Derivatives and standard errors of standardized parameters in structural equation models

Daniel Oberski
RECSM research paper, Universitat Pompeu Fabra

May 17, 2011

Abstract

In structural equation models, often not only the parameters of the model, but also the standardized coefficients are of interest to the researcher. The literature does not, however, provide explicit analytical standard errors for standardized parameters.

We derive an explicit expression in terms of the model parameters for derivatives and asymptotic standard errors of standardized parameters in structural equation models, that is straightforward to implement in standard SEM software. The expression is applied to an example application.

1 Introduction

Linear structural equation models (SEM) with latent variables have become a popular tool in different branches of science. Such models encompass as special cases a diverse range of common models of interest such as factor analysis, multivariate regression, errors-in-variables models, growth, and multilevel and multigroup models (Bollen, 1989). In addition, extensions for categorical, count, and censored dependent variables as well as complex sampling available (Muthen and Satorra, 1995; Muthén, 2002).

Although the main interest in such models lies in the parameters of the model, researchers' interest often also focuses on the so-called 'standardized' parameters (Bollen, 1989). Typical applications include examination of factor loadings and correlations in factor analysis, and the evaluation of the relative size of regression coefficients, possibly of latent variables.

In spite of the interest in standardized coefficients in structural equation models, standard errors and confidence intervals for these coefficients are usually not provided by the standard software[1]. To our knowledge the literature does not provide any explicit expression for the asymptotic standard errors of standardized coefficients. We remedy this situation by deriving an explicit expression for the asymptotic variance-covariance matrix of the standardized coefficients. The solution requires only the parameter estimates of the model and can be readily implemented in SEM software.

Section 2 provides a motivating example, using a SEM model with latent variables where standardized parameters and their standard errors or confidence intervals are of interest to the researcher. Section 3 derives the explicit expression for the asymptotic variance-covariance matrix of the standardized estimates. In section 4, this expression is then applied to the example. Finally, the last section discusses the scope and limitations of this proposal and suggests future research.

[1] At the time of writing, the exception is Mplus version 5.2 and above (Muthén and Muthén, 1998).

2 Example SEM with standardized parameters

The study of differences across societal groups in self-rated health is of interest to researchers in public health. **?**, for example, compare people with different incomes, age groups, and level of education. It is well-known, however, that in order to be able to compare correlations across groups it is necessary for the reliability of the measures to be the same. Therefore not only across-group comparison of the levels of self-rated health are of interest, but also the evaluation of differences between the groups in reliability (**?**).

Different types of designs exist to estimate the reliability of survey measures. One such design is the repeated measures design, wherein the same question is asked at least three times with a certain interval. One can then apply the so-called 'quasi-simplex' model to estimate reliability **??**. For an overview of other designs we refer to **?**.

The quasi-simplex model can be formulated and estimated as a structural equation model in different societal groups, such as groups with different levels of completed education. This model is shown for four repetitions in figure **??**. Following **?**, the unstandardized error variances are restricted to be equal across the four repetitions. The model is formulated for the observed variables with loadings set to unity. The reliability coefficients of interest are then the standardized loadings.

We estimate this model using data from the LISS panel study in the Netherlands. The LISS panel is a random probability sample of 8000 Dutch citizens. The respondents answer questionnaires over the web. For more information about the LISS we refer to **?**. The LISS panel contains a study that included the commonly used self-rated health question,

> All things considered, how healthy would you say you are nowadays? 1-5.

This question was asked of 3425 LISS respondents in the years 2007, 2008, 2009, and 2010. Using the data from all four years, we estimate the quasi-simplex model shown in figure **??** as a SEM to yield four standardized loadings for each educational group, which can be interpreted as the reliability coefficients for each group[2]. The results are shown in table 1.

	Education level					
	Primary $n = 279$	Lower secondary $n = 940$	Middle secondary $n = 782$	Upper secondary $n = 369$	Lower tertiary $n = 799$	Upper tertiary $n = 256$
Year						
2007	0.799	0.821	0.822	0.891	0.869	0.896
2008	0.818	0.821	0.820	0.879	0.878	0.898
2009	0.828	0.812	0.808	0.896	0.863	0.886
2010	0.816	0.822	0.805	0.877	0.874	0.887

Table 1: Reliability of self-rated health in the Netherlands 2007-2010 for groups with different levels of education.

