Scalars And Vectors

① Physics ② Maths ③ Computer Science {Data Science}

Defn: Scalar :

A Scalar is a single numerical Value. It represents a magnitude or quantity and has no direction.

Ex: Car Speed = 45 Km/hr → Magnitude

Temperature in Celsius T = 25°C

Application in Data Science

Dataset: Count of the Total No.of Records = 5

Average of the feature f1 = —

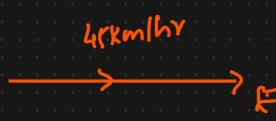
	Age		
	f1	f2	f3
	—	—	—
	—	—	—
	—	—	—
	—	—	—
	—	—	—

Simple Linear Regression ⇒ $y = mx + c$ (c → intercept → Scalar value; m → slope)

② **Vector** : Numerical Value which has both magnitude And Direction.

A vector is an ordered list of numbers. It can represent a point in space or quantity with both magnitude and direction.

Eg: Speed of the car is 45km/hr and is moving Toward East Direction



3 hrs → magnitude, units

Example : Student marks

IQ	No.of study hrs	Pass/Fail
→ [90	3hrs]	Fails

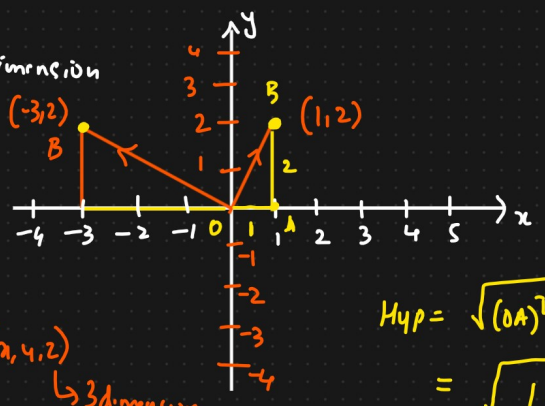
A vector representing person IQ and no.of study = [90, 3hrs]

→ [100 3hrs] Pass

A vector representing person's weight over time $[70, 72, 75, 73]$ ← 4 dimension

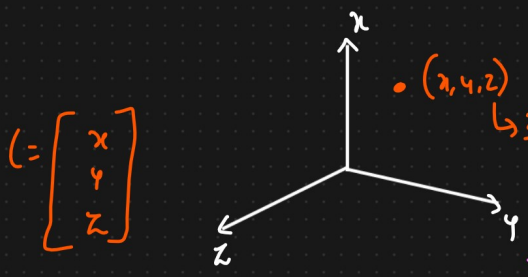
$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ⇒ 2 dimension

$B = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$



$Hyp = \sqrt{(OA)^2 + (AB)^2}$

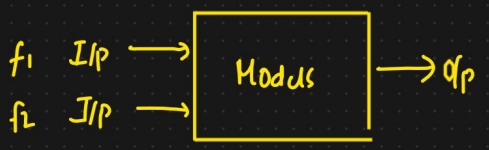
$= \sqrt{1+4} = \sqrt{5} = OB$

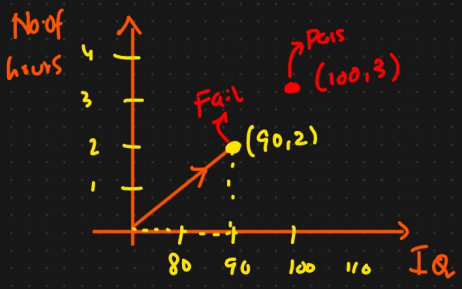


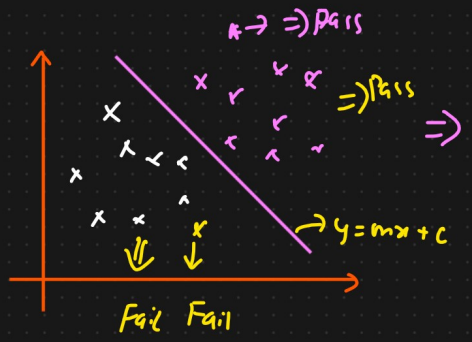
$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

