Astronomy 8824: Numerical and Statistical Methods in Astrophysics
Autumn 2017, David Weinberg

Course Objectives

The goal of this course is to provide basic background in numerical and statistical methods relevant to astrophysical research. By the end of the course, you should have a good gut-level sense for many of the statistical issues that will arise in your own research or when reading papers, listening to seminar talks, or following coffee discussions. You should also have an entry point for solving some of the most common types of numerical problems that arise in astronomical research. I would characterize the course as being rather elementary on the numerical side but fairly sophisticated on the statistics side.

I have decided to use Python as the common language for the course, in part because it is easier to get started in Python than in Fortran, C, or C++, but mostly because it has become a valuable language to have in your toolbox, as there are many useful, publicly available, scientific codes in Python. However, I should caution that I am not an expert Python programmer, and this is *not* a course in Python programming techniques. The code snippets I will provide for problem set starting points or solutions are designed to be easy to follow, and thus unsophisticated in structure.

Books

There are two primary references for this course. The first is *Numerical Recipes*, by Press, Teukolsky, Vetterling, and Flannery. The book is available in several editions in different programming languages. You can choose whichever edition seems like it will be most useful to you; we will be interested primarily in the explanations that it provides of different numerical methods rather than the routines themselves. The 3rd edition does have useful material that isn't in previous editions, but if you want to buy a used version of an older edition it will have the essentials. While one can often find better implementations of the methods elsewhere, and routines for many of the important numerical methods are available in Python, this book is an indispensable resource for an astronomer because of its explanation of how these methods work.

The other primary reference is *Statistics, Data Mining, and Machine Learning in Astronomy,* by Ivezic, Connolly, VanderPlas, and Gray. It is mainly the first five chapters of this book that we will use, i.e., the material on statistics rather than the material on data mining and machine learning. These chapters are as good a treatment of statistics for astronomers as I have come across — compact and practical — and the book is also a helpful introduction to Python. The second half of the book, which we will not cover in the course, has good introductions to many useful techniques, such as Principal Component Analysis, classsification schemes, and methods for time series analysis. I recommend buying this book because I think you will find it useful over the long haul, but for purposes of this course alone you could get away with borrowing a copy to read the parts you need.

Another excellent resource is *Effective Computation in Physics*, by Scopatz & Huff. This is the best introduction to scientific computing with Python I have seen, and it covers many other useful topics about linux, makefiles, version control, a little bit of LaTeX, and so on. We won't rely directly on this book for this course, but reading it will be a big asset to any computational work you do, and this course is the natural venue for you to read it and start taking advantage of what you learn.

Topics and Schedule

This is a 2-credit-hour course rather than a 3-credit-hour course, because when we changed the graduate curriculum from quarters to semesters we wanted to keep the total number of classroom hours in the graduate program the same. For the most part, therefore, the class will meet Tuesday and Thursday every *other* week. However, there will be some adjustments to accommodate both my travel schedule and university holidays, including one phase shift where we meet two consecutive weeks. I currently anticipate that we will have our class meetings on: 8/22 (short orientation only); 8/29, 8/31; 9/12, 9/14; 9/26, 9/28; 10/3, 10/5 (note two consecutive weeks); 10/17, 10/19; 10/26, 10/31 (note Thursday/Tuesday instead of Tuesday/Thursday); 11/14, 11/16; 11/28, 11;30. However, please don't schedule conflicts in other weeks, because the schedule could still shift.

The statistical and numerical topics will be intermingled throughout the course. The primary topics are

Statistical:

- Basic principles of probability and statistics
- Parameter estimation and hypothesis testing
- Correlated errors and multi-variate Gaussians
- χ^2 and other statistical tests
- Fisher matrix forecasting
- Markov Chain Monte Carlo estimation of parameter distributions

Numerical:

- Numerical integration
- Numerical solution of differential equations
- Root-finding and minimization
- Fourier transforms and their uses
- Time permitting: Matrices, eigenvalues, diagonalization, principal component analysis (I didn't make it to this topic last time)

We will take a mostly Bayesian outlook on statistical questions, though with some discussion of other viewpoints.

Assignments and Grading

There will be seven problem sets, primarily computational. These are a central element of the course; while I expect the lectures to be useful, you will probably learn more from doing the problems and reading my solution sets than you do from everything else.

I will usually hand out problem sets on Tuesdays, and they will be due at the start of the class session 2 weeks later. Each problem set should take 6-10 hours to complete. You shouldn't go over 10 hours on a given problem set, so once you hit the 8-9 hour mark focus on writing up what you have in hand.

Reading assignments will be given in class or specified together with the problem set assignment.

There are no exams. The course grade will be based on the problem sets and class participation.

I will post notes and useful links on the course web page,
`http://www.astronomy.ohio-state.edu/~dhw/A8824`.