# ENGR 421 / Homework 2: Discrimination by Regression

Umur Berkay Karakaş

March 24, 2021

In homework 2, we are given 1000 images which consist of 784 pixels where each pixel is continuous between 0 and 1.

First, I created trainingset variable from the first 500 images in the data set and testset variable from the last 500 points in the data set. I assigned their corresponding class labels to ytraining and ytest. I also calculated number of classes by using np.max.

I defined a sigmoid function which takes a data matrix, W matrix and w0 vector as parameters and returns the sigmoid calculated by those inputs. I also defined two other functions to calculate gradient of W and gradient of w0 in each iteration while minimizing the error. Gradient functions are from Section 10.8 called discrimination by regression of our textbook.

$$\text{sigmoid}(\boldsymbol{W}^\top \boldsymbol{X} + w_0) = \frac{1}{1 + \exp\left[-(\boldsymbol{W}^\top \boldsymbol{X} + w_0)\right]}$$

$$\begin{aligned} \Delta \boldsymbol{w}_i &= \eta \sum_t (r_i^t - y_i^t) y_i^t (1 - y_i^t) \boldsymbol{x}^t \\ \Delta w_{i0} &= \eta \sum_t (r_i^t - y_i^t) y_i^t (1 - y_i^t) \end{aligned}$$

Figure 1: Update functions for gradients

$$E(\{\boldsymbol{w}_i, w_{i0}\}_i | \mathcal{X}) = \frac{1}{2} \sum_t \|\boldsymbol{r}^t - \boldsymbol{y}^t\|^2 = \frac{1}{2} \sum_t \sum_i (r_i^t - y_i^t)^2$$

Figure 2: Error function

By using given eta, epsilon and max iteration values, I did the learning iteration using gradient descent to minimize the error whose function is given above. After the iteration is over, I used final W and w0 values to calculate sigmoid of trainingset and testset and the resulting confusion matrices for both are:

| y_train | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| y_predicted | | | | | |
| 1 | 101 | 2 | 11 | 0 | 1 |
| 2 | 0 | 79 | 1 | 0 | 0 |
| 3 | 6 | 5 | 100 | 0 | 1 |
| 4 | 0 | 0 | 0 | 99 | 2 |
| 5 | 0 | 0 | 0 | 1 | 91 |

Figure 3: Confusion matrix for training set

| y_test | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| y_predicted | | | | | |
| 1 | 82 | 1 | 4 | 0 | 2 |
| 2 | 0 | 87 | 0 | 0 | 0 |
| 3 | 13 | 3 | 98 | 0 | 3 |
| 4 | 0 | 0 | 0 | 103 | 1 |
| 5 | 4 | 0 | 3 | 0 | 96 |

Figure 4: Confusion matrix for test set