STATISTICS COLLOQUIUM

MONDAY, MARCH 19, 2012
TALK: 4:00 PM — SCIENCE CENTER RM. 309
RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

“Spectral Methods for Learning Graphical Models”

Sham Kakade
Microsoft Research, New England
University of Pennsylvania

ABSTRACT

This work presents a methodology for learning graphical models with hidden nodes through algebraic techniques (in particular, matrix decomposition and spectral methods), using independent samples of the observed variables. The talk focuses on tree models, and covers two aspects of the underlying learning problem: parameter estimation and structural learning. The underlying idea is to utilize the spectral decomposition of the second moment matrix to reveal the latent structure.

The first part is concerned with parameter estimation. Here, we present an efficient and provably correct algorithm for learning HMMs (i.e. recovering the correct HMM dynamics), with a sample complexity depending on some mild conditions of the underlying system. The algorithm is also simple, employing only a singular value decomposition and matrix multiplications, and does not suffer from local minimum issues in non-convex optimization (such as for more traditional approaches, including the EM algorithm), and it handles high dimensional observations and long range dependencies more easily. The method can be extended to estimating parameters for nonlinear systems and general tree structured graphical models with unobserved nodes.

The second part is concerned with structural learning, where we provide the Spectral Recursive Grouping algorithm, an efficient and simple procedure for recovering the underlying tree topology of a broad class of multivariate tree models with hidden nodes. Exact recovery of the tree structure can be established based on certain natural dependencies on statistical and structural properties of the underlying joint distribution.

Join work with: Daniel Hsu, Tong Zhang; Anima Anandkumar, Kamalika Chaudhuri, Le Song.