SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCSA3016 DATA SCIENCE

UNIT-I

LINEAR ALGEBRA

C.KAVITHA

ASSISTANT PROFESSOR

DEPARTMENT OF CSE

12/16/21

SYLLABUS

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.

UNIT 2 PROBABILITY AND STATISTICS

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.

SYLLABUS

UNIT 3 EXPLORATORY DATA ANALYSIS AND THE DATA SCIENCE PROCESS

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

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

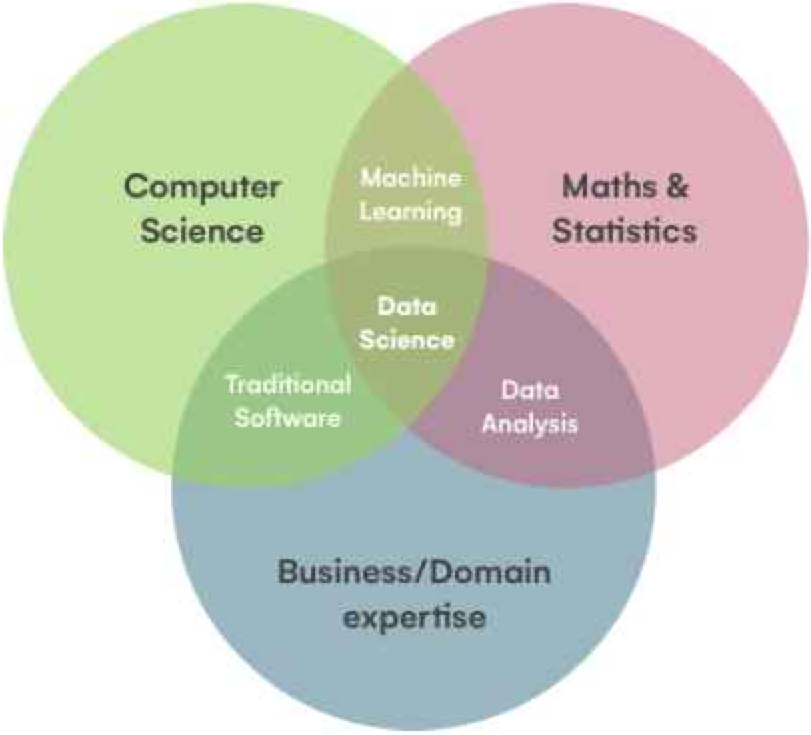
Data structures-Functions-Numpy-Matplotlib-Pandas- problems based on computational complexity-Simple case studies based on python (Binary search, common elements in list), Hash tables, Dictionary.

TEXT / REFERENCE BOOKS

1. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O'Reilly. 2014.
2. Introduction to Linear Algebra - By Gilbert Strang, Wellesley-Cambridge Press, 5th Edition.2016.
3. Applied Statistics and Probability For Engineers – By Douglas Montgomery.2016.
4. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press. 2014. (free online)
5. Avrim Blum, John Hopcroft and Ravindran Kannan. Foundations of Data Science.
6. Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, 3rd Edition. ISBN 0123814790, 2011.
7. Trevor Hastie, Robert Tibshirani and Jerome Friedman. Elements of Statistical Learning, 2nd Edition. ISBN 0387952845. 2009. (free online)

What is Data Science?

- **Data Science** is the science of analyzing **raw data** using statistics and machine learning techniques with the purpose of drawing insights from the data
- **Data Science** is used in many industries to allow them to make better business decisions, and in the sciences to test models or theories
- This requires a process of inspecting, cleaning, transforming, modeling, analyzing, and interpreting raw data


Computer
Science
Machine
Learning
Maths &
Statistics
Data
Science
Traditional
Software
Data
Analysis
Business/Domain
expertise

Data perspective

- Read data
- Data processing and cleaning
- Summarizing data
- Visualization
- Deriving insights from data

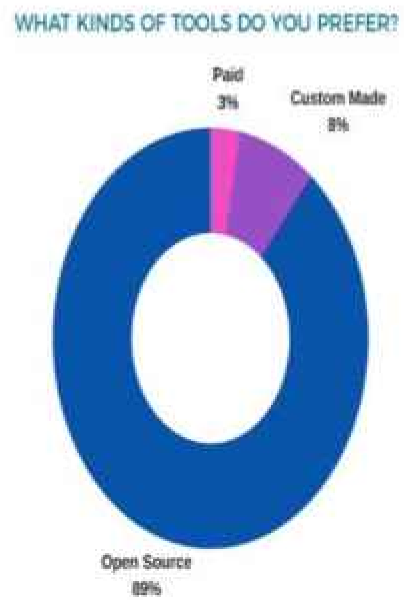
Data science using Python

- Python libraries provide key feature sets which are essential for data science
- Data manipulation and pre-processing
  - Python's 'pandas' library offers a variety of functions for data wrangling and manipulation
- Data summary
- Visualization
  - Plotting libraries like 'matplotlib' and 'seaborn' aid in condensing statistical information and help in identifying trends and relationships

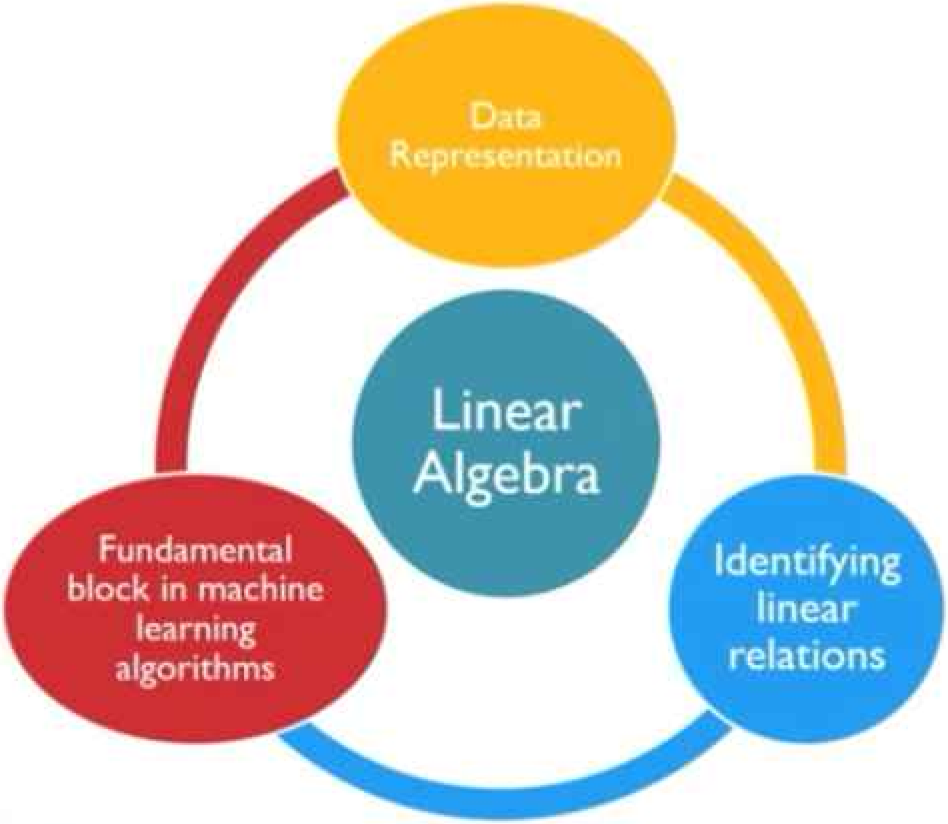
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Trends in tools used for data science

As of 2018, most Indian data scientists prefer to use open source tools over paid or custom made tools



Overview


Data Representation
Linear Algebra
Fundamental block in machine learning algorithms
Identifying linear relations

Matrix theory and linear algebra

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 understand and manipulate matrices to derive useful knowledge from data

Matrices for data science: Data representation

- Usually matrices are used to store and represent the data on machines
- Matrix is a very natural approach for organizing data
- In general, data is organized in the following fashion
  - Rows represent samples
  - Columns represent the values of the variables (or attributes)
  - It is also possible to use rows for variables and columns for samples
  - However, we will stick to rows as samples and columns as variables in all of the material that will be presented

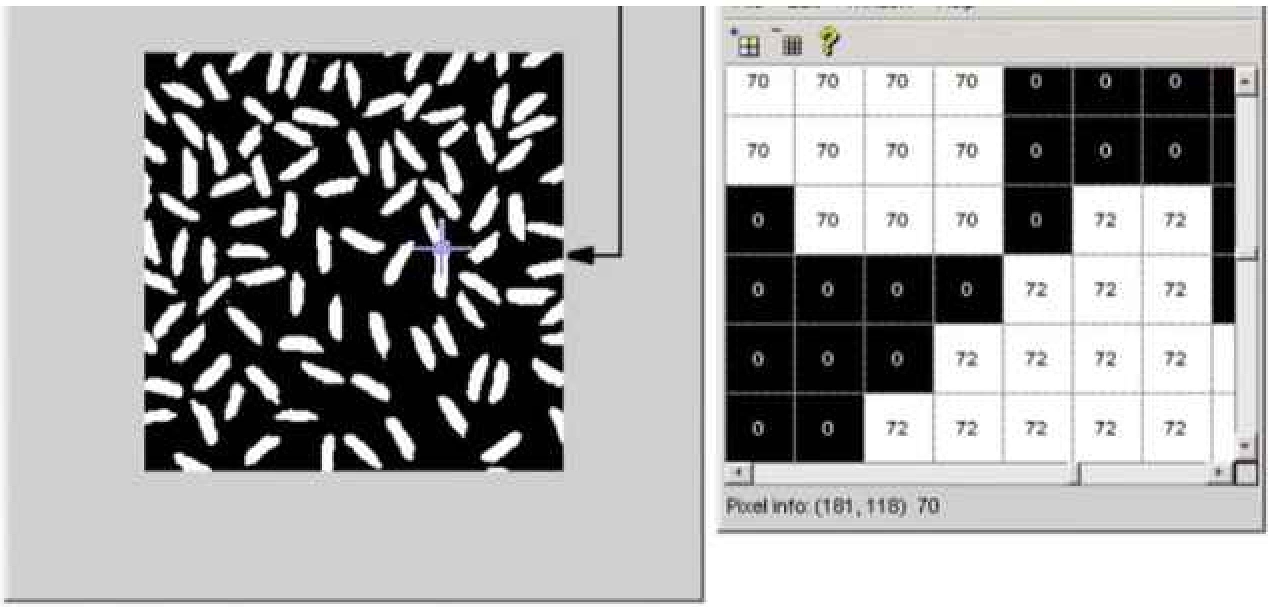
Data representation: Examples

- A real life example
  - Consider a reactor which needs to be controlled using multiple attributes from various sensors like Pressure (Pa), Temperature (K), Density (gm/m^3) etc.
  - Independently, the sensors have generated 1,000 data points
  - This complete set of information is contained in

