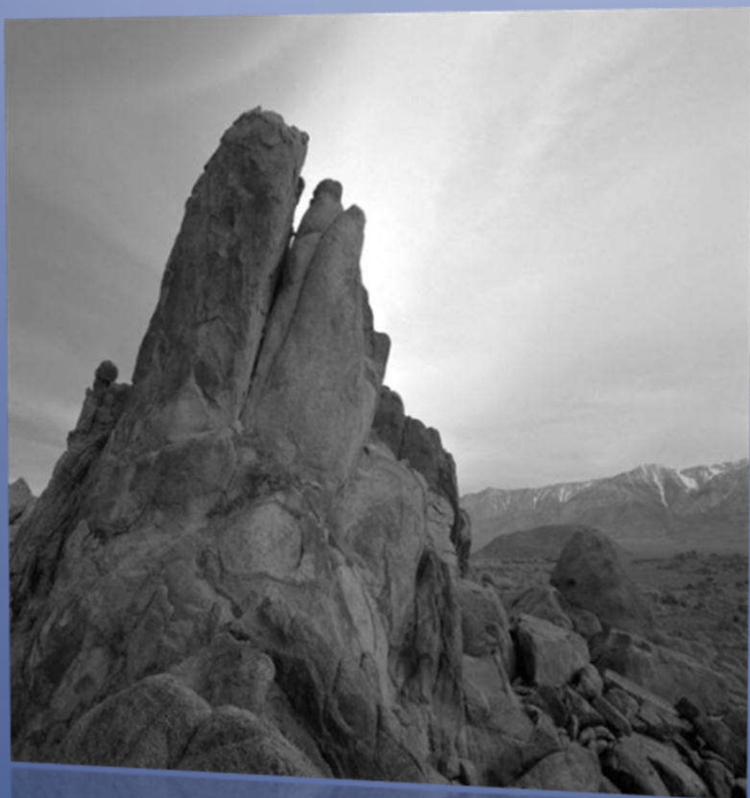


ELEMENTARY LINEAR ALGEBRA

MUHAMMAD USMAN HAMID

PREFACE (Scope of Linear Algebra)

Linear Algebra is the study of vectors and linear transformations. The main objective of this course is to help students learn in rigorous manner, the tools and methods essential for studying the solution spaces of problems in mathematics, engineering, the natural sciences and social sciences and develop mathematical skills needed to apply these to the problems arising within their field of study; and to various real world problems.

Course Contents:

- System of Linear Equations: Representation in matrix form, matrices, operations on matrices, echelon and reduced echelon form, inverse of a matrix (by elementary row operations), solution of linear system, Gauss-Jordan method, Gaussian elimination.
- Vector Spaces: Definition and examples, subspaces. Linear combination and spanning set. Linearly Independent sets. Finitely generated vector spaces. Bases and dimension of a vector space. Operations on subspaces, Intersections, sums and direct sums of subspaces. Quotient Spaces.
- Inner product Spaces: Definition and examples. Properties, Projection. Cauchy inequality. Orthogonal and orthonormal basis. Gram Schmidt Process.
- Determinants: Permutations of order two and three and definitions of determinants of the same order. Computing of determinants. Definition of higher order determinants. Properties. Expansion of determinants.
- Diagonalization, Eigen-values and eigenvectors
- Linear mappings: Definition and examples. Kernel and image of a linear mapping. Rank and nullity. Reflections, projections, and homotheties. Change of basis. Theorem of Hamilton-Cayley.

Recommended Books:

- Curtis C. W., Linear Algebra
- Apostol T., Multi Variable Calculus and Linear Algebra.
- Anton H., Rorres C., Elementary Linear Algebra: Applications Version
- Dr. Karamat Hussain, Linear Algebra
- Linear Algebra by Seymour Lipschutz

CHAPTER # 1

SYSTEMS OF LINEAR EQUATIONS

Systems of linear equations play an important and motivating role in the subject of linear algebra. In fact, many problems in linear algebra reduce to finding the solution of a system of linear equations. Thus, the techniques introduced in this chapter will be applicable to abstract ideas introduced later. On the other hand, some of the abstract results will give us new insights into the structure and properties of systems of linear equations. All our systems of linear equations involve scalars as both coefficients and constants, and such scalars may come from any number field **F**. There is almost no loss in generality if the reader assumes that all our scalars are real numbers — that is, that they come from the real field **R**.

Linear Equation: $(\boldsymbol{ax+b=0}\ ;\ \boldsymbol{a\neq 0})$

It is an algebraic equation in which each term has an exponent of one and graphing of equation results in a straight line.

Or A linear equation in unknowns $x_1, x_2, \dots, x_n$ is an equation that can be put in the standard form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ where $a_1, a_2, \dots, a_n$ and b are constants. The constant a_k is called the coefficient of x_k , and b is called the constant term of the equation. e.g $6x_1 + 7x_2 = 5$, $2x + 3y + 4z = -1$

Solutions of Linear Equation:

A **solution** of the linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is a list of values for the unknowns or, equivalently, a vector **u** in $\mathbf{R}^n$, say

$x_1 = k_1, x_2 = k_2, \dots, x_n = k_n$ or $\vec{u} = (k_1, k_2, \dots, k_n)$ such that the following statement (obtained by substituting k_i for x_i in the equation) is true:

$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$

In such a case we say that **u** satisfies the equation.

Example : Consider the following linear equation in three unknowns x, y, z: $x + 2y - 3z = 6$ We note that $x = 5; y = 2; z = 1$, or, equivalently, the vector $\vec{u} = (5,2,1)$ is a solution of the equation. That is, $5 + 2(2) - 3(1) = 6$

On the other hand, $\vec{v} = (1,2,3)$ is not a solution, because on substitution, we do not get a true statement: $1 + 2(2) - 3(3) = -4 \neq 6$

System of Linear Equations (System in which more than one linear equations involve)

A system of linear equations is a list of linear equations with the same unknowns. In particular, a system of 'm' linear equations $L_1, L_2, \ldots, L_m$ in 'n' unknowns $x_1, x_2, \ldots, x_n$ can be put in the standard form

$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$

::

$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$

$\boxed{a_{ij}x_j = b_i}$

m = No. of equations
n = No. of unknowns

where the a_{ij} and b_i are constants. The number a_{ij} is the coefficient of the unknown x_j in the equation L_i, and the number b_i is the constant term of the equation L_i.

- The system of linear equations is called an $m \times n$ system. It is called a **square system** if $m = n$ that is, if the number m of equations is equal to the number n of unknowns.
- The system is said to be **homogeneous** if all the constant terms are zero that is, if $b_1 = 0, b_2 = 0, \ldots\ldots, b_n = 0$ Otherwise the system is said to be **nonhomogeneous (inhomogeneous**.).
- A **solution** (or a particular solution) of the system (above) is a list of values for the unknowns or, equivalently, a vector **u** in $\mathbf{R}^n$, which is a solution of each of the equations in the system. The set of all solutions of the system is called **the solution set or the general solution of the system**.
- A finite set of linear equations is called a **system of linear equations**, or more briefly a **linear system**. The variables are called **unknown**.
- A linear equation does not involve any products or roots of variables. All variables occur only to the first power, and do not appear as arguments of trigonometric, logarithmic or exponential functions.

EXAMPLES FOR LINEAR AND NON – LINEAR EQUATIONS

- $x + 3y = 7$ linear
- $5x + 7y - 8yz = 16$ not linear
- $x + \pi y + ez = log5$ linear for constants π, e
- $\frac{1}{2}x - y + 3z = -1$ linear
- $x_1 - 2x_2 - 3x_3 + x_4 = 0$ linear
- $x_1 + x_2 + x_3 + \cdots ... + x_n = 1$ linear
- $x + 3y^2 = 4$ not linear
- $3x + 2y - xy = 5$ not linear
- $Sinx + y = 0$ not linear
- $\sqrt{x_1} + 2x_2 + x_3 = 1$ not linear
- $x_1 + 5x_2 - \sqrt{2}x_3 = 1$ linear
- $x_1 + 3x_2 + x_1x_3 = 2$ not linear
- $x_1 = -7x_2 + 3x_3$ linear
- ${x_1}^{-2} + x_2 + 8x_3 = 5$ not linear
- ${x_1}^{3/5} - 2x_2 + x_3 = 4$ not linear
- $\pi x_1 - \sqrt{2}x_2 = 7^{\frac{1}{3}}$ linear
- $2^{\frac{1}{3}}x + \sqrt{3}y = 1$ linear
- $Cos\left(\frac{\pi}{7}\right)x - 4y = log3$ not linear
- $xy = 1$ not linear
- $y + 7 = x$ linear
- $\frac{\pi}{7}Cosx - 4y = 0$ not linear

TRY OTHERS ALSO!!!!!!!!

لینئر کی پہچان: ویری ایبل کی پاور ایک سے کم زیادہ نہ ہو۔ ویری ایبل کے ساتھ اس کا ڈیری ویٹو نہ لکھا گیا ہو۔ دو ویری ایبل اکٹھے نہ ہوں۔

Variable not appears in this form $tanx, logx, sinx, cosx, \sqrt{x}, e^x$ etc.

Example: Consider the following system of linear equations:

$x_1 + x_2 + 4x_3 + 3x_4 = 5$

$2x_1 + 3x_2 + x_3 - 2x_4 = 1$

$x_1 + 2x_2 - 5x_3 + 4x_4 = 3$

It is a 3×4 system because it has three equations in four unknowns. Determine whether (a) $u = (-8, 6, 1, 1)$ and (b) $v = (-10, 5, 1, 2)$ are solutions of the system.

Solution:

(a) Substitute the values of **u** in each equation, obtaining

$-8 + 6 + 4(1) + 3(\ 1) = 5 \Rightarrow 5 = 5$

$2(-8) + 3(6) + 1 - 2(1) = 1 \Rightarrow 1 = 1$

$-8 + 2(6) - 5(1) + 4(1) = 3 \Rightarrow 3 = 3$

Yes, **u** is a solution of the system because it is a solution of each equation.

(b) Substitute the values of **v** into each successive equation, obtaining

$-10 + 5 + 4(1) + 3(\ 2) = 5 \Rightarrow 5 = 5$

$2(-10) + 3(5) + 1 - 2(2) = 1 \Rightarrow -8 \neq 1$

No, **v** is not a solution of the system, because it is not a solution of the second equation. (We do not need to substitute **v** into the third equation.)

Consistent and Inconsistent Solutions:

The system of linear equations is said to be **consistent** if it has one or more solutions, and it is said to be **inconsistent** if it has no solution.

Underdetermined: A system of linear equations is considered underdetermined if there are fewer equations than unknowns. $m < n$

Over determined: A system of linear equations is considered over determined if there are more equations than unknowns. $n < m$

PRACTICE: (Solution to system of linear equations)

1. Consider the following system of linear equations:

$$2x_1 - 4x_2 - x_3 = 1$$

$$x_1 - 3x_2 + x_3 = 1$$

$$3x_1 - 5x_2 - 3x_3 = 1$$

Determine whether given 3 – tuples are solutions of the system?

(a) $(3,1,1)$
(b) $(3,-1,1)$
(c) $(13,5,2)$
(d) $\left(\frac{13}{2},\frac{5}{2},2\right)$
(e) $(17,7,5)$

2. Consider the following system of linear equations:

$$x + 2y - 2z = 3$$

$$3x - y + z = 1$$

$$-x + 5y - 5z = 5$$

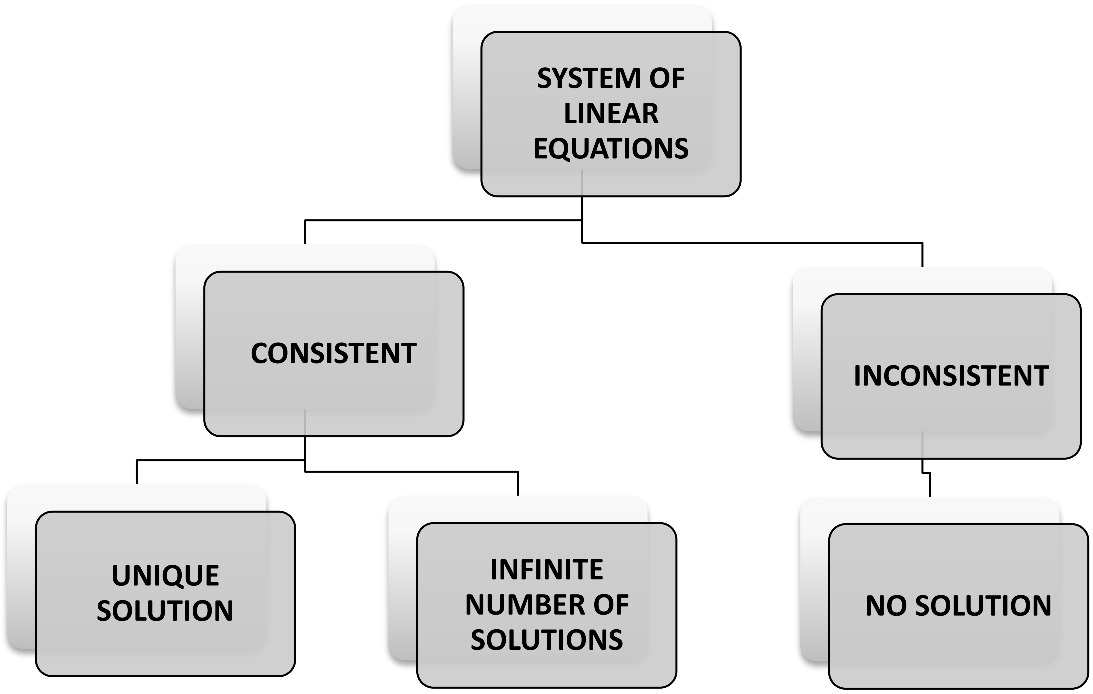
Determine whether given 3 – tuples are solutions of the system?

a) $\left(\frac{5}{7},\frac{8}{7},1\right)$
b) $\left(\frac{5}{7},\frac{8}{7},0\right)$
c) $(5,8,1)$
d) $\left(\frac{5}{7},\frac{10}{7},\frac{2}{7}\right)$
e) $\left(\frac{5}{7},\frac{22}{7},2\right)$

If the field **F** of scalars is infinite, such as when **F** is the real field **R** or the complex field **C**, then we have the following important result.

Result: Suppose the field **F** is infinite. Then any system of linear equations has

(i) a unique solution, (ii) no solution, or (iii) an infinite number of solutions.



Remark: (Geometrical Presentation / Graphics)

Linear system in two unknowns arise in connection with intersection of lines.

- The lines may be parallel and distinct, in which case there is no intersection and consequently no solution.
- The lines may be intersect at only one point, in which case the system has exactly one solution.
- The lines may coincide, in which case there are infinitely many points of intersection (the points on the common line) and consequently infinitely many solutions. (in such system, all equations will be same with few common factors)

Example (A Linear System with one Solution):

Solve the following system of linear equations:

$x - y = 1$ ………..(i)

$2x + y = 6$ ………..(ii)

Solution:

$(i) \Rightarrow -2x + 2y = -2$ multiplying with -2

Adding (i) with (ii) $\Rightarrow -2x + 2y + 2x + y = -2 + 6 \Rightarrow y = \frac{4}{3}$

$(i) \Rightarrow x - \frac{4}{3} = 1 \Rightarrow x = \frac{7}{3}$

$S.S = \left(x = \frac{7}{3}, y = \frac{4}{3}\right)$ Geometrically this means that the lines represented by the equations in the system intersect at the single point $\left(x = \frac{7}{3}, y = \frac{4}{3}\right)$

Example (A Linear System with No Solution):

Solve the following system of linear equations:

$x + y = 4$ ………..(i)

$3x + 3y = 6$ ………..(ii)

Solution:

$(i) \Rightarrow -3x - 3y = -12$ multiplying with -3

Adding (i) with (ii) $\Rightarrow -3x - 3y + 3x + 3y = -12 + 6 \Rightarrow 0 = -6$

The result is contradictory, so the given system has no solution. Geometrically this means that the lines may be parallel and distinct, in this case there is no intersection and consequently no solution.

Example (A Linear System with Infinitely many Solutions):

Solve the following system of linear equations:

$4x - 2y = 1$ ....(i) $16x - 8y = 4$ ....(ii)

Solution:

$(i) \Rightarrow -16x + 8y = -4$ multiplying with -4

Adding (i) with (ii) $\Rightarrow -16x + 8y + 16x - 8y = -4 + 4 \Rightarrow 0 = 0$

Equation $0 = 0$ does not impose any restriction on ‘x’ and ‘y’ and hence can be omitted. Thus the solution of the system are those values of ‘x’ and ‘y’ that satisfy the single equation $4x - 2y = 1$

Geometrically this means that the lines corresponding to the two equations in the original system coincide. And this system will have infinitely many solutions.

How to Find Few Solutions of Such System?

- Find the value of ‘x’ from Common equation.
- Put $y = t$ ‘t’ being **Parameter** (arbitrary value instead of actual value)
- Replace $y = t$ in given system.
- Use $t = 0,1,2,3,$ .... Upon your taste and get different answers.
- We may apply same procedure by replacing ‘x’ and ‘y’

Example: Want to find different solutions for problem as follows using **Parametric Equation** (arbitrary equation using Parameter instead of actual value).

$4x - 2y = 1$ ....(i) $16x - 8y = 4$ ....(ii)

Solution:

$\Rightarrow x = \frac{1}{4} + \frac{1}{2}y$ and put $y = t$ $\Rightarrow x = \frac{1}{4} + \frac{1}{2}t$

$S.S = \left(x = \frac{1}{4}, y = 0\right)$ $t = 0$, $S.S = \left(x = \frac{3}{4}, y = 1\right)$ $t = 1$

$S.S = \left(x = -\frac{1}{4}, y = -1\right)$ $t = -1$

Example: Want to find different solutions for problem as follows using **Parametric Equation** (arbitrary equation using Parameter instead of actual value).

$x - y + 2z = 5$ ……….(i)

$2x - y + 4z = 10$ ……….(ii)

$3x - 3y + 6z = 15$ ……….(ii)

Solution:

Since above all equations have same graphics or formation. Therefore will have infinitely many solutions. We will solve it using parametric equations.

In above all equations we have the parallel form $x - y + 2z = 5$

$\Rightarrow x = 5 + y - 2z$ and put $y = r, z = s$

$\Rightarrow x = 5 + r - 2s$

$S.S = (x = 5, y = 0, z = 0) \quad r = 0, s = 0$

$S.S = (x = 6, y = 1, z = 0) \quad r = 1, s = 0$

$S.S = (x = 4, y = 1, z = 1) \quad r = 1, s = 1$

$General\ Solution = \{(5,0,0), (6,1,0), (4,1,1)\}$

TRY OTHERS ALSO!!!!!!!!

How to find solution of more than two equations?

1st method: find x,y,z solving equations in pair (Lengthy Process)

2nd: solve two equations, find x,y and put in 3rd equation to get value of z.

3rd method: observe given equations and take common if possible and then check all equations are same or not, if same then solution will be infinite.

General Solution:

If a linear system has infinitely many solutions, then a set of parametric equations from which all solutions can be obtained by assigning numerical values to the parameters is called a **General Solution** of the system.

PRACTICE:

1. In each part, solve the linear system, if possible, and use the result to determine whether the lines represented by the equations in the system have zero, one, or infinitely many points of intersection. If there is a single point of intersection, give its coordinates, and if there are infinitely many, find parametric equations for them.
   a) $3x - 2y = 4$ and $6x - 4y = 9$
   b) $2x - 4y = 1$ and $4x - 8y = 2$
   c) $x - 2y = 0$ and $x - 4y = 8$

2. In each part use parametric equations to describe the solution set of linear equations.

a) $7x - 5y = 3$
b) $x + 10y = 2$
c) $3x_1 - 5x_2 + 4x_3 = 7$
d) $-8x_1 + 2x_2 - 5x_3 + 6x_4 = 1$
e) $3v - 8w + 2x - y + 4z = 0$
g) $4x_1 + 2x_2 + 3x_3 + x_4 = 20$
f) $x_1 + 3x_2 - 12x_3 = 3$
h) $v + w + x - 5y + 7z = 0$

3. In each part use parametric equations to describe the infinitely many solutions of linear equations.

a) $2x - 3y = 1$ and $6x - 9y = 3$
b) $x_1 + 3x_2 - x_3 = -4$, $3x_1 + 9x_2 - 3x_3 = -12$ and $-x_1 - 3x_2 + x_3 = 4$
c) $6x_1 + 2x_2 = -8$ and $3x_1 + x_2 = -4$
d) $2x - y + 2z = -4$, $6x - 3y + 6z = -12$ and $-4x + 2y - 4z = 8$

Matrices:

A matrix A over a field **F** or, simply, a matrix A (when **F** is implicit) is a rectangular array of scalars usually presented in the following form:

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

The numbers in the array are called the **entries** in the matrix.

Augmented and Coefficient Matrices of a System

Consider the general system of m equations in n unknowns.

$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$

::

$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$

Coefficient matrix

Consider a system of linear equations $a_{ij}x_j = b_i$ then coefficient matrix is defined as $A = [a_{ij}]$

Augmented matrix

Consider a system of linear equations $a_{ij}x_j = b_i$ then augmented matrix is defined as

$A = [a_{ij} \;\vdots\; b_i]$

Such a system has associated with it the following two matrices:

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix} \text{ and } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]$$

The first matrix M is called the **augmented matrix** of the system, and the second matrix A is called the **coefficient matrix**.

The coefficient matrix A is simply the matrix of coefficients, which is the augmented matrix M without the last column of constants. Some texts write $M = [A, B]$ to emphasize the two parts of M, where B denotes the column vector of constants.

Example:

Consider the general system of 3 equations in 3 unknowns.

$x_1 + x_2 + 2x_3 = 9$

$2x_1 + 4x_2 - 3x_3 = 1$

$3x_1 + 6x_2 - 5x_3 = 0$

Then $M = augmented\ matrix = \begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$

and $A = coefficient\ matrix = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix}$

Example:

Consider

$$M = augmented\ matrix = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Then the general system of 3 equations in 3 unknowns is as follows;

$x_1 = 1$

$x_2 = 2$

$x_3 = 3$

PRACTICE:

1. In each part, find a linear system in the unknowns $x_1, x_2, \ldots$ that corresponds to the given augmented matrix.

 a) $\begin{bmatrix} 3 & 0 & -2 & 5 \\ 7 & 1 & 4 & -3 \\ 0 & -2 & 1 & 7 \end{bmatrix}$

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

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

 d) $\begin{bmatrix} 3 & 0 & 1 & -4 & 3 \\ -4 & 0 & 4 & 1 & -3 \\ -1 & 3 & 0 & -2 & -9 \\ 0 & 0 & 0 & -1 & -2 \end{bmatrix}$

2. In each part, find the augmented matrix for the given linear system.

 a) $-2x_1 = 6$, $3x_1 = 8$, $9x_1 = -3$

 b) $6x_1 - x_2 + 3x_3 = 4$, $5x_2 - x_3 = 1$

 c) $2x_1 - 2x_2 = -1$, $4x_1 + 5x_2 = 3$, $7x_1 + 3x_2 = 0$

 d) $x_1 = 6$, $3x_1 = 8$, $9x_1 = -3$

 e) $2x_1 + 2x_3 = -1$
 $3x_1 - x_2 + 4x_3 = 7$
 $6x_1 + x_2 - x_3 = 0$

 f) $2x_2 - 3x_4 + x_5 = 0$
 $-3x_1 - x_2 + x_3 = -1$
 $6x_1 + 2x_2 - x_3 + 2x_4 - 3x_5 = 6$

Degenerate Linear Equations A linear equation is said to be degenerate if all the coefficients are zero, that is, if it has the form

$0x_1 + 0x_2 + \cdots + 0x_n = b$

The solution of such an equation depends only on the value of the constant b. Specifically,

(i) If b ≠ 0, then the equation has no solution.

(ii) If b = 0, then every vector $\vec{u} = (k_1, k_2, \dots, k_n)$ in $\mathbf{R}^{\mathbf{n}}$ is a solution.

The following theorem applies.

Theorem: Let ℓ be a system of linear equations that contains a degenerate equation L, say with constant b.

(i) If b ≠ 0, then the system ℓ has no solution.

(ii) If b $= 0$, then L may be deleted from the system without changing the solution set of the system.

Part (i) comes from the fact that the degenerate equation has no solution, so the system has no solution.

Part (ii) comes from the fact that every element in $\mathbf{R}^{\mathbf{n}}$ is a solution of the degenerate equation.

Leading Unknown in a Non-degenerate Linear Equation

Let 'L' be a non-degenerate linear equation. This means one or more of the coefficients of L are not zero. By the leading unknown of L, we mean the first unknown in L with a nonzero coefficient.

For example, x_3 and y are the leading unknowns, respectively, in the equations

$0x_1 + 0x_2 + 5x_3 + 6x_4 + 0x_5 + 8x_6 = 7$ and $0x + 2y - 4z = 5$

We frequently omit terms with zero coefficients, so the above equations would be written as $5x_3 + 6x_4 + 8x_6 = 7$ and $2y - 4z = 5$

In such a case, the leading unknown appears first.

Linear Combination of System of Linear equations

Consider the system of m linear equations in n unknowns. Let L be the linear equation obtained by multiplying the m equations by constants $c_1, c_2, \dots, c_n$ respectively, and then adding the resulting equations. Specifically, let L be the following linear equation:

$$(c_1 a_{11} + \cdots \dots + c_m a_{m1})x_1 + \cdots + (c_1 a_{1n} + \cdots + c_m a_{mn})x_n = c_1 b_1 + \cdots + c_m b_m$$

Then L is called a linear combination of the equations in the system.

EXAMPLE: Let L_1, L_2, L_3 denote, respectively, the three equations in

$$x_1 + x_2 + 4x_3 + 3x_4 = 5$$

$$2x_1 + 3x_2 + x_3 - 2x_4 = 1$$

$$x_1 + 2x_2 - 5x_3 + 4x_4 = 3$$

Let L be the equation obtained by multiplying L_1, L_2, L_3 by 3, −2, 4, respectively, and then adding. Namely,

$$3L_1:\ 3\,x_1 + 3x_2 + 12x_3 + 9x_4 = 15$$

$$-2L_2:\ -4x_1 - 6x_2 - 2x_3 + 4x_4 = -2$$

$$4L_3:\ 4x_1 + 8x_2 - 20x_3 + 16x_4 = 12$$

Then Sum will be L: $3x_1 + 5x_2 - 10x_3 + 29x_4 = 25$

Then L is a linear combination of L_1, L_2, L_3. As expected, the solution $u = (-8, 6, 1, 1)$ of the system is also a solution of L. That is, substituting 'u' in L, we obtain a true statement:

$$3(-8) + 5(6) - 1(1) + 29(\ 1) = 25 \Rightarrow 25 = 25$$

The following theorem holds.

Theorem (when two systems have same solution): Two systems of linear equations have the same solutions if and only if each equation in each system is a linear combination of the equations in the other system.

PRACTICE: Show that if the linear equations

$x_1 + kx_2 = c$

$x_1 + lx_2 = d$

Have the same solution set, then the two equations are identical. (i.e. $k = l, c = d$)

Equivalent Systems: Two systems of linear equations are said to be equivalent if they have the same solutions.

Elementary Operations (Elementary Row Operations)

The basic method for solving a linear system is to perform algebraic operations on the system that do not alter the solution set and that produce a succession of increasingly simpler system, until a point is reached where it can be ascertained whether the system is consistent, and if so, what its solutions are. Typically the algebraic operations are:

1. Multiply an equation through by a non – zero constant.
2. Interchange two equations.
3. Add a constant times one equation to another.

Since the rows (horizontal lines) of an augmented matrix correspond to the equations in the associated system, these three operations correspond to the following operations on the rows of the augmented matrix;

1. Multiply a row through by a non – zero constant.
2. Interchange two rows.
3. Add a constant times one row to another. Or replace an equation by the sum of a multiple of another equation and itself.

These are called **Elementary Row Operations** on a matrix.

The main property of the above elementary operations is contained in the following theorem.

Theorem: Suppose a system of $\mathcal{M}$ of linear equations is obtained from a system ℓ of linear equations by a finite sequence of elementary operations. Then $\mathcal{M}$ and ℓ have the same solutions.

Remark: Sometimes (say to avoid fractions when all the given scalars are integers) we may apply step 1 and 3 in one step.

EXAMPLE: In the left column we solve a system of linear equations by operating on the equations in the system, and in the right column we solve the system by operation on the rows of the augmented matrix.

$x + y + 2z = 9$

$2x + 4y - 3z = 1$

$3x + 6y - 5z = 0$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

$\Rightarrow \boldsymbol{E_2 - 2E_1}$ $\qquad$ $\sim \boldsymbol{R_2 - 2R_1}$

$x + y + 2z = 9$

$2y - 7z = -17$

$3x + 6y - 5z = 0$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

$\Rightarrow \boldsymbol{-3E_1 + E_3}$ $\qquad$ $\sim -\boldsymbol{3R_1 + R_3}$

$x + y + 2z = 9$

$2y - 7z = -17$

$3y - 11z = -27$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

$\Rightarrow \frac{1}{2}\boldsymbol{E_2}$ $\sim \frac{1}{2}\boldsymbol{R_2}$

$x + y + 2z = 9$

$y - \frac{7}{2}z = -\frac{17}{2}$

$3y - 11z = -27$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

$\Rightarrow -3\boldsymbol{E_2} + \boldsymbol{E_3}$ $\sim -3\boldsymbol{R_2} + \boldsymbol{R_3}$

$x + y + 2z = 9$

$y - \frac{7}{2}z = -\frac{17}{2}$

$-\frac{1}{2}z = -\frac{3}{2}$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

$\Rightarrow -2\boldsymbol{E_3}$ $\sim -2\boldsymbol{R_3}$

$x + y + 2z = 9$

$y - \frac{7}{2}z = -\frac{17}{2}$

$z = 3$

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$\Rightarrow -\boldsymbol{E_2} + \boldsymbol{E_1}$ $\sim -\boldsymbol{R_2} + \boldsymbol{R_1}$

$x + \frac{11}{2}z = \frac{35}{2}$

$y - \frac{7}{2}z = -\frac{17}{2}$

$z = 3$

$$\begin{bmatrix} 1 & 0 & \frac{11}{2} & \frac{35}{2} \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$\Rightarrow \left(-\frac{\boldsymbol{11}}{\boldsymbol{2}}\boldsymbol{E_3} + \boldsymbol{E_1}\right) \& \left(\frac{\boldsymbol{7}}{\boldsymbol{2}}\boldsymbol{E_3} + \boldsymbol{E_2}\right)$ $\sim \left(-\frac{\boldsymbol{11}}{\boldsymbol{2}}\boldsymbol{R_3} + \boldsymbol{R_1}\right) \& \left(\frac{\boldsymbol{7}}{\boldsymbol{2}}\boldsymbol{R_3} + \boldsymbol{R_2}\right)$

$x = 1$

$y = 2$

$z = 3$

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

Thus, the solution is $x = 1, y = 2, z = 3$

We may write $(1,2,3)$ as a required solution. In the order triple form.

PRACTICE:

1. Find a single elementary row operation that will create a '1' in the upper left corner of the given augmented matrices and will not create any fractions in its first rows.

a) $\begin{bmatrix} -3 & -1 & 2 & 4 \\ 2 & -3 & 4 & 2 \\ 0 & 2 & -3 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 0 & -1 & -5 & 0 \\ 2 & -9 & 3 & 2 \\ 1 & 4 & -3 & 3 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 4 & -6 & 8 \\ 7 & 1 & 4 & 3 \\ -5 & 4 & 2 & 7 \end{bmatrix}$

d) $\begin{bmatrix} 7 & -4 & -2 & 2 \\ 3 & -1 & 8 & 1 \\ -6 & 3 & -1 & 4 \end{bmatrix}$

2. Find all values of 'k' for which the given augmented matrices correspond to a consistent linear system.

a) $\begin{bmatrix} 1 & k & -4 \\ 4 & 8 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 1 & k & -1 \\ 4 & 8 & -4 \end{bmatrix}$

c) $\begin{bmatrix} 3 & -4 & k \\ -6 & 8 & 5 \end{bmatrix}$

d) $\begin{bmatrix} k & 1 & -2 \\ 4 & -1 & 2 \end{bmatrix}$

Systems in Triangular and Echelon Forms

The main method for solving systems of linear equations, Gaussian elimination, is treated in the next Section.

Here we consider two simple types of systems of linear equations: systems in triangular form and the more general systems in echelon form.

Triangular Form

Consider the following system of linear equations, which is in triangular form:

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 - 2x_4 &= 9 \\ 5x_2 - 1x_3 + 3x_4 &= 1 \\ 7x_3 - \quad x_4 &= 3 \\ 2x_4 \quad &= 8 \end{aligned}$$

That is, the first unknown x_1 is the leading unknown in the first equation, the second unknown x_2 is the leading unknown in the second equation, and so on.

Definition: The system in which the first unknown x_1 is the leading unknown in the first equation, the second unknown x_2 is the leading unknown in the second equation, and so on. Then such system is called **Triangular system**.

(Example given above)

Thus, in particular, the system is square and each leading unknown is directly to the right of the leading unknown in the preceding equation. Such a triangular system always has a unique solution, which may be obtained by back-substitution.

That is,

(1) First solve the last equation for the last unknown to get $x_4 = 4$.

(2) Then substitute this value $x_4 = 4$ in the next-to-last equation, and solve for the next-to-last unknown x_3 as follows:

$7x_3 - 4 = 3$ or $x_3 = 1$

(3) Now substitute $x_3 = 1$ and $x_4 = 4$ in the second equation, and solve for the second unknown x_2 as follows:

$5x_2 - 1(1) + 3(4) = 1$ or $x_2 = -2$

(4) Finally, substitute $x_2 = -2$, $x_3 = 1$ and $x_4 = 4$ in the first equation, and solve for the first unknown x_1 as follows:

$2x_1 - 3(-2) + 5(1) - 2(4) = 9$ or $x_1 = 3$

Thus, $x_1 = 3$, $x_2 = -2$, $x_3 = 1$ and $x_4 = 4$, or, equivalently, the vector $\boldsymbol{u} = (3, -2,1,4)$ is the unique solution of the system.

Remark: There is an alternative form for back-substitution (which will be used when solving a system using the matrix format). Namely, after first finding the value of the last unknown, we substitute this value for the last unknown in all the preceding equations before solving for the next-to-last unknown. This yields a triangular system with one less equation and one less unknown. For example, in the above triangular system, we substitute $x_4 = 4$ in all the preceding equations to obtain the triangular system

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 17 \\ 5x_2 - 1x_3 &= -1 \\ 7x_3 &= 7 \end{aligned}$$

We then repeat the process using the new last equation. And so on.

PRACTICE: Solve the following Triangular system: (by Back substitution)

$$\begin{aligned} 2x - 6y + 7z &= 1 \\ 4y + 3z &= 8 \\ 2z &= 4 \end{aligned}$$

Pivoting:

Changing the order of equations is called pivoting. It has two types.

1. Partial pivoting
2. Total pivoting

Partial pivoting:

In partial pivoting we interchange rows where pivotal element is zero.

In Partial Pivoting if the pivotal coefficient "$\boldsymbol{a_{ii}}$" happens to be zero or near to zero, the i^{th} column elements are searched for the numerically largest element. Let the j^{th} row $(j > i)$ contains this element, then we interchange the "i^{th}" equation with the "j^{th}" equation and proceed for elimination. This process is continued whenever pivotal coefficients become zero during elimination.

Total pivoting:

In Full (complete, total) pivoting we interchange rows as well as column.

In Total Pivoting we look for an absolutely largest coefficient in the entire system and start the elimination with the corresponding variable, using this coefficient as the pivotal coefficient (may change row and column). Similarly, in the further steps. It is more complicated than Partial Pivoting. Partial Pivoting is preferred for hand calculation.

Why is Pivoting important?:

Because Pivoting made the difference between non-sense and a perfect result.

Pivotal coefficient:

For elimination methods (Gauss's Elimination, Gauss's Jordan) the coefficient of the first unknown in the first equation is called Pivotal Coefficient.

Back substitution:

The analogous algorithm for upper triangular system "$AX = B$" of the form

$$\begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ 0 & a_{22} \dots\dots\dots a_{2n} \\ \vdots & \vdots \qquad\qquad \vdots \\ 0 & 0 \quad \cdots\cdots \quad a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Is called **Back Substitution**.

The solution "x_i" is computed by $\boldsymbol{x_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij} x_j}{a_{ii}}}$ $\quad ; \boldsymbol{i = 1, 2, 3, \dots\dots\dots n}$

Forward substitution

The analogous algorithm for lower triangular system "$LX = B$" of the form

$$\begin{pmatrix} l_{11} & 0 \quad \dots\dots 0 \\ l_{21} & l_{22} \dots\dots\dots 0 \\ \vdots & \vdots \qquad\qquad \vdots \\ l_{m1} & l_{m2} \cdots\cdots \quad l_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Is called **Forward Substitution**.

The solution "x_i" is computed by $\boldsymbol{x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}}}$ $\quad ; \boldsymbol{i = 1, 2, 3, \dots\dots\dots n}$

Echelon Form, Pivot and Free Variables

The following system of linear equations is said to be in echelon form:

$$2x_1 + 6x_2 - 1x_3 + 4x_4 - 2x_5 = 15$$

$$1x_3 + 2x_4 + 2x_5 \;\; = 5$$

$$3x_4 - 9x_5 \;\; = 6$$

That is, no equation is degenerate and the leading unknown in each equation other than the first is to the right of the leading unknown in the preceding equation. **The leading unknowns in the system, x_1, x_3, x_4, are called pivot variables, and the other unknowns, x_2 and x_5, are called free variables. Those positions in which leading 1 occur called pivot positions and pivot column.**

Generally speaking, an echelon system or a system in echelon form has the following form:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{2j_2}x_{j_2} + a_{2j_{2+1}}x_{j_{2+1}} + \cdots + a_{2n}x_n = b_2$$

::

$$a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n = b_r$$

where $1 < j_2 < \ldots. < j_r$ and a_{11}, a_{2j_2},…..,a_{rj_r} are not zero. The pivot variables are x_1, x_{j_2},…, x_{j_r} Note that $r \leq n$.

The solution set of any echelon system is described in the following theorem

Theorem: Consider a system of linear equations in echelon form, say with ‘r equations in ‘n’ unknowns. There are two cases:

(i) r = n. That is, there are as many equations as unknowns (triangular form). Then

the system has a unique solution.

(ii) r < n. That is, there are more unknowns than equations. Then we can arbitrarily assign values to the $n - r$ free variables and solve uniquely for the ‘r’ pivot variables, obtaining a solution of the system.

Suppose an echelon system contains more unknowns than equations. Assuming the field **F** is infinite, the system has an infinite number of solutions, because each of the $n - r$ free variables may be assigned any scalar.

The general solution of a system with free variables may be described in either of two equivalent ways. One description is called the **''Parametric Form''** of the solution, and the other description is called the **''Free –Variable Form.''**

Parametric Form

Consider we have the system

$$2x_1 + 6x_2 - 1x_3 + 4x_4 - 2x_5 = 15$$

$$1x_3 + 2x_4 + 2x_5 = 5$$

$$3x_4 - 9x_5 = 6$$

Procedure

i. Write given system
ii. Solve last equation using parameter
iii. Using back substitution find variables

Assign arbitrary values, called parameters, to the free variables x_2 and x_5, say $x_2 = a$ and $x_5 = b$, and then use back-substitution to obtain values for the pivot variables x_1, x_3, x_5 in terms of the parameters 'a' and 'b'. Specifically,

(1) Substitute $x_5 = b$ in the last equation, and solve for x_4:

$$3x_4 - 9b = 6 \Rightarrow x_4 = 2 + 3b$$

(2) Substitute $x_4 = 2 + 3b$ and $x_5 = b$ into the second equation, and solve for x_3:

$$1x_3 + 2(2 + 3b) + 2(b) = 5 \Rightarrow x_3 = 1 - 8b$$

(3) Substitute $x_2 = a$, $x_3 = 1 - 8b$, $x_4 = 2 + 3b$, $x_5 = b$ into the first equation, and solve for x_1:

$$2x_1 + 6a - 1(1 - 8b) + 4(2 + 3b) - 2b = 15 \Rightarrow x_1 = 4 - 3a - 9b$$

Accordingly, the general solution in parametric form is

$x_1 = 4 - 3a - 9b$; $x_2 = a$; $x_3 = 1 - 8b$; $x_4 = 2 + 3b$; $x_5 = b$ or, equivalently, $u = (4 - 3a - 9b, a, 1 - 8b, 2 + 3b, b)$ where a and b are arbitrary numbers.

Free –Variable Form

Consider we have the system

$2x_1 + 6x_2 - 1x_3 + 4x_4 - 2x_5 = 15$

$$1x_3 + 2x_4 + 2x_5 \quad = 5$$

$$3x_4 - 9x_5 \quad = 6$$

Use back-substitution to solve for the pivot variables x_1, x_3, x_4 directly in terms of the free variables x_2 and x_5.

That is, the last equation gives $x_4 = 2 + 3x_5$

Substitution in the second equation yields $x_3 = 1 - x_5$

and then substitution in the first equation yields $x_1 = 4 - 3x_2 - 9x_5$

Accordingly, $x_1 = 4 - 3x_2 - 9x_5$; $x_2 =$ free variable;

$x_3 = 1 - 8x_5$; $x_4 = 2 + 3x_5$; $x_5 =$free variable

or, equivalently, $u = (4 - 3x_2 - 9x_5, x_2, 1 - 8x_5, 2 + 3x_5, x_5)$ is the free –variable form for the general solution of the system.

We emphasize that there is no difference between the above two forms of the general solution, and the use of one or the other to represent the general solution is simply a matter of taste.

Remark: A particular solution of the above system can be found by assigning any values to the free variables and then solving for the pivot variables by back substitution. For example, setting $x_2 = 1$ and $x_5 = 1$ we obtain

$x_4 = 2 + 3(1) = 5$; $x_3 = 1 - 8(1) = -7$; $x_1 = 4 - 3(1) - 9(1) = -8$

Thus, $u = (-8,1,-7,5,1)$ is the particular solution corresponding to $x_2 = 1$ and $x_5 = 1$.

PRACTICE:

1. Determine the 'Pivot" and 'Free variables' in each of the followings;

$$2x_1 + 6x_2 - 1x_3 + 4x_4 - 2x_5 = 15$$

$$1x_3 + 2x_4 + 2x_5 \quad = 5$$

$$3x_4 - 9x_5 \quad = 6$$

$$2x - 6y + 7z = 1$$

$$4y + 3z = 8$$

$$2z = 4$$

$$x + 2y - 3z = 2$$

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

$$3x + 4y + 5z = 8$$

---

2. Solve using **parametric form** as well as **free variable** form assigning Pivot.

$$2x_1 + 6x_2 - 1x_3 + 4x_4 - 2x_5 = 15$$

$$1x_3 + 2x_4 + 2x_5 \quad = 5$$

$$3x_4 - 9x_5 \quad = 6$$

Echelon Form of a Matrix:

A matrix is said to be in echelon form if it has the following structure;

i. All the non – zero rows proceed the zero rows.
ii. The first non – zero element in each row is **1**.
iii. The preceding number of zeros before the first non – zero element **1** in each row should be greater than its previous row.

For example followings are in echelon form.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 3 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Reduced Echelon Form of a Matrix:

A matrix is said to be in reduced echelon form if it has the following structure;

i. Matrix should be in echelon form.
ii. If the first non – zero element **1** in the i^{th} row of matrix lies in the j^{th} column then all other elements in the j^{th} column are zero.

For example followings are in reduced echelon form.

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

Remark about echelon forms:

i. Every matrix has a unique reduced row echelon form.
ii. Row echelon forms are not unique.
iii. Although row echelon forms are not unique, the reduced row echelon form and all row echelon forms of a matrix A have the same number of zero rows, and the leading 1’s always occur in the same positions.

Row Echelon Form of a Matrix:

A matrix is said to be in row echelon form if it has the following structure;

i. If a row does not consists entirely of zeros, then the first non – zero number in the row is a 1. We call this a **leading 1**.
ii. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
iii. In any two successive rows that do not consists entirely of zeros, the leading **1** in the lower row occurs farther to the right than the leading **1** in the higher row.

For example followings are in row echelon form.

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 4 & -3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Row Reduced Echelon Form of a Matrix:

A matrix is said to be in row reduced echelon form if it has the following structure;

i. If a row does not consists entirely of zeros, then the first non – zero number in the row is a **1**. We call this a **leading 1**.
ii. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
iii. In any two successive rows that do not consists entirely of zeros, the leading **1** in the lower row occurs farther to the right than the leading **1** in the higher row.
iv. Each column that contains a leading **1** has zeros everywhere else in that column.

For example followings are in row reduced echelon form.

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

PRACTICE:

Determine whether the matrix is in row echelon form, reduced row echelon form, both or neither.

i. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

ii. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

iii. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

iv. $\begin{bmatrix} 1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 4 \end{bmatrix}$

v. $\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

vi. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

vii. $\begin{bmatrix} 1 & -7 & 5 & 5 \\ 0 & 1 & 3 & 2 \end{bmatrix}$

viii. $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

ix. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{bmatrix}$

x. $\begin{bmatrix} 1 & 3 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

xi. $\begin{bmatrix} 1 & 5 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

xii. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

xiii. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 7 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

xiv. $\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

Gaussian Elimination

The main method for solving the general system of linear equations is called Gaussian elimination. It essentially consists of two parts:

Part A. (Forward Elimination) Step-by-step reduction of the system yielding either a degenerate equation with no solution (which indicates the system has no solution) or an equivalent simpler system in triangular or echelon form.

Part B. (Backward Elimination) Step-by-step back-substitution to find the solution of the simpler system.

Gaussian Elimination steps (Procedure):

i. Reduce the augmented matrix into echelon form. In this way, the value of last variable is calculated.
ii. Then by backward substitution, the values of remaining unknown can be calculated.

Example: Solve the matrix using Gauss's Elimination method.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Solution: Firstly we reduce the given matrix in echelon form.

Step – I: locate the left most column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step – II: interchange the top row with another row, if necessary, to bring a non – zero entry to the top of the column found in step – I.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim R_{12}$$

Step – III: if the entry that is now at the top of the column found in step – I is ‘a’, multiply the first row by ‘1/a’ in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim \frac{1}{2}R_1$$

Step – IV: add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zero.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim R_3 - 2R_1$$

Step – V: now cover the top row in the matrix and begin again with step – I applied to the submatrix that remains or remaining rows. Continue in this way until the entire matrix is in row echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim -\frac{1}{2}R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \qquad \sim R_3 - 5R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim 2R_3$$

Hence above matrix is in row echelon form.

Thus corresponding system is

$x_5 = 2$

$x_3 - \frac{7}{2}x_5 = -6$

$x_1 + 2x_2 - 5x_3 + 3x_4 + 6x_5 = 14$

Solving for leading variables we obtain and in next line solution using free variable

$x_5 = 2,\ x_3 = -6 + \frac{7}{2}x_5,\ x_1 = 14 - 2x_2 + 5x_3 - 3x_4 - 6x_5$

$x_5 = 2,\ x_3 = -6 + \frac{7}{2}(2) = 1$

$x_1 = 2 - 2x_2 + 5(1) - 3x_4 - 6(2) = 7 - 2x_2 - 3x_4$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4 arbitrary values 'r' and 's' respectively. This yield

$x_1 = 7 - 2r - 3s\ , x_2 = r\ \ , x_3 = 1\ , x_4 = s\ \ ,\ \ x_5 = 2$

Above is our required solution.

PRACTICE:

1. Solve the linear system by Gauss's Elimination method.

i. $\begin{aligned} x_1 + x_2 + 2x_3 &= 8 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 3x_1 - 7x_2 + 4x_3 &= 10 \end{aligned}$

ii. $\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$

iii. $\begin{aligned} x - y + 2z - w &= -1 \\ 2x + y - 2z - 2w &= -2 \\ -x + 2y - 4z + w &= 1 \\ 3x \qquad\qquad - 3w &= -3 \end{aligned}$

iv. $\begin{aligned} -2b + 3c &= 1 \\ 3a + 6b - 3c &= -2 \\ 6a + 6b + 3c &= 5 \end{aligned}$

2. Find two different row echelon forms of $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$

Gauss Jordan Elimination

Procedure:

i. In this method we reduce the augmented matrix into <u>reduced echelon</u> form. In this way, the value of last variable is calculated.
ii. Then by backward substitution, the values of remaining unknown can be calculated.

Example: Solve the matrix using Gauss's Elimination method.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Solution: Firstly we reduce the given matrix in reduced echelon form.

Step – I: locate the left most column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step – II: interchange the top row with another row, if necessary, to bring a non – zero entry to the top of the column found in step – I.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim R_{12}$$

Step – III: if the entry that is now at the top of the column found in step – I is 'a', multiply the first row by '1/a' in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim \frac{1}{2} R_1$$

Step – IV: add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zero.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim -2R_1 + R_3$$

Step – V: now cover the top row in the matrix and begin again with step – I applied to the submatrix that remains or remaining rows. Continue in this way until the entire matrix is in row echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim -\frac{1}{2}R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \qquad \sim R_3 - 5R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim 2R_3$$

Step – VI: Beginning with the last non – zero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1’s.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim R_2 + \frac{7}{2}R_3$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim R_1 - 6R_3$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim R_1 + 5R_2$$

Hence above matrix is in row reduced echelon form. Thus corresponding system is

$x_5 = 2,\ x_3 = 1,\ x_1 + 2x_2 + 3x_4 = 7$

Solving for leading variables we obtain and in next line solution using free variable

$x_5 = 2,\ x_3 = 1,\ x_1 = -2x_2 - 3x_4 - 7$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4 arbitrary values 'r' and 's' respectively. These yields

$x_1 = -2r - 3s - 7$, $x_2 = r$, $x_3 = 1$, $x_4 = s$, $x_5 = 2$

Above is our required solution.

Example:

Solve the linear system by Gauss's Jordan Elimination method.

$x_1 + 3x_2 - 2x_3 + 2x_5 = 0$

$2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = -1$

$5x_3 + 10x_4 + 15x_6 = 5$

$2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 = 6$

Solution:

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} \quad \sim R_2 - 2R_1 \;\; and \;\; R_2 - 2R_4$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} \quad \sim -1R_2$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \end{bmatrix} \quad \sim R_3 - 5R_2 \; and \; R_4 - 4R_2$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_{34}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \frac{1}{6}R_3$$ Echelon form.

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_2 - 3R_3$$

$$\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_1 + 2R_2$$ reduced echelon form.

Thus corresponding system is

$x_6 = \frac{1}{3},\ x_3 + 2x_4 = 0,\ x_1 + 3x_2 + 4x_4 + 2x_5 = 0$

Solving for leading variables we obtain and in next line solution using free variable

$x_6 = \frac{1}{3},\ x_3 = -2x_4\ ,\ x_1 = -3x_2 - 4x_4 - 2x_5$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4, x_5 arbitrary values 'r', 's' and 't' respectively. These yields

$x_1 = -3r - 4s - 2t\ , x_2 = \mathrm{r}\ ,\ x_3 = -2s,\ x_4 = s\ ,\ x_5 = t,\ x_6 = \frac{1}{3}$

Above is our required solution.

☼ Find reduced row echelon forms of $\begin{bmatrix} 2 & 1 & 3 \\ 0 & -2 & -29 \\ 3 & 4 & 5 \end{bmatrix}$ without introducing fractions at any intermediate stages.

PRACTICE:

1. Solve the linear system by Gauss's Jordan Elimination method.

i.
$$2x_1 + 2x_2 + 2x_3 = 0$$
$$-2x_1 + 5x_2 + 2x_3 = 1$$
$$8x_1 + x_2 + 4x_3 = -1$$

ii.
$$x - y + 2z - w = -1$$
$$2x + y - 2z - 2w = -2$$
$$-x + 2y - 4z + w = 1$$
$$3x \qquad\qquad - 3w = -3$$

iii.
$$x_1 + x_2 + 2x_3 = 8$$
$$-x_1 - 2x_2 + 3x_3 = 1$$
$$3x_1 - 7x_2 + 4x_3 = 10$$

iv.
$$-2b + 3c = 1$$
$$3a + 6b - 3c = -2$$
$$6a + 6b + 3c = 5$$

2. Solve the following system for x, y and z. using any method.
$$\frac{1}{x} + \frac{2}{y} - \frac{4}{z} = 1$$
$$\frac{2}{x} + \frac{3}{y} + \frac{8}{z} = 0$$
$$-\frac{1}{x} + \frac{9}{y} + \frac{10}{z} = 5$$

Example: Solve the linear system by Gauss Elimination method. Or show that system has no solution.

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 & 7 \end{bmatrix} \quad \sim R_2 - 2R_1 \; also \sim R_3 - 3R_1$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \quad \sim R_3 - \frac{3}{2}R_2$$

The matrix is now in echelon form. The third row of the echelon matrix corresponds to the degenerate equation $0x_1 + 0x_2 + 0x_3 + 0x_4 = -2$ which has no solution, thus the system has no solution.

Homogeneous system of linear equations

A system of linear equations $AX = B$ is said to be Homogeneous if $AX = 0$.i.e. $B = 0$. **or** A system of linear equations is said to be Homogeneous if the constant terms are all zeros; i.e. the system has the form;

$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0$

::

$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0$

Remark:

- Every homogeneous system of linear equations is consistent because all such systems have $x_1 = 0, x_2 = 0, \ldots\ldots, x_n = 0$ as a solution. This solution is called the **Trivial Solution**. If there are others solutions, they are called **nontrivial solutions**.
- Because a homogeneous linear system always has the trivial solution, there are only two possibilities for its solutions:
  - i. The system has only the trivial solution.
  - ii. The system has infinitely many solutions in addition to the trivial solution.

Example:

Using Gauss Jordan's Elimination method matrix $\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 & 0 \\ 2 & 6 & 0 & 8 & 4 & 18 & 0 \end{bmatrix}$ can be converted into row reduced echelon form as $\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Thus corresponding system is $x_6 = 0,\ x_3 + 2x_4 = 0,\ x_1 + 3x_2 + 4x_4 + 2x_5 = 0$

These yields $x_1 = -3r - 4s - 2t$, $x_2 = \mathrm{r}$, $x_3 = -2s,\ x_4 = s$, $x_5 = t,\ x_6 = 0$

PRACTICE:

1) Solve the linear system <u>by any method</u>, Gauss elimination method or Gauss's Jordan Elimination method

i. $2x_1 + x_2 + 3x_3 = 0$
$x_1 + 2x_2 \qquad = 0$
$\qquad x_2 + x_3 = 0$

ii. $2x - y - 3z = 0$
$-x + 2y - 3z = 0$
$x + y + 4z \quad = 0$

iii. $3x_1 + x_2 + x_3 + x_4 = 0$
$5x_1 - x_2 + x_3 - x_4 = 0$

iv. $v + 3w - 2x = 0$
$2u + v - 4w + 3x = 0$
$2u + 3v + 2w - x = 0$
$-4u - 3v + 5w - 4x = 0$

v. $2x + 2y + 4z = 0$
$w - y - 3z = 0$
$2w + 3x + y + z = 0$
$-2w + x + 3y - 2z = 0$

vi. $3x_1 + 3x_2 + x_4 = 0$
$x_1 + 4x_2 + 2x_3 = 0$
$-2x_2 - 2x_3 - x_4 = 0$
$2x_1 - 4x_2 + x_3 + x_4 = 0$
$x_1 - 2x_2 - x_3 + x_4 = 0$

vii. $2I_1 - I_2 + 3I_3 + 4I_4 = 0$
$I_1 - 2I_3 + 7I_4 = 0$
$3I_1 - 3I_2 + I_3 + 5I_4 = 0$
$2I_1 + I_2 + 4I_3 + 4I_4 = 0$

viii. $Z_3 + Z_4 + Z_5 = 0$
$-Z_1 - Z_2 + 2Z_3 - 3Z_4 + Z_5 = 0$
$Z_1 + Z_2 - 2Z_3 - Z_5 = 0$
$2Z_1 + 2Z_2 - Z_3 + Z_5 = 0$

ix. $x^2 + y^2 + z^2 = 6$
$x^2 - y^2 + 2z^2 = 2$
$2x^2 + y^2 - z^2 = 3$

x. $2Sin \propto -Cos\beta + 3tan\gamma = 3$
$4Sin \propto +2Cos\beta - 2tan\gamma = 2$ find $\propto, \beta, \gamma$
$6Sin \propto -3Cos\beta + tan\gamma = 9$

2) Solve the following systems where ‘a’ , ‘b’ and ‘c’ are constants.
   a) $x_1 + x_2 + x_3 = a$, $2x_1 + 2x_3 = b$, $3x_2 + 3x_3 = c$
   b) $2x + y = a$, $3x + 6y = b$

Remark: A homogeneous system AX = 0 with more unknowns than equations has a nonzero solution.

Nonhomogeneous and Associated Homogeneous Systems

Let AX = B be a nonhomogeneous system of linear equations. Then AX = 0 is called the associated homogeneous system. For example,

$x + 2y - 4z = 7$	$x + 2y - 4z = 0$
$3x - 5y + 6z = 8$	$3x - 5y + 6z = 0$

show a nonhomogeneous system and its associated homogeneous system.

Free variable theorem for Homogeneous system:

If a homogeneous linear system has ‘n’ unknowns, and if the reduced row echelon form of its augmented matrix has ‘r’ non – zero rows, then the system has $n - r$ free variables.

CHAPTER #2

MATRICES AND MATRIX OPERATIONS

Matrices:

A matrix **A** over a field **K** or, simply, a matrix **A** (when **K** is implicit) is a rectangular array of scalars usually presented in the following form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]$$

The numbers in the array are called the **entries** in the matrix.

Note that

- the element a_{ij} , called the *ij*-entry or *ij*-element, appears in row *i* and column j. We frequently denote such a matrix by simply writing $A = [a_{ij}]$
- A matrix with 'm' rows and 'n' columns is called an m – by –n matrix, written $m \times n$. The pair of numbers m and n is called the **size** of the matrix.
- Two matrices **A** and **B** are **equal**, written **A = B**, if they have the same size and if corresponding elements are equal. Thus, the equality of two $m \times n$ matrices is equivalent to a system of mn equalities, one for each corresponding pair of elements.
- A matrix with only one row is called a **row matrix** or **row vector**, and a matrix with only one column is called a **column matrix** or **column vector**.
- A matrix whose entries are all zero is called a **zero matrix** and will usually be denoted by **0** or $\vec{0}$.
- Matrices whose entries are all real numbers are called **real matrices** and are said to be matrices over **R**.
- Analogously, matrices whose entries are all complex numbers are called **complex matrices** and are said to be matrices over **C**. This text will be mainly concerned with such real and complex matrices.

Square matrix:

A square matrix is a matrix with the same number of rows as columns. An $n \times n$ square matrix is said to be of order 'n' and is sometimes called an n-square matrix.

Diagonal and Trace:

Let $A = [a_{ij}]$ be an n-square matrix. The diagonal or main diagonal of A consists of the elements with the same subscripts—that is, $a_{11}, a_{22}, \dots, a_{nn}$

The **trace** of A, written as $tr(A)$, is the sum of the diagonal elements.
Namely, $tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$

For example $A = \begin{bmatrix} -1 & 2 & 7 & 0 \\ 3 & 5 & -8 & 4 \\ 1 & 2 & 7 & -3 \\ 4 & -2 & 1 & 0 \end{bmatrix}$

Then $diag(A) = \{-1,5,7,0\}$ and $tr(A) = -1 + 5 + 7 + 0 = 11$

Remark: (Prove Yourself)

Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are n – square matrices and 'k' is scalar then

1) $tr(A + B) = tr(A) + tr(B)$
2) $tr(A^T) = tr(A)$
3) $tr(kA) = ktr(A)$
4) $tr(AB) = tr(BA)$

Addition and Subtraction of matrices:

If A and B are matrices of same size, then the **sum** $A+\ B$ is the matrix obtained by adding the entries of B to the corresponding entries of A, and the **difference** $A-B$ is the matrix obtained by subtracting the entries of B to the corresponding entries of A.

For example

$$A=\begin{bmatrix} 2 & 1 & 0 & 3 \\ -1 & 0 & 2 & 4 \\ 4 & -2 & 7 & 0 \end{bmatrix} \text{ and } B=\begin{bmatrix} -4 & 3 & 5 & 1 \\ 2 & 2 & 0 & -1 \\ 3 & 2 & -4 & 5 \end{bmatrix}$$

Then $A+B=\begin{bmatrix} -2 & 4 & 5 & 4 \\ 1 & 2 & 2 & 3 \\ 7 & 0 & 3 & 5 \end{bmatrix}$ and $A-B=\begin{bmatrix} 6 & -2 & -5 & 2 \\ -3 & -2 & 2 & 5 \\ 1 & -4 & 11 & -5 \end{bmatrix}$

Remark: addition and Subtraction of two matrices will be defined if both matrices are equal i.e. having same size and corresponding entries.

e.g $A=\begin{bmatrix} 2 & 1 & 0 & 3 \\ -1 & 0 & 2 & 4 \\ 4 & -2 & 7 & 0 \end{bmatrix}$ and $C=\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ are not defined for above operations.

Scalar Multiplication:

If A is any matrix and c is any scalar, then the product cA is the matrix obtained by multiplying each entry of the matrix A by c. the matrix cA is said to be a **Scalar Multiple** of A.

For example

$$A=\begin{bmatrix} 2 & 1 & 0 & 3 \\ -1 & 0 & 2 & 4 \\ 4 & -2 & 7 & 0 \end{bmatrix} \text{ and } B=\begin{bmatrix} -4 & 3 & 5 & 1 \\ 2 & 2 & 0 & -1 \\ 3 & 2 & -4 & 5 \end{bmatrix}$$

Then $(-1)A=\begin{bmatrix} -2 & -1 & 0 & 3 \\ 1 & 0 & -2 & -4 \\ -4 & 2 & -7 & 0 \end{bmatrix}$ and $2B=\begin{bmatrix} -8 & 6 & 10 & 2 \\ 4 & 4 & 0 & -2 \\ 6 & 4 & -8 & 10 \end{bmatrix}$

Matrix Multiplication (Entry method):

If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the product AB is then $m \times n$ matrix whose entries are determined as follows;

To find the entry in the row 'i' and column 'j' of AB, single out row 'i' from the matrix A and column 'j' from the matrix B. multiply the corresponding entries from the row and column together, and then add up the resulting products.

For example if $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

Then $AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

$$\begin{bmatrix} (1)(4)+(2)(0)+(4)(2) & (1)(1)+(2)(-1)+(4)(7) & (1)(4)+(2)(3)+(4)(5) & (1)(3)+(2)(1)+(4)(2) \\ (2)(4)+(6)(0)+(0)(2) & (2)(1)+(6)(-1)+(0)(7) & (2)(4)+(6)(3)+(0)(5) & (2)(3)+(6)(1)+(0)(2) \end{bmatrix}$$

$AB = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$

Row Column Rule (General Product definition):

If $A = [a_{ij}]$ is an $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then the product AB is then $m \times n$ matrix given as follows;

$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj}$

Remark: multiplication of two matrices will be defined if number of column of 1st matrix equals to the number of rows of 2nd matrix.

Transpose of a Matrix: If $A = [a_{ij}]$ is an $m \times n$ matrix, then the **transpose of A**, denoted by A^T is defined to be the $n \times m$ matrix that results by interchanging the rows and columns of A. i.e. $(A^T)_{ij} = (A)_{ji}$

Like $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & 2 \\ 2 & 6 \\ 4 & 0 \end{bmatrix}$

PRACTICE:

1. Suppose that A, B, C, D and E are matrices with the following sizes;

A	B	C	D	E
(4×5)	(4×5)	(5×2)	(4×2)	(5×4)

In each part, determine whether the given matrix expression is defined. For those that are defined, give the size of the resulting matrix.

i. BA
ii. AB^T
iii. $AC + D$
iv. $E(AC)$
v. $A - 3E^T$
vi. $E(5B + A)$
vii. CD^T
viii. DC
ix. $BC - 3D$
x. $D^T(BE)$
xi. $B^TD + ED$
xii. $BA^T + D$

2. Use the following matrices to compute the indicated expression if it is defined.

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

$$E = \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$$

i. $D + E$
ii. $D - E$
iii. $5A$
iv. $2B - C$
v. $4E - 2D$
vi. $A - A$
vii. $-3(D + 2E)$
viii. $tr(D)$
ix. $tr(D - 3E)$
x. $4tr(7B)$
xi. $tr(A)$
xii. $2A^T + C$
xiii. $D^T - E^T$
xiv. $(D - E)^T$
xv. $B^T + 5C^T$
xvi. $\frac{1}{2}C^T - \frac{1}{4}A$
xvii. $B - B^T$
xviii. $2E^T - 3D^T$
xix. $(CD)E$
xx. $(C^TB)A^T$
xxi. $tr(C^TA^T + 2E^T)$
xxii. $B^T(CC^T - A^TA)$
xxiii. $(2E^T - 3D^T)^T$

Matrix form of a linear System (Already discussed)

Consider the general system of m equations in n unknowns.

$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$

$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$

::

$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$

Such a system has associated with it the following form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \Rightarrow AX = B$$

Also

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix} \text{ and } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The first matrix M is called the **augmented matrix** of the system, and the second matrix A is called the **coefficient matrix**.

PRACTICE:

1. Express the given linear system as a single matrix equation $AX = B$

a) $2x_1 - 3x_2 + 5x_3 = 7$
$9x_1 - x_2 + x_3 = -1$
$x_1 + 5x_2 + 4x_3 = 0$

b) $4x_1 - 3x_3 + x_4 = 1$
$5x_1 + x_2 - 8x_4 = 0$
$2x_1 - 5x_2 + 9x_3 - x_4 = 0$
$3x_2 - x_3 + 7x_3 = 2$

c) $x_1 - 2x_2 + 3x_3 = -3$
$2x_1 + x_2 = 0$
$-3x_2 + 4x_3 = 1$
$x_1 + x_3 = 5$

d) $3x_1 + 3x_2 + 3x_3 = -3$
$-x_1 - 5x_2 - 2x_3 = 3$
$-4x_2 + x_3 = 0$

2. Express the matrix equation as a system of linear equation.

a) $\begin{bmatrix} 5 & 6 & -7 \\ -1 & -2 & 3 \\ 0 & 4 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$

b) $\begin{bmatrix} 3 & -1 & 2 \\ 4 & 3 & 7 \\ -2 & 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 0 \\ 5 & -3 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 9 \end{bmatrix}$

d) $\begin{bmatrix} 3 & -2 & 0 & 1 \\ 5 & 0 & 2 & -2 \\ 3 & 1 & 4 & 7 \\ -2 & 5 & 1 & 6 \end{bmatrix}\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

3. Solve the matrices for ‘a’, ‘b’, ‘c’ and ‘d’.

a) $\begin{bmatrix} a & 3 \\ -1 & a+b \end{bmatrix} = \begin{bmatrix} 4 & \mathrm{d} - 2\mathrm{c} \\ d + 2c & -2 \end{bmatrix}$

b) $\begin{bmatrix} a - b & \mathrm{b} + \mathrm{a} \\ 3d + c & 2d - c \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 7 & 6 \end{bmatrix}$

4. Solve the matrices for ‘k’ if any that satisfy the equation.

a) $\begin{bmatrix} k & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & -3 \end{bmatrix}\begin{bmatrix} k \\ 1 \\ 1 \end{bmatrix} = 0$

b) $\begin{bmatrix} 2 & 2 & k \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \\ k \end{bmatrix} = 0$

Partitioned Matrices

A matrix can be subdivided or Partitioned into smaller matrices by inserting horizontal and vertical rules between selected rows and columns.

For example, the following are three possible partitions of a general 3×4 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$.

