BRAC University

CSE427: Machine Learning Summer 2023

Assignment 2: Gradient Descent

August 6, 2023

Regression: Let's say, we want to perform linear regression on a dataset containing m examples and n features. Our output is a linear funcion as follows:

$$\bar{y}_i = w_1 x_{i,1} + w_2 x_{i,2} + ...+ w_n x_{i,n} + b$$

Now, if the error is E, then the gradient descent weight update rules should be as follows:

$$w_i = w_i - \lambda \frac{\delta E}{\delta w_i} \text{ for } i \in \{1, 2, \cdots, n\}$$
$$b = b - \lambda \frac{\delta E}{\delta w_i}$$

For the following loss functions E, find $\frac{\delta E}{\delta w_i}$ and $\frac{\delta E}{\delta b}$.

1. **Mean Squared Error:**

$$E = \frac{1}{m} \sum_{i=1}^{m} (y_i - \bar{y}_i)^2$$

2. **Sum of Squared Error:**

$$E = \sum_{i=1}^{m} (y_i - \bar{y}_i)^2$$

3. **Mean Squared Logged Error:** Sometimes, y_i and $\bar{y}_i$ can be too large. So, we use the following loss function.

$$E = \frac{1}{m} \sum_{i=1}^{m} (\log y_i - \log \bar{y}_i)^2$$

4. **Mean Absolute Error:**

$$E = \frac{1}{m} \sum_{i=1}^{m} |y_i - \bar{y}_i|$$

5. **Huber Loss:**

$$E = \frac{1}{m}\sum_{i=1}^{m}\begin{cases} \frac{1}{2}(y_i - \bar{y}_i)^2, & \text{if } |y_i - \bar{y}_i| \le \delta. \\ \delta(|y_i - \bar{y}_i| - \frac{1}{2}\delta), & \text{if } |y_i - \bar{y}_i| > \delta. \end{cases} \tag{1}$$

6. **Log Cosh Loss:**

$$E = \tfrac{1}{m}\textstyle\sum_{i=1}^{m} \log\left(\cosh\left(y_i - \bar{y}_i\right)\right)$$

Where,

$$\cosh(x) = \tfrac{e^x + e^{-x}}{2}$$