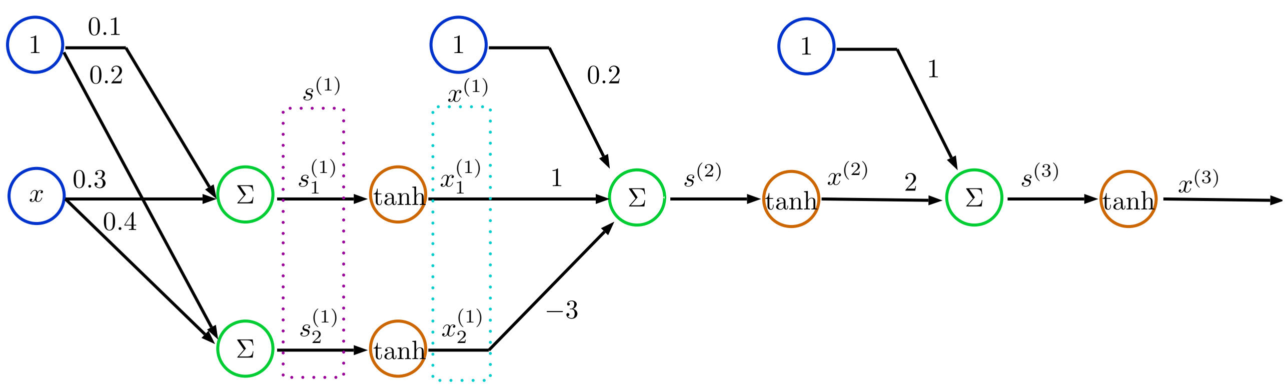
ECE421 - Winter 2022
Homework Problems - Tutorial #6

Theme: Neural Networks & Backpropagation

Due: March 8, 2022 11:59 PM

Question 1 (Based on Example 7.1 from LFD)

Consider the following neural network.



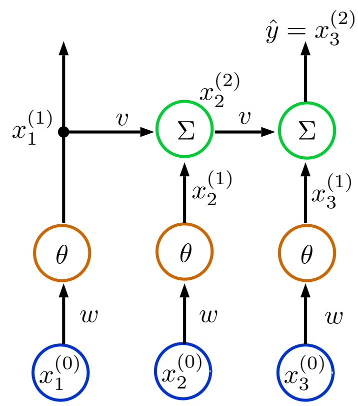
Letting $x = 2$ and $y = 1$, we would like to calculate the gradients as in Example 7.1 from LFD. Clearly show the intermediate steps in calculations (note these steps are not presented in Example 7.1) and perform one round of forward and backward propagation, and use the results to calculate the gradients. Specifically, fill-in all the missing intermediate steps in the computation of

$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}), s^{(2)}, x^{(2)}, s^{(3)}, x^{(3)}, \delta^{(1)}, \delta^{(2)}, \delta^{(3)}, \frac{\partial e}{\partial W^{(2)}}, \frac{\partial e}{\partial W^{(3)}}.$$

For any missing problem specifications/parameters, assume those provided in Example 7.1.

After you have worked out this problem, it may be useful to repeat the steps for Exercise 7.8, where the final (output) tanh is replaced with the identity (note that this part is not considered for grading).

Question 2 (Problem 3 Final 2018)



Given an input (x, y), where $x = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}) \in \mathbb{R}^3$, the above network computes $\hat{y}$ as shown. More specifically, the intermediate computations are

$$\begin{aligned} x_1^{(1)} &= \theta(w \cdot x_1^{(0)}) & \qquad x_2^{(2)} &= x_2^{(1)} + v \cdot x_1^{(1)} \\ x_2^{(1)} &= \theta(w \cdot x_2^{(0)}) & \qquad x_3^{(2)} &= x_3^{(1)} + v \cdot x_2^{(2)} \\ x_3^{(1)} &= \theta(w \cdot x_3^{(0)}) & \qquad \hat{y} &= x_3^{(2)}. \end{aligned} \tag{1}$$

Note that v and w are shared weights on the horizontal edges and vertical edges as shown in the figure. Assume that $\theta(\cdot)$ is some arbitrary activation function with derivative denoted by $\theta'(\cdot)$. For the input (x, y) and model parameter $\Omega = (w, v)$ of the network, assume that the loss is given by

$$e(\Omega) = (\hat{y} - y)^2. \tag{2}$$

(a) Find an expression for $\dfrac{de}{dv}$. Express your answer in terms of $x_1^{(1)}, v, x_2^{(2)}$ and $\Delta = \hat{y} - y$.

(b) Find the expressions for $\dfrac{de}{dx_2^{(2)}}, \dfrac{de}{dx_1^{(1)}}, \dfrac{de}{dx_2^{(1)}}$ and $\dfrac{de}{dx_3^{(1)}}$. Express your answer in the simplest possible form (with as few variables as possible). (Hint: at-least one of your answer should be in terms of Δ and v.)

(c) Using parts (a) and (b) find an expression for $\dfrac{de}{dw}$.