STAT 221: Statistical Computing and Computational Statistics

Instructor: Jun Liu; Office: 715 Science Center; Phone: 495-1600
Office Hour: Thursday 1-2PM.

COURSE DESCRIPTION: The ultimate goal of this course is to help you learn computational tools and methodologies necessary for your statistical researcher and possibly other quantitative researches. These techniques can be classified into three main classes: linear algebra methods for dealing with matrices, numerical integration and optimization methods, and Monte Carlo methods.

PREREQUISITE & REQUIREMENT: First-year graduate level probability and statistics, linear algebra and calculus. You will be required to implement most of the methods discussed in class as part of the problem sets. Your projects at the end of the term will also require some amount of programming. So, some knowledge of programming, especially in R/S-PLUS, is absolutely essential. Officially, programming in R/S-PLUS will be supported, although, you could use any other computing language MATLAB (or, more ideally, C or C++) of your choice.

There will be four to five assignments spread throughout the semester. At the end of the term you will be asked to do a project, which will culminate in a presentation and submission of a project report. This might involve reading a paper and implementing the methods in there, or even better, inventing new ones by modifying the existing methods.

Meetings every TF 2:30-4:00 in SC-222. Expected to participate in class discussions (5%) and to work on homework (50%) and projects (presentation 20%, report 20%).

TOPICS TO BE COVERED:
The following topics are to be covered (not in chronological order and subject to changes).

- Newton-Raphson method, Scoring method
- Other mode finding methods, linear programming.
- Numerical integration
- Regression, QR, Sweep operator, Cholesky.
- Singular Value Decomposition, LDA, SVM (optional)
- EM, GEM, ECM, SEM algorithms
- Monte Carlo methods
  - Random number generation
  - Acceptance-Rejection, Antithetic variates, Control variates
  - Importance Sampling, Sequential Importance Sampling
  - Markov chain Monte Carlo
  - Sequential Monte Carlo
- Examples and other advanced MC techniques.

- Other topics: spline smoothing, density estimation, FFT, bootstrap, etc.

REFERENCES


