Assignment 1

Alexander Falk

DD2424, May 2024

Exercise 1) I have conducted 4 experiments in compliance with the instructions from Assignment 1:

Experiment 1: lambda=0, n epochs=40, n batch=100, eta=.1
Experiment 2: lambda=0, n epochs=40, n batch=100, eta=.001
Experiment 3: lambda=.1, n epochs=40, n batch=100, eta=.001
Experiment 4: lambda=1, n epochs=40, n batch=100, eta=.001

To check if my analytical gradient computations were correct I compared the optimal weight and bias values calculated from my functions with the numerically calculated weights and biases using the centered difference method (Equation 1) on a subset of the available data. The subset was 20 images, lambda was set to 1 and value of h set to 0.0001. Table 1 shows the results. Relative errors are very small, which indicates that my analytical gradient computation is correct.

$$\frac{|g_a - g_n|}{\max(\text{eps}, |g_a| + |g_n|)}$$

Equation 1. Centered difference method

Relative error W	Relative error b
1.8275e-10	8.5772e-10

Table 1. Relative errors between weights and biases for a subset of the available data

The generated plots from the 4 aforementioned experiments can be found in Figure 1-6 below. For the first 2 experiments, lambda was set to 0 so I only included the loss graph in these instances due to loss and cost function being the same. The test accuracies for each experiment can be found in Table 2.

Studying the test accuracies as well as the graphs generated we can see the importance of having a correct learning rate set and also the effect an increasing amount of regularization has on the accuracy of the model. With eta set too high, the model has a hard time improving over the epochs (Figure 1). With lambda set to 0, there is no regularization. We can see that the training loss is steadily decreasing, but after around 20 epochs the validation loss is almost constant. This could be a sign of overfitting due to the lack of regularization (Figure 2).

	Test Accuracy
Experiment 1	29.91 %
Experiment 2	39.40 %
Experiment 3	39.06 %
Experiment 4	37.52 %

