

Übung 3 DeepLearning für Visual Computing SoSe 25

a) Geg: vereinfachtes Netzwerk, Eingabe $x \in \mathbb{R}$, n versteckte Schichten $h_i = \sigma(w_i \cdot h_{i-1})$, w_i Gewicht der Schicht
 $\sigma: \mathbb{R} \rightarrow \mathbb{R}$ beliebige elementweise Aktivierungsfunktion
 $h_0 = x$, h_n Ausgabe des Netzes

1. Berechne $\frac{\partial h_i}{\partial w_i}$ für $i \in \{1, \dots, n\}$

Lsg:

Kettenregel $(u(v(x)))'(x) = u'(v(x)) \cdot v'(x)$

$$\text{I. } \frac{\partial h_j}{\partial h_{j-1}} \stackrel{\text{Def}}{=} \frac{\partial \sigma(w_j \cdot h_{j-1})}{\partial h_{j-1}} \stackrel{\downarrow}{=} \sigma'(w_j \cdot h_{j-1}) \cdot w_j$$

$$\text{II. } \frac{\partial h_i}{\partial w_i} \stackrel{\text{Kettenregel}}{=} \sigma'(w_i \cdot h_{i-1}) \cdot h_{i-1}$$

$$\begin{aligned} \frac{\partial h_n}{\partial w_i} &\stackrel{\text{Kettenregel}}{=} \frac{\partial h_n}{\partial h_i} \cdot \frac{\partial h_i}{\partial w_i} \stackrel{\text{Kettenregel}}{=} \frac{\partial h_n}{\partial h_{n-1}} \cdot \frac{\partial h_{n-1}}{\partial h_i} \cdot \frac{\partial h_i}{\partial w_i} \\ &\stackrel{\text{Kettenregel}}{=} \frac{\partial h_n}{\partial h_{n-1}} \cdot \frac{\partial h_{n-1}}{\partial h_{n-2}} \cdot \frac{\partial h_{n-2}}{\partial h_i} \cdot \frac{\partial h_i}{\partial w_i} = \frac{\partial h_n}{\partial h_{n-1}} \cdot \dots \cdot \frac{\partial h_{i+1}}{\partial h_i} \cdot \frac{\partial h_i}{\partial w_i} \\ &= \prod_{j=i+1}^n \frac{\partial h_j}{\partial h_{j-1}} \cdot \frac{\partial h_i}{\partial w_i} \stackrel{\text{I}}{=} \prod_{j=i+1}^n [\sigma'(w_j \cdot h_{j-1}) w_j] \cdot \frac{\partial h_i}{\partial w_i} \end{aligned}$$

$$\text{II} = \prod_{j=i+1}^n [\sigma'(w_j \cdot h_{j-1}) w_j] \cdot \sigma'(w_i \cdot h_{i-1}) \cdot h_{i-1}$$

