Online Low-Rank Subspace Clustering

Jie Shen
Dept. of Computer Science

10/6/2016 at 11:00 am
Hill 482

Abstract

Low-Rank Representation (LRR) has been a significant method for segmenting data that are generated from a union of subspaces. It is also known that solving LRR is challenging in terms of time complexity and memory footprint, in that the size of the nuclear norm regularized matrix is n-by-n (where n is the number of samples). In this paper, we thereby develop a novel online implementation of LRR that reduces the memory cost from $O(n^2)$ to $O(pd)$, with p being the ambient dimension and d being some estimated rank $(d < p << n)$. We also establish the theoretical guarantee that the sequence of solutions produced by our algorithm converges to a stationary point of the expected loss function asymptotically. There are two crucial techniques involved: non-convex matrix factorization and basis dictionary pursuit. Interestingly, after we transform the LRR problem to its non-convex form, it turns out that it is essential to explore sparse patterns for the underlying data for the sake of an effective and discriminate clustering. Extensive experiments on synthetic and realistic datasets further substantiate that our algorithm is fast, robust and memory efficient.

Examination Committee: Prof. Ping Li (Chair), Prof. Pranjal Awasthi, Prof. Abdeslam Boularias and Prof. Zheng Zhang