

Übungen zur Vorlesung  
**Ausgewählte Kapitel der Computational Intelligence**  
 SS 2008  
 Blatt 3

**Aufgabe 3.1: Fuzzy Clustering vom Typ 1** (10 Punkte)

Implementieren Sie den Fuzzy-C-Means Algorithmus (FCM) oder besorgen Sie sich eine lauffähige Version, die für die hier vorgesehenen Tests angepasst werden muss.

- a) Testen Sie den Algorithmus für die Daten in den Dateien `uniform.txt`, `uniform4.txt` und `olympic.txt` für  $c = 2, \dots, 6$  Cluster und Kontrastparameter  $p = \frac{11}{10}, 2, 4, 6$ . Dokumentieren Sie die Positionen der Clusterzentren und interpretieren Sie die Zugehörigkeitsmatrix. Die Dateien sind auf der Webseite der Übung verfügbar. Die 2-dimensionalen Datenpunkte sind zeilenweise im Textformat gespeichert.
- b) Berechnen Sie die Gütemaße namens Partitionskoeffizient

$$p_C = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K u_{ik}^2,$$

Partitionsentropie

$$p_E = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K u_{ik} \cdot \log u_{ik}$$

und Xie-Beni-Index

$$p_{XB} = \frac{\sum_{i=1}^N \sum_{k=1}^K u_{ik}^2 \cdot d_k^2(x_i)}{K \cdot \min\{d(m_k, m_\ell) : 1 \leq k, \ell \leq K, k \neq \ell\}}$$

für die obigen Experimente und dokumentieren Sie die Werte. Lässt sich daraus eine „richtige“ Clustereinteilung ablesen?

**Aufgabe 3.2: Theorie: Varianten des Fuzzy-C-Means Clusterverfahren** (10 Punkte)

Was muss am FCM-Algorithmus verändert werden, wenn als Abstandsnorm die  $\ell_1$ -Norm

$$d(x, y) = \|x - y\|_1 = \sum_{i=1}^n |x_i - y_i|$$

bzw. die  $\ell_q$ -Norm

$$d(x, y) = \|x - y\|_q = \left( \sum_{i=1}^n (x_i - y_i)^q \right)^{\frac{1}{q}}$$

verwendet wird? Begründen Sie Ihre Aussagen.

---

Abgabe bis zum 02.07.2008, 12.00 Uhr an Günter Rudolph per E-Mail. Sie haben freie Auswahl bei der Programmiersprache.