Quaderni di STATISTICA

VOLUME 14 - 2012

LIGUORI EDITORE

Quaderni di Statistica
Vol. 14, 2012

Globally-optimized latent variable extraction in formative-reflective models

Marco Fattore Giorgio Vittadini
Department of Quantitative Methods, University of Milano-Bicocca
E-mail: marco.fattore@unimib.it, giorgio.vittadini@unimib.it

Matteo Pelagatti
Department of Statistics, University of Milano-Bicocca
E-mail: matteo.pelagatti@unimib.it

Summary: In this paper, we propose a novel globally optimal procedure to extract exogenous and endogenous latent variables (LVs) in formative-reflective structural equation models. The procedure is a valuable alternative to PLS-PM and Lisrel, since it is fully consistent with the causal structure of formative-reflective schemes and extracts both the structural parameters and the factor scores without identification or indeterminacy problems. The algorithm estimates the structural model taking into acccount both the capability of the exogenous LVs to represent their manifest formative blocks and the capability of the endogenous LVs to explain the manifest reflective blocks. It can be applied to virtually any kind of formative-reflective scheme and can be easily implemented in any programming language with numerical optimisation capabilities.

Keywords: Structural Equation Models, Formative-reflective scheme, PLS-PM.

1. Introduction

Formative-reflective models are a standard tool in socio-economic research, particularly in the fields of causal modeling and multidimensional evaluation. Despite the relevance of the topic there are still unsolved methodological problems when dealing with formative constructs (Howell et al., 2007; Wilcox, 2008). In fact, these models are usually estimared using the Lisrel algorithm, which is affected by indeterminacy problems (Vittadini, 1989), or using the PLS-PM algorithm, which cannot handle reflective relationships properly (Vittadini et al., 2007). In a formative-reflective scheme, the exogenous latent variables play a double role. On the one hand, they should summarize their formative blocks; on the other hand, they should mediate, *via* the system of endogenous latent variables, the causal relationships linking the formative side to the reflective side. Realizing this, the proposed procedure extracts the exogenous latent variables balancing between these two aspects.

2. Formative-reflective models

Let $\boldsymbol{x}_i$, $i = 1, \ldots, p$, and $\boldsymbol{y}_j$, $j = 1, \ldots, q$, be vectors of zero-mean manifest variables and let $\boldsymbol{\omega}_i$, $i = 1, \ldots, p$, be vectors of real coefficients. According to the formative-reflective scheme, each exogenous LV ξ_i is expressed as a linear combination of the MVs of the corresponding formative group:

$$\xi_i = \boldsymbol{\omega}_i' \boldsymbol{x}_i, \qquad i = 1, \ldots, p. \tag{1}$$

By stacking $\xi_1, \ldots, \xi_p$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p$ into the vectors $\boldsymbol{\xi}$ and $\boldsymbol{x}$ respectively, definitions (1) can be cast in the following compact form

$$\boldsymbol{\xi} = \boldsymbol{\Omega} \boldsymbol{x}, \tag{2}$$

with

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\omega}_1' & \mathbf{0}' & \ldots & \mathbf{0}' \\ \mathbf{0}' & \boldsymbol{\omega}_2' & \ldots & \mathbf{0}' \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}' & \mathbf{0}' & \ldots & \boldsymbol{\omega}_p' \end{bmatrix}, \quad \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_p \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_p \end{bmatrix}.$$

In an analogous way, the vector $\boldsymbol{y}_j$ of MVs of the j-th reflective block is assumed to be built as sums of a rescaled common scalar endogenous LV η_j plus a residual ε_j,

$$\boldsymbol{y}_j = \boldsymbol{\lambda}_j \eta_j + \boldsymbol{\varepsilon}_j, \qquad j = 1, \ldots, q,$$

where $\boldsymbol{\lambda}_j$, $j = 1, \ldots, q$, are real vectors. Again, stacking $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_q$ and $\varepsilon_1, \ldots, \varepsilon_q$ into the vectors $\boldsymbol{y}$ and ε, respectively, we get

$$\boldsymbol{y} = \boldsymbol{\Lambda} \boldsymbol{\eta} + \varepsilon, \tag{3}$$

with

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\lambda}_q \end{bmatrix}, \quad \boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_p \end{bmatrix}.$$

Finally, the endogenous LVs stacked in vector $\boldsymbol{\eta}$ are built as linear combinations of the exogenous LVs stacked in vector $\boldsymbol{\xi}$, namely

$$\boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{\xi}, \tag{4}$$

where $\boldsymbol{\Gamma}$ is a conformable matrix of real coefficients. By putting (2), (3) and (4) together we obtain the final model, linking the formative MVs and the reflective MVs, *via* the constrained latent structure expressed by the matrices $\boldsymbol{\Lambda}, \boldsymbol{\Gamma}, \boldsymbol{\Omega}$:

$$\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{x} + \boldsymbol{\varepsilon}.$$

