EE550 Project-4

Halil İbrahim Güllük[1]

[1]Department of Mathematics, Boğaziçi University

April 28, 2020

1 Introduction

In this project, as programing tools PyTorch,Numpy and Matplotlib libraries are used. Numpy is used in order to generate random numbers, while PyTorch is used for matrix manipulations.Also,Matplotlib is used for plotting data points. Multi Layer Perceptron algorithm is implemented as it is covered in class in order to function approximations and classification problems.

2 Methodology

As we covered in lectures we denote the error function by$E$, value of ith row and jth column of kth weight matrix by w_{ij}^k, value of jth node of kth layer by s_j^k, output value of jth node of kth layer after nonlinear function by o_j^k, jth node value of real output values by t_j.Then as it is covered in class the following equations are going to be used.

$$E = \frac{1}{2}\sum_j (t_j - o_j^L)^2$$

$$\frac{\partial E}{\partial w_{ij}^k} = \frac{\partial E}{\partial s_j^{k+1}} \frac{\partial s_j^{k+1}}{\partial w_{ij}^k} = \frac{\partial E}{\partial s_j^{k+1}} o_i^k$$

Also we define a new variable δ_j^k such that,

$$\delta_j^k = -\frac{\partial E}{\partial s_j^{k+1}}$$

Then as it is covered in class by backpropagation we get the following equation,

$$\delta_j^k = \sum_l \delta_l^{k+1} w_{jl}^k f'(s_j^k)$$

After this backpropagation, our weights update rule is,

$$\Delta w_{ij}^k = \eta \delta_j^{k+1} o_i^k$$

,where η is the learning rate.

3 Problem 1

For XOR problem it is decided to have 3 layers with 2,10,1 number of nodes respectively.As nonlinear function, sigmoid is used. So our outputs will be between 0 and 1. For training we have 4 samples (1,1),(1,0),(0,1),(0,0) and their labels 0,1,1,0 respectively. The error function vs number of epoch graph is provided below.

And for (1,1),(1,0),(0,1),(0,0) our model predicts 0.0727, 0.9230, 0.9228, 0.0796 respectively.

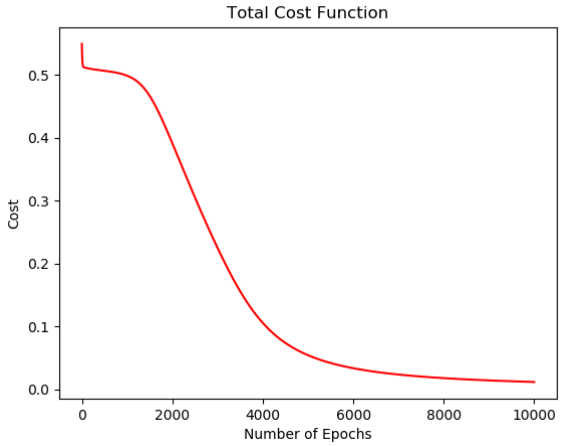

Figure 1: Error Function Graph

4 Problem 2

In sin(x) problem 1,10,7,6,1 numbers of nodes in layers are used respectively. However, as it is expected to have output between -1 and 1, tanh nonlinear function is used. The error function vs number of epoch graph is provided below.

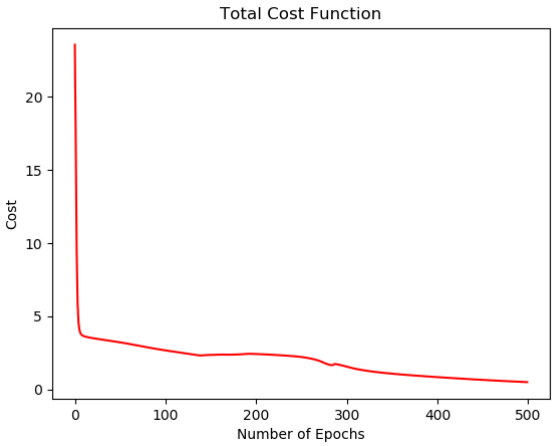

Figure 2: Error Function Graph

In order to evaluate whether we have a good approximation or not we plot sin(x) function and 25 test points results, which is provided below.

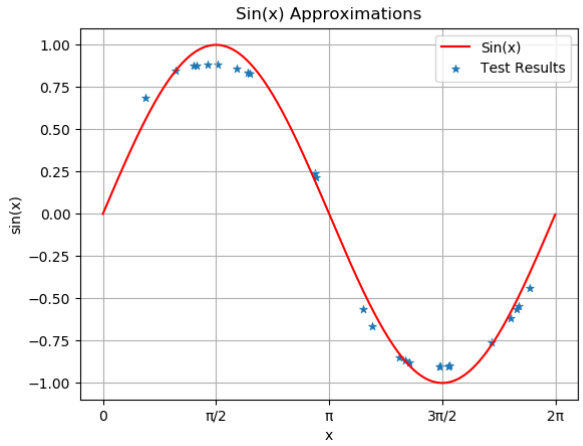

Figure 3: Appraoximation of sin(x)

5 Problem 3

In image classification problem 64,40,35,20,10 numbers of nodes in layers are used respectively.As we have vectors of length 64 as inputs and 10 different classes, in the first layer 64 nodes and in the output layer we have 10 nodes representing the classes. As nonlinear function sigmoid is used. Node with maximum value among all other output nodes will be its class.The error function vs number of epoch graph is provided below.

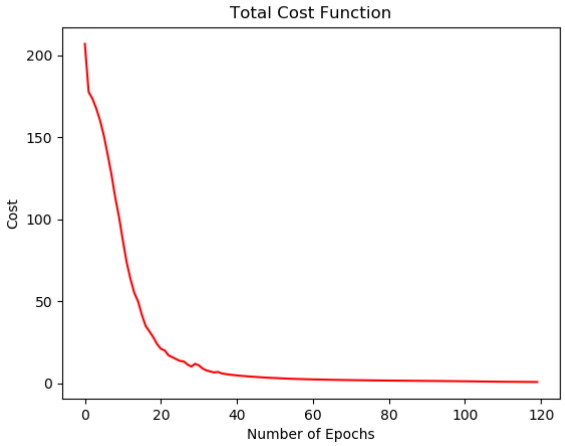

Figure 4: Error Function Graph

When we do classification by taking nodes with maximum values of the output layer, we get the true positive accuracy(TP) %83.In other words our model classified 83 out of 100 test pattern correctly.