What We Will Learn

3 Main Important things

$\boxed{23+\text{hrs}}$ → Maths

1. Linear Algebra
2. Statistics → Basics to Advanced ⇒ Applications of all these Topics In Data Science.
3. Differential Calculus.

① Linear Algebra : Scalers, Vectors, Vectors Operation, Matrices, Matrix Operation functions, Linear Transformations, Inverse function, Eigen values and Eigen vectors

Neural N/W : Forward propogation → Matrices operations

⇓

Applications In Data Science

② Statistics → ML, Deep Learning → Models ⇒ Huge Dataset

↳ Tools to learn from these Data

Descriptive

1. Measure of Central Tendency
2. Measure of Dispersion
3. Histograms, Box plot
4. Types of distribution of DATA.
5. PDF, PMF, Normal Distr, LogNor

Applications ⇐

Inferential

1. Hypothesis Testing & P Value
2. Z test, t-test
3. Chi Square Test
4. ANNOVA Test

} Multiple problems

3 Differential Calculus

1. Derivatives, Slope => Visual Diagrams => Deriving Equations ←
2. Tangent lines
3. polynomial Expressions [Derivative of this Expression]
4. Trigonometric Expression
5. Chain Rule of Derivative
6. Composite Function => Optimizations => $\boxed{\text{Chain Rule.}}$

Applications of Linear Algebra, Stats Differential In Data Science

1. Simple Linear Regression, Multiple Linear Regression ← Application
2. Dimensionality Reduction [Principal Component Analysis] → Eigen Values, Eigen Vector
3. Neural N/W IS TRAINED → ANN → Multi Layered NN.

ANN ↓ Artificial NN

Linear Algebra

Linear algebra is a branch of mathematics that focuses on the study of vectors, vector spaces (also called linear spaces), linear transformations, and systems of linear equations. It provides a framework for understanding the properties and operations of these mathematical objects, which can be represented using matrices and vectors.

(1) Foundational Concepts → ML, DL, NLP, Images

↓

Scalers, Vectors, Matrices, Mathematical Operation of Matrices, Linear Transformation

Eigen Value Eigen Vector

(1) Physics (2) Mathematics (3) Computer Science Student. [Data Science].

Applications of Linear Algebra

1) Data Representation And Manipulation

DATASET → Create Model which will be able to predict. $\vec{v} = [1200]$ $\vec{v} = [1200\ 2]$

House price dataset I/P features O/P feature ⇒ $\boxed{\text{Model}}$ → O/P Price.

Area	No. of Rooms	Location	Price
[1200	2	Bangalore	45 lakhs INR]
—	—	—	—
—	—	—	—
—	—	—	—

Location ⟷ Price

⇒ Quantify the relationship

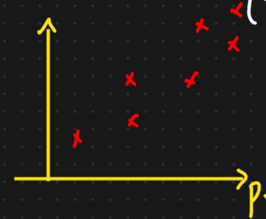
X, Y

Covariance, Correlation

X↑ Y↑
X↓ Y↓
X↑ Y↓
X↓ Y↑
⇒ Stats

(x, y) ⇒ (1200, 45 lakhs) ⇒ 2 dimension vector

X↑ Y↑
X↓ Y↓
X↑ Y↓
X↓ Y↑



(x, y, z)

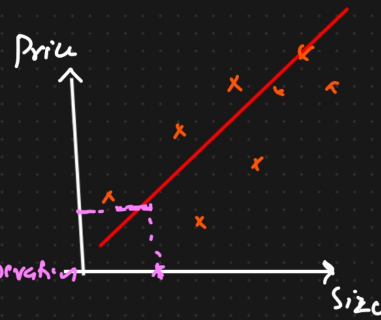
3 dimension

500 dimension ⇓ Dimensionality Reduction ⇑ PCA.

2 dimension

(1) Linear algebra works higher dimension data.

(2) Machine Learning And Artificial Intelligence



(1) Model Train : Linear Algebra → Matrix Multiplication

⇓ Matrix Arithmetic operation

⇓

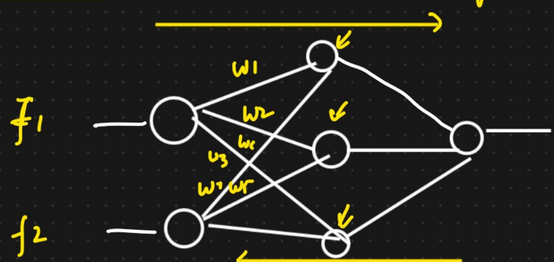
$\boxed{\text{Linear Equation}}$

Equation of a straight line ⇒ $ax+by+c=0$ ⇒ $y=mx+c$

(2) Dimensionality Reduction ⇒ PCA → Linear Algebra → Eigen Value And Eigen Vector

Reduce from higher dimension → Lower Dimension.

(3) Neural N/W : Forward propogation and Backward propogation



$\boxed{2\times3}$

f_1	f_2	O/P
Area	No of Rooms	Price

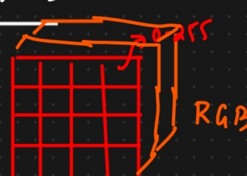
⇒ GPU → Cores → Parallelly

Tensorflow → Tensors

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \end{bmatrix}$$ ⇒ Matrix Multiplication

(3) Computer Graphics

Image



⇒ Scaling, Rotate, Black & White ⇒ Linear ⇒ Transforming

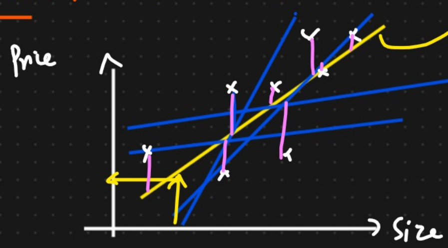
(4) Optimization

(1) Solving Equations : Linear Equation

slope or coefficient, intercept

$y=mx+c$ ⇒ Regression

↳ $f(x)$ ⇒ Maximize f_n ⇒ Minimize the Error



⇒ Right slope and intercept

⇓
GRADIENT DESCENT // $\Rightarrow$ Optimizer.

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

$$\begin{array}{cccc} & \text{Age} & & \\ & f1 & f2 & f3 \\ & - & - & - \\ & - & - & - \\ & - & - & - \\ & - & - & - \\ & - & - & - \end{array}$$

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

Average of the feature f1 = —

Simple Linear Regression ⇒ $y = mx + c$ → intercept ↳ Scalar value

↳ 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

$\xrightarrow{45km/hr}$ E

3 hrs → units

↗ magnitude

Example : Student marks

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

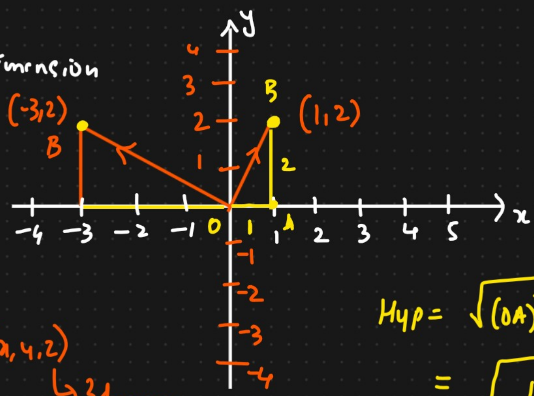
A vector representing person IQ and no.of Study = [90 , 3hrs]

→ [100 3hrs] Pass

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

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

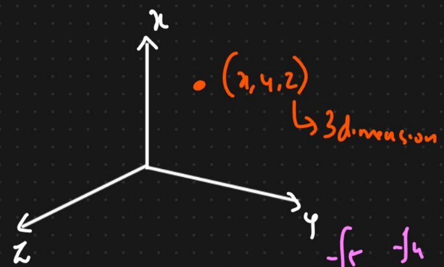
$B = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$



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

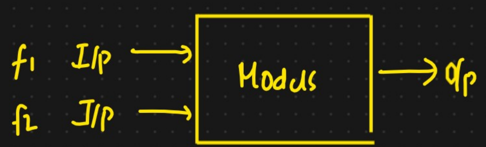
$= \sqrt{1+4} = \sqrt{5} = OB$

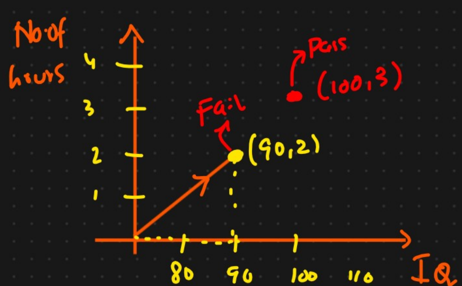
$C = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

