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?

$a < b$

$2a < 2b$

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?

find two points that get closer after applying f

0.1 and 0.2

$\xrightarrow{f}$ 0.01 and 0.04 (closer)

$$\boxed{\begin{array}{l} a < b \\ a^2 < b^2 \end{array} \text{ from groupwork 1} \quad a < b \Rightarrow a \cdot a < a \cdot b < b \cdot b}$$

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

order can change

find $a < b$ where $b^2 < a^2$

$-3 < 2$ $\qquad 2^2 < (-3)^2$

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

$f(x) = x \pm c$

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?

$f(x) = cx$ where $0 < c < 1$

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?

$f(x) = -x$

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 minimizer *of $g(x) = 5f(x) + 3$.*

def. of minimizer

$f(x^*) \leq f(x)$ for any x

$5f(x^*) \leq 5f(x)$ for any x

$5f(x^*)+3 \leq 5f(x)+3$ for any x

goal $g(x^*) \leq g(x)$ for any x

b) Fill in the blank and prove your result:

If x^* is a minimizer of $f(x)$ then it's a neither *of $g(x) = -(f(x))^2$.*

try it

Let's find a function $g(x)$ such that minimizer of f is a maximizer of g.

$\frac{1}{f(x)}$?? $\frac{1}{e^{f(x)}}$

$g(x) = -f(x)$, $g(x) = -3f(x)+2$

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:

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

← normal approach

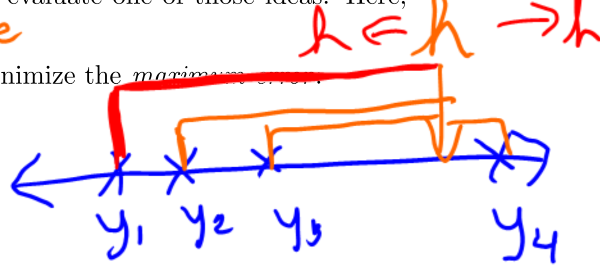
We saw that the median of $y_1, \ldots, y_n$ is the prediction with the smallest mean error. Your friend Max has many ideas for other ways to make predictions. In Homework 1, you'll evaluate one of those ideas. Here, we'll evaluate another.

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

distance

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

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



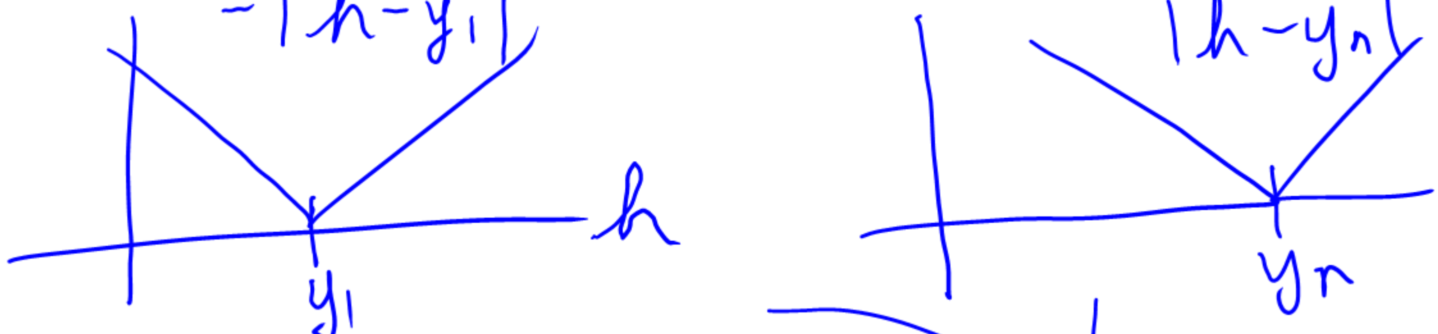
a) Suppose that the data set is arranged in increasing order, so $y_1 \leq y_2 \leq \cdots \leq y_n$. Argue that $M(h) = \max(|y_1 - h|, |y_n - h|)$.

If y_i with $i \neq 1, i \neq n$ where h is furthest from y_i: if $h > y_i$, then y_i isn't furthest from h because y_1 is even further.

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

$f(h) = |y_1 - h|$ for arbitrary y_1

$= |h - y_1|$



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

$\boxed{\dfrac{y_1 + y_n}{2}}$ (halfway between y_1 and y_n)

combine on same axes