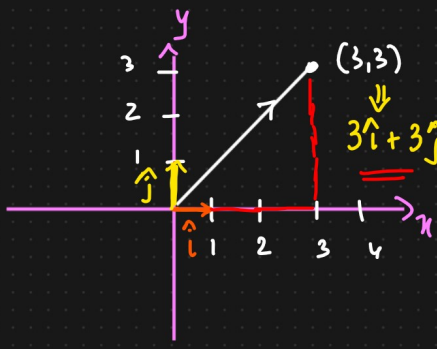
f_1 f_2 f_3 f_4 f_5

f_1	f_2	op
IQ	No. of hours	Pass/Fail
→ 90	2	Fail ⇒ 0
→ 100	3	Pass ⇒ 1







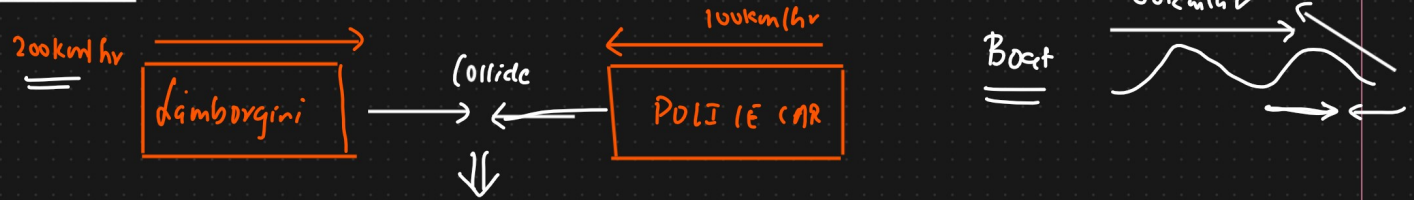


Unit Vector

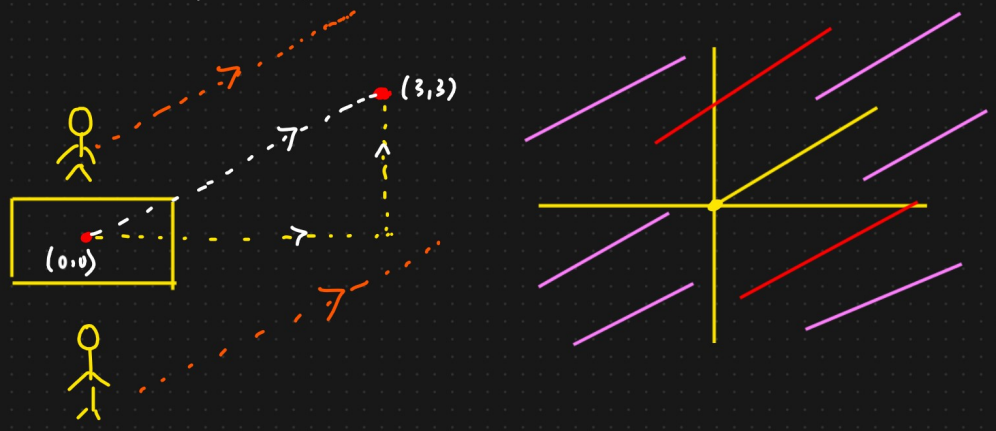
$\hat{\mu} = 1$

$\hat{i}$ $\hat{j}$ ⇒ unit vector towards x And y axis ⇒ 1

Gaming Industry ⇒ GTA 6



Advance Effect => CAR => Blown Up

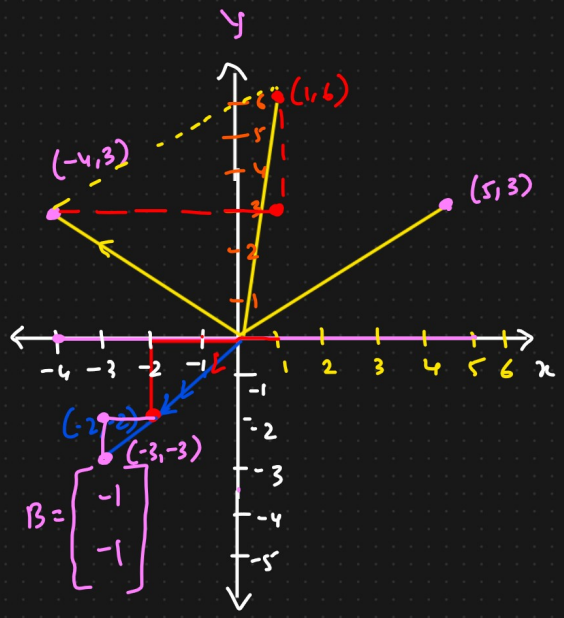


③ Addition of 2 Vectors

$$P_1 = \begin{bmatrix} -4 \\ 3 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$P_1 + P_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix} + \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$



$$A = \begin{bmatrix} -2 \\ -2 \end{bmatrix} \qquad B = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$A = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \qquad B = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \qquad A+B = \begin{bmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{bmatrix} = \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix}$$

Example

Solving a use case

EDA And F.E

Sensor 1 → $[3, 5, 7]$

Sensor 2 $[2, 4, 6]$

Final Sensor Reading

Sensor 1 + Sensor 2 $[5, 9, 13]$.

1) DATA Aggregation Task

2) Feature Engineering

NLP : {Natural Language Processing}

