

---

# Assignment 1. MLPs, CNNs and Backpropagation

---

Maurice Frank  
11650656  
maurice.frank@posteo.de

## 1 MLP backprop

### 1.1

#### 1.1.a

$$\frac{\partial x^{(N)}}{\partial L} : \quad \frac{\partial L}{\partial x_i^{(N)}} = -t_i \cdot \frac{1}{x_i^{(N)}} \quad (1)$$

$$\Rightarrow \frac{\partial x^{(N)}}{\partial L} = -[\dots \frac{t_i}{x_i^{(N)}} \dots] \quad (2)$$

$$\frac{\partial x^{(N)}}{\partial \tilde{x}^{(N)}} : \quad \frac{\partial x_i^{(N)}}{\partial \tilde{x}_j^{(N)}} = \frac{\partial}{\partial \tilde{x}_j^{(N)}} \frac{\exp \tilde{x}_i^{(N)}}{\sum_k \exp \tilde{x}_k^{(N)}} \quad (3)$$

$$= \frac{(\frac{\partial}{\partial \tilde{x}_j^{(N)}} \exp \tilde{x}_i^{(N)}) \cdot \sum_k \exp \tilde{x}_k^{(N)} - \exp \tilde{x}_i^{(N)} \cdot \frac{\partial}{\partial \tilde{x}_j^{(N)}} \sum_k \exp \tilde{x}_k^{(N)}}{(\sum_k \exp \tilde{x}_k^{(N)})^2} \quad (4)$$

$$= \frac{\delta_{ij} \exp \tilde{x}_j^{(N)}}{\sum_k \exp \tilde{x}_k^{(N)}} - \frac{\exp \tilde{x}_i^{(N)} \cdot \exp \tilde{x}_j^{(N)}}{(\sum_k \exp \tilde{x}_k^{(N)})^2} \quad (5)$$

$$= \text{softmax}(\tilde{x}_j^{(N)}) \cdot (\delta_{ij} - \text{softmax}(\tilde{x}_i^{(N)})) \quad (6)$$

$$\Rightarrow \quad (7)$$

$$\frac{\partial x^{(N)}}{\partial \tilde{x}^{(N)}} = \begin{bmatrix} \vdots \\ \dots \text{softmax}(\tilde{x}_j^{(N)}) \cdot (\delta_{ij} - \text{softmax}(\tilde{x}_i^{(N)})) \dots \\ \vdots \end{bmatrix} \quad (8)$$

$$\frac{\partial x^{(l < N)}}{\partial \tilde{x}^{(l < N)}} : \quad \frac{\partial x^{(l < N)}}{\partial \tilde{x}^{(l < N)}} = x^{(l < N)} \odot \tilde{x}^{(l < N)} \quad (9)$$

$$\frac{\partial \tilde{x}^{(l)}}{\partial x^{(l-1)}} : \quad \frac{\partial \tilde{x}^{(l)}}{\partial x^{(l-1)}} = \frac{\partial}{\partial x^{(l-1)}} W^{(l)} x^{(l-1)} + b^{(l)} \quad (10)$$

$$= W^{(l)} \quad (11)$$

$$\frac{\partial \tilde{x}^{(l)}}{\partial W^{(l)}} : \quad \frac{\partial \tilde{x}^{(l)}}{\partial W^{(l)}} = \frac{\partial}{\partial W^{(l)}} W^{(l)} x^{(l-1)} = \begin{bmatrix} \vdots \\ \frac{\partial \tilde{x}_i^{(l)}}{\partial W^{(l)}} \\ \vdots \end{bmatrix} \in \mathbb{R}^{M \times (M \times N)} \quad (12)$$

$$\frac{\partial \tilde{x}^{(l)}}{\partial b^{(l)}} : \quad \frac{\partial \tilde{x}^{(l)}}{\partial b^{(l)}} = \frac{\partial}{\partial b^{(l)}} b^{(l)} = b^{(l)} \otimes b^{(l)} \quad (13)$$

Note the use of  $\oslash$  for element-wise division.

**1.1.b**

(14)