

Computer Science Seminar Series

National Capital Region

Discovering Latent Factors in High-dimensional Data via Spectral Methods

Speaker: Dr. Furong Huang
University of Maryland
Department of Computer Science
Friday, November 10, 2017
1:00PM- 2:00PM, NVC T3

Abstract

Latent variable models have a broad set of applications in domains such as social networks, natural language processing, computer vision and computational biology. Training them on a large scale is challenging due to non-convexity of the objective function. We propose a unified framework that exploits tensor algebraic constraints of the (low order) moments of the models. This versatile framework is guaranteed to estimate the correct model consistently and the spectral decomposition (matrix/tensor decomposition) proposed are embarrassingly parallel and has global convergence guarantees using SGD despite the non-convexity of the objective function. Topic Modeling will be discussed extensively, as well as user commonality inference in the large-scale social network using Mixed Membership Stochastic Blockmodel, and convolutional dictionary learning for text paraphrase embedding learning.

Biography



Furong Huang is an assistant professor at University of Maryland, Computer Science department. Her research interests lie in developing scalable and parallel algorithms for large-scale data using statistical models. She has worked on non-convex function optimization such as finding tensor decomposition with global convergence guarantee using stochastic gradient descent; developing fast detection algorithm to discover hidden and overlapping user communities in social networks; designing a parallel spectral tensor decomposition algorithm for detecting hidden topics in articles on Map-Reduce frameworks; and learning convolutional sparse coding models using tensor methods for extracting text sequence embeddings and image filter-bank. Beside pure statistical computation, Furong has applied her machine learning techniques to biology.