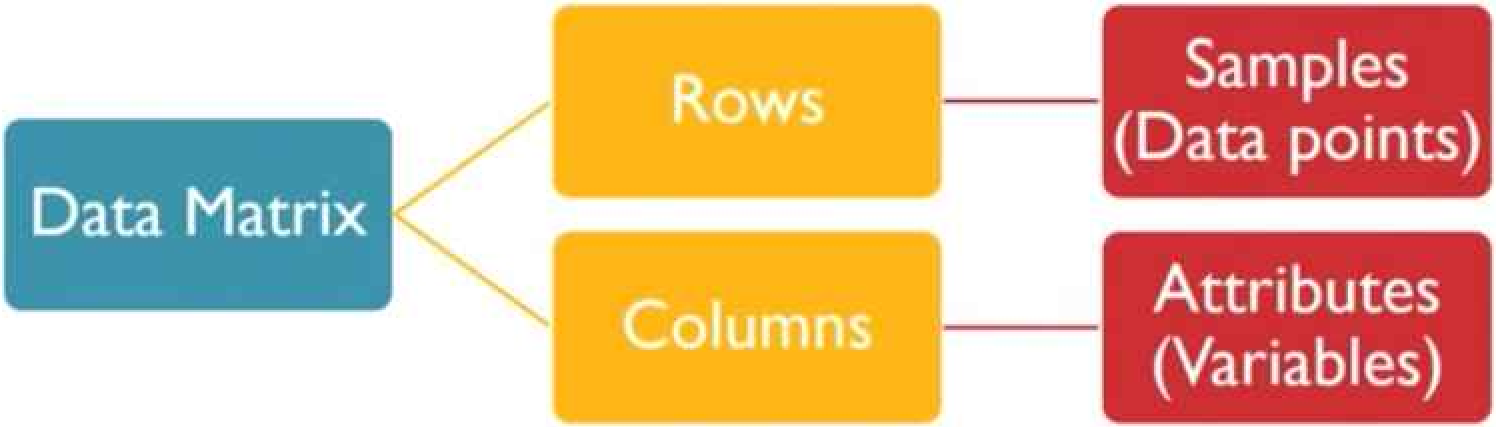
$$\begin{array}{c c} & \begin{array}{ccc} P & T & \rho \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ 1000 \end{array} & \begin{bmatrix} 300 & 300 & 1000 \\ \vdots & \vdots & \vdots \\ 500 & 1000 & 5000 \end{bmatrix} \end{array}$$

Data representation: Examples

- The simplicity in representation will become apparent when the image below is considered



Data as matrix: Summary



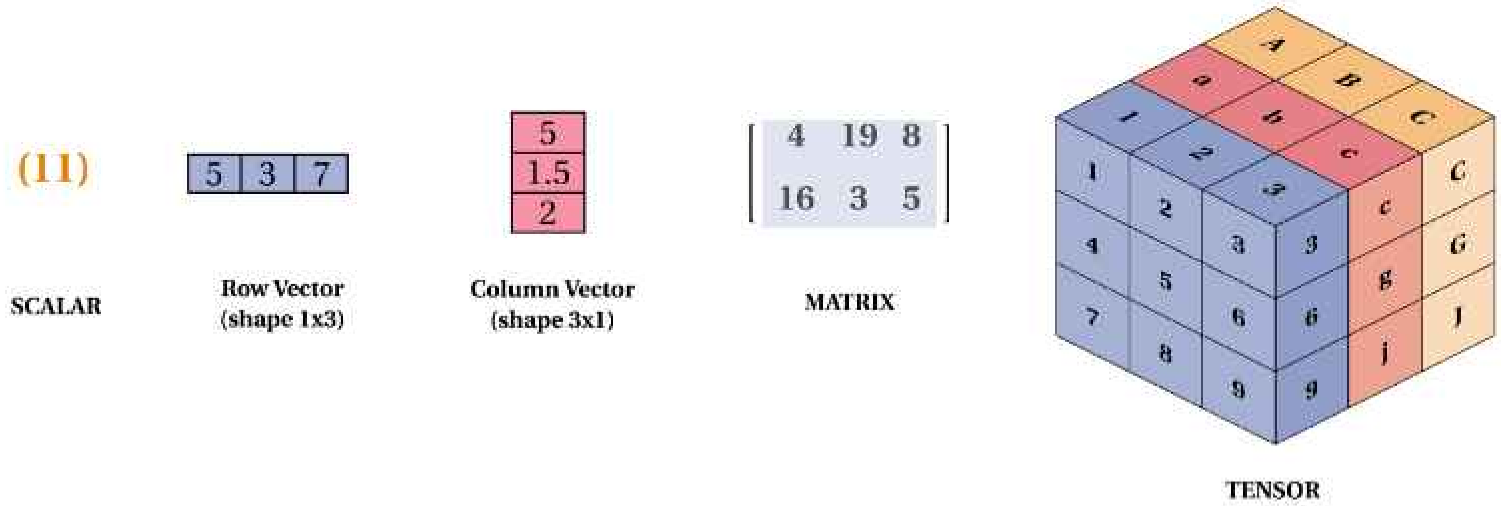
- **Scalar:** Any single numerical value is a scalar as shown in the image above. It is simply denoted by lowercase and italics. For example: n
- **Vector:** An array of numbers(data) is a vector. You can assume a column in a dataset to be a feature vector.

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

- **Matrix**: A matrix is a 2-D array of shape (m×n) with m rows and n columns.

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}$$

- **Tensor:** Generally, an n-dimensional array where n>2 is called a Tensor. But a matrix or a vector is also a valid tensor.


(11)
SCALAR
5 3 7
Row Vector
(shape 1x3)
5
1.5
2
Column Vector
(shape 3x1)
4 19 8
16 3 5
MATRIX
TENSOR

What is a 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.
- In **machine learning**, feature **vectors** are used to represent numeric or symbolic characteristics, called features, of an object in a mathematical, easily analyzable way.

Types of Vectors

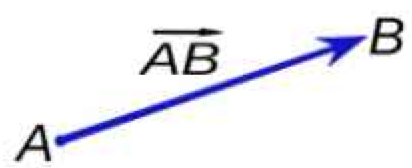
We can represent a vector as a line segment oriented from an initial point, called the tail, to a final point, called the head.

Geometric Vectors

Geometric vectors are not related to any coordinate system.

Algebraic vectors

Algebraic vectors are related to a coordinate system.



A geometric vector is not related to any coordinate system. A is the tail, B is the head.

$$\vec{v} = (1, 2)$$

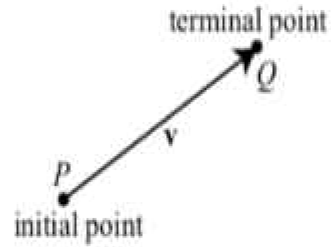
(1,2) is the point located in the coordinate system, where 1 corresponds to x and 2 to y.

Within this algebraic vectors, we can have a **position vector** that connects the origin of the coordinate system (O) with any point (P), and we write it like this:

$$\vec{OP}$$

Geometric Representation of Vectors

Vectors can be represented geometrically by arrows (directed line segments). The **arrowhead** indicates the direction of the vector, and the **length** of the arrow describes the magnitude of the vector.



A vector with initial point P (the tail of the arrow) and terminal point Q (the tip of the arrowhead) can be represented by

$$\overrightarrow{PQ}, \mathbf{v}, \text{ or } \vec{v}. \tag{3.5.1}$$

We often write $v = \overrightarrow{PQ}$. In this text, we will use boldface font to designate a vector. When writing with pencil and paper, we always use an arrow above the letter (such as $\vec{v}$) to designate a vector. The **magnitude** (or **norm** or **length**) of the vector $\mathbf{v}$ is designated by $|\mathbf{v}|$. It is important to remember that $|\mathbf{v}|$ is a number that represents the magnitude or length of the vector $\mathbf{v}$.

According to our definition, a vector possesses the attributes of length (magnitude) and direction, but position is not mentioned. So we will consider two vectors to be equal if they have the same magnitude and direction. For example, if two different cars are both traveling at 45 miles per hour northwest (but in different locations), they have equal velocity vectors. We make a more formal definition.

Perpendicular vectors and Orthogonal vectors

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

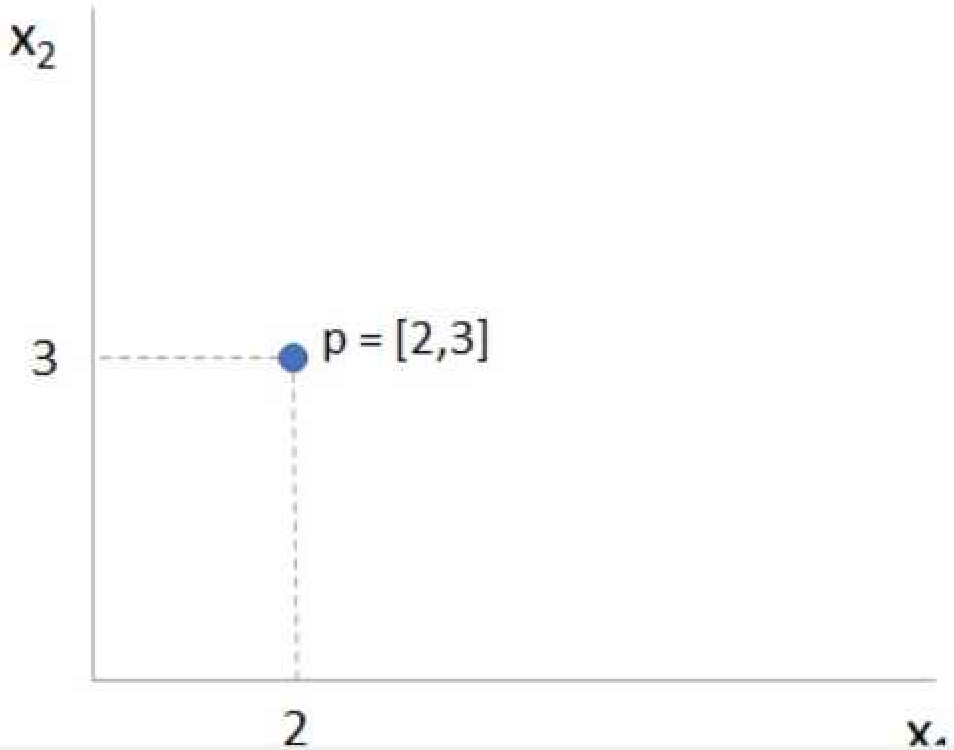
$$\vec{x} \text{ and } \vec{y} \text{ are perpendicular} \rightarrow \vec{x}.\vec{y} = 0$$

$$\vec{x}.\vec{y} = 0 \,,\ \vec{x},\ \vec{y} \neq \vec{0} \rightarrow \vec{x} \text{ and } \vec{y} \text{ are perpendicular}$$

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

$$\vec{0}.\vec{x} = 0$$

Defining a 2D point/Vector:

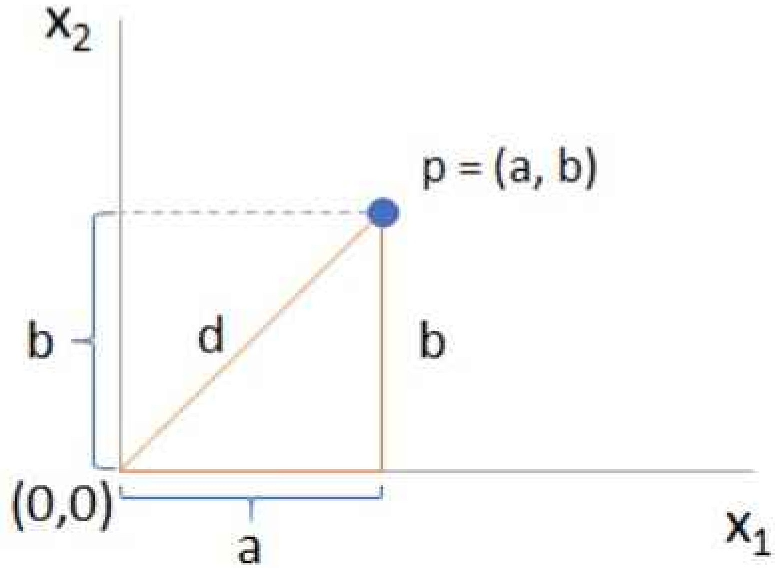


In 2D space, a point is defined as the (x,y) coordinates as shown above.
Here, the x_1 coordinate (x coordinate) is 2, and the x_2 coordinate (y coordinate) is 3.

