Χριστίνα Μάη 4612

Ζήτημα 4.1

```
Running numeric gradient check for fc
  grad_x difference:  2.0078061435668815e-10
  grad_w difference:  1.6502310629107342e-10
  grad_b difference:  1.203377397729355e-10
Running numeric gradient check for relu
  grad_x difference:  5.855738116622433e-11
Running numeric gradient check for softmax loss
  grad_x difference:  1.8800439156863025e-10
Running numeric gradient check for L2 regularization
  grad_w difference:  3.4616209204640924e-11
```
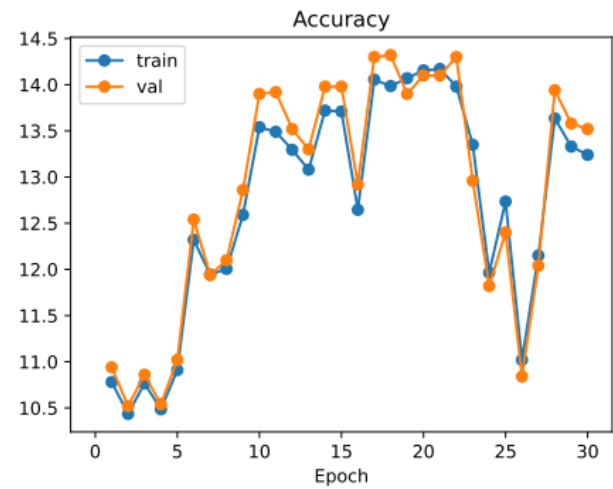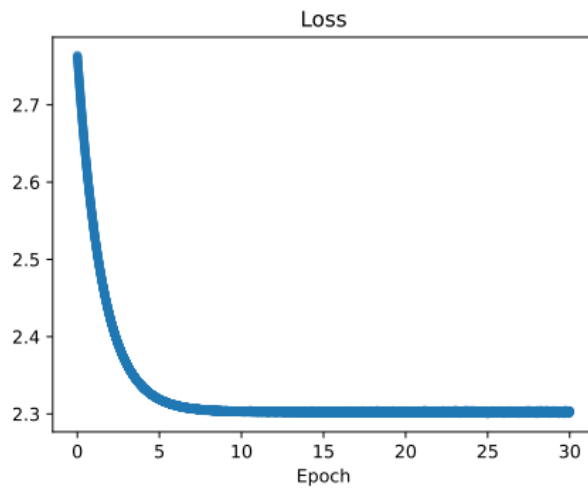
Παρατηρούμε ότι όλες οι διαφορές είναι μικρότερες του 10^{-9}

Ζήτημα 4.2

```
Running numeric gradient check for LinearClassifier
  Max diff for grad_W:  1.538769112130467e-13
  Max diff for grad_b:  8.970602038971265e-14
Running numeric gradient check for TwoLayerNet
  Max diff for grad_W1:  9.697104230710352e-16
  Max diff for grad_b1:  2.8362728832220796e-16
  Max diff for grad_W2:  5.412337245047638e-16
  Max diff for grad_b2:  8.881784197001252e-16
```

Παρατηρούμε ότι όλες οι διαφορές είναι μικρότερες του 10^{-10}

Ζήτημα 4.3

Με τις τιμές των παραμέτρων:

num_train = 20000

hidden_dim = 300

batch_size = 60

num_epochs=30

learning_rate = 1e-3

reg=1

To train=13.24

To val=13.52