Q. Compute the the linear Discriminant projection for the following two dimensional dataset.

Samples for class ω_1,

$X_1 = (x_1, x_2) = \{(4,2), (2,4), (2,3), (3,6), (4,4)\}$

Samples for class ω_2,

$X_2 = (x_1, x_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$

Our task is to find a new axis to project the data values.

→ Find class means,

$$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x = \frac{1}{5} \begin{bmatrix} 15 \\ 19 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.8 \end{bmatrix}$$

$$\mu_2 = \frac{1}{N_2} \sum_{x \in \omega_2} x = \frac{1}{5} \begin{bmatrix} 42 \\ 38 \end{bmatrix} = \begin{bmatrix} 8.4 \\ 7.6 \end{bmatrix}$$

Covariance matrix of the first class:

$$\Sigma_1 = \sum_{x \in \omega_1} \frac{(x_1 - \mu_1)(x_1 - \mu_1)^T}{N-1}$$

$$= \frac{1}{4}\left(\begin{bmatrix} 1 \\ -1.8 \end{bmatrix} \begin{bmatrix} 1 & -1.8 \end{bmatrix} + \begin{bmatrix} -1 \\ 0.2 \end{bmatrix} \begin{bmatrix} -1 & 0.2 \end{bmatrix} + \begin{bmatrix} -1 \\ -0.8 \end{bmatrix} \begin{bmatrix} -1 & -0.8 \end{bmatrix} + \begin{bmatrix} 0 \\ 2.2 \end{bmatrix} \begin{bmatrix} 0 & 2.2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0.2 \end{bmatrix} \begin{bmatrix} 1 & 0.2 \end{bmatrix} \right)$$

$$= \frac{1}{4}\left(\begin{bmatrix} .1 & -1.8 \\ -1.8 & 3.24 \end{bmatrix} + \begin{bmatrix} 1 & -0.2 \\ -0.2 & 0.04 \end{bmatrix} + \begin{bmatrix} 1 & 0.8 \\ 0.8 & 0.64 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 4.84 \end{bmatrix} + \begin{bmatrix} 1 & 0.2 \\ 0.2 & 0.04 \end{bmatrix}\right)$$

$$\Rightarrow \frac{1}{4}\begin{bmatrix} 4 & -1 \\ -1 & 8.8 \end{bmatrix} = \begin{bmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{bmatrix}$$

$$\therefore \Sigma_1 = \begin{bmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{bmatrix}$$

Covariance matrix of the second class,

$$\Sigma_2 = \sum_{x \in \omega_2} \frac{(x_2 - \mu_2)(x_2 - \mu_2)^T}{N-1}$$

$$= \frac{1}{4}\left(\begin{bmatrix} 0.6 \\ 2.4 \end{bmatrix}\begin{bmatrix} 0.6 & 2.4 \end{bmatrix} + \begin{bmatrix} -2.4 \\ 0.4 \end{bmatrix}\begin{bmatrix} -2.4 & 0.4 \end{bmatrix} + \begin{bmatrix} 0.6 \\ -2.6 \end{bmatrix}\begin{bmatrix} 0.6 & -2.6 \end{bmatrix} + \begin{bmatrix} -0.4 \\ -0.6 \end{bmatrix}\begin{bmatrix} -0.4 & -0.6 \end{bmatrix} + \begin{bmatrix} 1.6 \\ 0.4 \end{bmatrix}\begin{bmatrix} 1.6 & 0.4 \end{bmatrix}\right)$$

$$= \frac{1}{4}\left(\begin{bmatrix} 0.36 & 1.44 \\ 1.44 & 5.76 \end{bmatrix} + \begin{bmatrix} 5.76 & -0.96 \\ -0.96 & 0.16 \end{bmatrix}\right.$$

$$+ \begin{bmatrix} 0.36 & -1.56 \\ -1.56 & 6.76 \end{bmatrix} + \begin{bmatrix} 0.16 & 0.24 \\ 0.24 & 0.36 \end{bmatrix}$$

$$\left. + \begin{bmatrix} 2.56 & 0.64 \\ 0.64 & 0.16 \end{bmatrix}\right)$$

$$S_2 = \begin{bmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{bmatrix}$$

$$S_W = S_1 + S_2 = \begin{bmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{bmatrix}$$

