Neural Networks

Homework 2

1 Forward Pass

$$z_1^1 = \sum_{i=1}^{n} W_{z1,i} \cdot X_{z1,i} = 1 \cdot 2 + 0.5 \cdot 4 - 0.5 \cdot 4 = 2$$

$$z_2^1 = \sum_{i=1}^{n} W_{z2,i} \cdot X_{z2,i} = 1 \cdot (-1) + 0.5 \cdot 1 - 0.5 \cdot (-2) = 0.5$$

Application of sigmoid function:

$$a_1^1 = \text{sigmoid}(z_1^1) = \frac{1}{1+e^{-z_1^1}} = 0{,}881$$

$$a_2^1 = \text{sigmoid}(z_2^1) = \frac{1}{1+e^{-z_1^2}} = 0{,}622$$

$$z_1^2 = W_{z_1^{(2)},1} \cdot a_1^1 + W_{z_1^2,2} \cdot a_2^1 + 1 \cdot 2 = -1{,}265$$

$$p(x) = e^{z_1^2} = e^{W_{z_1^2,1} \cdot a_1^1 + W_{z_1^2,2} \cdot a_2^1 + 1 \cdot 2} = e^{-3 \cdot 0.881 - 1 \cdot 0{,}622 + 1 \cdot 2} = 0.282$$

Backward Pass

First, start from $\frac{\partial J}{\partial W_{11}^2}$:

$$\frac{\partial J}{\partial W_{11}^2} = \frac{\partial J}{\partial p} \cdot \frac{\partial p}{\partial z_1^2} \cdot \frac{\partial z_1^2}{\partial W_{11}^2}$$

The derivative of p with respect to z_1^2 is needed, that's simply: $\frac{\partial p}{\partial z_1^2} = e^{z_1^2}$

The derivative of J with respect to p: $\frac{\partial J}{\partial p} = p - y$

The derivatives of z_1^2 with respect to W_{11}^2: $\frac{\partial z_1^2}{\partial W_{11}^2} = a_1^1$

Then, substitute this into the expression for $\frac{\partial J}{\partial W_{11}^2}$: $\frac{\partial J}{\partial W_{11}^2} = (p-y) \cdot e^{z_1^2} \cdot a_1^1$

Alter the values with the solutions from previous exercises, we have that:

$$\frac{\partial J}{\partial W_{11}^2} = (0{,}282 - y) \cdot 0{,}282 \cdot 0{,}881 = -0{,}248442 \cdot y + 0{,}0700606$$

Now, for the $\frac{\partial J}{\partial W_{11}^1}$: $\frac{\partial J}{\partial W_{11}^1} = \frac{\partial J}{\partial p} \cdot \frac{\partial p}{\partial z_1^2} \cdot \frac{\partial z_1^2}{\partial a_1^1} \cdot \frac{\partial a_1^1}{\partial z_1^1} \cdot \frac{\partial z_1^1}{\partial W_{11}^1}$

The values of the derivative of J to p, and the derivative of p with respect to z_1^2, so let us calculate:

$$\frac{\partial z_1^2}{\partial a_1^1} = W_{11}^2$$

$$\frac{\partial a_1^1}{\partial z_1^1} = \sigma(z_1^1) \cdot (1 - \sigma(z_1^1))$$

$$\frac{\partial z_1^1}{\partial W_{11}^1} = X_{11}$$

Now, put everything together in our equation:

$$\frac{\partial J}{\partial W_{11}^1} = (p-y) \cdot e^{z_1^2} \cdot W_{11}^2 \cdot \sigma(z_1^1) \cdot (1 - \sigma(z_1^1)) \cdot X_{11}$$

Replacing the equations with all known values, final answer:

$$\frac{\partial J}{\partial W_{11}^1} = (0{,}282 - y) \cdot 0.282 \cdot 0.881 \cdot (-3) \cdot (1 - 0{,}881) \cdot 0.5 =$$

$$= -0.01250582 + 0.04434689y$$