A Weighted Average

Andrew Candela

September 19, 2016

Motivation

Imagine that you are considering opening a gas station. You know you can secure a regular fuel delivery for \$2.5 per gallon. But can you justify charging your future customers more than that? In order to quickly estimate what you might be able to charge, you look up the cost per gallon at 10 nearby gas stations. Your challenge now is to come up with a reasonable estimate for what you should charge based on what everyone around you is charging. Your estimate should satisfy a few properties. It should:

- Behave like a weighted average of the prices of the other stations (the sum of the weights must equal 1)
- Give greater weight to stations that are closer to you and less to stations that are farther away.

Specifically, we want the following two conditions to hold:

$$\sum_{i=1}^{n} W_i = 1 \text{ and } W_i \in (0, 1] \text{ for all } i \tag{1}$$

$$\frac{W_i}{W_j} = \frac{D_i}{D_j} \tag{2}$$

We can combine the two conditions above to arrive at the following assignment for the wieghts:

$$W_i = \frac{D_i}{\sum_{i=1}^{n} D_i}$$

In the process of writing this I realized that it is far simpler than I originally thought it would be. It's almost not even worth sharing... but I suppose I need the practice.

Algorithm

The inputs here are the point $P = (x, y)$ at which we wish to have our estimate evaluated, and a dataset composed of 3 vectors:

- the x coordinate $\boldsymbol{X}$
- the y coordinate $\boldsymbol{Y}$
- the value to be estimated $\boldsymbol{V}$

We can now build a distance vector $\boldsymbol{D}$ as

$$D_i = \sqrt{(X_i - x)^2 + (Y_i - y)^2}$$

and a weight vector $\boldsymbol{W}$ as

$$W_i = \frac{D_i}{\sum_{i=1}^{n} D_i}$$

and then the estimate E as $E = \boldsymbol{W} \cdot \boldsymbol{V}$