Practical Class 6
Neural Networks and Deep Learning, 2020

September 13, 2020

1. Consider the following neural network. The circled nodes denote variables (*e.g.*, x_1 is an input variable, $\hat{y}$ is the output variable), and rectangular nodes denote functions (*e.g.*, $\oplus$ takes the sum of its inputs, and σ denotes the logistic function $\sigma(x) = 1/\big(1 + \exp(-x)\big)$).
   In the network below, let:

$$
\begin{aligned}
z_1 &= w_1 x_1 + w_2 x_2 \\
z_2 &= w_3 x_3 + w_4 x_4 \\
h_1 &= \sigma(z_1) = \frac{1}{1 + \exp(-w_1 x_1 - w_2 x_2)} \\
h_2 &= \sigma(z_2) = \frac{1}{1 + \exp(-w_3 x_3 - w_4 x_4)} \\
z_3 &= w_5 h_1 + w_6 h_2 \\
\hat{y} &= \sigma(z_3)
\end{aligned}
$$

   We would like to train the NN using the L2 loss:

$$L(y, \hat{y}) = \|y - \hat{y}\|^2 .$$

   Suppose the weights are initialised as:

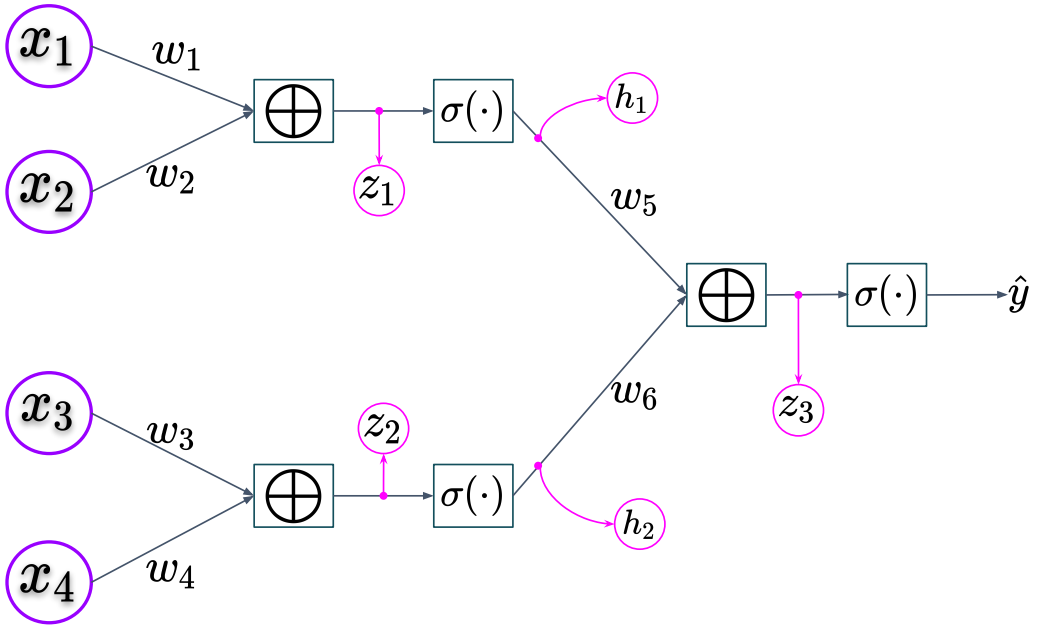
$$
\begin{aligned}
w_1 &= -1.7 \\
w_2 &= 0.1 \\
w_3 &= -0.6 \\
w_4 &= -1.8 \\
w_5 &= -0.2 \\
w_6 &= 0.5
\end{aligned}
$$

   And we are given a data point:

$$
\begin{aligned}
x_1 &= -0.7 \\
x_2 &= 1.2 \\
x_3 &= 1.1 \\
x_4 &= -2.0,
\end{aligned}
$$

   With the true label $y = 0.5$.

   Use the backpropagation algorithm to compute the partial derivative $\partial L / \partial w_1$.


x_1
x_2
x_3
x_4
w_1
w_2
w_3
w_4
w_5
w_6
$\sigma(\cdot)$
$\sigma(\cdot)$
$\sigma(\cdot)$
h_1
h_2
z_1
z_2
z_3
$\hat{y}$