Principal Component Analysis

10 June 2023 07:13 PM

x_1	x_2	x_3	x_4	x_5	Y

x_1	x_3	Y

→ x_2, x_4, x_5 not considering

x_1	x_2	x_3	x_4	x_5	Y

Accuracy

0.91

By consideration of all the feature

independepen

PC1	PC_2	PC_3	PC_4	PC_5	Y
78%	21	0.01	0.03	0 ----	

Unsupervised

By considering Only 2 PC 0.92 ← same accuracy

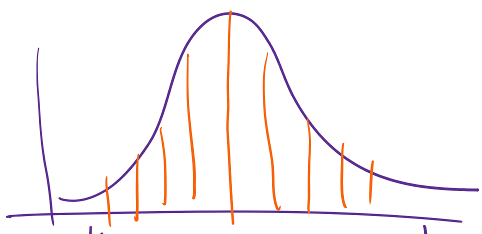
5D → 2D

Feature extraction

Dimensionality Reduction

Variance :-

var is more information is more



Var = spread

Var_thresh=0 → drop

Spread

X_1 more var

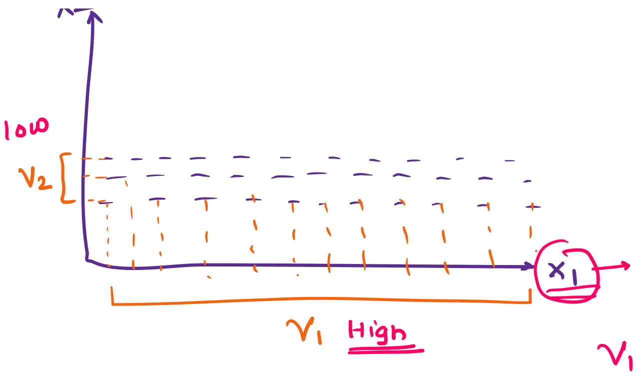
X_2

Var X_1 > Var X_2

X_3

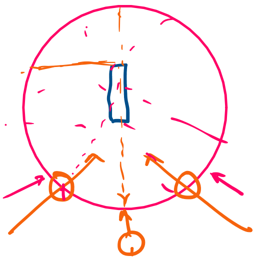
Var=0 | spread=0 → drop

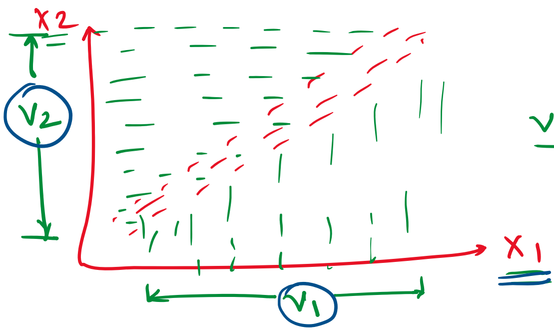
X_2 ↑



x_1 giving more info. than x_2

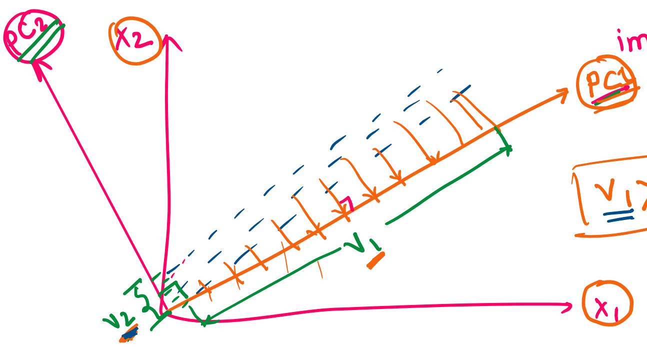
$V_1 > V_2$





$V_1 \cong V_2$

PCA → Ultimate ~~Goal~~ → To get maximum Variance



imaginary ⊥r orthogonal projections

$\boxed{V_1 > V_2}$

x_1 and x_2 → PC_1 and PC_2

$\boxed{\text{No. of features = No. of principal components}}$

X_1	X_2	X_2	X_3	X_4	Y

5D

⟹

PC_1	PC_2	PC_3	PC_4	PC_5	Y
81%	15%	1%	0.2	-	