Table 2. Test accuracies for the different experiments

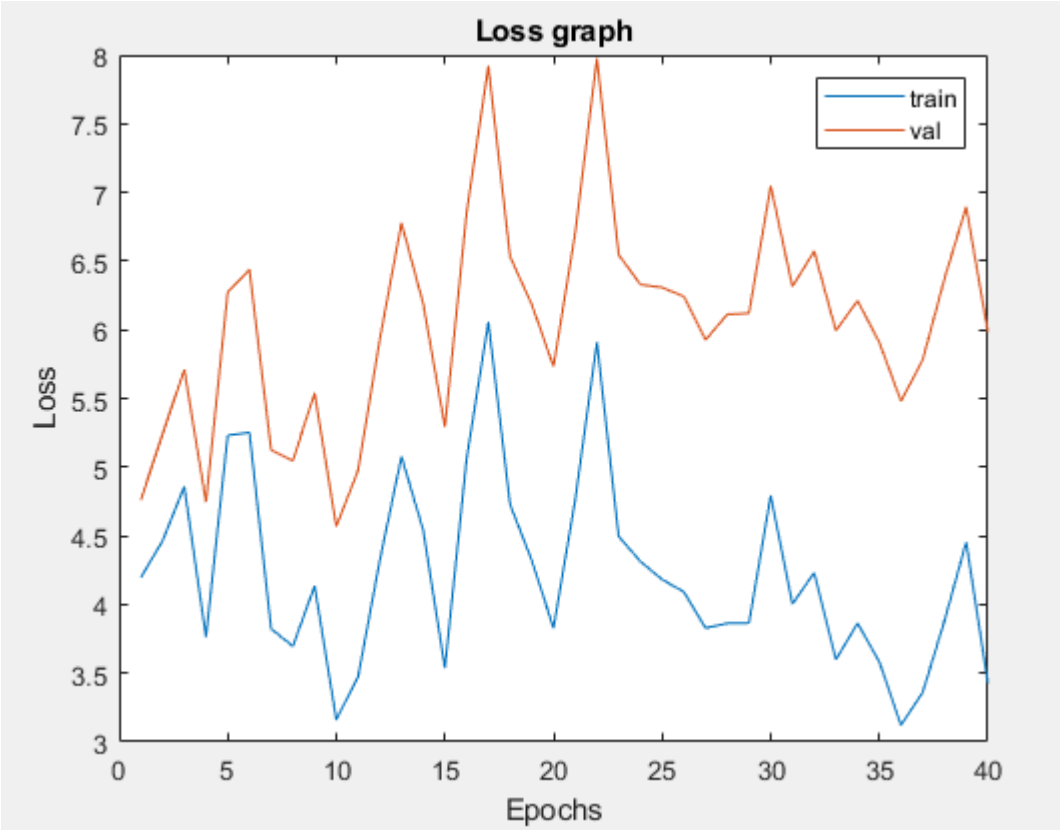

Figure 1. Training and validation cost/loss graph for experiment 1.

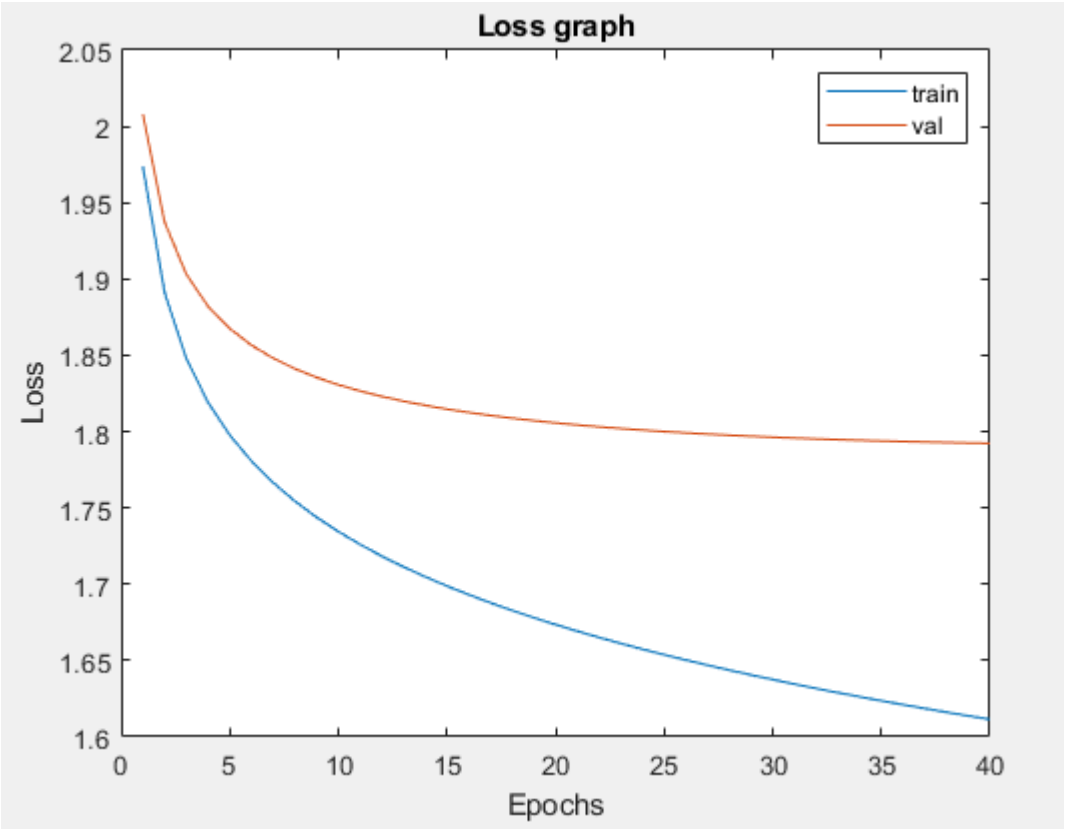

Figure 2. Training and validation cost/loss graph for experiment 2.

