Multivariate analysis and structural equation modeling

M.S. in Behavior and Cognition

Alberto Maydeu Olivares, amaydeu@ub.edu

OBJECTIVES

In Regression Analysis we obtained a solid understanding of what statistics is. We also covered some useful techniques, regression analysis and analysis of variance. In this course we focus on larger, more complex, and hence more realistic problems. By the end of this course students will be able to handle highly sophisticated research questions using some of the most advanced methods.

METHODOLOGY

The course is divided into two parts: 1) Classical multivariate analysis and test theory, and 2) structural equation modeling.

When many observations and variables are to be modeled it is necessary to reduce the amount of information available. The first part of this course is devoted to some useful techniques for data reduction: principal components analysis, exploratory factor analysis and cluster analysis. Along with factor analysis we will also cover test theory, that is, how to construct tests and how to assess their quality.

The second part of the course is devoted to multivariate modeling using systems of equations (i.e. structural equation modeling). Many techniques that we have seen can be described as a special case of this framework.

The first part of the course will be based on the familiar software SPSS, whereas for structural equation modeling we will use a more advanced package, Mplus. Information about this program can be obtained from http://www.statmodel.com There students can download a short overview of what the program is able of doing. **Students should download the program manual** from those web pages.

Each week we will have a seminar/lecture. As needed, additional computer laboratories will be scheduled. Before each seminar, students must read the material to be covered Seminars will intermix theory as well as applications. As a result, **students are encouraged to bring laptops to the seminars to reproduce what we do in class.**

BASIC TEXT AND MATERIALS

REQUIRED TEXTS

For the first and second part you are to follow

Raykov, T. & Marcoulides, G.A. (2008). *An introduction to applied multivariate analysis*. New York: Routledge.

Raykov, T. & Marcoulides, G.A. (2006). *A first course in structural equation modeling*. New York: Routledge.

respectively. In addition, you should read the relevant chapters from

Norusis, M. J. (2008). *SPSS 16. Statistical procedures companion*. Upper Saddle River, NJ: Prentice-Hall.

FURTHER REFERENCES

For your applied work, you may need a more in-depth coverage of the material. To that aim, the following list of textbooks is given. Within each topic, the books are arranged in order of difficulty:

Multivariate statistics

Hair, J.F, Black, W.C., Babin, B., Anderson, R.E., & Tatham, R.L. (2006). *Multivariate data analysis*. (6th ed). Englewood Cliffs, CA: Prentice Hall.
(very hands on, no formulae, even simpler than Tabachnick & Fidell)

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed). Thousand Oaks, CA: Sage.
(basic)

Tabachnick, B.G. & Fidell, L.S. (2006). *Using multivariate statistics* (5th ed). Allyn and Bacon
(easy reading)

Johnson, R.A. & Wichern, D. W. (2007). *Applied multivariate data analysis*. (6th ed). Englewood Cliffs, CA: Prentice Hall.
(the standard textbook for multivariate statistics)

Timm, N. H. (2006). *Applied Multivariate Analysis.* New York: Springer.
(advanced)

Mardia, K.V., Kent, J.T. & Bibby, J.M. (1979). *Multivariate analysis*. London: Academic Press.
(the most technical textbook, but very concise)

Structural equations modeling

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.
(the classic reference)

Kaplan, D. (2009). *Structural equation modeling. Foundations and extensions* (2nd ed). Thousand Oaks, CA: Sage.

McDonald, R.P & Maydeu-Olivares (forthcoming). *Structural models for multivariate data.*. New York: Routledge.
(may be made available to students, but it is quite technical)

PROGRAM

Week 1
Overview. Classification of univariate and multivariate techniques
Principal components
- Principal components as a data reduction technique
- How many components to retain
- Rotation
- Component scores

Chapter 17 of Norusis (2008)
Chapter 7 of Raykov and Marcoulides multivariate analysis book
Tech note: Principal components

Week 2
Factor analysis
- Factor analysis as a statistical model
- Methods of estimation (unweighted least squares, maximum likelihood)
- Goodness of fit assessment
- Factor scores and sum scores

Reliability analysis
- Coefficient alpha

How to construct a test
Chapter 8 of Raykov and Marcoulides multivariate analysis book
Tech note: Factor analysis and alpha
Tech note: The reliability of a scale score
Chapter 18 of Norusis (2008) –only description of alpha

Week 3
Cluster analysis
- Introduction
- Hierarchical and partitioning methods
- Assessing cluster solutions
- Two-step clustering

Chapter 16 of Norusis (2008)
Tech note: Cluster analysis

Week 4
Introduction to Structural equation modeling
- Introduction to Mplus
- Introduction to SEM
- Estimation methods
- Some basic applications
- Regression analysis as a SEM model

Chapters 1 and 2 of Raykov and Marcoulides SEM book
Tech note: Regression in Mplus

Week 5
The common factor model

- Identifiability and model choice
- Estimation and testing fit
- Improper solutions
- Some methodology for applications
- Exploratory factor analysis with Mplus
- Confirmatory factor analysis with Mplus
- Higher-order factors and hierarchical solutions

Chapter 9 of Raykov and Marcoulides multivariate analysis book
Tech note: Factor analysis using Mplus

Week 6
Path models
- Path diagrams
- The algebra of path models
- Identification
- Testing for mediational effects
- The problem of omitted variables
- Total effects, direct and indirect effects
- Standardized solution

Chapter 3 of Raykov and Marcoulides SEM book
Tech note: Path analysis
Tech note: Using summary statistics from published studies

Week 7
Structural equation modeling
- Estimation
- Asymptotically distribution free estimation
- Assessing model fit
- Tests for nested models
- Testing for close fit
- Modification indices

Tech note: Goodness of fit in structural equation models

Week 8
Ordinal factor analysis
Introduction to Item Response Theory
Analysis with multiple populations
- Assessing measurement invariance

Formative vs. reflexive indicators

Week 9
Further topics in structural equation modeling
- Analysis with missing data
- Bootstrapping

Week 10
Buffer
Exam

EVALUATION CRITERIA

Weekly exercises will be assigned. In addition, students are asked to turn in a short paper (5 pp maximum) where they replicate a structural equation model previously published in the literature. More specifically students must search for a paper of their interest where SEM is applied and the covariance (or correlation) matrix is provided. Students are asked to reproduce the model reported and explore additional competing models. When turning in the short paper, a copy of the manuscript must be enclosed, along with the datafile suitable for input into Mplus. Additionally, the different Mplus input files must be provided and the paper must provide as an appendix a description of each Mplus file.

Grades will be assigned as follows:

Class discussions and assigned exercises	50%
Final paper	50%