1. Data ingestion
2. EDA
3. Preprocessing ⇒ feature engineering
4. Model building
5. Model evalution

Dimension Reduction → Reducing the dimension

↓

Feature

Feature Selection

Feature extraction

Curse of Dimensionality :- Data

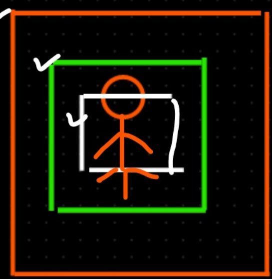
(Row x column)

$f_1 \ldots f_{10} \; Y \rightarrow M_1 \Rightarrow 75\%$ — less

$f_1 \ldots f_{100} \; Y \rightarrow M_2 \Rightarrow 80\%$ — optimal

$f_1 \ldots f_{1000} \; Y \rightarrow M_3 \Rightarrow 72\%$ — over ⇒ overfitting ⇒ Sparsity

Underfitting



① Feature selection ⇒ Subset of the feature

$\{1, 2, 3, 4, 5, 6, 7\}$

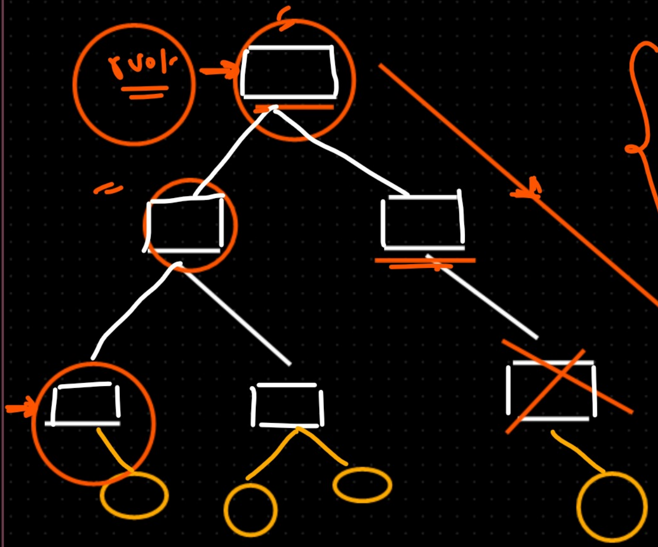
↓

Set theory $\{1, 2, 3\}$

$f_1 ----- f_{100}$ ⇒ Model ⇒ 75%

↓

50, 60, 70, 80



Scikitlearn

① Filter method ⇒ Corr

Info gain → feature importance

② wrapper method.

③ lasso

Wafer fault Detection

⇓

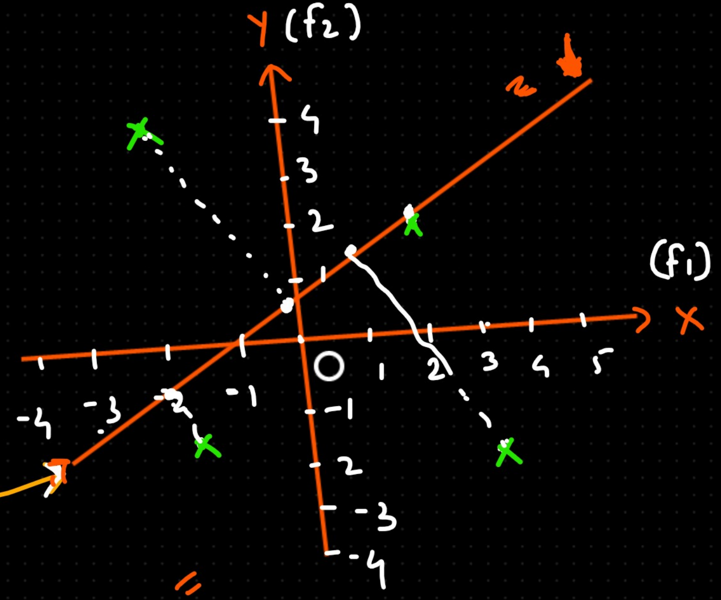
500

= $\boxed{\text{feature selection}}$

Dimension Reduction | Feature extraction | feature transformation

f_1	f_2
2	2
3	−2
−2	−1
−2	4

$\Rightarrow F$



this vector having essance of both feature?

