

Quadratic Euclidean 1-Mean and 1-Median 2-Clustering Problem with Constraints on the Size of the Clusters: Complexity and Approximability

A. V. Kel'manov[†], A. V. Pyatkin^{1,2,*}, and V. I. Khandeev^{1,2,**}

Received August 12, 2019; revised September 10, 2019; accepted September 16, 2019

Abstract—We consider the problem of partitioning a set of N points in d -dimensional Euclidean space into two clusters minimizing the sum of the squared distances between each element and the center of the cluster to which it belongs. The center of the first cluster is its centroid (the geometric center). The center of the second cluster should be chosen among the points of the input set. We analyze the variant of the problem with given sizes (cardinalities) of the clusters; the sum of the sizes equals the cardinality of the input set. We prove that the problem is strongly NP-hard and there is no fully polynomial-time approximation scheme for it.

Keywords: Euclidean space, clustering, 2-partition, quadratic variation, center, centroid, median, strong NP-hardness, nonexistence of FPTAS, approximation-preserving reduction.

DOI: 10.1134/S0081543821030123

[†]Deceased.

¹Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

²Novosibirsk State University, Novosibirsk, 630090 Russia
e-mail: *artem@math.nsc.ru, **khandeev@math.nsc.ru