

Princeton COS 495: Introduction to Deep Learning

Homework 4

Lecturer: Yingyu Liang
TA: Bochao Wang

Due date: April 13th, 2016
Office: Electrical Engineering Department, C319B

1. (Math) Consider how to find the first principal component v for data points x_1, x_2, \dots, x_n . One can maximize the variance of the projections (suppose the data have mean $\frac{1}{n} \sum_{i=1}^n x_i = 0$):

$$\max_{v: \|v\|_2=1} \sum_{i=1}^n (v^\top x_i)^2.$$

One can also minimize the mean square error of the projections:

$$\min_{v: \|v\|_2=1} \frac{1}{n} \sum_{i=1}^n \|x_i - vv^\top x_i\|_2^2.$$

Show that the optimal solution to the two is the same. Hint: see the intuition on slide 18 of “Deep Learning Basics Lecture 7: factor analysis”.

2. (Coding) Add convolutional layers to the network in Problem 4 in homework 3, and compare the results with those without convolution.

Comments: 1) The number of convolutional layers can be 1/2/3. You only need to test one of them. Check the tutorials for some of the parameters like number of feature maps to use (e.g., <https://github.com/tensorflow/tensorflow/blob/master/tensorflow/models/image/mnist/convolutional.py>).

2) Early stopping may be a bit hard to tune. So in this experiment, you don't need to apply early stopping. Just need to plot the training loss and the validation error.