Defining a point in 3D space:

Extending the 2D concept in 3D space point 'p' is defined by (x,y,z) coordinates, where 2 is x_1 coordinate(x coordinate), 3 is x_2 coordinate (y coordinate), and 5 is x_3 coordinate (z coordinate).

Distance of a point from Origin:
(a) In 2D:

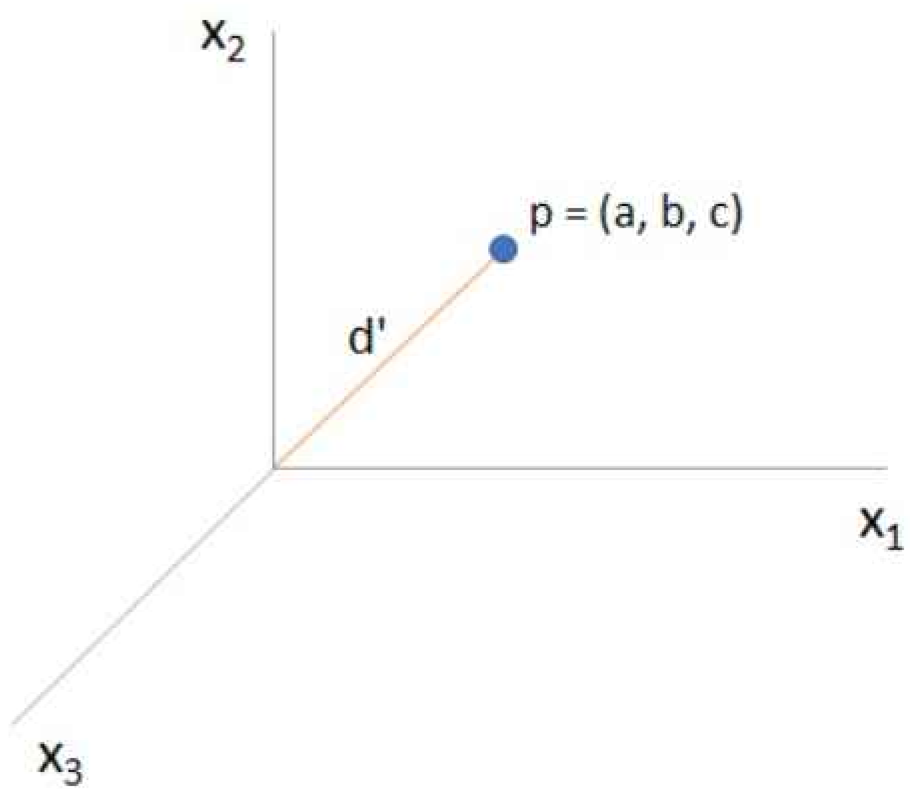


In 2D space, the distance d is given by,

$$d = \sqrt{(a^2 + b^2)} \left\{ \quad \text{where, } p = (a, b) \right.$$

This can be extended to 3D space and beyond for nD space.

(b) In 3D:



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

$$d' = \sqrt{(a^2 + b^2 + c^2)} \left\{ \quad \text{where } p = (a, b, c) \right.$$

(c) In nD:

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

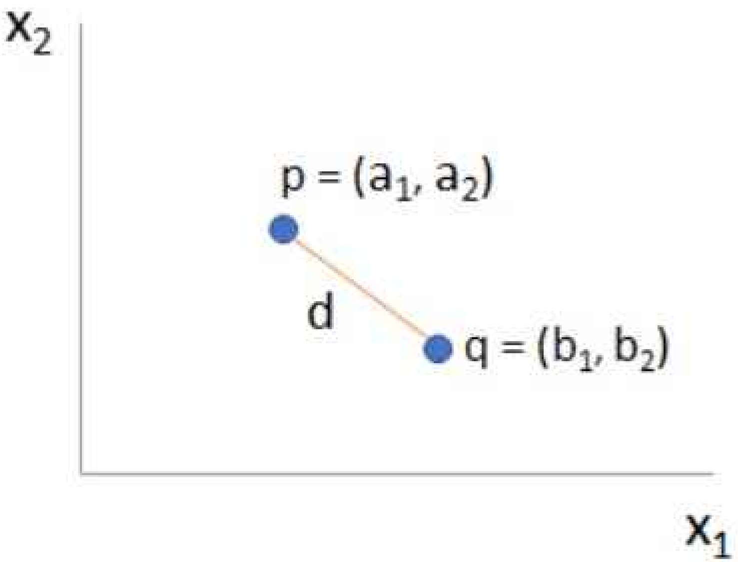
$$d = \sqrt{(a_1^2 + a_2^2 + ... + a_n^2)} \Big\{ \quad \text{where } p = (a_1, a_2, .., a_n)$$

Distance between two points

Consider we have two points say, p and q then the distance d for p=(a,b) is given by

$$d = \sqrt{((a_1 - b_1)^2 + (a_2 - b_2)^2)} \begin{cases} \text{where } p = (a_1, a_2) \\ \text{and } q = (b_1, b_2) \end{cases}$$

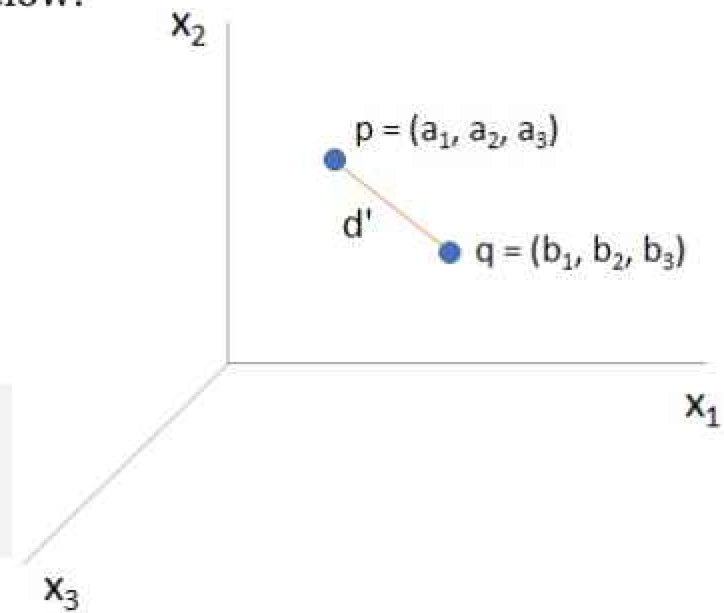
The image for 2D space is as shown below:



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

$$d = \sqrt{((a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2)} \begin{cases} \text{where } p = (a_1, a_2, a_3) \\ \text{and } q = (b_1, b_2, b_3) \end{cases}$$

The image for 3D space is as shown below:



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

$$d_{pq} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \begin{cases} \text{where } p = (a_1, a_2, ..., a_n) \\ \text{and } q = (b_1, b_2, ..., b_n) \end{cases}$$

Row-Vector Representation:

The row vector is a (1xn) matrix where the number of rows is 1 and the number of columns is ‘n’.

$$\begin{bmatrix} a_1 & a_2 & a_3 \ldots a_n \end{bmatrix}$$

Column-Vector Representation:

The column vector is an (nx1) matrix where the number of rows is ‘n’ and the number of columns is 1.

$$\begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{bmatrix}$$

Dot product of Vectors

Dot product

If we have two vectors **u** and **v**:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The dot product (or inner product) of these vectors is defined as the transpose of **u** multiplied by **v**:

$$\mathbf{u}.\mathbf{v} = \mathbf{u}^T\mathbf{v} = [u_1 u_2 \cdots u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Based on this definition the dot product is commutative so:

$$\mathbf{u}.\mathbf{v} = \mathbf{v}.\mathbf{u}$$

Dot product of Vectors

Lengh of vector **u**,**v**

Symbol for inner product

Angle between **u** and **v**

$$\mathbf{u} \bullet \mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos(\theta) \quad (1)$$

$$= x_1 \times x_2 + y_1 \times y_2 \quad (2)$$

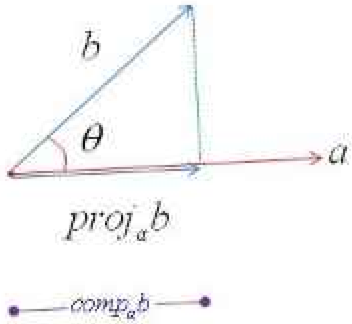
$$= \mathbf{u}\mathbf{v}^T \quad (3)$$

Transpose of vector **v**
(Why do we have to transpose ?)

Projection

Vector Projection

$$proj_a b = \frac{a \cdot b}{|a|^2} a$$

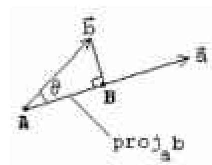


Scalar Projection

$$comp_a b = \frac{a \cdot b}{|a|}$$

Projections

One important use of dot products is in projections. The scalar projection of **b** onto **a** is the *length* of the segment AB shown in the figure below. The vector projection of **b** onto **a** is the *vector* with this length that begins at the point A points in the same direction (or opposite direction if the scalar projection is negative) as **a**.



Thus, mathematically, the scalar projection of **b** onto **a** is |**b**|cos(theta) (where theta is the angle between **a** and **b**) which from (*) is given by

$$comp_{\vec{a}}\vec{b} = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}|}$$

This quantity is also called the component of **b** in the **a** direction (hence the notation comp). And, the vector projection is merely the unit vector **a**/|**a**| times the scalar projection of **b** onto **a**:

$$proj_{\vec{a}}\vec{b} = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}|}\frac{\vec{a}}{|\vec{a}|} = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}|^2}\vec{a}$$

Thus, the scalar projection of **b** onto **a** is the magnitude of the vector projection of **b** onto **a**.

