Intelligent Data Analysis Summer 2013
C. Braune

Exercise Sheet 8

Exercise 30 Fuzzy Clustering

Consider the objective function of fuzzy clustering with a fuzzifier $w = 1$, that is,

$$J(\mathbf{X}, \mathbf{B}, \mathbf{U}) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} \, d^2(\beta_i, \vec{x}_j),$$

which is to be minimized under the constraint

$$\forall j \in \{1, \dots, n\} : \quad \sum_{i=1}^{c} u_{ij} = 1.$$

Show that one obtains a hard/crisp assignment of the data points even if the membership degrees u_{ij} may come from the interval $[0, 1]$. That is, show that for the minimum of the objective function J it is $\forall i \in \{1, \dots, c\} : \forall j \in \{1, \dots, n\} : u_{ij} \in \{0, 1\}$.
(Hint: You may find it easier to consider the special case $c = 2$ (two clusters) and to examine the term for a single data point $\vec{x}_j$. Then generalize the result.)

Exercise 31 Agglomerative Clustering

Let the following one-dimensional data set be given:

$$2, 5, 11, 12, 17, 21, 32.$$

Process this data set with hierarchical agglomerative clustering using

a) the centroid method,
b) the single linkage methode,
c) the complete linkage methode!

Draw a dendrogram for each case!

Exercise 32 Fuzzy Clustering

Consider the objective function

$$J(\mathbf{X}, \mathbf{B}, \mathbf{U}) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} \, d^2(\beta_i, \vec{x}_j),$$

which is to be minimized under the constraint

$$\forall j \in \{1, \ldots, n\} : \quad \sum_{i=1}^{c} \sqrt{u_{ij}} = 1.$$

Derive the update formulae for the membership degrees and the cluster centers using the Euclidean distance. How does the result differ from standard fuzzy clustering with a fuzzifier $w = 2$? (In particular, consider the cluster centers.)