• Problems on PCA:

Q.1 Apply PCA on following matrix for dimensionality reduction

$$A^T = \begin{bmatrix} 2 & 1 & 0 & -1 \\ 4 & 3 & 1 & 0.5 \end{bmatrix}$$

⇒ STEP 1: Find mean of each feature

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

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

STEP 2: Find $x_i - \bar{x}$ & $y_i - \bar{y}$ i.e. deviation from mean

$A = x_i - \bar{x}$	$B = y_i - \bar{y}$	AB	A^2	B^2
1.5	1.875	2.8125	2.25	3.5156
0.5	0.875	0.4375	0.25	0.7656
-0.5	-1.125	0.5625	0.25	1.2656
-1.5	-1.625	2.4375	2.25	2.6401
		6.25	5	8.1874

STEP 3: Calculate co-variance matrix using the formula :-

$$\text{cov}[x, y] = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$\text{cov} = \begin{bmatrix} \text{cov}[x,x] & \text{cov}[x,y] \\ \text{cov}[y,x] & \text{cov}[y,y] \end{bmatrix}$$

$$\therefore \text{cov}[x,x] = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})}{n-1} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{4-1} = \frac{A^2}{3}$$

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

$$\text{cov}[x,y] = \text{cov}[y,x] = \frac{AB}{3} = \frac{6.25}{3} = 2.083$$

$$\text{cov}[y,y] = \frac{B^2}{3} = 2.7289$$

$$\therefore \text{COV} = \begin{bmatrix} 1.67 & 2.083 \\ 2.083 & 2.73 \end{bmatrix}$$

STEP 4: Calculate Eigen values & Eigen vectors of the co-variance matrix.

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

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

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

$$4.56 - 1.67\lambda - 2.73\lambda + \lambda^2 - 4.34 = 0$$

$$\lambda^2 - 4.4\lambda + 0.22 = 0$$

$$\therefore \lambda_1 = 4.3494$$

$$\lambda_2 = 0.0506$$

To find eigen vectors use:-

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

$$\begin{bmatrix} -2.6794 & 2.083 \\ 2.083 & -1.619 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = 0$$

$$\therefore -2.6794\,x_{11} + 2.083\,x_{12} = 0$$

$$2.083\,x_{11} - 1.619\,x_{12} = 0$$

$$\therefore x_{11} = 0 \qquad x_{12} = 0$$

For orthogonal transformation:

$$x_{11}^2 + x_{12}^2 = \cancel{0}\ 1$$

$x_{11} = 0.61$ and $x_{12} = 0.79$

Similarly, $x_{21} = 0.79$ and $x_{22} = 0.61$

STEP 5: Find out principle component using following formula for j^{th} principle component

$$\frac{\lambda_j}{trace(S)} \; ; \; Trace(S) = \Sigma \lambda_j$$

$$\Sigma \lambda_j = 4.3494 + 0.0506 = 4.4$$

For λ_1,

$$\text{principle component} = \frac{4.3494}{4.4} = 0.9885$$

For λ_2,

$$\text{principle component} = \frac{0.0506}{4.4} = 0.0115$$

2. From the following obs. of 10 years of rainfall and run-off

$$\cancel{\begin{matrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 5 & 5 & 3 \end{matrix}} \qquad \cancel{\begin{matrix} 105 & 94 & 120 \\ 115 & 95 & 121 \\ 103 & 104 & \end{matrix}}$$

$$A = \begin{bmatrix} 105 & 42 \\ 115 & 46 \\ 103 & 26 \\ 94 & 39 \\ 95 & 29 \\ 104 & 33 \\ 120 & 48 \\ 121 & 58 \\ 127 & 45 \\ 79 & 20 \end{bmatrix}$$

STEP 1: $\bar{x} = \frac{105+115+103+94+95+104+120+121+127\ldots}{10}$

$\therefore \bar{x} = 106.3$

Similarly, $\bar{y} = 38.6$

STEP 2:

$A = x_i - \bar{x}$	$B = y_i - \bar{y}$	AB	A^2	B^2
-1.3	3.4	-4.42	1.69	11.56
8.7	7.4	64.38	75.69	54.76
-3.3	-12.6	41.58	10.89	158.76
-12.3	0.4	-4.92	151.29	0.16
-11.3	-9.6	108.48	127.69	92.16
-2.3	-5.6	12.88	5.29	31.36
13.7	9.4	128.78	187.69	88.36
14.7	19.4	285.18	216.09	376.36
20.7	6.4	132.48	428.49	40.96
-27.3	-18.6	507.78	745.29	345.96
		1272.2	1950.1	1200.4

STEP 3:

$$cov[x, y] = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$COV = \begin{bmatrix} cov[x,x] & cov[x,y] \\ cov[y,x] & cov[y,y] \end{bmatrix}$$

$$cov[x,x] = \frac{A^2}{n-1} = \frac{1950.1}{9} = 216.67$$

$$cov[x,y] = cov[y,x] = \frac{AB}{9} = \frac{1272.2}{9} = 141.35$$

$$cov[y,y] = \frac{B^2}{9} = \frac{1200.4}{9} = 133.37$$

$\therefore$

$$COV = \begin{bmatrix} 216.67 & 141.35 \\ 141.35 & 133.37 \end{bmatrix}$$

STEP 4:

