UNIVERSITY OF TWENTE.

Research Seminar
2019, Ghent

Non-iterative estimation via method-of-moments and Croon's method

Florian Schuberth[1] Theo K. Dijkstra[2]
[1]University of Twente
[2]University of Groningen

February 7, 2019

Overview

1. A non-linear factor model
2. Non-iterative method-of-moments
3. Croon's method
4. Monte Carlo simulation

Non-linear factor model: structural model

Example of a polynomial/non-linear factor model:

$$
\begin{aligned}
\eta_3 = &\gamma_1\eta_1 + \gamma_2\eta_2 + \\
&\gamma_{11}(\eta_1^2 - 1) + \gamma_{22}(\eta_2^2 - 1) + \\
&\gamma_{12}(\eta_1\eta_2 - \mathsf{E}(\eta_1\eta_2)) + \zeta_3
\end{aligned} \tag{1}
$$

All latent variables are standardized, and $\mathsf{E}(\zeta_3|\eta_1, \eta_2) = 0$.

Non-linear factor model: measurement model

Each latent variable is measured by at least two indicators and each block of indicators is connected to one LV only:

$$\boldsymbol{y}_i = \boldsymbol{\lambda}_i \eta_i + \boldsymbol{\epsilon}_i. \quad (2)$$

The indicators are standardized and the measurement errors are mutually independent and independent of the ηs. The correlation of the indicators of one block can be calculated as

$$\mathsf{E}(\boldsymbol{y}_i \boldsymbol{y}_i') = \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' + \boldsymbol{\Theta}_i, \quad (3)$$

where the covariance matrix of the measurement errors $\boldsymbol{\Theta}_i$ is a diagonal matrix.
The correlation between the indicators of different block $(i \neq j)$:

$$\mathsf{E}(\boldsymbol{y}_i \boldsymbol{y}_j') = \rho_{ij} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j'. \quad (4)$$

How to estimate this type model

The literature suggests several ways to estimate such models, e.g.,

- Latent Moderated Structural Equations (LMS)
  [Klein, A. & Moosbrugger, H., 2000]
- Quasi-Maximum Likelihood (QML)
  [Klein, A. & Muthén, B. O., 2007]
- Consistent Partial Least Squares (PLSc)
  [Dijkstra, T. K. & Schermelleh-Engel, K., 2014]
- Product Indicator Approach
  [Kenny, D. A. & Judd, C. M., 1984]
- Two-stage Method-of-Moments (2SMM)
  [Wall, M. M. & Amemiya, Y., 2000]
- non-iterative method-of-moments
  [Dijkstra, T. K., 2014, Schuberth, F. et al., in progress]
- ...

Proxies and weights

To estimate the model parameters, we build proxies for each LV as weighted linear combination of its indicators. The weights used to build proxy i are obtained as

$$\hat{\boldsymbol{w}}_i \propto \sum_{j \neq i} e_{ij} \boldsymbol{S}_{ij} \boldsymbol{w}_j, \tag{5}$$

where $\boldsymbol{w}_j$ is an arbitrary vector of the same length as $\boldsymbol{y}_j$ and $e_{ij} = \text{sign}(\boldsymbol{w}_i' \boldsymbol{S}_{ij} \boldsymbol{w}_j)$. Both weight vectors $\boldsymbol{w}_j$ and $\hat{\boldsymbol{w}}_i$ are scaled: $\boldsymbol{w}_j' \boldsymbol{S}_{jj} \boldsymbol{w}_j$ and $\hat{\boldsymbol{w}}_i' \boldsymbol{S}_{ii} \hat{\boldsymbol{w}}_i = 1$.