The first is a partition of A into four submatrices $A_{11}, A_{12}, A_{21}, A_{22}$

$$A = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

The second is a partition of A into its row vectors r_1, r_2, r_3

$$A = \left[\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

The third is a partition of A into its column vectors c_1, c_2, c_3

$$A = \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}$$

Row method

This will be computed by following procedure. If the system is given as follows;

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \Rightarrow AX = B$$

Then $\quad j^{th}\ column\ vector\ of\ AB = A[j^{th}\ column\ vector\ of\ B]$

Or

$$AB\ (column\ by\ column) = A[b_1 \quad b_2 \quad \cdots \quad b_n] = [Ab_1 \quad Ab_2 \quad \cdots \quad Ab_n]$$

For example if $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

Then $$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} (1)(1) + (2)(-1) + (4)(7) \\ (2)(1) + (6)(-1) + (0)(7) \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$$

Column method

This will be computed by following procedure. If the system is given as above;

Then $\quad i^{th}\ row\ vector\ of\ AB = [i^{th}\ row\ vector\ of\ A]B$

Or $$AB\ (row\ by\ row) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_n B \end{bmatrix}$$

For example if $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$

Then $$[1 \quad 2 \quad 4] \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = [12 \quad 27 \quad 30 \quad 13]$$

PRACTICE:

Use the following matrices and either the row method or column method, as appropriate, to find the indicated row or column.

$$A = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix}$$

i. The first row of AB
ii. The third row of AB
iii. The second column of AB
iv. The first column of BA
v. The third row of AA
vi. The third column of AA
vii. The first column of AB
viii. The third column of BB
ix. The second row of BB
x. The first column of AA
xi. The third column of AB
xii. The first row of BA

Linear combination of matrices:

If $A_1, A_2, \dots, A_n$ are matrices of the same sizes and $k_1, k_2, \dots, k_n$ are scalars then an expression of the form $k_1A_1 + k_2A_2 + \cdots + k_nA_n$ is called a **linear combination** of the matrices $A_1, A_2, \dots, A_n$ with coefficients $k_1, k_2, \dots, k_n$

General form: If $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Then

$$Ax = \begin{bmatrix} a_{11}x_1 & a_{12}x_2 & \cdots & a_{1n}x_n \\ a_{21}x_1 & a_{22}x_2 & \cdots & a_{2n}x_n \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & a_{m2}x_2 & \cdots & a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Matrix product as a Linear combination:

For example $\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$ can be written as in linear combination $2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$

Column of a product as a Linear combination:

Since $\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$

then $\begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 27 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 7\begin{bmatrix} 4 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 30 \\ 26 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 5\begin{bmatrix} 4 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 13 \\ 8 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix}$

PRACTICE: Use the following matrices and linear combination, to find the indicated operations.

$A = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix}$

i. Each column vector of AA as a linear combination of the column vector of A
ii. Each column vector of BB as a linear combination of the column vector of B
iii. Each column vector of AB as a linear combination of the column vector of A
iv. Each column vector of BA as a linear combination of the column vector of B

Column Row expansion

Suppose that an $m \times r$ matrix A is partitioned into its 'r' column vectors $c_1, c_2, \dots, c_r$ (each of size $m \times 1$) and an $r \times n$ matrix B is partitioned into its 'r' row vectors $r_1, r_2, \dots, r_r$ (each of size $1 \times n$). Each term in the sum

$c_1 r_1 + c_2 r_2 + \cdots .. + c_r r_r$ has size $m \times n$ so that sum itself is an $m \times n$ matrix.

So from above discussion we write the **column row expansion** of AB as follows;

$$AB = c_1 r_1 + c_2 r_2 + \cdots + c_r r_r$$

Question: find the column row expansion of the product;

$$AB = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 4 \\ -3 & 5 & 1 \end{bmatrix}$$

Solution: using the column of A and rows of B as follows;

$$AB = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [2 \quad 0 \quad 4] + \begin{bmatrix} 3 \\ -1 \end{bmatrix} [-3 \quad 5 \quad 1]$$

$$AB = \begin{bmatrix} 2 & 0 & 4 \\ 4 & 0 & 8 \end{bmatrix} + \begin{bmatrix} -9 & 15 & 3 \\ 3 & -5 & -1 \end{bmatrix} = \begin{bmatrix} -7 & 15 & 7 \\ 7 & -5 & 7 \end{bmatrix}$$

PRACTICE:

Use the column row expansion of AB to express this product as a sum of matrices.

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

Square Root of a Matrix:

A matrix B is said to be square root of a matrix A if $BB = A$

1) Find two square roots of $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$
2) How many different square roots can you find of $A = \begin{bmatrix} 5 & 0 \\ 0 & 9 \end{bmatrix}$?
3) Do you think that every 2×2 matrix has at least one square root? Explain your reasoning.

Properties of matrix arithmetic (addition and scalar multiplication):

i. $A + B = B + A$ (Commutative law for matrix addition)
ii. $A + (B + C) = (A + B) + C$ (Associative law for matrix addition)
iii. $A(BC) = (AB)C$ (Associative law for matrix multiplication)
iv. $A(B + C) = AB + AC$ (Left distributive law)
v. $(B + C)A = BA + CA$ (Right distributive law)
vi. $A(B - C) = AB - AC$
vii. $(B - C)A = BA - CA$
viii. $a(B + C) = aB + aC$
ix. $a(B - C) = aB - aC$
x. $(a + b)C = aC + bC$
xi. $(a - b)C = aC - bC$
xii. $a(bC) = (ab)C$
xiii. $a(BC) = (aB)C = B(aC)$

Question: Verify the property $A(BC) = (AB)C$ for the following matrices

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

Solution:

$$L.H.S = A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}$$

$$L.H.S = A(BC) = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

$$R.H.S = (AB)C = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

$$R.H.S = (AB)C = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

Hence verified that $A(BC) = (AB)C$

PRACTICE:

Verify the law given in following lines for given matrices and scalars.

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}, B = \begin{bmatrix} 0 & 2 \\ 1 & -4 \end{bmatrix}, C = \begin{bmatrix} 4 & 1 \\ -3 & -2 \end{bmatrix}, a = 4 \ , b = -7$$

i. Associative law for matrix addition
ii. Associative law for matrix multiplication
iii. Left distributive law
iv. $A(B - C) = AB - AC$
v. $(B - C)A = BA - CA$
vi. $a(B + C) = aB + aC$
vii. $(a + b)C = aC + bC$
viii. $a(bC) = (ab)C$
ix. $a(BC) = (aB)C = B(aC)$
x. $(A^T)^T = A$
xi. $(AB)^T = B^T A^T$
xii. $(A + B)^T = A^T + B^T$
xiii. $(aC)^T = aC^T$

Important Remark:

In matrix arithmetic, the equality of AB and BA can fail for three reasons;

i. AB may be defined and BA may not defined (for example, if A is 2×3 and B is 3×4)
ii. AB and BA may both be defined, but they may have different sizes (for example, if A is 2×3 and B is 3×2)
iii. AB and BA may both be defined and have the same size, but the two products may be different. (as given below)

Example: for given matrices

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

$AB = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix}$ and $BA = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$

Clearly $AB \neq BA$

Zero Matrix: A matrix whose entries are all zero is called a **zero matrix**.

For example: $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, [0]$

Remark:

i. We will denote a zero matrix by O. i.e. $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
ii. If we want to mentions size then write as $O_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ i.e. $O_{m\times n}$
iii. 'O' plays the same role in matrix equation as the number '0' in the numerical equations. i.e. $O + A = A + O = A$
iv. $A - O = A$
v. $A - A = A + (-A) = O$
vi. $OA = AO = O$
vii. If $cA = O \Rightarrow c = O \ or \ A = O$
viii. If $AB = AC$ and $A \neq O$ then $B = C$ but this law does not hold in general.
ix. If $AB = O$ then $A = O$ or $B = O$ but this law does not hold in general.

Failure of Cancellation law: for given matrices

$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}, C = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$ we have $AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$

Although $A \neq O$ but cancelling of A from both sides of $AB = AC$ does not lead to the statement $B = C$

Failure of Zero Product with non – zero factors:

For given matrices $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}$

We have $AB = O$ but $A \neq O$ also $B \neq O$

Identity Matrix: A matrix in which all diagonal elements are 1 and other are zero is called the identity matrix. **Or** the n – square identity or unit matrix, denoted by I_n, or simply I, is the n-square matrix with 1's on the diagonal and 0's elsewhere. The identity matrix I is similar to the scalar 1 in that, for any n-square matrix A, $AI = IA = A$

For example: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, [1]$

Remark:

i. We will denote an identity matrix by I. i.e. $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
ii. If we want to mentions size then write as $I_{2\times2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ i.e. $I_{m\times n}$
iii. $AI_n = A$ for example $AI_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A$
iv. $I_m A = A$ for example $I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$
v. If R is the reduced row echelon form of an $n \times n$ matrix A , then either R has a row of zeros or R is the identity matrix I_n.
vi. $A^0 = I$ and $A^n = AA \dots A(n\, factor)$
vii. $A^{-n} = (A^{-1})^n = A^{-1}A^{-1} \dots A^{-1}(n\, factor)$

PRACTICE:

1. Compute the given operation using A and B both.

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

   i. A^3 and B^3

   ii. $A^2 - 2A + I$ and $B^2 - 2B + I$

2. Show that the matrix $A = \begin{bmatrix} a & \mathrm{b} \\ c & d \end{bmatrix}$ satisfies the equation

   $A^2 - (a + d)A + (ad - bc)I = 0$

3. A square matrix A is said to be **idempotent** if $A^2 = A$

   Then Show that if A is idempotent, then so is $I - A$

Scalar Matrices:

For any scalar k, the matrix kI that contains k's on the diagonal and 0's elsewhere is called the scalar matrix corresponding to the scalar k. Observe that $(kI)A = k(IA) = kA$ That is, multiplying a matrix A by the scalar matrix kI is equivalent to multiplying A by the scalar k.

Invertible (Nonsingular) Matrices:

A square matrix A is said to be **invertible or nonsingular** if there exists a matrix B such that $AB = BA = I$ where I is the identity matrix. If no such matrix B can be found, then A is said to be **Singular**.

Example: let $A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$

Then $AB = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

Also $BA = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

Remark:

i. If B and C are both inverse of a matrix A then $B = C$

ii. An invertible matrix has exactly one inverse. Denoted as A^{-1}

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

iv. $A^{-1} \neq \frac{1}{A}$ for matrices.

Theorem (necessary and sufficient condition for the existence of invertible matrix):

The matrix $A = \begin{bmatrix} a & \mathrm{b} \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$ in this case $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -\mathrm{b} \\ -c & a \end{bmatrix} = \frac{1}{|A|}(adjA)$

Inverse of a 2×2 Matrix: Let A be an arbitrary 2×2 matrix, then its inverse can be defined as follows; $A^{-1} = \frac{1}{|A|}(adjA)$

In other words, when $|A| \neq 0$, the inverse of a 2×2 matrix A may be obtained from A as follows:

(1) Interchange the two elements on the diagonal.

(2) Take the negatives of the other two elements.

(3) Multiply the resulting matrix by $\frac{1}{|A|}$ or, equivalently, divide each element by $|A|$
In case $|A| = 0$, the matrix A is not invertible.

Example: let $A = \begin{bmatrix} 6 & 1 \\ 5 & 2 \end{bmatrix}, B = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$ then find A^{-1}, B^{-1}

$$A^{-1} = \frac{1}{|A|}(adjA) = \frac{1}{7}\begin{bmatrix} 2 & -1 \\ -5 & 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{5}{7} & \frac{6}{7} \end{bmatrix}$$

For B^{-1} since $|B| = 0$ therefore B is not invertible.

PRACTICE: Compute the inverse of the following matrices.

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

ii. $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$

iii. $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

iv. $\begin{bmatrix} 6 & 4 \\ -2 & -1 \end{bmatrix}$

v. $\begin{bmatrix} Cos\theta & \mathrm{Sin}\theta \\ -\mathrm{Sin}\theta & Cos\theta \end{bmatrix}$

vi. $\begin{bmatrix} \frac{1}{2}(e^x + e^{-x}) & \frac{1}{2}(e^x - e^{-x}) \\ \frac{1}{2}(e^x - e^{-x}) & \frac{1}{2}(e^x + e^{-x}) \end{bmatrix}$

Solution of a Linear System by Matrix Inversion:

- Consider we have a system $AX = B$
- Fins A^{-1} using formula $A^{-1} = \frac{1}{|A|}(adjA)$
- Use formula $X = A^{-1}B$

PRACTICE:

Find the unique solution of given linear system.

i. $3x_1 - 2x_2 = -1$ and $4x_1 + 5x_2 = 3$
ii. $-x_1 + x_2 = 4$ and $-x_1 - 3x_2 = 1$
iii. $6x_1 + x_2 = 0$ and $4x_1 - 3x_2 = -2$
iv. $2x_1 - 2x_2 = 4$ and $x_1 + 4x_2 = 4$

Theorem:

If A and B are invertible matrices with the same size then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

Proof:

Consider $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$

Similarly $(B^{-1}A^{-1})(AB) = B^{-1}(AA^{-1})B = B^{-1}IB^{-1} = B^{-1}B = I$

Then $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I \Rightarrow (AB)^{-1} = B^{-1}A^{-1}$

Or $(AB)(B^{-1}A^{-1}) = I \Rightarrow (AB)^{-1}(AB)(B^{-1}A^{-1}) = (AB)^{-1}I$

$\Rightarrow (AB)^{-1}A(BB^{-1})A^{-1} = (AB)^{-1}I \Rightarrow (AB)^{-1}AA^{-1} = (AB)^{-1}I$

$\Rightarrow (AB)^{-1} = (AB)^{-1} \Rightarrow (AB)^{-1} = B^{-1}A^{-1}$

In general $(A_1A_2 \dots A_n)^{-1} = {A_n}^{-1}{A_{n-1}}^{-1} \dots {A_1}^{-1}$

Polynomial Matrix:

If A is a square matrix, say $n \times n$ and if $P(x) = a_0 + a_1 x + \cdots + a_m x^m$ is any polynomial, then we define the $n \times n$ matrix $P(A)$ to be

$P(A) = a_0 I + a_1 A + \cdots + a_m A^m$

Where I is the $n \times n$ identity matrix. Above expression is called **a matrix polynomial in A**.

Example:

Let $A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}$ and $P(x) = x^2 - 2x - 3$ then find $P(A)$

Solution: given polynomial is $P(x) = x^2 - 2x - 3$

$\Rightarrow P(A) = A^2 - 2A - 3I = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}^2 - 2\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\Rightarrow P(A) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O$ after solving.

Remark:

if $P(x) = P_1(x)P_2(x)$ then for square matrix A we can write $P(A) = P_1(A)P_2(A)$

PRACTICE:

1) Compute $P(A)$ for the given matrix A and the following polynomials;
   $P(x) = x - 2,\ \ P(x) = x^2 - x + 1,\ \ P(x) = x^3 - 2x + 1$
   $A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$
2) Verify the statement $P(A) = P_1(A)P_2(A)$ for the stated matrix A and given polynomials;
   $P(x) = x^2 - 9, P_1(x) = x + 3, P_2(x) = x - 3$

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

ii. $A = \begin{bmatrix} a & \mathrm{b} \\ c & d \end{bmatrix}$

Properties of exponents:

If A is invertible and 'n' is a non – negative integer, then

i. A^{-1} is invertible and $(A^{-1})^{-1} = A$
ii. A^n is invertible and $(A^n)^{-1} = A^{-n} = (A^{-1})^n$
iii. kA is invertible for any non – zero scalar 'k', then $(kA)^{-1} = k^{-1}A^{-1}$

Example: let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$ then

$A^{-3} = (A^{-1})^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$ …….(i)

Also $A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$

Then $(A^3)^{-1} = \frac{1}{|A^3|}(adjA^3) = \frac{1}{(11)(41)-(-30)(-15)}\begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$

$(A^3)^{-1} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$ …….(ii)

Thus from (i) and (ii) $(A^3)^{-1} = A^{-3} = (A^{-1})^3$

Theorem:

If A is invertible matrix, then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$

Proof: since we know that $I^T = I$

Then $A^T(A^{-1})^T = (AA^{-1})^T = I^T = I$ Also $(A^{-1})^T A^T = (A^{-1}A)^T = I^T = I$

Thus $A^T(A^{-1})^T = (A^{-1})^T A^T = I \Rightarrow A^T(A^{-1})^T = I$ and $(A^{-1})^T A^T = I^T$

This implies $(A^T)^{-1} = (A^{-1})^T$

REMEMBER: $(A^T)^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-c}{ad-bc} \\ \frac{-b}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = (A^{-1})^T$

PRACTICE (MIXED):

1) Compute then given operation for following matrices;

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

ii. $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$

iii. $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

iv. $\begin{bmatrix} 6 & 4 \\ -2 & -1 \end{bmatrix}$

a) $(A^T)^{-1} = (A^{-1})^T$
b) $(A^{-1})^{-1} = A$
c) $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$
d) $(ABC)^T = C^T B^T A^T$

2) Use the given information to find A

i. $(7A)^{-1} = \begin{bmatrix} -3 & 7 \\ 1 & -2 \end{bmatrix}$

ii. $(5A^T)^{-1} = \begin{bmatrix} -3 & -1 \\ 5 & 2 \end{bmatrix}$

iii. $(I + 2A)^{-1} = \begin{bmatrix} -1 & 2 \\ 4 & 5 \end{bmatrix}$

iv. $A^{-1} = \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix}$

3) Compute A^{-3} and B^{-3} using A and B both.

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

4) A square matrix A is said to be **idempotent** if $A^2 = A$

Then Show that if A is idempotent, then $2A - I$ is invertible and is its own inverse.

5) Determine whether given matrices are invertible, and if so, find the inverse.

(**Hint:** solve AX = I for X by equating corresponding entries on the two sides)

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

6) Give an example of 2×2 matrices such that $(A + B)(A - B) \neq A^2 - B^2$

Elementary Matrices Let e denotes an elementary row operation and let $e(A)$ denote the results of applying the operation e to a matrix A. Now let E be the matrix obtained by applying e to the identity matrix I ; that is, $E = e(I)$ Then E is called the **elementary matrix** corresponding to the elementary row operation e. Note that E is always a square matrix.

OR A matrix E is called an **elementary matrix** if it can be obtained from an identity matrix by performing a single elementary row operation.

Examples: Below are four elementary matrices and the operations that produce them;

$\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ $\longrightarrow$ Multiply the 2^{nd} row of I_2 by -3

$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ $\longrightarrow$ interchange the 2^{nd} and 4^{th} rows of I_4

$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\longrightarrow$ add 3 time the third row of I_3 to the first row.

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\longrightarrow$ multiply the 1^{st} row of I_3 by 1

PRACTICE: Determine whether the given matrix is elementary.

i. $\begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix}$

ii. $\begin{bmatrix} -5 & 1 \\ 1 & 0 \end{bmatrix}$

iii. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

iv. $\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

v. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 9 \\ 0 & 0 & 1 \end{bmatrix}$

vi. $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

vii. $\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix}$

A Method for Inverting Matrices (Inversion Algorithm)

To find the inverse of an invertible matrix A

- Find a sequence of elementary row operations that reduces A to the identity.
- Then perform that same sequence of operations on I_n to obtain A^{-1}
- For this we will change $[A \mid I]$ to $[I \mid A^{-1}]$

Example: Find the inverse of $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$

Solution:

$$\left[\begin{array}{ccc:ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] \sim R_2 - 2R_1 \,\&\, R_3 - R_1$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] \sim R_3 + 2R_2$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \sim -1R_3$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \sim R_2 + 3R_3 \,\&\, R_1 - 3R_3$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \sim R_1 - 2R_2$$

$$\Rightarrow A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

PROCEDURE

Write given matrix A

Make form $[A: I]$

Use row operation and shift identity matrix on left side. i.e. $[I: A^{-1}]$

Remark:

Often it will not be known in advance if a given $n \times n$ matrix A is invertible. However, if it is not, then by results (part i, iii);

A is invertible. Then the reduced row echelon form of A is I_n

It will be impossible to reduce A to I_n by elementary row operations. This will be signaled by a row zeros appearing on the left side of the partition at some stage of the inversion algorithm. If this occurs, then you can stop the computations and conclude that A is not invertible.

Example:

Find the inverse of $A = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$

Solution:

$$\left[\begin{array}{ccc:ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 2 & 4 & -1 & 0 & 1 & 0 \\ -1 & 2 & 5 & 0 & 0 & 1 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 8 & 9 & 1 & 0 & 1 \end{array}\right] \sim R_2 - 2R_1 \,\&\, R_3 + R_1$$

$$\Rightarrow \left[\begin{array}{ccc:ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right] \sim R_3 + R_2$$

Since we obtain a row of zeros on the left side, A is not invertible.

PRACTICE:

1. Find inverse of given matrices, if exists.

i. $\begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix}$

ii. $\begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix}$

iii. $\begin{bmatrix} 1 & -5 \\ 3 & -16 \end{bmatrix}$

iv. $\begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix}$

v. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$

vi. $\begin{bmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$

vii. $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{10} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{10} \end{bmatrix}$

viii. $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & -\frac{3}{5} & -\frac{3}{10} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{10} \end{bmatrix}$

ix. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

x. $\begin{bmatrix} \sqrt{2} & 3\sqrt{2} & 0 \\ -4\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

xi. $\begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \\ 2 & 7 & 7 \end{bmatrix}$

xii. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 1 & 3 & 5 & 7 \end{bmatrix}$

xiii. $\begin{bmatrix} 2 & -4 & 0 & 0 \\ 1 & 2 & 12 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & -1 & -4 & -5 \end{bmatrix}$

xiv. $\begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 3 & 0 \\ 2 & 1 & 5 & -3 \end{bmatrix}$

2. Find the inverse of each of the following 4×4 matrices, where k_1, k_2, k_3, k_4 and k are all non – zeros.

i. $\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$ iii. $\begin{bmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{bmatrix}$

ii. $\begin{bmatrix} k & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & k & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ iv. $\begin{bmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{bmatrix}$

3..Find all values of ‘c’, if any, for which the given matrix is invertible.

i. $\begin{bmatrix} c & c & c \\ 1 & c & c \\ 1 & 1 & c \end{bmatrix}$ ii. $\begin{bmatrix} c & 1 & 0 \\ 1 & c & 1 \\ 0 & 1 & c \end{bmatrix}$

4..Express the matrix and its inverse as products of elementary matrices.

i. $\begin{bmatrix} -3 & 1 \\ 2 & 2 \end{bmatrix}$ ii. $\begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix}$

iii. $\begin{bmatrix} 1 & 0 & -2 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{bmatrix}$ iv. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

5.. Show that the matrices A and B are row equivalent by finding a sequence of elementary row operations that produces B from A, and then use that result to find a matrix C such that CA = B.

Diagonal Matrix:

A square matrix in which all the entries off the main diagonal (without main diagonal) are zero is called a **Diagonal Matrix**. Examples are given as follows;

$$\begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Remark:

- A general $n \times n$ diagonal matrix D can be written as
$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & d_n \end{bmatrix}$$
- A diagonal matrix is invertible if and only if all of its diagonal entries are non – zero. In this case inverse is
$$D^{-1} = \begin{bmatrix} \frac{1}{d_1} = d_1^{-1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} = d_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{d_n} = d_n^{-1} \end{bmatrix}$$
- If D is the diagonal matrix and ‘k’ is a positive integer, then
$$D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & d_n^k \end{bmatrix}$$

Is Null (Zero) matrix a diagonal matrix? Or why Null (Zero) matrix a diagonal matrix?

A diagonal matrix is one in which all non – diagonal entries are zero. Entries on the main diagonal may or may not be zero. Clearly this is also satisfied. Hence, a zero square matrix is upper and lower triangular as well as a diagonal matrix.

Example: if $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ then

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}, A^{5} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -243 & 0 \\ 0 & 0 & 32 \end{bmatrix}, A^{-5} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{243} & 0 \\ 0 & 0 & \frac{1}{32} \end{bmatrix}$$

Triangular Matrix: A square matrix $A = [a_{ij}]$ that is either upper triangular or lower triangular is called Triangular matrix.

For example; $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{mn} \end{bmatrix}$

Lower triangulation matrix: A matrix having only <u>zeros above the diagonal</u> is called Lower Triangular matrix.

(Or)

A "$n \times n$" matrix “L” is lower triangular if its entries satisfy $l_{ij} = 0 \ for \ i < j$

i.e. $\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$

Upper triangulation matrix: A matrix having only <u>zeros below the diagonal</u> is called Upper Triangular matrix.

(Or)

A "$n \times n$" matrix “U” is upper triangular if its entries satisfy $u_{ij} = 0 \ for \ i > j$

i.e. $\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$

REMARKS:

- A square matrix $A = [a_{ij}]$ is upper triangular iff all entries to the left of the main diagonal are zero. i.e. $a_{ij} = 0$ if $i > j$
- A square matrix $A = [a_{ij}]$ is lower triangular iff all entries to the right of the main diagonal are zero. i.e. $a_{ij} = 0$ if $i < j$
- A square matrix $A = [a_{ij}]$ is upper triangular iff the i^{th} row starts with at least $i - 1$ zeros for every i.
- A square matrix $A = [a_{ij}]$ is lower triangular iff the j^{th} column starts with at least $j - 1$ zeros for every j.

PRACTICE:

1. Classify the matrix as upper triangular, lower triangular, or diagonal, and decide by inspection whether the matrix is invertible. (Recall that: diagonal matrix is both upper and lower triangular, so there may be more than one answer in some parts.)

i. $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$

ii. $\begin{bmatrix} 0 & 0 \\ 4 & 0 \end{bmatrix}$

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

iv. $\begin{bmatrix} 3 & -2 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 8 \end{bmatrix}$

v. $\begin{bmatrix} 4 & 0 \\ 1 & 7 \end{bmatrix}$

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

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

viii. $\begin{bmatrix} 3 & 0 & 0 \\ 3 & 1 & 0 \\ 7 & 0 & 0 \end{bmatrix}$

2.. Find the product by inspection.

i. $\begin{bmatrix}3 & 0 & 0\\0 & -1 & 0\\0 & 0 & 2\end{bmatrix}\begin{bmatrix}2 & 1\\-4 & 1\\2 & 5\end{bmatrix}$

ii. $\begin{bmatrix}1 & 2 & -5\\-3 & -1 & 0\end{bmatrix}\begin{bmatrix}-4 & 0 & 0\\0 & 3 & 0\\0 & 0 & 2\end{bmatrix}$

iii. $\begin{bmatrix}5 & 0 & 0\\0 & 0 & 0\\0 & 0 & -3\end{bmatrix}\begin{bmatrix}-3 & 2 & 0 & 4 & -4\\1 & -5 & 3 & 0 & 3\\-6 & 2 & 2 & 2 & 2\end{bmatrix}$

iv. $\begin{bmatrix}2 & 0 & 0\\0 & -1 & 0\\0 & 0 & 4\end{bmatrix}\begin{bmatrix}4 & -1 & 3\\1 & 2 & 0\\-5 & 1 & -2\end{bmatrix}\begin{bmatrix}-3 & 0 & 0\\0 & 5 & 0\\0 & 0 & 2\end{bmatrix}$

3.. Find A^2, A^{-2} and A^{-k} (where 'k' is any integer) by inspection.

i. $\begin{bmatrix}1 & 0\\0 & -2\end{bmatrix}$

ii. $\begin{bmatrix}-6 & 0 & 0\\0 & 3 & 0\\0 & 0 & 5\end{bmatrix}$

iii. $\begin{bmatrix}\frac{1}{2} & 0 & 0\\0 & \frac{1}{3} & 0\\0 & 0 & \frac{1}{4}\end{bmatrix}$

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

4.. Compute the product by inspection.

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

ii. $\begin{bmatrix}-1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 4\end{bmatrix}\begin{bmatrix}3 & 0 & 0\\0 & 5 & 0\\0 & 0 & 7\end{bmatrix}\begin{bmatrix}5 & 0 & 0\\0 & -2 & 0\\0 & 0 & 3\end{bmatrix}$

5..Compute the indicated quantity.

i. $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}^{39}$

ii. $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}^{1000}$

6.. Multiplying by diagonal matrices compute the product by inspection.

i. $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} u & v \\ w & x \\ y & z \end{bmatrix}$

ii. $\begin{bmatrix} r & s & t \\ u & v & w \\ x & y & z \end{bmatrix}\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$

iii. $\begin{bmatrix} u & v \\ w & x \\ y & z \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

iv. $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} r & s & t \\ u & v & w \\ x & y & z \end{bmatrix}$

7... Determine by inspection whether the matrix is invertible.

i. $\begin{bmatrix} 0 & 6 & -1 \\ 0 & 7 & -4 \\ 0 & 0 & -2 \end{bmatrix}$

ii. $\begin{bmatrix} -1 & 2 & 4 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

iii. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -5 & 0 & 0 \\ 4 & -3 & 4 & 0 \\ 1 & -2 & 1 & 3 \end{bmatrix}$

iv. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ -3 & -1 & 0 & 0 \\ -4 & -6 & 0 & 0 \\ 0 & 3 & 8 & -5 \end{bmatrix}$

8… Find the diagonal entries of AB by inspection.

i. $A = \begin{bmatrix} 3 & 2 & 6 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 2 & 7 \\ 0 & 5 & 3 \\ 0 & 0 & 6 \end{bmatrix}$

ii. $A = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 0 & 0 \\ -3 & 0 & 7 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 0 & 0 \\ 1 & 5 & 0 \\ 3 & 2 & 6 \end{bmatrix}$

Symmetric Matrices:

A matrix A is symmetric if $A^T = A$. Equivalently, $A = [a_{ij}]$ is symmetric if symmetric elements (mirror elements with respect to the diagonal) are equal, that is, if each $a_{ij} = a_{ji}$

Examples: $\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix}$, $\begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix}$, $\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & d_n \end{bmatrix}$

Skew Symmetric Matrices:

A matrix A is skew-symmetric if $A^T = -A$ or, equivalently, if each $a_{ij} = -a_{ji}$

Clearly, the diagonal elements of such a matrix must be zero, because $a_{ii} = -a_{ii}$ implies $a_{ii} = 0$

(Note that a matrix A must be square if $A^T = A$ or $A^T = -A$)

Remark:

- The product of two symmetric matrices is symmetric iff the matrices commute.
- If A is an invertible symmetric matrix, then A^{-1} is symmetric.
- If A is an invertible, then AA^T and A^TA are also invertible.

Theorem:

If A and B are symmetric matrices with the same size, and if 'k' is any scalar, then

a) A^T is symmetric.
b) $A + B$ and $A - B$ are symmetric.
c) kA is symmetric.

PRACTICE:

1) Find all values of unknown constant(s) for which A is symmetric.

i. $A = \begin{bmatrix} 4 & -3 \\ \mathrm{a}+5 & -1 \end{bmatrix}$ ii. $A = \begin{bmatrix} 4 & \mathrm{x}+2 \\ 2\mathrm{x}-3 & \mathrm{x}+1 \end{bmatrix}$

iii. $A = \begin{bmatrix} 2 & a-2b+2c & 2a+b+c \\ 3 & 5 & a+c \\ 0 & -2 & 7 \end{bmatrix}$

2) Find all values of 'x' for which A is invertible.

i. $A = \begin{bmatrix} x-1 & x^2 & x^4 \\ 0 & x+2 & x^3 \\ 0 & 0 & x-4 \end{bmatrix}$

ii. $A = \begin{bmatrix} x-\frac{1}{2} & 0 & 0 \\ x & x-\frac{1}{3} & 0 \\ x^2 & x^3 & x+\frac{1}{4} \end{bmatrix}$

3) Find a diagonal matrix that satisfies the given conditions.

i. $A^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

ii. $A^{-2} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4) Let A be $n \times n$ symmetric matrix, then
   i. Show that A^2 is symmetric.
   ii. Show that $2A^2 - 3A + I$ is symmetric.

5) Find an upper triangular matrix that satisfies $A^3 = \begin{bmatrix} 1 & 30 \\ 0 & -8 \end{bmatrix}$

6) Find all values of a,b,c and d for which A is skew symmetric.

$$A = \begin{bmatrix} 0 & 2a-3b+c & 3a-5b+5c \\ -2 & 0 & 5a-8b+6c \\ -3 & -5 & d \end{bmatrix}$$

Orthogonal Matrices:

A real matrix A is orthogonal if $A^T = A^{-1}$; that is, if $AA^T = A^TA = I$. Thus, A must necessarily be square and invertible.

Examples: $\begin{bmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{bmatrix}$

Remark (discussed later):

Vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3 \dots\dots\dots, \vec{u}_n$ in R^n are said to form an orthonormal set of vectors if the vectors are unit vectors and are orthogonal to each other. i.e.

$\vec{u}_i . \vec{u}_j = \delta_{ij} = \begin{cases} 0 & if \;\; i \neq j \\ 1 & if \;\; i = j \end{cases}$ where δ_{ij} is known as Kronecker delta function.

Theorem: Let A be a real matrix then following are equivalent;

a) A is orthogonal.
b) The rows of A form an orthonormal set.
c) The columns of A form an orthonormal set

Normal Matrices:

A real matrix A is normal if it commutes with its transpose A^T—that is, if

$AA^T = A^TA$ .If A is symmetric, orthogonal, or skew-symmetric, then A is normal. There are also other normal matrices.

Examples: let $A = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix}$

$\Rightarrow AA^T = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix}$

$\Rightarrow A^TA = \begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix}$

$\Rightarrow AA^T = A^TA$

COMPLEX MATRICES

A matrix with complex entries is called a complex matrix. Since we know that $z = a + ib$ is a complex number then $\bar{z} = \overline{a + ib} = a - ib$ is its conjugate. Then the conjugate of a complex matrix A is written as $\bar{A}$ is the matrix obtain from A by taking the conjugate of each entry in A. i.e. if $A = [a_{ij}]$ then $\bar{A} = [\bar{a}_{ij}]$

Remark:

- The two operations of transpose and conjugation commute for any complex matrix.
- The special notation A^H is used for the conjugate transpose of A. i.e. $A^H = (\bar{A})^T = \overline{A^T}$
- If A is real then $A^H = A^T$ (some author use A^* instead of A^H)

Examples:

let $A = \begin{bmatrix} 2+8i & 5-3i & 4-7i \\ 6i & 1-4i & 3+2i \end{bmatrix}$ then $A^H = \begin{bmatrix} 2-8i & -6i \\ 5+3i & 1+4i \\ 4+7i & 3-2i \end{bmatrix}$

Hermitian Matrix:

A complex matrix A is said to be Hermitian if $A^H = A$

Remember:

$A = [a_{ij}]$ is Hermitian iff symmetric elements are conjugate. i.e. if each $a_{ij} = \bar{a}_{ij}$, in which case each diagonal element a_{ii} must be real.

Examples: let $A = \begin{bmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 5 \end{bmatrix}$

Clearly the diagonal elements of A are real and the symmetric elements $1-2i$ and $1+2i$, $4-7i$ and $4+7i$, $-2i$ and $2i$ are conjugate. Thus A is Hermitian.

Skew Hermitian Matrix:

A complex matrix A is said to be Skew Hermitian if $A^H = -A$

Unitary Matrix:

A complex matrix A is said to be Hermitian if $A^H A^{-1} = A^{-1} A^H = I$ i.e. $A^H = A^{-1}$

Examples: let $A = \frac{1}{2}\begin{bmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{bmatrix}$

Clearly $A^H A^{-1} = A^{-1} A^H = I$ i.e. $A^H = A^{-1}$ this yields A is Unitary matrix.

Normal Matrices:

A real matrix A is normal if it commutes with its transpose A^H—that is, if

$AA^H = A^H A$.

Examples: let $A = \begin{bmatrix} 2+3i & 1 \\ i & 1+2i \end{bmatrix}$

Clearly $AA^H = A^H A$ this yields A is Normal matrix.

Note: When a matrix A is real, Hermitian is the same as Symmetric and Unitary is the same as Orthogonal.

PRACTICE:

1) Find A^H where (a) $A = \begin{bmatrix} 3-5i & 2+4i \\ 6+7i & 1+8i \end{bmatrix}$ (b) $A = \begin{bmatrix} 2-3i & 5+8i \\ -4 & 3-7i \\ -6-i & 5i \end{bmatrix}$
2) Show that $A = \begin{bmatrix} \frac{1}{3}-\frac{2}{3}i & \frac{2}{3}i \\ -\frac{2}{3}i & -\frac{1}{3}-\frac{2}{3}i \end{bmatrix}$ is unitary.
3) Determine which of the following matrices are normal;
   (a) $A = \begin{bmatrix} 3+4i & 1 \\ i & 2+3i \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 0 \\ 1-i & i \end{bmatrix}$

BLOCK MATRICES

Using a system of horizontal and vertical (dashed) lines, we can partition a matrix A into submatrices called blocks (or cells) of A. Clearly a given matrix may be divided into blocks in different ways. For example,

$$\left[\begin{array}{cc:cc:c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hdashline 3 & 1 & 4 & 5 & 9 \\ 4 & 6 & -3 & 1 & 8 \end{array}\right], \left[\begin{array}{cc:ccc} 1 & -2 & 0 & 1 & 3 \\ \hdashline 2 & 3 & 5 & 7 & -2 \\ 3 & 1 & 4 & 5 & 9 \\ \hdashline 4 & 6 & -3 & 1 & 8 \end{array}\right], \left[\begin{array}{ccc:cc} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hdashline 3 & 1 & 4 & 5 & 9 \\ 4 & 6 & -3 & 1 & 8 \end{array}\right]$$

The convenience of the partition of matrices, say A and B, into blocks is that the result of operations on A and B can be obtained by carrying out the computation with the blocks, just as if they were the actual elements of the matrices. This is illustrated below, where the notation $A = [A_{ij}]$ will be used for a block matrix A with blocks A_{ij}

Suppose that $A = [A_{ij}]$ and $B = [B_{ij}]$ are block matrices with the same numbers of row and column blocks, and suppose that corresponding blocks have the same size. Then adding the corresponding blocks of A and B also adds the corresponding elements of A and B, and multiplying each block of A by a scalar 'k' multiplies each element of A by 'k'. Thus,

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \cdots & A_{1n} + B_{1n} \\ A_{21} + B_{21} & A_{22} + B_{22} & \cdots & A_{2n} + B_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} + B_{m1} & A_{m2} + B_{m2} & \cdots & A_{mn} + B_{mn} \end{bmatrix}$$

And $$kA = \begin{bmatrix} kA_{11} & kA_{12} & \cdots & kA_{1n} \\ kA_{21} & kA_{22} & \cdots & kA_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ kA_{m1} & kA_{m2} & \cdots & kA_{mn} \end{bmatrix}$$

The case of matrix multiplication is less obvious, but still true. That is, suppose that $U = [U_{ik}]$ and $V = [V_{kj}]$ are block matrices such that the number of columns of each block U_{ik} is equal to the number of rows of each block V_{kj}

(Thus, each product $U_{ik}V_{kj}$ is defined.)

Then

$$UV = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ W_{m1} & W_{m2} & \cdots & W_{mn} \end{bmatrix} \text{ where } \quad W_{ij} = U_{i1}V_{1j} + U_{i2}V_{2j} + \cdots + U_{ip}V_{pj}$$

Square Block Matrices

Let M be a block matrix. Then M is called a square block matrix if

(i) M is a square matrix.

(ii) The blocks form a square matrix.

(iii) The diagonal blocks are also square matrices.

The latter two conditions will occur if and only if there are the same numbers of horizontal and vertical lines and they are placed symmetrically.

Consider the following two block matrices:

$$A = \left[\begin{array}{cc:cc:c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ \hdashline 4 & 4 & 4 & 4 & 4 \\ 3 & 5 & 3 & 5 & 3 \end{array}\right] \quad , \quad B = \left[\begin{array}{cc:cc:c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ \hdashline 3 & 5 & 3 & 5 & 3 \end{array}\right]$$

The block matrix A is not a square block matrix, because the second and third diagonal blocks are not square. On the other hand, the block matrix B is a square block matrix.

Block Diagonal Matrices

Let $M = [A_{ij}]$ be a square block matrix such that the non-diagonal blocks are all zero matrices; that is, $A_{ij} = 0$ when $i \neq j$. Then M is called a block diagonal matrix. We sometimes denote such a block diagonal matrix by writing

$$M = \, diag\,(A_{11}, A_{22}, \dots, A_{rr}) \quad \text{or} \quad M = A_{11} \oplus A_{22} \oplus \dots \oplus A_{rr}$$

The importance of block diagonal matrices is that the algebra of the block matrix is frequently reduced to the algebra of the individual blocks. Specifically, suppose $f(x)$ is a polynomial and M is the above block diagonal matrix. Then $f(M)$ is a block diagonal matrix, and $f(M) = diag\left(f(A_{11}), f(A_{22}), \dots, f(A_{rr})\right)$

Also, M is invertible if and only if each A_{ii} is invertible, and, in such a case, M^{-1} is a block diagonal matrix, and $M^{-1} = diag\,(A_{11}^{-1}, A_{22}^{-1}, \dots, A_{rr}^{-1})$

Analogously, <u>a square block matrix is called a block upper triangular matrix if the blocks below the diagonal are zero matrices and a block lower triangular matrix if the blocks above the diagonal are zero matrices</u>.

Example: Determine which of the following square block matrices are upper diagonal, lower diagonal, or diagonal:

$$A = \left[\begin{array}{cc:c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hdashline 0 & 0 & 6 \end{array}\right] \qquad B = \left[\begin{array}{c:cc:c} 1 & 0 & 0 & 0 \\ \hdashline 2 & 3 & 4 & 0 \\ 5 & 0 & 6 & 0 \\ \hdashline 0 & 7 & 8 & 9 \end{array}\right] \qquad C = \left[\begin{array}{c:cc} 1 & 0 & 0 \\ \hdashline 0 & 2 & 3 \\ 0 & 4 & 5 \end{array}\right]$$

$$D = \left[\begin{array}{cc:c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hdashline 0 & 6 & 7 \end{array}\right]$$

(a) A is upper triangular because the block below the diagonal is a zero block.

(b) B is lower triangular because all blocks above the diagonal are zero blocks.

(c) C is diagonal because the blocks above and below the diagonal are zero blocks.

(d) D is neither upper triangular nor lower triangular. Also, no other partitioning of D will make it into either a block upper triangular matrix or a block lower triangular matrix.

Periodic Matrix: A square matrix A is said to be periodic matrix of period k, where k is the least positive integer such that $A^{k+1} = A$

Idempotent Matrix: A square matrix A is said to be idempotent matrix if $A^2 = A$

Nilpotent Matrix: A square matrix A is said to be Nilpotent matrix of index k, where k is the least positive integer such that $A^k = O$

Involutory: A square matrix A is said to be Involutory matrix if $A^2 = I$

Rank of a Matrix: The rank of a matrix A, written rank A , is equal to the number of pivots in an echelon form of A. It is equal to the number of non – zero rows in its echelon form.

The rank is a very important property of a matrix and, depending on the context in which the matrix is used; it will be defined in many different ways. Of course, all the definitions lead to the same number.

Example 1: *Find the rank of* $A=\begin{bmatrix}1&2&3&1&2\\2&1&2&3&1\\3&3&5&4&3\\1&-1&-1&2&-1\end{bmatrix}$.

Solution: $A=\begin{bmatrix}1&2&3&1&2\\2&1&2&3&1\\3&3&5&4&3\\1&-1&-1&2&-1\end{bmatrix}\sim\begin{bmatrix}1&2&3&1&2\\0&-3&-4&1&-3\\0&-3&-4&1&-3\\0&-3&-4&1&-3\end{bmatrix}$, R_4-R_1, R_3-3R_1, R_2-2R_1

$$\underset{\sim}{R_4-R_2,\ R_3-R_2}\begin{bmatrix}1&2&3&1&2\\0&-3&-4&1&-3\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}\underset{\sim}{-\frac{1}{3}R_2}\begin{bmatrix}1&2&3&1&2\\0&1&\frac{4}{3}&-\frac{1}{3}&1\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$$

This is echelon form of matrix A and the number of its nonzero rows is 2, therefore, the rank of A is 2.

Canonical Form of a Matrix

If A is a nonzero $m \times n$ matrix then it is equivalent to an $m \times n$ matrix D, where

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

This matrix D is called the ***canonical form*** or ***normal form*** of the matrix A.

In reducing a matrix A to canonical form we shall consider two identity matrices, then each ow operation performed on A is also applied on the first identity matrix, and each column operation perfo d on A is also applied on the second identity matrix. When we get the canonical form of the given matrix A as a result of these row and column operations, the two identity matrices reduce to matrices P and Q. The matrices P and Q are non-singular and satisfy the condition $PAQ = D$. In the following example we show that non-singular matrices P and Q so obtained are note unique.

Example 1: *Reduce the matrix* $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$ *to canonical form. Also find the non-singular matrices P and Q such that PAQ is in the canonical for.*

Solution:

$$\begin{array}{cccc}
A & \text{Row operations} & \text{Column operations} & \\
\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \\
\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & R_2 - R_1 \\
\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & R_3 + R_2 \\
\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & R_1 - R_2 \\
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} & C_3 - (C_1 + C_2) \\
D & P & Q &
\end{array}$$

$$\Rightarrow D=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}, P=\begin{bmatrix}2&-1&0\\-1&1&0\\-1&1&1\end{bmatrix}, Q=\begin{bmatrix}1&0&-1\\0&1&-1\\0&0&1\end{bmatrix}$$

where D is the canonical form of A. We can also verify that

$$PAQ=\begin{bmatrix}2&-1&0\\-1&1&0\\-1&1&1\end{bmatrix}\begin{bmatrix}1&1&2\\1&2&3\\0&-1&-1\end{bmatrix}\begin{bmatrix}1&0&-1\\0&1&-1\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=D$$

Once again we reduce the matrix A to canonical form using different operations.

A	Row operations	Column operations
$\begin{bmatrix}1&1&2\\1&2&3\\0&-1&-1\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$
$\begin{bmatrix}1&1&2\\0&1&1\\0&-1&-1\end{bmatrix}$	$\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$, R_2-R_1
$\begin{bmatrix}1&0&0\\0&1&1\\0&-1&-1\end{bmatrix}$	$\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}$	$\begin{bmatrix}1&-1&-2\\0&1&0\\0&0&1\end{bmatrix}$, C_2-C_1, C_3-2C_1
$\begin{bmatrix}1&0&0\\0&1&0\\0&-1&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}$	$\begin{bmatrix}1&-1&-1\\0&1&-1\\0&0&1\end{bmatrix}$, C_3-C_2
$\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\-1&1&0\\-1&1&1\end{bmatrix}$	$\begin{bmatrix}1&-1&-1\\0&1&-1\\0&0&1\end{bmatrix}$, R_3+R_2
D	P_1	Q_1

$$\Rightarrow D=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}, P_1=\begin{bmatrix}1&0&0\\-1&1&0\\-1&1&1\end{bmatrix}, Q_1=\begin{bmatrix}1&-1&-1\\0&1&-1\\0&0&1\end{bmatrix}$$

where D is the canonical form of A. We can also verify that

$$P_1AQ_1=\begin{bmatrix}1&0&0\\-1&1&0\\-1&1&1\end{bmatrix}\begin{bmatrix}1&1&2\\1&2&3\\0&-1&-1\end{bmatrix}\begin{bmatrix}1&-1&-1\\0&1&-1\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=D$$

This example shows that the non-singular matrices are not unique.

Example 2: *Reduce the matrix* $A=\begin{bmatrix}1 & -1 & 3\\ 2 & -4 & 1\\ 0 & 3 & 2\end{bmatrix}$ *to canonical form. Also find the non-singular matrices P and Q such that PAQ is in the canonical for.*

Solution:

A	Row operations	Column operations	
$\begin{bmatrix}1 & -1 & 3\\ 2 & -4 & 1\\ 0 & 3 & 2\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	
$\begin{bmatrix}1 & -1 & 3\\ 0 & -2 & -5\\ 0 & 3 & 2\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$, R_2-2R_1$
$\begin{bmatrix}1 & -1 & 3\\ 0 & 1 & -3\\ 0 & 3 & 2\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$, R_2+R_3$
$\begin{bmatrix}1 & -1 & 3\\ 0 & 1 & -3\\ 0 & 0 & 11\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ 6 & -3 & -2\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$, R_3-3R_2$
$\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 11\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ 6 & -3 & -2\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$	$, C_3+3C_2$
$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 11\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ 6 & -3 & -2\end{bmatrix}$	$\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$	$, C_2+C_1$
$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ \frac{6}{11} & -\frac{3}{11} & -\frac{2}{11}\end{bmatrix}$	$\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$	$, \frac{1}{11}R_3$
D	P	Q	

$$\Rightarrow D=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, P=\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ \frac{6}{11} & -\frac{3}{11} & -\frac{2}{11}\end{bmatrix}, Q=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$$

where D is the canonical form of A. We can also verify that

$$PAQ=\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 1\\ \frac{6}{11} & -\frac{3}{11} & -\frac{2}{11}\end{bmatrix}\begin{bmatrix}1 & -1 & 3\\ 2 & -4 & 1\\ 0 & 3 & 2\end{bmatrix}\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=D$$

CHAPTER # 3

VECTOR SPACES

This chapter introduces the underlying structure of linear algebra that of a finite dimensional vector space. Where vector space is a collection of objects, called vectors, which may be added together and multiplied by numbers, called scalars.

Vectors: Many physical quantities, such as temperature and speed, possess only ‘‘magnitude.’’ These quantities can be represented by real numbers and are called scalars. On the other hand, there are also quantities, such as force and velocity, that possess both ‘‘magnitude’’ and ‘‘direction.’’ These quantities, which can be represented by arrows having appropriate lengths and directions and emanating (originating) from some given reference point O, are called vectors. The tail of the arrow is called **initial point** of the vector and the tip the **terminal point** of the vector.

Remark: Mathematically, we identify the vector $\vec{v}$ with its (a, b, c) and write $\vec{v} = (a, b, c)$ Moreover, we call the ordered triple (a, b, c) of real numbers a point or vector depending upon its interpretation. We generalize this notion and call an n- tuple $(v_1, v_2, \dots, v_n)$ of real numbers a vector. However, special notation may be used for the vectors in R^3 called spatial vectors.

Vectors in R^n: The set of all n-tuples of real numbers, denoted by R^n is called n-space. A particular n-tuple in R^n say $\vec{v} = (v_1, v_2, \dots, v_n)$ is called a point or vector. The numbers v_i are called the coordinates, components, entries or elements of $\vec{v}$. Moreover, when discussing the space R^n

- we use the term **scalar** for the elements of R.
- Two vectors, **u** and **v**, are **equal**, written **u** = **v**, if they have the same number of components and if the corresponding components are equal. Although the vectors (1,2,3) and (2,3,1) contain the same three numbers, these vectors are not equal because corresponding entries are not equal.
- The vector $(0,0,0, \dots..0)$ whose entries are all 0 is called the **zero vector** and is usually denoted by $\vec{0}$ or **0**.

VECTOR ADDITION AND SCALAR MULTIPLICATION

Vector addition: Consider two vectors **u** and **v** in R^n, say $\vec{u} = (u_1, u_2, \dots, u_n)$ and $\vec{v} = (v_1, v_2, \dots, v_n)$ then Their sum written $\vec{u} + \vec{v}$, is the vector obtained by adding corresponding components from **u** and **v**. That is, $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

Scalar Product: The scalar product or, simply, product, of the vector **v** by a real number k, written k**v**, is the vector obtained by multiplying each component of **v** by k. That is, $k\vec{v} = (kv_1, kv_2, \dots, kv_n)$

- Observe that **u** + **v** and k**v** are also vectors in R^n
- The sum of vectors with different numbers of components is not defined.
- **Negatives and subtraction** are defined in R^n as follows:
  $-\vec{v} = (-1)\vec{v}$ and $\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$ The vector –**v** is called the **negative** of **v**, and **u** – **v** is called the difference of **u** and **v**.
- **Linear combination of the vectors** : suppose we are given vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ in R^n and scalars $k_1, k_2, \dots, k_n$ in R. We can multiply the vectors by the corresponding scalars and then add the resultant scalar products to form the vector

  $$\vec{v} = k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_n\vec{v}_n$$

Such a vector **v** is called a **linear combination** of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$

Vectors in R^3 (Spatial Vectors), $\boldsymbol{ijk}$ Notation

Vectors in R^3, called spatial vectors, appear in many applications, especially in physics. In fact, a special notation is frequently used for such vectors as follows:

$\mathbf{i} = (1, 0, 0)$ denotes the unit vector in the 'x' direction:

$\mathbf{j} = (0, 1, 0)$ denotes the unit vector in the 'y' direction:

$\mathbf{k} = (0, 0, 1)$denotes the unit vector in the 'z' direction:

Then any vector $\vec{v} = (a, b, c)$ in R^3 can be expressed uniquely in the form

$$\vec{v} = (a, b, c) = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}$$

n – Space: If ‘n’ is a positive integer, then an **ordered n – tuple** is a sequence of ‘n’ real numbers$(v_1, v_2, \dots, v_n)$. The set of all ordered n – tuples is called **n – space** and is denoted by $\mathbf{R}^n$

Finding the components of vectors:

If a vector in 2 – space or 3 – space is positioned with its initial point at the origin of a rectangular coordinate system, then the vector is completely determined by the coordinates of its terminal point. We call these coordinates the **components** of vector **v** relative to the coordinate system.

If $\boldsymbol{v} = \overrightarrow{P_1P_2}$ denote the vector with initial point $P_1(x_1, x_2, x_3)$ and terminal point $P_2(y_1, y_2, y_3)$ then the components of the vector are given by the formula $\boldsymbol{v} = \overrightarrow{P_1P_2} = (y_1 - x_1, y_2 - x_2, y_3 - x_3)$

Example: The components of the vector $\boldsymbol{v} = \overrightarrow{P_1P_2}$ with initial point $P_1(2, -1, 4)$ and terminal point $P_2(7, 5, -8)$ are

$\boldsymbol{v} = \overrightarrow{P_1P_2} = (7 - 2, 5 - (-1), -8 - 4) = (5, 6, -12)$

PRACTICE:

1) Find the components of the vector $\boldsymbol{v} = \overrightarrow{P_1P_2}$
   i. $P_1(3,5)$; $P_2(2,8)$
   ii. $P_1(5, -2, 1)$; $P_2(2, 4, 2)$
   iii. $P_1(-6, 2)$; $P_2(-4, -1)$
   iv. $P_1(0,0,0)$; $P_2(-1, 6, 1)$

2) Let $\vec{u} = (-4, 1)$, $\vec{v} = (0, 5)$ and $\vec{w} = (-3, -3)$ then find the components of
   i. $\vec{u} + \vec{w}$
   ii. $\vec{v} - 3\vec{u}$
   iii. $2(\vec{u} - 5\vec{w})$
   iv. $3\vec{v} - 2(\vec{u} + 2\vec{w})$
3) Let $\vec{u} = (1, -1, 3, 5)$, $\vec{v} = (2, 1, 0, -3)$ find scalar ‘a’ and ‘b’ so that $a\vec{u} + b\vec{v} = (1, -4, 9, 18)$

Vector Space:

Let V be a nonempty set with two operations:

(i) **Vector Addition**: This assigns to any $\vec{u},\vec{v}\epsilon\boldsymbol{V}$ a sum $\vec{u}+\vec{v}\ \epsilon\ \boldsymbol{V}$.

(ii) **Scalar Multiplication**: This assigns to any $\vec{v}\epsilon\boldsymbol{V}$, $k\epsilon\boldsymbol{K}$ a product $k\vec{v}\ \epsilon\boldsymbol{V}$.

Then **V** is called a vector space (over the field **K**) if the following axioms hold for any vectors $\vec{u},\vec{v},\vec{w}\epsilon\boldsymbol{V}$:

- $\vec{u}+\vec{v}=\vec{v}+\vec{u}$
- $\vec{u}+(\vec{v}+\vec{w})=(\vec{u}+\vec{v})+\vec{w}$
- There is a vector in V, denoted by $\vec{0}$ and called the zero vector, such that, for any $\vec{v}\epsilon\boldsymbol{V}$ **;** $\vec{v}+\vec{0}=\vec{0}+\vec{v}=\vec{v}$
- For each $\vec{v}\epsilon\boldsymbol{V}$; there is a vector in **V**, denoted by $-\vec{v}$, and called the negative of $\vec{u}$, such that $\vec{v}+(-\vec{v})=(-\vec{v})+\vec{v}=\vec{0}$
- If ‘k’ is any scalar and $\vec{v}$ is in **V** then $k\vec{v}$ is in **V**.
- $k(\vec{u}+\vec{v})=k\vec{u}+k\vec{v}$ for $\vec{u},\vec{v}\epsilon\boldsymbol{V}$ and $k\epsilon\boldsymbol{K}$
- $(k+m)\vec{v}=k\vec{v}+m\vec{v}$ for $\vec{v}\epsilon\boldsymbol{V}$ and $k,m\ \epsilon\boldsymbol{K}$
- $k(m\vec{v})=(km)\vec{v}$ for $\vec{v}\epsilon\boldsymbol{V}$ and $k,m\ \epsilon\boldsymbol{K}$
- $1(\vec{v})=\vec{v}$, for the unit scalar 1 in K.

The above axioms naturally split into two sets (as indicated by the labeling of the axioms). The first four are concerned only with the additive structure of **V** and can be summarized by saying **V** is a **commutative group (Abelian group)** under addition. This means

- Any sum $v_1,v_2,\dots,v_n$ of vectors requires no parentheses and does not depend on the order of the summands.
- The zero vector $\vec{0}$ is unique, and the negative $-\vec{v}$ of a vector $\vec{v}$ is unique.
- (Cancellation Law) If $\vec{u}+\vec{w}=\vec{v}+\vec{w}$, then $\vec{u}=\vec{v}$

Also, subtraction in **V** is defined by $\vec{v}-\vec{v}=\vec{v}+(-\vec{v})$, where $-\vec{v}$ is the unique negative of $\vec{v}$. On the other hand, the remaining four axioms are concerned with the ‘‘action’’ of the field K of scalars on the vector space **V**.

To Show that a Set with two operation is a Vector Space

- Identify the set **V** of objects that will become vectors.
- Identify the addition and scalar multiplication operation on **V**.
- Verify remaining axioms.

EXAMPLES OF VECTOR SPACES

This section lists important examples of vector spaces that will be used throughout the text.

The Zero Vector Space:

Let $\vec{0}\,\epsilon\boldsymbol{V}$ and define $\vec{0}+\vec{0}=\vec{0}$ also $k\vec{0}=\vec{0}$ for all scalars 'k' then given space **V** will a be vector space and called **zero vector space**.

Space $\mathbf{K}^{\mathbf{n}}$

Let **K** be an arbitrary field. The notation $\mathbf{K}^{\mathbf{n}}$ is frequently used to denote the set of all n-tuples of elements in **K**. Here $\mathbf{K}^{\mathbf{n}}$ is a vector space over **K** using the following operations:

(i) **Vector Addition:**

$$(u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$$

(ii) **Scalar Multiplication:**

$$k(v_1, v_2, \dots, v_n) = (kv_1, kv_2, \dots, kv_n)$$

The zero vector in $\mathbf{K}^{\mathbf{n}}$ is the n-tuple of zeros, $\vec{0} = (0,0,, \dots.,0)$ and the negative of a vector is defined by $-(v_1, v_2, \dots, v_n) = (-v_1, -v_2, \dots, -v_n)$

Theorem 1: *Every field is a vector space over itself.*

Proof: Let F be any field, then for $\alpha, \beta, u, v, w \in F$

1. $u + v \in F$
2. $u + (v + w) = (u + v) + w$
3. There exists $0 \in F$ such that $u + 0 = u = 0 + u$
4. There exists $-u \in F$ such that $u + (-u) = 0 = (-u) + u$
5. $u + v = v + u$
6. $\alpha u \in F$
7. $\alpha(\beta u) = (\alpha\beta)u$
8. $1u = u$
9. $\alpha(u + v) = \alpha u + \alpha v$
10. $(\alpha + \beta)u = \alpha u + \beta u$

This shows that F is a vector space over F, so every field is a vector space over itself.

Example 1: *Show that $V = \{a + b\sqrt{2} : a, b \in Q\}$ is a vector space over Q.*

Solution: Let $\alpha, \beta \in Q$, $u, v, w \in V$, then

$$u = a + b\sqrt{2}, v = a' + b'\sqrt{2}, w = a'' + b''\sqrt{2}, \; a, a', a'', b, b', b'' \in Q$$

1. $u + v = (a + b\sqrt{2}) + (a' + b'\sqrt{2}) = (a + a') + (b + b')\sqrt{2} \in V$
2. $$\begin{aligned} u + (v + w) &= (a + b\sqrt{2}) + \{(a' + b'\sqrt{2}) + (a'' + b''\sqrt{2})\} \\ &= (a + b\sqrt{2}) + \{(a' + a'') + (b' + b'')\sqrt{2}\} \\ &= (a + a' + a'') + (b + b' + b'')\sqrt{2} \\ &= \{(a + a') + (b + b')\sqrt{2}\} + (a'' + b''\sqrt{2}) \\ &= \{(a + b\sqrt{2}) + (a' + b'\sqrt{2})\} + (a'' + b''\sqrt{2}) \\ &= (u + v) + w \end{aligned}$$
3. $0 = 0 + 0\sqrt{2} \in V$ such that $u + 0 = u = 0 + u$
4. $-u = -(a + b\sqrt{2}) = -a - b\sqrt{2} \in V$ such that $u + (-u) = 0 = (-u) + u$
5. $$\begin{aligned} u + v &= (a + b\sqrt{2}) + (a' + b'\sqrt{2}) = (a + a') + (b + b')\sqrt{2} \\ &= (a' + a) + (b' + b)\sqrt{2} = (a' + b'\sqrt{2}) + (a + b\sqrt{2}) = v + u \end{aligned}$$
6. $\alpha u = \alpha(a + b\sqrt{2}) = \alpha a + \alpha b\sqrt{2} \in V$
7. $\alpha(\beta u) = \alpha(\beta(a + b\sqrt{2})) = (\alpha\beta)(a + b\sqrt{2}) = (\alpha\beta)u$
8. $1u = 1(a + b\sqrt{2}) = a + b\sqrt{2} = u$
9. $$\begin{aligned} \alpha(u + v) &= \alpha\{(a + b\sqrt{2}) + (a' + b'\sqrt{2})\} \\ &= \alpha(a + b\sqrt{2}) + \alpha(a' + b'\sqrt{2}) = \alpha u + \alpha v \end{aligned}$$
10. $(\alpha + \beta)u = (\alpha + \beta)(a + b\sqrt{2}) = \alpha(a + b\sqrt{2}) + \beta(a + b\sqrt{2}) = \alpha u + \beta u$

This shows that V is a vector space over Q.

Example 3: *Show that the set of all solutions of the differential equation $y'' - 5y' + 6y = 0$ is a real vector space.*

Solution: Let V be the set of all solutions of the differential equation $y'' - 5y' + 6y = 0$, then

$$V = \{ae^{2X} + be^{3X} : a, b \in R\}$$

Let $\alpha, \beta \in R$, $u, v, w \in V$, then

$$u = ae^{2X} + be^{3X},$$
$$v = a'e^{2X} + b'e^{3X},$$
$$w = a''e^{2X} + b''e^{3X}, \quad a, b, a', b', a'', b'' \in R$$

1. $u + v = (ae^{2X} + be^{3X}) + (a'e^{2X} + b'e^{3X}) = (a + a')e^{2X} + (b + b')e^{3X} \in V$

2. $u + (v + w) = (ae^{2X} + be^{3X}) + \{(a'e^{2X} + b'e^{3X}) + (a''e^{2X} + b''e^{3X})\}$
$= (ae^{2X} + be^{3X}) + \{(a' + a'')e^{2X} + (b' + b'')e^{3X}\}$
$= (a + a' + a'')e^{2X} + (b + b' + b'')e^{3X}$
$= (a + a')e^{2X} + (b + b')e^{3X} + (a''e^{2X} + b''e^{3X})$
$= \{(ae^{2X} + be^{3X}) + (a'e^{2X} + b'e^{3X})\} + (a''e^{2X} + b''e^{3X})$
$= (u + v) + w$

3. $0 = 0e^{2X} + 0e^{3X} \in V$ such that $u + 0 = u = 0 + u$

4 $-u = -(ae^{2X} + be^{3X}) = -ae^{2X} - be^{3X} \in V$
such that $u + (-u) = 0 = (-u) + u$

5 $u + v = (ae^{2X} + be^{3X}) + (a'e^{2X} + b'e^{3X})$
$= (a'e^{2X} + b'e^{3X}) + (ae^{2X} + be^{3X}) = v + u$

6 $\alpha u = \alpha(ae^{2X} + be^{3X}) = \alpha ae^{2X} + \alpha be^{3X} \in V$

7 $\alpha(\beta u) = \alpha(\beta(ae^{2X} + be^{3X})) = (\alpha\beta)(ae^{2X} + be^{3X}) = (\alpha\beta)u$

8 $1u = 1(ae^{2X} + be^{3X}) = (ae^{2X} + be^{3X}) = u$

9 $\alpha(u + v) = \alpha\{(ae^{2X} + be^{3X}) + (a'e^{2X} + b'e^{3X})\}$
$= \alpha(ae^{2X} + be^{3X}) + \alpha(a'e^{2X} + b'e^{3X}) = \alpha u + \alpha v$

10 $(\alpha + \beta)u = (\alpha + \beta)(ae^{2X} + be^{3X})$
$= \alpha(ae^{2X} + be^{3X}) + \beta(ae^{2X} + be^{3X}) = \alpha u + \beta u$

This shows that V is a vector space over R, i.e. V is a real vector space.

Space R^n

We have to show that is a vector space. Since we know that

$V = R^n = \{(a_1, a_2, \dots, a_n): a_1, a_2, \dots, a_n \in R\}$

Let us define addition and scalar multiplication of n – tuples as follows for $\vec{u}, \vec{v} \epsilon \boldsymbol{V}$

$\vec{u} + \vec{v} = (u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

And $k(u_1, u_2, \dots, u_n) = (ku_1, ku_2, \dots, ku_n)$

Firstly we will show that $\boldsymbol{V}$ is an Abelian Group.

Closure Law: that is $\forall \vec{u}, \vec{v} \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v}\ \epsilon \boldsymbol{V}$

Let $\vec{u} = (u_1, u_2, \dots, u_n)$, $\vec{v} = (v_1, v_2, \dots, v_n)$ then

$\vec{u} + \vec{v} = (u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n)$

$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)\ \epsilon \boldsymbol{V}$

⇒ Closure Law holds.

Associative Law: that is $\forall \vec{u}, \vec{v}, \vec{w}\ \epsilon \boldsymbol{V}$ we will have $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

Let $\vec{u} = (u_1, u_2, \dots, u_n)$, $\vec{v} = (v_1, v_2, \dots, v_n)$, $\vec{w} = (w_1, w_2, \dots, w_n)$ then

$\vec{u} + (\vec{v} + \vec{w}) = (u_1, u_2, \dots, u_n) + [(v_1, v_2, \dots, v_n) + (w_1, w_2, \dots, w_n)]$

$\vec{u} + (\vec{v} + \vec{w}) = (u_1, u_2, \dots, u_n) + (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n)$

$\vec{u} + (\vec{v} + \vec{w}) = [u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \dots, u_n + (v_n + w_n)]$

$\vec{u} + (\vec{v} + \vec{w}) = [(u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \dots, (u_n + v_n) + w_n]$

$\vec{u} + (\vec{v} + \vec{w}) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) + (w_1, w_2, \dots, w_n)$

$\vec{u} + (\vec{v} + \vec{w}) = [(u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n)] + (w_1, w_2, \dots, w_n)$

$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

⇒ Association Law holds.