Ecommerce Website

Reviews	Sentiment
The product is good	1
The product is bad	0
⋮	

Ip → Model → Sentiment

Text → Vector → OHE
TFIDF
BOW
Word2Vec

[- - - - -]
↓
Numerical Value

Word Embeddings

1) DATA : [0.2, 0.1, 0.4]

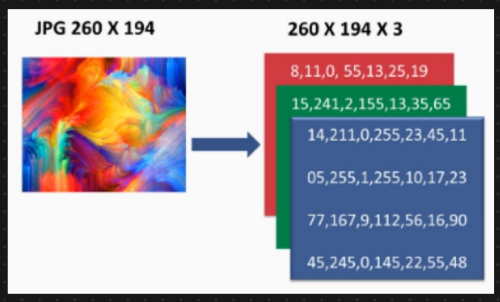
2) Science : [0.3, 0.7, 0.2]

Data Science =

$V_{DATA} + V_{SCIENCE} = [0.2, 0.1, 0.4] + [0.3, 0.7, 0.2]$

$= [0.5, 0.8, 0.6]$ ⇒ DATA SCIENCE

3) Image Processing



Color Image [R, G, B] =

- Red channel R = [255, 128, 0] ✓
- GREEN channel G = [128, 255, 0] ✓
- Blue channel B = [64, 64, 255] ✓

RGB → Greyscale = $\left[\frac{255+128+64}{3}, \frac{128+255+64}{3}, \frac{0+0+255}{3}\right]$ = [149, 149, 85].

Black And White

④ Multiplication of Vectors

3 Types

1) Dot Product (Inner Product) ✓
2) Element Wise multiplication ✓
3) Scalar Multiplication ✓

① Dot Product :

Defn : The dot product of 2 vectors results in a scalar and is calculated as the sum of the products of their corresponding components.

$A = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $B = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$

$A \cdot B = 2*4 + 3*5$

$= 8 + 15 = 23$ // $\Rightarrow$ Scalar value

$A \cdot B^T = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 & 5 \end{bmatrix}$

