



THE WEIZMANN INSTITUTE OF SCIENCE
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 155 ,Ziskind Building
on Monday, Jun 25, 2018
at 16:00

NOTE SPECIAL DATE, TIME AND ROOM

Ariel Jaffe / Phd defence
Weizmann

Spectral methods for unsupervised ensemble learning and latent variable models

Abstract:

The use of latent variables in probabilistic modeling is a standard approach in numerous data analysis applications. In recent years, there has been a surge of interest in spectral methods for latent variable models, where inference is done by analyzing the lower order moments of the observed data. In contrast to iterative approaches such as the EM algorithm, under appropriate conditions spectral methods are guaranteed to converge to the true model parameters given enough data samples.

The focus of the seminar is the development of novel spectral based methods for two problems in statistical machine learning. In the first part, we address unsupervised ensemble learning, where one obtains predictions from different sources or classifiers, yet without knowing the reliability and expertise of each source, and with no labeled data to directly assess it. We develop algorithms to estimate the reliability of the classifiers based on a common assumption that different classifiers make statistically independent errors. In addition, we show how one can detect subsets of classifiers that strongly violate the model of independent errors, in a fully unsupervised manner.

In the second part of the seminar we show how one can use spectral methods to learn the parameters of binary latent variable models. This model has many applications such as overlapping clustering and Gaussian-Bernoulli restricted Boltzmann machines. Our methods are based on computing the eigenvectors of both the second and third moments of the observed variables.

For both problems, we show that spectral based methods can be applied effectively, achieving results that are state of the art in various problems in computational biology and population genetics.