Projection of a Vector on another vector

The projection of a vector onto another vector is given as $proj_{\vec{v}}(\vec{u}) = \frac{\vec{u}.\vec{v}}{||\vec{v}||^2}\vec{v}$

```
# import numpy to perform operations on vector
import numpy as np
 u = np.array([1, 2, 3])   # vector u
v = np.array([5, 6, 2])   # vector v:

# Task: Project vector u on vector v

# finding norm of the vector v
v_norm = np.sqrt(sum(v**2))

# Apply the formula as mentioned above
# for projecting a vector onto another vector
# find dot product using np.dot()
proj_of_u_on_v = (np.dot(u, v)/v_norm**2)*v

print("Projection of Vector u on Vector v is: ", proj_of_u_on_v)
```

Output:

Projection of Vector u on Vector v is: [1.76923077 2.12307692 0.7076923

Projection of a Vector onto a Plane

The projection of a vector onto a plane is calculated by subtracting the component of which is orthogonal to the plane from $proj_{Plane}(\vec{u}) = \vec{u} - proj_{\vec{n}}(\vec{u}) = \vec{u} - \frac{\vec{u}.\vec{n}}{||\vec{n}||^2}\vec{n}$

where, $\vec{n}$ is the plane normal vector.

```
# import numpy to perform operations on vector
import numpy as np

u = np.array([1, 2, 3])   # vector u
v = np.array([5, 6, 2])   # vector v:

# Task: Project vector u on vector v

# finding norm of the vector v
v_norm = np.sqrt(sum(v**2))

# Apply the formula as mentioned above
# for projecting a vector onto another vector
# find dot product using np.dot()
proj_of_u_on_v = (np.dot(u, v)/v_norm**2)*v

print("Projection of Vector u on Vector v is: ", proj_of_u_on_v)
```

Output:

Projection of Vector u on Vector v is: [1.76923077 2.12307692 0.7076

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

Matrix Addition, Subtraction, Multiplication

Given the following: $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix}$

Matrix **Addition**:

$$A + \mathrm{B} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 0+1 & 1+0 \\ -2+3 & -3+(-2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -5 \end{bmatrix}$$

Matrix **Subtraction**:

$$A - \mathrm{B} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 0-1 & 1-0 \\ -2-3 & -3-(-2) \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -5 & -1 \end{bmatrix}$$

Matrix **Multiplication**:

$$A\mathrm{B} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix} = = \begin{bmatrix} 0\cdot 1 + 1\cdot 3 & 0\cdot 0 + 1\cdot(-2) \\ -2\cdot 1 - 3\cdot 3 & -2\cdot 0 - 3\cdot(-2) \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -11 & 6 \end{bmatrix}$$

Basic operations on matrix

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

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.

$$A = \begin{matrix} 1 & 0 \\ 2 & 3 \end{matrix}$$

$$B = \begin{matrix} 4 & -1 \\ 0 & 5 \end{matrix} \qquad C = \begin{matrix} 5 & -1 \\ 2 & 8 \end{matrix}$$

Matrix Multiplication

- In matrix multiplication the matrices don't need to be quadratic, but the inner dimensions need to be the same.
- The size of the resulting matrix will be the outer dimensions.

$$[A] \times [B] = [C]$$

$$(n \times m) \qquad (m \times p) \qquad (n \times p)$$

Inner dimensions need to be the same

The resulting matrix will be the **outer** dimensions

A

A(0,0)	A(0,1)
A(1,0)	A(1,1)

B

B(0,0)	B(0,1)
B(1,0)	B(1,1)

A(0,0) * B(0,0)	A(0,1) * B(0,1)
A(1,0) * B(1,0)	A(1,1) * B(1,1)

numpy.multiply(A, B)

```
import numpy as np

# first 2-D array arr1
arr1 = np.array([[3, 0], [0, 4]])
print("first array is :")
print(arr1)
print("Shape of first array is: ", arr1.shape)

# second 2-D array arr1
arr2 = np.array([1, 2])
print("second array is :")
print(arr2)
print("Shape of second array is: ", arr2.shape)

# calculating matrix product
res = np.matmul(arr1, arr2)
print("Resultant array is :")
print(res)
print("Shape of resultant array is: ", res.shape)
```

Dot Product

$$\begin{bmatrix} a & b \end{bmatrix} \bullet \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \end{bmatrix}$$

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

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \bullet \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{bmatrix}$$

Dot Product

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \bullet \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} (1\cdot 5)+(2\cdot 7) & (1\cdot 6)+(2\cdot 8) \\ (3\cdot 5)+(4\cdot 7) & (3\cdot 6)+(4\cdot 8) \end{bmatrix}$$

$$= \begin{bmatrix} 5+14 & 6+16 \\ 15+28 & 18+32 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

DoT Matrix

```
A = np.random.randint(5, size=(3,2))
B = np.random.randint(5, size=(2,3))
```

```
A
```

```
array([[2, 2],
       [0, 3],
       [0, 4]])
```

```
B
```

```
array([[2, 1, 2],
       [3, 2, 4]])
```

```
np.dot(A,B)
```

```
array([[10,  6, 12],
       [ 9,  6, 12],
       [12,  8, 16]])
```

Transpose of a Matrix

A general matrix is given by:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Where n is number of rows and m is number of columns

$(n \times m)$

The transpose of matrix A is then:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}$$

$(m \times n)$

Transpose of a Matrix - Examples

$$A = \begin{bmatrix} 1 & 3 & 7 & 2 \\ 5 & 8 & -9 & 0 \\ 6 & -7 & 11 & 12 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 1 & 5 & 6 \\ 3 & 8 & -7 \\ 7 & -9 & 11 \\ 2 & 0 & 12 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 5 \\ 4 & 5 \\ 3 & 2 \\ 7 & 8 \end{bmatrix} \Rightarrow B^T = \begin{bmatrix} 1 & 4 & 3 & 7 \\ 5 & 5 & 2 & 8 \end{bmatrix}$$

Determinant of a Matrix

- The Determinant of a matrix is a special number that can be calculated from square matrices

What is the Determinant used for?

- The determinant helps us find the inverse matrix (which we will cover later)
- The Determinant will give us useful information when dealing with Systems of Linear Equations (which we will cover later)
- Used in advanced Control Engineering theory
- Etc.

Determinant of a Matrix

Given a matrix A the Determinant is given by:

$$det(A) = |A|$$

For a $2x2$ matrix we have:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \quad det(A) = |A| = a_{11}\,a_{22} - a_{21}\,a_{12}$$

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \Rightarrow \quad det(A) = |A| = 1 \cdot 4 - 3 \cdot 2 = 4 - 6 = \underline{\underline{-2}}$$

Determinant of a Matrix

Given the following Matrices:

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

$$det(A) = -2$$

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

$$det(B) = -21$$

Python Solution:

```
-2.0000000000000004
-21.00000000000001
```

```
import numpy as np
import numpy.linalg as la

A = np.array([[1, 2],
              [3, 4]])

Adet = la.det(A)

print(Adet)

B = np.array([[-1, 3, 0],
              [2, 1, -5],
              [1, 4, -2]])

Bdet = la.det(B)

print(Bdet)
```

Determinant of a Matrix

For a $3x3$ matrix we have:

We develop the determinant along a row or a column.

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

$$\Rightarrow \quad det(A) = |A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Here we have developed the determinant along the first column

We see that the determinant of a higher order system can be expressed as a sum of lower order determinants

It will be a little more complicated and time-consuming work for systems with larger order, so we need a programming language like python to handle this.

Inverse Matrices

The **inverse** of a quadratic matrix A is defined by: A^{-1}

Note: $AA^{-1} = A^{-1}A = I$

For a $2x2$ matrix we have:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

The inverse A^{-1} is then given by

$$A^{-1} = \frac{1}{det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \Rightarrow A^{-1} = \frac{1}{det(A)} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$$

Where:

$$det(A) = |A| = 1 \cdot 4 - 3 \cdot 2 = 4 - 6 = -2$$

This gives:

$$A^{-1} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$$

It will be more complicated for systems with larger order, so we need a programming language like python to handle this.

Linear Equations

Given the following linear equations:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots = b_1$$
$$a_{21}x_1 + a_{21}x_2 + a_{23}x_3 + \cdots = b_2$$
$$\cdots$$

These equations can be set on the following general form:

$$Ax = b$$

Where A is a matrix, x is a vector with the unknowns and b is a vector of constants

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Solution:

$$x = A^{-1}b$$

(assuming A^{-1} is possible)

Null space for data science

- The null space of a matrix **A** consists of all vectors $\boldsymbol{\beta}$ such that $\boldsymbol{A\beta} = \mathbf{0}$ and $\boldsymbol{\beta} \neq \mathbf{0}$
- Nullity of a matrix is the number of vectors in the null space of the given matrix
- The size of the null space of a matrix provides us with the number of linear relations among the attributes
- And the null space vectors $\boldsymbol{\beta}$ are useful to identify these linear relationships

Null space: The idea

- Notice that if $\boldsymbol{A\beta} = \mathbf{0}$, every row of A when multiplied by $\boldsymbol{\beta}$ goes to zero
- This implies that variable values in each sample (represented by a row) behave the same
- This helps in identifying the linear relationships in the attributes
- Every null space vector corresponds to one linear relationship
- This idea is demonstrated further using examples

Null space : general description

- Let us suppose
- $A = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$ is a data matrix and there is one vector in the null space of A, i.e, $\beta = [\beta_1 \ \ldots \beta_m]^T$, then as per the definition, β satisfies all the equations given below

- $x_{11}\beta_1 + x_{12}\beta_2 + \cdots x_{1n}\beta_n = 0$

$$\vdots$$

- $x_{m1}\beta_1 + x_{m2}\beta_2 + \cdots x_{mn}\beta_n = 0$

Null space: An Example

- Consider the matrix A with attributes $\{x_1, x_2\}$

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

Number of columns in A =2

Rank of A =2

Thus, nullity =0

- This implies that the null space of the matrix A does not contain any vectors
- Thus we can claim that all the attributes are linearly independent

Null space: Another example

- Now consider A with attributes $\{x_1, x_2, x_3\}$ such that

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

Number of columns in A =3

Rank of A =2

Thus, nullity =1

- Thus, we need to identify the vectors in the null space of A which is non-zero in this case

Null space: Further Example

$$A\boldsymbol{\beta} = 0$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- Thus we obtain,

$$b_1 + 2b_2 = 0$$
$$b_3 = 0$$

- The null vector is $\boldsymbol{B} = [b_1 \ \ b_2 \ \ b_3 \ ]^T = [-2b_2 \ \ b_2 \ \ 0]^T = k[-2 \ 1 \ 0]^T$
- We see that we obtain a direct linear relationship between the attributes of A using null space and rank-nullity theorem
- The same concept can be extended for bigger data set

- **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.

Consider the following matrix as an example that we want to calculate its *Null-Space*:

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 2 & 2 & -1 \\ 2 & 4 & 0 & 6 \end{bmatrix}$$

To find $N(A)$ we should find all vectors x *such that:*

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \ A\vec{x} = \vec{0} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 2 & 2 & -1 \\ 2 & 4 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

By multiplying matrix A to vector x we have:

$$\begin{bmatrix} x_1 + 2x_2 + x_3 + x_4 \\ x_1 + 2x_2 + 2x_3 - x_4 \\ 2x_1 + 4x_2 + 6x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, we should solve this equation by considering it as a system of linear equations:

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 = 0 \\ x_1 + 2x_2 + 2x_3 - x_4 = 0 \\ 2x_1 + 4x_2 + 6x_4 = 0 \end{cases}$$

Therefore, the solution to the above equation is:

$$\begin{cases} x_1 + 2x_2 + 3x_4 = 0 \\ x_3 - 2x_4 = 0 \end{cases} \rightarrow \begin{cases} x_1 = -2x_2 - 3x_4 \\ x_3 = 2x_4 \end{cases}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

The x2 and x4 are free variables in $\mathbb{R}$. So, all of the vectors in *Null-Space(A)*, which satisfy the original equation, *Ax=0,* can be represented as a linear combination of these two vectors.