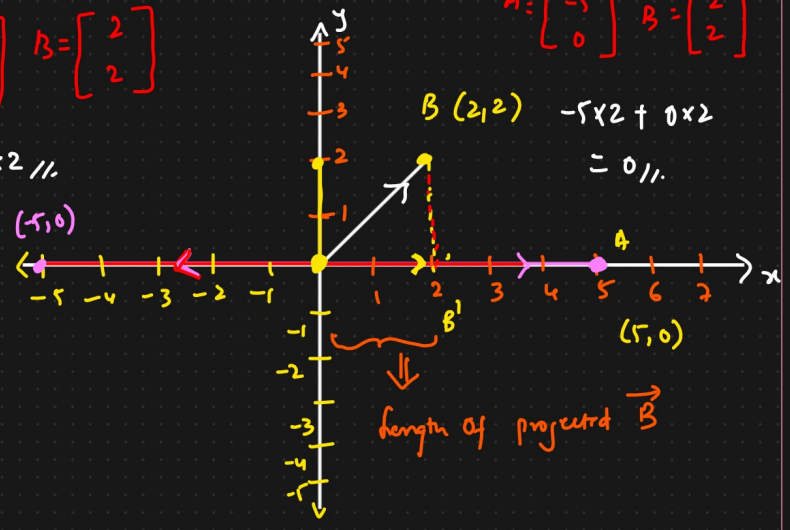
$= 2*4 + 3*5$

$= 8 + 15 = 23$ //

$A = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ $B = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$\vec{A} \cdot \vec{B} = 5*2 + 0*2$ //

$= 10$



$A = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$ $B = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$-5*2 + 0*2$

$= 0$ //.

$\vec{A} \cdot \vec{B}$ = (Length of projected $\vec{B}$) $\cdot$ (Length of vector $\vec{A}$)

$(2) \cdot (5) = 10$ // = +ve

$\vec{A} \cdot \vec{B} = 2 * (-5) = -10$ // = -ve //

$\vec{A} \cdot \vec{B} = 0$ // $\Rightarrow$ Project the vector to the origin. //

Application of Dot product in DATA SCIENCE $\Rightarrow$ Gen AI App $\Rightarrow$ RAG

1) Cosine Similarity

Defn : It is a measure used to determine how similar 2 vectors are. It calculates the cosine of the angle between 2 vectors, providing a similarity

Score that ranges -1 (dissimilar) to 1 (complete similar)

$$\cos\theta = \frac{A \cdot B}{\|A\| \|B\|} \Rightarrow \text{Dot Product } A \cdot B$$

Recommendation System : Netflix Account ⟶ Action Movie

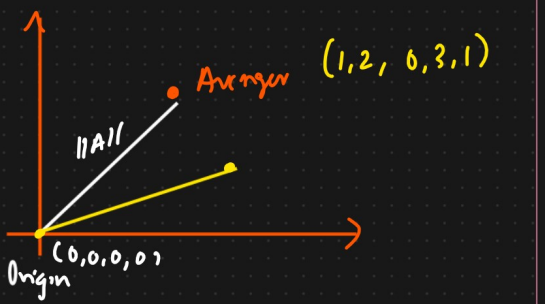
↳ Recommendation of other Action Movies.

⇓

Avengers ⇒ [1, 2, 0, 3, 1] (Drama, Romance)

(Action, Comedy)

⇒ B ⇒ [2, 0, 1, 1, 1]



