CS 547 Homework 1

Name: Jiashuo Tong
Date Submitted: Sep., 3, 2019

Neural Network Implementation

Function Definitions

1. def loadData(url, nimage): loadData reads a image and formats it into a 28x28 long array
2. def loadLabels(url, nimage): loadLabels reads the corresponding label data, one for each image
3. def download(dataurl, labelsurl, nimage): download uses loadData and loadLabels to generate raw data to be used later
4. def relu(M): relu is an activation function defined as 0 for x < 0 and x for x > 0
5. def relu_p(M): relu_p is the derivative of relu function
6. def sigmoid(M): sigmoid is an activation function defined as exp(x)/(exp(x)+1)
7. def sigmoid_p(M): sigmoid_p is the derivative of sigmoid function
8. def softmax(z): softmax is a function that converts the raw scores for each category to probabilities that sum up to 1
9. def feed_forward(X,Y): feed_forward feeds MNIST data to the neural network and computes values at each intermediate layers.
10. def back_propagate(Z,H,U,f,X,Y): back_propagate uses the results from feed_forward to compute the gradients of the loss to each parameter of the neural network.
11. def predict(probability, labels): predict function predicts the categories for a dataset using the trained neural network and calculate the accuracy on the dataset by comparing the results to the labels.

Main Block

The main block starts by downloading data from LeCun's website. Then, it prepares the data so they can be fed into the neural network. For the labels data, It convert the original labels (0-9) to vectors containing 1's and 0's (one-hot encoding). Then, it sets trainable parameters "W", "b1", "C" and "b2" for the neural network that contains one hidden layer. The matrices "W" and "C" are initialized to random numbers spanning from -sqrt(1/784) to + sqrt(1/784), while the bias "b1" and "b2" are set to zeros. At last comes the training code which sets all the tunable hyperparameters. The learning rate is initialized to 1e-4, and will reduce to 1e-5 at epoch 8. Since the optimizing scheme is Stochastic Gradient Descent, the batch size is set to 1. Thus, there will be 60000 iterations in each epoch. For each iteration a feed forward and back propagation operations are done through the network with a hidden layer of size 100. A total of 20 epochs are done on the training set. The accuracy reached 97% at the 10^{th} epoch, and finally settled down at around 97.18%. Please see the training log attached below for more details.

Results

I have achieved a test accuracy of > 97% at the 10th epoch. The accuracy for the rest epochs stays above 97%.

```
Epoch  1 / 21  is complete with a loss of  5.421095825220706e-06
   Accuracy on the training set is  0.9240166666666667
   Accuracy on the test set is  0.9214
   Current learning rate is  0.0001
Epoch  2 / 21  is complete with a loss of  1.8044077224836576e-06
   Accuracy on the training set is  0.9467833333333333
   Accuracy on the test set is  0.9414
   Current learning rate is  0.0001
Epoch  3 / 21  is complete with a loss of  2.8969534326487525e-07
   Accuracy on the training set is  0.9511333333333334
   Accuracy on the test set is  0.9412
   Current learning rate is  0.0001
Epoch  4 / 21  is complete with a loss of  1.9576114201149124e-05
   Accuracy on the training set is  0.954
   Accuracy on the test set is  0.9426
   Current learning rate is  0.0001
Epoch  5 / 21  is complete with a loss of  2.2403207092351586e-09
   Accuracy on the training set is  0.9648833333333333
   Accuracy on the test set is  0.9529
   Current learning rate is  0.0001
Epoch  6 / 21  is complete with a loss of  5.750747657506261e-10
   Accuracy on the training set is  0.96735
   Accuracy on the test set is  0.9543
   Current learning rate is  0.0001
Epoch  7 / 21  is complete with a loss of  1.3813777908870717e-09
   Accuracy on the training set is  0.9707333333333333
   Accuracy on the test set is  0.9569
   Current learning rate is  0.0001
Epoch  8 / 21  is complete with a loss of  4.24993373827413e-12
   Accuracy on the training set is  0.9752333333333333
   Accuracy on the test set is  0.9617
   Current learning rate is  0.0001
Epoch  9 / 21  is complete with a loss of  7.426947945560027e-12
   Accuracy on the training set is  0.9866666666666667
   Accuracy on the test set is  0.9697
   Current learning rate is  1e-05
Epoch  10 / 21  is complete with a loss of  1.8806622925814685e-11
   Accuracy on the training set is  0.98865
   Accuracy on the test set is  0.9708
   Current learning rate is  1e-05
```

```
Epoch  11 / 21  is complete with a loss of  1.5267453967854314e-11
   Accuracy on the training set is  0.9896666666666667
   Accuracy on the test set is  0.9705
   Current learning rate is  1e-05
Epoch  12 / 21  is complete with a loss of  1.1608269900943088e-11
   Accuracy on the training set is  0.99035
   Accuracy on the test set is  0.9709
   Current learning rate is  1e-05
Epoch  13 / 21  is complete with a loss of  8.93862761590201e-12
   Accuracy on the training set is  0.99085
   Accuracy on the test set is  0.9712
   Current learning rate is  1e-05
Epoch  14 / 21  is complete with a loss of  6.4291905133224864e-12
   Accuracy on the training set is  0.99135
   Accuracy on the test set is  0.9715
   Current learning rate is  1e-05
Epoch  15 / 21  is complete with a loss of  5.415778936438642e-12
   Accuracy on the training set is  0.9917833333333334
   Accuracy on the test set is  0.9716
   Current learning rate is  1e-05
Epoch  16 / 21  is complete with a loss of  4.154676602760891e-12
   Accuracy on the training set is  0.9920166666666667
   Accuracy on the test set is  0.9718
   Current learning rate is  1e-05
Epoch  17 / 21  is complete with a loss of  2.7792212975480164e-12
   Accuracy on the training set is  0.9924666666666667
   Accuracy on the test set is  0.9717
   Current learning rate is  1e-05
Epoch  18 / 21  is complete with a loss of  2.159050715990873e-12
   Accuracy on the training set is  0.9928166666666667
   Accuracy on the test set is  0.9718
   Current learning rate is  1e-05
Epoch  19 / 21  is complete with a loss of  1.7609247393595111e-12
   Accuracy on the training set is  0.9932666666666666
   Accuracy on the test set is  0.9718
   Current learning rate is  1e-05
Epoch  20 / 21  is complete with a loss of  1.3391510123037605e-12
   Accuracy on the training set is  0.99365
   Accuracy on the test set is  0.9719
   Current learning rate is  1e-05
```