COURSE INFORMATION: Applied Multivariate Analysis

Course web:
http://faculty.chicagobooth.edu/ruey.tsay/teaching/ama/sp2016

Course Objective:
- To learn basic techniques for analyzing multi-dimensional data (including visualization)
- To study multivariate distributions and their properties, especially Gaussian distribution
- To understand multivariate statistical inference and applications in scientific fields
- To discuss various methods for dimension reduction, including principal component analysis, factor model, multi-dimensional scaling, sliced inverse regression, independent component analysis, etc.
- To introduce recent developments in high-dimensional statistical analysis, including machine learning

Textbooks:
- *An Introduction to Statistical Learning with Applications in R* by G. James, D. Witten, T. Hastie, and R. Tibshirani (2013), Springer, ISBN: 978-1-4614-7137-0

Lecture Notes: Some handouts will be posted on the course web before lectures.

References: (Optional)

4. *Statistics for High-Dimensional Data* by Peter Bühlmann and Sara van de Geer, (2011),

5. *Introduction to High-Dimensional Statistics* by Christophe Giraud, (2015), CRC Press,

**Articles:** A list of recent articles for dimension reduction will be given and students are
expected to read over these articles (Download from e-journals from the library).

**Office hour:**
Thursday: 4:00 pm to 5:00 pm or by appointment.
My phone number 702-6750, My office: HPC 455.
E-mail: ruey.tsay@chicagobooth.edu
(this is the easiest way to make contact with me)

**Teaching assistant:** Grading only
   - e-mail: TBA

**Grading:**
Mid-term (30%), Final Exam (45%), and Homework Assignments (25%).

**Special notes:**
- Homework is due before the class one week after assigned.
- No late homework assignments will be accepted. Solutions or discussions will follow
  after the assignments are handed in.
- You may discuss assignments with each other, but must turn in your own answers.
- No e-mail submission of any assignment will be accepted.
- Mid-term: Week 6, close book. (May 5)

**Computing:**
The main package is R (see www.r-project.org). You may use any software of your choice.
You may use regular R or RStudio. To install RStudio, you need to install R first.

**Course Outline:** All topics include applications
1. Multivariate Normal distributions (multivariate-$t$)
2. Inferences about a mean vector and comparisons
3. Multivariate linear regression
4. Principal component analysis & Independent component analysis
5. New developments in dimension reduction
6. Factor analysis and discriminant analysis
7. Canonical Correlation Analysis: prediction
8. Discrimination and Classification
9. Clustering and introduction to data mining
10. LASSO regression and variable selection

Chicago Booth Honor Code
This course requires students to follow the Chicago Booth Honor Code and Standards of Scholarship in examination, final project and assignments. The Chicago Booth Honor Code requires students to sign the following pledge, “I pledge my honor that I have not violated the Honor Code during this examination”, on every examination.