Identity Law: that is $\forall \vec{0}, \vec{v}\, \epsilon \boldsymbol{V}$ we will have $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$

Let $\vec{0} = (0,0, \dots .,0)$, $\vec{v} = (v_1, v_2, \dots , v_n)$ then

$\vec{0} + \vec{v} = (0,0, \dots .,0) + (v_1, v_2, \dots , v_n) = (0 + v_1, 0 + v_2, \dots ,0 + v_n)$

$\vec{0} + \vec{v} = (v_1, v_2, \dots , v_n) = \vec{v}$

$\vec{v} + \vec{0} = (v_1, v_2, \dots , v_n) + (0,0, \dots .,0) = (v_1 + 0, v_2 + 0, \dots , v_n + 0)$

$\vec{v} + \vec{0} = (v_1, v_2, \dots \dots , v_n) = \vec{v}$

$\Rightarrow$ Identity Law holds.

Inverse Law: that is $\forall \vec{v}, -\vec{v}\, \epsilon \boldsymbol{V}$ we will have $\vec{v} + (-\vec{v}) = -\vec{v} + \vec{v} = \vec{0}$

Let $\vec{v} = (v_1, v_2, \dots , v_n)$, $-\vec{v} = (-v_1, -v_2, \dots , -v_n)$ then

$\vec{v} + (-\vec{v}) = (v_1, v_2, \dots , v_n) + (-v_1, -v_2, \dots , -v_n) = (v_1 - v_1, v_2 - v_2 \dots v_n - v_n)$

$\vec{v} + (-\vec{v}) = (0,0, \dots .,0) = \vec{0}$

$-\vec{v} + \vec{v} = (-v_1, -v_2, \dots , -v_n) + (v_1, v_2, \dots , v_n)$

$-\vec{v} + \vec{v} = (-v_1 + v_1, -v_2 + v_2 \dots - v_n + v_n) = (0,0, \dots .,0) = \vec{0}$

$\Rightarrow$ Inverse Law holds.

Commutative Law: that is $\forall \vec{u}, \vec{v}\, \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

Let $\vec{u} = (u_1, u_2, \dots , u_n)$, $\vec{v} = (v_1, v_2, \dots , v_n)$ then

$\vec{u} + \vec{v} = (u_1, u_2, \dots , u_n) + (v_1, v_2, \dots , v_n)$

$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \dots , u_n + v_n) = (v_1 + u_1, v_2 + u_2, \dots , v_n + u_n)$

$\vec{u} + \vec{v} = (v_1, v_2, \dots , v_n) + (u_1, u_2, \dots , u_n) = \vec{v} + \vec{u}$

$\Rightarrow$ Commutative Law holds.

Hence given space is Abelian group under addition.

Now we will show Scalar multiplication properties.

- If 'k' is any scalar and $\vec{v}$ is in **V** then $k\vec{v}$ is in **V**.

Let 'k' is any scalar and $\vec{v} = (v_1, v_2, \dots, v_n)$ then

$k(v_1, v_2, \dots, v_n) = (kv_1, kv_2, \dots, kv_n)\ \epsilon\boldsymbol{V}$

- $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$ for $\vec{u}, \vec{v}\epsilon\boldsymbol{V}$ and $k\epsilon\boldsymbol{K}$
  $k(\vec{u} + \vec{v}) = k[(u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n)]$
  $k(\vec{u} + \vec{v}) = [k(u_1, u_2, \dots, u_n) + k(v_1, v_2, \dots, v_n)] = k\vec{u} + k\vec{v}$
  $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$

- $(k + m)\vec{v} = k\vec{v} + m\vec{v}$ for $\vec{v}\epsilon\boldsymbol{V}$ and $k, m\ \epsilon\boldsymbol{K}$
  $(k + m)\vec{v} = (k + m)(v_1, v_2, \dots, v_n)$
  $(k + m)\vec{v} = k(v_1, v_2, \dots, v_n) + m(v_1, v_2, \dots, v_n)$
  $(k + m)\vec{v} = k\vec{v} + m\vec{v}$

- $k(m\vec{v}) = (km)\vec{v}$ for $\vec{v}\epsilon\boldsymbol{V}$ and $k, m\ \epsilon\boldsymbol{K}$
  $k(m\vec{v}) = k[m(v_1, v_2, \dots, v_n)] = (km)(v_1, v_2, \dots, v_n)$
  $k(m\vec{v}) = (km)\vec{v}$

- $1(\vec{v}) = \vec{v}$, for the unit scalar 1 in K.
  $1(\vec{v}) = 1(v_1, v_2, \dots, v_n) = (1v_1, 1v_2, \dots, 1v_n) = (v_1, v_2, \dots, v_n) = \vec{v}$

Hence above all conditions show that $V = R^n$ is a vector space.

Matrix Space $M_{m\times n}$

The notation $M_{m\times n}$ or simply M; will be used to denote the set of all $m \times n$ matrices with entries in a field **K**. Then $M_{m\times n}$ is a vector space over **K** with respect to the usual operations of matrix addition and scalar multiplication of matrices.

We will prove it as follows; since we know that

$$V = M_{m\times n} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{bmatrix} = [v_{ij}]_{m\times n}$$

Let us define addition and scalar multiplication for $\vec{u}, \vec{v} \epsilon \boldsymbol{V}$

$$\vec{u} + \vec{v} = [u_{ij}]_{m\times n} + [v_{ij}]_{m\times n} = [u_{ij} + v_{ij}]_{m\times n}$$

And $k\vec{v} = k[v_{ij}]_{m\times n} = [kv_{ij}]_{m\times n}$

Firstly we will show that $\boldsymbol{V}$ is an Abelian group.

Closure Law: that is $\forall \vec{u}, \vec{v} \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v}\ \epsilon \boldsymbol{V}$

Let $\vec{u} = [u_{ij}]_{m\times n}$, $\vec{v} = [v_{ij}]_{m\times n}$ then

$$\vec{u} + \vec{v} = [u_{ij}]_{m\times n} + [v_{ij}]_{m\times n} = [u_{ij} + v_{ij}]_{m\times n}\ \epsilon \boldsymbol{V}$$

⇒ Closure Law holds.

Associative Law: that is $\forall \vec{u}, \vec{v}, \vec{w}\ \epsilon \boldsymbol{V}$ we will have $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

Let $\vec{u} = [u_{ij}]_{m\times n}$, $\vec{v} = [v_{ij}]_{m\times n}$, $\vec{w} = [w_{ij}]_{m\times n}$ then

$$\vec{u} + (\vec{v} + \vec{w}) = [u_{ij}]_{m\times n} + \left[[v_{ij}]_{m\times n} + [w_{ij}]_{m\times n}\right]$$

$$\vec{u} + (\vec{v} + \vec{w}) = [u_{ij}]_{m\times n} + [v_{ij} + w_{ij}]_{m\times n} = [u_{ij} + (v_{ij} + w_{ij})]_{m\times n}$$

$$\vec{u} + (\vec{v} + \vec{w}) = [(u_{ij} + v_{ij}) + w_{ij}]_{m\times n} = \left[[u_{ij}]_{m\times n} + [v_{ij}]_{m\times n}\right] + [w_{ij}]_{m\times n}$$

$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ ⇒ Association Law holds.

Identity Law:

that is $\forall \vec{0}, \vec{v}\ \epsilon \boldsymbol{V}$ we will have $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$

Let $\vec{0} = [0_{ij}]_{m\times n}$, $\vec{v} = [v_{ij}]_{m\times n}$ then

$$\vec{0} + \vec{v} = [0_{ij}]_{m\times n} + [v_{ij}]_{m\times n} = [0_{ij} + v_{ij}]_{m\times n} = [v_{ij}]_{m\times n} = \vec{v}$$

$$\vec{v} + \vec{0} = [v_{ij}]_{m\times n} + [0_{ij}]_{m\times n} = [v_{ij} + 0_{ij}]_{m\times n} = [v_{ij}]_{m\times n} = \vec{v}$$

$\Rightarrow$ Identity Law holds.

Inverse Law:

that is $\forall \vec{v}, -\vec{v}\ \epsilon \boldsymbol{V}$ we will have $\vec{v} + (-\vec{v}) = -\vec{v} + \vec{v} = \vec{0}$

Let $\vec{v} = [v_{ij}]_{m\times n}$, $-\vec{v} = [-v_{ij}]_{m\times n}$ then

$$\vec{v} + (-\vec{v}) = [v_{ij}]_{m\times n} + [-v_{ij}]_{m\times n} = [v_{ij} + (-v_{ij})]_{m\times n} = [0_{ij}]_{m\times n} = \vec{0}$$

$$-\vec{v} + \vec{v} = [-v_{ij}]_{m\times n} + [v_{ij}]_{m\times n} = [-v_{ij} + v_{ij}]_{m\times n} = [0_{ij}]_{m\times n} = \vec{0}$$

$\Rightarrow$ Inverse Law holds.

Commutative Law:

that is $\forall \vec{u}, \vec{v}\ \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

Let $\vec{u} = [u_{ij}]_{m\times n}$, $\vec{v} = [v_{ij}]_{m\times n}$ then

$$\vec{u} + \vec{v} = [u_{ij}]_{m\times n} + [v_{ij}]_{m\times n} = [u_{ij} + v_{ij}]_{m\times n} = [v_{ij} + u_{ij}]_{m\times n}$$

$$\vec{u} + \vec{v} = [v_{ij}]_{m\times n} + [u_{ij}]_{m\times n} = \vec{v} + \vec{u}$$

$\Rightarrow$ Commutative Law holds.

Hence given space is Abelian group under addition.

Now we will show Scalar multiplication properties.

- If ‘k’ is any scalar and $\vec{v}$ is in **V** then $k\vec{v}$ is in **V**.

Let ‘k’ is any scalar and $\vec{v} = [v_{ij}]_{m\times n}$ then

$k\vec{v} = k[v_{ij}]_{m\times n} = [kv_{ij}]_{m\times n}\, \epsilon \boldsymbol{V}$

- $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$ for $\vec{u}, \vec{v} \epsilon \boldsymbol{V}$ and $k \epsilon \boldsymbol{K}$

  $k(\vec{u} + \vec{v}) = k\left[[u_{ij}]_{m\times n} + [v_{ij}]_{m\times n}\right]$

  $k(\vec{u} + \vec{v}) = \left[k[u_{ij}]_{m\times n} + k[v_{ij}]_{m\times n}\right] = k\vec{u} + k\vec{v}$

  $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$

- $(k + m)\vec{v} = k\vec{v} + m\vec{v}$ for $\vec{v} \epsilon \boldsymbol{V}$ and $k, m\ \epsilon \boldsymbol{K}$

  $(k + m)\vec{v} = (k + m)[v_{ij}]_{m\times n}$

  $(k + m)\vec{v} = k[v_{ij}]_{m\times n} + m[v_{ij}]_{m\times n}$

  $(k + m)\vec{v} = k\vec{v} + m\vec{v}$

- $k(m\vec{v}) = (km)\vec{v}$ for $\vec{v} \epsilon \boldsymbol{V}$ and $k, m\ \epsilon \boldsymbol{K}$

  $k(m\vec{v}) = k\left[m[v_{ij}]_{m\times n}\right] = (km)[v_{ij}]_{m\times n}$

  $k(m\vec{v}) = (km)\vec{v}$

- $1(\vec{v}) = \vec{v}$, for the unit scalar 1 in K.

  $1(\vec{v}) = 1[v_{ij}]_{m\times n} = [1.v_{ij}]_{m\times n} = [v_{ij}]_{m\times n} = \vec{v}$

Hence above all conditions show that $V = M_{m\times n}$ is a vector space.

The Vector Space of real valued Functions defined on $(-\infty, \infty)$

We will prove it as follows; since we know that

$V = F(-\infty, \infty) = \{f : f \ is \ real \ valued \ function \ on \ (-\infty, \infty)\}$

Let us define addition and scalar multiplication for $\vec{u}, \vec{v} \epsilon \boldsymbol{V}$

$\vec{u} + \vec{v} = (f + g)(x) = f(x) + g(x)$

And $k\vec{v} = (kf)(x) = kf(x)$

Firstly we will show that $\boldsymbol{V}$ is an Abelian Group.

Closure Law: that is $\forall \vec{u}, \vec{v} \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v} \ \epsilon \boldsymbol{V}$

Let $\vec{u} = f$, $\vec{v} = g$ then

$\vec{u} + \vec{v} = (f + g)(x) = f(x) + g(x) \ \epsilon \boldsymbol{V}$

⇒ Closure Law holds.

Associative Law: that is $\forall \vec{u}, \vec{v}, \vec{w} \ \epsilon \boldsymbol{V}$ we will have $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

Let $\vec{u} = f$, $\vec{v} = g$, $\vec{w} = h$ then

$\vec{u} + (\vec{v} + \vec{w}) = f(x) + [g + h](x) = f(x) + g(x) + h(x)$

$\vec{u} + (\vec{v} + \vec{w}) = [f + g](x) + h(x) = (\vec{u} + \vec{v}) + \vec{w}$

$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ ⇒ Association Law holds.

Identity Law:

that is $\forall \vec{0}, \vec{v} \ \epsilon \boldsymbol{V}$ we will have $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$

Let $\vec{0} = 0$, $\vec{v} = f$ then

$\vec{0} + \vec{v} = (0 + f)(x) = 0(x) + f(x) = f(x) = \vec{v}$

$\vec{v} + \vec{0} = (f + 0)(x) = f(x) + 0(x) = f(x) = \vec{v}$

⇒ Identity Law holds.

Inverse Law: that is $\forall \vec{v}, -\vec{v}\ \epsilon \boldsymbol{V}$ we will have $\vec{v} + (-\vec{v}) = -\vec{v} + \vec{v} = \vec{0}$

Let $\vec{v} = f\ , -\vec{v} = -f$ then

$\vec{v} + (-\vec{v}) = \big(f + (-f)\big)(x) = f(x) - f(x) = \vec{0}$

$-\vec{v} + \vec{v} = (-f + f)(x) = -f(x) + f(x) = \vec{0}$ $\Rightarrow$ Inverse Law holds.

Commutative Law: that is $\forall \vec{u}, \vec{v}\ \epsilon \boldsymbol{V}$ we will have $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

Let $\vec{u} = f\ , \vec{v} = g$ then

$\vec{u} + \vec{v} = (f + g)(x) = f(x) + g(x) = g(x) + f(x) = \vec{v} + \vec{u}$

$\vec{u} + \vec{v} = \vec{v} + \vec{u}$ $\Rightarrow$ Commutative Law holds.

Hence given space is Abelian group under addition.

Now we will show Scalar multiplication properties.

- If ‘k’ is any scalar and $\vec{v}$ is in **V** then $k\vec{v}$ is in **V**.

  Let ‘k’ is any scalar and $\vec{v} = f$ then $k\vec{v} = (kf)(x) = kf(x)\ \epsilon \boldsymbol{V}$

- $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$ for $\vec{u}, \vec{v} \epsilon \boldsymbol{V}$ and $k \epsilon \boldsymbol{K}$
  $k(\vec{u} + \vec{v}) = [k(f + g)](x) = (kf + kg)(x) = kf(x) + kg(x)$
  $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$

- $(k + m)\vec{v} = k\vec{v} + m\vec{v}$ for $\vec{v} \epsilon \boldsymbol{V}$ and $k, m\ \epsilon \boldsymbol{K}$
  $(k + m)\vec{v} = [(k + m)f](x) = (kf + mf)(x) = kf(x) + mf(x)$
  $(k + m)\vec{v} = k\vec{v} + m\vec{v}$

- $k(m\vec{v}) = (km)\vec{v}$ for $\vec{v} \epsilon \boldsymbol{V}$ and $k, m\ \epsilon \boldsymbol{K}$
  $k(m\vec{v}) = k[mf](x) = k[mf(x)] = (km)f(x) = (km)f$
  $k(m\vec{v}) = (km)\vec{v}$

- $1(\vec{v}) = \vec{v}$, for the unit scalar 1 in K.
  $1(\vec{v}) = 1f = [1f](x) = 1f(x) = f(x) = f = \vec{v}$
  Hence above all conditions show that $V = F(-\infty, \infty)$ is a vector space.

Polynomial Space $\boldsymbol{P(x)}$

Let $P(x)$ denote the set of all polynomials of the form

$P(x) = a_0 + a_1x + \cdots + a_nx^n \ ; n = 1,2,3, \ldots$

where the coefficients a_i belong to a field **K**. Then $P(x)$ is a vector space over **K** using the following operations:

(i) Vector Addition: Here $p(x) + q(x)$ in $P(x)$ is the usual operation of addition of polynomials.

(ii) Scalar Multiplication: Here $kp(x)$in $P(x)$ is the usual operation of the product of a scalar 'k' and a polynomial $p(x)$.

The zero polynomial $0(x)$ is the zero vector in $P(x)$.

Polynomial Space $\boldsymbol{P_n(x)}$

Let $P_n(x)$denote the set of all polynomials $p(x)$ over a field **K**, where the degree of $p(x)$ is less than or equal to n; that is,

$P(x) = a_0 + a_1x + \cdots + a_sx^s$ where $s \leq n$.

Then $P_n(x)$ is a vector space over **K** with respect to the usual operations of addition of polynomials and of multiplication of a polynomial by a constant (just like the vector space $P(x)$above). We include the zero polynomial 0 as an element of $P_n(x)$, even though its degree is undefined.

Fields and Subfields

Suppose a field **E** is an extension of a field **K**; that is, suppose **E** is a field that contains **K** as a subfield. Then **E** may be viewed as a vector space over **K** using the following operations:

(i) Vector Addition: Here $\boldsymbol{u} + \boldsymbol{v}$ in **E** is the usual addition in **E**.

(ii) Scalar Multiplication: Here $k\boldsymbol{u}$ in E, where $k\epsilon\boldsymbol{K}$ and $\boldsymbol{u}\epsilon\boldsymbol{E}$, is the usual product of 'k' and **u** as elements of E.
That is, the eight axioms of a vector space are satisfied by E and its subfield K with respect to the above two operations.

A Set that is Not a Vector Space

Show that $\boldsymbol{V} = R^2$ is not a vector space under the operation

$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2)$ and $k\vec{u} = (ku_1, 0)$

We have to show that is a vector space. Since we know that

Let us consider $\vec{u} = (2,4)$, $\vec{v} = (-3,5)$, $k = 7$

$\vec{u} + \vec{v} = (2,4) + (-3,5) = (2 - 3,4 + 5) = (-1,9)$

And $k\vec{u} = (ku_1, 0)$

$\Rightarrow 7\vec{u} = (7u_1, 0) = (7 \times 2,0) = (14,0)$

$\Rightarrow 7\vec{u} = 7(u_1, u_2) = 7(2,7) = (14, -21)$ actual value.

$\Rightarrow 7\vec{u} \neq (7u_1, 0)$

Not satisfy given operation.

Hence $\boldsymbol{V} = R^2$ is not a vector space under the given operation

Some Standard Operations:

- For R = Set of real numbers
  $\Rightarrow u + v$ addition and $k\vec{u}$ scalar multiplication.
- For R^2 = Set of ordered pairs of real numbers
  $\Rightarrow u + v = (u_1 + v_1, u_2 + v_2)$ and $k\vec{u} = (ku_1, ku_2)$
- For R^n = Set of n – tuples of real numbers

$\Rightarrow \vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$

And $k\vec{u} = (ku_1, ku_2, \dots, ku_n)$

PRACTICE:

1) Show that $R^{\infty} = \{(a_1, a_2, \dots, a_n \dots): a_1, a_2, \dots, a_n \dots \in R\}$ is a vector space. **Or** show that space of infinite sequence of real number is vector space.
2) Show that $V = set\ of\ positive\ real\ numbers$ is a vector space under defined operation as $u + v = uv$ and $ku = u^k$
3) Show that $\boldsymbol{V} = R^2$ is not a vector space under the operation $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2)$ and $k\vec{u} = (ku_1, 0)$
4) Show that $\boldsymbol{V} = R = set\ of\ real\ numbers$ is a vector space under the standard operation of addition and scalar multiplication.
5) Let Show that $\boldsymbol{V} = set\ of\ all\ ordered\ pairs\ of\ real\ numbers$ with defined operations $u + v = (u_1 + v_1, u_2 + v_2)$ and $k\vec{u} = (0, ku_2)$
   i. Show that not a vector space.
   ii. Compute $u + v$ and ku for $u = (-1,2), v = (3,4), k = 3$
6) Show that $\boldsymbol{V} = R = set\ of\ real\ numbers$ is a vector space or not under the operation of addition and scalar multiplication as follows $u + v = (u_1 + v_1 + 1, u_2 + v_2 + 1)$ and $k\vec{u} = (ku_1, ku_2)$
   i. Compute $u + v$ and ku for $u = (0,4), v = (1,-3), k = 2$
   ii. Show that $(0,0) \neq \vec{0}$
   iii. Show that $(-1,-1) = \vec{0}$
7) Show that the set of all pairs of real numbers of the form $(x, 0)$ is a vector space or not with standard operations on R^2.
8) Show that the set of all pairs of real numbers of the form $(1, x)$ is a vector space or not with the operations $(1, y) + (1, y') = (1, y + y')$ and $k(1, y) = (1, ky)$
9) Show that the set of all pairs of real numbers of the form (x, y) with $x \geq 0$ is a vector space or not with standard operations on R^2.
10) Show that the set of all n – tuples of the real numbers of the form $(x, x, \dots, x)$ is a vector space or not with standard operations on R^n.
11) Show that the set of all triples of the real numbers is a vector space or not with standard vector addition but with scalar multiplication defined by $k(x, y, z) = (k^2x, k^2y, k^2z)$

12) Show that $M_{2\times 2}$ is a vector space.
Or show that space of 2×2 matrices is a vector space.

13) Show that set of all 2×2 invertible matrices with the standard matrix addition and scalar multiplication is a vector space or not.

14) Show that set of all 2×2 non singular matrices with the standard matrix addition and scalar multiplication is a vector space or not.

15) Show that set of all 2×2 matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with the standard matrix addition and scalar multiplication is a vector space or not.

16) Show that set of all 2×2 real matrices of the form $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ with the standard matrix addition and scalar multiplication is a vector space or not.

17) Show that Function Space $F[X]$ i.e. set of all function of X into K (an arbitrary field) is vector space.
Or show that $F[X] = \{f : f\ is\ real\ valued\ function\}$ is a vector space.

18) Show that $C[a,b] = \{f : f\ is\ continuous\ real\ valued\ function\ on\ [a,b]\}$ is a vector space.

19) Show that $C'[a,b] = \{f : f \in C[a,b]\ and\ \ f(a) = f(b)\}$ is a vector space.

20) Show that the set of all real valued functions f defined everywhere on the real line and such that $f(1) = 0$ is a vector space or not with the operations $(f + g)(x) = f(x) + g(x)$ And $(kf)(x) = kf(x)$

21) Show that the set of polynomials of the form $a_0 + a_1x$ is a vector space or not with the operation
$(a_0 + a_1x) + (b_0 + b_1x) = (a_0 + b_0) + (a_1 + b_1)x$
And $k(a_0 + a_1x) = (ka_0) + (ka_1)x$

Theorem: Let **V** be a vector space over a field **K**.

i. For any scalar $k \in \boldsymbol{K}$ and $\mathbf{0} \in \boldsymbol{V}$; $k\mathbf{0} = \mathbf{0}$
ii. For $0 \in \boldsymbol{K}$ and any vector $\boldsymbol{u} \in \boldsymbol{V}$; $0\boldsymbol{u} = \mathbf{0}$
iii. If $k\boldsymbol{u} = \mathbf{0}$, where $k \in \boldsymbol{K}$ and $\boldsymbol{u} \in \boldsymbol{V}$, then $k = \mathbf{0}$ or $\boldsymbol{u} = \mathbf{0}$
iv. For any $k \in \boldsymbol{K}$ and any $\boldsymbol{u} \in \boldsymbol{V}$; $(-k)\boldsymbol{u} = k(-\boldsymbol{u}) = -k\boldsymbol{u}$
   And in particular $(-1)\boldsymbol{u} = -\boldsymbol{u}$

Proof:

Part = i: For any scalar $k \in \boldsymbol{K}$ and $\mathbf{0} \in \boldsymbol{V}$; $k\mathbf{0} = \mathbf{0}$

Since we know that $\mathbf{0} + \mathbf{0} = \mathbf{0}$ therefor $k\mathbf{0} = k(\mathbf{0} + \mathbf{0}) = k\mathbf{0} + k\mathbf{0}$

Adding$-k\mathbf{0}$ on both sides $-k\mathbf{0} + k\mathbf{0} = -k\mathbf{0} + k\mathbf{0} + k\mathbf{0} \Rightarrow \mathbf{0} = k\mathbf{0} \Rightarrow k\mathbf{0} = \mathbf{0}$

Part = ii: For $0 \in \boldsymbol{K}$ and any vector $\boldsymbol{u} \in \boldsymbol{V}$; $0\boldsymbol{u} = \mathbf{0}$

We can write $0\boldsymbol{u} + 0\boldsymbol{u} = (0 + 0)\boldsymbol{u} = 0\boldsymbol{u}$

$\Rightarrow [0\boldsymbol{u} + 0\boldsymbol{u}] + (-0\boldsymbol{u}) = 0\boldsymbol{u} + (-0\boldsymbol{u}) \Rightarrow 0\boldsymbol{u} + [0\boldsymbol{u} + (-0\boldsymbol{u})] = 0\boldsymbol{u} + (-0\boldsymbol{u})$

$\Rightarrow 0\boldsymbol{u} + \mathbf{0} = \mathbf{0} \Rightarrow 0\boldsymbol{u} = \mathbf{0}$

Part = iii: If $k\boldsymbol{u} = \mathbf{0}$, where $k \in \boldsymbol{K}$ and $\boldsymbol{u} \in \boldsymbol{V}$, then $k = \mathbf{0}$ or $\boldsymbol{u} = \mathbf{0}$

Suppose $k\boldsymbol{u} = \mathbf{0}$ and $k \neq 0$ then there exists a scalar k^{-1} such that $k^{-1}k = 1$

Thus $\boldsymbol{u} = 1\boldsymbol{u} = (k^{-1}k)\boldsymbol{u} = k^{-1}(k\boldsymbol{u}) = k^{-1}(\mathbf{0}) = \mathbf{0}$

Part = iv: For any $k \in \boldsymbol{K}$ and any $\boldsymbol{u} \in \boldsymbol{V}$; $(-k)\boldsymbol{u} = k(-\boldsymbol{u}) = -k\boldsymbol{u}$

And in particular $(-1)\boldsymbol{u} = -\boldsymbol{u}$

Using $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}$ and $k + (-k) = 0$

$\Rightarrow \mathbf{0} = k\mathbf{0} = k[\boldsymbol{u} + (-\boldsymbol{u})] = k\boldsymbol{u} + k(-\boldsymbol{u}) \Rightarrow -k\boldsymbol{u} = k(-\boldsymbol{u})$

and $\Rightarrow 0 = 0\boldsymbol{u} = [k + (-k)]\boldsymbol{u} = k\boldsymbol{u} + (-k)\boldsymbol{u} \Rightarrow -k\boldsymbol{u} = (-k)\boldsymbol{u}$

Thus $(-k)\boldsymbol{u} = k(-\boldsymbol{u}) = -k\boldsymbol{u}$ and for k = 1 we get $(-1)\boldsymbol{u} = -\boldsymbol{u}$

Linear Combinations

Let **V** be a vector space over a field **K**. A vector $\vec{v}$ in **V** is a linear combination of vectors $v_1, v_2, \dots, v_n$ in **V** if there exist scalars $k_1, k_2, \dots, k_n$ in **K** such that

$$\vec{v} = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$$

Alternatively, $\vec{v}$ is a linear combination of $k_1, k_2, \dots, k_n$ if there is a solution to the vector equation $\vec{v} = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$ where $x_1, x_2, \dots, x_n$ are unknown scalars. These scalar are called **coefficients** of linear combination.

Example:(Linear Combinations in $\mathbf{R}^n$)

Suppose we want to express $\vec{v} = (3,7,-4)$ in $\mathbf{R}^3$ as a linear combination of the vectors $u_1 = (1,2,3)$; $u_2 = (2,3,7)$; $u_3 = (3,5,6)$

We seek scalars x, y, z such that $\vec{v} = xu_1 + yu_2 + zu_3$ that is,

$$\begin{bmatrix} 3 \\ 7 \\ -4 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} + z\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$$

or $x + 2y + 3z = 3$, $2x + 3y + 5z = 7$, $3x + 7y + 6z = -4$

(For notational convenience, we have written the vectors in $\mathbf{R}^3$ as columns, because it is then easier to find the equivalent system of linear equations.) Reducing the system to echelon form yields

$$\begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ y - 3z &= -13 \end{aligned} \qquad \text{And then} \qquad \begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ -4z &= -12 \end{aligned}$$

Back-substitution yields the solution $x = 2, y = -4, z = 3$.

Thus $\vec{v} = 2u_1 - 4u_2 + 3u_3$

Remark:

Generally speaking, the question of expressing a given vector $\vec{v}$ in $\mathbf{K^n}$ as a linear combination of vectors $v_1, v_2, \ldots, v_n$ in $\mathbf{K^n}$ is equivalent to solving a system $AX = B$ of linear equations, where $\vec{v}$ is the column B of constants, and the **v's** are the columns of the coefficient matrix A. Such a system may have a unique solution (as above), many solutions, or no solution. The last case—no solution—means that $\vec{v}$ cannot be written as a linear combination of the **v's**.

Example:

Suppose that the vectors $\vec{u} = (1,2,-1)$ and $\vec{v} = (6,4,2)$ in $\mathbf{R^3}$ show that $\vec{w} = (9,2,7)$ is a linear combination of $\vec{u}$ and $\vec{v}$ and $\vec{w}' = (4,-1,8)$ is not a linear combination of $\vec{u}$ and $\vec{v}$

Solution: In order for $\vec{w}$ to be a linear combination of $\vec{u}$ and $\vec{v}$, there must be scalars k_1, k_2 such that $\vec{w} = k_1\vec{u} + k_2\vec{v}$ that is

$(9,2,7) = k_1(1,2,-1) + k_2(6,4,2) = (1k_1 + 6k_2, 2k_1 + 4k_2, -1k_1 + 2k_2)$

Equating corresponding components gives

$k_1 + 6k_2 = 9 \quad , \quad 2k_1 + 4k_2 = 2 \quad , \quad -k_1 + 2k_2 = 7$

Solving the system using Gaussian Elimination yields $k_1 = -3$, $k_2 = 2$

So $\vec{w} = -3\vec{u} + 2\vec{v}$ hence $\vec{w}$ is a linear combination of $\vec{u}$ and $\vec{v}$

Similarly In order for $\vec{w}'$ to be a linear combination of $\vec{u}$ and $\vec{v}$, there must be scalars k_1, k_2 such that $\vec{w}' = k_1\vec{u} + k_2\vec{v}$ that is

$(4,-1,8) = k_1(1,2,-1) + k_2(6,4,2) = (1k_1 + 6k_2, 2k_1 + 4k_2, -1k_1 + 2k_2)$

Equating corresponding components gives

$k_1 + 6k_2 = 4 \quad , \quad 2k_1 + 4k_2 = -1 \quad , \quad -k_1 + 2k_2 = 8$

Solving the system of equation we notify that this is inconsistent. So no such k_1, k_2 exists. Consequently $\vec{w}'$ is not a linear combination of $\vec{u}$ and $\vec{v}$

Example: (Linear combinations in $P(t)$)

Suppose we want to express the polynomial $\vec{v} = 3t^2 + 5t - 5$ as a linear combination of the polynomials

$p_1 = t^2 + 2t + 1$, $p_2 = 2t^2 + 5t + 4$, $p_3 = t^2 + 3t + 6$

We seek scalars x, y, z such that $\vec{v} = xp_1 + yp_2 + zp_3$ that is,

$3t^2 + 5t - 5 = x(t^2 + 2t + 1) + y(2t^2 + 5t + 4) + z(t^2 + 3t + 6)$ …….(i)

<u>There are two ways to proceed from here.</u>

(1) Expand the right-hand side of (i) obtaining:

$3t^2 + 5t - 5 = (x + 2y + z)t^2 + (2x + 5y + 3z)t + (x + 4y + 6z)$

Set coefficients of the same powers of 't' equal to each other, and reduce the system to echelon form:

$x + 2y + z = 3$, $2x + 5y + 3z = 5$ and $x + 4y + 6z = -5$

Or $x + 2y + z = 3$, $y + z = -1$ and $2y + 5z = -8$

Or $x + 4y + 6z = -5$, $2y + 5z = -8$ and $3z = -6$

The system is in triangular form and has a solution. Back-substitution yields the solution $x = 3, y = 1, z = -2.$ Thus, $\vec{v} = 3p_1 + p_2 - 2p_3$

(2) The equation (i) is actually an identity in the variable 't'; that is, the equation holds for any value of 't'. We can obtain three equations in the unknowns x, y, z by setting 't' equal to any three values.

For example, Set $t = 0$ in $(\mathbf{1})$ to obtain: $x + 4y + 6z = -5$

Set $t = 1$ in $(\mathbf{1})$ to obtain: $4x + 11y + 10z = 3$

Set $t = -1$ in $(\mathbf{1})$ to obtain: $y + 4z = -7$

Reducing this system to echelon form and solving by back-substitution again yields the solution $x = 3, y = 1, z = -2$

Thus (again), $\vec{v} = 3p_1 + p_2 - 2p_3$

PRACTICE:

1) Which of the following are linear combination of $\vec{u} = (0, -2, 2)$ and $\vec{v} = (1, 3, -1)$?

   i. $(2,2,2)$ ii. $(0,4,5)$ iii. $(0,0,0)$

2) Express the following as linear combination of $\vec{u} = (2,1,4)$ $\vec{v} = (1, -1, 3)$ and $\vec{w} = (3,2,5)$

   i. $(-9, -7, -15)$ ii. $(6,11,6)$ iii. $(0,0,0)$

3) Which of the following are linear combination of $A = \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ and $C = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}$?

   i. $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$

   ii. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

   iii. $\begin{bmatrix} -1 & 5 \\ 7 & 1 \end{bmatrix}$

4) For what value of 'k' will the vector $(1, -2, k)$ in $\mathbf{R}^3$ be a linear combination of the vectors $(3,0,-2), (2,-1,-5)$

5) In each part express the vector as a linear combination of $p_1 = 2 + x + 4x^2$, $p_2 = 1 - x + 3x^2$, $p_3 = 3 + 2x + 5x^2$

   i. $-9 - 7x - 15x^2$

   ii. $6 + 11x + 6x^2$

   iii. 0

   iv. $7 + 8x + 9x^2$

6) Let **V** be a vector space over a field **K**. Then show that for every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$ and $k \in \boldsymbol{K}$: we have $k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$

Subspaces

Let **V** be a vector space over a field **K** and let **W** be a subset of **V**. Then **W** is a subspace of **V** if **W** is itself a vector space over **K** with respect to the operations of vector addition and scalar multiplication on **V**.

The way in which one shows that any set **W** is a vector space is to show that **W** satisfies the eight axioms of a vector space. However, if **W** is a subset of a vector space **V**, then some of the axioms automatically hold in **W**, because they already hold in **V**. Simple criteria for identifying subspaces follow.

Theorem: Suppose **W** is a subset of a vector space **V**. Then **W** is a subspace of **V** iff the following two conditions hold:

a) The zero vector **0** belongs to **W**
b) For every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$ and $k \in \boldsymbol{K}$:
   (i) The sum $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$
   (ii) The multiple $k\boldsymbol{u} \in \boldsymbol{W}$

Property (i) in (b) states that **W** is *closed under vector addition* and property (ii) in (b) states that **W** is *closed under scalar multiplication*. Both properties may be combined into the following equivalent single statement:

(b') For every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$; $a, b \in \boldsymbol{K}$, the linear combination $a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{W}$

Now let **V** be any vector space. Then **V** automatically contains two subspaces: <u>the set {0} consisting of the zero vector alone and the whole space **V** itself</u>. These are sometimes called the <u>trivial subspaces</u> of **V**. this means that every vector space has at least two subspaces.

The Zero Subspace:

If **V** is any vector space and $W = \{0\}$ is the subspace of **V** that contains the zero vector only, then **W** is closed under addition and scalar multiplication

Since $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $k\mathbf{0} = \mathbf{0}$ for any scalar ‘k’

Then we call **W** the ***zero subspace*** of **V**.

Remember that smallest vector space is $\{0\}$.

Example 2: *Show that Q is not a subspace of R.*

Solution: Let $x, y \in Q$, then for irrational numbers $\sqrt{2}, \sqrt{3} \in R$

$$\sqrt{2}x + \sqrt{3}y \notin Q$$

This shows that Q is not a subspace of *R*.

Example 3: *Show that the set of all irrational numbers is not a subspace of R.*

Solution: Let Q^C be the set of all irrational numbers, then for $\sqrt{2}, \sqrt{3} \in Q^C$ and $\sqrt{2}, \sqrt{3} \in Q^C \subset R$,

$$\sqrt{2}\sqrt{2} + \sqrt{3}\sqrt{3} = 2 + 3 = 5 \notin Q^C$$

This shows that Q^C is not a subspace of *R*.

Example 4: *Show that the set of all 2×2 non-singular matrices is not a subspace of M_2.*

Solution: Let *A* and *B* be two 2×2 non-singular matrices, then $0A + 0B = 0$ is a singular matrix, so the set of all 2×2 non-singular matrices does not contain this matrix. Hence the set of all 2×2 non-singular matrices is not a subspace of M_2.

Example 6: Show that $W = \{(x, y, z) : x + y + z = 0\}$ is a subspace of R^3.

PU. 2016. 2015. 2014. Mathematics A-III. BS (Math/Stat/Chem)

Solution: Let $w_1, w_2 \in W$, $\alpha, \beta \in R$ then

$$w_1 = (x, y, z), w_2 = (x', y', z'); \; x + y + z = 0, \; x' + y' + z' = 0$$

Consider

$$\begin{aligned} \alpha w_1 + \beta w_2 &= \alpha(x, y, z) + \beta(x', y', z') \\ &= (\alpha x, \alpha y, \alpha z) + (\beta x', \beta y', \beta z') \\ &= (\alpha x + \beta x', \alpha y + \beta y', \alpha z + \beta z') \end{aligned}$$

Now $(\alpha x + \beta x') + (\alpha y + \beta y') + (\alpha z + \beta z') = \alpha(x + y + z) + \beta(x' + y' + z')$

$$= \alpha(0) + \beta(0) = 0$$

Thus, $\alpha w_1 + \beta w_2 \in W$. This shows that *W* is a subspace of R^3.

Theorem: If $S = \{w_1, w_2, \dots \dots, w_r\}$ is a non-empty set of vectors in a vector space **V** then the set **W** of all possible linear combinations of the vectors in **S** is a subspace of **V**.

Proof: Let **W** be the set of all possible linear combinations of the vectors in **S**

We must show that **W** is closed under addition and scalar multiplication.

To prove closure under addition let $\vec{u}, \vec{v} \in$ **W** as

$\vec{u} = c_1 w_1 + c_2 w_2 + \cdots + c_r w_r$ and $\vec{v} = k_1 w_1 + k_2 w_2 + \cdots + k_r w_r$

Then their sum can be written as

$\vec{u} + \vec{v} = (c_1 + k_1) w_1 + (c_2 + k_2) w_2 + \cdots + (c_r + k_r) w_r$

Which is the linear combination of the vectors in **S**.

To prove closure under multiplication let $\vec{u} \in$ **W**and k ∈ **K** as

$\vec{u} = c_1 w_1 + c_2 w_2 + \cdots + c_r w_r$

then $k\vec{u} = (kc_1) w_1 + (kc_2) w_2 + \cdots + (kc_r) w_r$

$k\vec{u} = a_1 w_1 + a_2 w_2 + \cdots + a_r w_r$

Which is the linear combination of the vectors in **S**.

Then **W** is closed under multiplication.

Hence **W** is a subspace of **V**

Theorem: If $S = \{w_1, w_2, \dots, w_r\}$ is a non-empty set of vectors in a vector space **V** and if the set **W** of all possible linear combinations of the vectors in **S** is a subspace of **V** then set **W** is the smallest subspace of **V** that contains all of the vectors in **S** in the sense that any other subspace that contains those vectors contains **W**.

Proof: Let $\boldsymbol{W'}$ be any subspace of **V** that contains all of the vectors in **S**. Since $\boldsymbol{W'}$ is closed under addition and scalar multiplication, it contains all linear combinations of the vectors in **S** and hence contains **W**.

Theorem: Suppose **W** is a subset of a vector space **V**. Then **W** is a subspace of **V** iff the following two conditions hold:

For every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$ and $k \in \boldsymbol{K}$:

i. The sum $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$
ii. The multiple $k\boldsymbol{u} \in \boldsymbol{W}$

Proof:

Suppose that **W** is a subspace of **V** then by definition of a subspace **W** is a vector space over field **K** and hence given conditions must hold.

Conversely:

Suppose that **W** is a non – empty subset of **V** such that the conditions (i) and (ii) are satisfied then we have to show that **W** is a subspace of **V**

For this let $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$ and since $\boldsymbol{v} \in \boldsymbol{W}$ also $-1 \in \boldsymbol{K} \quad \Rightarrow -1\boldsymbol{v} = -\boldsymbol{v} \in \boldsymbol{W}$ by (ii)

So for $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W} \Rightarrow \boldsymbol{u}, -\boldsymbol{v} \in \boldsymbol{W} \Rightarrow \boldsymbol{u} - \boldsymbol{v} \in \boldsymbol{W}$ by (i)

Which shows that **W** is a subspace of **V** under addition.

Now because **V** is an Abelian group under addition and $\boldsymbol{W} \subseteq \boldsymbol{V}$so **W** is also an Abelian group under addition.

Also for $k \in \boldsymbol{K}$ and $\boldsymbol{u} \in \boldsymbol{W} \Rightarrow k\boldsymbol{u} \in \boldsymbol{W}$ by (ii)

The remaining four conditions of scalar multiplication holds in **W** because they hold in **V** therefore **W** is subspace of **V**.

Theorem: Suppose **W** is a non – empty subset a vector space **V**. Then **W** is a subspace of a vector space **V**. If and only if for every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$; $a, b \in \boldsymbol{K}$, the linear combination $a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{W}$

Proof: Suppose that **W** is a subspace of **V** then by definition of a subspace **W** is a vector space over field **K** and hence given condition must hold.

Conversely: Suppose that **W** is a non – empty subset of **V** such that the condition $a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{W}$ is satisfied then we have to show that **W** is a subspace of **V**

For this let $a = b = 1 \in \boldsymbol{K}$ then according to condition

$a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{W} \Rightarrow \boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$

Now take $b = 0 \in \boldsymbol{K}$ so that $a\boldsymbol{u} \in \boldsymbol{W}$

This means that for every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$; $a \in \boldsymbol{K}$, we have $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$and $a\boldsymbol{u} \in \boldsymbol{W}$

Thus by theorem "**W** is a subspace of **V** iff the following two conditions hold:

For every $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$ and $k \in \boldsymbol{K}$: The sum $\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{W}$ and the multiple $k\boldsymbol{u} \in \boldsymbol{W}$ "

$\Rightarrow$ **W** is a subspace of **V**

Theorem:

Suppose **U** and **W** are subspaces of a vector space **V**. then show that $\boldsymbol{U} \cap \boldsymbol{W}$ is also a subspace of **V.**

Proof: Suppose that $a, b \in \boldsymbol{K}$ and $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U} \cap \boldsymbol{W}$

$\Rightarrow \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$ also $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{W}$

$\Rightarrow a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{U}$ also $a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{W}$ being Subspace.

$\Rightarrow a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{U} \cap \boldsymbol{W}$

Hence for $a, b \in \boldsymbol{K}$ and $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U} \cap \boldsymbol{W} \Rightarrow a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{U} \cap \boldsymbol{W}$

This implies $\boldsymbol{U} \cap \boldsymbol{W}$ is a subspace of **V**

Theorem:

Show that the intersection of any number of subspaces of a vector space **V** is a subspace of **V.**

Proof: Suppose that $\{\boldsymbol{U}_\propto\ ; \propto \in I\}$ be any subcollection of subspaces of a vector space **V** over the field **K**. Then we have to show that is also a subspace of **V**.

For this Suppose that $a, b \in \boldsymbol{K}$ and $\boldsymbol{u}, \boldsymbol{v} \in \cap_{\propto \in I} \boldsymbol{U}_\propto$

$\Rightarrow \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}_\propto$ for all $\propto \in I$

$\Rightarrow a\boldsymbol{u} + b\boldsymbol{v} \in \boldsymbol{U}_\propto$ for each $\propto \in I$ being Subspace. $\Rightarrow a\boldsymbol{u} + b\boldsymbol{v} \in \cap_{\propto \in I} \boldsymbol{U}_\propto$

This implies $\cap_{\propto \in I} \boldsymbol{U}_\propto$ is a subspace of **V**

Theorem:

Suppose **U** and **W** are subspaces of a vector space **V**. then $\boldsymbol{U} + \boldsymbol{W}$ is also a subspace of **V** containing both **U** and **W.** Further $\boldsymbol{U} + \boldsymbol{W}$ is a smallest subspace of **V** containing both **U** and **W**

Proof: Given that **U** and **W** are subspaces of a vector space **V** then we can define

$\boldsymbol{U} + \boldsymbol{W} = \{\boldsymbol{u} + \boldsymbol{w} : \boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{w} \in \boldsymbol{W}\}$

We will prove that $\boldsymbol{U} + \boldsymbol{W}$ is also a subspace of **V**

For this Suppose that $a, b \in \boldsymbol{K}$ and $\boldsymbol{v_1}, \boldsymbol{v_2} \in \boldsymbol{U} + \boldsymbol{W}$

$\Rightarrow \boldsymbol{v_1} = \boldsymbol{u_1} + \boldsymbol{w_1},\ \boldsymbol{v_2} = \boldsymbol{u_2} + \boldsymbol{w_2}$ where $\boldsymbol{u_1}, \boldsymbol{u_2} \in \boldsymbol{U} \quad \boldsymbol{and} \quad \boldsymbol{w_1}, \boldsymbol{w_2} \in \boldsymbol{W}$

$\Rightarrow a\boldsymbol{u_1} + b\boldsymbol{u_2} \in \boldsymbol{U}$ also $a\boldsymbol{w_1} + b\boldsymbol{w_2} \in \boldsymbol{W}$ being Subspace.

$\Rightarrow (a\boldsymbol{u_1} + b\boldsymbol{u_2}) + (a\boldsymbol{w_1} + b\boldsymbol{w_2}) \in \boldsymbol{U} + \boldsymbol{W}$

$\Rightarrow (a\boldsymbol{u_1} + a\boldsymbol{w_1}) + (b\boldsymbol{u_2} + b\boldsymbol{w_2}) \in \boldsymbol{U} + \boldsymbol{W}$

$\Rightarrow a(\boldsymbol{u_1} + \boldsymbol{w_1}) + b(\boldsymbol{u_2} + \boldsymbol{w_2}) \in \boldsymbol{U} + \boldsymbol{W} \Rightarrow a\boldsymbol{v_1} + b\boldsymbol{v_2} \in \boldsymbol{U} + \boldsymbol{W}$

This implies $\boldsymbol{U} + \boldsymbol{W}$ is a subspace of **V**

Next we will prove that $\boldsymbol{U+W}$ is a smallest subspace of **V** containing both **U** and **W** i.e. $\boldsymbol{U \subseteq U+W}$ and $\boldsymbol{W \subseteq U+W}$

Since $\boldsymbol{u} \in \boldsymbol{U}$ $\quad \boldsymbol{and}\ \ \boldsymbol{0} \in \boldsymbol{W} \ \Rightarrow \boldsymbol{u}+\boldsymbol{0}=\boldsymbol{u} \in \boldsymbol{U}+\boldsymbol{W}$ for all $\boldsymbol{u} \in \boldsymbol{U}$

$\Rightarrow \boldsymbol{U \subseteq U+W}$ and similarly $\boldsymbol{W \subseteq U+W}$

Hence $\boldsymbol{U+W}$ is a subspace of **V** containing both **U** and **W**

Now

we will prove that $\boldsymbol{U+W}$ is a smallest subspace of **V** containing both **U** and **W**

let $\boldsymbol{S}$ be any subspace of **V** containing both **U** and **W**

then for every $\boldsymbol{u} \in \boldsymbol{U}\ \boldsymbol{and}\ \ \boldsymbol{w} \in \boldsymbol{W}$ we have $\boldsymbol{u} \in \boldsymbol{S}\ \ \boldsymbol{and}\ \boldsymbol{w} \in \boldsymbol{S}$ so that $\boldsymbol{u}+\boldsymbol{w} \in \boldsymbol{S}$

but $\boldsymbol{u}+\boldsymbol{w} \in \boldsymbol{U}+\boldsymbol{W}$ So $\boldsymbol{U+W \subseteq S}$

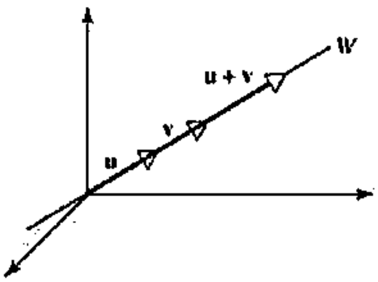
Hence $\boldsymbol{U+W}$ is a smallest subspace of **V** containing both **U** and **W**

Only way to learn mathematics is to do mathematics

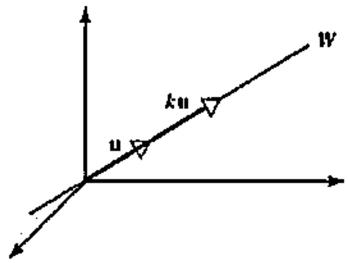
Examples of *nontrivial subspaces* follow.

Lines through the origin are subspace of R^2 and of R^3

If **W** is the line through the origin of either R^2 or R^3, then adding two vectors on line or multiplying a vector on the line by a scalar produces another vector on the line, so **W** is closed under addition and scalar multiplication. Hence a subspace.



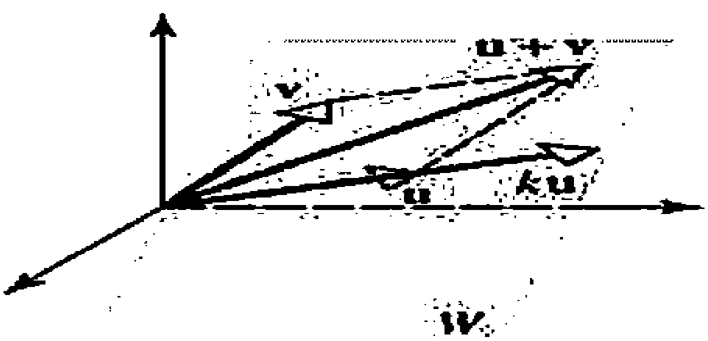
(*a*) *W* is closed under addition.



(*b*) *W* is closed under scalar multiplication.

Planes through the origin are subspace of R^3

If **u** and **v** are vectors in a plane **W** through the origin of R^3, then it is evident geometrically that **u** + **v** and k**u** also lie in the same plane **W** for any scalar 'k'. thus **W** is closed under addition and scalar multiplication.



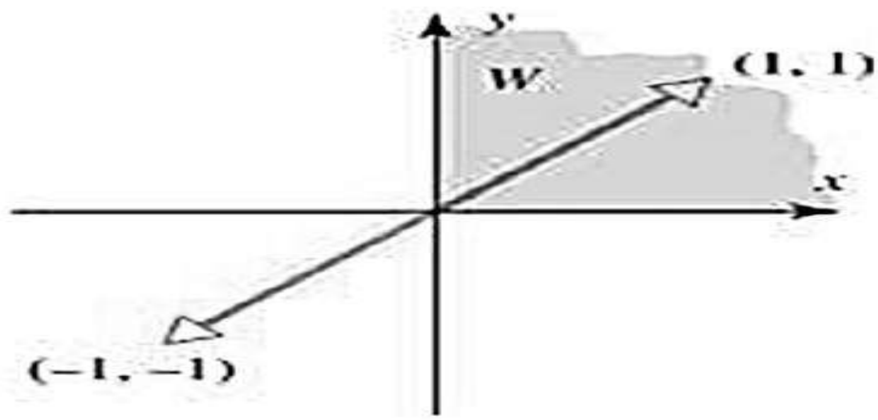
A list of subspace of R^2 and of R^3

Subspace of R^2	**Subspace of R^3**
{0}	{0}
Lines through the origin	Lines through the origin Planes through the origin
R^2	R^3

A subset of $\boldsymbol{R^2}$ that is not a subspace of $\boldsymbol{R^2}$

Consider $\boldsymbol{W} = \{(x, y): x \geq 0, \geq 0\}$ in $\boldsymbol{R^2}$ This is not a subspace of $\boldsymbol{R^2}$ because it is not closed under scalar multiplication.

For example: $\boldsymbol{v} = (1,1) \in \boldsymbol{W}$ but $-1\boldsymbol{v} = (-1, -1) \notin \boldsymbol{W}$



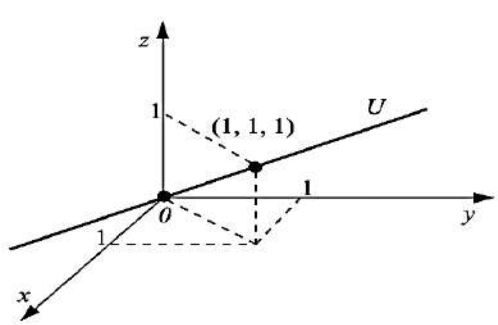
Example: Consider the vector space $\boldsymbol{V} = \boldsymbol{R^3}$ and Let **U** consist of all vectors in $\mathbf{R}^3$ whose entries are equal; that is, $\boldsymbol{U} = \{(a, b, c): a = b = c\}$

For example, $(1, 1, 1), (-3, -3, -3), (7, 7, 7), (-2, -2, -2)$ are vectors in U. Geometrically, U is the line through the origin O and the point (1, 1, 1) as shown in Figure. Clearly $\mathbf{0} = (0,0, 0)$ belongs to U, because all entries in $\mathbf{0}$ are equal. Further, suppose **u** and **v** are arbitrary vectors in U, say, $\boldsymbol{u} = (a, a, a)$ and $\boldsymbol{v} = (b, b, b)$.

Then, for any scalar $k \in \boldsymbol{R}$, the following are also vectors in **U**:

$\boldsymbol{u} + \boldsymbol{v} = (a + b, a + b, a + b)$ and $k\boldsymbol{u} = (ka, ka, ka)$

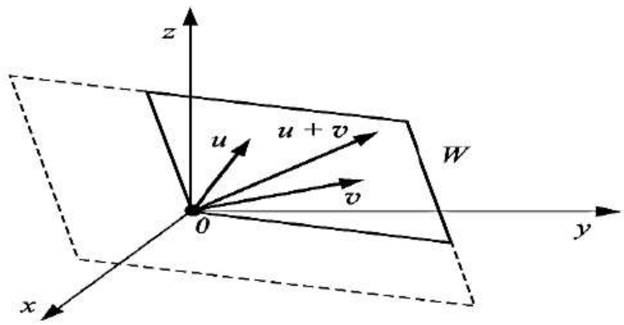
Thus, U is a subspace of $\boldsymbol{R^3}$.



Example: Consider the vector space $\boldsymbol{V} = \boldsymbol{R^3}$ And Let **W** be any plane in $\mathbf{R}^3$

passing through the origin, as pictured in Figure. Then $\mathbf{0} = (0,0,0)$belongs to **W**,

because we assumed **W** passes through, the origin O. Further, suppose **u** and **v** are vectors in **W**. Then **u** and **v** may be viewed as arrows in the plane **W** emanating from the origin O, as in Figure. The sum $\boldsymbol{u} + \boldsymbol{v}$ and any multiple k**u** of **u** also lie in the plane **W**. Thus, **W** is a subspace of $\boldsymbol{R^3}$



PRACTICE:

1) Use subspace criteria to determine which of the following are subspaces of $\boldsymbol{R^3}$?
   a) All vectors of the form $(a, 0{,}0)$
   b) All vectors of the form $(a, 1{,}1)$
   c) All vectors of the form (a, b, c) where $b = a + c$
   d) All vectors of the form (a, b, c) where $b = a + c + 1$
   e) All vectors of the form $(a, b, 0)$
   f) All vectors of the form (x, y, z) where $x + y + z = 0$
   g) All vectors of the form (x, y, z) where $x \geq 0$
   h) All vectors of the form (x, y, z) where $x^2 + y^2 + z^2 \leq 0$ and $x^2 + y^2 + z^2 \leq 1$
   i) All vectors of the form (x, y, z) where x, y, z are rationals
   j) All vectors of the form (x, y, z) where $x, z \in \boldsymbol{R}$
   k) All vectors of the form (x, y, z) where $y^2 = x^2 + z^2$

2) Show that set of rational numbers **Q** is not a subspace of **R**.
3) The union of any number of subspaces need not to be a subspace. Prove!

Subspaces of $\boldsymbol{M_{nn}}$

We know that the sum of two symmetric $n \times n$ matrices is symmetric and that a scalar multiple of a symmetric $n \times n$ matrix is symmetric. Thus the set of symmetric $n \times n$ matrices is closed under addition and scalar multiplication and hence is a subspace of $\boldsymbol{M_{nn}}$.

Similarly the sets of upper triangular matrices, lower triangular matrices and diagonal matrices are subspaces of $\boldsymbol{M_{nn}}$.

Let $\boldsymbol{V} = \boldsymbol{M_{nn}}$ the vector space of $n \times n$ matrices. Let $\mathbf{W_1}$ be the subset of all (upper) triangular matrices then $\mathbf{W_1}$ is a subspace of $\mathbf{V}$, because $\mathbf{W_1}$ contains the zero matrix $\mathbf{0}$ and $\mathbf{W_1}$ is closed under matrix addition and scalar multiplication; that is, the sum and scalar multiple of such triangular matrices are also triangular.

A subset of $\boldsymbol{M_{nn}}$ that is not a Subspace

The set W of invertible $n \times n$ matrices is not a subspace of $\boldsymbol{M_{nn}}$, failing on two counts;

- It is not closed under addition
- It is not closed under scalar multiplication

We will illustrate this with an example;

Let us consider two matrices $U = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ and $V = \begin{bmatrix} -1 & 2 \\ -2 & 5 \end{bmatrix}$ in $\boldsymbol{M_{22}}$ then the matrix $0U$ is the 2×2 zero matrix and hence is not invertible and then $U + V$ has a column of zeros, so it also is not invertible.

Remark:

Matrices whose determinant is zero are not subspaces.

PRACTICE:

Use subspace criteria to determine which of the following are subspaces of $\boldsymbol{M_{nn}}$?

a) The set of all diagonal $n \times n$ matrices.
b) The set of all $n \times n$ matrices A such that $det(A) = 0$
c) The set of all $n \times n$ matrices A such that $tr(A) = 0$
d) The set of all symmetric $n \times n$ matrices.
e) The set of all $n \times n$ matrices A such that $A^T = -A$
f) The set of all $n \times n$ matrices A for which $Ax = 0$ has only the trivial solution.
g) The set of all $n \times n$ matrices A such that $AB = BA$ for some fixed $n \times n$ matrices B.

The Subspace $\boldsymbol{C}(-\infty, \infty)$

Since we know that a theorem in calculus "a sum of continuous functions is continuous and that a constant times a continuous function is continuous" Rephrased in vector language, the set of continuous functions on $(-\infty, \infty)$ is a subspace of $\boldsymbol{F}(-\infty, \infty)$ we will denote this subspace by $\boldsymbol{C}(-\infty, \infty)$

Remark:

- A function with continuous derivative is said to be continuously differentiable.
- Sum of two continuous differentiable functions is continuously differentiable and that a constant times a continuous differentiable function is continuously differentiable.
- Functions that are continuously differentiable on $(-\infty, \infty)$ form a subspace of $\boldsymbol{F}(-\infty, \infty)$**.** we will denote this subspace by $\boldsymbol{C}(-\infty, \infty)$
- We will denote this subspace by $\boldsymbol{C'}(-\infty, \infty)$ where the superscript emphasizes that the first derivatives are continuous.
- We will denote this subspace by $\boldsymbol{C^m}(-\infty, \infty), \boldsymbol{C^\infty}(-\infty, \infty)$ where the superscript emphasizes that the m continuous derivatives and the derivatives of all orders respectively.

The Subspace of all Polynomials:

Since we know that a p**olynomial** is a function that can be expressed in the form $P(t) = a_0 + a_1 t + \cdots + a_n t^n$ where $a_0, a_1, \ldots, a_n$ are constants.

Also we know that "the sum of two polynomials is a polynomial and that a constant time polynomial is polynomial." Thus the set **W** of all polynomials is closed under addition and scalar multiplication and hence is a subspace of $\boldsymbol{F}(-\infty, \infty)$. We will denote it by $\boldsymbol{P}_\infty$

Degree of Polynomial:

The highest power of the variable that occurs with a non – zero coefficient.

For example the Polynomial $\boldsymbol{P}(\boldsymbol{t}) = a_0 + a_1 t + \cdots + a_n t^n$ with $a_n \neq 0$ has degree 'n'

The Subspace of all Polynomials of degree ≤ n:

It is not true that the set **W** of polynomials with positive degree 'n' is a subspace of $\boldsymbol{F}(-\infty, \infty)$ because the set is not closed under addition.

For example the Polynomials $1 + 2t + 3t^2$ and $5 + 7t - 3t^2$ both have degree 2 but their sum has degree 1.

But for each non – negative integer 'n' the polynomials of degree 'n' or less form a subspace of $\boldsymbol{F}(-\infty, \infty)$. We will denote it by $\boldsymbol{P}_{\boldsymbol{n}}$

Example:

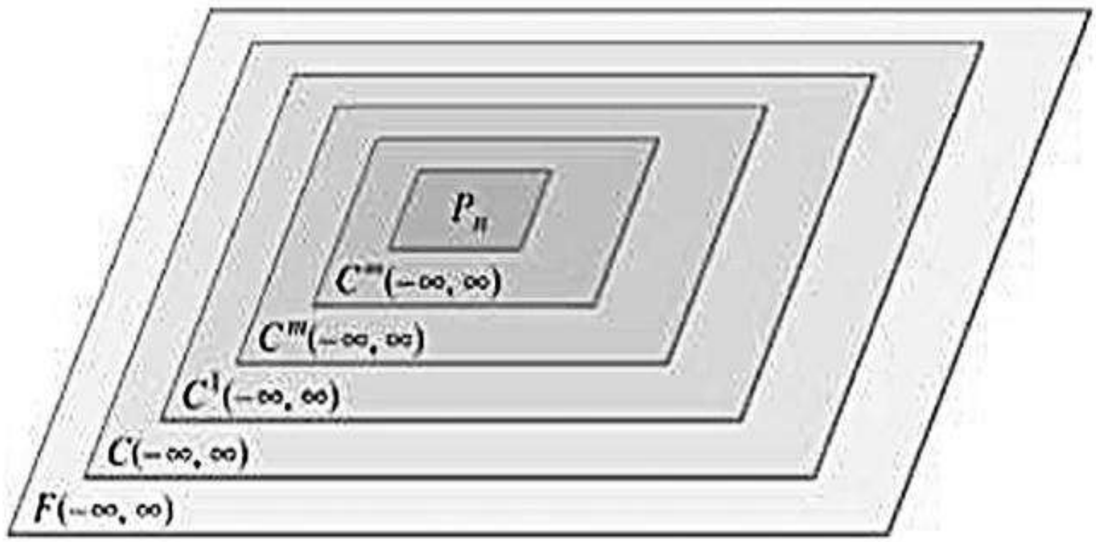
Let $\boldsymbol{V} = \boldsymbol{P}(\boldsymbol{t})$, the vector space $\boldsymbol{P}(\boldsymbol{t})$of polynomials. Then the space $\boldsymbol{P}_{\boldsymbol{n}}(\boldsymbol{t})$ of polynomials of degree at most 'n' may be viewed as a subspace of $\boldsymbol{P}(\boldsymbol{t})$**.** Let $\boldsymbol{Q}(\boldsymbol{t})$ be the collection of polynomials with only even powers of 't'. For example, the following are polynomials in $\boldsymbol{Q}(\boldsymbol{t})$:

$p_1 = 3 + 4t^2 - 5t^6$ and $p_2 = 6 - 7t^4 + 9t^6 + 3t^{12}$

(We assume that any constant $k = kt^0$ is an even power of 't') Then $\boldsymbol{Q}(\boldsymbol{t})$ is a subspace of $\boldsymbol{P}(\boldsymbol{t})$.

Remark: It is proved in calculus that polynomials are continuous functions and have continuous derivatives of all orders on $(-\infty,\infty)$, thus it follows that P_∞ is not only a subspace of $F(-\infty,\infty)$ but is also a subspace of $C^\infty(-\infty,\infty)$

All spaces discussed previously are nested. See follow;



PRACTICE:

1) **(Calculus Required)** Show that followings set of functions are subspaces of $\boldsymbol{F}(-\infty,\infty)$?

i. All functions f in $\boldsymbol{F}(-\infty,\infty)$ that satisfy $f(0)=0$
ii. All functions f in $\boldsymbol{F}(-\infty,\infty)$ that satisfy $f(0)=1$
iii. All functions f in $\boldsymbol{F}(-\infty,\infty)$ that satisfy $f(-x)=f(x)$
iv. All polynomials of degree 2.
v. All continuous functions on $(-\infty,\infty)$
vi. All differentiable functions on $(-\infty,\infty)$
vii. All differentiable functions on $(-\infty,\infty)$ that satisfy $f'+2f=0$

2) **(Calculus Required)** Show that the set of continuous functions $f=f(x)$ on $[a,b]$ such that $\int_a^b f(x)\,dx=0$ is a subspace of $C[a,b]$

3) Use subspace criteria to determine which of the following are subspaces of $\boldsymbol{P_3}$?

   i. All polynomials $a_0 + a_1x + a_2x^2 + a_3x^3$ for which $a_0 = 0$

   ii. All polynomials $a_0 + a_1x + a_2x^2 + a_3x^3$ for which $a_0 + a_1 + a_2 + a_3 = 0$

   iii. All polynomials $a_0 + a_1x + a_2x^2 + a_3x^3$ in which a_0, a_1, a_2, a_3 are rational

   iv. All polynomials $a_0 + a_1x$ in which a_0, a_1 are real numbers.

4) Use subspace criteria to determine which of the following are subspaces of $\boldsymbol{R}^\infty$?

   i. All sequences **v** in R^∞ of the form $\boldsymbol{v} = (v, 0, v, 0, v, 0, \dots \dots ..)$

   ii. All sequences **v** in R^∞ of the form $\boldsymbol{v} = (v, 1, v, 1, v, 1, \dots \dots ..)$

   iii. All sequences **v** in R^∞ of the form $\boldsymbol{v} = (v, 2v, 4v, 8v, 16v, \dots \dots ..)$

   iv. All sequences **v** in R^∞ whose components are zero from some point on.

5) Let **V** be a vector space of functions $f: \boldsymbol{R} \to \boldsymbol{R}$. Show that **W** is subspace of **V** where;

   i. $W = \{f(x): f(1) = 0\}$ all functions whose values at 1 is 0.

   ii. $W = \{f(x): f(3) = f(1)\}$ all functions assigning to same value to 3 and 1.

   iii. $W = \{f(t): f(x) = -f(x)\}$ all odd functions.

