Course Description: Simple and multiple correspondence analysis

In the social sciences, multiple correspondence analysis (MCA) is a statistical technique that first and foremost has become known through the work of the late Pierre Bourdieu (1930-2002), in particular “Distinction” (Bourdieu 1984), “Homo Academicus” (Bourdieu 1988) and “The State Nobility” (Bourdieu 1996).

As a counterpart to principal component analysis (PCA), a geometric technique for the analysis of metric variables, MCA is a geometric technique for the analysis of categorical or categorized variables. Originating in the early 1960s and the French statistician Jean-Paul Benzécri’s work in mathematical linguistics, MCA represents and models data sets as clouds of points in a multidimensional Euclidean space. The interpretation of the data is based on these clouds of points. By combining MCA with inferential techniques and variance analysis, we arrive at an integrated framework of interpretation that also is known under the name of Geometric Data Analysis (GDA).

In a combination of lectures and laboratory exercises, this course will introduce students to the fundamental properties, procedures and rules of interpretation of the most commonly used forms of correspondence analysis, i.e. simple correspondence analysis (CA) and MCA, and also to the most commonly used software.

Particular attention will be paid to how MCA can be used in the construction of social spaces.

The course starts with an historical introduction to Benzécri’s work on contingency tables, and to the key ideas and basic properties in geometric data analysis. A first
explanation of the procedures, the key concepts and the fundamental rules of interpretation will be given through a simple correspondence analysis (CA) of a standard contingency table.

Thereafter, and using Pierre Bourdieu’s work as a point of reference, we will go through the generalisation from CA to MCA by analysing an Individuals x Variables table. Particular emphasis will be put on the definition of distances between individuals, of distances between categories or modalities, the fundamental rules for the interpretation of axes in MCA, on how MCA can be integrated with variance analysis, and also on more general guidelines and coding principles.

We will then proceed to the more detailed exploration of the cloud of individuals, the introduction of supplementary variables, the use of concentration ellipses and how MCA also can be used in a confirmatory or explanatory mode by introducing variables as structuring factors in the constructed space. Tools for statistical inference, i.e. confidence ellipses around mean modality points in factorial planes, will also be presented, as will also the integration of MCA and hierarchical clustering as a way of further validation of the results.

Lectures and reading list:

Day 1: May 17th 2011

Lecture 1:

1. General Introduction. Correspondence Analysis in the Social Sciences. Some Basic Properties in CA


Lecture 2, Day 1.

2. (Simple) Correspondence Analysis. Basic Concepts and Properties. Interpretation of Results.


Lecture 3, Day 1:

3. From Simple to Multiple Correspondence Analysis.


Chapter 5: Multiple Correspondence Analysis. Pp. 214-241.


Day 2:

Lecture 1:

4. Multiple Correspondence Analysis. Specific Multiple Correspondence Analysis. Examination of The Cloud of Individuals.


**Day 2, Lecture 2**

5. **Stability Issues, Coding and Statistical Significance. Integration of Cluster Analysis and MCA.**


**Day 2, Lecture 3:**

6. **Constructing Social Spaces.**


**Lab sessions:**

Brigitte Le Roux, Mikael Börjesson, Philippe Bonnet (2006): *Performing Multiple Correspondence Analysis (MCA) USING SPAD 1 (VERSION 6.5)* Can be downloaded from [http://www.skeptron.uu.se/broady/sec/p-gda-0609-spadguide-mca.pdf](http://www.skeptron.uu.se/broady/sec/p-gda-0609-spadguide-mca.pdf) (more material can be downloaded from the same site)
Data sets: TBA

SPSS 13.0. Categories. SPSS Inc.

Outline – lab sessions:

1. Introduction to SPAD
   - SPAD – import files
   - setting up first analyses
   - getting to know the software
   - CA
   - a first introduction to the output file
   - interpretation of results
   - Own analyses of other variables

2. SPAD - MCA.
   - MCA/MCAspe module in SPAD
   - Detailed interpretation of results
   - The graphic module in SPAD – how to manipulate the graphs
   - First examination of the cloud of individuals
   - Graphical tools of interpretation
   - Use of supplementary variables

3. SPAD – Cluster Analysis and MCA.
   - More on SPAD – MCA/MCAspe module
   - More detailed examination of the cloud of individuals
   - Concentration, Indicator and confidence ellipses
   - More on supplementary variable
   - Integration of cluster analysis and MCA
   - Interpretation of cluster analyses in the factorial planes