Übungsblatt 3

3.1 a)

$$\sigma'(z) = \left(\frac{1}{1+\exp(-z)}\right)'$$

$$= \left(\left(1+e^{-z}\right)^{-1}\right)'$$

Kettenregel $= \left(1+e^{-z}\right)' \cdot (-1)\left(1+e^{-z}\right)^{-2}$

$$= \frac{+e^{-z}}{\left(1+e^{-z}\right)^2}$$

$$= \frac{1-1+e^{-z}}{\left(1+e^{-z}\right)^2}$$

$$= \frac{1 \cdot \left(1+e^{-z}\right)}{\left(1+e^{-z}\right)\left(1+e^{-z}\right)} - \frac{1}{\left(1+e^{-z}\right)^2}$$

$$= \frac{1}{1+e^{-z}} - \frac{1}{\left(1+e^{-z}\right)^2}$$

$$= \frac{1}{1+e^{-z}} \cdot \left(1 - \frac{1}{1+e^{-z}}\right)$$

$$= \sigma(z) \cdot \left(1 - \sigma(z)\right)$$

3.1 b)

$$\mathrm{softmax}(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{10} e^{z_k}}$$

$$\frac{\partial\, \mathrm{softmax}(z)_i}{\partial z_j} = \frac{\partial \dfrac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}}{\partial z_j}$$

$$f(x) = \frac{g(x)}{h(x)} \Rightarrow f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{h(x)^2}$$

hier: $g_i = e^{z_i}$

$h_i = \sum_{k=1}^{N} e^{z_k}$

$h_i' = e^{z_j}$

Wenn $i = j$ dann $g_i' = e^{z_i}$

Ausonsten $g_i' = 0$

uzrzy - uqjwv

Wenn $i=j$

$$= \frac{e^{z_j}\cdot\sum_{k=1}^{N} e^{z_k} - e^{z_j}e^{z_i}}{\left(\sum_{k=1}^{N} e^{z_k}\right)^2}$$

$$= \frac{e^{z_i}}{\sum_{k=1}^{N} e^{z_k}} \cdot \frac{\sum_{k=1}^{N} e^{z_k} - e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}$$

$$= \text{softmax}(z)_i \cdot \left(1 - \text{softmax}(z)_j\right)$$

Wenn $i \neq j$

$$= \frac{\cancel{0\cdot\sum_{k=1}^{N} e^{z_k}} - e^{z_j}e^{z_i}}{\left(\sum_{k=1}^{N} e^{z_k}\right)^2}$$

$$= \frac{e^{z_i}}{\sum_{k=1}^{N} e^{z_k}} \cdot \left(- \frac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}\right)$$

$$= \text{softmax}(z)_i \cdot \left(-\text{softmax}(z)_j\right)$$

3.1 c)

$$\delta_n^H = \frac{\partial L}{\partial z_n^H} = \frac{\partial L}{\partial a_n} \cdot \frac{\partial a_n}{\partial z_n^H} \overset{\substack{\delta_{ij}=1, \text{da} \\ i=j=n}}{=} \frac{\partial L}{\partial a_n} \cdot \left(a_i \cdot (1 - a_n)\right)$$

$$= \sum_{i=1} \left(-\frac{y_i}{a_i^H}\right) \cdot \left(a_i^H (1 - a_n^H)\right) = -\sum_{i=1} y_i (1 - a_n^H)$$

$$= -y_n + \sum_{i=1} y_i a_n^H \underset{\substack{1 \\ \text{One-Hot-Encoding} \\ \Rightarrow \sum_{i=1} y_i = 1}}{=} -y_n + a_n^H = a_n^H - y_n$$

3.1 d)

$$\delta_n^L = \frac{\partial L}{\partial z_n^L} = \sum_m \frac{\partial L}{\partial z_m^{L+1}} \cdot \frac{\partial z_m^{L+1}}{\partial z_n^L} = \sum_m \delta_m^{L+1} \cdot \frac{\partial z_m^{L+1}}{\partial z_n^L}$$

$$= \sum_m \delta_m^{L+1} \cdot \sigma(z_n^L)(1 - \sigma(z_n^L)) \cdot w_{mn}^{L+1}$$

$$= \underbrace{\sigma(z_n^L)(1 - \sigma(z_n^L))}_{= \sigma'(z_n^L) = h'(z_n^L)} \sum_m \delta_m^{L+1} \cdot w_{mn}^{L+1}$$

$$= h'(z_n^L) \sum_m \delta_m^{L+1} \cdot w_{nm}^{L+1} \qquad (10)$$