

**SCHOOL OF COMPUTING**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**SCSA3016 DATA SCIENCE**

# COURSE OBJECTIVES

* To understand the mathematical foundations required for data science.
* To describe a flow process for data science problems.
* To introduce basic data science algorithms and data visualization.
* To learn machine tools and techniques.
* To learn the ideas and tools for data visualization.

# UNIT 1 LINEARALGEBRA 9 Hrs.

Algebraic view – vectors 2D, 3D and nD, matrices, product of matrix & vector, rank, null space, solution of over determined set of equations and pseudo-inverse. Geometric view - vectors, distance, projections, eigenvalue decomposition, Equations of line, plane, hyperplane, circle, sphere, Hypersphere.

**UNIT 2 PROBABILITY AND STATISTICS** **9 Hrs.**

Introduction to probability and statistics, Population and sample, Normal and Gaussian distributions, Probability Density Function, Descriptive statistics, notion of probability, distributions, mean, variance, covariance, covariance matrix, understanding univariate and multivariate normal distributions, introduction to hypothesis testing, confidence interval for estimates.

**UNIT 3 EXPLORATORY DATA ANALYSIS AND THE DATA SCIENCE PROCESS 9 Hrs.**

Exploratory Data Analysis and the Data Science Process - Basic tools (plots, graphs and summary statistics) of EDA - Philosophy of EDA - The Data Science Process - Data Visualization - Basic principles, ideas and tools for data visualization

- Examples of exciting projects- Data Visualization using Tableau.

**UNIT 4 MACHINE LEARNING TOOLS, TECHNIQUES AND APPLICATIONS 9 Hrs**.

Supervised Learning, Unsupervised Learning, Reinforcement Learning, Dimensionality Reduction, Principal Component Analysis, Classification and Regression models, Tree and Bayesian network models, Neural Networks, Testing, Evaluation and Validation of Models.

**UNIT 5 INTRODUCTION TO PYTHON** 9 Hrs.

Data structures-Functions-Numpy-Matplotlib-Pandas- problems based on computational complexity-Simple case studies based on python (Binary search, common elements in list), Hash tables, Dictionary.

# Max. 45 Hrs.

**Course Outcome**

On completion of the course, student will be able to

CO1 - Explain the basic terms of Linear Algebra and Statistical Inference.

CO2 - Describe the Data Science process and how its components interact.

CO3 - Apply EDA and the Data Science process in a case study.

CO4 - Classify Data Science problems.

CO5 - Analyse and correlate the results to the solutions.

CO6 - Simulate Data Visualization in exciting projects.

**UNIT 1 LINEAR ALGEBRA**

Algebraic view – vectors 2D, 3D and nD, matrices, product of matrix & vector, rank, null space, solution of over determined set of equations and pseudo-inverse. Geometric view - vectors, distance, projections, eigenvalue decomposition, Equations of line, plane, hyperplane, circle, sphere, Hypersphere.

* 1. **Introduction to Data science**
* **Data science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data.
* It is an analyzing method to extract accurate and deep understanding of a raw data using methods in statistics, Machine Learning etc.
* Different process includes in data science are inspecting, cleaning, transforming, modeling, analyzing and interpreting raw data.

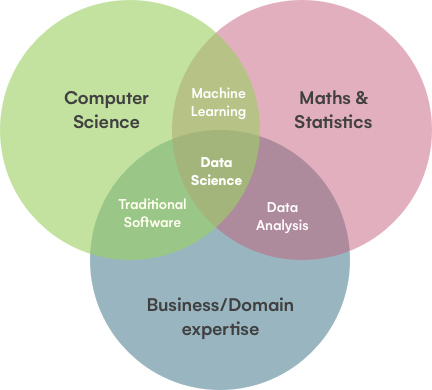


Figure 1 : Collaboration of data science with other branch of studies

**Data Science Life Cycle**

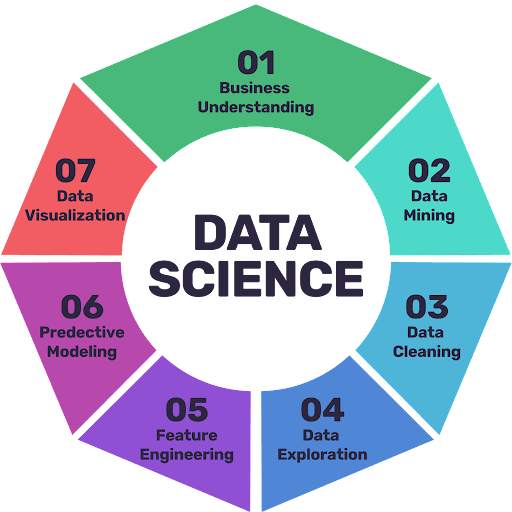


Figure 2 :Data science life cycle process

**Applications of Data Science**

* Fraud and risk detection
* Healthcare
* Airline Route Planning
* Image and speech recognition
* Augmented reality
  1. **Linear Algebra**
* Linear algebra is one of the most important mathematical and computational tools in data science.
* It is a branch of mathematics that deals with the theory of systems of linear equations, matrices, vector spaces, determinants, and linear transformations.

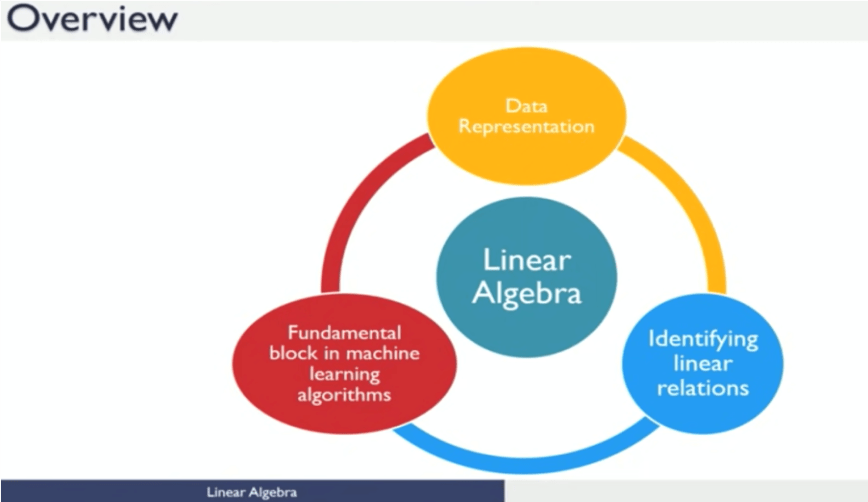


Figure 3 : Process in Linear Algebra

* 1. **Algebraic View: Vector**
* A vector is an object that has both a magnitude and a direction.
* Geometrically, we can picture a vector as a directed line segment, whose length is the magnitude of the vector and with an arrow indicating the direction. The direction of the vector is from its tail to its head.
* Vectors are used to represent numeric or symbolic characteristics, called features, of an object in a mathematical, easily analyzable way.

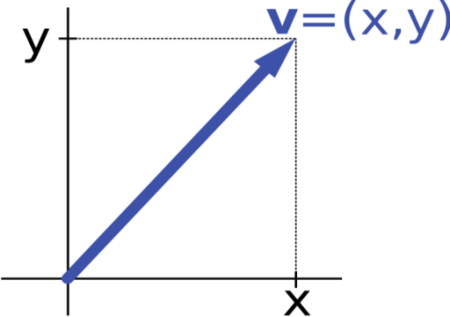
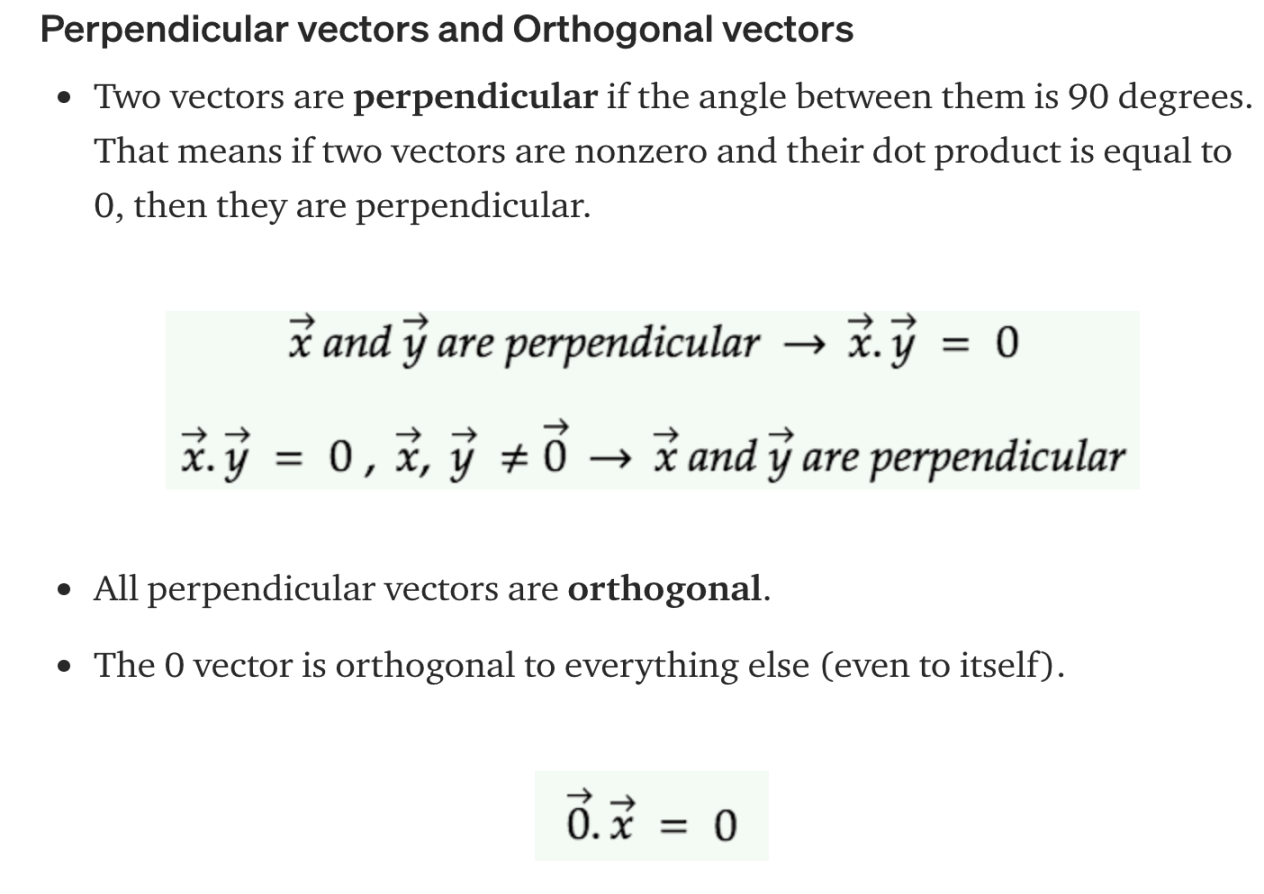
****

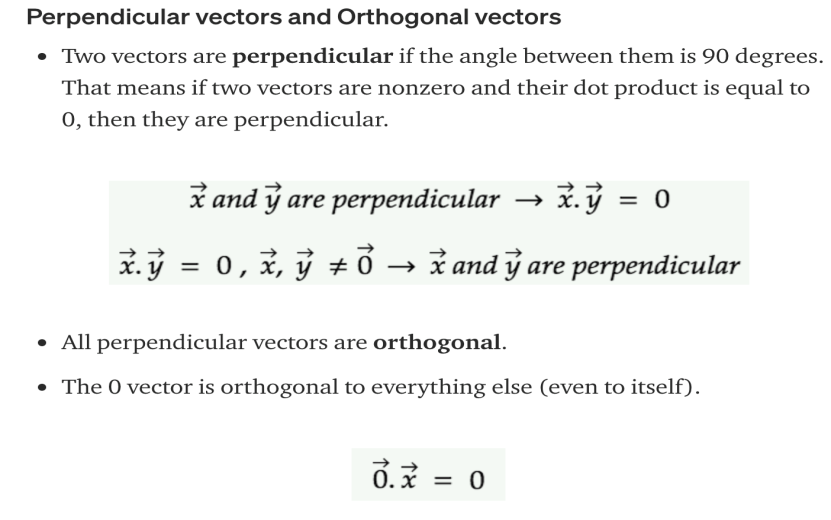
Figure 4 : Vector representation

**Perpendicular vector and orthogonal vector**

* Two vectors are perpendicular if the angle between them is 90 degrees.
* If two vectors are nonzero and their dot product is equal to 0, then they are perpendicular.



* All perpendicular vectors are orthogonal.
* The 0 vector is orthogonal to everything else (even to itself)



* 1. **Defining a 2D point/Vector:**

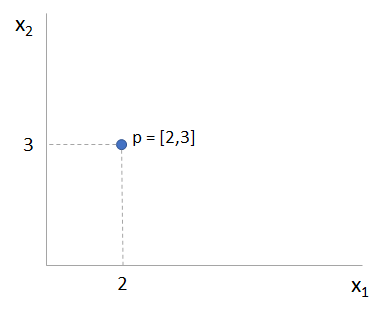
****

Figure 5: 2D vector plot

In 2D space, a point is defined as the (x,y) coordinates as shown above. Here, the x1 coordinate (x coordinate) is 2, and the x2 coordinate (y coordinate) is 3.

**Defining a 3D point/Vector:**

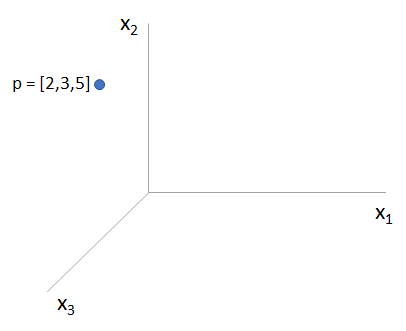
****

Figure 6 : 3D vector plot

Extending the 2D concept in 3D space point ‘p’ is defined by (x,y,z) coordinates, where 2 is x1 coordinate(x coordinate), 3 is x2 coordinate (y coordinate), and 5 is x3 coordinate (z coordinate).

