Nave - Bayes Model.

PAGE NO. :
DATE : / /

→ Bayes Theorem :-

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Ex : of Wrenches.

→ Mach 1: 30 wrenches / hr
Mach 2: 20 wrenches / hr

→ Out of all produced parts :
we can see that 1% are defective

→ Out of all defective parts :-
we can see that 50% came from [mach 1]
and 50% came from [Mach 2]

Q. What is the probability that a part produced by Mach 2 is defective = ?

↳ $P(M1) = 30/50 = 0.6$

↳ Pro. of picking wrench that was produced by M1

↳ $P(M_2) = 20/50 = 0.4$

↳ $P(\text{Defect}) = 1\%$

PAGE NO. :
DATE : / /

↳ P (M1 / Defect) = 50%

↳ P (M2 / Defect) = 50% → { This machine is producing more defective parts }

↳ The Question that was asked ⇒

{ Probability by Mach 2 is defective } = ?

⇓

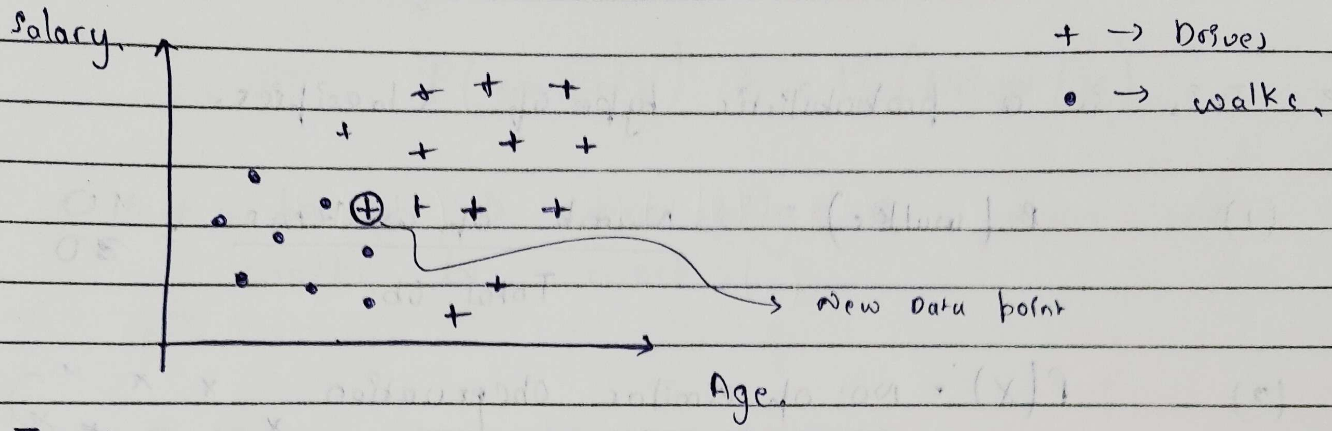
$$P(\text{Defect} \mid M2) = ?$$

$$P(\text{Defect} / M2) = \frac{P(M2 / \text{Defect}) \cdot P(\text{Defect})}{P(M2)}$$

$$= \frac{(0.5) \cdot (0.01)}{0.4} = 0.0125 = 1.25\%$$

So we can say if machine produces 1000 parts 125 will be defective.

Intuition :-



This data represents Salary and Age of the people and How do they go to their office.

$$① \left\{ \underbrace{P(\text{walks}/X)}_{\text{Posterior Prob.}} = \frac{\overbrace{P(X/\text{walks})}^{\text{Likelihood}} \cdot \overbrace{P(\text{walks})}^{\text{Prior Prob.}}}{\underbrace{P(X)}_{\text{marginal}}} \right\}$$

$$② \quad P(\text{Drives}/X) = \frac{P(X/\text{Drives}) \cdot P(\text{Drives})}{P(X)}$$

Then compare Both.

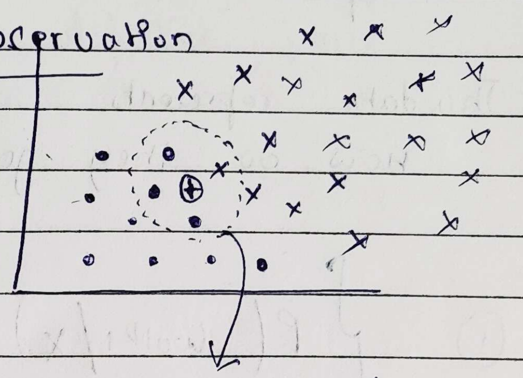
$$P(\text{walks} \mid X) \quad \text{V.s.} \quad P(\text{Drives} \mid X)$$

→ then we will Decide in which category we should put this Data point in.

