Probabilistic Graphical Models

CSCI 2950-P: Special Topics in Machine Learning, Spring 2013

Probabilistic graphical models provide a flexible framework for modeling large, complex, heterogeneous collections of random variables. Graphs are used to decompose multivariate, joint distributions into a set of local interactions among small subsets of variables. These local relationships produce conditional independencies which lead to efficient learning and inference algorithms. Moreover, their modular structure provides an intuitive language for expressing domain-specific knowledge, and facilitates the transfer of modeling advances to new applications.

After a brief introduction to their representational power, this course will provide a comprehensive survey of state-of-the-art methods for statistical learning and inference in graphical models. Our primary focus will be variational methods, which adapt tools from optimization theory to develop efficient, possibly approximate, inference algorithms. We will also discuss a complementary family of Monte Carlo methods, based on stochastic simulation.

Many course readings will be drawn from the draft textbook *An Introduction to Probabilistic Graphical Models*, in preparation by Michael Jordan. Advanced topics will be supported by tutorial and survey articles, and illustrated with state-of-the-art research results and applications. Overall grades will be assigned based on homework assignments combining statistical analysis and implementation of learning algorithms, as well as a final research project involving probabilistic graphical models. Students who took CSCI 2950-P in the Fall of 2011 may repeat for credit, as the topic has changed.

Prerequisites: Completion of an introductory course in statistical machine learning, such as Brown's *CSCI 1950-F: Introduction to Machine Learning* or *APMA 1690: Computational Probability and Statistics*. Sufficient comfort with calculus, linear algebra, and probability to read mathematically sophisticated research papers. Programming experience for homeworks and projects.

Administrative Information

Lectures: Tuesdays and Thursdays from 1:00-2:20pm, CIT room 506, 115 Waterman St.

Instructor:

Erik Sudderth (sudderth@cs.brown.edu; 401-863-7660)
Office Hours: Tuesdays 2:30-4:00pm, CIT room 509.

Graduate Teaching Assistant:

Jason Pacheco (pachecoj@cs.brown.edu)
Office Hours: Thursdays 2:30-4:00pm, CIT room 361.

Grading: Homework Assignments

Homework assignments will count towards 60% of overall grades. There will be four equally weighted assignments, each of which will be available for two weeks before its due date. Homeworks will involve a combination of mathematical derivations, algorithm design, programming, and real data analysis. Further details, and homework due dates, will be announced early in the semester.

Grading: Final Projects

The final project will count towards 40% of overall grades. Of these points, 5% will be based on a 1-3 page project proposal, due on March 22; 10% will be based on a short oral presentation, given on May 7; and 25% will be based on a technical report describing the results, due on May 13. This technical report should be between 6-10 pages long, in the style of top machine learning conferences. Although the results need not be sufficiently novel for publication, the presentation and experimental protocols should be of high quality. Projects which apply graphical models to the student's own research interests are particularly encouraged.

Tentative Syllabus

Graphical models directed, undirected, and factor graph representations; factorization, Markov properties; common temporal, spatial, hierarchical, and relational models

Exact inference variable elimination; message passing algorithms; junction tree algorithm

Exponential families sufficient statistics, ML estimation; parameter estimation in directed and undirected graphical models; iterative scaling algorithms; conjugate priors

Gaussian graphical models Bayesian networks, Markov random fields; inference, Kalman filters; linear dynamical systems, factor analysis, probabilistic PCA

Sequential Monte Carlo importance sampling; particle filters; non-sequential models

Markov Chain Monte Carlo Metropolis-Hastings algorithm, detailed balance; Gibbs samplers; Rao-Blackwellization, blocking, auxiliary variables

Variational methods inference as optimization; entropy, information; mean field methods; parameter estimation, variational EM, variational Bayes

Loopy belief propagation variational interpretations, Bethe entropy approximation; historical perspectives, implementation and applications; reparameterization and convergence

Advanced variational methods convex relaxations, reweighted BP algorithms

MAP Estimation max-product algorithm; loopy max-product BP; variational interpretations, linear programming relaxations