**Distance of a point from Origin:**

1. **In 2D**

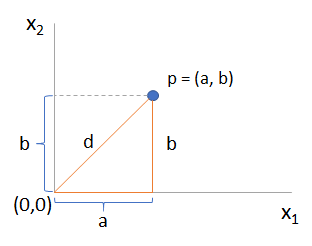
****

Figure 7 : Distance of a point from origin in 2D space

In 2D space, the distance d is given by,



1. **In 3D**

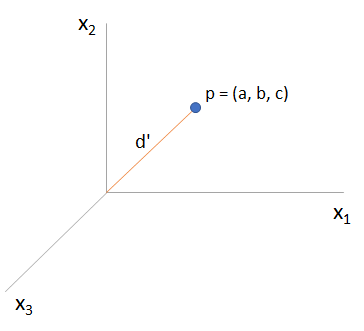
****

Figure 8 : Distance of a point from origin in 3D space

In 3D space, the distance d is given by,



1. **In nD:**

In n-Dimensional space applying Pythagoras theorem on point ‘p’ we get,



**Distance between two points**

1. In 2D

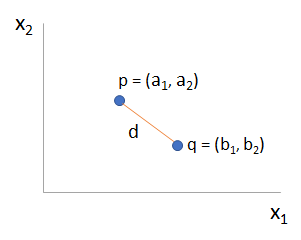
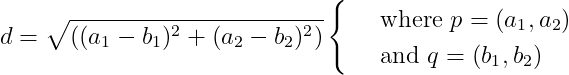


Figure 9 : Distance between two points in 2D space

Consider we have two points say, p and q then the distance d is given by,



1. In 3D

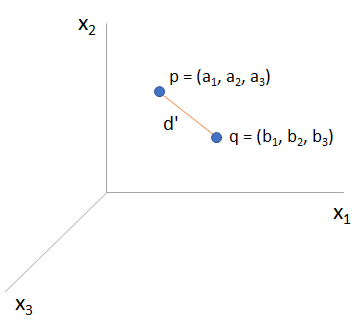
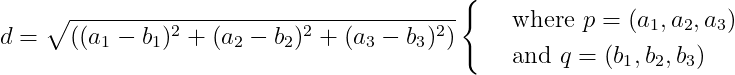


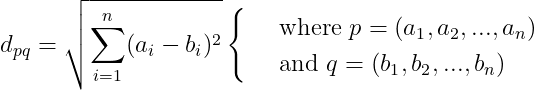
Figure 10 : Distance between two points in 3D space

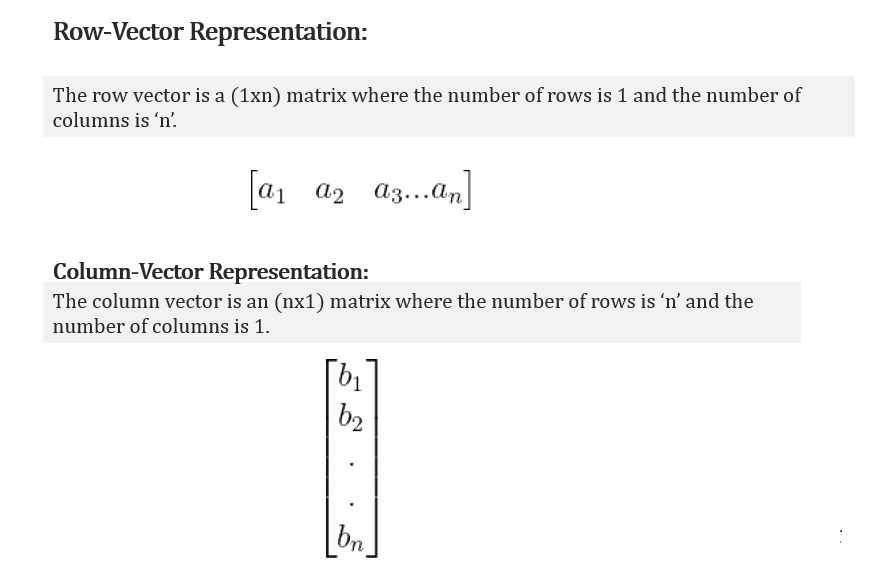
Extending the same concept in 3D space we get the distance d’ for the points p and q as follows:



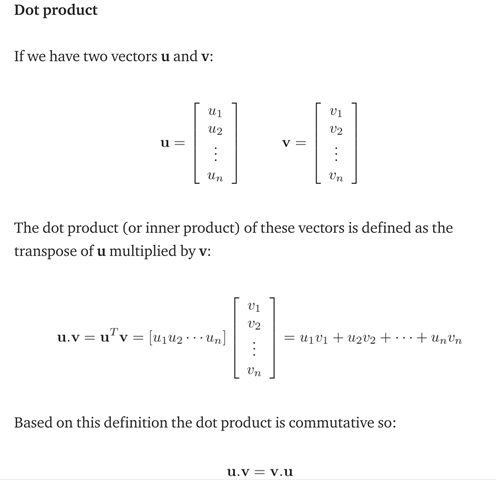
1. In nD

Extending the above concept in nD space, we get the distance formulae as,





* 1. **Dot Product of Vectors**



* 1. **Matrix Theory and Linear Algebra**
* Matrices can be used to represent samples with multiple attributes in a compact form.
* Matrices can also be used to represent linear equations in a compact and simple fashion.
* Linear algebra provides tools to understood and manipulate matrices to derive useful knowledge from data.

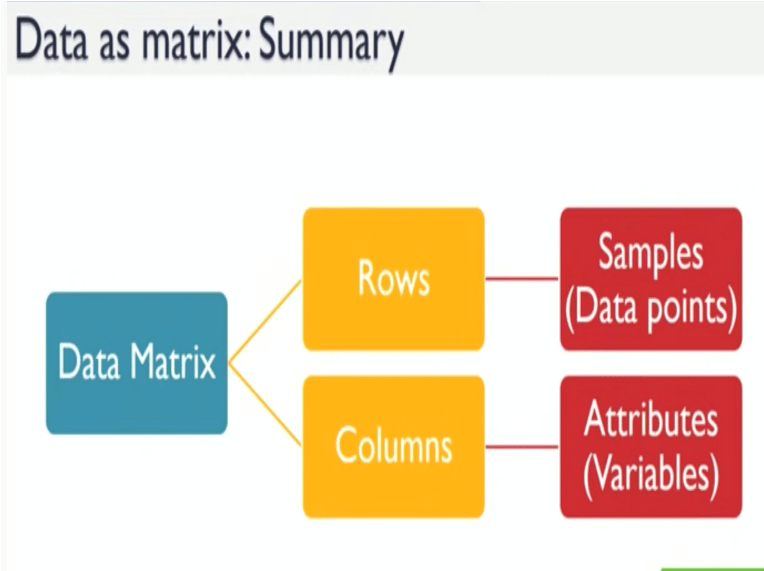
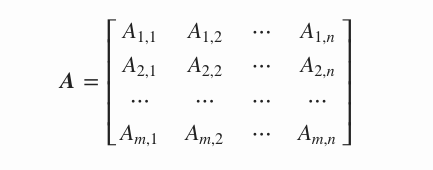


Figure 11 : Matrix elements

* A matrix is a rectangular array of numbers, symbols, or expressions, arranged in rows and columns.
* A matrix with m rows and n columns is called an m×n matrix or m-by-n matrix, where m and n are called the matrix dimensions.
* Matrices can be used to compactly write and work with multiple linear equations, that is, a system of linear equations.
* A matrix is a 2-D array of shape (m×n) with m rows and n columns is as given below:



**Tensor**

Generally, an n-dimensional array where n>2 is called a Tensor. But a matrix or a vector is also a valid tensor. A tensor is an algebraic object that describes a multilinear relationship between sets of algebraic objects related to a vector space.

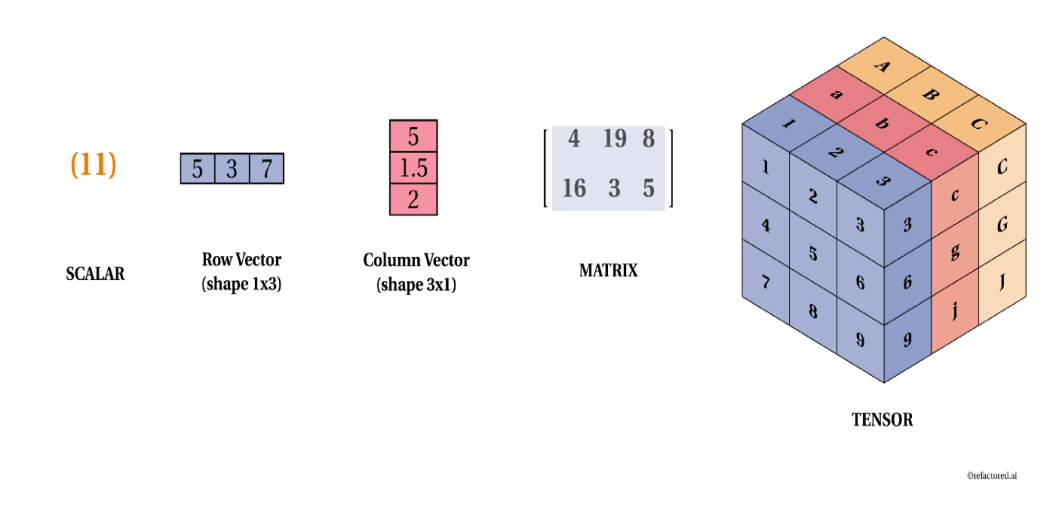


Figure 12 : Matrix and tensor

**Terms related to Matrix**

* **Order of matrix** – If a matrix has 3 rows and 4 columns, order of the matrix is 3\*4

i.e. row\*column.

* **Square matrix** – The matrix in which the number of rows is equal to the number of columns.
* **Diagonal matrix** – A matrix with all the non-diagonal elements equal to 0 is called a diagonal matrix.
* **Upper triangular matrix** – Square matrix with all the elements below diagonal equal to 0.
* **Lower triangular matrix** – Square matrix with all the elements above the diagonal equal to 0.
* **Scalar matrix** – Square matrix with all the diagonal elements equal to some constant k.
* **Identity matrix** – Square matrix with all the diagonal elements equal to 1 and all the non-diagonal elements equal to 0.
* **Column matrix** – The matrix which consists of only 1 column. Sometimes, it is used to represent a vector.
* **Row matrix** – A matrix consisting only of row.
* **Trace** – It is the sum of all the diagonal elements of a square matrix.

**Basic operations on matrix**

* **Addition** – Addition of matrices is almost similar to basic arithmetic addition.

Eg : Suppose we have 2 matrices ‘A’ and ‘B’ and the resultant matrix after the addition is ‘C’. Then

Cij = Aij + Bij

For example, let’s take two matrices and solve them.

A = 1 0

2 3

B = 4 -1

0 5 Then C = 5 -1

2 8

* **Subtraction** – Subtraction of matrices is almost similar to basic arithmetic subtraction.

Eg : Suppose we have 2 matrices ‘A’ and ‘B’ and the resultant matrix after the subtraction is ‘D’. Then

Dij = Aij - Bij

For example, let’s take two matrices and solve them.

A = 1 0

2 3

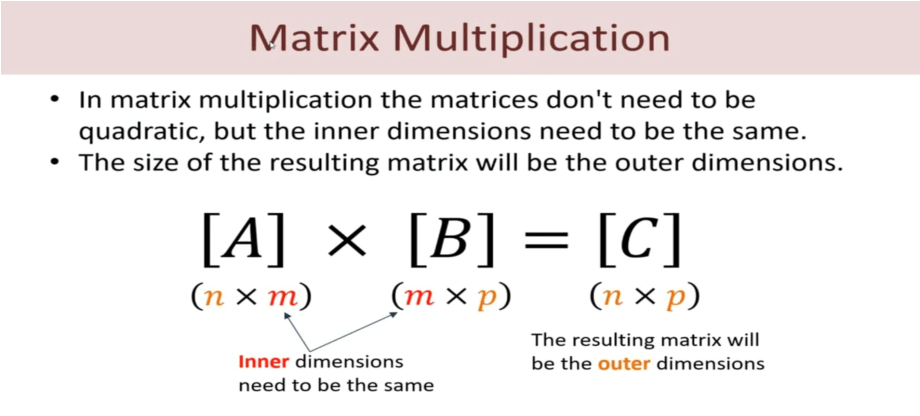
B = 4 -1

1. 5

Then C = -3 1

2 -2

* **Multiplication** – In matrix multiplication the matrices don’t need to be quadric, but the inner dimension needs to be same. The size of the resulting matrix will be the outer dimensions.



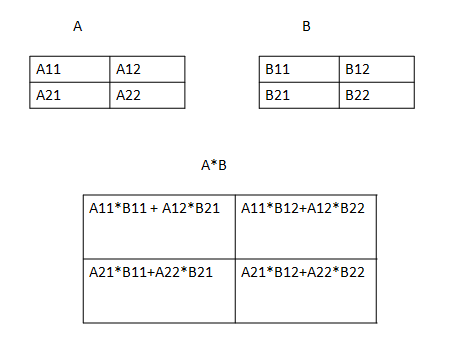
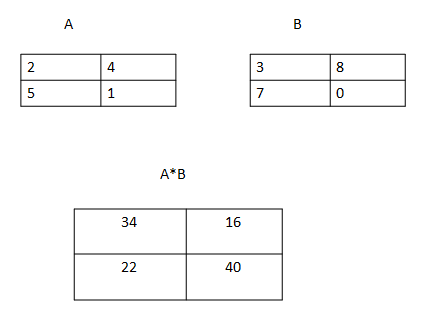


Figure 13 : Matrix multiplication

Eg :



**Applying Python on matrices**

* In python, matrix can be implemented as 2D list or 2D Array.
* Matrix operations and array are defines in module “**numpy**“.
* **add() :-** This function is used to perform element wise matrix addition.
* **subtract()** :- This function is used to perform element wise matrix subtraction.
* **multiply() :-** This function is used to perform element wise matrix multiplication.
* **dot()** :- This function is used to compute the matrix multiplication, rather than element wise multiplication.

**Applying Python on matrices – Addition and Subtraction**

# importing numpy for matrix operations

**import** numpy

# initializing matrices

x **=** numpy.array([[1, 2], [4, 5]])

y **=** numpy.array([[7, 8], [9, 10]])

# using add() to add matrices

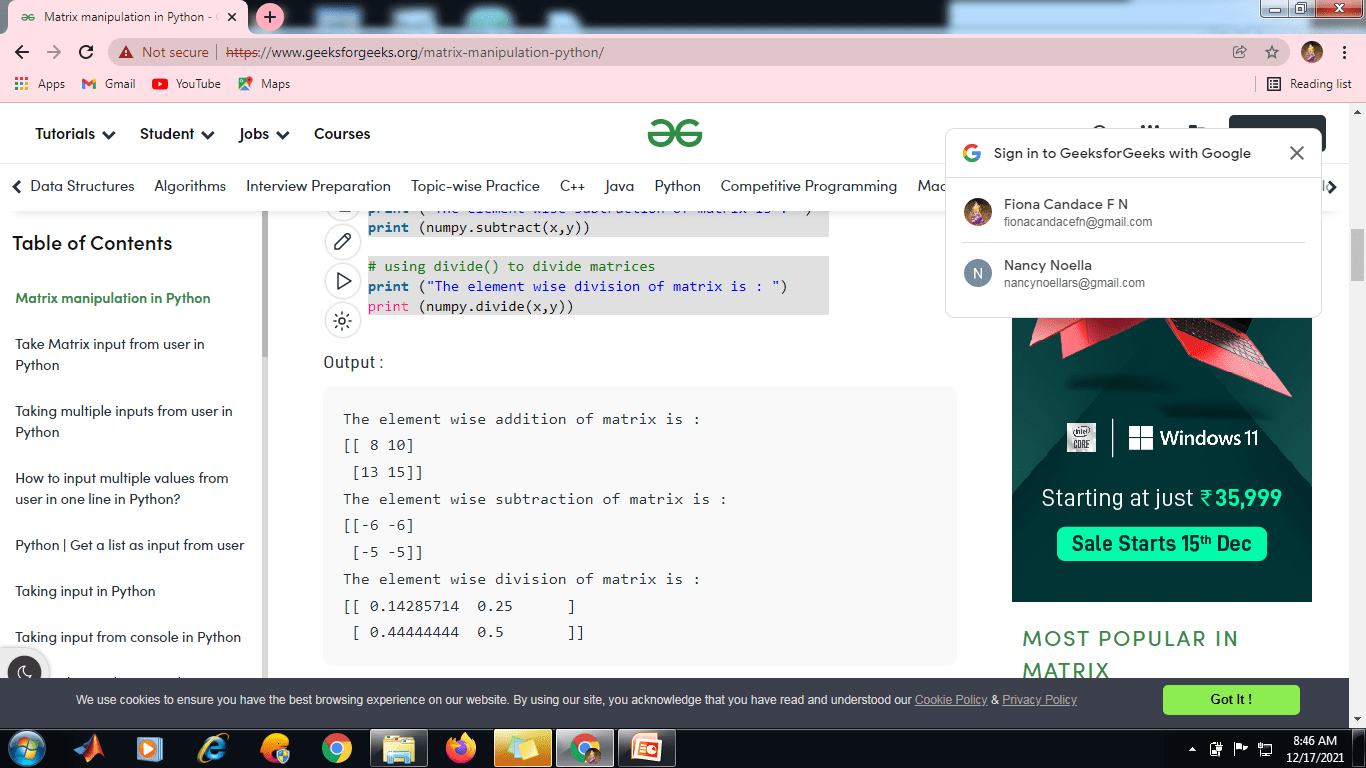
**print** ("The element wise addition of matrix is : ")

**print** (numpy.add(x,y))

# using subtract() to subtract matrices

**print** ("The element wise subtraction of matrix is : ")

**print** (numpy.subtract(x,y))



**Applying Python on matrices – Multiplication and Dot Product**

# importing numpy for matrix operations

import numpy

# initializing matrices

x = numpy.array([[1, 2], [4, 5]])

y = numpy.array([[7, 8], [9, 10]])

