STAT 6560 Applied Multivariate Analysis
Autumn 2023

Lecture: MWF 11:30am–12:25pm in Cockins Hall 232

Instructor: Yoonkyung Lee
Office: 440H Cockins Hall
Phone: 292-9495
Office Hours: M 3:00-3:55pm, R 2:00-2:55pm or by appointment
Email: yklee@stat.osu.edu or lee.2272@osu.edu

Grader: Kiljae Lee
Office Hours: by appointment only
Email: lee.10428@osu.edu

Course web page:
The course has a web page on Carmen (`https://carmen.osu.edu/`). You will find the class schedule, course announcements, homework assignments, and other information about the class on the web page. Please check it out on a regular basis.

Text: *Applied Multivariate Statistical Analysis*, 6th Edition by Richard A. Johnson and Dean W. Wichern (required)

Prerequisites: Stat 6450 (Applied Regression Analysis) or equivalent, Math 2568 (Linear Algebra) or equivalent, and some experience with statistical computing packages are required.

Course Description

Statistics 6560 is an introductory multivariate statistical analysis course designed for graduate students in the Department of Statistics. The aim of the course is to introduce a variety of standard statistical methods used to analyze multivariate data, emphasizing the implementation and interpretations of these methods. Topics covered include matrix computation of summary statistics, graphical techniques, the geometry of sample data, the multivariate normal distribution, inferences on a mean vector, principal component analysis, factor analysis, classification/discrimination, as well as cluster analysis and canonical correlation analysis if time permits.

Computing

Students will be required to use the R software environment for statistical computing and graphics. R can be downloaded for free at `http://www.r-project.org`. More information about computing with R can be found on the course web page. Instructions for data analysis using the software will be given in class. Most homework assignments will require some computing.

Grading

Homework: 30%
Midterm: 35% (tentatively around November 3, Friday)
Final take-home project: 35%

While the standard grading scale generally applies, final grades may be curved upwards.

Homework Assignments

Homework will be assigned regularly (about every other week) throughout the semester. Typically, homework assignments will involve computing and data analysis. Your report should include properly formatted computer output and graphs embedded in text for justification of your answer. Homework will be posted using Assignments on Carmen, and you should submit your report online. Due dates will be announced in class and on Carmen as well. For homework assignments requiring computing, computer code must be separately submitted online via Carmen as an appendix to each assignment. Your code should include comment statements that indicate what sections of the code correspond to the specific homework questions so that, if needed, the grader can read and check your code for its accuracy.

Late assignments: No late homework assignments will be accepted with few exceptions. If you have documented reasons for missing work or needing extra time, please contact me as soon as possible prior to the due dates. Where appropriate, due dates could be extended.

Final Project

The final project will involve data analysis using multivariate techniques. The data set for the project will be provided by the instructor. In your project report, you summarize results from your data analysis and describe findings in the style of a research article. Instructions on the project will be given on December 7 via Carmen, and the project report will be due on December 13 (Wednesday) by noon.

Academic Integrity Policy

Although students are encouraged to work together on assignments, each student is expected to write and submit individual solutions to homework problems. The midterm and final project are to be completed on your own without any external help or communication.

Academic misconduct will not be tolerated and will be dealt with procedurally in accordance with university policy. It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct at `http://studentlife.osu.edu/csc/`.

Accessibility accommodations for students with disabilities

The university strives to maintain a healthy and accessible environment to support student learning in and out of the classroom. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion.

If you are isolating while waiting for a COVID-19 test result, please let me know immediately. Those testing positive for COVID-19 should refer to the Safe and Healthy Buckeyes site (`https://safeandhealthy.osu.edu/tracing-isolation-quarantine`) for resources. Beyond five days of the required COVID-19 isolation period, I may rely on Student Life Disability Services to establish further reasonable accommodations. You can connect with them at slds@osu.edu; 614-292-3307; or `http://slds.osu.edu`.