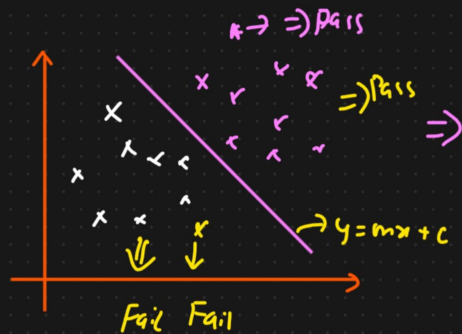


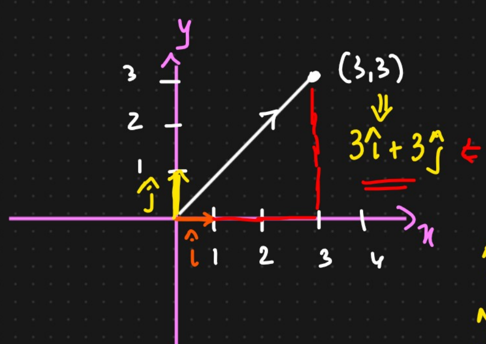
f5 f4 f3

f1	f2	o/p ←
IQ	No of hours	Pass/fail
→ [90	2]	Fail ⇒ 0
→ [100	3]	Pass ⇒ 1







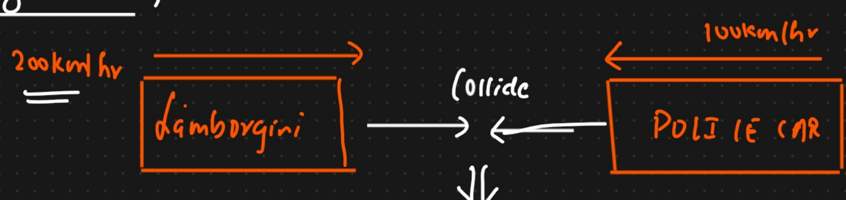


Unit Vector

$\hat{\mu} = 1$

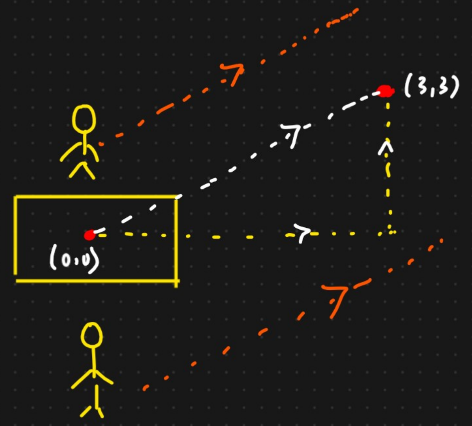
$\hat{i}$, $\hat{j}$ ⇒ unit vector towards x And y axis ⇒ 1

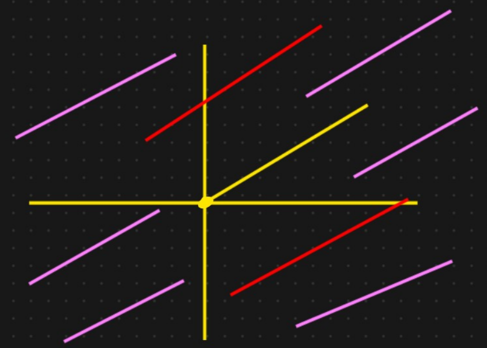
Gaming Industry ⇒ GTA 6



Boat

Adverse Effect => CAR => Blown Up





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

Age

$f1$	$f2$	$f3$
–	–	–
–	–	–
–	–	–
–	–	–
–	–	–

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

Average of the feature $f1$ = —

Simple Linear Regression ⇒ $y = mx + c$ → intercept

↳ Slope ↳ Scalar value

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

$\xrightarrow{45km/hr}$ E

3 hrs → units

↑ magnitude

Example : Student marks

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

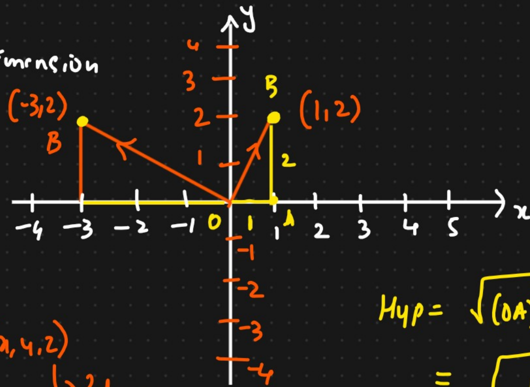
A vector representing person IQ and no.of Study = [90 , 3hrs]

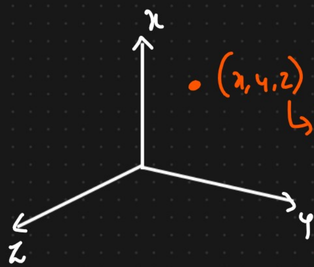
→ [100 3hrs] Pass

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

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

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





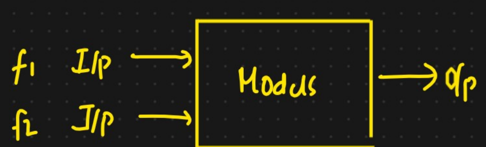
$C = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

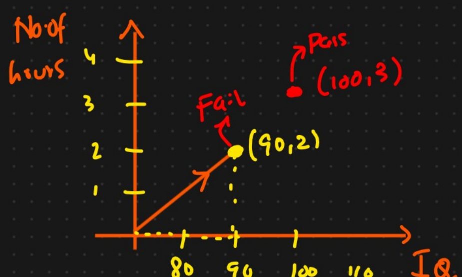
$Hyp = \sqrt{(OA)^2 + (AB)^2}$

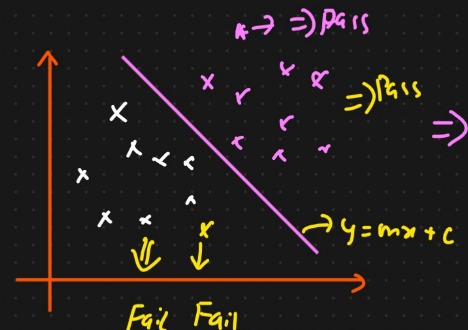
$= \sqrt{1+4} = \sqrt{5} = OB$

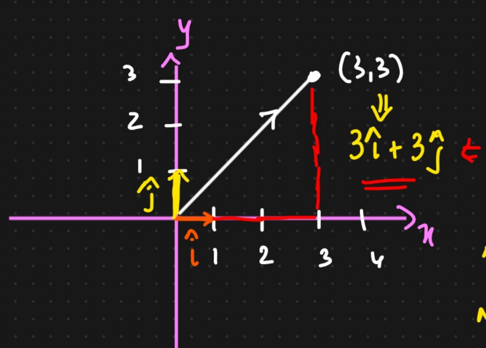
f1 f2 f3

f1	f2	o/p ←
IQ	No. of hours	Pass/fail
→ [90	2]	Fail ⇒ 0
→ [100	3]	Pass ⇒ 1







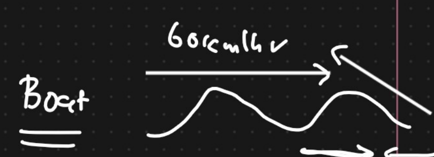


Unit Vector

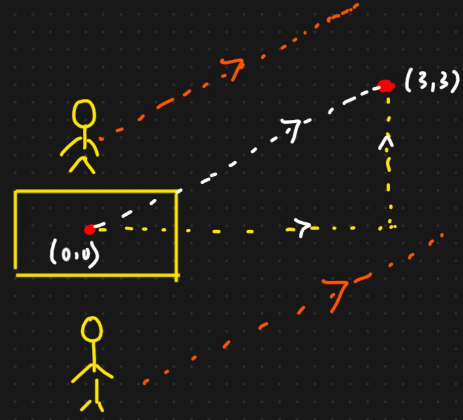
$\hat{\mu} = 1$

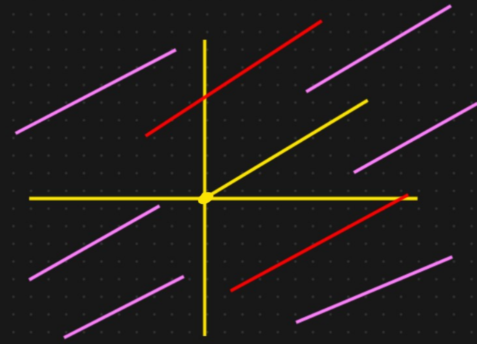
$\hat{i}$, $\hat{j}$ ⇒ unit vector towards x And y axis ⇒ 1

Gaming Industry ⇒ GTA 6



Adverse Effect => CAR => Blown Up



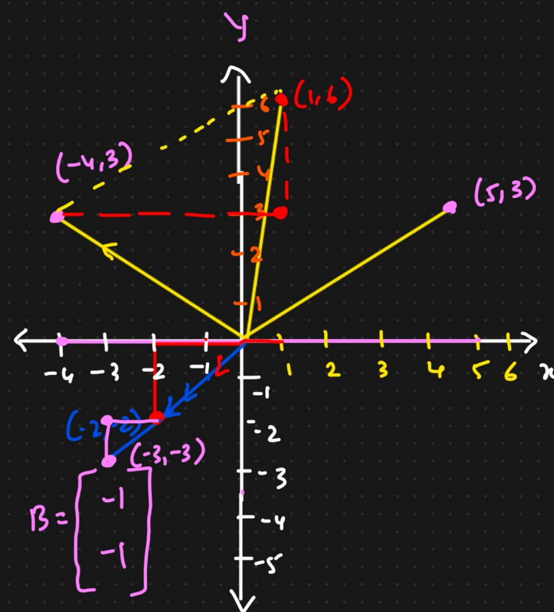


③ Addition of 2 Vectors

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

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

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



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

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

Example

Solving a usecase

Sensor 1   Sensor 2

→ [3, 5, 7]   [2, 4, 6]

EDA And F.E

Final Sensor Reading

Sensor1 + Sensor2 [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
⋮	

I/p ⟶ Model ⟶ Sentiment

Text → Vector → OHE
TFIDF
BoW
Word2Vec

[- - - - -]
↓
Numerical Value

Word Embeddings

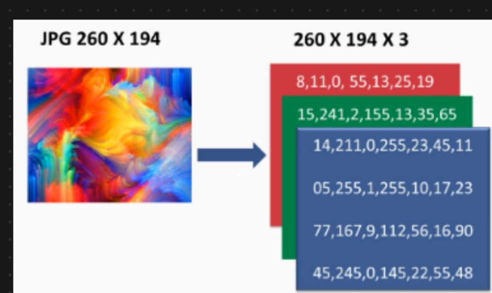
1) DATA : [0.2, 0.1, 0.4]
2) Science : [0.3, 0.7, 0.2]

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] ✓

White And Black

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

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.

