n

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are permutation matrices.

Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.

(a) Then, A is called **normal** if $A^*A = AA^*$. For example, $A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$ is a normal matrix

(b) Then, A is called **Hermitian** if $A* = A$. For example,

$$A = \begin{bmatrix} 1 & 1+\mathbf{i} \\ 1-\mathbf{i} & 2 \end{bmatrix}.$$

(b) Then, A is called **Hermitian** if $A* = A$. For example,

$$A = \begin{bmatrix} 0 & 1+\mathbf{i} \\ -1+\mathbf{i} & 0 \end{bmatrix}.$$

(c) Then, A is called **skew-Hermitian** if $A* = -A$. For example,

$$A = \begin{bmatrix} 0 & 1+\mathbf{i} \\ -1+\mathbf{i} & 0 \end{bmatrix}.$$

(d) Then, A is called **unitary** if $AA* = A*A = I$. For example,

$$A = \frac{1}{\sqrt{3}} \begin{bmatrix} 1+\mathbf{i} & 1 \\ -1 & 1-\mathbf{i} \end{bmatrix}.$$

4. Then, A is called **idempotent** if $A^2 = A$. For example, is idempotent.

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

5. A vector $\mathbf{u} \in \mathbb{M}_{n,1}$ ($\mathbb{C}$).such that $\mathbf{u}*\mathbf{u} = 1$ is called a **unit vector**.
6. A matrix that is symmetric and idempotent is called a **projection** matrix. For example, let $\mathbf{u} \in \mathbb{M}_{n,1}$ ($\mathbb{R}$).be a unit vector. Then, $A = \mathbf{uu}^T$ is a symmetric and an idempotent matrix. Hence, A is a projection matrix. In particular, let $\boldsymbol{u} = \frac{1}{\sqrt{5}}[1,2]^T$ and $A = \mathbf{uu}^T$ .Then, $\mathbf{u}^T\mathbf{u} = 1$ and for any vector $\mathbf{x} = [x_1, x_2]^T \in \mathbb{M}_{2,1}$ ($\mathbb{R}$) note that

$$A\mathbf{x} = (\mathbf{uu}^T)\mathbf{x} = \mathbf{u}(\mathbf{u}^T\mathbf{x}) = \frac{x_1 + 2x_2}{\sqrt{5}}\mathbf{u} = \left[\frac{x_1 + 2x_2}{5}, \frac{2x_1 + 4x_2}{5}\right]^T.$$

Thus, $A\mathbf{x}$ is the foot of the perpendicular from the point $\mathbf{x}$ on the vector $[1\ 2]^T$.

7. Fix a unit vector $\mathbf{a} \in \mathbb{M}_{n,1}$ ($\mathbb{R}$) and let $A = 2\mathbf{aa}^T - I_n$. Then, verify that $A \in \mathbb{M}_n$ ($\mathbb{R}$) and $A\mathbf{y} = 2(\mathbf{a}^T\mathbf{y})\mathbf{a} - \mathbf{y}$, for all $\mathbf{y} \in \mathbb{R}^n$. This matrix is called the **reflection** matrix about the line containing the points $\mathbf{0}$ and $\mathbf{a}$.

8. Let $A \in \mathbb{M}_n$ ($\mathbb{C}$) Then, A is said to be **nilpotent** if there exists a positive integer n such that $An = \mathbf{0}$. The least positive integer k for which

Ak = **0** is called the **order of nilpotency**. For example, if $A = [a_{ij}] \in \mathbb{M}_n$ ($\mathbb{C}$) with a_{ij} equal to 1 if $i - j = 1$ and 0, otherwise then A_n = **0** and $A^{`} \neq$ **0** for $1 \leq l \leq n - 1$.

Check your progress

3. Explain inverse of a matrix [HINT: Provide definition and example]

4. Define skew symmetric and permutation matrix[HINT: Provide the definition and example]

1.8 SUMMARY

We learnt different types of matrices and how different operations on Matrices are performed. Specially product and inverse of matrices. We also learnt product of two matrices. Even though it seemed complicated, it basically tells that multiplying by a matrix on the

1. left to a matrix A is same as operating on the rows of A.
2. right to a matrix A is same as operating on the columns of A.

1. 9 KEYWORDS

1. Vector- a quantity having direction as well as magnitude, especially as determining the position of one point in space relative to another.
2. Perpendicular- **Perpendicular** means "at right angles". A line meeting another at a right angle, or 90° is said to be **perpendicular** to it
3. Non-zero entry - A quantity which does **not** equal **zero** is said to be **nonzero**. A real **nonzero** number must be either positive or negative, and a complex **nonzero** number can have either real or imaginary part **nonzero**
4. **Invertible:** A function f that has an inverse is called invertible;

1.10 QUESTION FOR REVIEW

1. Suppose A = $[a_{ij}]$ and B are matrices such that A+B = **0**. Then, show that B = (−1)A = $[−a_{ij}]$.

2. Let A be a square matrix satisfying $A^3 + A - 2I = \mathbf{0}$. Prove that A – 1 = ½ $(A^2 + I)$

3. Let A be an upper triangular matrix. If $A * A = AA^*$ then prove that A is a diagonal matrix. The same holds for lower triangular matrix.

4. Let A, B ∈ Mm,n(C). If A**x** = B**x**, for all **x** ∈ Mn,1(C) then prove that A = B.

5. Find 2 × 2 nonzero matrices A, B and C satisfying AB = AC but B ≠ C. That is, thecancelation law doesn’t hold.

1.11 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.

2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.

3.S. Lang, Linear Algebra, Springer, 1989.

4.David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.

5.R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.

6.Thomas Hungerford, Algebra, Springer GTM.

7.I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.

8.D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.

1.12 ANSWER TO CHECK YOUR PROGRESS

1.Provide the definition with example --1.2.1 – 2 & 5

2.Provide the definition--1.3.4 and state the properties and their proof from theorem- 1.3.5

3.Provide the explanation, lemma and remark with example – 1.5.1, 1.5.2 & 1.5.3

4.Provide definition and example 1.6.2 – (b) and (d)

UNIT 2: SYSTEM OF LINEAR EQUATIONS

STRUCTURE

2.0 Objective

2.1 Introduction

2.2 Linear Equations

2.3 Elementary Row Operations

2.4 Elementary Matrices And The Row-Reduced Echelon Form

2.5 Row-Reduced Echelon Form (Rref)

2.6 Summary

2.7 Keywords

2.8 Questions for review

2.9 Suggested Readings

2.10 Answers To Check Your Progress

2.0 OBJECTIVE

Learn the concept of linear equations and their representation using matrix form

Understand the elementary row operation

Comprehend the concept of Elementary Matrices and the Row-Reduced Echelon Form & Row-Reduced Echelon Form (RREF)

2.1 INTRODUCTION

Systems of **linear equations** arose in Europe with the introduction in 1637 by René Descartes of coordinates in geometry. In fact, in this new geometry, now called Cartesian geometry, lines and planes are represented by **linear equations**, and computing their intersections amounts to solving systems of **linear equations.**

2.2 LINEAR EQUATIONS

Example Let us look at some examples of linear systems.

1. Suppose $a, b \in$ R. Consider the system $ax = b$ in the variable x. If

(a) $a \neq 0$ then the system has a unique solution $x = a\ b$.

(b) $a = 0$ and

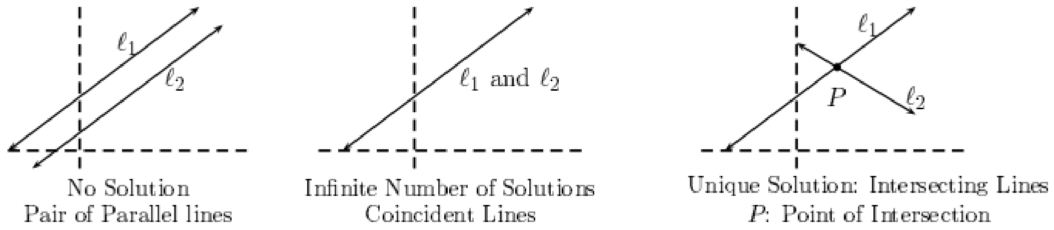
i. $b \neq 0$ then the system has no solution.

ii. $b = 0$ then the system has infinite number of solutions, namely all $x \in$ R.

2. Consider a linear system with 2 equations in 2 variables. The equation $ax + by = c$ in the

variables x and y represents a line in R2 if either $a \neq 0$ or $b \neq 0$. Thus the solution set of the system

$$a_1x + b_1y = c_1,\ a_2x + b_2y = c_2$$



is given by the points of intersection of the two lines (see Figure 2.1 for illustration of different cases)

Figure 2.1: Examples in 2 dimension.

(a) Unique Solution:

$x - y = 3$ and $2x + 3y = 11$. The unique solution is $[x, y]^T = [4, 1]^T$.

Observe that in this case, $a_1b_2 - a_2b_1 \neq 0$.

(b) Infinite Number of Solutions:

$x + 2y = 1$ and $2x + 4y = 2$. As both equations represent the same line, the solution set is $[x, y]^T = [1 - 2y, y]^T = [1, 0]^T + y[-2, 1]^T$ with y arbitrary. Observe that

i. $a_1b_2 - a_2b_1 = 0$, $a_1c_2 - a_2c_1 = 0$ and $b_1c_2 - b_2c_1 = 0$.
ii. the vector $[1, 0]^T$ corresponds to the solution $x = 1, y = 0$ of the given system.
iii. the vector $[-2, 1]^T$ corresponds to the solution $x = -2, y = 1$ of the system
$x + 2y = 0, 2x + 4y = 0$.

(c) No Solution
$x + 2y = 1$ and $2x + 4y = 3$. The equations represent a pair of parallel lines and hence there is no point of intersection. Observe that in this case, $a_1b_2 - a_2b_1 = 0$ but $a_1c_2 - a_2c_1 \neq 0$.

3. As a last example, consider 3 equations in 3 variables.
A linear equation $ax + by + cz = d$ represents a plane in R3 provided $[a, b, c] \neq [0, 0, 0]$. Here, we have to look at the points of intersection of the three given planes. It turns out that there are seven different ways in which the three planes can intersect. We present only three ways which correspond to different cases.

(a) Unique Solution
Consider the system $x + y + z = 3$, $x+4y+2z = 7$ and $4x+10y-z = 13$. The unique solution to this system is $[x, y, z]T = [1, 1, 1]T$, *i.e.*, the three planes intersect at a point.

(b) Infinite Number of Solutions
Consider the system $x + y + z = 3$, $x + 2y + 2z = 5$ and $3x + 4y + 4z = 11$. The solution set is $[x, y, z]T = [1, 2 - z, z]T = [1, 2, 0]T + z[0, -1, 1]T$, with z arbitrary.

Observe the following:
i. Here, the three planes intersect in a line.

ii. The vector $[1, 2, 0]T$ corresponds to the solution $x = 1$, $y = 2$ and $z = 0$ of the linear system $x + y + z = 3$, $x + 2y + 2z = 5$ and $3x + 4y + 4z = 11$. Also, the vector $[0, -1, 1]T$ corresponds to the solution $x = 0$, $y = -1$ and $z = 1$ of the linear system $x + y + z = 0$, $x + 2y + 2z = 0$ and $3x + 4y + 4z = 0$.

(c) No Solution

The system $x + y + z = 3$, $2x + 2y + 2z = 5$ and $3x + 3y + 3z = 3$ has no solution. In this case, we have three parallel planes. The readers are advised to supply the proof.

Before we start with the general set up for the linear system of equations, we give different

interpretations of the examples considered above.

Example 1. Recall above Example a, where we have verified that the solution of the linear system $x - y = 3$ and $2x + 3y = 11$ equals $[4, 1]T$. Now, observe the following:

(a) The solution $[4, 1]T$ corresponds to the point of intersection of the corresponding two

lines.

(b) Using matrix multiplication the linear system equals $A\mathbf{x} = \mathbf{b}$, where $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$, $x = \begin{bmatrix} x \\ y \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 11 \end{bmatrix}$. So, the solution is

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{5}\begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 11 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

(c) Re-writing

$$A\mathbf{x} = \mathbf{b} \text{ as } \begin{bmatrix} 1 \\ 2 \end{bmatrix} x + \begin{bmatrix} -1 \\ 3 \end{bmatrix} y = \begin{bmatrix} 3 \\ 11 \end{bmatrix}$$

gives us $4 \cdot (1, 2) + 1 \cdot (-1, 3) = (3, 11)$.

This corresponds to addition of vectors in the Euclidean plane.

2. Recall Example 3.3a, where the point of intersection of the three

planes is the point (1, 1, 1) in the Euclidean space. Note that in matrix notation, the system reduces to $A\mathbf{x} = \mathbf{b}$, where]

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 2 \\ 4 & 10 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 3 \\ 7 \\ 13 \end{bmatrix}.$$

Then

$$\text{(a) } \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}\mathbf{b} = \frac{1}{21}\begin{bmatrix} 24 & -11 & 2 \\ -9 & 5 & 1 \\ 6 & 6 & -3 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \\ 13 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(b) $1 \cdot (1, 1, 4) + 1 \cdot (1, 4, 10) + 1 \cdot (1, 2, -1) = (3, 5, 13)$. This corresponds to addition of vectors in the Euclidean space.

Thus, there are three ways of looking at the linear system $A\mathbf{x} = \mathbf{b}$, where, as the name suggests,
one of the ways is looking at the point of intersection of planes, the other is the vector sum approach and the third is the matrix multiplication approach. All of three approaches are important as they give different insight to the study of matrices.

Definition 2.1.1 [Linear System] A system of m linear equations in n variables $x_1, x_2, \ldots, x_n$
is a set of equations of the form

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

where for $1 \leq i \leq m$ and $1 \leq j \leq n$; aij, $bi \in$ R. Linear System (1) is called **homogeneous**
if $b_1 = 0 = b_2 = \cdots = b_m$ and **non-homogeneous**, otherwise.

Definition 2.1.2. [Coefficient and Augmented Matrices] Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

(1)

Then, (1) can be re-written as $A\mathbf{x} = \mathbf{b}$. In this setup, the matrix A is called the **coefficient** matrix and the block matrix $[A\ \mathbf{b}]$ is called the **augmented** matrix of the linear system (1).

Remark 2.1.3. *Consider the linear system* $A\mathbf{x} = \mathbf{b}$, *where* $A \in \mathbb{M}_{m,n}(\mathbb{C})$, $\mathbf{b} \in \mathbb{M}_{m,1}(\mathbb{C})$ *and* $\mathbf{x} \in \mathbb{M}_{m,1}(\mathbb{C})$ *.If* $[A\ \mathbf{b}]$ *is the augmented matrix and* $\mathbf{x}^T = [x1, \ldots, xn]$ *then,1. for* $j = 1, 2, \ldots, n$, *the variable* xj *corresponds to the column* $([A\ \mathbf{b}])[:, j]$.

2. the vector $\mathbf{b} = ([A\ \mathbf{b}])[:, n + 1]$.

3. for $i = 1, 2, \ldots, m$, *the ith equation corresponds to the row* $([A\ \mathbf{b}])[i,:]$.

Definition 2.1.4. [Solution of a Linear System] A **solution** of $A\mathbf{x} = \mathbf{b}$ is a vector $\mathbf{y}$ such that $A\mathbf{y}$ indeed equals $\mathbf{b}$. The set of all solutions is called the **solution set** of the system. For example, the solution set of $A\mathbf{x} = \mathbf{b}$, with

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 2 \\ 4 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ equals } \left\{ \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \right\}.$$

Definition 2.1.5. [Consistent, Inconsistent] Consider a linear system $A\mathbf{x} = \mathbf{b}$. Then, this linear system is called **consistent** if it admits a solution and is called **inconsistent** if it admits no solution. For example, the homogeneous system $A\mathbf{x} = \mathbf{0}$ is always consistent as $\mathbf{0}$ is a solution whereas, verify that the system $x + y = 2$, $2x + 2y = 3$ is inconsistent.

Definition 2.1.6. [Associated Homogeneous System] Consider a linear system $A\mathbf{x} = \mathbf{b}$. Then, the corresponding linear system $A\mathbf{x} = \mathbf{0}$ is called the **associated homogeneous system**. **0** is always a solution of the associated homogeneous system.

Theorem 2.1.7. Consider a homogeneous linear system $A\mathbf{x} = \mathbf{0}$.
1. Then, $\mathbf{x} = \mathbf{0}$, the zero vector, is always a solution, called the **trivial** solution.
2. Let $\mathbf{u} \neq \mathbf{0}$ be a solution of $A\mathbf{x} = \mathbf{0}$. Then, $\mathbf{y} = c\mathbf{u}$ is also a solution, for all $c \in C$.
A nonzero solution is called a **non-trivial** solution. Note that, in this case, the system $A\mathbf{x} = \mathbf{0}$ has an infinite number of solutions.

3. Let $\mathbf{u_1} \ldots \mathbf{u}_k$ be solutions of $A\mathbf{x} = \mathbf{0}$. Then, $\sum_{i=1}^{k} a_i u_i$ is also a solution of $A\mathbf{x} = \mathbf{0}$, for
each choice of $a_i \in C$, $1 \leq i \leq k$.

Remark 2.1.8. 1. Let A = [1 1 1 1]. Then, **x** = [− 11] is a non-trivial solution of $A\mathbf{x} = \mathbf{0}$.
2. Let $\mathbf{u} \neq \mathbf{v}$ be solutions of a non-homogeneous system $A\mathbf{x} = \mathbf{b}$. Then, **x**h = $\mathbf{u} - \mathbf{v}$ is a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$. That is, any two distinct solutions of $A\mathbf{x} = \mathbf{b}$ differ by a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$. Or equivalently, the solution set of $A\mathbf{x} = \mathbf{b}$ is of the form, $\{\mathbf{x}_0 + \mathbf{x}_h\}$, where $\mathbf{x}_0$ is a particular solution of $A\mathbf{x} = \mathbf{b}$ and **x**h is a solution of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

2.3 ELEMENTARY ROW OPERATIONS

Example 2.2.1. Solve the linear system y + z = 2, 2x + 3z = 5, x + y + z = 3.
Solution: Let B0 = [A **b**], the augmented matrix. Then systematically proceed to get the solution.

$$B_0 = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

1. Interchange 1-st and 2-nd equations (interchange $B_0[1, :]$ and $B_0[2, :]$ to get B_1).

$$\begin{aligned} 2x + 3z &= 5 \\ y + z &= 2 \\ x + y + z &= 3 \end{aligned} \qquad B_1 = \begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

2. In the new system, multiply 1-st equation by ½ (multiply $B_1[1, :]$ by 1/2 to get B_2).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ x + y + z &= 3 \end{aligned} \qquad B_2 = \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

3. In the new system, replace 3-rd equation by 3-rd equation minus 1-st equation (replace $B_2[3, :]$ by $B_2[3, :] - B_2[1, :]$ to get B_3)

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ y - \tfrac{1}{2}z &= \tfrac{1}{2} \end{aligned} \qquad B_3 = \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

4. In the new system, replace 3-rd equation by 3-rd equation minus 2-nd equation (replace $B_3[3, :]$ by $B_3[3, :] - B_3[2, :]$ to get B_4).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ -\tfrac{3}{2}z &= -\tfrac{3}{2} \end{aligned} \qquad B_4 = \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix}.$$

5. In the new system, multiply 3-rd equation by −2/3 (multiply $B_4[3, :]$ by −2/3 to get B_5).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ z &= 1 \end{aligned} \qquad B_5 = \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

.

The last equation gives z = 1. Using this, the second equation gives y = 1. Finally, the first equation gives x = 1. Hence, the solution set is {[x, y, z]T | [x, y, z] = [1, 1, 1]}, a unique solution. In Example 2.2.1, observe how each operation on the linear system corresponds to a similar operation on the rows of the augmented matrix. We use this idea to define elementary row operations and the equivalence of two linear systems.

Definition 2.2.2. [Elementary Row Operations] Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **elementary row operations** are

1. E_{ij}: Interchange the i-th and j-th rows, namely, interchange A[i, :] and A[j, :].
2. $E_k(c)$ for $c \neq 0$: Multiply the k-th row by c, namely, multiply A[k, :] by c.
3. $E_{ij}(c)$ for $c \neq 0$: Replace the i-th row by i-th row plus c-times the j-th row, namely, replace A[i, :] by A[i, :] + cA[j, :].

Definition 2.2.3. [Row Equivalent Matrices] Two matrices are said to be **row equivalent** if one can be obtained from the other by a finite number of elementary row operations.

Definition 2.2.4. [Row Equivalent Linear Systems] The linear systems A**x** = **b** and C**x** = **d** are said to be **row equivalent** if their respective augmented matrices, [A **b**] and [C **d**], are row equivalent. Thus, note that the linear systems at each step in Example 2.2.1 are row equivalent to each other. We now prove that the solution set of two row equivalent linear systems are same.

Lemma 2.2.5. Let C**x** = **d** be the linear system obtained from A**x** = **b** by application of a single elementary row operation. Then, $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same solution set.

Proof. We prove the result for the elementary row operation $E_{jk}(c)$ with $c \neq 0$.

In this case, the systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ vary only in the j^{th} equation. So, we need to show that $\mathbf{y}$ satisfies the jth equation of $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{y}$ satisfies the j^{th} equation of $C\mathbf{x} = \mathbf{d}$. So, let $\mathbf{y}^T = [\alpha_1, . . . , \alpha_n]$. Then, the j^{th} and k^{th} equations of $A\mathbf{x} = \mathbf{b}$ are

$$a_{j1}\alpha_1 + \cdot \cdot \cdot + a_{jn}\alpha_n = b_j \text{ and } a_{k1}\alpha_1 + \cdot \cdot \cdot + a_{kn}\alpha_n = b_k.$$

Therefore, we see that αi's satisfy

$$(a_{j1} + ca_{k1})\alpha 1 + \cdot \cdot \cdot + (a_{jn} + ca_{kn})\alpha_n = b_j + c_{bk}. \qquad (1)$$

Also, by definition the jth equation of $C\mathbf{x} = \mathbf{d}$ equals

$$(a_{j1} + ca_{k1})x_1 + \cdot \cdot \cdot + (a_{jn} + ca_{kn})x_n = b_j + c_{bk} \qquad (2)$$

Therefore, using Equation (1), we see that $\mathbf{y}^T = [\alpha_1, . . . , \alpha_n]$ is also a solution for Equation

(2). Now, use a similar argument to show that if $\mathbf{z}^T = [\beta_1, . . . , \beta_n]$ is a solution of $C\mathbf{x} = \mathbf{d}$

then it is also a solution of $A\mathbf{x} = \mathbf{b}$. Hence, the required result follows.

Theorem 2.2.6: Let $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ be two row equivalent linear systems. Then, they have the same solution set.

Check you progress

1. Explain Elementary row operations

---

2. Define Coefficient and Augmented matrices

---

2. 4 ELEMENTARY MATRICES AND THE ROW-REDUCED ECHELON FORM (RREF)

Definition 2.3.1. [Elementary Matrix] A matrix $E \in \mathbb{M}_n(\mathbb{C})$ is called an **elementary matrix** if it is obtained by applying exactly one elementary row operation to the identity matrix I_n.

Remark 2.3.2. The elementary matrices are of three types and they correspond to elementary row operations.

1. E_{ij}: Matrix obtained by applying elementary row operation Eij to I_n.
2. $E_k(c)$ for $c \neq 0$: Matrix obtained by applying elementary row operation $E_k(c)$ to I_n.
3. $E_{ij}(c)$ for $c \neq 0$: Matrix obtained by applying elementary row operation $E_{ij}(c)$ to I_n.

When an elementary matrix is multiplied on the left of a matrix A, it gives the same result as that of applying the corresponding elementary row operation on A.

When an elementary matrix is multiplied on the left of a matrix A, it gives the same result as that of applying the corresponding elementary row operation on A.

Remark 2.3.3. *Observe that*

1. $(E_{ij})^{-1} = E_{ij}$ as $E_{ij}E_{ij} = I = E_{ij}E_{ij}$.
2. Let $c \neq 0$. Then, $(E_k(c))^{-1} = E_k(1/c)$ as $E_k(c)Ek(1/c) = I = E_k(1/c)E_k(c)$.
3. Let $c \neq 0$. Then, $(E_{ij}(c))^{-1} = E_{ij}(-c)$ as $E_{ij}(c)E_{ij}(-c) = I = E_{ij}(-c)E_{ij}(c)$.

Thus, each elementary matrix is invertible. Also, the inverse is an elementary matrix of the same type.

Proposition 2.3.4. *Let A and B be two row equivalent matrices. Then, prove that* $B = E_1 \cdots E_{kA}$, *for some elementary matrices* $E_1, \ldots, E_k$.

Proof. By definition of row equivalence, the matrix *B* can be obtained from *A* by a finite number of elementary row operations. But by Remark 2.3.2, each elementary row operation on *A* corresponds to left multiplication by an elementary matrix to *A*. Thus, the required result follows.

Theorem 2.3.5. Let A**x** = **b** and C**x** = **d** be two row equivalent linear systems. Then, they have the same solution set.

Proof. Let E1, . . . , Ek be the elementary matrices such that $E_1 \cdots E_k$[A **b**] = [C **d**]. Put
$E = E_1 \cdots E_k$. Then, by Remark 2.3.3

$$EA = C,\ E\mathbf{b} = \mathbf{d},\ A = E{-}1C \text{ and } \mathbf{b} = E{-}1\mathbf{d} \qquad .\ (1)$$

Now assume that A**y** = **b** holds. Then, by Equation (1)

$$C\mathbf{y} = EA\mathbf{y} = E\mathbf{b} = \mathbf{d}. \qquad (2)$$

On the other hand if C**z** = **d** holds then using Equation (1), we have

$$A\mathbf{z} = E{-}1C\mathbf{z} = E{-}1\mathbf{d} = \mathbf{b}.\ (2.2.3)$$

Therefore, using Equations (2) and (3) the required result follows.
The following result is a particular case of Theorem 2.3.5

Corollary 2.3.6. Let A and B be two row equivalent matrices. Then, the systems A**x** = **0** and
B**x** = **0** have the same solution set

Example. Are the matrices row equivalent?