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

$$\therefore \begin{vmatrix} 216.67-\lambda & 141.35 \\ 141.35 & 133.37-\lambda \end{vmatrix} = 0$$

$$\therefore (216.67-\lambda)(133.37-\lambda) - 19979.82 = 0$$

$$28897.27 - 350.04\lambda + \lambda^2 - 19979.82 = 0$$

$$\therefore \lambda^2 - 350.04 + 8917.45 = 0$$

$$\lambda_1 = 322.37$$

$$\lambda_2 = 27.6614$$

Finding Eigen Vectors:-

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

$\therefore$ for $\lambda_1 = 322.37$

$$\begin{bmatrix} -105.7 & 141.35 \\ 141.35 & -189 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = 0$$

For orthogonal transformation:

$$x_{11}^2 + x_{12}^2 = 1$$

Here,

$$-105.7x_{11} + 141.35x_{12} = 0$$

$$\therefore 141.35x_{12} = 105.7x_{11}$$

$$\therefore 19979.82x_{12}^2 = 11172.49x_{11}^2$$

$$19979.82(1 - x_{11}^2) = 11172.49x_{11}^2$$

$$\therefore 19979.82 - 19979.82x_{11}^2 = 11172.49x_{11}^2$$

$$19979.82 = 31152.31 x_{11}^2$$

$$\therefore x_{11} = 0.8$$

$$\therefore x_{12} = 0.59$$

Similarly, for $\lambda_2 = 27.66$

$$\begin{bmatrix} 189.01 & 141.35 \\ 141.35 & 105.71 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = 0$$

$$189.01 x_{21} = -141.35 x_{22}$$

$$\therefore 35724.78 (1 - x_{22}^2) = 19979.82 x_{22}^2$$

$$35724.78 = 55704.6 x_{22}^2$$

$$\therefore x_{22} = 0.8$$

$$\therefore x_{21} = -0.59$$

STEP 5: For j^{th} principle component,

$$\frac{\lambda_j}{trace(S)};$$

$$Trace(S) = \Sigma \lambda_j = 322.37 + 27.66 = 350.03$$

For λ_1,

$$\text{principle component} = \frac{322.37}{350.03} = 0.9209$$

For λ_2,

$$\text{principle component} = \frac{27.66}{350.03} = 0.079$$

Problems on SVD :

Compute SVD for following matrix

$$A = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$$

STEP 1: Compute A^T and find A^TA

$$\therefore A^T = \begin{bmatrix} 4 & 3 \\ 0 & -5 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 25 & -15 \\ -15 & 25 \end{bmatrix}$$

STEP 2: Determine eigen values of A^TA matrix and sort them in descending order.

Square root of eigen values will be used to find singular values of A.

$$\therefore |A^TA - \lambda I| = 0$$

$$\begin{bmatrix} 25-\lambda & -15 \\ -15 & 25-\lambda \end{bmatrix} = 0$$

$$\therefore (25-\lambda)^2 - 225 = 0$$

$$\therefore \lambda^2 - 50\lambda + 400 = 0$$

$$\therefore \lambda_1 = 40$$

$$\lambda_2 = 10$$

Singular values, say s_1 and s_2 can be calculated by taking the square root of eigen values.

$$\therefore s_1 = \sqrt{40} = 6.3245$$

$$s_2 = \sqrt{10} = 3.1622$$

Sigma matrix, $\Sigma = \begin{bmatrix} 6.3245 & 0 \\ 0 & 3.1622 \end{bmatrix}$

STEP 3: ~~Find~~ Compute the inverse of sigma matrix, Σ^{-1}
From constructed diagonal matrix in which the values are placed in a descending order along the diagonal.

$$\therefore \Sigma^{-1} = \begin{bmatrix} 0.1581 & 0 \\ 0 & 0.3162 \end{bmatrix}$$

STEP 4: Use the ordered eigen values from step 2 and compute eigen vectors of A^TA.
Place these eigen vectors along the columns of V and compute its transpose.
For $\lambda_1 = 40$

$$A^TA - \lambda I = \begin{bmatrix} 25-40 & -15 \\ -15 & 25-40 \end{bmatrix}$$

$$= \begin{bmatrix} -15 & -15 \\ -15 & -15 \end{bmatrix}$$

$$(A^TA - \lambda_1 I) X_1 = 0$$

$$\therefore \begin{bmatrix} -15 & -15 \\ -15 & -15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$x_2 = -x_1$$

We have computed value of x_2 from above equations and dividing these values by their length where, $L = \sqrt{x_1^2 + x_2^2}$

$$\therefore L = \sqrt{x_1^2 + (-x_1)^2} = x_1\sqrt{2}$$

$$X_1 = \begin{bmatrix} \frac{x_1}{L} \\ \frac{x_1}{L} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.7072 \\ -0.7072 \end{bmatrix}$$

Similarly, for $\lambda_2 = 10$

$$X_2 = \begin{bmatrix} 0.7071 \\ 0.7071 \end{bmatrix}$$

$$V = [X_1 \quad X_2]$$

$$V = \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$

$$V^T = \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$$

STEP 5: Compute U as, $U = AV\Sigma^{-1}$

$$U = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix} \begin{bmatrix} 0.1581 & 0 \\ 0 & 0.3162 \end{bmatrix}$$

$$U = \begin{bmatrix} 0.4471 & 0.8943 \\ 0.8943 & -0.4471 \end{bmatrix}$$

~~Co~~

Compute the full SVD using $A = U\Sigma V^T$

$$A = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$$

The orthogonal nature of U and V matrices ~~b~~ is evident by inspecting their eigen vectors. This can be demonstrated by computing dot products between column vectors. All dot products are equal to 0.

Alternatively, we can plot this and see, they are orthogonal.