Aufgabe 2:

1a) $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \frac{e^{cu_i}}{\sum_{j=1}^{n} e^{cu_j}} = \frac{\sum_{j=1}^{n} e^{cu_j}}{\sum_{j=1}^{n} e^{cu_j}} = 1$ ✓

$y_i = \frac{\overbrace{e^{cu_i}}^{\geq 0}}{\underbrace{\underbrace{\sum_{j=1}^{n} e^{cu_j}}_{\geq 0}}_{\geq 0}} \geq 0 \quad ; \forall i$ ✓

$\Rightarrow$ Es ist gültige W'keitsdichte

---

b) $y_1 = \frac{e^{cu_1}}{\sum_{j=1}^{3} e^{cu_j}} = \frac{1}{e^{-cu_1}(e^{cu_1} + e^{cu_2} + e^{cu_3})}$

$= \frac{1}{(1 + e^{c(u_2-u_1)} + e^{c(u_3-u_1)})}$

---

c) i) $u_1 > u_2 > u_3$

$\Rightarrow y_1 = \frac{1}{1 + \underbrace{e^{c(u_2-u_1)}}_{\to 0} + \underbrace{e^{c(u_3-u_1)}}_{\to 0}} \xrightarrow{c \to \infty} \frac{1}{1+0+0} = 1$

(da $(u_2-u_1) < 0$ und $(u_3-u_1) < 0$)

ii) $u_2 > u_1 > u_3$

$\Rightarrow (u_2 - u_1) > 0$ und $(u_3 - u_1) < 0$

$\Rightarrow \lim_{c \to \infty} \frac{1}{1 + \underbrace{e^{c(u_2-u_1)}}_{\to \infty} + \underbrace{e^{c(u_3-u_1)}}_{\to 0}} \left(\text{"} = \frac{1}{1+\infty+0} \text{"} \right) = 0$

iii) $u_2 > u_3 > u_1$

$\Rightarrow \lim_{c \to \infty} y_1 = \lim_{c \to \infty} \frac{1}{1 + \underbrace{e^{c(u_2-u_1)}}_{\to \infty} + \underbrace{e^{c(u_3-u_1)}}_{\to \infty}} = 0$

d) i) $c>0$:

das Ergebnis ist wie erwartet. Die dendritischen Potenziale werden also „gedämpft", so dass hohe Werte danach immernoch relativ hoch bleiben, aber die Eigenschaften der W'keitsfunktion erfüllt sind.

ii) $c=0$:

Ergebnisverteilung entartet zu einer Gleichverteilung unabhängig von den dendritischen Potenzialen.

iii) $c<0$:

Ergebnisverteilung verhält sich jetzt gegenteilig zu den dendritischen Potenzialen. Niedrige u_i werden zu danach relativ großen y_i.

2a)

$$\frac{\partial E}{\partial y_1} = \frac{\partial}{\partial y_1} -t_1 \cdot \ln(y_1) = \frac{-t_1}{y_1}$$

$$\frac{\partial E}{\partial y_2} = \frac{\partial}{\partial y_2} -t_2 \ln(y_2) = \frac{-t_2}{y_2}$$

---

b)

$$\frac{\partial y_1}{\partial u_2} = \frac{\partial}{\partial u_2} \frac{e^{u_1}}{e^{u_2}+e^{u_1}} = e^{u_1} \frac{\partial}{\partial u_2} \frac{1}{e^{u_2}+e^{u_1}}$$

$$\stackrel{\text{Kettenr.}}{=} e^{u_1}\left((e^{u_2}+e^{u_1})^{-2}(-1)\left(\frac{\partial}{\partial u_2}(e^{u_2}+e^{u_1})\right)\right)$$

$$= \frac{-e^{u_1}}{e^{u_2}+e^{u_1}} \frac{e^{u_2}}{e^{u_2}+e^{u_1}} = -y_1 y_2$$

$$\frac{\partial y_2}{\partial u_2} = \frac{\partial}{\partial u_2} \frac{e^{u_2}}{e^{u_2}+e^{u_1}} \stackrel{\text{Produktr.}}{=} \frac{e^{u_2}}{e^{u_2}+e^{u_1}} + e^{u_2}\left(\frac{\partial}{\partial u_2} \frac{1}{e^{u_1}+e^{u_2}}\right)$$

$$\stackrel{\text{Kettenregel}}{=} y_2 + \underbrace{e^{u_2}(-1)(e^{u_1}+e^{u_2})^{-2}\, e^{u_2}}_{=y_2^2} = y_2 + y_2^2$$

$$= y_2(1+y_2)$$

c) $\frac{\partial u_2}{\partial w_2} = \frac{\partial}{\partial w_2} w_2 x + b_2 = x$

---

d) $\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial y_1} \frac{\partial y_1}{\partial u_2} \frac{\partial u_2}{\partial w_2} + \frac{\partial E}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial w_2}$

$= \left(-\frac{t_1}{y_1}\right)(-y_1 \cdot y_2) \cdot x + \left(-\frac{t_2}{y_2}\right) y_2 (1 - y_2) \cdot x$

$= t_1 y_2 x + t_2 y_2 x - t_2 x$

$= x(y_2(t_1 + t_2) - t_2) = x(y_2 - t_2)$

---

e) Ableitung der quadratischen Fehlerfunktion:

$\frac{\partial E}{\partial w} = \sum_{\mu=1}^{M} (T_\mu - y(x_\mu))(-y'(x_\mu)) x_\mu$

Hauptunterschied zu unserer Fehlerfunktion ist, dass ~~sie~~ ihre Ableitung von der Ableitung der Transferfunktion abhängt. Bei unserer Fehlerfunktion fällt dieser Teil weg.