Spanning Sets

Let **V** be a vector space over **K**. Vectors $v_1, v_2, \dots, v_n$ in **V** are said to <u>span **V**</u> or to form a <u>spanning set of **V**</u> if every $\vec{v}$ in **V** is a linear combination of the vectors $v_1, v_2, \dots, v_n$. That is, if there exist scalars $k_1, k_2, \dots, k_n$ in **K** such that

$$\vec{v} = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$$

Or if $S = \{w_1, w_2, \dots \dots, w_r\}$ is a non – empty set of a vector in a vector space **V** then the subspace **W** of **V** that consists of all possible linear combinations of the of the vectors in **S** is called the subspace of **V** generated by **S** and we say that the vectors $w_1, w_2, \dots, w_r$ span **W**. we denote this subspace as

$$W = span\{w_1, w_2, \dots, w_r\} \qquad \text{or} \qquad W = span(S)$$

Or Let **S** be a non – empty subset of a vector space **V**. then the set of all linear combinations of finite number of elements of **S** is called the linear span of **S** and is denoted by $\langle S \rangle$ or $L(S)$ or $[S]$

Remarks:

- Suppose $v_1, v_2, \dots, v_n$ span **V**. Then, for any vector **w**, the set $w, v_1, v_2, \dots, v_n$ also spans **V**.
- Suppose $v_1, v_2, \dots, v_n$ span **V** and suppose v_k is a linear combination of some of the other v's. Then the v's without v_k also span **V**.
- Suppose $v_1, v_2, \dots, v_n$ span **V** and suppose one of the v's is the zero vector. Then the v's without the zero vector also span **V**.
- If $S = \{v_1, v_2, \dots, v_r\}$ and $S' = \{w_1, w_2, \dots, w_k\}$ are non – empty sets of a vector in a vector space **V**, then $Span\{v_1, v_2, \dots, v_r\} = Span\{w_1, w_2, \dots, w_k\}$ iff each vector in S is a linear combination of those in S', and each vector in S' is a linear combination of those in S

Theorem: If **S** and **T** are subsets of **V** then $S \subset T \Rightarrow \langle S \rangle \subset \langle T \rangle$

Proof: Let $S = \{v_1, v_2, \dots, v_r\}$ and $T = \{v_1, v_2, \dots, v_r, v_{r+1}, \dots., v_n\}$

Let $v \in \langle S \rangle$ then $v = a_1 v_1 + a_2 v_2 + \cdots \dots + a_r v_r$ a linear combination of vectors in S.

We may write $v = a_1 v_1 + a_2 v_2 + \cdots + a_r v_r + 0 v_{r+1} + \cdots + 0 v_n$ a linear combination of vectors in T. then this implies $v \in \langle T \rangle$

Hence $S \subset T \Rightarrow \langle S \rangle \subset \langle T \rangle$

Theorem:

Let **S** be a non – empty set of vectors in a vector space **V** over a field **K**. then $\langle S\rangle$ is a subspace of **V** containing **S** and it is the smallest subspace of **V** containing **S**.

Proof:

Let $a, b \in \boldsymbol{K}$ and $u, v \in \langle S\rangle$ also $u_i, v_j \in \boldsymbol{S}$ and $a_i, b_j \in \boldsymbol{K}$ then

$$u = a_1u_1 + a_2u_2 + \cdots + a_nu_n = \sum_{i=1}^{n} a_iu_i$$

$$v = b_1v_1 + b_2v_2 + \cdots + b_mv_m = \sum_{j=1}^{m} b_jv_j$$

Now $au + bv = a(\sum_{i=1}^{n} a_iu_i) + b\left(\sum_{j=1}^{m} b_jv_j\right) = \sum_{i=1}^{n} a(a_iu_i) + \sum_{j=1}^{m} b\left(b_jv_j\right)$

$au + bv = \sum_{i=1}^{n}(aa_i)u_i + \sum_{j=1}^{m} b\left(b_jv_j\right) \qquad \therefore a(bv) = (ab)v$

Which shows that $au + bv$ is a linear combination of vectors in **S**.

So $au + bv \in \langle S\rangle$

So $a, b \in \boldsymbol{K}$ and $u, v \in \langle S\rangle \Rightarrow au + bv \in \langle S\rangle \Rightarrow \langle S\rangle$ is a subspace of **V**.

Now we prove that $\langle S\rangle$ is the smallest subspace of **V** containing **S**.

If **W** is any other subspace of **V** containing **S** then it contains all vectors of the form $u = \sum_{i=1}^{n} a_iu_i$ where $u_i \in \boldsymbol{S}$ and $a_i \in \boldsymbol{K}$

$\Rightarrow \langle S\rangle \subseteq W$

Thus $\langle S\rangle$ is the smallest subspace of **V** containing **S**.

The standard unit vectors spans $\mathbf{R}^n$

Since we know that standard unit vectors are as follows;

$\hat{e}_1 = (1,0,0, \dots ,0)$, $\hat{e}_2 = (0,1,0, \dots ,0)$, … … … … .. $\hat{e}_n = (0,0,0, \dots ,1)$

These vectors span $\mathbf{R}^n$ since every vector $\vec{v} = (v_1, v_2, \dots \dots, v_n)$ in $\mathbf{R}^n$ can be expressed as $\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2 + \cdots + v_n\hat{e}_n$ which is a linear combination of $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$

Example: Consider $\vec{v} = (a, b, c)$ in $\mathbf{R}^3$ then it can be written as a linear combination of standard unit vectors in $\mathbf{R}^3$ as follows;

$\vec{v} = (a, b, c) = a(1,0,0) + b(0,1,0) + c(0,0,1) = a\hat{e}_1 + b\hat{e}_2 + c\hat{e}_3$

Example: Consider the vector space **V** = $\mathbf{R}^3$ then show that

$\hat{e}_1 = (1,0,0), \hat{e}_2 = (0,1,0), \hat{e}_3 = (0,0,1)$ span $\vec{v} = (5, -6, 2)$ in $\mathbf{R}^3$

Solution: Since we know that "The standard unit vectors spans R^n" then we can write $\vec{v} = (5, -6, 2)$ as a linear combination of standard unit vectors in $\mathbf{R}^3$

$\vec{v} = (5, -6, 2) = 5(1,0,0) - 6(0,1,0) + 2(0,0,1) = 5\hat{e}_1 - 6\hat{e}_2 + 2\hat{e}_3$

Example: Consider the vector space **V** = $\mathbf{R}^3$ then we claim that the following vectors also form a spanning set of $\mathbf{R}^3$

$\boldsymbol{w_1} = (1,1,1), \boldsymbol{w_2} = (1,1,0), \boldsymbol{w_3} = (1,0,0)$

Specifically, if $\vec{v} = (a, b, c)$ is any vector in $\mathbf{R}^3$, then

$\vec{v} = (a, b, c) = c\boldsymbol{w_1} + (b - c)\boldsymbol{w_2} + (a - b)\boldsymbol{w_3}$

For example, $\vec{v} = (5, -6, 2) = 2\boldsymbol{w_1} - 8\boldsymbol{w_2} + 11\boldsymbol{w_3}$

Example: Consider the vector space **V** = $\mathbf{R}^3$ then One can show that $\vec{v} = (2,7,8)$ cannot be written as a linear combination of the vectors

$\boldsymbol{u_1} = (1,2,3), \boldsymbol{u_2} = (1,3,5), \boldsymbol{u_3} = (1,5,9)$ Accordingly, $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ do not span $\mathbf{R}^3$

Testing for spanning:

Determine whether the vectors $\boldsymbol{v_1} = (1,1,2), \boldsymbol{v_2} = (1,0,1), \boldsymbol{v_3} = (2,1,3)$ span the vector space $\mathbf{R}^3$

Solution:

must determine whether an arbitrary vector $\boldsymbol{b} = (b_1, b_2, b_3)$ in $\mathbf{R}^3$ can be expressed as linear combination $\boldsymbol{b} = k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + k_3\boldsymbol{v_3}$ of the vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$

Expressing above equation in terms of components gives

$(b_1, b_2, b_3) = k_1(1,1,2) + k_2(1,0,1) + k_3(2,1,3)$

$(b_1, b_2, b_3) = (k_1 + k_2 + 2k_3,\ k_1 + 0k_2 + k_3,\ 2k_1 + k_2 + 3k_3)$

$b_1 = k_1 + k_2 + 2k_3$

$b_2 = k_1 + 0k_2 + k_3$

$b_3 = 2k_1 + k_2 + 3k_3$

This system is consistent if and only if its coefficient matrix

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

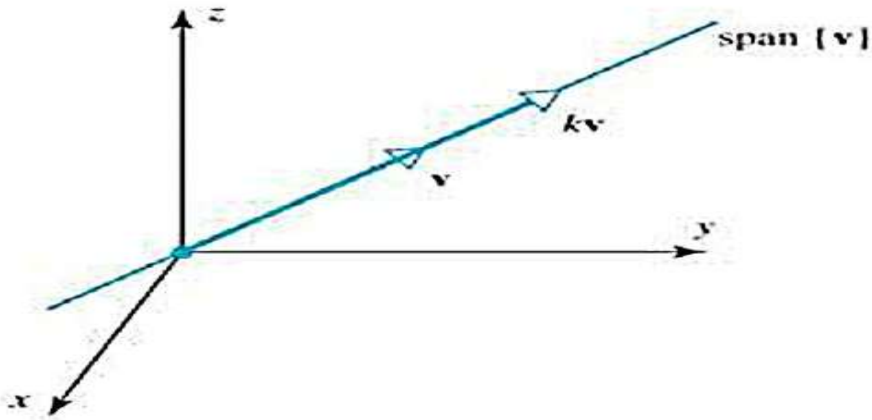
has a non - zero determinant.

But this is not the case here since $det(A) = 0$

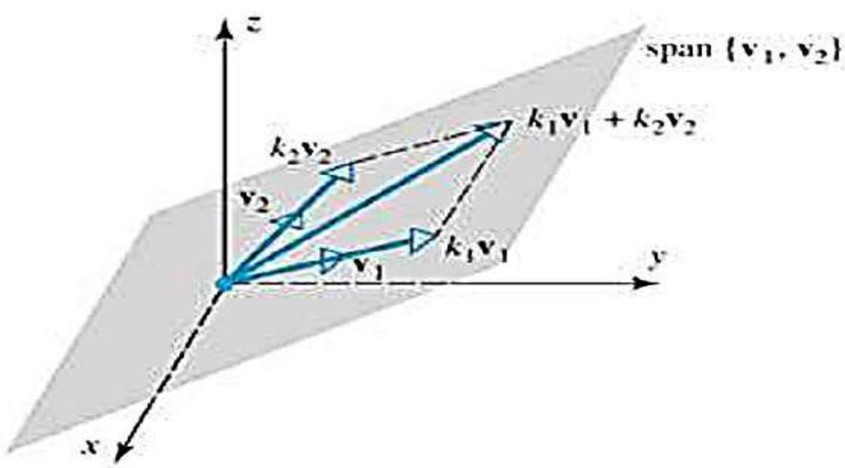
So $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ do not span $\mathbf{R}^3$

A geometric view of spanning in $\mathbf{R}^2$ and $\mathbf{R}^3$

a) If $\mathbf{v}$ is a non – zero vector in $\mathbf{R}^2$ or $\mathbf{R}^3$ that has its initial points at the origin, then span$\{\mathbf{v}\}$, which is the set of all scalar multiples of $\mathbf{v}$, is the line through the origin determined by $\mathbf{v}$.



b) If $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ are non-zero vectors in $\mathbf{R}^3$ that have their initial points at the origin, then span$\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$, which consists of all linear combinations of $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$, is the plane through the origin determined by these two vectors.



PRACTICE:

1) In each part determine whether the vectors span $\mathbf{R}^3$
   a) $\boldsymbol{v_1} = (2,2,2), \boldsymbol{v_2} = (0,0,3), \boldsymbol{v_3} = (0,1,1)$
   b) $\boldsymbol{v_1} = (2,-1,3), \boldsymbol{v_2} = (4,1,2), \boldsymbol{v_3} = (8,-1,8)$

2) Suppose $\boldsymbol{v_1} = (2,1,0,3), \boldsymbol{v_2} = (3,-1,5,2), \boldsymbol{v_3} = (-1,0,2,1)$ then which of the following vectors are in span $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$
   a) $(2,3,-7,3)$
   b) $(0,0,0,0)$
   c) $(1,1,1,1)$
   d) $(-4,6,-13,4)$

3) Show that the yz – plane $W = \{(0, y, z): y, z \in R\}$ is spanned by $(0,1,1)$ and $(0,2,-1)$
4) Find an equation of the subspace **W** of $\mathbf{R}^3$ generated by $\{(1,-3,5), (-2,6,-10)\}$
5) Show that the complex numbers $2 + 3i$ and $1 - 2i$ generate the vector space **C** over **R**.
6) Let $\boldsymbol{v_1} = (1,6,4), \boldsymbol{v_2} = (2,4,-1), \boldsymbol{v_3} = (-1,2,5)$ and $\boldsymbol{w_1} = (1,-2,5), \boldsymbol{w_2} = (0,8,9)$ then show that $Span\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\} = Span\{\boldsymbol{w_1}, \boldsymbol{w_2}\}$
7) Show that $\boldsymbol{u_1} = (1,1,1), \boldsymbol{u_2} = (1,2,3), \boldsymbol{u_3} = (1,5,8)$ span $\mathbf{R}^3$
8) Find conditions on a, b, c so that $\vec{v} = (a, b, c)$ in $\mathbf{R}^3$ belongs to $W = Span(\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3})$ where $\boldsymbol{u_1} = (1,2,0), \boldsymbol{u_2} = (-1,1,2), \boldsymbol{u_3} = (3,0,-4)$
9) Let S and T be subsets of a vector space V. then show that
   i. $\langle S\rangle \cup \langle T\rangle \subset \langle S \cup T\rangle$ but equality does not hold.
   ii. $\langle S \cap T\rangle \subset \langle S\rangle \cap \langle T\rangle$ but equality does not hold.
   iii. $\langle S\rangle \cup \langle T\rangle = \langle S\rangle + \langle T\rangle$
   iv. $\langle\langle S\rangle\rangle = \langle S\rangle$
10) Suppose $\{u_1, u_2, \dots, u_r, w_1, w_2, \dots, w_s\}$ is linearly independent subset of V then show that $\langle u_i\rangle \cap \langle w_j\rangle = \{0\}$. That is $span\{u_i\} \cap span\{w_j\} = \{0\}$

A spanning set for $\boldsymbol{P_n(x)}$

Consider the vector space $\boldsymbol{V} = \boldsymbol{P_n(x)}$ consisting of all polynomials of degree n.

(a) Clearly every polynomial in $\boldsymbol{P_n(x)}$ can be expressed as a linear combination of the $x + 1$ polynomials $1, x, x^2, \dots, x^n$
Thus, these powers of x (where $1 = x^0$) form a spanning set for $\boldsymbol{P_n(x)}$.
We can denote this by writing $P_n(x) = Span\{1, x, x^2, \dots, x^n\}$

(b) One can also show that, for any scalar c, the following $x + 1$ powers of $x - c$, that is $1, x - c, (x - c)^2, \dots, (x - c)^n$ (where $(x - c)^0 = 1$), also form a spanning set for $\boldsymbol{P_n(x)}$.

We can denote this by writing $P_n(x) = Span\{1, x - c, (x - c)^2, \dots, (x - c)^n\}$

(c) Consider the vector space $\boldsymbol{M} = \boldsymbol{M_{2,2}}$ consisting of all 2×2 matrices, and consider the following four matrices in M:
$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Then clearly any matrix **A** in **M** can be written as a linear combination of the four matrices. For example,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Accordingly, the four matrices E_{11}, E_{12}, E_{21}, E_{22} span M.

PRACTICE:

1) Show that a vector space $\boldsymbol{V} = \boldsymbol{P(t)}$ of real polynomials cannot be spanned by a finite number of polynomials.
2) Determine whether the following polynomials span $\boldsymbol{P_2}$
   i. $\boldsymbol{p_1} = 1 - x + 2x^2, \boldsymbol{p_2} = 3 + x$
   $\boldsymbol{p_3} = 5 - x + 4x^2, \boldsymbol{p_4} = -2 - 2x + 2x^2$

Linear Dependence and Independence

Let **V** be a vector space over a field **K**. The following defines the notion of linear dependence and independence of vectors over **K**. (One usually suppresses mentioning **K** when the field is understood.) This concept plays an essential role in the theory of linear algebra and in mathematics in general.

Definition: We say that the vectors $v_1, v_2, \dots, v_n$ in **V** are **linearly dependent** if there exist scalars $a_1, a_2, \dots \dots, a_n$ in **K**, not all of them 0, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$$

On the other hand we say that the vectors $v_1, v_2, \dots, v_n$ in **V** are **linearly independent** if there exist scalars $a_1, a_2, \dots, a_n$ in **K**, all of them 0, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$$

Another definition: If $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r \}$ is a set of two or more vectors in a vector space **V** , then **S** is said to be **linearly independent** set if no vector in **S** can be expressed as a linear combination of the others.

A set that is not linearly independent is said to be **linearly dependent**.

Remark:

- A set $S = \{ v_1, v_2, \dots, v_n \}$ of vectors in **V** is linearly dependent or independent according to whether the vectors $v_1, v_2, \dots, v_n$are linearly dependent or independent.
- An infinite set S of vectors is linearly dependent or independent according to whether there do or do not exist vectors $v_1, v_2, \dots, v_n$in **S** that are linearly dependent.
- **Warning:** The set $S = \{ v_1, v_2, \dots, v_n \}$ above represents a list or, in other words, a finite sequence of vectors where the vectors are ordered and repetition is permitted.
- Suppose **0** is one of the vectors $v_1, v_2, \dots, v_n$, say $v_1 = 0$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$;
  $1v_1 + a_2 v_2 + \cdots + a_n v_n = 1.0 + 0 + \cdots + 0$

- Suppose **v** is a nonzero vector. Then **v**, by itself, is linearly independent, because $kv = 0;\ v \neq 0$ implies $k = 0$
  Implies a single non – zero vector always linearly independent.
- Suppose two of the vectors $v_1, v_2, \dots, v_n$ are equal or one is a scalar multiple of the other, say $v_1 = kv_2$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$;
  $v_1 - kv_2 + 0v_3 + \cdots + 0v_n = 0$
- Two vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ are linearly dependent if and only if one of them is a multiple of the other.
- If any two vectors out of $v_1, v_2, \dots, v_n$ are equal say $v_3 = v_4$ then $v_1, v_2, \dots, v_n$ are linearly independent because
  $0v_1 + 0v_2 + 1.v_3 + (-1)v_4 + 0v_5 + \cdots + 0v_n = 0$
- If the set $\{ v_1, v_2, \dots, v_n\}$ is linearly independent, then any rearrangement of the vectors $\left\{ v_{i_1}, v_{i_2}, \dots, v_{i_n}\right\}$ is also linearly independent.
- If a set **S** of vectors is linearly independent, then any **subset** of **S** is linearly independent. Alternatively, if **S** contains a linearly dependent subset, then **S** is linearly dependent.
- A set $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r\}$ in a vector space **V** is said to be **linearly independent** set iff the only coefficients satisfying the vector equation $a_1v_1 + a_2v_2 + \cdots + a_rv_r = 0$ are $a_1 = 0, a_2 = 0, \dots, a_r = 0$
- Let **V** be a vector space over a field **F** and $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ be a set of vectors in **V**, then if **S** is linearly independent then any subset of **S** is also linearly independent.
- Let **V** be a vector space over a field **F** and $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ be a set of vectors in **V**, then if **S** is linearly dependent then any subset of **S** is also linearly dependent.
- A finite set that contains **0** is linearly dependent.
- A set with exactly one vector is linearly independent iff that vector is not **0**.
- A set with exactly two vectors is linearly independent iff neither vector is a scalar multiple of other.
- If $\{ v_1, v_2, v_3\}$ is linearly independent set of vectors, then so are $\{ v_1, v_2\}, \{ v_1, v_3\}, \{ v_2, v_3\}, \{ v_1\}, \{ v_2\}$ and $\{ v_3\}$

Theorem:

A set $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r\}$ in a vector space **V** is said to be **linearly independent** set iff the only coefficients satisfying the vector equation

$a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathbf{0}$ are $a_1 = 0, a_2 = 0, \dots, a_r = 0$

Proof:

Suppose that $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r\}$ is linearly independent. Then we will show that if the equation $a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathbf{0}$ can be satisfied with coefficients that are not all zero, then at least one of the vectors in **S** must be expressible as the linear combination of the others, thereby contradicting the assumption of linear independence.

To be specific suppose that $a_1 \neq 0$ then we can rewrite the above equation as

$v_1 = -\frac{a_2}{a_1}v_2 - \frac{a_3}{a_1}v_3 - \cdots - \frac{a_r}{a_1}v_r$

Which expresses v_1 as the linear combination of the other vectors in **S**.

Conversely: We must show that if the only coefficients satisfying $a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathbf{0}$ are $a_1 = 0, a_2 = 0, \dots, a_r = 0$ then the vectors in **S** must be linearly independent. But if this were true of the coefficients and the vectors were not linearly independent, then at least one of them would be expressible as a linear combination of the other, say

$v_1 + (-c_2)v_2 + (-c_3)v_3 + \cdots + (-c_r)v_r = \mathbf{0}$

But this contradict our assumption that $a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathbf{0}$ can only be satisfied by coefficients that are all zero.

Thus the vectors in **S** must be linearly independent.

Theorem: (Just Statement): Suppose $\{v_1, v_2, \dots, v_n\}$ spans **V**, and suppose $\{w_1, w_2, \dots, w_m\}$ is linearly independent. Then $m \leq n$, and **V** is spanned by a set of the form $\{w_1, w_2, \dots, w_m, v_{i_1}, v_{i_2}, \dots, v_{i_{n-m}}\}$

Thus, in particular, $n + 1$ or more vectors in **V** are linearly dependent.

Theorem:

Let **V** be a vector space over a field **F** and $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ be a set of vectors in **V**, then if **S** is linearly independent then any subset of **S** is also linearly independent.

Proof:

Here $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ and $\{ v_1, v_2, \dots, v_i\}$; $i < n$ is a subset of **S**.

And let $a_1 v_1 + a_2 v_2 + \cdots \dots + a_i v_i = 0$ where a_i are scalars.

We may write $a_1 v_1 + a_2 v_2 + \cdots + a_i v_i + a_{i+1} v_{i+1} + \cdots \dots + a_n v_n = 0$

But $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ is linearly independent then $a_1 = a_2 =, \dots = a_i = 0$

Hence $\{ v_1, v_2, \dots, v_i\}$; $i < n$ is linearly independent.

Theorem:

Let **V** be a vector space over a field **F** and $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ be a set of vectors in **V**, then if **S** is linearly dependent then any subset of **S** is also linearly dependent.

Proof:

Here $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ and $\{v, v_1, v_2, \dots, v_n\}$ is a subset of **S**.

Since $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ is linearly dependent then

$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$ where $a_i \neq 0$ for some i

Now $0v + a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$ where $a_i \neq 0$ for some i

Then $\{v, v_1, v_2, \dots, v_n\}$ is linearly dependent.

Theorem:

A set $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ of 'n' vectors $(n \geq 2)$ in a vector space V is linearly dependent iff atleast one of the vectors in **S** is a linear combination of the remaining vectors of the set.

Proof: Suppose the set $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ is linearly dependent. Then there exists scalars $a_1, a_2, \dots \dots, a_n$ at least one of them say a_i is non – zero such that

$$a_1v_1 + a_2v_2 + \cdots + a_iv_i + \cdots + a_nv_n = 0$$

Or $\quad a_iv_i = -a_1v_1 - a_2v_2 - \cdots - a_{i-1}v_{i-1} - a_{i+1}v_{i+1} - \cdots - a_nv_n$

Or $\quad v_i = -\frac{a_1}{a_i}v_1 - \frac{a_2}{a_i}v_2 - \cdots - \frac{a_{i-1}}{a_i}v_{i-1} - \frac{a_{i+1}}{a_i}v_{i+1} - \cdots - \frac{a_n}{a_i}v_n$

Which shows that v_i is a linear combination of the remaining vectors of the set.

Conversely:

Suppose that some vector v_j of the given set is a linear combination of the remaining vectors of the set. i.e.

$$v_j = a_1v_1 + a_2v_2 + \cdots + a_{j-1}v_{j-1} + a_{j+1}v_{j+1} + \cdots + a_nv_n$$

Then above equation can be written as

$$a_1v_1 + a_2v_2 + \cdots + a_{j-1}v_{j-1} + (-1)v_j + a_{j+1}v_{j+1} + \cdots + a_nv_n = 0$$

Here there is at least one coefficient namely -1 of v_j which is non – zero and so that $\left\{ v_1, v_2, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n\right\}$ is linearly dependent.

Theorem:

A set $\boldsymbol{S} = \{ v_1, v_2, \dots \dots, v_n\}$ in a vector space V is linearly dependent iff some of the vectors say v_k is a linear combination of the vectors preceding it.

Proof: Suppose the set $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$ is linearly dependent. Then there exists scalars $a_1, a_2, \dots, a_n$ at least one of them say a_i is non – zero such that

$a_1v_1 + a_2v_2 + \cdots + a_iv_i + \cdots + a_nv_n = 0$ ……….(i)

Let a_k be the last non – zero scalar in (i) then the terms $a_{k+1}v_{k+1}, a_{k+2}v_{k+2}, \dots., a_nv_n$ are all zeros.

So the equation (i) becomes

$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$ where $a_k \neq 0$

Or $\quad v_k = -\frac{a_1}{a_k}v_1 - \frac{a_2}{a_k}v_2 - \cdots - \frac{a_{k-1}}{a_k}v_{k-1}$

Which shows that v_k is a linear combination of the vectors preceding it.

Conversely:

Suppose that in $\boldsymbol{S} = \{ v_1, v_2, \dots, v_n\}$, some of the vectors say v_k is a linear combination of the vectors preceding it.

$v_k = b_1v_1 + b_2v_2 + \cdots + \cdots + b_nv_n$

Then above equation can be written as

$b_1v_1 + b_2v_2 + \cdots + b_{k-1}v_{k-1} + (-1)v_k = 0$

Or $\quad b_1v_1 + b_2v_2 + \cdots + b_{k-1}v_{k-1} + (-1)v_k + 0v_{k+1} + \cdots + 0v_n = 0$

Here there is at least one coefficient namely -1 of v_k which is non – zero and so that $\{ v_1, v_2, \dots, v_n\}$ is linearly dependent.

Theorem:

Let $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r\}$ be a set of vector in $\mathbf{R}^{\mathbf{n}}$. if $r > n$ then S is linearly dependent.

Proof: Suppose that

$\boldsymbol{v_1} = (v_{11}, v_{12}, \dots, v_{1n})$

$\boldsymbol{v_2} = (v_{21}, v_{22}, \dots, v_{2n})$

$\vdots$

$\boldsymbol{v_r} = (v_{r1}, v_{r2}, \dots, v_{rn})$

And consider the equation $k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + \cdots + k_r\boldsymbol{v_r} = \mathbf{0}$

If we express both sides of this equation in terms of components and then equate the corresponding components, we obtain the system;

$$
\begin{array}{l}
k_1 v_{11} + k_2 v_{21} + \cdots + k_r v_{r1} = 0 \\
k_1 v_{12} + k_2 v_{22} + \cdots + k_r v_{r2} = 0 \\
\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\
k_1 v_{1n} + k_2 v_{2n} + \cdots + k_r v_{rn} = 0
\end{array}
$$

This is the homogeneous system of ‘n’ equations in ‘r’ unknowns $k_1, k_2, \dots \dots, k_r$

Since $r > n$.

It follows from theorem “ *a homogeneous linear system with more unknowns than equations has infinitely many solution*” that the system has non – trivial solution. Therefore $\boldsymbol{S} = \{ v_1, v_2, \dots, v_r\}$ is linearly dependent.

Linear independence of the Standard unit vectors in $\mathbf{R}^n$

In $\mathbf{R}^n$ the most basic linearly independent set is the set of standard unit vectors

$\boldsymbol{e_1} = (1,0,,\dots,0)$, $\boldsymbol{e_2} = (0,1,0,\dots,0)$, ………$\boldsymbol{e_n} = (0,0,0,\dots,1)$

Now we consider the standard unit vectors in $\mathbf{R}^3$

$\boldsymbol{i} = (1,0,0)$, $\boldsymbol{j} = (0,1,0)$, $\boldsymbol{k} = (0,0,1)$

To prove linearly independent we must show that the only coefficient satisfying the vector equation $k_1\boldsymbol{i} + k_2\boldsymbol{j} + k_3\boldsymbol{k} = 0$ are $k_1 = 0, k_2 = 0, k_3 = 0$ But this become evident by writing the equation in its component form $(k_1, k_2, k_3) = (0,0,0)$

Linear independence in $\mathbf{R}^3$

Determine whether the vectors $\boldsymbol{v_1} = (1,-2,3), \boldsymbol{v_2} = (5,6,-1), \boldsymbol{v_3} = (3,2,1)$ are linearly independent or linearly dependent in $\mathbf{R}^3$

Solution: Consider $k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + k_3\boldsymbol{v_3} = 0$

Rewriting in component form $k_1(1,-2,3) + k_2(5,6,-1) + k_3(3,2,1) = (0,0,0)$

Equating corresponding components on the two sides yields the homogeneous linear system

$$k_1 + 5k_2 + 3k_3 = 0$$

$$-2k_1 + 6k_2 + 2k_3 = 0$$

$$3k_1 - k_2 + k_3 = 0$$

After solving the system we get $k_1 = -\frac{1}{2}t$, $k_2 = -\frac{1}{2}t$, $k_3 = t$

This shows that the system has non – trivial solution and hence that the vectors are linearly dependent.

Linear independence in $\mathbf{R}^4$

Determine whether the vectors

$\boldsymbol{v_1} = (1,2,2,-1), \boldsymbol{v_2} = (4,9,9,-4), \boldsymbol{v_3} = (5,8,9,-5)$ are linearly independent or linearly dependent in $\mathbf{R}^4$

Solution: Consider $k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + k_3\boldsymbol{v_3} = 0$

Rewriting in component form

$k_1(1,2,2,-1) + k_2(4,9,9,-4) + k_3(5,8,9,-5) = (0,0,0,0)$

Equating corresponding components on the two sides yields the homogeneous linear system

$k_1 + 4k_2 + 5k_3 = 0$

$2k_1 + 9k_2 + 8k_3 = 0$

$2k_1 + 9\ k_2\ + 9k_3 = 0$

$-k_1 - 4\ k_2\ - 5k_3 = 0$

After solving the system we get $k_1 = 0\ ,\ k_2 = 0\ ,\ k_3 = 0$

This shows that the system has trivial solution and hence that the vectors are linearly independent.

Example:

Let $\boldsymbol{u} = (1,1,0), \boldsymbol{v} = (1,3,2), \boldsymbol{w} = (4,9,5)$ .Then u, v, w are linearly dependent, because $3\boldsymbol{u} + 5\boldsymbol{v} - 2\boldsymbol{w} = 3(1,1,0) + 5(1,3,2) - 2(4,9,5) = (0,0,0) = 0$

Example: We show that the vectors $\boldsymbol{u} = (1,2,3), \boldsymbol{v} = (2,5,7), \boldsymbol{w} = (1,3,5)$ are linearly independent. We form the vector equation $x\boldsymbol{u} + y\boldsymbol{v} + z\boldsymbol{w} = 0$,
where x, y, z are unknown scalars. This yield

$$x\begin{bmatrix}1\\2\\3\end{bmatrix} + y\begin{bmatrix}2\\5\\7\end{bmatrix} + z\begin{bmatrix}1\\3\\5\end{bmatrix} = 0$$

Or

$$x + 2y + z = 0$$
$$2x + 5y + 3z = 0$$
$$3x + 7y + 5z = 0$$

Or

$$x + 2y + z = 0$$
$$y + z = 0$$
$$2z = 0$$

Back-substitution yields $x = 0, y = 0, z = 0$. We have shown that

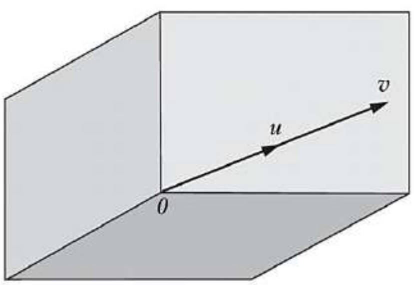
$x\boldsymbol{u} + y\boldsymbol{v} + z\boldsymbol{w} = 0$ implies $x = 0, y = 0, z = 0$

Accordingly, u, v, w are linearly independent.

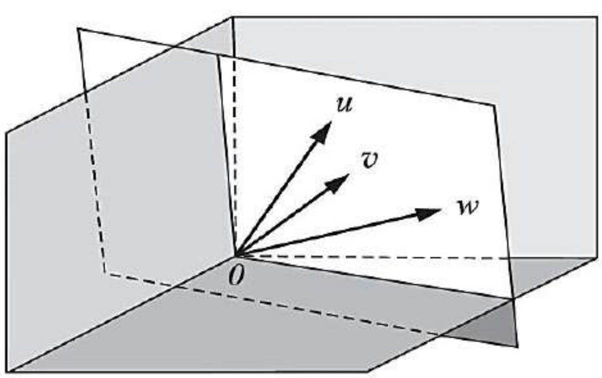
Linear Dependence in R^3

Linear dependence in the vector space $\boldsymbol{V} = \boldsymbol{R^3}$ can be described geometrically as follows:

(a) Any two vectors **u** and **v** in $\boldsymbol{R^3}$ are linearly dependent (not independent) if and only if they lie on the same line through the origin O, as shown in Fig.



(b) Any three vectors **u**, **v**, **w** in $\boldsymbol{R^3}$ are linearly dependent (not independent) if and only if they lie on the same plane through the origin O, as shown in Fig.



Practice:

1) In each part, determine whether the vectors are linearly independent or are linearly dependent in $\boldsymbol{R}^3$

   a) $(-3,0,4),(5,-1,2),(1,1,3)$
   b) $(-2,0,1),(3,2,5),(6,-1,1),(7,0,-2)$

2) In each part, determine whether the vectors are linearly independent or are linearly dependent in $\boldsymbol{R}^4$

   a) $(3,8,7,-3),(1,5,3,-1),(2,-1,2,6)$
   b) $(3,0,-3,6),(0,2,3,1),(0,-2,-2,0),(-2,1,2,1)$
   c) $(0,3,1,-1),(6,0,5,1),(4,-7,1,3)$
   d) $(1,2,3,4),(0,1,0,-1),(1,3,3,3)$

3) Express each vector in $(3,8,7,-3),(1,5,3,-1),(2,-1,2,6)$ as a linear combination of other.
4) For which real value of λ do the following vectors form a linearly dependent set in $\boldsymbol{R}^3$?
   $\vec{v}_1 = \left(\lambda, -\frac{1}{2}, -\frac{1}{2}\right), \vec{v}_2 = \left(-\frac{1}{2}, \lambda, -\frac{1}{2}\right), \vec{v}_3 = \left(-\frac{1}{2}, -\frac{1}{2}, \lambda\right)$

Or determine λ so that the above vectors are linearly dependent in $\boldsymbol{R}^3$.

5) Show that the vectors $(1-i, i),(2,-1+i)$ in $\mathbf{C}^2$ are linearly dependent over **C** but linearly independent over **R**.
6) Show that the vectors $\left(3+\sqrt{2}, 1+\sqrt{2}\right),\left(7, 1+2\sqrt{2}\right)$ in $\mathbf{R}^2$ are linearly dependent over **R** but linearly independent over **Q**.
7) Show that the vectors $(1+i, 2i),(1, 1+i)$ in $\mathbf{C}^2$ are linearly dependent over the complex field **C** but linearly independent over the real field **R**.
8) Suppose that $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly independent vectors.
   Prove that $\boldsymbol{u}+\boldsymbol{v}-\mathbf{2}\boldsymbol{w}$ is linearly independent.
9) Show that for any vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in a vector space **V**, the vectors $\boldsymbol{u}-\boldsymbol{v}$, $\boldsymbol{v}-\boldsymbol{w}, \boldsymbol{w}-\boldsymbol{u}$ form a linearly dependent set.
10) Under what conditions is a set with one vector linearly independent?

Linear independence for Polynomials

Consider the vector space $\boldsymbol{V} = \boldsymbol{P_n(x)}$ consisting of all polynomials of degree n. Show that the polynomials $1, x, x^2, \dots., x^n$ in $\boldsymbol{P_n(x)}$ form a linearly independent set.

Solution: Let $\boldsymbol{p_0} = 1, \boldsymbol{p_1} = x, \boldsymbol{p_2} = x^2, \dots., \boldsymbol{p_n} = x^n$

And consider $a_0\boldsymbol{p_0} + a_1\boldsymbol{p_1} + \cdots + a_n\boldsymbol{p_n} = 0$

Equivalently $a_0 + a_1x + \cdots + a_nx^n = 0$ for all 'x' in $(-\infty, \infty)$

Since we know that a non – zero polynomial of degree 'n' has atmost 'n' distinct roots. Then in this case all coefficients in above expression must be zero. For otherwise the left side of the equation would be a non – zero polynomial with infinity many roots. Thus above equation has only the trivial solution.

Implies $\boldsymbol{P_n(x)}$ is the linearly independent set.

Example:

Determine whether the polynomials

$\boldsymbol{p_1} = 1 - x, \boldsymbol{p_2} = 5 + 3x - 2x^2, \boldsymbol{p_3} = 1 + 3x - x^2$

Are linearly dependent or linearly independent in $\boldsymbol{P_2}$

Solution: Consider $k_1\boldsymbol{p_1} + k_2\boldsymbol{p_2} + k_3\boldsymbol{p_3} = 0$

Equivalently $k_1(1 - x) + k_2(5 + 3x - 2x^2) + k_3(1 + 3x - x^2) = 0$

$\Rightarrow (k_1 + 5k_2 + k_3) + (-k_1 + 3k_2 + 3k_3)x + (0k_1 - 2k_2 - k_3)x^2 = 0$

Since this equation must be satisfied by all 'x' in $(-\infty, \infty)$, each coefficient must be zero. Thus the linear independence or linear dependence of the given polynomials hinges on whether the following linear system has a non – trivial solution;

$k_1 + 5k_2 + k_3 = 0,\ -k_1 + 3k_2 + 3k_3 = 0,\ k_1 - 2k_2 - k_3 = 0$

After solving we will get $k_1 = k_2 = k_3 = 0$. And hence given polynomials form linearly dependent set.

Example: Determine whether the polynomials

$\boldsymbol{p_1} = x^3 - 4x^2 + 2x + 3, \boldsymbol{p_2} = x^3 + 2x^2 + 4x - 1, \boldsymbol{p_3} = 2x^3 - x^2 - 3x + 3$

Are linearly dependent or linearly independent in $\boldsymbol{P_2}$

Solution: Consider $k_1\boldsymbol{p_1} + k_2\boldsymbol{p_2} + k_3\boldsymbol{p_3} = 0$

$k_1(x^3 - 4x^2 + 2x + 3) + k_2(x^3 + 2x^2 + 4x - 1) + k_3(2x^3 - x^2 - 3x + 3) = 0$

$\Rightarrow (k_1 + k_2 + 2k_3)x^3 + (-4k_1 + 2k_2 - k_3)x^2 + (2k_1 + 4k_2 - 3k_3)x$

$+(3k_1 - k_2 + 3k_3) = 0$

$k_1 + k_2 + 2k_3 = 0$ ....(i)

$-4k_1 + 2k_2 - k_3 = 0$ ....(ii)

$2k_1 + 4k_2 - 3k_3 = 0$ ....(iii)

$3k_1 - k_2 + 3k_3 = 0$ ....(iv)

After solving we will get $k_1 = k_2 = k_3 = 0$. As follows;

$\frac{k_1}{-1-4} = \frac{-k_2}{-1+8} = \frac{k_3}{2+4}$

$\frac{k_1}{-5} = \frac{k_2}{-7} = \frac{k_3}{6} = k$

Implies $k_1 = -5k, \; k_2 = -7k, \; k_3 = 6k$

Putting these values in (iii) and (iv) we see equations are not satisfied. They are satisfied only when $k_1 = k_2 = k_3 = 0$.

Hence given polynomials are linearly independent.

Remember: In $\boldsymbol{P_2}$ every set with more than three vectors is linearly dependent.

Practice: In each part, determine whether the vectors are linearly independent or are linearly dependent in $\boldsymbol{P_2}$

a) $2 - x + 4x^2$, $3 + 6x + 2x^2$, $2 + 10x - 4x^2$
b) $1 + 3x + 3x^2$, $5 + 6x + 3x^2$, $7 + 2x - x^2$

Linear independence of functions:

Example: Let **V** be the real vector space of all functions defined on **R** into **R**. determine whether the given vectors $f(t) = Sint, g(t) = e^t, h(t) = t^2$ are linearly independent or linearly depended in **V**.

Solution: Consider $x\boldsymbol{f} + y\boldsymbol{g} + z\boldsymbol{h} = 0$ where x,y,z are unknown scalar.

This implies $xSint + ye^t + zt^2 = 0$

Thus in this equation we choose appropriate values of 't' to easily get $x = 0, y = 0, z = 0$ for example

i. Substitute $t = 0$ to obtain $x(0) + y(1) + z(0) = 0$ or $y = 0$
ii. Substitute $t = \pi$ to obtain $x(0) + y(e^{\pi}) + z(\pi^2) = 0$ or $z = 0$
iii. Substitute $t = \frac{\pi}{2}$ to obtain $x(0) + y\left(e^{\pi/2}\right) + z\left(\frac{\pi^2}{4}\right) = 0$ or $x = 0$

We have shown $xSint + ye^t + zt^2 = 0$ implies $x = 0, y = 0, z = 0$

Thus given vectors are linearly independent.

Wronskian of Functions:

If $\boldsymbol{f_1} = f_1(x), \boldsymbol{f_2} = f_2(x), \dots, \boldsymbol{f_n} = f_n(x)$ are functions that are n – 1 time differentiable on the interval $(-\infty, \infty)$, then the determinant

$$\boldsymbol{W}(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1{}'(x) & f_2{}'(x) & \cdots & f_n{}'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1{}^{n-1}(x) & f_2{}^{n-1}(x) & \cdots & f_n{}^{n-1}(x) \end{vmatrix}$$

Is called **Wronskian** of $\boldsymbol{f_1} = f_1(x), \boldsymbol{f_2} = f_2(x), \dots, \boldsymbol{f_n} = f_n(x)$

Remember:

Sometime linear dependence of functions can be deduced from known identities. However, it is relatively rare that linear independence or dependence of functions can be ascertained by algebraic or trigonometric methods. To make matter worse, there is no general method for doing that either.

Theorem:

If the functions $\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_n}$ have n – 1 continuous derivatives on the interval $(-\infty, \infty)$ and if the Wronskian of these functions is not identically zero on $(-\infty, \infty)$, then these functions form a linearly independent set of vectors in $C^{n-1}(-\infty, \infty)$ but the converse of this theorem is false.

Example:

Use the Wronskian to show that $\boldsymbol{f_1} = x, \boldsymbol{f_2} = Sinx$ are linearly independent vectors in $C^{\infty}(-\infty, \infty)$.

Solution:

The Wronskian of given functions is as follows;

$$\boldsymbol{W}(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix} = \begin{vmatrix} x & Sinx \\ 1 & Cosx \end{vmatrix} = xCosx - Sinx$$

Consider $\boldsymbol{W}\left(\frac{\pi}{2}\right) = \frac{\pi}{2} Cos\left(\frac{\pi}{2}\right) - Sin\left(\frac{\pi}{2}\right) = \frac{\pi}{2} \neq 0$

This function is not identically zero on the interval $(-\infty, \infty)$, thus , the functions are linearly independent.

Example:

Use the Wronskian to show that $\boldsymbol{f_1} = 1, \boldsymbol{f_2} = e^x, \boldsymbol{f_3} = e^{2x}$ are linearly independent vectors in $C^{\infty}(-\infty, \infty)$.

Solution:

The Wronskian of given functions is as follows;

$$\boldsymbol{W}(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) \end{vmatrix} = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x} \neq 0$$

This function is not identically zero on the interval $(-\infty, \infty)$, thus , the functions are linearly independent.

Practice:

1) Let **V** be the real vector space of all functions defined on **R** into **R**. determine whether the given vectors are linearly independent or linearly depended in **V**.
   i. $x,\ Cosx$
   ii. $Sin^2x,\ Cos^2x,\ Cos2x$
   iii. $Sin^2x,\ Cos^2x,\ 5$
   iv. $Sinx,\ Cosx,\ Sinhx,\ Coshx$
   v. $Sinx,\ Sinx + Cosx,\ Sinx - Cosx$
2) By using appropriate identities, where required, determine which of the following sets of vectors in $F(-\infty, \infty)$ are linearly dependent.
   i. $6,\ \ 3Sin^2x,\ \ 2Cos^2x$
   ii. $1,\ Sinx,\ \ Sin2x$
   iii. $(3-x)^2,\ \ 5,\ \ x^2 - 6x$
   iv. $0,\ \ Sin^5 3\pi x,\ \ Cos^3 \pi x$

3) Use the Wronskian to show that given functions are linearly independent vectors in $C^\infty(-\infty, \infty)$.
   i. $x\ ,\ Cosx$
   ii. $Sinx\ ,\ Cosx$
   iii. $1\ ,\ x\ ,\ e^x$
   iv. $1\ ,\ x\ ,\ x^2$
   v. $e^x\ ,\ xe^x\ ,\ x^2e^x$
   vi. $Sinx\ ,\ Cosx\ ,\ xCosx$

4) Using the technique of casting out vectors which are linear combination of others, find a linearly independent subset of the given set spanning the same subspace;
   i. $\{(1,-3,1), (2,1,-4), (-2,6,-2), (-1,10,-7)\}$ in $\mathbf{R}^3$
   ii. $\{1, Sin^2x, Cos^2x, Cos2x\}$ in the space of all functions from **R** to **R**
   iii. $\{1, 3x-4, 4x+3, x^2+2, x-x^2\}$ in the space $\boldsymbol{P_2}$ of all polynomials

Basis

A set $\boldsymbol{S} = \{v_1, v_2, \dots, v_n\}$ of vectors is a basis of a finite dimensional vector space **V** if it has the following two properties:

i. **S** is linearly independent.
ii. **S** spans **V**.

Or

A set $\boldsymbol{S} = \{v_1, v_2, \dots, v_n\}$ of vectors is a basis of a finite dimensional vector space **V** if every $\vec{v}\epsilon\boldsymbol{V}$ can be written uniquely as a linear combination of the basis vectors.

Examples of Bases

This subsection presents important examples of bases of some of the main vector spaces appearing in this text.

Usual or the Standard basis for $\mathbf{R^n}$

In $\mathbf{R^n}$ the most basic **linearly independent** set is the set of standard unit vectors

$\boldsymbol{e_1} = (1,0,0, \dots ,0)$, $\boldsymbol{e_2} = (0,1,0, \dots ,0)$, ………., $\boldsymbol{e_n} = (0,0,0, \dots ,1)$

Thus they form a basis for $\mathbf{R^n}$ that we call **the Standard basis for $\mathbf{R^n}$**

For example, any vector $\boldsymbol{u} = (a_1, a_2, \dots, a_n)$in $\mathbf{R^n}$ can be written as a linear combination of the above vectors. i.e. $\boldsymbol{u} = a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2} + \dots + a_n\boldsymbol{e_n}$

In particular we consider the standard unit vectors in $\mathbf{R^3}$

$\boldsymbol{i} = (1,0,0)$, $\boldsymbol{j} = (0,1,0)$, $\boldsymbol{k} = (0,0,1)$

And we call **the Standard basis for $\mathbf{R^3}$**

Remark:

- The number of elements in a basis of a vector space **V** over **F** is called dimension of **V**. It is denoted by $dim\boldsymbol{V}$.
- For n – dimensional vector space **V**, every set $\boldsymbol{S} = \{v_1, v_2, \dots, v_n\}$ of 'n' linearly independent vectors forms a basis for **V**.

Example: Show that the vectors $\boldsymbol{v_1} = (1,2,-1), \boldsymbol{v_2} = (0,3,1), \boldsymbol{v_3} = (1,-5,3)$ form a basis for $\mathbf{R}^3$

Solution: We must show that these vectors are linearly independent and span $\mathbf{R}^3$ To prove linear independence we must show that the vector equation $k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + k_3\boldsymbol{v_3} = 0$ has only the trivial solution.

Then $k_1(1,2,-1) + k_2(0,3,1) + k_3(1,-5,3) = 0$

By equating corresponding components on the two sides, we get a linear system;

$k_1 + 0k_2 + k_3 = 0$ …………..(i)

$2k_1 + 3k_2 - 5k_3 = 0$ …………..(ii)

$-k_1 + k_2 + 3k_3 = 0$ …………..(iii)

From (ii) and (iii)

$\frac{k_1}{9+5} = \frac{-k_2}{6-5} = \frac{k_3}{2+3} = k \Rightarrow \frac{k_1}{14} = \frac{k_2}{-1} = \frac{k_3}{5} = k \Rightarrow k_1 = 14k, k_2 = -k, k_3 = 5k$

Putting these values in equation (i), we see equation (i) is not satisfied. It is satisfied only when k = 0. i.e. $k_1 = 0, k_2 = 0, k_3 = 0$ Hence given vectors $\boldsymbol{v_1} = (1,2,-1), \boldsymbol{v_2} = (0,3,1), \boldsymbol{v_3} = (1,-5,3)$ are linearly independent.

Since dimension of $\mathbf{R}^3$ is '3' and the number of linearly independent vectors in $\mathbf{R}^3$ is also '3', so the vectors $\boldsymbol{v_1} = (1,2,-1), \boldsymbol{v_2} = (0,3,1), \boldsymbol{v_3} = (1,-5,3)$ forms a basis for $\mathbf{R}^3$

Practice:

1) Show that the given vectors may or may not form a basis for $\mathbf{R}^2$ or $\mathbf{R}^3$.
   i. $\boldsymbol{v_1} = (1,2,1), \boldsymbol{v_2} = (2,9,0), \boldsymbol{v_3} = (3,3,4)$
   ii. $\boldsymbol{v_1} = (2,1), \boldsymbol{v_2} = (3,0)$
   iii. $\boldsymbol{v_1} = (3,1,-4), \boldsymbol{v_2} = (2,5,6), \boldsymbol{v_3} = (1,4,8)$
   iv. $\boldsymbol{v_1} = (2,-3,1), \boldsymbol{v_2} = (4,1,1), \boldsymbol{v_3} = (0,-7,1)$
   v. $\boldsymbol{v_1} = (1,6,4), \boldsymbol{v_2} = (2,4,-1), \boldsymbol{v_3} = (-1,2,5)$
2) In words explain why the matrices $(1,2), (0,3), (1,5)$ and $(-1,3,2), (6,1,1)$ are not basis for $\mathbf{R}^2$ and $\mathbf{R}^3$ respectively.

Usual or the Standard basis for $P_n(x)$

Consider the vector space $V = P_n(x)$ consisting of all polynomials of degree n or less then the set $\{1, x, x^2, \dots, x^n\}$ is a basis for $V = P_n(x)$

For this we must show that given polynomials in **S** are linearly independent and span $P_n(x)$.

Clearly every polynomial p in $P_n(x)$ can be expressed as a linear combination of the $x + 1$ polynomials $1, x, x^2, \dots, x^n$ that is $p = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$. Thus, these powers of x (where $1 = x^0$) form a spanning set for $P_n(x)$ .We can denote this by writing $P_n(x) = Span\{1, x, x^2, \dots, x^n\}$

Now we have to show that the polynomials $1, x, x^2, \dots, x^n$ in $P_n(x)$ form a linearly independent set.

For this let $p_0 = 1, p_1 = x, p_2 = x^2, \dots, p_n = x^n$

And consider $\quad a_0p_0 + a_1p_1 + \dots + a_np_n = 0$

Equivalently $\quad a_0 + a_1x + \dots + a_nx^n = 0$ for all 'x' in $(-\infty, \infty)$

Since we know that a non – zero polynomial of degree 'n' has atmost 'n' distinct roots. Then in this case all coefficients in above expression must be zero. For otherwise the left side of the equation would be a non – zero polynomial with infinity many roots. Thus above equation has only the trivial solution. Implies $P_n(x)$ is the linearly independent set.

Thus the set $\{1, x, x^2, \dots, x^n\}$ is a basis for $V = P_n(x)$

Practice:

1) Show that the given polynomials may or may not form a basis for P_2 or P_3.
   i. $1 - 3x + 2x^2$, $1 + x + 4x^2$, $1 - 7x$
   ii. 1 , $2x$, $-2 + 4x^2$, $-12x + 8x^3$ **Hermite Polynomials**
   iii. 1 , $1 - x$, $2 - 4x + x^2$, $6 - 18x + 9x^2 - x^3$ **Laguerre Polynomials**
2) Show that $\{Cos^2x, Sin^2x, Cos2x\}$ is not a basis. Find a basis for vector space **V** spanned by these polynomials.
3) In words explain why the polynomials $1 + x + x^2$, x are not basis for P_2

Usual or the Standard basis for $\boldsymbol{M_{mn}}$

Show that the following matrices form a basis of the vector space $\boldsymbol{M_{22}}$ of all 2×2 matrices over **K**:

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Solution:

We must show that the given matrices are linearly independent and span $\boldsymbol{M_{22}}$

To prove linear independence we must show that the equation
$k_1\boldsymbol{M_1} + k_2\boldsymbol{M_2} + k_3\boldsymbol{M_3} + k_4\boldsymbol{M_4} = \boldsymbol{0}$ has only the trivial solution, where $\boldsymbol{0}$ is the 2×2 zero matrix.

Consider $\quad k_1\boldsymbol{M_1} + k_2\boldsymbol{M_2} + k_3\boldsymbol{M_3} + k_4\boldsymbol{M_4} = \boldsymbol{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$$k_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + k_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + k_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + k_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Above equation has only trivial solution. i.e. $k_1 = k_2 = k_3 = k_4 = 0$

Given matrices are linearly independent.

To prove matrices span $\boldsymbol{M_{22}}$

Consider $\quad k_1\boldsymbol{M_1} + k_2\boldsymbol{M_2} + k_3\boldsymbol{M_3} + k_4\boldsymbol{M_4} = \boldsymbol{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$k_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + k_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + k_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + k_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\qquad$ i.e. given matrices span $\boldsymbol{M_{22}}$

This show that $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Form a basis of the vector space $\boldsymbol{M_{22}}$ of all 2×2 matrices over **K**

Remark:

- Vector space $M = \boldsymbol{M_{mn}}$ of all $m \times n$ matrices: The following six matrices form a basis of the vector space $\boldsymbol{M_{23}}$ of all 2×3 matrices over **K**:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad M_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$M_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad M_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad M_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- More generally, in the vector space $M = \boldsymbol{M_{mn}}$ of all $m \times n$ matrices, let E_{ij} be the matrix with ij-entry 1 and 0's elsewhere. Then all such matrices form a basis of $\boldsymbol{M_{mn}}$ called the **usual or standard basis** of $\boldsymbol{M_{mn}}$

Practice:

1) Show that the given matrices may or may not form a basis for $\boldsymbol{M_{22}}$ **.**

   i. $\begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

   ii. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

   iii. $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$

2) In words explain why the matrices $\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 6 & 0 \\ -1 & 4 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 7 \end{bmatrix}, \begin{bmatrix} 5 & 0 \\ 4 & 2 \end{bmatrix}$ are not basis for $\boldsymbol{M_{22}}$

Theorem (Uniqueness of Basis representation):

If $\boldsymbol{S} = \{v_1, v_2, \ldots, v_n\}$ is a basis for a vector space **V**, then every vector **v** in **V** can be expressed in the form $\boldsymbol{v} = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$ in *exactly one way*.

Or let $\boldsymbol{S} = \{v_1, v_2, \ldots, v_n\}$ in **V** is linearly independent then each **v** of $V \subseteq L(S)$ of **S** is *uniquely expressible*.

Proof:

Since **S** spans **V**, if follows from the definition of a spanning set that every vector in **V** is expressible as a linear combination of the vectors in **S**. To see that there is only one way to express a vector as a linear combination of the vector in **S**, suppose that some vector **v** can be written as;

$\boldsymbol{v} = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$ ……………..(i)

And also as $\boldsymbol{v} = k_1 v_1 + k_2 v_2 + \ldots + k_n v_n$ ……………..(ii)

Subtracting the second equation from the first;

$\mathbf{0} = (c_1 - k_1)v_1 + (c_2 - k_2)v_2 + \ldots + (c_n - k_n)v_n$

Since the right side of the equation is the linear combination of vectors in **S**, and **S** is linearly independent then;

$(c_1 - k_1) = 0, (c_2 - k_2) = 0, \ldots, (c_n - k_n) = 0$

That is $c_1 = k_1, \quad c_2 = k_2, \quad \ldots, \quad c_n = k_n$

Thus the two expressions for **v** are the same (unique).

Theorem:

Any finite dimensional vector space contains a basis.

Proof:

Let **V** be a finite dimensional vector space then **V** should be linear span of some finite set. Let $\{v_1, v_2, \dots, v_r\}$ be a finite spanning set of **V**. in case $v_1, v_2, \dots, v_n$ are linearly independent, then they form a basis of **V** and the proof is complete.

Suppose $v_1, v_2, \dots, v_r$ are not linearly independent. i.e. they are linearly dependent, so one of the vectors v_i is a linear combination of the preceding vectors. We drop this vector v_i from the set and obtain a set of $r-1$ vectors, $v_1, v_2, \dots, v_{r-1}$. Then clearly any linear combination of $v_1, v_2, \dots, v_r$ is also a linear combination of $v_1, v_2, \dots, v_{r-1}$. So $\{v_1, v_2, \dots, v_{r-1}\}$ is also a spanning set for **V**. continuing in this way, we arrive at a linearly independent spanning set $\{v_1, v_2, \dots, v_n\}$ such that $1 \le n \le r$ and so it forms a basis for **V**.

Thus every finite dimensional vector space contains a basis.

Theorem:

Let **V** be a vector space of finite dimension 'n'. Then, any $n+1$ or more vectors in **V** are linearly dependent.

Proof:

Suppose $\boldsymbol{B} = \{w_1, w_2, \dots, w_n\}$ is a basis of **V**. because **B** spans **V**, then by lemma **'Suppose $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}\}$ spans V, and suppose $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_m}\}$ is linearly independent. Then $\boldsymbol{m} \le \boldsymbol{n}$, and V is spanned by a set of the form $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_m}, \boldsymbol{v_{i_1}}, \boldsymbol{v_{i_2}}, \dots, \boldsymbol{v_{i_{n-m}}}\}$'**

Thus, in particular, $n+1$ or more vectors in **V** are linearly dependent.

Theorem:

Let **V** be a vector space of finite dimension ‘n’. Then, any linearly independent set $\boldsymbol{S} = \{v_1, v_2, \ldots, v_n\}$ with ‘n’ elements is a basis of **V**.

Proof:

Suppose $\boldsymbol{B} = \{w_1, w_2, \ldots, w_n\}$ is a basis of **V**. then by lemma

‘Suppose $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ spans V, and suppose $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_m}\}$ is linearly independent. Then $\boldsymbol{m} \leq \boldsymbol{n}$, and V is spanned by a set of the form $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_m}, \boldsymbol{v_{i_1}}, \boldsymbol{v_{i_2}}, \ldots, \boldsymbol{v_{i_{n-m}}}\}$’

Elements from **B** can be adjoined to **S** to form a spanning set of **V** with ‘n’ elements. Because **S** already has ‘n’ elements, **S** itself is a spanning set of **V**.

Thus **S** is a basis of **V**.

Theorem:

Let **V** be a vector space of finite dimension ‘n’. Then, any spanning set $\boldsymbol{T} = \{v_1, v_2, \ldots, v_n\}$ of **V** with ‘n’ elements is a basis of **V**.

Proof:

Suppose $\boldsymbol{B} = \{w_1, w_2, \ldots, w_n\}$ is a basis of **V**. and suppose $\boldsymbol{T} = \{v_1, v_2, \ldots, v_n\}$ is linearly dependent. Then some v_i is a linear combination of the preceding vectors. By problem **“if $\boldsymbol{T} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ spans V then for $\boldsymbol{w \epsilon V}$ the set $\{\boldsymbol{w}, \boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ will be linearly dependent and spans V and if $\boldsymbol{v_i}$ is a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_{i-1}}$ then $\boldsymbol{T}$ without $\boldsymbol{v_i}$ spans V”**

Thus **V** is spanned by vectors in **T** without v_i and there are $n-1$ of them. By Lemma **‘‘Suppose $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ spans V, and suppose $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_m}\}$ is linearly independent. Then $\boldsymbol{m} \leq \boldsymbol{n}$, and V is spanned by a set of the form $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_m}, \boldsymbol{v_{i_1}}, \boldsymbol{v_{i_2}}, \ldots, \boldsymbol{v_{i_{n-m}}}\}$’’** the independent set **B** cannot have more than $n-1$ elements. This contradict the fact **B** has ‘n’ elements.

Thus **T** is linearly independent and hence **T** is a basis of **V**.

Theorem: Suppose **S** spans a vector space **V**. Then: Any maximum number of linearly independent vectors in S form a basis of **V**.

Proof: Suppose $\{v_1, v_2, \ldots, v_n\}$ is maximum linearly independent subset of **S**, and suppose $w \epsilon \boldsymbol{S}$. Accordingly, $\{v_1, v_2, \ldots, v_n, w\}$ is linearly independent. No v_k can be linear combination of preceding vectors. Hence is w is a linear combination of the v_i. Thus $w \epsilon\ Span(v_i)$ and hence $\boldsymbol{S} \subseteq\ Span(v_i)$

This leads to $\boldsymbol{V} = Span\boldsymbol{S} \subseteq\ Span(v_i) \subseteq \boldsymbol{V}$

Thus $\{v_i\}$ spans **V** and as it is linearly independent, it is a basis of **V**.

Theorem: Suppose **S** spans a vector space **V**. Then: Suppose one deletes from **S** every vector that is a linear combination of preceding vectors in **S**. Then the remaining vectors form a basis of **V**.

Proof: The remaining vectors form a maximum linearly independent subset of **S**; hence by theorem **"Suppose S spans a vector space V. Then: Any maximum number of linearly independent vectors in S form a basis of V"** it is a basis of **V**.

Theorem: Let **V** be a vector space of finite dimension and let $\boldsymbol{S} = \{v_1, v_2, \ldots, v_r\}$ be a set of linearly independent vectors in **V**. Then **S** is part of a basis of **V**; that is, **S** may be extended to a basis of **V**.

Proof: Suppose $\boldsymbol{B} = \{w_1, w_2, \ldots, w_n\}$ is a basis of **V**. then **B** spans **V** and hence **V** is spanned by $\boldsymbol{S} \cup \boldsymbol{B} = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_n\}$By theorem **"Suppose S spans a vector space V. Then: Any maximum number of linearly independent vectors in S form a basis of V. And Suppose one deletes from S every vector that is a linear combination of preceding vectors in S. Then the remaining vectors form a basis of V "** we can delete from $\boldsymbol{S} \cup \boldsymbol{B}$ each vector that is the linear combination of preceding vectors to obtain a basis $\boldsymbol{B}'$ for **V**. because **S** is linearly independent, no v_k is a linear combination of preceding vectors. Thus $\boldsymbol{B}'$ contains every vector in **S**, and **S** is the part of the basis $\boldsymbol{B}'$ for **V**

Examples:

(a) The following four vectors in $\mathbf{R}^4$ form a matrix in echelon form:
$(1,1,1,1)$, $(0,1,1,1)$, $(0,0,1,1)$, $(0,0,0,1)$

Thus, the vectors are linearly independent, and, because dim$\mathbf{R}^4$ = 4, the four vectors form a basis of $\mathbf{R}^4$

(b) The following $n+1$ polynomials in $\boldsymbol{P_n}(t)$ are of increasing degree:
1 , $t-1$, $(t-1)^2$, ... $(t-1)^n$

Therefore, no polynomial is a linear combination of preceding polynomials; hence, the polynomials are linear independent. Furthermore, they form a basis of $\boldsymbol{P_n}(t)$, because dim$\boldsymbol{P_n}(t) = n+1$.

(c) Consider any four vectors in $\mathbf{R}^3$, say
$(257,-132,58)$, $(43,0,-17)$, $(521,-317,94)$, $(328,-512,-731)$

By Theorem **"Let V be a vector space of finite dimension 'n'. Then, any $\boldsymbol{n+1}$ or more vectors in V are linearly dependent"**, the four vectors must be linearly dependent, because they come from the three-dimensional vector space $\mathbf{R}^3$.

Dimension

The number of elements (vectors) in a basis of a vector space **V** over **F** is called dimension of **V**. It is denoted by $\dim(\boldsymbol{V})$.

Engineers often use the term **degree of freedom** as a synonym for dimension.

Remark:

- The vector space {0} is defined to have dimension 0. It is known as **zero vector space**.
- The simplest of all vector spaces is the zero vector space **V** = {0}. This space is the finite dimensional because it is spanned by vector **0**.Since {0} is not linearly independent set, thats why **V** = {0} has no basis. However we will find it useful to define empty set φ to be a basis for this vector space.
- Suppose a vector space **V** does not have a finite basis. Then **V** is said to be of **infinite dimension** or to be **infinite-dimensional**.
- $\dim(\boldsymbol{R^n}) = n$ (The standard basis has 'n' vectors)
- **Example:** $\{(1,0,0),(0,1,0),(0,0,1)\}$ is the standard basis set of $\boldsymbol{R^3}$ and $\dim(\boldsymbol{R^3}) = 3$
- $\dim(\boldsymbol{P_n}) = n + 1$ (The standard basis has 'n + 1' vectors)
- $\dim(\boldsymbol{M_{mn}}) = mn$ (The standard basis has 'mn' vectors)
- $\boldsymbol{R^n}, \boldsymbol{P_n}, \boldsymbol{M_{mn}}$ are finite dimensional vector space. While $\boldsymbol{R}^\infty, \boldsymbol{P}_\infty, \boldsymbol{F}(-\infty,\infty), \boldsymbol{C}(-\infty,\infty), \boldsymbol{C^n}(-\infty,\infty), \boldsymbol{C}^\infty(-\infty,\infty)$ are infinite dimensional vector spaces
- $\dim[Span\{v_1, v_2, \dots, v_r\}] = r$ it means, the dimension of the space spanned by a linearly independent set of vectors is equal to the number of vectors in that set.
- For two finite dimensional subspaces **U** and **W** of a vector space **V** over a field **F** we have $\dim(\boldsymbol{U}+\boldsymbol{W}) = \dim(\boldsymbol{U}) + \dim(\boldsymbol{W}) - \dim(\boldsymbol{U}\cap\boldsymbol{W})$
- For two finite dimensional subspaces **U** and **W** of a vector space **V** over a field **F** with $\boldsymbol{U}\cap\boldsymbol{W} = \{0\}$ and $\boldsymbol{V} = \boldsymbol{U}\oplus\boldsymbol{W}$ we have $\dim(\boldsymbol{U}+\boldsymbol{W}) = \dim(\boldsymbol{U}) + \dim(\boldsymbol{W})$

Keep in mind:

- Let **V** be a vector space such that one basis has 'm' elements and another basis has 'n' elements. Then $m = n$.
- Every basis of the finite dimensional vector space has the same number of elements (vectors).
- Let **V** be an n – dimensional vector space, and let $\{v_1, v_2, \dots, v_n\}$ be any basis then if a set in **V** has more than 'n' vectors, then it is linearly dependent.
- Let **V** be an n – dimensional vector space, and let $\{v_1, v_2, \dots, v_n\}$ be any basis then if a set in **V** has fewer than 'n' vectors, then it does not span **V**.
- **(Plus Theorem)** Let **S** be a non – empty set of a vectors in a vector space **V** then if **S** is linearly independent set, and if $\boldsymbol{v}$ is a vector in **V** that is outside of span(**S**), then the set $\boldsymbol{S} \cup \{\boldsymbol{v}\}$ that results by inserting $\boldsymbol{v}$ into **S** is still linearly independent.
- **(Minus Theorem)** Let **S** be a non – empty set of a vectors in a vector space **V** then if $\boldsymbol{v}$ is a vector in **S** that is expressible as a linear combination of other vectors in **S**, and if $\boldsymbol{S} - \{\boldsymbol{v}\}$ denotes the set obtained by removing $\boldsymbol{v}$ from **S**, then **S** and $\boldsymbol{S} \cup \{\boldsymbol{v}\}$ span the same space; that is,$span(\boldsymbol{S}) = span(\boldsymbol{S} - \{\boldsymbol{v}\})$
- Let **V** be an n – dimensional vector space, and let **S** be a set in **V** with exactly 'n' vectors. Then **S** is a basis for **V** if and only if S spans **V** or **S** is linearly independent.
- Let **S** be a finite set of vectors in a finite dimensional vector space **V** then if **S** spans **V** but is not a basis for **V**, then **S** can be reduced to a basis for **V** by removing appropriate vectors from **S**. (This theorem tells that; every spanning set for a subspace is either a basis for that subspace or has a basis as a subset)
- Let **S** be a finite set of vectors in a finite dimensional vector space **V** then if **S** is linearly independent set that is not already a basis for **V**, then **S** can be enlarged to a basis for **V** by inserting appropriate vectors into **S**. (This theorem tells that; every linearly independent set in a subspace is either a basis for that subspace or can be extended to a basis for it.)
- If **W** is subspace of a finite dimensional vector space **V**, then **W** is finite dimensional , $\dim(\boldsymbol{W}) \le \dim(\boldsymbol{V})$ and $\boldsymbol{W} = \boldsymbol{V} \Leftrightarrow \dim(\boldsymbol{W}) = \dim(\boldsymbol{V})$

An infinite dimensional vector space

Show that $\boldsymbol{P}_\infty$ is an infinite dimensional vector space as it has no finite spanning set.

