ENGN 2520 Spring 2024
Homework 2

TURNING IN: Upload a PDF file to the course website.

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

In this problem you will experiment with the least absolute deviation method for robust 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 differences. For (1) you should solve a linear system. For (2) you should use linear programming.

In Matlab you can use 'linprog' to solve a linear program. Type `'help linprog'` in the Matlab prompt to learn how to use that package. In python you can use various packages such as `cvxopt` or `scipy` (`scipy.optimize.linprog`) for linear programming. If you use `scipy.optimize.linprog` you should include the parameter `'bounds = (None,None)'` to avoid putting a lowerbound of 0 on each variable.

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

Problem 2

Consider linear regression with absolute error loss and ℓ_1 regularization. We have a fixed feature map $\phi : \mathbb{X} \to \mathbb{R}^M$, a parameter λ, and a training dataset with N examples $T = \{(x_1, y_1), \ldots, (x_N, y_N)\}$. Our goal is to select w minimizing,

$$\left(\sum_{i=1}^{N} |y_i - w^T \phi(x_i)| \right) + \lambda ||w||_1.$$

Show how this optimization can be posed as a linear programming problem.