Lecture 0

Course Information

DSA 8070 Multivariate Analysis

Whitney Huang
Clemson University

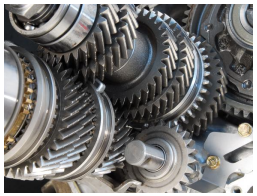
About the Instructor
Class Policies

About the Instructor

About the Instructor

- Associate Professor of Applied Statistics and Data Science
- Born in Laramie, WY, and grew up in Taiwan

- Obtained a B.S. in Mechanical Engineering and switched to Statistics in graduate school



- Earned a Ph.D. in Statistics in 2017 from Purdue University.

How to reach me?

- **Email:** wkhuang@clemson.edu
- **Office**: O-221 Martin Hall
- **Office Hours**: TBD

Class Policies

Logistics

- There will be three projects. The due dates are:
  - **Project I**: Sep. 25, Thursday
  - **Project I**: Nov. 6, Thursday
  - **Project II**: Dec. 11, Thursday
- There will be weekly R Labs:
  - To be uploaded to Canvas by 11:59 pm ET on the due dates
  - Worst grade will be dropped
- No lectures during Thanksgiving week (Nov. 24-28)

Course Materials at CANVAS

- Course syllabus / Announcements
- Lecture slides/notes/videos
- R Labs/Projects
- Data sets for lectures and labs

Reference Books

- *Applied Multivariate Statistics with* R, **Daniel Zelterman, 2015** [Link]
- *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, **Alan Izenman, 2008**, [Link]
- *Methods of Multivariate Analysis*, 3_{rd} Edition, **Alvin Rencher and William Christensen, 2012** [Link]
- *Applied Multivariate Statistical Methods,* 6_{th} Edition, **Richard Johnson and Dean Wichern, 2008** [Link]

Evaluation

Grades will be weighted as follows:

R Labs	25%
Project I	25%
Project II	25%
Project III	25%

Final course grades will be assigned using the following grading scheme:

>= 90.00	A
88.00 ~ 89.99	A-
85.00 ~ 87.99	B+
80.00 ~ 84.99	B
78.00 ~ 79.99	B-
75.00 ~ 77.99	C+
70.00 ~ 74.99	C
68.00 ~ 69.99	C-
<= 67.99	F

Computing

We will use software to perform statistical analyses.
Specifically, we will be using `R/Rstudio` 

- a **free/open-source** programming language for statistical analysis

- available at `https://www.r-project.org/` (`R`); `https://rstudio.com/` (`Rstudio`)

Topics

Week	Dates	Topic
1	8/20 - 8/22	Introduction and Multivariate Data Exploration
2	8/25 - 8/28	A Short Review of Matrix Algebra
3	9/2 - 9/5	Multivariate Normal Distributions and Copula Models
4	9/9 - 9/12	Inference and Comparison of Mean Vectors
5	9/15 - 9/19	Multivariate Regression I
6	9/22 - 9/26	Multivariate Regression II
7	9/29 - 10/3	Inference for Covariance Matrix
8	10/6 - 10/19	Principal Components Analysis
9	10/13 - 10/17	Factor Analysis
10	10/20 - 10/24	Canonical Correlation Analysis
11	10/27 - 10/31	Discrimination and Classification
12	11/3 - 11/7	Cluster Analysis
13	11/10 - 11/15	Multidimensional Scaling and Distance Embedding
14	11/17 - 11/21	Manifold Learning and Nonlinear Embedding Methods
15	11/24 - 11/28	**No Class–Thanksgiving**
16	12/1 - 12/5	Review