The University of California - Los Angeles
Department of Statistics

Monte Carlo Methods for Optimization

Statistics 202C
Professor: Mark S. Handcock

Syllabus and Description

Professor:
Mark S. Handcock
8141 Mathematical Sciences Bldg
Office Hours:
see the Bruin Learn *Home* page
held in-person and on zoom simultaneously
Other times by arrangement. Clearly
composed questions sent to the
"handcock@stat.ucla.edu"
will receive written replies

Motivation and Synopsis

During the twentieth century, the development of statistical computing played a crucial facilitating role for the growth of the statistics discipline and the adoption of statistical methods within the scientific community and beyond. In the twenty-first century digital age, the amounts of data available for statistical analysis has grown tremendously, yielding new opportunities for statistical computing, as well as new challenges. Statistical computing constitutes an important part of a statistics education, and is highly valuable for statisticians in both academia and industry.

This graduate level course introduces Monte Carlo methods for simulation, optimization, estimation, learning and complex landscape visualization, including: Importance sampling; Sequential importance sampling; Markov chain Monte Carlo (MCMC) sampling techniques including Gibbs samplers, Metropolis/Hastings and various improvements; Simulated annealing; Exact sampling techniques; Convergence analysis; Data augmentation; Cluster sampling, such as Swendsen-Wang and SW-cuts; Hamiltonian and Langevin Monte Carlo; Equi-energy and multi-domain sampler; and Techniques for mapping complex energy landscapes.

Prerequisites

- Stat 202B Matrix Algebra and Optimization.
- People who didn't take 202B before may still take this class as long as they have background on matrix algebra, probability theory, and programming skills. To do this attend the first classes and we can assess if this is advisable.

Structure of the Course

There will be two lectures per week.

Textbooks

[**BZ**] Barbu, Adrian and Zhu, Song-Chung
Monte Carlo Methods (2020).
Springer
Required and available online as a ebook from the library, free for UCLA students.
`https://link.springer.com/book/10.1007/978-981-13-2971-5`

[**G**] Robert, Christian P. and Casella, George
Monte Carlo Statistical Methods (2004).
Springer
Required and available online as a ebook from the library, free for UCLA students.
`https://link.springer.com/book/10.1007%2F978-1-4757-4145-2`

[**EH**] Liu, Jun S.
Monte Carlo Strategies in Scientific Computing (2008).
Springer
Required and available online as a ebook from the library, free for UCLA students.
`http://link.springer.com/book/10.1007%2F978-0-387-76371-2`

Students must be connected to the UCLA network to obtain their free download. Students who would like to download the textbook off-campus may do so by connecting the the UCLA network via VPN `https://www.bol.ucla.edu/services/vpn/all.html`.

In addition to these books, there are a multitude of books covering pieces of the course content and from varying perspectives. I suggest you use one or more of the following.

Background on `R`

[**W**] Wickham, Hadley.
Advanced R. (2014).
Chapman and Hall/CRC.
Required and available online as a ebook from the library, free for UCLA students.
`http://www.crcnetbase.com/doi/book/10.1201/b17487`

[**M**] Maillardet, Robert, Owen Jones, and Andrew Robinson.
Introduction to Scientific Programming and Simulation Using R. (2014).
Chapman and Hall/CRC.

Required and available online as a ebook from the library, free for UCLA students.
`http://www.crcnetbase.com/doi/book/10.1201/9781420068740`

[**JC**] Chambers, John M.
Software for Data Analysis: Programming with R. (2008).
Springer: New York.
Required and available online as a ebook from the library, free for UCLA students.
`http://link.springer.com/book/10.1007/978-0-387-75936-4/page/1`

Other Resources

You can read the other books with different perspectives online for free from any UCLA account, starting from:
`proquest.safaribooksonline.com/search?q=BOOKTITLE%20r`

Course Webpage and Discussion Forum

The course has a webpage through the UCLA Bruin Learn The webpage will be continuously updated throughout the course with handouts, homework assignments and solutions. Users sign in to Bruin Learn with their UCLA Logon IDs.

I will be using Bruin Learn Discussion Board to provide discussion of issues in class and related questions. For questions that might be of interest to other students, please use Bruin Learn rather than solely emailing me. There other students and the TA can answer questions in addition to me. Example of questions are about interesting articles you have seen in the media, problems with access to resources, homework or computer questions. Enjoy!

Please regularly consult this class's Bruin Learn Bruin Learn home page, and the archive of the Announcements mailing list. It will contain lecture notes, homework, solutions and course information.

Computer Usage and Software

The computer is the scientific laboratory of the applied researcher in quantitative fields. As such this course requires the student to develop a degree of comfort and competence "in the lab".

Course Requirements and Grades

- 20% Homeworks (2 assignments, 10% each, none are dropped)
- 45% Three small projects
- 35% Final project (written report)

Syllabus of the Course

The syllabus of the course will develop on the following weekly schedule. The some later topics may not be reached and we will make choice among them toward the end of the quarter.