% Variance

information

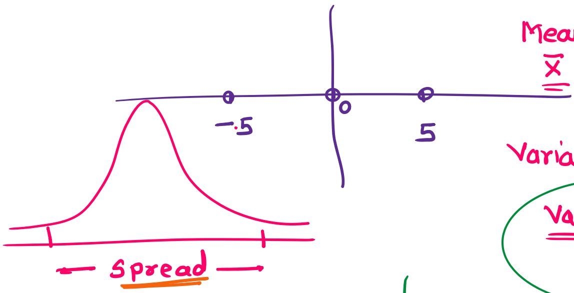
information

2D

Accuracy 90%

Variance: $= \dfrac{\sum (x-\bar{x})^2}{N}$

Mean $= \dfrac{\sum x}{N}$



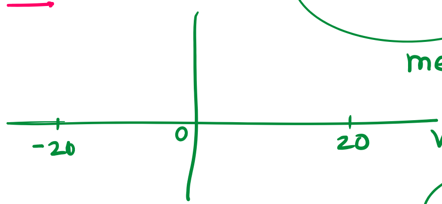
Mean $\bar{X} = \dfrac{-5+0+5}{3}$

$= \dfrac{0}{3} = 0$

Variance $= \dfrac{(-5-0)^2+(0-0)^2+(5-0)^2}{3}$

Var $= \dfrac{50}{3}$ ← magnitude

mean = 0



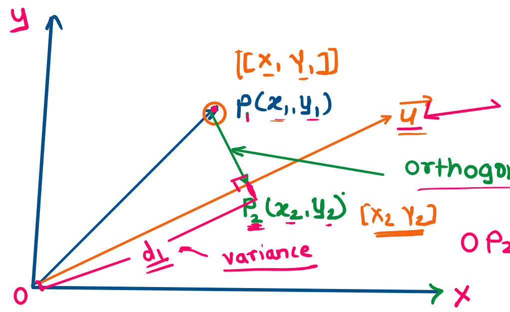
Var $= \dfrac{(-20)^2+(20)^2}{3}$

Var $= \dfrac{800}{3}$

magnitude

Var ∝ spread

'Variance' is scalar term.



unitvector giving maximum variance

orthogonal projection

$OP_2 = d_1 =$ variance

$d_1 = \dfrac{\vec{P_1} \cdot \vec{u}^T}{|\vec{u}|} = \dfrac{\vec{P_1} \cdot \vec{u}^T}{1}$

$d_1 = \vec{P_1} \cdot \vec{u}^T$

$\vec{P_1} = [x_1 \ y_1]_{1\times 2}$

$\vec{u} = [x_2 \ y_2]_{1\times 2}$

$d_1 = [x_1 \ y_1]_{1\times 2} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}_{2\times 1}$

$\vec{P_1} \cdot \vec{u} =$

$[1\times 2] \ [1\times 2]$

X

$d_1 = [x_1 x_2 + y_1 y_2]_{1\times 1}$

$\vec{P_1} \cdot \vec{u}^T = [1\times 2] \begin{bmatrix} 2 \\ \times \\ 1 \end{bmatrix}$

