CMPE 58N: Monte Carlo methods for data analysis and scientific computing

Instructor: Ali Taylan Cemgil
Boğaziçi University, Department of Computer Engineering
Istanbul, Turkey

Fall 2009-2010

Course Homepage

`http://www.cmpe.boun.edu.tr/courses/cmpe58N/fall2009`

Catalog Description

Generating random variates, Basic principles, Rejection, Reweighting and variance reduction, Importance sampling, Rejection control, Sequential Monte Carlo, Metropolis algorithm, Reversible Jump, Gibbs sampler, Ising models and Boltzman machines, Clustering methods and Swedsen-Wang, General conditional sampling, Hybrid Monte Carlo, Multilevel sampling and Optimisation, Simulated annealing and Bridging, Population Monte Carlo, Markov chains and convergence, Propp-Wilson algorithm, Perfect sampling

Course Description

Monte Carlo methods are stochastic simulation based algorithms designed to compute answers to problems where exact solutions are intractable and take exponential time to compute. In strict computer science sense, those methods are not algorithms as they provide an answer only asymptotically – often we don't know when and if they converge and provide the desired solution. Yet, these techniques perform remarkably well in practice. Historically, the first application of Monte Carlo was during the Manhattan project for development of the atomic bomb. With the increase of computational power, these techniques are now used in more "constructive" branches of science and engineering, including aerospace, computer vision, network analysis, speech recognition, robot navigation, bioinformatics to name a few.

The scope of this course is to review the following fundamental aspects in Monte Carlo computations

1. Model construction
2. Design of strategies for inference
3. theoretical aspects (convergence proofs, performance analysis)

Our exposure will be primarily slanted towards inference strategies.

In particular, we will study Markov Chain Monte Carlo methods and Sequential Monte Carlo. Our ultimate aim is to provide a basic understanding of computational techniques based on Monte Carlo simulations and associated concepts such that the students can orient themselves in the relevant literature and understand the current state of the art.

Topics

```
Introduction, Course structure, Motivating Examples, Applications, Law of Large numbers, CLT
Probability and Statistics Review, Random Number generation, Rejection sampling
Importance sampling
Markov Chains, Stochastic Processes, Discrete State Space Markov Chains
The Metropolis-Hastings Algorithm, The Gibbs Sampler
Example applications with MH and Gibbs
```

```
Model selection, Change Point models
Reversible Jump,
State space models, Sequential Monte Carlo
Annealed Importance sampling, SMC Samplers,
Optimisation, Simulated Annealing, Iterative Improvement,
```

Textbooks

Handouts and relevant chapters from the following books:

- Liu, J.S., Monte Carlo Strategies in Scientific Computing, Springer.
- Adam M. Johansen and Ludger Evers (edited by Nick Whiteley), Monte Carlo Methods, Lecture notes, University of Bristol
  `http://www.maths.bris.ac.uk/~manpw/teaching/notes.pdf`
- David MacKay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press
  `http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html`

Other relevant books for reference are

- C. M. Grinstead and J. L. Snell. Introduction to probability
  `http://www.dartmouth.edu/~chance`
- Gilks, W.R., Richardson, S. and Spiegelhalter, D. Markov Chain Monte Carlo in Practice, Chapman and Hall.
- Arnaud Doucet, Nando De Freitas and Neil J. Gordon (eds), Sequential Monte Carlo in Practice, Springer.
- Robert, C.P. and Casella, G., Monte Carlo Statistical Methods, Springer-Verlag.
- Jean-Michel Marin and Christian P. Robert, Bayesian Core: A Practical Approach to Computational Bayesian Statistics, Springer

Prerequisite

CmpE 343 (Introductory Probability and Statistics) or equivalent

Cmpe 58K (Bayesian Stats. and Machine learning) would be helpful but it is not a prerequiste

Administrative (Tentative)

- Grading
  - % 25 1 Midterm
  - % 25 1 Take Home Final
  - % 50 Assignments, Quizzes and Final Project
- Total Credits 3