

Séminaire

pour les professeurs et les étudiants de 2^e et 3^e cycles

La détermination du nombre de dimensions en analyse factorielle exploratoire : bien mieux que l'analyse parallèle

André Achim
Professeur (UQAM)

En analyse factorielle exploratoire, l'analyse parallèle (AP) (Horn, 1965) a longtemps été considérée égale ou supérieure à toute autre méthode pour déterminer le nombre de dimensions communes (exprimées par plus d'une variable) d'un ensemble de données. L'AP compare les valeurs propres de l'ensemble de données à celles produites de façon aléatoire pour des variables linéairement indépendantes les unes des autres. Toutefois, lorsqu'on a constaté, par exemple, qu'il doit y avoir au moins deux dimensions, il n'est pas optimal de comparer la troisième valeur propre de l'ensemble de données à celles obtenues à l'aide de l'AP, qui ne reconnaît, pour sa part, aucune dimension commune. Ruscio et Roche (2012), dans une approche appelée *Comparison Data* (CD), ont proposé l'idée de produire des données parallèles ayant une structure factorielle sous diverses hypothèses quant au nombre de dimensions communes. Leurs simulations ont montré une nette amélioration par rapport à l'AP. L'algorithme qu'ils ont utilisé détermine le nombre requis de dimensions, k , lorsque l'erreur quadratique moyenne des simulations obtenue avec $k+1$ facteurs ne diffère pratiquement plus de celle obtenue avec k facteurs. Une nouvelle procédure, appelée *Factor Space Dimensionality Assessment Algorithm* (FSDAA), utilise un critère alternatif de rang de la $k+1$ i^{ème} composante principale de l'ensemble de données parmi celles obtenues par simulation reproduisant la structure avec k facteurs. La comparaison sur diverses structures factorielles confirme que la stratégie CD est supérieure à la stratégie AP, mais indique que la stratégie FSDAA l'est encore davantage. En absence de variable orpheline (ne dépendant d'aucun facteur commun), la stratégie FSDAA ne surévalue jamais le nombre de dimensions et sous-évalue celui-ci seulement par manque de puissance statistique. La stratégie FSDAA peut même permettre d'identifier des dimensions avec des valeurs propres nettement inférieures à l'unité. Le code nécessaire à l'application de la stratégie FSDAA est disponible sous R, Excel et Matlab.

Cdame

Collectif pour le développement et les applications en
mesure et évaluation

Coordonnées:
www.cdame.uqam.ca
n.talbot@videotron.ca

Lundi 17 mars 2014 | 14:00 à 16:00 | N-6320 |