

PR ASSIGNMENT - 2
Design of Bayes Classifier
Deadline: 31 March 2022

Deliverables for this assignment:

1. Programming Assignment (MATLAB or Python)
 2. Code file and output screenshots for all. **You can make use of built in command to find the covariance matrix, where normalization is done using $1/n-1$.**
-

Q1. Find and plot the decision boundary between class ω_1 and ω_2 . Assume $P(\omega_1) = P(\omega_2)$.

$$\omega_1 = [1,6; 3,4; 3,8; 5,6]$$

$$\omega_2 = [3,0; 1,-2; 3,-4; 5,-2]$$

Q2. Find and plot the decision boundary between class ω_1 and ω_2 . Assume $P(\omega_1)=0.3$; $P(\omega_2)=0.7$

$$\omega_1 = [1,-1; 2,-5; 3,-6; 4,-10; 5,-12; 6,-15]$$

$$\omega_2 = [-1,1; -2,5; -3,6; -4,10; -5,12; -6,15]$$

Q3. Find and plot the decision boundary between class ω_1 and ω_2 . Assume $P(\omega_1) = P(\omega_2)$.

$$\omega_1 = [2,6; 3,4; 3,8; 4,6]$$

$$\omega_2 = [3,0; 1,-2; 3,-4; 5,-2]$$

Q4. Implement Bayes Classifier for Iris Dataset.

Dataset Specifications:

Total number of samples = 150

Number of classes = 3 (Iris setosa, Iris virginica, and Iris versicolor)

Number of samples in each class = 50

Use the following information to design classifier:

Number of training feature vectors (first 40 in each class) = 40

Number of test feature vectors (remaining 10 in each class) = 10

Number of dimensions = 4

Feature vector = <sepal length, sepal width, petal length, petal width>

If the samples follow a multivariate normal density, find the accuracy of classification for the test feature vectors.

Q5. Use only two features: Petal Length and Petal Width, for 3 class classification and draw the decision boundary between them (2 dimension, 3 regions also called as multi-class problem)

Q6. Consider the 128- dimensional feature vectors given in the “face feature vectors.csv” file. Use the following information to design and implement a Bayes Classifier.

Dataset Specifications:

Total number of samples = 800

Number of classes = 2 (labelled as “male” and “female”)

Samples from “1 to 400” belongs to class “male”

Samples from “401 to 800” belongs to class “female”

Number of samples per class = 400

Use the following information to design classifier:

Number of test samples (first 5 in each class) = 5

Number of training samples (remaining 395 in each class) = 395

Number of dimensions = 128

Design of Bayes Classifier

Given,

Iris dataset

$X = \langle x_1, x_2, x_3, x_4 \rangle$

Number of classes = $\omega_1, \omega_2, \omega_3$; $c=3$

$N=150$; $n(\omega_1)=n(\omega_2)=n(\omega_3)=50$

Bayes Rule:

$$\text{Find } P(\omega_i|X) = \frac{P(X|\omega_i).P(\omega_i)}{P(X)}$$

$P(X)$ is a constant for all classes; so it can be ignored.

Steps to follow in Iris Classification:

1. Find apriori probability $P(\omega_i) = \frac{n(\omega_i)}{N} = \frac{50}{150}$
2. Find $P(X|\omega_i)$, it's multivariate class, by following normal density

$$P(X|\omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2} \{ (X - \mu_i)^t \Sigma_i^{-1} (X - \mu_i) \} \right]$$

- 2 a. Find the mean vector
- 2 b. Find the covariance matrix, Σ_i
- 2 c. Find the $|\Sigma_i|$ and $|\Sigma_i|^{-1}$
3. Find $P(\omega_1|X)$, $P(\omega_2|X)$ and $P(\omega_3|X)$. Find the maximum and assign X to that class.
Also, plot the accuracy for :
 - i) Separate classes
 - ii) Overall performance
4. Find the discriminant function and draw the decision surface between the classes.

Note: The same steps can be followed for Q6.
