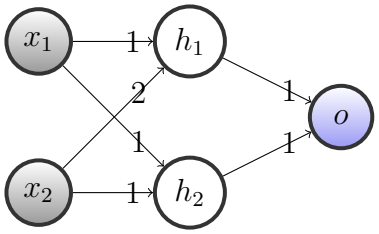
STAT3007/7007 Deep Learning, Tutorial 5

2022 Semester 2

1. (MLP) Consider the following MLP from Lecture 11 slide 9.



The MLP has sigmoid hidden units and identity output activation, with weights shown on the edges. The output o is obtained using the following computation

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \sigma\left(W_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right), \qquad o = W_2 \begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

where $W_1 = \begin{pmatrix} w_{1,11} & w_{1,12} \\ w_{1,21} & w_{1,22} \end{pmatrix}$, and $W_2 = \begin{pmatrix} w_{2,1} & w_{2,2} \end{pmatrix}$.

(a) For arbitrary weight matrices W_1 and W_2 and arbitrary input (x_1, x_2), write down the output o as a function of W_1, W_2 and the input (x_1, x_2).

(b) Assume that when $(x_1, x_2) = (1, 2)$, the true output is $y = 1$. Consider the loss $L = \frac{1}{2}(o - y)^2$. When $W_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, and $W_2 = \begin{pmatrix} 1 & 1 \end{pmatrix}$, compute the partial derivatives in (b) using the backpropagation algorithm.

2. (Automatic differentiation)

(a) Numerical differentiation, symbolic differentiation and automatic differentiation are often used to automatically compute the gradient of a function. Explain how these methods are different from each other.

(b) Draw a computational graph for evaluating the following function

$$f(x_1, x_2) = e^{x_1} \sin x_2 + e^{x_2} \cos x_1.$$

Show how forward mode autodiff can be used to compute the partial derivative $\frac{\partial f}{\partial x_1}$.