The probability limit of $\hat{\boldsymbol{w}}_i$ is

$$\text{plim}(\hat{\boldsymbol{w}}_i) = \bar{\boldsymbol{w}}_i = \boldsymbol{\lambda}_i / \sqrt{\boldsymbol{\lambda}_i' \boldsymbol{\Sigma}_{ii} \boldsymbol{\lambda}_i}. \tag{6}$$

Consistent estimation of the loadings

Calculate a factor $\hat{c}_i$ such that the squared Euclidean difference between the off-diagonal elements of

$$\boldsymbol{S}_{ii} \text{ and } \hat{c}_i \hat{\boldsymbol{w}}_i \hat{c}_i \hat{\boldsymbol{w}}_i{}' \tag{7}$$

is minimized. As a result, we obtain

$$\hat{c}_i = \sqrt{\frac{\hat{\boldsymbol{w}}_i'(\boldsymbol{S}_{ii} - \text{diag}(\boldsymbol{S}_{ii}))\hat{\boldsymbol{w}}_i}{\hat{\boldsymbol{w}}_i'(\hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i' - \text{diag}(\hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i'))\hat{\boldsymbol{w}}_i}}. \tag{8}$$

The factor loadings can be consistently estimated as $\hat{\boldsymbol{\lambda}}_i = \hat{c}_i \hat{\boldsymbol{w}}_i$.

The probability limit of $\hat{c}_i$ is denoted as $\bar{c}_i = \text{plim}(\hat{c}_i) = \sqrt{\boldsymbol{\lambda}_i' \boldsymbol{\Sigma}_{ii} \boldsymbol{\lambda}_i}$.

Relationship between latent variables and proxies

We define a population proxy:

$$\bar{\eta}_i = \bar{\boldsymbol{w}}_i' \boldsymbol{y}_i = (\bar{\boldsymbol{w}}_i' \boldsymbol{\lambda}_i)\eta_i + \bar{\boldsymbol{w}}_i' \boldsymbol{\epsilon}_i = Q_i \eta_i + \delta_i, \tag{9}$$

where Q_i (quality) is the correlation between the proxy and the latent variable. The δ_i's have zero mean and are mutually independent and independent of the η_i's.

Replacing $\boldsymbol{\lambda}_i$ by $\bar{c}_i \bar{\boldsymbol{w}}_i$, we obtain

$$Q_i = \bar{\boldsymbol{w}}_i' \boldsymbol{\lambda}_i = \bar{c}_i \bar{\boldsymbol{w}}_i' \bar{\boldsymbol{w}}_i. \tag{10}$$

We can estimate the quality by

$$\widehat{Q}_i = \widehat{c}_i \widehat{\boldsymbol{w}}_i' \widehat{\boldsymbol{w}}_i. \tag{11}$$

Relationship between latent variables and proxies

Relationship between the proxies' correlation and latent variables' correlation:

$$\mathrm{E}(\bar{\eta}_i \bar{\eta}_j) = \mathrm{E}((\bar{\boldsymbol{w}}_i' \boldsymbol{y}_i)(\bar{\boldsymbol{w}}_j' \boldsymbol{y}_j)) = \tag{12}$$

$$\mathrm{E}((\bar{\boldsymbol{w}}_i' \boldsymbol{\lambda}_i \eta_i + \bar{\boldsymbol{w}}_i' \boldsymbol{\epsilon}_i)(Q_j \eta_j + \boldsymbol{\delta}_j)) = \tag{13}$$

$$\mathrm{E}(Q_i Q_j \eta_i \eta_j) + \mathrm{E}(Q_i \eta_i \delta_j) + \mathrm{E}(\delta_i Q_j \eta_j) + \mathrm{E}(\delta_i \delta_j) = \tag{14}$$

$$Q_i Q_j \, \mathrm{E}(\eta_i \eta_j), \tag{15}$$

where $\mathrm{E}(\bar{\eta}_i \bar{\eta}_j)$ is estimated by the sample covariance between the proxies i and j.

Interaction terms

Starting point is a model with a two-way interaction term:

$$\eta_3 = \gamma_1\eta_1 + \gamma_2\eta_2 + \gamma_{12}(\eta_1\eta_2 - \mathrm{E}(\eta_1\eta_2)) + \zeta_3. \tag{16}$$