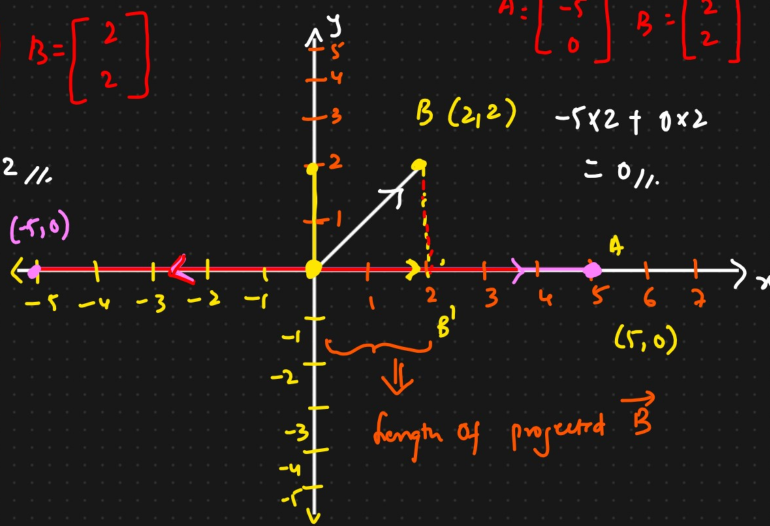
$A = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad B = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$

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

$\vec{A} \cdot \vec{B} = 5 \times 2 + 0 \times 2$ //
$= 10$

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

$-5 \times 2 + 0 \times 2$
$= 0$ //.



$A \cdot B = 2 * 4 + 3 * 5$
$= 8 + 15 = 23$ // $\Rightarrow$ Scalar value

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

$= 2 \times 4 + 3 \times 5$
$= 8 + 15 = 23$ //.

$\vec{A} \cdot \vec{B}$ = (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)

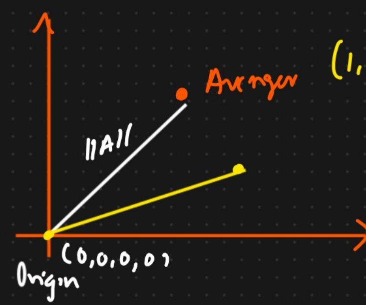
$$\cos\theta = \frac{A \cdot B}{\|A\|\,\|B\|} \Rightarrow \text{Dot Product } \underline{A \cdot B}$$

Recommendation System : Netflix Account ⟶ Action Movie

↳ Recommendation of other Action Movies.

$$\text{Avengers} \Rightarrow \begin{array}{ccccc} & & & \text{Drama} & \\ [1, & 2, & 0, & 3, & 1] \\ \text{Action} & \text{Comedy} & & & \text{Romance} \end{array}$$

$$\Rightarrow B \Rightarrow [2, 0, 1, 1, 1]$$



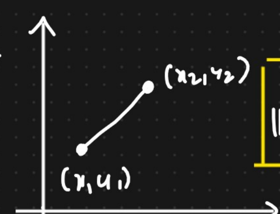
Step 1 : Dot Product of A·B

$$A \cdot B = 1\cdot2 + 2\cdot0 + 0\cdot1 + 3\cdot1 + 1\cdot1 = 6$$

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

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

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



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

= ⇓

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]

{ Videos pdf }

Cosine Similarity Search [- - - -] ⇒ [- - - - - -] ⇒ 90% // ↑Highest

② Element Wise Multiplication

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

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

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

Application Data Science

Feature Engineering

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

Deep Learning : $\text{RNN}, \underbrace{\text{LSTM RNN, GRU RNN}}$

$\Downarrow$

$\otimes \oplus \Rightarrow$ forget gate, input gate

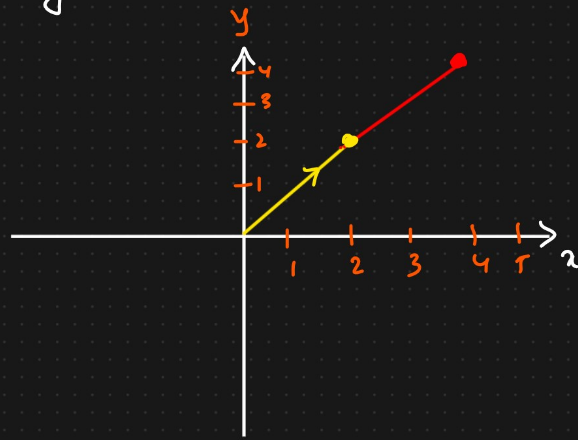
$$\begin{bmatrix} 0.5 \\ 0.6 \\ 0.3 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ 0.3 \end{bmatrix}$$

gate $\Rightarrow$ 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

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



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

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

Eg: Normalization And Standardization ⇓ Scaling data ⇒ units.

R, G, B

0-255 0-255 0-255

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

1/255	1/255	1/255	1/255
1/255	1/255		

Eg: Machine

Height = [160, 170, 180]

Scale (to meters) = $C = 0.01$

1cm = 0.01 m

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

Matrices

A matrix is a rectangular array of numbers, symbols or expressions arranged in rows and columns

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

$\Rightarrow$ has a_{ij} where i denotes the row and j denotes the columns

Example of Matrices in Data Science

① Data Representation

Dataset

	Math Score	Physic Score	Biology Score
$\rightarrow$	[55	65	75]
$\rightarrow$	[65	60	55]
$\rightarrow$	[70	45	80]

$$\begin{array}{cc} & \begin{array}{ccc} f_1 & f_2 & f_3 \\ \downarrow & \downarrow & \downarrow \end{array} \\ \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} & \begin{bmatrix} 55 & 65 & 75 \\ 65 & 60 & 55 \\ 70 & 45 & 80 \end{bmatrix} \end{array}$$

$\boxed{3 \times 3}$

② Images in Computer Vision

0 (black) $\longleftrightarrow$ 255 (white)

3×3 grayscale Image

Image =

0	128	255
255	128	0
128	255	128

$$\Rightarrow \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \begin{bmatrix} 0 & 128 & 255 \\ 255 & 128 & 0 \\ 128 & 255 & 128 \end{bmatrix}_{3 \times 3}$$

③ Confusion Matrix : Accuracy of the Model

$f_1 \longrightarrow$ / $f_2 \longrightarrow$ $\boxed{\text{Model}} \longrightarrow$ o/p $\rightarrow \hat{y} \Rightarrow$ predicted output.

→ y ⇒ actual o/p

Confusion Matrix = $\begin{bmatrix} 50 & 10 \\ 5 & 35 \end{bmatrix}_{2\times2}$ ⇒

50 → True positive

10 → false negatives

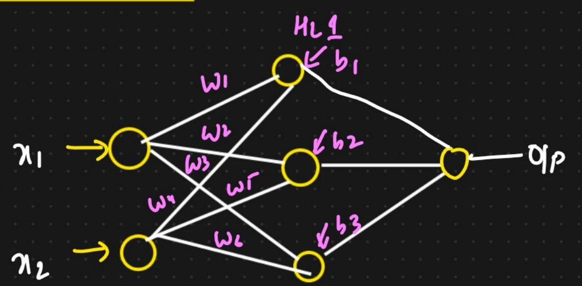
5 → false positive.

35 → True negative

$$\frac{TP + TN}{TP + FN + FP + TN} = \text{Accuracy}.$$

4) Neural N/W : Matrix Operation

[Linear Regression]



2x3

Forward propogation

$z = w^T x + b$

x_1	x_2
—	—
—	—
—	—
—	—

$W = \begin{bmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \end{bmatrix}$

x

$x = [x_1 \;\; x_2]$

+

$[b_1, b_2, b_3]$

Matrix Multiplication ⇐

x_1	x_2		o/p feature
⇓	⇓	⇑	⇓
No of study hours	IQ	Score	dependent feature
4	100	90	
5	90	85	
—	—	-	

Regression

$y = mx + c$

$y = m_1 x_1 + m_2 x_2 + c$

↳ slope ↳ slope or Coefficient

$\Rightarrow m^T x + c$

$\left\{ \begin{matrix} m = [m_1, m_2] \\ x = [x_1, x_2] \end{matrix} \right\}$

5) NLP :

Dataset

Review	Positive / Negative
→ The food is bad	0
→ The food is Good	1

$\begin{bmatrix} [0.1 & 0.2 & 0.3] \\ [0.4 & 0.5 & 0.6] \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Matrices Operations

⇒ To manipulate and analyze multidimension data efficiently.

