

Übungsblatt 7A1

① a)  $w_2 = 0.6 \quad b_2 = 0.9$

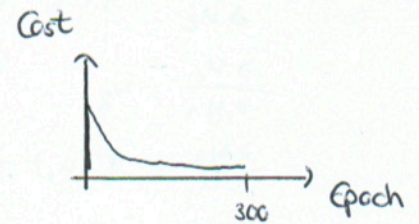
$$\frac{\partial E}{\partial b_2} \stackrel{T=0}{=} -2 (T - y_2) \cdot f'(u_2)$$

$$\stackrel{(*)}{=} -2 y_2 \frac{1}{1 + e^{-1.5}} \left( 1 - \frac{1}{1 + e^{-1.5}} \right)$$

$$= 0.2983 y_2$$

$$y_2 = \frac{1}{1 + e^{-1.5}} = 0.2439$$

$$(*) \quad u_2 = 0.6 y_1 - 0.9 \quad y_1 = 1.5$$



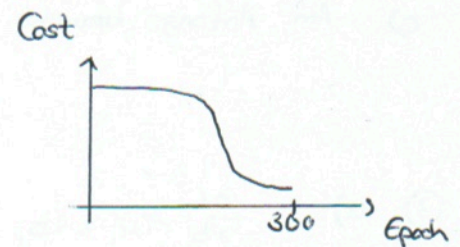
$$(w_2 = 0.6 \quad b_2 = 0.9 \rightarrow w_2 = -1.28 \quad b_2 = -0.98)$$

b)  $w_2 = b_2 = 2$

$$u_2 = w_2 y_1 + b_2 = 2 y_1 + 2 \quad y_1 = 4$$

$$\frac{\partial E}{\partial b_2} = 0.0353 y_2 \quad = 0.0347$$

$$y_2 = \frac{1}{1 + e^{-1.5}}$$



c) Problem: Das Lernen beginnt sehr viel langsamer.

Der Faktor  $f'(u_2)$  aus Gleichung (3) ist hauptsächlich dafür verantwortlich.



$$\textcircled{2} \text{ a) } \frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial y_1} \frac{\partial y_1}{\partial u_1} \frac{\partial u_1}{\partial b_1}$$

$$\frac{\partial E}{\partial y_2} = -2(T - y_2)$$

$$\frac{\partial y_2}{\partial u_2} = f(u_2)(1 - f(u_2)) = f'(u_2)$$

$$\frac{\partial u_2}{\partial y_1} = w_2$$

$$\frac{\partial y_1}{\partial u_1} = f'(u_1) = f(u_1)(1 - f(u_1))$$

$$\frac{\partial u_1}{\partial b_1} = 1$$

$$\frac{\partial E}{\partial b_1} = -2(T - y_2) f'(u_2) w_2 f'(u_1) \cdot 1$$

Das Problem wird verstärkt, da wir  $2 \cdot f'(u_i)$  haben.

b) Die Probleme verstärken sich, da noch mehr Ableitungen  $f'(u_i)$  hinzukommen.

c) Am Anfang braucht das Lernen <sup>länger</sup> ~~länger~~, somit ist der Lernfortschritt langsam.

$$\textcircled{3} \text{ a) } u_1 = w_1 x + b_1 = 100x - 100 \stackrel{x=1}{=} 0$$

$$u_2 = w_2 y_1 + b_2 = 100 \cdot f(u_1) - 50 = 100 \cdot \frac{1}{2} - 50 = 0$$

$$f(u_1) = f(u_2) = \frac{1}{2}$$

$$f'(u_1) = f'(u_2) = \frac{1}{4}$$

$$\begin{aligned} \frac{\partial E}{\partial b_1} &= -2(0 - y_2) f'(0) \cdot 100 f'(0) \\ &= 2 \cdot \frac{1}{2} \cdot \frac{1}{4} \cdot 100 \cdot \frac{1}{4} = \frac{100}{16} = \frac{50}{8} = \frac{25}{4} \end{aligned}$$

$$\frac{\partial E}{\partial b_2} = -2(T - y_2) f'(u_2) = -2(0 - \frac{1}{2}) \frac{1}{4} = \frac{1}{4}$$

$$\text{b) } \frac{\partial E}{\partial b_i} = \frac{\partial E}{\partial b_{i+1}} \cdot w_{i+1} f'(u_i) = \frac{\partial E}{\partial b_i} \cdot \frac{100}{4}$$

c) bei schlechtem / langsamem Lernen  $\Rightarrow$  Kettenreaktion  
(die anderen lernen langsamer)

$\textcircled{4}$  Die cross-entropy Funktion verwendet weniger Ableitungen ( $f'$ ).