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T = \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix}(-5.4 \quad -3.8)$$

$$= \begin{bmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{bmatrix}$$

$$S_W^{-1} S_B W = \lambda$$

$$\begin{bmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{bmatrix}\begin{bmatrix} -5.4 \\ 3.8 \end{bmatrix} = \begin{bmatrix} 1.5813 \\ -0.6 \end{bmatrix}$$

Q. Find LD for given dataset

$C_1 = \{(4,2), (2,4), (2,3)\}$

$C_2 = \{(3,6), (4,4), (5,5)\}$

$$\mu_1 = \frac{1}{3}\left[\begin{bmatrix}4\\2\end{bmatrix} + \begin{bmatrix}2\\4\end{bmatrix} + \begin{bmatrix}2\\3\end{bmatrix}\right] = \begin{bmatrix}2.67\\3\end{bmatrix}$$

$$\mu_2 = \frac{1}{3}\left[\begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}4\\4\end{pmatrix} + \begin{pmatrix}5\\5\end{pmatrix}\right] = \begin{bmatrix}4\\5\end{bmatrix}$$

~~$S_1 = \frac{1}{N_1}(x - \mu) = \frac{1}{2}[$~~

$$S_1 = \frac{(x_1 - \mu_1)(x_1 - \mu_1)^T}{N-1}$$

$$\Rightarrow \frac{1}{2}\left[\begin{bmatrix}1.33\\-1\end{bmatrix}\begin{bmatrix}1.33 & -1\end{bmatrix} + \begin{bmatrix}-0.67\\1\end{bmatrix}\begin{bmatrix}-0.67 & 1\end{bmatrix}\right.$$

$$\left. + \begin{bmatrix}-0.67\\0\end{bmatrix}\begin{bmatrix}-0.67 & 0\end{bmatrix}\right]$$

$$\Rightarrow \frac{1}{2}\left[\begin{pmatrix}1.7689 & -1.33\\-1.33 & 1\end{pmatrix} + \begin{pmatrix}0.4489 & -0.67\\-0.67 & 1\end{pmatrix}\right.$$

$$\left. + \begin{pmatrix}0.4489 & 0\\0 & 0\end{pmatrix}\right]$$

$$= \frac{1}{2}\begin{bmatrix} 2.66 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 1.33 & -1 \\ -1 & 1 \end{bmatrix}$$

$$S_2 = \frac{(x_2 - \mu_2)(x_2 - \mu_2)^T}{N-1}$$

$$\Rightarrow \frac{1}{2}\left[\begin{pmatrix} -1 \\ 1 \end{pmatrix}(-1 \quad 1) + \begin{pmatrix} 0 \\ -1 \end{pmatrix}(0 \quad -1) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1 \quad 0)\right]$$

$$\Rightarrow \frac{1}{2}\left[\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]$$

$$\Rightarrow \frac{1}{2}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$$

$$S_W = S_1 + S_2 = \begin{bmatrix} 2.33 & -1.5 \\ -1.5 & 2 \end{bmatrix}$$

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T = \begin{bmatrix} -2.67 \\ -2 \end{bmatrix}\begin{bmatrix} -2.67 & -2 \end{bmatrix}$$

__/__/____

Feature Selection

It can be supervised or unsupervised

supervised

1. Filter methods
   - chi square test
   - correlation coefficient
   - mutual
2. Wrapper method
   - Forward selection
   - Backward elimination
   - Recursive method
3. Embedded method
   - Lasso Lasso Regularization

Unsupervised

1. Variance threshold
2. Dimensionality Reduction
3. Correlation Analysis

Chi-Square Test

Steps

1. Define the Null and Alternate Hypothesis.
2. Calculate the contingency table for our feature

eg:

	Subscribed (O)	Not Subscribed	Total
Low Income	20	30	50
Medium Income	40	25	65
High Income	10	15	25
Total	70	70	140

