Provable and Efficient Algorithms for Subspace Learning and Tracking

Praneeth Narayanamurthy

ABSTRACT

In the past decades, there has been an explosion in the amount of data that is generated. This calls for development of efficient algorithms to uncover useful information from massive datasets. Although several recent advances in computation allows for faster processing, efficient communication and storage and so on, the need of the hour is to develop intelligent algorithms that minimize resource utilization, and does so in a near real-time fashion. A commonly observed theme in the Signal Processing and Machine Learning literature is to exploit the fact that most real-world (extremely high dimensional) data exhibits a simple, succinct, low-dimensional representation. In other words, the data lies close to some low-dimensional structure of the ambient space. In this thesis, we consider two such low-dimensional structures: sparsity and low-rank. Specifically, we develop provable algorithms for the problem of Subspace Tracking (ST) under several constraints. First we study robust ST wherein the data is corrupted by arbitrary outliers. Next, we consider the setting where part of the data is missing (due to issues in transmission or storage). Finally, we develop algorithms that also deal with distributed data.