# using multiply() to multiply matrices element wise

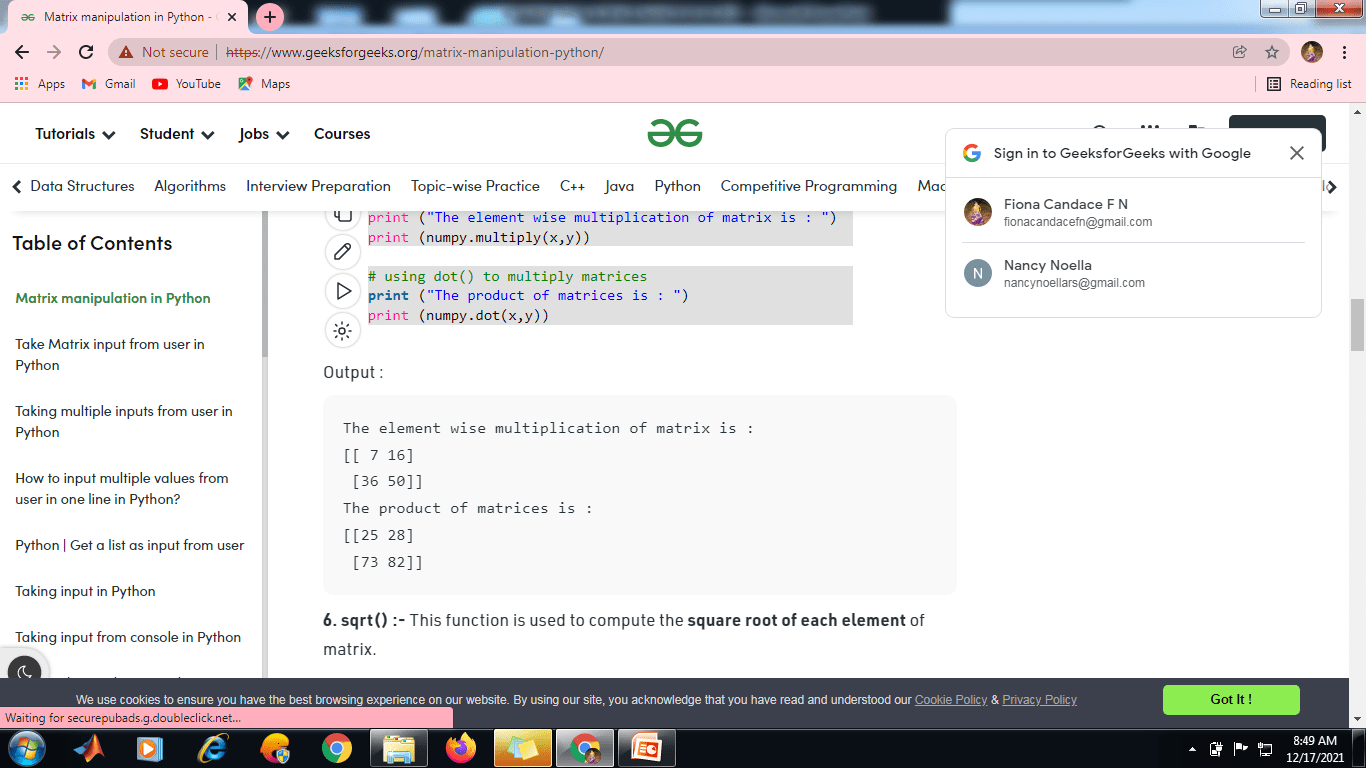
print ("The element wise multiplication of matrix is : ")

print (numpy.multiply(x,y))

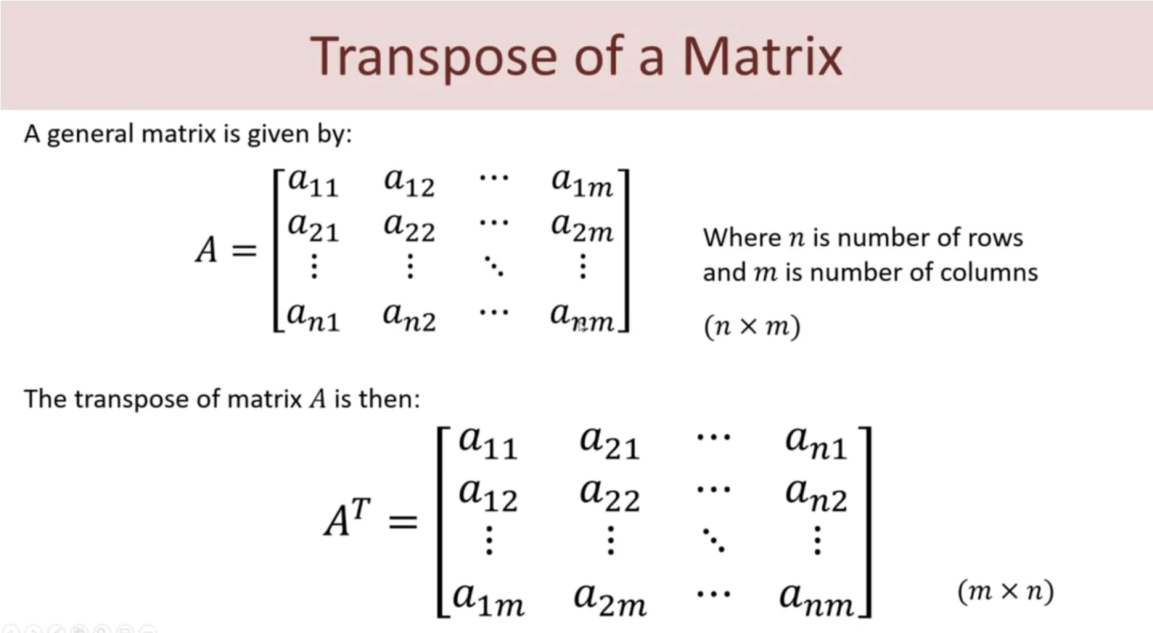
# using dot() to multiply matrices

print ("The product of matrices is : ")

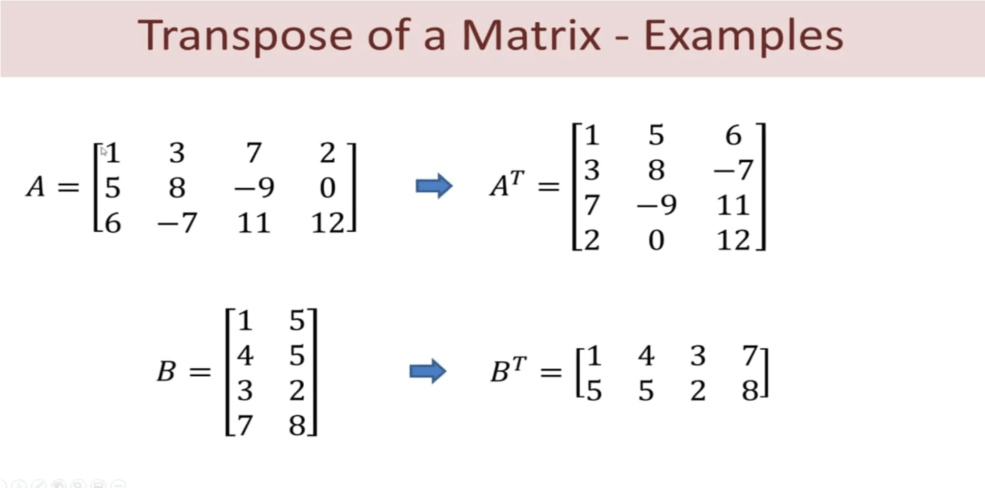
print (numpy.dot(x,y))

****

**Transpose of a matrix**

****

Eg :

****

**Applying Python on matrices – Transpose**

“T” :- This argument is used to transpose the specified matrix.

# importing numpy for matrix operations

import numpy

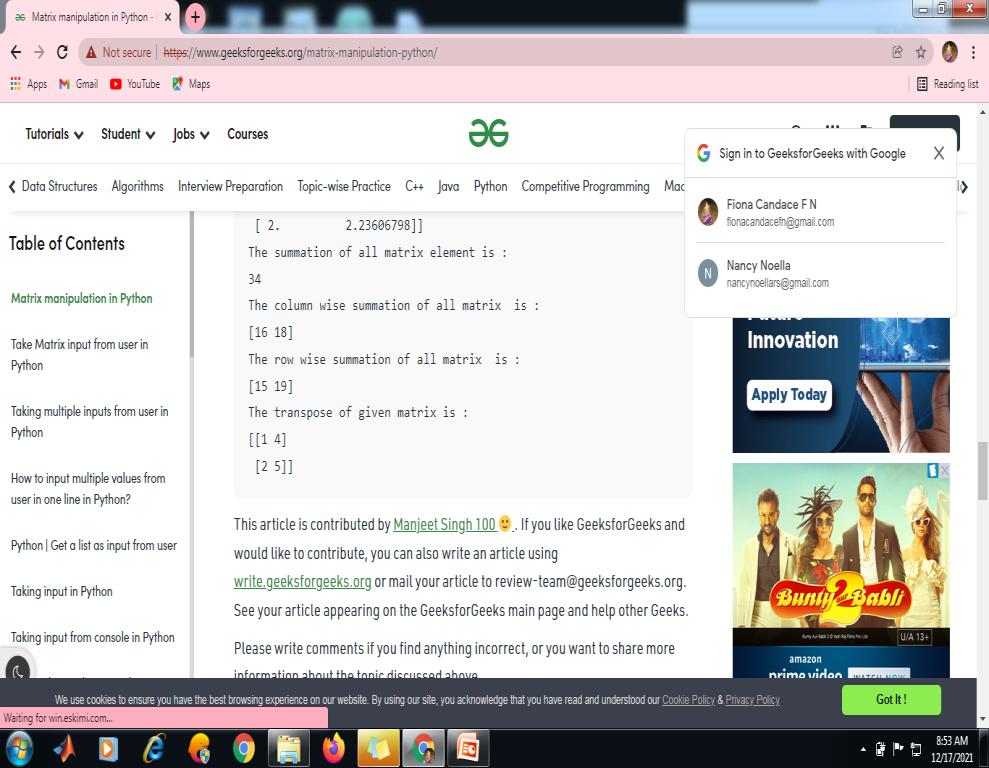
# initializing matrices

x = numpy.array([[1, 2], [4, 5]])

# using "T" to transpose the matrix

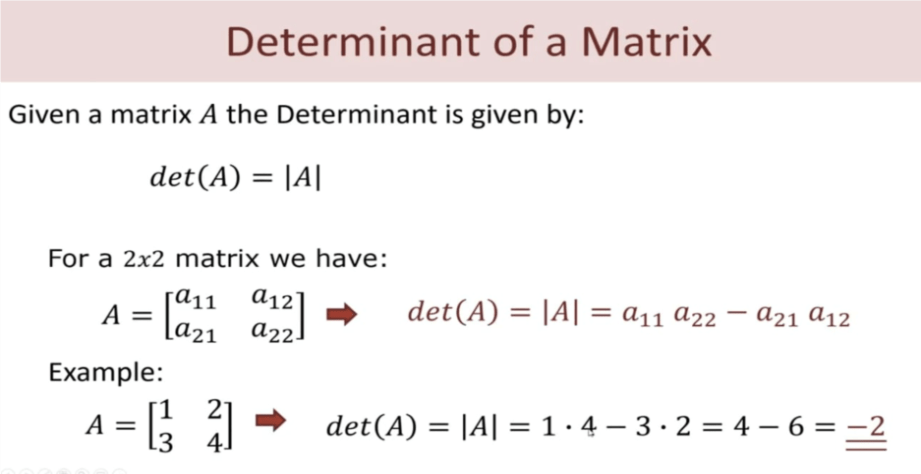
print ("The transpose of given matrix is : ")

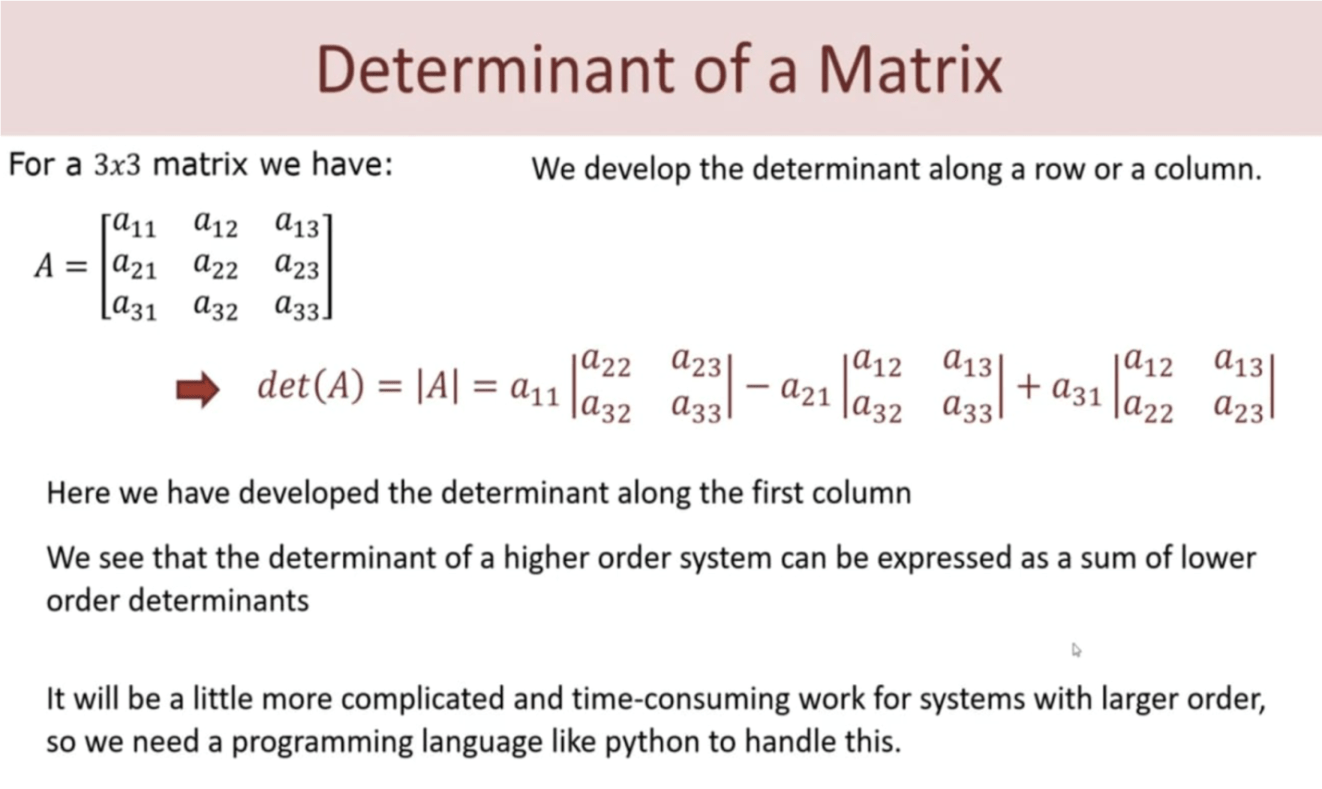
print (x.T)

****

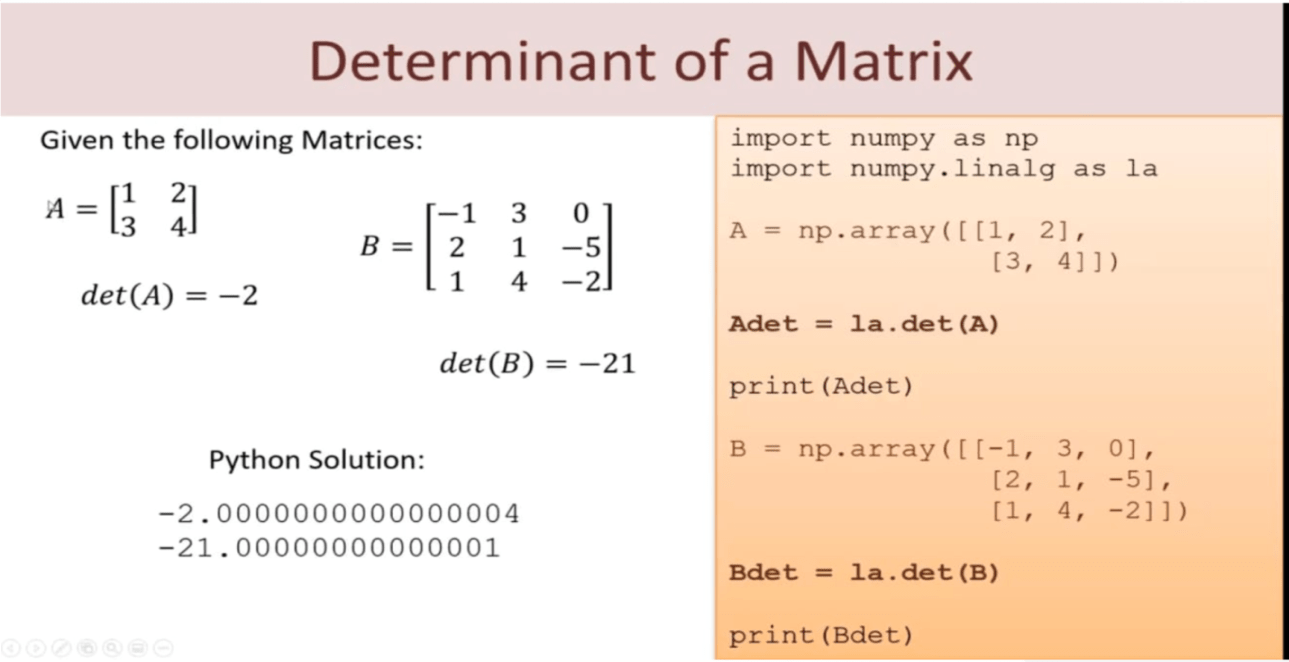
**Determinant of a matrix**

The determinant of a matrix can be calculated from square matrices.

