Collab Worksheet 5

CS440/ECE448, Spring 2021

Week of 3/8 - 3/12, 2021

Question 1

A particular two-layer neural net has input vector $x = [x[1], x[2]].T$, hidden layer activations $h[1] = [h[1][1], h[1][2]].T$, and a scalar output $h[2]$. Its weights and biases are stored in a pair of matrices $w[1]$ and $w[2]$ and a pair of vectors $b[1]$ and $b[2]$, respectively. Each of these variables may be indexed using either superscripts and subscripts, or using a python-like notation, as shown in Table 1. The weights and biases are given to you; their values are also provided in Table 1. The hidden layer nonlineary is ReLU; the output nonlinearity is a logistic sigmoid.

Table 1: Variables used in Problem 1.

Subscript Notation	Python-Like Notation
$x = [x_1, x_2]^T$	x = [x[1],x[2]].T
$h^{(1)} = [h_1^{(1)}, h_2^{(1)}]^T$	h = [h[1,1],h[1,2]].T
$h^{(2)}$	h[2]
$w^{(1)} = \begin{bmatrix} 3 & 4 \\ 0 & 9 \end{bmatrix}$	w[1]=[[3,4],[0,9]]
$b^{(1)} = [-3, 3]^T$	b[1] = [-3,3].T
$w^{(2)} = [5, 4]$	w[2] = [5,4]
$b^{(2)} = -7$	b[2]=-7

(a) Suppose the input is $x = [9, -6].T$. What is $h[1]$? Write your answer as a vector of sums of products; do not simplify.

(b) Suppose the hidden layer is $h[1] = [4, 5].T$. What is $h[2]$? Write your answer as a ratio of terms involving the exponential of a sum of products; do not simplify.

Question 2

You have a two-layer neural network trained as an animal classifier. The input feature vector is $x = [x[1], x[2], x[3]].T$, where $x[1]$, $x[2]$, and $x[3]$ are some features. There are two hidden nodes $h[1] = [h[1][1], h[1][2]].T$, and three output nodes, $h[2] = [h[2][1], h[2][2], h[2][3]].T$, corresponding to the three output classes $h[2][1]$ =Pr(Y=dog|X=x), $h[2][2]$ =Pr(Y=cat|X=x), and $h[2][3]$ =Pr(Y=skunk|X=x). The hidden layer uses a sigmoid nonlinearity, the output layer uses a softmax. Each of these variables, and the weight matrices $w[l][j,k]$ and bias vectors $b[l][j]$, may be indexed using either superscripts and subscripts, or using a python-like notation, as shown in Table 2.

Table 2: Variables used in Problem 2.

Subscript Notation	Python-Like Notation
$x = [x_1, x, x_3]^T$	x = [x[1],x[2],x[3]].T
$h^{(1)} = [h_1^{(1)}, h_2^{(1)}]^T$	h = [h[1,1],h[1,2]].T
$h^{(2)} = [h_1^{(2)}, h_2^{(2)}, h_3^{(2)}]^T$	h[2]=[h[2][1],h[2][2],h[2][3]].T
$w^{(1)} = \begin{bmatrix} w_{1,1}^{(1)} & w_{1,2}^{(1)} & w_{1,3}^{(1)} \\ w_{2,1}^{(1)} & w_{2,2}^{(1)} & w_{2,3}^{(1)} \end{bmatrix}$	w[1]=[[w[1][1,1],...],...,[...,w[1][2,3]]]
$b^{(1)} = [b_1^{(1)}, b_2^{(1)}]^T$	b[1] = [b[1][1],b[1][2]].T
$w^{(2)} = \begin{bmatrix} w_{1,1}^{(2)} & w_{1,2}^{(2)} \\ w_{2,1}^{(2)} & w_{2,2}^{(2)} \\ w_{3,1}^{(2)} & w_{3,2}^{(2)} \end{bmatrix}$	w[2]=[[w[1][1,1],...],...,[...,w[2][3,2]]]
$b^{(2)} = [b_1^{(2)}, ..., b_3^{(2)}]^T$	b[2]=[b[2][1],...,b[2][3]].T

(a) A Maltese puppy has the feature vector $x = [2, 20, -1].T$. Suppose all weights and biases are initialized to zero. What is $h[2]$?

(b) Let $w[2][i,j]$ be the weight connecting the i^{th} output node to the j^{th} hidden node. What is $dh[2][2]/dw[2][2,1]$? Write your answer in terms of $h[2][i]$, $w[2][i,j]$, and/or the hidden node activations $h[1][j]$, for any appropriate values of i and/or j.

(c) Suppose that you are presented with an all-zero feature vector $x = [0, 0, 0].T$. Suppose that the first-layer weight matrix is also all zero, $w[1][j,k] = 0$, but the bias is nonzero, specifically, it has the value $b[1] = [12, 13].T$. Suppose that, for this particular training token, $dh[2][2]/dh[1][1] = 15$. What is $dh[2][2]/db[1][1]$? Write your answer as a product of fractions involving exponentials of integers; there should be only constants in your answer, no variables, but you need not simplify.