Blatt 5 Neuronale Netze

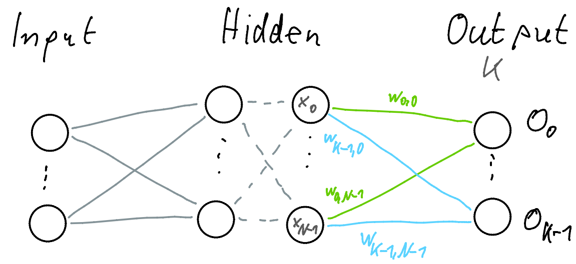
Donnerstag, 13. August 2020 20:15

Aufgabe 1 (6 TP)
Bestimmen Sie $\frac{\partial E}{\partial w_k}$ für

Wertung k-ter Output Zielwert

$$E(w) = \frac{1}{2}\sum_k (\mathcal{O}_k - t_k)^2 \tag{1}$$

und wobei K die Menge der Knoten im Outputlayer eines Neuronalen Netzwerk ist.



Die Output-Ebene habe K Neuronen.
Sei $0 \leq i < K$, dann ist O_i der Wert des i-ten Neuronen der letzten Ebene.

Die vorletzte Ebene habe N Neuronen und es sei $0 \leq l < N$.

Dann berechnet sich O_i durch:

$$\boxed{O_i = \phi\left(\sum_{n=0}^{N-1} w_{i,n} \cdot x_n + \theta_i\right)}$$

neuer Wert des aktuellen Neurons — Sigmoid Funktion — Gewicht — Wert des vorherigen Neurons — Bias

$$\boxed{\phi(x) = \frac{1}{1+e^{-x}}}$$

Wir suchen nun $\frac{\partial E}{\partial w_{i,l}}$.

$$\frac{\partial E}{\partial w_{i,l}} = \frac{\partial}{\partial w_{i,l}}\left(\frac{1}{2}\sum_{k=0}^{K-1}(O_k - t_k)^2\right)$$

$$= \frac{1}{2}\sum_{k=0}^{K-1} \underbrace{\frac{\partial}{\partial w_{i,l}}(O_k - t_k)^2}_{0 \text{ falls } i \neq k \text{ ist.}}$$

$$= \frac{1}{2}\frac{\partial}{\partial w_{i,l}}(O_i - t_i)^2$$

$$\overset{K.R.}{=} \frac{1}{2} \cdot 2(O_i - t_i)\frac{\partial}{\partial w_{i,l}} O_i$$

$$= (O_i - t_i)\frac{\partial}{\partial w_{i,l}} \phi(s_i)$$

Beachte:

$$O_i = O_i(w_{i,0}, \dots, w_{i,l}, \dots, w_{i,N-1}, \theta_i)$$

K.R. = Kettenregel

$$\frac{\partial f(g(x))}{\partial x} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$$

$$s_i := \left(\sum_{n=0}^{N-1} w_{i,n} \cdot x_n\right) + \theta_i$$

Werte der Knoten in der vorherigen Ebene.

$\overset{K.R.}{=} (O_i - t_i) \frac{\partial}{\partial w_{i,l}} \frac{1}{1+e^{-s_i}}$

$= (O_i - t_i) \left(\frac{\partial}{\partial s_i} \frac{1}{1+e^{-s_i}} \right) \left(\frac{\partial s_i}{\partial w_{i,l}} \right)$

Q.R.: Quotientenregel

$\frac{\partial}{\partial x} \frac{f}{g} = \frac{g f' - f g'}{g^2}$

$\overset{Q.R.}{=} (O_i - t_i) \left(\frac{e^{-s_i} - 1}{(1+e^{-s_i})^2} \right) \frac{\partial}{\partial w_{i,l}} \left[\sum_{n=0}^{N-1} w_{i,n} \cdot x_n \right] + \theta_i$

(*) bis hier unabhängig von n, identisch mit $\frac{\partial E}{\partial \theta_i}$

$$\boxed{\frac{\partial E}{\partial w_{i,l}} = (O_i - t_i) \left(\frac{e^{-s_i} - 1}{(1+e^{-s_i})^2} \right) x_{i,l}}$$