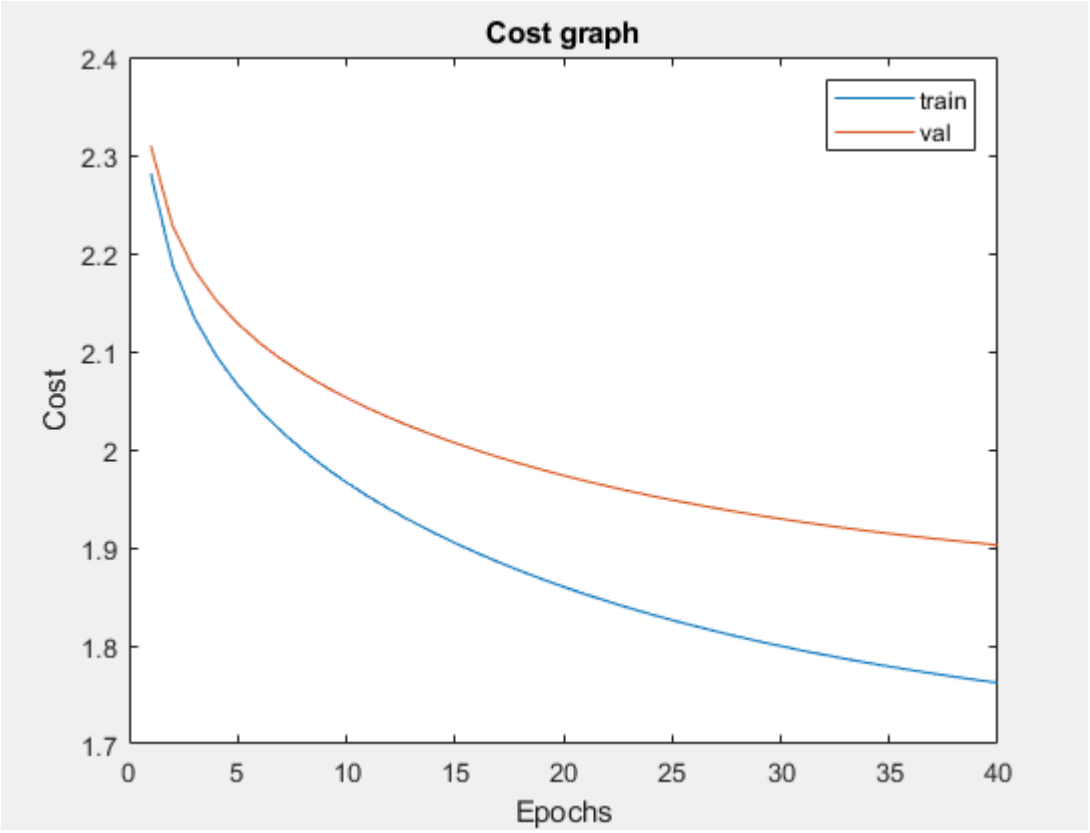

Figure 3. Training and validation cost graph for experiment 3.