Variance = Scalar quantity

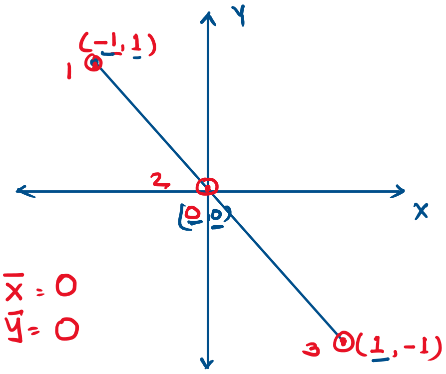
Var = 5

Covariance $= \dfrac{\sum (x_i-\bar{x}) \cdot (y_i-\bar{y})}{N}$

Range → $-\infty$ to $+\infty$

Y

Range $\rightarrow -\infty$ to $+\infty$



$\bar{x} = 0$
$\bar{y} = 0$

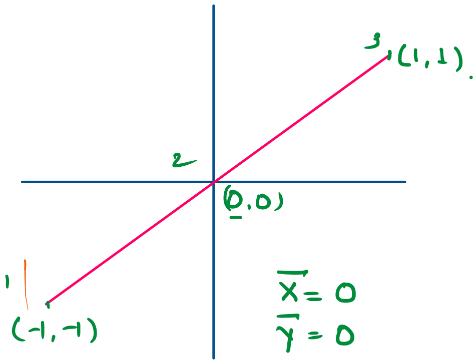
$$cov = \frac{\sum(x-\bar{x})(y-\bar{y})}{N}$$

$$= \frac{(1)\times(-1) + 0 + (-1)(1)}{3}$$

$$= \frac{(-1) + 0 - 1}{3}$$

$$= -\frac{2}{3}$$

direction of spread of data



$\bar{x} = 0$
$\bar{y} = 0$

$$cov = \frac{\sum(x-\bar{x})(y-\bar{y})}{N}$$

$$= \frac{(-1)(-1) + 0 + (1)(1)}{3}$$

$$= \frac{+1 + 0 + 1}{3}$$

$$+\frac{2}{3}$$

$$\text{Variance} = \frac{\sum(x-\bar{x})^2}{N} = \frac{1+0+1}{3}$$

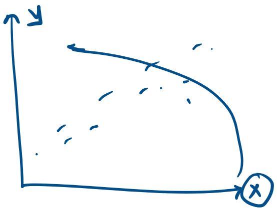
$$\text{Variance} = \frac{2}{3}$$

$$\text{Variance} = \frac{2}{3}$$

Var $\rightarrow$ spread

Covar $\rightarrow$ spread + direction

correlation :- $= \dfrac{cov(x, y)}{std(x_1) \cdot std(x_2)}$



Strength

$$\begin{array}{ccc} & x_1 & x_2 \\ x_1 & \left[\begin{matrix} \underset{Var(x_1)}{cov(x_1, x_1)} & cov(x_1, x_2) \\ cov(x_2, x_1) & \underset{Var(x_2)}{cov(x_2, x_2)} \end{matrix}\right] & \\ x_2 & & \end{array}$$

$$cov(x_1, x_1) = \frac{\sum(x_i-\bar{x})(x_i-\bar{x})}{N}$$

$$= \sum(x_i - \bar{x})^2$$

$x_2 \begin{vmatrix} cov(x_2, x_1) & cov(x_2, x_2) \\ & var(x_2) \end{vmatrix}$

$cov(x_1, x_1) = \frac{\ldots}{N}$

Var $\rightarrow$ $= \frac{\sum (x_i - \bar{x})^2}{N}$

- Covariance Matrix

$A = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}_{2\times2}$ $\quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times2}$

✓ $AX = \lambda I X$

Identity

$\lambda \rightarrow$ Scalar value 5, 10, ...

$AX - \lambda I . X = 0$

$(A - \lambda I) . X = 0$

Eigen vector

Eigen values

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

$A - \lambda I = \begin{vmatrix} 2-\lambda & 3 \\ 4 & 3-\lambda \end{vmatrix}$

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

$|A| = [a_{11} . a_{22} - a_{12} . a_{21}]$

$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 3 \\ 4 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda) - (4 \times 3)$

$= 6 - 2\lambda - 3\lambda + \lambda^2 - 12$

$= \lambda^2 - 5\lambda - 6 = 0$

$-6 \rightarrow -6, +1$

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

$= \lambda(\lambda - 6) + (\lambda - 6) = 0$

$(\lambda + 1)(\lambda - 6) = 0$

$(\lambda + 1) = 0 \quad$ or $(\lambda - 6) = 0$

Eigen Values $\quad \boxed{\lambda = -1 \quad \text{or} \quad \lambda = 6}$

① Sort eigen values in descending order

$\lambda_1 = 6$ and $\lambda_2 = -1$

Eigen ...

