Labwork3's report

Hai Nguyen Ngoc

May 3 2024

1 Introduction

Logistic regression is method for classification task. We calculate the predicted value by:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{1}$$

With $z = (w_1 \times x_i^{(1)} + w_2 \times x_i^{(2)} + w_0)$. Because we have 2 possible outputs, so that the error (or loss) functions can be calculated by:

$$L_i = \begin{cases} -\log(y_i) & \text{if } y_i = 1 \\ -\log(1 - y_i) & \text{if } y_i = 0 \end{cases} \tag{2}$$

And finally we have combination of above equation for 1 data point:

$$L_i = -(y_i \cdot \log(y_i) + (1 - y_i) \cdot \log(1 - y_i)) \tag{3}$$

For n data points, we have Loss function:

$$J = -\frac{1}{N} \sum_{i=1}^{N} (y_i \cdot \log(y_i') + (1 - y_i) \cdot \log(1 - y_i')) \tag{4}$$

2 Implementation

2.1 Functions

- First, we need to write predicted value as equation (1)
- Second, write function for loss function as equation (4)
- Functions for gradient descent as below equations:

$$\frac{\partial L_i}{\partial w_0} = \left(\frac{y_i}{y_i'} + \frac{-1 + y_i}{1 - y_i'}\right) \cdot \left[-e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)} / \left(1 + e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)}\right)^2\right] \tag{5}$$

$$\frac{\partial L_i}{\partial w_1} = \left(\frac{y_i}{y_i'} + \frac{-1+y_i}{1-y_i'}\right) \cdot \left[-e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)} / \left(1 + e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)}\right)^2\right] \cdot x_i^{(1)} \tag{6}$$

$$\frac{\partial L_i}{\partial w_2} = \left(\frac{y_i}{y_i'} + \frac{-1+y_i}{1-y_i'}\right) \cdot \left[-e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)} / \left(1 + e^{-(w_1 x_i^{(1)} + w_2 x_i^{(2)} + w_0)}\right)^2\right] \cdot x_i^{(2)} \tag{7}$$

- Write function to optimize parameters by Gradient Descent:

$$w_0 = w_0 - \alpha * \frac{\partial L_i}{\partial w_0} \tag{8}$$

$$w_1 = w_1 - \alpha * \frac{\partial L_i}{\partial w_1} \tag{9}$$

$$w_2 = w_2 - \alpha * \frac{\partial L_i}{\partial w_1} \tag{10}$$

2.2 Main

Create initial value for w_0, w_1, W_2: $w_0 = 1$, $w_1 = 2$, $w_2 = 0$, Then loop until reach expected value or max number of iterations.

3 Evaluation

4 Conclusion

- Logistic regression work well in binary classification.
- Easy to implement and need to choose appropriate value of learning rate
- Need to do carefully in partial derivative of w_0, w_1, w_2

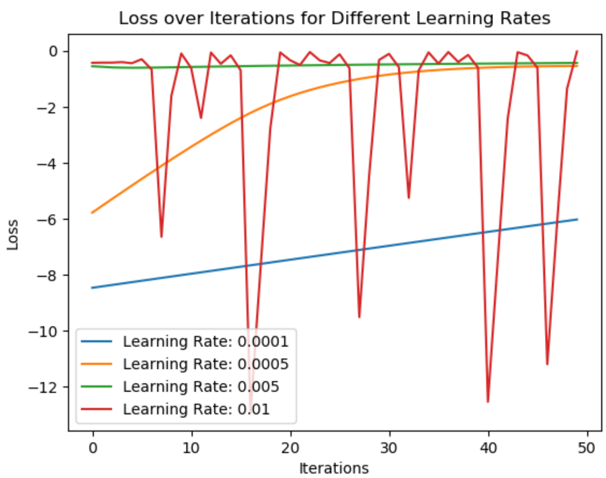

Figure 1: Value of Loss function through iterations with different values of learning rate