Stevens Institute of Technology Department of Electrical and Computer Engineering

Spring Semester 2022

CpE 646 Pattern Recognition and Classification

Midterm Exam

Mar 11, 12:00 – 4:00 PM

Instructions:

- Provide all necessary intermediate steps in your work. You will get zero credit if you only provide the final result without justification or explanation.
- All steps and intermediate results should be composed into one single PDF or DOC file, which should be submitted through Canvas Assignments as your midterm solutions.
- You may use Matlab or Python for math calculations. Only the basic math functions are allowed, which include probability density functions, matrix inversion, eigen-decomposition. You may not use other built-in functions or library functions, such as PCA or LDA etc.
- If you are using Matlab, you can copy your command window script into a text file. If you are using Python, you can save your Python script in a text file. Your saved Matlab or Python script file should be submitted through Canvas Assignments separately, which will only be used for reference and will not be graded.

Name (print):	
tame (print).	

In a 3-dimensional 2-class classification problem, the following training sample vectors are given:

$$\mathbf{D}_{1} = \left\{ \begin{bmatrix} 7 \\ 3 \\ 13 \end{bmatrix}, \begin{bmatrix} 12 \\ 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 10 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ 4 \end{bmatrix} \right\} \text{ and } \mathbf{D}_{2} = \left\{ \begin{bmatrix} 0 \\ 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 3 \end{bmatrix} \right\}$$

Following are the steps to be taken to process this dataset.

Step 1. (20 points) Use Fisher linear discriminant analysis method to calculate a) calculate a transform vector w to reduce the data dimension to 1 dimension, and then b) calculate the projections of all the data samples in the resulting 1-dimensional space. (Show your steps and immediate results.)

Step 2. (20 points) Continue from Step 1. In the resulting 1-dimensional space of Step 1, assume class 1 and class 2 data samples follow two specific distributions as shown.

class 1:
$$p(x \mid \omega_1) = \begin{cases} \theta_1^2(x+1)e^{-\theta_1(x+1)} & x \ge -1 \\ 0 & \text{otherwise} \end{cases}$$

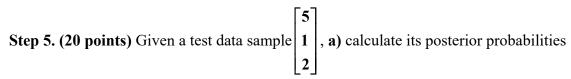
$$\text{class 2:} \quad p(x \mid \omega_2) = \frac{1}{\sqrt{2\pi} \sigma_2} \exp\left[-\frac{1}{2} \left(\frac{x-\mu_2}{\sigma_2}\right)^2\right]$$

For the class 1 distribution $p(x|\omega_1)$, given N independent and identically distributed training samples $x_1, x_2, ..., x_N$, use maximum likelihood estimation to estimate the parameter θ_1 . (This derivation should be done by hand.) (ML estimation of class 2 distribution parameters can be found in lecture notes CPE646-4)

Step 3. (20 points) Use the maximum likelihood estimation formula from Step 2 and lecture notes, estimate the parameters of these two distributions $p(x|\omega_1)$ and $p(x|\omega_2)$ based on the 1-diementional projections of the training samples from Step 1.

Step 4. (20 points) Based on the results from **Step 3**, derive the Bayesian decision boundary with the following risk matrix. Assume the two classes have equal prior probabilities. (Simplify your expression as much as you can, you will be graded accordingly.)

	ω_1	@ 2
a_1	0	3
a_2	1	0



 $P(\omega_1|x)$ and $P(\omega_2|x)$. b) If the risk matrix in Step 4 is used, calculate the associated Bayesian risk, and c) classify this data sample. (Show your steps and immediate results.)