

Matrix Sketching over Streams

Mina Ghashami
DIMACS, Rutgers University

2/7/2018 at 11:00 am
CoRE 301

Abstract

It is common to represent data in the form of a matrix, and a large set of data analytic tasks rely on obtaining a low-rank approximation of the data matrix. Such approximations can be computed using the Singular Value Decompositions (SVD). In many scenarios, however, data matrices are extremely large and computing their SVD exactly is infeasible. Efficient approximate solutions exist for distributed setting or when data access otherwise is limited. In the data streaming model, the data points are presented to the algorithm one by one in an arbitrary order. The algorithm is tasked with processing the stream in one pass while being severely restricted in its memory footprint. At the end of the stream, the algorithm must provide a sketch matrix which is a good approximation of the original data. In this talk, we will discuss two recent matrix sketching methods over data streams.

Bio

Prior to postdoc, I was a PhD student at University of Utah, advised by Dr. Jeff Phillips. During PhD, I worked on designing low-rank approximations to matrices, mostly in streaming/online data models. My other research interests are dimensionality reduction, machine learning, and approximation algorithms.

I received my B.E. and M.E. in Software Engineering, specializing in Data Mining in Computer Engineering department of Sharif University of Technology.

Organizer(s): Pranjal Awasthi and Shubhangi Saraf