1) Matrix Addition And Substraction
2) Scalar Matrix Multiplication.
3) Matrix Multiplication

1) Matrix Addition And Substraction

Add or Substract corresponding elements of 2 matrices of the same dimension

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}_{3\times3} \Rightarrow \text{STORE A} \qquad B = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{bmatrix}_{3\times3} \Rightarrow \text{STORE B}$$

DATASET

	Prod A	Prod B	Prod C
Day 1	1	2	3
Day 2	4	5	6
Day 3	7	8	9

$$A + B = \begin{bmatrix} 1+4 & 2+5 & 3+6 \\ 4+7 & 5+8 & 6+9 \\ 7+1 & 8+2 & 9+3 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 7 & 9 \\ 11 & 13 & 15 \\ 8 & 10 & 12 \end{bmatrix}$$

② Scalar Multiplication

Scalar Multiplication involves multiplying every element of a matrix by a scalar value.

$$B = cA$$

Eg: Suppose we have a matrix representing product prices in dollars and we want to adjust these prices for inflation by a factor of 1.05.

Original :

$$P = \begin{bmatrix} 10 & 20 & 30 \\ 15 & 25 & 35 \\ 20 & 30 & 40 \end{bmatrix}$$

Scaler Multiplication

$$P_{adjusted} = 1.05 \cdot P = 1.05 \begin{bmatrix} 10 & 20 & 30 \\ 15 & 25 & 35 \\ 20 & 30 & 40 \end{bmatrix} = \begin{bmatrix} 10.5 & 21 & 31.5 \\ 15.75 & 26.25 & 36.75 \\ 21 & 31.5 & 42 \end{bmatrix}$$

Eg : DATASET : IT Firms

2024 ⇒ 2025 ⇒ Inflation

6%

$$\begin{array}{ccc} \text{Base salary S/W Eng} & \text{Base Salary HR} & \text{Base Salary Accountant} \end{array}$$

$$\begin{bmatrix} 45K & 30K & 40K \\ 50K & 35K & 45K \\ - & - & - \\ - & - & - \end{bmatrix} \times \boxed{1.06\%}$$

③ Matrix Multiplication

Operation : It involves the dot product of rows of the first matrix with columns of the second matrix.

For 2 matrices $A(m \times n)$ and $B(n \times p)$, the result is a matrix $C(m \times p)$

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \Rightarrow 1 \times 2 + 2 \times 3 + 3 \times 4 = 2 + 6 + 12$$
$$= 20 / .$$

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ 2×3 $\underline{\underline{m\times n}}$

$B = \begin{bmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{bmatrix}$ 2×3 $\underline{\underline{n\times p}}$

$\underline{\underline{\text{TRANSPOSE}}}$

$B^T = \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix}$ 3×2 $\Rightarrow$ $\underline{\underline{2\times2}}$

$$C = A \cdot B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}_{2\times2} = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix}$$

$C_{11} = (1\times7) + (2\times9) + (3\times11) = 7 + 18 + 33 = 58$

$C_{12} = (1\times8) + (2\times10) + (3\times12) = 8 + 20 + 36 = 64$

$C_{21} = (4\times7) + (5\times9) + (6\times11) = 139$

$C_{22} = (4\times8) + (5\times10) + (6\times12) = 154$

Functions And Linear Transformations

Functions

A function is a mathematical relationship that uniquely associates elements of one set (called the domain) with elements of another set (called the codomain). In simpler terms, a function maps inputs to outputs in a specific way.

Notation: A function f mapping elements from set X (domain) to set Y (codomain) is denoted by $f: X \to Y$

If x is an element of X, then $f(x)$ is the corresponding element in Y.

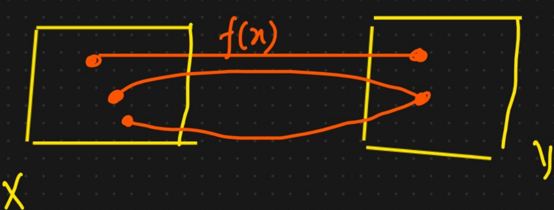
Example: $f(x) = 2x + 3 \Rightarrow$ maps each real number x to a real Number

$2x+3$

$$x \xrightarrow{f:} y.$$

$x = 5$

$f(2) = 2 \times 2 + 3 = 7 \Rightarrow f(x) \to$ Mapping $2 \in R$ to $7 \in R$



$g(x)$ $\quad f: X \to Y$

$$f(x) = \begin{bmatrix} x+y \\ 6z \end{bmatrix}$$

$$f: \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R^3 \xrightarrow{f(x)} \begin{bmatrix} x+y \\ 6z \end{bmatrix} \in R^2$$

$\Rightarrow$ 3 dimension $\Rightarrow$ 2 dimension vector

$f: R^3 \to R^2$

R^3 → Domain

R^2 → Codomain.

$f(x) \Rightarrow$ Transformation.

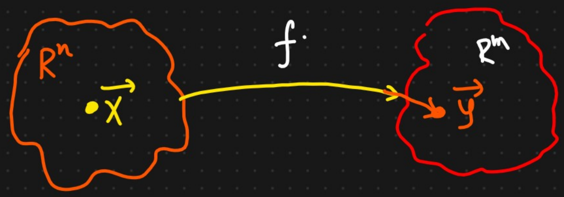
Eg: Dimensionality Reduction

Vector Transformations

$f: X \to Y$

$\vec{x} \to \vec{y}$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$ $\quad x_1, x_2, x_3 \cdots x_n \in R$

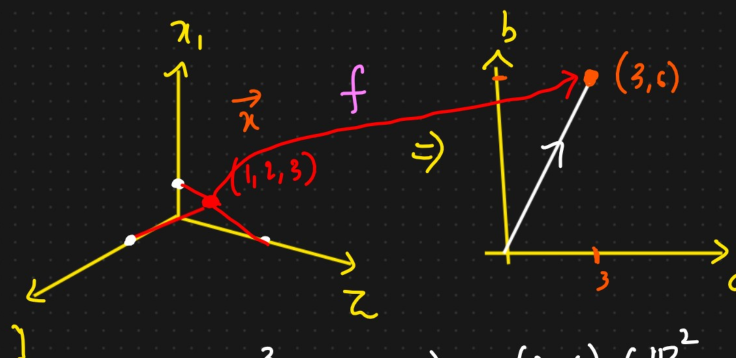


$f: \mathbb{R}^n \to \mathbb{R}^m$

$f(x, y, z) = (x+y, 2z)$

$f\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

Vector Transformation



$(1,2,3) \in \mathbb{R}^3 \quad \Rightarrow \quad (3,6) \in \mathbb{R}^2$

$\boxed{T: \mathbb{R}^3 \to \mathbb{R}^2}$

Eg:

Defn

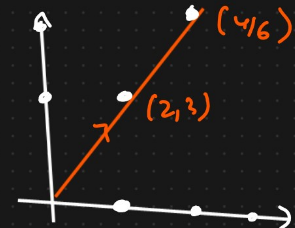
Vector transformations refer to operations that map vectors from one space to another, often changing their magnitude, direction, or both. These transformations are typically described using matrices and are fundamental in various fields, including physics, engineering, computer graphics, and data science.

Eg:

1) Scaling

Scaling is a transformation that changes the magnitude of vector while keeping their direction same.

$v' = 2v = 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$



Application

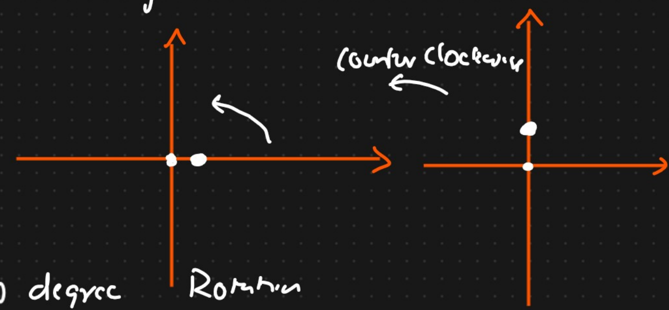
① DATA Normalization

② Computer graphic to resize Objects ⇒ Paint ⇒ Image ⇒ Resize

② Rotation

Transformation that turns vectors around the origin.

$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in R^2$



$v' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ⇒ Showing a 90 degree Rotation

Eg:- Rotation will be used in Image processing ⇒ Rotating Image.

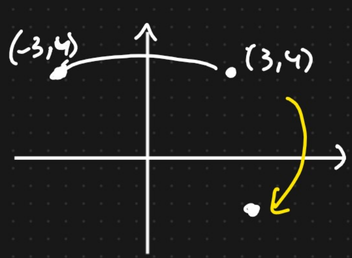
Robotics ⇒ Adjusting Robot orientation

3D graphics ⇒ Rotating objects.

③ Reflection

Transformation that flips vectors over a specified axis or plane.

$v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ ⇒ Across the y axis. ↗ f(x).



$\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

④ Mirroring Images

Analyzing wave reflections.