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{bmatrix}$$

Notes

Solution: No, as $\begin{bmatrix} a \\ b \\ -1 \end{bmatrix}$ is a solution of $B\mathbf{x} = \mathbf{0}$ but it isn't a solution of $A\mathbf{x} = \mathbf{0}$.

Definition 2.3.7. [Pivot/Leading Entry] Let A be a nonzero matrix. Then, in each nonzero

row of A, the left most nonzero entry is called a **pivot/leading entry**. The $\boxed{a_{ij}}$ column containing

the pivot is called a **pivotal column**. If *aij* is a pivot then we denote it by For example,

the entries a_{12} and a_{23} are pivots in

$$A = \begin{bmatrix} 0 & \boxed{3} & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{2} & 1 \end{bmatrix}.$$

Thus, columns 2 and 3 are pivotal columns.

Definition 2.3.8. [Row Echelon Form] A matrix is in **row echelon form (REF)** (ladder like)

1. if the zero rows are at the bottom;
2. if the pivot of the $(i + 1)$-th row, if it exists, comes to the right of the pivot of the i-th row.
3. if the entries below the pivot in a pivotal column are 0.

Example. The following matrices are in echelon form

$$\begin{bmatrix} 0 & \boxed{2} & 4 & 2 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \boxed{1} & 1 & 0 & 2 & 3 \\ 0 & 0 & 0 & \boxed{3} & 4 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}, \begin{bmatrix} \boxed{1} & 2 & 0 & 5 \\ 0 & \boxed{2} & 0 & 6 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$$

2.5 ROW-REDUCED ECHELON FORM (RREF)]

Definition 2.4.1. [Row-Reduced Echelon Form (RREF)] A matrix C is said to be in **row-reduced echelon form (RREF)**

1. if C is already in echelon form,
2. if the pivot of each nonzero row is 1,
3. if every other entry in each pivotal column is zero.

A matrix in RREF is also called a row-reduced echelon matrix.

Example The following matrices are in RREF.

$$\begin{bmatrix} 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \boxed{1} & 0 & 0 & 5 \\ 0 & \boxed{1} & 0 & 6 \\ 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \boxed{1} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}.$$

Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$.

2.4.2 Gauss-Jordan Elimination

We now present an algorithm, commonly known as the Gauss-Jordan Elimination (GJE), to compute the RREF of A.

1. Input: A.
2. Output: a matrix B in RREF such that A is row equivalent to B.
3. **Step** 1**:** Put 'Region' = A.
4. **Step** 2**:** If all entries in the Region are 0, STOP. Else, in the Region, find the left most
nonzero column and find its topmost nonzero entry. Suppose this nonzero entry is $a_{ij} = c$ (say). Box it. This is a pivot.
5. **Step** 3**:** Interchange the row containing the pivot with the top row of the region. Also,
make the pivot entry 1 by dividing this top row by c. Use this pivot to make other entries
in the pivotal column as 0.
6. **Step** 4**:** Put Region = the submatrix below and to the right of the

current pivot. Now,

go to step 2.

Important: The process will stop, as we can get at most min$\{m, n\}$ pivots.

Example. Apply GJE to

$$\begin{bmatrix} 0 & 2 & 3 & 7 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

1. Region = A as $A \neq \mathbf{0}$

2. Then

$$E_{12}A = \begin{bmatrix} \boxed{1} & 1 & 1 & 1 \\ 0 & 2 & 3 & 7 \\ 1 & 3 & 4 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Also,

$$E_{31}(-1)E_{12}A = \begin{bmatrix} \boxed{1} & 1 & 1 & 1 \\ 0 & 2 & 3 & 7 \\ 0 & 2 & 3 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix} = B \text{ (say)}.$$

3. Now $\quad \text{Region} = \begin{bmatrix} 2 & 3 & 7 \\ 2 & 3 & 7 \\ 0 & 0 & 1 \end{bmatrix} \neq \mathbf{0}.$

$$E_2(\tfrac{1}{2})B = \begin{bmatrix} \boxed{1} & 1 & 1 & 1 \\ 0 & \boxed{1} & \frac{3}{2} & \frac{7}{2} \\ 0 & 2 & 3 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C\text{(say)}.$$

Then,

$$E_{12}(-1)E_{32}(-2)C = \begin{bmatrix} \boxed{1} & 0 & \frac{-1}{2} & \frac{-5}{2} \\ 0 & \boxed{1} & \frac{3}{2} & \frac{7}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = D(\text{say}).$$

4. Now, Then

$$E_{34}D = \begin{bmatrix} \boxed{1} & 0 & \frac{-1}{2} & \frac{-5}{2} \\ 0 & \boxed{1} & \frac{3}{2} & \frac{7}{2} \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now, multiply on the left by $E_{13}(5/2)$ and $E_{23}(-2/7)$ to get

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

$$F = \text{RREF}(A) = \begin{bmatrix} \boxed{1} & 0 & -\frac{1}{2} & 0 \\ 0 & \boxed{1} & \frac{3}{2} & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Proposition 2.4.3. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then, A is invertible if and only if RREF(A) = I_n. That is, every invertible matrix is a product of elementary matrices.*

Proof. If RREF(A) = In then $In = E1 \cdots EkA$, for some elementary matrices $E_1, \ldots, E_k$. As Ei's are invertible, $E_1^{-1} = E_2 \cdots E_kA$, $E_2^{-1}\ E_1^{-1} = E_3 \cdots E_kA$ and so on. Finally, one obtains $A = E_k^{-1} \cdots E_1^{-1}$. A similar calculation now gives $AE_1 \cdots E_k = I_n$. Hence, by definition of invertibility $A^{-1} = E_1 \cdots E_k$.

Now, let A be invertible with B = RREF(A) = $E_1 \cdots E_k A$, for some elementary matrices $E_1, \ldots, E_k$. As A and Ei's are invertible, the matrix B is invertible. Hence, B doesn't have any zero row. Thus, all the n rows of B have pivots. Therefore, B has n pivotal columns. As B has exactly n columns, each column is a pivotal column and hence $B = In$. Thus, the required result follows.

As a direct application of Proposition 2.4.3 one obtains the following.

Theorem 2.4.5. *Let $A \in M_{m,n}(\mathbb{C})$. Then, for any invertible matrix S, RREF(SA) = RREF(A).*

Proposition 2.4.6. *Let $A \in \mathrm{M}_n(\mathbb{C})$ be an invertible matrix. Then, for any matrix B, define* $\mathbf{C} = [\boldsymbol{A} \quad \boldsymbol{B}]$ *and* $D = \begin{bmatrix} A \\ B \end{bmatrix}$. *Then, RREF(C) =* $[I_n \quad A^{-1}B]$ *and RREF(D) =* $\begin{bmatrix} I_n \\ 0 \end{bmatrix}$.

Proof. Using matrix product, $A^{-1}C = [\boldsymbol{A^{-1}A} \quad \boldsymbol{A^{-1}B}] = [\boldsymbol{I_n} \quad \boldsymbol{A^{-1}B}]$.

As $[\boldsymbol{I_n} \quad \boldsymbol{A^{-1}B}]$ is in RREF, RREF(C) = $[\boldsymbol{I_n} \quad \boldsymbol{A^{-1}B}]$. For the second part, note that the matrix $X = \begin{bmatrix} A^{-1} & 0 \\ -BA^{-1} & I_n \end{bmatrix}$ is an invertible matrix. Thus, by Proposition 2.4.3, X is a product of elementary matrices. Now, verify that $XD = \begin{bmatrix} I_n \\ 0 \end{bmatrix}$. As $\begin{bmatrix} I_n \\ 0 \end{bmatrix}$ is in RREF. Let $A \in \mathbb{M}_n(\mathbb{C})$. Suppose we start with $C = [A\ I_n]$ and compute RREF(C). If RREF(C) = $[G\ H]$ then, either $G = In$ or $G \neq I_n$. Thus, if $G = I_n$ then we must have $F = A^{-1}$. If $G \neq I_n$ then, A is not invertible.

Example Use GJE to find the inverse of

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Solution : Applying GJE to

$$[A \mid I_3] = \left[\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \text{ gives}$$

$$[A \mid I_3] \xrightarrow{E_{13}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right] \xrightarrow{E_{13}(-1),E_{23}(-2)} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{E_{12}(-1)} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right].$$

Thus,

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

Check your progress

3. Define **Row Echelon Form and Row Reduced Echelon Form**

4. Explain Gauss – Jordan Elimination algorithm

2.6 SUMMARY

we started with a system of m linear equations in n variables and formally wrote it as $A\mathbf{x} = \mathbf{b}$ and in turn to the augmented matrix $[A \mid \mathbf{b}]$. Then, the basic operations on equations led to multiplication by elementary matrices on the right of $[A \mid \mathbf{b}]$. These elementary matrices are invertible and applying the GJE on a matrix A, resulted in getting the RREF of A.

2.7 KEYWORDS

1. Distinct Solution - A **distinct** real **solution** is a **solution** to an equation that occurs once, and differs in value from other **solutions**.
2. Non-Trivial Solution - The system of equation in which the determinant of the coefficient is zero is called non-trivial solution.
3. Solution Set - In **mathematics**, a **solution set** is the **set** of values that satisfy a given **set** of equations or inequalities.
4. Intersection - **Intersection** of two given sets is the largest set which contains all the elements that are common to both the sets or The point where two lines meet or cross
5. Verify - to establish the truth, accuracy, or reality of.

2.8 QUESTIONS FOR REVIEW

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

1. Which of the following matrices are elementary?

2. Find the inverse of the following matrices using GJE.

$$(i) \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 2 & 4 & 7 \end{bmatrix}$$

2.9 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.
2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.
3. S. Lang, Linear Algebra, Springer, 1989.
4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.
5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.
6. Thomas Hungerford, Algebra, Springer GTM.
7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.
8. D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.

2.10 ANSWER TO CHECK YOUR PROGRESS

1. [HINT: Provide the steps -2.2.2]
2. [HINT: Provide definition with example –2.1.2]
3. [HINT: Provide definition with example—2.3.8 & 2.4.1]
4. [HINT: Provide the steps of algorithm—2.4.2]

UNIT-3 SOLUTION SET OF LINEAR EQUATION

STRUCTURE

3.0 Objective

3.1 Introduction

3.2 Rank of a Matrix

3.3 Solution Set of Linear System

3.4 Square Matrices and Linear System

3.5 Summary

3.6 Keywords

3.7 Questions for review

3.8 Suggested Readings

3.9 Answers to Check your Progress

3.0 OBJECTIVE

Understand the concept of Rank of Matrix

Comprehend the Solution Set of Linear System

Understand the relationship between Square Matrices and Linear System

3.1 INTRODUCTION

A concept closely connected with the concept of a basis. Usually rank is defined either as the minimal cardinality of a generating set (in this way, for example, one introduces the basis rank of an algebraic system), or as the maximal cardinality of a subsystem of elements which are independent in a certain sense.

The rank of a system of a vectors in a vector space over a skew-field is the maximal number of linearly independent vectors in this system (see Linear independence). The rank, or dimension, of a vector space, in particular, is equal to the number of elements in a basis of this space (the rank does not depend on the choice of the basics: all bases have the same cardinality).

The rank of a matrix is defined as the rank of the system of vectors forming its rows (row rank) or of the system of columns (column rank). For matrices over a commutative ring with a unit these two concepts of rank coincide. For a matrix over a field the rank is also equal to the maximal order of a non-zero minor. The rank of a product of matrices is not greater than the rank of each of the factors. The rank of a matrix does not change under multiplication by a non-singular matrix.

3.2 RANK OF A MATRIX

Definition 3.1.1 [Rank of a Matrix] Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$. Then, the **rank** of *A*, denoted Rank(*A*), is the number of pivots in the RREF(*A*).

For example, Rank (I_n) = *n* and Rank (**0**) = 0.

Remark 3.1.2 before proceeding further, for$A \in \mathbb{M}_{m,n}(\mathbb{C})$, we observe the following.

1. The number of pivots in the RREF (A) is same as the number of pivots in REF of A. Hence, we need not compute the RREF (A) to determine the rank of A.

2. Since, the number of pivots cannot be more than the number of rows or the number of columns, one has Rank(A) ≤ min{m, n}.

3. If B = $\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$ Then Rank(B) = Rank(A) as RREF(B) = $\begin{bmatrix} \mathrm{RREF}(A) & 0 \\ 0 & 0 \end{bmatrix}$

4. If A = $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ then, by definition

$$Rank(A) \leq Rank([\boldsymbol{A_{11}} \quad \boldsymbol{A_{12}}]) + Rank([\boldsymbol{A_{21}} \quad \boldsymbol{A_{22}}])$$

(a) $Rank(A) \geq Rank([\boldsymbol{A_{11}} \quad \boldsymbol{A_{12}}])$

(b) $Rank(A) \geq Rank([\boldsymbol{A_{21}} \quad \boldsymbol{A_{22}}])$

(c) $Rank(A) \geq Rank\left(\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}\right)$

Lemma 3.1.2. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$. *If S is an invertible matrix then* Rank(*SA*) = Rank(*A*).

Proof. By Theorem 2.2.22, RREF(*A*) = RREF(*SA*). Hence, Rank(*SA*) = Rank(*A*).

We now have the following result.

Corollary 3.1.3. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$. *and* $\boldsymbol{B} \in \mathbb{M}_{n,q}(\mathbb{C})$. *Then,* Rank(*AB*) ≤ Rank(*A*). *In particular, if* $\boldsymbol{B} \in \mathbb{M}_n(\mathbb{C})$. *is invertible then* Rank(*AB*) = Rank(*A*).

Proof. Let Rank(*A*) = *r*. Then, there exists an invertible matrix *P* and $\boldsymbol{A_1} \in \mathbb{M}_{n,q}(\mathbb{C})$. such that *PA* = RREF(*A*) = $\begin{bmatrix} A_1 \\ 0 \end{bmatrix}$. Then, *PAB* = $\begin{bmatrix} A_1 \\ 0 \end{bmatrix}$. *B* = $\begin{bmatrix} A_1 B \\ 0 \end{bmatrix}$. So, using Lemma 3.1.2 and Remark 3.1.2.2, we get

$$\text{Rank}(AB) = \text{Rank}(PAB) = \text{Rank}\left(\begin{bmatrix} A_1 B \\ 0 \end{bmatrix}\right) = \text{Rank}(A_1 B) \leq r = \text{Rank}(A).$$

(A)

In particular, if *B* is invertible then, using Equation (A), we get]

$$\text{Rank}(A) = \text{Rank}(ABB^{-1}) \leq \text{Rank}(AB)$$

and hence the required result follows.

Theorem 3.1.4 *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$ If* Rank $(A) = r$ *then, there exist invertible matrices P and Q such that*

$$PAQ = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Proof. Let C = RREF(A). Then, by Remark 2.2.19.4 there exists as invertible matrix P such that $C = PA$. Note that C has r pivots and they appear in columns, say $i_1 < i_2 < \cdots < i_r$. Now, let $D = CE_{1i1}E_{2i2} \cdots E_{rir}$. As E_{jij} 's are elementary matrices that interchange the columns of C, one has

$$D = \begin{bmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $B \in \mathbb{M}_{r,n-r}(\mathbb{C})$

Put $Q_1 = E_{1i1}E_{2i2} \cdots E_{rir}$. Then, Q_1 is invertible. Let

$$Q_2 = \begin{bmatrix} I_r & -B \\ \mathbf{0} & I_{n-r} \end{bmatrix}.$$

Then, verify that Q_2 is invertible and

$$CQ_1Q_2 = DQ_2 = \begin{bmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} I_r & -B \\ 0 & I_{n-r} \end{bmatrix} = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Thus, if we put $Q = Q_1Q_2$ then Q is invertible and $PAQ = CQ = CQ_1Q_2 =$ $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ and hence, the required result follows.

.

Proposition 3.1.5 . *Let $A \in \mathbb{M}_n(\mathbb{C})$ be an invertible matrix.*

1. If $A = [A_1\ A_2]$ with $A_1 \in \mathbb{M}_{n,r}(\mathbb{C})$ and $A2 \in Mn,n-r(C)$ then Rank($A1$) *= r and*

Rank(A_2) = $n - r$.

2. If A = $\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ with B1 ∈ Ms,n(C) and B2 ∈ $\mathbb{M}_{n-s,n}(\mathbb{C})$ then Rank(B_1) = s and Rank(B_2) =
n − s. In particular, if B = A[S, :] and C = A[:, T], for some subsets S, T of [n] then Rank(B) = |S|
and Rank(C) = |T |.

Proof. Since A is invertible, RREF(A) = In. Hence, by Remark 2.2.19.4, there exists an
invertible matrix P such that PA = I
n. Thus,

$$\begin{bmatrix} PA_1 & PA_2 \end{bmatrix} = P\begin{bmatrix} A_1 & A_2 \end{bmatrix} = PA = I_n = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & I_{n-r} \end{bmatrix}.$$

Thus, $PA_1 = \begin{bmatrix} I_r \\ 0 \end{bmatrix}$ and $PA2 = \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix}$ So, using Corollary 3.1.3, Rank($A1$) = r. Also, note
that $\begin{bmatrix} 0 & I_{n-r} \\ I_r & 0 \end{bmatrix}$ is an invertible matrix and

$$\begin{bmatrix} \mathbf{0} & I_{n-r} \\ I_r & \mathbf{0} \end{bmatrix} PA_2 = \begin{bmatrix} \mathbf{0} & I_{n-r} \\ I_r & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ I_{n-r} \end{bmatrix} = \begin{bmatrix} I_{n-r} \\ \mathbf{0} \end{bmatrix}.$$

So, again by using Corollary 3.1.3,, Rank(A_2) = $n - r$, completing the proof of the first part.
For the second part, let us assume that Rank(B_1) = $t < s$. There exists an invertible matrix Q such that

$$QB_1 = \mathrm{RREF}(B_1) = \begin{bmatrix} C \\ \mathbf{0} \end{bmatrix}, \qquad \text{(A)}$$

for some matrix C, where C is in RREF and has exactly t pivots. Since $t < s$, $QB1$ has at least
one zero row. As $PA = I_n$, one has $AP = I_n$. Hence

$$\begin{bmatrix} B_1P \\ B_2P \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} P = AP = I_n = \begin{bmatrix} I_s & \mathbf{0} \\ \mathbf{0} & I_{n-s} \end{bmatrix}.$$

$$B_1P = \begin{bmatrix} I_s & \mathbf{0} \end{bmatrix} \text{ and } B_2P = \begin{bmatrix} \mathbf{0} & I_{n-s} \end{bmatrix}. \quad \text{(B)}$$

Further, using Equations (A) and (B), we see that

$$\begin{bmatrix} CP \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} C \\ \mathbf{0} \end{bmatrix} P = QB_1P = Q\begin{bmatrix} I_s & \mathbf{0} \end{bmatrix} = \begin{bmatrix} Q & \mathbf{0} \end{bmatrix}.$$

Thus, Q has a zero row, contradicting the assumption that Q is invertible.
Hence, Rank(B_1) = s.
Similarly, Rank(B_2) = $n - s$ and thus, the required result follows.

Corollary .3.1.6. *Let* $A \in M_{m,n}(\mathbb{C})$. *If* Rank($A$) = $r < n$ *then, there exists an invertible matrix* Q *and* $B \in M_{m,r}(\mathbb{C})$ *such that* $AQ = [B\ \mathbf{0}]$, *where* Rank(B) = r.

Proof. By Theorem 3.1.4, there exist invertible matrices P and Q such that $PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. If P^{-1} = hB Ci, where $B \in M_{m,r}(\mathbb{C})$ and $C \in M_{m,m-r}(\mathbb{C})$ then,

$$AQ = P^{-1}\begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} B & C \end{bmatrix}\begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} B & \mathbf{0} \end{bmatrix}.$$

Now, by Proposition 3.1.5, Rank(B) = r = Rank(A) as the matrix P^{-1} = [B C] is an invertible matrix. Thus, the required result follows.
As an application of Corollary 3.1.6, we have the following result.
Corollary 3.1.7 *Let* $A \in Mm,n$(C) *and* $B \in Mn,p$(C). *Then,* Rank(AB) ≤ Rank(B).

Proof. Let Rank(B) = r.
Then, by Corollary 3.1.6, there exists an invertible matrix Q and a matrix $C \in M_{n,r}(\mathbb{C})$ such that $BQ = [C\ \mathbf{0}]$ and Rank(C) = r.
Hence, $ABQ = [AC\quad \mathbf{0}] = [AC\quad \mathbf{0}]$.

Notes

Thus, using Corollary 3.1.3 and Remark, we get

Rank (AB) = Rank(ABQ) = Rank [AC **0**] = Rank(AC) $\leq r$ = Rank(B).

Proposition 3.1.8. *Let A, B* $\in M_{m,n}$(C). *Then, prove that* Rank($A + B$) $\leq$ Rank(A) + Rank(B). *In particular, if* $A = \sum_{i=1}^{k} x_i y_i^*$ *for some* $\mathbf{x}_i$, $\mathbf{y}_i \in$ C, *for* $1 \leq i \leq k$, *then* Rank(A) $\leq k$.

Proof. Let Rank(A) = r. Then, there exists an invertible matrix P and a matrix $A_1 \in M_{r,n}(\mathbb{C})$ such that PA = RREF(A) = $\begin{bmatrix} A_1 \\ 0 \end{bmatrix}$. Then,

$$P(A+B) = PA + PB = \begin{bmatrix} A_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1 + B_1 \\ B_2 \end{bmatrix}.$$

Now using Corollary 3.1.3 and the condition Rank(A) = Rank(A_1) = r, the number of rows of $A1$, we have

Rank($A + B$) = Rank(P ($A + B$)) $\leq r$ + Rank($B2$) $\leq r$ + Rank(B) = Rank(A) + Rank(B).

Thus, the required result follows. The other part follows, as Rank($x_i y_i^*$) = 1, for $1 \leq i \leq k$.

3.3 SOLUTION SET OF A LINEAR SYSTEM

Definition 3.2.1 [Basic, Free Variables] Consider the linear system $A\mathbf{x} = \mathbf{b}$. If RREF([A **b**]) =
[C **d**]. Then, the variables corresponding to the pivotal columns of C are called the **basic** variables and the variables that are not basic are called **free** variables.

Example: 1. If the system $A\mathbf{x} = \mathbf{b}$ in n variables is consistent and

RREF(A) has r

nonzero rows then, $A\mathbf{x} = \mathbf{b}$ has r basic variables and $n - r$ free variables.

2. Let

$$\mathrm{RREF}([A \quad \mathbf{b}]) = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence, x and y are basic variables and z is the free variable. Thus, the solution set of $A\mathbf{x} = \mathbf{b}$ is given by

$\{[x, y, z]^T \mid [x, y, z] = [1, 2 - z, z] = [1, 2, 0] + z[0, -1, 1]$, with z arbitrary$\}$.

3. Let

$$\mathrm{RREF}([A \quad \mathbf{b}]) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then, the system $A\mathbf{x} = \mathbf{b}$ has no solution as (RREF([A **b**]))[3, :] = [0 0 0 1].

Example: Consider a linear system $A\mathbf{x} = \mathbf{b}$. Suppose RREF([A **b**]) = [C **d**], where

$$[C \quad \mathbf{d}] = \begin{bmatrix} \boxed{1} & 0 & 2 & -1 & 0 & 0 & 2 & 8 \\ 0 & \boxed{1} & 1 & 3 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then to get the solution set, we observe the following.

1. C has 4 pivotal columns, namely, the columns 1, 2, 5 and 6. Thus, x_1, x_2, x_5 and x_6 are basic variables.

2. Hence, the remaining variables, x_3, x_4 and x_7 are free variables.

Therefore, the solution set is given by

$$\begin{bmatrix} x_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \\ \mathbf{x}_5 \\ \mathbf{x}_6 \\ \mathbf{x}_7 \end{bmatrix} = \begin{bmatrix} 8 - 2x_3 + x_4 - 2x_7 \\ 1 - x_3 - 3x_4 - 5x_7 \\ x_3 \\ x_4 \\ 2 + x_7 \\ 4 - x_7 \\ x_7 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ 0 \\ 0 \\ 2 \\ 4 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} -2 \\ -5 \\ 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix},$$

where x_3, x_4 and x_7 are arbitrary.

$$\text{Let } \mathbf{x}_0 = \begin{bmatrix} 8 \\ 1 \\ 0 \\ 0 \\ 2 \\ 4 \\ 0 \end{bmatrix}, \ \mathbf{u}_1 = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{u}_3 = \begin{bmatrix} -2 \\ -5 \\ 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

In this example, verify that

$C\mathbf{x}_0 = \mathbf{d}$, and for $1 \le i \le 3$, $C\mathbf{u}_i = \mathbf{0}$. Hence, it follows that $A\mathbf{x}_0 = \mathbf{d}$, and for $1 \le i \le 3, A\mathbf{u}_i = \mathbf{0}$.

Theorem 3.2.2: *Let* $A\mathbf{x} = \mathbf{b}$ *be a linear system in n variables with* *RREF*$([A\ \mathbf{b}]) = [C\ \mathbf{d}]$ *with* $\text{Rank}(A) = r$ *and* $\text{Rank}([A\ \mathbf{b}]) = r_a$.

1. Then, the system $A\mathbf{x} = \mathbf{b}$ is inconsistent if r < ra

2. Then, the system $A\mathbf{x} = \mathbf{b}$ is consistent if $r = r_a$.

(a) Further, A**x** = **b** has a unique solution if r = n.

(b) Further, A**x** = **b** has infinite number of solutions if r < n. In this case, there exist vectors $\mathbf{x}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-r} \in R_n$ with $A\mathbf{x}_0 = \mathbf{b}$ and $A\mathbf{u}_i = \mathbf{0}$, for $1 \le i \le n - r$. Furthermore, the solution set is given by $\{\mathbf{x}_0 + k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_{n-r}\mathbf{u}_{n-r} \mid k_i \in C, 1 \le i \le n-r\}$.

Proof. Part 1: As $r < r_a$, $([C\ \mathbf{d}])[r + 1, :] = [\mathbf{0}^T\ 1]$. Note that this row corresponds to the linear equation

$$0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = 1$$

which clearly has no solution. Thus, by definition and Theorem 2.1.17, $A\mathbf{x} = \mathbf{b}$ is inconsistent.

