Math 240 — Hw 7
Sara Jamshidi, Mar 21, 2025

Consider a small dataset with the following points:

x	y
1	2
2	3
3	5
4	4

We want to fit a linear model $y = mx + b$ to this data using gradient descent.

1. Identify a good set of initial parameters for your model.

   Solution:

   We can start with initial parameters $m = \frac{2}{3}$ and $b = \frac{4}{3}$. This is the rough estimate of the "rise" over "run" for the dataset (looking at the first and last points) with an intercept needed to connect them.

2. Compute the mean absolute error of your initial parameters:

$$E(m, b) = \frac{1}{n}\sum_{i=1}^{n} |y_i - (mx_i + b)|$$

   Solution: Using $m = \frac{2}{3}$ and $b = \frac{4}{3}$:

$$E\left(\frac{2}{3}, \frac{4}{3}\right) = \frac{1}{4}\left(\left|2 - \left(\frac{2}{3}\cdot 1 + \frac{4}{3}\right)\right| + \left|3 - \left(\frac{2}{3}\cdot 2 + \frac{4}{3}\right)\right| + \left|5 - \left(\frac{2}{3}\cdot 3 + \frac{4}{3}\right)\right| + \left|4 - \left(\frac{2}{3}\cdot 4 + \frac{4}{3}\right)\right|\right)$$

$$E\left(\frac{2}{3}, \frac{4}{3}\right) = \frac{1}{4}\left(|2-2| + |3-3| + |5-4| + |4-4|\right) = \frac{1}{4}(0+0+1+0) = \frac{1}{4} = 0.25$$

3. Make one adjustment to m and b using the derivative of E with respect to m and b. What is your updated model?

   Solution:

   The derivatives of E with respect to m and b are:

$$\frac{\partial E}{\partial m} = -\frac{1}{n}\sum_{i=1}^{n} \operatorname{sign}(y_i - (mx_i + b))x_i$$

$$\frac{\partial E}{\partial b} = -\frac{1}{n}\sum_{i=1}^{n} \operatorname{sign}(y_i - (mx_i + b))$$

Plugging in the values of the dataset give us:

$$\frac{\partial E}{\partial m} = -\frac{1}{4}\left(\text{sign}(0)\cdot 1 + \text{sign}(0)\cdot 2 + \text{sign}(1)\cdot 3 + \text{sign}(0)\cdot 4\right) = -\frac{1}{4}\left(0+0+3+0\right) = -0.75$$

$$\frac{\partial E}{\partial b} = -\frac{1}{4}\left(\text{sign}(0) + \text{sign}(0) + \text{sign}(1) + \text{sign}(0)\right) = -\frac{1}{4}\left(0+0+1+0\right) = -0.25$$

Updating m and b with a learning rate of 0.1 (optional):

$$m_{\text{new}} = m - \alpha\frac{\partial E}{\partial m} = \frac{2}{3} - 0.1\times(-0.75) = \frac{2}{3} + 0.075 \approx 0.7420$$

$$b_{\text{new}} = b - \alpha\frac{\partial E}{\partial b} = \frac{4}{3} - 0.1\times(-0.25) = \frac{4}{3} + 0.025 \approx 1.36$$

The updated model is $y = 0.7420x + 1.36$.

4. Compute the mean absolute error of your updated model.

$$\begin{aligned}
E\left(0.742, 1.36\right) &= \frac{1}{4}\left(|2-(0.742\cdot 1+1.36)|\right.\\
&\quad + |3-(0.742\cdot 2+1.36)|\\
&\quad + |5-(0.742\cdot 3+1.36)|\\
&\quad \left.+ |4-(0.742\cdot 4+1.36)|\right)\\
&= \frac{1}{4}\left(|2-2.102| + |3-2.844| + |5-3.586| + |4-4.328|\right)\\
&= \frac{1}{4}\left(0.102+0.156+1.414+0.328\right)\\
&= \frac{1}{4}\times 2.000\\
&= 0.5
\end{aligned}$$

Note that we saw the error grow! Remember, this can happen when we have a learning rate that is too large.