④ Shearing

⊛ Linear Transformation

A **linear transformation** is a function between two vector spaces that **preserves the operations of vector addition and scalar multiplication**. This means that if **T is a linear transformation from a vector space V to a vector space W**, then for any vectors

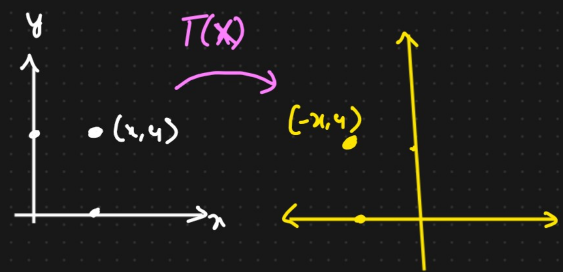
2 important properties

$T: V \to W \} \Rightarrow$ Linear Transformation

① Additivity $T(u+v) = T(u) + T(v)$

② Homogeneity $T(cu) = cT(u)$

$u, v \in V$ and c is a scalar value



Eg ÷ Reflection

The reflection transformation T across the y axis maps a vector

$X = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ $\quad T(x) = \begin{bmatrix} -x \\ y \end{bmatrix}$

Transformation can be expressed as $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

$\rightarrow \boxed{T(x) = AX} \Rightarrow$ Linear Transformation

$$= \underset{1\times 2}{\begin{bmatrix} x & y \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}} = \underset{1\times 2}{\begin{bmatrix} -x \\ y \end{bmatrix}}$$

① Checking Additivity

Let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be two vector in R^2

$$T(u+v) = T(u) + T(v)\}$$

$$u+v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}$$

$$T(u+v) = A(u+v) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} -(u_1+v_1) \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} -u_1-v_2 \\ u_2+v_2 \end{bmatrix}$$

$$T(u) = Au = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix}$$

$$\boxed{LHS = RHS}$$

$$T(v) = Av = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix}$$

$$RHS \Rightarrow T(u) + T(v) = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -u_1-v_1 \\ u_2+v_2 \end{bmatrix}$$

2) Checking Homogenity

Let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in R^2$ and c be a scaler

Homogenity Requirement

$$T(cu) = c\,T(u)$$

$$cu = c \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}$$

$$T(cu) = A(cu) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix} = \begin{bmatrix} -cu_1 \\ cu_2 \end{bmatrix} \Rightarrow LHS$$

Equal

RHS

$$c\,T(u) = c(A \cdot u) = c \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \Rightarrow c \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -cu_1 \\ cu_2 \end{bmatrix}$$

Example that dont follow Linear Transformation

$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad T(x) = x + b$ $\quad T: R^2 \to R^2$

$u \quad v$

vector ⇒ fixed vector

$$T(x) = x + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

① Check Additivity

$$T(u+v) = T(u) + T(v)$$

$$u = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad v = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$T(u+v) = T\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 6 \\ 2 \end{bmatrix}\right)$$

$$T\left(\underbrace{\begin{bmatrix} 6 \\ 2 \end{bmatrix}}_{x}\right) = \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix} \Rightarrow LHS$$

$$T(u) + T(v)$$

$$T\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$T(u) + T(v) = \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \Rightarrow RHS$$

$$LHS \neq RHS$$

Check Homogeneity:

$$T(cu) = c\,T(u)$$

$$u = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad c = 2$$

$$T(cu) = T\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \Rightarrow \text{LHS}$$

$$\neq$$

$$cT(u) = 2\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 2\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \Rightarrow \text{RHS}$$

$$T(x) = x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \text{Not a Linear Transformation}$$

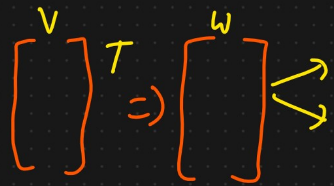
Fails both additivity and homogenity properties.

Why Linear Transformation?

Why Linear Transformation

Linear transformations are fundamental in data science for several reasons. They provide a mathematical framework for manipulating and analyzing data, which is crucial for various data processing tasks, model building, and interpretation. Here are some key reasons why linear transformations are important in data science:

1. Dimensionality Reduction
Principal Component Analysis (PCA): $\Rightarrow$ 5000 $\Rightarrow$ Lower Dimension $\Rightarrow$ Variance
PCA is a widely used technique for reducing the dimensionality of datasets while retaining as much variance as possible. It involves finding a set of orthogonal axes (principal components) and projecting the data onto these axes. The transformation of data points in the original space to the new space defined by the principal components is a linear transformation. This helps in:



Reducing computational cost.
Mitigating the curse of dimensionality.
Visualizing high-dimensional data.

2. **Feature Engineering**
Linear transformations can be used to create new features from existing ones. For example, interactions between features can be captured through linear combinations, which can then be used in machine learning models to improve predictive performance. Techniques like linear regression, ridge regression, and linear discriminant analysis (LDA) all rely on linear transformations to find meaningful feature representations.

3. **Data Preprocessing**
Normalization and Standardization: }
Linear transformations are used to scale data, making it suitable for machine learning models. Standardization transforms data to have a mean of zero and a standard deviation of one, while normalization scales data to a specific range (e.g., [0, 1]). These transformations are essential for ensuring that all features contribute equally to the model, especially in algorithms like gradient descent.

4. **Neural Networks** { Forward , Activation }
In neural networks, especially deep learning models, the layers consist of linear transformations followed by non-linear activation functions. The weights in a neural network can be seen as a series of linear transformations that map input data to intermediate layers and, eventually, to the output layer. This linear aspect is crucial for the network's ability to learn complex patterns in data.

5. **Image and Signal Processing**
In image and signal processing, linear transformations are used extensively. For example:

Convolutional filters in image processing can be seen as linear transformations applied to local regions of an image.
Fourier transforms, which decompose signals into sinusoidal components, are linear transformations that convert time-domain signals into frequency-domain representations.
6. Understanding and Interpretation
Linear transformations simplify complex relationships between variables into linear relationships, which are easier to understand and interpret. For example, linear regression provides a clear model of how each feature affects the target variable through linear coefficients, making it easier to explain to stakeholders.

7. **Optimization and Solving Systems of Equations**
Linear transformations are used to solve systems of linear equations, which is a common problem in data analysis and optimization. Techniques like matrix inversion and the use of pseudo-inverses are essential for finding solutions in linear regression and other linear models.

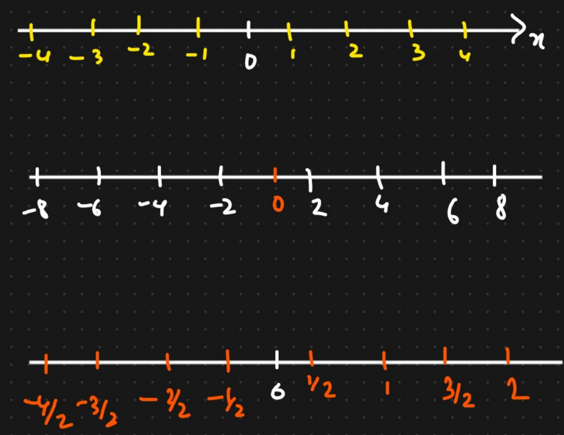
8. Theoretical Foundations
Many advanced machine learning algorithms and statistical techniques have linear algebra and linear transformations at their core. Understanding these fundamentals is crucial for grasping more complex topics like support vector machines

Linear Transformations Visualization

$T: R \rightarrow R$ 1 dimension

2d Matrix

$T(x) = 2x$ [Linear Transformation]



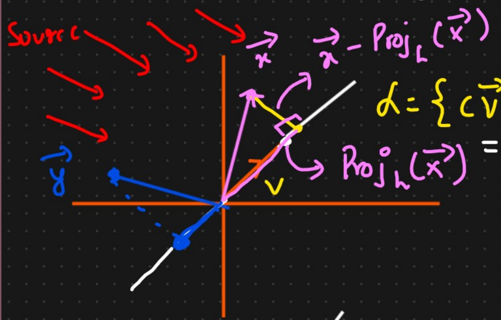
$T(x) = \frac{1}{2}x$ $f(x) = 2x$

$T(x) = -3x$

Property

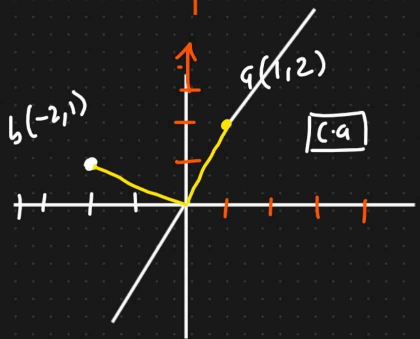
1. Origin must be fixed
2. All lines must remain lines

Introduction To Projections



$Proj_L(\vec{x}) \Rightarrow$ Project the $\vec{x}$ on the line L.

$Proj_L(\vec{x}) \Rightarrow$ Some Vector in Line where $\vec{x} - Proj_L(\vec{x})$ is perpendicular to L



$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$b = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

$$a \cdot b = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$a \cdot b = a_1 b_1 + a_2 b_2.$

$$a \cdot b = (1)(-2) + 2(1)$$

$$\boxed{a \cdot b = 0}$$ $\Rightarrow$ Dot product is always 0.