Part 2: As $r = r_a$, $[C\ \mathbf{d}]$ doesn't have a row of the form $[\mathbf{0}^T\ 1]$. Further, the number of pivots in $[C\ \mathbf{d}]$ and that in C is same, namely, r pivots. Suppose the pivots appear in columns $i_1, \ldots, i_r$ with $1 \le i_1 < \cdots < i_r \le n$. Thus, the variables x_{i_j}, for $1 \le j \le r$, are basic variables and the remaining $n - r$ variables, say $x_{t_1}, \ldots, x_{t_{n-r}}$, are free variables with $t_1 < \cdots < t_{n-r}$. Since C is in RREF, in terms of the free variables and basic variables, the l-th row of $[C\ \mathbf{d}]$, for $1 \le l \le r$, corresponds to the equation

$$x_{i_\ell} + \sum_{k=1}^{n-r} c_{\ell t_k} x_{t_k} = d_\ell \Leftrightarrow x_{i_\ell} = d_\ell - \sum_{k=1}^{n-r} c_{\ell t_k} x_{t_k}.$$

Thus, the system $C\mathbf{x} = \mathbf{d}$ is consistent. Hence, the system $A\mathbf{x} = \mathbf{b}$ is consistent and the solution set of the system $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ are the same. Therefore, the solution set of the system $C\mathbf{x} = \mathbf{d}$ (or equivalently $A\mathbf{x} = \mathbf{b}$) is given by

$$\begin{bmatrix} x_{i_1} \\ \vdots \\ x_{i_r} \\ x_{t_1} \\ x_{t_2} \\ \vdots \\ x_{t_{n-r}} \end{bmatrix} = \begin{bmatrix} d_1 - \sum\limits_{k=1}^{n-r} c_{1t_k} x_{t_k} \\ \vdots \\ d_r - \sum\limits_{k=1}^{n-r} c_{rt_k} x_{t_k} \\ x_{t_1} \\ x_{t_2} \\ \vdots \\ x_{t_{n-r}} \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{t_1} \begin{bmatrix} c_{1t_1} \\ \vdots \\ c_{rt_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{t_2} \begin{bmatrix} c_{1t_2} \\ \vdots \\ c_{rt_2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_{t_{n-r}} \begin{bmatrix} c_{1t_{n-r}} \\ \vdots \\ c_{rt_{n-r}} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

(A)

Part 2a: As $r = n$, there are no free variables. Hence, $x_i = d_i$, for $1 \le i \le n$, is the unique solution

Part 2b: Define

$$\mathbf{x}_0 = \begin{bmatrix} d_1 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and } \mathbf{u}_1 = \begin{bmatrix} c_{1t_1} \\ \vdots \\ c_{rt_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{u}_{n-r} = \begin{bmatrix} c_{1t_{n-r}} \\ \vdots \\ c_{rt_{n-r}} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Then, it can be easily verified that $A\mathbf{x}_0 = \mathbf{b}$ and, for $1 \le i \le n-r$, $A\mathbf{u}_i = \mathbf{0}$. Also, by Equation (A) the solution set has indeed the required form, where k_i corresponds to the free variable x_{t_i}. As there is at least one free variable the system has infinite number of solutions. Thus, the proof of the theorem is complete.

Corollary 3.2.3. *Let* $A \in M_{m,n}(\mathbb{C})$. *If* Rank($A$) $= r < \min\{m, n\}$ *then* $A\mathbf{x} = \mathbf{0}$ *has infinitely many solutions. In particular, if* $m < n$, *then* $A\mathbf{x} = \mathbf{0}$ *has infinitely many solutions. Hence, in either case, the homogeneous system* $A\mathbf{x} = \mathbf{0}$ *has at least one non-trivial solution.*

Remark 3.2.4. *Let* $A \in M_{m,n}(\mathbb{C})$. *Then, Theorem 3.2.2 implies that* $A\mathbf{x} = \mathbf{b}$ *is consistent if and only if* Rank(A) = Rank([A $\mathbf{b}$]). *Further, the vectors associated to the free variables in Equation (A) are solutions to the associated homogeneous system* $A\mathbf{x} = \mathbf{0}$.

Example.1 Determine the equation of the line/circle that passes through the points (−1, 4), (0, 1) and (1, 4).

Solution: The general equation of a line/circle in Euclidean plane is given by $a(x^2 + y^2) + bx + cy + d = 0$, where *a*, *b*, *c* and *d* are variables. Since this curve passes through the given points, we get a homogeneous system in 3 equations and 4 variables, namely

$$\begin{bmatrix} (-1)^2+4^2 & -1 & 4 & 1 \\ (0)^2+1^2 & 0 & 1 & 1 \\ 1^2+4^2 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{0}.$$

Solving this system, we get

$$[a, b, c, d] = [\tfrac{3}{13}d, 0, -\tfrac{16}{13}d, d].$$

Hence, choosing $d = 13$, the required circle is given by $3(x^2 + y^2) - 16y + 13 = 0$.

2. Determine the equation of the plane that contains the points (1, 1, 1), (1, 3, 2) and (2, −1, 2).

Solution: The general equation of a plane in space is given by $ax + by + cz + d = 0$, where *a*, *b*, *c* and *d* are variables. Since this plane passes through the 3 given points, we get a homogeneous system in 3 equations and 4 variables. So, it has a non-trivial solution, namely

$$[a, b, c, d] = [-\tfrac{4}{3}d, -\tfrac{d}{3}, -\tfrac{2}{3}d, d].$$

Hence, choosing $d = 3$, the required plane is given by $-4x - y + 2z + 3 = 0$.

3. Let

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 0 & -1 & 0 \\ 0 & -3 & 4 \end{bmatrix}$$

Notes

Then, find a non-trivial solution of $A\mathbf{x} = 2\mathbf{x}$. Does there exist a nonzero vector $\mathbf{y} \in R^3$ such that $A\mathbf{y} = 4\mathbf{y}$?

Solution: Solving for $A\mathbf{x} = 2\mathbf{x}$ is equivalent to solving $(A - 2I)\mathbf{x} = \mathbf{0}$. The augmented matrix of this system equals

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

Verify that $\mathbf{x}^T = [1, 0, 0]$ is a non-zero solution. For the other part, the augmented matrix for solving $(A - 4I)\mathbf{y} = \mathbf{0}$ equals

$$\begin{bmatrix} -2 & 3 & 4 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & -3 & 0 & 0 \end{bmatrix}.$$

Thus, verify that $\mathbf{y}^T = [2, 0, 1]$ is a nonzero solution.

3.4 SQUARE MATRICES AND LINEAR SYSTEMS

In this section the coefficient matrix of the linear system $A\mathbf{x} = \mathbf{b}$ will be a square matrix.

Theorem 3.3.1. Let $A \in M_n(\mathbb{C})$. Then, the following statements are equivalent.

1. A is invertible.

2. RREF(A) = In.

3. A is a product of elementary matrices.

4. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

5. Rank(A) = n.

Proof. 1 ⇔ 2 Already done in Proposition 2.4.3.

2 ⇔ 3 Again, done in Proposition 2.4.3.

3 ⇒ 4 Let $A = E_1 \cdot\cdot\cdot E_k$, for some elementary matrices $E_1, \ldots, E_k$. Then, by previous equivalence A is invertible. So, A^{-1} exists and $A^{-1} A = I_n$. Hence, if $\mathbf{x}_0$ is any solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$ then, $\mathbf{x}_0 = I_n \cdot \mathbf{x}_0 = (A^{-1} A)\mathbf{x}_0 = A^{-1}(A\mathbf{x}_0) = A^{-1} \mathbf{0} = \mathbf{0}$.
Thus, $\mathbf{0}$ is the only solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$.

4 ⇒ 5 Let if possible Rank(A) = r < n. Then, by Corollary 3.2.3, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has infinitely many solution. A contradiction. Thus, A has full rank.

5 ⇒ 2 Suppose Rank(A) = n. So, RREF(A) has n pivotal columns. But, RREF(A) has exactly n columns and hence each column is a pivotal column. Thus, RREF(A) = In.
We end this section by giving two more equivalent conditions for a matrix to be invertible.

Theorem 3.3.2. The following statements are equivalent for $A \in M_n(\mathbb{C})$.
1. A is invertible.
2. The system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.
3. The system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$.

Proof. 1 ⇒ 2 Note that $\mathbf{x}_0 = A^{-1} \mathbf{b}$ is the unique solution of $A\mathbf{x} = \mathbf{b}$.
2 ⇒ 3 The system is consistent as $A\mathbf{x} = \mathbf{b}$ has a solution.
3 ⇒ 1 For $1 \leq i \leq n$, define $\mathbf{e}_\mathbf{i}^\mathbf{T}$ = In[i, :]. By assumption, the linear system $A\mathbf{x} = \mathbf{e}i$ has a solution, say $\mathbf{x}i$, for $1 \leq i \leq n$. Define a matrix $B = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$. Then,

$$AB = A[\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_n] = [A\mathbf{x}_1, A\mathbf{x}_2 \ldots, A\mathbf{x}_n] = [\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n] = I_n.$$

Therefore, $n = \text{Rank}(I_n) = \text{Rank}(AB) \leq \text{Rank}(A)$ and hence Rank(A) = n. Thus, by Theorem 3.3.1, A is invertible.

Theorem 3.3.3. The following two statements cannot hold together for A $\in M_n(\mathbb{C})$.

1. The system A**x** = **b** has a unique solution for every **b**.
2. The system A**x** = **0** has a non-trivial solution.

Corollary 3.3.4. Let A $\in M_n(\mathbb{C})$. Then, the following holds.

1. Suppose there exists C such that CA = In. Then, A^{-1} exists.
2. Suppose there exists B such that AB = In. Then, A^{-1} exists.

Check your progress

1. Explain Rank of a Matrix with example

2. Define basic and free variable

3. The following statements are equivalent for A ∈ Mn(ℂ).

1. A is invertible.

2. The system Ax = b has a unique solution for every b.

3. The system Ax = b is consistent for every b.

3.5 SUMMARY

Useful **application** of calculating the **rank** of a **matrix** is the computation of the number of solutions of a system of linear equations.

3.6 KEYWORDS

1. Statement - a **mathematical statement** is a sentence which is either true or false
2. Consistent - In **mathematics** and in particularly in algebra, a linear or nonlinear system of equations is called as **consistent** if there is at least one set of values for the unknowns that satisfies each equation in the system—that is, that when substituted into each of the equations makes each equation hold true as an identity.
3. Homogeneous system - A linear **system** of equations Ax = b is called **homogeneous** if b = 0, and non-**homogeneous** if b = 0. Notice that x = 0 is always solution of the **homogeneous equation**.
4. In linear **algebra**, an **augmented matrix** is a **matrix** obtained by appending the columns of two given **matrices**, usually for the purpose of performing the same elementary row operations on each of the given **matrices**.

3.7 QUESTION FOR REVIEW

1. Let P and Q be invertible matrices. Then, prove that Rank(PAQ) = Rank(A).

2. Prove that if Rank(A) = Rank(AB) then A = ABX, for some matrix X. Similarly, if Rank(A) = Rank(BA) then A = Y BA, for some matrix Y .

3. Let **u** = $(1, 1, -2)^T$ and **v** = (−1, 2, 3)T . Find condition on x, y and z such that the systemc**u** + d**v** = (x, y, z)T in the variables c and d is consistent.

4.. Find the condition(s) on x, y, z so that the systems given below (in the variables a, b and
c) is consistent?
(a) a + 2b − 3c = x, 2a + 6b − 11c = y, a − 2b + 7c = z.

5. Let A and B be two matrices having positive entries and of orders 1 × n and n × 1, respectively. Which of BA or AB is invertible? Give reasons.

3.8 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.

2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.

3. S. Lang, Linear Algebra, Springer, 1989.

4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.

5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.

6. Thomas Hungerford, Algebra, Springer GTM.

7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.

D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc

3.9 ANSWER TO CHECK YOUR PROGRESS

1. [HINT : Provide definition and example 3.1.1]

2. [HINT: Provide definition -3.2.1]

3. Provide the proof of the theorem 3.3.2

UNIT-4 LINEAR TRANSFORMATION AND DETERMINANT

STRUCTURE

4.0 Objective

4.1 Introduction

4.2 Formation Of Determinant

4.3 Determinant Of Third Order

4.4 Properties Of Determinant

4.5 Cramer's Rule

4.6 Inverse Of Matrix Using Determinant

4.7 Summary

4.8 Keywords

4.9 Questions for review

4.10 Suggested Readings

4.11 Answers To Check Your Progress

4.0 OBJECTIVE

Understand the concept of determinants and its application in Linear Transformation

Understand Cramer's Rule

4.1 INTRODUCTION

In linear algebra, the **determinant** is a scalar value that can be computed from the elements of a square matrix and encodes certain properties of the linear transformation described by the matrix. Geometrically, it can be viewed as the volume scaling factor of the linear transformation described by the matrix. This is also the signed volume of the *n*-dimensional parallelepiped spanned by the column or row vectors of the matrix.

It can be computed from the entries of the matrix by a specific arithmetic expression. The determinant provides important information about a matrix of coefficients of a system of linear equations, or about a matrix that corresponds to a linear transformation of a vector space.

The determinant of a matrix A is denoted by det(A), det A, or $|A|$. The determinant is denoted by surrounding the matrix entries by vertical bars.

Example:

$\begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix}$ OR det $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$

4.2 FORMATION OF DETERMINANTS:

Let us consider the following equations:

$$a_1x + b_1y = d_1 \quad \ldots (1)$$

$$a_2x + b_2y = d_2 \quad \ldots (2)$$

Multiply equation (1) by b_2 and equation (2) by b_1 we get

$$a_1b_2x + b_1b_2y = d_1b_2 \quad \ldots(3)$$

$$a_2b_1x + b_1b_2y = d_2b_1 \quad \ldots(4)$$

Now, subtracting (4) from (3) we get

$$\Rightarrow (a_1b_2 - a_2b_1)x = d_1b_2 - d_2b_1 \quad \ldots(A)$$

Similarly, multiply equation (1) by a_1 and equation (2) by a_2 we get the following equations

$$a_1a_2x + b_1a_2y = d_1a_2 \quad \ldots(5)$$

$$a_2a_1x + a_1b_2y = d_2a_1 \quad \ldots(6)$$

Now, subtracting (5) from (6) we get

$$\Rightarrow (a_1b_2 - a_2b_1)y = a_1d_2 - a_2d_1 \quad \ldots(B)$$

If $a_1b_2 - a_2b_1 \neq 0$, then from equation (A) and (B) we can write

$$x = \frac{d_1b_2 - d_2b_1}{a_1b_2 - a_2b_1}$$

And

$$y = \frac{a_1d_2 - a_2d_1}{a_1b_2 - a_2b_1}$$

Now, we can write the common denominator of x and y i.e. $a_1b_2 - a_2b_1$ as $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ and it is called as a determinant of second order.

The value of this determinant is $a_1b_2 - a_2b_1$

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

Here blue colour arrow $\longrightarrow$ indicates +ve direction and red colour arrow $\longrightarrow$ indicates –ve direction. When we multiply in the direction of blue arrow we take product as positive while in the direction of red colour we consider the product as negative.

Hence, we get $a_1b_2 - a_2b_1$

Here a_1, a_2, b_1 *and* b_2 are known as the elements of the determinant.

Example: $\begin{vmatrix} 2 & 4 \\ 1 & 5 \end{vmatrix} = 2 \times 5 - 4 \times 1 = 10 - 4 = 6$

$\begin{vmatrix} 4 & 2 \\ -1 & -3 \end{vmatrix} = 4 \times (-3) - 2 \times (-1) = -12 + 2 = -10$

We can represent the elements in a row of determinat between round brackets ( ) and elements in a column of determinant between square bracket[ ].

Consider $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$

We can represent the elements in a row as (a_1, b_1) and (a_2, b_2) respectively.

We can represent the elements in a column as $[a_1, a_2]$ and $[b_1, b_2]$ respectively.

4.3 DETERMINANT OF A THIRD ORDER:

Let us consider the three set of linear equations as:

$a_1x + b_1y + c_1z = d_1$ … (1)

$a_2x + b_2y + c_2z = d_2$ … (2)

$a_3x + b_3y + c_3z = d_3$ … (3)

If we solve simultaneously the above three equations, we will get a common denominator of the value x, y and z as:

$$a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - c_3a_2) + c_1(a_2b_3 - a_3 b_2)$$

We can represent the above equation in the following manner:

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

It is known as a determinant of third order.

It is evaluated as $a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$

4.3.1Rule for Evaluating the Third order Determinant:

We will consider the example:

$$\begin{vmatrix} 1 & 2 & -1 \\ -3 & 2 & 4 \\ 2 & -1 & 5 \end{vmatrix} = 1 \begin{vmatrix} 2 & 4 \\ -1 & 5 \end{vmatrix} - 2 \begin{vmatrix} -3 & 4 \\ 2 & 5 \end{vmatrix} - 1 \begin{vmatrix} -3 & 2 \\ 2 & -1 \end{vmatrix}$$

1.Write the elements from the first row with alternate positive and negative sign, the first element with positive sign, second with negative and third is positive

2.Then we will consider a determinant of second order which we obtained after omitting the corresponding row and column related to the elements of 1^{st} row.

Like, if we consider a_1 element i.e. 1 in the above example, we will omit all the first row and first column and would be left with a determinant of four elements i.e. b_2 , b_3 , c_2 *and* c_3 which are 2, – 1 , 4 and 5 in the above example.

Similarly, when we consider element a_2 we will omit 2nd row and 2nd column and for a_3 omit 3rd row and 3rd column.

Let's solve the above example.

= $1\,[2 \times 5 - 4 \times (-1)] - 2[(-3) \times 5 - 4 \times 2] - 1[(-3) \times (-1) - 2 \times 2]$

= $1\,[10 + 4] - 2[-15 - 8] - 1[3 - 4]$

= $1\,(14) - 2(-23) - 1(-1)$

= $14 + 46 + 1$

=61

[NOTE: In case of a determinant Number of rows = Number of columns always]

A determinant of **n**$^{\text{th}}$order contains **n**^{2} elements arranged in **n** rows and **n** columns illustrated below:

$$\begin{vmatrix} a_{11} & a_{12} \ldots & a_{1n} \\ a_{21} & a_{22} \ldots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} \ldots & a_{nn} \end{vmatrix}$$

We can represent the elements of the determinant in generalized form as $\mathbf{a_{ij}}$where **i**ndicates the row and **j** indicates the column position. [i = 1, 2, 3,…,n and j = 1,2, 3, …, n].

Notation: A determinant is denoted by Δ. A second order determinant may be denoted by Δ_2 and third order by Δ_3.

Example:

1. Find the value of the following determinants by expanding them.

   a. $\begin{vmatrix} b & 0 \\ b & -b \end{vmatrix}$

Solution: $\begin{vmatrix} b & 0 \\ b & -b \end{vmatrix} =$

$= b \times (-b) - 0 \times b$

$= -b^2 - 0$

$= -b^2$

b. $\begin{vmatrix} 0 & -h & g \\ h & 0 & -f \\ -g & f & 0 \end{vmatrix}$

Solution: $\begin{vmatrix} 0 & -h & g \\ h & 0 & -f \\ -g & f & 0 \end{vmatrix} =$

$= 0 \begin{vmatrix} 0 & -f \\ f & 0 \end{vmatrix} - (-h) \begin{vmatrix} h & -f \\ -g & 0 \end{vmatrix} + g \begin{vmatrix} h & 0 \\ -g & f \end{vmatrix}$

$= 0\left(0 — f^2\right) + h(0 - fg) + g\,(hf - 0)$

$= 0 - hfg + ghf$

$= 0.$

Check Your Progress

1. Explain formation of determinant

---

2 Explain Rule for Evaluating the Third order Determinant

4.4 PROPERTIES OF A DETERMINANT:

Property 1:If every element of a row (or a column) of a determinant is zero, then the determinant vanishes.

Proof:

$$\Delta = \begin{vmatrix} 0 & 0 & 0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

$$= 0 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - 0 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + 0 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= 0.\,(b_2c_3 - b_3c_2) - 0\,(a_2c_3 - a_3c_2) + 0\,(a_2b_3 - a_3b_2)$$

$$= 0$$

Similarly, we can show that if all elements of any other row or column are zeros then the value of determinant will be zero.

Property 2: The value of a determinant remains unaltered if rows and columns are interchanged.

Proof:

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

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

$$= a_1\,(b_2c_3 - b_3c_2) - b_1\,(a_2c_3 - a_3c_2) + c_1\,(a_2b_3 - a_3b_2)$$

$$= a_1b_2c_3 - a_1b_3c_2 - b_1a_2c_3 + b_1a_3c_2 + c_1\,a_2b_3 - c_1\,a_3b_2 \quad \ldots(1)$$

Now we will interchange the rows and columns of the determinant Δ and denote it as Δ'

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

Similarly, expanding the elements of Δ' we get

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

$= a_1\,(b_2c_3 - b_3c_2) - b_1\,(a_2c_3 - a_3c_2) + c_1\,(a_2b_3 - a_3b_2) =$ $a_1b_2c_3 - a_1b_3c_2 - b_1a_2c_3 + b_1a_3c_2 + c_1\,a_2b_3 - c_1\,a_3b_2$ …(2)

From (1) and (2) we find that $\Delta = \Delta'$ and hence the property is proved.

Property 3: If two rows (or columns) of a determinant are interchanged then the value of the new determinant become (– 1) time the value of original determinant.

Proof:

$$\text{Let } \Delta_1 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Let Δ_2 be the determinant that is obtained from Δ_1 by interchanging the first row and the second row, we can write

$$\Delta_2 = \begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Expanding Δ_1 we get $= a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_3b_1c_2 + a_2b_3c_1 - a_3b_2c_1$ …(1)

Expanding Δ_2 we get,

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

$= a_2\,(b_1c_3 - b_3c_1) - b_2\,(a_1c_3 - a_3c_1) + c_2\,(a_1b_3 - a_3b_1)$

$= a_2b_1c_3 - a_2b_3c_1 - a_1b_2c_3 + a_3b_2c_1 + a_1b_3c_2 - a_3b_1c_2$

Rearranging the terms

$= -a_1b_2c_3 + a_1b_3c_2 + a_2b_1c_3 - a_3b_1c_2 - a_2b_3c_1 + a_3b_2c_1$

$= -(a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_3b_1c_2 + a_2b_3c_1 - a_3b_2c_1)$ …(2)

From (1) and (2) we can say that

$$\Delta_1 = -\Delta_2$$

Property 4: ***If two rows (or columns) of a determinant are identical , the value of the determinant is zero.***

Proof:

Let $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}$

Here, first two rows of the above determinant are identical.

Expanding the above determinant we get

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

$$= a_1 (b_1c_3 - b_3c_1) - b_1 (a_1c_3 - a_3c_1) + c_1 (a_1b_3 - a_3b_1)$$

$$= a_1b_1c_3 - a_1b_3c_1 - a_1b_1c_3 + a_3b_1c_1 + a_1b_3c_1 - a_3b_1c_1 \ldots(1)$$

Rearranging the above equation as

$$= a_1b_1c_3 - a_1b_1c_3 - a_1b_3c_1 + a_1b_3c_1 + a_3b_1c_1 - a_3b_1c_1$$

$$= 0 + 0 + 0$$

$$= 0$$

Property 5: If all elements of one row (or column) of a determinant be multiplied by the same constant number k(say) then the determinant itself is multiplied by that constant k.

Proof: Let

$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ and $\Delta' = \begin{vmatrix} ka_1 & kb_1 & kc_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

Here we need to show that $\Delta' = k\Delta$

Now, we will expand the determinant Δ'

$$\Delta' = ka_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - kb_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + kc_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

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

$$= k \left[\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \right]$$

$$\Delta' = k\,\Delta$$

4.4.1 Minors And Cofactors Of The Element Ofa Determinant:

4.4.1.1 Minor: The minor of an element in a determinant is the determinant obtained by deleting the row and the column in which that element appears.

Let us consider the example as:

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

Minor of a_1 i.e. $M_{11} = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$

Minor of b_1 $M_{12} = \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}$

Minor of $c_1 M_{13} = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$

Thus, in other words we can say that a minor of an element in a third order determinant is a second order determinant and the sign before it is always positive.

4.4.1.2 COFACTOR:

The cofactor of any element in a determinant is the minor of that element with + or – sign according to the formula

$\mathbf{C_{ij} = (-1)^{i+j} M_{ij}}$

Where $\mathbf{C_{ij}}$is the Cofactor of the element**,**

i = denotes the number of row of that specific element which is considered.

j = denotes the number of row of that specific element which is considered.

