Multivariate Analysis – Honours 2024
Course Outline

Department of Statistical Sciences
University of Cape Town

Overview

Michael Greenacre describes multivariate analysis as "the area of statistics that deals with observations made on many variables. The main objective is to study how the variables are related to one another, and how they work in combination to distinguish between the cases on which the observations are made.

The analysis of multivariate data permeates every research discipline: biology, medicine, environmental science, sociology, economics, education, linguistics, archaeology, anthropology, psychology and behavioural science, to name a few, and has even been applied in philosophy. All natural and physical processes are essentially multivariate in nature – the challenge is to understand the process in a multivariate way, where variables are connected and their relationships understood, as opposed to a bunch of univariate processes, i.e. single variables at a time, isolated from one another."

This honours module will introduce students to the underlying theory and the application of several multivariate techniques.

Course structure and Class times

The course consists of two sections – which can be broadly considered as theoretical and applied, respectively – although there is some overlap. The two sections are delivered in different formats, to best suit the type of content each covers.

Section A will be presented in a *blended learning* format: Lessons (videos) will be released on Vula at the start of each week, and you are required to work through the material and complete the class exercises and online quizzes in your own time before Thursday of that week. There will also be dedicated forums for each week's content, where you can post questions before the Thursday lecture. During this face-to-face lecture, we will revise the most important aspects of the week's topic, answer questions, and go over the homework exercises. The second hour will be dedicated to each week's continuous assessment (see below).

Section B will entail 2-hour face-to-face lectures. You will also receive weekly continuous assessment to complete for this section.

Weekly lecture details:

Day and time: Thursdays, 09:00 - 11:00
Venue: Jordan 1C

Lecturers

Section	Lecturer	Email
A	Mr Stefan Britz (convener)	stefan.britz@uct.ac.za
B	Mr Miguel Rodo	miguel.rodo@uct.ac.za

Consultation hours, where applicable, will be communicated.

Prescribed Textbook

Sufficient material will be made available in the form of course notes and slides, much of which is based on the following book; students are encouraged to obtain a copy:

RA Johnson & DW Wichern, "Applied Multivariate Statistical Analysis", 6th edition, Pearson International Edition, 2007.

Additional Resources:

B Everitt, "An R and S-Plus Companion to Multivariate Analysis:", Springer, 2005.
AC Rencher, "Methods of Multivariate Analysis", second edition, Wiley, 2002.

Course Content

The syllabus is divided into two broad sections, covering the following topics:

Section A: Theory and Inference

Week 1: Introduction to and Visualisation of Multivariate Data
Week 2: Singular Value Decomposition, Eigenvalue Decomposition and Spectral Decomposition Revisited
Week 3: The Multivariate Normal Distribution
Week 4: Multivariate Maximum Likelihood Estimation
Week 5: Multivariate Inference
Week 6: MANOVA

Section B: Applications of Multivariate Analysis

Week 7: Multivariate Regression
Week 8: Principal Component Analysis
Week 9: Factor Analysis
Week 10: Canonical Correlation Analysis
Week 11: Discriminant Analysis
Week 12: Correspondence Analysis

Assessment

Sections A and B each contribute 50% to your final mark, with the breakdown of each deliverable as a **proportion of the final mark** as follows:

Section A

- Quizzes – 1% each x 6 = 6%
- Continuous assessments – 3% each x 6 = 18%
- Invigilated exam = 26%

The continuous assessments (CAs) in Section A will be done in groups of three, to which students will be randomly assigned each week. The second hour of each class will be "flipped", wherein the pairs will sit together and work on the week's CA, which is due the following Monday morning. Note that the bulk of each week's CA can realistically be completed within that hour. Therefore, students are encouraged to bring laptops/notebooks to class.

The 3-hour invigilated exam will take place during the first semester exam block, and will cover the theoretical aspects of Section A.

Section B

- Continuous assessments – 3% each x 5 = 15%
- Lab exam = 35%

The lab exam will also take place during the first semester exam block. The length and exact arrangements will be confirmed closer to the time, since this depends on the rest of the exam schedule and load shedding.

Duly Performed Requirements

Submission of all quizzes and continuous assessments.