$Proj_L(\vec{x}) \Rightarrow$ Some Vector in Line where $\vec{x} - Proj_L(\vec{x})$ is perpendicular to L

$$(\vec{x} - c\vec{v}) \cdot \vec{v} = 0$$

$$\vec{x} \cdot \vec{v} - c\,\vec{v} \cdot \vec{v} = 0$$

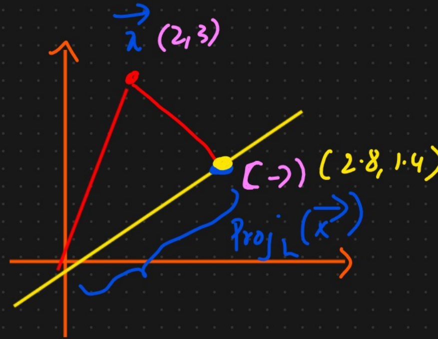
$$\vec{x} \cdot \vec{v} = c\,\vec{v} \cdot \vec{v}$$

$$\boxed{c = \frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}}$$

$$Proj_L(\vec{x}) = c\vec{v}$$
$$= \left(\frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\right) \cdot \vec{v}$$

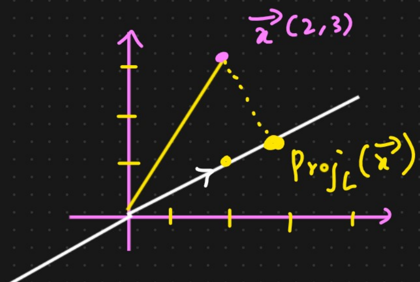
$$\text{Proj}_b(a) = \left(\frac{a \cdot b}{b \cdot b}\right) \cdot b$$

$$\boxed{\text{Proj}_L(\vec{x}) = \left(\frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\right) \cdot \vec{v}}$$



$$L = \{ c \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{} \; c \in \mathbb{R} \}. \qquad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$c\vec{v}$



$$\text{Proj}_L(\vec{x}) = \left(\frac{\begin{bmatrix} 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}}\right) \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \frac{7}{5} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\boxed{\text{Proj}_L(\vec{x}) = \begin{bmatrix} 14/5 \\ 7/5 \end{bmatrix} = \begin{bmatrix} 2.8 \\ 1.4 \end{bmatrix}}$$

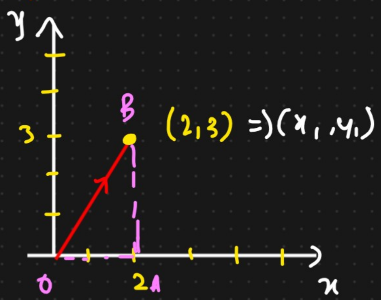
Magnitude and Unit Vectors

Vector length

Unit vector → vector has a length of 1

$\vec{X} \in R^n$ $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$

$\vec{A} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \in R^2$ ($2 \to x$, $3 \to y$)



$\vec{B} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$OB = \|\vec{A}\| = \sqrt{(OA)^2 + (OB)^2}$

$= \sqrt{x_1^2 + y_1^2}$

$\|\vec{B}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$

$\|\vec{A}\| = \sqrt{4+9} = \sqrt{13}.$

$$\boxed{\|\vec{X}\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}}$$ ⇒ Vector length

Unit vector

$\|\vec{u}\| = 1 \Rightarrow \hat{u}$

$\vec{V} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \longrightarrow \vec{u} \Rightarrow \|\vec{u}\| = 1$

$$\boxed{\vec{u} = \frac{1}{\|\vec{V}\|} \cdot \vec{V}}$$ → Scalar Multiplicat.

$\vec{V} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \in R^3$

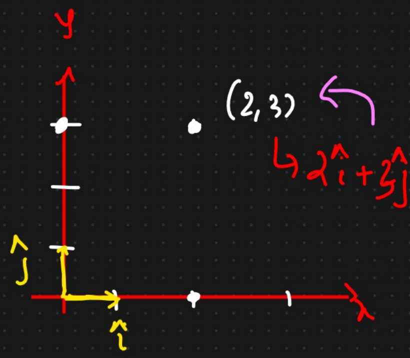
$\|\vec{V}\| = \sqrt{1^2 + 2^2 + 0}$

$= \sqrt{5}$

$\vec{u} = \frac{1}{\sqrt{5}} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$

$$\|\vec{u}\| = \sqrt{\left(\frac{1}{\sqrt{5}}\right)^2 + \left(\frac{2}{\sqrt{5}}\right)^2 + 0^2}$$

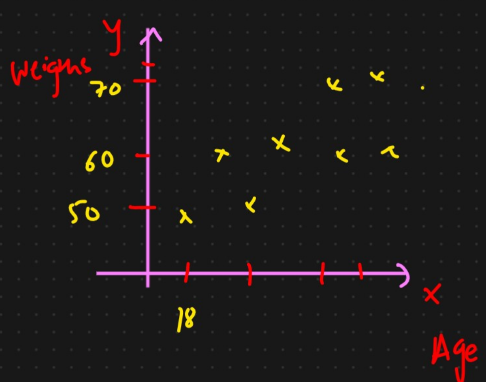
$$= \sqrt{1/5 + 4/5 + 0} = \sqrt{5/5} = 1 = \hat{u}$$ ⇒ unit vector



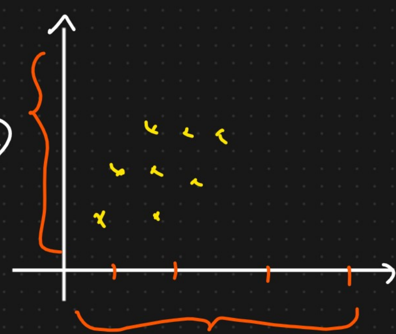
(2,3)

$\hat{i}, \hat{j}$ ⇒ unit vector

Normalization : Vector size ⇒ Vector length = 1



⇒



⇒ Improving the optimization process.

Inverse of a function

A inverse of a function is a function that "reverses" the effect of the original function.

If you have a function f that maps an element x from set X to an element y in a set Y, the inverse function f^{-1} map y back to x.

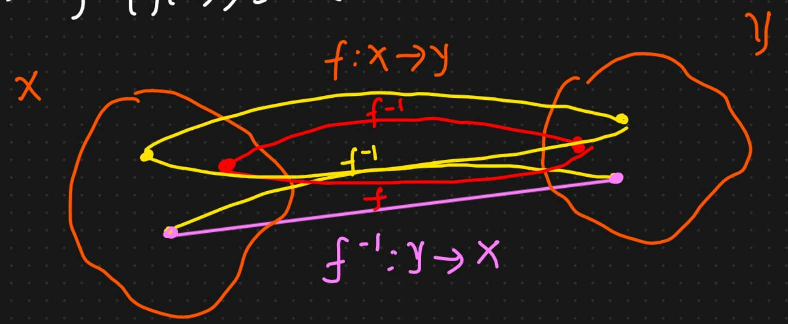
Defn:

Given a function $f: X \to Y$, then inverse function $f^{-1}: Y \to X$ for every $y \in Y$, there is a unique $x \in X$ such that

$$f(x) = y$$

The inverse function f^{-1} satisfies the following condition

1) For all $x \in X$: $f(f^{-1}(y)) = y$
2) For all $y \in Y$: $f^{-1}(f(x)) = x$



These conditions imply that applying the function and then it inverse will return the original value.

Identity function

$$I_X : X \to X \quad \Rightarrow \quad I_X(a) = a$$

$a \in X$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

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

For a set X, the identity function I_X is defined as:

$$I_X(a) = a \quad \text{for all } a \in X$$

I_X is the identity function on the set X and it maps every element x in X to itself

Properties of Identity function

1) **Preservation**: Does not alter any element. If x is the domain, then the image of x under the identity fn is x.

2) **Linearity**: Identify fn is a linear transformation

- $I(u+v) = I(u) + I(v)$
- $I(cu) = c\,I(u) = cu$ //.

3) **Identity Matrix**: $n \times n \Rightarrow$ All the diagonal elements will be $\underline{1}$ and 0's elsewhere.

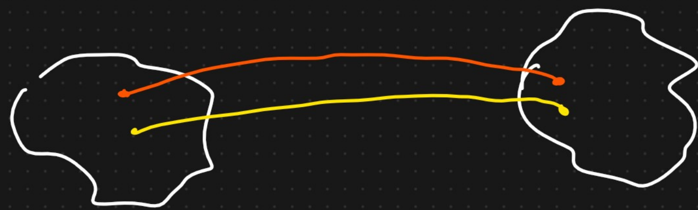
$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3\times3} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

④ Inverse: The Identify fn is its own inverse

Existence And Uniqueness

A function f has an inverse if and only if it is bijective.

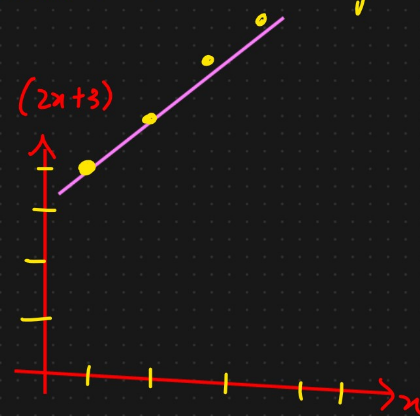
1) Injective (One to one) : Different elements in the domain map to different elements in the codomain.



2) Surjective (Onto) : Every element in the codomain is the image of atleast one element in the domain.