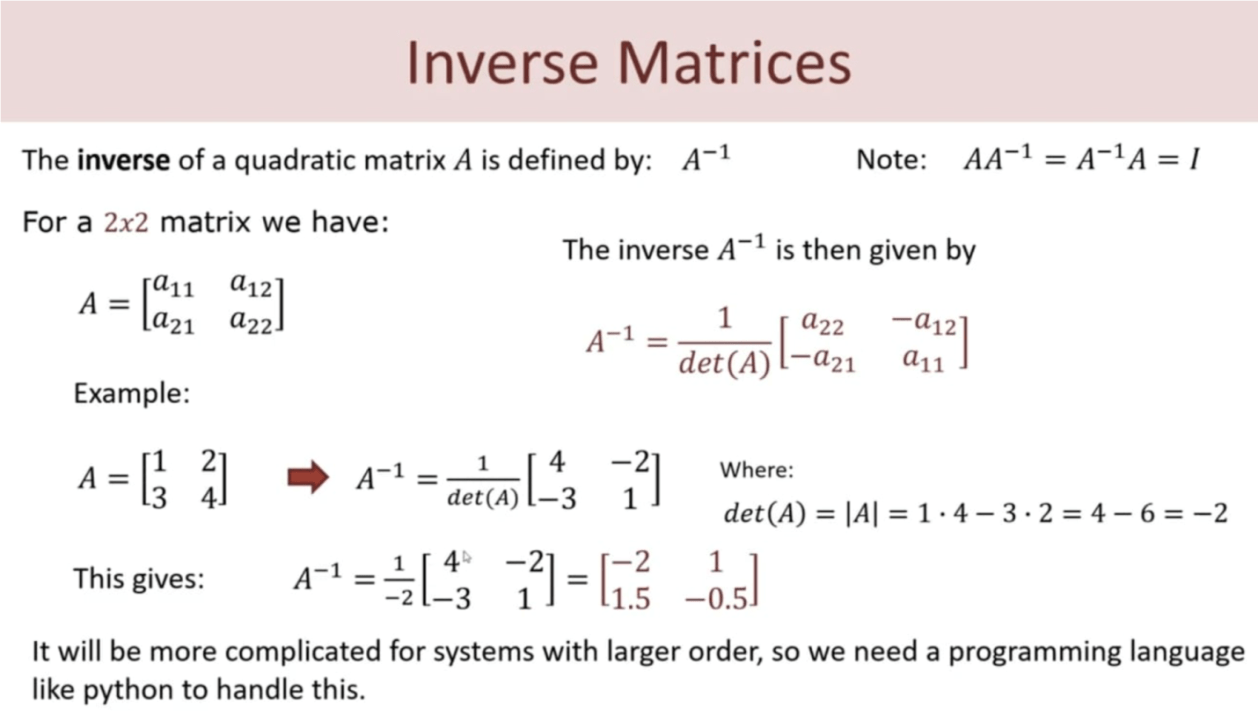
****

****

**Applying Python on matrices – Determinant**

****

**Inverse of a Matrix**

****

The formula to find the inverse of a matrix is

**A-1  = 1/det(A) \*Adj(A)**

**Applying Python on matrices – Inverse**

# Import required package

import numpy as np

import numpy.linalg as la

# Taking a 3 \* 3 matrix

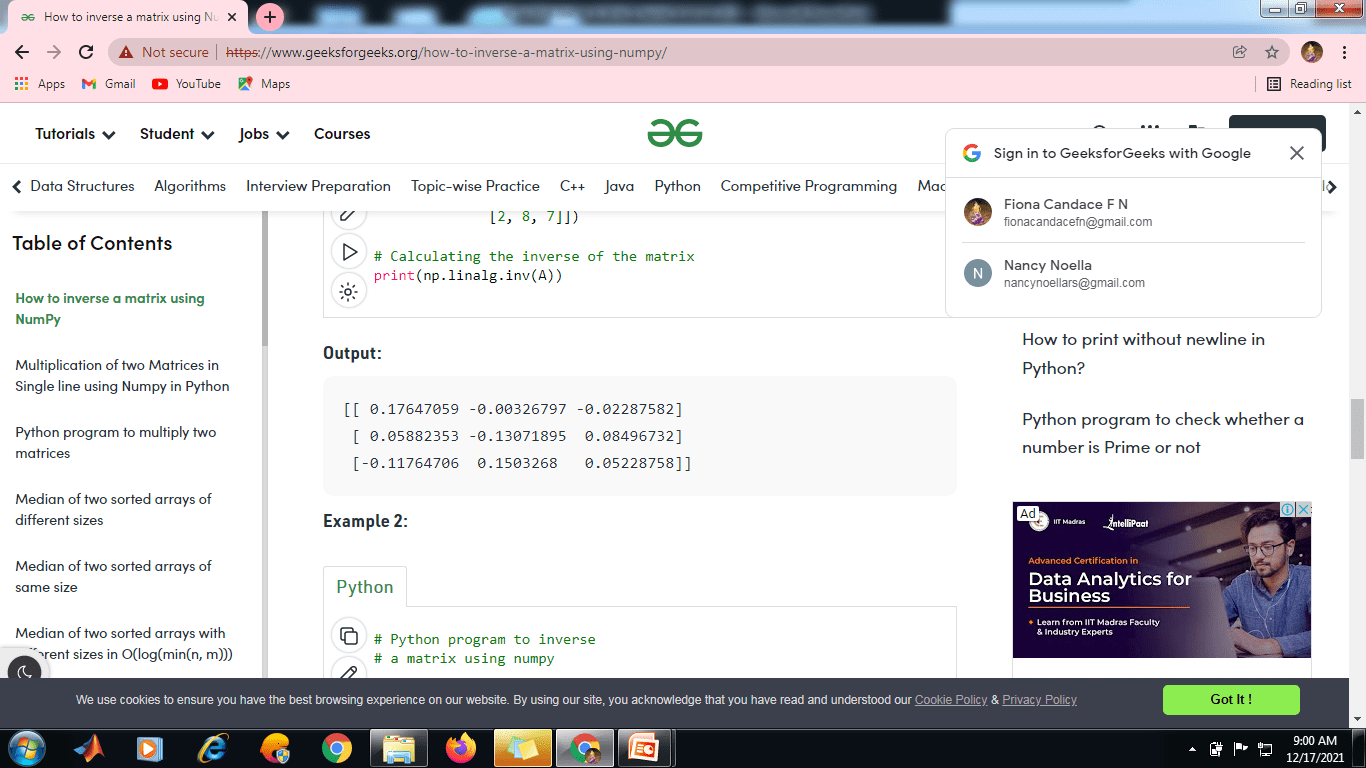
A = np.array([[6, 1, 1],

              [4, -2, 5],

              [2, 8, 7]])

# Calculating the inverse of the matrix

print(la.inv(A))



* 1. **Rank of a matrix**
* Rank of a matrix is equal to the maximum number of linearly independent row vectors in a matrix.
* A set of vectors is linearly dependent if we can express at least one of the vectors as a linear combination of remaining vectors in the set.
* Note : Rank is the number of rows with non zero vectors.
* **Rank of a matrix** – Rank of a matrix is equal to the maximum number of linearly independent row vectors in a matrix.
* A set of vectors is linearly dependent if we can express at least one of the vectors as a linear combination of remaining vectors in the set.
* **To Calculate Rank of Matrix There are Two Methods:**

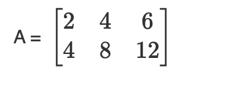
1.  Minor method

2.  Echelon form

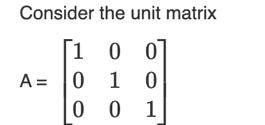
* The maximum number of linearly independent rows in a matrix *A* is called the **row rank** of *A*, and the maximum number of linearly independent columns in *A* is called the **column rank** of *A*.
* If *A* is an *m* by *n* matrix, that is, if *A* has *m* rows and *n* columns, then it is obvious that



* To find the rank of a matrix, we will transform that [matrix](https://byjus.com/jee/matrices/) into its echelon form.
* Then determine the rank by the number of non zero rows.
* Consider the following matrix.



* While observing the rows, we can see that the second row is two times the first row. Here we have two rows. But it does not count. The rank is considered as 1.

****

We can see that the rows are independent. Hence the rank of this matrix is 3.

The rank of a unit matrix of order m is m.

If A matrix is of order m×n, then ρ(A ) ≤ min{m, n } = minimum of m, n.

If A is of order n×n and |A| ≠ 0, then the rank of A = n.

If A is of order n×n and |A| = 0, then the rank of A will be less than n

**Rank of a Matrix by Row- Echelon Form**

* We can transform a given non-zero matrix to a simplified form called a Row-echelon form, using the row elementary operations. In this form, we may have rows all of whose entries are zero. Such rows are called zero rows. A non-zero row is one in which at least one of the elements is not zero.

A matrix is said to be in **row-echelon form** if the following rules are satisfied.

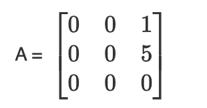
* All the leading entries in each row of the matrix is 1
* If a column contains a leading entry then all the entries below the leading entry should be zero
* If any two consecutive non-zero rows, the leading entry in the upper row should occur to the left of the leading entry in the lower row.
* All rows which consist only of zeros should occur in the bottom of the matrix

A matrix *A* of order *m* × *n* is said to be in echelon form if

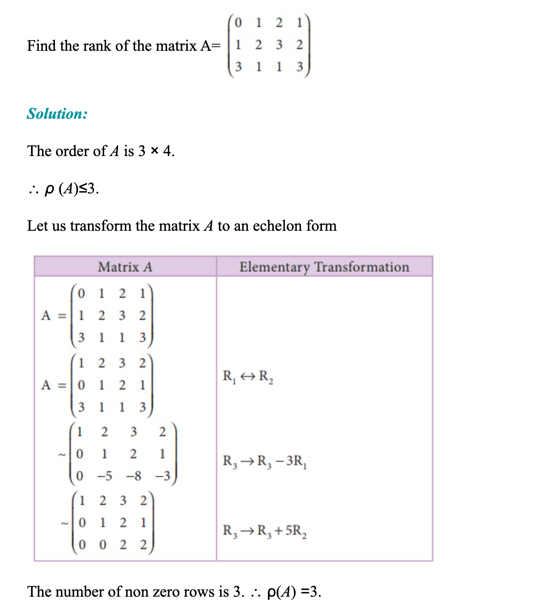
(i) Every row of *A* which has all its entries 0 occurs below every row which has a non-zero entry.

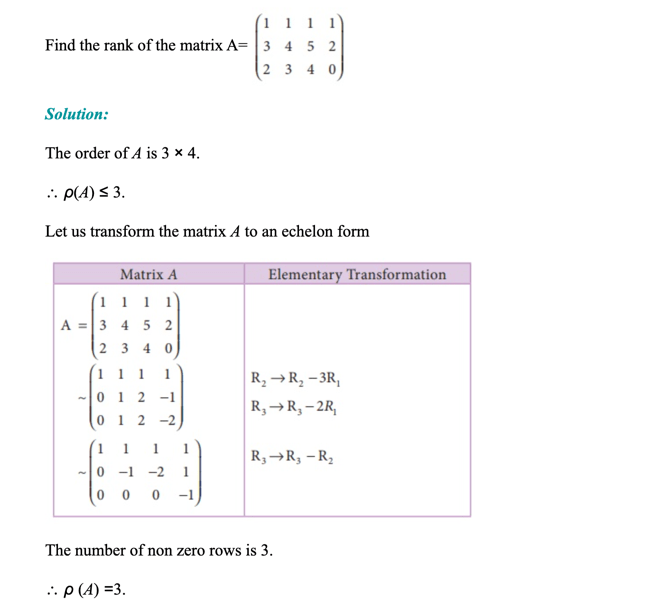
(ii) The number of zeros before the first non-zero element in a row is less then the number of such zeros in the next row.

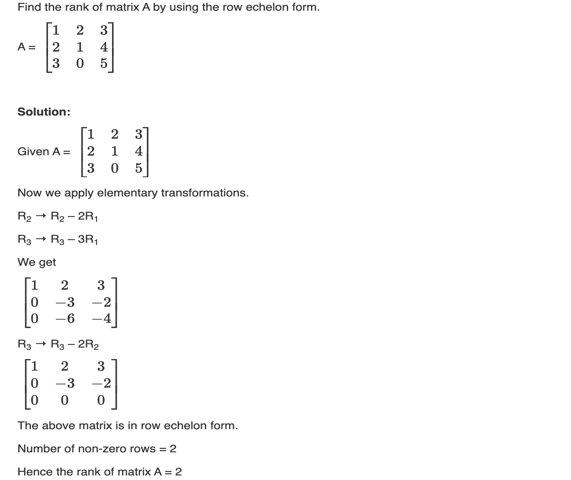
* For example, consider the following matrix.
* Here R1 and R2 are non zero rows.
* R3 is a zero row.
* Note: A non-zero matrix is said to be in a row-echelon form, if all zero rows occur as bottom rows of the matrix and if the first non-zero element in any lower row occurs to the right of the first non-zero entry in the higher row.
* If a matrix is in row-echelon form, then all elements below the leading diagonal are zeros.
* Consider the following matrix.

