EECE5644 Fall 2019 – Homework 2

Submit: Monday, 2019-October-07 before 09:00ET

Please submit your solutions to Blackboard in a single PDF file that includes all math and numerical results in the main body. Make sure that you cite all resources you benefit from (books, papers, software packages). Also include your code in one of the following ways: (Acceptable) upload a ZIP file containing all code files, (Preferred) keep your code in an online version control repository (such as GitHub) and provide a link to the relevant online repository. This is a graded assignment and the entirety of your submission must contain only your own work. You may benefit from literature including software (as allowed by specific restrictions in questions), as long as these sources are properly acknowledged in your submission.

Question 1 (30%)

Problem 2.13 from the textbook (included in the next page).

Question 2 (35%)

Write a function that generates a specified number of independent and identically distributed samples paired with the class labels that generated these samples. Specifically, the data distribution is a mixture of Gaussians with specified prior probabilities for each Gaussian class conditional pdf, as well as respective mean vectors and covariance matrices. Generate and visualize data in the form of scatter plots, with a color/marker based identification of the class label for each sample for each of the following cases (using Matlab syntax for 2×2 matrices):

1. Number of samples = 400; class means $[0,0]^T$ and $[3,3]^T$; class covariance matrices both set to $\mathbf{I}$; equal class priors.
2. All parameters same as (1), but both covariance matrices changed to $[3,1;1,0.8]$.
3. Number of samples = 400; class means $[0,0]^T$ and $[2,2]^T$; class covariance matrices $[2,0.5;0.5,1]$ and $[2,-1.9;-1.9,5]$; equal class priors.
4. Same (1), but prior for class priors are 0.05 and 0.95.
5. Same (2), but prior for class priors are 0.05 and 0.95.
6. Same (3), but prior for class priors are 0.05 and 0.95.

Make sure your plots include axis labels, titles, and data legends. Describe how your sampling procedure works, using zero-mean identity-covariance Gaussian sample generators.

Additionally, for each of these datasets, use the maximum-a-priori (MAP) classification rule (using full knowledge of the respective data pdfs) and produce inferred class labels for each data samples. In accompanying visualizations, demonstrate scatter plots of the data for each case along with their inferred (decision) labels. For each case, count the number of errors and estimate the probability of error based on these counts.

Question 3 (35%)

For the datasets you generated in Question 2, implement and apply the Fisher Linear Discriminant Analysis classifier with the decision threshold for the linear discriminant score set to minimize the smallest probability error you can achieve on the specific data sets generated for each case. Visualize the one-dimensional Fisher LDA discriminant scores and decision labels for each sample in separate plots for each case. *Note: We will soon discuss the principle of cross-validation that dictates parameter selection and performance assessment must use independent datasets.*

Section 2.4

13. In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0 & i = j \qquad i, j = 1, \ldots, c \\ \lambda_r & i = c + 1 \\ \lambda_s & \text{otherwise,} \end{cases}$$

where λ_r is the loss incurred for choosing the $(c + 1)$th action, rejection, and λ_s is the loss incurred for making any substitution error. Show that the minimum risk is obtained if we decide ω_i if $P(\omega_i|\mathbf{x}) \geq P(\omega_j|\mathbf{x})$ for all j and if $P(\omega_i|\mathbf{x}) \geq 1 - \lambda_r/\lambda_s$, and reject otherwise. What happens if $\lambda_r = 0$? What happens if $\lambda_r > \lambda_s$?