Recall: The linear combinations of two vectors is the span of those two vectors.

Hence, it is true to say that the *Null-Space(A)* is:

$$Null-Space(A) = Span\left(\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-3\\0\\2\\1\end{bmatrix}\right)$$

Using Null-Space to show linearly independence of the column vectors of a matrix

If the column vectors of a matrix are ***linearly independent***, then the *Null-Space* of that matrix is only going to consist of the zero vector. Also, If the *Null-Space* of a matrix only contains the zero vector, that means that the columns of that matrix are linearly independent. To better understanding let us do a little math! Consider matrix *A(n*m)*. We can rewrite it as its column vectors (A1 to An) such as below:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left[\overrightarrow{A_1}\ \overrightarrow{A_2} \ldots \overrightarrow{A_n}\right]$$

To calculate the *Null-Space*(A), we should find all vectors *x* that satisfy *Ax=0:*

$$A\vec{x} = \vec{0} \rightarrow \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\overrightarrow{A_1} + x_2\overrightarrow{A_2} + \ldots + x_n\overrightarrow{A_n} = 0$$

Recall: The vectors v1,v2, …,vn are ***linearly independent*** if and only if the only solution to the equation of their linear combination to be zero is c1=c2=…=cn=0.

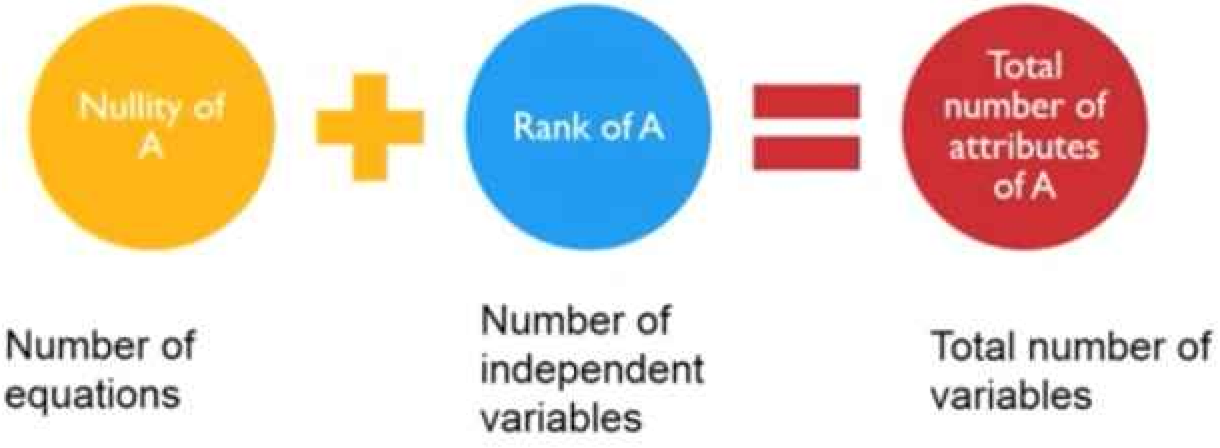
$$\{v_1, v_2, \ldots, v_n\} \text{ are linearly independent}$$
$$\Updownarrow$$
$$c_1 v_1 + c_2 v_2 + \ldots, c_n v_n = \vec{0}$$
$$\downarrow$$
$$c_i = 0 \text{ for } 1 \leq i \leq n \; (c_1 = c_2 = \ \ldots \ = c_n = 0)$$

Therefore, if the only solution for *Ax=0 is* x_i*=0 (for 1≤i≤n),* we can say that the column vectors of *A* are linearly independent. In other words:

$$Null - Space(A) = \{\vec{0}\}$$
$$\Updownarrow$$
$$x_1 = x_2 = \ldots = x_n = 0$$
$$\Updownarrow$$
$$\overrightarrow{A_1}, \overrightarrow{A_2}, \ldots, \overrightarrow{A_n} \text{ are linearly independent}$$

Rank nullity theorem

- Consider the data matrix A with the null space and nullity as defined before
- 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
- According to the rank-nullity theorem



Summary till now

Data Matrix

- The available data is expressed in the form of a data matrix
- This data matrix is further used to do the necessary operations

Null Space

- Defined as a collection of vectors satisfying $A\beta = 0$
- Helps in identifying the linear relationships between the attributes directly

Nullity

- Nullity is the size of the null space of the data matrix
- Useful to identify the number of linear relationships in the attributes
- Rank- Nullity theorem

Hyperplanes

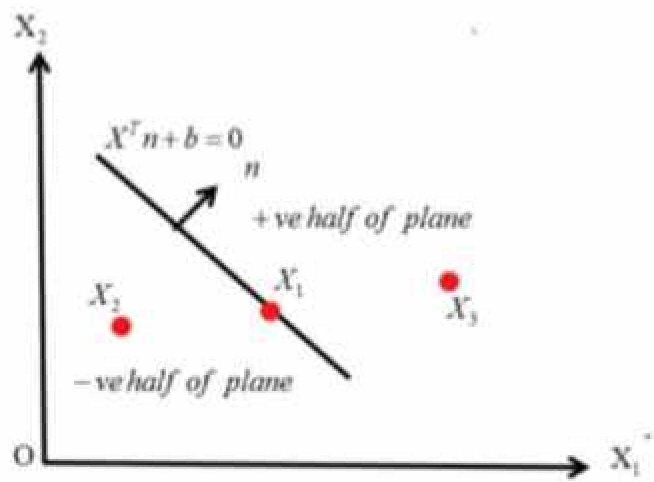
- Geometrically, hyperplane is a geometric entity whose dimension is one less than that of its ambient space.
- For instance, the hyperplanes for a 3D space are 2D planes and hyperplanes for a 2D space are 1D lines and so on.
- The hyperplane is usually described by an equation as follows

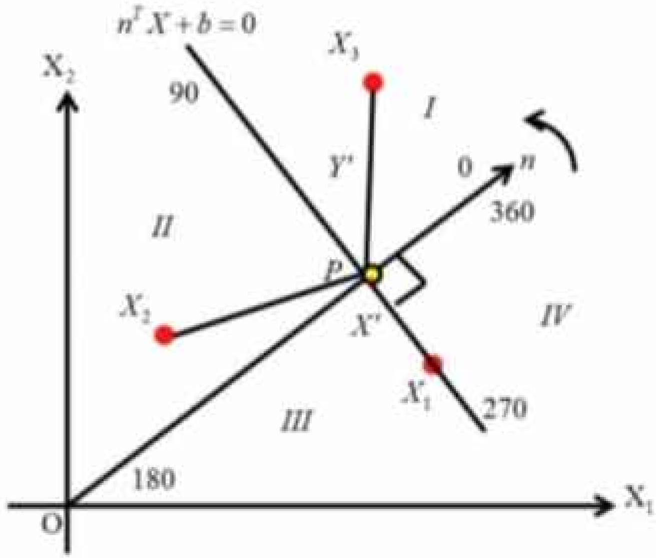
$$X^T n + b = 0$$

Halfspace

- We can observe that the equation can be evaluated for the two halfspaces
- It can be seen that

$$X^T n + b = 0 \ \forall\ X \in line$$
$$X^T n + b > 0 \ \forall\ X \in subspace\ in\ the\ n\ direction\ (X_3)$$
$$X^T n + b < 0 \ \forall\ X \in subspace\ in\ the\ -n\ direction\ (X_2)$$





Hyperplanes and halfspaces: Example

- Let us consider a 2D geometry with $n = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $b = 4$

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$X^T n + b = 0$$

$$[x_1 \; x_2] \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 4 = 0$$

$$x_1 + 3x_2 + 4 = 0$$

- The hyperplane is the equation of a line
- The halfspaces corresponding to this hyperplane are

$$x_1 + 3x_2 + 4 > 0 : \text{Positive halfspace}$$

$$x_1 + 3x_2 + 4 < 0 : \text{Negative halfspace}$$

Eigenvalues and eigenvectors

- We have previously seen linear equations of the form $Ax = b$
- What is the geometrical interpretation of this equation?
- We can make an interpretation as follows
  - When vector x is operated on by A, we obtain a new vector b with a different orientation

Eigenvalues and eigenvectors

- Operator representation

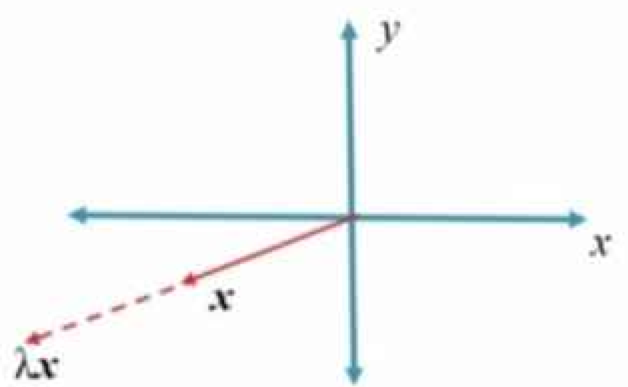
$$\xrightarrow{x} \boxed{\mathbf{A}} \longrightarrow \mathbf{A}x = \mathbf{b}$$

- The newly obtained b vector represents a new orientation. So we ask the following question
- Are there directions for a matrix A such that when the matrix operates on these directions they maintain their orientation save for multiplication by a scalar (positive or negative)?
- That is

$$\xrightarrow{x} \boxed{\mathbf{A}} \longrightarrow \mathbf{A}x = \mathbf{b} = \lambda x$$

Eigenvalues and eigenvectors

- The mathematical formulation of our question is
$$Ax = \lambda x$$
- The constant λ ($positive$) represents the amount of stretch or shrinkage the attributes x go through in the x direction
- The solutions (x) are known as eigenvectors and their corresponding λ are eigenvalues



Eigenvalues and eigenvectors

- We can find the eigenvalues as follows

$$Ax = \lambda x \quad A(n \times n); x(n \times 1)$$
$$Ax - \lambda I x = 0$$
$$(A - \lambda I)x = 0$$

- Thus the eigenvalues of the equation can be identified using

$$|A - \lambda I| = 0$$

- Substituting the eigenvalues in the original equation will help us find solutions for the eigenvector x

Eigenvalues and eigenvectors: Examples

- Consider the following example with the given A matrix

$$A = \begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$$

$$|A - \lambda I| = \left| \begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \left| \begin{bmatrix} 8-\lambda & 7 \\ 2 & 3-\lambda \end{bmatrix} \right|$$
$$= 0$$

$$(8 - \lambda)(3 - \lambda) - 14 = 0$$
$$\lambda^2 - 11\lambda + 10 = 0$$
$$\lambda = (10,1)$$

