

Acceleration and quantification with dimension reduction in fuzzy clustering

Tatsuki Akaki*, Yuichi mori, Masahiro Kuroda**, Masaya
Iizuka*****

*** Graduate School of Management, Okayama University of Science, Japan**

**** Okayama University of Science, Japan**

***** Okayama University, Japan**

Fuzzy clustering, e.g., fuzzy c-means clustering (FCM), is a useful tool in the variety of fields, but we sometimes meet the following two problems: computational time due to iterative convergence and difficulty of finding the latent clusters due to high dimensionality. For the former case, we have developed an acceleration algorithm for FCM using Wynn's vector epsilon algorithm (ve-FCM), which generates a new accelerated convergent sequence based on the original linearly convergent sequence in estimating two parameters, the membership matrix and the cluster centroid matrix, alternately. Some numerical experiments demonstrate that the ve-FCM accelerates the computation twice or more as faster as the original one. For the latter case, although the reduced k-means method is a well-known method to find clusters with the k-means clustering and dimension reduction simultaneously, this method is hard clustering for numerical data, but not soft clustering, not for categorical data. So we propose a modified fuzzy clustering method with dimension reduction for categorical data (catRFCM). The catRFCM quantifies the original categorical data and estimates low-dimensional cluster centers by implementing fuzzy c-means (FCM) with quantification and FCM with dimension reduction simultaneously. Numerical experiments are conducted to evaluate the performance of the catRFCM by comparing with, e.g., a ordinary FCM with the original categorical data as a numerical data, and a tandem method using quantification, dimension reduction, and fuzzy clustering in sequence.