1. PCA
2. t-SNE
3. LDA

} state of art

1. Standarized the Data
2. $\boxed{\text{Covariance matrix}}$
3. Eigen value and eigen vector
4. Principle component

PCA

1. Scaling of the data

Min-Max ⇒ [0-1]

Standard Scalar ⇒

Z-Score

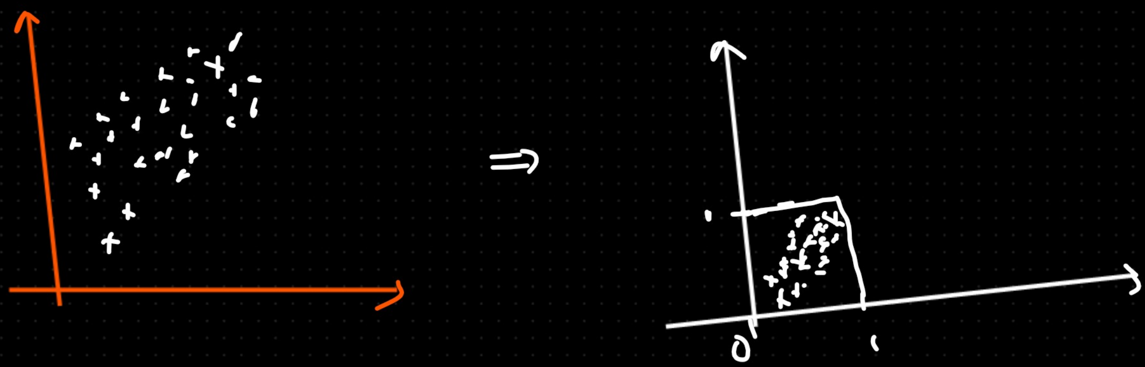
$$\text{Min-Max} = \boxed{\frac{X - \text{Min}}{\text{Max} - \text{Min}}}$$

$[a_1\, a_2\, a_3\, a_4\, \cdots\, a_n]$

Min   Max

$$= \frac{\cancel{\text{Min}} - \cancel{\text{Min}}}{\text{Max} - \text{Min}} = 0$$

$$= \frac{\text{Max} \cancel{-} \text{Min}}{\cancel{\text{Max} - \text{Min}}} = 1$$



SND

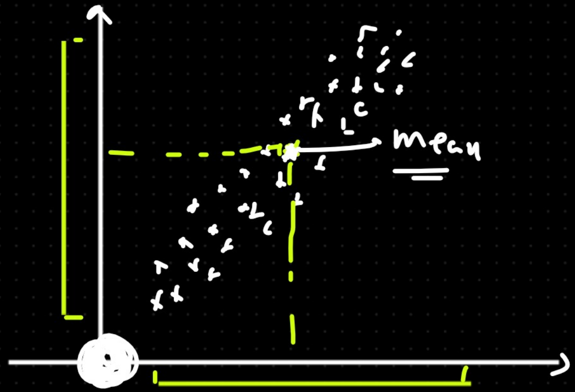
Standardization $\Rightarrow \dfrac{x-\mu}{\sigma}$

Z-Score

$$\left\{ \begin{array}{l} \mu = \text{mean} \\ \sigma = \text{std} \end{array} \right\}$$

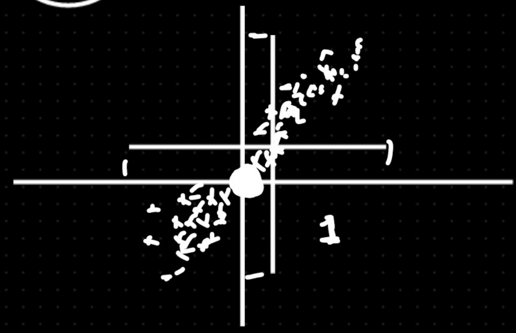
$$\Downarrow$$

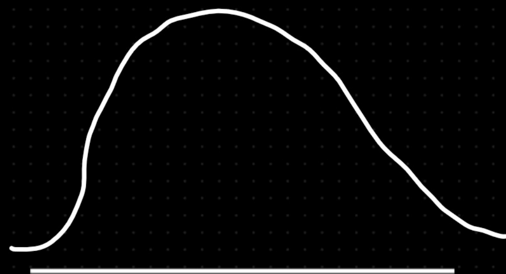
Data $\rightarrow$ standardization $\Rightarrow$ $\left\{ \begin{array}{c} \text{Standard Data} \\ \downarrow \quad \downarrow \\ \mu = 0 \quad \sigma = 1 \end{array} \right\}$



SND

mean = 0 std = 1





Variance =>

Spread of the data around the mean

2, 8, 10, 12, 6, 5, 13, 14

↓

Mean

↓

deviation => (Mean - DP)^2

↓

= mean of deviation

$$\boxed{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Variance → Cov-variance

x, x

x → f_1	y → f_2

$$\boxed{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})} \leftarrow Cov(x,y)$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})(x_i-\bar{x}) \qquad cov(x,x)$$

$$\Downarrow$$

$$= \boxed{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

$$\mathrm{cov}(x,y) = \mathrm{cov}(y,x)$$

$x \mid y$

$$= \begin{bmatrix} \mathrm{cov}(x,x) & \mathrm{cov}(x,y) \\ \mathrm{cov}(y,x) & \mathrm{cov}(y,y) \end{bmatrix} \Rightarrow$$ Covariance matrix of x,y feature