Table 1 shows that there appear to be quite some differences across the groups in

[2]For the estimation we used `R` 2.13.0 64-bit (**?**) and `lavaan` 0.4.9 (**?**)

3 Standard errors of standardized parameters

Let y be a p-vector of observed variables, from which a sample is obtained. The following SEM for y is specified:

$$y = \Lambda\eta + \epsilon \tag{1}$$
$$\eta = B_0\eta + \zeta, \tag{2}$$

where η is a vector of unobserved variables, ζ is a vector of disturbance terms and ϵ is a vector of measurement errors. Model 2 implies the following model $\Sigma_\eta(\theta)$ for the variance-covariance matrix of the unobserved variables as a function of a parameter vector θ:

$$\Sigma_\eta(\theta) = B^{-1}\Phi B^{-T}, \tag{3}$$

where $B \triangleq I - B_0$ is positive definite, and Φ is the variance-covariance matrix of ζ. Model 1 can then be seen to imply the following model $\Sigma_y(\theta)$ for the variance-covariance matrix of the observed variables:

$$\Sigma_y(\theta) = \Lambda B^{-1}\Phi B^{-T}\Lambda' + \Psi, \tag{4}$$

where Ψ is the variance-covariance matrix of ϵ. We assume throughout that both Σ_y and Σ_η are positive definite. In what follows we will write $\Sigma_.$ for $\Sigma_.(\theta)$ in the interest of clarity.

The parameters of the model are collected in a parameter vector

$$\theta \triangleq [\mathrm{vec}\,\Lambda, \mathrm{vech}\,\Phi, \mathrm{vec}\,B_0, \mathrm{vech}\,\Psi].$$

Often, interest focuses not only on the parameters θ, but also on the so-called "standardized" matrices, which we denote $\tilde{\Lambda}$ and $\tilde{B}_0$. These are defined as:

$$\tilde{\Lambda} \triangleq D_y^{-1}\Lambda D_\eta \tag{5}$$
$$\tilde{B}_0 \triangleq D_\eta^{-1}B_0 D_\eta, \tag{6}$$

where $D_y \triangleq \sqrt{I \circ \Sigma_y}$, and $D_\eta \triangleq \sqrt{I \circ \Sigma_\eta}$.

TODO: Define $\tilde{\theta}$.

By standard application of the Delta method (e.g. Oehlert, 1992), the asymptotic variance of $\tilde{\theta}$ is

$$\left(\frac{d\tilde{\theta}}{d\theta}\right) \mathrm{var}(\theta) \left(\frac{d\tilde{\theta}}{d\theta}\right)', \tag{7}$$

where $\mathrm{var}(\theta)$ is the appropriate asymptotic variance matrix of the free model parameters θ (e.g. Satorra, 1989).

3.1 Derivatives of the standardized parameters

To obtain the full expression for the variance-covariance matrix of standardized parameters in equation 7, the needed derivatives of the standardized parameters with respect to the free parameters of the model are derived here.

From definition 5,

$$\mathrm{d}\,\mathrm{vec}\,\tilde{\Lambda} = (D_\eta\Lambda' \otimes I_p)\,\mathrm{d}\,\mathrm{vec}\,D_y^{-1} + (D_\eta \otimes D_y^{-1})\,\mathrm{d}\,\mathrm{vec}\,\Lambda + (I_q \otimes D_y^{-1}\Lambda)\,\mathrm{d}\,\mathrm{vec}\,D_\eta, \tag{8}$$

and from definition 6,

$$\mathrm{d}\,\mathrm{vec}\,\tilde{B}_0 = (D_\eta B_0' \otimes I_p)\,\mathrm{d}\,\mathrm{vec}\,D_\eta^{-1} + (D_\eta \otimes D_\eta^{-1})\,\mathrm{d}\,\mathrm{vec}\,B_0 + \\ (I_q \otimes D_\eta^{-1} B_0)\,\mathrm{d}\,\mathrm{vec}\,D_\eta. \quad (9)$$