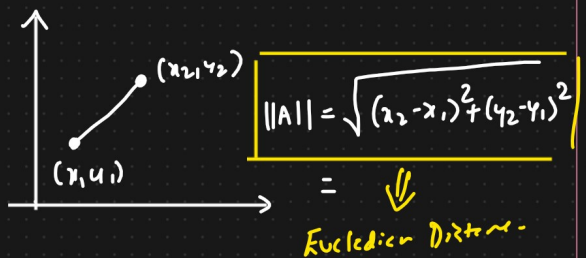
Step 1 : Dot Product of A · B

$$A \cdot B = 1 \cdot 2 + 2 \cdot 0 + 0 \cdot 1 + 3 \cdot 1 + 1 \cdot 1 = 6$$

Step 2 : ||A|| ||B||

$$\|A\| = \sqrt{1^2 + 2^2 + 0^2 + 3^2 + 1^2} = \sqrt{15} = \approx 3.872$$

$$\|B\| = \sqrt{2^2 + 0^2 + 1^2 + 1^2 + 1^2} = \sqrt{7} = 2.646$$



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

⇓

Euclidean Distance

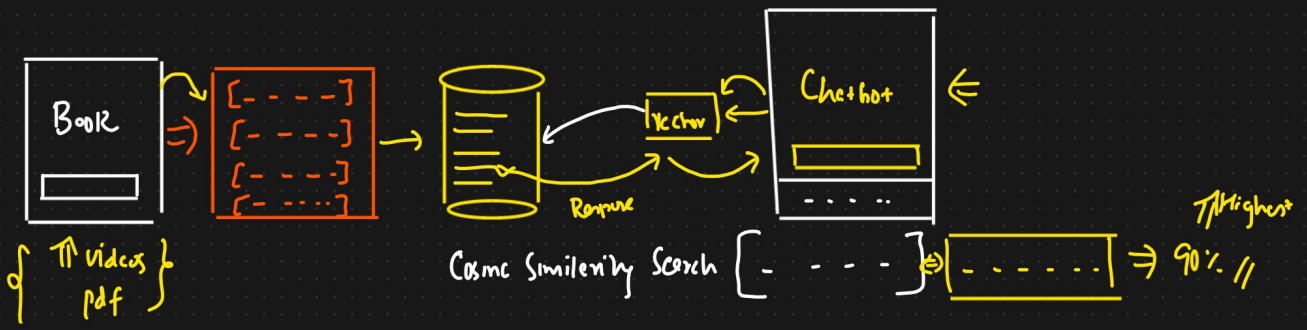
$$\cos\theta = \frac{6}{3.872 * 2.646} = \approx = 0.586 \Rightarrow 0.58$$

⇓

58.6 % +ve Similar.

Vector Databases ⇒ GenAI LLM Models

[RAG System]



② Element Wise Multiplication

In element wise multiplication, corresponding elements of 2 vectors are multiplied to form a new vector of the same dimension.

$$A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

$$A \otimes B = \begin{bmatrix} 3 \\ 8 \\ 15 \end{bmatrix} \Rightarrow \text{Dimension.}$$

Application Data Science

Feature Engineering

Product	Cost		Discount			Discounted Price	Final Price
A	1000	⊗	0.1	−		100	900
B	500	⊗	0.2	−		100	400
C	200	⊗	0.15	−		30	170.

Deep learning : RNN, LSTM RNN, GRU RNN

⊗ ⊕ ⇒ forget gate, input gate

$$\begin{bmatrix} 0.5 \\ 0.6 \\ 0.3 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ 0.3 \end{bmatrix}$$

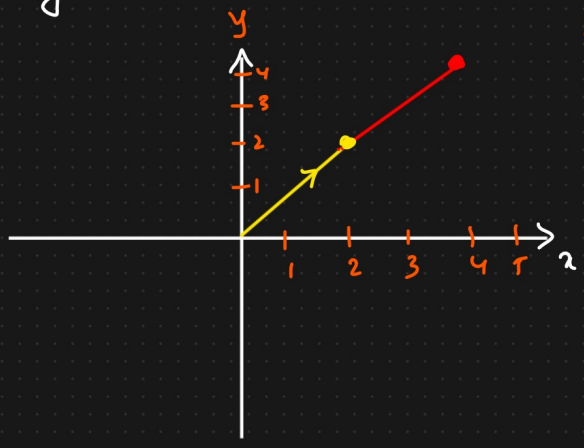
gate ⇒ pass info or not

③ Scalar Multiplication

It involves multiplying vector by a scalar, resulting in a vector where each component is scaled by the vector

$$A = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \qquad C = 4$$

$$CA = \begin{bmatrix} 12 \\ 20 \\ 28 \end{bmatrix}$$

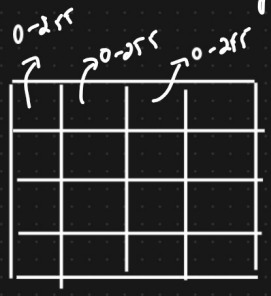


$$A = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \qquad C = 2$$

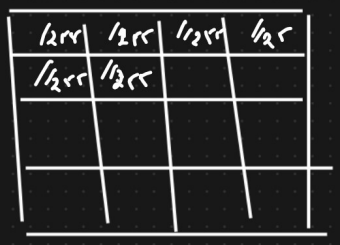
Eg: Normalization And Standardization — R, G, B

⇓

Scaling data ⇒ units.



⇒ Image Processing ⇒ Normalize pixel [0-1] ⇒



Eg: Machine

1cm = 0.01 m

Height = [160, 170, 180]

Scale (to meters) = C = 0.01

Ch = 0.01 [160, 170, 180] = [1.6, 1.7, 1.8] m