Chennai Mathematical Institute
Computer Vision

Homework Assignment: Lectures 19-21

Note: Questions one and two are from Chapter 1 of the book "Introduction to Deep Learning" by Eugene Charniak @ https://cs.brown.edu/courses/csci1460/assets/files/deep-learning.pdf

1) This question is in reference to the stochastic gradient descent algorithm for digit recognition given in Figure 1.10, Page 20 (see below). The algorithm uses the cross-entropy loss function, and the softmax operator for determining class probabilities.

1. for j from 0 to 9 set b_j randomly (but close to zero)
2. for j from 0 to 9 and for i from 0 to 783 set $w_{j,i}$ similarly
3. until development accuracy stops increasing
   (a) for each training example k in batches of m examples
      i. do the forward pass using Equations 1.7 1.8, and 1.9
      ii. do the backward pass using Equations 1.20, 1.17, and 1.12
      iii. every m examples, modify all Φ's with the summed updates
   (b) compute the accuracy of the model by running the forward pass on all examples in the development corpus
4. output the Φ from the iteration *before* the decrease in development accuracy.

Figure 1.10: Pseudo code for simple feed-forward digit recognition

Consider the above feed-forward algorithm for MNIST with a batch size of 1. If we are to examine the bias variables before and after the first training example is processed, describe how the bias variables may change as a result.

2) We simplify our MNIST computation by assuming our "image" has two binary-valued pixels, 0 and 1. Further, we assume that there are no bias parameters and that we are performing a binary classification problem.

a) Compute the forward-pass logits and probabilities when the pixel values are $[0,1]$, and the weights are $\begin{bmatrix} 0.2 & -0.3 \\ -0.1 & 0.4 \end{bmatrix}$. Here $w[i,j]$ is the weight on the connection between the i^{th} pixel and the j^{th} unit. For example, $w[0,1] = -0.3$.

b) Assume that the correct answer is 1 (not 0) and that the learning rate $\mathcal{L} = 0.1$. What is the loss?

c) Compute $\Delta w_{0,0}$ on the backward pass.

3) Assume a ConvNet with an initial convolution layer and a pooling layer that follows.

a) Given the following:

i. Input size: 96 x 96 x 3
ii. Kernel size: 5 x 5 x 3
iii. Stride: 1
iv. Max pooling layer: 4 x 4

What is the size of the resulting feature map?

b) Given the following:

v. Input size: 96 x 96 x 3
vi. Kernel size: 3 x 3 x 3
vii. Stride: 3
viii. Max pooling layer: 8 x 8

What is the size of the resulting feature map?