The differentials of the standardized parameter matrices are, thus, functions of the differentials of the covariance structure models Σ_y and Σ_η. Neudecker and Satorra (1991) derived the differential of the implied variance matrix Σ_y of the observed variables. To ensure completeness of the treatment, we repeat it here:

$$\begin{aligned} \mathrm{d}\,\mathrm{vec}\,\Sigma_y = (I+K)(\Lambda B^{-1}\Phi B^{-T} \otimes I)\,\mathrm{d}\,\mathrm{vec}\,\Lambda + \\ (I+K)(\Lambda B^{-1}\Phi B^{-T} \otimes \Lambda B^{-1})\,\mathrm{d}\,\mathrm{vec}\,B_0 + \\ (\Lambda B^{-1} \otimes \Lambda B^{-1})P\,\mathrm{d}\,\mathrm{vech}\,\Phi + \\ \mathrm{d}\,\mathrm{vech}\,\Psi, \end{aligned} \quad (10)$$

where the commutation matrix K, the duplication matrix P, and the operators vec and vech are defined in Magnus and Neudecker (1988) Let Δ_λ equal the coefficient matrix of $\mathrm{d}\,\mathrm{vec}\,\Lambda$ in equation 10, Δ_β the coefficient of $\mathrm{d}\,\mathrm{vec}\,B_0$, and Δ_ϕ the coefficient of $\mathrm{d}\,\mathrm{vech}\,\Phi$.

The differential of the variance matrix of η can be obtained as:

$$\begin{aligned} \mathrm{d}\,\mathrm{vec}\,\Sigma_\eta = (I+K)(B^{-1}\Phi B^{-T} \otimes B^{-1})\,\mathrm{d}\,\mathrm{vec}\,B_0 + \\ (B^{-1} \otimes B^{-1})P\,\mathrm{d}\,\mathrm{vech}\,\Phi. \end{aligned} \quad (11)$$

Let Δ_β^* and Δ_ϕ^* be the coefficients of $\mathrm{d}\,\mathrm{vec}\,B_0$ and $\mathrm{d}\,\mathrm{vech}\,\Phi$ respectively in equation 11.

Also, let

$$C_{y2} \triangleq -(I_p \otimes \frac{1}{2(I_p \circ \Sigma_y)^{\frac{3}{2}}})\mathrm{diag}[\mathrm{vec}\,(I_p)], \quad (12)$$

$$C_{\eta 1} \triangleq (I_q \otimes \frac{1}{2(I_q \circ \Sigma_\eta)^{\frac{1}{2}}})\mathrm{diag}[\mathrm{vec}\,(I_q)], \quad (13)$$

$$C_{\eta 2} \triangleq -(I_q \otimes \frac{1}{2(I_q \circ \Sigma_\eta)^{\frac{3}{2}}})\mathrm{diag}[\mathrm{vec}\,(I_q)]. \quad (14)$$

Then, applying standard operations on 8 and rearranging terms, the differential of the standardized $\tilde{\Lambda}$ matrix is

$$\begin{aligned} \mathrm{d}\,\mathrm{vec}\,\tilde{\Lambda} =& [(D_\eta \Lambda' \otimes I_p)C_{y2}\Delta_\lambda + (D_\eta \otimes D_y^{-1})]\,\mathrm{d}\,\mathrm{vec}\,\Lambda + \\ & [(D_\eta \Lambda' \otimes I_p)C_{y2}\Delta_\beta + (I_q \otimes D_y^{-1}\Lambda)C_{\eta 1}\Delta_\beta^*]\,\mathrm{d}\,\mathrm{vec}\,B_0 + \\ & [(D_\eta \Lambda' \otimes I_p)C_{y2}\Delta_\phi + (I_q \otimes D_y^{-1}\Lambda)C_{\eta 1}\Delta_\phi^*]\,\mathrm{d}\,\mathrm{vec}\,\Phi + \\ & (D_\eta \Lambda' \otimes I_p)C_{y2}\,\mathrm{d}\,\mathrm{vec}\,\Psi. \end{aligned} \quad (15)$$

And similarly, the differential of the standardized $\tilde{B}_0$ matrix from equation 9 is

$$\begin{aligned} \mathrm{d}\,\mathrm{vec}\,\tilde{B}_0 =& [\{(D_\eta B_0' \otimes I_q)C_{\eta 2} + (I_q \otimes D_\eta^{-1}B_0)C_{\eta 1}\}\Delta_\beta^* + (D_\eta \otimes D_\eta^{-1})]\,\mathrm{d}\,\mathrm{vec}\,B_0 + \\ & [(D_\eta B_0' \otimes I_q)C_{\eta 2} + (I_q \otimes D_\eta^{-1}B_0)C_{\eta 1}]\Delta_\phi^*\,\mathrm{d}\,\mathrm{vec}\,\Phi. \end{aligned} \quad (16)$$

TODO: Lose the G? Just write $d\tilde{\theta}/d\theta$?

