Barbara Hammer, FB Mathematik/Informatik, Universität Osnabrück

Neuronale Netze (SS 2002)
5. Übungsblatt

Abgabe: Donnerstag, 16.5.02, $12^{\underline{00}}$ Uhr, Briefkasten ‚Neuronale Netze' im 4. Stock des AVZ

---

1. (5 Punkte) Noch einmal zurück zu Blatt 1, Aufgabe 1. Finden Sie je ein vorwärtsgerichtetes neuronales Perzeptronnetz, das die schwarzen Flächen auf 1 und den Rest auf 0 abbildet?

   Für welche aller möglichen binären Funktionen $f : \{0,1\}^n \to \{0,1\}$ kann man ein vorwärtsgerichtetes neuronales Perzeptronnetz finden, das f implementiert?

2. (5 Punkte) Sigmoide contra Perzeptron:

   Sei $P \subset \mathbb{R}^n \times \{0,1\}$ eine endliche Trainingsmenge. Wir sagen, daß ein sigmoides neuronales vorwärtsgerichtetes Netz die Punkte P implementiert, sofern die Ausgabe des Netzes bei Eingabe $\vec{x}$ nur um 0.01 von der gewünschten Ausgabe y für alle $(\vec{x}, y) \in P$ abweicht. Was gilt dann; warum?

   - Kann P durch ein Perzeptronnetz implementiert werden, dann auch durch ein sigmoides Netz derselben Architektur.
   - Kann P durch ein sigmoides Netz implementiert werden, dann auch durch ein Perzeptronnetz derselben Architektur.

3. (5 Punkte) Wir hatten den quadratischen Fehler $E = 0.5 \cdot \sum_{p,i}(o_i(\vec{x}^p) - y_i^p)^2$ für Ausgabeneuronen i und Muster $(\vec{x}^p, y^p)$ eingeführt. Training eines Netzes bedeutete, diesen Fehler durch Gradientenabstieg zu minimieren. Könnten wir anstelle des quadratischen Fehler einen der folgenden Fehler durch Gradientenabstieg minimieren? Was könnten jeweils Vorteile/Nachteile sein?

   - $E_1 = 0.25 \cdot \sum_{p,i}(o_i(\vec{x}^p) - y_i^p)^4$
   - $E_2 = \sum_{p,i} |o_i(\vec{x}^p) - y_i^p|$
   - $E_3 = \begin{cases} 0 & \text{falls } o_i(\vec{x}^p) = y_i^p \ \forall i, p \\ 42 & \text{sonst} \end{cases}$
   - $E_4 = 1 - \exp(-E)$, wobei E der normale quadratische Fehler ist.
   - $E_5 = \sum_i o_i(\vec{x}^p)^3$

4. (5 Punkte) Implementieren Sie einen einfachen Gradientenabstieg für eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ und minimieren Sie die folgenden Funktionen:

   - $f_1(x,y) = (x-1)^2 + (y-2)^2$
   - $f_2(x,y) = \sqrt{(18x+100)^2 + (5y-50)^2}$
   - $f_3(x,y) = e(0,0,0) + e(0,1,1) + e(1,0,1) + e(1,1,1)$
     mit $e(a,b,c) := (\mathrm{sgd}(4ax + 4by - 26.5) - c)^2$
   - $f_4(x,y) = f_3(x,y) + f_3(15-x, 15-y)$,
   - $f_5(x,y) = \sin(x - 0.5)\sin(y)\mathrm{e}^{-\frac{(x-10)^2+y^2}{40}}$

   Dokumentieren Sie Ihre Ergebisse und die jeweiligen Schwierigkeiten eines Gradientenabstiegs in den Fällen.

   (P.s.: Ableitungen der Funktionen stehen unter blatt5/ableitungen im Netz.)