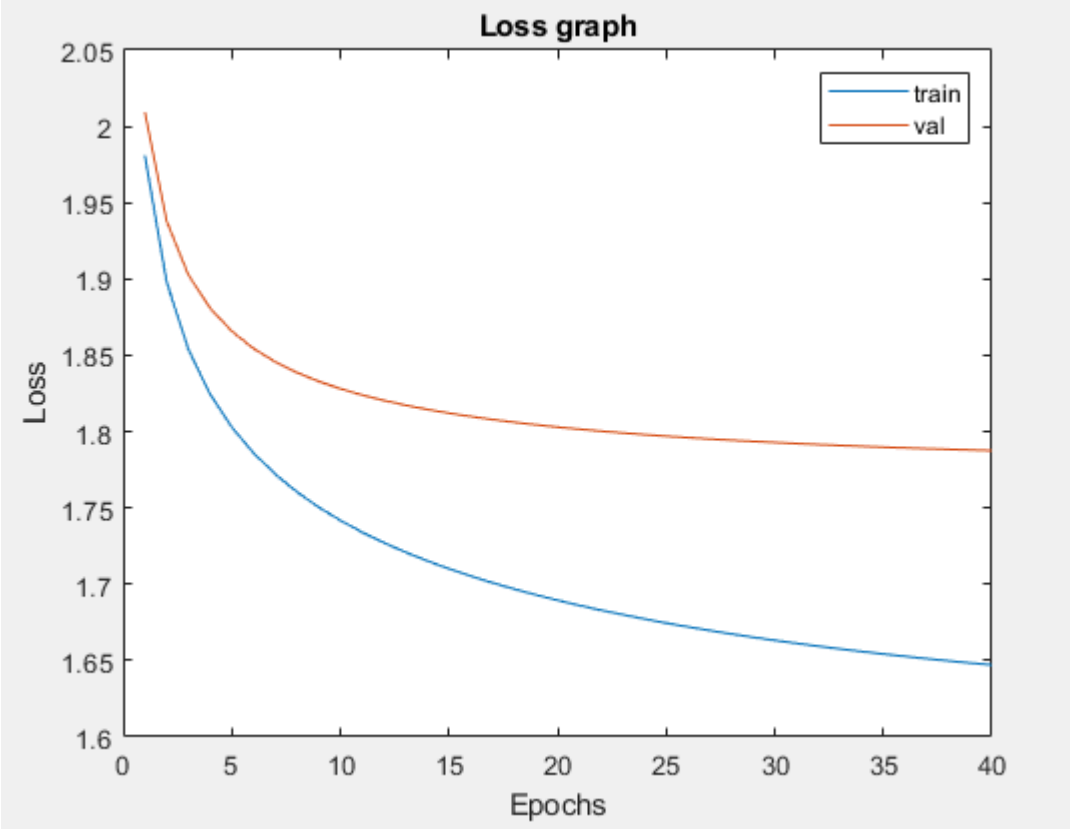

Figure 4. Training and validation loss graph for experiment 3.

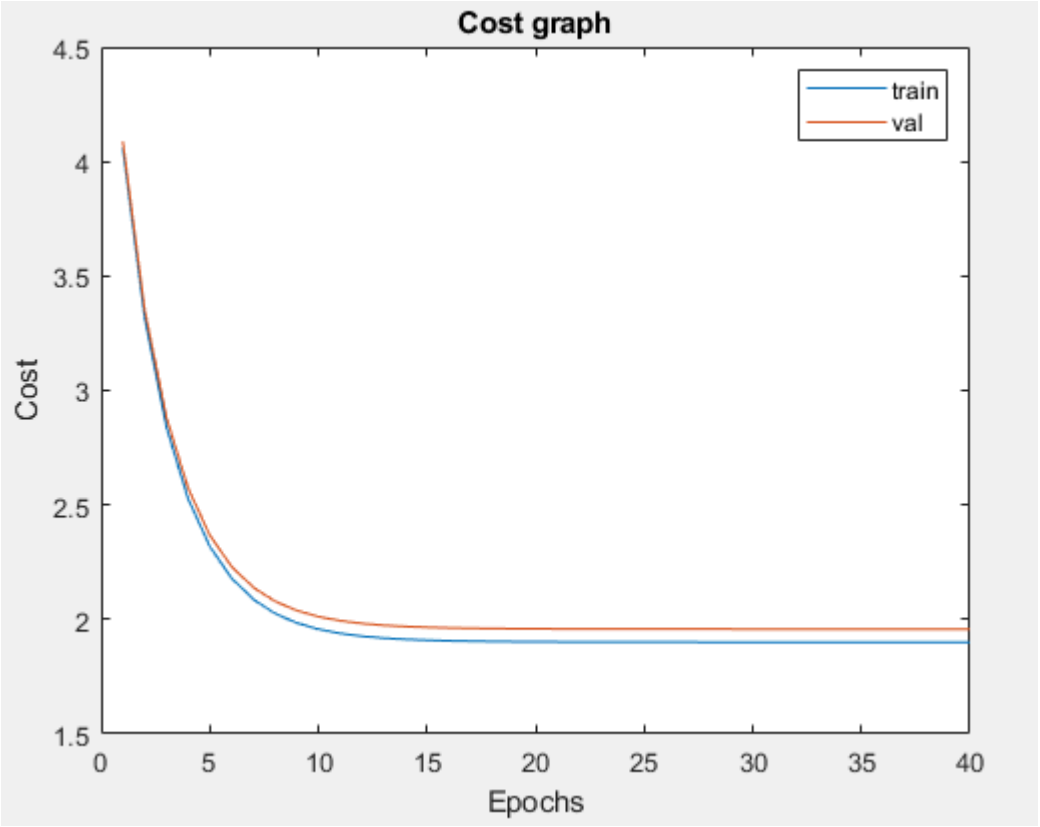

Figure 5. Training and validation cost graph for experiment 4.

