Neural network derivation

Machine Learning I (2023-2024)
UMONS

Let $x \in \mathbb{R}^2$ denote the input of a neural network and $y \in \mathbb{R}$ the target value. We define the parameters of the neural network, the forward pass and the loss function below. Can you determine the derivative of the loss function with respect to each parameter for the backward pass?

Parameters

- $u \in \mathbb{R}^{3\times 2}$: Weight matrix for the input-to-hidden layer.
- $b \in \mathbb{R}^3$: Bias vector for the hidden layer.
- $w \in \mathbb{R}^3$: Weight vector for the hidden-to-output layer.
- $c \in \mathbb{R}$: Bias for the output.

Forward pass

$$
\begin{aligned}
g_j &= \sum_{k=1}^{2} u_{jk} x_k + b_j \quad \text{for } j = 1, 2, 3 \\
h_j &= \sigma(g_j) \quad \text{for } j = 1, 2, 3 \\
z &= \sum_{j=1}^{3} w_j h_j + c
\end{aligned}
$$

Note that σ is the logistic function defined as

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

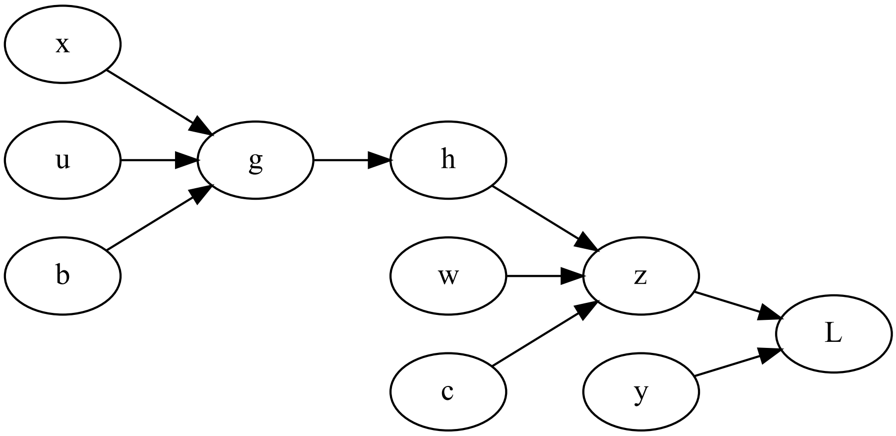
and that its derivative is

$$\frac{\partial \sigma(x)}{\partial x} = \sigma(x)(1 - \sigma(x)).$$

Loss Function

$$L = (z - y)^2$$

Computation graph



Backward Pass

$$
\begin{aligned}
\frac{\partial L}{\partial z} &= 2(z - y) \\
\frac{\partial L}{\partial w_j} &= \frac{\partial L}{\partial z} \cdot h_j \quad \text{for } j = 1, 2, 3 \\
\frac{\partial L}{\partial c} &= \frac{\partial L}{\partial z} \\
\frac{\partial L}{\partial h_j} &= \frac{\partial L}{\partial z} \cdot w_j \quad \text{for } j = 1, 2, 3 \\
\frac{\partial L}{\partial g_j} &= \frac{\partial L}{\partial h_j} \cdot \sigma(g_j)(1 - \sigma(g_j)) \quad \text{for } j = 1, 2, 3 \\
\frac{\partial L}{\partial u_{jk}} &= \frac{\partial L}{\partial g_j} \cdot x_k \quad \text{for } j = 1, 2, 3 \text{ and } k = 1, 2 \\
\frac{\partial L}{\partial b_j} &= \frac{\partial L}{\partial g_j} \quad \text{for } j = 1, 2, 3
\end{aligned}
$$