COMPUTATIONAL METHODS FOR DATA SCIENCE
DATA 5009 SYLLABUS (FALL SEMESTER 2022)

General Information

Instructor: Frederick Kin Hing Phoa
Telephone: (02) 2787-5700
E-mail: fredphoa@stat.sinica.edu.tw
WWW: http://staff.stat.sinica.edu.tw/fredphoa
Lecture: Wednesday 13:20 - 16:20 Room 402 of Xinsheng Lecture Building

Course Outline

In this course students will be exposed to advanced methodologies in statistics computing and optimization methods, and get hand-on experience by applying various techniques on real-world data. For some unpredicted reasons, the class will be available on Youtube channel if necessary.

Part 1: Introduction and Reviews			
Week 01	(2022/09/07)	L01:	Basics on Statistics and Optimization
Part 2: Linear Algebra Techniques			
Week 02	(2022/09/14)	L02:	Linear Algebra and Matrix Decomposition
Week 03	(2022/09/21)	L03:	Eigenvalue Problems, PCA and SVD
Week 04	(2022/09/28)	L04:	Low-Rank Approximation and ICA
Week 05	(2022/10/05)		Homework 1 (No Class)
Part 3: Modern Methods in Optimization			
Week 06	(2022/10/12)	L05:	Linear Programming and Gradient Methods
Week 07	(2022/10/19)	L06:	Stochastic Local Search Methods
Week 08	(2022/10/26)	L07:	Genetic Algorithms and Evolution Strategy
Week 09	(2022/11/02)	L08:	Methods in Group Intelligence
Week 10	(2022/11/09)		Homework 2 (No Class)
Week 11	(2022/11/16)		Midterm Examination
Part 4: Statistical Computing Techniques			
Week 12	(2022/11/23)	L09:	Simulation, Sampling and Monte Carlo Methods
Week 13	(2022/11/30)	L10:	EM Algorithms and Methods of Bootstrapping
Week 14	(2022/12/07)	L11:	MCMC: Metropolis-Hasting and Gibbs Sampling
Week 15	(2022/12/14)	L12:	Nonparametric Density Estimation and Multivariate Smoothing
Week 16	(2022/12/21)		Homework 3 (No Class)
	(2022/12/31)		Final Project due

Statistical Software

You are expected to be familiar with one statistical software (e.g. Python, R, S+, SAS, Stata, MATLAB, etc.).

Textbooks and Course Materials

There is no textbooks in this course. Course materials are posted in my personal website every week after the class ends. The course website requires username and password to log in. Students who have registered in this class can request these information via email.

Course Policies and Academic Integrity

Please remember to turn off cell phones during lecture. You are expected to adhere to the honor code and code of conduct. As a student and memeber of National Central University community, you are here to get an education and are expected to demonstrate integrity in your academic endeavors. Students are subject to disciplinary action for several types of misconduct, including but not limited to, cheating, multiple submissions, plagiarism, prohibited collaboration, facilitating academic dishonesty, or knowingly furnishing false information. You may have assignments or projects in which you work with a partner or with a group. For example, you are welcome, and even encouraged, to work with others to solve homework problems. Even though you are working together, the assignment you submit for a grade must be in your own words, unless you receive specific instructions to the contrary.

Course Grades

There will be your attendance, a midterm examination, and a final project in the calculation of the final course scores:

$$\text{Final Score} = 0.15 \times \text{Attendance} + 0.30 \times \text{Homeworks} + 0.25 \times \text{Midterm Exam} + 0.30 \times \text{Final Project}$$

You are expected to attend all lectures of this course. Don't be late. Our TA will call your name at a specific moment of the class.

Homeworks

There will be three assigned homeworks, each results in a maximum of 10 points in the final grade.

Homework	Assigned Date	Due Date	Topics
1	2022/09/28	2022/10/11	Linear Algebra (Lecture 01-04)
2	2022/11/02	2022/11/15	Optimization (Lecture 05-08)
3	2022/12/14	2022/12/27	Statistics Computing (Lecture 09-12)

Midterm and Final Projects

Detailed information of the midterm exam and the final project will be given during the course.