3. Calculate the expected frequencies.

$$\text{Expected Frequency} = \frac{\text{Row total} \times \text{Col total}}{\text{Total}}$$

$$E_{LI,S} = \frac{50 \times 70}{140} = 25$$

$$E_{LI,NS} = \frac{50 \times 70}{140} = 25$$

$$E_{MI,S} = \frac{65 \times 70}{140} = 32.5$$

$$E_{MI,NS} = \frac{65 \times 70}{140} = 32.5$$

$$E_{HI,S} = \frac{25 \times 70}{140} = 12.5$$

$$E_{HI,NS} = \frac{25 \times 70}{140} = 12.5$$

4. Calculate the chi-square value and compare with critical value,

$$\lambda^2 = \sum \frac{(O - E)^2}{E}$$

$$\Rightarrow \frac{(20-25)^2}{25} + \frac{(30-25)^2}{25} + \frac{(40-32.5)^2}{32.5} +$$

$$\frac{(25-32.5)^2}{32.5} + \frac{(10-12.5)^2}{12.5} + \frac{(15-12.5)^2}{12.5}$$

$$\Rightarrow 1 + 1 + 1.7307 + 1.7307$$

5. Compare the chi square value or with critical value to accept or reject.

$$\lambda^2 = \qquad \lambda_{2-0.05} =$$

If the calculated value is higher than the one obtained from the table, then the feature is relevant and we reject the null hypothesis.

__/__/__

Regularization

Lasso Regularization Regression

- It uses L1 regularization technique called LASSO regression.
- It adds the absolute value of magnitude of the coefficient as a penalty term to the loss function.
- We use Lasso if many features in the dataset are irrelevant or not strongly related to the target.

$$\text{cost} \Rightarrow \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^{m} |w_i|$$

$\lambda \Rightarrow$ Regularization parameter
(how much penalty is added for the weights)

higher $\lambda \Rightarrow$ it applies stronger regularization and shrinks more coefficients to zero.

Ridge Regression

- It uses L2 regularization technique.
- it adds the "squared magnitude" of the coefficient as a penalty
- Used when all features contribute to the

target but to varying degrees and when one want retain all features with smaller coefficients. i.e. if you have a dataset where all features have some relevance but we want to reduce the impact of noise or to multicollinearity.

$$\text{cost} \Rightarrow \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^{m} w_i^2$$

Elastic Net Regression

- Combination of L1 and L2 regularization.
- We add

Q. Find PCA for the given dataset

x_1	2.5	0.5	2.2	1.9	3.1	2.3	2.0	1.0	1.5	1.1
x_2	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

$$\bar{x}_1 = \frac{1}{10}(2.5 + 0.5 + 2.2 + 1.9 + 3.1 + 2.3 + 2.0 + 1.0 + 1.5 + 1.1)$$

$$\Rightarrow \underline{\underline{1.81}}$$

$$\bar{x}_2 = \frac{1}{10}(2.4 + 0.7 + 2.9 + 2.2 + 3.0 + 2.7 + 1.6 + 1.1 + 1.6 + 0.9)$$

$$\Rightarrow \underline{\underline{1.91}}$$

Finding covariance matrix,

$$S = \begin{bmatrix} \text{cov}(x_1, x_1) & \text{cov}(x_1, x_2) \\ \text{cov}(x_2, x_1) & \text{cov}(x_2, x_2) \end{bmatrix}$$

$$\text{cov}(x_1, x_1) = \frac{1}{9}\left[(0.69)^2 + (-1.31)^2 + (0.39)^2 + (0.09)^2 + (1.29)^2 + (0.49)^2 + (0.19)^2 + (-0.81)^2 + (-0.31)^2 + (-0.71)^2\right]$$

$$\Rightarrow \frac{0}{9} = \frac{0}{4}$$

$$\Rightarrow \frac{5.549}{9} = \underline{0.6165}$$

x_1 new	0.69	-1.31	0.39	0.09	1.29	0.49	0.19	-0.81	-0.31	-0.71
x_2 new	0.49	-1.21	0.99	0.29	1.09	0.79	-0.31	-0.81	-0.31	-1.01

$$\text{cov}(x_1, x_2) = \frac{1}{9}[5.539] = \underline{0.6154}$$

$$\text{cov}(x_2, x_1) = 0.6154$$

$$\text{cov}(x_2, x_2) = \frac{1}{9}[6.449] = \underline{0.7165}$$

$$S = \begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix}$$

Z score based Outlier on Anomaly Detection

Suppose we have a dataset of daily sales revenue for a retail store over the past 30 days.