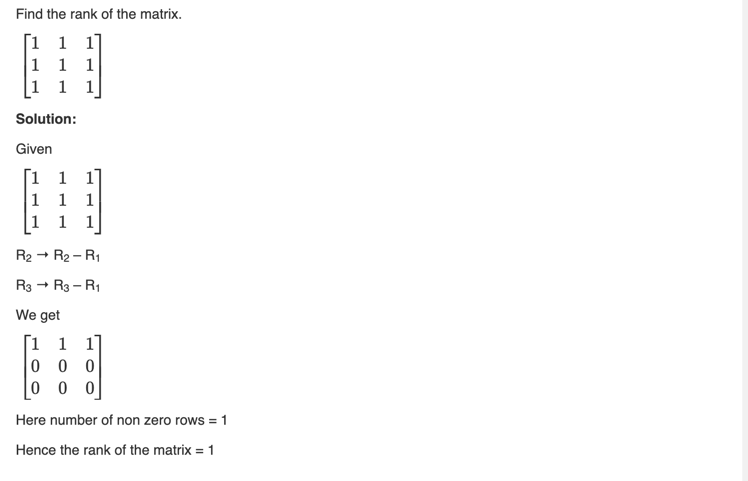


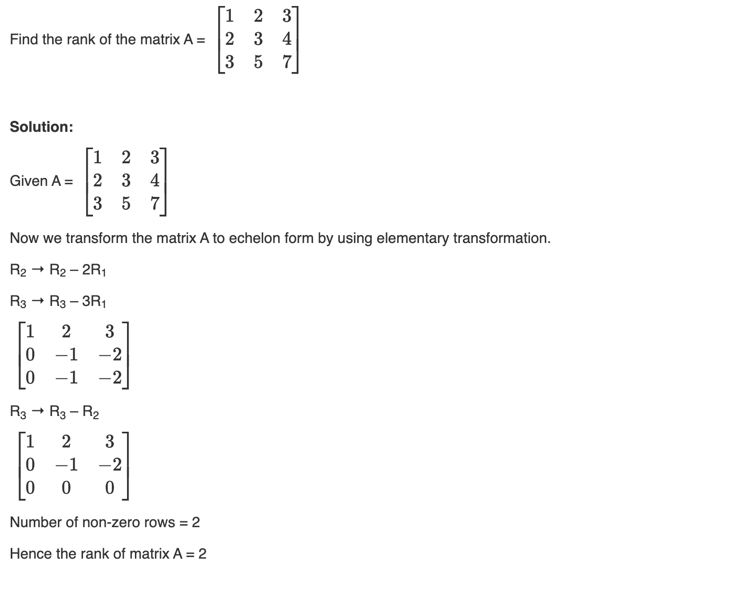
* Check the rows from the last row of the matrix. The third row is a zero row. The first non-zero element in the second row occurs in the third column and it lies to the right of the first non-zero element in the first row which occurs in the second column. Hence the matrix A is in row echelon form.

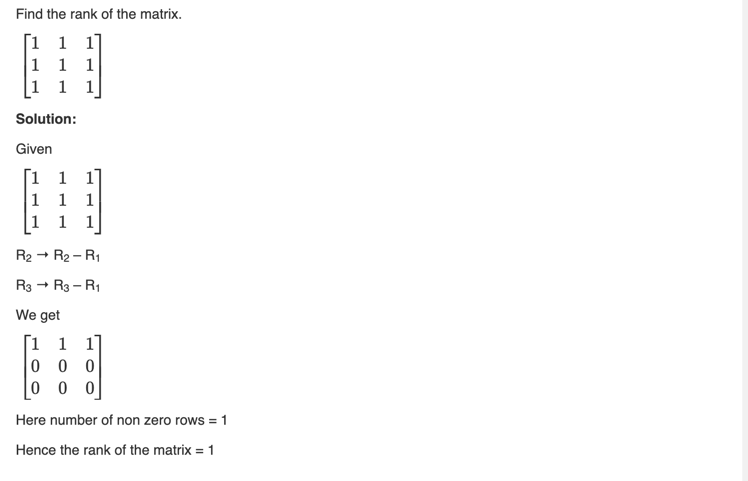


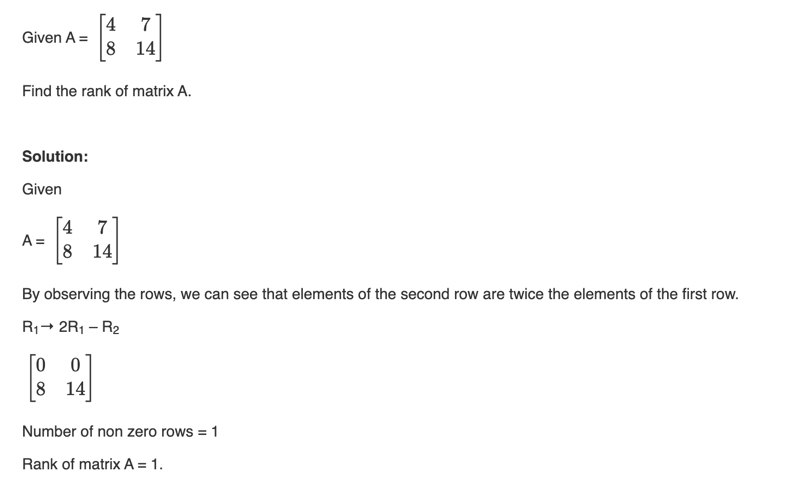


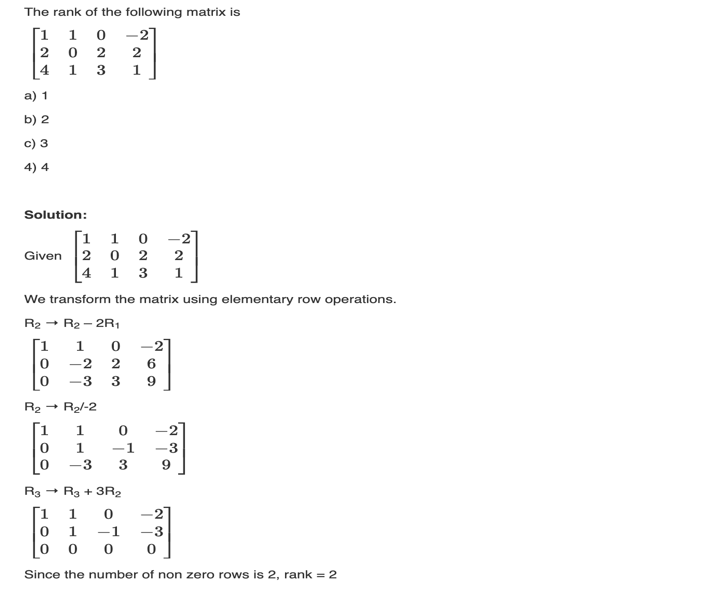






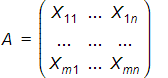


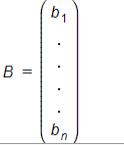




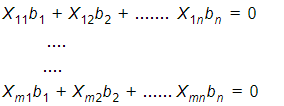
* 1. **Null Space and Nullity of a Matrix**
* Null space is a concept in linear algebra identifies the linear relationship among attributes.
* The null space of a matrix A consists of all vectors B such that AB -= 0 and B ≠ 0.
* Size of null space of matrix – number of linear relation among attributes

Consider

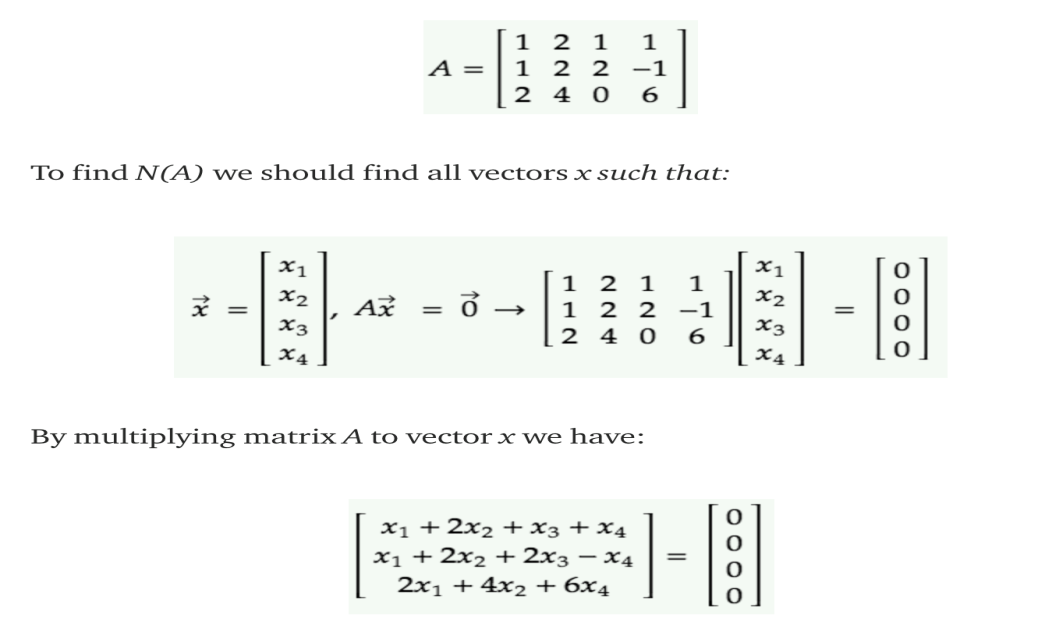
 Size of A is m \* n and

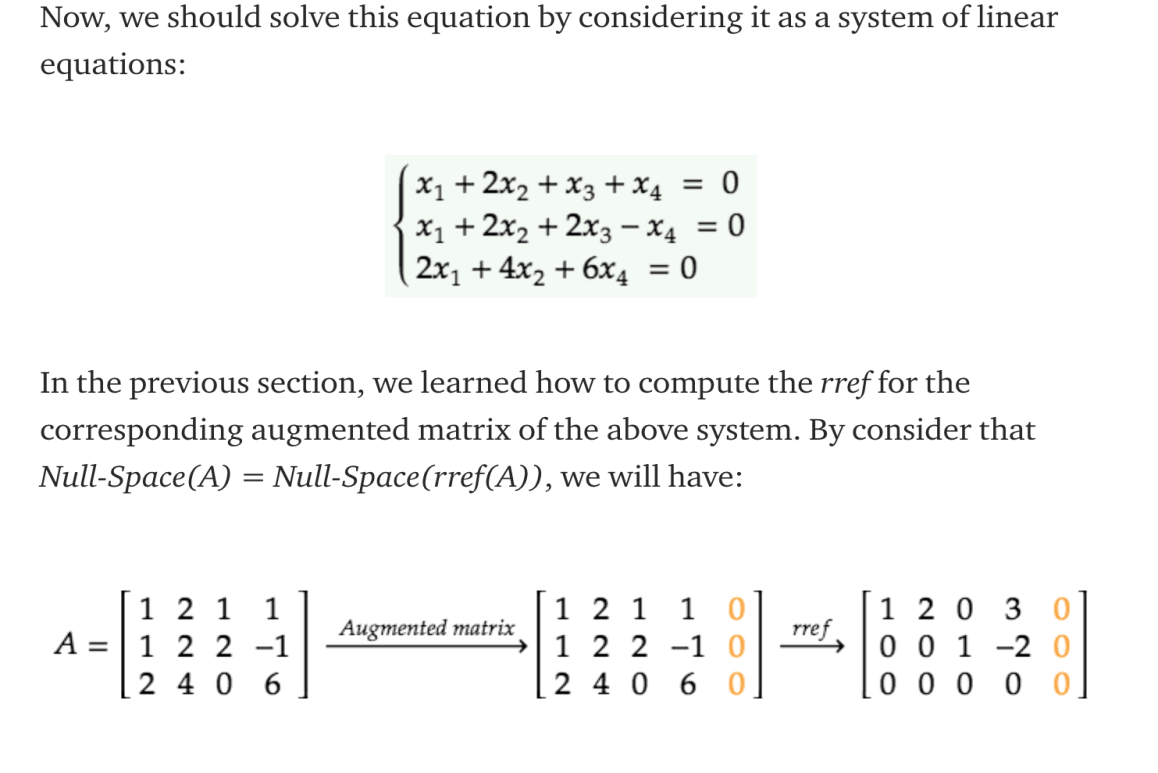
Size of B is n \* 1

Then the set of linear equations are

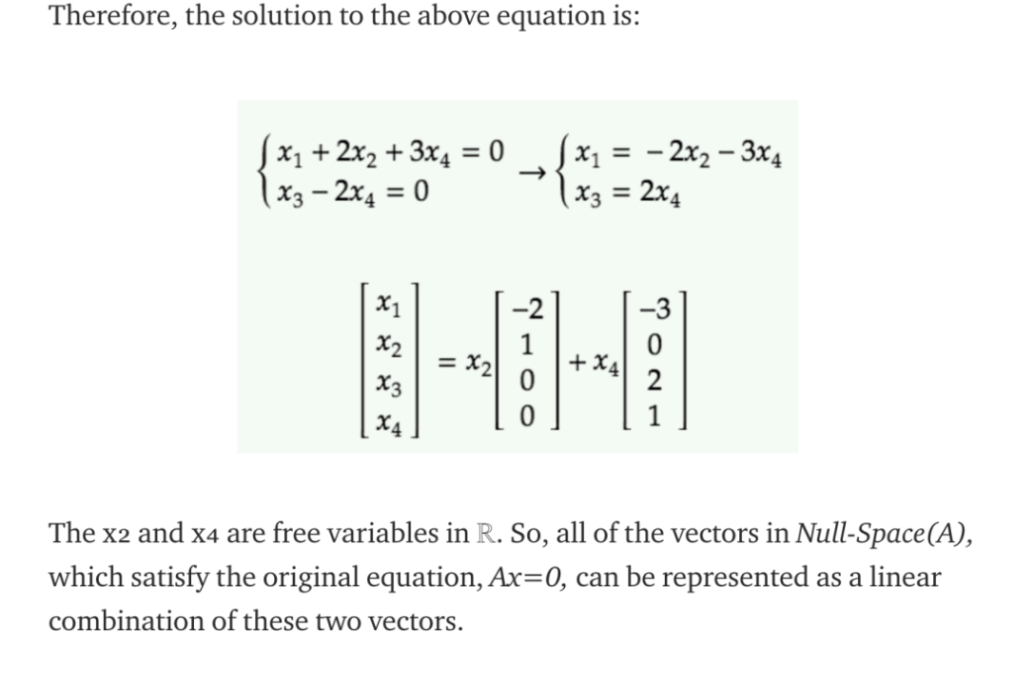


Eg : **Find the null space for given Matrix A**





Reduce the equation using Echelon form.



* 1. **Rank - Nullity Theorem**

The rank – nullity theorem helps us to relate the nullity of the data matrix to the rank and the number of attributes in the data

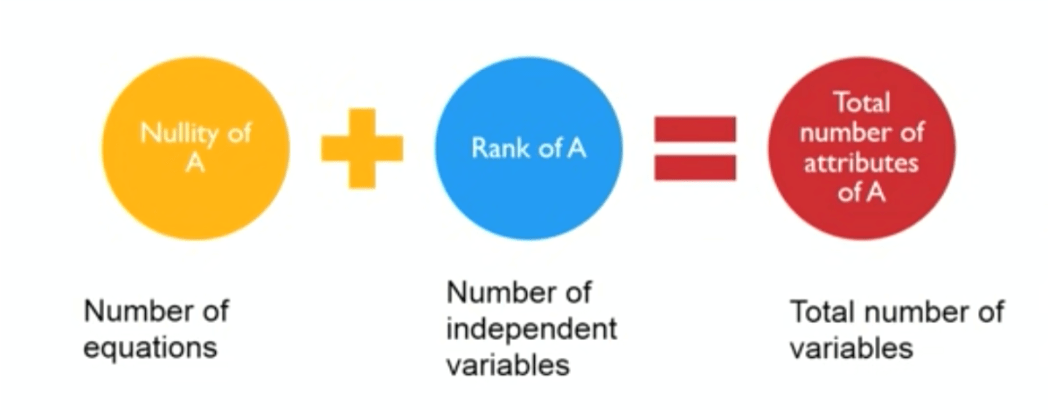
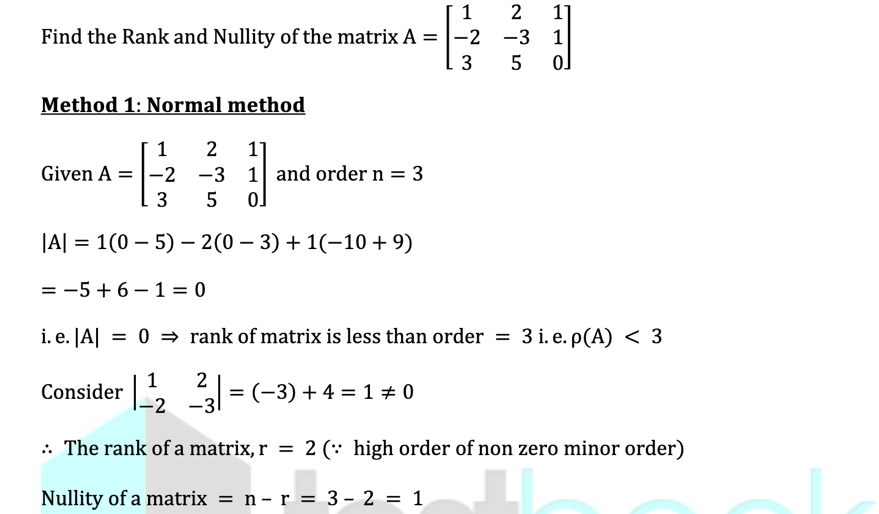


