Least-squares Consensus Clustering versus: (a) other Consensus Approaches and (b) K-Means

A. Shestakov[1] and B. Mirkin[2]

Abstract

We take on two criteria for consensus clustering proposed by Mirkin and Muchnik (1981, in Russian) and optimize them with similarity clustering approaches described in (Mirkin, 2005, 2012). Given a set of partitions R on the same entity set, one criterion is to find a partition r, that is behind those in R, which is akin to current concepts of ensemble consensus clustering. The other criterion is to build a partition r from R. Both can be equivalently reformulated as similarity clustering criteria; the first working over the conventional consensus matrix, the second over the summary projection matrix. We consider a number of recent clustering consensus methods: Voting Scheme (Weingessel, Dimitriadou, Hornik 2002), Borda Voting (Sevillano, Claudi Socoro, Alias 2009), Bayesian (Wang, Shan, Banerjee 2009), Fusion-Transfer consensus (Guenoche 2011), MCLA, CSPA and HGPA (Strehl, Ghosh 2002), and cVote (Ayad, Kamel 2010). For experiments, we take all three types of data: (a) UCI repository datasets, (b) specially drawn two-dimensional "ornaments", and (c) generated Gaussian cluster datasets. We evaluate found cluster partitions according to their similarity to the partition hidden in data. We address two issues:

1. How least-squares consensus algorithms fare in comparison with the others? Answer: The least-squares consensus algorithms outperform the others, usually up to a large margin.
2. Is it true that the least-squares k-means clustering criterion is a better criterion than consensus? Answer: No. in most situations, least squares consensus partition is closer to the hidden partition than that minimizing the k-means criterion. This shows that developing algorithms for reaching deep minima of k-means criterion may be a wrong idea.

Keywords

CONSENSUS CLUSTERING, LEAST SQUARES CONSENSUS, ONE-BY-ONE CLUSTERING, K-MEANS

NRU Higher School of Economics, Moscow, RF shestakoffandreyi@gmail.com · NRU Higher School of Economics, Moscow, RF bmirkin@hse.ru