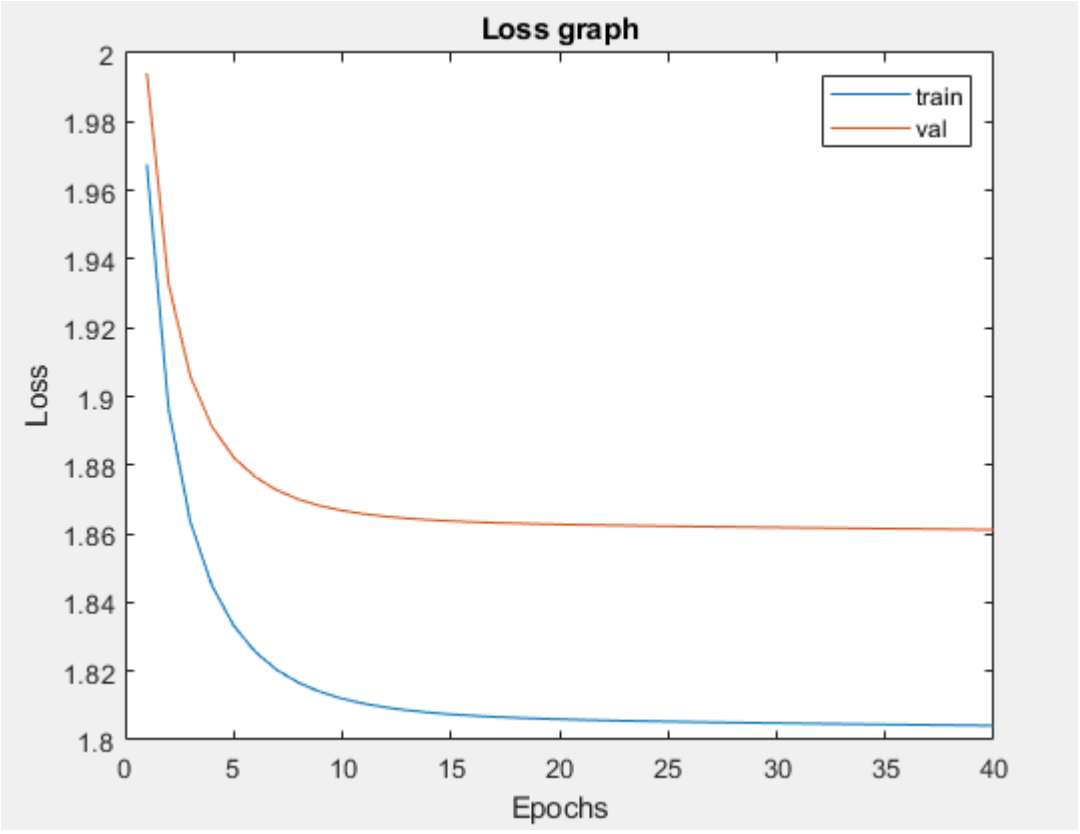

Figure 6. Training and validation loss graph for experiment 4.

In Figures 7-10 the learnt weight matrix for each experimental model is displayed.

Figure 7. Visualization of the weight matrix for experiment 1.

Figure 8. Visualization of the weight matrix for experiment 2.

Figure 9. Visualization of the weight matrix for experiment 3.

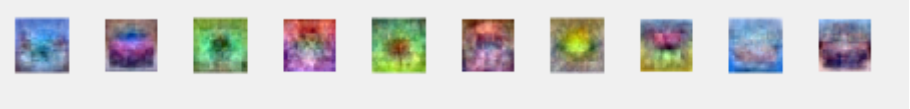

Figure 10. Visualization of the weight matrix for experiment 4.