- Thus we identify two eigenvalues and now we proceed to find the corresponding eigenvectors

Eigenvalues and eigenvectors: Examples

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- $\lambda = 1$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8x_1 + 7x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x_1 + x_2 = 0$$

- Thus the eigenvector (unit) corresponding to $\lambda = 1$is

$$X = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$$

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Eigenvalues and eigenvectors: Examples

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- $\lambda = 1$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8x_1 + 7x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x_1 + x_2 = 0$$

$8x_1 + 7x_2 = x_1$

$7x_1 + 7x_2 = 0 \Rightarrow x_1 + x_2 = 0$ ✓

$2x_1 + 3x_2 = x_2$

$2x_1 + 2x_2 = 0 \to x_1 + x_2 = 0$

- Thus the eigenvector (unit) corresponding to $\lambda = 1$is

$$X = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \qquad \begin{bmatrix} k \\ -k \end{bmatrix}$$

$\sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{-1}{\sqrt{2}}\right)^2}$

$\sqrt{y_1^2 + y_2^2} =$

Eigenvalues and eigenvectors: Examples

- $\lambda = 10$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10x_1 \\ 10x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8x_1 + 7x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} 10x_1 \\ 10x_2 \end{bmatrix}$$

$$7x_2 = 2x_1$$

- Thus the eigenvector (unit) corresponding to $\lambda = 10$

$$X = \begin{bmatrix} \frac{7}{\sqrt{53}} \\ \frac{2}{\sqrt{53}} \end{bmatrix}$$

Code for finding EigenVectors in python

```
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)
```

Singular Value Decomposition

In linear algebra, the Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices.

$$A=UDV^T$$

where the columns of U and V are orthonormal and the matrix D is diagonal with positive real entries.

- Only the square matrix has eigenvalues, the rectangular matrix can have ***singular values***, a square matrix whose determinant = 0 will be a ***singular matrix***, or in another word, it is non-invertible.
- A complex matrix will be ***unitary*** if its conjugate transpose happens to be its inverse. U*U = UU* = I. The real value counterpart is the ***orthogonal matrix***.
- So a rectangular matrix A can be decomposed into three components: U, S, and Vh, diagonal of S matrix contain ***singular values*** for matrix A. This decomposition is called ***Singular Value Decomposition (SVD)***.

How do you find *eigenvalues* of a matrix? Could you provide an example?

Answer

Given a matrix **A**, we find its *eigenvalues* λ by solving the equation:

$$\det(\lambda I - A) = 0$$

For example, given the following matrix,

$$A = \begin{pmatrix} -5 & 2 \\ -7 & 4 \end{pmatrix}$$

we determine its eigenvalues in the following way:

$$\det(\lambda I - A) = \left| \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -5 & 2 \\ -7 & 4 \end{pmatrix} \right|$$

$$= \left| \begin{pmatrix} \lambda + 5 & -2 \\ 7 & \lambda - 4 \end{pmatrix} \right|$$

Now the characteristic polynomial is:

$\lambda^2 + \lambda - 6 = 0$

The solutions of this equation and therefore the eigenvalues are then,

$\lambda_1 = 2, \lambda_2 = -3$

How will you calculate eigenvalues and eigenvectors of the following 3x3 matrix?

-2	-4	2
-2	1	2
4	2	5

The characteristic equation is as shown:

Expanding determinant:

$(-2 - \lambda) [(1-\lambda) (5-\lambda)-2x2] + 4[(-2) \times (5-\lambda) -4x2] + 2[(-2) \times 2-4(1-\lambda)] =0$

$- \lambda 3 + 4\lambda 2 + 27\lambda - 90 = 0,$

$\lambda 3 - 4 \lambda 2 -27 \lambda + 90 = 0$

Here we have an algebraic equation built from the eigenvectors.

By hit and trial:

$33 - 4 \times 32 - 27 \times 3 +90 = 0$

Hence, $(\lambda - 3)$ is a factor:

$\lambda 3 - 4 \lambda 2 - 27 \lambda +90 = (\lambda - 3) (\lambda 2 - \lambda - 30)$

Eigenvalues are 3,-5,6:

Calculate eigenvector for $\lambda = 3$

For $X = 1$,

$-5 - 4Y + 2Z = 0$,

$-2 - 2Y + 2Z = 0$

Subtracting the two equations:

$3 + 2Y = 0$,

Subtracting back into second equation:

$Y = -(3/2)$

$Z = -(1/2)$

Similarly, we can calculate the eigenvectors for -5 and 6.

What is $Ax = b$? When does $Ax = b$ has a *unique* solution?

Answer

Ax = b is a system of linear equations expressed in matrix notation, in which:

- **A** is the coefficient matrix of order *m x n*.
- **x** is the incognite variables vector of order *n x 1*.
- **b** is the vector formed by the constants and its order is *m x 1*.

The system **Ax = b** has a unique solution if and only if

$$rank[A] = rank[A|b] = n$$

where the matrix `A|b` is matrix **A** with **b** *appended* as an extra column.

When are two vectors `x` and `y` orthogonal?

Answer

Two vectors are said to be orthogonal if the dot product of them is equal to zero,

$$\vec{x} \cdot \vec{y} = 0$$

This is because the definition of the dot product:

$$\vec{x} \cdot \vec{y} = |x||y| \cos(\theta)$$

where `θ` is the angle between the two vectors, therefore if *x* and *y* are orthogonal, the angle between them is `90` and `cos(90) = 0`.

Find the *inverse* of the following matrix

Problem

Consider the matrix:

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

Answer

We will compute inverse using the following equation: $A^{-1} = \frac{1}{det(A)} adj(A)$ where *adj(A)* is the *adjugate* of matrix *A*. Now we follow the next steps:

1. **Calculate the determinant of *A*:**

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

$$= 4(0-6) + 2(30+3) + 1(10-0)$$

$$= 52 \neq 0 \therefore \text{ the matrix is invertible}$$

2. **Calculate the cofactor** of each element:

$$C_{11} = (-1)^{1+1}\begin{vmatrix} 0 & 3 \\ 2 & 6 \end{vmatrix} = -6,\ C_{12} = (-1)^{1+2}\begin{vmatrix} 5 & 3 \\ -1 & 6 \end{vmatrix} = -33,$$

$$C_{13} = (-1)^{1+3}\begin{vmatrix} 5 & 0 \\ -1 & 2 \end{vmatrix} = 10,\ C_{21} = (-1)^{2+1}\begin{vmatrix} -2 & 1 \\ 2 & 6 \end{vmatrix} = 14,$$

$$C_{22} = (-1)^{2+2}\begin{vmatrix} 4 & 1 \\ -1 & 6 \end{vmatrix} = 25,\ C_{23} = (-1)^{2+3}\begin{vmatrix} 4 & -2 \\ -1 & 2 \end{vmatrix} = -6,$$

$$C_{31} = (-1)^{3+1}\begin{vmatrix} -2 & 1 \\ 0 & 3 \end{vmatrix} = -6,\ C_{32} = (-1)^{3+2}\begin{vmatrix} 4 & 1 \\ 5 & 3 \end{vmatrix} = -7,$$

$$C_{33} = (-1)^{3+3}\begin{vmatrix} 4 & -2 \\ 5 & 0 \end{vmatrix} = 10.$$

Thus, the cofactor matrix is: $C = \begin{bmatrix} -6 & -33 & 10 \\ 14 & 25 & -6 \\ -6 & -7 & 10 \end{bmatrix}$

3. **Obtain the adjugate matrix** by transposing cofactor matrix

$$adj(A) = \begin{bmatrix} -6 & 14 & -6 \\ -33 & 25 & -7 \\ 10 & -6 & 10 \end{bmatrix}$$

3. **Obtain the adjugate matrix** by transposing cofactor matrix

$$adj(A) = \begin{bmatrix} -6 & 14 & -6 \\ -33 & 25 & -7 \\ 10 & -6 & 10 \end{bmatrix}$$

4. Finally, the inverse matrix is the **adjugate matrix divided by the determinant**:

$$A^{-1} = \frac{1}{52} \cdot \begin{bmatrix} -6 & 14 & -6 \\ -33 & 25 & -7 \\ 10 & -6 & 10 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -3/26 & 7/26 & -3/26 \\ -33/52 & 25/52 & -7/52 \\ 5/26 & -3/26 & 5/26 \end{bmatrix}$$

How do you *diagonalize* a matrix?

Given an nxn matrix **A**, to find its diagonal matrix **D** we must follow the next steps:

1. Find the characteristic polynomial of **A**.
2. Find the roots of the characteristic polynomial to obtain the eigenvalues λ of **A**.
3. For each eigenvalue λ of **A**, find their correspondent eigenvectors.
4. If the total number of eigenvectors *m* found in step 3 is not equal to *n* (the numbers of rows and columns of **A**), then the matrix is not diagonalizable, but if $m = n$ then the diagonal matrix **D** is given by:

$$D = P^{-1}AP,$$

where **P** is a matrix which columns are the eigenvectors of the matrix **A**.

How do you find the inverse of a 2x2 matrix?

Answer

For an arbitrary *A* matrix, we can derive it's inverse by following the next steps:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1. Check if the matrix if invertible by finding its determinant :

$$|A| = ad - bc.$$

   If $|A| \neq 0$ then the matrix is invertible.

2. Interchange the two elements on the diagonal.
3. Take the negatives of the other two elements out of the diagonal.
4. Divide each element of the matrix by $|A|$. The result of the inverse of the matrix *A* is then:

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ c & a \end{pmatrix}$$

What is the *determinant* of a square matrix? How is it calculated?

Answer

The **determinant** is a scalar value that is a *function of the entries* of a square matrix. The determinant of a matrix **A** is denoted **det(A)**, **det A**, or **|A|**. Geometrically, it can be viewed as the *volume scaling factor* of the *linear transformation* described by the matrix.

In the case of a 2×2 matrix the determinant is calculated following the next diagram:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$$

That is, the determinant is equal to the *product* of the elements along the *plus-labeled arrow* minus the *product* of the elements along the *minus-labeled arrow*.

Similarly, for a 3×3 matrix **A**, its determinant is

$$\begin{aligned} |A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= aei + bfg + cdh - ceg - bdi - afh. \end{aligned}$$

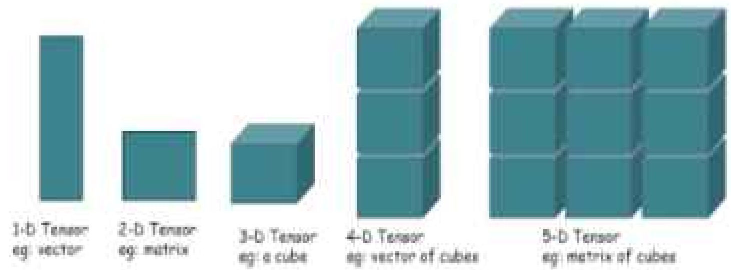
Each determinant of a 2×2 matrix in the equation above is called a *minor* of the matrix **A**.

For an $n \times n$ matrix, the previous procedure is extended and provides a *recursive definition* for the determinant, known as a *Laplace expansion*.

