Practical 1

Vasileios Charatsidis

June 18, 2018

1 Excersice 1

1.1 Compute (generalized) derivatives of the loss function, with respect to the weights.

$$\mathcal{L} = 0.5 * (y_{out} - y_{gt})^2 \tag{1}$$

$$\textbf{for } \boldsymbol{\frac{\partial \mathcal{L}}{\partial W_{out}}}$$

$$\frac{\partial \mathcal{L}}{\partial W_{out}} = \frac{\partial \mathcal{L}}{\partial y_{out}} \cdot \frac{\partial y_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial W_{out}} \Rightarrow \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial W_{out}} = (y_{out} - s_{out}) \cdot \frac{\partial f_3(s_{out})}{\partial s_{out}} \cdot \frac{\partial (W_{out} \cdot z_2)}{\partial W_{out}} \Rightarrow$$

$$\frac{\partial \mathcal{L}}{\partial W_{out}} = (y_{out} - s_{out}) \cdot f\,'_3(s_{out}) \cdot z_2$$

$$\textbf{for } \boldsymbol{\frac{\partial \mathcal{L}}{\partial W_2}}$$

$$\frac{\partial \mathcal{L}}{\partial W_2} = \frac{\partial \mathcal{L}}{\partial y_{out}} \cdot \frac{\partial y_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial z_2} \cdot \frac{\partial z_2}{\partial s_2} \cdot \frac{\partial s_2}{\partial W_2} \Rightarrow \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial W_2} = (y_{out} - s_{out}) \cdot f\,'_3(s_{out}) \cdot W_{out} \cdot f\,'_2(s_2) \cdot z_1$$

$$\textbf{similarly for } \boldsymbol{\frac{\partial \mathcal{L}}{\partial W_1}}$$

$$\frac{\partial \mathcal{L}}{\partial W_1} = \frac{\partial \mathcal{L}}{\partial y_{out}} \cdot \frac{\partial y_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial z_2} \cdot \frac{\partial z_2}{\partial s_2} \cdot \frac{\partial s_2}{\partial z_1} \cdot \frac{\partial z_1}{\partial s_1} \cdot \frac{\partial s_1}{\partial W_1} \Rightarrow \tag{4}$$

$$\frac{\partial \mathcal{L}}{\partial W_1} = (y_{out} - s_{out}) \cdot f\,'_3(s_{out}) \cdot W_{out} \cdot f\,'_2(s_2) \cdot W_2 \cdot f\,'_1(s_1) \cdot x_{in}$$

1.2 Write down ΔW_k in terms of δ_k

$$\text{We know that: } \delta_k = \frac{\partial \mathcal{L}}{\partial s_k} \text{ and that: } \frac{\partial \mathcal{L}}{\partial s_k} = \frac{\partial \mathcal{L}}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_k}$$

$$\text{So: } \delta_k = \frac{\partial \mathcal{L}}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_k} \tag{5}$$

We can rewrite equations (2), (3) and (4) as:

$$\frac{\partial \mathcal{L}}{\partial W_{out}} = \underbrace{\frac{\partial \mathcal{L}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial s_{out}}}_{\delta_{out}} \cdot \frac{\partial s_{out}}{\partial W_{out}} \Rightarrow$$

$$\Delta W_{out} = \frac{\partial \mathcal{L}}{\partial W_{out}} = \delta_{out} \cdot \frac{\partial s_{out}}{\partial W_{out}} = \boldsymbol{\delta_{out} \cdot z_2} \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial W_2} = \frac{\partial \mathcal{L}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial z_2} \cdot \frac{\partial z_2}{\partial s_2} \cdot \frac{\partial s_2}{\partial W_2} \tag{7}$$

$$\delta_2 = \frac{\partial \mathcal{L}}{\partial s_2} = \frac{\partial \mathcal{L}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial z_2} \cdot \frac{\partial z_2}{\partial s_2} \tag{8}$$

$$\text{from (7) and (8) is implied that } \Delta W_2 = \frac{\partial \mathcal{L}}{\partial W_2} = \delta_2 \cdot \frac{\partial s_2}{\partial W_2} = \boldsymbol{\delta_2 \cdot z_1} \tag{9}$$

similarly for δ_1

$$\frac{\partial \mathcal{L}}{\partial W_1} = \underbrace{\frac{\partial \mathcal{L}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial s_{out}} \cdot \frac{\partial s_{out}}{\partial z_2} \cdot \frac{\partial z_2}{\partial s_2} \cdot \frac{\partial s_2}{\partial z_1} \cdot \frac{\partial z_1}{\partial s_1}}_{\delta_1} \cdot \frac{\partial s_1}{\partial W_1} \Rightarrow$$

$$\text{is implied that } \Delta W_1 = \frac{\partial \mathcal{L}}{\partial W_1} = \delta_1 \cdot \frac{\partial s_1}{\partial W_1} = \boldsymbol{\delta_1 \cdot x_{in}} \tag{10}$$