The γ's can be obtained by solving the following moment equations:

$$\begin{pmatrix} \mathrm{E}(\eta_1\eta_3) \\ \mathrm{E}(\eta_2\eta_3) \\ \mathrm{E}(\eta_1\eta_2\eta_3) \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{E}(\eta_1\eta_2) & \mathrm{E}(\eta_1^2\eta_2) \\ & 1 & \mathrm{E}(\eta_1\eta_2^2) \\ & & \mathrm{E}(\eta_1^2\eta_2^2) - \mathrm{E}(\eta_1\eta_2)^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_{12} \end{pmatrix}, \tag{17}$$

where the moments are given by:

$$\mathrm{E}(\bar{\eta}_i\bar{\eta}_j) = Q_iQ_j\mathrm{E}(\eta_i\eta_j), \tag{18}$$

$$\mathrm{E}(\bar{\eta}_i^2\bar{\eta}_j) = Q_i^2Q_j\mathrm{E}(\eta_i^2\eta_j), \tag{19}$$

$$\mathrm{E}(\bar{\eta}_i^2\bar{\eta}_j^2) = Q_i^2Q_j^2(\mathrm{E}(\eta_i^2\eta_j^2) - 1) + 1, \text{ and} \tag{20}$$

$$\mathrm{E}(\bar{\eta}_i\bar{\eta}_j\bar{\eta}_k) = Q_iQ_jQ_k\mathrm{E}(\eta_i\eta_j\eta_k). \tag{21}$$

Distributional assumptions?

So far, **no distributional assumptions are necessary**. We only assume that

- the moments exist,
- the measurement errors are mutually independent and independent from the η's, and
- that the structural error term is independent from the η's of the right-hand side.

Quadratic terms

Adding quadratic terms to the equation with the one-way interaction term:

$$\begin{aligned}\eta_3 =&\gamma_1\eta_1 + \gamma_2\eta_2 + \\ &\gamma_{11}(\eta_1^2 - 1) + \gamma_{12}(\eta_1\eta_2 - \mathsf{E}(\eta_1\eta_2)) + \gamma_{22}(\eta_2^2 - 1) + \zeta_3.\end{aligned} \tag{22}$$

Now higher moments are required, and therefore more assumptions are necessary, in particular, assumptions about the higher moments of the error terms

$$\mathsf{E}(\bar{\eta}_i{}^3) = Q_i^3\,\mathsf{E}(\eta_i^3) + \mathsf{E}(\delta_i^3), \tag{23}$$

$$\mathsf{E}(\bar{\eta}_i^4) = Q_i^4\,\mathsf{E}(\eta_i^4) + 6Q_i^2(1 - Q_i^2) + \mathsf{E}(\delta_i^4), \text{ and} \tag{24}$$

$$\mathsf{E}(\bar{\eta}_i^3\bar{\eta}_j) = Q_i^3 Q_j\,\mathsf{E}(\eta_i^3\eta_j) + 3\,\mathsf{E}(\bar{\eta}_i\bar{\eta}_j)(1 - Q_i^2). \tag{25}$$

Additional assumptions are required

Way out: we assume that δ_i, has the same higher moments as the normal distribution, i.e.,

$$\begin{aligned} \mathrm{E}(\delta_i^3) =& 0, \text{ and} && (26) \\ \mathrm{E}(\delta_i^4) =& 3\left[\mathrm{var}(\delta_i)\right]^2 = 3(1 - Q_i^2)^2. && (27) \end{aligned}$$

Presentation of Yves Rosseel & Ines Devlieger

Why we may not need SEM after all

Yves Rosseel & Ines Devlieger
Department of Data Analysis
Ghent University – Belgium

March 15, 2018
Meeting of the SEM Working Group – Amsterdam

Presentation of Yves Rosseel & Ines Devlieger

future plans and challenges

- challenge: (analytical) standard errors that perform well in the presence of missing indicators and/or non-normal (but continuous) indicators
- challenge: categorical indicators
- challenge: nonlinear/interaction effects (involving latent variables)
- challenge: models where the distinction between the measurement part and the structural part of the model is not clear
- solved: extension to multilevel SEM (see talk by Ines on EAM in Jena)
- future plans: study the relationship with other related approaches:
  - consistent PLS
  - model-implied instrumental variables estimation
  - two-step approaches
  - ...

