SMD-Blatt 7 - Aufgabe 21

Lars Volk
Julia Sobolewski
Jannine Salewski

a) - K Klassen
- m Beispiele x_i mit jeweils M Komponenten

$$\begin{pmatrix} \omega_{11} & \dots & \omega_{1,M} \\ \vdots & \ddots & \vdots \\ \omega_{K,1} & \dots & \omega_{K,M} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_K \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_K \end{pmatrix}$$

x_i: Spaltenvektor $x_i \in \mathbb{R}^{M \times 1}$

C: Kostenfunktion $C \in \mathbb{R}$

W: Gewichtungsmatrix $W \in \mathbb{R}^{K \times M}$

b: Bias-Vektor $b \in \mathbb{R}^{K \times 1}$

$$\nabla_W \hat{C} = \sum_{k=1}^{K} \underbrace{\frac{\partial \hat{C}}{\partial f_{ki}}}_{\in \mathbb{R}} \cdot \underbrace{\frac{\partial f_{ki}}{\partial W}}_{\in \mathbb{R}^{K \times M}} \Rightarrow \nabla_W \hat{C} \in \mathbb{R}^{K \times M}$$

$$\nabla_{f_i} \hat{C} = \sum_{k=1}^{K} \underbrace{\frac{\partial \hat{C}}{\partial f_{ki}}}_{\mathbb{R}} \underbrace{\frac{\partial f_{ki}}{\partial f_i}}_{\mathbb{R}^{K \times 1}} \Rightarrow \nabla_{f_i} \hat{C} \in \mathbb{R}^{K \times 1}$$

$$\frac{\partial f_{ki}}{\partial W} \in \mathbb{R}^{K \times M}$$

$$\frac{\partial f_{ki}}{\partial b} \in \mathbb{R}^{K \times 1}$$

b) $\nabla_{f_a} C(f) = \frac{1}{m} \sum_{i=1}^{m} \nabla_{f_a} \hat{C}(f)$

$$\nabla_{f_a} \hat{C}(f) = -\nabla_{f_a} \cdot \sum_{k=1}^{K} \mathbb{1}(y_i = k) \log\left(\frac{e^{f_{ki}}}{\sum_j e^{f_{ji}}}\right)$$

$$= -\nabla_{f_a} \log\left(\frac{e^{f_{y_i i}}}{\sum_j e^{f_{ji}}}\right)$$

$$= \underbrace{\frac{-\sum_j e^{f_{ji}}}{e^{f_{y_i i}}}}_{\text{äußere Ableitung}} \cdot \underbrace{\left[\frac{(\nabla_{f_a} e^{f_{y_i i}}) \cdot \sum_j e^{f_{ji}} - e^{f_{y_i i}} \cdot (\nabla_{f_a} \sum_j e^{f_{ji}})}{(\sum_j e^{f_{ji}})^2}\right]}_{\text{innere Ableitung}}$$

$$= -\frac{1}{e^{f_{y_i i}}} \cdot \left[\frac{e^{f_{ai}} \mathbb{1}(y_i = a) \sum_j e^{f_{ji}} - e^{f_{y_i i}} e^{f_{ai}}}{\sum_j e^{f_{ji}}}\right]$$

$$= -\frac{e^{f_{ai}} \mathbb{1}(y_i = a)}{e^{f_{y_i i}}} + \frac{e^{f_{y_i i}} e^{f_{ai}}}{e^{f_{y_i i}} \sum_j e^{f_{ji}}} = \frac{e^{f_{ai}}}{\sum_j e^{f_{ji}}} - \mathbb{1}(y_i = a)$$

$$\Rightarrow \nabla_{f_a} C(f) = \frac{1}{m} \sum_{i=1}^{m} \left[\frac{e^{f_{ai}}}{\sum_j e^{f_{ji}}} - \mathbb{1}(y_i = a)\right] \quad \text{qed.}$$