Markov Mix and Match

(On Markov chain mixing and applications to algorithms)

ORIE 7591: Syllabus

Spring 2018

Essential Course information:

- *Physical and Virtual Location*
  Class location: Phillips 213
  Website: http://people.orie.cornell.edu/sbanerjee/ORIE7591/orie7591s18.html
- *Lecture Schedule*
  The official slot for the class is MF 8.50-955am. However, due to travel constraints, **we will start on 29th January, and front-load the course by having 2 hour lectures (from 8.40am to 11am, with a break in between) for the first 11 lectures** (till March 5th, with a possible extra lecture during/after February break). I am hoping this also indirectly helps research, as you will have seen all the main topics by March, and have the rest of the semester to focus on your projects.
- *Instructor*: Sid Banerjee
  E-mail: sbanerjee@cornell.edu

Course description:

The goal of this course is to provide an introduction to the theory of Markov chains, with an emphasis on their applications to algorithms, learning and control. The main tool we will focus on is the *mixing* properties of finite-state, discrete-time, reversible Markov chains, i.e., how 'fast' a given chain converges to its stationary distribution from any starting state. This may seem like a narrow question; however, it is central to the analysis of *Markov Chain Monte Carlo*, with connections to classical questions in statistical physics, numerical integration, randomized algorithms and combinatorial optimization, and increasingly, applications in distributed computing, large-scale graph mining, reinforcement learning and high-dimensional optimization.

For the major part of this course, we will survey some techniques for bounding the mixing time of random walks, and see how they are used to analyze some classical MCMC problems.

1. **Fundamentals of Markov Chains**: Markov chains, stationary distributions, ergodicity, reversibility, mixing time, cover time.
2. **Introduction to MCMC and applications**: Metropolis-Hastings, Gibbs sampler and Glauber dynamics, shuffling, approximate counting, combinatorial optimization
3. **Probabilistic methods for mixing time bounds**: Strong stationary times, coupling, path coupling, perfect sampling (coupling from the past)
4. **Combinatorial methods for mixing**: Multi-commodity flows, canonical paths
5. **Analytic and geometric methods for mixing**: the Perron-Frobenius theorem, conductance and isoperimetric inequalities, the Jerrum-Sinclair theorem, expanders.
6. **Excursions of random walks**: Spitzer's lemma, Martingale bounds, evolving sets

In the remaining time, we will see how these techniques are used in various applications. The following are a list of potential topics, which we will sample from based on the interest of the class (alternately, they form a good set of topics for student projects):

1. **Algorithms for approximate counting**: Matchings and colorings, approximating the permanent, network reliability
2. **Data mining in large networks**: Random walk algorithms for graph counting, network centrality metrics
3. **Inference in graphical models**: Sampling from Log-concave distributions, large-scale Gibbs sampling
4. **Applications in optimization and control**: Volume and integration of convex bodies, randomized linear algebra, fastest mixing Markov Chains, randomized algorithms for reinforcement learning
5. **Connections to statistical physics**: The Ising model and the partition function, belief propagation and MCMC, local algorithms and correlation decay

Prerequisites:

This is a topics course, and so may be more advanced than an intro graduate probability/algorithms course; that said, I will not assume extensive background knowledge. You should have some prior exposure to basic probability and optimization and algorithms (ideally at the level of ORIE 6500 and ORIE 6300/CS 6820; however you can survive with less background). In particular, we will mostly focus on discrete-state Markov chains, so you do not need to know measure theoretic probability. Send me a mail if you are concerned about having the appropriate prerequisites.

Course Grading

Grades for the course will be based on assignments, scribe notes and a project. During the first half of the semester, I will maintain a list of homework problems, which will be updated after each class – these are mainly to clarify topics we cover in class, and point out related topics. Students will also be required to scribe one or two lectures. Finally, all students should work on a project on topics related to what we cover in class; the project can either be a survey of some of the applications that we do not cover, or (ideally!) original research. Ideally, students should relate the project to their own research interests; I am always available and happy to discuss ways of doing so.

For the project, students should discuss with me and submit a 1 page proposal by **Friday, March 9, 2018** (this is when we hope to finish the main topics in the class). The final class (May 7th, 2016) will be kept aside for student project presentations.

References:

There are two excellent sources for a lot of the techniques we will cover:

- Reversible Markov Chains and Random Walks on Graphs by Aldous and Fill
- Markov Chains and Mixing Times by Peres, Levin and Wilmer

The first is a remarkable unfinished monograph, which has been online since the 1990s, and inspired a lot of the work in this field. The second is a more recent book which gives a nice and accessible introduction to many of the topics we will study.

For more recent material and applications, I will maintain a list of papers and online notes – some of these we will cover in class, while others can serve as a starting point for projects.