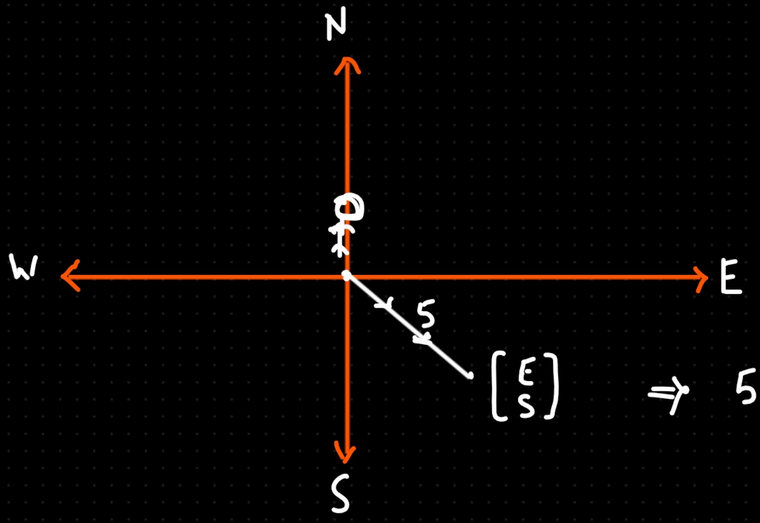
1 Std = mean = 0

$$\Rightarrow \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$$

mean mean

$$\Rightarrow \frac{1}{N} \sum_{i=1}^{N} (x_i)(y_i)$$

$$\Rightarrow \frac{\boxed{x_1 y_1 + x_2 y_2 + x_3 y_3 \cdots x_n y_n}}{N} \Rightarrow$$ Dot Product of two vector

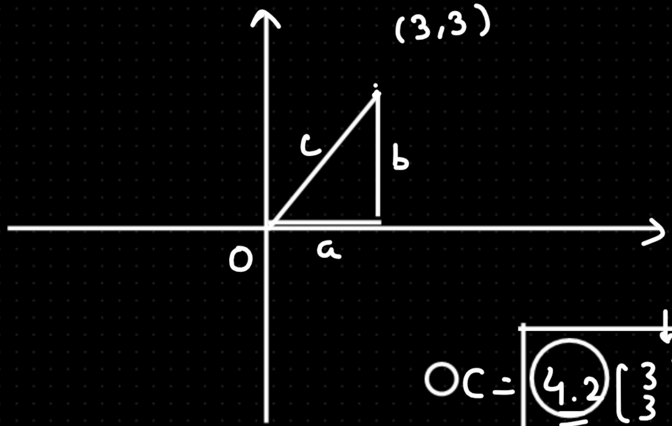


$\Rightarrow 5\,\text{km} \begin{bmatrix} E \\ S \end{bmatrix}$

Scalar value (Magnitude)

Vector

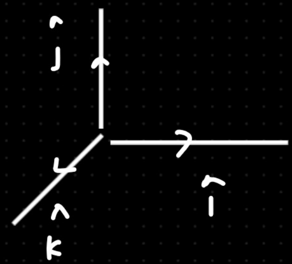
5 KM   1hr



$c^2 = a^2 + b^2$

$c = \sqrt{a^2+b^2} = \sqrt{3^2+3^2}$

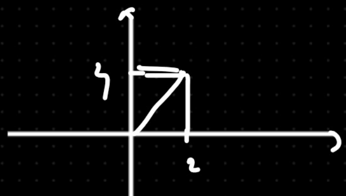
$= \sqrt{18} = 4.2$



$\boxed{OC = 4.2 \begin{bmatrix} 3 \\ 3 \end{bmatrix}}$

$3\hat{i} + 3\hat{j}$

$V = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$



$\sqrt{2^2+4^2}$

$= \sqrt{4+16} = \sqrt{20}$

Eigen value & eigen vector

$$\boxed{AV = \lambda V}$$

A = Squre Matrix

V = Vector

λ = Scalar value

if we multiplying a squre Matrix A with vector V and if we get a New vector $\boxed{\lambda V}$ that is λ times of vector V, then vector $\boxed{V}$ is called $\boxed{\text{eigen vector of}}$ Matrix $\boxed{A}$