$\mathbf{M_{ij}}$= minor of that considered element

So, the cofactors of a_1, b_1& c_1 in the above example are as follows,

$$C_{11} = (-1)^{1+1}\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$$

$$C_{12} = (-1)^{1+2}\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}$$

$$C_{13} = (-1)^{1+3}\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

Also, if we expand the determinant Δ we get,

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

$$\Delta = a_1\, C_{11} + b_1\, C_{12} + c_1\, C_{13}$$

For quick working, the sign of the different cofactors are as follows:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Example: Find the minor and cofactor of each element of the following determinant

$$\begin{vmatrix} 3 & -2 & 1 \\ 5 & 4 & -3 \\ -2 & 6 & -1 \end{vmatrix}$$

Solution:

$$M_{11} = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} = \begin{vmatrix} 4 & -3 \\ 6 & -1 \end{vmatrix} = -4 + 18 = 14$$

$$M_{12} = \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ -2 & -1 \end{vmatrix} = -3 + 2 = -1$$

$$M_{13} = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = \begin{vmatrix} 5 & 4 \\ -2 & 6 \end{vmatrix} = 30 + 8 = 38$$

Similarly,

$M_{21} = \begin{vmatrix} -2 & 1 \\ 6 & -1 \end{vmatrix} = 2 - 6 = -4$ $\qquad$ $M_{31} = \begin{vmatrix} -2 & 1 \\ 4 & -3 \end{vmatrix} =$
$6 - 4 = 2$

$M_{22} = \begin{vmatrix} 3 & 1 \\ -2 & -1 \end{vmatrix} = -3 + 2 = -1$ $M_{32} = \begin{vmatrix} 3 & 1 \\ 5 & -3 \end{vmatrix} = -9 - 5 = -14$

$M_{23} = \begin{vmatrix} 3 & -2 \\ -2 & 6 \end{vmatrix} = 18 - 4 = 14$ $M_{33} = \begin{vmatrix} 3 & -2 \\ 5 & 4 \end{vmatrix} = 12 + 10 = 22$

Thus, Cofactors of the elements are

$C_{11} = (-1)^{1+1} \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} = (-1)^2 \begin{vmatrix} 4 & -3 \\ 6 & -1 \end{vmatrix} = 14$

$C_{12} = (-1)^{1+2} \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} = (-1)^3 \begin{vmatrix} 3 & 1 \\ -2 & -1 \end{vmatrix} = (-1)(-1) = 1$

$C_{13} = (-1)^{1+3} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = (-1)^4 \begin{vmatrix} 5 & 4 \\ -2 & 6 \end{vmatrix} = 38$

Similarly,

$C_{21} = (-1)^{2+1} M_{21} = (-1)^3(-4) = 4$ $C_{31} = (-1)^{3+1} M_{31} = (-1)^4(2) = 2$

$C_{22} = (-1)^{2+2} M_{22} = (-1)^4(-1) = -1$ $C_{32} = (-1)^{3+2} M_{32} = (-1)^5(-14) = 14$

$C_{23} = (-1)^{2+3} M_{23} = (-1)^5(14) = -14$ $C_{33} = (-1)^{3+3} M_{33} = (-1)^6(22) = 22$

Property 6:If every element of a row (or column) of a determinant is the sum of two terms then the determinant can be expressed as the sum of two determinants.

Proof: Let $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

and the cofactors of a_1 , b_1 and c_1 is C_{11}, C_{12} and C_{13} respectively.

Consider $\Delta' = \begin{vmatrix} a_1 + \alpha_1 & b_1 + \alpha_2 & c_1 + \alpha_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

Now, the cofactors of $a_1 + \alpha_1$, $b_1 + \alpha_2$ and $c_1 + \alpha_3$ is C_{11}, C_{12} and C_{13} respectively.

Expand Δ' determinant with respect to its cofactor we get following expression;

$\Delta' = (a_1 + \alpha_1)\, C_{11} + (b_1 + \alpha_2)\, C_{12} + (c_1 + \alpha_3)\, C_{13}$

$= a_1 C_{11} + \alpha_1 C_{11} + b_1 C_{12} + \alpha_2 C_{12} + c_1 C_{13} + \alpha_3 C_{13}$

$= a_1 C_{11} + b_1 C_{12} + c_1 C_{13} + \alpha_1 C_{11} + \alpha_2 C_{12} + \alpha_3 C_{13}$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Also, from the above theorem we can write,

$$\Delta = a_1C_{11} + b_1C_{12} + c_1C_{13} = a_2C_{21} + b_2C_{22} + c_2C_{23}$$
$$= a_3C_{31} + b_3C_{32} + c_3C_{33}$$

$$= a_1C_{11} + a_2C_{21} + a_3C_{31} = b_1C_{12} + b_2C_{22} + b_3C_{32}$$
$$= c_1C_{13} + c_2C_{23} + c_3C_{33}$$

Property 7: If the elements of a certain row(or column) of a determinant are multiplied by the cofactors of the corresponding elements of another row(or column) then the sum of these product will be zero.

Proof: Let $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

Now, we will multiply the element of first row by the cofactors of the corresponding elements of the second row, we get

$\Delta = a_1 C_{21} + b_1 C_{22} + c_1 C_{23}$

$$= a_1(-1)^{2+1} \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + (-1)^{2+2} b_1 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} + (-1)^{2+3} c_1 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

$= -a_1(b_1c_3 - b_3c_1) + b_1(a_1c_3 - a_3c_1) - c_1(a_1b_3 - a_3b_1)$

$= -a_1b_1c_3 + a_1b_3c_1 + b_1a_1c_3 - b_1a_3c_1 - c_1a_1b_3 + c_1a_3b_1)$

$= 0$

Similarly, we can show that $a_1 C_{31} + b_1 C_{32} + c_1 C_{33} = a_2 C_{11} + b_2 C_{12} + c_2 C_{13} = 0$

Property 8: ***The value of a determinant remains unaltered if to all elements of any row (or column) are added the same multiples of the corresponding elements of any number of other rows (columns).***

Proof: We have to show that

$$\begin{vmatrix} a_1 + ma_2 + na_3 & b_1 + mb_2 + nb_3 & c_1 + mc_2 + nc_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

With reference to property 6, we can expand LHS as follows

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} ma_2 & mb_2 & mc_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} na_2 & nb_2 & nc_2 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

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

[Property 5]

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + m.\,0 + n.\,0$$

[Property 4]

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

= RHS

Similarly, we can prove that

$$\begin{vmatrix} a_1 + mb_1 + nc_1 & b_1 & c_1 \\ a_2 + mb_2 + nc_2 & b_2 & c_2 \\ a_3 + mb_3 + nc_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Examples:

1. Show that $\begin{vmatrix} a^2 & a & bc \\ b^2 & b & ca \\ c^2 & c & ab \end{vmatrix} = \begin{vmatrix} a^3 & a^2 & 1 \\ b^3 & b^2 & 1 \\ c^3 & c^2 & 1 \end{vmatrix}$

Proof: $LHS = \begin{vmatrix} a^2 & a & bc \\ b^2 & b & ca \\ c^2 & c & ab \end{vmatrix}$

Multiply first row by a , second by b and third row by c, we get

$$= \frac{1}{abc}\begin{vmatrix} a^3 & a^2 & abc \\ b^3 & b^2 & bca \\ c^3 & c^2 & cab \end{vmatrix}$$

$$= \frac{abc}{abc}\begin{vmatrix} a^3 & a^2 & 1 \\ b^3 & b^2 & 1 \\ c^3 & c^2 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a^3 & a^2 & 1 \\ b^3 & b^2 & 1 \\ c^3 & c^2 & 1 \end{vmatrix}$$

= RHS

2. Show that $\Delta = \begin{vmatrix} x+y & z & z-x \\ y+z & x & x-y \\ z+x & y & y-z \end{vmatrix} = x^3 + y^3 + z^3 - 3xyz$

Proof: $\Delta = \begin{vmatrix} x+y & z & z-x \\ y+z & x & x-y \\ z+x & y & y-z \end{vmatrix}$

Now we will perform the following operation on the above determinant

$C_1 \to C_1 + C_2$ and $C_3 \to C_3 - C_2$ we get

$$= \begin{vmatrix} x+y+z & z & -x \\ x+y+z & x & -y \\ z+x+y & y & -z \end{vmatrix}$$

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

Now, $R_1 \to R_1 - R_3$ and $R_2 \to R_2 - R_3$

$$= (x+y+z)\begin{vmatrix} 0 & z-y & z-x \\ 0 & x-y & z-y \\ 1 & y & -z \end{vmatrix}$$

Expanding along the elements of the 1st column

$= (x+y+z)\{(z-y)^2 - (x-y)(z-x)\}$

$= (x+y+z)\{z^2 + y^2 - 2yz - xz + x^2 + yz - xy\}$

$= (x+y+z)\{x^2 + y^2 + z^2 - yz - zx - xy\}$

$= x^3 + y^3 + z^3 - 3xyz.$

= RHS

3. Show that $\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)^3$

Solution: Consider LHS $= \begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$

$R_1 \rightarrow R_1 + R_2 + R_3$ we get

$$= \begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$$= (a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$C_2 \rightarrow C_2 - C_1$ and $C_3 \rightarrow C_3 - C_1$

$$= (a+b+c)\begin{vmatrix} 1 & 0 & 0 \\ 2b & -(a+b+c) & 0 \\ 2c & 0 & -(a+b+c) \end{vmatrix}$$

Expanding along the elements of the first row

$= (a+b+c)(a+b+c)^2$

$= (a+b+c)^3$

4. Solve $\begin{vmatrix} a & a & a \\ b & x & x \\ c & b & a \end{vmatrix} = 0$

Solution: Perform $C_2 \to C_2 - C_1$ and $C_3 \to C_3 - C_1$, we will get

$$\to \begin{vmatrix} a & 0 & 0 \\ b & x-b & x-b \\ c & b-c & a-c \end{vmatrix} = 0$$

$\to a[(x-b)(a-c) - (x-b)(b-c)] = 0$

$\to (x-b)[a(a-c) - a(b-c)] = 0$

$\to (x-b) = 0 \; or \; [a(a-c) - a(b-c)] = 0$

Assuming $[a(a-c) - a(b-c)] \neq 0$

$\to (x-b) = 0$

$\to x = b$

4.5 CRAMER'S RULE:

We will consider the example

Example 1: Related to two variable

Solve the following system by using determinants.

$$\begin{cases} 4x - 3y = -14 \\ 3x - 5y = -5 \end{cases}$$

To solve this system, three determinants are created. One is called the denominator determinant, labelled as *D;* another is the *x*-numerator determinant, labelled D_x; and the third is the *y*-numerator determinant, labelled D_y.

The denominator determinant, *D,* is formed by taking the coefficients of *x* and *y* from the equations written in standard form.

$$D = \begin{vmatrix} 4 & -3 \\ 3 & -5 \end{vmatrix} = 4(-5) - 3(-3)$$

$= -20 + 9$

$= -11$

The x-numerator determinant is formed by taking the constant terms from the system and placing them in the x-coefficient positions and retaining the y-coefficients.

$$D_x = \begin{vmatrix} -14 & -3 \\ -5 & -5 \end{vmatrix}$$

$$= -14(-5) - (-3)(-5)$$

$$= 70 - 15$$

$$= 55$$

The y-numerator determinant is formed by taking the constant terms from the system and placing them in the y-coefficient positions and retaining the x-coefficients.

$$D_y = \begin{vmatrix} 4 & -14 \\ 3 & -5 \end{vmatrix}$$

$$= 4(-5) - 3(-14)$$

$$= -20 - (-42)$$

$$= 22$$

The answers for x and y is as follows:

$$x = \frac{D_x}{D} = \frac{55}{-11} = -5$$

$$y = \frac{D_y}{D} = \frac{22}{-11} = -2$$

The solution is $x = -5$, $y = -2$.

Finding solutions by using determinants is referred to as **Cramer's Rule**, named after the mathematician who devised this method. Cramer's Rule could hardly be considered a "shortcut," but it is a rather neat way to solve systems of equations by using determinants.

For 3 variables

To use determinants to solve a system of three equations with three variables (Cramer's Rule), say *x*, *y*, and *z*, four determinants must be formed following this procedure:

1. Write all equations in standard form.
2. Create the denominator determinant, *D*, by using the coefficients of *x*, *y*, and *z* from the equations and evaluate it.
3. Create the *x*-numerator determinant, D_x, the *y*-numerator determinant, D_y, and the *z*-numerator determinant, D_z, by replacing the respective *x*, *y*, and *z* coefficients with the constants from the equations in standard form and evaluate each determinant.

The answers for *x*, *y*, and *z* are as follows:

$$x = \frac{D_x}{D}, y = \frac{D_y}{D}, z = \frac{D_z}{D}$$

Example 2: Solve this system of equations, using Cramer's Rule.

$$\begin{cases} 3x + 2y - z = 2 \\ 2x - y - 3z = 13 \\ x + 3y - 2z = 1 \end{cases}$$

Find the minor determinants.

x-coefficients, *y*-coefficients, *z*-coefficients

$$D = \begin{vmatrix} 3 & 2 & -1 \\ 2 & -1 & -3 \\ 1 & 3 & -2 \end{vmatrix} = 3\begin{vmatrix} -1 & -3 \\ 3 & -2 \end{vmatrix} - 2\begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} + 1\begin{vmatrix} 2 & -1 \\ -1 & -3 \end{vmatrix}$$

$$= 3[2-(-9)] - 2[-4-(-3)] + 1(-6-1)$$

$$= 3(11) - 2(-1) + 1(-7)$$

$$= 33 + 2 - 7 = 28$$

Notes

Use the constants to replace the *x*-coefficients.

Constants
replacing the
x-coefficient

$$D_x = \begin{vmatrix} 2 & 2 & -1 \\ 13 & -1 & -3 \\ 1 & 3 & -2 \end{vmatrix}$$

$$= 2\begin{vmatrix} -1 & -3 \\ 3 & -2 \end{vmatrix} - 13\begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} + 1\begin{vmatrix} 2 & -1 \\ -1 & -3 \end{vmatrix}$$

$$= 2[2 - (-9)] - 13[-4 - (-3)] + 1[(-6) - 1]$$

$$= 2[11] - 13[-1] + 1[-7]$$

$$= 22 + 13 - 7 = 28$$

Use the constants to replace the *y*-coefficients.

Constants
replacing the
y-coefficient

$$D_y = \begin{vmatrix} 3 & 2 & -1 \\ 2 & 13 & -3 \\ 1 & 1 & -2 \end{vmatrix}$$

$$= 3\begin{vmatrix} 13 & -3 \\ 1 & -2 \end{vmatrix} - 2\begin{vmatrix} 2 & -1 \\ 1 & -2 \end{vmatrix} + 1\begin{vmatrix} 2 & -1 \\ 13 & -3 \end{vmatrix}$$

$$= 3[-26 - (-3)] - 2[-4 - (-1)] + 1[(-6) - (-13)]$$

$$= 3[-23] - 2[-3] + 1[7]$$

$$= -69 + 6 + 7 = -56$$

Use the constants to replace the *z*-coefficients.

Constants replacing the z-coefficient ↓

$$D_z = \begin{vmatrix} 3 & 2 & 2 \\ 2 & -1 & 13 \\ 1 & 3 & 1 \end{vmatrix}$$

$$= 3\begin{vmatrix} -1 & 13 \\ 3 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} + 1\begin{vmatrix} 2 & 2 \\ -1 & 13 \end{vmatrix}$$

$$= 3[-1 - 39] - 2[2 - 6] + 1[26 - (-2)]$$

$$= 3[-40] - 2[-4] + 1[28]$$

$$= -120 + 8 + 28 = -84$$

Therefore,

$$x = \frac{D_x}{D} = \frac{28}{28} = 1, y = \frac{D_y}{D} = \frac{-56}{28} = -2 \; and \; z = \frac{D_z}{D} = \frac{-84}{28} = -3$$

The solution is $x = 1$, $y = -2$, $z = -3$.

[**NOTE:** If the denominator determinant, *D*, has a value of zero, then the system is either inconsistent or dependent.
The system is ***dependent*** if all the determinants have a value of zero.
The system is ***inconsistent*** if at least one of the determinants, D_x, D_y, or D_z, has a value not equal to zero and the denominator determinant has a value of zero.]

Example 3: Solve the following by Cramer's rule

$$3x - 5y = 7,$$

$$4x + y = 17$$

Now, we will find out the determinant D as

Notes

$D = \begin{vmatrix} 3 & -5 \\ 4 & 1 \end{vmatrix}$

$= 3\,(1) - 4(-5)$

$= 3 + 20$

$= 23 \neq 0$

We will find D_x and D_y as follows:

$D_x = \begin{vmatrix} 7 & -5 \\ 17 & 1 \end{vmatrix}$

$= 7(1) - 17(-5)$

$= 7 + 85$

$= 92$

$D_y = \begin{vmatrix} 3 & 7 \\ 4 & 17 \end{vmatrix}$

$= 3(17) - 7(4)$

$= 51 - 28$

$= 23$

$$x = \frac{D_x}{D} = \frac{92}{23} = 4$$

$$y = \frac{D_y}{D} = \frac{23}{23} = 1$$

The solution is $x = 4$ and $y = 1$.

Example 4: Solve the following equation by Cramer's Rule:

$x + y + z = 3,$

$2x + 3y + 4z = 9$

$x + 2y - 4z = -1$

Solution:

$$D = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 1 & 2 & -4 \end{vmatrix}$$

$$= 1\begin{vmatrix} 3 & 4 \\ 2 & -4 \end{vmatrix} - 1\begin{vmatrix} 2 & 4 \\ 1 & -4 \end{vmatrix} + 1\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}$$

$$= 1\,[3(-4) - 4(2)] - 1[2(-4) - 4(1)] + 1[2(2) - 3(1)]$$

$$= [-12 - 8] - [-8 - 4] + [4 - 3]$$

$$= -20 + 12 + 1$$

$$= -7$$

$$D_x = \begin{vmatrix} 3 & 1 & 1 \\ 9 & 3 & 4 \\ -1 & 2 & -4 \end{vmatrix}$$

$$= 3\begin{vmatrix} 3 & 4 \\ 2 & -4 \end{vmatrix} - 1\begin{vmatrix} 9 & 4 \\ -1 & -4 \end{vmatrix} + 1\begin{vmatrix} 9 & 3 \\ -1 & 2 \end{vmatrix}$$

$$= 3[-12 - (8)] - 1[-36 - (-4)] + 1[18 - 1(-3)]$$

$$= 3[-20] - 1[-32] + 1[21]$$

$$= -60 + 32 + 21 = -7$$

$$D_y = \begin{vmatrix} 1 & 3 & 1 \\ 2 & 9 & 4 \\ 1 & -1 & -4 \end{vmatrix}$$

$$= 1\begin{vmatrix} 9 & 4 \\ -1 & -4 \end{vmatrix} - 3\begin{vmatrix} 2 & 4 \\ 1 & -4 \end{vmatrix} + 1\begin{vmatrix} 2 & 9 \\ 1 & -1 \end{vmatrix}$$

$$= 1[-36 - (-4)] - 3[-8 - (4)] + 1[-2 - 9]$$

$$= 1[-32] - 3[-12] + 1[-11]$$

$$= -32 + 36 - 11 = -7$$

$$D_z = \begin{vmatrix} 1 & 1 & 3 \\ 2 & 3 & 9 \\ 1 & 2 & -1 \end{vmatrix}$$

$$= 1\begin{vmatrix} 3 & 9 \\ 2 & -1 \end{vmatrix} - 1\begin{vmatrix} 2 & 9 \\ 1 & -1 \end{vmatrix} + 3\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}$$

$$= 1[-3 - 18] - 1[-2 - 9] + 3[4 - 3]$$

$$= 1[-21] - 1[-11] + 3[1]$$

$$= -21 + 11 + 3 = -7$$

Therefore,

$$x = \frac{D_x}{D} = \frac{-7}{-7} = 1, y = \frac{D_y}{D} = \frac{-7}{-7} = 1 \; and \; z = \frac{D_z}{D} = \frac{-7}{-7} = 1$$

The solution is x = 1, y = 1, and z = 1.

4.6 INVERSE OF A MATRIX USING DETERMINANT:

We will consider the following example to understand the method of finding inverse using determinant.

Example1: Find the inverse of a matrix A = $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}$ if it exist.

Solution: Let us find the determinant of matrix A

$$|A| = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}$$

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

= 1 (12 – 6) – 2(4 – 3) + 3 (2 – 3)

= 6 – 2 – 3

= 1 ≠ 0

which indicates the inverse of given matrix A exists.

Now let's find out the adjoint of a matrix A

Calculate minor for each element of the given matrix A

$M_{11} = \begin{vmatrix} 3 & 3 \\ 2 & 4 \end{vmatrix} = 6$ $M_{21} = \begin{vmatrix} 2 & 3 \\ 2 & 4 \end{vmatrix} = 2$ $M_{31} = \begin{vmatrix} 2 & 3 \\ 3 & 3 \end{vmatrix}$

$= -3$

$M_{12} = \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1$ $M_{22} = \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1$ $M_{32} = \begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix}$

$= 0$

$M_{13} = \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = -1$ $M_{23} = \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0$ $M_{33} = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix}$

$= 1$

We have cofactor matrix as

$$C = \begin{bmatrix} M_{11} & -M_{12} & M_{13} \\ -M_{21} & M_{22} & -M_{23} \\ M_{31} & -M_{32} & M_{33} \end{bmatrix}$$

$$\text{adj } A = C^T = \begin{bmatrix} M_{11} & -M_{21} & M_{31} \\ -M_{12} & M_{22} & -M_{32} \\ M_{13} & -M_{23} & M_{33} \end{bmatrix}$$

$$\text{adj } A = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\therefore A^{-1} = \frac{1}{|A|} \text{adj } A$$

$$= \frac{1}{1}\begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Example 2: Find the inverse of a matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$ if it exist.

Solution: Let us find the determinant of matrix A

$$|A| = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$$

$|A| = 1\begin{vmatrix}4 & 5\\0 & 6\end{vmatrix} - 2\begin{vmatrix}0 & 5\\1 & 6\end{vmatrix} + 3\begin{vmatrix}0 & 4\\1 & 0\end{vmatrix}$

= 1 (24 – 0) – 2(0 – 5) + 3 (0 –4)

= 24 +10 – 12

= 22 $\neq$ 0

which indicates the inverse of given matrix A exists.

Now let's find out the adjoint of a matrix A

Calculate minor for each element of the given matrix A

$M_{11} = \begin{vmatrix}4 & 5\\0 & 6\end{vmatrix} = 24$ $M_{21} = \begin{vmatrix}2 & 3\\0 & 6\end{vmatrix} = 12$ $M_{31} = \begin{vmatrix}1 & 2\\0 & 4\end{vmatrix}$

= 4

$M_{12} = \begin{vmatrix}0 & 5\\1 & 6\end{vmatrix} = -5$ $M_{22} = \begin{vmatrix}1 & 3\\1 & 6\end{vmatrix} = 3$ $M_{32} = \begin{vmatrix}1 & 3\\0 & 5\end{vmatrix}$

= 5

$M_{13} = \begin{vmatrix}0 & 4\\1 & 0\end{vmatrix} = -4$ $M_{23} = \begin{vmatrix}1 & 2\\1 & 0\end{vmatrix} = -2$ $M_{33} = \begin{vmatrix}1 & 2\\0 & 4\end{vmatrix}$

= 4

We have cofactor matrix as

$$C = \begin{bmatrix} M_{11} & -M_{12} & M_{13}\\ -M_{21} & M_{22} & -M_{23}\\ M_{31} & -M_{32} & M_{33}\end{bmatrix}$$

$$\text{adj } A = C^T = \begin{bmatrix} M_{11} & -M_{21} & M_{31}\\ -M_{12} & M_{22} & -M_{32}\\ M_{13} & -M_{23} & M_{33}\end{bmatrix}$$

$$\text{adj } A = \begin{bmatrix} 24 & -12 & 4\\ 5 & 3 & -5\\ -4 & 2 & 4\end{bmatrix}$$

$\therefore A^{-1} = \frac{1}{|A|}\text{ adj } A$

$$A^{-1} = \frac{1}{22}\begin{bmatrix} 24 & -12 & 4 \\ 5 & 3 & -5 \\ -4 & 2 & 4 \end{bmatrix}$$

Check Your Progress

1. State property 8 with proof

2. State the steps to solve a system of three equations with three variables (Cramer's Rule)

4.7 SUMMARY

A matrix is often used to represent the coefficients in a system of linear equations, and the determinant can be used to solve those equations, although other methods of solution are much more computationally efficient. In linear algebra, a matrix (with entries in a field) is singular (not invertible) if and only if its determinant is zero. This leads to the use of determinants in defining the characteristic polynomial of a matrix, whose roots are the eigenvalues. In analytic geometry, determinants express the signed *n*-dimensional volumes of *n*-dimensional parallelepipeds. This leads to the use of determinants in calculus, the Jacobian determinant in the change of variables rule for integrals of functions of several variables. Determinants appear frequently in algebraic identities such as the Vandermonde identity.

4.8 KEYWORDS

1.Inverse - something that is the opposite or reverse of something else.

2. Coefficient - In a mathematical equation, a coefficient is a constant by which a variable is multiplied.

3. Replacing - a set of elements any one of which may be used **to replace** a given variable or placeholder in a **mathematical** sentence or expression

4. Constant - a **constant** is a number on its own, or sometimes a letter such as a, b or c to stand for a fixed number.

4.9 QUESTION FOR REVIEW

1. *Let* $A \in \mathrm{M}n(\mathrm{C})$ *be an upper triangular matrix with nonzero entries on the diagonal. Then, prove that* $A-1$ *is also an upper triangular matrix.*

2. Let

$$A = \begin{bmatrix} a & b & c \\ e & f & g \\ h & j & \ell \end{bmatrix} \text{ and } B = \begin{bmatrix} a & e & 10^2a + 10e + h \\ b & f & 10^2b + 10f + j \\ c & g & 10^2c + 10g + \ell \end{bmatrix}$$

where $a, b \dots, l \in \mathbb{C}$*. Without computing deduce that* $\det(A) = \det(B)$.

3. Solve $A\mathbf{x} = \mathbf{b}$ using Cramer's rule, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

4. *Solve the following linear system by Cramer's rule.*

i) $x + y + z - w = 1,\ x + y - z + w = 2,\ 2x + y + z - w = 7,\ x + y + z + w = 3.$

ii) $x - y + z - w = 1,\ x + y - z + w = 2,\ 2x + y - z - w = 7,\ x - y - z + w = 3.$

4.10 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.

2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.

3. S. Lang, Linear Algebra, Springer, 1989.

4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.

5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.

6. Thomas Hungerford, Algebra, Springer GTM.

7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.

