AVANT-PROPOS

Due to the extensive field of application, Multiple Factor Analysis (MFA) is today widely used. This book is the first comprehensive account of the method in English : it brings together the theoretical and methodological aspects with application examples and details how to implement them using an R package (**FactoMineR**).

In the same way as for principal component analysis (PCA) or multiple correspondence analysis (MCA), MFA is applied to tables in which a set of individuals (one individual = one row) are described by a set of variables (one variable = one column). The particularity of MFA lies in the fact that within the active variables, it can account for a group structure defined by the user. Such data tables are named « individuals × variables organised into groups ».

This data format is widely used, firstly because it corresponds to the user's approach when designing data-collection. For example, someone making an opinion poll organises the questionnaire into themes, each of which are then developed using several questions (the questions are the variables and the themes the groups of variables). This structure must of course be present when analysing the results of the poll. Secondly, it is widespread because users often want to bring together data collected for the same statistical individuals but in different contexts (for example geographical and/or temporal). Thus, in the food industry, for a given set of products, we often have sensory profiles from tastings conducted in different countries. These sets of data need to be analysed simultaneously while at the same time preserving their individuality both during the statistical analysis and the interpretation phase.

Experience in working with many diverse users has shown that multiple tables are in fact the standard data format used today. Along with this complex structure (into groups of variables), the nature of the data is also complex as the variables can be quantitative and/or qualitative. It is therefore necessary for users to have access to methodology for analysing individuals × variables tables in which the variables are structured into quantitative, qualitative or mixed groups. This is precisely the field of application of MFA.

MFA is the result of joint research by Brigitte Escofier and Jérôme Pagès in the early 80's. This method is now well established if we consider the wide range of software available. To name but a few software including an MFA procedure : **FactoMiner** (R Package) **ade4** (R Package), SPAD, Uniwin (Statgraphics), XLStat.

Having achieved widespread availability of the method, and with the data format justifying its implementation, one question remains : What exactly does « account for a group structure of the variables in an overall analysis » mean ? In other words, why do we not simply conduct, for example, a principal component analysis and take the group structure of the variables into account solely during interpretation ? We might summarise by saying that this book answers this question first and foremost.

The first two chapters look back at basic factorial analysis methods for indivi-

duals × variables tables, PCA and MCA.

Chapter 3 presents Factor Analysis for Mixed Data (FAMD), a little-known method for simultaneously analysing quantitative and qualitative variables without group distinction. FAMD contains the technical elements required for taking into account both types of variables within one single analysis.

The following chapters, numbered 4 to 9, describe multiple factor analysis. The first four look in turn at the key points of MFA in the context of quantitative variables. In addition, one chapter is given over to qualitative and mixed data. Finally, one chapter compares MFA and Procrustes analysis.

Chapter 10 presents a natural extension of MFA : hierarchical MFA (HMFA). In this method, the variables are not structured by a simple partition, but by a hierarchy. A typical example of this data is surveys, for which the questionnaire is organised into themes and sub-themes.

The final chapter presents several elements of matrix calculation and metric spaces used in the book in the form of two technical appendices.

***

To conclude this work, it gives me great pleasure to thank Sophie Puyo and Magalie Houée-Bigot, the statistical engineers who were responsible for the majority of the layout for this book. I also thank Eric Matzner-Løber, editor for the French version of this book for his contributions. Thanks also go to Rob Calver at Chapman & Hall for his friendly assistance. A special mention goes to Rebecca Clayton for her invaluable help in translating from the French to the English version. I would also finally like to thank my wife Annie who brightens my life and therefore, indirectly, this book.

***

The data and the R scripts used in this book are available on the website of the applied mathematics department at Agrocampus Ouest.

Chapters 3, 8 and 9 are adapted from works first published in the *Revue de Statistique Appliquée* (*Journal of Applied Statistics*, which ceased publication in 2006). This is an excellent occasion to thank Pierre Cazes, director of the journal, first for his enthusiastic reception of the work and then for his continued encouragement concerning the adaptation for the book.

Thanks also go to Richard Delécolle and his talent for calligraphy.