Croon's approach

Croon's approach [Croon, M. A., 2002] can be used to estimate non-linear factor models by conducting the following steps:

- Estimate each measurement model by CFA to obtain the factor loading estimates $\hat{\boldsymbol{\lambda}}_i$.
- Build proxies by sum scores, i.e., unit weights $\hat{\boldsymbol{w}}_i = \boldsymbol{\iota}/\sqrt{\boldsymbol{\iota}'\boldsymbol{S}_{ii}\boldsymbol{\iota}}$.
- Estimate the quality of the proxy as $\hat{Q}_i = \hat{\boldsymbol{w}}_i'\hat{\boldsymbol{\lambda}}_i$.
- Estimate the moments using these quality estimates.

Design of the Monte Carlo simulation

- Structural model:

$$\eta_3 = 0.3\eta_1 + 0.4\eta_2 + 0.12(\eta_1^2 - 1) + 0.15(\eta_1\eta_2 - 0.3) + 0.1(\eta_2^2 - 1) + \zeta_3. \quad (28)$$

- Correlation between η_1 and η_2 is set to $\rho_{12} = 0.3$
- Each latent variable is measured by 3 indicators, $\lambda_i' = \begin{pmatrix} 0.9 & 0.85 & 0.8 \end{pmatrix}$
- Exogenous variables are normally distributed
- Sample size of $N = 400$ and 500 runs
- Estimators: Croon's approach, Non-iterative method-of-moments, and LMS

Results of the Monte Carlo simulation

Para.	true	Croon[1]	Non-iter.[2]	LMS[1]
γ_1	0.300	0.297	0.297	0.297
		(0.048)	(0.048)	(0.047)
γ_2	0.400	0.404	0.403	0.402
		(0.047)	(0.047)	(0.045)
γ_{11}	0.120	0.120	0.119	0.117
		(0.045)	(0.044)	(0.041)
γ_{12}	0.150	0.148	0.146	0.147
		(0.063)	(0.062)	(0.059)
γ_{22}	0.100	0.106	0.105	0.101
		(0.043)	(0.043)	(0.039)
ρ_{12}	0.300	0.299	0.300	0.299
		(0.052)	(0.052)	(0.052)

[1]No inadmissible solutions were produced
[2]Inadmissible solutions are removed, therefore the results are based on 484 estimations

Outlook

What we have done so far in case of the non-iterative method-of-moments estimator:

- Implementation in R (cSEM package ↗)
- Allow for correlated measurement errors within a block of indicators
- Assume normality of all exogenous variables $\Rightarrow$ facilitates calculation of the moments

For future research:

- Generation of non-normally distributed data maintaining the covariance structure
- Estimate non-recursive models, e.g., by 2SLS
- Deal with categorical indicators
- Apply approach to other methods, e.g., MIIV-SEM
- Test for overall model fit

Thank you!
Questions/Comments?

References

Bollen, A. K. (1996)
An Alternative 2SLS Estimator for Latent Variable Models
Psychometrika 61(1) 109 – 121.

Croon, M. A., (2002)
Using predicted latent scores in general latent structure models. In Marcoulides, G. & Moustaki, I. (eds)
Latent Variable and Latent Structure Models 195 – 224.

Dijkstra, T. K. (2014)
Very simple estimators for a class of polynomial factor models
The VI European Congress of Methodology, Utrecht, The Netherlands, July 24, 2014.

Dijkstra, T. K. & Schermelleh-Engel, K. (2014)
Consistent partial least squares for nonlinear structural equation models
Psychometrika 79(4) 585 – 604.

References

Klein, A. & Moosbrugger, H. (2000)
Maximum likelihood estimation of latent interaction effects with the LMS method
Psychometrika 65(4) 457 – 474.

Klein, A. & Muthén, B. O. (2007)
Quasi-Maximum Likelihood Estimation of Structural Equation Models With Multiple Interaction and Quadratic Effects
Multivariate Behavioral Research 42(4) 647 – 673.

Kenny, D. A. & Judd, C. M. (1984)
Estimating the nonlinear and interactive effects of latent variables
Psychological Bulletin, 96(1), 201 – 210.

Schuberth, F., Büchner, R., Schermelleh-Engel, K. & Dijkstra, T. K. (in progress)
Polynomial factor models: non-iterative estimation via method-of-moments
Working Paper.

References

Wall, M. M. & Amemiya, Y. (2000)
Estimation for polynomial structural equation models
Journal of the American Statistical Association, 95(451), 929 – 940.