From the differentials in equations 15 and 16, we conclude that the derivative matrices $G_{\tilde{\lambda}}$ and $G_{\tilde{\beta}}$ of the standardized parameters $\mathrm{vec}\,\tilde{B}$ and $\mathrm{vec}\,\tilde{\Lambda}$ with respect to the free parameters of the model θ will be the partitioned matrices

$$\begin{aligned} G_{\tilde{\lambda}} = [(D_\eta \Lambda' \otimes I_p) C_{y2} \Delta_\lambda + (D_\eta \otimes D_y^{-1}), \\ (D_\eta \Lambda' \otimes I_p) C_{y2} \Delta_\phi + (I_q \otimes D_y^{-1} \Lambda) C_{\eta 1} \Delta_\phi^*, \\ (D_\eta \Lambda' \otimes I_p) C_{y2} \Delta_\beta + (I_q \otimes D_y^{-1} \Lambda) C_{\eta 1} \Delta_\beta^*, \\ (D_\eta \Lambda' \otimes I_p) C_{y2}] \end{aligned} \quad (17)$$

and

$$\begin{aligned} G_{\tilde{\beta}} = [\mathbf{0}, \{(D_\eta B_0' \otimes I_q) C_{\eta 2} + (I_q \otimes D_\eta^{-1} B_0) C_{\eta 1}\} \Delta_\phi^*, \\ \{(D_\eta B_0' \otimes I_q) C_{\eta 2} + (I_q \otimes D_\eta^{-1} B_0) C_{\eta 1}\} \Delta_\beta^* + (D_\eta \otimes D_\eta^{-1}), \mathbf{0}]. \end{aligned} \quad (18)$$

TODO: Add Φ and Ψ; put it all together in one matrix

4 Application: standard errors of standardized parameters

	Education level					
	Primary	VMBO	MBO	HAVO/VWO	HBO	WO
	$n = 279$	$n = 940$	$n = 782$	$n = 369$	$n = 799$	$n = 256$
Year						
2007	0.799	0.821	0.822	0.891	0.869	0.896
	(0.029)	(0.014)	(0.015)	(0.016)	(0.013)	(0.017)
2008	0.818	0.821	0.820	0.879	0.878	0.898
	(0.024)	(0.013)	(0.014)	(0.017)	(0.012)	(0.017)
2009	0.828	0.812	0.808	0.896	0.863	0.886
	(0.023)	(0.014)	(0.015)	(0.015)	(0.013)	(0.018)
2010	0.816	0.822	0.805	0.877	0.874	0.887
	(0.027)	(0.014)	(0.017)	(0.018)	(0.013)	(0.019)

Table 2: Reliability of self-rated health in the Netherlands 2007-2010 for different educational groups. Asymptotic standard errors based on the method described in this paper are also given. For direct tests of differences between the groups in reliability, see table 3.

5 Discussion

5.1 Scope

Limited, categorical variables, effects/correlations between latent variables, factor loadings, error correlations.

5.2 Others

Bollen and Stine (1990) note that an approximate method exists for regression analysis that does not take the estimation of the model-implied standard deviations into account.

Comparison	t_{dif}	p_{dif}	$p_{\text{dif,adj}}$	
Primary <> vmbo	-0.68	0.25	0.96	
Primary <> havo/vwo	-2.76	0.00	0.03	*
Primary <> mbo	-0.71	0.24	0.96	
Primary <> hbo	-2.18	0.01	0.10	
Primary <> wo	-2.86	0.00	0.02	*
vmbo <> havo/vwo	-3.25	0.00	0.01	*
vmbo <> mbo	-0.06	0.48	0.96	
vmbo <> hbo	-2.46	0.01	0.06	
vmbo <> wo	-3.35	0.00	0.01	*
havo/vwo <> mbo	3.12	0.00	0.01	*
havo/vwo <> hbo	1.07	0.14	0.71	
havo/vwo <> wo	-0.22	0.41	0.96	
mbo <> hbo	-2.33	0.01	0.08	
mbo <> wo	-3.22	0.00	0.01	*
hbo <> wo	-1.26	0.10	0.62	

Table 3: Multiple comparison between educational groups for the reliability in 2007. The groups that are compared on reliability are shown in the first column. The second column shows the t-value for the test of no difference between the groups in reliability. The last two columns show the corresponding p-value and p-value adjusted for multiple comparisons (Holm, 1979), respectively.

