CS698 - Winter 2017

Assignment 1 - Q3a

January 19, 2017

1 Problem Statement

Theory. In class, we discussed several loss functions for linear regression. However all the loss functions that we discussed assume that the error contributed by each data point have the same importance. Consider a scenario where we would like to give more weight to some data points. Our goal is to fit the data points (x_n, y_n) in proportion to their weights r_n by minimizing the following objective:

$$L(w, b) = \sum_{n=1}^{m} r_n(y_n w x_n + b)^2$$

2 Solution

2.1 Calculation of the partial derivatives

Since the loss function needs to be minimized with respect to 2 variables, the partial derivative of the loss function L with respect to each variable is calculated independently.

$$\begin{aligned}
\frac{\partial L}{\partial w} &= 0 \\
2 * \sum_{n=1}^{m} r_n(y_n w x_n + b) * (x_n) &= 0 \\
\sum_{n=1}^{m} r_n(y_n w x_n + b) * (x_n) &= 0 \\
\sum_{n=1}^{m} r_n(y_n x_n w x_n x_n^T + b x_n) &= 0 \\
\sum_{n=1}^{m} r_n y_n x_n r_n w x_n x_n^T + r_n b x_n &= 0 \\
\sum_{n=1}^{m} r_n y_n x_n w \sum_{n=1}^{m} r_n x_n x_n^T + b \sum_{n=1}^{m} r_n x_n &= 0 \\
w \sum_{n=1}^{m} r_n x_n x_n^T &= \sum_{n=1}^{m} r_n y_n x_n + b \sum_{n=1}^{m} r_n x_n
\end{aligned}$$

$$\boxed{w = \frac{b \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n}{\sum_{n=1}^{m} r_n x_n x_n^T}} \tag{1}$$

Also,

$$\frac{\partial L}{\partial b} = 0$$

$$2 * \sum_{n=1}^{m} r_n(y_n w x_n + b) = 0$$

$$\sum_{n=1}^{m} r_n(y_n w x_n + b) = 0$$

$$\sum_{n=1}^{m} r_n(y_n w x_n + b) = 0$$

$$\sum_{n=1}^{m} r_n y_n r_n w x_n + r_n b = 0$$

$$\sum_{n=1}^{m} r_n y_n w \sum_{n=1}^{m} r_n x_n + b \sum_{n=1}^{m} r_n = 0$$

$$b \sum_{n=1}^{m} r_n = w \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n$$

$$\boxed{b = \frac{w \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n}{\sum_{n=1}^{m} r_n}} \tag{2}$$

2.2 Solution of the system of linear equations

The value of b, as obtained in Equation 2 is substituted in Equation 1.

$$w = \frac{\frac{w \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n}{\sum_{n=1}^{m} r_n} \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n}{\sum_{n=1}^{m} r_n x_n x_n^T}$$

$$w = \frac{(w \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n) \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n}$$

$$w = \frac{w (\sum_{n=1}^{m} r_n x_n)^2 - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n}$$

$$w \sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n = w(\sum_{n=1}^{m} r_n x_n)^2 - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n$$

$$w \sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - w(\sum_{n=1}^{m} r_n x_n)^2 = - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n$$

$$w(\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - (\sum_{n=1}^{m} r_n x_n)^2) = - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n + \sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n$$

$$w = \frac{\sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - (\sum_{n=1}^{m} r_n x_n)^2} \tag{3}$$

Now that the value of w has been expressed in terms of r_n, x_n and y_n, we can substitute this value of w in Equation 2 to obtain a closed form expression for b.

$$b = \frac{w \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n}{\sum_{n=1}^{m} r_n}$$

$$b\sum_{n=1}^{m} r_n = w\sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n$$

$$b\sum_{n=1}^{m} r_n = \frac{\sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - (\sum_{n=1}^{m} r_n x_n)^2} \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n$$

$$b\sum_{n=1}^{m} r_n = \frac{\sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n - \sum_{n=1}^{m} r_n y_n \sum_{n=1}^{m} r_n x_n}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - (\sum_{n=1}^{m} r_n x_n)^2} \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n$$

$$b\sum_{n=1}^{m} r_n = \frac{\sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n (\sum_{n=1}^{m} r_n x_n)^2}{\sum_{n=1}^{m} r_n x_n x_n^T \sum_{n=1}^{m} r_n - (\sum_{n=1}^{m} r_n x_n)^2} - \sum_{n=1}^{m} r_n y_n$$

$$b = \frac{\sum_{n=1}^{m} r_n y_n x_n \sum_{n=1}^{m} r_n \sum_{n=1}^{m} r_n x_n - \sum_{n=1}^{m} r_n y_n (\sum_{n=1}^{m} r_n x_n)^2}{\sum_{n=1}^{m} r_n x_n x_n^T (\sum_{n=1}^{m} r_n)^2 - \sum_{n=1}^{m} r_n (\sum_{n=1}^{m} r_n x_n)^2} - \frac{\sum_{n=1}^{m} r_n y_n}{\sum_{n=1}^{m} r_n} \tag{4}$$

3 Conclusion

Equations 3 and 4 are the closed form solutions for minimizing the loss function L, in terms of the input variable x, the target variable w, and the local weight r.