

Modèle libre généralisé pour équations structurelles

Christian Derquenne (EDF R&D)

Résumé

Les modèles à équations structurelles sont souvent utilisés en marketing pour construire des indices de satisfaction [5]. Les relations structurelles sont définies par des liens "causaux" entre des variables non observables, dites variables latentes et des variables observées dites manifestes. Les variables latentes correspondent aux facettes ou aux composantes de la satisfaction des consommateurs ; les variables manifestes ont trait aux réponses des clients aux enquêtes de satisfaction. L'étude des liens entre les variables latentes s'appuie sur différents modèles conceptuel marketing provenant de la théorie du processus de décision des consommateurs. Le choix du modèle est primordial, car il conditionne par la suite toute l'étude des liaisons. Dans de nombreuses applications, le modèle initial est créé sur des dires d'experts. Par exemple, dans le cas de la psychologie quantitative, un expert mettra en évidence des relations entre variables manifestes et variables latentes, et entre variables latentes, elles-mêmes. En marketing, c'est cette approche qui est préférée. Les deux méthodes statistiques traditionnelles utilisées pour estimer ces modèles sont la méthode LISREL (LInear Structural RELationships) [7] et l'approche PLS (Partial Least Squares) [9,10,11]. Une autre approche nommée RFPC (Regression on First Principal Components) a été introduite [1,2]. Les deux dernières méthodes ont deux avantages par rapport à la première : la capacité à traiter les données manquantes et la convergence assurée dans l'étape d'estimation du modèle.

Cependant, ces modèles sont figés au moins à quatre niveaux dans :

- le choix des variables manifestes correspondant aux questions de l'enquête, par exemple,
- la détermination des blocks de variables manifestes constituant les variables latentes,
- la présence ou l'absence de liens entre les blocks,
- le choix du sens de ces liaisons.

Si l'on considère que le premier niveau reste imposé, les trois autres peuvent en revanche être remis en question. Nous pouvons laisser "parler les données", notamment en utilisant l'analyse exploratoire. Dans ce cadre, nous avons développé une approche pour construire des "modèles libres", lorsque les variables manifestes sont mesurées sur une échelle numérique. Cette approche se déroule en quatre étapes : classification des variables manifestes en utilisant une analyse factorielle avec rotation oblique pour construire les blocks, sélection de la première composante principale pour obtenir une variable latente par block, recherche des liens entre les variables latentes à l'aide des corrélations linéaires partielles et décision par l'expert du sens des liens (solution préférable, mais il est également possible d'utiliser des critères statistiques). Ce modèle libre est alors estimé par l'approche RFPC. L'objectif de la présentation est de généraliser cette approche [3,4] à d'autres échelles de mesure des variables manifestes (booléenne, ordinale, nominale, comptage, ...). Cette approche est appliquée sur des données réelles en marketing, puis les forces et les faiblesses de celle-ci sont établies par rapport à d'autres méthodes [6,8]. Enfin, nous discutons des voies potentielles de recherches.

Bibliographie

[1] Derquenne, Ch., (2001), Une méthode alternative à l'approche PLS, rapport technique EDF R&D, Clamart.

[2] Derquenne Ch. et Hallais C., (2003), Une méthode alternative à l'approche PLS : comparaison et application aux modèles conceptuels marketing, Revue de Statistique Appliquée, LII, 37-72.

- [3] Derquenne Ch., (2005), Generalized Path Modelling based on the Partial Maximum Likelihood Approach, PLS'05, Fourth International Symposium on PLS and Related Methods, Barcelona – Spain.
- [4] Derquenne, Ch., (2007), Free model for Generalized Path Modelling, PLS'07, Fifth International Symposium in Causalities explored by indirect observation, Oslo, Norge.
- [5] Fornell, C., (1982), A National Customer Satisfaction Barometer : The Swedish Experience, *Journal of Marketing*, 56, pp. 5-21.
- [6] Jensen F.V., (2001), *Bayesian Networks and Decision Graphs*, Springer, New-York.
- [7] Joreskog, K.G. and Sorbum, D., (1979), *Advances in Factor Analysis and Structural Equation Models*, Abt Books, Cambridge.
- [8] Hui B.S., (1982), On Building Partial Least Squares Models with Interdependent Inner Relations, pp. 249-271 in K.G. Jöreskog & H. Wold (eds.), *Systems under Indirect Observations: Causality, Structure, Prediction*, Vol 2, Amsterdam, North-Holland.
- [9] Lohmöller, J.B., (1989), *Latent Variable Path Modeling with Partial Least Squares*, Physica-Verlag, Heildelberg.
- [10] Tenenhaus, M. et al., (2005), *PLS Path Modeling*, *Computational Statistics & Data Analysis*, Elsevier (Ed.).
- [11] Wold, H., (1975), *Modelling in Complex Situations with Soft Information*, Third World Congress of Econometrics Society, August 21-26, Toronto, Canada.