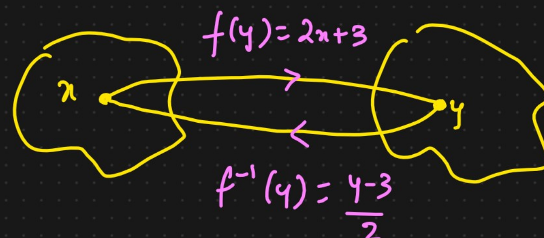
Eg:

Linear function $f(x) = 2x + 3$



Find the Inverse

$y = 2x + 3$ for x:



$y = 2x + 3$

$y - 3 = 2x$

$$\boxed{x = \frac{y-3}{2}}$$

The inverse function $\boxed{f^{-1}(y) = \frac{y-3}{2}}$

Verification

1) $f(f^{-1}(y)) = f\left(\frac{y-3}{2}\right) = \cancel{2}\left(\frac{y-3}{\cancel{2}}\right) + 3 = y - 3 + 3 \Rightarrow y$

2) $f^{-1}(f(x)) = f^{-1}(2x+3) = \frac{(2x+3)-3}{2} = x$

Application of Inverse function In Data Science

1) Normalization And Standardization

Eg: House Price Dataset

Independent features

Dependent feature

(sqft) Carpet Area	No. of Rooms	Area (sqft)	Price (INR or USD)
1500	4	1800	50 lakhs
1800	3	2100	60 lakhs

$\Downarrow$

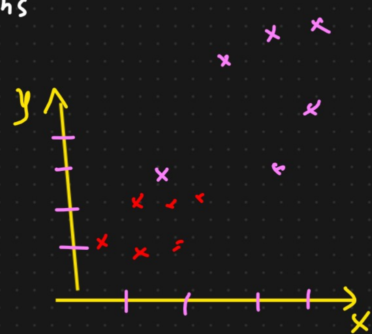
Same Scale

CA, NR, Area $\rightarrow$ ML Model $\rightarrow$ O/p $\Rightarrow$ Price

$\Downarrow$

$\left\{ \text{Mathematical Operation} \right\} \Rightarrow$ Quickly $\Rightarrow$ optimization

[0-1]



$\boxed{1500 \times 1800} \Rightarrow$

$\boxed{1.5 \times 1.5} \Rightarrow$

Standardization $\Rightarrow$ Data feature $\Rightarrow$ feature $\Rightarrow$ $\boxed{\mu = 0 \ \& \ \sigma = 1}$

$\Downarrow$

Standard Normal Distribution

No. of Rooms

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \in R$$

$\Rightarrow$ Standardization Transformation $\Rightarrow$ $Z = \frac{x_i - \mu}{\sigma}$

No. of Rooms $\xrightarrow{f(x)}$ No. of Rooms'

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \in I \quad \underset{f^{-1}(x)}{\overset{f(x)}{\rightleftarrows}} \quad \begin{bmatrix} -1.5/2 \\ -0.5/2 \\ 0.5/2 \\ 1.5/2 \end{bmatrix}$$

$$z_1 = \frac{1-\mu}{\sigma}$$

$$\frac{1-2.5}{2} = \frac{-1.5}{2}$$

$$\mu = \frac{1+2+3+4}{4} = \frac{10}{4} = 2.5$$

$$\sigma = 2$$

$$x = \frac{-1.5}{\cancel{2}} \times \cancel{2} + 2.5$$

$$x_i = 1$$

Original Transformation $\boxed{z = \frac{x-\mu}{\sigma}}$

Inverse Transformation $\boxed{x = z\sigma + \mu}$

$\mu = 0, \sigma = 1$

$\rightarrow$ Standard Scaling $\rightarrow$ ML Model $\rightarrow$ O/P [Scaled Data] 2.5

$\Downarrow$ Inverse Scaling

45 lakhs

Price y

New Test Data $\rightarrow$ Standard Scaling

Inverse Scaling

Use Case:

After training a machine learning model on standardized data, the predictions are often rescaled back to the original scale to interpret the results in a meaningful way. For instance, if house prices were standardized, the inverse transformation would convert the standardized predictions back to the original price scale.

f_1 f_2 f_3 O/P

$\Downarrow$

Standard Scaling

$\downarrow\downarrow\downarrow$

Price O/P $\rightarrow$ Model Train $\rightarrow$ O/P [Scaled O/P] $\Rightarrow$ Normal O/P

Inverse Scaling.

New I/p

Normalization

Feature Scaling with Min Max Normalization

→ Original Transformation : $z = \frac{x - min(x)}{max(x) - min(x)}$ → $T : X \to Y$

→ Inverse Transformation : $x = \frac{z(max(x) - min(x))}{x - min(x)}$ → $T^{-1} : Y \to X$

Distribution of Data

Logarithmic Distribution

Original Transformation $y = \log(x)$

Inverse Transformation $x = e^{y}$

Use Case:
In financial data analysis, income or sales data often exhibit skewness. Applying a log transformation can stabilize the variance and make patterns more visible. After model prediction, the inverse log transformation is applied to interpret the results on the original scale.

5 Data Encryption And Decryption

Encryption Function : $E(p) = c$ (where p is plaintext and c is cipher Text)

Decryption Function : $D(c) = p$.

Use Case:
Sensitive data like personal information, financial records, and medical data are encrypted before storage or transmission. Decryption is applied to retrieve the original information.

(*) How to find Inverse of a Matrix

1. Determinant
2. How To Inverse

Eg: 2x2 Matrix

$A \Rightarrow$ find its inverse and also verify inverse using a transformation

$$A = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}$$

$T : R \rightarrow R$

$\Downarrow$

$A \cdot x = y$

$\Downarrow$

A^{-1}

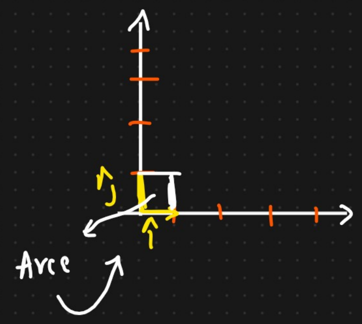
Find the Inverse of A

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

$$\boxed{A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}$$

$\Rightarrow$ High School.

Determinant



The determinant is a scalar value that can be computed from a square matrix. It provides important information about the matrix, such as whether the matrix is invertible (i.e., has an inverse), and it also has geometric interpretations, such as describing the scaling factor of linear transformations represented by the matrix.

Determinant

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \boxed{ad - cb} \Rightarrow \text{scalar} \Rightarrow \text{Determinant.}$$

$$A = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}$$

$\det(A) = 24 - 14 = 10$ → Step 1

$\det(A)$ = Non zero ⇒ Inverse of the matrix

Since the determinant is non zero the matrix A is invertible

② Find the Inverse of A

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \Rightarrow \text{adjacent}$$

$$A^{-1} = \frac{1}{10} \begin{bmatrix} 6 & -7 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{bmatrix}$$

③ Verify using a vector

$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ⇒ x using A then use A^{-1} to recover the Original vector

Transformation using A

$$y = Ax = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4+7 \\ 2+6 \end{bmatrix} = \begin{bmatrix} 11 \\ 8 \end{bmatrix}$$

Recovering x using A^{-1}

$$x = A^{-1} y = \begin{bmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 11 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus A^{-1} successfully recover the Original vector X.

Eigen Vectors And Eigen Values

Eigenvalues and eigenvectors are fundamental concepts in linear algebra that have numerous applications in various fields such as physics, computer science, and data science. They provide insights into the properties of linear transformations represented by matrices.

Defn

Eigen value (λ) : A scaler that indicates how much an eigen vector is stretched or compressed during linear transformation.

Eigen vector (v) : A non zero vector that only changes in scale (not direction) when a linear transformation is applied.

$$Av = \lambda v$$

For square matrix A, an eigen vector and its corresponding eigen value λ satisfy the above equation

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

1) Find Eigen Values

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

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

$$\det(A - \lambda I) = (4-\lambda)(3-\lambda) - (2)(1)$$

$$= (4-\lambda)(3-\lambda) - 2$$

$$= 12 - 4\lambda - 3\lambda + \lambda^2 - 2$$

$= \lambda^2 - 7\lambda + 10$

3) <u>Solve this equation</u>

$$\lambda^2 - 7\lambda + 10 = 0$$

Solve quadratic equation for λ

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$a = 1$, $b = -7$, $c = 10$

$$\lambda = \frac{7 \pm \sqrt{49 - 40}}{2}$$

$$\lambda = \frac{7 \pm \sqrt{9}}{2}$$

$$\lambda_1 = \frac{7+3}{2} \qquad \lambda_2 = \frac{7-3}{2}$$

Eigen values of A. $\boxed{\lambda_1 = 5 \qquad \lambda_2 = 2}$

2) <u>Find the Eigen Vectors</u> :

$$(A - \lambda I)\, v = 0$$

For $\lambda_1 = 5$

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

$$= \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$-x + y = 0$$

$\boxed{y = x}$

For
$\lambda_2 = 2$

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

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\Downarrow$$

$$2x + y = 0$$

An eigen vector corresponding to $\lambda_2 = 2$ $\boxed{y = -2x}$

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} \qquad \lambda_1 = 5 \qquad \lambda_2 = 2$$

For $\lambda_1 = 5$ $\quad v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ //

For $\lambda_2 = 2$ $\quad v_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ //.