→ This is a probabilistic type of classifier.

(1) $P(\text{walks}) = \dfrac{\text{Numb. of walkers}}{\text{Total ob.}} = \dfrac{10}{30}$

(2) $P(X) = \dfrac{\text{No. of similar observation}}{\text{Total obs.}}$



$P(X) = \dfrac{4}{30}$

$P(X \mid \text{walks}) = \dfrac{3}{10}$

anything that falls in that circle will be considered similar to the new point that we are adding to the circle.

$P(\text{walks}/X) = \dfrac{\frac{3}{10} \cdot \frac{10}{30}}{\frac{4}{30}} = 0.75$

Similarly

$P(\text{Drives}/X) = \dfrac{\frac{1}{20} \cdot \frac{20}{30}}{\frac{4}{30}} = 0.25$

So the Prob. of people who exhibits features X bean the person who Drives to work is 0.25

PAGE NO. :
DATE : / /

so clearly,

$$P(\text{walks}/X) > P(\text{Drives}/X)$$

that person is probably the person who walks to work.

Extras.

* Principle Component Analysis - (PCA)

→ most popular unsupervised Algorithm, Best Dimentionality redution Algorithm.

Used for operations such as
- Noise filtering
- feature Extraction
- stock market preditions
- Gene data analysis.

Work of PCA is →

(1) Identify Patterns in data.
(2) detect the correlation b/w variables.

The goal is to Reduce the dimensions of a d-dimensional dataset by projecting it onto a k-dimensional subspace (where $k < d$)

Scanned with ACE Scanner

PAGE NO. :
DATE : / /

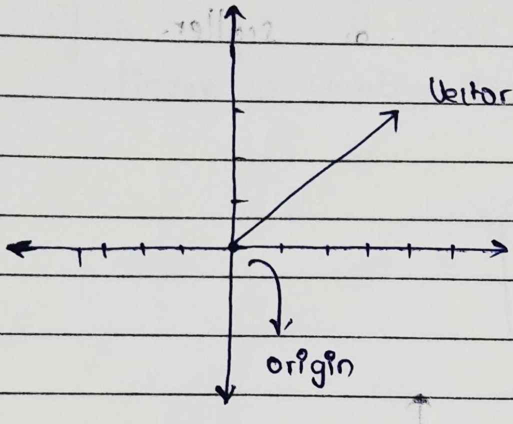
Process is

- standardize the data.
- obtain the Eigenvectors and Eigenvalues from the covariance matrix or correlation matrix or perform singular Vector Decomposition.
- sort eigenvalues in descending order and choose the K is the number of dimensions of the new features subspaces ($K \le d$)
- construct the projection matrix W from the selected K eigenvectors.
- Transform the original dataset X via W to obtain a K - dimentional features subspace Y.

PCA tries to
- learn abt the relationship b/w X and Y values
- find list of principal axes.

linear Algebra

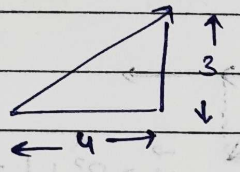
- Vectors.



Vector $\begin{array}{c|c} \begin{bmatrix}4\\3\end{bmatrix} & \begin{matrix}x\\y\end{matrix}\end{array}$

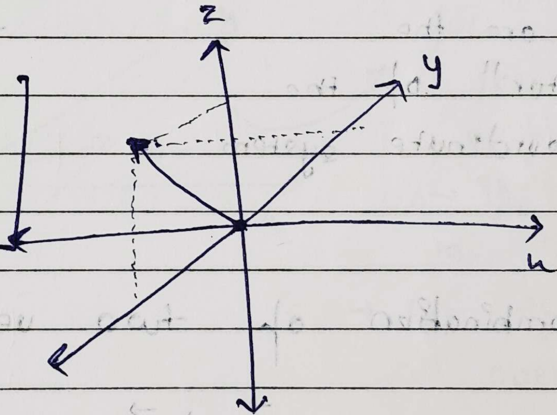
Notation. (2 dimentional)

- in linear algebra in most of the cases vector will originate from origin

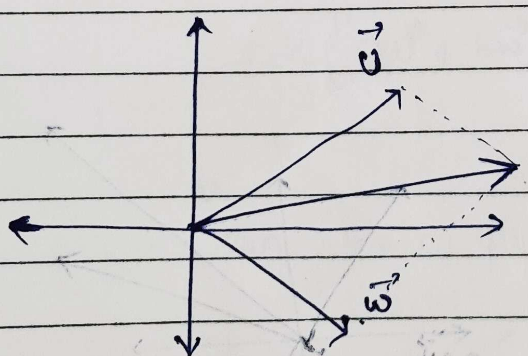


- It can also be in 3 dimentional.

