Fuzzy c-means clustering

Recall the membership function $c(i, j)$:

$$c(i,j) = \begin{cases} 1 & x_i \text{ belongs to class } j \\ 0 & x_i \text{ does not belong to class } j \end{cases}$$

Observe that it allows cases in which x_i may belong to more than one class. It also allows for a "fuzzy" description, where $0 \leq c(i, j) \leq 1$, and for a probabilistic description. We impose the conditions: $c(i, j) \geq 0$, and for all i, $\sum_j c(i, j) = 1$. Here the value of $c(i, j)$ is the likelihood, that x_i belongs to class j.

The following error criterion generalizes the k-means error criterion:

$$J = \sum_{i=1}^{m} \sum_{j=1}^{k} c(i,j) \|x_i - u_j\|^2$$

where $c(i, j)$ is the fuzzy degree of membership of x_i in Cluster j.

Given the u_j, consider the membership values of x_j. It should be inversely related to $\|x_i - u_j\|$. It can be computed as follows:

$$\begin{aligned} &\text{for all } i, j: && \tilde{c}(i,j) = \frac{1}{\|x_i - u_j\|} \\ &\text{for all } i: && z_i = \sum_{j=1}^{k} \tilde{c}(i,j) \\ &\text{for all } i, j: && c(i,j) = \frac{\tilde{c}(i,j)}{z_i} \end{aligned}$$

Given the $c(i, j)$ the new c-means are computed by:

$$\text{for all } j: \quad u_j = \frac{\sum_{i=1}^{m} c(i,j) x_i}{\sum_{i=1}^{m} c(i,j)}$$

This algorithm can be shown to converge to a local minimum of J.