$\boxed{A\vec{v} = \lambda \vec{v}}$

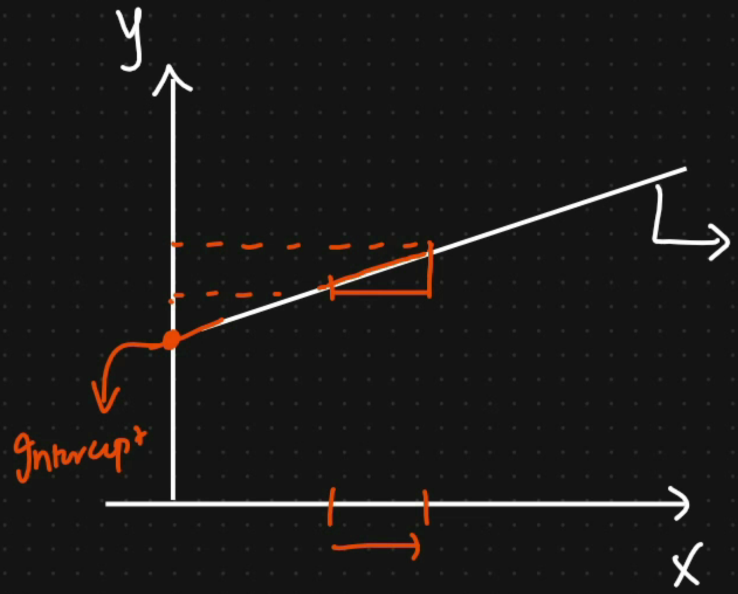
Eigen value

These eigenvectors and eigenvalues describe how the matrix A A scales and rotates vectors in its transformation. Eigenvalues indicate the factor by which the eigenvectors are stretched or compressed, and eigenvectors provide the directions in which this stretching or compression occurs.

Application :

Principal Component Analysis → Dimensionality Reduction

Equation of Line, 3d plane and Hyperplane (n Dimension)



$x = 0$

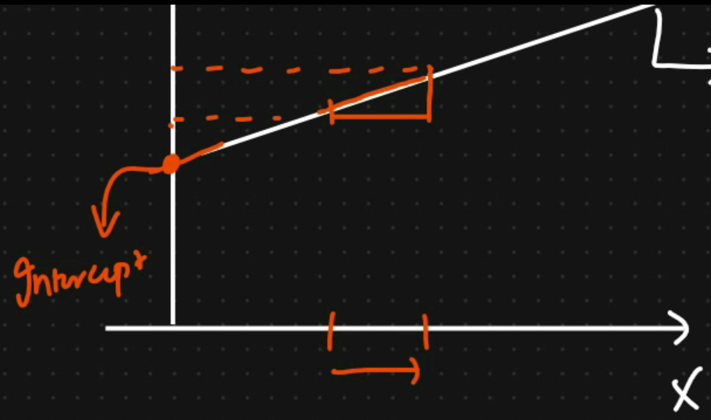
$y = mx + c$

m = Slope

C = Intercept

$y = \beta_0 + \beta_1 x$

$ax + by + c = 0$

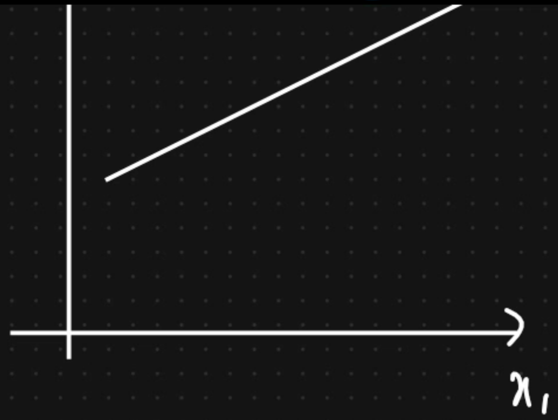


$$y = mx + c \}$$

C = Intercept

$$y = \beta_0 + \beta_1 x$$

$$ax + by + c = 0 \}$$



$$by + c = -ax$$

$$by = -ax - c$$

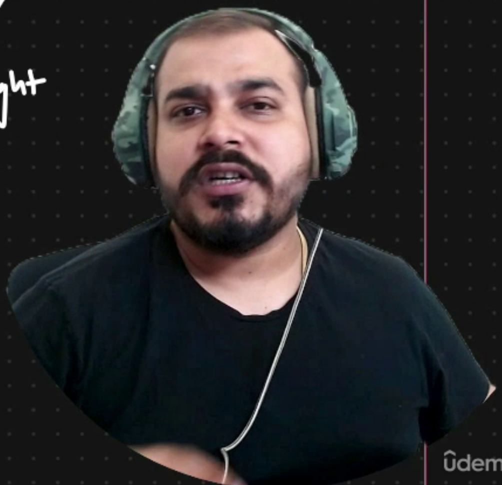
$$y = mx + c \Leftarrow y = \underbrace{\boxed{\frac{-a}{b}}}_{m} x - \boxed{\frac{c}{b}} \rightarrow c$$

$$w_1 x_1 + w_2 x_2 + b = 0$$

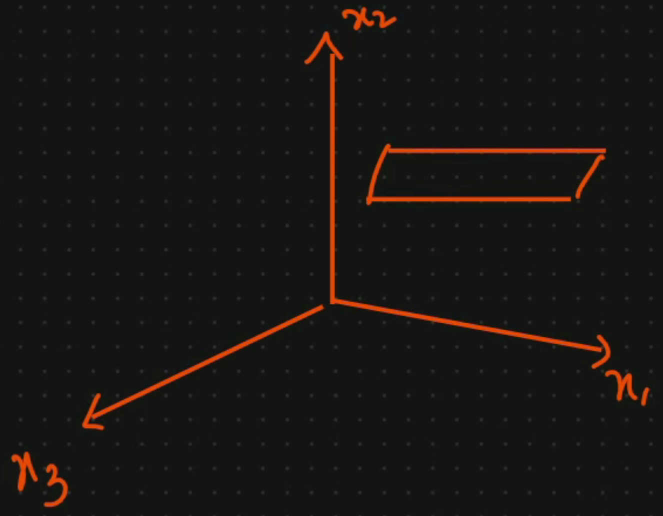
$$\boxed{w^T x + b = 0}$$

$$\Downarrow$$

Eq of a straight line



m



$$w_1x_1 + w_2x_2 + w_3x_3 + b = 0$$

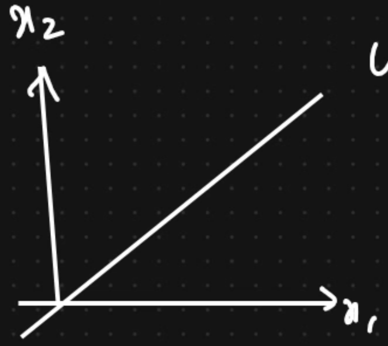
$$\boxed{w^Tx + b = 0}$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \cdot x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

n-Dimension plane

$$w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n + b = 0$$

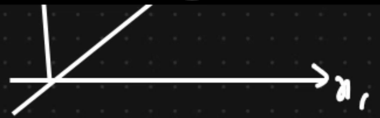
$$\boxed{w^Tx + b = 0}$$



$$w_1x_1 + w_2x_2 + b = 0$$

$$w_1x_1 + w_2x_2 = 0$$

$$w^Tx = 0$$



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

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

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \cdot x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

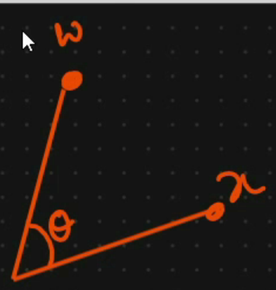
Equation of a Straight passing through an origin

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

Equation of a plane = Π_n : w^T

Untitled (2/2)

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

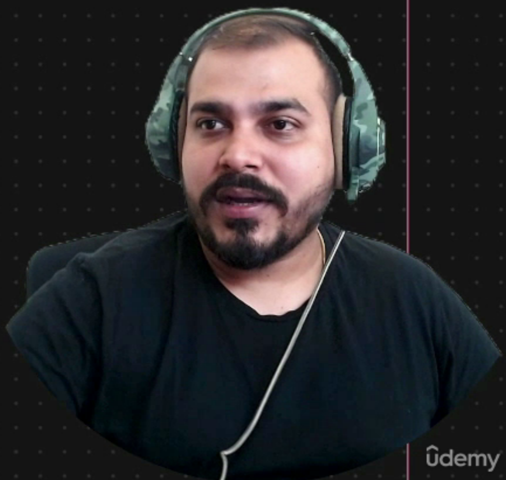


$$w^T x = 0$$

$$w \cdot x = w^T x = \|w\| \, \|x\| \cos\theta = 0$$

$$\theta = 90$$

$$\cos\theta = 0$$



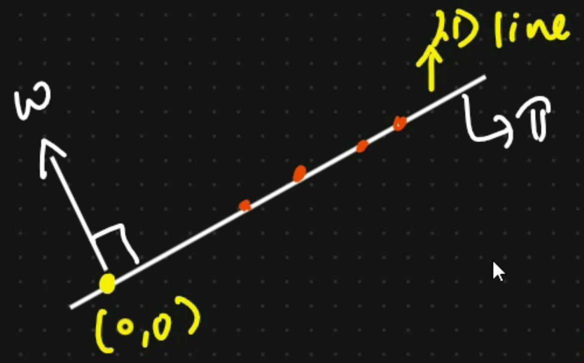
Untitled (2/2)

$$w^T x = 0$$

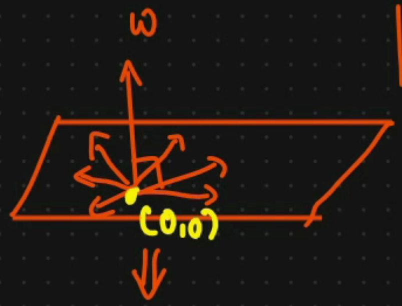
$$w \cdot x = w^T x = \|w\| \, \|x\| \cos\theta = 0$$

$\theta = 90$

$\cos\theta = 0$



intercept = 0



$$\boxed{w \perp \pi}$$

$$w^T x = 0$$