

Homework 2

CSE 802 - Pattern Recognition and Analysis

Total Points: 150

Instructor: Dr. Arun Ross

Due Date: March 9, 12:40pm

Please read the following instructions carefully:

1. You are permitted to discuss the following questions with others in the class. However, you must write up your own solutions to these questions. Any indication to the contrary will be considered an act of academic dishonesty.
 2. A hard-copy of this assignment must be submitted by March 9, 12:40 pm in class. In this copy, please include the names of individuals you discussed this homework with and the list of external resources (e.g., websites, other books, articles, etc.) that you used to complete the assignment (if any). Late submissions will not be graded.
 3. When solving equations or reducing expressions you must explicitly show every step in your computation and/or include the code that was used to perform the computation. Missing steps or code will lead to a deduction of points.
 4. Code developed as part of this assignment must be included as an appendix to your submission or inline with your solution.
-

1. Consider a set of 1-dimensional feature values (i.e., points) pertaining to a class that is available [here](#).
 - (a) [3 points] Plot the histogram of these points using a bin size of 2, i.e., each bin should have a range of 2.
 - (b) [4 points] Compute and report the mean and the biased variance of these points.
 - (c) [3 points] Assuming that the given points are generated by an underlying Gaussian distribution, plot the pdf function on the same graph as (a).
2. Consider the three-dimensional normal distribution $p(\mathbf{x}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (1, 1, 1)^t$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{bmatrix}$.
 - (a) [2 points] Compute and report the determinant of the covariance matrix, i.e., $|\boldsymbol{\Sigma}|$.
 - (b) [2 points] Compute and report the inverse of the covariance matrix, i.e., $\boldsymbol{\Sigma}^{-1}$.
 - (c) [3 points] Compute and report the eigen-vectors and eigen-values of the covariance matrix.
 - (d) [3 points] Compute and report the probability density at $(0, 0, 0)^t$ and at $(5, 5, 5)^t$.
 - (e) [2 points] Compute the Euclidean Distance between $\boldsymbol{\mu}$ and the point $(5, 5, 5)^t$.
 - (f) [3 points] Compute the Mahalanobis Distance between $\boldsymbol{\mu}$ and the point $(5, 5, 5)^t$.

3. [10 points] Suppose we have two normal distributions, corresponding to two classes (ω_1 and ω_2) with the same covariances but different means: $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$. In terms of their prior probabilities $P(\omega_1)$ and $P(\omega_2)$, state the condition that the Bayes decision boundary *not* pass between the two means.
4. [10 points] Consider a two-class problem with the following class-conditional probability density functions (pdfs):

$$p(x | \omega_1) \sim N(0, \sigma^2)$$

and

$$p(x | \omega_2) \sim N(1, \sigma^2).$$

Show that the threshold, τ , corresponding to the Bayes decision boundary is:

$$\tau = \frac{1}{2} - \sigma^2 \ln \left[\frac{\lambda_{12} P(\omega_2)}{\lambda_{21} P(\omega_1)} \right],$$

where we have assumed that $\lambda_{11} = \lambda_{22} = 0$.

5. [15 points] Consider a two-category classification problem involving one feature, x . Let both class-conditional densities conform to a Cauchy distribution as follows:

$$p(x | \omega_i) = \frac{1}{\pi b} \cdot \frac{1}{1 + \left(\frac{x - a_i}{b}\right)^2}, \quad i = 1, 2.$$

Assume a 0-1 loss function and equal priors.

- (a) Compute the Bayes decision boundary.
(b) Show that the probability of misclassification according to the Bayes decision rule is

$$P(error) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \left| \frac{a_2 - a_1}{2b} \right|.$$

- (c) Plot $P(error)$ as a function of $|a_2 - a_1|/b$.
(d) What is the maximum value of $P(error)$. Under what conditions will this occur?

6. [10 points] Consider the following class-conditional densities for a three-class problem involving two-dimensional features:

$$p(\mathbf{x} | \omega_1) \sim N((-1, -1)^t, I);$$

$$p(\mathbf{x} | \omega_2) \sim N((1, 1)^t, I);$$

$$p(\mathbf{x} | \omega_3) \sim \frac{1}{2} N((0.5, 0.5)^t, I) + \frac{1}{2} N((-0.5, -0.5)^t, I).$$

(Here, class ω_3 conforms to a **Gaussian Mixture Model (GMM)** with two components - one component is $N((0.5, 0.5)^t, I)$ and the other component is $N((-0.5, -0.5)^t, I)$ - whose weights are equal (i.e., $\frac{1}{2}$)).

