Einführung in die Neuroinformatik SoSe 2012
Institut für Neuroinformatik
Dr. F. Schwenker

3. Aufgabenblatt (Abgabe am 22.05.2012 zur Vorlesung)

---

4. Aufgabe (3+3): Lineare Mehrschichtnetze

1. Gegeben sei ein zweischichtiges neuronales Netzwerk aus jeweils linearen Neuronen mit den Kopplungsmatrizen $C_1 \in \mathbb{R}^n \times \mathbb{R}^m$ und $C_2 \in \mathbb{R}^m \times \mathbb{R}^p$. Nehmen Sie als Konstante $\theta_j = 0$ für alle Neuronen an. Zeigen Sie, dass sich das zweischichtige Netz aus linearen Neuronen durch ein einschichtiges Netz, definiert durch eine Kopplungsmatrix $A \in \mathbb{R}^n \times \mathbb{R}^p$ und mit p linearen Neuronen darstellen lässt, geben Sie A an?

2. Voraussetzungen wie im Aufgabenteil 1. Seien nun aber $\theta_j \in \mathbb{R}$ beliebige Konstanten. Lässt sich dieses zweischichtige Netz aus linearen Neuronen in ein einschichtiges Netz aus linearen Neuronen transformieren? Begründen Sie ihre Antwort! Geben Sie ggf. die Kopplungsmatrix des Netzes und die Konstanten θ_j der linearen Neuronen an.

5. Aufgabe (4): Lernregel

In der Vorlesung wurde eine **batch modus** Lernregel (als Gradientenverfahren) für ein einschichtiges neuronales Netzwerk mit n nichtlinearen Neuronen hergeleitet.
Gegeben sei dazu eine Trainingsmenge

$$T = \{(x^\mu, T^\mu) \mid x^\mu \in \mathbb{R}^d, T^\mu \in \mathbb{R}^n,\ \mu = 1, \dots, M\}$$

Das j-te Neuron berechnet dann bei Eingabe von $x^\mu \in \mathbb{R}^d$:

1. das dendritische Potential: $u_j^\mu = \sum_{i=1}^d x_i^\mu c_{ij}$

2. das axonale Potential: $y_j^\mu = f(u_j^\mu) = f\left(\sum_{i=1}^d x_i^\mu c_{ij}\right)$

Zur Fehlerbewertung der Netzwerkausgaben wurde in der Vorlesung als Zielfunktion die Summe der quadratischen Abweichungen zwischen der Netzausgabe und dem Lehrersignal in der Trainingsmenge benutzt, d.h.

$$E_\delta(C) = \sum_{\mu=1}^{M} \sum_{j=1}^{n} \delta(T_j^\mu - y_j^\mu)$$

mit $\delta : \mathbb{R} \to \mathbb{R}$ definiert durch $\delta(s) = s^2$.
Verwenden Sie nun als Zielfunktion $E_{\tilde{\delta}}(C) = \sum_{\mu=1}^M \sum_{j=1}^n \tilde{\delta}(T_j^\mu - y_j^\mu)$ mit

$$\tilde{\delta}(s) = \beta^2 \ln[\cosh(s/\beta)] \quad \text{mit} \quad \beta > 0 \quad \text{und} \quad \cosh(x) = \frac{e^x + e^{-x}}{2}$$

und leiten Sie hierfür die Lernregeln für die synaptischen Kopplungen c_{ij} her.