Pseudo maximum likelihood estimation in SEM – Scale score versus multiple indicator –

Yutaka Kano
Graduate School of Human Sciences
Osaka University, Japan, kano@hus.osaka-u.ac.jp

Keywords: Structural equation modeling, identification, pseudo MLE.

Abstract

In structural equation modeling, one often introduces fixed parameters to achieve identification of the model under consideration. For example, one path coefficient from a latent variable to an observed one is fixed to be one to determine the scale of the latent variable in a measurement model. It is also useful to make a constraint among parameters to achieve identification. In a more general setting, one may fix a parameter at a pre-calculated statistic, or one may make a functional constraint among parameters where the function depends on pre-calculated statistics.

The general theory for the situation described above has been developed which considers the maximum likelihood inference based on a density function $f(x|\theta_1, \theta_2)$ with an estimated θ_2. See e.g., Gong and Samaniego (1981). Almost all the literature we know assumes that the probability model $f(x|\theta_1, \theta_2)$ is identifiable. Lee and Bentler (1980), Lee and Ho (1993) and Kano et al (1993) among others have developed similar methodologies in the analysis of covariance structures.

In this talk, we first present an example that has motivated us to consider this problem. The example is related with the traditional debate on a scale score (or factor score) versus multiple-indicator (effect-indicator) model. Many similarities and distinctions have been discussed, among which we shall pay attentions to the fact that the scale score includes errors so that attenuation can take place, whereas the multiple-indicator model excludes specific factors as well as errors. One can suggest an alternative model where the error terms are also used as explanatory variables in the multiple-indicator model, to make up for the shortcomings. The model, however, is not identifiable and one needs to use a pre-calculated statistic. Using this model as an example, we shall discuss the identification problem in the pseudo estimation in the general setting described above in the scope of structural equation modeling.

References

Gong, G. and Samaniego, F. J. (1981). Pseudo-maximum likelihood estimation: Theory and applications. *Ann. Statist.*, **9**, 861-869.

Kano, Y., Berkane, M. and Bentler, P. M. (1993). Statistical inference based on pseudo-maximum likelihood estimators in elliptical populations. *J. Amer. Statist. Assoc.*, **88**, 135-143.

Lee, Sik-Yum and Ho, Wai-Tung (1993). Analysis of multisample identified and nonidentified structural equation models with stochastic constraints. *Comput. Statist. Data Anal.*, **16**, 441-453.

Lee, Sik-Yum and Bentler, P. M. (1980). Some asymptotic properties of constrained generalized least squares estimation in covariance structure modules. *South African Statist. J.* **14**, 121-136.