# Cluster stability in a partition and applications

Talk

Kristian Sabo
Department of Mathematics, University of Osijek
ksabo@mathos.hr

(joint work with Rudolf Scitovski and Ivan Vazler)

A partition of the set $\mathcal{A} = \{a_i \in \mathbb{R}^n : i = 1, \ldots, m\} \subset \mathbb{R}^n$ into $k$ disjoint subsets $\pi_1, \ldots, \pi_k$, $1 \leq k \leq m$, such that

$$\bigcup_{i=1}^{k} \pi_i = \mathcal{A}, \qquad \pi_i \cap \pi_j = \emptyset, \quad i \neq j, \qquad |\pi_j| \geq 1, \quad j = 1, \ldots, k, \tag{1}$$

will be denoted by $\Pi(\mathcal{A}) = \{\pi_1, \ldots, \pi_k\}$, and the elements $\pi_1, \ldots, \pi_k$ of such partition are called *clusters in* $\mathbb{R}^n$. The problem of cluster stability in a partition $\Pi(\mathcal{A})$ is considered. A characterization of a well-separated partition is given and an operational criterion is provided that gives the possibility to measure the quality of cluster separability in a partition. We also give an application to a definition of a new separation index for estimation of the right number of clusters in a partition. In comparison with some other known separation indexes, this index showed its superiority. Several illustrative examples are mentioned as well.

MSC2010: 62H30, 68T10, 90C26, 90C27, 91C20, 47N10.

Keywords: clustering, data mining, cluster stability, stability balls, number of clusters, separation index.

Section: Numerical Analysis and Scientific Computing.

.