[100, 150, 120, 125, 140, 130, 110, 135, 130, 150, 140, 100, 95, 80, 120, 125, 130, 100, 190, 135, 130, 145, 110, 120, 130, 135, 140, 125, 130, 120]

We want to identify any days where the sales revenue is significantly different from the other days, which may indicate an anomaly or outlier.

$$\mu = \frac{3740}{30} = 124.67$$

$$z = \frac{x - \mu}{\sigma}$$

$$\sigma = \sqrt{\frac{8346.687}{9}} = \sqrt{927.407} = 30.453$$

16.68

LOF scores

Q. Consider the 2D dataset representing points in a 2 dimensional space

A(1,2) B(2,3) C(3,4) D(10,10)

Calculate LOF scores for each point using K=2 nearest neighbours.

$AB = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} = 1.414$

$AC = \sqrt{(-2)^2 + (-2)^2} = \sqrt{8} = 2.828$

$AD = \sqrt{(-9)^2 + (-8)^2} = \sqrt{145} = 12.041$

$BA = 1.414$

$BC = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} = 1.414$

$BD = \sqrt{(-8)^2 + (-7)^2} = \sqrt{113} = 10.630$

$CA = 2.828$

$CB = 1.414$

$CD = \sqrt{(-7)^2 + (-6)^2} = \sqrt{85} = 9.2195$

Calculate reachability distance (before that find K=2 nearest neighbours for each point)

for A,

B ; AB = 1.414

C ; AC = 2.828

for B,

A, AB = 1.414

C, BC = 1.414

for C,

A, AC = 2.828

B, BC = 1.414

for D,

B, BD = 10.630

C, CD = 9.2195

Now calculate the Reachability distance,

for A,

⇒ max(AB, BA) ⇒ max(1.414, 1.414)

⇒ 1.414

(nearest neighbour dist of B)

⇒ max(AC, CB) ⇒ max(2.828, 1.414)

⇒ 2.828

for B,

⇒ max(BA, AB) = max(1.414, 1.414)

⇒ 1.414

⇒ max(BC, CB) = max(1.414, 1.414)

⇒ 1.414

for C,

⇒ max(AC, AB) = max(2.828, 1.414)

⇒ 2.828

$\Rightarrow \max(BC, BA) = \max(1.414, 1.414)$

$\Rightarrow 1.414$

for D,

$\Rightarrow \max(BCD, CB) = \max(9.2195, 1.414)$

$\rightarrow 9.2195$

by order

$\Rightarrow \max(BD, BA) = \max(10.630, 1.414)$

$\Rightarrow 10.630$

Modified z-score

usually we take ± 3.5

$$M_i \text{ score} = \frac{0.6745 \times (X_i - \text{Median})}{\text{MAD}}$$

where X_i = datapoint

$$\text{MAD} = \text{Median}(|X_i - \text{Median}|)$$

(Median Absolute Deviation)

