ACM 206: Monte Carlo methods for scientific computing

Course Syllabus – Spring / 2023
Computing & Mathematical Sciences, California Institute of Technology

Course Instructor

Robert J. Webber (he/him), in collaboration with Joel A. Tropp (he/him)
Contact Information: *rwebber@caltech.edu*
Course Hours: T/Th 1:00 – 2:25 in the Annenberg lab (ANB 104)
Office Hours: Th 2:30 – 3:30 in the Annenberg conference room (ANB 121)
TA: Joseph Cachaldora, *jcachald@caltech.edu*

Course Description

This course introduces Monte Carlo methods with applications in Bayesian computing and rare event sampling. Topics include Markov chain Monte Carlo (MCMC), Gibbs samplers, Langevin samplers, MCMC for infinite-dimensional problems, convergence of MCMC, parallel tempering, umbrella sampling, forward flux sampling, and sequential Monte Carlo. Emphasis is placed both on rigorous mathematical development and on practical coding experience.

Recommended Prerequisites

Linear algebra at the level of ACM104 or ACM107; probability theory at the level of ACM116 or ACM117; some programming experience.

Course Welcome

Welcome to ACM 206. I'm delighted that you've chosen to participate in this experimental course on Monte Carlo methods!

There are two parts to the course: Tuesday lectures and Thursday problem sessions. On Tuesdays, I will give a blackboard lecture in which I teach you about contemporary Monte Carlo algorithms. On Thursdays, the class becomes a choose-your-own-adventure, in which you can either code up a Monte Carlo algorithm or solve math problems to analyze a Monte Carlo algorithm. We'll compare and discuss solutions at the end.

If you're a theoretician at heart, don't worry – you can participate in all the lectures and the problem sessions without writing a single line of code. If you enjoy coding, don't worry – there'll be opportunities to practice coding on Thursdays, and I'll be sharing my own code + simulations as well.

I hope this course will be fun, engaging, and helpful for your research / industrial endeavors. Don't hesitate to talk to me after class, and I'd be happy to answer any questions about Monte Carlo algorithms or work through any mathematical / coding problems together.

Learning Outcomes

By the end of this course, students will be able to:

- Identify the standard Monte Carlo algorithms that can used for high-dimensional inference problems.
- Understand when and why Monte Carlo algorithms can be expensive to run.
- Gain experience with enhanced methods (parallel tempering, umbrella sampling, forward flux sampling, etc.) for speeding up Monte Carlo.
- Apply the major limit theorems (Central Limit Theorem; Law of the Iterated Logarithm; Law of Large Numbers), which govern the convergence rate of Monte Carlo algorithms.

Required Text

Our main textbook is "Monte Carlo Strategies in Scientific Computing" by Jun Liu. If you enjoy owning physical copies of your favorite reference books, then I would recommend buying a copy from Springer Link for $40; otherwise, you can download a PDF version of this textbook from Springer Link for free. I will be supplementing this textbook with additional readings and with my own notes.

Course Website or Learning Management System

Additional recommended readings will be posted on Canvas.

Assessment Rubric

Grades are 100% participation based. If you attend the lectures and participate actively in the problem sessions, I will assign you the highest grade.

Attendance and Participation

Since this is a participation-based course, you are expected to attend class (particularly, the problem sessions) in order to earn the highest grade. These problem sessions are the setting in which you will do the most learning, since math and Monte Carlo are not spectator sports.

However, I recognize that life happens (weddings, funerals, can't-get-out-of-bed days, etc.) and don't expect 100% attendance. If you attend 80% of the lectures and problem sessions, I won't raise an eyebrow. If attending 80% of the lectures and problem sessions becomes a challenge, come talk to me and let me know what's happening: we can work it out.

Collaboration Policy

I am delighted when you work together with your classmates during the problem sessions. It is often easier to present your solutions clearly when you have already been discussing a problem with one of your classmates.

Course Schedule

Course schedules including lecture topics, associated readings and homework (if *appropriate to include in this table for your course) can be clearly laid out in a table format. Include key dates for exams, project milestones, etc.*

Week	Date	Lecture Topic	Associated Readings
1	April 4	Simulating random variables: quantile functions, Box Muller, multivariate Gaussian, rejection sampling, importance sampling	Liu Chs. 1-2
1	April 6	Problem session	
2	April 11	Simulating random geometric processes: (taught by Eliza O'Reilly)	
2	April 13	Problem session	
3	April 18	Simulating random continuous-time processes: Brownian motions, diffusion processes, continuous-time Markov chains, fractional Brownian motions	
3	April 20	Problem session	
4	April 25	Splitting and killing: Russian roulette sampling, self-avoiding walks, forward flux sampling	Liu Chs. 3-4
4	April 27	Problem session	
5	May 2	Discrete MCMC: Metropolis samplers, Gibbs samplers, cluster samplers	Liu Chs. 5-8
5	May 4	Problem session	
6	May 9	Theory of MCMC: Markov chain convergence theory	Liu Chs. 12-13
6	May 11	Problem session	
7	May 16	Continuous MCMC: Langevin dynamics, hybrid Monte Carlo, No-U-Turn-Sampler (NUTS)	Liu Ch. 9
7	May 18	Problem session	
8	May 23	Strange MCMC: trans-dimensional samplers, functional samplers, path samplers	
8	May 25	Problem session	
9	May 30	Ensemble MCMC: parallel tempering, umbrella sampling, AIES	Liu Chs. 10-11
9	June 1	Problem session	
10	June 6	Optional lecture: randomized numerical linear algebra	
10	June 8	Optional problem session	