Bayesian Theory and Computation

Spring 2024

Instructor: Cheng Zhang

Time: Monday 10:00-12:00pm, Odd Wednesday 3:10-5:00pm

Location: Classroom Building No.3, Room 106

Contact: chengzhang@math.pku.edu.cn

Office Hours: Thursday 3:00-5:00pm or by appointment, 315 Building No.20

Web Page: `https://zcrabbit.github.io/courses/btc-s24.html`. Visit this page regularly. It will contain homework assignments, lectures, etc.

Description and Objectives

The objective of this course is to explore Bayesian statistical theories and methods, and discuss their application in real life problems. Students would learn how to formulate a scientific question by constructing a Bayesian model, and perform Bayesian statistical inference to answer that question. Throughout this course, students would be exposed to the theory of Bayesian inference. They would also learn several computational techniques, such as importance sampling, sequential Monte Carlo, Markov Chain Mote Carlo (MCMC) algorithms, variational inference (VI), and use these techniques for Bayesian analysis of real data. Additional topics may vary. Coursework will include computer assignments.

Prerequisites

Some background in probability and statistical inference equivalent to two quarters of upper division or graduate coursework. Prior programming experience in **python** or R is helpful.

Assignments and Grading Policy

There will be 4 problem sets ($4 \times 15\% = 60\%$), and a final project (40%) which includes a midterm proposal (5%), an oral presentation (10%) and a final write-up (25%). There will be 7 free late days in total, use them in your own ways. Afterwards, late homework will be discounted by 25% for each additional day. Not accepted after 3 late days per problem set (PS). Late policy does not apply to the final project, please submit it on time. Discussing assignments verbally with classmates is allowed and encouraged. However, you should finish your work independently. Identified cheating incidents will be reported and will result in zero grades.

Computer and Technical Requirements

We will use python during the course. A good Python tutorial is available at `http://www.scipy-lectures.org/`. You may also find another shorter tutorial useful at `http://cs231n.github.io/python-numpy-tutorial/`. If you have never used Python before, I recommend using Anaconda Python 3.7 `https://www.continuum.io/`.

References

Main textbook

- Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003). Bayesian Data Analysis, 2nd Edition, Chapman & Hall.

Other interesting references

- Liu, J. (2001). Monte Carlo Strategies in Scientific Computing, Springer-Verlag.
- Lange, K. (2002). Numerical Analysis for Statisticians, Springer-Verlag, 2nd Edition.
- Keener, R.W. (2010). Theoretical Statistics: Topics for a Core Course, Springer.
- Christian, P.R. (2004). The Bayesian Choice, Springer.
- MacKay, D. (2003). Information Theory, Inference, and Learning Algorithms, Cambridge University Press.

Tentative Outline of Topics

1. Basic concepts in Bayesian statistics: prior; likelihood; posterior; Bayes' theorem; exponential family; conjugate priors; posterior predictive distribution; Bayesian hypothesis testing and model evaluation, etc.
2. Decision theory: utility; loss; posterior risk; formal Bayes rule; classical decision theory; risk function; Bayes rule.
3. Monte Carlo integration; simple simulation methods; rejecting sampling; importance sampling; variance reduction techniques; sequential Monte Carlo.
4. Markov chain Monte Carlo: Metropolis-Hasting; Gibbs sampling; slice sampling; Hamiltonian Monte Carlo; stochastic gradient MCMC; Convergence analysis.
5. Variational inference: mean-field; stochastic gradient optimization (control variate and the reparameterization trick); scalable approaches; choice of training objectives.
6. Bayesian regression and classification models: linear regression; logistic regression; Poisson regression; multinomial logistic regression.
7. Some advanced topics: Gaussian process for regression and classification; Dirichlet process mixtures for density estimation and clustering.