DSC 40A - Extra Practice Session 1

Wednesday, January 12, 2022

Problem 1. Visualizing Transformations

a) Suppose we have a dataset $x_1, x_2, \ldots, x_n$ and we transform the data set by $f(x) = 2x$. How does applying this transformation affect the order of the data? The spacing of the data?

b) Suppose we have a dataset $x_1, x_2, \ldots, x_n$ of positive numbers and we transform the data set by $f(x) = x^2$. How does applying this transformation affect the order of the data? The spacing of the data?

c) For the previous transformation, what would happen to the order of the data if the data values were not all positive?

d) Can you think of a transformation $f(x)$ for which the order and spacing of the data values is preserved?

e) Can you think of a transformation $f(x)$ for which the order of the data is unchanged and the spacing of the data is decreased?

f) Can you think of a transformation $f(x)$ for which the order of the data is reversed and the spacing of the data is unchanged?

Problem 2. Minimizers and Maximizers

a) Fill in the blank and prove your result:

If x^* is a minimizer of $f(x)$ then it's a ________________ *of $g(x) = 5f(x) + 3$.*

b) Fill in the blank and prove your result:

If x^* is a minimizer of $f(x)$ then it's a ________________ *of $g(x) = -(f(x))^2$.*

Problem 3. Max's Other Idea

In our lecture, we argued that one way to make a good prediction h is to minimize the mean absolute error:

$$R(h) = \frac{1}{n} \sum_{i=1}^{n} |h - y_i|.$$

We saw that the median of $y_1, \ldots, y_n$ is the prediction with the smallest mean error. Your friend Max has many ideas for other ways to make predictions. In Homework 1, you'll evaluate one of those ideas. Here, we'll evaluate another.

Max thinks that instead of minimizing the mean error, it is better to minimize the *maximum error*:

$$M(h) = \max_{i=1,\ldots,n} |y_i - h|$$

In this problem, we'll see if Max has a good idea.

a) Suppose that the data set is arranged in increasing order, so $y_1 \leq y_2 \leq \cdots \leq y_n$. Argue that $M(h) = \max(|y_1 - h|, |y_n - h|)$.

b) For arbitrary values of y_1 and y_n, show how to draw the graph of $M(h) = \max(|y_1 - h|, |y_n - h|)$.

c) At what value h^* is $M(h)$ minimized? Did Max have a reasonable idea?