

STAT 7730 Syllabus
Advanced Computational Statistics
Autumn 2014

Instructor: Dr. Sebastian Kurtek

E-Mail: kurtek.1@stat.osu.edu

Office: CH 440H

Office Phone: 292-0463

Class Times: MWF 11:30AM-12:25PM

Class Location: Cockins Hall 228

Office Hours: TR 9:30-10:30AM, other times by appointment

Main Text: *Computational Statistics* by G. Givens and J. Hoeting

Other Useful Texts:

Monte Carlo Statistical Methods by C. Robert and G. Casella

An Introduction to the Bootstrap by B. Efron and R. Tibhsirani

Numerical Methods of Statistics by J. Monahan

Statistical Computing with R by M. Rizzo

Tentative Schedule:

1. Linear Methods for Regression Analysis/Matrix Decomposition: basic numerical analysis, multiple regression analysis, orthogonalization by Householder transformation, singular value decomposition, QR decomposition, principal component analysis, linear discriminant analysis – 2 weeks

2. Numerical Methods for Maximum Likelihood Estimation: univariate/multivariate numerical optimization, maximum likelihood estimation, expectation-maximization (EM) algorithm and extensions – 3 weeks

3. Random Number and Variable Generation: uniform random number generators, modular arithmetic, combination generators, discrete and continuous random variables, inverse transform method, acceptance-rejection method, tilted sampling – 2 weeks

4. Monte Carlo Integration: general formulation, importance sampling, variance reduction, numerical integration and differentiation – 2 weeks

5. Markov Chain Monte Carlo (MCMC) Methods: properties of Markov chains, Metropolis-Hastings algorithm, Gibbs sampler – 2 weeks

6. Bootstrap: plug-in estimator, non-parametric/parametric bootstrap, bootstrap estimate of standard error, confidence intervals based on bootstrap – 1 week

8. Additional Topics: statistics of directional data - time permitting

Prerequisites: 6802 (622) and 6950 (645) or permission of instructor. Not open to students with credit for 773. Additionally, working knowledge of linear algebra, advanced calculus, and some programming background is helpful.

Course Description: STAT 7730 is a graduate level course in modern statistical computing methods. This course is not about the use of pre-packaged statistical software. The main goal of this course is to gain an understanding of advanced techniques and ideas used in implementing mathematical/statistical formulations on computers, with a focus on common statistical approaches. Students will be expected to implement the methods we cover in class by programming in a language of their choice (preferably R, although Matlab is acceptable). I will provide example R code that goes along with the material covered in class. Students are expected to be able to analyze the code, and apply the basic structure of the code to new problems assigned as homework. **Note:** Students who have had no prior programming experience should expect to spend extra time outside of class reviewing the example code and familiarizing themselves with a statistical programming environment.

Computing: Students are expected to have basic familiarity with a scientific computing environment such as R, S-PLUS or Matlab. We will use the R language and environment for statistical computing and graphics as an aid to learning about statistical computing methods. R is available for free at <http://www.r-project.org>. I will provide example R code in class. Students are highly encouraged to use R; check with the instructor to see if use of another language/environment is acceptable. Some students in the past have successfully completed the course using Matlab.

Grading: The final course grade will be based on homework assignments (40%), a midterm project (30%) and a final project (30%).

Homework Assignments: Homework assignments will be given approximately once every two weeks on a Friday and be due the following Friday at the beginning of class. *Late assignments will not be accepted.* The assignments will require the derivation of analytical results as well as the implementation of the computational methods we discuss in class. Please write clear and detailed answers to the homework problems and provide a statement interpreting the obtained results. If a problem involves writing a program, submit a printout of the code with the solution. It is important to provide illustrative outputs of your programs to accompany the homework solutions. For instance, all graphs should be labeled and placed close to the associated written part. Points are allocated to both the correctness of the solution and the level of presentation. Students may consult with each other on the homework problems, but each student must complete and turn in his or her own work. **DO NOT** copy or use computer code written by another student.

Midterm and Final Projects: Both the mid-term and the final projects will be take-home projects. You will be asked to write a report on the project similar to a research paper. These projects are graded for both the solutions and the quality of presentation. I encourage all students to use Latex to generate their project reports (this is not a class requirement). The Latex typesetting language is very popular for creating mathematical documents, and I hope that you will become at least somewhat familiar with Latex by the end of this class. I will provide an example of a document written in Latex as a reference.

Academic Misconduct: Academic misconduct will not be tolerated and will be dealt with procedurally in accordance with university policy, which can be found at <http://oaa.osu.edu/coam.html>. Students are allowed to consult with each other on homework assignments. However, each student must complete and turn in his or her own work. The discussions should be at the ideas level and not the details level. DO NOT copy or use computer code written by another student.

Addressing Issues of Differing Abilities: If you have a documented disability please register with the Office for Disability Services (ODS). After registration, make arrangements with me as soon as possible to discuss your accommodations so that they can be implemented in a timely fashion. If you have any questions about this process please contact ODS at (614) 292-3307.

Note: Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advanced notice.