ISyE 6416 - Computational Statistics - Spring 2017
Tentative Syllabus

Computational statistics is an interface between statistics and computing. We will cover algorithms arising from parametric and nonparametric statistical problems, developed with optimization and other modern computational tools. Course will cover the following topics:

Optimization tools for statistics
Parametric approaches

- EM algorithm
- Gaussian mixture model (GMM)
- Hidden Markov models (HMM)
- Model selection and cross validation

Nonparametric approaches

- Principle component analysis and low-rank models
- Splines and approximation of functions
- Bootstrap
- Monte Carlo methods

And other optional selected topics (e.g., convex and non-convex optimization for statistics, neural networks).

Class Time and Location: TR, 12:05-1:25pm, Clough 102. From Jan. 9 to April 25

Instructor: Prof. Yao Xie, Groseclose #339, email: yao.xie@isye.gatech.edu
Instructor Office Hour: Wednesday 1-2pm, Groseclose 339

Class TA: Xi He, xi.he@gatech.edu
TA Office Hour: Tuesday 3-4pm, Main Building 324
(May have additional office hours before exams.)

Class Website: T-square
Class material available on our website includes
- Announcements
- Course syllabus
- Homework assignments and solutions
- Slides and other lecture material
- Practical exams
- Your course grades on exams and homework
- Any important announcements

Make sure the scores in T-square are consistent with what you got. We will not make any change in grading for works older than 2 weeks.

Class Mailing List: Registered students are automatically subscribed to the class mailing list.

Textbook: the course material will be based on lectures and slides posted on T-square.
References*:*
Computational statistics. G. H. Givens and J. A. Hoeting, 2013.
Computational Statistics. James E. Gentle. 2009.
Numerical Analysis for Statisticians, Kenneth Lange.
An introduction to statistical learning: with applications in R, G. James, D. Witten, T. Hastie, R. Tibshirani.
Computational Statistics handbook with MATLAB, W. L. Martinez and A. R. Martinez.
Elements of Computational Statistics, J. E. Gentle.
Numerical linear algebra, Lloyd N. Trefethen and David Bau III.
Numerical Recipes, 3rd edition, William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery.
The elements of Statistical Learning: Data Mining, Inference, and Predictions, 2nd edition, Trevor Hastie, Robert Tibshirani, and Jerome Friedman.
Machine learning: A probabilistic perspective, K. P. Murphy.

Prerequisites: undergraduate level basic probability, linear algebra, and statistics.

Honor Code: For any question involving Academic Honor Code issues, please consult
www.honor.gatech.edu

Software: MATLAB and/or R. You will use this software for homework assignment.

Grading Policy: Class Participation 8% (there will be a few pop-up quizzes), Submitting Teaching Evaluation - 2%, Homework - 10%, Midterm 1 - 20%, Midterm 2 - 20%, Project - 10%, Final – 30%.

Homework: The homework should be handed in **before the end of the class on the due date**. The lowest homework-score will be dropped. Late Homework will NOT be accepted. Assignments will include both exercises and computer problems; the computer problems will ask you to carry out statistical analysis using computer statistical software. Keep in mind that you should not hand in raw computer output. Conclusions and interpretation of results are more important than good printouts. You are allowed to work together with other students on homework, as long as you write up and turn in your own solutions. You are also allowed (and encouraged) to ask me questions, although you should try to think about the problems before asking. Request for re-grading the Homework/Exams/Quizzes should be made within a week of returning Homework/Exams/Quizzes.

Midterms: There will be two in-class midterm exams during the class. The midterms are close notes (including assignment solutions) and close textbook but two and respectively, four two-sided pages with formulas will be allowed. Do not write homework solutions on the formula sheet. You are not allowed to use your cell phone. The notes have to be self-made. **No make-ups.**

Midterm 1: Week of Feb 12, 2017, in class, TBA.

Midterm 2: Week of March 12, 2017, in class, TBA.

Final Date: May 4, Thursday, 8:00-10:50am.

Project: by group, each group consists of 2-3 students. Details TBA.