Foundations of Graphical Models

David M. Blei
Columbia University
Fall, 2015

Description

This is a PhD-level course about how to develop and use probability models. We will study their mathematical properties, algorithms for computing with them, and applications to real problems. We will study both the foundations and modern methods in this field, such as large-scale inference 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 to develop good practices for specifying and applying probabilistic models.

Prerequisites. You know basic probability and statistics, calculus, and some optimization. You are comfortable writing software to analyze data and learning about new tools for that purpose. You are familiar with a good programming language for statistics, such as R or Python.

Requirements and Grades

There are three requirements: a weekly paper, occasional homework, and a final project.

- Each week, you 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.
- The homework assigments involve problems and data analysis. There are about three assignments during the semester.
- You complete a class project. Most projects 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, your project connects to your research.

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

Please prepare all written work using LaTeX. We will distribute LaTeX templates.

Your course grade is mainly based on your final project report and homework. (But note the weekly papers are required. If you miss many weekly papers then you will get a low grade.)

Syllabus and readings

Many readings will be from Jordan (2003) 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 and lecture notes 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).

Introduction

1. Introduction (Blei, 2014)
2. Probability: Basic concepts and review

The basics of graphical models

3. 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

Latent variable models

8. Models, data, and statistical concepts I
9. 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. Exponential families, conjugacy, and mixtures of exponential families I
13. Exponential families, conjugacy, and mixtures of exponential families II
14. Mixed-membership, topic models, and variational inference I
15. Mixed-membership, topic models, and variational inference II
16. Matrix factorization and recommendation systems I
17. Matrix factorization and recommendation systems II

Conditional models

18. Regression: Linear and logistic
19. Generalized linear models
20. Hierarchical models, robust models, and empirical Bayes I
21. Hierarchical models, robust models, and empirical Bayes II

Advanced ideas in approximate posterior inference

22. Markov chain Monte Carlo I
23. Markov chain Monte Carlo II
24. Variational inference I

25. Variational inference II

Other topics and summary

26. An brief introduction to Bayesian nonparametrics
27. Summary (and wiggle room)

References

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

Blei, D. (2014). Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Application*, 1:203–232.

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

Jordan, M. (2003). *An Introduction to Probabilistic Graphical Models*.

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.