2.1. Latent variables extraction

The extraction of variables $\xi_1, \ldots, \xi_p$ requires a compromise between two goals: on one hand, each of them should summarize effectively its own formative block; on the other hand, as a whole, they should indirectly predict (*via* the set of endogenous LVs) the variables in the reflective groups. The first goal would be achieved extracting exogenous LVs through the minimization of the following loss function:

$$L_x(\boldsymbol{\Pi}, \boldsymbol{\Omega}) = \frac{1}{p}\sum_{i=1}^{p} \frac{\mathbb{Tr}\{\mathbb{E}[(\boldsymbol{x}_i - \boldsymbol{\pi}_i\boldsymbol{\omega}_i'\boldsymbol{x}_i)(\boldsymbol{x}_i - \boldsymbol{\pi}_i\boldsymbol{\omega}_i'\boldsymbol{x}_i)']\}}{\mathbb{Tr}\{\mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i']\}} \tag{5}$$

where $\boldsymbol{\pi}_i, i = 1, \ldots, p$, are vectors of regression coefficients of $\boldsymbol{x}_i$ on $\boldsymbol{\omega}'\boldsymbol{x}_i$, and

$$\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{\pi}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\pi}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\pi}_p \end{bmatrix}.$$

The second goal would be achieved by extracting the exogenous (and thus the endogenous) LVs through the minimization of the following loss function

$$L_y(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Lambda}) = \frac{1}{q}\sum_{j=1}^{q} \frac{\mathbb{Tr}\{\mathbb{E}[(\boldsymbol{y}_j - \boldsymbol{\lambda}_j\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\boldsymbol{x})(\boldsymbol{y}_j - \boldsymbol{\lambda}_j\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\boldsymbol{x})']\}}{\mathbb{Tr}\{\mathbb{E}[\boldsymbol{y}_j\boldsymbol{y}_j']\}}, \tag{6}$$

where $\boldsymbol{\gamma}_j'$ is the j-th row of matrix $\boldsymbol{\Gamma}$ and the vectors $\boldsymbol{\lambda}_j$ are regression coefficients of $\boldsymbol{y}_j$ on $\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\boldsymbol{x}$. If the matrix $\boldsymbol{\Omega}$ were given, and so the latent variables $\boldsymbol{\xi} = \boldsymbol{\Omega}\boldsymbol{x}$, then L_y would be minimised by a rank-one reduced rank regression of each $\boldsymbol{y}_j$ on $\boldsymbol{\xi}$. The problem of extracting the exogenous and endogenous LVs taking into account both competitive goals can be solved by minimizing the following global loss function

$$L^{(\alpha)}(\boldsymbol{\Pi}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Lambda}) = (1 - \alpha)L_x(\boldsymbol{\Pi}, \boldsymbol{\Omega}) + \alpha L_y(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Lambda}). \tag{7}$$

where $\alpha \in [0, 1]$ determines the relative weight given to each goal. When $\alpha = 0$, ω_i is just the first eigenvector of $E[\boldsymbol{x}_i\boldsymbol{x}_i']$, ξ_i is the first principal component of the variables in the i-th formative block and $\boldsymbol{y}_j = \boldsymbol{\lambda}_j\gamma_j'\boldsymbol{\xi} + \boldsymbol{\varepsilon}_j$ is an ordinary rank-one reduced rank regression of $\boldsymbol{y}_j$ on $\boldsymbol{\xi}$. On the contrary, when $\alpha = 1$ the latent variables ξ_i's are built as the linear combinations of the respective $\boldsymbol{x}_i$ that best fit, *via* the endogenous latent variables, the vector $\boldsymbol{y}$, and the whole problem reduces to a multivariate regression with many constraints, implied by the form of matrix $\boldsymbol{B} = \boldsymbol{\Lambda\Gamma\Omega}$.

3. First results and conclusion

The extraction methodology has been applied to both simulated and real datasets and its performance has been compared to that of PLS-PM. In general terms, the methodology proves effective in extracting latent variables and capturing causal links among them. Letting the parameter α moving in $[0, 1]$, and comparing the extracted exogenous latent variables as α varies, the methodology also reveals whether the formative side and the reflective side of the model are consistently specified, or whether the goals of representing the formative blocks and explainng the reflective manifest variables cannot be jointly achieved. Real data application shows that our methodology produces results matching the causal structure of the model much better than PLS-PM and reveals the existence of causal links even when PLS-PM does not, with comparable predictive power. In fact, PLS-PM expresses endogenous latent variables as linear combinations of reflective manifest variables. This eventually leads to extracting exogenous latent variables without accounting properly for the formative blocks and weakening the causal links from the formative to the reflective side of the model. Differently, our methodology is designed exactly to take into account both sides, resulting in more balanced and interpretable results.

References

Howell R. D., Breivik, E., Wilcox, J.B. (2007), Reconsidering formative measurement, *Psychological Methods*, 12, 205-218.

Vittadini G. (1989), Intederminacy problems in the Lisrel model, *Multivariate Behavioral Research*, 24, 397-414.

Vittadini G., Minotti S., Fattore M., Lovaglio P.G. (2007), On the relationships among latent variables and residuals in PLS path modeling: the formative-reflective scheme, *Computational Statistics & Data Analysis*, 51, 5828-5846.

Wilcox J.B., Howell R.D., Breivik E. (2008), Questions about formative measurement, *Journal of Business Research*, 61, 1219-1228.