Aufgabe 2 (6 TP)
Bestimmen Sie $\frac{\partial E}{\partial \theta_j}$ für

$$E(\theta) = \frac{1}{2} \sum_k (\mathcal{O}_k - t_k)^2 \tag{2}$$

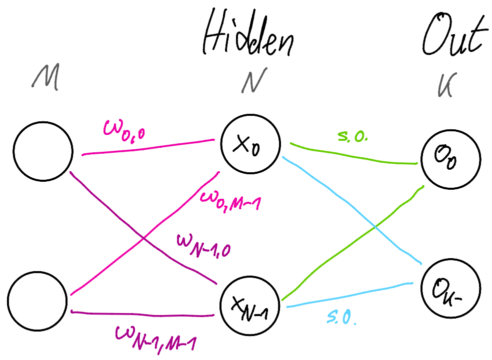
und wobei J die Menge der Knoten im Hiddenlayer eines Neuronalen Netzwerk ist (es gibt nur ein hidden layer).

θ_j bezeichne den Bias eines Knoten in der vorletzten Ebene.

Wir gehen davon aus, dass alle Verbindungen zwischen den letzten beiden Ebenen existieren. $\Rightarrow$ alle O_k hängen von θ_j ab.

$\frac{\partial E}{\partial \theta_j} = \frac{\partial}{\partial \theta_j} \frac{1}{2} \sum_{k=0}^{K-1} (O_k - t_k)^2$

$= \sum_{k=0}^{K-1} \underbrace{\frac{\partial}{\partial \theta_j} \frac{1}{2} (O_k - t_k)^2}_{\text{siehe oben}}$



$\overset{(*)}{=} \sum_{k=0}^{K-1} (O_k - t_k) \left(\frac{e^{-s_k} - 1}{(1+e^{-s_k})^2} \right) \frac{\partial}{\partial \theta_j} \left[\sum_{n=0}^{N-1} w_{k,n} \cdot x_n \right] + \theta_k$

Nur $x_{k,j}$ ist von θ_j abhängig.

$= \sum_{k=0}^{K-1} (O_k - t_k) \left(\frac{e^{-s_k} - 1}{(1+e^{-s_k})^2} \right) w_{k,j} \frac{\partial}{\partial \theta_j} x_j$

$x_j = \sum_{m=1}^{M-1} \omega_{j,m} \, x_{j,m} + \theta_j$

$$= \sum_{k=0}^{K-1} (O_k - t_k) \left(\frac{e^{-s_k}}{(1+e^{-s_k})^2} \right) W_{k,j} \underbrace{\frac{\partial}{\partial \theta_j} \left(\sum_{m=1}^{M-1} w_{j,m} \, x_{j,m} + \theta_j \right)}_{1}$$

$$\boxed{\frac{\partial E}{\partial \theta_j} = \sum_{k=0}^{K-1} (O_k - t_k) \left(\frac{e^{-s_k}}{(1+e^{-s_k})^2} \right) W_{k,j}}$$

Bemerkungen zur Notation:

Ich habe mich bemüht eindeutige Notation zu nutzen, aber gleichzeitig nicht zu weit von der Vorlesung abzuweichen.
Eine bessere Notation würde jedem Gewicht 3 Indizes zuordnen:

- Einen dafür, in welcher Ebene das Gewicht vorkommt,
- einen dafür, an welchem Knoten in der Ebene das Gewicht vorkommt
- und einen letzten Index für den Index des Knotens der Vorebene, dem das Gewicht zugeordnet ist.

Jeder Knoten (Neuron) und jeder zugehörige Bias bräuchte zwei Indizes:

- Einen, der beschreibt, in welcher Ebene das Neuron ist,
- und einen, der beschreibt an welcher Stelle das Neuron in der Ebene ist.

Die hier und in der Vorlesung verwendete Notation ersetzt jeweils einen dieser Indizes durch Verwendung anderer Buchstaben, was aber für unwissende Leser nicht immer verständlich ist.