Assignment 2
DD2424 Oscar Azrak

1. Introduction

This assignment aimed to develop, train and evaluate a two-layer neural network for image classification using the CIFAR-10 dataset. The network was trained using mini-batch gradient descent which was applied to a cost function that consisted of the cross-entropy loss and an L2 regularization term on the weight matrices. The network's performance was optimized by fine tuning the regularization parameter and utilizing a cyclical learning rate. This allowed me to achieve best test accuracy for this assignment.

2. Method

1. Download and extract the contents of CIFAR-10 dataset
2. Load the dataset and split the data to train, validation and test sets.
3. Normalize the relevant datasets.
4. Initialize model parameters
5. Implement relevant functions
6. Compute and compare the numerical and analytical gradients
7. Train the model using default settings
8. Find the best parameters
9. Evaluate the model on the test set

3. Results

The central difference, h of 10^{-5}, was used where these errors were observed:

Relative error W1	Relative error b1	Relative error W2	Relative error b2
$1.66*10^{-5}$	$2.66*10^{-8}$	$5.4*10^{-10}$	$4.64*10^{-9}$

The relative errors were obtained by comparing the difference of the analytical and numerical gradients for each variable. The fact that the errors were small indicates high accuracy of the gradient calculations. It is important to note that this is not an absolute guarantee, but it provides a strong indication of the implementation's accuracy.

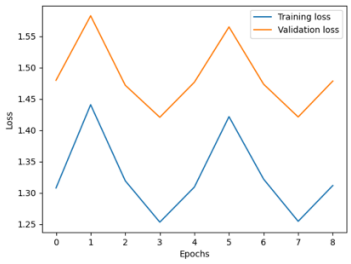
The range for retrieving best lambdas for the coarse search used values generated on a log space in the range of 10^{-5} to 10^{-1}.

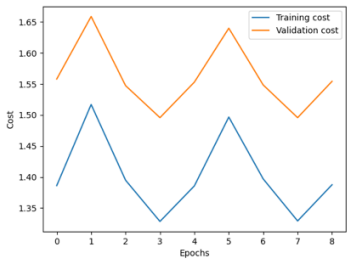
Lambda	Accuracy
0.00464	51.62%
$2.783*10^{-5}$	50.72%
$1*10^{-5}$	50.55%

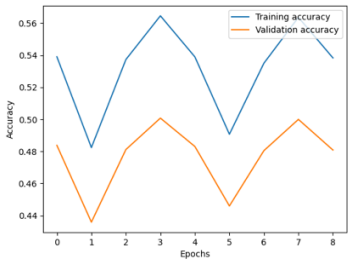
The range for retrieving the best lambdas for the fine search used values generated using the best performing lambda.

Lambda	Accuracy
0.00438	50.96%
0. 004177	50.94%
0.00459	50.84%

Plots of the loss, costs and accuracy:







Both the losses and costs have a down going trend while the training has an up going trend. The best accuracy for the coarse search was better than the best for the fine search but all of the fine searches were better than the second and third best coarse search accuracies.