$\lambda_1 = 6$ and $\lambda_2 = -1$

Eigen vector

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

First eigen Vector

$$\left(\begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix} - 6 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\left(\begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

How to Solve $\Rightarrow \begin{bmatrix} -4 & 3 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

identity Row → colum

$R_1 \rightarrow -4$

Row Echolen Method $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$

$\lambda_1 = 6$

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

Eigen vector ⇒ PC1

$\lambda_2 = -1$

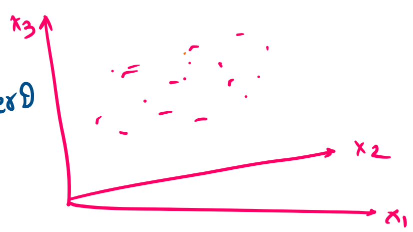
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix}$$

Eigenvector2 ⇒ PC2

$x_1 \quad x_2 \longrightarrow PC_2 \quad PC_2$

Data → 1000 × 3
Rows colowns

Higher D → lower D



3D → 2D → PCA

x_1

$3D \to 2D \to$ PCA $\to$ Variance

① Find Covariance $\to$ 3×3

$$\begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline x_1 & Var(x_1) & cov(x_1,x_2) & cov(x_1,x_3) \\ x_2 & cov(x_2,x_1) & Var(x_2) & cov(x_2x_3) \\ x_3 & cov(x_3x_1) & cov(x_3x_2) & Var(x_3) \end{array}$$

② Eigen Decomposition < Eigen Values, Eigen vector

$$Ax = \lambda I x = 0$$

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

Eigen values

$\lambda_1 = 5$ $\quad\quad \lambda_2 = -2 \quad\quad \lambda_3 = 4$

$\downarrow$

Eigen vector

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 9 \\ 8 \end{bmatrix}$$

$= \lambda^2 - 5\lambda - 6 = 0$ polynomial

$$ax^2 + bx + c = 0$$

$a = 1, \; b = -5, \; c = -6$

Indivisual Eigen vect = 1×3

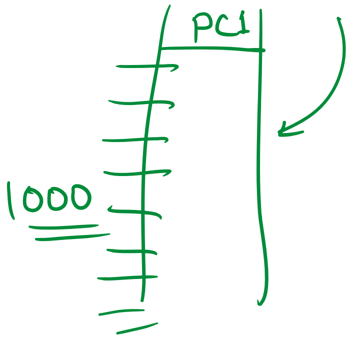
③ Transformation

$$\vec{P} = 1000 \times 3$$

$$\vec{\hat{u}} = 1 \times 3$$

$$\vec{P} \cdot \vec{\hat{u}}^T = \underbrace{[1000 \times 3]}_{\text{size}} \times [3 \times 1]$$