Answer

- In simple terms, a **matrix** is a grid of `n × m` (say, `3 × 3`) numbers surrounded by brackets. We can add and subtract matrices of the same size, multiply one matrix with another as long as the sizes are *compatible*, and multiply an entire matrix by a constant.

- A *tensor* is a *generalization of matrices* to N-dimensional space. That is, it could be a `1-D` matrix (a vector), a `3-D` matrix (something like a cube of numbers), or even a `0-D` matrix (a single number), etc. The dimension of the tensor is called *rank*.



Tensors with different dimensions

- A **tensor** can be also seen as a mathematical entity that lives in a structure and interacts with other mathematical entities. If one *transforms* the other entities in the structure in a *regular way*, then the tensor obeys a related *transformation rule*. This *dynamical* property of a tensor is also a key to distinguish it from a mere matrix. For example, any rank- `2` tensor can be represented as a matrix, but not every matrix is really a rank- `2` tensor. The difference depends on the *transformation rules* that have been applied to the entire system.

1. What is Ax=b? How to solve it?
2. How do we multiply matrices?
3. What is an Eigenvalue? And what is an Eigenvector? What is Eigenvalue Decomposition or The Spectral Theorem?
4. What is Singular Value Decomposition?

Basic Equations of Lines

- 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

$$\mathbf{y = mx + b.}$$

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 evenhandedly. The general equation for a line (normal form) is

$$\mathbf{ax + by = c,}$$

- This can easily be converted to slope-intercept form by solving for y:

$$\mathbf{y = (-a/b) + c/b,}$$

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

- If the coefficients on the normal form are multiplied by a nonzero constant, the set of solutions is exactly the same, so, for example, all these equations have the same line as solution.

$$2x + 3y = 4$$
$$4x + 6y = 8$$
$$-x - (3/2)y = -2$$
$$(1/2)x + (3/4)y = 1$$

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 = c of line PQ.

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

$$a1 + b2 = c, \text{ or } a + 2b = c$$

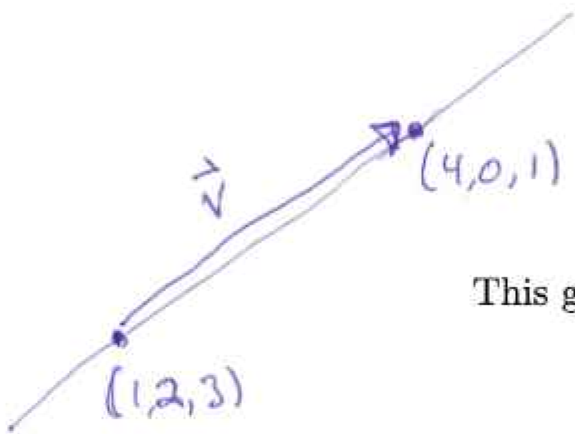
Since Q is on the line, its coordinates satisfy the equation: a(-2) + b5 = c,

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**.

Example

- Find the equation of the line that contains the points(1,2,3)and(4,0,1). We have a point on the line (actually two, but it suffices to pick one). To find the direction of the line, we find the displacement vector from (1,2,3) to (4,0,1).



This is $\vec{v} = \langle 4-1, 0-2, 1-3 \rangle = \langle 3, -2, -2 \rangle$.

This gives us the equation

$$\langle x, y, z \rangle = \langle 1, 2, 3 \rangle + t\langle 3, -2, -2 \rangle$$

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

Equation of a Plane

- A plane in 3-space has the equation

$$\mathbf{ax + by + cz = d,}$$

- 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:

$$\mathbf{z = -(a/c)x + (-b/c)\,y + d/c}$$

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

$$\mathbf{(a/d)x + (b/d)y + (c/d)z = 1.}$$

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 first 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

$$x + 2y + 2z = 5$$

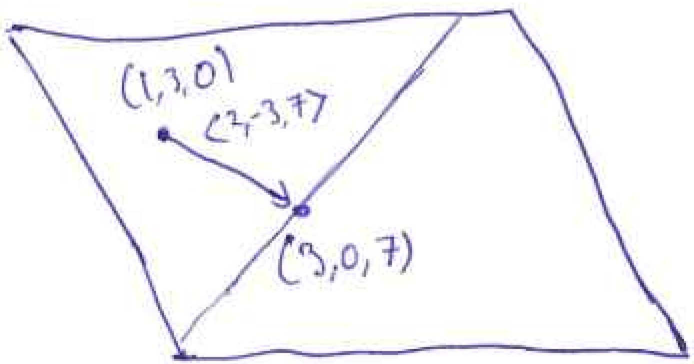
- or another choice would be $(1/5)x + (2/5)y + (2/5)z = 1$

Example

- Find the equation of the plane that contains the point(1,3,0)and the line given by

$$x= 3 + 2t, y=-4t, z= 7-t.$$

We know a point on the line is (1,3,0). The line has direction $\langle 2,-4,-1\rangle$, so this lies parallel to the plane. Now we need another direction vector parallel to the plane. Plugging t= 0 into the line equation gives the point (3,0,7), which is also on the plane. Then the vector from(1,3,0) to (3,0,7) will also lie parallel to the plane; this vector is $\langle 2,-3,7\rangle$



This gives us the following parametric equation of a plane

$$\langle x, y, z\rangle = \langle 1, 3, 0\rangle + t\langle 2, -4, -1\rangle + s\langle 2, -3, 7\rangle$$

To find the implicit formula, we must find a vector orthogonal/normal to the plane. This vector can be computed via the cross product

$$\vec{n} = \langle 2, -4, -1\rangle \times \langle 2, -3, -7\rangle = \langle 31, 12, 2\rangle$$

This gives us with the implicit equation

$$\begin{aligned} \langle 31, 12, 2\rangle \cdot \langle x-1, y-3, z\rangle &= 0 \\ 31(x-1) + 12(y-3) + 2z &= 0 \\ 31x + 12y + 2z &= 67 \end{aligned}$$

$$\langle x, y, z\rangle = \langle 3, 0, 7\rangle + t\langle 4, -7, 6\rangle + s\langle -2, 3, -7\rangle$$

and

$$-31(x-5) + -12(y+4) + -2(z-6) = 0$$

give the same plane. This is harder than in #1 because planes give more options. Notice, however, that analyzing the implicit equation makes this clearer.

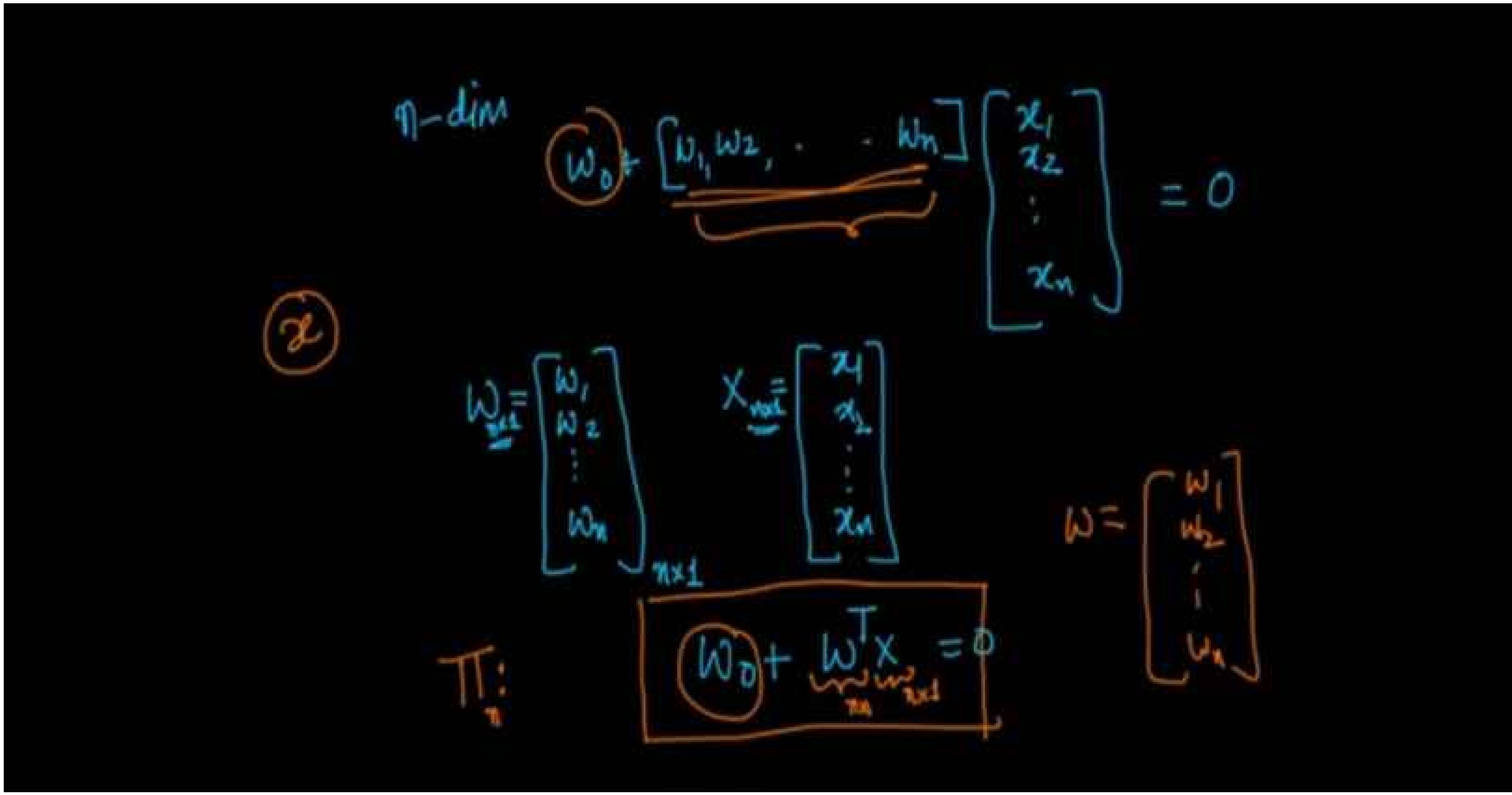
Equation of Hyperplane

- A hyperplane in an n−dimensional vector space satisfying the equation:

$$a_1x_1+\cdots+a_nx_n=b$$

Where $a_1,\ldots,a_n$ and b are real numbers with at least $a_1,\ldots,a_n$ non-zero.

