Special Topics Course

Statistics 700: Probabilistic graphical models

Instructor: XuanLong Nguyen
Department of Statistics
Fall 2009
1:00–3:00 pm Thursday

Course description. This course provides an introduction to the area of probabilistic and computational methods for the statistical modeling of complex, multivariate data. The emphasis will be on the unifying framework provided by graphical models, a formalism that merges aspects of graph theory and probability theory. Key topics include exponential families, hidden Markov, mixture and other latent variable models, Markov random fields, models for space-time data. A strong emphasis will be on the development of efficient algorithms for statistical inference using Markov Chain Monte Carlo (MCMC) and variational methods. These concepts will be illustrated using examples drawn from various application domains, including machine learning, signal and image processing, text analysis, communication theory, computational biology, etc.

Prerequisites. The prerequisites are previous coursework in linear algebra, multivariate calculus, and basic probability and statistics. Previous coursework in graph theory, information theory, optimization theory and statistical physics would be helpful but is not required. Familiarity with R, Matlab, Splus or a related matrix-oriented programming language will be necessary.

Structure/Evaluation. The course will meet once a week and will follow a regular lecture format. There will be bi-weekly homework assignments, due one week after being passed out. There will be a final course project in forms of either a survey paper, an application research or a methodological research project. Details will be available in the course's homepage.

Course homepage. `http://www.stat.lsa.umich.edu/~xuanlong/courses/stat700-f09/`
All announcements, homeworks, project information and data sets will be posted at this site.

Textbook. Book chapters from "An introduction to probabilistic graphical models" by M. I. Jordan will be provided. Additional reading assignments will be posted at the course homepage.

Tentative outline.

- Basics on graphical models, Markov properties, recursive decomposability, elimination algorithms
- Sum-product algorithm, factor graphs, semi-rings
- Frequentist and Bayesian methods
- Bayesian classification, linear models and generalized linear (GLIM) models, on-line methods
- Exponential family, sufficiency, conjugacy

- The EM algorithm
- Mixture models, conditional mixture models, hierarchical mixture models
- Hidden Markov models (HMM) and forward-backward algorithms
- Factor analysis, principal component analysis (PCA), canonical correlation analysis (CCA), independent component analysis (ICA)
- Kalman filtering and Rauch-Tung-Striebel smoothing
- Markov properties of graphical models
- Junction tree algorithm
- Chains, trees, factorial models, coupled models, layered models
- Importance sampling, Markov Chain Monte Carlo methods (Gibbs/ Metropolis-Hastings sampling)
- Variational inference algorithms: mean field, belief propagation, convex relaxations
- Model selection, marginal likelihood, AIC, BIC and MDL
- Nonparametric methods: Gaussian processes, Dirichlet processes, infinite mixture models
- Applications to signal processing, bioinformatics, communication, natural language processing, computer vision etc.