

Approximate Clusters, Biclusters and n -Clusters in the Analysis of Binary and General Data Matrices

Boris G. Mirkin

National Research University Higher School of Economics, Moscow, Russia

Abstract. Approximate cluster structures are those of formal concepts and n -concepts with added numerical intensity weights. The talk presents theoretical results and computational methods for approximate clustering and n -clustering as extensions of the algebraic-geometrical properties of numerical matrices (SVD and the like) to the situations where one or most of elements of the solutions to be found are expressed by binary vectors. The theory embraces such methods as k-means, consensus clustering, network clustering, biclusters and triclusters and provides natural data analysis criteria, effective algorithms and interpretation tools.

Keywords: Approximate clusters, biclusters, n -clusters, Formal Concept Analysis

References

1. Mirkin, B.G., Rostovtsev, P.S.: Method for revealing associated feature subsets,. In Mirkin, B., ed.: Models for Summarization of SocioEconomic Data (Metody Agregirovania Sotsial'no-Economicheskoi Informatsii), Novosibirsk: Institute of Economics Press (1978) 107–112 (in Russian).
2. Ignatov, D.I., Gnatyshak, D.V., Kuznetsov, S.O., Mirkin, B.G.: Triadic formal concept analysis and triclustering: searching for optimal patterns. Machine Learning **101**(1-3) (2015) 271–302