0.6745 is a scaling factor

Q. Consider a small dataset with 7 points in 2D plane

$X = \{A(1,1), B(2,2), C(2,3), D(3,3), E(3,4), F(8,8), G(100,100)\}$

__/__/__

19602
19208

9604
9409
19013

R464

9409
9216

16928

1. Compute distance between the points.

$AB = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} = 1.414$

$AC = \sqrt{(-1)^2 + (-2)^2} = \sqrt{5} = 2.236$

$AD = \sqrt{(-2)^2 + (-2)^2} = \sqrt{8} = 2.8281$

$AE = \sqrt{(-2)^2 + (-3)^2} = \sqrt{13} = 3.605$

$AF = \sqrt{(-7)^2 + (-7)^2} = \sqrt{98} = 9.899$

$AG = \sqrt{(-99)^2 + (-99)^2} = \sqrt{19602} = 140.007$

$BA = 1.414$

$BC = \sqrt{0^2 + (-1)^2} = 1$

$BD = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} = 1.414$

$BE = \sqrt{(-1)^2 + (-2)^2} = \sqrt{5} = 2.236$

$BF = \sqrt{(-6)^2 + (-6)^2} = \sqrt{72} = 8.485$

$BG = \sqrt{(-98)^2 + (-98)^2} = \sqrt{19208} = 138.592$

$CA = 2.236$

$CB = 1$

$CD = \sqrt{(-1)^2 + 0^2} = \sqrt{1} = 1$

$CE = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} = 1.414$

$CF = \sqrt{(-6)^2 + (-5)^2} = \sqrt{61} = 7.810$

$CG = \sqrt{(-98)^2 + (-97)^2} = \sqrt{19013} = 137.88$

$DA = 2.828$

$DB = 1.414$

$DC = 1$

$DE = \sqrt{0^2 + (-1)^2} = \sqrt{1} = 1$

$DF = \sqrt{(-5)^2 + (-5)^2} = \sqrt{50} = 7.071$

$DG = \sqrt{(-97)^2 + (-97)^2} = \sqrt{18,818} = 137.178$

EA = 3.605

EB = 2.236

EC = 1.414

ED = 1

EF = $\sqrt{(-5)^2+(-4)^2}$ = 6.403

EG = $\sqrt{(-97)^2+(-96)^2}$ = $\sqrt{18625}$ = 136

GA = FA = 9.899

GB = FB = 8.485

GC = FC = 7.810

GD = FD = 7.071

GE = FE = 6.403

GF = FG = $\sqrt{(-92)^2+(-92)^2}$ = $\sqrt{16928}$ = ~~130.10~~ 130.107

GA = 140.007

GB = 138.592

GC = 137.88

GD = 137.178

GE = 136

GF = 130.107

Finding K=2 nearest neighbours & ~~reachability~~ values

for A ⇒ B, C,

AB, AC (1.414, 2.236)

~~Rch = max(AB, B~~

for B ⇒ C, A (1, 1.414)

for C ⇒ B, D [1, 1)

for D ⇒ C, E [1, 1)

for E ⇒ C, D [1.414, 1]

for F ⇒ E, D [6.403, 7.071]

for G ⇒ E, F [136, 130.107]

Finding Reachability values,

for A

⇒ max(AB, BC) = max(1.414, 1)
= 1.414

⇒ max(AC, CB) = max(2.236, 1)
= 2.236

for B

⇒ max(BC, CB) = ~~[1, 1]~~ max(1, 1)
= 1

⇒ max(BA, AB) = max(1.414, 1.414)
= 1.414

for C,

⇒ max(CB, BC) = max (1, 1) = 1

⇒ max(CD, DC) = max (1, 1) = 1

for D,

⇒ max (DC, CB) = max(1, 1) = 1

⇒ max (DE, ED) = max (1, 1) = 1

for E,

⇒ max (ED, DC) = max(1, 1) = 1

⇒ max (EC, CB) = max (1.414, 1) = 1.414

for F,

⇒ max (FE, ED) = max (6.403, 1) = 6.403

⇒ max (FD, DC) = max (7.071, 1) = 7.071

for G,

⇒ max (GE, ED) = max (136, 1) = 136

⇒ max (GF, FE) = max (130.107, 6.403)
= 130.107

Next we calculate LRD values.

$$LRD = \frac{1}{\frac{1}{k}\sum_{q \in N_k(p)} \text{reach dist}(p,q)}$$

LRD of A $= \frac{1}{\frac{1}{2}(1.414 + 2.23)} = 1.822 = \frac{1}{1.822} = 0.54$

LRD of B $= \frac{1}{\frac{1}{2}(1 + 1.414)} = \frac{1}{1.207} = 0.8285$ (maybe

LRD of C $= \frac{1}{\frac{1}{2}(1+1)} = 1$

LRD of D $= \frac{1}{\frac{1}{2}(1+1)} = 1$

LRD of E $= \frac{1}{\frac{1}{2}(1+1)} = 1$

LRD of E $= \frac{1}{\frac{1}{2}(1 + 1.414)} = \frac{1}{1.207} = 0.8285$

LRD of F $= \frac{1}{\frac{1}{2}(6.403 + 7.071)} = \frac{1}{6.737} = 0.1484$

LRD of G $= \frac{1}{\frac{1}{2}(136 + 130.107)} = \frac{1}{133.0535} = 0.00751$

(mostly outlier)

$$LOF = \frac{1}{k} \sum_{q \in N_k(p)} LRD(q)$$

LOF of A = ~~* Σ scaled~~

$$= \frac{\Sigma \text{ LRD's of A's neighbours}}{k \times \text{LRD of A}}$$

point with highest LOF value is the outlier.