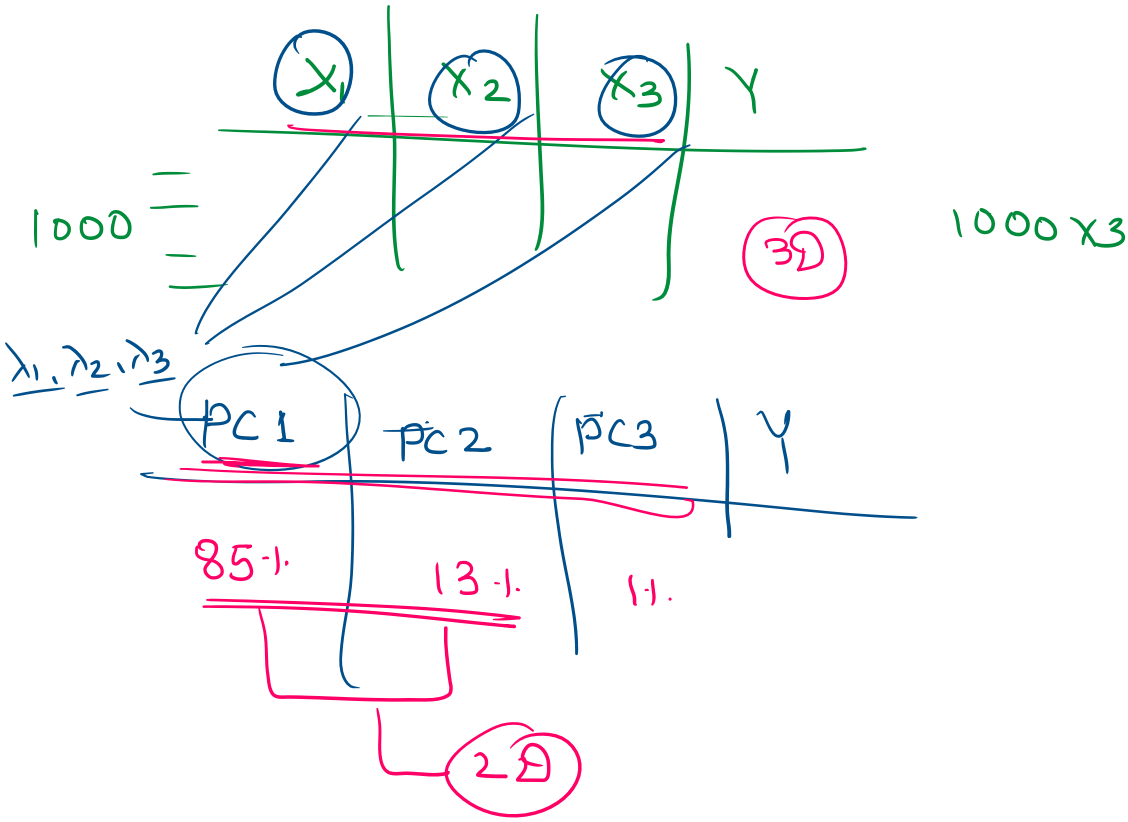
PC1 = [1000 x 1]