Solution:

Arbitrary if we consider a finite spanning set, say $\boldsymbol{S} = \{\boldsymbol{p_1}, \boldsymbol{p_2}, \ldots, \boldsymbol{p_r}\}$ then the degree of the polynomials in S would have a maximum value say 'n' and this in turn would imply that any linear combination of the polynomials in **S** would have degree at most 'n'. Thus there we would be no way to express the polynomial x^{n+1} as a linear combination of the polynomials in **S**, contradicting the fact that the vectors in **S** span $\boldsymbol{P}_\infty$

Example: Let **W** be a subspace of the real space $\mathbf{R}^3$. Note that dim $\mathbf{R}^3$ = 3. The following cases apply:

(a) If $dim\boldsymbol{W} = 0$, then $\boldsymbol{W} = \{0\}$, a point.
(b) If $dim\boldsymbol{W} = 1$, then **W** is a line through the origin 0.
(c) If $dim\boldsymbol{W} = 2$, then **W** is a plane through the origin 0.
(d) If $dim\boldsymbol{W} = 3$, then **W** is the entire space $\mathbf{R}^3$

Example: Find a basis and dimension of the subspace **W** of $\mathbf{R}^3$ where $\boldsymbol{W} = \{(a, b, c): a + b + c = 0\}$

Solution: Note that $\boldsymbol{W} \neq \boldsymbol{R}^3$, because, for example $(1,2,3) \notin \boldsymbol{W}$ thus $dim(\boldsymbol{W}) < 3$. Note that $(1,0,-1), (0,1,-1)$ are two independent vectors in **W** thus $dim(\boldsymbol{W}) = 2$ and so both vectors $(1,0,-1), (0,1,-1)$ form a basis of **W**.

Example: Find a basis and dimension of the subspace **W** of $\mathbf{R}^3$ where $\boldsymbol{W} = \{(a, b, c): a = b = c\}$

Solution: The vector $\boldsymbol{u} = (1,1,1)\epsilon\boldsymbol{W}$ any vector $\boldsymbol{w}\epsilon\boldsymbol{W}$ has the form $\boldsymbol{w} = (k, k, k)$. Hence $\boldsymbol{w} = k\boldsymbol{u}$. Thus $\boldsymbol{u}$ spans **W** and $dim(\boldsymbol{W}) = 1$

Example:

Find a basis and dimension of the solution space of the homogeneous system

$$x_1 + 3x_2 - 2x_3 + 0x_4 + 2x_5 + 0x_6 = 0$$

$$2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = 0$$

$$0x_1 + 0x_2 + 5x_3 + 10x_4 + 0x_5 + 15x_6 = 0$$

$$2x_1 + 6x_2 + 0x_3 + 8x_4 + 4x_5 + 18x_6 = 0$$

Solution:

Using Gauss Jordan's Elimination method matrix $\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 & 0 \\ 2 & 6 & 0 & 8 & 4 & 18 & 0 \end{bmatrix}$ can be converted into row reduced echelon form as $\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Thus corresponding system is $x_6 = 0,\ x_3 + 2x_4 = 0,\ x_1 + 3x_2 + 4x_4 + 2x_5 = 0$

These yields $x_1 = -3r - 4s - 2t$, $x_2 = \text{r}$, $x_3 = -2s,\ x_4 = s$, $x_5 = t,\ x_6 = 0$

Which can be written in vector form as

$(x_1, x_2, x_3, x_4, x_5, x_6) = (-3r - 4s - 2t, r, -2s, s, t, 0)$

Or alternatively as

$(x_1, x_2, x_3, x_4, x_5, x_6) = r(-3,1,0,0,0,0) + s(-4,0,-2,1,0,0) + t(-2,0,0,0,1,0)$

This shows that the vectors

$\boldsymbol{v_1} = (-3,1,0,0,0,0), \boldsymbol{v_2} = (-4,0,-2,1,0,0), \boldsymbol{v_3} = (-2,0,0,0,1,0)$

Span the solution space and are linearly independent (Check!). Thus the solution space has dimension 3.

Example:**(Applying the plus minus theorem)**

Show that $\boldsymbol{p_1} = 1 - x^2$, $\boldsymbol{p_2} = 2 - x^2$, $\boldsymbol{p_3} = x^3$ are linearly independent vectors.

Solution: The set $\boldsymbol{S} = \{\boldsymbol{p_1}, \boldsymbol{p_2}\}$ is linearly independent since neither vector in **S** is a scalar multiple of the other. Since the vector $\boldsymbol{p_3}$ cannot be expressed as a linear combination of the vectors in **S**, it can be adjoined to **S** to produce a linearly independent set $\boldsymbol{S} \cup \{\boldsymbol{p_3}\} = \{\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}\}$

Example: (**Basses by inspection)**

Explain why the vectors $\boldsymbol{v_1} = (-3,7)$, $\boldsymbol{v_2} = (5,5)$ form a basis for $\mathbf{R}^2$

Solution: Since neither vector is a scalar multiple of the other, the two vectors form a linearly independent set in the two dimensional space $\mathbf{R}^2$ and hence they form a basis by theorem **'Let V be an n – dimensional vector space, and let S be a set in V with exactly 'n' vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.'**

Example: (**Basses by inspection):** Explain why the vectors $\boldsymbol{v_1} = (2,0,-1)$, $\boldsymbol{v_2} = (4,0,7)$ and $\boldsymbol{v_3} = (-1,1,4)$ form a basis for $\mathbf{R}^3$

Solution: The vectors $\boldsymbol{v_1}, \boldsymbol{v_2}$ form a linearly independent set in the xz – plane. The vector $\boldsymbol{v_3}$ is outside of the xz – plane, so the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is also linearly independent. Since $\mathbf{R}^3$ is there dimensional theorem **'Let V be an n – dimensional vector space, and let S be a set in V with exactly 'n' vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.'** Implies that $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is the basis for $\mathbf{R}^3$

Example: Determine a basis for the subspace $the\ plane\ x - 2y + 5z = 0$ of $\mathbf{R}^3$

Solution: Given equation of is $the\ plane\ x - 2y + 5z = 0$ or $x = 2y - 5z$ where 'y' and 'z' are free variables. Then the above equation in vector form can be written as $(x, y, z) = (2y - 5z, y, z) = (2y - 5z, y + 0, z + 0)$

$(x, y, z) = (2y, y, 0) + (-5z, 0, z) = y(2,1,0) + z(-5,0,1)$ thus given plane is spanned by vectors $(2,1,0), (-5,0,1)$ and as none of the vector is multiple of the other. So that set $\{(2,1,0), (-5,0,1)\}$ is linearly independent. Hence $\{(2,1,0), (-5,0,1)\}$ and dimension of subspace is 2.

Practice:

1) Determine whether $(1,1,1,1),(1,2,3,2),(2,6,8,5)$ form a basis of $\mathbf{R}^4$. If not, find the dimension of the subspace they span.

2) Find a basis and dimension of the subspace **W** of $\mathbf{R}^4$ where
   i. $\boldsymbol{W} =$ All vectors of the form $(a,b,c,0)$
   ii. $\boldsymbol{W} =$ All vectors of the form (a,b,c,d) where $d = a + b$ and $c = a - b$
   iii. $\boldsymbol{W} = \{(a,b,c,d): a = b = c = d\}$
   iv. $\boldsymbol{W} = \{(x_1,x_2,x_3,x_4): x_2, = x_3\}$

3) Find a basis and dimension of the solution space of the homogeneous system.

i.
$$x_1 + x_2 - x_3 = 0$$
$$-2x_1 - x_2 + 2x_3 = 0$$
$$-x_1 + 0x_2 + x_3 = 0$$

ii.
$$3x_1 + x_2 + x_3 + x_4 = 0$$
$$5x_1 - x_2 + x_3 - x_4 = 0$$

iii.
$$2x_1 + x_2 + 3x_3 = 0$$
$$x_1 + 0x_2 + 5x_3 = 0$$
$$0x_1 + x_2 + x_3 = 0$$

iv.
$$x_1 - 4x_2 + 3x_3 - x_4 = 0$$
$$2x_1 - 8x_2 + 6x_3 - 2x_4 = 0$$

v.
$$x_1 - 3x_2 + x_3 = 0$$
$$2x_1 - 6x_2 + 2x_3 = 0$$
$$3x_1 - 9x_2 + 3x_3 = 0$$

vi.
$$x + y + z = 0$$
$$3x + 2y - 2z = 0$$
$$4x + 3y - z = 0$$
$$6x + 5y + z = 0$$

4) Determine a basis for the subspace of $\mathbf{R}^3$ and state its dimension.
   i. $\boldsymbol{W} = \{(a,b,c): 3a - 2b + c\}$
   ii. $\boldsymbol{W} = \{(a,b,c): b = a + c\}$
   iii. $the\ plane\ 3x - 2y + 5z = 0$
   iv. $the\ plane\ x - y = 0$
   v. $the\ line\ x = 2t, y = -t, z = 4t$
   vi. $the\ line\ \frac{x}{-2} = \frac{y}{1} = \frac{z}{6}$
5) Find the dimension of each of the following vector spaces:
   i. The vector space of all diagonal $n \times n$ matrices
   ii. The vector space of all symmetric $n \times n$ matrices
   iii. The vector space of all upper triangular $n \times n$ matrices
6) Find the dimension of the subspace of $\mathbf{P}_3$ consisting of all polynomials $a_0 + a_1x + a_2x^2 + a_3x^3$ for which $a_0 = 0$
7) Show that the set **W** of all polynomials in $\mathbf{P}_2$ such that $\boldsymbol{p}(1) = 0$ is a subspace of $\mathbf{P}_2$. Then make a conjecture about the dimension of **W**, and conform your conjecture by finding a basis for **W**.
8) Find a standard basis vector for $\mathbf{R}^3$ that can be added to the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to produce a basis for $\mathbf{R}^3$
   i. $\boldsymbol{v_1} = (-1,2,3)$, $\boldsymbol{v_2} = (1,-2,-2)$
   ii. $\boldsymbol{v_1} = (1,-1,0)$, $\boldsymbol{v_2} = (3,1,-2)$
9) Find a standard basis vector for $\mathbf{R}^4$ that can be added to the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to produce a basis for $\mathbf{R}^4$ where $\boldsymbol{v_1} = (1,-4,2,-3)$, $\boldsymbol{v_2} = (-3,8,-4,6)$
10) Let $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ be a basis for a vector space **V**. Show that $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ is also a basis, where $\boldsymbol{u_1} = \boldsymbol{v_1}, \boldsymbol{u_2} = \boldsymbol{v_1} + \boldsymbol{v_2}, \boldsymbol{u_3} = \boldsymbol{v_1} + \boldsymbol{v_2} + \boldsymbol{v_3}$
11) The vectors $\boldsymbol{v_1} = (1,-2,3)$, $\boldsymbol{v_2} = (0,5,-3)$ are linearly independent. Enlarge $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to a basis for $\mathbf{R}^3$
12) The vectors $\boldsymbol{v_1} = (1,0,0,0)$, $\boldsymbol{v_2} = (1,1,0,0)$ are linearly independent. Enlarge $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ to a basis for $\mathbf{R}^4$
13) Find a basis for the subspace of $\mathbf{R}^3$ that is spanned by the vectors;

$$\boldsymbol{v_1} = (1,0,0), \boldsymbol{v_2} = (1,0,1), \boldsymbol{v_3} = (2,0,1), \boldsymbol{v_4} = (0,0,-1)$$

14) Find a basis for the subspace of $\mathbf{R}^4$ that is spanned by the vectors;

$$\boldsymbol{v_1} = (1,1,1,1), \boldsymbol{v_2} = (2,2,2,0), \boldsymbol{v_3} = (0,0,0,3), \boldsymbol{v_4} = (3,3,3,4)$$

The following is a fundamental result in linear algebra.

Theorem (Omit Proof): Let **V** be a vector space such that one basis has 'm' elements and another basis has 'n' elements. Then $m = n$.

(A vector space **V** is said to be of finite dimension 'n' or n-dimensional, written $dim\boldsymbol{V} = n$ if V has a basis with 'n' elements. Theorem tells us that all bases of V have the same number of elements, so this definition is well defined.)

Theorem:

Every basis of the finite dimensional vector space has the same number of elements (vectors).

Proof:

Let a vector space **V** over field **F** has two basis $\boldsymbol{A}$ and $\boldsymbol{B}$ with 'm' and 'n' number of elements. Since $\boldsymbol{A}$ spans **V** and $\boldsymbol{B}$ is a linearly independent subset in **V**, so $\boldsymbol{B}$ cannot have more than 'm' number of elements

i.e. $n \leq m$ ……………(i)

now Since $\boldsymbol{B}$ spans **V** and $\boldsymbol{A}$ is a linearly independent subset in **V**, so $\boldsymbol{A}$ cannot have more than 'n' number of elements

i.e. $m \leq n$ ……………(ii)

from (i) and (ii) $m = n$

Hence the theorem.

Theorem:

Let **V** be an n – dimensional vector space, and let $\{v_1, v_2, \dots, v_n\}$ be any basis then if a set in **V** has more than 'n' vectors, then it is linearly dependent.

Proof:

Let $\boldsymbol{S'} = \{w_1, w_2, \dots, w_m\}$ be any set of 'm' vectors in **V**, where $m > n$. We want to show that $\boldsymbol{S'}$ is linearly dependent. Since $\boldsymbol{S} = \{v_1, v_2, \dots, v_n\}$ is a basis, each w_i can be expressed as linear combination of the vectors in **S**, say

$$\boldsymbol{w_1} = a_{11}\boldsymbol{v_1} + a_{21}\boldsymbol{v_2} + \cdots + a_{n1}\boldsymbol{v_n}$$

$$\boldsymbol{w_2} = a_{12}\boldsymbol{v_1} + a_{22}\boldsymbol{v_2} + \cdots + a_{n2}\boldsymbol{v_n} \qquad \dots\dots\dots\dots\dots\dots\text{(i)}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\boldsymbol{w_m} = a_{1m}\boldsymbol{v_1} + a_{2m}\boldsymbol{v_2} + \cdots + a_{nm}\boldsymbol{v_n}$$

To show that $\boldsymbol{S'}$ is linearly dependent, we must find scalar $k_1, k_2, \dots, k_m$, not all zero, such that $k_1\boldsymbol{w_1} + k_2\boldsymbol{w_2} + \cdots + k_m\boldsymbol{w_m} = 0 \quad \dots\dots\dots\dots\dots\dots$(ii)

We can write equation (i) in partition as follows

$$[\boldsymbol{w_1}|\boldsymbol{w_2}|\cdots|\boldsymbol{w_m}] = [\boldsymbol{v_1}|\boldsymbol{v_2}|\cdots|\boldsymbol{v_n}]\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \qquad \dots\dots\dots\dots\dots\dots\text{(iii)}$$

Since $m > n$, the linear system

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \dots\dots\dots\dots\dots\dots\text{(iv)}$$

Has more equations than unknowns and hence has a non – trivial solution

$$x_1 = k_1, x_2 = k_2, \dots, x_m = k_m$$

Creating a column vector from this solution and multiplying both sides of (iii) on the right by this vector yields

$$[\boldsymbol{w_1}|\boldsymbol{w_2}|\cdots|\boldsymbol{w_m}]\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}=[\boldsymbol{v_1}|\boldsymbol{v_2}|\cdots|\boldsymbol{v_n}]\begin{bmatrix}a_{11}&a_{21}&\cdots&a_{m1}\\a_{12}&a_{22}&\cdots&a_{m2}\\\vdots&\vdots&\vdots&\vdots\\a_{1n}&a_{2n}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}$$

By (iv) this simplifies to $$[\boldsymbol{w_1}|\boldsymbol{w_2}|\cdots|\boldsymbol{w_m}]\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}=[\boldsymbol{v_1}|\boldsymbol{v_2}|\cdots|\boldsymbol{v_n}]\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$$

$$[\boldsymbol{w_1}|\boldsymbol{w_2}|\cdots|\boldsymbol{w_m}]\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$$

Which we can writes as $k_1\boldsymbol{w_1}+k_2\boldsymbol{w_2}+\cdots+k_m\boldsymbol{w_m}=0$

Since the scalar coefficients in this equation are not all zero, we have proved that $\boldsymbol{S'}=\{w_1,w_2,\dots,w_m\}$ is linearly dependent

Theorem:

Let **S** be a non – empty set of a vectors in a vector space **V** then if **S** is linearly independent set, and if $\boldsymbol{v}$ is a vector in **V** that is outside of span(**S**), then the set $\boldsymbol{S}\cup\{\boldsymbol{v}\}$ that results by inserting $\boldsymbol{v}$ into **S** is still linearly independent.

Proof:

Assume that $\boldsymbol{S}=\{v_1,v_2,\dots,v_r\}$ is a linearly independent set of vectors in **V** and $\boldsymbol{v}$ is a vector in **V** that is outside of span(**S**). To show that $\boldsymbol{S'}=\{v_1,v_2,\dots,v_r,v\}$ is linearly independent set we must show that the only scalar that satisfy $k_1\boldsymbol{v_1}+k_2\boldsymbol{v_2}+\cdots+k_r\boldsymbol{v_r}+k_{r+1}\boldsymbol{v}=\boldsymbol{0}$ are $k_1=k_2=\cdots=k_r=k_{r+1}=0$
But it must be true that $k_{r+1}=0$ for otherwise we could solve $k_1\boldsymbol{v_1}+k_2\boldsymbol{v_2}+\cdots+k_r\boldsymbol{v_r}+k_{r+1}\boldsymbol{v}=\boldsymbol{0}$ for $\boldsymbol{v}$ as a linear combination of $v_1,v_2,\dots,v_r$, contradicting the assumption that $\boldsymbol{v}$ is outside of span(**S**). Thus $k_1\boldsymbol{v_1}+k_2\boldsymbol{v_2}+\cdots+k_r\boldsymbol{v_r}+k_{r+1}\boldsymbol{v}=\boldsymbol{0}$ simplifies to $k_1\boldsymbol{v_1}+k_2\boldsymbol{v_2}+\cdots+k_r\boldsymbol{v_r}=\boldsymbol{0}$ which, by linear independence of $\{v_1,v_2,\dots,v_r,v\}$ implies that $k_1=k_2=\cdots=k_r=0$

Hence the theorem.

Theorem:

Let **S** be a non – empty set of a vectors in a vector space **V** then if $\boldsymbol{v}$ is a vector in **S** that is expressible as a linear combination of other vectors in **S**, and if $\boldsymbol{S} - \{\boldsymbol{v}\}$ denotes the set obtained by removing $\boldsymbol{v}$ from **S**, then **S** and $\boldsymbol{S} \cup \{\boldsymbol{v}\}$ span the same space; that is,$span(\boldsymbol{S}) = span(\boldsymbol{S} - \{\boldsymbol{v}\})$

Proof:

Assume that $\boldsymbol{S} = \{v_1, v_2, \dots, v_r\}$ is a set of vectors in **V** and (to be specific) suppose that v_r is a linear combination of $v_1, v_2, \dots, v_{r-1}$, say

$$\boldsymbol{v_r} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_{r-1}\boldsymbol{v_{r-1}}$$

We want to show that if $\boldsymbol{v_r}$ is removed from $\boldsymbol{S}$, then the remaining set of vectors $\{v_1, v_2, \dots, v_{r-1}\}$ still spans **S**; that is , we must show that every vector $\boldsymbol{w}$ in $span(\boldsymbol{S})$ is expressible as a linear combination of $\{v_1, v_2, \dots, v_{r-1}\}$. But if $\boldsymbol{w}$ is in $span(\boldsymbol{S})$, then $\boldsymbol{w}$ is expressible in the form

$$\boldsymbol{w} = k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + \cdots + k_{r-1}\boldsymbol{v_{r-1}} + k_r\boldsymbol{v_r}$$

Or, on substituting $\boldsymbol{v_r} = k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + \cdots + k_{r-1}\boldsymbol{v_{r-1}}$ in above

$$\boldsymbol{w} = k_1\boldsymbol{v_1} + k_2\boldsymbol{v_2} + \cdots + k_{r-1}\boldsymbol{v_{r-1}} + k_r(c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_{r-1}\boldsymbol{v_{r-1}})$$

Which expresses $\boldsymbol{w}$ as a linear combination of $\{v_1, v_2, \dots, v_{r-1}\}$.Hence the theorem.

In general, to show that a set of vectors $\{v_1, v_2, \dots, v_n\}$ is a basis for a vector space **V**, one must show that the vectors are linearly independent and span **V**, However, if we happen to know that **V** has dimension 'n' (so that $\{v_1, v_2, \dots, v_n\}$ contains the right number of vectors for a basis), the it is suffices to check either linear independence or spanning, the remaining condition will hold automatically. This is the content of the following theorem.

Theorem:

Let **W** be a subspace of an n-dimensional vector space **V**. Then $dim\boldsymbol{W} \leq n$. In particular, if $dim\boldsymbol{W} = n$, then $\boldsymbol{W} = \boldsymbol{V}$.

Proof:

Because **V** is of dimension 'n' , any $n + 1$ or more vectors are linearly dependent. Furthermore, because a basis of **W** consists of linearly independent vectors, it cannot contain more than 'n' elements. Accordingly, $dim\boldsymbol{W} \leq n$ in particular, if $\{w_1, w_2, \dots, w_n\}$ is a basis of **W**, then because it is an independent set with *n* elements, it is also a basis of **V**. Thus $\boldsymbol{W} = \boldsymbol{V}$ when $dim\boldsymbol{W} = n$

Quotient Space

Let **W** be a subspace of **V** i.e. $\boldsymbol{W} \subset \boldsymbol{V}$ then $\frac{\boldsymbol{V}}{\boldsymbol{W}} = \{w + x : w\epsilon\boldsymbol{W} \ \& \ x\epsilon\boldsymbol{V}\}$ then $w + x$ is a subset of **V** for $x\epsilon\boldsymbol{V}$ then $\frac{\boldsymbol{V}}{\boldsymbol{W}}$ form a vector space over the same field of **V** with respect to operation defined as;

i. $(w + x) + (w + y) = w + (x + y)$ where $w\epsilon\boldsymbol{W} \ \& \ x, y\epsilon\boldsymbol{V}$
ii. $\propto (w + x) = w + \propto x$ where $w\epsilon\boldsymbol{W}$, $x\epsilon\boldsymbol{V}$ & $\propto \epsilon\boldsymbol{F}$

Remember:

i. **W** is the additive identity of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$ i.e. **W** is zero vector of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$
ii. $w + x = w + y \Leftrightarrow x - y \ \epsilon\boldsymbol{W}$

Examples:

i. Let $\mathbf{W}$ be a subspace of a $\mathbf{R^3}$ spanned by the vector $(1,1,1)$ that is $\boldsymbol{W} = span\{(1,1,1)\} = \{k(1,1,1): k \in \boldsymbol{R}\}$ then $\mathbf{W}$ is the straight line through origin and the point $(1,1,1)$. For any vector $(x, y, z) \in \boldsymbol{R^3}$ we can regard the coset $(x, y, z) + \boldsymbol{W}$ as the set of vectors obtained by adding the vector (x, y, z) to each vector of $\boldsymbol{W}$. This coset is therefore the set of all vectors on the line through the point (x, y, z) parallel to the line $\boldsymbol{W}$. Hence $\frac{\boldsymbol{R^3}}{\boldsymbol{W}}$ is the collection of lines parallel to $\boldsymbol{W}$

ii. Let $\boldsymbol{W} = span\{(1,0,0), (0,1,0)\}$ then $\mathbf{W}$ is the set of all vectors in xy – plane and the cosets are the planes parallel to the xy – plane. Thus the quotient space $\frac{\boldsymbol{R^3}}{\boldsymbol{W}}$ is the collection of planes parallel to xy – plane.

iii. Let $\boldsymbol{W} = span\{(1,0,0)\}$ then for any vector $(x, y, z) \in \boldsymbol{R^3}$ we have $(x, y, z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$ and therefore since $x(1,0,0) \in \boldsymbol{W}$

$(x, y, z) + \boldsymbol{W} = \boldsymbol{W} + y[(0,1,0) + \boldsymbol{W}] + z[(0,0,1) + \boldsymbol{W}]$

$(x, y, z) + \boldsymbol{W} = y(0,1,0) + \boldsymbol{W} + z(0,0,1) + \boldsymbol{W}$

$(x, y, z) + \boldsymbol{W} = (0, y, z) + \boldsymbol{W}$

The vectors $(0,1,0) + \boldsymbol{W}$ and $(0,0,1) + \boldsymbol{W}$ are therefore also independent and hence they for a basis of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$

iv. The set $\mathbf{P_1}$ is a subspace of $\mathbf{P_4}$ form the quotient space $\frac{\boldsymbol{P_4}}{\boldsymbol{P_1}}$ for this we consider $\boldsymbol{p}(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \epsilon \boldsymbol{P_4}$

Then $\boldsymbol{p}(x) + \boldsymbol{P_1} = a_4(x^4 + \boldsymbol{P_1}) + a_3(x^3 + \boldsymbol{P_1}) + a_2(x^2 + \boldsymbol{P_1}) + \boldsymbol{P_1}$

$\boldsymbol{p}(x) + \boldsymbol{P_1} = a_4x^4 + a_3x^3 + a_2x^2 + \boldsymbol{P_1}$

So $(x^4 + \boldsymbol{P_1}), (x^3 + \boldsymbol{P_1}), (x^2 + \boldsymbol{P_1})$ spans $\frac{\boldsymbol{P_4}}{\boldsymbol{P_1}}$

Moreover these are linearly independent so a basis for $\frac{\boldsymbol{P_4}}{\boldsymbol{P_1}}$ is

$\{x^4 + \boldsymbol{P_1}, x^3 + \boldsymbol{P_1}, x^2 + \boldsymbol{P_1}\}$

v. Let $\boldsymbol{B} = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_n}\}$ be a basis for a subspace $\mathbf{W}$ of $\mathbf{V}$, and extended it to a basis $\boldsymbol{B'} = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_n}, \boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_k}\}$ of $\mathbf{V}$, then $\{\boldsymbol{v_1} + \mathbf{W}, \boldsymbol{v_2} + \mathbf{W}, \dots, \boldsymbol{v_k} + \mathbf{W}\}$ is a basis of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$

Theorem: Dimension of Quotient Space:

If **W** is subspace of a finite dimensional vector space **V**, then

i. **W** is finite dimensional
ii. $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V})$
iii. $dim\frac{\boldsymbol{V}}{\boldsymbol{W}} = dim\boldsymbol{V} - dim\boldsymbol{W}$

Proof:

i. Since $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V}) = n$ and if **W** has a basis with n elements. Then by theorem **"Let W be a vector space such that one basis has *m* elements and another basis has *n* elements, then m = n."** all basis of W have the same number of elements and hence **W** is finite dimensional.
ii. If W is finite dimensional, so it has a basis $\boldsymbol{S} = \{w_1, w_2, \dots, w_m\}$. Either **S** is also a basis for **V** or it is not. If so, then $\dim(\boldsymbol{V}) = m$, which means that $\dim(\boldsymbol{W}) = \dim(\boldsymbol{V})$. If not, then because **S** is a linearly independent set it can be enlarged to a basis for **V** by theorem **"Let S be a finite set of vectors in a finite dimensional vector space V then if S is linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S"**

But this implies that $\dim(\boldsymbol{W}) < \dim(\boldsymbol{V})$. So we have shown that $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V})$ in all cases.

iii. Let $\{w_1, w_2, \dots, w_m\}$ be a basis of **W** with $dim\boldsymbol{W} = m$ and $\{w_1, w_2, \dots, w_n, v_1, v_2, \dots, v_n\}$ be an extended basis of **V** with $dim\boldsymbol{V} = m + n$.
Then we have to prove that $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \dots, v_n + \mathbf{W}\}$ is a basis of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$

For this firstly we will show that $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \dots, v_n + \mathbf{W}\}$ is linearly independent. Consider;

$$\propto_1 (v_1 + \mathbf{W}) + \propto_2 (v_2 + \mathbf{W}) + \cdots + \propto_n (v_n + \mathbf{W}) = 0_v + \boldsymbol{W}$$

$$(\propto_1 v_1 + \mathbf{W}) + (\propto_2 v_2 + \mathbf{W}) + \cdots + (\propto_n v_n + \mathbf{W}) = 0_v + \boldsymbol{W}$$

$$(\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n) + \mathbf{W} = 0_v + \boldsymbol{W}$$

$(\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n) \epsilon \mathbf{W}$

Implies $\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n$ is a linear combination of $w_1, w_2 + \cdots, w_m$

$\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m$

$\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n - \beta_1 w_1 - \beta_2 w_2 - \cdots - \beta_m w_m = 0$

Since $\{w_1, w_2, \ldots, w_m, v_1, v_2, \ldots, v_n\}$ be an extended basis of **V** so $w_1, w_2, \ldots, w_m, v_1, v_2, \ldots, v_n$ are linearly independent.

Implies $\propto_1 = \propto_2 = \cdots = \propto_n = 0 \;\; and \;\; \beta_1 = \beta_2 w_2 = \cdots = \beta_m = 0$

Thus $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \ldots, v_n + \mathbf{W}\}$ is linearly independent.

Now we will show that $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \ldots, v_n + \mathbf{W}\}$ is spans $\frac{\boldsymbol{V}}{\boldsymbol{W}}$ Consider; $v + \mathbf{W}$ be any element of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$. Here $v \epsilon \boldsymbol{V}$ can be expressed as a linear combination of $w_1, w_2, \ldots, w_n, v_1, v_2, \ldots, v_n$ .i.e.

$v = \propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m$

$v + \mathbf{W} = (\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_n v_n) + (\beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m) + \mathbf{W}$

$v + \mathbf{W} = (\beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m) + \mathbf{W}$

$v + \mathbf{W} = (\beta_1 w_1 + \mathbf{W}) + (\beta_2 w_2 + \mathbf{W}) + \cdots + (\beta_m w_m + \mathbf{W})$

$v + \mathbf{W} = \beta_1 (w_1 + \mathbf{W}) + \beta_2 (w_2 + \mathbf{W}) + \cdots + \beta_m (w_m + \mathbf{W})$

This shows that $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \ldots, v_n + \mathbf{W}\}$ is spans $\frac{\boldsymbol{V}}{\boldsymbol{W}}$.

Thus $\{v_1 + \mathbf{W}, v_2 + \mathbf{W}, \ldots, v_n + \mathbf{W}\}$ is a basis of $\frac{\boldsymbol{V}}{\boldsymbol{W}}$ and $dim \frac{\boldsymbol{V}}{\boldsymbol{W}} = n$

Hence $\qquad dim \frac{\boldsymbol{V}}{\boldsymbol{W}} = dim \boldsymbol{V} - dim \boldsymbol{W}$

Theorem: If **W** is subspace of a finite dimensional vector space **V**, then

i. W is finite dimensional
ii. $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V})$
iii. $\boldsymbol{W} = \boldsymbol{V} \Leftrightarrow \dim(\boldsymbol{W}) = \dim(\boldsymbol{V})$

Proof:

i. Since $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V}) = n$ and if **W** has a basis with n elements. Then by theorem **"Let W be a vector space such that one basis has *m* elements and another basis has *n* elements, then m = n."** all basis of W have the same number of elements and hence **W** is finite dimensional.
ii. If W is finite dimensional, so it has a basis $\boldsymbol{S} = \{w_1, w_2, \dots, w_m\}$. Either **S** is also a basis for **V** or it is not. If so, then $\dim(\boldsymbol{V}) = m$, which means that $\dim(\boldsymbol{W}) = \dim(\boldsymbol{V})$. If not, then because **S** is a linearly independent set it can be enlarged to a basis for **V** by theorem **"Let S be a finite set of vectors in a finite dimensional vector space V then if S is linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S"**

But this implies that $\dim(\boldsymbol{W}) < \dim(\boldsymbol{V})$. So we have shown that $\dim(\boldsymbol{W}) \leq \dim(\boldsymbol{V})$ in all cases.

iii. Assume that $\dim(\boldsymbol{W}) = \dim(\boldsymbol{V})$ and that $\boldsymbol{S} = \{w_1, w_2, \dots, w_m\}$ is a basis for **W**. If **S** is not also a basis for **V**, then being linearly independent **S** can be extended to a basis for **V** by theorem **"Let S be a finite set of vectors in a finite dimensional vector space V then if S is linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S"**. But this would mean that $\dim(\boldsymbol{W}) < \dim(\boldsymbol{V})$, which contradiction our hypothesis. Thus **S** must also be a basis for **V**, which means that $\boldsymbol{W} = \boldsymbol{V}$ The converse is obvious.

Theorem:

Let **V** be an n – dimensional vector space, and let **S** be a set in **V** with exactly ‘n’ vectors. Then **S** is a basis for **V** if and only if S spans **V** or **S** is linearly independent.

Proof:

Assume that **S** has exactly n vectors and spans **V**. To prove that **S** is a basis, we must show that **S** is linearly independent set. But if this is not so, then some vector $\boldsymbol{v}$ in S is a linear combination of the remaining vectors. If we remove this vector from **S**, then it follows from theorem;

“Let S be a non – empty set of a vectors in a vector space V then if $\boldsymbol{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, and if $\boldsymbol{S}-\{\boldsymbol{v}\}$ denotes the set obtained by removing $\boldsymbol{v}$ from S, then S and $\boldsymbol{S} \cup \{\boldsymbol{v}\}$ span the same space; that is,$\boldsymbol{span}(\boldsymbol{S}) = \boldsymbol{span}(\boldsymbol{S}-\{\boldsymbol{v}\})$”

That the remaining set of $n-1$vectors still spans V. But this is impossible since theorem **“Let V be an n – dimensional vector space, and let $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}\}$ be any basis then if a set in V has fewer than ‘n’ vectors, then it does not span V”** states that no set with fewer than n vectors can span an n – dimensional vector space. Thus **S** is linearly independent.

Assume that **S** has exactly n vectors and is a linearly independent set. To prove that **S** is a basis, we must show that S spans **V**. But if this is not so, then there is some vector $\boldsymbol{v}$ in **V** that is not in span(**S**). If we insert this vector into **S**, then it follows from theorem **“Let S be a non – empty set of a vectors in a vector space V then if S is linearly independent set, and if $\boldsymbol{v}$ is a vector in V that is outside of span(S), then the set $\boldsymbol{S} \cup \{\boldsymbol{v}\}$ that results by inserting $\boldsymbol{v}$ into S is still linearly independent”** that this set of $n+1$ vectors is still linearly independent. But this is impossible, since theorem

“Let V be an n – dimensional vector space, and let $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}\}$ be any basis then if a set in V has more than ‘n’ vectors, then it is linearly dependent”

states that no set with more than n vectors in an n – dimensional vector space can be linearly independent. Thus **S** spans **V**.

Theorem:

Let **S** be a finite set of vectors in a finite dimensional vector space **V** then if **S** spans **V** but is not a basis for **V**, then **S** can be reduced to a basis for **V** by removing appropriate vectors from **S**. (This theorem tells that; every spanning set for a subspace is either a basis for that subspace or has a basis as a subset)

Proof:

If **S** is a set of vectors that span **V** but is not a basis for **V**, then **S** is a linearly dependent set. Thus some vector $\boldsymbol{v}$ in **S** is expressible as a linear combination of the other vectors in **S**. By theorem;

"Let S be a non – empty set of a vectors in a vector space V then if *v* is a vector in S that is expressible as a linear combination of other vectors in S, and if $\boldsymbol{S}-\{\boldsymbol{v}\}$ denotes the set obtained by removing *v* from S, then S and $\boldsymbol{S}\cup\{\boldsymbol{v}\}$ span the same space; that is,$\boldsymbol{span}(\boldsymbol{S}) = \boldsymbol{span}(\boldsymbol{S}-\{\boldsymbol{v}\})$"

We can remove $\boldsymbol{v}$ from **S**, and the resulting set $\boldsymbol{S'}$ will still span **V**. If $\boldsymbol{S'}$ is linearly independent, then $\boldsymbol{S'}$ is a basis for V, and we are done. If $\boldsymbol{S'}$ is linearly dependent, then we can remove some appropriate vector from $\boldsymbol{S'}$ to produce a set $\boldsymbol{S''}$ that still span **V**. We can continue removing vectors in this way until we finally arrive at a set of vectors in $\boldsymbol{S}$ that is linearly independent and span **V**. This subset of $\boldsymbol{S}$ is a basis for **V**.

Theorem:

Let **S** be a finite set of vectors in a finite dimensional vector space **V** then if **S** is linearly independent set that is not already a basis for **V**, then **S** can be enlarged to a basis for **V** by inserting appropriate vectors into **S**. (This theorem tells that; every linearly independent set in a subspace is either a basis for that subspace or can be extended to a basis for it)

Proof:

Suppose that $\dim(\boldsymbol{V}) = n$. If **S** is a linearly independent set that is not already a basis for **V**, then **S** fails to span **V**, so there is some vector ***v*** in **V** that is not in span(**S**). By the theorem **"Let S be a non – empty set of a vectors in a vector space V then if S is linearly independent set, and if *v* is a vector in V that is outside of span(S), then the set $S \cup \{v\}$ that results by inserting v into S is still linearly independent."** We can inset ***v*** into **S**, and the resulting set $\boldsymbol{S'}$ will still be linearly independent. If $\boldsymbol{S'}$ spans **V**, then $\boldsymbol{S'}$ is a basis for **V**, and we are finished. If $\boldsymbol{S'}$ does not spans **V**, then we insert an appropriate vector into $\boldsymbol{S'}$ to produce a set $\boldsymbol{S''}$ that is still linearly independent. We can continue inserting vectors in this way until we reach a set with *n* already independent vectors in **V**. this set will be a basis for **V** by theorem **"Let V be an n – dimensional vector space, and let S be a set in V with exactly 'n' vectors. Then S is a basis for V if and only if S spans V or S is linearly independent."**

Sums

Let **U** and **W** be subsets of a vector space **V**. The sum of **U** and **W**, written **U** + **W**, consists of all sums $\boldsymbol{u}$ + $\boldsymbol{w}$ where $\boldsymbol{u}\epsilon\boldsymbol{U}$ and $\boldsymbol{w}\epsilon\boldsymbol{W}$. That is,

$\boldsymbol{U}+\boldsymbol{W}=\{\boldsymbol{v}\text{:}\ \boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w};\boldsymbol{u}\epsilon\boldsymbol{U}\text{ and }\boldsymbol{w}\epsilon\boldsymbol{W}\}$

Now suppose **U** and **W** are subspaces of **V**. Then one can easily show that **U** + **W** is a subspace of **V**. Recall that $\boldsymbol{U}\cap\boldsymbol{W}$ is also a subspace of V. The following theorem relates the dimensions of these subspaces.

Theorem:

Suppose **U** and **W** are finite-dimensional subspaces of a vector space **V**. Then $dim(\mathbf{U}+\mathbf{W})=dim\boldsymbol{U}+dim\boldsymbol{W}-dim(\boldsymbol{U}\cap\boldsymbol{W})$

Proof:

Suppose $\{v_1,v_2,\dots,v_r\}$ be a basis of $\boldsymbol{U}\cap\boldsymbol{W}$, $\{v_1,v_2,\dots,v_r,u_1,u_2,\dots,u_m\}$ be a basis of $\boldsymbol{U}$ and $\{v_1,v_2,\dots,v_r,w_1,w_2,\dots,w_n\}$ be a basis of $\boldsymbol{W}$

Then we have to show that $\{v_1,v_2,\dots,v_r,u_1,u_2,\dots,u_m,w_1,w_2,\dots,w_n\}$ be a basis of $\mathbf{U}+\mathbf{W}$

Firstly we will show that linearity condition. For this consider;

$\propto_1 v_1+\propto_2 v_2+\cdots+\propto_r v_r+\beta_1u_1+\beta_2u_2+\cdots+\beta_mu_m+\gamma_1w_1+\gamma_2w_2+\cdots+\gamma_nw_n=0_v$

$\Rightarrow\propto_1 v_1+\propto_2 v_2+\cdots+\propto_r v_r+\beta_1u_1+\beta_2u_2+\cdots+\beta_mu_m=-\gamma_1w_1-\gamma_2w_2-\cdots-\gamma_nw_n$ ……………..(i)

Since LHS of (i) is in $\boldsymbol{U}$ so does RHS also will be in $\boldsymbol{U}$

i.e. $-\gamma_1w_1-\gamma_2w_2-\cdots-\gamma_nw_n\in\boldsymbol{U}$

also $-\gamma_1w_1-\gamma_2w_2-\cdots-\gamma_nw_n\in\boldsymbol{W}$

therefore $-\gamma_1w_1-\gamma_2w_2-\cdots-\gamma_nw_n\in\boldsymbol{U}\cap\boldsymbol{W}$

as $\{v_1,v_2,\dots,v_r\}$ be a basis of $\boldsymbol{U}\cap\boldsymbol{W}$ then for $\delta_i\in\boldsymbol{F}$ we have

$\delta_1 v_1 + \delta_2 v_2 + \cdots + \delta_r v_r + \gamma_1 w_1 + \gamma_2 w_2 + \cdots + \gamma_n w_n = 0_v$

Since $\{v_1, v_2, \dots, v_r, w_1, w_2, \dots, w_n\}$ be a basis of $\boldsymbol{W}$

$\Rightarrow \delta_1 = \delta_2 = \cdots = \delta_r = \gamma_1 = \gamma_2 = \cdots = \gamma_n = 0$

So that (i) becomes

$\propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_r v_r + \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m = 0_v$

But $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_m\}$ be a basis of $\boldsymbol{U}$

$\Rightarrow \propto_1 = \propto_2 = \cdots = \propto_r = \beta_1 = \beta_2 = \cdots = \beta_m = 0$

i.e. each $\propto_i = \beta_i = \gamma_i = 0$

Hence $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_m, w_1, w_2, \dots, w_n\}$ is linearly independent.

Now we to show spanning condition. For this suppose $x + y \in \boldsymbol{U} + \boldsymbol{W}$

i.e. $x \in \boldsymbol{U}, y \in \boldsymbol{W}$. Also we know that $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_m\}$ be a basis of $\boldsymbol{U}$ and $\{v_1, v_2, \dots, v_r, w_1, w_2, \dots, w_n\}$ be a basis of $\boldsymbol{W}$

Then $x = \propto_1 v_1 + \propto_2 v_2 + \cdots + \propto_r v_r + \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m$

And $y = \propto_1' v_1 + \propto_2' v_2 + \cdots + \propto_r' v_r + \beta_1' w_1 + \beta_2' w_2 + \cdots + \beta_n' w_n$

Adding both we get

$x + y = (\propto_1 + \propto_1')v_1 + (\propto_2 + \propto_2')v_2 + \cdots + (\propto_r + \propto_r')v_r + \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m + \beta_1' w_1 + \beta_2' w_2 + \cdots + \beta_n' w_n$

This implies that $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_m, w_1, w_2, \dots, w_n\}$ spans $\boldsymbol{U} + \boldsymbol{W}$

From both conditions we conclude that $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_m, w_1, w_2, \dots, w_n\}$ is a basis of $\boldsymbol{U} + \boldsymbol{W}$

Therefore $\boldsymbol{U} + \boldsymbol{W}$ is finite dimensional. And

$dim(\mathbf{U} + \mathbf{W}) = r + m - n = (r + m) + (r + n) - r$

$dim(\mathbf{U} + \mathbf{W}) = dim\boldsymbol{U} + dim\boldsymbol{W} - dim(\boldsymbol{U} \cap \boldsymbol{W})$

Example: Let $\boldsymbol{V} = \boldsymbol{M}_{2;2}$, the vector space of 2×2 matrices. Let **U** consist of those matrices whose second row is zero, and let **W** consist of those matrices whose second column is zero. Then

$$\mathbf{U} = \left\{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}\right\} \quad , \quad \mathbf{W} = \left\{\begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}\right\} \ , \ \mathbf{U} + \mathbf{W} = \left\{\begin{bmatrix} a & b \\ c & 0 \end{bmatrix}\right\} \ , \quad \mathbf{U} \cap \mathbf{W} = \left\{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\right\}$$

That is, $\mathbf{U} + \mathbf{W}$ consists of those matrices whose lower right entry is 0, and $\mathbf{U} \cap \mathbf{W}$ consists of those matrices whose second row and second column are zero. Note that $dim(\mathbf{U}) = 2$, $dim(\mathbf{W}) = 2$, $dim(\mathbf{U} + \mathbf{W}) = 1$. Also, $dim(\mathbf{U} \cap \mathbf{W}) = 3$, which is expected from Theorem. That is,

$$dim(\mathbf{U} + \mathbf{W}) = dim\boldsymbol{U} + dim\boldsymbol{W} - dim(\boldsymbol{U} \ \cap \ \boldsymbol{W}) = 2 + 2 - 1 = 3$$

Practice:

1. Give an example of a vector space **V** and its subspace **W** such that **V**,**W** and $\frac{\boldsymbol{V}}{\boldsymbol{W}}$ are infinite dimensional and $dim\frac{\boldsymbol{V}}{\boldsymbol{W}} = dim\boldsymbol{V} - dim\boldsymbol{W}$
2. Let **U** and **W** be subspaces of a vector space **V**. Then show that;

i. $\mathbf{U} + \mathbf{V}$ is subspace of **V**

ii. **U** and **W**are contained in $\mathbf{U} + \mathbf{W}$

iii. $\mathbf{U} + \mathbf{W}$ is the smallest subspace containing **U** and **W**
i.e. $\mathbf{U} + \mathbf{W} = \text{span}(\mathbf{U}, \mathbf{W})$

iv. $\mathbf{W} + \mathbf{W} = \mathbf{W}$

3. Consider the following subspaces of $\mathbf{R}^5$;
$\mathbf{U} = \text{span}(u_1, u_2, u_3) = span\{(1,3,-2,2,3), (1,4,-3,4,2), (2,3,-1,-2,9)\}$
$\mathbf{W} = \text{span}(w_1, w_2, w_3) = span\{(1,3,0,2,1), (1,5,-6,6,3), (2,5,3,2,1)\}$

   Find a basis and the dimension of (a) $\mathbf{U} + \mathbf{W}$ (b) $\boldsymbol{U} \ \cap \ \boldsymbol{W}$

4. Suppose **U** and **W** are distinct four dimensional subspaces of a vector space **V**, where $dim\boldsymbol{V} = 6$. Find the possible dimension of $\boldsymbol{U} \ \cap \ \boldsymbol{W}$

5. Let $\boldsymbol{W_1}$ *and* $\boldsymbol{W_2}$ be the subset of all upper and lower triangular matrices of $\boldsymbol{F}^{\boldsymbol{n}\times\boldsymbol{n}}$ show that both are subspaces of $\boldsymbol{F}^{\boldsymbol{n}\times\boldsymbol{n}}$,
find $dim(\boldsymbol{W_1}, \boldsymbol{W_2})$, $dim(\boldsymbol{W_1} + \boldsymbol{W_2})$, $dim(\boldsymbol{W_1} \cap \boldsymbol{W_2})$
also verify $dim(\boldsymbol{W_1} + \boldsymbol{W_2}) = dim\boldsymbol{W_1} + dim\boldsymbol{W_2} - dim(\boldsymbol{W_1} \cap \boldsymbol{W_2})$

Direct Sums

The vector space **V** is said to be the direct sum of its subspaces **U** and **W**, denoted by $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$ if every $\boldsymbol{v} \in$ **V** can be written in one and only one way as $v = u + w$ where $\boldsymbol{u} \in$ **U** and $\boldsymbol{w} \in$ **W**.

The following theorem characterizes such a decomposition.

Theorem:

The vector space **V** is the direct sum of its subspaces **U** and **W** if and only if:

(i) $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$, (ii) $\boldsymbol{U} \cap \boldsymbol{W} = \{0\}$.

Proof: Suppose $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$ then every $\boldsymbol{v} \in$ **V** can be uniquely written in the form $v = u + w$ where $\boldsymbol{u} \in$ **U** and $\boldsymbol{w} \in$ **W**. Thus in particular, $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$

Now suppose $\boldsymbol{v} \in \boldsymbol{U} \cap \boldsymbol{W}$ then

(i) $v = v + 0$ where $\boldsymbol{v} \in$ **U** and $\boldsymbol{0} \in$ **W**.
(ii) $v = 0 + v$ where $\boldsymbol{0} \in$ **U** and $\boldsymbol{v} \in$ **W**.

Thus $v = 0 + 0 = 0$ and $\boldsymbol{U} \cap \boldsymbol{W} = \{0\}$

On the other hand, suppose that $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$, and $\boldsymbol{U} \cap \boldsymbol{W} = \{0\}$. Let $\boldsymbol{v} \in$ **V** because $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$, there exists $\boldsymbol{u} \in$ **U** and $\boldsymbol{w} \in$ **W** such that $v = u + w$. We need to show that such a sum is unique. Suppose also that $v = u' + w'$ where $\boldsymbol{u}' \in$ **U** and $\boldsymbol{w}' \in$ **W** then

$u + w = u' + w'$ also $u - u' = w' - w$

But $u - u' \in$ **U** and $w' - w \in$ **W** then by $\boldsymbol{U} \cap \boldsymbol{W} = \{0\}$

$u - u' = 0$ and $w' - w = 0$ and so $u = u'$ and $w' = w$

Thus such a sum for $\boldsymbol{v} \in$ **V** is unique, and $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$

Example: Consider the vector space $\boldsymbol{V} = \boldsymbol{R^3}$

(a) Let **U** be the xy-plane and let **W** be the yz-plane; that is,

$\mathbf{U} = \{(a, b, 0): a, b \in \boldsymbol{R}\}$ and $\mathbf{U} = \{(0, b, c): b, c \in \boldsymbol{R}\}$

Then $\boldsymbol{R^3} = \boldsymbol{U} + \boldsymbol{W}$, because every vector in $\boldsymbol{R^3}$ is the sum of a vector in **U** and a vector in **W**. However, $\boldsymbol{R^3}$ is not the direct sum of **U** and **W**, because such sums are not unique. For example,

$(3,5,7) = (3,1,0) + (0,4,7)$ and also $(3,5,7) = (3,-4,0) + (0,9,7)$

(b) Let **U** be the xy-plane and let **W** be the z-axis; that is,

$\mathbf{U} = \{(a, b, 0): a, b \in \boldsymbol{R}\}$ and $\mathbf{U} = \{(0,0, c): c \in \boldsymbol{R}\}$

Now any vector $(a, b, c) \in \boldsymbol{R^3}$ can be written as the sum of a vector in **U** and a vector in **V** in one and only one way:

$(a, b, c) = (a, b, 0) + (0,0, c)$

Accordingly, $\boldsymbol{R^3}$ is the direct sum of **U** and **W**; that is, $\boldsymbol{R^3} = \boldsymbol{U} \oplus \boldsymbol{W}$

General Direct Sums

The notion of a direct sum is extended to more than one factor in the obvious way. That is, **V** is the direct sum of subspaces $\boldsymbol{W_1}, \boldsymbol{W_2}, \ldots, \boldsymbol{W_r}$, written

$\boldsymbol{V} = \boldsymbol{W_1} \oplus \boldsymbol{W_2} \oplus \ldots \oplus \boldsymbol{W_r}$ if every vector $\boldsymbol{v} \in \mathbf{V}$ can be written in one and only one way as $v = w_1 + w_2 + \cdots + w_r$ where $w_1 \in \boldsymbol{W_1}, w_2 \in \boldsymbol{W_2}, \ldots, w_r \in \boldsymbol{W_r}$

Theorem:

Suppose $\boldsymbol{V} = \boldsymbol{W_1} \oplus \boldsymbol{W_2} \oplus \ldots \oplus \boldsymbol{W_r}$. Also, for each k, suppose $\mathbf{S_k}$ is a linearly independent subset of $\mathbf{W_k}$ .Then

(a) The union $S = \cup_k S_k$ is linearly independent in **V**.

(b) If each $\mathbf{S_k}$ is a basis of $\mathbf{W_k}$, then $\cup_k S_k$ is a basis of V.

(c) $dim\boldsymbol{V} = dim\boldsymbol{W_1} + dim\boldsymbol{W_2} + \cdots + dim\,\boldsymbol{W_r}$

Theorem: (For two factors)

Suppose $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$. Suppose $\boldsymbol{S} = \{u_1, u_2, \dots, u_m\}$ and $\boldsymbol{S'} = \{w_1, w_2, \dots, w_n\}$ are linearly independent subset of $\boldsymbol{U}\ and\ \boldsymbol{W}$.Then

(a) The union $\mathbf{S} \cup \boldsymbol{S'}$ is linearly independent in **V**.

(b) If each $\boldsymbol{S} = \{u_1, u_2, \dots, u_m\}$ and $\boldsymbol{S'} = \{w_1, w_2, \dots, w_n\}$ are basis of $\boldsymbol{U}\ and\ \boldsymbol{W}$, then $\mathbf{S} \cup \boldsymbol{S'}$ is a basis of V.

(c) $dim\boldsymbol{V} = dim\boldsymbol{U} + dim\boldsymbol{W}$

Proof:

a) Suppose that $\propto_1 u_1 + \propto_2 u_2 + \dots + \propto_m u_m + \beta_1 w + \beta_2 w_2 + \dots + \beta_n w_n = 0$
Where $\propto_i, \beta_j$ are scalars then
$(\propto_1 u_1 + \propto_2 u_2 + \dots + \propto_m u_m) + (\beta_1 w + \beta_2 w_2 + \dots + \beta_n w_n) = 0 = 0 + 0$
Where $0 , \propto_1 u_1 + \dots + \propto_m u_m \in \boldsymbol{U}$ and $0 , \beta_1 w + \beta_2 w_2 + \dots + \beta_n w_n \in \boldsymbol{W}$
Because such a sum for 0 is unique, this leads to
$\propto_1 u_1 + \propto_2 u_2 + \dots + \propto_m u_m = 0$ and $\beta_1 w + \beta_2 w_2 + \dots + \beta_n w_n = 0$
Because **S** is linearly independent for each $\propto_i = 0$ and $\boldsymbol{S'}$ is linearly independent for each $\beta_j = 0$ therefore is $\mathbf{S} \cup \boldsymbol{S'}$ is linearly independent.

b) By theorem "Suppose $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$. Suppose $\boldsymbol{S} = \{u_1, u_2, \dots, u_m\}$ and $\boldsymbol{S'} = \{w_1, w_2, \dots, w_n\}$ are linearly independent subset of $\boldsymbol{U}\ and\ \boldsymbol{W}$.Then the union $\mathbf{S} \cup \boldsymbol{S'}$ is linearly independent in **V**." $\mathbf{S} \cup \boldsymbol{S'}$ is linearly independent, and by problem "Suppose that $\boldsymbol{U}\ and\ \boldsymbol{W}$are subspaces of a vector space **V** and that $\boldsymbol{S} = \{u_i\}$ spans **U** and $\boldsymbol{S'} = \{u_i\}$ spans **W**. This show that $\mathbf{S} \cup \boldsymbol{S'}$ spans $\boldsymbol{U} + \boldsymbol{W}$"
Thus $\mathbf{S} \cup \boldsymbol{S'}$ is a basis of V.

c) Suppose $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$. Suppose $\boldsymbol{S} = \{u_1, u_2, \dots, u_m\}$ and $\boldsymbol{S'} = \{w_1, w_2, \dots, w_n\}$ are linearly independent subset of $\boldsymbol{U}\ and\ \boldsymbol{W}$.Then If each $\boldsymbol{S} = \{u_1, u_2, \dots, u_m\}$ and $\boldsymbol{S'} = \{w_1, w_2, \dots, w_n\}$ are basis of $\boldsymbol{U}\ and\ \boldsymbol{W}$, then $\mathbf{S} \cup \boldsymbol{S'}$ is a basis of V. then it follows directly $\quad dim\boldsymbol{V} = dim\boldsymbol{U} + dim$

Theorem: (Just read): Suppose $\boldsymbol{V} = \boldsymbol{W_1} + \boldsymbol{W_2} + \dots + \boldsymbol{W_r}$.

And $dim\boldsymbol{V} = \sum_k dim\boldsymbol{W_k}$ then $\boldsymbol{V} = \boldsymbol{W_1} \oplus \boldsymbol{W_2} \oplus \dots \oplus \boldsymbol{W_r}$

Practice:

1. Consider the following subspaces of $\mathbf{R^3}$;
   $\mathbf{U} = \{(a, b, c): a = b = c\}$, $\mathbf{W} = \{(0, b, c)\}$ (**W** is yz – plane)
   Show that $\boldsymbol{R^3} = \boldsymbol{U} \oplus \boldsymbol{W}$
2. Suppose that $\boldsymbol{U}$ *and* $\boldsymbol{W}$are subspaces of a vector space **V** and that $\boldsymbol{S} = \{u_i\}$ spans **U** and $\boldsymbol{S}' = \{u_i\}$ spans **W**. Then show that $\boldsymbol{S} \cup \boldsymbol{S}'$ spans $\boldsymbol{U} + \boldsymbol{W}$.

Cardinality of a vector space V(F):

Let $dim(\mathbf{V}) = n$ and $\boldsymbol{B} = \{x_1, x_2, \dots, x_n\}$ is any basis of **V(F)** then cardinality of vector space is given as $|\boldsymbol{V}(\boldsymbol{F})| = p^n$ where p is any prime.

Number of Basis of a vector space V(F):

Let $dim(\mathbf{V}) = n$ then;

i. Number of distinct basis $= \frac{(p^n-1)(p^n-p)(p^n-p^2)\dots(p^n-p^{n-1})}{n!}$

ii. Number of ordered basis $= (p^n - 1)(p^n - p)(p^n - p^2) \dots (p^n - p^{n-1})$

Number of subspaces of a vector space V(F):

Let **V(F)** be a vector space and $\boldsymbol{W}$ be a subspace of **V(F)** . Let $dim(\mathbf{W}) = r$ and $\boldsymbol{B_W} = \{x_1, x_2, \dots, x_r\}$ then

i. Number of basis in $\mathbf{W} = \frac{(p^r-1)(p^r-p)(p^r-p^2)\dots(p^r-p^{r-1})}{r!}$

ii. Number of basis of V selective 'r' linearly independent vectors from **V** $= \frac{(p^n-1)(p^n-p)(p^n-p^2)\dots(p^n-p^{r-1})}{r!}$

iii. Number of subspaces of $dim(\text{r}) = t_r$ $= \frac{(p^n-1)(p^n-p)(p^n-p^2)\dots(p^n-p^{r-1})}{(p^r-1)(p^r-p)(p^r-p^2)\dots(p^r-p^{r-1})}$
where $t_n = 1 = t_0$

Coordinates

If $\boldsymbol{S} = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space **V**, and $\vec{v} = k_1v_1 + k_2v_2 + \cdots + k_nv_n$ is the expression for a vector $\vec{v}$ in terms of the basis **S**, then the scalars $k_1, k_2, \dots, k_n$ are called the **coordinates** of $\vec{v}$ relative to the basis **S**.

The vector $(k_1, k_2, \dots, k_n)$ in $\mathbf{R}^n$ constructed from these coordinates is called the **coordinate vector $\vec{v}$ relative to S**; it is denoted by $(\vec{v})_{\boldsymbol{S}} = (k_1, k_2, \dots, k_n)$

Remark:

The above 'n' scalars $k_1, k_2, \dots, k_n$ also form the coordinate column vector $(k_1, k_2, \dots, k_n)^T$ of $\vec{v}$ relative to **S**. The choice of the column vector rather than the row vector to represent v depends on the context in which it is used.

Coordinates relative to the standard basis for $\mathbf{R}^n$

In the special case where **V** = $\mathbf{R}^n$ and **S** is the standard basis, the coordinate vector $(\vec{v})_{\boldsymbol{S}}$ and the vector $\vec{v}$ are the same; that is , $\vec{v} = (\vec{v})_{\boldsymbol{S}}$

For example in $\mathbf{R}^3$ the representation of a vector $\vec{v} = (a, b, c)$ as a linear combination of the vectors in the standard basis $\boldsymbol{S} = \{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is $\vec{v} = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}$ so the coordinate vector relative to the basis is $(\vec{v})_{\boldsymbol{S}} = (a, b, c)$ which is the same as the vector $\vec{v}$.

Example:

Consider real space $\mathbf{R}^3$. The following vectors form a basis S of $\mathbf{R}^3$:

$\boldsymbol{v_1} = (1, -1, 0)$; $\boldsymbol{v_2} = (1,1,0)$; $\boldsymbol{v_3} = (0,1,1)$

The coordinates of $\boldsymbol{v} = (5,3,4)$ relative to the basis **S** are obtained as follows.

Set $\boldsymbol{v} = x\boldsymbol{v_1} + y\boldsymbol{v_2} + z\boldsymbol{v_3}$; that is, set **v** as a linear combination of the basis vectors using unknown scalars x, y, z. This yield

$$\begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix} = x\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

The equivalent system of linear equations is as follows:

$x + y = 5$

$-x + y + z = 3$

$z = 4$

The solution of the system is $x = 3, y = 2, z = 4$. Thus,

$\boldsymbol{v} = 3\boldsymbol{v_1} + 2\boldsymbol{v_2} + 4\boldsymbol{v_3}$

And so $(\vec{v})_{\boldsymbol{S}} = (3,2,4)$

Practice:

1) Find the coordinate vectors of **w** relative to the basis $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ for $\mathbf{R}^2$
   i. $\boldsymbol{v_1} = (2,-4), \boldsymbol{v_2} = (3,8)$, $\mathbf{w} = (1,1)$
   ii. $\boldsymbol{v_1} = (1,1), \boldsymbol{v_2} = (0,2)$, $\mathbf{w} = (a,b)$
   iii. $\boldsymbol{v_1} = (1,-1), \boldsymbol{v_2} = (1,1)$, $\mathbf{w} = (1,0)$
   iv. $\boldsymbol{v_1} = (1,-1), \boldsymbol{v_2} = (1,1)$, $\mathbf{w} = (0,1)$
2) Find the coordinate vectors of **v** relative to the basis $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ for $\mathbf{R}^3$
   i. $\boldsymbol{v_1} = (1,2,1), \boldsymbol{v_2} = (2,9,0)$, $\boldsymbol{v_3} = (3,3,4)$, $\mathbf{v} = (5,-1,9)$
   ii. $\boldsymbol{v_1} = (1,0,0), \boldsymbol{v_2} = (2,2,0)$, $\boldsymbol{v_3} = (3,3,3)$, $\mathbf{v} = (2,-1,3)$
   iii. $\boldsymbol{v_1} = (1,2,3), \boldsymbol{v_2} = (-4,5,6)$, $\boldsymbol{v_3} = (7,-8,9)$, $\mathbf{v} = (5,-12,3)$

3) Relative to the basis $\boldsymbol{S} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\} = \{(1,1),(2,3)\}$ of $\mathbf{R}^2$, find the coordinate vector of $\mathbf{v} = (4,-3)$

Coordinates relative to the standard basis for $\mathbf{P_n}$

In the special case where $\mathbf{V} = \mathbf{P_n}$ the given formula for

$\boldsymbol{p}(x) = k_0 + k_1 x + k_2 x^2 \dots + k_n x^n$ expresses the polynomials as a linear combination of the standard basis vectors $\mathbf{S} = \{1, x, x^2, \dots, x^n\}$. Thus the coordinate vectors for $\boldsymbol{p}$ relative to **S** is $(\boldsymbol{p})_{\boldsymbol{S}} = (k_1, k_2, \dots, k_n)$

Example: Consider the vector space $\mathbf{P_2}$ of polynomials of degree ≤ 2. The polynomials $\boldsymbol{p_1} = x + 1$, $\boldsymbol{p_2} = x - 1$, $\boldsymbol{p_3} = (x-1)^2 = x^2 - 2x + 1$

form a basis **S** of $\mathbf{P_2}$. The coordinate vector $(\vec{v})$of $\vec{v} = 2x^2 - 5x + 9$ relative to **S** is obtained as follows.

Set $\boldsymbol{v} = a\boldsymbol{p_1} + b\boldsymbol{p_2} + c\boldsymbol{p_3}$ using unknown scalars a,b,c, and simplify:

$$2x^2 - 5x + 9 = a(x+1) + b(x-1) + c(x^2 - 2x + 1)$$

$$2x^2 - 5x + 9 = cx^2 + (a + b - 2c)x + (a - b + c)$$

Then set the coefficients of the same powers of ‘x’ equal to each other to obtain the system $c = 2$, $a + b - 2c = -5$, $a - b + c = 9$

The solution of the system is $a = 3$, $b = -4$, $c = 2$. Thus,

$\boldsymbol{v} = 3\boldsymbol{p_1} - 4\boldsymbol{p_2} + 2\boldsymbol{p_3}$ and hence; $(\boldsymbol{p})_{\boldsymbol{S}} = (3, -4, 2)$

Practice: Find the coordinate vectors of **p** relative to the basis $\boldsymbol{S} = \{\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}\}$ for $\mathbf{P_2}$ **,** $\mathbf{P_3}$ and $\mathbf{P_4}$

i. $\boldsymbol{p} = x^2 - 3x + 4$, $\boldsymbol{p_1} = 1$, $\boldsymbol{p_2} = x$, $\boldsymbol{p_3} = x^2$
ii. $\boldsymbol{p} = x^2 - x + 2$, $\boldsymbol{p_1} = 1 + x$, $\boldsymbol{p_2} = 1 + x^2$, $\boldsymbol{p_3} = x + x^2$
iii. $\boldsymbol{p} = 2x^2 - x + 7$, $\boldsymbol{p_1} = 1 + x + x^2$, $\boldsymbol{p_2} = x + x^2$, $\boldsymbol{p_3} = x^2$
iv. $\boldsymbol{p} = -3x^2 + 17x + 2$, $\boldsymbol{p_1} = 1 + 2x + x^2$, $\boldsymbol{p_2} = 2 + 9x$, $\boldsymbol{p_3} = x^2$
v. $\boldsymbol{p} = 8x^3 + 8x^2 - 4x - 1$, $\boldsymbol{p_1} = 1$, $\boldsymbol{p_2} = 2x$, $\boldsymbol{p_3} = -2 + 4x^2$, $\boldsymbol{p_4} = -12x + 8x^3$
vi. $\boldsymbol{p} = -x^3 + 9x^2 - 10x$, $\boldsymbol{p_1} = 1$, $\boldsymbol{p_2} = 1 - x$, $\boldsymbol{p_3} = 2 - 4x + x^2$, $\boldsymbol{p_4} = 6 - 18x + 9x^2 - x^3$

Coordinates relative to the standard basis for M_{22}

In the special case where **V = M_{22}** the representation of a vector $\boldsymbol{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ could be expressible as the linear combination of the standard basis vectors in matrices as follows;

$$\boldsymbol{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

So the coordinate vector of **B** relative to **S** is $(\boldsymbol{B})_{\boldsymbol{S}} = (a, b, c, d)$

Example:

Consider the vector space M_{22} . The matrices $\boldsymbol{A_1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \boldsymbol{A_2} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \boldsymbol{A_3} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{A_4} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

form a basis **S** of M_{22} . The coordinate vector $(\boldsymbol{A})$of $\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix}$ relative to **S** is obtained as follows.