- (a) In a 2D graph, mark the mean of ω_1 , ω_2 , and the two components of ω_3 . In the same graph, mark the point $\mathbf{x} = (0.1, 0.1)^t$.

- (b) Assuming a 0-1 loss function and equal priors, determine the class to which you will assign the two-dimensional point $\mathbf{x} = (0.1, 0.1)^t$ based on the Bayes decision rule.

7. Consider the following bivariate density function for the random variable \mathbf{x} :

$$p_{original}(\mathbf{x}) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 20 & 10 \\ 10 & 30 \end{bmatrix}\right).$$

- [5 points] What is the whitening transform, \mathbf{A}_w , of \mathbf{x} ? (You can use matlab (or any other programming language) to compute the eigenvectors/values).
 - [3 points] When the whitening transformation is applied to \mathbf{x} , what is the density function, $p_{transform}$ of the resulting random variable?
 - [5 points] Generate 10,000 bivariate *random* patterns from $p_{original}$ (if you are using matlab, then the *mvnrnd* function can be used to generate these patterns). Plot these patterns in a graph.
 - [5 points] Apply the whitening transform, \mathbf{A}_w , to the 10,000 bivariate patterns generated above. Plot the transformed patterns in a separate graph.
 - [2 points] Compare the patterns in 7c and 7d. **What do you observe?**
8. Consider a two-category (ω_1 and ω_2) classification problem with equal priors. Each feature is a two-dimensional vector $\mathbf{x} = (x_1, x_2)^t$. The class-conditional densities are:
 $p(\mathbf{x}|\omega_1) \sim N(\mu_1 = (0, 0)^t, \Sigma_1 = 2I)$,
 $p(\mathbf{x}|\omega_2) \sim N(\mu_2 = (2, 2)^t, \Sigma_2 = I)$.
- [7 points] Compute the Bayes decision rule and the Bayes decision boundary.
 - [3 points] Generate 10,000 bivariate *random* patterns from each of the two densities (if you are using matlab, then the *mvnrnd* function can be used to generate these patterns). Plot these patterns in a graph using different markers to distinguish the two classes. On the same graph, plot the Bayes decision boundary.
 - [5 points] Compute and report the empirical error rate and the confusion matrix when classifying these 10,000 patterns using the Bayes decision policy.
9. Consider a two-category classification problem involving two-dimensional feature vectors of the form $\mathbf{x} = (x_1, x_2)^t$. The two categories are ω_1 and ω_2 , and

$$p(\mathbf{x} | \omega_1) \sim N((-1, -1)^t, I),$$

$$p(\mathbf{x} | \omega_2) \sim N((1, 1)^t, I),$$

$$P(\omega_1) = P(\omega_2) = \frac{1}{2}.$$

- [10 points] Calculate the Bayes decision boundary and write down the Bayes decision rule assuming a 0-1 loss function. The Bayes decision rule has to be written in terms of the Bayes decision boundary.
- [5 points] What are the Bhattacharyya and Chernoff theoretical bounds on the probability of misclassification, $P(\text{error})$?

- (c) [5 points] Generate $n = 25$ test patterns from *each* of the two class-conditional densities and plot them in a two-dimensional feature space using different markers for the two categories (if you are using matlab, then the *mvnrnd* function can be used to generate these patterns). Draw the Bayes decision boundary on this plot for visualization purposes.
- (d) [5 points] What is the confusion matrix and empirical error rate when classifying the generated patterns using the Bayes decision rule?
- (e) [5 points] Can the empirical error rate exceed the theoretical bounds on the probability of misclassification? Why or why not? (Hint: Compute the empirical error rate multiple times by *repeatedly* generating $n = 25$ test patterns from *each* of the two class-conditional densities.)
10. [15 points] Consider a 1-dimensional classification problem involving two categories ω_1 and ω_2 such that $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$. Assume that the classification process can result in one of three actions:
- α_1 - choose ω_1 ;
 - α_2 - choose ω_2 ;
 - α_3 - do not classify.
- Consider the following loss function, λ :
- $\lambda(\alpha_1|\omega_1) = \lambda(\alpha_2|\omega_2) = 0$;
 - $\lambda(\alpha_2|\omega_1) = \lambda(\alpha_1|\omega_2) = 1$;
 - $\lambda(\alpha_3|\omega_1) = \lambda(\alpha_3|\omega_2) = 1/4$.
- For a given feature value x , assume that $p(x|\omega_1) = \frac{2-x}{2}$ and $p(x|\omega_2) = 1/2$. Here, $0 \leq x \leq 2$.
- Based on the Bayes minimum risk rule, what action will be undertaken when encountering the value $x = 0.5$?
-