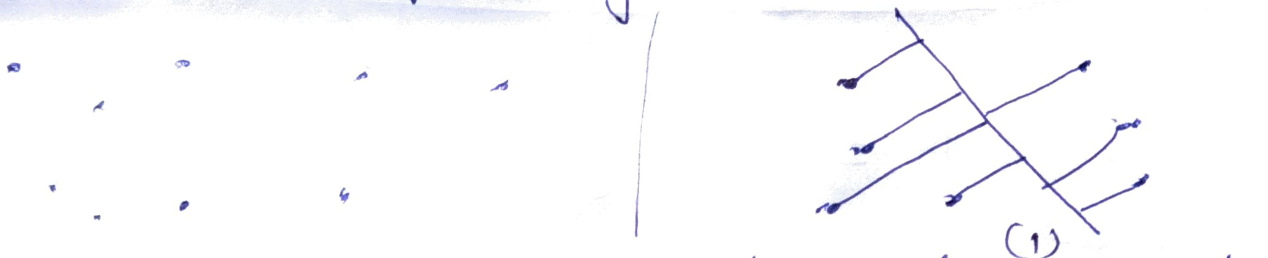
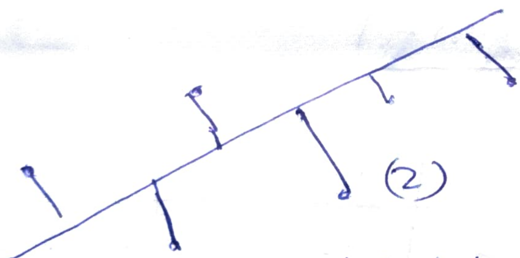
Feature Extraction:→

In Feature Extraction we have (higher) ①
n dimensional feature we want to map it to a lower dimentional space.
Suppose there are two features x_1 and x_2 and in that space you we have these instances



Suppose we want to map 2 dimensional feature space in to a one dimentional feature space

this is possible axis.

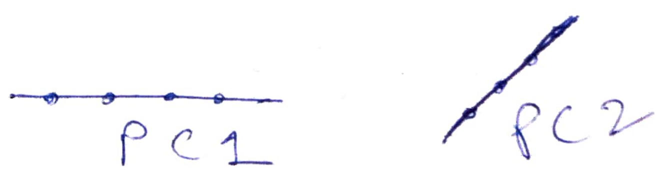


out of these two possible projections. Which one would you prefer?

Here we notice that in 2 there is a larger variation; this is a larger variance among the features so we would prefer (1) axis.

Principle Component Analysis→

To reduce the overfitting Problem we use Principle Component Analysis. [To reduce Extra attributes]



No of Principal Component Can be less than or Equal to Number of Attributes given in a data for building a Model.

$$\boxed{PC \leq \text{Attributes}}$$

for Example

$$A^T = \begin{bmatrix} 2 & 1 & 0 & -1 \\ 4 & 3 & 1 & 0.5 \end{bmatrix} \begin{array}{l} \to x \\ - y \end{array}$$

Orthogonal transformation.

Covariance Matrix.

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

this is for (x, y) & (y, x)

X	Y	$x_i - \bar{x}$ (A)	$y_i - \bar{y}$ (B)	$(x_i-\bar{x})(y_i-\bar{y})$ (AB)	$(x_i-\bar{x})^2$ A^2	$(y_i-\bar{y})^2$ B^2
2	4	1.5	1.875	2.8125	2.25	3.5156
1	3	0.5	0.875	0.4375	0.25	0.7656
0	1	−0.5	−1.125	0.5625	0.25	1.2656
−1	0.5	−1.5	−1.625	2.4375	2.25	2.6406
				6.25	5	8.1874

Sum

$$\bar{x} = \frac{2+1+0-1}{4} \Rightarrow 0.5$$

$$\bar{y} = \frac{4+3+1+0.5}{4} \Rightarrow 2.125$$

$$Cov(x, x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(x_i - \bar{x})}{n-1} = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

$$\Rightarrow \frac{5}{3} \Rightarrow 1.67$$

$$\mathrm{Cov}(y,y) = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1} = \frac{8.1874}{3} \Rightarrow 2.73 \quad (3)$$

$\mathrm{cov}(x,y) = \mathrm{cov}(y,x) \Rightarrow$ both will be same

so $\underline{2.083}$

$$\underline{S} = \mathrm{Cov} = \begin{array}{c|cc} & x & y \\ x & 1.67 & 2.083 \\ y & 2.083 & 2.73 \end{array}$$

Matrix

we have to find out 2 PCA because it's two dimensional.

$$|S - \lambda I| = 0$$

$$\left| \begin{bmatrix} 1.67 & 2.083 \\ 2.083 & 2.73 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$$\begin{vmatrix} 1.67-\lambda & 2.083 \\ 2.083 & 2.73-\lambda \end{vmatrix} = 0$$

$$(1.67-\lambda)(2.73-\lambda) - (2.083)^2 = 0$$

$$\lambda^2 - 4.4\lambda + 0.2202 = 0$$

$$\boxed{\begin{aligned} \lambda_1 &= 4.3494 \\ \lambda_2 &= 0.0506 \end{aligned}}$$

$$Z_1 = a_{11} x_1 + a_{12} x_2$$

$$Z_2 = a_{21} x_1 + a_{22} x_2$$

for $\lambda_1 = 4.3494$

$$\begin{bmatrix} 1.67 - 4.3494 & 2.083 \\ 2.083 & 2.73 - 4.3494 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = 0$$

$$\begin{bmatrix} -2.6794 & 2.083 \\ 2.083 & -1.6194 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = 0$$

$\Rightarrow -2.6794 a_{11} + 2.083 a_{12} = 0$ — (1)

$\Rightarrow 2.083 a_{11} - 1.6194 a_{12} = 0$ — (2)

$\boxed{a_{11}^2 + a_{12}^2 = 1}$ — (3)

from Equation 1

$$a_{12} = \frac{2.6794}{2.083} \times a_{11}$$

$a_{12} \Rightarrow 1.2867 \times a_{11}$ — (4)

put the value of a_{12} in Equation (3)

(4) → (3) $\boxed{a_{11} = 0.61}$

$\boxed{a_{12} = 0.79}$

for $\lambda_2 = 0.0506$

$$\begin{bmatrix} 1.67 - 0.0506 & 2.083 \\ 2.083 & 2.73 - 0.0506 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = 0$$

$\boxed{a_{21}^2 + a_{22}^2 = 1}$

$\boxed{a_{21} = 0.79}$ $\boxed{a_{22} = 0.61}$

The Principal Component are

$$z_1 = a_{11} x_1 + a_{12} x_2$$

$$\boxed{z_1 = 0.61 x_1 + 0.79 x_2}$$

$$z_2 = a_{21} x_2 + a_{22} x_2$$

$$\boxed{z_2 \Rightarrow 0.79 x_2 + 0.61 x_2}$$