Set $\boldsymbol{A} = a\boldsymbol{A_1} + b\boldsymbol{A_2} + c\boldsymbol{A_3} + d\boldsymbol{A_4}$ using unknown scalars a,b,c,d and simplify:

$$\begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix} = a\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + d\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & -7 \end{bmatrix} = \begin{bmatrix} a+c+d & a-b-c \\ a+b & a \end{bmatrix}$$

Then set the coefficients of the same powers of 'x' equal to each other to obtain the system $a = -7$, $a + c + d = 2,$ $a - b - c = 3$, $a + b = 4$

The solution of the system is $a = -7$, $b = 11$, $c = -21$, $d = 30$. Thus,

$\boldsymbol{A} = -7\boldsymbol{A_1} + 11\boldsymbol{A_2} - 21\boldsymbol{A_3} + 30\boldsymbol{A_4}$ and hence; $(\boldsymbol{A})_{\boldsymbol{S}} = (-7, 11, -21, 30)$
(Note that the coordinate vectors of **A** is a vector in $\mathbf{R}^4$, because dim $\boldsymbol{M_{22}} = 4$)

Practice:

1) Find the coordinate vectors of **A** relative to the basis $\boldsymbol{S} = \{\boldsymbol{A_1}, \boldsymbol{A_2}, \boldsymbol{A_3}, \boldsymbol{A_4}\}$ for $\boldsymbol{M_{22}}$

   i. $\boldsymbol{A_1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \boldsymbol{A_2} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \boldsymbol{A_3} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \boldsymbol{A_4} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

   ii. $\boldsymbol{A_1} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{A_2} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{A_3} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{A_4} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} 6 & 2 \\ 5 & 3 \end{bmatrix}$

2) Consider the coordinate vector $(\boldsymbol{B})_{\boldsymbol{S}} = \begin{bmatrix} -8 \\ 7 \\ 6 \\ 3 \end{bmatrix}$ find **B** if **S** is the basis in

   $\begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

Geometrical interpretation of the coordinates of a vector relative to a basis S:

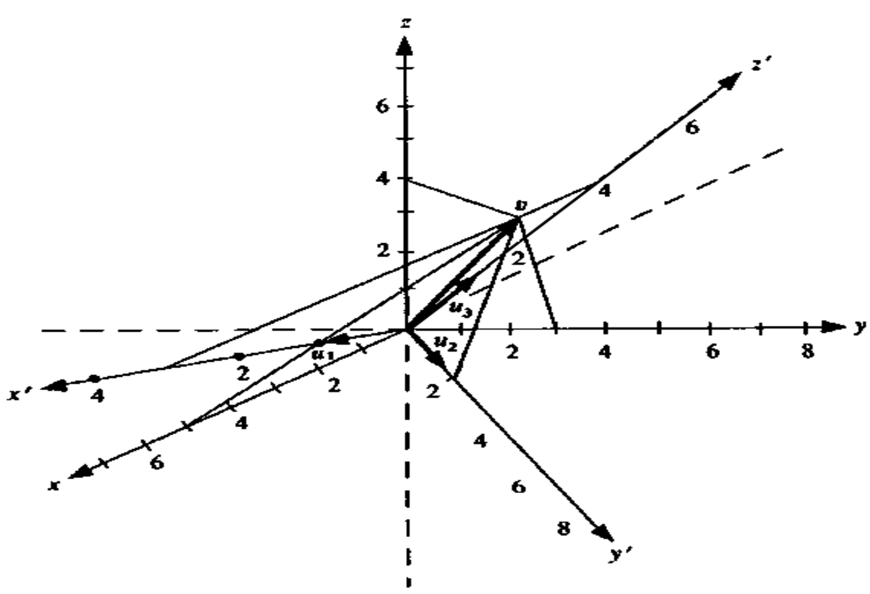
There is a geometrical interpretation of the coordinates of a vector $\vec{v}$ relative to a basis **S** for the real space $\mathbf{R}^n$, which we illustrate using the basis S of $\mathbf{R}^3$ in Example the following vectors form a basis **S** of $\mathbf{R}^3$:

$\boldsymbol{v_1} = (1,-1,0)$; $\boldsymbol{v_2} = (1,1,0)$; $\boldsymbol{v_3} = (0,1,1)$

The coordinates of $\boldsymbol{v} = (5,3,4)$ relative to the basis **S** are $(\vec{v})_S = (3,2,4)$. First consider the space $\mathbf{R}^3$ with the usual x, y, z axes. Then the basis vectors determine a new coordinate system of $\mathbf{R}^3$, say with x', y', z' axes as shown in **following figure**. That is;

i. The $x' -$ axis is in the direction of $\boldsymbol{v_1}$ with unit length $|\boldsymbol{v_1}|$
ii. The $y' -$ axis is in the direction of $\boldsymbol{v_2}$ with unit length $|\boldsymbol{v_2}|$
iii. The $z' -$ axis is in the direction of $\boldsymbol{v_3}$ with unit length $|\boldsymbol{v_3}|$

Then each vector $\boldsymbol{v} = (a, b, c)$, or equivalently the point $P(a, b, c)$ in $\mathbf{R}^3$ will have new coordinates with respect to the new x', y', z' axes. These new coordinates are precisely $(\vec{v})_S$, the coordinates of $\vec{v}$ with respect to the basis **S**, thus as shown in example, the coordinates of the point $P(5,3,4)$ with the new axes form the vector $(\vec{v})_S = (3,2,4)$



$\boldsymbol{v} = (5, 3, 4) = [3, 2, 4]$

CHAPTER # 4

INNER PRODUCT SPACES

In this chapter we will generalize the ideas of length, angle, distance and orthogonality. We will also discuss various applications of these ideas.

Norm of Vector:

Consider a vectors **v** in $\mathbf{R^n}$, then norm, length or size of vector is non – negative square root of $\vec{v}.\vec{v}$ denoted by $\|\vec{v}\|$ and defined as follows;

$$\|\vec{v}\| = \sqrt{\vec{v}.\vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \quad \text{for } \vec{v} = (v_1, v_2, \dots, v_n)$$

Note:

a) $\|\vec{v}\| \geq 0$
b) $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = 0$
c) $\|k\vec{v}\| = |k|\|\vec{v}\|$ for any scalar 'k'
d) a vector is said to be **unit vector** if $\|\vec{v}\| = 1$ or equivalently $\vec{v}.\vec{v} = 1$
e) for any non – zero vector **v** in $\mathbf{R^n}$ the vector $\hat{v} = \frac{\vec{v}}{\|\vec{v}\|}$ is the unit vector in the same direction as $\vec{v}$ and the process of finding $\hat{v}$ is called **normalizing** $\vec{v}$

Example:

- Consider $\vec{v} = (-3,2,1)$ in $\boldsymbol{R^3}$ then $\|\vec{v}\| = \sqrt{(-3)^2 + 2^2 + 1^2} = \sqrt{14}$
- Also $\vec{v} = (2,-1,3,-5)$ in $\boldsymbol{R^4}$
  Then $\|\vec{v}\| = \sqrt{2^2 + (-1)^2 + 3^2 + (-5)^2} = \sqrt{39}$

Example (Normalizing a vector):

Consider $\vec{v} = (2,2,-1)$ Then $\|\vec{v}\| = \sqrt{2^2 + 2^2 + (-1)^2} = \sqrt{9} = 3$

Then clearly $\vec{u} = \hat{v} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{3}(2,2,-1)$
Thus $\vec{u}$ is the unit vector that has the same direction as $\vec{v} = (2,2,-1)$

PRACTICE:

Find the norm of $\vec{v}$ and a unit vector that is oppositely directed to $\vec{v}$

i. $\vec{v} = (2,2,2)$
ii. $\vec{v} = (1,-1,2)$
iii. $\vec{v} = (1,0,2,1,3)$
iv. $\vec{v} = (-2,3,3,-1)$

Find $\hat{v}$ then change its sign.

The Standard Unit Vectors in $\boldsymbol{R^n}$

Vectors given as follows are called Standard Unit Vectors in $\mathbf{R^n}$

$\hat{e}_1 = (1,0,0,\dots,0), \hat{e}_2 = (0,1,0,\dots,0), \dots, \quad \hat{e}_n = (0,0,0,\dots,1)$

Note:

f) $\hat{\imath} = (1,0), \hat{\jmath} = (0,1)$ are called standard unit vectors in $\boldsymbol{R^2}$
g) $\hat{\imath} = (1,0,0), \hat{\jmath} = (0,1,0), \hat{k} = (0,0,1)$ are called Standard unit vectors in $\boldsymbol{R^3}$
h) Every vector $\vec{v} = (v_1, v_2, \dots, v_n)$ can be expressed as the **linear combination** of standard unit vectors. e.g.
$\vec{v} = (v_1, v_2, \dots, v_n) = v_1\hat{e}_1 + v_2\hat{e}_2 + \cdots + v_n\hat{e}_3$
Like
$(2,-3,4) = 2(1,0,0) - 3(0,1,0) + 4(0,0,1) = 2\hat{\imath} - 3\hat{\jmath} + 4\hat{k}$
Also
$(7,3,-4,5) = 7(1,0,0,0) + 3(0,1,0,0) - 4(0,0,1,0) + 5(0,0,0,1)$
$(7,3,-4,5) = 7\hat{e}_1 + 3\hat{e}_2 - 4\hat{e}_3 + 5\hat{e}_4$

Distance between vectors:

Consider two non – zero vectors **u** and **v** in $\mathbf{R^n}$, say $\vec{u} = (u_1, u_2, \dots, u_n)$ and $\vec{v} = (v_1, v_2, \dots, v_n)$ then distance between them is given as follows;

$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$

Example: if $\vec{u} = (1,3,-2,7)$ and $\vec{v} = (0,7,2,2)$ then

$$d(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\| = \sqrt{(u_1-v_1)^2+(u_2-v_2)^2+(u_3-v_3)^2+(u_4-v_4)^2}$$

$$d(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\| = \sqrt{(1-0)^2+(3-7)^2+(-2-2)^2+(7-2)^2} = \sqrt{58}$$

PRACTICE:

1) Evaluate the given expression with $\vec{u} = (2,-2,3)$, $\vec{v} = (1,-3,4)$ and $\vec{w} = (3,6,-4)$
   i. $\|\vec{u}+\vec{v}\|$
   ii. $\|\vec{u}\| + \|\vec{v}\|$
   iii. $\|-2\vec{u}+2\vec{v}\|$
   iv. $\|3\vec{u}-5\vec{v}+\vec{w}\|$
   v. $\|\vec{u}+\vec{v}+\vec{w}\|$
   vi. $\|\vec{u}-\vec{v}\|$
   vii. $\|3\vec{v}\| - 3\|\vec{v}\|$
   viii. $\|\vec{u}\| - \|\vec{v}\|$

2) Evaluate the given expression with $\vec{u} = (-2,-1,4,5)$, $\vec{v} = (3,1,-5,7)$ and $\vec{w} = (-6,2,1,1)$
   i. $\|3\vec{u}-5\vec{v}+\vec{w}\|$
   ii. $\|3\vec{u}\| - 5\|\vec{v}\| + \|\vec{w}\|$
   iii. $\|\vec{u}\|\|\vec{v}\|$
   iv. $\|\vec{u}\| + \|-2\vec{v}\| + \|-3\vec{w}\|$
   v. $\big\|\|\vec{u}-\vec{v}\|\vec{w}\big\|$

3) Let $\vec{v} = (-2,3,0,6)$. Find all scalars ‘k’ such that $\|k\vec{v}\| = 5$
4) Let $\vec{v} = (1,1,2,-3,1)$. Find all scalars ‘k’ such that $\|k\vec{v}\| = 4$

Projection of vectors onto (along) another vector:

Consider two non – zero vectors **u** and **v** in $\mathbf{R}^{\mathbf{n}}$, say $\vec{u} = (u_1, u_2, \dots, u_n)$ and $\vec{v} = (v_1, v_2, \dots, v_n)$ then projection of vector $\vec{u}$ onto (along) a non – zero vector $\vec{v}$ is given as follows;

$\boldsymbol{proj}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) = \boldsymbol{proj}_{\vec{v}}\vec{\boldsymbol{u}} = \frac{\vec{\boldsymbol{u}}.\vec{\boldsymbol{v}}}{\|\vec{\boldsymbol{v}}\|^2}\vec{\boldsymbol{v}} = \frac{\vec{\boldsymbol{u}}.\vec{\boldsymbol{v}}}{\vec{\boldsymbol{v}}.\vec{\boldsymbol{v}}}\vec{\boldsymbol{v}}$ (vector component of $\vec{u}$ onto $\vec{v}$)

This is also called orthogonal projection of $\vec{u}$ on $\vec{v}$

Projection of vectors orthogonal to another vector:

Consider two non – zero vectors **u** and **v** in $\mathbf{R}^{\mathbf{n}}$, say $\vec{u} = (u_1, u_2, \dots, u_n)$ and $\vec{v} = (v_1, v_2, \dots, v_n)$ then projection of vector $\vec{u}$ orthogonal to a non – zero vector $\vec{v}$ is given as follows;

$\vec{\boldsymbol{u}} - \boldsymbol{proj}_{\vec{v}}\vec{\boldsymbol{u}} = \vec{\boldsymbol{u}} - \frac{\vec{\boldsymbol{u}}.\vec{\boldsymbol{v}}}{\|\vec{\boldsymbol{v}}\|^2}\vec{\boldsymbol{v}}$ (vector component of $\vec{u}$ orthogonal to $\vec{v}$)

Example:

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

Solution: We have $\vec{u} = (2, -1, 3)$ and $\vec{v} = (4, -1, 2)$

$\vec{u}.\vec{v} = (2, -1, 3).(4, -1, 2) = (2)(4) + (-1)(-1) + (3)(2) = 15$

$\|\vec{v}\|^2 = (4)^2 + (-1)^2 + (2)^2 = 21$

Then the vector component of $\vec{u}$ along (onto) $\vec{v}$ is as follows;

$$proj_{\vec{v}}\vec{u} = \frac{\vec{u}.\vec{v}}{\|\vec{v}\|^2}\vec{v} = \frac{15}{21}(4, -1, 2) = \frac{5}{7}(4, -1, 2) = \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right)$$

Also the vector component of $\vec{u}$ orthogonal $\vec{v}$ is as follows;

$$\vec{u} - proj_{\vec{v}}\vec{u} = \vec{u} - \frac{\vec{u}.\vec{v}}{\|\vec{v}\|^2}\vec{v} = (2, -1, 3) - \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right) = \left(-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7}\right)$$

PRACTICE:

1) Find $\|proj_{\vec{a}}\vec{u}\|$ for given vectors.

   i. $\vec{u} = (1,-2)$ and $\vec{a} = (-4,-3)$
   ii. $\vec{u} = (3,0,4)$ and $\vec{a} = (2,3,3)$
   iii. $\vec{u} = (5,6)$ and $\vec{a} = (2,-1)$
   iv. $\vec{u} = (3,-2,6)$ and $\vec{a} = (1,2,-7)$

2) Find the vector component of $\vec{u}$ along $\vec{a}$ and the vector component of $\vec{u}$ orthogonal to $\vec{a}$.
   i. $\vec{u} = (6,2)$ and $\vec{a} = (3,-9)$
   ii. $\vec{u} = (-1,-2)$ and $\vec{a} = (-2,3)$
   iii. $\vec{u} = (2,0,1)$ and $\vec{a} = (1,2,3)$
   iv. $\vec{u} = (2,1,1,2)$ and $\vec{a} = (4,-4,2,-2)$
   v. $\vec{u} = (5,0,-3,7)$ and $\vec{a} = (2,1,-1,-1)$

Field: A non-empty set **F** is called a field if

- **F** is Abelian group under addition.
- **F – {0}** is Abelian group under multiplication.
- Distributive law holds in **F**.

Note: Elements of a Field called Scalars.

Inner Product Spaces

Let **V** be a real vector space. Suppose to each pair of vectors $\vec{u},\vec{v}\epsilon\boldsymbol{V}$ there is assigned a real number, denoted by $\langle\vec{u},\vec{v}\rangle$. This function is called a (real) inner product on **V** if it satisfies the following axioms:

i. **(Linear Property):** $\langle a\vec{u}_1+b\vec{u}_2,\vec{v}\rangle = a\langle\vec{u}_1,\vec{v}\rangle + b\langle\vec{u}_2,\vec{v}\rangle$.
ii. **(Symmetric Property):** $\langle\vec{u},\vec{v}\rangle = \langle\vec{v},\vec{u}\rangle$.
iii. **(Positive Definite Property):** $\langle\vec{u},\vec{u}\rangle \geq 0$; and $\langle\vec{u},\vec{u}\rangle = 0$ if and only if $\vec{u} = 0$.

The vector space **V** with an inner product is called **a (real) inner product space**.

Axiom (i) states that an inner product function is linear in the first position. Using (i) and the symmetry axiom (ii) , we obtain

$$\langle\vec{u},c\vec{v}_1+d\vec{v}_2\rangle = \langle c\vec{v}_1+d\vec{v}_2,\vec{u}\rangle = c\langle\vec{v}_1,\vec{u}\rangle + d\langle\vec{v}_2,\vec{u}\rangle = c\langle\vec{u},\vec{v}_1\rangle + d\langle\vec{u},\vec{v}_2\rangle$$

That is, the inner product function is also linear in its second position. Combining these two properties and using induction yields the following general formula:

$$\langle\sum_i a_i\vec{u}_i\,,\sum_j b_j\vec{v}_j\rangle = \sum_i\sum_j a_ib_j\langle\vec{u}_i,\vec{v}_j\rangle$$

That is, an inner product of linear combinations of vectors is equal to a linear combination of the inner products of the vectors.

We may define above axioms as follows

i. **(Additivity Axiom):** $\langle\vec{u}_1+\vec{u}_2,\vec{v}\rangle = \langle\vec{u}_1,\vec{v}\rangle + \langle\vec{u}_2,\vec{v}\rangle$.
ii. **(Symmetry Axiom):** $\langle\vec{u},\vec{v}\rangle = \langle\vec{v},\vec{u}\rangle$.
iii. **(Homogeneity Axiom):** $\langle k\vec{u},\vec{v}\rangle = k\langle\vec{u},\vec{v}\rangle$.
iv. **(Positivity Axiom):** $\langle\vec{u},\vec{u}\rangle \geq 0$; and $\langle\vec{u},\vec{u}\rangle = 0$ if and only if $\vec{u} = 0$.

Because the axiom for a real inner product space are based on properties of the dot product, these inner product space axioms will be satisfied automatically if we define the inner product of two vectors **u** and **v** in $\mathbf{R^n}$ to be

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}.\vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

This inner product is commonly called the **Euclidean Inner Product** (or the **Standard Inner Product**) on $\mathbf{R^n}$ to distinguish it from other possible inner products that might be defined on $\mathbf{R^n}$. We call $\mathbf{R^n}$ with the Euclidean Inner Product **Euclidean – n space**.

Example:

Let **V** be a real inner product space. Then, by linearity,

$$\langle 3\boldsymbol{u_1} - 4\boldsymbol{u_2}, 2\boldsymbol{v_1} - 5\boldsymbol{v_2} + 6\boldsymbol{v_3} \rangle = 6\langle \boldsymbol{u_1}, \boldsymbol{v_1} \rangle - 15\langle \boldsymbol{u_1}, \boldsymbol{v_2} \rangle + 18\langle \boldsymbol{u_1}, \boldsymbol{v_3} \rangle - 8\langle \boldsymbol{u_2}, \boldsymbol{v_1} \rangle + 20\langle \boldsymbol{u_2}, \boldsymbol{v_2} \rangle - 24\langle \boldsymbol{u_2}, \boldsymbol{v_3} \rangle$$

$$\langle 2\boldsymbol{u} - 5\boldsymbol{v}, 4\boldsymbol{u} + 6\boldsymbol{v} \rangle = 8\langle \boldsymbol{u}, \boldsymbol{u} \rangle + 12\langle \boldsymbol{u}, \boldsymbol{v} \rangle - 20\langle \boldsymbol{v}, \boldsymbol{u} \rangle - 30\langle \boldsymbol{v}, \boldsymbol{v} \rangle$$

$$= 8\langle \boldsymbol{u}, \boldsymbol{u} \rangle - 8\langle \boldsymbol{v}, \boldsymbol{u} \rangle - 30\langle \boldsymbol{v}, \boldsymbol{v} \rangle$$

Observe that in the last equation we have used the symmetry property that $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$.

Norm (length) of a Vector

By the axiom $\langle \vec{u}, \vec{u} \rangle \geq 0$; and $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if $\vec{u} = 0.$, $\langle \vec{u}, \vec{u} \rangle$ is nonnegative for any vector $\vec{u}$. Thus, its positive square root exists. We use the notation $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}$ and This nonnegative number is called the *norm* or *length* of $\vec{u}$. The relation $\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle$ will be used frequently. Also remember a vector of norm 1 is called *unit vector*.

Distance between two Vectors

For this we use the notation $d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{\langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle}$

Examples of Inner Product Spaces

This section lists the main examples of inner product spaces used in this text.

Although the Euclidean inner product is the most important inner product on $\mathbf{R}^n$. However, there are various applications in which it is desirable to modify the Euclidean inner product by weighting its terms differently. More precisely, if $w_1, w_2, \dots, w_n$ are positive real numbers, which we shall call ***weights***, and if $\vec{u} = (u_1, u_2, \dots, u_n)$ and $\vec{v} = (v_1, v_2, \dots, v_n)$ are vectors in $\mathbf{R}^n$, then it can be shown that the formula $\langle \vec{u}, \vec{v} \rangle = w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n$ defines an inner product on $\mathbf{R}^n$; it is called the ***weighted Euclidean inner product*** with weights $w_1, w_2, \dots, w_n$.

Example: Weighted Euclidean Inner Product

Let $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$ be vectors in $\mathbf{R}^2$. Verify that the weighted Euclidean inner product $\langle \vec{u}, \vec{v} \rangle = 3u_1v_1 + 2u_2v_2$ satisfies the four inner product axioms.

Solution

Axiom 1: If $\vec{u}$ and $\vec{v}$ are interchanged in this equation, the right side remains the same. Therefore, $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$

Axiom 2: If $\vec{w} = (w_1, w_2)$, then

$$\langle \vec{u} + \vec{v}, \vec{w} \rangle = 3(u_1 + v_1)w_1 + 2(u_2 + v_2)w_2$$

$$\langle \vec{u} + \vec{v}, \vec{w} \rangle = 3(u_1w_1 + v_1w_1) + 2(u_2w_2 + v_2w_2)$$

$$\langle \vec{u} + \vec{v}, \vec{w} \rangle = (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2)$$

$$\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$$

Axiom 3: $\langle k\vec{u}, \vec{v} \rangle = 3(ku_1)v_1 + 2(ku_2)v_2 = k(3u_1v_1 + 2u_2v_2) = k\langle \vec{u}, \vec{v} \rangle$

Axiom 4: $\langle \vec{v}, \vec{v} \rangle = 3(v_1v_1) + 2(v_2v_2) = 3v_1^2 + 2v_2^2 \geq 0$; and $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = 0$

Example: Using a Weighted Euclidean Inner Product

It is important to keep in mind that norm and distance depend on the inner product being used. If the inner product is changed, then the norms and distances between vectors also change. For example, for the vectors $\vec{u} = (1,0)$ and $\vec{v} = (0,1)$ in $\mathbf{R}^2$ with the Euclidean inner product, we have $\|\vec{u}\| = \sqrt{1^2 + 0^2} = 1$

and $d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \|(1, -1)\| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$ However, if we change to the weighted Euclidean inner product $\langle \vec{u}, \vec{v} \rangle = 3u_1v_1 + 2u_2v_2$ then we obtain $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle} = \sqrt{3(1)(1) + 2(0)(0)} = \sqrt{3}$

and $d(\vec{u}, \vec{v}) = \|\vec{u}\|\|\vec{u} - \vec{v}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle} = \sqrt{\langle (1, -1), (1, -1) \rangle}$

$d(\vec{u}, \vec{v}) = \sqrt{3(1)(1) + 2(-1)(-1)} = \sqrt{5}$

Practice:

1. Let $\mathbf{R}^2$ have the weighted Euclidean inner product $\langle \vec{u}, \vec{v} \rangle = 2u_1v_1 + 3u_2v_2$ and let $\vec{u} = (1,1)$, $\vec{v} = (3,2)$, $\vec{w} = (0, -1)$ and $k = 3$ then compute the stated quantities;
   a) $\langle \vec{u}, \vec{v} \rangle$
   b) $\langle k\vec{v}, \vec{w} \rangle$
   c) $\langle \vec{u} + \vec{v}, \vec{w} \rangle$
   d) $\|\vec{v}\|$
   e) $d(\vec{u}, \vec{v})$
   f) $\|\vec{u} - k\vec{v}\|$

2. Let $\mathbf{R}^2$ have the weighted Euclidean inner product $\langle \vec{u}, \vec{v} \rangle = \frac{1}{2}u_1v_1 + 5u_2v_2$ and let $\vec{u} = (1,1)$, $\vec{v} = (3,2)$, $\vec{w} = (0, -1)$ and $k = 3$ then compute the stated quantities;
   a) $\langle \vec{u}, \vec{v} \rangle$
   b) $\langle k\vec{v}, \vec{w} \rangle$
   c) $\langle \vec{u} + \vec{v}, \vec{w} \rangle$
   d) $\|\vec{v}\|$
   e) $d(\vec{u}, \vec{v})$
   f) $\|\vec{u} - k\vec{v}\|$

Euclidean n-Space $\mathbf{R^n}$

Consider the vector space $\mathbf{R^n}$. The dot product or scalar product in $\mathbf{R^n}$ is defined by

$\vec{u}.\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$. This function defines an inner product on $\mathbf{R^n}$. The norm $\|\vec{u}\|$ of the vector $\vec{u}$ in this space is as follows:

$$\|\vec{u}\| = \sqrt{\vec{u}.\vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

On the other hand, by the Pythagorean theorem, the distance from the origin O in $\mathbf{R^3}$ to a point $P(u_1, u_2, u_3)$ is given by $\sqrt{u_1^2 + u_2^2 + u_3^2}$. This is precisely the same as the above-defined norm of the vector $\vec{u} = (u_1, u_2, u_3)$ in $\mathbf{R^3}$. Because the Pythagorean Theorem is a consequence of the axioms of Euclidean geometry, the vector space $\mathbf{R^n}$ with the above inner product and norm is called ***Euclidean n-space***. Although there are many ways to define an inner product on $\mathbf{R^n}$, we shall assume this inner product unless otherwise stated or implied. It is called ***the usual (or standard) inner product*** on $\mathbf{R^n}$.

Remark: Frequently the vectors in $\mathbf{R^n}$ will be represented by column vectors—that is, by $n \times 1$ column matrices. In such a case, the formula $\langle \vec{u}, \vec{v} \rangle = \vec{u}^T \vec{v}$ defines the usual inner product on $\mathbf{R^n}$.

Example: Let $\vec{u} = (1,3,-4,2), \vec{v} = (4,-2,2,1), \vec{w} = (5,-1,-2,6)$ in $\mathbf{R^4}$.

(a) Show $\langle 3\vec{u} - 2\vec{v}, \vec{w} \rangle = 3\langle \vec{u}, \vec{w} \rangle - 2\langle \vec{v}, \vec{w} \rangle$

By definition, $\langle \vec{u}, \vec{w} \rangle = 5 - 3 + 8 + 12 = 22$ and $\langle \vec{v}, \vec{w} \rangle = 20 + 2 - 4 + 6 = 24$

As $3\vec{u} - 2\vec{v} = (-5,13,-16,4)$ Thus, $\langle 3\vec{u} - 2\vec{v}, \vec{w} \rangle = -25 - 13 + 32 + 24 = 18$

Then $3\langle \vec{u}, \vec{w} \rangle - 2\langle \vec{v}, \vec{w} \rangle = 3(22) - 2(24) = 18 = \langle 3\vec{u} - 2\vec{v}, \vec{w} \rangle$

(b) Normalize $\vec{u}$ and $\vec{v}$:

Since $\|\vec{u}\| = \sqrt{1 + 9 + 16 + 4} = \sqrt{30}$ and $\|\vec{v}\| = \sqrt{16 + 4 + 4 + 1} = 5$

We normalize $\vec{u}$ and $\vec{v}$ to obtain the following unit vectors in the directions of $\vec{u}$ and $\vec{v}$, respectively:

$$\hat{u} = \frac{\vec{u}}{\|\vec{u}\|} = \left(\frac{1}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{-4}{\sqrt{30}}, \frac{2}{\sqrt{30}}\right) \text{ and } \hat{v} = \frac{\vec{v}}{\|\vec{v}\|} = \left(\frac{4}{5}, \frac{-2}{5}, \frac{2}{5}, \frac{1}{5}\right)$$

Unit Circles and Spheres in Inner Product Spaces

If **V** is an inner product space, then the set of points in **V** that satisfy $\|\vec{v}\| = 1$ is called the ***unit sphere*** or sometimes the ***unit circle*** in **V**. In $\mathbf{R}^2$ and $\mathbf{R}^3$ these are the points that lie 1 unit away from the origin.

Example: Unusual Unit Circles in

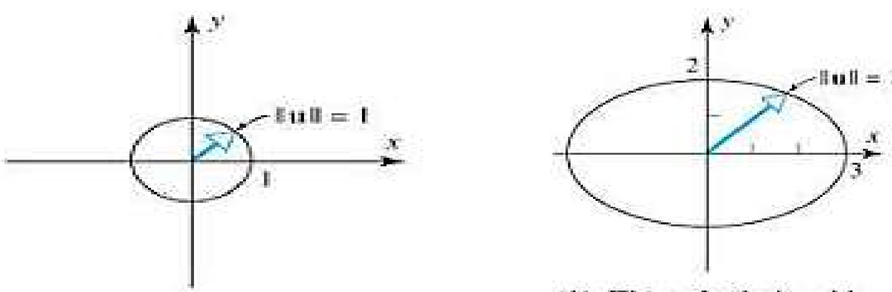
(a) Sketch the unit circle in an xy −coordinate system in $\mathbf{R}^2$ sing the Euclidean inner product $\langle\vec{u}, \vec{v}\rangle = u_1v_1 + u_2v_2$.

(b) Sketch the unit circle in an xy −coordinate system in $\mathbf{R}^2$ using the weighted Euclidean inner product $\langle\vec{u}, \vec{v}\rangle = \frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2$

Solution (a)

If $\vec{u} = (x, y)$, then $\|\vec{u}\| = \sqrt{x^2 + y^2}$, so the equation of the unit circle is $\sqrt{x^2 + y^2} = 1$, or, on squaring both sides, $x^2 + y^2 = 1$

As expected, the graph of this equation is a circle of radius 1 centered at the origin



(a) The unit circle with the Euclidian norm $\|\mathbf{u}\| = \sqrt{x^2 + y^2}$

(b) The unit circle with the norm $\|\mathbf{u}\| = \sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2}$

Solution (b)

If $\vec{u} = (x, y)$, then $\|\vec{u}\| = \sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2}$, so the equation of the unit circle is

$\sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2} = 1$, or, on squaring both sides, $\frac{1}{9}x^2 + \frac{1}{4}y^2 = 1$

Practice:

1. Compute the quantities using the inner product on $\mathbf{R}^2$ generated by
   $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$
   a) $\langle \vec{u}, \vec{v} \rangle$
   b) $\langle k\vec{v}, \vec{w} \rangle$
   c) $\langle \vec{u} + \vec{v}, \vec{w} \rangle$
   d) $\|\vec{v}\|$
   e) $d(\vec{u}, \vec{v})$
   f) $\|\vec{u} - k\vec{v}\|$
2. Find $\|\vec{u}\|$ and $d(\vec{u}, \vec{v})$ relative to the weighted Euclidian inner product $\langle \vec{u}, \vec{v} \rangle = 2u_1v_1 + 3u_2v_2$ on $\mathbf{R}^2$
   a) $\vec{u} = (-3,2)$ and $\vec{v} = (1,7)$
   b) $\vec{u} = (-1,2)$ and $\vec{v} = (2,5)$
3. Sketch the unit circle in $\mathbf{R}^2$ using the given inner products
   a) $\langle \vec{u}, \vec{v} \rangle = \frac{1}{4}u_1v_1 + \frac{1}{16}u_2v_2$
   b) $\langle \vec{u}, \vec{v} \rangle = 2u_1v_1 + u_2v_2$

The Standard Inner Product on $\boldsymbol{P_n}$

If $\boldsymbol{p} = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ and $\boldsymbol{q} = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n$ are polynomials in $\boldsymbol{P_n}$ then the following formula defines an inner product on $\boldsymbol{P_n}$ that we call the ***standard inner product*** on $\boldsymbol{P_n}$

$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$

The norm of polynomial $\boldsymbol{p}$ relative to this inner product is

$\|\boldsymbol{p}\| = \sqrt{\langle \boldsymbol{p}, \boldsymbol{q} \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$

The Evaluation Inner Product on $\boldsymbol{P_n}$

If $\boldsymbol{p} = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ and $\boldsymbol{q} = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n$ are polynomials in $\boldsymbol{P_n}$ and if $x_0, x_1, \dots, x_n$ are distinct real numbers (called ***sample point***) then the formula $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n)$ defines an inner product on $\boldsymbol{P_n}$ called the ***evaluation inner product*** at $x_0, x_1, \dots, x_n$ Algebraically this can be viewed as the dot product in $\mathbf{R^n}$ of the n – tuples $\big(p(x_0), p(x_1), \dots, p(x_n)\big)$ and $\big(q(x_0), q(x_1), \dots, q(x_n)\big)$ and hence the first three inner product axioms follow from properties of the product.

The fourth inner product axiom follows from the fact that

$\langle \boldsymbol{p}, \boldsymbol{p} \rangle = [p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2 \geq 0$

With equality holding if and only if $p(x_0) = p(x_1) = \cdots + p(x_n) = 0$

But a non – zero polynomial of degree ‘n’ or less can have at most ‘n’ distinct roots, so it must be that $\boldsymbol{p} = \boldsymbol{0}$ which proves that the fourth inner product axiom holds.

The norm of a polynomial $\boldsymbol{p}$ relative to the evaluation inner product is

$\|\boldsymbol{p}\| = \sqrt{\langle \boldsymbol{p}, \boldsymbol{p} \rangle} = \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2}$

Example

Let P_2 have the evaluation inner product at the points $x_0 = -2, x_1 = 0, x_2 = 2$ then compute $\langle \boldsymbol{p}, \boldsymbol{q} \rangle$ and $\|\boldsymbol{p}\|$ for the polynomials

$\boldsymbol{p} = p(x) = x^2$ and $\boldsymbol{q} = q(x) = 1 + x$

Solution

$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n)$

$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = p(-2)q(-2) + p(0)q(0) + p(2)q(2) = 8$

$\|\boldsymbol{p}\| = \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2} = \sqrt{4^2 + 0^2 + 4^2} = 4\sqrt{2}$

Practice:

1. Find the standard inner product on P_2 of the given polynomials
   a) $\boldsymbol{p} = -2 + x + 3x^2$ and $\boldsymbol{q} = 4 - 7x^2$
   b) $\boldsymbol{p} = -5 + 2x + x^2$ and $\boldsymbol{q} = 3 + 2x - 4x^2$
2. In the following exercise, a sequence of a sample points is given. Use the evaluation inner product on P_3 at those sample points to find $\langle \boldsymbol{p}, \boldsymbol{q} \rangle$ for the polynomials $\boldsymbol{p} = x + x^3$ and $\boldsymbol{q} = 1 + x^2$
   a) $x_0 = -2, x_1 = -1, x_2 = 0, x_3 = 1$
   b) $x_0 = -1, x_1 = 0, x_2 = 1, x_3 = 2$
3. Find $\boldsymbol{d}(\boldsymbol{p}, \boldsymbol{q})$ and $\|\boldsymbol{p}\|$ relative to the evaluation inner product on P_2 at the stated sample points.
   a) $\boldsymbol{p} = -2 + x + 3x^2$ and $\boldsymbol{q} = 4 - 7x^2$
   b) $\boldsymbol{p} = -5 + 2x + x^2$ and $\boldsymbol{q} = 3 + 2x - 4x^2$
4. Find $\boldsymbol{d}(\boldsymbol{p}, \boldsymbol{q})$ and $\|\boldsymbol{p}\|$ relative to the evaluation inner product on P_3 at the stated sample points.
   a) $x_0 = -2, x_1 = -1, x_2 = 0, x_3 = 1$
   b) $x_0 = -1, x_1 = 0, x_2 = 1, x_3 = 2$

Function Space $\boldsymbol{C} = [\boldsymbol{a}, \boldsymbol{b}]$ and Polynomial Space $\boldsymbol{P}(\boldsymbol{x})$

The notation $C = [a, b]$ is used to denote the vector space of all continuous functions on the closed interval $[a, b]$ that is, where $a \leq x \leq b$. The following defines an inner product on $C = [a, b]$, where $f(x)$ and $g(x)$ are functions in $C = [a, b]$:

$\langle f, g \rangle = \int_a^b f(x)g(x)dx$ And it is called the *usual inner product* on $C = [a, b]$.

The vector space $P(x)$ of all polynomials is a subspace of $C = [a, b]$ for any interval $[a, b]$, and hence, the above is also an inner product on $P(x)$.

Example

Show that $\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \int_a^b f(x)g(x)dx$ defines an inner product on $\boldsymbol{C} = [\boldsymbol{a}, \boldsymbol{b}]$

Solution: **Axiom 1:** $\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle \boldsymbol{g}, \boldsymbol{f} \rangle$

Axiom 2: $\langle \boldsymbol{f} + \boldsymbol{g}, \boldsymbol{h} \rangle = \int_a^b \big(f(x) + h(x)\big)g(x)dx$

$\langle \boldsymbol{f} + \boldsymbol{g}, \boldsymbol{h} \rangle = \int_a^b f(x)h(x)dx + \int_a^b g(x)h(x)dx = \langle \boldsymbol{f}, \boldsymbol{h} \rangle + \langle \boldsymbol{g}, \boldsymbol{h} \rangle$

Axiom 3: $\langle k\boldsymbol{f}, \boldsymbol{g} \rangle = \int_a^b kf(x)g(x)dx = k\int_a^b f(x)g(x)dx = k\langle \boldsymbol{f}, \boldsymbol{g} \rangle$

Axiom 4: $\langle \boldsymbol{f}, \boldsymbol{f} \rangle = \int_a^b f(x)f(x)dx = \int_a^b f^2(x)dx \geq 0$

And $\langle \boldsymbol{f}, \boldsymbol{f} \rangle = 0 \Leftrightarrow \boldsymbol{f} = \boldsymbol{0}$

Norm of a vector in $\boldsymbol{C} = [\boldsymbol{a}, \boldsymbol{b}]$

If $C = [a, b]$ has the inner product $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ then the norm of a function $f = f(x)$ relative to this inner product is

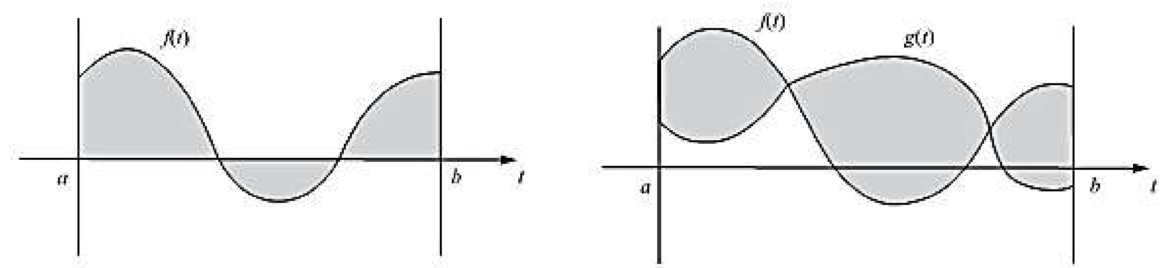
$$\|\boldsymbol{f}\|_2 = \|\boldsymbol{f}\| = \sqrt{\langle \boldsymbol{f}, \boldsymbol{f} \rangle} = \sqrt{\int_a^b f(x)f(x)dx} = \sqrt{\int_a^b f^2(x)dx}$$

And the unit sphere in this space consists of all functions $\boldsymbol{f}$ in $C = [a, b]$ that satisfy the equation $\int_a^b f^2(x)dx = 1$

Remember The following define the other norms on $C = [a, b]$:

- $\|\boldsymbol{f}\|_1 = \int_a^b |f(x)|dx$ = area between the function f and the t-axis
- $d_1(\boldsymbol{f}, \boldsymbol{g})$ =area between the functions f and **g**

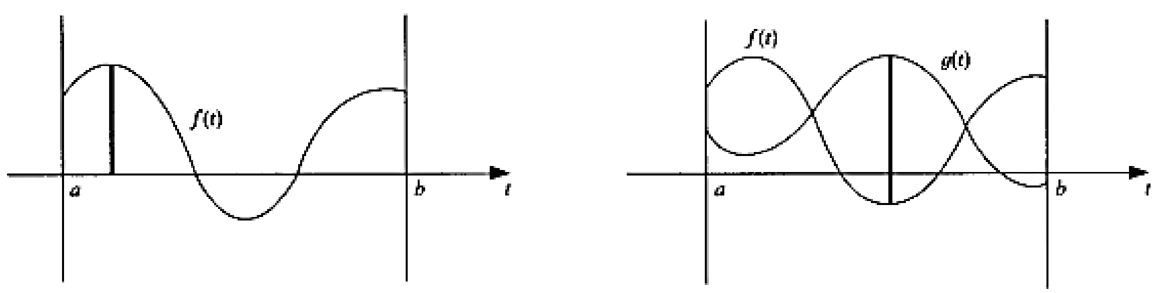
The geometrical descriptions of this norm and its corresponding distance function is described below.



(a) $\|f\|_1$ is shaded. (b) $d_1(f, g)$ is shaded.

- $\|\boldsymbol{f}\|_\infty = max(|f(x)|)$ = maximum distance between f and the t-axis
- $d_\infty(\boldsymbol{f}, \boldsymbol{g})$ = maximum distance between the functions f and **g**

The geometrical descriptions of this norm and its corresponding distance function is described below.



(a) $\|f\|_\infty$ (b) $d_\infty(f, g)$

Example:

Consider $f(x) = 3x - 5$ and $g(x) = x^2$ in the polynomial space $P(x)$ with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$ then find $\langle f, g\rangle$, $\|f\|$ and $\|g\|$

Solution: Using defined inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx = \int_0^1 (3x-5)(x^2)dx = \int_0^1 (3x^3 - 5x^2)dx$$

$$\langle f, g\rangle = \left|\frac{3}{4}x^4 - \frac{5}{3}x^3\right|_0^1 = -\frac{11}{12}$$

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 f(x)f(x)dx = \int_0^1 (9x^2 - 30x + 25)dx$$

$$\|f\|^2 = |3x^3 - 15x^2 + 25x|_0^1 = 13$$

$$\|g\|^2 = \langle g, g\rangle = \int_0^1 g(x)g(x)dx = \int_0^1 x^4 dx = \left|\frac{1}{5}x^5\right|_0^1 = \frac{1}{5}$$

Then clearly $\|f\| = \sqrt{13}$ and $\|g\| = \sqrt{\frac{1}{5}} = \frac{1}{5}\sqrt{5}$

Practice:

1. Let the vector space P_2 have the inner product $\langle p, q\rangle = \int_{-1}^1 p(x)q(x)dx$ then find the following for $\boldsymbol{p} = 1$ and $\boldsymbol{q} = x^2$
   a) $\langle \boldsymbol{p}, \boldsymbol{q}\rangle$
   b) $\|\boldsymbol{p}\|$
   c) $\|\boldsymbol{q}\|$
   d) $\boldsymbol{d}(\boldsymbol{p}, \boldsymbol{q})$
2. Let the vector space P_3 have the inner product $\langle p, q\rangle = \int_{-1}^1 p(x)q(x)dx$ then find the following for $\boldsymbol{p} = 2x^3$ and $\boldsymbol{q} = 1 - x^3$
   a) $\langle \boldsymbol{p}, \boldsymbol{q}\rangle$
   b) $\|\boldsymbol{p}\|$
   c) $\|\boldsymbol{q}\|$
   d) $\boldsymbol{d}(\boldsymbol{p}, \boldsymbol{q})$
3. Use the inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$ to compute $\langle \boldsymbol{f}, \boldsymbol{g}\rangle$
   a) $\boldsymbol{p} = Cos2\pi x$ and $\boldsymbol{q} = Sin2\pi x$
   b) $\boldsymbol{p} = x$ and $\boldsymbol{q} = e^x$

Matrix Space $M = Mm, n$

Let $M = Mm, n$, the vector space of all real $m \times n$ matrices. An inner product is defined on M by $\langle A, B \rangle = tr(B^T A)$ where, as usual, $tr(\quad)$ is the trace—the sum of the diagonal elements.

If $A = [a_{ij}]$ and $B = [b_{ij}]$, then $\langle A, B \rangle = tr(B^T A) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}$

And $\|A\|^2 = \langle A, A \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2$

That is, $\langle A, B \rangle$ is the sum of the products of the corresponding entries in A and B and, in particular, $\langle A, A \rangle$ is the sum of the squares of the entries of A.

Practice

1. Find a matrix that generate the stated weighted inner product on $\mathbf{R}^2$

$\langle \vec{u}, \vec{v} \rangle = 2u_1 v_1 + 3u_2 v_2$ and $\langle \vec{u}, \vec{v} \rangle = \frac{1}{2} u_1 v_1 + 5u_2 v_2$

2. Use the inner product on $\mathbf{R}^2$ generated by the matrix $A = \begin{bmatrix} 4 & 1 \\ 2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$ to find $\langle \vec{u}, \vec{v} \rangle$ for the vectors $\vec{u} = (0, -3)$ and $\vec{v} = (6,2)$

Hilbert Space

Let V be the vector space of all infinite sequences of real numbers $(a_1, a_2, a_3, \dots)$ satisfying $\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + a_3^2 + \cdots < \infty$

That is, the sum converges. Addition and scalar multiplication are defined in V component wise; that is, if $\boldsymbol{u} = (a_1, a_2, a_3, \dots)$ and $\boldsymbol{v} = (b_1, b_2, b_3, \dots)$

Then $\boldsymbol{u} + \boldsymbol{v} = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \dots)$and $k\boldsymbol{u} = (ka_1, ka_2, ka_3, \dots)$

An inner product is defined in V by $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = a_1 b_1, a_2 b_2, a_3 b_3, \dots$

The above sum converges absolutely for any pair of points in V. Hence, the inner product is well defined. This inner product space is called l_2 −space or Hilbert space.

Theorem (Algebraic Properties of inner products):

If $\vec{u}$ and $\vec{v}$ are vectors in a real inner product space **V** and if k is a scalar then

- **a)** $\langle \vec{0}, \vec{v} \rangle = \langle \vec{v}, \vec{0} \rangle = 0$
- **b)** $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$
- **c)** $\langle \vec{u}, \vec{v} - \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle - \langle \vec{u}, \vec{w} \rangle$
- **d)** $\langle \vec{u} - \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle - \langle \vec{v}, \vec{w} \rangle$
- **e)** $k\langle \vec{u}, \vec{v} \rangle = \langle \vec{u}, k\vec{v} \rangle$

Example: Evaluate $\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle$

Solution

$\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle = \langle \vec{u}, 3\vec{u} + 4\vec{v} \rangle - \langle 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle$

$\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle = \langle \vec{u}, 3\vec{u} \rangle + \langle \vec{u}, 4\vec{v} \rangle - \langle 2\vec{v}, 3\vec{u} \rangle - \langle 2\vec{v}, 4\vec{v} \rangle$

$\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle = 3\langle \vec{u}, \vec{u} \rangle + 4\langle \vec{u}, \vec{v} \rangle - 6\langle \vec{v}, \vec{u} \rangle - 8\langle \vec{v}, \vec{v} \rangle$

$\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle = 3\|\vec{u}\|^2 + 4\langle \vec{u}, \vec{v} \rangle - 6\langle \vec{u}, \vec{v} \rangle - 8\|\vec{v}\|^2$

$\langle \vec{u} - 2\vec{v}, 3\vec{u} + 4\vec{v} \rangle = 3\|\vec{u}\|^2 - 2\langle \vec{u}, \vec{v} \rangle - 8\|\vec{v}\|^2$

Practice

1. Suppose that $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors in an inner product space such that
   $\langle \vec{u}, \vec{v} \rangle = 2, \langle \vec{v}, \vec{w} \rangle = -6, \langle \vec{u}, \vec{w} \rangle = -3, \|\vec{u}\| = 1, \|\vec{v}\| = 2, \|\vec{w}\| = 7$
   Then evaluate the given expressions
   a) $\langle 2\vec{v} - \vec{w}, 3\vec{u} + 2\vec{w} \rangle$
   b) $\|\vec{u} + \vec{v}\|$
   c) $\langle \vec{u} - \vec{v} - 2\vec{w}, 4\vec{u} + \vec{v} \rangle$
   d) $\|2\vec{w} - \vec{v}\|$
2. Expand the followings;
   a) $\langle 5\vec{u}_1 + 8\vec{u}_2, 6\vec{v}_1 - 7\vec{v}_2 \rangle$
   b) $\langle 3\vec{u} + 5\vec{v}, 4\vec{u} - 6\vec{v} \rangle$
   c) $\|2\vec{u} - 3\vec{v}\|^2$

Cauchy–Schwarz Inequality

If $\vec{u}$ and $\vec{v}$ are vectors in a real inner product space **V**, then

$\langle\vec{u},\vec{v}\rangle^2 \leq \langle\vec{u},\vec{u}\rangle^2\langle\vec{v},\vec{v}\rangle^2$ or $|\langle\vec{u},\vec{v}\rangle| \leq \|\vec{u}\|\|\vec{v}\|$

Proof

For any real number 't' consider $\langle t\vec{u}-\vec{v}, t\vec{u}-\vec{v}\rangle \geq 0$

$= t\langle\vec{u}, t\vec{u}-\vec{v}\rangle - \langle\vec{v}, t\vec{u}-\vec{v}\rangle = t\langle t\vec{u}-\vec{v},\vec{u}\rangle - \langle t\vec{u}-\vec{v},\vec{v}\rangle \geq 0$

$= t\{t\langle\vec{u},\vec{u}\rangle - 1\langle\vec{v},\vec{u}\rangle\} - \{t\langle\vec{u},\vec{v}\rangle - 1\langle\vec{v},\vec{v}\rangle\} \geq 0$

$= t^2\langle\vec{u},\vec{u}\rangle - t\langle\vec{v},\vec{u}\rangle - t\langle\vec{u},\vec{v}\rangle + 1\langle\vec{v},\vec{v}\rangle \geq 0$

$= t^2\|\vec{u}\|^2 - 2t\langle\vec{u},\vec{v}\rangle + \|\vec{v}\|^2 \geq 0$

Let $t = \frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|^2}$ then

$= \frac{\langle\vec{u},\vec{v}\rangle^2}{\|\vec{u}\|^4}\|\vec{u}\|^2 - 2\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{u}\|^2}\langle\vec{u},\vec{v}\rangle + \|\vec{v}\|^2 \geq 0$

$= \frac{\langle\vec{u},\vec{v}\rangle^2}{\|\vec{u}\|^2} - 2\frac{\langle\vec{u},\vec{v}\rangle^2}{\|\vec{u}\|^2} + \|\vec{v}\|^2 \geq 0$

$\Rightarrow -\frac{\langle\vec{u},\vec{v}\rangle^2}{\|\vec{u}\|^2} + \|\vec{v}\|^2 \geq 0$

$\Rightarrow \|\vec{v}\|^2 \geq \frac{\langle\vec{u},\vec{v}\rangle^2}{\|\vec{u}\|^2}$

$\Rightarrow \|\vec{u}\|^2\|\vec{v}\|^2 \geq \langle\vec{u},\vec{v}\rangle^2$

$\Rightarrow \langle\vec{u},\vec{v}\rangle^2 \leq \|\vec{u}\|^2\|\vec{v}\|^2$

$\Rightarrow |\langle\vec{u},\vec{v}\rangle| \leq \|\vec{u}\|\|\vec{v}\|$ or $\langle\vec{u},\vec{v}\rangle^2 \leq \langle\vec{u},\vec{u}\rangle^2\langle\vec{v},\vec{v}\rangle^2$

Theorem: Let V be an inner product space, V also a normed space if following axioms are true;

- $\|\vec{u}\| \geq 0$; $\|\vec{u}\| = 0 \Leftrightarrow \vec{u} = 0$
- $\|k\vec{u}\| = |k|\|\vec{u}\|$; $\forall k \in F$
- $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Proof

- $\|\boldsymbol{\vec{u}}\| \geq \mathbf{0}$; $\|\boldsymbol{\vec{u}}\| = \mathbf{0} \Leftrightarrow \boldsymbol{\vec{u}} = \mathbf{0}$

Let $\vec{u} \neq 0$ then $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle} > 0 \Rightarrow \|\vec{u}\| > 0$

If $\vec{u} = 0$ then $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle} = 0 \Rightarrow \|\vec{u}\| = 0$

$\Rightarrow \|\vec{u}\| \geq 0$; $\|\vec{u}\| = 0 \Leftrightarrow \vec{u} = 0$

- $\|\boldsymbol{k\vec{u}}\| = |\boldsymbol{k}|\|\boldsymbol{\vec{u}}\|$; $\forall \boldsymbol{k} \in \boldsymbol{F}$

Let $\|k\vec{u}\|^2 = \langle k\vec{u}, k\vec{u} \rangle = kk\langle \vec{u}, \vec{u} \rangle = k^2\|\vec{u}\|^2 \Rightarrow \|k\vec{u}\| = |k|\|\vec{u}\|$; $\forall k \in F$

- $\|\boldsymbol{\vec{u}} + \boldsymbol{\vec{v}}\| \leq \|\boldsymbol{\vec{u}}\| + \|\boldsymbol{\vec{v}}\|$

$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u} + \vec{v}, \vec{u} + \vec{v} \rangle = \langle \vec{u}, \vec{u} + \vec{v} \rangle + \langle \vec{v}, \vec{u} + \vec{v} \rangle$

$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle = \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle$

$\|\vec{u} + \vec{v}\|^2 \leq \langle \vec{u}, \vec{u} \rangle + 2|\langle \vec{u}, \vec{v} \rangle| + \langle \vec{v}, \vec{v} \rangle = \|\vec{u}\|^2 + 2\|\vec{u}\|\|\vec{v}\| + \|\vec{v}\|^2$

$\|\vec{u} + \vec{v}\|^2 \leq (\|\vec{u}\| + \|\vec{v}\|)^2 \Rightarrow \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Theorem of Pythagoras: If $\vec{u}$ and $\vec{v}$ are orthogonal vectors in a real inner product space then $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$

Proof: Since $\vec{u}$ and $\vec{v}$ are orthogonal therefore $\langle \vec{u}, \vec{v} \rangle = 0$

$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u} + \vec{v}, \vec{u} + \vec{v} \rangle = \langle \vec{u}, \vec{u} + \vec{v} \rangle + \langle \vec{v}, \vec{u} + \vec{v} \rangle$

$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle = \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle$

$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u}, \vec{u} \rangle + 2\langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle = \|\vec{u}\|^2 + \|\vec{v}\|^2 \qquad \therefore \vec{u}.\vec{v} = 0$

Inner Product Space Satisfies Parallelogram Equality

$\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = 2(\|\vec{u}\|^2 + \|\vec{v}\|^2)$

Proof: $L.H.S = \|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2$

$= \langle \vec{u}+\vec{v}, \vec{u}+\vec{v}\rangle + \langle \vec{u}-\vec{v}, \vec{u}-\vec{v}\rangle$

$= \langle \vec{u}, \vec{u}+\vec{v}\rangle + \langle \vec{v}, \vec{u}+\vec{v}\rangle + \langle \vec{u}, \vec{u}-\vec{v}\rangle - \langle \vec{v}, \vec{u}-\vec{v}\rangle$

$= \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{u}\rangle + \langle \vec{u}, \vec{u}\rangle - \langle \vec{u}, \vec{v}\rangle - \langle \vec{v}, \vec{u}\rangle + \langle \vec{v}, \vec{v}\rangle$

$= \|\vec{u}\|^2 + 2\langle \vec{u}, \vec{v}\rangle + \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\langle \vec{u}, \vec{v}\rangle + \|\vec{v}\|^2$

$= 2(\|\vec{u}\|^2 + \|\vec{v}\|^2) = R.H.S$

Polarization Inequality $\langle \vec{u}, \vec{v}\rangle = \frac{1}{4}(\|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2)$

Proof: **Consider** $\|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2$

$= \langle \vec{u}+\vec{v}, \vec{u}+\vec{v}\rangle - \langle \vec{u}-\vec{v}, \vec{u}-\vec{v}\rangle$

$= \langle \vec{u}, \vec{u}+\vec{v}\rangle + \langle \vec{v}, \vec{u}+\vec{v}\rangle - \langle \vec{u}, \vec{u}-\vec{v}\rangle + \langle \vec{v}, \vec{u}-\vec{v}\rangle$

$= \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{u}\rangle - \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{u}\rangle - \langle \vec{v}, \vec{v}\rangle$

$= \|\vec{u}\|^2 + 2\langle \vec{u}, \vec{v}\rangle + \|\vec{v}\|^2 - \|\vec{u}\|^2 + 2\langle \vec{u}, \vec{v}\rangle - \|\vec{v}\|^2 = 4\langle \vec{u}, \vec{v}\rangle$

$\langle \vec{u}, \vec{v}\rangle = \frac{1}{4}(\|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2)$

Triangular Inequality for Vectors

$\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Proof

$\|\vec{u}+\vec{v}\|^2 = \langle \vec{u}+\vec{v}, \vec{u}+\vec{v}\rangle = \langle \vec{u}, \vec{u}+\vec{v}\rangle + \langle \vec{v}, \vec{u}+\vec{v}\rangle$

$\|\vec{u}+\vec{v}\|^2 = \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{u}\rangle + \langle \vec{v}, \vec{v}\rangle = \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{v}\rangle$

$\|\vec{u}+\vec{v}\|^2 \leq \langle \vec{u}, \vec{u}\rangle + 2|\langle \vec{u}, \vec{v}\rangle| + \langle \vec{v}, \vec{v}\rangle = \|\vec{u}\|^2 + 2\|\vec{u}\|\|\vec{v}\| + \|\vec{v}\|^2$

$\|\vec{u}+\vec{v}\|^2 \leq (\|\vec{u}\| + \|\vec{v}\|)^2 \Rightarrow \|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Appolonius Inequality

$$\|z-x\|^2+\|z-y\|^2=\frac{1}{2}\|x-y\|^2+2\left\|z-\frac{1}{2}(x+y)\right\|^2$$

Proof:

$$\|x+y\|^2=\langle x+y,x+y\rangle=\langle x,x+y\rangle+\langle y,x+y\rangle$$

$$\|x+y\|^2=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle$$

$$\|x+y\|^2=\|x\|^2+2\langle x,y\rangle+\|y\|^2 \qquad \ldots\ldots\ldots\ldots\ldots..(1)$$

$$\|x-y\|^2=\langle x-y,x-y\rangle=\langle x,x-y\rangle-\langle y,x-y\rangle$$

$$\|x-y\|^2=\langle x,x\rangle-\langle x,y\rangle-\langle y,x\rangle+\langle y,y\rangle$$

$$\|x-y\|^2=\|x\|^2-2\langle x,y\rangle+\|y\|^2 \qquad \ldots\ldots\ldots\ldots\ldots..(2)$$

Adding (1) and (2) we have

$$\|x+y\|^2+\|x-y\|^2=\|x\|^2+2\langle x,y\rangle+\|y\|^2+\|x\|^2-2\langle x,y\rangle+\|y\|^2$$

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2 \qquad \ldots\ldots\ldots\ldots\ldots..(3)$$

Put $x=z-x, y=z-y$ in (3)

$$\|z-x+z-y\|^2+\|z-x-z+y\|^2=2\|z-x\|^2+2\|z-y\|^2$$

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

$$4\left\|z-\frac{1}{2}(x+y)\right\|^2+\|-x+y\|^2=2\|z-x\|^2+2\|z-y\|^2$$

$$2\left\|z-\frac{1}{2}(x+y)\right\|^2+\frac{1}{2}\|x-y\|^2=\|z-x\|^2+\|z-y\|^2$$

Hence we get

$$\|z-x\|^2+\|z-y\|^2=\frac{1}{2}\|x-y\|^2+2\left\|z-\frac{1}{2}(x+y)\right\|^2$$

Examples

(a) Consider any real numbers $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$. Then, by the Cauchy–Schwarz inequality,
$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \leq (a_1^2 + a_2^2 + \cdots + a_n^2)(b_1^2 + b_2^2 + \cdots + b_n^2)$
$\Rightarrow (\vec{u}.\vec{v})^2 \leq \|\vec{u}\|^2\|\vec{v}\|^2$ where $\vec{u} = (a_i)$, $\vec{v} = (b_i)$

(b) Let f and g be continuous functions on the unit interval $[0,1]$. Then, by the Cauchy–Schwarz inequality,
$\left[\int_0^1 f(t)g(t)dt\right]^2 \leq \int_0^1 f^2(t)dt \int_0^1 g^2(t)dt$
$\Rightarrow (\langle f.g\rangle)^2 \leq \|f\|^2\|g\|^2$. Here V is the inner product space $C[0,1]$.

Angle between Vectors

For any nonzero vectors $\vec{u}$ and $\vec{v}$ in an inner product space **V**, the angle between $\vec{u}$ and $\vec{v}$ is defined to be the angle θ such that $0 \leq \theta \leq \pi$ and $\theta = Cos^{-1}\left(\frac{\langle\vec{u}.\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}\right)$ and cosine of the angle between vectors is defined as $Cos\theta = \frac{\langle\vec{u}.\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}$

By the Cauchy–Schwartz inequality, $-1 \leq Cos\theta \leq 1$, and so the angle exists and is unique.

Example

(a) Consider vectors $\vec{u} = (2,3,5)$ and $\vec{v} = (1,-4,3)$ in $\mathbf{R}^3$. Then
$\langle\vec{u}.\vec{v}\rangle = 2 - 12 + 15 = 5, \|\vec{u}\| = \sqrt{38}\,, \|\vec{v}\| = \sqrt{26}$
Then the angle θ between $\vec{u}$ and $\vec{v}$ is given by $\theta = Cos^{-1}\left(\frac{5}{\sqrt{38}\sqrt{26}}\right)$
Note that θ is an acute angle, because $Cos\theta$ is positive.

(b) Let $f(t) = 3t - 5$ and $g(t) = t^2$ in the polynomial space $\boldsymbol{P(t)}$ with inner product $\langle f.g\rangle = \int_0^1 f(t)g(t)dt$ then
$\langle f.g\rangle = -\frac{11}{12}$, $\|f\| = \sqrt{13}$, $\|g\| = \frac{1}{5}\sqrt{5}$
Then the angle θ between $\vec{u}$ and $\vec{v}$ is given by $\theta = Cos^{-1}\left(\frac{-\frac{11}{12}}{\sqrt{13}\left(\frac{1}{5}\sqrt{5}\right)}\right)$
i.e. $\theta = Cos^{-1}\left(-\frac{55}{12\sqrt{13}\sqrt{5}}\right)$
Note that θ is an obtuse angle, because $Cos\theta$ is negative.

(c) Consider vectors $\vec{u} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$ in $\mathbf{M_{22}}$. Then

$\langle \vec{u}.\vec{v} \rangle = 16, \|\vec{u}\| = \sqrt{30}\,, \|\vec{v}\| = \sqrt{14}$

Then the cosine of angle θ between $\vec{u}$ and $\vec{v}$ is given by $Cos\theta = \left(\frac{16}{\sqrt{30}\sqrt{14}}\right)$

i.e. $Cos\theta \approx 0.78$

Practice

1) Find the cosine of the angles between the vector spaces with respect to the Euclidean inner product.
   a) $\vec{u} = (1, -3)$ and $\vec{v} = (2,4)$
   b) $\vec{u} = (-1,5,2)$ and $\vec{v} = (2,4,-9)$
   c) $\vec{u} = (1,0,1,0)$ and $\vec{v} = (-3,-3,-3,-3)$
   d) $\vec{u} = (-1,0)$ and $\vec{v} = (3,8)$
   e) $\vec{u} = (4,1,8)$ and $\vec{v} = (1,0,-3)$
   f) $\vec{u} = (2,1,7,-1)$ and $\vec{v} = (4,0,0,0)$
2) Find the cosine of the angles between the vector spaces with respect to the Standard inner product on $\boldsymbol{P_2}$.
   a) $\boldsymbol{p} = -1 + 5x + 2x^2$ and $\boldsymbol{q} = 2 + 4x - 9x^2$
   b) $\boldsymbol{p} = x - x^2$ and $\boldsymbol{q} = 7 + 3x + 3x^2$
3) Find the cosine of the angles between the vector spaces with respect to the Standard inner product on $\boldsymbol{M_{22}}$.
   a) $\vec{u} = \begin{bmatrix} 2 & 6 \\ 1 & -3 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$
   b) $\vec{u} = \begin{bmatrix} 2 & 4 \\ -1 & 3 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -3 & 1 \\ 4 & 2 \end{bmatrix}$

Orthogonality

Let **V** be an inner product space. The vectors $\vec{u}, \vec{v} \in V$ are said to be orthogonal and $\vec{u}$ is said to be orthogonal to $\vec{v}$ if $\langle \vec{u}.\vec{v} \rangle = 0$

Orthogonality depends on inner Product

The vectors are orthogonal with respect to the Euclidean inner product on $\mathbf{R}^2$ for example for $\vec{u} = (1,1)$ and $\vec{v} = (1,-1)$ we have $\vec{u}.\vec{v} = 0$ but

The vectors are not orthogonal with respect to the weighted Euclidean inner product $\langle \vec{u}.\vec{v} \rangle = 3u_1v_1 + 2u_2v_2$ for example for $\vec{u} = (1,1)$ and $\vec{v} = (1,-1)$ we have $\langle \vec{u}.\vec{v} \rangle = 3(1)(1) + 2(1)(-1) = 1 \neq 0$

Orthogonal Vectors in M_{22}

Let $\vec{u} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ then $\langle \vec{u}.\vec{v} \rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$

Orthogonal Vectors in P_2

Let $\mathbf{P_2}$ have the inner product $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_{-1}^{1} p(x)q(x)dx$ and let $\boldsymbol{p} = x$ and $\boldsymbol{q} = x^2$ then $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_{-1}^{1} (x)(x^2)dx = \int_{-1}^{1} x^3 dx = 0$. Hence given vectors are orthogonal.

