Computational Statistics and Data Analysis 120 (2018) 84–97

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A least squares approach to latent variables extraction in formative–reflective models

Marco Fattore, Matteo Pelagatti *, Giorgio Vittadini

University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, Italy

HIGHLIGHTS

- An effective way to extract latent variables in formative–reflective path models.
- The objective function of the model is clearly stated, as opposed to PLS-PM.
- Performs better than PLS-PM and SEM in formative–reflective schemes.
- The R package "pathmod" makes the application of the method straightforward.

ARTICLE INFO

Article history:
Received 18 October 2016
Received in revised form 12 November 2017
Accepted 13 November 2017
Available online 22 November 2017

Keywords:
Path model
Formative–reflective model
Least squares
Reduced rank regression
PLS-PM
SEM

ABSTRACT

A new least-squares based procedure for the extraction of latent variables in structural equation models with formative–reflective schemes is developed and illustrated. The procedure is a valuable alternative to PLS-PM and SEM since it is fully consistent with the causal structure of formative–reflective schemes and it extracts the factor scores without substantial identification or indeterminacy problems. Moreover, the new methodology involves the optimization of an explicit and simple to interpret objective function, provides a natural way to check the correct specification of the model and is computationally light. The superiority of the new algorithm over its competitors is proved through examples involving both simulated and real data.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

We propose a novel least-squares based procedure for the extraction of latent variables in formative–reflective models, which overcomes the main limitations of PLS-PM and of other mainstream tools, like SEM, when applied to this kind of causal structures. The new procedure is based on a simple and easy to understand optimization criterion, which takes into account both the capability of exogenous latent variables to summarize their formative manifest blocks and the ability of the endogenous latent variables to predict their own reflective manifest blocks.

Formative–reflective constructs are, explicitly or implicitly, quite spread throughout socio-economic data analysis, where sets of manifest variables are (often hierarchically) aggregated into composite indicators, in order to get synthetic indexes. As examples, consider the evaluation of multi-dimensional poverty and well-being, the assessment of service quality or the computation of the Regional Competitiveness Index (Annoni et al., 2016), which is based on 73 elementary indicators, combined together to get various intermediate layers of sub-indexes. By complementing the formative side (aggregation process) with a reflective side, one removes much of the potential arbitrariness of the construction of composite indicators

* Corresponding author.
E-mail addresses: marco.fattore@unimib.it (M. Fattore), matteo.pelagatti@unimib.it (M. Pelagatti), giorgio.vittadini@unimib.it (G. Vittadini).

https://doi.org/10.1016/j.csda.2017.11.006
0167-9473/© 2017 Elsevier B.V. All rights reserved.

(e.g., due to the discretionary choice of the aggregation weights), getting much more coherent and reliable models, which can be used not only for descriptive purposes, but also for prediction and causal analysis. This is the reason why formative–reflective models are of interest in practical data analysis. As a matter of fact, only the lack of effective tools to treat them algorithmically has somehow limited their practical use insofar, thus motivating the present paper. Deeper discussions of the formative–reflective scheme can be found in Tenenhaus et al. (2005), Bollen and Bauldry (2011), Bainter and Bollen (2014), Lovaglio and Vittadini (2014), Bollen and Diamantopoulos (2017a), Hardin (2017) and Bollen and Diamantopoulos (2017b).

