Statistics SEMINAR

Title:

Statistical and Computational Guarantees for the EM Algorithm

Speaker:

Dr. Sivaraman Balakrishnan
Postdoctoral Researcher
Department of Statistics
University of California at Berkeley

Time & Place:

Thursday, January 29, 2015, 4:00-5:00pm
Room 133 SMI
(Cookies & Coffee @ 3:30 in Rm 1210 MSC)

Abstract and Bio:

The expectation-maximization (EM) algorithm is an iterative method for finding maximum-likelihood estimates of parameters in statistical models with unobserved latent variables. Along with Markov Chain Monte Carlo (MCMC) it is one of the two computational workhorses that provided much impetus for statistics in entering its modern "computation-intensive" phase. Much is known about the EM algorithm, its convergence properties, and its susceptibility to local optima. However, despite the existence of multiple fixed points, on a variety of statistical problems the EM algorithm has been observed empirically to perform well given either a reasonable initialization or with several random starts.

In this talk, I will introduce novel techniques to theoretically characterize some of the empirically observed behavior of the EM algorithm and give conditions under which the EM algorithm converges to near globally optimal parameter estimates. In particular, I will show the surprising result that for several canonical latent variable models there are large regions of the parameter space over which every fixed point of the EM algorithm is close to a population global optimum. I will conclude with a discussion of some of my other research interests, presenting a few vignettes from my work on clustering and topological data analysis.

Sivaraman Balakrishnan is a postdoctoral researcher in the Department of Statistics at the University of California at Berkeley, working with Martin J. Wainwright and Bin Yu. Prior to this he received his Ph.D. from the School of Computer Science at Carnegie Mellon University in 2013. His Ph.D. work was supported by several fellowships including the Richard King Mellon Fellowship and a grant from the Gates Foundation. He is broadly interested in problems that lie at the interface between computer science and statistics. Some particular areas that have provided motivation for his past and current research include the applications of statistical methods in computational biology, clustering, topological data analysis and non-parametric statistics.

THE UNIVERSITY of WISCONSIN MADISON