

Consider dataset,

Feature	Sample 1	Sample 2	Sample 3	Sample 4
a	4	8	13	7
b	11	4	5	14

Step 1:

No. of features, $n = 2$ (a, b)

No. of samples, $N = 4$ (sample 1, sample 2, sample 3, sample 4)

Step 2:

calculating mean,

$$\bar{a} = \frac{4+8+13+7}{4} = 8$$

$$\bar{b} = \frac{11+4+5+14}{4} = 8.5$$

Step 3:

calculating covariance matrix, between features,

In the given dataset, ordered features are as,

(a, a), (a, b), (b, a), (b, b)

$$\begin{aligned}\text{cov}(a, a) &= \frac{1}{N-1} \sum_{k=1}^N (a_i - \bar{a})(a_i - \bar{a}) \\&= \frac{1}{N-1} \sum_{k=1}^N (a_i - \bar{a})^2 \rightarrow \text{for same feature} \\&= \frac{1}{4-1} [(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2] \\&= \frac{4^2 + 0 + 5^2 + 1^2}{3} = \frac{16 + 0 + 25 + 1}{3} \\&= \frac{42}{3} = \underline{\underline{14}}.\end{aligned}$$

$$\begin{aligned}\text{cov}(a, b) &= \frac{1}{N-1} \sum_{k=1}^N (a_i - \bar{a})(b_i - \bar{b}) \\&= \frac{1}{4-1} [(4-8)(11-8.5) + (8-8)(4-8.5) + \\&\quad (13-8)(5-8.5) + (7-8)(14-8.5)] \\&= \frac{1}{3} [(-4)(2.5) + (0) + (5)(-3.5) + (-1)(5.5)] \\&= \frac{1}{3} [-10 - 17.5 - 5.5] = \frac{-33}{3} = \underline{\underline{-11}}\end{aligned}$$

$$\text{cov}(b, a) = \frac{1}{N-1} \sum_{k=1}^N (b_i^p - \bar{b})(a_i^p - \bar{a})$$

$$= \text{cov}(a, b)$$

$$= -11$$

$$\text{cov}(b, b) = \frac{1}{N-1} \sum_{k=1}^N (b_i^p - \bar{b})(b_i^p - \bar{b})$$

$$= \frac{1}{N-1} \sum_{k=1}^N (b_i^p - \bar{b})^2$$

$$= \frac{1}{4-1} [(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2]$$

$$= \frac{1}{3} [(2.5)^2 + (-4.5)^2 + (-3.5)^2 + (5.5)^2]$$

$$= \frac{1}{3} [6.25 + 20.25 + 12.25 + 30.25]$$

$$= \frac{69}{3} = \underline{\underline{23}}$$

Hence covariance matrix can be

$$S = \begin{bmatrix} \text{cov}(a, a) & \text{cov}(a, b) \\ \text{cov}(b, a) & \text{cov}(b, b) \end{bmatrix} = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Step 4:-

calculate Eigen value, Eigen vector, Normalized Eigen vector.

In order calculate Eigen value,

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

$$I (\text{Identity matrix}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\lambda I = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

$$\det \left(\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = 0$$

$$\det \begin{pmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{pmatrix} = 0.$$

$$(14-\lambda)(23-\lambda) - (-11 \times -11) = 0$$

$$322 - 14\lambda - 23\lambda + \lambda^2 - 121 = 0$$

$$201 - 37\lambda + \lambda^2 = 0.$$

After rearranging,

$$\lambda^2 - 37\lambda + 201 = 0.$$

' λ ' can be calculated by quadratic eqn,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \begin{array}{l} a=1 \\ b=-37 \\ c=201 \end{array}$$

$$= \frac{-(-37) \pm \sqrt{(-37)^2 - 4(1)(201)}}{2(1)}$$

$$= \frac{37 \pm \sqrt{1369 - 804}}{2} = \frac{37 \pm \sqrt{565}}{2}$$

$$= \frac{37 \pm 23.76}{2} \Rightarrow \frac{37+23.76}{2}, \frac{37-23.76}{2}$$

$$= \frac{60.76}{2}, \frac{13.24}{2}$$

Eigen values.

$$\boxed{\lambda_1 = 30.38, \lambda_2 = 6.62}$$

So, while arranging in descending order,

$$\lambda_1 > \lambda_2 > \dots$$

Hence, $\lambda_1 = 30.38$

$$\lambda_2 = 6.62$$

We are going to find out Eigenvectors for Eigen value, $\lambda = 30.38$.