Although at theoretical level the debate on the role and the use of formative models is still very fired, in practice these models are involved in many applied studies (for instance, Vittadini and Lovaglio, 2004; Diamantopoulos et al., 2008; Bruhn et al., 2008; Ruiz et al., 2008; Bollen and Bauldry, 2011), making it urgent to develop effective and reliable tools to deal with them. As a matter of fact, there are still unsolved methodological problems in the theory of structural equation models with formative constructs, particularly when formative–reflective schemes are adopted. Such schemes, often used in causal modelling, are typically addressed using structural equation models (SEM; Jöreskog, 1970, 1981; Jöreskog and van Thillo, 1972) or partial least squares procedures (PLS-PM; Wold, 1982, 1975a). However, a wide literature has progressively underlined the limitations of both approaches, particularly in formative–reflective schemes. For instance, Chin and Todd (1995) and Chin (1998) hold that SEMs support only reflective manifest variables, while Williams et al. (2003) concedes that formative schemes are possible under restrictive conditions. Furthermore, Jöreskog (1981, p.90), in a reply to comments to his seminal article, admits that there is no sufficient and necessary condition for the identification of the model and that the identification problem may be too difficult to resolve for practitioners. As for the possible and often overlooked indeterminacy or non-uniqueness of SEM latent scores, the reader can refer to Schönemann (1971), Schönemann and Steiger (1976), Vittadini (1989), Hwang and Takane (2004), Vittadini et al. (2007) and Lovaglio and Vittadini (2014).

The complexity and limitations of the SEM framework has encouraged the adoption of alternative "soft modelling" approaches, in particular spreading the use of PLS-PM in applied studies. A deep criticism to this procedure can be found in Vittadini et al. (2007), where the essential inconsistencies of PLS-PM in formative–reflective schemes are made explicit and an original algorithm, called RA-PM, is proposed as a valuable alternative. RA-PM overcomes many of the issues of SEM and PLS-PM, being consistent with the formative and the reflective relationships in the model and providing unique latent scores; unfortunately, RA-PM is only conditionally optimal and, like PLS-PM, it does not take into account the formative side of the model properly.

Trying to fill the gap and provide an effective tool for latent variable extraction in formative–reflective schemes, here we propose a new soft modelling procedure, which shares the same benefits of RA-PM, but is globally optimal, takes into account both the formative and reflective sides of the model in a balanced way and, differently from PLS-PM, involves the least-squares optimization of an explicit and simple objective function. 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 to the reflective side. The procedure we propose extracts the exogenous latent variables balancing between these two aspects and provides a way to check the correct specification of the model by verifying if the formative and the reflective sides of the model can be consistently connected through the inner model structure. The extraction procedure has been designed for a formative–reflective scheme with unidimensional blocks of manifest variables, but it can be easily extended and adapted to the case where blocks are multidimensional as we show in Section 2.4. The algorithm can be effortlessly implemented in any programming language with matrix algebra and numerical optimization capabilities. We implemented it in a user-friendly Ox object class (Doornik, 2007) and in an R package (R Core Team, 2017), both freely available from the second author (the R package can be easily installed directly from its GitHub repository by writing `install_github("matteopelagatti/pathmod")` after loading the `devtools` package).

The paper is organized as follows: Section 2 introduces the new technique, Section 3 discusses our approach in comparison to PLS-PM and SEM, Section 4 applies our method to artificial covariance structures to illustrate it and compare it with other methods in a controlled environment, Section 5 applies our technique and its competitors to a well known dataset and Section 6 concludes.

2. Least-squares latent variable extraction in formative–reflective models

In this section, we give a formal definition of formative–reflective models and develop the new least-squares extraction procedure.

2.1. The structure of formative–reflective models

Formative–reflective models are an extension of the formative first-order, formative second-order model cited in Diamantopoulos et al. (2008), where more than one reflective endogenous latent variable is allowed. In such models, p blocks of formative manifest variables (MVs) form p different exogenous latent variables (LVs), which, in turn, form q endogenous LVs that are reflected by q blocks of MVs (see Fig. 1). Both formative and reflective blocks of MVs are usually retained as unidimensional; this restricts the field of application of such models, but makes them easier to handle from the statistical point of view. However, in Section 2.4 we show how to handle multidimensional exogenous blocks (i.e., groups of exogenous MVs forming two or more LVs).

Download English Version:

https://daneshyari.com/en/article/6868848

Download Persian Version:

https://daneshyari.com/article/6868848

Daneshyari.com