$$\begin{array}{cc} x & \\ y & \begin{bmatrix}2\\4\\6\end{bmatrix} \\ z & \end{array}$$



- addition of 2 vector:-

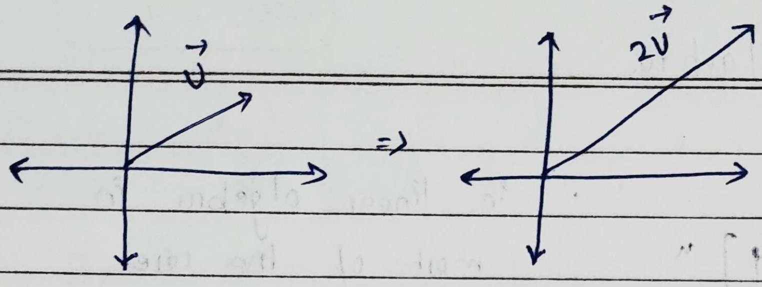


if $\vec{v} = \begin{bmatrix}1\\2\end{bmatrix}$ $\vec{w} = \begin{bmatrix}3\\-1\end{bmatrix}$

$\vec{v}+\vec{w} = \begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}3\\-1\end{bmatrix}$

$= \begin{bmatrix}4\\1\end{bmatrix}$

- multiplication of vector.

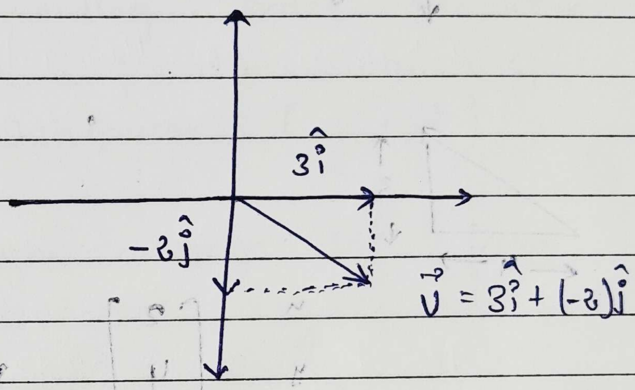


Here '2' is called as scaller.

→ $2 \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

*.

let $\vec{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$



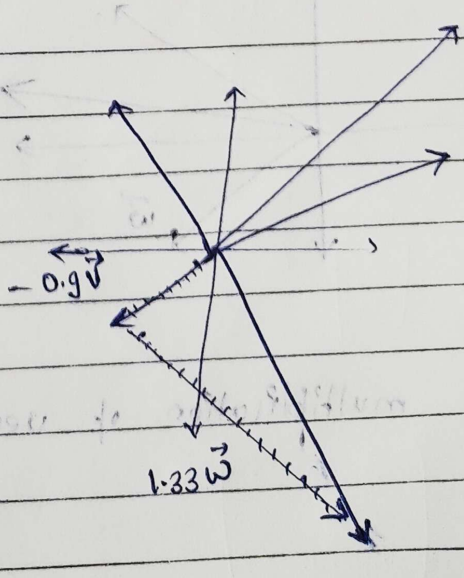
$\hat{i}$ and $\hat{j}$ are the "basis vector" of the xy coordinate system.

linear combination of two vector

$= a\vec{v} + b\vec{w}$

→ if one of the a or b is fixed then resulting vector of combination will form a "straight line"

→ if non is fixed then this combination is reach every vector possible in xy plane

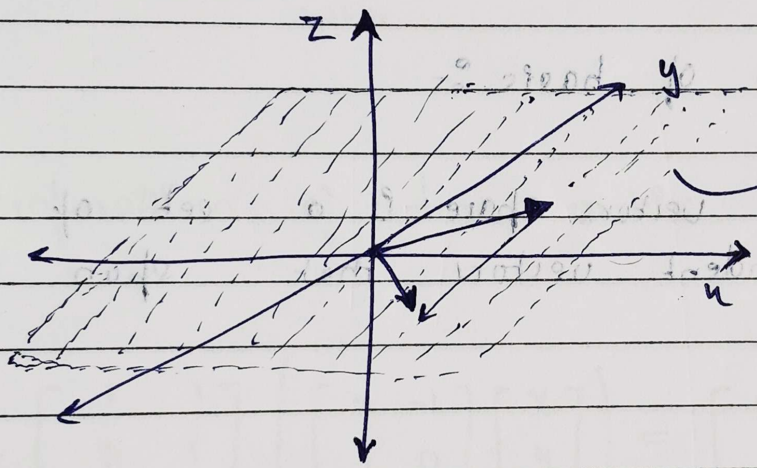


- Span of $\vec{v}$ and $\vec{w}$ is the set of all their linear combinations

$$a\vec{v} + b\vec{w}$$

let a and b vary over all real numbers.