Example

(a) Consider the vectors $\vec{u} = (1,1,1), \vec{v} = (1,2,-3)$ and $\vec{w} = (1,-4,3)$ in $\mathbf{R}^3$.
Then $\langle \vec{u}, \vec{v} \rangle = 0, \langle \vec{u}, \vec{w} \rangle = 0, \langle \vec{v}, \vec{w} \rangle = -16 \neq 0$
Thus, $\vec{u}$ is orthogonal to $\vec{v}$ and $\vec{w}$, but $\vec{v}$ and $\vec{w}$ are not orthogonal.

(b) Consider the functions $sint$ and $cost$ in the vector space $C[-\pi, \pi]$ of continuous functions on the closed interval $[-\pi, \pi]$. Then
$\langle Sint, Cost \rangle = \int_{-\pi}^{\pi} SintCostdt = \left| \frac{1}{2} Sin^2 t \right|_{-\pi}^{\pi} = 0$
Thus, $sint$ and $cost$ are orthogonal functions in the vector space $C[-\pi, \pi]$.

Practice

1) Determine whether the vectors are orthogonal with respect to the Euclidean Inner Product.
   a) $\vec{u} = (-1,3,2), \vec{v} = (4,2,-1)$
   b) $\vec{u} = (-2,-2,-2), \vec{v} = (1,1,1)$
   c) $\vec{u} = (a,b), \vec{v} = (-b,a)$
   d) $\vec{u} = (u_1,u_2,u_3), \vec{v} = (0,0,0)$
   e) $\vec{u} = (-4,6,-10,1), \vec{v} = (2,1,-2,9)$
   f) $\vec{u} = (a,b,c), \vec{v} = (-c,0,a)$

2) Determine whether the vectors are orthogonal with respect to the Standard Inner Product on $\mathbf{P_2}$.
   a) $\boldsymbol{p} = -1 - x + 2x^2$ and $\boldsymbol{q} = 2x + x^2$
   b) $\boldsymbol{p} = 2 - 3x + x^2$ and $\boldsymbol{q} = 4 + 2x - 2x^2$

3) Determine whether the vectors are orthogonal with respect to the Standard Inner Product on $\mathbf{M_{22}}$.
   a) $U = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$ and $V = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$
   b) $U = \begin{bmatrix} 5 & -1 \\ 2 & -2 \end{bmatrix}$ and $V = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix}$

4) Show that the vectors are not orthogonal with respect to the Euclidean Inner Product on R^2 and then find a value of 'k' for which the vectors are orthogonal with respect to the weighted Euclidean inner product
   $\langle \vec{u}.\vec{v} \rangle = 2u_1v_1 + ku_2v_2$
   a) $\vec{u} = (1,3,), \vec{v} = (2,-1)$
   b) $\vec{u} = (2,-4), \vec{v} = (0,3)$

5) Find 'k' so that the vectors $\vec{u} = (1,2,k,3)$ and $\vec{v} = (3,k,7,-5)$ in $\mathbf{R^4}$ are orthogonal.

Remark

A vector $\vec{w} = (x_1, x_2, \dots, x_n)$ is orthogonal to $\vec{u} = (a_1, a_2, \dots, a_n)$ in $\mathbf{R^n}$ if $\langle \vec{u}, \vec{w} \rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$

That is, $\vec{w}$ is orthogonal to $\vec{u}$ if $\vec{w}$ satisfies a homogeneous equation whose coefficients are the elements of $\vec{u}$.

Example

Find a nonzero vector $\vec{w}$ that is orthogonal to $\vec{u} = (1,2,1)$ and $\vec{v} = (2,5,4)$ in $\mathbf{R^3}$.

Solution

Let $\vec{w} = (x, y, z)$. Then we want $\langle \vec{u}, \vec{w} \rangle = 0$ and $\langle \vec{v}, \vec{w} \rangle = 0$. This yields the homogeneous system

$$x + 2y + z = 0$$

$$2x + 5y + 4z = 0$$

Or

$$x + 2y + z = 0$$

$$y + 2z = 0$$

Here z is the only free variable in the echelon system. Set $z = 1$ to obtain $y = -2$ and $x = 3$. Thus, $\vec{w} = (3, -2,1)$ is a desired nonzero vector orthogonal to $\vec{u}$ and $\vec{v}$.

Orthogonal Complements

Let S be a subset of an inner product space **V**. The orthogonal complement of S, denoted by $S^{\perp}$ (read ‘‘S perp’’) consists of those vectors in **V** that are orthogonal to every vector $\vec{u} \in S$; that is, $S^{\perp} = \{\vec{v} \in V : \langle \vec{v}, \vec{u} \rangle = 0 \ \ for\ every\ \vec{u} \in S\}$

Or If S be a subspace of real inner product space **V**. then the set of all vectors in **V** that are orthogonal to every vector in S is called the orthogonal compliment of S. that is, $S^{\perp} = \{\vec{v} \in V : \langle \vec{v}, \vec{u} \rangle = 0 \ \ for\ every\ \vec{u} \in S\}$

Proposition Let S be a subset of a vector space **V**. Then $S^{\perp}$ is a subspace of **V**.

Proof

Choose $\vec{u}, \vec{v} \in S$ and $\propto, \beta \in F$ then we have to show that $\propto \vec{u} + \beta\vec{v} \in S$

Consider for $\vec{w} \in S$

$\langle \propto \vec{u} + \beta\vec{v}, \vec{w} \rangle = \propto \langle \vec{u}, \vec{w} \rangle + \beta\langle \vec{v}, \vec{w} \rangle = \propto (0) + \beta(0) = 0$

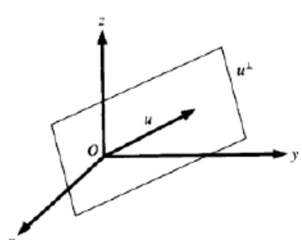
Implies $\propto \vec{u} + \beta\vec{v} \in S$ and hence $S^{\perp}$ is a subspace of **V**.

Annihilator Let Y be a subset of a Hilbert space H, then the set of all vectors of H which are orthogonal to Y is called the annihilator of Y and is denoted by $Y^{\perp}$. i.e. $Y^{\perp} = \{x \in H : x \perp Y\}$

Remember

- The annihilator of $Y^{\perp}$ and is denoted by $Y^{\perp\perp}$. i.e. $Y^{\perp\perp} = \{x \in H : x \perp Y^{\perp}\}$
- $\{0\}^{\perp} = \{x \in H : x \perp \{0\}\} = H$ and $H^{\perp} = \{x \in H : x \perp H\} = \{0\}$

Remark: Suppose $u^{\perp}$ is a nonzero vector in R^3. Then there is a geometrical description of $u^{\perp}$. Specifically, $u^{\perp}$ is the plane in R^3 through the origin O and perpendicular to the vector u. This is shown in Figure.



Example

Find a basis for the subspace $u^{\perp}$ of $\mathbf{R}^3$, where $u = (1,3,-4)$.

Solution

Consider $w = (x, y, z)$ and $u = (1,3,-4)$ then $\langle \vec{u}, \vec{w} \rangle = 0$

Implies $\langle \vec{u}, \vec{w} \rangle x + 3y - 4z = 0$ where The free variables are y and z.

(1) Set $y = 1, z = 0$ to obtain the solution $w_1 = (-3,1,0)$.

(2) Set $y = 0, z = 1$ to obtain the solution $w_2 = (4,0,1)$.

The vectors w_1 and w_2 form a basis for the solution space of the equation, and hence a basis for $u^{\perp}$.

Theorem

Let W be a subspace of V. Then V is the direct sum of W and $W^{\perp}$; that is,

$V = W \oplus W^{\perp}$

Proof

We know that every set of simple basis can be converted to orthogonal basis by Gram Schmidt Orthogonal process. Therefore we can say that there exists an orthogonal basis $\{u_1, u_2, \dots, u_r\}$ of W. therefore by Basis Extension Theorem process we can extend the orthogonal basis set to orthogonal basis set off vector space V. i.e. $\{u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_n\}$.

If $v \in V$ then $v = \sum_1^n \propto_i u_i = \propto_1 u_1 + \cdots + \propto_r u_r + \propto_{r+1} u_{r+1} + \cdots + \propto_n u_n$

Where $\propto_1 u_1 + \propto_2 u_2 + \cdots + \propto_r u_r \in W$

$\Rightarrow \propto_{r+1} u_{r+1} + \cdots + \propto_n u_n \in W^{\perp}$

$\Rightarrow \propto_1 u_1 + \cdots + \propto_r u_r + \propto_{r+1} u_{r+1} + \cdots + \propto_n u_n \in W + W^{\perp}$

$\Rightarrow V = W + W^{\perp}$

Now choose $x \in W \cap W^{\perp}$

$\Rightarrow x \in W$ and $x \in W^{\perp}$

$\Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0 \Rightarrow W \cap W^{\perp} = \{0\}$

$\Rightarrow V = W \oplus W^{\perp}$

Theorem

Let W be a subspace of V. Then show that $W \cap W^{\perp} = \{0\}$

Proof

Choose $x \in W \cap W^{\perp}$

$\Rightarrow x \in W$ and $x \in W^{\perp}$

$\Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0 \Rightarrow W \cap W^{\perp} = \{0\}$

Theorem

Let Y be a subset of a Hilbert space H. Then $Y \subseteq Y^{\perp\perp}$

Proof

Let $x \in Y$ then $\langle x, y \rangle = 0$ for all $y \in Y^{\perp}$

$\Rightarrow x \perp Y^{\perp} \Rightarrow x \in Y^{\perp\perp} \Rightarrow Y \subseteq Y^{\perp\perp}$

Theorem

Let A and B be subset of a Hilbert space H. And $A \subseteq B$ Then $B^{\perp} \subseteq A^{\perp}$

Proof

Let $A \subseteq B$ and $x \in B^{\perp}$ then $\langle x, y \rangle = 0$ for all $y \in B$

$\Rightarrow \langle x, y \rangle = 0$ for all $y \in A \Rightarrow x \in A^{\perp} \Rightarrow B^{\perp} \subseteq A^{\perp}$

Theorem

Let A and B be subset of a Hilbert space H. Then $(A \cup B)^{\perp} = A^{\perp} \cap B^{\perp}$.

Proof

Since $A \subseteq A \cup B$ and $B \subseteq A \cup B$ then $(A \cup B)^{\perp} \subseteq A^{\perp}$ and $(A \cup B)^{\perp} \subseteq B^{\perp}$

$\Rightarrow (A \cup B)^{\perp} \subseteq A^{\perp} \cap B^{\perp}$ ……………..(1)

Let $x \in A^{\perp} \cap B^{\perp}$ this means that $x \in A^{\perp}$ and $x \in B^{\perp}$ then by definition

$\langle x, u \rangle = 0$ for all $u \in A$ and $\langle x, v \rangle = 0$ for all $v \in B$

Hence $\langle x, v \rangle = 0$ for every $v \in A \cup B$ and so by definition $x \in (A \cup B)^{\perp}$

$\Rightarrow A^{\perp} \cap B^{\perp} \subseteq (A \cup B)^{\perp}$ ……………..(2)

From (1) and (2) $(A \cup B)^{\perp} = A^{\perp} \cap B^{\perp}$

Theorem

Let A and B be subset of a Hilbert space H. Then $A^{\perp} \cup B^{\perp} \subseteq (A \cap B)^{\perp}$.

Proof

Since $A \cap B \subseteq A$ and $A \cap B \subseteq B$ then $A^{\perp} \subseteq (A \cap B)^{\perp}$ and $B^{\perp} \subseteq (A \cap B)^{\perp}$

$\Rightarrow A^{\perp} \cup B^{\perp} \subseteq (A \cap B)^{\perp}$

Theorem

Let A be a subset of a Hilbert space H. Then $A^{\perp} = A^{\perp\perp\perp}$.

Proof

Since $A \subseteq A^{\perp\perp}$ then $(A^{\perp\perp})^{\perp} \subseteq A^{\perp} \Rightarrow A^{\perp\perp\perp} \subseteq A^{\perp}$

Also $A^{\perp} \subseteq (A^{\perp})^{\perp\perp} \Rightarrow A^{\perp} \subseteq A^{\perp\perp\perp}$

Hence $A^{\perp} = A^{\perp\perp\perp}$

Theorem

Let Y be a subset of a Hilbert space H. Then $Y \cap Y^{\perp} \subseteq \{0\}$.

Proof

If $Y \cap Y^{\perp} = \varphi$ then clearly $Y \cap Y^{\perp} = \varphi \subseteq \{0\}$ so the condition is true.

If $Y \cap Y^{\perp} \neq \varphi$ then let $x \in Y \cap Y^{\perp}$ implies $x \in Y$ and $x \in Y^{\perp}$ and so $\langle x, x\rangle = 0$ i.e. $\|x\|^2 = 0 \Rightarrow x = 0 \in \{0\} \Rightarrow x \in \{0\} \Rightarrow \boldsymbol{Y \cap Y^{\perp} \subseteq \{0\}}$

Theorem

Let Y be a subset of a Hilbert space H. Then $Y^{\perp}$ is closed linear subspace of H.

Proof

Let $x, y \in Y^{\perp}$ and $\propto, \beta \in F$ then we have to show that $\propto x + \beta y \in Y^{\perp}$

Since $x, y \in Y^{\perp}$ therefore $\langle x, u\rangle = 0$ and $\langle y, u\rangle = 0$ for every $u \in Y$ then

$\langle \propto x + \beta y, u\rangle = \propto \langle x, u\rangle + \beta\langle y, u\rangle = 0 \Rightarrow \propto x + \beta y \perp Y \Rightarrow \propto x + \beta y \in Y^{\perp}$.

This shows that $Y^{\perp}$ is linear subspace of H. Now we have to show that $Y^{\perp}$ is closed. For this we just show $Y^{\perp} = \overline{Y^{\perp}}$.

We already know that $\quad Y^{\perp} \subseteq \overline{Y^{\perp}} \quad$ …………..(1)

Now let $x \in \overline{Y^{\perp}}$ then there exists a sequence (x_n) in $Y^{\perp}$ such that $x_n \to x$

Now by using continuity of inner producsts for any $u \in Y$ we have

$\langle x, u\rangle = \langle \lim_{n\to\infty} x_n , u\rangle = \lim\limits_{n\to\infty} \langle x_n, u\rangle = 0 \Rightarrow x \perp Y \Rightarrow x \in Y^{\perp}$

Implies that $\quad \overline{Y^{\perp}} \subseteq Y^{\perp} \quad$ …………..(2)

Then $Y^{\perp} = \overline{Y^{\perp}}$ and $Y^{\perp}$ is closed subspace of H.

Theorem

Let Y be a closed linear subspace of a Hilbert space H. Then $Y \cap Y^{\perp} = \{0\}$.

Proof

Since we know that if Y be a subset of a Hilbert space H. Then

$Y \cap Y^{\perp} \subseteq \{0\}$ …………….(1)

Given that Y is closed linear subspace of H and we also know that $Y^{\perp}$ is closed linear subspace of H. Let $x \in Y \cap Y^{\perp}$ implies $x \in Y$ and $x \in Y^{\perp}$ and so $\langle x, x\rangle = 0$ i.e. $\|x\|^2 = 0 \Rightarrow x = 0 \Rightarrow 0 \in Y$ and $0 \in Y^{\perp} \Rightarrow 0 \in Y \cap Y^{\perp}$

$\Rightarrow \{0\} \subseteq Y \cap Y^{\perp}$ …………….(2)

Combining (1) and (2) we get $Y \cap Y^{\perp} = \{0\}$

Projection Theorem

Let Y be any closed subspace of a Hilbert space H. Then $H = Y \oplus Y^{\perp}$

Proof

Suppose $Y + Y^{\perp}$ is proper subspace of H then there is a non – zero vector $z \in H$ such that $z \perp (Y + Y^{\perp})$. i.e. $z \in (Y + Y^{\perp})^{\perp}$

Now $Y \subseteq (Y + Y^{\perp})$ implies $(Y + Y^{\perp})^{\perp} \subseteq Y^{\perp}$

Also we know $Y^{\perp} \subseteq (Y + Y^{\perp})$ implies $(Y + Y^{\perp})^{\perp} \subseteq Y^{\perp\perp}$

Then $z \in (Y + Y^{\perp})^{\perp} \subseteq Y^{\perp} \cap Y^{\perp\perp} = \{0\} \Rightarrow z = 0$ a contradiction.

Hence $Y + Y^{\perp}$ is the whole of H. i.e. $H = Y + Y^{\perp}$ since $Y \cap Y^{\perp} = \{0\}$

Thus $H = Y \oplus Y^{\perp}$

Theorem

Let Y be a closed subset of a Hilbert space H. Then $Y = Y^{\perp\perp}$

Or Let W be a subspace of a real finite dimensional inner product space V. Then show that $(W^{\perp})^{\perp} = W$

Proof

Let $x \in Y$ then $\langle x, y \rangle = 0$ for all $y \in Y^{\perp}$

$\Rightarrow x \perp Y^{\perp} \Rightarrow x \in Y^{\perp\perp} \Rightarrow Y \subseteq Y^{\perp\perp}$

Now let $x \in Y^{\perp\perp}$ and Y be a closed subset H also $H = Y \oplus Y^{\perp}$ so;

For each $x \in Y^{\perp\perp} \subseteq H$; $x = y + z$: $y \in Y, z \in Y^{\perp}$

But $Y \subseteq Y^{\perp\perp}$ therefore $y \in Y^{\perp\perp}$

$\Rightarrow z = x - y \in Y^{\perp\perp} \Rightarrow z \perp Y^{\perp}$

But $z \in Y^{\perp} \Rightarrow z \perp z \Rightarrow z = 0 \Rightarrow z = x - y = 0 \Rightarrow x = y \Rightarrow x \in Y \Rightarrow Y^{\perp\perp} \subseteq Y$

Hence from both cases $Y = Y^{\perp\perp}$.

Theorem

For any complete subspace Y of an inner product space V, Prove that $Y = Y^{\perp\perp}$.

Proof

Let $x \in Y$ then $\langle x, y \rangle = 0$ for all $y \in Y^{\perp}$

$\Rightarrow x \perp Y^{\perp} \Rightarrow x \in Y^{\perp\perp} \Rightarrow Y \subseteq Y^{\perp\perp}$

Now let $x \in Y^{\perp\perp}$ and Y be a complete subspace of V also $V = Y \oplus Y^{\perp}$ so;

For each $x \in Y^{\perp\perp} \subseteq V$; $x = y + z$: $y \in Y \subseteq Y^{\perp\perp}, z \in Y^{\perp}$

$\Rightarrow z = x - y \in Y^{\perp\perp} \Rightarrow z \perp Y^{\perp}$

But $z \in Y^{\perp} \Rightarrow z \perp z \Rightarrow z = 0 \Rightarrow z = x - y = 0 \Rightarrow x = y \Rightarrow x \in Y \Rightarrow Y^{\perp\perp} \subseteq Y$

Hence from both cases $Y = Y^{\perp\perp}$.

Theorem

Let S be a subspace of an inner product space V. Then $S^{\perp} = [span(S)]^{\perp} = \langle S\rangle^{\perp}$

Proof

Since $S \subseteq \langle S\rangle$ therefore $\langle S\rangle^{\perp} \subseteq S^{\perp}$ ………..(1)

Let $u \in S^{\perp}$ and suppose $S = \{u_1, u_2, \dots, u_r\}$ then

$span(S) = \langle S\rangle = \{\propto_1 u_1 + \propto_2 u_2 + \cdots + \propto_r u_r : \propto_i \in F\}$

Consider $\langle \propto_1 u_1 + \cdots + \propto_r u_r, u_i\rangle$

$= \propto_1 \langle u_1, u\rangle + \propto_2 \langle u_2, u\rangle + \cdots + \propto_r \langle u_r, u\rangle$

$= \propto_1 (0) + \propto_2 (0) + \cdots + \propto_r (0) = 0$

$u \in \langle S\rangle^{\perp}$

Then $\qquad S^{\perp} \subseteq \langle S\rangle^{\perp}$ ………..(2)

From (1) and (2)

$S^{\perp} = [span(S)]^{\perp} = \langle S\rangle^{\perp}$

Orthogonal Set

Consider a set $S = \{u_1, u_2, \dots, u_r\}$ of nonzero vectors in an inner product space V. S is called orthogonal if each pair of vectors in S are orthogonal, That is,

$$\langle u_i, u_j \rangle = 0 \quad ; i \neq j$$

For example $\hat{\imath} + 2\hat{\jmath} + \hat{k}$ is orthogonal.

Orthonormal Set

Consider a set $S = \{u_1, u_2, \dots, u_r\}$ of nonzero vectors in an inner product space V. S is called orthonormal if S is orthogonal and each vector in S has unit length. i.e.

$$\langle u_i, u_j \rangle = \begin{cases} 0 & ; i \neq j \\ 1 & ; i = j \end{cases}$$

For example $\hat{\imath} + \hat{\jmath} + \hat{k}$ is orthonormal.

Example

(a) Let $E = \{e_1, e_2, e_3\} = \{(1,0,0), (0,1,0), (0,0,1)\}$ be the usual basis of Euclidean space R^3. It is clear that
$\langle e_1, e_2 \rangle = \langle e_1, e_3 \rangle = \langle e_2, e_3 \rangle = 0$ and $\langle e_1, e_1 \rangle = \langle e_2, e_3 \rangle = \langle e_3, e_3 \rangle = 1$
Namely, E is an orthonormal basis of R^3. More generally, the usual basis of R^n is orthonormal for every n.

(b) Let $V = C\,[-\pi, \pi]$ be the vector space of continuous functions on the interval $-\pi \leq t \leq \pi$ with inner product defined by $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$.
Then the following is a classical example of an orthogonal set in V: $\{1, cos\ t, cos\ 2t, cos\ 3t, \dots, sin\ t, sin\ 2t, sin\ 3t, \dots\}$
This orthogonal set plays a fundamental role in the theory of Fourier series.

Theorem

Suppose S is an orthogonal set of nonzero vectors. Then S is linearly independent.

Proof

Let $S = \{u_1, u_2, \dots, u_r\}$ be an orthogonal set of nonzero vectors, consider for all $\propto_i \in F$ we have $\propto_1 u_1 + \cdots + \propto_r u_r = 0_v$

$\langle \propto_1 u_1 + \cdots + \propto_r u_r, u_i \rangle = \langle 0_v, u_i \rangle$

$\propto_1 \langle u_1, u_i \rangle + \propto_2 \langle u_2, u_i \rangle + \cdots + \propto_i \langle u_i, u_i \rangle + \ldots + \propto_r \langle u_r, u_i \rangle = 0$

$\propto_i \langle u_i, u_i \rangle = 0 \Rightarrow \propto_i = 0$. Hence S is linearly independent.

Theorem of Pythagoras

If x and y are orthogonal vectors in a real inner product space then

$\|x + y\|^2 = \|x\|^2 + \|y\|^2$

Proof: Since x and y are orthogonal therefore $\langle x, y \rangle = 0$

$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x + y \rangle + \langle y, x + y \rangle$

$\|x + y\|^2 = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle x, y \rangle + \langle y, y \rangle$

$\|x + y\|^2 = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2 \qquad \therefore \langle x, y \rangle = 0$

Generalized Pythagoras Theorem

If $x_1, x_2, x_3, \dots, x_n$ are piecewise orthogonal vectors in a real inner product space then $\|\sum_{i=1}^{n} x_i\|^2 = \sum_{i=1}^{n} \|x_i\|^2$

Proof

$\|\sum_{i=1}^{n} x_i\|^2 = \langle \sum_{i=1}^{n} x_i, \sum_{j=1}^{n} x_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x_i, x_j \rangle$

$\|\sum_{i=1}^{n} x_i\|^2 = \sum_{i=1}^{n} \langle x_i, x_i \rangle \quad \because \langle x_i, x_j \rangle = 0 \,; i \neq j$

$\|\sum_{i=1}^{n} x_i\|^2 = \langle \sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i \rangle$

$\|\sum_{i=1}^{n} x_i\|^2 = \sum_{i=1}^{n} \|x_i\|^2$

Orthogonal Basis Set

Let $S = \{u_1, u_2, \dots, u_r\}$ be a basis set in an inner product space. Then S is said to be orthogonal basis set of vectors if $\langle u_i, u_j \rangle \;\; ; \forall i \neq j$.

Example Let S consist of the following three vectors in R^3

$u_1 = (1, 2,1)$; $u_2 = (2,1,4)$; $u_3 = (3,2,1)$

These vectors are orthogonal; hence, they are linearly independent. Thus, S is an orthogonal basis of R^3.

Suppose we want to write $v = (7,1,9)$ as a linear combination of u_1, u_2, u_3. First we set v as a linear combination of u_1, u_2, u_3 using unknowns x_1, x_2, x_3 as follows:

$v = x_1 u_1 + x_2 u_2 + x_3\, u_3$

$(7,1,9) = x_1(1, 2,1) + x_2(2,1,4) + x_3(3,2,1)$

We can proceed in two ways.

METHOD 1: Expand equation to obtain the system

$x_1 + 2x_2 + 3x_3 = 7$; $2\,x_1 + x_2 - 2x_3 = 1$; $x_1 - 4x_2 + x_3 = 7$

Solve the system by Gaussian elimination to obtain $x_1 = 3, x_2 = -1, x_3 = 2$. Thus, $v = 3u_1 - u_2 + 2\,u_3$

METHOD 2: (This method uses the fact that the basis vectors are orthogonal, and the arithmetic is much simpler.) If we take the inner product of each side of equation with respect to u_i, we get

$\langle v, u_i \rangle = \langle x_1 u_1 + x_2 u_2 + x_3\, u_3, u_i \rangle$

$\langle v, u_i \rangle = x_i \langle u_i, u_i \rangle$ or $x_i = \frac{\langle v, u_i \rangle}{\langle u_i, u_i \rangle}$

Here two terms drop out, because u_1, u_2, u_3 are orthogonal. Accordingly,

$x_1 = \frac{\langle v, u_1 \rangle}{\langle u_1, u_1 \rangle} = \frac{7+2+9}{1+4+1} = \frac{18}{6} = 3$, $x_2 = \frac{\langle v, u_2 \rangle}{\langle u_2, u_2 \rangle} = -1$, $x_3 = \frac{\langle v, u_3 \rangle}{\langle u_3, u_3 \rangle} = 2$

Thus, again, we get $v = 3u_1 - u_2 + 2\,u_3$.

Theorem

Let $u_1, u_2, \ldots, u_n$ be an orthogonal basis of V. Then, for any $v \in V$,

$$v = \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle} u_1 + \frac{\langle v,u_2\rangle}{\langle u_2,u_2\rangle} u_2 + \cdots + \frac{\langle v,u_n\rangle}{\langle u_n,u_n\rangle} u_n$$

Proof

Let $\{u_1, u_2, \ldots, u_n\}$ be a basis of V then

$v = \propto_1 u_1 + \cdots + \propto_n u_n$ ……………..(1)

Consider $\langle v, u_i\rangle = \langle \propto_1 u_1 + \cdots + \propto_i u_i + \cdots + \propto_n u_n, u_i\rangle$

$\langle v, u_i\rangle = \propto_1 \langle u_1, u_i\rangle + \propto_2 \langle u_2, u_i\rangle + \cdots + \propto_i \langle u_i, u_i\rangle + \ldots + \propto_n \langle u_n, u_n\rangle$

$\langle v, u_i\rangle = \propto_i \langle u_i, u_i\rangle$ or $\propto_i = \frac{\langle v,u_i\rangle}{\langle u_i,u_i\rangle}$; $\forall i = 1,2,3, \ldots$

$(1) \Rightarrow v = \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle} u_1 + \frac{\langle v,u_2\rangle}{\langle u_2,u_2\rangle} u_2 + \cdots + \frac{\langle v,u_n\rangle}{\langle u_n,u_n\rangle} u_n$

Theorem

Let $\{u_1, u_2, \ldots, u_n\}$ form an orthogonal set of non – zero vectors in V. Let $v \in V$ Then define $v' = v - (\propto_1 u_1 + \cdots + \propto_n u_n)$, where

$\propto_1 = \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle}, \propto_2 = \frac{\langle v,u_2\rangle}{\langle u_2,u_2\rangle}, \ldots, \propto_n = \frac{\langle v,u_n\rangle}{\langle u_n,u_n\rangle}$ then v' is orthogonal to $\{u_1, u_2, \ldots, u_n\}$

Proof

Consider $\langle v', u_i\rangle = \langle v - (\propto_1 u_1 + \cdots + \propto_n u_n), u_i\rangle$

$\langle v', u_i\rangle = \langle v, u_i\rangle - \propto_1 \langle u_1, u_i\rangle - \propto_2 \langle u_2, u_i\rangle - \cdots - \propto_i \langle u_i, u_i\rangle - \ldots - \propto_n \langle u_n, u_n\rangle$

$\langle v', u_i\rangle =$
$\langle v, u_i\rangle - \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle}\langle u_1, u_i\rangle - \frac{\langle v,u_2\rangle}{\langle u_2,u_2\rangle}\langle u_2, u_i\rangle - \cdots - \frac{\langle v,u_i\rangle}{\langle u_i,u_i\rangle}\langle u_i, u_i\rangle - \ldots - \frac{\langle v,u_n\rangle}{\langle u_n,u_n\rangle}\langle u_n, u_n\rangle$

$\langle v', u_i\rangle = \langle v, u_i\rangle - \langle v, u_i\rangle = 0$

Thus v' is orthogonal to $\{u_1, u_2, \ldots, u_n\}$

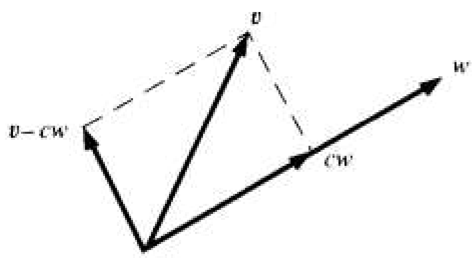
Fourier coefficient

The scalar $\frac{\langle v,u_i\rangle}{\langle u_i,u_i\rangle}$ is called the Fourier coefficient of v with respect to u_i, because it is analogous to a coefficient in the Fourier series of a function. This scalar also has a geometric interpretation, which is discussed below.

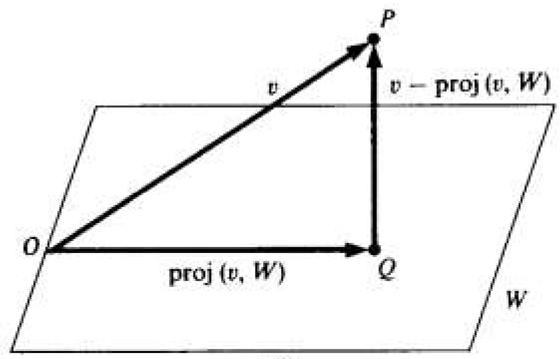
Projections

Let V be an inner product space. Suppose w is a given nonzero vector in V, and suppose v is another vector. We seek the ''projection of v along w,'' which, as indicated in Fig. (a), will be the multiple cw of w such that $v' = v - cw$ is orthogonal to w. This means

$$\langle v - cw, w\rangle = 0 \Rightarrow \langle v, w\rangle - c\langle w, w\rangle = 0 \Rightarrow c = \frac{\langle v,w\rangle}{\langle w,w\rangle}$$



(a)



(b)

Accordingly, the **projection of v along w** is denoted and defined by

$$proj(v, w) = cw = \frac{\langle v,w\rangle}{\langle w,w\rangle} w$$

Such a scalar c is unique, and it is called the Fourier coefficient of v with respect to w or the component of v along w.

We define w to be the projection of v along W, and denote it by $proj(v, W)$, as pictured in Fig. 7-2(b). In particular, if $W = span\ (w_1, w_2, \ldots, w_r)$, where the w_i form an orthogonal set, then $proj(v, W) = c_1w_1 + c_2w_2 + \cdots + c_rw_r$

Here c_i is the component of v along w_i, as above.

Theorem Suppose $w \neq 0$, let v be any vector in V. Show that $c = \frac{\langle v,w \rangle}{\langle w,w \rangle}, = \frac{\langle v,w \rangle}{\|w\|^2}$ is the unique scalar such that $v' = v - cw$ is orthogonal to w.

Proof

In order for v' to be orthogonal to w we must have

$$\langle v - cw,\ w \rangle = 0 \quad \text{or} \quad \langle v, w \rangle - c\langle w, w \rangle = 0 \quad \text{or} \quad \langle v, w \rangle = c\langle w, w \rangle$$

Thus, $c\frac{\langle v, w \rangle}{\langle w, w \rangle}$. Conversely, suppose $c = \frac{\langle v, w \rangle}{\langle w, w \rangle}$. Then

$$\langle v - cw,\ w \rangle = \langle v, w \rangle - c\langle w, w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle}\langle w, w \rangle = 0$$

Theorem

Suppose $w_1, w_2, \dots, w_r$ form an orthogonal set of nonzero vectors in V. Let v be any vector in V. Define $v' = v - (c_1 w_1 + c_2 w_2 + \cdots + c_r w_r)$ where

$c_1 = \frac{\langle v,w_1 \rangle}{\langle w_1,w_1 \rangle}, c_2 = \frac{\langle v,w_2 \rangle}{\langle w_2,w_2 \rangle}, \dots, c_r = \frac{\langle v,w_r \rangle}{\langle w_r,w_r \rangle}$

Then v' is orthogonal to $w_1, w_2, \dots, w_r$.

Proof

Suppose $w_1, w_2, \dots, w_r$ form an orthogonal set of nonzero vectors in V. Let v be any vector in V. Define $v' = v - (c_1 w_1 + c_2 w_2 + \cdots + c_r w_r)$ where

$c_1 = \frac{\langle v,w_1 \rangle}{\langle w_1,w_1 \rangle}, c_2 = \frac{\langle v,w_2 \rangle}{\langle w_2,w_2 \rangle}, \dots, c_r = \frac{\langle v,w_r \rangle}{\langle w_r,w_r \rangle}$ Then

For $i = 1, 2, \dots, r$ and using $\langle w_i, w_j \rangle = 0$ for $i \neq j$, we have

$$\begin{aligned}
\langle v - c_1 w_1 - c_2 x_2 - \cdots - c_r w_r,\ w_i \rangle &= \langle v, w_i \rangle - c_1\langle w_1, w_i \rangle - \cdots - c_i\langle w_i, w_i \rangle - \cdots - c_r\langle w_r, w_i \rangle \\
&= \langle v, w_i \rangle - c_1 \cdot 0 - \cdots - c_i\langle w_i, w_i \rangle - \cdots - c_r \cdot 0 \\
&= \langle v, w_i \rangle - c_i\langle w_i, w_i \rangle = \langle v, w_i \rangle - \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle}\langle w_i, w_i \rangle = 0
\end{aligned}$$

The theorem is proved.

Question

Find the Fourier coefficient c and the projection of $v = (1,2,3,4)$ along $w = (1,2,1,2)$ in R[4].

Solution

Compute $\langle v, w\rangle = 1 - 4 + 3 - 8 = -8$ and $\|w\|^2 = 1 + 4 + 1 + 4 = 10$. Then

$$c = -\frac{8}{10} = -\frac{4}{5} \quad \text{and} \quad \text{proj}(v, w) = cw = \left(-\frac{4}{5}, -\frac{8}{5}, -\frac{4}{5}, -\frac{8}{5}\right)$$

Gram–Schmidt Orthogonalization Process

Suppose $\{v_1, v_2, \dots, v_n\}$ is a basis of an inner product space V. One can use this basis to construct an orthogonal basis $\{w_1, w_2, \dots, w_n\}$ of V as follows. Set

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$$

$$w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2$$

....

$$w_n = v_n - \frac{\langle v_n, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_n, w_2\rangle}{\langle w_2, w_2\rangle} w_2 - \cdots - \frac{\langle v_n, w_{n-1}\rangle}{\langle w_{n-1}, w_{n-1}\rangle} w_{n-1}$$

In other words, for $k = 2, 3, \dots, n$, we define

$$w_k = v_k - \left(c_{k1}w_1 + c_{k2}w_2 + \cdots + c_{k,k-1}w_{k-1}\right)$$

Where $c_{ki} = \frac{\langle v, w_r\rangle}{\langle w_r, w_r\rangle}$ is the component of v_k along w_i. Each w_k is orthogonal to the preceeding w's. Thus, $w_1, w_2, \dots, w_n$ form an orthogonal basis for V as claimed. Normalizing each w_i will then yield an orthonormal basis for V. The above construction is known as the Gram–Schmidt orthogonalization process.

Remark

- Each vector w_k is a linear combination of v_k and the preceding w's. Hence, one can easily show, by induction, that each w_k is a linear combination of $v_1, v_2, \dots, v_n$.
- Because taking multiples of vectors does not affect orthogonality, it may be simpler in hand calculations to clear fractions in any new w_k, by multiplying w_k by an appropriate scalar, before obtaining the next w_{k+1}.
- Suppose $u_1, u_2, \dots, u_r$ are linearly independent, and so they form a basis for $U = span\ (u_i)$. Applying the Gram–Schmidt orthogonalization process to the u's yields an orthogonal basis for U.

Theorem (Proof Ommited)

Let $\{v_1, v_2, \dots, v_n\}$ be any basis of an inner product space V. Then there exists an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ of V such that the change-of-basis matrix from $\{v_i\}$ to $\{u_i\}$ is triangular; that is, for $k = 1,2, \dots, n$,

$$u_k = \propto_{k1} v_1 + \propto_{k2} v_2 + \cdots + \propto_{kk} v_k$$

Theorem

Suppose $S = \{w_1, w_2, \dots, w_r\}$ is an orthogonal basis for a subspace W of a vector space V. Then one may extend S to an orthogonal basis for V; that is, one may find vectors $w_{r+1}, , \dots, w_n$ such that $\{w_1, w_2, \dots, w_n\}$ is an orthogonal basis for V.

Proof

Extend S to a basis $S' = \{w_1, \dots, w_r, v_{r+1}, \dots, v_n\}$ for V. Applying the Gram–Schmidt algorithm to S', we first obtain $w_1, w_2, \dots, w_r$ because S is orthogonal, and then we obtain vectors $w_{r+1}, \dots, w_n$, where $\{w_1, w_2, \dots, w_n\}$ is an orthogonal basis for V. Thus, the theorem is proved.

Example

Apply the Gram–Schmidt orthogonalization process to find an orthogonal basis and then an orthonormal basis for the subspace U of R^4 spanned by

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, 2, 4, 5), \qquad v_3 = (1, -3, -4, -2)$$

(1) First set $w_1 = v_1 = (1, 1, 1, 1)$.

(2) Compute

$$v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = v_2 - \frac{12}{4} w_1 = (-2, -1, 1, 2)$$

Set $w_2 = (-2, -1, 1, 2)$.

(3) Compute

$$v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = v_3 - \frac{(-8)}{4} w_1 - \frac{(-7)}{10} w_2 = \left(\tfrac{8}{5}, -\tfrac{17}{10}, -\tfrac{13}{10}, \tfrac{7}{5}\right)$$

Clear fractions to obtain $w_3 = (-6, -17, -13, 14)$.

Thus, w_1, w_2, w_3 form an orthogonal basis for U. Normalize these vectors to obtain an orthonormal basis $\{u_1, u_2, u_3\}$ of U. We have $\|w_1\|^2 = 4$, $\|w_2\|^2 = 10$, $\|w_3\|^2 = 910$, so

$$u_1 = \frac{1}{2}(1, 1, 1, 1), \qquad u_2 = \frac{1}{\sqrt{10}}(-2, -1, 1, 2), \qquad u_3 = \frac{1}{\sqrt{910}}(16, -17, -13, 14)$$

Example

Let V be the vector space of polynomials $f(t)$ with inner product

$\langle f, g \rangle = \int_{-1}^{1} f(t) g(t) dt$

Apply the Gram–Schmidt orthogonalization process to $\{1, t, t^2, t^3\}$ to find an orthogonal basis $\{f_0, f_1, f_2, f_3\}$ with integer coefficients for $P_3(t)$.

Solution

Here we use the fact that, for $r + s = n$,

$$\langle t^r, t^s \rangle = \int_{-1}^{1} t^n dt = \left| \frac{t^{n+1}}{n+1} \right|_{-1}^{1} = \begin{cases} \dfrac{2}{n+1} & ; when\ n\ is\ even \\ 0 & ; when\ n\ is\ odd \end{cases}$$

(1) First set $f_0 = 1$.

(2) Compute $t = \frac{\langle t, 1\rangle}{\langle 1, 1\rangle}(1) = t - 0 = t$. Set $f_1 = t$.

(3) Compute

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^2, t\rangle}{\langle t, t\rangle}(t) = t^2 - \frac{\frac{2}{3}}{2}(1) + 0(t) = t^2 - \tfrac{1}{3}$$

Multiply by 3 to obtain $f_2 = 3t^2 = 1$.

(4) Compute

$$t^3 - \frac{\langle t^3, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^3, t\rangle}{\langle t, t\rangle}(t) - \frac{\langle t^3,\ 3t^2 - 1\rangle}{\langle 3t^2 - 1,\ 3t^2 - 1\rangle}(3t^2 - 1)$$
$$= t^3 - 0(1) - \frac{\frac{2}{5}}{\frac{2}{3}}(t) - 0(3t^2 - 1) = t^3 - \tfrac{3}{5}t$$

Multiply by 5 to obtain $f_3 = 5t^3 - 3t$.

Thus, $\{1,\ t,\ 3t^2 - 1,\ 5t^3 - 3t\}$ is the required orthogonal basis.

Question

Consider the vector space $\mathbf{P}(t)$ with inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Apply the Gram–Schmidt algorithm to the set $\{1, t, t^2\}$ to obtain an orthogonal set $\{f_0, f_1, f_2\}$ with integer coefficients.

First set $f_0 = 1$. Then find

$$t - \frac{\langle t, 1\rangle}{\langle 1, 1\rangle}\cdot 1 = t - \frac{\frac{1}{2}}{1}\cdot 1 = t - \frac{1}{2}$$

Clear fractions to obtain $f_1 = 2t - 1$. Then find

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^2,\ 2t - 1\rangle}{\langle 2t - 1,\ 2t - 1\rangle}(2t - 1) = t^2 - \frac{\frac{1}{3}}{1}(1) - \frac{\frac{1}{6}}{\frac{1}{3}}(2t - 1) = t^2 - t + \frac{1}{6}$$

Clear fractions to obtain $f_2 = 6t^2 - 6t + 1$. Thus, $\{1,\ 2t - 1,\ 6t^2 - 6t + 1\}$ is the required orthogonal set.

Question

. Consider the subspace U of $\mathbf{R}^4$ spanned by the vectors:

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, 1, 2, 4), \qquad v_3 = (1, 2, -4, -3)$$

Find (a) an orthogonal basis of U; (b) an orthonormal basis of U.

(a) Use the Gram–Schmidt algorithm. Begin by setting $w_1 = u = (1, 1, 1, 1)$. Next find

$$v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = (1, 1, 2, 4) - \frac{8}{4}(1, 1, 1, 1) = (-1, -1, 0, 2)$$

Set $w_2 = (-1, -1, 0, 2)$. Then find

$$v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = (1, 2, -4, -3) - \frac{(-4)}{4}(1, 1, 1, 1) - \frac{(-9)}{6}(-1, -1, 0, 2)$$
$$= \left(\tfrac{1}{2}, \tfrac{3}{2}, -3, 1\right)$$

Clear fractions to obtain $w_3 = (1, 3, -6, 2)$. Then w_1, w_2, w_3 form an orthogonal basis of U.

Geometrical Interpreatation of the Bessel Inequality اسکا حل آگے آرہا

A Geometrical Interpreatation of the Bessel Inequality is that the sum of the squares of the projections of a vector x onto a set of mutually perpendicular directions can not exceed the square of the length of the vector itself.

Bessel Inequality

Let (e_k) be anorthonormal sequence in an inner product space X. Then for every $x \in X$ we have $\sum_{k=1}^{\infty}|\langle x, e_k\rangle|^2 \leq \|x\|^2$

Proof

Let $Y_n = span\{e_1, e_2, \dots, e_n\}$ then for every $y \in Y_n$ we can express

$y = \sum_{k=1}^{n} \propto_k e_k \quad ; \propto_k = \langle y, e_k\rangle$

We claim that for a particular choice of $\propto_k$. i.e. $\propto_k = \langle x, e_k\rangle \; : x \in X$ but $x \notin Y_n$ then we can obtain $y \in Y_n$such that $z = (x - y) \perp y$ (will show this)

We first note that

$\|y\|^2 = \langle y, y\rangle = \langle \sum_{k=1}^{n} \propto_k e_k , \sum_{m=1}^{n} \propto_m e_m\rangle = \langle \sum_{k=1}^{n}\langle x, e_k\rangle e_k , \sum_{m=1}^{n}\langle x, e_m\rangle e_m\rangle$

$\|y\|^2 = \sum_{k=1}^{n}\langle x, e_k\rangle \langle e_k, \sum_{m=1}^{n}\langle x, e_m\rangle e_m\rangle = \sum_{k=1}^{n}\langle x, e_k\rangle \overline{\sum_{m=1}^{n}\langle x, e_m\rangle} \langle e_k, e_m\rangle$

$\|y\|^2 = \sum_{k=1}^{n}\langle x, e_k\rangle \sum_{m=1}^{n}\overline{\langle x, e_m\rangle} \langle e_k, e_m\rangle$

$\|y\|^2 = \sum_{k=1}^{n}\langle x, e_k\rangle \sum_{k=1}^{n}\overline{\langle x, e_k\rangle} \langle e_k, e_k\rangle$ taking $m = k$

$\|y\|^2 = \sum_{k=1}^{n}\langle x, e_k\rangle \overline{\langle x, e_k\rangle}(1)$

$\|y\|^2 = \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 = \sum_{k=1}^{n}|\langle x, e_k\rangle|^2$ ………(1)

Now consider

$\langle z, y\rangle = \langle x - y, y\rangle = \langle x, y\rangle - \langle y, y\rangle = \langle x, \sum_{k=1}^{n} \propto_k e_k\rangle - \|y\|^2$

$\langle z, y\rangle = \langle x, \sum_{k=1}^{n}\langle x, e_k\rangle e_k\rangle - \|y\|^2 = \overline{\sum_{k=1}^{n}\langle x, e_k\rangle} \langle x, e_k\rangle - \sum_{k=1}^{n}|\langle x, e_k\rangle|^2$

$\langle z, y\rangle = \sum_{k=1}^{n}\overline{\langle x, e_k\rangle} \langle x, e_k\rangle - \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 = \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 - \sum_{k=1}^{n}|\langle x, e_k\rangle|^2$

$\langle z, y\rangle = 0$ Implies $z \perp y$

Now $z = x - y$ then using Pyhtagorian Theorem $\|z\|^2 = \|x\|^2 - \|y\|^2$

$0 \leq \|z\|^2 = \|x\|^2 - \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 \Rightarrow \sum_{k=1}^{n}|\langle x, e_k\rangle|^2 \leq \|x\|^2$

$\Rightarrow \sum_{k=1}^{\infty}|\langle x, e_k\rangle|^2 \leq \|x\|^2$ if $n \to \infty$. Hence proved.

Bessel Inequality (Another Form)

Suppose $\{e_1, e_2, \dots, e_k\}$ is an orthonormal set of vectors in an inner product space X. Let $x \in X$ be any arbitrary vector and c_k be the fourier coefficients of vector x with respect to e_k then $\sum_{k=1}^{r} c_k^2 \le \|x\|^2$

Proof

Consider $\langle x - \sum_{k=1}^{r} c_k e_k , x - \sum_{k=1}^{r} c_k e_k \rangle \ge 0$

$\langle x, x\rangle - \langle x, \sum_{k=1}^{r} c_k e_k\rangle - \langle \sum_{k=1}^{r} c_k e_k , x\rangle + \langle \sum_{k=1}^{r} c_k e_k , \sum_{k=1}^{r} c_k e_k\rangle \ge 0$

$\|x\|^2 - 2\langle x, \sum_{k=1}^{r} c_k e_k\rangle + \sum_{k=1}^{r} c_k^2 \langle e_k, e_k\rangle \ge 0$

$\|x\|^2 - 2\sum_{k=1}^{r} c_k \langle x, e_k\rangle + \sum_{k=1}^{r} c_k^2 \langle e_k, e_k\rangle \ge 0$

$\|x\|^2 - 2\sum_{k=1}^{r} c_k \frac{\langle x, e_k\rangle}{\langle e_k, e_k\rangle} + \sum_{k=1}^{r} c_k^2 \ge 0$

$\|x\|^2 - 2\sum_{k=1}^{r} c_k^2 + \sum_{k=1}^{r} c_k^2 \ge 0$

$\|x\|^2 - \sum_{k=1}^{r} c_k^2 \ge 0$

$\sum_{k=1}^{r} c_k^2 \le \|x\|^2$ Hence proved.

Orthogonal Matrices

A square matrix P is orthogonal if P is nonsingular and its transpose is the same as its inverse, that is $P^{-1} = P^T$, or, equivalently, if $PP^T = P^T P = I$.

Question

Show that the matrix $P = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$ is orthogonal.

Solution

$$PP^T = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{2}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{6}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{6}{7} & 0 \end{bmatrix}$$

Theorem (Keep in mind)

Let P be a real matrix. Then the following are equivalent:

(a) P is orthogonal;

(b) the rows of P form an orthonormal set;

(c) the columns of P form an orthonormal set.

(This theorem is true only using the usual inner product on R^n. It is not true if R^n is given any other inner product.)

Example

(a) Let $P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \end{bmatrix}$. The rows of P are orthogonal to each other and are unit vectors. Thus P is an orthogonal matrix.

(b) **Rotation and Reflection matrices are orthogonal:** Let P be a 2×2 orthogonal matrix. Then, for some real number θ, we have

$P = \begin{bmatrix} Cos\theta & Sin\theta \\ -Sin\theta & Cos\theta \end{bmatrix}$ or $P = \begin{bmatrix} Cos\theta & Sin\theta \\ Sin\theta & -Cos\theta \end{bmatrix}$

The following two theorems show important relationships between orthogonal matrices and orthonormal bases of a real inner product space V.

Theorem

Suppose $E = \{e_i\}$ and $E' = \{e'_i\}$ are orthonormal bases of V. Let P be the change-of-basis matrix from the basis E to the basis E . Then P is orthogonal.

Suppose

$$e'_i = b_{i1}e_1 + b_{i2}e_2 + \cdots + b_{in}e_n, \qquad i = 1, \ldots, n \tag{1}$$

Using Problem 7.18(b) and the fact that E' is orthonormal, we get

$$\delta_{ij} = \langle e'_i, e'_j \rangle = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \tag{2}$$

Let $B = [b_{ij}]$ be the matrix of the coefficients in (1). (Then $P = B^T$.) Suppose $BB^T = [c_{ij}]$. Then

$$c_{ij} = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \tag{3}$$

By (2) and (3), we have $c_{ij} = \delta_{ij}$. Thus, $BB^T = I$. Accordingly, B is orthogonal, and hence, $P = B^T$ is orthogonal.

Remember

7.18. Suppose $E = \{e_1, e_2, \ldots, e_n\}$ is an orthonormal basis of V. Prove

(a) For any $u \in V$, we have $u = \langle u, e_1 \rangle e_1 + \langle u, e_2 \rangle e_2 + \cdots + \langle u, e_n \rangle e_n$.

(b) $\langle a_1 e_1 + \cdots + a_n e_n, \ b_1 e_1 + \cdots + b_n e_n \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.

(c) For any $u, v \in V$, we have $\langle u, v \rangle = \langle u, e_1 \rangle \langle v, e_1 \rangle + \cdots + \langle u, e_n \rangle \langle v, e_n \rangle$.

Theorem

Let $\{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of an inner product space V. Let $P = \{a_{ij}\}$ be an orthogonal matrix. Then the following n vectors form an orthonormal basis for V:

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n \quad ; i = 1,2, \ldots, n$$

Proof

Prove Theorem 7.13: Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of an inner product space V. Let $P = [a_{ij}]$ be an orthogonal matrix. Then the following n vectors form an orthonormal basis for V:

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, 2, \ldots, n$$

Because $\{e_i\}$ is orthonormal, we get, by Problem 7.18(b),

$$\langle e'_i, e'_j \rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \langle C_i, C_j \rangle$$

where C_i denotes the ith column of the orthogonal matrix $P = [a_{ij}]$. Because P is orthogonal, its columns form an orthonormal set. This implies $\langle e'_i, e'_j \rangle = \langle C_i, C_j \rangle = \delta_{ij}$. Thus, $\{e'_i\}$ is an orthonormal basis.

Question

Find an orthogonal matrix P whose first row is $u_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$

Solution

First find a nonzero vector $w_2 = (x, y, z)$ that is orthogonal to u_1—that is, for which

$$0 = \langle u_1, w_2 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{or} \qquad x + 2y + 2z = 0$$

One such solution is $w_2 = (0, 1, -1)$. Normalize w_2 to obtain the second row of P:

$$u_2 = (0, 1/\sqrt{2}, -1/\sqrt{2})$$

Next find a nonzero vector $w_3 = (x, y, z)$ that is orthogonal to both u_1 and u_2—that is, for which

$$0 = \langle u_1, w_3 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{or} \qquad x + 2y + 2z = 0$$

$$0 = \langle u_2, w_3 \rangle = \frac{y}{\sqrt{2}} - \frac{y}{\sqrt{2}} = 0 \qquad \text{or} \qquad y - z = 0$$

Set $z = -1$ and find the solution $w_3 = (4, -1, -1)$. Normalize w_3 and obtain the third row of P; that is

$$u_3 = (4/\sqrt{18}, -1/\sqrt{18}, -1/\sqrt{18}).$$

Thus,

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 4/3\sqrt{2} & -1/3\sqrt{2} & -1/3\sqrt{2} \end{bmatrix}$$

We emphasize that the above matrix P is not unique.

Question

Let $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 4 \\ 7 & -5 & 2 \end{bmatrix}$. Determine whether or not: (a) the rows of A are orthogonal;

(b) A is an orthogonal matrix; (c) the columns of A are orthogonal.

(a) Yes, because $(1, 1, -1) \cdot (1, 3, 4) = 1 + 3 - 4 = 0$, $(1, 1 - 1) \cdot (7, -5, 2) = 7 - 5 - 2 = 0$, and $(1, 3, 4) \cdot (7, -5, 2) = 7 - 15 + 8 = 0$.

(b) No, because the rows of A are not unit vectors, for example, $(1, 1, -1)^2 = 1 + 1 + 1 = 3$.

(c) No; for example, $(1, 1, 7) \cdot (1, 3, -5) = 1 + 3 - 35 = -31 \neq 0$.

Theorem Prove each of the following:

(a) P is orthogonal if and only if P^T is orthogonal.

(b) If P is orthogonal, then P^{-1} is orthogonal.

(c) If P and Q are orthogonal, then PQ is orthogonal.

Proof

(a) We have $(P^T)^T = P$. Thus, P is orthogonal if and only if $PP^T = I$ if and only if $P^{TT}P^T = I$ if and only if P^T is orthogonal.

(b) We have $P^T = P^{-1}$, because P is orthogonal. Thus, by part (a), P^{-1} is orthogonal.

(c) We have $P^T = P^{-1}$ and $Q^T = Q^{-1}$. Thus, $(PQ)(PQ)^T = PQQ^TP^T = PQQ^{-1}P^{-1} = I$. Therefore, $(PQ)^T = (PQ)^{-1}$, and so PQ is orthogonal.

Practice

Determine whether that matrix is orthogonal. If so find its inverse.

1. (a) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ Theorem 7.1 (b) $\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

2. (a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$

3. (a) $\begin{bmatrix} 0 & 1 & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$ (b) $\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

4. (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{5}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{6} & -\frac{5}{6} \\ \frac{1}{2} & \frac{1}{6} & -\frac{5}{6} & \frac{1}{6} \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{3}} & \frac{1}{2} & 0 \end{bmatrix}$

Positive Definite Matrices

Let A be a real symmetric matrix; that is, $A^T = A$. Then A is said to be positive definite if, for every nonzero vector **u** in $\mathbf{R}^n$, $\langle u, Au\rangle = u^T A u > 0$

Theorem

A 2×2 real symmetric matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ is positive definite if and only if the diagonal entries a and d are positive and the determinant

$|A| = ad - bc = ad - b^2$ is positive.

Example Consider the following symmetric matrices:

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

A is not positive definite, because $|A| = 4 - 9 = -5$ is negative. B is not positive definite, because the diagonal entry -3 is negative. However, C is positive definite, because the diagonal entries 1 and 5 are positive, and the determinant $|C| = 5 - 4 = 1$ is also positive.

Question

Which of the following symmetric matrices are positive definite?

(a) $A = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 8 & -3 \\ -3 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$, (d) $D = \begin{bmatrix} 3 & 5 \\ 5 & 9 \end{bmatrix}$

Use Theorem 7.14 that a 2×2 real symmetric matrix is positive definite if and only if its diagonal entries are positive and if its determinant is positive.

(a) No, because $|A| = 15 - 16 = -1$ is negative.

(b) Yes.

(c) No, because the diagonal entry -3 is negative.

(d) Yes.

Question

Find the values of k that make each of the following matrices positive definite:

(a) $A = \begin{bmatrix} 2 & -4 \\ -4 & k \end{bmatrix}$, (b) $B = \begin{bmatrix} 4 & k \\ k & 9 \end{bmatrix}$, (c) $C = \begin{bmatrix} k & 5 \\ 5 & -2 \end{bmatrix}$

(a) First, k must be positive. Also, $|A| = 2k - 16$ must be positive; that is, $2k - 16 > 0$. Hence, $k > 8$.

(b) We need $|B| = 36 - k^2$ positive; that is, $36 - k^2 > 0$. Hence, $k^2 < 36$ or $-6 < k < 6$.

(c) C can never be positive definite, because C has a negative diagonal entry -2.

Question

Find the matrix A that represents the usual inner product on $\mathbf{R}^2$ relative to each of the following bases of $\mathbf{R}^2$: (a) $\{v_1 = (1,3),\ v_2 = (2,5)\}$; (b) $\{w_1 = (1,2),\ w_2 = (4,-2)\}$.

(a) Compute $\langle v_1, v_1 \rangle = 1 + 9 = 10$, $\langle v_1, v_2 \rangle = 2 + 15 = 17$, $\langle v_2, v_2 \rangle = 4 + 25 = 29$. Thus,

$$A = \begin{bmatrix} 10 & 17 \\ 17 & 29 \end{bmatrix}.$$

(b) Compute $\langle w_1, w_1 \rangle = 1 + 4 = 5$, $\langle w_1, w_2 \rangle = 4 - 4 = 0$, $\langle w_2, w_2 \rangle = 16 + 4 = 20$. Thus, $A = \begin{bmatrix} 5 & 0 \\ 0 & 20 \end{bmatrix}$.
(Because the basis vectors are orthogonal, the matrix A is diagonal.)

Theorem

Let A be a real positive definite matrix. Then the function $\langle u, v \rangle = u^T A v$ is an inner product on $\mathbf{R}^n$.

For any vectors u_1, u_2, and v,

$$\langle u_1 + u_2,\ v \rangle = (u_1 + u_2)^T A v = (u_1^T + u_2^T) A v = u_1^T A v + u_2^T A v = \langle u_1, v \rangle + \langle u_2, v \rangle$$

and, for any scalar k and vectors u, v,

$$\langle ku, v \rangle = (ku)^T A v = k u^T A v = k \langle u, v \rangle$$

Thus $[I_1]$ is satisfied.

Because $u^T A v$ is a scalar, $(u^T A v)^T = u^T A v$. Also, $A^T = A$ because A is symmetric. Therefore,

$$\langle u, v \rangle = u^T A v = (u^T A v)^T = v^T A^T u^{TT} = v^T A u = \langle v, u \rangle$$

Thus, $[I_2]$ is satisfied.

Last, because A is positive definite, $X^T A X > 0$ for any nonzero $X \in \mathbf{R}^n$. Thus, for any nonzero vector v, $\langle v, v \rangle = v^T A v > 0$. Also, $\langle 0, 0 \rangle = 0^T A 0 = 0$. Thus, $[I_3]$ is satisfied. Accordingly, the function $\langle u, v \rangle = Av$ is an inner product.

Matrix Representation of an Inner Product (Optional)

Every positive definite matrix A determines an inner product on $\mathbf{R}^n$. This subsection may be viewed as giving the converse of this result.

Let V be a real inner product space with basis $S = \{u_1, u_2, \dots, u_n\}$. The matrix $A = \{a_{ij}\}$; where $a_{ij} = \langle u_i, u_j\rangle$ is called the matrix representation of the inner product on V relative to the basis S.

Observe that A is symmetric, because the inner product is symmetric; that is, $\langle u_i, u_j\rangle = \langle u_j, u_i\rangle$. Also, A depends on both the inner product on V and the basis S for V. Moreover, if S is an orthogonal basis, then A is diagonal, and if S is an orthonormal basis, then A is the identity matrix.

Example

The vectors $u_1 = (1,1,0)\,, u_2 = (1,2,3)\,, u_3 = (1,3,5)$ form a basis S for Euclidean space $\mathbf{R}^3$. Find the matrix A that represents the inner product in $\mathbf{R}^3$ relative to this basis S.

Solution

First compute each $\langle u_i, u_j\rangle$ to obtain

$$\langle u_1, u_1\rangle = 1+1+0 = 2, \qquad \langle u_1, u_2\rangle = 1+2+0 = 3, \qquad \langle u_1, u_3\rangle = 1+3+0 = 4$$
$$\langle u_2, u_2\rangle = 1+4+9 = 14, \qquad \langle u_2, u_3\rangle = 1+6+15 = 22, \qquad \langle u_3, u_3\rangle = 1+9+25 = 35$$

Then $A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 14 & 22 \\ 4 & 22 & 35 \end{bmatrix}$. As expected, A is symmetric.

Theorem

Let A be the matrix representation of an inner product relative to basis S for V.

Then, for any vectors $u, v \in V$, we have $\langle u, u\rangle = [u]^T A[v]$

where $[u]$ and $[v]$ denote the (column) coordinate vectors relative to the basis S.

Prove Theorem 7.16: Let A be the matrix representation of an inner product relative to a basis S of V. Then, for any vectors $u, v \in V$, we have

$$\langle u, v\rangle = [u]^T A[v]$$

Suppose $S = \{w_1, w_2, \ldots, w_n\}$ and $A = [k_{ij}]$. Hence, $k_{ij} = \langle w_i, w_j\rangle$. Suppose

$$u = a_1w_1 + a_2w_2 + \cdots + a_nw_n \quad \text{and} \quad v = b_1w_1 + b_2w_2 + \cdots + b_nw_n$$

Then

$$\langle u, v\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j\langle w_i, w_j\rangle \tag{1}$$

On the other hand,

$$[u]^T A[v] = (a_1, a_2, \ldots, a_n)\begin{bmatrix} k_{11} & k_{12} & \ldots & k_{1n} \\ k_{21} & k_{22} & \ldots & k_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ k_{n1} & k_{n2} & \ldots & k_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= \left(\sum_{i=1}^{n} a_ik_{i1}, \ \sum_{i=1}^{n} a_ik_{i2}, \ \ldots, \ \sum_{i=1}^{n} a_ik_{in}\right)\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{j=1}^{n}\sum_{i=1}^{n} a_ib_jk_{ij} \tag{2}$$

Equations (1) and (2) give us our result.

Theorem

Let A be the matrix representation of any inner product on V. Then A is a positive definite matrix.

Prove Theorem 7.17: Let A be the matrix representation of any inner product on V. Then A is a positive definite matrix.

Because $\langle w_i, w_j\rangle = \langle w_j, w_i\rangle$ for any basis vectors w_i and w_j, the matrix A is symmetric. Let X be any nonzero vector in $\mathbf{R}^n$. Then $[u] = X$ for some nonzero vector $u \in V$. Theorem 7.16 tells us that $X^T AX = [u]^T A[u] = \langle u, u\rangle > 0$. Thus, A is positive definite.

CHAPTER # 5

DETERMINANTS

Determinant of a matrix

Each n-square matrix $A = [a_{ij}]$ is assigned a special scalar called the determinant of A, denoted by $detA$or $|A|$ or $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

Or Let A be a 2×2 matrix, i.e. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then its determinant could be define as follows which is a scalar number

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

Remember: We emphasize that an $n \times n$ array of scalars enclosed by straight lines, called a determinant of order n, is not a matrix but denotes the determinant of the enclosed array of scalars (i.e., the enclosed matrix).

The determinant function was first discovered during the investigation of systems of linear equations.

Determinants of Orders 1 and 2

Determinants of orders 1 and 2 are defined as follows:

$|a_{11}| = a_{11}$ and $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

Example

(a) Because the determinant of order 1 is the scalar itself, we have:

$$\det(27) = 27, \qquad \det(-7) = -7, \qquad \det(t-3) = t-3$$

(b) $\begin{vmatrix} 5 & 3 \\ 4 & 6 \end{vmatrix} = 5(6) - 3(4) = 30 - 12 = 18,$ $\qquad \begin{vmatrix} 3 & 2 \\ -5 & 7 \end{vmatrix} = 21 + 10 = 31$

Determinants of Order 3

Consider an arbitrary 3×3 matrix $A = [a_{ij}]$. The determinant of A is defined as follows:

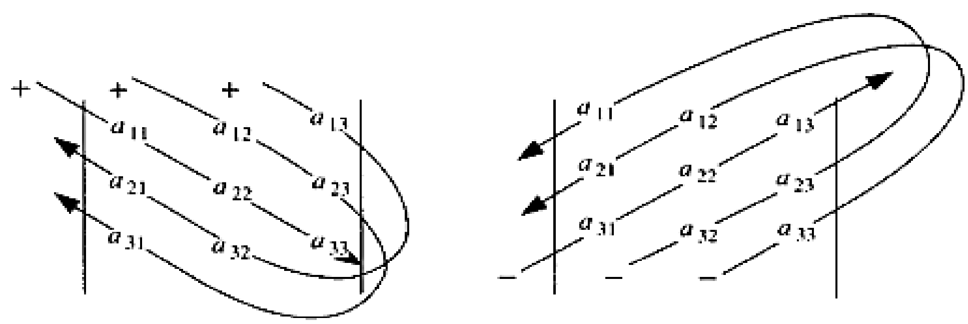
$$det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Observe that there are six products, each product consisting of three elements of the original matrix.

Three of the products are plus-labeled (keep their sign) and three of the products are minus-labeled (change their sign).

The diagrams in Fig. 8-1 may help us to remember the above six products in detA. That is, the determinant is equal to the sum of the products of the elements along the three plus-labeled arrows in Fig. 8-1 plus the sum of the negatives of the products of the elements along the three minus-labeled arrows. We emphasize that there are no such diagrammatic devices with which to remember determinants of higher order.



EXAMPLE 8.3 Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2 & 1 \\ -4 & 5 & -1 \\ 2 & -3 & 4 \end{bmatrix}$. Find $\det(A)$ and $\det(B)$.

Use the diagrams in Fig. 8-1:

$$\det(A) = 2(5)(4) + 1(-2)(1) + 1(-3)(0) - 1(5)(1) - (-3)(-2)(2) - 4(1)(0)$$
$$= 40 - 2 + 0 - 5 - 12 - 0 = 21$$
$$\det(B) = 60 - 4 + 12 - 10 - 9 + 32 = 81$$

Alternative Form for a Determinant of Order 3

The determinant of the 3×3 matrix $A = [a_{ij}]$ may be rewritten as follows:

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$det(A) = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Example

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

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

$$= 1(2 - 15) - 2(-4 + 0) + 3(20 + 0) = -13 + 8 + 60 = 55$$

Permutation

A permutation σ of the set $\{1, 2, \ldots, n\}$ is a one-to-one mapping of the set onto itself or, equivalently, a rearrangement of the numbers $1, 2, \ldots, n$. Such a permutation σ is denoted by

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} \qquad \text{or} \qquad \sigma = j_1 j_2 \cdots j_n, \qquad \text{where } j_i = \sigma(i)$$

The set of all such permutations is denoted by S_n, and the number of such permutations is $n!$. If $\sigma \in S_n$, then the inverse mapping $\sigma^{-1} \in S_n$; and if $\sigma, \tau \in S_n$, then the composition mapping $\sigma \circ \tau \in S_n$. Also, the identity mapping $\varepsilon = \sigma \circ \sigma^{-1} \in S_n$. (In fact, $\varepsilon = 123 \ldots n$.)

