STAT598 Monte Carlo Methods for Complex Data

Instructor: Dr. Faming Liang, http://www.stat.purdue.edu/~fmliang/

Time & Location: Tuesday & Thursday 10:30–11:45am, Recitation Building 113

Course Overview: This course is to introduce to students statistical theory and methods for analysis of large complex data. The course will cover fundamental topics of computational statistics, such as Markov chain Monte Carlo, stochastic optimization, multiple hypothesis testing, high-dimensional variable selection, graphical models, and deep learning.

Prerequisites: One year graduate course of Probability and Statistical Inference

Course Objectives: Upon successful completion of the course, students should:

- Understand the role of computational statistics in complex data analysis.
- Understand the nature of high-dimensional and big data.
- Be aware of fundamental concepts of Markov chain Monte Carlo and stochastic optimization.
- Be aware of fundamental concepts of data reduction, model selection, large-scale hypothesis testing, and deep learning.
- Be able to analyze high-dimensional data and big data using statistical methods recently developed in the literature.

Topics to be covered:

1. Markov Chain Monte Carlo and Stochastic Optimization
   (a) Basic algorithms: Metropolis-Hastings algorithm, Gibbs sampler.
   (b) Advanced MCMC algorithms: parallel tempering, evolutionary Monte Carlo, stochastic approximation Monte Carlo, adaptive MCMC algorithms.
   (c) Sequential Monte Carlo.
   (d) Stochastic optimization: simulated annealing, genetic algorithms, simulated stochastic approximation annealing algorithms.
   (e) The EM, stochastic approximation, and imputation-regularized-optimization algorithms for missing data analysis.

2. High-Dimensional Data Analysis
   (a) Dimension reduction: principal component analysis, single value decomposition.
   (b) Regularized regression: Lasso and related methods.
(c) Bayesian variable selection, Bayesian model averaging.

3. Large-Scale Hypothesis Testing
   (a) False discovery rate (FDR)
   (b) Empirical Bayes methods
   (c) Benjamini and Hochberg’s FDR control methods.

4. Graphical Models
   (a) regularization methods: graphical Lasso and nodewise regression
   (b) multiple hypothesis test-based methods: ψ-learning.

5. Deep Learning
   (a) Bayesian neural networks.
   (b) Boltzmann machine.
   (c) Deep neural networks.

Recommended texts/references:


Course Website: http://www.stat.purdue.edu/~fmliang/. This site will be used to provide you with information relevant to the course. Such information includes this page, announcements, lecture notes, homework assignments, reading assignments, data sets, dates of exams, and review sheets.

Blackboard Learn: A Learn site has been created for this course. Students enrolled in this course will automatically be given access to this site (auto-populated). The site is limited to enrolled students and thus will be used to provide restricted information such as grades.

Course Requirements/Evaluation/Grading
The assessment will include class participation, assignments, and one course project. Class participation will include weekly attendance and participation in discussions. Students are responsible for all course material, including reading required materials prior to each class. Failure to complete assignments will result in a failing grade.

Class participation: 10% Assignments: 40% Course Project: 50%

The grading scale for this course consists of the standard scale, including minus grades, below. The conversion factors for grade point values that are assigned to each grade are also included (in parentheses):

- 90% - 100% = A (4.00)
- 80% - 89% = B (3.00)
- 70% - 79% = C (2.00)
- 60% - 69% = D (1.0)
- Below 60% = F (0.00)

Class Demeanor Expected by the Professor (late to class, cell phones): Students are expected to show up for class prepared and on time. Cell phones are to be silenced during class unless there is an emergency, in which case please inform the instructor.

Academic Integrity: We take academic integrity very seriously in this course. The only true way to get an education is through hard work and striving to understand concepts on your own. The penalty for academic misconduct on any assignment, exam, or final project is failure for the course with referral to the Dean of Students for further sanctions. Cheating on the assignments, midterm, final project, or final exam results in an F for the course. Note that we punish not only the person who cheats but also the person who enables the cheater. When it comes to academic misconduct we have zero tolerance.

Policy Related to Class Attendance and Late or Missed Assignments: Attendance of all class sessions is required. Please see the instructor as early as possible regarding possible absences. All assignments need to be handed in on time. Grading will penalize late assignments. Missed assignments will receive a zero score. Personal issues with respect to class attendance or fulfillment of course requirements (assignments, final presentation, class discussion) will be handled on an individual basis.

Emergency Preparedness: In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructors control. Here are ways to get information about changes in this course.

- Course web page (www.stat.purdue.edu/~fmliang/)
- Instructors email (fmliang@purdue.edu)
Emergency preparedness is a personal responsibility. Purdue University is continuously preparing for natural disasters or human-caused incidents with the ultimate goal of maintaining a safe and secure campus. Let's review the following procedures:

- To report an emergency, call 911.
- To obtain updates regarding an ongoing emergency, and to sign up for Purdue Alert text messages, view www.purdue.edu/ea.
- There are nearly 300 Emergency Telephones outdoors across campus and in parking garages that connect directly to the Purdue Police Department (PUPD). If you feel threatened or need help, push the button and you will be connected immediately.
- If we hear a fire alarm, we will immediately suspend class, evacuate the building, and proceed outdoors, and away from the building. Do not use the elevator.
- If we are notified of a Shelter in Place requirement for a tornado warning, we will suspend class and shelter in the lowest level of this building away from windows and doors.
- If we are notified of a Shelter in Place requirement for a hazardous materials release, or a civil disturbance, including a shooting or other use of weapons, we will suspend class and shelter in our classroom, shutting any open doors or windows, locking or securing the door, and turning off the lights.