EECE5644 Summer 1 2022 – Practice

This is a set of practice questions to give you an idea of the homework format and what to expect. Please note that you should NOT submit solutions on Canvas. Also note that some of the question topics used in this practice sheet may not have yet been covered in lectures.

Below is what a typical homework cover sheet will look like for further reference:

Please submit your solutions at the classroom assignments page in Canvas. Please upload a single PDF file that includes all math, numerical and visual results, and code (as appendix or as a link to your online code repository for this assignment). If you point to an online repository, do NOT edit the contents after the deadline, because TAs may interpret a last-modified timestamp past the deadline as a late submission of the assignment. Only the contents of the PDF will be graded. Please do NOT link from the pdf to external documents online where results may be presented (e.g. online notebooks of any kind).

This is a graded assignment and the entirety of your submission must contain only your own work. You may benefit from publicly available literature including software (not from classmates), as long as these sources are properly acknowledged in your submission. Copying math or code from each other are not allowed and will be considered as academic dishonesty. While there cannot be any written material exchange between classmates, verbal discussions to help each other are acceptable. Discussing with the instructor, the TA, and classmates at open office periods to get clarification or to eliminate doubts is acceptable.

By submitting a PDF file in response to this take home assignment you are declaring that the contents of your submission, and the associated code is your own work, except as noted in your citations to resources.

Question 1 (30%)

Design a classifier that achieves minimum probability of error for a three-class problem where the class priors are respectively $P(L=1)=0.15$, $P(L=2)=0.35$, $P(L=3)=0.5$ and the class-conditional data distributions are all Gaussians for two-dimensional data vectors:

$$N(\begin{bmatrix}-1\\0\end{bmatrix},\begin{bmatrix}1 & -0.4\\-0.4 & 0.5\end{bmatrix})\quad N(\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0.5 & 0\\0 & 0.2\end{bmatrix})\quad N(\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}0.1 & 0\\0 & 0.1\end{bmatrix})$$

Generate 10000 samples according to this data distribution, keep track of the true class labels for each sample. Apply your optimal classifier designed as described above to this dataset and obtain decision labels for each sample. Report the following:

- actual number of samples that were generated from each class;
- the confusion matrix for your classifier consisting of number of samples decided as class $r \in \{1,2,3\}$ when their true labels were class $c \in \{1,2,3\}$, using r,c as row/column indices;
- the total number of samples misclassified by your classifier;
- an estimate of the probability of error your classifier will achieves, based on these samples;
- a visualization of the data as a 2-dimensional scatter plot, with true labels and decision labels indicated using two separate visulization cues, such as marker shape and marker color;
- a clear but brief description of the results presented as described above.

Question 2 (30%)

Consider a scalar real-valued feature x that has the following probability distributions under two class labels as follows:

$$p(x|l=1) = \begin{cases} x & \text{if } 0 \leq x \leq 1 \\ 2-x & \text{if } 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases} \qquad p(x|l=2) = \begin{cases} 2x-3 & \text{if } \frac{3}{2} \leq x \leq \frac{5}{2} \\ 0 & \text{otherwise} \end{cases}$$

Minimum Expected Loss Classification

Let loss values be set to $Loss(\text{Decide } i \text{ when truth is } j) = \lambda_{ij} \geq 0$ for $i, j \in 1, 2$, with loss of erroneous decisions assigned to be greater than corresponding correct decisions. Let the class priors be $q_1 = p(l=1)$ and $q_2 = p(l=2)$, respectively. Express the minimum expected loss decision rule with a discriminant function that is simplified as much as possible. Show your steps.

Maximum a Posteriori Classification

For the case when 0-1 loss assignments are used, the minimum expected risk classifier reduces to the maximum a posteriori classification rule. For this case, express the maximum a posteriori classification rule.

Maximum Likelihood Classification

In addition to 0-1 loss assignments, assume that the class priors are equal. In this case, ML classification achieves minimum expected risk. For this case, express the maximum likelihood classification rule.

Question 3 (40%)

Generate N iid random 2-dimensional samples from two Gaussian pdfs $N(\boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i)$ with specified prior class probabilities q_i for $i \in 0, 1$. Set

$$\boldsymbol{\mu}_0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \boldsymbol{\Lambda}_0 = \begin{bmatrix} 16 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{\Lambda}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix}$$

Make sure to keep track of the true label of each sample. In the numerical results use $N = 10000$, $q_0 = 0.35$, and $q_1 = 0.65$. For the following two classification methods, generate two sets of classification decisions for different values of their respective thresholds. For each classifier, plot the estimated false-positive and true-positive probabilities for each value of the threshold; overlay the two curves in the same plot for easy visual comparison.

Minimum Expected Loss Classifier for Two Gaussian Classes

Determine and implement the minimum expected loss classifier parametrized by a threshold γ in the following form:

$$ln\ p(\mathbf{x}; \boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1) - ln\ p(\mathbf{x}; \boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0) \underset{Decide 0}{\overset{Decide 1}{\gtrless}} ln\ \gamma$$

where $\gamma > 0$ is a scalar that depends on loss and prior values. Make this threshold take many values along the positive real axis, and for each threshold value, classify every sample, empirically estimate the true positive and false positive probabilities (by counting samples that fall in each category), then plotting the true-positive versus false-positive performance at each threshold in a plot. Sweep the positive real axis by sampling densely to see the true-positive vs false-positive curves for the classifier in detail. Here class 1 is *positive* and class 0 is *negative*.

Fisher Linear Discriminant Analysis for Two Gaussian Classes

Implement the Fisher LDA classifier using the true mean vectors and covariance matrices provided above to obtain a classifier γ in the following form:

$$\mathbf{w}^{\mathrm{T}}\mathbf{x} \underset{Decide 0}{\overset{Decide 1}{\gtrless}} \gamma$$

where $\gamma \in (-\inf, \inf)$. Repeat the exercise in the previous section, this time assigning many values to this threshold parameter, estimating the true- and false-positive probabilities from data classification label matches and mismatches, and plotting the true-positive versus false-positive performance curves as the threshold changes. Overlay this curve on top of the previous one.