Figure 14 : Rank – Nullity Theorem

**Exercise 1: **

Rank of the matrix r=2

Nullity of a matrix =n-r= 3-2=1

**Applying Python on matrices – Null space**

# Sympy is a library in python for symbolic Mathematics

from sympy import Matrix

# List A

A = [[1, 2, 0], [2, 4, 0], [3, 6, 1]]

# Matrix A

A = Matrix(A)

# Null Space of A

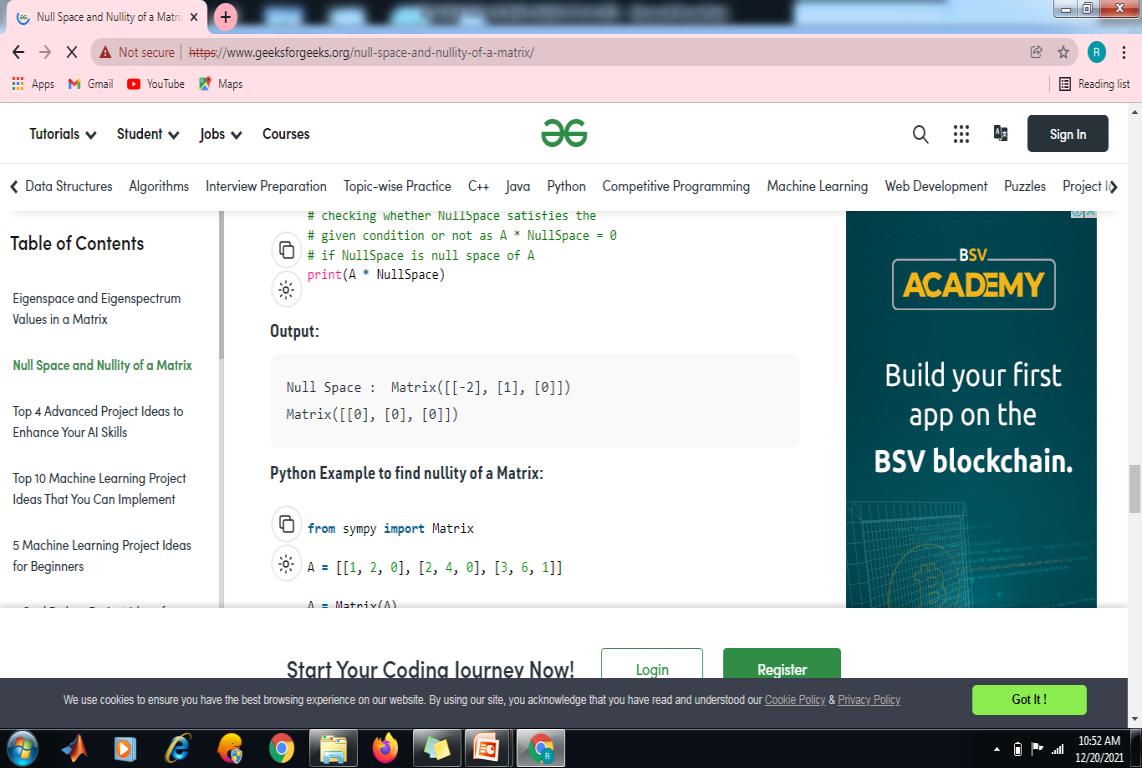
NullSpace = A.nullspace() # Here NullSpace is a list

NullSpace = Matrix(NullSpace) # Here NullSpace is a Matrix

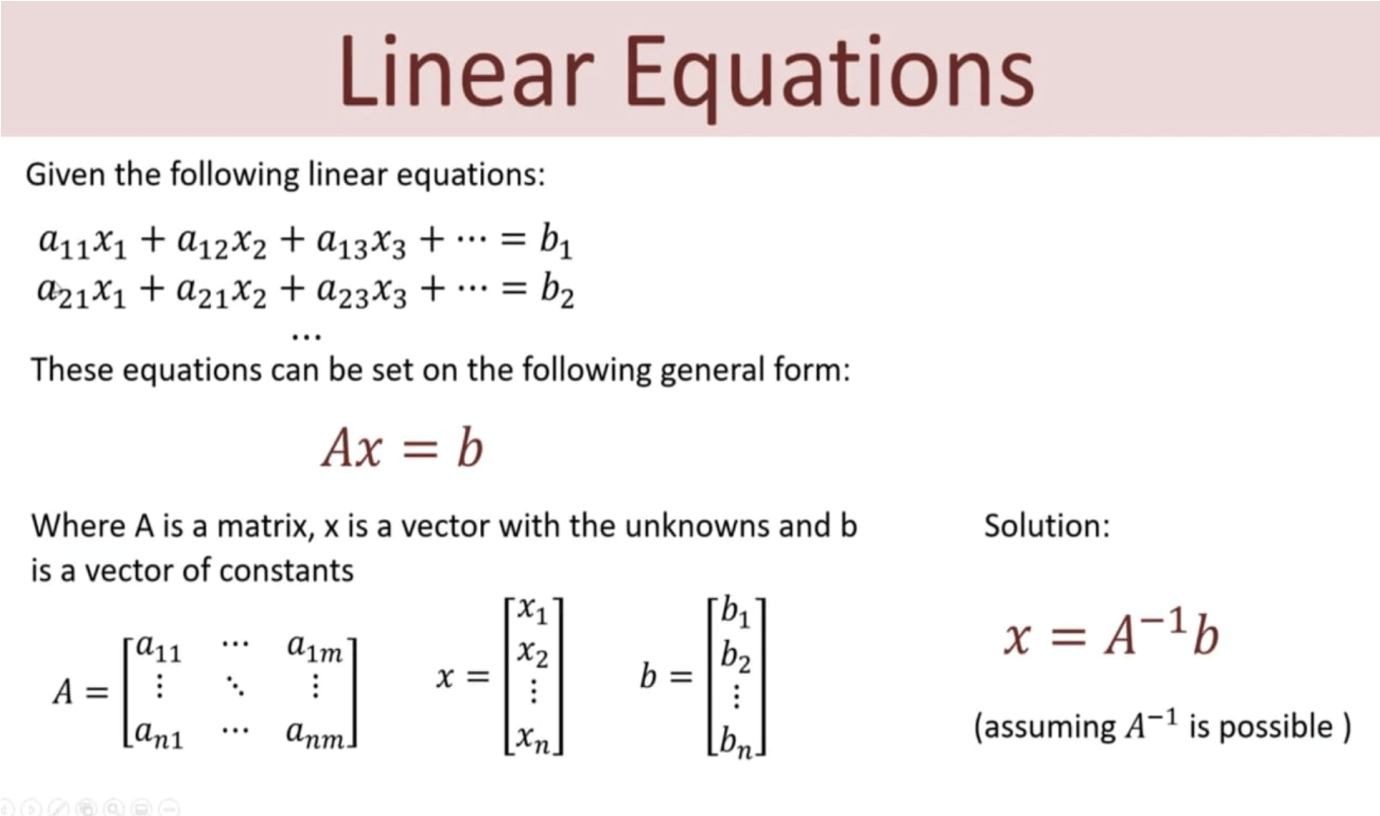
print("Null Space : ", NullSpace)

# checking whether NullSpace satisfies the given condition or not as A \* NullSpace = 0

print(A \* NullSpace)

****

* 1. **Linear Equations**

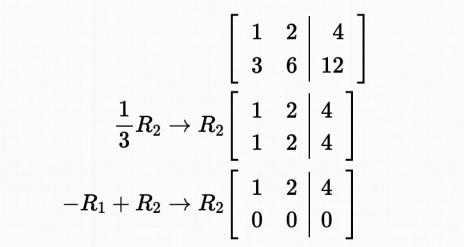
****

* 1. **Solution to over determined set of equations**
* If there are fewer equations than variables, then the system is called **underdetermined** and cannot have a unique solution. In this case, there are either infinitely many or no solutions.
* A system with more equations than variables is called **overdetermined.**
* If the number of equations equals the number of variables, it is a **balanced or square** system.
* A balanced system or an overdetermined system may have a unique solution.

**Example 1:** This system has infinitely many solutions. You can tell because these two lines are the same. (The second one is scaled by a factor of 3.)

x+2y=4

3x+6y=12



Notice that there is only one pivot column in this row-reduced matrix. The second column is not a pivot column, so we call y a free variable. A system with free variables is called dependent. Variables that do correspond to a pivot column are called fixed variables. In general, we can express the solution of the system as the fixed variables in terms of the free variables.

x=4−2yx=4−2y

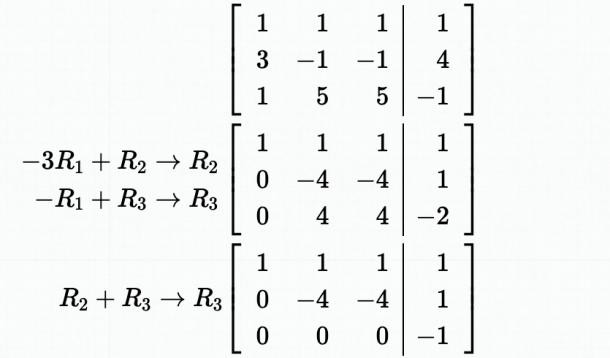
This is a dependent system, and there are infintely many solutions, depending on the value of the free variable y. Once a value of yy is selected, the value of x is automatically fixed.

**Example 2:**This system has no solution, so we call it inconsistent.

x+y+z=13

x−y−z=4

x+5y+5z=−1



We don't need to go any further than this. The last row reads 0x+0y+0z=−10x+0y+0z=−1, that is, 0=−10=−1. Such a false statement reveals that this system of equations has no solution. It is inconsistent.

* 1. **Eigen Values and Eigen Vectors**

The mathematical formulation is, Ax = λx

* The constant λ (positive) represents the amount of stretch or shrinkage the attributes x go through in the x direction.
* The solution x are known as eigen vectors and λ is eigen values.

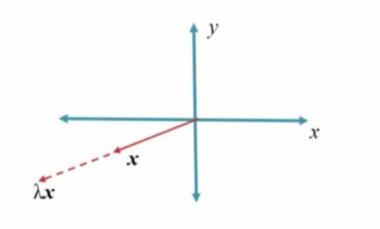
****

Figure 15: Eigen value λ

##### **Characteristic Equation**

The equation  is called the characteristic equation of the matrix A

##### **Note:**

* + 1. Solving , we get n roots for and these roots are called characteristic roots or eigen values or latent values of the matrix A
    2. Corresponding to each value of ,the equation AX = has a non-zero solution vector X.

If  be the non-zero vector satisfying AX= , when,  is said to be the latent vector or eigen vector of a matrix A corresponding to .

**Working rule to find characteristic equation:**

##### **Fora 3x3matrix:**

**Method1:**

The characteristic equation is 

##### **Method2:**

Its characteristic equation can be written as where S1 = sum of the main diagonal elements, S2=sum of the minors of the main diagonal elements,

S3=Determinant of A = |A|

##### **Fora2x2matrix**

**Method1:**

The characteristic equation is 

##### **Method2:**

Its characteristic equation can be written as  where S1= sum of the main diagonal elements, S2=Determinant of A = A

**Problems:**

1. **Find the characteristic equation of Solution**:

Its characteristic equation is 

Where S1=sum of the main diagonal elements=8+7+3=18,

S2=sum of the minors of the main diagonal elements=45

S3=Determinant of A= A=0

Therefore, the characteristic equation is .

##### **Find the characteristic equation of**

##### **Solution:**

##### The characteristic equation of A is



.

S1 = 3+2=5and=3(2)–1(-1)=7

Therefore, the characteristic equation is =0.

**Steps to find out the eigen value and eigen vector for variable x,**

1. Find the characteristic equation 
2. Solve the characteristic equation to get characteristic roots. They are called Eigen values
3. To find the Eigen vectors, solve for different values of

##### **Problems:**

1. **Find the eigen values and eigenvectors of the matrix** 

**Solution:**

Let A=  which is a non-symmetric matrix

##### To find the characteristic equation:

The characteristic equation of A is 

Where,

,

=1(-1)–1(3)=-4

Therefore, the characteristic equation is 

i.e., or

Therefore, the eigen values are 2,-2

##### To find the eigen vectors:



---------------(1)

##### Case1: If From(1)]

i.e.,

i.e.,,

i.e., we get only one equation



Therefore

##### Case2:If From(1)]

i.e.,

i.e.,



i.e., we get only one equation



Hence,

##### **Find the eigen values and eigen vectors of**

****

**Solution:**

##### Let A=

To find the characteristic equation:

Its characteristic equation can be written as 

where

,

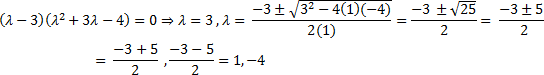
=2(-5)-2(-6)-7(2)=-10+12–14=-12

Therefore, the characteristic equation of A is 

3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 0 | -13 | 12 | |
| 1 | 3 | -4 | 0 |  |





Therefore, the eigenvaluesare3,1,and-4

##### To find the eigen vectors: Let



Case1:If

i.e.,

 (1)

 (2)

 (3)

Considering equations (1) and (2) and using method of cross-multiplication, we get, x1 x2 x3

2 -7 1 2



0 2 2 0

*x x x x x x*

## 123123

4 16

## 4 1

1

##  

4 1

Therefore,

*X*1=4

## 1

 

##### Case 2:If ,

i.e.,

 (1)

 (2)

 (3)

Considering equations (1) and (2) and using method of cross-multiplication, we get,x1 x2 x3

2 -7 -1 2



-2 2 2 -2

*x x x x x x*

## 123123

10 12 2 5 6 1



Therefore,

##### Case3:If

 (1)

 (2)

 (3)

Considering equations (1) and (2) and using method of cross-multiplication, we get,x1 x2 x3

2 -7 6 2



5 2 2 5



Therefore,

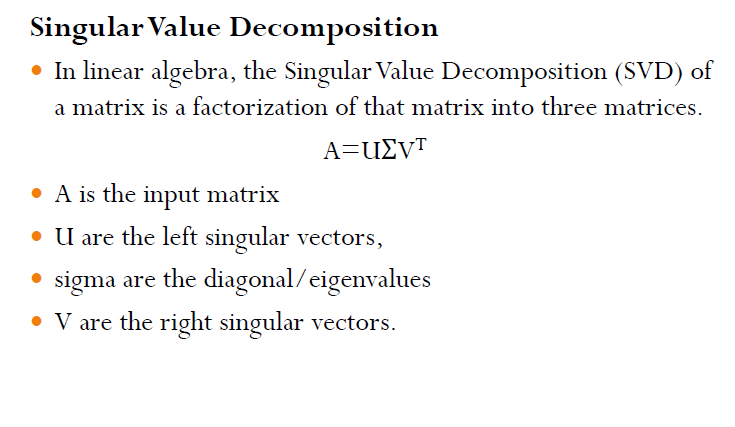
**Python Code for finding EigenVectors**

