

# **k-MEANS CLUSTERING THROUGH CONTINUOUS OPTIMIZATION**

**J. Kogan**<sup>1</sup>

kogan@math.umbc.edu

*Dept. of Mathematics and Statistics,  
University of Maryland Baltimore County, USA*

The classical  $k$ -means clustering algorithm is probably one of the most popular techniques for data clustering. While partitioning a dataset into a prescribed number of clusters, the algorithm attempts to optimize a real valued function  $Q$  whose domain is the set of partitions of the dataset. This, in particular, prevents one from applying standard optimization techniques to the  $k$ -means clustering.

Alternatively the  $k$ -means clustering can be stated as an optimization of a non-convex and non-smooth real valued function  $F$  of a vector argument. Therefore, the  $k$ -means clustering problem belongs to the class of continuous optimization problems which combine two of the most difficult characteristics one encounters in optimization: non-convexity and non-smoothness. In this talk we focus on an approach handling the second difficulty, namely non-smoothness.

---

<sup>1</sup>Joint work with M. Teboulle (teboulle@post.tau.ac.il), School of Mathematical Sciences, Tel-Aviv University, Israel