STAT546 Computational Statistics

Instructor: Dr. Faming Liang, http://www.stat.purdue.edu/∼fmliang/
Time & Location: MWF 1:30–2:20pm, Helen B. Schleman Hall 114

Course Overview: This course introduces students to the statistical theory and methods for analyzing large, complex datasets. It covers core topics in computational statistics, including 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 be able to:

- Understand the role of computational statistics in modern data science.
- 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.
- Analyze high-dimensional and/or big data using statistical computing and deep learning methods.

Topics to be covered:

1. Markov Chain Monte Carlo and Stochastic Optimization
   (a) Basic algorithms: Metropolis-Hastings algorithm, Gibbs sampler, stochastic gradient Langevin dynamics (SGLD)
   (b) Advanced MCMC algorithms: parallel tempering, evolutionary Monte Carlo, stochastic approximation Monte Carlo (SAMC), stochastic gradient MCMC algorithms, 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, sufficient dimension reduction
   (b) Regularized regression: Lasso and related methods.
   (c) Bayesian variable selection, Bayesian model averaging.
   (d) High-dimensional inference: post-selection inference, de-sparsified Lasso, Markov neighborhood regression

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) Deep neural networks, convolutional neuarl networks, Bayesian neural networks, stochastic neural networks
   (b) Generative adversarial network (GAN), variational autoenconder (VAE), diffusion model
   (c) Uncertainty quantification: extended fiducial inference, physics-informed neural networks

Recommended texts/references:

1. Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications.* Springer.

2. Drori, I. (2022). *The Science of Deep Learning.* Cambridge University Press.

3. Grohs, P. and Kutyniok, G. (2023). *Mathematical Aspects of Deep Learning.* Cambridge University Press.

4. Liang, F. and Jia, B. (2023). *Sparse Graphical Modeling for High Dimensional Data: A Paradigm of Conditional Independence Tests.* CRC Press.

5. Liang, F., Liu, C. and Carroll, R.J. (2010). *Advanced Markov chain Monte Carlo Methods: Learning from Past Samples.* Wiley.

Exams Homework (20%), project (30%), Final (50%). The final percentages needed for a particular grade are as follows: 90—100=A, 80—89 = B, 70—79 = C, 55—69 = D, 0—54 = F. The minimum score needed for a given letter grade could be lowered if necessary but will not be raised. +/- grades are only given in special circumstances.

Policy Related to Class Attendance and Late or Missed Assignments Students are expected to show up for class prepared and on time. Please see the instructor as early as possible regarding possible absences. Cell phones are to be silenced during class unless there is an emergency, in which case please inform the instructor. 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.

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.

Purdue prohibits "dishonesty in connection with any University activity. Cheating, plagiarism, or knowingly furnishing false information to the University are examples of dishonesty." [Part 5, Section III-B-2-a, Student Regulations] Furthermore, the University Senate has stipulated that "the commitment of acts of cheating, lying, and deceit in any of their diverse forms (such as the use of substitutes for taking examinations, the use of illegal cribs, plagiarism, and copying during examinations) is dishonest and must not be tolerated. Moreover, knowingly to aid and abet, directly or indirectly, other parties in committing dishonest acts is in itself dishonest." [University Senate Document 72-18, December 15, 1972]

Use of Copyrighted Materials Students are expected, within the context of the Regulations Governing Student Conduct and other applicable University policies, to act responsibly and ethically by applying the appropriate exception under the Copyright Act to the use of copyrighted works in their activities and studies. The University does not assume legal responsibility for violations of copyright law by students who are not employees of the University.

A Copyrightable Work created by any person subject to this policy primarily to express and preserve scholarship as evidence of academic advancement or academic accomplishment. Such works may include, but are not limited to, scholarly publications, journal articles, research bulletins, monographs, books, plays, poems, musical compositions and other works of artistic imagination, and works of students created in the course of their education, such as exams, projects, theses or dissertations, papers and articles.

Violent Behavior Policy Purdue University is committed to providing a safe and secure campus environment for members of the university community. Purdue strives to create an educational environment for students and a work environment for employees that promote educational and career goals. Violent Behavior impedes such goals. Therefore, Violent Behavior is prohibited in or on any University Facility or while participating in any university activity. See the University's website for additional information: http://www.purdue.edu/policies/facilities-safety/iva3.html.

Emergencies 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 instructor's control. Relevant changes to this course will be posted onto the course website or can be obtained by contacting the instructors or TAs via email or phone. You are expected to read your @purdue.edu email on a frequent basis. See the University's website for additional information: https://www.purdue.edu/ehps/emergency_preparedness/.

Accessibility and Accommodations Purdue University strives to make learning experiences as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, you are welcome to let me know so that we can discuss options. You are also encouraged to contact the Disability Resource Center at: drc@purdue.edu or by phone: 765-494-1247.

Nondiscrimination Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life.

Purdue University views, evaluates, and treats all persons in any University related activity or circumstance in which they may be involved, solely as individuals on the basis of their own personal abilities, qualifications, and other relevant characteristics.

Purdue University prohibits discrimination against any member of the University community on the basis of race, religion, color, sex, age, national origin or ancestry, genetic information, marital status, parental status, sexual orientation, gender identity and expression, disability, or status as a veteran. The University will conduct its programs, services and activities consistent with applicable federal, state and local laws, regulations and orders and in conformance with the procedures and limitations as set forth in Purdue's Equal Opportunity, Equal Access and Affirmative Action policy which provides specific contractual rights and remedies. Additionally, the University promotes the full realization of equal employment opportunity for women, minorities, persons with disabilities and veterans through its affirmative action program.

Any question of interpretation regarding this Nondiscrimination Policy Statement shall be referred to the Vice President for Ethics and Compliance for final determination.

Disclaimer This syllabus is subject to change.