Statistics and Numerical Methods

Course Syllabus for Fall 2019

Instructors

- Xuening Bai (xbai@tsinghua.edu.cn; Science Building 312 and Mong Man-wai South Building S-617B)
- Dandan Xu (dandan@tsinghua.edu.cn; Mong Man-wai South Building S-613B)

Times

- Mondays, 9:50am-12:15 pm

Location

- Science Building (科学馆) 322.

Suggested Prerequisites

- Calculus, linear algebra, complex analysis, probability theory
- Experience with at least one programming language

Course Web Page

- We use the Tsinghua University web-learning platform (网络学堂) to post lectures and problem sets, and send out announcements.
- You are welcome to post questions of common interests to the platform, though the preferred means is via office hours.

Office Hours

- Xuening Bai: 1:30 pm to 3:00pm on Wednesdays at Science Building 312 (exception: Sep. 11, Oct. 30, Nov. 27), or by email appointment.
- Dandan Xu: 1:30 pm to 3:00pm on Mondays, or by email appointment.

Summary

This course provides an overview to statistics and computational methods primarily for astronomers. Topics include numerical linear algebra (mostly on matrix computations), non-linear system of equations, optimization techniques, probability and statistical distributions, classical and Bayesian inference, regression and model fitting, data mining, stochastic processes and MCMC, Fourier analysis, and time series analysis. Specific applications are mostly drawn from astronomy and astrophysics.

We do not aim to be rigorous in math, but rather aim to cover broad range of topics at relatively high-level. The goal is to provide you an overview of major algorithms, techniques and methodologies that are practically most useful, as well as the underlying ideas behind (rather than giving full proofs). Furthermore, we anticipate you are able to master these subjects through homework and final projects at application level.

Due to limited time and very broad range of subjects, we will not cover some other important topics, particularly on methods for solving ordinary and partial differential equations. We anticipate some of these subjects will be covered in future extensions of this course. Also, while we do not explicitly cover machine learning, many of the concepts and techniques in machine learning are addressed in this course.

The materials in this course will be presented a series of lectures. All lectures will also be posted to the online platform. Besides, in class, you are strongly encouraged to take notes and ask questions.

Course Materials

There is no required textbook in this course. The primary references where majority of our lectures are built upon include the following

1. *Numerical Recipes, the Art of Scientific Computing* by W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (Cambridge University Press 1992 for 2nd edition, 2007 for 3rd edition). The 2nd edition is available online: http://www.numerical.recipes/

2. *Statistics, Data Mining, and Machine Learning in Astronomy* by Z. Ivezic, A.J. Connolly, J.T. Vanderplas and A. Gray (Princeton University Press, 2014).

The following resources contain additional information on numerical analysis, scientific computing, and astrostatistics which are good secondary references.

1. *Numerical Algorithms* by Justin Solomon (AK Peters/CRC Press, 2015). Available online: http://people.csail.mit.edu/jsolomon/#book
2. *Practical Statistics for Astronomers* by J.V. Wall and C.R. Jenkins (Cambridge University Press, 2nd edition, 2012).

3. *Summer schools in Statistics for Astronomers* (I-XV), run by center for Astrostatistics at Penn State University, with well documented lectures online: https://astrostatistics.psu.edu/

Course Schedule

The following is the approximate course schedule. These dates may change.

Dates	Content	Lecturer	Comments
9/9	Introduction; basic numerical analysis	DX	
9/16	Numerical linear algebra 1	XB	HW 1
9/23	Numerical linear algebra 2	XB	
9/30	Nonlinear systems	XB	HW1 due; HW2
10/7	No lecture		
10/14	Optimization methods	XB	
10/21	Probability and statistical distributions	XB	HW2 due; HW3
10/28	Classical statistical inference	XB	
11/4	Bayesian statistical inference	DX	HW3 due; HW4
11/11	Regression and model fitting	DX	
11/18	Monte Carlo method and Stochastic processes	DX	HW4 due; HW5
11/25	Fourier analysis	DX	
12/2	Structures in point data; dimensionality reduction	DX	HW5 due; HW6
12/9	Classification;	DX	
12/16	Time series analysis	XB	HW6 due; HW7
12/23	TBD/Summary	XB, DX	HW7 due

Problem Sets

There will be seven problem sets assigned throughout the semester (once every two weeks). These problem sets are designed to help you become familiar with the concepts from the prior two lectures. The course schedule lists the dates when the problem sets will be assigned, and the target dates for them to be handed in. Students may either hand in hardcopies of their solutions to one of the instructors (preferably during the class), or email electronic versions to both.

Each problem (and sub-problems within it) weighs a certain number of points. Grading will be based on the accuracy of the solution, the inclusion of intermediate steps, and communication of how the solution was found.

Many of the problems require certain level of coding. You must include your original code that are adequately commented in the solution sets.

Problems sets are due on the dates listed. However, since it is our top priority that students have sufficient time to learn from the problem sets, an extension (generally up to 1 week) will be granted on special cases (such as illness, conference travel, family obligations), if the students inform the instructors before the due date. If instructors are not consulted, then each problem set will incur a penalty of 10% lower grade per day that it is late.

Grading and Exams

The final grade will be broken down into contributions from problem sets (45%), final project (25%), and a final take-home exam (30%).

The final project will enable you to explore a chosen topic in more detail and gain experience with scientific writing. We will provide a potential list of topics on Nov 4, and you are welcome/encourage to choose your own topics in consultation with the instructors. Some projects can be challenging and you are encouraged to form groups of 2-3 people to work together (and indicate individual contributions in the end). Two (or two groups of) students are allowed work on the same topic (must be independent). You must decide on the topics for by Nov 18 (though it may be changed at a later time without penalty). We anticipate the project would require substantial but not overwhelming time commitment. The tentative date of the final project will be Monday, Jan 6th, 2020.

Policy on Collaboration

Discussion and the exchange of ideas are essential to doing academic work. For assignments in this course, you are encouraged to consult with your classmates as you work on problem sets. However, after discussions with peers, make sure that you can work through the problem yourself and ensure that any answers you submit for evaluation are the result of your own efforts. Similarly, you must list the names of students with whom you have collaborated on problem sets.