$$(S - \lambda_1 I) \vec{u}_1 = 0$$

Covariance matrix
30.38
Identity matrix
Eigenvectors of λ_1

$$\left(\begin{pmatrix} 14 & -11 \\ -11 & 23 \end{pmatrix} - 30.38 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \vec{u}_1 = 0$$

$$\text{Assume } \vec{u}_1 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Hence,

$$\begin{pmatrix} 14 & -11 \\ -11 & 23 \end{pmatrix} - \begin{pmatrix} 30.38 & 0 \\ 0 & 30.38 \end{pmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{pmatrix} 14 - 30.38 & -11 \\ -11 & 23 - 30.38 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -16.38 & -11 \\ -11 & -7.38 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-16.38 v_1 - 11 v_2 = 0 \rightarrow (1)$$

$$-11 v_1 - 7.38 v_2 = 0 \rightarrow (2)$$

So, from this if we need to calculate v_1, v_2

$$(1) \times 7.38 \Rightarrow 120.88 v_1 - 81.18 v_2 = 0$$

$$(2) \times -11 \Rightarrow +121 v_1 + 81.18 v_2 = 0$$

$$\hline 0.12 v_1 = 0$$

$$\boxed{v_1 = 0}$$

then, apply v_1 in (1), then

$$-16.38 \times 0 - 11 v_2 = 0$$

$$\boxed{v_2 = 0}$$

this can't be possible, hence

$$\begin{bmatrix} (14 - \lambda_1) & (-11) \\ (-11) & (23 - \lambda_1) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(14 - \lambda_1) v_1 - 11 v_2 = 0 \rightarrow (a)$$

$$-11 v_1 + (23 - \lambda_1) v_2 = 0 \rightarrow (b)$$

from (a),

$$(14 - \lambda_1) v_1 - 11 v_2 = 0$$

$$(14 - \lambda_1) v_1 = 11 v_2$$

$$\frac{v_1}{11} = \frac{v_2}{14 - \lambda_1} = A \text{ (Assigning)}$$

Assume $A=1$,

$$\frac{v_1}{11} = \frac{v_2}{14 - \lambda_1} = A = 1$$

$$\text{Hence, } \frac{v_1}{11} = 1 \Rightarrow v_1 = 11$$

$$\begin{aligned} \frac{v_2}{14 - \lambda_1} &= 1 \Rightarrow v_2 = 14 - \lambda_1 \\ &= 14 - 30.38 \\ &= -16.38 \end{aligned}$$

Hence Eigenvectors
for $\lambda_1 \Rightarrow$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 11 \\ -16.38 \end{bmatrix}$$

Then if we want to normalise the eigen vectors,

$$n_1 = \begin{bmatrix} 11 / \sqrt{11^2 + 16.38^2} \\ -16.38 / \sqrt{11^2 + 16.38^2} \end{bmatrix} \quad \text{dividing by} \\ \text{the long m.}$$
$$= \begin{bmatrix} \frac{11}{19.73} \\ -16.38/19.73 \end{bmatrix} = \begin{bmatrix} 0.5575 \\ -0.8302 \end{bmatrix}$$

Now, calculate eigen vectors for $\lambda_2 = 6.62$

$$(S - \lambda_2 I) v_2 = 0$$

$$\begin{bmatrix} (14 - \lambda_2) & -11 \\ -11 & (23 - \lambda_2) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(14 - \lambda_2) v_1 - 11 v_2 = 0 \rightarrow (c)$$

$$-11 v_1 - (23 - \lambda_2) v_2 = 0 \rightarrow (d)$$

from (c),

$$(14 - \lambda_2) v_1 - 11 v_2 = 0$$

$$(14 - \lambda_2) v_1 = 11 v_2$$

$$\frac{v_1}{11} = \frac{v_2}{14 - \lambda_2} = B \text{ (Assume)}$$

Assume $B = 1$,

$$\frac{v_1}{11} = \frac{v_2}{14 - \lambda_2} = B = 1$$

$$\text{Hence, } \frac{v_1}{11} = 1 \Rightarrow v_1 = 11$$

$$\frac{v_2}{14 - \lambda_2} = 1 \Rightarrow v_2 = 14 - \lambda_2$$
$$= 14 - 6.62$$
$$= 7.38$$

$$\text{Hence, Eigen Vectors for } \lambda_2 = \begin{bmatrix} 11 \\ 7.38 \end{bmatrix}$$

If we want to normalise eigen vectors,

$$n_2 = \begin{bmatrix} 11 / \sqrt{11^2 + 7.38^2} \\ 7.38 / \sqrt{11^2 + 7.38^2} \end{bmatrix} = \begin{bmatrix} 11 / 13.24 \\ 7.38 / 13.24 \end{bmatrix}$$
$$= \begin{bmatrix} 0.8308 \\ 0.5574 \end{bmatrix}$$

