

Introduction to belief functions

Clustering– project

Thierry Dencœux

1 ECM algorithm

Apply the ECM algorithm to

- The `Somalia companies` dataset (with variables `Growth_Potential_Index` and `Monthly_Revenue`);
- The `Globalization` dataset (2014 data).

Visualize the results and experiment with different parameter values (type of partition, α , etc.). For the `Globalization dataset`, print the names of the countries in the lower and upper approximations of each cluster. Find the pairs of contiguous clusters and choose them as focal sets. Compare the results to those of HCM and FCM.

2 EVCLUS

- Apply the EVCLUS algorithm to the `Somalia companies` and `Globalization` datasets. As before, visualize the results and experiment with different parameter values.
- For the `Somalia companies` dataset, compute the Gower distance matrix (taking into account all quantitative and qualitative attributes) and apply EVCLUS to this matrix.
- Apply EVCLUS algorithm to the `Synthetic control chart` dataset with Dynamic Time Warping (DTW) dissimilarity. Try introducing different numbers of must-link and cannot-link constraints and see how it improves the results.