8. D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.

4.11 ANSWER TO CHECK YOUR PROGRESS

1. Explain with example – 4.2

2. Provide steps – . **4.3.1**

3. Provide statement and proof of property 8– 4.4

4. Provide definition – 4.5 – Cramer's Rule for 3 variable

UNIT-5 VECTOR SPACES IN LINEAR TRANSFORMATION

STRUCTURE

5.0 Objective

5.1 Introduction

5.2 Vector Spaces: Definition

5.3 Subspaces

5.4 Linear Span

5.5 Fundamental Subspaces Associated with a Matrix:

5.6 Linear Independence

5.7 Summary

5.8 Keywords

5.9 Questions for review

5.10 Suggested readings

5.11 Answers to check your progress

5.0 OBJECTIVE

Understand the concept of Vector subspaces and sub spaces

Understand the Fundamental Subspaces Associated with a Matrix

Comprehend Linear Independence

5.1 INTRODUCTION

In this chapter, we will mainly be concerned with finite dimensional vector spaces over or

ℂ. Please note that the real and complex numbers have the property that any pair of elements can be added, subtracted or multiplied. Also, division is allowed by a nonzero element. Such

sets in mathematics are called field. So, ℝ and ℂ are examples of field. The fields ℝ and ℂ have infinite number of elements. But, in mathematics, we do have fields that have only finitely many elements. For example, consider the set $\mathbb{Z}^5 = \{0,1,2,3,4\}$. In Z5, we respectively, define addition and multiplication, as

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

and

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Then, we see that the elements of $\mathbb{Z}^5$ can be added, subtracted and multiplied. Note that 4 behaves as −1 and 3 behaves as −2. Thus, 1 behaves as −4 and 2 behaves as −3. Also, we see that in this multiplication $2 \cdot 3 = 1$ and $4 \cdot 4 = 1$. Hence,

1. the division by 2 is similar to multiplying by 3,
2. the division by 3 is similar to multiplying by 2, and
3. the division by 4 is similar to multiplying by 4.

Thus, $\mathbb{Z}^5$ indeed behaves like a field. So, in this chapter,𝔽 will represent a field.

5.2 VECTOR SPACES: DEFINITION

Let $A \in$ M

m,n(F) and let V denote the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$. Then, by Theorem 2.1.9, V satisfies:

1. $\mathbf{0} \in$ V as $A\mathbf{0} = \mathbf{0}$.
2. if $\mathbf{x} \in$ V then $\alpha\mathbf{x} \in$ V, for all $\alpha \in$ F. In particular, for $\alpha = -1$, $-\mathbf{x} \in$ V.
3. if $\mathbf{x}, \mathbf{y} \in$ V then, for any $\alpha, \beta \in$ F, $\alpha\mathbf{x} + \beta\mathbf{y} \in$ V.

That is, the solution set of a homogeneous linear system satisfies some nice properties. The Euclidean plane, R2, and the Euclidean space, R3, also satisfy the above properties. In this chapter, our aim is to understand sets that satisfy such properties. We start with the following definition.

Definition 5.1.1. [Vector Space] A **vector space** V over F, denoted V(F) or in short V (if the field F is clear from the context), is a non-empty set, satisfying the following conditions:

1. **Vector Addition:** To every pair **u**, **v** $\in$ V there corresponds a unique element **u** $\oplus$ **v** $\in$ V (called the **addition of vectors**) such that

(a) **u** $\oplus$ **v** = **v** $\oplus$ **u** (Commutative law).

(b) (**u** $\oplus$ **v**) $\oplus$ **w** = **u** $\oplus$ (**v** $\oplus$ **w**) (Associative law).

(c) V has a unique element, denoted **0**, called **the zero vector** that satisfies **u** $\oplus$ **0** = **u**, for every **u** $\in$ V (called **the additive identity**).

(d) For every **u** $\in$ V there is an element **w** $\in$ V that satisfies **u** $\oplus$ **w** = **0**.

2. **Scalar Multiplication:** For each **u** $\in$ V and $\alpha \in$ F, there corresponds a unique element α **u** in V (called the **scalar multiplication**) such that

(a) $\alpha \cdot (\beta$ **u**) = $(\alpha\beta)$ **u** for every $\alpha, \beta \in$ F and **u** $\in$ V (· is multiplication in F).

(b) 1 **u** = **u** for every **u** $\in$ V, where 1 $\in$ F.

3. **Distributive Laws: relating vector addition with scalar multiplication**

For any $\alpha, \beta \in$ F and **u**, **v** $\in$ V, the following **distributive laws** hold:

(a) α (**u** $\oplus$ **v**) = (α **u**) $\oplus$ (α **v**).

(b) $(\alpha + \beta)$ **u** = (α **u**) $\oplus$ (β **u**) (+ is addition in F).

Remark 5.1.2. [Real / Complex Vector Space]

1. The elements of F are called **scalars**.

2. The elements of V are called **vectors**.

3. We denote the zero element of F by 0, whereas the zero element of V will be denoted by **0**.

4. Observe that Condition 3.1.1.1d implies that for every **u** ∈ V, the vector **w** ∈ V such that **u** + **w** = **0** holds, is unique. For if, **w**1, **w**2 ∈ V

with $\mathbf{u} + \mathbf{w}i = \mathbf{0}$*, for* $i = 1, 2$ *then by commutativity of vector addition, we see that*

$\mathbf{w}_1 = \mathbf{w}_1 + \mathbf{0} = \mathbf{w}_1 + (\mathbf{u} + \mathbf{w}_2) = (\mathbf{w}_1 + \mathbf{u}) + \mathbf{w}2 = \mathbf{0} + \mathbf{w}_2 = \mathbf{w}_2$.

Hence, we represent this unique vector by $-\mathbf{u}$ and call it **the additive inverse**.

5. If V *is a vector space over* R *then,* V *is called a* **real vector space**.

6. If V *is a vector space over* C *then* V *is called a* **complex vector space**.

7. In general, a vector space over R *or* C *is called a* **linear space**.

Some interesting consequences of Definition 5.1.1 is stated next.

Theorem 5.1.3. *Let* V *be a vector space over* F*. Then,*

1. $\boldsymbol{u} \oplus \boldsymbol{v} = \boldsymbol{u}$ *implies* $\mathbf{v} = \mathbf{0}$.

2. $\alpha \odot \boldsymbol{u} = \mathbf{0}$ *if and only if either* $\mathbf{u} = \mathbf{0}$ *or* $\alpha = 0$.

3. $(-1) \odot u = -\boldsymbol{u}$*, for every* $\mathbf{u} \in$ V.

Proof. Part 1: By Condition 5.1.1.1d, for each $\mathbf{u} \in$ V there exists $-\mathbf{u} \in$ V such that $-\boldsymbol{u} \oplus \boldsymbol{u} = \mathbf{0}$. Hence, $\boldsymbol{u} \oplus \boldsymbol{v} = \boldsymbol{u}$ is equivalent to

$$-\boldsymbol{u} \oplus (\boldsymbol{u} \oplus \boldsymbol{v}) = -\boldsymbol{u} \oplus \boldsymbol{u} \Leftrightarrow (-\boldsymbol{u} \oplus \boldsymbol{u}) \oplus \boldsymbol{v} = \mathbf{0}$$
$$\Leftrightarrow \mathbf{0} \oplus \boldsymbol{v} = \mathbf{0} \Leftrightarrow \boldsymbol{v} =$$

Part 2: As $\mathbf{0} = \mathbf{0} \oplus \mathbf{0}$, using Condition 5.1.1.3, we have

$$\alpha \odot \mathbf{0} = \alpha \odot (\mathbf{0} \oplus \mathbf{0}) = (\alpha \odot \mathbf{0}) \oplus (\alpha \odot \mathbf{0}).$$

Thus, using Part 1, $\alpha \odot \mathbf{0} = \mathbf{0}$ for any $\alpha \in$ F. In the same way, using Condition 5.1.1.3b,

$$0 \odot \boldsymbol{u} = (0 + 0) \odot \boldsymbol{u} = (0 \odot \boldsymbol{u}) \oplus (0 \odot \boldsymbol{u}).$$

Hence, using Part 1, one has $0 \ \mathbf{u} = \mathbf{0}$ for any $\mathbf{u} \in$ V.

Now suppose $\alpha \odot \boldsymbol{u} = \mathbf{0}$. If $\alpha = 0$ then the proof is over. Therefore, assume that $\alpha \neq 0$, $\alpha \in$ F.

Then, $(\alpha)-1 \in$ F and

$$\mathbf{0} = (\alpha)^{-1} \odot \mathbf{0} = (\alpha)^{-1} \odot (\alpha \odot \boldsymbol{u}) = ((\alpha)^{-1} . \alpha) \odot \boldsymbol{u} = 1 \odot \boldsymbol{u}$$
$$= \boldsymbol{u}$$

as $1 \odot \boldsymbol{u} = \boldsymbol{u}$ for every vector **u** $\in$ V. Thus, if $\alpha \neq 0$ and α **u** = **0** then **u** = **0**.

Part 3: As $\mathbf{0} = 0 \cdot \boldsymbol{u} = (1 + (-1))\boldsymbol{u} = \boldsymbol{u} \oplus (-1) \cdot \boldsymbol{u}$, one has $(-1) \cdot \boldsymbol{u} = -\boldsymbol{u}$.

5.3 SUBSPACES

Definition 5.2.1. [Vector Subspace] Let V be a vector space over F. Then, a non-empty subset *S* of V is called a **subspace** of V if *S* is also a vector space with vector addition and scalar multiplication inherited from V.

Theorem 5.2.2. *Let* V(F) *be a vector space and* W $\subseteq$ V, W $\neq \emptyset$. *Then,* W *is a subspace of* V *if and only if* α**u** + β**v** $\in$ W *whenever* $\alpha, \beta \in$ F *and* **u**, **v** $\in$ W.

Proof. Let W be a subspace of V and let **u**, **v** $\in$ W. Then, for every $\alpha, \beta \in$ F, α**u**, β**v** $\in$ W and hence α**u** + β**v** $\in$ W.

Now, we assume that α**u** + β**v** $\in$ W, whenever $\alpha, \beta \in$ F and **u**, **v** $\in$ W. To show, W is a subspace of V:

1. Taking $\alpha = 1$ and $\beta = 1$, we see that **u** + **v** $\in$ W, for every **u**, **v** $\in$ W.
2. Taking $\alpha = 0$ and $\beta = 0$, we see that **0** $\in$ W.
3. Taking $\beta = 0$, we see that α**u** $\in$ W, for every $\alpha \in$ F and **u** $\in$ W. Hence, using Theorem 3.1.3.3, −**u** = (−1)**u** $\in$ Was well.
4. The commutative and associative laws of vector addition hold as they hold in V.
5. The conditions related with scalar multiplication and the distributive laws also hold as

they hold in V.

Check Your Progress

1. Define Vector addition and state its properties

2. Explain vector subspace and state relevant theorem

5.4 LINEAR SPAN

Definition 5.3.1 [Linear Combination] Let V be a vector space over F. Then, for any

$\mathbf{u}_1, \ldots, \mathbf{u}_n \in V$ and $\alpha_1, \ldots, \alpha_n \in F$, the vector $\alpha_1\mathbf{u}_1 + \cdots + \alpha_n\mathbf{u}_n = \sum_{i=1}^{n} \alpha_i u_i$ is said to be a

linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$.

Example 1. (3, 4, 3) is a linear combination of (1, 1, 1) and (1, 2, 1) as (3, 4, 3) = 2(1, 1, 1) + (1, 2, 1).

2. (3, 4, 5) is not a linear combination of (1, 1, 1) and (1, 2, 1) as the linear system

(3, 4, 5) = a(1, 1, 1) + b(1, 2, 1), in the variables a and b has no solution.

Definition 5.3.2. [Linear Span] Let V be a vector space over F and $S \subseteq V$. Then, the

linear span of S, denoted $LS(S)$, is defined as

$LS(S) = \{\alpha_1\mathbf{u}_1 + \cdots + \alpha_n\mathbf{u}_n \mid \alpha_i \in F, \mathbf{u}_i \in S, \text{ for } 1 \le i \le n\}$.

That is, $LS(S)$ is the set of all possible linear combinations of finitely many vectors of S. If S

is an empty set, we define $LS(S) = \{\mathbf{0}\}$.

Example: For the set S given below, determine $LS(S)$.

1. $S = \{(1, 0)^T, (0, 1)^T\} \subseteq \mathbb{R}^2$.

Solution: $LS(S) = \{a(1, 0)^T + b(0, 1)^T \mid a, b \in \mathbb{R}\} = \{(a, b)^T \mid a, b \in \mathbb{R}\} = \mathbb{R}^2$.

2. S = $\{(1, 1, 1)^T, (2, 1, 3)^T\}$. What does *LS*(*S*) represent in $\mathbb{R}^3$?

Solution: LS(S) = $\{a(1, 1, 1)^T + b(2, 1, 3)^T \mid a, b \in \mathbb{R}\} = \{(a + 2b, a + b, a + 3b)^T \mid a, b \in \mathbb{R}\}$.

Note that *LS*(*S*) represents a plane passing through the points $(0, 0, 0)^T$, $(1, 1, 1)^T$and $(2, 1, 3)^T$. To get the equation of the plane, we proceed as follows:

Find conditions on *x*, *y* and *z* such that $(a + 2b, a + b, a + 3b) = (x, y, z)$. Or equivalently, find conditions on *x*, *y* and *z* such that $a + 2b = x$, $a + b = y$ and $a + 3b = z$ has a solution for all $a, b \in$ R. The RREF of the augmented matrix equals

$$\begin{bmatrix} 1 & 0 & 2y - x \\ 0 & 1 & x - y \\ 0 & 0 & z + y - 2x \end{bmatrix}$$

Thus, the required condition on *x*, *y* and *z* is given by $z + y - 2x = 0$.

$$LS(S) = \{a(1,1,1)^T + b(2,1,3)^T \mid a, b \in \mathbb{R}\} = \{(x,y,z)^T \in \mathbb{R}^3 \mid 2x - y - z = 0\}.$$

Hence

3. $S = \{(1, 2, 1)^T, (1, 0, -1)^T, (1, 1, 0)^T\}$. What does *LS*(*S*) represent?

Solution: As above, *LS*(*S*) is a plane passing through the given points and $(0, 0, 0)^T$. To get the equation of the plane, we need to find condition(s) on *x*, *y*, *z* such that the linear system

$a\,(1, 2, 1) + b(1, 0, -1) + c(1, 1, 0) = (x, y, z)$ (A)

in the variables *a*, *b*, *c* is always consistent. An application of GJE to Equation (A) gives

$$\begin{bmatrix} 1 & 0 & 1 & \frac{x+y}{3} \\ 0 & 1 & \frac{1}{2} & \frac{2x-y}{3} \\ 0 & 0 & 0 & x - y + z \end{bmatrix}.$$

Thus, $LS(S) = \{(x, y, z)^T \in \mathbb{R}^3 \mid x - y + z = 0\}$.

4. $S = \{1 + 2x + 3x2, 1 + x + 2x2, 1 + 2x + x3\}$.

Solution: To understand $LS(S)$, we need to find condition(s) on $\alpha, \beta, \gamma, \delta$ such that the linear system

$$a(1 + 2x + 3x^2) + b(1 + x + 2x^2) + c(1 + 2x + x^3) = \alpha + \beta x + \gamma x^2 + \delta x^3$$

in the variables a, b, c is always consistent. An application of GJE method gives $\alpha + \beta - \gamma - 3\delta = 0$ as the required condition. Thus, $LS(S) = \{\alpha + \beta x + \gamma x^2 + \delta x^3 \in \mathbb{R}[x] \mid \alpha + \beta - \gamma - 3\delta = 0\}$.

5. $$S = \left\{ I_3, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 4 \end{bmatrix} \right\} \subseteq \mathbb{M}_3(\mathbb{R}).$$

Solution: To get the equation, we need to find conditions of *aij*'s such that the system

$$\begin{bmatrix} \alpha & \beta+\gamma & \beta+2\gamma \\ \beta+\gamma & \alpha+\beta & 2\beta+2\gamma \\ \beta+2\gamma & 2\beta+2\gamma & \alpha+2\gamma \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

in the variables α, β, γ is always consistent. Now, verify that the required condition equals

$$LS(S) = \{A = [a_{ij}] \in \mathbb{M}_3(\mathbb{R}) \mid A = A^T, a_{11} = \frac{a_{22} + a_{33} - a_{13}}{2}, a_{12} = \frac{a_{22} - a_{33} + 3a_{13}}{4}, a_{23} = \frac{a_{22} - a_{33} + 3a_{13}}{2}\}.$$

Definition 5.3.3. [Finite Dimensional Vector Space] Let V be a vector space over F. Then, V is called **finite dimensional** if there exists $S \subseteq V$, such that S has finite number of elements and V = $LS(S)$. If such an S does not exist then V is called **infinite dimensional**.

Example. 1. $\{(1, 2)^T, (2, 1)^T\}$ spans R2. Thus, $\mathbb{R}^2$ is finite dimensional.

2. $\{1, 1 + x, 1 - x + x^2, x^3, x^4, x^5\}$ spans C[x; 5]. Thus, C[x; 5] is finite dimensional.

Lemma 5.3.4 (Linear Span is a Subspace)**.** *Let* V *be a vector space over* F *and S* $\subseteq$ V. *Then,*

LS(*S*) *is a subspace of* V.

Proof. By definition, **0** ∈ LS(S). So, LS(S) is non-empty. Let **u**, **v** ∈ LS(S). To show, a**u** + b**v** ∈ LS(S) for all a, b ∈ F. As **u**, **v** ∈ LS(S), there exist n ∈ N, vectors $\mathbf{w}_i$ ∈ S and scalars α_i, β_i ∈ F such that $\mathbf{u} = \alpha_1\mathbf{w}_1 + \cdots + \alpha_n\mathbf{w}_n$ and $\mathbf{v} = \beta_1\mathbf{w}_1 + \cdots + \beta_n\mathbf{w}_n$. Hence,

$$a\mathbf{u} + b\mathbf{v} = (a\alpha_1 + b\beta_1)\mathbf{w}_1 + \cdots + (a\alpha_n + b\beta_n)\mathbf{w}_n \in LS(S)$$

as $a\alpha_i + b\beta_i \in$ F for $1 \le i \le n$. Thus, by *LS*(*S*) is a vector subspace

Theorem 5.3.5 *Let* V *be a vector space over* F *and S* $\subseteq$ V. *Then, LS*(*S*) *is the smallest*

subspace of V *containing S.*

Proof. For every **u** $\in S$, **u** = 1 · **u** $\in LS(S)$. Thus, $S \subseteq LS(S)$. Need to show that *LS*(*S*) is the

smallest subspace of V containing *S*. So, let W be any subspace of V containing *S*. Then, by

Exercise 3.1.20, *LS*(*S*) $\subseteq$ W and hence the result follows.

Definition 5.3.6. [Sum of two subsets] Let V be a vector space over F.

1. Let *S* and *T* be two subsets of V. Then, the **sum** of *S* and *T* , denoted *S* + *T* equals $\{s + t | s \in S, t \in T\}$. For example,

(a) if V = R, *S* = {0, 1, 2, 3, 4, 5, 6} and *T* = {5, 10, 15} then *S* + *T* = {5, 6, . . . , 21}.

(b) if $\mathbb{V} = \mathbb{R}^2$, $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ and $T = \left\{\begin{bmatrix}-1\\1\end{bmatrix}\right\}$ then $S + T = \left\{\begin{bmatrix}0\\2\end{bmatrix}\right\}$.

(c) if $\mathbb{V} = \mathbb{R}^2$, $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ and $T = LS\left(\begin{bmatrix}-1\\1\end{bmatrix}\right)$ then $S + T = \left\{\begin{bmatrix}1\\1\end{bmatrix} + c\begin{bmatrix}-1\\1\end{bmatrix} | c \in \mathbb{R}\right\}$.

2. Let *P* and *Q* be two subspaces of R2. Then, *P* + *Q* = R2, if

$$(x, y) = \frac{2y - x}{3}(1, 2) + \frac{2x - y}{3}(2, 1).$$

(a) $P = \{(x, 0)T \mid x \in R\}$ and $Q = \{(0, x)T \mid x \in R\}$ as $(x, y) = (x, 0) + (0, y)$.

(b) $P = \{(x, 0)T \mid x \in R\}$ and $Q = \{(x, x)T \mid x \in R\}$ as $(x, y) = (x - y, 0) + (y, y)$.

(c) $P = LS((1, 2)T)$ and $Q = LS((2, 1)T)$ as

Lemma 5.3.7. Let P and Q be two subspaces of a vector space V over F. Then, P + Q is a subspace of V. Furthermore, P + Q is the smallest subspace of V containing both P and Q.

5.5 FUNDAMENTAL SUBSPACES ASSOCIATED WITH A MATRIX:

Definition 5.4.1. [Fundamental Subspaces] Let $A \in M_{m,n}(\mathbb{C})$. Then, we define the four fundamental subspaces associated with A as

1. $\text{Col}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n\} \subseteq \mathbb{C}^m$, called the **Column space**. Observe that Col (A) is the linear span of the columns of A.
2. $\text{Row}(A) = \{\mathbf{x}^T A \mid \mathbf{x} \in \mathbb{C}^m\}$, called the **row space** of A. Observe that Row(A) is the linear span of the rows of A.
3. $\text{Null}(A) = \{\mathbf{x} \in \mathbb{C}n \mid A\mathbf{x} = \mathbf{0}\}$, called the **Null space** of A.
4. $\text{Null}(A^*) = \{\mathbf{x} \in Cm \mid A^*\mathbf{x} = \mathbf{0}\}$.

Remark 5.4.2. *Let* $A \in M_{m,n}(\mathbb{C})$.

1. Then, Col(A) *is a subspace of* C^m *and* Col(A^*) *is a subspace of* C*n.*

2. Then, Null(A) *is a subspace of* C^n *and* Null(A^*) *is a subspace of* C*m.*

5.6 LINEAR INDEPENDENCE

Definition 5.5.1. Linear Independence and Dependence

Let S = {**u**1, . . **u**m} be a

non-empty subset of a vector space 𝕍 over 𝔽. Then, S is said to be

linearly independent if

the linear system

$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m = \mathbf{0}$ (A) in the variables αi's, $1 \le i \le m$, has only the trivial solution. If Equation (A) has a non-trivial solution then S is said to be **linearly dependent**.

If S has infinitely many vectors then S is said to be **linearly independent** if for every finite subset T of S, T is linearly independent.
Observe that we are solving a linear system over $\mathbb{F}$. Hence, linear independence and dependence depend on $\mathbb{F}$, the set of scalars.

Example 1. Is the set S a linear independent set? Give reasons.

(a) Let $S = \{1 + 2x + x2, 2 + x + 4x2, 3 + 3x + 5x2\} \subseteq R[x; 2]$.

Solution: Consider the system

$$\begin{bmatrix} 1+2x+x^2 & 2+x+4x^2 & 3+3x+5x^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0,$$

or equivalently $a\,(1 + 2x + x^2) + b(2 + x + 4x^2) + c(3 + 3x + 5x^2) = 0$, in the variables a, b and c.
As two polynomials are equal if and only if their coefficients are equal, the above system reduces to the homogeneous system

$$a + 2b + 3c = 0,\ 2a + b + 3c = 0,\ a + 4b + 5c = 0.$$

The corresponding coefficient matrix has rank $2 < 3$, the number of variables. Hence, the system has a non-trivial solution. Thus, S is a linearly dependent subset of $\mathbb{R}[x; 2]$.
(b) $S = \{1, \sin(x), \cos(x)\}$ is a linearly independent subset of $C([-\pi, \pi]$, R) over R as the system

$$\begin{bmatrix} 1 & \sin(x) & \cos(x) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow a \cdot 1 + b \cdot \sin(x) + c \cdot \cos(x) = 0,$$

(B)

in the variables a, b and c has only the trivial solution. To verify this, evaluate Equation (3.3.2) at $-\pi/2$, 0 and $\pi/2$ to get the homogeneous system $a - b = 0$, $a + c = 0$, $a + b = 0$. Clearly, this system has only the trivial solution.

(c) Let $S = \{(0, 1, 1)T, (1, 1, 0)T, (1, 0, 1)T\}$.

Solution: Consider the system

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

variables a, b and c. As rank of coefficient matrix is 3 = the number of variables, the system has only the trivial solution. Hence, S is a linearly independent subset of $\mathbb{R}^3$.

(d) Consider C as a complex vector space and let $S = \{1, i\}$.

Solution: Since C is a complex vector space, $i \cdot 1 + (-1)i = i - i = 0$. So, S is a linear dependent subset of the complex vector space C.

(e) Consider C as a real vector space and let $S = \{1, i\}$.

Solution: Consider the linear system $a \cdot 1 + b \cdot i = 0$, in the variables $a, b \in \mathbb{R}$. Since $a, b \in \mathbb{R}$, equating real and imaginary parts, we get $a = b = 0$. So, S is a linear independent subset of the real vector space C.

5.5.2 Basic Results on Linear Independence:

Proposition 5.5.2.1. *Let* V *be a vector space over* F.

1. Then, **0**, *the zero-vector, cannot belong to a linearly independent set.*

2. Then, every subset of a linearly independent set in V *is also linearly independent.*

3. Then, a set containing a linearly dependent set of V *is also linearly dependent.*

Proof. Let $\mathbf{0} \in S$. Then, $1 \cdot \mathbf{0} = \mathbf{0}$. That is, a non-trivial linear combination of some vectors in S is **0**. Thus, the set S is linearly dependent.

Proposition 5.5.2.2. *Let S be a linearly independent subset of a vector space* $\mathbb{V}$ *over* $\mathbb{F}$. *If* T_1

, T_2 are two subsets of S such that $T_1 \cap T_2 = \emptyset$ then, $LS(T_1) \cap LS(T_2) = \{\mathbf{0}\}$. That is, if $\mathbf{v} \in LS(T_1) \cap LS(T_2)$ then $\mathbf{v} = \mathbf{0}$.

Proof. Let $\mathbf{v} \in LS(T1) \cap LS(T2)$. Then, there exist vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in T_1$, $\mathbf{w}_1, \ldots, \mathbf{w}_l \in T_2$ and scalars αi's and βj's (not all zero) such that $v = \sum_{i=1}^{k} \alpha_i u_i$ and $v = \sum_{j=1}^{l} \beta_j w_j$. Thus, we see that

$$\sum_{i=1}^{k} \alpha_i \mathbf{u}_i + \sum_{j=1}^{\ell} (-\beta_j) \mathbf{w}_j = \mathbf{0}.$$

As the scalars αi's and βj's are not all zero, we see that a non-trivial linear combination of some vectors in $T1 \cup T2 \subseteq S$ is $\mathbf{0}$. This contradicts the assumption that S is a linearly independent subset of V. Hence, each of α's and βj's is zero. That is $\mathbf{v} = \mathbf{0}$.

