Statistics and Numerical Methods

Course Syllabus for Fall 2024

Instructors

- Xuening Bai (xbai@tsinghua.edu.cn; Science Building 312, Physics Building E-316)
- Dandan Xu (dandanxu@tsinghua.edu.cn; Physics Building E-221)

Times

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

Location

- Teaching building #3-1200
- NO online component, but special request will be considered.

Suggested Prerequisites

- Calculus, linear algebra, 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, or to our Wechat group.
- Homework should be submitted through this platform.

Office Hours

- Xuening Bai: 1:30 pm to 3:00pm on Tuesdays, by email appointment.
- Dandan Xu: 1:30 pm to 3:00pm on Tuesdays, by email appointment.

Summary

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

We do not aim at being 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). While we do not explicitly cover machine learning, certain key concepts and techniques in will be addressed in this course. Furthermore, we anticipate you are able to master these subjects through the problem sets at application level.

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).

3. 天文学中的概率统计，陈黎/著（科学出版社，2020）

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-XVII), 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
9/10	Introduction; basic scientific computing	XB	
9/17	No lecture		
9/24	Interpolation/differentiation/integration	XB	
9/29	Numerical linear algebra I	XB	HW1
10/8	Numerical linear algebra II, Nonlinear systems	Zhuo Chen	
10/15	Optimization methods	XB	HW1 due; HW2
10/22	Ordinary differential equations	XB	
10/29	Partial differential equations	XB	HW2 due; HW3
11/5	Probability and statistical distributions	XB	
11/12	Classical statistical inference	XB	HW3 due; HW4
11/19	Bayesian statistical inference	DX	
11/23*	Monte Carlo method and Stochastic processes	DX	HW4 due; HW5
12/3	Model Regression	DX	
12/10	Fourier analysis and Application	DX	HW5 due; HW6
12/17	Data Structure and Classification	DX	
12/24	Applications	TBD	HW6 due

*: Rescheduled class. Time and location TBD.

Problem Sets

There will be six problem sets assigned throughout the semester (about once every two weeks). The course schedule lists the dates when the problem sets will be assigned, and the target dates for them to be submitted through the online system. When submitting your homework, please **pack them into a .zip file**, with the following **naming convention**:

YourStudentID_YourName_HWx, where x=1-6.

For example, 2024311234_ZhangSan_HW2.zip

These problem sets are designed to help you become familiar with the concepts from the previous lectures, some of which might be challenging. **Never wait till the last moment to start working on the problem sets!** 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.

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.

The problem sets generally require certain (from minimum to substantial) level of programming. While most problems can be worked out using scripted language such as Python or Matlab, and generally you do need them to make plots, we note that these languages are not designed for scientific computing (and can be exceedingly slow). It is encouraged that you work with a compiled language (C, C++, Fortran) for some problems to gain some proficiency. Normally, we ask that you include your original code with adequate comments in the solution sets.

Discussions and exchange of ideas are essential for academic work. You are encouraged to consult with your classmates as you work on problem sets. However, after discussion 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.

Grading and Final Eaxm

The final grade will be broken down into contributions from problem sets (65%), and a final take-home exam (35%). We expect to distribute the final exam questions during our final lecture, and the exam is due one week later.