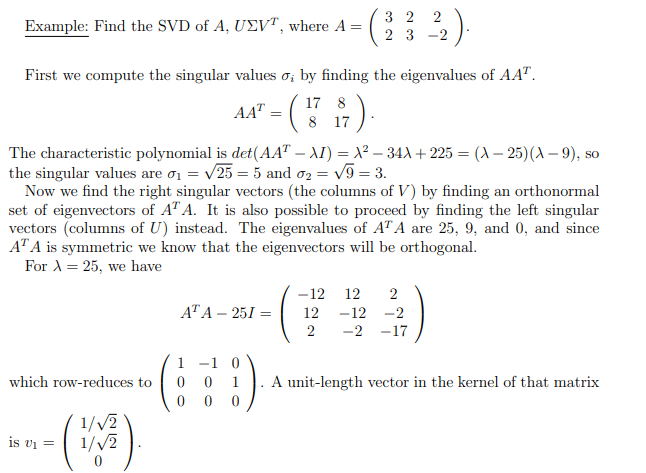
import  numpy as np

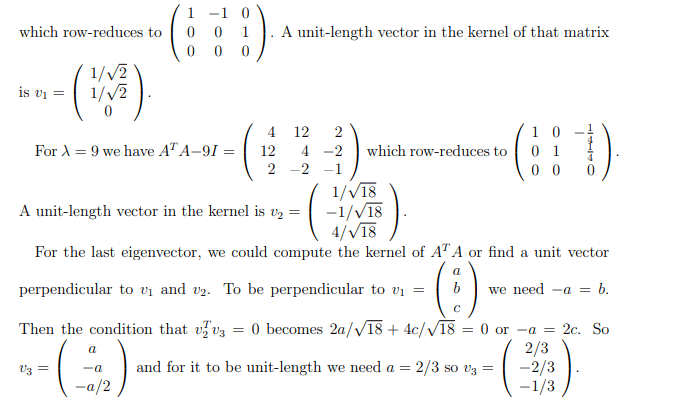
#create an array  
arr = np.arange(1,10).reshape(3,3)

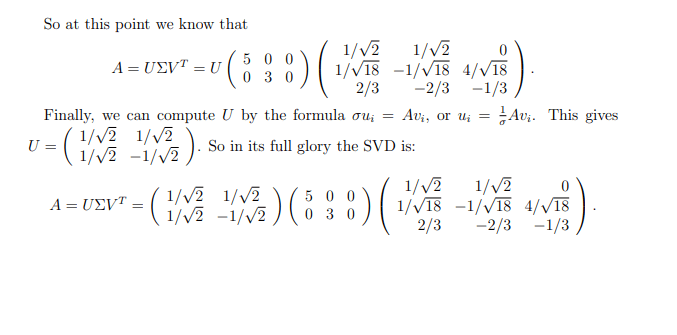
#finding the Eigenvalue and Eigenvectors of arr  
np.linalg.eig(arr)

* 1. **Singular Value Decomposition**



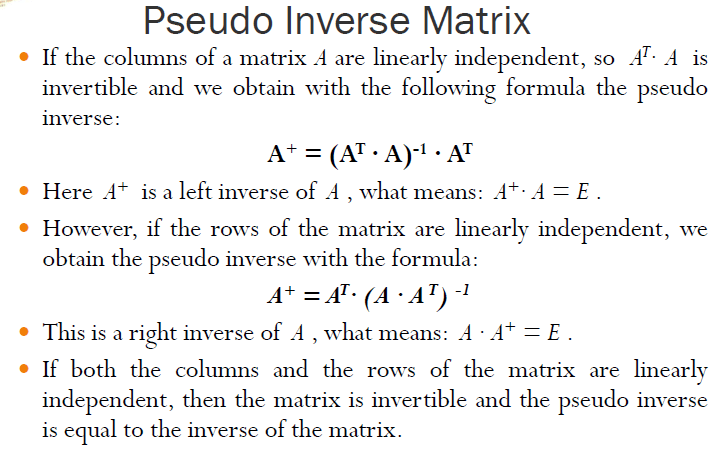


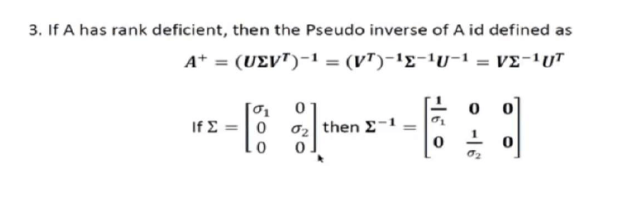




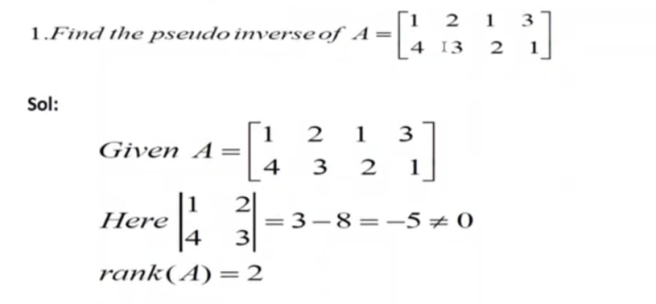
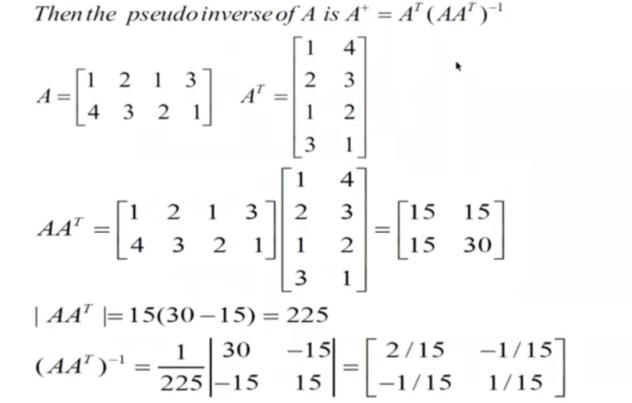
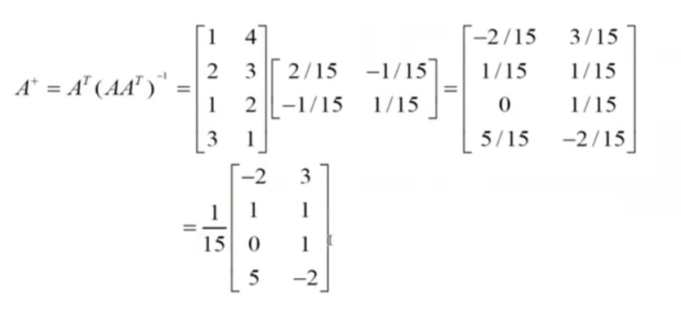
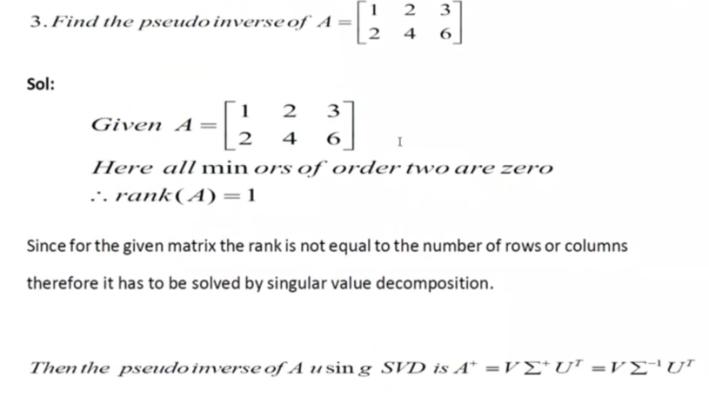
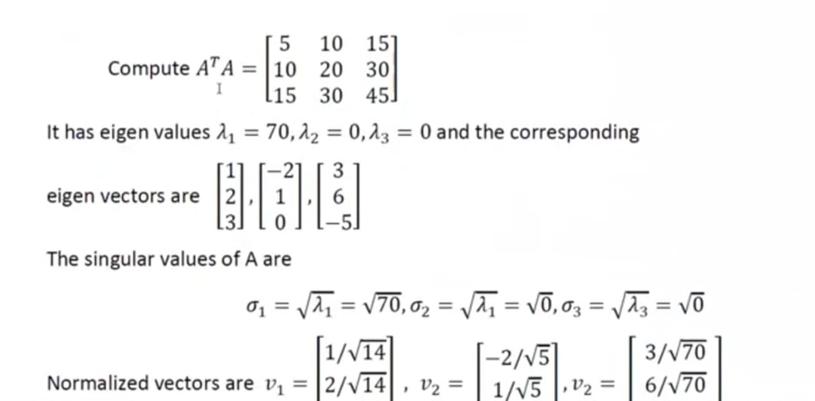
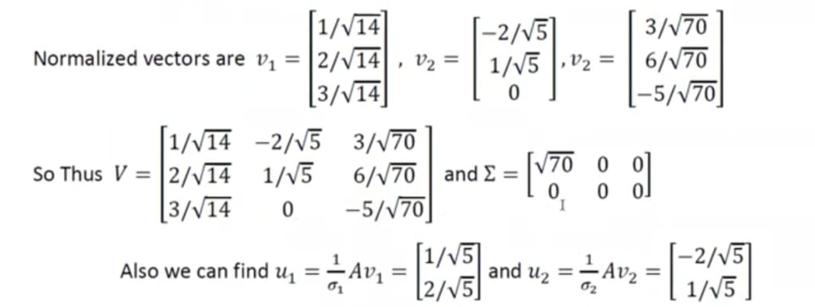
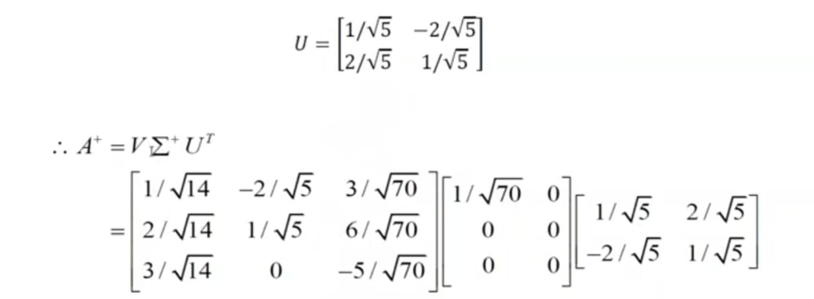
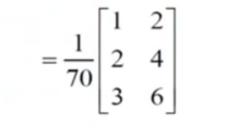
* 1. **Pseudo Inverse**

Pseudo inverse or Moore – Penrose inverse is the generalization of the matrix inverse that may not be invertible. If the matrix is invertible then its inverse will be equal to pseudo inverse and denoted by A+.

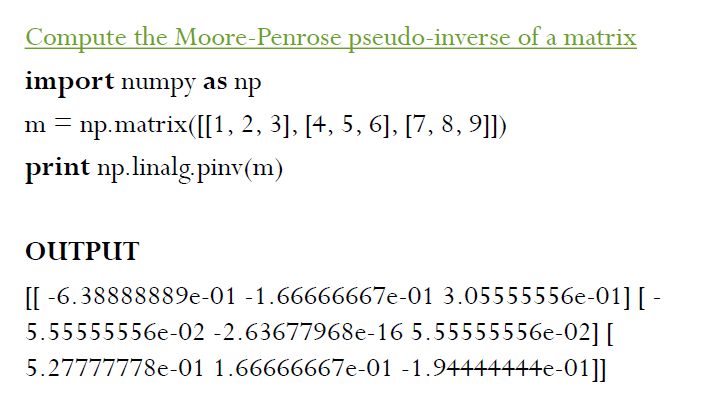




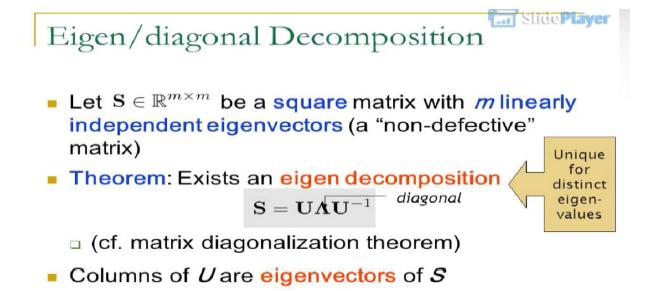
**Problems:**

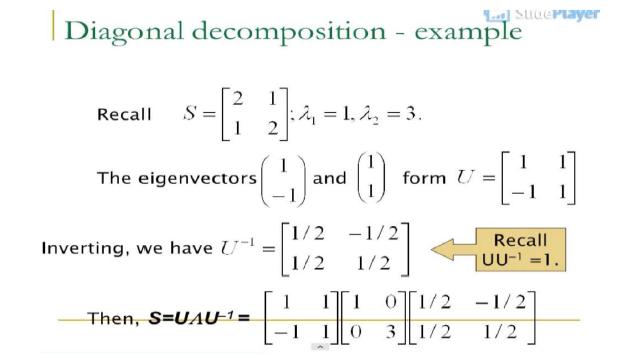
**Python code for Pseudo-inverse**

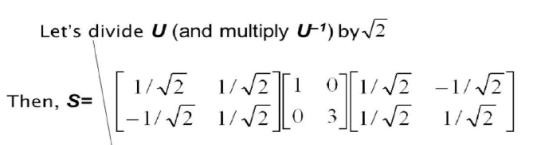


* 1. **Eigen Value Decomposition:**
* A few applications of eigen values and eigenvectors that are very useful when handing the data in a matrix form because you could decompose them into matrices that are easy to manipulate.
* In order for the matrix “A” to be either diagonalized or eigen decomposed, it has to meet the following criteria:
  + Must be a Square matrix
  + Has to have linearly independent eigenvectors



**Problem:**





* 1. **Equations of Line**
* The equation of a line means an equation in x and y whose solution set is a line in the (x,y) plane.
* The most popular form in algebra is the "slope-intercept" form **y = mx + b.**
* This in effect uses x as a parameter and writes y as a function of x: y = f(x) = mx+b. When x = 0, y = b and the point (0,b) is the intersection of the line with the y-axis.
* Line as a geometrical object and not the graph of a function, it makes sense to treat x and y more even handedly. The general equation for a line (normal form) is **ax + by = c,**
* This can easily be converted to slope-intercept form by solving for y:

**y = (-a/b) + c/b,**

except for the special case b = 0, when the line is parallel to the y-axis.

**Finding the equation of a line through 2 points in the plane**

* For any two points P and Q, there is exactly one line PQ through the points. If the coordinates of P and Q are known, then the coefficients a, b, c of an equation for the line can be found by solving a system of linear equations.

**Example**: For P = (1, 2), Q = (-2, 5), find the equation ax + by = cof line PQ.

* Since P is on the line, its coordinates satisfy the equation:

a(1) + b(2) = c, or a + 2b = c

Since Q is on the line, its coordinates satisfy the equation: a(-2) + b5 = c,

Or -2 a + 5b = c.

* Multiply the first equation by 2 and add to eliminate a from the equation:

4b + 5b = 9b = 2c + c = 3c, so b = (1/3)c.

* Then substituting into the first equation, a = c - 2b = c - (2/3)c = (1/3)c.
* This gives the equation**[(1/3)c]x + [(1/3)c}y = c**.
  1. **Equations of Plane**
* A plane in 3D-space has the equation

ax + by + cz = d,

* where at least one of the numbers a, b, c must be nonzero.
* If c is not zero, it is often useful to think of the plane as the graph of a function z of x and y. The equation can be rearranged like this:

z = -(a/c)x + (-b/c) y + d/c

* Another useful choice, when d is not zero, is to divide by d so that the constant term = 1.

(a/d)x + (b/d)y + (c/d)z = 1.

**Example: Finding the equation of a plane through 3 points in space**

* Given points P, Q, R in space, find the equation of the plane through the 3 points.