EXAMPLE 8.5

(a) There are $2! = 2 \cdot 1 = 2$ permutations in S_2; they are 12 and 21.

(b) There are $3! = 3 \cdot 2 \cdot 1 = 6$ permutations in S_3; they are 123, 132, 213, 231, 312, 321.

Sign (Parity) of a Permutation

Consider an arbitrary permutation σ in S_n, say $\sigma = j_1 j_2 \cdots j_n$. We say σ is an even or odd permutation according to whether there is an even or odd number of inversions in σ. By an *inversion* in σ we mean a pair of integers (i, k) such that $i > k$, but i precedes k in σ. We then define the sign or parity of σ, written sgn σ, by

$$\operatorname{sgn} \sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

EXAMPLE 8.6

(a) Find the sign of $\sigma = 35142$ in S_5.

For each element k, we count the number of elements i such that $i > k$ and i precedes k in σ. There are

2 numbers (3 and 5) greater than and preceding 1,
3 numbers (3, 5, and 4) greater than and preceding 2,
1 number (5) greater than and preceding 4.

(There are no numbers greater than and preceding either 3 or 5.) Because there are, in all, six inversions, σ is even and sgn $\sigma = 1$.

(b) The identity permutation $\varepsilon = 123\ldots n$ is even because there are no inversions in ε.

(c) In S_2, the permutation 12 is even and 21 is odd. In S_3, the permutations 123, 231, 312 are even and the permutations 132, 213, 321 are odd.

(d) Let τ be the permutation that interchanges two numbers i and j and leaves the other numbers fixed. That is,

$$\tau(i) = j, \qquad \tau(j) = i, \qquad \tau(k) = k, \quad \text{where} \quad k \neq i, j$$

We call τ a *transposition*. If $i < j$, then there are $2(j - i) - 1$ inversions in τ, and hence, the transposition τ is odd.

Properties of Determinants

Theorem: The determinant of a matrix A and its transpose A^T are equal; that is, $|A| = |A^T|$.

Proof

If $A = [a_{ij}]$ then $A^T = [b_{ij}]$ with $b_{ij} = a_{ji}$, hence

$$|A^T| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

Let $\tau = \sigma^{-1}$. By Problem 8.21 sgn τ = sgn σ, and $a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} = a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$. Hence,

$$|A^T| = \sum_{\sigma \in S_n} (\operatorname{sgn} \tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$$

However, as s runs through all the elements of S_n; $\tau = \sigma^{-1}$ also runs through all the elements of Sn. Thus $|A| = |A^T|$

Theorem: Let A be a square matrix.

(i) If A has a row (column) of zeros, then $|A| = 0$.
(ii) If A has two identical rows (columns), then $|A| = 0$.
(iii) If A is triangular (i.e., A has zeros above or below the diagonal), then $|A| =$ product of diagonal elements. Thus, in particular, $|I| = 1$, where I is the identity matrix.

Proof

(i) Each term in $|A|$ contains a factor from every row, and so from the row of zeros. Thus, each term of $|A|$ is zero, and so $|A| = 0$.

(ii) Suppose $1 + 1 \neq 0$ in K. If we interchange the two identical rows of A, we still obtain the matrix A. Hence, by Problem 8.23, $|A| = -|A|$, and so $|A| = 0$.

Now suppose $1 + 1 = 0$ in K. Then sgn $\sigma = 1$ for every $\sigma \in S_n$. Because A has two identical rows, we can arrange the terms of A into pairs of equal terms. Because each pair is 0, the determinant of A is zero.

(iii) Suppose $A = [a_{ij}]$ is lower triangular; that is, the entries above the diagonal are all zero: $a_{ij} = 0$ whenever $i < j$. Consider a term t of the determinant of A:

$$t = (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n}, \qquad \text{where} \qquad \sigma = i_1 i_2 \cdots i_n$$

Suppose $i_1 \neq 1$. Then $1 < i_1$ and so $a_{1i_1} = 0$; hence, $t = 0$. That is, each term for which $i_1 \neq 1$ is zero.

Now suppose $i_1 = 1$ but $i_2 \neq 2$. Then $2 < i_2$, and so $a_{2i_2} = 0$; hence, $t = 0$. Thus, each term for which $i_1 \neq 1$ or $i_2 \neq 2$ is zero.

Similarly, we obtain that each term for which $i_1 \neq 1$ or $i_2 \neq 2$ or ... or $i_n \neq n$ is zero. Accordingly, $|A| = a_{11} a_{22} \cdots a_{nn} =$ product of diagonal elements.

Theorem

Suppose B is obtained from A by an elementary row (column) operation. If two rows (columns) of A were interchanged, then $|B| = -|A|$.

Proof

We prove the theorem for the case that two columns are interchanged. Let τ be the transposition that interchanges the two numbers corresponding to the two columns of A that are interchanged. If $A = [a_{ij}]$ and $B = [b_{ij}]$, then $b_{ij} = a_{i\tau(j)}$. Hence, for any permutation σ,

$$b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = a_{1(\tau\circ\sigma)(1)} a_{2(\tau\circ\sigma)(2)} \cdots a_{n(\tau\circ\sigma)(n)}$$

Thus,

$$|B| = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1(\tau\circ\sigma)(1)} a_{2(\tau\circ\sigma)(2)} \cdots a_{n(\tau\circ\sigma)(n)}$$

Because the transposition τ is an odd permutation, $\operatorname{sgn}(\tau\circ\sigma) = (\operatorname{sgn}\tau)(\operatorname{sgn}\sigma) = -\operatorname{sgn}\sigma$. Accordingly, $\operatorname{sgn}\sigma = -\operatorname{sgn}(\tau\circ\sigma)$, and so

$$|B| = -\sum_{\sigma\in S_n} [\operatorname{sgn}(\tau\circ\sigma)] a_{1(\tau\circ\sigma)(1)} a_{2(\tau\circ\sigma)(2)} \cdots a_{n(\tau\circ\sigma)(n)}$$

But as σ runs through all the elements of S_n, $\tau\circ\sigma$ also runs through all the elements of S_n. Hence, $|B| = -|A|$.

Theorem

The determinant of a product of two matrices A and B is the product of their determinants; that is, $det(AB) = det(A)det(B)$

The above theorem says that the determinant is a multiplicative function.

8.36. Prove Theorem 8.12: Suppose M is an upper (lower) triangular block matrix with diagonal blocks $A_1, A_2, \ldots, A_n$. Then

$$\det(M) = \det(A_1)\det(A_2)\cdots\det(A_n)$$

We need only prove the theorem for $n = 2$—that is, when M is a square matrix of the form $M = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$. The proof of the general theorem follows easily by induction.

Suppose $A = [a_{ij}]$ is r-square, $B = [b_{ij}]$ is s-square, and $M = [m_{ij}]$ is n-square, where $n = r + s$. By definition,

$$\det(M) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}$$

If $i > r$ and $j \leq r$, then $m_{ij} = 0$. Thus, we need only consider those permutations σ such that

$$\sigma\{r+1, r+2, \ldots, r+s\} = \{r+1, r+2, \ldots, r+s\} \qquad \text{and} \qquad \sigma\{1, 2, \ldots, r\} = \{1, 2, \ldots, r\}$$

Let $\sigma_1(k) = \sigma(k)$ for $k \leq r$, and let $\sigma_2(k) = \sigma(r+k) - r$ for $k \leq s$. Then

$$(\operatorname{sgn} \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)} = (\operatorname{sgn} \sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} \cdots a_{r\sigma_1(r)} (\operatorname{sgn} \sigma_2) b_{1\sigma_2(1)} b_{2\sigma_2(2)} \cdots b_{s\sigma_2(s)}$$

which implies $\det(M) = \det(A)\det(B)$.

Remember (8.27) Suppose B is row equivalent to a square matrix A. then $|B| = 0$ if and only if $|A| = 0$.

Theorem

Let A be a square matrix. Then the following are equivalent:

(i) A is invertible; that is, A has an inverse A^{-1}
(ii) $AX = 0$ has only the zero solution.
(iii) The determinant of A is not zero; that is, $detA \neq 0$.

The proof is by the Gaussian algorithm. If A is invertible, it is row equivalent to I. But $|I| \neq 0$. Hence, by Problem 8.27, $|A| \neq 0$. If A is not invertible, it is row equivalent to a matrix with a zero row. Hence, $\det(A) = 0$. Thus, (i) and (iii) are equivalent.

If $AX = 0$ has only the solution $X = 0$, then A is row equivalent to I and A is invertible. Conversely, if A is invertible with inverse A^{-1}, then

$$X = IX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}0 = 0$$

is the only solution of $AX = 0$. Thus, (i) and (ii) are equivalent.

Lemma

Let E be an elementary matrix. Then, for any matrix A; $|EA| = |E||A|$.

Consider the elementary row operations: (i) Multiply a row by a constant $k \neq 0$,
(ii) Interchange two rows, (iii) Add a multiple of one row to another.

Let E_1, E_2, E_3 be the corresponding elementary matrices That is, E_1, E_2, E_3 are obtained by applying the above operations to the identity matrix I. By Problem 8.25,

$$|E_1| = k|I| = k, \qquad |E_2| = -|I| = -1, \qquad |E_3| = |I| = 1$$

Recall (Theorem 3.11) that E_iA is identical to the matrix obtained by applying the corresponding operation to A. Thus, by Theorem 8.3, we obtain the following which proves our lemma:

$$|E_1A| = k|A| = |E_1||A|, \qquad |E_2A| = -|A| = |E_2||A|, \qquad |E_3A| = |A| = 1|A| = |E_3||A|$$

Theorem

Suppose B is obtained from A by an elementary row (column) operation.

(i) If a row (column) of A were multiplied by a scalar k, then $|B| = k|A|$.

(ii) If a multiple of a row (column) of A were added to another row (column) of A, then $|B| = |A|$.

Proof

(ii) If the jth row of A is multiplied by k, then every term in $|A|$ is multiplied by k, and so $|B| = k|A|$. That is,

$$|B| = \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots (k a_{ji_j}) \cdots a_{ni_n} = k \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n} = k|A|$$

(iii) Suppose c times the kth row is added to the jth row of A. Using the symbol $\hat{}$ to denote the jth position in a determinant term, we have

$$\begin{aligned} |B| &= \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots (\widehat{c a_{ki_k} + a_{ji_j}}) \ldots a_{ni_n} \\ &= c \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots \widehat{a_{ki_k}} \cdots a_{ni_n} + \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ji_j} \cdots a_{ni_n} \end{aligned}$$

The first sum is the determinant of a matrix whose kth and jth rows are identical. Accordingly, by Theorem 8.2(ii), the sum is zero. The second sum is the determinant of A. Thus, $|B| = c \cdot 0 + |A| = |A|$.

8.29. Prove Theorem 8.4: $|AB| = |A||B|$.

If A is singular, then AB is also singular, and so $|AB| = 0 = |A||B|$. On the other hand, if A is nonsingular, then $A = E_n \cdots E_2 E_1$, a product of elementary matrices. Then, Lemma 8.6 and induction yields

$$|AB| = |E_n \cdots E_2 E_1 B| = |E_n| \cdots |E_2||E_1||B| = |A||B|$$

8.30. Suppose P is invertible. Prove that $|P^{-1}| = |P|^{-1}$.

$P^{-1}P = I$. Hence, $1 = |I| = |P^{-1}P| = |P^{-1}||P|$, and so $|P^{-1}| = |P|^{-1}$.

8.31. Prove Theorem 8.7: Suppose A and B are similar matrices. Then $|A| = |B|$.

Because A and B are similar, there exists an invertible matrix P such that $B = P^{-1}AP$. Therefore, using Problem 8.30, we get $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A||P^{-1}||P = |A|$.

We remark that although the matrices P^{-1} and A may not commute, their determinants $|P^{-1}|$ and $|A|$ do commute, because they are scalars in the field K.

8.34. Prove Theorem 8.9: $A(\text{adj } A) = (\text{adj } A)A = |A|I$.

Let $A = [a_{ij}]$ and let $A(\text{adj } A) = [b_{ij}]$. The ith row of A is

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \tag{1}$$

Because adj A is the transpose of the matrix of cofactors, the jth column of adj A is the tranpose of the cofactors of the jth row of A:

$$(A_j, A_{j2}, \ldots, A_{jn})^T \tag{2}$$

Now b_{ij}, the ij entry in $A(\text{adj } A)$, is obtained by multiplying expressions (1) and (2):

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

By Theorem 8.8 and Problem 8.33,

$$b_{ij} = \begin{cases} |A| & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Accordingly, $A(\text{adj } A)$ is the diagonal matrix with each diagonal element $|A|$. In other words, $A(\text{adj } A) = |A|I$. Similarly, $(\text{adj } A)A = |A|I$.

Minors and Cofactors

Consider an n-square matrix $A = [a_{ij}]$. Let M_{ij} denote the $(n-1)$ –square submatrix of A obtained by deleting its *ith* row and *jth* column. The determinant $|M_{ij}|$ is called the minor of the element a_{ij} of A.

Minors and Cofactors

Consider an n-square matrix $A = [a_{ij}]$. Let M_{ij} denote the $(n-1)$ –square submatrix of A obtained by deleting its *ith* row and *jth* column. The determinant $|M_{ij}|$ is called the minor of the element a_{ij} of A, and we define the cofactor of a_{ij}, denoted by A_{ij}; to be the ‘‘signed’’ minor:

$$A_{ij} = (-1)^{i+j}|M_{ij}|$$

We emphasize that M_{ij} denotes a matrix, whereas A_{ij} denotes a scalar.

Example

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Find the following minors and cofactors:

(a) $|M_{23}|$ and A_{23}
(b) $|M_{31}|$ and A_{31}

Solution

(a) $|M_{23}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = 8 - 14 = -6$, and so $A_{23} = (-1)^{2+3}|M_{23}| = -(-6) = 6$

(b) $|M_{31}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 12 - 15 = -3$, and so $A_{31} = (-1)^{1+3}|M_{31}| = +(-3) = -3$

Practice

1. Let

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 6 & 7 & -1 \\ -3 & 1 & 4 \end{bmatrix}$$

(a) Find all the minors of A.

(b) Find all the cofactors.

2. Let

$$A = \begin{bmatrix} 4 & -1 & 1 & 6 \\ 0 & 0 & -3 & 3 \\ 4 & 1 & 0 & 14 \\ 4 & 1 & 3 & 2 \end{bmatrix}$$

Find

(a) M_{13} and C_{13}

(b) M_{23} and C_{23}

(c) M_{22} and C_{22}

(d) M_{21} and C_{21}

Cofactor Expansion of a Matrix A / Determinant using Cofactor

If A is $n \times n$ matrix, then the number obtained by multiplying the entries in any row or column of A by the corresponding cofactors and adding the resulting products is called the determinant of A and the sums themselves is called cofactor expansion of A, that is

$detA = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$ (cofactor expansion along the j^{th} columns)

$detA = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$ (cofactor expansion along the i^{th} rows)

Example (Cofactor Expansion along the First Row)

Find the determinant of the matrix $A = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$ by cofactor expansion along the first row.

Solution

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

$$detA = \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} = 3(-4) - 1(-11) + 0 = -1$$

Example (Cofactor Expansion along the Column)

Find the determinant of the matrix $A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 3 & 1 & 2 & 2 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & 0 & 1 \end{bmatrix}$ by cofactor expansion along the second column.

Solution

Since second column has the most zeros we will expand along the second column,

$$detA = 1.\begin{vmatrix} 1 & 0 & -1 \\ 1 & -2 & 1 \\ 2 & 0 & 1 \end{vmatrix} = 1.-2.\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = -2(1+2) = -6$$

Classical Adjoint

Let $A = [a_{ij}]$ be an $n \times n$ matrix over a field K and let A_{ij} denote the cofactor of a_{ij}. The classical adjoint of A, denoted by $adjA$, is the transpose of the matrix of cofactors of A. Namely, $adjA = [A_{ij}]^T$

We say ‘‘classical adjoint’’ instead of simply ‘‘adjoint’’ because the term ‘‘adjoint’’ is currently used for an entirely different concept.

Example

Let $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$. The cofactors of the nine elements of A follow:

$$A_{11} = +\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -18, \qquad A_{12} = -\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} = 2, \qquad A_{13} = +\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} = 4$$

$$A_{21} = -\begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} = -11, \qquad A_{22} = +\begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} = 14, \qquad A_{23} = -\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 5$$

$$A_{31} = +\begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} = -10, \qquad A_{32} = -\begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} = -4, \qquad A_{33} = +\begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8$$

The transpose of the above matrix of cofactors yields the classical adjoint of A; that is,

$$\operatorname{adj} A = \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix}$$

Theorem

Let A be any square matrix. Then

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = |A|I$$

where I is the identity matrix. Thus, if $|A| \neq 0$,

$$A^{-1} = \frac{1}{|A|}(\operatorname{adj} A)$$

Proof

Let $A = [a_{ij}]$ and let $A(\text{adj } A) = [b_{ij}]$. The ith row of A is

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \tag{1}$$

Because adj A is the transpose of the matrix of cofactors, the jth column of adj A is the tranpose of the cofactors of the jth row of A:

$$(A_{j}, A_{j2}, \ldots, A_{jn})^T \tag{2}$$

Now b_{ij}, the ij entry in $A(\text{adj } A)$, is obtained by multiplying expressions (1) and (2):

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

By Theorem 8.8 and Problem 8.33,

$$b_{ij} = \begin{cases} |A| & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Accordingly, $A(\text{adj } A)$ is the diagonal matrix with each diagonal element $|A|$. In other words, $A(\text{adj } A) = |A|I$. Similarly, $(\text{adj } A)A = |A|I$.

Example

Let $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$. **then we have**

$$\det(A) = -40 + 6 + 0 - 16 + 4 + 0 = -46$$

Thus, A does have an inverse, and, by Theorem 8.9,

$$A^{-1} = \frac{1}{|A|}(\text{adj } A) = -\frac{1}{46}\begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix} = \begin{bmatrix} \frac{9}{23} & \frac{11}{46} & \frac{5}{23} \\ -\frac{1}{23} & -\frac{7}{23} & \frac{2}{23} \\ -\frac{2}{23} & -\frac{5}{46} & \frac{4}{23} \end{bmatrix}$$

Cramer's Rule

If $A\mathbf{x}=\mathbf{b}$ is a system of n linear equations in n unknowns such that $\det(A)\neq 0$, then the system has a unique solution. This solution is

$$x_1=\frac{\det(A_1)}{\det(A)},\qquad x_2=\frac{\det(A_2)}{\det(A)},\ \ldots,\qquad x_n=\frac{\det(A_n)}{\det(A)}$$

where A_j is the matrix obtained by replacing the entries in the jth column of A by the entries in the matrix

$$\mathbf{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Proof If $\det(A)\neq 0$, then A is invertible, and by Theorem 1.6.2, $\mathbf{x}=A^{-1}\mathbf{b}$ is the unique solution of $A\mathbf{x}=\mathbf{b}$. Therefore, by Theorem 2.1.2 we have

$$\mathbf{x}=A^{-1}\mathbf{b}=\frac{1}{\det(A)}\operatorname{adj}(A)\mathbf{b}=\frac{1}{\det(A)}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Multiplying the matrices out gives

$$\mathbf{x}=\frac{1}{\det(A)}\begin{bmatrix} b_1C_{11}+b_2C_{21}+\cdots+b_nC_{n1} \\ b_1C_{12}+b_2C_{22}+\cdots+b_nC_{n2} \\ \vdots \qquad \vdots \qquad\qquad \vdots \\ b_1C_{1n}+b_2C_{2n}+\cdots+b_nC_{nn} \end{bmatrix}$$

The entry in the jth row of $\mathbf{x}$ is therefore

$$x_j=\frac{b_1C_{1j}+b_2C_{2j}+\cdots+b_nC_{nj}}{\det(A)} \tag{11}$$

Now let

$$A_j = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{bmatrix}$$

Since A_j differs from A only in the jth column, it follows that the cofactors of entries $b_1, b_2, \cdots, b_n$ in A_j are the same as the cofactors of the corresponding entries in the jth column of A. The cofactor expansion of $\det(A_j)$ along the jth column is therefore

$$\det(A_j) = b_1C_{1j} + b_2C_{2j} + \cdots + b_nC_{nj}$$

Substituting this result in 11 gives

$$x_j = \frac{\det(A_j)}{\det(A)}$$

Procedure

- Write given system $AX = B$
- Find $|A| \neq 0, |A_x|, |A_y|$
- Find solution set using $x = \frac{|A_x|}{|A|}, y = \frac{|A_y|}{|A|}$

Example

Solve the system using Cramer's Rule.

$x_1 + 2x_3 = 6$; $-3x_1 + 4x_2 + 6x_3 = 30$; $-x_1 - 2x_2 + 3x_3 = 8$

Solution

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 31 \end{bmatrix}, A_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 31 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 31 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

$$x_1 = \frac{|A_1|}{|A|} = -\frac{40}{44} = \frac{10}{11}\ ;\ x_2 = \frac{|A_2|}{|A|} = \frac{72}{44} = \frac{18}{11};\ \ x_3 = \frac{|A_3|}{|A|} = \frac{152}{44} = \frac{38}{11}$$

Theorem

A square homogeneous system $AX = 0$ has a nonzero solution if and only if $D = |A| = 0$.

EXAMPLE 8.12 Solve the system using determinants $\begin{cases} x + y + z = 5 \\ x - 2y - 3z = -1 \\ 2x + y - z = 3 \end{cases}$

First compute the determinant D of the matrix of coefficients:

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

Because $D \neq 0$, the system has a unique solution. To compute N_x, N_y, N_z, we replace, respectively, the coefficients of x, y, z in the matrix of coefficients by the constant terms. This yields

$$N_x = \begin{vmatrix} 5 & 1 & 1 \\ -1 & -2 & -3 \\ 3 & 1 & -1 \end{vmatrix} = 20, \qquad N_y = \begin{vmatrix} 1 & 5 & 1 \\ 1 & -1 & -3 \\ 2 & 3 & -1 \end{vmatrix} = -10, \qquad N_z = \begin{vmatrix} 1 & 1 & 5 \\ 1 & -2 & -1 \\ 2 & 1 & 3 \end{vmatrix} = 15$$

Thus, the unique solution of the system is $x = N_x/D = 4$, $y = N_y/D = -2$, $z = N_z/D = 3$; that is, the vector $u = (4, -2, 3)$.

8.10. Consider the system $\begin{cases} kx + y + z = 1 \\ x + ky + z = 1 \\ x + y + kz = 1 \end{cases}$

Use determinants to find those values of k for which the system has

(a) a unique solution, (b) more than one solution, (c) no solution.

(a) The system has a unique solution when $D \neq 0$, where D is the determinant of the matrix of coefficients. Compute

$$D = \begin{vmatrix} k & 1 & 1 \\ 1 & k & 1 \\ 1 & 1 & k \end{vmatrix} = k^3 + 1 + 1 - k - k - k = k^3 - 3k + 2 = (k-1)^2(k+2)$$

Thus, the system has a unique solution when

$$(k-1)^2(k+2) \neq 0, \quad \text{when } k \neq 1 \text{ and } k \neq 2$$

(b and c) Gaussian elimination shows that the system has more than one solution when $k = 1$, and the system has no solution when $k = -2$.

CHAPTER # 6

DIAGONALIZATION

Diagonalizable Matrix

Suppose an n-square matrix A is given. The matrix A is said to be diagonalizable if there exists a nonsingular matrix P such that $B = P^{-1}AP$ is diagonal.

Similar Matrix

A matrix B is similar to a matrix A if there exists a nonsingular matrix P such that $B = P^{-1}AP$ implies $PB = AP$.

Diagonalizable Operator

Suppose a linear operator $T: V \to V$ is given. The linear operator T is said to be diagonalizable if there exists a basis S of V such that the matrix representation of T relative to the basis S is a diagonal matrix D.

Characteristic Polynomial/ Characteristic Matrix/ Characteristic Equation

Let $A = [a_{ij}]$ be an n-square matrix over field F. Then the matrix $M = tI - A$ is called characteristic matrix of A and $\Delta_A(t) = |tI - A|$ is called the characteristic polynomial of A. And $\Delta_A(t) = |tI - A| = 0$ is called the characteristic equation of A.

Example

Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$. Its characteristic polynomial is

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 \\ -4 & t-5 \end{vmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

Characteristic Polynomials of Degrees 2 and 3

There are simple formulas for the characteristic polynomials of matrices of orders 2 and 3.

(a) Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \det(A) = t^2 - \mathrm{tr}(A)\, t + \det(A)$$

Here $\mathrm{tr}(A)$ denotes the trace of A—that is, the sum of the diagonal elements of A.

(b) Suppose $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Then

$$\Delta(t) = t^3 - \mathrm{tr}(A)\, t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

(Here A_{11}, A_{22}, A_{33} denote, respectively, the cofactors of a_{11}, a_{22}, a_{33}.)

EXAMPLE 9.3 Find the characteristic polynomial of each of the following matrices:

(a) $A = \begin{bmatrix} 5 & 3 \\ 2 & 10 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -1 \\ 6 & 2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 5 & -2 \\ 4 & -4 \end{bmatrix}$.

(a) We have $\mathrm{tr}(A) = 5 + 10 = 15$ and $|A| = 50 - 6 = 44$; hence, $\Delta(t) + t^2 - 15t + 44$.

(b) We have $\mathrm{tr}(B) = 7 + 2 = 9$ and $|B| = 14 + 6 = 20$; hence, $\Delta(t) = t^2 - 9t + 20$.

(c) We have $\mathrm{tr}(C) = 5 - 4 = 1$ and $|C| = -20 + 8 = -12$; hence, $\Delta(t) = t^2 - t - 12$.

EXAMPLE 9.4 Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$.

We have $\mathrm{tr}(A) = 1 + 3 + 9 = 13$. The cofactors of the diagonal elements are as follows:

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 3 & 9 \end{vmatrix} = 21, \qquad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 9 \end{vmatrix} = 7, \qquad A_{33} = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3$$

Thus, $A_{11} + A_{22} + A_{33} = 31$. Also, $|A| = 27 + 2 + 0 - 6 - 6 - 0 = 17$. Accordingly,

$$\Delta(t) = t^3 - 13t^2 + 31t - 17$$

9.1. Let $A = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}$. Find $f(A)$, where

(a) $f(t) = t^2 - 3t + 7$, (b) $f(t) = t^2 - 6t + 13$

First find $A^2 = \begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix}$. Then

(a) $f(A) = A^2 - 3A + 7I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -3 & 6 \\ -12 & -15 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -3 & -6 \\ 12 & 9 \end{bmatrix}$

(b) $f(A) = A^2 - 6A + 13I = \begin{bmatrix} -7 & -12 \\ 24 & 17 \end{bmatrix} + \begin{bmatrix} -6 & 12 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

[Thus, A is a root of $f(t)$.]

9.2. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 7 & -3 \\ 5 & -2 \end{bmatrix}$, (c) $C = \begin{bmatrix} 3 & -2 \\ 9 & -3 \end{bmatrix}$

Use the formula $(t) = t^2 - \mathrm{tr}(M)\, t + |M|$ for a 2×2 matrix M:

(a) $\mathrm{tr}(A) = 2 + 1 = 3$, $|A| = 2 - 20 = -18$, so $\Delta(t) = t^2 - 3t - 18$

(b) $\mathrm{tr}(B) = 7 - 2 = 5$, $|B| = -14 + 15 = 1$, so $\Delta(t) = t^2 - 5t + 1$

(c) $\mathrm{tr}(C) = 3 - 3 = 0$, $|C| = -9 + 18 = 9$, so $\Delta(t) = t^2 + 9$

9.3. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 4 \\ 6 & 4 & 5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 6 & -2 \\ -3 & 2 & 0 \\ 0 & 3 & -4 \end{bmatrix}$

Use the formula $\Delta(t) = t^3 - \mathrm{tr}(A)t^2 + (A_{11} + A_{22} + A_{33})t - |A|$, where A_{ii} is the cofactor of a_{ii} in the 3×3 matrix $A = [a_{ij}]$.

(a) $\mathrm{tr}(A) = 1 + 0 + 5 = 6$,

$$A_{11} = \begin{vmatrix} 0 & 4 \\ 4 & 5 \end{vmatrix} = -16, \qquad A_{22} = \begin{vmatrix} 1 & 3 \\ 6 & 5 \end{vmatrix} = -13, \qquad A_{33} = \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = -6$$

$$A_{11} + A_{22} + A_{33} = -35, \quad \text{and} \quad |A| = 48 + 36 - 16 - 30 = 38$$

Thus, $$\Delta(t) = t^3 - 6t^2 - 35t - 38$$

(b) $\mathrm{tr}(B) = 1 + 2 - 4 = -1$

$$B_{11} = \begin{vmatrix} 2 & 0 \\ 3 & -4 \end{vmatrix} = -8, \qquad B_{22} = \begin{vmatrix} 1 & -2 \\ 0 & -4 \end{vmatrix} = -4, \qquad B_{33} = \begin{vmatrix} 1 & 6 \\ -3 & 2 \end{vmatrix} = 20$$

$$B_{11} + B_{22} + B_{33} = 8, \quad \text{and} \quad |B| = -8 + 18 - 72 = -62$$

Thus, $$\Delta(t) = t^3 + t^2 - 8t + 62$$

Remark: Suppose $A = [a_{ij}]$ is a triangular matrix. Then $tI - A$ is a triangular matrix with diagonal entries $t - a_{ii}$; hence,

$$\Delta(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

Observe that the roots of $\Delta(t)$ are the diagonal elements of A.

EXAMPLE 9.2 Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$. Its characteristic polynomial is

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 \\ -4 & t-5 \end{vmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

As expected from the Cayley–Hamilton theorem, A is a root of $\Delta(t)$; that is,

$$\Delta(A) = A^2 - 6A - 7I = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix} + \begin{bmatrix} -6 & -18 \\ -24 & -30 \end{bmatrix} + \begin{bmatrix} -7 & 0 \\ 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

9.4. Find the characteristic polynomial $\Delta(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 2 & 5 & 1 & 1 \\ 1 & 4 & 2 & 2 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 2 & 3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 3 & 4 \\ 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 6 \end{bmatrix}$

(a) A is block triangular with diagonal blocks

$$A_1 = \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 6 & -5 \\ 2 & 3 \end{bmatrix}$$

Thus, $$\Delta(t) = \Delta_{A_1}(t)\Delta_{A_2}(t) = (t^2 - 6t + 3)(t^2 - 9t + 28)$$

(b) Because B is triangular, $\Delta(t) = (t-1)(t-3)(t-5)(t-6)$.

Theorem

Similar matrices have the same characteristic polynomial.

Proof

Let A and B are similar matrix then $B = P^{-1}AP$ and using $tI = P^{-1}tIP$

$|tI - B| = |P^{-1}tIP - P^{-1}AP| = |P^{-1}(tI - A)P| = |P^{-1}P||tI - A|$

$|tI - B| = |tI - A|$

Thus Similar matrices have the same characteristic polynomial.

Cayley – Hamilton Theorem

Every matrix A is a root of its characteristic polynomial.

Or Every Square matrix is zero of its characteristic polynomial.

Or Every Square matrix satisfies its characteristic equation.

Or if Δt characteristic polynomial of a square matrix A, then A is root of Δt.

Proof

Let A be an arbitrary n-square matrix and let $\Delta(t)$ be its characteristic polynomial, say,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Now let $B(t)$ denote the classical adjoint of the matrix $tI - A$. The elements of $B(t)$ are cofactors of the matrix $tI - A$ and hence are polynomials in t of degree not exceeding $n - 1$. Thus,

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

where the B_i are n-square matrices over K which are independent of t. By the fundamental property of the classical adjoint (Theorem 8.9), $(tI - A)B(t) = |tI - A|I$, or

$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Removing the parentheses and equating corresponding powers of t yields

$$B_{n-1} = I, \quad B_{n-2} - AB_{n-1} = a_{n-1}I, \quad \ldots, \quad B_0 - AB_1 = a_1 I, \quad -AB_0 = a_0 I$$

Multiplying the above equations by $A^n, A^{n-1}, \ldots, A, I$, respectively, yields

$$A^n B_{n-1} = A_n I, \quad A^{n-1}B_{n-2} - A^n B_{n-1} = a_{n-1}A^{n-1}, \quad \ldots, \quad AB_0 - A^2 B_1 = a_1 A, \quad -AB_0 = a_0 I$$

Adding the above matrix equations yields 0 on the left-hand side and $\Delta(A)$ on the right-hand side; that is,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

Therefore, $\Delta(A) = 0$, which is the Cayley–Hamilton theorem.

Eigenvalue

Let A be any square matrix. A scalar λ is called an eigenvalue of A if there exists a nonzero (column) vector v such that $Av = \lambda v$

Any vector satisfying this relation is called an eigenvector of A belonging to the eigenvalue l λ.

Note

- Each scalar multiple kv of an eigenvector v belonging to λ is also such an eigenvector, because $A(kv) = k(Av) = k\ (\lambda v) = \lambda(kv)$
- The set E_λ of all eigenvectors is a subspace of V, called the **eigenspace** of λ.
- If $dimE_\lambda = 1$,then E_λ is called an **eigenline** and λ is called a **scaling factor**.
- The terms characteristic value and characteristic vector (or proper value and proper vector) are sometimes used instead of eigenvalue and eigenvector.

EXAMPLE 9.5 Let $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ and let $v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

$$Av_1 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = v_1 \quad \text{and} \quad Av_2 = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4v_2$$

Thus, v_1 and v_2 are eigenvectors of A belonging, respectively, to the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 4$.

Question

Show that a matrix A and its transpose A^T have the same characteristic polynomial.

By the transpose operation, $(tI - A)^T = tI^T - A^T = tI - A^T$. Because a matrix and its transpose have the same determinant,

$$\Delta_A(t) = |tI - A| = |(tI - A)^T| = |tI - A^T| = \Delta_{A^T}(t)$$

Theorem: Let A be a square matrix. Then the following are equivalent.

(i) A scalar λ is an eigenvalue of A.
(ii) The matrix $M = A - \lambda I$ is singular.
(iii) The scalar λ is a root of the characteristic polynomial Δt of A.

Proof

The scalar λ is an eigenvalue of A if and only if there exists a nonzero vector v such that

$$Av = \lambda v \quad \text{or} \quad (\lambda I)v - Av = 0 \quad \text{or} \quad (\lambda I - A)v = 0$$

or $\lambda I - A$ is singular. In such a case, λ is a root of $\Delta(t) = |tI - A|$. Also, v is in the eigenspace E_λ of λ if and only if the above relations hold. Hence, v is a solution of $(\lambda I - A)X = 0$.

Theorem: Suppose $v_1, v_2, \dots, v_n$ are nonzero eigenvectors of a matrix A belonging to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then $v_1, v_2, \dots, v_n$ are linearly independent.

Suppose the theorem is not true. Let $v_1, v_2, \dots, v_s$ be a minimal set of vectors for which the theorem is not true. We have $s > 1$, because $v_1 \neq 0$. Also, by the minimality condition, $v_2, \dots, v_s$ are linearly independent. Thus, v_1 is a linear combination of $v_2, \dots, v_s$, say,

$$v_1 = a_2 v_2 + a_3 v_3 + \cdots + a_s v_s \tag{1}$$

(where some $a_k \neq 0$). Applying T to (1) and using the linearity of T yields

$$T(v_1) = T(a_2 v_2 + a_3 v_3 + \cdots + a_s v_s) = a_2 T(v_2) + a_3 T(v_3) + \cdots + a_s T(v_s) \tag{2}$$

Because v_j is an eigenvector of T belonging to λ_j, we have $T(v_j) = \lambda_j v_j$. Substituting in (2) yields

$$\lambda_1 v_1 = a_2 \lambda_2 v_2 + a_3 \lambda_3 v_3 + \cdots + a_s \lambda_s v_s \tag{3}$$

Multiplying (1) by λ_1 yields

$$\lambda_1 v_1 = a_2 \lambda_1 v_2 + a_3 \lambda_1 v_3 + \cdots + a_s \lambda_1 v_s \tag{4}$$

Setting the right-hand sides of (3) and (4) equal to each other, or subtracting (3) from (4) yields

$$a_2(\lambda_1 - \lambda_2)v_2 + a_3(\lambda_1 - \lambda_3)v_3 + \cdots + a_s(\lambda_1 - \lambda_s)v_s = 0 \tag{5}$$

Because $v_2, v_3, \dots, v_s$ are linearly independent, the coefficients in (5) must all be zero. That is,

$$a_2(\lambda_1 - \lambda_2) = 0, \qquad a_3(\lambda_1 - \lambda_3) = 0, \qquad \dots, \qquad a_s(\lambda_1 - \lambda_s) = 0$$

However, the λ_i are distinct. Hence $\lambda_1 - \lambda_j \neq 0$ for $j > 1$. Hence, $a_2 = 0$, $a_3 = 0, \dots, a_s = 0$. This contradicts the fact that some $a_k \neq 0$. The theorem is proved.

Algebraic multiplicity

If λ is an eigenvalue of a matrix A, then the algebraic multiplicity of λ is defined to be the multiplicity of λ as a root of the characteristic polynomial of A, and the geometric multiplicity of λ is defined to be the dimension of its eigenspace, $dimE_\lambda$.

Theorem

The geometric multiplicity of an eigenvalue λ of a matrix A does not exceed its algebraic multiplicity.

Suppose the geometric multiplicity of λ is r. Then its eigenspace E_λ contains r linearly independent eigenvectors $v_1, \dots, v_r$. Extend the set $\{v_i\}$ to a basis of V, say, $\{v_i, \dots, v_r, w_1, \dots, w_s\}$. We have

$$T(v_1) = \lambda v_1, \qquad T(v_2) = \lambda v_2, \qquad \dots, \qquad T(v_r) = \lambda v_r,$$

$$\begin{aligned} T(w_1) &= a_{11}v_1 + \cdots + a_{1r}v_r + b_{11}w_1 + \cdots + b_{1s}w_s \\ T(w_2) &= a_{21}v_1 + \cdots + a_{2r}v_r + b_{21}w_1 + \cdots + b_{2s}w_s \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ T(w_s) &= a_{s1}v_1 + \cdots + a_{sr}v_r + b_{s1}w_1 + \cdots + b_{ss}w_s \end{aligned}$$

Then $M = \begin{bmatrix} \lambda I_r & A \\ 0 & B \end{bmatrix}$ is the matrix of T in the above basis, where $A = [a_{ij}]^T$ and $B = [b_{ij}]^T$.

Because M is block diagonal, the characteristic polynomial $(t - \lambda)^r$ of the block λI_r must divide the characteristic polynomial of M and hence of T. Thus, the algebraic multiplicity of λ for T is at least r, as required.

Minimal Polynomial

A polynomial $m(t)$ is called minimal polynomial of the matrix A if;

i. $m(t)$ divides the characteristic polynomial $\Delta(t)$
ii. Each irreducible factor of $\Delta(t)$ divides $m(t)$
iii. $m(A) = 0$

Theorem

The minimal polynomial $m(t)$ of a matrix (linear operator) A divides every polynomial that has A as a zero. In particular, $m(t)$ divides the characteristic polynomial $\Delta(t)$ of A.

Proof

Suppose $f(t)$ is a polynomial for which $f(A) = 0$. By the division algorithm, there exist polynomials $q(t)$ and $r(t)$ for which $f(t) = m(t)q(t) + r(t)$ and $r(t) = 0$ or $\deg r(t) < \deg m(t)$. Substituting $t = A$ in this equation, and using that $f(A) = 0$ and $m(A) = 0$, we obtain $r(A) = 0$. If $r(t) \neq 0$, then $r(t)$ is a polynomial of degree less than $m(t)$ that has A as a zero. This contradicts the definition of the minimal polynomial. Thus, $r(t) = 0$, and so $f(t) = m(t)q(t)$; that is, $m(t)$ divides $f(t)$.

9.34. Let $m(t)$ be the minimal polynomial of an n-square matrix A. Prove that the characteristic polynomial $\Delta(t)$ of A divides $[m(t)]^n$.

Suppose $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Define matrices B_j as follows:

$$\begin{array}{lll}
B_0 = I & \text{so} & I = B_0 \\
B_1 = A + c_1 I & \text{so} & c_1 I = B_1 - A = B_1 - AB_0 \\
B_2 = A^2 + c_1 A + c_2 I & \text{so} & c_2 I = B_2 - A(A + c_1 I) = B_2 - AB_1 \\
\cdots\cdots & & \cdots\cdots \\
B_{r-1} = A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I & \text{so} & c_{r-1} I = B_{r-1} - AB_{r-2}
\end{array}$$

Then

$$-AB_{r-1} = c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) = c_r I - m(A) = c_r I$$

Set

$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

Then

$$\begin{aligned}
(tI - A)B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\
&= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\
&= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} tI + c_r I = m(t)I
\end{aligned}$$

Taking the determinant of both sides gives $|tI - A||B(t)| = |m(t)I| = [m(t)]^n$. Because $|B(t)|$ is a polynomial, $|tI - A|$ divides $[m(t)]^n$; that is, the characteristic polynomial of A divides $[m(t)]^n$.

Theorem

The characteristic polynomial $\Delta(t)$ and the minimal polynomial $m(t)$ of a matrix A have the same irreducible factors.

Roof

Suppose $f(t)$ is an irreducible polynomial. If $f(t)$ divides $m(t)$, then $f(t)$ also divides $\Delta(t)$ [because $m(t)$ divides $\Delta(t)$]. On the other hand, if $f(t)$ divides $\Delta(t)$, then by Problem 9.34, $f(t)$ also divides $[m(t)]^n$. But $f(t)$ is irreducible; hence, $f(t)$ also divides $m(t)$. Thus, $m(t)$ and $\Delta(t)$ have the same irreducible factors.

EXAMPLE 9.11 Find the minimal polynomial $m(t)$ of $A = \begin{bmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{bmatrix}$.

First find the characteristic polynomial $\Delta(t)$ of A. We have

$$\text{tr}(A) = 5, \qquad A_{11} + A_{22} + A_{33} = 2 - 3 + 8 = 7, \qquad \text{and} \qquad |A| = 3$$

Hence,

$$\Delta(t) = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$$

The minimal polynomial $m(t)$ must divide $\Delta(t)$. Also, each irreducible factor of $\Delta(t)$ (i.e., $t-1$ and $t-3$) must also be a factor of $m(t)$. Thus, $m(t)$ is exactly one of the following:

$$f(t) = (t-3)(t-1) \qquad \text{or} \qquad g(t) = (t-3)(t-1)^2$$

We know, by the Cayley–Hamilton theorem, that $g(A) = \Delta(A) = 0$. Hence, we need only test $f(t)$. We have

$$f(A) = (A - I)(A - 3I) = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $f(t) = m(t) = (t-1)(t-3) = t^2 - 4t + 3$ is the minimal polynomial of A.

9.28. Find the minimal polynomial $m(t)$ of each of the following matrices:

(a) $A = \begin{bmatrix} 5 & 1 \\ 3 & 7 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$, (c) $C = \begin{bmatrix} 4 & -1 \\ 1 & 2 \end{bmatrix}$

(a) The characteristic polynomial of A is $\Delta(t) = t^2 - 12t + 32 = (t-4)(t-8)$. Because $\Delta(t)$ has distinct factors, the minimal polynomial $m(t) = \Delta(t) = t^2 - 12t + 32$.

(b) Because B is triangular, its eigenvalues are the diagonal elements 1, 2, 3; and so its characteristic polynomial is $\Delta(t) = (t-1)(t-2)(t-3)$. Because $\Delta(t)$ has distinct factors, $m(t) = \Delta(t)$.

(c) The characteristic polynomial of C is $\Delta(t) = t^2 - 6t + 9 = (t-3)^2$. Hence the minimal polynomial of C is $f(t) = t - 3$ or $g(t) = (t-3)^2$. However, $f(C) \neq 0$; that is, $C - 3I \neq 0$. Hence,

$$m(t) = g(t) = \Delta(t) = (t-3)^2.$$

EXAMPLE 9.9 Let $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$, a real symmetric matrix. Find an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

First we find the characteristic polynomial $\Delta(t)$ of A. We have

$$\mathrm{tr}(A) = 2 + 5 = 7, \qquad |A| = 10 - 4 = 6; \qquad \text{so} \qquad \Delta(t) = t^2 - 7t + 6 = (t-6)(t-1)$$

Accordingly, $\lambda_1 = 6$ and $\lambda_2 = 1$ are the eigenvalues of A.

(a) Subtracting $\lambda_1 = 6$ down the diagonal of A yields the matrix

$$M = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} \quad \text{and the homogeneous system} \quad \begin{aligned} -4x - 2y &= 0 \\ -2x - y &= 0 \end{aligned} \quad \text{or} \quad 2x + y = 0$$

A nonzero solution is $u_1 = (1, -2)$.

Question (PP)

Find a real orthogonal matrix P for which $P^{-1}AP$ is diagonal, where $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

CHAPTER # 7

LINEAR TRANSFORMATIONS

DEFINITION: Let V and U be vector spaces over the same field K. A mapping $F: V \to U$ is called a *linear mapping* or *linear transformation* if it satisfies the following two conditions:

(1) For any vectors $v, w \in V$, $F(v+w) = F(v) + F(w)$.

(2) For any scalar k and vector $v \in V$, $F(kv) = kF(v)$.

Namely, $F: V \to U$ is linear if it "preserves" the two basic operations of a vector space, that of vector addition and that of scalar multiplication.

Examples

(a) Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the "projection" mapping into the xy-plane; that is, F is the mapping defined by $F(x,y,z) = (x,y,0)$. We show that F is linear. Let $v = (a,b,c)$ and $w = (a',b',c')$. Then

$$F(v+w) = F(a+a',\ b+b',\ c+c') = (a+a',\ b+b',\ 0)$$
$$= (a,b,0) + (a',b',0) = F(v) + F(w)$$

and, for any scalar k,

$$F(kv) = F(ka,kb,kc) = (ka,kb,0) = k(a,b,0) = kF(v)$$

Thus, F is linear.

(b) Let $G: \mathbf{R}^2 \to \mathbf{R}^2$ be the "translation" mapping defined by $G(x,y) = (x+1,\ y+2)$. [That is, G adds the vector $(1,2)$ to any vector $v = (x,y)$ in $\mathbf{R}^2$.] Note that

$$G(0) = G(0,0) = (1,2) \neq 0$$

Thus, the zero vector is not mapped into the zero vector. Hence, G is not linear.

Example

Suppose the mapping $F : \mathbf{R}^2 \to \mathbf{R}^2$ is defined by $F(x, y) = (x + y, \ x)$. Show that F is linear.

We need to show that $F(v + w) = F(v) + F(w)$ and $F(kv) = kF(v)$, where u and v are any elements of $\mathbf{R}^2$ and k is any scalar. Let $v = (a, b)$ and $w = (a', b')$. Then

$$v + w = (a + a', \ b + b') \qquad \text{and} \qquad kv = (ka, kb)$$

We have $F(v) = (a + b, a)$ and $F(w) = (a' + b', a')$. Thus,

$$\begin{aligned} F(v + w) &= F(a + a', \ b + b') = (a + a' + b + b', \ a + a') \\ &= (a + b, \ a) + (a' + b', \ a') = F(v) + F(w) \end{aligned}$$

and

$$F(kv) = F(ka, kb) = (ka + kb, \ ka) = k(a + b, \ a) = kF(v)$$

Because v, w, k were arbitrary, F is linear.

Example

. Suppose $F : \mathbf{R}^3 \to \mathbf{R}^2$ is defined by $F(x, y, z) = (x + y + z, \ 2x - 3y + 4z)$. Show that F is linear.

We argue via matrices. Writing vectors as columns, the mapping F may be written in the form $F(v) = Av$, where $v = [x, y, z]^T$ and

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

Then, using properties of matrices, we have

$$F(v + w) = A(v + w) = Av + Aw = F(v) + F(w)$$

and

$$F(kv) = A(kv) = k(Av) = kF(v)$$

Thus, F is linear.

Example

. Show that the following mappings are not linear:

(a) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (xy, x)$

(b) $F : \mathbf{R}^2 \to \mathbf{R}^3$ defined by $F(x, y) = (x + 3,\ 2y,\ x + y)$

(c) $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (|x|,\ y + z)$

(a) Let $v = (1, 2)$ and $w = (3, 4)$; then $v + w = (4, 6)$. Also,

$$F(v) = (1(2), 1) = (2, 1) \qquad \text{and} \qquad F(w) = (3(4), 3) = (12, 3)$$

Hence,

$$F(v + w) = (4(6), 4) = (24, 6) \neq F(v) + F(w)$$

(b) Because $F(0, 0) = (3, 0, 0) \neq (0, 0, 0)$, F cannot be linear.

(c) Let $v = (1, 2, 3)$ and $k = -3$. Then $kv = (-3, -6, -9)$. We have

$$F(v) = (1, 5) \quad \text{and} \quad kF(v) = -3(1, 5) = (-3, -15).$$

Thus,

$$F(kv) = F(-3, -6, -9) = (3, -15) \neq kF(v)$$

Accordingly, F is not linear.

Examples (Zero and Identity Transformations)

(a) Let $F : V \to U$ be the mapping that assigns the zero vector $0 \in U$ to every vector $v \in V$. Then, for any vectors $v, w \in V$ and any scalar $k \in K$, we have

$$F(v + w) = 0 = 0 + 0 = F(v) + F(w) \qquad \text{and} \qquad F(kv) = 0 = k0 = kF(v)$$

Thus, F is linear. We call F the *zero mapping*, and we usually denote it by 0.

(b) Consider the identity mapping $I : V \to V$, which maps each $v \in V$ into itself. Then, for any vectors $v, w \in V$ and any scalars $a, b \in K$, we have

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Thus, I is linear.

Theorem

Suppose a linear mapping $F: V \to U$ is one-to-one and onto. Show that the inverse mapping $F^{-1}: U \to V$ is also linear.

Suppose $u, u' \in U$. Because F is one-to-one and onto, there exist unique vectors $v, v' \in V$ for which $F(v) = u$ and $F(v') = u'$. Because F is linear, we also have

$$F(v + v') = F(v) + F(v') = u + u' \qquad \text{and} \qquad F(kv) = kF(v) = ku$$

By definition of the inverse mapping,

$$F^{-1}(u) = v, \; F^{-1}(u') = v', \; F^{-1}(u + u') = v + v', \; F^{-1}(ku) = kv.$$

Then

$$F^{-1}(u + u') = v + v' = F^{-1}(u) + F^{-1}(u') \qquad \text{and} \qquad F^{-1}(ku) = kv = kF^{-1}(u)$$

Thus, F^{-1} is linear.

Vector space Isomorphism

DEFINITION: Two vector spaces V and U over K are *isomorphic*, written $V \cong U$, if there exists a bijective (one-to-one and onto) linear mapping $F: V \to U$. The mapping F is then called an *isomorphism* between V and U.

Consider any vector space V of dimension n and let S be any basis of V. Then the mapping

$$v \mapsto [v]_S$$

which maps each vector $v \in V$ into its coordinate vector $[v]_S$, is an isomorphism between V and K^n.

Projection Operators

Consider the operator $T: R^2 \longrightarrow R^2$ that maps each vector into its orthogonal projection on the x-axis (Figure 4.2.3). The equations relating the components of x and $\mathbf{w} = T(\mathbf{x})$ are

$$\begin{aligned} w_1 &= x = x + 0y \\ w_2 &= 0 = 0x + 0y \end{aligned} \qquad (12)$$

or, in matrix form,

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (13)$$

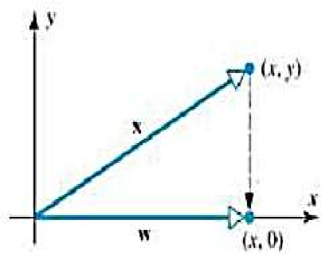


Figure 4.2.3

The equations in 12 are linear, so T is a linear operator, and from 13 the standard matrix for T is

$$[T] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

In general, a ***projection operator*** (more precisely, an ***orthogonal projection operator***) on R^2 or R^3 is any operator that maps each vector into its orthogonal projection on a line or plane through the origin. It can be shown that such operators are linear. Some of the basic projection operators on R^2 and R^3 are listed in Tables 4 and 5.

Kernel and Image of a Linear Mapping

Let $F: V \to U$ be a linear mapping. The *kernel* of F, written Ker F, is the set of elements in V that map into the zero vector 0 in U; that is,

$$\operatorname{Ker} F = \{v \in V : F(v) = 0\}$$

The *image* (or *range*) of F, written Im F, is the set of image points in U; that is,

$$\operatorname{Im} F = \{u \in U : \text{there exists } v \in V \text{ for which } F(v) = u\}$$

Example

(a) Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the projection of a vector v into the xy-plane [as pictured in Fig. 5-2(a)]; that is,

$$F(x,y,z) = (x,y,0)$$

Clearly the image of F is the entire xy-plane—that is, points of the form $(x,y,0)$. Moreover, the kernel of F is the z-axis—that is, points of the form $(0,0,c)$. That is,

$$\operatorname{Im} F = \{(a,b,c) : c = 0\} = xy\text{-plane} \qquad \text{and} \qquad \operatorname{Ker} F = \{(a,b,c) : a = 0, b = 0\} = z\text{-axis}$$

(b) Let $G: \mathbf{R}^3 \to \mathbf{R}^3$ be the linear mapping that rotates a vector v about the z-axis through an angle θ [as pictured in Fig. 5-2(b)]; that is,

$$G(x,y,z) = (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta,\ z)$$

Observe that the distance of a vector v from the origin O does not change under the rotation, and so only the zero vector 0 is mapped into the zero vector 0. Thus, $\operatorname{Ker} G = \{0\}$. On the other hand, every vector u in $\mathbf{R}^3$ is the image of a vector v in $\mathbf{R}^3$ that can be obtained by rotating u back by an angle of θ. Thus, $\operatorname{Im} G = \mathbf{R}^3$, the entire space.

Theorem

Let $F: V \to U$ be a linear mapping. Then the kernel of F is a subspace of V and the image of F is a subspace of U.

Proof

Now suppose that $v_1, v_2, \ldots, v_m$ span a vector space V and that $F: V \to U$ is linear. We show that $F(v_1), F(v_2), \ldots, F(v_m)$ span $\operatorname{Im} F$. Let $u \in \operatorname{Im} F$. Then there exists $v \in V$ such that $F(v) = u$. Because the v_i's span V and $v \in V$, there exist scalars $a_1, a_2, \ldots, a_m$ for which

$$v = a_1v_1 + a_2v_2 + \cdots + a_mv_m$$

Therefore,

$$u = F(v) = F(a_1v_1 + a_2v_2 + \cdots + a_mv_m) = a_1F(v_1) + a_2F(v_2) + \cdots + a_mF(v_m)$$

Thus, the vectors $F(v_1), F(v_2), \ldots, F(v_m)$ span $\operatorname{Im} F$.

Example

Let $F : \mathbf{R}^4 \to \mathbf{R}^3$ be the linear mapping defined by

$$F(x, y, z, t) = (x - y + z + t, \quad x + 2z - t, \quad x + y + 3z - 3t)$$

Find a basis and the dimension of (a) the image of F, (b) the kernel of F.

(a) Find the images of the usual basis of $\mathbf{R}^4$:

$$F(1,0,0,0) = (1,1,1), \qquad F(0,0,1,0) = (1,2,3)$$
$$F(0,1,0,0) = (-1,0,1), \qquad F(0,0,0,1) = (1,-1,-3)$$

By Proposition 5.4, the image vectors span Im F. Hence, form the matrix whose rows are these image vectors, and row reduce to echelon form:

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

Thus, $(1,1,1)$ and $(0,1,2)$ form a basis for Im F; hence, $\dim(\text{Im } F) = 2$.

(b) Set $F(v) = 0$, where $v = (x, y, z, t)$; that is, set

$$F(x, y, z, t) = (x - y + z + t, \ x + 2z - t, \ x + y + 3z - 3t) = (0, 0, 0)$$

Set corresponding entries equal to each other to form the following homogeneous system whose solution space is Ker F:

$$\begin{aligned} x - y + z + t &= 0 \\ x \quad + 2z - t &= 0 \\ x + y + 3z - 3t &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x - y + z + t &= 0 \\ y + z - 2t &= 0 \\ 2y + 2z - 4t &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x - y + z + t &= 0 \\ y + z - 2t &= 0 \end{aligned}$$

The free variables are z and t. Hence, $\dim(\text{Ker } F) = 2$.

(i) Set $z = -1$, $t = 0$ to obtain the solution $(2, 1, -1, 0)$.

(ii) Set $z = 0$, $t = 1$ to obtain the solution $(1, 2, 0, 1)$.

Thus, $(2, 1, -1, 0)$ and $(1, 2, 0, 1)$ form a basis of Ker F.
[As expected, $\dim(\text{Im } F) + \dim(\text{Ker } F) = 2 + 2 = 4 = \dim \mathbf{R}^4$, the domain of F.]

Difference between transformation and operator

If the domain of a function f is R^n and the codomain is R^m (m and n possibly the same), then f is called a **map or transformation** from R^n to R^m, and we say that the function f maps R^n into R^m. We denote this by writing $f: R^n \to R^m$. When $m = n$ the transformation $f: R^n \to R^m$ is called an **operator** on R^n.

Example

The equations

$$w_1 = x_1 + x_2$$
$$w_2 = 3x_1x_2$$
$$w_3 = x_1^2 - x_2^2$$

define a transformation $T: R^2 \longrightarrow R^3$. With this transformation, the image of the point (x_1, x_2) is

$$T(x_1, x_2) = (x_1 + x_2, 3x_1x_2, x_1^2 - x_2^2)$$

Thus, for example, $T(1, -2) = (-1, -6, -3)$

Rank and Nullity of a Linear Mapping

Let $F : V \to U$ be a linear mapping. The *rank* of F is defined to be the dimension of its image, and the *nullity* of F is defined to be the dimension of its kernel; namely,

$$\text{rank}(F) = \dim(\text{Im } F) \quad \text{and} \quad \text{nullity}(F) = \dim(\text{Ker } F)$$

Theorem

Let V be of finite dimension, and let $F : V \to U$ be linear. Then

$$\dim V = \dim(\text{Ker } F) + \dim(\text{Im } F) = \text{nullity}(F) + \text{rank}(F)$$

Suppose $\dim(\text{Ker } F) = r$ and $\{w_1, \ldots, w_r\}$ is a basis of Ker F, and suppose $\dim(\text{Im } F) = s$ and $\{u_1, \ldots, u_s\}$ is a basis of Im F. (By Proposition 5.4, Im F has finite dimension.) Because every $u_j \in \text{Im } F$, there exist vectors $v_1, \ldots, v_s$ in V such that $F(v_1) = u_1, \ldots, F(v_s) = u_s$. We claim that the set

$$B = \{w_1, \ldots, w_r, v_1, \ldots, v_s\}$$

is a basis of V; that is, (i) B spans V, and (ii) B is linearly independent. Once we prove (i) and (ii), then $\dim V = r + s = \dim(\text{Ker } F) + \dim(\text{Im } F)$.

(i) *B spans V.* Let $v \in V$. Then $F(v) \in \text{Im } F$. Because the u_j span Im F, there exist scalars $a_1, \ldots, a_s$ such that $F(v) = a_1u_1 + \cdots + a_su_s$. Set $\hat{v} = a_1v_1 + \cdots + a_sv_s - v$. Then

$$F(\hat{v}) = F(a_1v_1 + \cdots + a_sv_s - v) = a_1F(v_1) + \cdots + a_sF(v_s) - F(v)$$
$$= a_1u_1 + \cdots + a_su_s - F(v) = 0$$

Thus, $\hat{v} \in \text{Ker } F$. Because the w_i span Ker F, there exist scalars $b_1, \ldots, b_r$, such that

$$\hat{v} = b_1w_1 + \cdots + b_rw_r = a_1v_1 + \cdots + a_sv_s - v$$

Accordingly,

$$v = a_1v_1 + \cdots + a_sv_s - b_1w_1 - \cdots - b_rw_r$$

Thus, B spans V.

(ii) *B is linearly independent.* Suppose

$$x_1w_1 + \cdots + x_rw_r + y_1v_1 + \cdots + y_sv_s = 0 \qquad (1)$$

where $x_i, y_j \in K$. Then

$$\begin{aligned} 0 = F(0) &= F(x_1w_1 + \cdots + x_rw_r + y_1v_1 + \cdots + y_sv_s) \\ &= x_1F(w_1) + \cdots + x_rF(w_r) + y_1F(v_1) + \cdots + y_sF(v_s) \qquad (2) \end{aligned}$$

But $F(w_i) = 0$, since $w_i \in \operatorname{Ker} F$, and $F(v_j) = u_j$. Substituting into (2), we will obtain $y_1u_1 + \cdots + y_su_s = 0$. Since the u_j are linearly independent, each $y_j = 0$. Substitution into (1) gives $x_1w_1 + \cdots + x_rw_r = 0$. Since the w_i are linearly independent, each $x_i = 0$. Thus B is linearly independent.

EXAMPLE 5.9 Let $F : \mathbf{R}^4 \to \mathbf{R}^3$ be the linear mapping defined by

$$F(x, y, z, t) = (x - y + z + t, \quad 2x - 2y + 3z + 4t, \quad 3x - 3y + 4z + 5t)$$

(a) Find a basis and the dimension of the image of F.

First find the image of the usual basis vectors of $\mathbf{R}^4$,

$$F(1, 0, 0, 0) = (1, 2, 3), \qquad F(0, 0, 1, 0) = (1, 3, 4)$$
$$F(0, 1, 0, 0) = (-1, -2, -3), \qquad F(0, 0, 0, 1) = (1, 4, 5)$$

By Proposition 5.4, the image vectors span Im F. Hence, form the matrix M whose rows are these image vectors and row reduce to echelon form:

$$M = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 1 & 3 & 4 \\ 1 & 4 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $(1, 2, 3)$ and $(0, 1, 1)$ form a basis of Im F. Hence, $\dim(\text{Im } F) = 2$ and $\text{rank}(F) = 2$.

(b) Find a basis and the dimension of the kernel of the map F.

Set $F(v) = 0$, where $v = (x, y, z, t)$,

$$F(x, y, z, t) = (x - y + z + t, \quad 2x - 2y + 3z + 4t, \quad 3x - 3y + 4z + 5t) = (0, 0, 0)$$

Set corresponding components equal to each other to form the following homogeneous system whose solution space is Ker F:

$$\begin{aligned} x - y + z + t &= 0 \\ 2x - 2y + 3z + 4t &= 0 \\ 3x - 3y + 4z + 5t &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + z + t &= 0 \\ z + 2t &= 0 \\ z + 2t &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + z + t &= 0 \\ z + 2t &= 0 \end{aligned}$$

The free variables are y and t. Hence, $\dim(\text{Ker } F) = 2$ or $\text{nullity}(F) = 2$.

(i) Set $y = 1$, $t = 0$ to obtain the solution $(-1, 1, 0, 0)$,

(ii) Set $y = 0$, $t = 1$ to obtain the solution $(1, 0, -2, 1)$.

Thus, $(-1, 1, 0, 0)$ and $(1, 0, -2, 1)$ form a basis for Ker F.

As expected from Theorem 5.6, $\dim(\text{Im } F) + \dim(\text{Ker } F) = 4 = \dim \mathbf{R}^4$.

Singular and Nonsingular Linear Mappings, Isomorphisms

Let $F: V \to U$ be a linear mapping. Recall that $F(0) = 0$. F is said to be *singular* if the image of some nonzero vector v is 0—that is, if there exists $v \neq 0$ such that $F(v) = 0$. Thus, $F: V \to U$ is *nonsingular* if the zero vector 0 is the only vector whose image under F is 0 or, in other words, if $\text{Ker } F = \{0\}$.

5.24. Determine whether or not each of the following linear maps is nonsingular. If not, find a nonzero vector v whose image is 0.

(a) $F: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (x - y,\ x - 2y)$.

(b) $G: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $G(x, y) = (2x - 4y,\ 3x - 6y)$.

(a) Find $\text{Ker } F$ by setting $F(v) = 0$, where $v = (x, y)$,

$$(x - y,\ x - 2y) = (0, 0) \quad \text{or} \quad \begin{aligned} x - y &= 0 \\ x - 2y &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x - y &= 0 \\ -y &= 0 \end{aligned}$$

The only solution is $x = 0$, $y = 0$. Hence, F is nonsingular.

(b) Set $G(x, y) = (0, 0)$ to find $\text{Ker } G$:

$$(2x - 4y,\ 3x - 6y) = (0, 0) \quad \text{or} \quad \begin{aligned} 2x - 4y &= 0 \\ 3x - 6y &= 0 \end{aligned} \quad \text{or} \quad x - 2y = 0$$

The system has nonzero solutions, because y is a free variable. Hence, G is singular. Let $y = 1$ to obtain the solution $v = (2, 1)$, which is a nonzero vector, such that $G(v) = 0$.

5.26. Let $G: \mathbf{R}^2 \to \mathbf{R}^3$ be defined by $G(x, y) = (x + y,\ x - 2y,\ 3x + y)$.

(a) Show that G is nonsingular. (b) Find a formula for G^{-1}.

(a) Set $G(x, y) = (0, 0, 0)$ to find $\text{Ker } G$. We have

$$(x + y,\ x - 2y,\ 3x + y) = (0, 0, 0) \quad \text{or} \quad x + y = 0,\ x - 2y = 0,\ 3x + y = 0$$

The only solution is $x = 0$, $y = 0$; hence, G is nonsingular.

(b) Although G is nonsingular, it is not invertible, because $\mathbf{R}^2$ and $\mathbf{R}^3$ have different dimensions. (Thus, Theorem 5.9 does not apply.) Accordingly, G^{-1} does not exist.

Theorem

Prove Theorem 5.7: Let $F : V \to U$ be a nonsingular linear mapping. Then the image of any linearly independent set is linearly independent.

Suppose $v_1, v_2, \ldots, v_n$ are linearly independent vectors in V. We claim that $F(v_1), F(v_2), \ldots, F(v_n)$ are also linearly independent. Suppose $a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n) = 0$, where $a_i \in K$. Because F is linear, $F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = 0$. Hence,

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n \in \operatorname{Ker} F$$

But F is nonsingular—that is, $\operatorname{Ker} F = \{0\}$. Hence, $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Because the v_i are linearly independent, all the a_i are 0. Accordingly, the $F(v_i)$ are linearly independent. Thus, the theorem is proved.

Theorem

Prove Theorem 5.9: Suppose V has finite dimension and $\dim V = \dim U$. Suppose $F : V \to U$ is linear. Then F is an isomorphism if and only if F is nonsingular.

If F is an isomorphism, then only 0 maps to 0; hence, F is nonsingular. Conversely, suppose F is nonsingular. Then $\dim(\operatorname{Ker} F) = 0$. By Theorem 5.6, $\dim V = \dim(\operatorname{Ker} F) + \dim(\operatorname{Im} F)$. Thus,

$$\dim U = \dim V = \dim(\operatorname{Im} F)$$

Because U has finite dimension, $\operatorname{Im} F = U$. This means F maps V onto U. Thus, F is one-to-one and onto; that is, F is an isomorphism.

حرفِ آخر (08-10-2022)

خوش رہیں خوشیاں بانٹیں اور جہاں تک ہو سکے دوسروں کے لیے آسانیاں پیدا کریں۔

اللہ تعالیٰ آپ کو زندگی کے ہر موڑ پر کامیابیوں اور خوشیوں سے نوازے۔ (آمین)

محمد عثمان حامد

چک نمبر 105 شمالی (گودھے والا) سرگودھا

UNIVERSITY OF SARGODHA

PUNJAB, PAKISTAN