Theorem 5.5.2.3. *Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a non-empty subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. If*
$T \subseteq LS(S)$ having more than k vectors then, T is a linearly dependent subset in $\mathbb{V}$.

Proof. Let $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. As $\mathbf{w}_i \in LS(S)$, there exist $a_{ij} \in F$ such that $\mathbf{w}_i = a_{i1}\mathbf{u}_1 + \cdots + a_{ik}\mathbf{u}_k$, for $1 \le i \le m$.

$$\begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix} = \begin{bmatrix} a_{11}\mathbf{u}_1 + \cdots + a_{1k}\mathbf{u}_k \\ \vdots \\ a_{m1}\mathbf{u}_1 + \cdots + a_{mk}\mathbf{u}_k \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mk} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix}.$$

As $m > k$, the linear system $\mathbf{x}T\,A = \mathbf{0}T$ has a non-trivial solution, say $\mathbf{Y} \neq \mathbf{0}^T$. That is, $\mathbf{Y}^T A = \mathbf{0}^T$. Thus,

$$\mathbf{Y}^T \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix} = \mathbf{Y}^T \left(A \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} \right) = (\mathbf{Y}^T A) \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} = \mathbf{0}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} = \mathbf{0}^T.$$

As $\mathbf{Y} \neq \mathbf{0}$, a non-trivial linear combination of vectors in T is $\mathbf{0}$. Thus, the set T is linearly dependent subset of V.

Corollary 5.5.2.4. Fix n ∈ ℕ. Then, any subset S of R_n with | S | ≥ n+1 is linearly dependent.

Proof. Observe that R_n = LS({**e**1, . . . , **e**n}), where **e**i = In[:, i], is the i-th column of In. Hence,
using Theorem 3.3.5, the required result follows.

Theorem 5.5.2.5. Let S be a linearly independent subset of a vector space 𝕍 over 𝔽. Then, for
any **v** ∈ 𝕍 the set S ∪ {**v**} is linearly dependent if and only if **v** ∈ LS(S).

Proof. Let us assume that S ∪ {**v**} is linearly dependent. Then, there exist **v**i's in S such that
the linear system

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p + \alpha_{p+}1\mathbf{v} = \mathbf{0} \qquad \text{(C)}$$

in the variables αi's has a non-trivial solution, say αi = ci, for 1 ≤ i ≤ p + 1. We claim that $c_{p+1} \neq 0$. For, if $c_{p+1} = 0$ then, Equation (C) has a non-trivial solution corresponds to having a non-trivial solution of the linear system $\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p = \mathbf{0}$ in the variables $\alpha_1, \ldots, \alpha_p$. This contradicts Proposition 5.5.2.1 (2) as $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} \subseteq S$, a linearly independent set. Thus, $c_{p+1} \neq 0$ and we get

$$\mathbf{v} = -\frac{1}{c_{p+1}}(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) \in LS(\mathbf{v}_1, \ldots, \mathbf{v}_p)$$

as $-\frac{c_i}{c_{p+1}} \in \mathbb{F}$, for 1 ≤ i ≤ p. That is, **v** is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Now, assume that **v** ∈ LS(S). Then, there exists **v**i ∈ S and ci ∈ F, not all zero, such that
$v = \sum_{i=1}^{p} c_i v_i$Thus, the linear system $\alpha_1\mathbf{v}_1 + \cdots + \alpha_p\mathbf{v}_p + \alpha_{p+1}\mathbf{v} = \mathbf{0}$ in the variables α_i's has a
non-trivial solution $[c_1, \ldots, c_p, -1]$. Hence, S ∪ {**v**} is linearly dependent.

[We now state a very important corollary of above Theorem without proof. This result can also be used as an alternative definition of linear independence and dependence.]

Corollary 5.5.2.6. *Let* $\mathbb{V}$ *be a vector space over* $\mathbb{F}$ *and let S be a subset of* $\mathbb{V}$ *containing a non-zero vector* $\mathbf{u}_1$.

1. If S is linearly dependent then, there exists k such that LS $(\mathbf{u}_1, \ldots, \mathbf{u}_k) = LS(\mathbf{u}_1, \ldots, \mathbf{u}_{k-1})$.

2. If S linearly independent then, $\mathbf{v} \in V \setminus LS(S)$ if and only if $S \cup \{\mathbf{v}\}$ is also a linearly independent subset of V.

3. If S is linearly independent then, LS(S) = V if and only if each proper superset of S is

linearly dependent.

Check Your Progress

4. Explain **Finite Dimensional Vector Space**

5. Define Linear Independence and Dependence

5.7 SUMMARY

This unit deals with all basic concept required to build the clear understanding of vector spaces. Even we got clear understanding of Vector subspace and linear Dependence.

5.8 KEYWORDS

Proper superset – A **proper superset** of a set A is a **superset** of A that is not equal to A. In other words, if B is a **proper superset** of A, then all elements of A are in B but B contains at least one element that is not in A.

Corresponds – When two things correspond, they match up or are equivalent to one another.

Assumption -- An **assumption** is something that you assume to be the case, even without proof.

5.9 QUESTIONS FOR REVIEW

1. Does the set V given below form a real/complex or both real and complex vector space?
Give reasons for your answer.

Let V = R with $x \oplus y = x - y$ and $\alpha\ \ x = -\alpha x$, for all $x, y \in V$ and $\alpha \in R$

2. Determine all the subspaces of $\mathbb{R}$ and $\mathbb{R}^2$.

3. Let A = [B C]. Then, determine the condition under which Col(A) = Col(C).

4. Prove that a line in $\mathbb{R}^2$ is a subspace if and only if it passes through $(0, 0) \in \mathbb{R}^2$

5. Find condition(s) on $x, y, z \in \mathbb{R}$ such that (x, y, z) is a linear combination of (1, 2, 3), (−1, 1, 4) and (3, 3, 2).

5.10 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.
2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.
3. S. Lang, Linear Algebra, Springer, 1989.
4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.
5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.
6. Thomas Hungerford, Algebra, Springer GTM.
7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.
8. D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.
9.

5.11ANSWER TO CHECK YOUR PROGRESS

1. Provide definition and properties – 5.1.1

2. Provide definition, theorem and proof – 5.2.1 & 5.2.2

3. Provide definition **- 5.3.3.**

4. Provide definition of 5.5.1

UNIT-6 VECTOR SPACE IN LINEAR TRANSFORMATION II

STRUCTURE

6.0 Objective

6.1 Introduction

6.2 Application To Matrices

6.3 Basis Of Vector Space

6.4 Main Results Associated With Vector Space

6.5Application To the Subspaces Of $\mathbb{C}^n$

6.6 Ordered Bases- Change Of Basis Matrix

6.7 Summary

6.8 Keywords

6.9 Questions for review

6.10 Suggested Readings

6.11 Answers To Check Your Progress

6.0 OBJECTIVE

Understand the concept of basis and its important results associated with vector space.

Understand its application to the sub spaces

Enumerate the concept of ordered bases

6.1 INTRODUCTION

In this section, we will study results that are intrinsic to the understanding of linear algebra from the point of view of matrices, especially the fundamental subspaces associated with matrices.

6.2 APPLICATION TO MATRICES

Theorem 6.1.1. Let $A \in M_{m,n}(\mathbb{C})$. Then, the rows of A corresponding to the pivotal rows of RREF (A) are linearly independent. Also, the columns of A corresponding to the pivotal columns of RREF (A) are linearly independent.

Proof. Let RREF (A) = B. Then, the pivotal rows of Bare linearly independent due to the pivotal 1's. Now, let B_1 be the submatrix of B consisting of the pivotal rows of B. Also, let A_1 be the submatrix of A whose rows corresponds to the rows of B_1. As the RREF of a matrix is unique there exists an invertible matrix Q such that $QA_1 = B_1$. So, if there exists $\mathbf{c} \neq \mathbf{0}$ such that $\mathbf{c}^T A_1 = \mathbf{0}^T$ then $\mathbf{0}^T = \mathbf{c}^T A_1 = \mathbf{c}^T (Q^{-1} B1) = (\mathbf{c}T Q^{-1})B_1 = \mathbf{d}^T B_1$,with $\mathbf{d}^T = \mathbf{c}^T Q^{-1} \neq \mathbf{0}^T$ as Q is an invertible matrix . This contradicts the linear independence of the rows of B_1.

Let $B[:, i_1], \ldots, B[:, i_r]$ be the pivotal columns of B. Then, they are linearly independent due to pivotal 1's. As B = RREF(A), there exists an invertible matrix P such that B = PA.

Then, the corresponding columns of A satisfy

$[A[:, i_1], \ldots, A[:, i_r]] = [P^{-1} B[:, i_1], \ldots, P^{-1} B[:, i_r]] = P^{-1}[B[:, i1], \ldots, B[:, i_r]]$.

As P is invertible, the systems

$$[A[:, i_1], \ldots, A[:, i_r]] \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \mathbf{0} \text{ and } [B[:, i_1], \ldots, B[:, i_r]] \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \mathbf{0}$$

are row-equivalent. Thus, they have the same solution set. Hence, $\{A[:, i_1], \ldots, A[:, i_r]\}$ is linearly independent if and only if $\{B[:, i_1], \ldots, B[:, i_r]\}$ is linear independent. Thus, the required result follows.

The next result follows directly from Theorem 6.6.1and hence the proof is left to readers.

Corollary 6.1.2 *The following statements are equivalent for* $A \in M_n(\mathbb{C})$.

1. A is invertible.

2. The columns of A are linearly independent.

3. The rows of A are linearly independent.

Linear Independence and Uniqueness of Linear Combination

Lemma 6.1.3. *Let S be a linearly independent subset of a vector space* $\mathbb{V}$ *over* $\mathbb{F}$. *Then, each* $\mathbf{v} \in LS(S)$ *is a unique linear combination of vectors from S.*

Proof. Suppose there exists $\mathbf{v} \in LS(S)$ with $\mathbf{v} \in LS(T_1), LS(T_2)$ with $T_1, T_2 \subseteq S$. Let $T_1 = \{\mathbf{v}_1, . . . , \mathbf{v}_k\}$ and $T2 = \{\mathbf{w}_1, . . . , \mathbf{w}_l\}$, for some $\mathbf{v}i$'s and $\mathbf{w}j$'s in S. Define $T = T_1 \cup T_2$. Then, T is a subset of S. Hence the set T is linearly independent. Let $T = \{\mathbf{u}_1, . . . , \mathbf{u}_p\}$. Then, there exist α_i's and β_j's in F, not all zero, such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_p\mathbf{u}_p$ as well as $\mathbf{v} = \beta_1\mathbf{u}_1 + \cdots + \beta_p\mathbf{u}_p$.

Equating the two expressions for $\mathbf{v}$ gives

$$(\alpha_1 - \beta_1)\mathbf{u}_1 + \cdots + (\alpha_p - \beta_p)\mathbf{u}_p = \mathbf{0}. \qquad \text{(A)}$$

As T is a linearly independent subset of V, the system $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$, in the variables $c_1, . . . , c_p$, has only the trivial solution. Thus, in Equation (A), $\alpha_i - \beta_i = 0$, for $1 \le i \le p$. Thus, for $1 \le i \le p$, $\alpha_i = \beta_i$ and the required result follows.

6.3 BASIS OF A VECTOR SPACE

Definition 6.2.1. [Maximality] Let S be a subset of a set T. Then, S is said to be a **maximal subset** of T having property P if

1. S has property P and
2. no proper superset of S in T has property P .

Example: Let $T = \{2, 3, 4, 7, 8, 10, 12, 13, 14, 15\}$. Then, a maximal subset of T of consecutive integers is $S = \{2, 3, 4\}$. Other maximal subsets are $\{7, 8\}$, $\{10\}$ and $\{12, 13, 14, 15\}$. Note that $\{12, 13\}$ is not maximal.

Definition 6.2.2. [Maximal linearly independent set] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then,

Notes

S is called a **maximal linearly independent** subset of V if

1. S is linearly independent and

2. no proper superset of S in $\mathbb{V}$ is linearly independent.

Example. 1. In $\mathbb{R}^3$, the set $S = \{\mathbf{e}_1, \mathbf{e}_2\}$ is linearly independent but not maximal as $S \cup \{(1, 1, 1)^T\}$ is a linearly independent set containing S.

2. In $\mathbb{R}^3$, $S = \{(1, 0, 0)T, (1, 1, 0)T, (1, 1, -1)T\}$ is a maximal linearly independent set as S is
linearly independent and any collection of 4 or more vectors from R3 is linearly dependent.

3. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathrm{R}n$. Now, form the matrix $A = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ and let $B = \mathrm{RREF}(A)$.

Then, using Theorem 6.1.1 we see that if $B[:, i1], \ldots, B[:, ir]$ are the pivotal columns of B then $\{\mathbf{v}i_1, \ldots, \mathbf{v}i_r\}$ is a maximal linearly independent subset of S.

Theorem 6.2.3. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and S a linearly independent set in $\mathbb{V}$. Then,*
S is maximal linearly independent if and only if $LS(S) = \mathbb{V}$.

Proof. Let $\mathbf{v} \in \mathbb{V}$. As S is linearly independent, using Corollary 3.3.8.2, the set $S \cup \{\mathbf{v}\}$ is
linearly independent if and only if $\mathbf{v} \in \mathbb{V} \setminus LS(S)$. Thus, the required result follows. Let $\mathbb{V} = LS(S)$ for some set S with $|S| = k$. If $T \subseteq \mathbb{V}$ is linearly independent then $|T| \leq k$. Hence, a maximal linearly independent subset of $\mathbb{V}$ can have at most k vectors. Thus, we arrive at the following important result.

Theorem 6.2.4. *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and let S and T be two finite maximal linearly*
independent subsets of $\mathbb{V}$. Then, $|S| = |T|$.

Proof. By Theorem 6.2.3, S and T are maximal linearly independent if and only if $LS(S) = \mathbb{V} = LS(T)$. Now, use the previous paragraph to get the required result.

Let $\mathbb{V}$ be a finite dimensional vector space. Then, by Theorem 6.2.4, the number of vectors in any two maximal linearly independent set is the same.

Definition 6.2.5. [Dimension of a finite dimensional vector space] Let V be a finite dimensional vector space over F. Then, the number of vectors in any maximal linearly independent set is called the **dimension** of $\mathbb{V}$, denoted dim(V). By convention, $\dim(\{\mathbf{0}\}) = 0$.

Example. 1. As $\{1\}$ is a maximal linearly independent subset of R, dim(R) = 1.

2. As $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \subseteq \mathbb{R}^3$ is maximal linearly independent, $\dim(\mathbb{R}^3) = 3$.

3. As $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a maximal linearly independent subset in $\mathbb{R}^n$, $\dim(\mathbb{R}^n) = n$.

4. As $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a maximal linearly independent subset in $\mathbb{C}^n$ over $\mathbb{C}$, $\dim(\mathbb{C}^n) = n$.

Definition 6.2.6: Let V be a vector space over $\mathbb{F}$. Then, a maximal linearly independent subset of V is called a **basis/Hamel basis** of $\mathbb{V}$. The vectors in a basis are called **basis** vectors. By convention, a basis of $\{\mathbf{0}\}$ is the empty set.

Existence of Hamel basis

Definition 6.2.7. [Minimal Spanning Set] Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then, a subset S of $\mathbb{V}$ is called **minimal spanning** if $LS(S) = \mathbb{V}$ and no proper subset of S spans $\mathbb{V}$.

Remark 6.2.8 (Standard Basis)**:**

1. All the maximal linearly independent set given in Example 3.4.8 form the standard basis
of the respective vector space.

2. $\{1, x, x^2, \ldots\}$ is the standard basis of $\mathbb{R}[x]$ over $\mathbb{R}$.

3. Fix a positive integer n. Then, $\{1, x, x^2, \ldots, x^n\}$ is the standard basis

of ℝ [x; n] over ℝ.

4. Let V = {A ∈ M_n(ℝ) | A = AT }. Then,𝕍 is a vector space over ℝ with standard basis {Eii, Eij + Eji | 1 ≤ i < j ≤ n}.

5. Let V = {A ∈ M_n(ℝ) | AT = −A}. Then, 𝕍 is a vector space over ℝ with standard basis {$E_{ij} - E_{ji}$ | 1 ≤ i < j ≤ n}.

Example: 1. Note that *{*−2*}* is a basis and a minimal spanning subset in ℝ.

2. Let $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in \mathbb{R}^2$. Then, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ can neither be a basis nor a minimal spanning subset of $\mathbb{R}^2$.

3. $\{(1, 1, -1)^T, (1, -1, 1)^T, (-1, 1, 1)^T\}$ is a basis and a minimal spanning subset of $\mathbb{R}^3$.

4. Let V = $\{(x, y, 0)^T \mid x, y \in \mathbb{R}\} \subseteq \mathbb{R}^3$. Then, $B = \{(1, 0, 0)^T, (1, 3, 0)^T\}$ is a basis of V.

5. Let V = $\{(x, y, z)^T \in \mathbb{R}^3 \mid x + y - z = 0\} \subseteq \mathbb{R}^3$. As each element $(x, y, z)^T \in$ V satisfies $x + y - z = 0$. Or equivalently $z = x + y$, we see that $(x, y, z) = (x, y, x + y) = (x, 0, x) + (0, y, y) = x(1, 0, 1) + y(0, 1, 1)$. Hence, *{*(1, 0, 1)*T* , (0, 1, 1)*T }* forms a basis of V.

6.4 MAIN RESULTS ASSOCIATED WITH BASES

Theorem 6.3.1. Let 𝕍 be a non-zero vector space over 𝔽. Then, the following statements are equivalent.

1. B is a basis (maximal linearly independent subset) of 𝕍.
2. B is linearly independent and spans 𝕍.
3. B is a minimal spanning set in 𝕍.

Proof. 1 ⇒ 2 By definition, every basis is a maximal linearly independent subset of 𝕍. Thus, using Corollary 5.5.2.6 (2), we see that B spans 𝕍.

2 ⇒ 3 Let S be a linearly independent set that spans 𝕍. As S is linearly independent, for any **x** ∈ S, **x** ∈ / LS (S − {**x**}). Hence LS (S − {**x**}) $ LS(S) = 𝕍.

3 ⇒ 1 If B is linearly dependent then using Corollary 5.5.2.6 (1), B is not minimal spanning. A contradiction. Hence, B is linearly independent.
We now need to show that B is a maximal linearly independent set. Since LS(B) = 𝕍, for any **x** ∈ 𝕍 \ B, using Corollary 5.5.2.6 (2), the set B ∪ {**x**} is linearly dependent. That is, every proper superset of B is linearly dependent. Hence, the required result follows.
Now, using Lemma 3.3.12, we get the following result.

Remark 6.3.2. Let B be a basis of a vector space 𝕍 over 𝔽. Then, for each **v** ∈ 𝕍, there exist unique **u**i ∈ B and unique αi ∈ 𝔽, for 1 ≤ i ≤ n, such that **v** = $\sum_{i=1}^{n} \alpha_i u_i$
The next result is generally known as "every linearly independent set can be extended to form a basis of a finite dimensional vector space".

Theorem 6.3.3. Let V be a vector space over F with dim(V) = n. If S is a linearly independent subset of V then there exists a basis T of 𝕍 such that S ⊆ T.
Proof. If LS(S) = V, done. Else, choose **u**1 ∈ 𝕍 \ LS(S). Thus, by Corollary 3.3.8.2, the set
S∪{**u**1} is linearly independent. We repeat this process till we get n vectors in T as dim(𝕍) = n.
By Theorem 3.4.13, this T is indeed a required basis.

6.3.4 CONSTRUCTING A BASIS OF A FINITE DIMENSIONAL VECTOR SPACE

Step 1: Let **v**1 ∈ 𝕍 with **v**1 ≠ **0**. Then, {$\mathbf{v}_1$} is linearly independent.
Step 2: If V = LS(**v**1), we have got a basis of V. Else, pick **v**2 ∈ 𝕍 \ LS(**v**1). Then, by
Corollary 5.5.2.6 (2), {**v**1, **v**2} is linearly independent.
Step i**:** Either 𝕍 = LS(**v**1, . . . , **v**i) or LS(**v**1, . . . , **v**i) ⊊ 𝕍. In the first case, {**v**1, . . . , **v**i} is
a basis of V. Else, pick **v**i+1 ∈ 𝕍 \ LS(**v**1, . . . , **v**i). Then, by Corollary 5.5.2.6 (2), the set
{**v**1, . . . , **v**i+1} is linearly independent.
This process will finally end as V is a finite dimensional vector space.

Notes

Check your progress

1. Explain Dimension of a finite dimensional vector space

2. Explain the steps for constructing A Basis Of A Finite Dimensional Vector Space.

6.5 APPLICATION TO THE SUBSPACES OF ℂN

Example:

Let

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

Find a basis and dimension of Null(*A*).

Solution: Writing the basic vairables *x*1, *x*3 and *x*6 in terms of the free variables x_2, x_4, x_5 and x_7, we get $x_1 = x_7 - x_2 - x_4 - x_5$, $x_3 = 2x_7 - 2x_4 - 3x_5$ and $x_6 = -x_7$. Hence, the solution set has the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} x_7 - x_2 - x_4 - x_5 \\ x_2 \\ 2x_7 - 2x_4 - 3x_5 \\ x_4 \\ x_5 \\ -x_7 \\ x_7 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ 0 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}. \quad (3.5.1)$$

Now, let $u_1^T = [-1, 1, 0, 0, 0, 0, 0]$, $u_2^T = [-1, 0, -2, 1, 0, 0, 0]$, $u_3^T =$ h−1, 0, −3, 0, 1, 0, 0i and $u_4^T = [1, 0, 2, 0, 0, -1, 1]$. Then, $S = \{\mathbf{u}1, \mathbf{u}2, \mathbf{u}3, \mathbf{u}4\}$ is a basis of Null(A). The reasons for S to be a basis are as follows:

(a) By Equation (3.5.1) Null(A) = $LS(S)$.

(b) For Linear independence, the homogeneous system $c1\mathbf{u}1 + c2\mathbf{u}2 + c3\mathbf{u}3 + c4\mathbf{u}4 = \mathbf{0}$ in

the variables $c1$, $c2$, $c3$ and $c4$ has only the trivial solution as i. $\mathbf{u}_4$ is the only vector with a nonzero entry at the 7-th place ($\mathbf{u}4$ corresponds to x_7) and hence $c_4 = 0$.

ii. $\mathbf{u}3$ is the only vector with a nonzero entry at the 5-th place ($\mathbf{u}3$ corresponds to $x5$) and hence $c3 = 0$.

iii. Similar arguments hold for the variables $c2$ and $c1$.

Lemma 6.4.1. *Let* $A \in M_{m\times n}(\mathbb{C})$ *and let* E *be an elementary matrix. If*

1. $B = EA$ *then*

(a) Null(A) = Null(B), Row(A) = Row(B). *Thus, the dimensions of the corresponding spaces are equal.*

(b) Null ($\bar{A}$) = Null($\bar{B}$), Row($\bar{A}$) = Row($\bar{B}$). *Thus, the dimensions of the corresponding spaces are equal.*

2. $B = AE$ *then*

(a) Null(A*) = Null(B*), Col(A) = Col(B). *Thus, the dimensions of the corresponding spaces are equal.*

(b) Null(A^T) = Null(B^T), Col(A) = Col(B). *Thus, the dimensions of the corresponding spaces are equal.*

Proof. Part 1a: Let $\mathbf{x} \in$ Null(A). Then, $B\mathbf{x} = EA\mathbf{x} = E\mathbf{0} = \mathbf{0}$. So, Null($A$) $\subseteq$ Null(B).

Further, if $\mathbf{x} \in$ Null(B), then $A\mathbf{x} = (E^{-1} E)A\mathbf{x} = E-1(EA)\mathbf{x} = E^{-1} B\mathbf{x} = E^{-1}$ $\mathbf{0} = \mathbf{0}$. Hence, Null(B) $\subseteq$ Null(A). Thus, Null(A) = Null(B).

Let us now prove Row(A) = Row(B). So, let $\mathbf{x}T \in$ Row(A). Then, there exists $\mathbf{y} \in \mathbb{C}^m$ such that $\mathbf{x}^T = \mathbf{y}^T A$. Thus, $\mathbf{x}^T = \mathbf{y}^T E^{-1} EA = \mathbf{y}^T E^{-1} B$ and hence $\mathbf{x}G \in \text{Row}(B)$. That is, $\text{Row}(A) \subseteq \text{Row}(B)$. A similar argument gives $\text{Row}(B) \subseteq \text{Row}(A)$ and hence the required result follows.

Part 1b: E is invertible implies E is invertible and $B = EA$. Thus, an argument similar to the previous part gives us the required result.
For Part 2, note that $B^* = E^*A^*$ and E^* is invertible. Hence, an argument similar to the first part gives the required result.

Let $A \in M_{m\times n}(\mathbb{C})$ and let $B = \text{RREF}(A)$. Then, as an immediate application of Lemma 6.4.1, we get $\dim(\text{Row}(A)) = \text{Row rank}(A)$. We now prove that $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$.

Theorem 6.4.2. *Let* $A \in M{m\times n}(\mathbb{C})$. *Then,* $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$.
Proof. Let $\dim(\text{Row}(A)) = r$. Then, there exist $i1, \ldots, ir$ such that $\{A[i1, :], \ldots, A[ir, :]\}$ forms a basis of $\text{Row}(A)$
Then,

$$B = \begin{bmatrix} A[i_1, :] \\ \vdots \\ A[i_r, :] \end{bmatrix}$$

is an $r\times n$ matrix and it's rows are a basis of $\text{Row}(A)$.
Therefore, there exist $\alpha ij \in \mathbb{C}$, $1 \le i \le m$, $1 \le j \le r$ such that $A[t, :] = [\alpha t_1, \ldots, \alpha t_r]B$, for $1 \le t \le m$. So, using matrix multiplication

$$A = \begin{bmatrix} A[1,:] \\ \vdots \\ A[m,:] \end{bmatrix} = \begin{bmatrix} [\alpha_{11}, \ldots, \alpha_{1r}]B \\ \vdots \\ [\alpha_{m1}, \ldots, \alpha_{mr}]B \end{bmatrix} = CB = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mr} \end{bmatrix} B,$$

where $C = [\alpha ij]$ is an $m\times r$ matrix. Thus, using matrix multiplication, we see that each column of A is a linear combination of r columns of C. Hence, $\dim(\text{Col}(A)) \le r = \dim(\text{Row}(A))$. A similar argument gives $\dim(\text{Row}(A)) \le \dim(\text{Col}(A))$. Hence, we have the required result.