Equation of Hyperplane



l passing through origin

line: 2D: $w_1x_1 + w_2x_2 = 0$

plane: 3D: $w_1x_1 + w_2x_2 + w_3x_3 = 0$

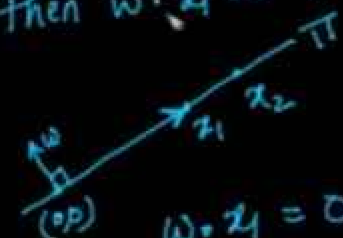
hyperplane: nD: $w_1x_1 + w_2x_2 + \cdots + w_nx_n = 0$

$$\boxed{w^T x = 0}$$ — eqn. of a plane passing through origin

$$\boxed{w^T x + w_0 = 0}$$

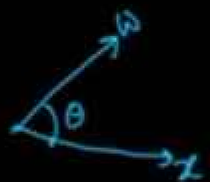
$$\Pi_n : \boxed{w^T x = 0}$$

If $w \perp \Pi$
then $w \cdot x_i = 0 \;\; \forall \, x_i \in \Pi$



$$w \cdot x_1 = 0$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

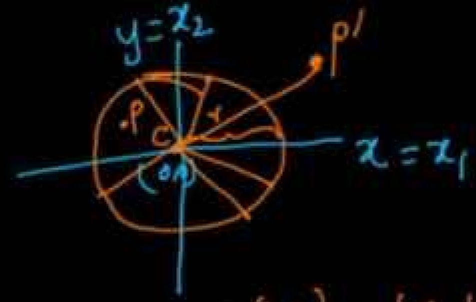


$$w \cdot x = w^T x = \|w\| \, \|x\| \cos \theta_{w,x} = 0$$

$$w \perp x \Rightarrow \theta_{w,x} = 90^\circ$$

Equation of Circle

Circle:



$x^2 + y^2 = r^2$

$c : (h, k)$; $(x-h)^2 + (y-k)^2 = r^2$

$p(x_1, x_2)$

$$\begin{cases} x_1^2 + x_2^2 < r^2 \Rightarrow p \text{ lies inside the circle} \\ x_1^2 + x_2^2 > r^2 \Rightarrow p \text{ lies outside the circle} \\ x_1^2 + x_2^2 = r^2 \Rightarrow p \text{ lies on the circle} \end{cases}$$

Sphere:

Sphere is the locus of points which are at constant distance from a fixed point known as Centre of the Sphere and the constant distance is known as Radius of the sphere.

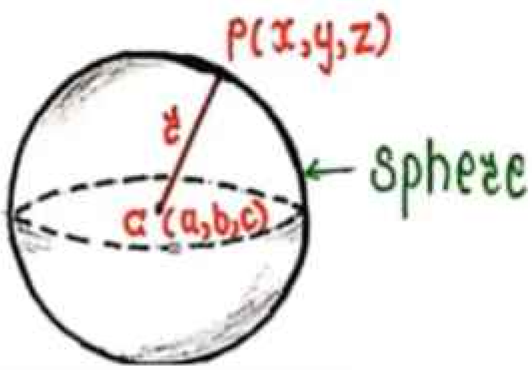
Equations of Sphere:

A) Centre-Radius form:

Let P(x, y, z) be any point on the sphere, C(a, b, c) be the centre of the sphere and r be the radius.

Then by distance formula,

$$(x-a)^2+(y-b)^2+(z-c)^2=r^2$$

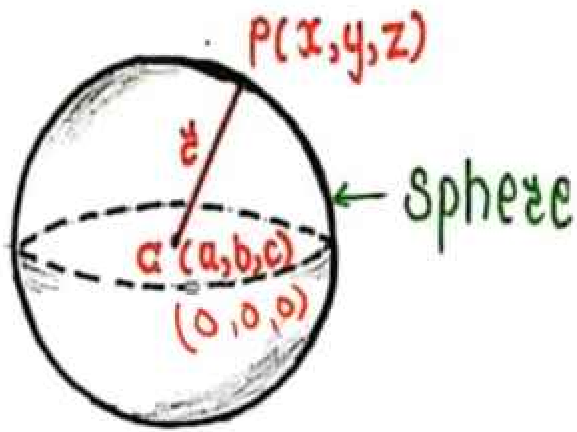


Equations of Sphere:

A) Centre-Radius form:

Note: If C$\equiv (a, b, c) \equiv (0,0,0)$ i.e. be the centre of the sphere is at origin and r be the radius.

Then equation of the sphere is,



$$x^2 + y^2 + z^2 = r^2$$

Equations of Sphere:

B) General form:

Let $P(x, y, z)$ be any point on the sphere, $C(a, b, c)$ be the centre of the sphere and r be the radius.

Then equation of sphere by Centre-Radius form is:

$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$

$\Rightarrow x^2 + y^2 + z^2 - 2ax - 2by - 2cz + (a^2 + b^2 + c^2 - r^2) = 0$

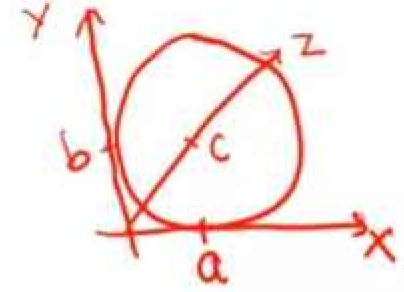
Put $a = -u, b = -v, c = -w$ & $d = a^2 + b^2 + c^2 - r^2$

$\therefore$ Equation of the sphere becomes,

$$\boxed{x^2 + y^2 + z^2 + 2ux + 2vy + 2wz + d = 0}$$

Centre $C \equiv (a, b, c) \equiv (-u, -v, -w)$ & Radius: $r = \sqrt{u^2 + v^2 + w^2 - d}$

Equations of Sphere:



C) Intercept form:

Consider the equation of the sphere in general form

$$\boldsymbol{x^2 + y^2 + z^2 + 2ux + 2vy + 2wz + d = 0} \quad \ldots . (1)$$

Suppose this sphere cuts off the co-ordinate axes at a, b, c respectively and also it is passing through the Origin i.e. $O \equiv (0,0,0)$.

$\therefore O \equiv (0,0,0)$ satisfies eq.(1), which gives $d = 0$.

$\therefore$ eq.(1) becomes:

$$\boldsymbol{x^2 + y^2 + z^2 + 2ux + 2vy + 2wz = 0} \quad \ldots .(2)$$

Equations of Sphere:

C) Intercept form:

$\therefore$ eq.(1) becomes:

$$\boldsymbol{x^2 + y^2 + z^2 + 2ux + 2vy + 2wz = 0} \quad ....(2)$$

Which is the equation of sphere passing through Origin.

Since, the sphere intercepts X-axis at $(a, 0,0)$, Y-axis at $(0, \mathrm{b}, 0)$ and Z-axis at $(0,0, \mathrm{c})$, we get,

For point $(a, 0,0)$, $\Rightarrow a^2 + 2ua = 0 \qquad \Rightarrow u = -\frac{a}{2}$

For point $(0, b, 0)$, $\Rightarrow b^2 + 2vb = 0 \qquad \Rightarrow v = -\frac{b}{2}$

For point $(0,0, c)$, $\Rightarrow c^2 + 2wc = 0 \qquad \Rightarrow u = -\frac{c}{2}$

$\therefore$ eq.(2) reduces to,

Equations of Sphere:

C) Intercept form:

$$\boldsymbol{x^2 + y^2 + z^2 + 2ux + 2vy + 2wz = 0} \quad ....(2)$$

$$\Rightarrow x^2 + y^2 + z^2 + 2\left(-\frac{a}{2}\right)x + 2\left(-\frac{b}{2}\right)y + 2\left(-\frac{c}{2}\right)z = 0$$

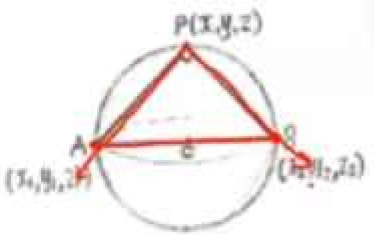
$$\Rightarrow \quad \boldsymbol{x^2 + y^2 + z^2 - ax - by - cz = 0}$$

Hence, The equation of sphere in Intercept Form is:

$$\boxed{x^2 + y^2 + z^2 - ax - by - cz = 0}$$

Equations of Sphere:

D) Diameter form:



Consider the sphere is defined on the line joining points A & B.
Let A (x_1, y_1, z_1) & B (x_2, y_2, z_2) be endpoints of the diameter. Also let P (x, y, z) be any point on the sphere.

$\therefore \angle APB = 90°$

$\Rightarrow$ d.r.s of AP $\equiv (x - x_1, y - y_1, z - z_1)$ &

d.r.s of BP $\equiv (x - x_2, y - y_2, z - z_2)$

As AP $\perp$ BP, By using condition for perpendicularity of two line,

Equations of Sphere:

E) Four Point form:

To find the equation of the sphere passing through four points $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4)$.

Consider the equation of sphere in general form,

$$x^2 + y^2 + z^2 + \underline{2ux} + \underline{2vy} + \underline{2wz} + \underline{d} = 0 \quad \text{....(1)}$$

As it passes through above four points, we get,

$$x_1^2 + y_1^2 + z_1^2 + 2ux_1 + 2vy_1 + 2wz_1 + d = 0 \quad \text{....(2)}$$

$$x_2^2 + y_2^2 + z_2^2 + 2ux_2 + 2vy_2 + 2wz_2 + d = 0 \quad \text{....(3)}$$

$$x_3^2 + y_3^2 + z_3^2 + 2ux_3 + 2vy_3 + 2w_3 + d = 0 \quad \text{....(4)}$$

$$x_4^2 + y_4^2 + z_4^2 + 2ux_4 + 2vy_4 + 2w_4 + d = 0 \quad \text{....(5)}$$

On solving eq.(2), (3), (4), (5) simultaneously to find u, v, w & d,

Example:

Find the equation of the sphere which passes through the points $(2, 1, 1)$ and $(0, 3, 2)$ and has its centre on the line $2x + y + 3z = 0 = x + 2y + 2z$.

Solution:

Let the equation of required sphere be:

$$x^2 + y^2 + z^2 + 2ux + 2vy + 2wz + d = 0 \ldots (1)$$

Since the sphere passes through points (2,1,1) and (0,3,2),

$\Rightarrow$ They satisfy eq.(1),

$\therefore$ for(2,1,1) $\Rightarrow 6 + 4u + 2v + 2w + d = 0$ ....(2)

$\therefore$ for(0,3,2) $\Rightarrow 13 + 0u + 6v + 4w + d = 0$ ....(3)

Example:

Also,

Centre of the sphere $(-u, -v, -w)$ lies on the line $2x + y + 3z = 0 = x + 2y + 2z$.

$\Rightarrow -2u - v - 3w = 0 \quad \& \quad -u - 2v - 2w = 0$

$\Rightarrow 2u + v + 3w = 0 \quad \quad .... (4)$

$\& \; u + 2v + 2w = 0 \quad \quad .... (5)$

On solving eq. (2), (3), (4), (5), we get,

$$u = \frac{28}{18}, \quad v = \frac{7}{18}, \quad w = -\frac{21}{18}$$

$\therefore$ from eq.(2) $\Rightarrow \; d = -\frac{96}{9}$

Example:

As $u = \frac{28}{18}$, $v = \frac{7}{18}$, $w = -\frac{21}{18}$, $d = -\frac{96}{9}$

$\therefore$ The required equation of the sphere is

$$x^2 + y^2 + z^2 + 2\left(\frac{28}{18}\right)x + 2\left(\frac{7}{18}\right)y + 2\left(-\frac{21}{18}\right)z + \left(-\frac{96}{9}\right) = 0$$

$$\Rightarrow \mathbf{9\left(x^2 + y^2 + z^2\right) + 28x + 7y - 21z - 96 = 0}$$

THANKYOU