

*Analyse exploratoire de tableaux multiples quantitatifs et/ou qualitatifs.
Application de l'Analyse Factorielle Multiple Hiérarchique à des données
sensorielles recueillies par la méthode du napping catégorisé.*

Jérôme Pagès et Sébastien Lê

Laboratoire de mathématiques appliquées

Agrocampus, Rennes.

Notre cadre est l'analyse exploratoire d'un tableau individu×variables par une méthode factorielle (analyse en composantes principales lorsque les variables sont quantitatives, analyse des correspondances multiples lorsque les variables sont qualitatives).

En pratique, les variables sont très souvent structurées en groupe, l'exemple type étant celui d'une enquête dont les questions (i.e. les variables) sont structurées en thèmes. Prendre en compte une telle structure en groupe de variables dans une analyse factorielle pose le problème technique de l'équilibre entre ces groupes mais aussi induit des questions spécifiques qui enrichissent les résultats. A cette complexité de structure (en groupes de variables), s'ajoute une complexité de nature, les variables pouvant être quantitatives et/ou qualitatives. Il est donc nécessaire de disposer d'une méthodologie d'analyse de tableaux individu × variables, dans lesquels les variables sont structurées en groupes quantitatifs, qualitatifs ou mixtes. Tel est précisément le champ d'application de l'Analyse Factorielle Multiple (AFM).

De plus en plus souvent, on doit analyser des tableaux dans lesquels les variables ne sont pas structurées seulement en groupe (i.e. selon une partition) mais selon une hiérarchie ou, si l'on préfère, une suite de partitions emboîtées. Un exemple typique de ces données est fourni par les enquêtes dont le questionnaire est structuré en thèmes et en sous-thèmes. L'Analyse Factorielle Multiple Hiérarchique (AFMH) traite précisément des tableaux dans lesquels les variables, quantitatives et/ou qualitatives, sont structurées selon une hiérarchie.

Dans un premier temps, nous présenterons, à partir d'exemples, les points méthodologiques clés soit :

- 1) Que veut dire prendre en compte une structure des variables en groupes ;
- 2) Comment analyser simultanément des variables quantitatives et qualitatives.

Dans un second temps, nous décrirons la mise en œuvre des méthodes présentées à l'aide du package R FactoMineR. Pour illustrer cette mise en œuvre, nous utiliserons des données issues de la méthode du napping® catégorisé. Dans cette méthode on demande à des sujets (e.g. des consommateurs) d'évaluer un ensemble de stimuli (e.g. des produits alimentaires) de la façon suivante.

- 1) Chaque sujet place, selon ses propres critères, l'ensemble des stimuli sur une nappe de façon telle que deux stimuli se trouvent d'autant plus proches qu'ils se ressemblent.
- 2) Une fois les stimuli placés, le sujet effectue des regroupements (i.e. une partition des stimuli) et associe à chacun des groupes de stimuli ainsi créés un ou plusieurs termes qui lui semble bien le caractériser.

Enfin, pour chaque sujet on dispose de données quantitatives (les coordonnées des stimuli sur la nappe), qualitatives (une partition) et textuelles (les termes décrivant les groupes de stimuli). On souhaite visualiser ces données en équilibrant les données de types différents au sein d'un même sujet et en équilibrant les sujets entre eux. L'AFMH est adaptée au traitement de telles données.

Un exemple réel de telles données sera présenté et analysé. A cette occasion, on présentera un logiciel d'acquisition des données (selon la procédure du napping® catégorisé) sur tablette.

Références

Pagès Jérôme (2013) Analyse factorielle multiple avec R. 253 p. *EDP sciences*, Paris.

Francois Husson, Julie Josse, Sebastien Le and Jeremy Mazet (2013). FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.25. <http://CRAN.R-project.org/package=FactoMineR>