



Doctoral Program
Department of Statistics
University of Padua

CORRESPONDENCE ANALYSIS

Michael Greenacre
Universitat Pompeu Fabra, Barcelona

Correspondence analysis is part of a family of multivariate statistical methods which aims to visualize the main dimensions of a rectangular table of data, facilitating interpretation and understanding of patterns in the data. Within this family it is probably the most versatile technique, because it is primarily aimed at categorical data and incorporates the idea of weighting the rows and columns of the table. This method has long been used in the social and environmental sciences and is increasingly popular in marketing research, archaeology, psychology and linguistics.

In this course, the first day will be spent giving a comprehensive introduction to simple correspondence analysis, emphasising the geometric concepts that underlie the method. On the second day we shall discover multiple correspondence analysis, which analyses several categorical variables simultaneously and which is used extensively in survey research. Finally, on the third day, some related topics are discussed, such as data coding for other types of variables (preference, paired comparison and continuous data) as well as the analysis of square tables.

Throughout the course the program R will be used, specifically the **ca** package developed by Nenadic and Greenacre (2007).

Bibliography

Greenacre, M.J. and Blasius, J., eds (2006) *Multiple Correspondence Analysis and Related Methods*. Chapman & Hall/CRC, London.

Greenacre, M.J. (2007) *Correspondence Analysis in Practice (Second Edition)*. Chapman & Hall/CRC, London.

Nenadic, O. & Greenacre, M.J. (2007) Correspondence analysis in R, with two- and three-dimensional graphics: the **ca** package. *Journal of Statistical Software*, **20**(3).

URL <http://www.jstatsoft.org/v20/i03/>

(Note: *download this article*)

Program

Correspondence Analysis

Michael Greenacre

- 16 June 2008, 10:00-12:00 **Introduction to correspondence analysis (CA)**
Main concepts: profile, mass and chi-square distance.
Dimension reduction: singular value decomposition.
- 16 June 2008, 14:30-16:30 **Graphical interpretation and diagnostics**
Supplementary points.
Measuring and decomposing total variance.
Using the **ca** package in R for simple CA.
- 17 June 2008, 10:00-12:00 **Multiple correspondence analysis (MCA)**
Indicator and Burt matrices.
Measuring and decomposing inertia.
- 17 June 2008, 14:30-16:30 **Variants of MCA**
Joint correspondence analysis (JCA).
Subset MCA.
Using the **ca** package in R for MCA, JCA and subset analyses.
- 18 June 2008, 10:00-12:00 **Related topics**
Data coding: doubling, and fuzzy coding in the analysis of preferences, ratings and continuous-scale data.
Analysis of square tables.

Michael Greenacre is Professor of Statistics in the Department of Economics and Business at the Universitat Pompeu Fabra in Barcelona. He has given courses on correspondence analysis and multivariate analysis to students and researchers in 15 countries, most recently to marine biologists at the Norwegian College of Fishery Science, Tromsø, Norway, and to postgraduate students at BI School of Management in Oslo. He has participated as a statistician in various environmental research projects and the ARCTOS network. His research interests are in the analysis of large data sets in the social and environmental sciences, having authored and co-edited six books and numerous journal articles on correspondence analysis and data visualization. He is also a musician and composer.

Web: www.econ.upf.edu/~michael
 www.globalsong.net

Email: michael.greenacre@upf.edu
 michael.greenacre@gmail.com