Step 5: New dataset,

Feature	sample1	sample2	sample3	sample4
a	4	8	13	7
b	11	4	5	14

1st PC	P ₁₁	P ₁₂	P ₁₃	P ₁₄
	sample1	sample2	sample3	sample4

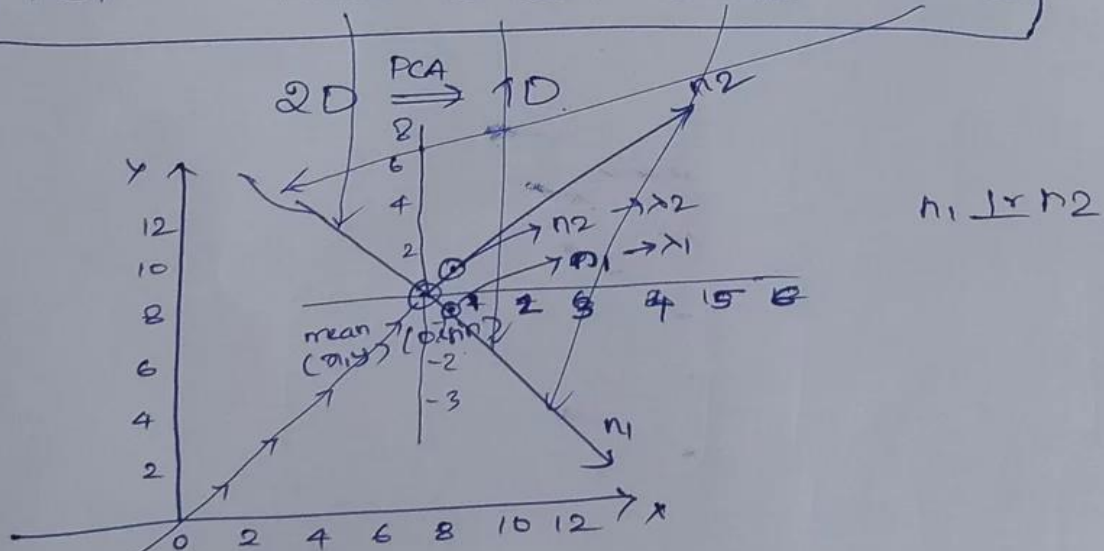
$$\begin{aligned}
 P_{11} &= \eta_1^T \begin{bmatrix} 4-8 \\ 11-8.5 \end{bmatrix} \\
 &= \begin{bmatrix} 0.5575 & -0.8302 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix} \\
 &= (-2.23 - 2.0755) \\
 &= -4.305\%
 \end{aligned}$$

$$\begin{aligned}
 P_{12} &= \eta_1^T \begin{bmatrix} 8-8 \\ 4-8.5 \end{bmatrix} \Rightarrow (0.5575 - 0.8302) \begin{pmatrix} 0 \\ -4.5 \end{pmatrix} \\
 &= 0 + 3.7359 = 3.7359\%
 \end{aligned}$$

$$\begin{aligned}
 P_{13} &= \eta_1^T \begin{bmatrix} 13-8 \\ 5-8.5 \end{bmatrix} \Rightarrow (0.5575 - 0.8302) \begin{pmatrix} 5 \\ -3.5 \end{pmatrix} \\
 &= 2.787 + 2.905 = \underline{5.692}
 \end{aligned}$$

$$\begin{aligned}
 P_{14} &= \eta_1^T \begin{bmatrix} 7-8 \\ 14-8.5 \end{bmatrix} = (0.5575 - 0.8302) \begin{pmatrix} -1 \\ 5.5 \end{pmatrix} \\
 &= -0.5575 - 4.5661 \\
 &= \underline{-5.123\%}
 \end{aligned}$$

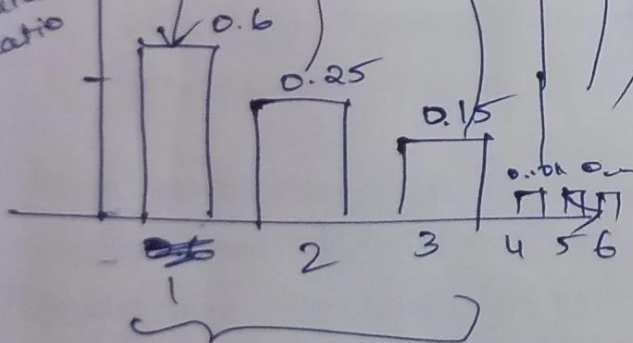
PC1	-4.305	3.7359	5.692	-5.123
-----	--------	--------	-------	--------



	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						

⇒ score plot

Explained variance ratio



$$0.6 + 0.25 + 0.15 \Rightarrow 0.90$$

Hence these PC's are OK here.

In general 80% criteria is feasible to achieve
 ↓ this is not sure

⇒ 90% of info / variance achieved

PCA scores is essential to calculate classification / regression