Kernel Induced Feature Expansion

Different types of Kernel;

1) Linear Kernel

$$k(x, y) = x^T y$$

x and y are vector,

Therefore, $k(x, y) = \phi(x) \cdot \phi(y) = x^T y$

2) Polynomial Kernel

$$k(x, y) = (x^T y)^q$$

q is the degree of the polynomial.

3) Gaussian Kernel (Radial Basis Function)

~~k(x, y) =~~

$$k(x, y) = e^{-(x-y)^2/2\sigma^2}$$

σ is the constant

4) ~~Sigmoid Kernel~~

Consider 2 data points
$x = (1, 2)$ and $y = (2, 3)$ with $\sigma = 1$.
Apply RBF kernel and find the value of RBF kernel for these points.

$$k(x, y) = e^{-(x-y)^2/2\sigma^2}$$

$$(1-2)^2 + (2-3)^2 = (-1)^2 + (-1)^2 = 2$$

If $\sigma = 1$,

$$k((1,2),(2,3)) = e^{-\{(2)^2 / \{2(1)^2\}\}}$$

Q Consider 2 data points $x = (1, 2)$ and $y = (2, 3)$ with $c = 1$. Apply linear, homogeneous and inhomogeneous kernels:-

4) Inhomogeneous kernel

$$k(x, y) = (c + x^T y)^q$$

c is a constant and q is the degree of the polynomial.

Linear kernel ($q = 1$)

$$k(x, y) = \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}^T (2\ 3) \right)^2 = 2 + 6 = 8$$

when $q = 2$,

__/__/___

$$k(x,y) = \left(\begin{bmatrix}1\\2\end{bmatrix}^T \begin{pmatrix}2 & 3\end{pmatrix}\right)^2 = 8^2 = \underline{\underline{64}}$$

If $c=1$

$$k(x,y) = (c + x^T y)^2$$

$$= \left(1 + \begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}2 & 3\end{bmatrix}\right)^2 = (1+8)^2 = \underline{\underline{81}}$$

__/__/____

Q Find PCA for the given dataset

	1	2	3	4	5	6	7	8	9	10
x_1	2.5	0.5	2.2	1.9	3.1	2.3	2.0	1.0	1.5	1.1
x_2	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

No. of features (n) = 2
No. of samples (N) = 10

$\mu_{x_1} = 1.81$
$\mu_{x_2} = 1.91$

Covariance matrix :
Ordered pair: (x,x) (x,y) (y,x) (y,y)

$$cov(x,x) = \frac{1}{N-1} \sum (x_{ik} - x_i)(x_{jk} - x_j)$$

$$= \frac{1}{9}[(2.5-1.81)^2 + (0.5-1.81)^2 + (2.28-1.81)^2 + (1.9-81)^2 + (3.1-1.81)^2 + (2.3-1.81)^2 + (2-1.81)^2 + (1-1.81)^2 + (1.5-1.81)^2 + (1.1-1.81)^2]$$

$$\Rightarrow \frac{1}{9}[0.4761 + 1.7161 + 0.1521 + 0.0081 + 1.66414 + 0.28014 + 0.0361 + 0.6561 + 0.0961 + 0.5041]$$

$$\Rightarrow \frac{1}{9}(5.549) = 0.6165 = 0.62$$

$$cov(x,y) = \frac{1}{N-1} \sum (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$$

$$= \frac{1}{9}[0.34 + 1.6 + 0.386 + 0.03 + 1.41 + 0.387 - 0.059 + 0.656 + 0.096 + 0.714]$$