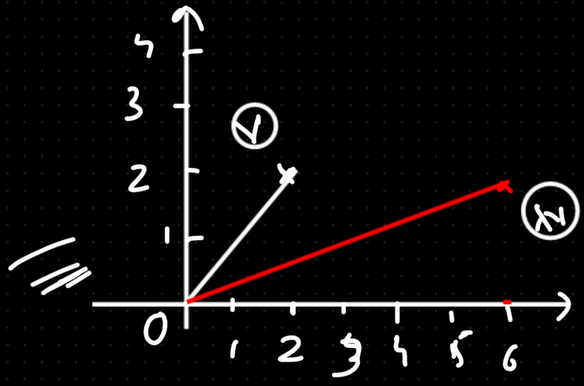
= if we multiply matrix A by vector V the new vector $\boxed{\lambda V}$ does not change the direction. than only it is eigen vector?

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

$V = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$$\boxed{A \times V = \lambda \times V}$$

$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} 2 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1\times 2 + 2\times 2 \\ 1\times 2 + 2\times 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2+4 \\ 2+0 \end{bmatrix} \Rightarrow \begin{bmatrix} 6 \\ 2 \end{bmatrix}$



$$A\underline{v} = \begin{bmatrix}6\\2\end{bmatrix} = \lambda v$$

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

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

$$A\,\underline{v} = \lambda\,\underline{v}$$

$V \rightarrow A$

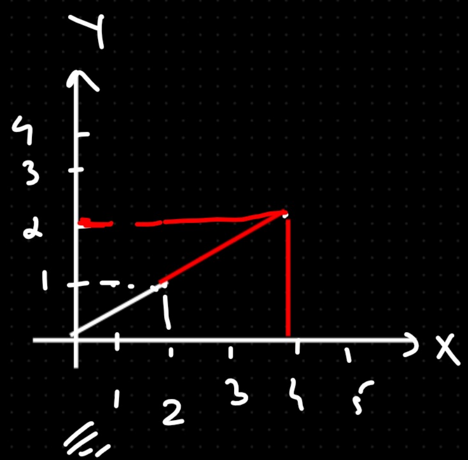
No

$$\begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix} \times \begin{bmatrix}2\\ 1\end{bmatrix} = \begin{bmatrix}1\times 2 + 2\times 1\\ 1\times 2 + 0\times 1\end{bmatrix} = \begin{bmatrix}2+2\\ 2+0\end{bmatrix} \Rightarrow \begin{bmatrix}4\\ 2\end{bmatrix}$$

λv

$$\boxed{A \times V = \lambda \times V}$$

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



$$\Rightarrow \begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix} \boxed{\begin{bmatrix}2\\ 1\end{bmatrix}} = 2 \begin{bmatrix}2\\ 1\end{bmatrix}$$

λ = eigen value

are this vector is eigen vector of matrix A

$$\boxed{\text{Principle component}}$$

Eigen value tells us how much the eigen vector changes in size when we multiply with matrix

$A \rightarrow v$

λv

$$\boxed{Av = \lambda v}$$

$$\boxed{Av - \lambda v = 0}$$

$$\boxed{(A - \lambda I)v = 0}$$

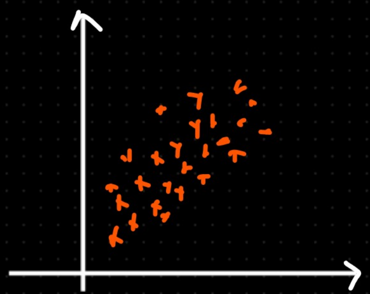
$$\boxed{\text{Det } |A - \lambda I|}$$

I = Identity matrix

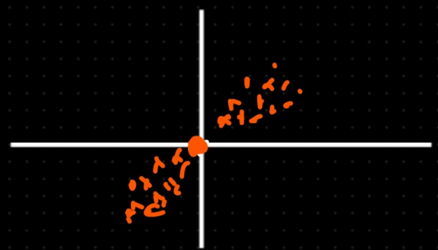
f_1	f_2

$\Rightarrow$

1. Standarization
2. Cov Matrix
3. eigen value & eigen vector
4. Principle component



$\Rightarrow$



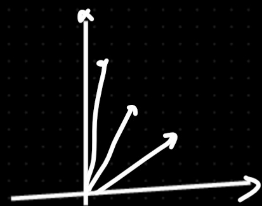
$\Rightarrow$

$$\begin{bmatrix} cov(f_1 f_2) & cov(f_1 f_1) \\ cov(f_2, f_2) & cov(f_2 f_1) \end{bmatrix}$$

$$\boxed{Av = \lambda v}$$

eigen vector

eigen value



μ_i

$\mu_i \times$ Data

Principle component