STA 4006 Fundamentals in high dimensional data analysis
Fall 2022

Classes: Monday 9:00 AM - Noon, **Details: In-person class will be from 10:30 to 11:45**

Instructor: Changryong Baek
Email: crbaek@skku.edu
Office: Dasan 32410 (**Tel** 760-0602)
Office Hours: By appointment.

Course description:

- This course aims to learn various latest statistical techniques for high dimensional data and apply them to real applications. In particular, high-dimensional data analysis with time dependence will be considered intensively. Also, presentation on recent advances in statistical methods are required to promote developing new methods.

Recommended Prerequisite:

- Good understanding of matrix algebra, regression theory, mathematical statistics and basics of machine learning. Basic time series analysis background is also assumed.

Course Webpage: Course webpage is located at "icampus"(http://www.icampus.ac.kr).

Topics considered: Topics are quite flexible. Tentatively, I am trying to cover multivariate (possibly high dimensional) time series, adaptive lasso in high-dimensional time series (HDTS), graphical and fused lasso in HDTS, bootstrapping, HDTS testing and change-point detection, dynamic factor model, extreme value theory, deep learning and LSTM, etc..

Some references:

- *Time Series: Theory and Methods.* Second edition by Brockwell and Davis.
- *New Introduction to Multiple Time Series Analysis* by Lütkepohl. Comprehensive treatments for multiple time series analysis.
- *Analysis of Financial Time Series Analysis* by Tsay, Broad overview of financial time series analysis techniques. Covers most of topics in this course, but in the financial side.
- *Statistical Learning with Sparsity: The Lasso and Generalizations* by Hastie and Tibshirani
- *Deep Learning (Adaptive Computation and Machine Learning series)* by Ian Goodfellow et al.

Data analysis projects & Grading:

- First project (replacing midterm) is in-class presentation. Find recent papers (published less than 5 year ago) about HD data analysis, and present them in class.
- Second project is to find real data applications and apply the method you presented in the first project. You are required to turn in 5-page-report on your findings including the goal of the stdudy, data description, methods, empirical results and conclusion.
- However, you cannot present your own work/thesis, and any other materials already presented in other class/competition.

Schedule and Grading:

- Observe University policy based on the following criteria:

Midterm project	50%
Final project	50%
Total	100%

Academic misconduct:

- Academic Honesty / Honor Code / Student Code of Conduct will be observed at all times in this course. In case of any academic misconduct, you will get an F and may subject to file on University Judiciaries.

Problems/ Suggestions:

- Each student should feel comfortable approaching the Instructor with any problems s/he has with the course. Please feel free to visit my office during the office hour or e-mail me for any further questions/concerns/suggestions.

Disclaimer:

- Instructor reserves the right to change the syllabus if it is academically advisable and necessary.