Example: P = (1, 1, 1), Q = (1, 2, 0), R = (-1, 2, 1). We seek the coefficients of an equation ax + by + cz = d, where P, Q and R satisfy the equations, thus:

a + b + c = d

a + 2b + 0c = d

-a + 2b + c = d

* Subtracting the first equation from the second and then adding the second equation to the third, we eliminate a to get

b - c = 0  
4b + c = 2d

* Adding the equations gives 5b = 2d, or b = (2/5)d, then solving for c = b = (2/5)d and then a = d - b - c = (1/5)d.
* So the equation (with a nonzero constant left in to choose) is d(1/5)x + d(2/5)y + d(2/5)z = d, so one choice of constant gives

x + 2y + 2z = 5

* or another choice would be (1/5)x + (2/5)y + (2/5)z = 1

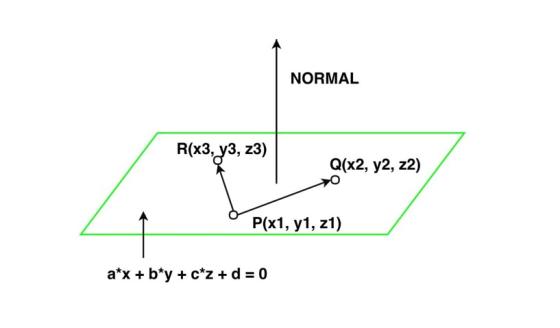
****

Figure 16. Equation of a plane

**Example 2: P(x1, y1, z1), Q(x2, y2, z2), and R (x3, y3, z3) are three non-collinear points on a plane.  Find equation of plane.**

We know that: ax + by + cz + d = 0 —————(i)

By plugging in the values of the points P, Q, and R into equation (i), we get the following:

a(x1) + b(y1) + c(z1) + d = 0

a(x2) + b(y2) + c(z2) + d = 0

a(x3) + b(y3) + c(z3) + d = 0

Suppose, P = (1,0,2), Q = (2,1,1), and R = (−1,2,1)

Then, by substituting the values in the above equations, we get the following:

a(1) + b(0) + c(2) + d = 0

a(2) + b(1) + c(1) + d = 0

a(-1) + b(2) + c(1) + d = 0

Solving these equations gives us b = 3a, c = 4a, and d = (-9)a——————(ii)

By plugging in the values from (ii) into (i), we end up with the following:

ax + by + cz + d = 0

ax + 3ay + 4az−9a

x + 3y + 4z−9

Therefore, the equation of the plane with the three non-collinear points P, Q, and R is x + 3y + 4z−9.

**Example 3: A (3,1,2), B (6,1,2), and C (0,2,0) are three non-collinear points on a plane. Find the equation of the plane.**

**Solution:**

We know that: ax + by + cz + d = 0 —————(i)

By plugging in the values of the points A, B, and C into equation (i), we get the following:

a(3) + b(1) + c(2) + d = 0

a(6) + b(1) + c(2) + d = 0

a(0) + b(2) + c(0) + d = 0

Solving these equations gives us

a = 0, c = 1/2b, d = —2b ———————(ii)

By plugging in the values from (ii) into (i), we end up with the following:

ax + by + cz + d = 0

(0)x + (—by) + ½ bz — 2b = 0

x - y + ½ z —2 = 0

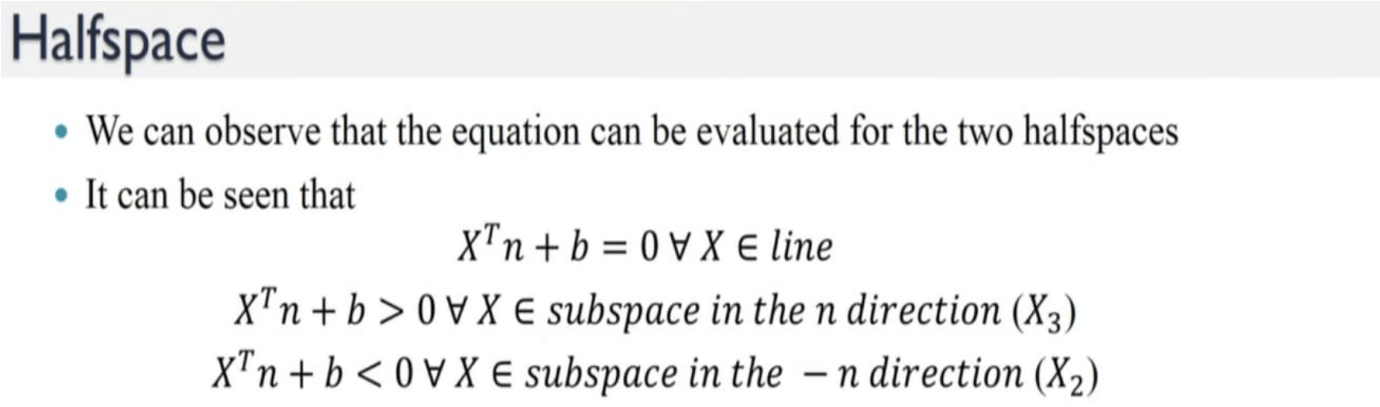
2x-2y + z-4 = 0

Therefore, the equation of the plane with the three non-collinear points A, B and C is

2x-2y + z-4 = 0.

* 1. **Equation of hyperplane**
* A hyperplane is a higher-dimensional generalization of lines and planes.
* The equation of a hyperplane is w · x + b = 0, where w is a vector normal to the hyperplane and b is an offset.

**Half Space**



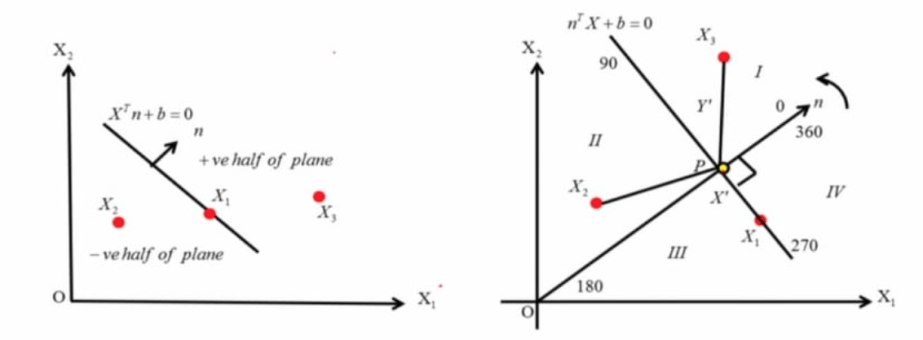
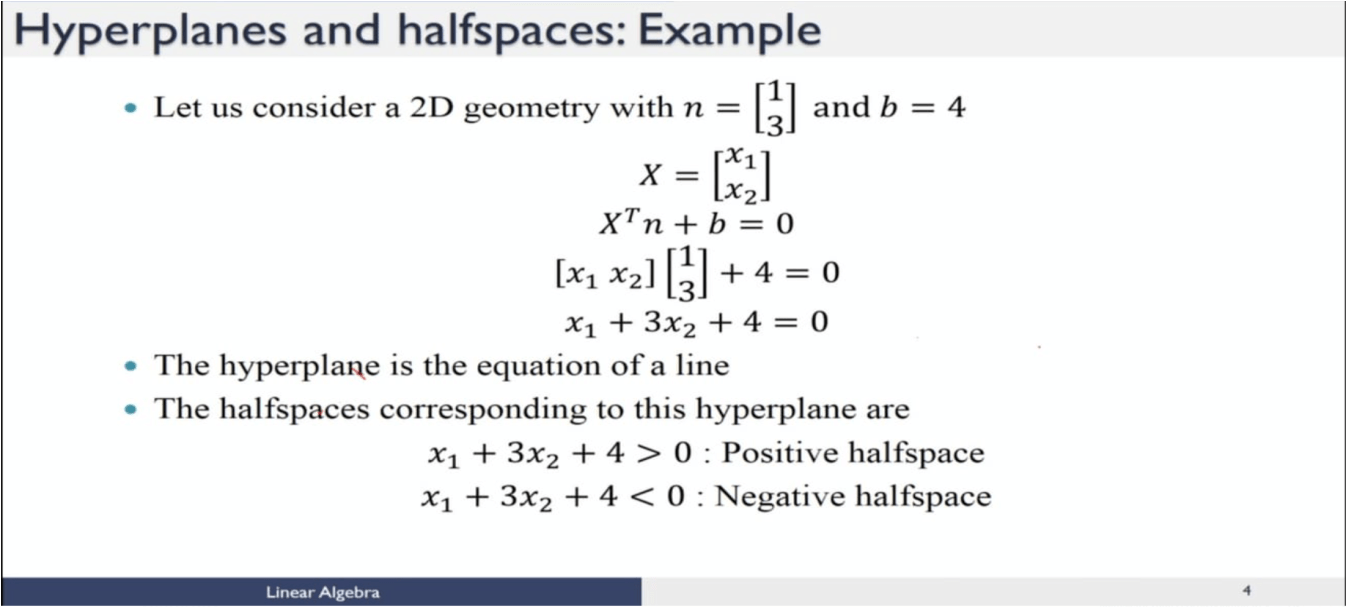


Figure 17: Halfspace

**Example**



* 1. **Equation of circle**

A circle is a closed curve that is drawn from the fixed point called the center, in which all the points on the curve are having the same distance from the center point of the center. The equation of a circle with (h, k) center and r radius is given by:

(x-h)2 + (y-k)2 = r2

This is the standard form of the equation.Thus, if we know the coordinates of the center of the circle and its radius as well, we can easily find its equation.

**Example**:

1. Consider a circle whose center is at the origin and radius is equal to 8 units.

**Solution:**

Given: Centre is (0, 0), radius is 8 units.

We know that the equation of a circle when the center is origin:

**x2+ y2 = a2**

For the given condition, the equation of a circle is given as

x2+ y2 = 82

x2+ y2= 64, which is the equation of a circle

1. Find the equation of the circle whose center is (3,5) and the radius is 4 units.

**Solution:**

Here, the center of the circle is not an origin.

Therefore, the general equation of the circle is,

(x-3)2 + (y-5)2 = 42

x2 – 6x + 9 + y2 -10y +25 = 16

x2 +y2 -6x -10y + 18 =0 is the equation of circle

1. Equation of a circle is x2+y2−12x−16y+19=0. Find the center and radius of the circle.

**Solution:**

Given equation is of the form x2+ y2 + 2gx + 2fy + c = 0,

2g = −12, 2f = −16,c = 19

g = −6,f = −8

Centre of the circle is (6,8)

Radius of the circle = √[(−6)2 + (−8)2 − 19 ]= √[100 − 19] = √81 = 9 units.

Therefore, the radius of the circle is 9 units.

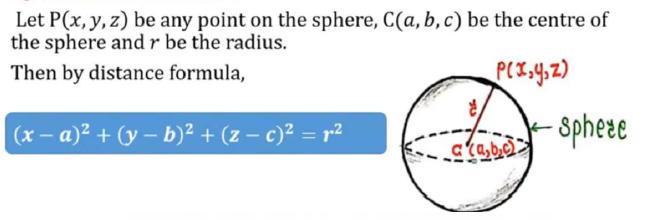
* 1. **Equation of sphere:**

A sphere is a geometrical object in three-dimensional space that resembles the surface of a ball. Similar to a circle in two-dimensional space, a sphere can be mathematically defined as the set of all points that are at the same distance from a given point. This given point is called the center of the sphere. The distance between the center and any point on the surface of the sphere is called the radius, represented by r.

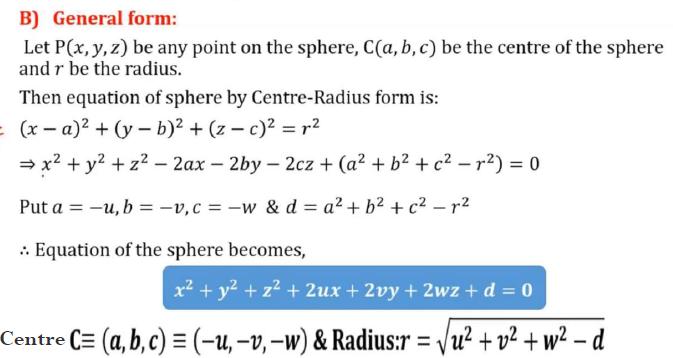
x2 + y2 + z2= r2

This is called the equation of a sphere, also known as the general equation of a sphere or the equation of a sphere through the circle.

1. Center - Radius form

****

1. **General Form**

****

* 1. **Equation of hypersphere**
* A hypersphere is a four-dimensional analog of a sphere; also known as a 4-sphere.
* The intersection of a sphere with a plane is a circle; the intersection of a hypersphere with a hyperplane is a sphere. These analogies are reflected in the underlying mathematics.
* *x*2 + *y*2 = *r* 2 is the Cartesian equation of a circle of radius *r*; *x*2 + *y*2 + *z*2 = *r* 2 is the corresponding equation of a sphere; *x*2 + *y*2 + *z*2 + *w*2 = *r* 2 is the equation of a hypersphere, where *w* is measured along a fourth dimension at right angles to the *x-*, *y*-, and *z*-axes.
* The n-hypersphere (often simply called the n-sphere) is a generalization of the circle (called by geometers the 2-sphere) and usual sphere (called by geometers the 3-sphere) to dimensions n>=4. The n-sphere is therefore defined (again, to a geometer; see below) as the set of n-tuples of points (x1, x2, ..., xn) such that

x12+x22+...+xn2=R2,

where R is the radius of the hypersphere.

* The hypersphere has a **hypervolume** (analogous to the volume of a sphere) of π2*r*4/2, and a surface volume (analogous to the sphere's surface area) of 2π2*r*3.
* A solid angle of a hypersphere is measured in **hypersteradians**, of which the hypersphere contains a total of 2π2. The apparent pattern of 2π [**radians**](https://www.daviddarling.info/encyclopedia/A/angular_units.html) in a circle and 4π [**steradians**](https://www.daviddarling.info/encyclopedia/A/angular_units.html) in a sphere does not continue with 8π hypersteradians because the *n*-volume, *n*-area, and number of *n*-radians of an *n*-sphere are all related to gamma function and the way it can cancel out powers of π halfway between integers. In general, the term "hypersphere" may be used to refer to any *n*-sphere.

**Question Bank**

|  |  |  |  |
| --- | --- | --- | --- |
| **Part-A** | | | |
| **Q.No** | **Questions** | **Competence** | **BT Level** |
|  | Define Data science. | Remember | BTL 1 |
|  | Distinguish between between a vector and a scalar point | Analysis | BTL 4 |
|  | List out the applications of data science. | Remember | BTL 1 |
|  | Differentiate between matrices and tensors. | Analysis | BTL 4 |
|  | When two vector will become perpendicular? | Analysis | BTL 4 |
|  | Define Linear algebra. | Understand | BTL 2 |
|  | Illustrate relationship diagram between data science and other branch of studies. | Apply | BTL 3 |
|  | Illustrate the geometrical representation of a vector point. | Apply | BTL 3 |
|  | Is all orthogonal vector are perpendicular vectors? Justify your answer. | Analysis | BTL 4 |
|  | Given two points A and B with coordinates (3,5) and (6,8) respectively. What is the distance between A and B? | Analysis | BTL 4 |
|  | Write down a python program to find out the determinant and inverse of a matrix. | Create | BTL6 |
|  | Write a python code to perform element wise multiplication and dot product on two matrices. | Create | BTL6 |
|  | Write a python code to find out null space of a matrix. | Create | BTL6 |
|  | Define Rank of a matrix with an example. | Understand | BTL 2 |
|  | Describe Rank-Nullity theorem. | Understand | BTL 2 |
|  | Differentiate between underdetermined and overdetermined set of equations. | Analysis | BTL 4 |
|  | When a set of equations can be termed as balanced system? | Understand | BTL 2 |
|  | Compare and contrast hyperplane and halfspace. | Analysis | BTL 4 |
|  | Find out the Eigen value for the matrix | Apply | BTL 3 |
|  | What is the equation of a circle when the center is at the origin? | Analysis | BTL 4 |
| **PART B** | | | |
| **Q.No** | **Questions** | **Competence** | **BT Level** |
|  | Find the inverse of the given matrix A = | Analysis | BTL 4 |
|  | Find the rank and nullity of the given matrix | Apply | BTL 3 |
|  | Calculate the Null Space for the matrix A | Analysis | BTL 4 |
|  | Find out the eigen values and eigen vectors for the given matrix | Apply | BTL 3 |
|  | Perform diagonal decomposition to the given matrix A to make the sub matrices.  A = | Analysis | BTL 4 |
|  | Do Singular Value Decomposition in the given matrix | Analysis | BTL 4 |
|  | Calculate the pseudo inverse for the matrix B = | Analysis | BTL 4 |
|  | Find the equation of the sphere which passes through the points (2,1,1) and (0,3,2) and has its center on the line  2x + y + 3z = 0 = x + 2y + 2z | Apply | BTL 3 |