**ECE 8527: Introduction to Machine Learning and Pattern Recognition**

# HW No. 2: ML vs. Bayesian Estimation

The tasks to be accomplished in this homework assignment are:

1. Generate two 2D GRVs with a mean of (1,1) and (-1,-1) respectively. Plot the theoretical probability of error for an ML classifier as a function of the prior probabilities and the covariance matrices. Since there are a number of degrees of freedom, determine the best way to visualize the results.
2. Compare the theoretical results in (1) to those obtained when you construct an ML classifier by generating [100, 1,000, 10,000, 100,000] random variables for each class. Estimate the means and variances from the data. Only consider the equal priors case for this example, and focus on a small representative set of covariances.
3. Generate two 2D Gaussian GRVs with means of [-2,-5] and [3, 6] and covariance matrices of  and .
   1. Construct an ML estimator and measure the error rate.
   2. Convert each GRV to a Gaussian GRV with an identity covariance matrix by performing Principal Components Analysis (PCA).
   3. Classify each data point by transforming it to the PCA space using a whitening transformation and computing the distance from the mean. Select the class assignment by choosing the class that has the smallest distance. This is essentially an ML classifier, but implemented in a slightly different way. Do your results match part (a)?
   4. Examine the eigenvectors of each covariance matrix and relate those to the support region for each GRV. It is preferable to visualize this with a graph of the vectors overlaid on the support region.
4. Following the example presented in the notes, assume you have a 1D GRV with mean and variance. Demonstrate estimation of the mean and variance using the theoretical results derived in class for a Bayesian estimation. Compare this to an ML estimation. Show convergence as the number of data points is increased from 100 to 100K.
5. Consider the Bayesian estimation of the mean of a one-dimensional Gaussian distribution. Suppose you are given the prior for the mean as  . Generate 1,000 1D GRVs for GRV[1,1]. Assume  and  (assume the variance is known). Plot the Bayesian estimate of  and  as a function of the number of data points, n. Do this for 100, 1,000, 10,000 and 100,000 points. Explain your results.