Remark 6.4.3 *The proof also shows that for every* $A \in M_{m\times n}(\mathbb{C})$ *of rank r there exists matrices B* $r\times n$ *and C*$m\times r$, *each of rank r, such that A =*

CB.

Let W1 and W1 be two subspaces of a vector space V over F. Then, recall that (see Exercise 3.1.24.4d) W1 + W2 = *{***u** + **v** | **u** $\in$ W1, **v** $\in$ W2*}* = *LS*(W1 $\cup$ W2) is the smallest subspace of V containing both W1 and W2. We now state a result similar to a result in Venn diagram that states $|A| + |B| = |A \cup B| + |A \cap B|$, whenever the sets *A* and *B* are finite

Theorem 6.4.4. *Let V be a finite dimensional vector space over* F. *If* W1 *and* W2 *are two subspaces of V then*

$$\dim(W1) + \dim(W2) = \dim(W1 + W2) + \dim(W1 \cap W2). \qquad (A)$$

For better understanding, we give an example for finite subsets of $\mathbb{R}^n$.
The example uses.
Theorem 3.3.9 to obtain bases of *LS*(*S*), for different choices *S*. The readers are advised to see Example 3.3.9 before proceeding further.

Example 6.4.5. Let V and W be two spaces with V = *{*(*v*, *w*, *x*, *y*, *z*)*T* $\in$ R5 | $v + x + z = 3y$*}*
and W = $\{(v, w, x, y, z)^T \in \mathbb{R}^5 \mid w - x = z, v = y\}$. Find bases of V and W containing a basis
of V ∩ W.
Solution: One can first find a basis of V ∩ W and then heuristically add a few vectors to get
bases for V and W, separately

$$\left[\begin{array}{ccccc} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ \hline -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccccc} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Thus, a required basis of V is $\{(1, 2, 0, 1, 2)^T, (0, 0, 1, 0, -1)^T, (0, 1, 0, 0, 0)^T, (0, 0, 0, 1, 3)^T\}$. Similarly, a required basis of W is $\{(1, 2, 0, 1, 2)^T, (0, 0, 1, 0, -1)^T, (0, 1, 0, 0, 1)^T\}$.

Theorem 6.4.6 (Rank-Nullity Theorem)**.** *Let* $A \in M_{m\times n}(\mathbb{C})$. *Then,*

$$\dim(\mathrm{Col}(A)) + \dim(\mathrm{Null}(A)) = n. \qquad \text{(B)}$$

Proof. Let $\dim(\mathrm{Null}(A)) = r \leq n$ and let $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be a basis of $\mathrm{Null}(A)$. Since B is a linearly independent set in $\mathbb{R}^n$, extend it to get $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ as a basis of $\mathbb{R}^n$. Then,

$$\mathrm{Col}(A) = LS(B) = LS(A\mathbf{u}_1, \ldots, A\mathbf{u}_n)$$

$$= LS(\mathbf{0}, \ldots, \mathbf{0}, A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n) = LS(A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n).$$

So, $C = \{A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n\}$ spans $\mathrm{Col}(A)$. We further need to show that C is linearly independent. So, consider the linear system

$$\alpha_1 A\mathbf{u}_{r+1} + \cdots + \alpha_{n-r} A\mathbf{u}_n = \mathbf{0} \Leftrightarrow A(\alpha_1\mathbf{u}_{r+1} + \cdots + \alpha_{n-r}\mathbf{u}_n) = \mathbf{0} \qquad \text{(C)}$$

in the variables $\alpha_1, \ldots, \alpha_{n-r}$. Thus, $\alpha_1\mathbf{u}_{r+1} + \cdots + \alpha_{n-r}\mathbf{u}_n \in \mathrm{Null}(A) = LS(B)$. Therefore, there exist scalars β_i, $1 \leq i \leq r$, such that

$$\sum_{i=1}^{n-r} \alpha_i \mathbf{u}_{r+i} = \sum_{j=1}^{r} \beta_j \mathbf{u}_j.$$

Or equivalently,

$$\beta_1\mathbf{u}_1 + \cdots + \beta_r\mathbf{u}_r - \alpha_1\mathbf{u}_{r+1} - \cdots - \alpha_{n-r}\mathbf{u}_n = \mathbf{0}. \qquad \text{(D)}$$

As B is a linearly independent set, the only solution of Equation (D) is $\alpha_i = 0$, for $1 \leq i \leq n - r$ and $\beta_j = 0$, for $1 \leq j \leq r$.

In other words, we have shown that the only solution of Equation (C) is the trivial solution.

Hence, $\{A\mathbf{u}_{r+1}, \ldots, A\mathbf{u}_n\}$ is a basis of $\mathrm{Col}(A)$. Thus, the required result follows. Theorem **6.4.6** is part of what is known as the fundamental theorem of linear algebra. The following are some of the consequences of the rank-nullity theorem. The proofs are left as an exercise for the reader.

6.6 ORDERED BASES

Let 𝕍 be a vector space over ℂ with dim(V) = n, for some positive integer n. Also, let W be a subspace of 𝕍 with dim(W) = k. Then, a basis of 𝕎 may not look like a standard basis. Our problem may force us to look for some other basis. In such a case, it is always helpful to fix the vectors in a particular order and then concentrate only on the coefficients of the vectors as was done for the system of linear equations where we didn't worry about the variables. It may also happen that k is very-very small as compared to n in which case it is better to work with k vectors in place of n vectors.

Definition 6.5.1. [Ordered Basis, Basis Matrix] Let 𝕎 be a vector space over 𝔽 with a basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then, an **ordered basis** for 𝕎 is a basis B together with a one-to-one correspondence between B and $\{1, 2, \ldots, m\}$. Since there is an order among the elements of B, we write $B = (\mathbf{u}1, \ldots, \mathbf{u}m)$. The vector $B = [\mathbf{u}1, \ldots, \mathbf{u}m]$ is an element of $\mathbb{W}^m$ and is generally called the **basis matrix**.

Definition 6.5.2. [Coordinate Vector] Let $B = [\mathbf{v}1, \ldots, \mathbf{v}m]$ be the basis matrix corresponding to an ordered basis B of W. Since B is a basis of W, for each $\mathbf{v} \in \mathbb{W}$, there exist βi, $1 \leq i \leq m$, such that

$$\mathbf{v} = \sum_{i=1}^{m} \beta_i \mathbf{v}_i = B \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

The vector $\begin{bmatrix} \beta_1 \\ : \\ \beta_m \end{bmatrix}$, denoted $[\mathbf{v}]B$, is called **the coordinate vector of v with respect to** B. Thus

$$\mathbf{v} = B[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}_1, \ldots, \mathbf{v}_m][\mathbf{v}]_{\mathcal{B}}, \text{ or equivalently, } \mathbf{v} = [\mathbf{v}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_m \end{bmatrix}.$$

(A)

The last expression is generally viewed as a symbolic expression.

Remark 6.5.3. [Basis representation of v]

1. Let $\mathcal{B}$ be an ordered basis of a vector space $\mathbb{V}$ over $\mathbb{F}$ of dimension n.

(a) Then,

$$[\alpha\mathbf{v} + \mathbf{w}]_B = \alpha[\mathbf{v}]_B + [\mathbf{w}]_B, \textit{ for all } \alpha \in \mathrm{F} \textit{ and } \mathbf{v}, \mathbf{w} \in \mathbb{V}.$$

(b) Further, let $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\} \subseteq \mathbb{V}$. Then, observe that S is linearly independent if and only if $\{[\mathbf{w}1]_B, \ldots, [\mathbf{w}m]_B\}$ is linearly independent in $\mathbb{F}^n$.

2. Suppose $\mathbb{V} = \mathbb{F}^n$. in Definition 6.5.2. Then, $B = [\mathbf{v}1, \ldots, \mathbf{v}n]$ is an $n \times n$ invertible matrix. Thus, using Equation (A), we have

$$B[\mathbf{v}]B = \mathbf{v} = (BB{-}1)\mathbf{v} = B\,B{-}1\mathbf{v}, \text{ for every } \mathbf{v} \in V. \qquad \text{(B)}$$

As $\mathcal{B}$ is invertible, $[\mathbf{v}]B = B{-}1\mathbf{v}$, for every $\mathbf{v} \in V$.

Definition 6.5.4 . [Change of Basis Matrix] Let V be a vector space over F with dim(V) = *n*.

Let $A = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $B = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ be basis matrices corresponding to the ordered bases *A* and *B*, respectively, of $\mathbb{V}$. Thus, using Equation (A), we have

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n] = [B[\mathbf{v}_1]_B, \ldots, B[\mathbf{v}_n]^B] = B\,[[\mathbf{v}_1]_B, \ldots, [\mathbf{v}_n]_B] = B[A]_B,$$

where $[A]_B = [[\mathbf{v}_1]_B, \ldots, [\mathbf{v}_n]_B]$. Or equivalently, verify the symbolic equality

$$\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix}.$$

The matrix $[A]_B$ is called the matrix of A **with respect to the ordered basis** $\mathcal{B}$ or the **change of basis matrix** from $\mathcal{A}$ to $\mathcal{B}$.

Theorem 6.5.5 *Let* $\mathbb{V}$ *be a vector space over* $\mathbb{F}$ *with* $\dim(\mathbb{V}) = n$. *Further,* let A = ($\mathbf{v}$1, . . . , $\mathbf{v}$n) and B = ($\mathbf{u}$1, . . . , $\mathbf{u}$n) be two ordered bases of $\mathbb{V}$

1. Then, the matrix $[\mathcal{A}]_B$ is invertible.
2. Similarly, the matrix $[\mathcal{B}]_A$ is invertible.
3. Moreover, $[\mathbf{x}]_B = [A]_B[\mathbf{x}]_A$, for all $\mathbf{x} \in \mathbb{V}$. Thus, again note that the matrix $[A]_B$ takes coordinate vector of $\mathbf{x}$ with respect to A to the coordinate vector of $\mathbf{x}$ with respect to $\mathcal{B}$.

Hence, $[A]_B$ was called the change of basis matrix from A to B.

4. Similarly, $[\mathbf{x}]_A = [B]_A[\mathbf{x}]_B$, for all $\mathbf{x} \in V$.
5. Furthermore, $([A]_B)^{-1} = [B]A$.

Proof. Part 1: Note that using Equation (3.6.3), we have

$$\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \qquad \text{(C)}$$

and hence the matrix $[A]T\,B$ or equivalently $[A]B$ is invertible, which proves Part 1. A similar argument gives Part 2.

Part 3: Using Equations (A) and (3.6.3). for any $\mathbf{x} \in V$, we have

$$[\mathbf{x}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} = \mathbf{x} = [\mathbf{x}]_{\mathcal{A}}^T \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} = [\mathbf{x}]_{\mathcal{A}}^T[\mathcal{A}]_{\mathcal{B}}^T \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix}.$$

Since the basis representation of an element is unique, we get $[\mathbf{x}]T\,B =$ $[\mathbf{x}]T\,A[A]T\,B$. Or equivalently,

$[\mathbf{x}]B = [A]B[\mathbf{x}]A$. This completes the proof of Part 3.

Remark 6.5.6 . Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ with A = ($\mathbf{v}$1, . . . , $\mathbf{v}$n) as an ordered basis.

Then, by Theorem 3.6.7, $[\mathbf{v}]$A is an element of Fn, for each $\mathbf{v} \in$ V.

Therefore,

Notes

1. if F = $\mathbb{R}$ then, the elements of V correspond to vectors in $\mathbb{R}^n$.

2. if F = $\mathbb{C}$ then, the elements of V correspond to vectors in $\mathbb{F}^n$.

Check your progress

3. What do you understand by basis matrix

4. Define Change of Basis

6.7 SUMMARY

We defined vector spaces over $\mathbb{F}$. The set $\mathbb{F}$ was either $\mathbb{R}$ or $\mathbb{C}$. To define a vector space, we start with a non-empty set $\mathbb{V}$ of vectors and $\mathbb{F}$ the set of scalars. We also needed to do the following:

We then learnt linear combination of vectors and the linear span of vectors. It was also shown that the linear span of a subset *S* of a vector space $\mathbb{V}$ is the smallest subspace of V containing *S*. Then, we learnt linear independence and dependence. We then talked about the maximal linearly independent set (coming from the homogeneous system) and the minimal spanning set (coming from the non-homogeneous system) and culminating in the notion of the basis of a finite dimensional vector space $\mathbb{V}$ over $\mathbb{F}$.

6.8 KEYWORDS

1. Representation - a **representation** is a very general relationship that expresses similarities (or equivalences) between **mathematical** objects or structures.

2. A **matrix** is a rectangular array of numbers or other **mathematical** objects for which operations such as addition and multiplication are **defined**.

3. Similar - having the same shape; having corresponding sides proportional and corresponding angles equal: related by means of a similarity transformation

4. A **variable** is a quantity that may change within the context of a **mathematical** problem or experiment.

6.9 QUESTION FOR REVIEW

1.. Prove that $S = \{1, i, x, x + x^2\}$ is a linearly independent subset of the vector space C[x; 2] over R. Whereas, it is linearly dependent subset of the vector space C[x; 2] over C.

2.Let $\mathbb{V}$ be a vector space of dimension n. Then,

(a) prove that any set consisting of n linearly independent vectors forms a basis of $\mathbb{V}$.

(b) prove that if S is a subset of $\mathbb{V}$ having n vectors with LS(S) = $\mathbb{V}$ then, S forms a basis of $\mathbb{V}$.

3. Find a basis of $\mathbb{R}^3$containing the vector $(1, 1, -2)^T$.

4.

$$Let\ A = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 2 & 2 & 2 & 4 \\ 2 & -2 & 4 & 0 & 8 \\ 4 & 2 & 5 & 6 & 10 \end{bmatrix} \ and\ B = \begin{bmatrix} 2 & 4 & 0 & 6 \\ -1 & 0 & -2 & 5 \\ -3 & -5 & 1 & -4 \\ -1 & -1 & 1 & 2 \end{bmatrix}$$

- ✓ Find RREF(A) and RREF(B).
- ✓ Find invertible matrices P_1 and P_2 such that P_1A = RREF(A) and P_2B = RREF(B).
- ✓ Find bases for Col(A), Row(A), Col(B) and Row(B).
- ✓ Find bases of Null(A), Null(A^T), Null(B) and Null(B^T).
- ✓ Find the dimensions of all the vector subspaces so obtained.

Notes

5. Let V = {(v, w, x, y, z)T ∈ R5 | w − x = z, v = y, v + x + z = 3y}.

Then, verify that

B = (1, 2, 0, 1, 2)T , (0, 0, 1, 0, −1)T can be taken as an ordered basis of V. In this case,

[(3, 6, 0, 3, 1)]B = "3 5#.

6.10 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.
2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.
3. S. Lang, Linear Algebra, Springer, 1989.
4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.
5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.
6. Thomas Hungerford, Algebra, Springer GTM.
7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.
8. D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.

6.11 ANSWER TO CHECK YOUR PROGRESS

1. Provide definition and example – 6.2.5

2. Provide steps – 6.3.4

3. Provide definition – 6.5.1

4. Provide definition – 6.5.4

UNIT-7 LINEAR TRANSFORMATION

STRUCTURE

7.0 Objective

7.1 Introduction

7.2 Definition

7.3 Range And Kernel Of A Linear Transformation

7.4 Rank-Nullity Theorem

7.5 Algebra Of Linear Transformations

7.6 Summary

7.7 Keywords

7.8 Questions

7.9 Suggested Readings

7.10 Answers To Check Your Progress

7.0 OBJECTIVE

Understand the Concept of Linear Transformation

Understand the Algebra of Linear Transformation

Understand Rank Nullity Theorem

7.1 NTRODUCTION

A **linear transformation** is a function from one vector space to another that respects the underlying (linear) structure of each vector space. A linear transformation is also known as a linear operator or map. The range of the transformation may be the same as the domain, and when that happens, the transformation is known as an endomorphism or, if invertible, an automorphism. The two vector spaces must have the same underlying field.

Linear transformations are useful because they preserve the structure of a vector space. So, many qualitative assessments of a vector space that is the domain of a linear transformation may, under certain conditions, automatically hold in the image of the linear transformation. For instance, the structure immediately gives that the kernel and image are both subspaces (not just subsets) of the range of the linear transformation.

Most linear functions can probably be seen as linear transformations in the proper setting. Transformations in the change of basis formulas are linear, and most geometric operations, including rotations, reflections, and contractions/dilations, are linear transformations. Even more powerfully, linear algebra techniques could apply to certain very non-linear functions through either approximation by linear functions or reinterpretation as linear functions in unusual vector spaces. A comprehensive, grounded understanding of linear transformations reveals many connections between areas and objects of mathematics.

7.2 DEFINITIONS AND BASIC PROPERTIES

Definition 7.1.1. [Linear Transformation, Linear Operator] Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$. A function (map) $T : \mathbb{V} \to W$ is called a **linear transformation** if for all $\alpha \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ the function T satisfies

$T(\alpha \cdot \mathbf{u}) = \alpha \; T(\mathbf{u})$ and $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) \oplus T(\mathbf{v})$,

where $+, \cdot$ are binary operations in $\mathbb{V}$ and $\oplus$, are the binary operations in W. By $L(\mathbb{V}, \mathbb{W})$, we denote the set of all linear transformations from $\mathbb{V}$ to $\mathbb{W}$. In particular, if $\mathbb{W} = \mathbb{V}$ then the linear transformation T is called a **linear operator** and the corresponding set of linear operators is denoted by $\mathcal{L}(\mathbb{V})$.

Definition 7.1.2. [Equality of Linear Transformation] Let $S, T \in \mathcal{L}$ ($\mathbb{V}$, 𝕎). Then, S and T

are said to be **equal** if $S(\mathbf{x}) = T(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{V}$.

We now give examples of linear transformations.

Example: 1. Let $\mathbb{V}$ be a vector space. Then, the maps Id, $\mathbf{0} \in \mathcal{L}(\mathbb{V})$, where

(a) $\text{Id}(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in \mathbb{V}$, is commonly called the **identity operator**.

(b) $\mathbf{0}(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v} \in \mathbb{V}$, is commonly called the **zero operator**.

2. Let $\mathbb{V}$ and W be two vector spaces over F. Then, $\mathbf{0} \in \mathcal{L}(\mathbb{V}, \mathbb{W})$, where $\mathbf{0}(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v} \in \mathbb{V}$, is commonly called the **zero transformation**.

3. The map $T(\mathbf{x}) = \mathbf{x}$, for all $\mathbf{x} \in$ R, is an element of L(R) as $T(a\mathbf{x}) = a\mathbf{x} = aT(\mathbf{x})$ and

$$T(\mathbf{x} + \mathbf{y}) = \mathbf{x} + \mathbf{y} = T(\mathbf{x}) + T(\mathbf{y}).$$

The map $T(x) = (x, 3x)^T$, for all $x \in \mathbb{R}$, is an element of $\mathcal{L}(\mathbb{R}, \mathbb{R}^2)$ as $T(\lambda x) = (\lambda x, 3\lambda x)^T = \lambda(x, 3x)^T = \lambda^T(x)$ and $T(x + y) = (x + y, 3(x + y)^T = (x, 3x)^T + (y, 3y)^T = T(x) + T(y)$.

5. Let $\mathbb{V}$, $\mathbb{W}$ and $\mathbb{Z}$ be vector spaces over F. Then, for any $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $S \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$, the map $S \circ T \in \mathcal{L}(\mathbb{V}, \mathbb{Z})$, where $(S \circ T)(\mathbf{v}) = ST(v)$, *for all* $v \in \mathbb{V}$, *is called the composition of maps. Observe that for each* v $\in \mathbb{V}$,

$(S \circ T)(\alpha v + \beta u) = ST(\alpha v + \beta u) = S(\alpha T(v) + \beta T(u))$

$= \alpha S(T(v)) + \beta S(T(u)) = \alpha(S \circ T)(v) + \beta(S \circ T)(u)$

and hence $S \circ T$, *in short* ST, *is an element of* $\mathcal{L}(\mathbb{V}, \mathbb{Z})$.

6. Fix $a \in \mathbb{R}^n$ and define $T(x) = aT\,x$, for all $x \in \mathbb{R}^n$. Then $T \in L(\mathbb{R}^n, \mathbb{R})$.

For example,

(a) if $a = (1, \ldots, 1)^T$ then $T(x) = \sum_{i=1}^{n} x_i$ for all $x \in \mathbb{R}^n$.

(b) if a = ei, for a fixed i, $1 \le i \le n$, then $Ti(x) = xi$, for all $x \in \mathbb{R}^n$.

Remark 7.1.3. Let $A \in M_n(\mathbb{C})$ and define $T_A : \mathbb{C}^n \to \mathbb{C}^n$ by $TA(\mathbf{x}) = A\mathbf{x}$, for every $\mathbf{x} \in \mathbb{C}^n$.

Then, verify that

$$T_A^k(\mathbf{x}) = \underbrace{(T_A \circ T_A \circ \cdots \circ T_A)}_{k\ times}(\mathbf{x}) = A^k\mathbf{x},$$

for any positive integer k.

Also, for any two linear transformations $S \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $T \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$,we will interchangeably use $T \circ S$ and TS, for the corresponding linear transformation in $\mathcal{L}(\mathbb{V}, \mathbb{Z})$, We now prove that any linear transformation sends the zero vector to a zero vector.

Proposition 7.1.4. Let $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ Suppose that $\mathbf{0}_{\mathbb{V}}$ is the zero vector in $\mathbb{V}$ and $\mathbf{0}_{\mathbb{W}}$.is the zero vector of W. Then $T(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$.

Proof. Since $\mathbf{0}_{\mathbb{V}} = \mathbf{0}_{\mathbb{V}} + \mathbf{0}_{\mathbb{V}}$, we get

$$T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}} + \mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) + T(\mathbf{0}_{\mathbb{V}}).$$

As $T(\mathbf{0}_{\mathbb{V}}) \in \mathbb{W}$,

$$\mathbf{0}_{\mathbb{W}}. + T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) = T(\mathbf{0}_{\mathbb{V}}) + T(\mathbf{0}_{\mathbb{V}}).$$

Hence, $T(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$.

From now on **0** will be used as the zero vector of the domain and co-domain.

Example: Does there exist a linear transformation $T : \mathbb{R} \to \mathbb{R}$ such that $T(x) = x^2$, for all $x \in \mathbb{R}$?

Solution: No, as $T(ax) = (ax)^2 = a^2x^2 = a^2T(x) \neq aT(x)$, unless $a = 0, 1$.

Example: Does there exist a linear transformation $T : \mathbb{R} \to \mathbb{R}$ such that $T(x) = \sqrt{x}$, for all $x \in$ R?

Solution: No, as $T(ax) = \sqrt{ax} = \sqrt{a}\,\sqrt{x} \neq a\sqrt{x} = aT(x)$, unless $a = 0, 1$.

The next result states that a linear transformation is known if we know its image on a basis of the domain space.

Lemma 7.1.5 *Let* 𝕍 *and* 𝕎 *be two vector spaces over* F *and let* $T \in L$(V, W). *Then T is known, if the image of T on basis vectors of* 𝕍 *are known. In particular, if* 𝕍 *is finite dimensional and* $B = (\mathbf{v}_1, . . . , \mathbf{v}_n)$ *is an ordered basis of* 𝕍 *over* F *then,* $T(\mathbf{v})$ = h$T(\mathbf{v}_1) \cdots T(\mathbf{v}_n)$i[**v**]*B.*

Proof. Let B be a basis of 𝕍 over F. Then, for each **v** ∈ 𝕍, there exist vectors $\mathbf{u}_1, . . . , \mathbf{u}_k$ in *B* and scalars $c_1, . . . , c_k \in$ F such that **v** = $\sum_{i=1}^{k} c_i u_i$.Thus, by definition $T(\mathbf{v}) = \sum_{i=1}^{k} c_i T(u_i)$ Or equivalently, whenever

(A)

$$\mathbf{v} = [\mathbf{u}_1, \ldots, \mathbf{u}_k] \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \text{ then, } T(\mathbf{v}) = \begin{bmatrix} T(\mathbf{u}_1) & \cdots & T(\mathbf{u}_k) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}.$$

Corollary 7.1.6. *Let* 𝕍 *and* 𝕎 *be vector spaces over* F *and let* $T : 𝕍 \to 𝕎$ *be a linear transformation. If B is a basis of* 𝕍 *then,* RNG(*T*) = *LS*(*T* (**x**)|**x** ∈ *B*).

Corollary 7.1.7. [Reisz Representation Theorem] *Let* $T \in L(\mathbb{R}^n, \mathbb{R})$. *Then, there exists* $\mathbf{a} \in \mathbb{R}^n$ *such that* $T(\mathbf{x}) = \mathbf{a}T\,\mathbf{x}$.

Proof. By Lemma 7.1.5, *T* is known if we know the image of *T* on {**e**1, . . . , **e***n*}, the standard basis of $\mathbb{R}^n$. As *T* is given, for $1 \le i \le n$, $T(\mathbf{e}_i) = a_i$, for some $a_i \in \mathbb{R}$. So, consider the vector **a** = [*a*1, . . . , *an*]*T* . Then, for **x** = [*x*1, . . . , *xn*]$T \in \mathbb{R}^n$, we see that

$$T(\mathbf{x}) = T\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = \sum_{i=1}^{n} x_i T(\mathbf{e}_i) = \sum_{i=1}^{n} x_i a_i = \mathbf{a}^T \mathbf{x}.$$

Thus, the required result follows

7.3 RANGE AND KERNEL OF A LINEAR TRANSFORMATION

Notes

Definition 7.2.1. [Range and Kernel of a Linear Transformation] Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ and let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then,

1. the set $\{T(\mathbf{v}) | \mathbf{v} \in \mathbb{V}\}$ is called the **range space** of T, denoted Rng(T).
2. the set $\{\mathbf{v} \in \mathbb{V} | T(\mathbf{v}) = \mathbf{0}\}$ is called the **kernel** of T, denoted Ker(T). In certain books, it is also called the **null space** of T.

