Statistics 606: Computational Statistics

Instructor: Kerby Shedden (kshedden@umich.edu), 461 West Hall, 764-0438.

Meeting time & place: T/Th 10-11:30, B760 East Hall.

Prerequisites: Students should have statistical background equivalent to Stat 425/426. Some experience in writing computer code is strongly recommended.

Office hours: Make an appointment or drop by.

Overview: This course covers the use of computer algorithms in statistics and data analysis. The emphasis is on the practical aspects of implementing algorithms from the ground up. Upon completion of the course the student will be able to implement any algorithm that arises in routine statistical research or scientific data analysis. Note that this course will not cover statistical packages such as SAS and SPSS.

Grading: Grades will be determined based on approximately four problem sets. There will be no exam. The problem sets will primarily involve writing computer programs.

Computing issues: The class will be taught using example programs primarily written in Matlab (Octave) and C. Students may write their problem set solutions using R, Matlab or C. SAS, SPSS, and other very high level statistical packages are not suitable for the problem sets.

Outline:

1. Brief overview of the Matlab (Octave) and C programming languages.

2. Elementary algorithms: sorting, interpolation, root finding and maximization in one dimension.


7. Applications of linear algebra: least squares problems, smoothing splines, multivariate analysis and time series analysis.

8. Optimization: Newton-Raphson, Fisher scoring, IRLS, steepest descent, conjugate gradient, constrained optimization, simulated annealing, EM algorithm; applications to maximum likelihood and other estimation problems including finite mixtures, ridged logistic regression, mixed effects models, hierarchical models, robust regression, missing data techniques, Kalman filtering, hidden Markov Models (HMM’s), and models based on linear differential equations.
9. Simulation I: random number generation, random permutations, subsampling and resampling, inversion method, rejection sampling, importance sampling.

10. Simulation II: Markov chain methods, perfect sampling, applications to Bayesian inference.