Religious accommodations

It is Ohio State's policy to reasonably accommodate the sincerely held religious beliefs and practices of all students. The policy permits a student to be absent for up to three days each academic semester for reasons of faith or religious or spiritual belief.

Students planning to use religious beliefs or practices accommodations for course requirements must inform the instructor in writing no later than 14 days after the course begins. The instructor is then responsible for scheduling an alternative time and date for the course requirement, which may be before or after the original time and date of the course requirement. These alternative accommodations will remain confidential. It is the student's responsibility to ensure that all course assignments are completed.

Disclaimer

This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular we reserve the right to change due dates or the methods of grading and/or assessment if necessary. Any changes will be communicated to you through official course announcements.

Course Schedule (tentative)

Week	Dates	Topics, Readings, Assignments, Deadlines
1	8/22-8/25	Introduction, Data Organization, Summary Statistics, Graphics and Distance (Chapter 1)
2	8/28-9/1	Matrix Algebra and Geometry: Inner product, Projection, Spectral Decomposition, Quadratic Forms, Positive Definite Matrices (2.1-2.3)
3	9/4-9/8	9/4 (M): Labor Day Powers of Matrices, Random Vectors and Matrices, Mean Vectors and Covariance Matrices (2.4-2.6). **Homework 1 Due**: 9/8 (F)
4	9/11-9/15	Matrix Inequalities (2.7), Sample Geometry, Random Sampling, Expectation of Sample Mean and Covariance Matrix (3.2-3.3), Matrix Operations for Descriptive Statistics (3.5), Generalized Variance (3.4)
5	9/18-9/22	Multivariate Normal Distribution and Its Properties, MLE (4.1-4.3), Sampling Distribution of Sample Mean and Covariance (4.4). **Homework 2 Due**: 9/22 (F)
6	9/25-9/29	Large-Sample Behavior of Sample Mean and Covariance (4.5), Assessing Normality Assumption, Detecting Outliers, and Transformations to Near Normality (4.6-4.8)
7	10/2-10/6	Principal Component Analysis, Population Principal Components, Population Principal Components (8.1-8.2), Summarizing the Sample Variation by Principal Components (8.3). **Homework 3 Due**: 10/6 (F)
8	10/9-10/13	Selection of the Number of Components, Interpretation of the Sample Principal Components (8.3), Graphing Principal Components (8.4), Numerical Examples of PCA. 10/12 (R) - 10/13 (F): **Autumn Break**
9	10/16-10/20	Factor Analysis, Orthogonal Factor Model (9.1-9.2) Methods of Estimation - Principal Component Method (9.3). **Homework 4 Due**: 10/20 (F)
10	10/23-10/27	Maximum Likelihood Method, Likelihood Ratio Test (9.3), Factor Rotation (9.4)
11	10/30-11/3	Factor Scores (9.5), Data Examples 11/3 (F): **Midterm**
12	11/6-11/10	Introduction to Classification, Classification for Two Populations (11.1-11.2), Optimal Classification Rule, Bayes Error Rate (11.2), Expected Cost of Misclassification (11.2) 11/10 (F): **Veterans Day**
13	11/13-11/17	Classification with Two Normal Populations (11.3), Fisher's Linear Discriminant Analysis, Quadratic Discriminant Analysis (11.3), Classification with Several Populations (11.5). **Homework 5 Due**: 11/17 (F)
14	11/20-11/24	Fisher's Linear Discriminants (11.6), Evaluation of Classification Rules (11.4), Numerical Examples of Classification. 11/22–11/24: **Thanksgiving Break**
15	11/27-12/1	Inferences about a Mean Vector (5.1-5.2), Hotelling's T-square (5.3), Confidence Regions for a Mean Vector (5.4)
16	12/4-12/6	Likelihood Ratio Test for a Mean Vector (5.4) 12/7 (R): **Final project assigned**
17	12/13 (W)	**Final project report due by noon**