$$|HW \# 2|$$

EEL 6825 - Fall 2001

Due Friday, September 21, 2001 in class. Late homework will lose $e^{\# of days late} - 1$ percentage points. Click on http://www.cnel.ufl.edu/hybrid/harris/latepoints.html to see the current penalty.

Your homework should be in two distinct sections. The first section should show the answers, explanations, plots, hand calculations etc. that you need to answer the questions in parts A-C. The second section should contain all of the Matlab code that you have written to generate the answers in the first section. You don't need a computer for most of parts A and B but of course you may use one to check your work if you like.

PART A: Short Answer Questions

A1 (5 points) You are given the heights and weights of a certain set of individuals in unknown units. Which one of the following six matrices is the most likely to be the sampled covariance matrix? Explain your reasoning.

1.232 0.013	1.232 0.867	1.232	-0.867
0.013 2.791	-0.867 2.791	-0.867	3.307
1.232 3.307	1.232 0.867	1.232	3.307
0.013 2.791	0.867 2.791	3.307	2.791

A2 (5 points) Sample data from one class are given as:

$\begin{bmatrix} 3 \end{bmatrix}$	[-1]	$\begin{bmatrix} -1 \end{bmatrix}$	$\begin{bmatrix} -1 \end{bmatrix}$
$\left\lfloor 2 \right\rfloor$	$\left\lfloor -2 \right\rfloor$	$\left[\begin{array}{c}2\end{array}\right]$	$\left\lfloor -2 \right\rfloor$

Compute the sampled mean and sampled covariance matrix (by hand). Make sure to use estimators that are unbiased. Show your work.

- A3 (5 points) Problem 2-21 in DHS
- A4 (5 points) Problem 2-39 in DHS
- A5 (5 points) Problem 3-2 in DHS

PART B: Maximum Likelihood Problem

Consider the following probability distribution:

$$p(x) = \begin{cases} (k+1)x^k & \text{for } 0 \le x \le 1\\ 0 & \text{else} \end{cases}$$

- B1 (5 points) For what values of k is this distribution valid? Verify that this distribution integrates to one for all valid values of k.
- B2 (15 points) Given N data points $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$ sampled from a distribution, derive a formula for the maximum likelihood value of k for the distribution.
- B3 (5 points) Suppose N = 3 and $x^{(1)} = e^0 = 1$, $x^{(2)} = e^{-1}$, $x^{(3)} = e^{-2}$. What the numerical value of the maximum likelihood estimate of k?

PART C: Computer Experiment: Bayes Classifier with two classes

Two normal distributions are characterized by:

$$P(\omega_1) = P(\omega_2) = 0.5$$

$$\mu_1 = \begin{bmatrix} 0\\0 \end{bmatrix}, \mu_2 = \begin{bmatrix} 2\\3 \end{bmatrix}, \Sigma_1 = \begin{bmatrix} 1&0\\0&2 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 1&0.8\\0.8&1 \end{bmatrix}$$

- C1 (5 points) Generate 200 points from each of these distributions. Show a scatter plot of the data using two different symbols (and colors) to identify the classes.
- C2 (5 points) Compute the sampled mean and covariance matrix for each class. How close are these values to the true mean and covariance matrix?
- C3 (20 points) Write a matlab program that classifies the data using a Bayes Classifier. Assume the data comes from Normal distributions but use your estimated parameters in the classifier. What is the percent error that you find? Run your program several times and average your results to improve your estimate of the expected error. Also, calculate the standard deviation of your error estimate.
- C4 (10 points) Compute the Bhatacharrya bound for this problem. How does it compare to your actual error from [C3]. Remember from class that the Bhatacharrya bound is an upper bound on the expected Bayes error for Normal distributions and is given by:

$$BOUND = e^{-K} \sqrt{P(\omega_1)P(\omega_2)}$$

where K is given by:

$$K = \frac{1}{8}(\mu_2 - \mu_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2}\right]^{-1}(\mu_2 - \mu_1) + \frac{1}{2}\ln\left(\frac{\left|\frac{\Sigma_1 + \Sigma_2}{2}\right|}{\sqrt{\left|\Sigma_1\right|\left|\Sigma_2\right|}}\right)$$

Again use the estimated parameters in your calculation.

- C5 (10 points) Have the computer draw the decision surface on the scatter plot. (Do not hard code the particular boundary for this problem, you should write a general program.) Hint: think about using the functions meshgrid and contour in matlab.
- Extra Credit Quantify what happens to the error calculated in part [C3] as you change the number of samples in each class. Plot a graph of average error vs. the number of sample points as you vary the number of samples from very low to very high numbers. Include error bars in your plots. Explain your results.