Example. Determine Rng(T) and Ker(T) of the following linear transformations.. $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^4)$, where $T((x, y, z)^T) = (x - y + z, y - z, x, 2x - 5y + 5z)^T$.

Solution: Consider the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$. Then

$$\begin{aligned} \text{Rng}(T) &= LS(T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)) = LS((1,0,1,2)^T, (-1,1,0,-5)^T, (1,-1,0,5)^T) \\ &= LS((1,0,1,2)^T, (1,-1,0,5)^T) = \{\lambda(1,0,1,2)^T + \beta(1,-1,0,5)^T \mid \lambda, \beta \in \mathbb{R}\} \\ &= \{(\lambda+\beta, -\beta, \lambda, 2\lambda + 5\beta) : \lambda, \beta \in \mathbb{R}\} \\ &= \{(x,y,z,w)^T \in \mathbb{R}^4 \mid x + y - z = 0, 5y - 2z + w = 0\} \end{aligned}$$

and

$$\begin{aligned} \text{Ker}(T) &= \{(x,y,z)^T \in \mathbb{R}^3 : T((x,y,z)^T) = \mathbf{0}\} \\ &= \{(x,y,z)^T \in \mathbb{R}^3 : (x - y + z, y - z, x, 2x - 5y + 5z)^T = \mathbf{0}\} \\ &= \{(x,y,z)^T \in \mathbb{R}^3 : x - y + z = 0, y - z = 0, x = 0, 2x - 5y + 5z = 0\} \\ &= \{(x,y,z)^T \in \mathbb{R}^3 : y - z = 0, x = 0\} \\ &= \{(0,z,z)^T \in \mathbb{R}^3 : z \in \mathbb{R}\} = LS((0,1,1)^T) \end{aligned}$$

Example. In each of the examples given below, state whether a linear transformation exists or not. If yes, give at least one linear transformation. If not, then give the condition due to which a linear transformation doesn't exist.

1. $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T((1, 1)^T) = (1, 2)^T$ and $T((1, -1)^T) = (5, 10)^T$?

Solution: Yes, as the set $\{(1, 1), (1, -1)\}$ is a basis of $\mathbb{R}^2$, the matrix

$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is invertible.

Also, $T\left(\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}\right) = T\left(a\begin{bmatrix}1\\ 1\end{bmatrix} + b\begin{bmatrix}1\\ -1\end{bmatrix}\right) = a\begin{bmatrix}1\\ 2\end{bmatrix} + b\begin{bmatrix}5\\ 10\end{bmatrix} = \begin{bmatrix}1 & 5\\ 2 & 10\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}$. So,

$$T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right) = T\left(\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\left(\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}^{-1}\begin{bmatrix}x\\ y\end{bmatrix}\right)\right) = \begin{bmatrix}1 & 5\\ 2 & 10\end{bmatrix}\left(\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}^{-1}\begin{bmatrix}x\\ y\end{bmatrix}\right)$$

$$= \begin{bmatrix}1 & 5\\ 2 & 10\end{bmatrix}\begin{bmatrix}\frac{x+y}{2}\\ \frac{x-y}{2}\end{bmatrix} = \begin{bmatrix}3x-2y\\ 6x-4y\end{bmatrix}.$$

Check your progress

1. Explain Linear Transformation and Linear Operator.

---

2. Define the concept of Range space and kernel of linear transformation

---

7.4 RANK-NULLITY THEOREM

Theorem 7.3.1. *Let* $\mathbb{V}$ *and* $\mathbb{W}$ *be two vector spaces over* $\mathbb{F}$ *and let* $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$.

1. If $S \subseteq \mathbb{V}$ *is linearly dependent then* $T(S) = \{T(\mathbf{v}) \mid \mathbf{v} \in \mathbb{V}\}$ *is linearly dependent.*

2. Suppose $S \subseteq \mathbb{V}$ *such that* $T(S)$ *is linearly independent then* S *is linearly independent.*

Proof. As S is linearly dependent, there exist $k \in \mathbb{N}$ and $\mathbf{v}i \in S$, for $1 \le i \le k$, such that the system $\sum_{i=1}^{k} x_i v_i = 0$, in the variable xi's, has a non-trivial solution, say $xi = ai \in \mathbb{F}$, $1 \le i \le k$.

Thus, $\sum_{i=1}^{k} a_i v_i = 0$. Now, consider the system $\sum_{i=1}^{k} y_i T(v_i) = 0$,in the variable yi's. Then,

$$\sum_{i=1}^{k} a_i T(\mathbf{v}_i) = \sum_{i=1}^{k} T(a_i \mathbf{v}_i) = T\left(\sum_{i=1}^{k} a_i \mathbf{v}_i\right) = T(\mathbf{0}) = \mathbf{0}.$$

Thus, *ai*'s give a non-trivial solution of $\sum_{i=1}^{k} y_i T(v_i) = 0$ and hence the required result follows.

Definition 7.3.2. [Rank and Nullity] Let V and W be two vector spaces over F. If $T \in L$(V, W) and dim(V) is finite then we define

Rank(T) = dim(RNG (T)) and Nullity(T) = dim(KER (T)).

Theorem 7.3.3 (Rank-Nullity Theorem)**.** *Let* V *and* W *be two vector spaces over* F. *If* dim(V) *is finite and* $T \in L$(V, W) *then,*

Rank(T) + Nullity(T) = dim(RNG (T)) + dim(KER (T)) = dim($\mathbb{V}$).

Proof. As dim(Ker(T)) ≤ dim($\mathbb{V}$). Let B be a basis of Ker(T). We extend it to form a basis C of V. As, T (**v**) = **0**, for all **v** $\in B$, using Corollary 7.1.6, we get

RNG(T) = LS({T (**v**)|**v** $\in C$}) = LS({T (**v**)|**v** $\in C \setminus B$}).

We claim that {T (**v**)|**v** $\in C \setminus B$$ is linearly independent subset of W.

Let, if possible, the claim be false. Then, there exists **v**1, . . . , **v***k* $\in C\setminus B$ and **a** = [*a*1, . . . , *ak*]*T*

such that **a** ≠ and $\sum_{i=1}^{k} a_i T(v_i) = 0$. Thus, we see that

$$T\left(\sum_{i=1}^{k} a_i \mathbf{v}_i\right) = \sum_{i=1}^{k} a_i T(\mathbf{v}_i) = \mathbf{0}.$$

That is, $\sum_{i=1}^{k} a_i v_i \in$ KER(T). Hence, there exists *b*1, . . . , *bl* ∈ $\mathbb{F}$ and **u**1, . . . , **u***l* $\in \mathcal{B}$ such that Or equivalently, the system

$$\sum_{i=1}^{k} a_i \mathbf{v}_i = \sum_{j=1}^{k} b_j \mathbf{u}_j.$$

,

$$\sum_{i=1}^{k} x_i \mathbf{v}_i + \sum_{j=1}^{k} y_j \mathbf{u}_j = \mathbf{0},$$

in the variables *xi*'s and *yj*'s, has a non-trivial solution [a_1, . . . , a_k, −*b*1, . . . , −*bl*]*T* (non-trivial as **a** ≠ **0**). Hence, *S* = {**v**1, . . . , **v***k*, **u**1, . . . , **u**`} is linearly dependent subset in V. A contradiction to *S* ⊆ *C*. Thatis, dim (RNG (*T*)) + dim(KER (*T*)) = |*C* \ *B*| + |*B*| = |*C*| = dim(𝕍).

Thus, we have proved the required result.

Corollary 7.3.4. Let 𝕍 and 𝕎 be vector spaces over 𝔽 and let T ∈ ℒ(𝕍, 𝕎). If dim(𝕍) = dim(𝕎) then, the following statements are equivalent.

1. T is one-one.

2. KER (*T*) = {**0**}.

3. T is onto.

4. dim(RNG (*T*)) = dim(𝕍).

Corollary 7.3.5. *Let* 𝕍 *be a vector space over* F *with* dim(𝕍) = *n. If S, T* ∈ *L*(𝕍). *Then*

1. Nullity(*T*) + Nullity(*S*) ≥ Nullity(*ST*) ≥ max{Nullity(*T*), Nullity(*S*)}.

2. min{Rank(*S*), Rank(*T*)} ≥ Rank(*ST*) ≥ *n* − Rank(*S*) − Rank(*T*).

Proof. The prove of Part 2 is omitted as it directly follows from Part 1 and Theorem 7.3.3. Part 1: We first prove the second inequality. Suppose **v** ∈ KER (T). Then

(*ST*)(**v**) = *S*(*T* (**v**)) = *S*(**0**) = **0**

implies KER (*T*) ⊆ KER (*ST*). Thus, NULLITY (*T*) ≤ NULLITY (*ST*).

By Theorem 7.3.3, NULLITY(*S*) ≤ NULLITY (*ST*) is equivalent to RNG (*ST*) ⊆ RNG (*S*). And

this holds as RNG (*T*) ⊆ V implies RNG (*ST*) = *S*(Rng(*T*)) ⊆ *S*(V) = Rng(*S*).

To prove the first inequality, let {**v**1, . . . , **v***k*} be a basis of KER (*T*).

Then {**v**1, . . . , **v***k*} ⊆

Notes

KER (ST). So, let us extend it to get a basis {**v**1, . . . , **v**k, **u**1, . . . , **u**l} of KER (ST).

Claim: {T (**u**1), T (**u**2), . . . , T (**u**`)} is a linearly independent subset of KER (S).

Clearly, {T (**u**1), . . . , T (**u**`)} $\subseteq$ KER (S). Now, consider the system $c1T$ (**u**1)+· · ·+$c_l T$ ($\mathbf{u}_l$) = **0**

in the variables $c1, . . . , c_l$. As $T \in \mathcal{L}(\mathbb{V})$, we get ($\sum_{i=1}^{l} c_i u_i$)= **0**. Thus, $\sum_{i=1}^{l} c_i u_i \in$ KER (T).

Hence, $\sum_{i=1}^{l} c_i u_i$ is a unique linear combination of **v**1, . . . , **v**k, a basis of KER (T). Therefore,

$$c_1\mathbf{u}_1 + \cdots + c_l\mathbf{u}_l = \alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k \qquad (1)$$

for some scalars $\alpha_1, . . . , \alpha_k$. But by assumption, {$\mathbf{v}_1, . . . , \mathbf{v}_k, \mathbf{u}_1, . . . , \mathbf{u}_l$} is a basis of KER ($ST$)

and hence linearly independent. Therefore, the only solution of Equation (1) is given by c_i = 0, for 1 ≤ i ≤ ` and αj = 0, for 1 ≤ j ≤ k. Thus, we have proved the claim. Hence,

NULLITY (S) ≥ l and NULLITY ST) = $k + l$ ≤ NULLITY (T) + NULLITY (S).

7.5 ALGEBRA OF LINEAR TRANSFORMATIONS

Definition 7.4.1 . [Sum and Scalar Multiplication of Linear Transformations]: Let $\mathbb{V}$, $\mathbb{W}$ be

vector spaces over F and let $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Then, we define the point-wise

1. **sum** of S and T , denoted $S + T$, by $(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$, for all **v** $\in \mathbb{V}$.

2. **scalar multiplication**, denoted cT for $c \in \mathbb{F}$, by $(cT)(\mathbf{v}) = c\,(T(\mathbf{v}))$, for all $\mathbf{v} \in \mathbb{V}$.

Theorem 7.4.2. *Let* $\mathbb{V}$ *and* $\mathbb{W}$ *be vector spaces over* F. *Then* $\mathcal{L}(\mathbb{V}, \mathbb{W})$. *is a vector space over*

F. *Furthermore, if* dim $\mathbb{V}$ = n *and* dim $\mathbb{W}$ = m, *then* dim $\mathcal{L}(\mathbb{V}, \mathbb{W})$. = mn.

Proof. It can be easily verified that for $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$., if we define ($S + \alpha T$)(**v**) = S(**v**)+αT (**v**)

(point-wise addition and scalar multiplication) then $\mathcal{L}(\mathbb{V}, \mathbb{W})$. is indeed a vector space over F.

We now prove the other part. So, let us assume that $B = \{\mathbf{v}1, \ldots, \mathbf{v}n\}$ and $C = \{\mathbf{w}1, \ldots, \mathbf{w}m\}$ are bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. For $1 \leq i \leq n$, $1 \leq j \leq m$, we define the functions **f***ij* on the basis vectors of $\mathbb{V}$ by

$$\mathbf{f}_{ij}(\mathbf{v}_k) = \begin{cases} \mathbf{w}_j, & \text{if } k = i \\ \mathbf{0}, & \text{if } k \neq i. \end{cases}$$

For other vectors of V, we extend the definition by linearity. That is, if $v = \sum_{s=1}^{n} \alpha_s v_s$ then,

$$\mathbf{f}_{ij}(\mathbf{v}) = \mathbf{f}_{ij}\left(\sum_{s=1}^{n} \alpha_s \mathbf{v}_s\right) = \sum_{s=1}^{n} \alpha_s \mathbf{f}_{ij}(\mathbf{v}_s) = \alpha_i \mathbf{f}_{ij}(\mathbf{v}_i) = \alpha_i \mathbf{w}_j. \quad (2)$$

Thus, **f***ij* $\in \mathcal{L}(\mathbb{V}, \mathbb{W})$.

Claim: $\{\mathbf{f}ij | 1 \leq i \leq n,\ 1 \leq j \leq m\}$ is a basis of $\mathcal{L}(\mathbb{V}, \mathbb{W})$.

So, consider the linear system

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mathbf{f}_{ij} = \mathbf{0},$$

in the variables *cij*'s, for $1 \leq i \leq n$, $1 \leq j \leq m$. Using the point-wise addition and scalar multiplication, we get

$$\mathbf{0} = \mathbf{0}(\mathbf{v}_k) = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mathbf{f}_{ij}\right)(\mathbf{v}_k) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \mathbf{f}_{ij}(\mathbf{v}_k) = \sum_{j=1}^{m} c_{kj} \mathbf{w}_j.$$

But, the set $\{\mathbf{w}1, \ldots, \mathbf{w}m\}$ is linearly independent and hence the only solution equals $ckj = 0$, for $1 \leq j \leq m$. Now, as we vary **v***k* from **v**1 to **v***n*, we see that $cij = 0$, for $1 \leq j \leq m$ and $1 \leq i \leq n$. Thus, we have proved the linear independence.

Now, let us prove that LS ({**f**ij|$1 \le i \le n,\ 1 \le j \le m$}) = $\mathcal{L}(\mathbb{V}, \mathbb{W})$.. So, let $f \in \mathcal{L}(\mathbb{V}, \mathbb{W})$.

Then, for $1 \le s \le n$, f(**v**s) $\in \mathbb{W}$ and hence there exists βst's such that f(**v**s) $f(v_s) = \sum_{t=1}^{m} \beta_{st} v_{wt}$. So, if $v = \sum_{s=1}^{n} \alpha_s v_s \in \mathbb{V}$ then,

$$\begin{aligned} f(\mathbf{v}) &= f\left(\sum_{s=1}^{n} \alpha_s \mathbf{v}_s\right) = \sum_{s=1}^{n} \alpha_s f(\mathbf{v}_s) = \sum_{s=1}^{n} \alpha_s \left(\sum_{t=1}^{m} \beta_{st} \mathbf{w}_t\right) = \sum_{s=1}^{n} \sum_{t=1}^{m} \beta_{st}(\alpha_s \mathbf{w}_t) \\ &= \sum_{s=1}^{n} \sum_{t=1}^{m} \beta_{st} \mathbf{f}_{st}(\mathbf{v}) = \left(\sum_{s=1}^{n} \sum_{t=1}^{m} \beta_{st} \mathbf{f}_{st}\right)(\mathbf{v}). \end{aligned}$$

Since the above is true for every **v** $\in \mathbb{V}$, LS ({**f**ij|$1 \le i \le n,\ 1 \le j \le m$}) = $\mathcal{L}(\mathbb{V}, \mathbb{W})$. and thus the required result follows.

Definition 7.4.3 [Inverse of a Function] Let $f : S \to T$ be any function.

1. Then, a function $g : T \to S$ is called a **left inverse** off if $(g \circ f)(x) = x$, for all $x \in S$.

That is, $g \circ f$ = Id, the identity function on S.

2. Then, a function $h : T \to S$ is called a **right inverse** of f if $(f \circ h)(y) = y$, for all $y \in T$.

That is, $f \circ h$ = Id, the identity function on T .

3. Then f is said to be **invertible** if it has a right inverse and a left inverse.

Remark 7.4.4. Let f : S → T be invertible. Then, it can be easily shown that any right inverse and any left inverse are the same. Thus, the inverse function is unique and is denoted by f^{-1}. It is well known that f is invertible if and only if f is both one-one and onto.

Lemma 7.4.5. Let V and W be vector spaces over F and let T ∈ $\mathcal{L}(\mathbb{V}, \mathbb{W})$. If T is one-one and onto then, the map $T^{-1} : W \to V$ is also a linear transformation. The map T^{-1} is called the **inverse linear transform** of T and is defined by T^{-1}(**w**) = **v**, whenever T (**v**) = **w**.

Proof. Part 1: As T is one-one and onto, by Theorem 4.2.3, dim($\mathbb{V}$) = dim($\mathbb{W}$). So, by Corollary 7.3.4, for each **w** ∈ $\mathbb{W}$ there exists a unique **v** ∈ $\mathbb{V}$ such that T (**v**) = **w**. Thus, one defines T^{-1}(**w**) = **v**. We need to show that $T^{-1}(\alpha 1\mathbf{w}1 + \alpha 2\mathbf{w}2) = \alpha 1T^{-1}(\mathbf{w}1) + \alpha 2T^{-1}(\mathbf{w}2)$, for all $\alpha 1, \alpha 2 \in \mathbb{F}$ and $\mathbf{w}1,\mathbf{w}2 \in \mathbb{W}$. Note that by previous paragraph, there exist unique vectors $\mathbf{v}1,\mathbf{v}2 \in \mathbb{V}$ such that $T^{-1}(\mathbf{w}1) = \mathbf{v}1$ and $T^{-1}(\mathbf{w}2) = \mathbf{v}2$. Or equivalently, $T(\mathbf{v}1) = \mathbf{w}1$ and $T(\mathbf{v}2) = \mathbf{w}2$. So, $T(\alpha 1\mathbf{v}1 + \alpha 2\mathbf{v}2) = \alpha 1\mathbf{w}1 + \alpha 2\mathbf{w}2$, for all $\alpha 1, \alpha 2 \in \mathbb{F}$. Hence, for all $\alpha 1, \alpha 2 \in$ F, we get

$$T^{-1}(\alpha 1\mathbf{w}1 + \alpha 2\mathbf{w}2) = \alpha 1\mathbf{v}1 + \alpha 2\mathbf{v}2 = \alpha 1T^{-1}(\mathbf{w}1) + \alpha 2T^{-1}(\mathbf{w}2).$$

Thus, the required result follows.

Definition 7.4.6. [Singular, Non-singular Transformations] Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over F and let $T \in \mathcal{L}(\mathbb{V},\mathbb{W})$. Then, T is said to be **singular** if **0** $ KER (T). That is, KER (T) contains a non-zero vector. If KER (T) = $\{\mathbf{0}\}$ then, T is called **non-singular**.

Theorem 7.4.7. *Let* V *and* W *be vector spaces over* F *and let* $T \in \mathcal{L}(\mathbb{V},\mathbb{W})$. *Then the following statements are equivalent.*

1. T is one-one.

2. T is non-singular.

3. Whenever $S \subseteq \mathbb{V}$ *is linearly independent then* $T(S)$ *is necessarily linearly independent.*

Proof. 1⇒2 Let T be singular. Then, there exists $\mathbf{v} \neq \mathbf{0}$ such that $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$. This implies that T is not one-one, a contradiction.

2⇒3 Let $S \subseteq \mathbb{V}$ be linearly independent. Let if possible $T(S)$ be linearly dependent. Then, there exists $\mathbf{v}1, \ldots, \mathbf{v}k \in S$ and $\alpha = (\alpha 1, \ldots, \alpha k)T \neq \mathbf{0}$ such that $\sum_{i=1}^{k} \alpha_i T v_i = 0$. Thus, $T(\sum_{i=1}^{k} \alpha_i v_i) = \mathbf{0}$. But T is nonsingular and hence we get $\sum_{i=1}^{k} \alpha_i v_i = \mathbf{0}$ with $\alpha \neq \mathbf{0}$, a contradiction to S being a linearly independent set.

3⇒1 Suppose that T is not one-one. Then, there exists $\mathbf{x}, \mathbf{y} \in \mathbb{V}$such that $\mathbf{x} \neq \mathbf{y}$ but $T(\mathbf{x}) = T(\mathbf{y})$. Thus, we have obtained $S = \{\mathbf{x} - \mathbf{y}\}$, a linearly independent subset of $\mathbb{V}$ with $T(S) = \{\mathbf{0}\}$, a linearly dependent set. A contradiction to our assumption. Thus, the required result follows.

Definition 7.4.8. [Isomorphism of Vector Spaces] Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathbb{V},\mathbb{W})$Then, T is said to be an **isomorphism** if T is one-one and onto. The vector spaces $\mathbb{V}$ and W are said to be **isomorphic**, denoted $\mathbb{V} \cong \mathbb{W}$, if there is an isomorphism from $\mathbb{V}$ to $\mathbb{W}$.

Theorem 7.4.9. *Let* V *be an n-dimensional vector space over* F. *Then* $\mathbb{V} \cong \mathbb{F}^n$

Proof. Let $\{\mathbf{v}1, \ldots, \mathbf{v}n\}$ be a basis of V and $\{\mathbf{e}1, \ldots, \mathbf{e}n\}$, the standard basis of $\mathbb{F}^n$. Now define

$T(\mathbf{v}i) = \mathbf{e}i$, for $1 \le i \le n$ and $T\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i e_i$ for $\alpha 1, \ldots, \alpha n \in \mathbb{F}$. Then, it is easy to

observe that $T \in \mathcal{L}(\mathbb{V}, \mathbb{F}^n) T$ is one-one and onto. Hence, T is an isomorphism.

Corollary 7.4.10. *The vector space* $\mathbb{R}$ *over* $\mathbb{Q}$ *is not finite dimensional. Similarly, the vector*

space $\mathbb{C}$ *over* $\mathbb{Q}$ *is not finite dimensional.*

Theorem 7.4.11 Let V be a vector space over F with dim V = n. Then the following statements are equivalent for $T \in \mathcal{L}(\mathbb{V})$.

1. T is one-one.

2. KER (T) = {**0**}.

3. Rank(T) = n.

4. T is onto.

5. T is an isomorphism.

6. If {**v**1, . . . , **v**n} is a basis for V then so is {T (**v**1), . . . , T (**v**n)}.

7. T is non-singular.

8. T is invertible.

Check your progress

3. Explain inverse linear transform with lemma and proof.

---

4. Define

a. Isomorphic Vector Space

b. Singular and Non-Singular Transformation

7.6 SUMMARY

We have shown that matrices give rise to functions between two finite dimensional vector spaces. To do so, we start with the definition of functions over vector spaces that commute with the operations of vector addition and scalar multiplication.

7.7 KEYWORDS

5. Invertible -- If y = f (x), then the inverse relation is written as $y = f^{-1}(x)$. If the inverse is also a function, then we say that the function f is **invertible**
6. Non-singular – whose solution is not zero
7. Dimensional - **Dimensions in mathematics** are the measure of the size or distance of an object or region or space in one direction. In simpler terms, it is the measurement of the length, width, and height of anything.
8. **Equivalent** - means equal in value, function, or **meaning**

7.8 QUESTION FOR REVIEW

1. *Let* V *and* W *be two vector spaces over* F *and let* $T \in L(V, W)$. *If* dim(V) *is finite then prove that*

1. T cannot be onto if dim(V) < dim(W).

2. T cannot be one-one if dim(V) > *dim (W)*

2. Define $T \in L(R3, R2)$ *by*

$$\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x - y + z \\ x + 2z \end{bmatrix}.$$

Find a basis and the dimension of RNG(T) and KER(T).

3. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ *be defined by* $T((x, y, z)^T) = (2x - 2y + 2z, -2x + 5y + 2z, 8x + y + 4z)^T$. *Find* $\mathbf{x} \in \mathbb{R}^3$ *such that* $T(\mathbf{x}) = (1, 1, -1)^T$.

4. Let n $\in$ N. *Does there exist a linear transformation* $T : \mathbb{R}^3 \to \mathbb{R}^n$ *such that* $T((1, 1, -2)^T) =$

$\mathbf{x}$, $T((-1, 2, 3)^T) = \mathbf{y}$ *and* $T((1, 10, 1)T) = \mathbf{z}$

(a) with $\mathbf{z} = \mathbf{x} + \mathbf{y}$?

(b) with $\mathbf{z} = c\mathbf{x} + d\mathbf{y}$, *for some* $c, d \in \mathbb{R}$?

7.9 SUGGESTED READINGS

1. K. Hauffman and R. Kunz, Linear Algebra, Pearson Education (INDIA), 2003.
2. G. Strang, Linear Algebra And Its Applications, 4th Edition, Brooks/Cole, 2006.
3. S. Lang, Linear Algebra, Springer, 1989.
4. David S. Dummit and Richard M. Foote, Abstract Algebra (3e), John Wiley and Sons.
5. R. Gallian Joseph, Contemporary Abstract Algebra, Narosa Publishing House.
6. Thomas Hungerford, Algebra, Springer GTM.
7. I.N. Herstein, Topics in Abstract Algebra, Wiley Eastern Limited.
8. D.S. Malik, J.M. Mordesen, M.K. Sen, Fundamentals of Abstract Algebra, The McGraw-Hill Companies, Inc.

7.10 ANSWERS TO CHECK YOUR PROGRESS

1. Refer 7.1.1

2. Refer 7.2

3. Refer 7.4.5

4. a – Refer 7.4.6

b. Refer – 7.4.8