
Summary: Given a set of entities associated with points in Euclidean space, minimum sum-of-squares clustering (MSSC) consists in partitioning this set into clusters such that the sum of squared distances from each point to the centroid of its cluster is minimized. A column generation algorithm for MSSC was given by O. du Merle et al. [SIAM J. Sci. Comput. 21, No. 4, 1485–1505 (1999; Zbl 1049.90129)]. The bottleneck of that algorithm is the resolution of the auxiliary problem of finding a column with negative reduced cost. We propose a new way to solve this auxiliary problem based on geometric arguments. This greatly improves the efficiency of the whole algorithm and leads to exact solution of instances with over 2,300 entities, i.e., more than 10 times as much as previously done.

MSC:
90C27 Combinatorial optimization
65K05 Numerical mathematical programming methods
91C20 Clustering in the social and behavioral sciences

Keywords: clustering; sum-of-squares; column generation; ACCPM

Software:
TSPLIB; minlpBB; UCI-ml; J-MEANS; Bonmin

Full Text: DOI Link

References:


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.