ENGN 2520 / CSCI 1420
Spring 2017
Homework 2
Due Wednesday February 22 at 4pm

INCLUDE THIS COVER PAGE WITH YOUR HOMEWORK

NAME:

BANNER ID:

BROWN EMAIL:

COLLABORATED WITH:

(LEAVE THIS BLANK)

problem	grade	memo
1		
2		
3		
total		

IMPORTANT: Students may discuss and work on homework problems in groups. However, each student must write down their solutions independently. All of the work submitted should be your own. Each student should write on the problem set the set of people with whom they collaborated.

Problem 1

Let x be a real valued random variable with a uniform distribution $p(x)$ on some unknown interval $[a, b]$. Suppose we have a training set T with k independent samples from $p(x)$. What is the maximum likelihood estimator for $[a, b]$? Justify your answer.

Problem 2

Note that (Q1) and (Q2) below are hypothetical questions. You should answer questions (a), (b), (c) and (d).

Alice and Bob are working together to estimate a function mapping the chemical composition of a solar array to the power output. To collect training data Alice and Bob work in turns in the lab, making new compositions and measuring power output. Suppose Bob is not as carefull as Alice in making his measurements. This leads to some questions:

(Q1) Should Alice and Bob ignore Bob's measurements in estimating their function?

(Q2) How can they incorporate both sets of measurements in a reasonable way?

We can capture the situation with a simple mathematical model. Let $f_w : X \to \mathbb{R}$ be a function defined by a feature map $\phi : X \to \mathbb{R}^M$ and a vector of parameters $w \in \mathbb{R}^M$,

$$f_w(x) = w^T \phi(x)$$

Let T_A and T_B be two sets of training examples. We assume the errors in the training examples are independent but are larger in T_B compared to T_A.

For (x, y) in T_A we assume $y = f_w(x) + e$ with error e distributed according to a Normal distribution $N(0, \sigma_A^2)$. For (x, y) in T_B we assume $y = f_w(x) + e$ with error e distributed according to a Normal distribution $N(0, \sigma_B^2)$. The errors are independent and $\sigma_B^2 > \sigma_A^2$.

Suppose we know σ_A^2 and σ_B^2. What is the maximum likelihood estimate of w?

$$w_{\text{ML}} = \max_w p(T_A, T_B | w)$$

(a) Show that w_{ML} minimizes a sum of *weighted* squared differences. The sum should have one term per example in T_A and one term per example in T_B. Justify your answer.

(b) Show how to compute w_{ML} by solving a linear system. Justify your answer.

(c) What does this mathematical model say about questions (Q1) and (Q2) above?

(d) Suppose we don't know σ_A and σ_B. How can we estimate w?

Problem 3

In this problem you will experiment the least absolute deviation method for regression. The data for this problem is available on the course website. The data is similar to what you used for Homework 1, but there are a few outliers in the training set. You should review the notes on robust regression from class.

You will use polynomial basis functions to estimate a polynomial $f_w(x)$ using (1) sum of squared differences and (2) sum of absolute deviations. For (1) you should solve a linear system. For (2) you should use 'linprog' in Matlab to solve the resulting linear program. Type 'help linprog' in the Matlab prompt to learn how to use that package.

(a) Use the training data to estimate two degree 2 polynomials, one with each regression method. Make a plot showing the training set and the two polynomials you estimate. You should clearly label the polynomials in the plot according to which regression method was used for each one.

(b) Repeat part (a) using degree 4 polynomials.

(c) What can you say about the differences between the two approaches for regression based on these experiments?

Submit your Matlab source code along with your homework. You should include the plots for parts (a) and (b) in your writeup.