Foundations of Graphical Models

David M. Blei
Columbia University

October 14, 2014

Description and Prerequisites

We will study how to use probability models to analyze data, focusing on the mathematical details of the models and the algorithms for computing with them. We will study both the foundations and advanced methods, such as large scale inference, model diagnostics and selection, and Bayesian nonparametrics. Our goals are to understand the cutting edge of modern probabilistic modeling, to begin research that makes contributions to this field, and develop good practices for specifying and applying probabilistic models to analyze real-world data.

Prerequisites. You should know basic probability and statistics, calculus, and some optimization. You should be comfortable writing software to analyze data and learning about new tools for that purpose. For example, you should be familiar with a statistical programming language such as R and a scripting language such as Python.

Requirements and Grades

There are two requirements for the course: a weekly paper and a project report.

- Each week, you will write what you thought about the reading. (Later, you will also write about your progress on the class project.) These papers can be up to one page; they can be as short as one paragraph. (Short is good.) They are required, but not graded. They must be handed in during Wednesday's class. Late papers are not accepted.
- Each student will complete a class project. Most projects will involve using and possibly developing probabilistic models to analyze real-world data. (Though less common, some projects will develop novel theoretical research in graphical models.) Ideally, the project will be connected to your graduate research.

> A report about your project is due at the end of the semester. I grade reports on both content and writing quality. Two good books about writing are Strunk and White (1979) and Williams (1981).

Your course grade will be mostly based on the final report. But note the weekly papers are required. Missing a significant number of papers will result in a low grade. Prepare all written work using LaTeX; I will distribute LaTeX templates.

Logistics

The instructor is David Blei (`david.blei@columbia.edu`).

The course meets M and W from 1:10PM - 2:25PM.

Prof. Blei's office hours are W 2:30PM - 4:00PM in his office (703 CEPSR).

Syllabus and readings

Many readings will be from Jordan (2009) or Bishop (2006). Some readings will be from the other excellent books on these subjects: Gelman et al. (1995), Koller and Friedman (2009), and Murphy (2013). We will also read research papers.

Readings will be posted to the website. When available, we will provide PDFs online. Otherwise, photocopies will be available at Prof. Blei's office (703 CEPSR).

Below is a tentative schedule of the lectures.

1. Introduction
2. Basic concepts in probability
3. The semantics of graphical models
4. D-separation and conditional independence
5. The elimination algorithm
6. Tree propagation (and hidden Markov models) I
7. Tree propagation (and hidden Markov models) II
8. Probability models, data, and statistical concepts I
9. Probability models, data, and statistical concepts II
10. Bayesian mixtures of Gaussians and the Gibbs sampler I
11. Bayesian mixtures of Gaussians and the Gibbs sampler II
12. Markov chain Monte Carlo I
13. Markov chain Monte Carlo II
14. The exponential family, conjugacy, and mixtures of exponential families

15. Mixtures and mixed membership models (including topic models)
16. Matrix factorization: Gaussian, Poisson, exponential family
17. Special guest, Bob Carpenter: Applied modeling in Stan
18. Time series models: Hidden Markov models and state-space models
19. Spatial models
20. Regression: Linear and logistic, generalized linear models, regularization I
21. Regression: Linear and logistic, generalized linear models, regularization II
22. Hierarchical models, multi-level models, empirical Bayes
23. Variational inference (and a word about EM) I
24. Variational inference (and a word about EM) II
25. An introduction to Bayesian nonparametrics
26. Scalable inference with stochastic variational inference
27. Model checking and revision with posterior predictive checks

Lecture notes and slides will be posted to the website.

References

Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer New York.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.

Jordan, M. (2009). *An Introduction to Probabilistic Graphical Models*. In preparation.

Koller, D. and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Murphy, K. (2013). *Machine Learning: A Probabilistic Approach*. MIT Press.

Strunk, W. and White, E. (1979). *Elements of Style*. Longman Press.

Williams, J. (1981). *Style: Towards Clarity and Grace*. University of Chicago Press.