$$\Rightarrow \frac{1}{9}(5.563) = 0.618 = 0.62$$

$$\text{cov}(y,x) = 0.618 = 0.62$$

$$\text{cov}(y,y) = \frac{1}{9}[0.24 + 1.46 + 0.98 + 0.084 + 1.188 + 0.624 + 0.096 + 0.66 + 0.016 + 1.02]$$

$$= 0.7164$$

$$\text{cov} = \begin{bmatrix} 0.6185 & 0.62 \\ 0.62 & 0.7169 \end{bmatrix}$$

Eigen values : $\det \begin{Bmatrix} 0.62-\lambda & 0.62 \\ 0.62 & 0.72-\lambda \end{Bmatrix}$

$$\Rightarrow (0.62-\lambda)(0.72-\lambda) - 0.3844$$
$$\Rightarrow 0.4464 - 0.62\lambda - 0.72\lambda + \lambda^2 - 0.3844$$
$$\Rightarrow \lambda^2 - 1.34\lambda + 0.062$$

$$0 = \lambda^2 - 1.34\lambda + 0.062$$
$$\sqrt{b^2-4ac} = 1.8 - 4(1)(0.062)$$
$$\Rightarrow 1.552$$

$$\lambda = \frac{1.34 \pm 1.246}{2} = 0.67 \pm 1.246$$

$$\lambda_1 = 1.9 \qquad \lambda_2 = -0.576$$

$$(cov - \lambda_1 I)\, u_1 = 0 \Rightarrow \begin{bmatrix} -1.28 & 0.62 \\ 0.62 & -1.18 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$1.28x - 0.62y = 0$$
$$-1.28x + 0.62y = 0$$
$$0.62x - 1.18y = 0$$
$$x = \frac{0.62y}{1.28}$$
$$x = 0.48y$$

$$1.28 - 0.62y = 0.62x - 1.18y \Rightarrow x(1.28 + 1.18)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.48 \\ 1 \end{bmatrix}$$

Normalize the eigen vector,

$$e_1 = \begin{bmatrix} \frac{0.48}{\sqrt{0.48^2+1^2}} \\ \frac{1}{\sqrt{0.48^2+1^2}} \end{bmatrix} = \begin{bmatrix} 0.43 \\ 0.9 \end{bmatrix}$$

Derive new data

$$PC_1 = e_1^T \begin{bmatrix} 2.5 - 1.81 \\ 2.4 - 1.91 \end{bmatrix} = \begin{bmatrix} 0.43 & 0.9 \end{bmatrix} \begin{bmatrix} 0.69 \\ 0.49 \end{bmatrix} = 0.7377$$

$$PC_2 = \begin{bmatrix} 0.43 & 0.9 \end{bmatrix} \begin{bmatrix} -1.31 \\ -1.21 \end{bmatrix} = -1.65$$

$$PC_5 = [0.43 \quad 0.9] \begin{bmatrix} 0.369 \\ 0.99 \end{bmatrix} = 1.06$$

$$PC_9 = [0.43 \quad 0.9] \begin{bmatrix} 0.09 \\ 0.29 \end{bmatrix} = 0.3$$

$$PC_5 = [0.43 \quad 0.9] \begin{bmatrix} 1.29 \\ 1.09 \end{bmatrix} = 1.54$$

$$PC_6 = [0.43 \quad 0.9] \begin{bmatrix} 0.49 \\ 0.79 \end{bmatrix} = 0.92$$

$$PC_7 = [0.43 \quad 0.9] \begin{bmatrix} 0.19 \\ -0.31 \end{bmatrix} = -0.2$$

$$PC_8 = [0.43 \quad 0.9] \begin{bmatrix} -0.81 \\ -0.81 \end{bmatrix} = -1.1$$

$$PC_9 = [0.43 \quad 0.9] \begin{bmatrix} -0.31 \\ -0.31 \end{bmatrix} = -0.41$$

$$PC_{10} = [0.43 \quad 0.9] \begin{bmatrix} -0.71 \\ -1.01 \end{bmatrix} = -1.2$$

$$\therefore PC = [0.7 \quad -1.65 \quad 1.06 \quad 0.3 \quad 1.54 \quad 0.92 \quad -0.2 \quad -1.1 \quad -0.41 \quad -1.2]$$