UGBA 198 Homework 1

September 13, 2018

1 Deriving Mode, Median, and Mean Under a Unifying Framework

Take a set of n numbers $x_1, x_2...x_n$. How do we summarize this data? In grade school, we've seen three possibilities: mode, median, and mean. As it turns out, we can interpret all three under one framework! The mode, median, and mean minimize the L0, L1, and L2 distances to all the numbers, respectively. More formally, they are the solutions to the following objective problem, for $m = 0, 1, 2$.

$$\min_u \sum_{i=1}^n |x_i - u|^m \tag{1}$$

To make this argument simpler, **define** $0^0 = 0$.

1. **Show that "mean" is the solution to (1) when $m = 2$.** Hint: Use what you know from calculus to minimize this objective.

$$\min_u \sum_{i=1}^n (x_i - u)^2$$

2. **Show that "mode" is the solution to (1) when $m = 0$.** Hint: Use intuition.

$$\min_u \sum_{i=1}^n (x_i - u)^0$$

3. Start with two points–x_1 and x_2. **How do we choose u to minimize (1)?** Is there only one possible value or many? Hint: Consider two people standing on a line. Each person pays \$1 to walk one foot. How do you pick a point for them to meet, so that they minimize the total cost they pay?

$$\min_u \sum_{i=1}^n |x_i - u|^1 = |x_1 - u| + |x_2 - u|$$

4. Now, take three points–x_1, x_2 and x_3. **How do we choose u to minimize (1)?** Hint: Again, consider three people. Minimize the total distance all of them, collectively, need to walk.

5. Generally, how do we pick u to minimize (1) for an even number of points? For an odd number of points?