Instructions:

In this homework you will practice putting together a simple image classification pipeline to classify **handwritten digit 0 and 1**, based on the logistic regression or the SVM classifier. Please keep the batch size untouched. The goals of this homework are as follows:

- understand the basic Image classification pipeline and the data-driven approach (train/predict stages)

- understand how to use pytorch to build binary classifiers.

- understand how to exploit pytorch's autograd mechanics to do optimization.

- implement and apply a logistic regression classifier.

- implement and apply a binary Support Vector Machine (SVM) classifier.

Problem Description.

In this homework, you are asked to implement and solve linear Logistic Regression model and Linear SVM model (without regularization term) on MNIST dataset. In this task, you only need to perform binary classification on digit 0 and 1. Details of these models could be found in lecture 2 slides. We provide a skeleton code for data loading and iterations of training data. You are asked to implement the rest of training in Pytorch code. You are required to optimize the model by using SGD and Momentum methods. Detailed submission requirements are written in the final section.

Resources

You can follow the setup instructions at `https://pytorch.org/get-started/locally/`.
A useful tutorial on learning pytorch by examples at `https://pytorch.org/tutorials/beginner/pytorch_with_examples.html`.
More illustrations of different optimizers could be found here: `http://ruder.io/optimizing-gradient-descent/`.

Data

We use MNIST digit classification dataset. Pytorch/torchvision has provide a useful dataloader to automatically download and load the data into batches. In this homework, we need two class, digit 0 and digit 1, for binary classification. We have written the data loader for you as follow. You can find it in the attached file.

Submitting your assignment

Upload a PDF report using LATEX containing your experiment results to CCLE (remember to include your name and ID number in the report). Some requirements for your report:

- Limit the number of pages to 2.

- For each of the model, report the $(\sum_{b=1}^{B} \sum_{d=1}^{D_b} \frac{loss(label_{b,d}, f_b(data_{b,d}))}{D_b})/B$ for each training epoch, where B is the total number of batches, f_b is the model (Logistic regression or Linear SVM) after updated by b-th batch and D_b is the number of data points in b-th batch. An epoch is defined as one iteration of all dataset. Essentially, during a training epoch, you record down the average training loss of that batch after you update the model, and then report the average of all such batch-averaged losses after one iteration of whole dataset. You could plot the results as a figure our simply list down. Please at least report 10 epochs.

- Report the final testing accuracy of trained model.

- Please compare results for 2 optimizer (SGD and SGD-Momentum)).

- Try different step sizes and discuss your findings.

Also, upload your code in a zip file and show how to run your code in README.