Foundations of Graphical Models

David M. Blei
Columbia University

September 8, 2020

Description. *Foundations of Graphical Models* is a PhD-level course about how to design and use probability models. We study their mathematical properties, algorithms for computing with them, and applications to real problems. We study both the foundations and modern methods in this field. Our goals are to understand probabilistic modeling, to begin research that makes contributions to this field, and to develop good practices for building and applying probabilistic models.

Prerequisites. The prerequisites are: knowledge of basic probability and statistics, calculus, and some optimization; comfort writing software to analyze data; familiarity with a good programming language for statistics and machine learning, such as R or Python.

Requirements and Grades. The requirements are weekly reader reports, several homework assignments, and a final project.

- Each week, you write what you thought about the reading. (Later in the semester, you 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 Tuesday's class. Late papers are not accepted.

- The homework assigments involve problems, programming, and data analysis. There will be about three assignments throughout the semester.

- The main requirement is the class project. Most projects involve using and developing probabilistic models to analyze real-world data. Some projects will develop novel theoretical research in probabilistic models. Ideally, your project connects to your doctoral research.

  A project report is due at the end of the semester. There will also be some intermediate assignments to check your progress. We grade your project on both content and writing quality.

Please prepare all written work using the LaTeX templates we provide.

Your grade is mainly based on your final project. Also contributing to your grade are your homework, completion of the weekly papers, and participation in the class.

Syllabus

Below are the subjects we cover and in what order. (It may change.)

Introduction

1. Introduction
2. Probability: Basic concepts and review
3. The ingredients of probabilistic models I
4. The ingredients of probabilistic models II

Latent variable models

5. Bayesian mixtures and the Gibbs sampler I
6. Bayesian mixtures and the Gibbs sampler II
7. Mixed-membership, topic models, and variational inference I
8. Mixed-membership, topic models, and variational inference II
9. Matrix factorization, recommendation systems, and efficient MAP I
10. Matrix factorization, recommendation systems, and efficient MAP II
11. Exponential families, conjugacy, and generalized linear models I
12. Exponential families, conjugacy, and generalized linear models II
13. Deep learning: A probabilistic perspective I
14. Deep learning: A probabilistic perspective II
15. Deep generative models and black box variational inference
16. Hierarchical models, robust models, and empirical Bayes

The basics of graphical models

17. Graphical models I : Semantics
18. Graphical models II: d-separation and independence
19. Graphical models III: Tree propagation

Advanced ideas

20. Advanced topics in variational inference
21. Model checking and model diagnostics
22. An introduction to causality I
23. An introduction to causality II
24. Summary (and wiggle room)