→ lets talk about span of two vector in 3-dimensions



it will form a plane.
→ if we consider the tip
→ the tip can move all over a plane

→ linear combination of $\vec{v}$, $\vec{w}$, and $\vec{u}$

$$\{a\vec{v} + b\vec{w} + c\vec{u}\}$$

for span, let there constants vary.

can cover all the 3-D space.

- linearly dependent :-

$$\vec{u} = a\vec{v} + b\vec{w}$$

for all values of a and b

- linearly in-dependent :-

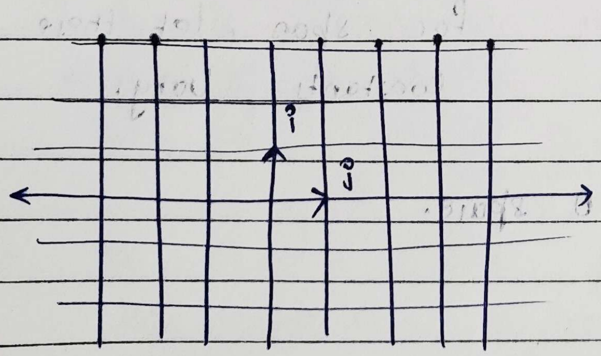
$$\vec{u} \neq a\vec{v} + b\vec{w}$$

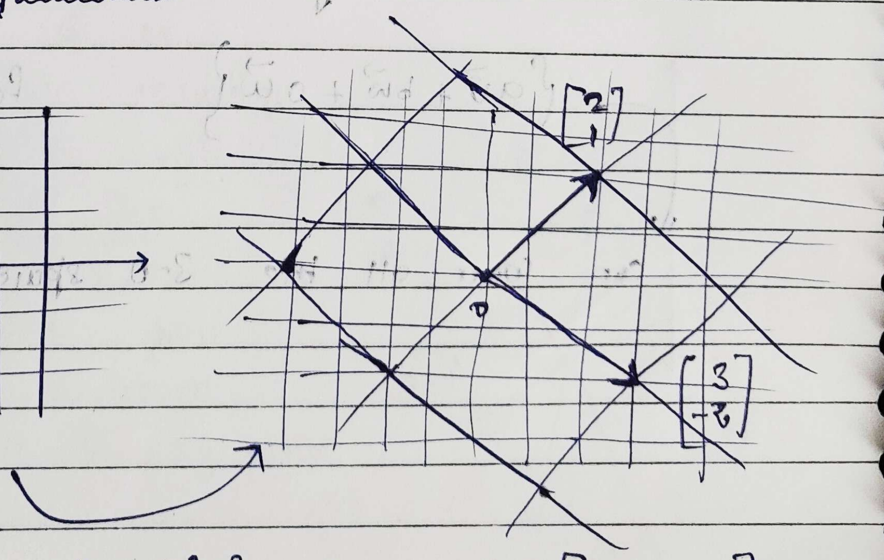
for all of a and b

* Technical definition of basis :-

→ The basis of a vector space is a set of linearly independent vectors that span the full space.

• * linear transformation :-





notice

where does the i, j vector went after transformation $= \begin{bmatrix} 3 & 2 \\ -2 & 1 \end{bmatrix}$

PAGE NO. :
DATE : / /

• 2x2 matrix

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

where $\hat{i}$ lands   where $\hat{j}$ lands.

→ and vector multiplication computes what that ~~vector~~ transformation does to a vector.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\text{vector.}} = x\begin{bmatrix} a \\ c \end{bmatrix} + y\begin{bmatrix} b \\ d \end{bmatrix} = \underbrace{\begin{bmatrix} ax + by \\ ax + dy \end{bmatrix}}_{\text{Transformed vector}}$$

* Composition of two matrix :-

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{\substack{\text{Shear} \\ 2}}\left(\underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\substack{\text{Rotation} \\ 1}}\begin{bmatrix} x \\ y \end{bmatrix}\right) = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}}_{\text{composition}}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{2}\underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{1} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

Readed second.   Readed first

$(M_1 M_2 \neq M_2 M_1)$ → order matter's

$A(BC) = (AB)C \longrightarrow$ Associative Properties