$\vec{p}$ $\qquad$ $\vec{u}^T$

✓ λ_2 = 2^{nd} Eigen vect → [1000 x 3] [3 x 1] = [1000 x 1]
PC2

✓ λ_3 = 3^{rd} Eigen vect = [1000 x 3] [3 x 1] = [1000 x 1]
PC3



original ⟶ PC →

50 D ⟶ 4 D

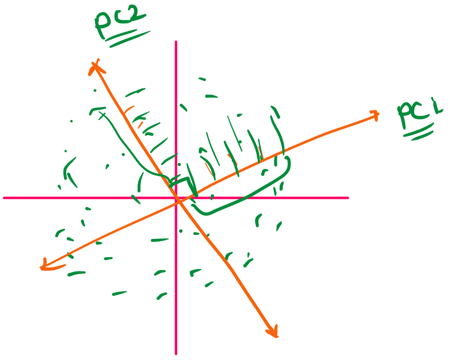
94% $\quad$ 94%

94% — 94%

Reducing dimensions of data

PCA → Fails

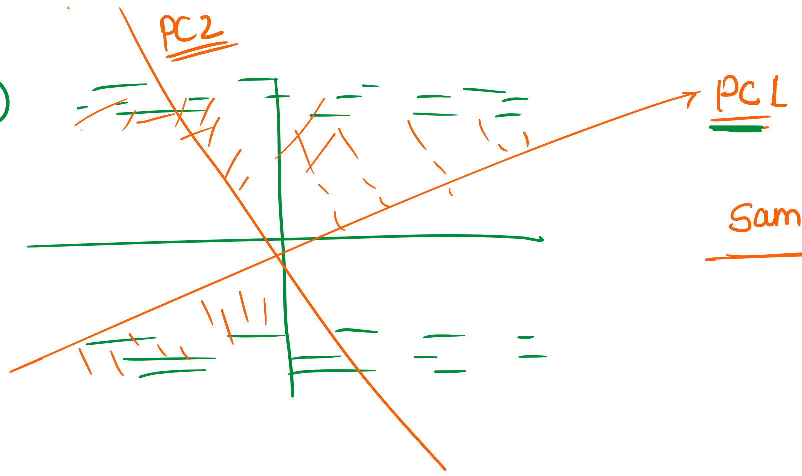
①



Varince = same

No impact of PCA

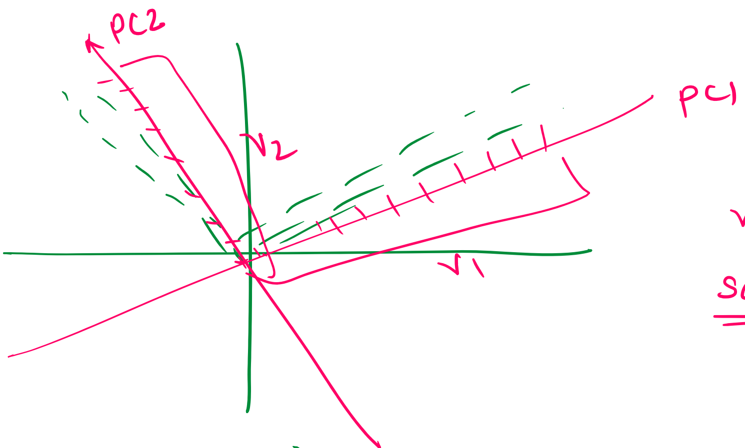
②



symmetricdata

Same variance

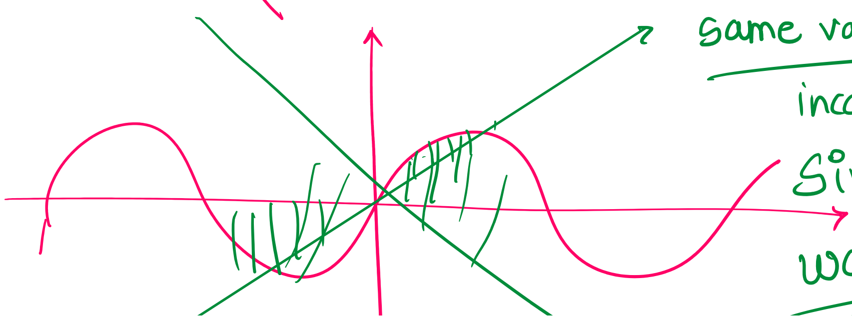
③



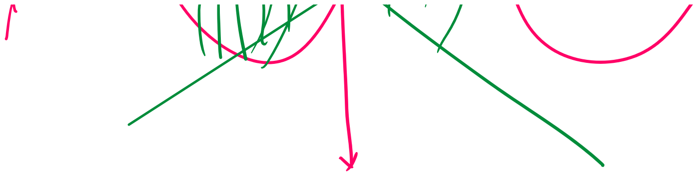
$V_1 = V_2$

Same variance

④

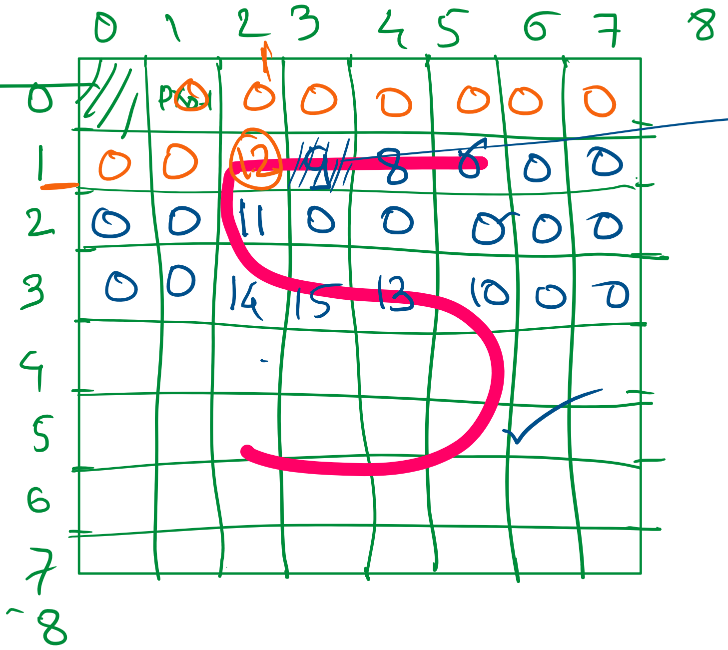


Same variance incase Sine wave



wave

Pix0_0



Density → 0 – 16

8 × 8 = 64