References to Delta method but without providing the derivatives.
Likelihood based confidence intervals, bootstrapping.
Our method dispenses with the need for approximate methods or numerical derivatives.

5.3 Future

Examine relative performance of confidence intervals, similar to work done for indirect effects.

References

Bollen, K. (1989). *Structural equations with latent variables*. Wiley New York.

Bollen, K. and Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological methodology*, 20(1):15–140.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70.

Magnus, J. and Neudecker, H. (1988). Matrix differential calculus with applications in statistics and econometrics.

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29(1; ISSU 51):81–118.

Muthen, B. and Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological methodology*, 25:267–316.

	Unstandardized free parameters θ										
	ψ_{11}	ψ_{22}	ψ_{33}	ψ_{44}	ϕ_{11}	ϕ_{22}	ϕ_{33}	ϕ_{44}	β_{21}	β_{32}	β_{43}
$\lambda_{s,11}$	-0.733	.	.	.	0.277	.	.	.	.	.	.
$\lambda_{s,22}$	.	-0.739	.	.	0.217	0.282	.	.	0.209	.	.
$\lambda_{s,33}$	.	.	-0.755	.	0.188	0.244	0.300	.	0.180	0.225	.
$\lambda_{s,33}$	.	.	.	-0.754	0.173	0.225	0.276	0.298	0.166	0.207	0.230
$\beta_{s,21}$	.	.	.	.	0.232	-1.059	.	.	0.223	.	.
$\beta_{s,32}$	.	.	.	.	0.131	0.171	-1.147	.	0.126	0.157	.
$\beta_{s,43}$	.	.	.	.	0.058	0.076	0.093	-1.193	0.056	0.070	0.077

Table 4: Derivatives of the standardized parameters Λ_s and B_s with respect to the free parameters of the model in the quasi simplex example.

	ψ	ϕ_{11}	ϕ_{22}	ϕ_{33}	ϕ_{44}	β_{21}	β_{32}	β_{43}
ψ	0.2×10^{-4}							
ϕ_{11}	-0.2×10^{-4}	2.2×10^{-4}						
ϕ_{22}	-0.3×10^{-4}	0.4×10^{-4}	0.9×10^{-4}					
ϕ_{33}	-0.2×10^{-4}	0.2×10^{-4}	0.1×10^{-4}	0.5×10^{-4}				
ϕ_{44}	-0.3×10^{-4}	0.3×10^{-4}	0.3×10^{-4}	0.1×10^{-4}	0.8×10^{-4}			
β_{21}	0.4×10^{-4}	-1.5×10^{-4}	-0.9×10^{-4}	-0.1×10^{-4}	-0.5×10^{-4}	4.3×10^{-4}		
β_{32}	0.2×10^{-4}	-0.2×10^{-4}	-0.4×10^{-4}	-0.4×10^{-4}	-0.1×10^{-4}	-0.9×10^{-4}	3.3×10^{-4}	
β_{43}	0.3×10^{-4}	-0.3×10^{-4}	-0.3×10^{-4}	-0.3×10^{-4}	-0.7×10^{-4}	0.4×10^{-4}	-1.0×10^{-4}	3.4×10^{-4}

Table 5: Variance-covariance matrix of the parameter estimates.

Muthén, L. and Muthén, B. (1998). Mplus users guide.

Neudecker, H. and Satorra, A. (1991). Linear structural relations: Gradient and Hessian of the fitting function. *Statistics & Probability Letters*, 11(1):57–61.

Oehlert, G. (1992). A note on the delta method. *The American Statistician*, 46(1):27–29.

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, 54(1):131–151.

A Calculations for the example analysis

B Extension to multigroup SEM

In multigroup SEM, let each matrix in the above equations equal the block diagonal of the corresponding matrices in each group.

Then define the stacking matrix $Z_{p \times pg} \triangleq [I_p, I_p, ..., I_p]$. The derivatives are then

$$(I_q \otimes Z)G$$

Parameter	Estimate	s.e.	z
$\lambda_{s,11}$	0.852	0.006	134
$\lambda_{s,22}$	0.851	0.006	139
$\lambda_{s,33}$	0.846	0.006	136
$\lambda_{s,44}$	0.847	0.007	128
$\beta_{s,21}$	0.882	0.012	75
$\beta_{s,32}$	0.920	0.009	103
$\beta_{s,43}$	0.960	0.011	85

Table 6: Standardized parameter estimates and asymptotic standard errors for the example ($n = 3425$).