Chapter	Contents discussed
1	Introduction to Monte Carlo Methods 1. Monte Carlo methods in science and engineering – Simulation, estimation, sampling, optimization, learning, and visualization. 2. Examples of the use of Monte Carlo 3. Ideas of Monte Carlo
2	Sequential Monte Carlo 1. Importance sampling and weighted samples 2. Advanced importance sampling techniques 3. Framework for sequential Monte Carlo (selection, pruning, resampling, ...) 4. Application: particle filtering in object tracking, Monte Carlo Tree Search
3	Backgrounds on Markov Chains 1. The transition matrix 2. Topology of transition matrix: communication and period 3. Positive recurrence and invariant measures 4. Ergodicity theorem
4	Metropolis methods and its variants 1. Metropolis algorithm and the Hastings's generalization 2. Special case: Metropolized independence sampler 3. Reversible jumps and trans-dimensional MCMC
5	Gibbs sampler and its variants 1. Gibbs sampler 2. generalizations: Hit-and-run, Multi-grid, generalized Gibbs, Metropolized Gibbs 3. Data association and data augmentation 4. Slice sampling

Course Requirements and Grades

- 20% Homeworks (2 assignments, 10% each, none are dropped)
- 45% Three small projects
- 35% Final project (written report)

Homework

There will be regular homeworks and small projects both the theory and real computation.

None of the scores will be dropped. It is your responsibility to verify that your work successfully uploaded by the deadline. All homework assignments will be posted. Students will submit the solution files electronically via *gradescope* from the Bruin Learn *Assignments* page.

Students are free to discuss homework problems and solutions. Discussing the contents of the course with fellow students can be a valuable element of the learning process, and doing so is therefore generally encouraged. However, each student must hand in their own solutions, and the student should, if asked, be able to explain the solutions.

Chapter	Contents discussed
6	Clustering sampling
	1. Ising/Potts models
	2. Swendsen-Wang and clustering sampling
	3. Three interpretations of the SW method
7	Langevin Dynamics
	1. Hamiltonian Monte Carlo
	2. Langevin dynamics used in machine learning
	Gibbs Reaction and Diffusion equations, Alternative Back-propagation
8	Convergence analysis
	1. Monitoring and diagnosing convergence
	2*. Contraction coefficient
	3. Puskin's order
	4*. Eigen-structures of the transition matrix
	(Perron-Frobenius theorem, spectral theorem)
	5. Geometric bounds
	6*. Exact analysis on independence Metropolized Sampler (IMS)
	7*. First hitting time analysis and bounds for IMS (paper)
	8. Path coupling techniques.
	Bounds for Gibbs sampler and Swendson-Wang algorithm (paper).
	* discussed in previous Chapters.
9	Exact sampling
	1. Coupling from the past (CFTP)
	2. Bounding chains
10	Advanced topics
	1. Equi-energy and mult-domain sampler
	2. Wang-Landau algorithm
	3. Attraction-Diffusion Algorithm
	4. Mapping the energy landscape and case studies
	5. Visualization of object recognition and the image universe
	6. Landscapes for curriculum learning

Late Policy for Homework (silly but necessary)

There is a 10 minute grace period. Submissions up to 10 minutes late will be accepted with no penalty.

Homework assignments submitted 11 minutes late or more will be accepted with penalty. There is a minimum deduction of 1 point for being less than one hour late. An additional 0.5-point deduction will be taken for each additional hour it is late.

Thus, an assignment that is between 11 and 59 minutes late will receive a 1-point deduction. An assignment that is between 1 hour and 1:59 late will receive a 1.5-point deduction. An assignment that is between 2 hours and 2:59 late will receive a 2-point deduction, and so on.

I strongly advise uploading the homework to the Bruin Learn well in advance of the deadline in case there are connectivity problems or server issues.

Academic Integrity

As a student and member of the University community, you are here to get an education and are, therefore, expected to demonstrate integrity in your academic endeavors. All students must uphold University of California Standards of Student Conduct as administered by the Office of the Dean of Students. Students are subject to disciplinary action for several types of misconduct, including but not limited to: cheating, multiple submissions, plagiarism, prohibited collaboration, facilitating academic dishonesty, or knowingly furnishing false information. You may have assignments or projects in which you work with a partner or with a group. For example, you are welcome, and even encouraged, to work with others to solve homework problems. Even though you are working together, the assignment you submit for a grade must be IN YOUR OWN WORDS, unless you receive specific instructions to the contrary. For more information about academic integrity, please go to www.deanofstudents.ucla.edu.

I welcome comments or suggestions about the course at any time, either in person, by letter, or by email. Please feel free to use these ways make comments to me about any aspect of the course.

Support

Title IX prohibits gender discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking. Students who have experienced sexual harassment or sexual violence can receive confidential support and advocacy at the CARE Advocacy Office for Sexual and Gender-Based Violence, 1st Floor Wooden Center West, CAREadvocate@caps.ucla.edu, (310) 206-2465. You can also report sexual violence or sexual harassment directly to the University's Title IX Coordinator, Kathleen Salvaty, 2241 Murphy Hall, titleix@conet.ucla.edu, (310) 206-3417.

I welcome comments or suggestions about the course at any time, either in person, by letter, or by email. Please feel free to use these ways make comments to me about any aspect of the course.